JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 86, NO. C10, PAGES 9707-9714, OCTOBER 20, 1981

Turbulence and Stress Owing to Gravity Wave and Tidal Breakdown

R. S. LINDZEN

Naval Research Laboratory, Washington, D. C. 20390
Center for Earth and Planetary Physics, Harvard University, Cambridge, Massachusetts 02139

It has been suggested (Lindzen, 1967, 1968*a, b;* Lindzen and Blake, 1971; Hodges, 1969) that turbulence in the upper mesosphere arises from the unstable breakdown of tides and gravity waves. Crudely speaking, it was expected that sufficient turbulence would be generated to prevent the growth of wave amplitude with height (roughly as (basic pressure)$^{-1/2}$). This work has been extended to allow for the generation of turbulence by smaller amplitude waves, the effects of mean winds on the waves, and the effects of the waves on the mean momentum budget. The effects of mean winds, while of relatively small importance for tides, are crucial for internal gravity waves originating in the troposphere. Winds in the troposphere and stratosphere sharply limit the phase speeds of waves capable of reaching the upper mesosphere. In addition, the existence of critical levels in the mesosphere significantly limits the ability of gravity waves to generate turbulence, while the breakdown of gravity waves contributes to the development of critical levels. The results of the present study suggest that at middle latitudes in winter, eddy coefficients may peak at relatively low altitudes (50 km) and at higher altitudes in summer and during sudden warmings (70–80 km), and decrease with height rather sharply above these levels. Rocket observations are used to estimate momentum deposition by gravity waves. Accelerations of about 100 m/s/day are suggested. Such accelerations are entirely capable of producing the warm winter and cold summer mesopauses.

1. INTRODUCTION

The possibility of breaking waves generating turbulence in the mesosphere was noted some years ago by *Lindzen* [1967; 1968*a*] and *Hodges* [1969] for tides and gravity waves, respectively. The idea, in these papers, was simply that vertically propagating gravity waves that, in the absence of damping, grow in amplitude as $p_0^{-1/2}$ (where p_0 is the basic unperturbed pressure) could at some height reach amplitudes for which the wave fields themselves would be strongly unstable (i.e., the combination of mean and wave fields would have a negative static stability or at least its Richardson number would drop below 0.25). Above such a height it was suggested that the waves would generate sufficient turbulence, on the average, to prevent further wave growth with height. Theoretical results implied that among tidal modes, only the first propagating diurnal mode would prove important in this respect. *Lindzen and Blake* [1971] estimated that this mode would break down between 80 and 90 km, generating turbulence up to a height of about 108 km above which molecular viscosity and conductivity are sufficient to inhibit further wave growth. The turbulence generated by this tidal mode is restricted to within about 30° latitude of the equator as is the mode itself. The cessation of turbulence at 108 km corresponds closely to the height usually designated for the turbopause and in a loose fashion supports the notion of tidal generation for turbulence.

The situation with respect to internal gravity waves is less certain. Structures with moderately short vertical wavelengths (0(12 km)) are commonly observed in rocket soundings, poleward of 30° (thus excluding diurnal tides, at least in some measure), especially in winter. Examples of such soundings are shown in Figure 1. On the basis of dispersion relations given by *Hines* [1960], one commonly—but arbitrarily—identified the gravity waves as oscillations of relatively short period (0(3 hours)). Observations suggested that such gravity waves also broke down, and *Hodges* [1969] estimated the resulting generation of turbulence along similar lines to those discussed in connection with the diurnal tide. Again, the turbulence was assumed to persist up to some turbopause height. These studies tended to ignore the fact that the relevant frequencies were the Doppler shifted frequencies and that intrinsic frequencies might be very different.

Lindzen [1971] presented a crude model of turbulence in the mesosphere that summarized the above mechanisms and, in addition, argued that 'turbulent' diffusion could result from nonbreaking waves (NB tidal modes are simply special cases of internal gravity waves [viz. *Lindzen,* 1970]. The specific argument was based on wave transiency, but other arguments have been advanced by *Weinstock* [1976]. This matter is discussed further in a separate note by R. S. Lindzen and J. Forbes (manuscript in preparation, 1981) wherein we consider the cascade of energy from stable waves to waves of sufficiently small vertical wavelengths to permit unstable breakdown. Returning to *Lindzen* [1971], the model of eddy diffusion presented consisted in a diffusion coefficient that increased as $(p_0)^{-1/2}$ up to a level at which wave breaking occurred, increased markedly at this level (50–90 km), and remained constant up to some turbopause level. Both the magnitude and the distribution of the diffusion coefficient were reasonably compatible with what was called for by measurements of composition [*Hunten,* 1975; *von Zahn et al.,* 1980] but the range of uncertainty was large. The notion of exponentially increasing (with height) eddy coefficients has, however, gained fairly general acceptance.

Recently, *Holton and Wehrbein* [1980] have concentrated on the momentum deposition by breaking waves, modeling this by Rayleigh friction (acting to bring zonal flow to zero). To be sure, wave breaking does lead to deposition of wave momentum flux, quite apart from the generation of diffusive turbulence; the two effects are, however, distinct as will be shown in section 2 of this paper. Moreover, in the case of breaking tides, the deposited momentum is not attempting to bring the mean flow to zero [*Fels and Lindzen,* 1974]. Nevertheless, as was noted by *Leovy* [1964] and *Lindzen* [1968*b,* 1973], if one anticipates that friction is responsible for the reversal of the latitudinal temperature gradient at the meso-

This paper is not subject to U.S. copyright. Published in 1981 by the American Geophysical Union.

Paper number 1C0154.

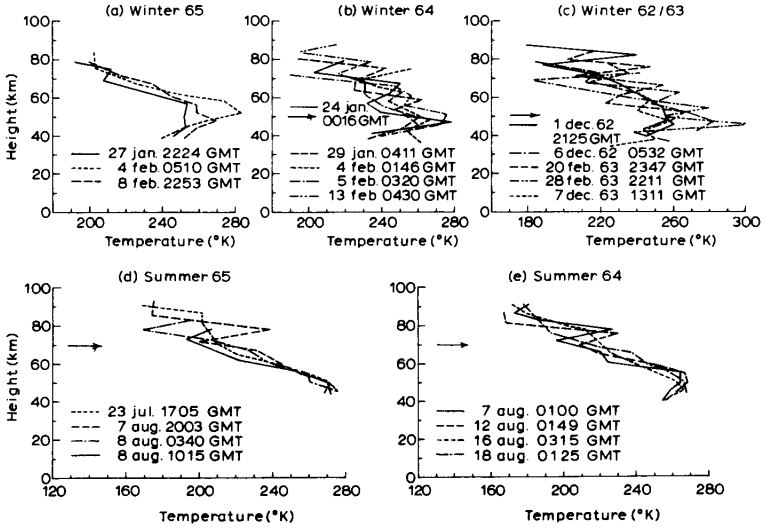


Fig. 1. Winter and summer temperature sounding at Wallops Island (38°N).

pause, then one really does need a mechanism that acts to reduce zonal velocities rather than gradients.

The purpose of the present paper is to reassess carefully, but simply, the role of wavebreaking in the mesosphere, emphasizing the role of the mean flow in determining the behavior of gravity waves. Assuming that most high latitude gravity waves originate in the troposphere, it is shown in section 3 that changes in the mean zonal flow accompanying changing seasons, sudden warmings, etc., dramatically alter the nature of gravity waves in the mesosphere. Such changes lead to simple explanations of various mesospheric observations.

2. Gravity Waves in the Mesosphere—Simple WKB Analysis

In this section, gravity waves typical of middle to high latitudes will be emphasized, though theoretical results are easily extended to tidal gravity waves. It will be assumed that the reader is familiar with the mathematical theory of internal gravity waves. Only selected results will be presented here. The following references are but a few of the references where more complete treatments are presented: *Hines* [1960], *Lindzen* [1970], *Gossard and Hooke* [1975], and *Holton* [1979].

We will begin by presenting some data relevant to our subsequent discussion. Figure 1 shows various temperature profiles over Wallops Island (38°N) taken in winter and summer by using rocket grenades [*Theon et al.*, 1967]. We will choose to interpret the pronounced waviness in these profiles as being due to internal gravity waves. Winter 1965 results appear anomalous. However, the remaining profiles suggest a fairly clear picture. Vertical wavelengths are typically 0(12 km) in both winter and summer (the coarse resolution of the grenade method leaves such estimates fairly uncertain). Below breaking levels there is uncertainty over wavelength because wave amplitudes are weak.

In winter the profiles begin to manifest neutral or even unstable lapse rates above 50 km; in summer such lapse rates are found above 70 km. Such levels are marked in Figures 1 and 3. We will identify these lapse rates with the onset of wave breaking. In Figure 2 we show the 'typical' distribution of zonal wind with the latitude and height for winter and summer (CIRA, 1972). Of course, it is by now recognized that the results in Figure 2 are not likely to be applicable to particular winters and summers. For example, in Figure 3 we show winter and summer wind profiles for the latitude band 30°–45°N based on rocket data for the period 1960–1964 [*U.S. Standard Atmosphere Supplement*, 1966]. Also shown are geostrophic winds based on observed temperatures. There is qualitative similarity to Figure 2, but the quantitative differences are obvious. In particular, Figure 3 shows wind magnitudes peaking at the levels of wavebreaking deduced from Figure 1. For our calculations we will use the results in Figure 3. For reference purposes, we show in Figure 4 standard models for the height distribution of temperature in winter and summer (CIRA, 1972). The information in Figure 4 will be needed in subsequent calculations. However, it should be noted that in the absence of wave perturbations, mean lapse rates are only 0 (± 2.4°/km) which are (for many purposes) small when compared with $(g/c_p) = 9.8°/\text{km}$.

We now proceed to use the mathematical theory of internal gravity waves to exploit the above observationally derived features. The crudeness of these features leads us to restrict ourselves to only the simplest, approximate, theoretical results. Our equation for the vertical velocity perturbation, w, associated with internal gravity waves is

$$\tilde{w}_{zz} + \left(\frac{N^2(1 + l^2/k^2)}{(\bar{u} - c)^2} - \frac{1}{4H^2} \right) \tilde{w} = 0 \tag{1}$$

where

w vertical velocity perturbation;
$\equiv \tilde{w}(z) e^{z/2\bar{H}} e^{ik(x-ct)} \cos(ly + \phi)$; (2)
z height;
H scale height $= R\bar{T}/g$;
$\bar{T}(z)$ basic temperature;
$\bar{u}(z)$ basic zonal wind;
k wave number in x direction;
l wave number in y direction;
x west-east distance;
y south-north distance;
c phase speed in x direction;
g gravitational acceleration;
N^2 (Brunt-Vaisala frequency)$^2 = g/\bar{T}\,\Gamma$;
Γ $(\partial\bar{T}/\partial z) + (g/c_p)$;
ϕ arbitrary phase.

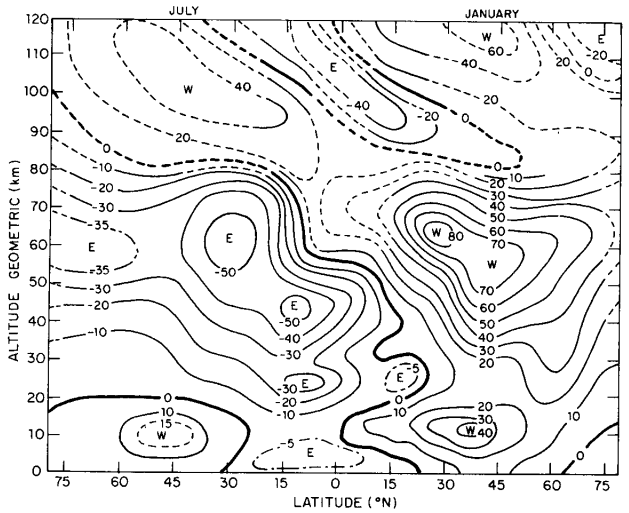


Fig. 2. Zonal wind contours as a function of height and latitude for July and January (CIRA, 1972).

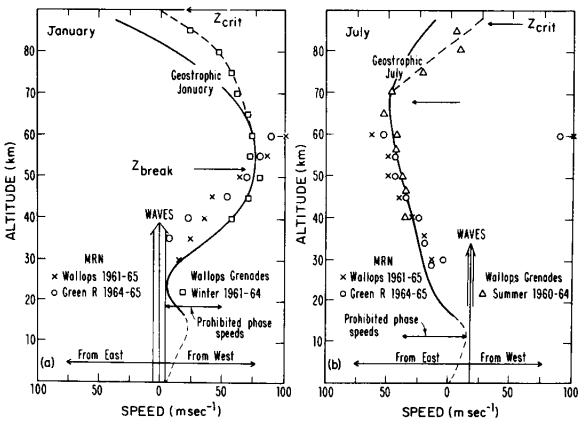


Fig. 3. Zonal wind as a function of height in the latitude belt 30°–45°N based on rocket probes. (*a*) Results for January; (*b*) results for July. Also shown are the geostrophic winds derived from temperature soundings and the phase speeds permitted and prohibited for mesospheric gravity waves of tropospheric origin.

Equation (1) assumes that $\bar{u}_{zz}$ is not too large. Equation (1) has the following approximate WKB solution [viz. *Bender and Orszag,* 1978]:

$$\tilde{w} \approx A\lambda^{-1/2} \exp\{i \int \lambda\, dz'\} \tag{3}$$

where $\lambda^2 = [N^2(1 + l^2/k^2)/(\bar{u} - c)^2] - (1/4H^2)$, and A = constant.

For the problems we are considering

$$\frac{1}{4H^2} \ll \frac{N^2(1 + l^2/k^2)}{(\bar{u} - c)^2}$$

and

$$\lambda^2 \approx \frac{N^2(1 + l^2/k^2)}{(\bar{u} - c)^2} \tag{4}$$

(2) and (3) now determine w. We next wish to obtain the temperature perturbation associated with w. From the thermodynamic energy equation one has

$$ik(c - \bar{U})\delta T \approx w\Gamma \tag{5}$$

or by using (3) and (4)

$$\delta T \approx \frac{i\Gamma^{1/2}T^{-1/2}A}{g^{1/2}k(1 + l^2/k^2)^{1/2}} \lambda^{1/2} e^{i\int\lambda dz} e^{z/2H} e^{ik(x-ct)} \cos(ly + \phi) \tag{6}$$

Equation (6) is appropriate below the level of wave breakdown. For some purposes it proves useful to allow the multiplicative constant in (6) to have a parametric dependence on latitude. It is also convenient to incorporate

$$-ig^{1/2}k\left(1 + \frac{l^2}{k^2}\right)^{1/2}$$

into the 'constant':

$$\delta T \approx A(\theta)\Gamma^{1/2}T^{-1/2}\lambda^{1/2} e^{i\int\lambda dz}\ e^{z/2H} e^{ik(x-ct)} \cos(ly + \phi) \tag{7}$$

where θ = latitude.

For the waves shown in Figure 1, the factor $e^{i\int\lambda dz}$ dominates the z variation of δT, so

$$\frac{d\delta T}{dz} \approx iA(\theta)\Gamma^{1/2}\bar{T}^{-1/2}\lambda^{3/2} e^{i\int\lambda dz}\, e^{z/2H} e^{ik(x-ct)} \cos(ly + \phi) \tag{8}$$

For simplicity, we will identify the level of wave breaking with that level at which

$$\left|\frac{d\delta T}{dz}\right| = \Gamma \tag{9}$$

or from (8)

$$A(\theta)\Gamma^{1/2}\bar{T}^{1/2}\lambda^{3/2} e^{z_{break}/2H} \approx \Gamma \tag{10}$$

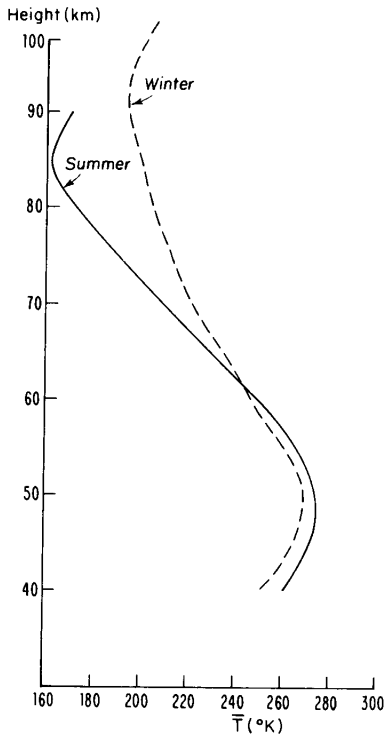


Fig. 4. Zonally averaged temperature as a function of height at 40°N for winter and summer (CIRA, 1972).

Given $A(\theta)$, (10) determines z_{break}, or, alternately, we may use z_{break} as found in Figure 1 to determine $A(\theta)$.

We now hypothesize as in *Lindzen* [1968] and *Hodges* [1969] that above z_{break}, sufficient diffusive turbulence is generated to prevent further growth of $|d\delta T/dz|$. Because of this turbulence, the momentum flux carried by the wave will also be deposited in the mean flow above z_{break}. Our determination of the turbulent eddy coefficient follows the treatment of *Hodges* [1969]. The most important factors in (8) governing the growth of $|d\delta T/dz|$ are $\lambda^{3/2}$ and $e^{z/2H}$, i.e.,

$$\left|\frac{d\delta T}{dz}\right| \alpha \lambda^{3/2} e^{z/2H} \tag{11}$$

The $\lambda^{3/2}$ factor causes growth as $|u - c| \to 0$, and diminution as $|u - c|$ increases. This factor, as a rule, is competitive with the $e^{z/2H}$ growth only as $|u - c| \to 0$. Here, $\lambda^{3/2}$ can be viewed as adding to the $1/2H$ exponential growth rate an amount given by

$$\frac{1}{\lambda^{3/2}} \frac{d\lambda^{3/2}}{dz} = -\frac{3}{2} \frac{1}{(\bar{u} - c)} \frac{d\bar{u}}{dz} \tag{12}$$

Thus, one wants to determine the value of eddy diffusion that will cancel exponential growth with a rate

$$\frac{1}{2H} - \frac{3}{2} \frac{1}{(\bar{u} - c)} \frac{d\bar{u}}{dz} \tag{13}$$

The effect of eddy diffusion is approximated, following *Pitteway and Hines* [1963], by linear damping. Thus in the equations for temperature and horizontal momentum, diffusion contributes terms of the form

$$D_{eddy} \frac{\partial^2}{\partial z^2} \begin{Bmatrix} u \\ \delta T \end{Bmatrix} \approx -\lambda^2 D_{eddy} \begin{Bmatrix} u \\ \delta T \end{Bmatrix} \tag{14}$$

The presence of the right-hand side of (14) is equivalent to including an imaginary contribution to c given by

$$kc_i \approx \lambda^2 D_{eddy} \tag{15}$$

Such a contribution leads to an imaginary contribution to the vertical wave number λ:

$$\lambda \approx \frac{N(1 + l^2/k^2)^{1/2}}{\bar{u} - c} \approx \frac{Nk(1 + l^2/k^2)^{1/2}}{k(\bar{u} - c) + i\lambda_r^2 D_{eddy}}$$

where

$$\lambda_r \approx \frac{N(1 + l^2/k^2)^{1/2}}{(\bar{u} - c)}$$

Solving for λ_i we get the following approximate expression

$$\lambda_i \approx \frac{D_{eddy}\lambda_r^3}{k(\bar{u} - c)} \approx \frac{D_{eddy}N^3(1 + l^2/k^2)^{1/2}}{k(\bar{u} - c)^4} \tag{16}$$

The condition that eddy diffusion prevent the growth of $|d\delta T/dz|$ implied by (13) is simply

$$|\lambda_i| \approx \frac{D_{eddy}N^3(1 + l^2/k^2)}{k|\bar{u} - c|^4} \approx \frac{1}{2H} - \frac{3}{2}\left(\frac{1}{\bar{u} - c} \frac{d\bar{u}}{dz}\right)$$

or

$$D_{eddy} \approx \frac{k|\bar{u} - c|^4}{N^3(1 + l^2/k^2)^{3/2}} \left[\frac{1}{2H} - \frac{3}{2}\left(\frac{1}{\bar{u} - c} \frac{d\bar{u}}{dz}\right)\right] \tag{17}$$

Here, (17) is appropriate above $z = z_{break}$. Below z_{break} we still anticipate that gravity waves can generate turbulence [viz. *Lindzen,* 1971; *Weinstock,* 1976], but at a lower rate. While we will discuss (17) in more detail in the following section, two points should be mentioned immediately:

1. If $|\bar{u} - c| \gg |\bar{u}|$ (as is the case for the diurnal tide) then we expect D_{eddy} to be almost constant above z_{break} (until molecular viscosity becomes sufficient to damp the wave, a level *Lindzen and Blake* [1971] associate with the turbopause).
2. If $|\bar{u} - c| \to 0$ above z_{break}, then we expect D_{eddy} to decrease rapidly as the critical level (where $|u - c| = 0$) is approached.

The last item in our theoretical development is the determination of the deposition of wave momentum flux. In the context of the present discussion, this is given simply by $-d/dz(\rho_0\overline{uw})$, where u and w are the wave zonal and vertical velocity components. We already have an expression for w. Moreover, u is simply related to w. For the vertical wavelengths in Figure 1 we can simplify the continuity equation to

$$\frac{\partial v}{\partial y} + \frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} \approx 0 \approx \text{horizontal divergence} + i\lambda w \tag{18}$$

Now horizontal divergence $\approx ik(1 + l^2/k^2)u$ [*Lindzen,* 1970]. Thus

$$u \approx \frac{\lambda}{k(1 + l^2/k^2)} w$$

and

$$\overline{wu} \approx \frac{1}{2} \frac{\lambda}{k(1 + l^2/k^2)} |w|^2 \tag{19}$$

From (5) and (7) we have

$$|w| \approx g^{1/2} k(1 + l^2/k^2)^{1/2} A(\theta)\lambda^{1/2} e^{z/2H} \tag{20}$$

for $z < z_{break}$, and

$$\overline{wu} \approx \frac{1}{2} gkA^2(\theta) \frac{\lambda}{|\lambda|} e^{z/H} \tag{21}$$

for $z < z_{break}$.

Equation (21) is consistent with the Eliassen-Palm Theorem, which states that $d/dz(\rho_0\overline{uw}) = 0$ in the absence of forcing, dissipation, and critical levels. As noted in connection with (10), δT and w are known at z_{break}. It is readily shown that

$$\overline{wu}|_{z=z_{break}} \approx \frac{k}{2} \frac{N^2}{\lambda^3} \tag{22}$$

Above z_{break}, diffusion essentially eliminates the exponential growth shown in (21). Thus, above z_{break}

$$\overline{wu} \approx \left.\frac{k}{2} \frac{N^2}{\lambda^3}\right|_{\text{evaluated at } z_{break}} \tag{23}$$

Acceleration of the mean zonal flow due to wave breakdown is given by

$$-\frac{1}{\rho_0} \frac{d}{dz} \rho_0 \overline{uw} \approx \left.\frac{k}{2H} \frac{N^2}{\lambda^3}\right|_{z_{break}} \approx \left.\frac{-k}{2H} \frac{(\bar{u} - c)^3}{N(1 + l^2/k^2)^{3/2}}\right|_{\text{eval at } z = z_{break}} \tag{24}$$

The acceleration will be roughly constant between z_{break} and z_{crit} or some other height beyond which wave action ceases.

While the sign of the acceleration will always be such as to bring $\bar{u}$ toward c, it is clear that the process is not properly described by Rayleigh friction. It is also clear that the mean flow acceleration does not result from approximating the effect of D_{eddy} on the mean flow. The two effects are related but separate manifestations of wavebreaking. Note that the distribution of wave acceleration cannot be fixed in space but must be determined on the basis of the particular positions of z_{break} and z_{crit} at any given time.

3. Quantitative Considerations

To evaluate the formulae in section 2, information is needed concerning the phase speeds and wave numbers of the relevant waves. Such quantities are, indeed, well known for tidal modes [*Chapman and Lindzen,* 1970]. This section, therefore, will concentrate primarily on estimating these quantities for gravity waves.

a. *Phase Speeds*

The determination of phase speeds depends, fairly obviously, on the source of gravity waves. (For tides, thermal forcing is reasonably well known [*Lindzen,* 1979; *Forbes and Garrett,* 1979], and the phase speed is essentially the linear rotation speed of the earth, ~ 465 m/s at the equator.) While the possibility of stratospheric and mesospheric origins for mesospheric gravity waves (excluding tides for the moment) exists, no explicit evidence has been found. On the other hand, we do know of very important tropospheric sources, some of which are (1) flow over fixed surface features (mountains, cities, or 'heat islands,' etc.); viz. *Queney* [1948] and many others); (2) frontal processes [*Hines,* 1968; *Peltier and Ley,* 1978]; (3) Kelvin Helmholtz instability [*Lindzen and Rosenthal,* 1976].

The last two are associated with tropospheric flow speeds, while the first is associated with zero phase speed. However, owing to the unsteadiness of the troposphere, one expects phase speeds actually to be spread somewhat about the above speeds. An explicit study of the gravity wave spectra generated goes well beyond the scope of the present study.

We will simply assume a tropospheric origin for mesospheric gravity waves, and, in light of the above remarks, we will assume the phase speeds to be limited roughly to the range of tropospheric wind speeds (i.e., from some neighborhood of zero to some neighborhood of the tropospheric flow speeds shown in Figure 2).

As concerns the mesosphere, we have an additional (and crucial) constraint. Namely, gravity waves cannot readily pass through critical surfaces (where the flow speed and phase speed are equal). Assuming the zonal flows shown in Figure 3 to be appropriate, at least for discussion purposes (and recognizing that this is not likely to be accurately true at any given time), we see that in winter only phase speeds less than 10 m/s can be transmitted, while in summer only phase speeds somewhat in excess of 20 m/s can be transmitted, at least in the neighborhood of 40°. From Figure 2 we see that lower phase speeds would be permitted at other latitudes. Recall that we do not expect the generation of waves with either large negative phase speeds or with phase speeds much in excess of 20 m/s. Intuitively, we expect much greater generation of gravity wave energy with phase speeds near zero than with phase speeds greater than 20 m/s. This is consistent with the fact that z_{break} occurs at a much greater height in summer than in winter (viz. Figure 1). On the other hand, during sudden warmings, when easterlies develop in the stratosphere, we expect a significant reduction in the transmission of gravity waves to the mesosphere and a consequent increase in z_{break}. The seasonal variations in z_{break} caused by the seasonal variation in gravity wave filtering by mean winds provides a reasonable explanation of the observed seasonal variation of height of radio reflections from 50 km in winter to 70 km in summer [*Balsley et al.,* 1980]. The above results also appear consistent with the observation of enhanced D region ionization during sudden warmings [*Koshelev,* 1976].

b. *Horizontal Wave Numbers*

It is clear from (17) and (24) that our estimates of eddy diffusion and wave-induced acceleration depend on our choice of horizontal wave numbers. Vertical wave numbers are independent of specific choices of horizontal wave numbers, but they do depend on the ratio l^2/k^2.

In the case of the main diurnal propagating mode, wave numbers are, of course, well known. From the equivalent gravity mode formulation [*Lindzen,* 1970]

$$k = 1/6400 \text{ km}$$

and

$$l \approx 6.8k$$

Taking

$$u - c \approx -c \approx 466 \text{ m/s}$$

$$\frac{\partial T}{\partial z} + \frac{g}{c_p} \approx 7.5°/\text{km}$$

and

$$T \approx 250°\text{K}$$

we get from (4)

$$\lambda \approx \left(\frac{g}{T}\left|\frac{\partial T}{\partial z} + \frac{g}{c_p}\right|\right)^{1/2}\left(1 + \frac{l^2}{k^2}\right)^{1/2} \approx 2.5 \times 10^{-4} \text{ m}^{-1} \qquad (25)$$

corresponding to a vertical wavelength of about 25 km as expected.

The situation for gravity waves at middle and high latitudes is more ambiguous. What follows are to some extent 'guesstimates.' We will attempt to find a reasonable value of $(l/k)^2$ by matching the 'observed' vertical wave lengths with those predicted by (4). For both winter and summer

$$|u - c| \sim 50 \text{ m/s}$$

$$\frac{\partial T}{\partial z} + \frac{g}{c_p} \approx 7.5°/\text{km}$$

and

$$T \sim 250°\text{K}$$

Now we want

$$\lambda \sim \frac{2\pi}{12 \times 10^3 \text{ m}} \sim 5.236 \times 10^{-4} \text{ m}^{-1}$$

Equation (4) yields

$$\lambda \sim 3.4 \times 10^{-4} \text{ m}^{-1}\left(1 + \frac{l^2}{k^2}\right)^{1/2}$$

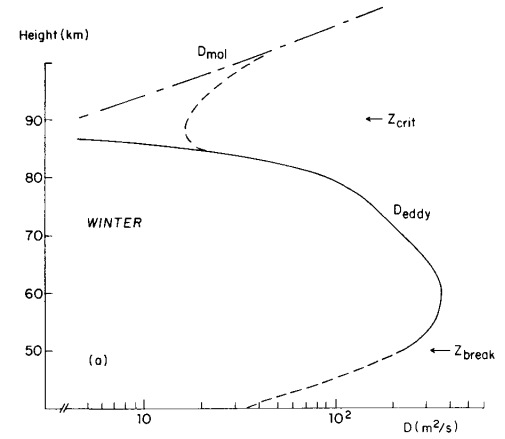


Fig. 5*a*

Fig. 5. Solid line shows predicted eddy diffusion coefficient based on breaking gravity waves. Simple dashed lines show schematic extrapolations as discussed in the text. The line consisting of alternating short and long dashes shows the molecular diffusion coefficient. (*a*) results for winter; (*b*) results for summer.

from which we obtain

$$\left(1 + \frac{l^2}{k^2}\right)^{1/2} \sim 1.54$$

or

$$\frac{l}{k} \sim 1.17 \tag{26}$$

While (26) ought not be taken too seriously, the fact that l and k are about equal is not implausible. As a practical matter, it should be noted that a modest increase in l/k can significantly decrease both D_{eddy} and acceleration.

The guidelines for determining k, itself, are comparably uncertain. There does exist a lower bound of sorts on k. Namely, in the presence of rotation, vertical propagation occurs only when the doppler shifted frequency $|k(\bar{u} - c)|$ exceeds the Coriolis frequency, $f = 2\Omega \sin\theta$ (where Ω = earth's rotation rate and θ = latitude). Now, through the bulk of the mesosphere (near 40° latitude) $|u - c| > 25$ m/s, and propagation will be possible for

$$k \gtrsim \frac{2\Omega \sin 40°}{25 \text{ m/s}} \approx 3.74 \times 10^{-6} \text{ m}^{-1} \tag{27}$$

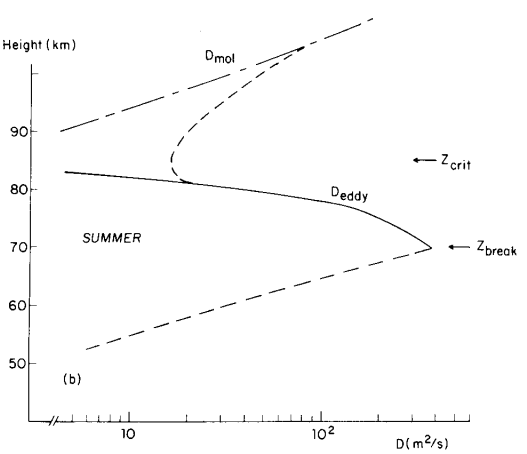


Fig. 5*b*

(40° is the approximate latitude for Figures 1 and 3). Finally, both surface topography and meteorological disturbances (both of which we have assumed are the basic sources of mesospheric gravity waves) tend to have red noise spectra. We will, therefore, take the lower bound for k given by (27) to also be the characteristic value of k. This value of k corresponds to a horizontal wavelength of 1680 km.

c. *Eddy Diffusion*

We evaluate (17) above z_{break} (appropriate to gravity waves) for the profiles of $\bar{u}$ and $\bar{T}$ shown in Figures 3 and 4. The results are shown in Figure 5. We have schematically continued D_{eddy} below z_{break} to indicate that nonbreaking gravity waves can still generate turbulence. We also show schematically an increase of D_{eddy} above the neighborhood of z_{crit}. The idea here is simply that while gravity waves may have a characteristic phase speed, c, they are not restricted to that phase speed. Weaker contributions at other phase speeds may break above that z_{break} appropriate to the bulk of the gravity wave energy. It is perhaps worth noting that the weaker gravity waves would break at a lower height if the bulk of the gravity wave spectrum were absent, since the weaker waves would not then be subject to damping by the turbulence generated by the stronger waves. Finally, one must recall that molecular diffusion increases as $1/\rho_0$ and reaches 0 (2×10^2 m^2/s) by about 110 km. Molecular processes, of course, tend to produce diffusive separation rather than turbulent mixing.

Turning to the explicitly calculated part of Figure 5, we see that for both summer and winter, D_{eddy} peaks near z_{break}. Here z_{break}, as determined from Figure 1, occurs near 50 km in winter and 70 km in summer. In contrast to most models for eddy diffusion, Figure 5 shows a decrease in D_{eddy} with height above the neighborhood of z_{break}. The more precipitous decrease in summer is simply associated with the more rapid approach of $\bar{u}$ to c in summer in the profiles shown in Figure 3. Whether such profiles are, in fact, applicable at any given time is open to question. In particular, during sudden warmings when an easterly flow regime sets on in the stratosphere, waves with $c = 0$ are effectively prevented from reaching the mesosphere and profiles for D_{eddy} should approach distributions similar to what is obtained for summer. This is what appears to be needed to produce enhanced D region ionization [*Koshelev*, 1976].

Interestingly, *Allen et al.* [1981] in a recent observation study of upper atmosphere composition, inferred a distribution of D_{eddy} very similar to the summer distribution in Figure 5. Their maximum value of D_{eddy} was about half of what is shown in Figure 5, but an increase in $(1 + l^2/k^2)^{1/2}$ of only 25% would eliminate this discrepancy.

Finally, we estimate D_{eddy} for the main propagating diurnal tidal mode. As already mentioned, the tide is expected to break at about 85 km. Since $c \cong -465$ m/s, $\bar{u}$ is too small to be of much significance. Thus (17) reduces to

$$D_{eddy} \approx \frac{kc^4}{N^3(1 + l^2/k^2)^{3/2}} \frac{1}{2H} \tag{28}$$

According to CIRA (1972), $\bar{T} \approx 200$°K at 85 km and $d\bar{T}/dz \approx 0$. With these values (28) yields

$$D_{eddy} \approx 1.83 \times 10^2 \text{ m}^2/\text{s}$$

The diminution of D_{eddy} with height due to the variation of $\bar{u}$ was not expected in this case. However, molecular viscosity

and conductivity rise rapidly in this region, and their role in damping the tide is similar to the role of D_{eddy}. Thus we must let D_{eddy} decrease with height so that the sum of molecular and eddy diffusion is given by the above value. This is shown in Figure 6. It should be repeated that although the roles of eddy and molecular diffusion in damping waves are essentially similar, the same cannot be said for their roles in transporting chemical constituents.

d. Wave-Induced Acceleration

As noted in section 2 and shown in Figure 1, $z_{break} \approx 50$ km in winter and $z_{break} \approx 70$ km in summer. Using the data in Figures 3 and 4 to evaluate (24) we obtain

$$-\frac{1}{\rho_0}\frac{d}{dz}\overline{\rho_0 uw} \approx 135 \text{ m/s/day}$$

for $z_{break} < z < z_{crit}$ in summer and

$$-\frac{1}{\rho_0}\frac{d}{dz}\overline{\rho_0 uw} \approx -102 \text{ m/s/day} \qquad (29)$$

for $z_{break} < z < z_{crit}$ in winter.

The values given by (29) seem large, and there is ample room for uncertainty. In particular, it was suggested in section 3*c* that it might be advisable to increase $(1 + l^2/k^2)^{1/2}$ 25% to reduce our estimates of D_{eddy} by a factor of 2. Such a change would also reduce the values in (29) by a factor of 2. That acceleration is constant in the layer $z_{break} < z < z_{crit}$ can, of course, only be approximately true. Total wave absorption in the neighborhood of the critical level, for example, implies the existence of a small neighborhood of the critical level where accelerations exceed the values given by (24).

Application of (24) to the diurnal tidal mode (using the same values as in section 3*c*) yields

$$-\frac{1}{\rho_0}\frac{d}{dz}\overline{\rho_0 uw} \approx -16.3 \text{ m/s/day} \qquad (30)$$

for $z_{break} < z \lesssim 110$ km. Recall $z_{break} \approx 85$ km while 110 km corresponds to a region within 30° latitude of the equator. However, as was noted by *Andrews and McIntyre* [1976], the detailed latitudinal distribution of the momentum deposition can be more complicated.

It was earlier remarked that the accelerations in (29) seemed 'large.' Clearly, such accelerations are not, as a rule, directly observed in the $\bar{u}$ field. However, it is not immediately clear whether the values in (29) are large compared to other contributors to the momentum budget. We might, for example, inquire what values of meridional velocity, $\bar{v}$, would allow Coriolis forces to balance the wave induced accelerations; that is,

$$2\Omega \sin\phi\, \bar{v} \sim 100 \text{ m/s/day}$$

or

$$\bar{v} \sim 0\left(\frac{100 \text{ m/s/day}}{2\Omega \sin\phi}\right)$$

$$\sim 0\left(\frac{8 \text{ m/s}}{\sin\phi}\right) \qquad (31)$$

Now according to CIRA (1972), such meridional winds are observed in the upper mesosphere (though one might question whether the observations really represent a zonal average). If this is so, then accelerations on the order of those given by (29) would be necessary to balance the 'observed' Coriolis torques.

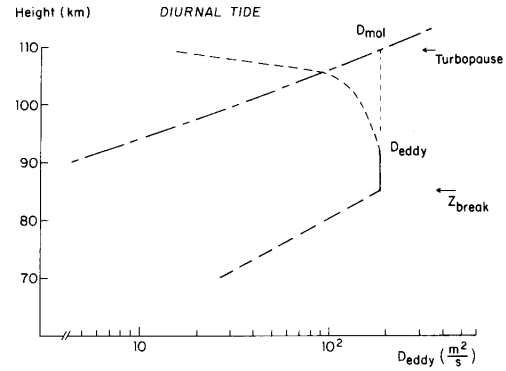


Fig. 6. Same as Figure 5, but for turbulence produced by the breaking of the diurnal tide.

Incidentally, the relatively small accelerations produced in the tropics could be balanced by the diffusion of momentum by the tidally induced eddy diffusion; this is not the case for the gravity wave-induced turbulence and acceleration in middle and high latitudes.

4. CONCLUDING REMARKS

For some years it has been accepted that tides and gravity waves propagating into the upper mesosphere from below are the major source of turbulence in the upper mesosphere. The purpose of the present paper has been to examine the implications of such a situation in some detail. The main propagating diurnal mode seems to be the primary contributor at tropical latitudes. Because of the high phase speed of this mode (−465 m/s), it is only slightly affected by the mean zonal flow of the atmosphere [*Lindzen,* 1971]. Wavebreaking appears to occur around 85 km, leading to a layer of enhanced eddy diffusion and wave induced acceleration (0(−16 m/s/day)) extending between 85 km and about 108 km above which height molecular transport dominates.

The situation at middle and high latitudes is more complicated. Gravity waves appear dominant at these latitudes, and if such waves originate in the troposphere (as seems likely) then their phase speeds will typically range from zero (appropriate to mountain waves, etc.) to typical tropospheric flow speeds. For such low phase speeds the effect of the mean zonal flow distribution will be crucial. The flow distribution will effectively determine which gravity waves (depending on phase speed) can reach the mesosphere and relatedly the amplitudes of mesospheric gravity waves and their breaking level. This effect, for example, implies that gravity wavebreaking will occur at greater altitude in summer than in winter (70 km versus 50 km). In addition, the turbulent diffusion and wave-induced accelerations no longer extend up to some level where molecular processes dominate. Instead, they are restricted to layers extending from the breaking level to some approximate critical level (where the mean flow, $\bar{u}$, equals or almost equals the wave phase speed). The lattr are likely to be developed by the wave-induced accelerations themselves. Between the critical levels and the thermosphere we expect a sharp minimum in eddy diffusion.

The wave-induced accelerations provide an explicit source for the 'friction' needed to reverse mesospheric shears and,

relatedly, to reverse the pole-to-pole temperature gradient at the mesopause. Models without such friction develop neither of these features [*Mahlman and Sinclair,* 1979]. The wave-induced acceleration acts to bring the mean flow toward the wave phase speed. In this respect the acceleration is similar to Rayleigh friction, provided the waves have zero phase speed. In contrast to Rayleigh friction the acceleration is not proportional to $(\bar{u} - c)$; indeed, it can increase as $(\bar{u} - c) \rightarrow 0$ (i.e., as a critical level is approached).

There are, in the present study, any number of shortcomings. For example, the existence of D_{eddy} below z_{break} requires that there be some momentum deposition below z_{break} in contrast to our restricting the deposition to $z \geq z_{\text{break}}$. Similarly, in a calculation where $A(\theta)$ is specified instead of z_{break} (viz. (10)), this will alter the value of z_{break}. It is not difficult to include such effects (at least approximately), but their inclusion seems unwarranted in the present paper, which is meant to be primarily illustrative. More problematic is the fact that our mathematical development considered only the vertical propagation of gravity waves (vis. (1)). A realistic application of the present concepts may require considering the propagation of gravity waves in a two-dimensional inhomogeneous medium. This, however, remains to be investigated. It is hoped that, for many purposes, a reasonable specification of $A(\phi)$, and the use of our present formulas, will suffice.

Acknowledgments. This research was supported in part by the National Science Foundation under grant ATM-78-23330 and by the National Aeronautics and Space Administration under grant NGL-22-007-228. Useful conversations with B. Balsley and U. von Zahn are also gratefully acknowledged.

REFERENCES

Allen, M., Y. L. Yung, and L. Waters, Vertical eddy diffusion, transport, and photochemistry in the terrestrial mesosphere and lower thermosphere, *J. Geophys. Res.*, this issue, 1981.

Andrews, D. G., and M. E. McIntyre, Planetary waves in horizontal and vertical shear: The generalized Eliassen-Palm relation and the mean zonal acceleration, *J. Atmos. Sci., 33,* 2031–2048, 1976.

Balsley, B. B., W. L. Ecklund, D. A. Carter, and P. E. Johnston, The MST radar at Poker Flat, Alaska, *Radio Sci., 15,* 213–223, 1980.

Bender, C. M., and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers,* McGraw-Hill, New York, 1978.

Chapman, S., and R. S. Lindzen, *Atmospheric Tides.,* D. Reidel, Hingham, Mass., 1970.

Fels, S, and R. S. Lindzen, Interaction of thermally excited gravity waves with mean flows, *Geophys. Fluid Dyn., 6,* 149–191, 1974.

Forbes, J., and H. B. Garrett, Thermal excitation of atmospheric tides due to insolation absorption by O_3 and H_2O, *Geophys. Res. Lett., 5,* 1013–1016, 1979.

Gossard, E. E., and W. H. Hooke, *Waves in the Atmosphere,* Elsevier, New York, 1975.

Hines, C. O., Internal gravity waves at ionospheric heights. *Can. J. Phys., 38,* 1441–1481, 1960.

Hines, C. O., A possible source of waves in noctilucent clouds, *J. Atmos. Sci., 25,* 937–942, 1968.

Hodges, R. R., Jr., Eddy diffusion coefficients due to instabilities in internal gravity waves, *J. Geophys. Res., 74,* 4087–4090, 1969.

Holton, J. R., *An Introduction to Dynamic Meteorology* (2nd ed.) Academic, New York, 1979.

Holton, J. R., and W. M. Wehrbein, A numerical model of the zonal mean circulation of the middle atmosphere, *Pure Appl. Geophys., 118,* 284–306, 1981.

Hunten, D. M., Vertical transport in atmospheres, in *Atmospheres of Earth and the Planets,* edited by B. M. McCormac, pp. 59–72, D. Reidel, Hingham, Mass., 1975.

Koshelev, V. V., Diurnal and seasonal variations of oxygen, hydrogen and nitrogen components at heights of mesosphere and lowere thermosphere, *J. Atmos. Phys., 38,* 991–998, 1976.

Leovy, C., Simple models of thermally driven mesospheric circulation, *J. Atmos. Sci., 21,* 327–241, 1964.

Lindzen, R. S., Thermally driven diurnal tide in the atmosphere, *Q. J. R. Meterol Soc., 93,* 18–42, 1967.

Lindzen, R. S., The application of classical atmospheric tidal theory, *Proc. R. Soc. London, Ser. A., 303,* 229–316, 1968*a*.

Lindzen, R. S., Lower atmospheric energy sources for the upper atmosphere, *Meteoral. Mono., 9,* 37–46, 1968*b*.

Lindzen, R. S., Internal gravity waves in atmospheres with realistic dissipation and temperature, I, Mathematical development and propagation of waves in the thermosphere. *Geophys. Fluid Dyn., 1,* 303–355, 1970.

Lindzen, R. S., Tides and gravity waves in the upper atmosphere, in *Mesospheric Models and Related Experiments,* edited by G. Fiocco, D. Reidel, Hingham, Mass., 1971.

Lindzen, R. S., Equatorial planetary waves in shear, II, *J. Atmos. Sci., 29,*1452–1463, 1972.

Lindzen, R. S., Wave-mean flow interaction in the upper atmosphere, *Boundary Layer Meteoral., 4,* 327–343, 1973.

Lindzen, R. S., Atmospheric tides, *Ann. Rev. Earth Planet. Sci., 7,* 199–225, 1979.

Lindzen, R. S., and D. Blake, Internal gravity waves in atmospheres with realistic dissipation and temperature, II, Thermal tides excited below the mesopause, *Geophys. Fluid Dyn., 2,* 31–61, 1971.

Lindzen, R. S., and A. Rosenthal, On the instability of Helmholtz velocity profiles in stably stratified fluids when a lower boundary is present, *J. Geophys. Res., 81,* 1561–1571, 1976.

Mahlman, J. D., and R. W. Sinclair, Recent results from the GFDL troposphere-stratosphere-mesosphere general circulation model, paper presented at IUGG-IAMAP XVII General Assembly, Canberra, Australia, 1979.

Peltier, W. R., and B. E. Ley, Wave generation and frontal collapse, *J. Atmos. Sci., 35,* 3–17, 1978.

Piteway, M. L. V., and C. O. Hines, The viscous damping of atmospheric gravity waves, *Can. J. Phys., 43,* 2222–2243, 1963.

Queney, P., The problem of airflow over mountains: A summary of theoretical studies, Bull. Am. Meteoral. Soc., 29, 16–26, 1948.

Theon, J. S., W. Nordberg, L. B. Katchen, and J. J. Horvath, Some observations on the thermal behavior of the mesosphere, *J. Atmos. Sci., 24,* 428–438, 1967.

U.S. Standard Atmosphere Supplements, U.S. Government Printing Office, Washington, D. C., 1966.

von Zahn, U., K. H. Fricke, D. M. Hunten, D. Krankowsky, K. Mauersberger, and A. O. Nier, The upper atmosphere of Venus during morning conditions, *J. Geophys. Res.,* in press, 1981.

Weinstock, J., Nonlinear theory of acoustic gravity waves, l, Saturation and enhanced diffusion, *J. Geophys. Res., 81,* 633–652, 1976.

(Received September 29, 1980;
revised December 1, 1980;
accepted January 19, 1981.)

Simulation of gravitational wave detectors

J. F. de Ronde,[a] G. D. van Albada, and P. M. A. Sloot
Faculty of Mathematics, Computer Science, Physics, and Astronomy, University of Amsterdam, 1098 SJ Amsterdam, The Netherlands

(Received 17 January 1997; accepted 11 June 1997)

A simulation program that provides insight into the vibrational properties of resonant mass gravitational radiation antennas is developed from scratch. The requirements that are set necessitate the use of an explicit finite element kernel. Since the computational complexity of this kernel requires significant computing power, it is tailored for execution on parallel computer systems. After validating the physical correctness of the program as well as the performance on distributed memory architectures, we present a number of ''sample'' simulation experiments to illustrate the simulation capabilities of the program. The development path of the code, consisting of problem definition, mathematical modeling, choosing an appropriate solution method, parallelization, physical validation, and performance validation, is argued to be typical for the design process of large-scale complex simulation codes. © *1997 American Institute of Physics.* [S0894-1866(97)01805-1]

INTRODUCTION

General background

Although the existence of gravitational radiation, predicted by Einstein's theory of general relativity,[1] is unquestioned, its detection is a long-standing problem in experimental physics. The aim of the GRAIL project[2] is to realize a spherical resonant mass detector with a sensitivity that is a few orders of magnitude higher than the present generation of detectors, thus offering the possibility of validating the existence of gravitational radiation that is emitted by astrophysical sources.[3] These sources include supernovae, stellar collapses to black hole states, and coalescence of binary neutron star systems.[4]

The quadrupole moment of a mass distribution $\rho(\mathbf{x})$ is given by

$$D_{ij}=\int_V dV\rho(\mathbf{x})\left(x_i x_j-\frac{1}{3}\,\delta_{ij}\mathbf{x}^2\right). \tag{1}$$

According to general relativity (GR), an oscillation of this quadrupole moment is the simplest mode of vibration that can generate gravitational waves. The counterstatement for absorption of gravitational waves is that in the simplest case it also takes place via the excitation of the quadrupole modes of vibration of a massive object. This principle is the main argument for constructing a gravitational radiation antenna, which essentially is a large resonant mass that is forced into oscillation by impinging gravitational radiation.

In GRAIL, the antenna will consist of a spherical resonant mass of a Cu alloy, possibly CuAl, with a mass of about 100,000 kg and a diameter of 3 m. It will be suspended in vacuum inside a large cryostat in such a way that the external vibrations at its resonant frequency ($\approx$700 Hz) are attenuated by at least a factor of 10^{16}. The sphere will be cooled to a temperature in the range of 10–20 mK in order to cancel out thermal noise. Transducers will be attached to the surface of the sphere to detect the vibrations induced by gravitational radiation. An analysis of the electric signals from these transducers must give information (such as source type and direction of incidence) about the gravitational radiation that interacts with the sphere. The gravitational waves have extremely weak interaction with matter; the typical deformations induced on a Cu sphere of 3 m will be of the order of 10^{-20} m, which explains why detection is such a difficult task and why it has not yet been realized. Several arguments can be given for using a spherically shaped antenna[5–8] instead of, for example, the cylinder or bar antennas in the pioneering days of gravitational radiation detection[9] and present-day experimental setups like the NAUTILUS detector at the Frascati National Laboratory in Rome.[10]

Since the estimated cost ($\approx$ 45 million Dutch guilders) of the project is quite high, a pilot study has been funded by NWO[11] to investigate the technological feasibility of GRAIL. In this project several critical aspects concerning the design of the antenna had to be addressed. The task of our simulation group was to construct detailed models of the antenna that provide insight into its vibrational properties. A schematic design of the GRAIL antenna concept is shown in Fig. 1.

The development and design of complex systems can be significantly enhanced by means of simulation. A recent example in which simulation was used in the analysis of the behavior of resonant mass antennas can be found in Ref. 8. In this work modal analysis was used to determine the frequency spectrum of a truncated icosahedral antenna. The accuracy of this simulation was fairly limited due to the low model resolution. The main result was that one can

[a]E-mail: janr@wins.uva.nl

 © 1997 AMERICAN INSTITUTE OF PHYSICS 0894-1866/97/11(5)/484/14/$10.00

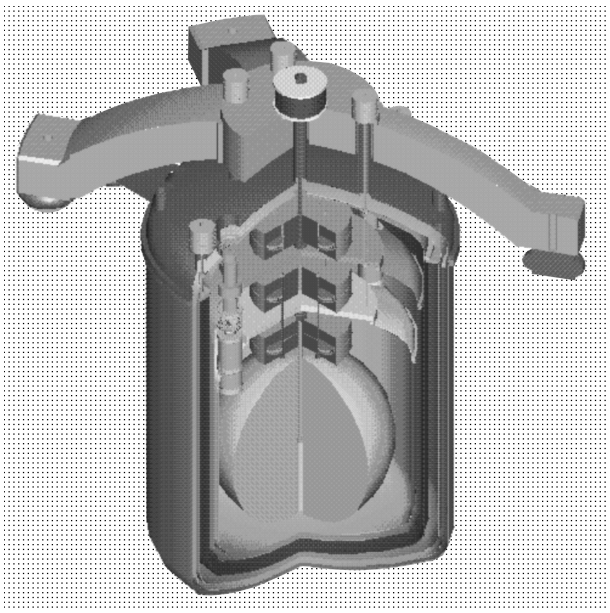


Figure 1. The GRAIL *antenna in a heat isolating vacuum chamber.*

conclude that the behavior of a truncated icosahedral antenna is very similar to that of a sphere.

In this article we will describe the development of a high-resolution simulation system for resonant mass gravitational radiation antennas that can be applied to a number of important design issues.

Design issues

The following questions need to be addressed by means of extensive simulation experiments.

(1) How does the suspension of the huge resonant mass affect the eigenmodes and eigenfrequency of the fundamental quadrupole modes of a sphere?
(2) How will material inhomogeneities affect these eigenmodes and eigenfrequencies?
(3) Given a perfect sphere, how will it deform under its own weight?
(4) Does the coupling between the suspension rod and the sphere induce additional modes that will cause interference with the modes we are looking for?
(5) What is the effect of mounting transducers on the antenna on the frequency spectrum of the system?
(6) How will seismic noise, entering the sphere through the framework in which it is suspended, influence the transducer signals?
(7) How will the fingerprint of typical gravitational wave sources be seen by the antenna?
(8) Can energy that has been deposited by cosmic rays or particles like muons induce system eigenmodes that are not distinguishable from those induced by gravitational waves?

We can identify two types of questions in this list. First, the design questions that need to be answered in order to tune the readout of the system, like the effects of the suspension hole, suspension rod, or material inhomogeneities on the frequency spectrum of the antenna. Second, we have those questions that pertain to the various sources of vibration: How can we discriminate whether the vibration that we observe is the result of a gravitational wave or of something else? Clearly a simulation that can help us in answering these questions can be of great value to the design and development phase of the GRAIL project. Eventually, when the system is operational, a (well-defined) simulation system can even be used to analyze the transducer output by means of reverse engineering of the transducer signal towards the original source. Incorporation of simulation in the experimentation cycle is sometimes referred to as ''living simulations'' or ''simulation in the loop.''

The remainder of this article is structured as follows. Section I begins with a description of two solution methods that can be used to answer a few basic questions but that are too restrictive for our overall goal. For this purpose we resort to development of a simulation that is based on the explicit finite-element (FE) method. The simulation accuracy required necessitates high-resolution FE models. The performance of the FE solver is therefore enhanced by exploiting the models' parallelism (Sec. II). Since the simulation is developed from scratch, we can circumvent several pitfalls that are known to hamper the migration of existing simulation codes to parallel computers.[12–14] In Sec. III the physical correctness of our simulation is validated by a comparison with an analytical solution as well as by experiments that were carried out with a small prototype antenna. It is further shown that we benefit greatly by exploiting computational concurrency. In Sec. IV we present a few typical simulation experiments that can be carried out with the simulation program. Some concluding remarks regarding the computational science approach taken in our study are given in Sec. V.

I. SIMULATION METHODS

A. Analytical solution

The equation of motion for elastic objects, also known as the Navier equation [Eq. (2)] describes the variation in time of the displacement field ($\mathbf{u}$) in homogeneous isotropic elastic objects. The material of such objects is parameterized by the two Lamé parameters (λ,μ) and the material density (ρ).

$$\mu\Delta\mathbf{u}+(\lambda+\mu)\nabla\nabla\cdot\mathbf{u}=\rho\ddot{\mathbf{u}}. \tag{2}$$

Two parameters that are frequently used to describe the elastic properties of a material are the Poisson ratio (ν) and Young's modulus (E). Both parameters can be expressed in terms of the Lamé parameters as follows:

$$\nu\equiv\frac{\lambda}{2(\lambda+\mu)}, \tag{3}$$

$$E\equiv\frac{\mu(3\lambda+2\mu)}{\lambda+\mu}. \tag{4}$$

The general solution of Eq. (2) consists of two different contributions that correspond to divergence-free ($\nabla\cdot\mathbf{u}=0$) and rotation-free ($\nabla\times\mathbf{u}=0$) waves that propagate through the material with unique velocities.

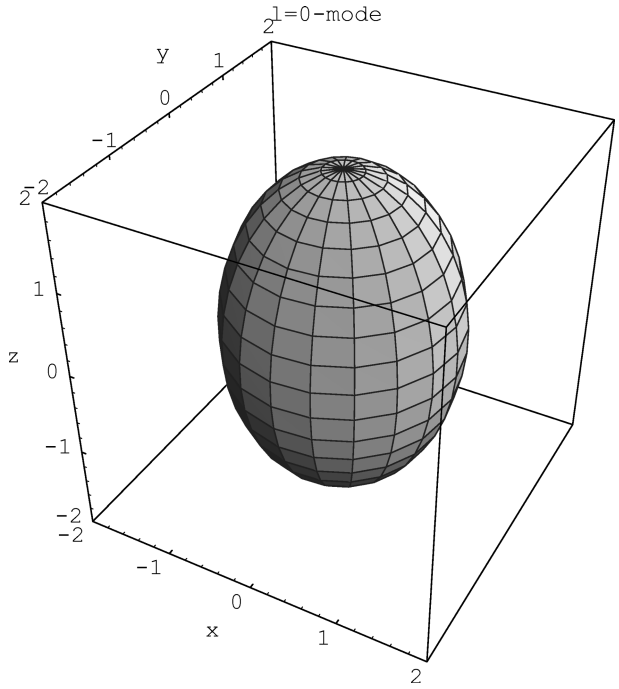


Figure 2. The prolate stage of the l=0 mode.

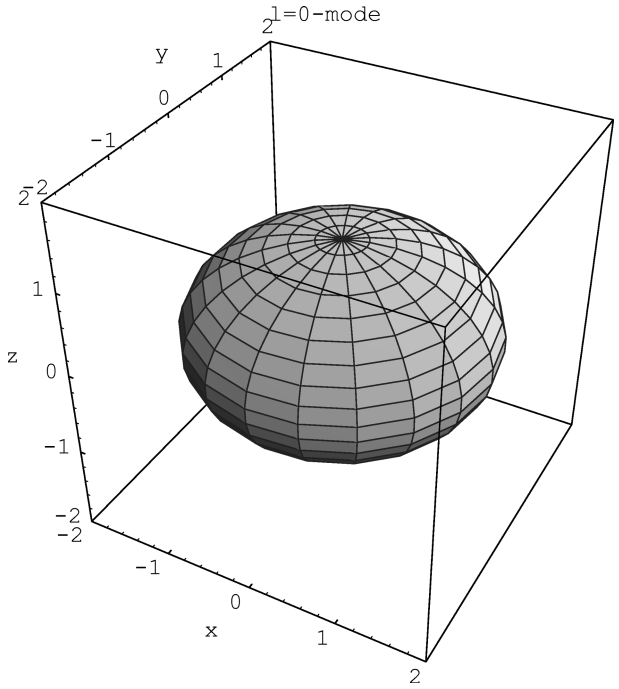


Figure 3. The oblate stage of the l=0 mode.

For both types of vibration a frequency equation can be derived for a freely vibrating (traction-free) elastic sphere and can be solved numerically.[15] The rotation-free vibrations are known as *spheroidal eigenmodes*. A general spheroidal eigenmode, Ψ^s, is given by Eq. (5), which is the superposition of all particular rotation-free solutions of Eq. (2) for a spherical object with traction-free boundary conditions; note that we use polar coordinates.

$$\Psi^s = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} [a_n(r)\hat{e}_r + b_n(r)R\nabla]Y_{nm}(\theta,\phi), \quad (5)$$

with R the sphere radius and $\hat{e}_r$ the unit vector in the radial direction; $a_n(r)$ and $b_n(r)$ are dimensionless radial eigenfunctions determined by the boundary conditions, and Y_{nm} are spherical harmonics. For each value of n the frequency equation contains infinitely many solutions. A general quadrupole mode of the sphere can be described as the superposition of the five modes with $n=2$ $(m=-2...2)$. Since (according to GR) only these modes interact with gravitational radiation, they are the primary subject of our investigations. Figures 2 and 3 show the shape of one of the quadrupole modes (denoted as the $l=0$ mode[16]) differing by half a period. This eigenmode oscillates between a prolate and an oblate shape.

The applicability of this analytical model fails as soon as we perturb the sphere in any way. However, it provides us with a useful gauge for a simulation program. First, we take a look at an elegant numerical model that is less restrictive than the analytical model described above.

B. Numerical method

If we write down the Lagrangian [Eq. (6)] for a freely vibrating elastic object (following Ref. 17) (the summation convention is used where appropriate)

$$L = \int_V \left(\frac{1}{2}\rho\omega^2 u_i u_i - \frac{1}{2} c_{ijkl} u_{i,j} u_{k,l} \right) dV, \quad (6)$$

with u_i the component of the displacement field $\mathbf{u}$ that is parallel to the i axis, $u_{i,j}$ the derivative of u_i in the j direction, c_{ijkl} the elastic tensor (dependent on λ and μ), and ω the eigen- (angular) frequency, and apply the principle of least action, setting $\delta L = 0$, we get

$$\delta L = \int_V (\rho\omega^2 u_i + c_{ijkl} u_{k,lj})\,\delta u_i dV - \int_S (\nu_j c_{ijkl} u_{k,l})\,\delta u_i dS = 0, \quad (7)$$

where ν_j denotes the component along the j axis of the outward-pointing normal vector on the surface of the object. Equation (7) exactly expresses the elastic wave equation (2), with traction-free boundary conditions, of Sec. I A. If we expand the displacement vector in a set of basis functions according to

$$\mathbf{u} = \sum_{lmn} \mathbf{a}_{lmn} x^l y^m z^n, \quad (8)$$

and truncate the expansion to a certain polynomial order, Eq. (7) can be rewritten in terms of a generalized eigenvalue problem (9).

Table I. The analytical and "numerical" frequencies in Hz for a few spheroidal modes of a Cu sphere. The dashes denote that no frequencies in the same range were found for specific N. This is due to the fact that the polynomial expansion in question is not able to represent the corresponding eigenmode properly.

n	Analytical solution	$N=5$ $R=168$	$N=6$ $R=252$	$N=7$ $R=360$	$N=8$ $R=495$	$N=9$ $R=660$	$N=10$ $R=858$
2	654.0828	654.090	654.090	654.085	654.085	654.085	654.085
1	891.7504	893.488	891.796	891.796	891.788	891.788	891.788
3	975.4565	988.374	975.550	975.550	975.464	975.464	975.464
4	1251.997	⋯	1281.98	1252.35	1252.35	1252.01	1252.01
2	1263.632	1270.78	1270.78	1263.78	1263.78	1263.68	1263.68
0	1395.698	1396.06	1396.06	1396.06	1396.06	1396.06	1396.06
5	1510.716	⋯	⋯	⋯	1511.6	1511.65	1510.83
3	1661.032	⋯	1681.20	1681.20	1661.6	1661.63	1661.04
6	1760.268	⋯	⋯	⋯	⋯	1762.42	1762.42
1	1793.638	⋯	1843.70	1843.70	1795.91	1795.81	1793.67

$$\omega^2 E\mathbf{a}=\Gamma\mathbf{a}. \tag{9}$$

The matrices E and Γ can be computed by integrating Eq. (7) properly. Solving for ω provides us with estimates of the eigenfrequencies of the system in question.[16,17]

1. Validation

Table I shows several eigenfrequencies found with the numerical method as well as the corresponding analytical values for a Cu sphere (material parameters: $\rho=8933$ kg/m^3, $E=129.810^9$ Pa, and $\nu=0.343$). N denotes the truncation order of the basis function expansion, and $R\times R$ is the matrix size of the corresponding eigenvalue problem. We have used a LAPACK[18] routine called DSYGV, which can be applied to symmetric generalized eigenvalue problems to solve the system.

2. Spheroidal perturbations

An ellipsoid is described by its three semi-axes d_x, d_y, and d_z:

$$\frac{x^2}{d_x^2}+\frac{y^2}{d_y^2}+\frac{z^2}{d_z^2}=1. \tag{10}$$

For a sphere these axes have equal length. When we make one semi-axis longer and keep the other two constant, the object is called a prolate spheroid. If we make one semi-axis shorter than the other two, we have an oblate spheroid. In Fig. 4 the dependence of the lowest quadrupole frequency on d_z is depicted [varied in the range (0.1, 5.0) m], while the other two axes are kept at 1.5 m for a Cu sphere. We can observe three lines that intersect at $d_z=1.5$. Since the spherical symmetry is broken in the z direction, the fivefold degeneracy is only partially removed. It splits into two doublets and one singlet. Additional perturbations in the x and y directions would break the symmetry completely, leaving us with five slightly different frequencies.

C. A finite-element model

The numerical method of Sec. I B also was applied to a sphere with a suspension hole. Again it can be observed that the five frequencies are split up (see Fig. 5). It leaves us with an uncomfortable feeling, however, since the sharp edges of the suspension hole possibly disrupt the numerical

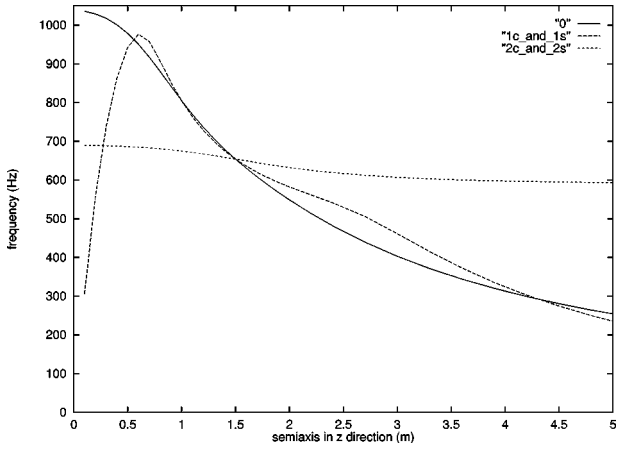


Figure 4. The fundamental quadrupole frequencies under variation of d_z. The spherical symmetry is broken in the $\hat{z}$ direction, resulting in a threefold splitting of the fivefold degenerate quadrupole frequency: two doublets and one singlet. The doublets are denoted by 1c_and_1s and 2c_and_2s and the singlet by 0.

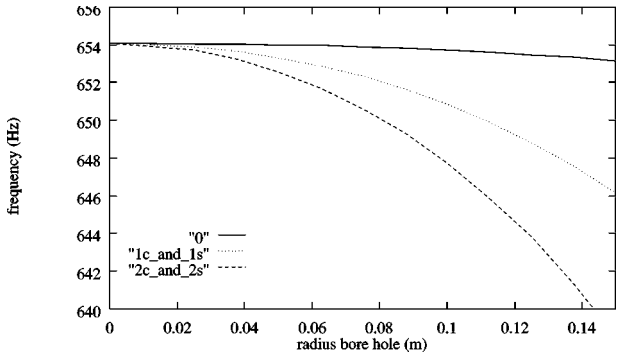


Figure 5. The fundamental quadrupole frequencies vs the radius of the suspension hole. The spherical symmetry is broken in the $\hat{z}$ direction, resulting in a threefold splitting of the fivefold degenerate quadrupole frequency: two doublets and one singlet. The doublets are denoted by 1c_and_1s and 2c_and_2s and the singlet by 0.

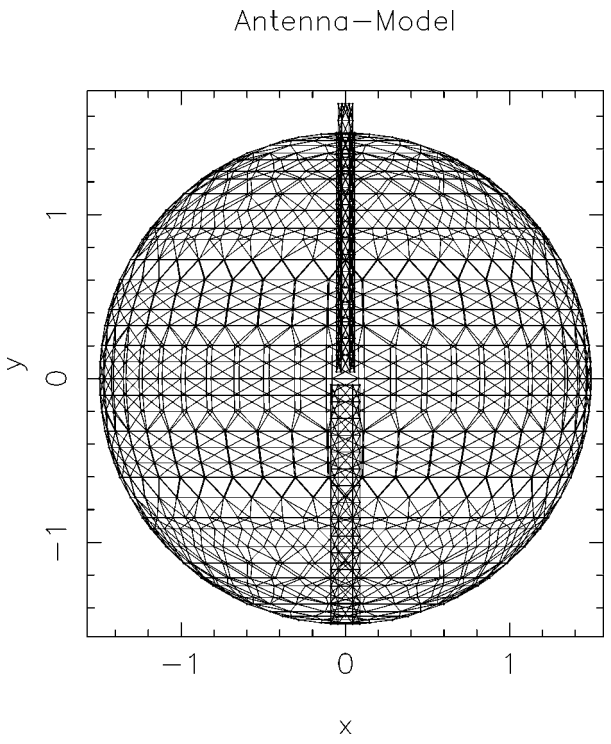


Figure 6. The hull of a finite-element antenna model with the suspension rod discretized into 39,684 tetrahedral elements.

process. Therefore, we do not know whether the frequency splitting is an artifact or is physically correct. Furthermore, we are stuck with a variety of questions that cannot be answered by the analytical and numerical approaches above. Next we will consider the finite-element method as a third alternative to answer our questions. We discretize our antenna model in first-order tetrahedral elements (each element has four nodal points). As an example, Fig. 6 shows the hull of such a finite-element model, consisting of 39,684 tetrahedral elements. One approach to the problem is by means of modal analysis, that is, by studying the following eigenvalue problem:

$$K\mathbf{d}=\omega^2 M\mathbf{d}, \tag{11}$$

with K the system's stiffness matrix, M the mass matrix, $\mathbf{d}$ an eigenmode, and ω an eigenfrequency of the system. Although it is theoretically possible to investigate a linear system such as ours on the basis of modal analysis, its practical realization is a different story. First of all, the model resolutions that we aim for are very high, necessitating solution of huge eigenvalue problems. These may be solvable using state-of-the-art numerical solvers. If we succeed in doing that, we have to cope with a storage problem, since for each nodal point a superposition of at least over 100 eigenmodes is necessary to characterize a general low-frequency solution of the system. In addition, it is difficult to model the effects that are due to coupling to external forces that are aperiodic, like seismic noise entering via the suspension or stochastic sources.

On the other hand we can choose for explicit time integration of the equation of motion (12).

$$M\mathbf{a}+C\mathbf{v}+K\mathbf{d}+\mathbf{F}=0. \tag{12}$$

The matrices M, C, and K are constructed by assembling the local mass, the viscosity, and the stiffness matrices of all individual finite elements in which the antenna model is divided; **a**, **v**, and **d** are vectors that describe, respectively, the acceleration, velocity, and displacement fields, and **F** is an external force vector. In order to integrate this system explicitly in time we adopt the following Newmark scheme:[19]

$$\begin{aligned}\mathbf{d}_{n+1}&=\mathbf{d}_n+\mathbf{v}_n\Delta t+\mathbf{a}_n\frac{(\Delta t)^2}{2},\\ \mathbf{a}_{n+1}&=M^{-1}K(\mathbf{d}_{n+1}),\\ \mathbf{v}_{n+1}&=\mathbf{v}_n+(\mathbf{a}_n+\mathbf{a}_{n+1})\frac{\Delta t}{2}\end{aligned} \tag{13}$$

with $\mathbf{a}_n$, $\mathbf{v}_n$, and $\mathbf{d}_n$ vectors that describe, respectively, the acceleration, the velocity, and the displacement fields in the FE system at time n. Viscous damping (C) and external forces (**F**) can easily be added to this algorithm.[19–21] The masses are ''lumped'' at the nodal points, resulting in a diagonal mass matrix. Therefore M^{-1} can be calculated quickly. For numerical stability Δt must be less than the Courant value, since the time integration scheme is conditionally stable. The choice of the algorithm is, on the one hand, motivated by the fact that it displays good numerical properties (low numerical dispersion and damping). This was found through a number of exploratory experiments on a one-dimensional finite-element model (data not shown). On the other hand, it is very well suited for exploiting computational concurrency, as we will see below. Although this explicit approach has the drawback that we have to solve our system in the time domain and afterwards perform a spectral analysis, it is much more flexible for our purposes, and is not hampered by the restrictions that apply to modal analysis. Furthermore, the simulation model can be extended such that it can incorporate nonlinear effects, such as *violin modes* in the suspension rod, that cannot be realized by modal analysis.

It was expected that the required simulation accuracy necessitates the use of a very-high-resolution FE model that is computationally and memory intensive. The simulation code was therefore designed and implemented for execution on distributed-memory computers.

II. PARALLELIZATION

In good tradition, we have chosen to base our parallelization method on the decomposition of the FE mesh. The consequences of the decomposition method for the time integration algorithm as well as the calculation of important system parameters (like energy) had to be considered in the design. In our discussion we will only consider the pure finite-element system. In the full-fledged simulation a variety of phenomena can be taken into account, such as transducer mounting, suspension rod, and sources of vibration. At this point it suffices to say that the parallelization method chosen does not have major consequences for the full-fledged simulation, and therefore may be left out of this discussion. We focus on two important aspects of the parallelization process: parallel time integration (the kernel) and parallel calculation of the energy.

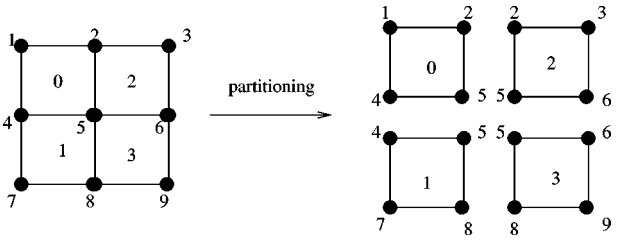


Figure 7. Partitioning of a quadrilateral mesh into four parts. The nodal points at the shared boundaries are replicated while elements are assigned to unique processes.

The code was implemented in C with PVM[22] message passing primitives according to the single-program, multiple-data paradigm.[23] The problem domain (the finite-element mesh consisting of N tetrahedral elements) is partitioned into P subdomains. Each subdomain is assigned to a single process and executes the same program applied to its local data. The process with identifier 0 is used to read in the whole mesh; it supplies the other $P-1$ processes with their locally essential data, and is used for input/output (I/O), a *primus inter pares* or *master* among the parallel processes.

A. Energy

In order to check energy conservation, we have to monitor the energy in the system. The total energy E_{tot} is given by

$$E_{\text{tot}} = \tfrac{1}{2}(\mathbf{v}^T M_{\mathbf{v}} + \mathbf{d}^T K \mathbf{d}), \tag{14}$$

with $\mathbf{v}$ the velocity vector in which the individual velocity vectors of all nodal points are concatenated, and $\mathbf{d}$ the displacement vector, also formed by concatenating the fields at all nodal points. How do we calculate this in parallel?

The FE mesh is decomposed along the surfaces of the tetrahedral elements that constitute the mesh. First, a dual graph of the mesh is constructed; this then can be partitioned using any graph-partitioning method.[24] Nodal points that lie on shared boundaries are replicated on each process with this boundary in its local domain, whereas elements belong to a unique process. This procedure is schematically depicted in Fig. 7 for a simple quadrilateral mesh.

The global mass and stiffness matrices are constructed by assembling the individual contributions from each element in a global matrix. If we denote the assembly of two-element stiffness matrices k_1 and k_2 by $k_1 \oplus k_2$, K is constructed (analogously for the mass matrix) as follows:

$$K = k_1 \oplus k_2 \oplus \ldots \oplus k_N, \tag{15}$$

with N the number of tetrahedral elements. If we monitor the energy values during a simulation run (it would be nice if this could be done completely separate), that is, for each subdomain (i), the energy content (E^i_{tot}) is calculated, and once in a while we assemble the subdomain energy contributions and accumulate them on process 0. This can be done as follows:

$$E^i_{\text{tot}} = \tfrac{1}{2}(\mathbf{v}^{\mathbf{i}T} M^i_r \mathbf{v}^{\mathbf{i}} + \mathbf{d}^{\mathbf{i}T} K_i \mathbf{d}^{\mathbf{i}}), \tag{16}$$

$\mathbf{v}^{\mathbf{i}}$ denotes the velocity vector for the nodes in subdomain i, $\mathbf{d}^{\mathbf{i}}$ is the displacement field, K_i is the stiffness matrix of subdomain i, which is simply the assembly of its local element stiffness matrices, and finally M^i_r is an adjusted version of the (diagonal) mass matrix of subdomain i. The ''normal'' mass matrix contains the mass for each nodal point at the appropriate entry. In this adjusted version the masses of the internal nodal points are untouched. However, for the shared nodal points this value is divided by its ''redundancy'' value, which denotes how many processes have a replica of it. For example, in Fig. 7 nodal point 5 has a redundancy value of 4. As a consequence, the corresponding entry in each local adjusted mass matrix, either being on process 0, 1, 2, or 3, has to be divided by 4. In this way we can assure that we are not counting any energy contributions twice.

B. Time evolution

If we have constructed the mass and stiffness matrices for each subdomain (or process) correctly, we can parallelize the Newmark scheme above as follows: Each process gets a unique process identifier (ID) equal to the subdomain identifier (i) above. Nodal points that are shared among two or more processes are considered to have one real owner (the owning process considers it as a node of type A), while the other processes only have the replica of this node (considering the node as type B). Assigning a node to a specific process is based on the process ID. The process with the lowest ID considers a shared nodal point as A type, while the other processes see it as B type. Consequently, process 0 can only have shared nodes of type A, while process $P-1$ (assuming we have P processes) can have only shared nodes of type B. The following construction parallelizes the Newmark scheme correctly:

$\mathbf{d}_{n+1} = \mathbf{d}_n + \mathbf{v}_n \Delta t + \mathbf{a}_n (\Delta t)^2/2$,
Send $\mathbf{d}_{n+1}$ from A nodes to B nodes,
Receive $\mathbf{d}_{n+1}$ for B nodes from A nodes,
$\mathbf{a}_{n+1} = M^{-1} K(\mathbf{d}_{n+1})$,
Send $\mathbf{a}_{n+1}$ from B nodes to A nodes,
Receive $\mathbf{a}_{n+1}$ for A nodes from B nodes,
Add contributions to $\mathbf{a}_{n+1}$ from B nodes,
$\mathbf{v}_{n+1} = \mathbf{v}_n + (\mathbf{a}_n + \mathbf{a}_{n+1})\Delta t/2$.

The A–B nodal point scheme was adopted from Lonsdale *et al.*[25]

III. VALIDATION

Here we show a selection of experimental results that were produced for validation of our parallel FE simulation. First, we investigate whether the model is physically correct. Next, we measure the performance benefits that we get from exploiting parallelism for different mesh resolutions. Meshes are decomposed by two different graph-partitioning methods: recursive spectral bisection (RSB) and a generalized version of orthogonal recursive bisection (ORB),[24] which we call *generalized* ORB (GORB). GORB operates in such a way that consecutive partitionings are not necessarily orthogonal and that partitioning cardinalities are not necessarily a power of 2 (as is the case in ORB). Figure 8 shows partitioning in the case of GORB in four parts, where the consecutive partitionings are not performed orthogonally but at an angle of 80°. GORB recursively bi-

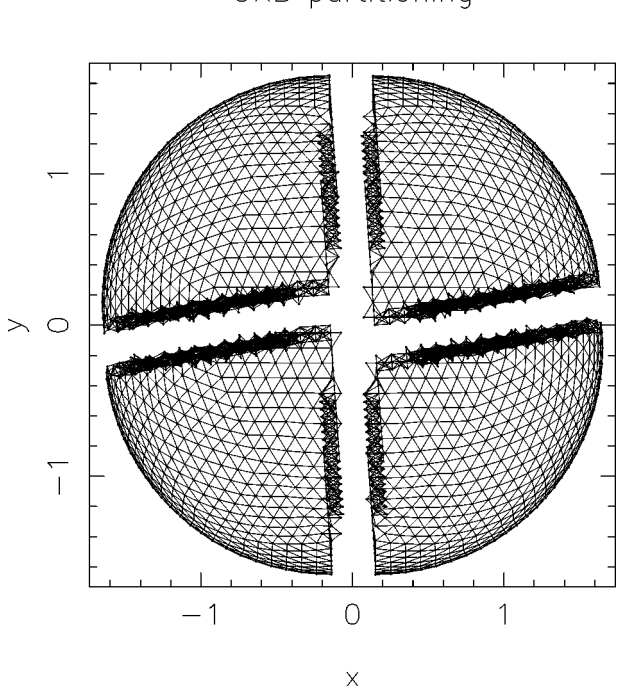


Figure 8. The hulls of the four partitions obtained for a sphere consisting of 79,556 tetrahedral elements obtained using the GORB *partitioning method. Each of the four hulls has about 3140 triangular faces. The first bisection is along the* $\hat{x}$ *direction. Next, the bisection surface is rotated 80° and the two subdomains are both bisected individually along this second direction.*

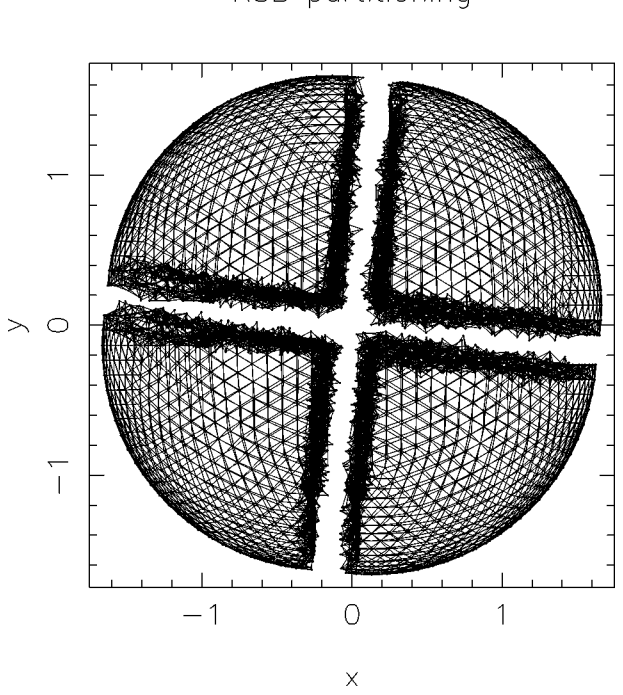


Figure 9. The hulls of the four partitions obtained for the same sphere as in Fig. 8, but now using the RSB *partitioning method. Although the edges in this partitioning appear more ragged than those obtained using* ORB, *they have only about 3010 faces.* RSB *optimizes the connectivity within partitions and minimizes the communication between them.*

sects the geometrical domain in equal partitions (work load balancing), in alternating directions, not paying any attention to the mesh connectivity (communication). RSB applies a more complex, and consequently computationally more expensive, method to partition the mesh. In contrast with (G)ORB it also takes into account the mesh connectivity. It recursively bisects a finite-element mesh into equal numbers of elements (work load balancing), while (approximately) minimizing the number of edges between the resulting subdomains (communication load balancing). Figures 8 and 9, respectively, show the partitioning in the cases of GORB and RSB into four parts, with both bisections applied to a FE sphere of 79,556 elements. The difference in partitioning quality between GORB and RSB will have to be determined by explicitly measuring the kernel execution times because it cannot be found by simple visual inspection.

The performance measurements are carried out on two target platforms both installed at the IC^{3}A in Amsterdam;[26] a Parsytec PowerXplorer [32 PowerPCs 601, with each 32 Mbytes of random access memory (RAM)] and a Parsytec CC (40 PowerPCs 604, each having 96 Mbytes of RAM). Our target material in the experiments below is the alloy CuAl (90–10) (with $\rho = 7534.0$ kg/m^3, $E = 154.0 \times 10^9$ Pa, and $\nu = 0.3265$), and only considers spheres with radii of 1.5 m. All results with respect to the physical experiments given below are independent of the number of processors used, which indicates the correctness of our parallel implementation.

A. Physical correctness

Let us consider simulations that are initiated with a pure quadrupole mode, that is, each nodal point in the system gets an initial velocity $\mathbf{v}^T = \dot{h}(x, -y, 0)$, with $\dot{h}$ the time derivative of h, a scale-free parameter denoting the deformation of space. Since the FE models are always slightly asymmetric, the fivefold degeneracy will be removed. Table II enumerates the average frequencies that are found around the analytical eigenfrequency f_0 ($f_0 = 780.25$ Hz)

Table II. The average frequency $\langle f \rangle$ and the approximated error $\sigma(f)$ in % for the quadrupole principal mode for various model resolutions.

n	$\langle f \rangle$	$\sigma(f)$ %
84	7.91e+02	8.72e−01
124	7.92e+02	1.70e+00
698	8.01e+02	4.87e−01
1086	7.97e+02	4.44e−01
2724	7.90e+02	3.23e−01
2851	7.96e+02	3.62e−01
6114	7.86e+02	2.25e−01
9412	7.85e+02	3.01e−01
51501	7.82e+02	2.85e−01
79556	7.82e+02	1.56e−01

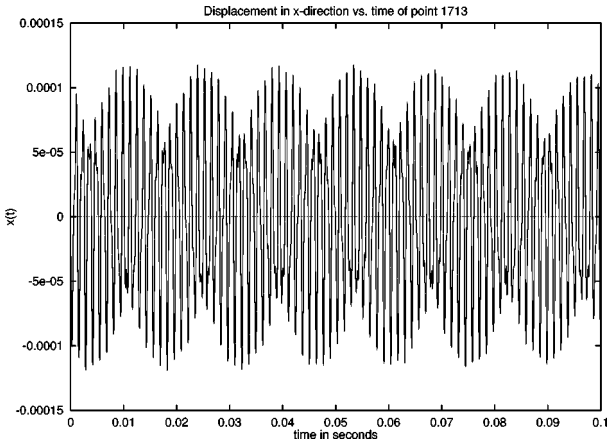


Figure 10. The time trajectory in the $\hat{x}$ direction of an arbitrary nodal point at the surface of a FE model of 79,556 elements initiated with a pure quadrupole velocity field.

and the deviation in % of the average value for 10 different FE models that consist of n tetrahedral elements. These results indicate that the dominating frequency for the FE models approaches the analytical frequency value for increasing model resolution (up to 0.16% from the analytical value for the highest resolution model).

Figure 10 shows the $\hat{x}$ displacement of an arbitrary nodal point at the sphere surface in the same simulation for the highest-resolution model. At first glance, it seems that the nodal point oscillates with a frequency of about 780 Hz, with a signal of $\approx$70 Hz superimposed. However, Fourier analysis (data not shown) has pointed out that this signal mainly results from superposition of the three oscillations that are most pronounced, that is, those oscillating at 781.8, 1493, and 1499 Hz, respectively. Possible beats that result from superposition of the principal quadrupole modes (around 780 Hz) are not visible in this picture.

We can also observe the system energy during simulation. It can be derived that the analytical energy (E_{quad}) for a sphere vibrating in a pure quadrupole mode can be expressed in terms of the radius (R) and the material density (ρ) as follows:

$$E_{\text{quad}} = \tfrac{4}{15}\,\pi\rho R^5 \dot{h}^2. \tag{17}$$

For the CuAl sphere with a radius of 1.5 m it follows that $E_{\text{quad}} = 47{,}929.2$ Nm. Within the finite-element simulations, it was found that the energy approaches this value very well for increasing model resolution (data not shown). For instance, the total energy for the highest-resolution model of 79,556 elements is approximately equal to 47,924.7, which approaches the analytical value up to 0.01%.

Within the GRAIL project, a prototype spherical antenna of the same CuAl alloy as that above, but with a radius of 0.075 m was fabricated. The following experiment was carried out. The prototype is excited by a radial impulse. Next, the time series of radial vibrations are recorded at several (arbitrarily chosen) places on the antenna surface. From these time series we have derived the power spectrum of the antenna vibrations by alternately applying a fast Fourier transform (FFT) and a nonlinear least-squares fitting procedure on the signal residue. Simultaneously, a simulation experiment was performed in which the same procedure was carried out. That is, a radial displacement is induced on an arbitrary surface point of a detailed finite-element model consisting of 191,310 tetrahedral elements. The model is realistic in the sense that it incorporates the suspension rod and the suspension hole. For several points at the model surface we record the time series, analogous to the real experiment, followed by the same procedure to produce a power spectrum. Table III lists a part of the frequency spectrum that is obtained for a radial impulse on the prototype antenna, accompanied by the frequencies found from the finite element simulation. The third column in Table III corresponds to the ratio of model frequency to experimental frequency. We can observe that the frequency spectrum found in the simulation model is quite close to experimental values. The residue of the experimental signal and the simulated signal turns out to be less than 0.1%.

Table III. Matches between frequencies computed using the explicit finite element code and experimental values for a 15 cm sphere that is excited by a purely radial excitation. In both sets of data, mode excitations may have been below the detection threshold, explaining frequencies occurring in a single set only. The analytical frequencies that are found for a pure sphere with the same material specifications are (within the same frequency range) 14830 (2,1), 15,708 (2,2), 21,243 (1,2), 22,915 (3,1), 23,407 (3,2), 30,025 (4,2) Hz, with, for instance, 14,830 (2,1) denoting a mode vibrating with 14,830 Hz, $n=2$ of toroidal type (1), and 15,708 (2,2) a mode vibrating with 15708 Hz, $n=2$ of spheroidal type (2).

Computed (Hz)	Measured (Hz)	Computed/ measured
13,650	13,687	0.997
13,740	13,720	1.001
	13,797	
13,849	13,815	1.002
14,391	14,424	0.998
14,500	14,470	1.002
14,626	14,587	1.003
19,580		
19,689	19,690	1.000
19,797	19,812	0.999
21,045		
21,171	21,138	1.002
21,207	21,289	0.996
21,316	21,307	1.000
	21,343	
	21,371	
21,623	21,647	0.999
	21,667	
21,750	21,777	0.999
21,840	21,809	1.000
27,716	27,719	1.000
	27,739	
	27,802	
27,843	27,863	0.999
	27,919	
	27,954	
	27,972	
27,987	27,991	1.000

Table IV. The kernel execution times in seconds on the PowerXplorer, for four different mesh sizes and two different decomposition methods (GORB denoted by suffix G and RSB denoted by suffix R) versus the partitioning cardinality.

	6114-G	6114-R	9412-G	9412-R	51501-G	51501-R	79556-G	79556-R
1	3.46e−01	3.46e−01	5.25e−01	⋯	⋯	⋯	⋯	
4	1.47e−01	1.45e−01	2.05e−01	2.02e−01	8.89e−01	8.85e−01	1.31e+00	1.30e+00
6	1.58e−01	⋯	2.02e−01	⋯	6.95e−01	⋯	1.01e+00	⋯
8	1.22e−01	1.13e−01	1.77e−01	1.72e−01	5.98e−01	6.00e−01	7.65e−01	7.84e−01
9	1.52e−01	⋯	2.17e−01	⋯	5.99e−01	⋯	7.80e−01	⋯
12	1.25e−01	⋯	1.84e−01	⋯	5.14e−01	⋯	6.55e−01	⋯
16	1.10e−01	1.04e−01	1.64e−01	1.63e−01	5.00e−01	4.53e−01	6.34e−01	6.41e−01
18	1.02e−01	⋯	1.57e−01	⋯	4.78e−01	⋯	6.71e−01	⋯
24	7.99e−02	⋯	1.31e−01	⋯	4.61e−01	⋯	6.93e−01	⋯
27	8.01e−02	⋯	9.75e−02	⋯	4.72e−01	⋯	6.59e−01	⋯
32	7.00e−02	8.58e−02	9.07e−02	1.21e−01	4.42e−01	4.59e−01	6.11e−01	6.36e−01

B. Parallel performance

Next we consider the parallel performance of the simulation kernel. We apply GORB (partitioning cardinalities of 2, 4, 6, 8, 9, 12, 16, 18, 24, 27, and 32) and RSB (partitioning cardinalities of 4, 8, 16, and 32) to four different meshes consisting of 6114, 9412, 51,501, and 79,556 elements, respectively. Table IV displays the single-kernel execution times (one time step in the Newmark scheme) that were obtained for simulations on the PowerXplorer for a varying number of processors. Table V shows similar results obtained on the Parsytec CC. The execution times are the best case, that is, energy and external force calculations (modeling gravitational radiation sources) are left out.

In both Tables IV and V we can observe that the difference in execution speed between GORB and RSB is not significant. In fact GORB is often slightly better than RSB for partitioning cardinalities that are powers of 2. This can be ascribed to the fact that both partitioning methods create compact three-dimensional subdomains that do not contain any disconnected parts, and this leads to an average surface–volume ratio (or, equivalently, communication–calculation ratio) of these subdomains that is approximately the same.

The working memory of the CC processors is much larger than that of the PowerXplorer. Therefore on the CC all problems can be executed on a single processor, in contrast to the PowerXplorer, where single-processor performance cannot be determined for high-resolution problems. The CC communication network and CPU are significantly faster than those of the PowerXplorer, which accounts for the fact that the kernel execution time for each problem is shorter on the CC.

For the PowerXplorer the kernel execution times are not significantly decreased for more than 16 partitions, whereas in case of the CC increasing partitioning cardinality consequently leads to faster execution times, at least for high-resolution models. Since we have considered the best performance, we can expect that the absolute performance of the simulation code in a more realistic simulation (taking other phenomena into account) will be degraded compared to the best case and, consequently, result in a relatively better ''scalability'' of the code, since the amount of work per subdomain increases whereas the communication pattern and size remain unchanged.

IV. "SAMPLE" SIMULATION EXPERIMENTS

Since the main theme of this article is the concept of a complex parallel simulation program, we will not present detailed quantitative results with respect to the feasibility of

Table V. The kernel execution times in seconds on the CC for four different mesh sizes and two different decomposition methods (GORB denoted by suffix G and RSB denoted by suffix R) versus the partitioning cardinality.

	6114-G	6114-R	9412-G	9412-R	51501-G	51501-R	79556-G	79556-R
1	1.38e−01	1.38e−01	2.10e−01	2.10e−01	1.12e+00	1.12e+00	1.71e+00	1.71e+00
2	8.68e−02	⋯	1.27e−01	⋯	6.23e−01	⋯	9.20e−01	⋯
4	5.41e−02	5.34e−02	7.63e−02	7.48e−02	3.47e−01	3.39e−01	4.99e−01	4.93e−01
6	6.26e−02	⋯	7.61e−02	⋯	2.39e−01	⋯	3.47e−01	⋯
8	5.49e−02	5.40e−02	7.94e−02	7.84e−02	2.23e−01	2.26e−01	2.92e−01	2.93e−01
9	4.85e−02	⋯	7.14e−02	⋯	2.17e−01	⋯	2.72e−01	⋯
12	4.26e−02	⋯	5.93e−02	⋯	1.83e−01	⋯	2.72e−01	⋯
16	3.74e−02	3.99e−02	5.02e−02	4.88e−02	1.83e−01	1.87e−01	2.64e−01	2.37e−01
18	3.66e−02	⋯	4.84e−02	⋯	1.64e−01	⋯	2.30e−01	⋯
24	3.73e−02	⋯	4.36e−02	⋯	1.17e−01	⋯	1.80e−01	⋯
27	3.75e−02	⋯	4.36e−02	⋯	1.24e−01	⋯	1.79e−01	⋯
32	3.98e−02	4.40e−02	4.42e−02	4.66e−02	1.15e−01	1.19e−01	1.62e−01	1.66e−01

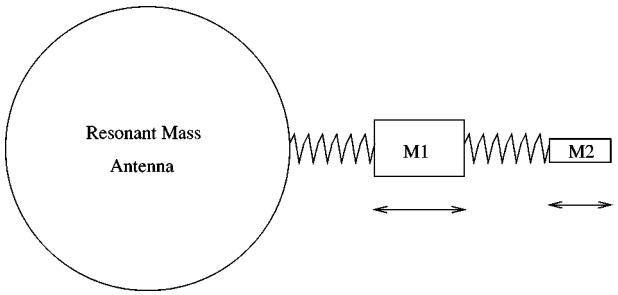


Figure 11. The spherical antenna with one transducer system mounted.

building a gravitational wave detector. Nevertheless, we will illustrate the capabilities of the FE simulation program using a few ''sample'' experiments.

All simulations are carried out with full-fledged antenna models of high resolution (approximately 130,000 elements) on the two high-performance distributed-memory platforms discussed in Sec. III B. The antenna material resembles CuAl (90–10). Six resonant mass transducers are mounted on the antenna in somewhat of an arrangement that places the transducers at the pentagons on the lower half of a truncated icosahedron; this was also done in the TIGA project.[27] Each transducer consists of two masses. The first mass of $M_1 = 100$ kg is attached to the antenna by a spring that is described by the diagonal stiffness matrix $K_1 = \mathrm{diag}(2.27\times10^9, 2.27\times10^{10}, 2.27\times10^{10})$ N/m. The second mass of $M_2 = 0.385$ kg is attached to the first mass by a spring with stiffness matrix $K_2 = \mathrm{diag}(8.70\times10^6, 8.70\times10^7, 8.70\times10^7)$ N/m.

Figure 11 schematically depicts a transducer that is attached to the antenna. The principal resonance frequency f_r of the transducer system is equal to $f_r = 780$ Hz, which coincides with the principal quadrupole frequency of the ideal detector (see Sec. III A). The two masses are chosen such that the amplitude of vibration detected at the antenna's surface will be amplified by two orders of magnitude. The addition of the multimode resonant transducers results in a significant splitting of the resonant frequencies of the system (data not shown). Although the transducer models that we have implemented are rather simplistic (point masses coupled with ideal springs), they have already given quite satisfactory results in several applications (see, for example, Sec. IV B).

A. Gravity

Earth's gravity will induce internal strain in the antenna. The sphere that is suspended is expected to deform due to this strain. We would like to quantify the amount of deformation that results due to the presence of gravity. In order to do this we perform the following simulation. At time zero we ''switch on'' gravity ($g = 9.8$ m/s^2), and let the antenna fall down. The suspension rod is fixed at the top to prevent the sphere from dramatically accelerating towards the Earth's surface. The velocities in the sphere are quenched by a small damping factor in order to get rid of the kinetic energy in the system. When the system has come to rest, we take a look at the displacement and strain distribution within the sphere. Figure 12 displays a contour

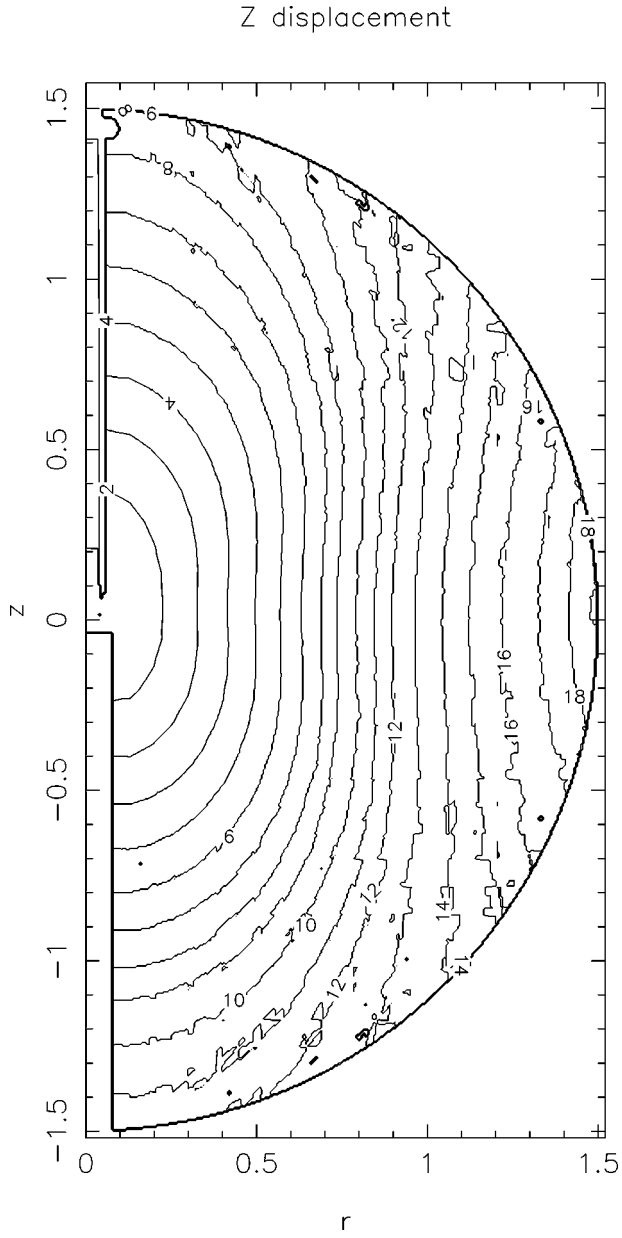


Figure 12. Contour plot of the displacement in the $\hat{z}$ direction of the suspended CuAl *sphere due to gravity ($g = 9.8$ m/s^2). The displacement values of the various contours correspond to (1) −0.35, (2) −1.96, (3) −2.18; further contours do not differ significantly from the value of contour (3); all values are in 10^{-3}* m. *The minus signs indicate displacements towards the Earth's surface.*

plot of the displacement field in the $\hat{z}$ direction within the antenna, after this deformation process. Figure 13 zooms in on the suspension region, again showing a contour plot of the displacement in the $\hat{z}$ direction. Overall, the sphere moves about 2 mm downwards, which results from the fact that the suspension rod is elongated by this amount in the $\hat{z}$ direction. The displacement values within the sphere vary from about 2×10^{-3} to 3×10^{-4} m when we traverse the sphere from the outside towards the suspension rod. Another interesting view is provided by the vertical strain

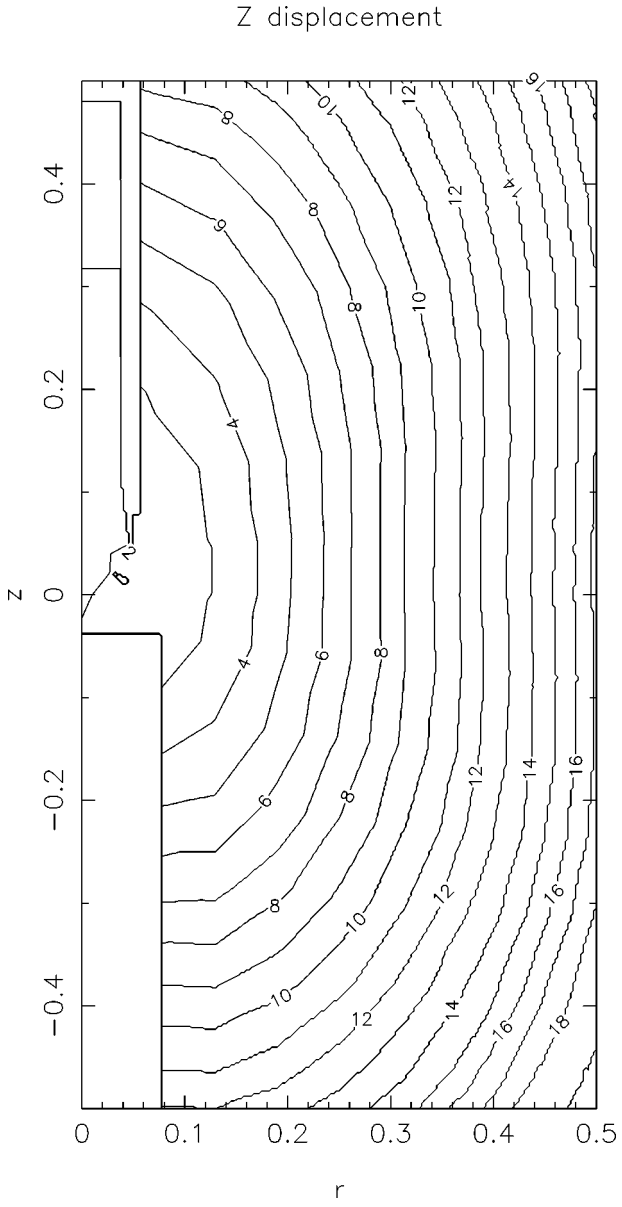


Figure 13. The same contour plot as that displayed in Fig. 12, but zoomed in on the suspension region. The displacement values of the various contours correspond to (1) −1.60, (2) −1.82, (3) −2.16, (4) −2.18; further contours do not differ significantly from the value of contour (4); all values are in 10^{-3} m. The minus signs indicate displacements towards the Earth's surface.

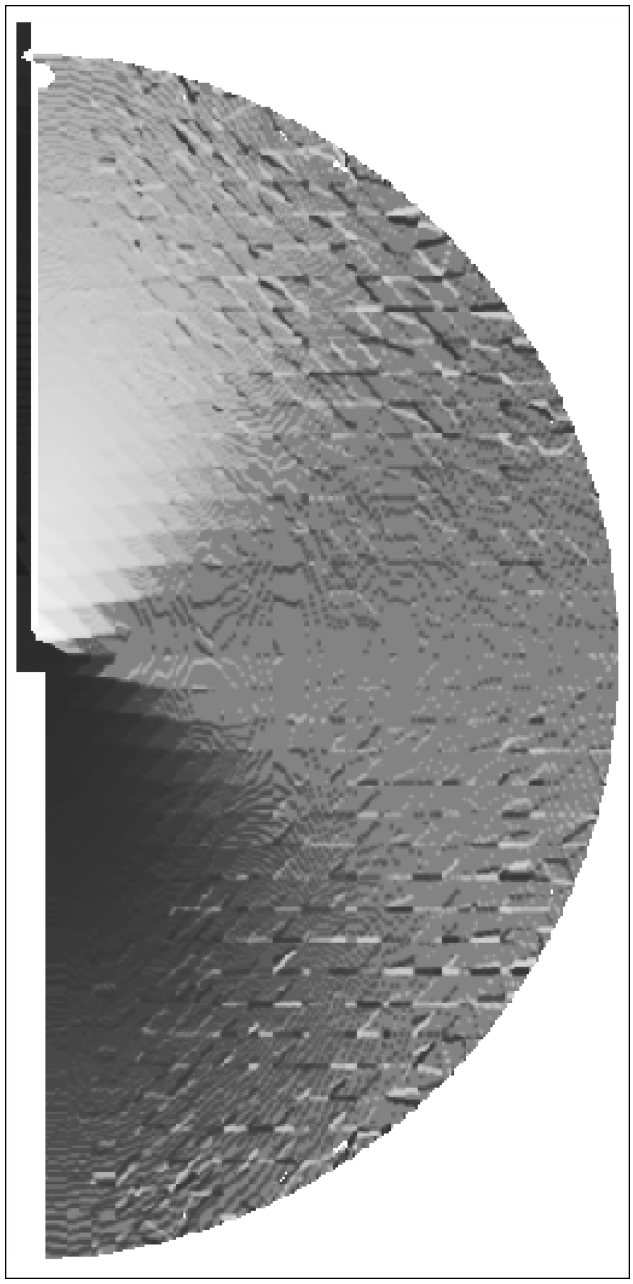


Figure 14. The vertical strain $(\partial u_z/\partial z)$ within the antenna. High strain values (dark areas) are found in the suspension rod, that is carrying the entire sphere, and near the contact region of the sphere and the rod. The strain in the sphere itself is much lower (light colored areas).

$(\partial u_z/\partial z)$ within the antenna, which is shown in Fig. 14. As expected, the suspension rod is subjected to the highest strain, as is the region where the sphere and rod make contact.

A shortcoming of our simulation program is that it relies on a linear finite element solver. As a consequence, we are not able to model nonlinear effects, such as so called *violin modes* in the suspension rod, or show how the internal strain of the sphere affects the eigenmodes. Incorporating these into our simulation program is a future challenge.

B. Seismic noise

The suspension system of the antenna must attenuate vibrations at the principal resonance frequency to a factor of 320 dB.[28] If we let seismic noise enter the system via the suspension, how does it appear at the transducer readout? The following experiment is carried out. The suspension points (connections to the outside world) of the suspension rod are driven by forced oscillation in the $\hat{z}$ direction. That is, the z coordinate of the displacement field $u_z(t)$ of each suspension point oscillates as follows:

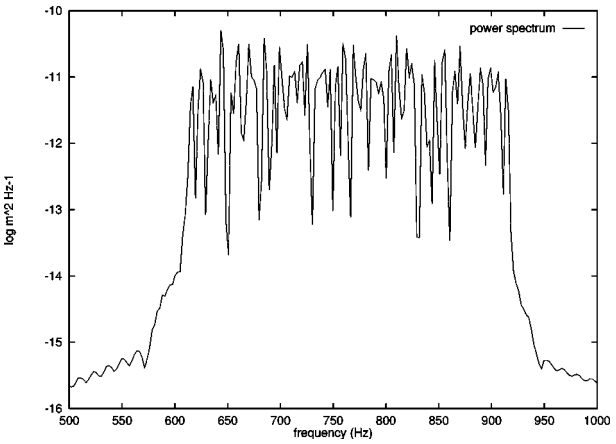


Figure 15. Narrow-banded seismic noise around the principal quadrupole frequency.

$$u_z(t)=\sum_{n=1}^{N}\left[A_n\cos(2\pi f_n t)+B_n\sin(2\pi f_n t)\right]. \quad (18)$$

The amplitudes A_n and B_n and the frequency spectrum f_n model the seismic noise. In the following experiment the narrow-banded noise spectrum depicted in Fig. 15 was used. The $\hat{z}$ displacement versus time, which is driving the suspension points according to this spectrum, is displayed in Fig. 16.

Figures 17 (inner mass) and 18 (outer mass) show the radial displacement of the two transducer masses starting at $t=0$ in one of the six transducers. The maximum amplitude in the noise spectrum (Fig. 15) has the same order of magnitude as the maximum amplitude of the outer mass. Consequently, it is of paramount importance that all seismic vibrations in the target frequency range are attenuated below the expected amplitude of vibrations that are induced by gravitational radiation in order to assure that we have a decent signal-to-noise ratio.

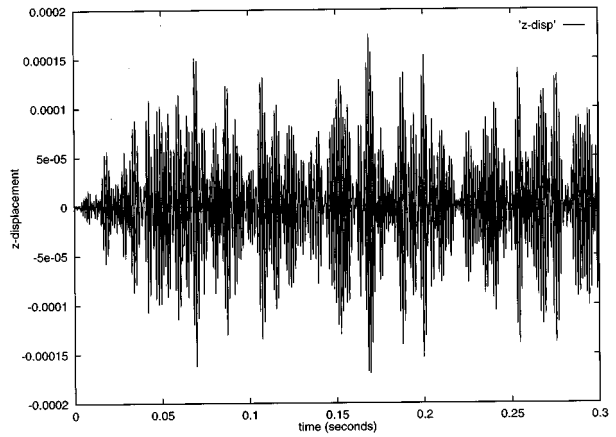


Figure 16. The $\hat{z}$ displacement for the suspension point driven by the noise spectrum of Fig. 15.

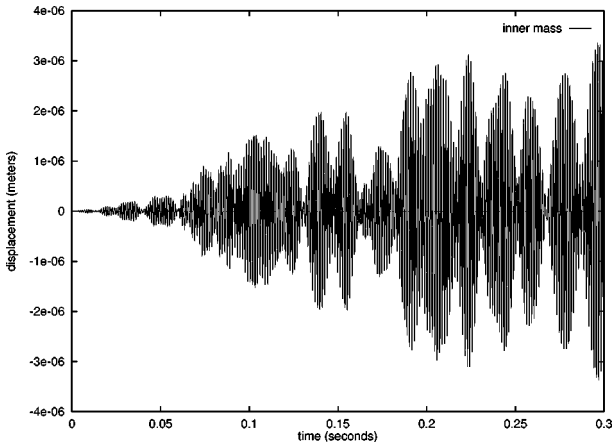


Figure 17. Time trajectory of the inner transducer mass on one transducer. The system is forced into oscillation by the seismic noise spectrum of Fig. 15.

C. Chirps

A so-called gravitational radiation ''chirp'' will be emitted (expected theoretically) just before coalescence of a binary neutron-star system.[29] The effects of this phenomenon on the detector can be modeled by applying a time-varying ''chirp force'' $\mathbf{f}(t)$ on the antenna, which acts along a pure quadrupole field, according to

$$\mathbf{f}^T(t)\propto\ddot{h}(x,-y,0), \quad (19)$$

with $\ddot{h}$ the second-order time derivative of the deformation of space, derived in a Newtonian approximation.

Let us consider a chirp that enters the antenna at time $t=-0.04$ s. Figure 19 models its time varying amplitude $\ddot{h}$. The normalization of the chirp signal has an order of magnitude that is representative for an event in the Virgo cluster. Since the antenna responds purely linearly, the shape of the response signal resulting from interaction with any other identically shaped incoming signal will be exactly the same.

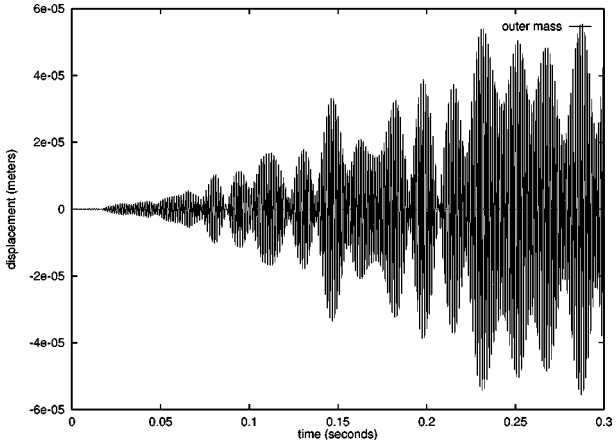


Figure 18. Time trajectory of the outer transducer mass on one transducer. The system is forced into oscillation by the seismic noise spectrum of Fig. 15.

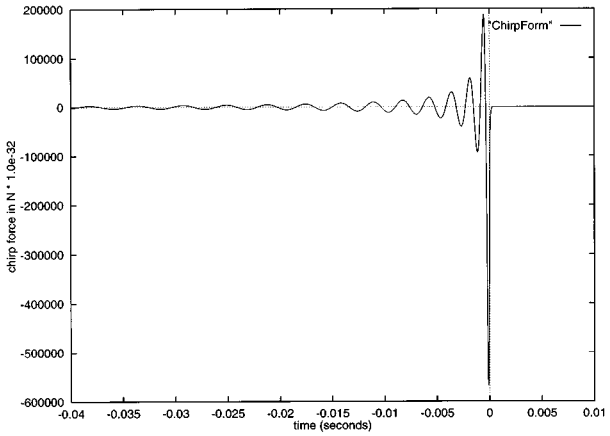


Figure 19. The time dependent amplitude $\ddot{h}$ of a chirp initiating at time $t=-0.04$ s. The chirp ends at $t=0$.

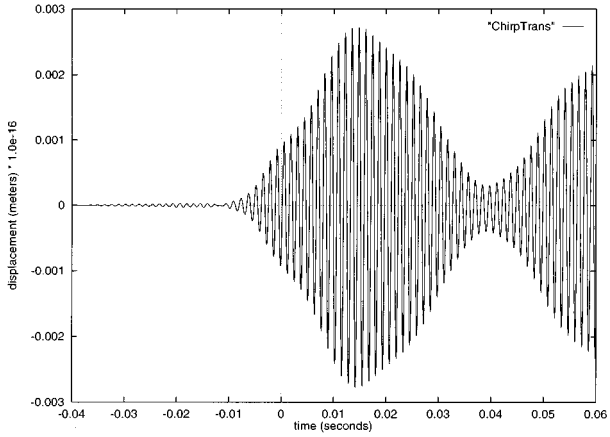


Figure 21. The $\hat{x}$ displacement of one of the outer transducer masses. The transducer already ''notices'' the chirp's presence before $t=0$.

Figure 20 shows the evolution of the energy in the antenna from the moment that the chirp enters until it has passed. The energy of the chirp gets deposited in the antenna. After the chirp has passed, the energy stabilizes (at $t=0$ s).

The response to this chirp can again be measured. Figure 21 shows the time trajectory in the $\ddot{x}$ direction of one of the outer transducer masses. We observe the tiny ($\approx 10^{-19}$ m) displacements of the transducer masses.

V. CONCLUSIONS

In this article we have sketched the complete path along which a complex high-performance simulation was developed. We were able to identify the problem definition, modeling choices, the code design for parallel systems, validation, and experimentation. In order to realize such a system, it is mandatory that we have knowledge of a variety of disciplines, like computer science, numerical mathematics, and (computational) physics. It is our feeling that the whole process that is described above is typical for the development of new simulation systems that utilize high-performance methodology and therefore can serve as an example of *computational science*, which inherently is an interdisciplinary research field. Mixing the best of several worlds allowed us to realize a full-fledged simulation system that meets our purposes.

In the future we will focus on embedding the simulation program within the experimentation cycle of GRAIL, referred to as ''simulation in the loop'' in the Introduction.

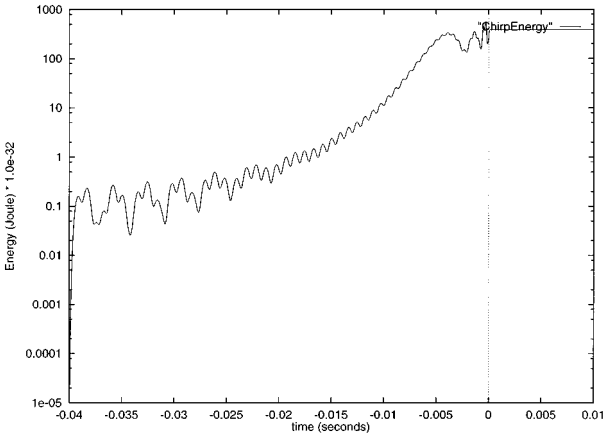


Figure 20. The time evolution of the total energy deposited by the chirp of Fig. 19. After $t=0$ the energy stabilizes. Note that we have used a log scale on the energy axis.

ACKNOWLEDGMENTS

The work presented in this article was made possible by funding from NWO of the GRAIL project pilot study. The contributions of J. de Rue to Sec. I B are gratefully acknowledged.

REFERENCES

1. C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973).
2. http://qv3pluto.leidenuniv.nl/grail/grail.html. GRAIL stands for ''gravitational radiation antenna in Leiden.''
3. The GRAIL Project Team. GRAIL: A proposal for a gravitational radiation antenna in The Netherlands. Technical Report No. NIKHEF/95-005, Nationaal Instituut voor Kernfysica en Hoge-Energie Fysica, 1995.
4. K. S. Thorn, *300 Years of Gravitation*, edited by S. W. Hawking and W. Israel (Cambridge University Press, Cambridge, 1987).
5. N. Ashby and J. Dreitlein, Phys. Rev. D **12**, 336 (July 1975).
6. J. A. Lobo, Phys. Rev. D **52**, 591 (July 1995).
7. S. M. Merkowitz and W. W. Johnson, Phys. Rev. D **51**, 2546 (March 1995).
8. S. M. Merkowitz, Ph.D. thesis, Lousiana State University, 1995.
9. J. Weber, Phys. Rev. **117**, 307 (1960).
10. http://www.roma1.infn.it/rog/rogmain.html. Rome gravitational wave experiment, NAUTILUS, 1996.
11. Dutch organization of scientific research.

12. P. M. A. Sloot and J. Reeve, Camas-tr-2.3.7 executive report on the Camas workbench; Technical report, University of Amsterdam and University of Southampton, October 1995.
13. B. D. Kandhai, P. M. A. Sloot, and J. P. Huot, in *High-performance Computing and Networking*, edited by B. Hertzberger, H. Liddel, A. Colbrook, and P. M. A. Sloot, No. 1067 in ISBN 3-540-61142-8 (Springer, Berlin, 1996), pp. 269–275.
14. P. M. A. Sloot, Eurosim '95, Simulation Congress, Amsterdam, The Netherlands, 1995, pp. 29–44.
15. A. E. H. Love, *A Treatise on the Mathematical Theory of Elasticity* (Dover, New York, 1944).
16. J. de Rue, MS. thesis, University of Amsterdam, 1996.
17. T. M. Bell, W. M. Visscher, A. Migliori, and R. A. Reinert, J. Acoust. Soc. Am. **90**, 2154 (1991).
18. E. Andersen, Z. Bai, and C. Bischof, LAPACK User's Guide, 2nd ed. 1994.
19. O. C. Zienkiewicz and R. L. Taylor, *The Finite Element Method* (McGraw-Hill, New York, 1994).
20. T. J. R. Hughes, *The Finite Element Method, Linear and Dynamic Finite Element Analysis* (Prentice-Hall, Englewood Cliffs, NJ, 1987).
21. R. H. MacNeal, *Finite Elements: Their Design and Performance* (Dekker, New York, 1994).
22. A. Geist, A. Beguelin, J. Dongarra, W. Jiang, R. Manchek, and V. Sunderam, PVM 3 User's Guide and Reference Manual, Oak Ridge National Laboratory, May 1993.
23. G. Fox, M. Johnson, G. Lyzenga, S. Otto, J. Salmon, and D. Walker, *Solving Problems on Concurrent Processors* (Prentice-Hall, Englewood Cliffs, NJ, 1988), Vol. 1.
24. H. D. Simon, Partitioning of unstructured problems for parallel processing. Comput. Syst. Engr. **2**, 135 (1991).
25. G. Lonsdale, J. Clinckemaillie, S. Vlachoutsis, J. F. de Ronde, P. M. A. Sloot, N. Floros, and J. Reeve, Conference on Supercomputing Applications in the Automotive Industries of the 26th ISATA, September 1993.
26. Interdisciplinary Center for Complex Computer Facilities. http://www.wins.uva.nl/research/ic3a/. Amsterdam, 1996.
27. W. W. Johnson and S. M. Merkowitz, Phys. Rev. Lett. **70**, 2367 (April 1993).
28. T. L. Aldcroft, Rev. Sci. Instrum. **63**, 3815 (1992).
29. L. Blanchet, B. R. Iyer, C. M. Clifford, and A. G. Wiseman, Class. Quantum Grav. **13**, 575 (1996).

Detection of Gravitational Waves with High Frequency Square Waves from Crystal Oscillator

Dhananjay P. Mehendale
Sir Parashurambhau College, Tilak Road, Pune 411030, India

Abstract

We propose novel technique for detecting gravitational waves using setup like LIGO and a robust square wave generator as source of waves generating stable square waves with high frequency using an appropriate crystal oscillator. This change in source will prove very effective for the process of detecting gravitational waves because even if with extremely small change in phase of two interfering waves reaching the beam splitter there will be detectable change in the amplitude due to the suddenly rising sharp edge of these almost but not exactly destructively interfering waves upon the passing of gravitational waves. This idea of achieving detectable amplitude of a portion of signal reaching the detector will make it possible to record extremely small phase difference generated due to passing of some gravitational waves through apparatus in the two interfering square wave beams. The change suggested in the gravitational wave detection experiment using high frequency well regulated square wave source will produce the following effects. When no gravitational waves will reach the apparatus the square waves from crystal oscillator will move towards beam splitter and will split into two square waves heading towards reflectors kept orthogonally to each other at equal distance from beam splitter at the ends of two interferometer arms and these waves will reflect back to interfere destructively so that no signal will reach the detector. But when some gravitational waves will be reaching the apparatus and further suppose there is squeezing along arm A and stretching along arm B of the interferometer as an effect these gravitational waves will be producing the (very small) phase difference in the interfering square waves reaching back to beam splitter. Now, however small may be this phase difference its undetectably small amount still will remain detectable because of sharp rise and fall of amplitude at edges of square wave because of its typical shape and so some detectable part of these interfering waves will pass towards detector to offer a conclusive proof for the existence of gravitational waves!

1. **Introduction:** Experimental detection of gravitational waves is a big challenge of this time and enormous efforts are on world over by people working in highly sophisticated gravitational wave detection laboratories. Gravitational wave laboratories will be leading laboratories in the coming future to offer new important insights in our study of large scale phenomena. Detection and study of gravitational waves of different types and of different intensity and frequency will make revolutionary contributions to our knowledge about galactic dynamics. It will add greatly to our knowledge about astrophysical sources and about processes driven by strong gravitational fields. Objects of fundamental importance, such as astrophysical black holes, merge and radiate with luminosity larger than the entire electromagnetic universe, and these events will become clearly detectable only through a tool for detection of gravitational waves that are mainly associated with detectable amplitude with such unimaginably huge events [1]. When observed with gravitational waves these intrinsically interesting astronomical sources such as massive black holes and their merger, extremely compact stellar binaries and their collisions, supernovae events etc will surely yield many new surprises. Thus, the discovery potential associated with detection of gravitational waves is immense.

   Gravitational radiation was detected indirectly in 1974 by J. Taylor and R. Hulse, who observed its effects on the orbital period of a binary system containing two neutron stars, one of them a pulsar (PSR 1913 + 16). Efforts to detect gravitational waves directly have been severely challenged by the extreme weakness of the waves impinging on the Earth. However, as the 21st century begins, observations of the gravitational waves from astrophysical sources such as black holes, neutron stars, and stellar collapse are expected to open a new window on the universe [2].

   There are two major gravitational wave detection concepts: acoustic and interferometric detection [3]. The acoustic method deals with a resonance response of massive elastic bodies on gravitational wave excitations. Historically the acoustic method was proposed first by J. Weber [4] where he suggested to use long and narrow elastic cylinders as Gravitational Wave Antennas. Although a significant progress has been achieved in fabrication and increasing sensitivity of such type of detectors [4, 5, 6] the interpretation of obtained data is still far to claim undoubtedly the detection of gravitational waves. On the other hand a considerable attention has been shifted recently to more promising interferometric detection methods. The interferometric gravitational wave detector like Laser Interferometer Gravitational Wave Observatory (LIGO) and VIRGO [7, 8] represents a Michelson interferometer with a laser beam split between two perpendicular arms of interferometer. The principles of operation of such type of detectors are reviewed in Refs [9, 10, 11, 12, 13]. The action of gravitational waves on an interferometer can be presented as relative deformation of both interferometer arms. A gravitational wave with dimensionless amplitude h induces the opposite length changes $\frac{\delta l}{l} = \frac{1}{2} h \cos(\Omega t)$ in each arm of the Michelson interferometer, where l stands for the length of each of the arm, Ω for the gravitational wave frequency. These

length changes produce opposite phase shifts between two light beams in interferometer arms, when interference occurs at the beam splitter of Michelson interferometer. The resulting phase shift of a single beam of light spending time τ in the interferometer can be written as [13]

$$\delta\phi = h\frac{\omega}{\Omega}\sin\left(\frac{\Omega\tau}{2}\right), \qquad (1)$$

where, ω is the light frequency. This phase shift results an intensity signal change of the light from interferometer beam splitter hitting the photo detector.

The main problem of the acoustic and interferometric methods that they both deal with gravitational waves with extremely small amplitudes of the order $h \sim 10^{-21}$ [14] reached the Earth from deep space. One can see from equation (1) that for gravitational wave frequencies in the 1 kHz range, $\Omega \sim 10^3$ Hz, and for the light in visual frequency range, $\omega \sim 10^{14}$ Hz, one has the maximum phase shift of the order $\delta\phi \sim 10^{-10}$ for interferometer arms length of the order 150 kilometers. Such extraordinarily weak effect requires exceedingly high detector sensitivity for both acoustic as well as interferometric detectors.

2. **The Novel Technique:** As seen above one requires extraordinarily high sensitivity of detectors to conclusively capture signal called gravitational wave and this may be one of the important reasons that we have not yet succeeded in this task even though many gravitational waves would have passed through our apparatus installed at different locations.

In this paper we suggest a new idea of using as source a square wave generator generating stable high frequency square waves from a crystal oscillator. This change in the source will cause passing of detectable value of the part of interfering square waves due to the typical steep rise and fall in the amplitude of waves at the locations of completion of distance equal to wavelength. Due to this typical shape of waves the undetectably small phase difference created in the interfering waves reaching the beam splitter due to passing of gravitational waves in the beams reflected back from two perfect reflectors, fixed at ends of two orthogonal interferometer arms, due to which these beams come back to interfere almost destructively at beam splitter will produce detectable portion of amplitude reaching the detector. We take high energy well regulated square wave generator as new source and perform usual experiment. Now, suppose when the beams of square waves were running through two mutually orthogonal arms and were coming back to beam splitter there was squeezing in the arm, A say, of the interferometer and reciprocally suppose there was stretching in the interferometer arm, B say, of the interferometer. As an effect of this specially chosen source generating square waves the phase difference, however small it may be, will produce detectable overlap in the interfering square waves because of their typical shape having a sharp rise, equal to full amplitude of the waves. It will cause reaching of sufficient amplitude towards detector which essentially will lead to reaching of detectable amplitude at the detector. This will give us a confirmation that though the phase difference is

extremely small still our new choice of the shape of the waves makes it detectable to our instrument and thus will give a confirmation that beams are in fact not interfering destructively as they are sending in effect sizable amplitude towards detector.

The main idea in brief behind the suggestion in this paper is to capture the undetectably small phase difference that should result due to passing of gravitational waves by making it sizable by choosing the square shape for the waves from the source. Thus, we are suggesting a way out to achieve sizable value of the amplitude of waves reaching the detector by choosing typical square shape for the signal from the source to make the undetectably small phase difference detectable.

References

1. Tom Prince (Lead Author for Members of the LISA International Science Team), The Promise of Low-Frequency Gravitational Wave Astronomy, 2010.
2. Joan M. Centrella, Laboratory for High Energy Astrophysics, Resource Letter GrW-1: Gravitational Waves, 2003.
3. G. B. Lesovika, A. V. Lebedeva, V. Mounutcharyana, T. Martinb, Detection of gravity waves by phase modulation of the light from a distant star, arXiv: astro-ph/0506602v1, 2005.
4. E. Amaldi et al., Nuovo Cimento, 7C, 338 (1984).
5. E. Amaldi et al., Nuovo Cimento 9C, 829 (1986).
6. P. Aston et al., Phys. Rev. D 47, 362 (1993).
7. R. E. Vogt, in Sixth Marcel Grossmann Meeting on General Relativity, Kyoto, Japan, 1991, edited by H. Sato and T. Nakamura (World Scientific, Singapore, 1992).
8. C. Bradaschia et al., Nucl. Instrum. Methods A 289, 518 (1990).
9. R. L. Forward, Phys. Rev. D 17, 379 (1978).
10. K. S. Thorne, in 300 Years of Gravitation, edited by S.W. Hawking and W. Israel (Cambridge University Press, Cambridge, 1987).
11. J. Hough, B. J. Meers, G. P. Newton, N. A. Robertson, H. Ward, B. F. Schultz, I. F. Corbett and R. W. P. Drever, Vistas Astron. 30, 109 (1987).
12. R. W. P. Drever, in Gravitational Radiation, edited by N. Deruelle and T. Piran (North-Holland, Amsterdam, 1983).
13. B. J. Meers, Phys. Rev. D 38, 2317 (1988).
14. B. F. Schutz, in Fourteenth Texas Symposium on Relativistic Astrophysics Proceedings, Dallas, Texas, 1988, edited by E. J. Fenyves [Ann. N. Y. Acad. Sci. 571, 27 (1980)].

Email: dhananjay.p.mehendale@gmail.com

Generation and Detection of Low-Frequency Squeezing for Gravitational-Wave Detection

Michael Steve Stefszky

A thesis submitted for the degree of
doctor of philosophy of
The Australian National University

April 2012

Declaration

This thesis is an account of research undertaken between March 2008 and April 2012 at The Department of Physics, Faculty of Science, The Australian National University, Canberra, Australia.

Except where acknowledged in the customary manner, the material presented in this thesis is, to the best of my knowledge, original and has not been submitted in whole or part for a degree in any university.

Michael S. Stefszky
April, 2012

Acknowledgements

The last four years of my life have been amazing and I would like to thank (hopefully) everyone who has contributed and supported me over this time.

These years have been made fruitful and enjoyable by the support of an outstanding supervisory panel. For this, I would like to first give my sincerest thanks to Professor Ping Koy Lam, Doctor Ben Buchler and Doctor Thomas Symul from the Quantum Optics Group (now CQC2T), and Professor David McClelland and Doctor Daniel Shaddock from the Centre for Gravitational Physics.

The foremost thanks has to go to Ping Koy Lam, whom I have studied with for many years now. Much appreciation is given for my acceptance into the quantum optics group many years ago, first for my honours year and then for my PhD. The foresight and sense of direction that Ping Koy has showed over the years, both in individual experiments and for the entire group, has been, and continues to be, invaluable. I am very grateful for the opportunities I have had to travel both locally and overseas for conferences and collaborations. Ping Koy's outstanding abilities as a group leader have been made apparent to me many times and his enthusiasm for his work is very addictive.

My decision to work on the H1 LIGO squeezer project later into my PhD I can now look back and say was most certainly the correct one. For this opportunity I have firstly to thank David. Working on this project meant many things. It meant the support of a second research group, the Centre for Gravitational Physics, the opportunity to collaborate with many institutions and people on a truly international project, and the learning advantages from working in a new group with its own areas of expertise. Working on this project allowed me to work at MIT for nearly 6 months and at LIGO for nearly 3. The experience gained from these travels makes up a large portion of what I consider to be the most valuable training from my studies and for this opportunity I need to thank David. Additionally, David always found time in a busy schedule to help with experimental problems and provide direction when needed.

Sincerest thanks then goes to Ben and Daniel. Together with Ping Koy and David, I cannot imagine a better group of people to have available when a problem arose. Although by no means the limit of their knowledge, Daniel's expertise in electronics and locking loops, combined with Ben's expertise in just about anything related to optics, particularly non-linear optics, meant that solutions to any problem were quickly found. The willingness to help combined with the knowledge and good humour of both Daniel and Ben was both invaluable and greatly appreciated.

Onto my two lab mates, Sheon Chua and Conor Mow-Lowry. If it were not so much fun to work in the lab as it was, then I am sure things would not have progressed as far as they did. The environment in the lab was a relaxed one, where input from each person was valued, and I really appreciate the synergy we had going. Conor's knowledge ensured that progress was smooth and constantly provided alternative solutions to problems that arose. Sheon's problem solving techniques were different to mine and so together we attacked problems from many different angles.

From the ANU still, I need to thank the electronics and mechanical workshops, Shane

Grieves, Neil Hinchey, James Dickson, Paul McNamara, Paul Tant and Neil Devlin. The work these people do is of the highest quality and without their help, I might still be milling steel and soldering circuits. Also to everyone from both the quantum optics group and all of the graviteers, as well as everyone else from the physics department. How enjoyable a particular job is depends on both the work one is doing and the people involved. The people in the physics department have have certainly made the last few years enjoyable. A quick mention to Nicolai Grosse for his aid, wisdom and friendship during my honours year. A big thanks also goes out to Phil Threlfall, Richard Barry, and especially Conor Mow-Lowry and Timothy Lam, who along with Ping Koy, David, Daniel and Ben, drafted various sections of this thesis.

To all the people in the United States, thank you. To Nergis Mavalvala, Daniel Sigg, Lisa Barsotti, Sheila Dwyer and the rest of the LIGO team, thank you very much for the opportunity to visit a new country and work in new places on exciting experiments. Thank you to Nergis and Daniel for allowing me to work at MIT and at LIGO, and thank you to Lisa for her excellent organisational skills in organising both human resources and experiments. Thanks to Sheila for making my stay in Boston as enjoyable as possible and to Dick Gustafson for introducing me to Cognac and ice-cream in Hanford.

Onto the friends and help that I have received from people at the Max Planck Institute for Gravitational Physics. Firstly, thank you to Roman Schnabel, Henning Vahlbruch and Alexander Khalaidovski and Moritz Mehmet. Thank you for the many discussions that were had on detector design, balanced homodyne detection and squeezed light production. Your willingness to share information and aid us in our experiment was invaluable. Thank you to all the people at Hannover who made my short stay there incredibly fun. So much fun that I hope to return to Hannover some time in the near future. Thank you to Moritz for showing me around your home country and institute. I hope I can return the favour soon.

Then there are my friends outside of work. My housemates and friends for many years, Neil, Mitch, Sam, and Jason as well as my friends whom I see the most often; Conor, Andrew, Kieran, Neil and Phil to name but a few. Thank you for providing distractions when I needed it, and sometimes when I did not, and for making my home environment a fun place to be. Thank you for trying to keep me fit with the many mountain bike rides we went on. The ability to unwind and relax after writing for hours on end is not an easy task sometimes but you guys made it easier.

Onto my family. Thank you to Mum, Dad, Michael and Angelina for your never ending support. Thank you to Warren, Tanya, Michael, and Anthony for keeping things real. Finally, the most heartfelt thank you to Jenessa, whose understanding and support has been nothing short of amazing.

Abstract

Gravitational-wave astronomy promises a radically new method of investigating the universe, one that does not detect elementary particles, but rather ripples in space-time itself. One might expect that with this radically new method, our understanding of the universe will make an equally giant leap. One of the most promising methods for the eventual detection of gravitational-waves is that of interferometric gravitational-wave detection.

The sensitivity of these devices is nothing short of amazing, and their sensitivity continues to improve. The next generation of detectors are expected reach a point where increasing the laser power, previously used to increase the sensitivity, is no longer expected to provide any benefit. One of the most promising options for further increasing the sensitivity, through reduction of the quantum noise, is via the application of "squeezed states of light". Squeezed states of light are light fields for which the noise of some observable has been reduced below the quantum noise limit. These states have been shown for many years to offer increased sensitivity in interferometers and recently in interferometric gravitational-wave detectors. One of the most challenging tasks to make suitable squeezed light sources for these detectors is to produce low-frequency squeezing, corresponding to the detection band of these interferometers.

This thesis details the doubly resonant travelling wave bow-tie squeezer as a source of squeezing for interferometric gravitational-wave detection. This squeezer achieves record results of squeezing in the gravitational-wave detection band. For the first time, 10 dB of shot noise suppression at 10 Hz is directly observed and above 200 Hz, 11.6 dB is observed. The work presented in this thesis provides evidence that further substantiates previous progress indicating that squeezed states of light are now ready for full-time integration into interferometric gravitational-wave detectors. The new cavity design, described in detail throughout this thesis, is shown to outperform previous designs and provide benefits when integrating squeezing into these detectors, in particular, isolation to backscattered light from the interferometer.

The noise sources that typically lead to degradation in squeezing measurements are investigated and discussed. Additionally, a modification to the standard locking technique used to control these vacuum squeezed states, *coherent locking*, is presented and discussed. The modified technique reduces the required number of locking loops and provides a larger beat note from which to derive an error signal for one of the remaining loops.

A squeezer using these design philosophies is then constructed and used to inject squeezing into the LIGO gravitational-wave detector. Injection of the squeezed state provides enhancement of the sensitivity of the detector at frequencies around 200 Hz and above. The sensitivity seen was equivalent to approximately 2 dB of shot noise reduction. The enhancement is limited by the loss within the interferometer. Injection of the squeezed state did not show any degradation in the sensitivity of the device below 200 Hz, indicating for the first time that noise couplings between the interferometer and the squeezer could be sufficiently suppressed at these frequencies even for the most sensitive detectors.

Contents

Declaration iii

Acknowledgements v

Abstract vii

Summary of symbols and acronyms xix

1 Introduction 1
- 1.1 Thesis Structure . . . 2
- 1.2 Publications . . . 3

2 Background 5
- 2.1 Gravitational-Waves . . . 5
- 2.2 Detecting Gravitational-Waves . . . 6
  - 2.2.1 Resonant Bar Detectors . . . 6
  - 2.2.2 Interferometric Detectors . . . 7
- 2.3 aLIGO Noise Sources . . . 8
  - 2.3.1 Low-frequency Noise Sources . . . 8
  - 2.3.2 High Frequency Sensitivity . . . 11
- 2.4 Quantum Enhancement . . . 13
- 2.5 Optical Squeezing . . . 13
  - 2.5.1 Low-frequency Squeezing . . . 13
  - 2.5.2 Squeezing Enhancement of Interferometers . . . 14
  - 2.5.3 Squeezing in LIGO H1 . . . 15

3 Quantum States of Light 17
- 3.1 The Quantum Mechanical Electromagnetic Field . . . 17
  - 3.1.1 Quadratures . . . 17
  - 3.1.2 Quadrature Operators . . . 17
- 3.2 The Heisenberg Uncertainty Principle . . . 18
  - 3.2.1 State Variance . . . 18
  - 3.2.2 Minimum Uncertainty States . . . 19
  - 3.2.3 Squeezing Purity . . . 19
  - 3.2.4 The Number Operator . . . 19
- 3.3 Linearisation . . . 20
- 3.4 Representing the States of Light . . . 20
  - 3.4.1 Ball-on-Stick Representation . . . 20
  - 3.4.2 Sideband Representation . . . 20
- 3.5 Introducing the States of Light . . . 22
  - 3.5.1 Vacuum State . . . 22
  - 3.5.2 Coherent State . . . 22

3.5.3 Squeezed State . . . 23
3.6 Squeezed State Loss . . . 25
3.7 State Measurement . . . 27
3.7.1 Oscilloscope Measurements . . . 28
3.7.2 Spectrum Analysers . . . 28
3.8 Balanced Homodyne Detection . . . 28
3.8.1 Balanced Homodyne Detector Fringe Visibility . . . 31

4 Generation of Squeezed States 33
4.1 Methods of Squeezed State Production . . . 33
4.1.1 History of Optical Parametric Oscillator and Amplifier Performance 33
4.2 The Second Order Non-linearity . . . 34
4.3 Second Harmonic Generation . . . 35
4.4 Downconversion Processes . . . 35
4.4.1 The Degenerate Optical Parametric Amplifier . . . 35
4.4.2 The Non-Degenerate Optical Parametric Amplifier . . . 36
4.4.3 The Optical Parametric Oscillator . . . 37
4.5 Phase Matching . . . 39
4.5.1 Birefringent Phase Matching . . . 40
4.5.2 Quasi-Phase Matching and Periodic Poling . . . 40

5 Optical Cavities 43
5.1 Cavity Equations of Motion . . . 43
5.2 Classical Cavity Phase and Amplitude Response . . . 45
5.2.1 Circulating Power and Finesse . . . 46
5.2.2 Free Spectral Range . . . 47
5.2.3 Linewidth . . . 47
5.3 Semi-Classical Cavity Noise Response . . . 47
5.4 The Optical Parametric Oscillator . . . 49
5.4.1 Escape Efficiency . . . 53
5.5 The Classical Degenerate Optical Parametric Amplifier . . . 54

6 The Doubly Resonant Bow-Tie Optical Parametric Oscillator 57
6.1 Cavity Geometry . . . 57
6.1.1 Travelling Wave vs. Standing Wave . . . 58
6.1.2 Choosing a Bow-Tie Geometry . . . 60
6.2 Cavity Optical Design Considerations . . . 62
6.2.1 Doubly Resonant System . . . 62
6.2.2 Escape Efficiency and the Doubly Resonant System . . . 66
6.2.3 Linewidths and Finesse . . . 68
6.2.4 Threshold Power . . . 68
6.3 Non-Linear Crystal Considerations . . . 69
6.3.1 Periodic Poling and Non-Linear Strength . . . 69
6.3.2 Absorption and Scattering . . . 70
6.3.3 Green Induced Infra-Red Absorption (GRIIRA) . . . 70
6.3.4 Damage Threshold and Photothermal Effects . . . 70
6.3.5 Grey Tracking . . . 71
6.3.6 Wedged Crystal . . . 71

7 The ANU Squeezer 73
7.1 ANU Cavity Properties . . . 73
7.1.1 Optimising the Wedged Crystal . . . 75
7.1.2 Measuring Threshold . . . 77
7.1.3 Measuring Intra-Cavity Loss . . . 79
7.2 The ANU experiment . . . 82
7.3 Squeezed State Control . . . 84
7.4 Coherent Control . . . 85
7.4.1 Coherent Control Model . . . 86
7.4.2 Modified Coherent Locking . . . 90

8 Detection of Squeezed States 93
8.1 Balanced Homodyne Detection Noise Couplings . . . 93
8.2 Electronic Issues . . . 94
8.2.1 Balanced Homodyne Photo-Detector Electronics Designs . . . 94
8.2.2 Dark Noise . . . 97
8.3 Optical Issues . . . 97
8.3.1 Optical Balancing . . . 97
8.3.2 Scattering Loss . . . 99
8.3.3 Beam Pointing and Mode Shape . . . 101
8.3.4 Parasitic Interference . . . 102
8.3.5 Detecting the Presence of Parasitic Interference . . . 105
8.3.6 Phase Jitter . . . 107
8.4 Squeezing from the ANU Squeezer . . . 108
8.4.1 March 2009 Squeezing . . . 108
8.4.2 May 2009 Squeezing . . . 109
8.4.3 January 2010 Squeezing . . . 109
8.4.4 April 2011 Squeezing . . . 111
8.4.5 November 2011 Squeezing . . . 111
8.4.6 Correcting for Detection Losses . . . 113

9 The LIGO Injection 115
9.0.7 LIGO Injection Aims . . . 115
9.1 LIGO Squeezer Properties . . . 117
9.1.1 Mirror Properties . . . 117
9.1.2 Threshold . . . 118
9.1.3 Losses . . . 118
9.1.4 Cavity Parameters . . . 119
9.1.5 Squeezing . . . 119
9.2 LIGO Enhancement . . . 120
9.2.1 Injection Losses . . . 120
9.2.2 Travelling Wave Cavity Backscatter Isolation . . . 122

10 Conclusions and Further Work 123
10.1 Summary of Squeezed Light Generation and Control for Gravitational-Wave Detection . . . 123
10.2 Further Work . . . 124
10.2.1 Squeezing Stability . . . 124

10.2.2 Dithering to Remove Parasitic Interference 124
10.2.3 Noise Coupling of the Coherent Locking Field 125
10.2.4 Frequency Dependent Squeezing 125

Appendix 1 **127**

Bibliography **131**

List of Figures

2.1 The effect of a gravitational-wave passing through a ring of free falling test masses. . . . 6
2.2 AURIGA. . . . 6
2.3 The effect of a passing gravitational-wave on a Michelson interferometer. . . . 7
2.4 The Laser Inteferometric Gravitational-wave Observatory (LIGO) in Livingston, Louisiana. . . . 8
2.5 The expected aLIGO sensitivity. . . . 9
2.6 The effect of radiation pressure on a test mass. . . . 10
2.7 The squeezing experiment on the GEO interferometer. . . . 14
2.8 The squeezing experiment at the LIGO site. . . . 15

3.1 The ball-on-stick representation of a coherent state with coherent amplitude. 21
3.2 The sideband representation of a coherent state. . . . 21
3.3 The ball-on-stick and sideband representations of the vacuum state of light. 22
3.4 The ball-on-stick and sideband representations of a coherent state of light. . 23
3.5 The ball-on-stick and sideband representations of a squeezed state of light. . 24
3.6 The ball-on-stick and sideband representations of a squeezed vacuum state of light. . . . 25
3.7 Modelling loss on an optical state. . . . 25
3.8 The effect of loss on various initial squeezed states. . . . 26
3.9 The effect of loss on an initial pure 10 dB squeezed state. . . . 27
3.10 Detection of a an optical state with a photodetector. . . . 27
3.11 The idealised balanced homodyne detector. . . . 29

4.1 The SHG system. . . . 35
4.2 The degenerate optical parametric amplifier (DOPA). . . . 36
4.3 The non-degenerate optical parametric amplifier (NDOPA). . . . 37
4.4 The optical parametric oscillator (OPO). . . . 38
4.5 Two phase matched and phase mismatched waves. . . . 39
4.6 Phasor diagrams showing phase matched, quasi-phase matched and phase mismatched systems. . . . 40

5.1 The optical cavity. . . . 44
5.2 The classical response of the two-ended triangle cavity. . . . 46
5.3 Cavity properties as the input coupler reflectivity is varied. . . . 47
5.4 The noise response of the optical cavity. . . . 49
5.5 The variance of the squeezing and anti-squeezing at the output of an OPO as the system is driven towards threshold. . . . 52
5.6 The effect of the escape efficiency on the maximum amount of squeezing that can be produced from any cavity. . . . 53

5.7 Non-linear gain of an OPA system as the input pump amplitude is driven to threshold. 56

6.1 The various cavity geometries; linear, monolithic, bow-tie and hemilithic. . 57
6.2 Backscattered light from the interferometer travelling towards the squeezer. 59
6.3 The geometry of the bow-tie cavity with the distances involved and the location of the two waists illustrated. 61
6.4 The tangential and saggital beam radii of the two fundamental waists in the bow-tie cavity as the distance between the curved mirrors is varied. . . 63
6.5 The tangential and saggital beam radii of the two pump waists in the bow-tie cavity as the distance between the curved mirrors is varied. 64
6.6 Cavity dispersion as the pump and fundamental fields traverse one round-trip of the cavity. 65
6.7 Contour plot of the vacuum squeezing produced, in dB relative to shot noise, from a singly-ended doubly resonant OPO cavity as the reflectivities of the input coupler at the fundamental and pump wavelengths are varied. . 66
6.8 The fundamental finesse, fundamental linewidth, escape efficiency, and the threshold power of the cavity as the reflectivities of the input/output coupler for the fundamental is varied. 67
6.9 Contour plot of threshold power for a singly-ended doubly resonant OPO cavity as the reflectivities of the input coupler at the fundamental and pump wavelengths are varied. 68
6.10 The periodically poled wedged crystal. 71
6.11 Propagation of the pump field through a PPKTP crystal with a 1 degree wedge. 72

7.1 The doubly resonant, bow-tie optical parametric oscillator. 73
7.2 To-scale layout of the bow-tie optical parametric oscillator. 74
7.3 Schematic of the experimental setup used to optimise the crystal parameters within the OPO. 76
7.4 Schematic of the procedure used to optimise the temperature and position of the non-linear crystal. 77
7.5 Experimental setup for measuring parametric gain. 78
7.6 Parametric gain as a function of pump amplitude for the ANU squeezer. . . 79
7.7 Theoretical fundamental field power response ratio of a singly-ended OPO (off phase matching). 80
7.8 Reflected power from the OPO as the OPO length was scanned using a PZT. 81
7.9 The layout of the ANU experiment. 83
7.10 The electronics of the ANU experiment, showing the locking loops used. . . 84
7.11 The fields involved in the coherent locking scheme. 85
7.12 Simplified schematic of the locking loops that make up the coherent locking scheme. 86
7.13 The powers of the transmitted and reflected CSF and GSF fields as the pump is increased up to OPO threshold and the phase of the CSF and GSF fields as the phase of the pump field is varied. 89
7.14 The fields involved in the modified coherent locking scheme. 90
7.15 The modified coherent locking electronics scheme. 91

8.1 Conceptual view of the major issues arising in low-frequency measurements with a balanced homodyne detector setup. 94
8.2 The two balanced homodyne detector schemes: the variable gain design, and the current subtracting design. 95
8.3 Shot noise traces from both current subtracting and variable gain homodyne detector designs, highlighting the importance of flicker noise. 96
8.4 The optical layout of the experiment designed to detect low-frequency noise sources in the balanced homodyne detector setup. 98
8.5 Relative intensity noise of the local oscillator field incident on the balanced homodyne detector. 99
8.6 The DC output of the balanced homodyne detector as a function of time, showing the effect of dust. 100
8.7 Balanced homodyne detector output traces illustrating a beam pointing event. 102
8.8 Scatter in the balanced homodyne detector leading to parasitic interference. 103
8.9 Parasitic interference from the signal port of the balanced homodyne detector. 104
8.10 Simulated results of the cyclic averaging technique. 105
8.11 The cyclic averaging technique applied to the balanced homodyne detector with a scatter source placed in the signal path. 106
8.12 The effect of phase jitter on the measurement of a vacuum squeezed state. . 108
8.13 The variance at the output of an OPO as the system is driven towards threshold power. 109
8.14 A snapshot of squeezing in the ANU squeezer over the years. 110
8.15 The box encasing the ANU experiment. 111
8.16 Measured squeezing from the ANU squeezer using quantum noise limited balanced homodyne detection. 112

9.1 Schematic of the squeezed light injection experiment into H1 at LIGO. . . . 116
9.2 The LIGO squeezer, an almost exact replica of the ANU squeezer design. . 118
9.3 Parametric gain as a function of pump amplitude for the LIGO squeezer. . 119
9.4 Measured squeezing from the LIGO squeezer using a quantum noise limited balanced homodyne detector. 121
9.5 LIGO H1 detector squeezing sensitivity with and without squeezing enhancement. 122

10.1 The Advanced LIGO expected noise budget showing roughly how frequency dependent squeezing would integrate. 125

List of Tables

4.1 A brief list of relevant squeezing results, pioneering squeezing in terms of magnitude, frequency, cavity design and/or wavelength. 34

6.1 ABCD matrices for various propagation events. 61
6.2 Crystal properties of some of the more commonly used periodically poled materials. 69

7.1 The properties of the ANU squeezer. 75

9.1 Mirror properties for the curved HR OPO mirrors. 118
9.2 Properties of the LIGO squeezer. 120

Summary of symbols and acronyms

Symbol	Definition
h_s	Interferometer strain
l	Interferometer arm length
$(\Delta Z)_{rp}$	Interferometer error due to radiation pressure
$(\Delta Z)_{pc}$	Interferometer error due to photon counting or shot noise
S_{Th}	Root power spectral density due to thermal fluctuations on the test mass
M	Mass of interferometer test mass
I	Laser field intensity
i	Photocurrent
k	Optical mode
E	Energy
ω	Optical angular frequency or angular detection frequency
$h\,(\hbar)$	Planck's constant & (reduced Planck's constant)
c	Speed of light in vacuum
b	Bounce number
τ_s	Storage time
τ_m	Measurement time
P	Optical power
$\vec{k}$	Wavevector
Δk	Wavevector mismatch
L_C	Coherence Length
$w(z)$	Beam radius at position z
$w(0)$	Beam waist
z_R	Rayleigh Range
q	Complex beam parameter
M	ABCD Matrix
d	Physical distance
d_{eff}	Crystal Non-linear susceptibility strength
$\hat{n}$	Photon number operator
ϕ	State phase angle
$\hat{D}(\alpha)$	Displacement operator
$\hat{S}(\rho)$	Squeezing operator
r	Squeezing parameter
MM	Mode matching
η_{vis}	Efficiency of spatial overlap from visibility measure

Symbol	Definition
$\hat{a}, (\hat{a}^\dagger)$	Annihilation & (creation) operators of intra-cavity mode
$\hat{A}, (\hat{A}^\dagger)$	Annihilation & (creation) operators of travelling wave
$\check{A}$	Frequency domain annihilation operator of travelling wave
$a_3^{crit}, (A_3^{crit})$	Intra-cavity & (external) field threshold amplitude
P_3^{crit}	External field threshold power
$\delta\hat{a}, (\delta\hat{a})$	First order fluctuating term of annihilation & (creation) operators
$\hat{X}^+, (\hat{X}^-)$	Amplitude & (phase) quadrature operators
$\delta\hat{X}^+, (\delta\hat{X}^-)$	First order fluctuation term of amplitude & (phase) quadrature operators
$\hat{X}^\theta$	General quadrature operator
$\Delta\hat{G}$	Standard deviation of observable $\hat{G}$
$V(\hat{G}) \equiv \Delta^2(\hat{G})$	Variance of the observable $\hat{G}$
α	Classical coherent amplitude
η	Beamsplitter power transmissivity or propagation efficiency
η_{esc}	Escape Efficiency
τ	Cavity round-trip time
δ	Frequency offset
L	Cavity optical path length
κ	Cavity decay rate
$R, (T)$	Power Reflectivity & (Transmittance)
Δ	Cavity detuning
g	Intra-cavity non-linear coupling strength
G	Combined non-linear coupling strength and intra-cavity pump amplitude
λ	Optical field wavelength
n	Refractive index
$\mathcal{V}$	Visibility
$\chi^{(2)}$	Second order non-linear susceptibility
$\mathcal{F}$	Finesse
Δ_ν	Cavity linewidth
$1, 2, 3$	Defines parameters for signal, idler and pump fields, respectively

Acronym	Definition
aLIGO	advanced LIGO
ANU	The Australian National University
AR	Anti-Reflection
AURIGA	Ultracryogenic Antenna for Astronomical Gravitational Investigation
CLL	Coherent Locking Loop
CMRR	Common-Mode Rejection Ratio
CSF	Coherent Sideband Field
DOPA	Degenerate Optical Parametric Amplifier
EOM	Electro-Optic Modulator
EPR	Einstein-Podolsky-Rosen
ETX500	Epitaxx 500 photodiode
FFT	Fast Fourier Transform
FSR	Free Spectral Range
FWHM	Full-Width-at-Half-Maximum
GEO600	British-German Cooperation for Gravity Wave Experiment
GRIIRA	Green Induced Infra-Red Absorption
GSF	Generated Sideband Field
GW	Gravitational-Wave
GWINC	Gravitational-Wave Interferometer Noise Calculator
H1	4km arm length interferometer at enhanced LIGO Hanford site
H2	2km arm length interferometer at enhanced LIGO Hanford site
HD	Homodyne
HR	High Reflectivity
KAGRA	Kamioka Gravitational Wave Detector
KNB	Potassium Niobate
KTP	Potassium Titanyl Phosphate
LIGO	Laser Interferometer Gravitational-wave Observatory
LN	Lithium Niobate
LSC	LIGO Scientific Collaboration
MC	Modecleaner Cavity
MgO:LN	Magnesium Oxide doped Lithium Niobate
MIT	Massachusettes Institute of Technology
MZ	Mach-Zender
NDOPA	Non-Degenerate Parametric Amplifier
Nd:YAG	Neodymium doped Yttrium Aluminium Garnet
NGO	New Gravitational-wave Observatory
OPA	Degenerate Optical Parametric Amplifier
OPO	Optical Parametric Oscillator
PDH	Pound-Drever-Hall Locking Technique
PPKTP	Periodically Poled Potassium Titanyl Phosphate
PZT	Lead Zirconate-Titanate piezo/electric ceramic
RBW	Resolution BandWidth
RMS	Root Mean Square
RoC	Radius of Curvature
RP	Radiation Pressure
SHG	Second Harmonic Generation
SN	Shot Noise
TEM	Transverse Electro-Magnetic
VIRGO	Italian-French Laser Interferometer Collaboration

Chapter 1

Introduction

Our understanding of the universe around us, from subatomic particles to the vast cosmos, shapes the way we live our lives. As precision measurements improve and we unravel more about how the world works, we are constantly transforming society in ways that are nearly impossible to predict. A perfect example is that of light amplification by stimulated emission of radiation, or the laser [1]. In a mere fifty years, the laser has become a building block of modern society. From entertainment, to beauty treatments, to surgery and probing of the fundamental laws of our universe, the usefulness of lasers has shaped the way our modern world works [2]. Our understanding of things on the opposite side of the size-scale have arguably had just as large an impact. Many years ago, astronomy led to the realisation that the earth is indeed round and not the centre of the universe, contrary to prior beliefs. This type of knowledge impacts significantly on theological ideas and in turn, on society itself.

Arguably, the advent of the telescope was the catalyst for the evolution of astronomy as a modern science. Early in the 17^{th} century, Galileo was the first to utilise their potential in observing details of the surface of the moon and the motions of the planets and the stars. Whilst telescopes evolved, becoming larger and providing more precise images, all early telescopes were restricted by the same fundamental limitation. Detections were made with the user's eye, and thus the information gathered was limited to the visible spectrum of light. Visible light, however, is a very small portion of the electromagnetic spectrum. Astronomy progressed, and the advantages obtained by widening the detection band were revealed. In the 1800's infra-red radiation was discovered by Herschel and was later used to study properties of the moon [3]. The detection band has continued to increase over time, such that we now routinely observe many orders of wavelengths of electromagnetic radiation, from gamma rays to radio waves [4, 5]. However, these forms of astronomy are all founded upon a common principle, the detection of electromagnetic waves. As such they all share similar advantages, and of course, similar disadvantages. Their detection can be hampered by various complications; they are prone to absorption, and follow the local space-time trajectory. One example of particular importance is that this means that direct detections of black holes cannot be made using electromagnetic observations. For many of the most violent, massive and arguably interesting phenomenon, another measurement method is required to complement the often limited information that electromagnetic waves can provide.

This is the role that gravitational-wave detection is hoped to fulfil in the near future. This thesis presents work that aims to improve the sensitivity of current interferometric gravitational-wave detectors via the integration of squeezed optical states. With the enhanced sensitivity that squeezed states are to offer, along with other improvements to current detector designs, we are nearing the era of gravitational-wave astronomy.

1.1 Thesis Structure

The thesis is structured as follows;

- Chapter 2 - Background

  The exciting realm of ground based gravitational-wave astronomy is introduced. Noise sources in current and next generation detectors are detailed and in particular the role of quantum noise sources is highlighted. Methods for reducing quantum noise are discussed and optical squeezing is introduced as an exciting new method of reducing the quantum noise in future detector designs.

- Chapter 3 - Quantum States of Light

  The mathematical description of the states of light used in this work is presented in this chapter. The ball-on-stick representation and the sideband representation for visualising the states of light are detailed. Using these representations, the properties of squeezed states, vacuum states, and coherent states are all introduced.

- Chapter 4 - Generation of Squeezed States

  In this chapter the generation of squeezed states is introduced. The various methods available are first discussed before detailing the production of squeezing from the second order non-linearity. Second harmonic generation, non-degenerate optical parametric amplification and optical parametric oscillation are introduced as modes of operation for the second order non-linearity.

- Chapter 5 - Optical Cavities

  The properties of optical cavities are introduced in this chapter. The cavity equations of motion are first introduced and then used to derive the classical and quantum response of an empty cavity. The finesse, free spectral range, linewidth and circulating power are defined. The cavity equations of motion are then investigated for an optical cavity containing a non-linear medium. From this work, the escape efficiency, non-linear gain, and expected squeezing from such a cavity are investigated.

- Chapter 6 - The Doubly Resonant Bow-Tie Optical Parametric Oscillator

  The methods and philosophies for designing a travelling wave cavity cavity are detailed in this chapter. The available non-linear mediums, techniques for overcoming intra-cavity dispersion and the choices involved when deciding all remaining parameters are discussed. The effect that variation of the cavity parameters has is illustrated.

- Chapter 7 - The ANU Squeezer

  The details of the ANU squeezer are given in this chapter. All of the chosen cavity parameters are stated before the methods for accurately measuring threshold and intra-cavity loss are introduced. A layout of the optical and electronic components is given. The coherent locking scheme used to control the angle of the squeezed state is investigated and a modified coherent locking scheme is introduced as a means for simplifying and improving the original scheme. This work was jointly undertaken with Sheon Chua and Conor Mow-Lowry at the ANU under the supervision of David McClelland, Ping Koy Lam, Ben Buchler and Daniel Shaddock, with much appreciated input from Kirk McKenzie and which utilised a cavity design originally devised

by Nicolai Grosse. Improvements to the squeezer were also made by Sheila Dwyer, who was involved for a number of months on the project. The later measurements of the squeezing produced by this squeezer were taken using a homodyne detector supplied by Roman Schnabel and Henning Vahlbruch.

- Chapter 8 - Detection of Squeezed States

  This chapter details an investigation into understanding the low-frequency performance of the homodyne detector setup used to measure the squeezed states. Many noise sources are investigated and methods for detecting and overcoming these noise sources are presented. The homodyne detector supplied by Roman Schnabel and Henning Vahlbruch plays a large role in this work. A short history of the squeezing produced by the ANU squeezer illustrates the various improvements made in order to improve the squeezing over a timespan of nearly 3 years. The work on detecting low-frequency noise sources as presented in this chapter was primarily undertaken with Conor Mow-Lowry, with some input from Sheon Chua, and the modified coherent locking scheme was devised by Conor-Mow Lowry and Kirk McKenzie. The modified coherent locking scheme was implemented by Sheon Chua.

- Chapter 9 - The LIGO Injection

  In this chapter the details of the LIGO squeezed light source are introduced. The cavity properties, including threshold and intra-cavity loss are given. The squeezing produced by the LIGO cavity is presented and the results are discussed. The integration of the squeezer into LIGO is briefly discussed but greater detail pertaining to the integration of the squeezer will be presented in the PhD theses of Sheon Chua and Sheila Dwyer as well as future papers. The squeezer was constructed and tested by myself, Sheon Chua, Conor Mow-Lowry, Alexander Khalaidovski and Sheila Dwyer at MIT. The LIGO squeezed light injection project was a large collaborative project, with too many people involved to mention them all, but particular mention is made to Nergis Mavalvala, Daniel Sigg, Keita Kawabe, Sheila Dwyer and Lisa Barsotti, who I worked most closely with during my time on this project. A number of months were personally spent aiding in the integration of the squeezer with the LIGO detector.

- Chapter 10 - Conclusions and Further Work

  The conclusions from the work presented in this thesis are summarised. The future work on integration of squeezed light into interferometric gravitational-wave detectors that needs to be undertaken, in light of the work presented in this thesis, is then discussed.

1.2 Publications

- M. Stefszky, C. Mow-Lowry, K. McKenzie, S. Chua, B. Buchler, T. Symul, D. McClelland and P. K. Lam, "An investigation of doubly-resonant optical parametric oscillators and nonlinear crystals for squeezing," J. Phys. B: At. Mol. Opt Phys. **44**, (2011).

- M. S. Stefszky, C. M. Mow-Lowry, S. S. Y. Chua, D. A. Shaddock, B. C. Buchler, H. Vahlbruch, A. Khalaidovski, R. Schnabel, P. K. Lam, and D. E. McClelland, "Bal-

anced Homodyne Detection of Optical Quantum States at Audio-Band Frequencies and Below," Clas. Quantum Grav. (submitted), (2012).

- S. Chua, M. Stefszky, C. Mow-Lowry, B. Buchler, S. Dwyer, D. Shaddock, P. Lam, and D. McClelland, "Backscatter tolerant squeezed light source for advanced gravitational-wave detectors," Opt. Lett. **36**, 4680 (2011).
- The LIGO Scientific Collaboration, "A gravitational wave observatory operating beyond the quantum shot-noise limit," Nat. Phys. **2083**, DOI:10.1038 (2011).
- The LIGO Scientific Collaboration, "Implementation and testing of the first prompt search for gravitational wave transients with electromagnetic counterparts," ArXiv:1109.3498v2 (2011).
- The LIGO Scientific Collaboration, "Search for Gravitational Wave Bursts from Six Magnetars," Astrophys. J. **734**, (2011).
- The LIGO Scientific Collaboration, "Beating the Spin-Down Limit on Gravitational Wave Emission from the Vela Pulsar," Astrophys. J. **737**, (2011).
- S. Chua, M. Stefszky, C. M. Mow-Lowry, B. C. Buchler, K. McKenzie, D. A. Shaddock, P. K. Lam, and D. E. McClelland, "Quantum Squeezing in Advanced Gravitational Wave Detectors" IJMPD **20**, 2043 (2011).
- The LIGO Scientific Collaboration, "Enhancement of the astrophysical reach of a gravitational wave observatory using squeezed states of light," Unpublished Manuscript.

Chapter 2

Background

In this chapter, gravitational-waves are introduced. The properties of gravitational waves are briefly discussed and two methods for their detection, resonant bar detectors and interferometric detectors, are detailed. The noise sources in interferometric detectors are then listed and squeezed light is introduced as a solution to reducing the quantum noise in interferometric detectors.

2.1 Gravitational-Waves

Since the advent of general relativity it is now common to consider the three space dimensions and one time dimension as part of a single manifold, known as space-time. It was postulated by Albert Einstein early in the 20^{th} century that accelerating masses should release energy in the form of gravitational-waves, or ripples in space-time [6, 7]. These gravitational-waves oscillate perpendicular to the direction of propagation and act to shrink space in one direction, whilst simultaneously expanding in the perpendicular direction, as shown in Figure 2.1. The magnitude of these oscillations is incredibly small. An extensive review of the various sources and the characteristics of these waves is given in [8, 9]. By the time these waves propagate from their original sources and reach the Earth, we might expect the largest gravitational-wave sources to have a strain amplitude, h_s, on the order of 10^{-20} [10], where strain is defined as the change in length, Δl, over the length, l, in question

$$h_s \equiv \frac{\Delta l}{l}. \tag{2.1}$$

Many searches have been made [11, 12, 13], but direct detection of gravitational-waves has yet to be achieved. However, strong evidence for their existence has been identified. The binary pulsar system PSR 1913+16 was observed for many decades by Russell Hulse and Joseph Taylor [14]. The point of closest approach between these two pulsars, the periastron, was seen to decrease over the measurement time of 30 years. The decrease seen agreed with the change in periastron predicted by Einstein's theory of general relativity with startling accuracy. The challenge now is to directly detect these waves and retrieve from them the information that they carry about the systems from which they originated, thus opening to door into gravitational-wave astronomy.

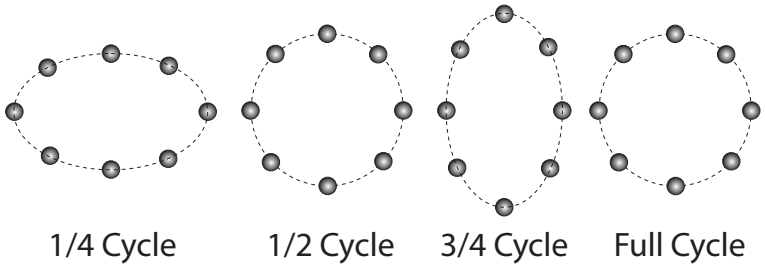


Figure 2.1: The effect of a gravitational-wave passing through a ring of free falling test masses. As the wave passes (into the paper), space-time is stretched along one axis and compressed in the orthogonal axis. This action oscillates at the frequency of the passing gravitational-wave.

2.2 Detecting Gravitational-Waves

The incredibly small amplitude of gravitational-waves makes their detection an overwhelmingly challenging task. When Albert Einstein first postulated their existence early in the 20^{th} century, there was conjecture, even from Einstein himself, as to whether these waves were mathematical artefacts or whether they would theoretically able to be detected [15]. However, it was not long before people were devising various methods that should, in theory, be capable of detecting the largest gravitational-waves here on Earth [16, 17].

2.2.1 Resonant Bar Detectors

Figure 2.2: AURIGA, cryogenically cooled resonant bar detector [18].

The first attempts at observing gravitational-waves were made by Joseph Weber in the 1960's [16]. Weber's design was that of a resonant bar. Resonant bars work on the principle that a gravitational-wave passing through it, with sufficient magnitude and the correct frequency, will excite a resonance mode within the bar. The resonance frequency of Weber's bar was tuned to a frequency of around 1660 Hz with a strain sensitivity of approximately $10^{-17}\frac{1}{\sqrt{Hz}}$ [19, 20]. The chance of detecting events with this detector was incredibly small, due to its very narrow bandwidth, of 10s of millihertz [19] and limited

sensitivity. Even now, the sensitivity of resonant bar detectors is limited to strains of approximately $10^{-21}\frac{1}{\sqrt{Hz}}$, which is believed to be enough to detect only the strongest, and as such the rarest, of sources. Bandwidths of a few hertz are typical in current resonant bar detectors and up to a few kilohertz is attainable in resonant sphere systems, with resonance frequencies typically in the kilohertz region or above [21].

2.2.2 Interferometric Detectors

An alternative method of detecting these waves, proposed by Thorne, Drever and Weiss [22], began construction in 1992 [23], after many years of research into design choices and various noise sources [24, 25, 26]. This new class of detector would use a laser interferometer design to increase both sensitivity and bandwidth over resonant bar designs. The geometry was to resemble that of a typical Michelson interferometer with arm lengths of 4 km. The effect of a passing gravitational-wave on a Michelson interferometer is shown in Figure 2.3. This interferometer is incredibly sensitive to path length differences in the two arms and it is this property that would be used to detect gravitational-waves. The two end mirrors are suspended such that they act as free falling test masses. A gravitational-wave passing through the plane of the detector acts to increase the length of one arm by displacing one of the test masses, whilst decreasing the length in the other. An overview of interferometric gravitational-wave detector design is given in [27].

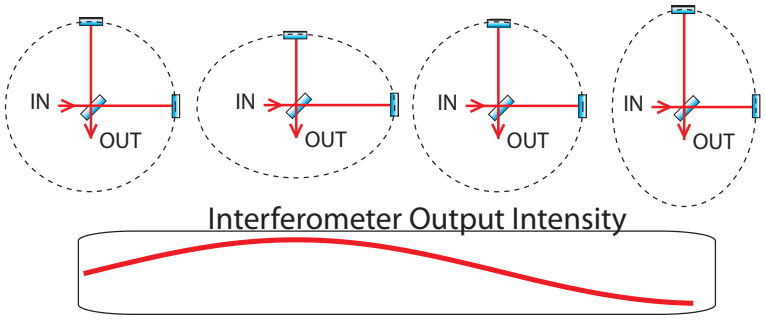


Figure 2.3: The effect of a passing gravitational-wave on a Michelson interferometer operating in the middle of an interference fringe. A gravitational-wave passing through the interferometer causes the length of the two arms to oscillate. This changes the relative phases of the laser fields from the two interferometer arms, 1 and 2, resulting in modulation of the intensity of the output field.

The detector shown in Figure 2.4, the Laser Interferometer Gravitational-wave Observatory (LIGO), is more sensitive and has a much larger bandwidth than current resonant bar detectors. LIGO's full detection band is from 10 Hz to 10 kHz, which is also known as the *audio detection band*. However, the the sensitivity of LIGO varies over this band. The device is currently undergoing upgrades but during its last science run, the device had a strain sensitivity of approximately $10^{-23}\frac{1}{\sqrt{Hz}}$ at around 100 Hz and greater than 10^{-22} from around 60 Hz up to approximately 1 kHz [28]. The device in this state was known as enhanced LIGO, further improvements are currently under way, with the next generation to be known as aLIGO [29]. Looking for gravitational wave sources at frequencies higher than these is not useful due to the fact that there is a limit as to how fast the massive bodies producing these waves can move. In LIGO's full detection band, the

most promising source of gravitational-waves are the inspiral of massive bodies, such as black holes and neutron stars, as well as the burst signature from supernovae, gravitational radiation background and pulsars [10, 8]. A comprehensive overview of the properties of these sources and how their magnitudes relate to the various detectors is given by Schutz in [30] and the expected detection rates are presented in [31].

A network of detectors is necessary for both triangulation and increasing the confidence of events [10]. The device in Figure 2.4 is one of two 4 km interferometers situated in the United States. There is also a 3 km detector in Italy known as VIRGO [32] and another one in Germany with an arm length of 600 m known as GEO600 [33]. Together, these make up the current network (although not all are currently in operation) of interferometric gravitational-wave detectors.

Figure 2.4: The Laser Inteferometric Gravitational-wave Observatory (LIGO) in Livingston, Louisiana.

2.3 aLIGO Noise Sources

Interferometric gravitational-wave detectors are sensitive to noise sources that either directly change the position of one of the test masses or noise sources that mimic such a change. The aLIGO noise budget, as illustrated in Figure 2.5, shows that the expected major contributing noise factors below approximately 100 Hz are; thermal noise, seismic noise and quantum noise. At these frequencies, the quantum noise is dominated by radiation pressure noise. Above 100 Hz, the dominant quantum noise source is shot noise, resulting from photon counting error at the detector readout.

2.3.1 Low-frequency Noise Sources

Radiation Pressure Noise

It is expected in the next generation of interferometric gravitational-wave detectors that seismic and thermal noise will be so small in the region from around 10 Hz to 100 Hz that quantum noise will become the dominating source in this frequency regime. This is shown in Figure 2.5. The quantum noise source in this frequency band is known as *radiation pressure noise*. Photons reflecting off the end mirrors, or test masses, will impart

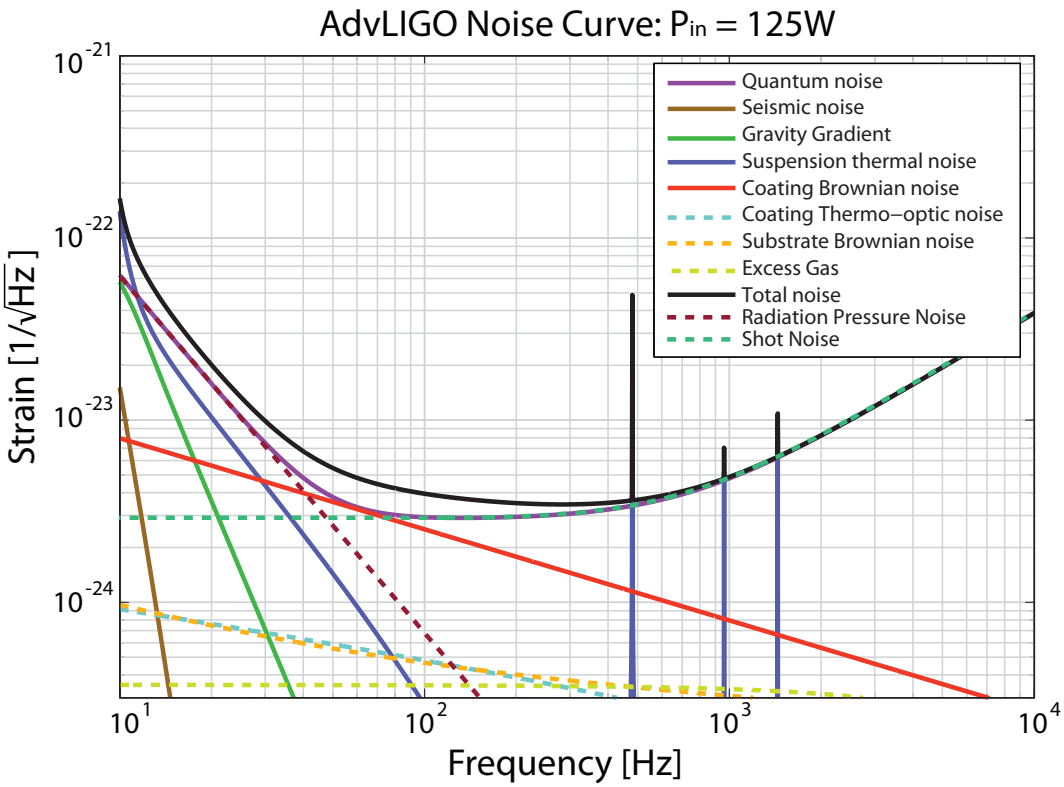


Figure 2.5: The expected aLIGO sensitivity. Plot produced using GWINC, developed by the LSC community.

a momentum kick to the mirror, conserving momentum in the collision. The momentum imparted is proportional to the number of photons, and hence proportional to the intensity of the incident field. If the intensity of the field were constant, then the constant stream of photons hitting the mirror would impart momentum that would simply alter the free resting position of the mirror. However, as the intensity of the field fluctuates, the position of the mirror will also fluctuate. If the intensity noise were classical, then the radiation pressure force felt by each test mass would be identical and this effect could be optically subtracted at the beamsplitter. However, due to the quantum mechanical nature of light, uncorrelated vacuum noise will be incident on each test mass. One can regard these fluctuations as vacuum fluctuations entering the empty input port of the interferometer beamsplitter [25], the same port where the interferometer output typically exits as in Figure 2.3. This uncorrelated noise cannot be subtracted and will be detected, masking the gravitational-wave signal. The radiation pressure error is written [34]

$$(\Delta Z)_{rp} \simeq \frac{b\hbar 2\pi f \tau_m}{mc} \left(\frac{P\tau_m}{\hbar 2\pi f}\right)^{\frac{1}{2}} e^r, \tag{2.2}$$

where $(\Delta Z)_{rp}$ is the radiation pressure error, or uncertainty due to radiation pressure, b is the number of bounces at each end mirror, τ_m is the measurement time, P is the optical

power, f is the frequency of the laser light and $\hbar$ and c are the reduced Planck constant and speed of light respectively. The squeezing factor, r, introduced in detail in § 3.5.3, is a measure of the magnitude and angle of a squeezed state injected into the second input port of the beamsplitter of the interferometer. By convention, a positive squeezing factor implies the use of a phase squeezed state. Therefore, radiation pressure error is reduced through injection of an amplitude squeezed state, with a negative r value, into the second input port of the interferometer.

Radiation pressure noise is reduced by either decreasing the number of photons hitting the mirror, decreasing the velocity imparted to the mirror by each photon, or by directly reducing the uncorrelated noise. The number of photons can be reduced by decreasing the circulating power by varying the parameters P, τ_m and b, the velocity imparted to the mirror by individual photons is reduced by increasing the mirror mass, m, or reducing the energy of the photons, f. The uncorrelated quantum noise can be reduced via the introduction of an amplitude squeezed state. The effect of radiation pressure on a test mass is shown in Figure 2.6. In Figure 2.6, the case where no laser field is incident,

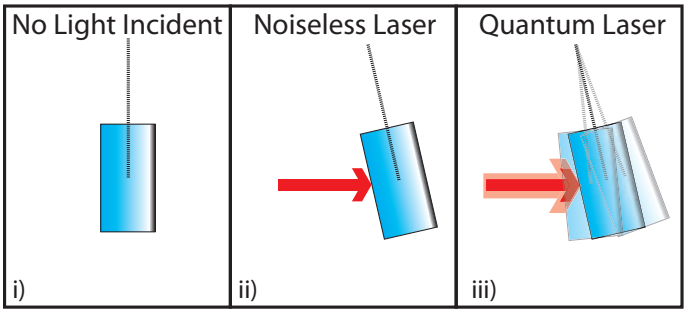


Figure 2.6: The effect of radiation pressure on a test mass for three cases: i) with no light incident; ii) with a theoretical incident laser that has no classical or quantum noise; and iii) with a laser that has quantum noise.

i), shows the test mass hanging freely. In the case of the completely noiseless laser, ii), which cannot physically exist, the photons impart momentum to the mirror resulting in a constant force. In the final case, there is a quantum noise limited laser source, a so called coherent state, incident on the test mass. The incident laser field no longer has a constant intensity, but rather fluctuates, a direct consequence of Heisenberg's uncertainty principle. The varying intensity imparts a varying force and the position of the mirror fluctuates. The forces imparted to the test masses in the two arms are uncorrelated due to the vacuum noise entering the second input port of the interferometer, resulting in a noise source which cannot be subtracted.

Thermal Noise

Thermal noise causes mechanical expansion and contraction that changes the length of the interferometer arms relative to one another. There are a few main sources of thermal noise; the suspension wires or fibres that hold the test masses in place [35], the internal motion of the test masses [36] and the thermal motion of the coatings on the optics [37, 38]. Using the fluctuation-dissipation theorem [39], the root power spectral density of thermally

induced fluctuations on the test mass as read by the interferometer is written [37]

$$S_{Th} = \frac{k_B T}{\pi^2 \Omega^2} \mathrm{Re}\{\mathbb{Y}(\Omega)\}, \tag{2.3}$$

where k_B is Boltzmann's constant, T is the temperature of the test mass, Ω is the frequency of the fluctuations and $\mathrm{Re}\{\mathbb{Y}(\Omega)\}$ is the mechanical admittance of the test mass to a cyclic pressure distribution that has the same form as the intensity profile of the interferometer beam. This term can be calculated with knowledge of the mechanical loss angle of the test mass response to the applied cyclic Gaussian pressure distribution [36].

One method of overcoming this noise source is to cool the masses. A new detector, to be built in Japan, known as KAGRA (previously called LCGT) [40, 41], is planning to use precisely this technique in order to minimise their thermal noise contribution, cooling their test masses and fibre suspensions to tens of Kelvins [42]. The other method for reducing thermal noise is to increase the quality factor and reduce the intrinsic loss of the materials involved. In aLIGO, the thermal noise in the test masses is decreased through the use of high quality factor fused silica test masses, the replacement of steel wires with silica fibres, and the use of advanced coatings [43, 28, 37].

Seismic Noise

Below frequencies of tens of hertz, seismic noise limits the performance of ground based detectors [44]. The magnitude of seismic noise increases rapidly at lower frequencies. Seismic noise causes motion in the test masses, resulting in interferometer noise.

Significantly improved seismic isolation systems have been produced for aLIGO that will reduce the coupling of seismic noise to test mass motion [45], but ultimately seismic noise will limit the very low-frequency performance of ground based gravitational-wave detection, below 10's of hertz. One method of reducing seismic noise is to build a detector underground, where seismic noise is reduced [46]. KAGRA is planned to be built underground for this purpose and is expected to have a sensitivity of 10^{-23} or greater down to 50 or 60 Hz depending upon the operating conditions [40]. Seismic noise can be bypasses altogether operating a gravitational-wave detector in space. This is the goal of what was originally called the Laser Interferometer Space Antenna (LISA) but is now known as the New Gravitational-Wave Observatory (NGO) [47, 48]. The operational band of this detector is expected to be from approximately 0.1 mHz up to around 1 Hz.

2.3.2 High Frequency Sensitivity

Shot Noise

In the high frequency region of the spectrum, frequencies above 100 Hz, the noise is almost entirely quantum in origin [49], as shown in Figure 2.5. All of the classical and technical noise sources drop off as the frequency increases. The photons arrive at the photodetector with a Poissonian distribution in the mean number of photons per unit time. The Fourier transform of this distribution is white noise across the entire spectrum. The optical response of the arm cavities shapes this noise source into that seen in Figure 2.5. The photon counting error, or error due to shot noise, of a Michelson interferometer is given by [25],

$$(\Delta Z)_{pc} \simeq \frac{c}{4b\pi f} \left(\frac{\hbar 2\pi f}{P\tau_m} \right)^{\frac{1}{2}} e^{-r}, \tag{2.4}$$

where $(\Delta Z)_{pc}$ is the photon-counting error, or uncertainty due to shot noise, b is the number of bounces at each end mirror, τ_m is the measurement time, P is the optical power, f is the frequency of the laser light, r is the squeezing factor, and $\hbar$ and c are the reduced Planck constant and speed of light respectively. As with Equation 2.2, a positive squeezing factor, which in this case will result in a negative exponent, implies the use of a phase squeezed state.

The strain sensitivity can be improved by increasing the values of b, f, l, P, or r. However, for each of these variables, increasing their values will eventually reach a practical limit. Perhaps the simplest method of reducing the photon counting error is to increase the optical power, P. However, increased power will eventually lead to thermal heating issues [50] and, as shown in § 2.6, increases the radiation pressure noise. Likewise, increasing the laser frequency requires investigation into new low loss materials and coatings and the development of high-power laser sources at these frequencies.

Equation 2.2 shows that most of the processes that reduce the shot noise increase the radiation pressure noise. This can be understood with a very simple principle. Error due to shot noise is reduced with increased power, whereas error due to radiation pressure noise increases with higher powers. Of note is that the same squeezing used to decrease the photon counting noise, a positive squeezing factor, will act to increase the radiation pressure noise. This is because squeezing in the phase quadrature results in anti-squeezing in the amplitude quadrature, conserving Heisenberg's uncertainty principle. The anti-squeezing increases the noise in the amplitude quadrature, resulting in an increase in the radiation pressure noise. Dependent upon the frequency band of interest and the dominating noise sources at these frequencies, phase squeezing, amplitude squeezing, or some squeezing angle between the two might be optimal. This is not made clear in the presented equations for the error due to radiation pressure noise and shot noise. It is shown in Figure 2.5 that radiation pressure noise dominates at low frequencies, and shot noise dominates at high frequencies. Progress has been made on producing frequency dependent squeezed states, which are squeezed in the amplitude quadrature at low frequencies, and squeezed in the phase quadrature at high frequencies [51]. In this way, squeezing can reduce the error due to both radiation pressure noise and shot noise in an interferometer simultaneously.

Increasing Laser Power

Equation 2.4 shows that the effect of shot noise can be reduced by increasing the laser power. The shot noise scales as the square root of the power of the optical field, whereas the signal scales as the power, therefore, an increase of the laser power by a factor of 4 will increase the signal to noise ratio of a shot noise limited system by 3 dB. However, increasing the laser power comes at a cost. Increased laser power results in larger radiation pressure noise and increased thermal effects. First, thermal lensing [43], and eventually parametric instabilities [52], will arise in the interferometer as the laser power is increased. These unwanted effects will reduce the efficiency and perhaps even more importantly, the stability of the detector.

The next generation of detectors will have the ability to operate at circulating powers of up to 0.5 MW. The original LIGO design had circulating powers of approximately 10 kW. The laser power in the initial LIGO setup was already large enough to require compensation for thermal effects [53]. With increased laser power, aLIGO is expected to require more sophisticated thermal compensation systems to ensure stable, reliable

operation [54, 50, 43]. Eventually a limit will be reached where the laser power cannot be increased, whilst maintaining reliable interferometer operation. It will then be necessary to decrease the shot noise of the device through other means.

2.4 Quantum Enhancement

In order to improve the sensitivity of ground based gravitational-wave detectors for and beyond the next generation, methods will need to be introduced that are capable of reducing the quantum noise in the interferometer without increasing the operational power. Many such methods have been proposed, such as variational readout, quantum non-demolition measurements and the use of squeezing [49]. Perhaps the simplest method is through the use of squeezed states due to the fact that it is fully compatible with current detector designs (See Chapter 8. Some of the other quantum enhancement techniques require complete redesigns of current interferometers.

2.5 Optical Squeezing

Optical squeezing is a means of reducing the quantum noise in a laser field, under certain conditions, below the standard limit. The usefulness of these states has been proven in various applications such as spatial displacement measurements [55] and atomic spectroscopy measurements [56]. In 1981 Caves showed that the introduction of squeezing to a Michelson interferometer was theoretically capable of improving the sensitivity of the system [25]. However, squeezed states had not yet been produced, let alone ones that were suitable for such an application. The first squeezed state, exhibiting a mere 7 % reduction in the shot noise at a measurement frequency of 422 MHz, was observed in 1985 by Slusher *et al.* using a four wave mixing process [57]. For many years thereafter, a multitude of systems were used to produce varying levels of squeezing [58]. A goal was set by some members of the squeezing community to produce squeezed states with 10 dB of quantum noise reduction down to frequencies as low as 10 Hz [59, 49, 60]. Squeezing with these properties would provide a substantial sensitivity enhancement to interferometric detectors and would display squeezing across the entire detection band of ground based detectors, marking a time when squeezing had matured to a point where the production of squeezed states was a solved problem. This thesis presents the first measurements that have achieved this 10 dB at 10 Hz goal for squeezed light production.

2.5.1 Low-frequency Squeezing

Early squeezed states could not be used for interferometric gravitational-wave detection due to the fact that they degraded at frequencies below a few MHz [61]. Recall that Figure 2.5 shows that ground based detectors require squeezing in the audio-band, from tens of hertz up to ten kilohertz. The majority of experiments aimed at producing low-frequency squeezing have utilised the second order non-linearity in crystalline materials placed in optical cavities to generate these states. Attempts were made in one of these systems to recover the squeezing at low-frequencies by building two identical squeezers and subtracting excess noise, and while it did improve the low-frequency performance, this technique still only resulted in squeezing down to approximately 200 kHz [62].

The breakthrough was made when McKenzie *et al.* determined that the presence of a bright seed field, traditionally used to control the phase of the squeezed state, was coupling

noise into the state [63]. It was determined that removing the bright seed, resulting in what is known as a vacuum squeezed state, should allow for the production of squeezed states in the audio-band. However, new control techniques would need to be introduced to control the phase of these squeezed states, which lack a coherent amplitude to use as a phase reference. Using a technique called "quantum noise locking", McKenzie *et al.* observed squeezing down to frequencies of a few hundred hertz [63]. Below these frequencies it was believed that beam pointing [64] and dust [65] were reducing the measured squeezing.

Progress was made regarding the control of these states. A new locking technique, later called "coherent locking", was developed in order to more stably control the vacuum squeezed states [66, 67]. However, degradation in the squeezing was still seen at frequencies below a couple of hundred hertz. In 2007 Vahlbruch *et al.* managed to measure large levels of squeezing, up to 6.5 dB, down to frequencies as low as tens of Hertz. The improvement was attributed to the reduction of parasitic interference (see § 8.3.4. A squeezing source had now been produced that was able to deliver a reasonable magnitude of squeezing that covered the complete frequency range of large scale interferometers. The demonstration of more squeezing, particularly at lower frequencies, was still desirable and additionally, the integration of squeezing into the detectors was a challenge that still needed to be addressed.

2.5.2 Squeezing Enhancement of Interferometers

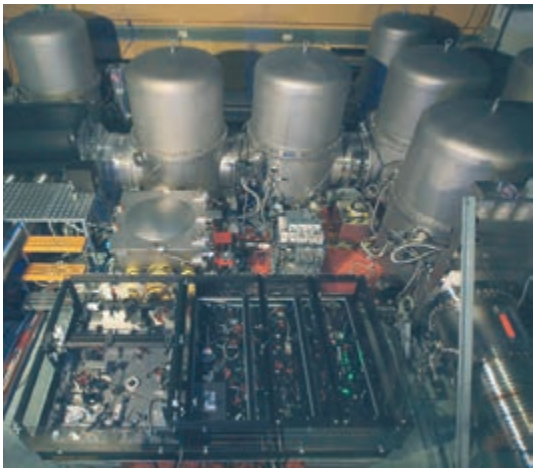


Figure 2.7: The squeezing experiment on the GEO interferometer. The squeezing is produced on the table at the bottom of the picture and injected into the interferometer located inside the vacuum chambers [68].

In the absence of low-frequency squeezing, the first experiments showcasing the improvement in sensitivity offered by squeezed states was demonstrated at higher frequencies. In 1987 Xiao *et al.* showed sensitivity improvement in a squeezing enhanced interferometer at a frequency of 1.6 MHz [69]. In 2002, McKenzie *et al.* showed a sensitivity improvement in a power-recycled squeezing enhanced Michelson interferometer at 5.5 MHz. In 2008 Goda *et al.* [70] showed a 3 dB improvement in a prototype gravitational-wave detector down to frequencies of tens of kilohertz. Only very recently has broadband improvement been demonstrated in a large scale interferometer, GEO600 [68], the experimental setup

of which is shown in Figure 2.7. Reduction in the shot noise of up to 3 dB was achieved across the entire shot noise limited region, from frequencies of a couple of hundred of hertz up to ten kilohertz. A duty cycle of 99 % was achieved and the squeezing did not degrade the performance of the detector at frequencies below hundreds of hertz. The squeezing table is modular in design, meaning it can be switched off and the interferometer is able to function as normal. This landmark achievement showed that squeezing is viable in large scale interferometers to improve the high frequency sensitivity of the device. However, a more sensitive device was needed to demonstrate the effect of squeezing on the lower frequency region.

2.5.3 Squeezing in LIGO H1

The work presented in this thesis was part of an international collaboration, involving the ANU, Caltech, the Max Planck Institute for Gravitational Physics and MIT, to inject squeezing into the LIGO detector situated in Hanford, Washington State, in the United States of America. Enhanced LIGO had just finished its sixth science run and the squeezed light injection occurred as upgrades towards aLIGO were beginning [71, 29]. The enhanced LIGO Hanford detector had two interferometers within its beam tubes, H1 and H2. H2 was a shorter detector with arm lengths of 2 km and H1 had arm lengths of 4 km, which, along with the LIGO Louisiana site, had the longest arm lengths of any current detector. These two LIGO detectors were the most sensitive interferometers ever made. In particular, their low-frequency sensitivity surpassed the performance of any other interferometer [72]. Figure 2.8 shows a picture of the squeezing table at the LIGO Hanford site.

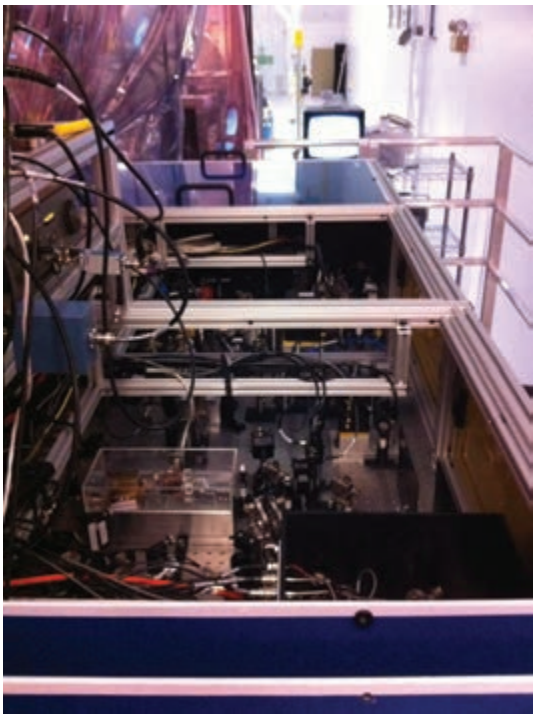


Figure 2.8: The squeezing experiment at the LIGO site.

It was previously noted that squeezing has already been shown to be viable in the GEO600 interferometric gravitational-wave detector [68]. The differentiating factors be-

tween this experiment and the LIGO injection are that the squeezer itself was to be a very different design (discussed in Chapter 6), and that the improved sensitivity of the LIGO H1 interferometer at lower frequencies would allow for the investigation of noise couplings at these frequencies. With the injection of squeezed light, it was also expected that the sensitivity of the device would exceed that previously seen by any gravitational wave detector.

Chapter 3

Quantum States of Light

This chapter introduces the formal description of the quantised electromagnetic field. This will form the basis for an investigation into squeezed states and their application in gravitational-wave detectors. First, the quadrature operators of the light field and their uncertainty relation are introduced. From here the necessary properties of these states of light and various methods of visualising these states are detailed. The effect of loss on squeezed states is investigated and the standard method for detecting these states of light, balanced homodyne detection, is analysed.

3.1 The Quantum Mechanical Electromagnetic Field

A quantum mechanical picture of light begins with quantization of the electromagnetic field. The details of this procedure are not described here but can be found in [73]. The result is that we can describe an optical mode, k, in terms of the annihilation and creation operators, $\hat{a}_k$ and $\hat{a}_k^\dagger$ respectively, in much the same way that one can describe a quantum harmonic oscillator [74]. As with the quantum harmonic oscillator, vacuum energy is present for every optical mode k, which is this case is given by $E = \frac{1}{2}\hbar\omega_k$. A classical description of the electromagnetic field does not include a description of these vacuum modes, and as such they are a defining feature of the quantum mechanical description. These vacuum fluctuations are the lowest energy state of the field and can be regarded as the instantaneous creation of particle and anti-particle pairs. In order to measure quantum mechanical states of light, including the vacuum state, it is first necessary to develop a framework for describing them.

3.1.1 Quadratures

Electric fields are often written in complex form, given by $E = Ce^{i\omega t}$, where C is the amplitude of the wave, ω is the angular frequency and E is the electric field [75]. In this form, equations are often compact and easy to manipulate. However, in some situations, this is not the most useful representation. The quadrature representation of the electromagnetic field is an alternative method of describing the electromagnetic field. The operators arising form this description are Hermitian, and therefore observable in the lab, providing a method by which quantum mechanical theories can be tested.

3.1.2 Quadrature Operators

The annihilation and creation operators, $\hat{a}$ and $\hat{a}^\dagger$ respectively, can be used to describe the electromagnetic field but they are not Hermitian, and are therefore not observable. In

order to relate measurements in the lab to the theory it is necessary to introduce Hermitian operators. This is the role that the amplitude, $\hat{X}^+$, and phase, $\hat{X}^-$, quadrature operators fulfil, they can be written [76]

$$\hat{X}^+ = \hat{a} + \hat{a}^\dagger \tag{3.1}$$

$$\hat{X}^- = i(\hat{a}^\dagger - \hat{a}). \tag{3.2}$$

By definition, the amplitude quadrature operator is chosen to have an angle that aligns with the coherent amplitude of the state. Under certain circumstances this can complicate issues because the chosen reference frame may rotate. We can also write a more general form of these operators

$$\hat{X}^\theta = \hat{a}e^{-i\theta} + \hat{a}^\dagger e^{i\theta}, \tag{3.3}$$

where θ is the angle of the field relative to a second field. The amplitude and phase quadrature operators are particular cases of the general quadrature operator. For $\theta = 0$ Equation 3.3 results in the amplitude quadrature operator and for $\theta = \frac{\pi}{2}$ it returns the phase quadrature operator. Provided that they are orthogonal to one another, which is to say that there is a $\frac{\pi}{2}$ difference between the angles, θ, of the operators, any pair of quadrature operators can be used to describe the various states of light.

3.2 The Heisenberg Uncertainty Principle

The uncertainty relation between the phase and amplitude quadrature operators can now be determined. Heisenberg's uncertainty principle for a pair of non-commuting observables is given by

$$\Delta(\hat{G})\Delta(\hat{H}) \geq \frac{1}{2i}\langle[\hat{G}, \hat{H}]\rangle, \tag{3.4}$$

where $\Delta(\hat{G})$ and $\Delta(\hat{H})$ are the standard deviation in the observables $\hat{G}$ and $\hat{H}$ and $\langle[\hat{G}, \hat{H}]\rangle$ is the expectation value of the commutation relation of the Hermitian operators $\hat{G}$ and $\hat{H}$. The commutation relation between the two operators, $\hat{G}$ and $\hat{H}$, is defined by

$$[\hat{G}, \hat{H}] = \hat{G}\hat{H} - \hat{H}\hat{G}. \tag{3.5}$$

Using the boson commutation relations, $[a, a^\dagger] = 1$ and $[a, a] = [a^\dagger, a^\dagger] = 0$ [73], we can now calculate the uncertainty relation for the any pair of orthogonal quadrature operators,

$$\Delta(\hat{X}^\theta)\Delta(\hat{X}^{\theta+\frac{\pi}{2}}) \geq 1. \tag{3.6}$$

Equation 3.6 is integral to the work presented in this thesis. In fact, it is direct manipulation of this equality that describes much of the work presented.

3.2.1 State Variance

It is typically the variance of the observable in question which is measured in the lab and not the standard deviation, as will be discussed in § 3.7.2. We define the variance, Δ^2, of an arbitrary operator $\hat{G}$ as,

$$\begin{aligned} V(\hat{G}) &= \Delta^2(\hat{G}) \\ &= (\Delta(\hat{G}))^2 \\ &= \langle\hat{G}^2\rangle - \langle\hat{G}\rangle^2. \end{aligned} \tag{3.7}$$

Notice that this also defines a relationship between the variances of the operators as set by the Heisenberg uncertainty relationship,

$$\Delta^2(\hat{X}^+)\Delta^2(\hat{X}^-) \geq 1. \tag{3.8}$$

3.2.2 Minimum Uncertainty States

A minimum uncertainty state is defined as any state where Equation 3.8 is a strict equality. A minimum uncertainty state can thus be written

$$\Delta^2(\hat{X}^+)\Delta^2(\hat{X}^-) = 1. \tag{3.9}$$

A minimum uncertainty state is one in which no classical noise is present. It is not uncommon to produce these states under certain conditions. For example, solid-state lasers are often in this minimum uncertainty state at frequencies where the low-frequency technical noise has rolled off. This typically occurs above a few megahertz. For example, a typical Nd:YAG system might be in a minimum uncertainty state at 20 MHz or lower [77]. Another way of saying this is that above these frequencies, the performance of many lasers is governed by Heisenberg's uncertainty principle.

3.2.3 Squeezing Purity

Imperfect processes, such as loss, will result in an output state that is not a minimum uncertainty state. We can define the squeezing purity, henceforth simply known as the purity, of a state, P as a means of quantifying this loss (the effect of loss is shown in §3.6. This value is not to be confused with the traditional quantum state purity [78], which is not used in this thesis. The purity is written as

$$\Delta^2(\hat{X}^+)\Delta^2(\hat{X}^-) = P. \tag{3.10}$$

A pure state will have a purity of $P = 1$, as does the vacuum. The value P, as described here, increases as loss is introduced but will eventually begin to decrease as the state approaches the vacuum state with very large losses [76]. This value is particularly important for squeezed states, as it is a measure of the efficiencies of the processes used to produce and measure the state. Excess noise in particular systems reduces Einstein-Podolsky-Rosen (EPR) entanglement [79] and as such is a limitation to many entanglement experiments [80, 81].

3.2.4 The Number Operator

The number operator, $\hat{n}$, is defined by

$$\hat{n} = \hat{a}^\dagger \hat{a}. \tag{3.11}$$

When operated upon a mode, $\langle\hat{\alpha}|\hat{n}|\hat{\alpha}\rangle$, returns the average number of particles in the mode. This operator is particularly useful when describing number, or Fock states. A Fock state is a state of light where the number of photons in the state is well known. It is particularly relevant to states with very small numbers of photons, such as the single photon state. The number operator can be used when determining the power in a state because the power is proportional to the number of photons.

3.3 Linearisation

Linearisation of the creation and annihilation operators can be used, under appropriate circumstances, to simplify calculations. The first order linearisation of the annihilation operator is written

$$\hat{a}(t) = \alpha + \delta\hat{a}(t), \tag{3.12}$$

where α is the complex classical coherent amplitude and $\delta\hat{a}(t)$ is the first order fluctuating term. The same process can be repeated for the creation operator. Linearisation requires that the average of the fluctuations in the field is zero and that the coherent amplitude of the state of light is much larger than the first order fluctuation term. These requirements are written

$$\langle \delta\hat{a}(t) \rangle = 0 \tag{3.13}$$

$$| \delta\hat{a}(t) | \ll | \alpha | . \tag{3.14}$$

It is also useful to linearise the number operator

$$\begin{aligned} \hat{n} &= \hat{a}^\dagger \hat{a} \\ &= (\alpha^* + \delta\hat{a}^\dagger)(\alpha + \delta\hat{a}) \\ &= \alpha^2 + \alpha(\delta\hat{a} + \delta\hat{a}^\dagger) + \delta\hat{a}^\dagger\delta\hat{a} \\ &\approx \alpha^2 + \alpha(\delta\hat{X}^+) \end{aligned} \tag{3.15}$$

where the second order fluctuation term has been removed.

3.4 Representing the States of Light

In this section the *ball-on-stick* and *sideband* representations of the states of light are introduced. These tools provide an intuitive way of visualising the various states of light.

3.4.1 Ball-on-Stick Representation

The ball-on-stick figure is the phase space diagram of the states of light. At a glance, the coherent amplitude, α, the standard deviation of the quadratures on the axes, $\Delta(\hat{X}^\theta)$ and $\Delta(\hat{X}^{\theta+\pi/2})$, and the phase of the state, ϕ, can be identified. Figure 3.1 shows the ball-on-stick representation for a coherent state, the details of which will be introduced in § 3.5.2. The ball-on-stick representation can be regarded as an ensemble of measurements of the amplitude and phase of the state. Each measurement of the state is plotted against the two quadratures and eventually, as many measurements are made, leads to the ball-on-stick diagram. The noise in the state will produce the ball and the stick is drawn from the origin to the average of these measurements. The length of the stick then defines the coherent amplitude and the angle of the stick, the phase of the state.

3.4.2 Sideband Representation

The sideband represenation is complementary to the ball-on-stick representation. Any modulation on some carrier field at some frequency, Δ, can be represented as sidebands around the carrier field at a frequency separation of Δ from the carrier. In this representation we move into the reference frame of the carrier. This means that when the sideband

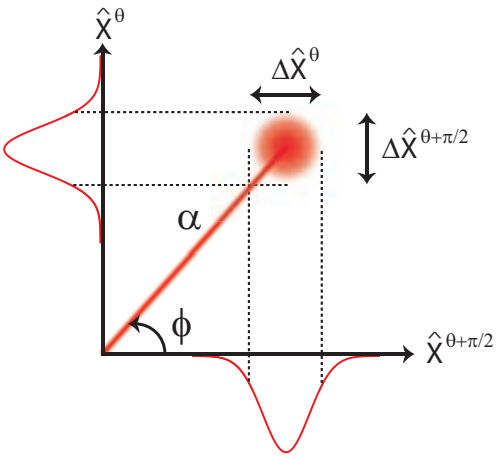


Figure 3.1: The ball-on-stick representation of a coherent state with coherent amplitude, α, and phase, ϕ. The projection of the uncertainty ball onto the two axes indicates the standard deviations, $\Delta(\hat{X}^{\theta,\theta+\pi/2})$ of the two quadrature operators.

representation of a state is drawn, this figure is of a snapshot in time. If one were to draw the representation at some later time, the figure will have evolved. Sidebands above the carrier frequency rotate in one direction and sideband fields below the carrier rotate in the opposite direction. These sidebands add together to modulate the carrier field and as such it is the relative alignment of the sidebands at frequencies Δ and $-\Delta$ that determine the dynamics of the carrier at the frequency Δ. Figure 3.2 shows the sideband representation for a coherent state.

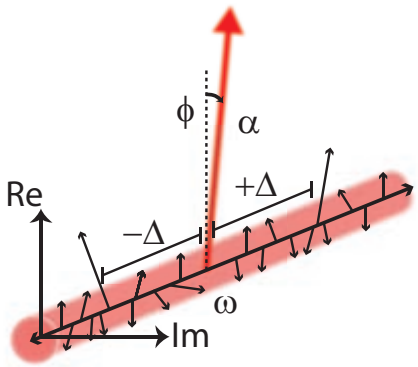


Figure 3.2: The sideband representation of a coherent state with coherent amplitude, α, and phase, ϕ. Quantum noise in the state is represented by a cylinder of quantum sidebands spanning the entire spectrum. The sideband pair at $\pm\Delta$, which are above the quantum noise level, define the motion of the carrier at the frequency Δ.

The quantum sidebands around the carrier frequency are equal in magnitude at all frequencies and beat with the carrier field. This causes the length and phase of the carrier field to fluctuate. If one were to take a slice of the sideband representation at a single

frequency, the ball-on-stick diagram could be constructed from this information.

3.5 Introducing the States of Light

Now that the framework for describing and visualising various states of light has been constructed, we can introduce the relevant states of light; the coherent state, the vacuum state, the squeezed state and the squeezed vacuum state.

3.5.1 Vacuum State

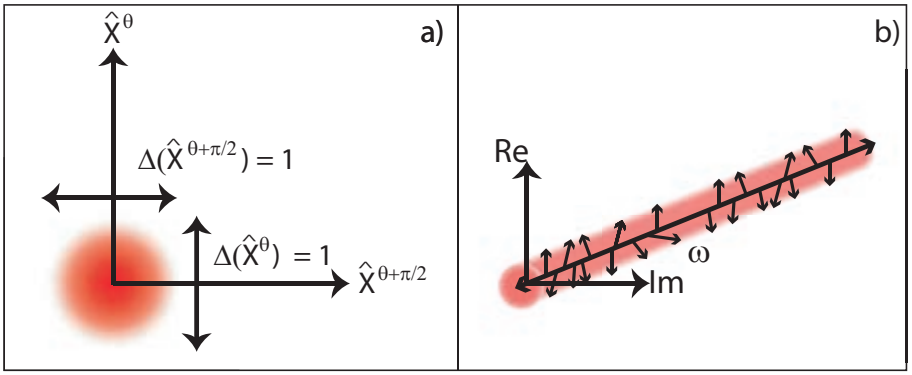


Figure 3.3: The ball-on-stick a) and sideband b) representations of the vacuum state of light.

The vacuum state has no coherent amplitude, determined by operating on this state with the number operator,

$$\langle 0 \mid \hat{n} \mid 0 \rangle = 0. \tag{3.16}$$

Even though there is no coherent amplitude, the vacuum state is not free from fluctuations. It is a minimum uncertainty state with equal noise in all quadratures,

$$\Delta(\hat{X}^{+}) = \Delta(\hat{X}^{-}) = 1. \tag{3.17}$$

These fluctuations are a consequence of the non-zero ground state energy of the quantum harmonic oscillator. It is commonly referred to as zero-point energy or vacuum fluctuations and can also be regarded as virtual particles. These fluctuations occur in every possible optical mode and this has important ramifications when considering optical loss, on squeezed states in particular.

3.5.2 Coherent State

The coherent state, $|\alpha\rangle$, can be produced through application of the displacement operator, $\hat{D}$, on the vacuum state

$$\hat{D}(\alpha)|0\rangle = |\alpha\rangle, \tag{3.18}$$

where the displacement operator is defined by

$$\hat{D}(\alpha) = e^{(\alpha\hat{a}^\dagger - \alpha^*\hat{a})}. \tag{3.19}$$

The representations of the coherent state are shown in Figure 3.4.

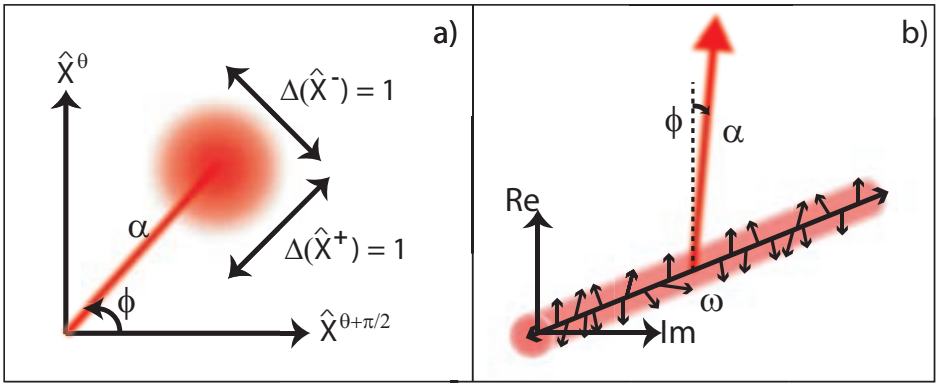


Figure 3.4: The ball-on-stick a) and sideband b) representations of a coherent state of light with phase θ and coherent amplitude α.

The coherent state is a minimum uncertainty state, and has the same noise properties as the vacuum state,

$$\Delta(\hat{X}^+) = \Delta(\hat{X}^-) = 1. \tag{3.20}$$

The coherent state is also defined as having a coherent amplitude that is greater than zero. An ideal laser source would produce a coherent state at all measurement frequencies. However, technical limitations in the low-frequency regime tend to add additional classical noise to the state. As mentioned in § 3.9, the lasers used in this work tend to be shot noise limited above a couple of megahertz. This is analogous to saying that these lasers produce coherent states above these frequencies.

3.5.3 Squeezed State

Squeezed states can be generated through the use of the squeeze operator, $\hat{S}(\rho)$, and the displacement operator,

$$\hat{D}(\alpha)\hat{S}(\rho)|0\rangle = |\alpha, \rho\rangle. \tag{3.21}$$

The squeezing operator is defined by

$$\hat{S}(\rho) = e^{\frac{1}{2}(\rho^*\hat{a}\hat{a} - \rho\hat{a}^\dagger\hat{a}^\dagger)} \tag{3.22}$$

$$\rho = re^{i\varphi}, \tag{3.23}$$

where $r \geq 0$ is known as the squeezing parameter and ψ is the angle of the squeezing.

The squeezing operator reduces the noise in some quadrature by a factor or e^{-r} and

increases it by a factor of e^r in the conjugate observable. The noise in the conjugate quadrature is necessarily increased such that the Heisenberg uncertainty relation between the two observables is not violated. A squeezed state of light is shown in Figure 3.5.

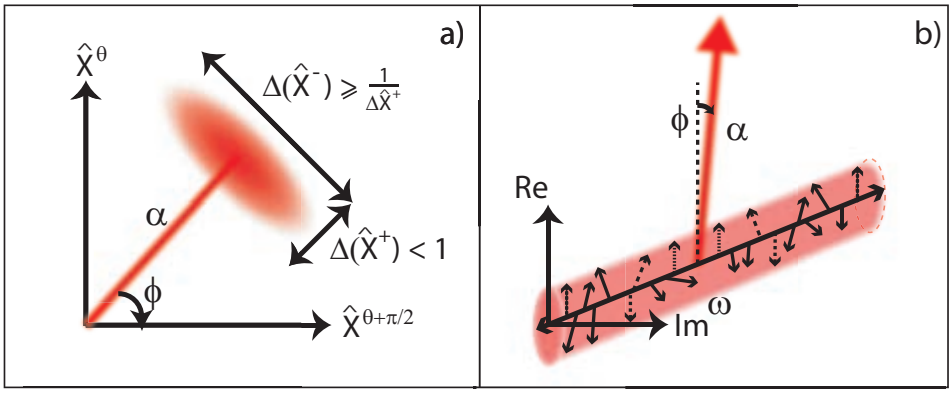


Figure 3.5: The ball-on-stick a) and sideband b) representations of a squeezed state of light with phase θ and coherent amplitude α. Correlated sideband pairs are represented by matching dashed and/or dotted lines in b)

Note that squeezed states need not be minimum uncertainty states. In fact, any real state will have some loss from any number of sources, such as scattering or absorption. This loss will reduce the purity of these states, as discussed in § 3.6.

For an amplitude squeezed state, application of the number operator will yield

$$\langle \alpha, \rho \mid \hat{a}^{\dagger}\hat{a} \mid \alpha, \rho \rangle = \sinh^2(r) + |\alpha|^2. \tag{3.24}$$

The variance of the quadrature is found to be

$$\Delta^2(\hat{X}^{+}) = e^{-2r} \tag{3.25}$$

$$\Delta^2(\hat{X}^{-}) = e^{2r}. \tag{3.26}$$

Inspection of the squeezing operator, Equation 3.22, shows that this operator acts to create or destroy photons in pairs. In fact it is these pairs of photons that are the source of squeezing.

Squeezed Vacuum States

Squeezed vacuum states are simply squeezed states with no coherent amplitude. However, Equation 3.24 shows that the squeezing operation produces photons. The photons produced are correlated photon pairs. The angle of the squeezing is determined by how these photon pairs align in the sideband representation. Whilst squeezed vacuum states have no coherent amplitude, they are still sensitive to loss because the removal of these correlated sidebands leaves more noisy vacuum fluctuations in its place. The various representations of the squeezed vacuum state are shown in Figure 3.6.

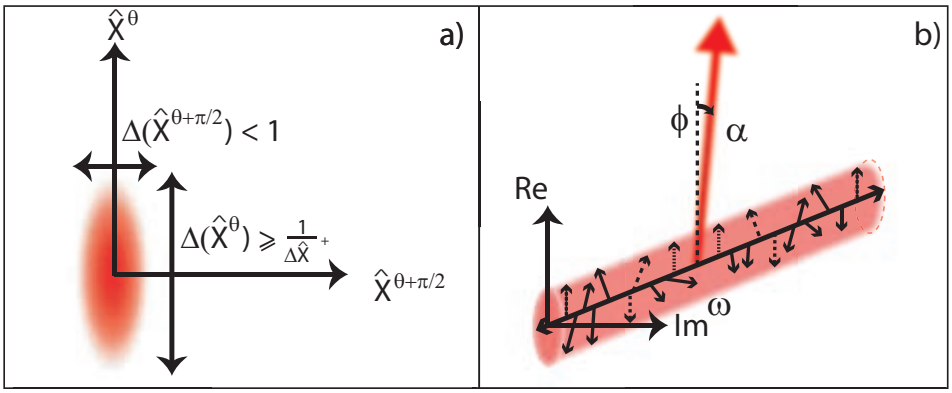


Figure 3.6: The ball-on-stick a) and sideband b) representations of a squeezed vacuum state of light. Correlated sideband pairs are represented by matching dashed and/or dotted lines.

3.6 Squeezed State Loss

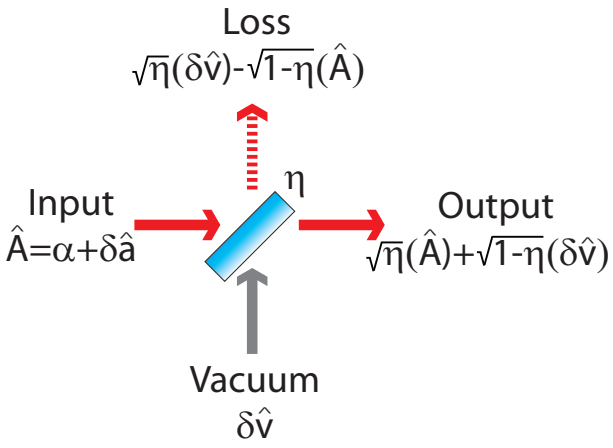


Figure 3.7: Modelling loss on an optical state. The power transmissivity of the beamsplitter, η, can be regarded as the efficiency of propagation of the field.

Optical loss can occur from many different processes such as absorption, scattering and transmission. Due to the presence of vacuum fields, the effect of loss on quantum states can have a large impact. To understand loss, we model the process as a beamsplitter with a power transmissivity of η, as shown in Figure 3.7. In this way, η can also be considered as a measure of the efficiency of some process. An η value of 1 corresponds to no state loss and the noise properties of the output of the process is equal to the input. The beamsplitter can be regarded as having two input ports and two output ports. One of the input ports has a bright field, $\mathrm{A} = \alpha + \delta a$, and the second input port is the vacuum field, $\delta\hat{v}$. The output field is then the field that has passed through the beamsplitter and has been attenuated by a factor of $\sqrt{\eta}$. The remaining field, labelled loss, is considered to be

removed from the system and is unmeasurable, thus representing loss. The output field, $\hat{A}_{out}$ is written

$$\hat{A}_{out} = \sqrt{\eta}\hat{A} + \sqrt{1-\eta}\delta\hat{v} \tag{3.27}$$

If η is not unity then we couple noise from the vacuum field, δv, into the output via beamsplitter power transmissitivity, η. In many situations this vacuum field is of little consequence because the noise of the vacuum field is either identical to, or much less than, the noise on the original input field. However, if the input field is squeezed then, by definition, the noise on the vacuum field is greater. The squeezing measurable at the output will now be less than the amount of squeezing that was measurable at the input due to the added noise from the vacuum field. This effect is extremely important as the loss of a system will set a hard upper bound on the improvement that squeezing can offer.

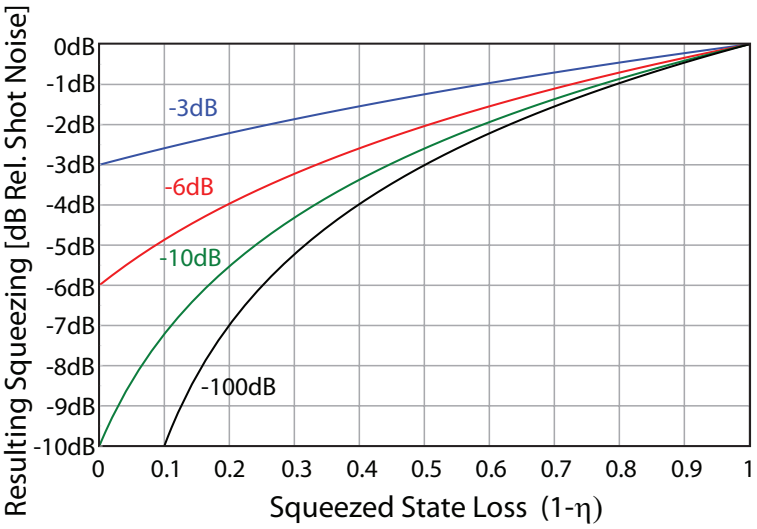


Figure 3.8: The effect of loss on various initial squeezed states.

It is important to have an intuition of the effect of loss on squeezed states. Figure 3.8 shows the resulting squeezing from four different initial squeezed states after having undergone various losses, $(1-\eta)$. The four initial squeezed states have 3, 6, 10, and 100 dB suppression of the quantum noise when compared to a field with equal intensity. Along the y-axis it can be seen that, as expected, a value of $\eta = 1$, or no loss, results in squeezing equal to that of the initial state. It can also be seen that with total loss of the initial squeezed state, $\eta = 1$, all traces converge to 0 dB, or the vacuum shot noise set by Heisenberg's uncertainty principle. The 100 dB state, although currently impossible to produce, illustrates the point that even with infinite squeezing, for a given loss there is a maximum amount of squeezing that can still be measured. With a 50% loss, we can only ever hope to measure 3 dB of squeezing, even if our original input state is infinitely squeezed. At higher losses, the dependence on the initial amount of squeezing also tends to be less important. With a 70% loss, a 6 dB initial state will result in a squeezing level of just over 1.1 dB, whereas a 100 dB initial state will result in approximately 1.5 dB.

Another way of looking at this is to consider what happens to a pure squeezed state after undergoing various levels of loss. The ball-on-stick representation of this situation

is illustrated in Figure 3.9. The resulting states after a pure, 10 dB squeezed state has undergone power losses of 25, 50, 75 and 100 percent are shown. Figure 3.9 shows that loss results in degradation of the squeezing, and also gives a feeling for the rate at which the squeezing and anti-squeezing degrades.

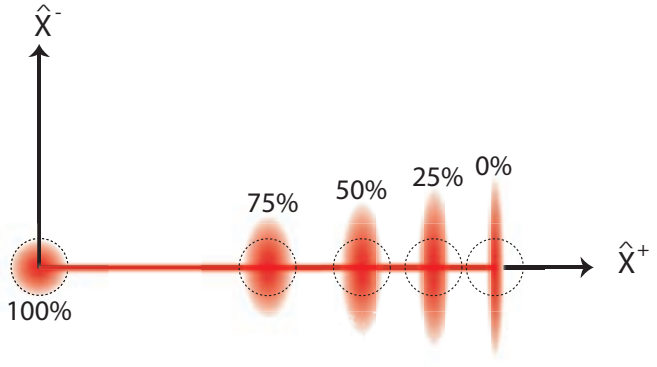


Figure 3.9: The effect of power losses of 25, 50, 75 and 100 percent on an initial pure 10 dB squeezed state are shown.

3.7 State Measurement

States of light are typically measured using a photodetector, consisting of a photodiode and a trans-impedance gain stage, and the resulting signal is then observed on either a spectrum analyser or an oscilloscope. The photodetector is shown in Figure 3.10.

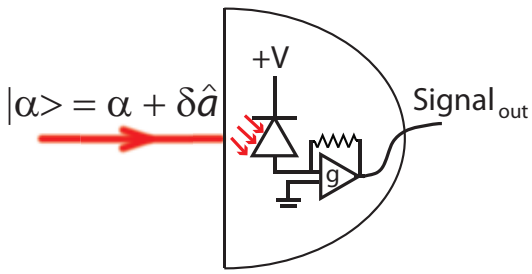


Figure 3.10: Detection of a an optical state, $|\alpha\rangle$, with a photodetector. This simple photodetector consists of a photodiode and a trans-impedance gain stage.

Ideally, the photodiode provides a one-to-one conversion of photons into electrons. Due to this, the photocurrent, i, immediately after the photodiode follows the intensity of the input field. The gain stage then converts this current signal into a voltage signal, with an amplification factor of g Ohms, resulting in

$$\begin{aligned} \text{Signal}_{out} &= g\text{i} \\ &= g(\alpha + \delta a)(\alpha^* + \delta a^\dagger) \\ &\approx g(\alpha^2 + \alpha(\delta \hat{X}^+(t))). \end{aligned} \tag{3.28}$$

3.7.1 Oscilloscope Measurements

Oscilloscopes are a means of measuring voltages in the time domain. Equation 3.28 is a voltage signal and hence oscilloscopes directly measure this value. The signal consists of a DC term proportional to the intensity, α^2, and a scaled fluctuation term, $\alpha(\delta\hat{X}^+)$.

3.7.2 Spectrum Analysers

In contrast to the oscilloscope, a spectrum analyser provides measurements in the frequency domain. It measures the variance of the input voltage signal for the frequency band in question. This variance in the voltage is proportional to the intensity spectrum of the incident laser field. For the output of our photodetector, given by Equation 3.28, we find

$$\begin{aligned} V(\text{Signal}_{\text{out}}) &\approx g(\alpha^2 + \alpha(\delta\hat{X}^+(t)) \\ &= g^2\alpha^2 V(X^+(t)), \end{aligned} \tag{3.29}$$

which is found making use of the identity,

$$V(aX + b) = a^2 V(X). \tag{3.30}$$

We can then transform this into the frequency domain to determine the signal measured with the spectrum analyser.

$$V(\text{Signal}_{\text{out}}(\omega)) = g^2\alpha^2 V(\check{X}^+(\omega)), \tag{3.31}$$

where the caron, or inverted hat, is used to signify operators in the frequency domain. The final step is to then determine the variance of the amplitude quadrature operator in the frequency domain. Using linearisation we can write

$$\begin{aligned} V(\check{X}^+(\omega)) &\equiv \Delta^2(\check{X}^+) \\ &= \langle(\check{X}^+)^2\rangle - \langle\check{X}^+\rangle^2 \\ &= \langle(\check{a} + \check{a}^\dagger)^2\rangle - \langle\check{a} + \check{a}^\dagger\rangle^2 \\ &= \langle(\alpha + \delta\check{a} + \alpha^* + \delta\check{a}^\dagger)^2\langle - \langle\alpha + \delta\check{a} + \alpha^* + \delta\check{a}^\dagger\rangle^2 \\ &= \langle(\delta\check{a} + \delta\check{a}^\dagger)^2\rangle \\ &= \langle(\delta\check{X}^+)^2\rangle. \end{aligned} \tag{3.32}$$

The variance for the phase quadrature operator is found in much the same way

$$\begin{aligned} V(\check{X}^-(\omega)) &\equiv \Delta^2(\check{X}^-) \\ &= \langle(\check{X}^-)^2\rangle - \langle\check{X}^-\rangle^2 \\ &= \langle(i(\delta\check{a}^\dagger - \delta\check{a}))^2\rangle \\ &= \langle(\delta\check{X}^-)^2\rangle. \end{aligned} \tag{3.33}$$

3.8 Balanced Homodyne Detection

Loss is one limitation to the measurement of many of these quantum states of light but the measurement itself is also a complicated issue. Measuring the quantum noise in these

states requires that all classical noise terms, such as laser intensity noise and electronic noise, are reduced to a level well below the noise level of interest (This is discussed in detail in § 8.3.1). In the case of squeezed light, where the noise in some quadrature may have been reduced by even more than one order of magnitude below the shot noise limit, achieving the required suppression of classical noise is non-trivial. A measurement system that allows for the detection of any general quadrature, including the amplitude and phase quadratures, is also necessary in order to fully characterise the measured states. All of these conditions can be met using the balanced homodyne detector [82], illustrated in Figure 3.11.

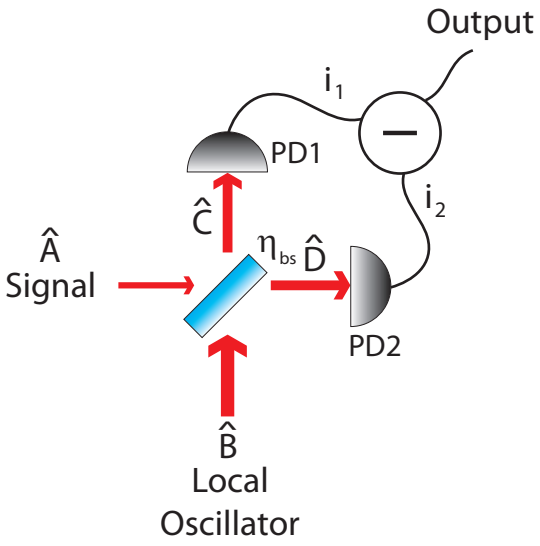


Figure 3.11: The idealised balanced homodyne detector. The signal and the local oscillator (LO) fields are both incident on the beamsplitter, of power transmissivity $\eta_{bs} = 0.5$, and the two output fields are detected on separate photodetectors, PD1 and PD2.

In the balanced homodyne detector the weak *signal field*, whose noise properties one wishes to measure, and a bright *local oscillator* (LO) field, used to amplify and probe the signal field, interfere on a beamsplitter with a power transmissivity of $\eta_{\rm bs}$, which is ideally 50 %. Using linearisation (§ 3.3) the signal field is labelled as $\hat{A} = \alpha_a + \delta\hat{a}$ and the local oscillator field as $\hat{B} = \alpha_b + \delta\hat{b}$, where the subscripts a and b on the coherent amplitudes, α, are used to distinguish between amplitudes of the signal and local oscillator fields respectively.

The fields incident on the two photo-detectors, $\hat{C}$ and $\hat{D}$, are written

$$\hat{C} = \sqrt{1-\eta_{bs}}\hat{A} + \sqrt{\eta_{bs}}\hat{B} \tag{3.34}$$
$$\hat{D} = \sqrt{\eta_{bs}}\hat{A} - \sqrt{1-\eta_{bs}}\hat{B}. \tag{3.35}$$

Here we will ensure that the coherent amplitude terms of the fields are real by separating the relative phase between the two fields into a phase term, $\hat{B}$, where $\hat{B} = \hat{B}e^{i\theta}$. We can now write

$$\hat{C} = \sqrt{1-\eta_{bs}}\hat{A} + \sqrt{\eta_{bs}}\hat{B}e^{i\theta} \tag{3.36}$$
$$\hat{D} = \sqrt{\eta_{bs}}\hat{A} - \sqrt{1-\eta_{bs}}\hat{B}e^{i\theta}. \tag{3.37}$$

Assuming that the photodetectors are perfect, every photon hitting the detectors will result in an electron at the output of the detector. The photocurrent produced by PD1 is equal to $\hat{C}^\dagger\hat{C}$ and the photocurrent from PD2 is equal to $\hat{D}^\dagger\hat{D}$. The two photocurrents, i_1 and i_2 can then be written,

$$\begin{aligned}\mathrm{i}_1 &= \hat{C}^\dagger\hat{C} \\ &= (1-\eta_{bs})\hat{A}^\dagger\hat{A} + \eta_{bs}\hat{B}^\dagger\hat{B} + \sqrt{\eta_{bs}(1-\eta_{bs})}(\hat{A}^\dagger\hat{B}e^{i\theta} + \hat{B}^\dagger\hat{A}e^{-i\theta}) \qquad (3.38)\\ \mathrm{i}_2 &= \hat{D}^\dagger\hat{D} \\ &= \eta_{bs}\hat{A}^\dagger\hat{A} + (1-\eta_{bs})\hat{B}^\dagger\hat{B} - \sqrt{\eta_{bs}(1-\eta_{bs})}(\hat{A}^\dagger\hat{B}e^{i\theta} + \hat{B}^\dagger\hat{A}e^{-i\theta}). \qquad (3.39)\end{aligned}$$

As shown in Figure 3.11, the output of the homodyne detector is then the resulting signal when these two photocurrents are subtracted. The result is

$$\mathrm{i}_1 - \mathrm{i}_2 = (1-2\eta_{bs})\hat{A}^\dagger\hat{A} + (2\eta_{bs}-1)\hat{B}^\dagger\hat{B} + 2\sqrt{\eta_{bs}(1-\eta_{bs})}(\hat{A}^\dagger\hat{B}e^{i\theta} + \hat{B}^\dagger\hat{A}e^{-i\theta}). \quad (3.40)$$

Utilising the linearisation technique described in § 3.3 allows for simplification of some of the terms in Equation 3.40. We begin by recalling that the problem was constructed such that the coherent fields are real, and we choose to drop terms of second order. Using these simplifications we can write

$$\begin{aligned}\hat{A}^\dagger\hat{A} &\approx (\alpha_a^* + \delta\hat{a}^\dagger)(\alpha_a + \delta\hat{a}) \\ &\approx \alpha_a^2 + \alpha_b\delta X_A^+ \qquad (3.41)\\ \hat{B}^\dagger\hat{B} &\approx (\alpha_b^* + \delta\hat{b}^\dagger)(\alpha_b + \delta\hat{b}) \\ &\approx \alpha_b^2 + \alpha_b\delta X_B^+. \qquad (3.42)\end{aligned}$$

where $\delta X_A^+ = \delta\hat{a} + \delta\hat{a}^\dagger$ is the amplitude quadrature operator for the field $\hat{A}$, and $\delta X_B^+ = \delta\hat{b} + \delta\hat{b}^\dagger$ is the amplitude quadrature operator for the field $\hat{B}$. Making use of Equations 3.41 and 3.42 we can now simplify Equation 3.40, resulting in

$$\begin{aligned}\mathrm{i}_1 - \mathrm{i}_2 \approx\ & (1-2\eta_{bs})(\alpha_a^2 + \alpha_a\delta X_A^+) + (2\eta_{bs}-1)(\alpha_b^2 + \alpha_b\delta X_B^+) + \\ & 2\sqrt{\eta_{bs}(1-\eta_{bs})}[(\alpha_a^* + \delta\hat{a}^\dagger)(\alpha_b + \delta\hat{b})e^{i\theta} + (\alpha_b^* + \delta\hat{b}^\dagger)(\alpha_a + \delta\hat{a})e^{-i\theta}].\end{aligned} \quad (3.43)$$

This can be rewritten (using various trigonometric identities on the exponential terms) as

$$\begin{aligned}\mathrm{i}_1 - \mathrm{i}_2 \approx\ & (1-2\eta_{bs})(\alpha_a^2 + \alpha_a\delta X_A^+) + (2\eta_{bs}-1)(\alpha_b^2 + \alpha_b\delta X_B^+) + 2\sqrt{\eta_{bs}(1-\eta_{bs})} \times \\ & \left[2\alpha_a\alpha_b\mathrm{Cos}\theta + \alpha_a\left(\delta\hat{X}_B^+\mathrm{Cos}\theta - \delta\hat{X}_B^-\mathrm{Sin}\theta\right) + \alpha_b\left(\delta\hat{X}_A^+\mathrm{Cos}\theta + \delta\hat{X}_A^-\mathrm{Sin}\theta\right)\right].\end{aligned} \quad (3.44)$$

It is typically required that the local oscillator field strength, α_b, is much brighter than the signal field in order to suppress the noise from the local oscillator. For most of the work presented in this thesis the signal field is a vacuum squeezed state, with no coherent amplitude, and hence this condition is easily met. This condition allows us to remove all

terms that do not contain the coherent amplitude of the LO field, α_b, resulting in

$$\begin{aligned} \mathrm{i}_1 - \mathrm{i}_2 &\approx (2\eta_{bs} - 1)(\alpha_b^2 + \alpha_b \delta X_B^+) + \\ &\quad 2\sqrt{\eta_{bs}(1-\eta_{bs})}\left[2\alpha_a\alpha_b \mathrm{Cos}\theta + \alpha_b\left(\delta\hat{X}_A^+ \mathrm{Cos}\theta + \delta\hat{X}_A^- \mathrm{Sin}\theta\right)\right]. \end{aligned} \tag{3.45}$$

If we now assume that our beamsplitter ratio is exactly 50 % we completely subtract the contributions from the local oscillator field and Equation 3.45 simplifies to

$$\mathrm{i}_1 - \mathrm{i}_2 \approx 2\alpha_a\alpha_b \mathrm{Cos}\theta + \alpha_b\left(\delta\hat{X}_A^+ \mathrm{Cos}\theta + \delta\hat{X}_A^- \mathrm{Sin}\theta\right). \tag{3.46}$$

It is worth taking a moment and considering the ramifications of this equation. The first term, $2\alpha_a\alpha_b \mathrm{Cos}\theta$, is a DC term that describes the interference between the two fields as the phase of the local oscillator, θ, is varied. If the two fields are in exactly the same spatial mode then this term will completely vanish, given an appropriate value for θ. In practice, however, the two fields are never perfectly matched in shape and size and hence this term will not completely vanish. The overlap between the two fields can be quantified using the *visibility* of the balanced homodyne detector, which will be introduced in § 3.8.1.

The remaining terms are the quantum noise terms. The quantum noise terms have been "amplified" by the coherent amplitude term of the LO field, which we see occurs without adding any noise from the LO field in the case of perfect subtraction. This gain factor allows us to measure vanishingly small quantum noise terms on the signal field. It is also evident that as the phase of the LO field is varied, the noise seen at the output is some mixture of amplitude and phase quadratures. The balanced homodyne detector, by varying the phase of the LO, works as a measuring device for any general quadrature operator that we wish to measure, allowing us to fully characterise the noise characteristics of any state.

3.8.1 Balanced Homodyne Detector Fringe Visibility

The visibility of the homodyne detection system is a measure of the overlap between the signal and local oscillator fields. In order to achieve complete destructive or constructive interference the waist size and waist position of the two fields must be exactly the same. If this is not the case then the wavefronts will not match and the same phase relationship will not be kept over the entire beam. The visibility is a way of quantifying the spatial mode mismatch between the signal and LO fields and is given by

$$\mathcal{V} = \frac{I_{Max} - I_{Min}}{I_{Max} + I_{Min}}. \tag{3.47}$$

I_{Max} is the maximum intensity seen as the phase is varied, which corresponds to constructive interference, and I_{Min} is the minimum intensity measured, corresponding to destructive interference. The values are determined by ensuring that the power of the local oscillator and signal fields are first equal. The phase of the local oscillator is then swept whilst examining the output DC term on an oscilloscope.

It can be shown that the fringe visibility is equivalent to optical loss on the signal field [83]. If the equivalent loss due to an imperfect fringe visibility is given by η_{vis}, it can be defined as

$$\eta_{vis} = \mathcal{V}^2. \tag{3.48}$$

It is perhaps not surprising that an inefficient spatial overlap between these two fields is equivalent to a loss on the signal field. Equation 3.46 shows that the noise terms from the signal field are amplified by the coherent amplitude of the LO field. If some part of the signal field does not overlap then it will not receive this gain and will be unmeasurable, thus acting as a loss source.

Chapter 4

Generation of Squeezed States

This chapter begins by discussing the various methods for producing squeezed states and from there looks at a brief history of squeezing results from cavity systems utilising the second order non-linearity from various materials. The second order process is then described in detail. First, the photon interaction in second harmonic generation is detailed, followed by the optical parametric amplifier and optical parametric oscillator systems. From there, some of the technical issues present in these systems are discussed such as phase matching and intra-cavity dispersion.

4.1 Methods of Squeezed State Production

The first observation of squeezed states was achieved by Slusher *et al.* using a four wave mixing process involving Na atoms in a gas cell [57]. Since then, optical squeezing has been produced in many systems, such as vapour cells [84] and fibres [85]. For many years there was no system that clearly outperformed the others [58], however, one of the stronger performers over the years has been the below threshold optical parametric oscillator (OPO) [58]. Additionally, the parameters of the optical cavity can be tuned for the specific application, providing flexibility. As opposed to three wave mixing processes such as those present in fibre or vapour squeezing, the strength of the second order non-linear interaction, $\chi^{(2)}$, utilised in the OPO is very strong. These properties result in the OPO being a versatile, reliable source of squeezing and currently holds the record for squeezing magnitude, of 12.7 dB [86], and low-frequency performance, 10 dB at 10 Hz [87]. These low-frequency results are presented in Chapter 8. However, squeezing results generated with four-wave mixing experiments are also progressing [88, 89].

4.1.1 History of Optical Parametric Oscillator and Amplifier Performance

A short list of relevant landmark squeezing results from optical cavity systems and some of the details of the systems used to produce these results are shown in Table 8.14.

Table 8.14 shows that the highest level of squeezing measured to date is 12.7 dB at a measurement frequency of 5 MHz [86]. This was achieved in a singly resonant monolithic Periodically Poled Potassium Titanyl Phosphate (PPKTP) cavity. The largest magnitude of squeezing at the lowest frequency is 10 dB at 10 Hz and was achieved in 2012 in a doubly resonant bow-tie cavity also using PPKTP. This work is presented in § 8.4.5. The same work, presented in this thesis showed squeezing of 11.6 dB down to 200 Hz. Two of the most influential results were the first to observe squeezing at audio frequencies, and the first to achieve a suppression of the shot noise by an order of magnitude. The first

Year	Group	Mag.	Freq.	Cavity	λ [nm]	Material
1986	Wu *et al.* [90]	3 dB	1.8 MHz	Linear	532 & 1064	MgO:LN
1992	Polzik *et al.* [91]	6.0 dB	1.4 MHz	Bow-Tie	856	KNB
1992	Ou *et al.* [92]	3.6 dB	1.1 MHz	Bow-Tie	540 & 1080	KTP
1996	Schneider *et al.* [93]	4.3 dB	5 MHz	Hemilithic	1064	MgO:LN
1998	Schneider *et al.* [94]	6.5 dB	6.5 MHz	Hemilithic	1064	MgO:LN
1999	Lam *et al.* [95]	7.1 dB	3 MHz	Monolithic	1064	MgO:LN
2002	Bowen *et al.* [62]	2.5 dB	220 kHz	Hemilithic	1064	MgO:LN
2004	McKenzie *et al.* [63]	2.0 dB	500 Hz	Hemilithic	1064	MgO:LN
2006	Suzuki *et al.* [96]	7 dB	1 MHz	Bow-Tie	860	PPKTP
2007	Takeno *et al.* [97]	9 dB	1 MHz	Bow-Tie	860	PPKTP
2007	Goda *et al.* [60]	7.4 dB	2 kHz	Linear	1064	PPKTP
2007	Vahlbruch *et al.* [98]	10 dB	5 MHz	Monolithic	1064	MgO:LN
2010	Mehmet *et al.* [99]	11.5 dB	5 MHz	Monolithic	1064	MgO:LN
2010	Eberle *et al.* [86]	12.7 dB	5 MHz	Monolithic	1064	PPKTP
2011	Khalaidovski *et al.* [100]	9.5 dB	3 kHz	Hemilithic	1064	PPKTP
2011	Mehmet *et al.* [101]	12 dB	60 kHz	Hemilithic	1550	PPKTP
2011	Chua *et al.* [102]	8.6 dB	10 Hz	Bow-Tie	1064 & 532	PPKTP
2012	Stefszky *et al.* [87]	10 dB	10 Hz	Bow-Tie	1064 & 532	PPKTP

Table 4.1: A brief list of relevant squeezing results, pioneering squeezing in terms of magnitude, frequency, cavity design and/or wavelength. The magnitude of the squeezing given (Mag.) is the directly observed squeezing level, with no corrections for imperfect detection efficiencies. The λ column indicates the wavelengths resonant in the squeezer, where finesses greater than 5 are considered resonant. The materials used by the various groups are labelled; Magnesium Oxide doped Lithium Niobate (MgO:LN), Potassium Niobate (KNB), Potassium Titanyl Phosphate (KTP) and Periodically Poled Potassium Titanyl Phosphate (PPKTP).

audio frequency squeezing was achieved in 2004 by McKenzie *et al.* in a singly resonant hemilithic cavity using MgO:LN [63]. The first directly observed 10 dB of squeezing below the shot noise level was achieved by Vahlbruch *et al.* in 2007 using a singly resonant monolithic MgO:LN cavity [98].

4.2 The Second Order Non-linearity

The OPO utilises the relatively strong second order non-linear interaction, $\chi^{(2)}$, in order to produce squeezing. We can write the induced polarisation, $\vec{P}$, in a non-linear dielectric medium due to an external electromagnetic wave, $\vec{E}$, as

$$\vec{P} = \varepsilon_0(\chi\vec{E} + \chi^{(2)}\vec{E}^2 + \chi^{(3)}\vec{E}^3 + \dots), \tag{4.1}$$

where the $\chi^{(x)}$ terms are the various orders of the non-linear susceptibility and ε_0 is the electric permittivity of free space. The third order susceptibility, $\chi^{(3)}$, is typically orders of magnitude lower than the second order term. It is responsible for four-wave processes such as the Kerr effect, third harmonic generation and self phase modulation [103]. The second order interaction can be used to drive processes such as second harmonic generation (SHG), optical parametric amplification (OPA) and optical parametric oscillation (OPO).

4.3 Second Harmonic Generation

The simplest of these processes is second harmonic generation, or SHG, and as such is a good introduction to the other non-linear processes. Here, only the photon interactions, and as such the energy conservation requirements, are investigated. Consideration of momentum conservation is left until § 4.5. SHG is an up-conversion process, meaning that photons at a lower frequency, ω, combine to produce photons at a higher frequency, in this case at 2ω. It was first experimentally achieved in 1961 by Franken *et al.* [104]. We can write

$$\omega_{sh} = \omega + \omega, \tag{4.2}$$

where ω_{sh} is the frequency of the second harmonic field and ω is the angular frequency of the fundamental field. This process is shown in Figure 4.1. These systems are often placed inside a resonator such that the fundamental field undergoes many passes through the material, substantially improving the conversion efficiency.

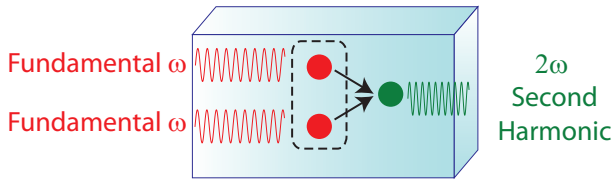


Figure 4.1: The SHG system. Two photons at frequency ω combine to produce one photon at frequency 2ω.

4.4 Downconversion Processes

Downconversion processes are ones in which pump photons are converted into lower energy photons. Although the basic setup is similar, we can identify three different downconversion modes of operation for our system, depending upon how the process is seeded. These are the degenerate optical parametric amplifier, the non-degenerate optical parametric amplifier, and the optical parametric oscillator. These various modes of operation are integral in squeezed state production from a second order system.

4.4.1 The Degenerate Optical Parametric Amplifier

It is often easier to first consider the more intuitive case of the DOPA before moving onto describing the processes in the OPO. The difference between the two, as shall soon be made clearer, is the presence of a bright seed field. The OPA is capable of producing exotic quantum states of light but in this instance we will only investigate the classical behaviour. The quantum behaviour will follow in the treatment of the OPO.

Figure 4.2 depicts the DOPA and the processes involved. The presence of the seed field, also known as the signal field, can be thought of as a catalyst for the reaction of the pump photon downconversion. In the DOPA these two frequencies are equal. We write

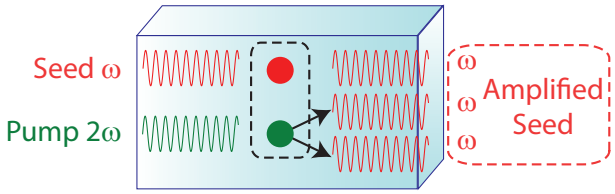


Figure 4.2: The degenerate optical parametric amplifier (DOPA). The system is driven with a pump field at frequency 2ω and a seed field at frequency ω. Inside the non-linear medium, the presence of a seed photon causes the pump photon to split into two photons of frequency ω, thus amplifying the seed field.

for the energy, E, of the process,

$$E \propto \omega_{pump} = \omega + \omega. \tag{4.3}$$

An idealised picture of the process is shown. In reality, only a portion of the pump photons will undergo the downconversion process. The remaining light will exit the system and is known as the residual pump. In many applications, the portion of the pump that is converted is very small due to the relative magnitude between the seed and pump fields, and an approximation is made that the power of the pump field is constant. Conservation of momentum must also be achieved and this is done through a technique known as phase matching, introduced in § 4.5.

Explicitly writing the origin of the two incident fields at angular frequencies of ω and 2ω as ω_{pump} and ω_{seed} respectively, the combined field present on the non-linear dielectric, $\vec{E}_{inc}(t)$, can thus be written

$$\vec{E}_{inc}(t) = \vec{E}_{seed}e^{-i\omega_{seed}t} + \vec{E}_{pump}e^{-i\omega_{pump}t} + \vec{E}^*_{seed}e^{i\omega_{seed}t} + \vec{E}^*_{pump}e^{i\omega_{pump}t}. \tag{4.4}$$

If we look at only the second order term from Equation 4.1 then we can investigate the second order processes involved. Dropping the vector notation on assumption that the two fields are co-propagating we can write

$$\begin{aligned} \chi^{(2)}E^2_{inc}(t) \quad = \quad & \chi^{(2)}[E_{seed}E^*_{seed}e^{-i2\omega_{seed}t} + E_{pump}E^*_{pump}e^{-i2\omega_{pump}t} + \\ & 2E_{seed}E_{pump}e^{-i(\omega_{seed}+\omega_{pump})t} + \\ & 2E_{seed}E^*_{pump}e^{-i(\omega_{pump}-\omega_{seed})t} + \\ & 2(|E_{seed}| + |E_{pump}|)]. \end{aligned} \tag{4.5}$$

In the DOPA, the term we are interested in is the difference frequency generation term

$$2E_{seed}E^*_{pump}e^{-i(\omega_{pump}-\omega_{seed})t}, \tag{4.6}$$

which describes the process of amplifying the field at ω_{seed} as seen in Equation 4.3.

4.4.2 The Non-Degenerate Optical Parametric Amplifier

The NDOPA can be used to create exotic states of light, particularly in cases where one might wish to separate the two output modes, such as separable entangled states [105, 106].

Here we only investigate the classical behaviour of the NDOPA. In this case, the frequency of the seed field is shifted from half of the pump frequency by an amount δ.

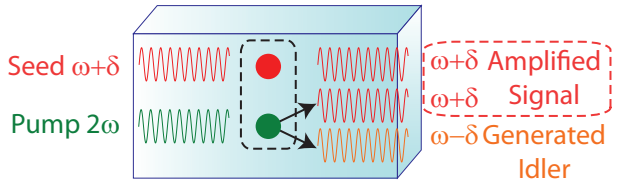


Figure 4.3: The non-degenerate optical parametric amplifier (NDOPA). The system is driven with a pump field at frequency 2ω and a seed field at frequency $\omega+\delta$. Inside the non-linear medium the presence of a seed photon causes the pump photon to split into two photons of frequency $\omega+\delta$ and $\omega-\delta$. This amplifies the seed field but also creates a second field, the idler or generated sideband field.

We see that the defining feature is the generation of the idler field. This field is generated much in the same way as in the DOPA 4.6. The idler field has the same frequency separation as the signal from the carrier, δ, but with a negative sign. This results in a beat signal at the frequency 2δ.

$$E \propto \omega_{pump} = \omega + \delta + \omega - \delta. \tag{4.7}$$

The classical behaviour of the NDOPA is important to the work presented in this thesis because it describes the coherent locking scheme, introduced in § 7.4. This scheme is used to control the vacuum squeezed states produced. In contrast to the DOPA, it is also shown in § 7.4 that the phase of the pump in this system does not affect the output field at the seed frequency, but instead alters the phase of the generated idler field relative to the seed.

4.4.3 The Optical Parametric Oscillator

The optical parametric oscillator can be regarded as an OPA with no bright seed field, or perhaps more informatively, as having a vacuum seed. The OPO is driven by this vacuum seed and by the pump field at 2ω. Vacuum fluctuations can be thought of as the spontaneous production of particle and anti-particle pairs that come into existence and swiftly annihilate. We can imagine that one of these particles seeds the OPO in exactly the same way as the classical seed does, resulting in the downconversion of a pump photon into a photon at the virtual particle frequency, δ, and a photon into the corresponding frequency necessary to conserve energy.

$$E \propto \omega_{pump} = \delta + (2\omega - \delta). \tag{4.8}$$

If we look at the non-linear response of the system then the relevant term is again the difference frequency generation

$$2E_{vacseed}E^*_{pump}e^{-i(\omega_{pump}-\omega_{vacseed})t} \tag{4.9}$$

where the seed field is now replaced with a vacuum seed term, $E_{vacseed}$. It has previously been mentioned that vacuum fluctuations can be regarded as the creation and annihilation of virtual particle pairs. It has also been mentioned that vacuum fluctuations exist at all frequencies and all spatial modes. What this means is that the vacuum seed frequency term, $\omega_{vacseed}$, in Equation 4.9 is constantly changing as a consequence of quantum mechanics. At some point in time the vacuum fluctuation will cause the pump photon to down-convert to a photon pair with some random frequency separation from the carrier. At the next instant in time, the photon pair will be created with a different frequency separation, but always conserving energy. In this way, the vacuum seed acts to populate all frequencies with these photon pairs. These photon pairs, seeded by the vacuum itself, reduce the quantum noise in the output field and are the source of squeezing.

A list is made detailing the important properties of this system that define the production and the spectral shape of the squeezing that exits the OPO system.

- Squeezing originates from the production of these sideband pairs. These pairs replace the vacuum fluctuations and tend to align in some quadrature and cancel in the orthogonal one. The quadrature where they cancel, and hence where the noise is reduced, is the squeezed quadrature.

- The alignment of the generated sidebands from these photon pairs is a function of the phase of the pump. Hence, relative to a carrier field, we can rotate our squeezing quadrature by rotating the phase of the pump.

- The squeezing as it is produced is equal at all frequencies, limited only by the phase matching range of the material. This is because the vacuum fluctuations are white, existing at all frequencies with equal magnitude. However, the OPO consists of a non-linear medium inside an optical cavity and the optical cavity will shape the squeezing through its spectral response.

- Squeezing is centred around the fundamental frequency. This is because the photon pairs are centred about half of the pump frequency due to conservation of energy.

Figure 4.4 shows the conceptual process of the OPO. Comparison with Figure 4.2 shows that the two systems share similarities.

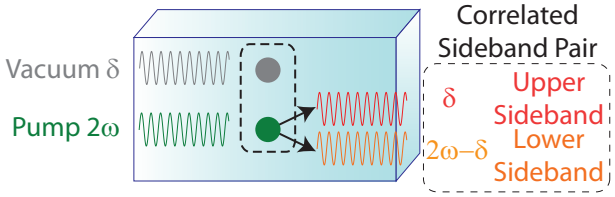


Figure 4.4: The optical parametric oscillator (OPO). The system is driven with a pump field at frequency 2ω and seeded by the vacuum field. Inside the non-linear medium the presence of a virtual photon causes the pump photon to split into two photons of frequency δ and $2\omega - \delta$. This process creates photon pairs across all frequencies.

4.5 Phase Matching

In the previous section, the various non-linear systems and the photon interactions which take place were introduced. Energy conservation was introduced but another quantity also has to be conserved, momentum. The technique of conserving the momentum is known as *phase matching*. The origin of phase mismatch is dispersion within the non-linear medium. In order for phase matching to occur, the wave vector of the pump field must be equal to twice the wave vector of the seed field, which can be written

$$\vec{k}_{pump} = 2\vec{k}_{seed}, \tag{4.10}$$

where $\vec{k}_{pump}$ is the wave vector of the pump field and $\vec{k}_{seed}$ is the wave vector of the seed field. The process typically occurs in a non-linear medium that has been placed inside an optical cavity. The two fields are thus co-propagating and the directional information is not important. We can then write

$$\frac{n_{pump}\,\omega_{pump}}{c} = 2\frac{n_{seed}\,\omega_{seed}}{c}, \tag{4.11}$$

where n_{pump} and n_{seed} are the refractive indices experienced by the pump and seed fields respectively and c is the speed of light in a vacuum.

Using Equation 4.2 we can simplify this to indicate that in order for phase matching to occur, the refractive index of the pump and seed fields must be equal. We can write

$$n_{pump} = n_{seed}. \tag{4.12}$$

The refractive index of the pump field is typically not equal to the refractive index of the seed field and so the phase relationship between the two fields varies, reducing the efficiency of the non-linear processes. Clever techniques are used to ensure that phase matching is satisfied. Figure 4.5 shows the phase relationship between a fundamental wave and its harmonic for the phase matched case and the phase mismatched case, respectively. In the phase matched case, a), the waves have a constant phase relationship whilst in the case where they are not phase matched, b), the waves do not maintain a constant relationship.

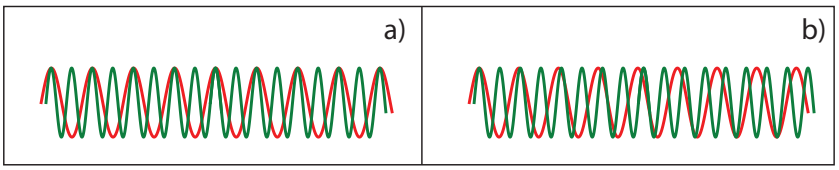


Figure 4.5: Two waves at ω, red, and 2ω, green, when the waves are phase matched a) and phase mismatched b).

We can also illustrate phase matching using a simple phasor diagram. Figure 4.6 shows the resultant second harmonic field in an SHG setup for the phase matched a), quasi-phase matched b) and phase mismatched c), cases. In the phase matched case, a), the wave vectors of the second harmonic and fundamental fields match (See § 4.5.1. In the quasi-phase matched case, b), the wave vectors do not match but periodic poling, marked by the dotted lines, ensures that the process does not reverse (See § 4.5.2). In the final

case, c), where the waves are phase mismatched, the energy oscillates between the second harmonic and fundamental fields.

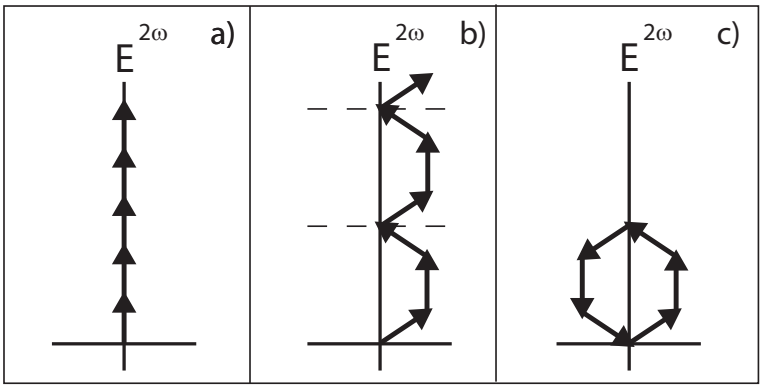


Figure 4.6: Phasor diagrams showing the generated second harmonic field, $E^{2\omega}$ in the phase matched, a), quasi-phase matched, b), and phase mismatched cases in an SHG setup.

4.5.1 Birefringent Phase Matching

One technique commonly used to achieve phase matching is *birefringent phase matching*, illustrated in Figure 4.6. The non-linear mediums used in these non-linear systems are often birefringent. This is in part due to the fact that the crystal has to lack inversion symmetry in order to exhibit the second order non-linearity [75]. The two fields will ordinarily experience non-equal refractive indices depending upon which crystal axis they are polarised along. Rotating the angle of polarization can give any refractive index between the refractive index of the ordinary and the extraordinary axis. These values can be altered by increasing the temperature of the crystal as described by the Sellmeier equation for the material. A crystal cut is then found where varying the angle of the input fields and the temperature of the crystal, will result in an operating regime where their refractive indices match.

The disadvantage to birefringent phase matching is that the angle where phase matching occurs is typically not in the direction of the strongest $\chi^{(2)}$ coefficient. For example, in lithium niobate the non-linear coefficient is 27 pm/V in the d_{33} crystal direction but at the birefringent phase matching angle, which lies along the d_{31} plane, the coefficient is 4.3 pm/V [107]. The conversion efficiency is proportional to the square of the non-linear coefficient for typical applications. Clearly, the birefringent phase matching technique is not harnessing the full potential of this material. Another disadvantage to the birefringent phase matching technique, particularly in this material, is that the temperature at which phase matching occurs cannot be tailored under most circumstances. This leaves lithium niobate with a phase matching temperature of typically around 100 degrees Celsius for a pump field of 532 nm, [108]. It is technically challenging to provide such a temperature environment for the material which is stable.

4.5.2 Quasi-Phase Matching and Periodic Poling

Periodic poling is an alternative method for achieving momentum conservation, one in which the phase matching temperature can be tailored and the largest non-linear coefficent

strength of the material can be utilised. This technique does not achieve a phase matching condition within the material, but instead compensates for dispersion by flipping the crystal structure at specific intervals, thus avoiding reversal of the non-linear processes, as illustrated in Figure 4.6. The most common method of achieving domain inversion is *electric field periodic poling*. In this technique, a very strong electric field is applied to metallic "fingers" that are placed on top of the crystal [109]. These large electric fields rearrange the crystal lattice, changing the sign of the non-linear interaction term. The spacing between these fingers and the electric field applied determines the width of the periodic poling.

The distance at which the non-linear process begins to reverse in a phase mismatched system is known as the coherence length and is defined as

$$L_C = \frac{2\pi}{|\Delta k|}, \tag{4.13}$$

where L_C is the coherence length and $\Delta k = \vec{k}_3 - \vec{k}_2 - \vec{k}_1$ is the phase mismatch term, and we have assumed that the fields are co-propagating. The three wavevectors, $\vec{k}_{1,2,3}$ correspond to the wave vectors of the three electric fields that interact in the three-wave-mixing process. For example, in the NDOPA we can write

$$\begin{aligned} \Delta k &= \vec{k}_3 - \vec{k}_2 - \vec{k}_1 \\ &= \vec{k}_3 - \vec{k}_1 - \vec{k}_1 \\ &= \vec{k}_{pump} - 2\vec{k}_{seed} \end{aligned} \tag{4.14}$$

For collinear beams and in a degenerate system, the phase matching condition is equivalent to writing

$$n_{pump} = n_{seed}, \tag{4.15}$$

where n_{pump} and n_{seed} are the refractive indices for the pump and seed fields.

The standard method of poling a non-linear medium is to pole the crystal at intervals equal to the coherence length. This is known as *first order periodic poling*. Sometimes, however, it is technically not feasible to achieve first order poling. This can be due to the non-linear material requiring very large electric fields to facilitate domain inversion or it can be due to a large phase mismatch in a particular crystal direction, which results in a small coherence length [109]. If first order phase matching is not possible then third order phase matching, where the domain is flipped every 3 coherence lengths, can be attempted.

Chapter 5

Optical Cavities

In the following chapter we introduce the framework for the investigation of optical cavities using the standard methods (detailed in such works as[110, 83, 108]) and use this framework to then investigate the theoretical amount of squeezing that the cavity can produce. First, the semi-classical cavity equations of motion are derived for a simple empty, single mode optical cavity. Through simplification of these equations, the classical amplitude and phase response of this cavity are derived. Simple extension of this model allows for the derivation of the semi-classical noise response of the system. From here, the cavity equations of motion are extended to allow for a second field and the exchange of photons between these two fields, facilitated by the second order response of a non-linear medium. The relevant cavity properties and the dependence of these properties upon one another are investigated.

5.1 Cavity Equations of Motion

We begin by introducing the cavity equations of motion for a simple single mode empty cavity. This is a semi-classical approach to the problem in which we describe the time behaviour of the intra-cavity and external fields involved using the methods put forth by Gardiner and Collett [111]. Two partially transmissive mirrors, the input and output couplers, with power reflectivities of R^{in} and R^{out} respectively, and a third mirror representing the loss, $T^l = 1 - R^l$, make up the optical components. The round trip time, τ, defined as the time taken for the light to travel one round-trip of the cavity, is defined as the optical path length, L, divided by the speed of light, c, written

$$\tau[s] = \frac{L}{c}. \tag{5.1}$$

A driving field, A^{dr}, is incident on the input coupling mirror, A^{ref} is the field that reflects off the input coupler and A^{trans} is the field that exits the output coupler. We treat the cavity as an open quantum system, a system into which photons can be added, through a driving field, and photons can be irreversibly lost, through intra-cavity field decay. Cavity decay rates are a measure of the rate at which photons exit, or enter, the system. The decay rate in its entirety is an infinite Taylor expansion, from which only the first two terms are generally retained. This approximation is most accurate for high reflectivities, but for moderate to high reflectivities, it has been shown by White [112] that we can achieve reasonable accuracy by defining the cavity decay rates, κ^x, as

$$\kappa^x[s^{-1}] = \frac{1 - \sqrt{R^x}}{\tau}, \tag{5.2}$$

for any mirror, x. The setup is shown in Figure 5.1.

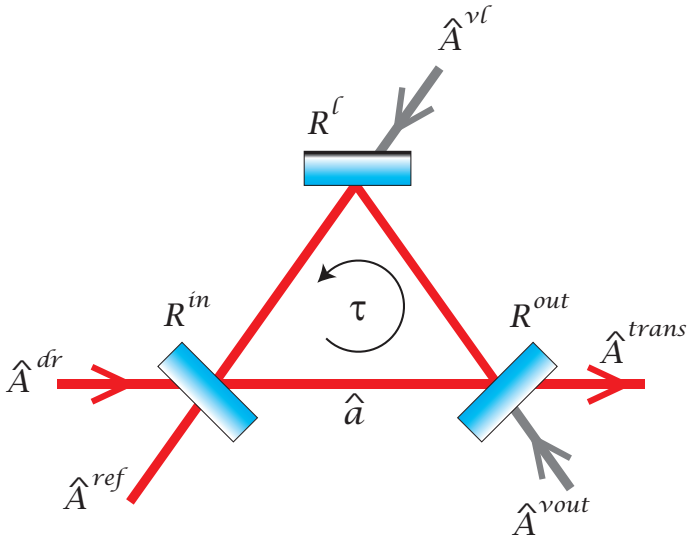


Figure 5.1: The optical cavity. Loss is modelled as transmission through one of the mirrors, $T^l = 1 - R^l$. A driving field, A^{dr}, enters the cavity through the input coupler, R^{in}. The reflected field, A^{out}, consists of the promptly reflected field and the escaping intra-cavity field, the transmitted field consists of the intra-cavity field escaping from the output mirror, with reflectivity R^{out}.

The equation of motion that describes the time evolution of the intra-cavity field, $\hat{a}$, for this system, is written [113]

$$\dot{\hat{a}} = i\Delta\hat{a} - \kappa^a\hat{a} + \sqrt{2\kappa^{in}}\hat{A}^{dr} + \sqrt{2\kappa^{out}}\hat{A}^{vout} + \sqrt{2\kappa^{l}}\hat{A}^{vl}, \tag{5.3}$$

where $\Delta = \omega_0 - \omega_A$ is the detuning of the driving field, ω_A, from the cavity resonance, ω_0, κ^{in} is the decay rate of the input coupler, κ^{out} is the decay rate of the output coupler, κ^l is the decay rate of the intra-cavity loss, $\hat{A}^{dr}$ is the external cavity driving field, and $\hat{A}^{vl}$ and $\hat{A}^{vout}$ are the external vacuum fields. The total cavity decay rate, κ^a, is defined as the sum of the individual cavity decay rates

$$\kappa^a = \kappa^{in} + \kappa^{out} + \kappa^l. \tag{5.4}$$

The external travelling fields have units of $\sqrt{\frac{\text{photons}}{s}}$ such that $\hat{A}^\dagger\hat{A}$, the number operator, gives the photons per second in the field and is thus proportional to the intensity. The intra-cavity fields have dimensions of $\sqrt{\text{photons}}$ such that $a^\dagger a$ gives the number of photons in the cavity mode.

The power in the external driving field, $\hat{A}^{dr}$, is proportional to the number of photons per second and can be determined via

$$P^{dr}[W] = \frac{\hat{A}^{\dagger dr}\hat{A}^{dr}hc}{\lambda^{dr}}, \tag{5.5}$$

where P^{dr} is the power in the external driving field, A^{dr}, in watts, h is Planck's constant, c is the speed of light, and λ^{dr} is the wavelength of the driving field.

To solve Equation 5.3 we move into the frequency domain by taking the Fourier trans-

form and find

$$i\omega\check{a} = i\Delta\check{a} - \kappa^a\check{a} + \sqrt{2\kappa^{in}}\check{A}^{dr} + \sqrt{2\kappa^{out}}\check{A}^{vout} + \sqrt{2\kappa^{l}}\check{A}^{vl}, \tag{5.6}$$

where the hats have been replaced with carons, or overturned hats, to signify that these operators are now in the frequency domain. The ω term arising from the Fourier transform is the frequency that a measurement of this state is made at, where a value of $\omega = 0$ indicates a measurement at the carrier frequency of the input field. A non-zero term implies investigation of the system when the cavity is off resonance, or equivalently, the frequency of the laser is shifted from the resonance condition of the cavity. To arrive at Equation 5.6 we make use of the identity

$$\mathcal{F}[\dot{X}(t)] = i\omega X(\omega), \tag{5.7}$$

where $\mathcal{F}[\dot{X}(t)]$ is the Fourier transform of the function $\dot{X}(t)$. Equation 5.6 can now be written

$$\check{a} = \frac{\sqrt{2\kappa^{in}}\check{A}^{dr} + \sqrt{2\kappa^{out}}\check{A}^{vout} + \sqrt{2\kappa^{l}}\check{A}^{vl}}{\kappa^a + i(\omega - \Delta)}, \tag{5.8}$$

from which the reflected and transmitted fields of the cavity can now be determined. The reflected and transmitted fields are determined using the input/output relations [111]

$$\check{A}^{ref} = \sqrt{2\kappa^{in}}\,\check{a} - \check{A}^{dr} \tag{5.9}$$

$$\check{A}^{trans} = \sqrt{2\kappa^{out}}\,\check{a} - \check{A}^{vout}. \tag{5.10}$$

This results in output fields that can be written

$$\begin{aligned}
\check{A}^{ref} &= \sqrt{2\kappa^{in}}\frac{\sqrt{2\kappa^{in}}\check{A}^{dr} + \sqrt{2\kappa^{out}}\check{A}^{vout} + \sqrt{2\kappa^{l}}\check{A}^{vl}}{\kappa^a + i(\omega - \Delta)} - \check{A}^{dr} \\
&= \frac{(2\kappa^{in} - \kappa^a - i\omega)\check{A}^{dr} + 2\sqrt{\kappa^{in}\kappa^{out}}\check{A}^{vout} + 2\sqrt{\kappa^{in}\kappa^{l}}\check{A}^{vl}}{\kappa^a + i(\omega - \Delta)}
\end{aligned} \tag{5.11}$$

$$\begin{aligned}
\check{A}^{trans} &= \sqrt{2\kappa^{out}}\frac{\sqrt{2\kappa^{in}}\check{A}^{dr} + \sqrt{2\kappa^{out}}\check{A}^{vout} + \sqrt{2\kappa^{l}}\check{A}^{vl}}{\kappa^a + i(\omega - \Delta)} - \check{A}^{vout} \\
&= \frac{2\sqrt{\kappa^{in}\kappa^{out}}\check{A}^{dr} + (2\kappa^{out} - \kappa^a - i\omega)\check{A}^{vout} + 2\sqrt{\kappa^{out}\kappa^{l}}\check{A}^{vl}}{\kappa^a + i(\omega - \Delta)}.
\end{aligned} \tag{5.12}$$

5.2 Classical Cavity Phase and Amplitude Response

Now that the equations describing the reflected and transmitted fields have been derived, we will first investigate the classical response of the cavity. This is done by ignoring the terms involving vacuum fields in Equation 5.12, $\check{A}^{vout} = \check{A}^{vl} = 0$. We also assume that the wavelength of the laser matches the resonance condition of the cavity, $\Delta = 0$. The reflected and transmitted fields can then be written

$$\check{A}^{ref} = \frac{(2\kappa^{in} - \kappa^a - i\omega)\check{A}^{dr}}{\kappa^a + i\omega} \tag{5.13}$$

$$\check{A}^{trans} = \frac{2\sqrt{\kappa^{in}\kappa^{out}}\check{A}^{dr}}{\kappa^a + i\omega}. \tag{5.14}$$

It was shown in Equation 5.5 that the intensities of these fields are proportional to the number operator. The phases of the fields can be found by taking their arguments. Figure 5.2 shows the classical response of the cavity.

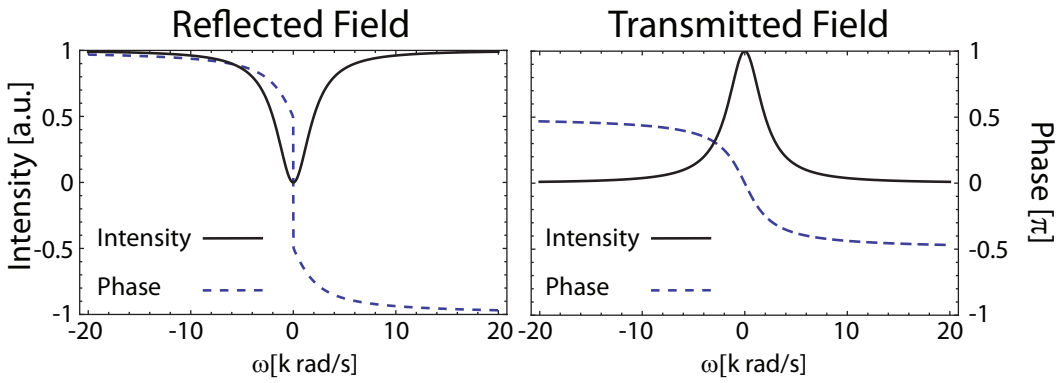


Figure 5.2: The classical response of the two-ended triangle cavity. A critically coupled, lossless cavity is chosen ($\kappa^{in} = \kappa^{out} = 1000$, $\kappa^{l} = 0$)

5.2.1 Circulating Power and Finesse

On resonance, the cavity provides enhancement of the optical field. A relationship between the incident field and the circulating field can be calculated, the result is [114]

$$\frac{P^{circ}}{P^{inc}} = \frac{1 - R^{in}}{(1-r)^2}, \tag{5.15}$$

where r is defined as the product of the square root of the reflectivity coefficients of all sources of cavity decay, in this case

$$r = \sqrt{R^{in} \times R^{out} \times R^{l}}. \tag{5.16}$$

The enhancement of the field results in a large circulating power. The circulating power is related to another cavity parameter, the finesse. For relatively high mirror reflectivities, we can define the unitless finesse of our cavity as [115]

$$\mathcal{F} = \frac{\pi\sqrt{r}}{1-r}. \tag{5.17}$$

The finesse is a measure of how efficiently the cavity resonates, which is equivalent to the number of times, on average, that the photons will bounce around the cavity before exiting the system. As can be seen in Equation 5.17, it has no dependence on the length of the cavity. In cavity systems where the intra-cavity loss is much less or similar to the intra-cavity mode loss due to the coupling mirrors, the finesse, although difficult to measure accurately, can provide a good indication of the intra-cavity loss.

The finesse is often used as a means of estimating the circulating power inside a cavity. Figure 5.3 shows the finesse and the circulating power as the power reflectivity of the input coupler is varied. The two values follow a similar trend but using the finesse as an estimate for the circulating power will result in higher circulating power, with the

discrepancy between the two values becoming larger as the finesse increases.

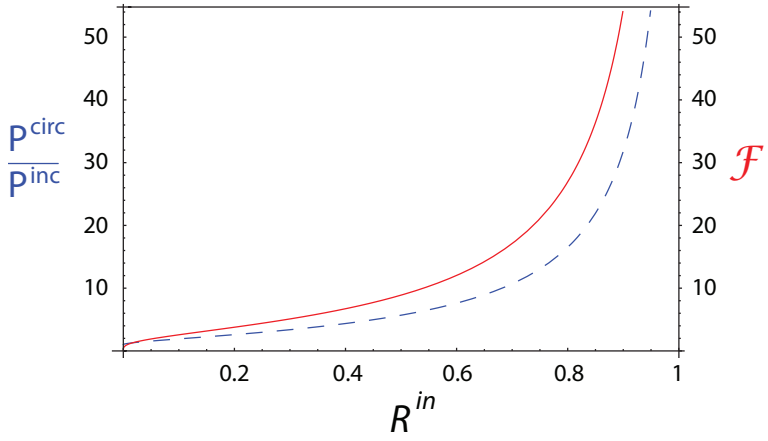


Figure 5.3: The finesse, solid red line, and the circulating power gain, dashed blue line, as the input coupler reflectivity, R^{in}, is varied.

5.2.2 Free Spectral Range

The Free Spectral Range, or FSR, of the cavity is a measure of the spacing between resonances and is defined as [75]

$$\text{FSR[Hz]} = \frac{c}{L}, \tag{5.18}$$

where c is the speed of light and L is the optical path length of one round trip of the cavity. The optical path length is the physical length multiplied by the refractive index of the material traversed through this length.

5.2.3 Linewidth

The linewidth, or Full Width at Half Maximum (FWHM), is a measure of the bandwidth of each resonance of the cavity, it can be defined as the ratio of the FSR to the finesse

$$\Delta_\nu\text{[Hz]} = \frac{\text{FSR}}{\mathcal{F}}. \tag{5.19}$$

This value is important to squeezed light generation because the bandwidth of the cavity limits the bandwidth of the squeezing produced. This is due to the fact that the bandwidth of the non-linear gain is limited by the bandwidth of the cavity. The linewidth also defines the transmission and reflection characteristics of the cavity, which plays a role in how the various locking loops are constructed.

5.3 Semi-Classical Cavity Noise Response

Investigation of Equations 5.12 in their entirety facilitates the analysis of the noise response of the cavity. This analysis includes effects originating from the vacuum fluctuations that enter the system through the various cavity decay mechanisms. The steps towards finding

an expression for the noise response of the cavity are as follows; the operators in the cavity equation of motion are first linearised, the quadrature operators are then calculated, and finally, the spectra of the fields are determined.

Linearisation follows the method described in 3.3. Note that the vacuum fields, $\check{A}^{vout}$ and $\check{A}^{vl}$, do not have any coherent amplitude and as such, linearisation of these terms results in only the first order fluctuating term. These steps results in

$$\check{A}^{ref} = \frac{(2\kappa^{in} - \kappa^a - i\omega)(\alpha^{dr} + \delta\check{A}^{dr}) + 2\sqrt{\kappa^{in}\kappa^{out}}\delta\check{A}^{vout} + 2\sqrt{\kappa^{in}\kappa^{l}}\delta\check{A}^{vl}}{\kappa^a + i\omega} \quad (5.20)$$

$$\check{A}^{trans} = \frac{2\sqrt{\kappa^{in}\kappa^{out}}(\alpha^{dr} + \delta\check{A}^{dr}) + (2\kappa^{out} - \kappa^a - i\omega)\delta\check{A}^{vout} + 2\sqrt{\kappa^{out}\kappa^{l}}\delta\check{A}^{vl}}{\kappa^a + i\omega}. \quad (5.21)$$

The next step is to calculate the quadrature operators. These were defined in Equations 3.1 and 3.2 but are written again in the frequency domain for clarity

$$\begin{aligned} \check{X}^{+} &= \mathcal{F}[\hat{X}^{+}] \\ &= \mathcal{F}[\hat{A}] + \mathcal{F}[\hat{A}^{\dagger}] \\ &= \check{A}(\omega) + \check{A}^{\dagger}(-\omega) \quad (5.22) \\ \check{X}^{-} &= \mathcal{F}[\hat{X}^{-}] \\ &= i(\mathcal{F}[\hat{A}^{\dagger}] - \mathcal{F}[\hat{A}]) \\ &= i(\check{A}^{\dagger}(-\omega) - \check{A}(\omega)). \quad (5.23) \end{aligned}$$

Equations 5.22 have been achieved using the identity [83]

$$\mathcal{F}[\hat{A}^{\dagger}(t)] = \check{A}^{\dagger}(-\omega). \quad (5.24)$$

These steps result in the quadrature operator equations. All terms without a fluctuating component are dropped because they do not contribute to the noise spectra and all second order terms are removed because, under the assumptions for linearisation (see § 3.3), these terms are small enough to be discarded, resulting in

$$\check{X}^{\pm ref} = \frac{(2\kappa^{in} - \kappa^a - i\omega)\delta\check{X}^{\pm dr} + 2\sqrt{\kappa^{in}\kappa^{out}}\delta\check{X}^{\pm vout} + 2\sqrt{\kappa^{in}\kappa^{l}}\delta\check{X}^{\pm vl}}{\kappa^a + i\omega} \quad (5.25)$$

$$\check{X}^{\pm trans} = \frac{2\sqrt{\kappa^{in}\kappa^{out}}\delta\check{X}^{\pm dr} + (2\kappa^{out} - \kappa^a - i\omega)\delta\check{X}^{\pm vout} + 2\sqrt{\kappa^{out}\kappa^{l}}\delta\check{X}^{\pm vl}}{\kappa^a + i\omega}. \quad (5.26)$$

Finally, we wish to calculate the noise spectra, $V[\omega]$, of these quadrature operators. The noise spectra is defined as the variance of the quadrature operators

$$V[\omega] = \langle \delta\check{X}^{\dagger}(\omega)\delta\check{X}(\omega)\rangle, \quad (5.27)$$

for any quadrature operator, $\check{X}$, in the frequency domain. Calculation of the noise spectra from Equations 5.25 and 5.26 results in

$$V^{\pm ref} = \frac{|2\kappa^{in} - \kappa^a - i\omega|^2 V^{\pm dr} + 4\kappa^{in}\kappa^{out} + 4\kappa^{in}\kappa^{l}}{|\kappa^a + i\omega|^2} \quad (5.28)$$

$$V^{\pm trans} = \frac{4\kappa^{in}\kappa^{out}V^{\pm dr} + |2\kappa^{out} - \kappa^a - i\omega|^2 + 4\kappa^{out}\kappa^{l}}{|\kappa^a + i\omega|^2}, \quad (5.29)$$

where the variances of the vacuum terms have been replaced with unity, the noise of the vacuum state. Using the fact that $\kappa^a = \kappa^l + \kappa^{in} + \kappa^{out}$, Equations 5.28 and 5.29 can be rearranged to the more succinct forms

$$V^{\pm ref} = 1 + \frac{((2\kappa^{in} - \kappa^a)^2 + \omega^2)(V^{\pm dr} - 1)}{(\kappa^a)^2 + \omega^2} \tag{5.30}$$

$$V^{\pm trans} = 1 + \frac{4\kappa^{in}\kappa^{out}(V^{\pm dr} - 1)}{(\kappa^a)^2 + \omega^2}. \tag{5.31}$$

The cavity noise response given by Equations 5.31 is illustrated in Figure 5.4. It can be seen that, in transmission, the cavity acts as a low pass filter around the resonance frequency. Consequently, the cavity acts as a high pass filter in reflection. Classical laser intensity noise above the characteristic frequency of the cavity can therefore be removed through the use of such a cavity. It can also be seen that the noise does not drop below the quantum noise level (QNL), where the noise is unity. This is a direct consequence of quantum mechanics that arises due to the presence of vacuum fields.

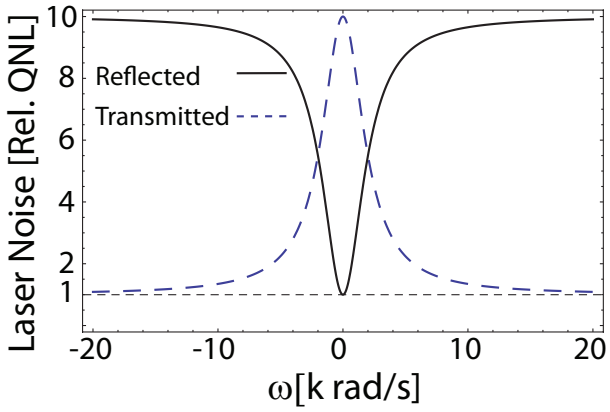


Figure 5.4: The noise response of the optical cavity. A critically coupled, lossless cavity was chosen, $\kappa^{in} = \kappa^{out} = 1000$, $\kappa^l = 0$, and the driving laser field, $\check{A}^{dr}$, has noise equal to ten times the quantum limit, $V^{dr} = 10$.

5.4 The Optical Parametric Oscillator

By extending the techniques presented up to this point, we can now investigate the squeezing response of an optical parametric oscillator. The next step is to take the equations of motion for the empty cavity (Equation 5.3) and introduce additional fields and a non-linear medium. The non-linear medium facilitates the exchange of photons between the seed and idler fields, A_1 and A_2 respectively, and the second harmonic, or pump field, A_3. The formalisms used here are those introduced by Drummond [110]. We begin by introducing the most general form of the equations that we will require and from these, derive the behaviour of the OPO. In Chapter 6 we will use these equations to investigate coherent locking. The first step is to define the equations of motion for three fields, a pump field and nondegenerate signal and idler fields, in much the same way as shown in

§ 5.1. They are written

$$\begin{aligned}
\dot{\hat{a}}_1 &= i\Delta_1\hat{a}_1 - \kappa_1\hat{a}_1 + g\hat{a}_3\hat{a}_2^\dagger + \sqrt{2\kappa_1^{in}}\hat{A}_1^{dr} + \sqrt{2\kappa_1^{out}}\hat{A}_1^{vout} + \sqrt{2\kappa_1^{l}}\hat{A}_1^{vl} \\
\dot{\hat{a}}_2 &= i\Delta_2\hat{a}_2 - \kappa_2\hat{a}_2 + g\hat{a}_3\hat{a}_1^\dagger + \sqrt{2\kappa_2^{in}}\hat{A}_2^{dr} + \sqrt{2\kappa_2^{out}}\hat{A}_2^{vout} + \sqrt{2\kappa_2^{l}}\hat{A}_2^{vl} \\
\dot{\hat{a}}_3 &= i\Delta_3\hat{a}_3 - \kappa_3\hat{a}_3 - g\hat{a}_1\hat{a}_2 + \sqrt{2\kappa_3^{in}}\hat{A}_3^{dr} + \sqrt{2\kappa_3^{out}}\delta\hat{A}_3^{vout} + \sqrt{2\kappa_3^{l}}\hat{A}_3^{vl}. \qquad (5.32)
\end{aligned}$$

Properties relating to the non-degenerate signal and idler fields are written with subscripts 1 and 2 respectively, and the pump field properties with subscript 3. The strength of the non-linear interaction is given by $g\left[s^{-\frac{1}{2}}\right]$ and describes the rate at which photons transfer between fields. The Equations in 5.32 can be used to describe the classical and quantum behaviour of OPO, NDOPA and DOPA systems. Assumptions for various operating regimes can be made which will result in simple analytical solutions to the intra-cavity field solutions. Most treatments of these equations assume that the OPO is not resonant at the pump frequency. This allows for the adiabatic elimination of the pump field, allowing one to assume $\dot{\hat{a}}_3 = 0$. This thesis investigates doubly resonant systems and as such this assumption is generally not valid.

The fact that we are operating the cavity as an OPO does however allow us to simplify the situation in other ways. First, the system is degenerate because it is vacuum seeded, with no coherent amplitude and a variance equal to unity. Finally, we choose to assume that the frequency of our fields match the resonant frequency of the cavity, removing the detuning terms, $\Delta_1 = \Delta_3 = 0$. This results in [110]

$$\begin{aligned}
\dot{\hat{a}}_1 &= -\kappa_1\hat{a}_1 + g\hat{a}_1^\dagger\hat{a}_3 + \sqrt{2\kappa_1^{in}}\delta\hat{A}_1^{dr} + \sqrt{2\kappa_1^{out}}\delta\hat{A}_1^{vout} + \sqrt{2\kappa_1^{l}}\delta\hat{A}_1^{vl} \\
\dot{\hat{a}}_3 &= -\kappa_3\hat{a}_3 - \frac{g}{2}\hat{a}_1\hat{a}_1 + \sqrt{2\kappa_3^{in}}\hat{A}_3^{dr} + \sqrt{2\kappa_3^{out}}\delta\hat{A}_3^{vout} + \sqrt{2\kappa_3^{l}}\delta\hat{A}_3^{vl}. \qquad (5.33)
\end{aligned}$$

Although we are operating as an OPO, the coherent amplitude of the seed field, $\alpha_1 = 0$, has not yet been removed because it will be used to illustrate noise couplings in Appendix 1. The next step is to linearise the field operators and remove all second order fluctuation terms and terms that have no fluctuating component. At this point we also make the assumption that the coherent amplitude terms are real. This assumption is valid assuming two things; that the driving fields are real, and that the OPO is operated below threshold. It can be shown that below threshold, the phase of the intra-cavity fields follows the phase of the driving fields [110]. In intra-cavity field units, the threshold of the OPO is defined as

$$a_3^{crit} = \frac{\kappa_1\kappa_3}{g}, \qquad (5.34)$$

whilst the external cavity threshold is written

$$A_3^{crit} = \frac{\kappa_1\kappa_3}{g\sqrt{2k_3^{in}}}, \qquad (5.35)$$

and the threshold in terms of pump field driving power is written

$$P_3^{crit}[W] = \left(\frac{\kappa_1\kappa_3}{g\sqrt{2k_3^{in}}}\right)^2 \frac{hc}{\lambda_3}. \tag{5.36}$$

The OPO threshold defines a point where the system enters a new mode of operation. It is the point where the round trip loss of the generated field at the fundamental frequency is equal to the round trip gain from the non-linear down-conversion of pump photons into fundamental photons. Operating near to this point, but not over it, produces stronger levels of squeezing as more photon pairs are produced. Above this point, the cavity can behave erratically and new noise couplings are introduced [116].

The next assumption made is that the non-linear interaction term is real. This assumes perfect phase matching within the non-linear medium which relies on sufficient temperature control and is a reasonable assumption for most OPO systems. Making these simplifications and calculating the hermitian conjugate of the intra cavity fields results in

$$\begin{aligned}
\delta\dot{\hat{a}}_1 &= -\kappa_1\delta\hat{a}_1 + g\alpha_1\delta\hat{a}_3 + g\alpha_3\delta\hat{a}_1^\dagger + \sqrt{2\kappa_1^{in}}\delta\hat{A}_1^{dr} + \sqrt{2\kappa_1^{out}}\delta\hat{A}_1^{vout} + \sqrt{2\kappa_1^{l}}\delta\hat{A}_1^{vl} \\
\delta\dot{\hat{a}}_1^\dagger &= -\kappa_1\delta\hat{a}_1^\dagger + g\alpha_1\delta\hat{a}_3 + g\alpha_3\delta\hat{a}_1 + \sqrt{2\kappa_1^{in}}\delta\hat{A}_1^{\dagger dr} + \sqrt{2\kappa_1^{out}}\delta\hat{A}_1^{\dagger vout} + \sqrt{2\kappa_1^{l}}\delta\hat{A}_1^{\dagger vl} \\
\delta\dot{\hat{a}}_3 &= -\kappa_3\delta\hat{a}_3 - g\alpha_1\delta\hat{a}_1 + \sqrt{2\kappa_3^{in}}\delta\hat{A}_3^{dr} + \sqrt{2\kappa_3^{out}}\delta\hat{A}_3^{vout} + \sqrt{2\kappa_3^{l}}\delta\hat{A}_3^{vl} \\
\delta\dot{\hat{a}}_3^\dagger &= -\kappa_3\delta\hat{a}_3^\dagger - g\alpha_1\delta\hat{a}_1^\dagger + \sqrt{2\kappa_3^{in}}\delta\hat{A}_3^{\dagger dr} + \sqrt{2\kappa_3^{out}}\delta\hat{A}_3^{\dagger vout} + \sqrt{2\kappa_3^{l}}\delta\hat{A}_3^{\dagger vl}.
\end{aligned} \tag{5.37}$$

From these equations we can now write the amplitude and phase quadrature operator equations of motion using Equations 3.1 and 3.2. For our purposes, it is sufficient to determine only one of the quadrature operators. It can be shown that the sign of the non-linear interaction term, g, completely determines whether one observes the amplitude or phase quadrature output variances [83]. The amplitude quadrature noise operators can be written

$$\begin{aligned}
\delta\dot{\hat{X}}_1 &= -\kappa_1\delta\hat{X}_1 + g\alpha_1\delta\hat{X}_3 + g\alpha_3\delta\hat{X}_1^\dagger + \sqrt{2\kappa_1^{in}}\delta\hat{X}_1^{dr} + \sqrt{2\kappa_1^{out}}\delta\hat{X}_1^{vout} + \sqrt{2\kappa_1^{l}}\delta\hat{X}_1^{vl} \\
\delta\dot{\hat{X}}_3 &= -\kappa_3\delta\hat{X}_3 - g\alpha_1\delta\hat{X}_1 + \sqrt{2\kappa_3^{in}}\delta\hat{X}_3^{dr} + \sqrt{2\kappa_3^{out}}\delta\hat{X}_3^{vout} + \sqrt{2\kappa_3^{l}}\delta\hat{X}_3^{vl}.
\end{aligned} \tag{5.38}$$

Now we take the Fourier transform of Equations 5.38 to transform the quadrature operators into the frequency domain such that we can determine the spectra. It is important to note that two of these operators are still in intra-cavity units, $\delta\check{X}_1$ and $\delta\check{X}_3$, whilst the rest are in travelling wave units. Taking the Fourier transform results in

$$\begin{aligned}
i\omega\delta\check{X}_1 &= -\kappa_1\delta\check{X}_1 + g\alpha_1\delta\check{X}_3 + g\alpha_3\delta\check{X}_1^\dagger + \sqrt{2\kappa_1^{in}}\delta\check{X}_1^{dr} + \sqrt{2\kappa_1^{out}}\delta\check{X}_1^{vout} + \sqrt{2\kappa_1^{l}}\delta\check{X}_1^{vl} \\
i\omega\delta\check{X}_3 &= -\kappa_3\delta\check{X}_3 - g\alpha_1\delta\check{X}_1 + \sqrt{2\kappa_3^{in}}\delta\check{X}_3^{dr} + \sqrt{2\kappa_3^{out}}\delta\check{X}_3^{vout} + \sqrt{2\kappa_3^{l}}\delta\check{X}_3^{vl}.
\end{aligned} \tag{5.39}$$

We want to solve these equations for the intra-cavity field quadratures, $\delta\check{X}_1$ and $\delta\check{X}_3$. At this point we will set the coherent amplitude of the seed field to zero, $\alpha_1 = 0$, and leave the details of the derivation of the intra-cavity field quadratures for Appendix 1.

The result is that we can write $\delta\check{X}_1$ as

$$\delta\check{X}_1 = \frac{\sqrt{2\kappa_1^{in}}(\kappa_3 + i\omega)}{\alpha_1\alpha_1 g^2 - (g\alpha_3 - \kappa_1 - i\omega)(\kappa_3 + i\omega)} \times \left(\delta\check{X}_1^{dr} + \frac{\sqrt{2\kappa_1^{out}}}{\sqrt{2\kappa_1^{in}}}\delta\check{X}_1^{vout} + \frac{\sqrt{2\kappa_1^{l}}}{\sqrt{2\kappa_1^{in}}}\delta\check{X}_1^{vl}\right). \quad (5.40)$$

The final steps are to use Equation 5.40 to determine the amplitude quadrature operator for the output field using the input-output relations, Equation 5.9, and then to determine the variance of the output field, using Equation 3.7. From the variance of the amplitude quadrature, one can determine the variance of the phase quadrature. The details of these steps are given in the appendix. The resulting amplitude and phase quadrature variances are

$$V_{out}^{+} = 1 + \frac{4g\kappa_1^{out}\alpha_3}{(\kappa_1 - g\alpha_3)^2 + \omega^2} \quad (5.41)$$

$$V_{out}^{-} = 1 - \frac{4g\kappa_1^{out}\alpha_3}{(\kappa_1 + g\alpha_3)^2 + \omega^2}. \quad (5.42)$$

Figure 5.5 shows the variance in the squeezed quadrature and the variance in the orthogonal anti-squeezed quadrature for various internal losses. The detection frequency, ω, is set to a value of zero because we assume that we are looking at frequencies well within the linewidth of the cavity and hence the filtering behaviour of the cavity is negligible.

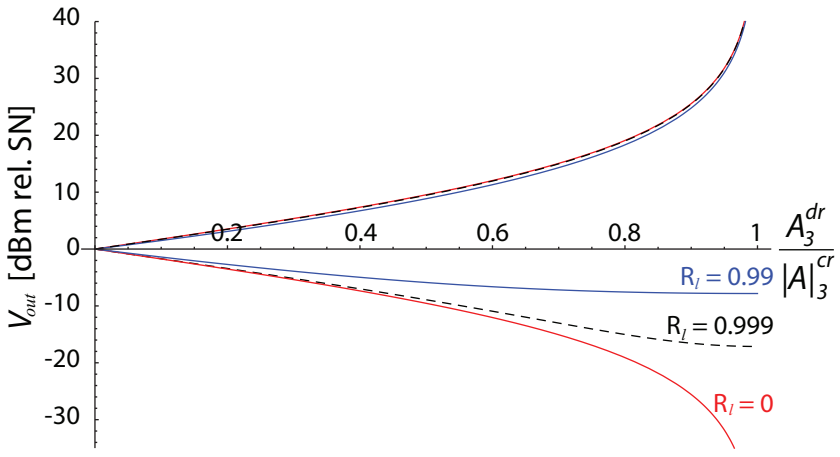


Figure 5.5: The variance of the squeezing and anti-squeezing at the output of an OPO as the system is driven towards threshold power, A_3^{crit}. The variance is shown for the same cavity with varying internal losses of; 0%, shown by the red lines, 0.1%, shown by the dotted black lines, and 1%, indicated by the blue lines. The cavity parameters are; cavity length, $L = 0.3\,\mathrm{m}$, output coupler reflectivity, $R_1^{out} = 0.95$ and input coupler reflectivity, $R_1^{out} = 1$. By setting $R_1^{out} = 1$ we have assumed that the OPO is single-ended for the fundamental field. Note that the losses and coupling reflectivities of the pump are irrelevant because we have normalised to threshold.

Figure 5.5 illustrates a number of important features of the OPO. Firstly, the squeezing gets greater the closer the OPO is to threshold. However, it is important to ensure that

the system does not drift above threshold as the system enters a new mode of operation. Second, intra-cavity loss affects the squeezing much more than it affects the anti-squeezing. This is simply due to the fact that the vacuum noise is relatively large compared to the noise of the squeezing but relatively small compared to the noise in the anti-squeezed quadrature. Mixing the state with vacuum therefore results in greater degradation of the squeezing than it does for the anti-squeezing. It is evident that the measured anti-squeezing is thus a relatively good measure of the proximity to threshold of an OPO system.

5.4.1 Escape Efficiency

The escape efficiency, η_{esc}, of an OPO is defined as the ratio of the decay rate of the squeezed field through the output coupler over the total decay rate of the cavity at the fundamental frequency

$$\eta_{esc} = \frac{\kappa_1^{out}}{\kappa_1}. \tag{5.43}$$

The escape efficiency of an OPO limits the maximum amount of squeezing that the system is capable of producing. It is essentially a measure of how many times the intra-cavity field will interact with the loss sources inside the cavity before it escapes through the output coupler. The number of round trips for the squeezed state can be quite high, therefore the intra-cavity field samples this loss many times. The effect of the escape efficiency is shown in Figure 5.6.

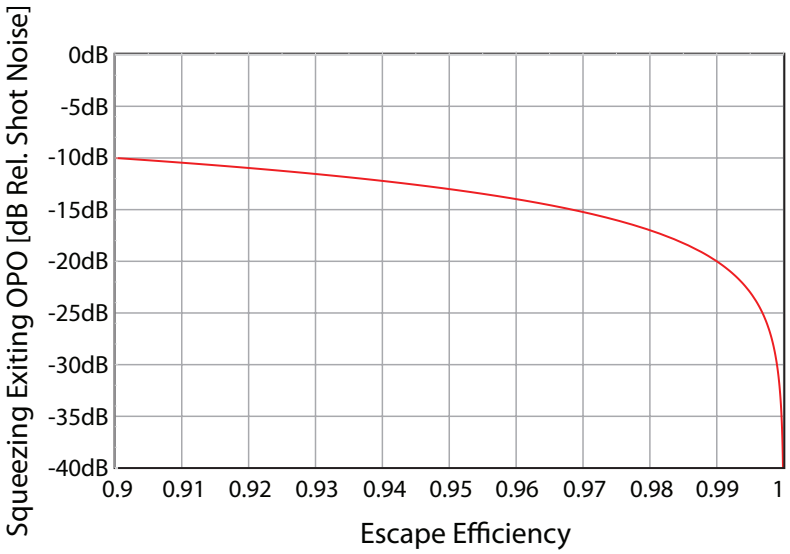


Figure 5.6: The effect of the escape efficiency on the maximum amount of squeezing that can be produced from any cavity.

There are two ways to improve the escape efficiency, one can reduce the intra-cavity losses, or one can reduce the reflectivity of the input/output coupler at the fundamental wavelength. From a design perspective, reducing the intra-cavity losses is achieved by reducing the number of surfaces that the intra-cavity field interacts with. In particular, anti-reflection (AR) coatings, tend to be relatively lossy. In this regard a monolithic cavity,

with no internal AR surfaces and one HR coating, is better than a hemilithic resonator, with two AR surfaces and one HR coating per round trip, which is better than a bow-tie cavity, with two AR surfaces and three HR coatings per round trip. However, the major source of loss is often absorption from the non-linear medium. This is why the various cavity designs can currently perform at similar levels. The best AR coatings can have a reflectivity less than 0.01% and the best HR coatings can have a reflectivity greater than 99.999%. Typical non-linear crystals currently have an absorption of approximately 0.1 % per centimetre (See Table 6.2). If advancements in crystal growing techniques result in reduced absorption, then loss from mirror coatings will become the limitation to improving the escape efficiency.

Reducing the reflectivity of the input coupling mirror in order to increase the escape efficiency unfortunately also changes other cavity parameters. Reducing the reflectivity of the input coupler will reduce the finesse of the cavity, which will increase the the linewidth and, more importantly, will also increase the threshold power of the system (this effect is later discussed in § 6.2.2). At some point it will not be possible to provide enough pump power to reach threshold. It has previously shown that the greatest levels of squeezing occur near to threshold and as such it is a requirement that we can produce an adequate amount of pump power. A solution to this problem is to move to a doubly resonant system. This enhances the pump field, reducing the required amount of driving pump power required to reach threshold. This will be explained in detail in § 6.2.1. However, large circulating powers may damage the crystal [117, 118, 119] and can lead to photothermal instabilities [60]. Squeezers typically operate in a regime where the squeezing is maximized whilst the stability of the OPO is kept at a level that is suitable for the application.

The escape efficiency is such an important parameter that it is worthwhile rearranging the squeezed field variances, Equations 5.41 and 5.42, to highlight the effect of the escape efficiency. Doing this results in

$$V_{out}^{+} = 1 + \eta_{esc}\frac{4g\alpha_3}{\kappa_1(1-\frac{g\alpha_3}{\kappa_1})^2+\frac{\omega^2}{\kappa_1}} \tag{5.44}$$

$$V_{out}^{-} = 1 - \eta_{esc}\frac{4g\alpha_3}{\kappa_1(1+\frac{g\alpha_3}{\kappa_1})^2+\frac{\omega^2}{\kappa_1}}. \tag{5.45}$$

Equations 5.44 and 5.45 illuminates the fact that the escape efficiency can be regarded as a beamsplitter operation directly after the OPO. It is often useful to consider the OPO as having produced a pure state which then passes through a beamsplitter of power transmittivity η_{esc}. The escape efficiency is then treated as another loss source on the squeezed state before detection. By treating it this way, it becomes evident that the escape efficiency completely determines the maximum amount of squeezing that a cavity system can produce.

5.5 The Classical Degenerate Optical Parametric Amplifier

Using Equations 5.32 we can also determine the behaviour of the system when operated as a degenerate optical parametric amplifier (DOPA or simply OPA). This is the case when a bright driving field is present at the fundamental frequency. The presence of a bright seed, which may temporarily be injected into an OPO setup, allows one to determine the threshold power of the system through investigation of the amplification and de-

amplification of the seed field. The equations describing the classical, degenerate OPA are written

$$\begin{aligned}\dot{\alpha}_1 &= -\kappa_1\alpha_1 + g\alpha_3\alpha_1^* + \sqrt{2\kappa_1^{in}}A_1^{dr} \\ \dot{\alpha}_3 &= -\kappa_3\alpha_3 - g\alpha_1\alpha_1 + \sqrt{2\kappa_3^{in}}A_3^{dr}.\end{aligned} \tag{5.46}$$

At this point we combine the non-linear gain term with the intra-cavity pump field and write $g\alpha_3 = G$. This is done under the assumption that the seed field is much smaller in magnitude than the pump field, $A_{dr}^1 << A_{dr}^3$, such that the pump field is not depleted by the parametric process to some approximation. This results in

$$\dot{\alpha}_1 = -\kappa_1\alpha_1 + G\alpha_1 + \sqrt{2\kappa_1^{in}}A_1^{dr}. \tag{5.47}$$

We are looking for the steady state solution, and as such set $\dot{\hat{a}}_1 = 0$ and solve for the intra-cavity field, α, resulting in

$$\alpha = \frac{\sqrt{2\kappa_1^{in}}\,(1 + G/\kappa_1)}{\kappa_1(1 - |G|^2/\kappa_1^2}A_1^{dr}. \tag{5.48}$$

This equation holds only for below threshold operation of the system. We now use the input/output relations given by Equation 5.9 to determine the output field, which is written

$$A_1^{out} = \frac{2\sqrt{\kappa_1^{in}\kappa_1^{out}}\,(1 + G/\kappa_1)}{\kappa_1\left(1 - |G|^2/\kappa_1^2\right)}A_1^{dr}. \tag{5.49}$$

The ratio of output power without the non-linear interaction to the output power with the non-linear interaction can then be calculated in order to determine the parametric gain of the system. Calculating the ratio we find

$$\begin{aligned}\frac{\mathrm{P}_1^{out}}{\mathrm{P}_1^{out}|_{G=0}} &= \frac{A_1^{out*}A_1^{out}}{A_1^{out*}|_{G=0}A_1^{out}|_{G=0}} \\ &= \frac{(1 + G/\kappa_1)^2}{\kappa_1\left(1 - |G|^2/\kappa_1^2\right)}.\end{aligned} \tag{5.50}$$

Noting that the steady state intra-cavity second harmonic field, $\alpha_3 = \frac{\sqrt{2\kappa_3^{in}}}{\kappa_3}A_3^{dr}$, divided by the intra-cavity threshold value, $\alpha_3^{crit} = \frac{\kappa_1\kappa_3}{g\sqrt{2\kappa_3^{in}}}$, results in one of the terms in this expression

$$\frac{\alpha_3}{\alpha_3^{crit}} = \frac{\kappa_1}{G}, \tag{5.51}$$

allows us to rearrange the non-linear gain into an expression written in terms of the external driving fields

$$\frac{\mathrm{P}_1^{out}}{\mathrm{P}_1^{out}|_{G=0}} = \frac{\left(1 + A_3^{dr}/|A_3^{crit}|\right)^2}{\left(1 - |A_3^{dr}|/A_3^{crit}|^2\right)}. \tag{5.52}$$

This expression will later be used to determine the threshold of our cavity systems. By changing the phase of the pump field by 180 degrees, the pump field term, A_3^{dr}, becomes negative and the system moves from amplification of the seed field to de-amplification. This effect is illustrated in Figure 5.7.

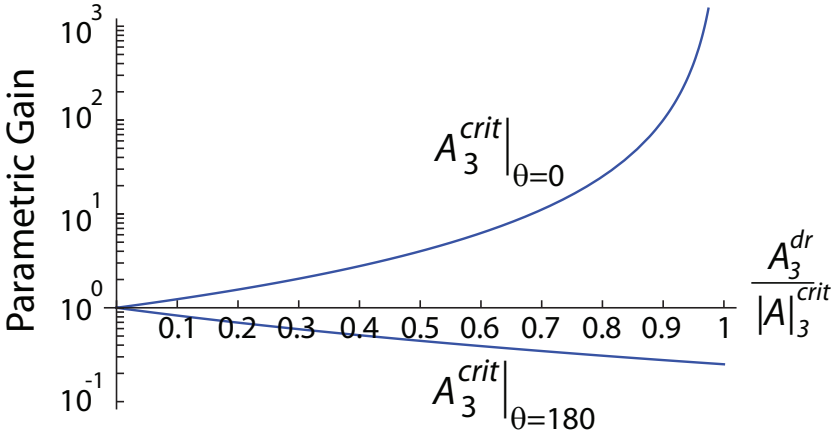


Figure 5.7: Non-linear gain of an OPA system as the input pump power, A_3^{dr}, is driven to threshold, A_3^{crit}.

Chapter 6

The Doubly Resonant Bow-Tie Optical Parametric Oscillator

In this chapter we introduce the details of the doubly resonant bow-tie OPO used to produce the results in this thesis and explain the decisions for arriving at the final design. We investigate the advantages and disadvantages of the various cavity geometries, non-linear media and cavity optical properties. The work in this chapter was undertaken by me in order to determine the optimal cavity design for the LIGO application.

6.1 Cavity Geometry

The first design consideration that we will investigate is the cavity geometry. The cavity geometry affects a number of things; the cavity mode shape, the number of HR and AR surfaces, the number of mirrors available for input and output coupling of external fields and whether the forward and reverse directions in the cavity are degenerate or not. The various cavity geometries are illustrated in Figure 6.1 for clarity.

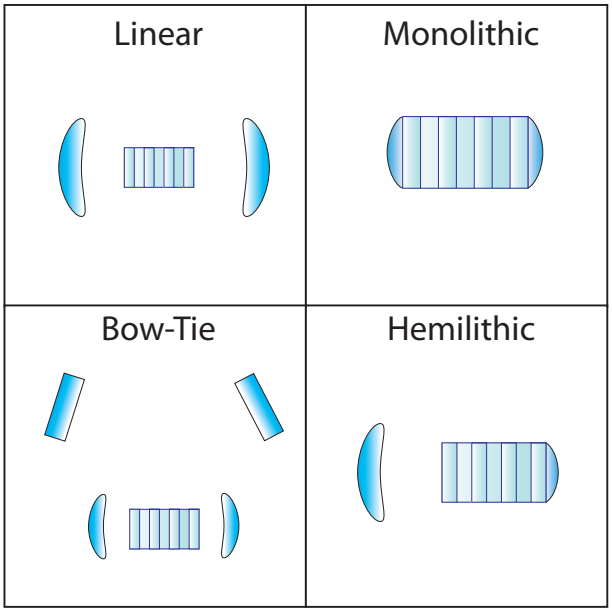


Figure 6.1: The various cavity geometries; linear, monolithic, bow-tie and hemilithic.

6.1.1 Travelling Wave vs. Standing Wave

The first decision we will investigate is whether to choose a standing wave or a travelling wave, ring cavity design. In a travelling wave design the forward and reverse directions of beam propagation exit the cavity at a different angle and hence are non-degenerate. In the standing wave case, light travelling in both the forward and reverse directions occupy the same mode when exiting the cavity. We can see from Figure 6.1 that the bow-tie cavity is a travelling wave cavity and that the three other cavities are standing wave designs.

The standing wave design has three major advantages. The first is the reduced number of surfaces per round trip. As discussed in § 5.6, reducing the number of HR and AR surfaces will reduce the intra-cavity loss and thus increase the escape efficiency. Thus, in theory standing wave cavities, particularly monolithic and hemilithic designs, are capable of producing higher levels of squeezing. However, this is only technically true provided that the loss due to the non-linear medium does not dominate the total intra-cavity loss.

The second advantage to standing wave cavities is that it is easier to make them mechanically stable due to the smaller number of optical elements and a fixed cavity geometry. An example of this is presented by Vahlbruch *et al.* [120], where a very solid modular hemilithic cavity is introduced. By design, the cavity is isolated from both air currents and thermal fluctuations. This aids in both the long term and short term stability of the cavity.

Finally, astigmatism is generally not an issue in these standing wave cavities. Due to the geometry of these cavities, there are no reflections from any curved mirrors which occur at an angle, hence astigmatism does not occur. Astigmatism can make it more difficult to match the mode of the squeezed light to the mode of the interferometer, an issue that should theoretically be relatively easy to overcome, but can also reduce the non-linear coupling between the intra-cavity fields in extreme cases.

The advantages of the bow-tie cavity are flexibility and isolation to backscatter. Access to an increased number of mirrors leaves more locations for the input and output of optical fields, providing flexibility in locking. Also, if a crystal degrades over time then a replacement crystal of the same material can be put in its place in less than a few hours without altering any of the mode matching. If a new material is desired in the setup then the separation of the curved mirrors can be altered to account for changes in refractive index of the non-linear medium and it is theoretically possible to return to a nearly identical cavity geometry.

The second advantage, and the deciding factor in choosing a bow-tie design for the work in this thesis, is the isolation to backscattered light that the travelling wave geometry affords, due to the ability of the cavity to separate the forward and reverse travelling modes. The effect is illustrated in Figure 6.2.

Figure 6.2 shows that scattered light from the interferometer may reflect at the Faraday isolator and thus be directed towards the squeezer. If this light is able to re-enter the interferometer, it will adversely impact on the sensitivity of the device. If the travelling wave cavity in the figure were to be replaced with a standing wave cavity, then virtually all of the scattered light that travels towards the squeezer will reflect off the linear cavity and re-enter the interferometer. This light will pick up random phase shifts as it travels towards the OPO, reflects off it, and travels back into the interferometer. These random phase shifts on the light are detected as a spurious signal. Additionally, this light will seed the OPO, making it operate as an OPA, which will couple pump noise into the squeezed field and degrade the low-frequency squeezing, as discussed in Appendix 1.

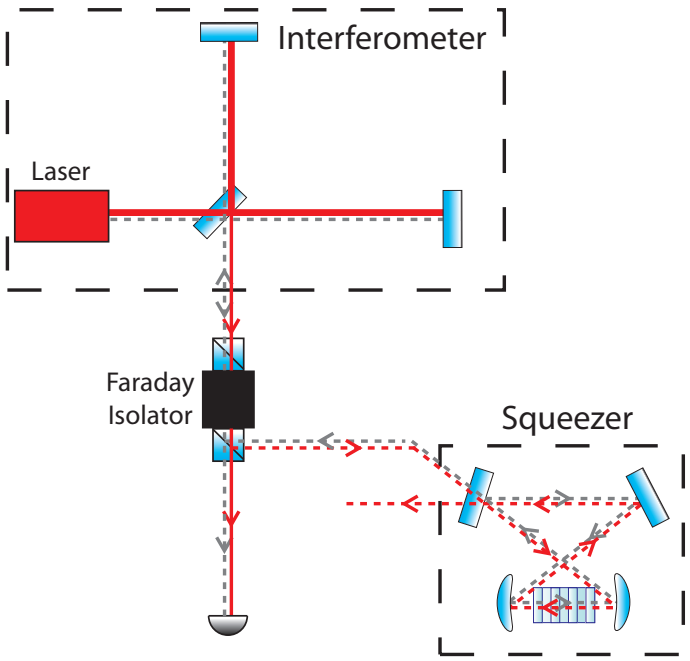


Figure 6.2: Backscattered light from the interferometer, dashed red trace, travelling towards the squeezer. The squeezed state, injected into the interferometer through a Faraday isolator in the signal path, is marked in dashed grey whilst the interferometer laser is marked in red. In the absence of scattering from within the cavity, all of the backscattered light from the interferometer does not re-enter the interferometer.

In order to remove this effect it is necessary to place a number of Faraday isolators between the interferometer and the squeezer. Unfortunately, this adds loss to the squeezed state, reducing the amount of squeezing entering the interferometer. A Faraday isolator can be expected to have an isolation factor of around 40 dB and a loss of anywhere from around 3 to 5 percent. With high levels of squeezing, this loss is substantial, particularly when multiple Faraday isolators are required.

Figure 6.2 shows that with the travelling wave design, the backscattered light from the interferometer will exit the cavity at some angle that does not direct it back towards the interferometer. The scattered light can then be dumped in order to ensure that it does not scatter again and end up being redirected towards the interferometer. Also, light that has scattered at the interferometer and enters the cavity is traversing the cavity in the opposite direction to the pump field and will not seed the OPO process.

However, it is useful to consider the experimental amount of backscattered light rejection that the travelling wave cavity affords. Scatter within the OPO may cause the backscattered light from the interferometer to enter the forward travelling mode of the cavity. If this occurs, then this light will seed the OPO and will be directed back towards the interferometer. The technical limitation to the isolation of forward (direction of propagation of the pump) and reverse modes is scatter off the intra-cavity surfaces, the mirrors and in particular the AR coatings on the crystal. The crystal is very close to normal to the beam direction and near a waist, light that reflects off this surface has a good chance of entering the cavity mode in the forward direction. The amount of isolation between the forward and reverse travelling cavity modes for the ANU squeezer has been measured at

41 dB [102]. This is nearly identical to the isolation offered by a good Faraday isolator. By using a travelling wave design, the loss outside of the cavity can therefore be reduced by an amount equal to the loss introduced by a single Faraday isolator. Different crystal geometries and super-polished cavity mirrors should increase the isolation offered.

6.1.2 Choosing a Bow-Tie Geometry

Due to the advantages detailed previously, it was decided that a bow-tie cavity was the preferred geometry. Using the standard treatment of ABCD matrices we can determine the spatial properties of the intra-cavity field for one of these cavities such that suitable cavity lengths can be found [114]. For Gaussian beams we can fully describe the beam by determining the complex beam parameter, $q(z)$, at some location z relative to the position of the beam waist z_0 as

$$q(z) = z + z_R i = \left(\frac{1}{R(z)} - \frac{i\lambda}{\pi n w(z)^2} \right)^{-1}, \tag{6.1}$$

where λ is the wavelength of the beam, $R(z)$ is the radius of curvature of the phase front, w_0 is the beam waist, n is the refractive index, $w(z)$ is the beam radius at position z and z_R is the Rayleigh range defined by

$$z_R = \frac{\pi w_0^2}{\lambda}. \tag{6.2}$$

The ABCD matrix method allows us to determine a relationship between an input field, and an output field that has undergone a system of optical elements. We can write

$$\begin{pmatrix} q_i \\ 1 \end{pmatrix} = k \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} q_o \\ 1 \end{pmatrix} \tag{6.3}$$

for the input and output beam parameters, q_i and q_o, respectively, passing through an optical system with some ray transfer matrix. The k term is merely a normalisation factor to aid in manipulation of the equations. Equation 6.3 allows us to write for the input and output fields

$$q_o = \frac{Aq_i + B}{Cq_i + D}. \tag{6.4}$$

In order for the cavity to resonate, we require our beam at some location to match the shape of the beam after one round trip. This is equivalent to assuming that the beam parameter at some location, the input, is equal to the beam parameter after traversing the cavity once, the output field. Setting this requirement results in a quadratic equation for the complex beam parameter that can be solved in order to determine the complex beam parameter that matches the cavity mode.

From the complex beam parameter, one can then determine any property of the beam. The beam radius at position z is given by

$$w(z) = \sqrt{\frac{\lambda \, \mathrm{Im}[\mathrm{q_o}]}{\pi}}, \tag{6.5}$$

where $\mathrm{Im}[q_o]$ is the imaginary part of the complex parameter for the cavity mode.

The ABCD matrices for propagation of distance d through a medium of refractive index

Propagation	**Interface**	**Tangential Reflection**	**Sagittal Reflection**
$M_d = \begin{pmatrix} 1 & \frac{d}{n} \\ 0 & 1 \end{pmatrix}$	$M_i = \begin{pmatrix} 1 & 0 \\ 0 & \frac{n_1}{n_2} \end{pmatrix}$	$M_{rt} = \begin{pmatrix} 1 & d \\ \frac{-R}{2\mathrm{Cos}\theta} & 1 \end{pmatrix}$	$M_{rs} = \begin{pmatrix} 1 & d \\ \frac{-R\mathrm{Cos}\theta}{2} & 1 \end{pmatrix}$

Table 6.1: ABCD matrices for various propagation events.

n, transmission through an interface from a refractive index of n_1 into a material with a refractive index of n_2, and reflection off a spherical mirror with a radius of R are shown in Table 6.1. The reflection off the concave mirror is shown for both the tangential (in the plane of incidence between the ray and the mirror) plane and the orthogonal sagittal plane for a mirror that has an angle of θ relative to the beam axis and the normal direction.

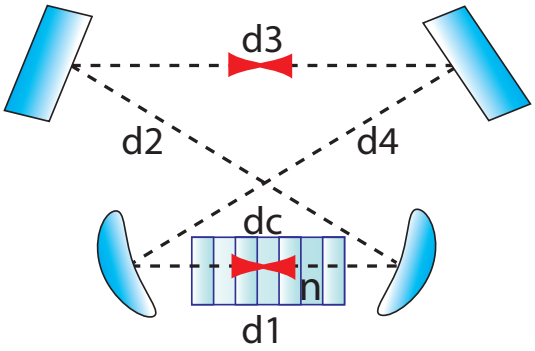


Figure 6.3: The geometry of the bow-tie cavity with the distances involved and the location of the two waists illustrated. The distance through the crystal, of refractive index n, is marked dc whilst d1 is the distance between the two curved mirrors.

The layout of the bowtie cavity is shown in Figure 6.3. The cavity consists of two curved mirrors, focussing the light to a small waist inside the non-linear crystal, and two flat mirrors. The term dc denotes the distance through the crystal whereas the remaining dx terms are used to represent the various distances from mirror to mirror in the cavity, whilst n is the refractive index of the non-linear medium. Reflection off a plane mirror is simply equal to unity. A full round-trip of the cavity for the tangential axis, beginning and ending mid-way through the non-linear medium, can thus be written

$$M = M_{dc/2} M_{i(1\to n)} M_{(d1-dc)/2} M_{rt} M_{d4} M_{d3} M_{d2} M_{rt} M_{(d1-dc)/2} M_{i(n\to 1)} M_{dc/2}, \tag{6.6}$$

where it is important to note that the refractive index as the light travels through the crystal, $M_{dc/2}$, must be accounted for and that the two interface terms vary depending on whether the material is exiting or entering the crystal. The same can be written for the sagittal plane, replacing all instances of M_{rt} with M_{rs}. The distances d2 and d4 are equal and can be determined from the angle of the bow-tie

$$d2 = d4 = \frac{d1 + d3}{2\mathrm{Cos}\theta}. \tag{6.7}$$

In order to solve this problem, we need to fix some of these parameters. The method

chosen was to set the distance between the two plane mirrors (90 millimetres), the angle of the bow-tie (twelve degrees), the radius of curvature of the curved mirrors (-38 millimetres) and then to vary the distance between the two curved mirrors. The angle of the cavity is chosen to be as small as possible in order to reduce astigmatism but not so small as to make clipping of the beam on mirror mounts an issue. The distance between the two curved mirrors was chosen after varying this value a number of times and finding a region that resulted in a suitable waist size (The waist size can be optimised using the Boyd-Kleinman method [121]). The radius of curvature of the curved mirrors was chosen to provide the necessary focussing to achieve the desired waist size without a large separation between the mirrors. A small radius of curvature is favoured such that the waist size within the crystal is small with a reasonable separation between the two mirrors.

After setting these values we can then plot the radii of the two waists of the bow-tie cavity. One waist is situated between the flat mirrors and the other between the curved mirrors. In order to determine the size of the waist outside the crystal we need only propagate the complex beam parameter through Equation 6.4 with a new transfer matrix that only transfers the beam the required distance. We can write for the propagation matrix, M2, from the crystal waist to the free-space waist

$$M2 = M_{d3/2} M_{d2} M_{rt} M_{(d1-dc)/2} M_{i(n\to 1)} M_{dc/2}, \tag{6.8}$$

where we note again that the transfer matrices have been applied in reverse order to the order of propagation of the beam.

Figure 6.4 shows the resulting waist sizes. It is evident that astigmatism within the non-linear medium is near zero for some separation between the curved mirrors. There is some astigmatism at the larger waist, situated outside the crystal, and this means that there will be astigmatism on the field exiting the cavity. This is because the portion of light that transmits through the coupling mirror has the same mode shape as the cavity mode and will therefore have an identical waist at the same distance from the mirror, albeit outside the cavity. It is important to remember that the squeezing will have this shape upon exiting the cavity when attempting to mode match the squeezed light to other fields or cavities.

Figure 6.4 shows the waist size for the cavity at the fundamental wavelength of 1064 nm. Due to the fact that we are using a quasi-phase matched system, the refractive indices for the fundamental and the pump fields are different. By changing the wavelength and the refractive index, we can repeat the process followed for the fundamental field to determine the cavity mode for the pump. The result is shown in Figure 6.5

6.2 Cavity Optical Design Considerations

Now that a cavity geometry is decided, the next step is to determine the properties of the mirror coatings. The coatings of the mirrors will dictate the loss dynamics of the optical field within the cavity, defining the escape efficiency, linewidths, finesse and threshold power for the system.

6.2.1 Doubly Resonant System

A choice must be made as to whether the system is operated in a doubly resonant configuration, one in which both the pump and fundamental fields resonate, or in a singly

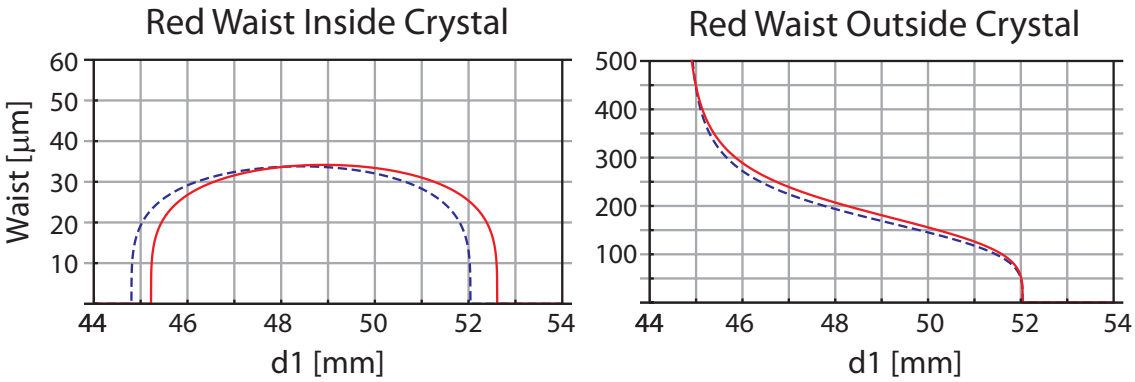


Figure 6.4: The tangential (dashed blue) and saggital (solid red) beam radii of the two fundamental waists in the bow-tie cavity as the distance between the curved mirrors, d1, is varied. The length of the crystal, dc, is 10.5 mm with a refractive index, n, of 1.830 [122]. The distance, d3, is set to 90 mm, the bow-tie cavity has an angle of 12 degrees and the curved mirrors have a radius of curvature of -38 mm.

resonant configuration, where only the fundamental field resonates. The advantages of a doubly resonant system are that the threshold pump power is reduced, the resonant green field provides an additional control field and the optimal spatial overlap between the fundamental and pump fields is ensured.

The optimal overlap between the pump field with waist w_p and fundamental field with waist w_f is given by the relation [121]

$$w_p = \sqrt{2} w_f. \tag{6.9}$$

Investigation of Equation 6.5 reveals that the cavity ensures that, on resonance, this condition is met for the two fields when using a phase matched material. This is not strictly the case for quasi-phase matched materials as will be shown in § 6.3.6. However, Figure 6.5 shows that the slight difference in refractive indices between the fundamental and pump fields has very little impact on the fundamental and pump field modes within the cavity. Figures 6.5 and 6.4 show that, with the correct separation between the curved mirrors, the waist sizes are very well matched, which is to say that Equation 6.9 holds.

As for cavity locking. The fact that the pump field is resonating allows for the pump field to be used as a reference for locking schemes, such as a Pound-Drever-Hall technique. Additionally, the pump field will likely have a different finesse to the fundamental field and may facilitate different modulation frequencies for such a lock.

Finally, the doubly resonant system is a means of reducing the threshold power. The threshold power of the cavity is given by Equation 5.36, but is written again for reference

$$P_3^{crit}[W] = \left(\frac{\kappa_1 \kappa_3}{g\sqrt{2k_3^{in}}} \right)^2 \frac{hc}{\lambda_3}. \tag{6.10}$$

By resonating the second harmonic field, we are decreasing the rate at which photons decay from the cavity, thus decreasing the value of κ_3^{in}. Expanding $\kappa_3 = \kappa_3^{in} + \kappa_3^{out} + \kappa_3^{l}$ in Equation 5.36 will result in a term proportional to $\frac{\kappa_3^{in}}{\sqrt{\kappa_3^{in}}}$, which reduces as the decay

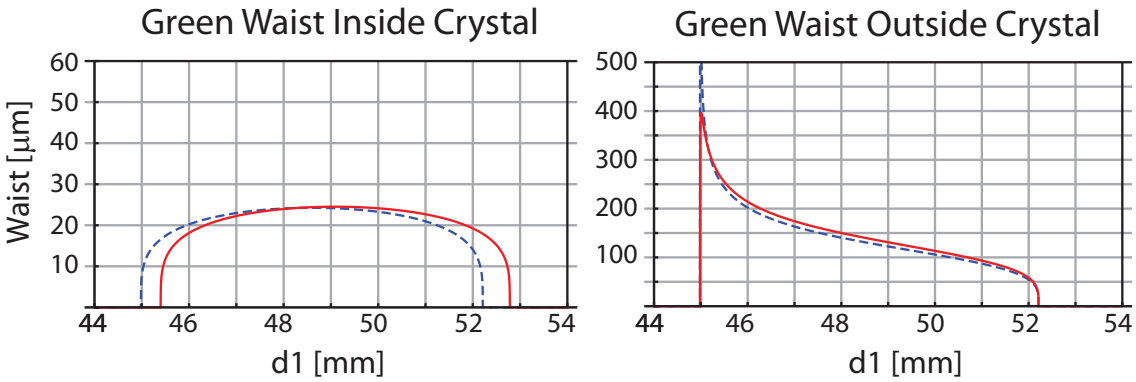


Figure 6.5: The tangential (dashed blue) and saggital (solid red) beam radii of the two pump waists in the bow-tie cavity as the distance between the curved mirrors, d1, is varied. The length of the crystal, dc, is 10.5 mm with a refractive index, n, of 1.830 [123]. The distance, d3, is set to 90 mm, the bow-tie cavity has an angle of 12 degrees and the curved mirrors have a radius of curvature of -38 mm.

rate of the input coupler reduces, thus decreasing threshold power.

The advantage to the singly resonant system is simplicity, and there are alternative solutions to many of the issues that the doubly resonant system offers. Techniques can be employed in singly resonant systems to ensure that the spatial overlap is close to optimal [97]. The threshold power can be lowered by reducing the intra-cavity loss or increasing the reflectivity of the fundamental output coupling mirror, rather than increasing the resonant enhancement of the pump. The phase matching bandwidth, the temperature over which the non-linear interaction is optimised, is also typically larger in the singly resonant case [108]. It is reduced in the doubly resonant system because a temperature shift will result in a differential phase shift on both fields. The cavity length is typically locked through deriving an error signal from one of these intra-cavity fields and hence the cavity will only track the resonance condition of that field. Finally, in the singly resonant system, no techniques are required to ensure that the pump field and fundamental field are co-resonant. Co-resonance is an issue because intra-cavity dispersion shifts the phase relationship between the two fields as they travel around the cavity.

Intra-Cavity Dispersion

The resonance condition of the cavity requires that the phase of a particular field at some location in the cavity be equal to the phase after the field has undergone a complete round-trip and returned to the same location. If the phase relationship between the two fields varies after one round trip then the two fields will now resonate at different cavity lengths. This effect is shown in Figure 6.6

We see that at location 5 the two fields now have a very different phase relationship to that when they started their journey, at location 1. A control loop designed to lock the pump field, for example, will now change the cavity length such that the phase of pump field at location 5 matches the initial phase. However, due to the changed relationship between the two fields, the fundamental field will no longer resonate. If this field is not resonant then many effects arise, but the final result is that squeezing is degraded. With this in mind, it is helpful to resonate the pump field by only a moderate amount, such

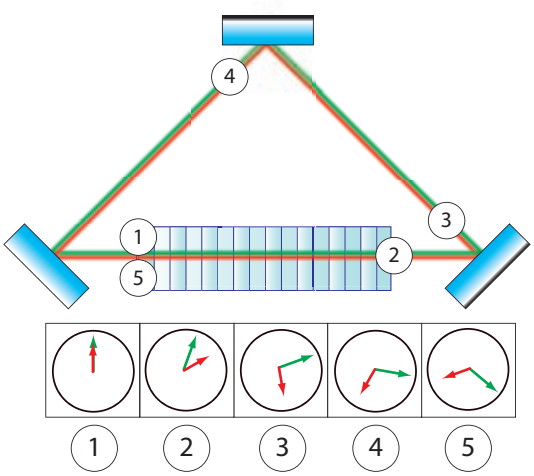


Figure 6.6: Cavity dispersion as the pump and fundamental fields traverse one round-trip of the cavity. As the fields complete one round-trip, they pick up phase shifts from each mirror surface and from the nonlinear medium. As they complete one round trip, from location 1 to location 5, the phase relationship between the two fields has changed.

that the pump field linewidth is much broader than the fundamental field linewidth, such that co-resonance is more easily achieved. By doing this, the sensitivity of the cavity to temperature and length fluctuations is reduced.

Historically, cavity dispersion between the two cavity modes has been overcome by detuning the temperature of the non-linear medium. This allows for dispersion within the material to compensate for dispersion from the various sources of dispersion. In a monolithic geometry, varying the temperature is the only method of overcoming intra-cavity dispersion. This technique is not ideal, as it forces the system away from the ideal phase matching temperature, reducing the efficiency of the non-linear process. The magnitude of this effect is dependant upon the FWHM of the non-linear conversion process, which varies significantly between materials [108].

A simple method for overcoming intra-cavity dispersion is to use a dispersion plate [124]. The dispersion plate is simply a "window" of dispersive material that has been anti-reflection coated at both wavelengths in order to minimise losses. The angle of incidence between this plate and the two fields, inside the cavity, is varied in order to change the path that the two fields travel through the material. By tuning the angle of the plate, a position can be found where the dispersion introduced by the plate compensates for that introduced by the crystal and the cavity mirrors. The disadvantage of this method is the extra loss introduced, decreasing the escape efficiency and hence the squeezing produced. Such a plate might will also likely reduce the forward to reverse direction rejection ratio of the bow-tie cavity.

An alternative method of overcoming intra-cavity dispersion, wedged quasi-phase matched crystals, is described in detail in 6.3.6. This method does not introduce additional loss or scattering into the OPO and allows the non-linear medium to operate at the optimum phase-matching temperature.

6.2.2 Escape Efficiency and the Doubly Resonant System

As mentioned previously, the bow-tie cavity has the disadvantage over other cavity designs of having a relatively larger number of HR and AR surfaces per round trip. This higher intra-cavity loss results in a decreased escape efficiency. The escape efficiency was introduced in § 5.6 and is defined

$$\eta_{esc} = \frac{\kappa_1^{out}}{\kappa_1}. \tag{6.11}$$

Seeing as the losses in the cavity cannot be easily tailored, the remaining means of increasing the escape efficiency is then to increase the decay rate of the output coupler by reducing the reflectivity of the output coupler for the squeezed field. This comes with the unwanted side-effect of increasing the threshold power, given by Equation 5.36. However, the threshold power depends upon the decay rates of both the fundamental and pump fields. By resonating the pump field, thereby decreasing the decay rate of the pump field, the threshold power can be tailored to suit the application. Figure 6.7 uses Equation 5.42 to illustrate the effect on the squeezing produced by the cavity as the reflectivity of the fundamental and pump fields are varied on a singly-ended cavity, $\kappa_1 = \kappa_1^{in/out} + \kappa_1^l$ and $\kappa_3 = \kappa_3^{in/out} + \kappa_3^l$. The sole coupling mirror has a power reflectivity for the pump of $R_3^{in/out}$ and a power reflectivity for the fundamental of $R_1^{in/out}$.

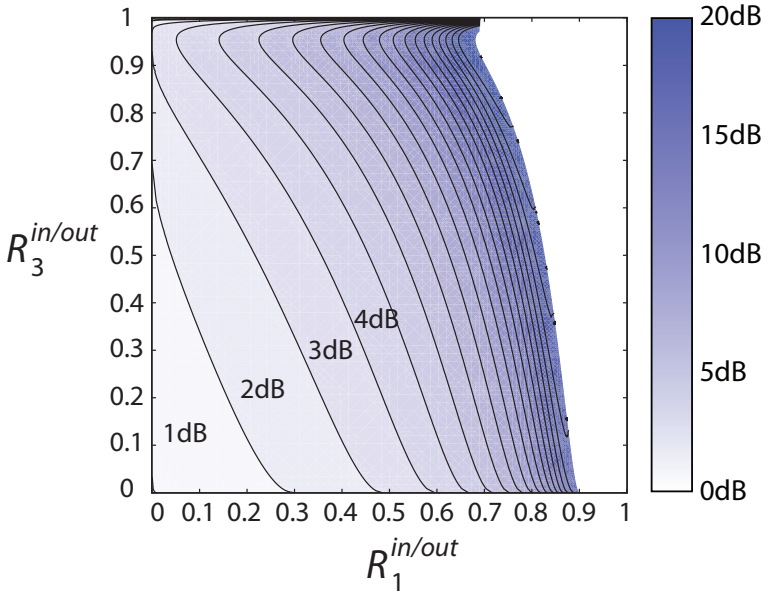


Figure 6.7: Contour plot of the vacuum squeezing produced, in dB relative to shot noise, from a singly-ended doubly resonant OPO cavity as the reflectivities of the input coupler at the fundamental, $R_1^{in/out}$, and pump, $R_3^{in/out}$, wavelengths are varied. The white region in the top right of the figure indicates the region where the system has moved above threshold. Darker regions indicate more squeezing, with each contour indicating an increase of the squeezing by 1 dB. Parameters are; cavity length $L = 0.273m$, non-linear interaction strength, $g = 1891$, intra-cavity loss at fundamental and pump, 0.26% and 4.6%, respectively, detection frequency, $\omega = 0$, and incident pump field power of 100 mW.

It is apparent that as the reflectivity of the coupling mirror at the fundamental is decreased and the escape efficiency increases, more squeezing becomes theoretically available. However, without changing the pump power, it is necessary to increase the reflectivity of the same mirror at the pump such that the incident power stays near to the threshold of the system. This can be seen by tracing the contour that separates the squeezing from the above threshold regime. This type of plot is particularly useful in situations where the maximum amount of pump power is known, as is often the case in experimental setups. The final thing to mention is that this figure does not indicate the amount of circulating pump field power. As the value of $R_3^{in/out}$ is increased, the circulating power also increases, by Equation 5.15. If the circulating power becomes too high then photothermal effects will reduce the stability of the system and may even lead to damage of the non-linear medium. This is explained in more detail in § 6.3.4.

It is important to understand the effect varying one property of the cavity system will have on the other properties. Improving one parameter often leads to the degradation of another. One of the most important and defining properties of the system is the reflectivity of the output coupler at the fundamental wavelength (or input/output coupler in the case of a singly-ended cavity). The interplay between the various parameters that depend upon the reflectivity of the input/output coupler at the fundamental wavelength for a singly ended cavity is shown in Figure 6.8. To find the right reflectivity is a balancing act between the various cavity parameters, particularly the escape efficiency and the threshold power. For example, lower threshold powers are generally preferred because of reduced photothermal effects. However, increasing $R_1^{in/out}$ to achieve this, decreases the escape efficiency and the linewidth, and increases the finesse.

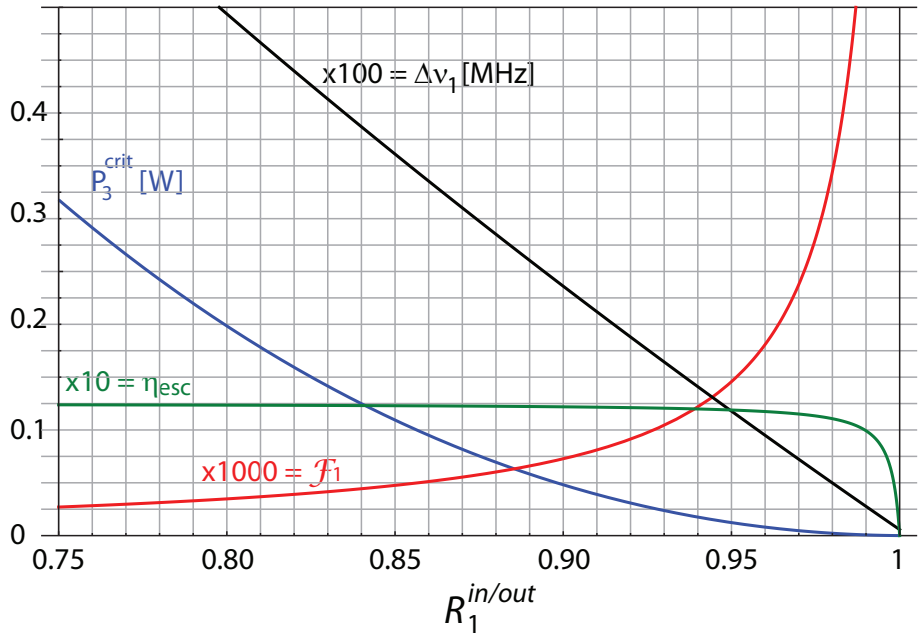


Figure 6.8: The fundamental finesse, $\mathcal{F}_1$, fundamental linewidth, $\Delta\nu_1$, escape efficiency, η_{esc}, and the threshold power P_3^{crit} of the cavity as the reflectivities of the input/output coupler for the fundamental, $R_1^{in/out}$, is varied. Parameters are; cavity length $L = 0.273m$, non-linear interaction strength, $g = 1891$, reflectivity of the input/output coupler at the pump wavelength, $R_3^{in/out} = 0.722$, and intra-cavity loss at fundamental and pump, 0.26% and 4.6%, respectively.

6.2.3 Linewidths and Finesse

The linewidth of the cavity at the fundamental wavelength was tailored specifically for the coherent locking field used. The cavity mirrors and length were chosen such that the linewidth was 25.9 MHz, placing the coherent locking field, which is detuned from the carrier frequency by 29.8 MHz, just outside the cavity linewidth. This will be explained in more detail in § 7.4 but the philosophy was to ensure that the coherent locking field inside the cavity was kept as small as possible. The linewidth of the cavity at the pump frequency was 71 MHz, chosen to be as broad as possible to ensure that the non-linear gain sensed by the coherent locking field was as large as possible. However, increasing the linewidth of the cavity at the pump frequency increases the necessary pump power to reach threshold, and the PDH lock controlling the cavity length on this field performs better with higher finesse, and hence smaller linewidth, cavities.

6.2.4 Threshold Power

As mentioned in the section on escape efficiency (§ 6.2.2), increasing the escape efficiency by reducing the reflectivity of the fundamental coupling mirror, $R_1^{in/out}$, results in a higher threshold power. This effect can be counteracted by increasing the reflectivity of this coupler at the pump frequency. Figure 6.9 has the same parameters as Figure 6.7, but this time shows the threshold power as the reflectivity at the fundamental and pump frequencies is varied.

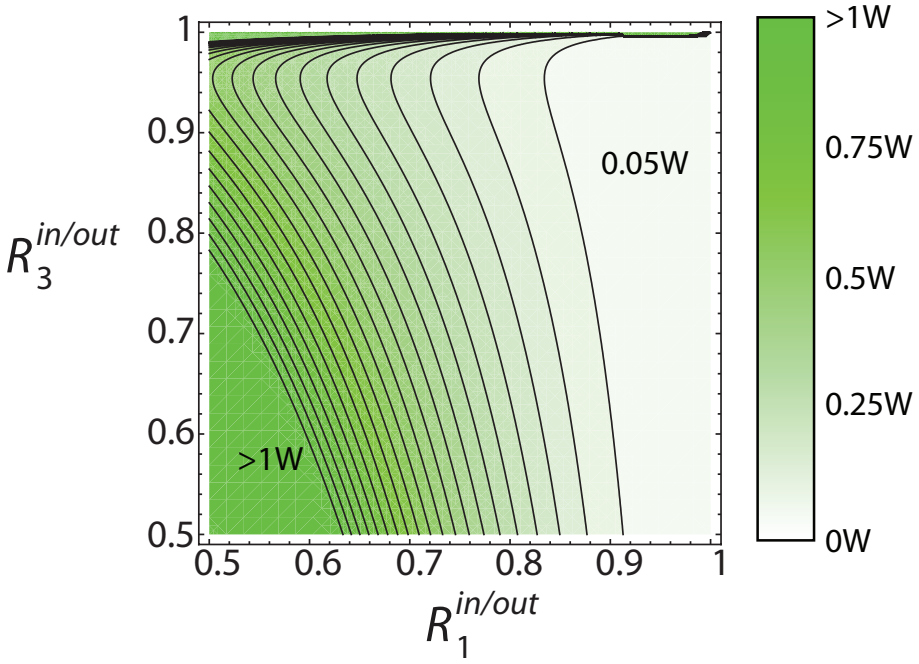


Figure 6.9: Contour plot of threshold power, P_3^{crit}[W], for a singly-ended doubly resonant OPO cavity as the reflectivities of the input coupler at the fundamental, $R_1^{in/out}$, and pump ,$R_3^{in/out}$, wavelengths are varied. Greener regions indicate a higher threshold, with each contour indicating an increase in the threshold power of 50 mW. Parameters are; cavity length $L = 0.273$m, non-linear interaction strength, $g = 1891$, intra-cavity loss at fundamental and pump, 0.26% and 4.6%, respectively.

It can be seen that the threshold power depends very strongly upon the reflectivity of the coupling mirror. As the reflectivity at the fundamental drops below a value of around 0.8, the threshold increases very quickly and the system becomes impractical. Increasing the reflectivity of this mirror at the second harmonic reduces the threshold value but will also increase the circulating power. An operating condition must be found where the escape efficiency and the threshold power are both reasonable.

6.3 Non-Linear Crystal Considerations

There are many choices of non-linear crystals available when building an OPO. The ideal material would possess high non-linear gain, low absorption and a high damage threshold. Below is a summary of the properties of some of the more commonly used periodically poled materials. Many of the absorption values are for materials that have not been periodically poled. A number of experiments have qualitatively found the absorption in periodically poled materials to be nearly identical to those without poling [109]. Multiplying the GRIIRA values by the absorption of the medium in the absence of green light results in the absorption value found in the presence of a 532 nm pump field.

Crystal	**n**	d_{eff} **[pm/V]**	α **[%/cm]**	**Dam. [W/cm^2]**	GRIIRA
PPMgO:LN	2.1494//2.2279	15.9	< 0.1 [125]	> 400 [126]	1 [127]
PPKTP	1.8296//1.8868	9.3	0.02 [109]	89k [128]	7 [109]
PPLT	2.1399//2.2078	8.8	0.1 [129]		1 [125]
PPKN	2.1201//2.2041	12.5	< 0.5 [130]		

Table 6.2: Crystal properties of some of the more commonly used periodically poled materials; the refractive indices, n, non-linear strength, d_{eff}, absorption, α, damage threshold, Dam., and green induced infra-red absorption, GRIIRA. Values for refractive indices and non-linear coefficients are from Shoji *et al.* [122]. Values which could not be found or verified have been left blank.

A periodically poled material was necessary for the cavity because a doubly resonant system was desired and dispersion compensation was to be achieved through the use of a wedged crystal. This is explained in detail in § 6.3.6.

6.3.1 Periodic Poling and Non-Linear Strength

Periodic poling was introduced in § 4.5.2. It is a means of achieving quasi-phase matching within a material. Periodically poled materials typically exhibit larger non-linear interaction strengths than their birefringent matched counterparts due to the fact that the poling can be designed for use with the optimal non-linear coefficient of the crystal. Birefringent phase matching [131], requires cutting the crystal at a particular angle such that the refractive indices of the fundamental and pump fields are equal at some temperature. The temperature at which birefringent phase matching occurs can also be rather inconvenient, as in the case of lithium niobate, whose phase matching temperature is 60 degrees Celsius or higher [108]. With periodic poling the quasi-phase matching temperature can be chosen by varying the domain size of the poling, choosing them to match the coherence length (defined in Equation 4.13) at a particular temperature. The limitation to this is when the domain periods required become too small to physically manufacture.

6.3.2 Absorption and Scattering

As discussed in § 5.6, the loss due to the crystal is often the limiting factor for the amount of squeezing that a system can produce. With this in mind, it is important to choose a non-linear material with as small an absorption coefficient as possible. Table 6.2 shows the expected absorption values for the various non-linear materials. These values are difficult to measure and as such there are limited papers detailing these properties. Quoted values for these materials can vary substantially, and is likely related to the specific growth conditions. This makes it hard to get a good estimate on the actual absorption one might expect. From experience, the loss can also vary greatly between different crystal growth batches. Local defects in the materials can have an enormous impact on the crystal loss. It is often necessary to translate individual crystals and search for the location where the intra-cavity loss is at a minimum.

6.3.3 Green Induced Infra-Red Absorption (GRIIRA)

Many non-linear materials exhibit a property known as GRIIRA. In the presence of green light, the absorption of these materials in the infra-red increases. The phenomenon has been studied in detail for many materials and is believed to be due to the formation of "colour centres" [127, 125, 132, 109]. The infrared absorption for some materials can increase by as much as a factor of 35 in the presence of green light [125] and as such it is worth understanding which materials exhibit this property and how it can be avoided.

Wang [109] meaured the GRIIRA response of KTP and PPKTP after illumination with a 532 nm pump intensity of 320 $\mathrm{MW/cm^2}$. In this material, the absorption of KTP was seen to increase by greater than a factor of 5 and the absorption of PPKTP was seen to increase by greater than a factor of 7. In the absence of GRIIRA the absorption of this material is that quoted in table 6.2, a value of 0.2 %/cm. GRIIRA in lithium niobate was a limiting factor until it became standard practice to dope the crystal with magnesium oxide. Furukawa *et al.* [127] showed that with sufficient doping, GRIIRA could essentially be eliminated from both congruent and stoichiometric growths of lithium niobate.

GRIIRA has been studied well enough that in most commonly used materials GRIIRA is no longer an issue under standard operating conditions. It is worth noting, however, that loss in PPMgO:LN may be less than in PPKTP under high circulating pump fields at 532 nm.

6.3.4 Damage Threshold and Photothermal Effects

The powers used in these experiments are typically very far from the damage threshold of the crystals. With a typical circulating power gain of 10 for the pump field in a doubly resonant system, a threshold power of 200 mW and a waist size of 30 μm, the intensity of the light hitting the non-linear medium is approximately 0.7 $\mathrm{kW/cm^2}$ at threshold. This is very close to the damage threshold of non-doped congruent lithium niobate, at 1 $\mathrm{kW/cm^2}$ [127], but most lithium niobate crystals are now doped in order to increase this value. Doped and/or stoichiometric growths can lead to lithium niobate damage thresholds greater than 8000 $\mathrm{kW/cm^2}$ [127]. For comparison, potassium titanyl phosphate is reported to have a damage threshold of 41 $\mathrm{MW/cm^2}$ [128]. Interestingly, periodic poling can also increase the damage threshold of both KTP and LN [128]. These power levels are generally not reached in OPO setups because photothermal effects, such as lensing and localised heating, from the relatively high absorption of the pump field in most of these

materials, will destroy the stability of the OPO well before the large laser powers needed to damage the crystal are ever reached.

6.3.5 Grey Tracking

Whilst the damage threshold of KTP is generally considered significantly higher than that for LN, it can exhibit a phenomenon known as grey tracking. Grey tracking, or photochromic damage, is not completely understood but it is an effect whereby the refractive index of the material and the loss changes after some time of exposure to bright radiation [118, 117, 119]. It has been shown that grey tracking only occurs in the presence of both fundamental and second harmonic fields [117] and that it also occurs in non-degenerate OPA systems. The effect can typically be removed through heating of the crystal, although long term effects of this damage have not been investigated. PPKTP is the only material from Table 6.2 that is known to exhibit this effect.

6.3.6 Wedged Crystal

In doubly resonant cavity systems, in which the pump and the fundamental field are both resonant, cavity dispersion must be overcome. Intra-cavity dispersion is introduced in detail in § 6.2.1. It was discussed that intra-cavity dispersion was historically overcome with the use of a dispersion plate but that this introduces additional intra-cavity losses.

An alternative solution, one that does not introduce additional intra-cavity loss, is to use a wedged quasi-phase matched crystals. This technique exploits the fact that quasi-phase matched materials inherently have different refractive indices for the fundamental and pump fields. At the phase matching temperature, the coherence length matches the poling length, compensating for dispersion. However, dispersion is still present. By leaving a wedged section at the end of the crystal that is free of poling, one can vary the phase relationship between the two fields as they exit the non-linear medium by shifting the crystal laterally to the path of the beams, thus varying the distance travelled through the dispersive medium. In this way, the intra-cavity dispersion can be compensated for such to ensure that the pump and the fundamental fields are co-resonant. The physical parameters of the wedged PPKTP crystals used throughout the work in this thesis are shown in Figure 6.10.

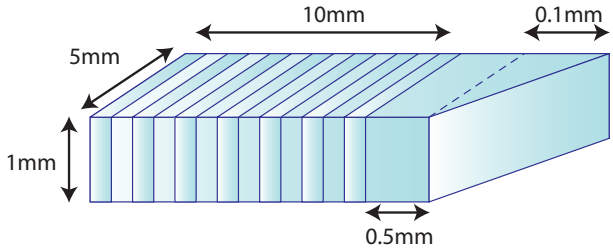


Figure 6.10: The periodically poled wedged crystal. The wedge is cut at an angle of approximately 1 degree.

One thing to note is that the wedge will cause the two fields to undergo different angles of refraction upon entering or exiting the medium due to the different refractive indices of the two fields. Using refractive indices for KTP given in Table 6.2 we can examine the

effect that the wedge will have on the propagation of the fields using Snell's law [114]. We choose to investigate only the effect on the pump field. The pump field is propagated through the non-linear medium and the cavity equations are then solved, taking into account the change in angle of incidence of the pump field on the first curved mirror it reflects off. Snell's law is written

$$n_1 \mathrm{Sin}\theta_1 = n_2 \mathrm{Sin}\theta_2, \tag{6.12}$$

where n_1 is the refractive index of air, n_2 is the refractive index of the medium, θ_1 is the angle between the normal of the mirror and the ray in free-space and θ_2 is the angle between the normal and the ray in the crystal. For the pump field we find, with a 1 degree wedge, that the refracted field has an angle of 0.53 degrees. If we then propagate this through the end of the crystal we find that the pump field exiting the crystal has an angle of 0.89 degrees from the normal. The situation is shown in Figure 6.11.

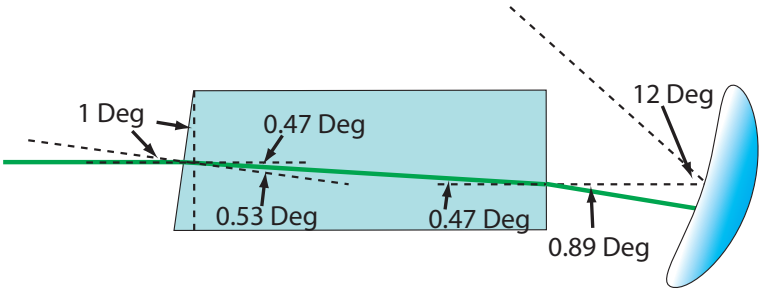


Figure 6.11: Propagation of the pump field through a PPKTP crystal with a 1 degree wedge.

To solve for the intra-cavity field then requires that the θ term describing the angle that the pump field hits the first curved mirror, matrix M_{rt} from Equation 6.6, have the additional angle from the wedge added to it before it. Without the wedge the intra-cavity pump field mode has a tangential waist of 23.81 μm that is 0.14 μm from the centre of the crystal. With the wedge, the pump mode has a tangential waist of 23.01 μm that is 0.55 μm from the centre of the crystal. The results for the sagittal waists are very similar. We therefore conclude that the wedge has only a minor effect on the intra-cavity modes.

Chapter 7

The ANU Squeezer

In this chapter the details of the squeezer that was constructed and operated at the ANU, the prototype for the LIGO squeezer, are discussed. The properties of this cavity are first introduced and the coherent locking scheme, used to control the angle of the squeezing, is modelled using the cavity equations of motion. A variant of the original coherent locking scheme, resulting in improved locking signals, is then investigated. The squeezing results from this system are presented and discussed in Chapter 8.

7.1 ANU Cavity Properties

An overview of many of the design decisions were addressed in Chapter 6 and included cavity geometry, choice of non-linear media, the escape efficiency and the physical distances between mirrors. With these in mind, the properties of the squeezer constructed at The Australian National University, henceforth referred to as the "ANU Squeezer" can be understood. A photo of the cavity is shown in Figure 7.1

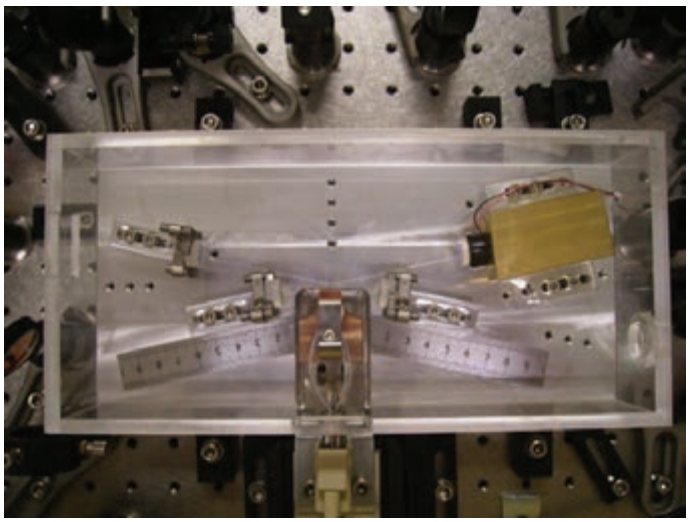


Figure 7.1: The doubly resonant, bow-tie optical parametric oscillator.

The original design for this cavity (the mirror mounts, optical layout, mounting block etc.) originates from the work of Nicolai Grosse for use in a harmonic entanglement experiment [116, 133, 80]. The general design suited our needs because of its modularity, the ability to customise the components and the fact that it was a travelling wave design.

The cavity length is controlled with a lead zirconate-titanate ceramic stack (or PZT). A number of modifications were made to the PZT mount, the crystal holder, the crystal translation stage, and the cavity length to improve upon the original setup. The design allows for easy exchange of non-linear materials, multiple input and output ports, and the ability to alter the mirror separations with ease. The layout of the cavity is shown in Figure 7.2. The design choices taken when constructing the bow-tie cavity were;

- Physical Size - One of the design philosophies was to design the OPO as small as possible after taking all other considerations into account. This was to reduce the bench space required for the device.
- Mechanical Stability - The entire cavity is constructed on a large steel block, shown in Figure 7.1. The steel block at the base of the cavity allows for easy construction and, once completed, shifting of the cavity. It allows the OPO to be a modular device and adds rigidity to the entire device.
- Angle - The folds in the bowtie cavity are designed to have an angle of 12 degrees. This angle is made as small as possible in order to reduce astigmatism. The limit on the reduction of this angle is the cavity size and clipping of mirror mounts by the beam.
- Cavity Length - The cavity length was chosen such that with the smallest angle possible, the mirror mounts would not interfere with the beam paths.

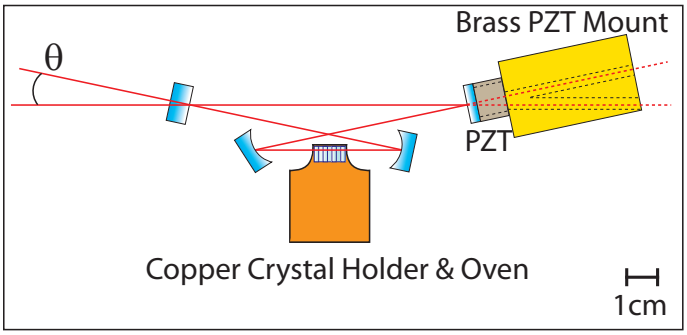


Figure 7.2: To-scale layout of the bow-tie optical parametric oscillator. The bow-tie angle, θ, is 12 degrees and the brass PZT block is cut out to allow for optical access to the back of the mirror attached to the PZT.

Three of the cavity mirrors were provided by Photon LaserOptik; the two curved HR mirrors (with radii of curvature of -38 mm) and the input/output coupler, all half an inch in diameter and 6.35 mm thick. The remaining mirror, from AT Films, was the 99.82 % reflective coherent locking mirror which was half an inch in diameter and 3.2 mm thick to reduce the weight that the PZT had to drive. The PZT is a standard ceramic stack from Piezomechanik, attached using Loctite E-30CL epoxy due to its large Young's modulus, onto a brass block. This block is hollowed such that the input and output ports of this mirror are accessible. This was done after finding that coupling through the curved mirrors had proved difficult in previous cavities. The coatings were all designed for an angle of incidence of 6 degrees to account for the 12 degree cavity angle. The cavity parameters are shown in Table 7.1

Cavity Parameter	**Symbol**	**Value**	**Uncert.**	**Units**
Fundamental Wavelength	λ_1	1064	-	nm
Second Harmonic Wavelength	λ_3	532	-	nm
Input/Output Coupler Fundamental Ref.	$R_1^{in/out}$	0.839	0.001	-
Input/Output Coupler Pump Ref.	$R_3^{in/out}$	0.722	0.001	-
Coherent Locking Mirror Fundamental Ref.	R_1^{cl}	0.9982	0.0002	-
Total Intra-Cavity Loss for Fundamental	T_1^l	0.0026	0.0001	-
Total Intra-Cavity Loss for Pump	T_3^l	0.046	0.001	-
Finesse at Fundamental	$\mathcal{F}_1$	36	3	-
Finesse at Pump	$\mathcal{F}_3$	16.8	0.6	-
Linewidth at Fundamental	$\Delta\nu_1$	31.2	1.5	MHz
Linewidth at Pump	$\Delta\nu_3$	65.3	2.5	MHz
Auxiliary Laser Detuning	Δ_{Aux}	29.8	-	MHz
Optical Path Length	L	0.279	0.001	m
Threshold Power	P_3^{crit}	118.0	0.1	mW
Escape Efficiency	η_{esc}	0.985	0.001	-
Non-linear Coupling Strength	g	1891	9	$\mathrm{s}^{-\frac{1}{2}}$
Curved Mirror Radius of Curvature	RoC	-38	-	mm

Table 7.1: The properties of the ANU squeezer. The loss values given are the transmission values of loss equivalent high reflectivity mirrors. Uncertainties are given for measured values.

The parameters were chosen to optimise squeezing without sacrificing stability. The locking frequencies were chosen to match the frequencies available at the LIGO Hannford site and as such the linewidths of the cavity were tailored, as much as possible, to work with these frequencies. The following sections detail the optimisation of the non-linear medium within the cavity and the two most important measurements for defining the cavity properties, the threshold, which describes the strength of the non-linear interaction, and the intra-cavity loss.

7.1.1 Optimising the Wedged Crystal

Determining the optimum operating conditions of the non-linear medium within the squeezer is non-trivial. The position of the crystal varies the intra-cavity dispersion as well as the loss due to local defects; the temperature sets the phase matching condition; and finally, the pump power determines the amount of localised heating. Figure 7.3 shows the experimental setup used for optimising the crystal parameters within the OPO.

Figure 7.3 shows that both the fundamental (red) and pump (green) fields enter the cavity through the input coupler after combining on a dichroic mirror. The cavity length can be scanned using the PZT attached to one of the cavity mirrors and a second PZT can be driven to scan the relative phase between the fundamental and pump fields. The crystal is located on a translation mount and temperature actuation is available through utilisation of a Peltier device. Photo-detectors are used to detect the transmitted and reflected components of both laser fields. The reflected fields are separated by a second dichroic mirror.

- 1) The bottom half of 1) in Figure 7.4 shows that the crystal was set up such that the fields travelled through it at one edge to provide a position reference. The pump power was then set to the expected final operating pump power so that localised

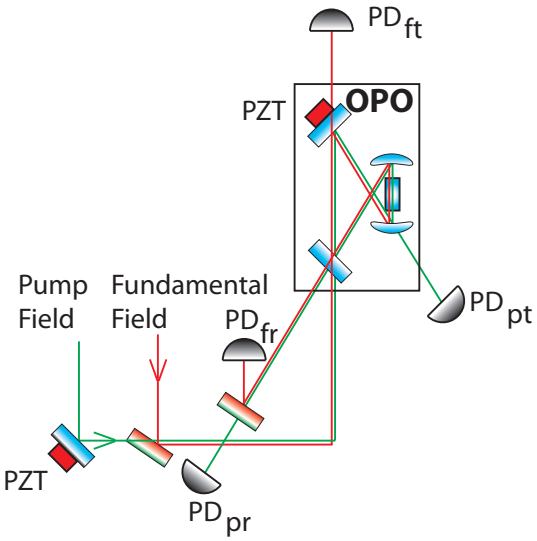


Figure 7.3: Schematic of the experimental setup used to optimise the crystal parameters within the OPO. PZT - piezo-electric transducer, PD_{ft} - transmitted fundamental field photo-detector, PD_{fr} - reflected fundamental field photo-detector, PD_{pt} - transmitted pump field photo-detector, PD_{pr} - reflected pump field photo-detector.

heating due to absorption from the pump was factored into the optimisation. A theoretical value for the threshold, allowing one to estimate the expected operating pump power, can be found using the technique described by Boyd and Kleinmann [121] or alternatively, one can estimate values from previous experiments by using information such as that shown in table 7.1. The top half of 1) illustrates the pump (green) and fundamental (red) fields detected on PD_{pt} and PD_{ft} photo-detectors respectively, as the cavity length was scanned. At this point, the temperature of the crystal had been set to the phase matching temperature specified by the manufacturer, of approximately 35 degrees Celsius. At this temperature, the two fields resonated at different cavity lengths.

- 2) To tune the dispersion such that the cavity was co-resonant for both circulating fields, the crystal was displaced relative to the circulating fields as shown in the bottom half of 2) in Figure 7.4. The crystal was translated until the transmitted fields, as monitored on PD_{pt} and PD_{pt}, showed that both fields were resonating at the same cavity length.

- 3) The co-resonance condition found by following the previous steps provides only a rough estimate for co-resonance. This is mainly due to the fact that co-resonance has been found as the cavity was being scanned and therefore the pump field, which causes localised heating and thereby introduces dispersion, does not have enough time to reach an equilibrium within the cavity. Additionally, it is not strictly co-resonance that one need to concern themselves with. To optimise the amount of squeezing, the non-linear interaction strength is the parameter that must be optimised. To optimise this, and to remove dynamical effects, the non-linear gain was then measured. This involved locking the cavity length, reducing the fundamental field power, and scanning the phase of the pump field using the second PZT. These

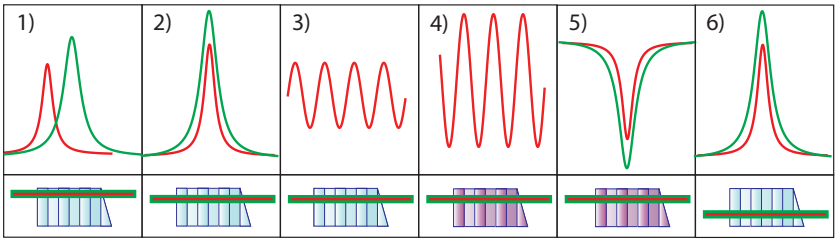


Figure 7.4: Schematic of the procedure used to optimise the temperature and position of the non-linear crystal. The top row illustrates the modelled signals (not to-scale) from various photo-detectors for the fundamental (red) and pump (green) fields (the magnitude of these signals is arbitrary and the x-axis is time). The bottom row shows the position of the laser beam through the crystal inside the cavity. Different colours for the crystal represent that the crystal temperature has changed. Full description in text.

steps are described in further detail in § 7.1.2. The signal on PD_{ft} then shows the non-linear gain of the fundamental field as the relative phase between the fundamental and pump fields was scanned.

- 4) To optimise the non-linear gain, the temperature of the crystal was then adjusted until the largest amount of gain was seen. The non-linear gain is recorded (as described in § 7.1.2.

- 5) Although the maximum non-linear interaction had then been determined, another parameter must also be addressed, loss within the crystal. The loss can change by very large amounts as the crystal is translated due to local defects in the crystal. Therefore the loss at this crystal position was measured, by unlocking and then scanning the cavity length and analysing the reflected field signals detected by PD_{pr} and PD_{fr}. The process is quite involved and so a detailed description is left for § 7.1.3.

- 6) The locally optimised non-linear interaction strength and the intra-cavity loss were then known for some position within the crystal. However, the crystal wedge was designed such that there were at least 3 locations where the beams could enter the crystal, at the phase matching temperature, and be co-resonant. The entire process was then repeated for these other locations, found by tuning the crystal to the phase matching temperature and translating the crystal to where the two fields were seen to co-resonate again. The entire process was then repeated, eventually resulting in a known crystal displacement with the lowest loss and highest non-linear interaction strength. The system was then brought back to this operating regime.

7.1.2 Measuring Threshold

The setup for measuring the non-linear gain of the OPA system is shown in Figure 7.5.

First, the pump field is blocked such that it does not enter the cavity, and the temperature of the non-linear crystal is set to approximately 20 degrees above the phase matching temperature (specified at approximately 35 degrees Celsius). The fundamental field enters the cavity, resonates and some portion is then transmitted through the cavity. This transmitted field reflects off a dichroic mirror, and is then detected on the photo-detector.

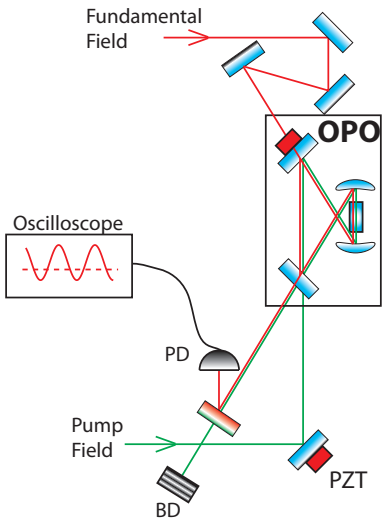


Figure 7.5: The setup for measuring the parametric gain. PD - Photo-Detector, BD - Beam Dump, PZT - Piezo-Electric Transducer.

This signal is measured on an oscilloscope and sets a reference level for the amount of transmitted cavity light in the absence of non-linear gain. Next, both the fundamental and second harmonic fields are made to enter the cavity and the temperature of the non-linear crystal is tuned to near phase matching. By scanning the phase of the pump field, using the piezo, the system is driven through amplification and de-amplification of the fundamental field (See § 5.5). The output field, consisting of light at both wavelengths, is then incident on a dichroic mirror, separating the two fields. The intensity of the fundamental field is once again detected on an oscilloscope. The resulting signal is a sinusoid showing the non-linear gain on the fundamental field. The amount of amplification and de-amplification is maximised by tuning the temperature of the crystal. By taking the ratio of the maximum and minimum voltage values of this sinusoid, with the initial reference value, one can determine the non-linear gain in both the amplification and de-amplification regimes respectively. By taking these measurements for many different pump powers, the threshold can be determined by using the theory introduced in § 5.5. The results for the ANU squeezer are shown in Figure 7.6.

The incident pump power was multiplied by the mode matching of this field to the cavity in order to get a more accurate figure on the amount of light that will actually enter the cavity and undergo cavity enhancement. The proportion of light that does not match the cavity mode simply reflects off the coupling mirror. The best fit to this data resulted in a threshold power of 118 mW for the amplification regime and a threshold of 85 mW for the de-amplification regime. The discrepancy is likely due to the fact that de-amplification is much harder to measure accurately. The signal to noise ratio is smaller because of the signal size and thus noise sources can add systematic offsets if not properly accounted for. With this in mind it was assumed that the amplification regime provided the best estimate and concluded that the squeezer had a threshold of approximately 118.0 ± 0.1 mW. Using this threshold power and the mirror properties in Table 7.1, Equation 5.36 can then be solved for the non-linear coupling term, g, to determine a value for the non-linear coupling term of $1891 \pm 9\,\mathrm{s}^{-\frac{1}{2}}$.

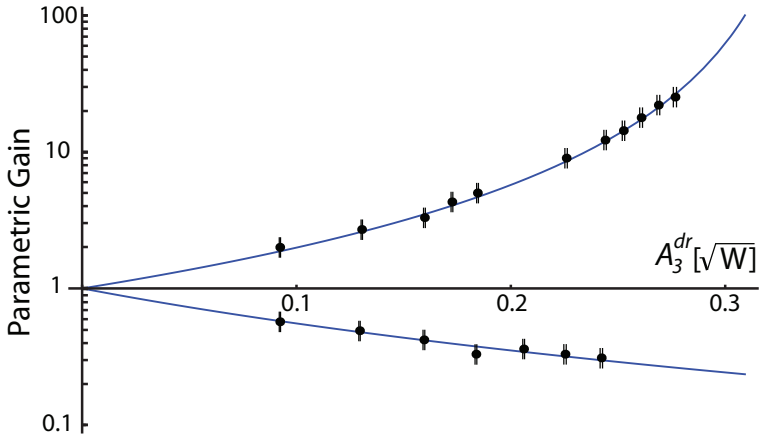


Figure 7.6: Parametric gain as a function of pump amplitude for the ANU squeezer. The points are the data and the blue curve is the best fit to Equation 5.52.

7.1.3 Measuring Intra-Cavity Loss

It has been emphasised many times that the total losses in the cavity, arising from sources such as scatter, absorption and non-unity mirror reflectivities, limit the amount of squeezing that can be produced. To get a reasonable understanding of the experiment, it is therefore necessary to characterise the intra-cavity losses as accurately as possible.

The simplest method to get an estimate for the intra-cavity losses is to measure the transmission of the individual components within the cavity, the mirrors and the crystal, in a single-pass setup. Whilst this will give a rough estimate on the losses, this method cannot provide information on scatter or absorption and the single-pass experiment cannot mimic the exact conditions in the cavity. Local defects, in particular within the non-linear medium, but also from defects in the mirror coatings, are spatially dependent and thus depend upon the beam radius as it interacts with the various loss sources, as well as the positioning of these beams on the mirrors and the positioning through the crystal. The cavity samples these loss sources many times before exiting the cavity so a method of measuring loss that utilises this fact can provide a more accurate measure than single-pass techniques.

Perhaps the most common method for determining intra-cavity losses is to measure the finesse of the cavity. Equation 5.17 shows that the finesse is dependant upon both the coupler reflectivity and the intra-cavity losses. The first step is to determine the properties of the coupler, specifically the power reflectivity. Measuring the transmissivity of an individual mirror, such as the coupler, is simple provided that the mirror is not highly reflective. This is a relative measurement (between input and output powers) and as such, any systematic error from the power meter is subtracted. If the reflectivity of the coupler were much higher, as is the case for the cavity HR mirrors, then the power reflectivity must be inferred from the transmission. By doing this, one must assume that loss from scatter and absorption is negligible, which is often valid, but as the reflectivity of the mirror increases, these effects become more important.

Now that the reflectivity of the input coupler is known, we can infer the losses of the OPO. Note that we infer the total losses of the intra-cavity system, and also that the process by which the loss occurs (such as scatter or absorption) is of no importance because

they all have the same effect on the intra-cavity field. We assume a single-ended cavity, feed a field at the fundamental frequency into the input/output coupler and investigate the reflected light as the cavity resonance frequency is scanned. Figure 7.7 shows the power of the reflected field for various intra-cavity losses, using Equation 5.14.

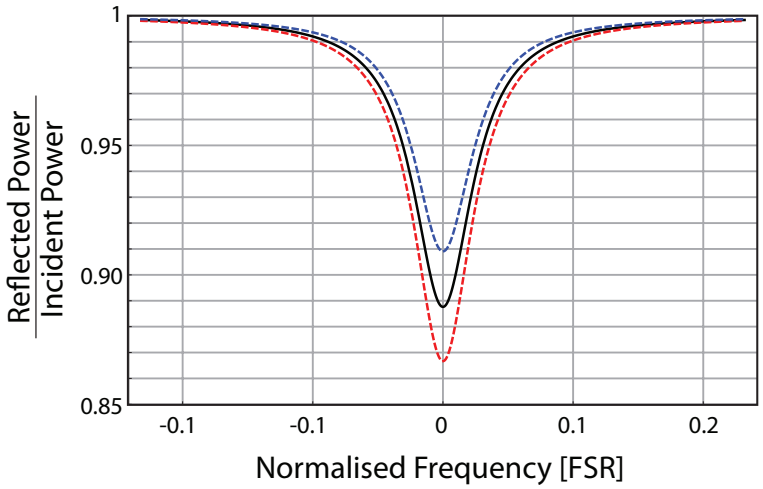


Figure 7.7: Theoretical fundamental field power response ratio of a singly-ended OPO (off phase matching). The cavity resonance conditions is scanned and the response is shown for 3 values of the total intra-cavity loss; 0.5% as the solid black line (finesse of 34.8), 0.4% as the dotted blue line (finesse of 34.6), and 0.6% as the dotted red line (finesse of 35.0).

Figure 7.7 reveals that the finesse is a poor measure of the loss. It is evident that the linewidth of the cavity response varies minutely as the intra-cavity loss changes. The reason for this is evident in how the finesse is defined, see Equation 5.17. By measuring the finesse, we are trying to measure a change in reflectivity of 0.1% on top of a reflectivity of 85% from the input/output coupler.

Additionally, resonance is typically scanned by driving a PZT. This needs to be done over an entire free spectral range in order to evaluate the finesse 5.2.3. The response of these PZTs is not linear, particularly not over a large scanning range [134]. So not only is the signal small, but the non-linear response of the PZT results in a completely unreliable measurement.

Figure 7.7 shows that the power of the reflected field on resonance is a much more sensitive measure of the intra-cavity loss. Working from Equations 5.14 we can determine the reflected power on resonance as a function of the cavity decay rates. We drop the operator notation as we are only interested in the classical behaviour of the fields and write

$$\begin{aligned} \frac{P_1^{ref}}{P_1^{dr}} &= \frac{A_1^{ref*} A_1^{ref}}{A_1^{dr*} A_1^{dr}} \\ &= \left(\frac{2\kappa_1^{in/out} - \kappa_1}{\kappa_1} \right)^2 \\ &= \left(\frac{\kappa_1^{in/out} - \kappa_1^{loss}}{\kappa_1^{in/out} + \kappa_1^{loss}} \right)^2 . \end{aligned} \tag{7.1}$$

where we have assumed that the driving power, P_1^{dr} is equal to the amount of reflected power when the cavity is off resonance. This is a valid assumption for small intra-cavity decay rates and when the cavity is far from resonance. The P_1^{ref} term, refers to the amount of power that is reflected from the cavity when it is on resonance.

The result is an equation that resembles the form of the fringe visibility, Equation 3.8.1, which is very sensitive to the loss term, $\kappa_1^{in/out}$. Determining the loss from the reflected power is very sensitive because the drop on reflection is completely due to the intra-cavity loss. In the absence of any intra-cavity loss, the reflected field equals the input field and no dip in reflection is seen. The limitation to determining the intra-cavity loss from this method is how well one can know the reflectivity of the input/output coupler. However, as noted previously, this mirror typically has a moderately low reflectivity and thus the error introduced by being unable to account for scatter and absorption is minimal.

The final note on measuring the intra-cavity loss using the reflected dip in power on resonance is the effect that imperfect mode-matching has. Imperfect mode-matching into the OPO can result in a low estimate for the loss because of the fact that any mode-mismatched light will enter the detection photodiode and provide a DC offset. The situation is shown in Figure 7.8, where the reflected power from the input/output coupler of the ANU cavity was measured as the cavity length was scanned.

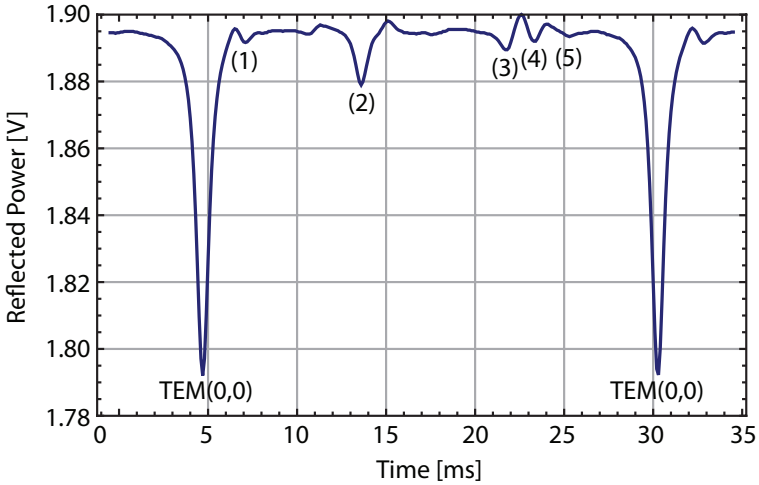


Figure 7.8: Reflected power from the OPO, in Volts, as the OPO length was scanned using a PZT. The higher order modes, modes (1)-(5), do not resonate when the cavity is locked to the TEM(0,0) mode.

We can quantify how well the light field is matched to the resonance of the cavity by defining the mode-matching value, MM, as

$$MM = \frac{MM_{(0,0)}}{MM_{(0,0)} + \sum MM_{(x,y)}}, \tag{7.2}$$

where $MM_{(0,0)}$ is the amount of light in the desired mode, in this case the (0,0) transverse electromagnetic mode, TEM$(0,0)$, and $\sum MM_{(x,y)}$ is the summation of the power in all the remaining modes. All values are measured in voltages and are relative to the maximum

reflected value. For Figure 7.8 the mode matching is written as

$$\frac{TEM(0,0)}{TEM(0,0)+((1)+(2)+(3)+(4)+(5))} = 0.784 \pm 0.003. \tag{7.3}$$

Let's assume that the cavity were off resonance, which is to say that we assume all modes $MM_{(x,y)}$ and $TEM(0,0)$ are off resonance. In this case, all of the light is reflected towards the detector. The power measured is therefore the total power of this field. When the cavity is on resonance, the reflected field consists of the reflected TEM(0,0) mode, *plus* the other modes that are not resonant at this particular cavity length. These other modes, in both cases, will result in a DC offset that needs to be subtracted. We can therefore rewrite Equation 7.1 for the case where the modematching is not perfect as

$$\frac{P_1^{ref} - \sum MM_{(x,y)}}{P_1^{dr} - \sum MM_{(x,y)}} = \frac{\left(\kappa_1^{in/out} - \kappa_1^{loss}\right)^2 - \sum MM_{(x,y)}}{\left(\kappa_1^{in/out} + \kappa_1^{loss}\right)^2 - \sum MM_{(x,y)}}. \tag{7.4}$$

This results in a loss value of 0.0026 ± 0.0001 per round-trip. Using this loss value and the coupling mirror transmission shown in Table 7.1, results in an escape efficiency of 0.985 ± 0.001.

For comparison, we can determine the estimated round-trip loss by multiplying all known loss sources from inside the cavity, the measured mirror transmission, and the expected AR coating and PPKTP absorption losses. The AR coatings on the non-linear crystal are specified at a transmission of 99.99% or higher, the PPKTP absorption was found by Wang [109] to be 0.02%/cm, and the two curved HR mirrors had a measured transmission of approximately 0.5%. The transmission of the coherent locking mirror was measured to be 0.9982 ± 0.0002. The resulting loss equivalent mirror power reflectivity is

$$\underbrace{0.9999 * 0.9999}_{2*ARCoatings} * \underbrace{0.9998}_{PPKTPAbsorption} * \underbrace{0.9995 * 0.9995}_{2*CurvedHR} * \underbrace{0.9985}_{Coh.Lock.HR} = 0.9971 \pm 0.0005. \tag{7.5}$$

Equation 7.5 shows that the resulting loss equivalent mirror is 99.71 %. Using the values for the coupling mirror and cavity length from Table 7.1, this results in an escape efficiency of 0.983 ± 0.003. As discussed, this value is not as reliable as found by measuring the ratio of the reflected power on and off resonance but the two methods do agree.

7.2 The ANU experiment

The optical layout and electronic schematic for the ANU squeezer are shown in Figures 7.9 and 7.10 respectively. The majority of the main laser power (Innolight Mephisto) is directed towards the SHG (SHG lock), with the remaining light passing through the modecleaner cavity (MC lock) and providing the light for the LO field at the homodyne detector. A beamsplitter in the LO path taps off some light to be used as a seed for measuring visibility and parametric gain. The SHG passes through a Mach-Zender setup (MZ lock) in order to stabilise the pump power before entering the OPO. The Mach-Zender was only locked when squeezing measurements spanning over many hours were desired. The reflected pump field is detected at the OPO Ref 532 photo-detector and used to control the OPO cavity length via the Pound-Drever-Hall locking technique (OPO lock). The auxiliary laser, phase locked to and frequency shifted from the main laser by

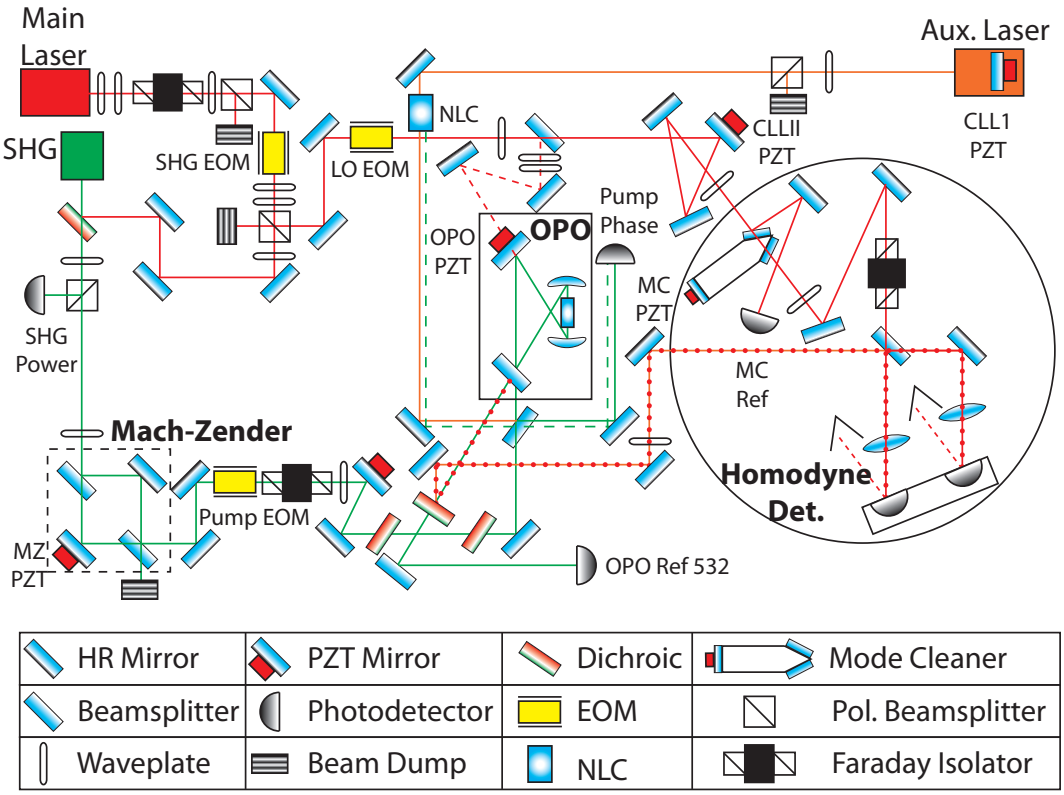


Figure 7.9: The layout of the ANU experiment. All of the optical elements apart from mode-matching lenses have been included. Red, green and orange fields are the fundamental, pump, and frequency offset auxiliary laser fields respectively. The dashed green beam is the auxiliary laser (Aux. Laser) second harmonic field, obtained after traversing a single pass through a non-linear crystal (NLC). The dotted red beam is the squeezed field. EOM is electro-optic modulator, dichroic indicates a dichroic mirror, and Pol. Beamsplitter represents a polarising beamsplitter.

29.8 MHz, makes a single pass through a PPKTP nonlinear crystal, generating a small amount of second harmonic. The second harmonic from the auxiliary laser beats, at 59.6 MHz, with some of the main laser second harmonic at the pump phase detector and is used to lock the phase of the pump field to the phase of the auxiliary laser (CLLI lock). The remaining power from the auxiliary laser enters the OPO, where it senses the nonlinear interaction and a second sideband field is generated, at -29.8 MHz. The beat between these two sidebands and the LO, at 29.8 MHz, is then used to derive an error signal to lock the phase of the LO to the phase of the auxiliary laser (CLLII lock).

The electronic loops for the ANU squeezer are shown in Figure 7.10. The second harmonic generation cavity length (SHG), Mach-Zender phase (MZ), optical parametric oscillator cavity length (OPO), the coherent locking loop phases (CLLI and CLLII), and the local oscillator mode cleaner cavity (MC) loops are all illustrated. The first half of the coherent locking loop, CLLI, locks the phase of the pump to the phase of the auxilliary laser, and the second half, CLLII, locks the phase of the local oscillator to the phase of the auxiliary laser. In the Mach-Zender loop, MZ, the SR560 preamplifier is set with a 6 dB per octave low pass filter at 3 Hz with a gain of 1. In the CLLII loop, the SR560 is set with a 6 dB per octave low pass filter at 100 Hz with a gain of 5. The servos, high voltage

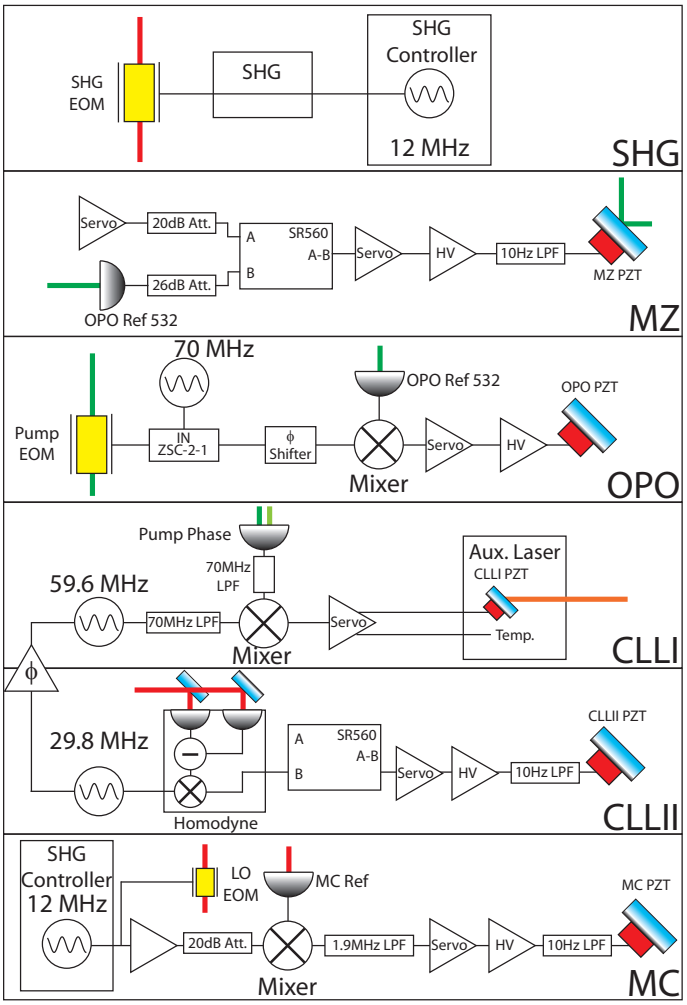


Figure 7.10: The electronics of the ANU experiment, showing the locking loops used. EOM - electro-optic modulator, HV - high voltage amplifier, SR560 - Stanford research systems preamplifier, LPF - low pass filter. Further explanation in text.

amplifiers (HV), and the phase shifter were all constructed in-house. The phase between the 29.8 MHz and 59.6 MHz frequency sources of the coherent locking loops are locked by a function within the frequency generators.

7.3 Squeezed State Control

It was shown by McKenzie *et al.* in 2004 that in order to produce squeezing in the audio detection band and below it was necessary to vacuum seed the OPO in order to remove various noise couplings [63]. This necessity is made apparent in Appendix 1, where the cavity equations of motion have been used to illustrate the relevant noise couplings. However, the presence of a bright seed allows for the use of standard modulation locking techniques in order to control the phase of the squeezing field [116]. Without a bright seed field centred about the squeezing, control of the squeezed state becomes a challenging task. In order to control the vacuum squeezing produced in the first audio-band squeezing

measurements, McKenzie used a technique known as quantum noise locking [135]. This technique was shown to work and was capable of controlling vacuum squeezed states. However, it was found that, when locking a bright squeezed state, the performance of the lock was worse than a reference technique, that of standard dither locking. The stability of the lock was shown to be limited by both the amount of squeezing and the squeezing bandwidth [135]. Unfortunately, these two parameters are not easily improved, and the amount by which they can be improved is limited. It was concluded that quantum noise locking was a useful technique but that its performance was limited.

7.4 Coherent Control

In 2006 Vahlbruch *et al.* introduced a new method for locking a vacuum squeezed state. It was a coherent control technique and is now commonly simply called *coherent locking* [66]. This locking technique derives its error signals from a bright field that enters the squeezer, but is detuned from the resonant condition of the cavity, which is at half of the pump frequency. This field can then provide information about the non-linear process within the cavity, and hence the phase of the squeezed vacuum field, without seeding the squeezing. The scheme is illustrated in Figure 7.11.

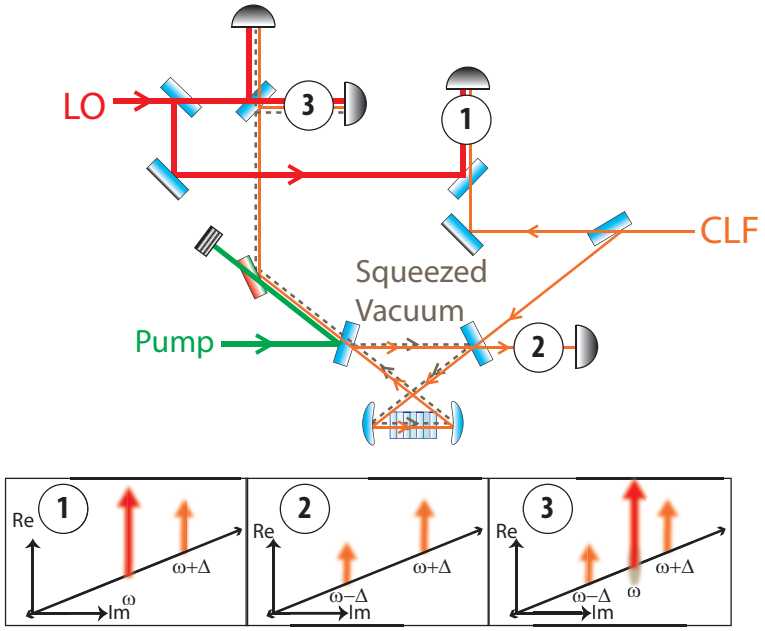


Figure 7.11: The coherent locking scheme. The fields important to the various locking loops are highlighted at locations 1,2,3. Full description of the fields and locking loops can be found in the text.

The final goal of the locking scheme is to lock the phase of the squeezing to the phase of the local oscillator. This occurs in many steps. At location 1 the auxiliary laser field, a single coherent sideband field (CSF), shifted from the fundamental frequency by an amount Δ, beats with a field from the main laser. This signal is used to offset phase lock the two lasers by the frequency Δ. This auxiliary laser is frequency offset so that the field can enter the OPO and undergo non-linear gain without seeding the process at the main laser

frequency. Inside the OPO, the CSF experiences the non-linear interaction, resulting in a generated sideband field (GSF) (See § 7.4.1). The beat between the CSF and GSF fields, at a frequency of 2Δ, is detected at location 2 and is used to lock the phase of the pump to the phase of the auxiliary laser. It will be shown in § 7.4.1 that the phase of the GSF depends upon the phase of the pump field, providing the necessary phase information from the pump field for this loop. At location 3, the CSF, GSF and co-propagating squeezing have interfered with the LO at the homodyne detector beamsplitter. The beat at the frequency Δ betweeen the CSF, GSF and LO is used to lock the phase of the LO field to the auxiliary laser. The pump field which exits the OPO is transmitted through a dichroic mirror and dumped such that it does not reach the homodyne detector.

In summary, the coherent locking scheme has one loop to lock the phase of the auxiliary laser to the main laser. It then has a second loop to lock the phase of the pump field to the phase of the auxiliary laser and a final loop to lock the phase of the LO to the phase of the auxiliary laser. This chain of loops results in the phase of the LO being locked to the phase of the pump field. In Chapter 5 it was identified that the phase of the pump field defines the angle of the squeezing. Therefore, the phase of the LO is locked to the angle of the squeezing, as desired.

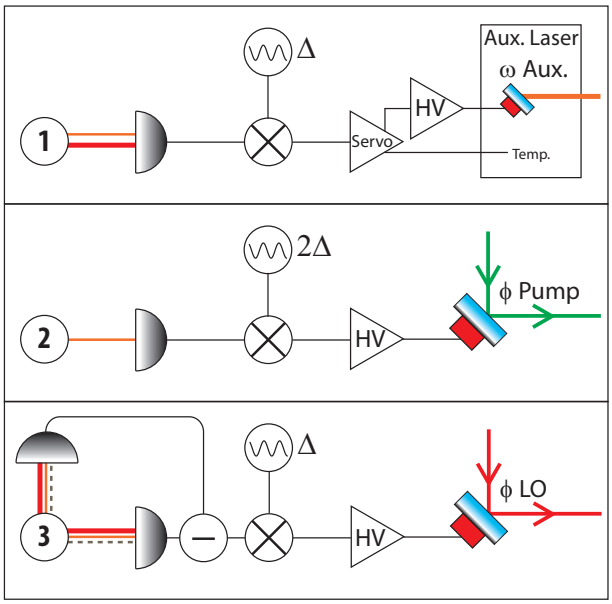


Figure 7.12: Simplified schematic of the locking loops that make up the coherent locking scheme. Fields 2 and 3 are the same as those illustrated in Figure 7.11. The two function generators at Δ and 2Δ are locked to one another to ensure phase stability between the two locking loops.

7.4.1 Coherent Control Model

We note that the coherent locking system is essentially a non-degenerate OPA, a system that we can model using the cavity equations of motion (Equations 5.32). We aim to model the classical behaviour of the CSF and GSF fields present in the coherent locking scheme. We begin by stating the classical equations of motion for the NDOPA and follow

similar methods as undertaken by [136, 137, 138]

$$\begin{aligned}
\dot{a}_1 &= i\Delta_1 a_1 - \kappa_1 a_1 + g a_3 a_2^\dagger + \sqrt{2\kappa_1^{in}} A_1^{dr} \\
\dot{a}_2 &= i\Delta_2 a_2 - \kappa_2 a_2 + g a_3 a_1^\dagger + \sqrt{2\kappa_2^{in}} A_2^{dr} \\
\dot{a}_3 &= i\Delta_3 a_3 - \kappa_3 a_3 - g a_1 a_2 + \sqrt{2\kappa_3^{in}} A_3^{dr}.
\end{aligned} \tag{7.6}$$

In this case, terms with a subscript 1 represent the properties relating to the CSF, subscript 2 relates to the GSF field and subscript 3 is, as before, the pump field. The first simplification we will make is that the decay rates for the two locking fields, the CSF and the GSF, are equal. This is reasonable provided that the cavity is locked at the squeezing frequency (half of the pump field) because the two fields are symmetric about the fundamental frequency. We will also assume that the pump field is resonant. The detuning terms for the CSF and GSF fields *cannot* be set to zero as in the case of the OPA because these terms describe the detuning of the coherent locking field from half the pump frequency. The CSF field entering the cavity under standard operating conditions is incredibly small, so it is valid to assume that the pump field is not depleted. Finally, the GSF is not driven by an external field and hence we set $A_2^{dr} = 0$. Solving for the steady state of the system allows us to write

$$\begin{aligned}
(\kappa_1 + i\Delta_1)\alpha_1 &= g\frac{\sqrt{2\kappa_3^{in}}\hat{A}_3^{dr}}{\kappa_3 + i\Delta_3}\alpha_2^* + \sqrt{2\kappa_1^{in}}\hat{A}_1^{dr} \\
(\kappa_2 + i\Delta_2)\alpha_2 &= g\frac{\sqrt{2\kappa_3^{in}}\hat{A}_3^{dr}}{\kappa_3 + i\Delta_3}\alpha_1^*.
\end{aligned} \tag{7.7}$$

We wish to allow for any pump phase because this is important to the coherent locking scheme. To achieve this, Equations 7.7 are simultaneously solved for the intra-cavity field terms, α_1 and α_2, and their complex conjugates. This results in

$$\begin{aligned}
\alpha_1 &= \frac{-\sqrt{2\kappa_1^{in}}A_1^{dr}\left(\kappa_2 + i\Delta_2\right)}{\frac{2\kappa_3^{in}g^2A_3^{dr*}A_3^{dr}}{\kappa_3^2} - \left(\kappa_1 + i\Delta_1\right)\left(\kappa_2 + i\Delta_2\right)} \\
\alpha_2 &= \frac{-2g\sqrt{\kappa_1^{in}\kappa_3^{out}}A_1^{dr}A_3^{dr}\left(\kappa_2 + i\Delta_2\right)}{\kappa_3\left(\frac{2\kappa_3^{in}g^2A_3^{dr*}A_3^{dr}}{\kappa_3^2} - \left(\kappa_1 + i\Delta_1\right)\left(\kappa_2 + i\Delta_2\right)\right)},
\end{aligned} \tag{7.8}$$

where the relationship between the detuning of the CSF and GSF fields is such that $\Delta_1 = -\Delta_2$.

Using the input/output relations (Equation 5.9), we can determine the output fields and from this the power of the fields exiting the system (using Equation 5.5). Figure 7.11 shows the setup for the coherent locking field.

The chosen parameters match those for the ANU cavity, given in Table 7.1. It is important to note that for the coherent locking scheme, the input coupler is now the HR mirror that the CLF enters the cavity through, with $R = 0.9982$, and the output coupler is, as before, the mirror where the squeezing and the coherent locking field exit the cavity and travel towards the homodyne detector, with $R = 0.839$. The input coupler was chosen to have a reduced reflectivity, when compared to the other HR mirrors, in order to slightly

increase the component of the GSF field at location 2. The reflectivity was chosen such that the additional intra-cavity loss, resulting in a lower escape efficiency, would be small compared to the total intra-cavity loss. The reflected and transmitted fields can thus be written

$$A_1^{ref} = \sqrt{2\kappa_1^{in}}\alpha_1 - A_1^{dr} \tag{7.9}$$

$$A_1^{trans} = \sqrt{2\kappa_1^{out}}\alpha_1 \tag{7.10}$$

$$A_2^{ref} = \sqrt{2\kappa_2^{in}}\alpha_1 \tag{7.11}$$

$$A_2^{trans} = \sqrt{2\kappa_2^{out}}\alpha_1, \tag{7.12}$$

from which the powers of the fields can be determined using Equations 5.5. Under most circumstances, we can assume $\kappa_1^{in} = \kappa_2^{in}$ and $\kappa_1^{out} = \kappa_2^{out}$, by requiring that $\Delta_1 = -\Delta 2$.

We are now in a position where we can investigate the properties of the fields present in the coherent locking scheme. Figure 7.13 shows the power and phase properties of the transmitted and reflected CSF and GSF fields as a function of pump power and phase. The figure illustrates many important features of the coherent locking scheme. With no pump power, there is no generated sideband field. This is expected as there is no non-linear interaction to transfer photons to this field. As the pump power is increased, the non-linear gain of the system increases and photons are added to the CSF and GSF fields. The gain in the figure is limited because of the large detuning of the fields. It is necessary to ensure that the powers in the CSF and GSF do not become too large, otherwise noise on these fields will be detected by the homodyne detector, adversely affecting the squeezed state measurement.

As the phase of the pump field is varied, the phase of the CSF does not change but the phase of the GSF varies linearly for both the reflected and transmitted fields. This is a very important feature for the coherent locking scheme. The phase of the GSF is linearly related to the phase of the pump and this is what facilitates the locking of the phase of the auxiliary laser to the phase of the pump field.

The power seen in the reflected CSF and GSF fields highlights the limitation to this coherent locking scheme as it was constructed in this experiment. The reflected field consists of two components, the promptly reflected field and a component that escapes the cavity. The issue is that due to the fact that we are coupling in via an HR mirror, the promptly reflected component is much larger than the small amount that escapes the cavity. This is very clear in Figure 7.13, where the reflected CSF has a power of 99 μW and the reflected GSF has a power of 0.8 nW. The beat between these fields is small due to the very minor contribution from the GSF. The small amount of intra-cavity field leaking through this HR mirror results in a lock with a suboptimal signal to noise ratio and limited bandwidth.

It was experimentally seen that when the power of the auxiliary laser (or CSF field) was increased above hundreds of micro-Watts the squeezing degraded. The mechanism by which this happens is believed to be due to direct detection of the noise from the locking field on the very sensitive homodyne detector but is not fully understood. The effect this has, however, is to set a limit on the strength of the coherent locking field that can enter the cavity. In particular, this sets a limit on the amount of GSF field that exits the cavity, limiting the size of the beatnote used for the various locking loops.

One might ask whether the locking loop which locks the phase of the pump to the

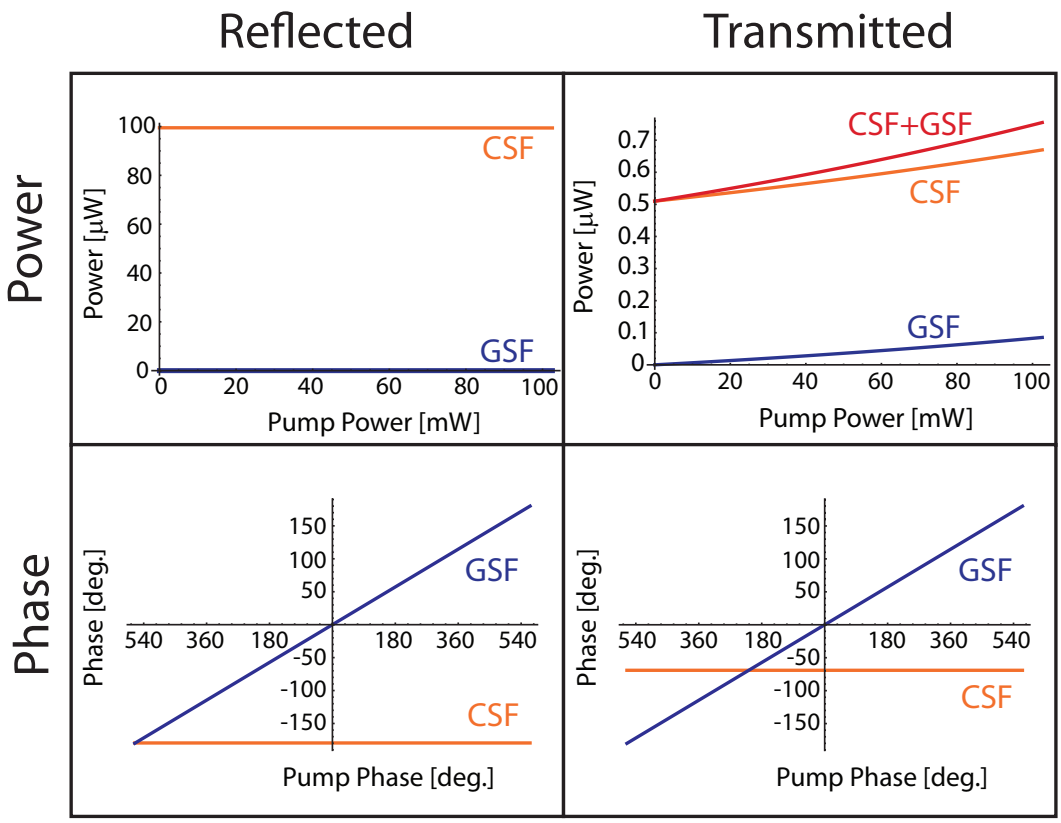


Figure 7.13: The powers of the transmitted and reflected CSF and GSF fields as the pump is increased up to OPO threshold and the phase of the CSF and GSF fields as the phase of the pump field is varied. The system is modelled on the ANU squeezer and as such all parameters are chosen to match those given in Table 7.1; cavity length $L = 0.273m$, non-linear interaction strength, $g = 1891$, reflectivity of the input/output coupler at the pump and fundamental wavelengths, $R_3^{in/out} = 0.722$ and $R_1^{in/out} = 0.839$, respectively, and intra-cavity loss at fundamental and pump, 0.26% and 4.6%, respectively. The coherent locking field power was set to 100 μW with a detuning of $\Delta_1 = 2\pi \times 29.8\,\text{MHz}$. For the phase diagrams, a pump power of 90 mW is assumed.

phase of the auxiliary laser, location 2 in Figure 7.11, could instead be derived from the field that exits the front coupler, where transmission is orders of magnitude larger than those for the HR mirrors. However, this field also carries the squeezed light and as such detection of this field before the homodyne detector cannot be made without degrading the squeezing. It might then be possible to derive the error signal for this lock at the homodyne detector. The homodyne detector is already used to lock the LO phase to the coherent locking fields but it could also be used to lock the CSF to the pump phase provided that the extra path length from the squeezer to the homodyne detector were stable. This should theoretically produce a much stronger lock because of the increased field strength exiting through the front coupler. The issue that arises, however, is that the homodyne detector must now detect in the gravitational-wave audio-band, at Δ_1 and at $2\Delta_1$. This is technically challenging to do without affecting the performance of the detection circuit in measuring the squeezed state and of the locks themselves. In particular, second harmonics of the lock at Δ_1 may affect the stability of the second lock at $2\Delta_1$.

The coherent locking scheme was the original scheme used to control the angle of the

squeezing in the ANU squeezer. With this scheme, the unity gain frequency of the loop that locked the phase of the auxiliary laser to the pump phase was approximately 20 Hz and the bandwidth of the loop that locked the LO phase to the CLF was 40 kHz. The very poor bandwidth of the fist loop was the limiting factor in our control scheme and so it was desirable to determine if a new method could be devised in order to improve the stability of our squeezed state control.

7.4.2 Modified Coherent Locking

The concept of the modified coherent locking scheme was that instead of locking the phase of the CLF to the pump phase by sensing the non-linear interaction, the phase between these two fields could be locked by first up-converting some small portion of the auxiliary laser field and then beating the pump field with this second harmonic CSF field. This would also lock the phase of the auxiliary laser to the phase of the main laser in the same locking loop, removing one of the locking loops from the original coherent locking scheme.

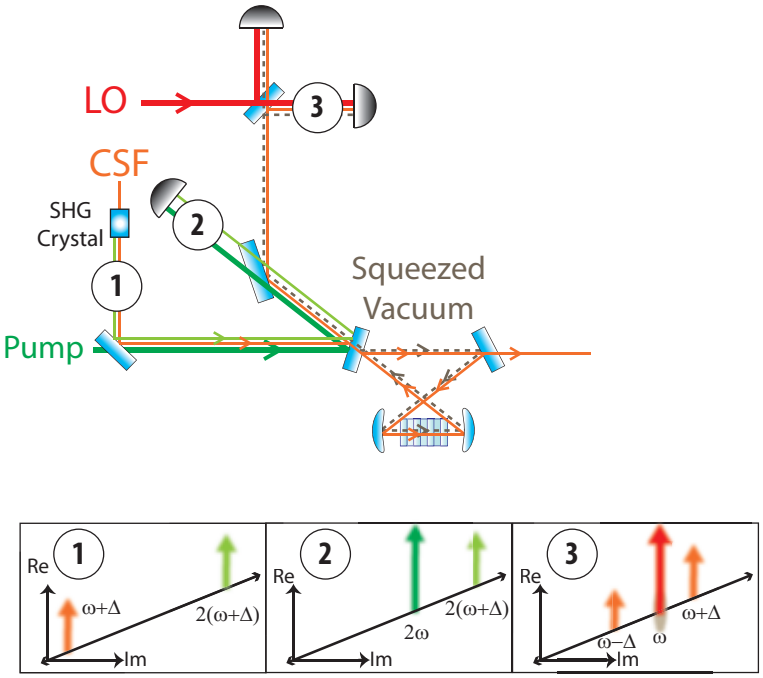


Figure 7.14: The modified coherent locking scheme. The important fields are highlighted at locations 1,2,3. A full description of the fields and locking loops can be found in the text.

The experimental setup for the modified coherent locking scheme is shown in Figure 7.14. At location 1 the coherent sideband field (CSF), shifted from the fundamental frequency by an amount Δ, and its second harmonic are present. The second harmonic is generated using a single-pass non-linear medium. At location 2 the green light reflected from the OPO and transmitted through a dichroic beamsplitter is detected. The beat, at a frequency of 2Δ, between the up-converted CSF and the pump field is used to offset phase lock the frequency of the auxiliary laser to the main laser using a standard offset phase lock [139]. At location 3 the fields are unchanged from the original coherent locking scheme, the CSF, GSF, squeezing and LO fields are all present at this location.

By moving to this modified coherent locking scheme, the first locking loop can now be optimised by varying parameters that do not affect the squeezing. In particular, the power in the frequency doubled CLF field can be tailored for the locking loop. The already strong second half of the lock, the loop that locks the phase of the LO to the phase of the auxiliary laser, remains unchanged. The electronic loop of the modified coherent lock, which locks the phase of the auxiliary laser to the phase of the pump, is shown in Figure 7.15.

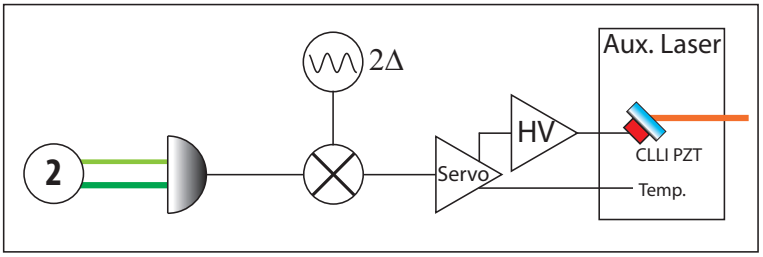


Figure 7.15: The modified coherent locking electronics scheme. The beat between the pump and up-converted auxiliary laser fields is used to derive an error signal. This error signal locks the phase of the the auxiliary laser field to the phase of the pump field. Note that this loop also locks the frequency of the auxiliary laser to the frequency of the main laser, with an offset of Δ.

With this modified coherent locking scheme, the bandwidth of the loop locking the auxiliary laser phase to the pump phase was 2.5 kHz. The larger bandwidth results in better suppression of the phase fluctuations on the squeezed field, reducing phase jitter. This lock is more flexible due to two reasons; the amount of frequency doubled CSF can be varied and; the fundamental CSF field power is only used for the stronger lock at the homodyne detector, allowing for reduction of the the power in the CSF is so desired. The CSF power used in the experiment for this locking loop was not changed from the original coherent locking scheme, retaining its bandwidth of 40 kHz.

Both of these coherent locking schemes share an inherent weakness. This is that the coherent locking schemes do not strictly force the lock to the angle of squeezing. It merely ensures that the phase of the local oscillator in the homodyne detection scheme is locked to the phase of the pump. In both the original and modified coherent locking schemes, if the temperature of the non-linear medium changes in a way that cannot be compensated for, as is the case with localised thermal heating due to drifts in pump power, then the cavity is no longer perfectly resonant for both the pump and squeezed fields. This condition results in a shift between the angle of the squeezed light and the local oscillator. Nonetheless, even with this limitation, stability has been shown over many hours after recognising local heating due to pump field fluctuations were a major source of noise, first identified and overcome through stabilisation of the pump field power by Khalaidovski [140]. Stable squeezing, controlled using the coherent locking scheme,in this paper was shown to have a duty cycle of 99 % over 20 hours.

Chapter 8

Detection of Squeezed States

This chapter introduces the challenge of the detection of the states described thus far. The first measurements of squeezing in the audio-frequency band were limited by excess noise at low frequencies [63]. A "roll-up" in the frequency spectrum was seen towards lower frequencies. This noise was evident on both the shot noise and squeezing measurements [63, 66]. This led to the interesting situation where it could not be determined whether the squeezing itself was degrading at low frequencies, or whether it was simply a limitation of the measurement device. It was later shown that it was indeed possible to measure white shot noise at frequencies down to below one hertz [141]. Beam jitter, electronics, local oscillator noise coupling [64], non-stationary events in the balanced homodyne detector output [65], and parasitic interference [141, 142] have all been shown to contribute to the excess noise at low frequencies. In this chapter we present our findings on low-frequency noise sources and the steps by which the various sources were identified and mitigated. This chapter is based largely on the work presented in a paper recently written on this topic [87].

8.1 Balanced Homodyne Detection Noise Couplings

We investigate the noise coupling by modelling the photocurrents in the balanced homodyne detector in the same way as introduced in § 3.8. The model is set up as shown in Figure 8.1. A bright local oscillator, $\hat{A} = \alpha + \delta\hat{a}$, and the weak signal beam, $\hat{B} = \delta\hat{b}$, of interest interfere on a beamsplitter of power splitting ratio, $\eta_{\rm bs}$, that is close to, but not precisely, 50 %. We assume that the signal field has no coherent amplitude, and is thus a vacuum or squeezed vacuum state. The two photo-detectors have differing quantum efficiencies, $\eta_1(x, y)$ and $\eta_2(x, y)$, that vary in the transverse plane. A loss term, η_l, is added to one arm of the balanced homodyne detector to simulate loss from sources such as dust (location 2). We define the δ_V terms as the vacuum fluctuation contributions entering due to the the losses from the various loss sources, the inefficient photodiodes, δV_1 and δV_2, and the loss introduced in one of the balanced homodyne detector arms, δV_0.

The fields incident on the two photo-detectors, $F_{1,2}$, are written

$$\hat{F}_1 = \sqrt{\eta_1}\left(\sqrt{1-\eta_l}\left(\sqrt{\eta_{\rm bs}}\hat{A} + \sqrt{1-\eta_{\rm bs}}\hat{B}\right) + \sqrt{\eta_l}\delta\hat{V}_0\right) + \sqrt{1-\eta_1}\delta\hat{V}_1, \quad (8.1)$$

$$\hat{F}_2 = \sqrt{\eta_2}\left(\sqrt{\eta_{\rm bs}}\hat{B} - \sqrt{1-\eta_{\rm bs}}\hat{A}\right) + \sqrt{1-\eta_2}\delta\hat{V}_2, \quad (8.2)$$

where the explicit spatial dependence of the photodiode efficiencies have been removed for succinctness. The ideal photocurrent produced by photodiode 1 is then proportional to $\hat{F}_1^\dagger\hat{F}_1$ and similarly for the photocurrent in photodiode 2. The two photocurrents can

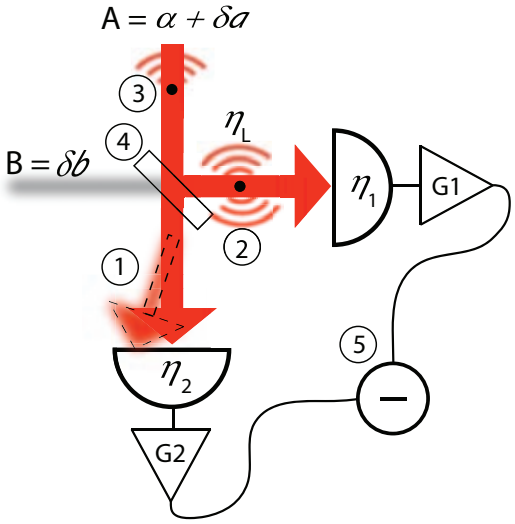


Figure 8.1: Conceptual view of the major issues arising in low-frequency measurements with a balanced homodyne detector setup; 1 - beam pointing, 2 - scatter resulting in loss and/or mode shape disortion, 3 - scatter leading to parasitic interference, 4 - the splitting ratio of the beamsplitter, 5 - the subtraction of the photocurrents. All other terms are described in the text.

thus be written

$$\begin{aligned} i_{HD1} &= \eta_1\eta_{bs}(1-\eta_l)\alpha^2 + \eta_1\eta_{bs}(1-\eta_l)\alpha\delta\hat{X}_a^+ + \\ &\quad \eta_1\sqrt{\eta_{bs}}\sqrt{1-\eta_{bs}}\sqrt{1-\eta_l}\alpha\delta\hat{X}_b^+ + \sqrt{\eta_1}\sqrt{\eta_{bs}}\sqrt{1-\eta_1}\sqrt{1-\eta_l}\alpha\delta\hat{X}_{V1}^+ + \\ &\quad \eta_1\sqrt{\eta_{bs}}\sqrt{1-\eta_l}\sqrt{\eta_l}\alpha\delta X_{V0}^+ \qquad (8.3) \\ i_{HD2} &= \eta_2\sqrt{1-\eta_{bs}}\alpha^2 + \eta_2(1-\eta_{bs})\alpha\delta\hat{X}_a^+ - \eta_2\sqrt{\eta_{bs}}\sqrt{1-\eta_{bs}}\alpha\delta\hat{X}_b^+ - \\ &\quad \sqrt{\eta_2}\sqrt{1-\eta_{bs}}\sqrt{1-\eta_2}\alpha\delta\hat{X}_{V2}^+, \qquad (8.4) \end{aligned}$$

where the $\delta\hat{X}_a^+ = \delta\hat{a}^\dagger + \delta\hat{a}$ terms are the amplitude quadrature fluctuation operators for the field fluctuation term $\delta\hat{a}$ and similarly for $\delta\hat{X}_b^+$. It is the subtraction of these two photocurrents, limited by experimental imperfections, that gives the output of the balanced homodyne detector. Many of the coupling mechanisms for the various noise sources can be seen in these equations and will be discussed in the following sections.

8.2 Electronic Issues

In an attempt to organise the various noise sources, the issues have been divided into those of an electronic nature, and those of an optical nature. However, the distinction is not always clear because many of these issues have a source in one of these origins but couple in via some mechanism in the other. We begin by discussing the issues of an electronic nature that limit the low-frequency measurement of squeezed states.

8.2.1 Balanced Homodyne Photo-Detector Electronics Designs

Two differing electronic designs for balanced homodyne photo-detectors were investigated. Their simplified schematics are shown in Figure 8.2. The first is a twin photo-detector,

variable gain design. In this design, the measured photocurrents are immediately amplified through a trans-impedance stage on separate electronic boards and an electronic subtractor then takes the difference between these two signals [63]. The gain between the two detectors can be varied to allow the subtraction of the photocurrents to be optimised electronically, allowing for the compensation of uneven optical powers and differences in the photodiode responses. This allows for the compensation of unequal $\eta_1(x, y)$ and $\eta_2(x, y)$ terms as well as an imperfect η_{bs} term from § 8.1. This design, henceforth known as the *variable gain* design, is given in detail in [108], and the first measurements of audio-band squeezing were made using this balanced homodyne detector design.

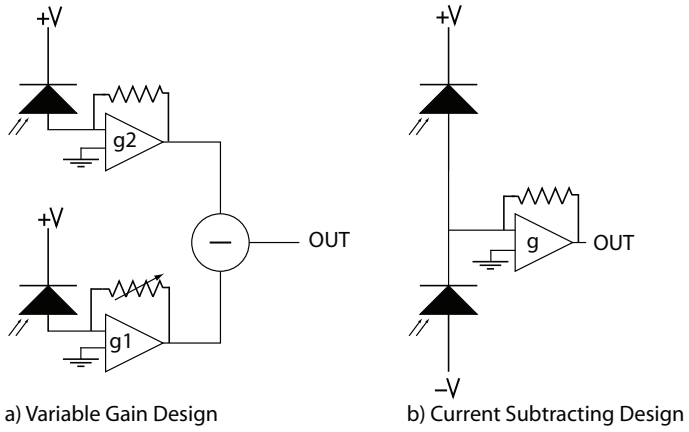


Figure 8.2: The two balanced homodyne detector schemes: a) The variable gain design. The photocurrents from the two photodiodes undergo separate transimpedance gain stage, g1 and g2, and the two output signals are then subtracted. b) The current subtracting design. The two photocurrents are subtracted before undergoing any electronic gain, g.

The second design is a single electronic board design, also shown in Figure 8.2. In this design the photocurrents of both photodiodes are immediately subtracted from each other before undergoing any amplification. There is no longer a variable electronic gain and thus, balancing of the subtraction must be done optically, through manipulation of the η_{bs} term. This is typically achieved by rotating the polarisation of the local oscillator and/or signal field. A homodyne detector utilising this design, henceforth known as the *current subtracting* design, is given in detail in [143]. A detector utilising this design was constructed at the Max Planck Institute for Gravitational Physics in Hannover, Germany and was provided to the ANU for use in the ANU squeezing experiment, to be later used in the LIGO squeezed light injection. This detector had already been shown to measure squeezing down to below 10 Hz [141] and was used in most of the homodyne detector measurements presented in this thesis. The current subtracting design has reduced classical noise due to the fact that all of the electronic components after the photodiodes are common.

The current subtracting design also eliminates a second noise source, flicker noise. This is not the case, however, for the variable gain design and as such it is necessary to address the issue. Flicker noise is a noise source found in resistors that is proportional to the amount of current passing through the resistor [144]. Flicker noise results in a pink noise spectrum, or a 1/f roll-up towards low frequencies. The source of flicker noise is not well known, but it can be mitigated using wire wound or metal film type resistors [145]. Figure

8.3 shows shot noise measurements for the variable gain design constructed with standard carbon film and with low flicker noise metal electrode leadless face (MELF) type resistors. An experiment was devised where shot noise was measured in three separate cases. The first two cases involved two identical homodyne photodetectors, the same as that presented in [108], set up to work as a variable gain homodyne detector. One measurement was taken with carbon film resistors and the other with MELF resistors. In the third case, one of these photodetectors was modified by adding a second photodiode and having current subtraction between these two photodiodes occur before the first gain stage, as shown in Figure 8.2. This was done on a board using standard carbon film resistors to show that the current subtracting design mitigates flicker noise. The results of these tests are shown in Figure 8.3.

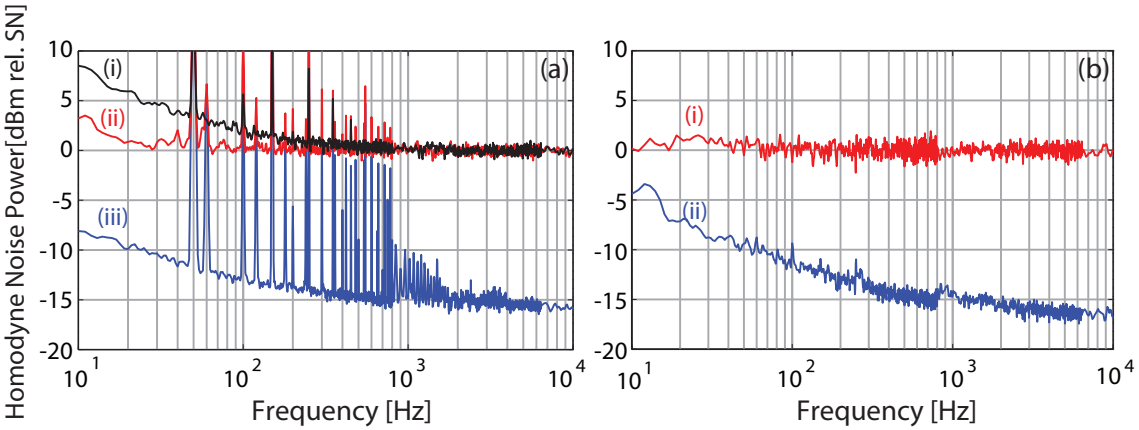


Figure 8.3: Figure (a) shows the dark noise, (iii), and two shot noise traces, (i) and (ii), for the variable gain balanced homodyne detector design with a local oscillator power of approximately 1.9 mW. In trace (i), standard carbon film resistors are used and the shot noise is seen to deviate strongly at lower frequencies. Replacing these resistors with low flicker noise resistors results in the near flat shot noise measurement seen in trace (ii). The deviation seen near 10 Hz is believed to be caused by parasitic interference, as it was seen to have a time dependence (See § 8.3.4). Figure 3(b) shows the normalised dark noise (ii) and shot noise (i) for the current subtracting design, constructed by adding a photodiode to one of the photodetectors used in the variable gain homodyne detector. As expected, even with standard carbon film resistors, flat shot noise was measured to 10 Hz. All traces are pieced together from 3 FFT windows: 0-800 Hz, 0-6.4 kHz, 0-102.4 kHz with 800 FFT lines, resulting in resolution bandwidths of 1 Hz, 8 Hz, and 128 Hz respectively. 200 RMS averages were taken for all traces. The noise spikes seen in (a) are due to mains harmonics that could not be removed.

The homodyne detector board design used to produce the traces in Figure 8.3 was the same as that used in the first measurement of squeezing in the audio-band [63]. Flicker noise had not yet been identified as a limiting noise source and as such, the measurements, taken with a variable gain board, had a very similar roll-up to that illustrated in Figure 8.3 (a) trace (i). Flicker noise was almost certainly the major noise contributor in this early experiment. As shown in Figure 8.3, the use of low flicker noise components is enough to reduce the flicker noise to a level that no longer affects the measurements at the frequencies we are interested in.

8.2.2 Dark Noise

Dark noise, the electronic noise of the detection system when no light is incident on the photodiodes, can also limit the measurement of shot noise and even more so squeezing. This is typically worse at lower frequencies where electronics tend to have increased noise. The problem is a technical one but it is worth briefly mentioning. Depending upon the operating conditions, the current and voltage noise in the op-amps, and less so the the dark current of the photodiode, are the sources of dark noise. Correct selection of readily available components ensures that the dark noise does not limit the measurement. The dark noise is typically measured relative to the shot noise level. The required clearance between the two is dependent upon the amount of squeezing one wishes to measure. Typical detectors can be made and bought with the shot noise approximately 10 dB above the dark noise for normal operating conditions, but more advanced designs have the shot noise around 20 dB or greater above the dark noise [143]. The clearance between shot noise and dark noise can be increased in the balanced homodyne detector by increasing the power of the LO field, thereby increasing the amount of measured shot noise. The limitations to this however, are the amount of power that the photodiode can receive before damage occurs and the fact that a greater LO field strength leads to a larger amount of classical noise. This increases the amount of subtraction the balanced homodyne detector needs to achieve in order to operate in the quantum noise limited regime.

Dark noise degrades shot noise and squeezing measurements because the noise of the electronics adds to the measured shot noise or squeezing and makes them appear noisier than they actually are. For the purpose of estimating the amount of squeezing incident on the detector, one can correct for dark noise by simply subtracting it from all other traces. For some signal, S, we can write its dark noise corrected value, S_{DNcor}, in dBm as

$$S_{\mathrm{DNCor}}[\mathrm{dBm}] = 10\log_{10}(10^{S/10} - 10^{\mathrm{DN}/10}), \tag{8.5}$$

where S[dBm] is the measured signal strength in dBm, and DN is the measured dark noise, in dBm.

8.3 Optical Issues

Many optical issues also need to be addressed in order to achieve flat shot noise across the audio-frequency spectrum. These optical issues and their solutions were investigated on a simple tabletop experiment, shown in Figure 8.4. The output of an ND:YAG 1064nm laser, after passing through a Faraday isolator, was directed into a steel chamber through an anti-reflection (AR) coated window. Within the chamber, the light underwent spatial filtering with a triangular ring mode-cleaner cavity or modecleaner with a linewidth of approximately 4.7 MHz. The modecleaner was locked using the Pound-Drever-Hall technique [146]. The light travels through a half-wave plate and is then incident on the beamsplitter which directs the light towards a current subtracting balanced homodyne detector [143]. The half-wave plate is used to tune the beamsplitter power splitting ratio, η_{bs}.

8.3.1 Optical Balancing

Balancing the optical splitting ratio of the beamsplitter is crucial to the performance of the balanced homodyne detector. A high common mode rejection ratio (CMRR), the

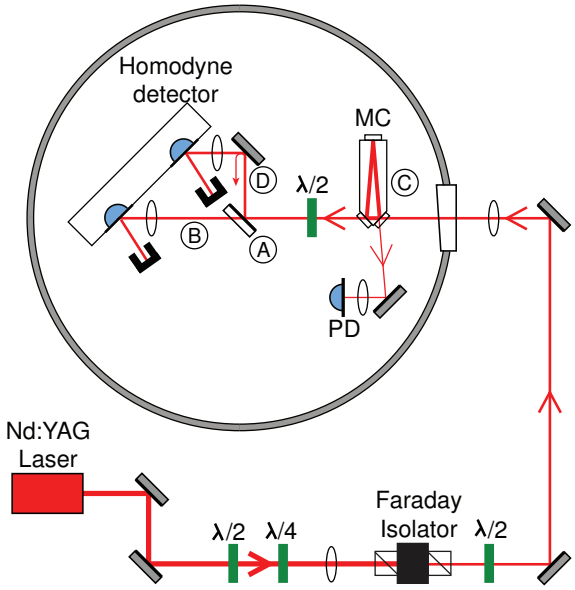


Figure 8.4: The optical layout of the experiment designed to detect low-frequency noise sources in the balanced homodyne detector setup. A high-speed photo-detector (PD) was used to lock the modecleaner cavity (MC). The core components are enclosed in a steel chamber with an AR coated window. The circled labels refer to the locations of issues discussed in following sections; (A) 8.3.1 - Optical Balancing; (B) 8.3.2 - Scattering Loss (C); 8.3.3 - Beam Pointing (D); 8.3.4 - Parasitic Interference.

measure of the ability of the balanced homodyne detector to reject signals common to both photodiodes, is necessary such that classical noise present on *both* fields entering the beamsplitter, the local oscillator and the signal field, is subtracted to a level well below the shot noise level by the balanced homodyne detector. This issue is highlighted in Equation 8.3. It is seen that with optimisation of the $\eta_{1,2}(x, y)$ and η_{bs} terms, the subtraction of these two photocurrents can completely remove all terms proportional to δX_a^+, which is the noise on the local oscillator field. This is a necessary condition for the balanced homodyne detector setup to ensure that one is accurately measuring the signal field, absent unwanted classical noise sources.

The importance of balancing the beam splitting ratio becomes even more critical when the signal field itself carries noise which needs to be subtracted, such as in the coherent locking scheme [66]. The optical balancing is firstly tuned with the splitting ratio of the beamsplitter, primarily through tuning of the angle. The coatings on the beamsplitters typically do not allow for a perfect splitting ratio, and as such, the final balancing is done through adjustment of the polarisation of the input fields using a half-wave plate. Alternatively, if a variable gain design were used, then this adjustment could be made through varying the gain. Optimising for the local oscillator arm, a stable CMRR of up to 80 dB was consistently achieved. Figure 8.5 shows that this level of subtraction should be enough to adequately to suppress the classical laser noise from the local oscillator across the entire spectrum. With this setup, a CMRR of 50 dB was obtained for the signal field. This is sufficient because the power in the signal field is many orders of magnitude less than the local oscillator.

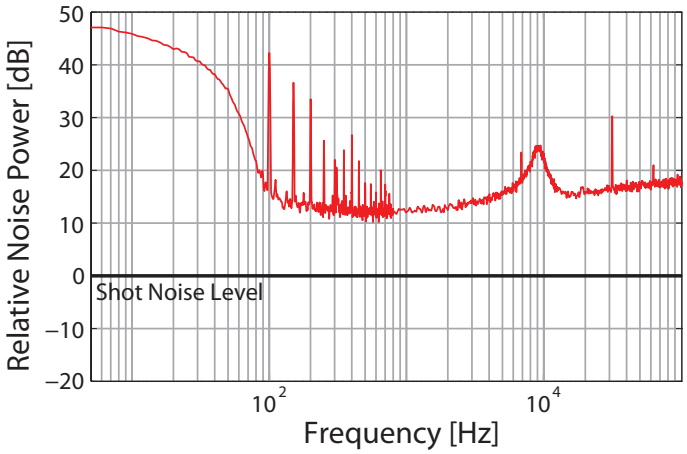


Figure 8.5: Relative intensity noise of the local oscillator field incident on the balanced homodyne detector. At low frequencies, local oscillator intensity noise is more than 40dB greater than shot noise. The trace is pieced together from 3 FFT frequency ranges: 0-800 Hz with 50 root mean square averages, 0-6.4 kHz and 0-102.4 kHz both with 100 averages. All traces taken with 800 FFT lines, resulting in resolution bandwidths of 1 Hz, 8 Hz and 128 Hz respectively.

8.3.2 Scattering Loss

We define scattering loss as an event occurring in the beam path which causes the scattered light to exit the beam path and be lost to the system. Scattering events before the beamsplitter will not affect the splitting ratio, but can couple noise in through changes in mode shape (this is further discussed in § 8.3.3). However, problems can arise if the loss originates in one of the balanced homodyne detector arms after the beamsplitter. This loss, η_l, is illustrated in Figure 8.1 and in Equations 8.3. Dust particles, originating from many sources such as clothing and skin, passing through the beam due to air currents or gravity, will result in a non-zero loss term and momentarily reduce the magnitude of the photocurrent, i_{HD1}. This momentary change in the photocurrent occurs at frequencies characteristic of the speed of the particle and will unbalance the homodyne detector setup. This will result in a reduction of the CMRR, the transient value of which may not be adequate to cancel the classical intensity noise of the laser.

If the dust particle in the beam path reduces the power in one of the beams by one percent, the subtraction is reduced from 80 dB to 40 dB. Figure 8.5 shows that this reduction in CMRR is enough to couple classical intensity noise into the output of the balanced homodyne detector. In addition, intensity noise at frequencies characteristic of the time taken for the dust particle to traverse the beam, will couple directly to the output. Prior work has shown that settling of dust does in fact improve the noise statistics of shot noise measured using a balanced homodyne detector [65].

The effect of loss due to scattering in the homodyne detector arm was investigated by first setting the optical power on each photodiode to 7V. With light on both photodiodes, the DC voltage was subtracted to less than 1 mV. An air puffer was then used to disturb the dust that had settled on various optics mounts. During this process, the subtracted output of the detector was monitored on an oscilloscope. With this configuration, substantial spikes in the subtracted output were observed for a few seconds after dust was excited with the puffer at a location after the beamsplitter.

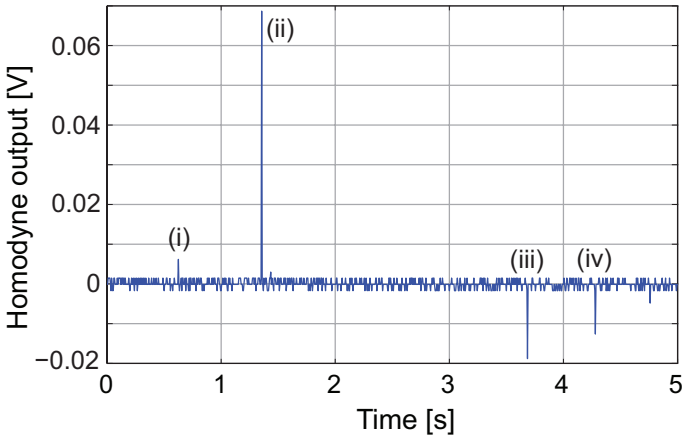


Figure 8.6: The DC output of the balanced homodyne detector as a function of time. At time t=0 a puffer was used to excite dust on a lens mount in the beam path. The spikes in the output, (i)-(iv), are individual dust events affecting the balancing of the homodyne detector.

Figure 8.6 shows the voltage measured on the balanced homodyne detector over time after one such excitation. Four spikes were observed, the largest causing a 0.07 V, or 1 %, change in the voltages between the two photodiodes. The typical voltage seen from these disturbances was approximately 0.01 V. These spikes would occur over a few seconds following the excitation.

We also excited dust activity before the beamsplitter to see what effect this has on the balanced homodyne detector. As expected, no spikes in the subtracted signal were observed because the dust-induced loss is now common to both arms. Isolation against scattering loss post homodyne detector beamsplitter is therefore crucial. The best way to stop dust is to ensure the cleanliness of the environment. Clean rooms, suitable clothing and enclosures all aid in reducing the amount of dust present in an experiment. However, there always will be some dust present in air, therefore eliminating the air currents that carry the dust through the beam paths, such as with an enclosure, is the next step. The steel tank placed around the detection optics in our experiment both reduced the dust present and the air currents. Once the tank was in place and sealed, the spikes in the DC output of the balanced homodyne detector were no longer seen.

8.3.3 Beam Pointing and Mode Shape

A fundamental limitation to the photocurrent subtraction of the balanced homodyne detector is provided by the spatially non-uniform response of the two photodiodes in the detection scheme. This is represented by photodiode quantum efficiencies that are a function of position, $\eta_{1,2}(x, y)$. An investigation into the response of typical photodiodes has shown that there are differences in the quantum efficiency across the photodiode surface and also indicated that small "dead regions" are common [147, 148]. This inhomogeneity allows for beam pointing and mode shape changes to couple into the measurement as a noise source.

Beam pointing, also known as beam jitter, is a change in the trajectory of the beam, which results in a translation of the intensity profile on the photodiode surface. Beam pointing originates from forces acting on optical components, such as acoustic vibrations

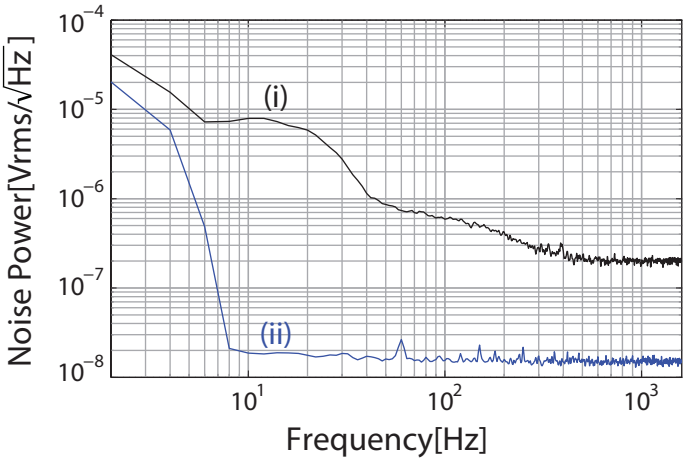


Figure 8.7: Balanced homodyne detector output traces illustrating a beam pointing event. Trace (ii) is the regular shot noise level and trace (i) is the output of the balanced homodyne detector during a typical blow-out event. The experimental setup is that shown in Figure 8.4, with the lid of the steel tank removed. Blow-outs such as that seen in trace (i) occurred every few seconds, with flat traces seen for most of the remaining time. All traces are measured with a span of 0-1.6 kHz with 800 FFT lines, resulting in a resolution bandwidth of 2 Hz. 100 root mean square averages were taken for both traces.

and air currents, or from changes in air density in the beam path resulting in varying refractive indices - the Schlieren effect [75]. If the transverse profile or positioning of the beam changes (see § 8.3.3), then the photocurrent from the photodiode will vary due to the spatial dependence of the diode efficiencies. Writing the spatial dependency on the quantum efficiency in radial coordinates, $\eta(r,\theta)$, the intensity measured by the photodiode is the integral over the entire diode surface, assumed to be circular, of the quantum efficiency multiplied by the intensity profile, and can be written

$$\int_0^{d/2}\int_0^{2\pi} I(r,\theta)\eta(r,\theta)\,d\theta\,dr, \tag{8.6}$$

where d is the diameter of the photodiode surface and $I(r,\theta)$ is the intensity profile of the incident field in radial coordinates.

This will directly modulate the photocurrent of the detectors and will be uncommon to both signals due to the differing responses of the photodiodes, hence it cannot be subtracted. The frequency dependence of this noise will be determined by the specifics of the photodiode inhomogeneities and the speed of the beam translation. The frequency characteristics are illustrated in [64] and show that the noise becomes larger at lower frequencies. The effect of beam pointing can be seen by disturbing the air near to the beam path, anywhere on the table. Many large blow-outs in the spectrum at frequencies up to about 1 kHz can be seen in the output spectrum when this disturbance is introduced. One of these events is shown in Figure 8.7.

These effects can be reduced by designing the experiment to be robust to pointing fluctuations by using short, shielded beam paths, and mechanically stable optics mounts. This effect can also be reduced by enlarging the spot size on the photodiode surface in

order to better average these inhomogeneities, whilst ensuring that all of the light is still detected. To completely solve the issue, a mode-cleaning optical cavity was placed in the balanced homodyne detector setup [141]. Beam pointing can be thought of as the coupling of some fraction of the original beam power into different spatial modes. The cavity is designed such that these modes are non-degenerate, and as such, the higher order modes do not resonate and are reflected from the cavity's input coupler. By this mechanism, pointing noise is converted into intensity noise. This intensity noise has occurred *before* the beamsplitter and is now common to both arms of the homodyne detector, allowing large CMRR of the balanced homodyne detector to remove this intensity modulation.

The modecleaner cavity also filters changes in mode shape, which may occur due to dust passing through the beam. Changes in the mode shape will also couple into intensity noise through the photodiode inhomogeneity. As with beam pointing, the modecleaner cavity will convert mode shape fluctuations into intensity fluctuations which can then be subtracted.

Once the mode cleaner was introduced, it was seen that air disturbances before the modecleaner did not result in low-frequency blow-outs. The experiment is still sensitive to pointing noise which occurs downstream of the modecleaner. By placing the modecleaner inside the homodyne detector system's steel tank, the steel tank then stabilised air currents and reduce dust in this section of the experiment, solving the issues of beam pointing and varying mode shape.

8.3.4 Parasitic Interference

Parasitic interference, also commonly known as a parasitic interferometer, was recognised as the limiting factor in the first measurements of flat shot noise across the entire audio-band [141]. Parasitic interference occurs when some light is scattered from the original beam path, through dust or surface imperfections, reflects off an object such as a mirror mount or another optical element, and re-enters the original beam path. Interference will occur between the original field and the scattered field at the point where these two fields overlap. Any change in the path length of the scattered field, caused by vibrations in the object that the scattered light is reflecting off or by the Schlieren effect will modulate the intensity of the light. If this modulation occurs after the beamsplitter then, much like beam pointing, it cannot be subtracted. Since the balanced homodyne detector is capable of measuring vacuum fluctuations, small fractions of a single photon on average per bandwidth Hz are clearly visible.

It is possible for the scattered light to exit the beam path completely upon scattering and it is possible for the scattered light to reverse direction but travel within the original beam path. If the scattered light exits the beam path, then careful placement of beam dumps will be enough to ensure that parasitic interference does not occur. However, if the scattered light does not exit the beam path, then it is more difficult to correct. This situation is shown schematically in Figure 8.8 for different scattering locations. The scattering is assumed to originate from either arm of the balanced homodyne detector and could occur at the photodiode or from one of the focussing lenses, illustrated in the Figure 8.8 as location 'S'. This *backscattered light* will traverse down the path of the main beam in the opposite direction. When the backscattered field scatters for a second time, it may again scatter into the beam path and then travel in the forward direction, resulting in interference between what we call the *forward scattered* field and the local oscillator field. The spurious interference converts motion of optics and air currents in the scattered

beam path into intensity noise. Figure 8.8 shows the possible locations where scattering can occur. We discuss the effect that parasitic interference, originating from different locations, will have on the shot noise.

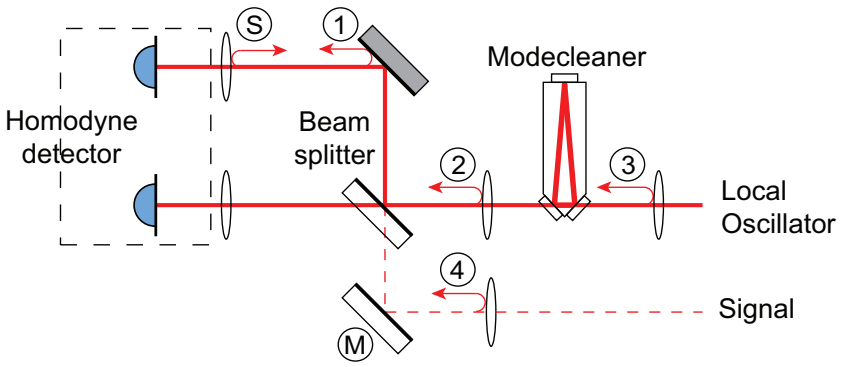


Figure 8.8: Scatter in the balanced homodyne detector leading to parasitic interference. The backscattering occurs in one of the detector arms, chosen here to be location S, and the scattered light travels along the beam path in the opposite direction. Forward scattering can then occur off a second surface, locations 1-4. The location of the forward scattering determines the effect on the balanced homodyne detector output, described in the text.

First we note that interferometric effects which occur between the beamsplitter and the laser source will result in intensity noise that is common to both arms of the balanced homodyne detector, which will be suppressed by the CMRR. This is the case for the forward scattering locations 2 and 3. Additionally, location 3 may be afforded some filtering effect by the modecleaner cavity. Parasitic interference in the experiment was detected here, and although it was not seen to negatively affect homodyne detector measurement (due to the CMRR being high enough to subtract these fluctuations) the ANU squeezer eventually had a Faraday isolator located near position 2 to remove interference at this location and ensure the CMRR was as high above the required level as possible. The forward scattering seen at location 1 will set up a parasitic interference in one arm of the balanced homodyne detector. The intensity noise on the optical field due to this parasitic interference is incident on only one photodiode, and hence cannot be subtracted. If the magnitude of this noise is large enough compared to the shot noise, then it will be seen in the output spectrum. The final location is forward scattering from location 4, in the path where the signal field enters the balanced homodyne detector setup. Scattering from this location will have a substantial impact since the balanced homodyne detector system is designed to be maximally sensitive to any light that enters through the signal port. Scattering in this path will interfere at the beamsplitter with the local oscillator field and the resulting intensity fluctuations will be anti-correlated in the two arms of the balanced homodyne detector, resulting in anti-correlated noise in each of the two photodiodes.

To determine whether parasitic interference was a limiting source of noise in our setup, shot noise spectra were taken with a beam dump in the signal port at location 4 in Figure 8.8 and with the signal port open but with no laser fields present in this path, shown in Figure 8.9. The beam dump was placed such that there were no optical components between it and the beamsplitter.

The difference between the shot noise for the two cases illustrated in Figure 8.9, with and without a beam dump in the signal path, is strong evidence for the presence of parasitic

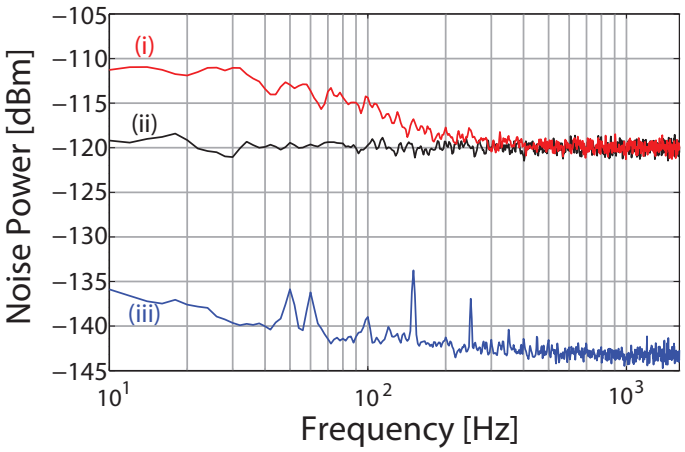


Figure 8.9: Parasitic interference from the signal port of the balanced homodyne detector. Only the local oscillator field was present. Trace (i) is the homodyne detector output with the signal port open, trace (ii) is the balanced homodyne detector output with the beam dump in place and trace (iii) is the dark noise of the balanced homodyne detector.

interference. A more conclusive method for determining the presence of scattered light is presented in § 8.3.5. Flat shot noise is seen when the beam dump is in place but excess noise is seen when it is absent. A Faraday isolator with approximately 22dB of isolation was inserted in place of the beam dump but the excess noise did not noticeably change, indicating that perhaps greater isolation is required or that scatter off the Faraday isolator itself was enough to cause these effects. The beam dump was moved up the signal path to try and determine if there was a single point where most of the forward scattering was occurring but it seemed to be a cumulative effect from several locations in this path.

8.3.5 Detecting the Presence of Parasitic Interference

Using a method known as opto-mechanical frequency shifting, or cyclic averaging, it is possible to determine with certainty whether a low-frequency roll-up is due to the presence of scattered light or otherwise [149]. This technique has been used in many situations to remove the noise from scattered light out of the detection band of interest [150, 151, 152]. The mirror M in Figure 8.8 is actuated with a PZT that is driven with a triangle wave. Now consider what happens to light that is scattered from position S, backwards to position 4, where a scattering plate has been introduced, and then forwards to the balanced homodyne detection. The PZT is set up such that it sweeps the phase of the scattered light through an integer number of cycles, time-averaging the scattered light to zero. The signal due to the motion of the scatter sources now appears at the dither frequency of the PZT and its harmonics. A simple classical model illustrates this effect. We start with an electric field, E, and write

$$E = A + Ce^{i\phi_s}, \tag{8.7}$$

where A represents the local oscillator field and we have chosen a reference frame such that we do not require a phase term for this field. The amplitude of the scattered light, C with phase ϕ_s, will interfere with the local oscillator field. The phase of the scattered

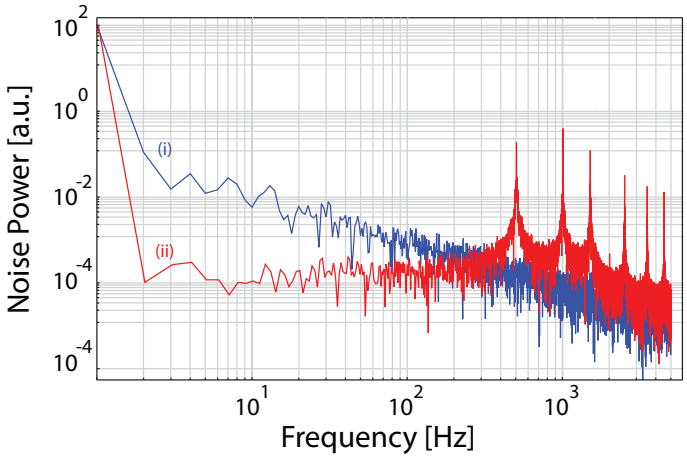


Figure 8.10: Simulated results of the cyclic averaging technique. Trace (i) shows the intensity noise of the LO field with some scattering present, resulting in a low-frequency roll-up. Trace (ii) shows that the low-frequency noise can be shifted up to the dither frequency and harmonics above it by driving a PZT in the scattering path with a triangle wave. Parameters are; local oscillator amplitude, $A = 1$, scattered light amplitude, $C = 0.01$, dithering frequency, $f = 500\,Hz$, and a standard deviation of the scattered phase of 0.01.

light is written

$$\phi_s = \sum R, \tag{8.8}$$

where $\sum R$ is the cumulative sum of normally distributed random noise about zero, with the standard deviation of this noise defining the degree of phase modulation due to air currents or mirror motion. The spectrum of such a field is shown in Figure 8.10.

In order to shift the intensity noise due to scattering out of the band of interest requires one to modulate the scattered field with a triangle wave. This provides a linear sweep on individual ramps, a necessary feature. The amplitude of this modulation is set such that one sweep of the dither progresses the phase an integer number of fringes. We can thus write for the dithered field

$$E_{\text{dither}} = A + Ce^{i\phi_d}, \tag{8.9}$$

where ϕ_d is the scattered phase plus the phase due to the dither. The phase of the field including the dithering is written

$$\phi_d = \phi_s + nD(f), \tag{8.10}$$

where n is the number of fringes wrapped by the dither and $D(f)$ is a triangle wave dither with some frequency, f. The spectra of the dithered scattered field is shown in Figure 8.10.

Figure 8.10 shows the spectra for the local oscillator field with scattered light present, trace (i), and with a PZT, dithered with a triangle wave present in the scattering path, trace (ii). The figure shows that even with scattering present, by dithering the scattered light we should be able to recover flat shot noise below the dithering frequency.

The experiment was then constructed as shown in Figure 8.8, with the mirror M being driven by a triangle wave. The amplitude of the dither was varied until the noise below the dithering frequency was at a minimum, this is the point at which the dither is shifting the phase by an integer number of fringes.

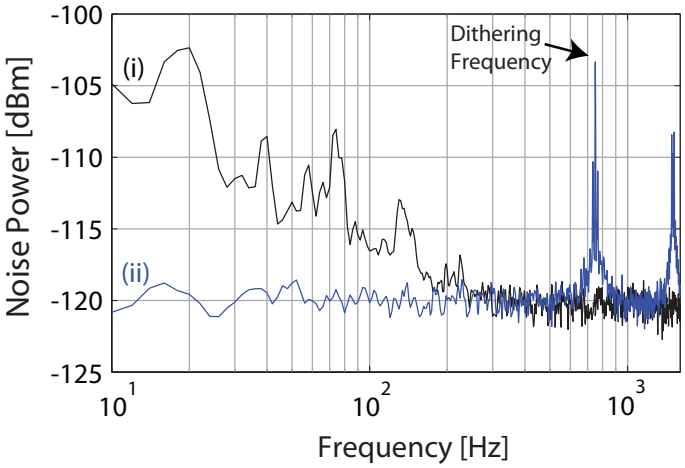


Figure 8.11: The cyclic averaging technique applied to the balanced homodyne detector with a scatter source placed in the signal path. Trace (i) shows the measured shot noise when a scatterer was placed in location 4 in Figure 8.8. With this scatterer in place, the PZT placed between this scatterer and the detector's beamsplitter was dithered at a frequency of 750 Hz. The amplitude of the dither was adjusted until all of the low-frequency noise was seen to shift up to the dither frequency and its harmonic as seen in trace (ii).

Figure 8.11 shows that the opto-mechanical frequency shifting-method has indeed confirmed the presence of scattered light in our setup. The scattered light noise has been mixed up to the dithering frequency and the shot noise is now flat below this frequency. This method has also provided additional information. The fact that the dithering results in flat shot noise has provided information on the location of the scattering and it signifies that the forward scatter originates in this beam path. If one were to move the position of the PZT in this beam path progressively further away from the balanced homodyne detector at some point it might be seen that the signal at the dither frequency decreased. This would indicate that scattering has not undergone the phase shifts introduced by the PZT and hence scattering has occurred nearer to the balanced homodyne detector than the current location of the PZT.

Parasitic interference can be overcome in 3 ways; by reducing the amount of scattering through the use of components with minimal surface roughness, by dumping the scattered photons through careful placement of beam dumps, and by reducing the phase fluctuations in the scattered fields through reduction of vibrations in the experimental setup. The use of super-polished optics and an enclosure around the entire experiment aids greatly in all of these endeavours.

8.3.6 Phase Jitter

The final detection issue is phase jitter. The effect that this has on the squeezing has been recognised previously by Takeno *et al.* [97]. Phase fluctuations on the vacuum squeezed

field result in a rotation of the squeezed field relative to the local oscillator field. The phase control loop designed to control the squeezing angle, such as coherent locking, will provide some finite suppression of this rotation but some rotation will still occur. This rotation results in a suboptimal detection angle for the homodyne detector setup, resulting in the measurement of a noise projection that does not align with the squeezed quadrature. This is shown in Figure 8.12.

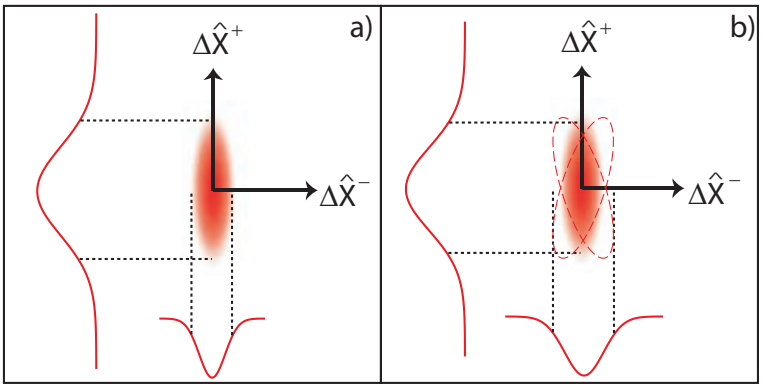


Figure 8.12: The effect of phase jitter on the measurement of a vacuum squeezed state. In case a) the phase is steady and measurement of either the amplitude or phase quadrature squeezed state, which is the projection onto one the quadrature axes, yields its properties. In case b), where the phase is fluctuating, and hence the squeezing ellipse is rotating, the projection of the state onto the two axes is averaged, resulting in different results to those found in the absence of jitter. In particular, the projection of the squeezing onto its axis, in this case the phase quadrature, increases. In the presence of phase jitter, the measured squeezing does not reflect the statistics of the original state.

Assuming some root mean square (RMS) fluctuations of the angle, determining the effect this has on the measured squeezing and anti-squeezing, $V^{\pm'}$, can be found via [97]

$$V^{\pm'} \approx V^{\pm}\mathrm{Cos}^2\theta + V^{\mp}\mathrm{Sin}^2\theta, \tag{8.11}$$

where $V^{\pm}$ is the squeezing/anti-squeezing level in the absence of phase jitter and θ is the rms phase fluctuations of the state. The larger the anti-squeezing, the greater the effect that phase fluctuations have. This is due to the fact that when rotated at the same angle, the projection of a longer ellipse will be proportionally larger for that rotation. This means that as the system is driven closer to threshold, and the anti-squeezing increases, it become increasingly important to stably control the phase of the squeezing relative to the local oscillator. The effect on measured squeezing due to phase fluctuations for an OPO with the ANU squeezer parameters is shown in Figure 8.13.

Figure 8.13 shows the degradation of measured squeezing for various rms phase jitter values. The squeezed state produced by this system mimics that we would expect from our ANU squeezer as the parameters have been chosen to match the ANU squeezer properties. We note that as the system gets closer to threshold, the rms phase jitter has a greater effect due to the increased anti-squeezing. The main source of phase jitter is changes in the refractive index of free space due to air currents, the Schlieren effect [75]. Hence the best way of reducing phase jitter is by reducing these air currents. The steel isolation tank placed around the detection part of the experiment helps in this regard but it is also

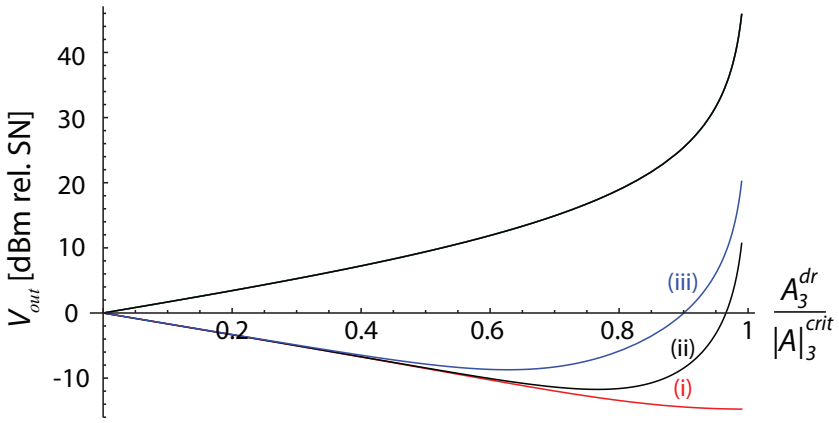


Figure 8.13: The variance at the output of an OPO as the system is driven towards threshold power, A_3^{crit}. The variance of the squeezing and anti-squeezing is shown for the same system with varying phase jitter of 0 degrees, trace (i), 1 degree, trace (ii) and 3 degrees, trace (iii). Parameters are; optical path length $L = 0.279\,\mathrm{m}$, non-linear interaction strength, $g = 1891$, reflectivity of the input/output coupler at the pump and fundamental wavelengths, $R^b_{co} = 0.722$ and $R^b_{co} = 0.839$, respectively, and intra-cavity loss at fundamental and pump, 0.26% and 4.6%, respectively.

necessary to place the rest of the experiment in an airtight (or close to airtight) enclosure in order to isolate the experiment from the typical air currents in the lab, as shown in Figure 8.15. Furthermore, optimisation of the locking loops, in particular increasing the gain at low frequencies, can aid in reducing the rms phase jitter of these states.

8.4 Squeezing from the ANU Squeezer

Now that the technical limitations to measuring shot noise and squeezing in the low-frequency regime are understood, we can move onto measuring these states of light. In this section, we will investigate the history of squeezing results from the ANU squeezer. In this way, we can directly see the effect that various changes to the setup has had. Shown in Figure 8.14 is the squeezing produced by the ANU squeezer over the course of two and a half years.

8.4.1 March 2009 Squeezing

The experiment at this time was in its early stages and many of the high quality components were not yet integrated into the table. Of particular importance is that the homodyne detector from the Max Planck Institute for Gravitational Physics had not yet arrived and in its place a voltage subtracting homodyne detector was used to measure all results. Low flicker noise resistors had not yet replaced the noisy carbon film resistors that were present from its original design. As such, low-frequency flicker noise accounted for much of the roll-up seen in the squeezing trace below about 2 kHz and also resulted in a shot noise measurement that was not flat across the entire spectrum (See § 8.2).

The squeezing magnitude was poor because a PPKTP crystal from Raicol with cheap in-house AR coatings was used. The loss on this crystal, using the method described in § 7.1.3 was found to be around 1.4 % per round trip at the fundamental frequency. This

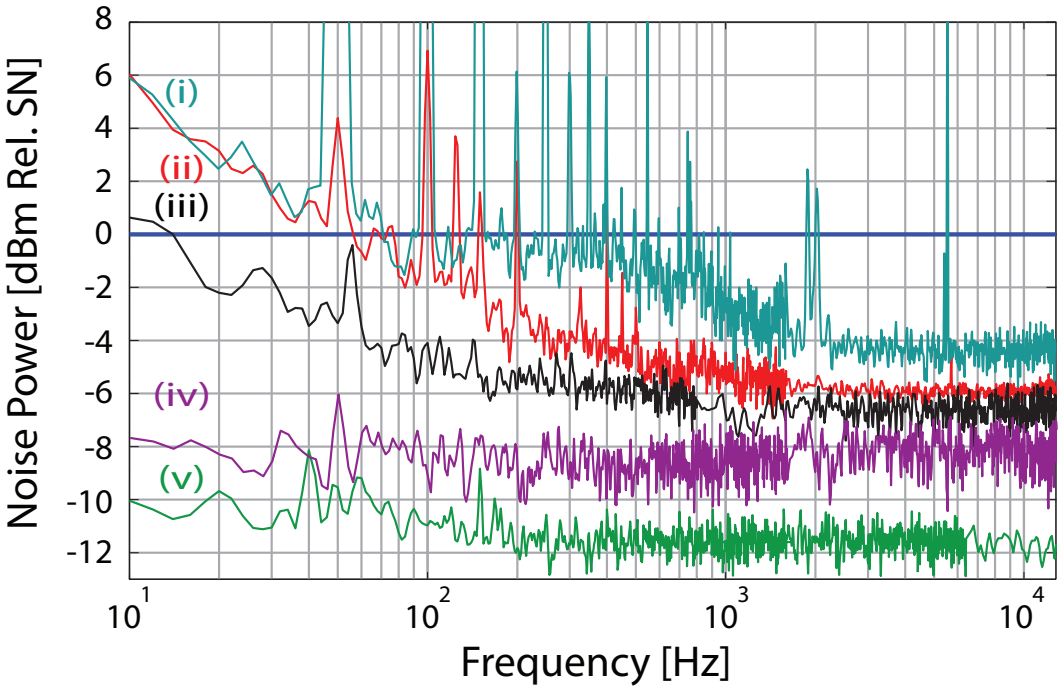


Figure 8.14: A snapshot of squeezing in the ANU squeezer over the years. (i) is the squeezing produced in March 2009 with a squeezing plateau (the lowest squeezing measured over some span) of 4.4 dB, (ii) is the squeezing produced in May 2009 that has a squeezing plateau of 6.0 dB, (iii) is the squeezing produced in January 2010 that has a squeezing plateau of 6.7 dB, (iv) is the squeezing produced in April 2011 that has a squeezing plateau of 8.6 dB and trace (v) is the squeezing produced in November 2011 with a plateau of 11.6 dB. Traces (i)-(iv) are constructed from two spans of 0-1.6 kHz (RBW = 2 Hz) and 0-12.8 kHz (RBW = 16 Hz). Trace (v) is constructed from three spans of 0-1.6 kHz (RBW = 2 Hz), 0-6.4 kHz (RBW = 8 Hz) and 0-102.4 kHz (RBW = 128 Hz). All traces have at least 100 averages except for traces (i) and (iv) which have 50. Dark noise has not been subtracted from any of the data. The dark noise clearance of traces (i) and (ii) was approximately 6 dB and the dark noise clearance of the remaining traces was greater than 10 dB as illustrated in Figure 8.16.

large loss value is enough to explain much of the very poor squeezing seen at this time, the theoretical effect of which can is discussed in § 5.6.

8.4.2 May 2009 Squeezing

The new components had still not arrived but much work had gone into increasing the stability of the locks. In particular, it was seen that noise was coupled into the squeezing at low frequencies via the coherent locking field. Low incident power levels, of 10s of micro-Watts hitting the CLF coupling mirror, were required in order to reduce this coupling to a level where the coherent field was not seen to degrade the squeezing (Theoretically investigated in § 7.4). Ground loops in these locks and the homodyne detector were also worked on, by having the detectors run from their own power supply which was plugged in at different locations until an optimum was found, such that the mains coupling into the measurements was drastically reduced. The phase lock between the primary and CLF lasers was also optimised by varying optical powers and electronic gains. These

improvements resulted in much less low-frequency noise coupling, reducing the roll-up seen at low frequencies, and the increased stability of the locking reduced in an improvement in squeezing magnitude due to a reduction in phase jitter. At this stage, the homodyne detector performance was still limited primarily by the flicker noise. The shot noise at this stage still exhibited a low-frequency roll-up.

8.4.3 January 2010 Squeezing

Many changes were made at this time. A new OPO was constructed that had a higher beam height to replicate the needs of the LIGO squeezing injection. Following advice from Roman Schnabel, new PPKTP crystals that were provided by Raicol, superpolished by Photon LaserOptik GMBH (with a micro-roughness < 0.2rms) and coated by LaserOptik GMBH (with reflectivity $< 0.1\%$ at 1064 nm and $< 0.5\%$ at 532 nm) were ordered. One of these crystals was placed in the new cavity and the intra-cavity loss due to the crystal dropped to the value shown in Table 7.1 of 0.26%.

Much work had been done on a separate experiment to determine the sources of the roll-up seen in the shot noise at low frequencies as discussed previously in this Chapter. The experiment was very successful, and flat shot noise was achieved to below 10 Hz. Flicker noise (§ 8.2), beam jitter (§ 8.3.3), dust and parasitic interference (§ 8.3.4) were found to be the major contributors to low-frequency noise, agreeing with previous findings from prior experiments. At this stage, a new homodyne detector, designed and supplied by Henning Vahlbruch from the Max Planck Institute for Gravitational Physics [143] was swapped in for the previous homodyne detector.

With the introduction of a modecleaner cavity in the LO path, and an isolation tank around the detection components of the setup, the shot noise was finally measured flat down to below 10 Hz. Parasitic interference was present in the squeezing spectrum but was not seen in the shot noise spectrum due to the beam path of the squeezing being blocked when measuring the squeezing.

8.4.4 April 2011 Squeezing

In order to measure flat squeezing then required comprehensive work on the reduction of parasitic interference locations which were initially believed, and later confirmed, to be a major source of noise (See § 8.3.4). It was also postulated that phase jitter may be playing a role in degradation of squeezing, however, it was difficult to confirm this postulate the second harmonic generator would often not produce enough power to reach threshold, where the phase jitter can be measured (See § 8.3.6). A semi-air-tight enclosure was built around the optical table in order to simultaneously reduce phase noise, beam jitter and protect the entire experiment from dust. The enclosure is shown in Figure 8.15.

At this stage, one of the ETX500 diodes in the homodyne detector had failed and was replaced. The efficiency of a number of these diodes were compared to one another and the one with the highest quantum efficiency was chosen. Between diodes from the same batch, only minor variance in efficiency was found, but a recently ordered batch of diodes showed an efficiency of approximately 10 % less than the ones in the older batch (ordered at least 1 year prior). The reduction of parasitic interference, and the box around the table, resulted in the first measurements of flat squeezing. Unfortunately, phase jitter on the beam was not measured at this time so it is difficult to say how much of the improvement in squeezing was due to the new diode, which could not be compared with the broken one, and how much was due to reduced phase jitter, as there was not enough power to get

Figure 8.15: The box encasing the ANU experiment. The primary purpose is to reduce air currents, which in turn reduces noise coupling of beam pointing and phase jitter. It also protects the experiment from dust.

close enough to threshold to characterise it sufficiently. In any case the enclosure certainly aided in stabilising the locks and reducing phase jitter.

8.4.5 November 2011 Squeezing

The final improvement to the experiment was the replacement of Epitaxx ETX500 photodiodes with high quantum efficiency (quoted at $> 99\%$) photodiodes from Laser Components GmbH. These photodiodes were measured to have a quantum efficiency which was approximately 4 % greater than that of the ETX500's. This was determined by constructing a small experiment in which a laser beam was incident on a beamsplitter, with one output field directed towards a detector where the diodes were switched in and out of the same electronic circuit, and the second output field directed towards a reference detector. As expected, this reduction in loss resulted in the detection of much more squeezing with greater purity. The full measurement of this squeezing is shown in Figure 8.16, illustrating the purity, flat shot noise, and dark noise clearance. The optical parametric oscillator was operated with 90 mW of input power (76 % of threshold). The expected propagation loss on the squeezed state is approximately 0.7 % and the balanced homodyne detector had a visibility of 99.4±0.2 %. The squeezing is seen to degrade slightly below 200 Hz due to the presence of parasitic interference that could not be removed. Small factors, such as the slightly altered detector alignment after the diodes were replaced, can have a large effect on the parasitic interference in the experiment and after this change there was difficulty in reducing interference such that it did not degrade the squeezing measurement.

Figure 8.16 shows that with a local oscillator power of approximately 1.9 mW, the shot noise of our detector is greater than 20 dB above the electronic noise. The anti-squeezing is 17.5±0.4 dB above the shot noise and is flat. Below 200 Hz, 10 dB of squeezing is seen and above this frequency 11.6±0.4 dB of squeezing is observed. The measurement shows that noise due to parasitic interference was suppressed down to a level approximately 10 dB below the shot noise. Correcting for dark noise (see § 8.5), the squeezing level becomes 11.9±0.4 dB.

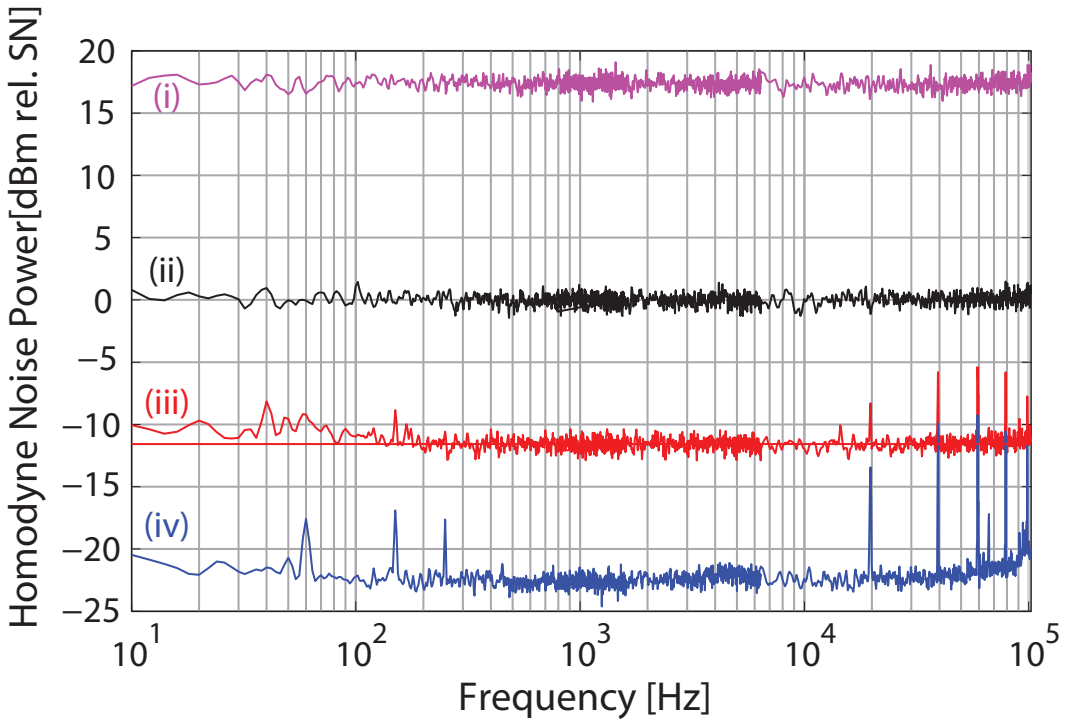


Figure 8.16: Measured squeezing from the ANU squeezer using quantum noise limited balanced homodyne detection showing the anti-squeezing (i), shot noise (ii), squeezing (iii), and dark noise (iv). A red line at -11.6 dB has been added to show the average of the squeezing from 1.6 kHz to 6.4 kHz. All traces are pieced together from 3 FFT windows: 0-1.6 kHz, 0-6.4 kHz, 0-102.4 kHz with 800 FFT lines, resulting in resolution bandwidths of 2 Hz, 8 Hz, and 128 Hz respectively. 100 RMS averages were taken for all traces. Dark noise has not been subtracted from the data.

8.4.6 Correcting for Detection Losses

Using the measured squeezing and anti-squeezing levels, we can infer the total loss on the state by assuming production of a pure state and determining the losses that result in the measured data. This value provides an estimate to the amount of squeezing injected into an interferometer, as the detection scheme is bypassed when the squeezed state is directed towards the interferometer. We can write for the measured squeezing, Sqz[dBm], and measured anti-squeezing, Asqz[dBm],

$$\begin{aligned} \text{Sqz[dBm]} &= 10 * \text{Log}_{10}\left[10^{\text{Sqz}_0/10}\eta_{tot} + (1-\eta_{tot})\right] \\ \text{Asqz[dBm]} &= 10 * \text{Log}_{10}\left[\frac{1}{10^{\text{Sqz}_0/10}}\eta_{tot} + (1-\eta_{tot})\right]. \end{aligned} \tag{8.12}$$

We then simultaneously solve these equations for the total detection efficiency, η_{tot}, and pure intra-cavity squeezing, Sqz_0. This equation accounts only for propagation losses and does not account for other detection issues such as phase jitter. The measured squeezing, Sqz[dBm], should be corrected for dark noise and the initial squeezing is the theoretical value of the squeezing before exiting the cavity. Multiplying Sqz_0 by the escape efficiency results in the amount of squeezing exiting the OPO. Using the measured values from

Figure 8.16, a total loss on the squeezed state of 0.048±0.004 is found, which agrees well with the expected losses. After accounting for the escape efficiency of 0.985 ± 0.001, the amount of squeezing exiting the cavity is found to be approximately 15.0±0.1 dB.

Chapter 9

The LIGO Injection

In this section the LIGO injection experiment is introduced. This project was a large international effort most notably involving the LIGO Scientific Collaboration, the ANU, the Max Planck Institute for Gravitational Physics, Caltech and MIT. The main focus of the input from the ANU was to supply a suitable squeezer, and as such, detail is directed towards the performance and properties of the squeezed light source. This squeezed light source was designed and constructed by Sheon Chua, Conor Mow-Lowry, Sheila Dwyer and I at the ANU. My personal responsibilities were in the design of the cavity geometry, modelling of the non-linear interactions, and investigation of noise sources in the balanced homodyne detection scheme.

The results of the LIGO injection experiment are then revealed. The injection experiment involved many people from all of the aforementioned institutions and will be detailed in the paper soon to be released on this work [153]. My contribution to this experiment was to firstly produce, as part of the group at the ANU, a travelling-wave squeezer that met strict performance benchmarks. Once this was achieved at the ANU, I then spent 6 months at MIT with Sheila Dwyer under the supervision of Prof. N. Mavalvala and Dr. D. Sigg building and characterising a replica of the ANU system. Finally, I spent approximately 2 months aiding in the integration of this squeezer with LIGO in Washington state. During this time I aided in integrating and characterising the squeezer, integrating the control electronics and taking the first measurements of the squeezing enhanced interferometer.

Further information on noise sources, scattered light measurements and interferometer performance with squeezed light injection will be given in more detail by Sheon Chua and Sheila Dwyer in their theses as well as future LIGO working notes and additional papers that are currently under construction.

9.0.7 LIGO Injection Aims

Squeezed light has already been shown to reduce the shot noise in interferometers, including the GEO600 device [68]. The aim of the LIGO H1 injection was not to demonstrate that squeezing could reduce shot noise, instead the goals were to investigate noise, scattered light and integration issues in the more sensitive LIGO interferometer. In particular, the sensitivity of LIGO at frequencies below about 500 Hz allows for investigation of these phenomena at low frequencies, inaccessible in previous experiments. With LIGO H1, it could be ensured that the reduction of shot noise through the injection of squeezed light did not degrade the performance of the device at lower frequencies. It was also decided that it would be beneficial to test a travelling wave bow-tie OPO geometry to complement the findings from GEO600, where a linear OPO geometry was used [100, 120]. The

experimental setup for the injection of squeezed states into H1 is shown in Figure 9.1.

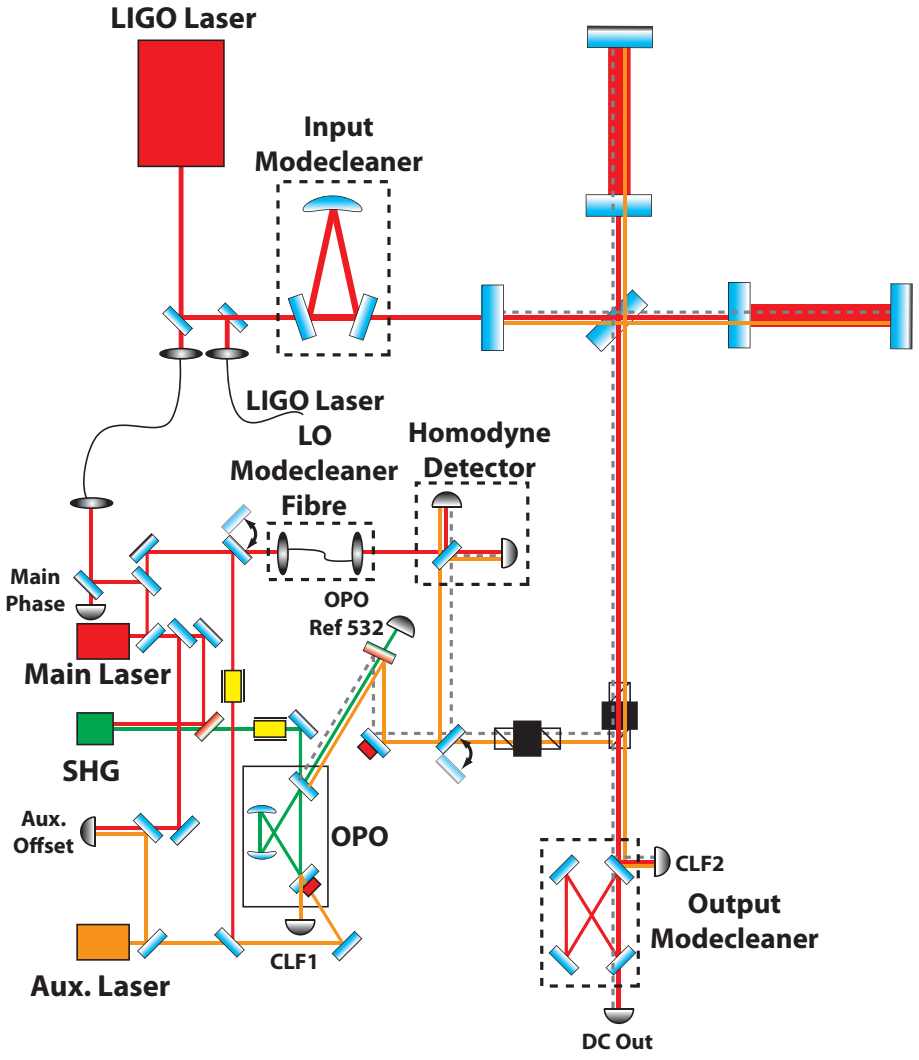


Figure 9.1: Schematic of the squeezed light injection experiment into H1 at LIGO. Pump field is in green, main laser and LIGO laser are both in red, the auxiliary laser is in orange and the squeezing is denoted by a dashed grey line. Full description in text.

The LIGO injection scheme involves three laser sources. The LIGO laser [154], the main squeezing laser (a decommissioned LIGO laser source operating without the amplifier [155]), and an auxiliary laser (Lightwave) for coherent control of the vacuum squeezed state. The phase of the main laser is locked to the phase of the LIGO laser by directing a tap-off from the LIGO laser into a fibre and beating this with a small amount of light from the main laser. Most of the main laser power is directed towards an SHG cavity, where the 1064 nm light is frequency doubled to 532 nm. The SHG cavity was provided by the Max Planck Institute for Gravitational Physics using the same design as that described in [143]. The pump field is incident on the coupling mirror of the OPO. The 532 nm field is then transmitted through a dichroic mirror while the squeezing and auxiliary beams are directed towards either the homdodyne detector or the LIGO interferometer. The transmitted 532 nm light is detected in order to derive a PDH locking signal for the cavity

length. The light from the main laser which is not used for SHG can be passed through a fibre optic cable and used as the local oscillator for the balanced homodyne detector. The fibre optic cable works in much the same way as the modecleaning cavity introduced in Chapter 8 to reduce beam pointing noise. Alternatively, a tap-off from the LIGO laser, directed through a second 100 m long fibre, could be used as the local oscillator for the homodyne detector by swapping the fibre connections at the end of modecleaner fibre setup, this fibre is marked as the LIGO laser LO. The LO from the LIGO laser could ideally be used to characterise the phase lock between the main laser and the LIGO laser when measuring squeezing, although the fibre noise in this LO plays a significant role. With the use of a flipper mirror the main laser field, instead of being used as the LO, can be directed towards the coherent locking coupling mirror in order to seed the OPO for gain and visibility measurements. A tap-off is also used to lock the frequency offset of the auxiliary laser to the main laser.

The auxiliary laser, which is frequency offset locked to the main laser, enters the OPO cavity through the coherent locking mirror. Within the cavity, it senses the non-linear interaction and a sideband field is generated. The reflected auxiliary laser and some small portion of the intra-cavity field is then detected at CLF1 and used to derive the first half of the coherent locking scheme. The auxiliary field and the generated sideband field also transmit through the OPO input coupler, along with the squeezed field, propagating through one Faraday isolator and reflecting off a second isolator, entering the interferometer. When it returns from the interferometer, the coherent locking field reflects off the output mode cleaner and is detected at CLF2. This field is used to derive the locking signal for the second half of the coherent locking. The squeezing passes through the output mode cleaner and is detected at the DC output of the interferometer. This detector is where the gravitational wave signals are also detected. The presence of squeezed light at this detector reduces the shot noise, thus increasing the sensitivity of the interferometer.

9.1 LIGO Squeezer Properties

The LIGO cavity was constructed taking into account the experience provided by the ANU prototype. The major differences were the purchase of higher quality cavity mirrors and the reflectivity of both the pump and squeezed fields on the input/output coupler. It was decided that the high escape efficiency of the ANU squeezer was not necessary for the LIGO cavity, seeing as losses inside the interferometer would limit the enhancement of the squeezing anyway, and as such the reflectivity of the input coupler at the fundamental was increased slightly. Whilst this decreases the escape efficiency, it also decreases the threshold power, reducing photothermal effects and increasing the thermal stability of the OPO slightly. Unfortunately, the modified coherent locking scheme was not developed in time to implement it in this experiment.

9.1.1 Mirror Properties

For reference, the properties of some of the intra-cavity mirrors and coatings bought from Advanced Thin Films are shown in Table 9.1. These mirrors were coated using the Ion Beam Sputtering (IBS) method and are of higher quality than those used in the ANU experiment.

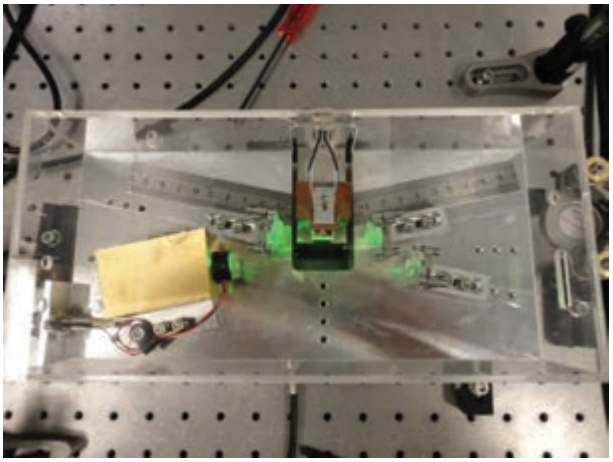


Figure 9.2: The LIGO squeezer, an almost exact replica of the ANU squeezer design.

Mirror Parameter	Value	Units
532 nm HR Reflectivity	>99.99%	-
1064 nm HR Reflectivity	>99.9%	-
532 nm AR Reflectivity	<0.2%	-
1064 nm AR Reflectivity	<0.1%	-
1064 HR Coating Absorption	<0.001%	-
RoC	-38	mm
Substrate Surface Roughness (Curved Surface)	<0.1	nm

Table 9.1: Specified mirror properties for the curved HR OPO mirrors from the manufacturer. The back of the mirrors are AR coated to allow for the injection of fields through these mirrors. All coatings are for an angle of incidence of 6 degrees.

9.1.2 Threshold

The threshold power of the system was measured using the same method as introduced in § 7.1.2. The mode matching of the pump field to the cavity was 74.5 %. The best fit to this data results in a threshold of 75.0±0.1 mW for the amplification regime and a threshold of 74±0.1 mW for the de-amplification regime. Using this information, we can determine the non-linear interaction strength, g, using Equation 5.36. We find that the non-linear interaction strength is 1880 $\pm 9\, s^{-\frac{1}{2}}$, which is very similar to the value found for the ANU squeezer. We expect this to be the case owing to the fact that the geometry of the two squeezers is nearly identical.

9.1.3 Losses

Using the same method as that shown for the ANU squeezer (see § 7.1.3), the reflected power of a seed directed into the front coupler was found to be 84.8 % after correcting for mode matching and dark noise. This results in an intra-cavity loss of 0.0056 ± 0.0001 for the squeezed field. The loss of this cavity was also measured by constructing the cavity without the crystal. The reflected power from the cavity without the crystal loss was 88.8 % after correcting for mode matching and dark noise, resulting in an intra-cavity loss of 0.0041 ± 0.0001. From these values, we can infer that the total crystal loss, consisting

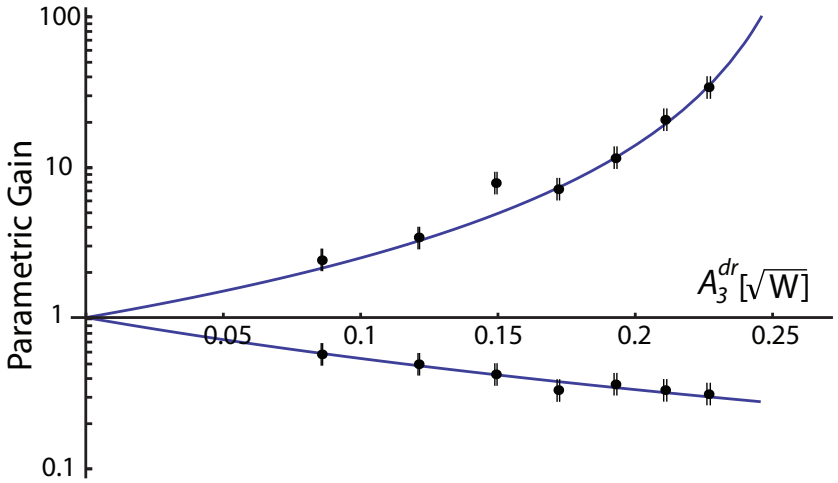


Figure 9.3: Parametric gain as a function of pump amplitude for the LIGO squeezer. The points are the data and the blue curve is the best fit to Equation 5.52.

of the two AR coatings and the absorption, is 0.0015 ± 0.0002. These losses, coupled with the reflectivity of the coupling mirror, of 0.868 ± 0.009, result in an escape efficiency of the cavity of 0.961 ± 0.002. The uncertainty is dominated by the uncertainty in the measurement of the coupling mirror reflectivity by the power meter.

The same process was repeated for the pump field. The reflected power of the pump directed into the front coupler was found to be 61.8 % after correcting for mode matching and dark noise. This implies an intra-cavity loss of 0.039 ± 0.001 for the pump field. The reflected power of the cavity without the crystal loss was 72.7 % after correcting for mode matching and dark noise. This results in an intra-cavity loss of 0.026 ± 0.001. From these values, we can infer that the total crystal loss for the pump, consisting of the two AR coatings and the absorption, is 0.013 ± 0.002.

9.1.4 Cavity Parameters

Table 9.2 shows the parameters of the LIGO squeezed light source. The main difference between this cavity and the prototype ANU cavity is the increase in reflectivity of the input/output coupler at the fundamental wavelength. As expected, this reduced the threshold of the LIGO cavity to 75 mW but also decreased the escape efficiency to 96.1 %.

9.1.5 Squeezing

The squeezing produced by the LIGO squeezed light source is illustrated in Figure 9.4. The squeezer was operated with 850 μW of local oscillator and 50 mW of pump light. The homodyne detector is the same as that used in the ANU squeezing experiment, with ETX500 diodes, and the dark noise lies greater than 20 dB below shot noise.

Figure 9.4 shows that the squeezing plateaus at 5.6 ± 0.4 dB below the shot noise above 1 kHz and the anti-squeezing is 8.8 ± 0.2 dB above shot noise. Using the methods detailed in §8.4.6, the total detection efficiency is found to be 81.4 %. This implies that there is 8.4 ± 0.1 dB of squeezing exiting the cavity. Much of this loss can be attributed to the ETX photodiodes, which we expect to have a quantum efficiency anywhere from 86 % to

Cavity Parameter	Symbol	Value	Uncert.	Units
Fundamental Wavelength	λ_1	1064	-	nm
Second Harmonic Wavelength	λ_3	532	-	nm
Input/Output Coupler Fundamental Reflectivity	$R_1^{in/out}$	0.868	0.009	-
Input/Output Coupler Pump Reflectivity	$R_3^{in/out}$	0.696	0.002	-
Coherent Locking Mirror Reflectivity	R_1^{cl}	0.998	-	-
PPKTP Loss for Fundamental	T_1^{cr}	0.0015	0.0002	-
PPKTP Loss for Pump	T_3^{cr}	0.013	0.002	-
Total Intra-Cavity Loss for Fundamental	T_l^a	0.0056	0.0001	-
Total Intra-Cavity Loss for Pump	T_l^b	0.039	0.001	-
Finesse at Fundamental	$\mathcal{F}_1$	42.5	1.5	-
Finesse at Pump	$\mathcal{F}_3$	15.6	0.1	-
Linewidth at Fundamental	$\Delta\nu_1$	26	1	MHz
Linewidth at Pump	$\Delta\nu_3$	70.0	0.7	MHz
Auxiliary Laser Detuning	Δ_{Aux}	29.8	-	MHz
Optical Path Length	L	0.279	0.001	m
Threshold Power	P_3^{crit}	75	0.1	mW
Escape Efficiency	η_{esc}	0.961	0.002	-
Non-linear Coupling Strength	g	1880	9	$s^{-\frac{1}{2}}$
Curved Mirror Radius of Curvature	RoC	-38	-	mm

Table 9.2: Properties of the LIGO squeezer. The loss values given are the transmission values of loss equivalent high reflectivity mirrors. Uncertainties are given for measured values.

96 % (See § 8.4.5). The degradation in squeezing below about 1 kHz is believed to be due to parasitic interference, however, due to time constraints on the project there was not enough time to hunt down the noise sources at these frequencies.

9.2 LIGO Enhancement

Figure 9.5 shows the LIGO strain sensitivity both with, blue trace (ii), and without, red trace (i), squeezed light enhancement. The sensitivity of LIGO has been increased by up to 2 dB in the shot noise limited regime. As expected, the amount of sensitivity improvement diminishes at lower frequencies. Importantly, the sensitivity of the device is not seen to decrease in the low-frequency regime, below about 200 Hz. This is an important result. It shows that scatter, electronic noise couplings and any other noise coupling mechanism has not degraded LIGO's performance at low frequencies. It also indicates that the increase in radiation pressure, the dominating quantum noise source at these frequencies, due to the presence of phase quadrature squeezing is still well below other noise sources. This is not expected to be the case for aLIGO, as shown in Figure 2.5.

9.2.1 Injection Losses

The main limitation to the amount of sensitivity improvement in this experiment was the losses on the squeezed state within the interferometer. One of the major sources of loss were the Faraday isolators in the path of the squeezed field. Each of the Faraday isolators, shown in Figure 9.1, introduced a 6 % loss upon transmission. The Faraday isolator in the LIGO output path is also traversed twice by the squeezed field, doubling this loss.

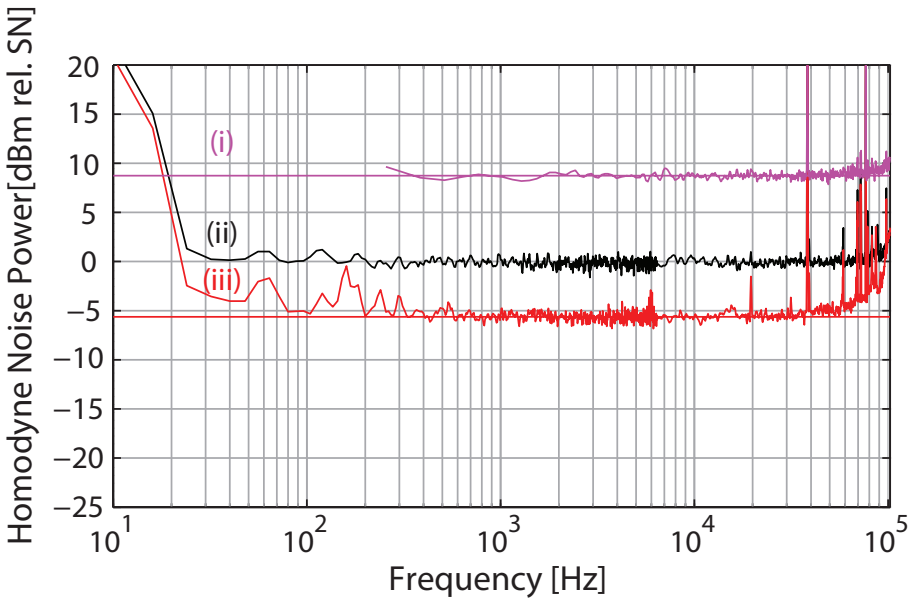


Figure 9.4: Measured squeezing from the LIGO squeezer using a quantum noise limited balanced homodyne detector showing the anti-squeezing (i), shot noise (ii), and squeezing (iii). A red line at -5.6 dB has been added to show the average of the squeezing from 1 kHz to 6 kHz. The shot noise and squeezing traces are pieced together from 2 FFT windows: 0-6.4 kHz, and 0-102.4 kHz with 800 FFT lines, resulting in resolution bandwidths of 8 Hz and 128 Hz respectively. The additional noise seen below 30 Hz is due to DC noise entering the measurements due to the large resolution bandwidths. The anti-squeezing trace has a 0-102.4 kHz bandwidth with 800 FFT lines and a bandwidth of 128 Hz. The first few data points have been removed. 100 RMS averages were taken for all traces. Dark noise has not been subtracted from the data.

The propagation efficiency between the squeezing cavity and the output mode cleaner was measured both in chamber with a power meter at various locations and using the power calibrated CLF2 diode from Figure 9.1 whilst under vacuum. The propagation efficiency determined by both of these measurements agreed and was found to be approximately 80 %.

The element producing the most loss was the output mode cleaner. The transmission through this cavity was measured for a well matched mode and was found to be 82 %. Mode matching to this cavity is difficult because it is near impossible to mode match with the interferometer at full power, therefore it is done at low power and then the power is increased to operational levels. Thermal effects, particularly thermal lensing, from the increased circulating laser power will alter the mode matching when moving from low power to standard operating conditions. The mode matching was measured in a low power configuration to be 74 %. These losses result in a total loss on the squeezed state of approximately 48 %. As shown in Figure 3.8, the most squeezing we could ever hope to measure with a 50 % loss is 3 dB. A more complete summary of these results is expected to be available in the near future [156].

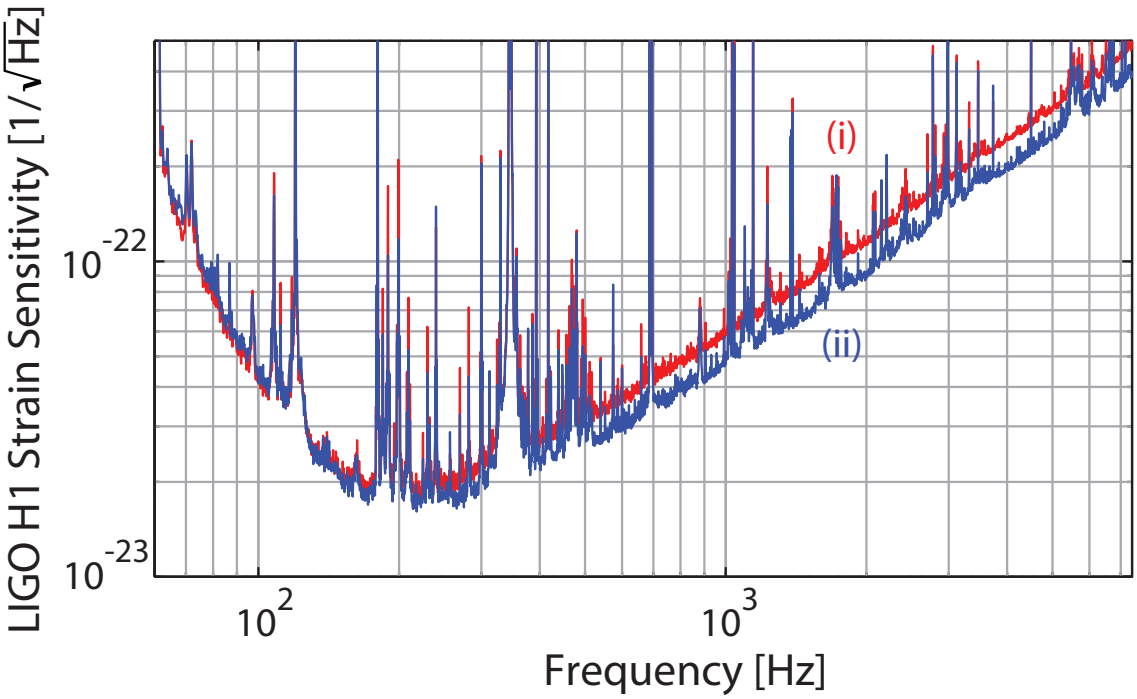


Figure 9.5: LIGO H1 detector squeezing sensitivity without, (i), and with, (ii), squeezing enhancement. Trace (i) shows the typical 6$^{\text{th}}$ science run (S6) operating sensitivity.

9.2.2 Travelling Wave Cavity Backscatter Isolation

It was previously mentioned that each pass through the Faraday isolator resulted in an approximately 6 % loss on the squeezed state. Figure 9.1 shows that the squeezing passes through one Faraday isolator before it enters the beam path of the interferometer's field. This isolator is placed in this location in order to reduce backscatter from the interferometer travelling towards the squeezer. The isolator is critical in ensuring that no scattered light re-enters the interferometer, which would reduce the sensitivity of the device.

In the GEO600 squeezed light experiment, where a linear OPO was used, two Faraday isolators were required in this path to reduce the scattering to an acceptable level [68]. The second Faraday isolator adds additional loss to the squeezed state, reducing the amount of squeezing injected into the interferometer. The travelling wave design used in the LIGO H1 experiment provides isolation to backscattering of greater than 40 dB, which is on par with the isolation offered by a single high-quality large-aperture Faraday isolator [102, 157]. The advantage of the travelling wave cavity is that this isolation is gained without introducing the additional losses of the Faraday isolator. As squeezing becomes a common technique for enhancing the sensitivity of gravitational-wave detectors, effort will be focussed towards decreasing the loss on the squeezed state. As the losses within the interferometer are reduced, the loss due to the Faraday isolators will become increasingly significant.

Chapter 10

Conclusions and Further Work

In this chapter, brief summary of the key results resulting from the ANU squeezer experiment and the H1 LIGO squeezed light injection are provided. From these results, a number of future goals and investigations are identified.

10.1 Summary of Squeezed Light Generation and Control for Gravitational-Wave Detection

- An optical parametric oscillator was designed and constructed as a prototype for the H1 squeezing experiment. The cavity was different to previous designs used in the enhancement of interferometers in that it was a doubly resonant, travelling wave system. Importantly, the travelling wave design was expected to provide a high level of immunity to backscattered light from integration into an interferometer and indeed this was found to be the case. A detailed discussion of the final cavity parameters was presented, and with them the reasoning behind choosing these parameters. Even with the additional intra-cavity loss introduced by the travelling wave design, this squeezer attained record levels of squeezing in the audio band. Down to 200 Hz the squeezing was directly observed at 11.6 dB below shot noise. Parasitic interference degraded the squeezing below this frequency but even so, 10 dB was still directly observed at 10 Hz. The progression of the squeezing results over the life of the experiment was presented and the steps taken to improve these results were discussed.

- A brief investigation of coherent locking was presented. Using the cavity equations of motion, the coherent locking scheme was modelled as a non-degenerate optical parametric oscillator and, using this method, the field amplitudes and phases were determined. It was found that one half of the coherent locking scheme was limiting the stability of the entire system and an alternative solution to this locking loop proposed. This alternative solution required second harmonic generation of the coherent locking field but allowed for the optimisation of this lock without negatively affecting the squeezing, as was found to be the case for the original scheme.

- An investigation into the various noise sources limiting low-frequency balanced homodyne detection was presented. Most of these noise sources had previously been identified but this work discussed all known noise sources and how they relate to each other. A method for detecting and eliminating parasitic interference was presented and the technique was experimentally shown to work as expected. The methods used to overcome the remaining noise sources was presented.

- Finally, a replica of the ANU squeezer was built in the United States and later moved to the LIGO Hannford site. The team at the LIGO site then installed and operated this squeezer in order to enhance the sensitivity of the device to levels never before seen. The noise couplings and integration issues were investigated by the science team at Hanford and will be detailed in later work and the theses of Sheon Chua and Sheila Dwyer.

10.2 Further Work

With the work presented in this thesis, a squeezing goal set around a decade ago has been reached [60, 49]. Squeezing of 10 dB down to frequencies of 10 Hz have been directly observed and should soon become readily reproducible. With the current amount of loss in gravitational-wave detectors, squeezed states of greater magnitude provide very little additional benefit over states such as the one presented here. This illuminates the fact that progression in this field of squeezed states for gravitational-wave detection needs to move towards the investigation of integration and stability, rather than squeezing magnitude. A number of issues that will need to be addressed in the near future are discussed below.

10.2.1 Squeezing Stability

As explained in § 7.4.2, the coherent locking field does not strictly lock to the angle of squeezing. It locks to some arbitrary angle that we choose to align to the angle of the squeezing. It was also discussed that the phase reference can drift, with changes in the temperature of the non-linear medium for example, and this will reduce the amount of squeezing enhancement in the interferometer. Ideally, the squeezer should provide stable squeezing for as long as possible in order to minimise the impact of integration on a fully operational gravitational wave detector. This will require further investigation along the lines of work previously conducted by Khalaidovski [140], possible introduction of a digital control loop to ensure the squeezed light angle does not drift over time, or perhaps the development of a new locking scheme. Further investigation into the coherent locking model presented in this thesis will aid in understanding the direction that needs to be taken.

One possible avenue for improving squeezed state control is to investigate the integration of a phase-matching locking loop into the system [158]. This technique reads out the phase matching condition from within the non-linear medium and if it were somehow integrated, perhaps by use of the frequency shifted auxiliary laser, then it might be able to increase the temperature stability of the system, thus stabilising the angle of the squeezing [140].

10.2.2 Dithering to Remove Parasitic Interference

It was shown in § 8.3.4 that noise due to parasitic interference could be averaged out at lower frequencies and shift the noise up to higher frequencies by dithering the phase of the scattered light. In this case, a mirror was placed in the beam path of the squeezing and therefore dithering of this mirror whilst attempting to measure squeezing would introduce a source of phase jitter, reducing the measured squeezing. By placing a second mirror in the path and feeding-forward the dither to the local oscillator phase, it should be possible to have the squeezing and local oscillator follow each other. This would allow one to

measure the full amount of squeezing, whilst still removing the noise due to the scattered light from the detection band. An investigation into whether this technique would be viable in a gravitational wave detector would be beneficial.

10.2.3 Noise Coupling of the Coherent Locking Field

It was mentioned in § 7.4.1 that as the coherent locking field power was increased, the squeezing was seen to degrade. The reason for this is believed to be due to the presence of intensity noise on the coherent locking field, but further investigation into the matter is required. By understanding how the power of this field leads to degradation in the squeezing, it might be possible to increase the performance of the coherent locking scheme by finding ways in which the generated field can be maximised.

10.2.4 Frequency Dependent Squeezing

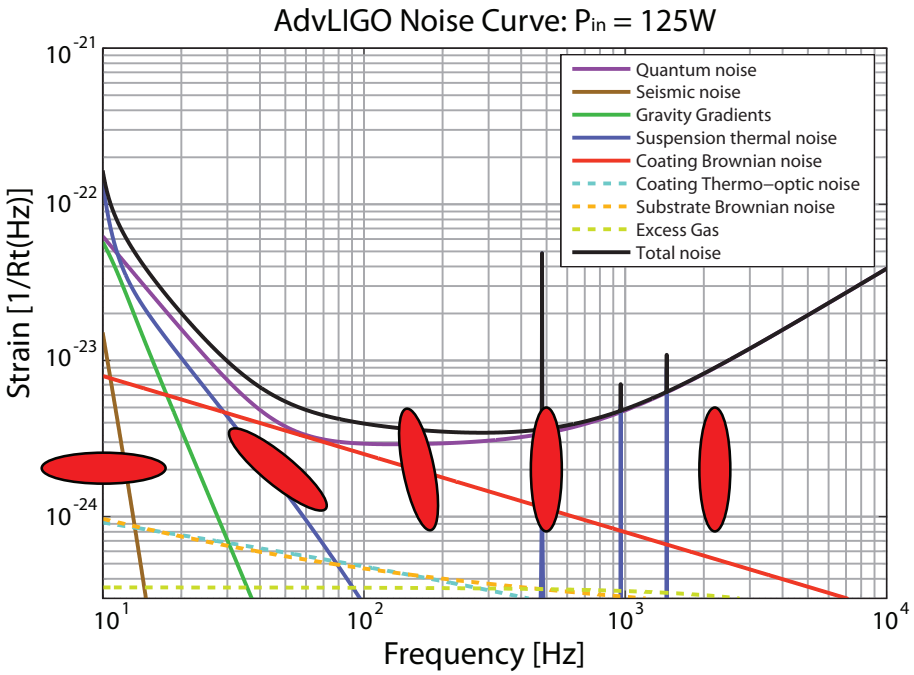


Figure 10.1: The Advanced LIGO expected noise budget showing roughly how frequency dependent squeezing would integrate. The red ellipses indicate the angle of the injected squeezing at the corresponding frequencies, where the phase quadrature is in the x-axis. At high frequencies the squeezed state is squeezed in the phase quadrature, whilst at lower frequencies the squeezing rotates and eventually aligns in the amplitude quadrature. With such a frequency dependent state, the quantum noise can be reduced at all frequencies.

Finally, now that ideal squeezing magnitudes have been achieved, and as the stability and noise couplings are better understood, the next logical step is the production of frequency dependent squeezing. In this work, the squeezing was uniform across all frequencies

and as such, decreasing the shot noise at high frequencies also increases the radiation pressure at low frequencies. Currently, radiation pressure noise is not the limiting noise source at low frequencies and as such increasing its magnitude by introducing phase squeezing has little effect. In the next generation of detectors. however, the radiation pressure noise is expected to be the dominant noise source below many tens of Hertz. Hence, the introduction of a phase squeezed state would improve the high frequency performance, as shown in this thesis, but would also decrease the low-frequency performance of the device.

The exception to this is if frequency dependent squeezed light is utilised, as illustrated in Figure 10.1. Such a state would exhibit amplitude squeezing below some frequency, and would gradually rotate to phase squeezing at higher frequencies. It seems as though this is the next big challenge facing squeezing for gravitational wave detection. It has been shown that filter cavities can rotate the squeezing in the necessary way, but that this will require very large cavities [51]. It needs to either be shown that this technique is practical or another method for rotating the squeezing needs to be developed. A new technique will require a system that can provide dispersion at very low frequencies without introducing substantial loss, and very few, if any, systems currently exist with both of these properties. However, with the squeezing magnitude problem solved, the stability currently being worked on, and the noise coupling currently under investigation from the H1 squeezing experiment and also at GEO600, effort now needs to be directed towards producing these frequency dependent states.

Appendix 1

We begin at Equations 5.39

$$\begin{aligned} i\omega\delta\check{X}_1 &= -\kappa_1\delta\check{X}_1 + g\alpha_1\delta\check{X}_3 + g\alpha_3\delta\check{X}_1^{\dagger} + \sqrt{2\kappa_1^{in}}\delta\check{X}_1^{dr} + \sqrt{2\kappa_1^{out}}\delta\check{X}_1^{vout} + \sqrt{2\kappa_1^{l}}\delta\check{X}_1^{vl} \\ i\omega\delta\check{X}_3 &= -\kappa_3\delta\check{X}_3 - g\alpha_1\delta\check{X}_1 + \sqrt{2\kappa_3^{in}}\delta\check{X}_3^{dr} + \sqrt{2\kappa_3^{out}}\delta\check{X}_3^{vout} + \sqrt{2\kappa_3^{l}}\delta\check{X}_3^{vl}. \end{aligned} \tag{10.1}$$

We solve this set of equations by first rearranging these equations into matrix form

$$\begin{aligned} \begin{pmatrix} \delta\check{X}_1^{dr} \\ \delta\check{X}_3^{dr} \end{pmatrix} &= \begin{pmatrix} \frac{i\omega+\kappa_1-g\alpha_3}{\sqrt{2\kappa_1^{in}}} & \frac{-g\alpha_1}{\sqrt{2\kappa_1^{in}}} \\ \frac{g\alpha_1}{\sqrt{2\kappa_1^{in}}} & \frac{i\omega+\kappa_3}{\sqrt{2\kappa_3^{in}}} \end{pmatrix} \begin{pmatrix} \delta\check{X}_1 \\ \delta\check{X}_3 \end{pmatrix} + \\ & \begin{pmatrix} -\frac{\sqrt{2\kappa_1^{out}}}{\sqrt{2\kappa_1^{in}}}\,\delta\check{X}_1^{vout} - \frac{\sqrt{2\kappa_1^{l}}}{\sqrt{2\kappa_1^{in}}}\,\delta\check{X}_1^{vl} \\ -\frac{\sqrt{2\kappa_3^{out}}}{\sqrt{2\kappa_3^{in}}}\,\delta\check{X}_3^{vout} - \frac{\sqrt{2\kappa_3^{l}}}{\sqrt{2\kappa_3^{in}}}\,\delta\check{X}_3^{vl} \end{pmatrix}. \end{aligned} \tag{10.2}$$

The next step is to solve for the intra-cavity fields, resulting in

$$\begin{aligned} \begin{pmatrix} \delta\check{X}_1 \\ \delta\check{X}_3 \end{pmatrix} &= \begin{pmatrix} \frac{\sqrt{2\kappa_1^{in}}(\kappa_3+i\omega)}{\alpha_1\alpha_1 g^2-(g\alpha_3-\kappa_1-i\omega)(\kappa_3+i\omega)} & \frac{\sqrt{2\kappa_3^{in}}\,g\alpha_1}{\alpha_1\alpha_1 g^2-(g\alpha_3-\kappa_1-i\omega)(\kappa_3+i\omega)} \\ -\frac{\sqrt{2\kappa_1^{in}}\,g\alpha_1}{\alpha_1\alpha_1 g^2-(g\alpha_3-\kappa_1-i\omega)(\kappa_3+i\omega)} & \frac{\sqrt{2\kappa_3^{in}}(g\alpha_3-\kappa_1-i\omega)}{-\alpha_1\alpha_1 g^2+(g\alpha_3-\kappa_1-i\omega)(\kappa_3+i\omega)} \end{pmatrix}. \\ & \begin{pmatrix} \delta\check{X}_1^{dr} + \frac{\sqrt{2\kappa_1^{out}}}{\sqrt{2\kappa_1^{in}}}\,\delta\check{X}_1^{vout} + \frac{\sqrt{2\kappa_1^{l}}}{\sqrt{2\kappa_1^{in}}}\,\delta\check{X}_1^{vl} \\ \delta\check{X}_3^{dr} + \frac{\sqrt{2\kappa_3^{out}}}{\sqrt{2\kappa_3^{in}}}\,\delta\check{X}_3^{vout} - \frac{\sqrt{2\kappa_3^{l}}}{\sqrt{2\kappa_3^{in}}}\,\delta\check{X}_3^{vl} \end{pmatrix}. \end{aligned} \tag{10.3}$$

We now have a solution for the intra-cavity fields, allowing us to determine the output fields using the input/output relations from Equation 5.9, which are rewritten as

$$\delta\check{X}^{ref} = \sqrt{2\kappa^{in}}\,\delta\check{X} - \delta\check{X}^{dr} \tag{10.4}$$

$$\delta\check{X}^{trans} = \sqrt{2\kappa^{out}}\,\delta\check{X} - \delta\check{X}^{vout}. \tag{10.5}$$

At this point we no longer need to investigate the properties of the pump field and so we will only calculate the output field properties of the squeezed field, resulting in

$$\begin{aligned} \delta\check{X}_1^{out} &= \begin{pmatrix} \frac{2\sqrt{\kappa_1^{in}\kappa_1^{out}}(\kappa_3+i\omega)}{\alpha_1\alpha_1 g^2-(g\alpha_3-\kappa_1-i\omega)(\kappa_3+i\omega)} \\ \frac{2\sqrt{\kappa_3^{in}\kappa_1^{out}}\,g\alpha_1}{\alpha_1\alpha_1 g^2-(g\alpha_3-\kappa_1-i\omega)(\kappa_3+i\omega)} \end{pmatrix}. \\ & \begin{pmatrix} \delta\check{X}_1^{dr} + \left(\frac{\sqrt{2\kappa_1^{out}}}{\sqrt{2\kappa_1^{in}}} - (\alpha_1\alpha_1 g^2 - (g\alpha_3-\kappa_1-i\omega)(\kappa_3+i\omega))\right)\delta\check{X}_1^{vout} + K \\ \delta\check{X}_3^{dr} + \frac{\sqrt{2\kappa_3^{out}}}{\sqrt{2\kappa_3^{in}}}\,\delta\check{X}_3^{vout} - \frac{\sqrt{2\kappa_3^{l}}}{\sqrt{2\kappa_3^{in}}}\,\delta\check{X}_3^{vl} \end{pmatrix}, \end{aligned}$$

where K is given by $\mathrm{K} = \frac{\sqrt{2\kappa_1^{l}}}{\sqrt{2\kappa_1^{in}}}\,\delta\check{X}_1^{vl}$.

At this point, these equations can be used to investigate how noise from the two driving fields, the pump and seed fields, couples into the output squeezed state. Up until now, the coherent amplitude of the seed field has not been removed in order to illustrate a point. From Equation 10.6 we see that the two elements in the first matrix can be regarded as coupling constants for the elements in the second matrix. Of note is that the first element of the second matrix consists entirely of noise terms from the seed field, and the second element of this matrix consists entirely of noise terms from the pump field. One can write

$$\delta\check{X}_1^{out} = \begin{pmatrix} C1 & \alpha_1 C2 \end{pmatrix} . \begin{pmatrix} \delta 1 \\ \delta 3 \end{pmatrix},$$

where;

$$\begin{aligned}
C1 &= \frac{2\sqrt{\kappa_1^{in}\kappa_1^{out}}(\kappa_3 + i\omega)}{\alpha_1\alpha_1 g^2 - (g\alpha_3 - \kappa_1 - i\omega)(\kappa_3 + i\omega)}, \\
C2 &= \frac{2\sqrt{\kappa_3^{in}\kappa_1^{out}}\, g}{\alpha_1\alpha_1 g^2 - (g\alpha_3 - \kappa_1 - i\omega)(\kappa_3 + i\omega)}, \\
\delta 1 &= \delta\check{X}_1^{dr} + \left(\frac{\sqrt{2\kappa_1^{out}}}{\sqrt{2\kappa_1^{in}}} - (\alpha_1\alpha_1 g^2 - (g\alpha_3 - \kappa_1 - i\omega)(\kappa_3 + i\omega)) \right) \delta\check{X}_1^{vout} + K, \\
\delta 2 &= \delta\check{X}_3^{dr} + \frac{\sqrt{2\kappa_3^{out}}}{\sqrt{2\kappa_3^{in}}}\, \delta\check{X}_3^{vout} - \frac{\sqrt{2\kappa_3^{l}}}{\sqrt{2\kappa_3^{in}}}\, \delta\check{X}_3^{vl}.
\end{aligned}$$

$C1$ can be considered as the coupling constant for the sum of all noise terms originating from fields at the fundamental frequency, $\delta 1$, and in much the same way, $C2$ can be considered as the coupling constant for all noise terms at the pump frequency $\delta 3$. The explicit dependence of the coupling of the noise terms at the pump frequency on α_1 has been highlighted. It becomes clear that the noise of the pump field is completely decoupled from the noise of the output squeezed field when the seed field has no coherent amplitude, as in the case of the OPO. It was this insight that led to the first measurements of squeezing in the audio band [63]. The details of this noise coupling are discussed in detail by McKenzie [108].

The final step is to calculate the spectrum of the amplitude quadrature noise operator. Using Equation 3.7 we can write for the variance of the output field at the fundamental frequency, V_1^{out},

$$V_1^{out} = < \delta\check{X}_1^{out\dagger} \delta\check{X}_1^{out} > . \tag{10.6}$$

This is simple to calculate due to the fact that the average of the first order noise terms is zero. Due to this fact, the average of the multiplication of uncorrelated noise terms will also result in an answer of zero. Also, the variances of all of the vacuum terms are replaced with their value of unity. At this point the coherent amplitude of the seed field is set to zero. This results in a variance for the amplitude quadrature of the squeezed field of

$$V_{out}^{+} = 1 + \frac{4g\kappa_1^{out}\alpha_3}{(\kappa_1 - g\alpha_3)^2 + \omega^2},$$

where the the subscript 1 has been dropped because it is clear that the equation is for the variance of the squeezed field. The superscript has been reintroduced to specify that this equation is for the amplitude quadrature variance. The result shown in Equation 10.7 is equivalent to the solution found for the case of an OPA under the assumption that the pump field can be treated classically (see [83]). This is because, as we have just shown, the noise from the pump field does not couple to the output state in the absence of a bright seed field. The fact that noise does not couple from this field means that we can treat it classically to simplify the situation.

The phase quadrature variance can be calculated by following the same methods presented here. However, it is possible to derive the phase quadrature variance from the variance of the amplitude quadrature. It can be shown that in the case of the OPA, where a bright seed is present, phase squeezing corresponds to amplification of the seed field, and amplitude squeezing corresponds to de-amplification of the seed field. The system can be moved from amplification to de-amplification by rotating the phase of the pump field by 180 degrees. This is equivalent to changing the sign of all the terms containing the coherent amplitude of the pump field, resulting in a variance for the phase quadrature of

$$V_{out}^{-} = 1 - \frac{4g\kappa_1^{out}\alpha_3}{(\kappa_1 + g\alpha_3)^2 + \omega^2}.$$

Bibliography

[1] T. H. Maiman. Stimulated optical radiation in ruby masers. *Nature*, 187:493, 1960.

[2] D. F. Welch. A brief history of high-power semiconductor lasers. *IEEE J. Sel. Topics Quantum Electron.*, 6:1470, 2000.

[3] J. L. E. Dreyer. *The Scientific Papers of Sir William Herschel, Volume 1.* Thoemmes Continuum, Bristol, 2003.

[4] R. A. Ong. Very high-energy gamma-ray astronomy. *Phys. Rep.*, 305:93, 1998.

[5] A. R. Thompson, J. M. Moran, and G. W. Swenson. *Interferometry and synthesis in radio astronomy.* Wiley-Interscience, New York, 1986.

[6] A. Einstein. *Ann. Phys. Lpz*, 49:769, 1916.

[7] A. Einstein. *Preuss. Akad. Wiss. Berlin*, page 688, 1916.

[8] The LIGO Scientific Collaboration and the Virgo Collaboration. An upper limit on the stochastic gravitational-wave background of cosmological origin. *Nature*, 460:990, 2009.

[9] B. S. Sathyaprakash and B. F. Schutz. Physics, astrophysics and cosmology with gravitational waves. *Living Reviews in Relativity*, 12, 2009.

[10] B. C. Barish and R. Weiss. LIGO and the detection of gravitational waves. *Phys. Today*, 52:44, 1999.

[11] B. P. Abbott et al. Searches for gravitational waves from known pulsars with science run 5 LIGO data. *Astrophys. J.*, 713:671, 2010.

[12] J. Abadie et al. All-sky search for gravitational-wave bursts in the first joint LIGO-GEO-Virgo run. *Phys. Rev. D*, 81:102001, 2010.

[13] J. Abadie et al. Search for gravitational waves from compact binary coalescence in LIGO and Virgo data from S5 and VSR1. *Phys. Rev. D*, 82:102001, 2010.

[14] R. A. Hulse and J. H. Taylor. Discovery of a pulsar in a binary system. *Astrophys. J. (Letters)*, 195:L51, 1975.

[15] A. Einstein and N. Rosen. *J. Franklin Inst.*, 223:43, 1937.

[16] J. Weber. Detection and generation of graviational waves. *Phys. Rev.*, 117:306, 1960.

[17] A. Giazatto. Interferometric detection of gravitational waves. *Phys. Rep.*, 182:365, 1989.

[18] http://www.auriga.lnl.infn.it.

[19] J. Weber. Evidence for discovery of gravitational radiation. *Phys. Rev. Lett.*, 22:1320, 1969.

[20] K. S. Thorne. Gravitational-wave research: Current status and future prospects. *Rev. Mod. Phys.*, 52:285, 1980.

[21] A. de Waard, L. Gottardi, J. van Houwelingen, A. Shumack, and G. Frosatti. Minigrail, the first spherical detector. *Class. Quantum Grav.*, 20:143, 2003.

[22] M. Bartusiak. *Einstein's Unfinished Symphony: Listening to the Sounds of Space-Time*. Joseph Henry Press, Washington DC, 2000.

[23] A. Abramovici, W. E. Althouse, R. W. P. Drever, Y. Gursel, S. Kawamura, F. J. Raab, D. Shoemaker, L. Sievers, R. E. Spero, K. S. Thorne, R. E. Vogt, R. Weiss, S. E. Whitcomb, and M. E. Zucker. LIGO: the laser interferometer gravitational-wave observatory. *Science*, 256:325, 1992.

[24] W. A. Edelstein, J. Hough, J. R. Pugh, and W. Martin. Limits to the measurement of displacement in an interferometric gravitational radiation detector. *J. Phys. E: Sci. Instrum.*, 11:710, 1978.

[25] C. M. Caves. Quantum-mechanical noise in an interferometer. *Phys. Rev. D.*, 23:1693, 1981.

[26] P. F. Michelson. On detecting stochastic background gravitational radiation with terrestrial detectors. *Mon. Not. R. astr. Soc.*, 227:933, 1987.

[27] S. Rowan and J. Hough. Gravitational wave detection by interferometry (ground and space). *Living Reviews in Relativity*, 3, 2000.

[28] J. R. Smith for the LIGO Scientific Collaboration. The path to the enhanced and advanced LIGO gravitational-wave detectors. *Class. Quantum Grav.*, 26:114013, 2009.

[29] G. M. Harry for the LIGO Scientific Collaboration. Advanced LIGO: the next generation of gravitational wave detectors. *Class. Quantum Grav.*, 27:084006, 2010.

[30] B. F. Schutz. Gravitational-wave sources. *Class. Quantum Grav.*, 13:A219, 1996.

[31] B. P. Abbott et al. Search for gravitational waves from low mass binary coalescences in the first year of LIGO's S5 data. *Phys. Rev. D*, 79:122001, 2009.

[32] F. Acernese et al. The VIRGO interferometric gravitational antenna. *Opt. Lasers Eng.*, 45:478, 2007.

[33] B. Willke *et al.* The GEO 600 gravitational wave detector. *Class. Quantum Grav.*, 19:1377, 2002.

[34] C. M. Caves. Quantum limits on noise in linear amplifiers. *Phys. Rev. D*, 26:1817, 1982.

[35] A. M. Gretarsson, G. M. Harry, S. D. Penn, P. R. Saulson, W. J. Startin, S. Rowan, G. Cagnoli, and J. Hough. Pendulum mode thermal noise in advanced interferometers: a comparison of fused silica fibers and ribbons in the presence of surface loss. *Phys. Lett. A*, 270:108, 2000.

[36] Y. Levin. Internal thermal noise in the ligo test masses: A direct approach. *Phys. Rev. D*, 57:659, 1998.

[37] G. M. Harry, A. M. Gretarsson, P. R. Saulson, S. E. Kittelberger, S. D. Penn, W. J. Startin, S. Rowan, M. M. Fejer, D. R. M. Crooks, G. Cagnoli, J. Hough, and N. Nakagawa. Thermal noise in interferometric gravitational wave detectors due to dielectric coatings. *Class. Quantum Grav.*, 19:897, 2002.

[38] G. M. Harry, H. Armandula, E. Black, D. R. M. Crooks, G. Cagnoli, J. Hough, P. Murray, S. Reid, S. Rowan, P. Sneddon, M. M. Fejer, R. Route, and S. D. Penn. Thermal noise from optical coatings in gravitational wave detectors. *Appl. Opt.*, 45:1569, 2006.

[39] H. B. Callen and R. F. Greene. On a theorem of irreversible thermodynamics. *Phys. Rev.*, 86:702, 1952.

[40] K. Kuroda. Status of LCGT. *Class. Quantum Grav.*, 27:084004, 2010.

[41] K. Somiya. Detector configuration of KAGRA - the Japanese cyrogenic graviatational-wave detector. *arXiv:gr-qc/1111.7185v2*, 2012.

[42] S. Miyoki *et al.* Status of the CLIO project. *Class. Quantum Grav.*, 21:S1173, 2004.

[43] P. Fritschel. Second generation instruments for the laser interferometer gravitational wave observatory (LIGO). *arXiv:gr-qc/0308090v1*, 2003.

[44] M. G. Beker, G. Cella, R. DeSalvo, M. Doets, H. Grote, J. Harms, E. Hennes, V. Mandic, D. S. Rabeling, J. F. J. van den Brand, and C. M. van Leeuwen. Improving the sensitivity of future GW observatories in the 1-10 Hz band: Newtonian and seismic noise. 43:623, 2011.

[45] R. Abbott, R. Adhikari, G. Allen, S. Cowley, E. Daw, D. Debra, J. Giaime, G. Hammond, M. Hammond, C. Hardham, J. How, W. Hua, W. Johnson, B. Lantz, K. Mason, R. Mittleman, J. Nichol, S. Richman, J. Rollins, D. Shoemaker, G. Stapfer, and R. Stebbins. Seismic isolation for Advanced LIGO. *Class. Quantum Grav.*, 19:1591, 2002.

[46] S. Sato, S. Miyoki, S. Telada, D. Tatsumi, A. Araya, M. Ohashi, Y. Totsuka, M. Fukushima, and M. Fujimoto. Ultrastable performance of an underground-based laser interferometer observatory for gravitational waves. *Class. Quantum Grav.*, 60:102005, 2004.

[47] P. Amaro-Seone *et al.* Low-frequency gravitational-wave science with eLISA/NGO. *arXiv:gr-qc/1202.0839.v1*, 2012.

[48] F. Antonucci *et al.* LISA pathfinder: mission and status. *Class. Quantum Grav.*, 2008:094001, 2011.

[49] T. Corbitt and N. Mavalavala. Quantum noise in gravitational-wave interferometers. *J. Opt. B: Quantum Semiclass. Opt.*, 6, 2004.

[50] R. Lawrence, D. Ottaway, M. Zucker, and P. Fritschel. Active correction of thermal lensing through external radiative thermal actuation. *Opt. Lett.*, 29:2635, 2004.

[51] S. Chelkowski, H. Vahlbruch, B. Hage, A. Franzen, N. Lastzka, K. Danzmann, and R. Schnabel. Experimental characterization of frequency-dependent squeezed light. *Phys. Rev. A*, 71:013806, 2007.

[52] V. B. Braginsky, S. E. Strigin, and S. P. Vyatchanin. Parametric oscillatory instability in Fabry-Perot interferometer. *Phys. Lett. A*, 287:331, 2001.

[53] S. Ballmer, V. Frolov, R. Lawrence, W. Kells, G. Moreno, K. Mason, D. Ottaway, M. Smith, C. Vorvick, P. Willems, and M. Zucker. Thermal compensation system description. *Open access LIGO document LIGO-T050064-00-R*, 2005.

[54] R. Lawrence, M. Zucker, P. Fritschel, P. Marfuta, and D. Shoemaker. Adaptive thermal compensation of test masses in advanced ligo. *Class. Quantum Grav.*, 19:1803, 2002.

[55] N. Treps, U. Andersen, B. Buchler, P. K. Lam, A. Maître, H. A. Bachor, and C. Fabre. Surpassing the standard quantum limit for optical imaging using nonclassical multimode light. *Phys. Rev. Lett*, 88, 2004.

[56] E. S. Polzik, J. C. Carri, and H. J. Kimble. Spectroscopy with squeezed light. *Phys. Rev. Lett.*, 68:3020, 1992.

[57] R. E. Slusher, L. W. Hollberg, B. Yurke, J. C. Mertz, and J. F. Valley. Observation of squeezed states generated by four-wave mixing in an optical cavity. *Phys. Rev. Lett.*, 55:2409, 1985.

[58] H. J. Kimble. Squeezed states of light: an (incomplete) survey of experimental progress and prospects. *Phys. Rep.*, 219:227, 1992.

[59] K. McKenzie, D. A. Shaddock, D. E. McClelland, B. C. Buchler, and P. K. Lam. Experimental demonstration of a squeezing-enhanced power-recycled michelson interferometer for gravitational wave detection. *Phys. Rev. Lett.*, 88:231102, 2002.

[60] K. Goda. *Development of Techniques for Quantum-Enhanced Laser-Interferometric Gravitational-Wave Detectors*. PhD thesis, Massachusetts Institute of Technology, Cambridge, 2007.

[61] K. Goda, K. McKenzie, E. E. Mikhailov, P. K. Lam, D.E. McClelland, and N. Mavalvala. Photothermal fluctuations as a fundamental limit to low-frequency squeezing in a degenerate optial parametric oscillator. *Phys. Rev. A*, 72:043819, 2005.

[62] W. P. Bowen, R. Schnabel, N. Treps, H-A. Bachor, and P. K. Lam. Recovery of continuous wave squeezing at low frequencies. *J. Opt. B: Quantum Semiclass. Opt.*, 4:421, 1992.

[63] K. McKenzie, N. Grosse, W. P. Bowen, S. E. Whitcomb, M. B. Gray, D. E. McClelland, and P. K. Lam. Squeezing in the audio gravitational-wave detection band. *Phys. Rev. Lett.*, 93:161105, 2004.

[64] K. McKenzie, M. B. Gray, P. K. Lam, and D. E. McClelland. Technical limitations to homodyne detection at audio frequencies. *Appl. Optics*, 46:3389, 2007.

[65] S. Chua, K. McKenzie, B.C. Buchler, and D.E. McClelland. Impact of non-stationary events on low frequency homodyne detection. *Journ. Phys.: Conf. Ser.*, 122:012023, 2008.

[66] H. Vahlbruch, S. Chelkowski, B. Hage, A. Franzen, K. Danzmann, and R. Schnabel. Coherent control of vacuum squeezing in the gravitational-wave detection band. *Phys. Rev. Lett.*, 97:011101, 2006.

[67] S. Chelkowski, H. Vahlbruch, K. Danzmann, and R. Schnabel. Coherent control of broadband vacuum squeezing. *Phys. Rev. A*, 75:043814, 2007.

[68] The LIGO Scientific Collaboration. A gravitational wave observatory operating beyond the quantum shot-noise limit. *Nat. Phys.*, 2083, 2011.

[69] M. Xiao, L.-A. Wu, and H. J. Kimble. Precision measurement beyond the shot-noise limit. *Phys. Rev. Lett.*, 59:279, 1987.

[70] K. Goda, O. Miyakawa, E. E. Mikhailov, S. Saraf, R. Adhikari, K. McKenzie, R. Ward, S. Vass, A. J. Weinstein, and N. Mavalvala. A quantum-enhanced prototype gravitational-wave detector. *Nat. Phys.*, 4:472, 2008.

[71] N. Christensen. LIGO S6 detector characterization studies. *Class. Quantum Grav.*, 27:194010, 2010.

[72] F. Acernese *et al.* Virgo status. *Class. Quantum Grav.*, 25:184001, 2008.

[73] D. F. Walls and G. J. Milburn. *Quantum Optics*. Springer-Verlag, Berlin, 1994.

[74] H. R. Lewis Jr. Classical and quantums systems with time-dependent harmonic-oscillator-type hamiltonians. *Phys. Rev. Lett.*, 18:510, 1967.

[75] E. Hecht. *Optics*. Addison-Wesley, Reading, 2 edition, 1987.

[76] H.-A. Bachor. *"A guide to experiments in quantum optics"*. VCH Publishers, New York, 1998.

[77] T. Kane. Intensity noise in diode-pumped single-frequency Nd:YAG lasers and its control by electronic feedback. *IEEE Photon. Technol. Lett.*, 2:244, 1990.

[78] G. Jaeger. *Quantum Information: An Overview*. Springer, New York, 2006.

[79] C. Silberhorn, P. K. Lam, O. Weiß, F. König, N. Korolkova, and G. Leuchs. Generation of continuous variable Einstein-Podolsky-Rosen entanglement via the kerr nonlinearity in an optical fiber. *Phys. Rev. Lett.*

[80] N. B. Grosse, S. Assad, M. Mehmet, R. Schnabel, T. Symul, and P. K. Lam. Observation of entanglement between two light beams spanning an octave in optical frequency. *Phys. Rev. Lett.*, 100:243601, 2008.

[81] W. P. Bowen, R. Schnabel, P. K. Lam, and T. C. Ralph. Experimental investigation of criteria for continuous variable entanglement. *Phys. Rev. Lett.*, 90:043601, 2003.

[82] M. J. Collett, R. Loudon, and C. W. Gardiner. Quantum theory of optical homodyne and heterodyne detection. *J. Mod. Optic.*, 34:881, 1987.

[83] B.E. Buchler. *Electro-optic control of quantum measurements.* PhD thesis, The Australian National University, 2001.

[84] C. F. McCormick, V. Boyer, E. Arimondo, and P. D. Lett. Strong relative intensity squeezing by four-wave mixing in rubidium vapor. *Opt. Lett.*, 32:178, 2007.

[85] R. M. Shelby, M. D. Levenson, S. H. Perlmutter, R. G. DeVoe, and D. F. Walls. Broad-band parametric deamplification of quantum noise in an optical fiber. *Phys. Rev. L.*, 57:692, 1986.

[86] T. Eberle, S. Steinlechner, J. Bauchrowitz, V. Händchen, H. Vahlbruch, M. Mehmet, H. Müller-Ebhardt, and R. Schnabel. Quantum enhancement of the zero-area sagnac interferometer topology for gravitational wave detection. *Phys. Rev. Lett.*, 104:251102, 2010.

[87] M. S. Stefszky, C. M. Mow-Lowry, S. S. Y. Chua, D. A. Shaddock, B. C. Buchler, H. Vahlbruch, A. Khalaidovski, R. Schnabel, P. K. Lam, and D. E. McClelland. Balanced homodyne detection of optical quantum states at audio-band frequenices and below. *Class. Quantum Grav. (submitted)*, 2012.

[88] C. F. McCorick, A. M. Marino, V. Boyer, and P. D. Lett. Strong low-frequency quantum correlations from a four wave-mixing amplifier. *Phys. Rev. A*, 78:043816, 2008.

[89] C. Liu, J. Jing, Z. Zhou, R. C. Pooser, F. Hudelist, L. Zhou, and W. Zhang. Realization of low frequency and controllable bandwidth squeezing based on a four-wave-mixing amplifier in rubidium vapor. *Opt. Lett.*, 36:2979, 2011.

[90] L.-A. Wu, H. J. Kimble, J. L. Hall, and H. Wu. Generation of squeezed states by parametric down conversion. *Phys. Rev. Lett.*, 57:2520, 1986.

[91] E. S. Polzik, J. Carri, and H. J. Kimble. Atomic spectroscopy with squeezed light for sensitivity beyond the vacuum-state limit. *Appl. Phys. B*, 55:279, 1992.

[92] Z. Y. Ou, S. F. Pereira, H. J. Kimble, and K. C. Peng. Realization of the einstein-podolsky-rosen paradox for continuous variables. *Phys. Rev. Letts.*, 68:3663, 1992.

[93] K. Schneider, R. Bruckmeier, H. Hansen, S. Schiller, and J. Mlynek. Bright squeezed-light generation by a continuous-wave semimonolithic parametric amplifier. *Opt. Lett.*, 21:1396, 1996.

[94] K. Schneider, M. Lang, J. Mlynek, and S. Schiller. Generation of strongly squeezed continuous-wave light at 1064 nm. *Opt. Ex.*, 2:59, 1998.

[95] P. K. Lam, T. C. Ralph, B. C. Buchler, D. E. McClelland, H.-A. Bachor, and J. Gao. Optimization and transfer of vacuum squeezing from an optical parametric oscillator. *J. Opt. B*, 1:469, 1999.

[96] S. Suzuki, H. Yonezawa, F. Kannari, M. Sasaki, and A. Furusawa. 7db quadrature squeezing at 860 nm with periodically poled $KTiOPO_4$. *Appl. Phys. Lett.*, 89:061116, 2006.

[97] Y. Takeno, M. Yukawa, H. Yonezawa, and A. Furusawa. Observation of -9 db quadrature squeezing with improvement of phase stability in homodyne measurement. *Opt. Exp.*, 15:4321, 2007.

[98] H. Vahlbruch, S. Chelkowski M. Mehmet, B. Hage, A. Franzn, N. Lastzka, S. Goßler, K. Danzmann, and R. Schnabel. Observation of squeezed light with 10-dB quantum-noise reduction. *Phsy. Rev. Lett.*, 100:033602, 2008.

[99] M. Mehmet, H. Vahlbruch, N. Lastzka, K. Danzmann, and R. Schnabel. Observation of squeezed states with strong photon-number oscillations. *Phys. Rev. A*, 81:013814, 2010.

[100] A. Khalaidovski, H. Vahlbruch, N. Lastzka, C. Gräf, H. Lück, K. Danzmann, H. Grote, and R. Schnabel. Status of the GEO 600 squeezed-light source. *arXiv:quant-ph/1112.0198v1*, 2011.

[101] M. Mehmet, S. Ast, T. Eberle, S. Steinlechner, H. Vahlbruch, and R. Schnabel. Squeezed light at 1550nm with a quantum noise reduction of 12.3 dB. *Opt. Exp.*, 19:25763, 2011.

[102] S.S.Y. Chua, M.S. Stefszky, C.M. Mow-Lowry, B.C. Buchler, S. Dwyer, D.A. Shaddock, P.K. Lam, and D.E. McClelland. Backscatter tolerant squeezed light source for advanced gravitational-wave detectors. *Opt. Lett.*, 36:4680, 2011.

[103] R. W. Boyd. *Nonlinear Optics*. Academic Press, Orlando, 2003.

[104] P. A. Franken, A. E. Hill, C. W. Peters, and G. Weinreich. Generation of optical harmonics. *Phys. Rev. Lett.*, 7:118, 1961.

[105] M. D. Reid. Demonstration of the Einstein-Podolsky-Rosen paradox using nondegenerate parametric amplification. *Phys. Rev. A*, 40:913, 1989.

[106] A. S. Villar, L. S. Cruz, K. N. Cassemiro, M. Martinelli, and P. Nussenzveig. Generation of bright two-color continuous variable entanglement. *Phys. Rev. Lett*, 95:243603, 2005.

[107] R.S. Klein, G.E. Kugel, A. Maillard, K. Polgár, and A. Péter. Absolute non-linear optical coefficients of $LiNbO_3$ for near stoichiometric crystal compositions.

[108] K. McKenzie. *Squeezing in the Audio Gravitational Wave Detection Band*. PhD thesis, Physics Department, Australian National University, Canberra, Australia, 2008.

[109] S. Wang. *Fabrication and characterization of periodcially-poled KTP and Rb-doped KTP for applications in the visible and UV*. PhD thesis, Laser Physics Division, Department of Physics, Royal Institute of Technology, Stockholm, Sweden, 2005.

[110] P. D. Drummond, K. J. McNeil, and D. F. Walls. Non-equilibrium transitions in sub/second harmonic generation I. semiclassical theory. *Optica Acta*, 27:321, 1980.

[111] C. W. Gardiner and M. J. Collett. Input and output in damped quantum systems: Quantum stochastic differential equations and the master equation. *Phys. Rev. A*, 31:3761, 1985.

[112] A. G. White. *Classical and quantum dynamics of optical frequency conversion.* PhD thesis, Physics Department, Australian National University, Canberra, Australia, 1995.

[113] M. J. Collett and C. W. Gardiner. Squeezing of intracavity and travelling-wave light fields produced in parametric amplification. *Phys. Rev. A*, 30:1386, 1984.

[114] A. E. Siegman. *Lasers.* University Science Books, Mill Valley, California, 1986.

[115] A. Yariv. *Quantum Electronics.* John Wiley & Sons., Singapore, 3 edition, 1989.

[116] N. B. Grosse. *Harmonic Entanglement & Photon Anti-Bunching.* PhD thesis, Physics Department, Australian National University, Canberra, Australia, 2009.

[117] R. Blachman, P. F. Bordui, and M. M. Fejer. Laser-induced photochromic damage in potassim titanyl phosphate. *Appl. Phys. Lett.*, 64:1318, 1994.

[118] J. C. Jacco, D. R. Rockafellow, and E. A. Teppo. Bulk-darkening threshold of flux-grown $KTiOPO_4$. *Opt. Lett.*, 16:1307, 1991.

[119] J. K. Tyminski. Photorefractive damage in KTP used as second-harmonic generator. *J. Appl. Phys.*, 70:5570, 1991.

[120] H. Vahlbruch, A. Khalaidovski, N. Lastzka, C. Gräf, K. Danzmann, and R. Schnabel. The GEO 600 squeezed light source. *Class. Quantum Grav.*, 27:084027, 2010.

[121] G. D. Boyd and D. A. Kleinman. Parametric interaction of focussed gaussian light beams. *J. Appl. Phys.*, 39:3597, 1968.

[122] I. Shoji, T. Kondo, A. Kitamoto, M. Shirane, and R. Ito. Absolute scale of seond-order nonlinear-optical coefficients. *J. Opt. Soc. Am. B*, 14:2268, 1997.

[123] R. C. Eckardt, C. D. Nabors, W. J. Kozlovsky, and R. L. Byer. Optical parametric oscillator frequency tuning and control. *J. Opt. Soc. Am. B*, 8:646, 1991.

[124] S. Pearl, H. Lotem, and Y. Shimony. Optimization of laser intracavity second-harmonic generation by a linear dispersion element. *J. Opt. Soc. Am. B*, 16:1705, 1999.

[125] A.L. Alexandrovski, M.M. Fejer, R.P. Route, and R.L. Byer. UV and visible absorption in $LiTaO_3$. In *SPIE Conference on Laser Material Crystal Growth and Nonlinear Materials and Devices (SPIE 1999)*, 1999.

[126] L. E. Myers, R. C. Eckardt, M. M. Fejer, and R. L. Byer. Quasi-phase-matched optical parametric oscillators in bulk periodically poled MgO:$LINbO_3$. *J. Opt. Soc. Am.*, 12:2102, 1995.

[127] Y. Furukawa, K. Kitamura, A. Alexandrovski, R. K. Route, M. M. Fejer, and G. Foulon. Green-induced infrared absorption in MgO doped $LiNbO_3$. *Appl. Phys. Lett.*, 78:1970, 2001.

[128] V. Pasiskevicius, S. Wang, J. A. Tellefsen, F. Laurell, and H. Karlsson. Efficient nd:yag laser frequency doubling with periodically poled KTP. *Appl. Opt.*, 37:7116, 1998.

[129] A. L. Alexandrovski, M. M. Fejer, R. P. Route, and R. L. Byer. Photothermal absorption measurements in optical materials. In *Lasers and Electro-Optics (CLEO 2000)*, 2000.

[130] L. E. Busse, L. Goldberg, and M. R. Surette. Absorption losses in Mg-O-doped and undoped potassium niobate. *J. Appl. Phys.*, 75:1102, 1993.

[131] P. K. Lam. *Applications of quantum electro-optic control and squeezed light.* PhD thesis, Physics Department, Australian National University, Canberra, Australia, 1998.

[132] R.G. Batchko, G.D.Miller, A.L. Alexandrovski, M.M. Fejer, and R.L. Byer. Limitations of high-power visible wavelength periodically poled lithium niobate devices due to green-induced infrared absorption and thermal lensing. In *(CLEO 1998)*, 1998.

[133] N. Grosse, W.P. Bowen, K. McKenzie, and P.K. Lam. Harmonic entanglement with second-order nonlinearity. *Phys. Rev. Lett.*, 96:063601, 2006.

[134] S. Li, W. Cao, and L. E. Cross. The extrinsic nature of nonlinear behavior observed in lead zirconate titanate ferroelectric ceramic. *J. Appl. Phys.*, 69:7219, 1991.

[135] K. McKenzie, E. E. Mikhailov, K. Goda, P. K. Lam, N. Grosse, M. B. Gray, N. Mavalvala, and D. E. McClelland. Quantum noise locking. *J. Opt. B*, 7, 2005.

[136] M. D. Reid and P. D. Drummond. Quantum correlations of phase in nondegenerate parametric oscillation. *Phys. Rev. Lett*, 60:2731, 1988.

[137] G. Björk and Y. Yamamoto. Generation and amplification of number states by nondegenerate parametric oscillators with idler-measurement feedback. *Phys. Rev. A*, 37:4229, 1988.

[138] P. D. Drummon and M. D. Reid. Correlations in nondegenerate parametric oscillation. II. Below threshold results. *Phys. Rev. A*, 41:3930, 1990.

[139] D. A. Shaddock. *Advanced Interferometry for Gravitational Wave Detection.* PhD thesis, Physics Department, Australian National University, Canberra, Australia, 2000.

[140] A. Khalaidovski, H. Vahlbruch, N. Lastzka, C. Gräf, K. Danzmann, H. Grote, and R. Schnabel. Long-term stable squeezed vacuum state of light for gravitational wave detectors. *arXiv:quant-ph/1109.3731v1*, 2011.

[141] H. Vahlbruch, S. Chelkowski, K. Danzmann, and R. Schnabel. Quantum engineering of squeezed states for quantum communication and metrology. *New J. Phys.*, 9:371, 2007.

[142] R. Schnabel, N. Mavalvala, D. E. McClelland, and P.K. Lam. Quantum metrology for gravitational wave astronomy. *Nat. Commun.*, 1.

[143] H. Vahlbruch. *Squeezed Light for Gravitational Wave Astronomy.* PhD thesis, The Albert Einstein Institute and the Institute of Gravitational Physics of Leibniz Universität Hannover, Hannover, Germany, 2008.

[144] P. Horrowitz and W. Hill. *The Art of Electronics.* Cambridge University Press, Cambridge, 2 edition, 1991.

[145] F. Seifert. Resistor current noise measurments. *Open access LIGO document LIGO-T0900200-v1*, 2009.

[146] E. D. Black. An introduction to pound-drever-hall laser frequency stabilization. *Am. J. Phys.*, 69:79, 2001.

[147] P. Kwee. *Laser characterization and stabilization for precision interferometry.* PhD thesis, Gottfried Wilhelm Leibniz Universität Hannover, Germany, 2010.

[148] F. Seifert. *Power stabilization of high power lasers for second generation gravitational wave detectors.* PhD thesis, Gottfried Wilhelm Leibniz Universität Hannover, Germany, 2010.

[149] H. Lück, J. Degallaix, H. Grote, M. Hewitson, S. Hild, B. Willke, and K. Danzmann. Opto-mechanical frequency shifting of scattered light. *J. Opt. A*, 10:085004, 2008.

[150] C. N. Man, A. Brillet, and P. Cerez. Suppression of optical feedback effects on saturated absorption signals by phase modulation of the reflected light. *J. Phys. E: Sci. Instrum.*, page 19, 1978.

[151] R. Schilling, L. Schnupp, W. Winkler, H. Billing, K. Maischberger, and A. Rüdiger. A method to blot out scattered light effects and its application to a gravitational wave detector. *J. Phys. E: Sci. Instrum.*, 14:65, 1981.

[152] G. A. Cranch, A. Dandridge, and C. K. Kirkendall. Suppression of double rayleigh scattering-induced excess noise in remotely interrogated fiber-optic interferometric sensors. *IEEE Photon. Technol. Lett.*, 15:1582, 2003.

[153] The LSC. Enhancement of the astrophysical reach of a gravitational wave observatory using squeezed states of light. Unpublished Manuscript.

[154] M. Frede, B. Schulz, R. Wilhelm, P. Kwee, F. Seifert, B. Willke, and D. Kracht. Fundamental mode, single-frequency laser amplifier for gravitational wave detectors. *Opt. Exp.*, 15:459, 2007.

[155] R. L. Savage Jr, P. J. King, and S. U. Seel. A highly stabilized 10-watt Nd:YAG laser for the laser interferometer gravitational-wave observatory (LIGO). *Laser Phys.*, 8:679, 1998.

[156] S. Dwyer. Summary of loss measurements for h1 squeezing test. *LIGO-T1200023-v1 (Unpublished)*, 2012.

[157] The LSC. Faraday isolator specifications for Advanced LIGO. *Open access LIGO document LIGO-T050226-00-D*, 2005.

[158] K. McKenzie, M. B. Gray, P. K. Lam, and D. E. McClelland. Nonlinear phase matching locking via optical readout. *Opt. Exp.*, 14:11256, 2005.

Electromagnetic Generators and Detectors of Gravitational Waves *

L. P. Grishchuk
Department of Physics and Astronomy, Cardiff University, Cardiff CF24 3YB, United Kingdom
and Sternberg Astronomical Institute, Moscow University, Moscow 119899, Russia
E-mail: grishchuk@astro.cf.ac.uk

The renewed serious interest to possible practical applications of gravitational waves is encouraging. Building on previous work, I am arguing that the strong variable electromagnetic fields are appropriate systems for the generation and detection of high-frequency gravitational waves (HFGW). The advantages of electromagnetic systems are clearly seen in the proposed complete laboratory experiment, where one has to ensure the efficiency of, both, the process of generation and the process of detection of HFGW. Within the family of electromagnetic systems, one still has a great variety of possible geometrical configurations, classical and quantum states of the electromagnetic field, detection strategies, etc. According to evaluations performed 30 years ago, the gap between the HFGW laboratory signal and its level of detectability is at least 4 orders of magnitude. Hopefully, new technologies of today can remove this gap and can make the laboratory experiment feasible. The laboratory experiment is bound to be expensive, but one should remember that a part of the cost is likely to be reimbursed from the Nobel prize money ! Electromagnetic systems seem also appropriate for the detection of high-frequency end of the spectrum of relic gravitational waves. Although the current effort to observe the stochastic background of relic gravitational waves is focused on the opposite, very low-frequency, end of the spectrum, it would be extremely valuable for fundamental science to detect, or put sensible upper limits on, the high-frequency relic gravitational waves. I will briefly discuss the origin of relic gravitational waves, the expected level of their high-frequency signal, and the existing estimates of its detectability.

I. INTRODUCTION

The engineers and enterpreneures are rightly seeing in gravitational waves (g.w.) a new opportunity, even though, at this moment of time, this opportunity does not look like having a chance of becoming practical in foreseeable future. However, we have many examples of wrong prognoses. It is known that when the discoverer of the nucleus, Lord Rutherford, was asked about the possible practical applications of nuclear energy, he answered "never". It is very likely that the gravitational radiation will eventually become useful, say, for purposes of communication. But in any case, it is difficult to imagine that we will go along without attempting to demonstrate in laboratory conditions the possibility of controllable generation and detection of gravitational waves. The laboratory experiment is a necessary first step to possible practical applications of g.w., and it is this first step that will be mostly discussed in this contribution.

Gravitational waves are less mysterious than one can gather from popular accounts. The relativistic Einstein's gravity is usually described in geometrical terms, as curvature of space-time. Correspondingly, gravitational waves are often described as "oscillations of space-time itself". This characterisation is quite puzzling and distressing for a practically-minded person. It makes defunct the usual physical intuition based on electromagnetic waves and other propagating physical fields. It makes one to suspect that there may be something artificial and unreal in the "waves of space-time itself". But this characterisation is only a matter of unfortunate language. Geometrical concepts are useful for some purposes, but they are greatly misleading for others. One should remember that although the Einstein gravity can be described in geometrical terms, it is also a normal physical field. The gravitational field is universal in all its interactions with itself and other fields and matter, and this is what makes the geometrical formulation of relativistic gravity possible, but the geometrical formulation is not necessary and is not obligatory. In fact, it would be a nightmare to try to discuss the laboratory gravitational-wave experiment in terms of differential geometry and "oscillating space-time", rather than in the engineering terms of the emitted and absorbed physical radiation.

The relativistic gravitational field is governed by the non-linear wave equations - the Einstein equations. However, in many situations one can neglect the non-linearities of the gravitational field. In particular, very often one can neglect the interaction of gravitational waves with other gravitational fields and with themselves. In this case, we

*Invited talk at the first conference on High-Frequency Gravitational Waves, May 2003, The MITRE Corporation, McLean, Virginia, USA

come to the notion of linearized gravity and weak gravitational waves. Certainly, in any laboratory experiment we will be dealing with extremely weak gravitational waves. Weak g.w. are very far from being mysterious; they may even appear quite boring and disappointing. Indeed, there is not so much difference with electromagnetic waves in terms of their general properties, but gravitational waves interact with matter and other fields much-much less effectively than electromagnetic waves. Gravitational waves carry their energy practically without scattering and absorption. This is why it is so difficult to detect astronomical gravitational waves. But the other side of this difficulty is the tremendous penetrating capability of gravitational waves. This is why they are so important as a tool of astronomical research, and this is why they attract attention as a possible unique means of communication.

The relativistic gravitational field can be described by 10 components of the 4×4 symmetric dimensionless tensor $h^{\mu\nu}$. The components of $h^{\mu\nu}$ are functions of time and spatial coordinates $x^\alpha = (ct, x, y, z)$. The functions $h^{\mu\nu}$ obey the nonlinear wave-like dynamical equations - the Einstein equations. The energy-momentum tensor $t^{\mu\nu}$ of the gravitational field is calculable from the field tensor $h^{\mu\nu}$.

The linearised gravitational waves satisfy the wave equation

$$h^{\mu\nu,\alpha}{}_{,\alpha} + \eta^{\mu\nu} h^{\alpha\beta}{}_{,\alpha,\beta} - h^{\nu\alpha,\mu}{}_{,\alpha} - h^{\mu\alpha,\nu}{}_{,\alpha} = 0, \tag{1}$$

where the ordinary derivative is denoted by a comma and $\eta^{\mu\nu}$ is the metric tensor of the Minkowski space-time:

$$\mathrm{d}\sigma^2 = \eta_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = c^2\mathrm{d}t^2 - \mathrm{d}x^2 - \mathrm{d}y^2 - \mathrm{d}z^2. \tag{2}$$

The first term in Eq. (1) is the familiar d'Alembert (wave) operator:

$$h^{\mu\nu,\alpha}{}_{,\alpha} = \left(\frac{\partial^2}{c^2 \partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2} \right) h^{\mu\nu}.$$

A plane-wave solution to Eq. (1) is given by

$$h^{\mu\nu} = a^{\mu\nu} e^{i k_\alpha x^\alpha} + c.c., \tag{3}$$

where $k_\alpha k^\alpha = 0$, reflecting the fact that a g.w. propagates with the speed of light. Because of this condition, the field equations (1) require the 10 components of the constant matrix $a^{\mu\nu}$ to satisfy 4 constraints: $a^{\mu\nu} k_\nu = 0$. Among the remaining 6 components, only 2 degrees of freedom (sometimes called the TT-components $\tilde{a}^{ij}$) are physically important, in the sense that it is only these degrees of freedom that fully determine the observational manifestations of the plane wave and its energy-momentum characteristics. Indeed, it is easy to show that the gravitational energy-momentum tensor

$$t_{\mu\nu} = \frac{c^4}{32\pi G} \left[h^{\alpha\beta}{}_{,\mu} h_{\alpha\beta,\nu} - \frac{1}{2} h_{,\mu} h_{,\nu} \right] \tag{4}$$

depends only on the TT-components of the field and its two independent amplitudes h_+ and $h_\times$:

$$t_{\mu\nu} = \frac{c^4}{32\pi G} k_\mu k_\nu \left[2\tilde{a}^{ij} \tilde{a}^*_{ij} \right] = \frac{c^4}{32\pi G} k_\mu k_\nu \left[h_+^2 + h_\times^2 \right]. \tag{5}$$

In this expression we have dropped (as we normally do in the case of energy-momentum tensor for the electromagnetic waves) the purely oscillatory terms.

The amplitudes $h_+, h_\times$ are determined by the source of the gravitational waves. To find the amplitudes, we should replace the zero in the right hand side of Eq. (1) by the source term $(16\pi G/c^4) T^{\mu\nu}$, where $T^{\mu\nu}$ is the energy-momentum tensor of the source, and seek the retarded solutions to the inhomogeneous wave equation. For example, for a pair of stars in a circular orbit, with masses of the stars M_1, M_2, located at the distance R_0 from us, and after averaging over the orbital period and orientation of the orbital plane, we obtain

$$h = \left(\langle h_+^2 \rangle + \langle h_\times^2 \rangle \right)^{1/2} = \left(\frac{32}{5} \right)^{1/2} \frac{1}{R_0} \frac{G^{5/3}}{c^4} \frac{M_1 M_2}{(M_1 + M_2)^{1/3}} (\pi f)^{2/3}, \tag{6}$$

where the emitted g.w. frequency f (in Hz) is twice the orbital frequency.

Roughly, the characteristic amplitude h from a given source is given by

$$h_{ij} \approx \frac{G}{c^4} \int \frac{T_{ij}(t_{ret})}{r} dV \qquad \text{and} \qquad h \approx \frac{1}{R_0} \frac{GM}{c^2} \left(\frac{v}{c} \right)^2, \tag{7}$$

where M is the total mass of the source and v is the characteristic velocity of the matter bulk motion. This formula is quite universal, and it can also be written in terms of the characteristic variable stresses σ within the source, and the source's volume V:

$$h \approx \frac{1}{R_0}\frac{G}{c^4}\sigma V. \tag{8}$$

As long as the retardation effects inside the source can be neglected, this formula is equally well applicable to the radiating systems of any nature - in cosmos and in laboratory, mechanical and electromagnetic.

II. CURRENT STATUS OF THE ASTRONOMICAL PROGRAM AND ITS COMPARISON WITH A HFGW LABORATORY PROGRAM

The dimensionless number h is a very convenient characterisation of the strenght of a given g.w. and its detectability. Under the action of a gravitational wave, a pair of free masses, initially separated by the distance l, experience relative oscillatory displacements proportional to the incoming wave amplitude h: $\delta l/l \approx h$. Very powerful astronomical sources, currently under intense experimental g.w. searches [1], [2], [3], produce something like $h \approx 10^{-22}$ at Earth. This is an increadibly small number. It enters any conceivable method of detection of gravitational waves and explains why it is so difficult to observe them. For example, in a $4km$-long laser interferometer, such as LIGO, we need to beat all the noises and measure the mirror's displacements at the level of $4 \times 10^{-17} cm$. At the same time, the flux of energy $F = ct^{0i}$ at Earth from the discussed sources is quite impressive by astronomical standards. At the representative frequency $f = 200Hz$, the flux reaches $F \approx 3 \times 10^{-2} erg/sec\ cm^2$. But most of this energy passes through the g.w. detectors without scattering and absorption. This is why, in the gravitational-wave physics, such characteristics as *watts* and *joules* may be misleading. The dimensionless amplitude h is more adequate. But in any case, whatever the units in which the analysis is being carried out, only a joint discussion of the complete system (emitter plus detector) gives the correct evaluation of what is "small" or "big", easy to detect or extremely difficult to detect.

The recently assembled LIGO interferometers are approaching their planned level of sensitivity. It appears that, at the time of writing, there still exists a gap in one order of magnitude, in terms of h, between the actually reached level of sensisitivity and the design sensitivity [1]. One should clearly understand what is likely to happen when the design level of sensitivity of the initial interferometers is finally achieved. Baring the extremely fortunate surprises, we will probably be able to see only the most powerful, but rare, sources, such as coalescing binary stellar-mass black holes. Less powerful, even if more numerous, sources - coalescing binary neutron stars, are unlikely to be seen, because the expected signal-to-noise ratio is somewhat smaller than 1. For a guaranteed detection of astronomical sources, the experimenters will have to improve the sensitivity by one further order of magnitude, as compared with the design sensitivity of the initial instruments. This is the goal of the so-called advanced interferometers, and this goal will probably be reached by 2007. (For more details on the current status of the gravitational wave astronomy see, for example, [2], [3].)

There is little doubt that the astronomical program will completely dominate, and rightly so, any other experimental efforts in the gravitational-wave physics for some time to come. Having said that, one can still wonder whether a laboratory experiment is much more difficult to realise than to build equipment for the observation of cosmic gravitational waves. Of course, the justifications for these efforts are totally different. In gravitational-wave astronomy we directly explore the fascinating Universe, whereas in the laboratory experiment we are likely to confirm a theory (true, very fundamental theory, but anyway tested also by other means), with very remote prospects for practical applications of gravitational radiation. However, the justification for the laboratory experiment is sufficiently convincing. A more difficult question is its feasibility. Here, one will have to admit that the both enterprises are very difficult. Surprisingly, the laboratory program does not appear to be unacceptably more difficult than the cosmic program. If we take as a benchmark the 4 orders of magnitude separating the detecting and generating capabilities in laboratory (see Sec. IV below), it is like having 1 dollar instead of required 10000. But, strictly speaking, in the cosmic program we also have, at the time of writing, only something like 1 dollar instead of required 100 (see above). The difference between the two programs is substantial, but not ridiculously large.

As an additional motivation for the experimental work on HFGW, there are arguments showing that it can be complementary and useful for the astronomical program. First, there exists the fundamentally important cosmological signal - relic gravitational waves. The high-frequency end of the spectrum is the most natural interval for the high-frequency techniques, and first of all for electromagnetic detectors (see Sec V below). Second, part of the high-frequency studies with electromagnetic detectors may eventually be useful for laser interferometers of the next generation. Indeed, it is quite likely that in order to reach the required level of sensitivity we will be forced to implement the sophisticated techniques such as squeezed light and quantum-nondemolition measurements (see, for example, [4]) and this is where the expertise of the HFGW community can be useful.

III. EFFICIENCY OF GRAVITATIONAL-WAVE EMITTERS

Astronomical sources are immensely more powerful than any conceivable sources at Earth, but they are very far away from us, and not under our control. In contrast, in laboratory, one can place the emitter and the detector very close to each other, choose the appropriate materials, implement the coherence and focusing of the emitted radiation, use the advantages of the resonant detection, etc. One usually illustrates the hopelessness of laboratory g.w. experiments by giving an example of the meager gravitational radiation generated by a rotating massive rod. Well, if you are so naive that you plan to rotate a rod, then the enterprise may indeed be hopeless. But surely there must exist something smarter. Let us give a general comparative analysis of possible mechanical and electromagnetic systems [5].

Let us consider an elementary mechanical emitter (m-emitter) and an elementary electromagnetic emitter (e-emitter). By the elementary we mean a source that occupies a volume of order λ_s^3 in the first case and λ_e^3 in the second case; λ_s and λ_e are the wavelengths of the acoustic and electromagnetic waves. A vibrating object and an oscillating electromagnetic wave in a cavity are examples of $m-$ and $e-$ emitters respectively. For the comparison to be fair, the sources are assumed to emit g.w. with one and the same wavelength λ_g and are placed at the same distance R_0 from the observer.

Let A be the amplitude of elastic vibrations in the m-emitter. Then the characteristic amplitude of the stress tensor T_{ij} is $\sigma_m \approx \rho_m v_s^2 (A/\lambda_s)$, where ρ_m is the density of the material and v_s^2 is the square of the sound speed. Since $\nu_g = c/\lambda_g \approx v_s/\lambda_s$, we have $\lambda_s \approx (v_s/c)\lambda_g \ll \lambda_g$ which means that the body of the elementary m-emitter is situated deeply inside the inductive zone of the gravitational radiation and, hence, the retardation effects within the source are negligibly small. For the g.w. amplitude we obtain from Eq. (8):

$$h_m \approx \frac{G}{c^4}\frac{1}{R_0}\sigma_m \lambda_s^3. \tag{9}$$

Let us now consider an elementary e-emitter. The amplitude σ_e of the electromagnetic stress has the order of magnitude of the energy density ϵ_e of the varying electromagnetic field, $\rho_e c^2 = \epsilon_e$, i.e. $\sigma_e \approx \epsilon_e$. Since $\nu_g = c/\lambda_g \approx c/\lambda_e$, the volume of the elementary e-emitter is of the order of λ_g^3, that is, it is still at the limit of applicability of Eq. (8) with the retardation effects ignored. The amplitude of the emitted gravitational wave is

$$h_e \approx \frac{G}{c^4}\frac{1}{R_0}\sigma_e \lambda_g^3. \tag{10}$$

The ratio of (9) and (10) is

$$\frac{h_m}{h_e} \approx \frac{\rho_m}{\rho_e}\left(\frac{v_s}{c}\right)^5 \frac{A}{\lambda_s}. \tag{11}$$

Using the reasonable parameters: $\rho_m \approx 1 g/cm^3$, $\rho_e \approx 10^{-18} g/cm^3$, $v_s/c \approx 10^{-5}$, $A/\lambda_s \approx 10^{-3}$, we find that $h_m/h_e \approx 10^{-10}$. In other words, an elementary e-emitter is much more efficient than an elementary m-emitter. However, the comparison is not entirely fair as the volume of the former, λ_g^3, is much larger than the volume of the later, λ_s^3. In the volume of an elementary e-emitter one can accomodate a large number $N = (\lambda_g/\lambda_s)^3 \approx (c/v_s)^3 \gg 1$ of elementary m-emitters. Under the condition that they all are phased to work coherently, the total g.w. amplitude is the sum of individual amplitudes, and it can be as large as $h_{mc} \approx N h_m$. Then, we finally obtain

$$\frac{h_{mc}}{h_e} \approx \frac{\sigma_m}{\sigma_e}. \tag{12}$$

This formula answers all the principal questions. We see that the most important parameter is the maximal reachable amplitude of the dynamical stresses. Extremely high stresses, at the limit of static breaking point, for very strong materials, is in the range of $\sigma_m \approx 10^9 dyne\ cm^2$. Similar variable stresses can be obtained by producing and maintaining variable electromagnetic fields with characteristic field strength $E \approx H \approx 10^5 gauss$. Then, an elementary e-emitter, having the volume $\lambda_g^3 \approx 10^6 cm^3$, produces a g.w. amplitude $h \approx 10^{-36}$ at the boundary of the wave zone, and emits the total power $W \approx 4 \times 10^{-13} erg/sec$. To raise the amplitude and power, one would have to implement large composite systems.

This analysis illustrates the advantages of the electromagnetic systems. First, the coherence of the source is automatically achieved in a large volume $\sim \lambda_g^3$. The huge number of mechanical emitters placed in the same volume would need to be specially phased in order to achieve the constructive interference of the emitted gravitational radiation. Second, it seems that it is much easier to manufacture a simple electromagnetic emitter (essentially, an oscillating eigen-mode of the electromagnetic field in a cavity) than try to pack together a large number of specially phased mechanical emitters. Third, the efficient generation process suggests the similar (inverse) process of detection. It seems natural to use the electromagnetic systems also as the detection technique.

IV. COMPLETE LABORATORY EXPERIMENT

The important concern for the succcess of the laboratory experiment is to ensure the focusing, as much as possible, of the emitted g.w. power in one place, rather than to let it be dispersed over all directions. This was one of motivations for the specific geometrical configuration suggested in [6]. It is proposed that the emitter is a torus-like electromagnetic resonator with a rectangular cross-section (see Fig. 1). The oscillating electromagnetic field in the resonator produces a standing gravitational wave in the focal region, near the axis of symmetry. The gravitational wave is standing because the emitted cylindrical gravitational wave passes through the axis of symmetry and interferes with itself. The gravitational-wave frequency Ω is twice the frequency of the variable electromagnetic field in the generator, $\Omega = 2\omega$. Another resonator is placed in the focal region and plays the role of the detector. Its resonant frequency is tuned to the frequency of the gravitational wave Ω. Among the advantages of this particular configuration is its geometrical simplicity, which allows one to find exact solutions to the Maxwell equations, both, in the generator and in the detector.

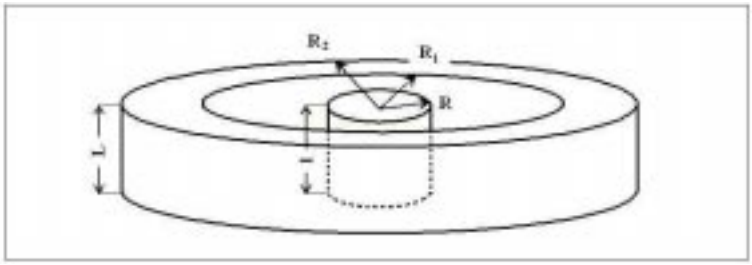


FIG. 1. Scheme of a possible laboratory experiment

Although almost all the calculations [6] have been performed exactly, it is convenient to use the order of magnitude evaluations which follow from the calculations, and which, of course, coincide with the evaluations outlined above. The very important estimate, which we did not discuss yet, is the detectabilty condition.

The analysis shows that the change of energy $\Delta\mathcal{E}$ in the resonator-detector depends on whether the initial electromagnetic field in the detector is realised as a constant (say, magnetic) field, or as a an oscillating resonant electromagnetic eigen-mode. It is assumed that the accumulation of the signal is taking place during the relaxation time $\tau^* = Q/\Omega$, where Q is the quality factor of the resonator-detector. Then, in the first case, $\Delta\mathcal{E} \approx (hQ)^2\mathcal{E}$, where $\mathcal{E} = (H^2/8\pi)V$. And in the second case, $\Delta\mathcal{E} \approx (hQ)\mathcal{E}$, where $\mathcal{E} = (E^2/4\pi)V$. These formulas show that the responses of the detector are different in these two options. But the natural electromagnetic noises are also different. One should impose different criteria in order to determine whether the one and the same g.w. signal is detectable or not. In the first case, one can think that one can distinguish one "photon" (a resonant mode exitation with energy $\mathcal{E} \sim \hbar\Omega$) on the background of a constant (magnetic) field. In the second case, one should typically assume that only the number $\sqrt{N}$ of new "photons" can be distinguished on the background of the already existing N "photons", $N = \mathcal{E}/\hbar\Omega$. Combining each of responses with the corresponding detectability condition, one finds that there is in fact no much difference between these two cases in terms of the measurable amplitude h. The order of magnitude evaluation shows that the detectable h is

$$h_{det} \approx \sqrt{\frac{\hbar\Omega}{\mathcal{E}}}\frac{1}{Q}. \tag{13}$$

The situation can, however, change considerably in favor of the second option, if one succeeds in realising a special (quantum) state of the electromagnetic field in the detector, such that the variance of the number of quanta N is much smaller than $\sqrt{N}$. Then, assuming that the quantum non-demolition measurement can distinguish the energy of, say, a few new quanta on the backgound of a huge (mean) number of quanta in the detector, the detectable g.w. amplitude h_{det} can be significantly lowered. We will first continue our estimates without taking this possibility into account, but will return to it later.

The geometry of the system allows one to choose the following optimal parameters: $L \approx l \approx \lambda$, where $\lambda = 2\pi c/\Omega$, $R \approx 2\lambda/3,\ R_1 \approx 2\lambda,\ R_2 \approx 7\lambda/2$. Then the detectability condition can be written as

$$\lambda^4 E^2 HQ \approx 30\frac{c^4}{G}\sqrt{\hbar c}, \tag{14}$$

where E is the amplitude of the oscillating field in the generator and H is the typical (possibly, constant magnetic) field in the detector. Let us take for illustration $\lambda \approx 10^2 cm,\ E \approx H \approx 3\times 10^5 gauss,\ Q \approx 7\times 10^{13}$. Then, the

left-hand side of Eq. (14) is 4 orders of magnitude smaller than the right-hand side of the same formula. It is this gap of 4 orders of magnitude that we referred to in the Abstract.

One possibility to satisfy Eq. (14) is to simply increase the size of the system, going from $\lambda \approx 10^2 cm$ to $\lambda \approx 10^3 cm$. Another possibility would be to raise the product E^2HQ by the same 4 orders of magnitude. Probably the most elegant and realistic possibility is to try and improve the detectability condition by using the sophisticated quantum states of the electromagnetic field in the detector, that we mentioned earlier. Certainly, the laboratory experiment is going to be difficult, but it seems worth of trying.

V. ELECTROMAGNETIC DETECTORS FOR RELIC GRAVITATIONAL WAVES

In laboratory conditions, and in almost all astrophysical situations, one can completely neglect the non-linearities of gravity, that is, the interaction of gravitational waves with other gravitational fields. However, this is not always the case. The most important example is the interaction of gravitational waves with the strong variable gravitational field of the very early Universe. One can use the engineering intuition in order to understand what is going on here. A gravitational wave can be thought of as a harmonic oscillator, while the smooth variable gravitational field of the surrounding Universe as a gravitational pump field. The g.w. oscillator is parametrically coupled to the gravitational pump field. This specific coupling follows from the non-linear structure of the Einstein equations. This coupling provides a mechanism for the superadiabatic (parametric) amplification of classical waves and for the quantum-mechanical generation of waves from their zero-point quantum oscillations [3]. The word "superadiabatic" emphasizes the fact that this effect takes place over and above whetever happens to the wave during very slow (adiabatic) changes of the pump field. That is, we are interested in the increase of occupation numbers, rather than in the gradual shift of energy levels. The word "parametric" emphasizes the mathematical structure of the wave equation. A parameter of the oscillator, namely its frequency, is being changed by the variable pump field. It is this sufficiently rapid change of frequency of the oscillator that is responsible for the considerable increase of energy of that oscillator.

The parametric amplification of the inevitable zero-point quantum oscillations leads to the generation of a stochastic background of relic gravitational waves. The mechanism itself is based on the fundamental physics only, but the generated signal depends on the pump field. In other words, the amount of relic gravitational waves depend on a concrete cosmological model of the very early Universe. Combining the theory with the available cosmological data, one can evaluate the expected level of high-frequency relic gravitational waves [3]. For example, the root-mean-square (r.m.s.) amplitude is expected to reach $h_{r.m.s.} \approx 10^{-30}$ at $\nu = 10^7 Hz$. The amplitude is smaller at higher frequencies. It may reach $h_{r.m.s.} \approx 10^{-32}$ at $\nu = 10^{11} Hz$, but then it should quickly decrease as a function of higher frequencies. There is no much sense to expect any relic gravitational waves at frequencies higher than that.

We may be lucky (although it does not seem very likely) if the thermal background of gravitational waves survived until now. Then, in the vicinity of $\nu = 10^{11} Hz$ there will be a maximum of the Planck spectrum. The amplitude of the Planck spectrum can be in the region of $h_{r.m.s.} \approx 3 \times 10^{-32}$, but not very much larger than this. The quoted numbers for relic and thermal backgrounds give the feeling of what we can expect of high-frequency cosmic gravitational waves.

What can be said about the detectability of relic gravitational waves ? The first impression is that they may be easier to detect than the gravitational waves produced in laboratory. For example, at $\nu = 10^7 Hz$, the amplitude of relic gravitational waves, $h_{r.m.s.} \approx 10^{-30}$, is several orders of magnitude higher than the realistic amplitude of laboratory gravitational waves (see Sec. IV). Unfortunately, this does not mean that the relic gravitational waves are easier to detect. The crucial difference is in the character of these two signals. The laboratory signal is essentially a deterministic monochromatic wave, which allows one to systematically accumulate the response amplitude. In contrast, relic gravitational waves form a random signal, which allows only an accumulation of energy. When it comes to evaluation of the detectable amplitude of relic gravitational waves [7], formula (13) should be modified. Under the same conditions that formula (13) was derived, but now for the stochastic signal, we obtain

$$h_{det} \approx \sqrt{\frac{\hbar\Omega}{\mathcal{E}}} \frac{1}{\sqrt{Q}}. \tag{15}$$

Applying this formula to a simple electromagnetic detector, we can expect to measure $h_{det} \approx 10^{-26}$ instead of the required level of the cosmic signal $h_{r.m.s.} \approx 10^{-30}$.

This gap is sufficiently wide to discount hopes to bridge this gap by longer observation time, or cross-correlation of several detectors, etc. It appears that the sensitivity of electromagnetic detectors can reach the required level only if the squeezed quantum states of the electromagnetic field in the detector, and quantum non-demolition measurements, are implemented. Aternatively, one can think of large composite systems ("large crystal" [7]), where the individual elements are arranged to work more coherently than simply absorbing gravitational radiation independently one from

other. Although these techniques seem possible in principle, so far, there is no concrete proposals how to implement them in practice. Regrettably, the detection of high-frequency relic gravitational waves will probably be even more challenging problem than their detection in low-frequency and very low-frequency bands. Smart ideas are badly needed.

---

[1] http://www.ligo.caltech.edu

[2] C. Cutler and K. S. Thorne. In: *General Relativity and Gravitation*, Eds. N.T.Bishop and S.D.Maharaj, (World Scientific, 2002) p. 72

[3] L. P. Grishchuk. *Update on gravitational-wave research*, ArXive: gr-qc/0305051

[4] V. B. Braginsky, M. L. Gorodetsky, F. Ya. Khalili, A. B. Matsko, K. S. Thorne, and S. P. Vyatchanin. ArXive: gr-qc/0109003

[5] L. P. Grishchuk. Phys. Lett. **56A**, 255 (1976)

[6] L. P. Grishchuk and M. V. Sazhin. Zh. Eksp. Teor. Fiz. **68**, 1569 (1975) [Sov. Phys. JETP. **41**, 787 (1976)]; L. P. Grishchuk. Usp. Fiz. Nauk. **121**, 629 (1977) [Sov. Phys. Usp. **20**, 319 (1977)]

[7] L. P. Grishchuk. Pis'ma Zh. Eks. Teor. Fiz. **23**, 326 (1976) [JETP Lett. **23**, 293 (1976)]; Usp. Fiz. Nauk. **156**, 297 (1988) [Sov. Phys. Usp. **31**, 940 (1989)]

Design of a Quantum Source of High-Frequency Gravitational Waves (HFGW) and Test Methodology.

Giorgio Fontana

Department of Information and Communication Technology, University of Trento, 38050 POVO (TN), Italy

Phone: +390461883906, e-mail giorgio.fontana@unitn.it

Abstract. The generation of High-Frequency Gravitational Waves (HFGW) has been identified as the required breakthrough that will lead to new forms of space propulsion. Many techniques have been devised to generate HFGW, but most of them exhibit marginal efficiency, therefore the power emitted in form of gravitational waves (GW) is orders of magnitude lower than the input power. The gravitational wave counterpart of the LASER, termed Gravitational-wave LASER or "GASER" is the quantum approach to the efficient generation of gravitational waves. Electrons, protons, muons, etc, all have charge and mass, if accelerated they usually lose energy through the very fast electric and magnetic channels, this causes a negligible emission through the gravitational channel. Quantum systems can be engineered to forbid electric and magnetic transitions, therefore the gravitational spin-2 transitions can take place. A class of active materials, suitable for making a GASER based on electronic transitions in the solid state, is identified along with their relevant physical properties. Means for creating coherence and population inversion and means to increase the emission probability are described. The expected performances of the device are derived from quantum gravitational theories. Additional properties of the active materials are considered to enforce the theoretical foundation of the device. A proof-of-concept device, operating at about 1 THz, is described. Experiments are proposed as a natural starting point of the research.

INTRODUCTION

Present space propulsion technology, mainly based on rocketry, suffers of serious limitations if the range of operations is extended above earth orbit.

Dynamical and speed limitations are accompanied by safety concerns. Rockets must transport the payload and the propellant used as reaction mass. Rockets require such a large amount of energy to accomplish their tasks that management of the required fuel may expose the passengers to a high risk. The energy requirement for a space transportation system can be reduced if a new technology is developed to employ distant masses as reaction mass. The propellant will be no longer needed and will not take away energy.

Gravitational Wave Propulsion

A remarkable advancement in space propulsion is a form of gravitational wave propulsion (GWP) (Fontana, 2000 and 2003a; Baker, 2000). This is still an untested technique, nevertheless it is based on a straightforward application of the theory of General Relativity (GR), which is trusted by the vast majority of scientists. The new technology is based on High Frequency Gravitational Waves (HFGWs).

A large literature exists on colliding gravitational waves (Szekeres, 1972; Ferrari, 1988a and 1988b), it has been found that the collision or focusing (Alekseev, 1995 and 1996) of gravitational waves produce curvature singularities. These singularities have properties very similar to those of a black-hole, an essential and

fundamentally simple object, which produces a gravitational field. GWP is the application of these theories to space travel.

Generators of GWs could be installed directly onboard or remotely to a spacecraft to induce curvature singularities near the spacecraft. The use of HFGW "… as a source of some additional gravitational field**…**" at a distance was suggested by L. D. Landau and E. M. Lifshitz (1975). According to GR, spacecraft mass interacts with spacetime curvature, therefore the spacecraft will move towards the singularity. In the Newtonian picture, because of the non-linearity of space, the wave at the focus is converted to a Coulomb-like gravitational field. The induced field accelerates the spacecraft and accelerates the reaction mass, this latter can be easily identified with distant masses (Fontana, 2003a). In Newtonian gravity the propagation of the gravitational field is instantaneous; this might not be true, nevertheless it must be accepted when dealing with this approach.

Generators of Gravitational Waves

To create spacetime singularities, powerful and efficient generators of GW are required. The produced GWs must have a wavelength of the order of millimeters or less, because it is necessary to collimate or focus the waves to a small spot to reach the high energy densities that produces non-linearity effects culminating with the creation of spacetime singularities. The wavelength of GWs must be much smaller than the size of the generator which, in turn, must be compatible with the size of a typical spacecraft.

A large literature describes different proposed devices capable of producing HFGWs (Baker , 2004; Pinto, 1988). Generators of high frequency gravitational waves based on quantum effects may offer higher efficiencies because only at the quantum level there is a precise distinction between gravitational and electromagnetic transitions. In addition, quantum generators may benefit from coherence and/or Bose condensation effects, which may strongly increase the emission probability of GW quanta. Quantum generators of HFGWs can potentially be very efficient in converting input power to GWs.

After recognizing that a quantum source can be developed (Halpern, 1964; Ford, 1982), the first serious difficulty in designing a quantum source of HFGW is the identification of a suitable active material.

The first interesting result consists of recognizing that the orthorombic cuprate High Temperature SuperConductor (HTSC) might become an active material in which gravitational transitions may take place (Fontana, 2000). Later a more comprehensive approach to the problem, including the details of the HTSC GASER, was developed and presented by Fontana and Baker (2003b).

According to the historical nomenclature the quantum source of gravitational waves has been named GASER and the quantum of gravitational wave, graviton.

In the article by Fontana and Baker (2003b), the instanton, with its theoretical foundations developed in Gibbons (1993) (Sanchez, 1982 and 1983), has been used to describe the wave-particle associated to a quantum of gravitational radiation. The (gravitational) instanton can be defined as a pseudoparticle solution of the Einstein equations in Euclidean spacetime (signature ++++) with a cosmological constant. The Euclidean signature helps to remove singularities found in the theory with Lorentzian signature (-+++) and has been introduced for the purpose of the analysis. The Euclidean signature is characterized by imaginary-valued time, the Lorentzian signature is characterized by real-valued time. Research programs have also found instantons solutions in our Lorentzian spacetime and, interestingly, some of them can be identified with colliding gravitational waves (Sanchez, 1982).

The well-known "graviton" is not so well defined. It is usually associated to a spin-2 field that has been somehow quantized. Recently it has been discovered that the graviton may have a mass, depending on boundary conditions like large scale, or local, properties of space-time associated to the Λ term, known as the cosmological constant (Novello, 2003), with the result that $m_g^2=-2\Lambda/3$. The possibly changing sign of the Λ term, while going from the vicinity to the source to the free space of our expanding universe, may give unexpected properties to the graviton.

Today, a generally accepted mathematical description of the quantum of GWs does not exist. The lack of experimental access to these objects and their interactions might explain the difficulties in finding an acceptable

model. In addition to space propulsion models, a quantum source of GWs is also the resource that must be developed for further progress towards a better understanding of gravitation and cosmology.

THE HTSC GASER

Different from GW quanta themselves, the mechanisms capable of producing GWs, even from microscopic systems, are better known. The theory describing the emission of gravitational radiation from quantum systems has been developed by Halpern (1964) and Ford (1982). Transitions for which the orbital quantum number L changes by ±2 and for which the total quantum number J changes by 0 or ±2 are gravitational quadrupolar transitions, for which the emission of photons is forbidden and the emission of gravitons is allowed. The selection rules for photons are those for which the orbital quantum number L changes by ±1 and for which the total quantum number J changes by 0 or ±1, etc. For instance, for an atom, gravitational transitions are those between orbitals 3d and one characterized by a lower energy among 3s, 2s and 1s. In typical atomic systems gravitational transitions compete with multiple photon transitions, therefore methods for counteracting photon transitions have to be developed. Detailed analysis comparing the emission of gravitons to the emission of photons has shown that the ratio R of the transition probabilities for matrix elements of equal structures is of the order of 1.6×10^{-36} for the proton (Halpern, 1964) and of the order of 4.8×10^{-43} for the electron. An intermediate value can be found for the muon.

To increase the efficiency, stimulated emission and coherence can improve the apparently negligible emission probability. The device that employs stimulated emission has been named GASER (Halpern, 1964). We adopted (Fontana, 2003b) the same name for the device that employs wavefunction coherence for increasing the emission probability.

According to Fontana and Baker (2003b), the ideal gaser material should provide:

1. A quantum system with two energy levels characterized by a difference in orbital quantum number of 2, in agreement with Halpern and Laurent selection rules.

2. A quantum system possibly composed of densely packed couples of closely interacting electrons.

3. A quantum system in which the two energy levels are respectively populated by objects with exactly the same wavefunction and the same energy in order to have negligible linewidths, thus permitting efficient stimulated emission.

4. A quantum system in which the two energy levels are populated by quantum objects, for which the two wavefunctions are orthogonal in order to prevent photon transitions and tunneling.

5. A quantum system in which population inversion can be achieved to initiate and sustain gaser action.

A candidate material with the enumerated requirements is a cuprate high temperature superconductor (HTSC) with orthorombic crystal structure, being the two quantum systems *s-wave* (l=0) and *d-wave* (l=2) Cooper-pair condensates (Kouznetsov, 1997) (Fontana, 2000).

In Fontana and Baker (2003b), it has been shown, with the help of three quantum gravitational theories, that the emission probability for gravitons is proportional to the square of the number of Cooper-pairs involved. Assuming Cooper-pair density of the order of $10^{20}/cm^3$ or slightly higher (there are $\sim 5 \times 10^{21}/cm^3$ elementary cells in YBCO) the emission probability looks interesting for a volume of the HTSC greater than few cubic centimeters. In addition to that analysis, the recent discovery that Cooper-pair formation in HTSC involves the *4s* to *3d* electron exchange amplitude in Cu atoms (Mishonov, 2003) that shows the electrons of the pairs posses the quantum numbers required for the gravitational transition. The main process allowing direct transitions between *s-wave* and *d-wave* condensates is therefore purely gravitational.

For direct transitions the energy of the emitted particle only depends on the energy gap. Taking into account the measured *s-wave/d-wave* energy gap, the frequency of 1.3THz can be estimated from BSCCO data (Moessle, 1999).

Like the well-known optical laser, a GASER could be pumped by another GW source, this possibility was not yet properly explored and does not solve the problem of producing GWs from a different form of energy. According to Fontana and Baker (2003b), the HTSC GASER can be pumped electrically, with the injection of *s-wave* Cooper-pairs by a Josephson junction between a Low Temperature SuperConductor (LTSC) and the HTSC, respectively connected to the poles of a current generator. LTSCs are characterized by pure *s-wave* pairing. The estimated power emitted in a GW is of the order of 10 W/ cm^2 with a current of 10 kA/cm^2.

FABRICATION OF A PROTOTYPE HTSC GASER

Generally, the fabrication of prototype devices is a complex procedure. It starts with the design of the device, structure and size, goes through the definition of the materials, nature and purity, the definition of the fabrication process, techniques and steps. Finally, the properties of the produced device will be measured to recognize if the results agree with expectation. The procedure is usually repeated many times.

The structure of the proposed HTSC GASER is quite simple; an orthorombic cuprate HTSC mono-crystal should be produced along with a Josephson junction with a LTSC. Both sides of the device should be metalized for connecting the GASER to a current generator. To define the preferable geometry of the HTSC mono-crystal we make some considerations on directivity patterns in wavefunction coherent systems.

Directivity of emission is generally determined by a combination of geometrical properties of the emitter, for instance the shape of the resonant cavity of a LASER. For a wavefunction coherent system in the solid state, composed of microscopic oscillators emitting waves, the directivity function is the product of two components: the directivity function of the single emitter, and the superposition function (interference) of many such emitters. The collective wavefunction is the mechanism that synchronizes the oscillators.

In Halpern (1964) it has been shown that there exists a directivity pattern for transitions emitting gravitons, the directivity here depends on the *d-wave* wavefunction orientation. In HTSC, the macroscopic wavefunction has been observed by interference effects in Josephson junctions. The macroscopic wavefunction is oriented according to the crystallographic orientation of the HTSC crystal. Even if this factor has a minima and maxima, its dependence on the angles is smooth, and it appears to be not critical, nevertheless crystals with different crystallographic orientation should be built to test the theory. The directivity factor related to the superposition function of many emitters is what mainly determines the directivity function and the directivity gain. The angular distribution of GWs for a quantum system is derived in Halpern (1964). For J=M=2 it takes the form:

$$\cos^8\left(\frac{\vartheta}{2}\right)+\sin^8\left(\frac{\vartheta}{2}\right). \tag{1}$$

It has been found that the maximum amount of gravitational radiation occurs in a direction for which the corresponding electromagnetic radiation is excluded or is a minimal. The *d-wave* wavefunction in HTSCs develops in the *a-b* crystallographic axes (Narlikar, 2003), making a maximum of the gravitational radiation emitted in the direction of the *c* axis. This emission is bi-directional.

Structure

The structure of the proposed HTSC GASER is the one proposed in Fontana and Baker (2003b).

The superposition of a large number of coherent sources is expressed by the interference pattern that they project on a detector at a large distance with respect to the wavelength λ. For instance, the maximum emission for a linear array is for $a \sin\vartheta = n\lambda$, where a is the spacing of the point sources, and n is an integer to which the maximum is associated. The amplitude of the maxima is N^2 times the amplitude produced by each source, where N is the number of sources. More generally a circular planar surface distribution of point sources will produce a beam of radiation orthogonal to the plane of the point sources provided that the wavelength is much shorter than the diameter of the

distribution. The emission is bi-directional. A binary neutron star pair is a celestial example in which the orbit plane is the source of bi-directional GW normal to that orbit plane.

Interestingly it is possible to conceive a distribution of coherent point sources on a spherical surface, as shown in FIGURE 1. For outgoing waves they mimic a point source located at the center of the sphere with a power equal to the coherent summation of the radiation emitted by all the point sources. The radiation directed towards the center is focused to the center of the sphere, therefore a large power density can be obtained. The order of magnitude of the power density gain for a given total input power is A/λ^2, where A is the total area covered by the point sources.

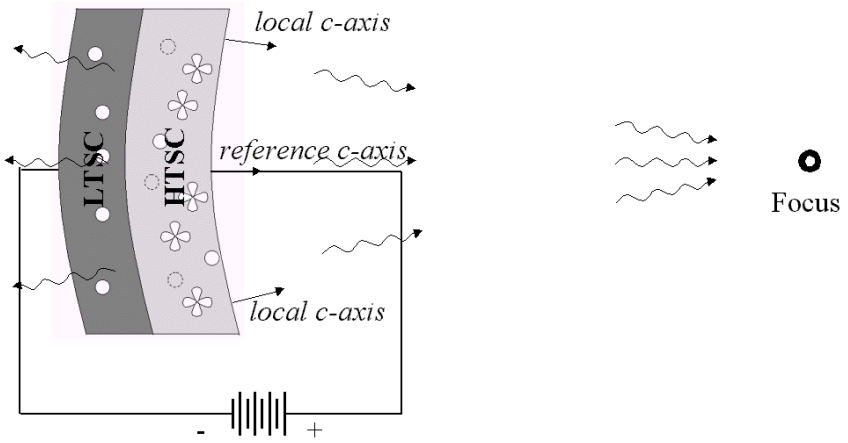


FIGURE 1. The Curved HTSC GASER (Section). The Section is Revolved around the *Reference c-axis*.

To produce curved wavefronts it is necessary to produce a thin HTSC film with the required crystallographic orientation, followed by bending the crystal as required to keep the wavefunction coherence on a curved spherical surface. Alternatively a more complex structure should be devised, based on a spherical surface tiled with planar polygons of HTSC interconnected among each other, each with the c-axis directed toward the focus.

The HTSC GASER is a dual layered device, the HTSC *s-wave/d-wave* layer is the one requiring the highest care, because the crystallographic orientation is important for the device to work properly. We therefore need to discuss the HTSC layer with more detail.

Production techniques and materials.

Orthorombic cuprate HTSC mono-crystals can be produced using different techniques. For bulk material, which appears suitable to planar devices, the melt-textured techniques seem acceptable. This technique requires the sintering of the precursor powders to produce unaligned, grainy HTSC. Afterward, a thermal process termed incongruent melting is performed. The grainy HTSC melts and reacts to form another solid plus a liquid phase. An oriented solidification is then performed, using a $SmBa_2Cu_3O_x$ seed to facilitate the crystallization with the required orientation. Very precise thermal gradients and time-temperature processing are involved. The melt-textured techniques and the related different methods are reviewed by Desgardin (1999).

Thin film techniques seem better suited for making curved GASERs, unfortunately there is no need in current technological applications of curved HTSC films similar to those required here, therefore a new production technology should be developed. Thin film depositions are reviewed by Narlikar (2003). Silicon or Aluminum substrates, buffered by thin layers of CeO_2 and Y-stabilized ZrO_2 can be coated with HTSCs. Thermal coevaporation may offer 20-30 nm/min growth rates. Pulsed laser deposition (PLD) offers similar rates. Magnetron sputtering is widely used for the deposition of layers of metals with high melting temperatures, it easily adapts to the deposition of HTSC with rates of about 10 nm/min. Metalorganic chemical vapor deposition (MOCVD) may offer 1 μm/min. Sol-gel techniques have also been successfully employed for the preparation of thin films of HTSCs.

Because the coherence gain is proportional to the volume of the HTSC, the melt-textured HTSC seems the best choice for a possible first prototype.

The production of HTSC/LTSC Josephson junction is a standard technique, see for instance the article by Moessle (1999). The Pb LTSC can be evaporated on the HTSC surface directly or with the interposition of an Au or Ag interlayer having thickness between 1 and 6 nm. The HTSC GASER must operate at temperatures below the lowest transition temperature of both HTSC and LTSC, i.e. liquid helium for Pb LTSC. The recently discovered superconductor MgB_2 is an extremely interesting option for the LTSC layer because of its high transition temperature of 39 K (Canfield, 2003) and low cost.

The isotopic purity of the raw materials used for sintering the HTSC layer appears a key factor for successfully manufacturing the HTSC GASER. Superconductors in which phonons play a role in inducing electron pairing exhibit a well known isotope effect (Bardeen, 1957; McMillan, 1968). The situation in HTSCs is not completely understood, with experiments giving controversial results, with the latest ones confirming the isotope effect (Bernhard, 2003). The HTSC GASER is based on a very precise transition between two energy levels (macroscopically related to critical temperatures $Tc \propto M^{-\alpha}$, with M mass of the element and α isotope effect). Isotopic mixtures are detrimental at the microscopic scale, where the wavefunction is perturbed with reduction of coherence.

Experimental setup and driving electronics.

It is possible to conceive a setup based on a planar, disk shaped, GASER with area of about 100 cm^2 and thickness of the order of 1 cm. A planar thyristor (SCR – semiconductor controlled rectifier) or a gas thyratron is connected in series with the HTSC GASER and a large set of parallel connected capacitors (C) close the circuit. The capacitors are distributed all around the sandwich. The configuration is intended for pulsed operation with high current. It offers low circuit inductance (L) and low series resistance (R), and is electrically similar to the configuration used for pulsed laser diodes. If the Q of the series equivalent RLC circuit is higher than 1, with a single discharge the GASER will emit a series of exponentially damped pulses of HFGWs with modulation frequency of $1/2\pi\sqrt{(LC)}$. To permit the complete discharge of the capacitors, the thyristor must conduct for the full discharge time and a recirculation diode should be connected in parallel with it. GaAs-based semiconductors can operate at liquid helium temperatures, instead Si-based semiconductors can operate at liquid nitrogen temperatures.

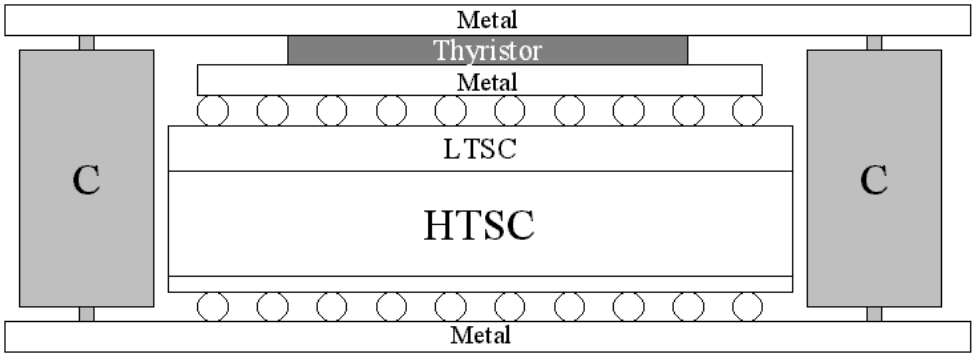


FIGURE 2. The Proposed Experimental Setup (Vertical Section).

FIGURE 2 shows a vertical section of the experimental setup for pulsed operations. Differently from standard Josephson junctions used in research instrumentation, the HTSC GASER is a high power device with current densities up to 10 kA/cm^2. Indium or Bismuth collapsible ball contacts may facilitate the connection of the device to the GASER cage and driving electronics. A voltage generator should be connected to the top and bottom electrodes for charging the capacitors, it should be disconnected before the discharge. Assuming $Q=R^{-1}\sqrt{(L/C)} >> 1$, the peak current is $I_P=V_C\sqrt{(C/L)}$. With capacitance of the order of 10 μF and inductance of the order of 0.1μH, peak currents of the order of 100 kA can be obtained with about 10000 V of charging voltage.

DETECTION OF HFGWs AND EXPECTED RESULTS.

Gravitational waves, possibly emitted by the HTSC GASER, can be detected by specialized detectors for HFGW in the GHz to THz range. It has been shown that the detection of 1.3 THz HFGW with $h \geq 10^{-30}$, as it is expected from the proposed planar HTSC GASER, could be possible with the synchroresonance converter described by Li (2003). Many different detector are also described in the literature as reported by Baker (2004) and in the book by Misner (1973) and many other articles (Thorne, 1980a-b), unfortunately many detectors have been developed for low frequency GWs and it is unclear if they can be adapted to the THz range. Additionally, it should be emphasized that the "classical" theory of gravitational wave propagation and detection is based on the solution of the linearized Einstein equations, intended to treat small perturbations on flat spacetime. Unfortunately for the approach, our spacetime is not flat but is characterized by a constant curvature (Λ term, de Sitter spacetime). Moreover, the energy density near the microscopic sources in the GASER may favor nonlinear effects. Usual simplification conditions seem violated in our approach, which has the potential to make the "classical" detection uncertain.

The expected instanton solutions (Gibbons, 1993; Sanchez, 1982), massive graviton solutions in spacetime with a cosmological constant (Novello, 2003) and self-interaction effects (Ferrari, 1988a and 1988b; Veneziano, 1987; Alekseev, 1996) might indeed simplify the detection of GASER radiation. Standard gravimeters and accelerometers could be the detection instruments. Invoking self-interaction effects, the curved GASER with the parameters already described might produce E≈1mJ of gravitational energy, which is able to accelerate test masses in the vicinity of the focus. The possibility of controlling the operations of the source allows the adoption of lock-in techniques for synchronous detection, thus improving the signal to noise ratio with long integration time.

Claims of observation of gravitational anomalies near HTSC excited by electromagnetic fields have already appeared in the scientific literature (Podkletnov, 1992). Unfortunately, it was not possible to replicate the results till now. This fact might be explained with the incomplete understanding of the phenomenology. Woods (2004) presents a review covering the subject and, more generally, the claims of interaction between gravitation and HTSC.

CONCLUSION

High Frequency Gravitational Waves have great potential for old, new and unexpected applications. Many interesting and new phenomena have been predicted to become possible by using HFGW. Among them we may cite the possibility to communicate directly through objects opaque to electromagnetic radiation and the possibility of 3D imaging of the internal structure of large and very dense objects like the Earth itself (Baker, 2000). Space travel is an old application that may strongly benefit from HFGW. As soon as powerful generators of HFGW will be made available, spacecrafts capable of fast interstellar travel could be constructed. It is superfluous to cite the extraordinary benefits of the discovery of extra-solar planets and their possible life forms. Commercial airlines could operate a fleet of planes with HFGW propulsion, permitting a faster and more comfortable trip to the destination. The HTSC GASER is the proposed device capable of generating ~1 THz HFGWs with high conversion efficiency and is intended to start the new era. This article has presented an approach for the construction of a proof-of-concept prototype, comprising the design of the GASER structure, choice of the materials and basic electronics designed for pulsed operations. Techniques for the detection of the emitted radiation have also been discussed.

ACKNOWLEDGMENTS

The author whishes to acknowledge Dr. Franklyn Mead and AFRL for a grant to attend the STAIF 2004 Conference.

REFERENCES

Alekseev, G. A. and Griffiths, J. B. "Gravitational waves with spherical wavefronts," *Classical and Quantum Gravity,* **12**,, pp. L13-L18 (1995).

Alekseev, G. A. and Griffiths J. B., "Exact solutions for gravitational waves with cylindrical, spherical and toroidal wavefronts," *Classical and Quantum Gravity* **13**, pp. 2191-2209 (1996).

Baker, R. M. L. Jr., "Preliminary Tests of Fundamental Concepts Associated with Gravitational-Wave Spacecraft Propulsion," *American Institute of Aeronautics and Astronautics: Space 2000 Conference and Exposition*, Paper Number 2000-5250, September 20, August 21 (2000).

Baker, R. M. L. Jr., " Precursor Proof-of-Concept Experiments for Various Categories of High-Frequency Gravitational Wave (HFGW) Generators," *STAIF 2004 Conference*, February 8-12 (2004).

Bardeen, J., Cooper, L.N. and Schrieffer,J.R., "Theory of Superconductivity," *Physical Review*, **108**, N. 5, December 1, pp. 1175-1204 (1957).

Bernhard, C. et al. "Anomalous oxygen isotope effect on the in-plane FIR conductivity of detwinned $YBa_2Cu_3{}^{16,18}O_{6.9}$," *arXIV:cond-mat/0306097 v1. 4* Jun (2003).

Canfield, P. C. and Crabtree, G. W., "Magnesium Diboride: Better Late then Never," *Physics Today*, March 34-40 (2003).

Desgardin, G., Monot, I., Raveau, B., "Texturing of high-Tc superconductors" *Supercond. Sci. Technol*. **12** R115–R133 (1999).

Fontana, G., "Gravitational Radiation and its Application to Space Travel," in proceedings *of Space Technology and Applications International Forum (STAIF-2000)* , edited by M. El-Genk, AIP Conference Proceedings 504, New York, 2000, pp. 1085-1092.

Fontana, G., "Gravitational Radiation Applied to Space Travel," *Proceedings of the Gravitational-Wave Conference*, edited by P. Murad and R. Baker, The MITRE Corporation, Mclean, Virginia, May 6-9, Paper HFGW-03-111 (2003a).

Fontana, G. and Baker, R. M. L., Jr., "The High-Temperature Superconductor (HTSC) Gravitational Laser (GASER)," *Proceedings of the Gravitational-Wave Conference*, edited by P. Murad and R. Baker, The MITRE Corporation, Mclean, Virginia, May 6-9, Paper HFGW-03-107 (2003b).

Ford, L. H. "Gravitational Radiation by Quantum Systems," *Annals of Physics,* **144**, pp. 238-248 (1982).

Ferrari V., Pendenza P., and Veneziano G., "Beam-like Gravitational Waves and Their Geodesics*," General Relativity and Gravitation*, **20**, No 11, pp. 1185-1191 (1988a).

Ferrari Valeria, "Focusing Process in the Collision of Gravitational Plane Waves," *Physical Review D*, **37**, No10, 15 May, pp. 3061-3064 (1988b).

Gibbons, G. W. and Hawking, S. W., *Euclidean Quantum Gravity,* World Scientific, Singapore, New Jersey, London, Hong Kong, ISBN 981-02-0516-3, pp. 495- (1993).

Halpern, L., Laurent, B., "On the Gravitational Radiation of Microscopic Systems," *Il Nuovo Cimento*, **XXXIII**, N. 3, pp. 728-751 (1964).

Kouznetsov K.A. et al., "c-axis Josephson Tunneling between $YBa_2Cu_3O_{7-x}$ and Pb: Direct Evidence for Mixed Order Parameter Symmetry in a High Tc Superconductor," *Physical Review Letters*, **79**, 20 Oct., pp. 3050-3053 (1997).

Landau, L. D. and. Lifshitz, E. M., *The Classical Theory of Fields*, Fourth Revised English Edition, Pergamon Press, p. 349, (1975).

Li, Fang-Yu, Tang, Meng-Xi and Shi, Dong-Ping , "Electromagnetic response for High-Frequency Gravitational Waves in the GHz to THz band," *Gravitational-Wave Conference*, edited by P. Murad and R. Baker*,* The MITRE Corporation, Mclean, Virginia, May 6-9, Paper HFGW-03-108 (2003).

McMillan, W. L., "Transition Temperature of Strong-Coupled Superconductors," *Physical Review,* **167**, 331-344 (1968).

Mishonov, T. M. , Indeku, J. O. and Penev, E. S., "Superconductivity of overdoped cuprates: the modern face of the ancestral two-electron exchange," *J. Phys.: Condens. Matter* **15** No 25, 4429-4456 (2003).

Misner, C. W., Thorne, K. S. and Wheeler, J. A., *Gravitation,* Freeman (1973).

Moessle, M. and Kleiner, R., "c-axis Josephson tunneling between $Bi_2Sr_2CaCu_2O_{8+x}$ and Pb," *Physical Review B*, **59** N. 6, 4486-4496 (1999).

Narlikar, A. V., "High Temperature Suprconductivity –1. Materials," Springer , Berlin, ISBN 3-540-40631-X (2003).

Novello, M. and Neves, R. P., "The mass of the graviton and the cosmological constant," *Classical and Quantum Gravity* **20** L67-L73 (2003).

Pinto, I.M. and. Rotoli, G., "Laboratory generation of gravitational waves?," *Proceedings of: 8th Italian Conference on General Relativity and Gravitational Physics*, Cavalese (Trento), August 30 – September 3, World Scientific (1988).

Podkletnov, E. and Nieminen, R., "A Possibility of Gravitational Force Shielding by Bulk $YBa_2Cu_3O_{7-x}$ Superconductor," *Physica C*, **203**, pp. 441-444 (1992).

Sanchez, N., "Connection between the nonlinear σ model and the Einstein equations of general relativity*," Physical Review D* **26**, 2589-2597 (1982).

Sanchez, N., "New solutions of Einstein equations from analytic mappings," *Physics Letters* **94A**, 125-130 (1983).

Szekeres, P., "Colliding Plane Gravitational Waves," *J. Math. Phys*, **13**, No. 3, March, pp. 286-294 (1972).

Thorne Kip S., "Gravitational-wave Research: Current Status and Future Prospects," *Reviews of Modern Physics*, **52**, No. 2, Part I, April, pp. 285-297 (1980a).

Thorne, Kip S., "Multipole Expansion of Gravitational Radiation," *Reviews of Modern Physics*, **52**, No. 2, Part I, April, pp. 299-339 (1980b).

Veneziano, G., "Mutual Focusing of Graviton Beams,",*Modern Physics Letters A*, **2**, No 11, pp. 899-903 (1987).

Woods, R. Clive, "Review of Claims of Interaction between Gravitation and High-Temperature Superconductors," *STAIF 2004 Conference*, February 8-12 (2004).

Earth Planets Space, **52**, 579–586, 2000

Amplification of gravity and Rayleigh waves in a layered water-soil model

Tatyana Novikova[1], Kuo-Liang Wen[2], and Bor-Shouh Huang[1]

[1]*Institute of Earth Sciences, Academia Sinica, Taiwan*
[2]*Institute of Geophysics, NCU, Taiwan*

(Received August 23, 1999; Revised August 18, 2000; Accepted August 25, 2000)

The coupled seismic to gravitational surface wave fields are analyzed in a liquid layer lying on the gravitating elastic, low-rigidity half-space. Solution is obtained within the framework of the normal mode formalism applied to the flat ocean-solid Earth model. From the theory of propagation of coupled surface waves (Rayleigh and Love) in layered media, we find the individual multipliers that determine the surface wave spectrum over the entire frequency range. Spectra of excitation functions are investigated for dip-slip point source in the half-space. Main results can be summarized as follows. When the half-space is filled with sediments, dip-slip excitation functions of gravity and Rayleigh waves are one order of magnitude larger than for the half-space composed of hard rocks. Including gravity in the elastic medium essentially changes the character of gravity wave spectrum, leading to an appearance of the third maximum. At the deepening of the source amplitude of this maximum increases. Theoretical marigrams show that including gravity in the half-space also increases period of the gravity wave excited by deep sources by a factor of two, up to 10 minutes. At the same time, presence of gravity force in the half-space has no effect on the spectrum of the Rayleigh wave.

1. Introduction

Considerable effort has been taken to develop both the analytical description of the propagation process of long period gravity (tsunami) and ocean Rayleigh waves for the case of uniform depth (Yamashita and Sato, 1974; Ward, 1980; Comer, 1984a, b; Okal, 1988; *et al.*) and the numerical one for actual bathymetry (e.g. Hwang, 1972; Satake, 1985).

Yamashita and Sato (1974, 1976), using the fully coupled ocean-solid Earth model, analyzed the influence of a number of focal parameters-dip-angle, fault length and focal depth and the rise time of the source time function- on the tsunami (long period gravity wave) and Rayleigh waves. Earthquake source was modeled by the point dislocation and a finite-moving source. Numerical computations have been performed for the dip-slip source model, because large submarine earthquakes near Japan are of this type. Their results can be summarized as follows. In the case of tsunami wave, such parameters as dip-angle, fault length and depth all play an important role in generation process. The maximum amplitudes in the wave trains at azimuths $\varphi = 90^\circ$ and $\varphi = 270^\circ$ become large for large dip-angles. In the case of $\varphi = 0^\circ$, the maximum amplitude is the largest when dip-angle is 45°. The fault length also significantly affects the directivity. When the ratio- of 'fault width to its length increases, the directivity at the azimuth $\varphi = 90^\circ$ and $\varphi = 270^\circ$ become very strong. Velocity of rupture propagation and rise time do not affect much the tsunami amplitude and waveform. In the case of Rayleigh wave, the spectral amplitude decreases monotonously, with as the increase of focal depth. The dip-angle factor has the entirely opposite effect compared to the tsunami wave. The spectral amplitude of Rayleigh wave predominates for smaller dip-angle. The rise time parameter also affects strongly the Rayleigh wave, which become essentially suppressed-as rise time increases.

Ward (1980) considered the tsunami generation problem in the context of a spherically symmetric self-gravitating and elastic ocean-Earth model using the normal mode formalism. He formulated and solved full linearized equations of motion in a manner similar to that originally applied to the Earth's free oscillations. Using moment tensor representation for both point and line seismic source he has derived expressions for tsunami wave displacement in near- and far-field zones.

Consequently Comer (1984a, b) introduced a solution for the tsunami mode excitation in the flat Earth by a point source. He emphasized that the excitation problem in the flat Earth differs substantially from the corresponding problem for the spherical Earth. There is a continuum of frequencies and wavenumbers for the flat Earth but frequencies and angular orders are discrete on the spherical Earth. Also, the normal modes of a finite body form a complete basis for the small oscillations of the body but those of an infinite body, like the flat Earth, do not. In common with Ward, Comer assumed the ocean to be non-viscid and only considered linearized equations of motion and boundary conditions. Alike the earlier authors investigations (Pod"yapol'sky, 1968, 1970; Alexeev and Gusiakov, 1976) he has shown that far-field tsunami in the fully coupled ocean-solid Earth model depend strongly on the source depth, duration, moment and mechanism.

Okal's study (1988) was largely based on Ward formalism considering tsunami wave as the superposition of the free oscillations of an elastic self-gravitating Earth, excited by a seismic sources. Extending the range of source depths to 250

Copy right© The Society of Geomagnetism and Earth, Planetary and Space Sciences (SGEPSS); The Seismological Society of Japan; The Volcanological Society of Japan; The Geodetic Society of Japan; The Japanese Society for Planetary Sciences.

km he showed that this parameter plays only a limited role in controlling the tsunami amplitude. More important are the effects of directivity due to rupture propagation along the fault and the possibility of enhanced tsunami excitation in an elastically weak material. Analyzing tsunami and Rayleigh wave generation by non-double couple sources, he found, in particular, that a landslide involving weak sediments could result in very large tsunami.

In the description of tsunami and Rayleigh waves propagation in most models, discussed above, the gravity force was included only in the liquid layer. Since in the present study we take gravity into account also in the solid Earth, we briefly review the previous works in this direction.

Surface wave propagation in the elastic medium under the influence of gravity was considered by a number of authors (Matuzawa, 1925; Gilbert, 1967; Lomnitz, 1970, 1990, 1991; De and Sengupta, 1976). Gilbert (1967) described a gravitationally perturbed Rayleigh wave in a low-rigidity medium which forms a linear continuum ranging from a solid rocks to water. In this approach everything is attributed to the variation of a single parameter- the rigidity. Gilbert has made energy estimations for surface waves and showed that in a low rigidity medium such as unconsolidated sediments the energy of gravity is comparable with the shear strain energy. In the case of Rayleigh wave propagated in the non-gravitating half-space the $\overline{S}$ is more dominant pulse. In the present of gravity, as β decreases, the $\overline{S}$ pulse becomes insignificant, and the $\overline{P}$ pulse becomes dispersive and approaches the behavior of the classical gravity wave.

Lomnitz (1990), analyzing the behavior of sediments during the Mexico 1985 earthquake, has supposed that a nonlinear mechanism could make the rheology of the Mexico City soft clay nearer to a fluid than to a solid. At large amplitudes, when the stress-strain relation becomes strongly nonlinear sediments may liquefy, and in such conditions the gravity wave propagates as in a liquid. It is worth mentioning here that originally Matuzawa (1925) proposed the idea of existence and propagation of waves of hydrodynamic origin in soils. Based on the study of Tokyo 1923 earthquake, he speculated that sediments could behave like solids at high frequencies and like fluids at low frequencies.

Dealing with gravity wave propagation in the system consisting of the liquid layer and elastic half-space it is possible to neglect the gravity in an elastic medium, since it is small in comparison with the elastic forces. General effect of gravity on the wave propagation in such a system is determined by gravity effect in the liquid. When the ocean bottom consists of rigid rocks the above assumption is reasonable. Although in this case too, including the gravity force in the motion equations of the elastic substratum, changes slightly the dispersion curve at the 10000 seconds period (Pod"yapol'sky, 1968, 1970; Alexeev and Gusiakov, 1974). In the present work, however, we focus on the wave excitation by seismic sources located in low-rigidity media. Therefore it is appropriate not to omit gravity from the description of this medium, since its effects can be evaluated to some extent even within the framework of linear theory.

Present study is focused on the developing an analytical model of gravity and Rayleigh waves excitation by a realistic earthquake source located in the gravitating low-rigidity half-space. We do not deal in detail with the effects of focal parameters on the gravity and Rayleigh wave generation, because it was already done in previous studies mentioned above. It is worth noting that in the case of gravity wave we analyze the wave spectrum in the whole range of frequencies, not limited by its long-period component (tsunami). We will describe the seismic response in function of source location in the half-space and compare the results obtained for a liquid layer and a gravitating half-space model with a liquid layer plus a non-gravitating half-space.

2. Statement of the Problem

Consider gravity and Rayleigh waves in a compressible liquid layer of uniform thickness H over a half-space. Gravity acts in both media. We assume: (1) both media are homogeneous; (2) in the solid medium stress and deformation are related through Hooke's law and in the liquid pressure is proportional to the degree of compression; (3) deformations are small and displacements in the liquid are small compared to the liquid layer thickness and characteristic wavelengths. The latter condition allows us to neglect the distinction between Eulerian and Lagrangian variables in the equations of motion (Pod"yapol'sky, 1968, 1970; Comer, 1984a).

Linear approximation is used to formulate the equations of motion. This means that we neglect the variation of the displacement field on distances comparable with displacement. We also neglect the displacements of the boundary, since these displacements are very small compare to the characteristic wavelengths.

Based on the above approximation and on the main postulates of the fluid dynamics (Landau and Lifschitz, 1980) the linearized equations for the dynamic displacement field in liquid and in elastic medium will be respectively (after Pod"yapol'sky, 1968, 1970):

$$c_f^2 \nabla \mathrm{div}\mathbf{u}_1 - g\mathbf{e}_z \mathrm{div}\mathbf{u}_1 = \frac{\partial \mathbf{u}_1}{\partial t^2} \qquad 0 < z < H \quad (1)$$

$$a^2 \nabla \mathrm{div}\mathbf{u}_2 - b^2 \nabla \times (\nabla \times \mathbf{u}_2) - g\mathbf{e}_z \mathrm{div}\mathbf{u}_2 = \frac{\partial^2 \mathbf{u}_2}{\partial t^2} \qquad z > H, \quad (2)$$

Boundary conditions at the ocean surface ($z = 0$) and the bottom ($z = H$) are:

$$c_f^2 \mathrm{div}\mathbf{u} - gW \Big|_{z=0} = 0, \quad (3)$$

$$W_1(H) = W_2(H), \quad (4)$$

$$p_f \left[c_f^2 \mathrm{div}\mathbf{u} - gW_1 \Big|_{z=H} \right] = \rho \left[\left(a^2 - 2b^2\right) \mathrm{div}\mathbf{u} + 2b^2 \frac{\partial W_2}{\partial z} - gW_2 \Big|_{z=H} \right]$$

$$\mu \left(\frac{\partial W_2}{\partial x} + \frac{\partial U_2}{\partial z} \right) \Big|_{z=H} = 0,$$

$$\mu \left(\frac{\partial V_2}{\partial z} + \frac{\partial W_2}{\partial y} \right) \Big|_{z=H} = 0, \quad (5)$$

At infinity the displacement vanishes:

$$\mathbf{u} \to 0 \quad \text{at} \quad z \to 0. \quad (6)$$

Here c_f is the velocity of the acoustic wave in the liquid layer, ρ_f is the density of the liquid, H is the thickness of the liquid layer, λ and μ are the Lame coefficients in the elastic half-space, ρ is the density of the elastic medium, a and b are the velocities of P- and S-waves, $\mathbf{u}_i = (U_i, 0, W_i)$ is the displacement vector ($i = 1$ liquid layer, $i = 2$ elastic half-space), g is gravity, is the angular frequency, and z is the vertical distance from the initial (undisturbed) position of the free surface (positive downwards).

3. Solution

The solution of Eqs. (1) and (2) may be written in the form of a stationary plane wave propagating along the x axis

$$\mathbf{u}_i(x, z, \omega, t) = \mathbf{v}_i(z, \omega) \exp\left[i\omega(t - x/c)\right]. \tag{7}$$

Substituting (7) into Eqs. (1) and (2) and introducing the boundary condition at $z \to \infty$, the wave amplitudes as a function of depth are found as

$$\begin{aligned} U_1(z, \omega) = {} & \frac{ic_f^2}{\omega c}\left[-B\exp(-\eta_2 z/c_f)\right. \\ & \left. + C\exp(-\eta_1 z/c_f)\right], \end{aligned} \tag{8}$$

$$\begin{aligned} W_1(z, \omega) = {} & -\frac{c_f}{\omega^2}\left[-\eta_1 B\exp(-\eta_2 z/c_f)\right. \\ & \left. + \eta_2 C\exp(-\eta_1 z/c_f)\right], \end{aligned} \tag{9}$$

where, for the liquid layer,

$$\eta_1 = -\omega\gamma - g/2c_f \tag{10}$$

$$\eta_2 = -\omega\gamma - g/2c_f \tag{11}$$

$$\gamma^2 = \frac{c_f^2}{c} - 1 + \frac{g^2}{4c_f^2\omega^2}, \tag{12}$$

and for the half-space,

$$\begin{aligned} U_2(z, \omega) = {} & \frac{ia^2}{\omega c} D\exp(-\omega\alpha(z - H)/a) \\ & - \frac{ib\beta}{\omega} F\exp(-\omega\beta(z - H)/b), \end{aligned} \tag{13}$$

$$\begin{aligned} W_2(z, \omega) = {} & \frac{a\alpha}{\omega}\left(1 + \frac{g}{2\omega\alpha a}\right) \\ & \cdot D\exp\left(\left(-\omega\alpha + \frac{g}{2a}\right)(z - H)/a\right) \\ & - \frac{b^2}{\omega c} F\exp(-\omega\beta(z - H)/b), \end{aligned} \tag{14}$$

where

$$\alpha^2 = a^2/c^2 - 1 + \frac{g^2}{4a^2\omega^2}$$

and

$$\beta^2 = b^2/c^2 - 1.$$

Equations (3) to (14) yield a set of four homogeneous equations in four unknown coefficients, namely B, C, D and F, which has a non-trivial solution when the determinant is zero.

The dispersion equation is obtained as follows:

$$\begin{aligned} & m\Big[-\gamma\cosh\left(\omega\gamma H/c_f\right) \\ & \quad + \frac{g}{2c_f\omega}\left(\frac{2c_f^2}{c^2} - 1\right)\sinh\left(\omega\gamma H/c_f\right)\Big] \\ & \quad \cdot\left[-a^2\left(1 - \frac{2b^2}{c^2} - \frac{g^2}{2\omega^2a^2}\right) + 4b^3a\alpha\beta/c^2\right] \\ & \quad - c_f\left(a\alpha - \frac{g\beta^2}{2\omega}\right) \\ & \quad \cdot\left[\frac{g\gamma}{c_f\omega}\cosh(\omega\gamma H/c_f) - \left(1 - \frac{g^2}{2c_f^2\omega^2}\right)\right. \\ & \qquad \left. \cdot\sinh\left(\omega\gamma H/c_f\right) - \frac{g\gamma}{c_f\omega}\exp(-gH/2c_f^2)\right] \\ & \quad - \frac{ga\alpha}{\omega^2\left(1 - \frac{c_f^2}{c^2}\right)} \\ & \quad \cdot\left[\omega\gamma\exp\left(-gH/2c_f^2\right) - \omega\gamma\cosh\left(\omega\gamma H/c_f\right)\right. \\ & \qquad \left. - \frac{g}{2c_f}\left(\frac{2c_f^2}{c^2} - 1\right)\sinh\left(\omega\gamma H/c_f\right)\right] = 0, \end{aligned} \tag{15}$$

where

$$m = \frac{\rho}{\rho_f}.$$

We follow the approach of surface-wave theory. Consider the displacement of a stationary surface wave excited by a point source in a homogeneous half-space (Aki, 1980; Keilis-Borok, 1989):

$$\begin{aligned} \mathbf{u}(r, \varphi, \omega, t) = {} & \frac{\exp(-i\pi/4)}{\sqrt{8\pi}}\frac{\exp(i\omega(t - r/c))}{\sqrt{\omega\ \ r/c}} \\ & \cdot\frac{\mathrm{U}(z, \omega)}{\sqrt{cuI_0}}\frac{\mathrm{Q}(h, \varphi, \omega)}{\sqrt{cuI_0}}; \end{aligned} \tag{16}$$

for a constant thickness of the liquid layer, or

$$\begin{aligned} \mathbf{u}(r, \varphi, \omega, t) = {} & \frac{\exp(-i\pi/4)}{\sqrt{8\pi}} \\ & \cdot\frac{\exp\left(i\omega\left(t - \int_M^N \frac{dx}{c(\omega, x)}\right)\right)}{\sqrt{\omega J(x)/c}} \\ & \cdot\left(\frac{\mathbf{U}(z, \omega)}{\sqrt{cuI_0}}\right)_N\left(\frac{\mathrm{Q}(h, \omega)}{\sqrt{cuI_0}}\right)_M, \end{aligned} \tag{17}$$

for a variable thickness of the liquid layer, where $\mathbf{U}(z, \omega) = U(z, \omega)\mathbf{e}_r + W(z, \omega)\mathbf{e}_z$, $\mathrm{Q}(h, \varphi, \omega) = m_{rs}(\omega)B_{rs}(h, \varphi, \omega)$ is the excitation function, $m_{rs}(\omega)$ is the spectrum of the seismic moment tensor, $B_{rs}(h, \varphi, \omega)$ is a tensor which can be expressed via the eigenfunctions $U(z, \omega)$ and $W(z, \omega)$ and their derivatives and which depends on the axis orientation of the source,

$$I_0 = \int_0^\infty \rho\left[U^2(z, \omega) + W^2(z, \omega)\right]dz$$

Table 1.

Media density [g/cm^3]		Wave velocities [km/s]		Layer thickness (H-for liquid) [km.]; depth of source location (h-for half-space) [km.]
Model I				
Liquid layer	1	Acoustic wave	1.45	$H = 4$
Half-space	2.03	P-wave	2.0	$h = 10$–30
		S-wave	1.15	
Model II				
Liquid layer	1	Acoustic wave	1.45	$H = 4$
Half-space	2.5	P-wave	3.9	$h = 10$–30
		S-wave	2.3	
Model III				
Liquid layer	1	Acoustic wave	1.45	$H = 4$
Half-space	3.1	P-wave	7.15	$h = 10$–30
		S-wave	4.1	

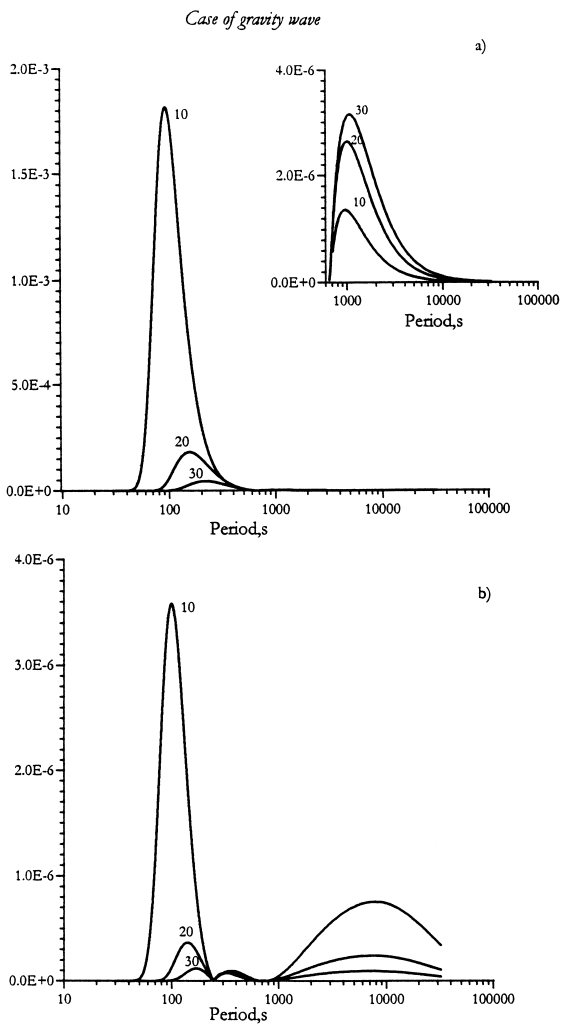


Fig. 1. The dependence of $|q(h, \omega, \varphi)|$ on period for the dip-slip source located a) in non-gravitating half-space, b) in gravitating half-space. Numbers at the curves indicate the source depth in half-space.

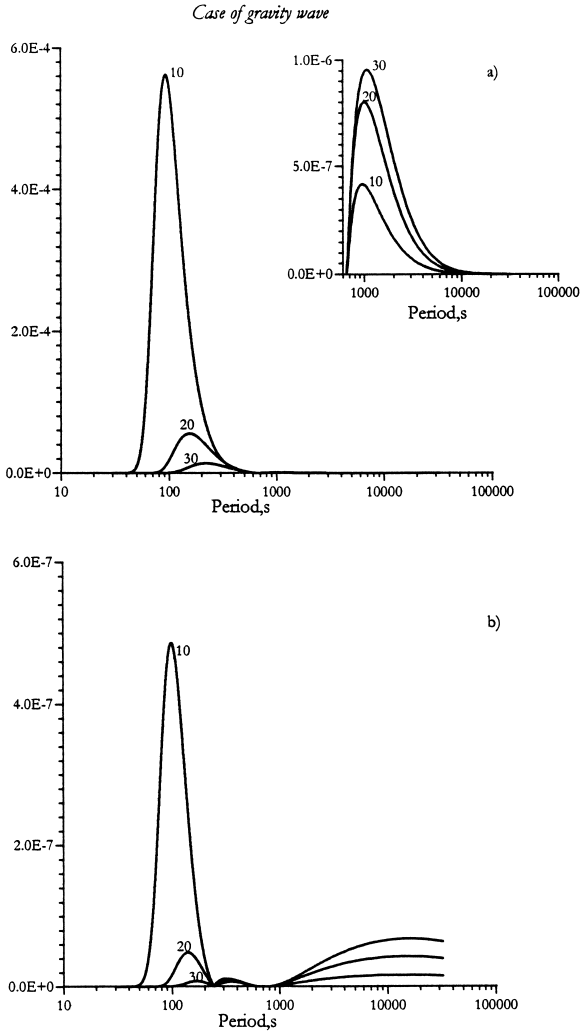


Fig. 2. The dependence of $|q(h, \omega, \varphi)|$ with period for dip-slip source located a) in non-gravitating half-space, b) in gravitating half-space. Numbers at the curves indicate the source depth in half-space.

is the energy integral, u is the group velocity and $J(x)$ is the geometrical divergence in the x-y plane.

In Eqs. (16) and (17), the second term describes the effect of geometrical divergence of the energy flow on the wave propagation. The third term depends on the depth of the receiver, and the fourth term depends on depth, focal mechanism and radiation spectrum in the half-space.

As we shall see, the presence of gravity in the half-space leads to new properties of the gravity wave spectrum. In the following we consider the case of surface waves excited by a dip-slip point source.

4. Dip-Slip Source

Let the fault plane be orthogonal to the x-axis and let displacement be vertical. Then the seismic moment tensor has two non-zero components:

$$m_{xz} = m_{zx} = M_0 F(\omega), \tag{18}$$

where M_0 is the seismic moment and $F(\omega)$ is the seismic moment spectrum. The corresponding components of tensor **B** are

$$B_{xz} = B_{zx} = \frac{i\cos\varphi}{2}\left(-W_2 + \frac{dU_2}{dz}\right). \tag{19}$$

Thus

$$Q(h,\omega,\varphi) = iM_0\cos\varphi\left(-\xi W_2 + \frac{dU_2}{dz}\right)F(\omega), \tag{20}$$

$$q(h,\omega,\varphi) = \frac{i\cos\varphi\left[-\xi W_2(h,\omega) + \left.\frac{dU_2(z,\omega)}{dz}\right|_{z=h}\right]}{\sqrt{cuI_0}}, \tag{21}$$

where $\xi = \omega/c$ is the wave number.

To estimate the source radiation function, some value of the seismic moment M_0 and its spectrum $F(\omega)$ must be assumed. We assume values of the radiation function up to the factor M_0 and for that azimuth along which seismic radiation is maximal. Thus

$$F(\omega) = \frac{1}{i\omega(i\omega\tau_0 + 1)}, \tag{22}$$

where τ_0 is the rise time.

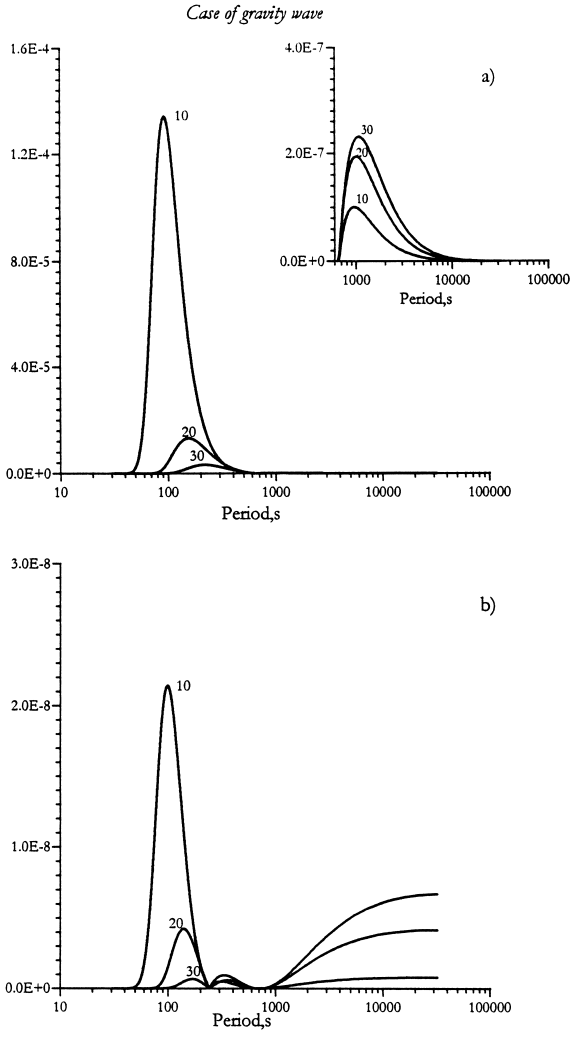


Fig. 3. The dependence of $|q(h,\omega,\varphi)|$ with period for dip-slip source located a) in non-gravitating half-space, b) in gravitating half-space. Numbers at the curves indicate the source depth in half-space.

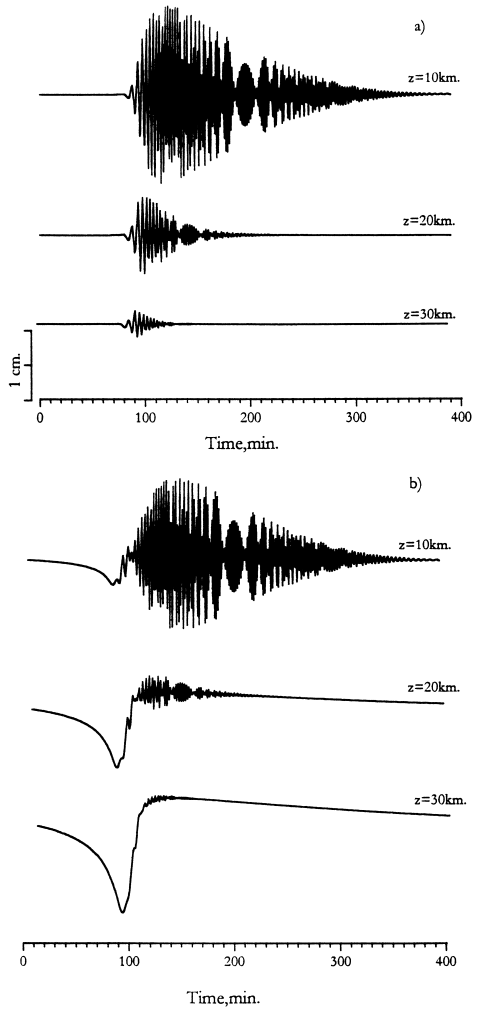


Fig. 4. A comparison of theoretical marigrams of gravity wave from the dip-slip source for model I a)- with non-gravitating half-space, b)- with gravitating half-space. The right numbers at the curves indicate the source depth in half-space.

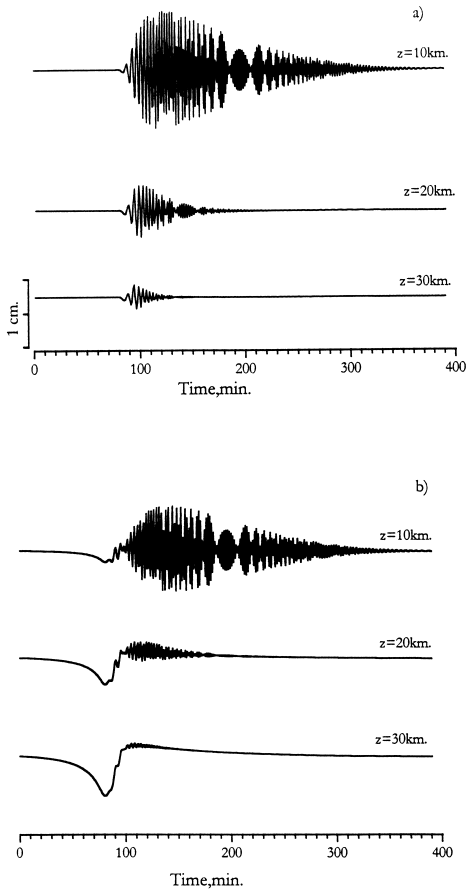


Fig. 5. A comparison of theoretical marigrams of gravity wave from the dip-slip source for model III a)- with non-gravitating half-space, b)- with gravitating half-space. The right numbers at the curves indicate the source depth in half-space.

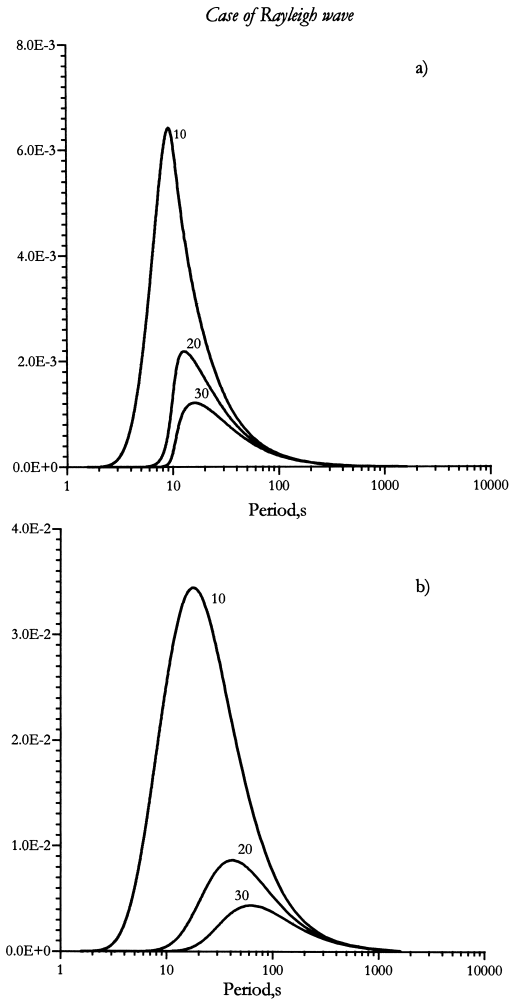


Fig. 6. The dependence of $|q(h, \omega, \varphi)|$ on period for the dip-slip source located a) in gravitating half space model III , b) in gravitating half-space model I. Results without gravitational force are identical. Numbers at the curves indicate the source depth in half-space.

5. Numerical Results

The parameters of our model are shown in Table 1 (Nafe and Drake, 1963).

We focused on the influence of gravitating sediments on excitation of gravity and Rayleigh waves. Although in real Earth thickness of ocean sediments seldom exceeds a few kilometers (Mooney *et al.*, 1998), for our numerical experiment we used structures with thick sedimentary mass like deep-sea trenches (Yoshii *et al.*, 1970; Westbrook *et al.*, 1973), and we vary the source location from shallow to deep. In spite of that we are using real paremeters of the ocean-Earth structure, proposed model is still a theoretical one, whose purpose is to explore some physical aspects of the role of sediments in the excitation of gravity and Rayleigh waves.

5.1 Case of gravity wave

As shown previously (Novikova, 1997), compressibility of the liquid leads to the appearance of two maxima (Figs. 1(a)–3(a)) in the spectrum of gravity waves, namely a high-frequency maximum at periods of 100–200 s and a low-frequency maximum around at 1000 s. For all liquid half-space models (Table 1) in the 10–30 km depth interval, the amplitude of the high-frequency maximum is larger than that of the low frequency maximum. When the half-space is composed of sediments excitation function of gravity wave is one order of magnitude larger compared with that in hard rock half-space (Figs. 1(a)–3(a)). Theoretical marigrams (vertical component) of gravity wave, calculated by integrating formula (16) at the epicentral distance of 1000 km, prove this as well (Figs. 4(a) and 5(a)). We assume that the seismic moment M_0 for the source model is 6×10^{21}Nm, corresponding roughly to an earthquake of magnitude 8.5.

By including the action of gravity in the half-space character of the excitation function spectrum changes. Third maximum appears (Figs. 1(b)–3(b)) and at the deepening of source its amplitude increases. Theoretical marigrams (Figs. 4(b) and 5(b)) show that including gravity in the half-space increases period of the gravity wave excited by deep sources by a factor of two, up to 10 minutes.

One possible explanation of this third maximum is that including gravity in the elastic medium results in the additional (gravity) root in the dispersion equation. This root corresponds to a gravity wave propagating at the liquid-elastic medium interface (Gusiakov, 1972) and cases the change in excitation spectrum.

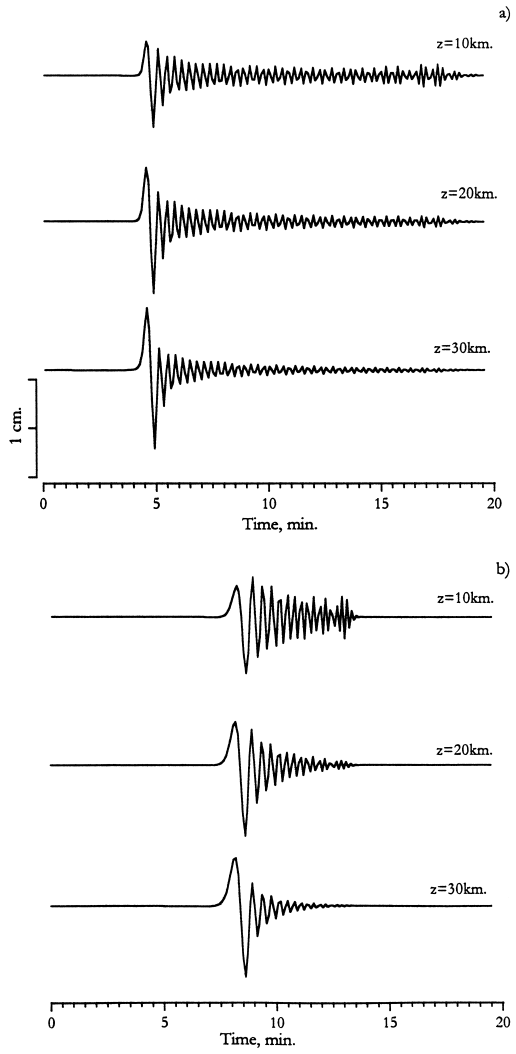


Fig. 7. A comparison of theoretical marigrams of Rayleigh wave from the dip-slip source a) for model III, b) for model II with gravitating half-space. Results without gravitational force are identical. The right numbers at the curves indicate the source depth in half-space.

5.2 Case of Rayleigh wave

Figures 6 and 7 demonstrate that low-rigidity media also amplify the excitation and propagation of Rayleigh waves.

At the deepening of source Rayleigh wave amplitude decreases, similar to the case of gravity wave. Including gravity in the half-space has no effect on the Raylegh wave (excitation function curves and theoretical marigrams with and without gravity coincide). It can be explained as follows. In case of Rayleigh wave the restoring force is elastic, not gravity. Gilbert (1967) showed that influence of gravity acting in an elastic medium, on a surface wave is important only if the medium is filled with soft sediments with wave velocities $b < 100$ m/s. In our model, however, s-wave velocities vary from 1.1 to 4.1 km/s, i.e. our model medium is rigid enough to suppress action of gravity on surface wave.

6. Conclusions

Excitation and propagation of gravity and Rayleigh waves are treated in the frame of model considering them as surface waves of the coupled type in a liquid layer overlying the gravitating half-space filled with sediments. Normal mode formalism was applied to calculate far-field displacement of surface waves generated by dip-slip point source located at different depths in the solid Earth underneath the ocean.

Using the liquid-solid model with different parameters of half-space we found that low-rigidity medium may be seen as an amplifier for both gravity and Rayleigh waves.

Including gravity in the half-space changes the character of excitation spectrum of gravity wave. The difference is more obvious for the gravity wave excited by a deep source in the weak media. In such conditions gravity wave becomes more long-period.

Proposed model is the theoretical one aimed to study the influence of gravitating sediments on the wave excitation. Consequently, for future investigations, it is assumed to use a more realistic structure with layered ocean crust. In particular, such model will be important for the study of Rayleigh waves that depend considerably on the layered structure (especially in the 1–10 seconds period range).

In further study of gravity waves, the results of the present work will be generalized for the case of "tsunami-earthquake" events, using a more realistic model of a finite size seismic source. Recent investigations (Houston, 1999; Bilek and Lay, 1999) suggested that some earthquakes rupture slowly because of the presence of shallow, unconsolidated sediments. The process of gravity wave excitation by a slow rupture on a shallow fault may be relevant, as slow rupture can generate large low-frequency gravity waves which take the form of tsunami waves at the ocean surface.

Acknowledgments. The authors would like to thank Dr. C. Lomnitz and an anonymous reviewer for suggestions that helped to improve an original version of the manuscript. This research was supported by a grant of Academia Sinica Institute of Earth Sciences, Taiwan.

References

Aki, K. and G. Richards, *Quantative Seismology*, v.1., 557 pp., W. H .Freeman, San Francisco, 1980.

Alexeev, A. and V. Gusiakov, Numerical modeling of tsunami and seismic surface wave generation by a submarine earthquake, in *Tsunami Research Symposium, Bull. R. Soc. N. Z.*, **15**, 243–251, 1976.

Bilek, S. and T. Lay, Rigidity variations with depth along interplate megathrust faults in subduction zones, *Nature*, **400**, 443–446, 1999.

Comer, R., The tsunami mode of a flat earth and its excitation by earthquake sources, *Geophys. J. R. astr. Soc.*, **77**, 1–27, 1984a.

Comer, R., Tsunami generation: a comparison of traditional and normal mode approaches, *Geophys. J. R. astr. Soc.*, **77**, 29–41, 1984b.

De, S. N. and P. R. Sengupta, Surface waves under the influence of gravity, *Gerlands Beitr. Geophys.*, **85**, 311–318, 1976.

Gilbert, F., Gravitationally perturbed elastic waves, *Bull. Seism. Soc. Am.*, **57**, 783–794, 1967.

Gusiakov, V., *Excitation of Tsunami and Oceanic Rayleigh Waves by Submarine Earthquake. Mathematical Problems in Geophysics*, pp. 250–267, Novosibirsk, 1972.

Houston, H., Slow ruptures, roaring tsunamis, *Nature*, **400**, 409, 1999.

Hwang, L.-S., H. L. Butler, and D. J. Divoky, Tsunami model: Generation and open-sea characteristics, *Bull. Seism. Soc. Am.*, **62**, 1579–1596, 1972.

Kanamori, H., Mechanism of tsunami earthquakes, *Phys. Earth Planet. Inter.*, **6**, 346–359, 1972.

Keilis-Borok, V., *Seismic Surface Waves in a Laterally Inhomogeneous Earth*, 293 pp., Kluwer Academic Publishers, 1989.

Landau, L. and E. Lifshitz, Fluid mechanics, in *Courses of Theoretical Physics*, v.6, 1980.

Lomnitz, C., Some observations of gravity waves in the 1960 Chile earthquake, *Bull. Seism. Soc. Am.*, **59**, 669–670, 1970.

Lomnitz, C., Mexico 1985: the case for gravity waves, *Geophys. J. Int.*,

102, 569–572, 1990.

Lomnitz, C., On the transition between Rayleigh waves and gravity waves, *Bull. Seism. Soc. Am.*, **81**, 273–275, 1991.

Matuzawa, T., On the possibility of the gravitational waves in soil and allied problems, *J. Inst. Astr. Geophys. Tokyo*, **3**, 161–174, 1925.

Mooney, W., G. Laske, and T. Masters, Crust 5.1: A global crustal model at 5^0x5^0, *J. Geophys. Res.*, **103**, 727–747, 1998.

Nafe, J. and C. Drake, Physical properties of marine sediments, in *The Sea*, vol. 3, edited by M. H. Hill, pp. 794–815, Interscience Publishers, New York, 1963.

Novikova, T., Numerical modeling of the tsunami generation by seismic sources, Ph.D.thesis Earth Physics Department, Institute of Physics, St. Petersburg University, 98 pp., 1997.

Okal, E., Seismic parameters controlling far-field tsunami amplitudes: a review, *Natural Hazards*, **1**, 67–96, 1988.

Pod"yapol'sky, G. S., Excitation of a long gravitational wave in the ocean from a seismic source in the crust, *Izv. AN SSSR, Fizika Zemli*, **1**, 1968 (in Russian).

Pod"yapol'sky, G. S., Generation of the tsunami wave by the earthquake in Tsunamis in the Pacific Ocean, edited by W. M. Adams, pp. 19–32, East-west Center Press, Honolulu, 1970.

Satake, K., The mechanism of the 1983 Japan Sea earthquake as inferred from long-period Surface waves and tsunamis, *Phys. Earth Planet. Inter.*, **37**, 249–260, 1985.

Ward, S., Relationships of tsunami generation and an earthquake source, *J. Phys. Earth*, **28**, 441–474, 1980.

Ward, S., On tsunami nucleation: a point source, *J. Geophys. Res.*, **86**, 7895–7900, 1981.

Ward, S., On tsunami nucleation: an instantaneous modulated line source. *Phys. Earth Planet. Inter.*, **27**, 273–285, 1982.

Weidner, D., Rayleigh waves from mid ocean ridge earthquakes: source and path effects, Ph.D. thesis, Harvard College, 253 pp., 1967.

Westbrook, G. *et al.*, Lasser Antilles subduction zone in the vicinity of Barbados, *Nature Phys. Sci.*, **244**, 118–120, 1973.

Yamashita, T. and R. Sato, Generation of tsunami by a fault model, *J. Phys. Earth*, **22**, 415–440, 1974.

Yamashita, T. and R. Sato, Correlation of tsunami and sub-oceanic Rayleigh wave amplitudes. Possibility of the use of Rayleigh wave in tsunami warning system, *J. Phys. Earth*, **24**, 397–416, 1976.

Yoshii, Y. *et al.*, Crustal structure of Tosa deep-sea terrace and Nankani trough (in Japan), in *Island Arc and Ocean*, edited by M. Hoshino and H. Aoki, pp. 93–103, Tokai University Press, Tokyo, 1970.

T. Novikova (e-mail: tatyana@earth.sinica.edu.tw), K.-L. Wen, and B.-S. Huang

Atmos. Chem. Phys., 11, 11913–11917, 2011
www.atmos-chem-phys.net/11/11913/2011/
doi:10.5194/acp-11-11913-2011
© Author(s) 2011. CC Attribution 3.0 License.

Atmospheric Chemistry and Physics

Observations of in-situ generated gravity waves during a stratospheric temperature enhancement (STE) event

A. J. Gerrard[1], **Y. Bhattacharya**[1], **and J. P. Thayer**[2]

[1]Center for Solar-Terrestrial Research, New Jersey Institute of Technology, 323 Martin Luther King Jr. Boulevard, Newark, NJ 07102-1982, USA
[2]Aerospace Engineering Sciences, University of Colorado, 429 UCB, Boulder, CO 80309-0429, USA

Received: 21 February 2011 – Published in Atmos. Chem. Phys. Discuss.: 9 May 2011
Revised: 25 October 2011 – Accepted: 11 November 2011 – Published: 30 November 2011

Abstract. Evidence for in situ generated atmospheric gravity waves associated with a stratospheric temperature enhancement (STE) are presented. The signatures of two sets of gravity waves are observed by molecular-aerosol lidar in conjunction with the early December 2000 STE event above Sondrestrom, Greenland. The first set of gravity waves shows downward phase progression with a vertical wavelength of ∼8 km while the second set shows upward phase progression with a vertical wavelength of ∼9 km. With estimates of the background wind fields from synoptic analyses, the various intrinsic gravity wave parameters of these two wave structures are found. The observed wave features compare well to previous numerical modeling predictions.

1 Introduction

Stratospheric temperature enhancements (STEs) have been experimentally observed, theoretically and numerically modeled, and discussed in the scientific literature for some time, often under differing nomenclature such as: stratospheric baroclinic zones, stratospheric sudden warmings (SSWs), winter stratospheric temperature anomalies, etc. SSW or a major warming is defined by the World Meteorological Organization as an increase in zonal-mean temperature of the polar region such that the net zonal-mean zonal winds become easterly north of 60° N at 10 hPa or below. An STE, such as the one discussed in this paper, leaves the lower stratosphere relatively undisturbed from a global context and is therefore considered distinct from an SSW. Meriwether and Gerrard (2004) have reviewed the phenomenology of STEs and have highlighted the interesting attributes of the structures, including (1) the nature of their development and evolution, (2) how STEs may serve as precursors to sudden stratospheric warmings, and (3) the role of STEs as in situ generators of middle atmospheric gravity waves. The first issue has been addressed by a number of authors, including a companion paper by Thayer and Livingston (2008). This paper addresses the third issue by analyzing a STE event in December 2000, when the height of the stratopause dropped to ∼42 km and temperatures were enhanced by ∼50 K, described in that earlier Thayer and Livingston (2008) paper.

As noted in the review by Meriwether and Gerrard (2004), numerical modeling of STE events by Fairlie et al. (1990) suggest that gravity waves are generated at the location of the STE. More recently, Vadas and Fritts (2001), and Fritts et al. (2006) have discussed strong temperature anomalies and subsequent local adjustment as potential generators of gravity waves. As reproduced from Fairlie et al. (1990) in Fig. 1, upward propagating gravity waves are seen in model runs with horizontal wavelengths of ∼1200 km, vertical wavelengths of ∼8 km, and intrinsic periods of ∼10 h. Though single large-scale wave such as this does not transport large amounts of horizontal momentum vertically, their presence may nonetheless assist smaller-scale gravity waves to move towards convective or dynamical instability due to localized wave-wave interactions (Fritts and Alexander, 2003). Furthermore, it is not clear to what extent the associated localized momentum deposition plays a role in the re-stabilization of the high latitude middle atmosphere system (i.e., the reformation of the polar vortex) and return to geostrophy.

Correspondence to: A. Gerrard
(gerrard@njit.edu)

Published by Copernicus Publications on behalf of the European Geosciences Union.

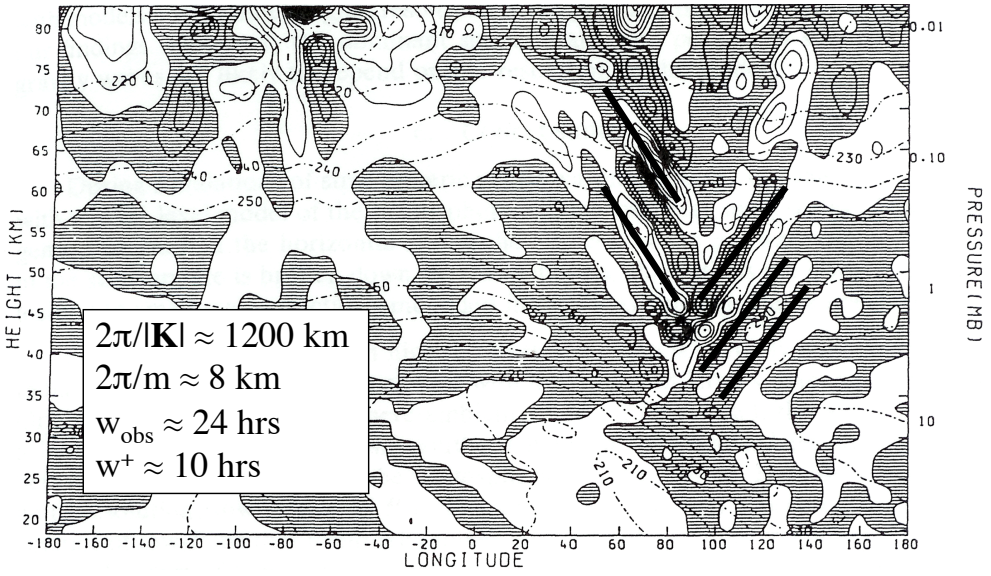


Fig. 1. From Fairlie et al. (1990). Simulated cross-section of temperature (dashed, units of K) and the divergence of the horizontal wind (solid, units of $10^{-6}\,s^{-1}$) at 57.5° N on 25 February 1979. Regions of convergence are shaded. Note that the the original pressure axis of millibars (written as mB) has been retained. The figure has been modified to highlight modeled, upwardly propagating gravity waves to both the west and east of the STE/baroclinic zone. Estimated gravity wave parameters are listed.

In situ observations of STE-associated gravity waves have not been reported (Meriwether and Gerrard, 2004). To our knowledge, this paper presents the first observations of such gravity wave signatures and associated wave structures generated during an STE event. The observations were taken during 11–12 December 2000 above Sondrestrom, Greenland. Synoptic scale conditions prevalent during this STE were presented in Thayer and Livingston (2008); this study deals exclusively with the period of peak STE temperatures over Sondrestrom. In Sect. 2 we present a brief overview of the instrumentation and analysis, in Sect. 3 we discuss the gravity wave characteristics, and in Sect. 4 we discuss these findings.

2 Instrumentation and analysis

Data used in this study originate from the ARCLITE molecular-aerosol lidar system located at Sondrestrom, Greenland. The ARCLITE system is discussed in detail in Thayer et al. (1997) and data from this system consist of range-time resolved temperature profiles of the upper stratosphere and lower mesosphere, relative gravity wave activity in the upper stratosphere, and aerosol volume backscatter coefficients. The data products have been used to explore mesospheric clouds (Thayer et al., 2003), relationship between gravity waves and mesospheric clouds (Gerrard et al., 1998, 2004a,b), thermal and gravity wave climatologies (Gerrard et al., 2000), and the synoptic development of the polar vortex (Gerrard et al., 2002). In Fig. 2 we present interpolated temperature profiles taken by the Sondrestrom lidar system in early December 2000 (Thayer and Livingston, 2008). Measurements were taken on the days indicated by black markers and interpolated across the entire duration of the STE event. Uncertainty of the lidar-deduced temperatures is less than 3 K below 70 km. Of particular note is the lidar data from 11–12 December [day of year 346–347] , whose time evolution is also portrayed in Fig. 3. The temperature data here was low pass filtered in the vertical with a cutoff at 7 km. Uncertainty on the measured temperatures is less than 5 K below 74 km. In these data the STE is clearly evident as a narrow ridge of temperatures exceeding 320 K; over 40 K above the climatological value (Gerrard et al., 2000).

From the ∼25 h long set of molecular – aerosol lidar data spanning this time period, it is possible to obtain time-resolved relative atmospheric density perturbations between 30 km and 70 km as outlined in Gerrard et al. (1998, 2004b). The density perturbation data (relative to an estimated atmospheric background), available in 5 min by 192 m vertical resolution bins, have been low pass filtered with cutoffs at 2 km in the vertical and 30 min in the temporal. These relative density perturbations are shown in Fig. 3.

In a perturbation analysis of the base-state parameters of an ideal gas, it can be shown that regions of high density perturbations correspond to regions of lower temperatures. This is the case on inspection of Figs. 2 and 3, where high density perturbations are 180° out of phase with the corresponding temperatures. As such, Figs. 2 and 3 essentially show the same thing, but gravity wave signatures are more readily extracted and analyzed in the relative density perturbations. From these figures, one sees that the STE peak temperatures

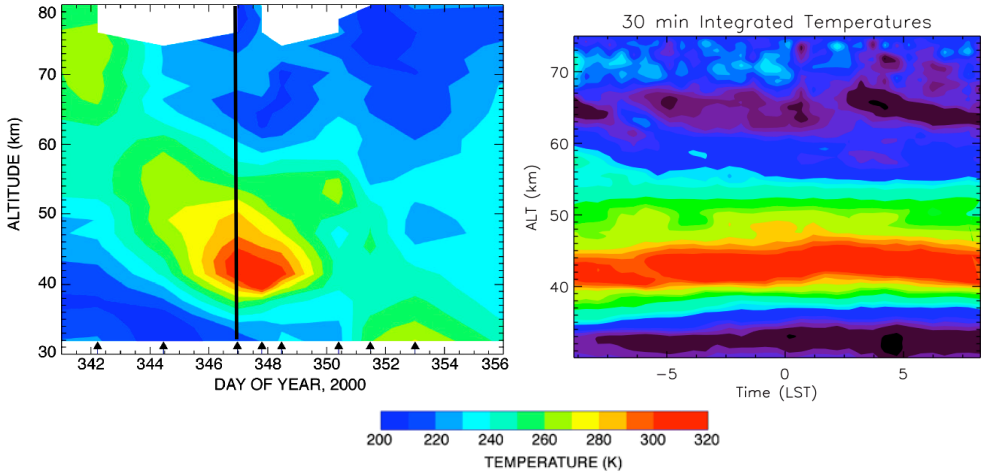


Fig. 2. From Thayer and Livingston (2008). Evolution of middle atmospheric temperatures over Sondrestrom, Greenland in December 2000 as measured by the molecular/aerosol lidar system. The vertical black line represents the day under study herein. (right) Range-time temperature evolution over Sondrestrom over the day near the estimated peak STE represented by the vertical black line (i.e., day of year 346–347, night of 11–12 December 2000, leap year). Time is local solar time (LST) with 00:00 representing midnight of 11 December.

at ∼42 km fluctuates by a few kilometers in height and in peak intensity through the course of the observations. This is presumably due to the time-varying nature of the baroclinic zone and other atmospheric structures, like gravity wave activity.

3 Gravity wave structures

Inspection of Fig. 3 shows at least two regions of relatively coherent wave-phase progression. These two areas of wave-like structure are highlighted in Fig. 3, where the waves in the upper left corner (centered at ∼04:00 LST and ∼65 km, and indicated by red streaks sketched by inspection of the data) are referred to as "set 1", and the waves on the right (centered at ∼09:00] LST and ∼50 km, and represented by blue streaks) are referred to as "set 2." No wave structures were apparent in the data below ∼45 km.

Addressing wave set 1, the wave structure has a downward phase progression in time, a vertical wavelength of ∼8 km and an observed period of ∼5 h. This type of structure is common over Sondrestrom (for e.g., seen in data presented in Gerrard et al., 1998 and Gerrard et al., 2000), and has the characteristic of an atmospheric gravity wave propagating upwards in height (i.e., the phase progression is downward, which is indicative of upward energy progression). Strictly speaking, downward phase propagation results in upward propagation of energy in a fluid at rest or in a frame moving with the background wind. This would hold approximately true if the polar vortex mass did not move significantly during the STE event. Such an assumption is not unreasonable; for instance, analysis presented in Bhattacharya and Gerrard (2010) using stratospheric assimilated data from

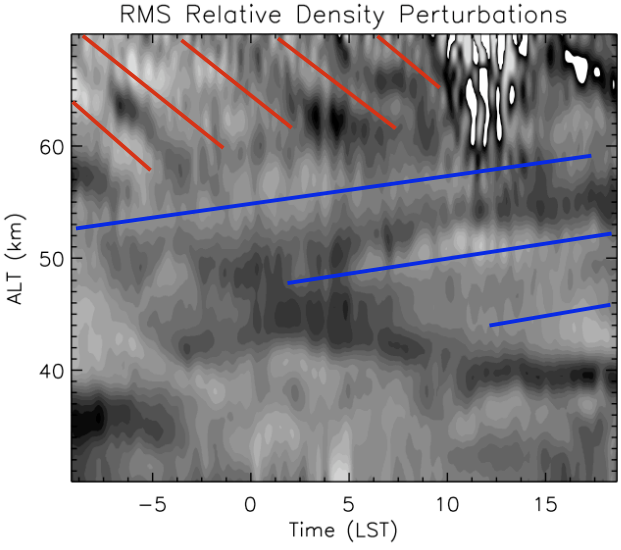


Fig. 3. Relative upper stratospheric and lower mesospheric density perturbations observed by the molecular/aerosol lidar system at Sondrestrom for the same range-time of Fig. 2. The relative density perturbations are scaled from −10 % to 10 % in 20 grey-scale intervals. Black represents the negative perturbations, while white represents the positive perturbations. Solar noise is present in the 60- to 70 km region of the figures around 12:00 LST which prohibited robust calculation of the temperatures in Fig. 2. Phase peaks of upward and downward phase lines corresponding to the two gravity wave sets are determined by inspection and highlighted.

United Kingdom Met Office (UKMO) shows that geopotential height at 10 hPa remained relatively constant (∼36 km) over Resolute Bay (75° N, 95° W) during the STE period

reported here. Assimilated data from National Center for Environmental Prediction (NCEP) gives the average horizontal wind velocity $\sim$50 m s^{-1} at $\sim$65 km.

If the wind velocity vector $\boldsymbol{U}$ is aligned with the gravity wave horizontal wavenumber vector $\boldsymbol{U}$ (in general, gravity waves with such aspect ratios propagate in the direction of wind as modeled in Gerrard et al., 2004a), and using the buoyancy frequency N deduced from the nightly averaged temperature profile, the horizontal wavelength and intrinsic period ω^+ of the wave can be estimated from the gravity wave dispersion relationship in the form:

$$(\omega^+)^2 = (\omega - \boldsymbol{U}\boldsymbol{k})^2 = \frac{N^2k^2 + f^2m^2}{k^2 + m^2} \quad (1)$$

where, aside from the variables already described earlier, ω is the ground-based frequency, f is the Coriolis parameter, and m is the gravity wave vertical wavenumber. The group velocity of the gravity wave can also be determined from this equation by taking the derivative of w with respect to the horizontal or vertical wavenumber to yield the horizontal group velocity or vertical group velocity, respectively.

Numerical solution of the dispersion relation using a Laguerre method yields two real roots and two imaginary roots. The imaginary roots are discarded as they represent evanescent structures. The first real root results in a gravity wave with a horizontal wavelength of $\sim$1700 km and an intrinsic wave period of $\sim$10.5 h, with positive horizontal and vertical group velocities indicating an upward propagating gravity wave moving in the direction of the background wind. The second root yields horizontal wavelength of $\sim$400 km and an intrinsic wave period of $\sim$4 h, with a positive horizontal group velocity and negative vertical group velocity indicating a downward propagating gravity wave moving in the direction of the background wind.

Addressing wave set 2, the wave structure has an upward phase progression in time, a vertical wavelength of $\sim$9 km, and an observed period of $\sim$26 h. Estimating that the background wind velocity from NCEP analyses to be +70 m s^{-1}, coupled with the assumption of wave vector k approximately aligned with U, the first real root yields a gravity wave with a horizontal wavelength of $\sim$1870 km and an intrinsic wave period of $\sim$10.4 h, with positive horizontal and vertical group velocities indicating an upwards propagating gravity wave moving in the direction of the background wind. The second real root yields a wave with a horizontal wavelength of $\sim$6200 km and an intrinsic wave period of $\sim$12.6 h, with a positive horizontal group velocity and negative vertical group velocity indicating a downward propagating gravity wave moving in the direction of the background wind.

It can be shown that uncertainties of $\pm$1 km in the vertical wavelengths, $\pm$1 h in the observed periods, and $\pm$10 m s^{-1} do not dramatically alter the estimated horizontal wavelengths and intrinsic periods.

4 Conclusions

Assuming that the first real root of set 1 and the first real root of set 2 are the physically meaningful roots (i.e., both indicating upwards gravity waves propagating from the STE region which fits the physical model), the observed gravity wave characteristics compare very well with model output of Fairlie et al. (1990). Furthermore, there was no indication of any wave structures below the STE temperature peak (i.e., between 30 km and 45 km) to indicate a gravity wave passing through the STE. As such, we conclude that this is strong evidence for an observation of upward propagating gravity waves generated in situ by a STE event.

It should be noted that higher than normal total ozone content were reported in the Arctic vortex (until 63° N latitude for which ozone data was available) for the 2000–2001 winter; for December 2000, total ozone was higher by $\sim$20 ,% compared to early 1980s (NCEP-NOAA Winter Bulletin, 2000–2001). An unusual and extensive $\sim$3 km thick aerosol cloud layer, spanning almost the entire polar vortex around $\sim$38 km altitude also formed in the first two weeks of December 2000, as reported by Gerding et al. (2003) observed from four different Arctic stations. Bhattacharya et al. (2004) have also reported increased gravity wave activity in the mesopause (hydroxyl airglow layer) polar vortex immediately following the STE events described in this paper. These observations in the current context increase the likelihood that the temperature enhancement, and the associated GW generation had their origins within a relatively shallow layer in the stratosphere. The long wavelength observed is also suggestive of a horizontally extensive forcing and a slower adjustment process (Fritts et al., 2006), unlike a vertically extended source typically found around the edges of a split vortex.

Based on this one observation, we are currently searching through the NCEP geopotential data archive for similar polar vortex configurations seen during this event. However, the combination of observing conditions, namely (a) the disruption of the polar vortex over Sondrestrom and actual occurrence of an STE when (b) there are clear sky conditions in which (c) the ARCLITE lidar system was collecting data is quite rare. Utilization of other lidar data archives, e.g., from the ALOMAR lidar system, the Poker Flat Lidar system, and the Eureka lidar system, would likely increase the probability of observing gravity waves associated with a STE.

In conclusion, in situ forcing of middle atmospheric gravity waves are of interest to the upper and lower atmospheric communities because, to date, little is known about the contribution that such waves may play in the global circulation. In this study we have demonstrated the likely origin of such in situ forcing associated with an observed STE. Further studies are necessary to better understand the role of such waves in the high latitude environment, how frequently such waves are generated, and the synoptic impacts on the surrounding atmospheric environment.

Acknowledgements. The authors of this paper would like to thank site members of the Sondrestrom research facility and J. Livingston (SRI) for their support. AG and YB were supported in part by the National Science Foundation through ATM-0735452 and NSF-ATM-0457277.

Edited by: F.-J. Lübken

References

Bhattacharya, Y. and Gerrard, A. J.: Wintertime mesopause region vertical winds from Resolute Bay, J. Geophys. Res, 115, D00N07, doi:10.1029/2010JD014113, 2010.

Bhattacharya, Y., Shepherd, G. G., and Brown, S.: Variability of atmospheric winds and waves in the Arctic polar mesosphere during a stratospheric sudden warming, Geophys. Res. Lett., 31, L23101, doi:10.1029/2004GL020389, 2004.

Fairlie, T. D. A., Fisher, M. and O'Neill, A.: The development of narrow baroclinic zones and other small-scale structure during simulated major warmings, Q. J. Roy. Meteorol. Soc., 116, 287–315, 1990.

Fritts, D. C. and Alexander M. J.: Gravity wave dynamics and effects in the middle atmosphere, Rev. Geophys., 41, 1003, doi:10.1029/2001RG000106, 2003.

Fritts, D. C., Vadas, S. L., Wan, K., and Werne, J. A.: Mean and variable forcing of the middle atmosphere by gravity waves, J. Atmos. Sol.-Terr. Phys., 68, 247–265, 2006.

Gerding, M., Baumgarten, G., Blum, U., Thayer, J. P., Fricke, K.-H., Neuber, R., and Fiedler, J.: Observation of an unusual mid-stratospheric aerosol layer in the Arctic: possible sources and implications for polar vortex dynamics, Ann. Geophys., 21, 1057–1069, doi:10.5194/angeo-21-1057-2003, 2003.

Gerrard, A. J., Kane, T. J., and Thayer, J. P: Noctilucent clouds and wave dynamics: Observations at Sondrestrom, Greenland, Geophys. Res. Lett., 25, 2817–2820, 1998.

Gerrard, A. J., Kane, T. J., and Thayer, J. P.: Year-round temperature and wave measurements of the arctic middle atmosphere for 1995-1998, in Geophysical Monograph Series 123: Atmospheric Science Across the Stratopause, edited by D. E. Siskind, S. D. Eckermann, and M. E. Summers, American Geophysical Union, Washington, DC, USA, 83, 488, 2000.

Gerrard, A. J., Kane, T. J., Thayer, J. P., Duck, T. J., Whiteway, J. A., and Fiedler, J.: Synoptic-scale study of the arctic polar vortex's influence on the middle atmosphere: I. observations, J. Geophys. Res., 107, 4276, doi:10.1029/2001JD000681, 2002.

Gerrard, A. J., Kane, T. J., Eckermann, S. D, and Thayer, J. P.: Gravity waves and mesospheric clouds in the summer middle atmosphere: A comparison of lidar measurements and ray modeling of gravity waves over Sondrestrom, Greenland, J. Geophys. Res., 109, D10103, doi:10.1029/2002JD002783, 2004a.

Gerrard, A. J., Kane, T. J., Thayer, J. P., and Eckermann, S. D.: Concerning the upper stratospheric gravity wave and mesospheric cloud relationship over Sondrestrom, Greenland, J. Atmos. Sol.-Terr. Phys., 66, 229–240, 2004b.

Meriwether, J. W. and Gerrard, A: Mesosphere inversion layers and stratosphere temperature enhancements, Rev. Geophys., 42, RG3003, doi:10.1029/2003RG000133, 2004.

Thayer, J. and Livingston, J: Observations of wintertime arctic mesospheric cooling associated with stratospheric baroclinic zones, Geophys. Res. Lett., 35, L10109, doi:10.1029/2008GL034955, 2008.

Thayer, J. P., Nielsen, N. B., Warren, R., Heinselman, C. J., and Sohn, J.: Rayleigh lidar system for middle atmosphere research in the arctic, Opt. Eng., 36, 2045–2061, 1997.

Thayer, J. P., Rapp, M., Gerrard, A. J., Gudmundsson, E., and Kane, T. J.: Gravity-wave influences on arctic mesospheric clouds as determined by a rayleigh lidar at Sondrestrom, Greenland, J. Geophys. Res., 108, 8449, doi:10.1029/2002JD002363, 2003.

Vadas, S.-L. and Fritts, D.-C.: Gravity wave radiation and mean responses to local body forces in the atmosphere, J. Atmos. Sci., 58, 2249–2279, 2001.

The Utilization of High-Frequency Gravitational Waves for Global Communications

Robert M L BAKER, Jr. and Bonnie S. BAKER
Transportation Sciences Corporation and GravWave® LLC
8123 Tuscany Avenue, Playa del Rey, CA 90293

ABSTRACT

For over 1000 years electromagnetic radiation has been utilized for long-distance communication. heliographs, telegraphs, telephones and radio have all served our previous communication needs. Nevertheless, electromagnetic radiation has one major difficulty: it is easily absorbed. In this paper we consider a totally different radiation, a radiation that is not easily absorbed: gravitational radiation. Such radiation, like gravity itself, is not absorbed by earth, water or any material substance. In particular we discuss herein means to generate and detect high-frequency gravitational waves or HFGWs, and how they can be utilized for communication. There are two barriers to their practical utilization: they are extremely difficult to generate (a large power required to generate very weak GWs) and it is extremely difficult to detect weak GWs. We intend to demonstrate theoretically in this paper their phase-coherent generation utilizing an array of in-phase microelectro-mechanical systems or MEMS resonator elements in which the HFGW flux is proportional to the square of the number of elements. This process solves the transmitter difficulty. Three HFGW detectors have previously been built; but their sensitivity is insufficient for meaningful HFGW reception; greater sensitivity is necessary. A new Li-Baker HFGW detector, discussed herein, is based upon a different measurement technique than the other detectors and is predicted to achieve a sensitivity to satisfy HFGW communication needs.

Keywords: Gravitational waves, communications, Li-Baker detector, microelectromechanical systems, high-frequency gravitational waves.

1. INTRODUCTION

Since the dawn of civilization electromagnetic radiation has been utilized for long-distance communication: heliographs, telegraphs, telephones and radio have all served our previous communication needs. Nevertheless, electromagnetic radiation has one major drawback: it is easily absorbed. In this paper we consider a totally different radiation, a radiation that is not easily absorbed: gravitational radiation. Such radiation, like gravity itself, is not absorbed by earth, water or any material substance. In particular we discuss herein a means to generate and detect high-frequency gravitational waves or HFGWs and how they can be utilized for communication. HFGWs are defined as GWs having frequencies in excess of 100 kHz (Douglas and Braginsky [1]) and long-wavelength GW detectors such as LIGO, Virgo and GEO600 cannot sense HFGWs [2]. Global communications by means of HFGWs would be the ultimate wireless system. HFGW communication would greatly reduce communications costs since it would not require the following:

- No satellite transponders
- No microwave relays
- No underwater cables
- No coaxial cables

Since the Nobel Prize winning observations of Hulse and Taylor in the 1970s no one has doubted the existence of gravitational waves. There are two barriers to their practical utilization: they are extremely difficult to generate (a large power required to generate very weak GWs) and it is extremely difficult to detect weak GWs. In the past several decades hundreds of peer-reviewed journal articles have addressed these issues, for example Beckwith [3] and Grishchuk [4]. We intend to demonstrate theoretically in this paper that their generation utilizing superradiance (Scully and Svidzinsky. [5]), which involves a linear double-helix array of in-phase micro-electromechanical systems (MEMS) resonator elements, in which the HFGW flux is proportional to the square of the number of elements, solves the HFGW generation or transmitter difficulty. The use of a new, but well documented in peer-reviewed literature, effect discovered by Fangyu Li (Chongqing University, China, [6]) solves the detection difficulty. This Li-effect is the basis for the very sensitive Li-Baker HFGW detector, designed by Robert Baker and developed jointly by United States and Chinese HFGW research teams. As documented in peer-reviewed literature [7, 8, 9, 10] such a detector has sensitivity more than sufficient to receive the transmitted HFGW signal at a significant distance from the transmitter. Dehnen in Germany showed in an article [11] that HFGWs could be generated in the laboratory, using General Relativity, through the use of crystal oscillators. His work is the basis for an efficient HFGW generator or transmitter. The critical element in Dehnen's HFGW generator or transmitter had been the large size and power requirements of his crystal oscillators. This difficulty is removed through the use of modern MEMS technology. There have been other challenges to HFGW communications based upon the mistaken belief that GW generators or transmitters can only be designed using spinning rods or the effect described by Gertsenshtein in 1962 [12] and analyzed by Eardley in 2008 in the JASON report [13]. Both of these methods for generating GWs are unsatisfactory and produce negligible GW power.

2. HFGW GENERATORS (Transmitters)

There exist several sources for HFGWs or means for their generation. The first generation means is the same for gravitational waves (GWs) of all frequencies and is based upon the quadrupole equation first derived by Einstein in 1918.[14] A formulation of the quadrupole that is easily related to the orbital motion of binary stars or black holes, rotating rods, laboratory HFGW generation, etc. is based upon the jerk or shake of mass (time rate of change of acceleration), such as the change in centrifugal force vector with time; for example as masses move around each other on a circular orbit. Figure 1 describes that situation. Recognize, however, that change in force Δf need NOT be a gravitational force (see Einstein, 1918 [14]; Infeld quoted by Weber 1964 [15] p. 97; Grishchuk [16]). Electromagnetic forces are more than 10^{35} larger than gravitational forces and should be employed in laboratory GW generation. As Weber ([15] p. 97) points out: "The non-gravitational forces play a decisive role in methods for detection and generation of gravitational waves ..." The quadrupole equation is also termed "quadrupole formalism" and holds in weak gravitational fields (but well over 100 g's), for speeds of the generator "components" less than the speed of light and for the distance between two masses r less than the GW

wavelength. Certainly there would be GW generated for r greater than the GW wavelength, but the quadrupole "formalism" or equation might not apply exactly. For very small time change Δt the GW wavelength, $\lambda_{GW} = c\,\Delta t$ (where $c \sim 3\times10^8$ m s^{-1}, the speed of light) is very small and the GW frequency ν_{GW} is high. The concept is to produce two equal and opposite jerks or Δf 's at two masses, such as MEMS, a distance $2r$ apart. This situation is completely analogous to binary stars on orbit as shown in Figs. 1 and 2**.**

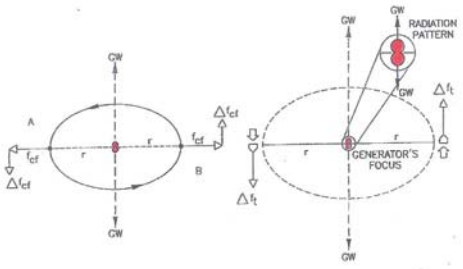


Figure 1. Change in Centrifugal Force of Orbiting Masses, Δf_{cf}, Creates Radiation.

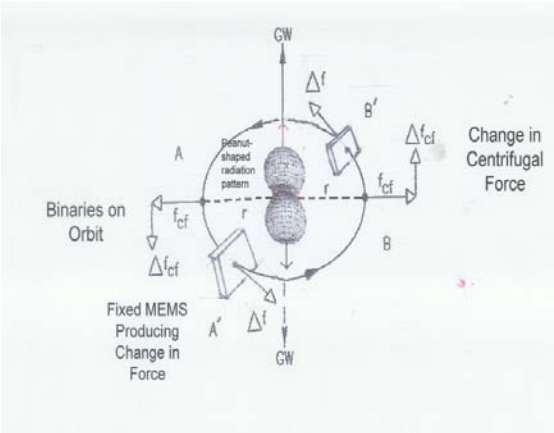


Figure 2. Radiation Pattern Calculated by Landau and Lifshitz [17] Section 110 Page 356.

Next we consider an array of GW sources. Consider a stack of orbit planes, each one involving a pair of masses circling each other on opposite sides of a circular orbit as in Fig. 3. Let the planes be stacked one light hour apart (that is, $60\times60\times3\times10^8 = 1.08\times10^{12}$ meters apart) and each orbit exactly on top of another (coaxial circles). According to Landau and Lifshitz [17] on each plane a GW will be generated that radiates from the center of each circular orbit. The details of that generation process are that as the masses orbit a radiation pattern is generated. In simplified terms (from the equations shown on page 356 of Landau and Lifshitz [17]) an elliptically shaped polarized arc of radiation is formed on each side of the orbit plane (mirror images). As the two masses orbit each other 180^0 the arcs sweep out figures of revolution. Together these figures of revolution become shaped like a peanut as shown in Fig. 2.

The general concept of the present HFGW generator is to utilize an array of force-producing elements arranged in pairs in a cylindrical formation such as a double helix as in Fig. 4. This is analogous to the binary-star arrays of Fig. 3 in which an imaginary cylinder could be formed or constructed from the collection of orbits. As a wavefront of energizing radiation proceeds along the cylindrical axis of symmetry of such a double-helix array, shown in Fig. 4 the force-producing element pairs (such as pairs of film-bulk acoustic resonators or FBARs) are energized simultaneously and jerk, that is they exhibit a third time

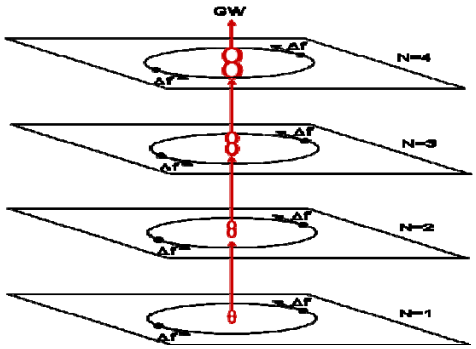


Figure 3. GW Flux Growth Analogous to Stack of *N* Orbital Planes.

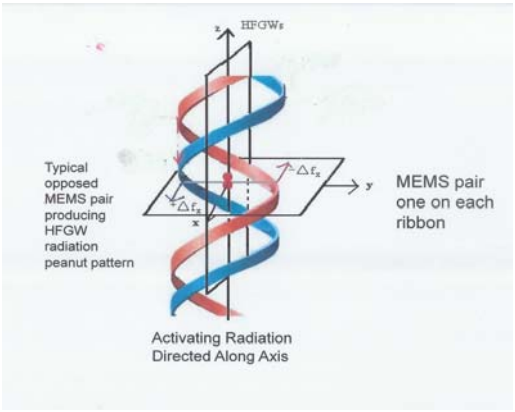


Figure 4. Double-Helix HFGW Generator Array (Patent Pending)

derivative of motion and the flux (W m^{-2}) thereby increased. Utilizing General Relativity, Dehnen and Romero-Borja [11] computed a superradiance build up of "… needle-like radiation …" HFGWs beam emanating from a closely packed but very long linear array of crystal oscillators. Their oscillators were essentially two vibrating masses a distance b apart whereas a pair of vibrating FBAR masses is a distance $2r$ apart as shown in Fig. 5, but operates in an analogous fashion as piezoelectric crystals.

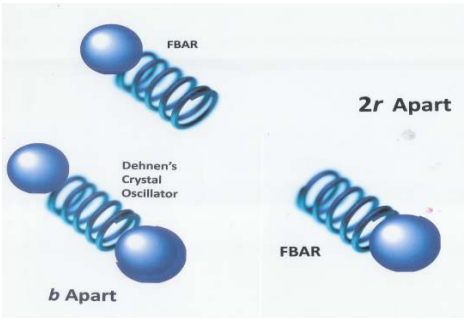


Figure 5. Comparison of Dehnen and Romero-Borja [11] crystal oscillator and FBAR-pair system.

Superradiance also occurs when emitting sources such as atoms "…are close together compared to the wavelength of the radiation …" (Scully and Svidzinsky [5] p.1510). Note that it is not necessary to have the FBAR elements perfectly aligned (that is, the FBARs *exactly* across from each other) since it is only necessary that the energizing wave front (from Magnetrons in the case of the MEMS or FBARs as in Baker, Woods and Li [18]) reaches a couple of nearly opposite FBARs at the same time so that a coherent radiation source or focus is produced

between the two FBARs. The energizing transmitters, such as Magnetrons, can be placed along the helixes' array axes between separate segments of the array or, more efficiently, at the base of the double helixes so that a superradiance force change, Δf, produced by energizing one off-the-shelf FBAR is 2 N microwave beam is projected up the axis of the helixes. The according to Woods and Baker [19], so that the power is given by the equation derived in Baker [20]:

$$P = 1.76\times10^{-52}\,(2r\,\Delta f/\Delta t)^2 \quad \text{W.} \qquad (1)$$

Let the activating radiation for the FBARs be conventional Magnetrons as employed in one-thousand watt microwave ovens. The frequency would be ν_{EM} = 2.5 GHz (thus Δt = 4×10^{-10} s and λ_{EM} = 12 cm**).** The HFGW frequency is twice that of the activating EM radiation or ν_{GW5} = 5 GHz**.** For Eq (1) the calculation of the combined Δf of all the pulsating MEMS or FBARs requires more consideration. We will set the length of a double-helix array cylinder as 20 m, but recognize that it can be separated into segments along the same axis with energizing transmitters, e.g., Magnetrons installed on the cylinder axis between the segments. The transmitters could also be phase coherent and arranged in a line along the double-helix axis at its base. If, for example, there were 1000 one-kilowatt Magnetrons feeding in on one hundred 12-cm, (λ_{EM}, wide levels) and each of their beams covered a 10-cm radius circle, then the energizing radiation flux would be 3.2×10^4 W m^{-2}. According to superradiance there would result a needle-like microwave radiation directed along the axis of the double helixes amounting to 32 gigawatts per square meter. In order to create a perfectly planner wave front, with no irregularities, the cylindrically symmetric MEMS array could be contained in a wave guide or possibly a very wide coaxial "cable," surrounded by a robust one megawatt heat sink. To increase instantaneous power to the array, bursts of gigawatt power, for example, every millisecond could be employed that would maintain a megawatt average power input.

The walls of the cylindrical array are taken to be 30-cm thick. Thus the volume of the twenty meter long array is $\pi(r_1^2 - r_2^2)$ ×20 m^3, where r_1 is the outside radius = 0.35 m and r_2 is the inside radius = 0.05 m. Thus the volume is 7.5 m^3. The FBAR is a mechanical (acoustic) resonator consisting of a vibrating membrane (typically about $100\times100\mu m^2$ in plan form, and about 1μm thickness), fabricated using well-established integrated circuit (IC) micro fabrication technology. A typical off-the-shelf FBAR as shown schematically in Fig. 6, usually has overall dimensions 500 μm by 500 μm by approximately 100 μm thick. For our purposes, in which a high number density is important, we will trim the FBARs to a minimum size. In order to account for fabrication margins we will take the dimensions as 110 μm by 110 μm by 20 μm for an FBAR volume of $2.42\times10^{-13}m^3$.

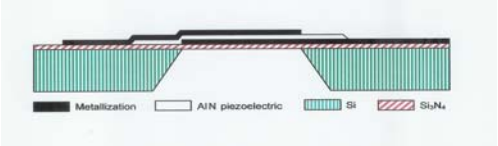


Figure 6. Basic FBAR Construction (cross-section side view, not to scale).

Thus the total number of FBARs in the double-helix cylindrical array is 2.85×10^{12} and the number of pairs is half of that. There would be $N = 1.425\times10^{12}$ FBAR pairs in the double-helix cylindrical array. Since each FBAR exhibits a jerking force of 2 N the combined Δf of all the jerking FBAR pairs is 2.85×10^{12} N. if the jerking pairs (or "orbits") moved in concert**.** From Eq. (1) the total power produced by the double-helix array is $P = 1.76\times10^{=52}(0.2\times2.85\times10^{12}/4\times10^{-10})^2 = 3.57\times10^{-10}$ W. But due to the N levels, each one of which represents an individual GW focus, there exists a "Superradiance" condition in which the HFGW beam becomes very narrow as shown schematically in Fig. B of [5]. Thus the HFGW flux, in W m^{-2}, becomes much larger at the cap of the radiation pattern. According to the analyses of Baker and Black [21] the area of the half-power cap is given by

$$A_{cap} = A_{1/2(N=1)} / N \quad m^2 \qquad (2)$$

A more conservative approach would be that there are N individual GW power sources each with a Δf = 2 N. Thus from Eq. (1), with $2r_{rms} = 2\sqrt{[(r_1^2 + r_2^2)/2]} = 0.5$ m, the total power produced by the double-helix array is $P = 1.55\times10^{13} \times1.76\times10^{-52}(0.5\times2/4\times10^{-10})^2 = 1.69\times10^{-20}$ W. But due to the N levels, each one of which represents an individual GW focus, there exists a "Superradiance" condition in which the HFGW beam becomes very narrow as shown schematically in Fig. B of Scully and Svidzinsky [5]. Thus the HFGW flux, in W m^{-2}, becomes much larger at the cap of the radiation pattern. According to the analyses of Baker and Black [21] the area of the half-power cap is proportional to $1/N$ and the GW flux is:

$$S(1) = (P/4)(1.71/N) = (1.69\times10^{-20}/4)/(1.71/1.55\times10^{13}) = 3.8\times10^{8}\ \text{W m}^{-2} \quad . \qquad (3)$$

From Baker, et al. [22], Eq. (6A) of the Appendix, the amplitude of the dimensionless strain in the fabric of spacetime is

$$A = 1.28\times10^{-18}\sqrt{S}/\nu_{GW} \quad \text{m/m} \quad . \qquad (4)$$

So that at a one-meter distance $A = 5\times10^{-32}$ m/m. If the FBARs in all of the helix levels are not activated as individual pairs, then the situation changes. For example, let all of the FBARs in a 6-cm wide level (½ λ_{EM}) be energized in concert. The number of levels would be reduced to N = 20 m/0.06 m = 333. But, because the FBAR-pairs in each level act together, Δf = (2 N)$(1.55\times10^{13}$ / 333). Thus the changes in Eq. (1) cancel out and there is no change in HFGW flux.

The HFGW beam is very narrow. From Eq. (4b) of [21] for $N = 1.55\times10^{13}$ it would be $\sin^{-1}(0.737)/\sqrt{1.55\times10^{13}} = 1.87\times10^{-7}$ radians. For N = 333 the angle is 0.0022 radians. This is still narrow, but the double helix configuration certainly reduces the width of the HFGW beam. Additionally multiple HFGW carrier frequencies can be used with modulation schemes e.g., pulse carrier phase shift key, so the signal is very difficult to intercept, and is therefore useful as a low-probability-of-intercept (LPI) signal, even with widespread adoption of the HFGW technology.

From Woods, et al. [23] the current estimated sensitivity of the Chinese Li-Baker HFGW Detector is $A = 1.0\times10^{-30}$ m/m to 1.0×10^{-32} m/m with a signal to noise ratio of over 1500 (Woods, et al [23] p. 511) or if we were at a $1.3x10^7$ m (diameter of Earth) distance, then $S = 1.33\times10^{-20}$ Wm^{-2}. and the amplitude A of the HFGW is given by $A = 3.8\times10^{-39}$ m/m. Although the best theoretical sensitivity of the Li-Baker HFGW detector is on the order of 10^{-32} m/m, its sensitivity can be increased dramatically (Li and Baker, [9] by introducing superconductor resonance chambers into the interaction volume (which also improves the Standard Quantum Limit; Stephenson [24]) and two others between the interaction volume and the two microwave receivers. Together they provide an increase in

sensitivity of five orders of magnitude and result in a theoretical sensitivity of the Li-Baker detector to HFGWs having amplitudes of 10^{-37} m/m. There also could be a HFGW superconductor lens, as described by Woods [25], which could concentrate very high frequency gravitational waves at the detector or receiver. Thus with Chinese Li-Baker HFGW detector program successful and the Wood's lens practical, **the Li-Baker detector will exhibit sufficient sensitivity to receive the generated HFGW signal globally.**

3. HFGW DETECTORS (Receivers)

Operational HFGW Receivers

In the past few years HFGW detectors, as exhibited in Figs. 7**,** 8 and 9 have been fabricated at *Birmingham University*, England, *INFN Genoa,* Italy and in Japan. These types of detectors may be promising for the detection of the HFGWs in the GHz band (MHz band for the Japanese) in the future, but currently, their sensitivities are orders of magnitude less than what is required. Such a detection capability is to be expected, however, utilizing the Li-Baker detector. Based upon the theory of Li, Tang and Zhao [6] termed the Li-effect, the detector was proposed by Baker during the period 1999-2000, a patent for it was filed in P. R. China in 2001, subsequently granted in 2007 [26.] (http://www.gravwave.com/docs/Chinese%20Detector%20Patent%2020081027.pdf).

Preliminary details were published later by Baker, Stephenson and Li [22]. This detector was conceived to be sensitive to relic HFGWs (or high-frequency relic gravitational waves, termed HFRGWs) having amplitudes as small as 10^{-32} to 10^{-30}, but using resonance chambers to 10^{-37} or possibly smaller [9].

The Birmingham HFGW detector measures changes in the polarization state of a microwave beam (indicating the presence of a GW) moving in a waveguide about one meter across as shown in Fig. 7. (Please see Cruise [27]; and Cruise and Ingley [28].) It is expected to be sensitive to HFGWs having spacetime strains whose amplitudes are $A \sim 2 \times 10^{-13}$.

Figure 7 *Birmingham University* **HFGW Detector**

The *INFN Genoa* HFGW resonant antenna consists of two coupled, superconducting, spherical, harmonic oscillators a few centimeters in diameter. Please see Fig. 8. The oscillators are designed to have (when uncoupled) almost equal resonant frequencies. In theory the system is expected to have a sensitivity to HFGWs with size of about $A \sim 2\times10^{-17}$ with an expectation to reach a sensitivity of $\sim 2\times10^{-20}$. (Bernard, Gemme, Parodi, and Picasso [29]); Chincarini and Gemme [30]). As of this date, however, there is no further development of the *INFN Genoa* HFGW detector.

Figure 8 . *INFN Genoa* **HFGW Detector**

The Kawamura 100 MHz HFGW detector has been built by the *Astronomical Observatory of Japan.* It consists of two synchronous interferometers exhibiting an arms length of 75 cm. Please see Fig. 9. Its sensitivity is now about $A \approx 10^{-16}$ (Nishizawa et al., [31]). According to Cruise [32]) of *Birmingham University* its frequency is limited to 100 MHz and at higher frequencies its sensitivity diminishes.

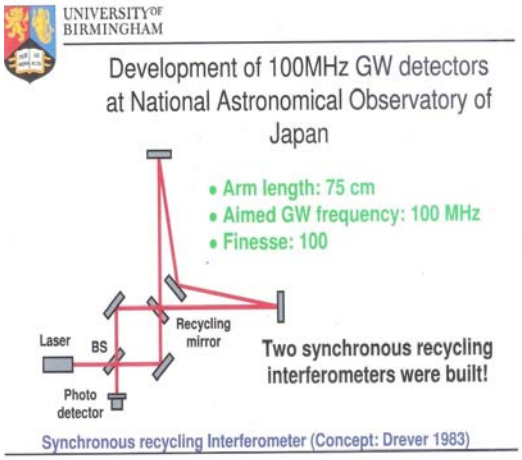


Figure 9. The National Astronomical Observatory of Japan 100MHz Detector. Cruise [31].

Concept (Li-Effect)

The Li-Effect or Li-Theory was first published in 1992 [6]. Subsequently the "Li Effect" was validated by secveral journal articles, independently peer reviewed by scientists well versed in General Relativity, [7, 8, 9, 10] including capstone paper, Li, et al [32]). The reader is encouraged to review the key results and formulas found in Li et al., [10] and the detailed discussion of the coupling among HFGWs, a magnetic field and a microwave beam found in Li et al. [10]. The Li-Effect is *very different* from the classical *(inverse)* Gertsenshtein- Effect. With the Li-Effect, a gravitational wave transfers energy to a separately generated electromagnetic (EM) wave in the presence of a static magnetic field. That EM wave has the same frequency as the GW and moves in the same direction. This is the "*synchro-resonance condition***,**" in which the EM and GW waves are synchronized and is **unlike the Gertsenshtein-Effect**.[12] The result of the intersection of the parallel and superimposed EM and GW beams, according to the Li-Effect, is *new EM photons moving off in a direction perpendicular to the*

beams and the magnetic field directions. These photons signal the presence of HFGWs and are termed a "perturbative photon flux" or PPF. Thus, these new photons occupy a separate region of space (see Fig. 10) that can be made essentially noise-free and the synchro-resonance EM beam itself (in this case a Gaussian beam) is not sensed there, so it does not interfere with detection of the photons. The existence of the transverse movement of new EM photons is a **fundamental physical requirement**; otherwise the EM fields will not satisfy the Helmholtz equation, the electrodynamics equation in curved spacetime, the non-divergence condition in free space, the boundary and will violate the laws of energy and total radiation power flux conservation. In this connection it should be recognized that *unlike the Gertsenshtein effect,* the Li-effect produces a *first-order* perturbative photon flux (PPF), proportional to the amplitude of the gravitational wave . A not A^2. In the case of the Gertsenshtein-Effect such photons are a second-order effect and according to Eq. (7) of Li, et al. [33] the number of EM photons are "…proportional to the amplitude squared of the relic HFGWs, A^2," … and that it would be necessary to accumulate such EM photons for at least 1.4×10^{16} seconds in order to achieve relic HFGW detection (Li et al., [33]) utilizing the Gertsenshtein-Effect. In the case of the Li theory the number of EM photons is proportional to the amplitude of the relic HFGWs, $A \approx 10^{-30}$, not the square, so that it would be necessary to accumulate such EM photons for less than about 1000 seconds in order to achieve relic HFGW detection (Li et al., [10]). The JASON report (Eardley, [13]) confuses the two effects and erroneously suggests that the Li-Baker HFGW Detector utilizes the inverse Gertsenshtein effect. It does not and does have a theoretical sensitivity that is about $A/A^2 = 10^{30}$ greater than that incorrectly assumed in the JASON report for the detection of relic HFGWs.

The Li-Baker HFGW detector operates as follows:

1. The perturbative photon flux (PPF), which signals the detection of a passing gravitational wave (GW), is generated when the two waves (EM and GW) have the same frequency, direction and suitable phase. This situation is termed "synchro-resonance." These PPF detection photons are generated (in the presence of a magnetic field) as the EM wave propagates along its *z*-axis path, which is also the path of the GWs, as shown in Figs. 10 and 11.

2. The magnetic field **B** is in the y-direction. According to the Li effect, the PPF detection photon flux (also called the "Poynting Vector") moves out along the x-axis in both directions.

3. The signal (the PPF) and the noise, or background photon flux (BPF) from the Gaussian beam have very different physical behaviors. The BPF (background noise photons) are from the synchro-resonant EM Gaussian beam and move in the z-direction, whereas the PPF (signal photons) move out in the x-direction along the x-axis and only occur when the magnet is on.

4. The PPF signal can be intercepted by microwave-absorbent shielded microwave receivers located on the *x*-axis (isolated from the synchro-resonance Gaussian EM field, which is along the *z*-axis).

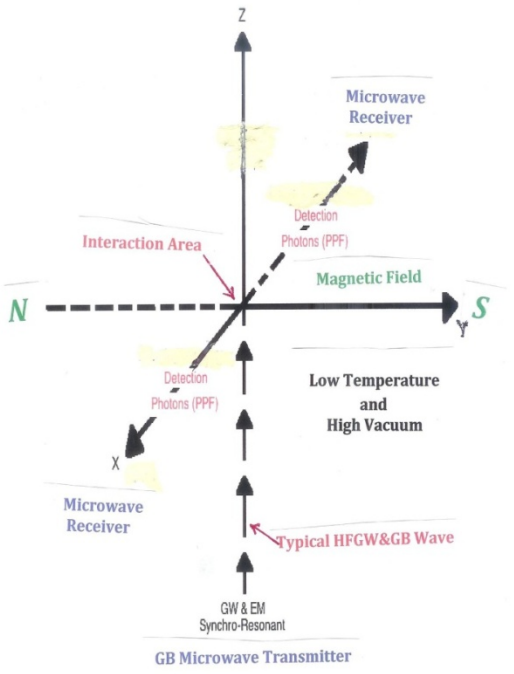


Figure 10.. Detection Photons Sent to Locations that are Less Affected by Noise.

5. The absorption is by means of off-the-shelf -40 db microwave pyramid reflectors/absorbers and by layers of metamaterial or MM absorbers (Landy, et al. 2008, Woods et al. [23] and Patent Pending) that also seal off out gassing. As discussed in detail by Woods, et al. [23] absorption of about –220 dB or an absorption coefficient of 10^{-22} for the two double MM layers, can be achieved. As noted by Landy, et al. [34] since "…impedance matching is possible, and with multiple layers, a *perfect* [absorbance] can be achieved." In addition, isolation is further improved by cooling the microwave receiver apparatus to reduce thermal noise background to a negligible amount as has been accomplished in single-photon receivers (Buller, [35]). In order to achieve a larger field of view and account for any curvature in the magnetic field, an array of microwave receivers having multiple horns (the two receivers having, for example, 12 cm by 12 cm horns (four such horns some two HFGW wavelengths or $2\lambda_{GW}$ on a side) could be installed at x = ± 100 cm (arrayed in planes parallel to the y-z plane). As noted in the following Table, all sources of noise in the Li-Baker HFGW detector such as *diffraction* from the intense Gaussian beam (Woods [36]), *dark-background shot noise, signal shot noise, Johnson noise, preamplifier noise, quantization noise, mechanical thermal noise, phase* or *frequency noise,* can be reduced to negligible amounts in a properly designed Li-Baker detector.

Noise Contributor	Brief Description of Noise source	Mitigation/Elimination Means	Nominal Computed Value photons s^{-1}, NEP W
Dark-background shot noise	GB noise especially diffraction	Wall geometry and absorbing wall materials	$4.2\times10^{-23}s^{-1}$,2.8 10^{-46}W
Signal shot noise	Noise in the signal itself	Part of useful data and not to be eliminated	--
Johnson noise	Thermal agitation in a power amplifier resistance	Refrigeration to low temperature	$5\times10^{-3}s^{-1}$,3×10^{-28} W
Preamplifier kTC noise	Stray capacitance and load resistance	Reducing bandwidth, load resistance and/or stray capacitance.	$1\times10^{-6}s^{-1}$, 8×10^{-30} W
Quantization noise	Analog to Digital Converter	Increasing the number of bits used	$2\times10^{-3}s^{-1}$, 1×10^{-28} W
Mechanical thermal noise	Brownian motion of sensor components.	Refrigeration to low temperature	$3\times10^{-4}s^{-1}$,2×10^{-29} W
Phase or Frequency noise	Fluctuations in the frequency of the microwave source for the GB.	Cavity-lock loop or a phase-compensating feedback loop	$5\times10^{-15}s^{-1}$,3×10^{-38} W

Summary Table of Li-Baker detector noise based upon experimental data concerning its components (Woods et al. [33]).

The total noise equivalent power or NEP is 1.02×10^{-26} W (noise flux is 1.54×10^{-3} photons per second). If need be the receivers could be further cooled and shielded from diffraction noise by baffles and optimum detector geometry as shown in Woods et al. [23]. Given a signal that exhibits the nominal value given in the Summary Table above of Woods et al. [23] of 99.2 s^{-1} photons, one quarter of which is focused on each of the microwave receivers, which is 24.8 s^{-1} photons or 1.6×10^{-22} W, the signal-to-noise ratio for each receiver is better than 1500:1 [23].

4. CONCLUSIONS

The utilization of modern MEMS technology and a double-helix array of them would allow for the construction of a HFGW generator or transmitter involving superradiance that exhibits sufficient strength to transmit HFGW signals globally. This is possible even though the conversion rate of EM power to GW power is exceedingly small and, like EM radiation, the GW signal power falls off as the inverse square of the distance. It is shown herein that a properly designed double-helix array of MEMS (or FBARS) can generate sufficient power to reach a receiver on the opposite side of the globe. Three HFGW detectors or HFGW receivers have previously been fabricated and others theoretically proposed, but analyses of their sensitivity suggest that for meaningful HFGW reception, greater sensitivity is necessary. The theoretical sensitivity of the Li-Baker HFGW detector discussed herein, that is based upon a different measurement technique than the other detectors, is predicted to satisfy HFGW communication needs The detector can be built from off-the-shelf, readily available components and, when coupled with the double-helix MEMS or FBAR array transmitter, could provide for transglobal HFGW communications.

Acknowledgments

This research was supported by *Transportation Sciences Corporation* in the United States and by *GravWave® LLC* internationally. The assistance of Professor Giorgio Fontana, *Trento University,* Italy, Professor Fangyu Li, *Chongqing University,* China and Gary V. Stephenson of *LinQuest Corporation,* Los Angeles, USA is gratefully acknowledged'

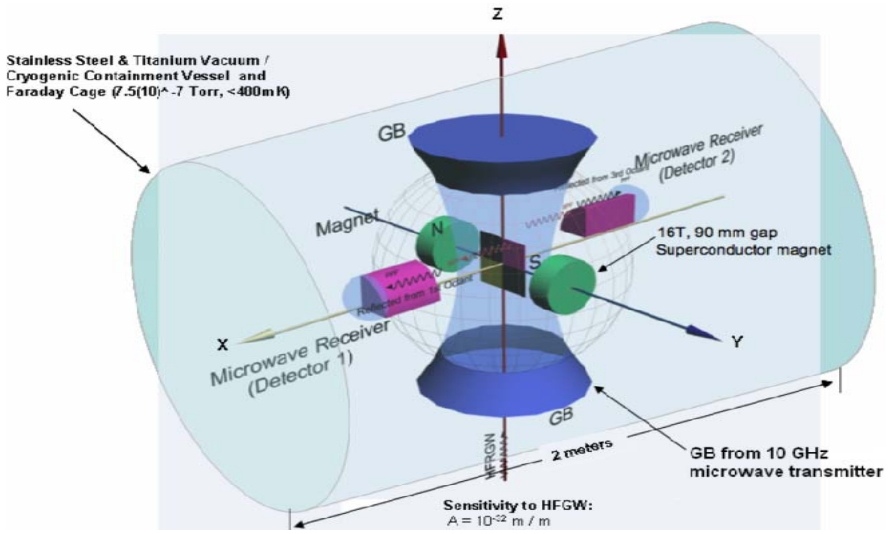


Figure 11. Schematic of Li-Baker HFGW Detector (Peoples Republic of China Patent Number 0510055882.2 [25])

. REFERENCES

[1] Douglass, D.H. and Braginsky, B. (1979), "Gravitational-radiation experiments," in "General relativity: an Einstein centenary survey" Ed. Hawking S.W. and Israel W. (CUP, UK), 90-137.

[2] Shawhan, P. S. (2004), "Gravitational Waves and the Effort to Detect them," *American Scientist* **92**, 356.

[3] Beckwith, A. W. (2009), "Relic High Frequency Gravitational waves from the Big Bang, and How to Detect them," *American Institute of Physics Conference Proceedings*, Melville, NY **1103, p.** 571... . arXiv:0809.1454v1 [physics.gen-ph] (Paper 031).

[4] Grishchuk, L. P. (2003), *Gravitational-Wave Conference,* The MITRE Corporation, May 6-9..

[5] Scully, M. O. and Svidzinsky, A. A. (2009), "The Super of Superradiance," *Science* **325**, pp.1510-1511.

[6] Li, F. Y., Tang M. and Zhao P. (1992), "Interaction Between Narrow Wave Beam-Type High Frequency Gravitational Radiation and Electromagnetic Fields," *Acta Physica Sinic*a **41,** pp. 1919-1928.

[7] Li,, F. Y., Meng-Xi Tang, Jun Luo, and Yi-Chuan Li (2000) "Electrodynamical response of a high-energy photon flux to a gravitational wave," *Physical Review D* **62**, July 21, pp. 044018-1 to 044018 -9.

[8] Li, F. Y.,, Meng-Xi Tang, and Dong-Ping Shi, (2003), "Electromagnetic response of a Gaussian beam to high-frequency relic gravitational waves in quintessential inflationary models," *Physical Review B* **67**, pp. 104006-1 to -17.

[9] Li, F. Y. and Baker, R. M L, Jr. (2007), "Detection of High-Frequency Gravitational Waves by Superconductors," *International Journal of Modern Physics* **21**, Nos. 18-19, pp. 3274-3278.

[10] Li F. Y., Baker R. M L, Jr., Fang Z., Stephenson G.V. and Chen, Z. (2008), "Perturbative Photon Fluxes Generated by High-Frequency Gravitational Waves and Their Physical Effects," *European Phys. J.* ***C*** **22**, Nos. 18-19, 30 July; peer reviews and manuscript available at: http://www.gravwave.com/docs/Li-Baker%206-22-08.pdf .

[11] Dehnen, H. and Fernando Romero-Borja (2003), "Generation of GHz – THz High-Frequency Gravitational Waves in the laboratory," paper HFGW-03-102, *Gravitational-Wave Conference,* The MITRE Corporation, May 6-9. Peer reviews and manuscript available at: http://www.gravwave.com/docs/Analysis%20of%20Lab%20HFGWs.pdf.

[12] Gertsenshtein, M. E, (1962), "Wave resonance of light and gravitational waves,"*Soviet Physics JETP*, Volume 14, Number 1, pp. 84-85.

[13] Eardley, et al. (2008) "High Frequency Gravitational Waves," JSR-08-506, October, the JASON defense science advisory panel and prepared for the Office of the Director of National Intelligence.

[14] Einstein, Albert, (1918) Über Gravitationswellen. In: Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften, Berlin (1918), 154–167.

[15] Weber, J. (1964), "Gravitational Waves" in *Gravitation and Relativity*, Chapter 5, pp. 90-105, W. A. Benjamin, Inc., New York.

[16] Grishchuk, L. P. and Sazhin M. V. (1974), "Emission of gravitational waves by an electromagnetic cavity." *Soviet Physics JETP* **38**, Number 2, pp. 215-221.

[17] Landau, L. D. and Lifshitz, E. M. (1975), *The Classical Theory of Fields*, Fourth Revised English Edition, Pergamon Press, pp. 348, 349, 355-357.

[18] Baker, R. M L, Jr., Woods, R. C. and Fangyu Li (2006), "Piezoelectric-Crystal-Resonator High-Frequency Gravitational Wave Generation and Synchro-Resonance Detection," American Institute of Physics Conference Proceedings, Melville NY **813**, pp. 1280-1289. Manuscript available at: http://www.drrobertbaker.com/docs/AIP;%20HFGW%20Piezoelectric%20Generator.pdf

[19] Woods,, R. C. and Baker, Jr., R. M L (2005), "Gravitational Wave Generation and Detection Using Acoustic Resonators and Coupled Resonance Chambers," in the proceedings of *Space Technology and Applications International Forum (STAIF-2005)*, edited by M.S. El-Genk, American Institute of Physics Conference Proceedings, Melville, NY **746**, 1298.

[20] Baker, R. M L, Jr. (2006) "Novel formulation of the quadrupole equation for potential stellar gravitational-wave power estimation" *Astronomische Nachrichten* **327**, No. 7, pp. 710-713. Peer reviews and manuscript available at: http://www.gravwave.com/docs/Astronomische%20Nachrichten%202006.pdf

[21] Baker, R. M L, Jr. and Black, C. S.(2009), "Radiation Pattern for a Multiple-Element HFGW Generator," American Institute of Physics Conference Proceedings, Melville, NY **1103.** pp. 582-590. Manuscript available at: http://www.drrobertbaker.com/docs/Analyses%20of%20HFGW%20Generators%20and%20Radiation%20Pattern.pdf

[22] Baker, R. M L, Jr., Stephenson, G. V. and Li, F. Y. (2008), "Analyses of the Frequency and Intensity of Laboratory Generated HFGWs," American Institute of Physics Conference Proceedings, Melville, NY **969**, pp. 1045-1054. Peer reviews and manuscript available at: at: http://www.gravwave.com/docs/Analysis%20of%20Lab%20HFGWs.pdf

[23] Woods, R. C., Baker, Jr. R. M L , Li, F. Y. Stephenson, G. V. Davis, E. W. and Beckwith, A. W. (2011), "A new theoretical technique for the measurement of high-frequency relic gravitational waves," *Journ Mod. Phys.* **2**, No. 6, pp. 498-518; manuscript available at: . http://www.gravwave.com/docs/J.%20of%20Mod.%20Phys%202011.pdf and abstract available at http://vixra.org/abs/1010.0062**.**

[24] Stephenson, G. V. (2009) "The standard quantum limit for the Li-Baker HFGW detector**,"** in the *Proceedings of the Space, Propulsion and Energy Sciences International Forum (SPESIF***),** 24-27 February, Edited by Glen Robertson. (Paper 023), American Institute of Physics Conference Proceedings, Melville, NY **1103** , 542-547**; manuscript available at:** http://www.gravwave.com/docs/HFGW%20Detector%20Sensitivity%20Limit.pdf

[25] Woods R. C . (2007), "Modified Design of Novel variable focus lens for VHFGW," in the proceedings of *Space Technology and Applications International Forum (STAIF-2007)*, edited by M.S. El-Genk, American Institute of Physics Conference Proceedings, Melville, NY 880:1011-1018.

[26] Baker, Jr. R. M L (2001)) Peoples Republic of China Patent Number 01814223.0, "Gravitational Wave Detector," issued September 19, 2007; patent claims available at: **http://www.gravwave.com/docs/Chinese%20Detector%20Patent%2020081027.pdf**

[27] Cruise, A. M. (2000), "An electromagnetic detector for very-high-frequency gravitational waves," *Class. Quantum Gravity* **17**, pp. 2525-2530.

[28] Cruise, A. M. and Ingley, R. M. J. (2005), "A correlation detector for very high frequency gravitational waves," *Class. Quantum Grav.* **22**, 5479-5481.

[29] Gemme, B. F.., Parodi, R. and Picasso, E. (2001), "A detector of small harmonic displacements based on two coupled microwave cavities," *Review of Scientific Instruments* **72,** Number 5, May, pp. 2428-2437.

[30] Chincarini, A and Gemme, G. (2003), "Micro-wave based High-Frequency Gravitational Wave detector," HFGW-03-103, *Gravitational-Wave Conference,* The MITRE Corporation, May 6-9.

[31] Nishizawa, A. et al. (2008), "Laser-interferometric detectors for gravitational wave backgrounds at 100 MHz: Detector design and sensitivity," *Phys. Rev. D* **77,** Issue 2**,** 022002.

[32] Cruise, A. M. (2008), "Very High Frequency Gravitational Waves," Gravitational Wave Advanced Detector Workshop (GWADW), Elba Conference, 17 May, slide presentation 132.

[33] Li, F.Y., Yang, N., Fang, Z., Baker, Jr., R. M L, Stephenson, G. V. and Wen, H.,(2009), "Signal photon flux and background noise in a coupling electromagnetic detecting system for high-frequency gravitational waves," *Phys. Rev. D.* **80**, 060413-1-14; manuscript available at: http://www.gravwave.com/docs/Li,%20et%20al.%20July%202009,%20HFGW%20Detector%20Phys.%20Rev.%20D.pdf

[34] Landy. N. I., Sajuyigbe, S,. Mock, J. J., Smith, D. R. and Padilla (2008), "Perfect Metamaterial Absorber," *Physical Review Letters* **100**, pp. 207402-1-4, May 23.

[35] Buller, G. S. *et al*. (2010) ,"Single-Photon Generation and Detection," *Measurement Science and Technology* **21**, No. 1, pp. 1-28.

[36] Woods, R. C. (2011), "Estimate of diffraction from Gaussian Beam in Li-Baker HFGW detector," *Proceedings of the Space, Propulsion and Energy Sciences International Forum (SPESIF 2011)***,** University of Maryland, Edited by Glen Robertson. (Paper 001); manuscript available at: http://www.gravwave.com/docs/Woods%202010.pdf

archived as **http://www.stealthskater.com/Documents/Baker_04.doc [pdf]**

more of Robert Baker at ***http://www.stealthskater.com/Science.htm#Baker***

note: because important websites are frequently "here today but gone tomorrow", the following was archived from
http://www.gravwave.com/pdf/Precursor%20HFGW%20Generator%20Experiments.pdf
on March 3, 2004. This is NOT an attempt to divert readers from the aforementioned website. Indeed, the reader should only read this back-up copy if it cannot be found at the original author's site.

Space Technology and Applications International Forum (STAIF-2004),
Albuquerque, New Mexico, February 8-12, 2004

Precursor Proof-of-Concept Experiments for Various Categories of High-Frequency Gravitational Wave (HFGW) Generators

Robert M. L. Baker, Jr.

Senior Consultant, GRAVWAVE® LLC and Transportation Sciences Corporation, 8123 Tuscany Avenue Playa del Rey, California 90293, USA, Telephone (310) 823-4143, Fax (310) 821-1694
e-mail robert.baker.jr@comcast.net.

Abstract

High-Frequency Gravitational Wave (HFGW) generators are separated into 3 general categories. Precursor, component-validation, laboratory experiments for each category except, possibly, the third are identified in general terms. The categories are:

(1) The electromechanical category includes micro- and nano-element, piezoelectric crystal, and multi-dielectric film HFGW generators.

(2) The high-temperature superconductor category includes gasers, impressed magnetic fields, and transformation of electromagnetic radiation into gravitational waves (Gertsenshtein effect) HFGWgenerators.

(3) The laser/plasma category includes laser-energized mirrors, synchrotron light, nuclear fusion, plasma toroid, and nonlinear optical-acoustical, molecular-level HFGW generators.

A perusal of HFGW literature reveals that since the 1960s, many authors have contributed designs of mechanisms and devices that relate to the terrestrial generation of gravitational waves. Only in the last few years, however, have any researchers demonstrated that their proposed devices were practical HFGW generators capable of producing kilowatts of power that were operational in a laboratory setting.

These recent devices make use of new technology and generate high-frequency (GHz and above) gravitational waves using non-gravitational forces. Most of the generators considered in this paper have been recently discussed at the May, 2003 Gravitational Wave Conference at The MITRE Corporation,

McLean, VA, which was the very first International Conference dedicated to HFGW and attracted 25 research papers from 9 countries. Although no detailed experimental tasks are discussed, experimental test objectives in the form of a roadmap are proposed for each category.

Introduction

There exist 3 general categories of High-Frequency Gravitational Wave (HFGW) generators: **(1) electromechanical**, **(2) high-temperature superconductor**, and **(3) laser/plasma**. We will not consider the very low-frequency (down to MHz) gravitational waves generated celestially by gravitational forces of interest to astrophysicists, which relate to a totally different technology (HFGW technology is different from that utilized in LIGO, GEO6000, VIRGO, et al).

For each of the devices comprising these general categories, there is a precursor experimental protocol that is generally described herein and is useful for experimental planning. By "precursor", I mean experimental tasks that are predecessors to the test of the capability of devices in question to actually generate detectable HFGW in the laboratory. These are component-validation, laboratory experiments. As an example of a precursor experiment, let us consider the micro- and nano-element devices included in the (1) electromechanical, HFGW-generator category.

The concept of these devices is to impart a large jerk to a mass or system of masses and thereby generate significant HFGW. In this case, the actual devices involve a myriad of energizing and energizable elements. Precursor experimental tasks would involve testing the efficacy of a single element or mechanism. That is, to measure the jerk produced in one single energizable element (in this case a small permanent magnet). If the jerk is successfully produced, then the mechanism can be replicated. Those replicated mechanisms -- when controlled by a computer logic system -- can be utilized as micro- or nano-elements in these devices, which are now capable of generating many kilowatts of HFGW in the laboratory.

This paper includes sections that are devoted to each of the 3 categories of HFGW generators. These sections then reference the specific research efforts involved in the particular category and explain the device's approach in each of the categories for HFGW generation in general terms. Most of the generators considered in this paper have been recently discussed at the May, 2003 Gravitational Wave Conference at The MITRE Corporation, McLean VA 22102, USA which was the very first International Conference dedicated to HFGW and attracted 25 researchers from 9 countries. Various precursor experiments (the experiment list is not intended to be comprehensive) are quite generally described as a guide to the device's development process.

(1) ElectroMechanical HFGW Generators

The precursor experimental tasks for the micro- and nano-device electromechanical HFGW generators (please see Baker, 2000a, 2000b, and 2003 for a description of a 380-kilowatt HFGW generator). In this category, the tasks have been discussed in the foregoing and in Baker (2004). In a little more detail regarding them, it is necessary to test the ability to produce very short (e.g., picosecond or less duration) energizing pulses (e.g., current, electromagnetic, laser, etc.). Next, the ability of various sensing or measuring instruments (e.g., strain gages, Doppler laser, interferometers, piezoelectric crystals, etc.) to measure the energizable element's jerk must be tested in order to calibrate such an experiment.

With regard to the utilization of an array of piezoelectric **electromechanical** crystals for HFGW generation, the recent paper by Dehnen and Romero-Borja (2003) describes such devices best. The concept here (as originally suggested by Weber, 1960) is to stress the crystals at high frequency (THz or

higher) by means of electrical pulses and utilize the resulting vibration, oscillation, or jerk (contractions and expansions of the crystals) to generate HFGW. An array of such crystals is connected to a computer logic system required to energize each crystal as the HFGW wavefront passes and generates a coherent and needle-like, super-radiant beam of HFGW.

At least 2 precursor experimental tasks would be of value here. First, the response of a piezoelectric crystal of various sizes and shapes to high-frequency pulses should be measured. Second, a single row of crystal vibrators (which are elements of the chain of such rows comprising the device) should be fabricated and tested at high frequency, and the motion and internal strain characteristics of the piezoelectric crystals studied.

As indicated by Portilla and Lapiedra (2001), "An electric charge shaken (or jerked) in a homogeneous stationary magnetic field produces both electromagnetic and gravitational waves (Gertsenshtein waves)." In addition, Portilla (2003) suggested that such a device could be "...utilized to construct a (**electromechanical**) generator of HFGW and ... by connecting the generator elements to a computer controlled logic system our device could be incorporated in a communications system."

The latest publication (Navarro, Portilla, and Valdes, 2004) suggests a test of such a concept through use of an electron paramagnetic resonant spectrometer. Precursor tests of the strong magnetic field, vacuum chamber, and microwave generator are also in order. The sensitivity of their antenna should be studied. Finally, several configurations of a waveguide contained multi-dielectric film should be tested.

In summary, the precursor experiments that are, in general, generic to this category are elucidated in Baker (2004) and Navarro, Portilla, and Valdes (2004).

(2) High-Temperature Superconductor HFGW Generators

The HFGW generator equivalent of the laser (the **gaser)** was originally proposed by Halpern and Laurent (1968); again suggested by Fontana (1998); and then presented in more detail by Fontana and Baker (2003). In this latter article, they adopted a combination of 3 theories on gravitational-wave emission at the quantum level (the jerk) to show that their approach can offer the required smallness of the transition line-width and the required increase of the emission probability.

In their gaser concept, coherence would be provided by the collective wave-function of Cooper pairs in a **High-Temperature SuperConductor** (HTSC) to generate high-frequency gravitational waves (HFGW). They calculated the output power of the HFGW (10 megawatts at 1.3 THz) and also proposed a configuration capable of producing a focused beam of radiation of HFGW. The focus can be dynamically adjusted by "...a mechanical connection of the generator elements to a computer controlled logic system." Fontana (2004) suggests a test methodology that would involve precursor tests of the electric current source, the gaser material, and the HTSC.

Ning Li (2003) looked at a curved space and at the gravitational scalar and vector potentials. The scalar potential is ignored because it is too difficult to deal with. When Li completes her algebra, it appears that High-Temperature SuperConductors (HTSC) -- by virtue of the jerk of the Cooper-pair electrons -- can generate **11 kilowatts** of gravitational-wave power under an impressed magnetic field controlled by a computer logic system similar to the concept of Fang-Yu Li, Tang, and Shi (2003). Precursor experiments would involve the test of an alternating (AC) magnetic field: its intensity, its highest frequency and variation of strength, and the ability to fabricate a YBCO HTSC of sufficient size to generate the HFGW.

Chiao (2002) describes an experiment in which he plans to convert electromagnetic (EM) waves into HFGW (Gertsenshtein effect) inside a device that is poised to go from a normally conducting state to a superconducting state. Specifically, "...that **Type II high-temperature superconductors** (such as YBCO) are considered to be Macroscopic quantum gravitational antennas which can simultaneously also be used as efficient transducers (by jerking electrons) for converting EM into HFGW, and vice versa." Precursor experiments would involve activating and testing the EM generator to define its characteristics and testing various designs of the YBCO HTSC. **[StealthSkater note: also see the UNITEL proposals to use special Type II semiconductors in a unique quantum laser that would -- in combination with a special Niobium-Tin SED-implanted "smart skin" achieve Macroscpic quantum tunneling =>** doc pdf URL **]**

In summary, the precursor experiments that are -- in general -- generic to this category are elucidated in Fontana (2004) and Chiao (2002).

(3) Laser/Plasma HFGW Generators

Baker (2004) describes the use of a pulsed laser beam (very short -- e.g., femtosecond or attosecond -- pulses at high repetition frequencies -- e.g., MHz and higher) that produce jerks in a target-mirror energizable element (Baker 2000b) and about 6 kilowatts of HFGW power. The precursor tests involve the generation of the pulses and the tests of various target mirrors as discussed in more detail in Baker (2004).

Rudenko (2003) suggests that for HFGW generation one should utilize strong non-gravitational forces. Possibly those that exist within a HTSC, or a large array of micro-electromechanical devices, or large numbers of coherently excited nano-piezoelectric crystals, etc., or "...to keep the resonance condition (i.e., maximum vibration or jerk amplitude) one has to decrease the mass (scale) of elementary radiator taking into account a possibility of using a natural mass-quadrupole at the molecular level. The deficit of mass in this case might be covered by increasing the total number of coherent micro radiators and detectors N_r , N_d ..

"It is obvious the maximum density of cells can not exceed the value A= 10^{22}-10^{24} per cm^3. After this argumentation, it is reasonable to seek a solution of the problem of effective GW generation by means of collective excitation of molecular oscillations at optical frequencies taking place in a nonlinear optic-acoustical media. One has to choose a molecular with largest quadrupole moment. An instrument for their excitation might be a powerful optical **laser** beam. A **coupling of optical, acoustical, and gravitational waves** must be provided by nonlinear optical permeability of the media. A ccumulation of the GW power as well as its beaming might be provided by the synchronism conditions typical for interaction of waves traveling through nonlinear optical media."

Rudenko concludes that traveling electromagnetic (please see also Grishchuk, 2003 in this same regard) and acoustic (quadrupole) waves interacting (jerking) in nonlinear medium looks quite reasonable and promising at the range of optical frequencies, 10^{14}-10^{15} HZ) for the generation of HFGW recursor experiments would include tests of powerful laser beams, of nonlinear optical permeable materials, and traveling EM waves.

Stephenson (2003) suggests the use of strong synchrotron light in a strong magnetic field to generate HFGW via the Gertsenshtein effect. The quadrupole (jerk) modulation of a synchrotron light source (SLS) and a low-hysteresis target for light generation are proposed. The precursor experiments would include test of various SLS and target designs. Stephenson also suggests a **plasma** toroid approach using plasma modulation to couple quadrupolar energy flow for HFGW generation. Precursor

experiments would involve tests of various mass/plasma designs, confinement designs (e.g., tokomak), and tests of modulation equipment.

Nuclear fusion generation of gravitational waves was originally studied by Chapline, Nuckolls, and Woods(1974). More recently, Rudenko suggested "... **laser**-stimulated thermonuclear fusion or laser micro-target explosions look more promising for the laboratory generation of HFGW." It is difficult to suggest precursor experiments here due to the safety concerns with nuclear reactions. However scientists at, for example, the Lawrence Livermore National Laboratory or similar nuclear research facilities in Russia or China might have some suggestions.

This 3rd category of HFGW generators is the least mature of the group and may not be ready for component-validation laboratory experiments. It is, therefore, difficult to establish generic guidelines for precursor experiments. As Paul Murad has suggested (May 8, 2003), one might consider "... using a laser (energizing element) to radiate a gas such that shock waves are formed to create a pressure pulse that induces a particle (energizable element) jerk." Measurement of such an imbedded particle (or gas molecule) jerk might reveal the capability of such a process -- with greatly increased power -- to generate HFGW.

Conclusions

It is self evident from the research efforts cited in this paper that significant theoretical analysis has been devoted of late to the laboratory or terrestrial generation of HFGW.

Now is the time to supplement these studies with a series of precursor experiments. Experiments that will incrementally lead to the fabrication and test of one or more HFGW generators and the detection of their emanations by the HFGW detectors that have already been fabricated in England (please see Cruise, 2000), Italy (please see Chincarini and Gemme, 2003) and China (please see Li, Tang, and Shi, 2003).

The time is right, *carpe diem* ... seize the moment!

Acknowledgment

I wish to express my great appreciation to Paul Murad for his editorial and substantive contributions to this paper.

[StealthSkater note: the reader might want to read the proposal of Prof. Ray Kramer (father of famous missing person Philip Taylor Kramer) to harness "unstable gravitational waves" for instantaneous teleportation of information (and possibly matter) to any point in the Universe => doc pdf URL]

References

Baker, R. M.L., Jr., "Gravitational Wave Generator," **United States Patent 6,417,597**, 50 claims, filed July 14, 2000a.

Baker, R. M.L., Jr., "Gravitational Wave Generator Utilizing Submicroscopic Energizable Elements," **United States Patent Application 09/752,975**, 100 claims, filed December 27, 2000b.

Baker, R. M.L., Jr., "Generation of High-Frequency Gravitational Waves (HFGW) by Means of an Array of Micro-and Nano-devices," Gravitational-Wave Conference, edited by P. Murad and R. Baker, The MITRE Corporation, Mclean, Virginia, May 6-9, 2003, Paper HFGW-03-117.

Baker, R. M.L., Jr., "Precursor Experiments Regarding the Generation of High-Frequency Gravitational Waves (HFGW) by Means of Using an Array of Micro- and Nano-Devices," Space Technology and Applications International Forum (STAIF-2004), edited by M.S. El-Genk, American Institute of Physics, Melville, New York, February 8-12, 2004, Paper F02-2-179.

Chapline, G.F., Nuckolls, J., and Woods, L.L., "Gravitational-Radiation Production Using Nuclear Explosions," *Physical Review D.*, 10, Number 4, August, 1974, pp. 1064-1065.

Chiao, R.Y., "Superconductors as Transducers and Antennas for Gravitational and Electromagnetic Radiation," arXiv:gr-qc/0204012 v2, April 11, 2002.

Chincarini A. and Gemme, G., "Microwave Based High-Frequency Gravitational Wave Detector," Gravitational-Wave Conference, edited by P. Murad and R. Baker, The MITRE Corporation, Mclean, Virginia, May 6-9, 2003, Paper HFGW-03-103.

Cruise, A.M. (2000), "An electromagnetic detector for very-high-frequency gravitational waves," *Class. Quantum Gravity*, 17, 2000, pp. 2525-2530.

Dehnen, H. and Romero-Borja, F., "Generation of GHz-THz High-Frequency Gravitational Waves in the Laboratory," Gravitational-Wave Conference, edited by P. Murad and R. Baker, The MITRE Corporation, Mclean, Virginia, May 6-9, 2003, Paper HFGW-03-102.

Fontana, G, "A possibility of Emission of High Frequency Gravitational Radiation from Junctions between d-wave and s-wave Superconductors," Preprint, Faculty of Science, University of Trento, 38050 Povo (TN), Italy, 1998, pp. 1-8.and **http://xxx.lanl.gov/html/cond-mat/9812070**.

Fontana, G. and Baker, R. M.L., Jr., "The High-Temperature Superconductor (HTSC) Gravitational Laser (GASER)," Gravitational-Wave Conference, edited by P. Murad and R. Baker, The MITRE Corporation, Mclean, Virginia, May 6-9, 2003, Paper HFGW-03-107.

Fontana, G., "Design of a Quantum Source of High-Frequency Gravitational Waves (HFGW) and Test Methodology," Space Technology and Applications International Forum (STAIF-2004), edited by M.S. El-Genk, American Institute of Physics, Melville, New York, February 8-12, 2004, Paper F02-1-143.

Grishchuk, L.P., "Electromagnetic generators and detectors of gravitational waves," Gravitational-Wave Conference, edited by P. Murad and R. Baker, The MITRE Corporation, Mclean, Virginia, May 6-9, 2003, Paper HFGW-03-119.

Halpren, L. and Jouvet, B., "On Stimulated Photon-Graviton Conversion by an Electromagnetic Field," *Annale H. Poincaré*, VII, NA1, pp. 25ff. (1968).

Li, Ning , "Measurability of AC gravity fields," Gravitational-Wave Conference, edited by P. Murad and R. Baker, The MITRE Corporation, Mclean, Virginia, May 6-9, 2003, Paper HFGW-03-106.

Li, Fang-Yu, Tang, Meng-Xi and Shi, Dong-Ping , "Electromagnetic response for High-Frequency Gravitational Waves in the GHz to THz band," Gravitational-Wave Conference, edited by P. Murad and R. Baker, The MITRE Corporation, Mclean, Virginia, May 6-9, 2003, Paper HFGW-03-108.

Navarro, E., Portilla, M., and Valdes, Jose Luis, "Test of the Generation of High-Frequency Gravitational Waves by Irradiating a Multidielectric Film," Space Technology and Applications International Forum (STAIF-2004), edited by M.S. El-Genk, American Institute of Physics, Melville, New York, February 8-12, 2004, Paper F02-3-141.

Portilla, M. and Lapiedra, R., "Generation of High Frequency Gravitational Waves," *Physical Review D 63*, pp. 044014-1 to 044014-7 (2001).

Portilla, M., “Generation of HFGW by Irradiating a Multidielectric Film,” Gravitational-Wave Conference, edited by P. Murad and R. Baker, The MITRE Corporation, Mclean, Virginia, May 6-9, 2003, Paper HFGW- 03-112.

Rudenko, V. N., “Optimization of Parameters of a Coupled Generator-Receiver for a Gravitational Hertz Experiment,” Gravitational-Wave Conference, edited by P. Murad and R. Baker, The MITRE Corporation, Mclean, Virginia, May 6-9, 2003, Paper HFGW-03-113.

Stephenson, G. V., “The Application of High-Frequency Gravitational Waves (HFGW) to Communications,” Gravitational-Wave Conference, edited by P. Murad and R. Baker, The MITRE Corporation, Mclean, Virginia, May 6-9, 2003, Paper HFGW-03-104.

Weber, J., “Detection and Generation of Gravitational Waves,” *Physics Review*, 117, Number 1, pp.306-313, (1960).

if on the Internet, Press <BACK> on your browser to return to the previous page (or go to www.stealthskater.com)

else if accessing these files from the CD in a MS-Word session, simply <CLOSE> this file's window-session; the previous window-session should still remain 'active'

Time and Matter
2007 Bled, Slovenia

Generation and Detection of Gravitational Waves at Microwave Frequencies by Means of a Superconducting Two-body System

R.Y. CHIAO*
School of Natural Sciences and School of Engineering, University of California, P.O. Box 2039, Merced, CA 95344, U.S.A.

Abstract: The 2-body system of a superconducting sphere levitated in the magnetic field generated by a persistent current in a superconducting ring, can possibly convert gravitational waves into electromagnetic waves, and vice versa. Faraday's law of induction implies that the time-varying distance between the sphere and the ring caused by the tidal force of an incident gravitational wave induces time-varying electrical currents, which are the source of an electromagnetic wave at the same frequency as the incident gravitational wave. At sufficiently low temperatures, the internal degrees of freedom of the superconductors are frozen out because of the superconducting energy gap, and only external degrees of freedom, which are coupled to the radiation fields, remain. Hence this wave-conversion process is loss-free and therefore efficient, and by time-reversal symmetry, so is the reverse process. A Hertz-like experiment at microwave frequencies should therefore be practical to perform. This would open up observations of the gravitational-wave analog of the Cosmic Microwave Background from the extremely early Big Bang, and also communications directly through the interior of the Earth.

Consider the configuration of a superconducting sphere levitated above a superconducting ring, as shown in Figure 1 [this was suggested by Clive Rowe, personal communication]. For levitation to occur, it is required that the downwards force of Earth's gravity $\mathbf{F}_{\mathrm{GR}}(0)$

on sphere be exactly balanced by the upwards electromagnetic force $\mathbf{F}_{\mathrm{EM}}(0)$

on the sphere due to the complete expulsion of the $\mathbf{B}$ field in the Meissner effect [1], so that

$$\mathbf{F}_{\mathrm{EM}}^{(0)} = -\mathbf{F}_{\mathrm{GR}}^{(0)}$$

* rchiao@merced.edu

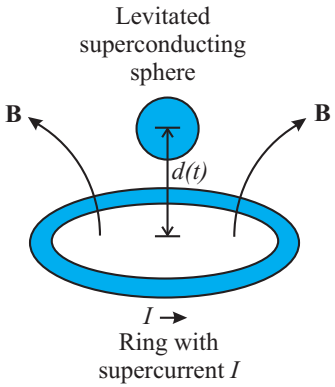


Figure 1: The DC magnetic **B** field generated by a superconducting ring with a persistent current I levitates a superconducting sphere. When a (+) polarized gravitational wave propagates radially inwards and converges upon the center of mass of this two-body system, the distance between the sphere and the ring $d(t)$ changes periodically with time, for example, at a microwave frequency, with small-amplitude excursions around an average distance of on the order of the microwave wavelength. This superconducting two-body configuration, viewed as a gravitational-wave antenna, can possibly efficiently convert this incident gravitational (GR) wave into an electromagnetic (EM) wave, and vice versa. Under the operation of time reversal, the current I and the magnetic field **B** of the ring in this Figure are reversed in direction, in order to achieve the time-reversed process of converting the EM wave back into the GR wave.

Note that the zero superscripts denote forces that are evaluated in the zeroth order of perturbation theory, i.e. in the absence of any perturbing radiation fields.

Now let a single mode of a weak, linearly polarized gravitational wave [2] at a microwave frequency (such as from the Big Bang), with one of its polarization axes along the vertical axis, be incident upon this two-body system. For simplicity, let us assume that the mass of the ring is much larger than that of the sphere, so that the reduced mass of the two-body system becomes simply that of the sphere, and the center of mass of the system approximately coincides with the center of the ring. The gravitational wave then exerts a time-varying tidal force upon the sphere relative to the ring. According to a distant inertial observer, the space between the

sphere and the ring is being periodically squeezed and stretched with time, so that the distance $d(t)$ is a sinusoidal function of time.

The mechanism for the wave conversion is the following: In the presence of the DC **B** field of the superconducting ring, Faraday's law of induction implies that the time-varying distance $d(t)$ between the sphere and the ring causes induced electrical currents to circulate azimuthally around the bottom pole of the sphere at a microwave frequency, and similarly, induced countercurrents around the ring. These time-varying induced currents are a source of electromagnetic radiation at the same frequency as that of the incident gravitational radiation. Thus a gravitational (GR) wave can in principle be converted into an electromagnetic (EM) wave [3]. Lenz's law implies that there exists an instantaneous EM force $\mathbf{F}_{\mathrm{EM}}$ acting on the sphere which opposes the instantaneous GR wave tidal force $\mathbf{F}_{\mathrm{GR}}$ acting on the sphere, such that under certain circumstances (to be discussed below) quasi-static mechanical equilibrium holds, i.e.,

$$\mathbf{F}_{\mathrm{EM}} = -\mathbf{F}_{\mathrm{GR}}.$$

Note that this equation implies an equality of the magnitudes of the electromagnetic and gravitational forces on the sphere.

Let us now examine whether or not there can exist circumstances under which a complete conversion of incoming GR wave power into outgoing EM wave power in Figure 1 can in principle occur. Under these circumstances, energy conservation demands that, in terms of quantities as measured by the distant inertial observer, the power absorbed from GR wave is equal to the power emitted into EM wave,

$$\langle \mathbf{F}_{\mathrm{GR}} \cdot \mathbf{v}_{\mathrm{rad}} \rangle = -\langle \mathbf{F}_{\mathrm{EM}} \cdot \mathbf{v}_{\mathrm{rad}} \rangle \qquad (1)$$

The angular brackets denote the time average over one period of the radiation fields, and $\mathbf{v}_{\mathrm{rad}}$ is the instantaneous radiation-damping velocity of the sphere undergoing rigid-body motion relative to the ring, which is parallel to, and in phase with, the instantaneous force $\mathbf{F}_{\mathrm{GR}}$, as seen by the distant observer. These circumstances can occur in a regime in which the electromagnetic radiation damping of the motion of the superconducting two-body system is the dominant damping mechanism as compared to all other loss mechanisms.

The superconducting sphere and ring can both undergo rigid-body motion in response to the gravitational wave, because the quantum adiabatic theorem [3, 4] applies to all perturbations due to any kind of radiation, whenever the frequency of these perturbations is below the BCS gap frequency [1]. Superconductors stay adiabatically, hence rigidly, in the BCS ground

state of the system, because no internal excitations are allowed within the BCS energy gap [1, 3, 5].

Also, there must be little energy dissipation, such as into heat, other than that due to radiation damping or conversion. This requirement is satisfied by the strictly zero resistance of the superconductors, which follows from the fact that no dissipative excitations into heat due to ohmic processes (or phonon generation to be discussed later) are allowed within the BCS energy gap, when the frequency of the radiation is lower than the BCS gap frequency [1, 3, 5].

Thus we see that a complete, loss-free conversion from GR wave energy to EM wave energy can indeed occur, provided that all internal degrees of freedom of the superconducting two-body system are frozen at low temperatures, so that only its external degrees of freedom remain. For these external degrees of freedom, in the case when electromagnetic radiation damping is dominant, there then can exist a quasi-static mechanical equilibrium at each instant of time, such that the instantaneous forces obey the equality $\mathbf{F}_{\mathrm{GR}} = -\mathbf{F}_{\mathrm{EM}}$, just like in the case of a linear induction motor (or generator). Then the power in a single, incoming GR wave mode can be completely converted into the power in a single, outgoing EM wave mode.

Under the operation of time reversal, the single, outgoing EM wave mode becomes a single, incoming EM wave mode incident upon the superconducting two-body system. This EM wave mode is then back-converted into an outgoing GR wave mode. This reverse (or reciprocal) process has the same complete, loss-free conversion efficiency as the forward process by time-reversal symmetry [3, 6]. Energy conservation demands that, the power absorbed from EM wave is equal to the power emitted into GR wave,

$$\langle \mathbf{F}_{\mathrm{EM}} \cdot \mathbf{v}_{\mathrm{rad}} \rangle = -\langle \mathbf{F}_{\mathrm{GR}} \cdot \mathbf{v}_{\mathrm{rad}} \rangle \tag{2}$$

But this is satisfied by Lenz's law and the equality $\mathbf{F}_{\mathrm{GR}} = -\mathbf{F}_{\mathrm{EM}}$ that is also valid for the reverse process. Here the force $\mathbf{F}_{\mathrm{EM}}$ is squeezing and stretching periodically the space between the sphere and the ring, so that the sphere is acting upon space, not moving through space, in a time-reversal of the forward process.

Under the action of the gravitational wave, the space between the sphere and the ring is undergoing small amplitude, anisotropic, periodic time-varying strains h_{ij} of the metric tensor [2], as if the space were an elastic medium. According to the Equivalence Principle [2], in both the forward and the reverse processes, the sphere is not accelerating at all with respect to a tiny, local inertial observer located at its center, so that the sphere is not moving, i.e., accelerating, through space. Hence the objection that its

inertia limits the amount of emitted GR radiation to extremely small values is invalid; see section 12 of [3]a. Similarly, the objection that the motion of the sphere through a flat-space background must be relativistic (i.e., with velocities close to c) in order to generate any appreciable amount of GR waves is also invalid.

Therefore we propose here to perform a Hertz-like experiment at 12 GHz, in which the EM-to-GR wave-conversion process becomes the source of GR waves, and the GR-to-EM wave-conversion process becomes the receiver of GR waves. Faraday cages consisting of normal metals at room temperature (the metallic casings of the two well-separated liquid helium storage dewars to be described below) prevent the transmission of EM waves, so that only GR waves, which can easily pass through all ordinary, classical matter, such as the normal (i.e., dissipative) metals of which standard, room-temperature Faraday cages are composed, are transmitted between the two halves of the apparatus that serve as the source and the receiver, respectively [3]. Such an experiment should be practical to perform using standard microwave sources and receivers, since the scattering cross-sections and the wave conversion efficiencies of the two-body superconducting system such as the one depicted in Figure 1, when viewed as a gravitational-wave antenna, should be large enough to be experimentally interesting [3]; see also below].

We plan to use lead (Pb) as the type I superconducting material for both the sphere and the ring. The sphere could consist of a light, low-density material (e.g., Styrofoam) coated on its surface with a Pb film which is much thicker than the penetration depth [5]. The ring could consist of a flat copper gasket electroplated with a thick coating of Pb. The central hole of the gasket could have a diameter of 34 mm, which should levitate the superconducting sphere above the hole of the ring at a distance of around 6 mm, i.e., approximately a quarter of a microwave wavelength at 12 GHz where we plan to work (12 GHz is over an order of magnitude below the BCS gap frequency of Pb, so that the losses at this frequency are negligible [3]). A persistent current in the ring sufficient to levitate the sphere could be induced by means of a nearby permanent magnet, which would then be removed after the ring has been cooled to become superconducting below the transition temperature of Pb at 7.2 K [5]. We plan also to experiment with other two-body geometries, such as with a pair of superconducting coaxial rings with opposing trapped magnetic flux [suggested by S. Minter], as realizations of the superconducting two-body system. We have chosen to use a type I superconductor instead of a type II superconductor, in order to avoid microwave losses arising from the Abrikosov vortex-motion degrees of freedom.

We propose to perform the first experiments in a standard 4 K liquid helium storage dewar using a dewar insert, which consists of a 50 mm diameter cylindrical structure with a detachable sample holder that is capable of containing a specific sphere-ring configuration (or other superconducting two-body geometries) at its bottom end. Such a dewar insert could be inserted directly into the liquid helium inside a storage dewar, in order to explore the parameters for the levitation of the sphere by the ring, by adjusting the dimensions of the sphere-ring configuration. Optical access and viewing could be made possible by means of glass prisms placed inside the sample can near the ring. Alternatively, a Hall probe sensor could detect the onset of the Meissner effect and levitation of the sphere. These dewar-insert experiments would enable us to test the computer simulations we are presently conducting for the levitation of type I superconducting spheres by various magnetic field configurations.

Two of these dewar inserts, each with a sphere-ring configuration (or some other superconducting two-body configuration) immersed deep inside two well separated 4 K liquid helium dewars, with the associated microwave electronics, would then allow us to perform the Hertz-like experiment. In preparation for this experiment, we plan to determine at room temperature the microwave antenna design parameters of the sphere-ring and the two-coaxial-ring configurations (and other possible configurations) using a network analyzer. We plan to use an appropriately designed magnetic loop antenna to couple to the TE_{01} microwave cylindrical mode of a superconducting ring for transferring energy from the GR wave mode to the EM wave mode, and vice versa.

We have already built and tested microwave sources and receivers at 12 GHz [6]. We have used a frequency-doubled 6 GHz microwave-cavity oscillator as the source, with an output power level at 12 GHz of around 10 mW, and have measured that a standard, commercial satellite-dish KU band receiver [Precision model PMJ-LNB KU gold label series from Astrotel Communications Corp.] has a noise figure of 0.6 dB at 12 GHz, so that our receiver system Noise-Equivalent-Power (NEP) sensitivity will be on the order of 10^{-25} W/Hz$^{1/2}$. This will allow us to achieve a very good signal-to-noise ratio for the Hertz-like experiment [3], assuming that our estimates for the cross-sections and wave-conversion efficiencies for the sphere-ring system are of the correct orders of magnitude.

If we should be successful in the Hertz-like experiment, one of the first tests to see if we truly have GR rather than EM coupling between the transmitter (or source) and receiver (or detector) halves of the apparatus, is to tilt the transmitter part of the apparatus by $+22.5^\circ$ with respect to the vertical

around the line of sight joining the transmitter and receiver, and to tilt the receiver part of the apparatus by -22.5° with respect to the line of sight. The signal should be extinguished at the resulting 45° relative orientation between the two halves of the apparatus, and not at 90°, as would the case for EM waves. This would be a clear signature that we have successfully generated and detected GR waves rather than EM waves [this was suggested by Kirk Wegter-McNelly, personal communication].

Another of our initial goals if we should be successful in detecting a signal in the Hertz-like experiment would be to directly measure the speed of gravitational waves for the first time. This can be done either by changing the distance between the two experimental dewars containing the two sphere-ring (or other) configurations, and then measuring the resulting change in phase of the received signal relative to that of the source signal, or by modulating the carrier signal, and measuring the delay in the arrival time of the demodulated signal. We will thereby be able to directly check experimentally in the laboratory for the first time the theoretical prediction by Einstein that the speed of these waves is identical to that of light in vacuum $c = 3.00 \cdot 10^8$ m/s [2] , to within $\pm 1\%$ accuracy.

It may be objected that sound waves will be generated as a dissipative mechanism in the proposed experiment. However, sound waves have a speed that is typically five orders of magnitude smaller than c. If Einstein were correct, then the typical generated sound wavelength of 250 nm or smaller would be much less than the gravitational wavelength of 25 mm, so that the typical sound wave would be badly mismatched to the incident GR wave. The generation of sound waves (or phonons) within the superconductor at microwave frequencies inside the BCS gap by the incident gravitational wave would therefore be forbidden due to the Mössbauer-like (or zero-phonon) effect described in section 9 of [3]a; see also [3]b. Hence the sphere and the ring would undergo rigid-body motion in response to the incident gravitational wave.

It may also be objected that the microwave-frequency electrical currents induced through Faraday's law in the GR to EM wave conversion process may be too feeble and too difficult to detect. However, we are not proposing to detect these currents, but rather to detect the power transmitted from the GR wave to the EM wave, and vice versa. With an antenna design which takes into account impedance-matching and cross-section considerations properly, the power transmitted from the source to the receiver should be readily measurable.

To better understand the cross-section estimate for the above proposed experiment, it is useful to introduce a weak-field representation of the lin-

earized Einstein equations for the gravito-electric field $\mathbf{E}_{\rm G} \equiv \mathbf{g}$ (i.e., the gravitational acceleration of a test mass), and for the gravito-magnetic field $\mathbf{B}_{\rm G}$ (i.e., the far-field, time-varying analog of the Lense-Thirring field), as observed in the coordinate system of a distant inertial observer. This representation takes the form of the Maxwell-like equations [7]

$$\nabla \cdot \mathbf{E}_{\rm G} = -\frac{\rho_{\rm G}}{\varepsilon_{\rm G}} \tag{3}$$

$$\nabla \times \mathbf{E}_{\rm G} = -\frac{\partial \mathbf{B}_{\rm G}}{\partial t} \tag{4}$$

$$\nabla \cdot \mathbf{B}_{\rm G} = 0 \tag{5}$$

$$\nabla \times \mathbf{B}_{\rm G} = \mu_{\rm G}\left(-\mathbf{j}_{\rm G} + \varepsilon_{\rm G}\frac{\partial \mathbf{E}_{\rm G}}{\partial t}\right) \tag{6}$$

where $\rho_{\rm G}$ is the mass density and $\mathbf{j}_{\rm G}$ is the mass current density of slowly-moving matter in the source, and the gravitational analog of the electric permittivity of free space ε_0 is

$$\varepsilon_{\rm G} = \frac{1}{4\pi G} = 1.19{\cdot}10^{9} \text{ units} \tag{7}$$

where G is Newton's constant, and the gravitational analog of the magnetic permeability μ_0 of free space is

$$\mu_{\rm G} = \frac{4\pi G}{c^2} = 9.31{\cdot}10^{-27} \text{ units} \tag{8}$$

where c is the speed of light.

The fields $\mathbf{E}_{\rm G} \equiv \mathbf{g}$ and $\mathbf{B}_{\rm G}$ are measurable quantities that obey the gravitational analog of the Lorentz force law [7]

$$\mathbf{F} = m(\mathbf{E}_{\rm G} + 4\mathbf{v} \times \mathbf{B}_{\rm G}) \tag{9}$$

where m is the mass of a test particle, and $\mathbf{v}$ is its velocity, as seen by the distant inertial observer.

These Maxwell-like equations are linear, so that the fields obey the superposition principle not only in the vacuum outside of the source, but also in the matter inside the source, provided that the field strengths are sufficiently weak and the matter is sufficiently slowly moving, so that there exists a regime of a linear response of the matter to the applied fields. The resulting optics for gravitational waves is therefore linear, just like the linear optics for electromagnetic waves.

Therefore any argument involving the characteristic power scale [2]

$$P_0 = \frac{c^5}{G} = 3.6{\cdot}10^{52}\,\mathrm{W} \tag{10}$$

is irrelevant in the context of linear optical devices such as linear mirrors, since it should always be possible to reduce the intensity of the incident GR waves so that there exists a linear regime of response of such devices to these sufficiently weak GR wave amplitudes. Then there cannot exist any characteristic scale of power in any linear response of such devices, such as that given by Equation (10).

Also, it is important to emphasize that the matter generating the GW waves in such linear devices as the two-body superconducting system described above, needs not be moving relativistically with velocities close to the speed of light c relative to each other, in order to generate any appreciable amount of GR waves. By the time-reversal symmetry argument given above, it is clear that the generation of GR waves by the two-body superconducting system is not due to the matter moving through space, but rather is due to the matter acting upon space directly. Hence it is irrelevant whether such matter is moving close to c through space or not.

In the case of the vacuum, where ρ_G and $\mathbf{j}_G$ both vanish, these equations lead to wave propagation at a wave speed exactly equal to the speed of light

$$c = \frac{1}{\sqrt{\varepsilon_G \mu_G}} = 3{\cdot}10^8\,\mathrm{SI\ units}. \tag{11}$$

A plane wave solution of the Maxwell-like equations possesses the gravitational characteristic impedance of free space given by [8, 3]

$$Z_G = \sqrt{\frac{\mu_G}{\varepsilon_G}} = 2.79{\cdot}10^{-18}\,\mathrm{SI\ units}, \tag{12}$$

which is the analog of the electromagnetic characteristic impedance of free space

$$Z_0 = \sqrt{\frac{\mu_0}{\varepsilon_0}} = 377\,\mathrm{Ohms}. \tag{13}$$

The gravitational characteristic impedance of free space Z_G, like Z_0, plays a central role in all radiation problems, such as in a comparison of the radiation resistance of gravitational-wave antennas to the value of this impedance, in order to estimate the coupling efficiency of these antennas to free space. The numerical value of Z_G is extremely small, but the impedance of all material objects must be "impedance matched" to this

extremely small quantity before significant power can be transferred efficiently from gravitational waves to these objects, or vice versa.

In contrast to the electromagnetic-wave case, in the gravitational-wave case, all ordinary, classical matter, such as Weber bars, possesses impedances much larger than that of the gravitational characteristic impedance of free space Z_G. It is therefore extremely difficult to impedance-match gravitational waves to any ordinary, classical matter. As a consequence, it is a general rule that all ordinary, classical matter, such as a Weber bar, is essentially completely transparent to these waves.

However, phase-coherent, loss-free quantum matter, such as superconductors, which can possess strictly zero dissipation due to the presence of the BCS energy gap, can be exceptions to this general rule. Experimental evidence for this "quantum dissipationlessness" is the fact that persistent currents in annular superconducting rings have been observed to last much longer than the age of the Universe [1]. One important consequence of the zero-dissipation property of a superconductor is that a mirror-like reflection of sufficiently low-frequency gravitational waves can occur at a planar interface between the vacuum and the superconductor. Note that by "low frequency" we mean frequencies much less than the BCS gap frequency.

In the electromagnetic case, the reflection coefficient R of a wave being transmitted down a transmission line with impedance Z, which is terminated by a resistance R which vanishes like that of a superconductor, is given by

$$R_G = \left| \frac{Z - R}{Z + R} \right|^2 \rightarrow 100\% \quad \text{as} \quad R \rightarrow 0. \tag{14}$$

This implies that a mirror-like reflection of the wave occurs from a superconductor, when it is used as a short-circuit termination of the transmission line.

Similarly, the reflection coefficient RG of a sufficiently low-frequency gravitational wave from a superconductor-vacuum interface is given by

$$R_G = \left| \frac{Z_G - R_G}{Z_G + R_G} \right|^2 \rightarrow 100\% \quad \text{as} \quad R_G \rightarrow 0. \tag{15}$$

where the gravitational characteristic impedance of free space Z_G is given by Equation (12), and R_G is the gravitational analog of the resistance R of the superconductor. This implies that a mirror-like reflection of a sufficiently low-frequency gravitational wave could in principle occur from the surface of the superconductor.

Therefore mirrors for gravitational waves can in principle exist. Curved, parabolic mirrors can focus these waves, and Newtonian telescopes for gravitational radiation can in principle be constructed. In the case of scattering of gravitational waves from superconducting bodies, the above mirror-like-reflection condition implies hard-wall boundary conditions at the surfaces of these superconducting bodies, so that the scattering cross-section of these waves from large superconducting bodies can in principle be geometric, e.g., hard-sphere, in size. For example, for a superconductor in the form of a sphere of radius which is much larger than the wavelength, the cross-section for GR wave scattering by the sphere is given by

$$\sigma_{\text{hardsphere}} = 2\pi a^2, \tag{16}$$

where a is the radius of the superconductor.

It may be objected that any kind of reflection of gravitational waves from matter, including any kind of partial reflection, would constitute a kind of "anti-gravity" effect. Now it is certainly true that only positive masses exist in nature, and therefore that the screening of longitudinal gravito-electric fields, like the Earth's gravitational field, including its partial screening by any kind of matter, including superconductors as has been falsely claimed in the so-called "Podkletnov effect", is impossible. However, both positive and negative mass currents can exist in nature, and therefore the screening of transverse gravito-magnetic fields, like those of a gravitational wave, including the partial screening by the reflection of these waves from matter, should indeed be possible. The strength of the reflection can in principle depend on the details of the nature of the matter.

A closely related objection is that, according to one interpretation of the Equivalence Principle, the response of all kinds of matter, whether classical or quantum mechanical in nature, should be universally the same to all gravitational fields. Locally, all dropped objects undergo exactly the same free-fall motion, which is independent of their mass, composition, or thermodynamic state. Specifically, the response of all matter to gravitational fields, such as the Earth's, is independent of whether the thermodynamic state of the matter is classical or quantum mechanical in nature.

Thus this interpretation of the Equivalence Principle would forbid any difference between the linear response of incoherent classical matter, such as Weber bars, and the linear response of coherent quantum mechanical matter, such as superconductors, to any kind of gravitational field. However, whilst this interpretation is valid for the local response of any kind of matter, whether coherent or incoherent, to all gravito-electric fields, like that of all dropped, freely-falling objects in their response to the gravitational field

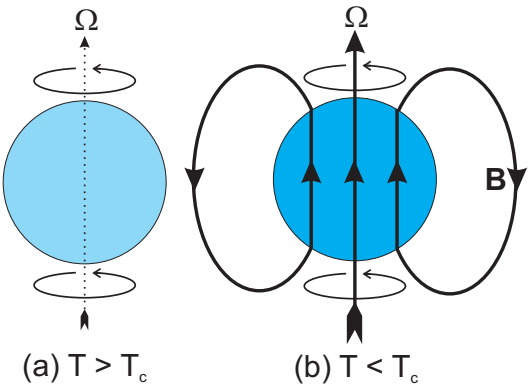


Figure 2: Comparison of a rotating normal metal with a rotating superconductor. In (a), the normal metal is a neutral sphere consisting of a rotating superconducting metal at a temperature above its transition temperature T_c. In (b), this body is slowly cooled down below T_c, whilst it is rotating. Now there is the onset of a magnetic field **B** proportional to its rate of rotation Ω with respect to the distant "fixed stars".

of the Earth, it is not valid for the nonlocal response of matter, whether coherent or incoherent, to gravito-magnetic fields, such as to the Lense-Thirring field arising from the distant matter of the Universe (i.e., from the distant "fixed stars").

Empirical evidence for this latter fact lies in the difference between the response of normal metals and the response of superconductors which are rotating with respect to the distant "fixed stars" [9]. Consider an experiment in which there is a spherical body rotating at a fixed angular speed Ω, which consists of a normal metal that is a superconductor above its transition temperature T_c. For all temperatures $T > T_c$, no magnetic field is produced by this neutral, rotating body.

Now consider what happens when this rotating body is slowly cooled, whilst it is undergoing rotation with respect to the distant "fixed stars," down below its transition temperature T_c. Now for all temperatures $T < T_c$, there exists a magnetic field B produced by this rotating superconducting body in the "London moment" effect. The London moment arises from the constructive quantum interference of the Cooper pairs of electrons in the superconductor near its surface after one round trip around the perimeter of the body, as seen by the distant inertial observer. This quantum interference effect implies that the Cooper pairs everywhere near

the surface of the body in its lowest energy state will come to a complete halt with respect to the distant "fixed stars". The nuclei near the surface, however, continue to rotate with respect to the distant "fixed stars". Thus this differential motion of the Cooper pairs with respect to the nuclei near the surface produces a surface electrical current. It is this current that produces the magnetic field B which is sketched in Figure 2(b).

However, according to the above interpretation of the Equivalence Principle, there cannot be any difference in the response of the rotating body above or below its transition temperature to any kind of gravitational field, including the Lense-Thirring field which arises from the distant "fixed stars". Specifically, the response of the rotating body to the Lense-Thirring field arising from these distant "fixed stars" should be independent of the thermodynamic state of this body, and in particular, it should be independent of whether the rotating body is composed of incoherent, classical matter, or coherent, quantum mechanical matter. This interpretation of the Equivalence Principle is contradicted by experiments that demonstrate the existence of the London moment [9].

Let us now generalize the above considerations to the case of the time-varying, transverse fields of gravitational radiation, when the frequency of the radiation is much lower than the BCS gap frequency, so that the quantum adiabatic theorem holds. Let us first consider the case of a normal metal, which consists of a superconducting metal above its transition temperature T_c. Let this neutral metallic body have the shape of a planar slab with a thickness of half a wavelength of the incident gravitational plane wave, which is propagating towards the slab at normal incidence along the x axis from the left, as depicted in Figure 3. The entrance face of the slab is at $x = 0$, and the exit face of the slab is at $x = \lambda/2$.

The linear response of a normal electron above the superconducting transition temperature Tc near the surface of the metal to the tidal field g of the incident wave is represented by the local, instantaneous velocity vector v, which would be the free-fall velocity of the normal electron, assuming the absence of any dissipation in the metal. (All measurable quantities are those that are being measured by a distant inertial observer represented by the eye in Figure 3). Since the normal electrons undergo local free fall together with the nearby nuclei (neglecting for the moment the weak restoring forces arising from the ionic lattice), no electrical currents are produced in the linear response of this neutral body to the local, instantaneous accelerations due to gravity g arising from tidal forces of the gravitational wave.

Let us now choose two rectangular loops, one above the center of mass (c.m.) of the system, and one below it, as shown in Figure 3, in order to

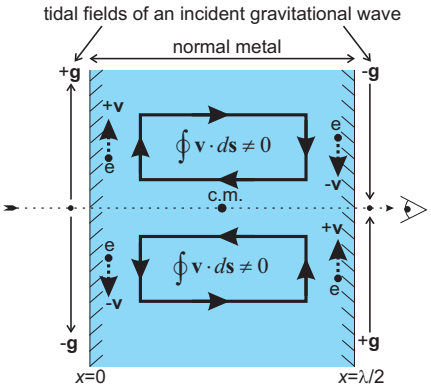


Figure 3: A slab of normal metal half a wavelength thick, which is a superconducting metal above its transition temperature T_c, in its linear response to the tidal fields of a normally incident gravitational plane wave. The center of mass of the slab is represented by the dot below the label "c.m." The eye on the right represents a distant inertial observer. However, below T_c, the upper and lower electrons form a Cooper pair in a Bohm singlet state, that is an entangled state which violates Bell's theorem and is nonlocal. The electrons local free-fall motion is thus suppressed.

evaluate the circulation of the normal electrons around these loops in their linear response to the incident wave. It is obvious upon inspection of Figure 3 that these two closed loop integrals of the velocity v of the normal electron, as seen by the distant inertial observer, obey the inequality

$$\oint \mathbf{v} \cdot \mathrm{d}\mathbf{s} \neq 0, \tag{17}$$

so that the circulation of the normal electrons around these loops does not vanish. This is because there is a reversal of the sign of the tidal forces as the wave propagates from the entrance face of the slab at $x = 0$ to the exit face at $x = \lambda/2$, and therefore there is also a reversal of the sign of the normal electron velocity in their linear response to these tidal forces.

Next, let us slowly cool this system down below the superconducting transition temperature T_c whilst the incident gravitational wave is still present. The normal electrons now become Cooper pairs, which possess quantum phase coherence over long distances. If the Cooper pairs near the surface of the metal in Figure 3 were to undergo free fall with exactly the same instantaneous, local velocity $\mathbf{v}$ in their linear response to the tidal fields g of the incident gravitational wave, as the instantaneous, local free-fall veloc-

ity $\mathbf{v}$ of the normal electrons in the normal state of the metal (as would be demanded by the above erroneous interpretation of the Equivalence Principle), then their macroscopic wavefunction (or the complex order parameter of Ginzburg and Landau) would no longer by single valued after one round trip around the closed loop, and the constructive quantum interference for these Cooper pairs after a round trip would be impossible. This is because the instantaneous velocity $\mathbf{v}$ is related to the phase ϕ of the Cooper pair macroscopic wavefunction (or of the complex order parameter) by

$$\mathbf{v} = \frac{\hbar}{m}\nabla\phi, \tag{18}$$

where $\hbar$ is Planck's constant$/2\pi$ and m is the mass of a Cooper pair.

However, experiments verifying the existence of the London moment [9] demonstrate that the Cooper pairs do in fact always possess a single-valued macroscopic wavefunction (or complex order parameter), as determined by the observations in the inertial frame of the distant observer, i.e., that of the distant "fixed stars". Hence it must be the case that below the temperature T_c, the circulation of the Cooper pairs vanishes identically, i.e.,

$$\oint \mathbf{v}\cdot d\mathbf{s} = 0, \tag{19}$$

in the case of the lowest energy state of the Cooper pairs. This implies that their local, instantaneous velocity vanishes everywhere along the left and the right vertical segments at $x = 0$ and $x = \lambda/2$ of the arbitrarily chosen closed loops in Figure 3, so that in the limit of arbitrarily skinny loops, it follows that

$$\mathbf{v} = 0 \tag{20}$$

everywhere near the surface of the superconductor. Therefore it follows that the time-dependent, transverse gravito-electric field

$$\mathbf{E}_G \equiv \mathbf{g} = \left(\frac{\partial \mathbf{v}}{\partial t}\right)_{\text{near surface}} = 0 \tag{21}$$

also vanishes everywhere near the surface of the superconductor. This also implies hard-wall boundary conditions that lead to a mirror-like reflection of gravitational waves. (See also footnotes 11 and 12 of [3a].)

The above conclusions also follow from the fact that the entire superconducting slab stays adiabatically in the BCS ground state of the system with a single value of the global phase ϕ of the macroscopic wavefunction (or of the complex order parameter) inside the entire superconducting body, i.e.,

$$\phi = \text{constant everywhere.} \tag{22}$$

This is true even in the presence of the weak, local perturbations due to the incident gravitational wave, provided that the wave frequency is well below the BCS gap frequency, so that the wave cannot cause any quantum transitions out of the BCS ground state.

The Cooper pairs in the BCS ground state are in entangled states of momentum and spin, which are nonlocal, so it is impossible to know whether it is the upper or the lower electron in Figure 3 that is moving upwards or downwards in response to the tidal g fields. The large number of these phase-coherent electrons in the system (on the order of Avogadro's number) in the case of superconductors, compensates for the usual weakness of the interaction between gravitation radiation and matter, and leads to a superradiant, mirror-like reflection [3b].

Therefore, just as in the case of the London moment below the transition temperature T_c, the Cooper pairs everywhere near the entrance and exit faces of the superconducting slab must come to a complete halt with respect to the distant inertial observer (i.e., with respect to the distant "fixed stars") below Tc. The vanishing of v everywhere near the surfaces of the slab must be independent of the size of the small amplitude of a sufficiently weak gravitational wave, in the regime of the linear response of the superconductor to the wave. This is the behavior of an extremely rigid material in its linear response to the gravitational wave, so that, once again, one concludes that a mirror-like reflection of a sufficiently weak incident gravitational wave should occur at the planar superconductor-vacuum interface. The resulting scattering cross-section for large superconducting bodies should therefore once again be of the order of magnitude of the geometric cross-section given by Equation (16). These counter-intuitive predictions will be tested in the above proposed experiment.

It is to be emphasized that throughout the above discussion, all measurable quantities in the above equations are those which are being observed and measured by the distant inertial observer. The coordinate system being used in these measurements is the one set up by means of light signals sent to and from this distant observer, and the time coordinate being used is the one set up by means of this observer's clock [11].

This experiment could lead to important applications in science and engineering. In science, it would open up the possibility of gravitational-wave astronomy at microwave frequencies. One important problem to explore would be observations of the analog of the Cosmic Microwave Background (CMB) in gravitational radiation. Since the Universe is much more transparent to gravitational waves than to electromagnetic waves, such observations would allow a much more penetrating look into the extremely

early Big Bang towards the Planck scale of time, than the presently well-studied CMB. Different cosmological models of the very early Universe give widely differing predictions of the spectrum of this penetrating radiation, so that by measurements of the spectrum, one could tell which model, if any, is close to the truth [10]. The anisotropy in this radiation would also be very important to observe.

In engineering, it could open up the possibility of intercontinental communications by means of microwave-frequency gravitational waves directly through the interior of the Earth, which is transparent to such waves. This would eliminate the need of communications satellites, and would allow an economical means of communication with people deep underground or underwater in submarines in the oceans. Such a new direction of gravitational-wave engineering could aptly be called "gravity radio" [3].

Acknowledgments

This work was supported in part by the CTNS-STARS program. I would like to thank Prof. Thomas Kibble for his helpful comments.

References

[1] M. Tinkham, *Introduction to Superconductivity*, 2nd edition (Dover Books on Physics, New York, 2004).

[2] C.W. Misner, K.S. Thorne, J.A. Wheeler, *Gravitation* (Freeman, San Francisco, 1972).

[3] R.Y. Chiao, *New directions for gravity-wave physics via Millikan oil drops*, to be published in Visions of Discovery (a Volume in honor of C.H. Townes) edited by R.Y. Chiao, W.D. Phillips, A.J. Leggett, M.L. Cohen and C.L. Harper Jr. (Cambridge University Press, Cambridge, 2007), [arχiv:gr-qc/0610146v16]; R.Y. Chiao, *The Interface between Quantum Mechanics and General Relativity*, Lamb Medal Lecture, J. Mod. Optics **53** (2007) 2349, [arχiv:quant-ph/0601193v7].

[4] L. Schiff, *Quantum Mechanics* (McGraw-Hill, New York, 1955).

[5] C. Kittel, *Introduction to Solid State Physics*, 8th edition (John Wiley & Sons, New York, 2005).

[6] R.Y. Chiao and W.J. Fitelson, *Time and Matter in the Interaction between Gravity and Quantum Fluids: Are There Macroscopic Quantum Transducers between Gravitational and Electromagnetic Waves?*, Proceedings of the Time and Matter Conference, Venice, Italy, 11-17 August 2002, eds. Ikaros Bigi and Martin Faessler (World Scientific, Singapore, 2006), p. 85, [arχiv:gr-qc/0303089].

[7] R.M. Wald, *General Relativity* (University of Chicago Press, Chicago, 1984), section 4.4, see also footnote 6 of [3a].

[8] C. Kiefer and C. Weber, *On the Interaction of Mesoscopic Quantum Systems with Gravity*, Annalen der Physik (Leipzig) **14** (2005) 253.
[9] J. Tate, B. Cabrera, S.B. Felch and J.T. Anderson, *Precise Determination of the Cooper-Pair Mass*, Phys. Rev. Lett. **62** (1989) 845; *Determination of the Cooper-Pair Mass in Niobium*, Phys. Rev. B **42** (1990) 7885.
[10] R.Y. Chiao, *Proposed Observations of Gravity Waves from the Early Universe via "Millikan Oil Drops"*, Proceedings of the NASA Conference Quantum to Cosmos, [arχiv:gr-qc/0606118], to be published in Int. J. Mod. Phys. D.
[11] J.L. Synge, *Relativity: The General Theory* (North-Holland Publishing Company, Amsterdam, 1960).

Draft version August 15, 2013
Preprint typeset using LaTeX style emulateapj v. 12/16/11

PULSAR GLITCHES AS A GRAVITATIONAL WAVE SOURCE

Corey Bathurst
University of South Florida
University of Florida IREU
Location:Seoul National University
Draft version August 15, 2013

ABSTRACT

This report describes research completed in in Dr. Hyung Mok Lee's group at Seoul National University through the University of Florida's International Research Experience for Undergraduates funded by the National Science Foundation. This project explores the possibility of pulsar glitches as a source of gravitational radiation through hydrodynamic code with applied perturbations that mimic the glitch event. The two perturbations considered mimic the star quake model and angular momentum transfer by vortex unpinning at crust-core interface model. We estimate the characteristic amplitude of the gravitational waves from the evolution of the quadrapole moment of the simulated star.

We use the weak field metric with a generalized source term to Poisson's equation to gain closer agreement with a full general relativistic treatment. This report discusses the validity of this method and some of the testing done to see how well our method agrees with other methods such as linear perturbative analysis. We find that our pseudo-Newtonian method gives better agreement with general relativity than similar methods without a generalized source term.

We find that current detectors are unable to reach the sensitivity to detect gravitational radiation from a glitch phenomena. For both perturbations, strain amplitudes were approximately 10^{-25}. We find that the vortex unpinning model excite the inertial mode more effectively than the star quake model.

1. INTRODUCTION

Neutron stars offer insight into physical states not achievable in any earth based laboratory today. Observations of these exotic stars allow for testing of many modern theories in physics. Unfortunately, their size, distance, and luminosities make it difficult to observe them under most conditions. Fortunately, we can observe neutron stars as pulsars. Neutron stars emit directed radiation along their magnetic axis and, if they are appropriately aligned with Earth, we can observe this radiation as regular pulses due to the rotation of the star.

Pulsars rotate at exceptionally regular rates but are constantly slowing down as they transfer energy to their environment. On occasion, though, pulsars are observed to suddenly speed up without explanation. This increase in rotational velocity if know as a glitch. Since the first pulsar glitch was observed over 100 glitches have been found. Typically, the fractional spin up rates ($\Delta\Omega/\Omega$) of a glitch range from 10^{-9} to 10^{-6} while the largest glitch ever observed had an amplitude of $\sim 3.3\times10^{-5}$(Manchester & Hobbs 2011). Glitches are the result of a sudden internal change of the state of the star and therefore give a unique look into the state of the star. More interestingly, a glitch should excite oscillations within the star which have the potential to emit gravitational radiation.

Observations of gravitational radiation from a glitch would have many benefits scientifically including the possibility of being the very first detection of gravitational radiation. In addition, the state of the star is highly dependent upon the equation of state (EoS) which is largely unknown for extremely dense regions such as those of a neutron star. Therefore, observations of gravitational radiation from pulsar glitches could offer further constraint on the candidates for a neutron star's EoS.

As previously mentioned, the exact mechanism of a pulsar glitch is still widely debated. Currently there exists two major hypotheses. The first is known as the star quake model (Ruderman 1969). In these models the star is treated as two regions, the core and the crust. As the neutron star slows down it adapts a more spherical shape. This model assumes the crust is rigid which causes it to maintain its ellipsoidal shape. The core on the other hand is adapting a spherical shape and is applying pressures on the crust to do the same. When the pressures become great enough, the crust breaks up to become more spherical and the moment of inertia of the star suddenly changes. It is this sudden change in the moment of inertia that leads to the observed change in rotational rate. The second model is vortex unpinning of a superfluid at the crust-core interface (Packard 1972; Anderson and Itoh 1975). The rate at which pulsars slow down is much lower than expected for a rotating normal fluid. For this reason, it is believed that the core of the star is a superfluid. In this model the core is considered a superfluid while the crust is considered a normal fluid. This leads to differential rotation between the core and the crust. Once the difference becomes too great, vortices coupling the core and crust unpin, causing the core to transfer angular momentum to the crust. This transfer of momentum to the crust is the cause of the observed spin up.

The goal of this project is to determine the modes and amplitudes of gravitational radiation resulting from a glitch event. The normal modes of a star can be predicted by linearized methods but these methods cannot predict

which modes are excited nor their amplitudes. Therefore the only realizable method for determining this information is through hydrodynamic simulations. In our simulations we create an equilibrium model of the star, apply a perturbation to mimic the glitch model, then follow the hydrodynamic evolution of the star. We use glitch perturbations that are much larger than real glitch magnitudes because of numerical limitations but use this data to extrapolate to real glitch magnitudes. From this information we can infer the maximum values for observed gravitational waves and the frequencies at which they occur.

This report will give some background on gravitational waves and their detection, information on what we know about pulsar glitches and the popular models for them, discuss the formulation used within our simulations and their validity, and finally predict the characteristic strain amplitudes of gravitational waves we might hope to one day see.

2. GRAVITATIONAL WAVES

In Einstein's theory of general relativity, gravity is considered to be the result of the curvature of spacetime due to mass and energy. Since the curvature is dependent on mass and energy any change in the location or density of mass and energy will result in a change in the curvature of spacetime. As one would expect, these changes propagate out from the source at the speed of light. This fact hints to the existence of gravitational waves. Just like electromagnetic radiation, gravitational waves depend on the acceleration of massive object, but, as gravitation is a much weaker force than the electromagnetic force, gravitational waves have very small amplitudes making them difficult to detect. In addition, due to the conservation of momentum, there can be no mass dipole radiation for gravitation unlike electromagnetic radiation. This means that not only does an object have to be very massive, it cannot be changing symmetrically for it to emit gravitational radiation. In other words, gravitational radiation with any chance of measurement must come from non-symmetric, massive systems. Sources commonly suggested for gravitational wave generation are binary systems consisting of neutron stars or black holes.

2.1. *Detection of Gravitational Radiation*

Since a gravitational wave is an actual fluctuation in the curvature of spacetime, when a wave passes through a portion of space, distances between two locations will appear to rhythmically increase and decrease with amplitudes depending on the gravitational wave. Therefore, to detect gravitational waves, one must detect these fluctuations in distance.

Gravitational radiation detectors must measure very precisely the distance between two points in space. Current detectors are all in the form of a laser interferometer. These interferometers take the form of a Michelson interferometer which splits the incoming laser into two different beams along orthogonal paths that reflect the light back to the splitting point and recombines the beams on the path to the detector. Depending on the path lengths of the split beams the detector will either see constructive or destructive interference. It is usually setup such that a dark fringe is registered at the detector normally, and if a gravitational wave were to pass through the detector, it would fluctuate the path length of the detector arms and break the destructive interference of light at the detector.

In addition, this interferometric setup is better utilized if it is placed in multiple locations. By having the similar systems in different locations it allows for a few different benefits. Firstly, it allows to confirm that a gravitational wave signal is real and not just noise in the system. Since the systems are far enough apart it would be uncommon for them to both be effected by the same transient noise signal at the same time. Also, by having largely seperated locations, the interferometers can time the signals when they start and stop. Due to their spatial seperation, the interferometers should observe the signals at slightly different times and allow for localization of the source of the signal.

In concept is a very simple setup but there are many obstacles that an interferometer must overcome to be able to detect the extremely weak spacetime fluctuations caused by a gravitational wave. The sensitivity of a detector must be so great that every source of noise must be minimized and accounted for if one is to properly detect gravitational waves. These sources of noise include things such as power fluctuation of the laser source, seismic noise, photon pressure on the mirror at the end of each arm, thermal noise, and an incredible number of other sources. Locating and removing these sources of noise is the focus of much gravitational wave detectors today. In fact, one the largest gravitational wave detector project today, LIGO, was commissioned under the assumption that it would be very unlikely for it to detect gravitational waves in its initial state (LIGO). Currently, LIGO is an upgrade phase to Advanced LIGO where scientists believe it will likely be able to detect some of the strongest sources of gravitational radiation. During this pivotal point in time, it is important to consider all possible sources of gravitational radiation, what their signals would look like, and how likely it would be to these signals with current detectors.

3. FORMULATION

The overall goal of our hydrodynamic simulations is to measure how the quadrapole moment of a neutron star evolves after a glitch. Our formalism uses the weak-field metric meaning all special relativistic effects are taken into account while general relativistic effects are ignored. This is done because our simulations evolve gravitation with each time step and solving Einstein's equation that many times would be too computationally expensive. Instead, to increase

our agreement with a full general relativistic solution, we used a generalized source term to Poisson's equation which takes into account all forms of energy instead of just rest mass density. The following outline of our formulation follows that presented in Kim et. al 2009 and Kim et. al 2012.

Our formalism uses the weak field metric in cylindrical coordinates (R,Z,Φ),

$$\begin{aligned} ds^2 &= g_{\mu\nu}dx^\mu dx^\nu \\ &= -(1+2\Phi)dt^2 + (1+2\Phi)^{-1}\delta_{ij}dx^i dx^j \\ &= -(1+2\Phi)dt^2 + \frac{1}{1+2\Phi}(dR^2 + dZ^2 + R^2 d\phi^2) \end{aligned} \tag{1}$$

As mentioned before, we take a single Newtonian gravitational potential:

$$\nabla^2\Phi = 4\pi\rho_{\text{active}}. \tag{2}$$

Where ρ_{active} is known as the active mass density which was first conceived by Tolman (cite) with the form:

$$\rho_{\text{active}} = T - 2T_0^0 \tag{3}$$

Where T is the trace of the stress-energy tensor. In a full general relativistic treatment, the stress-energy tensor describes the density of mass and energy. For this reason it can be considered the source term for gravity in general relativity. By bringing the stress-energy tensor of a perfect fluid into our formalism and treating all forms of energy as the source to our Poisson's equation, we get significantly better agreement with a full general relativistic treatment than just rest mass density as our source term.

As is standard in most numerical simulations we assume our star to be a perfect fluid. This is both for simplicity as viscosity and other effects of a normal fluid becomes too complex as well as the fact that stress energy tensor for a perfect fluid is easily computed:

$$T^{\mu\nu} = \rho_0 h u^\mu u^\nu + P g^{\mu\nu}, \tag{4}$$

where h is the specific enthalpy, u^μ is the four velocity with respect to a Eulerian observer, and P is the pressure. The specific enthalpy takes the form:

$$h = 1 + \epsilon + \frac{P}{\rho_0}, \tag{5}$$

where ϵ is the internal energy. Using this in combination with the stress energy tensor for a perfect fluid we can find the active mass density to be given by:

$$\begin{aligned} \rho_{\text{active}} &= T - 2T_0^0 \\ &= T_i^i - T_0^0 \\ &= \rho_0 h\frac{1+v^2}{1-v^2} + 2P. \end{aligned} \tag{6}$$

We enforce axisymmetry as well as equatorial symmetry. We define the conservative variables (q) and primitive variables (w) as

$$q = \begin{pmatrix} D \\ S_R \\ S_Z \\ S_\phi \\ \tau \end{pmatrix} = \begin{pmatrix} \rho_0 W \\ \rho_0 h W^2 v_R \\ \rho_0 h W^2 v_Z \\ \rho_0 h W^2 v_\phi \\ \rho_0 h W^2 - P - D \end{pmatrix}, \tag{7}$$

$$w = \begin{pmatrix} \rho_0 \\ v^R \\ v^Z \\ v^\phi \\ P \end{pmatrix}, \tag{8}$$

Which just stand for the conservation of mass and angular momentum. Then we can obtain the hydrodynamic equations:

$$\frac{\partial\left(\sqrt{\gamma}q\right)}{\partial t} + \frac{\partial\left(\sqrt{-g}f^R\right)}{\partial R} + \frac{\partial\left(\sqrt{-g}f^Z\right)}{\partial Z} = \sqrt{-g}\Sigma \tag{9}$$

where

$$f^R = \begin{bmatrix} Dv^R \\ S_R v^R + P \\ S_Z v^R \\ S_\phi v^R \\ \tau v^R + P v^R \end{bmatrix} \tag{10}$$

$$f^Z = \begin{bmatrix} Dv^Z \\ S_R v^Z \\ S_Z v^Z + P \\ S_\phi v^Z \\ \tau v^Z + P v^Z \end{bmatrix} \tag{11}$$

$$\Sigma = \begin{bmatrix} 0 \\ -\frac{\rho_{\rm active}}{1+2\Phi}\frac{\partial \Phi}{\partial R} + \frac{S_\phi v^\phi}{R} + \frac{P}{R} \\ -\frac{\rho_{\rm active}}{1+2\Phi}\frac{\partial \Phi}{\partial Z} \\ 0 \\ -\left(S_R \frac{\partial \Phi}{\partial R} + S_Z \frac{\partial \Phi}{\partial Z}\right) \end{bmatrix}. \tag{12}$$

Finally, the gravitational potential in cylindrical coordinates is

$$\frac{1}{R}\frac{\partial}{\partial R}\left(R\frac{\partial \Phi}{\partial R}\right) + \frac{\partial^2 \Phi}{\partial Z^2} = 4\pi \rho_{\rm active}. \tag{13}$$

These hydrodynamic formulas are used to evolve the star after the glitch perturbation using a finite volume where each grid enforces the conservation laws. The follow sections discuss how the initial model of the star is generated and the perturbations to the initial model used to mimic a glitch event.

4. INITIAL MODEL

4.1. *Equilibrium Models of Rotating Stars*

For our simulations we use the Hachisu Self Consistent Field method (HSCF) to generate our initial star (Hachisu 1989a). The HSCF method uses an integral form to the hydrostatic equation instead of directly solving the differential equations. This is more computationally efficient and generates stable star systems. The special feature of Hachisu's method is that instead of using a rotational velocity parameter as a condition to generate the star it uses an axis ratio (r_p/r_e) to define the star. This allows one to define the star by its elliptical shape caused by the rotation. This is a more stable way to calculate a stars equilibrium state than just using a velocity which can be more cumbersome as well as fail in generating a stable star if the velocity condition is insufficient.

In addition to this, to complete our formulation for hydrostatic equilibrium we need a relation for the pressure of the star (the equation of state). The equation of state of a neutron star is a highly sought after quantity in physics. The densities of a neutron star cannot be found anywhere in the universe and are not achievable in any laboratory today. Discovering the equation of state of a neutron star would be a huge achievement in the astronomical physics as well as nuclear physics. Today, most scientists use a piecewise polytropic equation of state while enforcing continuity at the boundaries to model the equation of state of a neutron star. This is primarily because a polytropic equation of state can be tuned to produce stars of the proper mass and radius as a neutron star. In our simulations we assume a simple polytropic equation of state:

$$P = \kappa \rho^{1+\frac{1}{N}} \tag{14}$$

where κ and N are the polytropic coefficient and index respectively. For our simulations we use $\kappa = 100$ and $N = 2$ because they produce stars of the appropriate mass and radius.

In addition to our assumption of a simple polytropic equation of state, we assume that the rotation to be a rigid one. In other words, our stars are not rotating differentially. This is valid because neutron stars must turn to uniform rotation relatively quickly after birth due to shear viscosity and magnetic tensions in the star. Using these two assumptions, the integral form of our equilibrium stars becomes

$$\ln\left[(N+1)\kappa\rho_0^{1/N} + 1\right] + \frac{1}{2}\ln(1+2\Phi) + \frac{1}{2}\ln(1-v^2) = C \tag{15}$$

where C is the integration constant to be found based on boundary conditions. A more complete treatment of the methods used to generate our initial stars can be found in Kim et. al 2009 (Kim et al. 2009).

This report takes a case study of a neutron star with the largest mass ever observed, $2.0M_\odot$, at a distance of 1kpc. Since this represents some of the best known conditions for a pulsar in which glitch induced gravitational radiation, the magnitudes of the characteristic strain amplitudes in this report are higher than would be observed from most known pulsars.

5. GLITCH PERTURBATIONS

Pulsars usually slow down at a very slow and predictable pace as they transfer energy to their environment but occasionally without explanation spin up. These spin ups are known as a glitch and are not at all well understood. The first pulsar glitches were observed in 1969, glitches have been observed in over 30 different pulsars. Even 40 years later, the exact mechanism that causes these events is still under debate. Because it is known that a pulsar glitch is caused by internal mechanism, proving the exact mechanism by which a glitch occurs would give valuable information about the state and properties of pulsars. The following section discusses two major glitch mechanism models: the star quake model and the vortex unpinning at crust-core interface model.

The following subsections discuss the two glitch mechanisms we consider in our simulations and the perturbations we apply to our equilibrium star to model these mechanisms. For both perturbations we consider the effective crust depth to be 10% the radius of the star and ignore all effects of tension and other complex properties the star might have. In addition we hold the conservation laws, the conservation of baryonic mass (M_0) as well as total angular momentum (J):

$$\begin{aligned} M_0 &= \int \rho_0 W dV \\ &= 2\pi \int \frac{\rho_0 W}{(1+2\Phi)^{3/2}} R dR dZ \end{aligned} \tag{16}$$

$$\begin{aligned} J &= \int S_\phi dV \\ &= 2\pi \int \frac{\rho_0 W^2 R^2 \Omega}{(1+2\Phi)^{3/2}} R dR dZ \end{aligned} \tag{17}$$

5.1. *Star Quake Model*

After the first pulsar glitch was observed in 1969 Ruderman developed a mechanism by which a glitch could occur due to a change in the moment of inertia of the star. This model assumes the star to be two components, the crust and the core. The core is assumed to be fluid while the crust has some rigidity to it. As the pulsar slows down it begins to adapt a more spherical shape but since the crust is rigid it resists this change. While this occurs the core is becoming more spherical and is placing stress on the crust. Once the stresses become strong enough the crust is forced to adapt to the core and this causes a change in the moment of inertia of the star. This change in moment of inertia, due to the conversation of angular momentum, causes the star to spin up and we observe a glitch.

This model has been widely explored and it has been found that even just a vertical surface motion of more than a centimeter would generate a glitch magnitude of 10−6. However this model has problems explaining pulsars with multiple glitches in short time periods. It was found by Baym and Pines that a significant amount of energy must be transferred per glitch (cite). So much that it would not be possible for a pulsar to build up enough energy for a glitches to occur in a single pulsar as have been observed. Nevertheless, we consider this mechanism for one of our perturbation for comparison with our second mechanism, vortex unpinning at the crust-core interface.

The above figure shows the perturbation we apply for the star quake model. The stellar surface is shrunken by the asymmetric gaussian formula,

$$\frac{\delta\rho_0}{\rho_0} = \begin{cases} \lambda\ \mathrm{e}^{-\left(\frac{r-c}{\sigma_1}\right)^2} & \text{if } r \le c \\ (1+\lambda)\ \mathrm{e}^{-\left(\frac{r-c}{\sigma_2}\right)^2} - 1 & \text{if } r > c, \end{cases} \tag{18}$$

where $\delta\rho_0$ is the perturbation to the rest mass density, and c is the center of the Gaussian profile, which is not necessarily the crust-core interface. σ_1 and σ_2 represent the width of the Gaussian and are chosen from the conservation of mass. λ is the perturbation magnitude that we apply to the system and should not be confused with the observed glitch magnitude we are simulating. The perturbation to the angular velocity follows from this perturbation by the conservation of angular momentum. From the resulting angular momentum perturbation we measure the simulated glitch magnitude.

5.2. *Vortex Unpinning at the Crust-Core Interface Model*

As the problems with the star quake model were realized, Packard (Packard 1972) and soon after Anderson and Itoh (Anderson & Itoh 1975). Considered if the superfluidity of the core could be responsible for the glitch mechanism. As previously mentioned, the interior of a neutron star is considered to be a superfluid due to the slow rate at which they spin down. Because of this, if the core is considered a superfluid while the crust is a normal fluid, the core can be considered an angular momentum reservoir. The basic consideration of this model is that vortices at the crust-core interface couple the crust and the core. As the star spins down, the crust spins down faster than the core and differential rotation develops. At the point when this difference in velocity becomes too great, the vortices unpin and results in the core transferring momentum to the crust. This transfer causes the observed spin up during a pulsar glitch. This theory is generally the most accepted today but has yet to be set in stone.

The second perturbation, for the vortex unpinning model, gives only a change in rotational velocity while enforcing no perturbation in the mass density:

$$\begin{aligned} \frac{\delta\rho_0}{\rho_0} &= 0, \\ \frac{\delta\Omega}{\Omega} &= \begin{cases} -\lambda_J \ \mathrm{e}^{-\left(\frac{r-c}{\sigma}\right)^2} & \text{if } r \leq c \\ \lambda & \text{if } r > c \end{cases} \end{aligned} \tag{19}$$

In this case, λ is both the perturbation and glitch magnitude since we directly perturb the angular velocity. While λ_J is how much the angular momentum of the interior is reduced, and σ is a fitting parameter. Both values are determined by the conservation laws.

6. RADIAL PULSATION MODE TEST

6.1. *Stellar Pulsation*

To test the validity of our formulation and simulation methods we preformed a radial pulsation mode test. In theory, if our stars are in equilibrium, we should not observe any oscillations within the star. Yet, as with all numerical simulations, there are intrinsic errors that cause the star to oscillate including our resolution, floor values, truncation and limitations in describing the stellar surface (Kim et al. 2012). For this test we generate a stationary (non-rotating), spherical star. All stars have normal modes of oscillation depending on their state and our goal in this test is to compare the radial modes measured in our simulations to another method for generating the normal mode frequencies, linear perturbative analysis.

Although numerical simulations will cause oscillations within our modeled star, we apply a radial perturbation to our system to further excite these modes,

$$\delta\rho_0 = B \sin\left(\pi \frac{r}{r_s}\right). \tag{20}$$

This radial perturbation disturbs the steady, equilibrium state of the star and causes the star to oscillate when we evolve it dynamically. We measure the oscillation of the density in the star by measuring the central density of the star at each time step. These oscillation are usually complex and a superposition of the fundamental mode and its overtones. To analyze which modes are excited we use the standard of all oscillation analysis, Fourier analysis.

6.2. *Fast Fourier Transform*

For our discrete Fourier transform calculation we utilize the Fast Fourier Transform of the West algorithm package(Frigo & Johnson 2005). In addition, we use the zero-padding method to increase our sampling in the frequency domain. This adds no new real information to the transform but merely decreases the sample spacing in the frequency domain so that actual maximums are not missed due to insufficient sampling.

In addition to problems due to sampling we also deal with spectral leakage in the frequency domain by multiplying the time domain series by a window function. Spectral leakage can cause significant error in amplitude and frequency of maximum. In addition, weaker components can be complete washed out by spectral leakage. Fast Fourier transform algorithms assume the time domain sample is periodic in natural. Discontinuities in the time domain require many more sinusoidal frequencies to represent such complex features. These discontinuities resulting from the non-periodic nature of a time domain sample are the source of spectral leakage in the frequency domain. To reduce these errors we apply a window function which suppresses the magnitudes of the discontinuities thereby suppressing the spectral leakage in the frequency domain. Specifically we use the Hamming window function:

$$w_j = 0.54 - 0.46 \cos\left(\frac{2\pi j}{N}\right) \tag{21}$$

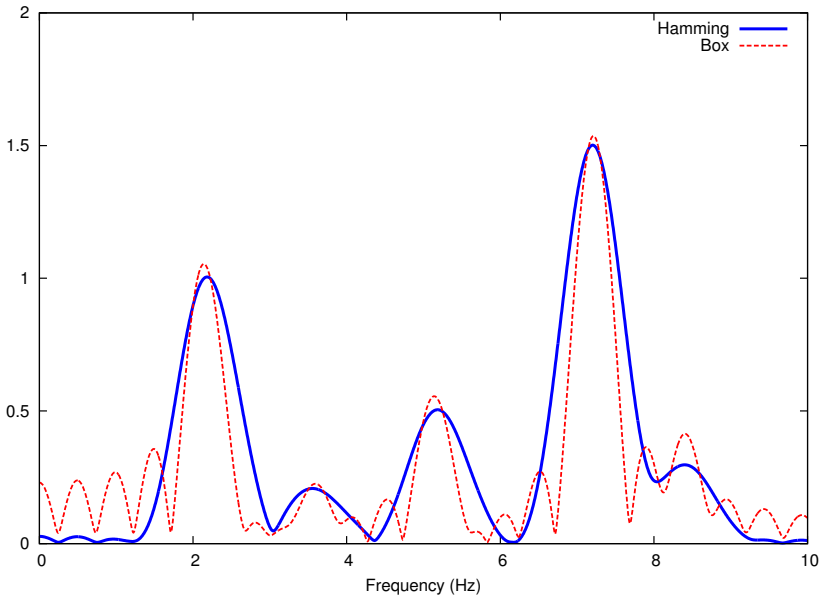


FIG. 1.— This figure shows the effects of a window function on a linear combination of simple sinusoidal waves. The box window is the Fourier transform without any window function. The Hamming window is with a window function modifying the time domain series. It is clear how effectively a window function reduces the error.

Figure one shows the effects of adding a window function to our time domain sample. It is important to note the addition of a window function decreases the overall spectral resolution of sample. So if two modes are located too close to each other, it may be difficult to accurately identify their amplitudes and frequencies.

6.3. *Radial Mode Test Results*

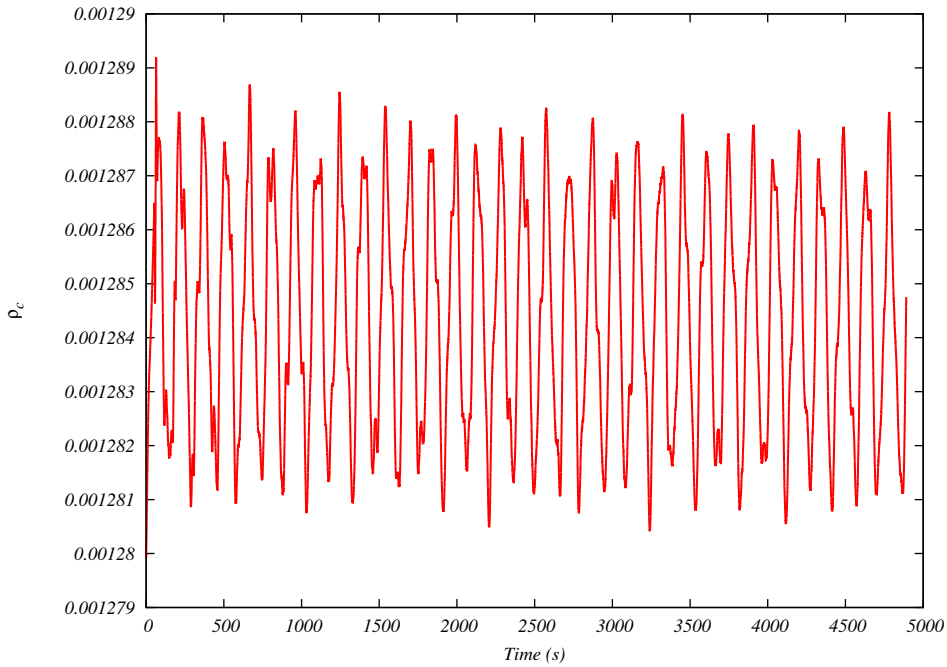


FIG. 2.— The above figure shows the oscillations of the central density of our simulated star in time.

The above figure shows the evolution of the central density of our simulated star in time. The simulation was ran for 100000 time steps to give a strong sample. The central density appears stable over the extent of the run. Below is the Fourier transform of the above central density.

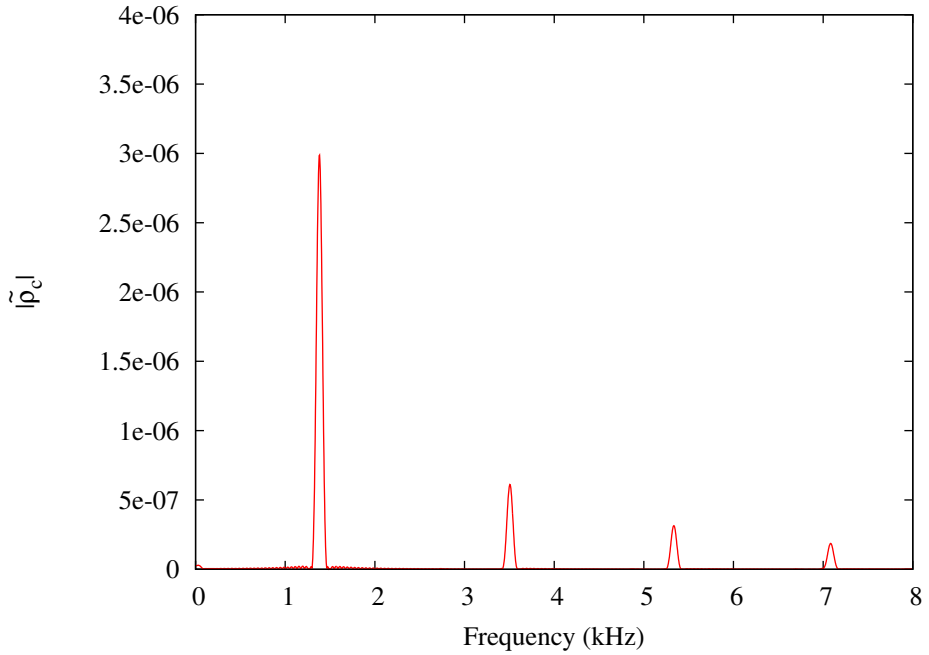


FIG. 3.— The above figure is the Fourier transform of our central density time sample. The transform clearly shows four excited radial modes, the fundamental and its three overtones.

	F	H_1	H_2	H_3
Simulations	1.382	3.505	5.330	7.083
σ	1.383	3.506	5.334	7.083
% Difference	0.07	0.03	0.07	0

The Fourier transform clearly shows four excited modes. We use the change in central density as opposed to the absolute central density so the modes stand out compared to low frequency components. We use a secondary code which utilizes linear perturbative analysis to compare to our simulations results. We find that our results agree very accurately with the linear analysis code.

7. MODE ANALYSIS

The goal of these simulations is to find the modes of gravitational radiation excited by the glitch event. Gravitational waves emitted by the star are related to the quadrapole moment of the star by

$$h_{ij}^{TT} = \frac{d^2 I_{kl}}{dt^2} P_{ik} P_{jl}. \tag{22}$$

Direct differentiation of the quadrapole moment introduces significant error. Instead, we take the Fourier transform of the quadrapole moment which allows us to find the amplitude and frequency of the gravitational radiation at the same time. In this form, the maximum characteristic strain amplitude is

$$h_c = \frac{64\pi^2}{r} f_c^2 \left| \tilde{I}_{xx}(f_c) \right| \tag{23}$$

Therefore, by taking the Fourier transform of quadrapole moment and multiplying by the mode frequency squared we can find the effective amplitude of the gravitational wave modes.

Because real glitch magnitudes are much too small to properly simulate, we run our simulations for many perturbation magnitudes on the same initial model. From this we apply a linear fit to each mode and extrapolate back to real glitch magnitudes. The below plots show the perturbation data used extrapolate back to real magnitudes.

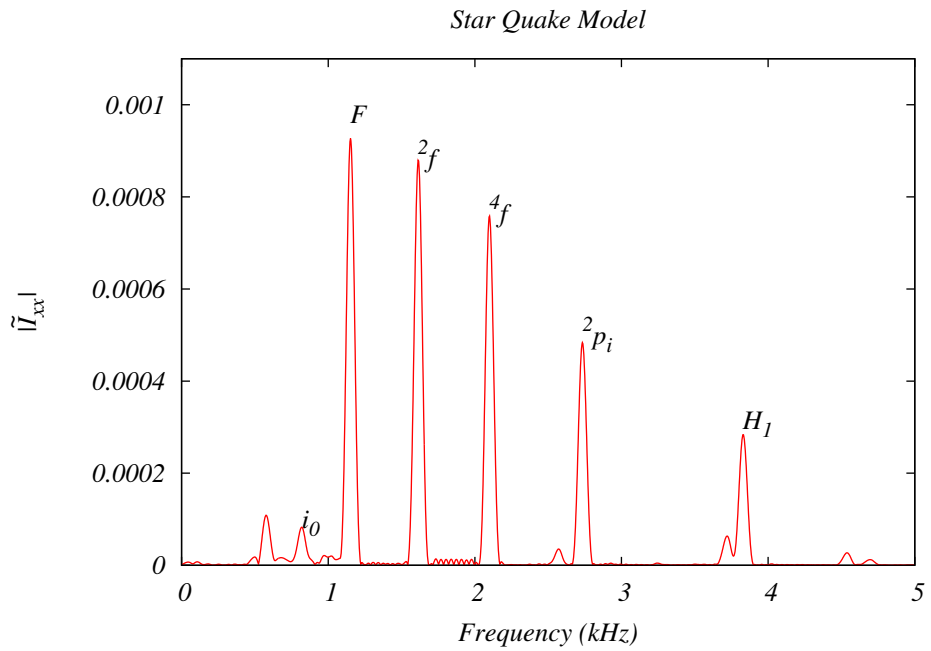


FIG. 4.— This plot is the Fourier transform of the quadrapole moment for a $2.0M_{\odot}$star for perturbation one with λ=0.001. The figure shows the modes excited. Although F and H_1 are radial modes, due to the nonspherical shape of the star they have a large quadrapole moment component.

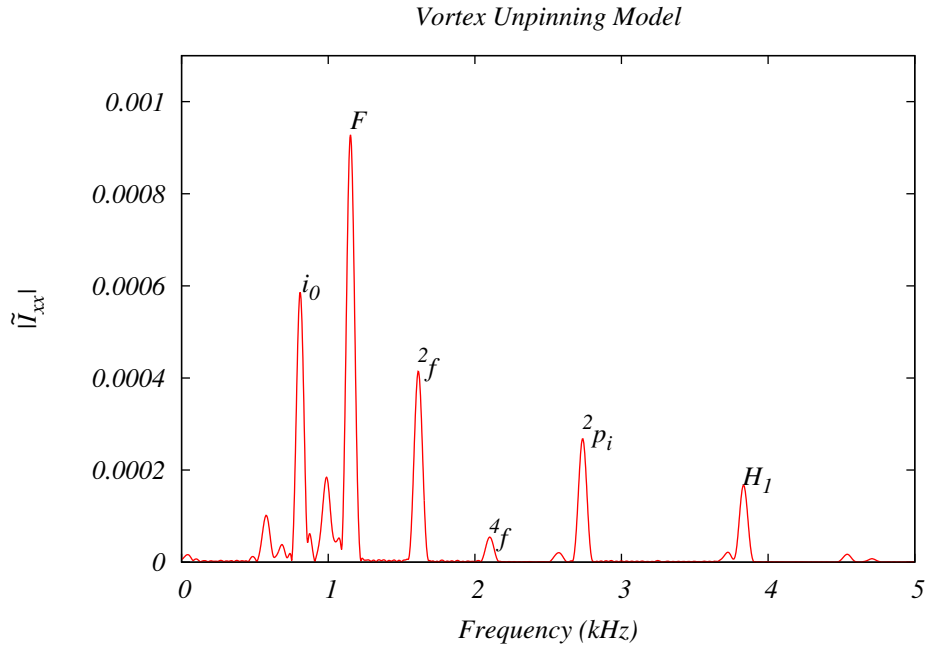


FIG. 5.— This plot is the Fourier transform of the quadrapole moment for a $2.0M_{\odot}$star for perturbation two with λ=0.001. The figure shows the modes excited.An interesting feature of this perturbation is that it highly excites the inertia mode of the star. This mode is known to have a long dampening time which can help increase the detect-ability of the modes.

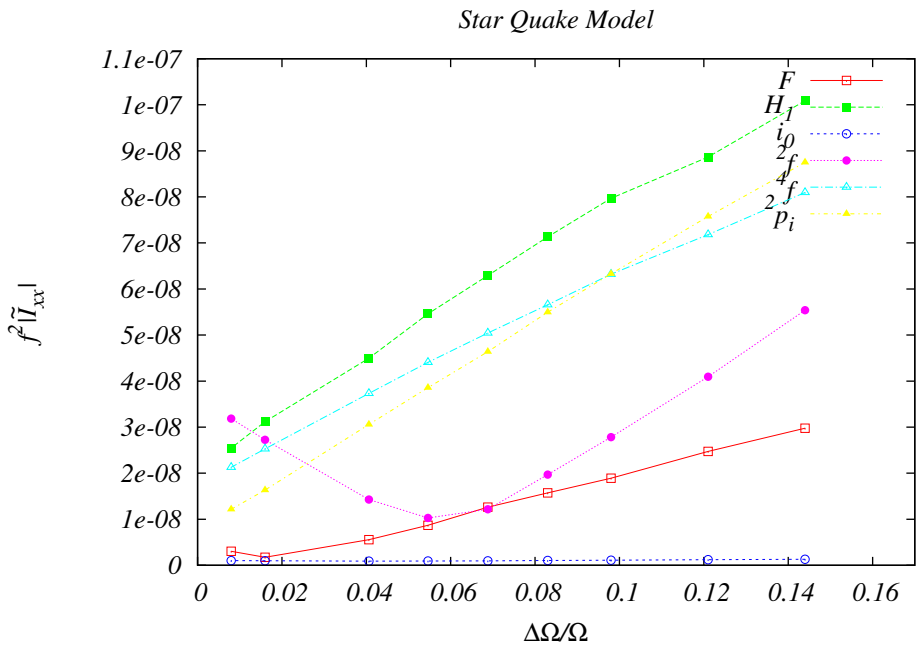


FIG. 6.— This figure shows the effective gravitational wave magnitudes for perturbation one. This data was used to extrapolate back to real glitch magnitudes. The nonphysical behavior exhibited by the ^{2}f mode is most likely caused by numerical errors.

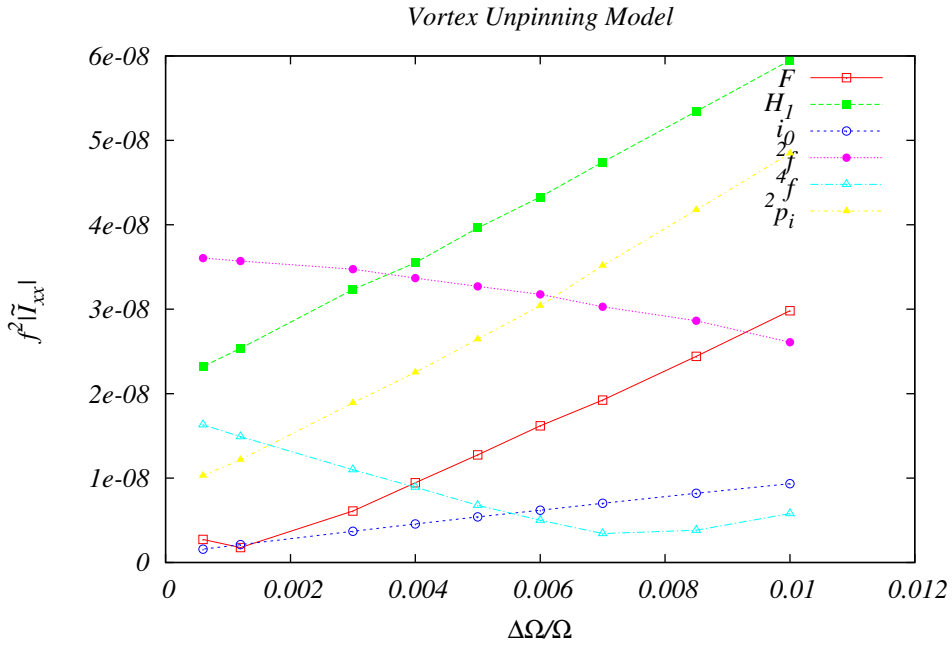


FIG. 7.— This figure shows the effective gravitational wave magnitudes for perturbation two for various glitch magnitudes. We use this data to extrapolate back to realistic glitch magnitudes. The nonphysical behavior of the ^{2}f and ^{4}f mode are currently under investigation.

8. DISCUSSION

	Frequency (kHz)	S.Q.	V.U.
F	1.152	1.11×10^{-27}	1.44×10^{-26}
H_1	3.828	1.10×10^{-24}	1.00×10^{-24}
i_0	0.808	4.20×10^{-26}	6.04×10^{-26}
^2f	1.613	6.58×10^{-25}	1.78×10^{-24}
^4f	2.100	9.12×10^{-25}	7.14×10^{-25}
^2p_i	2.734	3.78×10^{-25}	3.48×10^{-25}

TABLE 1
THE ABOVE TABLE SHOWS THE SUMMARY OF THE FINAL RESULTS FOR THE MAXIMUM CHARACTERISTIC STRAIN AMPLITUDE OF EACH MODE EXCITED BY THE DIFFERENT GLITCH MODELS. THESE VALUES ASSUME A GLITCH AMPLITUDE OF 10^{-5}, A NEUTRON STAR MASS OF $2.0M_\odot$, AND A DISTANCE OF 1 KPC.

Currently, detecting gravitational radiation from a pulsar glitch is very unlikely. The strongest strain amplitude calculated is located in a less sensitive frequency band as well as being an order of magnitude lower than the most sensitive band range of advanced LIGO. Future detectors such as the Einstein Telescope may reach the sensitivity necessary to measure such waves. It is important to reiterate that these strain values are for very good conditions; the highest mass neutron star observed, the highest glitch magnitude observed, and a rather close distance. In addition, we considered the maximum characteristic strain rate which assumes the best orientation of the star to us.

Overall, both perturbations excite the same modes to varying degrees. The vortex unpinning perturbation notably excites the inertial mode of the star which contains the majority of the kinetic energy and is located near the most sensitive frequency range of gravitational wave detectors (aLIGO). The star quake perturbation notably excites the ^4f mode more than the vortex unpinning model. Both perturbations exhibit nonphysical results in the ^4f and ^2f modes. Future work must be done to explain these features if these results are to be considered reputable.

Although the calculated strain amplitudes are below current detector sensitivity, they may be observable in the future depending on their dampening time. For periodic sources, the detection limit is inversely proportional to the observation time. Therefore, if the mode is excited for long enough, the low amplitudes may be overcame by the periodic nature of the signals. Future work could be done to explore dampening and incorporating it into our simulations.

9. ACKNOWLEDGMENTS

I'd like to thank both Bernard Whiting and Guido Mueller for choosing me for this year's IREU program. It was truly a once in a life time experience. I'd like to thank the National Science Foundation for funding such outstanding programs and motivating the next generation to get into science. I'd like to thank Dr. Hyung Mok Lee and Jinho Kim for welcoming me into their research group and helping me along the way. I'd also like to thank Peter Shaffery for being my partner in crime in everything Korean. I don't think I could have done it all without you.

10. REFERENCES

REFERENCES

[1]Anderson, P. W., & Itoh, N. 1975, Nature, 256, 25
[2]Baym, G., Pethick, C., & Pines, D. 1969, Nature, 224, 872
[3]Cox, J. P. 1980, Theory of Stellar Pulsation.
[4]Font, J. A., Dimmelmeier, H., Gupta, A., & Stergioulas, N. 2001, MNRAS, 325, 1463
[5]Hachisu, I., 1986a, ApJS, 61, 479.
[6]Hachisu, I., 1986b, ApJS, 62, 461.
[7]Kim, J., Kim, H. I., & Lee, H. M., 2009, MNRAS, 399, 229.
[8]Kim, J., Kim, H. I., & Lee, H. M., 2012, MNRAS, 424, 830
[9]Packard, R. E. 1972, Physical Review Letters, 28, 1080
[10]Ruderman, M. 1969, Nature, 223, 597
[11]Thorne, K. S. 1980, Rev. Mod. Phys., 52, 299

Squeezed quantum states of a harmonic oscillator in the problem of gravitational-wave detection

L. P. Grishchuk and M. V. Sazhin

P. K. Shternberg State Astronomical Institute
(Submitted 13 December 1982)
Zh. Eksp. Teor. Fiz. **84**, 1937–1950 (June 1983)

A quantum harmonic oscillator subjected to a resonant force action and to a parametric action is considered. Account is taken of the coupling of the oscillator with the thermostat (and with the zero-point oscillations). Explicit solutions of the Heisenberg equations are obtained for the creation and annihilation operators. Various methods of producing squeezed states in the oscillator are considered, and the most suitable for the registration of a weak (gravitational) force are chosen. Concrete estimates are obtained of the improvement of the sensitivity through the use of squeezed states and are obtained. The advantages and shortcomings of the squeezed-states technique and of quantum nondemolition measurements are compared. Possible uses of the squeezed-states technique are discussed.

PACS numbers: 03.65.Ge, 04.80. + z

The expected gravitational-wave signals of cosmic or laboratory origin are so weak that an analysis of detection properties with allowance for their quantum-mechanical character has become urgently necessary.[1] The usual observation criterion is taken to be an excess of the signal S over the noise N, i.e., $S/N \gtrsim 1$. Even if the thermal noise is neglected and there are no other imperfections in the detecting apparatus, a contribution to the noise is made by the quantum-mechanical variance of the measured quantity. A recent analysis of the so-called quantum nondemolition measurements (QND) consists precisely of a demonstration that for certain observable quantities (Hermitian operators) the variance can be made equal to zero and remains equal to zero, despite the action of the signal that leads to a change in the average value of the measured quantity. It assumed that the vanishing of the variance is due to the act of the first measurement. The successive continuous tracking of the measured quantities does not change its variance, since the quantity is QND-variable. In this idealized case the sensitivity of the detector is formally equal to infinity. The corresponding Hermitian operators, suitable for nondemolition measurement of a force F or of parametric action P were respectively named QNDF (Ref. 2) and QNDP operators (Ref. 3).

The studies of this subject have the following shortcomings. First, as a rule, they have ignored the presence of thermal noise, i.e., no account was taken of the fact that the variance of the QND variable, even when equal to zero at the initial instant of time, must inevitably increase with time because of the interaction of the oscillator with the thermostat. Second, it is not quite clear how to realize in practice instruments corresponding to the rather abstract QNDF and QNDP operators.

In the present paper we change somewhat the approach of the problem. We start from the fact that what is measured is a well defined quantity, not necessarily QND-variable. Such a quantity can be, e.g., the energy. As for the decrease of the noise N, i.e., of the variance of the measured quantity and (or) of the amplification of the signal S, i.e., the response of the detector to an external action, these are reached by a special "preparation" of the quantum state of the oscillator with the aid of really existing laboratory devices. We consider concretely the development and use of what are called squeezed states. Thus, an arriving gravitational signal interacts with an oscillator that is in a squeezed state and not, say, in an n-quantum or coherent state.

In quantum theory of light, squeezed states are known as "two-proton coherent states" and under other names.[4] In connection with detection of gravitational waves, squeezed states were considered by Caves,[5] who has shown that their use makes it possible to increase the sensitivity of a laser interferometer without increasing the laser power. It was also noted[6] that the use of squeezed states is promising in laboratory gravitation-wave experiments.

In the present paper we take consistent account of the connection between the oscillator and the thermostat. This enables us, in particular to generalize the theory of QND measurements to include the case of the presence of thermal noise and to compare the results with the method proposed here of using squeezed states.

§1. FUNDAMENTAL EQUATIONS AND THEIR SOLUTION

An arbitrary harmonic oscillator, mechanical or electrostatic, will be described in standard fashion with the aid of the creation and annihilation operators:

$$a^+=\left(\frac{m\omega}{2\hbar}\right)^{1/2}\left(\hat{q}-i\frac{\hat{p}}{m\omega}\right), \quad a=\left(\frac{m\omega}{2\hbar}\right)^{1/2}\left(q+i\frac{\hat{p}}{m\omega}\right),$$

$$[a, a^+]=1.$$

It is convenient also to introduce the operators X_1 and X_2 of the complex amplitude of the oscillator, in accordance with the definition

$$X_1=l(e^{-i\omega t}a^+ + e^{i\omega t}a), \quad X_2=il(e^{-i\omega t}a^+ - e^{i\omega t}a),$$

where $l=(\hbar/2m\omega)^{1/2}$.

It is known that a coherent state $|\alpha\rangle \equiv D(\alpha)|0\rangle$ is defined with the aid of the displacement operator

 0038-5646/83/061128-08$04.00 © 1984 American Institute of Physics

$D(\alpha)=\exp(\alpha a^+ - \alpha^* a)$, where α is a complex number. Participating in the construction of the squeezed state $|\alpha,\beta\rangle \equiv D(\alpha)S(\beta)|0\rangle$ is also the squeezing operator

$$S(\beta)=\exp\left[{}^1/_2(\beta^* a^2-\beta a^{+2})\right],$$

where β is a complex number. A squeezed state is thus characterized by two complex numbers.

A coherent state results from the vacuum state $|0\rangle$ under the influence of a force, i.e., on account of a term of the type $F(t)(a+a^+)$ in the Hamiltonian. One of the possible methods of obtaining a squeezed state is the use of a degenerate parametric amplifier,[4,7] i.e., of a term of the type $P(t)(a^{+2}+a^2)$ in the Hamiltonian. If the parametric pumping is effected at the frequency 2ω and is of the form

$$P(t)=|p|\sin(2\omega t+\psi),$$

the term of interest to us in the Hamiltonian can be rewritten in the form

$$P(t)\,(a^2+a^{+2})$$

$$=\frac{1}{2i}(pe^{i2\omega t}a^{+2}-p^*e^{-i2\omega t}a^{+2}+pe^{i2\omega t}a^2-p^*e^{-i2\omega t}a^2), \qquad (1)$$

$$p=|p|e^{i\psi}$$

inasmuch as for free evolution of the oscillator we have

$$a(t)=a(0)\,e^{-i\omega t} \text{ and } a^+(t)=a^+(0)\,e^{i\omega t}.$$

The first and fourth terms in (1) are rapidly oscillating and can be neglected compared with the resonant, second and third, terms. Similarly, retaining only the terms resonant at the frequency ω, the force term in the Hamiltonian can be written in the form

$$F(t)\,(a+a^+)=fe^{-i\omega t}a^+ + f^*e^{i\omega t}a, \qquad f=|f|e^{i\varphi}.$$

We shall thus work with a Hamiltonian

$$H=\hbar\omega a^+a+{}^1/_2 ik\omega\,(p^*e^{i2\omega t}a^2-pe^{-i2\omega t}a^{+2})$$
$$+\hbar\omega\,(fe^{-i\omega t}a^+ + f^*e^{i\omega t}a), \qquad (2)$$

where p and f are arbitrary complex numbers.

We note that the Hamiltonian (2) is among those that describe resonant action of a gravitational wave on a mode of the electromagnetic oscillations in the resonator.[3] To be sure, in the case of a gravitational signal the coupling constants $|p_g|$ and $|f_g|$ are determined by the amplitude of a gravitational wave h and are quite small under real conditions, whereas the coupling constants with laboratory generators, $|p_l|$ and $|f_l|$, depend to a considerable degree on the experimenter. In the general case, it can be assumed that $p=p_l+p_g$ and $f=f_l+f_g$, but hereafter we shall deal for simplicity only with estimates of observing a gravitational force and therefore put $p_g=0$ and $f_g\neq 0$.

The interaction of an oscillator with the thermostat can be described by adding to the Heisenberg equations of motion the term that takes the damping into account, and the operator of the random force (see, e.g., Ref. 8):

$$i\hbar\frac{da}{dt}=[a,H]-i\hbar\frac{\gamma}{2}a+i\hbar\theta. \qquad (3)$$

The constant $\gamma>0$ is connected with the relaxation time τ^* by the relation $\gamma^{-1}=\tau^*$. The random-force operator $\theta(t)$ satisfies the commutation relation $[\theta(t),\theta^+(t')]=\gamma\delta(t-t')$. The equation for a^+ is obtained from (3) by Hermitian conjugation.

Substituting in (3) the Hamiltonian (2) we obtain the equations of motion in explicit form

$$\frac{da}{dt}=-i\omega a-\omega pa^+e^{-i2\omega t}-i\omega fe^{-i\omega t}-\frac{\gamma}{2}a+\theta,$$
$$\frac{da^+}{dt}=i\omega a^+-\omega p^*ae^{i2\omega t}+i\omega f^*e^{i\omega t}-\frac{\gamma}{2}a^++\theta^+.$$

To solve them it is convenient to eliminate the free evolution by using the substitution

$$a(t)=e^{-i\omega t}e^{-\gamma t/2}A(t), \qquad a^+(t)=e^{i\omega t}e^{-\gamma t/2}A^+(t)$$

and carry out the Bogolyubov transformation

$$b(t)=e^{-i\psi/2}\operatorname{ch} rtA(t)+e^{i\psi/2}\operatorname{sh} rtA^+(t),$$
$$b^+(t)=e^{i\psi/2}\operatorname{ch} rtA^+(t)+e^{-i\psi/2}\operatorname{sh} rtA(t),$$

where $r=|p|\omega$. We then obtain for $b(t)$ the simple equation

$$db/dt=-i\omega e^{\gamma t/2}(f\operatorname{ch} rte^{-i\psi/2}-f^*\operatorname{sh} rte^{i\psi/2})$$
$$+e^{\gamma t/2}(e^{i(\omega t-\psi/2)}\operatorname{ch} rt\theta+e^{-i(\omega t-\psi/2)}\operatorname{sh} rt\theta^+).$$

The final solution for $a(t)$ is

$$a(t)=e^{-i\omega t}e^{-\gamma t/2}[m(t)+\Phi(t)+T(t)], \qquad (4)$$

where

$$m(t)=a_0\operatorname{ch} rt-a_0^+e^{i\psi}\operatorname{sh} rt,$$

$$\Phi(t)=-i\omega\int_0^t e^{\gamma\tau/2}[f\operatorname{ch} r(\tau-t)-f^*e^{i\psi}\operatorname{sh} r(\tau-t)]\,d\tau,$$

$$T(t)$$

$$=\int_0^t e^{\gamma\tau/2}[\theta(\tau)e^{i\omega\tau}\operatorname{ch} r(\tau-t)+\theta^+(\tau)e^{-i(\omega\tau-\psi)}\operatorname{sh} r(\tau-t)]\,d\tau.$$

An expression for $a^+(t)$ is obtained from (4) by Hermitian conjugation.

To calculate the mean values and the variances it is necessary to average using the density matrix. The complete oscillator + thermostat system is described by the direct product of the oscillator density matrix ρ and the thermostat density matrix ρ_T. The averaging of the operator $\theta(t)$ has the following properties[8] (the angle brackets $\langle\ \rangle$ denote averaging over the total density matrix, but the operator θ is in fact averaged over ρ_T):

$$\langle\theta(t)\rangle=\langle\theta^+(t)\rangle=0, \qquad \langle\theta^+(t)\theta(t')\rangle=\gamma n_\tau\delta(t-t'),$$
$$\langle\theta(t)\theta^+(t')\rangle=\gamma(n_\tau+1)\delta(t-t').$$

Here n_T is the average number of quanta in the thermostat at the frequency ω:

$n_T=[e^{\hbar\omega/kT}-1]^{-1}$.

The binary terms containing θ^2 and θ^{+2} yield zero on averaging.

When calculating the variance of the quantum-number operator $n=a^+a$ in the oscillator, use is made of fourth-order moments. Assuming that θ and θ^+ have Gaussian distributions, they reduce to second-order moments in accordance with the expansion[9]:

$$\langle\theta(t_1)\theta^+(t_2)\theta(t_3)\theta^+(t_4)\rangle=\langle\theta(t_1)\theta^+(t_2)\rangle\langle\theta(t_3)\theta^+(t_4)\rangle+\langle\theta(t_1)\theta^+(t_4)\rangle\langle\theta^+(t_2)\theta(t_3)\rangle.$$

We present expressions for the mean values of the operators X_1, X_2, and n and their variances, but do not specify concretely for the time being the initial value of the oscillator density matrix:

$$\langle X_1(t)\rangle=e^{-\gamma t/2}[(\operatorname{ch}rt-\cos\psi\operatorname{sh}rt)\langle X_1(0)\rangle-\sin\psi\operatorname{sh}rt\langle X_2(0)\rangle+l(\Phi(t)+\Phi^*(t)], \tag{5}$$

$$\langle X_2(t)\rangle=e^{-\gamma t/2}[(\operatorname{ch}rt+\cos\psi\operatorname{sh}rt)\langle X_2(0)\rangle-\sin\psi\operatorname{sh}rt\langle X_1(0)\rangle+il(\Phi^*(t)-\Phi(t))], \tag{6}$$

$$\langle n(t)\rangle=e^{-\gamma t}[\langle m^+(t)m(t)\rangle+\Phi(t)\langle m^+(t)\rangle+\Phi^*(t)\langle m(t)\rangle+|\Phi(t)|^2+\langle T^+T\rangle]. \tag{7}$$

We note that at $\sin\psi=0$ the mean values of X_1 and X_2 do not depend on the initial mean values X_2 and X_1, respectively.

We write down the variances of X_1 and X_2 by means of a single formula, in which the upper sign pertains to δX_1^2 and the lower to δX_2^2:

$$\delta X_{1,2}^2=\langle X_{1,2}^2\rangle-\langle X_{1,2}\rangle^2=e^{-\gamma t}[(\operatorname{ch}rt\mp\cos\psi\operatorname{sh}rt)^2\times\delta X_{1,2}^2(0)+\sin^2\psi\operatorname{sh}^2rt\,\delta X_{2,1}^2(0)-\sin\psi\operatorname{sh}rt(\operatorname{ch}rt\mp\cos\psi\operatorname{sh}rt)K(0)+l^2(\pm\langle T^+T^+\rangle\pm\langle TT\rangle+\langle T^+T\rangle+\langle TT^+\rangle)], \tag{8}$$

$$K(0)=\langle X_1(0)X_2(0)\rangle+\langle X_2(0)X_1(0)\rangle-2\langle X_1(0)\rangle\langle X_2(0)\rangle.$$

It can be seen that the variances δX_1^2 and δX_2^2 do not depend on the acting force. For the variance of the number of quanta we have the general formula

$$\begin{aligned}e^{2\gamma t}\delta n^2(t)=&[\langle(m^+m)^2\rangle-\langle m^+m\rangle^2]+\Phi[\langle m^+mm^++m^{+2}m\rangle-2\langle m^+\rangle\langle m^+m\rangle]+\Phi^*[\langle mm^+m+m^+m^2\rangle-2\langle m^+m\rangle\langle m\rangle]\\&+\Phi^2[\langle m^{+2}\rangle-\langle m^+\rangle^2]+\Phi^{*2}[\langle m^2\rangle-\langle m\rangle^2]\\&+|\Phi|^2[\langle m^+m+mm^+\rangle-2\langle m\rangle\langle m^+\rangle]+\langle T^2\rangle[\langle m^{+2}\rangle+2\Phi^*\langle m^+\rangle+\Phi^{*2}]+\langle T^{+2}\rangle[\langle m^2\rangle+2\Phi\langle m\rangle+\Phi^2]\\&+\langle TT^+\rangle[\langle m^+m\rangle+\Phi\langle m^+\rangle+\Phi^*\langle m\rangle+|\Phi|^2]\\&+\langle T^+T\rangle[\langle mm^+\rangle+\Phi\langle m^+\rangle+\Phi^*\langle m\rangle+|\Phi|^2]\\&+\langle(T^+T)^2\rangle-\langle T^+T\rangle^2.\end{aligned} \tag{9}$$

Expressions for the mean values containing products of T and T^+ are given in the Appendix.

If at the initial instant of time $t=0$ the oscillator was at equilibrium with the thermostat, it is described by the density matrix

$$\rho=[1-\exp(-\hbar\omega/kT)]\sum_{n=0}^{\infty}e^{-n\hbar\omega/kT}|n\rangle\langle n|,$$

and then

$$\langle X_1(0)\rangle=\langle X_2(0)\rangle=K(0)=0,\quad n(0)=n_T,$$
$$\delta X_1^2(0)=\delta X_2^2(0)=l^2(2n_T+1),\quad \delta n^2(0)=n_T(n_T+1).$$

Expressions for the mean values and the variances that arise under the action of f and p on an equilibrium oscillator are given in the Appendix.

Different aspects of the behavior of an oscillator in a thermostat under the action of a force and (or) parametric pumping were analyzed in Refs. 10 and 11.

§2. PRODUCTION OF SQUEEZED STATES

A decrease (squeezing) of the variance of the operator X_1 or X_2 is reached by a special choice of the phase ψ of the parametric generator. Assume that at $t=0$ the oscillator was in thermodynamic equilibrium with the thermostat. We put $\psi=0$. It follows then from (A3) that

$$\begin{aligned}\delta X_1^2(t)&=l^2(2n_T+1)(\gamma+2r)^{-1}(\gamma+2re^{-(\gamma+2r)t}),\\\delta X_2^2(t)&=l^2(2n_T+1)(\gamma-2r)^{-1}(\gamma-2re^{-(\gamma-2r)t}),\end{aligned} \tag{10}$$

i.e., in the variance of X_1 decreases with time, and the variance of X_2 increases, compared with the equilibrium value (we assume for the sake that $r>0$). The maximum squeezing of the variance of X_1 as $t\to\infty$ is always finite, and the equilibrium value $\delta X_1^2(0)$ decreases by a factor $\gamma/(\gamma+2r)$:

$$\delta X_1^2(t\to\infty)=l^2(2n_T+1)[\gamma/(\gamma+2r)]. \tag{11}$$

The increase of the variance of X_2 is finite if $2r<\gamma$, and increases exponentially with time if $2r>\gamma$. If the phase $\psi=\pi$ is chosen, the variances of X_1 and X_2 interchange roles (in (10) this is equivalent to replacing r by $-r$). At arbitrary ψ, the operators Y_1 and Y_2, which are certain linear combinations of X_1 and X_2, are subject to squeezing and to dilatation.

At arbitrary $2r\gg\gamma$, squeezing of the variance of X_1 to a value that practically reaches the limit takes place within a characteristic time

$$t_c=(2r)^{-1}\ln(2r/\gamma)\ll\gamma^{-1}.$$

Considerable squeezing is possible only at $2r$ comparable with γ or exceeding γ, and we shall therefore consider hereafter the cases $2r=\gamma$ and $2r\gg\gamma$.

The upper bound of the parameter $2r$ is 2ω. At this value $|p|=1$ and the Hamiltonian (2) would no longer be positive-definite. Consequently the oscillator would become unstable and the region of permissible values of the oscillator energy would range from $-\infty$ to $+\infty$.

As already noted, the variances of the operators X_1 and X_2 depend on the parameter r but not on f. The force enters only into the law that governs the mean values of X_1 and X_2. At $\psi=0$, as follows from (A.10), we obtain

$$\begin{aligned}\langle X_1(t)\rangle&=4l|f|\frac{\omega}{\gamma+2r}\sin\varphi[1-e^{-(r+\gamma/2)t}],\\\langle X_2(t)\rangle&=-4l|f|\frac{\omega}{\gamma-2r}\cos\varphi[1-e^{(r-\gamma/2)t}].\end{aligned} \tag{12}$$

We now turn to the mean value and to the variance of the number of quanta. Under a force generator and a parametric generator act jointly on an oscillator placed in a thermostat, $n(t)$ and $\delta n^2(t)$ vary in accordance with Eqs. (A2) and (A4). In contrast to δX_1^2 and δX_2^2, the variance δn^2 depends on f, so that the force increases not only n but also δn^2.

To observe a weak force (e.g., a gravitational wave) it is desirable to create states with low variance and with large number of the quanta. Large $\langle n(t)\rangle$ is desirable because in a squeezed state, just as in a coherent one, the response of the oscillator, i.e., the change Δn of the number of quanta under the action of the signal, increases in proportion to the square root of the number of quanta contained in the oscillator. Relegating the proof of this statement and a detailed analysis of the detection criteria to §3, we discuss the conditions under which the squeezing, i.e., the use of a parametric generator with $r\neq 0$, makes it possible to decrease the ratio $\delta n/n^{1/2}$ in the oscillator.

It can be seen from (A2) and (A4) that were we to have $r=0$ i.e., if a coherent state (plus thermal noise) were produced in the oscillator under the influence of a laboratory force, we would have by the instant of time $t=t_1$

$$\langle n(t_1)\rangle = n_T + |\alpha|^2, \tag{13}$$

$$\delta n^2(t_1) = n_T(n_T+1) + (2n_T+1)|\alpha|^2, \tag{14}$$

where

$$\alpha = (2\omega|f|/\gamma)(1-e^{-\gamma t/2})e^{i(\varphi-\pi/2)}.$$

The parameter α defines also the mean values of X_1 and X_2:

$$\langle X_1\rangle + i\langle X_2\rangle = 2l\alpha.$$

It can be seen therefore that even within a time exceeding the relaxation time $\tau^* = \gamma^{-1}$ it is impossible to obtain a better result than

$$\delta n/n^{1/2} \approx (2n_T+1)^{1/2}.$$

This quantity does not differ in practice from the equilibrium value

$$\delta n/n^{1/2} = (n_T+1)^{1/2}.$$

If, however, $r\neq 0$, the choice of the phase shift $\varphi - \psi/2$ becomes important. For the choice $\cos(\varphi-\psi/2)=0$ and under the assumption $2r\gg\gamma$, the joint action of the force and parametric generators during the time

$$t \gtrsim t_c = (2r)^{-1}\ln(2r/\gamma)$$

leads to

$$n(t_c) \approx \frac{|f|^2}{|p|^2}, \quad \delta n^2(t_c) \approx (2n_T+1)\frac{|f|^2}{|p|^2}\frac{\gamma}{2r}.$$

It was assumed here in the calculation that the deviation from the equilibrium state is appreciable, i.e., the terms containing $|f|$ predominate, and, concretely,

$$|f|^2/|p|^2 \gg (2n_T+1)(2r/\gamma)^3.$$

As a result we obtain a squeezed state with a value

$$\delta n/n^{1/2} \approx (2n_T+1)^{1/2}(\gamma/2r)^{1/2}, \tag{15}$$

which is better by a factor $(\gamma/2r)^{1/2}$ than in the coherent state. The obtained squeezed state can be used to record a weak force within a time interval that is short compared with τ^*, i.e., small compared with the decay time of this state.

Large deviations in the choice of the parameters used in the derivation of (15) make this estimate worse. The choice of the phase $\cos(\varphi-\psi/2)=0$ when considering $\langle n(t)\rangle$ and δn^2 follows the same purpose as the choice of the phases $\psi=0$ and $\cos\varphi=0$ when considering the most effective decrease of $\delta X_1^2(t)$ and the most effective increase of $\langle X_1(t)\rangle$.

In the derivation of (15) it was assumed that the force and parametric generators act jointly during the characteristic squeezing time t_c. A longer action of the generators is not advantageous, since it leads to dominance of terms that do not contain $|f|$, and then

$$\delta n/n^{1/2} \approx (2n_T+1)^{1/2}\exp[(r-\gamma/2)t].$$

At the same time, after a short time t_c the average number of the quanta does not increase strongly enough to permit it to grow under the influence of only the force generator in a time $\tau^* \gtrsim t \gg t_c$. Indeed, at $t\approx\tau^*$ we have

$$|\alpha|^2 \approx \frac{4\omega^2|f|^2}{\gamma^2} = \frac{|f|^2}{|p|^2}\left(\frac{2r}{\gamma}\right)^2 \gg \frac{|f|^2}{|p|^2}.$$

Under the experimental conditions it may be desirable to increase the oscillator response Δn that is proportional to the initial $\sqrt{n}$, with the preservation of the condition (15).

It is possible to obtain the same estimate (15), but with larger n, if the force and parametric generators act separately and in succession. The idea is to produce first with the aid of the force generator a large $\langle n(t)\rangle$, and then turn off the force action and turn on the parametric action. Then, at a suitable choice of the phases, after a short squeezing time t_c the variance δn^2 decreases strongly, and $\langle n(t)\rangle$ does not change so much within the same time.

Let us consider the proposed method in greater detail. Assume that an oscillator in equilibrium with the thermostat is acted upon by a force during a time t_1. Then by the instant t_1 we have

$$a(t_1) = e^{-i\omega t_1}e^{-\gamma t_1/2}[a_0 + \Phi(t_1) + T(t_1)]$$

and the Hermitian-adjoint expression for $a^+(t_1)$. The values of $\langle n(t_1)\rangle$ and $\delta n^2(t_1)$ are determined by Eqs. (13) and (14). At the instant t_1 the force is turned off and r is turned on. The subsequent evolution of the oscillator can be calculated from the general formulas (7) and (9) with the already known initial data. In this case $\Phi\equiv 0$ and it is convenient to reckon the time from the instant $t=t_1$. We then obtain for $\langle n(t)\rangle$

$$\langle n(t)\rangle = e^{-\gamma t}|\alpha|^2[\operatorname{ch} 2rt - \operatorname{sh} 2rt\cos(\psi-2\varphi_1)] + n_0,$$

where $\varphi_1 = \varphi - \omega t_1 - \pi/2$ and the term n_0 stands for expression (A2) in which we must put $|\Phi|^2=0$. For $\delta n^2(t)$, using the notation of the Appendix, we can obtain

$$\begin{aligned} e^{2\gamma t}\delta n^2(t) &= n_T(n_T+1)\operatorname{ch} 4rt \\ &+ |\alpha|^2(2n_T+1)[\operatorname{ch} 4rt - \operatorname{sh} 4rt\cos(\psi-2\varphi_1)] \\ &+ {}^1/_2\gamma|\alpha|^2(2n_T+1)\{e^{2rt}[1-\cos(2\varphi_1-\psi)](I_c-I_s) \\ &+ e^{-2rt}[1+\cos(\psi-2\varphi_1)](I_c+I_s)\} \\ &+ {}^1/_2[\operatorname{sh}^2 2rt - e^{\gamma t} + 1 + \gamma(2n_T+1)^2(I_c\operatorname{ch} 2rt - I_s\operatorname{sh} 2rt)]. \end{aligned}$$

The principal growing terms in $\langle n(t)\rangle$ and $\delta n^2(t)$ are eliminated by choosing the phase $\cos(2\varphi_1 - \psi) = 1$. We assume also that $2r \gg \gamma$ and

$$|\alpha|^2 \gg [n_T(n_T+1)/(2n_T+1)](2r/\gamma)^{1/2}.$$

By the instant of time t_c we then have

$$n(t_c) \approx |\alpha|^2 \frac{\gamma}{2r}, \quad \delta n^2(n_1, t_c) \approx |\alpha|^2 (2n_T+1) \left(\frac{\gamma}{2r}\right)^2, \quad (16)$$

i.e., $\delta n/\langle n\rangle^{-1/2}$ we return to Eq. (15) for, but now we have $n(t_c)$ larger by $2r/\gamma$ times than the value of n reached in the variant with joint action of f and r. We note that by that instant of time the average values of the operators a and a^+ are equal, apart from a phase factor, to the quantity $|\alpha|(\gamma/2r)^{-1/2}$, i.e., to $\langle n\rangle^{-1/2}$.

Worthy of special consideration is the case $2r = \gamma$. In Eqs. (12) and (10) for $\langle X_1(t)\rangle$ and $\delta X_1^2(t)$ this case is not at all remarkable, but in the expressions for $\langle X_2(t)\rangle$ and $\delta X_2^2(t)$ it leads to a singularity:

$$\langle X_2(t)\rangle = -2l|f|\omega t \cos\varphi \quad \delta X_2^2(t) = l^2(2n_T+1)(1+\gamma t). \quad (17)$$

If the experimenter can use time intervals that are long compared with the relaxation time, then at $t \gg \gamma^{-1}$ we can obtain (at $\cos\varphi = -1$)

$$\delta X_2/X_2 \approx [(2n_T+1)^{1/2}/2|f|\omega](\gamma/t)^{1/2},$$

i.e., $\delta X_2/X_2 \to 0$ at $t \to \infty$. We note that at $2r \gg \gamma$ and $t \to \infty$ we would obtain a nonvanishing and rather large quantity

$$\delta X_2/X_2 = (2n_T+1)^{1/2}(r/2|f|\omega).$$

For $\langle n(t)\rangle$ and $\delta n^2(t)$ at $2r = \gamma$, $\cos(\varphi - \psi/2) = \pm 1$ and $\gamma t \gg 1$ we have from (A2) and (A4)

$$\langle n(t)\rangle \approx |f|^2\omega^2 t^2, \quad \delta n^2 \approx (2n_T+1)|f|^2\omega^2\gamma t^3. \quad (18)$$

As will be shown in §3, the case $2r = \gamma$ can be used to improve the sensitivity if the signal can be accumulated during many relaxation times.

The variance δn of the number of quanta must be set in correspondence with the variance $\delta\hat{\varphi}$ of the phase operator, even though this operator is known to be defined subject to some stipulations. We present nevertheless illustrative arguments from which it can be seen that in squeezed states a decrease of δn is accompanied by an increase of $\delta\hat{\varphi}$ and an increase of δn is accompanied by a decrease of $\delta\hat{\varphi}$. We consider first coherent states. On a plane with coordinates $\langle X_1\rangle$ and $\langle X_2\rangle$ the coherent state $|\alpha\rangle$ is mapped by a circle with center at the points $2l\alpha$ in accordance with the equation $\langle X_1\rangle + i\langle X_2\rangle = 2l\alpha$. The diameter of the circle is determined by the variance $\delta X_1 = \delta X_2 = l(2n_T + 1)^{1/2}$. The uncertainty of the number of quanta at

$$|\alpha|^2(2n_T+1) \gg n_T(n_T+1)$$

is $\delta n \approx (2n_T + 1)^{1/2}|\alpha|$, i.e., $\delta n = S(m\omega/2\pi\hbar)$, where S is the area of a ring made up by rotating the circle around the origin (see Fig. 1). The phase uncertainty $\delta\hat{\varphi}$ is expressed in terms of the (doubled) angle from which the diameter is seen from the origin, so that

$$\delta\hat{\varphi} \approx (2n_T+1)^{1/2}/|\alpha|.$$

As a result we obtain

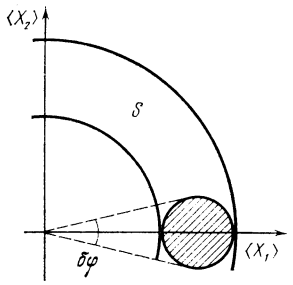


FIG. 1

$$\delta n\delta\hat{\varphi} \approx (2n_T+1).$$

In the squeezed state, the variance circle is transformed into an ellipse compressed along the axis for which the variance is a minimum. If the phases $\psi = 0$ and $\cos\varphi = 0$ are chosen, the ellipse is compressed along the X_1 axis, and its center is located on the X_1 axis; at $\psi = 1$ and $\cos\varphi = \pm 1$, the ellipse is compressed along the same axis, but its center is located on the X_2 axis (see Fig. 2). Accordingly, in the former case the area of the ring decreases, and the angle increases, so that [cf. (16)]

$$\delta n \approx (m\omega/\hbar)\langle X_1\rangle\delta X_1 \approx (2n_T+1)^{1/2}|\alpha|(\gamma/2r),$$
$$\delta\hat{\varphi} \approx 2\delta X_2/\langle X_2\rangle \approx (2n_T+1)^{1/2}|\alpha|^{-1}(2r/\gamma)$$

and

$$\delta n\delta\hat{\varphi} \approx (2n_T+1).$$

In the latter case δn increases and $\delta\hat{\varphi}$ decreases, so that again

$$\delta n\delta\hat{\varphi} \approx 2n_T+1.$$

§3. USE OF SQUEEZED STATES TO DETECT A WEAK FORCE

The possibilities of using squeezed states can be naturally separated into two versions, depending on whether the duration of the signal and other conditions of the experiment will permit accumulation of a signal during several relaxation times, or whether the accumulation time is much shorter than γ^{-1}. If $\gamma\tau \ll 1$, it is reasonable to first "prepare" the squeezed state, and then use it to observe the force before the variance of the measured quantity manages to increase substantially. The action of the signal itself in the course of preparation of the state is immaterial. We consider this technique for the measurement of $X_1(t)$ and of the number of quanta.

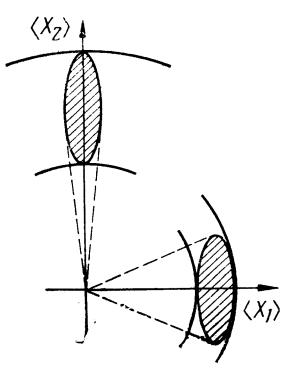


FIG. 2

Assume that at the instant when the laboratory generators are turned off (we choose this instant of time to be zero time) the squeezing reached (at $2r \gg \gamma$) is

$$\delta X_1^2(0) \approx l^2 (2n_\tau + 1)\,(\gamma/2r). \tag{19}$$

Subsequently $\delta X_1^2(t)$ increases in accordance with (8):

$$\delta X_1^2(t) = l^2(2n_\tau+1)\left[\frac{\gamma}{2r}e^{-\gamma t} + 1 - e^{-\gamma t}\right]$$

$$\approx l^2(2n_\tau+1)\left(\frac{\gamma}{2r} + \gamma t\right). \tag{20}$$

The expression (20), which is approximate at $\gamma t \ll 1$, makes it possible to refine the permissible observation time τ. It is desirable for this time to be such that the growing term in δX_1^2, i.e., the term proportional to γt, still has not exceeded the initial value $\delta X_1^2(0)$. In other words, if $\tau \lesssim (2r)^{-1}$, the variance did increase by not more than a factor of 2, but if $(2r)^{-1} \ll \tau < \gamma^{-1}$, the variance is still smaller than the equilibrium value, but comes close to it. The change of $\langle X_1(t)\rangle$ under the action of the gravitational force f_g to be observed proceeds in accordance with (5), where we must put $r = 0$, $f = f_g$, and $\varphi = \varphi_g$. This yields

$$\Delta X_1(t) = 4l|f_g|\frac{\omega}{\gamma}\sin\varphi_g(1 - e^{-\gamma t/2}) \approx 2l|f_g|\,\omega t \sin\varphi_g \tag{21}$$

Taking ΔX_1 to be the signal S and δX_1 the noise N and assuming also $\sin\varphi_g = 1$, we obtain within an observation time $\tau \lesssim 1/2r$

$$S/N \approx [2|f_g|\omega\tau/(2n_\tau+1)^{1/2}]\left(\frac{2r}{\gamma}\right)^{1/2} \tag{22}$$

and after a time $1/2r \ll \tau \lesssim 1/\gamma$

$$S/N \approx [2|f_g|\omega/(2n_\tau+1)^{1/2}]\left(\frac{\tau}{\gamma}\right)^{1/2}. \tag{23}$$

If the oscillator were to be in a coherent state, then after an observation time $0 < \tau \leqslant \gamma^{-1}$ we would obtain the estimate

$$S/N \approx 2|f_g|\,\omega\tau/(2n_\tau+1)^{1/2}, \tag{24}$$

which is worse by $(2r/\gamma)^{1/2}$ times than (22), and worse by $(\gamma\tau)^{-1}$ times than (23). Of course, if it is possible to accumulate a signal within a time on the order of γ^{-1}, the minimum observable force during that time is, according to (24),

$$f_{min} \approx (2n_\tau+1)^{1/2}\gamma/2\omega,$$

much better than the estimate

$$f_{min} \approx \frac{(2n_\tau+1)^{1/2}}{2\omega}(2r\gamma)^{1/2}$$

that follows from Eq. (22) at $\tau \approx 1/2r$, and somewhat better than the estimate that follows from (23). This is so because the time required to reach the maximum ΔX_1, as follows from (21), is γ^{-1} while the decay time of the squeezed state is also γ^{-1}. Thus, it is advantageous to use squeezed states that are prepared beforehand if the time of observation is strongly limited for some reason or another.

Perfectly analogous conclusions are obtained also by analyzing measurements of the number of quanta. The action of the force changes $\langle n(t)\rangle$ in accordance with (7):

$$\langle n(t)\rangle = n_0 e^{-\gamma t} + n_\tau(1 - e^{-\gamma t}) + \frac{2i\omega}{\gamma}e^{-\gamma t}(e^{\gamma t/2} - 1)$$

$$\times[\langle a_0\rangle f_g^* - \langle a_0^+\rangle f_g] + \frac{4\omega^2}{\gamma^2}|f_g|^2(1 - e^{-\gamma t/2})^2, \tag{25}$$

where n_0 is the initial average number of quanta.

Inasmuch as in the squeezed state $\langle a_0\rangle \neq 0$ and $\langle a_0^+\rangle \neq 0$, the term linear in the small quantity f_g is preserved, just as in the case of the coherent state (but not in the n-quantum or equilibrium state). We are interested in such squeezed states in which at a given average number of quanta the variance δn^2 is much less than in the coherent state with the same n. In such states, as can be seen from a discussion of Eqs. (16), $\langle a_0\rangle$ and $\langle a_0^+\rangle$ are equal, apart from phase factors, to $\sqrt{n_0}$. Thus, according to (25) the change of the number of quanta under the action of f_g and at $\gamma t \ll 1$ is

$$\Delta n(t) \approx 2|f_g|\sqrt{n_0}\,\omega t.$$

The variance δn^2 increases from an initial value δn_0^2, which we choose in accordance with (16) in the form

$$\delta n_0^2 = n_0(2n_\tau+1)\,\gamma/2r.$$

The law of variation of the variance is determined by the general formula (9) at $r = 0$, but we assume f_g to be so small that the increase of δn^2 is determined by the interaction with the thermostat and not with f_g. Then

$$\delta n^2(t) \approx n_0(2n_\tau+1)\,(\gamma/2r + \gamma t)$$

and for $\tau \lesssim (2r)^{-1}$ we have

$$\frac{S}{N} = \frac{\Delta n}{\delta n} \approx \frac{2|f_g|\,\omega\tau}{(2n_\tau+1)^{1/2}}\left(\frac{2r}{\gamma}\right)^{1/2}$$

which coincides with (22). At $1/2r \ll \tau \lesssim \gamma^{-1}$ we obtain a formula that coincides with (23).

We consider now the observation of prolonged signals with the aid of the variant $2r = \gamma$. As seen from (17) and (18), the mean values of X_2 and n and of their variances increase in the course of time, but the growth of the mean values is faster. The oscillator is acted upon jointly by the laboratory force f_l and by the gravitational f_g; i.e.,

$$f \equiv |f|e^{i\varphi} = |f_l|e^{i\varphi_l} + |f_g|e^{i\varphi_g}.$$

The increment of $\langle X_2(t)\rangle$ due to the force f_g is, if the phases are suitably chosen

$$\Delta X_2 \approx 2l|f_g|\,\omega t,$$

and the variance is

$$\delta X_2^2 \approx l^2(2n_\tau+1)\,\gamma t,$$

from which we obtain at an observation time $\tau \approx \gamma^{-1}$ or $\tau \gg \gamma^{-1}$

$$S/N = [2|f_g|\,\omega/(2n_\tau+1)^{1/2}]\,(t/\gamma)^{1/2}.$$

Exactly the same formula is obtained for $S/N = \Delta n/\delta n$. Thus, the minimum observable force decreases without limit with increasing observation time. The actual sensitivity limits will be determined by the accuracy with which the necessary relations between the phases and the parameters are satisfied.

§4. COMPARISON OF THE SQUEEZED-STATES TECHNIQUE AND QND MEASUREMENTS IN THE PROBLEM OF OBSERVING A WEAK FORCE

The squeezed-states technique and the use of QND measurements[1-3,12-14] have common features. In both cases one of the tasks of the method is to decrease the variance of the measured quantity. In QND measurements the decrease of the variance is reached by the very act of the first measurement, as a result of which a reduction of the wave function takes place, while in the squeezed-states technique the variance is decreased by specially chosen force and parametric action on the oscillator. Although the postulates of quantum mechanics permit the variance to be decreased after the measurement to zero, its actual value is determined by the coupling constants and by the measuring instrument, by the measurement time, etc. In addition, as already noted, the variance will inevitably increase with time as a result of the interaction of the system with the thermostat. In this sense, we lose the fundamental difference between the final variance obtained by squeezing, and that (final) dispersion which can be obtained after the first measurement performed in a realistic situation.

The initial requirement and theory of QND measurements is to choose quantities that can be measured with arbitrary accuracy without demolishing the state of the system. The system, by assumption is, after the first measurement in one of the eigenstates of the QNDF operator or of the QNDP operator, and it is precisely this quantity which is measured subsequently. The squeezed-states technique does not presuppose an obligatory measurement of the QND-variable. Prior to the arrival of the gravitational signal and during the time of its action, the oscillator is in a squeezed state and not, say, in an eigenstate $X_1(t)$ of the QNDF-operator. In this case the force can be detected, e.g., by the changes of the energy, i.e., of a quantity that is not a QNDF-variable. However, the experimenter must not claim too high an accuracy of the measurements and, consequently, too large a perturbation of the system. The accuracy of the measurements must not be better than the (*a priori*) variance of the measured quantity.

To compare the two considered methods, we assume that we measure the QNDF-variable $X_1(t)$. In accordance with the QND-measurement principle, we assume that at the initial instant of time an exact measurement of $X_1(t)$ was performed on the equilibrium oscillator, as a result of which the oscillator is brought to an eigenstate of the operator X_1 with a value $X_1(0)=0$; then $\delta X_1^2(0)=0$. Subsequently, the variance will increase in accordance with the general formula (8), and the average value will change under the influence of f_g in accordance with (5). The S/N ratio after a time τ of the action of the force will be

$$S/N=[2|f_g|\omega/(2n_T+1)^{1/2}](\tau/\gamma)^{1/2}. \tag{26}$$

We assume now that at the initial instant of time the oscillator is in a squeezed state. We measure again the quantity X_1, but not too accurately, within the limits of the variance. Then the S/N ratio is determined by Eqs. (22) and (23). We know that the measurement of the energy leads to similar equations. The difference from formula (26) takes place only at $0<\tau\lesssim 1/2r$. In this time interval the sensitivity of (26) is generally speaking better than that of (22). However, when $2r$ tends to a definite limit equal to 2ω, the difference vanishes, since the duration of the action of f_g is assumed at any rate to be not shorter than $2\pi/\omega$. Of course, as $r\to\omega$ the conclusions concerning the sensitivity are only qualitatively valid.

Summarizing, we can state that when account is taken of the inevitable damping in the oscillator ($\gamma\neq 0$, $n_T\neq 0$) the QND-measurement method loses its advantages to a considerable degree. At large squeezing parameters the state-squeezing technique is not inferior to that of QND measurements, having the advantage that the quantity measured must not necessarily be a QND variable.

The use of squeezed states can improve the realizability of laboratory experiments on the study and detection of gravitational waves. In Ref. 15 was considered a concrete scheme, in which the radiation and detection were proposed to be effected with electromagnetic fields. The estimate of the sensitivity started out with the fact that it is possible to observe the change of the number of quanta $\sqrt{n}$ against the background of the n quanta present in the cavity-detector. In other words, estimates were used for an oscillator in a coherent state. In this case the technical requirements on the facility are very stringent[15]: the total volume of the system $V=25\times 10^9$ cm^3, characteristic field intensity $E\sim H\sim 3\times 10^5$ G, signal frequency $\omega=2\times 10^{18}$ sec^{-1}, Q-factor of the cavity-detector $Q=\omega\tau^*=7\times 10^{13}$, time of signal accumulation (equal to the relaxation time) $\hat{\tau}=4\times 10^5$ sec. Were it possible to improve the Q of the resonator-detector, e.g., to $Q\approx 8\times 10^{15}$ and produced in it a squeezed state with value $2r/\gamma\approx 10^2$, the sensitivity could be increased during the same accumulation time by a factor of 10, and consequently relax the requirements on the other parameters of the system. These values of Q are apparently perfectly attainable.[16]

It is advantageous to use squeezed states also when detecting monochromatic cosmic radiation, e.g., from a pulsar in the Crab nebula. Experiment is being presently planned on detection of gravitational radiation from this pulsar.[17] The parameters of the resonant antenna are the following: $\nu=60$ Hz, $Q=2\times 10^8$ $M=1400$ kg, antenna length $L=1.65$ m, and the temperature at which the antenna will operate is $T=3\times 10^{-3}$ K. The signal accumulation time τ should amount to $\approx 5\times 10^5$ sec. At a frequency $\nu=60$ Hz, however, the Q of the antenna can possibly be raised to $Q\approx 10^{11}$ (Ref. 16). The relaxation time in such a system will be approximately 9 years. If the technique of squeezed states is used, the time of signal accumulation up to satisfaction of the observability criterion $S/N=1$ could be decreased to approximately 10^3 sec.

APPENDIX

We present first an expression for the quantities containing the operators T and T^+, which do not depend on the initial state of the oscillator. We introduce the notation

$$I_c=\int_0^t e^{\gamma\tau}\operatorname{ch}2r(\tau-t)\,d\tau,\quad I_s=\int_0^t e^{\gamma\tau}\operatorname{sh}2r(\tau-t)\,d\tau,$$

$$E=\int_0^t e^{\gamma\tau}\,d\tau.$$

Then

$$\langle T^+T\rangle=\gamma(n_\tau+{}^1/_2)I_c-{}^1/_2\gamma E,\quad \langle TT^+\rangle=\gamma(n_\tau+{}^1/_2)I_c+{}^1/_2\gamma E,$$
$$\langle T^+T^+\rangle=\langle TT\rangle^*=\gamma(n_\tau+{}^1/_2)e^{-i\psi}I_s,$$
$$\langle(T^+T)^2\rangle-\langle T^+T\rangle^2={}^1/_4\gamma^2[(2n_\tau+1)^2(I_c^2+I_s^2)-E^2].$$

We write out also the expression, which is used in the test for Φ:

$$\Phi(t)=2\omega|f|e^{i\psi/2}e^{\gamma t/2}\left\{-i\cos\left(\varphi-\frac{\psi}{2}\right)\frac{1-e^{(r-\gamma/2)t}}{\gamma-2r}\sin\left(\varphi-\frac{\psi}{2}\right)\times\frac{1-e^{-(r+\gamma/2)t}}{\gamma+2r}\right\}.$$

The general formulas (5)–(9) become simpler if at the initial instant of time the oscillator was at equilibrium with the thermostat. Then

$$\langle X_1(t)\rangle=le^{-\gamma t/2}[\Phi(t)+\Phi^*(t)],$$
$$\langle X_2(t)\rangle=le^{-\gamma t/2}i[\Phi^*(t)-\Phi(t)], \tag{A1}$$

$$\langle n(t)\rangle=e^{-\gamma t}\left[|\Phi(t)|^2+\frac{r}{2}(2n_\tau+1)\left(\frac{e^{-2rt}}{\gamma+2r}-\frac{e^{2rt}}{\gamma-2r}\right)\right]+\frac{n_\tau\gamma^2+2r^2}{\gamma^2-4r^2}. \tag{A2}$$

For the variances we have

$$\delta X_{1,2}(t)=l^2(2n_\tau+1)e^{-\gamma t}[\operatorname{ch}2rt\mp\cos\psi\operatorname{sh}2rt+\gamma(I_c\pm\cos\psi I_s)], \tag{A3}$$

$$\delta n^2(t)=e^{-2\gamma t}\left\{\frac{1}{4}(2n_\tau+1)^2\operatorname{ct}4rt-\frac{1}{4}e^{2\gamma t}+\frac{1}{4}\gamma(2n_\tau+1)^2[\gamma(I_c^2+I_s^2)+e^{2rt}(I_c-I_s)+e^{-2rt}(I_c+I_s)]+4(2n_\tau+1)|f|^2\omega^2\left[\cos^2\left(\varphi-\frac{\psi}{2}\right)\left(\frac{e^{\gamma t/2}-e^{rt}}{\gamma-2r}\right)^2\times(e^{2rt}+\gamma I_c-\gamma I_s)+\sin^2\left(\varphi-\frac{\psi}{2}\right)\left(\frac{e^{\gamma t/2}-e^{-rt}}{\gamma+2r}\right)^2(e^{-2rt}+\gamma I_c+\gamma I_s)\right]\right\}. \tag{A4}$$

[1]V. B. Braginsky, Y. I. Voronsov, and K. S. Thorne, Science **209**, 547 (1980).
[2]C. M. Caves, K. S. Thorne, R. W. Drever, V. D. Sandberg, and M. Zimmerman, Rev. Mod. Phys. **52**, 341 (1980).
[3]L. P. Grishchuk and M. V. Sazhin, Zh. Eksp. Teor. Fiz. **80**, 1249 (1981) [Sov. Phys. JETP **53**, 639 (1981)].
[4]H. P. Yuen, Phys. Rev. **A13**, 2226 (1976). D. Stoler, Phys. Rev. **D1**, 3217 (1970). E. Y. C. Lu, Lett. Novo Cim. **2**, 1241 (1971).
[5]C. M. Caves, Phys. Rev. **D23**, 1693 (1981).
[6]L. P. Grishchuk and M. V. Sazhin, Laboratory Experiment and Quantum Nondemolition Measurements. in: Abstracts, All-Union Conf. on Modern Theoretical and Experimental Problems of Relativity and Gravitation Theory (5th Soviet Gravitation Conf.), Moscow Univ., 1981.
[7]Louisell, Radiation and Noise in Quantum Electronics, McGraw, 1964.
[8]M. Lax, in: Statistical Physics, Phase Transitions, and Superfluidity, M. P. Chretien *et al.*, eds, Gordon and Breach, 1968.
[9]S. M. Rytov, Vvedenie v statisticheskuyu radiofiziku (Introduction to Statistical Radiophysics), Vol. 1, Nauka, 1976.
[10]B. Ya. Zel'dovich, A. M. Perelomov, and Yu. P. Popov, Zh. Eksp. Teor. Fiz. **55**, 589 (1968) [Sov. Phys. JETP **28**, 308 (1969)].
[11]V. N. Sazonov and A. A. Stuchebrukhov, Chem. Phys. **56**, 391 (1981); Zh. Eksp. Teor. Fiz. **84**, 3 (1983) [Sov. Phys. JETP **57**, 1 (1983)].
[12]W. G. Unruh, Phys. Rev. **D18**, 1764 (1978); **D19**, 2888 (1978).
[13]V. V. Dodonov, V. N. Rudenko, and V. I. Man'ko, Zh. Eksp. Teor. Fiz. **78**, 881 (1980) [Sov. Phys. JETP **51**, 443 (1980)].
[14]C. M. Caves, Quantum Nondemolition Measurements, in: Quantum Optics, Experimental Gravitation, and Measurement Theory, P. Meystre and M. O Scully, eds. (to be published).
[15]L. P. Grishchuk and M. V. Sazhin, Zh. Eksp. Teor. Fiz. **68**, 1569 (1975) [Sov. Phys. JETP **41**, 787 (1975)].
[16]V. B. Braginskiĭ, V. P. Mitrofanov, and V. I. Panov, Sistemy s maloĭ dissipatsieĭ (Systems with Low Dissipation), Nauka, 1981.
[17]H. Hirakawa, Detection of Gravitational Radiation from Pulsars, in: Gravitational Radiation, Collapsed Objects, and Exact Solutions, C. Edwards, ed., Lecture Notes in Physics, No. 124 1980.

Translated by J. G. Adashko

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 2 233 951 A1

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
29.09.2010 Bulletin 2010/39

(51) Int Cl.:
G01V 7/00 *(2006.01)*

(21) Application number: **09156549.9**

(22) Date of filing: **27.03.2009**

(84) Designated Contracting States:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO SE SI SK TR
Designated Extension States:
AL BA RS

(71) Applicant: **Elmarx Srl**
38017 Mezzolombardo (IT)

(72) Inventor: **Fontana, Giorgio**
38100, Trento (IT)

(74) Representative: **TBK-Patent**
Bavariaring 4-6
80336 München (DE)

(54) Gravitational wave generator utilizing coupled oscillators and methods for amplification of gravitational waves

(57) It is provided a use of a high vacuum chamber as an amplifier of a gravitational wave. Furthermore, it is provided a device for generating a gravitational wave, comprising a generator (10, 18) of an external energy; a primary vacuum chamber (20) with window (19) transparent for the external energy; a first vacuum pump (21) connected to the primary vacuum chamber (20) and configured to evacuate the primary vacuum chamber (20); a converting material (22, 12) for converting the external energy into gravitational waves, the converting material being disposed inside the primary vacuum chamber (20) such that the external energy illuminates the converting material (22); a second vacuum chamber (16, 23) used as an amplifier of the gravitational wave; and a second vacuum pump (24) connected to the second vacuum chamber (16, 23) and configured to evacuate the primary vacuum chamber (16, 23).

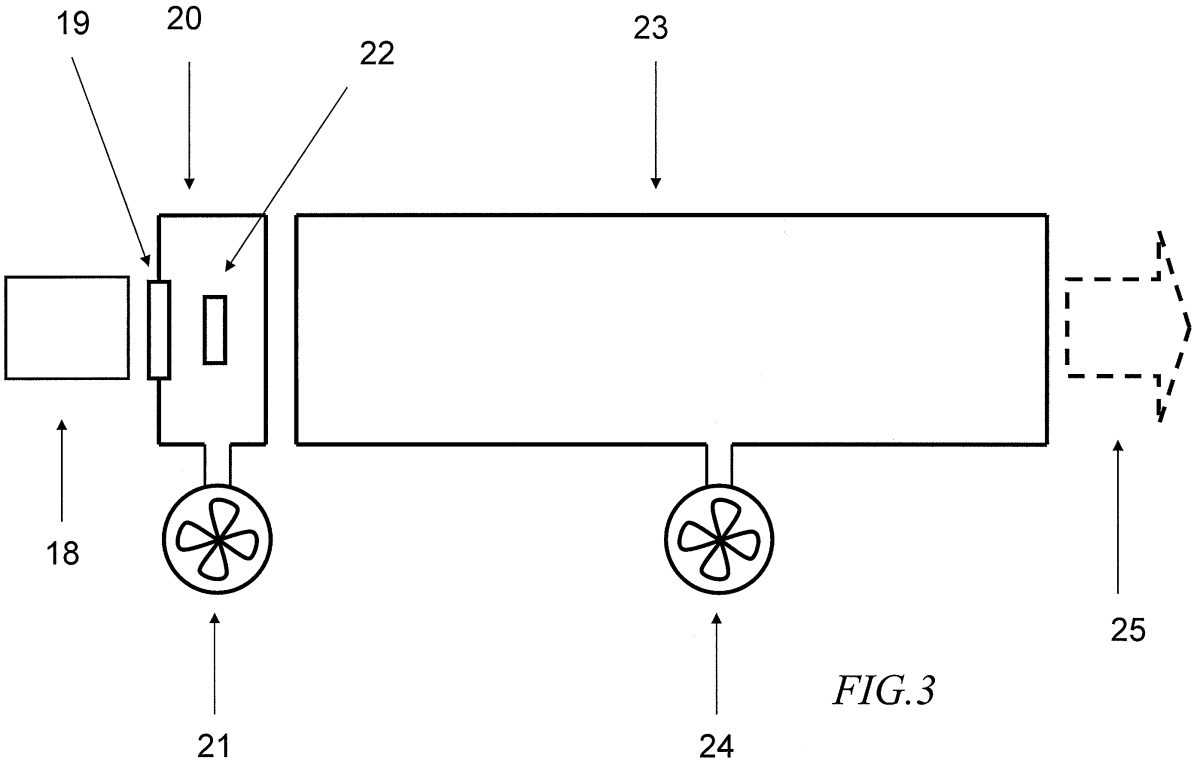


FIG.3

EP 2 233 951 A1

Description

FIELD OF THE INVENTION

[0001] The present invention relates to a use of a high vacuum chamber and a device for generating and amplifying gravitational waves.

BACKGROUND OF THE INVENTION

[0002] Einstein theory of relativity and astronomical observation made by Taylor and Hulse have given convincing proof of the existence of gravitational waves, Taylor and Hulse obtained a Nobel prize for this discovery. In fact the orbital decay of a binary stars have been analyzed using the formalism of Einstein's General Relativity with the result that the stars emit Gravitational Waves because of their orbital motion, this phenomenon causes a predictable decay of the orbit that fits observations.

[0003] Efficient generation of Gravitational Waves (GWs) is a century long endeavor. In this field there is a large number of contributions, where we refer to Landau and Lifshits (1975), Misner, Thorne and Wheeler (1973), Pinto and Rotoli (1988) and Baker (2000, 2004) for a brief exemplification of various traditional approaches.

[0004] The laboratory generation of GW was discussed by Pinto & Rotoli in General Relativity and Gravitational Physics, 1998, World Scientific, Singapore. They found (page 560) terrestrial laboratory GW generation to be "... at the limit of the state of the art ... ", but they did not consider any mechanism of amplification both internal and external like those disclosed by the present document. The general concept of the laboratory generation of GW is to emulate GW generated by energizable celestial systems (rotating binary stars, star explosions, collapse to black holes, etc) as observed by astrophysicists by the use of micro, terrestrial energizable systems.

[0005] Robert M. L. Baker, Jr. in U.S. Pat. No. 6,417,597, shows that a third time derivative or jerk of a mass generates gravitational waves (GW) or produces a quadrupole moment and that the GW energy radiates along the axis of the jerk or if a harmonic oscillation, then radiates in a plane normal to the axis of the oscillation. The force producing the jerk or oscillation can be gravitational attraction, centrifugal, electromagnetic, nuclear, or, in fact, any force. The magnitude of the jerk or, more specifically, the magnitude of the third time derivative of the moment of inertia of the mass squared, determines the magnitude of the generated GW determined, for example by a quadruple approximation as reported in textbooks on General Relativity. The force energizing mechanism can be a particle beam. The particle-beam frequency is that resulting from chopping the particle beam into bunches. The magnitude of the GW power is approximately proportional to the square of the kernel according to the general theory of relativity as discussed for instance in patent U.S. Pat. No. 6,417,597.

[0006] US Pat No. 6,784,591 describes a generator of GW based on submicroscopic energizable elements that does not employ the amplification techniques that are disclosed by the present document.

SUMMARY OF THE INVENTION

[0007] It is an object of the present invention to improve the prior art.

[0008] According to a first aspect of the present invention, there is provided a use of a high vacuum chamber as an amplifier of a gravitational wave.

Further details according to the first aspect may be as follows:

[0009] The pressure in the high vacuum chamber may be below 10^{-3} mbar.

[0010] The ratio of a wavelength of the gravitational wave by a length of the amplifier in a propagation direction of the gravitational wave may be between 10^8 and 10^{16} wavelengths per meter.

[0011] The amplifier may amplify the gravitational wave by a factor of 10 and 10^8 per meter of the amplifier in a propagation direction of the gravitational wave

[0012] The high vacuum chamber may have a length between 0.5 and 1000 meter in a propagation direction of the gravitational wave.

[0013] A non-linear material for the gravitational wave may be disposed in the high vacuum chamber.

[0014] According to a second aspect of the present invention, there is provided a device for generating a gravitational wave, comprising a generator of an external energy; a primary vacuum chamber with window transparent for the external energy; a first vacuum pump connected to the primary vacuum chamber and configured to evacuate the primary vacuum chamber; a converting material for converting the external energy into a gravitational wave, the converting material being disposed inside the primary vacuum chamber such that the external energy illuminates the converting material; a second vacuum chamber used as an amplifier of the gravitational wave according to the first aspect of the invention; and a

second vacuum pump connected to the second vacuum chamber and configured to evacuate the primary vacuum chamber.

Further details according to the second aspect may be as follows:

[0015] The converting material may be excitable into an excited state by the external energy, wherein the excited state comprises a first oscillator of a first characteristic frequency comprising a mass and emitting the gravitational wave, and a second oscillator coupled with the first oscillator, and wherein a characteristic frequency of the second oscillator is smaller than the first characteristic frequency.

[0016] The first oscillator may be not electrically charged.

[0017] Each of the first and the second oscillator may comprise a subnuclear structure.

[0018] The material in a state other than the excited state may comprise a stable or metastable nucleus, and the external energy may include at least one of an electromagnetic wave and an elementary particle.

[0019] The electromagnetic wave may comprise at least one of X-ray and gamma ray, and wherein the elementary particle may comprise at least one of a neutron, a proton, an electron, a positron, and an antiproton.

[0020] A number of protons in the nucleus may be between 90 and 125.

[0021] The converting material may comprise Gadolinium.

[0022] The first oscillator may be one of a superconducting electron pair and a crystal or molecular excitation in a frictionless paired state, the second oscillator may be one of a molecular structure, an atom, and a crystal lattice, and the external energy may be at least one of an electromagnetic excitation, a thermal excitation, and an acoustic excitation.

[0023] The first and the second oscillator may be coupled by at least one further oscillator.

[0024] It is to be understood that any of the above modifications can be applied singly or in combination to the respective aspects to which they refer, unless they are explicitly stated as excluding alternatives.

[0025] Further details, features, objects, and advantages are apparent from the following detailed description of the preferred embodiments of the present invention which is to be taken in conjunction with the appended drawings, in which

Fig. 1 represents a detailed description of the nuclear model and the dynamical model;

Fig. 2 shows an apparatus for generating and amplifying gravitational waves.

FIG. 3 shows another apparatus for generating and amplifying gravitational waves; and

FIG. 4 shows still another apparatus for generating and amplifying gravitational waves.

DETAILED DESCRIPTION OF CERTAIN EMBODIMENTS

[0026] Herein below, certain embodiments of the present invention are described in detail with reference to the accompanying drawings, wherein the features of the embodiments can be freely combined with each other unless otherwise described. However, it is to be expressly understood that the description of certain embodiments is given for by way of example only, and that it is by no way intended to be understood as limiting the invention to the disclosed details.

[0027] Moreover, it is to be understood that the apparatus is configured to perform the corresponding method, although in some cases only the apparatus or only the method are described.

[0028] In a gravitational wave generating device, a material may be used in which external or internally produced energy is transferred from a main oscillator to other oscillators with different resonant frequencies coupled one another and with the main oscillator. Because of the interconnected cascade of oscillators and possible non linear coupling, the efficiency of the gravitational wave emission can be greatly increased.

[0029] At the atomic and subatomic scale, it is well known that all material particles may have mass or, alternatively, a combination of mass and electric charge.
Using particles with mass and charge, like protons, it is possible to excite those particles with external energy, like lasers and/or particle beams or with internal energy, like a chemical or nuclear reaction. As a consequence of the excitation the particles emit electromagnetic waves as a function of acceleration and emit gravitational waves as a function of the derivative (change in time) of acceleration, named jerk. Therefore the higher the frequency the higher is the emission of gravitational waves compared to the emission of electromagnetic waves. Because of the weakness of gravity with respect to electromagnetism, it is known from theory (Halpern and Laurent) that the emission of electromagnetic waves overwhelms the emission of gravitational waves in most, if not all, known situations.

[0030] For generating gravitational wave, the ideal approach is to use emitting particles that have large mass and no electric charge, they are neutral particles like neutrons, neutrinos, etc, they will emit gravitational waves and will not emit electromagnetic waves.

[0031] A method is disclosed that permit to put in motion massive neutral particles coupled to electrically charged particles with a great additional increment of the frequency, therefore obtaining an upgeared converter, which increases the emission of gravitational waves by many orders of magnitude respect to a direct, non upgeared, converter.

[0032] In addition to the above embodiment the disclosed concept includes all systems of coupled oscillators that are intended to emit gravitational waves using the same principles.

[0033] An embodiment of the invention relates to the generation and amplification of gravitational waves that can be modulated, directed, collimated and focused to be utilized for communications, space propulsion, air propulsion, near-ground propulsion, marine propulsion, sub-marine propulsion, through earth digging, material processing nuclear processes, including igniting nuclear fusion for energy production, testing physical-cosmological concepts and conjectures. More particularly an embodiment relates to the generation of gravitational waves (GW) by the interaction of elements (molecules, atoms, nuclei, nuclear particles, electrons, photons, antiprotons, etc.) through a dynamical process in which the different elements exchange energy through a process in which the final emitter of GW reaches a frequency higher than the input frequency or the frequency of the energy internally generated by the internal energy source, thus improving the emission of GW to a level that make this emission useable for practical applications. Because the improvement according to the embodiment may be over 20 orders of magnitude (10^{20}) with respect to previous techniques, the present invention is critical for the practical use of GWs.

[0034] An embodiment also relates to the use a non-linear solution of GWs propagation to obtain an amplification of the GWs produced by the above described generator to reach the amplitudes required for the cited practical applications. The present invention adopts two additional methods of amplification, one is the Holonomy, spin-orbit coupling, that realizes what is already known as parametric amplification using non-linear dynamics of coupled oscillator, the other is the non-linear dynamics of GWs themselves that is a consequence of General Relativity (GR) the best model of gravity developed till now. The amplification can be obtained by a tube or box in which a high vacuum is present to permit the undisturbed non-linear growth of the GW.

[0035] According to an embodiment of the present invention it is provided the generation of gravitational waves (GW) caused by the interaction of submicroscopic (molecules, atoms, nuclei, nuclear particles, electrons, alpha particles, neutrons, dineutrons, photons, etc.) energizing and energizable elements, nuclear cycles including instability cycles, chaos bifurcations, holonomy processes and non linear processes, in which a cascade of energy transfers ends on electrically neutral elements that emit the gravitational waves. This is the first generations and amplification mechanism. An embodiment includes a non linear propagation mode for gravitational waves that can provide further amplification of the amplitude of the wave. The combination of the different contribution enables reaching the amplitudes required for practical applications.

[0036] The weak point of all traditional approaches of generating gravitational waves is that the emitting system, whatever it is, operates at the same frequency of the excitation energy. The end result is that the frequency of the emitted GWs is twice the frequency of the excitation energy and the conversion efficiency is extremely low. For instance, for an electro-magnetic (EM) to GWs converter (Pinto and Rotoli, 1988), no practical cavity can have a sufficiently high Q or can host an EM field strong enough to produce detectable GWs according to traditional generator and detector configurations. Going back to the elementary basic formula of the power P emitted with gravitational waves by a "rod" with mass *m,* length *L,* rotating ω times a second, we have:

$$P = \frac{2}{45}\frac{m^2 L^4 \omega^6}{(c^5/G)}, \qquad (1)$$

where *c* is the speed of light and *G* is Newton's gravitational constant. In some embodiments we achieve that the power of gravitational waves excited is at the highest possible frequency, with the goal that future gravitational wave optics will permit collimation and focusing for technological applications (Baker, 2005) in which those parameters are most relevant.

[0037] Looking at the formula we observe that for identical systems, the power emitted is proportional to the sixth power of the frequency. Therefore a "gear-up" mechanism between electromagnetic excitation and the "rod" is what is required to increase the power emitted in GWs. Frequency multiplication can be obtained by using "gearboxes" like those we use in cars. Nano-mechanical gearboxes could be made, but all of them still have electromagnetic coupling at the "fast" pinion, and they need a torque reaction coupling to operate. Both limitations need to be addressed to define a new converter that is highly efficient and capable of enabling new technologies.

[0038] The main requirements for such an elementary converter are:

1. Small size and highly reproducible structure with highest possible operating frequency leading to GWs,

2. Nuclear density GWs emitter for best impedance matching to space-time,

3. Capability to be used in large numbers and in synchronized modes,

4. High frequency multiplication ratio,

5. Electrically charged low frequency interface for easy energization,

6. Electrically neutral high frequency interface to avoid re-emission of EM waves,

7. Massive and with high gravitational quadrupole moment high frequency interface for GWs emission,

8. Capability to be used without external torque reaction (it should be based on autonomous coupled oscillators and spin-orbit coupling).

[0039] The requirements point to nuclear or molecular systems as the preferred choice. We therefore recall that the spheron model of the nucleus (Pauling, 1965) may suggest that some nuclei do indeed provide the full spectrum of features that is required to practically build an efficient EM to GWs converter. The spheron model of the nucleus has been proposed by the two times Nobel Laureate Linus Pauling in 1965 (Pauling, 1965) as a simple model capable of explaining many properties of the different nuclei. The spheron model was proposed after the drop model and the shell model and is probably the most practical and intuitive model of the nucleus capable of producing realistic predictions as an alternative to brute force quantum mechanical calculations.

[0040] Figure 1 represents a detailed description of the nuclear model and the dynamical model. Reference number **6** indicates the core of the nucleus that is composed by one or more spherons **1** and **5**. Each spheron can be a proton, a neutron, a dineutron **8** a triton or an alpha particle like in the specific spheron **5.** The alpha particle α is composed by two neutron **3** and two protons **4.** In the discolsesd invention the dineutron **8** orbits on the surface of the core of nucleus along an orbit **7,** the dineutron **8** orbits and spins **2** around its own axis producing gravitational waves **9.** Those subnuclear elements can be present as a real particle or as a hole for the device to operate properly.

[0041] The spheron model is characterized by clusters of protons and neutrons in the nucleus named spherons, that are alpha particle (helions), tritons and dineutrons: they are aggregates of protons and neutrons in localized 1*s* orbitals. Spherons may be present in the nucleus as "particles" and also as "holes".

[0042] Pauling shows that concentric layers of spherons allow the simple derivation of a subsubshell occupancy diagram and explanation of magic numbers (Pauling, 1965). The effective existence of dineutrons was discussed by Kamal and Parker (1991) and other authors with the clear conclusion that dineutrons are real objects.

[0043] Two qualities of the spheron model are of great relevance for GWs generation. The first one is that heavy isomers with high angular momentum are represented by a central core with magic number of protons and neutrons that does not rotate plus one or more spherons orbiting the core (Pauling, 1990). As the dynamics is driven by the strong nuclear force, the process is a proper gliding in an orbit. In one of his papers Pauling also suggests that $^{204}_{80}Hg_{124}$ has an α particle hole revolving in the sphere of the doubly magic $^{208}_{82}Pb_{126}$, and $^{206}_{82}Pb_{124}$ has a dineutron hole revolving in the sphere (Pauling, 1990).

[0044] The second quality of the spheron model is that each orbiting spheron may have its own rotational state that can be excited while gliding the magic number core. This is relevant if the spheron is a dineutron. In fact the dineutron obviously has nuclear density, it is electrically neutral and the two magnetic moments of the component neutrons are antiparallel, making it an ideal emitter of GWs, possibly the best microscopic emitter found so far.

[0045] Quantum mechanical analysis of nuclei is simpler for nuclei with less than ten nucleons, and various efforts are ongoing to produce detailed models. Itagaki et al. (2002) have studied the nuclear configuration of ^{10}Be, that has a couple of neutrons orbiting the α-α core; a lower energy state is found that has a bound neutron *n*-*n* (dineutron) structure orbiting the core, a higher energy state has the two neutrons separately orbiting the core in opposite directions (see page 044302-4 of Itagaki, 2002). Modem quantum mechanical analysis confirms the Pauling's view about the existence of nuclear substructures (Pauling, 1990). Nuclear substructures clearly have a rich dynamics and with our interdisciplinary view, we recognize that the dineutron rotational state excited by its orbital angular momentum around the core may enable a very efficient mechanism for the generation of GWs. Energy is transferred from orbital motion to the spin of the dineutrons by relativistic spin-orbit coupling. The interaction is mostly dynamical in nature, not involving nuclear reactions, and following the rules of nonlinearly coupled dynamical systems.

[0046] If the excitation energy is external, it can be a particle and/or electromagnetic and/or gravitational radiation that hit the nucleus and make it oscillate at some frequency; as a consequence the dineutron/s orbit the nucleus and also spins at a much higher frequency respect to the excitation one. As the emission of gravitational waves if proportional to the sixth power of frequency there will be a great increase in emitted power. Being electrically neutral the dineutrons will

not emit electromagnetic waves.

[0047] By adopting a collection of converting nuclei excited by coherent radiation a beamed coherent emission is possible, with additional gain due to coherence.

[0048] If the excitation energy is mostly internal, an external trigger particle, for example a proton is injected in the selected nucleus. According to theoretical analysis by Binder (2008), some heavy nuclei can enter a regime of dynamical instability related to chaos by a small increase of proton number Z; let's choose for example, but not limited to, Z=115. With the additional proton Z=115 may become a nucleus in the range Z=116, entering a region of instability, becoming unstable and highly energized trough Zero Point electromagnetic Fluctuations (ZPF). Afterwards a proton of the nucleus becomes a neutron plus a positron, following the usual beta+ decay channel. The positron is emitted and the neutron participates to the efficient emission of Gravitational Waves like we have in the externally excited converter. The emission of Gravitational Waves de-energizes the nucleus, the neutron is then released and the nucleus is ready for another cycle.

[0049] A vacuum enclosure permits further amplification because of non-liner phenomena related to Gravitational Waves, as determined by Fontana (2008).

[0050] A material according to an embodiment comprises one or more sets of coupled oscillators. Each set has a low frequency mode, which is excited/depleted by internal or external energy or by the high frequency mode, and a high frequency mode that is excited/depleted by internal energy, external energy or by the low frequency mode, being this last one the most relevant excitation path. The low frequency mode may couple to all known forces and fields to some degree, the high frequency mode couples mostly only to Gravitational Waves.

[0051] With external excitation an embodiment operates as converter from electromagnetic waves (EMW) to Gravitational Waves (GW); with internal excitation an embodiment operates as a generator of Gravitational Waves.

[0052] The invention is based on interdisciplinary application of nuclear dynamics within the Spheron model of the nucleus, non-linear dynamics of coupled oscillator, parametric amplification, holonomy, holonomy and nuclear instability in chaos theory, non-linear propagation of Gravitational Waves. It is based on the mainstream theories and on nuclei, nevertheless the same principles may be applied to similar systems of different nature like paired electrons in crystals, superconductors, polarons and other excitations in the solid state or plasmas that could manifest properties that may allow the here described phenomena to manifest as described in this document.

[0053] **Figure 2** shows a device for generating GW according to an embodiment of the present invention. The target plurality of nuclei **12** is excited or driven into instability by an external wave or exciting particles **11** produced by the generator **10.** The gravitational wave generating material **12** produces gravitational waves **13, 14, 15.** The gravitational wave **15** is amplified by the gravitational wave amplifying device **16** and emitted as high amplitude gravitational wave **17.**

[0054] More details of such an embodiment are shown in **Figure 3.** In the embodiment, an X-ray laser **18** sends a beam of X-ray electromagnetic radiation through an X-ray transparent window **19** inside a primary vacuum chamber **20** evacuated by the vacuum pump **21**. The X-ray illuminates the converting material **22** that as a consequence produces the gravitational waves. Gravitational waves enter the second vacuum chamber **23** whose pressure and/or gas content can be controlled by pump **24**. amplified gravitational waves **25** are emitted along the main axis of the vacuum chamber **23**. The selection of the window material of the second vacuum chamber is not critical because many known materials are transparent to gravitational waves.

[0055] The pressure in the primary and second vacuum chamber is below 10^{-3} mbar, preferably below 10^{-6} mbar, more preferably below 10^{-10} mbar. The x-ray energy is between 0.1 and 120 keV, preferably between 1 and 100 keV, more preferably between 30 and 70 keV.

Figure 4 shows another embodiment. Compared with the embodiment of Figure 3, the X-ray laser is replaced by a particle accelerator 26 which is in the same vacuum chamber as the converting material 22.

[0056] According to Binder (2008), that determined some general rules for spin-orbit coupling in nuclei, the spinning frequency of the dineutrons is about 8939 times higher than the excitation frequency of the core of the nucleus. The frequency ratio between electromagnetic waves and emitted Gravitational Waves is 2*8939=17878 times higher than the excitation frequency of the core of the nucleus. Thus we have defined a $8939^{GW}/1^{EMw}$ converter. According to the calculations of Halpern and Laurent (1964) The net gain in power emitted respect to a traditional $1^{GW}:1^{EMW}$ converter is therefore $8939^6=5*10^{23}$, with an additional gain of $2*10^7$ that can be obtained by a synchronized operation of coherent oscillators or by a high Q resonator, the unitary energy balance and 100% conversion efficiency, can be obtained. Please observe that the energy balance for the $1^{GW}:1^{EMW}$ converter is 10^{31} in favor to the electromagnetic interaction, which means that with the $1^{GW}:1^{EMW}$ converter a "graviton" is emitted every 10^{31} "photons" as described by Halpern and Laurent (1964). Instead with the disclosed invention $8939^{GW}/1^{EMw}$ converter, a "graviton" is emitted every $2*10^7$ "photons".

[0057] An ensemble of $2*10^7$ or more nuclei is required to operate with unitary (100%) conversion efficiency, that is about fractions of a picogram of a suitable element. According to Pauling and the periodic table of the elements a suitable element could be Gadolinium.

Coherent emission is possible by exciting the nuclei with coherent radiation, for instance in the X-ray frequency range and letting the stimulated emission of GW to build-up as described by Halpern and Laurent (1964). Frequency multipli-

cation in non linear coupled oscillators can be described as follows:

SPINNING AT HIGH HARMONIC FREQUENCIES IN NUCLEAR APPLICATIONS

[0058] Rolling, gliding, and revolving in the sphere provides for holonomy effects. In classical mechanics we have holonomy effects (Marsden, 1973), especially if the solid-body rotation angle is connected to the actual solid angle that the body-fixed axis describes as the body performs a conical motion (Ishlinskii, 1952, Malykin, 1999). There are mechanical toys modeling the dynamics of spin-precession coupling: the function of the Gyrotwister or Powerball (Mishler, 1973) is based on a mechanical gear-type coupling between precession and spin phase: if the spin axis rolls in the bearing (on the rim of the groove by friction) and the device has enough spin, the onset of linear spin-precession coupling enables to control spin by precession. If the conic motion of the external stimulation is such, that the number N of axis rotations for one precession period is reduced by j, the device auto-tunes recurrently and follows the external holonomy change. Now it can be powered-up to high spin frequencies with precession frequency multiplier N/j. Near dynamic equilibrium we get curves of constant precession (Darboux) in the Frenet description, which are rolling cone paths showing the dynamics of a precession axis pendulum.

[0059] "Magic Angle Precession" (MAP) (Binder, 2008) describes a special chaotic behavior in the precession angle of a precessing rotor (gyroscope), where the rotor angular velocity is linearly coupled by (an)holonomy to the precession angular velocity and angle. In the mechanical picture a linear coupling between spin and precession angular velocity provides for rolling cone paths and allows spinning up and controlling the rotor simply by forcing precession at special quantum magic precession angles. The principle can be experimentally verified with devices, where spin and precession of a rotor are coupled by a conic rolling mechanism. This linear relation models a recurrent chaotic holonomy, where precession as the effect of holonomy is also a cause of holonomy linearly related to the rotator spin. Gravitational waves can be regarded as a modulated holonomy signals obeying a hyperbolic wave equation. In a more recent approach, it is shown that a pure geometric attractor can promote a local hyperbolic (vortex) geodesic flow on a rotating spherically curved surface, where the vortex angular velocity is considerably higher than the angular velocity on the sphere (Binder, 2008). Here, we regard an ideal system of rotated rotations, where spherical harmonics store angular momentum while hyperbolic harmonics (traveling waves) extract or support angular momentum. Due to different curvature signs we have at the boundary a holonomy discontinuity, which could provide from a signal theoretic point of view for impedance jumps capsulating the spherical harmonics and reflecting external waves.

[0060] From referenced papers we know about magic angle precession constellations reducing the mismatch to zero and allowing for a lossless external angular momentum exchange. At special precession angles it is possible to inject and extract angular momentum from precession to and from the rotor spin, allowing spinning up the rotator to very high frequencies controlled by a slow precession frequency. Due to the apparent coupling between spin and precession it is a strictly nonlinear dynamics supporting fast auto-tuning and self-stabilization capabilities. The recurrent situation balances the linear coupling by a transcendental holonomy spin coupling effect. The emerging magic precession/nutation angle/phase patterns are singularities given by the intersections between linear and transcendental couplings. Near these points very strong and lossless angular momentum currents crossing the boundary between hyperbolic precession space and the spherical rotator space are observed. The strange attractor balancing the linear and holonomy spin-currents living on the tangential cone connects the sphere with the hyperbolic space, where the resulting precession or twister dynamics can be modeled.

[0061] With synergetic effects given by a neural-type recursive network based on coupled MAP elements, a strong long range spin-network synchronization dynamics should be possible (Binder, 2008). It has been proposed that this holonomy attractor could integrate a large number of quantum spins from Hopf fibrations and correspondent magnetic monopoles into one global higher-dimensional spin, where the local coupling network provides for electroweak and strong interactions (Binder, 2008).

[0062] With three or six MAP elements we get a rational exact solution without chaotic transcendental part and SU(3) symmetry, where the Dirac monopole charge condition $NM = \pm 1/\alpha$ fits very well to $Z = j/J = \pm 1/3, \pm 2/3$, the quark charge (Binder, 2008). This spin has some serious support in literature (Franklin, 1969). Finally, more complex networks near Z=115 could explain the heavy nuclei instability (Binder, 2003, 2009).

[0063] With the support of the Spheron model and MAP dynamics it is possible to identify a nucleus that could convert EMW radiation to GWs with high efficiency such as nuclei with a heavy core with highly relativistic dineutrons capable of strong spin-orbit coupling.

[0064] In the following analysis we compare the GWs emission properties of dineutrons dynamically coupled to a suitable core to a similar (according to Equation 1) GWs source that is the deuterium nucleus. The dineutron will be coupled to the electromagnetic field via the core and MAP dynamics, the deuterium nucleus will be directly coupled to the electromagnetic field.

NON LINEAR DYNAMICS OF DINEUTRONS

[0065] After recognizing that a nuclear model is suitable for producing GWs through angular frequency high order harmonics interplay between a charged body (the core) and an electrically neutral "dumbell" (the dineutron) a quantum mechanical relationship between the two frequencies, applicable to heavy nuclei, must be found. At STAIF 2008 Bernd Binder described MAP (Magic Angle Precession) dynamics for nuclei (Binder, 2008). The study of MAP dynamics applied to nuclei has provided values for some critical parameters capable of describing the efficiency of the EMW to GWs converter. Bernd's analysis is a powerful tool that we here apply to the spheron - core/dineutron model of suitable nuclei. The method involves non-linear dynamics of coupled oscillators. Coupled oscillators with non-linear terms exhibit a complex dynamics that goes from "regular motion" through "star patterns" to "chaotic motion" as a function of excitation energy. The "start pattern" dynamics is very peculiar for our approach because energy goes back and forth from a high frequency oscillation to a low frequency oscillation on different linear combination of physical coordinates. The model has widespread applications, for instance Earth atmosphere may be modeled by such a system of coupled oscillators (Lynch, 2002). Energy transfer between modes is also an example of parametric amplification and de-amplification. In such case the traditional approach is based on the Manley and Rowe equations (Brown, 1965). With a set of frequencies in the source and in the sink written as $f_{mn}=mf_1+nf_2$, with m, n integers and f_2/f_1 irrational, being P_{mn} the power at f_{mn}, positive if it enters the system and negative if it leaves the system, the Manley-Rowe equations are:

$$\sum_{m,n}\frac{mP_{m,n}}{mf_1+nf_2}=\sum_{m,n}\frac{nP_{m,n}}{mf_1+nf_2}=0 \qquad (1)$$

[0066] As a particular case, if two oscillators at frequency f_0 and Nf_0, where N can be irrational, are coupled by a non-linear non-dissipating element, such as a non-linear elastic constant, it is possible that energy goes from one oscillator to the other periodically, the direction of the energy transfer depends on the phase relationship between the phase of the higher frequency oscillator and the respective harmonic of the lower frequency oscillator, a changing phase is determined by the fact that N can be irrational, i.e. it is the nearest rational number plus a small irrational frequency difference (a changing phase). This is what Lynch describes as the star-pattern phenomenon. It is a form of degenerated parametric amplifier in which the mode at f_1 amplifies the mode at f_2 at the expense of its own energy, until is completely depleted, then the energy flows the other direction if no other mechanism can remove it. Usually, the physical coordinates of the system are linear combination of the orthogonal coordinates of the two modes. With higher initial energy the same system may enter the regime of chaotic motion. Chaotic motion is usually disruptive for nuclear systems because as soon as some binding energies are overcome the nucleus is interested by fission, which destroys the mechanism.

[0067] Another interesting example of parametric coupling applied to gravitating systems is satellite librations (Klemperer and Baker, 1957), in this case the time-scale of the phenomenon is much longer.

[0068] As soon as a dynamical system exhibits these properties (like non-linear or relativistic spin-orbit coupling), the periodic energy exchange takes place. In suitable nuclei, when energy peaks on the high frequency spinning dineutron, it can be emitted as GWs. The emission of a quantum of GWs removes power from the system and a new cycle can begin, thus partially depleting the "star pattern", whose dynamics do indeed operate to keep the system synchronized. If energy input is kept below some threshold, the converter can safely operate. With Binder's notation, parameter N is the integral number of nutations, or orbital waves (dineutron rotations on its axis) for one rotation of the base plane (one glide-orbit around the core, induced by EMW excitation of the core). Analyzing the Sommerfeld spin-orbit fine structure coupling conditions for M=137, best fit is obtained by averaging results with N about equal to 8939. N=8939 has been primarily applied to the numerical study of heavy nuclei bifurcation (Binder, 2008). For the analysis of the "gear up" mechanism in heavy nuclei we adopt N=8939 as gear ratio i.e. the "frequency ratio", an addition factor of 2 is required for the GWs frequency.

ESTIMATION OF THE CONVERSION EFFICIENCY

[0069] To estimate the conversion efficiency of the arrangement, the nucleus can be treated like a quantum system for which a ratio r between the emission probability for gravitons and photons from equal structures can be computed. Halpern and Laurent computed in their seminal paper (Halpern, 1964) the above mentioned ratio for electrons $r\sim10^{-43}$, for protons $r\sim10^{-36}$, and for larger molecular structures or heavy nuclei that are only slightly ionized $r\sim10^{-31}$.

[0070] If we introduce the concept that the EM excitation frequency is lower than the frequency of the high frequency GWs (HFGW) emission and the HFGWs emitting subsystem is "geared-up" by an electrically neutral mechanism, according to equation (1) r is increased by the sixth power of the frequency ratio, that is $8939^6=5*10^{23}$. Therefore for heavy

nuclei that are only slightly ionized and for which the MAP coupling effect ends on neutron structures $r \sim 5 \cdot 10^{-8}$.

[0071] To increase the interaction of the nucleus with the EMW field without affecting the emission of GWs, it is possible to adopt an EMW resonator tuned on the EMW absorption frequency of the nucleus. In an embodiment, a superconducting cavity with $Q=2 \cdot 10^{7}$, capable of accumulating energy till dissipation makes it leak to the external environment, makes unitary, which is the goal of any efficient converter from electromagnetic radiation to GWs. High power superconducting cavities with intrinsic Q exceeding 10^{10} are described by Padamsee, 2005, therefore nuclei designed to provide MAP coupling effect to dineutrons are excited with EM waves obtaining nearly complete conversion of EM energy to GWs. In some other embodiments photons may interact with sequences of many converting nuclei, to enhance the conversion probability, see Fontana and Baker (2006). In an embodiment, the traveling-wave converter is designed to perform coherent conversion, thus maintaining the spatial and temporal coherence of the EM excitation, mimicking the operation of non-linear crystals employed for optical frequency multiplication of Laser light. Discrepancies between the phase velocities of the two kinds of radiation are minimized by the geometry of the converter (Fontana and Baker, 2006).

[0072] The process can be described as a continuous excitation of the rotational state of the dineutron; alternatively it can be described as a transition to a higher energy state with unbound neutrons to a lower energy state with bound neutrons (pairing-condensation) (Itagaki et al., 2002).

ESTIMATION OF THE MAXIMUM OUTPUT POWER

[0073] In this section the maximum permissible output power in GWs per unit of converting material is estimated under a number of reasonable conditions. In an embodiment analyzed herinafter the dineutrons do not dissociate, therefore no nuclear reaction takes place. The estimate will be based on maximum centrifugal force counteracted by the strong nuclear force to determine the maximum angular velocity and therefore the power emitted in GWs. The classical approximation should be sufficient for an order of magnitude estimate. In some embodiments, the power is limited without overcoming the limit. Overcome the limit is possible by the usual fission phenomenon.

[0074] The measured binding energy of the deuteron nucleus is 2.226 MeV, for dineutrons the binding energy depends on the structure of the α-cluster core, because no binding is possible in isolated n-n systems. The early calculations made by Pauling himself give values of about 1.4 MeV (Pauling, 1967) for dineutrons belonging to heavy nuclei. The distance between the two neutrons of the couple is 2-3 fm, as reported by recent studies in neutron rich light nuclei (Kanada-En'yo, 2007), which also suggest that the α-cluster core is a Bose condensate of α-particles, the same kind of condensation apply to dineutrons for nuclei that have more than one dineutron orbiting the core. Adopting equation 1, the maximum power emitted by a single dineutron as GW is of the order of 10^{-33} W at frequencies of the order of 10^{22} Hz, that belong to the γ-ray region of the spectrum. A kilogram of dineutrons consists of $N \approx 3 \cdot 10^{26}$ dineutrons and the power emitted with incoherent summation (only applicable to spontaneous emission) is of the order of microwatts. In an embodiment with coherent amplitude summation is possible, the (Number of Emitters)2 factor pushes the maximum power to about 10^{20} W per kilogram of dineutrons, therefore the converter produces powerful beams of GWs according to the degree of coherence that can be achieved experimentally and the geometric configuration of the converter. The effective output power will be obviously less than the input power, that for the internally energized converter, includes Zero Point electromagnetic Fluctuations (ZPF).

[0075] The here presented concept of EM wave to GW converter with high order harmonic generation exploits the quantum relationship between the frequency excitation of EM energy and the frequency of the generated GWs. The process can be made coherent for GW output, with space and time coherent excitation in the X-ray region of the EM spectrum. In an embodiment, the excitation system is an X-ray laser.

[0076] If the excitation energy is mostly internal, with an external trigger particle, for example a proton, which is injected in the selected nucleus, some heavy nuclei can enter a regime of dynamical instability by a small increase of Z, let's choose for example, but not limited to, Z=115. With the additional proton Z=115 may become a nucleus in the range Z=116, entering a region of instability, becoming unstable and highly energized trough Zero Point electromagnetic Fluctuations (ZPF). Afterwards an energetic proton of the nucleus becomes a neutron plus a positron and a neutrino, following the usual beta+ decay channel. The positron and the neutrino are emitted and the neutron participates to the efficient emission of Gravitational Waves like we have in the externally excited converter. The emission of Gravitational Waves de-energizes the dineutron, the neutron is then released according to the core/dineutron energy balance already disclosed, and the nucleus is ready for another cycle.

AMPLIFICATION OF GRAVITATIONAL WAVES

[0077] When a strong GW propagates undisturbed in high vacuum it may grow to higher amplitude because of the following considerations:

[0078] As a simple rewriting of the metric of a GW of General Relativity it is possible to show that Gravitation can be easily modeled by resorting to the analogy to optical propagation in a three-space and introducing a gravitational refractive

index of the Hyperspace named *n* (Almeida, 2001):

$$c^2(dt)^2 = n^2\left[(d\tau)^2 + (dx)^2 + (dy)^2 + (dz)^2\right], \quad n \neq 0, \tag{2}$$

after switching to spherical coordinates with $d\varphi=0$ and $d\theta=0$, writing $ds^2=n(d\tau^2)$ and substituting:

$$\frac{1}{n} = \left(1 - \frac{2Gm}{r}\right), \tag{3}$$

we find the Schwarzschild's metric of General Relativity (Almeida, 2001), with G the gravitational constant, *m* is mass and *r* distance. Therefore there exist a formal link between General Relativity and the Theory of Gravity for the Hyperspace that is called Four Dimensional Optics (4DO).

[0079] For the wave described by Equation (2), with $a<<1$ we express Equation (2) as:

$$c^2(dt)^2 = n_\tau^2(d\tau)^2 + n_x^2(dx)^2 + n_y^2(dy)^2 + n_z^2(dz)^2, \tag{4}$$

Comparing Equation (1) with Equation (4), it is obvious that the weak gravitational wave is a wave of refractive index of the Hyperspace, which propagates at the speed of light in four dimensions. More generally, assuming isotropic propagation, a generic wave equation for weak waves traveling at speed c can be written:

$$\frac{1}{c^2}\frac{\partial^2 n}{\partial t^2} = \nabla^2 n, \tag{5}$$

Equation (5) is characterized by time *t* defined as an evolution variable as in the equations of electrodynamics and by having four space dimensions.

[0080] The linear wave Equation (5) can be integrated with results that will have integration constants. In fact in the flat background, the refractive index is unitary; this value must be added to the solutions of Equation (5). If the background value can be ignored in the derivatives, it must be included in the expression of speed. The nonlinear wave equation that includes the effect of the refractive index on speed is therefore:

$$\frac{(n+1)^2}{c^2}\frac{\partial^2 n}{\partial t^2} = \nabla^2 n, \tag{6}$$

whose nonlinear behavior expresses the non-linearity of gravitational waves (Christodoulou, 1991). The background value of *n* must be added to the solutions of Equation (6).

[0081] Equation (5) has the usual wave solutions for any value of *n*, instead Equation (6) has the usual wave solutions if $n^2 <<1$.

[0082] The solutions for weak waves in a unitary background may come from separations of variables in spherical (hyperspherical) coordinates. Writing the solutions of Equation (5) in the form $n(r, t)= T(t) R(r)$, the individual equations are:

$$r\frac{d^2R}{dr^2} + (i-1)\frac{dR}{dr} + k^2 rR = 0; \qquad \frac{d^2T}{dt^2} + \omega^2 T = 0, \tag{7}$$

where *i* is the space dimensionality of the problem, *k* has units of 1/distance and ω has units of 1/time. With *i*=4, they are waves in four dimensions. Please note that with odd dimensionality (*i*=1, 3, 5, ...) the traveling wave solutions have speed *c*; instead for even dimensionality the traveling wave solutions have a spread of speeds from zero to a maximum (Brown, 2007). This last result applies to gravitational waves in the Hyperspace.

[0083] Qualitatively, according to Equation (6), if a beam of gravitational waves is more intense at the center respect to the periphery, then the central part of the beam travels at lower four-speed, leading to self-focusing, in agreement with the classical GW treatment of the problem (Corkill and Stewart, 1983; Ferrari,1988).

[0084] The unitary background refractive index of Equation (6) needs some explanation as it could be attributed to a background gravitational wave that cosmologically supports our space-time. In fact removing the background the non-linear wave equation is:

$$\frac{n^2}{c^2}\frac{\partial^2 n}{\partial t^2} = \nabla^2 n\,, \tag{8}$$

[0085] Starting from the wave equation in *i* space dimensions (*i*=4 for the Hyperspace) and adopting again *n(r, t) = T (t) R(r)*:

$$T\frac{d^2R}{dr^2} + (\frac{i-1}{r})T\frac{dR}{dr} = R\frac{R^2T^2}{c^2}\frac{d^2T}{dt^2}\,, \tag{9}$$

we separate the variables:

$$\frac{1}{R^3}\frac{d^2R}{dr^2} + \frac{1}{R^3}(\frac{i-1}{r})\frac{dR}{dr} = \kappa = \frac{T}{c^2}\frac{d^2T}{dt^2}\,, \tag{10}$$

rewriting, the left hand equation is:

$$r\frac{d^2R}{dr^2} + (i-1)\frac{dR}{dr} - \kappa\, r\, R^3 = 0\,, \tag{11}$$

while the right hand equation gives:

$$\frac{d^2T}{dt^2} = \frac{\kappa c^2}{T}\,, \tag{12}$$

that is a case of the Emden-Fowler (EF) nonlinear ordinary differential equation with $A=\kappa c^2$ (Sachdev, 1991). Equation (12) has no known explicit solution in the form *T(t)=f(t),* instead it is possible to write:

$$t = \int \left(2A\ln T + C_2\right)^{-1/2} dT + C_1\,. \tag{13}$$

[0086] With $\kappa \neq 0$, Equation (12) can be integrated numerically with the approximate result: $T(t) \approx C_1 + C_2 t^{\alpha}$, with C_1 and C_2 real constants and α slightly larger than 1 ($\alpha \approx 1.03$). The solutions of Equation (11) are oscillating, while the solution of Equation (12) can be continuously increasing with time, which is the feature of our interest for amplification. Materials with nonlinear gravitational properties may enhance the amplification phenomenon.

[0087] With $\kappa \neq 0$, Equation (12) can be integrated numerically with the approximate result:

$$T(t) \approx C_1 + C_2 t^{\alpha}, \qquad (14)$$

with C_1 and C_2 real constants and α slightly larger than 1 ($\alpha \approx 1.03$). The solutions of Equation (11) are oscillating, while the solution of Equation (12) can be continuously increasing with time, which is the feature of our interest for amplification.

[0088] Materials with nonlinear gravitational properties may enhance the amplification phenomenon.

[0089] For calculating the physical dimensions of the amplifier it is necessary to consider the effect of the nonlinearity of GW propagation on gravitons emitted by the converter and propagating in high vacuum. Considering the time component of the refractive index, we can put it in the Schwarzschild's metric obtaining:

$$\frac{1}{C_1 + C_2 t^{\alpha}} = \left(1 - \frac{2Gm}{r}\right), \qquad (15)$$

[0090] In which the gravitational field associated to the wave is represented by *m,* which increases as time passes because the term in parentheses is positive when t=0.

With the proposed externally excited converter the reference peak power of the impulse is comparable to the peak power of the exciting X-ray laser and can be according to present technology about .6 MW at a typical wavelength of 47 nm. This power will be converted to gamma ray GWs by the converter at a typical wavelength of 5.2 pm by the N=8936 frequency relationship.

[0091] In an amplifier length of 1 m, the amplifier contains $1/5.2*10^{-12}$ wavelengths, that is $1.9*10^{11}$ wavelengths. More general, the ratio of the wavelength of the GW by a length of the amplifier in a propagation direction of the GW is usually between 10^6 and 10^{20} wavelengths per meter, preferably between 10^8 and 10^{16} wavelengths per meter. As C_2 of equation 14 represents the non-linearity of propagation in time and assuming an overall contribution of the order of 10^{-7} per wavelength for typical configurations, a high vacuum chamber with a 1 m length gives appreciable amplification of the order of $10^{4,}$ or more general of 10 to 10^8, preferably between 10^3 and 10^8.

[0092] In some embodiments, the amplifier may have a length between 0.5 and 1000 meter.

[0093] Implementations of any of the above described blocks, apparatuses, systems, techniques or methods include, as non limiting examples, implementations as hardware, software, firmware, special purpose circuits or logic, general purpose hardware or controller or other computing devices, or some combination thereof.

[0094] It is to be understood that what is described above is what is presently considered the preferred embodiments of the present invention. However, it should be noted that the description of the preferred embodiments is given by way of example only and that various modifications may be made without departing from the scope of the invention as defined by the appended claims.

[0095] Hereinafter is a short summary of the most important features relevant for generating gravitational waves according to some embodiments of the present invention:

[0096] A Gravitational Wave Generating device comprising:

One or a plurality of dynamical systems, each composed by two or more coupled oscillators with different characteristic frequencies. One of the oscillators is coupled to input energy and another emits Gravitational Waves, and a device for sending electromagnetic waves or excitation energy to the low frequency oscillator.

[0097] A method for avoiding the emission of electromagnetic energy from the high frequency oscillator by using electrically neutral and massive emitting particles, thus promoting the emission of Gravitational waves.

[0098] An amplification stage composed by a volume of ultra high vacuum and/or non-linear material for enhancing and/or guiding non-linear propagation and amplification of gravitational waves.

An optional excitation/triggering stage applicable to some specific oscillators near the chaotic instability region.

An optional particle accelerator to cycle the oscillators through the instability region.

[0099] A device in which the coupled oscillators are subnuclear structures like spheron cores or central fraction of

nuclei for coupling to external electromagnetic excitation and dineutrons for the emission of gravitational waves.
A device in which the oscillators are a stable or metastable nucleus that is made more unstable and/or excited by external energy.
A device in which the external energy is X-ray
A device in which the external energy is gamma-ray
A device in which the external energy is an elementary particle like a neutron, proton, electron, positron, antiproton or any subatomic particle.
A device in which the external energy is a generic electromagnetic wave.
A device in which, after the generation of a pulse of gravitational waves because the emitting nucleus has been hit the external energy, the nucleus returns to its original state.
A device in which, after the generation of a pulse of gravitational waves because the emitting nucleus has been hit the external energy, the nucleus does not return to its original state because of the nuclear reaction.
A device in which the coupled oscillators are molecular structures like atoms and/or crystal lattices for coupling to external electromagnetic excitation and superconducting electron for the emission of gravitational waves.
A device in which the coupled oscillators are molecular structures like atoms and/or crystal lattices for coupling to external thermal and/or acoustical excitation and superconducting electron for the emission of gravitational waves.
A device in which the coupled oscillators are molecular structures like atoms and/or crystal lattices for coupling to external electrical and/or thermal and/or acoustical excitation and crystal/molecular excitations in a frictionless paired state for the emission of gravitational waves.
A device in which a box with ultra high vacuum inside allows the amplification of the gravitational wave.
A device in which a box with non-linear material inside allows the amplification of the gravitational wave.
A device where the chaotic instability region is the Z=115 to Z=116 boundary.
A device where the chaotic instability region is the Z=X to Z=X+1 boundary. Where X is a number between 1 and 200.
A device where the chaotic instability region is the Z=X to Z=X+2 boundary. Where X is a number between 1 and 200.

REFERENCEES

U.S. Patent Documents

[0100]

2719944 Oct., 1955 Brailsford 318/254.
2814769 Nov., 1957 Williams 318/171.
3612630 Oct., 1971 Rosensweig 308/10.
3667019 May., 1972 Elliott et al. 318/254.
3903463 Sep., 1975 Kanamori 318/138.
3959700 May., 1976 Sugiura et al. 318/138.
3970936 Jul., 1976 Arnold 455/899.
4035658 Jul., 1977 Diggs 290/55.
4052134 Oct., 1977 Rumsey 416/119.
4086505 Apr., 1978 McDonald 310/74.
4205268 May., 1980 Eerkens 455/899.
4546264 Oct., 1985 Pinson 290/54.
4968475 Nov., 1990 Drukier et al. 376/153.
5276717 Jan., 1994 Weber 376/156.
5398571 Mar., 1995 Lewis 74/572.
5446018 Aug., 1995 Takahata et al. 310/90.
5514923 May., 1996 Gossler et al. 310/74.
5646728 Jul., 1997 Coutsomitros 356/352.
5721461 Feb., 1998 Taylor 310/268.
5831362 Nov., 1998 Chu et al. 310/90.
5929579 Jul., 1999 Hsu 318/439. 6784591 Aug, 2004, Baker

British Patent Documents

[0101]

1333343 Oct., 1973 GB

Other References:

[0102]

Almeida, J.B., "An Alternative to Minkowski Space-time," (2001), v2, http://arxiv.org/abs/gr-qc/0407022, accessed July 16, 2004.

Almeida, J.B., "An Hypersphere Model of the Universe - The Dismissal of Dark Matter," (2004), v3, http://arxiv.org/abs/Physics/0402075, accessed July 16, 2004.

Baker, R. M. L., Jr., "Gravitational Wave Generator Utilizing Submicroscopic Energizable Elements," *United States Patent* Number 6,784,591, 100 claims, filed December 27, 2000.

Baker, R. M. L. Jr., " Precursor Proof-of-Concept Experiments for Various Categories of High-Frequency Gravitational Wave (HFGW) Generators," in proceedings of Space Technology and Applications International Forum (STAIF-2004) , edited by M. El-Genk, AIP Conference Proceedings 699, Melville, New York, 2004, pp. 1093-1097.

Baker, R.M.L. Jr., "Applications of High-Frequency Gravitational Waves," in proceedings of Space Technological Applications International Forum (STAIF 2005), edited by M.S. El-Genk, AIP Conference Proceedings 746, Melville, New York, 2005 pp. 1311-1314.

Baker, R.M.L., Jr., Davis, E.W. and Woods, R.C., "Gravitational Wave Radiation Pattern at the Focus of a High-Frequency Gravitational Wave Generator and Aerospace Applications," in proceedings of the Space Technological Applications International Forum (STAIF 2005), edited by M. S. El-Genk, AIP Conference Proceedings 746, Melville, New York, 2005, pp 1315-1322.

Binder, B., "Geometric Phase Locked in Fine Structure," in Iterative Coupling and Balancing Currents, e-Book, ISBN: 3-00-010972-2, (2003).

Binder, B., "Magic Angle Precession," in the proceeding of Space Technology and Applications International Forum (STAIF-08), edited by M. El-Genk, AIP Conference Proceedings 969, New York, (2008a), pp. 1103-1110.

Binder, B., "Magic Angle Chaotic Precession," in the proceedings CHAOS2008, Chania, Greece, (2008b).

Brown, J., "Proof of the Manley-Rowe Relations from Quantum Considerations," Electronics Letters, 1, 23-24 (1965)

Fontana, G. and Baker, R. M. L. Jr., "Generation of Gravitational Waves with Nuclear Reactions" in proceedings of Space Technology and Applications International Forum (STAIF-2006) , edited by M. El-Genk, AIP Conference Proceedings 813, Melville, New York, 2006, pp. 1352-1358.

Franklin, J., "Fermion Quarks of Spin 3/2", Phys. Rev. 180, 1583-1587, (1969).

Halpern, L., Laurent, B., "On the Gravitational Radiation of Microscopic Systems," Il Nuovo Cimento, XXXIII, N. 3, 728-751 (1964).

Kamal, S., Parker, B., "Evidence for Dineutrons in Extremely Neutron-Rich Nuclei," Physical Review Letters, 66, 2448-2451 (1991).

Klemperer, W. B. and Baker, R. M. L. Jr., "Satellite Librations," Astronautica Acta, 3, 16-27 (1957). Ishlinskii, A., "Mechanics of Special Gyroscopic Systems", Izd. AN USSR, 1952, Kiev.

Itagaki, N. et al., "Triaxial Deformation in 10Be," Physical Review C, 65, 044302-1 - 044302-6 (2002). Landau, L. D. and Lifshitz, E. M., The Classical Theory of Fields, Fourth Revised English Edition, Pergamon Press, 1975.

Li, F., Baker, R. M. L. and Fang, Z., "Coupling of an Open Cavity to Microwave Beam: A Possible New Scheme of Detecting High-Frequency Gravitational Waves," in the proceedings of Space Technology and Applications International Forum (STAIF-2007), edited by M. El-Genk, AIP Conference Proceedings 880, Melville, New York, 2007, pp. 1139-1147.

Lynch, P., "Resonant Motion of the Three-dimensional Elastic Pendulum," International Journal of Nonlinear Mechanics, 37, 345-367 (2002).

Malykin, G. B., "The relation of Thomas precession to Ishlinskii's theorem as applied to the rotating image of a relativistically moving body", PHYS-USP 42, 5, (1999).

Marsden, J. E., T. Ratiu, Introduction to Mechanics and Symmetry, Texts in Applied Mathematics vol. 17, Springer-Verlag, 1994.

Mishler, L.A. Gyroscopic Device, *United States Patent* 3,726,146, (1973).

Misner, C. W., Thome, K. S. and Wheeler, J. A., Gravitation, Freeman (1973).

Padamsee, H., "Accelerating Applications of RF Superconductivity - Success Stories," IEEE Transactions on Applied Superconductivity, 15, 2432-2439 (2005).

Pauling, L., "The Close-Packed-Spheron Model of Atomic Nuclei and its Relation to the Shell Model," Proc. Natl. Acad. Sci. USA, 54, 989-994 (1965).

Pauling, L., "Binding Energy of Dineutrons N≈90," (1967), http://osufibrary.orst.edu/specialcollections/rnb/32/32-143.html, accessed June 4, 2008.

Pauling, L., "Orbiting Clusters in Atomic Nuclei," Proc. Natl. Acad. Sci. USA, 64, 807-809 (1969). Pauling, L., "Regularities in the Sequences of the Number of Nucleons in the Revolving Cluster for the Ground-state Energy Bands

of the Even-even Nuclei with Neutron Number Equal to or Greater than 126," Proc. Natl. Acad. Sci. USA, 87, 4435-4438 (1990).
Pinto I.M. and. Rotoli G., "Laboratory generation of gravitational waves?," Proceedings of 8th Italian Conference on General Relativity and Gravitational Physics, Cavalese (Trento), August 30 - September 3, World Scientific (1988).
Christodoulou, D., "Nonlinear Nature of Gravitation and Gravitational-Wave Experiment," Physical Review Letters, 67 N. 12, 1486-1489 (1991).
Corkill, R. W. and Stewart J. M., "Numerical Relativity. II. Numerical Methods for the Characteristic Initial Value Problem and the Solution of Vacuum Field Equations for Space-times with Two Killing Vectors," Proc. R. Soc. Lond. A, 386, 373-391 (1983).
Ferrari, V., "Focusing Process in the Collision of Gravitational Plane Waves," Physical Review D, 37, No 10, 3061-3064 (1988).
Fontana, G., "Gravitational Radiation Applied to Space Travel," proceedings of the Gravitational-Wave Conference, edited by P. Murad and R. Baker, The MITRE Corporation, Mclean, Virginia, May 6-9, 2003, Paper HFGW-03-111.
Fontana, G. and Baker, R. M. L. Jr., "The High-Temperature Superconductor (HTSC) Gravitational Laser (GASER)," in the proceedings of the Gravitational-Wave Conference, edited by P. Murad and R. Baker, The MITRE Corporation, Mclean, Virginia, 2003, Paper HFGW-03-107.
Fontana, G., "Design of a Quantum Source of High-Frequency Gravitational Waves (HFGW) and Test Methodology," in the proceedings of Space Technology and Applications International Forum (STAIF-2004), edited by M. S. El-Genk, AIP Conference Proceedings 699, Melville, New York, 2004, pp. 1114-1121.
Fontana, G., "Gravitational Wave Propulsion," in the proceedings of Space Technology and Applications International Forum (STAIF-2005), edited by M. S. El-Genk, AIP Conference Proceedings 746, Melville, New York, 2005a, pp. 1331-1338.
Fontana, G., "The Four Space-times Model of Reality," the in proceedings of Space Technology and Applications International Forum (STAIF-2005), edited by M. S. El-Genk, AIP CP 746, Melville, New York, 2005b, pp. 1403-1410.
Rudenko, V. N., "Optimization of Parameters of a Coupled Generator-Receiver for a Gravitational Hertz Experiment," in Proceedings of the Gravitational-Wave Conference, edited by P. Murad and R. Baker, The MITRE Corporation, McLean, Virginia, 2003, Paper HFGW-03-113.
Szekeres, P., "Colliding Plane Gravitational Waves," J. Math. Phys, 13, No. 3, March, 286-294,(1992). Taylor, J. H., Jr, "Binary pulsar and relativistic gravity," Reviews of Modern Physics, 66, No. 3, July, 711-719 (1994.)
Thorne, K. S., "Gravitational-wave Research: Current Status and Future Prospects," Reviews of Modern Physics, 52, No. 2, Part I, April, 285-297 (1980a).
Thorne, K. S., "Multipole Expansion of Gravitational Radiation," Reviews of Modern Physics, 52, No. 2, Part I, April, 299-339 (1980b).
Thorne, K.S., "Gravitational-Wave Bursts with Memory: the Christodoulou Effect," Physical Review D, 45, January, 520-524 (1992).
Veneziano, G., "Mutual Focusing of Graviton Beams," Modern Physics Letters A, 2, No 11, 899-903 (1987).

Claims

1. Use of a high vacuum chamber as an amplifier of a gravitational wave.

2. The use of claim 1, wherein the pressure in the high vacuum chamber is below 10^{-3} mbar.

3. The use of any of claims 1 to 2, wherein the ratio of a wavelength of the gravitational wave by a length of the amplifier in a propagation direction of the gravitational wave is between 10^8 and 10^{16} wavelengths per meter.

4. The use of any of claims 1 to 3, wherein the amplifier amplifies the gravitational wave by a factor of 10 and 10^8 per meter of the amplifier in a propagation direction of the gravitational wave

5. The use of any of claims 1 to 4, wherein the high vacuum chamber has a length between 0.5 and 1000 meter in a propagation direction of the gravitational wave.

6. The use of any of claims 1 to 5, wherein a non-linear material for the gravitational wave is disposed in the high vacuum chamber.

7. A device for generating a gravitational wave, comprising a generator (10, 18) of an external energy;
a primary vacuum chamber (20) with window (19) transparent for the external energy;

a first vacuum pump (21) connected to the primary vacuum chamber (20) and configured to evacuate the primary vacuum chamber (20);
a converting material (22, 12) for converting the external energy into a gravitational wave, the converting material being disposed inside the primary vacuum chamber (20) such that the external energy illuminates the converting material (22);
a second vacuum chamber (16, 23) used as an amplifier of the gravitational wave according to any of claims 1 to 6; and
a second vacuum pump (24) connected to the second vacuum chamber (16, 23) and configured to evacuate the primary vacuum chamber (16, 23).

8. A device according to claim 7, wherein the converting material is excitable into an excited state by the external energy, wherein the excited state comprises a first oscillator of a first characteristic frequency comprising a mass and emitting the gravitational wave, and a second oscillator coupled with the first oscillator, and wherein a characteristic frequency of the second oscillator is smaller than the first characteristic frequency.

9. The device according to claim 8, wherein the first oscillator is not electrically charged.

10. The device according to any of claims 8 to 9, wherein each of the first and the second oscillator comprises a subnuclear structure.

11. The device according to any of claims 8 to 10, wherein the converting material in a state other than the excited state comprises a stable or metastable nucleus, and the external energy includes at least one of an electromagnetic wave and an elementary particle.

12. The device according to claim 11, wherein the electromagnetic wave comprises at least one of X-ray and gamma ray, and wherein the elementary particle comprises at least one of a neutron, a proton, an electron, a positron, and an antiproton.

13. The device according to claim 11 or 12, wherein a number of protons in the nucleus is between 90 and 125.

14. The device according to any of claims 7 to 12, wherein the converting material comprises Gadolinium.

15. The device according to any of claims 8 to 12, wherein the first oscillator is one of a superconducting electron pair and a crystal or molecular excitation in a frictionless paired state, the second oscillator is one of a molecular structure, an atom, and a crystal lattice, and the external energy is at least one of an electromagnetic excitation, a thermal excitation, and an acoustic excitation.

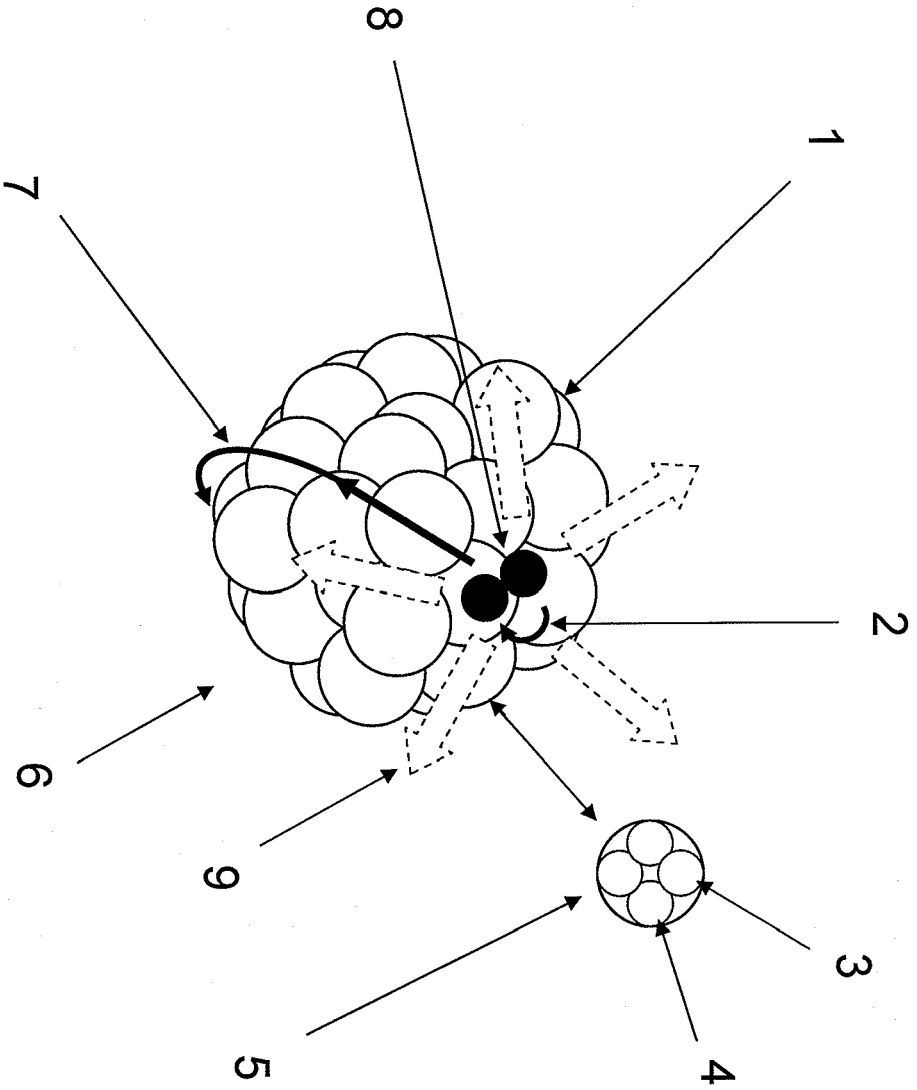


FIG. 1

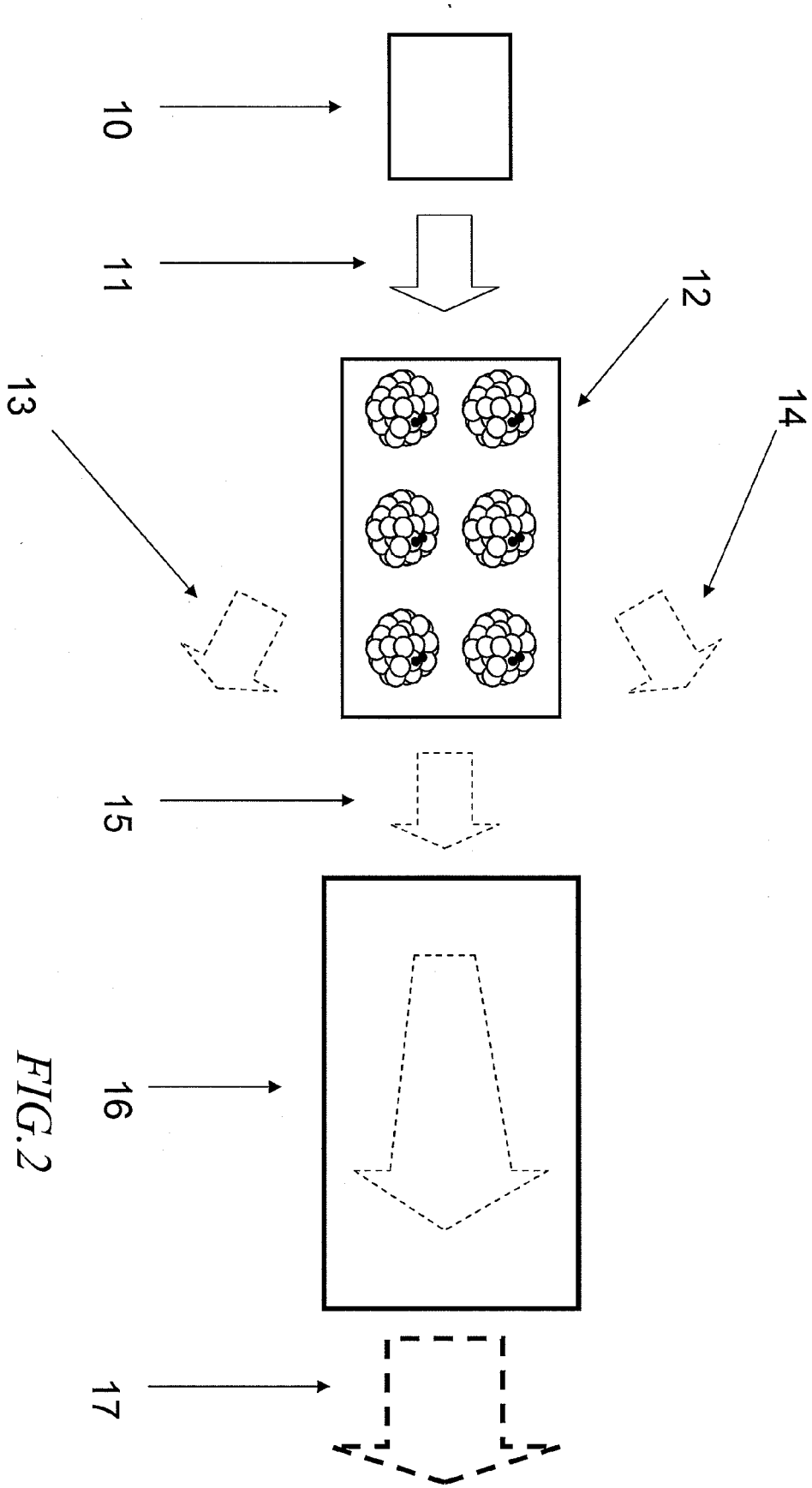


FIG.2

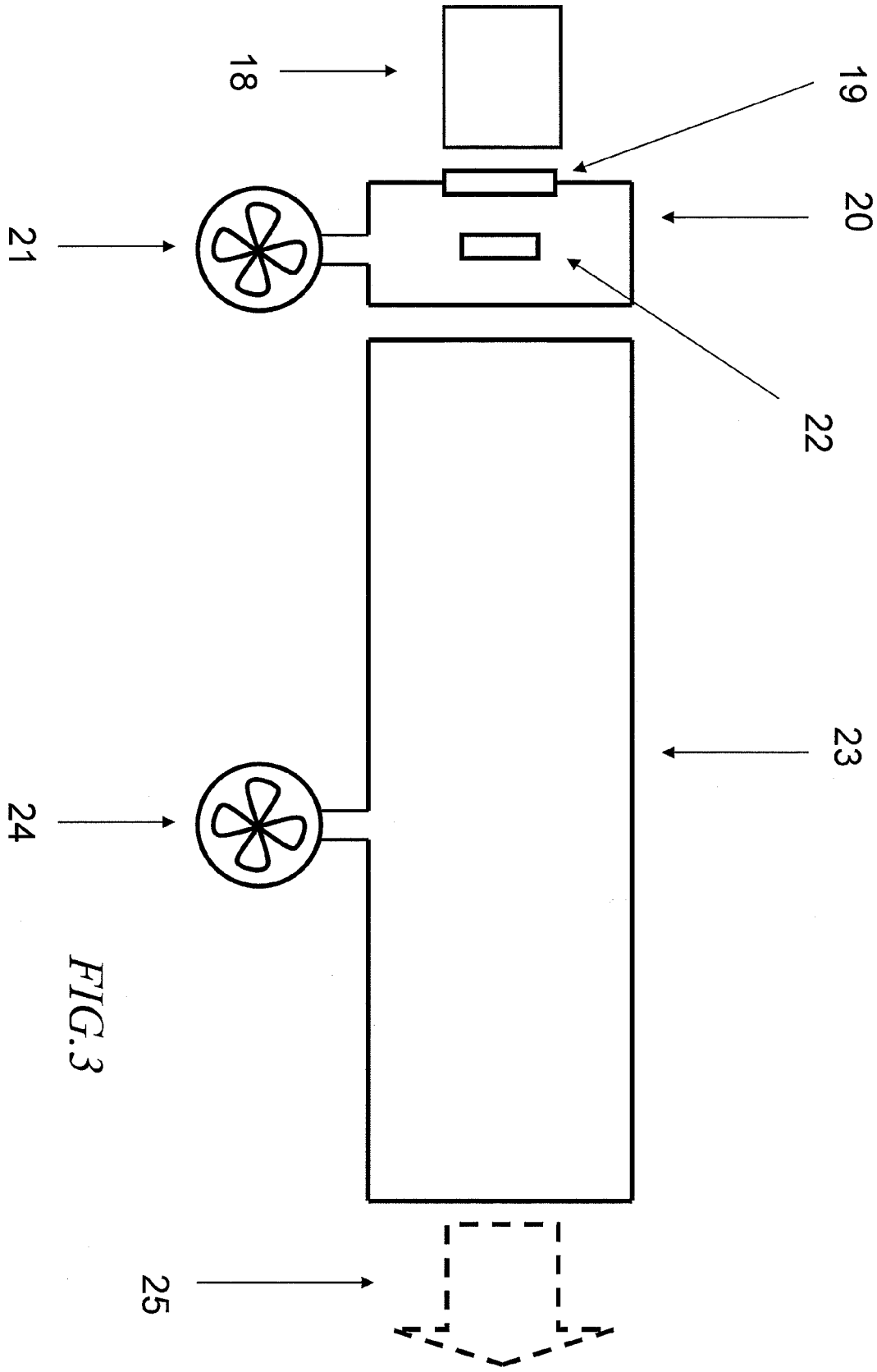


FIG.3

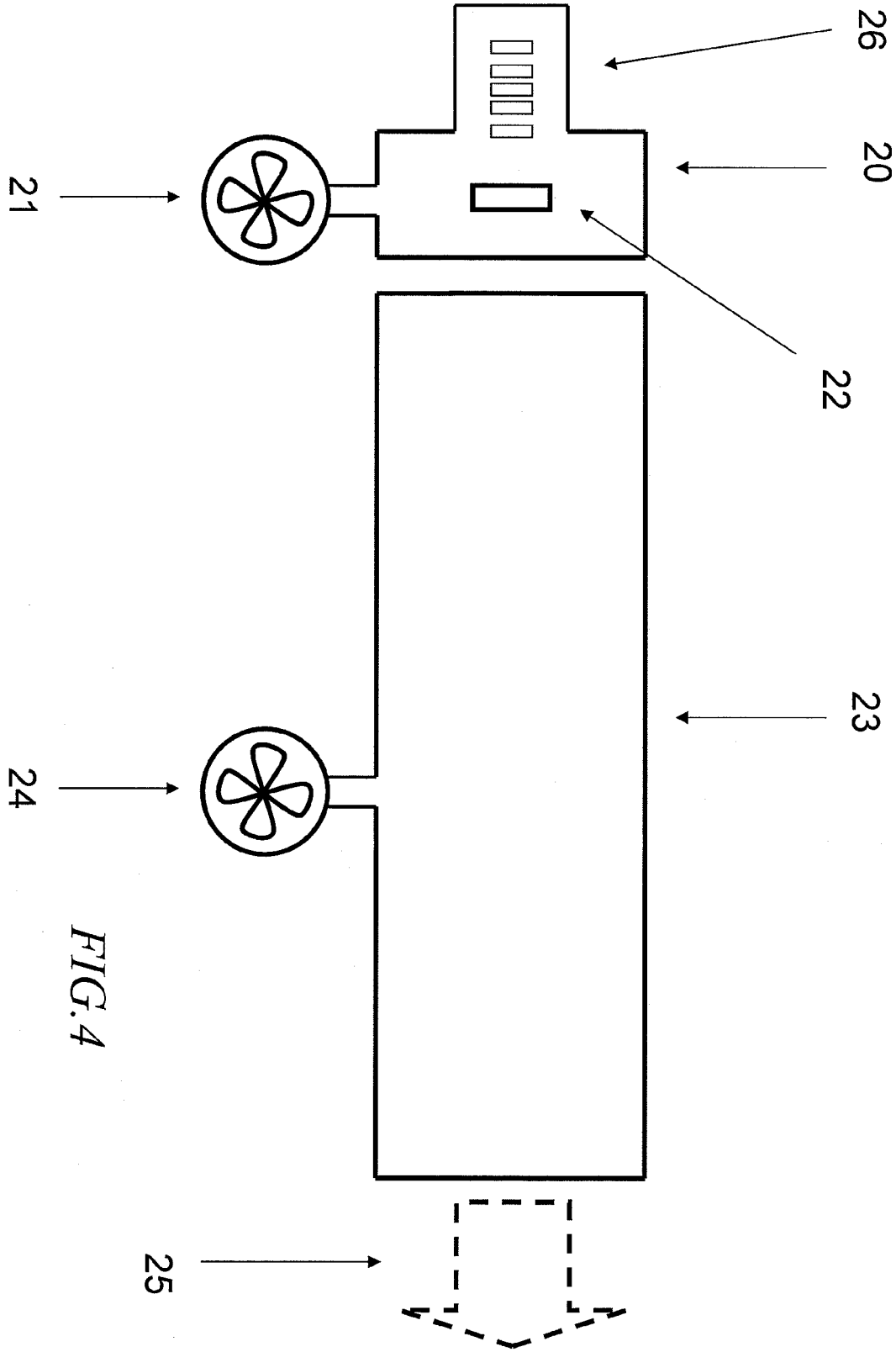


FIG.4

Europäisches Patentamt
European Patent Office
Office européen des brevets

DECLARATION

which under Rule 63 of the European Patent Convention shall be considered, for the purposes of subsequent proceedings, as the European search report

Application Number EP 09 15 6549

The Search Division considers that the present application, does not comply with the provisions of the EPC to such an extent that it is not possible to carry out a meaningful search into the state of the art on the basis of all claims

Reason:

All claims of the application include "amplification of gravitational waves". The possibility of amplification has so far not generally been recognized, prior art discusses amplification in combination with black holes. The application apparently targets amplification of gravitational waves by means of laboratory scale devices.
In the description, no clear description of an embodiment is present, only statements which say that certain materials or configuration may be used. In particular, the amplification is said to be made in a vacuum chamber, possibly by means of a nonlinear material (to gravitational waves). However no example of such a material is given in detail.
Further, reference is made to masses of nucleons, sometimes as nucleus of Gadolinium, but as well as a nucleus with Z=115, and a mass consisting of 10^7 nucleons. NB that the generation of the latter two has not yet been reported, yet confirmed.
Further, in one configuration an exciting coherent X-ray beam is to be used, and the description makes reference to and X-ray laser with energies preferable between 30 and 70 keV. Such an laser has not yet been realised.
In consequence, the application concerns entirely subject matter, which does not comply with the requirements of Art 83 EPC, and therefore a meaningful search is not possible.

The applicant's attention is drawn to the fact that a search may be carried out

-/--

CLASSIFICATION OF THE APPLICATION (IPC)

INV.
G01V7/00

Place of search	Date	Examiner
Munich	29 July 2009	Oestreich, Sebastian

2

EPO FORM 1504 (P04F37)

Europäisches Patentamt
European Patent Office
Office européen des brevets

DECLARATION

which under Rule 63 of the European Patent Convention shall be considered, for the purposes of subsequent proceedings, as the European search report

Application Number
EP 09 15 6549

The Search Division considers that the present application, does not comply with the provisions of the EPC to such an extent that it is not possible to carry out a meaningful search into the state of the art on the basis of all claims

Reason:

during examination following a declaration of no search under Rule 63 EPC, should the problems which led to the declaration being issued be overcome (see EPC Guideline C-VI, 8.2).

L GAL'TSOV D V ET AL: "Amplification of gravitational waves interacting with a binary system"
PHYSICS LETTERS A, NORTH-HOLLAND PUBLISHING CO., AMSTERDAM, NL,
vol. 90, no. 7, 26 July 1982 (1982-07-26),
pages 351-353, XP024623479
ISSN: 0375-9601 [retrieved on 1982-07-26]

CLASSIFICATION OF THE APPLICATION (IPC)

Place of search	Date	Examiner
Munich	29 July 2009	Oestreich, Sebastian

2

EPO FORM 1504 (P04F37)

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- US 6417597 B, Robert M. L. Baker, Jr. **[0005]**
- US 6784591 B **[0006] [0100] [0102]**
- US 2719944 A **[0100]**
- US 318254 A **[0100]**
- US 2814769 A **[0100]**
- US 318171 A **[0100]**
- US 3612630 A **[0100]**
- US 30810 A **[0100]**
- US 3667019 A **[0100]**
- US 3903463 A **[0100]**
- US 318138 A **[0100]**
- US 3959700 A **[0100]**
- US 3970936 A **[0100]**
- US 455899 A **[0100]**
- US 4035658 A **[0100]**
- US 29055 A **[0100]**
- US 4052134 A **[0100]**
- US 416119 A **[0100]**
- US 4086505 A **[0100]**
- US 31074 A **[0100]**
- US 4205268 A **[0100]**
- US 4546264 A **[0100]**
- US 29054 A **[0100]**
- US 4968475 A **[0100]**
- US 376153 A **[0100]**
- US 5276717 A **[0100]**
- US 376156 A **[0100]**
- US 5398571 A **[0100]**
- US 74572 A **[0100]**
- US 5446018 A **[0100]**
- US 5514923 A **[0100]**
- US 5646728 A **[0100]**
- US 356352 A **[0100]**
- US 5721461 A **[0100]**
- US 310268 A **[0100]**
- US 5831362 A **[0100]**
- US 31090 A **[0100]**
- US 5929579 A **[0100]**
- US 318439 A **[0100]**
- GB 1333343 A **[0101]**
- US 3726146 A **[0102]**

Non-patent literature cited in the description

- **Pinto ; Rotoli.** General Relativity and Gravitational Physics. World Scientific, Singapore, 1998 **[0004]**
- *Itagaki,* 2002, 044302-4 **[0045]**
- **Almeida, J.B.** An Alternative to Minkowski Space-time. *v2, http://arxiv.org/abs/gr-qc/0407022,* 16 July 2004 **[0102]**
- **Almeida, J.B.** An Hypersphere Model of the Universe - The Dismissal of Dark Matter. *v3, http://arxiv.org/abs/Physics/0402075,* 16 July 2004 **[0102]**
- Precursor Proof-of-Concept Experiments for Various Categories of High-Frequency Gravitational Wave (HFGW) Generators. **Baker, R. M. L. Jr.** proceedings of Space Technology and Applications International Forum (STAIF-2004). 2004, vol. 699, 1093-1097 **[0102]**
- Applications of High-Frequency Gravitational Waves. **Baker, R.M.L. Jr.** proceedings of Space Technological Applications International Forum (STAIF 2005). 2005, vol. 746, 1311-1314 **[0102]**
- Gravitational Wave Radiation Pattern at the Focus of a High-Frequency Gravitational Wave Generator and Aerospace Applications. **Baker, R.M.L., Jr. ; Davis, E.W. ; Woods, R.C.** proceedings of the Space Technological Applications International Forum (STAIF 2005). 2005, vol. 746, 1315-1322 **[0102]**
- Geometric Phase Locked in Fine Structure. **Binder, B.** Iterative Coupling and Balancing Currents, e-Book. 2003 **[0102]**
- Magic Angle Precession. **Binder, B.** the proceeding of Space Technology and Applications International Forum (STAIF-08). 2008, vol. 969, 1103-1110 **[0102]**
- **Binder, B.** Magic Angle Chaotic Precession. *the proceedings CHAOS2008,* 2008 **[0102]**
- **Brown, J.** Proof of the Manley-Rowe Relations from Quantum Considerations. *Electronics Letters,* 1965, vol. 1, 23-24 **[0102]**
- Generation of Gravitational Waves with Nuclear Reactions. **Fontana, G. ; Baker, R. M. L. Jr.** proceedings of Space Technology and Applications International Forum (STAIF-2006). 2006, vol. 813, 1352-1358 **[0102]**
- **Franklin, J.** Fermion Quarks of Spin 3/2. *Phys. Rev.,* 1969, vol. 180, 1583-1587 **[0102]**

- **Halpern, L. ; Laurent, B.** On the Gravitational Radiation of Microscopic Systems. *Il Nuovo Cimento,* 1964, vol. XXXIII (3), 728-751 **[0102]**
- **Kamal, S. ; Parker, B.** Evidence for Dineutrons in Extremely Neutron-Rich Nuclei. *Physical Review Letters,* 1991, vol. 66, 2448-2451 **[0102]**
- **Klemperer, W. B. ; Baker, R. M. L. Jr.** Satellite Librations. *Astronautica Acta,* 1957, vol. 3, 16-27 **[0102]**
- **Ishlinskii, A.** Mechanics of Special Gyroscopic Systems. *Izd. AN USSR,* 1952 **[0102]**
- **Itagaki, N. et al.** Triaxial Deformation in Be. *Physical Review C,* 2002, vol. 65, 044302-1044302-6 **[0102]**
- **Landau, L. D. ; Lifshitz, E. M.** The Classical Theory of Fields. Pergamon Press, 1975 **[0102]**
- Coupling of an Open Cavity to Microwave Beam: A Possible New Scheme of Detecting High-Frequency Gravitational Waves. **Li, F. ; Baker, R. M. L. ; Fang, Z.** the proceedings of Space Technology and Applications International Forum (STAIF-2007). 2007, vol. 880, 1139-1147 **[0102]**
- **Lynch, P.** Resonant Motion of the Three-dimensional Elastic Pendulum. *International Journal of Nonlinear Mechanics,* 2002, vol. 37, 345-367 **[0102]**
- **Malykin, G. B.** The relation of Thomas precession to Ishlinskii's theorem as applied to the rotating image of a relativistically moving body. *PHYS-USP,* 1999, vol. 42, 5 **[0102]**
- Introduction to Mechanics and Symmetry. **Marsden, J. E. ; T. Ratiu.** Texts in Applied Mathematics. Springer-Verlag, 1994, vol. 17 **[0102]**
- **Misner, C. W. ; Thome, K. S. ; Wheeler, J. A.** Gravitation. Freeman, 1973 **[0102]**
- **Padamsee, H.** Accelerating Applications of RF Superconductivity - Success Stories. *IEEE Transactions on Applied Superconductivity,* 2005, vol. 15, 2432-2439 **[0102]**
- **Pauling, L.** The Close-Packed-Spheron Model of Atomic Nuclei and its Relation to the Shell Model. *Proc. Natl. Acad. Sci. USA,* 1965, vol. 54, 989-994 **[0102]**
- **Pauling, L.** *Binding Energy of Dineutrons N≈90,* 1967, http://osufibrary.orst.edu/specialcollections/rnb/32/32-143.html **[0102]**
- **Pauling, L.** Orbiting Clusters in Atomic Nuclei. *Proc. Natl. Acad. Sci. USA,* 1969, vol. 64, 807-809 **[0102]**
- **Pauling, L.** Regularities in the Sequences of the Number of Nucleons in the Revolving Cluster for the Ground-state Energy Bands of the Even-even Nuclei with Neutron Number Equal to or Greater than 126. *Proc. Natl. Acad. Sci. USA,* 1990, vol. 87, 4435-4438 **[0102]**
- **Pinto I.M. ; Rotoli G.** Laboratory generation of gravitational waves?,'' Proceedings of 8th Italian Conference on General Relativity and Gravitational Physics. World Scientific, 30 August 1988 **[0102]**
- **Christodoulou, D.** Nonlinear Nature of Gravitation and Gravitational-Wave Experiment. *Physical Review Letters,* 1991, vol. 67 (12), 1486-1489 **[0102]**
- **Corkill, R. W. ; Stewart J. M.** Numerical Relativity. II. Numerical Methods for the Characteristic Initial Value Problem and the Solution of Vacuum Field Equations for Space-times with Two Killing Vectors. *Proc. R. Soc. Lond. A,* 1983, vol. 386, 373-391 **[0102]**
- **Ferrari, V.** Focusing Process in the Collision of Gravitational Plane Waves. *Physical Review D,* 1988, vol. 37 (10), 3061-3064 **[0102]**
- Gravitational Radiation Applied to Space Travel. **Fontana, G.** proceedings of the Gravitational-Wave Conference. The MITRE Corporation, 06 May 2003 **[0102]**
- The High-Temperature Superconductor (HTSC) Gravitational Laser (GASER). **Fontana, G. ; Baker, R. M. L. Jr.** the proceedings of the Gravitational-Wave Conference. The MITRE Corporation, 2003 **[0102]**
- Design of a Quantum Source of High-Frequency Gravitational Waves (HFGW) and Test Methodology. **Fontana, G.** the proceedings of Space Technology and Applications International Forum (STAIF-2004). 2004, vol. 699, 1114-1121 **[0102]**
- Gravitational Wave Propulsion. **Fontana, G.** the proceedings of Space Technology and Applications International Forum (STAIF-2005). 2005, vol. 746, 1331-1338 **[0102]**
- The Four Space-times Model of Reality. **Fontana, G.** proceedings of Space Technology and Applications International Forum (STAIF-2005). 2005, vol. 746, 1403-1410 **[0102]**
- Optimization of Parameters of a Coupled Generator-Receiver for a Gravitational Hertz Experiment. **Rudenko, V. N.** Proceedings of the Gravitational-Wave Conference. The MITRE Corporation, 2003 **[0102]**
- **Szekeres, P.** Colliding Plane Gravitational Waves. *J. Math. Phys,* March 1992, vol. 13 (3), 286-294 **[0102]**
- **Taylor, J. H., Jr.** Binary pulsar and relativistic gravity. *Reviews of Modern Physics,* July 1994, vol. 66 (3), 711-719 **[0102]**
- **Thorne, K. S.** Gravitational-wave Research: Current Status and Future Prospects. *Reviews of Modern Physics,* April 1980, vol. 52 (2), 285-297 **[0102]**
- **Thorne, K. S.** Multipole Expansion of Gravitational Radiation. *Reviews of Modern Physics,* April 1980, vol. 52 (2), 299-339 **[0102]**
- **Thorne, K.S.** Gravitational-Wave Bursts with Memory: the Christodoulou Effect. *Physical Review D,* January 1992, vol. 45, 520-524 **[0102]**
- **Veneziano, G.** Mutual Focusing of Graviton Beams. *Modern Physics Letters A,* 1987, vol. 2 (11), 899-903 **[0102]**

Exp Fluids (2007) 42:123–130
DOI 10.1007/s00348-006-0225-7

RESEARCH ARTICLE

A novel internal waves generator

L. Gostiaux · H. Didelle · S. Mercier · T. Dauxois

Received: 9 June 2006 / Revised: 20 September 2006 / Accepted: 28 September 2006 / Published online: 27 October 2006
© Springer-Verlag 2006

Abstract We present a new kind of generator of internal waves which has been designed for three purposes. First, the oscillating boundary conditions force the fluid particles to travel in the preferred direction of the wave ray, hence reducing the mixing due to forcing. Second, only one ray tube is produced so that all of the energy is in the beam of interest. Third, temporal and spatial frequency studies emphasize the high quality for temporal and spatial monochromaticity of the emitted beam. The greatest strength of this technique is therefore the ability to produce a large monochromatic and unidirectional beam.

1 Introduction

Stratified fluids are nowadays a field of paramount importance in fluid mechanics and are carefully studied theoretically, experimentally and numerically by a large and always increasing scientific community. The underlying reason is, of course, the identification of many crucial phenomena appearing in the oceanic and atmospheric environments, where the stratification cannot be overlooked. The possibility to reproduce in a laboratory experiment conditions close to theoretical hypotheses is, of course, a very important way to validate theoretical ideas developed in the last decades. Indeed, if one will always rely on in situ observations for the final decisive tests, new fundamental ideas are usually derived within an oversimplified framework that cannot be reproduced in a real environment. Among many difficulties, let us mention just a few. The linear stratification hypothesis is almost always used in theoretical works despite the complicated oceanic or atmospheric stratification (Thorpe 2005). Internal waves are also often considered as monochromatic plane waves, while oceanic measurements usually report a beam-like structure, with a width smaller than a wavelength (Lam et al. 2004; Gerkema et al. 2004). The paradox of internal waves impinging onto topography (Dauxois and Young 1999) was solved for a linear oblique slope which can only be a poor description of a real seamount (Eriksen 1998). It is not satisfactory to test this last theoretical example by considering a *beam* impinging onto a *curved* topography in a fluid with a *real and complicated* stratification (exponential, thermocline, mixed layer,...). Indeed, possible disagreements between theoretical and experimental results might always be attributed to one of these three hypothesis which are not fulfilled in the experimental test. Controlled experiments in a simplified framework are therefore of paramount importance.

In a laboratory experiment, one can identify three main difficulties: the preparation of a stratified fluid with a controlled *stratification*, the *generation* and the

L. Gostiaux · T. Dauxois (✉)
Laboratoire de Physique, UMR-CNRS 5672, ENS Lyon,
46 Allée d'Italie, 69364 Lyon Cédex 07, France
e-mail: Thierry.Dauxois@ens-lyon.fr

L. Gostiaux
e-mail: lgostiau@ens-lyon.fr

H. Didelle · S. Mercier
Laboratoire des Écoulements Géophysiques et Industriels
(LEGI), UMR 5519 CNRS-UJF-INPG,
21 rue des Martyrs, 38000 Grenoble, France

visualization of internal waves propagating within this fluid. The first difficulty was rapidly solved thanks to the well-known double bucket method and its improved versions (Oster 1965; Hill 2002; Benielli and Sommeria 1998). The visualization difficulty has been nicely solved in the last decade by two different techniques: the Schlieren method, on the one hand, later improved into the modern quantitative synthetic Schlieren method (Dalziel et al. 2000), gives time resolved density fields; on the other hand, the particle image velocimetry (PIV) technique provides time resolved velocity fields (Fincham and Delerce 2000). However, the generation of monochromatic plane waves remains a crucial problem from the experimental point of view. We propose in this article a new experimental set-up that is able to provide a single monochromatic and unidirectional large beam.

The paper is organized as follows. In Sect. 2, we briefly review existing generation methods, with a special emphasis on their limitations. In Sect. 3, we present the new internal waves generator. In Sect. 4, we analyze the internal waves which are generated with a particular emphasis on its temporal, and spatial properties. Section 5 concludes the paper.

2 Review of existing generation methods

2.1 Important properties of internal gravity waves

Internal waves in a continuously stratified fluid are shear waves that propagate with a phase velocity perpendicular to their group velocity. The density stratification allows the propagation of such waves by generating a restoring force. This latter is characterized by the Brunt–Väisälä frequency N, which corresponds to the vertical oscillation frequency of water masses in a local density gradient. It is defined as $N^2 = -(g/\rho)(\mathrm{d}\rho/\mathrm{d}z)$ where g is the gravity, ρ the density and z the vertical coordinate. The frequency N will be kept constant in the remainder of the paper.

In such a stratified fluid, it is well-known that the dispersion relation for linear internal gravity waves is

$$\omega = N \sin\theta, \tag{1}$$

where ω is the temporal frequency and θ the angle of the wave vector with the vertical direction. This dispersion relation emphasizes that for a fixed frequency, the direction in which the shear propagates is fixed and is denoted σ. This direction is also found to be orthogonal to the direction of propagation of the energy, denoted s.

The study of internal waves generation must therefore take into account these very unusual and restrictive geometrical conditions. Moreover, it is very important to notice that the spatial wavelength of the wave does not appear in the dispersion relation (1). There is no mechanism of wavelength selection other than the physical boundary conditions imposed by the forcing technique. This last point will be of first importance in the following.

2.2 Oscillating bodies

The simplest experiments on internal waves emission use oscillating bodies as radiation sources and, more precisely, in a two-dimensional set-up, a vertically oscillating cylinder (see the sketch in Fig. 1a). The first experiment was performed by Görtler (1943) in a two-dimensional setup. Nonetheless, as this experiment is often forgotten, one usually refers to the seminal experiment by Mowbray and Rarity (1967). This is nowadays a simple laboratory set-up, often used in laboratory tutorials at the undergraduate level. In the simplified two-dimensional set-up, these generators emit four beams of internal waves, making a constant angle with the horizontal. A typical resulting pattern is presented in Fig. 1b. Recent three-dimensional experiments with an oscillating sphere have reported striking double cones (Peacock and Weidman 2005).

As the wavelength does not enter the dispersion relation (1) of internal waves, it is absolutely non trivial, a priori, to predict the spatial structure of the beam. This question has been at the origin of several theoretical studies (Thomas and Stevenson 1972; Hurley 1997; Hurley and Keady 1997; Hurley and Hood 2001; Voisin 2003) which have proposed a non-trivial scaling of the beam spatial structure by the size of the object, the viscosity of the fluid and the emission frequency. In all cases, the beam envelope does not contain a full wavelength. Therefore, it is difficult to use these sources to study spatial properties of internal waves.

2.3 Paddle-like generators

It has also been proposed to generate directly a shear motion in the stratified fluid by using a multi-bladed folding paddle (McEwan 1973; Cacchione and Wunsch 1973; Teoh et al. 1997; Silva et al. 1997; Ivey et al. 2000; Gostiaux et al. 2006): see the sketch in Fig. 1c. The flow generated by such a device corresponds to the superposition of two shear waves propagating in opposite directions along the zig-zag paddle, each one being independently solution of the internal waves'

equations. Such a set-up thus generates two beams propagating symmetrically leftwards and rightwards (resp. upwards and downwards) when the paddle is set horizontally (resp. vertically) as in Teoh et al. (1997). Locating a vertical wall close to the paddle (see vertical black thick line in Fig. 1d) allows to avoid a second beam: moreover, because of the reflection of the leftward beam on this wall, one may increase artificially the width of the rightward beam.

The energy transmission is proportional to the sine of the angle between the paddle and the along-direction of the beam. To gain in amplitude, one might therefore propose to tilt the paddle orthogonally to the desired beam. One realizes immediately, however, that the generated shear is no more the superposition of two internal wave beams. Consequently it will excite strong nonlinearities. Intermediate position between the horizontal and this inclination can be proposed, this method, however, still generates higher harmonics. As shown in Fig. 1d (see Gostiaux et al. 2006 for additional details), the main part of the energy is transmitted to a beam of internal wave of frequency ω, while higher harmonics $n\omega$ are also excited by the oscillating paddle. Their frequencies being higher, their propagation angles are also larger (see Fig. 1d) and a screen can therefore be located appropriately to remove most harmonics (Gostiaux et al. 2006). Fortunately, in Ivey et al. (2000), this effect was avoided since the excitation frequency ω was larger than half of the Brunt–Väisälä. In conclusion, as avoiding harmonics generation is of primary importance for energy fluxes measurements and comparisons with theoretical predictions, especially for low frequency ratios ω/N, such a method is not ideal.

2.4 Tidal-like excitation

Finally, a third way of generating internal waves has been proposed and realized in a form of a global excitation of internal waves. A vertical periodic motion of the fluid container (see the sketch in Fig. 1e) acts as a perturbation of the gravity force on the whole fluid. Tidal-like excitations by barotropic forcing (Ivey and Nokes 1989; Gostiaux and Dauxois 2006) also excite globally the fluid. Resonant modes are selected and can be focused on internal waves attractors. This excitation

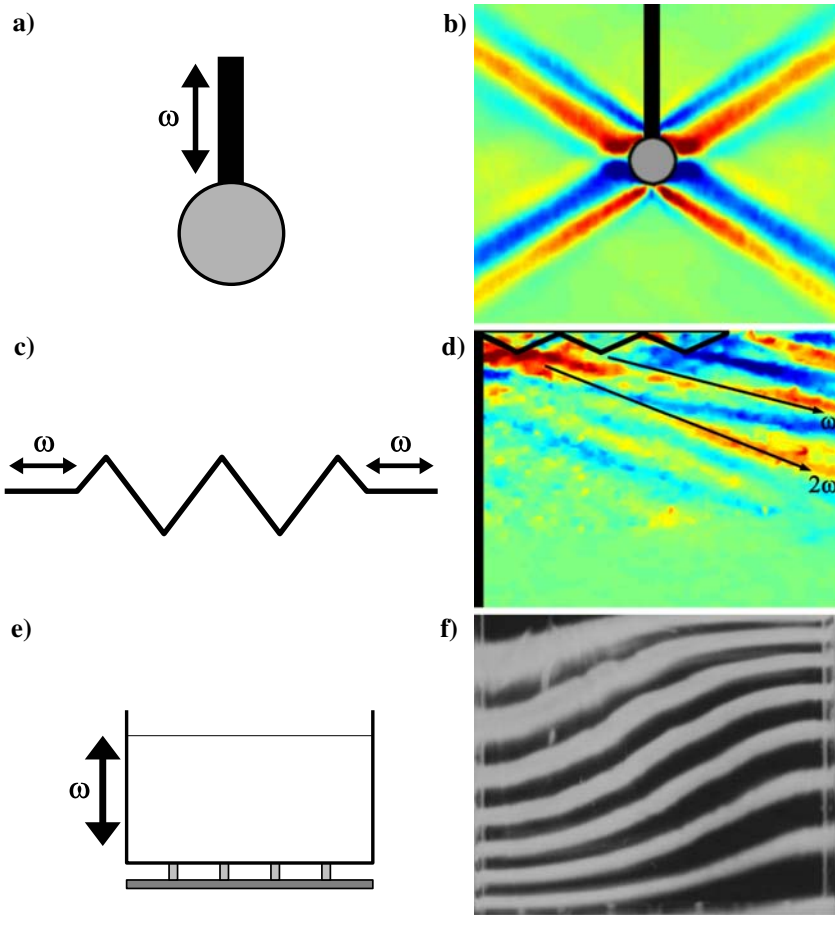


Fig. 1 Principles and internal waves patterns for the three previous internal gravity wave generators. **a** Presents the sketch of a vertically oscillating cylinder. **b** Shows the resulting four rays as captured with a synthetic Schlieren technique. **c** Presents the paddle-like generator, while **d** shows a typical instantaneous PIV measurement. Note harmonic waves propagating in different directions. **e** Presents the parametric excitation principle while **f** shows the resulting internal wave pattern (Benielli 1995)

is highly dependent of the size and shape of the fluid domain, and relies on resonances of the enclosed geometry itself (Maas et al. 1997) as shown in Fig. 1f.

The first very limitation is that oscillating vertically a large tank is, of course, very difficult and cannot be easily performed. In addition, the whole stratified fluid is thus excited so that it is not possible to discern how internal waves, generated in one part of the tank, might encounter independently a topography, or another internal wave, in another part of the tank.

3 The new solution for internal waves generation

As briefly motivated in the previous section, a new principle for internal wave generation has to be proposed, if a *monochromatic plane wave* is necessary for laboratory experiments. The key idea is to create physical boundary conditions that effectively propagate within the fluid interior, instead of a stationary forcing.

3.1 Mechanism of the generator

The generator consists of a pile of 24 expanded PVC sheets ($2 \times 36 \times 150$ cm) enclosed in a parallelepiped half-opened box and free to slip one over the other. The plates are weighted with thin lead sheets so that they are buoyantly neutral in water; it minimizes friction forces between plates. They are thus separated by 2.2 mm one from the other. Two rectangular holes in each plates allow two identical camshafts (see Fig. 2a)

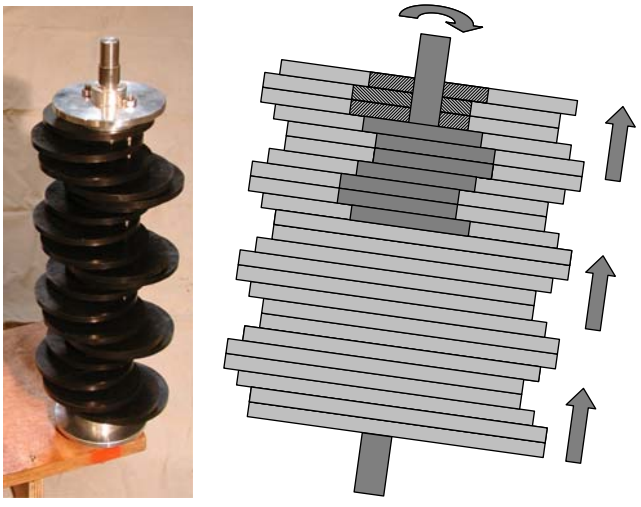


Fig. 2 Internal wave generator mechanism. The *left panel* presents one of the two identical camshafts, while the *right* one shows the cross section of the generator. One clearly identify the plates (*light grey*) and one of the camshafts (*dark grey*)

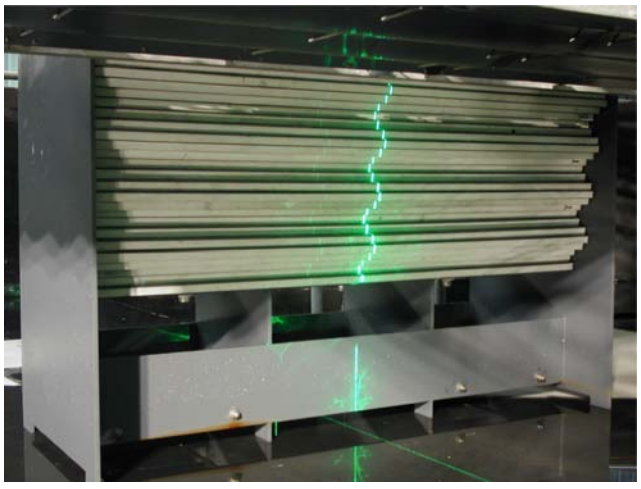


Fig. 3 Perspective view of the wave generator before the filling of the tank. The *thin vertical line* in the *middle* of the plates is the impact of the laser sheet, used for the PIV visualization. We can see the two *vertical* sidewalls of the generator that hinder to view the field very close to the generator

to go through the pile, imposing the relative position of the plates. At rest, the plates are sinusoidally shifted, due to the helicoidal repartition of the cams (see Fig. 2b). The rotation of the camshafts applies a periodic motion to the plates which propagates upward (resp. downward) for a clockwise (resp. anti-clockwise) rotation. The eccentricity of the camshafts defines the amplitude of oscillation of the plates, namely 6 cm peak to peak. A perspective view of the generator can be seen in Fig. 3.

To optimize the emission process, the generator is slightly tilted in direction of the emitted beam, i.e. the direction of shear propagation. In the following, we generate a downward propagating beam (defined by the direction of the group velocity) corresponding to an upward propagating shear (direction of the phase velocity). As anticipated, the device can easily generate upward propagating beams by inverting the sign of rotation of both camshaft.

3.2 Flow visualization

We performed experiments in the 13 m diameter circular tank of the Coriolis Plateform in Grenoble, filled with 90 cm of linearly stratified fluid at 3% corresponding to a Brunt–Väisälä period of $T_{\mathrm{BV}} = 11.7 \pm 0.2$ s. A vertical laser sheet illuminates the wavemaker (visible in Fig. 3) and the 400 μ diameter particles polystyrene beads seeded in the stratified fluid. PIV measurements of the velocity field were performed in order to characterize the waves generated by the new device. All experiments discussed in

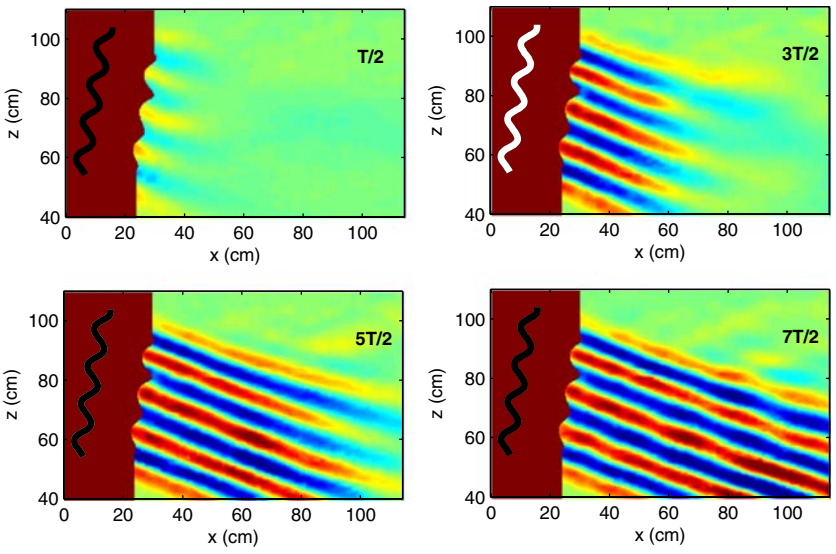


Fig. 4 Horizontal velocity pattern in the vertical plane after 1, 2, 3 and 4 periods T of excitation. The position of the wavemaker is indicated by the oscillating and *tilted white line*. On the *left* of each panel, the *shadowed region* corresponds to the invisible region hidden by the sidewalls of the generator

this paper were carried out with a period of excitation $T = 36$ s. It corresponds to an angle of emission $\theta = 19°$. Snapshots of the flow every 600 ms were recorded using a 12 bits $1{,}024 \times 1{,}024$ pixels camera. Subsequently, we compiled the CIV correlation algorithms (Fincham and Delerce 2000) between successive frames to get the velocity field induced by the wavemaker. Results are reported and analyzed in the next section.

4 Analysis of the emanating internal waves

4.1 Qualitative analysis

Figure 4 present four successive images of the emitted wave. Colors show the horizontal velocity which is the strongest component in this configuration since the angle of propagation is rather small. As can be seen, only one downward propagating beam is emitted: this is the first important improvement of this new generator. These four pictures clearly emphasize the generation of the internal plane waves propagating towards the fluid interior.

The transversal wave structure perfectly fits also the profile of the wavemaker, with four wavelengths emitted. It is the largest beam ever generated in a laboratory experiment. It is however important to stress that there is no theoretical limitation to have a larger beam. This is of crucial importance for experiments designed to test plane waves properties.

Figure 5 shows the velocity field. As expected, it is parallel to the direction of emission, emphasizing nicely the pure shear structure. Moreover, it clearly attests that iso-phase lines are propagating perpendicularly to the energy propagation. The maximal velocity value measured in that case is 1.8 mm/s, while amplitude displacements are of the order of 1 cm.

4.2 Temporal chromaticity

As stated in Sect. 1, *monochromatic* internal gravity waves are highly desirable for studying, for example, the role of nonlinearities involved in collisions with topography (Ivey and Nokes 1989; Dauxois and Young 1999; Gostiaux and Dauxois 2006) or with other internal waves (Teoh et al. 1997): the chromaticity quality of the incident beam is therefore of primary importance. With the present device, as the forcing is itself solution of the wave equation, no harmonics are measurable in the flow. To attest this result, Fig. 6a shows a temporal evolution of the horizontal velocity at a point situated in the centerline of the beam, 29 cm away from the generator.

First, note that the signal is vanishingly small before starting to increase at $t = 30$ s. This remark allows to compute the group velocity of the beam. As the generation and measurement points are distant by 29 cm, one gets a group velocity $c_g \simeq 0.97$ cm/s. This value has to be compared with the theoretical one, $c_g = \lambda/(T \tan\theta)$, where T is the excitation period, λ the wavelength of the internal beam and θ the angle of propagation. With $\lambda = 13.2$ cm, $T = 36$ s and $\theta = 19°$, the expected value $c_g = 1.06$ cm/s compares well with the measurement.

Moreover, one can notice that the transient regime is almost invisible as the wave reaches the measurement area. This allows to consider the forcing of such a wave as a suddenly switched on sinusoid (sinus function convoluted with a Heaviside distribution), which

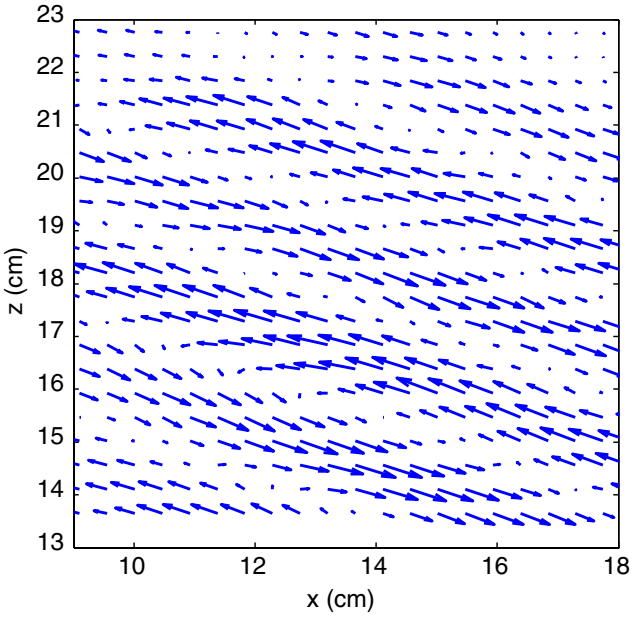


Fig. 5 Zoom of the velocity field generated by the generator

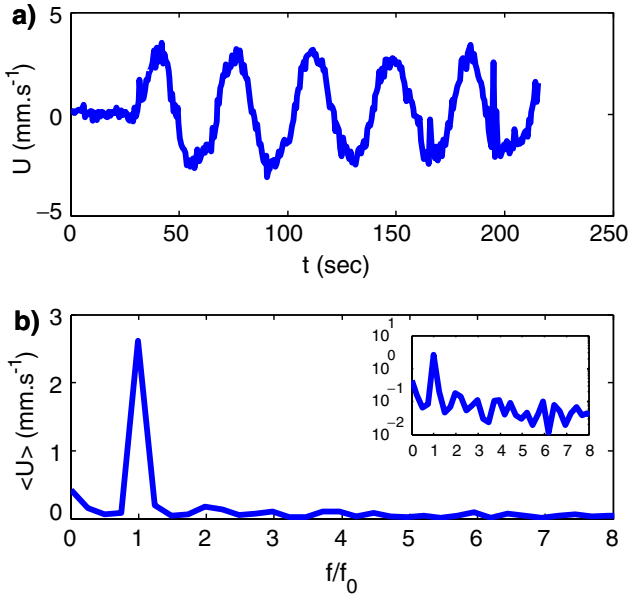


Fig. 6 Temporal analysis of the internal waves beam in a point located at 29 cm from the wavemaker and in the middle of the beam. **a** Presents the time evolution of the horizontal velocity field, while **b** shows the Fourier transform of this signal. Frequencies are renormalized with the excitation one, $f_0 = 1/T$. The *inset* in **b** presents the spectrum in logarithmic scale

allows fine comparisons with analytical results using such a forcing (Dauxois and Young 1999).

Once the beam reaches the measurement point, the velocity oscillates at a well defined frequency. The temporal spectrum of the wave, plotted in Fig. 6b, confirms that the frequency signal is, as expected, imposed by the generator. Moreover, it shows that amplitudes of higher harmonics (see the inset in logarithmic scale) are at least one order of magnitude smaller than the fundamental one. The emitted beam is thus highly temporally monochromatic. The latter characteristic allows to filter the velocity field at the excitation frequency in order to reduce measurement noise. The filtering technique is used in Figs. 7 and 8 which present spatial properties of the beam.

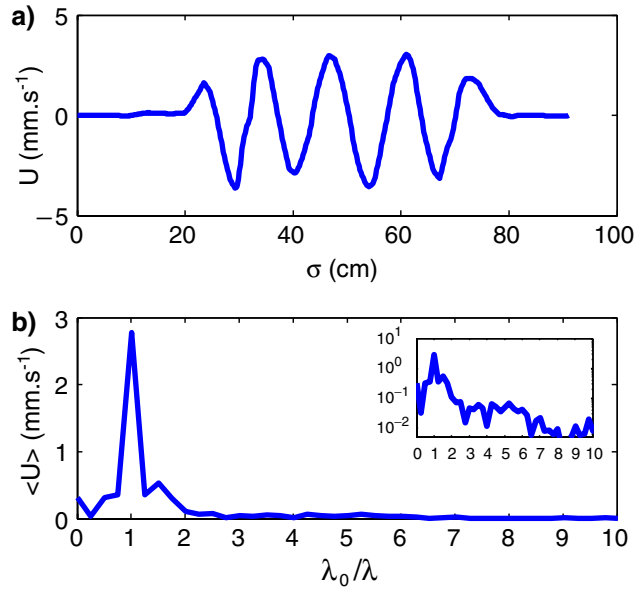


Fig. 7 Transversal structure of the internal waves beam, 29 cm from the wavemaker. **a** Presents the spatial evolution of the horizontal velocity field, while **b** shows the Fourier transform of this signal. Wavelengths are renormalized with the excitation one λ_0. The *inset* in **b** presents the spectrum in logarithmic scale

4.3 Spatial structure of the beam

Contrary to optical or acoustic waves, and owing to the unusual dispersion relation (1) of internal waves, temporal monochromatism does not imply any spatial monochromatism: it is necessary to study them separately. Figure 7a presents a cross section of the horizontal velocity field at 29 cm from the wavemaker. One may easily distinguish slightly more than four wavelengths. Its corresponding wavelength spectrum is shown in Fig. 7b. The spatial monochromatism of the produced beam is attested by the very well defined peak of the spectrum. Note that a careful study shows that end effects in the horizontal y-direction have very little, if no, consequences: one can only detect a very small transverse component of the velocity field.

Finally, Fig. 8 emphasizes that the viscous damping of the beam along its propagation is extremely weak.

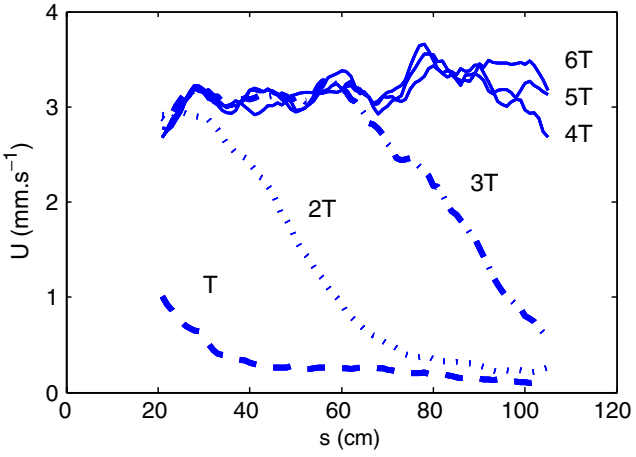


Fig. 8 Superposition of successive longitudinal profiles of the horizontal velocity after 1, 2, 3, 4, 5 and 6 excitation periods

One can notice the non-dispersive profile of the wave front: the beam does not disperse since it is spatially monochromatic. From this picture, it is possible to get another reliable estimate of the group velocity. Indeed, by measuring the location difference of the appearance of the velocity $U = 1$ mm/s or by trying to superpose the first three profiles, one gets $c_g \simeq 38$ cm per period i.e. 1.06 cm/s: the precise theoretical value derived previously.

This generator was used to excite waves up to 1.5 m away from the region of interest, no appreciable damping was measured at this point. The absence of nonlinear losses may explain this remarkable constancy of the wave amplitude.

5 Conclusion

We have reported the design of a new kind of internal waves generator that imposes a rigid boundary condition compatible with the equations governing the propagation of plane internal waves. By generating a propagating shear, we select a single direction of emission of the waves. The waves emitted are purely monochromatic, both spatially and temporally, which is of first importance for the study of nonlinear processes involved in internal waves dynamics.

Finally, it is important to emphasize that this experimental setup allows to generate any kind of wave form. Not only sinusoidal, but also gaussian beams or more complicated shapes can be easily generated, by modifying the camshafts eccentricity appropriately. This possibility opens new perspective for future studies of the dynamics of internal waves in stratified fluids.

Acknowledgments We warmly thank J. Sommeria for helpful discussions and S. Viboud for help during the experiments. Figure 1f was obtained by D. Benielli and J. Sommeria. Comments to the manuscript by Denis Martinand are deeply appreciated. This work has been partially supported by 2005-ANR Project TOPOGI-3D and by the 2006-PATOM CNRS program.

References

Benielli D (1995) Excitation paramétrique et déferlement d'ondes internes en fluide stratifié, PhD Thesis, ENS Lyon

Benielli D, Sommeria J (1998) Excitation and breaking of internal gravity waves by parametric instability. J Fluid Mech 374:117

Cacchione D, Wunsch C (1974) Experimental study of internal waves over a slope. J Fluid Mech 66:223

Dalziel SB, Hughes GO, Sutherland BR (2000) Whole field density measurements by synthetic schlieren. Exp Fluids 28:322

Dauxois T, Young WR (1999) Near-critical reflection of internal waves. J Fluid Mech 390:271

Eriksen CC (1998) Internal wave reflection and mixing at fieberling guyot. J Geophys Res 103:2977

Fincham A, Delerce G (2000) Advanced optimization of correlation imaging velocimetry algorithms. Exp Fluids 29:13

Gerkema T, Lam F-PA, Maas LRM (2004) Internal tides in the Bay of Biscay: conversion rates and seasonal effects. Deep Sea Res II 51:2995–3008

Görtler H (1943) Uber eine schwingungserscheinung in flussigkeiten mit stabiler dichteschichtung. Zeitschrift angewandte Mathematik Mechanik 23:165

Gostiaux L, Dauxois T (2006) Internal tides in laboratory experiments. Phys Fluids (Submitted)

Gostiaux L, Dauxois T, Didelle H, Sommeria J, Viboud S (2006) Quantitative laboratory observation of internal wave reflection on ascending slopes. Phys Fluids 18:056602

Hill DF (2002) General density gradients in general domains: the ''two-tank'' method revisited. Exp Fluids 32:434

Hurley DG (1997) The generation of internal waves by vibrating elliptic cylinders. Part 1. Inviscid solution. J Fluid Mech 351:105

Hurley DG, Hood MJ (2001) The generation of internal waves by vibrating elliptic cylinders. Part 3. Angular oscillations and comparison of theory with recent experimental observations. J Fluid Mech 433:61

Hurley DG, Keady G (1997) The generation of internal waves by vibrating elliptic cylinders. Part 2. Approximate viscous solution. J Fluid Mech 351:119

Ivey GN, Nokes RI (1989) Vertical mixing due to the breaking of critical internal waves on sloping boundaries. J Fluid Mech 204:479

Ivey GN, Winters KB, Silva IPDD (2000) Turbulent mixing in a sloping benthic boundary layer energized by internal waves. J Fluid Mech 418:59

Lam F-PA, Maas LRM, Gerkema T (2004) Spatial structure of tidal and residual currents as observed over the shelf break in the Bay of Biscay. Deep Sea Res I 51:1075–1096

Maas LRM, Benielli D, Sommeria J, Lam FPA (1997) Observation of an internal wave attractor in a confined stably stratified fluid. Nature 388:557

McEwan AD (1973) Interactions between internal gravity waves and their traumatic effect on a continuous stratification. Bound Layer Met 5:159

Mowbray DE, Rarity BSH (1967) A theoretical and experimental investigation of the phase configuration of internal waves of small amplitude in a density-stratified liquid. J Fluid Mech 28:1

Oster G (1965) Density Gradients. Sci Am 213:70

Peacock T, Weidman P (2005) The effect of rotation on conical wave beams in a stratified fluid. Exp Fluids 39:32

Silva IPDD, Imberger J, Ivey GN (1997) Localized mixing due to a breaking internal wave ray at a sloping bed. J Fluid Mech 350:1

Teoh SG, Imberger J, Ivey GN (1997) Laboratory study of the interactions between two internal wave rays. J Fluid Mech 336:91

Thomas NH, Stevenson TN (1972) A similarity solution for viscous internal wave. J Fluid Mech 54:495

Thorpe SA (2005) The turbulent ocean. Cambridge University Press, Cambridge

Voisin B (2003) Limit states of internal wave beams. J Fluid Mech 496:243

GRAVITY WAVE DYNAMICS AND EFFECTS IN THE MIDDLE ATMOSPHERE

David C. Fritts and M. Joan Alexander
Colorado Research Associates Division
Northwest Research Associates
Boulder, Colorado, USA

Received 6 July 2001; revised 29 August 2002; accepted 6 November 2002; published 16 April 2003.

[1] Atmospheric gravity waves have been a subject of intense research activity in recent years because of their myriad effects and their major contributions to atmospheric circulation, structure, and variability. Apart from occasionally strong lower-atmospheric effects, the major wave influences occur in the middle atmosphere, between ~ 10 and 110 km altitudes because of decreasing density and increasing wave amplitudes with altitude. Theoretical, numerical, and observational studies have advanced our understanding of gravity waves on many fronts since the review by *Fritts* [1984a]; the present review will focus on these more recent contributions. Progress includes a better appreciation of gravity wave sources and characteristics, the evolution of the gravity wave spectrum with altitude and with variations of wind and stability, the character and implications of observed climatologies, and the wave interaction and instability processes that constrain wave amplitudes and spectral shape. Recent studies have also expanded dramatically our understanding of gravity wave influences on the large-scale circulation and the thermal and constituent structures of the middle atmosphere. These advances have led to a number of parameterizations of gravity wave effects which are enabling ever more realistic descriptions of gravity wave forcing in large-scale models. There remain, nevertheless, a number of areas in which further progress is needed in refining our understanding of and our ability to describe and predict gravity wave influences in the middle atmosphere. Our view of these unknowns and needs is also offered. *INDEX TERMS:* 3319 Meteorology and Atmospheric Dynamics: General circulation; 3334 Meteorology and Atmospheric Dynamics: Middle atmosphere dynamics (0341, 0342); 3379 Meteorology and Atmospheric Dynamics: Turbulence; 3384 Meteorology and Atmospheric Dynamics: Waves and tides; *KEYWORDS:* gravity waves; middle atmosphere dynamics; wave-mean flow interactions; wave-wave interactions; gravity wave instability; turbulence

Citation: Fritts, D. C., and M. J. Alexander, Gravity wave dynamics and effects in the middle atmosphere, *Rev. Geophys., 41*(1), 1003, doi:10.1029/2001RG000106, 2003.

1. INTRODUCTION

[2] Considerable progress has occurred in our understanding of middle atmosphere gravity waves in the four decades since the pioneering study by *Hines* [1960]. Indeed, interest in the field continues to expand as the many effects of gravity waves become better understood. Early results were reviewed extensively by *Fritts* [1984a], while subsequent reviews addressed specific aspects of gravity wave instability [*Fritts and Rastogi*, 1985], saturation [*Dunkerton*, 1989; *Fritts*, 1989], or climatologies and effects [*Vincent*, 1984a; *Schoeberl*, 1987; *Hamilton*, 1996; *McLandress*, 1998]. Arguably, the greatest quantitative advances in our understanding have been the most recent and have accompanied equally dramatic advances in observational and computational techniques and capabilities. This review will span middle atmosphere gravity wave dynamics and effects but will emphasize results subsequent to the review by *Fritts* [1984a].

[3] In retrospect, our understanding of middle atmosphere gravity waves in 1984 was quite limited. We had recognized the role of gravity waves in transporting energy and momentum, in contributing turbulence and mixing, and in influencing the mean circulation and thermal structure of the middle atmosphere, based on limited theoretical, numerical, and observational studies [*Lindzen*, 1981; *Holton*, 1982, 1983; *Vincent and Reid*, 1983]. However, we had little or no understanding of the characteristics of the gravity wave spectrum and of its behavior with altitude, of the variability imposed by sources and variable mean and low-frequency motions, of the global climatology of gravity waves and their effects, or of the processes and interactions that constrain wave amplitudes, effect spectral energy transfers, and account for turbulence, fluxes, and mixing. We also had explored only the most rudimentary schemes for describing gravity wave effects in large-scale models [*Lindzen*, 1981; *Holton*, 1982; *Matsuno*, 1982; *Dunkerton*, 1982a] and had no appreciation for the rich range of

Copyright 2003 by the American Geophysical Union.

8755-1209/03/2001RG000106*$15.00* doi:10.1029/2001RG000106

behaviors accompanying gravity waves throughout the middle atmosphere.

[4] More recent studies have yielded a more detailed understanding of middle atmosphere gravity waves on many fronts. In situ, ground-based, and space-based observational studies have contributed greatly to our knowledge of gravity wave scales, amplitudes, fluxes, and spectra. They have also addressed instability dynamics, vertical propagation, variations with altitude, and seasonal and geographic variability. Recent theoretical and numerical studies have addressed source characteristics and scales, spectral character, evolution, and energy transfers, instability dynamics, wave–wave and wave–mean flow interactions, and the implications of various parameterizations for atmospheric circulation and structure.

[5] Our purpose here is to review the contributions since 1984 that have led to our more advanced understanding of middle atmosphere gravity waves and their effects. In doing so, it will be impossible to cite all of the work that has been published. Instead, we will restrict ourselves to what we consider to be the more significant contributions in each area. We also will note, on occasion, where errors in interpretation have occurred and where debate over observational or theoretical implications continues to allow the reader to steer clear of mistakes and identify topics of current interest.

[6] We begin by reviewing the linear inviscid theory of gravity wave propagation as well as departures due to nonlinearity, dissipation via wave breaking, transience, localization, and variable environments, in section 2. Sources, propagation, and scales relevant to the middle atmosphere will be addressed in section 3. Spectral character and evolution, including evidence for and against universality, are reviewed in section 4. Section 5 summarizes what is known of gravity wave climatologies and their implications for sources and effects. Gravity wave interactions, instability dynamics, and wave-breaking or saturation processes are examined in section 6. Sections 7 and 8 describe efforts to parameterize gravity wave effects to date and their implementation and effects on the large-scale circulation and structure of the middle atmosphere. Our summary and recommendations are presented in section 9.

2. EQUATIONS OF MOTION

[7] Atmospheric gravity waves can frequently be described with a simple linear theory that treats them as small departures from a stably stratified background state varying only in the vertical. The restoring force for gravity wave oscillations is the buoyancy that results from the adiabatic displacements of air parcels characteristic of these disturbances. Here we outline a general gravity wave solution to the linearized forms of the fundamental conservation equations. We then discuss the simpler forms of the dispersion and polarization relations that result when limited portions of the possible gravity wave spectrum are considered.

2.1. Linear Theory

[8] We begin with the fundamental fluid equations in Cartesian coordinates (x, y, z) that follow from conservation of momentum, mass, and energy [e.g., *Holton*, 1992]:

$$\frac{du}{dt} - fv + \frac{1}{\rho}\frac{\partial p}{\partial x} = X\,, \tag{1}$$

$$\frac{dv}{dt} + fu + \frac{1}{\rho}\frac{\partial p}{\partial y} = Y\,, \tag{2}$$

$$\frac{dw}{dt} + \frac{1}{\rho}\frac{\partial p}{\partial z} + g = 0\,, \tag{3}$$

$$\frac{1}{\rho}\frac{d\rho}{dt} + \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0\,, \tag{4}$$

$$\frac{d\theta}{dt} = Q\,, \tag{5}$$

where d/dt represents a total, or advective, derivative; (u, v, w) is the fluid velocity vector; and the terms X, Y, and Q represent unspecified forcings that could include wave-driven forces and diffusive mixing effects accompanying wave dissipation. The remaining symbols have the usual meanings: p is pressure, ρ is density ($\rho = \rho_0 \exp[-(z - z_0)/H]$ with $\rho_0 = \rho(z_0)$ the density at reference level z_0 and H the scale height), and $f = 2\Omega sin\phi$ is the Coriolis parameter (where Ω is the Earth rotation rate and ϕ is latitude). These five equations plus the definition of potential temperature θ,

$$\theta = \frac{p}{\rho R}\left(\frac{p_0}{p}\right)^{\kappa}\,, \tag{6}$$

define a complete set describing inviscid fluid motions. Here $p_0 = p(z_0)$, R is the ideal gas constant, and $\kappa = c_p/c_v$ is the ratio of specific heats at constant pressure and constant volume. Here θ represents the temperature a parcel of air would have if lowered adiabatically from p to p_0.

[9] The unforced forms of equations (1)–(6) linearized about a horizontally uniform hydrostatic basic state with background wind ($\bar{u}$, $\bar{v}$, 0), potential temperature $\bar{\theta}$, pressure $\bar{p}$, and density $\bar{\rho}$ varying only in z are

$$\frac{Du'}{Dt} + w'\frac{\partial \bar{u}}{\partial z} - fv' + \frac{\partial}{\partial x}\left(\frac{p'}{\bar{\rho}}\right) = 0\,, \tag{7}$$

$$\frac{Dv'}{Dt} + w'\frac{\partial \bar{v}}{\partial z} + fu' + \frac{\partial}{\partial y}\left(\frac{p'}{\bar{\rho}}\right) = 0\,, \tag{8}$$

$$\frac{Dw'}{Dt} + \frac{\partial}{\partial z}\left(\frac{p'}{\bar{\rho}}\right) - \frac{1}{H}\left(\frac{p'}{\bar{\rho}}\right) + g\frac{\rho'}{\bar{\rho}} = 0\,, \tag{9}$$

$$\frac{D}{Dt}\left(\frac{\theta'}{\bar{\theta}}\right) + w'\frac{N^2}{g} = 0\,, \tag{10}$$

$$\frac{D}{Dt}\left(\frac{\rho'}{\bar{\rho}}\right) + \frac{\partial u'}{\partial x} + \frac{\partial v'}{\partial y} + \frac{\partial w'}{\partial z} - \frac{w'}{H} = 0\,, \quad (11)$$

$$\frac{\theta'}{\bar{\theta}} = \frac{1}{c_s^2}\left(\frac{p'}{\bar{\rho}}\right) - \frac{\rho'}{\bar{\rho}}\,. \quad (12)$$

Here the derivative D/Dt is the linearized form of the time derivative,

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \bar{u}\,\frac{\partial}{\partial x} + \bar{v}\,\frac{\partial}{\partial y}\,, \quad (13)$$

primed quantities are perturbations to the background state, and $N = (g\partial \ln\theta/\partial z)^{1/2}$ is the buoyancy frequency.

[10] We now neglect the background shear terms in equations (7) and (8) by assuming $(\bar{u}, \bar{v})$ and N vary only slowly over a wave cycle in the vertical (the WKB approximation [e.g., *Gill*, 1982]). We also assume the gravity wave solutions have the form

$$\left(u',\, v',\, w',\, \frac{\theta'}{\bar{\theta}},\, \frac{p'}{\bar{\rho}},\, \frac{\rho'}{\bar{\rho}}\right) = (\tilde{u}, \tilde{v}, \tilde{w}, \tilde{\theta}, \tilde{p}, \tilde{\rho}) \cdot \exp\left[i(kx + ly + mz - \omega t) + \frac{z}{2H}\right]. \quad (14)$$

These describe a monochromatic wave perturbation with wave number components (k, l, m) and ground-relative (Eulerian) frequency ω. Substitution into equations (7)–(12) yields a set of six algebraic equations for $(\tilde{u}, \tilde{v}, \tilde{w}, \tilde{\theta}, \tilde{p}, \tilde{\rho})$:

$$-i\hat{\omega}\tilde{u} - f\tilde{v} + ik\tilde{p} = 0\,, \quad (15)$$

$$-i\hat{\omega}\tilde{v} + f\tilde{u} + il\tilde{p} = 0\,, \quad (16)$$

$$-i\hat{\omega}\tilde{w} + \left(im - \frac{1}{2H}\right)\tilde{p} = -g\tilde{\rho}\,, \quad (17)$$

$$-i\hat{\omega}\tilde{\theta} + (N^2/g)\tilde{w} = 0\,, \quad (18)$$

$$-i\hat{\omega}\tilde{\rho} + ik\tilde{u} + il\tilde{v} + \left(im - \frac{1}{2H}\right)\tilde{w} = 0\,, \quad (19)$$

$$\tilde{\theta} = \tilde{p}/c_s^2 - \tilde{\rho}\,, \quad (20)$$

where $\hat{\omega} = \omega - k\bar{u} - l\bar{v}$ is called the intrinsic frequency, i.e., the frequency that would be observed in a frame of reference moving with the background wind $(\bar{u}, \bar{v})$. The above equations can be combined to form a single equation for the perturbation vertical velocity amplitude. Demanding the imaginary coefficients of this equation go to zero gives,

$$\frac{g}{c_s^2} = \frac{1}{H} - \frac{N^2}{g}\,, \quad (21)$$

$$\hat{\omega}^2\left(k^2 + l^2 + m^2 + \frac{1}{4H^2} - \frac{(\hat{\omega}^2 - f^2)}{c_s^2}\right) = N^2(k^2 + l^2) + f^2\left(m^2 + \frac{1}{4H^2}\right). \quad (22)$$

The second equation is fourth-order in $\hat{\omega}$ and supports both acoustic and gravity waves. By letting the sound speed $c_s \to \infty$ we retain the compressibility term related to the background density gradient and obtain the gravity wave dispersion relation,

$$\hat{\omega}^2 = \frac{N^2(k^2 + l^2) + f^2\left(m^2 + \frac{1}{4H^2}\right)}{k^2 + l^2 + m^2 + \frac{1}{4H^2}}\,, \quad (23)$$

or alternatively for the vertical wave number as

$$m^2 = \frac{(k^2 + l^2)(N^2 - \hat{\omega}^2)}{(\hat{\omega}^2 - f^2)} - \frac{1}{4H^2}\,. \quad (24)$$

The dispersion relation relates the wave frequency to the wave's spatial characteristics (wave numbers) and to the background atmosphere properties N and $(\bar{u}, \bar{v})$. The gravity wave solutions above admit waves with a broad range of properties. For vertically propagating waves, (k, l, m) are real, and the intrinsic frequency is confined to the range $N > \hat{\omega} > |f|$. All parts of this spectrum can be important to middle-atmosphere dynamics. Because of this wide range of properties and because the intrinsic frequency and vertical wave number can vary strongly as a wave propagates vertically through background shear and stability gradients, models describing their propagation across a wide range of conditions may need to retain all terms in equation (23) in their analysis. Simpler forms representing waves in certain limited regions of this spectral space, however, provide considerable insight and will be summarized below.

[11] The wave group velocity describes energy transport and wave packet propagation and is written as

$$(c_{gx}, c_{gy}, c_{gz}) = \left(\frac{\partial\omega}{\partial k}, \frac{\partial\omega}{\partial l}, \frac{\partial\omega}{\partial m}\right) = (\bar{u}, \bar{v}, 0) + \frac{[k(N^2 - \hat{\omega}^2), l(N^2 - \hat{\omega}^2), -m(\hat{\omega}^2 - f^2)]}{\hat{\omega}\left(k^2 + l^2 + m^2 + \frac{1}{4H^2}\right)}. \quad (25)$$

Our sign convention here assumes that $\hat{\omega}$ is positive definite. The vector (k, l) defines the direction of horizontal propagation. Then m is negative for upward group velocity and positive for downward. Thus, for eastward $(k > 0)$ and upward $(m < 0)$ energy propagation relative to the wind, intrinsic phase speeds are eastward $(\hat{c}_x = \hat{\omega}/k)$ and downward $(\hat{c}_z = \hat{\omega}/m)$ relative to the wind. Note that the phase speed is not a vector quantity, although wave phase propagation has a direction given by the vector (k, l, m).

[12] Polarization relations relating perturbation amplitudes of the different variables to one another can also be derived from equations (15)–(20). Some useful examples are

$$\tilde{u} = \left(\frac{i\hat{\omega}k - fl}{i\hat{\omega}l + fk}\right)\tilde{v}, \tag{26}$$

$$\tilde{p} = \left(\frac{\hat{\omega}^2 - f^2}{\hat{\omega}k + ifl}\right)\tilde{u} = \left(\frac{\hat{\omega}^2 - f^2}{\hat{\omega}l - ifk}\right)\tilde{v}, \tag{27}$$

$$\tilde{w} = \frac{\left(m - \frac{i}{2H}\right)\hat{\omega}}{N^2 - \hat{\omega}^2}\tilde{p}, \tag{28}$$

and equation (18).

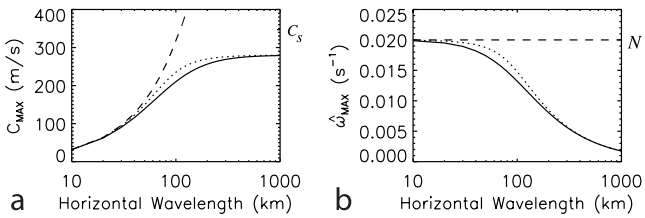


Figure 1. Analytical solutions derived from linear theory for the (a) maximum intrinsic phase speed (c_{MAX}) and (b) maximum intrinsic frequency (ω_{MAX}) at the point of total internal reflection, also known as the turning level. The dashed and solid lines are the incompressible and fully compressible solutions, respectively. The dotted line shows the approximation given by equation (29).

2.1.1. High-Frequency Waves: $\hat{\omega} \gg f$

[13] The dispersion relation (23) is derived by neglecting acoustic wave solutions, but the terms proportional to $1/4H^2$ retain one of the larger compressibility effects associated with gravity wave motions. These arise when the vertical wave number m becomes small so that $1/2H$ is a nonnegligible factor. With $H \sim$ 7 km in the middle atmosphere, this term becomes significant for waves with vertical wavelengths of ~30 km or longer, where the change in background density becomes significant over the vertical wavelength of the wave. This compressibility effect is important in the interpretation of airglow images of gravity wave perturbations near the mesopause [e.g., *Swenson et al.*, 2000]. As the vertical wavelength grows very large and $m \rightarrow 0$, the wave will undergo total internal reflection where the vertical group velocity can change sign (called a turning level). Above this level, the wave is evanscent. If turning levels occur both above and below, the wave is said to be ducted. Observational and modeling studies associated with these effects will be described further in section 3.2.2.

[14] Neglecting rotation effects but retaining the scale-height compressibility term and setting $m = 0$ in equation (23) yields an equation describing the maximum horizontal intrinsic phase speed $\hat{c}_{max}$ prior to total internal reflection as a function of horizontal wavelength:

$$|\hat{c}_{max}| = N\left(k^2 + l^2 + \frac{1}{4H^2}\right)^{-1/2}. \tag{29}$$

This gives a useful approximation to the fully compressible solution [*Marks and Eckermann*, 1995]. Figure 1 compares solutions for $\hat{c}_{max}$ and the corresponding maximum intrinsic frequency $\hat{\omega}_{max}$ versus horizontal wavelength for the fully compressible, incompressible, and approximate equation (29) solutions.

[15] Figure 1a shows that short horizontal wavelengths are much more easily reflected and will likely remain trapped at lower altitudes. Thus only waves with horizontal wavelengths >10 km are generally considered important to middle-atmosphere dynamics. Although the vertical wavelength theoretically goes to infinity at the turning level (and therefore the WKB approximation should be violated), the theory embodied in equation (29) appears to provide a good approximation even for waves observed near this theoretical limit [*Swenson et al.*, 2000].

[16] For high-frequency waves for which the Coriolis force can be neglected and for which $m^2 \gg 1/4H^2$, the dispersion relation simplifies to

$$\hat{\omega}^2 = \frac{N^2(k^2 + l^2)}{k^2 + l^2 + m^2} = N^2\cos^2\alpha, \tag{30}$$

where α is the angle between lines of constant phase and the vertical. These are simple plane waves with zero velocity perturbations perpendicular to the plane of propagation. Group velocity simplifies to

$$(c_{gh}, c_{gz}) = (\bar{u}_h, 0) + \frac{Nm}{k_h^2 + m^2}\frac{(m, -k_h)}{(k_h^2 + m^2)^{1/2}}, \tag{31}$$

where c_{gh} is the horizontal component of group velocity, $\bar{u}_h$ the horizontal wind in the direction of propagation, and $k_h = \sqrt{k^2 + l^2}$ is the horizontal wave number. Note that phase propagation is downward ($m < 0$) for upward group velocity. In this form it can be seen that group velocity is parallel to lines of constant phase and perpendicular to phase propagation.

[17] High intrinsic frequency waves are observed not only in mesopause airglow image data [*Swenson and Espy*, 1995; *Taylor and Garcia*, 1995; *Gardner et al.*, 1996; *Taylor et al.*, 1997; *Isler et al.*, 1997; *Hickey et al.*, 1998; *Haque and Swenson*, 1999] but also in sublimb viewing infrared and microwave satellite observations in the stratosphere [*Wu and Waters*, 1996a, 1996b; *Alexander*, 1998; *Dewan et al.*, 1998; *Picard et al.*, 1998; *McLandress et al.*, 2000] and in middle-atmosphere radar data [*Reid et al.*, 1988; *Fritts et al.*, 1990a, 1992; *Fritts and Wang*, 1991; *Hoppe and Fritts*, 1995; *Mitchell and Howells*, 1998].

2.1.2. Medium-Frequency Waves: $N \gg \hat{\omega} \gg f$

[18] For gravity waves with midrange intrinsic frequencies, very simplified relationships emerge that lend valuable insight into gravity wave properties and the effects due to changes in the background wind and stability. Letting c_h represent the horizontal phase speed and $k_h = \sqrt{k^2 + l^2}$, the dispersion relation then simplifies to

$$\hat{\omega} = N\left|\frac{k_h}{m}\right|, \tag{32}$$

and the vertical wave number can now be very simply related to the background wind and stability:

$$|m| = \frac{N}{|\hat{c}_h|}. \tag{33}$$

Note that $k_h > 0$ for $\hat{c}_h = (c_h - \bar{u}_h) > 0$, and $k_h < 0$ for $\hat{c}_h < 0$. Equations (32) and (33) demonstrate that vertical wavelength and intrinsic frequency are both proportional to the intrinsic phase speed $\hat{c}_h = c_h - \bar{u}_h$. The level where $c_h = \bar{u}_h(z)$ is a critical level for the wave where vertical wavelength shrinks to zero. This theoretical limit is never achieved in the real atmosphere because a host of instability and dissipation mechanisms become more likely as a wave approaches a critical level.

[19] The group velocity for medium-frequency gravity waves becomes

$$(c_{gh}, c_{gz}) = (\bar{u}_h, 0) + \left(\frac{\hat{\omega}}{k_h}, \frac{-\hat{\omega}}{m}\right) \tag{34}$$

and has the same magnitude as the phase propagation but opposite sign in the vertical. The time-dependent ray equations [*Lighthill*, 1978] show that a wave packet requires an infinite time to reach a critical level (with $m \to \infty$), implying that time-dependent and nonlinear effects will become important as a critical level is approached. Note, however, that wave transience and time dependence of the background flow can allow different interactions near the critical level. Departures from linear theory due to transience and nonlinearity are discussed in section 6.

[20] The polarization relations under the medium-frequency approximation take simple forms:

$$\tilde{p} = c_{gz}\tilde{w}, \tag{35}$$

$$\tilde{w} = -\frac{k_h}{m}\tilde{u}_h = \frac{k_h\hat{c}_h}{N}\tilde{u}_h = \frac{\hat{\omega}}{N}\tilde{u}_h. \tag{36}$$

Note that group velocity, vertical wavelength, and intrinsic frequency increase as a wave propagates "upshear," defined as an altitude region where the intrinsic phase speed $|\hat{c}_h|$ is increasing. These simplified forms demonstrate that vertical velocity will grow proportionately larger than the horizontal velocity perturbations in these conditions, and pressure perturbations become increasingly important. The opposite is true for waves propagating "downshear" such as those approaching a critical level ($|\hat{c}_h| \to 0$). In fact, it is the tendency for $\tilde{u}/|\hat{c}_h|$ to increase as the wave approaches a critical level that drives the tendency for instability and turbulence generation. These approximations for waves with midrange frequencies are inappropriate for most detailed analyses, but the set of very simple relationships that results provides an extremely valuable intuitive view of gravity wave properties and their relationship to background winds and stability.

2.1.3. Low-Frequency Waves: $\hat{\omega} \sim f$

[21] Low-frequency gravity waves, or inertia-gravity waves, are those for which the rotation of the Earth has an important influence. These wave perturbations have an interesting three-dimensional helical structure. Velocity perturbations in the direction perpendicular to the direction of propagation are no longer zero but grow proportionately larger as intrinsic frequency decreases toward f. For a zonally propagating wave the meridional velocity amplitude $\tilde{v}$ is

$$\tilde{v} = -i\frac{f}{\hat{\omega}}\tilde{u}. \tag{37}$$

An approximation to the dispersion relation appropriate for both low- and medium-frequency gravity waves is

$$\hat{\omega}^2 = N^2\frac{k_h^2}{m^2} + f^2. \tag{38}$$

Any wave nearing a critical level (where the background wind speed approaches the horizontal phase speed and $m \to \infty$) will move toward lower intrinsic frequency where rotation effects become important. The true critical level is therefore where $\hat{\omega} \to f$, somewhere below the critical level described for medium-frequency waves ($c_h = \bar{u}_h$).

[22] Because the ratio of vertical to horizontal group velocity,

$$\left|\frac{c_{gz}}{c_{gh}}\right| = \left|\frac{k_h}{m}\right| = \frac{(\hat{\omega}^2 - f^2)^{1/2}}{N}, \tag{39}$$

becomes quite small for inertia-gravity waves relative to their higher-frequency counterparts, they can be found at large horizontal distances from their sources [*Dunkerton*, 1984; *Eckermann*, 1992]. They are very commonly seen in lower stratosphere observations.

[23] A number of observational techniques are sensitive to inertia-gravity waves. Balloon-borne measurements of horizontal wind profiles in the lower stratosphere enable inertia-gravity wave hodograph and rotary spectral analyses [*Kitamura and Hirota*, 1989; *Barat and Cot*, 1992; *Tsuda et al.*, 1994a; *Ogino et al.*, 1995, *Hamilton and Vincent*, 1995; *de la Torre et al.*, 1996; *Sato et al.*, 1994; *Sato and Dunkerton*, 1997; *Shimizu and Tsuda*,

1997; *Vincent and Alexander*, 2000]. Limb viewing satellite measurements can also be sensitive to inertia-gravity waves [*Fetzer and Gille*, 1994, 1996; *Eckermann and Preusse*, 1999; *Tsuda et al.*, 2000], and they are prominent signals in stratospheric radar and lidar observations [*Sato*, 1994; *Mitchell et al.*, 1994; *Sato et al.*, 1997; *Riggin et al.*, 1997] and middle-atmosphere rocket sounding data [*Hirota and Niki*, 1985; *Hamilton*, 1991; *Tsuda et al.*, 1992; *Eckermann et al.*, 1995; *Fritts et al.*, 1997a].

2.2. Conservative Propagation and Departures

[24] Gravity waves conserve wave pseudomomentum flux if they propagate without dissipation through a background atmosphere that varies only in the vertical. The vertical flux of horizontal pseudomomentum

$$(F_{Px}, F_{Py}) = \bar{\rho} c_{gz} \frac{E}{\hat{\omega}} (k, l) \tag{40}$$

is related to the Reynolds stress and vertical flux of horizontal momentum under these assumptions as

$$(F_{Px}, F_{Py}) = \bar{\rho}(1 - f^2/\hat{\omega}^2)(\overline{u'w'}, \overline{v'w'}). \tag{41}$$

Dissipation results from processes such as radiative damping [*Fels*, 1984; *Zhu*, 1994], wave–wave and wave–mean flow interactions [*Broutman and Grimshaw*, 1988; *Sutherland*, 2000, 2001], and wave breaking and instability processes (see section 6). We note here only that nonlinearity and time dependence usually prevent the attainment of amplitudes anticipated by linear, inviscid, steady wave theory.

[25] Viscosity and molecular diffusion are generally not important dissipation mechanisms except in the thermosphere where the exponentially decreasing density and growth of kinematic viscosity with altitude ultimately allow these to dominate over other dissipation processes. (We note here that the "viscosity waves" proposed by *Hocking et al.* [1991] are, in reality, the viscous solutions of the Orr-Sommerfeld equation [e.g., *Fritts and Geller*, 1976]. They do not propagate, and they do not have a dispersion relation. Thus their role is to eliminate, not create, the artificial singularities that occur in the linear inviscid equations, and they cannot play the role proposed by *Hocking et al.* [1991] accounting for enhanced radar backscatter.) This growth of kinematic viscosity with altitude accounts for the turbopause at an altitude typically ~110 km.

[26] The negative of the zonal mean pseudomomentum flux is the gravity wave contribution to the more general Eliassen-Palm flux (EP flux) that is commonly used to describe planetary-scale wave effects [*Edmon et al.*, 1980; *Miyahara et al.*, 1986]. The divergence of the EP flux is proportional to the force exerted on the mean flow. In the case of gravity waves, the zonal average is replaced by an average over some smaller, but still large, area such as a GCM grid scale, and the force on the background wind on that scale becomes

$$(\bar{X}, \bar{Y}) = -\frac{\varepsilon}{\bar{\rho}} \frac{\partial}{\partial z} (F_{Px}, F_{Py}). \tag{42}$$

This is the wave driving (or wave-induced force) [*Andrews and McIntyre*, 1976] that would be used for studying wind evolution in the momentum conservation equation for the background flow as in equations (1) and (2). The force is a vector quantity such that the direction and sign of the force are always the direction and sign of the horizontal phase propagation in the frame of reference moving with the wind (i.e., the horizontal intrinsic phase speed of the wave).

[27] The factor ε has been called an efficiency [*Holton*, 1982] or intermittency [*Alexander and Dunkerton*, 1999] factor, and it can be thought of as describing the fractional coverage of the wave dissipation event within the larger-scale space- and/or time-averaging interval. The intermittency factor is a necessary ingredient in gravity wave parameterization estimates of $(\bar{X}, \bar{Y})$ when the input parameters are describing mean gravity wave properties. We note, however, that for observational estimates of (F_{Px}, F_{Py}) and $(\bar{X}, \bar{Y})$, such as derived from time-averaged radar data, the ε factor is already included.

[28] Gravity waves also have an associated energy flux, so departures from conservative propagation will lead to energy dissipation and local heating of the atmosphere. If wave dissipation involves turbulence generation, then mixing and turbulent heat, momentum, and constituent transport will also occur. Direct heating has been neglected in most GCMs that parameterize gravity wave effects. Models of the middle atmosphere, however, generally require parameterized vertical mixing that is commonly called eddy diffusion. Such mixing could, for example, result from convective overturning of material surfaces and subsequent turbulence following gravity wave breaking (see section 6.2.1).

[29] A momentum eddy diffusion coefficient associated with gravity wave dissipation may be approximated for a single wave [*Lindzen*, 1981; *Holton*, 1982] as

$$D_M = \frac{\hat{c}|(\bar{X}, \bar{Y})|}{N^2}. \tag{43}$$

This coefficient is related to the thermal eddy diffusion coefficient via the Prandtl number. The value of the Prandtl number associated with gravity wave breaking is poorly defined at present and may vary with height. It is expected to be large and highly variable based on numerical and theoretical models of gravity wave convective instability [*Chao and Schoeberl*, 1984; *Fritts and Dunkerton*, 1985; *Coy and Fritts*, 1988; *McIntyre*, 1989]. Assessments of turbulent thermal diffusion required to account for observed atmospheric structure in global models likewise imply large turbulent Prandtl numbers and smaller thermal diffusion coefficient than the momentum diffusion coefficient predicted by equation (43) [*Strobel et al.*, 1985, 1987; *Holton and Schoeberl*, 1988; *Smith and Brasseur*, 1991; *Huang and Smith*, 1991].

While equation (43) with an arbitrary scaling factor is very commonly applied, this relation remains to be confirmed and extended to multiple waves by direct numerical simulation.

[30] Departures from linear theory also occur because of the time dependence of the background flow and the spatial and temporal variations in wave activity (i.e., wave packets). In the case of a gravity wave packet that is localized in space and/or time, the result is a spread of horizontal and/or vertical wave numbers and a corresponding spread of intrinsic frequencies. A gravity wave packet in a shear flow may achieve large amplitudes, but it never attains a singular critical level solution. Spatial localization of a wave packet is also now understood to have some surprising effects. Gravity waves due to local topography comprise a continuous spectrum of wave numbers and propagation directions. When such a spectrum of gravity waves is incident on a veering wind, the result is a monotonic decrease in momentum flux and a turning of the direction of wave drag with altitude as successive portions of the spectrum are removed by critical levels [*Hines*, 1988a; *Shutts*, 1995, 1998; *Broad*, 1995, 1999]. At larger wave amplitudes, localization of a wave packet in the vertical can result in its penetration via nonlinear "self-acceleration" beyond its linear critical level [*Fritts and Dunkerton*, 1984], while localization horizontally and vertically can enable penetration through a turning level [*Sutherland*, 2000] (see section 6.2.3).

[31] Time dependence of the flow in which a gravity wave propagates also influences wave characteristics, enabling the formation of caustics, inducing wave number transport, and either accelerating or retarding the approach toward unstable local wave amplitudes [*Broutman and Young*, 1986; *Bruhwiler and Kaper*, 1995; *Eckermann*, 1997; *Buckley et al.*, 1999; *Walterscheid*, 2000]. Given the importance of (or lack of) these processes in the spectral evolution of gravity waves in the middle atmosphere and the current debate within the community over these issues, this topic will be reviewed more fully in section 6.

3. SOURCES AND PROPAGATION

3.1. Gravity Wave Sources

[32] A complete survey of the gravity wave sources important for the middle atmosphere is beyond the scope of this paper. Indeed, such a survey is beyond our current state of knowledge based on the extensive theoretical, modeling, and observational efforts to date. Instead, the purpose of this section is to identify the dominant sources, their spatial and temporal variability, and the general characteristics of gravity waves arising from them as we understand them at present. The most obvious sources include topography, convection, and wind shear, although other sources may also be significant at preferred sites or in association with specific larger-scale dynamics. Examples of other sources that are likely to be important statistically include adjustment of unbalanced flows in the vicinity of jet streams and frontal systems, body forcing accompanying localized wave dissipation, and wave–wave interactions. Less significant sources which may, nevertheless, be important at higher altitudes include auroral heating and eclipse cooling.

3.1.1. Topographic Generation

[33] Mountain waves have been studied extensively over the past three decades using theoretical, numerical, and observational methods. As a result, idealized two-dimensional (2-D) responses to 2-D topography are reasonably well understood in terms of nonlinear wave dynamics and hydraulic theory [*Long*, 1955; *Smith*, 1985; *Durran and Klemp*, 1987; *Nance and Durran*, 1998; *Lott*, 1998; *Farmer and Armi*, 1999]. Responses to three-dimensional (3-D) topography have also been quantified increasingly in a number of measurement programs and numerical studies, yielding estimates of mountain wave scales, amplitudes, and momentum fluxes. Horizontal wavelengths for vertically propagating waves are typically tens to hundreds of kilometers, while amplitudes vary from small to breaking [*Lilly and Kennedy*, 1973; *Nastrom and Fritts*, 1992; *Chan et al.*, 1993; *Dörnbrack et al.*, 1999; *Leutbecher and Volkert*, 2000]. Horizontal and vertical cross sections of a wave structure anticipated over northern Scandinavia in a 3-D mesoscale simulation by *Dörnbrack et al.* [1999] are shown in Figure 2 and exhibit the large-scale thermal structure (areas of low temperature in Figure 2) believed to trigger polar stratospheric clouds (PSC) and ozone depletion. Such simulations, however, do not yet capture the responses often observed over smaller-scale topography [*Nastrom and Fritts*, 1992].

[34] Momentum fluxes likewise vary greatly but are typically in the range of ~0.1 to a few $m^2\ s^{-2}$ (per unit mass, or ~0.01–0.5 Pa) in the upper troposphere and lower stratosphere based on in situ aircraft and radar measurements and large-scale cross-mountain pressure distributions [*Lilly and Kennedy*, 1973; *Lilly and Lester*, 1974; *Smith*, 1978; *Kennedy and Shapiro*, 1979; *Shutts et al.*, 1988; *Fritts et al.*, 1990b; *Sato*, 1994]. Momentum fluxes are also preferentially associated with the smaller horizontal scale waves that are vertically propagating because of their larger vertical velocities. Maximum (generally negative in Figure 2) momentum fluxes were inferred at horizontal wavelengths between ~10 and 100 km by *Nastrom and Fritts* [1992] (see Figure 3).

[35] The Global Atmospheric Sampling Program (GASP) and more recent aircraft observations also enabled assessments of the statistical importance of mountain waves over rough and smooth terrain and relative to other significant sources. These showed horizontal velocity and temperature variances to be ~2–3 times higher over significant topography compared to plains and oceans, independent of other sources [*Nastrom et*

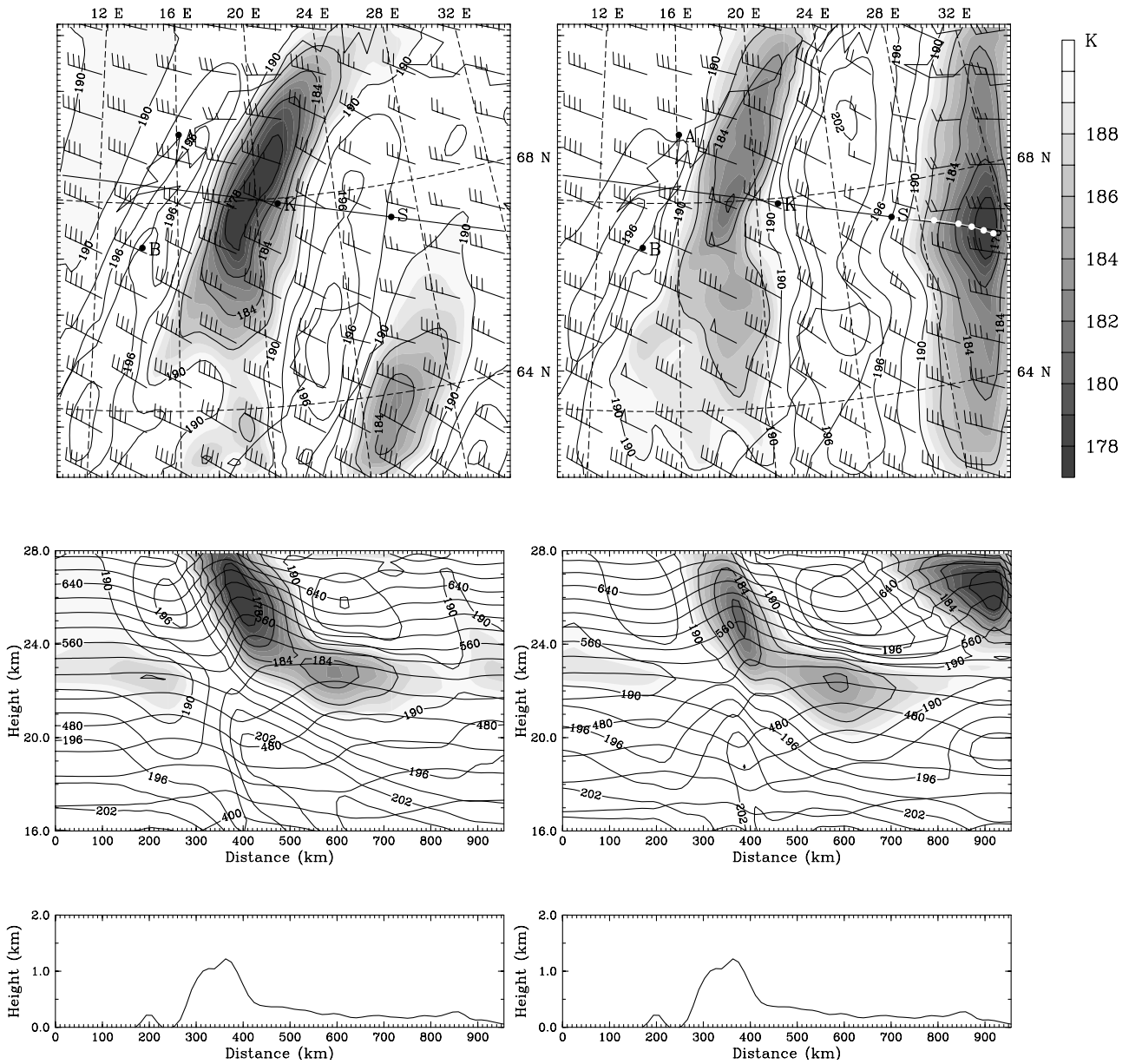


Figure 2. (top) Horizontal and (middle) vertical cross sections of temperature and horizontal wind in a 3-D mesoscale simulation of flow over northern Scandinavia. (bottom) Terrain height beneath the vertical cross section. Modified after *Dörnbrack et al.* [1999].

al., 1987; *Jasperson et al.*, 1990; *Bacmeister et al.*, 1990a] and ~5 times higher than regions having no obvious meteorological sources [*Fritts and Nastrom*, 1992]. Because of the importance of such waves for mean circulations a number of authors have attempted to parameterize their specific effects in large-scale models (see section 7). The importance of mountain waves relative to other significant sources is discussed further in section 5.

[36] Of relevance to the middle atmosphere, then, mountain waves have dominant scales (in terms of momentum fluxes) of ~10–100 km, phase speeds near zero and vertical wavelengths dictated by the local static stability and mean wind in the plane of wave propagation [*Bacmeister et al.*, 1990b]. Indeed, *Bacmeister* [1993] has argued that mountain waves likely account for a large fraction of zonally averaged wave-induced force (wave drag) in the mesosphere, although this result is based on predictions of Lindzen-type parameterizations (see section 7) and remains to be demonstrated observationally.

3.1.2. Convective Generation

[37] Although it has been known for decades that convection can excite gravity waves, it has only been in more recent years that observations and models have begun to characterize the waves generated by this source and the mechanisms within convection responsible for their generation. In fact, there is at the time of this

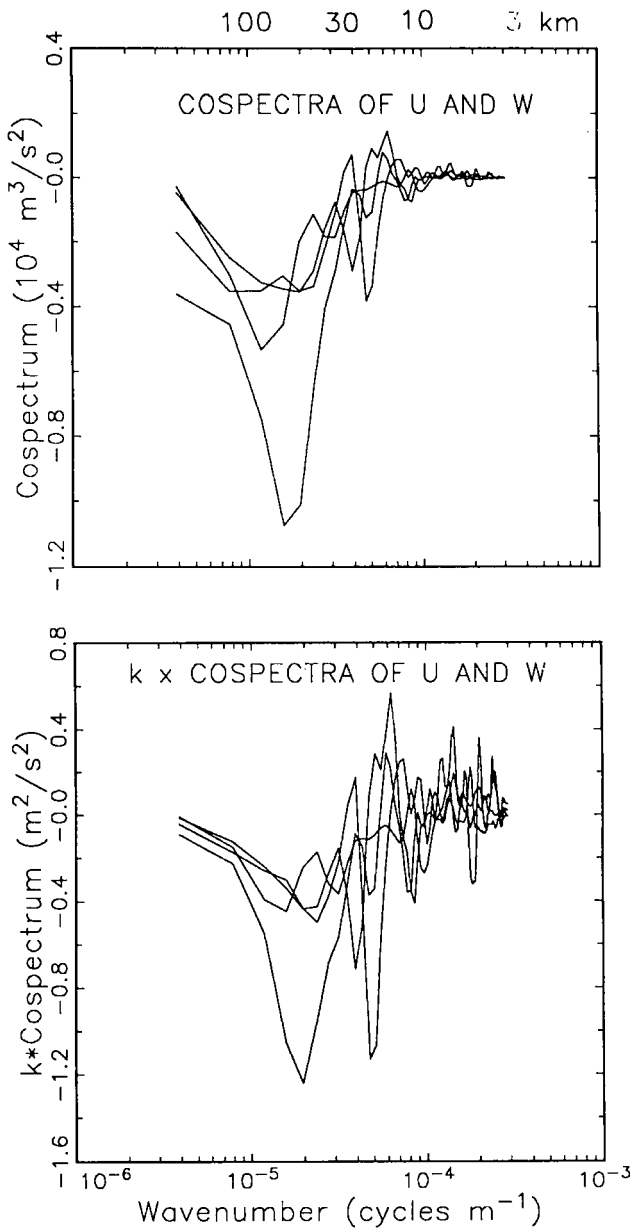


Figure 3. Spectra of momentum flux in (top) standard and (bottom) variance content forms computed for four flights over the central Rockies. Note that the dominant fluxes in each case are largely at horizontal scales between 10 and 100 km. After *Nastrom and Fritts* [1992] (reprinted with permission of the American Meteorological Society).

review still considerable controversy and ongoing research aimed at understanding this wave generation mechanism.

[38] One of the difficulties in characterizing this source is its inherent intermittency. Observations of high-frequency waves in the stratosphere have shown a close correspondence with deep convective clouds (see Figure 4) [*Sato*, 1992, 1993; *Alexander and Pfister*, 1995; *Sato et al.*, 1995; *Dewan et al.*, 1998; *McLandress et al.*, 2000; *Alexander et al.*, 2000]. Observational estimates of momentum flux are fewer but show significant variability. Local magnitudes range up to ~0.03–0.2 Pa [*Sato*, 1993; *Alexander and Pfister*, 1995; *Alexander et al.*, 2000], and longer-term averages peak at $\sim 1\text{–}5 \times 10^{-3}$ Pa [*Sato and Dunkerton*, 1997; *Sato et al.*, 1997; *Vincent and Alexander*, 2000].

[39] Waves generated by convection are not characterized by a single prominent phase speed or frequency as is the case for topographic waves. Instead, convection can generate waves throughout the full range of phase speeds, wave frequencies, and vertical and horizontal scales. The low-frequency waves, in particular, may be observed in the middle atmosphere at large horizontal distances from the convective source, making correlation with clouds or other indicators of convection more difficult. In the tropics, however, far from topography and regions of baroclinic instability, the occurrence of inertia-gravity waves has been linked to convection as the source [*Pfister et al.*, 1986; *Tsuda et al.*, 1994a; *Karoly et al.*, 1996; *Shimizu and Tsuda*, 1997; *Wada et al.*, 1999; *Vincent and Alexander*, 2000].

[40] Model studies of waves generated by convection (Figure 5) have greatly enhanced our understanding of these phenomena. Convection involves a time-varying thermal forcing associated with latent heat release that can interact with overlying stable layers and shear in complex ways that are not fully understood. Three simplified mechanisms have been proposed to describe convective generation: (1) pure thermal forcing, (2) an "obstacle" or "transient mountain" effect, and (3) a "mechanical oscillator" effect.

[41] For pure thermal forcing in the absence of strong shear the heating depth and the vertical wavelength of the gravity waves generated are surprisingly approximately equal [*Alexander et al.*, 1995; *Piani et al.*, 2000]. This relationship (demonstrated by *Salby and Garcia* [1987] for larger-scale waves) arises because the heating projects most strongly onto a wave with vertical wavelength twice the depth of the heating, but the wave is then refracted to half that vertical wavelength as it propagates across the twofold increase in buoyancy frequency at the tropopause (see equation (33)). *Bergman and Salby* [1994] used global cloud imagery at tropical latitudes to estimate the planetary-scale and long-wavelength, low-frequency gravity wave responses with this model, and the far-field effects of the waves generated have been extended through the mesosphere by *Garcia and Sassi* [1999]. The linear mechanism used in these studies has been validated for tropical planetary-scale waves, but the amplitudes are generally larger than those in corresponding nonlinear models [*Manzini and Hamilton*, 1993]. The gravity wave momentum fluxes derived from these cloud data sets are still highly uncertain. *Pandya and Alexander* [1999] further describe how the local mesoscale environment surrounding the storm can modify the smaller-scale gravity wave response in the stratosphere from a latent heat source. Observational evidence for the thermal forcing mechanism of gravity wave generation is given by *McLandress et al.* [2000]

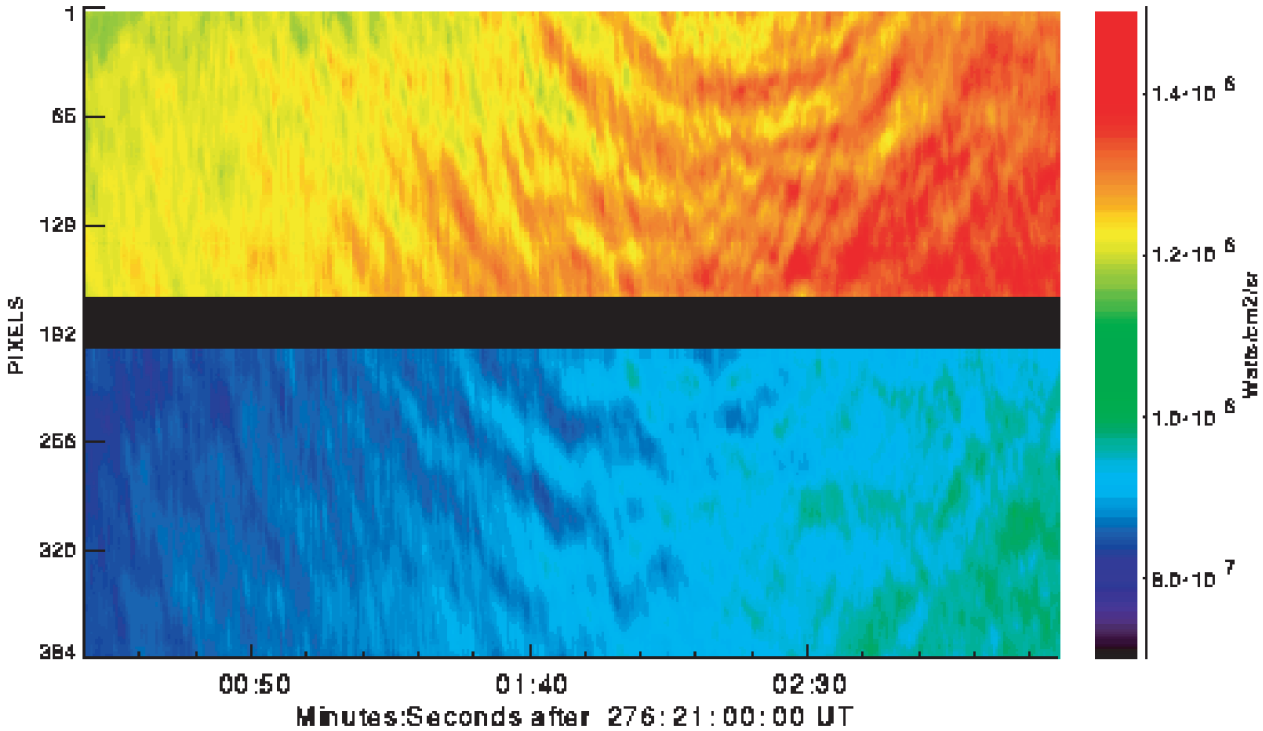


Figure 4. Images of infrared emissions from the stratosphere observed by the SPIRIT III radiometer from the MSX satellite. Circular phase fronts of waves with horizontal wavelength ~25 km can be seen in these images that were centered around a deep convective cloud below in the troposphere. After *Dewan et al.* [1998, Figure 1].

(Figure 6), although its importance relative to other mechanisms could not be quantified. Pure thermal forcing would be expected to generate an isotropic wave field, with anisotropies occurring primarily via background wind filtering effects.

[42] The obstacle effect was descibed by *Clark et al.* [1986] for short horizontal wavelength waves in the troposphere generated by boundary layer convection [*Kuettner et al.*, 1987], but the ideas developed in their study have also contributed to understanding the larger horizontal scale waves in the middle atmosphere generated by convection. The obstacle effect can be thought of as analogous to topographic wave generation where the convective heating modifies the shape of isentropes at the bottom of a stable layer which is moving relative to the convective obstacle. *Clark et al*. [1986] found that both the pure thermal forcing and the obstacle effect produced gravity waves in the overlying stable layer but that the obstacle effect generated much larger amplitude waves. *Pfister et al.* [1993a, 1993b] explicitly modeled a linear version of this mechanism and found evidence for it in stratospheric aircraft observations over convection. *Pfister et al.* [1993b] also described how the time dependence of the obstacle (or transient mountain) affects the phase speed spectrum of the waves that are generated. Observational evidence for the obstacle effect has appeared in radiosonde analyses of low-frequency waves [*Vincent and Alexander*, 2000; *Alexander and Vincent*, 2000]. This mechanism is associated with an anisotropic wave spectrum with a preference for generating waves propagating opposite to the mean wind relative to the obstacle.

[43] *Fovell et al.* [1992] described the "mechanical oscillator" mechanism as a periodic and localized momentum source term that generates waves with a frequency equal to the oscillation frequency. This mechanism is not distinct from the time-varying thermal source when that source oscillates with a regular period [*Clark et al.*, 1986; *Alexander et al.*, 1995; *Pandya and Alexander*, 1999]. *Lane et al.* [2001] describe a similar mechanism acting in their 3-D model that had little or no shear in the forcing region, with some added constraints. In their conceptual model the parcel oscillates about its level of neutral buoyancy at the local tropospheric buoyancy frequency which limits this mechanism to generating only very high frequency gravity waves in the upper troposphere at altitudes above any lower minima that occur in the N profile. For the mechanical oscillator mechanism, anisotropies in the wave spectrum can result from background wind effects in a manner similar to the obstacle effect [*Fovell et al.*, 1992].

[44] In summary, the three simplified excitation mechanisms that have been proposed may all be important, depending strongly upon the local shear and the vertical profile and time dependence of the latent heating. In reality, the three mechanisms are not distinct but cou-

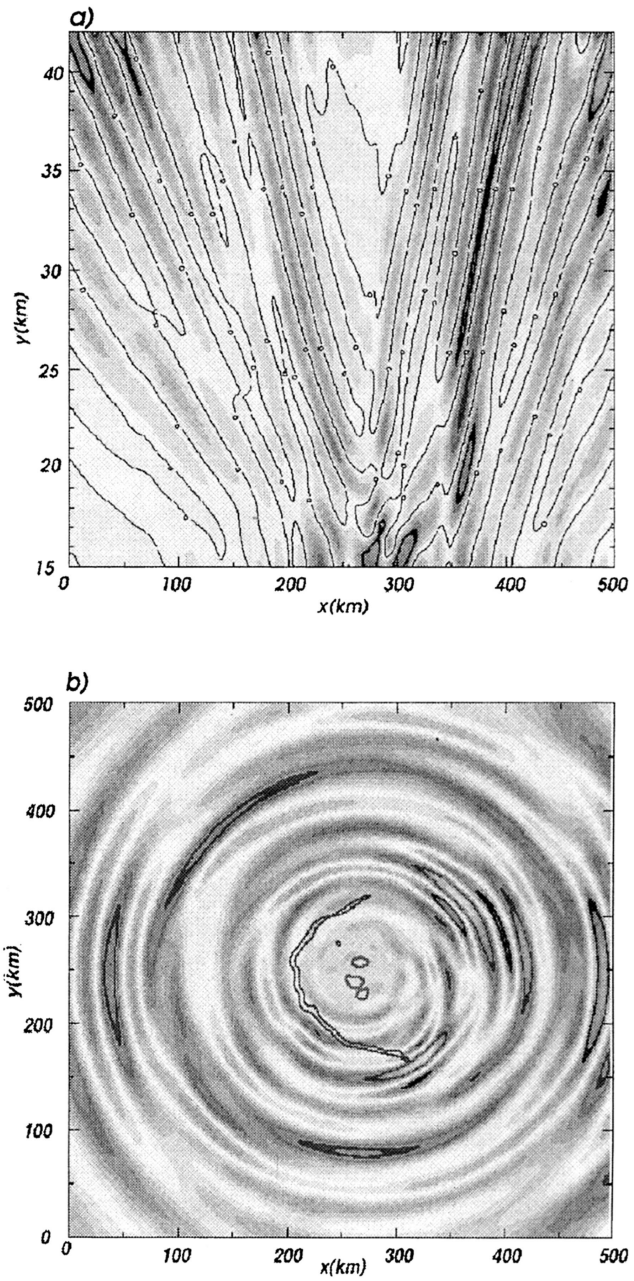


Figure 5. Three-dimensional study of gravity waves generated by convection in a mesoscale model with parameterized microphysics. (a) Vertical velocity patterns in a cross section in the vertical (z) and zonal (x) plane at $y = 250$ km. (b) The $x - y$ cross section of vertical velocity at $z = 40$ km. Also shown in Figure 5b are the surface gust front (arc-shaped solid line) and regions of strong latent heating in the troposphere (small solid contours). After *Piani et al.* [2000, Figure 4] (reprinted with permission of the American Meteorological Society.)

pled. However, one mechanism or another may serve to largely explain a set of observations with certain wind shear conditions or may serve to explain the generation of a certain class of waves, such as long vertical wavelength wave generation associated with deep convective heating.

3.1.3. Shear Generation

[45] The excitation of gravity waves by unstable shears has been studied for many years but remains one of the least quantified sources of gravity wave activity. A major challenge in these efforts was to account for the emergence of gravity waves that are propagating away from the shear layer on a timescale competitive with the most rapidly growing Kelvin-Helmholtz (KH) instability. This proved difficult for linear modes of instability for which the growth rate depends on the horizontal wave number. The subharmonic interaction, or pairing, of KH modes was proposed by *Davis and Peltier* [1979] to provide such excitation, but *Fritts* [1984b] showed pairing not to be an efficient mechanism when the exterior flow is stably stratified. By invoking a nonlinear interaction between KH and propagating modes, however, *Fritts* [1984b] and *Chimonas and Grant* [1984] were able to account for rapid excitation of propagating gravity waves. This "envelope" radiation is essentially excitation of a gravity wave by the packet-scale motions accompanying a packet of coherent KH billows evolving together in an unstable shear layer of finite horizontal extent. Two aspects of this mechanism have been assessed more recently in greater detail. *Bühler et al.* [1999] and *Bühler and McIntyre* [1999] examined wave radiation from the collapse of a mixed region due to KH billows of finite extent and concluded that this source could not be neglected in the momentum budget at greater altitudes. The mechanism envisioned by *Bühler et al.* [1999] is represented schematically in Figure 7. *Scinocca and Ford* [2000] reexamined the nonlinear radiation from spatially localized KH billows and concluded, like *Fritts* [1984b], that envelope radiation is a viable mechanism for gravity wave generation and likely an important contributor to the mesospheric momentum budget.

[46] Assuming envelope radiation to be the predominant shear excitation mechanism, we can infer characteristic horizontal scales of tens to ~100 KH wavelengths or a few to tens of kilometers, with horizontal phase speeds comparable to the mean wind at the unstable shear layer. There is, nevertheless, some observational evidence suggesting that certain linear modes of shear instability can be excited [*Mastrantonio et al.*, 1976; *Lalas and Einaudi*, 1976]. Under such circumstances the gravity wave scales are determined by the character of the shear flow and tend to be considerably larger than KH wavelengths.

3.1.4. Geostrophic Adjustment

[47] Most evidence for gravity wave excitation accompanying the restoration of balanced flow, commonly termed geostrophic adjustment, even though the balance need not be geostrophic, comes from theoretical or numerical treatments of adjustment processes. In such a process, an initial or evolving unbalanced flow relaxes to a new balanced state via both a redistribution of mean momentum, energy, and potential vorticity and a radiation of excess energy away as inertia-gravity waves. Such

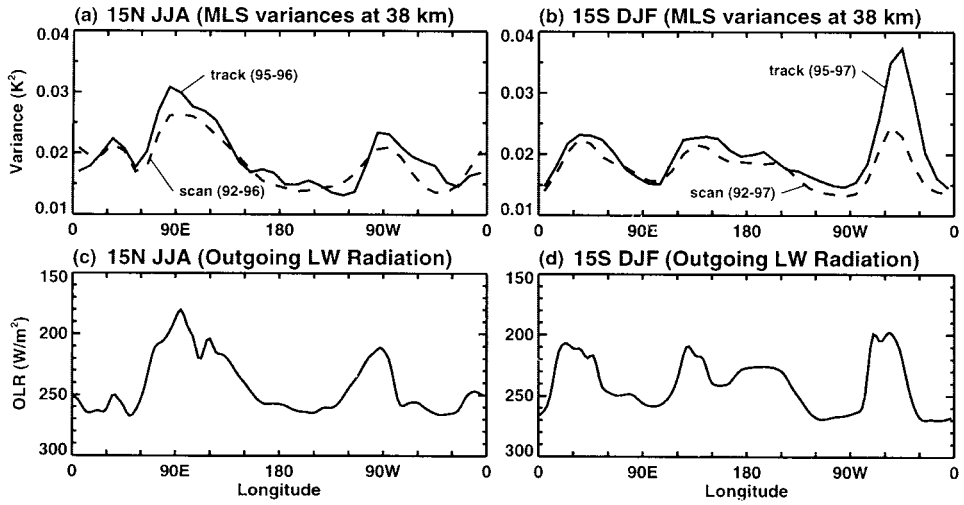


Figure 6. (a) Observations of gravity wave temperature variance as a function of longitude at 15°N from the Microwave Limb Sounder (MLS) instrument on the UARS satellite, averaged over the June-July-August season. (b) Observations of Outgoing Longwave Radiation for the same season and locations plotted on an inverse scale. The two plots show a close correspondence between gravity wave variance and infrared emission from deep, cold convective clouds. The gravity waves were also noted to have large horizontal phase speeds and very long vertical wavelengths. The close correspondence between Figures 6a and 6b provided observational evidence that long vertical wavelength gravity waves are preferentially excited by deep latent heating. After *McLandress et al.* [2000, Figure 8].

unbalanced initial flows can arise, for example, where cross-stream pressure gradients are varying along the fluid motion (requiring an acceleration or deceleration of the motion), accompanying frontal evolution and baroclinic instability, or in response to local body forces due to gravity wave dissipation and momentum flux divergence. The first of these was modeled in 2-D and 3-D by *Fritts and Luo* [1992] and *Luo and Fritts* [1993] as adjustments to unbalanced initial conditions and was found to lead to both new balanced 2-D or 3-D flows and a spectrum of inertia-gravity waves that owe their character to the geometry of the source perturbation. Observations of inertia-gravity wave activity by *Hirota and Niki* [1985] and *Thomas et al.* [1992] exhibiting downward phase progression above the jet stream and upward phase progression below the jet stream appear to provide some support for this conceptual model. Further support comes from the more recent study by *Guest et al.* [2000] who found both a high correlation of inertia-gravity wave observations with a synoptic pattern having a jet stream nearby and a tendency for reverse ray tracing to suggest the jet stream as the wave source. Our understanding of such adjustment processes is nevertheless poor at this time and will benefit from further research efforts.

[48] Frontogenesis and baroclinic instability were found by *O'Sullivan and Dunkerton* [1995], *Griffiths and Reeder* [1996], and *Reeder and Griffiths* [1996] to yield similar results, with both adjusted mean and waves scales largely defined by the character of the unbalanced flow. An example of the inertia-gravity waves emitted accompanying adjustment of an evolving baroclinic flow is shown in Figure 8 [*O'Sullivan and Dunkerton*, 1995]. As discussed above, gravity waves are seen to arise preferentially in regions of strong flow deceleration.

[49] Gravity waves arising from local body forces due to dissipating gravity waves were considered theoretically by *Zhu and Holton* [1987] and more recently by

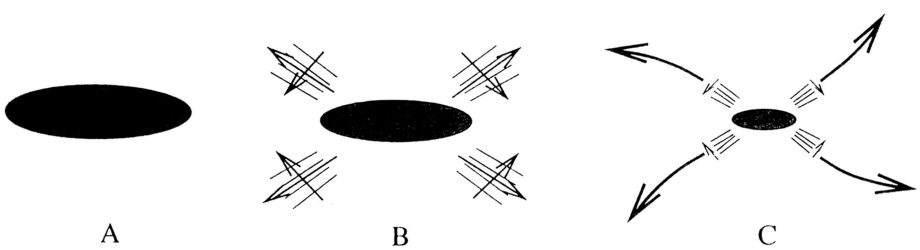


Figure 7. Schematic of the consequences of collapse of a train of KH billows which trigger radiation of gravity waves at larger scales. After *Bühler et al.* [1999] (reprinted with permission of the American Meteorological Society).

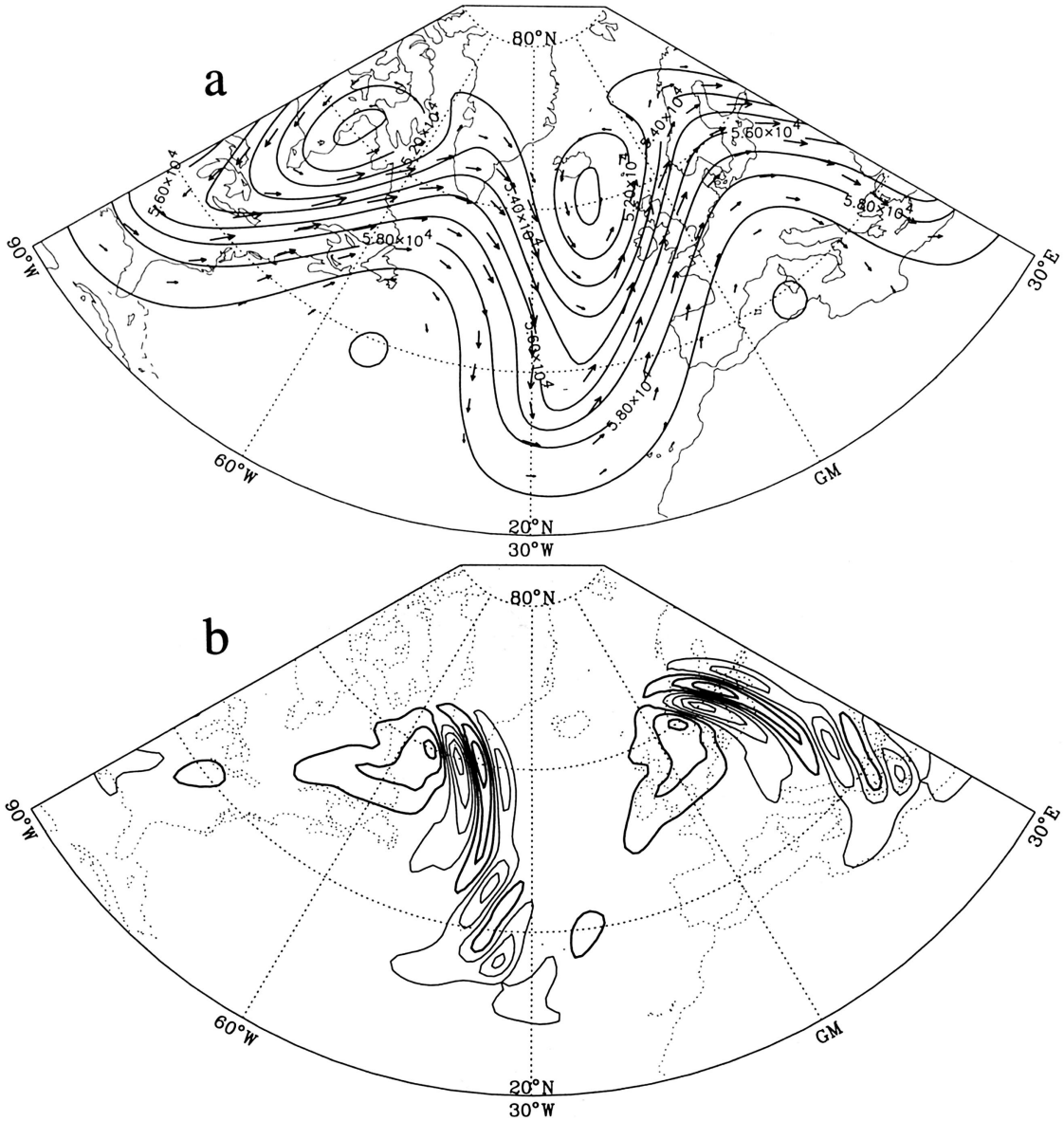


Figure 8. The (a) 500 hPA geopotential height and winds and (b) 130 hPA horizontal divergence (bottom) displaying the excitation of inertia-gravity waves by geostrophic adjustment processes. After *O'Sullivan and Dunkerton* [1995] (reprinted with permission the American Meteorological Society).

Vadas and Fritts [2001]. In all cases, the response is a combination of an adjusted mean (with redistributed momentum or potential vorticity), an altered mean thermal structure, and radiated inertia-gravity waves ensuring energy conservation in the absence of dissipation. The study by *Vadas and Fritts* [2001] considered a temporally evolving body force and showed (1) that the adjusted mean state depends on the momentum impulse and forcing geometry but not on the forcing duration and (2) that the radiated inertia-gravity waves are constrained by both the geometry and the temporal structure of the body forcing. Thus, while both long- and short-duration forcing (having the same geometry and momentum impulse) induce the same balanced mean flow, their gravity wave responses are quite different, with long-duration forcing strongly suppressing gravity waves having periods shorter than the forcing duration. The adjustment processes discussed here are also closely analogous to the radiation of gravity waves by a collapsing mixed layer due to shear instability examined by *Bühler et al.* [1999] and discussed above, though the source here is due to an addition, rather than a redistribution, of momentum.

[50] Scales of inertia-gravity waves arising from adjustment or body forcing processes are imposed by the source spatial scales and further constrained by the source timescales. Typical vertical scales are a few kilometers or more, with horizontal scales along the direction of propagation of ~10–100 times larger. Wave periods likely vary from an hour or so out to the inertial period. In cases where the source is elongated in one direction horizontally, the predominant direction of

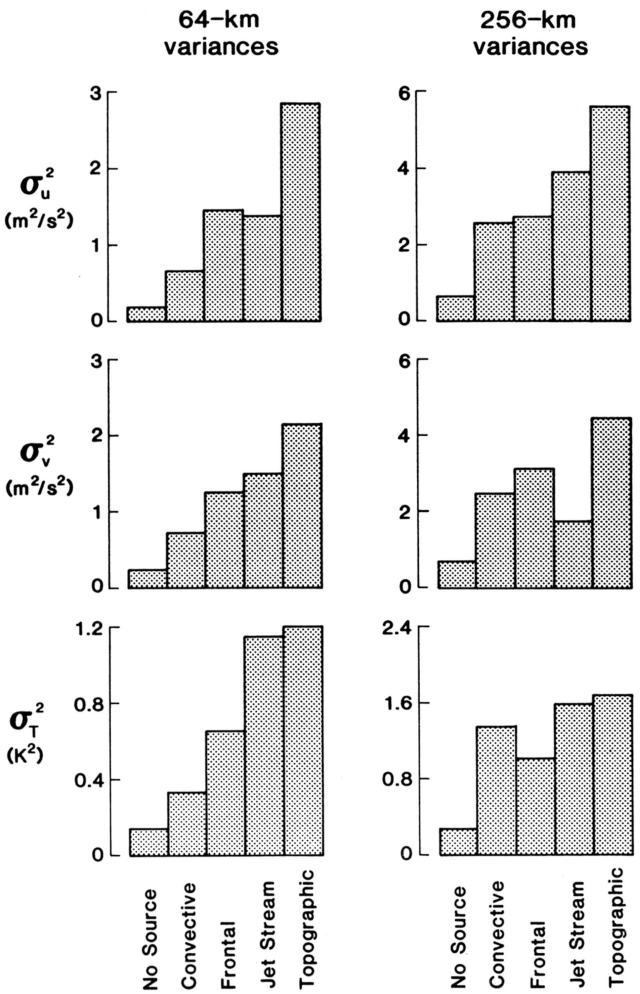


Figure 9. (top) Zonal velocity, (middle) meridional velocity, and (bottom) temperature variances in (left) 64-km and (right) 256-km data segments associated with various gravity wave sources. After *Fritts and Nastrom* [1992] (reprinted with permission of the American Meteorological Society).

propagation of the excited inertia-gravity waves is normal to the source axis.

[51] A comparison of gravity wave velocity and temperature variances associated with various sources computed using GASP data near the tropopause is shown in Figure 9 [*Fritts and Nastrom*, 1992]. These variances, for 64- and 256-km flight segments, suggest that topography, jet stream shears, and convective and frontal activity all contribute substantially to gravity wave excitation where these sources are prevalent.

3.1.5. Wave–Wave Interactions

[52] Nonlinear wave–wave interactions are arguably one of the more important mechanisms in energy exchanges, amplitude constraints, and spectral evolution for gravity waves in the middle atmosphere. Their roles and importance are not universally acknowledged, however, and there remains much to be done in quantifying their dynamics and effects.

[53] As a source of gravity waves the atmospheric community has taken guidance from oceanographers and attributed energy transfers to resonant and nonresonant interactions of various forms. There has been a long debate over the efficiency of such interactions and the validity of the various approaches in the oceanographic and atmospheric communities. Yet laboratory experiments and direct numerical simulations demonstrate that specific interactions, e.g., the parametric subharmonic instability (or PSI, see section 6.1), can be efficient and robust under a variety of circumstances [*McEwan*, 1971; *McEwan and Robinson*, 1975; *Mied*, 1976; *Dunkerton*, 1987; *Thorpe*, 1994; *Vanneste*, 1995; *Staquet and Sommeria*, 1996]. Indeed, *Klostermeyer* [1984, 1990] has argued for evidence of PSI in radar measurements of the middle atmosphere. Thus wave–wave interactions are likely an important source of waves at a range of scales not directly linked to source scales at lower altitudes. Because of their roles in spectral energy transfers and amplitude constraints, a more complete discussion of wave-wave interactions is deferred until section 6.1.

3.2. Propagation Issues

[54] Gravity wave propagation and inferences of gravity wave effects are complicated by a number of factors in the middle atmosphere. Spatially variable wind and stability profiles cause wave refraction, reflection, focusing, and ducting, while temporally variable winds alter wave phase speeds. Instability processes contribute to the generation of turbulence which may interact with or act as a source for additional wave activity. Gravity wave–vortical mode interactions may also exchange energy among these components of the motion spectrum. Our purpose in this section is to identify areas where these dynamics may confuse or obscure the inference of gravity waves and their effects in observational analyses.

3.2.1. Upward and Downward Propagation

[55] It is useful for many purposes to distinguish what fractions of the gravity wave spectrum are propagating upward and downward. Two approaches have usually been used to do this. One employs 2-D frequency–vertical wave number spectra or direct measurements of phase progression for individual waves to make the distinction, while a second employs a rotary spectral analysis of individual profiles. Each method is hampered by certain assumptions but yields insightful, if not precise, results.

[56] The 2-D spectral approach employed by *Fukao et al.* [1985], *Fritts and Cho* [1987], and *Wilson et al.* [1991a] assumes that observed and intrinsic frequencies are the same, thus confusing vertical directionality for motions having intrinsic phase speeds comparable to and opposed to mean winds and erring in the assignment of wave frequency. These are not generally serious errors when mean winds are not extreme because the total wave variance at small intrinsic phase speeds (hence small vertical scales and amplitudes) is not large. The

forms of the observed frequency spectra of horizontal velocity and temperature (slopes typically near $-5/3$) do not change dramatically for mean winds smaller than the intrinsic phase speeds of the most energetic waves [*Fritts and VanZandt*, 1987], however, while there is relatively less variance for $m \sim N/\bar{u}_h$ for which Doppler-shifting effects are larger. These and other observations suggest that the fraction of upward propagating gravity wave energy is much greater than 1/2 and that the upward propagating fraction is likely an underestimate because of uncertainties over wave frequencies [*Vincent*, 1984b; *Fukao et al.*, 1985; *Fritts and Chou*, 1987; *Wilson et al.*, 1991a; *Lintelman and Gardner*, 1994; *Gavrilov et al.*, 1996; *de la Torre et al.*, 1999].

[57] The second method, rotary spectral analysis, was first employed by *Vincent* [1984b] and is related to Stokes analysis used to assess the degree of polarization of the motion field [*Vincent and Fritts*, 1987; *Eckermann and Vincent*, 1989; *Eckermann and Hocking*, 1989; *Barat and Cot*, 1992; *Nakamura et al.*, 1993a; *de la Torre et al.*, 1999; *Zink and Vincent*, 2001a]. Rotary spectral analysis decomposes all oscillatory motions in the vertical into a superposition of upward and downward propagating motions assumed to be circularly polarized. This assumption causes a single upward propagating inertia-gravity wave with amplitude A and intrinsic frequency $\hat{\omega} > f$ (having elliptical polarization) to be misidentified as a sum of an upward propagating wave of amplitude $A_u = (1 + f/\hat{\omega})A/2$ and a downward propagating wave of amplitude $A_d = (1 - f/\hat{\omega})A/2$. As $\hat{\omega} \rightarrow f$, this yields $A_u \rightarrow 1$ and $A_d \rightarrow 0$ because the dynamics are then consistent with the assumptions. However, for $\hat{\omega} > f$ the inferred (and errant) downward propagating fraction of the total variance of $1/2 - f\hat{\omega}/(\hat{\omega}^2 + f^2)$ approaches 1/2 as $f/\hat{\omega} \rightarrow 0$ (or as $\hat{\omega} \rightarrow N$) and is entirely errant [*Eckermann and Vincent*, 1989]. Thus the rotary spectral analysis likewise leads to a conservative estimate of the upward propagating fraction of gravity wave energy. Effects of wave superposition, while not always significant, may also influence the results of a Stokes analysis and contribute to a lack of precision in partitioning of wave energy between upward and downward propagating components [*Eckermann and Hocking*, 1989]. Studies that use the technique of ellipse fitting to hodographs [*Tsuda et al.*, 1994a; *Hamilton and Vincent*, 1995; *Shimizu and Tsuda*, 1997] yield unambiguous estimates of the upward and downward propagating fractions, at least when a single dominant inertia-gravity wave is present in the profiles.

3.2.2. Vertical Propagation Versus Ducting

[58] Another issue that has arisen recently because of its implications both for estimates of gravity wave fluxes of heat and momentum and for the parameterization of gravity wave effects in large-scale models is the occurrence of gravity wave ducting. Waves that propagate vertically (with $m^2 > 0$) but having reflection (or turning) levels above and below are trapped or ducted. From the dispersion relation, equation (24) or equation (33), we see that variations in the thermal structure, as measured by the buoyancy frequency N, influence waves having any direction of propagation in the same manner, whereas the wind profile affects waves having different propagation directions very differently. In general, wind maxima in a particular direction favor ducting at that level (and in that direction) because of the relative maxima they imply in m^2, while maxima of N^2 favor ducting at this level for all propagation directions [*Chimonas and Hines*, 1986; *Fritts and Yuan*, 1989a]. Observational evidence of wind and thermal (N^2) ducts was obtained by *Isler et al.* [1997] and *Walterscheid et al.* [1999], while *Taylor et al.* [1995] and *Mitchell and Howells* [1998] inferred ducting from limited vertical wave extent. In all these cases, gravity waves having horizontal wavelengths of ~20 km or less were judged to have a high probability of being ducted by the wind and/or thermal structures of the atmosphere. The probability of ducting remains finite but decreases as horizontal wavelengths increase [*Swenson et al.*, 2000].

[59] Gravity wave structure within a duct typically exhibits very little vertical phase variation across the duct and is evanescent on either side, though more complex wave structures are also possible. Thus typical motions are largely in phase and vertical, have intrinsic frequencies near N, may have only a gradual decay of wave amplitude away from the duct, and have little or no associated momentum flux. Yet because of their structure such motions may make large contributions to vertical velocity, temperature, and airglow intensity variances.

[60] Despite our appreciation of the implications of gravity wave ducting, there have been several efforts to associate airglow intensity and/or temperature variances with gravity wave momentum fluxes (to which ducted waves contribute nothing, according to linear theory). These efforts by *Swenson et al.* [1999] and *Gardner et al.* [1999a] have been critiqued by *Fritts* [2000] and are flawed because of their neglect of ducting. *Swenson et al.* [2000] likewise recognized the lack of sensitivity to ducting behavior in earlier formulations and discussed the implications of the more complete dispersion relation, reaching similar conclusions to those of *Isler et al.* [1997] and *Walterscheid et al.* [1999] regarding susceptible horizontal wavelengths. The implications of these studies are that turning levels and ducting are important processes that impact gravity wave structure and inferred effects where wind and thermal variations are significant, particularly for shorter horizontal wavelengths.

3.2.3. Gravity Waves Versus Two-Dimensional Turbulence

[61] While quasi-2-D turbulence (hereafter 2DT, also called geostrophic or stratified turbulence under related assumptions) clearly makes contributions to atmospheric dynamics and velocity variances at some altitudes, the role it plays in the middle atmosphere appears

to be small, at least for scales well below synoptic and planetary scales (~500 km) [*Nastrom and Gage*, 1985]. *Gage* [1979] and *Lilly* [1983] have advocated 2DT as an explanation for the −5/3 power law in lower stratospheric horizontal wave number spectra, and *Lilly* [1983] has suggested that deep convection may be a viable source for 2DT. However, subsequent assessments using radar and balloon data by *Smith et al.* [1985], *Vincent and Eckermann* [1990], and *Eckermann and Vincent* [1993] suggested that only a small fraction of horizontal velocity variance could not be attributed to gravity waves. Increasing energy and momentum fluxes (per unit mass) and energy dissipation rates with altitude appear, at first sight, to imply a reduced role at greater altitudes. Without a more complete understanding of middle-atmosphere turbulence dynamics and energetics, however, it would be a mistake to assume that no strong middle-atmosphere sources of 2DT are present. This is particularly true given the potential role of gravity wave–vortical mode interactions as wave amplitudes increase [*Dong and Yeh*, 1988; *Yeh and Dong*, 1989], the departures of ratios of kinetic to potential energy from those predicted by gravity wave theories [*Tsuda et al.*, 1991; *Nastrom et al.*, 1997; *de la Torre et al.*, 1999], and the increasing scales and energies of turbulence with altitude (see section 4). We must be careful in attributing kinetic to potential energy ratios as evidence against a purely gravity wave interpretation, however, because such departures from theory can also arise from excess gravity wave energy near the inertial frequency [*Thompson*, 1978; *Sato et al.*, 1999].

4. SPECTRAL CHARACTER AND EVOLUTION

[62] Models of the gravity wave spectrum have evolved with time as we have identified spectral characteristics and constraints more precisely. As described in section 6, various theories constrain the gravity wave spectrum to behave in a particular manner over some range of wave numbers and/or frequencies. These observational and theoretical constraints have led to a canonical gravity wave spectrum that offers insights into mean properties of the gravity wave field and its variations with altitude. We emphasize, however, that this canonical spectrum cannot capture the true complexity of the gravity wave field or its evolution in altitude. There are, in addition, many reasons to expect that the spectrum will also exhibit considerable variability spatially and temporally because of various sources, filtering environments, quasi-discrete waves, and gravity wave interactions with larger scales of motion. We first review in this section the form and implications of the canonical mean spectrum. We then discuss some of the causes of departures from the canonical spectral forms.

4.1. Spectral Shape In *k, m,* and ω

[63] Frequency spectra of horizontal velocity and temperature typically vary as $E(\omega) \sim \omega^{-p}$ and exhibit slopes varying from $p \sim$ 1–2 for frequencies between f and N, with most slopes falling close to $p = 5/3$ [see, e.g., *VanZandt*, 1982; *Balsley and Carter*, 1982; *Nakamura et al.*, 1993b; *Collins et al.*, 1994]. Because horizontal wind and temperature variances tend to be concentrated near inertial frequencies, Doppler effects are small, and the intrinsic and Doppler-shifted (i.e., ground based) frequency spectra are expected to have nearly the same forms [*Scheffler and Liu*, 1986; *Fritts and VanZandt*, 1987]. This expectation is supported by the limited Lagrangian measurements that are available [*Quinn and Holzworth*, 1987; *Hertzog and Vial*, 2001]. Ground-based frequency spectra of vertical velocity, on the other hand, exhibit considerable variability with altitude, primarily because these spectra are highly sensitive to Doppler-shifting effects (see below).

[64] Vertical wave number spectra of velocities, temperatures, and various species concentrations have been measured with many instrument types. These spectra typically exhibit a distinct peak at the most energetic scales, with smaller power at lower and higher wave numbers. As noted above, a limited altitude range in observational data may result in a lack of sensitivity to motions with larger vertical scales. This causes those waves having larger vertical group velocities, thus potentially large energy and momentum fluxes, to be systematically underrepresented in such data sets. At the other extreme, resolution and measurement accuracy limits constrain the range of vertical scales that can be reliably measured. This suggests caution in assessments of characteristic scales, spectral slopes, and spectral amplitudes at large m. Nevertheless, multiple data sets suggest characteristic vertical wave numbers, m_* (i.e., the vertical wave number of the most energetic gravity waves), corresponding to vertical wavelengths of ~2–5 km in the lower stratosphere increasing to ~10–30 km near the mesopause.

[65] Spectral slope estimates at high m ($m > m_*$) also vary, but typical values range from −2.5 to −3 [*Dewan et al.*, 1984; *Tsuda et al.*, 1989, 1991]. Importantly, while there are departures from universal amplitudes and slopes (noted in section 4.2), particularly under strongly sheared conditions [*Eckermann*, 1995a], there is surprising conformity of spectral amplitudes and slopes at high m with the expectations of the various saturation theories over a broad range of altitudes [*Dewan and Good*, 1986; *Fritts and Chou*, 1987; *Smith et al.*, 1987; *Fritts et al.*, 1988a; *Sidi et al.*, 1988; *Tsuda et al.*, 1989, 1990a; *Wu and Widdel*, 1989, 1990, 1991; *Wilson et al*., 1991a; *Kuo et al.*, 1992; *Collins et al.*, 1994; *de la Torre et al.*, 1994; *Allen and Vincent*, 1995; *Mitchell et al.*, 1996]. An example of the vertical wave number spectra of radial velocities exhibiting both the decrease of m_* from the troposphere and stratosphere to the mesosphere and conformity to the spectral slopes expected for a satu-

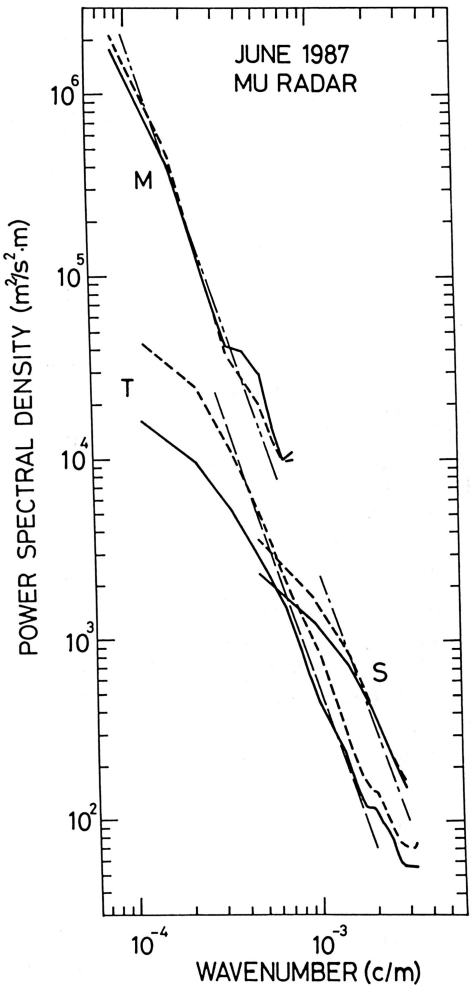


Figure 10. Radial velocity spectra assumed to represent largely horizontal velocities in the troposphere, stratosphere, and mesosphere (labeled T, S, and M) together with theoretical curves having amplitudes of $N^2/6m^3$. After *Tsuda et al.* [1989] (reprinted with permission of the American Meteorological Society).

rated gravity wave spectrum is shown in Figure 10 [*Tsuda et al.*, 1989].

[66] Saturated or nonlinear gravity waves are not the only explanation offered for the observed spectral character, however; a recurring alternative is buoyancy-range turbulence [*Weinstock*, 1985; *Sidi et al.*, 1988], which likely plays some role at the outer scales of turbulence accompanying initial instability processes. There is evidence, for example, that the ratio of kinetic to potential energy in the wave spectrum often exceeds that value predicted by the spectral model to be described later [*Tsuda et al.*, 1991; *Nastrom et al.*, 1997; *de la Torre et al.*, 1999], perhaps suggesting other 2DT turbulence or quasi-inertial components of the motion field not included in the gravity wave spectral formulation. Spectral amplitudes also exhibit, on occasion, departures from canonical values which cannot be attributed to the measurement technique. These departures and some of the reasons for them are discussed at greater length in sections 4.2 and 4.3.

[67] Horizontal wave number spectra are more challenging to measure unambiguously, as they mix spatial and temporal sampling. However, assuming that the speed of the measurement platform (typically, an aircraft) is large compared to wave horizontal phase speeds (this is not always true for the larger vertical scales), horizontal wave number spectra are obtained which are again in reasonable agreement with theory and other spectral measurements [*VanZandt*, 1982; *Fritts et al.*, 1989, 1993a; *Gardner et al.*, 1993; *Bacmeister et al.*, 1996; *Manson et al.*, 1998a]. Horizontal wavelengths range from ~10 km to thousands of kilometers; typical slopes are $\sim -5/3$, and different dynamics appear to account for the spectral slope at larger horizontal scales [*Nastrom and Gage*, 1985; *Nastrom et al.*, 1987; *Jasperson et al.*, 1990].

[68] A spectral description that captures much of the observed character of the gravity wave field is assumed (largely for convenience) to be separable in intrinsic frequency and vertical wave number and is given by

$$E(\mu, \omega, \phi) = E_0 A(\mu) B(\omega) \Phi(\phi) , \qquad (44)$$

where E is the total gravity wave energy,

$$A(\mu) = A_0 \frac{\mu^s}{1 + \mu^{s+t}}, \qquad (45)$$

$$B(\omega) = B_0 \omega^{-p} , \qquad (46)$$

$\Phi(\phi)$ expresses the azimuthal distribution of wave propagation directions, $\mu = m/m_*$, and the coefficients are chosen such that $\int_0^\infty A(\mu)d\mu = 1$, $\int_f^N B(\omega)d\omega = 1$, and $\int_0^{2\pi} \Phi(\phi)d\phi = 1$ [*Fritts and VanZandt*, 1993]. Typical values of the spectral slopes are $(s, t, p) = (1, 3, 5/3)$. The above form of the vertical wave number spectrum is more strongly peaked than was believed to be the case initially, while the requirement for a positive slope at small vertical wave numbers is imposed by the requirement for finite vertical energy fluxes (alternatively, one could impose a small-m cutoff to achieve a finite energy flux). Of these parameters, the slope at small m is perhaps the least well defined observationally or theoretically. The corresponding horizontal wave number spectrum can be derived from equations (45) and (46) and the gravity wave dispersion relation, equation (23); however, the overall spectrum in (k, ω) is then not separable. We also note that the spectrum given by equation (44) is specifically not separable in the presence of large-scale shears where vertical propagation implies differing evolutions for the m spectrum at different azimuths [*Fritts and Lu*, 1993].

[69] The gravity wave spectrum is assumed to be constrained in amplitude at $m > m_*$ by saturation processes, with all of the various saturation theories (see section 6) yielding limits for kinetic and potential energy within a factor of 2 or so of

$$E_k(m \gg m_*) \simeq \frac{pN^2}{10m^3} = \frac{N^2}{6m^3} \quad (47)$$

and

$$E_p(m \gg m_*) \simeq \frac{N^2}{10m^3}. \quad (48)$$

Variations of total wave energy density, $E(m)$, with variations in altitude and N are described approximately as [*VanZandt and Fritts*, 1989]

$$E(m) = E(\mu)d\mu/dm$$
$$= E_0 A(\mu)/m_* \sim (N/N_0)^{1/2} e^{z/H_E} A(\mu)/m_*, \quad (49)$$

where

$$m_* \sim (N/N_0)^{3/4} e^{-z/H_*} \quad (50)$$

for N increasing with height. Here $H_E \sim 2.3H$ is a value consistent with observations and linear, full-spectral propagation simulations by *Warner and McIntyre* [1996]. Corresponding values of H_* and E_0 are $H_* \simeq 2H_E$ and $E_0 = N^2/10m_*^2$. Equations (49) and (50) are also appropriate for N decreasing with height and for variable $\bar{u}(z)$, provided that the amplitude decreases implied by these effects are compensated by amplitude increases resulting from density decreasing with altitude.

[70] The corresponding energy flux by the wave field, assuming that all motions propagate upward, is

$$F_E \simeq \frac{NE_0}{18m_*}. \quad (51)$$

The factor $N/18m_* \simeq \overline{(1/m)}\bar{\omega}$ is appropriate at midlatitudes (with $f/N \simeq 1/200$) and may be interpreted as an effective vertical group velocity $\bar{c}_{gz}$ [*Fritts and VanZandt*, 1993]. Then $m_* \sim 3$ rad km^{-1} near the tropopause and $\sim 1/3$ rad km^{-1} near the mesopause, with corresponding $\bar{c}_{gz} \sim$ 0.5 and 5 m s^{-1}, respectively, and model variances are consistent with numerous radar measurements [*Vincent*, 1984a; *Fritts*, 1984a, 1989; *Balsley and Garello*, 1985; *Meek et al.*, 1985]. With $E_0 \sim 10$ and 10^3 m^2 s^{-2}, this implies $F_E \sim 5$ and 5×10^3 m^3 s^{-3}, respectively, near the tropopause and mesopause.

[71] With the above variations in E(m) and m_* and assuming, for simplicity, a constant N and that all energy flux divergence leads to turbulence, the corresponding turbulent energy dissipation rate may be written [*VanZandt and Fritts*, 1989] as

$$\epsilon \simeq \frac{NE_0}{18m_*}\left(\frac{1}{H} - \frac{3}{2H_E}\right), \quad (52)$$

with a corresponding dissipation timescale [*Fritts and Werne*, 2000] of

$$T_d \sim E_0/\epsilon \sim \frac{18m_*}{N}\left(\frac{1}{H} - \frac{3}{2H_E}\right)^{-1}. \quad (53)$$

There are, of course, uncertainties in E_0, m_*, and the fraction of wave energy that is directly dissipated as a result of wave interaction and instability processes. There are also variations that accompany changes of N, H_E, etc. with altitude. Nevertheless, these expressions provide useful guidance and/or limits for gravity influences throughout the middle atmosphere. The decreasing timescale for wave energy dissipation with increasing altitude, $T_d \sim E_0^{-1/2} \sim e^{-z/2H_E}$, is of relevance to the discussion of instability processes in section 6. The decrease with altitude implies a transition from a tendency for local dynamical (KH) instabilities in the lower stratosphere to a tendency for local convective instabilities in the mesosphere and lower thermosphere [*Fritts and Werne*, 2000] (also see section 6).

4.2. Spectral Universality, Separability, and Departures

[72] Despite the apparent role of various saturation processes in constraining wave amplitudes at large m, departures from the slopes and amplitudes predicted by the various saturation theories are common. In the troposphere, amplitudes are often larger by factors of ~ 2 to 3 [*Smith et al.*, 1987; *Tsuda et al.*, 1989, 1991; *Allen and Vincent*, 1995]. In the stratosphere, amplitudes vary from those expected from saturation theory due to wave filtering in mean shears [*Eckermann*, 1995a], enhanced saturation due to increasing N with altitude [*VanZandt and Fritts*, 1989], or other effects not yet quantified [*Cot and Barat*, 1990; *Senft and Gardner*, 1991; *Marsh et al.*, 1991; *Tsuda et al.*, 1991]. Spectral amplitudes may also exceed saturation values at any altitude (but especially in the mesosphere and lower thermosphere) where large vertical group velocities impose wave amplitude growth that is more rapid than the finite time required for instability dynamics to impose wave amplitude constraints (see section 6).

[73] Though departures from canonical frequency spectra of horizontal velocities or temperature occur in many, if not most, observations, the degree of uniformity in frequency spectra of horizontal velocity and temperature is initially surprising. The relative lack of variability of these spectra is due in part to the occurrence of the majority of the variance at (intrinsic and observed) frequencies near f and vertical wave numbers near or less than m_*. Large variances near f imply vertical scales near m_* (to enable large wave amplitudes) and corresponding large intrinsic phase speeds. This is especially the case in the mesosphere and lower thermosphere where m_* is small, vertical scales are large, and intrinsic phase speeds are often comparable to or exceed mean winds. In this case, observed frequencies cannot be Doppler shifted more than a factor of 2 or so away from their intrinsic frequency, and intrinsic and observed spectra bear a close resemblance [*Scheffler and Liu*, 1986; *Fritts and VanZandt*, 1987]. Even in the stratosphere, where vertical scales and intrinsic phase speeds are smaller, the variance is still largely dominated by

waves with large intrinsic phase speeds and low frequencies, and therefore Doppler shifting is not severe. In contrast, variances occurring at intrinsic frequencies near N are much smaller and may be Doppler spread more widely without significantly changing the shape of observed frequency spectra.

[74] The situation is quite different for frequency spectra of vertical velocity. For a frequency spectrum of horizontal gravity wave motions as described in section 4.1, we expect from equation (23) an intrinsic frequency spectrum of vertical velocities to have a slope of approximately $2 - p$ and a sharp peak (in variance content, $\hat{\omega}E_w(\hat{\omega})$) at frequencies just below N. Here large variances at high frequencies are associated primarily with (nonhydrostatic) wave motions having large k/m (see equation (30)) and vertical velocities comparable to or larger than horizontal velocities. Hence mean winds are often much larger than intrinsic phase speeds, Doppler-shifting effects may be large, and frequency spectra of vertical velocity exhibit significant variability and dependence on mean winds [*Fritts et al.*, 1990a; *Fritts and Wang*, 1991; *VanZandt et al.*, 1991]. Indeed, several studies using vertical velocity frequency spectra have shown a tendency for waves to be propagating preferentially against the local mean flow [*Fritts and Wang*, 1991; *VanZandt et al.*, 1991], in agreement with inferences from momentum flux measurements.

[75] Apart from theories which couple wave number and frequency spectra increasingly at higher altitudes [*Dewan*, 1994, 1997], there are no known reasons why frequency spectra of horizontal gravity wave motions should be universal. Indeed, specific sources impart to the wave field their own spectral attributes (see section 3). The tendency toward a canonical frequency spectrum as altitude increases more likely reflects a combination of (1) averaging over many wave sources and wave field realizations, (2) the variations of intrinsic frequency accompanying wave propagation in wind shear, and (3) systematic nonlinear interactions among spectral components which increase in importance with altitude.

[76] There is no reason to suppose that the gravity wave spectrum is separable among any two variables on theoretical grounds. Nor is separability fundamental to any saturation theory, despite claims to the contrary [*Gardner et al.*, 1999b]. Separability simply provides a convenient means of describing the dependence of gravity wave spectra on observed frequency, vertical wave number, and/or azimuth of propagation having some experimental justification [*Smith et al.*, 1987; *Fritts and Chou*, 1987; *Fritts and VanZandt*, 1987; *VanZandt and Fritts*, 1989; *Sidi et al.*, 1988].

[77] Finally, we note that estimates of gravity wave spectral parameters depend to some extent on the technique employed, as each technique has its own measurement biases and limitations. Thus care must be taken in interpreting and, in particular, intercomparing these estimates. Limited altitude coverage, limited temporal extent, spatial and temporal averaging, and windowing and linear trend removal additionally cause underestimates of spectral variance at lower wave numbers and frequencies [*Lefrere and Sidi*, 1990]. Geophysical variations of the mean wind and stability profiles with altitude likewise introduce biases in spectral amplitudes and slopes [*VanZandt and Fritts*, 1989; *Eckermann*, 1990, 1995a].

4.3. Broad Spectrum or Discrete Waves?

[78] Spectral characterization of the gravity wave field almost always gives the impression that the wave field is composed of many components at various scales. This is true even for spectra formed from a single time series or vertical profile [*Dewan*, 1979; *Reid et al.*, 1988; *Fritts et al.*, 1988a; *Wu and Widdel*, 1989, 1991; *Allen and Vincent*, 1995; *Sato and Yamada*, 1994]. There is ample evidence, however, that spectra are often composed of only a few waves of significant amplitude, lending support to the arguments advanced by *Hines* [2002a] (see section 6.1). Most vertical profiles of velocity or temperature exhibit in any altitude range a single dominant wave structure; the corresponding vertical wave number spectrum typically has ~70% of the variance within a factor of 2 of the characteristic vertical wave number, m_* (see equation (49)).

[79] Additional evidence for a discrete or "narrow spectral," rather than a "broad spectral" description of the local wave field comes from many studies. Airglow observations more often than not exhibit a single or very few wave structures extending over a few to many wavelengths [*Taylor et al.*, 1991, 1995, 1998; *Hecht et al.*, 1997; *Nakamura et al.*, 1998; *Swenson et al.*, 1999; *Yamada et al.*, 2001].

[80] Airglow and noctilucent cloud (NLC) observations have also yielded evidence of gravity wave instability processes accompanying single or a few large-amplitude wave motions at specific spatial scales [*Fritts et al.*, 1993b; *Swenson and Mende*, 1994; *Hecht et al.*, 1997; *Yamada et al.*, 2001]. Balloon, lidar, and rocket data provide evidence of large-scale overturning of apparently individual waves, or superpositions of a few large-amplitude waves [*Shutts et al.*, 1988; *Hecht et al.*, 1997; *Goldberg et al.*, 1997; *Williams et al.*, 2002]. Radar and optical estimates of gravity wave momentum fluxes provide enticing evidence of discrete, large-amplitude events [*Riggin et al.*, 1997; *Fritts et al.*, 2002] and suggest that such events may comprise a significant part of the mean forcing. Finally, *Sato and Yamada* [1994] found that a single wave approaching a critical level can yield a vertical wave number spectrum having a -3 slope, while *Eckermann* [1999] showed that such a spectrum can arise from as few as four discrete sets of wave scales. Thus we must be cautious of claims that the spectral character of the gravity wave field implies a broad mix of waves occupying a common volume, as this often appears not to be the case.

5. CLIMATOLOGIES OF GRAVITY WAVE ACTIVITY

[81] Because global model studies are increasingly recognizing the need to describe gravity wave effects via parameterization, efforts to observe and understand the climatology of gravity waves is an area of active research. Climatology in this context refers to the geographical and temporal variations in gravity wave activity (specifically momentum fluxes) and variations in gravity wave characteristics (wavelengths, phase speeds, etc.). If the mechanisms for gravity wave generation were completely understood and if their sources were fully characterized, we could in theory perfectly describe the temporal and geographical variations in gravity wave occurrence along with the specific wave characteristics for a given set of meteorological conditions. The largest obstacle to accurate parameterization of gravity wave effects in the middle atmosphere would then be overcome. Unfortunately, we instead have a disjointed set of observations that sometimes agree with one another and other times do not. We also have idealized model studies of gravity wave sources which are often poorly constrained by observations. Where such constraints do exist, models and observations do not always agree. Progress is being made, but the global characterization of the climatology of gravity wave occurrence remains a daunting observational challenge.

[82] In the last 15 years there have been a number of compilations of observations of small-scale temperature and horizontal wind structure in the middle atmosphere, with these structures interpreted as gravity wave perturbations. These compilations include long-term observations at single sites from which seasonal and interannual variations can be observed, global observations over some more limited period of time, and regional observations showing both temporal and geographical variations. The reasonably constant appearance of vertical wavelength and ground-based frequency spectra described in section 4 can be deceptive because of the logarithmic axes plotted. In fact, the gravity wave climatology information that has appeared in the literature shows substantial variations with time, with height, as well as with geography.

[83] Part of the difficulty in inferring a climatology of gravity waves from the existing set of observations is related to the fact that gravity waves are phenomena that occur with a broad and multidimensional spectrum of characteristics. Each observation technique tends to be sensitive only to some portion of that spectrum while being almost totally insensitive to other portions. These observational limitations can themselves be responsible for the climatological patterns observed in the data because the vertical wavelengths, intrinsic frequencies, and phase speeds of gravity waves can be dramatically altered by variations in the background atmosphere [*Alexander*, 1998]. These variations in gravity waves due to background atmosphere effects must then be separated from variations associated with gravity wave sources before we can achieve accurate interpretations of gravity wave climatological observations.

[84] Our understanding of the nature of gravity wave sources was summarized in section 3.1. In this section we summarize global patterns and seasonal variations of gravity wave variances and fluxes that have been reported using different observational techniques. We then review the theoretical studies that have helped to explain some of these observed variations and separate background atmosphere effects from source variations. Inferences thus gained about global variations in gravity wave sources are then summarized.

5.1. Observations

5.1.1. Remote Sensing Measurements From Space

[85] Observations from satellite platforms hold great promise for providing the kind of global coverage at frequent time intervals that is needed to understand gravity wave variability. However, gravity waves can have quite small vertical and horizontal scales and short periods. They can therefore be challenging to observe remotely from space. New techniques have emerged, yet each has been capable of observing only some fraction of the gravity wave spectrum that may be present in the atmosphere, while missing other portions of the spectrum.

[86] Infrared limb-viewing profiles like the Limb Infrared Monitor of the Stratosphere (LIMS) instrument [*Gille and Russell*, 1984] retrieve temperature profiles from 15 to 60 km at high vertical resolution (~1.5 km). However, line-of-sight (LOS) integration effects and the spacing between profiles limit these data to observation of only longer horizontal wavelengths. *Fetzer and Gille* [1994] retrieved global fields of gravity wave temperature perturbations by subtracting Kalman-filtered temperature fields from unfiltered temperature profiles. The Kalman-filtered data included planetary waves with zonal wave number ≤ 6. The difference field, assumed to be gravity wave perturbations, included information on waves with vertical wavelengths of ~6–50 km and horizontal wavelengths longer than ~200 km. (Note that aliasing from shorter unresolved horizontal wavelengths was evident, but these represented a small fraction of the total gravity wave variance observed [*Fetzer and Gille*, 1994].) *Fetzer and Gille* [1994] noted that the temperature variances derived from these data are likely dominated by low-frequency inertia-gravity waves. Maps of the LIMS gravity wave temperature variance (Figure 11a) as a function of latitude and height showed a peak near the equator in the lower stratosphere below 20 km and a peak in the winter stratospheric jet above 20 km. Seasonal variation above 40 km at the equator related to the semiannual oscillation (SAO) was also evident in the gravity wave variance.

[87] Maps of gravity wave temperature variance have been similarly derived from Cryogenic Infrared Spectrometers and Telescopes for the Atmosphere

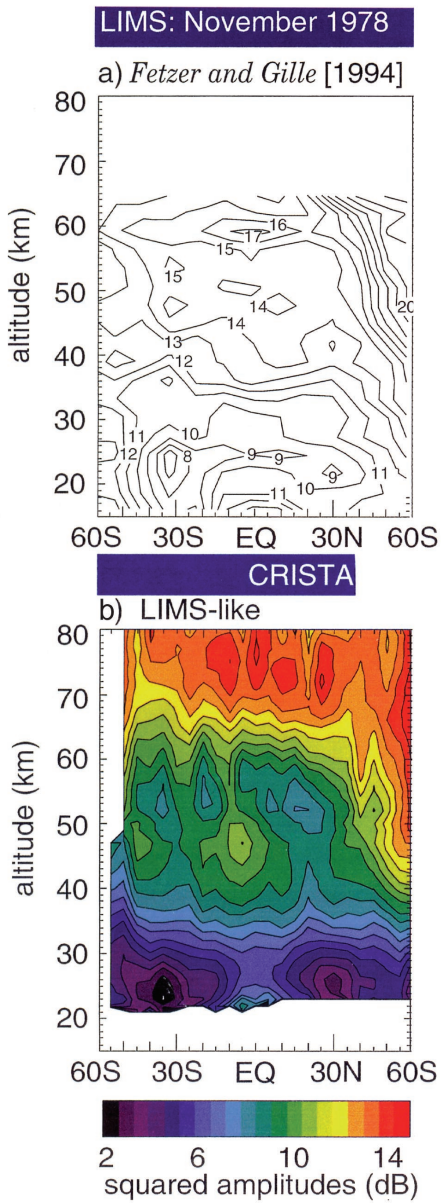


Figure 11. Comparison of zonal mean gravity wave temperature variance in November from (a) LIMS (after *Fetzer and Gille* [1994], reprinted with permission of the American Meteorological Society) and (b) CRISTA filtered to reflect the same vertical resolution as LIMS (after *Preusse et al.* [2000]).

(CRISTA) data [*Preusse et al.*, 1999]. These data give temperature profiles at finer horizontal resolution and slightly better vertical resolution than LIMS. CRISTA's altitude range varies with observational mode but can extend from ~20 to 80 km. The CRISTA data were used to infer that topography was the source for waves observed in the stratosphere from space [*Eckermann and Preusse*, 1999; *Preusse et al.*, 2002]. Global maps of the gravity wave temperature variance show similar features as those from LIMS (Figure 11b).

[88] *Tsuda et al.* [2000] derived global maps of gravity wave potential energy derived from GPS occultation data temperature profiles [*Rocken et al.*, 1997]. These observations include gravity waves with vertical wavelengths of ~2–10 km in the 15- to 40-km altitude range. The LOS integration path length is also ~200 km, so the observed wave characteristics are similar to LIMS and CRISTA. The gravity wave energy shows a similar equatorial maximum below ~25 km and midlatitude winter maxima at higher altitudes. At lower latitudes (≤30°), the gravity wave energy was particularly enhanced over regions of active convection. At midlatitudes, gravity wave energy was observed to be larger over continents than over oceans.

[89] Observations from the Microwave Limb Sounder (MLS) have been used to derive global maps of gravity wave temperature variance in the stratosphere and mesosphere [*Wu and Waters*, 1996a, 1996b]. These data include only gravity waves with long vertical wavelengths $\lambda_z > 12$ km because of the deeper weighting functions associated with microwave profilers compared to infrared profilers. One observation mode (the scanning mode) reportedly included gravity waves with horizontal wavelengths of ~60–150 km [*Wu and Waters*, 1996a], while another (the tracking mode) included gravity waves with horizontal wavelengths up to 1000 km. *McLandress et al.* [2000] noted that the integrated variance observed by the two modes is similar, suggesting that aliasing effects may be extending the range of wavelengths contributing to the variance in data from the scanning mode. *Wu and Waters* [1997] note that most of the observed waves propagate opposite to the direction of the wind. This, together with the long vertical wavelength restrictions, suggests these waves have much higher intrinsic frequencies than those observed via the other space-based observations described above. The global maps show peaks in variance at middle to high latitudes in winter, with a minor peak at subtropical latitudes in summer. Variances at the equator were indistinguishable from noise [*McLandress et al.*, 2000]. Longitudinal variations suggest deep convection as a source of the waves observed in the summertime and topography as the source for waves at southern midlatitudes in winter [*McLandress et al.*, 2000] (see also Figure 6).

5.1.2. Radiosonde Profiles

[90] Observations of temperature and horizontal wind profiles from radiosondes have provided information on short vertical wavelength gravity waves in the lower stratosphere below ~25 km. Several studies have inferred latitudinal, seasonal, and interannual variations in gravity wave activity in these data. The frequency of radiosonde launches rarely exceeds twice per day, so gravity wave perturbations are generally obtained by subtracting a low-order polynomial fit to the profiles and assuming the perturbations as a function of height are caused by gravity waves. In studies near the equator, however, where the inertial period is very long, radiosonde data have also been analyzed in the time domain

for information on 1- to 3-day period oscillations as inertia-gravity waves [*Sato et al.*, 1994; *Sato and Dunkerton*, 1997].

[91] Radiosonde profiles of horizontal wind plotted as hodographs generally show rotation indicative of upward propagating inertia-gravity waves [*Ogino et al.*, 1995; *Vincent and Alexander*, 2000]. In fact, there is abundant evidence from analyses of radiosonde horizontal wind and temperature profiles that the perturbations in these vertical profiles are dominated by very low intrinsic frequency inertia-gravity waves, specifically those with frequencies between ~1 and 3*f* [*Vincent et al.*, 1997; *Vincent and Alexander*, 2000; *Zink and Vincent*, 2001a].

[92] Seasonal variations at northern midlatitudes and high latitudes show a maximum in gravity wave activity in winter and minimum in summer [*Yoshiki and Sato*, 2000; *Kitamura and Hirota*, 1989]. At very high northern latitudes, significant gravity wave variances were observed only in the absence of mountain wave critical levels [*Bacmeister et al.*, 1990b; *Worthington and Thomas*, 1996; *Whiteway and Duck*, 1996; *Whiteway et al.*, 1997]. At midlatitudes in the Southern Hemisphere, observations also show peak wave activity in winter [*Allen and Vincent*, 1995; *Zink and Vincent*, 2001a]. At Antarctic latitudes a maximum in spring has instead been observed [*Allen and Vincent*, 1995; *Yoshiki and Sato*, 2000].

[93] At the tropical latitudes of northern Australia, Indonesia, and the Indian Ocean a seasonal cycle in wave activity has been observed with a maximum in December to February [*Allen and Vincent*, 1995; *Vincent and Alexander*, 2000]. This is a period of intense deep convection in that part of the world. Waves observed in radiosonde data from that area also often tend to have long horizontal wavelengths, ~1000 km [*Tsuda et al.*, 1994a; *Shimizu and Tsuda*, 1997; *Vincent and Alexander*, 2000]. Propagation directions of gravity waves observed in radiosonde data at low latitudes are influenced by the winds of the quasibiennial oscillation (QBO). An excess of eastward propagating waves has been observed most prominently during periods with strong westward winds in the lower stratosphere and eastward shear aloft [*Murayama et al.*, 1994; *Tsuda et al.*, 1994a; *Shimizu and Tsuda*, 1997; *Sato and Dunkerton*, 1997; *Vincent and Alexander*, 2000]. Cadet and Teitelbaum [1979] observed the analogous situation in the opposite phase of the QBO: westward propagating gravity waves in westward QBO wind shear.

[94] Latitudinal variations are also apparent in gravity waves observed in radiosonde profiles. The data show a trend toward increasing wave energy equatorward of ~30° [*Kitamura and Hirota*, 1989; *Allen and Vincent*, 1995; *Ogino et al.*, 1995]. Note that a similar equatorial maximum was observed in the GPS and IR limb sounding satellite observations in the lower stratosphere described above.

5.1.3. Rocket Soundings

[95] Rocket soundings of temperature and horizontal wind have been used to infer seasonal and latitudinal variations in gravity wave activity in the stratosphere below ~60 km at ~1-km resolution. Rocket profiles from launches in the 1970s and 1980s at sites in the Northern Hemisphere and in the tropics have provided information on the seasonal and latitudinal variations in gravity wave activity in the stratosphere [*Hirota*, 1984; *Hirota and Niki*, 1985; *Hamilton*, 1991; *Eckermann et al.*, 1995].

[96] The analysis method has varied between different studies, but the most common procedure is to subtract a background profile from the data and to assume that smaller-scale perturbations are gravity waves. In general, the rocket sounding analyses have shown a seasonal cycle in wave activity at high latitudes (~40°–80°N) with a maximum in winter, similar to results of radiosonde analyses. Hodograph or rotary spectral analyses of rocket profiles have also shown a dominance of upward propagation, again similar to the radiosondes. At low latitudes, different analyses gave different seasonal cycles in wave activity, suggesting a sensitivity of this result to the vertical scales of waves considered. Mean kinetic and potential energies in the 20- to 40-km altitude range both showed increases at latitudes <30° [*Eckermann et al.*, 1995]. A preference for eastward propagation at sites within 10° of the equator has been noted as a possible indication that the wave activity at these sites is dominated by equatorial Kelvin waves [*Hamilton*, 1991; *Eckermann*, 1995b; *Holton et al.*, 2001].

5.1.4. Lidar

[97] Lidar observations of middle atmosphere gravity waves have been restricted to a few specific sites. Among these, only a few have operated for long enough periods to provide information on seasonal variations in gravity wave activity [*Wilson et al.*, 1991b; *Marsh et al.*, 1991; *Mitchell et al.*, 1991; *Whiteway and Carswell*, 1995]. These are all Rayleigh lidar studies of gravity wave temperature or density perturbations at midlatitude sites. The data include gravity waves with vertical wavelengths as short as 1 km at altitudes in the upper stratosphere and lower mesosphere. The data have been analyzed both for vertical variations as well as in the time domain, though these observations are limited to nighttime only. The dominant pattern that has appeared in monthly mean averages of these data is a seasonal cycle with maximum in winter and minimum in summer: the same annual cycle that is observed in radiosonde and rocket sounding data at midlatitudes.

5.1.5. Radar

[98] Radar observations of gravity wave wind perturbations can be made in the lower stratosphere and in the mesosphere and lower thermosphere. Many of these observations have been made at a limited number of sites and on a short-term campaign basis. We focus here

on studies that have compiled observations over long enough time periods to study seasonal variations. These long-term studies have, to date, all been at midlatitude sites.

[99] Analyses of gravity wave observations in the mesosphere in the ~65- to 90-km region are generally reported for different ground-relative frequency bands for waves with periods ranging from 5 min to 24 hours. Gravity wave kinetic energy as well as variances ($\overline{u'^2}$, $\overline{v'^2}$, $\overline{w'^2}$) in each of the three component wind directions have shown a semiannual variation at these altitudes with maxima in summer and winter and minima in spring and fall [*Meek et al.*, 1985; *Vincent and Fritts*, 1987; *Tsuda et al.*, 1990a; *Manson and Meek*, 1993; *Nakamura et al.*, 1996]. These analyses included observations at both Northern and Southern Hemisphere sites, and the results show similar gravity wave activity in the two hemispheres. A seasonal asymmetry in the occurrence of polar mesosphere summer echos (PMSE) and the ice particles that cause them has been observed. Specifically, there is an absence of PMSE in the Southern Hemisphere that may be a result of warmer temperatures in the south due to a weaker mean meridional circulation and corresponding weaker gravity wave forcing [*Balsley et al.*, 1995; *Woodman et al.*, 1999]. Some mesopause temperature measurements support this idea [*Huaman and Balsley*, 1999] while others suggest similar thermal structures at high latitudes in each hemisphere [*Lübken et al.*, 1999]. A recent chemical-dynamical model study [*Siskind et al.*, 2003] supports the idea of hemispheric asymmetries in gravity wave forcing and an asymmetric narrow warm layer at the mesopause that may be difficult to resolve with some measurement techniques.

[100] Observations of horizontal and vertical wind covariance $\overline{u'w'}$ (momentum flux per unit density) in the mesosphere have shown an annual cycle, positive (i.e., eastward) in summer and negative (i.e., westward) in winter, so the flux tends to have the opposite sign as the winds at these altitudes [*Tsuda et al.*, 1990b; *Manson and Meek*, 1993; *Nakamura et al.*, 1993c, 1996]. The anticorrelation between flux direction and the background wind direction is a condition observed even more generally in shorter-term data records [*Vincent and Reid*, 1983; *Reid and Vincent*, 1987; *Fritts and Vincent*, 1987; *Reid et al.*, 1988; *Fritts and Yuan*, 1989b; *Meyer et al.*, 1989; *Wang and Fritts*, 1990; *Nakamura et al.*, 1993c]. The peak magnitudes of monthly mean fluxes (per unit mass) at mesospheric heights have been reported at ~1–4 $m^2 s^{-2}$. Reports of meridional fluxes $\overline{v'w'}$ have been variable, so no repeatable seasonal cycle is known, but reported magnitudes are often comparable to or slightly larger than the zonal fluxes.

[101] Reports of seasonal changes in the anisotropy of wave propagation directions have varied among different studies [*Manson and Meek*, 1988; *Nakamura et al.*, 1993d]. *Maekawa et al.* [1987] and *Fan et al.* [1991] inferred critical-level wind filtering as the mechanism responsible for anisotropies observed in the summer mesosphere wind field using the SOUSY and Arecibo VHF radars, respectively. Inferred propagation directions also depend greatly on the characteristics of gravity waves being observed. *Vincent and Fritts* [1987] and *Nakamura et al.* [1993b] observed a preference for meridional propagation of the dominant wave motions under winter conditions. *Nakamura et al.* [1993b] noted a tendency for propagation of the dominant motions toward either the NE or SW under summer conditions. *Taylor et al.* [1998] and *Nakamura et al.* [1998] observed propagation preferentially toward the NE for waves of shorter horizontal scales under summer conditions. Finally, *Nakamura et al.* [1993a] observed similar character in propagation directions at Adelaide and the MU radar at Kyoto in summer but with more isotropic propagation at Adelaide and more zonal propagation at Kyoto in winter.

[102] Observations in the stratosphere of the seasonal cycle of gravity wave activity measured by radar have been reported only for the MU radar near Kyoto in Japan for altitudes below 24 km [*Tsuda et al.*, 1994b; *Murayama et al.*, 1994; *Sato*, 1994]. These have shown an annual cycle in gravity wave variances ($\overline{u'^2}$, $\overline{v'^2}$, $\overline{w'^2}$) in all three wind components with minima in summer and maxima in winter to early spring. *Sato* [1994] and *Murayama et al.* [1994] also reported zonal and meridional momentum fluxes. The zonal component showed a clear annual cycle with large negative values ~ -0.1 $m^2 s^{-2}$ in winter and near zero in summer. The meridional component also tended to be negative but with smaller magnitudes and no clear annual cycle. The observed wave characteristics at this site have been generally consistent with topographic gravity wave generation. Specific campaigns have, however, also observed waves associated with convection (see section 3.1).

5.1.6. Aircraft

[103] Measurements of wind and temperature fluctuations by aircraft flying in the stratosphere have yielded information on the horizontal wavelength spectrum of gravity waves [*Nastrom et al.*, 1987; *Bacmeister et al.*, 1996] (see section 4.1). Some aircraft studies have related gravity wave characteristics in the stratosphere to the wave sources below, including mountains [*Gary*, 1989; *Bacmeister et al.*, 1990a, 1990b; *Nastrom and Fritts*, 1992; *Leutbecher and Volkert*, 2000] and convection and frontal systems [*Fritts and Nastrom*, 1992; *Pfister et al.*, 1993a, 1993b; *Alexander and Pfister*, 1995; *Alexander et al.*, 2000]. Several of these studies have provided in situ determination of gravity wave momentum fluxes. However, aircraft observations have been too limited in their duration to provide much information about the gravity wave climatology directly. Instead, they have contributed to our understanding of gravity wave variability by providing important quantitative constraints for theoretical and model studies of gravity wave sources (see section 3.1 and Figure 9).

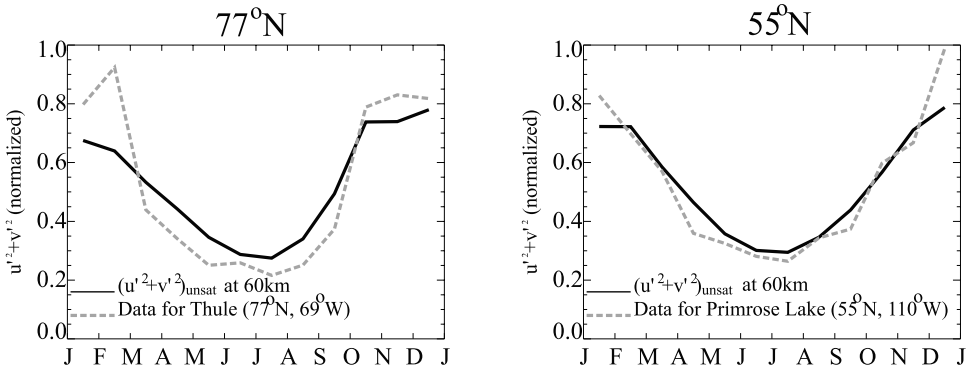


Figure 12. Seasonal variations observed in gravity wave horizontal wind variance in rocket soundings (dashed lines) and theoretical variations given by (54). The dotted line represents the seasonal variation that would be predicted if the wave amplitudes were saturated. After *Eckermann* [1995b] (reprinted with permission of Elsevier Science).

5.2. Theoretical Studies of the Global Patterns Observed in Gravity Waves

[104] At middle to high latitudes throughout the stratosphere the long-term measurements of gravity wave variances cited have shown an annual cycle with maximum in winter and minimum in summer. *Eckermann* [1995b] described a simple linear theoretical model that explained the annual cycle in both lidar and rocket sounding observations without any seasonal variation in gravity wave sources. This linear theory describes the change in amplitude of a gravity wave propagating vertically without dissipation from a source level z_0 where amplitudes are presumed constant to the observation level z. The theory predicts a simple proportionality for horizontal wind variance,

$$\overline{u'^2} \propto (N/N_0)(\rho_0/\rho)\,, \tag{54}$$

and for temperature variance,

$$\overline{(T'/\bar{T})^2} \propto (N^3/N_0)(\rho_0/\rho)\,, \tag{55}$$

that fit the observed annual cycles in variance quite well (Figure 12) despite the fact that shear effects are neglected. The theory yields larger variances in winter because the colder winter stratosphere (with correspondingly smaller-scale height and faster decrease in density) leads to faster amplitude growth with height. Note that the factor depending on stability N is more important for explaining temperature variance ($\propto N^3$) than wind variance ($\propto N$). Equations (54) and (55) may also be important for explaining the seasonal variations seen in radar and GPS observations. Note that in equations (54) and (55), z is the geometric altitude. For observations reported on pressure altitude levels $z_p = H\ln(p_0/p)$, such as radiosonde data, the density factor will be unimportant, and the effect greatly reduced.

[105] Seasonal and latitudinal variations observed in MLS gravity wave temperature variance were largely explained without variations in gravity wave sources by *Alexander* [1998]. Background wind effects were shown to have a controlling influence on the patterns in MLS gravity wave activity. *Alexander* [1998] showed how the MLS observation method essentially applies a low-pass filter to the vertical wave number spectrum of waves in the atmosphere that tends to allow the MLS to observe waves only at those places in the atmosphere where the background winds are very strong. In the stratosphere, these places are in the winter polar jet and in the summer subtropical jet. *McLandress et al.* [2000] further found that the waves seen by MLS in the winter hemisphere jet have, on average, smaller intrinsic phase speeds ($\sim$20 m s^{-1} or less), while those in the summer hemisphere subtropics tend to have larger intrinsic phase speeds ($\sim$50 m s^{-1} or less). This result is consistent with topographic sources being more important at high latitudes and deep convective sources being more important at low latitudes. *Alexander* [1998] also showed that seasonal changes in background winds can give rise to the late summer peak in variance that was observed in rocket sounding observations at low latitudes [*Eckermann et al.*, 1995] without any seasonal changes in sources.

[106] The December to February peak in wave activity observed in low-latitude radiosonde observations [*Vincent and Alexander*, 2000] was explained without any pronounced seasonal variation in wave sources in a companion model study [*Alexander and Vincent*, 2000]. However, the spectrum of waves input into their model required certain characteristics. The waves input needed to have slow phase speeds, < 5 m s^{-1} relative to the ground, and the source level had to be in the upper troposphere or at the tropopause. This is consistent with the spectrum of waves that would be generated if the source were slow-moving deep convective complexes that generated waves via the obstacle effect near the tropopause analogous to topographic wave forcing [*Clark et al.*, 1986; *Pfister et al.*, 1993b]. There is also a weak annual cycle in cloud top height near Cocos Island [*Vincent and Alexander*, 2000] that may imply some contribution to the gravity wave annual cycle due to source variations.

[107] GPS, CRISTA, radiosonde, and rocket sounding data have all shown a latitudinal variation in gravity wave energy in the lower stratosphere with increasing variance equatorward of ~30°. *Alexander et al.* [2002] proposed this peak is associated with the latitudinal variation in the Coriolis parameter f which defines the lowest gravity wave intrinsic frequency as a function of latitude. Each of these four observational techniques is sensitive to these low-frequency waves. The linear theory they applied to explain the peak accounts for the dependence of gravity wave vertical group velocity on intrinsic frequency (see equation (39)). If there is intermittency in the forcing of waves in the equatorial region (such as would be expected of waves generated by convection), then the slow vertical group velocities of the low-frequency waves make them the most probable waves to be observed. The theory then accounts for the nearly ubiquitous presence of low intrinsic frequency waves at low latitudes and also accounts for the low-latitude peak seen in long-term averages of gravity wave energy cited above. The results of *Alexander et al.* [2002] further imply that although higher-frequency waves appear relatively unimportant in the monthly mean zonally averaged gravity wave energy, they may dominate the gravity wave momentum flux spectrum.

[108] Seasonal variations in zonal momentum flux $\overline{u'w'}$ observed by radar in the midlatitude mesosphere, showing the mean flux opposes the background wind (positive in summer and negative in winter), are qualitatively consistent with our understanding of how the background winds would filter the gravity wave spectrum [e.g., *Maekawa et al.*, 1987; *Fan et al.*, 1991]: Summer westward winds would prevent most of the westward propagating waves from reaching the mesosphere, leaving primarily eastward propagating waves with positive momentum flux, and the reverse would be true in winter. Fluxes were observed to be small at equinoxes, but this is probably not due to averaging of large negative and positive fluxes together. Rather, the small equinoctal fluxes and kinetic energies, taken together, suggest that wave amplitudes are simply smaller at equinoxes. The equinoctal minima may be directly associated with the weak mesospheric winds at these times rather than any seasonal variation in sources. In weak wind conditions, waves are not Doppler shifted to high intrinsic phase speeds, and this results in smaller-amplitude limits inferred from linear instability theory (see section 6).

5.3. Summary: Global Characteristics of Gravity Waves and Their Sources and Variability

5.3.1. Sources

[109] There is convincing evidence for mountain waves in the stratosphere with fair quantitative agreement between high-resolution model studies and observations. Parameterized mountain waves, however, have a large effect in global models, and improved constraints on their properties and intermittency in occurrence are still needed. There is convincing evidence for convectively generated gravity waves in the stratosphere with qualitatively good agreement between observations and models. Observations that include both detailed information about the storm characteristics and about the waves that are generated are hard to find but are needed to make our understanding of convectively generated gravity waves more quantitative. There is evidence for low-frequency waves emanating from frontal zones in the jet stream and high-resolution model studies describing inertia-gravity wave generation under these conditions. Observational evidence detailed enough for a quantitative understanding of this wave source is lacking. A detailed understanding of the generation mechanisms and the resulting wave characteristics is needed for accurate description of gravity wave effects in global models.

5.3.2. Variability

[110] Observations of short vertical wavelength gravity waves in the lower stratosphere show increases in variance at latitudes $<30°$. This has been interpreted as evidence for very low intrinsic frequency waves at low latitudes that are excluded from higher latitudes by the variation in f and that have intermittent sources in time, such as would be expected from convection. Very close to the equator, at latitudes $<10°$, these measurements could also include contributions from short vertical wavelength Kelvin waves. Also in the tropics, there is evidence for geographical variations in long vertical wavelength gravity waves with high intrinsic phase speeds that are closely tied to deep convection. These observations support model studies that show how deep heat sources generate long vertical wavelength waves.

[111] The evidence for global variations in gravity wave sources from climatologies of gravity wave activity is currently weak because many of the most prominent variations observed can be explained by background atmosphere effects without any seasonal or geographical variations in sources. Yet we have clear evidence for the importance of topography, convection, and frontal generation of gravity waves, and these sources must clearly vary seasonally and geographically. The existing gravity wave climatologies that describe monthly mean gravity wave kinetic and potential energy have largely served their purpose. They have allowed us to test the linear theory of gravity waves and their interaction with the background atmosphere.

[112] Observational gravity wave studies in the future must begin to define the deviations in gravity wave activity from these climatological patterns. Observations and modeling tools must be used together to separate these variations due to the background atmosphere and observational biases from variations due to sources. Quantitative measures of the intermittency in wave activity are needed along with any monthly mean quantity since it is important to define whether a monthly mean

value represents intermittent large bursts of wave activity or more constant, smaller-amplitude wave activity.

[113] Satellite observations hold the most promise for providing the kind of geographical and temporal coverage needed to understand gravity wave variability. However, inherent observational limitations mean that continued in situ measurements from aircraft and high-resolution ground-based profiling instruments will still be needed to understand the full spectrum of gravity waves present in the atmosphere.

6. INTERACTIONS, INSTABILITY, AND SATURATION

[114] Nonlinearity is a ubiquitous aspect of gravity wave dynamics. At smaller wave amplitudes, nonlinearity leads to slow energy transfers via three-wave interactions satisfying resonance conditions for wave frequencies and wave number vectors. At larger amplitudes, multiple-wave, nonresonant, and wave–vortical mode interactions become possible, culminating in the local instability of the wave via generalized wave breaking or shear instability processes as amplitudes increase further. Wave–mean flow interactions may also play important roles at larger wave amplitudes, altering wave propagation, inducing instability, and/or leading to excitation of additional waves. Thus nonlinearity has pervasive influences and is always poised to intercede more vigorously when linear dynamics imply excessive wave amplitudes. A number of reviews of the earlier developments in these areas are available; several relevant to gravity wave instability processes include *Fritts and Rastogi* [1985], *Müller et al.* [1986], *Thorpe* [1987], and *Dunkerton* [1989].

[115] More recent theoretical, numerical, and laboratory studies have expanded our understanding of gravity wave instability processes considerably. Several studies have begun to clarify the diversity of instability character with wave amplitude and propagation angle, including identifying links between the resonant interactions arising at small wave amplitudes and the local instabilities accompanying wave breaking and turbulence generation at large amplitudes [*Klostermeyer*, 1991; *Lombard and Riley*, 1996; *Sonmor and Klaassen*, 1997; *Dunkerton*, 1997a]. Other efforts have demonstrated the role of wave–wave interactions at small and large wave amplitudes [*Dunkerton*, 1987; *Klostermeyer*, 1991; *Thorpe*, 1994; *Vanneste*, 1995] and defined the dynamics of the transition to turbulence for gravity waves that are locally convectively or dynamically unstable [*Winters and D'Asaro*, 1994; *Andreassen et al.*, 1994, 1998; *Fritts et al.*, 1994, 1996a, 1998; *Fritts and Werne*, 2000]. Further studies have addressed the processes accounting for more general spectral energy transfers, effects of finite amplitude and wave–mean flow interactions on wave propagation, dispersion, and instability, and the competition among and finite-amplitude responses to initial instabilities of varying amplitudes. Our purpose in this section is to review the more recent work in these areas.

6.1. Wave–Wave Interactions

[116] Wave–wave interactions as a source of middle atmosphere gravity waves were discussed briefly in section 3. The purpose here is to review their expected influences on the shape and evolution of the gravity wave spectrum more broadly. A number of authors have addressed wave–wave interactions and spectral evolution from a statistical perspective; others have approached the problem deterministically. There remain, however, substantial uncertainties and disagreements over the roles these interactions play in spectral evolution and wave amplitude constraints.

[117] Early studies of wave–wave interactions in the atmosphere identified the dominant interactions previously recognized in oceanic applications [*McComas and Bretherton*, 1977; *Yeh and Liu*, 1981; *Müller et al.*, 1986]. Three-wave (second order) resonant interactions require a match of wave numbers and frequencies among the participating waves satisfying

$$\mathbf{k} = \mathbf{k_1} + \mathbf{k_2} \tag{56}$$

and

$$\hat{\omega} = \hat{\omega}_1 + \hat{\omega}_2 , \tag{57}$$

where subscripts denote the secondary wave quantities, and include (1) elastic scattering, (2) induced diffusion, and (3) parametric subharmonic instability (PSI). Elastic scattering refers to the backscattering of an upward propagating wave into a downward propagating wave (or vice versa) of comparable vertical wave number by a low-frequency motion (or mean flow) having twice the vertical wave number (Bragg scattering). Induced diffusion refers to the transfer of energy from one wave to another having a nearly identical wave number through interaction with a low-frequency (or mean) structure at a much smaller vertical wave number. Alternatively, induced diffusion may be regarded as the evolution of a wave packet in a large-scale shear flow. The third interaction, and arguably the most important in terms of spectral broadening or wave amplitude constraints, is PSI. In its most studied form, PSI represents an exchange of energy from a dominant motion of intermediate frequency to two motions having approximately half the parent frequency and large and nearly opposite wave number vectors, though other wave number combinations exhibit preferred growth in certain circumstances, such as the presence of a mean shear.

[118] More recent efforts have identified other possibilities for energy exchange and further modes of interaction likely to occur among atmospheric waves and other motions. *Broutman and Young* [1986] and *Broutman et al.* [1997], for example, noted that an energetic resonant-triad member can result in an irreversible transfer of energy to small wave numbers not anticipated

by *McComas and Bretherton* [1977]. *Dong and Yeh* [1988] relaxed the assumptions of *Yeh and Liu* [1981] and considered wave–wave interactions not confined to a vertical plane or a resonance surface (nonresonant interactions were also discussed by *Müller et al.* [1986]). Further efforts revealed a potential for nonresonant gravity wave–vortical mode interactions for which the threshold amplitude is reduced or removed in the presence of rotation [*Yeh and Dong*, 1989]. As wave amplitudes increase, additional interactions become energetically viable. *Dunkerton* [1987] performed numerical studies to assess the impact of wave–wave interactions on wave momentum transport, with and without mean shear. He found a rapid energy transfer on the timescale of the forced waves following attainment of large amplitudes, followed by a nonlinear cascade toward increasing complexity as additional interactions became possible. When critical levels were present for the forced waves, *Dunkerton* [1987] found weaker resonant wave–wave interactions, the occurrence of additional nonresonant interactions, and the excitation of a wave having higher intrinsic frequency than the forced waves corresponding to a different branch of the PSI.

[119] *Klostermeyer* [1991] likewise found that larger primary wave amplitudes enable a cascade of interactions that quickly populate the spectrum, while waves that approach or exceed a convectively unstable amplitude support various modes of instability [*Hines*, 1971, 1988b; *Lombard and Riley*, 1996; *Sonmor and Klaassen*, 1997; *Dunkerton*, 1997a]. *Vanneste* [1995] performed both analytic and numerical studies of wave–wave interactions localized in space and noted good agreement between the two methods. In particular, *Vanneste* [1995] found that spatially localized interactions transferred significant energy on a timescale of ~10 primary wave periods for a primary wave amplitude of $a = u'/c \sim 0.25$. *Vanneste* [1995] also noted that mean shear suppresses the dominant unsheared wave–wave interaction but enables a "difference instability" which occurs near the primary wave critical level and has a secondary wave frequency larger than that of the primary wave (like *Dunkerton* [1987]), downward secondary wave propagation, and a growth rate that increases with shear strength. Such interactions are necessarily only resonant at specific locations but, nevertheless, enable rapid energy transfers under certain conditions [*Grimshaw*, 1988].

[120] An example of resonant excitation via PSI in the laboratory study by *Thorpe* [1994] is shown in Figure 13. Here a primary wave with phase aligned along the mean tilt of the tank transfers energy to a lower-frequency wave that becomes the predominant wave at later times. The wave–wave interaction proceeds quickly because of the large forcing wave amplitude, and the secondary wave achieves local convective instability. A numerical wave–wave interaction study performed by *Vanneste* [1995] is displayed in Figure 14 and exhibits the decay of a primary wave at higher frequency (Figure 14b) into two motions having lower frequencies (Figures 14c and 14d). The mean flow also experiences significant modifications in the presence of strong wave–wave interactions and can result in critical-level interactions and convective instability of the wave field [*Sutherland*, 2001]. Hence wave–wave and wave–mean flow interactions should be expected to play increasing roles in spectral evolution and energy and momentum fluxes as wave amplitudes increase with altitude.

[121] Theories of middle atmosphere spectral evolution employing statistical wave–wave interactions have also borrowed from the oceanic literature. The Lagrangian arguments by *Allen and Joseph* [1989] were employed in the Doppler spread theory by *Hines* [1991, 1993] to account for the form of the saturated "tail" spectrum. *Hines* [1996] extended these arguments and used the apparent success of this model to argue against the relevance of other existing saturation theories, most specifically "linear" saturation theory [*Dewan and Good*, 1986; *Smith et al.*, 1987] and the "nonlinear diffusion" theory by *Weinstock* [1976, 1982, 1984, 1990]. The Hines theory attributes the large majority of spectral energy transfers to Doppler spreading by a "broad" spectrum of waves, with wave amplitudes and interactions increasing with altitude. The theory presumes that dissipation generally does not occur except at small scales having vertical wave numbers $m > m_M$, where m_M is the maximum vertical wave number for which spectral character is determined largely by Doppler spreading, and infers a spectral shape approaching the m^{-3} value most often cited by other saturation theories and various observations cited above. Advantages of the Hines theory include a recognition of the increasing statistical importance of nonlinear wave–wave interactions with increasing wave amplitudes and a natural transition from "linear" source spectra at lower altitudes to a nonlinear "saturated" spectrum at higher altitudes.

[122] There has been considerable confusion over the *Allen and Joseph* [1989] results and their implications for Doppler-spread theory over time. *Hines* [1999] noted an error in the Allen and Joseph analysis overlooked in his earlier papers and suggested a corrected high-m spectral form of m^{-1} rather than m^{-3}. More recently, *Hines* [2001] and *Chunchuzov* [2002] argued for the removal of the m^{-1} term arising from the error correction to again infer a high-m spectrum of the form m^{-3} but with an amplitude ~2–3 times smaller than typically observed.

[123] The Hines Doppler spread theory has its detractors, however. Several authors have performed numerical studies that appear to seriously undermine the assumptions of Doppler spread theory [*Zhong et al.*, 1995; *Bruhwiler and Kaper*, 1995; *Broutman et al.*, 1997; *Eckermann*, 1997; *Buckley et al.*, 1999; *Walterscheid*, 2000]. Essentially, these studies show that when time dependence and vertical motions of the underlying wave field are accounted for, the tendency for transfer of spectral energy to ever smaller vertical scales is substantially reduced (or reversed). *Zhong et al.* [1995] employed

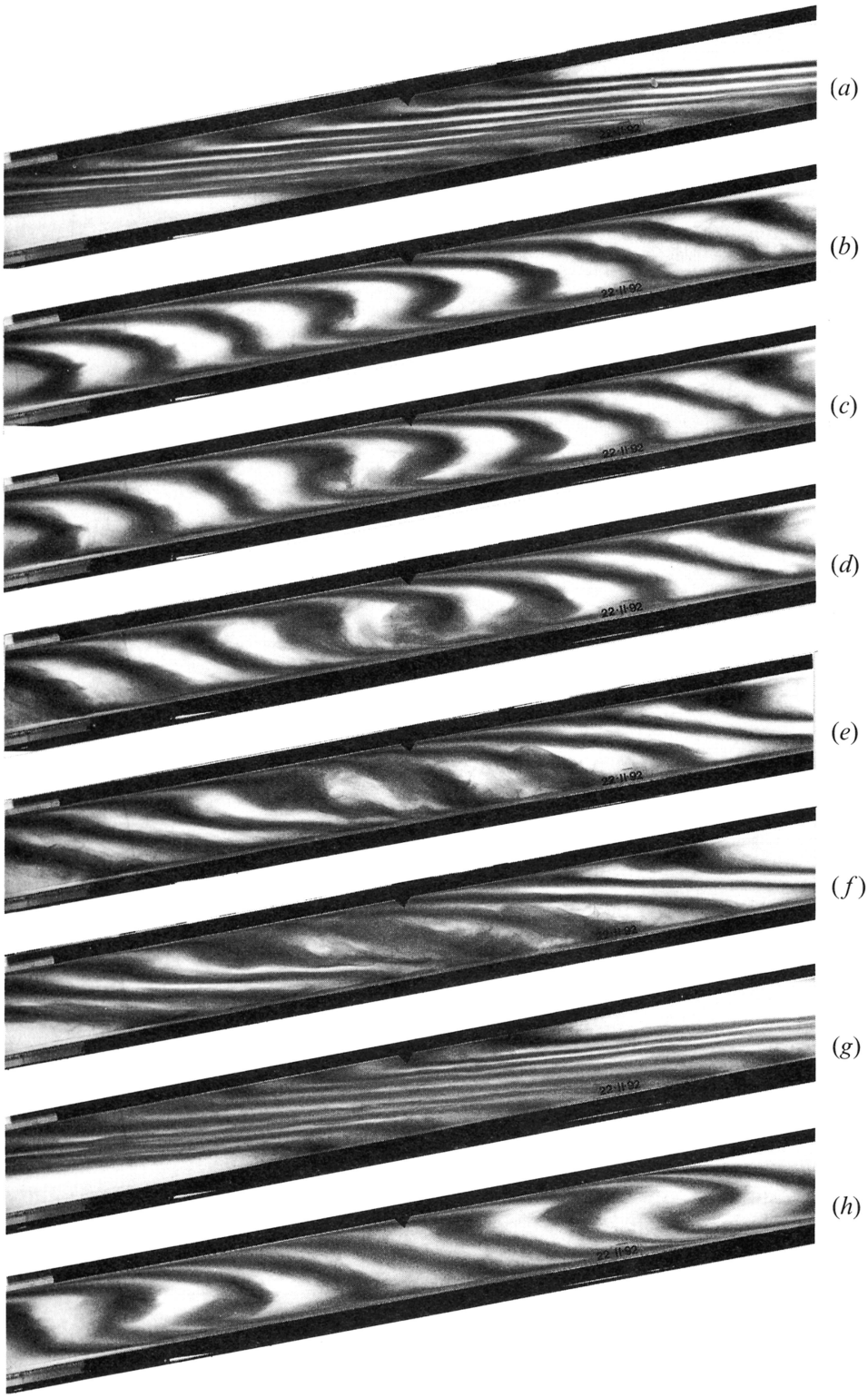


Figure 13. Dyed isopycnal surfaces in a rocked inclined tube showing the development and instability of a wave excited by the PSI. Note that the primary wave (with uniform structure along the tube) contributes largely at early times, whereas the subharmonic wave dominates the instability at later stages. After *Thorpe* [1994] (reprinted with permission of Cambridge University Press).

ray-tracing techniques to examine tidal modulation of wave propagation and noted that time dependence caused critical levels to become transient. *Bruhwiler and Kaper* [1995] formulated the problem from a Hamiltonian perspective and showed that the theory was in good agreement with ray-tracing results for a spectrum of waves with various background wave amplitudes. They found, in particular, a tendency for low-*m* waves to

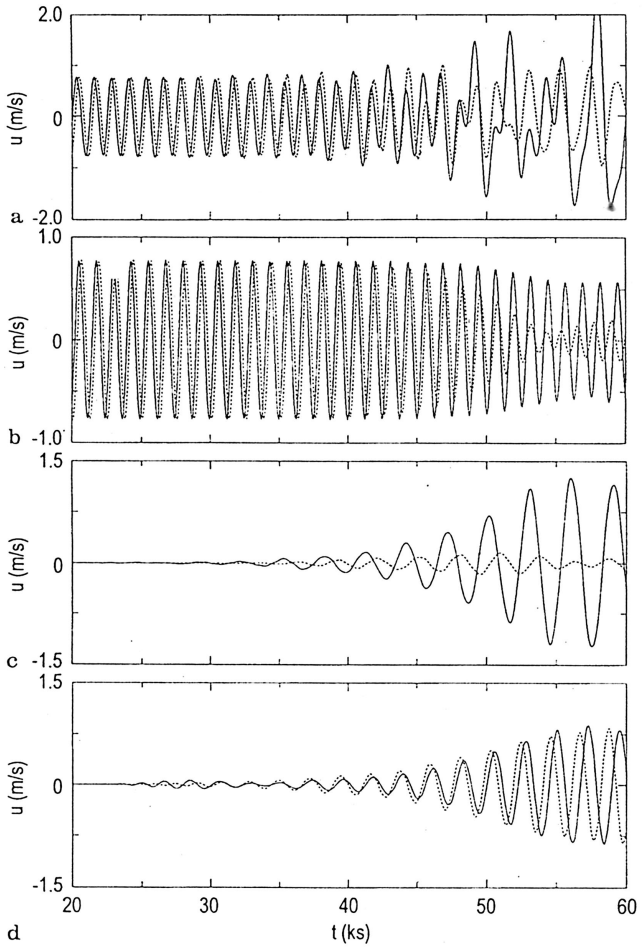


Figure 14. Numerical simulation of PSI in an unsheared atmosphere with a scale height of 7 km. Shown are the (first panel) total, (second panel) primary wave, and (third and fourth panels) two secondary wave vertical velocities at altitudes of 10.7 (solid lines) and 17.75 km (dashed lines). Note that both secondary waves grow to amplitudes exceeding that of the primary wave as the interaction proceeds. After *Vanneste* [1995] (reprinted with permission of Springer-Verlag).

remain at low m and for high-m waves to move preferentially to lower m, representing a sharp departure from expectations based on Doppler spread arguments.

[124] *Broutman et al.* [1997] performed both ray-tracing and full numerical simulations, demonstrating the equivalence of the two approaches and the various influences of spatial and temporal variability on short-wave propagation. These results are illustrated in Figure 15. In particular, they exhibit a scattering of short-wave (large m) energy to both smaller and larger wave numbers, consistent with previous assessments based only on ray tracing [*Broutman*, 1986; *Broutman and Young*, 1986; *Bruhwiler and Kaper*, 1995], and a spectral form close to m^{-3} that is consistent with the results of *Eckermann* [1999]. *Eckermann* [1997] and *Walterscheid* [2000] also considered single waves in backgrounds having various character. *Eckermann* [1997] addressed specifically the various assumptions underlying the Doppler spread theory by *Hines* [1991, 1993, 1996], concluding that wave field transience significantly reduces the tendency to transfer wave energy to smaller vertical scales and eliminates a cutoff in unspread vertical wave numbers near m_c, itself a key element of Doppler spread theory. *Walterscheid* [2000] obtained similar results, emphasizing both the mitigating effects of time dependence on Doppler spreading and the role of vertical velocities in allowing wave packets to penetrate anticipated critical levels. *Hines* [1999] raised strong objections to the analysis by *Eckermann* [1997], arguing that the multiple-wave background offset the results of a single-wave background to a significant degree.

[125] *Buckley et al.* [1999] and *Sonmor and Klaassen* [2000] extended the analyses by Eckermann, Broutman et al., and Walterscheid by considering the effects of a spectrum of background waves, isolated wave packets, and a mean shear. Both studies suggest a weaker tendency for refraction toward large m than with the steady assumptions underlying Doppler spread theory. *Sonmor and Klaassen* [2000] also found that multiple waves induce caustics under more general conditions than supported for single-wave environments, while both studies found that mean shears have significant cumulative refraction effects over timescales of a few inertial periods.

[126] Doppler spread theory, as it is presently developed, precludes competitive instability processes (i.e., linear saturation theory) from contributing to the shape of the vertical wave number spectrum except, to quote from *Hines* [1996], at wave numbers "a little less than m_M; but, arising as they almost certainly do in small-scale 'white-cap' regions, their effect is unlikely to extend very far". However, an increasing number of numerical and observational studies are providing evidence that local wave instability, via convective or shear instability, at large vertical scales is more the rule than the exception (see section 6.2.1), despite the clear potential for wave–wave interactions to contribute importantly to energy transfers as wave amplitudes increase. Additional observations suggesting that the atmospheric wave spectrum is often not broad but is composed of a single or a few dominant motions further undermine the ability of Doppler spreading among multiple waves to provide large enough spectral transfers and to reproduce the observed amplitude limits.

[127] Clearly, the jury is still out on these issues. There is clear evidence, as noted, of discrete, large-amplitude events or superpositions of a few waves which yield canonical spectral shapes in an Eulerian frame but which suggest a simpler Lagrangian viewpoint. Indeed, *Hines* [2001] stated "that the occurrence of the large-wave number Eulerian tail has nothing whatever to do with any physical process. . . . If one insists on defining 'waves' according to their Eulerian linear description, then one is forced to admit to the existence of 'nonlinear wave–wave interactions'. . .but these are mere mathematical artifacts and have no physical import." However, while transformation to a Lagrangian coordinate does remove

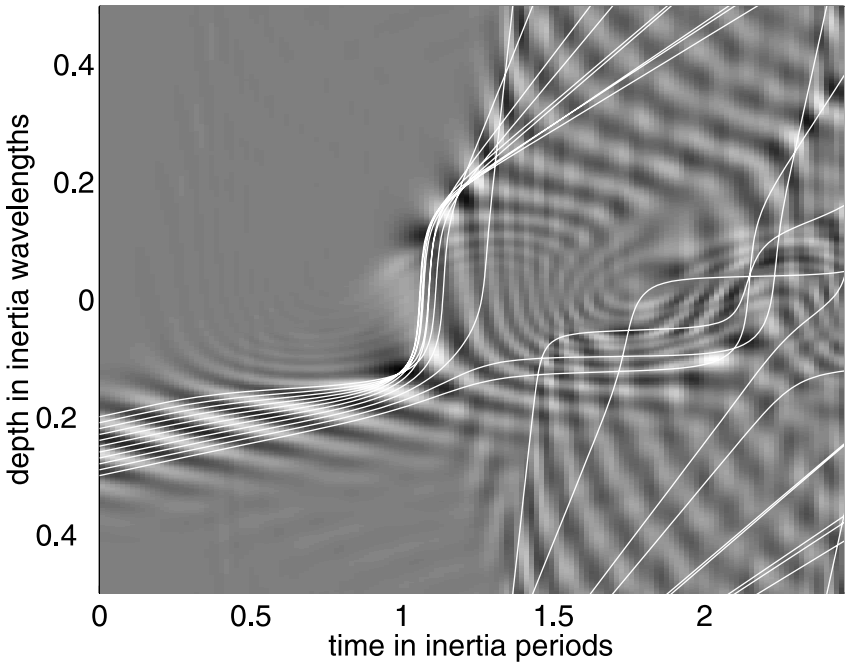


Figure 15. Perturbation density illustrating the encounter of a small-scale wave packet with an inertia-gravity wave of much larger scale. In this case, the small-scale wave has a horizontal phase speed 0.2 of the maximum of the inertia-gravity wave motion, no critical level trapping occurs, and both the vertical wavelength and horizontal phase speed of the small-scale wave increase dramatically. After *Broutman et al.* [1997].

the advective nonlinearity in the Navier-Stokes equations, it does not remove the coupling of wave scales that nonlinearity implies. For example, claims of Lagrangian linearity cannot explain the nonlinear coupling of scales observed in nonlinear simulations or laboratory studies revealing significant wave–wave interactions [*McEwan*, 1971; *Klostermeyer*, 1991; *Thorpe*, 1994; *Vanneste*, 1995]. Likewise, shears due to mean winds and large-scale waves imply critical levels, small vertical scales, and instability processes that are demonstrably not mathematical artifacts. Indeed, Broutman et al. (personal communication, 2002) have pointed out that the dispersion relation employed by *Hines* [2001, 2002a] does not include the influences of mean wind shear on vertical wave structure and thus does not describe the full effects of Doppler shifting. The important issue is not whether nonlinear wave–wave (and wave–mean flow) interactions are important, as they surely are, but whether they are the dominant mechanism in constraining wave amplitudes and shaping the spectrum for $m < m_M$, as argued by *Hines* [2001] and *Chunchuzov* [2002], or whether they are one of several key processes acting in this fashion, which seems a more defensible perspective, given the various evidence available to date. The identified links between wave instabilities at small and large amplitudes [*Sonmor and Klaassen*, 1997], however, may render such distinctions obsolete.

[128] The *Weinstock* [1976, 1982, 1985] theory was discussed in previous reviews and has itself served as the basis for additional saturation theories by *Zhu* [1994] and *Medvedev and Klaassen* [1995] having various attributes. All of these theories view nonlinearity as spectral diffusion of one form or another. There is no clearly defined link, however, with the predictions of wave–wave interaction theory or direct numerical studies [e.g., *Dunkerton*, 1987; *Klostermeyer*, 1991; *Vanneste*, 1995], so testing of the physical basis for the theory has proved challenging. Recently, *Hines* [2002b] has raised serious objections to assumptions in the original work by Weinstock and its successors by association. However, while Weinstock's theory is clearly an approximation to a class of nonlinear effects, like Doppler spread theory, it calls attention to the importance of nonlinear interactions in shaping the wave spectrum with increasing altitude. Indeed, Hines's criticism has itself been recently challenged (Klaassen and Medvedev, personal communication, 2000). In summary, nonlinear wave–wave interactions must be regarded, at present, as a viable means of exchanging energy among gravity waves at various scales and frequencies throughout the middle atmosphere. Evidence suggests, however, that they cannot account, by themselves, for the general shape or amplitude of the gravity wave spectrum. Rather, they represent one of several nonlinear processes that act jointly to define gravity wave spectral character and evolution with altitude. Wave–wave interactions likewise do not contribute directly to energy dissipation, except through their links to specific instabilities at larger wave amplitudes [*McEwan*, 1971].

6.2. Instability and Turbulence Dynamics

[129] Theoretical, numerical, and observational studies have made substantial contributions to our understanding of instability and turbulence dynamics accompanying gravity waves in recent years. Perhaps the greatest advances were made possible by the continuing evolution of high-performance computers, which are now capable of direct numerical simulations (DNS) of stratified and sheared flows having resolutions of $\sim 1000^3$ and above. Such simulations are capable of describing both the transition to turbulence in a geophysical flow and the vorticity dynamics driving the turbulence cascade. Observational capabilities have also progressed significantly over the last decade or so and have yielded high-resolution in situ and ground-based measurements of wave and turbulence structures. Other numerical and laboratory studies have contributed to our understanding of general turbulence dynamics in applications not specific to atmospheric gravity waves.

[130] Turbulence is generally believed to arise locally within a gravity wave (or a field of superposed waves) when the flow is either convectively or dynamically unstable and the timescale for instability growth is sufficiently shorter than that describing the evolution of the wave field. More recent analyses have shown, however, that a wide spectrum of instabilities is possible, with specific instability character depending on wave amplitude and intrinsic frequency, mean shear and stability profiles, and the form and amplitude of perturbations triggering flow instability. In particular, local instability can occur at wave amplitudes below that often considered necessary for convective instability in many instances [*McEwan*, 1971; *Hines*, 1988b; *Lombard and Riley*, 1996; *Sonmor and Klaassen*, 1997]. It is also believed, based on simple stability arguments, that dynamical instability should predominate for intrinsic frequencies $\hat{\omega} \sim f$ and that convective instabilities should predominate for intrinsic frequencies $\hat{\omega} \gg f$ [*Dunkerton*, 1984; *Fritts and Rastogi*, 1985] because of the very different relative amplitude thresholds in each case. These expectations are supported by the numerical simulations performed to date [*Andreassen et al.*, 1994, 1998; *Fritts et al.*, 1994, 1998, 2003; *LeLong and Dunkerton*, 1998a, 1998b].

[131] Because local instabilities at wave amplitudes below that normally identified as the "convective" instability limit,

$$a = u'/(c - \bar{u}) = 1.0\,, \tag{58}$$

generally exhibit slow growth, however [*Lombard and Riley*, 1996; C. Bizon, personal communication, 2001], we assume for our discussion that the nominal threshold amplitudes for dynamical and convective instability are those which yield $Ri < 1/4$ and $Ri < 0$, respectively, with the Richardson number (Ri) defined to be

$$Ri = \frac{N^2}{\bar{u}_z^2 + \bar{v}_z^2}\,, \tag{59}$$

where $\bar{u}_z$ and $\bar{v}_z$ are the local mean zonal and meridional wind shears and N is the local buoyancy frequency. For most intrinsic frequencies the two threshold amplitudes are nearly identical [*Dunkerton*, 1984; *Fritts and Rastogi*, 1985], departing significantly only for $\hat{\omega} \sim f$.

[132] With the high-resolution simulations now possible on current supercomputers, DNS studies of gravity wave breaking at sufficiently high intrinsic frequencies (i.e., more nearly comparable, rather than disparate, horizontal and vertical wavelengths) can capture both gravity wave scales and a broad range of turbulence scales simultaneously. On the other hand, the large disparity between inertia-gravity wave and turbulence scales ($\sim 10^6$ or more) prevents simulations that span this range of scales. Instead, relevant simulations have addressed the KH instability of inertia-gravity waves and the turbulence arising from KH instability separately.

6.2.1. Gravity Wave Breaking

[133] A number of numerical studies addressed gravity wave breaking and the accompanying wave amplitude limits in 2-D for lack of adequate computational resources to perform full 3-D studies. Because the instability processes are inherently 3-D, however, 2-D studies are either limited in their utility (e.g., in quantifying wave amplitude limits and momentum flux divergence) or complete misrepresentations of the relevant dynamics (e.g., in modeling wave breaking and turbulence dynamics) [*Andreassen et al.*, 1994]. Thus, we will review here only those studies addressing the 3-D character of wave instability or the occurrence and scales of such events in the atmosphere.

[134] The first 3-D numerical studies of gravity wave instability dynamics in the atmosphere were performed by *Andreassen et al.* [1994], *Fritts et al.* [1994], and *Isler et al.* [1994]. A parallel study examining similar dynamics in an oceanic context was performed by *Winters and D'Asaro* [1994]. These studies addressed gravity wave breaking via convective instability, which appears to be the preferred instability for gravity waves at relatively high intrinsic frequencies [*Dunkerton*, 1984; *Fritts and Rastogi*, 1985]. These simulations succeeded in describing the character of the primary wave instability process; this comprises counterrotating streamwise (along the flow) convective rolls or vortices (with spanwise, or normal, wave number). These convective rolls occupy the full depth of the convectively unstable region within the gravity wave, derive their eddy energy (or vorticity) from baroclinic and shear sources within the 2-D flow, and trigger a turbulence cascade via mutual vortex interactions as they interact with adjacent shear layers (or vortex sheets). Indeed, the initial study of gravity wave breaking in 3-D provided a plausible explanation for apparent streamwise instability structures observed in NLC near the summer mesopause [*Fritts et al.*, 1993b] (see Figure 16). Additional observational evidence of such instabilities has come from recent analyses of air-

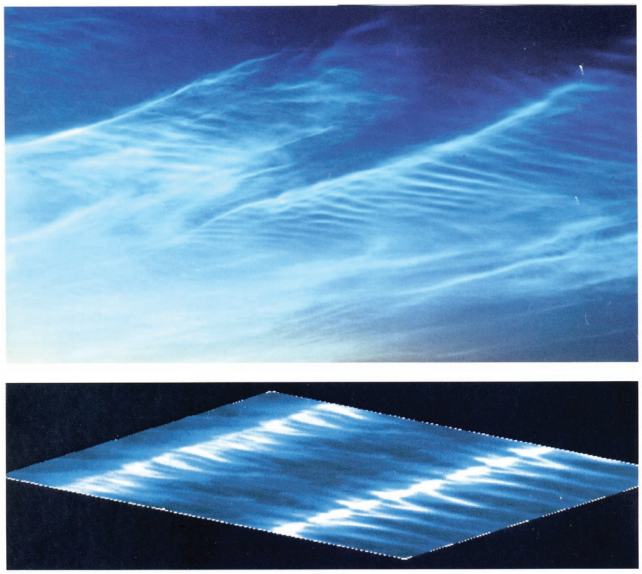


Figure 16. (top) Observed and (bottom) simulated gravity wave and instability structures in NLC. Photo was taken at Turku, Finland, on 21–22 July 1989, courtesy of Pekka Parvianien. Streamwise-aligned instability structures are believed to have accounted for the smaller-scale bright bands oriented approximately normal to the gravity wave phase fronts in the upper image. After *Fritts et al.* [1993b].

glow imaging data and correlative wind measurements [*Swenson and Mende*, 1994; *Hecht et al.*, 1997, 2000].

[135] Subsequent numerical studies at higher resolution addressed wave breaking in streamwise and transverse shear flows [*Fritts et al.*, 1996a], with superposed low- and high-frequency motions [*Fritts et al.*, 1997b], and the vorticity dynamics driving the turbulence cascade [*Andreassen et al.*, 1998; *Fritts et al.*, 1998]. The latter studies addressed in detail the vorticity dynamics of the initial shear-aligned instabilities as well as their subsequent interactions with adjacent vorticity sheets (the mean and 2-D gravity wave shears having spanwise vorticity).

[136] The vorticity field arising because of convective instability of the gravity wave, the subsequent transition to turbulence, and the dynamics within the turbulent flow are illustrated in Figure 17 at various times throughout the simulation described by *Andreassen et al.* [1998] and *Fritts et al.* [1998, 1999]. The simulation was performed for a gravity wave of horizontal wavelength ~30 km and intrinsic frequency at the level of wave breaking of $\sim N/6$. A mean shear was imposed to confine the breaking to the region below the initial critical level for the wave. The entire transition from a 2-D overturning gravity wave to quasi-isotropic turbulence occupies approximately a buoyancy period in this simulation. Vortex structures are displayed using a quantity representing the rotational or "tube-like" character of the motion field analogous to the minimum pressures within the flow (see *Jeong and Hussain* [1995] and *Andreassen et al.* [1998] for further details). The volumes are viewed from below, with the streamwise direction and wave propagation to the right.

[137] The transition from 2-D laminar flow to 3-D quasi-isotropic turbulence accompanying wave breaking involves several distinct phases: (1) the initial shear-aligned convective instability within the gravity wave (first two images in Figure 17), (2) a second phase in which divergent spanwise motions below adjacent streamwise vortices stretch and thin the adjacent spanwise vortex sheets, causing them to become locally dynamically unstable, (3) the formation of secondary, spanwise-localized KH billows on the intensified vortex sheets and their linkage to the overlying streamwise vortices to form series of intertwined vortex loops (second, third, and fourth images in Figure 17), and (4) the subsequent interactions of neighboring vortices, the excitation of twist waves on the various vortices, and the unraveling, fragmentation, and breakup of the vortices that comprises the cascade to smaller scales of motion (last four images in Figure 17). Indeed, the various stages in the transition to turbulence often occur nearly simultaneously in different (or the same) portions of the flow. Similar simulations having a spanwise mean shear component and more recent higher-resolution simulations having no mean shear confirm the general nature of the transition and turbulence dynamics described above [*Fritts et al.*, 1996a, 2003].

[138] Various stages in the turbulence evolution noted above also parallel in important respects vortex dynamics observed in other turbulent flows. Intertwined vortex loops, or similar "hairpin" or "horseshoe" vortices, are seen to arise in sheared boundary layers [*Gerz et al.*, 1994; *Adrian et al.*, 2000]; twist waves, so easily excited in the wave breaking simulations, were also observed, but not recognized as such, in laboratory studies of vortex dynamics [*Cadot et al.*, 1995]. Finally, the sequence of vortex dynamics seen in gravity wave breaking exhibits striking parallels to that observed in the transition to turbulence accompanying the KH instability to be discussed below [see also *Arendt et al.*, 1997, 1998; *Fritts et al.*, 1999]).

[139] The scales at which gravity wave breaking occurs are becoming better known based on both direct and indirect measurements. Such observations also address the current debate over the mechanisms constraining wave amplitudes and imposing saturation of the gravity wave spectrum (see sections 6.1 and 6.3). Direct measurements of density or temperature (via lidar and balloon- or rocket-borne instrumentation) reveal frequent occurrences of near-adiabatic or superadiabatic lapse rates, with occurrence frequency and vertical scale increasing with altitude. Recent examples include balloon and Rayleigh lidar measurements of large-amplitude wave motions at lower altitudes [*Hauchecorne et al.*, 1987; *Shutts et al.*, 1988; *Wilson et al.*, 1991a; *Whiteway and Carswell*, 1995; *Hoppe et al.*, 1999] and sodium resonance lidar measurements at greater altitudes

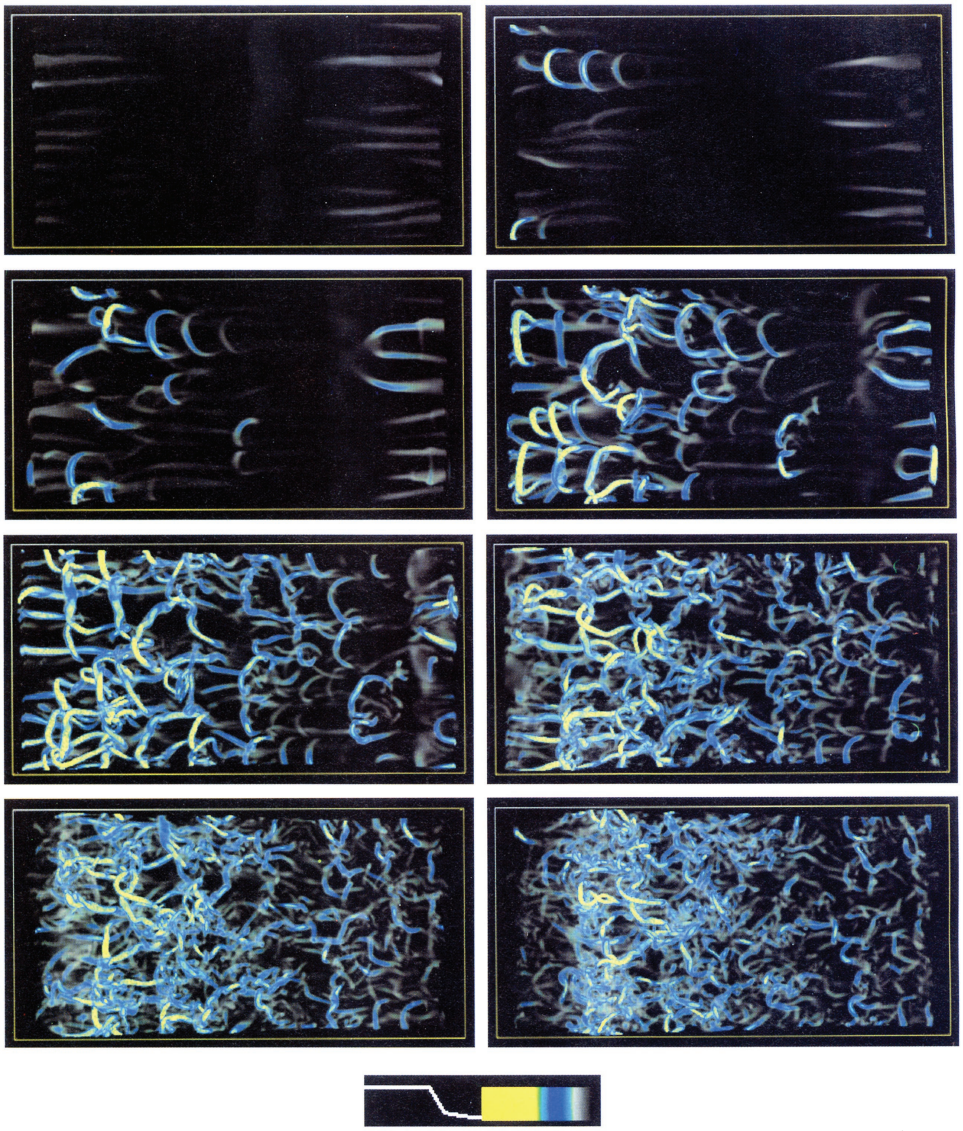


Figure 17. Vorticity structures in a breaking gravity wave viewed from below with the direction of wave propagation to the right. The images display vortex structures from left to right and top to bottom, spanning one buoyancy period following attainment of convective instability of the wave field. See text for details. After *Fritts et al.* [1999] (reprinted with permission of Terra Scientific Publishing Company).

[*Hecht et al.*, 1997; *Williams et al.*, 2002]. Rocket measurements employ various techniques, including falling spheres [*Fritts et al.*, 1988b], chaff [*Wu and Widdel*, 1991], ionization gauges [*Lübken*, 1997], and Rayleigh lidars [*Hoppe et al.*, 1999]. These observations reveal unstable or nearly unstable layers typically up to a few kilometers in depth, suggesting convectively unstable gravity waves having vertical wavelengths of ~3–10 km or more, with the larger scales more prevalent near the mesopause. The recent sodium lidar measurements by *Williams et al.* [2002] represent a particularly striking example, with multiple superadiabatic layers a few kilometers in depth within a near-adiabatic layer extending more than 10 km and persisting for more than 4 hours (see Figure 18). As noted above, these observations suggest, at least in this instance, wave saturation via local convective instability, though wave superposition or wave–wave interactions, specifically the PSI displayed in Figure 13, may play a role in the vertical fine structure within the overturning wave field. While this event is exceptional, the frequency of observation of such features suggests that they are far from the infrequent or pathological cases implied by *Hines* [1996]. Indirect inferences of convective overturning scales are provided by NLC and airglow measurements of horizontal instability scales [*Fritts et al.*, 1993b; *Swenson and Mende*, 1994; *Hecht et al.*, 1997, 2000], where the observed roll spacing is suggested to be comparable to the unstable layer depth in numerical simulations. Additional inferences of gravity wave amplitudes and instability depths are obtained from radar wind measurements via the gravity wave dispersion relation [*Muraoka et al.*, 1988].

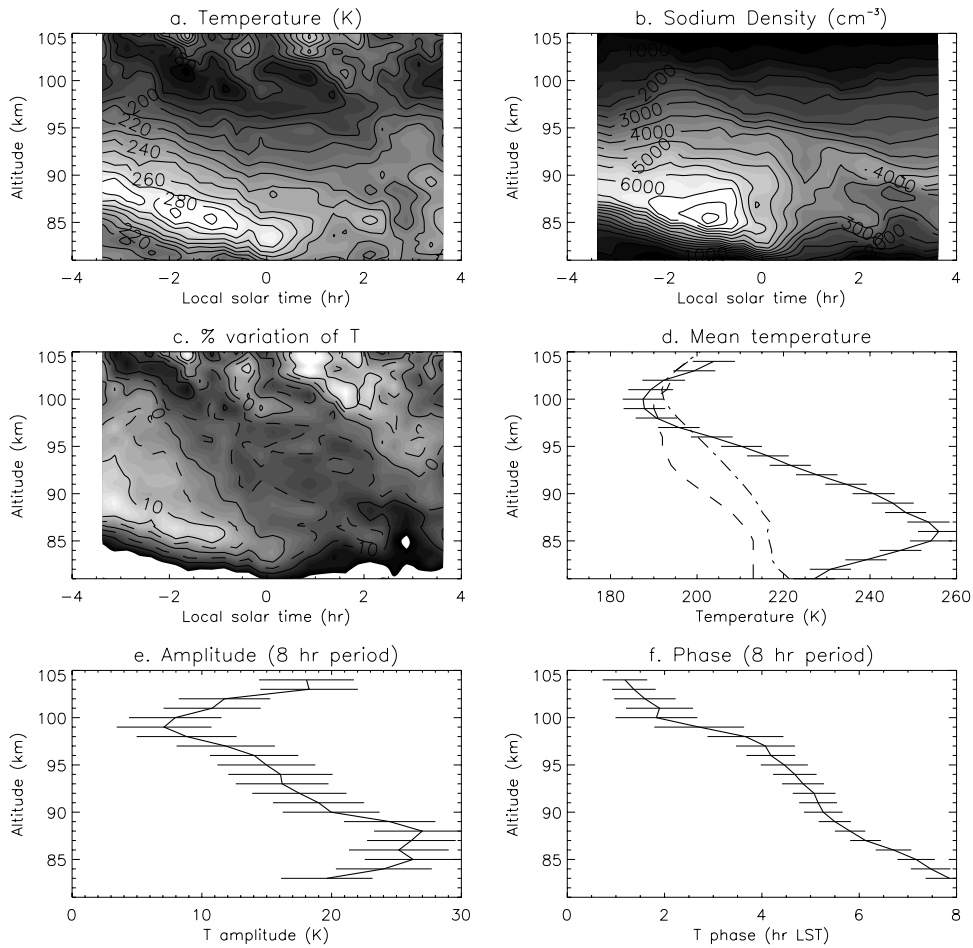


Figure 18. Contours of (a) temperature, (b) sodium density, (c) temperature perturbations as a percent of the mean value, and (d) profile of mean temperature for the 8-hour wave observation using the Fort Collins sodium lidar on day 307 of 1997. Also shown are the amplitude of the 8-hour wave in the temperature field (e) and the phase of the maximum temperature (f). The 1σ uncertainties are shown with horizontal bars and the dashed and dash-dotted lines in Figure 18d are the diurnal means for a 24-hour campaign in November 1998 and nightly mean for November averaged from 1990 to 1997. After *Williams et al.* [2002].

6.2.2. Kelvin-Helmholtz Instability

[140] KH instability is among the most common sources of turbulence in the atmosphere. In many instances it owes its existence, in part at least, to wind shears due to inertia-gravity waves because such motions contribute preferentially to wind shears throughout the atmosphere. Earlier evidence of KH instability due to inertia-gravity waves was reviewed by *Fritts and Rastogi* [1985]. More recent stability analyses have identified the preferred modes of instability of inertia-gravity waves with and without mean shears [*Fritts and Yuan*, 1989c; *Yuan and Fritts*, 1989; *Dunkerton*, 1997a]. Numerical simulations of unstable inertia-gravity waves revealed the character of the KH instability and its preferred orientation within the wave field [*LeLong and Dunkerton*, 1998a, 1998b]. The numerical studies showed a preferred direction consistent with the stability analysis by *Fritts and Yuan* [1989c] but with an increasing degree of isotropy as $\hat{\omega} \rightarrow f$. These simulations were also used to assess the theoretical prediction of the wave amplitude required for onset of shear instability by *Dunkerton* [1984] and *Fritts and Rastogi* [1985]. Results are displayed in Figure 19 and reveal that time dependence of the wave field increases the wave amplitude required to initiate instability, as anticipated by *Lombard and Riley* [1996] and *Sutherland* [2001] at higher intrinsic frequencies.

[141] Other evidence of the importance of KH instability and its relation to inertia-gravity waves comes from observations of wind and temperature profiles in the lower and middle stratosphere. Radar measurements at several sites have revealed persistent layers of enhanced radar reflectivity with spacings of a few kilometers, often exhibiting slow vertical motions and inferred unstable wave amplitudes [*Sato and Woodman*, 1982; *Yamamoto et al.*, 1987]. Likewise, high-resolution balloon measurements have revealed multiple layers yielding signatures of local turbulent mixing, with near-adiabatic layers sandwiched between sharp temperature inversions [*Cot and Barat*, 1986; *Coulman et al.*, 1995]. Examples of the

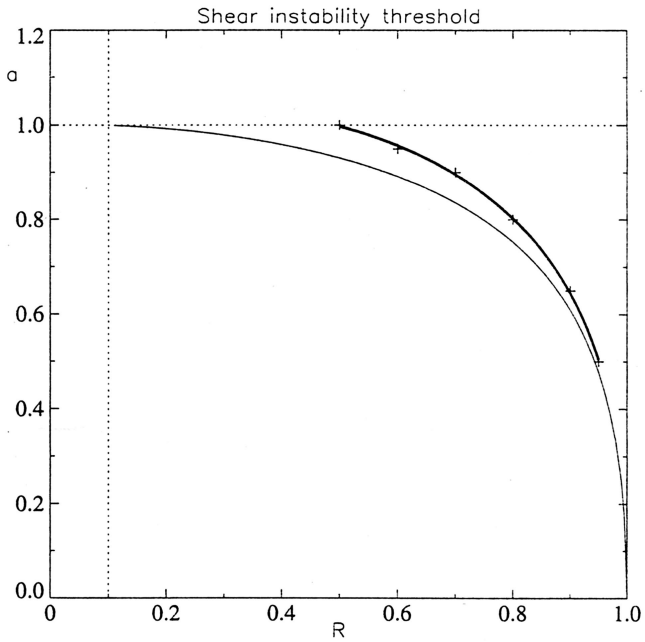


Figure 19. Threshold amplitude a for shear (KH) instability of inertia-gravity waves as a function of $R = f/\hat{\omega}$ in the absence of mean shear based on theory for a time-independent flow (solid line) and DNS simulations (thick solid line with data points). After *LeLong and Dunkerton* [1998a] (reprinted with permission of the American Meteorological Society).

latter, with profiles of the corresponding Richardson number, are displayed in the top and middle panels of Figure 20. Note, in particular, that the observations suggest a minimum Richardson number of ~1/4 following instability and mixing, well below what was often assumed to accompany restratification in the past. It will be seen below that these measurements are in good agreement with recent high-resolution DNS studies of KH instability.

[142] While observations help define the scales at which KH instability and mixing occur and the consequences of mixing for the temperature and wind fields, they cannot identify the dynamics of the turbulence transition or the evolution of the shear layer as KH breakdown and mixing occur. The only means of understanding these aspects of inertia-gravity wave instability is via DNS studies spanning KH growth, breakdown, and restratification.

[143] There have been many numerical studies of KH instability in the past. Only recently, however, have computational capabilities permitted 3-D studies of high enough resolution to confirm the predictions and laboratory observations of secondary instability [*Klaassen and Peltier*, 1985; *Thorpe*, 1987; *Palmer et al.*, 1994, 1996; *Caulfield and Peltier*, 1994; *Fritts et al.*, 1996b; *Smyth*, 1999]. Early 3-D KH simulations were not adequate to describe the transition to turbulence, the structure and anisotropy within the inertial-range of turbulence, and the implications for dissipation and mixing. However, recent simulations are now adequate to investigate these turbulence effects [*Werne and Fritts*, 1999, 2001; *Smyth and Moum*, 2001]. The vorticity dynamics accompanying this evolution and its relation to the vorticity dynamics of wave breaking were described by *Fritts and Werne* [2000].

[144] The KH simulations performed by *Werne and Fritts* [1999, 2001] have illustrated the dynamics of the transition to turbulence, the expansion of turbulence throughout the KH billow and the shear layer, and the turbulence statistics, dissipation, anisotropy, and decay. Most striking, perhaps, is the great degree of similarity of the turbulence dynamics to those observed in previous simulations of gravity wave breaking. Common features include (1) the initial streamwise, or shear-aligned, convective instability within the outer regions of the KH billow, (2) the stretching and wrapping of spanwise vortex sheets around the streamwise vortices, (3) the dynamical instability and rollup of the localized and intensified vortex sheets, and (4) the subsequent vortex interactions and twist wave excitation driving the cascade to smaller scales of motion [*Fritts and Werne*, 2000]. Major differences in the turbulence transitions accompanying wave breaking and shear instability are the timescale and the character of mixing. As noted above, the transition due to high-frequency wave breaking requires approximately one buoyancy period for the wave parameters simulated. The transition is somewhat longer due to KH instability, however, because turbulence requires time to expand from the initial site of instability throughout the KH billow and the full depth of the evolving shear layer. Implications for mixing are probably very different as well since KH instability mixes the shear layer vigorously prior to restratification, whereas turbulence due to wave breaking is quickly advected out of the unstable phase of the wave motion, allowing vigorous turbulence to act within the stably stratified portions of the wave field.

[145] The vorticity dynamics and impact on the thermal field of KH turbulence are illustrated in Figure 21 at four times throughout the simulation described by *Werne and Fritts* [1999]. In this simulation, a buoyancy period corresponds to 28 time units. Note here the gradual penetration of turbulence throughout the KH billow, the rapid obliteration of thermal gradients by turbulence shortly after it occurs, the homogenization of the expanded shear layer, the very sharp thermal gradients that evolve as a result of efficient mixing within the shear layer, and the continuing small-scale dynamics and mixing within the edge regions of the mixed layer at later times. Profiles obtained in the KH DNS at $t = 165$ are displayed in the fourth panel of Figure 20 and exhibit good agreement with those observed by *Coulman et al.* [1995] (first and second panels), despite the markedly different Reynolds numbers of the two flows. The implications for inertia-gravity waves include (1) instability, turbulence, and mixing on timescales short compared to wave periods, hence confined to the unstable phase of the wave, (2) reduced shears and wave amplitudes due to mixing, and (3) alternating layers of low and high static

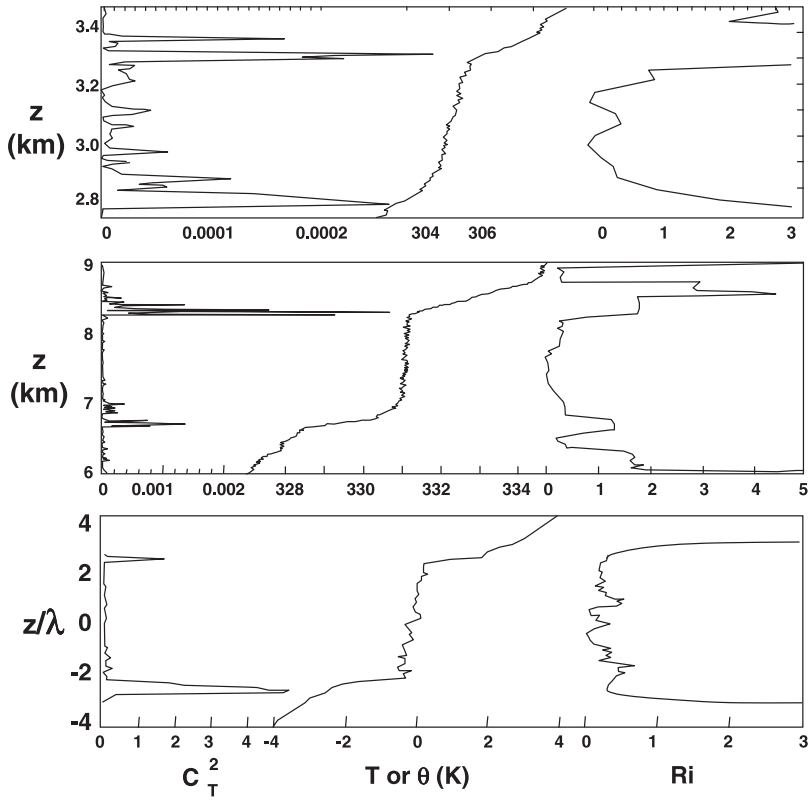


Figure 20. Profiles of C_T^2, θ, and *Ri* obtained by (top and middle) high-resolution balloon measurements and (bottom) DNS showing the aftermath of KH instability and mixing events. Note that the magnitude of C_T^2 and the gradient of θ maximize at the edges of the mixed region and that $Ri \approx 1/4$ throughout the mixed region. Modified after *Coulman et al.* [1995] and *Fritts et al.* [2003].

stability that may influence subsequent instability and mixing processes.

6.2.3. Synthesis and Other Instability and Large-Amplitude Processes

[146] Studies cited above have expanded greatly our understanding of gravity wave instability processes over the last decade. Importantly, these studies have also suggested links between specific modes of instability of small- and large-amplitude gravity waves. *Klostermeyer* [1991] found the PSI to be related to a 2-D parametric instability at finite amplitude and suggested, based on his analysis, that resonance may underlie all gravity wave instability processes. *Klostermeyer* [1991] and *Lombard and Riley* [1996] also identified transverse modes of instability at finite amplitudes below those required for local convective or dynamical instability.

[147] The links between instabilities at small and large wave amplitudes were further clarified and generalized by *Sonmor and Klaassen* [1997] to include all instability types and all wave amplitudes and intrinsic frequencies (or phase propagation angles) assuming no rotation. Specifically, they found a link between the PSI and the "slantwise static instability" (SSI) of *Hines* [1971, 1988b] at small wave amplitudes and with convective instability and high wave number parametric instability at amplitudes $a > 1$ (see equation (58)). *Sonmor and Klaassen* [1997] further linked the "branch-C" instability of *Yeh and Liu* [1981] to the resonant elastic scattering and induced diffusion interactions originally identified by *McComas and Bretherton* [1977] but noted quite different growth rates than inferred in the earlier study. The "shear-aligned instability" (SA, having spanwise wave number, as noted in the wave breaking simulations discussed in section 6.2.1) is itself linked to higher-order resonant instabilities at smaller wave amplitudes and exhibits an amplitude threshold near the convective limit for waves at high intrinsic frequencies [*Winters and Riley*, 1992; *Sonmor and Klaassen*, 1997]. A 2-D dual-mode instability represents a generalization of the KH instability within an inertia-gravity wave and competes favorably with SA or convective instability at large amplitudes [*Sonmor and Klaassen*, 1997] but is itself stabilized by weak environmental shear [*Thorpe*, 1994]. *Dunkerton* [1997a] examined the relative roles of these instabilities in the presence of rotation, employing a steady, plane-parallel approximation to the wave structure and obtained similar results concerning the dominant instabilities at large wave amplitudes.

[148] A further class of instability, "oblique instabilities" (having nonzero streamwise and spanwise wave numbers), was found by *Lombard and Riley* [1996] and *Sonmor and Klaassen* [1996, 1997] to represent the most rapidly growing instability for an important physical

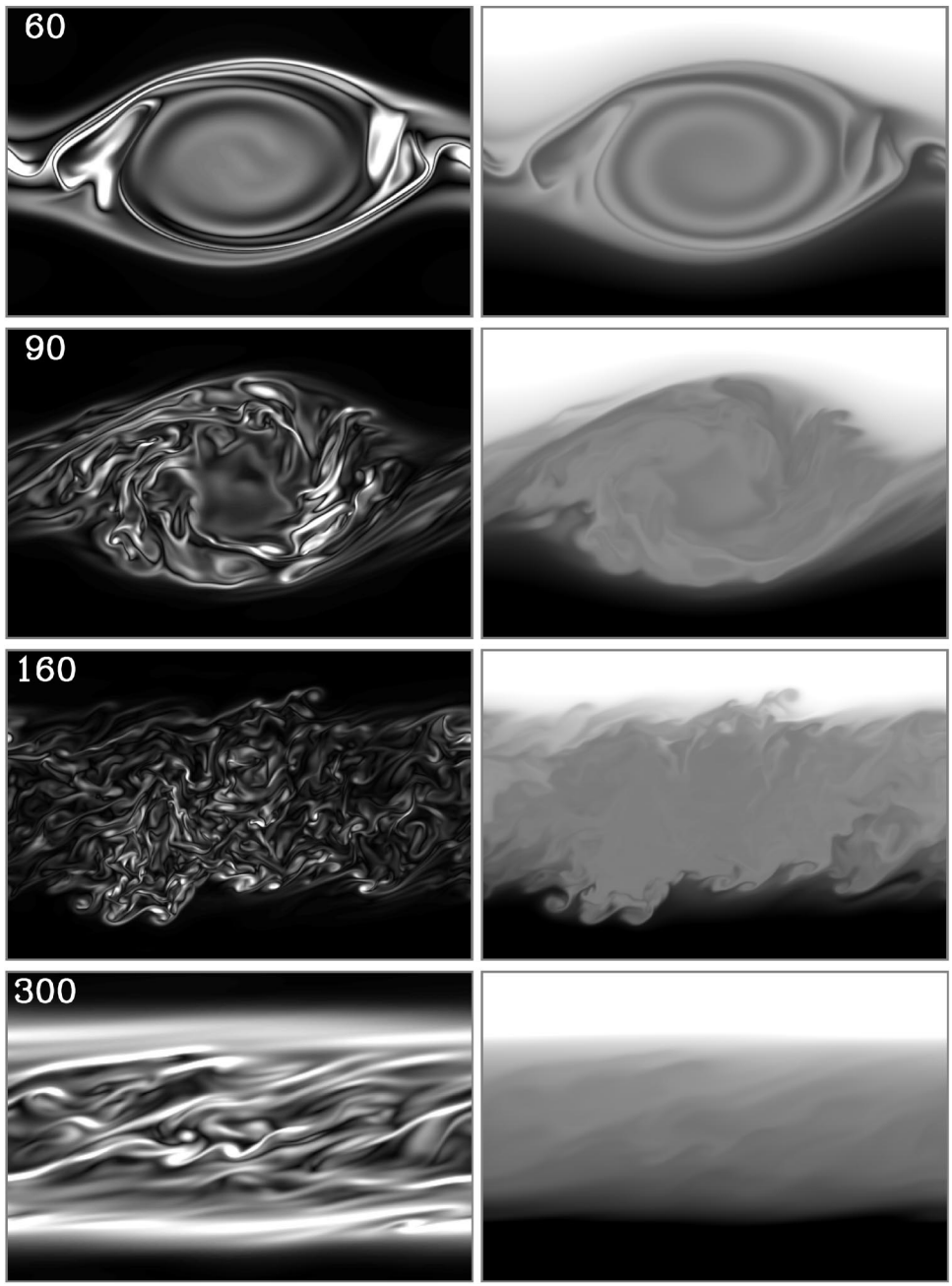


Figure 21. The 2-D streamwise-vertical cross sections of (left) 3-D vorticity magnitude and (right) temperature fields at four times throughout the turbulence transition and restratification of a KH billow. See text for details. After *Werne and Fritts* [2001] (reprinted with permission of Elsevier Science).

range of wave amplitudes at high intrinsic frequencies. *Sonmor and Klaassen* [1996, 1997] argue that oblique instabilities connect smoothly to SA instabilities at higher frequencies and larger wave amplitudes but are the preferred form when shear is the primary source of instability energy, whereas SA instabilities are preferred when buoyancy is the dominant source. The parameter ranges in which these various instabilities predominate are displayed in Figure 22. Here the dash-dotted line denotes an overturning amplitude, and high intrinsic frequencies have large phase elevation angles. We note, however, that recent DNS studies [*Fritts et al.*, 2003] reveal SA instabilities to predominate over 2-D dual-mode and oblique instabilities at finite amplitude for a wide range of the wave parameters displayed in Figure 22. This may be due either to the Sonmor and Klaassen analysis being inviscid (while the DNS studies are viscous) or to the selection of a finite-amplitude response in the DNS from among a variety of initial instabilities at infinitesimal amplitude.

[149] Other types of instabilities are enabled as a result of wave packet localization or by the induced mean flows accompanying waves of finite amplitude. *Sutherland* [2001] demonstrated numerically the modulational

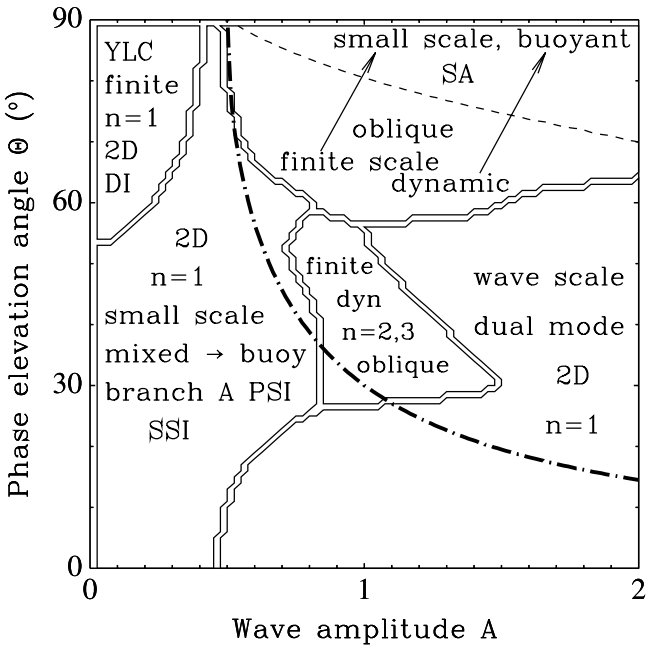


Figure 22. Fastest growing linear instabilities as a function of wave amplitude and phase elevation angle. The dash-dotted line corresponds to the amplitude of overturning ($a = 1$ in our terminology), and the dashed line represents the transition from oblique to shear-aligned instability and from shear to buoyancy as the dominant source of eddy energy. After *Sonmor and Klaassen* [1997] (reprinted with permission of the American Meteorological Society).

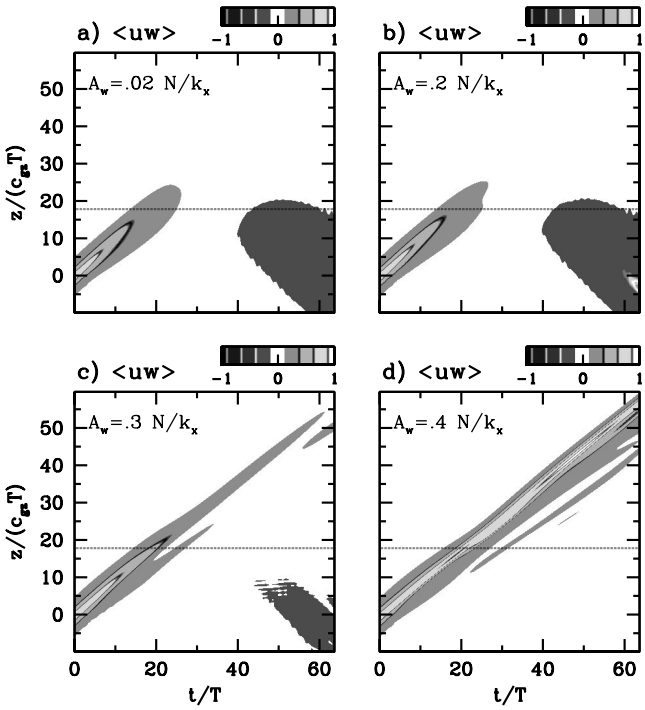


Figure 23. Time series of wave momentum flux for wave amplitudes of various vertical velocities incident on a turning level. The turning level anticipated by linear theory is shown by the horizontal line in each panel. After *Sutherland* [2000] (reprinted with permission of the Royal Meteorological Society).

instability (and stability) of a localized wave packet anticipated by *Whitham* [1965, 1974] but noted that this instability does not necessarily imply wave breaking. Sutherland also argued that self-acceleration of a vertically localized wave packet can lead to local convective instability at sufficiently large amplitudes and high intrinsic frequencies, providing a means by which wave localization can contribute to instability. The mechanism by which instability occurs is the differential tilting of surfaces of potential temperature by the wave-induced mean flow. Similar effects accompany localized wave packets incident on a turning level and may lead to wave instability via either self-acceleration (as above) or as a result of enhanced wave amplitude due to incident and reflected wave superposition [*Sutherland*, 2000]. Interesting additional consequences of wave packet localization and finite amplitude include tendencies (1) for wave packets that are vertically localized and of large amplitude to yield permanent momentum flux divergence near a turning level and (2) for wave packets that are both horizontally and vertically localized to penetrate substantial distances beyond a turning level based on linear theory [*Sutherland*, 1999, 2000]. Indeed, both of these mechanisms may have important but as yet unquantified implications for gravity wave momentum transport and redistribution in the middle atmosphere. The tendency for spatially localized, large-amplitude packets to penetrate turning levels may be particularly relevant for middle atmosphere momentum transport, where such packets may be more the rule than the exception. A demonstration of the tendency for penetration of a turning level as wave amplitude increases is shown in Figure 23. In these cases, total and partial reflection occur for small and intermediate wave amplitudes, while wave packet propagation appears to be increasingly unaffected by mean shear (and increasing intrinsic phase speed and frequency) as wave amplitude increases. Wave packet transmission and reflection for the largest wave amplitude are displayed at several stages throughout the event in Figure 24. Note, in particular, the phase distortions at the leading edge of the wave packet above the turning level which exhibit the self-acceleration effects discussed by *Fritts and Dunkerton* [1984] and *Sutherland* [1999, 2000].

[150] A final topic relevant to our understanding of instability dynamics in gravity waves is the optimal perturbation theory pioneered by *Farrell and Ioannou* [1996a, and references therein]. This theory has provided considerable insights into the occurrence of or competition among various instability modes in Couette, Poiseuille, and other time-independent shear flows known to be asymptotically stable below a threshold (or for any) Reynolds number. However, the theory also can be applied to time-dependent flows and to more general stratified problems including secondary KH and gravity wave instabilities for which the equations of motion are

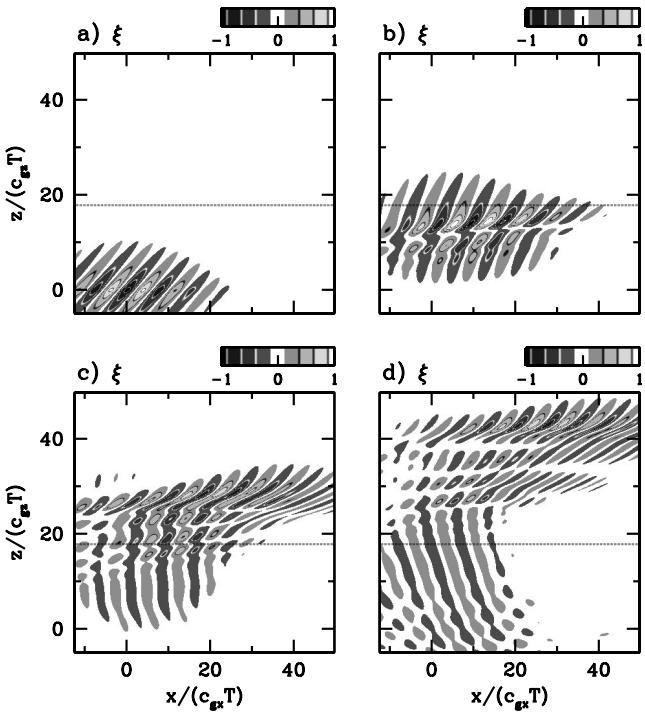


Figure 24. As in Figure 23 but showing vertical displacements at four times throughout the turning level encounter for the wave packet of largest amplitude. After *Sutherland* [2000] (reprinted with permission of the Royal Meteorological Society).

non-self-adjoint [*Farrell and Ioannou*, 1996b]. In such flows, particular superpositions of linear modes, termed "optimal perturbations", may exhibit rapid transient growth exceeding the growth rates of all eigenfunctions of the linear system. If this growth achieves a sufficient amplitude to have nonlinear consequences, either among the perturbations or for the mean flow, then it will be the optimal perturbation which dictates the finite-amplitude behavior of the nonlinear system. In applications to the KH instability problem, the optimal perturbation methodology has shown that the structures of both the initial 2-D instability and the secondary 3-D instabilities depend sensitively on the form and amplitude of the initial perturbations (J. Werne, personal communication, 2001). Optimal perturbation theory also reveals why the details of the turbulence transition and cascade discussed above depend so sensitively on the initial noise spectrum [*Werne and Fritts*, 1999; *Fritts and Werne*, 2000]. In the atmosphere, then, we anticipate that instability structures accompanying wave breaking or shear instability may often be dictated more by other finite amplitude perturbations than by the exact form of the most rapidly growing linear eigenmode of the unstable flow. While initial perturbations dictate the details of the transition to turbulence, they appear not to alter the statistical effects of turbulence on the larger-scale flow [*Fritts and Werne*, 2000].

6.3. Saturation Theories

[151] The dynamics underlying gravity wave saturation in the atmosphere, i.e., instability and wave–wave interaction processes, appear to impose a near-universal spectral form far from the dominant sources at lower altitudes (but subject to all of the caveats about departures from universality noted above). Our purpose in this section is to review the various theories for the shape and evolution of the wave spectrum and to identify how they are linked to the instability dynamics discussed above.

[152] The earliest wave amplitude limits attributed to instability dynamics were proposed by *Hodges* [1967] and involved the convective instability of a gravity wave due to its exponential growth with height. Subsequent studies by many authors considered the implications of such dissipation for momentum transports, turbulence, and mixing (see *Fritts* [1984a] for a review). Motivated by *VanZandt*'s [1982] suggestion of a universal gravity wave spectrum and their own observations of near universality in stratospheric wind spectra [*Dewan et al.*, 1984], *Dewan and Good* [1986] proposed the first theory of a saturated spectrum employing saturation (via local convective or shear instability) separately at each vertical wave number. Their dimensional analysis yielded a saturated range for gravity waves varying as CN^2/m^3 with C dependent on spectral bandwidth (and a bandwidth $dm \sim m$). The study by *Smith et al.* [1987] extended this theory to account for attainment of saturated wave amplitudes by the wave spectrum as a whole, yielding a saturation spectrum amplitude of $N^2/6m^3$ for velocities and a natural explanation for the evolution of the dominant vertical wave number, m_*, toward smaller m with increasing altitude. Further refinements of this theory yielded a corresponding saturated amplitude for gravity wave temperature spectra of $N^4/10g^2m^3$ and implications for enhanced saturation and dissipation where N increases with altitude [*Fritts et al.*, 1988a; *VanZandt and Fritts*, 1989]. Attributes of this "linear saturation theory" include a simple conceptual mechanism for amplitude limits for superposed waves of various scales and frequencies, general agreement of the predicted spectral form and amplitude and variations of wave energy density with altitude with observations throughout the atmosphere [*Allen and Vincent*, 1995], and a natural explanation for the increase in the dominant vertical wavelength with increasing altitude [*Tsuda et al.*, 1989]. Failings of the theory include an inability of the spectral description to account for specific wave sources having character quite different from the canonical spectral form, an inability to account for amplitudes in the stratosphere below the saturation limit because of wind shear effects on vertical wavelengths [*Eckermann*, 1995a; *Alexander*, 1996], relative amplitudes of velocity and temperature spectra that depart from predictions [*Tsuda et al.*, 1991; *Nastrom et al.*, 1997; *de la Torre et al.*, 1999], and the neglect of wave–wave interactions that are known to cause spectral energy transfers.

[153] An alternative explanation of the evolution of gravity wave spectral characteristics with increasing altitude was provided by *Weinstock* [1982, 1985, 1990] based on "nonlinear damping." Weinstock argued that this mechanism could both constrain wave amplitudes and account for spectral broadening with increasing altitude. Accounting for nonlinear influences within the wave spectrum, even without accounting explicitly for resonant interactions, was an appealing approach and spawned a number of extensions to the theory by *Zhu* [1994], *Gardner* [1994], and *Medvedev and Klaassen* [1995]. As with linear saturation theory, however, the various extensions to Weinstock's theory have their detractors (see section 6.1), and clarification of the mechanisms controlling spectral amplitude and shape remains a high priority. Appealing aspects of the nonlinear damping theories include a recognition of the central role of nonlinearity in shaping the wave spectrum and approximate agreement with observed spectral shape and evolution with altitude. Liabilities include an inability to account explicitly for local wave field instability and turbulence and concerns by *Hines* [2002b] over the legitimacy of the underlying assumptions.

[154] A further approach to gravity wave saturation and spectral evolution was taken by *Hines* [1991, 1993, 1996], who modeled wave-wave interactions by "Doppler spreading" of any one member of the spectrum by all of the remaining members. The Hines theory relies on the Lagrangian description of wave–wave interactions by *Allen and Joseph* [1989] and shares the attributes and liabilities of this formalism discussed at length in section 6.1. This approach accounts qualitatively for increasing interactions and effects with increasing altitude and yields predictions of a tail spectrum that is also in approximate agreement with observations, subject to correction of the earlier estimates of spectral slope [*Hines*, 2001; *Chunchuzov*, 2002]. The theory has also led to the development of a gravity wave parameterization that has experienced a number of successes (see section 7) but has been criticized (see section 6.1) for simplifications in the underlying assumptions that are inconsistent with the claimed cascade toward smaller vertical scales. Perhaps its major contribution has been a focusing of attention and renewed research activity on the role of wave-wave interactions in gravity wave spectral evolution in the atmosphere.

[155] A final approach to describing gravity wave spectral shape and evolution in the atmosphere is the "saturated-cascade similitude" theory by *Dewan* [1991, 1994, 1997]. The theory, as it has developed, embodies both the linear saturation ideas of *Dewan and Good* [1986] and *Smith et al.* [1987] as amplitude constraints and the notion that wave-wave interactions drive a cascade of wave energy from larger to smaller scales analogous to that observed in inertial-range turbulence. As in turbulence theory, spectral amplitudes are controlled by the energy dissipation rate ϵ, which represents both energy input at large scales via vertical wave propagation and energy dissipation at sufficiently small scales. The theory is appealing in its simplicity and its dependence on scaling arguments, and it makes a number of predictions of 1-D spectra that are either in reasonable agreement with observations or pose observational tests of the theory. The suggestion that saturation and cascade dynamics are jointly at work in shaping the gravity wave spectrum is more realistic than theories relying largely on one or the other. Dewan's theory is also more in line with recent results by *Sonmor and Klaassen* [1997] indicating that local wave field instability and spectral energy transfers via wave-wave interactions are manifestations of related dynamics at different wave amplitudes. As such, further advances merging or generalizing saturation and cascade processes and effects may help to unify the divergent views of wave saturation and spectral evolution prevailing at present. There are, however, a number of shortcomings or uncertainties in the present formulation, hence motivation for further improvements to the theory. It is unclear but likely not the case that wave energy is transferred from large to small scales without dissipation, while the direction and/or efficiency of the cascade are uncertain, based on the discussion in section 6.1. Finally, the current theory does not predict 2-D spectra, although these would be valuable in assessing the interaction and cascade dynamics, and the prediction of the vertical wave number spectrum of vertical velocities has a slope of $+1$ which differs significantly from the various observations to date of order -1 to -3 [*Kuo et al.*, 1985; *Larsen et al.*, 1986; *Wu and Widdel*, 1990; *Fritts and Hoppe*, 1995].

7. GRAVITY WAVE PARAMETERIZATIONS

[156] The effects of gravity waves on the larger-scale circulation must be parameterized in global models because the spatial resolution required to model them directly is prohibitively fine. The waves important to the circulation in the atmosphere have horizontal wavelengths ranging from approximately tens to thousands of kilometers. The largest of these can be resolved in some global models; however, the vertical wavelength of a gravity wave will vary substantially with height owing to the effects of wind shear. Vertical wavelengths as small as 1–2 km are important to resolve in the lower stratosphere. Gravity waves with periods as short as ~10 min can carry significant momentum flux vertically. Further, the sources of these waves include processes that are also parameterized and/or poorly resolved, namely, convective heating, fine-scale topography, localized shear zones, and frontal structures. The parameterizations for convection for example are designed for purposes other than generating a realistic wave spectrum [*Ricciardulli and Garcia*, 2000]. Direct modeling of a realistic gravity wave spectrum in global models is therefore still not feasible at the present time nor in the near future. Two effects of gravity wave dissipation that are currently

believed to be necessary in global models are a body force on the background flow (important at the tropopause and throughout the middle atmosphere) and mixing effects on temperature and trace constituent distributions (important in the upper stratosphere and above).

[157] The essential ingredients of gravity wave parameterizations include specification of input parameters describing the gravity wave sources and estimation of the wave dissipation as a function of height. Observational constraints on parameterization inputs are, however, still lacking in necessary detail. There is also no current consensus on how to compute the dissipation with height. Therefore middle atmosphere models currently have an uncomfortably large degree of freedom in the parameterization of gravity wave effects.

[158] There are currently at least eight parameterization schemes proposed or in various stages of implementation in global models [*Lindzen*, 1981; *Holton*, 1982; *Palmer et al.*, 1986; *McFarlane*, 1987; *Fritts and Lu*, 1993; *Medvedev and Klaasen*, 1995; *Hines*, 1997a, 1997b; *Lott and Miller*, 1997; *Alexander and Dunkerton*, 1999; *Scinocca and McFarlane*, 2000; *Warner and McIntyre*, 2001]. *McLandress* [1998] reviews the approach and basic mathematics for four of these. Here we will not embark on detailed descriptions of each parameterization. We will instead discuss the basic components of a gravity wave parameterization and describe some of the commonalities and differences among the various approaches in that context. The basic components are (1) specification of the characteristics of the waves at the source level, (2) wave propagation and/or spectral evolution as a function of height, and (3) wave dissipation and calculation of the effects on the background atmosphere.

7.1. Parameterization of Gravity Wave Sources

[159] Parameterizations of gravity wave effects all require input of the amplitudes and propagation properties of the waves at some altitude in the troposphere or at the model lower boundary. These would ideally be related to specific gravity wave sources as in the work of *Rind et al.* [1988], but this approach is insufficiently constrained by present observations. Stationary gravity waves ($c = 0$) generated by flow over topography are presently parameterized in many global models [e.g., *Palmer et al.*, 1986; *McFarlane*, 1987; *Miller et al.* 1989]; however, the properties of other nonstationary gravity waves ($c \neq 0$) and their sources are too poorly constrained to be included in this way. Instead, nonstationary waves, when they are included at all, are input with globally uniform properties or with properties that vary only as a function of latitude [*Garcia and Solomon*, 1985; *Holton and Schoeberl*, 1988; *Smith and Brasseur*, 1991; *McLandress and Ward*, 1994; *Roble and Ridley*, 1994; *Norton and Thuburn*, 1996; *Mayr et al.*, 1997a; *Manzini and McFarlane*, 1998; *Medvedev et al.*, 1998; *Scaife et al.*, 2000].

[160] Linear gravity wave theory tells us that the required information about gravity wave sources includes their phase speeds, horizontal wavelengths, propagation directions, and momentum fluxes. (Note that horizontal wavelength or phase speed could be replaced with vertical wavelength or frequency using the dispersion relation (23).) A factor describing gravity wave spatial and/or temporal intermittency must also be considered because the waves with the specified properties will not generally be occurring throughout the model area at all times.

[161] *Palmer et al.* [1986] and *McFarlane* [1987] defined parameterizations for topographic wave sources. Momentum fluxes are related to the subgrid-scale orographic variance, atmospheric stability, and surface wind. The waves are launched propagating opposite to the surface wind with phase speed relative to the ground $c = 0$. *Bacmeister* [1993] further considered the alignment of mountain ridges perpendicular to the surface wind. An efficiency factor is generally included because only a fraction of the surface stress will be carried by vertically propagating waves [*Schär and Durran*, 1997; *Lott*, 1998]. Horizontal wavelength is a free parameter, generally chosen to be ~100 km or considered to be a tunable parameter.

[162] *Kershaw* [1995] describes a parameterization for momentum flux due to convectively generated gravity waves by assuming a mechanism analogous to topographic waves, but the waves are assigned a phase speed equal to the wind speed at the center of the convective layer. *Chun and Baik* [1998] describe a convective heat source parameterization that models convective heating as a steady process that similarly generates only waves that are stationary relative to the heat source. The horizontal wavelength spectrum that results is also highly dependent on a set of assumed properties of the convective heating that are currently poorly constrained. By contrast, other model studies [e.g., *Piani et al.*, 2000] (also see Figure 5) instead suggest that in the absence of strong shear, convection may generate a very different spectrum of gravity waves with much larger phase speeds (see section 3.1).

[163] For nonstationary gravity waves ($c \neq 0$), there is no consensus on the value of momentum flux at the source level, so it is yet another free parameter. Many model studies have not reported the value that has been assumed, making comparisons difficult. The best constraints on the momentum flux are indirect and are derived from estimates of the zonal mean wave-driven force that must be present to explain global temperature, wind, and trace constituent patterns. These constrain the gravity wave momentum flux crossing the tropopause and are described in section 8.1. A wide variety of gravity wave parameterizations are currently in use in global models. Intercomparison of these studies will begin to be possible when distributions of the cross-tropopause gravity wave flux with latitude are reported, as *Manzini and McFarlane* [1998] have done.

[164] Other properties of the spectrum of nonstationary waves input to parameterizations currently vary widely, and these differences may be of fundamental importance in explaining the differences among the results obtained. Some input an isotropic energy spectrum that is separable in $\hat{\omega}$ and m with power law dependencies on these variables that resemble observations in the troposphere and stratosphere (see section 4) [*Fritts and VanZandt*, 1993; *Fritts and Lu*, 1993; *Hines*, 1997a, 1997b; *Medvedev et al.*, 1998; *Medvedev and Klaassen*, 2000; *Warner and McIntyre*, 2001]. *Hines* [1997a] argues that the fundamental difficulty with this approach is that the low-m portion of the spectrum is very poorly constrained by observations, yet it is this portion that has the largest effect in the mesosphere. The assumption of isotropy is also poorly constrained and probably wrong. *Alexander and Dunkerton* [1999] proposed a parameterization that can employ arbitrary source spectral shapes and anisotropies that are input as momentum flux versus phase speed. This approach suffers similar uncertainty in the input spectrum, particularly at high phase speeds (corresponding to low m), although model studies and observations of high phase speed waves generated by deep convection are beginning to constrain the phase speed spectrum of this source (see section 3.1.2).

[165] For nonstationary wave sources like convection the intermittency may be a small number, $\varepsilon \ll 1$, yet there are few observational constraints for this parameter. *Alexander* [1996] derived estimates of intermittency as a function of frequency and horizontal wavelength for high-frequency waves above deep convection in a cloud-resolving model, and *Zink and Vincent* [2001b] estimated intermittency for low-frequency waves in midlatitude radiosonde observations. The values will depend on the bandwidths of the wave properties considered (e.g., c, $\hat{\omega}$, etc.) [*Alexander and Dunkerton*, 1999].

7.2. Treatment of Wave Propagation and Dissipation

7.2.1. Commonalities and Differences

[166] All parameterizations assume a set of gravity wave properties, specified either as a continuous wave spectrum or as a collection of discrete waves at some source altitude. The spectral properties of the sources may vary considerably among the different parameterizations, and except for topographic waves (input at the ground), the source altitude for nonstationary waves will additionally vary even among different applications of the same parameterization. Differences in the specification of gravity wave sources among parameterizations may be responsible for some large part of the differences in their effects seen in models. However, no controlled comparisons have yet been made that address this issue.

[167] All parameterizations are one-dimensional in that they assume the waves considered propagate only vertically and that only vertical variations in the background atmosphere influence wave propagation. All parameterizations use linear theory in the absence of dissipation and use the midrange frequency approximation to the linear dispersion relation (32) for at least some portion of the calculation. Total internal reflection is included via equation (30) in one [*Alexander and Dunkerton*, 1999], but this is generally only important for gravity waves with horizontal wavelengths shorter than ~50 km.

[168] Aside from gravity wave source specification, the main differences among parameterizations reside in their treatment of nonlinearity and mechanisms of wave dissipation. These differences result in different vertical variations in gravity wave effects.

7.2.2. Treatment of Nonlinearity and Wave Dissipation

[169] Most global model studies that have included gravity wave effects parameterized nonlinear effects using a scheme related to that proposed by *Lindzen* [1981] and formalized by *Holton* [1982]. These "Lindzen-type" parameterizations specify some discrete set of wave properties at a source level, use the midrange frequency approximation to calculate wave propagation with height, and then use the linear convective instability threshold to determine where wave dissipation will occur. Wave dissipation continues above the unstable level via a saturation condition (see *Fritts* [1984a] for a review). This basic formula is the basis of the widely used mountain wave parameterizations by *Palmer et al.* [1986] and *McFarlane* [1987], who added a Froude number condition for finding the altitude regions where wave saturation would occur. There are also nonstationary wave analogs to these mountain wave parameterizations [e.g., *Kiehl et al.*, 1996; *Norton and Thuburn*, 1999].

[170] *Alexander and Dunkerton* [1999] used Lindzen's convective instability criterion but assumed total wave breakdown at that level rather than employing the wave saturation assumption. This simplification reduced computation time, allowing a treatment of a detailed spectrum of waves rather than the typical small set of 5–10 waves employed in Lindzen-type schemes. They also made explicit the separation of the parameters describing local wave momentum flux magnitude (which determines breaking levels) and net wave momentum flux (which is generally smaller if there is intermittency in the wave occurrence). The parameterization by *Norton and Thuburn* [1999] instead locked these two parameters together so that smaller fluxes gave higher breaking levels and vice versa. This is appropriate for mountain waves where amplitude and net flux are naturally linked because the intermittency in the source is modeled directly via topography and surface wind variations but is inappropriate for nonstationary wave sources with globally averaged input properties.

[171] *Warner and McIntyre* [2001] used the midrange frequency approximation to the linear dispersion relation and a saturation condition to limit wave amplitudes to a "quasi-saturation" spectrum. They developed a spectral analogue to the *Lindzen* [1981] saturation crite-

rion and simplified the computation with a three-part spectrum approximation. Short horizontal wavelength waves in their source spectrum [*Warner and McIntyre*, 1996] carry significant momentum flux, so their neglect of wave reflection is a potentially serious problem. *Fritts and VanZandt* [1993] and *Fritts and Lu* [1993] used saturation theory and empirical constraints on the gravity wave spectrum to describe the spectrally integrated energy density and momentum flux as functions of height.

[172] *Hines* [1991, 1997a, 1997b] used linear theory and the midrange frequency approximation to describe wave propagation with height but treated dissipation as resulting from wave-wave interactions which cause Doppler-shifting and refraction of portions of the wave spectrum to vertical wavelengths shorter than a prescribed cutoff wave number where dissipation is assumed to occur. *Medvedev and Klaassen* [1995, 2000] instead treated dissipation via the nonlinear diffusion ideas of *Weinstock* [1982, 1990]. Both of these mechanisms have, however, been questioned in the recent literature (see section 6.1).

[173] Note that although many of the spectral gravity wave parameterization schemes [*Fritts and Lu*, 1993; *Medvedev and Klaasen*, 1995, 2000; *Hines*, 1997a, 1997b; *Warner and McIntyre*, 1996, 2001] have been constrained to reproduce the "universal" m^{-3} gravity wave energy spectrum observed at high m (section 4), these parameterizations can produce very different results (see section 8.4). This leads to the natural conclusion that the shape and magnitude of the spectrum at high m alone do not sufficiently constrain parameterizations of gravity wave effects in the middle atmosphere.

[174] For a parameterization to be effective at both forcing the stratospheric QBO winds and driving the mesopause wind reversals, dissipation of nonstationary waves must occur for increasingly large intrinsic phase speeds and increasingly long vertical wavelengths at increasing altitudes. In other words, waves may dissipate fairly close to their critical levels in shear zones in the lower stratosphere but must dissipate far from their critical levels in the upper mesosphere. Intrinsic phase speed and vertical wavelength are related via equation (33). The result is that for nonstationary waves the force will tend to accelerate the mean wind (the force will have the same sign as the wind) in the lower stratosphere but decelerate the mean wind (the force will have opposite sign as the wind) in the upper mesosphere. (Note that this trend should not hold for the large amplitude mountain waves with phase speeds $c \sim 0$ that have been observed. Such waves can break and cause decelerations at all levels.) The potential for gravity waves to accelerate the mean wind is one reason to avoid the term "gravity wave drag" since this is an example of "gravity wave push."

[175] The vertical variation in wave dissipation described above is a natural consequence of linear instability theory for small amplitude waves propagating through exponentially decaying atmospheric density and the observed zonal mean middle atmosphere winds [*Alexander and Rosenlof*, 1996] (see also section 4). It can thus occur using any of the parameterizations above based on convective instability for some range of source spectrum properties. That includes the Lindzen-type, *Warner and McIntyre* [1996, 2001], and *Alexander and Dunkerton* [1999] parameterizations. It can also arise naturally in the *Hines* [1997a, 1997b] parameterization for wave dissipation because the amplitudes of the waves that drive the dissipation in this mechanism grow with altitude. They can thus shift waves with larger and larger intrinsic phase speeds (and vertical scales) to the prescribed dissipation scale as a function of height. The Hines dissipation scale is further allowed to vary with height adding another means of tuning this effect.

[176] Vertical gradients of gravity wave energy and momentum fluxes give rise to important terms in the energy and momentum conservation equations in global models (see section 8). All of the parameterizations will drive the deceleration of the winds in the mesosphere, but only a few have been shown to be effective at driving QBO circulations in global models. In addition to mean flow forcing effects of parameterized gravity waves, vertical mixing effects in the thermal energy and constituent transport equations are also important in the mesosphere of global models. Global model studies of both these effects are described in section 8.

8. INFLUENCES ON ATMOSPHERIC CIRCULATION AND STRUCTURE

8.1. Mean Flow Forcing Effects

[177] Around 20 years ago, gravity wave effects on the zonal mean winds in the mesosphere and lower thermosphere were becoming widely appreciated. The gravity wave zonal mean forces cause reversals of the zonal mean jets and drive a mean meridional transport circulation that leads to a warm winter mesopause, a cold summer mesopause, and a reversal of the latitudinal temperature gradient expected in the absence of wave driving [*Lindzen*, 1981; *Holton*, 1982, 1983; *Garcia and Solomon*, 1985]. Gravity wave effects at slightly lower altitudes such as those driving the mesopause SAO [*Dunkerton*, 1982b] and QBO [*Lindzen and Holton*, 1968] had been proposed but were not generally accepted. Since that time, gravity wave effects at even lower altitudes, such as mountain wave drag effects on the circulation in the troposphere, have become widely accepted and are now included via parameterization in most global models to provide a drag force on the jet stream [*Palmer et al.*, 1986; *McFarlane*, 1987; *Miller et al.*, 1989]. Most recently, direct gravity wave effects in the stratosphere have been inferred. These effects include contributions to driving the tropical QBO [*Gray and Pyle*, 1989; *Takahashi and Holton*, 1991; *Takahashi and Boville*, 1992; *Alexander and Holton*, 1997; *Dunkerton*,

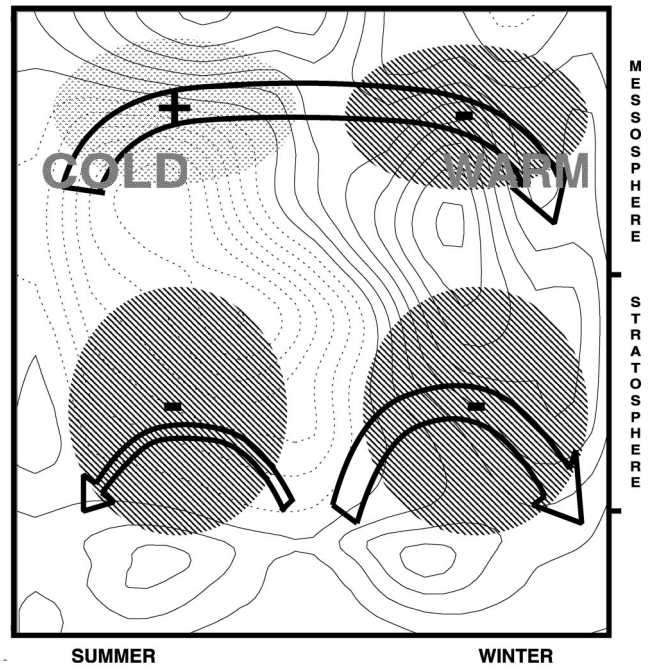


Figure 25. Schematic diagram illustrating gravity wave driving of the middle atmosphere transport circulation and effects on the zonal-mean extratropical winds and temperatures. The thin contours denote zonal mean wind at intervals of 10 m s^{-1} (dotted contours are for westward winds). The transport circulation is illustrated with thick arrows, and the gravity wave driving is illustrated with hatched areas with minus signs denote westward forcing and a plus sign denotes eastward forcing.

1997b; *Scaife et al.*, 2000; *Baldwin et al.*, 2001] and SAO (both in the stratosphere and mesosphere) [*Hitchman and Leovy*, 1988; *Hitchman et al.*, 1992; *Sassi and Garcia*, 1994, 1997; *Ray et al.*, 1998; *Garcia and Sassi*, 1999], contributions to driving the summer hemisphere equator to pole meridional transport circulation [*Rosenlof*, 1996; *Alexander and Rosenlof*, 1996], and effects on the winter polar vortex [*Hitchman et al.*, 1989; *O'Sullivan and Dunkerton*, 1994; *Kinnersley*, 1996; *Garcia and Boville*, 1994]. Figure 25 illustrates these mean-flow forcing effects. The role of waves in the middle atmosphere transport circulation was reviewed by *Holton and Alexander* [2000]. The theory underlying wave driving of the transport circulation is embodied in the downward control principle [*Haynes et al.*, 1991].

[178] In general, in the lower stratosphere the gravity wave mean body force will have the same sign as the local shear (and may or may not have the same sign as the local wind). Examples include the westward drag exerted by mountain waves at levels above the tropospheric jet peak winds and the QBO forcing that is eastward in eastward shear and westward in westward shear. In the mesosphere the force, in general, opposes the background winds. Wave amplitudes can grow large enough in the mesosphere that waves with large intrinsic phase speeds can break, while in the stratosphere, where wave amplitudes are smaller, breaking or dissipation will tend to occur only as waves approach their critical levels (see, e.g., equation (58)) unless the amplitudes are rather large.

[179] Several studies have estimated forces due to gravity waves in the stratosphere and mesosphere. Three basic techniques have been used: (1) Estimates of the zonal mean wave driving in the middle atmosphere, not explained by the observed planetary-scale wave driving, are presumed due to gravity waves. (2) Radar observations of zonal and meridional gravity wave momentum flux gradients employed in equation (42) give an estimate of gravity wave driving. (3) Atmospheric model studies with parameterized wave driving are constrained by observations to infer the contribution due to gravity waves.

8.1.1. Zonal Mean Momentum Balance Studies

[180] The first method estimates the gravity wave contribution to the wave driving through examination of the zonal mean momentum balance. The difference between the total momentum forcing and the resolved EP flux divergence contribution is presumed to be due to gravity wave driving. The total force is calculated using global satellite observations of temperature and the important radiatively active gases (O_3, H_2O, etc.) input to a radiative transfer model to estimate the zonal mean radiative heating rate. This total force is what drives the zonal mean residual circulation (or transport circulation) in the middle atmosphere. Satellite temperature observations are then also used to compute global winds with a balance model, and the temperature and wind fields together are then used to define the planetary-scale wave fluxes and EP flux divergence. Examples of this method are described by *Hamilton* [1983], *Smith and Lyjak* [1985], *Hitchman et al.* [1989], *Shine* [1989], *Marks* [1989], *Rosenlof* [1995; 1996], and *Fetzer and Gille* [1996]. These studies have improved in accuracy over the years, partly because of improvements in the temperature and constituent measurements and partly through improvements in the models used to calculate the wind fields.

[181] Calculations of EP flux divergence give inherently noisy fields, so it is not surprising that most studies looking for the residual gravity wave term have focused on high altitudes near the stratopause and above where the gravity wave forcing becomes large [*Hamilton*, 1983; *Smith and Lyjak*, 1985; *Shine*, 1989; *Marks*, 1989; *Fetzer and Gille*, 1996]. At lower altitudes in the winter stratosphere the calculations of the residual gravity wave forcing have been highly variable, which could result both from errors in the calculations and true variability.

[182] *Rosenlof* [1996] and *Alexander and Rosenlof* [1996] focused on the summer stratosphere where resolved EP flux divergence is small and gravity wave driving is believed to dominate. A simple model of gravity wave propagation and dissipation was found to generate gravity wave driving in the summer strato-

sphere that was consistent with the observational constraints. The model used a constant spectrum of nonstationary gravity waves input in the troposphere. *Ray et al.* [1998] applied the same methods in the equatorial stratophere and inferred that gravity wave driving is important to the eastward phase of the stratopause SAO. The gravity wave driving in the summer stratosphere is westward and below ~1 hPa has an estimated magnitude of order ~1 m s^{-1} d^{-1}. Similar westward forcing magnitudes have also been estimated near the tropopause in both summer and winter. The gravity wave driving at the onset of the stratopause SAO eastward phases (at equinoxes) has been estimated at ~1–3 m s^{-1} d^{-1}.

[183] Calculations of this type in the mesosphere extending up to ~80 km altitude show the annual cycle in extratropical gravity wave driving of the summer-to-winter pole-to-pole zonal mean circulation. The wave driving is westward in winter and eastward in summer [*Shine*, 1989; *Marks* 1989; *Fetzer and Gille*, 1996]. Two of these studies suggested that the mesopheric gravity wave driving in the Southern Hemisphere winter is significantly larger than that in the Northern Hemisphere winter. These estimates of the mesosphere force peak at magnitudes ~30–60 m s^{-1} d^{-1} at solstice seasons. This hemispheric asymmetry seems opposite to that suggested by PMSE and temperature observations described in section 5. The differences could suggest a high degree of variability or could be due to errors in the estimates.

8.1.2. Radar Observations of Momentum Flux Convergence

[184] Radar observations of gravity wave momentum fluxes at a number of sites have been used to estimate the body force due to gravity waves in the lower stratosphere and in the mesosphere and lower thermosphere. Mean momentum fluxes in the lower stratosphere over the MU radar (at 35°N latitude) were ~0.1–0.3 m^2 s^{-2}, with peak values of ~1 m^2 s^{-2} and implied mean flow forcing of ~1 m s^{-1} d^{-1} [*Fritts et al.*, 1990b]. Momentum fluxes in the mesosphere and lower thermosphere have been measured more widely. At extratropical sites, mean values of ~1–10 m^2 s^{-2} have been measured with mean flow forcing in the range ~10–70 m s^{-1} d^{-1} that generally oppose the mean winds [*Reid and Vincent*, 1987; *Fritts and Yuan*, 1989b; *Tsuda et al.*, 1990b; *Nakamura et al.*, 1993c]. At the tropical Jicamarca radar (12°S) [*Hitchman et al.*, 1992] fluxes of 1–4 × 10^{-4} Pa (~2–8 m^2 s^{-2}) were observed with mean flow forcing ~10–60 m s^{-1} d^{-1}. Measurements at 35° northern and southern latitudes have been compared [*Nakamura et al.*, 1996], and like the residual force calculations, these suggest larger gravity wave driving in the Southern Hemisphere winter than in the Northern Hemisphere winter. Momentum fluxes observed in the mesosphere and lower thermosphere are highly variable, with peak values of ~30–60 m^2 s^{-2} greatly exceeding the mean values. These imply correspondingly larger and possibly localized mean flow forcing and have been correlated with larger-scale motions [*Fritts and Vincent*, 1987; *Fritts et al.*, 1992; *Murphy and Vincent*, 1998].

8.1.3. Model Studies of Zonal Mean Wave Driving

[185] A third method for estimating the gravity wave mean flow driving uses models of the middle atmosphere with parameterized radiative transfer and wave driving. The model results are then constrained by observations so that inferences about the gravity wave driving can be made. *Holton* [1982, 1983], *Garcia and Solomon* [1985], *Holton and Schoeberl* [1988], *Hitchman et al.* [1989], and *Huang and Smith* [1991, 1995] examined gravity wave driving at extratropical latitudes in the mesosphere. The gravity wave mean flow force in these models peaked at midlatitudes at solstice seasons with values ranging from ~30 to 120 m s^{-1} d^{-1}, with more recent work tending to report smaller values.

[186] Other global and mechanistic model studies have focused specifically on the magnitude of the zonal mean transport circulation required to account for the summer mesopause thermal structure [*Garcia*, 1989; *McIntyre*, 1989; *Fritts and Luo*, 1995, *Zhu et al.*, 1997]. These have suggested vertical velocities near the mesopause of ~5 cm s^{-1} and meridional motions of ~20–30 m s^{-1}; values that are comparable to radar and satellite measurements of mean meridional winds near the polar summer mesopause. The observed meridional winds differ somewhat in their latitudinal distribution with peak magnitudes observed at higher and lower latitudes [*Nastrom et al.*, 1982; *Lieberman et al.*, 1998] than the model studies' inferred meridional transport circulation winds. Variability in wave driving can also account for abrupt temperature transitions that have been observed at the summer mesopause [*Luo et al.*, 1995; *Lübken*, 1999].

[187] *Dunkerton* [1997b] modeled the QBO and SAO and inferred that gravity waves likely contribute at least half of the wave momentum flux needed to drive the QBO. His estimates of the cross-tropopause momentum flux carried by gravity waves with phase speeds in the range of QBO wind speeds were ~2–3 × 10^{-3} Pa.

8.2. Gravity Wave–Tidal–Planetary Wave Interactions

[188] As discussed in section 2 and further in section 8.4, gravity waves play a central role in controlling the mean circulation and thermal structure through wave momentum transport, wave filtering, and the body forces arising from wave dissipation and momentum flux divergence. Such filtering and effects arise in response to mean shears as well as lower-frequency gravity waves (see section 6.1) and larger-scale tidal and planetary wave motions. The first studies along these lines were theoretical and addressed feedbacks on tidal structures due to selective gravity wave filtering near critical levels [*Walterscheid*, 1981], excitation of planetary waves at altitudes of asymmetric gravity wave body forcing imposed by variable source strengths [*Holton*, 1984], and

damping of planetary waves in response to modulation of gravity wave filtering processes [*Miyahara*, 1985; *Miyahara et al.*, 1986]. More recent model assessments of such interactions have revealed a number of potential influences on tidal and planetary wave structures and variability. Various studies suggest either tidal amplitude reductions [*Forbes et al.*, 1991; *Miyahara and Forbes*, 1991; *Lu and Fritts*, 1993; *Meyer*, 1999a] or increases [*Mayr et al.*, 1998], depending on which gravity wave parameterization scheme is used. However, *Akmaev* [2001] has argued that the *Miyahara and Forbes* [1991] results are due to an incorrect implementation of the Lindzen parameterization, while *McLandress* [2002] argues that gravity wave driving and induced diffusion are much less important in controlling tidal amplitudes than other nonlinear interaction and advection processes.

[189] Several studies considered the implications of tidal modulation of gravity wave-induced diffusion and/or momentum deposition, arguing that these may be important in minor species chemistry and the creation of temperature inversions [*Forbes et al.*, 1991; *Liu et al.*, 2000], in altering mean momentum deposition [*Lu and Fritts*, 1993; *McLandress and Ward*, 1994], and in enhancing planetary wave excitation in the mesosphere in response to spatially localized orographic sources [*McLandress and Ward*, 1994]. Other studies suggested that seasonal variations in mean winds enable a gravity wave modulation of tidal amplitudes that accounts for the observed semiannual variations [*Mayr et al.*, 1998; *Meyer*, 1999a] and that filtering of gravity waves by planetary waves at lower altitudes imposes a periodic forcing near the mesopause which enables 2-day and 16-day wave influences to penetrate well into the thermosphere [*Meyer*, 1999b; *Norton and Thuburn*, 1997].

[190] All of the above studies exhibited complex coupling between the imposed spectrum of gravity waves and the larger-scale motions accounting for wave filtering. In particular, tidal amplitude changes accompanying gravity wave-tidal interactions were found to be sensitive not only to geophysical parameters but also to the specific gravity wave parameterization employed. This is because tidal amplitudes (as well as any other variable wind field accounting for gravity wave filtering, i.e., mean winds, planetary waves, or the semiannual and quasi-biennial oscillations) respond to gravity wave filtering and momentum flux divergence differently depending on the phase of the gravity wave forcing relative to the phase of the filtering wind field, which can differ between parameterizations. If parameterized gravity waves are dissipated near and below the wind maxima (in the direction of gravity wave propagation), then the result is a wind acceleration and an induced downward phase motion (as discussed in section 7.2.2). If, on the other hand, parameterized gravity waves dissipate by breaking (due to amplitude growth with height where the winds are a minimum or where intrinsic phase speeds are large), then gravity wave forcing reduces wind amplitudes and the induced phase motion depends on where forcing occurs relative to the wind maxima. For these reasons, comparisons with observations may pose important constraints on the gravity wave parameterizations underlying such interactions in numerical models.

[191] Observational studies were slower to yield insights into gravity wave interactions with tidal and planetary wave motions. Nevertheless, several studies have shed light on the degree of modulation of gravity wave amplitudes and fluxes and their potential influences on the structures responsible for wave filtering. Evidence of gravity wave-tidal and gravity wave-planetary wave interactions has come primarily from various radars spanning equatorial to high latitudes. Measurements of gravity wave momentum fluxes and tidal amplitudes at Adelaide, Australia, and Poker Flat, Alaska, revealed a tendency for an out-of-phase relationship between fluxes and tidal winds [*Fritts and Vincent*, 1987; *Wang and Fritts*, 1991], a potential for very large momentum flux modulations correlated with large tidal amplitudes and shears (as large as ~ 40 m^2 s^{-2}), and a reduction of tidal amplitudes accompanying strong gravity wave filtering [*Fritts and Vincent*, 1987]. An example of the modulation of gravity wave momentum fluxes by the diurnal tide for 3 days during which the tide was strong is shown in Figure 26. In this example, the flux estimates are 8-hour averages, the modulation increases with altitude, and the peak values are ~ 10 times as large as the mean momentum flux [*Fritts and Vincent*, 1987]. A similar modulation of momentum fluxes by the 2-day wave was suggested by the observations of *Murphy and Vincent* [1998], while *Thayaparan et al.* [1995], *Isler and Fritts* [1996], and *Manson et al.* [1998b] noted significant peaks in gravity wave variances at the dominant tidal and planetary wave periods, suggestive of strong mutual interactions. Finally, indirect evidence of gravity wave filtering by planetary waves and of gravity wave influences on tidal structures has also been obtained. *Smith* [1996] inferred gravity wave filtering by and forcing of planetary waves from out-of-phase zonal wind anomolies at stratospheric and mesospheric altitudes, respectively, while *Nakamura et al.* [1997] inferred gravity wave-tidal interactions from tidal structure and variability at multiple equatorial and subtropical radar sites. There is insufficient data at this stage, however, to be able to distinguish between or guide the development of more quantitative gravity wave parameterizations.

8.3. Turbulent Mixing and Transport

[192] The impact of turbulence arising from gravity wave instability processes is one of the least quantified aspects of gravity wave forcing of the middle atmosphere at present. In fact, there are two separate issues. One is the energy dissipation rate ϵ accompanying gravity wave instability and its spatial and temporal variability; a second is the efficiency of mixing that occurs for a fixed energy dissipation rate. The former is becoming better known as measurements and theoretical inferences are beginning to converge. In the stratosphere, both ϵ and

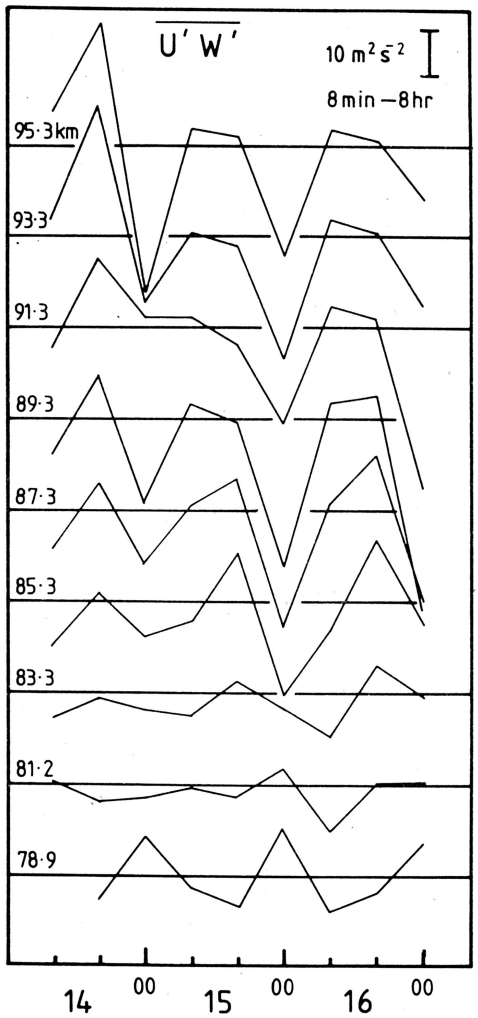


Figure 26. Zonal gravity wave momentum fluxes for 8-hour data segments during 3 days of large diurnal tidal winds over Adelaide, Australia. Note that peak momentum fluxes are ~10 times mean values. After *Fritts and Vincent* [1987] (reprinted with permission of the American Meteorological Society).

turbulent mixing and transport are relatively small and have minor impact on the large-scale thermal and constituent structures. Because, however, gravity wave energy flux and energy dissipation rate increase rapidly with altitude (see equations (51) and (52)), their influences are greater in the mesosphere and lower thermosphere. Early estimates of ϵ at higher altitudes based on radar spectral width and rocket trail diffusion measurements tended to be very high, but more recent estimates are typically ~10–300 mW kg^{-1} at altitudes from ~80 to 110 km [*CIRA*, 1986; *Hocking*, 1990]. More recent in situ measurements have refined these estimates even further, lowering mean ϵ estimates at high latitudes in winter below previous values and suggesting a sharp peak in ϵ at the high-latitude summer mesopause [*Blix et al.*, 1990; *Lübken et al.*, 1993; *Lübken*, 1997]. Peak values in these estimates are ~200 mW kg^{-1} and are in reasonable agreement with earlier CIRA 86 estimates. Peak values are also in reasonable agreement with recent theoretical estimates based on mean gravity wave energy densities [*Fritts and VanZandt*, 1993; *Hall and Hoppe*, 1998] and with estimates based on the gravity wave forcing required to account for the summer mesopause thermal structure [*Fritts and Luo*, 1995]. Comparable estimates have arisen from in situ measurements at equatorial latitudes [*Goldberg et al.*, 1997], while *Fuller-Rowell* [1994] attributed observed latitudinal variations in nitric oxide to a latitudinally varying turbulent transport with a maximum at midlatitudes.

[193] As noted in section 8.3, several studies have anticipated more efficient turbulent transport of momentum than of heat and constituents based on localization of turbulence within the wave field [*Fritts and Dunkerton*, 1985; *Coy and Fritts*, 1988; *McIntyre*, 1989; *Gavrilov and Yudin*, 1992] and assessments of the turbulent transport required to account for observed thermal and constituent profiles [*Strobel et al.*, 1985, 1987; *Holton and Schoeberl*, 1988; *Smith and Brasseur*, 1991]. More efficient turbulent transport of momentum than of heat implies a turbulent Prandtl number significantly larger than unity, with expected values in the range $Pr \sim 3$–10. The implication of a large Pr is a significantly reduced turbulent diffusivity, K_{zz}, relative to that anticipated from energy dissipation rates inferred from either discrete wave or spectral descriptions of gravity wave dissipation [*Lindzen*, 1981; *Fritts and VanZandt*, 1993]. These gravity wave dissipation studies together with more recent numerical modeling of wave breaking and turbulence dynamics [*Fritts et al.*, 2003] suggest that the turbulent Prandtl number will depend on wave amplitude and intrinsic frequency, with more efficient mixing accompanying larger wave amplitudes and more vigorous turbulence. We also anticipate that turbulent mixing accompanying KH instability of the wave field may imply smaller Pr since in this case strong mixing occurs in a sheared and stratified environment. Our understanding of the details of turbulence generation and mixing is insufficient, however, to be able to quantify middle atmosphere effects more completely at this time.

8.4. General Circulation Model Studies

[194] Several studies have used high-resolution GCMs to examine the properties of gravity waves resolved in these models and to examine their effects on the middle atmosphere circulation. Experiments with the SKYHI GCM [*Hamilton*, 1995, 1996; *Jones et al.*, 1997] have shown a tendency toward better simulation of the middle atmosphere circulation with increasing latitude and longitude resolution from 3° × 3.6° to 0.6° × 0.72°. The spectrum of wave momentum flux versus zonal wave number in these simulations is rather flat, suggesting that even wave numbers beyond 140 would need to be resolved to capture the wave fluxes important to the middle atmosphere circulation. The conclusion of these studies is that gravity wave effects in middle atmosphere GCMs must be included via parameterization at least for the foreseeable future.

[195] A number of studies have applied parameterizations of gravity wave effects in GCMs in order to study the role of gravity waves in various phenomena important to middle atmosphere dynamics. The results of such experiments are often difficult to interpret because they are complex in that they include many parameterized processes and modes of interaction. Insights gained from such experiments are described in this section.

8.4.1. Mesospheric Circulation

[196] Every gravity wave parameterization in use has been designed to yield the summer eastward and winter westward momentum forcing in the mesosphere that drives the pole-to-pole residual circulation (section 8.1). This role for gravity waves in the mesosphere was recognized decades ago [*Lindzen*, 1973, 1981; *Holton*, 1982] and has more recently been validated by full GCM studies [*Roble and Ridley*, 1994; *Norton and Thuburn*, 1996; *Hamilton*, 1997; *Manzini and McFarlane*, 1998].

[197] Two recent studies considered the effects of asymmetries in the eastward and westward gravity wave momentum fluxes input via parameterization but came to opposite conclusions. *Manzini and McFarlane* [1998] launched an isotropic spectrum of waves at two different source levels (at the ground and near the tropopause) in the MA/ECHAM4 model using the *Hines* [1997a, 1997b] parameterization. Launching the waves at the ground led to smaller eastward and larger westward momentum flux crossing the tropopause at midlatitudes because of filtering of eastward waves by the tropospheric jets (Figure 27). The larger westward fluxes in turn produced more realistic middle atmosphere zonal mean winds. *Medvedev et al.* [1998] applied their gravity wave parameterization [*Medvedev and Klaassen*, 1995] in the CMAM model. They compared results with an isotropic tropopause launch spectrum, with phase speeds in the range $0 < c < 60$ m s^{-1} in four azimuths, to results with only the eastward propagating waves with larger fluxes. The results with eastward waves alone showed weaker and more realistic zonal mean winds in both summer and winter (Figure 28). This is an apparently paradoxical result since the westward forcing needed to decrease the winter jet strength would be expected to decrease in the anisotropic case given the physics embodied in equation (42).

[198] *Norton and Thuburn* [1996, 1997] applied a Lindzen-type parameterization of gravity wave forcing in the UGAMP GCM. They found the model with gravity wave forcing developed a realistic 2-day wave while no 2-day wave developed in the model with only Rayleigh friction. They concluded that gravity wave forcing in their model was essential to maintaining the unstable zonal mean state that led to the 2-day wave generation.

[199] Several studies have employed mechanistic versions of GCMs to study the effects of various gravity wave parameterization schemes in the mesosphere. A comparison of the effects of the *Hines* [1997a, 1997b] and *Fritts and Lu* [1993] parameterizations on the diurnal tidal structure suggested that the Hines scheme tended to amplify tidal amplitudes, while the Fritts and Lu scheme caused excessive dissipation of the tides [*McLandress*, 1998]. Other mechanistic and simplified

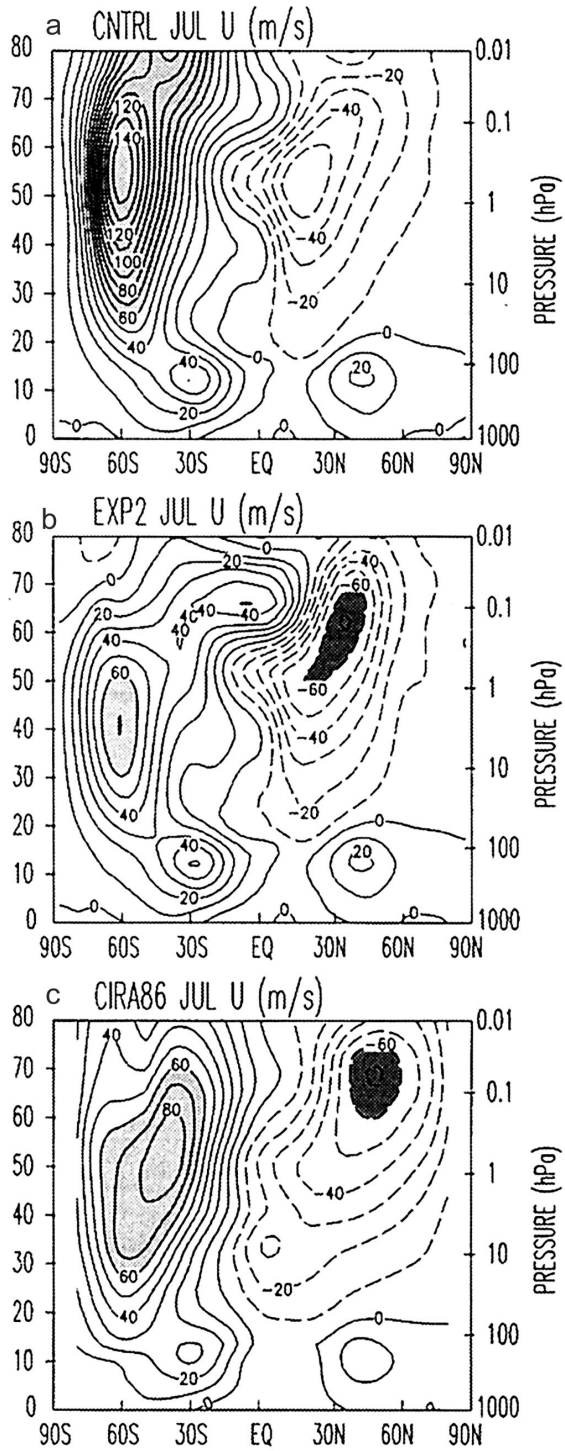


Figure 27. July zonal mean winds resulting from the ECHAM4 GCM with (a) approximately isotropic gravity wave fluxes input at the tropopause and (b) anisotropic gravity wave fluxes at the tropopause with smaller eastward gravity wave fluxes at midlatitudes. (c) The climatology from CIRA [*Fleming et al.*, 1988]. After *Manzini and McFarlane* [1998].

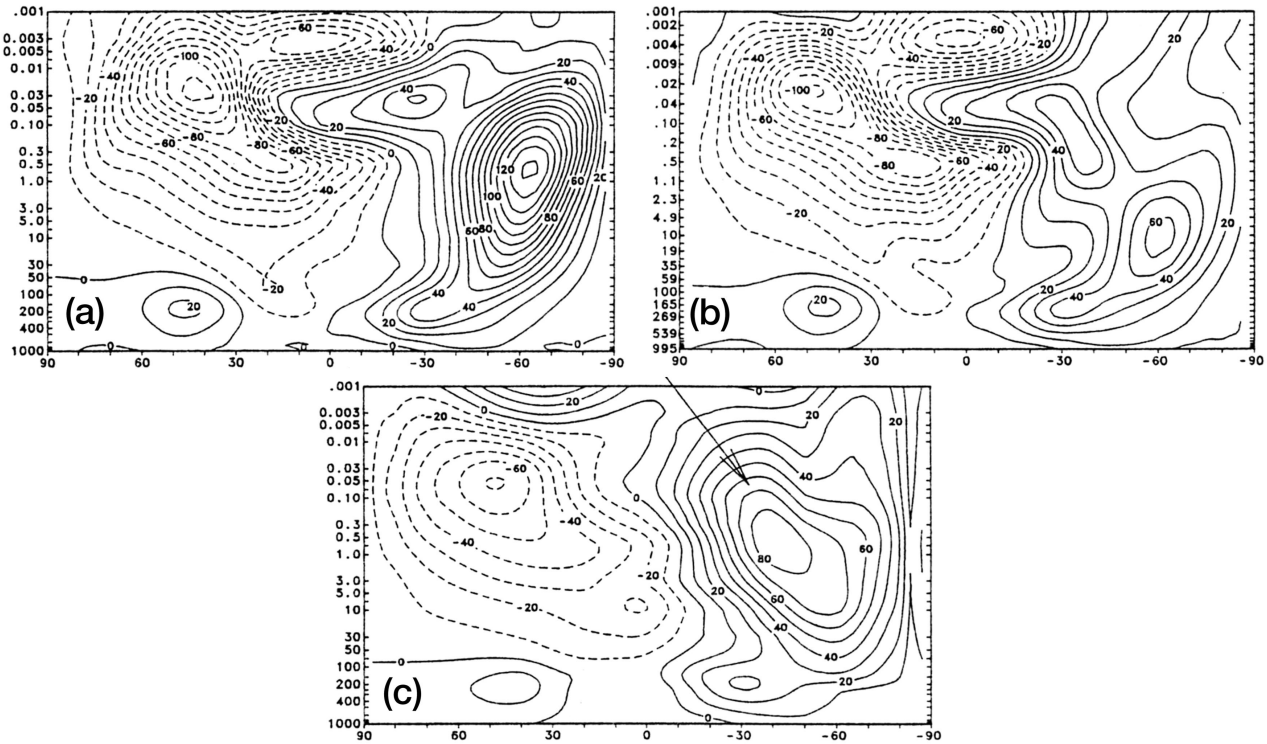


Figure 28. July zonal mean winds resulting from the CMAM GCM with (a) isotropic gravity wave sources input near the tropopause and (b) anisotropic gravity wave sources at the tropopause with larger eastward fluxes at all latitudes. (c) The climatology from CIRA [*Fleming et al.*, 1988]. After *Medvedev et al.* [1998].

GCMs have been used to test gravity wave parameterization effects in the mesosphere [*Hamilton*, 1997].

8.4.2. High-Latitude Stratosphere

[200] Through the concept of "downward control" of the mean meridional circulation by wave-induced mean zonal forces [*Haynes et al.*, 1991], *Garcia and Boville* [1994] stress that the winter mesospheric gravity wave forcing can have effects down to ~30 km in the high-latitude stratosphere that can help to alleviate the "cold pole" problem that plagues GCM studies of the winter stratosphere. *Hamilton* [1997], however, also suggests that gravity wave forcing at stratospheric altitudes, although smaller in magnitude, is also important to the winter circulation.

[201] The role of gravity wave forcing in sudden stratosphere warming (SSW) events has been examined in several global model studies. *Lawrence* [1997] compared effects of Hines and Fritts and Lu parameterization on the winter stratosphere. Both parameterizations generated reasonable zonal mean mesosphere winds, but the Fritts and Lu parameterization was found to inhibit the occurrence of SSW events while Hines did not. *Pawson* [1997] found that topographic gravity wave drag in the lower stratosphere influenced the onset of SSW events but found little or no sensitivity to mesospheric drag parameterized as Rayleigh friction. *Rind et al.* [1988] found that gravity wave forcing experiments, designed to describe waves generated by different sources, affected both the character and existence of SSW events in the GISS GCM.

8.4.3. Equatorial Oscillations

[202] At the equator, where the Coriolis force vanishes, the atmospheric response to a zonal force is simply zonal acceleration. Dissipation of zonally propagating waves at tropical latitudes provides the forcing that drives or helps to drive the three dominant equatorial oscillations: (1) the QBO in the lower stratosphere, (2) the stratopause SAO, and (3) the mesopause SAO. Both gravity waves and planetary-scale waves likely participate in driving all three of these oscillations to differing degrees. What remains to be quantified is the spectrum of wave momentum flux in the tropics. Global model studies are therefore free to include adjustable parameterized gravity waves without adequate constraints either on the gravity wave fluxes or on the resolved planetary-scale wave fluxes. *Mayr et al.* [1997b] demonstrated the potential of the gravity-wave-driving mechanism for all three of these equatorial oscillations in a simplified global model with parameterized gravity wave fluxes.

[203] The QBO in the lower stratosphere zonal winds is a wave-driven phenomenon believed to be predominantly driven by dissipation of planetary-scale waves [*Baldwin et al.*, 2001]. However, zonal forcing accompanying the dissipation of gravity waves is likely to contribute significantly, with recent estimates suggesting they may transport roughly half of the flux [*Dunkerton*, 1997;

Piani et al., 2000]. Until very recently, obtaining a QBO in GCM studies had proved difficult. Recent GCM studies with increased horizontal and vertical resolution and decreased numerical diffusion have since developed QBO-like oscillations [*Takahashi*, 1996; 1999; *Horinouchi and Yoden*, 1998; *Hamilton et al.*, 1999]. These successes are believed to be due to the increase in tropical wave fluxes allowed by the finer resolution and reduced numerical dissipation [*Boville and Randel*, 1992; *Nissen et al.*, 2000]. *Scaife et al.* [2000] instead modeled a QBO in a GCM by including parameterized small-scale gravity wave fluxes.

[204] The SAO at the stratopause is driven in part by the meridional advection of summer hemisphere westward (easterly) winds across the equator and in part due to zonal wave driving particularly important during the eastward (westerly) wind phases of the SAO [*Sassi et al.*, 1993; *Hamilton et al.*, 1995; *Müller et al.*, 1997; *Garcia et al.*, 1997]. *Rind et al.* [1988], *Jackson and Gray* [1994], and *Medvedev and Klaasen* [2001] modeled realistic SAOs by including parameterized gravity wave fluxes, the latter with a model deep enough to include an SAO at the mesopause.

9. SUMMARY, NEEDS, AND RECOMMENDATIONS

[205] Our intent in this review was to provide a broad overview of the current understanding of gravity wave dynamics and effects in the middle atmosphere. Considerable research activity in this area over the past two decades made this a daunting undertaking. In choosing to be comprehensive in scope we have necessarily provided less discussion and fewer references in many areas than we would have preferred, emphasizing the more significant contributions and directing readers, where appropriate, to related reviews having more limited foci.

[206] We began our review with a derivation of the dispersion and polarization relations using linear theory and with a discussion of the simplifications and dynamics appropriate for various ranges of intrinsic frequency. These offer, in most cases, significant insights into wave structure and effects at a small cost in accuracy. Also noted were departures from the expectations of linear theory accompanying wave transience, spatial localization, or finite amplitude effects that may have important implications for the middle atmosphere. Particularly relevant here are self-acceleration effects that enable wave packets to penetrate beyond the turning and critical levels anticipated by linear theory, with corresponding implications for wave dissipation and momentum transport. The impact of such effects in the middle atmosphere is not yet known, but studies addressing these effects are underway.

[207] We then summarized what is known of the more important sources of gravity waves, dominant wave scales, and factors influencing wave propagation in variable environments. Theoretical, modeling, and observational studies have yielded considerable advances in quantifying the mechanisms of and responses to wave forcing due to topography, convection, wind shear, geostrophic adjustment, body forcing, and wave–wave interaction sources. Mountain waves occur on scales dictated by the topography, may propagate to high altitudes, and have peak momentum fluxes at horizontal scales of ~10–100 km. Convection excites gravity waves having a wide range of spatial scales and phase speeds via several mechanisms that are dependent on the cloud environment and which exhibit significant temporal variability. Wind shear appears to support envelope radiation of gravity waves on the scale of patches of Kelvin–Helmholtz instability with attributes that favor their penetration to greater altitudes, while geostrophic adjustment favors inertia-gravity wave radiation due to the slow timescale and large spatial scale of the adjustment process. Body forcing due to local gravity wave dissipation and momentum flux divergence, in contrast, may occur on fast or slow timescales and on many spatial scales, resulting in a spectrum of radiated waves that mirror the source characteristics. Finally, wave–wave interactions are operative throughout the atmosphere, with interaction rates increasing with wave amplitude and energy transfers toward both larger and smaller scales. Dominant scales, amplitudes, and phase speeds are dictated largely by the various gravity wave sources at lower altitudes but are increasingly determined by propagation effects, filtering, and wave–wave interactions at greater altitudes. As altitude increases, so do the dominant vertical scales because wave amplitudes at smaller vertical scales are constrained by saturation processes whereas amplitudes at larger vertical scals are not. Despite significant advances, there remain many aspects of gravity wave sources and propagation that require further quantification. These are delineated below.

[208] Broad consistency in spectral shape and evolution with altitude led naturally to a spectral description of gravity wave motions. While far from the universal spectrum advocated initially, the frequency and vertical wave number spectra of horizontal velocity and temperature have near-universal forms throughout the middle atmosphere. Such descriptions constrain wave amplitudes, fluxes, and parameterizations and offer insights arising from spectral character and variability (or lack thereof). We also noted, however, that spectra often conceal, rather than reveal, the dynamics underlying spectral shape and pointed out where caution is warranted. Present unknowns include the degree to which adherence to or departures from specific spectral shapes are indicative of specific dynamics (e.g., saturation processes) or whether other spectral descriptions (e.g., wavelet spectra) might be more enlightening, given the localized nature of gravity waves in space and time that is often observed.

[209] Climatologies of gravity wave velocity and temperature perturbations have been compiled; however, these do not yet adequately constrain the inputs to

gravity wave parameterizations for global models. What is needed are descriptions of gravity wave sources and the detailed properties of the waves generated. The gravity wave climatologies instead have primarily served to partially validate the use of linear models for describing gravity wave propagation through the middle atmosphere and interaction with the background atmosphere with height. Departures from linear theory have been treated in these models with dissipation that is tied to linear instability or saturation of wave amplitudes. These treatments, of course, have serious limitations that are underscored by the more detailed studies of wave instability and mean flow interactions summarized below. The linear theory, however, remains a valuable research tool and is the starting point for most gravity wave parameterization approaches.

[210] Patterns observed in gravity wave climatologies are also influenced by the range of gravity wave properties each observation method is capable of detecting. The climatology of gravity waves that is ultimately needed will have to be constructed from a variety of measurement techniques, with modeling tools employed to aid in the interpretation of the observations.

[211] We next surveyed recent work addressing wave–wave interaction, wave instability, and wave saturation processes. These studies have yielded dramatic advances in our understanding of the myriad types of instabilities to which gravity waves are susceptible. Notable findings include the identification of wave–wave interactions that are facilitated by mean shear, the occurrence of wave–wave interactions (at large wave amplitudes) on a timescale competitive with local wave field instability, and characterization of the spectral energy transfers due to wave–wave interactions. Other studies identified links between instabilities occurring at small and large wave amplitudes, instability processes due to finite amplitude wave packets via wave–mean flow interactions, and definition of the instability dynamics leading to and occurring within a turbulence cascade. Despite these advances the competition between wave–wave interactions and local instability dynamics and their respective control over spectral shape and amplitude remain major unknowns.

[212] The last two decades also saw the development of several classes of saturation theories and a number of attempts to distinguish the most relevant dynamics. The major contenders are linear saturation theory (with amplitude constraints due to local instability), various nonlinear diffusion theories (relying on wave–wave interactions for spectral transfers and wave amplitude limits), Doppler spread theory (describing nonlinear interactions as advection by other spectral members, assuming they all occupy a common volume in physical space), and saturated cascade theory (employing elements of local instability and wave–wave interactions). Deciding which theories are most realistic has proven difficult, however, because all yield estimates of spectral amplitude and shape that are arguably consistent with the various measurements available. Perhaps more importantly, all of these theories are at present incomplete because they all fail to incorporate quantitatively the two nonlinear effects known to occur, i.e., wave–wave interactions and local wave field instability. As such, it is likely more valuable to recognize the insights we have gained from each rather than attempt to identify the "most correct" among deficient or incomplete alternatives. The result is a recognition of the dual role of nonlinearity, via both wave–wave interactions and local instability, in causing wave saturation and in determining the spectral shapes that arise from these dynamics. Thus a complete saturation theory will almost assuredly not be analytic but will rely instead on numerical studies of the competition among or merging effects of dynamics that are now believed to be linked theoretically.

[213] Parameterizations of middle atmosphere gravity wave effects presently suffer acutely from the lack of constraints on their inputs and tunable parameters. Global modelers can use these poorly constrained parameterizations, given the wide range of free parameters available, to get almost any answer they require. Until the needed constraints on parameterization inputs are forthcoming, the differences between different methods of parameterizing gravity wave dissipation and interaction effects will be difficult or impossible to quantify. The process of tuning gravity wave parameterizations in global models to attempt to infer the properties of the gravity wave input parameters is unwise at present because the range of parameters is too broad, and the solutions are not unique. As a result, GCMs tuned for one result (and at one resolution) using a gravity wave parameterization cannot confidently be used in any prognostic way.

[214] There has been significant progress in quantifying gravity wave–mean flow effects over the last two decades. The important mechanisms of wave dissipation are, however, still being debated and likely vary with altitude through the middle atmosphere. In the case of wave breaking, the mixing effects appear to be highly sensitive to the detailed properties of the wave at the point of breakdown. Tidal and planetary wave interactions likewise prove to be highly sensitive to details of the gravity wave field that are not known. To parameterize these effects therefore requires much more extensive knowledge of the properties of gravity waves in the middle atmosphere and of their sources than was imagined 20 years ago.

[215] Gravity waves almost certainly play a role in driving the prominent QBO and SAO equatorial wind oscillations. Quantifying that role remains illusive, however, until more complete knowledge is attained of the momentum flux spectrum of tropical wave motions ranging from planetary-scale to small-scale gravity waves. Similar lack of quantification of gravity wave fluxes and other gravity wave properties at high latitudes also hinders our ability to quantify their role in extratropical

processes such as sudden stratospheric warmings and the wave-driven transport circulation.

[216] Gravity wave interactions with larger-scale waves are strong and have been studied numerically and observationally. However, the details of the interactions appear to be highly sensitive to the relative phase of the gravity wave dissipation and the larger-scale wind variations with height. These remain poorly understood, as is evident from the fact that different gravity wave parameterizations yield very different results.

[217] Our field has experienced dramatic advances in understanding the dynamics and effects of middle atmosphere gravity waves over the last two decades. It is also clear that we require further advances in a number of areas before we can claim a comprehensive understanding of gravity wave dynamics and an ability to account for these dynamics statistically in large-scale operational or research models. We offer here our view of the further advances that are most needed to achieve these goals.

1. Observations and modeling studies that characterize and quantify the dominant gravity wave sources, especially scales, phase speeds, and momentum fluxes, as functions of meteorological conditions are a high priority, as this information underlies any attempt to quantify gravity wave influences at greater altitudes. More detailed observations and numerical studies are also required to quantify wave intermittency imposed by variable sources, propagation conditions, etc., as intermittency has major implications for wave effects at greater altitudes, including instability processes, wave–wave and wave–mean flow interactions, and radiation of secondary waves.

2. Further studies are also required to define the interplay of and competition among the nonlinear processes controlling spectral evolution with altitude in various environments, as these control gravity wave propagation, dissipation, and forcing at greater altitudes. These advances are likely to require numerical studies, as current and anticipated observational capabilities are unlikely to permit complete specification of the motion fields and their evolution in three dimensions. For the same reasons, numerical studies are required to characterize fully the mixing dynamics and transport accompanying the various wave instability processes leading to turbulence, including the statistical effects of multiple waves and events.

3. Finally, the accurate parameterization of gravity wave effects remains a critical need, as these may not ever be adequately described directly in large-scale models. New parameterizations will need to describe gravity wave sources, propagation and filtering, wave–wave (including tidal and planetary wave) interactions, spectral evolution, mean and variable climatologies, and instability and mixing far more quantitatively than present schemes. It seems unlikely, in fact, that any single scheme will satisfy all of these needs and that different schemes will be required to characterize different dynamics. Importantly, advanced parameterizations will rely to a large degree on the needed observational, numerical, and theoretical advances noted above and will advance together with successes in these areas.

[218] Recent theoretical and numerical activities have clarified key concepts, defined important dynamics and effects, and highlighted unresolved issues. Atmospheric and laboratory measurements have yielded similar advances in quantitative measurement techniques and definition of gravity wave sources, propagation characteristics, climatologies, instability processes and dissipation, mean and large-scale wave forcing, and variability. Ongoing and future research promises additional advances and further quantification of gravity wave dynamics and effects. Of the major needs identified above, there are none that seem insurmountable, given a continuation of our successes to date and the increasing awareness of the importance of the field, and we look forward to vigorous and stimulating interactions in achieving our collective objectives in the future.

[219] **ACKNOWLEDGMENTS.** D.C.F.'s review effort was supported by the National Science Foundation under grants ATM-9618004, ATM-9708633, ATM-9816160, the Air Force Office of Scientific Research under contracts F49620-98-C-0029 and F49620-00-C-0008, and the National Aeronautics and Space Administration under contracts S-46579-G and NAG5-02036. Support for M.J.A. came from the National Science Foundation grant ATM-9907501 and NASA contract NASW-01017. The authors would like to thank Jim Holton, Bob Vincent, and two anonymous reviewers for valuable comments on the first draft.

REFERENCES

Adrian, R. J., C. D. Meinhart, and C. D. Tomkins, Vortex organization in the outer region of the turbulent boundary layer, *J. Fluid Mech.*, *422*, 1–54, 2000.

Akmaev, R. A., Simulation of large-scale dynamics in the mesosphere and lower thermosphere with the Doppler-spread parameterization of gravity waves, 2, Eddy mixing and the diurnal tide, *J. Geophys. Res.*, *106*, 1205–1213, 2001.

Alexander, M. J., A simulated spectrum of convectively generated gravity waves: Propagation from the tropopause to the mesopause and effects on the middle Atmosphere, *J. Geophys. Res.*, *101*, 1571–1588, 1996.

Alexander, M. J., Interpretations of observed climatological patterns in stratospheric gavity wave variance, *J. Geophys. Res.*, *103*, 8627–8640, 1998.

Alexander, M. J., and T. J. Dunkerton, A spectral parameterization of mean-flow forcing due to breaking gravity waves, *J. Atmos. Sci.*, *56*, 4167–4182, 1999.

Alexander, M. J., and J. R. Holton, A model study of zonal forcing in the equatorial stratosphere by convectively induced gravity waves, *J. Atmos. Sci.*, *54*, 408–419, 1997.

Alexander, M. J., and L. Pfister, Gravity wave momentum flux in the lower stratosphere over convection, *Geophys. Res. Lett.*, *22*, 2029–2032, 1995.

Alexander, M. J., and K. H. Rosenlof, Nonstationary gravity wave forcing of the stratospheric zonal mean wind, *J. Geophys. Res.*, *101*, 23,465–23,474, 1996.

Alexander, M. J., and R. A. Vincent, Gravity waves in the

tropical lower stratosphere: A model study of seasonal and interannual variability, *J. Geophys. Res.*, *105*, 22,299–22,310, 2000.

Alexander, M. J., J. R. Holton, and D. R. Durran, The gravity wave response above deep convection in a squall line simulation, *J. Atmos. Sci.*, *52*, 2212–2226, 1995.

Alexander, M. J., J. H. Beres, and L. Pfister, Tropical stratospheric gravity wave activity and relationship to clouds, *J. Geophys. Res.*, *105*, 22,299–22,309, 2000.

Alexander, M. J., T. Tsuda, and R. A. Vincent, Latitudinal variations observed in gravity waves with short vertical wavelengths, *J. Atmos. Sci.*, *59*, 1394–1404, 2002.

Allen, K. R., and R. I. Joseph, A canonical statistical theory of oceanic internal waves, *J. Fluid Mech.*, *204*, 185–228, 1989.

Allen, S. J., and R. A. Vincent, Gravity wave activity in the lower atmosphere: Seasonal and latitudinal variations, *J. Geophys. Res*, *100*, 1327–1350, 1995.

Andreassen, Ø., C.-E. Wasberg, D. C. Fritts, and J. R. Isler, Gravity wave breaking in two and three dimensions, 1, Model description and comparison of two-dimensional evolutions, *J. Geophys. Res.*, *99*, 8095–8108, 1994.

Andreassen, Ø., P. Ø. Hvidsten, D. C. Fritts, and S. Arendt, Vorticity dynamics in a breaking gravity wave, 1, Initial instability evolution, *J. Fluid Mech.*, *367*, 27–46, 1998.

Andrews, D. G., and M. E. McIntyre, Planetary waves in horizontal and vertical shear: The generalized Eliassen-Palm relation and the mean zonal acceleration, *J. Atmos. Sci.*, *33*, 2031–2048, 1976.

Arendt, S., D. C. Fritts, and Ø. Andreassen, The initial value problem for Kelvin vortex waves, *J. Fluid Mech.*, *344*, 181–212, 1997.

Arendt, S., D. C. Fritts, and Ø. Andreassen, Kelvin twist waves in the transition to turbulence, *Eur. J. Mech.*, *17*, 595–604, 1998.

Bacmeister, J. T., Mountain-wave drag in the stratosphere and mesosphere inferred from observed winds and a simple mountain-wave parameterization scheme, *J. Atmos. Sci.*, *50*, 377–399, 1993.

Bacmeister, J. T., S. D. Eckermann, P. A. Newman, L. Lait, K. R. Chan, M. Loewenstein, M. H. Proffitt, and B. L. Gary, Stratospheric horizontal wavenumber spectra of winds, potential temperature, and atmospheric tracers observed by high-altitude aircraft, *J. Geophys. Res.*, *101*, 9441–9470, 1996.

Bacmeister, J. T., M. R. Schoeberl, L. R. Lait, P. A. Newman, and B. L. Gary, Small-scale waves encountered during AASE, *Geophys. Res. Lett.*, *17*, 349–352, 1990a.

Bacmeister, J. T., M. R. Schoeberl, L. R. Lait, P. A. Newman, and B. L. Gary, ER-2 mountain wave encounter over Antarctica: Evidence for blocking, *Geophys. Res. Lett.*, *17*, 81–84, 1990b.

Baldwin, M. P., et al., The quasi-biennial oscillation, *Rev. Geophys.*, *39*, 179–229, 2001.

Balsley, B. B., and D. A. Carter, The spectrum of atmospheric velocity fluctuations at 8 and 86 km, *Geophys. Res. Lett.*, *9*, 465–468, 1982.

Balsley, B. B., and R. Garello, The kinetic energy density in the troposphere, stratosphere and mesosphere: A preliminary study using the Poker Flat radar in Alaska, *Radio Sci.*, *20*, 1355–1362, 1985.

Balsley, B. B., R. F. Woodman, M. Sarango, R. Rodriguez, J. Urbina, E. Ragaini, J. Carey, M. Huaman, and A. Giraldez, On the lack of Southern Hemisphere polar mesosphere summer echoes, *J. Geophys. Res.*, *100*, 11,685–11,693, 1995.

Barat, J., and C. Cot, Wind shear rotary spectra in the atmosphere, *Geophys. Res. Lett.*, *19*, 103–106, 1992.

Bergman, J. W., and M. L. Salby, Equatorial wave activity derived from fluctuations in observed convection, *J. Atmos. Sci.*, *51*, 3791–3806, 1994.

Blix, T. A., E. V. Thrane, D. C. Fritts, U. von Zahn, F.-J. Lübken, W. Hillert, S. P. Blood, J. D. Mitchell, G. A. Kokin, and S. V. Pakhomov, Small-scale structure observed in-situ during MAC/EPSILON, *J. Atmos. Terres. Phys.*, *52*, 835–854, 1990.

Boville, B. A., and W. J. Randel, Equatorial waves in a stratospheric GCM: Effects of vertical resolution, *J. Atmos. Sci.*, *49*, 785–801, 1992.

Broad, A. S., Linear theory of momentum fluxes in 3-D flows with turning of the mean wind with height, *Q. J. R. Meteorol. Soc.*, *121*, 1891–1902, 1995.

Broad, A. S., Do orographic gravity waves break in flows with uniform wind direction turning with height?, *Q. J. R. Meteorol. Soc.*, *125*, 1695–1714, 1999.

Broutman, D., On internal wave caustics, *J. Phys. Oceanogr.*, *16*, 1625–1635, 1986.

Broutman, D., and R. Grimshaw, The energetics of the interaction between short small-amplitude internal waves and inertial waves, *J. Fluid Mech*, *196*, 93–106, 1988.

Broutman, D., and W. R. Young, On the interaction of small-scale internal waves with near-inertial waves, *J. Fluid Mech.*, *166*, 341–358, 1986.

Broutman, D., C. MaCaskill, M. E. McIntyre, and J. W. Rottman, On Doppler-spreading models of internal waves, *Geophys. Res. Lett.*, *24*, 2813–2816, 1997.

Bruhwiler, D. L., and T. J. Kaper, Wavenumber transport: Scattering of small-scale internal waves by large-scale wavepackets, *J. Fluid Mech.*, *289*, 379–405, 1995.

Buckley, G., D. Broutman, J. W. Rottman, and S. Eckermann, On the importance of weak steady shear in the refraction of short internal waves, *Geophys. Res. Lett.*, *26*, 2877–2880, 1999.

Bühler, O., and M. E. McIntyre, On shear-generated gravity waves that reach the mesosphere, part II, Wave propagation, *J. Atmos. Sci.*, *56*, 3764–3773, 1999.

Bühler, O., M. E. McIntyre, and J. F. Scinocca, On shear-generated gravity waves that reach the mesosphere, part I, Wave generation, *J. Atmos. Sci.*, *56*, 3749–3763, 1999.

Cadet, D., and H. Teitelbaum, Observational evidence of internal inertia-gravity waves in the tropical stratosphere, *J. Atmos. Sci.*, *36*, 892–907, 1979.

Cadot, O., S. Douady, and Y. Couder, Characterization of the low-pressure filaments in a three-dimensional turbulent shear flow, *Phys. Fluids*, *7*, 630–646, 1995.

Caulfield, C. P., and W. R. Peltier, Three dimensionalization of the stratified mixing layer, *Phys. Fluids*, *6*, 3803–3805, 1994.

Chan, K. R., L. Pfister, T. P. Buil, S. W. Bowen, J. Dean-Day, B. L. Gary, D. W. Fahey, K. K. Kelly, C. R. Webster, and R. D. May, A case study of the mountain lee wave event of January 6, 1992, *Geophys. Res. Lett.*, *20*, 2551–2554, 1993.

Chao, W. C., and M. R. Schoeberl, On the linear approximation of gravity wave saturation in the mesosphere, *J. Atmos. Sci.*, *41*, 1893–1898, 1984.

Chimonas, G., and J. R. Grant, Shear excitation of gravity waves, part II, Upscale scattering from Kelvin-Helmholtz waves, *J. Atmos. Sci.*, *41*, 2278–2288, 1984.

Chimonas, G., and C. O. Hines, Doppler ducting of atmospheric gravity waves, *J. Geophys. Res.*, *91*, 1219–1230, 1986.

Chun, H.-Y., and J.-J. Baik, Momentum flux by thermally induced internal gravity waves and its approximation for large-scale models, *J. Atmos. Sci.*, *55*, 3299–3310, 1998.

Chunchuzov, I. P., On the high-wavenumber form of the Eulerian internal wave spectrum in the atmosphere, *J. Atmos. Sci.*, *59*, 1753–1774, 2002.

Clark, T. L., T. Hauf, and J. P. Kuettner, Convectively forced internal gravity waves: Results from two-dimensional numerical experiments, *Q. J. R. Meteorol. Soc.*, *112*, 899–925, 1986.

Collins, R. L., A. Nomura, and C. S. Gardner, Gravity waves in

the upper mesosphere over Antarctica: Lidar observations at the South Pole and Syowa, *J. Geophys. Res.*, *99*, 5475–5485, 1994.

Cot, C. E., and J. Barat, Wave-turbulence interaction in the stratosphere: A case study, *J. Geophys. Res.*, *91*, 2749–2756, 1986.

Cot, C. E., and J. Barat, A "universal" wave spectrum for atmospheric temperature and velocity fluctuations in the stratosphere?, *Geophys. Res. Lett.*, *17*, 1577–1580, 1990.

Coulman, C. E., J. Vernin, and A. Fuchs, Optical seeing—Mechanism of formation of thin turbulent laminae in the atmosphere, *Appl. Opt.*, *34*, 5461–5474, 1995.

Coy, L., and D. C. Fritts, Gravity wave heat fluxes: A Lagrangian approach, *J. Atmos. Sci.*, *45*, 1770–1780, 1988.

Davis, P. A., and W. R. Peltier, Some characteristics of the Kelvin-Helmholtz and resonant overreflection modes of shear instability and of their interaction through vortex pairing, *J. Atmos. Sci.*, *36*, 2394–2412, 1979.

de la Torre, A., P. Alexander, and A. Giraldez, The kinetic to potential energy ratio and spectral separability from high-resolution balloon soundings near the Andes mountains, *Geophys. Res. Lett.*, *26*, 1413–1416, 1999.

de la Torre, A., A. Giraldez, and P. Alexander, Saturated gravity wave spectra measured with balloons in Mendoza (Argentina), *Geophys. Res. Lett.*, *21*, 2039–2042, 1994.

de la Torre, A., H. Teitelbaum, and F. Vial, Stratospheric and tropospheric wave measurements near the Andes mountains, *J. Atmos. Terres. Phys.*, *58*, 521–530, 1996.

Dewan, E. M., Stratospheric wave spectra resembling turbulence, *Science*, *204*, 832–835, 1979.

Dewan, E. M., Similitude modeling of internal gravity wave spectra, *Geophys. Res. Lett.*, *18*, 1473–1476, 1991.

Dewan, E. M., The saturated-cascade model for atmospheric gravity wave spectra, and the wavelength-period relations, *Geophys. Res. Lett.*, *21*, 817–820, 1994.

Dewan, E. M., Saturated-cascade similitude theory of gravity wave spectra, *J. Geophys. Res.*, *102*, 29,799–29,817, 1997.

Dewan, E. M., and R. E. Good, Saturation and the "universal" spectrum for vertical profiles of horizontal scalar winds in the atmosphere, *J. Geophys. Res.*, *91*, 2742–2748, 1986.

Dewan, E. M., N. Grossbard, A. F. Quesada, and R. E. Good, Spectral analysis of 10 m resolution scalar velocity profiles in the stratosphere, *Geophys. Res. Lett.*, *11*, 80–83, 1984.

Dewan, E. M., R. H. Picard, R. R. O'Neil, H. A. Gardiner, J. Gibson, J. D. Mill, E. Richards, M. Kendra, and W. O. Gallery, MSX satellite observations of thunderstorm-generated gravity waves in mid-wave infrared images of the upper stratosphere, *Geophys. Res. Lett.*, *25*, 939–942, 1998.

Dong, B., and K. C. Yeh, Resonant and nonresonant wave-wave interactions in an isothermal atmosphere, *J. Geophys. Res.*, *93*, 3729–3744, 1988.

Dörnbrack, A., M. Leutbecher, R. Kivi, and E. Kyrö, Mountain-wave induced record low stratospheric temperatures above northern Scandinavia, *Tellus, Ser. A*, *51*, 951–963, 1999.

Dunkerton, T. J., Stochastic parameterization of gravity wave stresses, *J. Atmos. Sci.*, *39*, 1711–1725, 1982a.

Dunkerton, T. J., Theory of the mesopause semiannual oscillation, *J. Atmos. Sci.*, *39*, 2681–2690, 1982b.

Dunkerton, T. J., Inertia-gravity waves in the stratosphere, *J. Atmos. Sci.*, *41*, 3396–3404, 1984.

Dunkerton, T. J., Effect of nonlinear instability on gravity wave momentum Transport, *J. Atmos. Sci.*, *44*, 3188–3209, 1987.

Dunkerton, T. J., Theory of internal gravity wave saturation, *Pure Appl. Geophys.*, *130*, 373–397, 1989.

Dunkerton, T. J., Shear instability of internal inertia-gravity waves, *J. Atmos. Sci.*, *54*, 1628–1641, 1997a.

Dunkerton, T. J., The role of gravity waves in the quasi-biennial oscillation, *J. Geophys. Res.*, *102*, 26,053–26,076, 1997b.

Durran, D. R., and J. B. Klemp, Another look at downslope winds, part II, Nonlinear amplification beneath wave overturning layers, *J. Atmos. Sci.*, *44*, 3402–3412, 1987.

Eckermann, S. D., Effects of nonstationarity on spectral analysis of mesoscale motions in the atmosphere, *J. Geophys. Res.*, *95*, 16,685–16,703, 1990.

Eckermann, S. D., Ray-tracing simulation of the global propagation of inertia gravity waves through the zonally averaged middle atmosphere, *J. Geophys. Res.*, *97*, 15,849–15,866, 1992.

Eckermann, S. D., Effect of background winds on vertical wavenumber spectra of atmospheric gravity waves, *J. Geophys. Res.*, *100*, 14,097–14,112, 1995a.

Eckermann, S. D., On the observed morphology of gravity-wave and equatorial-wave variance in the stratosphere, *J. Atmos. Terr. Phys.*, *57*, 105–134, 1995b.

Eckermann, S. D., Influence of wave propagation on the Doppler-spreading of atmospheric gravity waves, *J. Atmos. Sci.*, *54*, 2554–2573, 1997.

Eckermann, S. D., Isentropic advection by gravity waves: Quasi-universal M^{-3} vertical wavenumber spectra near the onset of instability, *Geophys. Res. Lett.*, *26*, 201–204, 1999.

Eckermann, S. D., and W. K. Hocking, Effect of superposition on measurements of atmospheric gravity waves: A cautionary note and some reinterpretations, *J. Geophys. Res.*, *94*, 6333–6339, 1989.

Eckermann, S. D., and P. Preusse, Global measurements of stratospheric mountain waves from space, *Science*, *286*, 1534–1537, 1999.

Eckermann, S. D., and R. A. Vincent, Falling sphere observations of anisotropic gravity wave motions in the upper stratosphere over Australia, *Pure Appl. Geophys.*, *130*, 509–532, 1989.

Eckermann, S. D., and R. A. Vincent, ST radar observations of gravity-wave production by cold fronts over southern Australia, *J. Atmos. Sci.*, *50*, 785–806, 1993.

Eckermann, S. D., I. Hirota, and W. K. Hocking, Gravity wave and equatorial wave morphology of the stratosphere derived from long-term rocket soundings, *Q. J. R. Meteorol. Soc.*, *121*, 149–186, 1995.

Edmon, H. J., Jr., B. J. Hoskins, and M. E. McIntyre, Eliassen-Palm cross sections for the troposphere, *J. Atmos. Sci.*, *37*, 2600–2616, 1980.

Fan, Y., J. Klostermeyer, and R. Rüster, VHF radar observation of gravity wave critical levels in the mid-latitude summer mesopause region, *Geophys. Res. Lett.*, *18*, 697–700, 1991.

Farmer, D., and L. Armi, Stratified flow over topography: The role of small-scale entrainment and mixing in flow establishment, *Proc. R. Soc. London, Ser. A*, *455*, 3221–3258, 1999.

Farrell, B. F., and P. J. Ioannou, Generalized stability theory, part I, Autonomous operators, *J. Atmos. Sci.*, *53*, 2025–2040, 1996a.

Farrell, B. F., and P. J. Ioannou, Generalized stability theory, part II, Nonautonomous operators, *J. Atmos. Sci.*, *53*, 2041–2053, 1996b.

Fels, S. B., The radiative damping of short vertical scale waves in the mesosphere, *J. Atmos. Sci.*, *41*, 1755–1764, 1984.

Fetzer, E. J., and J. C. Gille, Gravity wave variance in LIMS temperatures, part I, Variability and comparison with background winds, *J. Atmos. Sci.*, *51*, 2461–2483, 1994.

Fetzer, E. J., and J. C. Gille, Gravity wave variance in LIMS temperatures, part II, Comparison with the zonal-mean momentum balance, *J. Atmos. Sci.*, *53*, 398–410, 1996.

Fleming, E. L., S. Chandra, M. R. Schoeberl, and J. J. Barnett, Monthly mean global climatology of temperature, wind, geopotential height, and pressure for 0-120 km, *NASA Tech. Memo.*, *TM-100697*, 85 pp., 1988.

Forbes, J. M., J. Gu, and S. Miyahara, On the interactions between gravity waves and the diurnal propagating tide, *Planet. Space Sci.*, *39*, 1249–1257, 1991.

Fovell, R., D. Durran, and J. R. Holton, Numerical simulations of convectively generated stratospheric gravity waves, *J. Atmos. Sci.*, *49*, 1427–1442, 1992.

Fritts, D. C., Gravity wave saturation in the middle atmosphere: A review of theory and observations, *Rev. Geophys.*, *22*, 275–308, 1984a.

Fritts, D. C., Shear excitation of atmospheric gravity waves, part II, Nonlinear radiation from a free shear layer, *J. Atmos. Sci.*, *41*, 524–537, 1984b.

Fritts, D. C., A review of gravity wave saturation processes, effects, and variability in the middle atmosphere, *Pure Appl. Geophys.*, *130*, 343–371, 1989.

Fritts, D. C., Errant inferences of gravity wave momentum and heat fluxes using airglow and lidar instrumentation: Corrections and cautions, *J. Geophys. Res.*, *105*, 22,355–22,360, 2000.

Fritts, D. C., and H.-G. Chou, An investigation of the vertical wavenumber and frequency spectra of gravity wave motions in the lower stratosphere, *J. Atmos. Sci.*, *44*, 3610–3624, 1987.

Fritts, D. C., and T. J. Dunkerton, A quasi-linear study of gravity wave saturation and self-acceleration, *J. Atmos. Sci.*, *41*, 3272–3288, 1984.

Fritts, D. C., and T. J. Dunkerton, Fluxes of heat and constituents due to convectively unstable gravity waves, *J. Atmos. Sci.*, *42*, 549–556, 1985.

Fritts, D. C., and M. A. Geller, Viscous stabilization of gravity wave critical level flows, *J. Atmos. Sci.*, *33*, 2276–2284, 1976.

Fritts, D. C., and U.-P. Hoppe, High-resolution measurements of vertical velocity with the EISCAT VHF radar, 2, Spectral observations and model comparisons, *J. Geophys. Res.*, *100*, 16,827–16,838, 1995.

Fritts, D. C., and W. Lu, Spectral estimates of gravity wave energy and momentum fluxes, II, Parameterization of wave forcing and variability, *J. Atmos. Sci.*, *50*, 3695–3713, 1993.

Fritts, D. C., and Z. Luo, Gravity wave excitation by geostrophic adjustment of the jet stream, part I, Two-dimensional forcing, *J. Atmos. Sci.*, *49*, 681–697, 1992.

Fritts, D. C., and Z. Luo, Dynamical and radiative forcing of the summer mesopause circulation and thermal structure, 1, Mean solstice conditions, *J. Geophys. Res.*, *100*, 3119–3128, 1995.

Fritts, D. C., and G. D. Nastrom, Sources of mesoscale variability of gravity waves, II, Frontal, convective, and jet stream excitation, *J. Atmos. Sci.*, *49*, 111–127, 1992.

Fritts, D. C., and P. K. Rastogi, Convective and dynamical instabilities due to gravity wave motions in the lower and middle atmosphere: Theory and observations, *Radio Sci.*, *20*, 1247–1277, 1985.

Fritts, D. C., and T. E. VanZandt, Effects of Doppler shifting on the frequency spectra of atmospheric gravity waves, *J. Geophys. Res.*, *92*, 9723–9732, 1987.

Fritts, D. C., and T. E. VanZandt, Spectral estimates of gravity wave energy and momentum fluxes, I, Energy dissipation, acceleration, and constraints, *J. Atmos. Sci.*, *50*, 3685–3694, 1993.

Fritts, D. C., and R. A. Vincent, Mesospheric momentum flux studies at Adelaide, Australia: Observations and a gravity wave/tidal interaction model, *J. Atmos. Sci.*, *44*, 605–619, 1987.

Fritts, D. C., and D.-Y. Wang, Doppler-shifting effects on frequency spectra of gravity waves observed near the summer mesopause at high latitude, *J. Atmos. Sci.*, *48*, 1535–1544, 1991.

Fritts, D. C., and J. A. Werne, Turbulence dynamics and mixing due to gravity waves in the lower and middle atmosphere, in *Atmospheric Science Across the Stratopause, Geophys. Monogr. Ser.*, vol. 123, edited by D. E. Siskind, S. D. Eckermann, and M. E. Summers, pp. 143–159, AGU, Washington, D. C., 2000.

Fritts, D. C., and L. Yuan, An analysis of gravity wave ducting in the atmosphere: Eckarts resonances in thermal and Doppler ducts, *J. Geophys. Res.*, *94*, 18,455–18,466, 1989a.

Fritts, D. C., and L. Yuan, Measurement of momentum fluxes near the summer mesopause at Poker Flat, Alaska, *J. Atmos. Sci.*, *46*, 2569–2579, 1989b.

Fritts, D. C., and L. Yuan, Stability analysis of inertio-gravity wave structure in the middle atmosphere, *J. Atmos. Sci.*, *46*, 1738–1745, 1989c.

Fritts, D. C., T. Tsuda, T. Sato, S. Fukao, and S. Kato, Observational evidence of a saturated gravity wave spectrum in the troposphere and lower stratosphere, *J. Atmos. Sci.*, *45*, 1741–1759, 1988a.

Fritts, D. C., S. A. Smith, B. B. Balsley, and C. R. Philbrick, Evidence of gravity wave saturation and local turbulence production in the summer mesosphere and lower thermosphere during the STATE experiment, *J. Geophys. Res.*, *93*, 7015–7025, 1988b.

Fritts, D. C., R. C. Blanchard, and L. Coy, Gravity wave structure between 60 and 90 km inferred from space shuttle reentry data, *J. Atmos. Sci.*, *46*, 423–434, 1989.

Fritts, D. C., U.-P. Hoppe, and B. Inhester, A study of the vertical motion field near the high-latitude summer mesopause during MAC/SINE, *J. Atmos. Terr. Phys.*, *52*, 927–938, 1990a.

Fritts, D. C., T. Tsuda, T. E. VanZandt, S. A. Smith, T. Sato, S. Fukao, and S. Kato, Studies of velocity fluctuations in the lower atmosphere using the MU radar, II, Momentum fluxes and energy densities, *J. Atmos. Sci.*, *47*, 51–66, 1990b.

Fritts, D. C., L. Yuan, M. H. Hitchman, L. Coy, E. Kudeki, and R. F. Woodman, Dynamics of the equatorial mesosphere observed using the Jicamarca MST radar during June and August 1987, *J. Atmos. Sci.*, *49*, 2353–2371, 1992.

Fritts, D. C., D.-Y. Wang, R. C. Blanchard, Gravity wave and tidal structures between 60 and 140 km inferred from space shuttle re-entry data, *J. Atmos. Sci.*, *50*, 785–806, 1993a.

Fritts, D. C., J. R. Isler, G. E. Thomas, and Ø. Andreassen, Wave breaking signatures in noctilucent clouds, *Geophys. Res. Lett.*, *20*, 2039–2042, 1993b.

Fritts, D. C., J. R. Isler, and Ø. Andreassen, Gravity wave breaking in two and three dimensions, 2, Three-dimensional evolution and instability structure, *J. Geophys. Res.*, *99*, 8109–8123, 1994.

Fritts, D. C., J. F. Garten, and Ø. Andreassen, Wave breaking and transition to turbulence in stratified shear flows, *J. Atmos. Sci.*, *53*, 1057–1085, 1996a.

Fritts, D. C., T. L. Palmer, Ø. Andreassen, and I. Lie, Evolution and breakdown of Kelvin-Helmholtz billows in stratified compressible flows, I, Comparison of two- and three-dimensional flows, *J. Atmos. Sci.*, *53*, 3173–3191, 1996b.

Fritts, D. C., et al., Equatorial dynamics observed by rocket, radar, and satellite during the CADRE/MALTED campaign, 2, Mean and wave structures, coherence, and variability, *J. Geophys. Res.*, *102*, 26,11–26,216, 1997a.

Fritts, D. C., J. R. Isler, J. H. Hecht, R. L. Walterscheid, Ø. Andreassen, Wave breaking signatures in sodium densities and OH nightglow, part II, Simulation of wave and instability structures, *J. Geophys. Res.*, *102*, 6669–6684, 1997b.

Fritts, D. C., S. Arendt, and Ø. Andreassen, Vorticity dynamics in a breaking internal gravity wave, 2, Vortex interactions and transition to turbulence, *J. Fluid Mech.*, *367*, 47–65, 1998.

Fritts, D. C., S. Arendt, and Ø. Andreassen, The vorticity dynamics of instability and turbulence in a breaking internal gravity wave, *Earth Planets Space*, *51*, 457–473, 1999.

Fritts, D. C., S. L. Vadas, and Y. Yamada, An estimate of strong local body forcing and gravity wave radiation based on OH airglow and meteor radar observations, *Geophys. Res. Lett.*, *29*(10), 1429, doi:10.1029/2001GL013753, 2002.

Fritts, D. C., C. Bizon, J. A. Werne, and C. K. Meyer, Layering accompanying turbulence generation due to shear instability and gravity wave breaking, *J. Geophys. Res.*, *108*, doi: 10.1029/2002JD002406, in press 2003.

Fukao, S., Y. Maekawa, T. Sato, and S. Kato, Fine structure in mesospheric wind fluctuations observed by the Arecibo UHF Doppler radar, *J. Geophys. Res.*, *90*, 7547–7556, 1985.

Fuller-Rowell, T. J., The impact of gravity waves on nitric oxide in the lower Thermosphere, *J. Atmos. Terr. Phys.*, *56*, 1387–1398, 1994.

Gage, K. S., Evidence for $k^{-5/3}$ law inertial range in mesoscale two-dimensional turbulence, *J. Atmos. Sci.*, *36*, 1950–1954, 1979.

Garcia, R. R., Dynamics, radiation, and photochemistry in the mesosphere: Implications for the formation of noctilucent clouds, *J. Geophys. Res.*, *94*, 14,605–14,615, 1989.

Garcia, R. R., and B. Boville, "Downward control" of the mean meridional circulation and temperature distribution of the polar winter stratosphere, *J. Atmos. Sci.*, *51*, 2238–2245, 1994.

Garcia, R. R., and F. Sassi, Modulation of the Mesospheric Semiannual Oscillation by the Quasibiennial Oscillation, *Earth Planets Space*, *51*, 563–569, 1999.

Garcia, R. R., and S. Solomon, The effect of breaking gravity waves on the dynamics and chemical composition of the mesosphere and lower thermosphere, *J. Geophys. Res.*, *90*, 3850–3868, 1985.

Garcia, R. R., T. I. Dunkerton, R. S. Lieberman, and R. A. Vincent, Climatology of the semiannual oscillation of the tropical middle atmosphere, *J. Geophys. Res.*, *102*, 26,019–26,032, 1997.

Gardner, C. S., Diffusive filtering theory of gravity wave spectra in the atmosphere, *J. Geophys. Res.*, *99*, 20,601–20,622, 1994.

Gardner, C. S., C. A. Hostetler, and S. J. Franke, Gravity wave models for the horizontal wave number spectra of atmospheric velocity and density fluctuations, *J. Geophys. Res.*, *98*, 1035–1049, 1993.

Gardner, C. S., M. Coble, G. C. Papen, and G. S. Swenson, Observations of the unambiguous two-dimensional horizontal wave number spectrum of OH intensity perturbations, *Geophys. Res. Lett.*, *23*, 3739–3742, 1996.

Gardner, C. S., K. Gulati, Y. Zhao, and G. Swenson, Measuring gravity wave momentum fluxes with airglow imagers, *J. Geophys. Res.*, *104*, 11,903–11,915, 1999a.

Gardner, C. S., S. J. Franke, W. Yang, X. Tao, and J. R. Yu, Interpretation of gravity waves observed in the mesopause region at Starfire Optical Range, New Mexico: Strong evidence for nonseparable intrinsic (m, omega) spectra, *J. Geophys. Res.*, *103*, 8699–8713, 1999b.

Gary, B. L., Observational results using the microwave temperature profiler during the airborne antarctic ozone experiment, *J. Geophys. Res.*, *94*, 11,223–11,231, 1989.

Gavrilov, N. M., and V. A. Yudin, Model for coefficients of turbulence and effective Prandtl number produced by breaking gravity waves in the upper atmosphere, *J. Geophys. Res.*, *97*, 7619–7624, 1992.

Gavrilov, N. M., S. Fukao, T. Nakamura, T. Tsuda, M. D. Yamanaka, and M. Yamamoto, Statistical analysis of gravity waves observed with the middle and upper atmosphere radar in the middle atmosphere, 1, Method and general characteristics, *J. Geophys. Res.*, *101*, 29,511–29,521, 1996.

Gerz, T., J. Howell, and L. Mahrt, Vortex structures and microfronts, *Phys. Fluids*, *6*, 1242–1251, 1994.

Gill, A. E., *Atmosphere-Ocean Dynamics*, 662 pp., Academic, San Diego, Calif., 1982.

Gille, J. C., and J. M. Russell III, The limb infrared monitor of the stratosphere: Experiment description, performance, and results, *J. Geophys. Res.*, *89*, 5125–5140, 1984.

Goldberg, R. A., G. Lehmacher, F. J. Schmidlin, D. C. Fritts, J. D. Mitchell, C. L. Croskey, M. Friedrich, and W. E. Swartz, Equatorial dynamics observed by rocket, radar, and satellite during the CADRE/MALTED campaign, 1, Programmatics and small-scale fluctuations. *J. Geophys. Res.*, *102*, 26,179–26,190, 1997.

Gray, L. J., and J. A. Pyle, A two-dimensional model of the quasi-biennial oscillation of ozone, *J. Atmos. Sci.*, *46*, 203–220, 1989.

Griffiths, M., and M. J. Reeder, Stratospheric inertia-gravity waves generated in a numerical model of frontogenesis, I, Model solutions, *Q. J. R. Meteorol. Soc.*, *122*, 1153–1174, 1996.

Grimshaw, R. H. J., Resonant wave interactions in a stratified shear flow, *J. Fluid Mech.*, *190*, 357–374, 1988.

Guest, F. M., M. J. Reeder, C. J. Marks, and D. J. Karoly, Inertia-gravity waves observed in the lower stratosphere over Macquarie Island, *J. Atmos. Sci.*, *57*, 737–752, 2000.

Hall, C. M., and U.-P. Hoppe, Estimates of turbulent energy dissipation rates from determinations of characteristic vertical wavenumber by EISCAT, *Geophys. Res. Lett.*, *25*, 4075–4078, 1998.

Hamilton, K., Diagnostic study of the momentum balance in the Northern Hemisphere winter stratosphere, *Mon. Weather Rev.*, *111*, 1434–1441, 1983.

Hamilton, K., Climatological statistics of stratospheric intertia-gravity waves deduced from historical rocketsonde wind and temperature data, *J. Geophys. Res.*, *96*, 20,831–20,839, 1991.

Hamilton, K., Comprehensive simulation of the middle atmospheric climate: Some recent results, *Clim. Dyn.*, *11*, 223–241, 1995.

Hamilton, K., Comprehensive meteorological modelling of the middle atmosphere: A tutorial review, *J. Atmos. Terr. Phys.*, *58*, 1591–1627, 1996.

Hamilton, K. (Ed.), *Gravity Wave Processes and Their Parameterization in Global Climate Models*, 404 pp., Springer-Verlag, New York, 1997.

Hamilton, K., and R. A. Vincent, High-resolution radiosonde data offer new prospects for research, *Eos Trans. AGU*, *76*, 497, 1995.

Hamilton, K., R. J. Wilson, J. D. Mahlman, and L. J. Umscheid, Climatology of the SKYHI troposphere stratosphere mesosphere general circulation model, *J. Atmos. Sci.*, *52*, 5–43, 1995.

Hamilton, K., R. J. Wilson, and R. S. Hemler, Middle atmosphere simulated with high vertical and horizontal resolution versions of a GCM: Improvements in the cold pole bias and generation of a QBO-like oscillation in the tropics, *J. Atmos. Sci.*, *56*, 3829–3846, 1999.

Haque, R., and G. R. Swenson, Extraction of motion parameters of gravity-wave structures from all-sky OH image sequences, *Appl. Opt.*, *38*, 4433–4442, 1999.

Hauchecorne, A., M. L. Chanin, and R. Wilson, Mesopheric temperature inversion and gravity wave breaking, *Geophys. Res. Lett.*, *14*, 933–936, 1987.

Haynes, P. H., C. J. Marks, M. E. McIntyre, T. G. Shephard, and K. P. Shine, On the "downward control" of extratropical diabatic circulations by eddy-induced mean zonal forces, *J. Atmos. Sci.*, *48*, 657–678, 1991.

Hecht, J. H., R. L. Walterscheid, D. C. Fritts, J. R. Isler, D. C. Senft, C. S. Gardner, and S. J. Franke, Wave breaking signatures in OH airglow and sodium densities and temperatures, part I, Airglow imaging, Na lidar, and MF radar

observations, *J. Geophys. Res.*, *102*, 6655–6668, 1997.
Hecht, J. H., C. Fricke-Bergemann, and R. L. Walterscheid, Observations of the breakdown of an atmospheric gravity wave near the cold summer mesopause at 54N, *Geophys. Res. Lett.*, *27*, 879–882, 2000.
Hertzog, A., and F. Vial, A study of the dynamics of the equatorial lower atmosphere by use of ultra-long duration balloons, 2, Gravity waves, *J. Geophys. Res.*, *106*, 22,745–22,761, 2001.
Hickey, M. P., M. J. Taylor, C. S. Gardner, and C. R. Gibbons, Full-wave modeling of small-scale gravity waves using Airborne Lidar and Observations of the Hawaiian Airglow (ALOHA-93) O(1S) images and coincident Na wind/temperature lidar Measurements, *J. Geophys. Res.*, *103*, 6439–6453, 1998.
Hines, C. O., Internal atmospheric gravity waves at ionospheric heights, *Can. J. Phys.*, *38*, 1441–1481, 1960.
Hines, C. O., Generalizations of the Richardson criterion for the onset of atmospheric turbulence, *Q. J. R. Meteorol. Soc*, *97*, 429–439, 1971.
Hines, C. O., A modeling of atmospheric gravity waves and wave drag generated by isotropic and anisotropic terrain, *J. Atmos. Sci*, *45*, 309–322, 1988a.
Hines, C. O., The generation of turbulence by atmospheric gravity waves, *J. Atmos. Sci.*, *45*, 1269–1278, 1988b.
Hines, C. O., The saturation of gravity waves in the middle atmosphere, part I, Critique of linear-instability theory, *J. Atmos. Sci.*, *48*, 1348–1359, 1991.
Hines, C. O., The saturation of gravity waves in the middle atmosphere, part IV, Cutoff of the incident wave spectrum, *J. Atmos. Sci.*, *50*, 3045–3060, 1993.
Hines, C. O., Nonlinearity of gravity wave saturated spectra in the middle atmosphere, *Geophys. Res. Lett.*, *23*, 3309–3312, 1996.
Hines, C. O., Doppler-spread parameterization of gravity-wave momentum deposition in the middle atmosphere, 1, Basic formulation, *J. Atmos. Sol. Terr. Phys.*, *59*, 371–386, 1997a.
Hines, C. O., Doppler-spread parameterization of gravity-wave momentum deposition in the middle atmosphere, 2, Broad and quasi monochromatic spectra, and implementation, *J. Atmos. Sol. Terr. Phys.*, *59*, 387–400, 1997b.
Hines, C. O., Comments on paper by S.D. Eckermann, "Influence of wave propagation on the Doppler spreading of atmospheric gravity waves," *J. Atmos. Sci.*, *56*, 1094–1098, 1999.
Hines, C. O., Theory of the Eulerian tail in the spectra of atmospheric and oceanic internal gravity waves, *J. Fluid Mech.*, *448*, 289–313, 2001.
Hines, C. O., Nonlinearities and linearities in internal gravity waves of the atmosphere and oceans, *Geophys. Astrophys. Fluid Dyn.*, *96*, 1–30, 2002a.
Hines, C. O., Comments on the gravity wave theory of J. Weinstock concerning dissipation induced by nonlinear effects, *J. Atmos. Sci.*, *59*, 2024–2030, 2002b.
Hirota, I., Climatology of gravity waves in the middle atmosphere, *J. Atmos. Terr. Phys.*, *46*, 767–773, 1984.
Hirota, I., and T. Niki, A statistical study of inertia-gravity waves in the middle atmosphere, *J. Meteorol. Soc. Jpn.*, *63*, 1055–1066, 1985.
Hitchman, M. H., J. C. Gille, C. D. Rodgers, and G. Brasseur, The separated polar winter stratopause: A gravity wave driven climatological feature, *J. Atmos. Sci.*, *46*, 410–422, 1989.
Hitchman, M. H., K. W. Bywaters, D. C. Fritts, L. Coy, E. Kudeki, and F. Surucu, Mean winds and momentum fluxes over Jicamarca, Peru, during June and August 1987, *J. Atmos. Sci.*, *49*, 2372–2383, 1992.
Hitchman, M. H., and C. B. Leovy, Estimation of the Kelvin wave contribution to the semiannual oscillation, *J. Atmos. Sci.*, *45*, 1462–1475, 1988.
Hocking, W. K., Turbulence in the region 80–120 km, *Adv. Space Res.*, *10*, 153–161, 1990.
Hocking, W. K., S. Fukao, M. Yamamoto, T. Tsuda, and S. Kato, Viscosity waves and thermal conduction waves as a cause of "specular" reflectors in radar studies of the atmosphere, *Radio Sci.*, *26*, 1281–1303, 1991.
Hodges, R. R., Jr., Generation of turbulence in the upper atmosphere by internal gravity waves, *J. Geophys. Res.*, *72*, 3455–3458, 1967.
Holton, J. R., The role of gravity wave induced drag and diffusion in the momentum budget of the mesosphere, *J. Atmos. Sci.*, *39*, 791–799, 1982.
Holton, J. R., The influence of gravity wave breaking on the general circulation of the middle atmosphere, *J. Atmos. Sci.*, *40*, 2497–2507, 1983.
Holton, J. R., The generation of mesospheric planetary waves by zonally asymmetric gravity wave breaking, *J. Atmos. Sci.*, *41*, 3427–3430, 1984.
Holton, J. R., *An Introduction to Dynamic Meteorology*, 507 pp., Academic, San Diego, Calif., 1992.
Holton, J. R., and M. J. Alexander, The role of waves in the transport circulation of the middle atmosphere, in *Atmospheric Science Across the Stratopause, Geophys. Monogr. Ser.*, vol. 123, edited by D. E. Siskind, S. D. Eckermann, and M. E. Summers, pp. 21–35, AGU, Washington, D. C., 2000.
Holton, J. R., and M. R. Schoeberl, The role of gravity wave generated advection and diffusion in transport of tracers in the mesosphere, *J. Geophys. Res.*, *93*, 11,075–11,082, 1988.
Holton, J. R., M. J. Alexander, and M. T. Boehm, Evidence for short vertical wavelength Kelvin waves in the DOE-ARM Nauru99 radiosonde data, *J. Geophys. Res.*, *106*, 20,125–20,129, 2001.
Hoppe, U.-P., and D. C. Fritts, High-resolution measurements of vertical velocity with the EISCAT VHF radar, 1, Motion field characteristics and measurement biases, *J. Geophys. Res.*, *100*, 16,813–16,826, 1995.
Hoppe, U.-P., T. Eriksen, E. V. Thrane, T. A. Blix, J. Fiedler, and F.-J. Lübken, Observations in the polar middle atmosphere by rocket-borne Rayleigh lidar: First results, *Earth Planets Space*, *51*, 815–824, 1999.
Horinouchi, T., and S. Yoden, Wave-mean flow interaction associated with a QBO-like oscillation simulated in a simplified GCM, *J. Atmos. Sci.*, *55*, 502–526, 1998.
Huaman, M. M., and B. B. Balsley, Differences in near-mesopause summer winds, temperatures, and water vapor at northern and southern latitudes as possible causal factors for inter-hemispheric PMSE differences, *Geophys. Res. Lett.*, *26*, 1529–1532, 1999.
Huang, T. Y. W., and A. K. Smith, The mesospheric diabatic circulation and the parameterized thermal effect of gravity wave breaking on the circulation, *J. Atmos. Sci.*, *48*, 1093–1111, 1991.
Huang, T. Y. W., and A. K. Smith, Dynamical and chemical feedback in a two-dimensional interactive model of the middle atmosphere, *J. Geophys. Res.*, *100*, 11,085–11,104, 1995.
Isler, J. R., and D. C. Fritts, Gravity wave variability and interaction with lower-frequency motions in the mesosphere and lower thermosphere over Hawaii, *J. Atmos. Sci.*, *53*, 37–48, 1996.
Isler, J. R., D. C. Fritts, and Ø. Andreassen, Gravity wave breaking in two and three dimensions, 3, Vortex breakdown and transition to turbulence, *J. Geophys. Res.*, *99*, 8125–8137, 1994.
Isler, J. R., M. J. Taylor, and D. C. Fritts, Observational evidence of wave ducting and evanescence in the mesosphere, *J. Geophys. Res.*, *102*, 26,301–26,313, 1997.
Jackson, D. R., and L. J. Gray, Simulation of the semi-annual oscillation of the equatorial middle atmosphere using the

Extended UGAMP General Circulation Model, *Q. J. R. Meteorol. Soc.*, *120*, 1559–1588, 1994.

Jasperson, W. H., G. D. Nastrom, and D. C. Fritts, Further study of terrain effects on the mesoscale spectrum of atmospheric motions, *J. Atmos. Sci.*, *47*, 979–987, 1990.

Jeong, J., and F. Hussain, On the identification of a vortex, *J. Fluid Mech.*, *285*, 69–94, 1995.

Jones, P. W., K. Hamilton, and R. J. Wilson, A very high resolution general circulation model simulation of the global circulation in austral winter, *J. Atmos. Sci.*, *54*, 1107–1116, 1997.

Karoly, D. J., G. L. Roff, and M. J. Reeder, Gravity wave activity associated with tropical convection detected in TOGA COARE sounding data, *Geophys. Res. Lett.*, *23*, 261–264, 1996.

Kennedy, P. J., and M. A. Shapiro, Further encounters with clear air turbulence in research aircraft, *J. Atmos. Sci.*, *37*, 986–993, 1979.

Kershaw, R., Parameterization of momentum transport by convectively generated gravity waves, *Q. J. R. Meteorol. Soc.*, *121*, 1023–1040, 1995.

Kiehl, J. T., J. J. Hack, G. B. Bonan, B. A. Boville, B. P. Briegleb, D. L. Williamson, and P. J. Rasch, Description of the NCAR Community Climate Model (CCM3), *NCAR Tech. Note 420*, 96–103, Natl. Cent. for Atmos. Res., Boulder, Colo., 1996.

Kinnersley, J. S., The climatology of the stratospheric "THIN AIR" model, *Q. J. R. Meteorol. Soc.*, *122*, 219–252, 1996.

Kitamura, Y., and I. Hirota, Small-scale disturbances in the lower stratosphere revealed by daily rawin sonde observations, *J. Meteorol. Soc. Jpn.*, *67*, 817–830, 1989.

Klaassen, G. P., and W. R. Peltier, The onset of turbulence in finite-amplitude Kelvin-Helmholtz billows, *J. Fluid Mech.*, *155*, 1–35, 1985.

Klostermeyer, J., Observations indicating parametric instabilities of internal gravity waves at thermospheric heights, *Geophys. Astrophys. Fluid Dyn.*, *29*, 117–138, 1984.

Klostermeyer, J., On the role of parametric instability of internal gravity waves in atmospheric radar observations, *Radio Sci.*, *25*, 983–995, 1990.

Klostermeyer, J., Two- and three-dimensional parametric instabilities in finite amplitude internal gravity waves, *Geophys. Astrophys. Fluid Dyn.*, *64*, 1–25, 1991.

Kuettner, J. P., P. A. Hildebrand, and T. L. Clark, Convection waves: Observations of gravity wave systems over convectively active boundary layers, *Q. J. R. Meteorol. Soc.*, *113*, 445–467, 1987.

Kuo, F.-S., H.-W. Shen, I.-J. Fu, J.-K. Chao, J. Röttger, and C. H. Liu, Altitude dependence of vertical velocity spectra observed by VHF radar, *Radio Sci.*, *20*, 1349–1354, 1985.

Kuo, F. S., H. Y. Lue, C. M. Huang, C. L. Lo, C. H. Liu, and S. Fukao, A study of velocity fluctuation spectra in the troposphere and lower stratosphere using MU radar, *J. Atmos. Terres. Phys.*, *54*, 31–48, 1992.

Lalas, D. P., and F. Einaudi, On characteristics of gravity waves generated by atmospheric shear layers, *J. Atmos. Sci.*, *33*, 1248–1259, 1976.

Lane, T. P., M. J. Reeder, and T. L. Clark, Numerical modeling of gravity wave generation by deep tropical convection, *J. Atmos. Sci.*, *58*, 1249–1274, 2001.

Larsen, M. F., R. F. Woodman, T. Sato, and M. K. Davis, Power spectra of oblique velocities in the troposphere and lower stratosphere observed at Arecibo, Puerto Rico, *J. Atmos. Sci.*, *43*, 2230–2240, 1986.

Lawrence, B. N., The effect of parameterized gravity wave drag on simulations of the middle atmosphere during northern winter 1991/1992—General evolution, in *Gravity Wave Processes: Their Parameterization in Global Climate Models, NATO ASI Ser., Ser. I*, vol. 50, edited by K. Hamilton, pp. 291–307, Springer-Verlag, New York, 1997.

Lefrere, J., and C. Sidi, Distortion of gravity wave spectra measured from ground-based stations, *Geophys. Res. Lett.*, *17*, 1581–1584, 1990.

LeLong, M.-P., and T. J. Dunkerton, Inertia-gravity wave breaking in three dimensions, 1, Convectively stable waves, *J. Atmos. Sci.*, *55*, 2473–2488, 1998a.

LeLong, M.-P., and T. J. Dunkerton, Inertia-gravity wave breaking in three dimensions, 2, Convectively unstable waves, *J. Atmos. Sci.*, *55*, 2489–2501, 1998b.

Leutbecher, M., and H. Volkert, Propagation of mountain waves into the stratosphere: Quantitative evaluation of three-dimensional simulations, *J. Atmos. Sci.*, *57*, 3090–3108, 2000.

Lieberman, R. S., et al., HRDI observations of mean meridional winds at solstice, *J. Atmos. Sci.*, *55*, 1887–1896, 1998.

Lighthill, J., *Waves in Fluids*, 504 pp., Cambridge Univ. Press, New York, 1978.

Lilly, D. K., Stratified turbulence and the mesoscale variability of the atmosphere, *J. Atmos. Sci.*, *40*, 749–761, 1983.

Lilly, D. K., and P. J. Kennedy, Observations of a stationary mountain wave and its associated momentum flux and energy dissipation, *J. Atmos. Sci.*, *30*, 1135–1152, 1973.

Lilly, D. K., and P. F. Lester, Waves and turbulence in the stratosphere, *J. Atmos. Sci.*, *31*, 800–812, 1974.

Lindzen, R. S., Wave-mean flow interactions in the upper atmosphere, *Boundary Layer Meteorol.*, *4*, 327–343, 1973.

Lindzen, R. S., Turbulence and stress owing to gravity wave and tidal breakdown, *J. Geophys. Res.*, *86*, 9707–9714, 1981.

Lindzen, R. S., and J. R. Holton, A theory of the quasi-biennial oscillation, *J. Atmos. Sci.*, *25*, 1095–1107, 1968.

Lintelman, S. A., and C. S. Gardner, Observation and interpretation of spectra of atmospheric gravity wave perturbations with upward and downward phase progression, *J. Geophys. Res.*, *99*, 16,959–16,971, 1994.

Liu, H.-L., M. E. Hagan, and R. G. Roble, Local mean state changes due to gravity wave breaking modulated by the diurnal tide, *J. Geophys. Res.*, *105*, 12,381–12,396, 2000.

Lombard, P. N., and J. J. Riley, Instability and breakdown of internal gravity Waves, I, Linear stability analysis, *Phys. Fluids*, *8*, 3271–3287, 1996.

Long, R. R., Some aspects of the flow of stratified fluids, III, Continuous density gradients, *Tellus*, *7*, 341–357, 1955.

Lott, F., Linear mountain drag and averaged pseudo-momentum flux profiles in the presence of trapped lee waves, *Tellus, Ser. A*, *50*, 12–25, 1998.

Lott, F., and M. J. Miller, A new subgrid-scale orographic drag parametrization: Its formulation and testing, *Q. J. R. Meteorol. Soc.*, *123*, 101–127, 1997.

Lu, W., and D. C. Fritts, Spectral estimates of gravity wave enrgy and momentum fluxes, III, Gravity wave-tidal interactions, *J. Atmos. Sci.*, *50*, 3714–3727, 1993.

Lübken, F.-J., Seasonal variations of turbulent energy dissipation rates at high latitudes as determined by in situ measurements of neutral density fluctuations, *J. Geophys. Res.*, *102*, 13,441–13,456, 1997.

Lübken, F.-J., Thermal structure of the Arctic summer mesosphere, *J. Geophys. Res.*, *104*, 9135–9149, 1999.

Lübken, F.-J., W. Hillert, G. Lehmacher, and U. von Zahn, Experiments revealing small impact of turbulence on the energy budget of the mesosphere and lower thermosphere, *J. Geophys. Res.*, *98*, 20,369–20,384, 1993.

Lübken, F.-J., M. J. Jarvis, and G. O. L. Jones, First in situ temperature measurements at the Antarctic summer mesopause, *Geophys. Res. Lett.*, *26*, 3581–3584, 1999.

Luo, Z., and D. C. Fritts, Gravity-wave excitation by geostrophic adjustment of the jet stream, part II, Three-dimensional forcing, *J. Atmos. Sci.*, *50*, 105–115, 1993.

Luo, Z., D. C. Fritts, R. W. Portmann, and G. E. Thomas, Dynamical and radiative forcing of the summer mesopause circulation and thermal structure, 2, Seasonal variations, *J. Geophys. Res.*, *100*, 3129–3137, 1995.

Maekawa, Y., S. Fukao, and S. Kato, Vertical propagation characteristics of internal gravity waves around the mesopause observed by the Arecibo UHF radar, *J. Atmos. Terr. Phys.*, *49*, 73–80, 1987.

Manson, A. H., and C. E. Meek, Gravity wave propagation characteristics (60–120 km) as determined by the Saskatoon MF radar (Gravnet) system: 1983–85 at 52N, 107W, *J. Atmos. Sci.*, *45*, 932–946, 1988.

Manson, A. H., and C. E. Meek, Characteristics of gravity waves (10 min-6 hours) at Saskatoon (52°N, 107°W): Observations by the phase coherent medium frequency radar, *J. Geophys. Res.*, *98*, 20,357–20,367, 1993.

Manson, A. H., C. E. Meek, J. Qian, and C. S. Gardner, Spectra of gravity wave density and wind perturbations observed during Arctic Noctilucent Cloud (ANLC-93) campaign over the Canadian Prairies: Synergistic airborne Na lidar and MF radar observations, *J. Geophys. Res.*, *103*, 6455–6565, 1998a.

Manson, A. H., C. E. Meek, and G. E. Hall, Correlations of gravity waves and tides in the mesosphere over Saskatoon, *J. Atmos. Sol. Terr. Phys.*, *60*, 1089–1107, 1998b.

Manzini, E., and K. Hamilton, Middle atmospheric traveling waves forced by latent and convective heating, *J. Atmos. Sci.*, *50*, 2180–2200, 1993.

Manzini, E., and N. A. McFarlane, The effect of varying the source spectrum of a gravity wave parameterization in a middle atmosphere general circulation model, *J. Geophys. Res.*, *103*, 31,523–31,539, 1998.

Marks, C. J., Some features of the climatology of the middle atmosphere revealed by Nimbus 5 and 6, *J. Atmos. Sci.*, *46*, 2485–2508, 1989.

Marks, C. J., and S. D. Eckermann, A three-dimensional nonhydrostatic ray-tracing model for gravity waves: Formulation and preliminary results for the middle atmosphere, *J. Atmos. Sci.*, *52*, 1959–1984, 1995.

Marsh, A. K. P., N. J. Mitchell, and L. Thomas, Lidar studies of stratospheric gravity-wave spectra, *Planet. Space Sci.*, *39*, 1541–1548, 1991.

Mastrantonio, G., F. Einaudi, D. Fua, and D. P. Lalas, Generation of gravity waves by jet streams in the atmosphere, *J. Atmos. Sci.*, *33*, 1730–1738, 1976.

Matsuno, T., A quasi one-dimensional model of the middle atmosphere circulation interacting with internal gravity waves, *J. Meteorol. Soc. Jpn.*, *60*, 215–226, 1982.

Mayr, H. G., J. G. Mengel, C. O. Hines, K. L. Chan, N. F. Arnold, C. A. Reddy, and H. S. Porter, The gravity wave Doppler spread theory applied in a numerical spectral model of the middle atmosphere, 1, Model and global scale seasonal variations, *J. Geophys. Res.*, *102*, 26,077–26,091, 1997a.

Mayr, H. G., J. G. Mengel, C. O. Hines, K. L. Chan, N. F. Arnold, C. A. Reddy, and H. S. Porter, The gravity wave Doppler spread theory applied in a numerical spectral model of the middle atmosphere, 2, Equatorial oscillations, *J. Geophys. Res.*, *102*, 26,093–26,105, 1997b.

Mayr, H. G., J. G. Mengel, K. L. Chan, and H. S. Porter, Seasonal variations of the diurnal tide induced by gravity wave filtering, *Geophys. Res. Lett.*, *25*, 943–946, 1998.

McComas, C. H., and F. P. Bretherton, Resonant interaction of oceanic internal waves, *J. Geophys. Res.*, *82*, 1397–1412, 1977.

McEwan, A. D., Degeneration of resonantly-excited standing internal gravity waves, *J. Fluid Mech.*, *50*, 431–448, 1971.

McEwan, A. D., and R. M. Robinson, Parametric instability of internal gravity waves, *J. Fluid Mech.*, *67*, 667–687, 1975.

McFarlane, N. A., The effect of orographically excited gravity wave drag on the general circulation of the lower stratosphere and troposphere, *J. Atmos. Sci.*, *44*, 1775–1800, 1987.

McIntyre, M. E., On dynamics and transport near the polar mesopause in summer, *J. Geophys. Res.*, *94*, 14,617–14,628, 1989.

McLandress, C., On the importance of gravity waves in the middle atmosphere and their parameterization in general circulation models, *J. Atmos. Sol. Terr. Phys.*, *60*, 1357–1383, 1998.

McLandress, C., The seasonal variation of the propagating diurnal tide in the mesosphere and lower thermosphere, part I, The role of gravity waves and planetary waves, *J. Atmos. Sci*., *59*, 893–906, 2002.

McLandress, C., and W. E. Ward, Tidal/gravity wave interactions and their influence on the large-scale dynamics of the middle atmosphere: Model results, *J. Geophys. Res.*, *99*, 8139–8156, 1994.

McLandress, C., M. J. Alexander, and D. L. Wu, Microwave Limb Sounder observations of gravity waves in the stratosphere: A climatology and interpretation, *J. Geophys. Res.*, *105*, 11,947–11,967, 2000.

Medvedev, A. S., and G. P. Klaassen, Vertical evolution of gravity wave spectra and the parameterization of associated wave drag, *J. Geophys. Res.*, *100*, 25,841–25,853, 1995.

Medvedev, A. S., and G. P. Klaasen, Parameterization of gravity wave momentum deposition based on nonlinear wave interactions: Basic formulation and sensitivity tests, *J. Atmos. Sol. Terr. Phys.*, *62*, 1015–1033, 2000.

Medvedev, A. S., and G. P. Klaasen, Realistic semiannual oscillation simulated in a middle atmosphere general circulation model, *Geophys. Res. Lett.*, *28*, 733–736, 2001.

Medvedev, A. S., G. P. Klaassen, and S. R. Beagley, On the role of an anisotropic gravity wave spectrum in maintaining the circulation of the middle atmosphere, *Geophys. Res. Lett.*, *25*, 509–512, 1998.

Meek, C. E., I. M. Reid, and A. H. Manson, Observations of mesospheric wind velocities, 2, Cross sections of power spectral density for 48–8 hours, 8–1 hours, and 1 hour to 10 min over 60–110 km for 1981, *Radio Sci.*, *20*, 1383–1402, 1985.

Meyer, C. K., Gravity wave interactions with the diurnal propagating tide, *J. Geophys. Res.*, *104*, 4223–4239, 1999a.

Meyer, C. K., Gravity wave interactions with mesospheric planetary waves: A mechanism for penetration into the thermosphere-ionosphere system, *J. Geophys. Res.*, *104*, 28,181–28,196, 1999b.

Meyer, W., R. Siebenmorgen, and H.-U. Widdel, Estimates of gravity wave momentum fluxes in the winter and summer high mesosphere over northern Scandinavia, *J. Atmos. Terr. Phys.*, *51*, 311–319, 1989.

Mied, R. P., The occurrence of parametric instabilities in finite amplitude internal gravity waves, *J. Fluid Mech.*, *78*, 763–784, 1976.

Miller, M. J., T. N. Palmer, and R. Swinbank, Parameterization and influence of subgridscale orography in general circulation and numerical weather prediction models, *MAP Handbk.*, *40*, 84–109, 1989.

Mitchell, N. J., and V. S. C. Howells, Vertical velocities associated with gravity waves measured in the mesosphere and lower thermosphere with the EISCAT VHF radar, *Ann. Geophys.*, *16*, 1367–1379, 1998.

Mitchell, N. J., L. Thomas, and A. K. P. Marsh, Lidar observations of long-period gravity waves in the stratosphere, *Ann. Geophys.*, *9*, 588–596, 1991.

Mitchell, N. J., L. Thomas, and I. T. Prichard, Gravity waves in the stratosphere and troposphere observed by lidar and MST radar, *J. Atmos. Terr. Phys.*, *56*, 939–947, 1994.

Mitchell, N. J., A. J. McDonald, S. J. Reid, and J. D. Price,

Observations of gravity waves in the upper and lower stratosphere by lidar and ozonesondes, *Ann. Geophys.*, *14*, 309–314, 1996.

Miyahara, S., Suppression of stationary planetary waves by internal gravity waves in the mesosphere, *J. Atmos. Sci.*, *42*, 100–107, 1985.

Miyahara, S., and J. M. Forbes, Interactions between gravity waves and the diurnal tide in the mesosphere and lower thermosphere, *J. Meteorol. Soc. Jpn.*, *69*, 523–531, 1991.

Miyahara, S., Y. Hayashi, and J. D. Mahlman, Interactions between gravity waves and planetary-scale flow simulated by the GFDL "SKYHI" general circulation model, *J. Atmos. Sci.*, *43*, 1844–1861, 1986.

Müller, K. M., U. Langematz, and S. Pawson, The stratopause semiannual oscillation in the Berlin Troposphere-Stratosphere-Mesosphere GCM, *J. Atmos. Sci.*, *54*, 2749–2759, 1997.

Müller, P., G. Holloway, F. Henyey, and N. Pomphrey, Nonlinear interactions among internal gravity waves, *Rev. Geophys.*, *24*, 493–536, 1986.

Muraoka, Y., T. Sugiyama, K. Kawahira, T. Sato, T. Tsuda, S. Fukao, and S. Kato, Cause of monochromatic inertia-gravity wave breaking observed by the MU radar, *Geophys. Res. Lett.*, *15*, 1349–1352, 1988.

Murayama, Y., T. Tsuda, and S. Fukao, Seasonal variation of gravity wave activity in the lower atmosphere observed with the MU radar, *J. Geophys. Res.*, *99*, 23,057–23,069, 1994.

Murphy, D. J., and R. A. Vincent, Mesospheric momentum fluxes over Adelaide during the 2-day wave: Results and interpretation, *J. Geophys. Res.*, *103*, 28,627–28,636, 1998.

Nakamura, T., T. Tsuda, S. Fukao, and S. Kato, Comparison of the mesospheric gravity waves observed with the MU radar (35N) and the Adelaide MF radar (35S), *Geophy. Res. Lett.*, *20*, 803–806, 1993a.

Nakamura, T., T. Tsuda, S. Fukao, S. Kato, A. H. Manson, and C. E. Meek, Comparative observations of short-period gravity waves (10–100 min) in the mesosphere in 1989 by Saskatoon MF radar (52°N), Canada and the MU radar (35°N), Japan, *Radio Sci.*, *28*, 729–746, 1993b.

Nakamura, T., T. Tsuda, M. Yamamoto, S. Fukao, and S. Kato, Characteristics of gravity waves in the mesosphere observed with the middle and upper atmosphere radar, 1, Momentum flux, *J. Geophys. Res.*, *98*, 8899–8910, 1993c.

Nakamura, T., T. Tsuda, M. Yamamoto, S. Fukao, and S. Kato, Characteristics of gravity waves in the mesosphere observed with the middle and upper atmosphere radar, 2, Propagation direction, *J. Geophys. Res.*, *98*, 8911–8923, 1993d.

Nakamura, T., T. Tsuda, S. Fukao, A. H. Manson, C. E. Meek, R. A. Vincent, and I. M. Reid, Mesospheric gravity waves at Saskatoon (52°N), Kyoto (35°N), and Adelaide (35°S), *J. Geophys. Res.*, *101*, 7005–7012, 1996.

Nakamura, T., D. C. Fritts, J. R. Isler, T. Tsuda, and R. A. Vincent, Short-period fluctuations of the diurnal tide observed with low-latitude MF and meteor radars during CADRE: Evidence for gravity wave/tidal interactions, *J. Geophys. Res.*, *102*, 26,225–26,238, 1997.

Nakamura, T., T. Tsuda, H. Miyagawa, Y. Matsushita, H. Fukunishi, Y. Takahashi, and Y. Yamada, Propagation directions of gravity wave patterns observed in OH CCD images during the SEEK campaign, *Geophys. Res. Lett.*, *25*, 1793–1796, 1998.

Nance, L. B., and D. D. Durran, A modeling study of nonstationary trapped mountain lee waves, part II, Nonlinearity, *J. Atmos. Sci.*, *55*, 1429–1445, 1998.

Nastrom, G. D., and D. C. Fritts, Sources of mesoscale variability of gravity waves, part I, Topographic excitation, *J. Atmos. Sci.*, *49*, 101–110, 1992.

Nastrom, G. D., and K. S. Gage, A climatology of atmospheric wavenumber spectra of wind and temperature observed by commercial aircraft, *J. Atmos. Sci.*, *42*, 950–960, 1985.

Nastrom, G. D., B. B. Balsley, and D. A. Carter, Mean meridional winds in the mid- and high-latitude summer mesosphere, *Geophys. Res. Lett.*, *9*, 139–142, 1982.

Nastrom, G. D., D. C. Fritts, and K. S. Gage, An investigation of terrain effects on the mesoscale spectrum of atmospheric motions, *J. Atmos. Sci.*, *44*, 3087–3096, 1987.

Nastrom, G. D., T. E. VanZandt, and J. M. Warnock, Vertical wavenumber spectra of wind and temperature from high-resolution balloon soundings in the lower atmosphere over Illinois, *J. Geophys. Res.*, *102*, 6685–6702, 1997.

Nissen, K. M., P. Braesicke, and U. Langematz, QBO, SAO, and tropical waves in the Berlin TSM GCM: Sensitivity to radiation, vertical resolution, and convection, *J. Geophys. Res.*, *105*, 24,771–24,790, 2000.

Norton, W. A., and J. Thuburn, The two-day wave in a middle atmosphere GCM, *Geophys. Res. Lett.*, *23*, 2113–2116, 1996.

Norton, W. A., and J. Thuburn, The mesosphere in the extended UGAMP GCM, in *Gravity Wave Processes: Their Parameterization in Global Climate Models, NATO ASI Ser., Ser. I*, vol. *50*, edited by K. Hamilton, pp. 383–401, Springer-Verlag, New York, 1997.

Norton, W. A., and J. Thuburn (Eds.), *Sensitivity of mesospheric mean flow, planetary waves, and tides to strength of gravity wave drag*, *J. Geophys. Res.*, *104*, 30,897–30,911, 1999.

Ogino, S., M. D. Yamanaka, and S. Fukao, Meridional variation of lower stratospheric gravity wave activity: A quick look at Hakuho-Maru J-COARE cruise rawinsonde data, *J. Meteorol. Soc. Jpn.*, *73*, 407–413, 1995.

O'Sullivan, D., and T. J. Dunkerton, Seasonal development of the extratropical QBO in a numerical model of the middle atmosphere, *J. Atmos. Sci.*, *51*, 3706–3721, 1994.

O'Sullivan, D., and T. J. Dunkerton, Generation of inertia-gravity waves in a simulated life cycle of baroclinic instability, *J. Atmos. Sci.*, *52*, 3695–3716, 1995.

Palmer, T. L., D. C. Fritts, Ø. Andreassen, and I. Lie, Three-dimensional evolution of Kelvin-Helmholtz billows in stratified compressible flow, *Geophys. Res. Lett.*, *21*, 2287–2290, 1994.

Palmer, T. L., D. C. Fritts, and Ø. Andreassen, Evolution and breakdown of Kelvin-Helmholtz billows in stratified compressible flows, II, Instability structure, evolution, and energetics, *J. Atmos. Sci.*, *53*, 3192–3212, 1996.

Palmer, T. N., G. J. Shutts, and R. Swinbank, Alleviation of a systematic westerly bias in general circulation and numerical weather prediction models through an orographic gravity wave drag parameterization, *Q. J. R. Meteorol. Soc.*, *112*, 1001–10039, 1986.

Pandya, R. E., and M. J. Alexander, Linear stratospheric gravity waves above convective thermal forcing, *J. Atmos. Sci.*, *56*, 2434–2446, 1999.

Pawson, S., Effects of gravity wave drag in the Berlin troposphere-stratosphere-mesosphere GCM, in *Gravity Wave Processes: Their Parameterization in Global Climate Models, NATO ASI Ser., Ser. I*, vol. 50, edited by K. Hamilton, pp. 327–336, Springer-Verlag, New York, 1997.

Pfister, L., W. Starr, R. Craig, and M. Loewenstein, Small-scale motions observed by aircraft in the tropical lower stratosphere: Evidence for mixing and its relationship to large-scale flows, *J. Atmos. Sci.*, *43*, 3210–3225, 1986.

Pfister, L., S. Scott, M. Loewenstein, S. Bowen, and M. Legg, Mesoscale disturbances in the tropical stratosphere excited by convection: Observations and effects on the stratospheric momentum budget, *J. Atmos. Sci.*, *50*, 1058–1075, 1993a.

Pfister, L., K. R. Chan, T. P. Bui, S. Bowen, M. Legg, B. Gary, K. Kelly, M. Proffitt, and W. Starr, Gravity waves generated by a tropical cyclone during the STEP tropical feild pro-

gram: A case study, *J. Geophys. Res.*, *98*, 8611–8638, 1993b.
Piani, C., D. Durran, M. J. Alexander, and J. R. Holton, A numerical study of three-dimensional gravity waves triggered by deep tropical convection and their role in the dynamics of the QBO, *J. Atmos. Sci.*, *57*, 3689–3702, 2000.
Picard, R. H., R. R. O'Neil, H. A. Gardiner, J. Gibson, J. R. Winick, W. O. Gallery, A. T. Stair Jr., P. P. Wintersteiner, E. R. Hegblom, and E. Richards, Remote sensing of discrete stratospheric gravity-wave structure at 4.3-micron from the MSX satellite, *Geophys. Res. Lett.*, *25*, 2809–2812, 1998.
Preusse, P., B. Schaeler, J. T. Bacmeister, and D. Offermann, Evidence for gravity waves in CRISTA temperatures, *Adv. Space Res.*, *24*, 1601–1604, 1999.
Preusse, P., S. D. Eckermann, and D. Offermann, Comparison of global distributions of zonal-mean gravity wave variance inferred from different satellite instruments, *Geophys. Res. Lett.*, *27*, 3877–3880, 2000.
Preusse, P., A. Drnbrack, S. D. Eckermann, M. Riese, B. Schaeler, J. T. Bacmeister, D. Broutman, and K. U. Grossmann, Space-based measurements of stratospheric mountain waves by CRISTA, 1, Sensitivity, analysis method, and a case study, *J. Geophys. Res.*, *107*(D23), 8178, doi:10.1029/2001JD000699, 2002.
Quinn, E. P., and R. H. Holzworth, Quasi-Lagrangian measurements of density surface fluctuations and power spectra in the stratosphere, *J. Geophys. Res.*, *92*, 10,926–10,932, 1987.
Ray, E. A., M. J. Alexander, and J. R. Holton, An analysis of the structure and forcing of the equatorial semiannual oscillation in zonal wind, *J. Geophys. Res.*, *103*, 1759–1774, 1998.
Reeder, M. J., and M. Griffiths, Stratospheric inertia-gravity waves generated in a numerical model of frontogenesis, II, Wave sources and generation mechanisms, *Q. J. R. Meteorol. Soc.*, *122*, 1175–1195, 1996.
Reid, I. M., and R. A. Vincent, Measurements of mesospheric gravity wave momentum fluxes and mean flow accelerations at Adelaide, Australia, *J. Atmos. Terr. Phys.*, *49*, 443–460, 1987.
Reid, I. M., R. Rüster, P. Czechowsky, and G. Schmidt, VHF radar measurements of momentum flux in the summer polar mesosphere over Andenes (69°N, 16°E), Norway, *Geophys. Res. Lett.*, *15*, 1263–1266, 1988.
Ricciardulli, L., and R. R. Garcia, The excitation of equatorial waves by deep convection in the NCAR Community Climate Model (CCM3), *J. Atmos. Sci.*, *57*, 3461–3487, 2000.
Riggin, D. M., D. C. Fritts, C. D. Fawcett, E. Kudeki, and M. H. Hitchman, Radar observations of gravity waves over Jicamarca, Peru during the CADRE campaign, *J. Geophys. Res.*, *102*, 26,263–26,281, 1997.
Rind, D., R. Suozzo, and N. K. Balachandran, The GISS global climate-middle atmosphere model, part II, Model variability due to interactions between planetary waves, the mean circulation and gravity wave drag, *J. Atmos. Sci.*, *45*, 371–386, 1988.
Roble, R. G., and E. C. Ridley, A thermosphere-ionosphere-mesosphere- electrodynamics general circulation model (time-GCM): Equinox solar cycle minimum simulations (30–500 km), *Geophys. Res. Lett.*, *21*, 417–420, 1994.
Rocken, C., et al., Analysis and validation of GPS/MET data in the neutral atmosphere, *J. Geophys. Res.*, *102*, 29,849–29,866, 1997.
Rosenlof, K. H., Seasonal cycle of the residual mean meridional circulation in the stratosphere, *J. Geophys. Res.*, *100*, 5173–5191, 1995.
Rosenlof, K. H., Summer hemisphere differences in temperature and transport in the lower stratosphere, *J. Geophys. Res.*, *101*, 19,129–19,136, 1996.
Salby, M. L., and R. R. Garcia, Transient response to localized episodic heating in the tropics, part I, Excitation and short-time near-field behaviour, *J. Atmos. Sci.*, *44*, 458–498, 1987.
Sassi, F., and R. R. Garcia, A one-dimensional model of the semiannual oscillation driven by convectively forced gravity waves, *J. Atmos. Sci.*, *51*, 3167–3182, 1994.
Sassi, F., and R. R. Garcia, The role of equatorial waves forced by convection in the tropical semiannual oscillation, *J. Atmos. Sci.*, *54*, 1925–1942, 1997.
Sassi, F., R. R. Garcia, and B. A. Boville, The stratopause semiannual oscillation in the NCAR Community Climate Model, *J. Atmos. Sci.*, *50*, 3608–3624, 1993.
Sato, K., Vertical wind disturbances in the afternoon of midsummer revealed by the MU radar, *Geophys. Res. Lett.*, *19*, 1943–1946, 1992.
Sato, K., Small-scale wind disturbances observed by the MU radar during the passage of typhoon Kelly, *J. Atmos. Sci.*, *50*, 518–537, 1993.
Sato, K., A statistical study of the structure, saturation and sources of inertio-gravity waves in the lower stratosphere observed with the MU radar, *J. Atmos. Terr. Phys.*, *56*, 755–774, 1994.
Sato, K., and T. J. Dunkerton, Estimates of momentum flux associated with equatorial Kelvin and gravity waves, *J. Geophys. Res.*, *102*, 26,247–26,261, 1997.
Sato, T., and R. F. Woodman, Fine altitude resolution radar observations of stratospheric turbulent layers by the Arecibo 430 MHz radar, *J. Atmos. Sci.*, *39*, 2546–2552, 1982.
Sato, K., and M. Yamada, Vertical structure of atmospheric gravity waves revealed by the wavelet analysis, *J. Geophys. Res.*, *99*, 20,623–20,631, 1994.
Sato, K., F. Hasegawa, and I. Hirota, Short-period disturbances in the equatorial lower stratosphere, *J. Meteorol. Soc. Jpn.*, *72*, 859–872, 1994.
Sato, K., H. Hashiguchi, and S. Fukao, Gravity waves and turbulence associated with cumulus convection observed with the UHF/VHF clear-air Doppler radars, *J. Geophys. Res.*, *100*, 7111–7120, 1995.
Sato, K., D. O'Sullivan, and T. J. Dunkerton, Low-frequency inertia-gravity waves in the stratosphere revealed by three-week continuous observation with the MU radar, *Geophys. Res. Lett.*, *24*, 1739–1742, 1997.
Sato, K., T. Kumakura, and M. Takahashi, Gravity waves appearing in a high-resolution GCM simulation, *J. Atmos. Sci.*, *56*, 1005–1018, 1999.
Scaife, A. A., N. Butchart, C. D. Warner, D. Stainforth, and W. Norton, Realistic quasi-biennial osciallations in a simulation of the global climate, *Geophys. Res. Lett.*, *27*, 3481–3484, 2000.
Schar, C., and D. R. Durran, Vortex formation and vortex shedding in continuously stratified flows past isolated topography, *J. Atmos. Sci.*, *54*, 534–554, 1997.
Scheffler, A. O., and C. H. Liu, The effects of Doppler shift on the gravity wave spectra observed by MST radar, *J. Atmos. Terr. Phys.*, *48*, 1225–1231, 1986.
Schoeberl, M. R., Dynamics of the middle atmosphere, *Rev. Geophys.*, *25*, 501–507, 1987.
Scinocca, J. F., and R. Ford, The nonlinear forcing of large-scale internal gravity waves by stratified shear instability, *J. Atmos. Sci.*, *57*, 653–672, 2000.
Scinocca, J. F. and N. A. McFarlane, Anisotropy in the parameterization of drag due to freely propagating gravity waves and low-level dynamics, *Q. J. R. Meteorol. Soc.*, *126*, 2353–2393, 2000.
Senft, D. C., and C. S. Gardner, Seasonal variability of gravity wave activity and spectra in the mesopause region at Urbana, *J. Geophys. Res.*, *96*, 17,229–17,264, 1991.

Shimizu, A., and T. Tsuda, Characteristics of Kelvin waves and gravity waves observed with radiosondes over Indonesia, *J. Geophys. Res.*, *102*, 26,159–26,171, 1997.

Shine, K., Sources and sinks of zonal momentum in the middle atmosphere diagnosed using the diabatic circulation, *Q. J. R. Meteorol. Soc.*, *115*, 265–292, 1989.

Shutts, G. J., Gravity-wave parameterization over complex terrain: The effect of critical level absorption in directional wind shear, *Q. J. R. Meteorol. Soc.*, *121*, 1005–1021, 1995.

Shutts, G. J., Stationary gravity-wave structure in flows with directional wind Shear, *Q. J. R. Meteorol. Soc.*, *124*, 1421–1442, 1998.

Shutts, G. J., M. Kitchen, and P. H. Hoare, A large amplitude gravity wave detected by radiosonde, *Meteorol. Mag.*, *117*, 306–310, 1988.

Sidi, C., J. Lefrere, F. Dalaudier, and J. Barat, An improved atmospheric buoyancy wave model, *J. Geophys. Res.*, *93*, 774–790, 1988.

Siskind, D. E., S. Eckermann, J. McCormack, M. J. Alexander, and J. Bacmeister, Hemispheric differences in the temperature of the summertime stratosphere and mesosphere, *J. Geophys. Res.*, *108*, doi:10.1029/2002JD002095, in press, 2003.

Smith, A. K., Longitudinal variations in mesospheric winds: Evidence for gravity wave filtering by planetary waves, *J. Atmos. Sci.*, *53*, 1156–1173, 1996.

Smith, A. K., and G. P. Brasseur, Numerical simulation of the seasonal variation of mesospheric water vapor, *J. Geophys. Res.*, *96*, 7553–7563, 1991.

Smith, A. K., and L. V. Lyjak, An observational estimate of gravity wave drag from the momentum balance in the middle atmosphere, *J. Geophys. Res.*, *90*, 2233–2241, 1985.

Smith, R. B., A measurement of mountain drag, *J. Atmos. Sci.*, *35*, 1644–1654, 1978.

Smith, R. B., On severe downslope winds, *J. Atmos. Sci.*, *42*, 2597–2603, 1985.

Smith, S. A., D. C. Fritts, and T. E. VanZandt, Comparison of mesospheric wind spectra with a gravity wave model, *Radio Sci.*, *20*, 1331–1338, 1985.

Smith, S. A., D. C. Fritts, and T. E. VanZandt, Evidence for a saturated spectrum of atmospheric gravity waves, *J. Atmos. Sci.*, *44*, 1404–1410, 1987.

Smyth, W. D., Dissipation-range geometry and scalar spectra in sheared stratified turbulence, *J. Fluid Mech.*, *401*, 209–242, 1999.

Smyth, W. D., and J. N. Moum, Anisotropy of turbulence in stably stratified mixing layers, *Phys. Fluids*, *12*, 1343–1362, 2000.

Sonmor, L. J., and G. P. Klaassen, Higher-order resonant instabilities of internal gravity waves, *J. Fluid Mech.*, *324*, 1–23, 1996.

Sonmor, L. J., and G. P. Klaassen, Toward a unified theory of gravity wave stability, *J. Atmos. Sci.*, *54*, 2655–2680, 1997.

Sonmor, L. J., and G. P. Klaassen, Mechanisms of gravity wave focusing in the middle atmosphere, *J. Atmos. Sci.*, *57*, 493–510, 2000.

Staquet, C., and J. Sommeria, Internal waves, turbulence and mixing in stratified flows: A report on Euromech Colloquium 339, *J. Fluid Mech.*, *314*, 349–371, 1996.

Strobel, D. F., J. P. Apruzese, and M. R. Schoeberl, Energy balance constraints on gravity wave induced eddy diffusion in the mesosphere and lower thermosphere, *J. Geophys. Res.*, *90*, 13,067–13,072, 1985.

Strobel, D. F., M. E. Summers, R. M. Bevilacqua, M. T. DeLand, and M. Allen, Vertical constituent transport in the mesosphere, *J. Geophys. Res.*, *92*, 6691–6698, 1987.

Sutherland, B. R., Propagation and reflection of internal waves, *Phys. Fluids*, *11*, 1081–1090, 1999.

Sutherland, B. R., Internal wave reflection in uniform shear, *Q. J. R. Meteorol. Soc.*, *126*, 3255–3287, 2000.

Sutherland, B. R., Finite-amplitude internal wavepacket dispersion and breaking, *J. Fluid Mech.*, *429*, 343–380, 2001.

Swenson, G. R., and P. J. Espy, Observations of two-dimensional airglow structure and Na density from the ALOHA, October 9, 1993, "Storm Flight", *Geophys. Res. Lett.*, *22*, 2845–2848, 1995.

Swenson, G. R., and S. B. Mende, OH emission and gravity waves (including a breaking wave) in all-sky imagery from Bear Lake, UT, *Geophys. Res. Lett.*, *21*, 2239–2242, 1994.

Swenson, G. R., R. Haque, W. Yang, and C. S. Gardner, Momentum and energy fluxes of monochromatic gravity waves observed by an OH imager at Starfire Optical Range, New Mexico, *J. Geophys. Res.*, *104*, 6067–6080, 1999.

Swenson, G. R., M. J. Alexander, and R. Haque, Dispersion imposed limits on atmospheric gravity waves in the mesosphere: Observations from OH airglow, *Geophys. Res. Lett.*, *27*, 875–878, 2000.

Takahashi, M., Simulation of the stratospheric quasi-biennial oscillation using a general circulation model, *Geophys. Res. Lett.*, *23*, 661–664, 1996.

Takahashi, M., The first realistic simulation of the stratospheric quasi-biennial oscillation in a general circulation model, *Geophys. Res. Lett.*, *26*, 1307–1310, 1999.

Takahashi, M., and B. A. Boville, A three-dimensional simulation of the equatorial quasi-biennial oscillation, *J. Atmos. Sci.*, *49*, 1020–1035, 1992.

Takahashi, M., and J. R. Holton, The mean zonal flow response to Rossby wave and gravity wave forcing in the equatorial lower stratosphere: Relationship to the QBO, *J. Atmos. Sci.*, *48*, 2078–2087, 1991.

Taylor, M. J., and F. J. Garcia, A two-dimensional spectral analysis of short period gravity waves imaged in the OI(557.7 nm) and near infrared OH nightglow emissions over Arecibo, Puerto Rico, *Geophys. Res. Lett.*, *22*, 2473–2476, 1995.

Taylor, M. J., P. J. Espy, D. J. Baker, R. J. Sica, P. C. Neal, and W. R. Pendleton Jr., Simultaneous intensity, temperature and imaging measurements of short period wave structure in the OH nightglow emission, *Planet. Space Sci.*, *39*, 1171–1188, 1991.

Taylor, M. J., Y. Y. Gu, X. Tao, C. S. Gardner, and M. B. Bishop, An investigation of intrinsic gravity wave signatures using coordinated lidar and nightglow image measurements, *Geophys. Res. Lett.*, *22*, 2853–2856, 1995.

Taylor, M. J., W. R. Pendleton Jr., S. Clark, H. Takahashi, D. Gobbi, and R. A. Goldberg, Image measurements of short-period gravity waves at equatorial latitudes, *J. Geophys. Res.*, *102*, 26,283–26,299, 1997.

Taylor, M. J., S. H. Seo, T. Nakamura, T. Tsuda, H. Fukunishi, and Y. Takahashi, Long base-line measurements of short-period mesospheric gravity waves during the SEEK campaign, *Geophys. Res. Lett.*, *25*, 1797–1800, 1998.

Thayaparan, T., W. K. Hocking, and J. MacDougall, Observational evidence of tidal/gravity wave interactions using the UWO 2 MHz radar, *Geophys. Res. Lett.*, *22*, 373–376, 1995.

Thomas, L., I. T. Prichard, and I. Astin, Radar observations of an inertia-gravity wave in the troposphere and lower stratosphere, *Ann. Geophys.*, *10*, 690–697, 1992.

Thompson, R. O. R. Y., Observations of inertial gravity waves in the stratosphere, *Q. J. R. Meteorol. Soc.*, *104*, 691–698, 1978.

Thorpe, S. A., Transitional phenomena and the development of turbulence in stratified fluids: A review, *J. Geophys. Res.*, *92*, 5231–5248, 1987.

Thorpe, S. A., Observations of parametric instability and breaking waves in an oscillating tilted tube, *J. Fluid Mech.*, *261*, 33–45, 1994.

Tsuda, T., T. Inoue, D. C. Fritts, T. E. VanZandt, S. Kato, T. Sato, and S. Fukao, MST radar observations of a saturated gravity wave spectrum, *J. Atmos. Sci.*, *46*, 2440–2447, 1989.

Tsuda, T., S. Kato, T. Yokoi, T. Inoue, M. Yamamoto, T. E. VanZandt, S. Fukao, and T. Sato, Gravity waves in the mesosphere observed with the middle and upper atmosphere radar, *Radio Sci.*, *25*, 1005–1018, 1990a.

Tsuda, T., Y. Murayama, M. Yamamoto, S. Kato, and S. Fukao, Seasonal variation of momentum flux in the mesosphere observed with the MU radar, *Geophys. Res. Lett.*, *17*, 725–728, 1990b.

Tsuda, T., T. E. VanZandt, M. Mizumoto, S. Kato, and S. Fukao, Spectral analysis of temperature and Brunt-Vaisala frequency fluctuations observed by radiosondes, *J. Geophys. Res.*, *96*, 17,265–17,278, 1991.

Tsuda, T., Y. Murayama, K.-I. Oyama, F. J. Schmidlin, M. Bittner, H. Kanzawa, T. Nakamura, M. D. Yamanaka, S. Fukao, and S. Kato, Rocketsonde observations of the middle atmosphere dynamics at Uchinoura (31N, 131E) during the DYANA campaign, part II, Characteristics of gravity waves, *J. Geomag. Geoelectr.*, *44*, 1009–1023, 1992.

Tsuda, T., Y. Murayama, H. Wiryosumarto, S. W. B. Harijono, and S. Kato, Radiosonde obserations of equatorial atmosphere dynamics over Indonesia, 2, characteristics of gravity waves, *J. Geophys. Res.*, *99*, 10,507–10,516, 1994a.

Tsuda, T., Y. Murayama, T. Nakamura, R. A. Vincent, A. H. Manson, C. E. Meek, and R. L. Wilson, Variations of the gravity wave characteristics with height, season, and latitude revealed by comparative observations, *J. Atmos. Terr. Phys.*, *56*, 555–568, 1994b.

Tsuda, T., M. Nishida, and C. Rocken, A global morphology of gravity wave activity in the stratosphere revealed by the GPS occultation data (GPS/MET), *J. Geophys. Res.*, *105*, 7257–7274, 2000.

Vadas, S. L., and D. C. Fritts, Gravity wave radiation and mean responses to local body forces in the atmosphere, *J. Atmos. Sci.*, *58*, 2249–2279, 2001.

Vanneste, J., The instability of internal gravity waves to localized disturbances, *Ann. Geophys.*, *13*, 196–210, 1995.

VanZandt, T. E., A universal spectrum of buoyancy waves in the atmosphere, *Geophys. Res. Lett.*, *9*, 575–578, 1982.

VanZandt, T. E., and D. C. Fritts, A theory of enhanced saturation of the gravity wave spectrum due to increases in atmospheric stability, *Pure Appl. Geophys.*, *130*, 399–420, 1989.

VanZandt, T. E., G. D. Nastrom, and J. L. Green, Frequency spectra of vertical velocity from flatland VHF radar data, *J. Geophys. Res.*, *96*, 2845–2855, 1991.

Vincent, R. A., MF/HF radar measurements of the dynamics of the mesopause region—A review, *J. Atmos. Terr. Phys.*, *46*, 961–974, 1984a.

Vincent, R. A., Gravity wave motions in the mesosphere, *J. Atmos. Terr. Phys.*, *46*, 119–128, 1984b.

Vincent, R. A., and M. J. Alexander, Gravity waves in the tropical lower stratosphere: An observational study of seasonal and interannual variability, *J. Geophys. Res.*, *105*, 17,971–17,982, 2000.

Vincent, R. A., and S. D. Eckermann, VHF radar observations of mesoscale motions in the troposphere: Evidence for gravity wave Doppler shifting, *Radio Sci.*, *25*, 1019–1037, 1990.

Vincent, R. A., and D. C. Fritts, A climatology of gravity waves in the mesosphere and lower thermosphere over Adelaide, Australia, *J. Atmos. Sci.*, *44*, 748–760, 1987.

Vincent, R. A., and I. M. Reid, HF Doppler measurements of mesospheric momentum fluxes, *J. Atmos. Sci.*, *40*, 1321–1333, 1983.

Vincent, R. A., S. J. Allen, and S. D. Eckermann, Gravity-wave parameters in the lower stratosphere, in *Gravity Wave Processes: Their Parameterization in Global Climate Models, NATO ASI Ser., Ser. I,* vol. 50, edited by K. Hamilton, pp. 7–25, Springer-Verlag, New York, 1997.

Wada, K., T. Nitta, and K. Sato, Equatorial inertia-gravity waves in the lower stratosphere revealed by TOGA-COARE IOP data, *J. Meteorol. Soc. Jpn.*, *77*, 721–736, 1999.

Walterscheid, R. L., Inertio-gravity wave induced accelerations of mean flow having an imposed periodic component: Implications for tidal observations in the meteor region, *J. Geophys. Res.*, *86*, 9698–9706, 1981.

Walterscheid, R. L., Propagation of small-scale gravity waves through larg-scale internal wave fields: Eikonal effects at low-frequency approximation critical levels, *J. Geophys. Res.*, *105*, 18,027–18,037, 2000.

Walterscheid, R. L., J. H. Hecht, R. A. Vincent, I. M. Reid, J. Woithe, and M. J. Hickey, Analysis and interpretation of airglow and radar observations of quasi-monochromatic gravity waves in the upper mesosphere and lower thermosphere over Adelaide, Australia (35°S, 138°E), *J. Atmos. Sol. Terr. Phys.*, *61*, 461–478, 1999.

Wang, D.-Y., and D. C. Fritts, Mesospheric momentum fluxes observed by the MST Radar at Poker Flat, Alaska, *J. Atmos. Sci.*, *47*, 1511–1521, 1990.

Wang, D.-Y., and D. C. Fritts, Evidence of gravity wave-tidal interaction observed near the summer mesopause at Poker Flat, Alaska, *J. Atmos. Sci.*, *48*, 572–583, 1991.

Warner, C. D., and M. E. McIntyre, On the propagation and dissipation of gravity wave spectra through a realistic middle atmosphere, *J. Atmos. Sci.*, *53*, 3213–3235, 1996.

Warner, C. D., and M. E. McIntyre, An ultra-simple spectral parameterization for non-orographic gravity waves, *J. Atmos. Sci.*, *58*, 1837–1857, 2001.

Weinstock, J., Nonlinear theory of acoustic-gravity waves, 1, Saturation and enhanced diffusion, *J. Geophys. Res.*, *81*, 633–652, 1976.

Weinstock, J., Nonlinear theory of gravity waves: Momentum deposition, generalized Rayleigh friction, and diffusion, *J. Atmos. Sci.*, *39*, 1698–1710, 1982.

Weinstock, J., Simplified derivation of an algorithm for nonlinear gravity waves, *J. Geophys. Res.*, *89*, 345–350, 1984.

Weinstock, J., Theoretical gravity wave spectra in the atmosphere: Strong and weak interactions, *Radio Sci.*, *20*, 1295–1300, 1985.

Weinstock, J., Saturated and unsaturated spectra of gravity waves and scale-dependent diffusion, *J. Atmos. Sci.*, *47*, 2211–2225, 1990.

Werne, J., and D. C. Fritts, Stratified shear turbulence: Evolution and statistics, *Geophys. Res. Lett.*, *26*, 439–442, 1999.

Werne, J. A., and D. C. Fritts, Turbulence and mixing in a stratified shear layer: 3D K-H simulations at $Re = 24{,}000$, *Phys. Chem. Earth*, *26*, 263–268, 2001.

Whiteway, J. A., and A. I. Carswell, Lidar observations of gravity wave activity in the upper stratosphere over Toronto, *J. Geophys. Res.*, *100*, 14,113–14,124, 1995.

Whiteway, J. A., and T. J. Duck, Evidence for critical level filtering of atmospheric gravity waves, *Geophys. Res. Lett.*, *23*, 145–148, 1996.

Whiteway, J. A., T. J. Duck, D. P. Donovan, J. C. Bird, S. R. Pal, and A. I. Carswell, Measurements of gravity wave activity within and around the Arctic stratospheric vortex, *Geophys. Res. Lett.*, *24*, 1387–1390, 1997.

Whitham, G. B., A general approach to linear and nonlinear dispersive waves using a Lagrangian, *J. Fluid Mech.*, *22*, 273–283, 1965.

Whitham, G. B., *Linear and Nonlinear Waves*, John Wiley, New York, 1974.

Williams, B. P., M. A. White, D. A. Krueger, and C. Y. She, Observation of a large amplitude wave and inversion layer leading to convective instability in the mesopause region

over Fort Collins, CO (41°N, 105°W), *Geophys. Res. Lett.*, *29*(17), 1850, doi:10.1029/2001GL014514, 2002.

Wilson, R., M.-L. Chanin, and A. Hauchecorne, Gravity waves in the middle atmosphere observed by Rayleigh lidar, 1, Case studies, *J. Geophys. Res.*, *96*, 5153–5167, 1991a.

Wilson, R., M.-L. Chanin, and A. Hauchecorne, Gravity waves in the middle atmosphere observed by Rayleigh lidar, part 2, Climatology, *J. Geophys. Res.*, *96*, 5169–5183, 1991b.

Winters, K. B., and E. A. D Asaro, 3D wave breaking near a critical level, *J. Fluid Mech.*, *272*, 255–284, 1994.

Winters, K. B, and J. J. Riley, Instability of internal waves near a critical level, *Dyn. Atmos. Oceans*, *16*, 249–278, 1992.

Woodman, R. F., B. B. Balsley, F. Aquino, L. Flores, E. Vazquez, M. Sarango, M. M. Huaman, and H. Soldi, First observations of polar summer mesosphere echoes in Antarctic, *J. Geophys. Res.*, *104*, 22,577–22,590, 1999.

Worthington, R. M., and L. Thomas, Radar measurements of critical-layer absorption in mountain waves, *Q. J. R. Meteorol. Soc.*, *122*, 1263–1282, 1996.

Wu, D. L., and J. W. Waters, Satellite observations of atmospheric variances: A possible indication of gravity waves, *Geophys. Res. Lett.*, *23*, 3631–3634, 1996a.

Wu, D. L., and J. W. Waters, Gravity-wave-scale temperature fluctuations seen by the UARS MLS, *Geophys. Res. Lett.*, *23*, 3289–3292, 1996b.

Wu, D. L., and J. W. Waters, Observations of gravity waves with the UARS Microwave Limb Sounder, in *Gravity Processes: Their Parameterization in Global Climate Models, NATO ASI Ser., Ser. I*, vol. 50, edited by K. Hamilton, pp. 103–120, Springer-Verlag, New York, 1997.

Wu, Y.-F., and H.-U. Widdel, Observational evidence of a saturated gravity wave spectrum in the mesosphere, *J. Atmos. Terr. Phys.*, *51*, 991–996, 1989.

Wu, Y.-F., and H.-U. Widdel, Spectral analysis of atmospheric velocity fluctuations in the mesosphere, *J. Atmos. Terr. Phys.*, *52*, 23–33, 1990.

Wu, Y.-F., and H.-U. Widdel, Further study of a saturated gravity wave spectrum in the mesosphere, *J. Geophys. Res.*, *96*, 9263–9272, 1991.

Yamada, Y., H. Fukunishi, T. Nakamura, and T. Tsuda, Breakdown of small-scale quasi-stationary gravity wave and transition to turbulence observed in OH airglow, *Geophys. Res. Lett.*, *28*, 2153–2156, 2001.

Yamamoto, M., T. Tsuda, S. Kato, T. Sato, and S. Fukao, A saturated inertia gravity wave in the mesosphere observed by the middle and upper atmosphere radar, *J. Geophys. Res.*, *92*, 11,993–11,999, 1987.

Yeh, K. C., and B. Dong, The nonlinear interaction of a gravity wave with the vortical modes, *J. Atmos. Terr. Phys.*, *51*, 45–50, 1989.

Yeh, K. C., and C. H. Liu, The instability of atmospheric gravity waves through wave-wave interactions, *J. Geophys. Res.*, *86*, 9722–9728, 1981.

Yoshiki, M., and K. Sato, A statistical study of gravity waves in the polar regions based on operational radiosonde data, *J. Geophys. Res.*, *105*, 17,995–18,011, 2000.

Yuan, L., and D. C. Fritts, Influence of a mean shear on the dynamical instability of an inertio-gravity wave, *J. Atmos. Sci.*, *46*, 2562–2568, 1989.

Zhong, L., L. J. Sonmor, A. H. Manson, and C. E. Meek, The influence of time-dependent wind on gravity-wave propagation in the middle atmosphere, *Ann. Geophys.*, *13*, 375–394, 1995.

Zhu, X., A new theory of saturated gravity wave spectrum in the middle atmosphere, *J. Atmos. Sci.*, *51*, 3615–3626, 1994.

Zhu, X., and J. R. Holton, Mean fields induced by local gravity-wave forcing in the middle atmosphere, *J. Atmos. Sci.*, *44*, 620–630, 1987.

Zhu, X., P. K. Swaminathan, J.-H. Yee, D. F. Strobel, and D. Anderson, A globally balanced two-dimensional middle atmosphere model: dynamical studies of mesopause meridional circulation and stratosphere-mesosphere exchange, *J. Geophys. Res.*, *102*, 13,095–13,112, 1997.

Zink, F., and R. A. Vincent, Wavelet analysis of stratospheric gravity wave packets over Macquarie Island, 1, Wave parameters, *J. Geophys. Res.*, *106*, 10,275–10,288, 2001a.

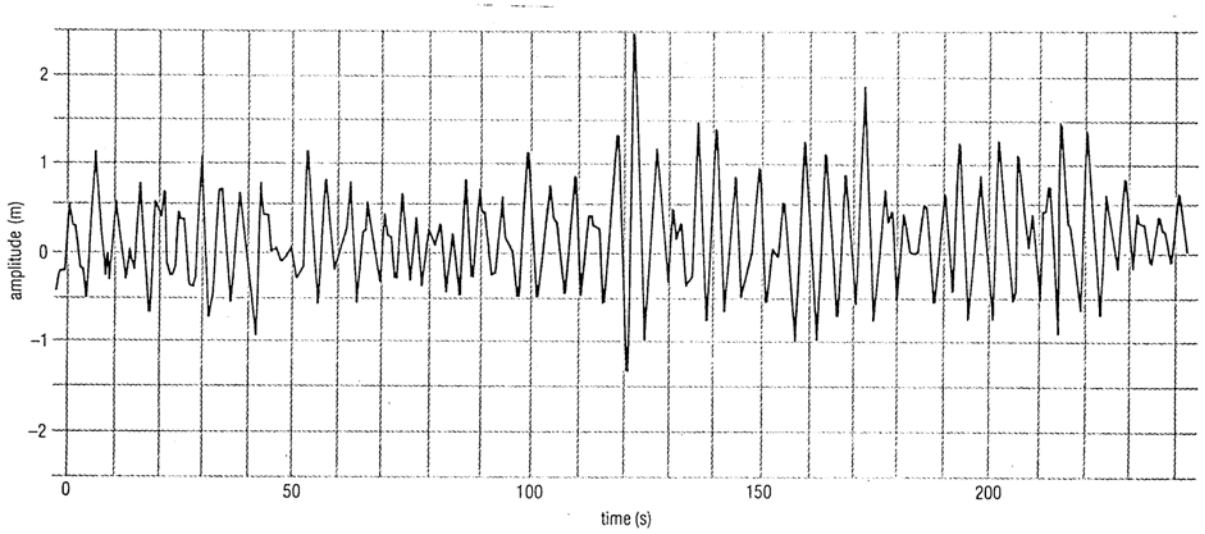
Zink, F., and R. A. Vincent, Wavelet analysis of stratospheric gravity wave packets over Macquarie Island, 2, Intermittency and mean-flow accelerations, *J. Geophys. Res.*, *106*, 10,289–10,297, 2001b.

M. J. Alexander and D. C. Fritts, Colorado Research Associates Division, Northwest Research Associates, 3380 Mitchell Lane, Boulder, CO 80301, USA. (dave@co-ra.com)

Surface Gravity Wave Generation

A typical wave record is shown below

Figure 1.5 A typical wave record, i.e. a record of variation in water level with time at one position.

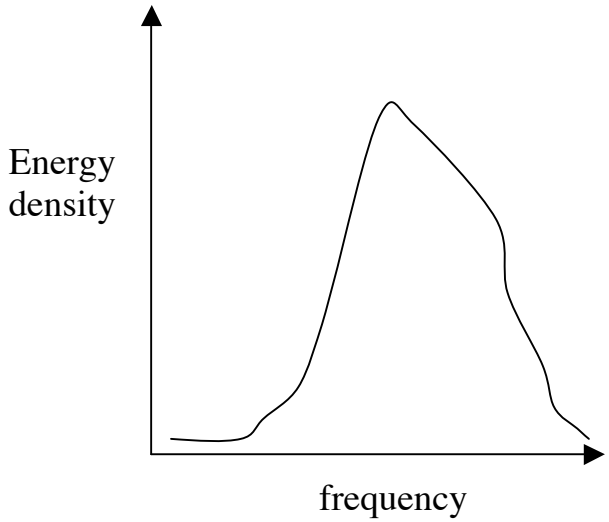


(from Waves, Tides and Shallow Water Processes).

From the record, we can see that the wave cannot be simply described as one sinusoid. Instead, it is characterized by many waves with different periods and phases. The spectrum of energy is usually plotted as energy density, (unit of energy/unit frequency interval, *Hz*). The energy density is given by the amount of energy in a particular frequency interval

$$\frac{1}{2}\rho g a^2 / Hz$$

To quantify the sea state, we make a plot of the energy density as a function of frequency.



Another way we can describe the waves, or the *sea state*, is with the *significant wave height*. The significant wave height $H_{1/3}$ is the average height of the highest 1/3 of all waves observed in a given period of time. This is the average height given by an experienced observer. The Beaufort scale gives a relationship between the wave height and the wind speed, so given the observed significant wave height, one can infer the wind speed.

Table 1.1 A selection of information from the Beaufort Wind Scale.

Beaufort No.	Name	Wind speed knots	Wind speed ms^{-1}	State of the sea-surface	Wave height* (m)
0	Calm	<1	0.0–0.2	Sea like a mirror.	0
1	Light air	1–3	0.3–1.5	Ripples with appearance of scales; no foam crests.	0.1–0.2
2	Light breeze	4–6	1.6–3.3	Small wavelets; crests have glassy appearance but do not break.	0.3–0.5
3	Gentle breeze	7–10	3.4–5.4	Large wavelets; crests begin to break; scattered white horses.	0.6–1.0
4	Moderate breeze	11–16	5.5–7.9	Small waves, becoming longer; fairly frequent white horses.	1.5
5	Fresh breeze	17–21	8.0–10.7	Moderate waves taking longer form; many white horses and chance of some spray.	2.0
6	Strong breeze	22–27	10.8–13.8	Large waves forming; white foam crests extensive everywhere and spray probable.	3.5
7	Moderate gale	28–33	13.9–17.1	Sea heaps up and white foam from breaking waves begins to be blown in streaks; spindrift begins to be seen.	5.0
8	Fresh gale	34–40	17.2–20.7	Moderately high waves of greater length; edges of crests break into sprindrift; foam is blown in well-marked streaks.	7.5
9	Strong gale	41–47	20.8–24.4	High waves; dense streaks of foam; sea begins to roll; spray may affect visibility.	9.5
10	Whole gale	48–55	24.5–28.4	Very high waves with overhanging crests; sea-surface takes on white appearance as foam in great patches is blown in very dense streaks; rolling of sea is heavy and visibility reduced.	12.0
11	Storm	56–64	28.5–32.7	Exceptionally high waves; sea covered with long white patches of foam; small and medium-sized ships might be lost to view behind waves for long times; visibility further reduced.	15.0
12	Hurricane	>64	>32.7	Air filled with foam and spray; sea completely white with driving spray; visibility greatly reduced.	>15

*$H_{1/3}$, i.e. the significant wave height.

11

While it is fairly intuitive how strong winds can generate larger waves, it is not as clear how the strength of the wind influences the dominant frequency of the waves, and thus the wave spectrum. To understand this, it is important to think about how waves grow. Initial disturbances are due to turbulent pressure fluctuations on small capillary waves.

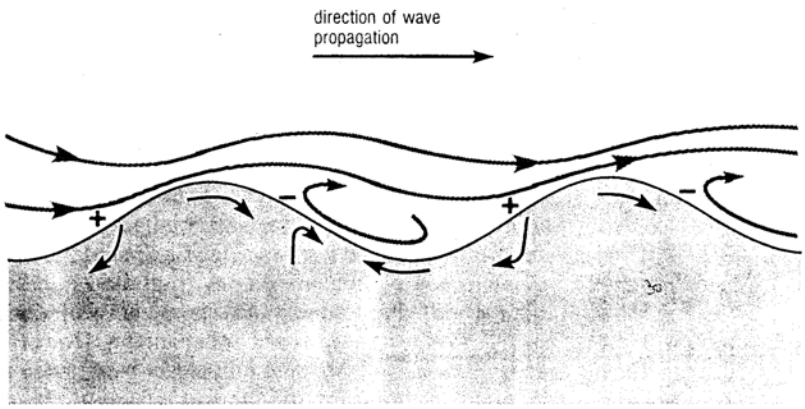


Figure 1.3 Jeffreys' 'sheltering' model of wave generation. Curved lines indicate air flow; short, straight arrows show water movement, which will be explained more fully in Section 1.2.1. The rear face of the wave against which the wind blows experiences a higher pressure than the front face, which is sheltered from the force of the wind. Air eddies are formed in front of each wave, leading to differences in air pressure. The excesses and deficiencies of pressure are shown by plus and minus signs respectively. The pressure difference pushes the wave along.

As in flow over topography, the wave causes air pressure fluctuations and eddies to form on the downwind side of the wave. Higher pressure on the upwind side of the wave creates a pressure gradient force on the wave and causes them to grow. As the waves grow, they steepen and non-linear effects become increasingly important and the waves break. In *deep water*, wave breaking tends to happen when $a/\Lambda = 1/12$.

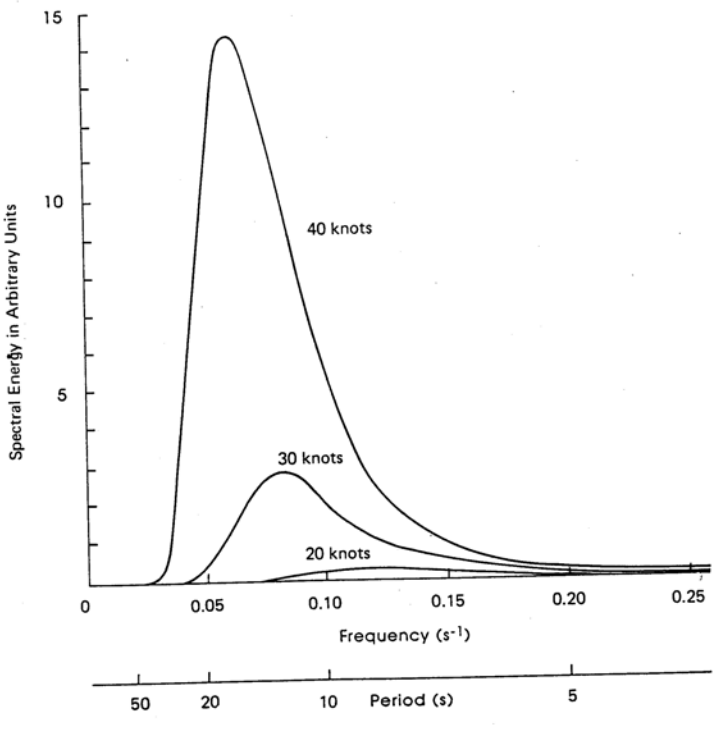
Now how does the speed of the wind effect the wave spectrum? First we need a definition for a *fully developed sea*. A fully developed sea occurs when the input by the wind is balanced by breaking of the waves.

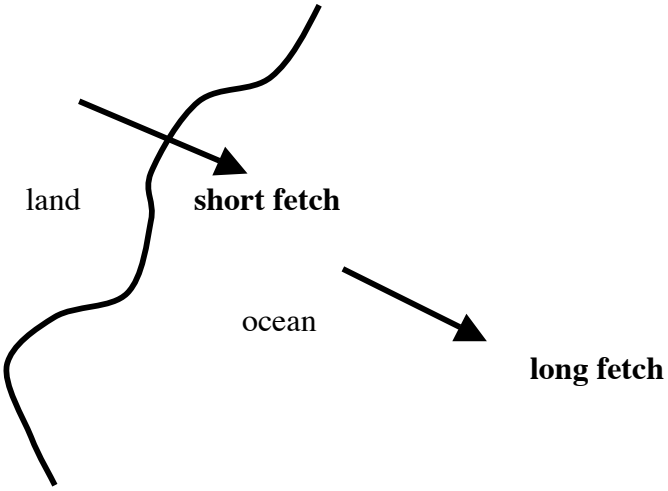
Wind energy in → Waves → Wave energy out (breaking)

The wind must blow as fast, if not faster, than the phase speed C of the wave to impart energy to it, where

$$C = \frac{\omega}{k} = \frac{\omega}{\omega^2/g} = \frac{g}{\omega} = \frac{gT}{2\pi}$$

Therefore, for a fully developed sea, the greater the wind speed, the longer the period of the most energetic waves (the peak in the spectrum). Also, since $\omega^2 = gk$, the longer the period T, the lower the frequency, and the longer the wavelength.





However, a fully developed sea does not always occur. Whether it does depends on the *fetch* (the distance over which the wind blows) and the *duration* (the length of time the wind blows) of the storm. *Young seas,* those with limited duration or fetch will tend to be peaked at shorter wavelengths. Waves tend to grow exponentially so that the seas quickly come into equilibrium with the winds, which is why the Beaufort scale is useful. See figure 9.12 in Knauss that shows the minimum fetch and duration required for a fully developed sea as a function of wind speed. In general, the *stronger the wind, the larger fetch and duration is needed for a fully developed sea to form.* Figure 9.13 shows the spectrum for a fully developed sea and for ones with more limited fetches.

Aside. The speed with which the ocean waves come into equilibrium with near-surface winds makes possible the measurement of winds using satellite-mounted radars called "scatterometers." The radar signal is scattered back from the ocean surface by waves with wavelengths of a few centimeters – their resonant backscatter, called "Bragg scatter", is used to estimate the speed of the wind that produces the short waves. By varying the look angle of the radar, wind direction can also be obtained.

Example: A storm is generated offshore of Washington. We measure the energy spectrum of surface gravity waves at Astoria on February 1 at noon and then again on February 3 at noon.

a. The peak energy occurs at a period of 12s on February 1, and 6 s on February 3. What is the wavelength for each of those waves?
We know that

$$\omega^2 = gk$$

Rearranging we have

$$\omega^2 = \frac{g2\pi}{\Lambda}$$

$$\Lambda = \frac{g2\pi}{\omega^2} = \frac{gT^2}{2\pi}$$

So plugging in for T we have 224 m for the 12 s wave and 56 m for the 6 s wave

b. What is the phase velocity and group velocity for each of the waves found in b?

$$C = \sqrt{g/k} = \sqrt{g\Lambda/2\pi}$$

18.7 m/s for the 12 s wave
9.35 m/s for the 10 s wave
The group velocity will be half that: 9.35 m/s and 4.68 m/s respectively.

c. Estimate *when* the storm occurred. Let d be the offshore distance (unknown, but the same for both cases)

$$d = C_1 t_1$$

$$d = C_2 t_2$$

and

$$t_2 = t_1 + 48 hours$$

where C_1 and C_2 are the group velocity for the 12s and 6 s waves, respectively. Then

$$C_1 t_1 = C_2 t_2 = C_2 (t_1 + 48 hours)$$

$$(C_1 - C_2) t_1 = C_2 48 hours$$

$$t_1 = \frac{C_2}{(C_1 - C_2)} 48 hours = 48 hours$$

d. Estimate *where* the storm occurred.

$$d = C_1 t_1 = 9.35 * 48 * 60 * 60$$

$$d = 1616\ km$$

© 2005,Susan Hautala, LuAnne Thompson, and Kathryn Kelly. All rights reserved.

Tests on Superconductor Gravitational Effects

by Alexander V. Frolov

1. Theoretical background

The high density fluctuations in Bose condensate is laboratory scale case to confirm theoretical conclusions about gravitational field cut-off frequency [1]. This experimental work was organized to examine if some resonance effects in 10-100 MHz range can be detected as mass (weight) anomalies.

Alexander V. Frolov is owner and director of research firm Faraday Lab, Russia, since 2001. He develops alternative energy and advanced aerospace projects. The publication of this article can be useful to find partners and investors for next stage of experiments. Contact email a2509@yahoo.com

2. Previous experimental data

Basically experimental approach in this area of research was described in [2] by Podkletnov:

1. Superconductor material was YBa2Cu2O7-x disk of 145 mm diameter and 6 mm high.
2. The effect is detected as 0.05% - 0.07% mass (weight) changes. It was detected for the case of non-rotating high temperature conductivity superconductor (HTCS) disk, which is leviting in 50-106Hz EM field.
3. Rotation of the disk increase the effect. Important fact is that during change of rotation velocity of the HTCS disk the effect was about 2-4% that is maximum data.
4. In the case of constant rotation velocity (about 5,000 rpm) the effect was detected about 0.3% - 0.5% change of the weight.
5. The effect was detected also for the case of rotation of the HTCS disk after EM field of solenoids was off.

Other experiment of 1995 was described by Podkletnov [3]. Two-phase material of the disk: in the operational mode the upper layer of the disk is superconductive the layer below is not superconductive. So, in this case the area of the phase transition between two layers was created especially.

One more important step in understanding of the effect was made by G. Modanese [4], who assumed that mechanical rotation of the HTCS disk produce motion of Bose condensate like electric current in HTCS material.

Next experiment by Podkletnov and Modanese was described as “impulse gravity generator” [5]. They used 50,000 A and 1MV electric discharge onto HTCS target to create non-dissipative “force beam” or gravity wave. This experiment is new step in understanding of process since

Podkletnov's idea of "shielding of gravity" is changed to conception of force action against gravity. This action is possible as result of artificial gravitational wave or impulse.

Static tests mainly were not effective but important example is experiment by John Schnurer [6]. The effect was detected in the case on non-rotating HTCS disk, which was leviting above permanent magnet. The effect was detected only during change of HTCS material phase from superconductor to non-superconductor phase (heating above Tk). This phase transition usually takes several seconds (2-3 seconds) when the effect can be detected.

Experiment with rotating HTCS ring described in [8] is example of gravimagnetic field produced by spinning superconductor. The results were presented at a one-day conference at ESA's European Space and Technology Research Centre (ESTEC), in the Netherlands , 21 March 2006. This experiment is the gravitational analogue of Faraday's electromagnetic induction experiment in 1831.

One more important aspect of experimenting with HTCS materials is their low temperature in superconductive state. This temperature is much low than environmental temperature and by this way the intensive heat transport present in all experiments. For the case of precise measurements flows of air produced by the temperature difference can be screened but there is aspect named as thermogravitation. For example, in Dotto ring [9] experiment it was demonstrated that intensive heat transfer along the ring produce gravimagnetic effects. For present experimental task it is not critical aspect since this effect is static i.e. it produce permanent force.

3. Related theories

Analysis of previous experimental and theoretical data allows to assume that gravity related effects resulting from changes of density of the Bose condensate. Maximum effect can be estimated for the case of correct frequency of oscillations of the external field, which resonate with natural high-density fluctuations in Bose condensate. In the case of correct frequency we can estimate full compensation of natural gravity field. Assume that the natural gravity field is not single frequency oscillation process. Complex frequency structure of the natural gravity field requires determine several main resonance frequencies to obtain full compensation.

Single phase transition in HTCS material [6] also is the case of change of the Bose condensate density (from maximum value to zero). Since in this case the change of the phase from superconductor to non-superconductor is gradual then the effects is weak and detected during several seconds. Experiment described in [5] is one of methods to produce rapid change of the phase in all Bose condensate of the HTCS target to create short but powerful gravitational pulse.

The nature of this gravitational pulse can be described here as longitudinal wave in aether. By this approach we can see analogy with Tesla experiments. Also we can see that impulse gravity generator by Podkletnov and Modanese [5] is development of Morton beam generator [7], which used electric spark between charged ball and metal plate to produce "Morton force beam". Powerful force effect in HTCS case [5] can be explained by coherent behavior of Bose condensate that produce "laser effect" since it is similar to coherent photon emission in laser.

Analysis of experimental data allows assume that:

1. Bose condensate currents in stationary HTCS disk involve aether in motion relative lattice in matter of the HTCS disk. This **relative motion** generate gravimagnetic field, which is responsible for the weight changes.

2. Rotation of the HTCS disk with Bose condensate currents produce more powerful effects due to **increase of the relative velocity** between Bose condensate and lattice of the matter of the disk.

3. **Change of rotation velocity** of the HTCS disk (i.e. **deceleration or acceleration**) produce maximum relative velocity between matter lattice and moving Bose condensate due to its inertial properties. Here is clear analogy with electromagnetic induction effects.

4. **Phase transition** from HTCS condition to non-superconductive phase is the case of maximum change of the Bose condensate density (zero – maximum or maximum – zero) that is responsible for generation of single **aether density wave.**

5. Special material of the HTCS disk [3] with **two-phase layers** demonstrated more powerful effects. It can be explained since in such disk there is **boundary layer** between superconductive and non-superconductive areas. External electromagnetic field make this layer produce **high frequency phase transitions** that generate high frequency aether density waves.

6. Impulse gravity generator [5] produce more powerful effect than Morton beam generator due to coherent behavior of the Bose condensate. The nature of this effect is the same **longitudinal aether wave** discovered by Tesla.

7. Experimental data from ESTEC [8] is confirmation of the above suggestions. Any rotation of mass by Einstein produce **gravimagentic field** that can be explained by aether dynamical theory. Spinning superconductor produces more powerful effect than non-superconductive matter due to physical properties of Bose condensate.

4. Conclusions

Considering matter as vortexes in aether it is possible to explain most of gravitational phenomenon. For example, inertial property of mass can be explained by behavior of aether, which is connected with this mass. Old theory [10] by Fatio (1690) and Le Sage (1700) can be confirmed in modern experiments. So called “gravitational waves” or “gravitational impulses” can be created as aether density waves, which are longitudinal waves. Bose condensate in superconductor can be presented as special physical state of matter when matter is connected with aether in different extent than usually. Phase transition of matter between superconductive state and non-superconductive state release or connect (join) some amount of aether and this phase transitions can be organized with high frequency to generate high frequency aether density waves. In the case of resonance (predicted in [1] range of frequencies is 10-100MHz) compensation of natural gravity forces can be obtained experimentally.

5. **Organization of tests**

HTCS disk was ordered from CAN superconductor producer [11]. Material is melt textured $YBa_2Cu_3O_{7-x}$ with Y_2BaCuO_5 excess. Critical temperature 90 K. Diameter 56 mm height 16 mm , Fig.1.

Fig.1 HTCS disk

Cooling of the HTCS was organized by liquid nitrogen, the HTCS disk was placed in plastic tank and immersed in nitrogen vapors, Fig.2.

Fig.2 One of the plastic tanks with the HTCS disk

Fig. 3 Cooling by liquid nitrogen

Detection of weight changes was made by digital scales HL-100 with accuracy 0.01 g . Balance rod with mass difference about 20g was organized in stable place of laboratory where any vibrations were minimized. The loads are 50g and 70g. In other experiment two loads were equal to 500g and balanced with small (about 20g) difference Fig.4. The loads made of plastic.

Fig. 4 Balance scales

Rotation of HTCS disk was organized mechanically by 3.000 rpm electromotor, Fig.5

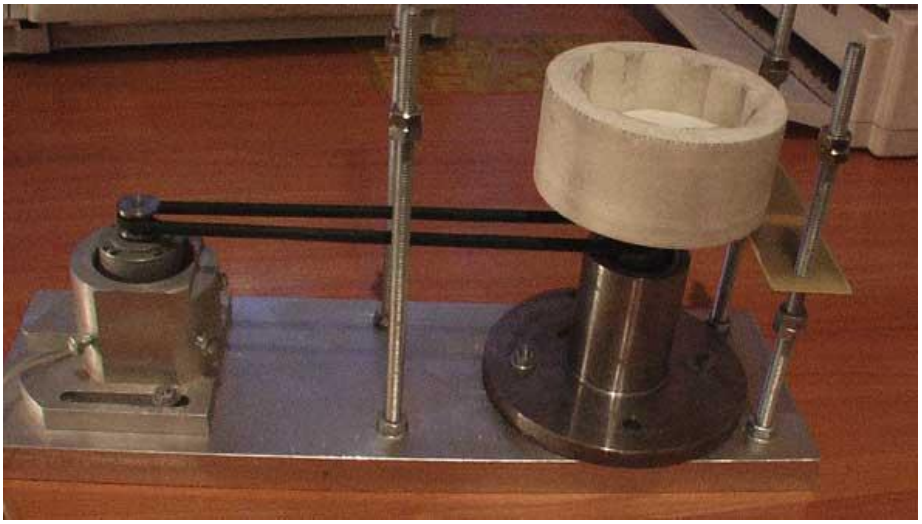


Fig.5 Electromotor and rotating plastic tank for HTCS disk

In this experimental setup the HTCS disk placed in the rotor and cooled by liquid nitrogen can be used in superconductive state only during 20-30 seconds. Due to this problem many measurements on rotational tests can not be reported here as reliable data.

6. **Logbook**

June 23, 2007. Reproduction of Schnurer experiment with balance scales. There are not visible effects for the phase transition from superconductive state to non-superconductive.

It was planned to build more precise rotational detector for more precise measurements. Production of low frequency and high frequency generators and experimental setup to rotate HTCS disk was started.

Other experiment was organized June 23: high voltage discharge to HTCS disk, which was immersed in liquid nitrogen, Fig.6.

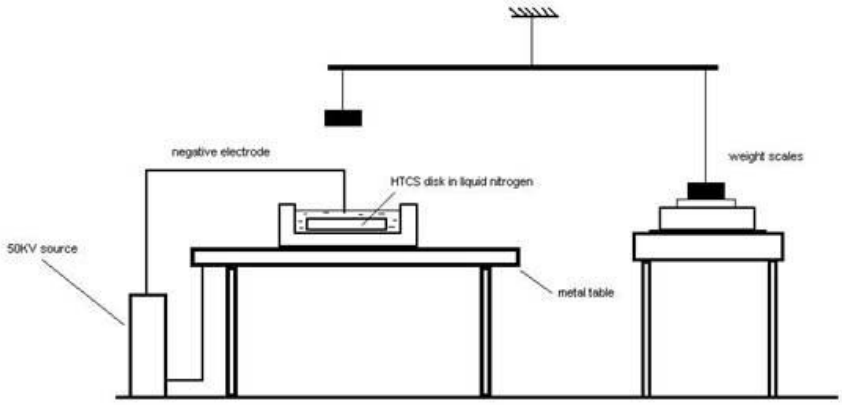


Fig.6

Initially significant weight changes (up to 0.3 g) were detected for the case of negative electrode connection to HTCS disk, which was immersed in liquid nitrogen. But future testing without HTCS disk also produced effects, which were identified as electrostatic interference to digital weight scales.

June 30, 2007. Test with rotational detector, Fig.7

The detector is made of wooden rods and plastic loads. Small glass plate in central point of the horizontal rod reflect red laser beam to the wall of the laboratory placed in 2 m distance that allow to detect small angle oscillations of the horizontal rod. Vertical axis is made of tungsten wire 0.05 mm diameter. All parts of the detector are placed under glass bell to avoid air flow interferences.

Experiment: cooled in plastic tank HTCS disk was placed near the detector. After 30-40 seconds when the disk is changing to non-superconductive state **the attraction of mass to the disk was detected.** After 3-5 min the detector is turning back to previous stable position. Maximum of the effect was measured if the HTCS disk was oriented by its flat side to the detector. Experiment was reproduced 4 times.

It seems to be impossible provide any quantitative data on this effect and future testing is necessary.

Possible mistakes here related with heat and cold flows, i.e. thermogravitation theory. To confirm or disprove this idea new tests with cold non-superconductive mass was organized. Metal disk of mass, which is equal to mass of tested HTCS disk, was cooled by liquid nitrogen and placed near of the rotational detector. Small effect of attraction of the load to the cold mass also was detected in this case. Values of effects for HTCS disk and simple metal disk are different. Conclusion: future testing is necessary to confirm if phase transition in superconductor generate gravity wave and produce attraction/repulsion of the detector.

Fig.7 Rotational detector

July 2, 2007. Experiment with permanent magnet installed near of rotating HTCS disk.

This experiment was planned to test if Lorenz force can be reason of gradient in Bose condensate that change its density and generate gravity wave. Fig. 8 is the case of radial magnetic field and Fig.9 show axial superposition of the permanent magnet (Faraday disk).

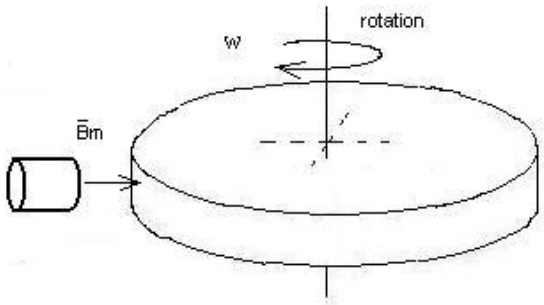


Fig.8 Superconductive disk and radial magnetic field

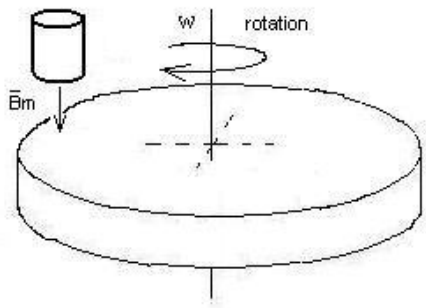


Fig.9 Superconductive disk and axial magnetic field

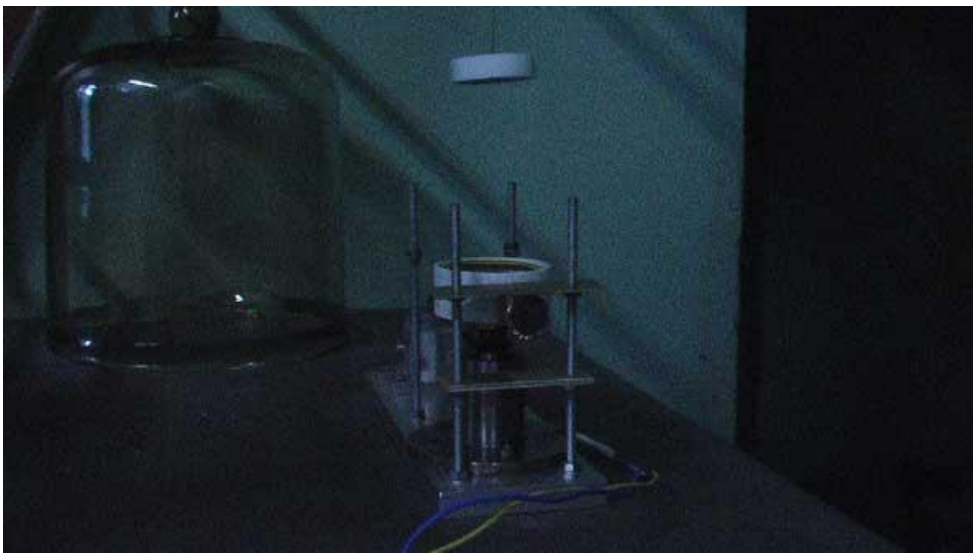


Fig.10 Radial permanent magnet installed near rotating HTSC disk.

Mass of the loads for this case was 50g (above the HTCS disk) and 70g (on the weight scales). Rotation velocity about 2000 rpm. Magnet of 1T field is made of NdFeB material, cylinder of 25 mm diameter and 24 mm height. Distance from magnet to HTCS disk is about 7 mm.

Weight changes were detected as 0.02 g only in experiments with axial superposition of the magnet (Fig.9). It is equal to 0.04% mass change that is too small to be considered as reliable data.

June 04, 2007

Tests with low frequency magnetic field was organized both for the case of stationary HTCS disk and for the case of rotation of the disk. Sinusoidal input signal with frequency from 10Hz up to 1KHz was connected to transistor current amplifier loaded on output coil. For frequencies between 10Hz – 100Hz the coil was made as 500 turns of 1 mm wire on U-shape transformator metal core, Fig. 11.

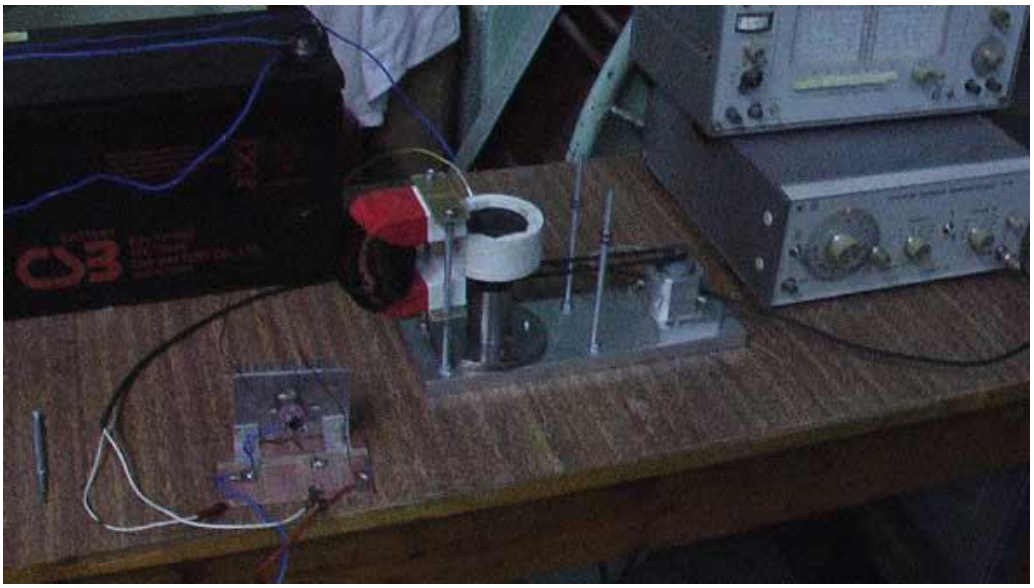


Fig.11 Low frequency tests

Frequency from 100Hz up to 10KHz was tested with other output coil and ferrite core, Fig.12 and Fig.13.

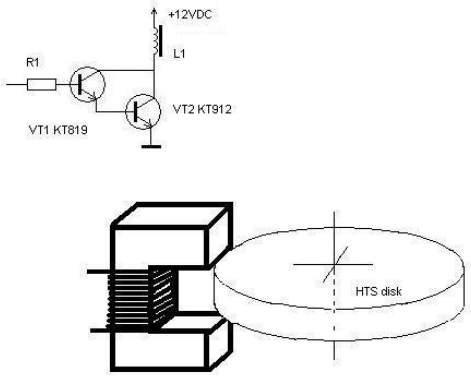


Fig.12

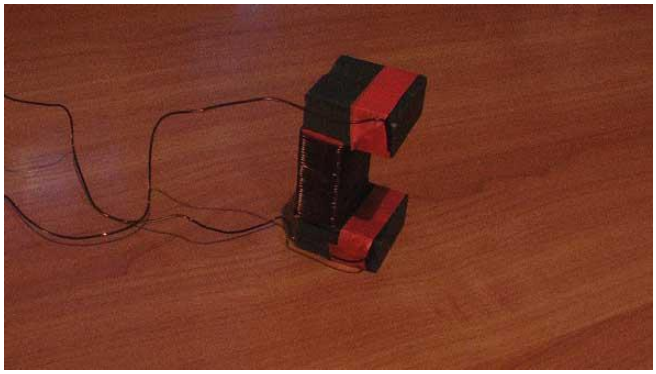


Fig.13

Small positive result was detected for the case of rotation in the field of 1kHz frequency. Weight changes was detected as 0.02 g for mass of the load 500g. Probable it was measurements mistake since percent ratio of the mass changes here is 0.004% only.

July 09, 2007

Tests for frequencies from 10KHz up to 3MHz were organized with air core output coil placed above HTCS disk, Fig. 14 and Fig.15.

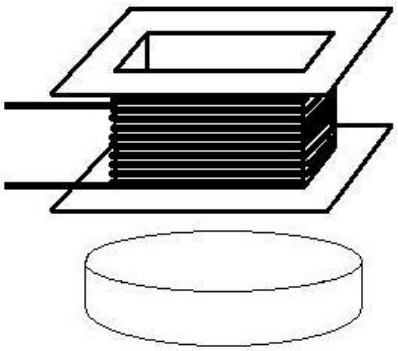


Fig.14

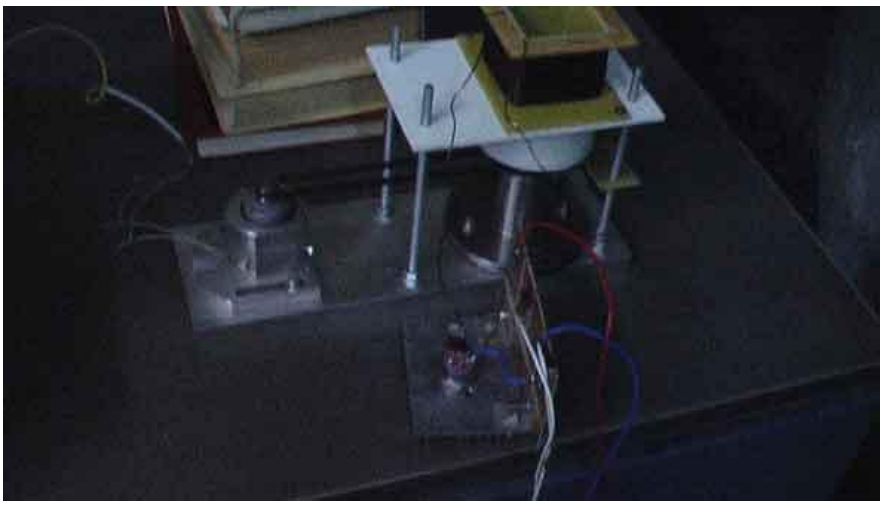


Fig.15

All tests in this case were negative, i.e. it was not confirmed that electromagnetic field in this case produced significant weight changes. Both stationary and rotational HTCS disk were tested.

July 12, 2007

High frequency generator was designed, Fig. 16 – Fig.17 for tests of 3MHz – 40MHz frequency band. Output power is about 10-30Wt.

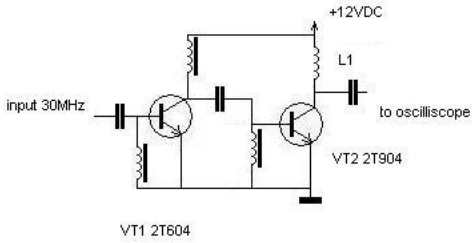


Fig.16

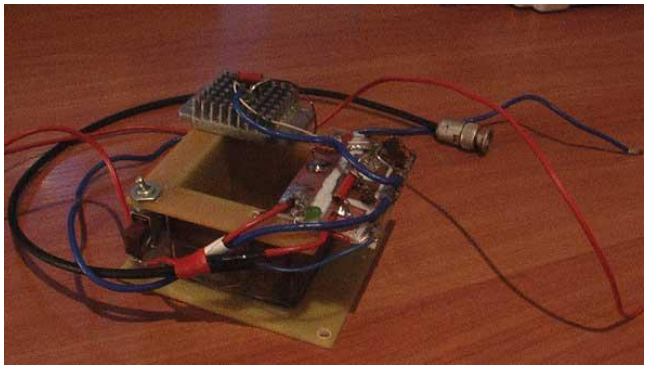


Fig.17

Weight changes were detected as 0.06 g for the case of stationary disk and frequency about 30MHz. High frequency generator was installed above HTCS disk, which was immersed in liquid nitrogen. It seems to be strange that the weight change were stable after the electromagnetic generator was OFF. Quantitative data: this weight change is about 0.01% only.

The case of rotation of the disk in high frequency electromagnetic field also was testes but without estimated effects, Fig.18. Perhaps that in this case important data was missed due to short time of superconductive state of the disk placed in the rotor. Other possible reason is that high frequency electromagnetic field was dissipated in metal parts of the rotor.

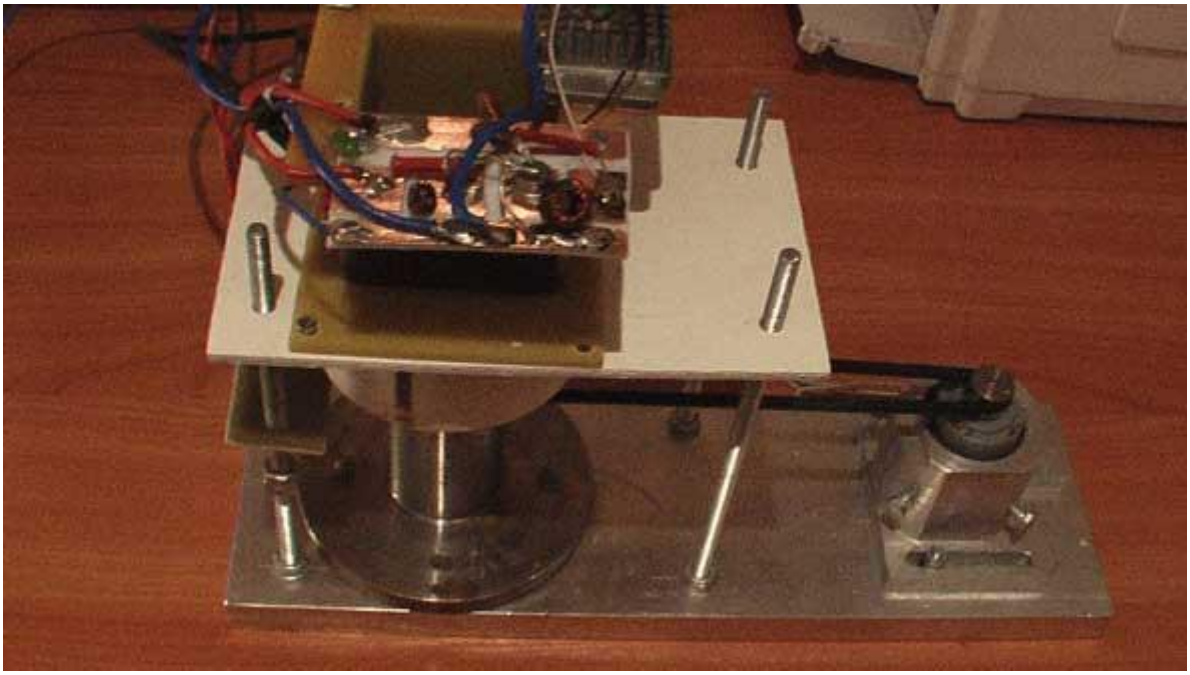


Fig.18

7. Conclusions

7.1. Experiments were organized with low power electromagnetic field. Due to short time of superconductivity state of rotating HTCS disk, reported effects for rotation tests can not be considered as reliable data and additional experimenting can be necessary.

7.2. There is positive effects in the case of rotation of the HTCS disk in permanent magnet field oriented cross the disk axially. If this effect is not a mistake then it can be explained by consideration of conditions created by this design for local gradient of Bose condensate density in the disk due to Lorenz force. Oscillations of this density due to rotation for the disk can generate gravity wave in axial (vertical) directions above and below the permanent magnet.

7.3. Main task of the project was to find resonance effects in 10-100MHz frequency range. Some effects were detected for 1KHz and 30MHz frequencies. To get more reliable data it is necessary to increase power of electromagnetic field.

8. **Planning of the future experiment**

Rotation or motion of HTCS matter in future tests is not planned. It is planned to test superconductor films instead of solid state disk. Instead of high frequency electromagnetic field producing induction currents in HTCS it is planned to use high frequency electric fields, which allow to create high frequency oscillations of Bose condensate and by this way to change its density to find reliable experimental data on resonance frequencies predicted in [1].

References:

1. Large- Scale Sakharov Condition, David Noever and Christopher Bremner, 35th AIAA/ASME/SAE/ASEE Joint Propulsion Conference, Los Angeles , CA , 20-23 June1999.
2. Podkletnov E. A Possibility of Gravitational Force Shielding by Bulk YBa2Cu2O7-x Superconductor, Physica C 203 1992, pp 441-444.
3. Los Alamos http://xxx.lanl.gov/abs/cond-mat/9701074
4. G. Modanese, Possible theoretical interpretations of the weak gravitational shielding effect by composite YBCO HTC superconductor, 1997, IAF.
5. Impulse Gravity Generator Based on Charged YBa_2Cu_3O_{7-y} Superconductor with Composite Crystal Structure, Evgeny Podkletnov, Giovanni Modanese, http://xxx.lanl.gov/abs/physics/0108005
6. Antigravity? http://www.businessweek.com/1997/07/b3514118.htm
7. Electric Spacecraft, Issue 22, 1997 pp.25-26
8. 8. Anti-gravity Effect? Gravitational Equivalent Of A Magnetic Field Measured In Lab http://www.sciencedaily.com/releases/2006/03/060325232140.htm
9. US Patent # 3,839,771, Method for Constructing a Thermionic Couple, October 8, 1974, Giani A. Dotto
10. http://en.wikipedia.org/wiki/Kinetic_theory_of_gravity
11. http://www.can.cz/

Contact

Mr. Alexander V. Frolov
Director, Faraday Lab Ltd, Russia
http://www.faraday.ru/
email:a2509@yahoo.com phone 7-920-794-4448

GravWave® LLC TEAM

(A group of individuals in the United States and abroad who have a common interest in
High-Frequency Gravitational Waves)

Robert M L Baker, Jr., PhD

Principal Investigator and Project Manager

Robert M L Baker, Jr. was born in Los Angeles on September 1, 1930. He has been married to his wife Bonnie since 1964 and has three grown children. Baker earned a bachelor's degree in Physics at UCLA with highest honors (*summa cum laude* – first in his class) was elected to *Phi Beta Kappa*, earned a master's degree in Physics and a Ph.D. in Engineering at UCLA—the Ph.D. degree with a specialization in space navigation was, according to UCLA officials, the first of its kind to be granted in the United States. Dr. Baker was on the faculty of the Department of Astronomy at UCLA from 1959 to 1963 and the Department of Engineering and Applied Science at UCLA from 1963 to 1971 as a Lecturer and Assistant Professor. During that time he was a Lecturer at the *United States Air Force Academy.* While on a two-year tour of active duty in the Air Force he worked on a variety of classified aerospace projects. He was the head of the Lockheed's *Astrodynamics Research Center* in Bel Air, California and in 1964 joined *Computer Sciences Corporation* as the Associate Manager for Mathematical Analysis. In 1980 he was elected President of *West Coast University*, an accredited university for the adult learner (Western Association of Schools and Colleges

or WASC and Accreditation Board for Engineering and Technology or ABET) now operating under the auspices of *American Career College* in Los Angeles. After retiring from West Coast University in 1997 as President and Professor of Engineering, Dr. Baker became the Senior Consultant for *Transportation Sciences Corporation* and *GRAVWAVE© LLC*. He won the UCLA *Physics Prize*, was recipient of the *Dirk Brouwer Award* for outstanding contributions in astrodynamics and orbital mechanics, was a recipient of the *Outstanding Man of the Year* Junior Chamber of Commerce award in 1965 presented to him by Ronald Reagan and the 2011 "Outstanding Lifetime Achievement in the Field of High-Frequency Gravitational Waves" Awarded during the joint Conference of the International Institute of Information and Systems (IIIS), International Conference on Information and Communication Technologies and Applications: ICTA 2011, the International Institute for Design Marketing in Science, Education and Technology (DeMset) and the 17th International Conference on Information Systems Analysis and Synthesis: ISAS 2011.. He is a *fellow* of the *American Association for the Advancement of Science.* He was national chairman of the Astrodynamics Technical Committee of the *American Institute of Aeronautics and Astronautics* (AIAA) from 1961 to 1964, was Editor of the *Journal of the Astronautical Sciences* from 1963 to 1975, was appointed by William Bennett to the *National Advisory Committee on Accreditation and Institutional Eligibility* of the Department of Education from 1987 to 1989, was appointed to the *Academic Review Committee on Gravitational Research* with the U. S. Army from 2001 to 2003, Vice Chairperson of the first International HFGW Workshop at the MITRE Corporation in 2003, Honorary Chairman of the second International HFGW Workshop in Austin Texas in 2007, Chairman of the third International HFGW Workshop in Huntsville, Alabama in 2009 , cochairman of the High-Frequency Gravitational Wave Symposium at the Johns Hopkins University Applied Physics Laboratory in 2010. and coauthor of papers presented at the Space Technology Applications International Forum in 2012. Dr. Baker is an Advisory Professor *Chongqing University*, China 2004, and was the author of several textbooks and over one hundred company reports, symposium papers, and journal articles in the area of astrodynamics, unmanned sailcraft, celestial mechanics, and High-Frequency Gravitational Waves (HFGWs) including *An Introduction to Astrodynamics* (1960) with Maud W. Makemson and *Astrodynamics: Applications and Advanced Topics* (1969). Dr. Baker has been Project Manager on three prototype development, fabrication, and test projects under contract to the U. S. Navy and Principal Investigator on several NASA and USAF projects while head of Lockheed's *Astrodynamic Research Center*. As President of *West Coast University* Dr. Baker coordinated the activities of six groups of scientists and engineers spread throughout Southern California. Dr. Baker has been interested in the dynamics of gravitational fields since the 1950's and in high-frequency gravitational-wave (HFGW) research since the early 1960's. He is the inventor of the Li-Baker HFGW detector and the double-helix HFGW generator/transmitter and holds six patents and several pending patents in the United States, Europe, Russia, and China in the area of gravitational-wave generation and detection in the laboratory. You are invited to visit: www.DrRobertBaker.com.

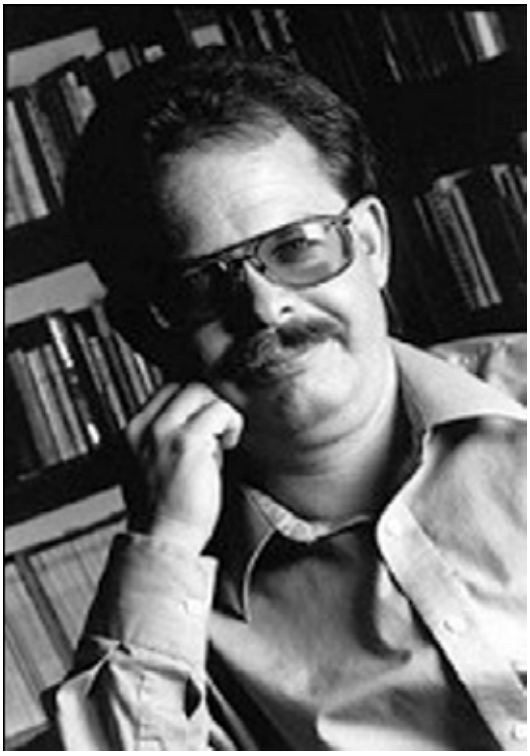
Roger Clive Woods, M.A., D.Phil., D.Sc. (Oxford), FIET

Senior Scientific Investigator

R. Clive Woods was born in Leicester, England on May 18, 1955. Currently he is Department Chairman and Voorhies Distinguished Professor, Department of Electrical and Computer Engineering, *Louisiana State University*. Woods earned a Master's degree and Doctorate at New College, *University of Oxford*, 1976 and 1980, for work on magnetic resonance in rare earth metal alloys in the Solid State Physics group of the Clarendon Laboratory in the *University of Oxford*. Dr. Woods was a Senior Scientist and Project Manager at Plessey Research (Caswell) Ltd. from 1982 to 1983, and Lecturer and Senior Lecturer on the faculty of the Electronic and Electrical Engineering Department of the *University of Sheffield* from 1983 to 2001. During this period he managed projects on the design, development, and test of microwave acoustic devices, III-V heterojunction bipolar transistors, avalanche photodiodes, high-temperature superconductors, and other solid-state devices. In 1989 he was appointed as a *British Association Media Fellow* and in 1995 he was Professeur Invité at the Laboratoire de Physique de la Matière, *Institut National des Science Appliquées* de Lyon, France. From January 2002 to June 2006 Dr. Woods was a Full Professor of Electrical and Computer Engineering at *Iowa State University*, Ames, Iowa. During 1992-1995 he was a Member of IEE Professional Group Committee S8 (Electromagnetics); 1999-2002 a Member of IEE Professional Group Committee E3 (Microelectronics and Superconductor Devices); 1999-2002 an Associate Editor of *IEE Electronics and Communication Journal*, and in 2003 a Member of the *National Science Foundation* SBIR/STTR Photonics (Lasers and

LEDs) panel. Dr. Woods has consulted for Barnsley Business and Innovation Centre Ltd., McLarens Ltd., Price Waterhouse, John Lovell Associates, Halpern & Ward Associates, the European Commission in Brussels, and Ashton Brown Associates Ltd. among others. He has authored over 70 technical papers and the book "Digital logic design" (with B. Holdsworth), Butterworth-Heinemann, 2002. Dr. Woods has been interested in the research associated with gravitational-waves for over a decade and participated in the first International HFGW Workshop at the MITRE Corporation in2003, the second International HFGW Workshop at Austin, Texas in 2007, the third at Huntsville, Alabama in 2009 and the High-Frequency Gravitational Wave Symposium at the Johns Hopkins University Applied Physics Laboratory in 2010. You are invited to visit: http://myprofile.cos.com/cwoods

Eric W. Davis, PhD



Senior Scientific Investigator

Eric Davis received his bachelor's degree in Physics in 1983 and a Ph.D. in Astrophysics with a specialization in the theory of relativity from the *University of Arizona* in 1991. His fields of specialization include relativity theory and cosmology, space mission engineering, gravitational-wave research, and the NASA Breakthrough Propulsion Physics program. Dr. Davis is a Senior Scientist at *The Institute for Advanced Studies at Austin.* He also participates in and serves as a consultant to the NASA Breakthrough Propulsion Physics program. For nearly six years Dr. Davis served as a research scientist and S&TI analyst at the National Inst. for Discovery Science-Bigelow Aerospace Co. where he conducted phenomenology research and forensic

investigations into transient optical phenomena and manned spaceflight research projects. Dr. Davis began his graduate work as a mission support and research assistant with the Infrared Astronomical Satellite group at the *Steward Observatory* in Tucson, AZ. In 1985, he joined the Voyager Ultraviolet Spectrometer Experiment group at the Lunar & Planetary Lab in Tucson where he conducted research on Jupiter's magnetosphere, the Uranus and Neptune planetary encounters, and participated in the Voyager 1 & 2 space mission support. Following completion of his doctorate, he became an associate faculty and interim director for the Arizona Astronomy Education Center at *Pima College* in Tucson. In 1995 Dr. Davis was appointed to the faculty of physical/computer sciences & mathematics, where he was involved with Air Force sponsored space mission engineering and Korean theater space reconnaissance training, at the *University of Maryland* Asian Division while he was stationed at the 8th Fighter Wing in Kunsan, South Korea. Dr. Davis developed experiment designs and cost estimates supporting the AFRL Advanced Concepts Office He has also been recognized by the *American Institute of Aeronautics and Astronautics* in their 1993-94 Public Policy Award for outstanding contributions to national defense and space public policy, and received recognition from the State of Arizona Economic Conversion Council for contributions to commercial conversion of small-medium space and defense businesses in Arizona. Dr. Davis is a Fellow of the *British Interplanetary Society*, senior member of the *American Institute of Aeronautics & Astronautics*, member of the *Association of Former Intelligence Officers*, and member of the *American Astronomical Society.* Dr. Davis participated in the first International HFGW Workshop at the MITRE Corporation in 2003 and the second International HFGW Workshop at Austin, Texas in 2007.

Andrew Beckwith, PhD

Scientific Investigator

Andrew (Andy) Beckwith was born on December 15, 1954. He has studied astrophysics, theoretical Physics and superconductivity. He received his PhD from the Texas Center of Superconductivity on June 5th, 2002. Dr. Beckwith is currently working in condensed matter CDW/ Domain wall physics problems, and early universe models of emergent gravity/ entropy generation. He joined *Fermi National Accelerator*

Laboratory (FNAL) in August of 2002, is a member of the American Institute of Beam Energy and was appointed senior scientific researcher for GravWave® LLC in February, 2009. Dr. Beckwith is a Fellow of the Space Physics Committee on Space Research (COSPAR) and on their Committee for Space Astrophysics as well as a Life Member of the American Institute of Beamed Energy Propulsion (AIBEP). He has attended several astrophysics conferences in the last few years (2008) including the following: *Space Technology Application International Forum (STAIF)* 2007 and 2008 (HFGW sessions), Quantum Atomic and Molecular Tunneling in Solids (QAMTS) 2007 (Texas), 23rd International Conference On Neutrino Physics And Astrophysics (Neutrino 2008) (New Zealand), and European Committee for Future Accelerators (ECFA) 2008 on behalf of *Fermi National Accelerator Laboratory (Poland).* He is also currently doing work in several areas including topics concerning the early universe and relic HFGWs. Dr. Beckwith participated in the third HFGW International Workshop at Huntsville, Alabama in 2009 and the High-Frequency Gravitational Wave Symposium at the Johns Hopkins University Applied Physics Laboratory in 2010.

.

Gary V. Stephenson

Scientific Investigator

Gary Stephenson received his Bachelor's of Science Degree in Physics at Montana State University in 1983, and his Masters of Engineering in Space Systems Engineering at Stevens Institute of Technology in 2011. From 1983 to 1986 he was a Member of the Technical Staff at Hughes Aircraft Company where as a systems engineer in optical and radar systems. In 1986 Mr. Stephenson joined the Aerospace Optical Division of ITT where he performed systems design of space borne meteorological Infrared imagers. In 1989 through 1997 he returned to Hughes as a Systems Engineer where he was responsible for the electro-optical systems engineering and on-site support of AST, an airborne infrared tracking sensor for the US Army. From 1997 to 2010 he was a

Systems Engineer at The Boeing Company where he performed the systems design of electromagnetic and electro-optical mission equipment, including TAPLOC, (TENCAP AWACS Precision Location), the satcom system for Connexion by Boeing, and several proprietary satellite projects. Mr. Stephenson is named as inventor or co-inventor in six US patents in satellite communications. Since 2010 he has been with LinQuest, and as a government crew chief manages the integration of a number of GPS II-F spacecraft for the US Air Force. Since 1989 Mr. Stephenson has also been the president and chief investigator for Seculine Consulting, an applied physics consultancy. He has several publications in the communications applications of high-frequency gravitational waves, including the possible application of the Gertsenshtein effect. He has participated in the first International HFGW Workshop at the MITRE Corporation in 2003, the second International HFGW Workshop at Austin, Texas in 2007, as cochairman of the third at Huntsville, Alabama in 2009, and as cochairman of the High Frequency Gravitational Wave Symposium at the Johns Hopkins University Applied Physics Laboratory in 2010.

Lawrence S. Moy, M.D.

Medical/Biological Investigator

Lawrence Moy graduated from Medical School at the University of California at Los Angeles in 1981 and was certified by the American Board of Dermatology. For the last several years Dr. Moy has worked with Dr. Baker on the applications and influences of high-frequency gravitational waves on biological matter especially on the muscles underlying the skin. They have jointly published peer-reviewed scientific papers on the subject.

Colby Harper

Telecommunications Specialist

Colby Harper graduated from the *University of Washington* with a BA with College Honors in Economics in 1993 and from 1996 to 1997 performed graduate studies *University of Washington Graduate School of Business Administration.* In 1993 he was the recipient of the Certificate in International Economics from the *University of Washington*. Between 1985 and 1990 Mr. Harper was in contract software development and network design and administration for companies including *Westinghouse Corporation*. He was involved in systems engineering at *Digital Communications Associates* in 1990-1991, was Senior Architect, Project Manager and Network Engineer at *Energy Northwest* from 1991 to 1995 and in late 1995 he served in the National Telecommunications and Information (NTIA) branch of the United States Department of Commerce and was a contributor to the Telecommunications Act of 1996. From 1996 to 2001 Harper was the National Director, Wireless Internet Service Provider & Technical Infrastructure at *Inteliant Corporation*, and in 2002 he joined *Washington Mutual* as a Product Manager involving, among other tasks, network systems and information handling. Harper is an IBM-Certified Client/Server Integration Specialist. In summary Mr. Harper is a hands-on wireless telecommunications executive having in-depth technology, operations, and process–improvement experience. In cooperation with Paul Murad he prepared a future concepts paper entitled: “Cost-based Global Optimization of Tomorrow’s High-Frequency Gravitational Wave (HFGW) Communications System: Value-added HFGW Device Requirements for Today,” and coauthored with Gary Stevenson a paper presented at the Space Technology and Applications International Forum (STAIF2007), sponsored by the American Institute of Physics, entitled “The

Value Estimation of an HFGW Frequency Time Standard for Telecommunications Network Optimization."

Buzz Aldrin, Sci. Dr.

Senior Scientific Advisor

Edwin "Buzz" Aldrin was educated at *West Point*, graduating with honors in 1951, and was third in his class. He earned a Doctorate in Astronautics from the *Massachusetts Institute of Technology* in Manned Space Rendezvous. On July 20, 1969, Buzz and Neil Armstrong made their historic Apollo XI moon walk, thus becoming the first two humans to set foot on another world. Since retiring from *NASA*, the *Air Force*, and his position as Commander of the Test Pilot School at Edwards Air Force Base, Dr. Aldrin has remained at the forefront of efforts to ensure a continued leading role for America in manned space exploration. He has been interested in High-Frequency Gravitational Waves since 1999 when he was briefed on the subject by Dr. Baker. Dr. Aldrin heads *Starcraft Boosters, Inc.,* Starcraft *Enterprises* and the *ShareSpace Foundation,* and is Chairman of the *National Space Society.* In 2011 he was awarded the Congressional Gold Medal, the United States highest civilian award. You are invited to visit: www.buzzaldrin.com.

Jerah D. Cameron

Senior Systems Engineer

Jerah Cameron attended the *University of Northern Colorado* and *CCU* and was awarded a Bachelors of Science degree in Computer Information Systems-Management in 2002. He has Certification as a *Microsoft Systems Engineer* (MCSE NT/2000/2003/2010), *Microsoft Professional Internet* (MCP+NT/2000/2001) and *Microsoft Office SharePoint Server* [SharePoint Server 2007 and 2010 (MCTS)]. He is designated as a Senior Systems Engineer, Solutions Architect & Web Design/Programmer. At *Raytheon*, from 2003 to 2005, Mr. Cameron was a Systems Administrator/ Integration Specialist (SSB!/Top Secret Clearance). At *IBM*, from 2005 to 2006, he was a Senior Systems Administrator and SharePoint Architect, At *Newsgator Technologies,* from 2006 to 2011; he was Solutions Architect/Senior Systems Engineer. Most recently, 2011, Mr. Cameron is Solutions Architect/Level V Senior Technical Consultant for Hewlett Packard. He provides the technical direction required to resolve complex issues to ensure the on-time delivery of solutions that meet customer expectations. He regularly leads large functional teams to include technical management of client staff assigned to the implementation team in the completion of one or more solution requirements, architecture or implementation deliverables. Mr. Cameron works with peers outside his immediate organization to define and characterize complex technology or process problems and/or develops new solutions, yet he works independently from others to drive technical problems to a solution.

Recently (2012), he has become very interested in theories and research in the area of high-frequency gravitational waves and its applications to Information Technologies (IT).

Peter Guernsey

Engineering Specialist

Peter Guernsey was born in Kansas City, Kansas, youngest of five children. His father was an executive in the aerospace reprographics industry. In 1967, the family moved to Southern California where Peter spent his grade school years. In high school he focused on shop classes, photography, auto mechanics, wood shop, metal shop, photo offset lithography, photography, astronomy, and horticulture. During his high school tenure Peter was introduced to optical coating. Learning basic principals of vacuum pumps and ultra high vacuum pumps, and the deposition of simple materials such as aluminum, gold, SIO, for the application of astronomical mirrors and optics for precision instruments such as mirrors for galvanometers for seismic measurements. In 1982, he founded Guernsey Coating Laboratories, Inc. (GCL) in Ventura, California, as CEO and President. For the past 27 years GCL has specialized in precision optical coatings for lasers, medical instruments and military applications. GCL is currently involved in the manufacture and deposition of ION beam assisted, ION beam sputtered; electron beam deposited, and thermally deposited films for the optical and industrial community. Mr. Guernsey is most interested in gravitational-wave detection and, in particular, in the highly absorbent material coating to minimize microwave noise in the Li-Baker detector.

His interest In Einstein's general theory of relativity has intrigued him since early child hood along with the grail of physic's, which is to combine particle and wave theory.

Robin Fell

International Business Specialist

Robin Fell received her Bachelor of Arts degree in Political Science at the *University of California at Berkley* in 1990, Certificate at the *American University in Paris* in 1989, and a MBA in International Business at *Loyola Marymount University*, Los Angeles in 1997. She was the Night Reporter at television station *KMIR* in Palm Springs, California in 1990, directed domestic video distribution at *Vidmark Pictures* 1992-1994, and Director of International Sales at *Rysher Entertainmen*t 1994-1998. Robin is an inventor (of "Drip Free"), an entrepreneur and an international marketing and manufacturing specialist. She has made numerous sales trips to Europe and Southeast Asia and is the Media Director for GravWave® LLC..

Thomas R. Lane

Scientific Advisor

Thomas R. Lane received his BS *cum laude* in Physics at the *United States Air Force Academy* in 1988. He completed Squadron Officer School in 1994 and graduated from Air Command and Staff College in 2002. From 2000 to 2004 he was Assistant Director of Operations and from 2004 to his retirement was Chief of Plans at the 28th Bomb Squadron in the *United States Air Force*. Lane is a B-1/T-37 Instructor Pilot and holds a Top Secret security clearance. In 2004 he assisted Dr. Baker in the preparation of the "Layperson's Description of High-Frequency Gravitational Waves (HFGWs)" and has been an active, although unofficial, contributor to HFGW research since that time. He and his wife Cathy live in Abilene, Texas. Currently Mr. Lane is a member of J9 Future Capabilities Division, Force Management Directorate, U.S. Special Operations Command.

Frederick W. Noble

Prototype Analyst

The co-inventor with Dr. Baker of the parent patented HFGW invention, Frederick W. Noble, was born in Oak Park, Illinois, May 1, 1939. He earned a bachelor's degree at *UCLA* with a major in English and a minor in Naval Science. In the *United States Navy* from 1961 to 1963 he served as a junior officer on a destroyer based in Japan. In 1966 he earned a Doctor of Law degree from the *UCLA School of Law* and is a Member of the California and Federal Bar. Mr. Noble was a practicing attorney from 1966 to 1978. From 1979 to the present he is President and CEO of Frederick W. Noble, Inc., a personal investment company. From 1982 to the present he is President and CEO of *Wintec Energy*, Inc. and its affiliates, which operate wind farms in Palm Springs, California and in the Altamont Pass area in Northern California. In July of 2001 he co-founded GRAVWAVE® LLC with Dr. Baker. Mr. Noble's *Wintec* corporation owns a one-thousand-acre property in Palm Springs, California, with some 600 wind turbines. The property has available several 10-acre to 40-acre flat, vacant sites for the construction of the prototype embodiments of the HFGW generator/detector in the US. Mr. Noble and *Wintec* can make available several civil and electrical engineers and construction personnel for HFGW prototype fabrication and test.

Bonnie Sue Baker

Scientific Investigator

Bonnie Sue Baker was born in Dunsmuir, California on August 17, 1942. She attended *Glendale City College* 1960-61, studied Physics at *San Jose State College* 1962, Certified Microsoft Office Specialists 2012. She was Technical Assistant *Lockheed Astrodynamics Research Center* 1962-1963 (where she first became interested in high-frequency gravitational waves or HFGWs), Technical Associate, *Planning Research Corporation* 1963-1965, Scientific Associate, *Astrophysics Research Corporation* 1965-1968, Scientific Investigator, *Online Data Processing* 1968-1980, Director of Special Scientific Projects, *West Coast University* 1981-1986 and Scientific Associate *Transportation Sciences Corporation* 1987 to date. Attendee of the first International HFGW Workshop at the *MITRE Corporation* in 2003, the second International HFGW Workshop in Austin Texas in 2007, the third International HFGW Workshop in Huntsville, Alabama in 2009 , the High-Frequency Gravitational Wave (HFGW) Symposium at the *Johns Hopkins University Applied Physics Laboratory* in 2010 and coauthor of a paper on HFGW Global Communications at the the 17th International Conference on Information Systems Analysis and Synthesis: ISAS 2011.in Orlando, Florida, 2011 .and coauthor of papers presented at the Space Technology Applications International Forum in 2012.. She is the co-inventor of two high-frequency gravitational wave United States Patents and coauthor of several high-frequency gravitational wave technical papers.

Christine Black

Scientific Associate

Christine Black graduated from *The University of Michigan* in 2009 majoring in Astronomy and Astrophysics with an interdisciplinary major in Physics. She has taken courses in EM Theory, Quantum Mechanics, Celestial Mechanics (Stars and Galaxies), Optics, Special Relativity and Astronomical Techniques. She has carried out research with at the University of Michigan involving variable stars in the nearby galaxy Carina and accomplished graduate research at the *University of Tasmania*. Ms Black was involved in research projects at the UCLA Astronomy Department in 2011. She is currently pursuing post-graduate studies in the Department of Physics and Astronomy at *Dartmouth College*. She has a keen interest in HFGW research and has been an assistant of Dr. Robert M L Baker, Jr. as a Project Scientist at *GravWave® LLC* since 2007 and presented a paper with him concerning high-frequency gravitational waves at the third HFGW Workshop in February, 2009 at Huntsville, Alabama..

Andrew Moe

Student Apprentice

Andrew Moe is an undergraduate student at *North Dakota State University*, with a major in Mechanical Engineering. Andrew takes an interest in some of the more obscure fields of science and is especially interested in the creation and application of HGFWs. He believes that they are going to become a key technology in the future. Andrew plans on testing the theories of the GravWave Team with the intention of engineering practical, and useful devices.

Jeff Schuss

Student Apprentice

Jeff Schuss is an undergraduate (2007) at *Loyola Marymount University*, with a hybrid major of Engineering Physics (EE/Physics). Besides having “gone the extra mile” in investigating many principles and theories that would have been otherwise briefly mentioned in a class, Mr. Schuss plans on taking classes in **pending**. He is fascinated by the work being done with HFGW and looks forward to broadening his horizons. He is also the grandson of Dr. Buzz Aldrin.

Maximilian Collins

Student Apprentice

Maximilian Collins (Max) is a 14 year old student in *Midlothian, Virginia*. He has loved science since the age of four months when he would stare fixatedly at the gears of the baby swing. He has been home schooled since the 6th grade and is currently enrolled in a nearby junior college. He will finish high school by the end of next year and plans to attend a four-year university in the fall of 2009. He first came across the subject of HFGWs in a scientific forum and began studying them. As a profession, he hopes to be a particle physicist, a research chemist, or a HFGW researcher.

IN MEMORIUM

One of the most important and intelligent members of the GRAVWAVE® Team has passed away. Bob Anderson was a Captain of Industry, perceived to be assertive, even gruff on the outside, but gentle, kind and generous on the inside – we will miss his business acumen and advice.

Robert Anderson
Senior Management Advisor

Robert Anderson was the Chairman of the Board, President, and CEO of *Rockwell*. He received his bachelor's degree in mechanical engineering from *Colorado State* University in 1943. He was a knowledgeable and aggressive business man who was an expert in defense and commercial applications of new technology. He has been referred to as "... the father of the B-1 Bomber." In fact, in 1977 he accepted the prestigious *Robert J. Collier* Trophy for Rockwell International's work on the B-1 bomber. Mr. Anderson was a Lifetime Trustee of the *California Institute of Technology* and was a Consulting member of the Caltech Board of Trustees Committee on the *Jet Propulsion Laboratory.*

INTERNATIONAL CORRESPONDING SCIENTIFIC ADVISORS

Giorgio Fontana, PhD

Senior Scientific Advisor, Italy

Giorgio Fontana was born in Trento, Italy on June 11, 1957. He received his doctorate with a specialization in electronics at the *University of Padua* in Italy. In 1984 Dr. Fontana became the Head of the Electronics Laboratory of the Department of Physics of the *University of Trento*. In 1995 he became the head of the computer center of the Faculty of Science (physics and mathematics) of the *University of Trento* and in 2000 he was advanced to the Head of the Electronics Laboratory of the Department of Information and Communication Technology at the *University of Trento*. Dr. Fontana has developed scientific instrumentation in the field of laser measurement and characterization of materials, ion transport and analysis, superconductors and related devices, cryogenic semiconductor electronics, optics, gravitational-wave Weber bar detectors and computer systems. Recently he has been involved in the development of technologies for atomic time transfer and synchronization, high speed data switching and fiber optic amplifiers and sensors. Dr. Fontana is interested in the development of new technologies for space travel with gravitational waves and is involved in the development of a suitable theoretical/engineering model for FTL travel through the Hyperspace. Currently he instructs in computer simulation and electronic circuit courses at the *University of Trento.* Dr. Fontana is a member of the *Istituto Nazionale di Fisica Nucleare (INFN).* He has authored well over a dozen scientific papers related to High-Frequency Gravitational Waves (HFGW) in internationally recognized technical journals. He participated in the first International HFGW Workshop at the MITRE Corporation in 2003, the second International HFGW Workshop at Austin, Texas in 2007 and the third at Huntsville, Alabama in 2009.

Gloria Garcia-Cuadrado

Scientific Advisor, Spain

Gloria Garcia-Cuadrado was born in Barcelona, Spain on January 14, 1974. She received a Master in Theoretical Physics (Research Proficiency) by the Physics Faculty of the *University of Barcelona*, SPAIN; in 2002 she was awarded a Diploma in Space Studies by the *International Space University; Polytechnical University of Pomona*; California; USA and in 2000 received a Bachelor in Physics by the University of Barcelona; SPAIN. Director of the Catalan Aerospace Cluster BAIE (*Barcelona Aeronautics & Space Association*), gathering up to 90 members devoted to Aerospace Research & Technology in Catalonia and Spain, including Universities and other Research Centres, Companies, Financing Institutions and Government (both at regional level in Catalonia and at national level, in Spain). The Area focuses also in identifying advanced concepts and technologies to pursue. Projects carried out so far include Space Tourism; Life Support Systems for Human Spaceflight and Gravitational Wave research. As Head of Communication main responsibilities include: representation of the Centre; science outreach; institutional and governmental relations; and media relations and networking. She received a grant by the *European Space Agency* (ESA) to design, implement and conduct an experiment in microgravity conditions on board an Airbus 300; and a joint grant from ESA and the Government of Catalonia - Generalitat de Catalunya - to attend the *International Space University* Course, ending in the attainment of the Diploma in Space Studies; California – USA. Ms. Garcia-Cuadrado has expressed a great interest in High-Frequency Gravitational Wave research and she has been previously involved with low frequency Gravitational

Wave projects such as LISA and LISA Pathfinder. She was a participant in the third International HFGW Workshop at Huntsville, Alabama in 2009.

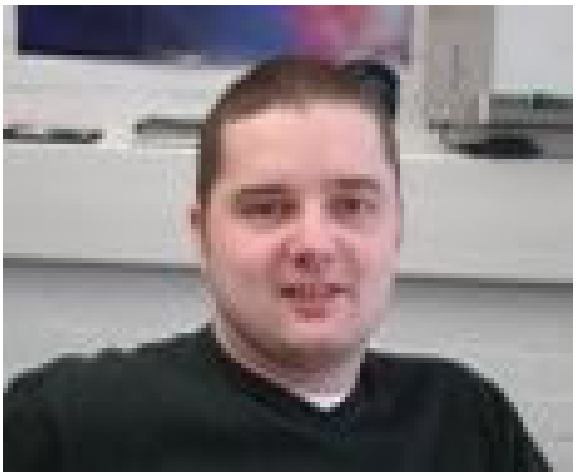
Christian Corda, PhD

Scientific Advisor, Italy

Christian Corda was born in Nuoro Italy on the March, 7, 1969. He obtained a Graduate Degree in Physics at the *Pisa University* in the 2001 and a specialist degree in Physical Sciences at the Pisa University in the 2003. He received his Ph.D in Physics at the Pisa University in 2003-2007, obtaining a research contract (Scientific Association) from the *Istituto Nazionale di Fisica Nucleare* (INFN) within the INFN-Pisa Section. Dr. Corda is a life member of the "Società Italiana della Relatività Generale e della Gravitazione" (SIGRAV). In 2007 Dr. Corda was the invited researcher at the *Albert Einstein Institute for Gravitational Physics*, in Hannover, Germany and invited Visiting Professor at the *Instituto de Cosmologia, Relatividade e Astrofisica* (ICRA-BR), Centro Brasilero de Pesquisas Fisicas, in Rio de Janeiro, Brazil. At the present time, Dr. Corda works within the Virgo Collaboration. He is Chairperson of the "Scientific Association Galileo Galilei", in Prato near Florence, Italy. He is also First Investigator of the Cosmology, Astrophysics and Gravitation Program of the *Centro di Scienze Naturali,* in Prato. Dr. Corda is author of more than 40 scientific publications in international journals in the fields of gravitation, gravitational-waves, astrophysics and cosmology. In 13 of these publications he is the sole author. Dr. Corda is also author of a number of scientific talks at various international conferences and/or workshops. At the end of the 2007, Dr. Corda became the world's most cited author in the *Official Astroparticle Publication Review* of ASPERA with 13 citations. ASPERA, the network of national government agencies responsible for coordinating and funding national research efforts in Astroparticle Physics. The Mayor and the city council of Nuoro will award Dr. Corda a solemn encomium in 2008 and also the nomination of world international ambassador for the city of Nuoro. Because of his recent research on the "magnetic" component of gravitational waves and on the stochastic background of relic gravitational

waves, Dr. Corda is strongly interested in the high-frequency gravitational waves.

Richard Ingley, PhD



Scientific Advisor, England

Richard Ingley received his Ph D in 2005 at *Birmingham University*, Birmingham, England. His thesis was entitled: “Implementation and Cross Correlation of Two High Frequency Gravitational Wave Detectors” and reports on his fabrication of the first High-Frequency Gravitational Wave (HFGW) detector. The detector was based upon the concept of Mike Cruise who found that in the presence of a gravitational wave, an electromagnetic wave will experience changes in its amplitude, frequency, direction of propagation and polarization. If an electromagnetic wave is confined to move in a circular path in the presence of a gravitational wave, then a resonant condition exists where the *polarization shift* is cumulative with successive passes of the path circumference and thus the HFGW can be detected. Dr. Ingley has published several peer-reviewed papers concerning HFGWs, presented papers at the Amaldi Conferences and was an invitee to the first International HFGW Workshop at the MITRE Corporation in 2003 and the second International HFGW Workshop at Austin, Texas in 2007. He is now associated with the Centre for Electronic Imaging School of Engineering and Design, *Brunel University*, Uxbridge Middlesex, England.

Fangyu Li, PhD

Senior Scientific Advisor, China

Fangyu Li was born on October 28, 1943. He was a student in the department of Physics at *Northwestern Normal University*, China from 1961 to 1965. From 1978 to 1990 he was a Lecturer. Associate Professor in the Department of Physics at *Chongqing University,* China and from 1990 to 1991 he was a visiting Scientist at the Gravitational Laboratory of the Sternberg State Astronomical Institute of the Moscow University in Russia. He was an Associate Professor, Department of Physics, *Chongqing University* from 1991 to 1994, was appointed Head of the Physics Department of *Chongqing University* from 1996 to 1998, and was appointed Dean of the college of Science at *Chongqing University* during the period 1998 to 2000. From 1994 to date, he has been a professor of Physics and Doctorate Tutor at the University. Now he is a member of the Council of the Chinese Physics Society, a member of the Council of the Chinese Gravitational and Relativity Astrophysical Society, a member of the World Laboratory, Chairman of *Chongqing Physics Society*, China and Head of the *Gravitational Physics Institute of Chongqing University*. His research fields include General Relativity and gravitation, classical and quantum electrodynamics in curved spacetime, theories of gravitational waves and gravitational radiation, positive definite problems of energy-momentum tensor of gravitational field, exact and approximate wave solutions of the Einstein field equations, gravitational perturbation effect in topological phonon space, interaction of gravitational waves with electromagnetic fields, and detection of high-frequency gravitational waves. Dr. Li has published more than sixty papers concerning gravitational waves in internationally recognized scientific journals and authored a paper at the first International HFGW Conference and

Workshop at The MITRE Corporation in 2003 participated in the second International HFGW Workshop at Austin, Texas in 2007.

Zhenyun Fang, PhD

Senior Scientific Advisor, China

Zhenyun Fang was born in Anhui China on May 4, 1945, graduate from Sichuan University in 1967 in Physics, received master's degree in Physics in 1982 at *Chongqing University (CQU)*. From 1983 to 1987 performed graduate studies at the *Unversite Catholique de Louvain (UCL)* in Belgium, and earned the PhD in Physics in 1987. From 1987 to 1990, was on the faculty of Department of Physics at Chongqing University (CQU) as an Associate Professor and Professor. From 1991, once as the Dean of the Graduate School, Assistant President of CQU. Currently, has been the Director of Institute of Theoretical Physics of CQU, Managing Director of *High Energy Society* of China, and has been interested in theoretical Physics, especially, in the research associate with the particle physics and related issues such as QED_QCD_SM (Standard Model) and beyond of electro-weak gauge-interaction, and high frequency gravitational-wave research, etc. for several decades. Dr. Fang participated in the second International HFGW Workshop at Austin, Texas in 2007.

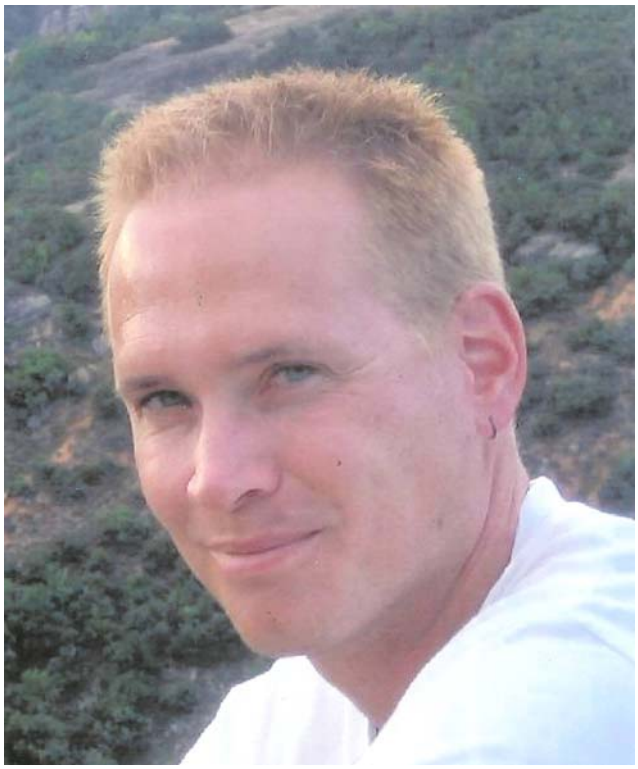
Ammar Sakaji, PhD

Senior Scientific Advisor, UAE

Ammar Sakaji was born in Jordan on August 28, 1967. Graduated from the *University of Jordan* with bachelor's degree in Physics in 1989 (first in his class), and Master Degree which was a research degree at that time in Theoretical Condensed Matter Physics in 1994 (first in his class), and finished the PhD course work and research on Theoretical Physics. He was a faculty member (lecturer and senior lecturer) in the physics department at *Zarqa Private University* and *Hashemite University* in Jordan, he came to UAE as a visiting professor at *Ajman University*, and then faculty member at Naval College. He is author or co-author of dozens of articles on Theoretical and Mathematical Physics, and editor or co-editor of several books and projects such as; "Majorana Legacy in Contemporary Physics," Di Renzo, Roma 2006, "Physics Of Emergence and Organization," World Scientific Publishing Company 2008, "Lev Davidovich Landau and his Impact on Contemporary Theoretical Physics" (Horizons in World Physics, Volume 264), Nova Science Publisher 2008, and "Quantum Field Theory," Springer 2009, and Quantum Information Science Project 2009. He also actively participated in several international conferences and workshops in the fields of Theoretical Physics, Quality Assurance in Higher Education, Intellectual Property, and E-Commerce and Legal Vacuum. He is the Editor in Chief of the *Electronic Journal of Theoretical Physics* (EJTP) www.ejtp.com, member of the *International Association of Mathematical Physics* (IAMP), member of the International Network for Quality Assurance Agencies in Higher Education (INQAAHE), referee of the *International Journal of Theoretical Physics*, he is a member of the *Planetary Society Berkeley*,

Jordan Physics Society, *American Institute of Physics*, and member of *Majorana Prize Committee* (Email: info@ejtp.com).

Harald Dimmelmeier, PhD



Senior Scientific Advisor, Germany

Harry Dimmelmeier was born in 1972 in Regensburg, Germany. He received his PhD in Physics from the Technical University Munich in 2001. He worked as a student at the Max Planck Institute for Gravitational Physics (Albert Einstein Institute) in Potsdam and as a post-doctoral scientist at the Max Planck Institute for Astrophysics in Garching. There he conducted research on numerical simulations of compact objects like stellar cores, neutron stars and black holes as astrophysical sources for gravitational waves. He received a Marie Curie fellowship from the EU to work at the Aristotle University in Thessaloniki on gravitational wave asteroseismology. He was also a part time technical editor for the free electronic scientific journal 'Living Reviews in Relativity'. He published many peer-reviewed papers on the numerical modeling of gravitational wave sources, participated in numerous international conferences on that subject, and closely collaborated with scientists working in data analysis for the large gravitational wave interferometer detectors. In 2002 he sponsored Dr. Baker's Lecture at the Max Planck Institute for Gravitational Physics (Albert Einstein Institute) in Potsdam. He recently moved to a position in industry and now works as a computational

engineer at AREVA NP in Erlangen in the field of safety analysis for nuclear power plants.

Kai Lin

Scientific Associate, China

Kai Lin was born in March, 1984 in P.R. China. In July 2006 he received a Bachelor of Science, in Physics at China West Normal University and in ·July 2009 he was awarded a Master of Science in Physics also at China West Normal University. Currently Mr. Lin is a doctorial student of Professor Li Fangyu and Professor Fang Zhenyun at the Institute of Theoretical Physics of Chongqing University and a visitor at Baylor University. His interests include: gravitational wave's theory and detection, cosmology, black hole thermodynamics, perturbations of black holes and computational physics. He has published in Physics Review D, Physics Letters B, Europhysics Letters, Modern Physics Letters A as well as other journals and is a member of the Chinese Physics Society.

CHAPTER

17 Gravitational Waves

Edmund Bertschinger & Edwin F. Taylor *

A modern physicist is a quantum theorist on Monday, Wednesday, and Friday, and a student of gravitational relativity theory on Tuesday, Thursday, and Saturday. On Sunday the physicist is neither, but is praying to his God that someone, preferably himself, will find the reconciliation between these two views.

—Norbert Wiener

I ask you to look both ways. For the road to a knowledge of the stars leads through the atom; and important knowledge of the atom has been reached through the stars."

—Arthur Eddington

1 INTRODUCTION

Gravity wave: a tidal force that propagates through spacetime.

General relativity differs from Newtonian gravity in several important ways. One way is in the behavior of light and matter in strong gravitational fields, especially near black holes. The black hole was predicted by Michell and Laplace on the basis of Newtonian gravity more than a century before Schwarzschild discovered his famous metric. However, the event horizon, singularity, and no-hair theorems are all consequences of general relativity that could not have been predicted from Newtonian physics.

Newton: Gravity propagates instantaneously.

Gravitational radiation is another phenomenon that has no counterpart in Newtonian physics. According to Newton, the gravitational interaction propagates instantaneously: When the Earth moves around the Sun, the Earth's gravitational field changes all at once throughout space, according to Newton.

Einstein: No signal propagates faster than light.

When Einstein formulated special relativity and recognized its requirement that no information can travel faster than the speed of light, he

*Draft of Second Edition of *Exploring Black Holes: Introduction to General Relativity.* Copyright ©2010 Edmund Bertschinger, Edwin F. Taylor, & John Archibald Wheeler. All rights reserved. Latest drafts at dropsite exploringblackholes.com, with a request for comments.

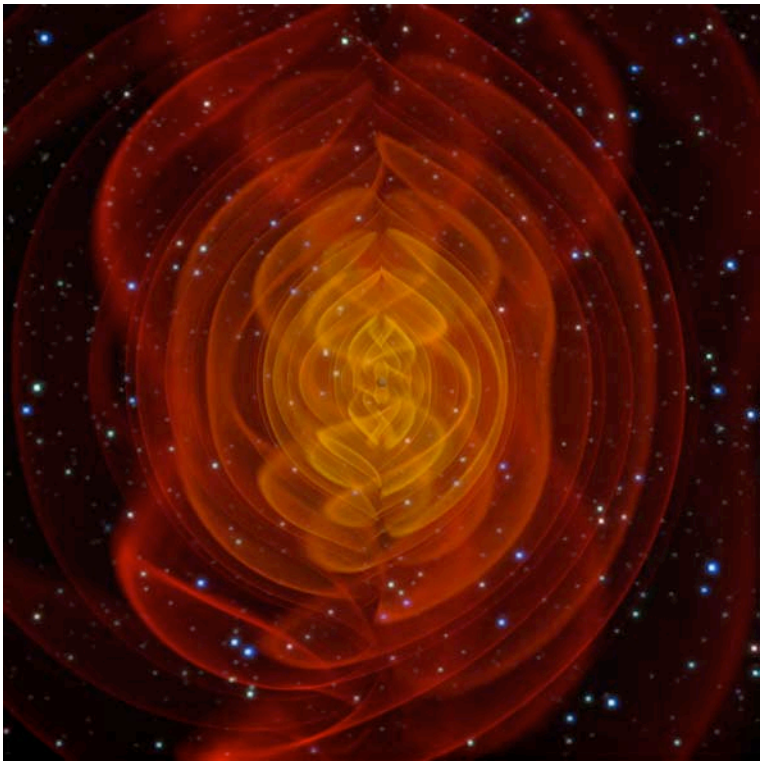


FIGURE 1 Computed emission of gravity waves. The tiny dot at the center of this image is two black holes churning spacetime as they combine into one. The swirling patterns represent distortions of spacetime that propagate outward as gravity waves. Close to the coalescing black holes, the gravity waves—essentially nothing but traveling tidal forces—are lethal. In contrast, we expect that gravity waves that could be detected on Earth are extremely small.

realized that Newtonian gravity would have to be modified. Not only would static gravitational fields differ from the Newtonian prediction in the vicinity of compact masses, but also time-varying fields would have to propagate. He showed that these fields would move with the speed of light, so gravity could not be used to send information faster than the speed of light, which would have destroyed the fundamental basis of all relativity.

Gravity waves like electromagnetic waves?

Einstein had a conceptual prototype for gravity waves: electromagnetic radiation. James Clerk Maxwell predicted electromagnetic radiation in 1873 and Heinrich Hertz demonstrated it experimentally in 1888. (Einstein was born in 1879.) When he grew up, Einstein quickly realized that a general relativity theory based on curved spacetime would not look like Maxwell's electromagnetic theory. After his theory was completed, Einstein and others were able to compute how gravitational fields propagate.

Gravity waves are essentially tidal forces that vary with time and position; that is all they are. As a gravity wave passes over you, you are alternately stretched and compressed in ways that depend on the particular form of the

wave. In principle there is no limit to the size of gravity waves. Figure 1 pictures the calculated result of two black holes emitting gravity waves as they combine into one. In the vicinity of the coalescence, gravity-wave-induced tidal forces would be dangerous to life.

We predict that gravity waves from various sources are continually sweeping over us on Earth's surface. Sections 3 and 7 describe some of these sources. Basically we hope to observe these waves by detecting changes in separation between two test masses suspended near to one another—changes in separation caused by the traveling tidal force that constitutes a gravity wave. We expect this change in separation to be *extremely* small for gravity waves detectable on Earth.

Gravity wave on Earth: An extremely small traveling tidal force.

Current gravity wave detectors on Earth are interferometers in which light is reflected back and forth between free test masses along two perpendicular directions, and the time difference measured between round-trip times in the two directions. The "free" test masses are hung from wires that are in turn supported with elaborate shock-absorbers so as to minimize the vibrations due to passing trucks and even waves crashing on a distant shore. But the back-and-forth pendulum-like motions of these test masses are free enough to permit measurement of their change in separation due to tidal effects resulting from a passing gravity wave, caused by some gigantic distant gravitational event, for example the coalescence of two black holes modeled in Figure 1.

Does the change in separation induced by gravity waves affect everything, for example a meter stick or the concrete slab on which a gravity wave detector rests? Answer: Only by an amount that is entirely negligible. The structure of meter sticks and concrete slabs is determined by electromagnetic forces mediated by quantum mechanics. The two ends of a meter stick are not freely-floating test masses. The tidal force of a passing gravity wave is much weaker than the internal forces that maintain the shape of solids. The meter stick—or the concrete slab underlying the vacuum chamber and detectors of a gravitational-wave observatory—is stiff enough to be negligibly affected by a passing gravity wave.

2. GRAVITY WAVE METRIC

Tiny but significant departure from the inertial metric

Our analysis uses a particular gravity wave: a certain kind of plane wave arriving from a very distant source and moving in the z-direction. This wave (and almost all of the gravity waves we discuss in this chapter) represents a very small perturbation of flat spacetime. Here is the timelike metric for such a particular wave that propagates along the z-axis.

Gravity wave metric

$$d\tau^2 = dt^2 - (1+h)dx^2 - (1-h)dy^2 - dz^2 \qquad (h \ll 1) \tag{1}$$

In this metric h is a dimensionless function of time and space. *Numerically, h is a fractional deviation from the flat-spacetime coefficient of dx^2 or dy^2 in the metric.* Another name for fractional deviation of length is **strain**, so h is also

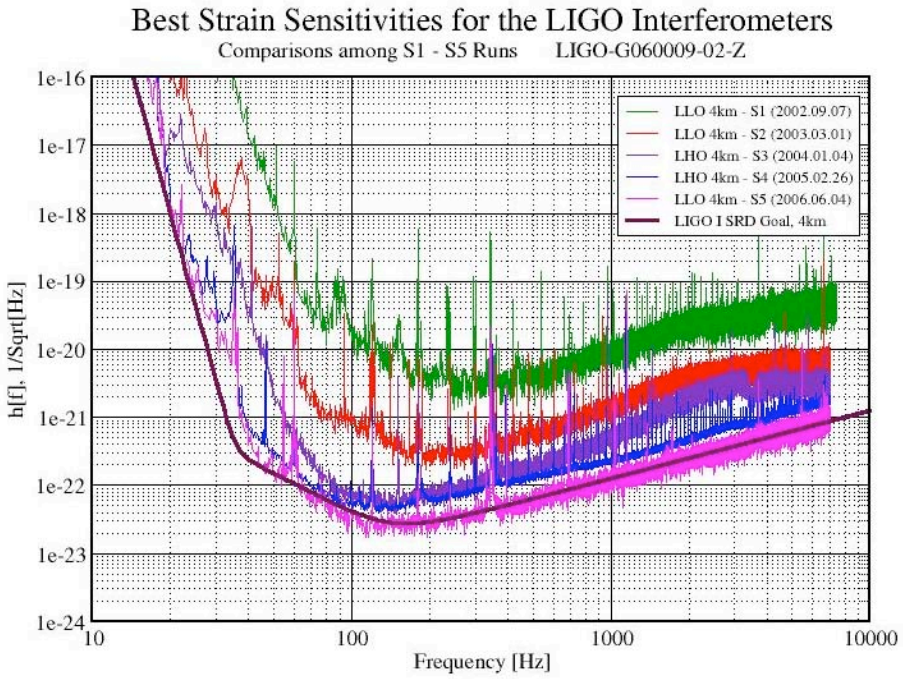


FIGURE 2 Progressive improvements in sensitivity of LIGO interferometers. On the vertical axis 1e-19, for example, means a fractional change in separation of 10^{-19} between test masses. The bottom solid line is the current goal. Spikes occur at frequencies of electrical or acoustical noise. To be detectable, gravity wave signals must cause greater displacement than what is represented by these noise curves. As of mid-2009, gravity wave signals have yet to be detected. FIND MOST RECENT UPDATE THAT SHOWS EVOLUTION OF SENSITIVITY: SCOTT HUGHES?

h = gravity wave strain

called the **gravity wave strain**. The wave leading to (1) is a transverse wave, since h describes a perturbation of space only in the x and y directions transverse to the z-direction of propagation. The strain h varies with both position and time. Its maximum value is very much less than one. Let two free test masses be at rest a distance D apart in the x or y direction. When a z-directed gravity wave passes over them, the change in their separation, called the **displacement**, equals hD, which follows directly from the definition of h as a "fractional deviation."

FOLLOWING IS NUMERICALLY INCONSISTENT. GET LATEST LIGO PARAMETERS AND SENSITIVITIES. One can use Einstein's field equations to make predictions about the magnitude of the function h in equation (1) for various kinds of astronomical phenomena. Currently, gravity wave detectors use laser interferometry and go by the full name **Laser Interferometer Gravitational Wave Observatory** or **LIGO** for short. The first-generation LIGO, called Initial LIGO, was able to detect waves with (approximately) $h > 10^{-19}$ for frequencies within a range of about 100 hertz.

LIGO gravity wave detector

(Abbreviation: Hz. Recall that one hertz is one cycle per second.) The second-generation LIGO, called Advanced LIGO, is about 10 times more sensitive; it is planned to be operational around 2014. Advanced LIGO can also be tuned in frequency to achieve higher sensitivity in frequency bands of interest.

LIGO sensitivity

Figure 2 compares the gradually-improving sensitivities of LIGO over time. The displacement sensitivity is expressed in the units of meter/$(\text{hertz})^{1/2}$ because the amount of noise limiting the measurement grows with the frequency range being sampled. Note that the instruments are designed to be most sensitive near 150 hertz. This frequency is determined by the different kinds of noise faced by experimenters: Quantum noise limits the sensitivity at high frequencies, while seismic noise is the largest problem at low frequencies. If the range of sampled frequencies—*bandwidth*—is 100 hertz, then the best sensitivity is about $10^{-22} \times 100^{1/2} = 10^{-21}$. This means that along a length of 4 kilometers $= 4 \times 10^3$ meters, the change in length is approximately $10^{-21} \times 4 \times 10^3 = 4 \times 10^{-18}$ meters, which is approximately one thousand times smaller than a proton, or a hundred million times smaller than a single atom!

Hold on! Your gravity wave detector sits on Earth's surface, but equation (1) says nothing about curved spacetime described, for example, by the Schwarzschild metric. The expression $2M/r$ measures departure from flatness in the Schwarzschild metric. At Earth's surface, $2M/r \approx 1.4 \times 10^{-9}$, which is 10^{13}—ten million million!—times greater than the corresponding gravity wave factor $h \sim 10^{-22}$. Why doesn't the quantity $2M/r$—which is much larger than h—appear in (1)?

First, the factor $2M/r$ is essentially constant over the size of LIGO, so we can ignore it. Second—and more important—the LIGO detector is "tuned" to detect a time-varying gravity wave of frequency near 150 hertz. LIGO is totally insensitive to the small *static* curvature introduced by the factor $2M/r$ at Earth's surface. For this reason, we simply omit static curvature factors from equation (1), effectively describing gravity waves "in free space" as well as for the predicted $h \ll 1$.

Einstein's equations become a wave equation.

In free space and for small values of h, Einstein's field equations actually reduce to a wave equation for h. For the most general case, this wave has the form $h = h(x, y, z, t)$. When x, y, z, and t are all in geometric units (for example meters), this wave equation takes the form:

$$\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2} + \frac{\partial^2 h}{\partial z^2} = \frac{\partial^2 h}{\partial t^2} \qquad \text{(free space and } h \ll 1\text{)} \tag{2}$$

For simplicity, think of a wave moving along the z-axis. The most general solution to the wave equation under these circumstances is

$$h = h_{+z}(z - t) + h_{-z}(z + t) \tag{3}$$

Assume gravity wave moves in $+z$ direction.

The expression $h_{+z}(z - t)$ means a function h of the variable $z - t$ and *not* some constant h *times* the quantity $(z - t)$. The function $h_{+z}(z - t)$ describes

a wave moving in the positive z-direction and the function $h_{-z}(z+t)$ describes a wave moving in the negative z-direction. In this chapter we deal only with a gravity wave propagating in the positive z-direction and hereafter use

$$h \equiv h(z-t) \equiv h_{+z}(z-t) \qquad \text{(wave moves in } +z \text{ direction)} \tag{4}$$

The argument $z-t$ means that h is a function of *only* the combined variable $z-t$. Indeed, h can be *any function whatsoever* of the variable $(z-t)$. The form of this variable tells us that, whatever the profile of the gravity wave at any given time; as time passes, that profile displaces itself in the positive z-direction with the speed of light (one in our units) .

LIGO sensitive 75 to 500 hertz

Figure 2 shows that the LIGO gravity wave detector has maximum sensitivity to gravity waves of frequencies between 75 and 500 hertz, with a peak sensitivity at around 150 hertz. Even at 500 hertz, the wavelength of the gravity wave is very much longer than the overall 4-kilometer dimensions of the LIGO detector. Therefore *we can assume in the following that at any given time the value of h is spatially uniform over the entire LIGO detector.*

QUERY 1. Uniform h?
Using numerical values, verify the claim in the preceding paragraph that h is effectively uniform over the LIGO detector.

Draw global map coordinates on rubber sheet.

It is important to understand the meaning of the coordinates in metric (1). These are global *map* coordinates; global coordinates are always fictions that we choose to reveal aspects of a spacetime we cannot visualize. For $h \neq 0$, these global coordinates are invariably distorted. Think of the three mutually perpendicular planes formed by pairs of space coordinates (x, y), (y, z), and (z, x). Draw a grid of lines on a rubber sheet lying in each corresponding plane. The gravity wave distorts the rubber sheet as it passes through it.

Glue map clocks to the intersections of these grid lines on the rubber sheet so that they move as the rubber sheet distorts. A gravitational wave moving in the $+z$ direction (Figure 5) passes through a rubber sheet lying in the xy plane, so that the grid ruled on the rubber sheet stretches and contracts in different directions within the plane of the sheet (Figures 3 and 4). The map clocks, glued at intersections of map coordinate grid lines, ride along with the grid as the sheet distorts,so that the map coordinates of any clock does not change.

Map time t read on clocks glued to the rubber sheet.

Think of two ticks on a single map clock. Between ticks the map coordinates of the clock do not change: $dx = dy = dz = 0$. Therefore metric (1) tells us that the wristwatch time $d\tau$ between two ticks is also map time dt between ticks. Therefore map time t corresponds to the time measured on the clocks glued to the rubber sheet, even when the strain h varies at their locations.

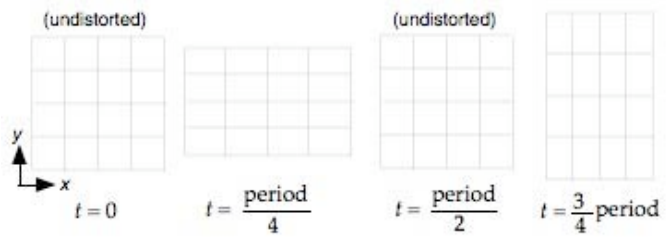


FIGURE 3 Change in shape (greatly exaggerated!) of the map coordinate grid at four times as a periodic wave passes through in the z-direction (perpendicular to the page). NOTE carefully!: The x-axis is stretched while the y-axis is compressed and vice versa. The areas of the panels remain the same.

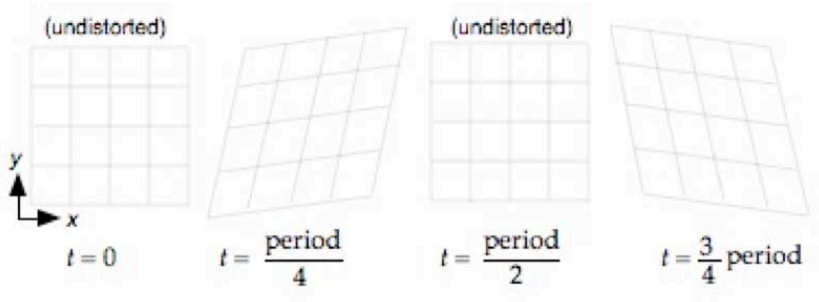


FIGURE 4 Effects of a periodic gravity wave with polarization "orthogonal" to that of Figure 3 on the map grid in the xy plane. Note that the axes of compression and expansion are at 45 degrees from the x and y axes. All grids stay in the xy plane as they distort. As in Figure 3, the areas of the panels are all the same.

Figure 3 represents the map time variation of the space distortion of the rubber sheet at a given location due to a particular polarization of the gravity wave. Although gravity waves are transverse like electromagnetic waves, the polarization forms of gravity waves are different from those of electromagnetic waves. Figure 4 shows the distortion caused by the wave "orthogonal" to that shown in Figure 3.

3 ■ SOURCES OF GRAVITY WAVES

Many sources; only one with clear prediction

Sources of gravity waves include collapsing stars, exploding stars, stars in orbit around one another, and the big bang. Neither electromagnetic waves nor gravity waves result from a spherically symmetric distribution of charge (for electromagnetic waves) or matter (for gravitational waves), even when that spherical distribution pulses symmetrically in and out. Therefore, *symmetric* collapses or explosions emit no waves, either electromagnetic or gravitational. The most efficient source of electromagnetic radiation is oscillating pairs of electric charges of opposite sign, the result technically called **dipole radiation**. But mass has only one "polarity;" there is no gravity dipole radiation from masses that oscillate back and forth along a line. Emission of gravity waves requires *asymmetric* movement or oscillation; the technical name for the simplest result is **quadrupole radiation**. Happily, most collapses and explosions are asymmetric; even the motion in a binary system is sufficiently asymmetric to emit gravitational waves.

We study here gravity waves emitted by a binary system consisting of two neutron stars—or a neutron star and a black hole—orbiting about one another (Section 6). All such pairs that we have detected are too far away to see directly; at least one neutron star needs to be a **pulsar** that emits a steady stream of pulses that we can receive at a great distance. Pulsars turn out to be extremely stable clocks. As the two objects orbit, they also emit gravity waves that cause the binary system to lose energy, so that the orbiting objects gradually spiral in toward one another. These orbits are fairly well described by Newtonian mechanics until about one millisecond before the two objects coalesce.

Binary system emits gravity waves . . .

Emitted gravity waves are nearly periodic during the Newtonian phase of orbital motion. As a result, these particular gravity waves are easy to predict and therefore easy to search for. When the two objects coalesce, they emit a burst of gravity waves (Figure 11). After coalescence the resulting structure vibrates ("rings down"), emitting more gravity waves as it settles into its final state as a black hole. Initial LIGO has already completed its efforts and would have been sensitive enough (Figure 2) to detect binary neutron star systems coalescing at a distance of about 26 million light years. Unfortunately, no such coalescences were detected during more than one year of observation. Advanced LIGO extends the detection radius to 200 Megaparsecs $\approx$ 650 million light years. The volume of space increases approximately as the cube of the distance, so the improved sensitivity will vastly increase the number of galaxies that can be "seen" by LIGO from about one thousand to millions, increasing the odds of success thousands of times.

. . . whose amplitude is predictable.

QUERY 2. Increased volume of detection

Use numerical values given in the preceding paragraph to calculate to two significant figures the increased "odds of success" of Advanced LIGO compared with Initial LIGO.

From other sources: hard to predict.

Binary coalescence is the only source for which we can currently make a clear prediction of the signal (and therefore of the detection distance limit). Other conceivable sources include supernovae and the collapse of a massive star to form a black hole—the event that triggers so called **gamma-ray bursts**. But we have only speculations about how far away any of these can be and still be detectable by either Initial LIGO or Advanced LIGO.

DETECTORS DO NOT AFFECT GRAVITY WAVES

We are used to the fact that metal structures can distort or reduce the amplitude of electromagnetic waves passing across them. Even the presence of a receiving antenna can distort an electromagnetic wave in its vicinity. The same is not true of gravity waves, whose generation or modification requires massive moving structures. Gravity wave detectors have negligible effect on the waves that they are designed to detect. Indeed, it is the smallness of the influence that gravity waves have on mechanical structures that makes gravity waves so difficult to detect.

QUERY 3. Electromagnetic waves *vs.* gravity waves. Discussion.
What property of electromagnetic waves makes their interaction with conductors so huge compared with the interaction of gravity waves with matter of any kind?

4 ■ MOTION OF LIGHT IN MAP COORDINATES

Light reflected back and forth between mirrored test masses.

The LIGO detector is an *interferometer* that employs mirrors mounted on "test masses" suspended at rest at the ends of an L-shaped vacuum cavity. The length of each leg of the L is 4 kilometers. Detection of the gravity wave is accomplished by measuring the relative round-trip *time delay* between light sent down one leg of the detector and light sent down the other, perpendicular leg.

LIGO is an interferometer.

Suppose that a gravity wave of the polarization illustrated in Figure 3 moves in the z-direction as shown in Figure 5 and that one leg of the detector lies along the x-direction and the other leg along the y-direction. In order to analyze the operation of LIGO, we need to know (a) how light propagates along the x and y legs of the interferometer and (b) how the test masses at the ends of the legs move when the z-directed gravity wave passes over them. In the present section we analyze the motion of light in map coordinates; Section 5 begins the description of the motion of test masses in map coordinates.

Motion of light in map coordinates.

With what map speed does light move in the x-direction in the presence of a gravity wave implied by metric (1)? To answer this question, set $dy = dz = 0$ in that equation, yielding

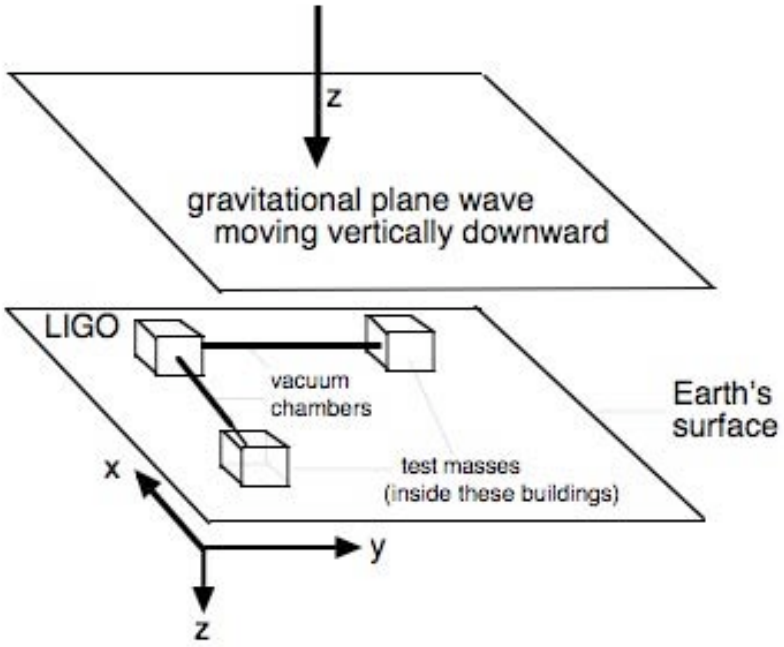


FIGURE 5 Perspective drawing of the relative orientation of legs of the LIGO interferometer lying in the x and y directions on the surface of Earth and the z-direction of the incident gravity wave descending vertically. [Illustrator: Rotate lower plate and contents CCW 90 degrees, so corner box is above the origin of the coordinate system. Same for Figure 10.]

$$d\tau^2 = dt^2 - (1+h)dx^2 \tag{5}$$

As always, the proper time is zero between two adjacent events on the worldline of a light pulse. Set $d\tau = 0$ to find the speed of light in the x-direction.

$$\frac{dx}{dt} = \pm(1+h)^{-1/2} \qquad \text{(light moving in } x \text{ direction)} \tag{6}$$

The plus and minus signs correspond to a pulse traveling in the positive or negative x-direction, respectively—that is, in the plane of LIGO in Figure 5. Remember that the magnitude of h is very much smaller than one, so we use the approximation inside the front cover. To first order:

$$(1+\epsilon)^n \approx 1 + n\epsilon \qquad |\epsilon| \ll 1 \text{ and } |n\epsilon| \ll 1 \tag{7}$$

Apply this approximation to (6) to obtain

$$\frac{dx}{dt} \approx \pm(1 - \frac{h}{2}) \qquad \text{(light moving in } x \text{ direction)} \tag{8}$$

Gravity wave modifies map speed of light.

In words, the map speed of light is changed (slightly!) by the presence of our gravity wave. Since h is a function of time as well as position, the map speed of light in the x-direction is not constant, but varies as the wave passes through. (Should we worry that the speed in (8) does not have the standard value one? No! This is a *map speed*—a mythical beast—measured directly by no one.)

By similar arguments, the map speeds of light in the y and z directions for the wave described by the metric (1) are:

$$\frac{dy}{dt} \approx \pm(1 + \frac{h}{2}) \qquad \text{(light moving in } y \text{ direction)} \tag{9}$$

$$\frac{dz}{dt} = \pm 1 \qquad \text{(light moving in } z \text{ direction)} \tag{10}$$

5 MOTION OF LIGO TEST MASSES IN MAP COORDINATES

"Obey the Principle of Maximal Aging!"

Consider two test masses with mirrors suspended at opposite ends of the x-leg of the detector. The signal of the interferometer due to the motion of light along this leg will be influenced only by the x-motion of the test masses due to the gravity wave. In this case the metric is the same as (5).

How does a test mass move as the gravity wave passes over it? As always, we answer this question with the Principle of Maximal Aging, maximizing the wristwatch time of the test mass across two adjoining segments of its worldline between fixed end-events. In what follows we verify the surprising result anticipated in Section 2 above, namely that a test mass initially at rest in map coordinates rides with the expanding and contracting map coordinates drawn on the distorting rubber sheet, so this test mass does not move with respect to map coordinates as a gravity wave passes over it. This result comes from showing that an out-and-back jog in the vertical worldline in map coordinates leads to smaller aging and therefore does not occur for a free test mass

Figure 6 pictures this idealized case: an incremental linear deviation from a vertical worldline from origin 0 to the event at $t = 2t_0$. Along Segment A the displacement x increases linearly with time: $x = v_0 t$, where the speed v_0 is a constant. Along segment B the displacement returns to zero at the same constant rate. The strain h has average values $\bar{h}_{\rm A}$ and $\bar{h}_{\rm B}$ along segments A and B respectively. We use the Principle of Maximal Aging to find the value of the speed v_0 that maximizes the wristwatch time along this worldline. We will find that $v_0 = 0$. In other words, the free test mass initially at rest in map coordinates stays at rest in map coordinates; it does not deviate from the vertical worldline in Figure 6. Now for the details.

Idealized case: Linear jogs out and back.

Write the metric (5) in approximate form for one of the segments:

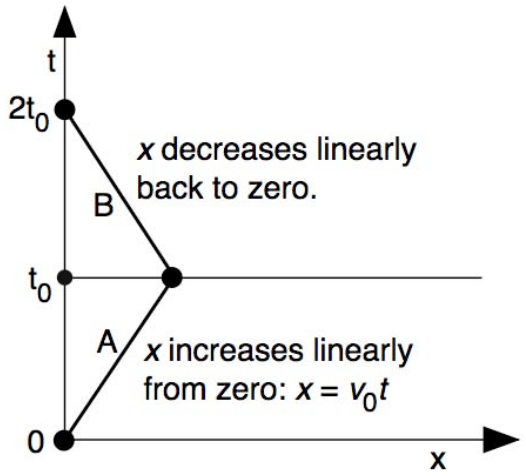


FIGURE 6 Trial worldline for a test mass; incremental departure from vertical line of a particle at rest. Segments A and B are very short.

$$\Delta\tau^2 \approx \Delta t^2 - (1+\bar{h})\Delta x^2 \tag{11}$$

where $\bar{h}$ is an average value of the strain h across that segment. Now we apply (11) first to Segment A in Figure 6, then to Segment B. We are going to take derivatives of these expressions, which will look awkward applied to Δ symbols. Therefore we temporarily ignore the Δ symbols in (12) and let τ stand for $\Delta\tau$, t for Δt, and x for Δx, holding in mind that these symbols actually represent increments, so equations in which they appear are approximations.
With these substitutions, equation (11) becomes, for the two adjoining worldline segments:

$$\tau_{\mathrm{A}} \approx \left[t_0^2 - \left(1+\bar{h}_{\mathrm{A}}\right)\left(v_0 t_0\right)^2\right]^{1/2} \qquad \text{Segment A} \tag{12}$$

$$\tau_{\mathrm{B}} \approx \left[t_0^2 - \left(1+\bar{h}_{\mathrm{B}}\right)\left(v_0 t_0\right)^2\right]^{1/2} \qquad \text{Segment B}$$

so that the total wristwatch time along the bent worldline from $t = 0$ to $t = 2t_0$ is the sum of the right sides of equations (12).

We want to know what value of v_0 (the out-and-back speed of the test mass) will lead to a maximal value of the total wristwatch time. To find this, take the derivative with respect to v_0 of the sum of individual proper times and set the result equal to zero.

$$\frac{d\tau_{\mathrm{A}}}{dv_0} + \frac{d\tau_{\mathrm{B}}}{dv_0} \approx -\frac{(1+\bar{h}_{\mathrm{A}})v_0 t_0^2}{\tau_{\mathrm{A}}} - \frac{(1+\bar{h}_{\mathrm{B}})v_0 t_0^2}{\tau_{\mathrm{B}}} = 0 \tag{13}$$

Initially at rest in map coordinates? Then stays at rest in map coordinates.

In middle expression of (13), all quantities are fixed except for v_0. The only way that (13) can be satisfied is if $v_0 = 0$. *The test mass initially at rest does not change its map x-coordinate as the gravity wave passes over.*

Our result seems rather specialized in two senses: First, it treats only the vertical worldline traced out by a test mass at rest. Second, it deals only with a very short segment of the worldline, along which $\bar{h}$ is considered to be nearly constant. Concerning the second point, you can think of (13) as a tiny out-and-back "jog" *anywhere* on a much longer vertical worldline. Then our result implies that *any* jog in the vertical worldline does not lead to an increased value of the wristwatch time, even if h varies a lot over a longer stretch of the worldline.

The first specialization, the vertical worldline, is important: The gravity wave does not cause a kink in a *vertical* map worldline. The same is typically *not* true for a particle that is moving in map coordinates before the gravity wave arrives. (We say "typically" because the kink may not appear for some directions of motion of the test mass and for some polarization forms and directions of propagation of the gravity wave.) In this more general case, a kink in the worldline corresponds to a change of velocity. In other words, a passing gravity wave can change the map velocity of a moving particle just as if it were a velocity-dependent force. If the particle velocity is zero, then the force is zero: a particle at rest in map coordinates remains at rest.

Not at rest in map coordinates? Maybe kink in map worldline.

QUERY 4. Disproof of relativity? (optional)
"AHA!" declares Kristin Burgess. "Now I can disprove relativity once and for all. If the test mass *moves*, a passing gravity wave can cause a kink in the worldline of the test mass as observed in the local inertial Earth frame. No kink appears in its worldline if the test mass is at rest. But if a worldline has a kink in it as observed in one inertial frame, it will have a kink in it as observed in all overlapping relatively-moving inertial frames. An observer in any such frame can detect this kink. So the *absence* of a kink tells me *and every other inertial observer* that the test mass is 'at rest'? We have found a way to determine absolute rest using a local experiment. Goodbye relativity!" Is Kristin right? (A detailed answer is beyond the scope of this book, but you can use some relevant generalizations drawn from what we already know to think about this paradox. As an analogy from flat-spacetime electromagnetism, think of a charged particle at rest in a purely magnetic field: The particle experiences no magnetic force. In contrast, when the same charged particle moves in the same frame, it may experience a magnetic force for some directions of motion.)

In this book we make every measurement in a local inertial frame, not using differences in global map coordinates. So of what possible use is our result that a particle at rest in global coordinates does not move in those coordinates when a gravity wave passes over it? Answer: Just because something is at rest in map coordinates does not mean that it is at rest in local inertial Earth coordinates. In the following section we find that a gravity wave *does* move a test mass as observed in the Earth coordinates. LIGO—attached to the Earth—can detect gravity waves!

At rest in map coordinates? Still can move in Earth coordinates.

6 ■ DETECTION OF A GRAVITY WAVE BY LIGO

Make measurement in the local Earth frame.

Suppose that the gravity wave that satisfies metric (1) passes over the LIGO detector oriented as in Figure 5. We know how the test masses at the two ends of the legs of the detector respond to the gravity wave: they remain at rest in map coordinates (Section 5). We know how light propagates along both legs: as the gravity wave passes through, the map speed of light varies slightly from the value one, as given by equations (8) through (10) in Section 4.

Earth frame tied to LIGO slab

The trouble with map coordinates is that they are arbitrary and need not correspond to what an observer measures. Recall that we require all measurements to take place in a local inertial frame. So think of a local reference frame anchored to the concrete slab on which LIGO rests. As explained in the Introduction (Section 1), the gravity wave has essentially no effect on this slab. Call the coordinates in the resulting local coordinate system **Earth coordinates**. Earth coordinates are analogous to shell coordinates for the Schwarzschild black hole; useful only locally but yielding the numbers that predict results of measurements. The metric for the local inertial frame then has the form:

$$\Delta\tau^2 \approx \Delta t^2_{\rm Earth} - \Delta x^2_{\rm Earth} - \Delta y^2_{\rm Earth} - \Delta z^2_{\rm Earth} \tag{14}$$

Compare this with the approximate version of (1):

$$\Delta\tau^2 \approx \Delta t^2 - (1+h)\Delta x^2 - (1-h)\Delta y^2 - \Delta z^2 \qquad (h \ll 1) \tag{15}$$

Legalistically, in order to make the coefficients in (15) constants we should use the symbol $\bar{h}$, with a bar over the h, to indicate the average value of the gravity wave amplitude over the detector. However, in Query 1 you showed that for the frequencies at which LIGO is sensitive, the wavelength is very much greater than the dimensions of the detector, so the amplitude h of the gravity wave is effectively uniform across the LIGO detector. Therefore it is not necessary to take an average, and we use the symbol h without a superscript bar.

Earth frame coordinate differences

Comparing (14) with (15) yields:

$$\begin{aligned}
\Delta t_{\rm Earth} &= \Delta t \\
\Delta x_{\rm Earth} &= (1+h)^{1/2}\Delta x \approx (1+\frac{h}{2})\Delta x \qquad h \ll 1 \\
\Delta y_{\rm Earth} &= (1-h)^{1/2}\Delta y \approx (1-\frac{h}{2})\Delta y \qquad h \ll 1 \\
\Delta z_{\rm Earth} &= \Delta z
\end{aligned} \tag{16}$$

where we use approximation (7). Notice, first, that Earth time lapse $\Delta t_{\rm Earth}$ between two events is identical to their map time lapse Δt and the z component of their space separation in Earth coordinates, $\Delta z_{\rm Earth}$, is identical to the z component of their separation in map coordinates, Δz.

Now for the differences! Let Δx be the map x-coordinate separation between the pair of mirrors in the x-leg of the LIGO interferometer and Δy be the map separation between the corresponding pair of mirrors in the y-leg. As the z-directed wave passes through the LIGO detector, the test masses at rest at the ends of the legs stay at rest in map coordinates, as Section 5 showed.

Test masses move in Earth coordinates.

Therefore the value of Δx remains the same during this passage, as does the value of Δy. But the presence of the time-varying strain $h(t)$ in (16) tells us that these test masses move when observed in Earth coordinates. More: when the distance between test masses increases (say) along the Earth x-axis, it decreases along the perpendicular Earth y-axis; and vice versa. Perfect for detection of a gravity wave by an interferometer!

Light speed = 1 in Earth frame.

Earth metric (14) is that of an inertial frame in which the speed of light has the value one in whatever direction it moves. With light we have the opposite weirdness to that of the motion of test masses initially at rest: In map coordinates light moves at speeds different from one in the presence of this gravity wave—equations (8) through (10)—but in Earth coordinates light moves with speed one. This is reminiscent of the corresponding case near a Schwarzschild black hole: Light moves at speeds different from one in Schwarzschild map coordinates but at speed one in shell coordinates.

In summary the situation is this: As the gravity wave passes over the LIGO detector, the speed of light propagating down the two legs of the

Different Earth times along different legs

detector has the usual value one as measured by the Earth observer. However, for the Earth observer the separations between the test masses along the x-leg and the y-leg change: one increases while the other decreases, as given by equations (16). The result is a difference in the round-trip times of light along the two legs. It is this difference that LIGO is designed to measure and thereby to detect the gravity wave.

What will be the value of this difference in round-trip times between light propagation along the two legs? Let D be the length of each leg in the absence of the gravity wave. The round-trip time is twice this length divided by the speed of light, which has the value one in Earth coordinates. From equations (16) we find that the difference in round-trip times between light propagated along the two legs is

$$\Delta t_{\text{Earth}} = 2D\left(\frac{h}{2} + \frac{h}{2}\right) = 2Dh \qquad \text{(one round trip of light)} \tag{17}$$

Time difference after N round trips.

Using the latest interferometer techniques, LIGO reflects the light back and forth down each leg approximately $N = 140$ times. That is, light executes approximately 140 round trips, which multiplies the detected time shift, increasing the sensitivity of the detector by the same factor. Equation (17) becomes

$$\Delta t_{\text{Earth}} = 2NDh \qquad (N \text{ round trips of light}) \tag{18}$$

Quantities N and h have no units, so the unit of time in (18) is the same as the unit of D, for example meters.

QUERY 5. LIGO fast enough?
Do the 140 round trips of light take place in a time small compared with one period of the gravity wave being detected? (If it does not, then LIGO detection is not fast enough to track the change in gravity strain.)

QUERY 6. Application to LIGO.
Each leg of the LIGO interferometer is of length $D = 4$ kilometers. Assume that the laser emits light of wavelength 1000 nanometer $= 10^{-6}$ meter (infrared light from a NdYAG laser). Suppose that we want LIGO to reach a sensitivity of $h = 10^{-22}$. For $N = 140$, find the corresponding value of Δt_{Earth}. Express your answer as a decimal fraction of the period T of the laser light used in the experiment. (For background see `http://www.ligo.caltech.edu/LIGO_web/about/`)

QUERY 7. Faster derivation?
In this book we insist that global map coordinates are arbitrary human choices, so that we cannot depend on map coordinate differences to be directly measurable quantities. However, the value of h in (1) is so small that the metric differs only slightly from an inertial metric. This once, therefore, we treat map coordinates as directly measurable and ask you to redo the derivation of equations (17) and (18) using only map coordinates.

Remember that test masses initially at rest in map coordinates do not change their coordinates as the gravity wave passes over them (Section 4), but the gravity wave alters the map speeds of light, and differently in the x-direction, equation (8), than in the y-direction, equation (9). Assume that each leg of the interferometer has the length D_{map} in map coordinates.

A. Find an expression for the difference Δt in map time between the two legs for one round trip of the light.

B, How great do you expect the difference to be between times Δt and Δt_{Earth} and the difference between distances D (in Earth coordinates) and D_{map}? Taken together, will these differences be great enough so that the result of your prediction and that of equation (18) could be distinguished experimentally?

QUERY 8. Different directions of propagation of the gravity wave
Thus far we have assumed that the gravitational plane wave of the polarization described by equation (1) descends vertically onto the LIGO detector, as shown in Figure 5. Of course the observers cannot prearrange in what direction an incident gravity wave will move. Suppose that the wave described by (1) propagates along the direction of, say, the y-leg of the interferometer, while the x-direction lies along the other leg, as before. What is the equation that replaces (18) in this case?

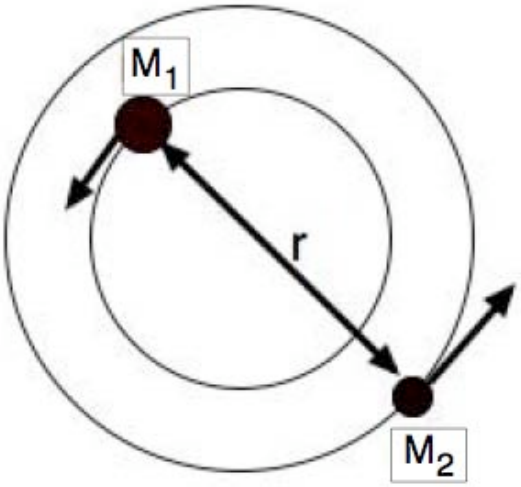


FIGURE 7 A binary system with each object in a circular path.

QUERY 9. LIGO fails to detect a gravity wave?
Think of various directions of propagation of the gravity wave pictured in Figure 3, together with different directions of x and y in equation (1) with respect to the LIGO detector. Give the name **orientation** to a given set of directions x and y—the transverse directions in (1)—plus z (the direction of propagation) in (1) relative to the LIGO detector. How many orientations are there for which LIGO will detect *no signal whatever*, even when its sensitivity is 10 times better than that needed to detect the wave arriving in the orientation shown in Figure 5? Are there zero such orientations? one? two? three? some other number less than 10? an infinite number?

7 ■ BINARY SYSTEM AS A SOURCE OF GRAVITY WAVES

Proof of the existence of gravity waves?

Unequal masses, each in circular orbit.

Now we consider in more detail gravity waves generated by a binary system consisting of two neutron stars, each in circular orbit around their center of mass. The binary system is the only known example of a stellar system for which we can explicitly calculate the emitted gravity waves. Suppose that the stars of the binary system have masses M_1 and M_2 and are assumed to orbit at a constant distance r apart, as shown in Figure 7.

Energy of the system.

The basic parameters of the orbit are adequately computed using Newtonian mechanics, according to which the energy of the system in conventional units is given by the expression:

$$E_{\rm conv} = -\frac{GM_{1,\rm kg}M_{2,\rm kg}}{2r} \qquad \text{(Newton)} \tag{19}$$

Rate of energy loss . . .

As these neutron stars orbit, they generate gravity waves. General relativity predicts the rate at which the orbital energy is lost to this radiation. In conventional units, this rate is:

$$\frac{dE_{\rm conv}}{dt_{\rm conv}} = -\frac{32G^4}{5c^5r^5}\left(M_{1,\rm kg}M_{2,\rm kg}\right)^2\left(M_{1,\rm kg}+M_{2,\rm kg}\right) \qquad \text{(Newtonian circular orbits)} \tag{20}$$

. . . derived from Einstein's equations.

Equation (20) assumes that the two stars are separated by much more than their Schwarzschild radii and that they are moving at nonrelativistic speeds. Deriving equation (20) involves a lengthy and difficult calculation starting from Einstein's field equations. The same is true of the derivation of the metric (1) for a gravity wave. These are two of only three equations in this chapter that we simply quote from a more advanced treatment of general relativity.

QUERY 10. Energy and rate of energy loss in geometric units

Convert equations (19) and (20) to geometric units to be consistent with our notation and to get rid of the constants G and c. Use the sloppy professional shortcut, "Let $G = c = 1$."

A. Show that (19) and (20) become:

$$E = -\frac{M_1M_2}{2r} \qquad \text{(Newton: geometric units, } E \text{ in units of length)} \tag{21}$$

$$\frac{dE}{dt} = -\frac{32}{5r^5}\left(M_1M_2\right)^2\left(M_1+M_2\right) \qquad \text{(geometric units, } E \text{ in units of length)} \tag{22}$$

B. Verify that in both of these equations E has the unit of length.

C. Suppose you are given the value of E in meters. Show how you would convert this value first to kilograms and then to joules.

QUERY 11. Rate of change of radius

Derive an expression for the rate at which the radius changes as a result of this energy loss. Show that the result is:

$$\frac{dr}{dt} = -\frac{64}{5r^3}M_1M_2\left(M_1+M_2\right) \qquad \text{(Newton: circular orbits)} \tag{23}$$

---

8 BINARY PULSAR PSR1913+16

Proof of gravity waves?

Hulse and Taylor discover binary.

On July 2, 1974 Russell A. Hulse was carrying out observations at the world's largest radio telescope at Arecibo, Puerto Rico. Hulse—a graduate student working under the direction of Joseph H. Taylor, then at the University of Massachusetts, Amherst—detected signals from a pulsar later named PSR1913+16. (PSR stands for "pulsar" and the numbers denote its celestial coordinates.) Here is an account of the discovery, excerpted from the Nobel Foundation website (which also has wonderful illustrations)
`http://www.nobel.se/physics/laureates/1993/illpres/discovery.html`
(Copyright ©2001 The Nobel Foundation)

> **THE DISCOVERY OF THE BINARY PULSAR**
> During 1974 Joseph Taylor and Russell Hulse were searching for new pulsars with the Arecibo telescope. They discovered 40, one of which was to be very important.
>
> When Hulse was observing the new pulsar, which has been named PSR1913+16, he found that the pulses arrived sometimes more often and sometimes less. The simplest interpretation was that the pulsar was orbiting another star very closely and at high velocity: Here one "pulsar year" is only about eight hours.
>
> By observing the shift in the pulses, Hulse and Taylor found that the stars were equally heavy, each weighing about 1.4 times as much as the Sun. Since they were not visible on any photographs either, it was concluded that the other body, somewhat unexpectedly, was also a neutron star. Seen from Earth, however, it does not show up as a pulsar.
> . . . .
>
> **MEASURING gravity waves**
> Since the two neutron stars in PSR1913+16 are moving so fast and close together they should, according to General Relativity, emit large amounts of gravity waves. This makes them lose energy: Their orbits will therefore shrink and their orbiting period will shorten.
>
> **Indirect evidence:** The binary pulsar has been observed continuously since its discovery, and the orbiting period has in fact decreased. Agreement with the prediction of General Relativity is better than 1/2%. This is considered to prove that gravity waves really exists. In turn, this result is currently one of our strongest supports for the validity of the General Theory of Relativity.

Data from binary pulsar

The signal from the pulsar constituted a very stable clock, stable to 10 significant figures. As a result, Hulse and Taylor were able to use general relativity to analyze the motion of the system in detail, verifying many general relativity predictions, some of which allowed them to determine the individual orbiting masses M_1 and M_2 (given below), which Newtonian mechanics does

not reveal. Their results show that the binary system PSR1913+16 has the following parameters:

$$
\begin{aligned}
M_1 &= (1.442 \pm 0.003) M_{\text{Sun}} && \text{(pulsar)} \\
M_2 &= (1.386 \pm 0.003) M_{\text{Sun}} && \text{(companion)} \\
a &= 2.3418 \pm 0.0001 \quad \text{light seconds} && \text{(Semi-major axis of both)} \\
e &= 0.617127 \pm 0.000003 && \text{(Eccentricity of both)}
\end{aligned}
\tag{24}
$$

Orbital period, ≈ 7.75 hours

Rate of advance of the periastron ≈ 4.2 degrees per Earth-year

Distance from Earth ≈ 7 kiloparsecs or about 20 000 light years. (This distance is quite uncertain.)

Each neutron star follows its own elliptical path about the center of mass. The semi-major axis of the elliptical orbit for a neutron star—label it a—is half of the major axis, the longest distance from one side of its orbit to the other. The semi-minor axis—label it b—is half of the minor axis. Then the eccentricity $e \equiv (a^2 - b^2)^{1/2}/a$. The word **periastron** refers to the point of closest approach of these "astron"omical objects (just as the word *perihelion* refers to the point of closest approach of an orbiting object to our Sun: Greek, "Helios"). Note how large the rate of this periastron advance is compared with 43 arcseconds of advance of the perihelion of the planet Mercury *per Earth-century*.

Meaning of *periastron*.

The non-zero eccentricity in equation (24) tells us that the neutron stars in PSR1913+16 are *not* in circular orbits. General relativity predicts that when a binary system has non-circular orbits it will radiate gravity waves at a greater rate than when the orbits are circular. Nevertheless, in the following QUERIES we assume for simplicity that the orbits are effectively circular, as in Figure 7. That is, we assume a binary system in which each companion is in a circular orbit with constant radial separation r equal to the major axis, twice the value of the semi-major axis given in (24). This is equivalent to setting to zero the eccentricity of each neutron star orbit.

We assume circular orbits.

QUERY 12. Shrinkage of r per orbit

For a single orbit, the separation r between the orbiting neutron stars (assumed to be in circular orbits) does not change much, but it does change a little. For one orbit, what is the approximate value of the change in this separation r? Express your answer in millimeters. (*Hint:* No integration is needed for an approximate calculation of this incremental change.)

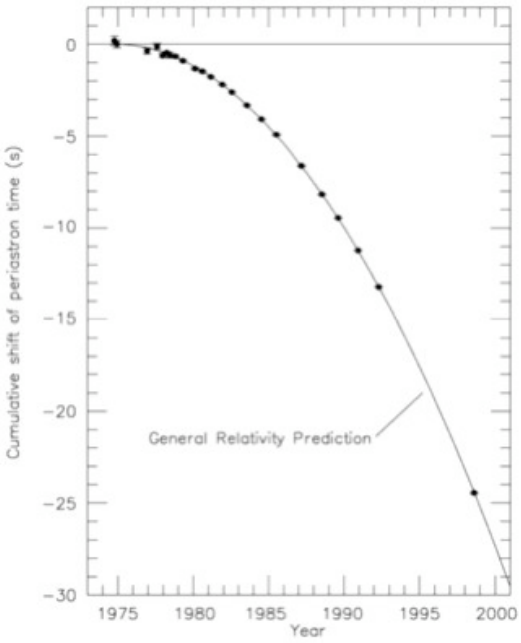


FIGURE 8 Decrease in the period in seconds (vertical axis) over the years 1975 to 1998 (horizontal axis) of binary system PSR1913+16. Agreement with the prediction of general relativity, assuming the change is due to emission of gravity waves, is now within 0.3 percent. This agreement appears to eliminate any other possible explanation for the change in orbits. From a paper (and Copyright ©2000) by J.H. Taylor and J. M. Weisberg.

QUERY 13. Energy radiated by idealized binary PSR1913+16

A. What is the power currently being radiated in gravity waves? Express your answer as a unitless measure (energy in meters divided by time in meters) and also in watts (joules per second).

B. Use equation (19) or (21) to calculate how much total energy in joules will be radiated in gravity waves from the present year to the future time when the two companions are separated by $r = 20$ kilometers (approximately the sum of their radii)? This total energy corresponds to how many kilograms of mass converted entirely to energy?

C. How long a time in years will it be before the two neutrons stars in PSR1913+16 are separated by only $r = 20$ kilometers, so that coalescence is imminent? (Only in the last millisecond or so before coalescence does the Newtonian description become completely useless.)

ADD QUERY ABOUT RADIATION RATE OF SUN-MERCURY BINARY SYSTEM AND LENGTH OF TIME TO COALESCE DUE TO GRAV RADIATION.

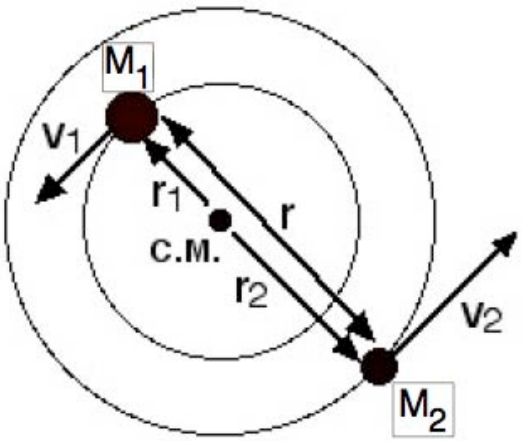


FIGURE 9 Figure 7 augmented to show the center of mass and orbit radii of individual components of PSR1913+16.

9 ■ GRAVITY WAVE AT EARTH DUE TO DISTANT BINARY SYSTEM

How far away from a binary system can we detect gravity waves?

Can LIGO on Earth's surface detect the gravity waves emitted by the distant binary system PSR1913+16 (idealized as one in which the neutron stars move in circular orbits as shown in Figure 7)? To answer this question we need to calculate the magnitude of h in the metric of equation (1).

Gravity waveform . . .

Here is the third and final result of general relativity quoted without proof in this chapter. The function $h(z,t)$ is given by the equation (in conventional units)

$$h(z,t) = -\frac{4G^2 M_1 M_2}{c^4 r z} \cos\left[\frac{2\pi f(z-ct)}{c}\right] \qquad \text{(conventional units)} \quad (25)$$

where f is the frequency of the binary orbit, r is the (constant!) distance between orbiters in Figures 7 and 9, and z is the distance from source to detector. Convert (25) to geometric units by setting $G = c = 1$. Note that $h(z,t)$ is a function of z and t.

Figure 10 schematically displays the notation of equation (25), along with relative orientations and relative magnitudes assumed in the equation. This equation makes the Newtonian assumptions that (a) the two stars are separated by a distance r much larger than their Schwarzschild radii and (b) they move at nonrelativistic speeds. Additional assumptions are:

(c) The distance z between the binary system and Earth is very much greater than a wavelength of the gravity wave. This assumption assures that the radiation at Earth constitutes the so-called "far radiation field" where it assumes the form of a plane wave given in equation (4).

(d) The binary stars are orbiting in the xy plane, so that from Earth the orbits would appear as circles if we could see them (which we cannot, because

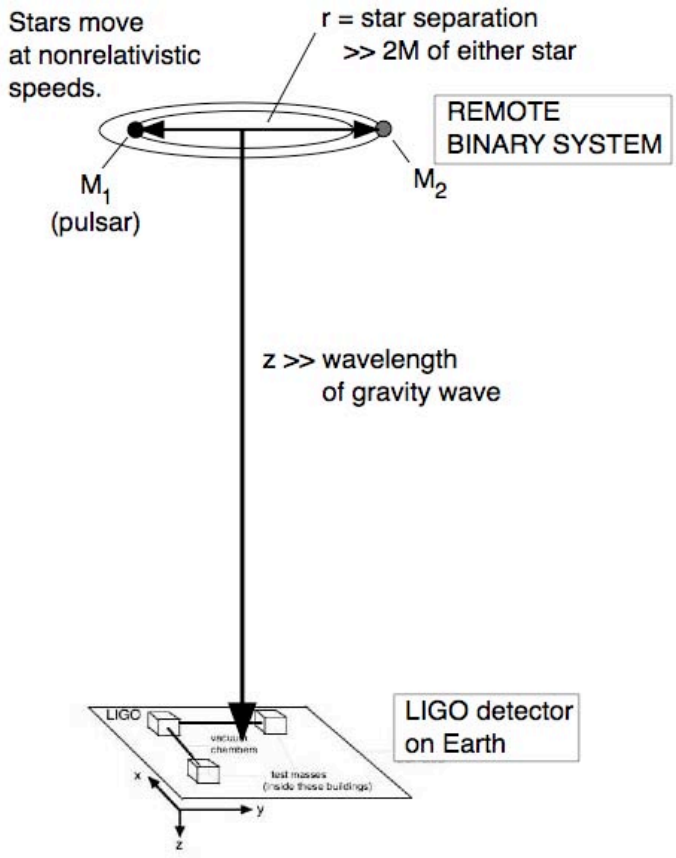


FIGURE 10 Schematic diagram, *not to scale*, showing notation and relative magnitudes for equation (25). The binary system and the LIGO detector lie in parallel planes.[Illustrator: See note in caption to Figure 5.]

they are too far away). Unfortunately this assumption is not true of the plane of the orbit of binary PSR1913+16, as we know from Doppler shifts of signals from the orbiting pulsar.

. . . for one case

Equation (25) describes only one linear polarization at Earth, the one generated by metric (1) and shown in Figure 3. The orthogonal polarization shown in Figure 4 is also transverse and equally strong, with components proportional to $(1 \pm h)$. The formula for the magnitude of h in that orthogonally polarized wave is identical to (25) with a sine function replacing the cosine function. We have not displayed the metric for that orthogonal polarization.

Detection requirements

In order for LIGO to detect a gravity wave, two conditions must be met: (a) the amplitude h of the gravity wave must be sufficiently large, and (b) the frequency of the wave must be in the range in which LIGO is most sensitive

(100 to 400 hertz). QUERY 13 deals with the amplitude of the wave. The frequency of gravity waves, discussed in QUERY 14, contains a surprise.

QUERY 14. Amplitude of gravity wave from PSR1913+16 at Earth

A. Use (25) to calculate the maximum amplitude of h at Earth due to the radiation from the "idealized circular-orbit" binary system PSR1913+16. Consider this amplitude to be positive.

B. Can either Initial LIGO or Advanced LIGO detect the gravity waves whose amplitude is given in part A?

C. What is the maximum amplitude of h at Earth just before coalescence of PSR1913+16, when the neutron stars are separated by a distance $r = 20$ kilometers (but with orbits still described approximately by Newtonian mechanics)?

QUERY 15. Frequency of gravity waves emitted from PSR1913+16

A. In order for either Initial LIGO or Advanced LIGO to detect the gravity waves whose amplitude is given in Query 13, the frequency of the gravity wave must be in the range 100 to 400 hertz. In Figure 9 the point C. M. is the stationary center of mass of the pulsar system. Using the symbols in Figure 9, fill in the steps to complete the following derivation.

$$\frac{v_1^2}{r_1} = \frac{GM_1}{r_1^2} \qquad \text{(for } M_1\text{, Newton, conventional units)} \tag{26}$$

$$\frac{v_2^2}{r_1} = \frac{GM_2}{r_2^2} \qquad \text{(for } M_2\text{, Newton, conventional units)} \tag{27}$$

$$M_1 r_1 = M_2 r_2 \qquad \text{(center-of-mass condition)} \tag{28}$$

$$f_{\text{orbit}} \equiv \frac{1}{T_{\text{orbit}}} = \frac{v_1}{2\pi r_1} = \frac{v_2}{2\pi r_2} \qquad \text{(common orbital frequency)} \tag{29}$$

where f_{orbit} and T_{orbit} are the frequency and period of the orbit, respectively. From these equations, show that for $r \equiv r_1 + r_2$ the frequency of the orbit is

$$f_{\text{orbit}} = \frac{1}{2\pi}\left[\frac{G\left(M_1 + M_2\right)}{r^3}\right]^{1/2} \tag{30}$$

B. Here is a surprise: The frequency f of the gravity wave generated by this binary pair and appearing in (25) is twice the orbital frequency.

$$f_{\text{gravity wave}} = 2 f_{\text{orbit}} \tag{31}$$

Why this doubling? Essentially it is because gravity waves are waves of tides. Just as there are two high tides and two low tides per day caused by the moon's gravity acting on the Earth, there are two peaks and two troughs of gravity waves generated per binary orbit.

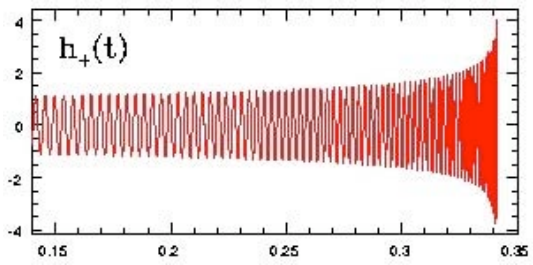


FIGURE 11
The "chirp" of one polarization of gravity wave as the two components of a binary system coalesce. (The plot for the other polarization is indistinguishable from this one.)

C. Approximate the average of the component masses in (24) by the value $M = 1.4M_{\rm Sun}$. Find the distance r between the binary stars when the orbital frequency is 75 hertz, so that the frequency of the gravity wave is 150 hertz. [ANS: Approximately 100 km.]

D. Using results quoted earlier in this chapter, estimate the time for the binary system to decay from the current radial separation to the radial separation calculated in part C.
ANS: $t_2 - t_1 = 5(r_2^4 - r_1^4)/(256M^3)$, everything in unit meter.

"Chirp" at coalescence

Newtonian mechanics predicts the motion of the binary system surprisingly accurately until the two components touch, a few milliseconds before they coalescence. As this happens, the gravity wave sweeps upward in both frequency and amplitude in what is called a **chirp**. Figure 11 is a predicted wave form for such a chirp.

To hear an audio simulation of the chirp, visit one of the following websites:
`http://www.lsc-group.phys.uwm.edu/~patrick/work/talks/itp/chirp.002.au`
`http://www.lsc-group.phys.uwm.edu/~patrick/work/talks/itp/chirp.002.wav`

Detection of such a waveform sweeping through the frequencies for which LIGO is sensitive would be a "smoking gun" for the coalescence of a binary source. Although LIGO cannot detect emission from PSR1913+16, we expect that many other binary systems are close to provide a detectable signal for Advanced LIGO.

10 ■ REFERENCES

First initial quote: Norbert Wiener, *I am a Mathematician*, MIT Press, Cambridge, MA, 1956, page 109

Second initial quote: Arthur Eddington, *Stars and Atoms* (1928), Lecture 1

LIGO sensitivity, Figure 2, at
`http://www.ligo.caltech.edu/advLIGO/scripts/summary.shtml`

Chirp wave shape, Figure 11, at
`http://www.lsc-group.phys.uwm.edu/~patrick/work/talks/itp/itp0008.gif`
Websites for updates:
`www.ligo.org`
`www.ligo.caltech.edu`

It took Engineer and Scientist Maurice Cotterell just 6-weeks to discover how electricity and magnetism work together to produce the force of Gravity, and 3 years to find a way of explaining it. Gravity, along with 11 other 'great mysteries of Physics', is explained in his book *FUTURESCIENCE—forbidden science of the 21st-century*. The third Hardback edition is now available, exclusively, from www.MauriceCotterell.com

How Gravity Works

Maurice Cotterell B.A.(Hons) MCMI I.Eng MIET

© 2007 – 2012 Maurice Cotterell SunspotUK@aol.com
www.MauriceCotterell.com

Abstract

How Gravity Works-a Unified field theory

This *explanation of the gravitational mechanism* provides a causal mechanism that reconciles physical phenomena with Newton's equation for the force of Gravity. It shows how electricity and magnetism work together to produce the force of Gravity and in so doing represents the first Unified Field Theory. It explains the reason why the gravitational force is proportional to the masses of the attracting bodies and why the force decreases in accordance with a square law scale. The Gravitational Constant is defined. The reason why objects accelerate to Earth in accordance with a square law scale is explained. The architecture of the atom is explained—the reason why electrons orbit the atom in up to 8 shells/sub-shells; why the shells contain the number of electrons that they do; and why the shells are offset by at least 45°, are all explained.

Conclusion; The hydrogen atom [and helium atom] generate *helically polarised electromagnetic radiation* (gravity radiation) from polar regions that bombards neighbouring atoms drawing them towards the source of the radiation. Gravity radiation then causes the nucleus of the neighbouring atom to spin axially (the 'motor effect') and, at the same time, synchronises the spin of the electrons in both atoms. The neighbouring atom then, in turn, generates helically polarised electromagnetic energy (the 'generator effect'). Both atoms spin axially in the same direction. Hence, the gravitational forces from both atoms pull in the same direction and the forces are additive. The gravitational Constant G is shown to be the instantaneous alternating magnetic force between any two electron-magnets in neighbouring atoms and, hence, the gravitational force F is proportional to G x m1 x M2 (where m1 and M2 represent the electron count of neighbouring atoms).

electricity gravity and magnetism

three sides of the same coin

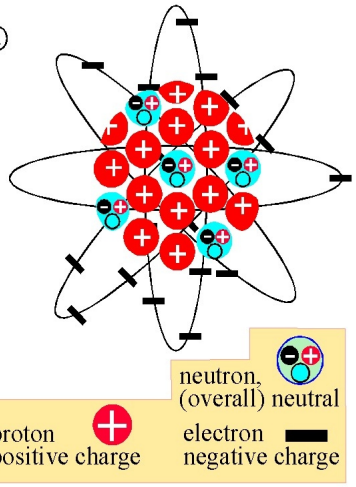
How Gravity Works

Orthodox Science does not understand the force that causes atoms to attract each other—because they do not understand the nature of the force, or the nature of the atom.

The accepted Model of the Atom

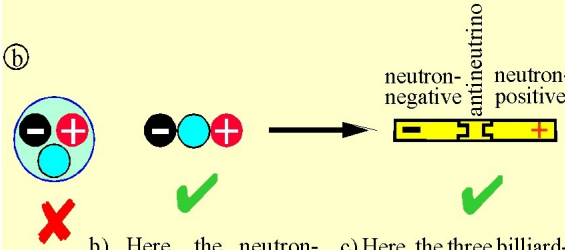
The accepted model of the atom (*right*) recognises that complex atoms are made of electric positive charges, protons (shown in red), electric negative charges (black), and neutrons (blue). The neutrons have no overall electrical charge but are known to be made of ½ of a proton, ½ of an electron and a neutral particle named the antineutrino. But this model *cannot* be correct because it raises three crucial questions; 1. Why—given that positive repels positive—do the positive charges simply not just spring apart? 2. Given that positive attracts negative, what stops the orbiting negative charges from being sucked-in to the positive charges in the nucleus? and, 3. What stops the ½ proton and the ½ electron inside the neutron from being drawn together and annihilating each other?

ⓐ



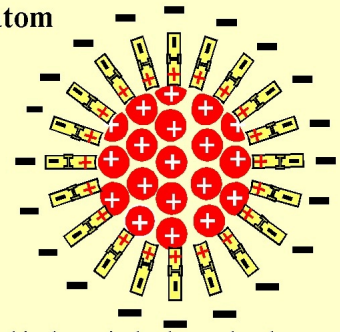
So what makes the complex atom stable (!)?—the new atom

ⓑ



b) Here, the neutron-negative particle is shown separated from the neutron-positive particle by the neutron-neutral particle (the antineutrino) using the billiard-ball-style convention.

c) Here, the three billiard-ball-style particles have been replaced with a tiny positive battery-like symbol and a tiny negative battery-like symbol that are kept apart by the neutral anti-neutrino. This new representation will from now on, herein, be referred to as the 'spiked-neutron'.



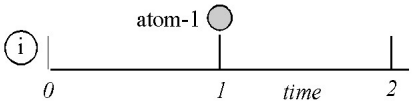
d) In this theoretical scheme the electrons and protons are kept apart by the spiked-neutrons. The positive pole of each spiked-neutron is embedded in the mass of protons contained in the nucleus. The neutron-positive poles push against the positive protons preventing them from springing apart. The neutron-negative poles repel the electrons preventing them from getting closer to the nucleus. The spiked-neutrons act like springs, forcing-out the electrons and forcing-in the protons.

Supporting evidence: Researchers, at the Hahn-Meitner Institute, Berlin, announced that 'neutrons behave like compass needles' (Science Daily, 31.3.2008)

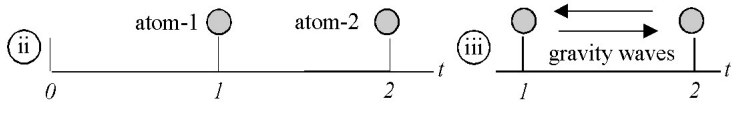
What's the nature of the force?

Isaac Newton defined Gravity as 'the force between 2 bodies' [atoms]—although he never understood what caused it, how it worked, or where it came from. c) However, when we recognise that the Universe began with just 1 atom, it becomes clear that *the force of gravity must travel*. This means that each atom must give-off some form of energy. The only energy known to travel through space is electromagnetic energy, radio waves.

ⓒ Before the creation of the Universe, at time zero, no physical atoms existed. Then, a moment later, at time interval *1*, the first atom, atom number 1, must have appeared.



i) At *time-1* gravity could not exist because only 1 atom exists, and Newtonian gravity is the force between 'two' atoms.



ii) At *time-2* atom-2 appears and gravity exists. Thus, a gravitational force must *travel* from atom-1 to atom-2, and from atom-2 to atom-1, between time intervals *1* and *2*. iii) This means that gravity must *travel* to establish itself between the atoms.

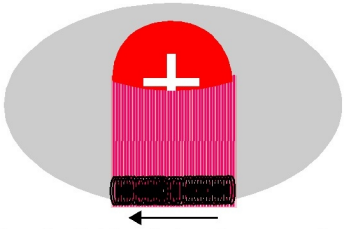
figure 1.

But how could an atom give off radio waves?

In 1831, Michael Faraday demonstrated that when a coil is swiped through a magnetic field, an electric field [a voltage] is produced across the coil. The corollary is also true; whenever a coil is swiped through an electric field [a voltage] a magnetic field is produced [induced] into the coil. In order to make the atom stable, it was firstly necessary to change the shape of the neutron. In order to enable a gravitational mechanism, we must now change the shape of the electron—into a coil-shape.

ⓐ

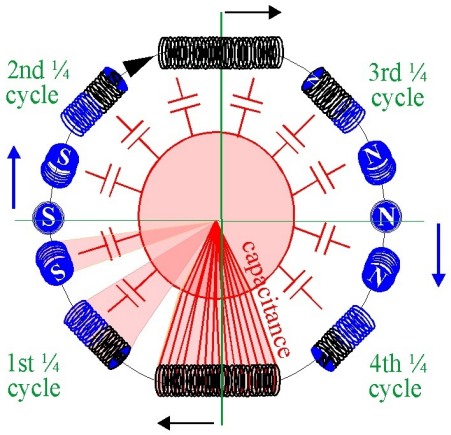
here, hydrogen, the simplest of atoms, is used to explain the principles involved



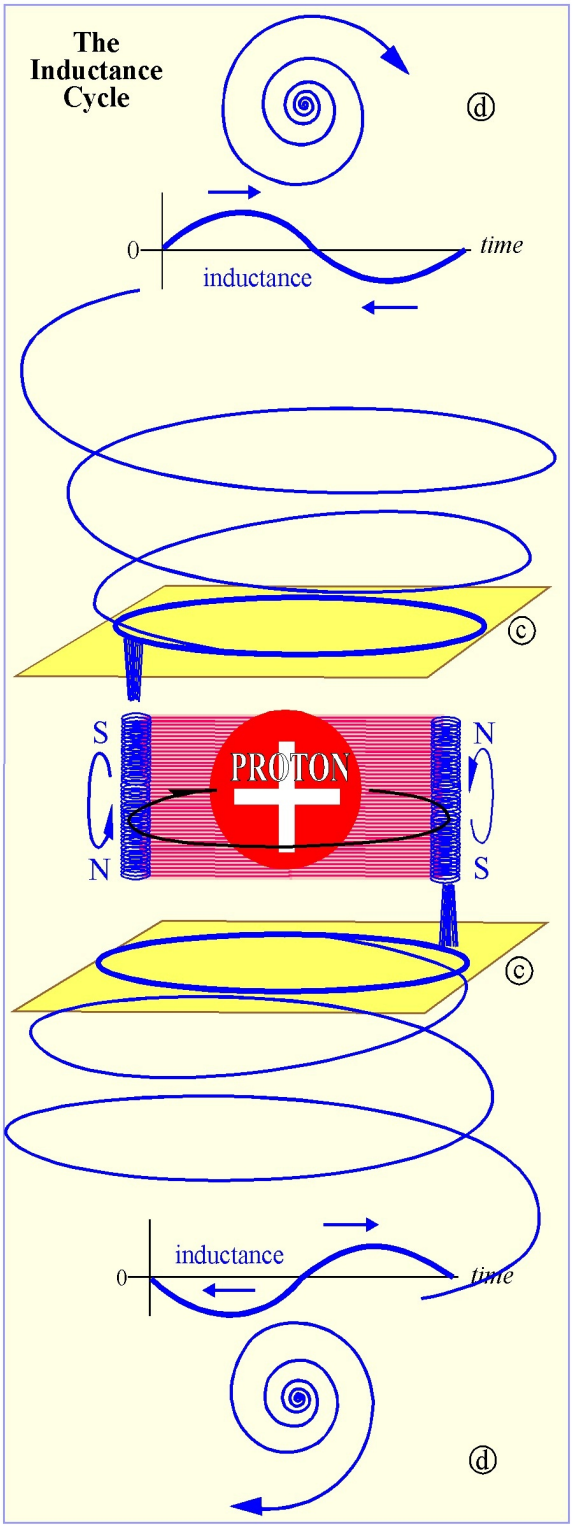
hydrogen atom

a) An electric field subsists between the orbiting negative electron and the positive proton. The amount of electric field contained in the space between the particles is known as its 'Capacitance'.

ⓑ



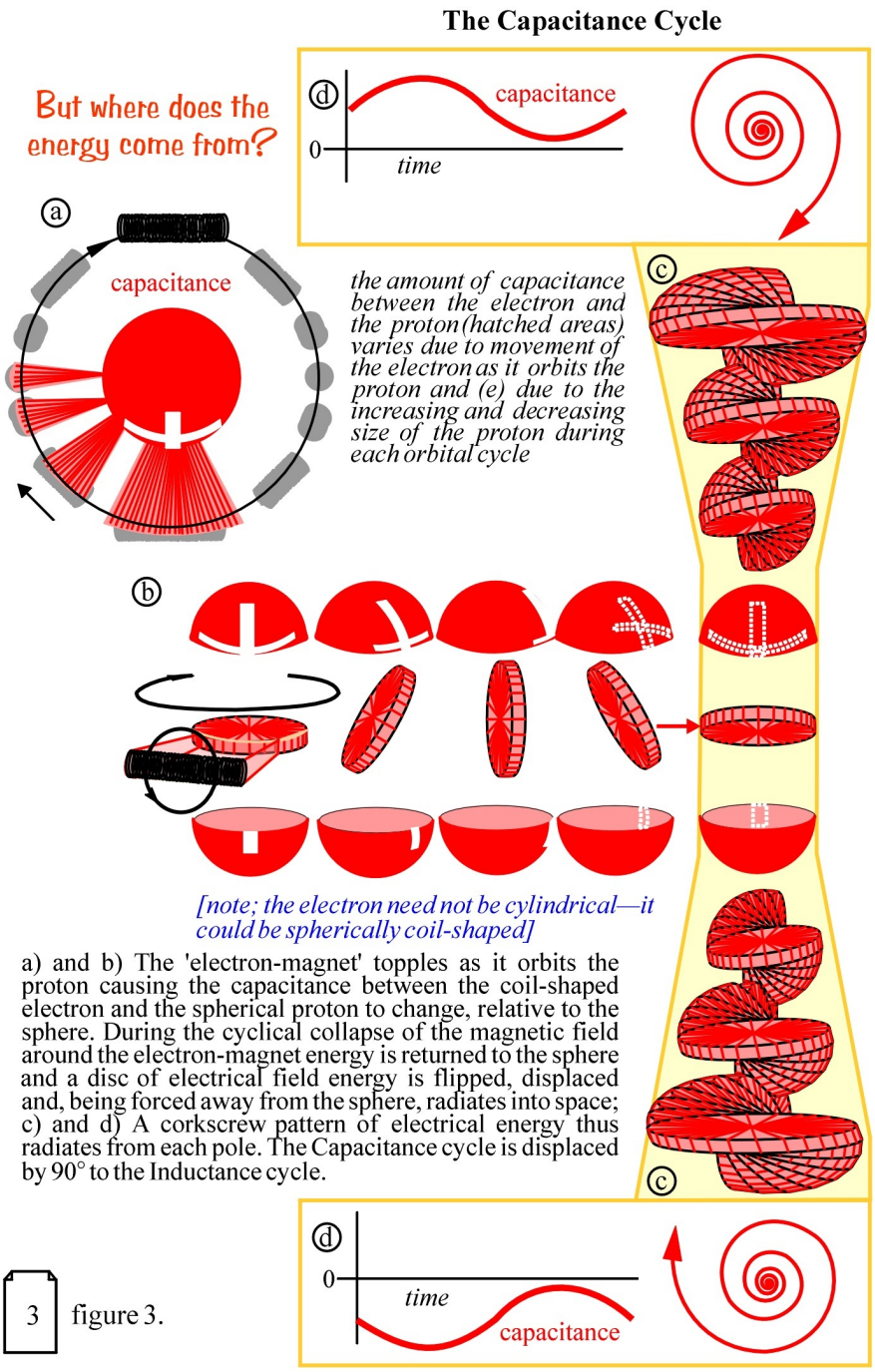
b) As the electron slices through the electric field, a magnetic field is induced into the electron. Because an electric field and an associated magnetic field are always displaced by 90°, the induced magnetic field parries against the associated electric field causing the elevation of the electron to twist to a more vertical position. Once the electron becomes vertical the coils of the electron no longer slice through the electric field. But the electron continues to spin through its own inertia and as it twists to a more horizontal elevation it once again begins to slice through the electric field. More magnetism is again induced into the electron during the second quarter cycle—at the same time as the magnetic field uploaded during the first quarter cycle begins to collapse. The magnetic field around the electron [although it should now be referred to as an 'electron-magnet'] is thus pushed-out into space; it 'radiates' into space. The electron-magnet changes direction as it slices through the electric field during quadrants 3 and 4, thus the associated magnetic field around the electron-magnet reverses, relative to the first and second quarter cycles.



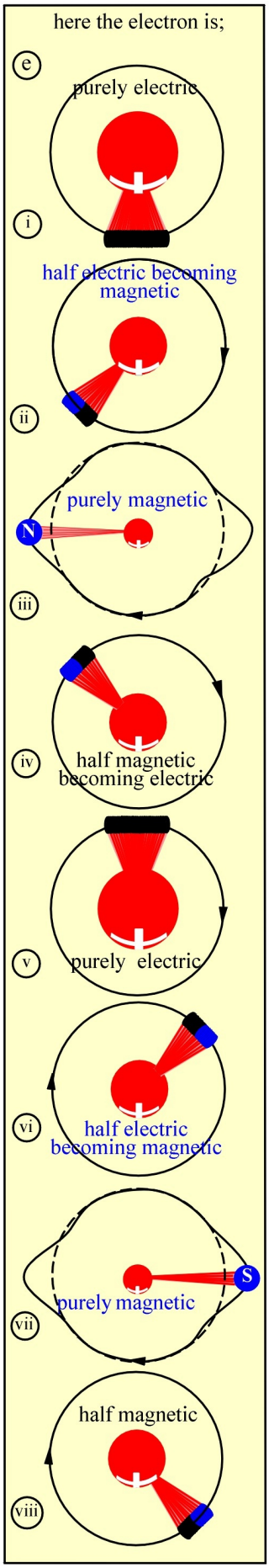
The electromagnetic interaction between the particles causes the electron-magnet to spin once [through 360°] like the opposing blades of a windmill, during each orbit of the proton. c) and d) It can be seen that the electron-magnet radiates helically polarised (corkscrew-style) magnetic energy from the equatorial region to the polar regions. c) Radiation from the northern sector is displaced by 180° from radiation from the southern sector.

figure 2.

The Capacitance Cycle



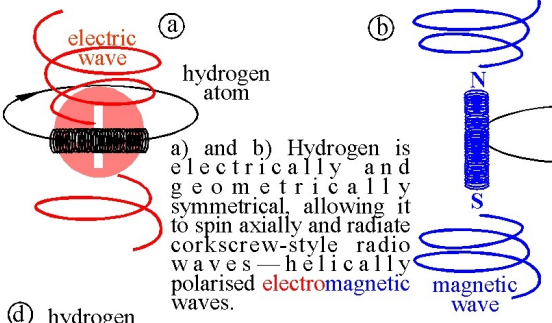
3 figure 3.



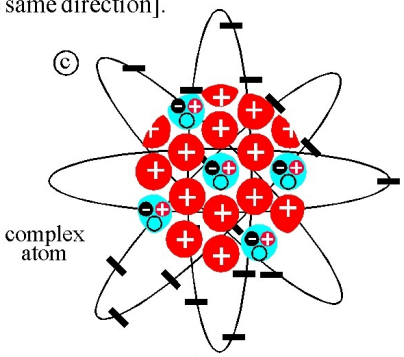
e) As the electron-magnet sucks-in energy from the proton the proton rapidly compresses (shrinks) and cools, as the induced magnetism rises in the electron-magnet. When the electron-magnet becomes vertical and stops sucking-in energy, the proton stops compressing. As soon as that happens the super-cold miniature proton sucks-in ambient heat (above –273.15° C) causing the proton to heat-up and expand rapidly. The surface area of the proton thus gets larger and smaller as the electron-magnet orbits the proton, so the amount of electric field (capacitance) between the particles must vary in a cyclical way [because Capacitance is proportional to the surface area between the particles]. Below –273.15° C the electron cannot access sources of heat to sustain electromagnetic oscillation and the electron, starved of input energy, ceases to orbit the proton. Oscillations cease. The atom ceases radiating corkscrew-style electromagnetic energy and the atom ceases to attract other atoms. e(i) to e(viii). Whenever the electron absorbs energy, it drains the electric field of some energy and the tension between the two particles diminishes, allowing the electron to increase the size of its orbit. Whenever energy is returned, the renewed stronger field pulls the electron back to a position inside the original orbit, momentarily, before returning to its original orbit, e(iii) and e(vii).

How do corkscrew-style electromagnetic radio-waves affect other atoms?

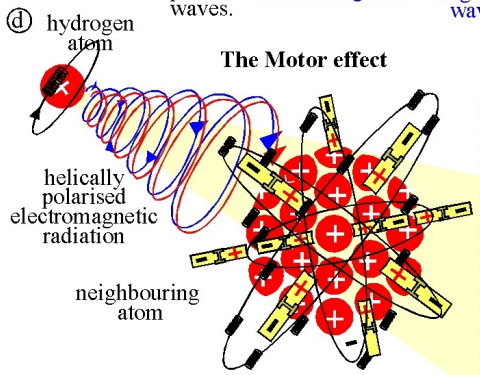
Gravity requires that every atom attracts every other atom in every different direction—and 93% of atoms in the Universe are hydrogen atoms. Hence, for a gravitational mechanism to be enabled by corkscrew-style radiating waves three conditions must be met: 1. All hydrogen atoms throughout the Universe must be randomly orientated. 2. Corkscrew-style radio-waves from a hydrogen atom must *not* interfere with corkscrew-style radio-waves from other hydrogen atoms, and 3. If corkscrew-style radio-waves from the hydrogen atom are the prime-mover in the gravitational mechanism then other [non-hydrogen] atoms *must* be affected [in some way *switched-on*] by the corkscrew-style radio waves from the hydrogen atom [in order to satisfy Newton's observations that bodies in alignment, like the Sun and Moon pull in the same direction].



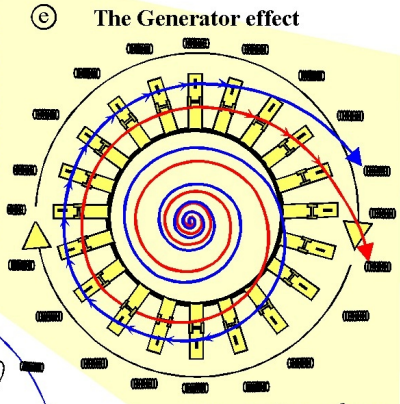
a) and b) Hydrogen is electrically and geometrically symmetrical, allowing it to spin axially and radiate corkscrew-style radio waves—helically polarised electromagnetic waves.



c) More complex atoms are not electrically-geometrically-symmetrical—electrons orbit in shells off-set by at least 45°. Electrons thus bunch together as they orbit the atom. The mass distribution of the atom therefore changes, causing the atom to 'tumble', not to spin autonomously. More complex atoms therefore do not *normally* spin or radiate radio waves.

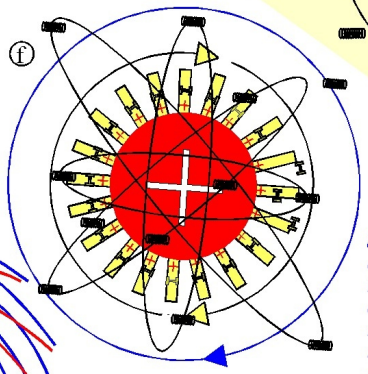


d) and e) However, when helically polarised electromagnetic radiation from the hydrogen atom bombards more complex atoms it acts on the negative charges of the neighbouring atom causing the nucleus and the electron-cage to spin axially. The mass of the electron is twice that of the neutron-negative charge, hence the torque on the electron-cage is twice that on the neutron-negative charges in the nucleus. Thus, the electron-cage spins faster than the nucleus.

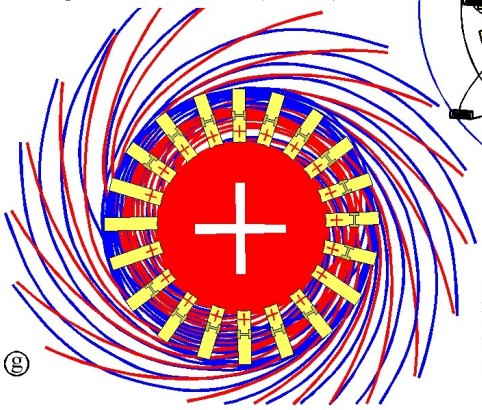


cross-section of new model with spiked-neutrons

f) Orbiting electron-magnets, in axially spinning atoms, generate helically polarised electromagnetic energy that radiates from the atom. The power output of the atomic generator is proportional to the differential speed between the nucleus and electron cage and also proportional to the number of orbiting electrons (atomic mass). Elements with more mass thus generate more electromagnetic energy which pulls them towards the Earth with more force, making them more difficult (heavier) to lift.



Supporting evidence: 'when electrons are bombarded by helically polarised radiation they align themselves—like compass needles—and follow the corkscrew spin of the waves' [Nature, 2009; 458 (7238):610DOI:10.1038.07871 (persistent spin helix)].

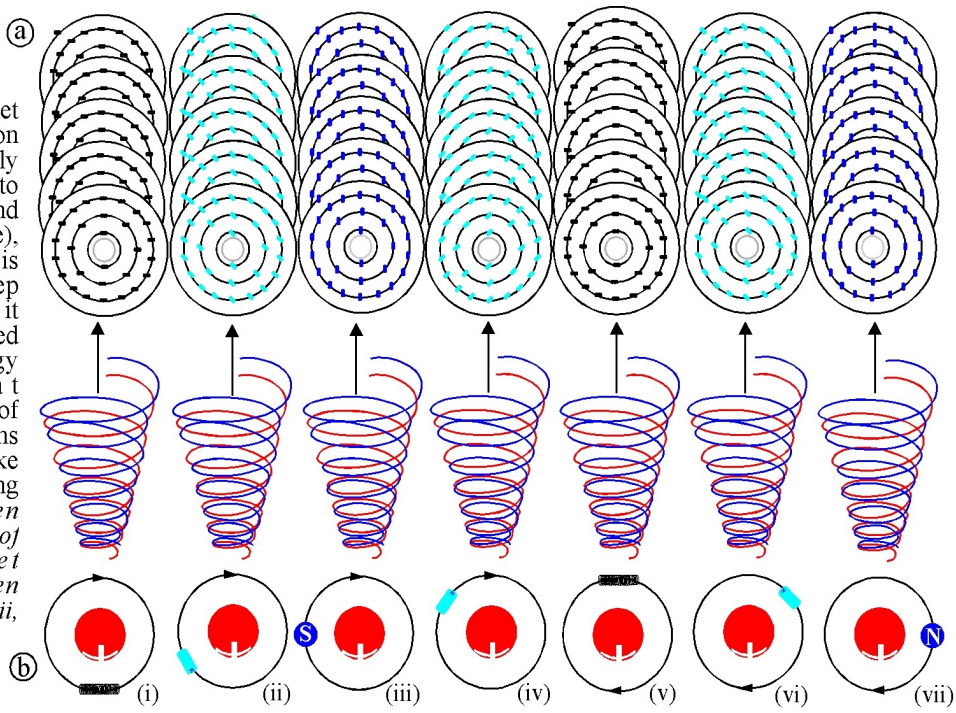


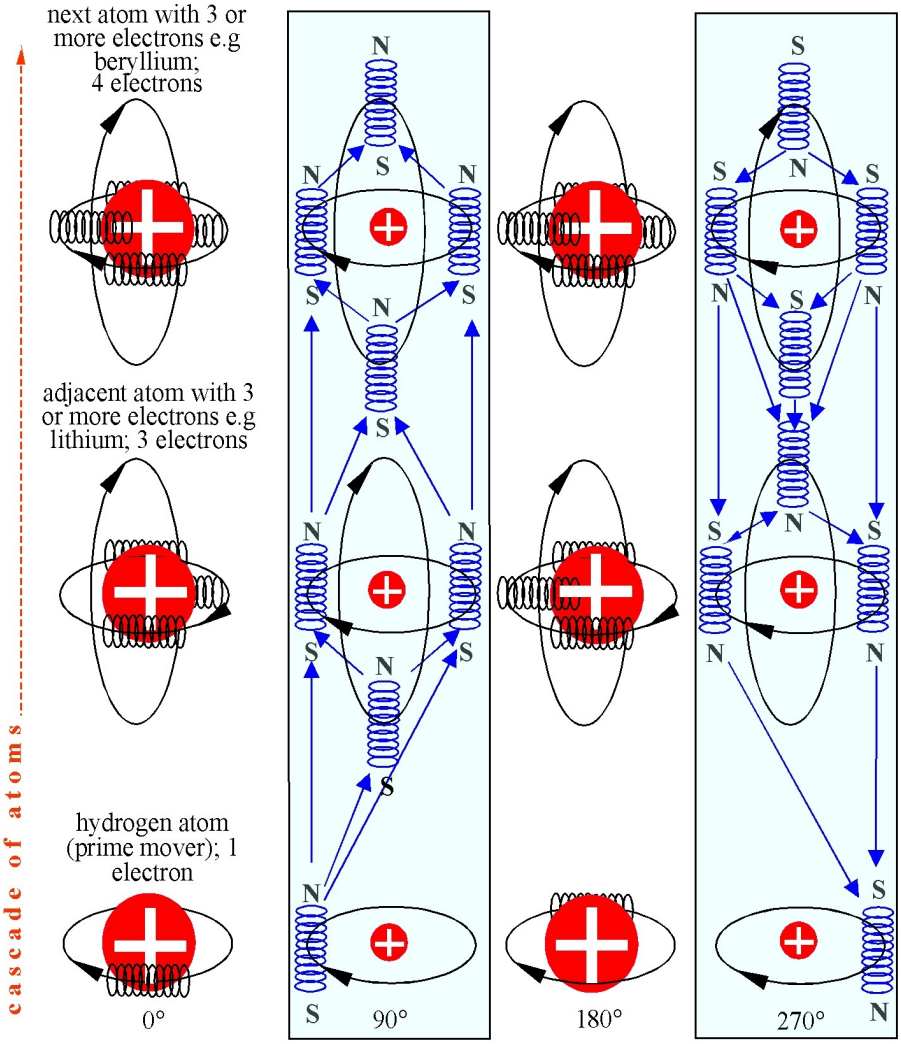
g) The neighbouring atom now radiates helically polarised electromagnetic radiation that bombards other neighbouring atoms in alignment.

figure 4.

Hydrogen synchro-spin

b) As the electron-magnet orbits the hydrogen proton it changes from a purely electric particle (black) to one that is half electric and half magnetic (light blue), and then to one that is purely magnetic (deep blue). At the same time it radiates helically polarised electromagnetic energy into space that synchronises the spin of electrons in nearby atoms causing them to spin—like windscreen wipers rotating through 360°. [*Seven consecutive snapshots of the electron-magnet orbiting the hydrogen proton are shown; time i, ii, iii, iv, v, vi and vii*].





no gravity when electron-magnets are at 0°

electron-magnetic force (gravity) when electron-magnets are at 90° (vertical)

no gravity when electron-magnets are at 180° (horizontal)

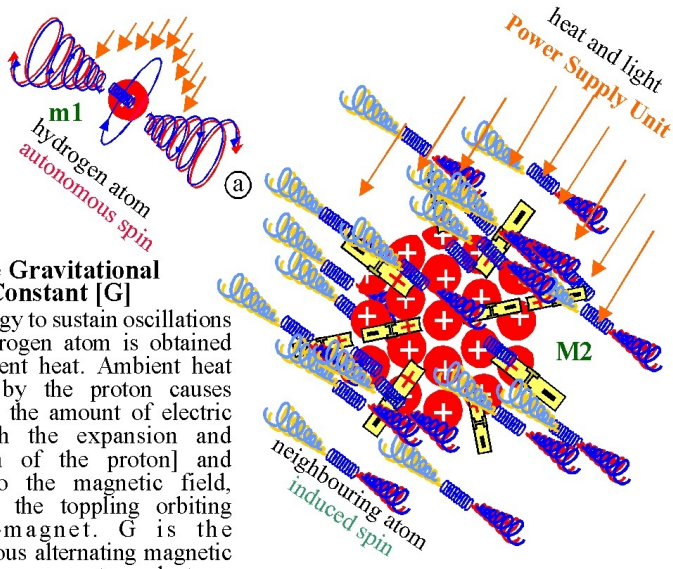
electron-magnetic force (gravity) when electron-magnets are at 270° (vertical)

How gravity works

c) A vertical cascade of atoms; helically polarised electromagnetic energy (not shown here) from the hydrogen atom bombards the lithium atom synchronising the spin of its electron-magnets with that of the hydrogen electron-magnet. The lithium electron-magnets then spin axially and radiate helically polarised electromagnetic radiation that bombards the beryllium atom, synchronising the spin of its electron-magnets with that of the hydrogen atom. The beryllium electron-magnets then spin axially and the beryllium electron-magnets radiate helically polarised electromagnetic energy etc. It can be seen that, in a cascade of atoms, helically polarised electromagnetic radiation from the hydrogen atom synchronises the spin of all electrons in atoms in alignment with the hydrogen atom. As a result the magnetic moments of orbiting electron-magnets in a cascade of atoms attract each other. This is the force of gravity. The electron-magnetic moments alternate during each orbital cycle of the hydrogen electron-magnet, hence gravity cannot be measured.

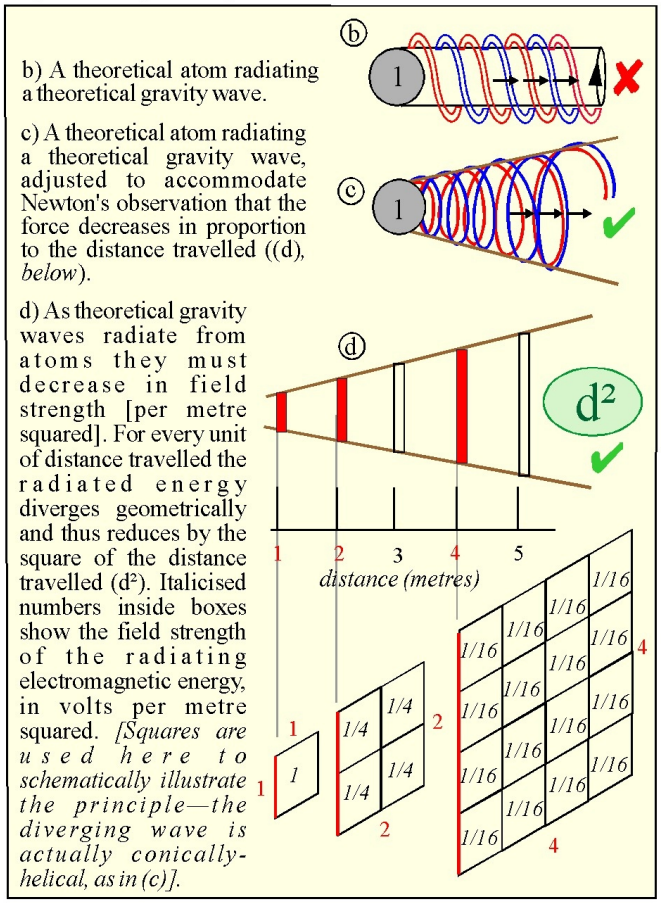
©

figure 5.



The Gravitational Constant [G]

a) The energy to sustain oscillations in the hydrogen atom is obtained from ambient heat. Ambient heat sucked-in by the proton causes changes in the amount of electric field [with the expansion and contraction of the proton] and changes to the magnetic field, caused by the toppling orbiting electron-magnet. G is the 'instantaneous alternating magnetic force between any two electron-magnets in neighbouring atoms'.



b) A theoretical atom radiating a theoretical gravity wave.

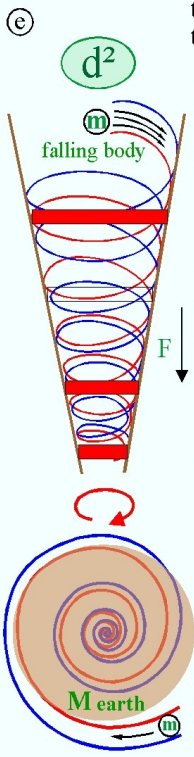
c) A theoretical atom radiating a theoretical gravity wave, adjusted to accommodate Newton's observation that the force decreases in proportion to the distance travelled ((d), *below*).

d) As theoretical gravity waves radiate from atoms they must decrease in field strength [per metre squared]. For every unit of distance travelled the radiated energy diverges geometrically and thus reduces by the square of the distance travelled (d^2). Italicised numbers inside boxes show the field strength of the radiating electromagnetic energy, in volts per metre squared. *[Squares are used here to schematically illustrate the principle—the diverging wave is actually conically-helical, as in (c)].*

Derivation of Newton's equation

The magnetic force between any two magnets is proportional to the magnetic force of one magnet multiplied by the magnetic force of the second magnet. The magnetic force between two atoms can therefore be calculated by multiplying the electron-magnetic force of one electron-magnet [G, the gravitational Constant] by the number of electron-magnets in atom 1 (a proportion of the mass of atom 1, m1) multiplied by the number of electron-magnets in atom 2 (a proportion of the mass of atom 2, M2); or as Newton said, the force (F) can be calculated by multiplying G [the magnetic force of 1 electron-magnet] multiplied by m1 x M2. (b) – (d) explain why the result must be divided by the distance (between the two atoms) squared.

Why objects accelerate to Earth in proportion to d^2



e) The frequency of the helically polarised EM radiation from the Earth remains constant. However, as m approaches M the spiraling EM radiation accelerates the differential rotation between each atomic nucleus and electron cage of which m is made. As a result the 'relative' atomic frequency increases and hence the output of the 'atomic generator' increases, increasing uniformly the attracting EM Force between m and M in accordance with a square-law scale as m proceeds along the EM spiral. M thus attracts m with square-law [d^2] uniform acceleration.

[m is shown spiralling towards M with the EM wave stationary but, in actuality, m approaches M in a straight line as the wave spirals across m]

Hence Isaac Newton's equation for the force of gravity; where the force is proportional to the masses (**m** & **M**) of the two attracting bodies and the strength of the force decreases inversely with the square of the distance between them (d^2).

$$F = GmM/d^2$$

G is Newton's gravitational Constant

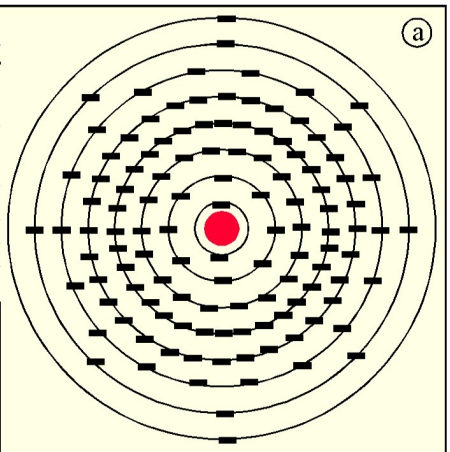
6.672×10^{-11} N m^2 kg^{-2}

—the instantaneous alternating magnetic force between any two electron-magnets in neighbouring atoms

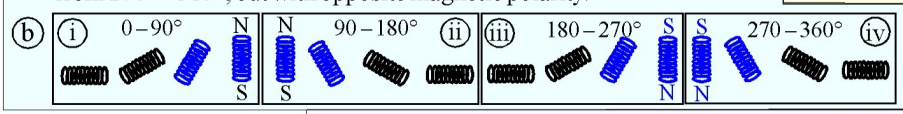
figure 6.

Proof that the electron is a coil-shaped electron-magnet—and electron-shell architecture explained

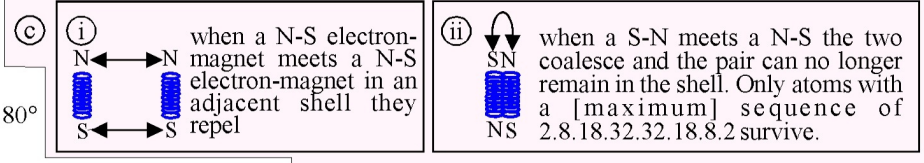
a) Schematic showing the theoretical maximum numbers of electrons filling the maximum number of available shells in an atom. The maximum number of electrons in the shell closest to the nucleus is 2. The maximum number of electrons in the next outermost shell/subshell is 8, followed by shells/subshells containing a maximum of 18, 32, 32, 18, 8, 2. It can be seen that, theoretically, the heaviest atom (illustrated) contains 120 electrons. The number of electrons orbiting an atom is usually balanced [but there are exceptions] with the same number of protons in the centre and (in our 'new atom') the same number of spiked-neutrons protruding from the centre (not shown). This also means that there are a maximum of only 120 different fundamental materials (elements). The heavier ones are more '*mass*ive' and hence said to contain more *mass* (electrons, protons and neutrons).

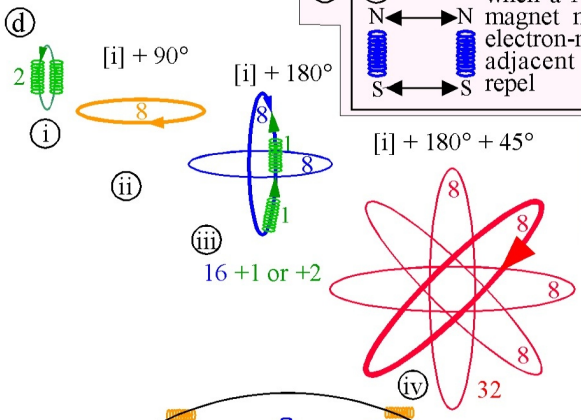


b(i – iv) illustrate the electric – magnetic tipping point of the orbiting electron-magnet [EM] at 45° interval: bi) From 45° – 90° (1/8 of the time) the magnetic field of the EM rises [and it is more magnetic than electric]. bii) From 90° – 135° it falls [but is still more magnetic than electric]. b(iii – iv) The same thing happens between 225°– 270°, and from 270° – 315°, but with opposite magnetic polarity.

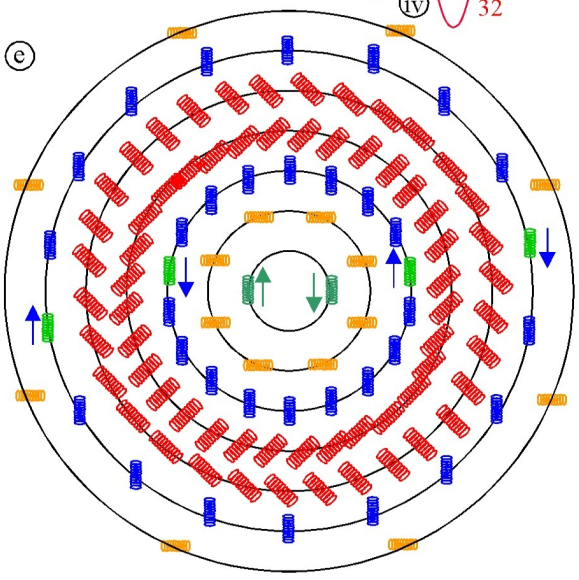


orthodox Science does not understand why the atom is structured the way it is—the reasons are given, below





c(i – ii) when a northsouth [N-S] EM meets a northsouth [N-S] EM the two repel each other, and when a northsouth [N-S] meets a southnorth [S-N] the two attract, stick together, and exit the host shell. Hence, as subatomic particles accrete into atoms, EMs inside shells/sub-shells must be separated by at least 45° to avoid annihilation and EMs in one shell must be separated by at least 45° from those in adjacent shells. Hence the constraint of up to 8 EMs per shell/sub-shell [8 x 45° = 360] and the requirement for different shell/sub-shell planar orientations.



schematic only—EMs are actually synchronised by gravity waves and [instead] the inclination of each shell/sub-shell is offset as shown in d(i – iv)

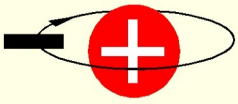
d) and e) As a general rule orbital shells [or sub-shells of equidistant radii] cannot sustain more then 8 EMs because of considerations set down in b(i – iv) and, to avoid magnetic conflict *between* shells/subshells, the plane of successive shells/subshells must be progressively offset by at least 45°. However, shells 3 and 6 can sustain up to 2 more EMs in the scheme proposed in (d) and (e) because the magnetic moments from the 2 EMs in shell 1 (being equal and opposite to 2 EMs in shell 3) cancel, and magnetic moments from 2 EMs in shell 3 cancel those of 2 EMs in shell 6, allowing those shells/sub-shells to sustain up to 18, rather than 16, EMs.

here, to further illustrate the relationship between magnetic moments, EMs in successive shells are shown reorientated; in shell 2 by 90°, in shell 3 by a further 90°, in shell 4, by a further 45°, in shell 5 by a further 90°, in shell 6 by a further 45° and in shell 7 by a further 90°. It can be seen that no magnetic conflicts occur in such a scheme and that the 2 EMs in shell 1 (green) influence the magnetic moments of those in shells 3 and 6, so that those shells can sustain up to 2 more EMs (green) than generally possible. This defined structure confirms that the electron must be coil-shaped and that it behaves as an electromagnetic particle

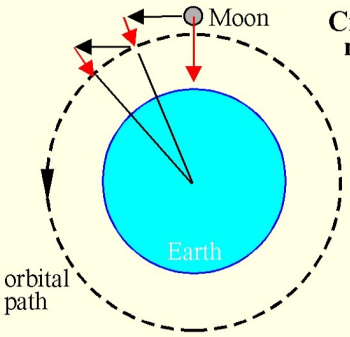
figure 7.

What makes the atom stable (II)?

ⓐ



a) Hydrogen, the simplest of atoms, contains only 1 proton that is orbited by 1 electron. The electron is attracted to the positive proton but does not crash into the atom—for the same reason that the Moon does not crash into the Earth (*right*)



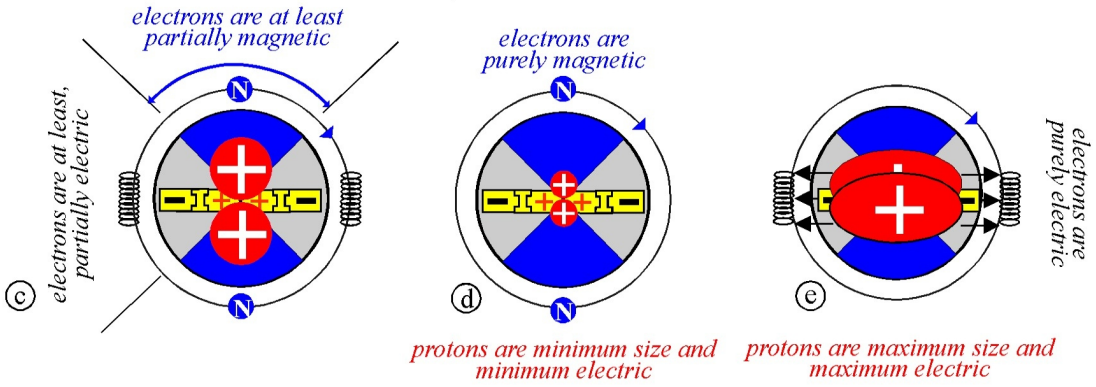
Circular orbital motion of the Moon

Newton reasoned that the Moon is continually falling towards the Earth with the same force and acceleration as a falling apple. At the same time, it is moving away from the Earth in its journey through space. The resulting motion captures the Moon in Earth's orbit (shown by the broken circular line).

Why atoms with spiked-neutrons are stable

note; helium, like hydrogen, is electrically and geometrically symmetrical and hence spins axially and autonomously and radiates helically polarised electromagnetic energy, however, helium is subordinate to hydrogen because hydrogen has no spiked-neutrons and is therefore unaffected by any radiation from helium

ⓑ



plan view of a helium atom with 2 electrons, two protons and 2 spiked-neutrons. The electron-magnets alternately possess an increasing and decreasing magnetic field that alternates with an increase and decrease of atomic capacitance. The two forces unite in self-sustained oscillation. The spiked-neutrons prevent the electric-electrons from crashing into the atom

Figure 3e) showed how the electron behaves as a permanent magnet at 90° and 270° from its starting position and how it does not possess a magnetic field at 0° and 180°.The electron is magnetic therefore during only half of the time it orbits the nucleus. We note further that, because a purely magnetic field and a purely electric field are displaced by 90°, a particle cannot be both purely electric and purely magnetic simultaneously, the two are mutually exclusive. It follows that if an electrically negative electron possesses magnetic qualities for half of the time it cannot be simultaneously electrical during that same part of the cycle. The electron, therefore, can only be considered as having electrical properties for half of the time. c) Shows the electric and magnetic possibilities, in plan view, of two electrons as they orbit the nucleus of a helium atom. d) When the electrons are each purely magnetic they cannot be attracted to an electric particle (the proton). Hence the magnetic electron is not attracted towards the nucleus during the quadrants shaded in blue. e) 90° later, the electron is maximum electric, but the charge on the nearby neutron-negative prevents them moving closer to the nucleus. Thus, the electrons cannot be sucked-into the nucleus. b) We also note that during the blue quadrants the protons, drained of energy, physically shrink, compress rapidly, release heat and cool. Hence the capacitance between electrons and protons reduces during the blue quadrants [because capacitance varies with distance between particles and the surface area of each particle]. In regard to the proton; during the blue quadrants, when the electron is maximum-magnetic, the positive electrical charge of each proton (and the force of repulsion between protons) is minimal, hence the protons have little propensity to move apart during the blue quadrants. e) As the electrons move into the grey quadrants the protons begin to suck-in ambient heat and rapidly expand. As the protons expand the distance between them and the electrons reduces and their surface area increases, increasing the amount of capacitance between them. It can thus be seen that when the electric particle is maximum-electric, the proton is also maximum-electric and attraction between the electrons and protons is maximum. The protons are thus stretched between the pair of electrons and have no propensity to spring apart; which explains why the maximum, and minimum, number of electrons in the first orbital shells of atoms containing spiked-neutrons must always be 2.

figure 8.

Reprinted from AMERICAN JOURNAL OF PHYSICS, Vol. 31, No. 3, 166–170, March, 1963
Printed in U. S. A.

Guidelines to Antigravity*

ROBERT L. FORWARD
Hughes Research Laboratories, Malibu, California
(Received 12 September 1962)

This paper emphasizes certain little known aspects of Einstein's general theory of relativity. Although these features are of minor theoretical importance, their understanding and use can lead to the generation and control of gravitational forces. Three distinctly different non-Newtonian gravitational forces are described. The research areas which might lead to methods for the control of gravitation are pointed out and guidelines for initial investigation into these areas are given.

INTRODUCTION

EINSTEIN'S general theory of relativity provides a number of ways to generate non-Newtonian gravitational forces. Theoretically, all of these forces could be used to counteract the gravitational field of the earth, thus acting as a form of antigravity. The three outlined here were probably known by Einstein before he published his paper on the principle of general relativity in 1916.[1] They were first specifically derived by Thirring[2] in 1918, and since then have been contained in nearly every text on general relativity.[3–5]

Although non-Newtonian forces are well known to the theorists in general relativity, they are little known to those outside the field, and it was for this reason that it was felt that a simplified discussion of these unusual gravitational effects would be of interest to the nonspecialist.

The equations of general relativity not only predict the usual radial Newtonian gravitational force behavior of a stationary mass on a stationary test body, but they also predict that a moving mass can create forces on a test body which are similar to the usual centrifugal and Coriolis forces, although much smaller. In addition, when the general relativity field equations are linearized, they result in a set of dynamic gravitational field relations similar to the Maxwell relations. Thus one can use intuitive pictures from electromagnetic theory to design theoretical models. Whether the effects predicted by the linearized theory really exist, will, of course, have to be checked by repeating the calculations with the nonlinearized field equations.

The essential point is that all of these unusual forces create accelerations which are independent of the mass of the test body and the forces are, therefore, indistinguishable from the usual Newtonian gravitational force.

NON-NEWTONIAN GRAVITATIONAL FORCES

Effect of Rotating Masses on Stationary Bodies

By using Einstein's general theory of relativity for a rotating system of masses, it can be shown that in addition to the usual Newtonian term, the gravitational scalar potential contains terms which arise from the rotation of the body. One of the shapes which has been rigorously investigated is the rotating massive ring.[3]

For a massive ring rotating in the x-y plane, the non-Newtonian acceleration on a stationary test body near the origin is approximately

$$\ddot{x} = (MG\omega^2/2c^2R)x,$$
$$\ddot{y} = (MG\omega^2/2c^2R)y, \qquad (1)$$

and

$$\ddot{z} = -(MG\omega^2/c^2R)z.$$

Where M and R are the mass and radius of the ring, ω is the angular velocity, $G = 6.67 \times 10^{-11}$ m³/kg-sec² is the Newtonian gravitational constant, c is the speed of light, and x, y, z are the coordinates of the test body with respect to the origin of the rotating mass. From these equations it is evident that the rotating mass not only

* This paper was submitted as an essay to the Gravity Research Foundation in 1962 and received the Second Award.

[1] A. Einstein, Ann. Phys. (New York) **49**, 769 (1916); see also A. Einstein, *Principle of Relativity* (Dover Publications, Inc., New York, 1923).

[2] H. Thirring, Z. Phys. **19**, 33 (1918); and **22**, 29 (1921).

[3] C. Möller, *The Theory of Relativity* (Oxford University Press, London, 1952), pp. 317ff.

[4] J. Weber, *General Relativity and Gravitational Waves* (Interscience Publishers, Inc., New York, 1961), p. 160.

[5] J. Landau and E. Lifshitz, *The Classical Theory of Fields* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1959).

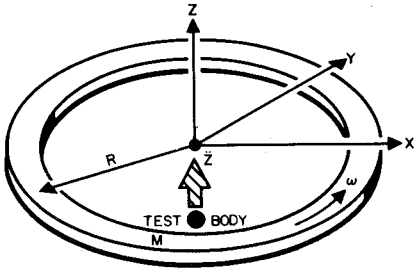


FIG. 1. Non-Newtonian component of acceleration on a test body below the center of a rotating massive ring.

forces the test body away from the axis in an imitation of centrifugal force, but also pulls it upward into the plane of rotation as shown in Fig. 1.

Effect of Rotating Masses on Moving Bodies

It has been pointed out that a rotating mass will exert a force similar to centrifugal force on a stationary test body; also, if the test body is moving at some constant velocity **v**, it will experience an additional force which is proportional to the cross product of the angular velocity of the rotating mass and the linear velocity of the test body. The additional force can be compared with two others: mechanically, it acts like a very weak Coriolis force; electrically,[4,6] it acts like the gravitational equivalent of the Lorentz force on a charged particle moving through a magnetic field.

One of the shapes which has been investigated is the rotating massive spherical shell. The acceleration on a test body moving with a velocity **v** inside the shell is approximately[4]

$$\ddot{x} = (GM/3c^2R)[\tfrac{4}{5}\omega^2 x - 8\omega v_y],$$
$$\ddot{y} = (GM/3c^2R)[\tfrac{4}{5}\omega^2 y + 8\omega v_x], \qquad (2)$$

and

$$\ddot{z} = -(8GM\omega^2 z/15c^3R),$$

where v_x and v_y are the x and y components of the velocity of the test body and M and R are the mass and radius of the spherical shell. The first term in each expression is the centrifugal-type force on a stationary test body described in the previous section. The second term in the x and y components of the acceleration depends upon the velocity of the test body as shown in Fig. 2.

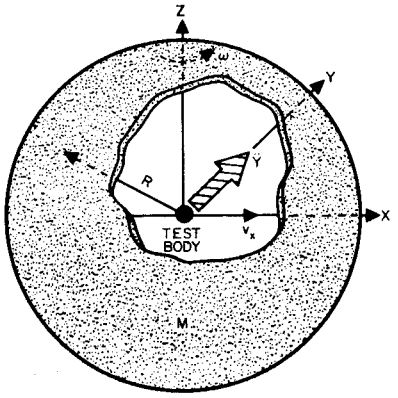


FIG. 2. Non-Newtonian component of acceleration on a test body with a velocity v_x at the center of a rotating massive sphere.

Effect of Accelerated Masses on Stationary Bodies

In employing Einstein's Theory to investigate the effect of a large accelerated mass on a small test body, it is found that the accelerated body drags the test body along with it. The exact equations are[4]

$$\ddot{\mathbf{r}} = \nabla(GM/R) + (4GM/c^2R)\mathbf{a} + (4GM/c^3)(\mathbf{a}\cdot\mathbf{R}/R^2)\mathbf{v}, \qquad (3)$$
$$= F_1/m + F_2/m + F_3/m,$$

where M, **a**, and **v** are the mass, acceleration, and velocity of the large body, respectively, and **R** is the distance from the small body to the large body. In addition to the usual Newtonian attraction, the test body experiences forces in the direction of the acceleration and the velocity of the large body as shown in Fig. 3.

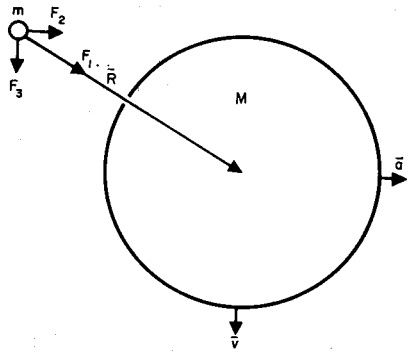


FIG. 3. Forces on a test body near a massive, accelerating, moving body.

[6] Robert L. Forward, Proc. IRE 49, 892 (1961).

CREATING NON-NEWTONIAN GRAVITATIONAL FORCES

Devices Using Moving Masses

The equations in the previous section contain two common factors. One is the Newtonian gravitational field,

$$GM/r^2 \approx (4\pi/3)G\rho r, \tag{4}$$

and the other is the ratio of the characteristic system velocity to the velocity of light such as

$$v^2/c^2 \quad \text{or} \quad ar/c^2. \tag{5}$$

In order to obtain measurable amounts of gravitational force, these quantities must be as large as possible. To obtain a high gravitational field, either a large mass or a high density is required. The greater the density, the less total mass necessary to achieve the same gravitational field.

To obtain high rotational velocities, we cannot use the mechanical strength of materials since this limits the obtainable equatorial velocities to approximately the speed of sound.[7] Because of this, it will be necessary to use fields to hold the systems together under inertial stresses. One example would be the high gravitational fields obtainable with dense matter. However, any practical device which might be constructed would probably use electric or magnetic fields. By using electromagnetic forces to contain rotating systems, it would be possible for the masses to reach relativistic velocities; thus a comparatively small amount of matter, if dense enough and moving fast enough, could produce usable gravitational effects.

An example of a system held together by gravitational fields is a contact binary dwarf star system. The force equation describing the mutual rotation of the stars is

$$-(GMM/(2r)^2) = -M\omega^2 r = -(Mv^2/r). \tag{6}$$

A particle coming near the system will experience not only a radial acceleration[4]

$$\ddot{x}_r = -GM\left[\frac{1}{(b-r)^2}+\frac{1}{(b+r)^2}\right] - \frac{4GMa}{c^2}\left[\frac{1}{b-r}-\frac{1}{b+r}\right], \tag{7}$$

but also a tangential acceleration

$$\ddot{x}_t = \frac{4GMva}{c^3}\left[\frac{1}{b-r}+\frac{1}{b+r}\right] \approx \frac{1}{c^3 b}\left(\frac{GM}{r}\right)^{\frac{5}{2}}, \tag{8}$$

where b is the distance of the particle from the center of the system.

In applying these equations to a test object (such as a space vehicle) passing by this star system, it is evident that, in general, the radial acceleration will not introduce any net change in velocity, but that the tangential acceleration will transfer energy and momentum to the vehicle. For a neutron binary star, this acceleration can be greater than one million g's; although these accelerations seem very high, there are no stresses on the human body because the forces are gravitational. Such systems could be used to accelerate space vehicles to nearly the speed of light.

Devices Based on Analogies between Electromagnetic and Gravitational Fields

There are other types of devices for obtaining non-Newtonian gravitational forces which follow from the known electromagnetic analogies based on Einstein's general theory of relativity; however, these devices have not been analyzed using the complete field equations. For instance, two rotating gyroscopes should repel each other if oriented properly, and two pipes with massive liquid flowing through them should exhibit a pinch effect.

Although the rotating gyroscopes and flowing masses should exhibit non-Newtonian forces, the linearized theory does not predict an interaction with a stationary test body. A device which electromagnetic analogies predict might be able to create a non-Newtonian gravitational force that will accelerate a nonrotating, nonmoving body is one that contains accelerated masses whose mass flow is like the current flow in a wire wound torus.[8] If we look at the electromagnetic model, a current **I** through the wire causes a magnetic field in the torus. If the current constantly increases, then the magnetic field also increases with time. This time-varying magnetic field then creates a dipole electric field. (Fig. 4.)

[7] J. W. Beams, Sci. Am. 204, No. 4, 134 (1961).

[8] Robert L. Forward, Proc. IRE 44, 1402 (1961).

The value of the electric field at the center of the torus is

$$E=-\dot{B}=+(d/dt)(\mu NIr^2/4\pi R^2), \qquad (9)$$

where R is radius of the torus, r is the radius of one of the loops of wire wound around it, and N is the total number of turns.

If we replace the wires with pipes carrying a massive liquid, then the known analogy between the electromagnetic and gravitational fields can now be used.[6] All the electromagnetic quantities are transformed to their equivalent gravitational quantities to obtain

$$G=-\dot{K}=-\frac{d}{dt}\left(\frac{\eta NTr^2}{4\pi R^2}\right), \qquad (10)$$

where G is the gravitational field generated by the total mass current NT, and $\eta=3.73\times10^{-26}$ m/kg is the gravitational equivalent to the magnetic permeability.[6] (See Fig. 5.)

It is important to notice in the above equation that the time derivative operates on the entire quantity in the brackets. One would normally say that all quantities are independent of time except the mass flow T. If the amount of mass flow increases with time, then $\dot{T}>0$, and thus a gravitational field G can be generated. However, in electromagnetism, the permeability of some materials such as iron is nonlinear, which allows the construction of highly efficient electromagnetic field generators. A material with a highly nonlinear η would also be useful in the construction of efficient gravitational field generators.

RESEARCH AREAS

Dense Materials

It has been emphasized that in order to achieve measurable gravitational effects with moderate amounts of mass, dense matter is required. Thus, the study of degenerate matter could lead to the generation and control of gravitation. Methods must be found to manufacture, contain, and control matter with densities from 10^8 to 10^{15} g/cm^3. The best starting point appears to be an investigation of the neutron–neutron interaction. By using extremely low temperatures and strong magnetic fields, one should be able to cool the thermal neutrons from a pile and concentrate them into a small region through the interaction of the magnetic field with the magnetic moment of the neutron.[9] The Fermi energy will limit the density to about 10^{-6} g/cm^3, but the formation of tetraneutrons (n^4) or the existence of a superconductive-type phase space condensation will create bosons which do not have this limitation. The results of a more comprehensive study of the properties of a cold neutron gas and the methods for containment will be given in a future paper.

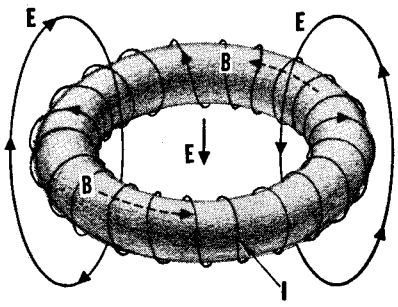


FIG. 4. Generator of a dipole electric field.

Gravitational Properties of Matter

In studying analogies between electromagnetism and gravitation, it can be seen that one analogous quantity has not been investigated. This is the gravitational equivalent to the magnetic permeability. Electrical power distribution systems depend upon the anomalously large and nonlinear permeability of iron and other magnetic materials. Since all atoms have spin, all materials will have a gravitational permeability which is different from that of free space. Rough calculations show that this difference is very small, but experimental investigation may

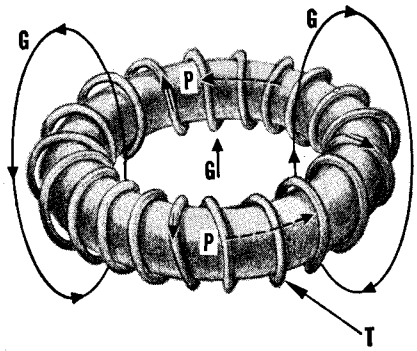


FIG. 5. Generator of a dipole gravitational field.

[9] V. V. Vladimirskii, Soviet Phys.—JETP 12, 740 (1961).

find materials with anomalously large or nonlinear properties that can be used to enhance time-varying gravitational fields. Also, since the magnetic moment and the inertial moment are combined in an atom, it may be possible to use this property to convert time-varying electromagnetic fields into time-varying gravitational fields. At present, the only way to search for such materials is to intersperse wedges of material between gravitational wave generators and detectors, such as those described by J. Weber,[10] and look for a change in amplitude or direction of propagation. The first efforts in this direction[11] have been carried out by the Russian workers Braginsky, Rudenko, and Rukman with negative results.

It is obvious that research in the field of gravitation will be very difficult since even the most optimistic calculations indicate that very large devices will be required to create usable gravitational forces. Antigravity, like space travel, will probably have no direct effect on the daily life of the average person. Future progress in the control of gravitation, like all modern sciences, will require special projects involving large sums of money, men, and energy.

[10] J. Weber, Phys. Rev. **117**, 306 (1960).

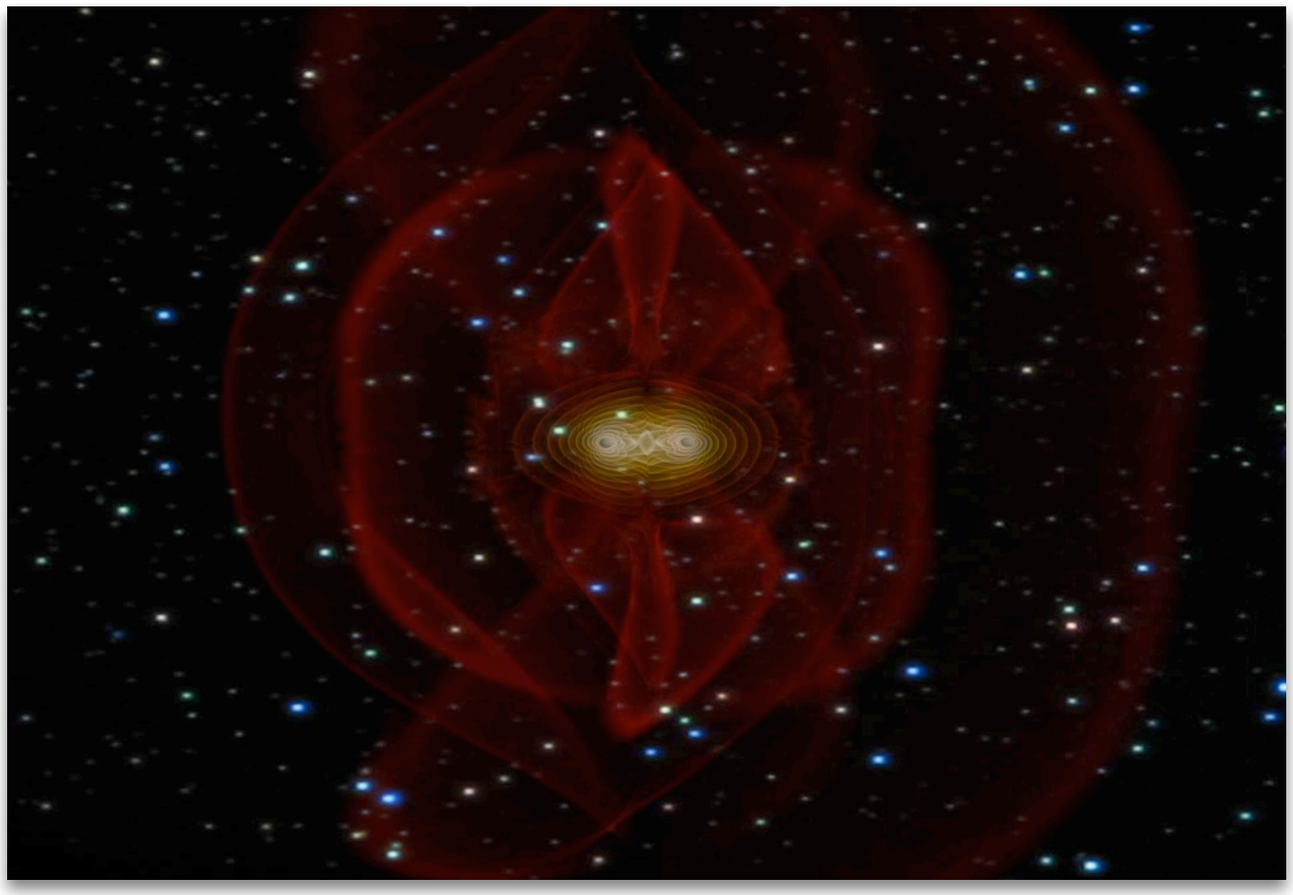
[11] V. B. Braginsky, V. N. Rudenko, and G. I. Rukman, Zhur. Eksp. i Theoret. Fiz. **43**, 51 (1962) (in Russian with English abstract).

Gravitational Waves

Theory - Sources - Detection

Kostas Glampedakis

UNIVERSIDAD DE MURCIA



Contents

- **Part I: Theory of gravitational waves.**

✓ Properties.

✓ Wave generation/the quadrupole formula.

✓ Basic estimates.

- **Part II: Gravitational wave sources.**

✓ Binary systems/the Hulse-Taylor pulsar.

- **Part III: Detection of gravitational waves.**

✓ Resonant detectors/Laser interferometers.

✓ Ground/Space-based detectors.

Part I

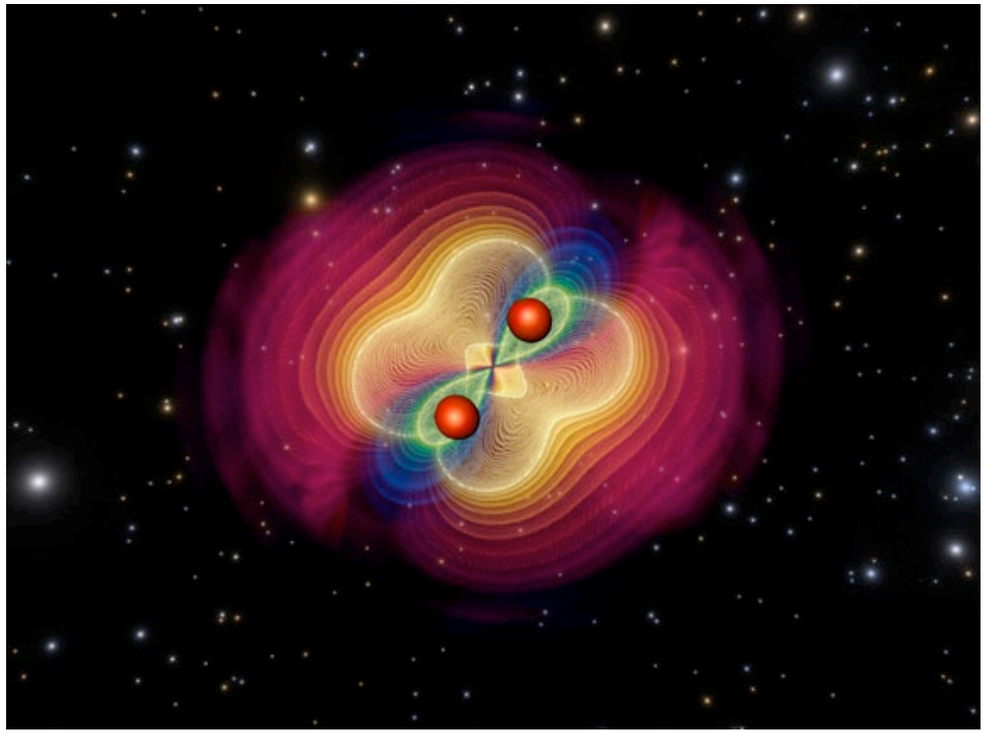
GW theory

Newtonian vs General Relativistic gravity

Newtonian field equations	**GR field equations**
$\nabla^2\Phi = 4\,\pi G\rho$	$G^{ab} = \dfrac{8\pi G}{c^4} T^{ab}$
Source: mass density Gravitational field: scalar Φ	Source: energy-momentum tensor (includes mass densities/currents) Gravitational field: metric tensor g_{ab}

GWs: origins

- Electromagnetism: accelerating charges produce EM radiation.



- Gravitation: accelerating masses produce gravitational radiation. (another hint: gravity has finite speed.)

GWs in linear gravity

- We consider weak gravitational fields:

$$g_{\mu\nu} \approx \eta_{\mu\nu} + h_{\mu\nu} + \mathcal{O}(h_{\mu\nu}^2)$$

↑

flat Minkowski metric

- The GR field equations in vacuum reduce to the standard wave equation:

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2 \right) h^{\mu\nu} = \Box h^{\mu\nu} = 0$$

- Comment: GR gravity like electromagnetism has a "gauge" freedom associated with the choice of coordinate system. The above equation applies in the so-called "transverse-traceless (TT)" gauge where

$$h_{0\mu} = 0, \qquad h^{\mu}_{\mu} = 0$$

Wave solutions

- Solving the previous wave equation in weak gravity is easy. The solutions represent “plane waves”:

$$h_{\mu\nu} = A_{\mu\nu} e^{i k_a x^a}$$

wave-vector

- Basic properties: $A_{\mu\nu} k^\mu = 0, \qquad k_a k^a = 0$

transverse waves

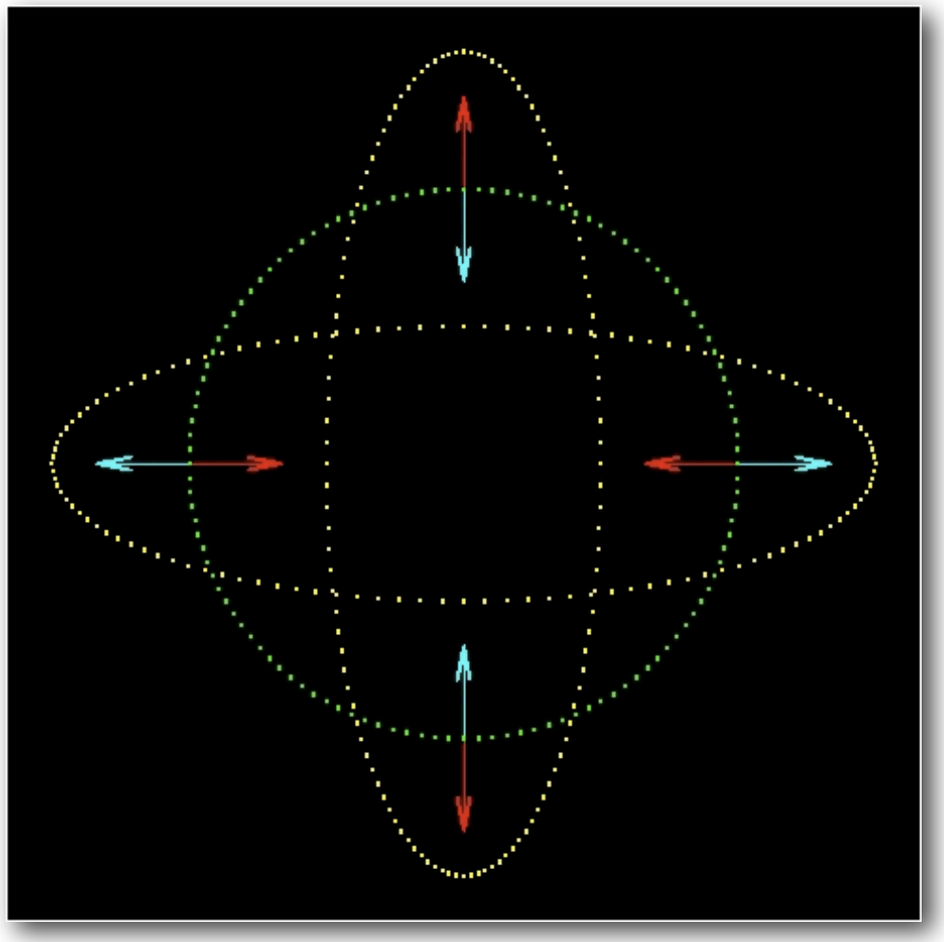
null vector = propagation along light rays

- Amplitude: $A^{\mu\nu} = h_+ e_+^{\mu\nu} + h_{\rm x} e_{\rm x}^{\mu\nu}$

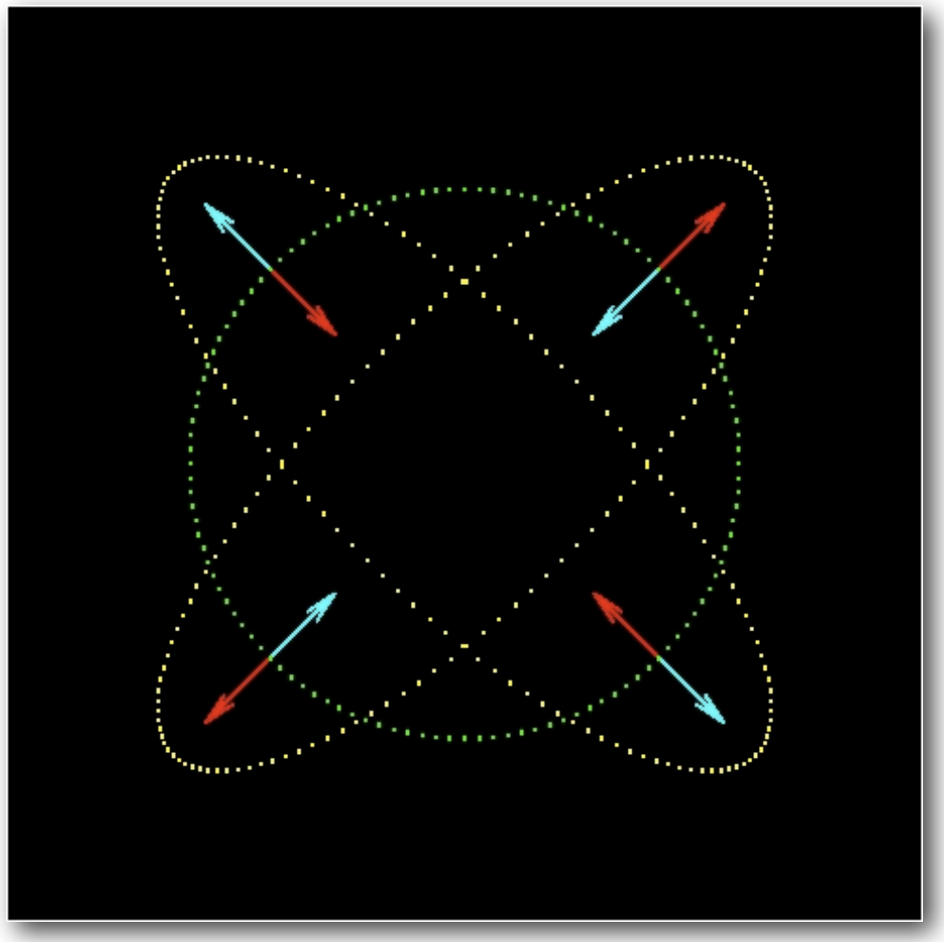
two polarizations

GWs: polarization

- GWs come in two polarizations:



"+" polarization



"x" polarization

GWs: more properties

- EM waves: at lowest order the radiation can be emitted by a dipole source (l=1). Monopolar radiation is forbidden as a result of charge conservation.

- GWs: the lowest allowed multipole is the quadrupole (l=2). The monopole is forbidden as a result of mass conservation. Similarly, dipole radiation is absent as a result of momentum conservation.

- GWs represents propagating “ripples in spacetime” or, more accurately, a propagating curvature perturbation. The perturbed curvature (Riemann tensor) is given by (in the TT gauge):

$$R^{\rm TT}_{j0k0} = -\frac{1}{2}\,\partial_t^2 h^{\rm TT}_{jk}\,, \qquad\qquad j,k = 1,2,3$$

GWs and curvature

- As we mentioned, GWs represent a fluctuating curvature field.
- Their effect on test particles is of tidal nature.
- Equation of geodetic deviation (in weak gravity):

$$\frac{d^2\xi^k}{dt^2} = -R^{k\ \mathrm{TT}}_{0j0}\,\xi^j$$

distance between geodesics (test particles)

- Newtonian limit: $R^{\mathrm{TT}}_{k0j0} \approx \dfrac{\partial^2\Phi}{\partial x^k \partial x^j}$

Newtonian grav. potential

GWs vs EM waves

- Similarities:

✓ Propagation with the speed of light.

✓ Amplitude decreases as ~ 1/r.

✓ Frequency redshift (Doppler, gravitational, cosmological).

- Differences:

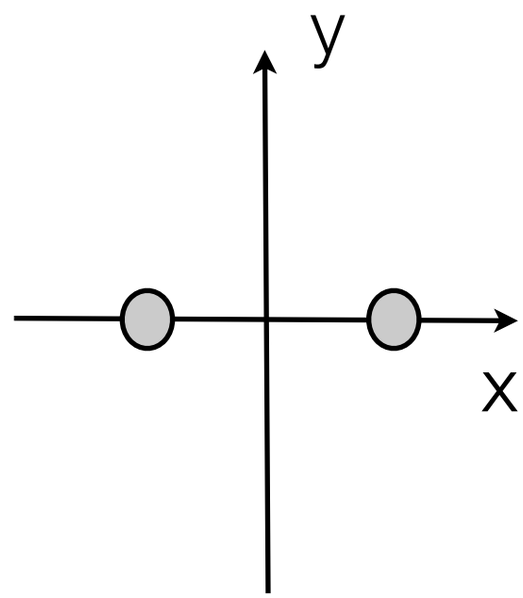
✓ GWs propagate through matter with little interaction. Hard to detect, but they carry uncontaminated information about their sources.

✓ Strong GWs are generated by bulk (coherent) motion. They require strong gravity/high velocities (compact objects like black holes and neutron star).

✓ EM waves originate from small-scale, incoherent motion of charged particles. They are subject to “environmental” contamination (interstellar absorption etc.).

Effect on test particles (I)

- We consider a pair of test particles on the cartesian axis Ox at distances $\pm x_0$ from the origin and we assume a GW traveling in the z-direction.



- Their distance will be given by the relation:

$$dl^2 = g_{\mu\nu} dx^\mu dx^\nu = \ldots = -g_{11} dx^2 =$$

$$= (1 - h_{11})(2x_0)^2 = [1 - h_+ \cos(\omega t)]\,(2x_0)^2$$

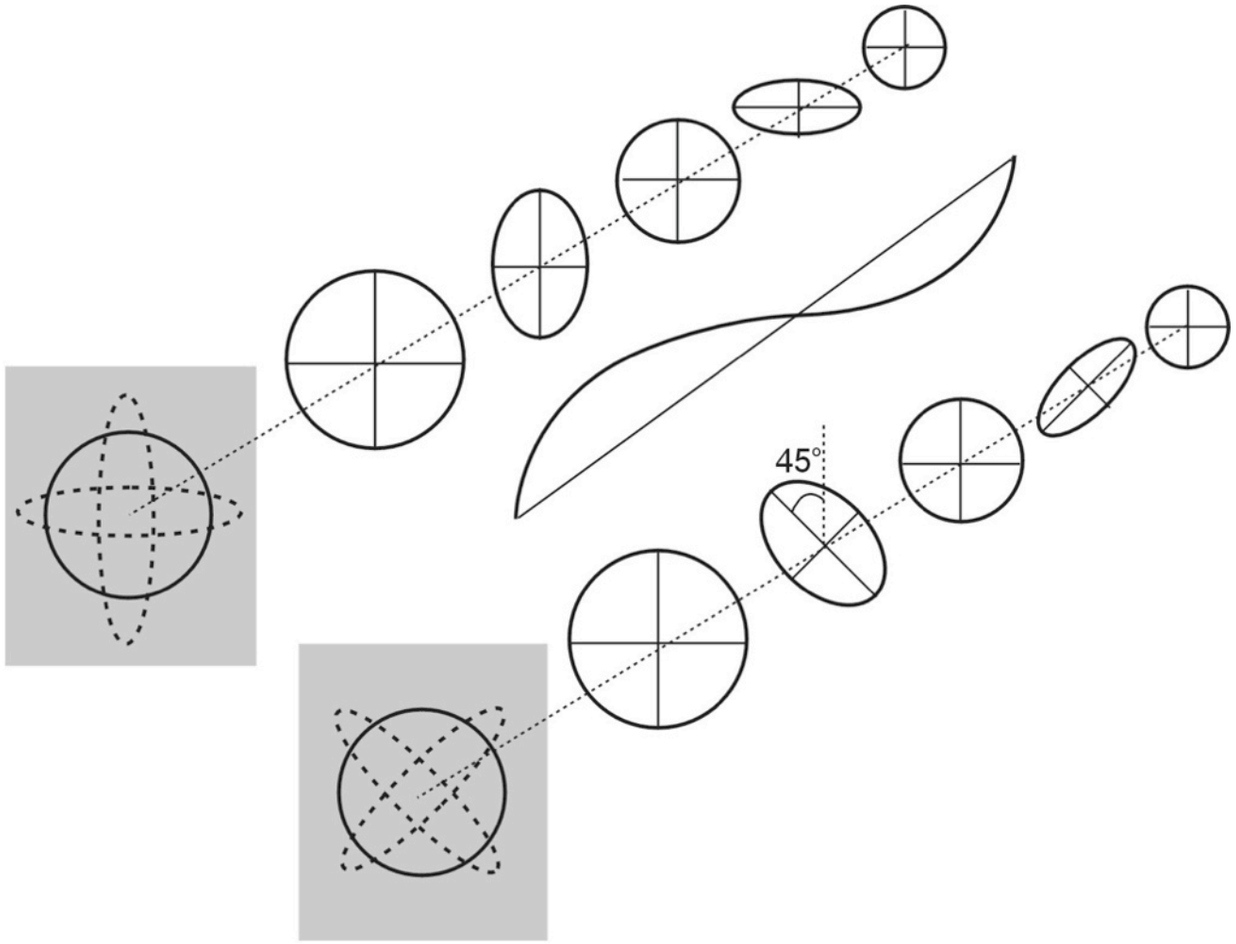
$$dl \approx \left[1 - \frac{1}{2} h_+ \cos(\omega t)\right](2x_0)$$

Effect on test particles (II)

- Similarly for a pair of particles placed on the y-axis:

- Comment: the same result can be derived using the geodetic deviation equation.

$$dl \approx \left[1 + \frac{1}{2} h_+ \cos(\omega t)\right] (2y_0)$$



The quadrupole formula

- Einstein (1918) derived the quadrupole formula for gravitational radiation by solving the linearized field equations with a source term:

$$\Box h^{\mu\nu}(t,\vec{x}) = -\kappa T^{\mu\nu}(t,\vec{x}) \quad \longrightarrow \quad h^{\mu\nu} = -\frac{\kappa}{4\pi}\int_V d^3x' \frac{T^{\mu\nu}(t-|\vec{x}-\vec{x}'|,\vec{x}')}{|\vec{x}-\vec{x}'|}$$

- This solution suggests that the wave amplitude is proportional to the second time derivative of the quadrupole moment of the source:

$$h^{\mu\nu} = \frac{2}{r}\frac{G}{c^4}\ddot{Q}^{\mu\nu}_{\rm TT}(t-r/c) \qquad\qquad Q^{\mu\nu}_{\rm TT} = \int d^3x\,\rho\left(x^\mu x^\nu - \frac{1}{3}\delta^{\mu\nu}r^2\right)$$

(quadrupole moment in the “TT gauge” and at the retarded time t-r/c)

- This result is quite accurate for all sources, as long as the wavelength is much longer than the source size R.

GW luminosity

- GWs carry energy. The stress-energy carried by GWs cannot be localized within a wavelength. Instead, one can say that a certain amount of stress-energy is contained in a region of the space which extends over several wavelengths. The stress-energy tensor can be written as:

$$T^{\rm GW}_{\mu\nu} = \frac{c^4}{32\pi G}\langle \partial_\mu h^{\rm TT}_{ij} \partial_\nu h^{ij}_{\rm TT}\rangle$$

- Using the previous quadrupole formula we obtain the GW luminosity:

$$L_{\rm GW} = \frac{dE_{\rm GW}}{dt} = \int dA\, T^{\rm GW}_{0j}\hat{n}^j \quad \longrightarrow \quad L_{\rm GW} = \frac{1}{5}\frac{G}{c^5}\langle \dddot{Q}^{\rm TT}_{\mu\nu}\, \dddot{Q}^{\mu\nu}_{\rm TT}\rangle$$

Basic estimates (I)

- The quadrupole moment of a system is approximately equal to the mass M of the part of the system that moves, times the square of the size R of the system. This means that the 3rd-order time derivative of the quadrupole moment is:

$$\dddot{Q} \sim \frac{MR^2}{T^3} \sim \frac{Mv^2}{T} \sim \frac{E_{\rm ns}}{T}$$

v = mean velocity of source's non-spherical motion,

E_{ns} = kinetic energy of non-spherical motion

T = timescale for a mass to move from one side of the system to the other.

- For a self gravitating system: $T \sim \sqrt{R^3/GM}$

- This relation provides a rough estimate of the characteristic frequency of the system $f \sim 2\pi/T$.

Basic estimates (II)

- The luminosity of GWs from a given source is approximately:

$$L_{\rm GW} \sim \frac{G^4}{c^5}\left(\frac{M}{R}\right)^5 \sim \frac{G}{c^5}\left(\frac{M}{R}\right)^2 v^6 \sim \frac{c^5}{G}\left(\frac{R_{\rm Sch}}{R}\right)^2 \left(\frac{v}{c}\right)^6$$

where $R_{\rm Sch} = 2GM/c^2$ is the Schwarzschild radius of the source. It is obvious that the maximum GW luminosity can be achieved if $R \sim R_{\rm Sch}$ and v ~ c.
That is, the source needs to be compact and relativistic.

- Using the above order-of-magnitude estimates, we can get a rough estimate of the amplitude of GWs at a distance r from the source:

$$h \sim \frac{G}{c^4}\frac{E_{\rm ns}}{r} \sim \frac{G}{c^4}\frac{\epsilon E_{\rm kin}}{r}$$

$\epsilon =$ the kinetic energy fraction that is able to produce GWs.

Basic estimates (III)

- Another estimate for the GW amplitude can be derived from the flux formula

$$F_{\rm GW} = \frac{L_{\rm GW}}{4\pi r^2} = \frac{c^3}{16\pi G}|\partial_t h|^2$$

- We obtain:

$$h \approx 10^{-22}\left(\frac{E_{\rm GW}}{10^{-4}\,M_\odot}\right)^{1/2}\left(\frac{1\,\mathrm{kHz}}{f_{\rm GW}}\right)\left(\frac{\tau}{1\,\mathrm{ms}}\right)^{-1/2}\left(\frac{15\,\mathrm{Mpc}}{r}\right)$$

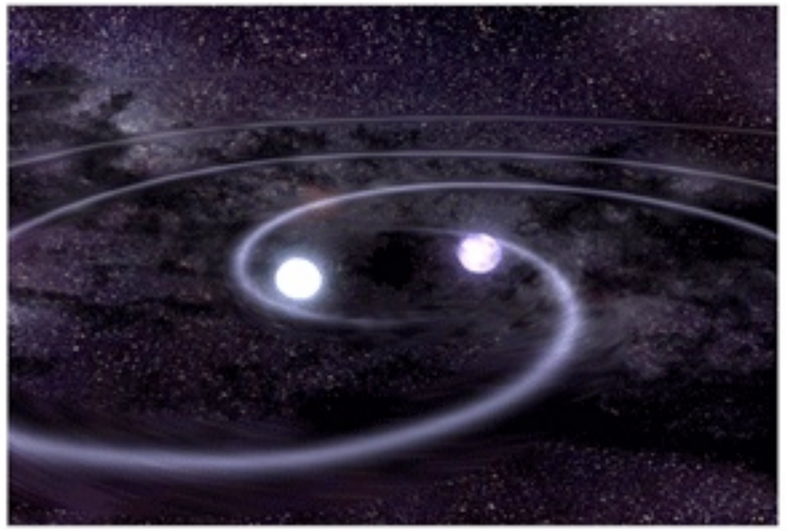
for example, this formula could describe the GW strain from a supernova explosion at the Virgo cluster during which the energy $E_{\rm GW}$ is released in GWs at a frequency of 1 kHz, and with signal duration of the order of 1 ms.

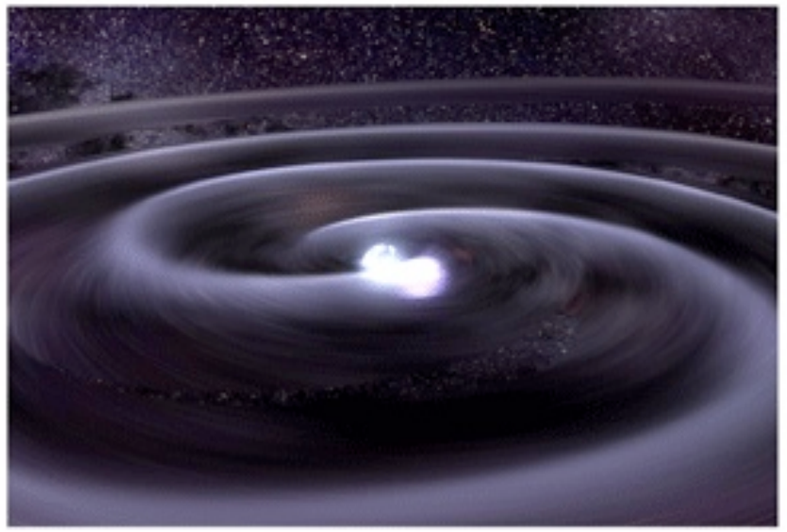
- This is why GWs are hard to detect: for a GW detector with arm length of $l = 4\,\mathrm{km}$ we are looking for changes in the arm-length of the order of

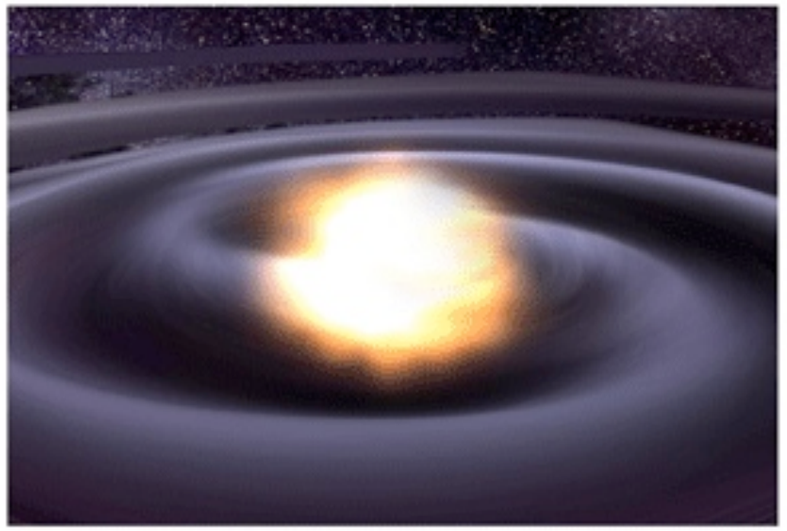
$$\Delta l = hl = 4\times 10^{-17}\,\mathrm{cm} \quad !!$$

Part II

GW sources

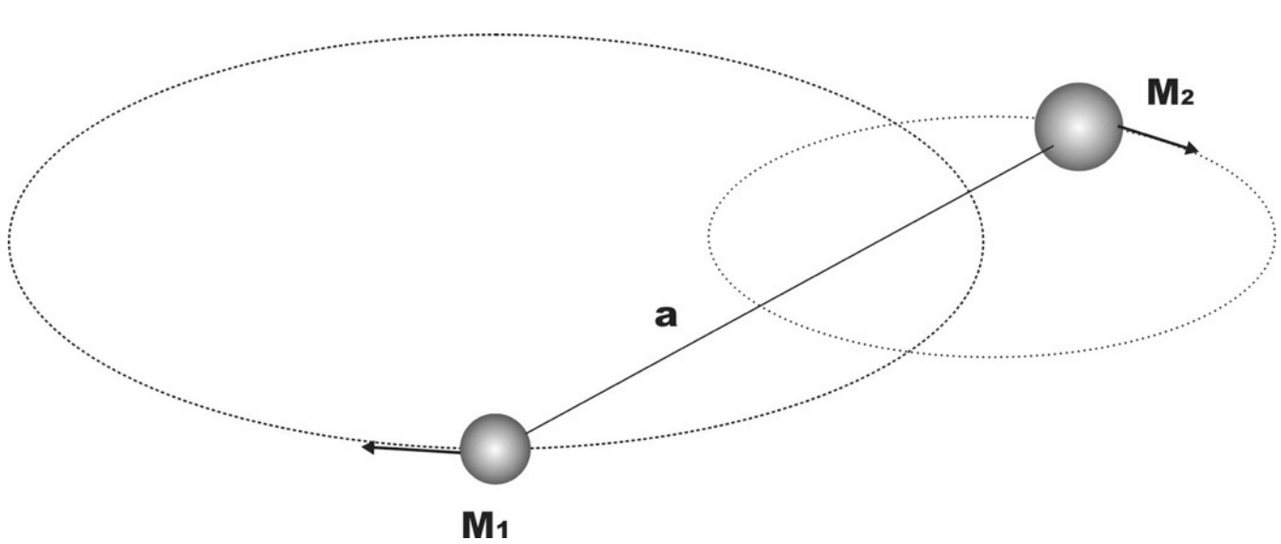






GW emission from a binary system (I)

- The binary consists of the two bodies M1 and M2 at distances a_1 and a_2 from the center of mass. The orbits are circular and lie on the x-y plane. The orbital angular frequency is Ω.



- We also define: $a = a_1 + a_2, \qquad \mu = M_1 M_2 / M, \qquad M = M_1 + M_2$

GW emission from a binary system (II)

- The only non-vanishing components of the quadrupole tensor are :

$$Q_{xx} = -Q_{yy} = (a_1^2 M_1 + a_2^2 M_2)\cos^2(\Omega t) = \frac{1}{2}\mu a^2 \cos(2\Omega t)$$

$$Q_{xy} = Q_{yx} = \frac{1}{2}\mu a^2 \sin(2\Omega t)$$

(GW frequency = 2Ω)

- And the GW luminosity of the system is (we use Kepler's 3rd law $\Omega^2 = GM/a^3$)

$$L_{\rm GW} = -\frac{dE}{dt} = \frac{G}{5c^5}(\mu\Omega a^2)^2 \langle\, 2\sin^2(2\Omega t) + 2\cos^2(2\Omega t)\,\rangle$$

$$= \frac{32}{5}\frac{G}{c^5}\mu^2 a^4 \Omega^6 = \frac{32}{5}\frac{G^4}{c^5}\frac{M^3\mu^2}{a^5}$$

GW emission from a binary system (III)

- The total energy of the binary system can be written as :

$$E = \frac{1}{2}\Omega^2 \left(M_1 a_1^2 + M_2 a_2^2\right) - \frac{G M_1 M_2}{a} = -\frac{1}{2}\frac{G\mu M}{a}$$

- As the gravitating system loses energy by emitting radiation, the distance between the two bodies shrinks at a rate:

$$\frac{dE}{dt} = \frac{G\mu M}{2a^2}\frac{da}{dt} \quad \rightarrow \quad \frac{da}{dt} = -\frac{64}{5}\frac{G^3}{c^5}\frac{\mu M}{a^3}$$

- The orbital frequency increases accordingly $\dot{T}/T = (3/2)\dot{a}/a$.

- The system will coalesce after a time: $\tau = \frac{5}{256}\frac{c^5}{G^3}\frac{a_0^4}{\mu M^4}$ (initial separation)

GW emission from a binary system (IV)

- In this analysis we have assumed circular orbits. In general the orbits can be elliptical, but it has been shown that GW emission circularizes them faster than the coalescence timescale.

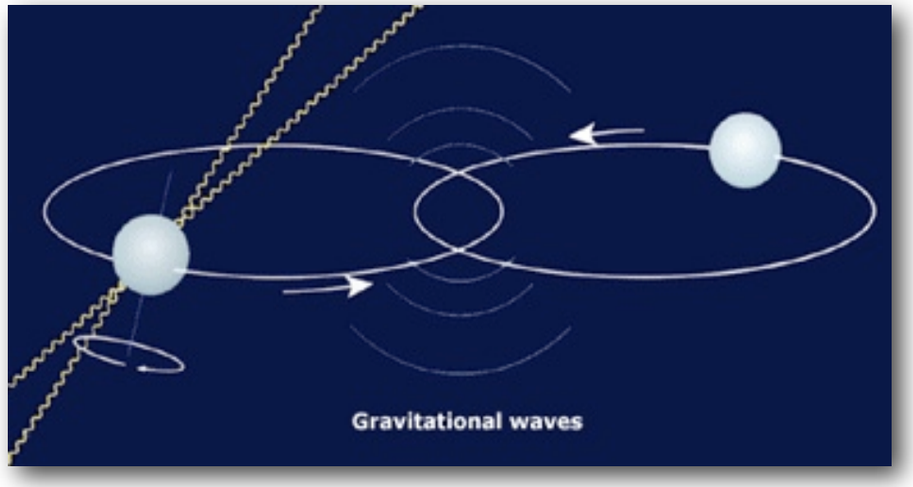
- The GW amplitude is (ignoring geometrical factors):

$$h \approx 5 \times 10^{-22} \left(\frac{M}{2.8 M_\odot}\right)^{2/3} \left(\frac{\mu}{0.7 M_\odot}\right) \left(\frac{f}{100\,\mathrm{Hz}}\right)^{2/3} \left(\frac{15\,\mathrm{Mpc}}{r}\right)$$

(set distance to the Virgo cluster, why?)

PSR 1913+16: a Nobel-prize GW source

- The now famous Hulse & Taylor binary neutron star system provided the first astrophysical evidence of the existence of GWs !



- The system's parameters: $r = 5\,\mathrm{Kpc}, \quad M_1 \approx M_2 \approx 1.4\,M_\odot, \quad T = 7\,\mathrm{h}\ 45\,\mathrm{min}$

- Using the previous equations we can predict:

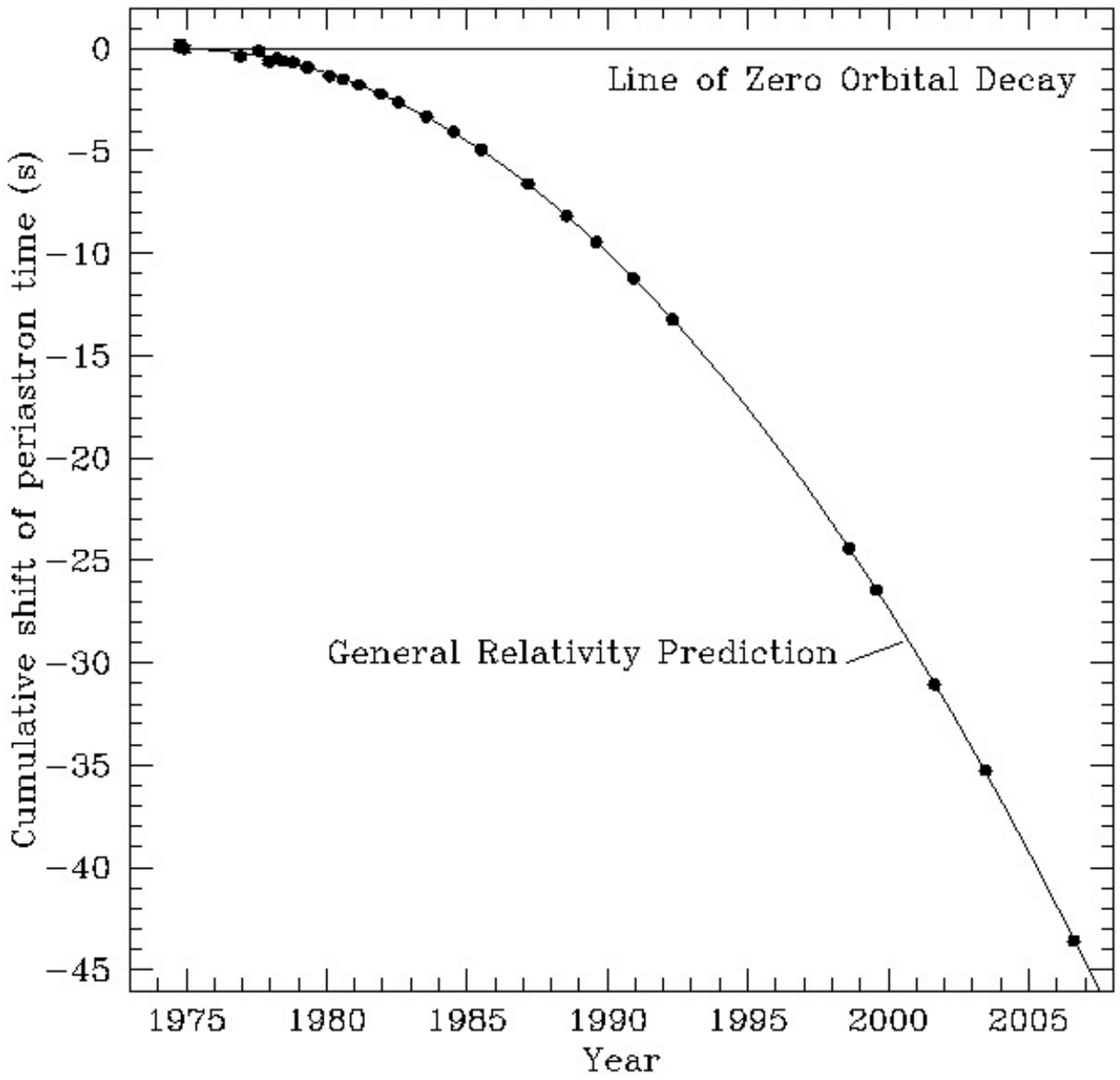
$$\dot{T} = -2.4 \times 10^{-12}\,\mathrm{sec/sec}, \qquad f_{\mathrm{GW}} = 7 \times 10^{-5}\,\mathrm{Hz}, \qquad h \sim 10^{-23}, \quad \tau \approx 3.5 \times 10^{8}\,\mathrm{yr}$$

Theory vs observations

- How can the orbital parameters be measured with such high precision?

- One of the neutron stars is a pulsar, emitting extremely stable periodic radio pulses. The emission is modulated by the orbital motion.

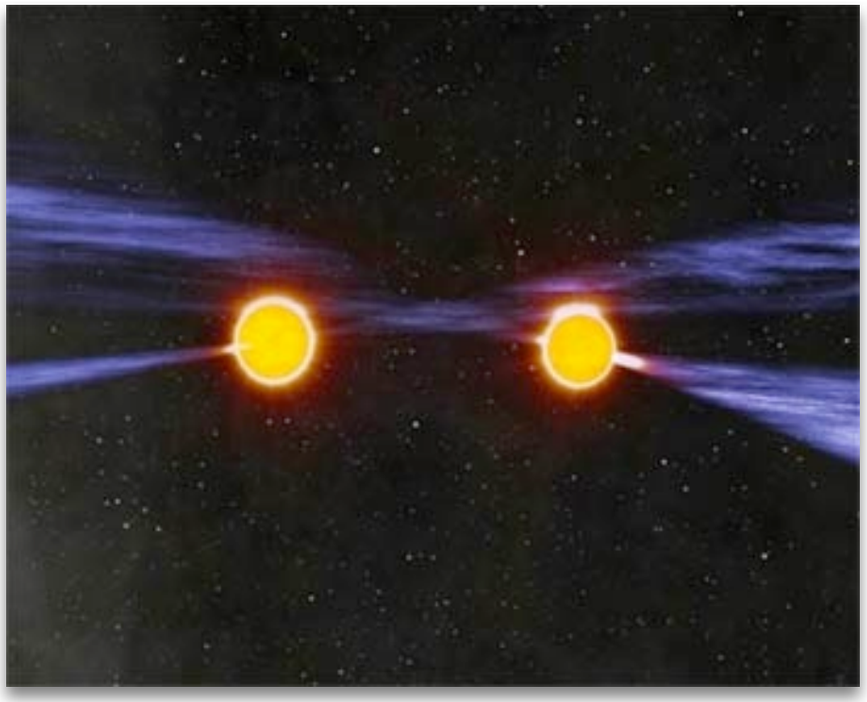
- Since the discovery of the H-T system in 1974 more such binaries were found by astronomers.

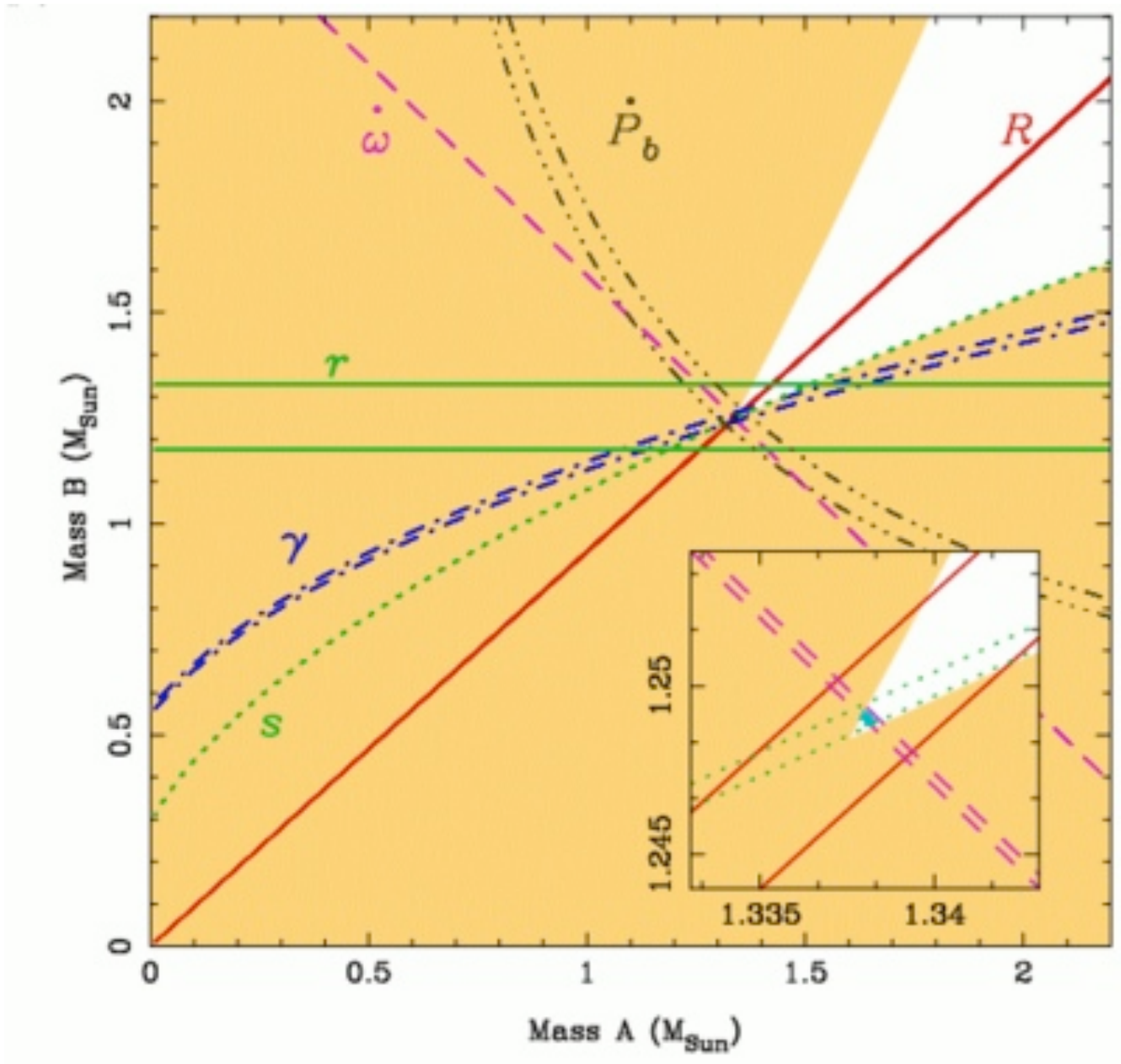


The double pulsar system

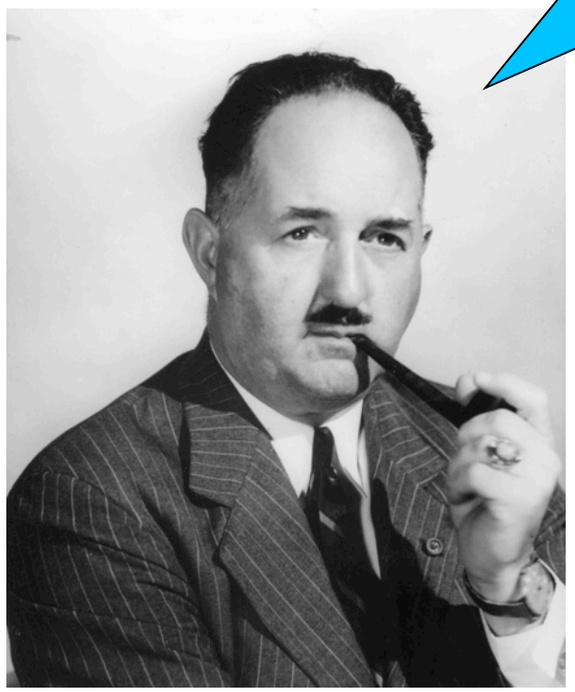
- Discovered in 2003, this binary system consists of two pulsars: PSR J0737–3039A & B.

- This rare system allows for high- precision tests of GR.



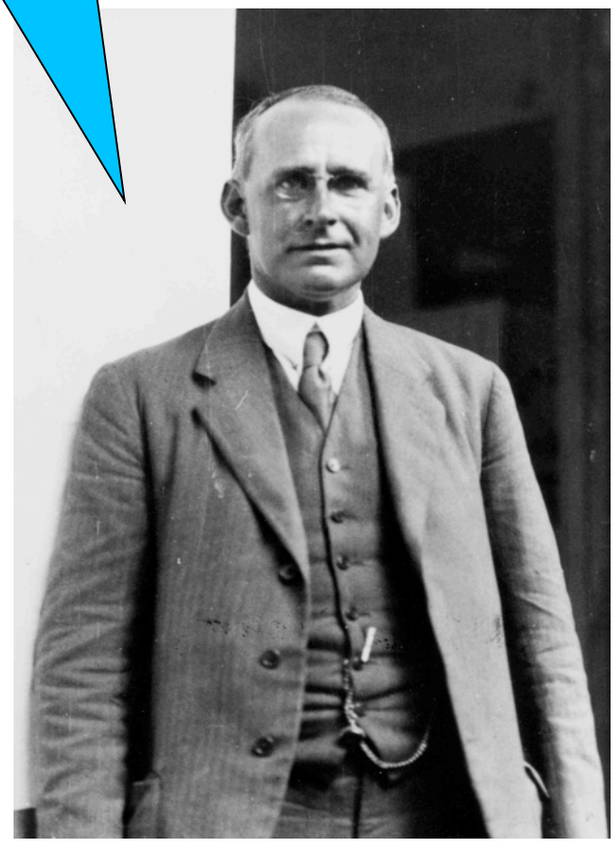


Do GWs exist? a historical footnote



They do Exist

Howard P. Robertson

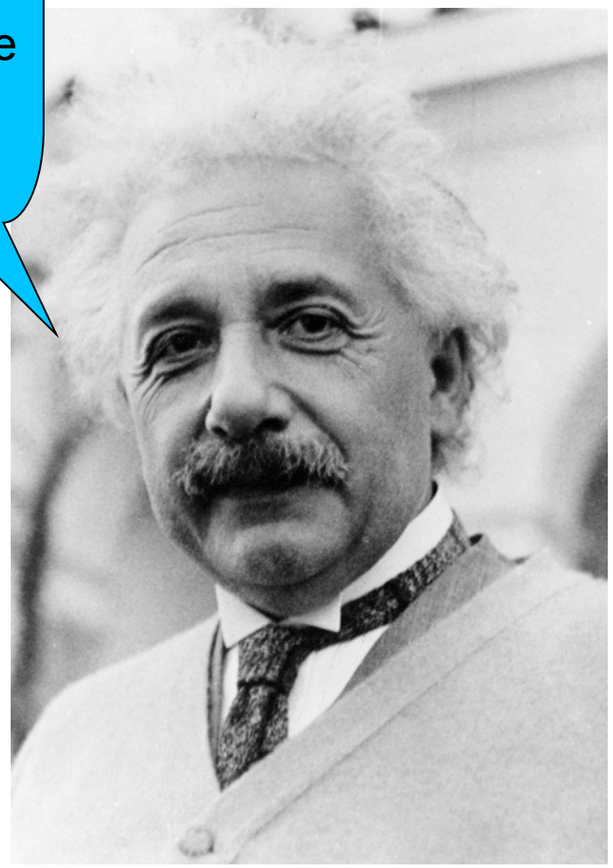


I'm skeptical

Arthur S. Eddington

They don't … They do!

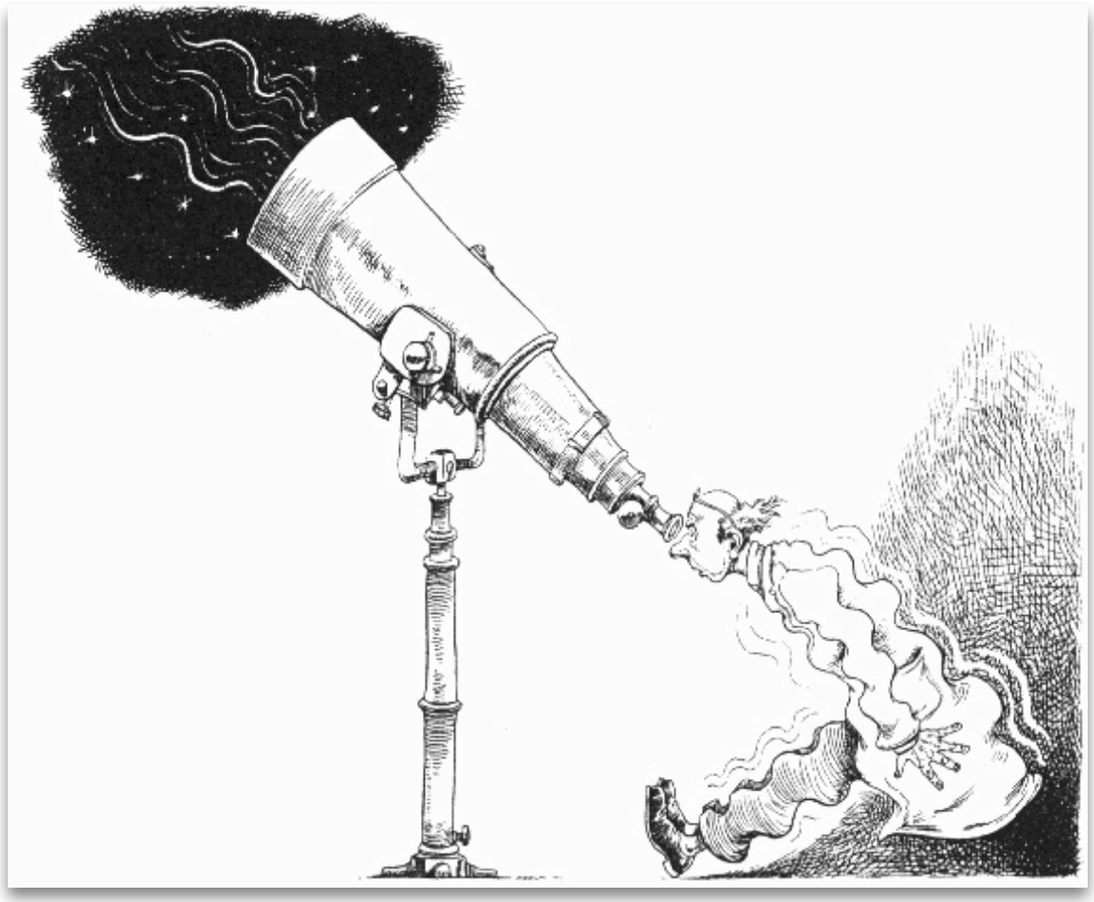
No, wait, they don't … or maybe they do



Albert Einstein

Part III

Detection of GWs



GW detectors: prehistory

- For decades after the formulation of Einstein's GR the notion of GWs was a topic for speculations and remote from real astrophysics.

- Joe Weber pioneered the construction of the first "primitive" bar detector. However, his claims of a GW detection were never verified ...

- Theoretical work in the 1970s-1990s (and the discovery of the Hulse-Taylor pulsar) advanced the popularity of GWs.

- GW astronomy is expected to become reality in the present decade.

A toy model GW detector

- Consider a GW propagating along the z-axis (with a "+" polarization and frequency ω), impinging on an idealized detector consisting of two masses joined by a spring (of length L) along the x-axis

- The resulting motion is that of a forced oscillator (with friction τ, natural frequency ω_0):

$$\ddot{\xi} + \dot{\xi}/\tau + \omega_0^2 \xi = -\frac{1}{2}\omega^2 L h_+ e^{i\omega t}$$

- The solution is:

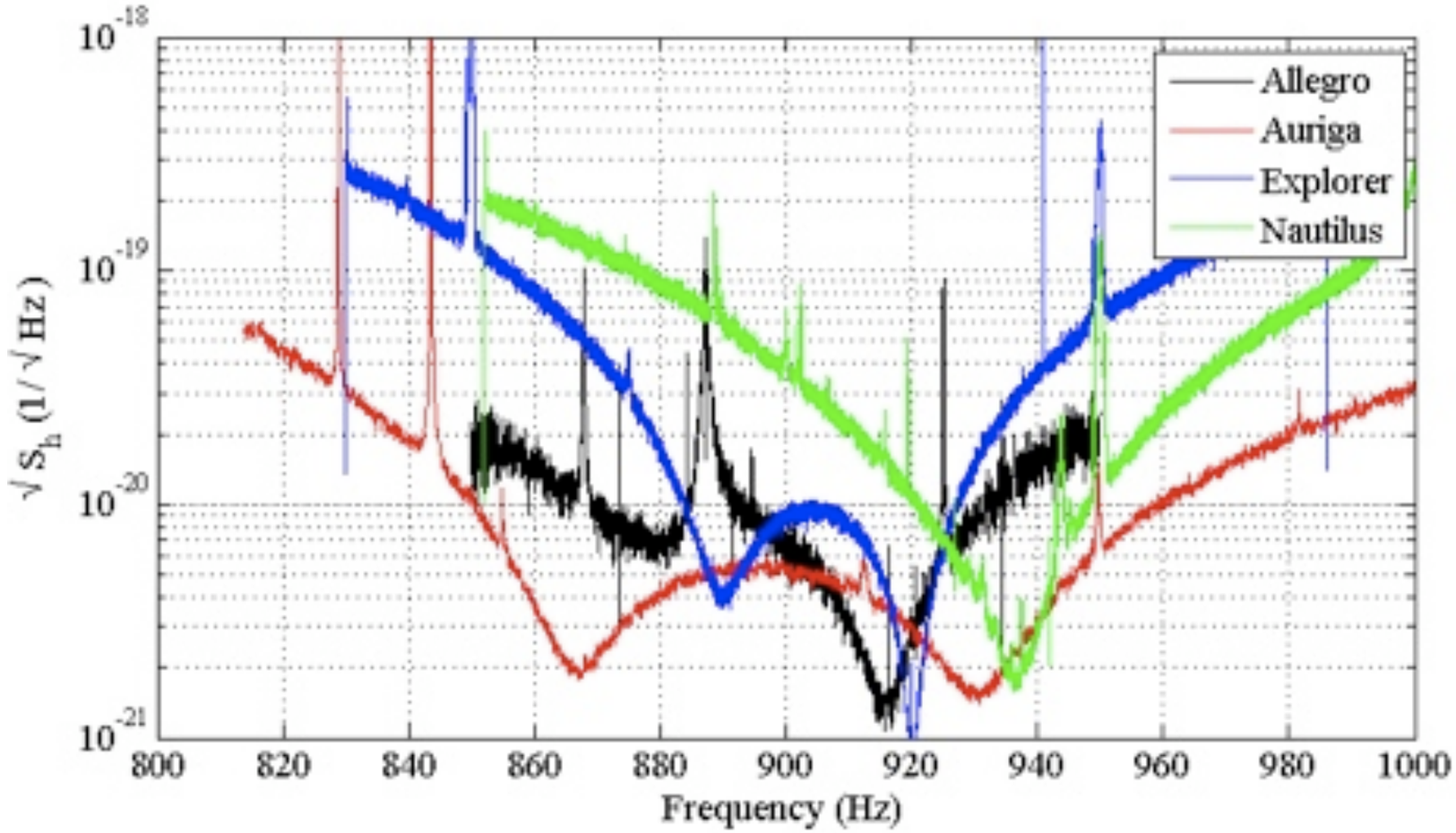
$$\xi = \frac{\omega^2 L h_+}{2(\omega_0^2 - \omega^2 + i\omega/\tau)} e^{i\omega t}$$

- The maximum amplitude is achieved at $\omega \approx \omega_0$ and has a size: $\xi_{\max} = \frac{1}{2}\omega_0 \tau L h_+$

- The detector can be optimized by increasing $\omega_0 \tau L$.

Bar detectors

- Bar detectors are narrow bandwidth instruments (like the previous toy-model)

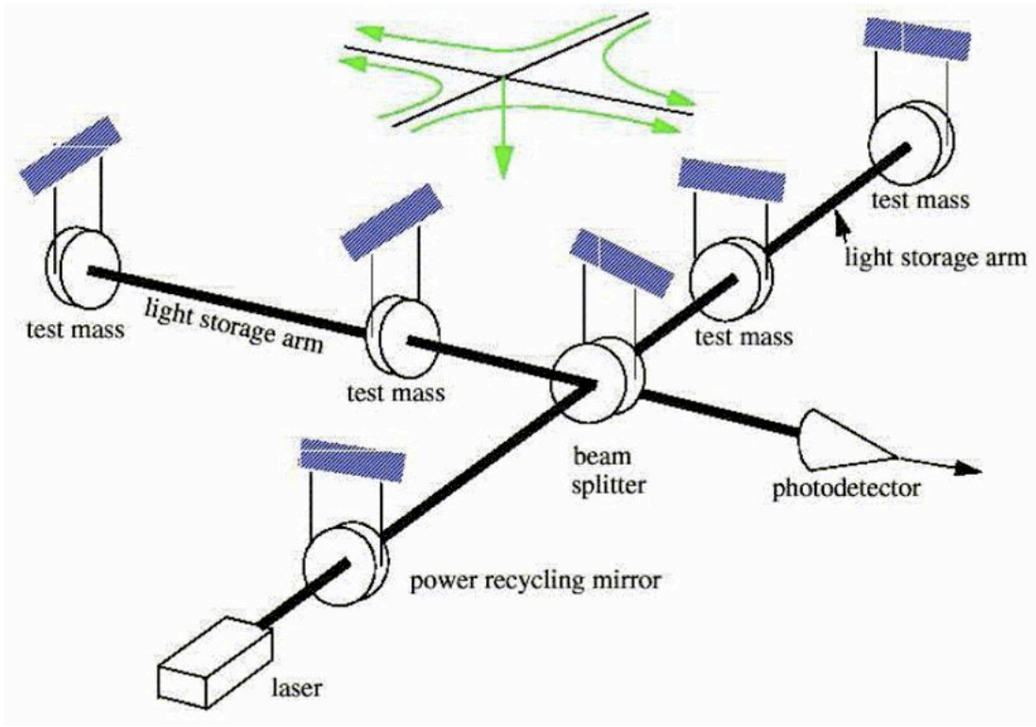


Sensitivity curves of various bar detectors

Detectors: laser interferometry

- A laser interferometer is an alternative choice for GW detection, offering a combination of very high sensitivities over a broad frequency band.

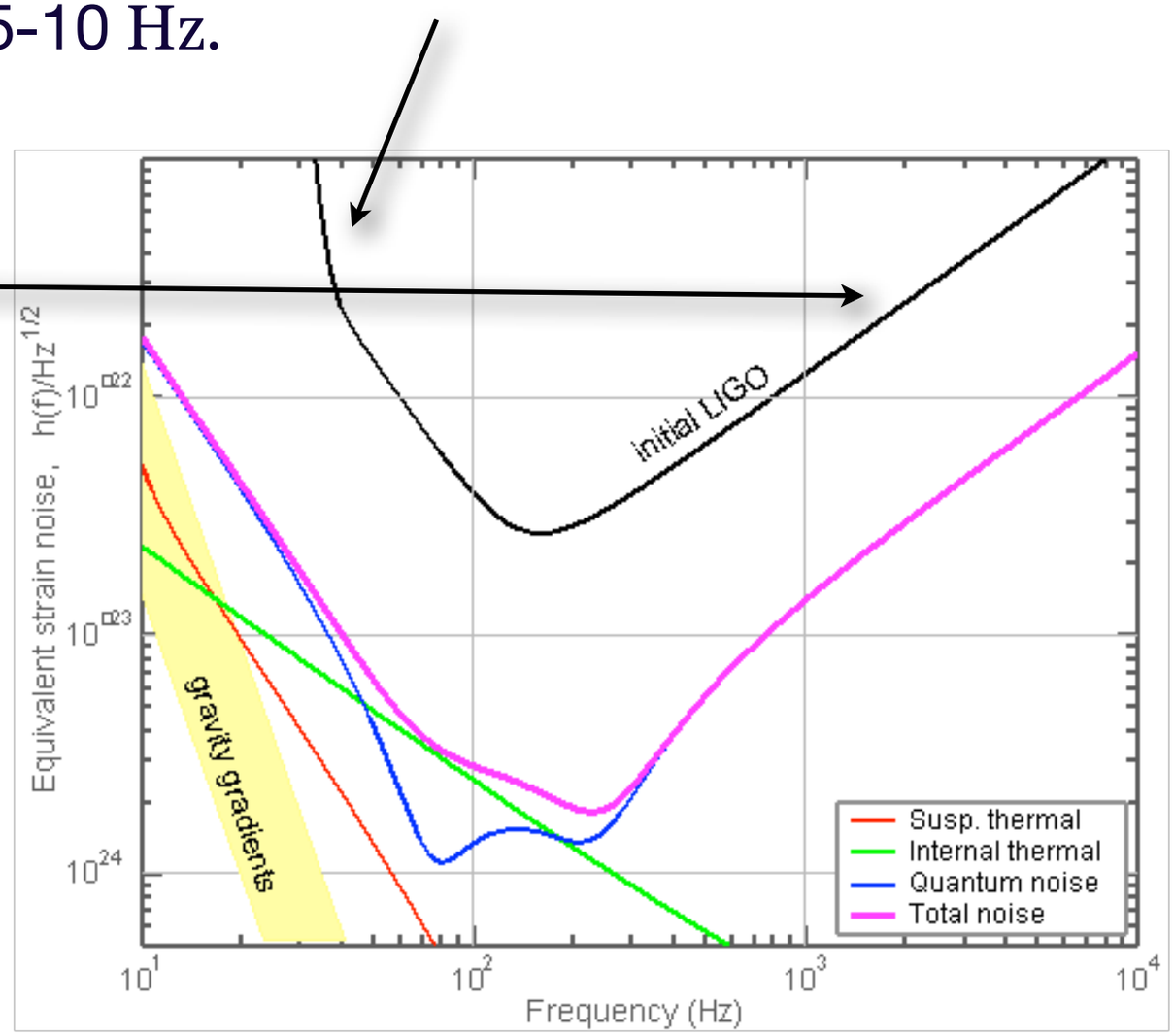
- Suspended mirrors play the role of “test-particles”, placed in perpendicular directions. The light is reflected on the mirrors and returns back to the beam splitter and then to a photodetector where the fringe pattern is monitored.



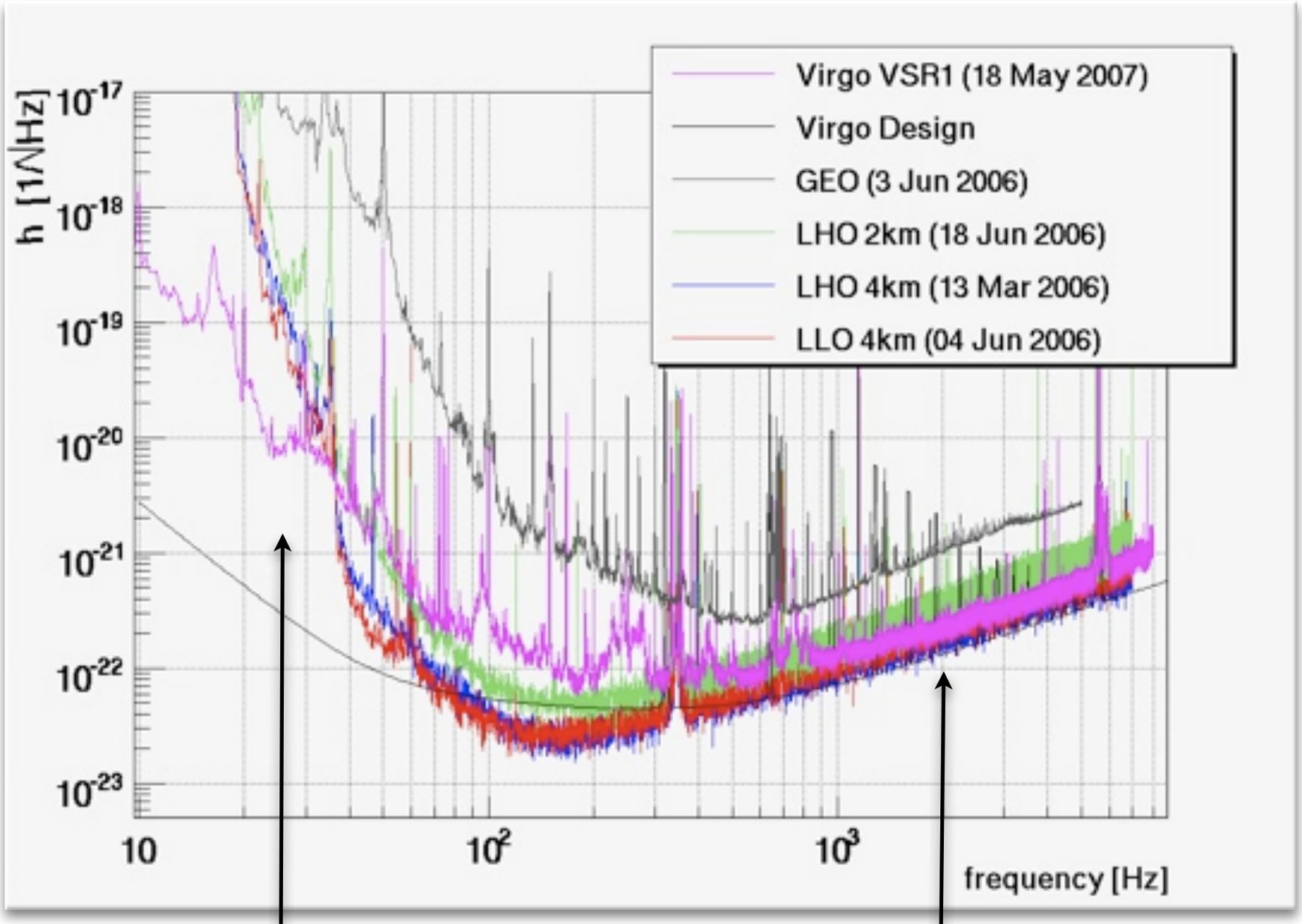
Noise in interferometric detectors

- Seismic noise (low frequencies). At frequencies below 60 Hz, the noise in the interferometers is dominated by seismic noise. The vibrations of the ground couple to the mirrors via the wire suspensions which support them. This effect is strongly suppressed by properly designed suspension systems. Still, seismic noise is very difficult to eliminate at frequencies below 5-10 Hz.

- Photon shot noise (high frequencies). The precision of the measurements is restricted by fluctuations in the fringe pattern due to fluctuations in the number of detected photons. The number of detected photons is proportional to the intensity of the laser beam. Statistical fluctuations in the number of detected photons imply an uncertainty in the measurement of the arm length.



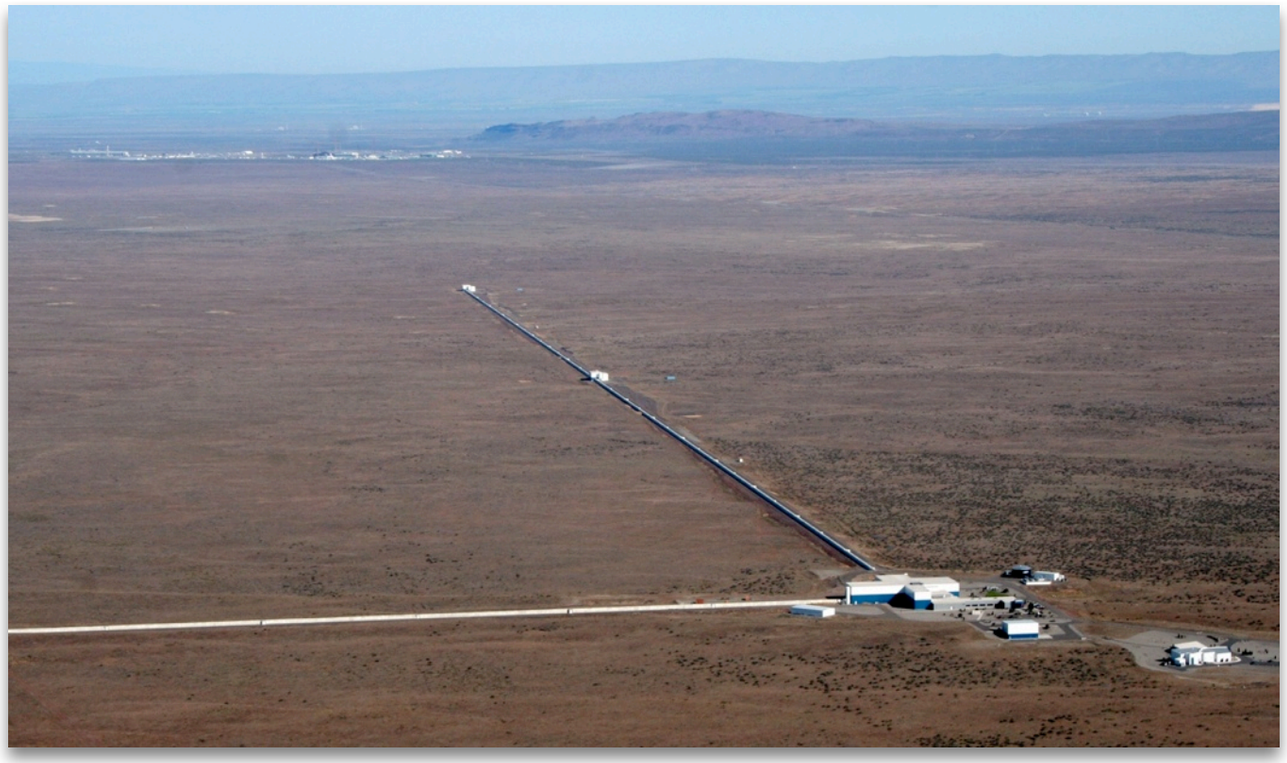
Detectors: real-life sensitivity



Seismic noise

laser photon noise

Detectors: the present (I)



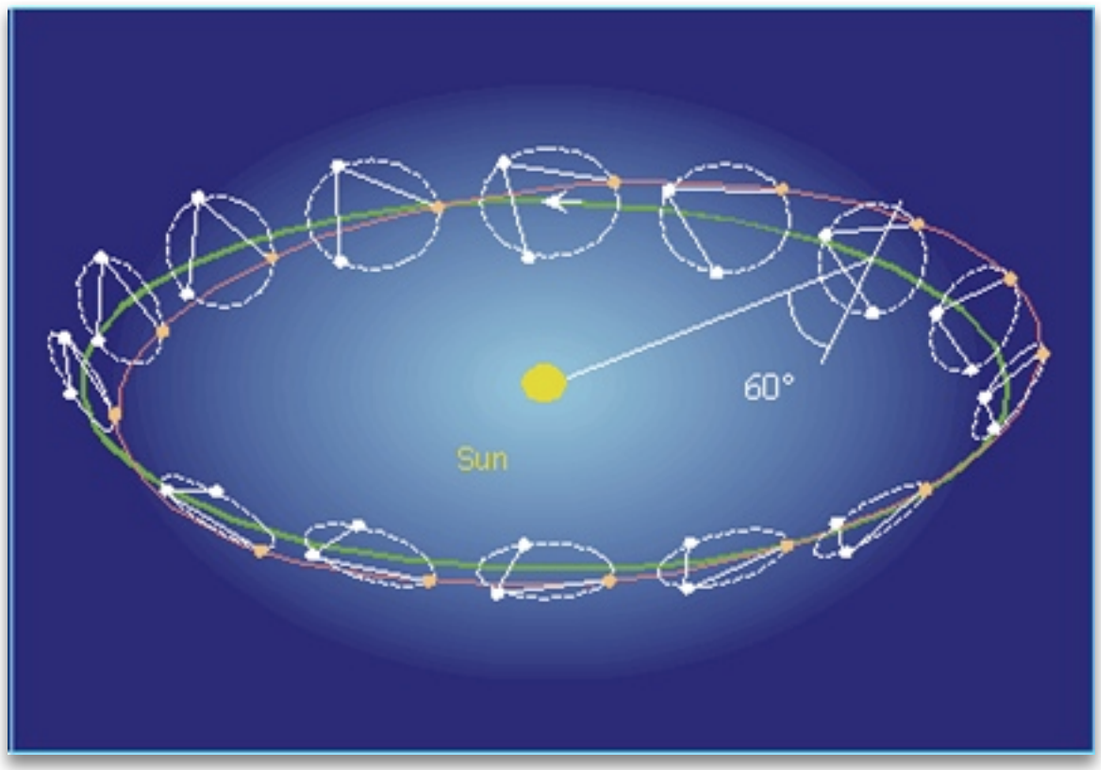
The twin LIGO detectors (L = 4 km) at Livingston Louisiana and Hanford Washington (US).

Detectors: the present (II)

The VIRGO detector (L= 3 km) near Pisa, Italy

Going to space: the LISA detector

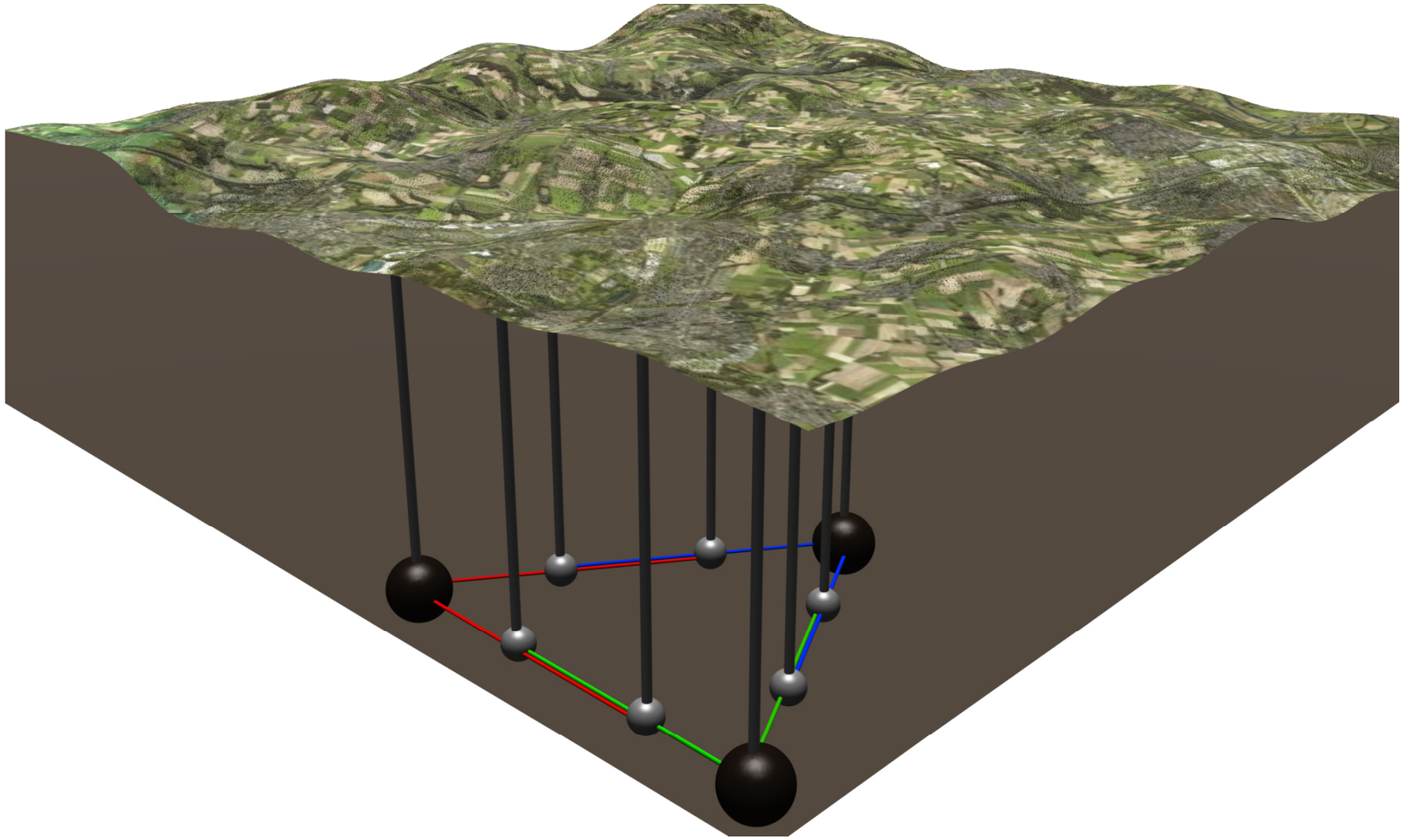
- Space-based detectors: “noise-free” environment, abundance of space!
- Long-arm baseline, low frequency sensitivity
- LISA: Up until recently a joint NASA/ESA mission, now an ESA mission only. To be launched around 2020.



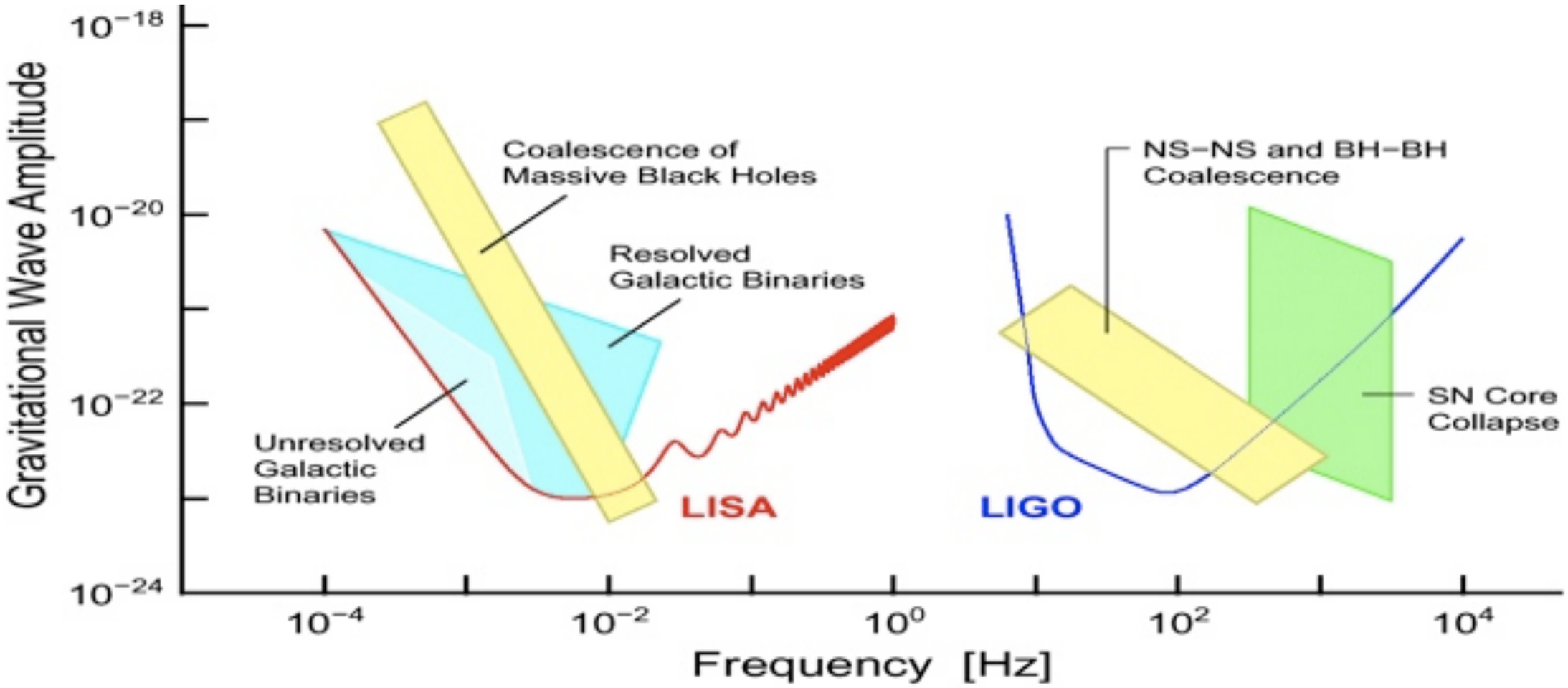
LISA (NASA-ESA)

Going underground: the ET

- The Einstein Telescope will be the next generation underground detector.



GWs detectors: ground and space



GW-related research in Spain

- Barcelona: binary BH systems, GW data analysis.

- Valencia: numerical GR/hydrodynamics, neutron star physics.

- Palma de Mallorca: GW data analysis, binary BHs.

- Alicante: neutron star physics.

Preprint typeset in JHEP style - PAPER VERSION

GRAVITATIONAL WAVE PHYSICS

Kostas D. Kokkotas *

Department of Physics, Aristotle University of Thessaloniki
Thessaloniki 54124, Greece.

Abstract: Gravitational waves are propagating fluctuations of gravitational fields, that is, "ripples" in spacetime, generated mainly by moving massive bodies. These distortions of spacetime travel with the speed of light. Every body in the path of such a wave feels a tidal gravitational force that acts perpendicular to the waves direction of propagation; these forces change the distance between points, and the size of the changes is proportional to the distance between the points. Gravitational waves can be detected by devices which measure the induced length changes. The frequencies and the amplitudes of the waves are related to the motion of the masses involved. Thus, the analysis of gravitational waveforms allows us to learn about their source and, if there are more than two detectors involved in observation, to estimate the distance and position of their source on the sky.

*Article for the Encyclopedia of Physical Science and Technology, 3rd Edition, Volume 7 Academic Press, (2002)

Contents

1. INTRODUCTION **1**
- 1.1 Why are gravitational waves interesting? 2
- 1.2 How we will detect them? 2

2. THEORY OF GRAVITATIONAL WAVES **3**
- 2.1 Linearized theory 5
- 2.2 Properties of gravitational waves 11
- 2.3 Energy flux carried by gravitational waves 12
- 2.4 Generation of gravitational waves 13
- 2.5 Rotating binary system 16

3. DETECTION OF GRAVITATIONAL WAVES **18**
- 3.1 Resonant detectors 19
- 3.2 Beam detectors 22
  - 3.2.1 Laser Interferometers 22
  - 3.2.2 Space Detectors 26
  - 3.2.3 Satellite Tracking 27
  - 3.2.4 Pulsar Timing 28

4. ASTRONOMICAL SOURCES OF GRAV. WAVES **28**
- 4.1 Radiation from gravitational collapse 29
- 4.2 Radiation from binary systems 30
- 4.3 Radiation from spinning neutron stars 31
- 4.4 Cosmological gravitational waves 31

1. INTRODUCTION

Einstein first postulated the existence of gravitational waves in 1916 as a consequence of his theory of General Relativity, but no direct detection of such waves has been made yet. The best evidence thus far for their existence is due to the work of 1993 Nobel laureates Joseph Taylor and Russell Hulse. They observed, in 1974, two neutron stars orbiting faster and faster around each other, exactly what would be expected if the binary neutron star was losing energy in the form of emitted gravitational waves. The predicted rate of orbital acceleration caused by gravitational

radiation emission according to general relativity was verified observationally, with high precision.

Cosmic gravitational waves, upon arriving on earth, are much weaker than the corresponding electromagnetic waves. The reason is that strong gravitational waves are emitted by *very* massive compact sources undergoing *very* violent dynamics. These kinds of sources are not very common and so the corresponding gravitational waves come from large astronomical distances. On the other hand, the waves thus produced propagate essentially unscathed through space, without being scattered or absorbed from intervening matter.

1.1 Why are gravitational waves interesting?

Detection of gravitational waves is important for two reasons: First, their detection is expected to open up a new window for observational astronomy since the information carried by gravitational waves is very different from that carried by electromagnetic waves. This new window onto universe will complement our view of the cosmos and will help us unveil the fabric of spacetime around black-holes, observe directly the formation of black holes or the merging of binary systems consisting of black holes or neutron stars, search for rapidly spinning neutron stars, dig deep into the very early moments of the origin of the universe, and look at the very center of the galaxies where supermassive black holes weighting millions of solar masses are hidden. These are only a few of the great scientific discoveries that scientists will witness during the first decade of the 21st century. Second, detecting gravitational waves is important for our understanding of the fundamental laws of physics; the proof that gravitational waves exist will verify a fundamental 85-year-old prediction of general relativity. Also, by comparing the arrival times of light and gravitational waves, from, e.g., supernovae, Einstein's prediction that light and gravitational waves travel at the same speed could be checked. Finally, we could verify that they have the polarization predicted by general relativity.

1.2 How we will detect them?

Up to now, the only indication of the existence of gravitational waves is the indirect evidence that the orbital energy in the Hulse-Taylor binary pulsar is drained away at a rate consistent with the prediction of general relativity. The gravitational wave is a signal, the shape of which depends upon the changes in the gravitational field of its source. As it has been mentioned earlier, any body in the path of the wave will feel an oscillating tidal gravitational force that acts in a plane perpendicular to the waves direction of propagation. This means that a group of freely moving masses, placed on a plane perpendicular to the direction of propagation of the wave, will oscillate as long as the wave passes through them, and the distance between them will vary as a function of time as in Figure 1. Thus, the detection of gravitational waves can be accomplished by monitoring the tiny changes in the distance between

freely moving test masses. These changes are extremely small; for example, when the Hulse-Taylor binary system finally merges, the strong gravitational wave signal that will be emitted will induce changes in the distance of two particles on earth, that are 1 km apart much smaller than the diameter of the atomic nucleus! To measure such motions of macroscopic objects is a tremendous challenge for experimentalists. As early as the mid-1960s, Joseph Weber designed and constructed heavy metal bars, seismically isolated, to which a set of piezoelectric strain transducers were bonded in such a way that they could detect vibrations of the bar if it had been excited by a gravitational wave. Today, there are a number of such apparatuses operating around the world which have achieved unprecedented sensitivities, but they still are not sensitive enough to detect gravitational waves. Another form of gravitational wave detector that is more promising uses laser beams to measure the distance between two well-separated masses. Such devices are basically kilometer sized laser interferometers consisting of three masses placed in an L-shaped configuration. The laser beams are reflected back and forth between the mirrors attached to the three masses, the mirrors lying several kilometers away from each other. A gravitational wave passing by will cause the lengths of the two arms to oscillate with time. When one arm contracts, the other expands, and this pattern alternates. The result is that the interference pattern of the two laser beams changes with time. With this technique, higher sensitivities could be achieved than are possible with the bar detectors. It is expected that laser interferometric detectors are the ones that will provide us with the first direct detection of gravitational waves.

2. THEORY OF GRAVITATIONAL WAVES

Newton's theory of gravity has enjoyed great success in describing many aspects of our every-day life and additionally explains most of the motions of celestial bodies in the universe. General relativity corrected Newton's theory and is recognized as one of the most ingenious creations of the human mind. The laws of general relativity, though, in the case of slowly moving bodies and weak gravitational fields reduce to the standard laws of Newtonian theory. Nevertheless, general relativity is conceptually different from Newton's theory as it introduces the notion of spacetime and its geometry. One of the basic differences of the two theories concerns the speed of propagation of any change in a gravitational field. As the apple falls from the tree, we have a rearrangement of the distribution of mass of the earth, the gravitational field changes, and a distant observer with a high-precision instrument will detect this change. According to Newton, the changes of the field are instantaneous, i.e., they propagate with infinite speed; if this were true, however the principle of causality would break down. No information can travel faster than the speed of light. In Einstein's theory there is no such ambiguity; the information of the varying gravitational field propagates with finite speed, the speed of light, as a ripple in the

Figure 1: The effects of a gravitational wave travelling perpendicular the plane of a circular ring of particles, is sketched as a series of snapshots. The deformations due the two polarizations are shown.

fabric of spacetime. These are the gravitational waves. The existence of gravitational waves is an immediate consequence of any relativistic theory of gravity. However, the strength and the form of the waves depend on the details of the gravitational theory. This means that the detection of gravitational waves will also serve as a test of basic gravitational theory.

The fundamental geometrical framework of relativistic metric theories of gravity is spacetime, which mathematically can be described as a four-dimensional manifold whose points are called *events*. Every event is labeled by four coordinates x^μ ($\mu = 0, 1, 2, 3$); the three coordinates x^i ($i = 1, 2, 3$) give the spatial position of the event, while x^0 is related to the coordinate time t ($x^0 = ct$, where c is the speed of light, which unless otherwise stated will be set equal to 1). The choice of the coordinate system is quite arbitrary and coordinate transformations of the form $\tilde{x}^\mu = f^\mu(x^\lambda)$ are allowed. The motion of a test particle is described by a curve in spacetime. The distance ds between two neighboring events, one with coordinates x^μ and the other with coordinates $x^\mu + dx^\mu$, can be expressed as a function of the coordinates via a symmetric tensor $g_{\mu\nu}(x^\lambda) = g_{\nu\mu}(x^\lambda)$, i.e.,

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu \tag{2.1}$$

This is a generalization of the standard measure of distance between two points in Euclidian space. For the Minkowski spacetime (the spacetime of special relativity),

$g_{\mu\nu} \equiv \eta_{\mu\nu} = \text{diag}(-1, 1, 1, 1)$. The symmetric tensor is called the *metric tensor* or simply the metric of the spacetime. In general relativity the gravitational field is described by the metric tensor alone, but in many other theories one or more supplementary fields may be needed as well. In what follows, we will consider only the general relativistic description of gravitational fields, since most of the alternative theories fail to pass the experimental tests.

The information about the degree of curvature (i.e., the deviation from flatness) of a spacetime is encoded in the metric of the spacetime. According to general relativity, any distribution of mass bends the spacetime fabric and the Riemann tensor $R_{\kappa\lambda\mu\nu}$ (that is a function of the metric tensor $g_{\mu\nu}$ and of its first and second derivatives) is a measure of the spacetime curvature. The Riemann tensor has 20 independent components. When it vanishes the corresponding spacetime is flat.

In the following presentation, we will consider mass distributions, which we will describe by the stress-energy tensor $T^{\mu\nu}(x^\lambda)$. For a perfect fluid (a fluid or gas with isotropic pressure but without viscosity or shear stresses) the stress-energy tensor is given by the following expression

$$T^{\mu\nu}(x^\lambda) = (\rho + p)u^\mu u^\nu + pg^{\mu\nu}, \tag{2.2}$$

where $p(x^\lambda)$ is the local pressure, $\rho(x^\lambda)$ is the local energy density and $u^\mu(x^\lambda)$ is the four velocity of the infinitesimal fluid element characterized by the event x^λ.

Einstein's gravitational field equations connect the curvature tensor (see below) and the stress-energy tensor through the fundamental relation

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = kT_{\mu\nu} \tag{2.3}$$

This means that the gravitational field, which is directly connected to the geometry of spacetime, is related to the distribution of matter and radiation in the universe. By solving the field equations, both the gravitational field (the $g_{\mu\nu}$) and the motion of matter is determined. $R_{\mu\nu}$ is the so- called Ricci tensor and comes from a contraction of the Riemann tensor ($R_{\mu\nu} = g^{\alpha\sigma}R_{\sigma\mu\rho\nu}$), R is the scalar curvature ($R = g^{\rho\sigma}R_{\rho\sigma}$), while $G_{\mu\nu}$ is the so-called Einstein tensor, $k = 8\pi G/c^4$ is the coupling constant of the theory and is the gravitational constant, which, unless otherwise stated will be considered equal to 1. The vanishing of the Ricci tensor corresponds to a spacetime free of any matter distribution. However, this does not imply that the Riemann tensor is zero. As a consequence, in the empty space far from any matter distribution, the Ricci tensor will vanish while the Riemann tensor can be nonzero; this means that the effects of a propagating gravitational wave in an empty spacetime will be described via the Riemann tensor.

2.1 Linearized theory

Now lets assume that an observer is far away from a given static matter distribution, and the spacetime in which he or she lives is described by a metric $g_{\mu\nu}$. Any change

in the matter distribution, i.e., in $T_{\mu\nu}$, will induce a change in the gravitational field, which will be recorded as a change in metric. The new metric will be

$$\tilde{g}_{\mu\nu} = g_{\mu\nu} + h_{\mu\nu} \tag{2.4}$$

where $h_{\mu\nu}$ is a tensor describing the variations induced in the spacetime metric. As we will describe analytically later, this new tensor describes the propagation of ripples in spacetime curvature, i.e., *the gravitational waves*. In order to calculate the new tensor we have to solve Einsteins equations for the varying matter distribution. This is not an easy task in general. However, there is a convenient, yet powerful, way to proceed, namely to assume that $h_{\mu\nu}$ is small ($|h_{\mu\nu}| \ll 1$), so that we need only keep terms linear in $h_{\mu\nu}$ in our calculations. In making this approximation we are effectively assuming that the disturbances produced in spacetime are not huge. This *linearization approach* has proved extremely useful for calculations, and for weak fields at least, gives accurate results for the generation of the waves and for their propagation.

The first attempt to prove that in general relativity gravitational perturbations propagate as waves with the speed of light is due to Einstein himself. Shortly after the formulation of his theory - the year after - he proved that by assuming linearized perturbations around a flat metric, i.e., $g_{\mu\nu} = \eta_{\mu\nu}$, then the tensor

$$\tilde{h}_{\mu\nu} \equiv h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h^{\alpha}_{\alpha} \tag{2.5}$$

is governed by a wave equation, which admits plane wave solutions similar to the ones of electromagnetism; here $h_{\mu\nu}$ is the metric perturbation and $\tilde{h}_{\mu\nu}$ is the gravitational field (or the trace reverse of $h_{\mu\nu}$). Then the linear field equations in vacuum have the form

$$\left(-\frac{\partial^2}{\partial t^2} + \nabla^2\right)\tilde{h}^{\mu\nu} \equiv \partial_\lambda \partial^\lambda \tilde{h}^{\mu\nu} = 0 \tag{2.6}$$

($\partial_\mu k^\alpha \equiv \partial k^\alpha / \partial x^\mu$), which is the three-dimensional wave equation. To obtain the above simplified form, the condition $\partial_\mu \tilde{h}^{\mu\nu} = 0$, known as *Hilbert's gauge condition* (equivalent to the Lorentz gauge condition of electromagnetism), has been assumed. A gauge transformation is a suitable change of coordinates defined by

$$x'^{\mu} \equiv x^{\mu} + \xi^{\mu} \tag{2.7}$$

which induces a redefinition of the gravitational field tensor

$$\tilde{h}'_{\mu\nu} \equiv \tilde{h}_{\mu\nu} - \partial_\mu \xi_\nu - \partial_\nu \xi_\mu + \eta_{\mu\nu}\partial_\lambda \xi^\lambda. \tag{2.8}$$

It can be easily proved that ξ^μ must satisfy the condition

$$\partial_\nu \partial^\nu \xi^\mu = 0, \tag{2.9}$$

so that the new gravitational field is in agreement with the Hilberts gauge condition. The solution of equation (2.9) defines the four components of ξ^μ so that the new tensor $\tilde{h}'_{\mu\nu}$ is also a solution of the wave equation (2.6), and thus, have the same physical meaning as $\tilde{h}_{\mu\nu}$. In general, gauge transformations correspond to symmetries of the field equations, which means that the field equations are invariant under such transformations. This implies that the field equations do not determine the field uniquely; however, this ambiguity in determining the field is devoid of any physical meaning.

The simplest solution to the wave equation (2.6) is a plane wave solution of the form

$$\tilde{h}^{\mu\nu} = A^{\mu\nu} e^{ik_\alpha x^\alpha}, \tag{2.10}$$

where $A^{\mu\nu}$ is a constant symmetric tensor, the *polarization tensor*, in which information about the amplitude and the polarization of the waves is encoded, while k_α is a constant vector, the *wave vector*, that determines the propagation direction of the wave and its frequency. In physical applications we will use only the real part of the above wave solution. By applying the Hilbert gauge condition on the plane wave solution we obtain the relation $A^{\mu\nu}k_\mu = 0$, the geometrical meaning of which is that $A^{\mu\nu}$ and k_μ are orthogonal. This relation can be written as four equations that impose four conditions on $A^{\mu\nu}$, and this is the first step in reducing the number of its independent components. As a consequence, $A^{\mu\nu}$, instead of having 10 independent components (as has every symmetric second rank tensor in a four dimensional space), has only 6 independent ones. Further substitution of the plane wave solution in the wave equation leads to the important equation $k_\mu k^\mu = k_0^2 - (k_x^2 + k_y^2 + k_z^2) = 0$, which means that k_μ is a *lightlike* or *null vector*, i.e., the wave propagates on the light-cone. This means that the speed of the wave is 1, i.e., equal to the speed of light. The frequency of the wave is $\omega = k^0$.

Up to this point it has been proven that $A^{\mu\nu}$ has six arbitrary components, but due to the gauge freedom, i.e., the freedom in choosing the four components of the vector ξ^μ, the actual number of its independent components can be reduced to **two**, in a suitable chosen gauge. The *transverse–traceless* or *TT gauge* is an example of such a gauge. In this gauge, only the spatial components of $\tilde{h}_{\mu\nu}$ are non-zero (hence $\tilde{h}_{\mu 0} = 0$), which means that the wave is *transverse* to its own direction of propagation, and, additionally, the sum of the diagonal components is zero ($\tilde{h}^\mu_\mu = h^0_0 + h^1_1 + h^2_2 + h^3_3 \equiv \tilde{h} = h = 0$) (*traceless*). Due to this last property and equation (5), in this gauge there is no difference between $h_{\mu\nu}$(the perturbation of the metric) and $\tilde{h}_{\mu\nu}$ (the gravitational field). It is customary to write the gravitational wave solution in the TT gauge as $h^{TT}_{\mu\nu}$. That $A_{\mu\nu}$ has only two independent components means that a gravitational wave is completely described by two dimensionless amplitudes, h_+ and $h_\times$, say. If, for example, we assume a wave propagating along the z-direction,

then the amplitude $A^{\mu\nu}$ can be written as

$$A^{\mu\nu} = h_+ \epsilon_+^{\mu\nu} + h_\times \epsilon_\times^{\mu\nu} \tag{2.11}$$

where $\epsilon_+^{\mu\nu}$ and $\epsilon_\times^{\mu\nu}$ are the so-called *unit polarization tensors* defined by

$$\epsilon_+^{\mu\nu} \equiv \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad \epsilon_\times^{\mu\nu} \equiv \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{2.12}$$

As mentioned earlier, the Riemann tensor is a measure of the curvature of spacetime. A gravitational wave, propagating in a flat spacetime, generates periodic distortions, which can be described in terms of the Riemann tensor. In linearized theory the Riemann tensor takes the following gauge-independent form:

$$R_{\kappa\lambda\mu\nu} = \frac{1}{2}\left(\partial_{\nu\kappa} h_{\lambda\mu} + \partial_{\lambda\mu} h_{\kappa\nu} - \partial_{\kappa\mu} h_{\lambda\nu} - \partial_{\lambda\nu} h_{\kappa\mu}\right), \tag{2.13}$$

which is considerably simplified by choosing the TT gauge:

$$R^{TT}_{j0k0} = -\frac{1}{2}\frac{\partial^2}{\partial t^2} h^{TT}_{jk}, \quad j,k = 1,2,3. \tag{2.14}$$

Furthermore, in the Newtonian limit

$$R^{TT}_{j0k0} \approx \frac{\partial^2 \Phi}{\partial x^j \partial x^k}, \tag{2.15}$$

where Φ describes the gravitational potential in Newtonian theory. Earlier we defined the Riemann tensor as a geometrical object, but this tensor has a simple physical interpretation: it is the tidal force field and describes the relative acceleration between two particles in free fall. If we assume two particles moving freely along geodesics of a curved spacetime with coordinates $x^\mu(\tau)$ and $x^\mu(\tau) + \xi^\mu(\tau)$ (for a given value of the proper time τ, $\xi^\mu(\tau)$ is the displacement vector connecting the two events) it can be shown that, in the case of slowly moving particles,

$$\frac{d^2 \xi^k}{dt^2} \approx -R^k_{0j0}{}^{TT} \xi^j. \tag{2.16}$$

This is a simplified form of the equation of *geodesic deviation*. Hence, the tidal force acting on a particle is:

$$f^k \approx -m R^k_{0j0} \xi^j, \tag{2.17}$$

where m is the mass of the particle. Equation (2.17) corresponds to the standard Newtonian relation for the tidal force acting on a particle in a field Φ.

Keeping this in mind, we will try to visualize the effect of a gravitational wave. Let us first consider two freely falling particles hit by a gravitational wave travelling along the z-direction, with the (+) polarization present only, i.e.,

$$h^{\mu\nu} = h_+ \epsilon^{\mu\nu} \cos[\omega(t-z)]. \tag{2.18}$$

Then, the measured distance ξ^x between the two particles, originally at a distance ξ_0^x, along the x direction, will be

$$\frac{\xi^x}{\xi_0^x} = 1 - \frac{1}{2} h_+ \cos[\omega(t-z)] \quad \text{or} \quad \delta\xi^x = \xi^x - \xi_0^x = -\frac{1}{2} h_+ \cos[\omega(t-z)]\xi_0^x \tag{2.19}$$

which implies that the relative distance $\delta\xi^x$ between the two particles will oscillate with frequency ω. This does not mean that the particles coordinate positions change; instead they remain at rest relative to the coordinates, but the coordinate distance oscillates. If the particles were placed originally along the y-direction the coordinate distance would oscillate according to

$$\frac{\xi^y}{\xi_0^y} = 1 + \frac{1}{2} h_+ \cos[\omega(t-z)] \quad \text{or} \quad \delta\xi^y = \xi^y - \xi_0^y = \frac{1}{2} h_+ \cos[\omega(t-z)]\xi_0^y \tag{2.20}$$

In other words, the coordinate distances along the two axes oscillate out of phase, that is, when the distance between two particles along the x direction is maximum the distance of two other particles along the y direction is minimum, and after half a period, it is the other way around. The effects are similar for the other polarization, where the axes along which the oscillations are out of phase are at an angle of 45^o with respect to the first ones. This can be visualized in Figure 1, where the effect of a passing gravitational wave on a ring of particles is shown as a series of snapshots closely separated in time.

Another way of understanding the effects of gravitational waves is to study the tidal force field lines. In the TT gauge the equation of the geodesic deviation (2.16) takes the simple form

$$\frac{d^2\xi^k}{dt^2} \approx \frac{1}{2} \frac{d^2 h_{jk}^{TT}}{dt^2} \xi^j \tag{2.21}$$

and the corresponding tidal force is

$$f^k \approx \frac{m}{2} \frac{d^2 h_{jk}^{TT}}{dt^2} \xi^j \tag{2.22}$$

For the wave given by equation (2.18) the two nonzero components of the tidal force are

$$f^x \approx \frac{m}{2} h_+ \omega^2 \cos[\omega(t-z)]\xi_0^x, \quad \text{and} \quad f^y \approx -\frac{m}{2} h_+ \omega^2 \cos[\omega(t-z)]\xi_0^y. \tag{2.23}$$

It can be easily proved that the divergence of the tidal force is zero ($\partial f^x/\partial\xi_0^x + \partial f^y/\partial\xi_0^y = 0$). It can therefore be represented graphically by field lines as in Figure 2.

Figure 2: The tidal field lines of force for a gravitational wave with polarization (+) (left panel) and (×) (right panel). The orientation of the field lines changes every half period producing the deformations as seen in Figure 1. Any point accelerates in the directions of the arrows, and the denser are the lines, the strongest is the acceleration. Since the acceleration is proportional to the distance from the center of mass, the force lines get denser as one moves away from the origin. For the polarization (×) the force lines undergo a 45^0 rotation.

Let us now return back to the two polarization states represented by the two matrices $\epsilon_+^{\mu\nu}$ and $\epsilon_\times^{\mu\nu}$. It is impossible to construct the (+) pattern from the (×) pattern and vice versa; they are orthogonal polarization states. By analogy with electromagnetic waves, the two polarizations could be added with phase difference $(\pm\pi/2)$ to obtain circularly polarized waves. The effect of circularly polarized waves on a ring of particles is to deform the ring into a rotating ellipse with either positive or negative helicity. The particles themselves do not rotate; they only oscillate in and out around their initial positions. These circularly polarized waves carry angular momentum the amount of which is $(2/\omega)$ times the energy carried by the wave. If we consider the gravitational field, then, according to quantum field theory, the waves are associated with fundamental particles responsible for the gravitational interaction, and a quantum of the field will have energy $\hbar\omega$ and consequently spin $2\hbar$. This means that the quanta of the gravitational field, the *gravitons*, are spin-2, massless particles (since they travel with the speed of light). Another way of explaining why a graviton should be a spin-2 particle comes from observing Figure 1. One can see that a gravitational wave is invariant under rotations of 180^0 about its direction of propagation; the pattern repeats itself after half period. For comparison, we mention that electromagnetic waves are invariant under rotations of 360^0. In the quantum mechanical description of massless particles, the wavefunction of a particle

is invariant under rotations of $360^0/s$, where s is the spin of the particle. Thus the photon is a spin-1 particle and the graviton is a spin-2 particle. In other relativistic theories of gravity, the wave field has other symmetries and therefore they attribute different spins to the gravitons.

2.2 Properties of gravitational waves

Gravitational waves, once they are generated, propagate almost unimpeded. Indeed, it has been proven that they are even harder to stop than neutrinos! The only significant change they suffer as they propagate is the decrease in amplitude while they travel away from their source, and the *redshift* they feel (cosmological, gravitational or Doppler), as is the case for electromagnetic waves.

There are other effects that marginally influence the gravitational waveforms, for instance, *absorption* by interstellar or intergalactic matter intervening between the observer and the source, which is extremely weak (actually, the extremely weak coupling of gravitational waves with matter is the main reason that gravitational waves have not been observed). *Scattering* and *dispersion* of gravitational waves are also practically unimportant, although they may have been important during the early phases of the universe (this is also true for the absorption). Gravitational waves can be *focused* by strong gravitational fields and also can be *diffracted*, exactly as it happens with the electromagnetic waves.

There are also a number of "exotic" effects that gravitational waves can experience, that are due to the nonlinear nature of Einsteins equations (purely general-relativistic effects), such as: *scattering* by the background curvature, the existence of *tails* of the waves that interact with the waves themselves, *parametric amplification* by the background curvature, *nonlinear coupling of the waves with themselves* (creation of *geons*, that is, bundles of gravitational waves held together by their own self- generated curvature) and even *formation of singularities* by colliding waves (for such exotic phenomena see the extensive review by Thorne[4]). These aspects of nonlinearity affect the majority of the gravitational wave sources and from this point of view our understanding of gravitational-wave generation is based on approximations. However, it is expected that the error in these approximations for most of the processes that generate gravitational waves is quite small. Powerful numerical codes, using state-of-the-art computer software and hardware, have been developed (and continue to be developed) for minimizing all possible sources of error in order to have as accurate as possible an understanding of the processes that generate gravitational waves and of the waveforms produced.

For most of the properties mentioned above there is a correspondence with electromagnetic waves. Gravitational waves are fundamentally different, however, even though they share similar wave properties away from the source. Gravitational waves are emitted by coherent bulk motions of matter (for example, by the implosion of the core of a star during a supernova explosion) or by coherent oscillations of spacetime

curvature, and thus they serve as a probe of such phenomena. By contrast, cosmic electromagnetic waves are mainly the result of incoherent radiation by individual atoms or charged particles. As a consequence, from the cosmic electromagnetic radiation we mainly learn about the form of matter in various regions of the universe, especially about its temperature and density, or about the existence of magnetic fields. Strong gravitational waves, are emitted from regions of spacetime where gravity is very strong and the velocities of the bulk motions of matter are near the speed of light. Since most of the time these areas are either surrounded by thick layers of matter that absorb electromagnetic radiation or they do not emit any electromagnetic radiation at all (black holes), the only way to study these regions of the universe is via gravitational waves.

2.3 Energy flux carried by gravitational waves

Gravitational waves carry energy and cause a deformation of spacetime. The stress-energy carried by gravitational waves cannot be localized within a wavelength. Instead, one can say that a certain amount of stress-energy is contained in a region of the space which extends over several wavelengths. It can be proven that in the TT gauge of linearized theory the stress-energy tensor of a gravitational wave (in analogy with the stress-energy tensor of a perfect fluid that we have defined earlier) is given by

$$t_{\mu\nu}^{GW} = \frac{1}{32\pi}\langle\left(\partial_\mu h_{ij}^{TT}\right)\left(\partial_\nu h_{ij}^{TT}\right)\rangle. \tag{2.24}$$

where the angular brackets are used to indicate averaging over several wavelengths. For the special case of a plane wave propagating in the z direction, which we considered earlier, the stress-energy tensor has only three non-zero components, which take the simple form

$$t_{00}^{GW} = \frac{t_{zz}^{GW}}{c^2} = -\frac{t_{0z}^{GW}}{c} = \frac{1}{32\pi}\frac{c^2}{G}\omega^2\left(h_+^2 + h_\times^2\right), \tag{2.25}$$

where t_{00}^{GW} is the energy density, t_{zz}^{GW} is the momentum flux and t_{0z}^{GW} the energy flow along the z direction per unit area and unit time (for practical reasons we have restored the normal units). The energy flux has all the properties one would anticipate by analogy with electromagnetic waves: (a) it is conserved (the amplitude dies out as $1/r$, the flux as $1/r^2$), (b) it can be absorbed by detectors, and (c) it can generate curvature like any other energy source in Einstein's formulation of relativity. As an example, by using the above relation, we will estimate the energy flux in gravitational waves from the collapse of the core of a supernova to create a 10 $M_\odot$ black hole at a distance of 50-million-light-years ($\sim$15 Mpc) from the earth (at the distance of the Virgo cluster of galaxies). A conservative estimate of the amplitude of the waves on earth (as we will show later) is of the order of 10^{-22} (at a frequency of about 1kHz). This corresponds to a flux of about 3 ergs/cm^2 sec. This

is an enormous amount of energy flux and is about ten orders of magnitude larger than the observed energy flux in electromagnetic waves! The basic difference is the duration of the two signals; gravitational wave signal will last a few milliseconds, whereas an electromagnetic signal lasts many days. This example provides us with a useful numerical formula for the energy flux:

$$F = 3\left(\frac{f}{1\text{kHz}}\right)^2\left(\frac{h}{10^{-22}}\right)^2\frac{\text{ergs}}{\text{cm}^2\text{sec}}, \tag{2.26}$$

from which one can easily estimate the flux on earth, given the amplitude (on earth) and the frequency of the waves.

2.4 Generation of gravitational waves

As we have mentioned earlier, when the gravitational field is strong there are a number on nonlinear effects that influence the generation and propagation of gravitational waves. For example, nonlinear effects are significant during the last phases of black hole formation. The analytic description of such a dynamically changing spacetime is impossible, and until numerical relativity provides us with accurate estimates of the dynamics of gravitational fields under such extreme conditions we have to be content with order of magnitude estimates. Furthermore, there are differences in the predictions of various relativistic theories of gravity in the case of high concentrations of rapidly varying energy distributions. However, all metric theories of gravity, as long as they admit the correct Newtonian limit, make similar predictions for the total amount of gravitational radiation emitted by "weak" gravitational wave sources, that is, sources where the energy content is small enough to produce only small deformations of the flat spacetime and where all motions are slow compared to the velocity of light.

Let us now try to understand the nature of gravitational radiation, by starting from the production of electromagnetic radiation. Electromagnetic radiation emitted by slowly varying charge distributions can be decomposed into a series of multipoles, where the amplitude of the 2^ℓ-pole ($\ell = 0, 1, 2, ...$) contains a small factor a^ℓ, with a equal to the ratio of the diameter of the source to the typical wavelength, namely, a number typically much smaller than 1. From this point of view the strongest electromagnetic radiation would be expected for monopolar radiation ($\ell = 0$), but this is completely absent, because the electromagnetic monopole moment is proportional to the total charge, which does not change with time (it is a conserved quantity). Therefore, electromagnetic radiation consists only of $\ell \geq 1$ multipoles, the strongest being the electric dipole radiation ($\ell = 1$), followed by the weaker magnetic dipole and electric quadrupole radiation ($\ell = 2$). One could proceed with a similar analysis for gravitational waves and by following the same arguments show that mass conservation (which is equivalent to charge conservation in electromagnetic theory) will exclude monopole radiation. Also, the rate of change of the mass dipole moment is

proportional to the linear momentum of the system, which is a conserved quantity, and therefore there cannot be any mass dipole radiation in Einsteins relativity theory. The next strongest form of electromagnetic radiation is the magnetic dipole. For the case of gravity, the change of the magnetic dipole is proportional to the angular momentum of the system, which is also a conserved quantity and thus there is no dipolar gravitational radiation of *any* sort. It follows that gravitational radiation is of quadrupolar or higher nature and is directly linked to the quadrupole moment of the mass distribution.

As early as 1918, Einstein derived the quadrupole formula for gravitational radiation. This formula states that the wave amplitude h_{ij} is proportional to the second time derivative of the quadrupole moment of the source:

$$h_{ij} = \frac{2}{r}\frac{G}{c^4}\ddot{Q}_{ij}^{TT}\left(t - \frac{r}{c}\right) \tag{2.27}$$

where

$$Q_{ij}^{TT}(x) = \int \rho \left(x^i x^j - \frac{1}{3}\delta^{ij} r^2\right) d^3x \tag{2.28}$$

is the quadrupole moment in the TT gauge, evaluated at the retarded time $t - r/c$ and ρ is the matter density in a volume element d^3x at the position x^i. This result is quite accurate for all sources, as long as the reduced wavelength $\tilde{\lambda} = \lambda/2\pi$ is much longer than the source size R. It should be pointed out that the above result can be derived via a quite cumbersome calculation in which we solve the wave equation (2.6) with a source term $T_{\mu\nu}$ on the right-hand side. In the course of such a derivation, a number of assumptions must be used. In particular, the observer must be located at a distance $r \gg \tilde{\lambda}$, far greater than the reduced wavelength (in what is called the radiation zone) and $T_{\mu\nu}$ must not change very quickly.

Using the formulae (2.24) and (2.25) for the energy carried by gravitational waves, one can derive the luminosity in gravitational waves as a function of the third-order time derivative of the quadrupole moment tensor. This is the quadrupole formula

$$L_{GW} = -\frac{dE}{dt} = \frac{1}{5}\frac{G}{c^5}\langle \frac{\partial^3 Q_{ij}}{\partial t^3}\ \frac{\partial^3 Q_{ij}}{\partial t^3}\rangle. \tag{2.29}$$

Based on this formula, we derive some additional formulas, which provide order of magnitude estimates for the amplitude of the gravitational waves and the corresponding power output of a source. First, the quadrupole moment of a system is approximately equal to the mass M of the part of the system that moves, times the square of the size R of the system. This means that the third-order time derivative of the quadrupole moment is

$$\frac{\partial^3 Q_{ij}}{\partial t^3} \sim \frac{MR^2}{T^3} \sim \frac{Mv^2}{T} \sim \frac{E_{\rm ns}}{T}, \tag{2.30}$$

where v is the mean velocity of the moving parts, $E_{\rm ns}$ is the kinetic energy of the component of the source's internal motion which is non spherical, and T is the time

scale for a mass to move from one side of the system to the other. The time scale (or period) is actually proportional to the inverse of the square root of the mean density of the system

$$T \sim \sqrt{R^3/GM}. \tag{2.31}$$

This relation provides a rough estimate of the characteristic frequency of the system $f = 2\pi/T$. Then, the luminosity of gravitational waves of a given source is approximately

$$L_{GW} \sim \frac{G}{c^5}\left(\frac{M}{R}\right)^5 \sim \frac{G}{c^5}\left(\frac{M}{R}\right)^2 v^6 \sim \frac{c^5}{G}\left(\frac{R_{\rm Sch}}{R}\right)^2 \left(\frac{v}{c}\right)^6 \tag{2.32}$$

where $R_{\rm Sch} = 2GM/c^2$ is the Schwarzschild radius of the source. It is obvious that the maximum value of the luminosity in gravitational waves can be achieved if the source's dimensions are of the order of its Schwarzschild radius and the typical velocities of the components of the system are of the order of the speed of light. This explains why we expect the best gravitational wave sources to be highly relativistic compact objects. The above formula sets also an upper limit on the power emitted by a source, which for $R \sim R_{\rm Sch}$ and $v \sim c$ is

$$L_{GW} \sim c^5/G = 3.6 \times 10^{59} \text{ergs/sec}. \tag{2.33}$$

This is an immense power, often called the *luminosity of the universe*.

Using the above order-of-magnitude estimates, we can get a rough estimate of the amplitude of gravitational waves at a distance r from the source:

$$h \sim \frac{G}{c^4}\frac{E_{\rm ns}}{r} \sim \frac{G}{c^4}\frac{\varepsilon E_{\rm kin}}{r} \tag{2.34}$$

where $\varepsilon E_{\rm kin}$ (with $0 \leq \varepsilon \leq 1$), is the fraction of kinetic energy of the source that is able to produce gravitational waves. The factor ε is a measure of the asymmetry of the source and implies that only a time varying quadrupole moment will emit gravitational waves. For example, even if a huge amount of kinetic energy is involved in a given explosion and/or implosion, if the event takes place in a spherically symmetric manner, there will be no gravitational radiation.

Another formula for the amplitude of gravitational waves relation can be derived from the flux formula (2.26). If, for example, we consider an event (perhaps a supernovae explosion) at the Virgo cluster during which the energy equivalent of $10^{-4}M_\odot$ is released in gravitational waves at a frequency of 1 kHz, and with signal duration of the order of 1 msec, the amplitude of the gravitational waves on Earth will be

$$h \approx 10^{-22}\left(\frac{E_{GW}}{10^{-4}M_\odot}\right)^{1/2}\left(\frac{f}{1\text{kHz}}\right)^{-1}\left(\frac{\tau}{1\text{msec}}\right)^{-1/2}\left(\frac{r}{15\text{Mpc}}\right)^{-1}. \tag{2.35}$$

For a detector with arm length of 4 km we are looking for changes in the arm length of the order of

$$\Delta\ell = h \cdot \ell = 10^{-22} \cdot 4\text{km} = 4 \times 10^{-17}\text{!!!}$$

These numbers shows why experimenters are trying so hard to build ultra-sensitive detectors and explains why all detection efforts till today were not successful.

Finally, it is useful to know the damping time, that is, the time it takes for a source to transform a fraction $1/e$ of its energy into gravitational radiation. One can obtain a rough estimate from the following formula

$$\tau = \frac{E_{\text{kin}}}{L_{\text{GW}}} \sim \frac{1}{c} R \left(\frac{R}{R_{\text{Sch}}} \right)^3 . \tag{2.36}$$

For example, for a non-radially oscillating neutron star with a mass of roughly $1.4\text{M}_\odot$ and a radius of 12Km, the damping time will be of the order of $\sim$50msec. Also, by using formula (2.31), we get an estimate for the frequency of oscillation which is directly related to the frequency of the emitted gravitational waves, roughly 2kHz for the above case.

2.5 Rotating binary system

Among the most interesting sources of gravitational waves are binaries. The inspiralling of such systems, consisting of black holes or neutron stars, is, as we will discuss later, the most promising source for the gravitational wave detectors. Binary systems are also the sources of gravitational waves whose dynamics we understand the best. If we assume that the two bodies making up the binary lie in the $x-y$ plane and their orbits are circular (see Figure 3), then the only non-vanishing components of the quadrupole tensor are

$$Q_{xx} = -Q_{yy} = \frac{1}{2}\mu a^2 \cos 2\Omega t, \quad \text{and} \quad Q_{xy} = Q_{yx} = \frac{1}{2}\mu a^2 \sin 2\Omega t, \tag{2.37}$$

where Ω is the orbital angular velocity, $\mu = M_1 M_2/M$ is the reduced mass of the system and $M = M_1 + M_2$ its total mass.

According to equation (2.29) the gravitational radiation luminosity of the system is

$$L^{\text{GW}} = \frac{32}{5}\frac{G}{c^5}\mu^2 a^4 \Omega^6 = \frac{32}{5}\frac{G^4}{c^5}\frac{M^3\mu^2}{a^5}, \tag{2.38}$$

where, in order to obtain the last part of the relation, we have used Kepler's third law, $\Omega^2 = GM/a^3$. As the gravitating system loses energy by emitting radiation, the distance between the two bodies shrinks at a rate

$$\frac{da}{dt} = -\frac{64}{5}\frac{G^3}{c^5}\frac{\mu M^2}{a^3}, \tag{2.39}$$

Figure 3: A system of two bodies orbiting around their common center of gravity. Binary systems are the best sources of gravitational waves.

and the orbital frequency increases accordingly ($\dot{T}/T = 1.5\dot{a}/a$). If, for example, the present separation of the two stars is a_0, then the binary system will coalesce after a time

$$\tau = \frac{5}{256}\frac{c^5}{G^3}\frac{a_0^4}{\mu M^4} \tag{2.40}$$

Finally, the amplitude of the gravitational waves is

$$h = 5 \times 10^{-22} \left(\frac{M}{2.8M_\odot}\right)^{2/3} \left(\frac{\mu}{0.7M_\odot}\right) \left(\frac{f}{100\text{Hz}}\right)^{2/3} \left(\frac{15\text{Mpc}}{r}\right). \tag{2.41}$$

In all these formulae we have assumed that the orbits are circular. In general, the orbits of the two bodies are approximately ellipses, but it has been shown that long before the coalescence of the two bodies, the orbits become circular, at least for long-lived binaries, due to gravitational radiation. Also, the amplitude of the emitted gravitational waves depends on the angle between the line of sight and the axis of angular momentum; formula (2.41) refers to an observer along the axis of the orbital angular momentum. The complete formula for the amplitude contains angular factors of order 1. The relative strength of the two polarizations depends on that angle as well.

If three or more detectors observe the same signal it is possible to reconstruct the full waveform and deduce many details of the orbit of the binary system.

As an example, we will provide some details of the well-studied pulsar PSR 1913+16 (the Hulse-Taylor pulsar), which is expected to coalesce after $\sim 3.5 \times 10^8$ years. The binary system is roughly 5kpc away from Earth, the masses of the two neutron stars are estimated to be $\sim 1.4 M_\odot$ each, and the present period of the system is $\sim$7h and 45min. The predicted rate of period change is $\dot{T} = -2.4 \times 10^{-12}$sec/sec, while the corresponding observed value is in excellent agreement with the predictions, i.e., $\dot{T} = (-2.30 \pm 0.22) \times 10^{-12}$sec/sec; finally the present amplitude of gravitational waves is of the order of $h \sim 10^{-23}$ at a frequency of $\sim 7 \times 10^{-5}$Hz.

Figure 4: A pair of masses joined by a spring can be viewed as the simplest conceivable detector.

3. DETECTION OF GRAVITATIONAL WAVES

The first attempt to detect gravitational waves was undertaken by the pioneer Joseph Weber during the early 1960s. He developed the first resonant mass detector and inspired many other physicists to build new detectors and to explore from a theoretical viewpoint possible cosmic sources of gravitational radiation.

A pair of masses joined by a spring can be viewed as the simplest conceivable detector; see Figure 4. In practice, a cylindrical massive metal bar or even a massive sphere is used instead of this simple system. When a gravitational wave hits such a device, it causes the bar to vibrate. By monitoring this vibration, we can reconstruct the true waveform. The next step, following the idea of resonant mass detectors, was the replacement of the spring by pendulums. In this new detector the motions induced by a passing-by gravitational wave would be detected by monitoring, via laser interferometry, the relative change in the distance of two freely suspended bodies. The use of interferometry is probably the most decisive step in our attempt to detect gravitational wave signals. In what follows, we will discuss both resonant bars and laser interferometric detectors.

Although the basic principle of such detectors is very simple, the sensitivity of detectors is limited by various sources of noise. The internal noise of the detectors can be Gaussian or non-Gaussian. The non-Gaussian noise may occur several times per day such as strain releases in the suspension systems which isolate the detector from any environmental mechanical source of noise, and the only way to remove this type of noise is via comparisons of the data streams from various detectors. The so-called Gaussian noise obeys the probability distribution of Gaussian statistics and can be characterized by a *spectral density* $S_n(f)$. The observed signal at the output of a detector consists of the true gravitational wave strain h and Gaussian noise. The optimal method to detect a gravitational wave signal leads to the following signal-to-noise ratio:

$$\left(\frac{S}{N}\right)^2_{\rm opt} = 2\int_0^\infty \frac{|\tilde{h}(f)|^2}{S_n(f)} df, \tag{3.1}$$

where $\tilde{h}(f)$ is the Fourier transform of the signal waveform. It is clear from this expression that the sensitivity of gravitational wave detectors is limited by noise.

3.1 Resonant detectors

Suppose that a gravitational wave propagating along the z axis with pure (+) polarization impinges on an idealized detector, two masses joined by a spring along the x axis as in Figure 4. We will assume that $h_{\mu\nu}$ describes the strain produced in the spacetime by the passing wave. We will try to calculate the amplitude of the oscillations induced on the spring detector by the wave and the amount of energy absorbed by this detector. The tidal force induced on the detector is given by equation (2.23), and the masses will move according to the following equation of motion:

$$\ddot{\xi} + \dot{\xi}/\tau + \omega_0^2 \xi = -\frac{1}{2}\omega^2 L h_+ e^{i\omega t}, \tag{3.2}$$

where ω_0 is the natural vibration frequency of our detector, τ is the damping time of the oscillator due to frictional forces, L is the separation between the two masses and ξ is the relative change in the distance of the two masses. The gravitational wave plays the role of the driving force for our ideal oscillator, and the solution to the above equation is

$$\xi = \frac{\frac{1}{2}\omega^2 L h_+ e^{i\omega t}}{\omega_0^2 - \omega^2 + i\omega/\tau} \tag{3.3}$$

If the frequency ω of the impinging wave is near the natural frequency ω_0 of the oscillator (near resonance) the detector is excited into large-amplitude motions and it rings like a bell. Actually, in the case of $\omega = \omega_0$, we get the maximum amplitude $\xi_{\max} = \omega_0 \tau L H_+/2$. Since the size of our detector L and the amplitude of the gravitational waves h_+ are fixed, large-amplitude motions can be achieved only by increasing the quality factor $Q = \omega_0 \tau$ of the detector. In practice, the frequency of the detector is fixed by its size and the only improvement we can get is by choosing the type of material so that long relaxation times are achieved.

The *cross section* is a measure of the interception ability of a detector. For the special case of resonance, the average cross section of our test detector, assuming any possible direction of the wave, is

$$\sigma = \frac{32\pi}{15}\frac{G}{c^3}\omega_0 Q M L^2. \tag{3.4}$$

This formula is general; it applies even if we replace our toy detector with a massive metal cylinder, for example Weber's first detector. That detector had the following characteristics: Mass M=1410 kg, length L=1.5 m, diameter 66 cm, resonant frequency ω_0=1660Hz, and quality factor $Q = \omega_0 \tau = 2 \times 10^5$. For these values the calculated cross section is roughly $3 \times 10^{-19}\text{cm}^2$, which is quite small, and even worse, it can be reached only when the frequency of the impinging wave is very close to resonance frequency (the typical resonance width is usually of the order of 0.1-1 Hz).

In reality, the efficiency of a resonant bar detector depends on several other parameters. Here, we will discuss only the more fundamental ones. Assuming perfect

Figure 5: A graph of NAUTILUS in Frascati near Rome. Nautilus is probably the most sensitive resonant detector available.

isolation of the resonant bar detector from any external source of noise (acoustical, seismic, electromagnetic), the thermal noise is the only factor limiting our ability to detect gravitational waves. Thus, in order to detect a signal, the energy deposited by the gravitational wave every τ seconds should be larger than the energy kT due to thermal fluctuations. This leads to a formula for the minimum detectable energy flux of gravitational waves, which, following equation (2.25), leads into a minimum detectable strain amplitude

$$h_{\min} \leq \frac{1}{\omega_0 LQ}\sqrt{\frac{15kT}{M}} \tag{3.5}$$

For Weber's detector, at room temperature this yields a minimum detectable strain of the order of 10^{-20}. However, this estimate of the minimum sensitivity applies only to gravitational waves whose duration is at least as long as the damping time of the bar's vibrations and whose frequency perfectly matches the resonant frequency of the detector. For bursts or for periodic signals which are off-resonance (with regard to the detector) the sensitivity of a resonant bar detector decreases further by several orders of magnitude.

In reality, modern resonant bar detectors are quite complicated devices, consisting of a solid metallic cylinder weighing a few tons and suspended *in vacuo* by a cable that is wrapped under its center of gravity (Figure 5). This suspension system protects the antenna from external mechanical shocks. The whole system is cooled down to temperatures of a few kelvins or even millikelvins. To monitor the vibrations of the bar, piezoelectric transducers (or the more modern capacitive ones) are

attached to the bar. The transducers convert the bar's mechanical energy into electrical energy. The signal is amplified by an ultra-low-frequency amplifier, by using a device called a SQUID (Super-conducting QUantum Interference Device) before it becomes available for data analysis. Transducers and amplifiers of electronic signals require careful design to achieve low noise combined with adequate signal transfer.

The above description of the resonant bar detectors shows that, in order to achieve high sensitivity, one has to:

1. *Create more massive antennas.* Today, most of the antennas are about 50% more massive than the early Webers antenna. There are studies and research plans for future construction of spherical antennae weighing up to tenths of tons.

2. *Obtain higher quality factor Q.* Modern antennas generally use aluminum alloy 5056 ($Q \sim 4 \times 10^7$); although niobium (which is used in the NIOBE detector) is even better ($Q \sim 10^8$) but much more expensive. Silicon or sapphire bars would enhance the quality factor even more, but experimenters must first find a way to produce large single pieces of these crystals.

3. *Lower the temperature of the antenna as much as possible.* Advanced cryogenic techniques have been used and the resonant bar detectors are probably the coolest places in the Universe. Typical cooling temperatures for the most advanced antennae are below the temperature of liquid helium.

4. *Achieve strong coupling between the antenna and the electronics and low electrical noise.* The bar detectors include the best available technology related to transducers and integrate the most recent advances in SQUID technology.

Since the early 1990s, a number of resonant bar detectors have been in nearly continuous operation in several places around the world. They have achieved sensitivities of a few times 10^{-21}, but still there has been no clear evidence of gravitational wave detection. As we will discuss later, they will have a good chance of detecting a gravitational wave signal from a supernova explosion in our galaxy, although, this is a rather rare event (2-5 per century).

The technical specifications of the most sensitive cryogenic bar detectors in operation are as follows:

- ALLEGRO (Baton Rouge, USA) Mass 2296 Kg (Aluminium 5056), length 3 m, bar temperature 4.2 K, mode frequency 896 Hz.

- AURIGA (Legrano, Italy) Mass 2230 Kg (Aluminium 5056), length 2.9 m, bar temperature 0.2 K, mode frequency 913 Hz.

- EXPLORER (CERN, Switzerland) Mass 2270 Kg (Aluminium 5056), length 3 m, bar temperature 2.6K, mode frequency 906Hz.

- NAUTILUS (Frascati, Italy) Mass 2260 Kg (Aluminium 5056), length 3 m, bar temperature 0.1 K, mode frequency 908 Hz.

- NIOBE (Perth, Australia) Mass 1500 Kg (Niobium), length 1.5 m, bar temperature 5K, mode frequency 695Hz.

Also, there are plans for construction of *massive spherical resonant detectors*, the advantages of which will be their high mass, their broader sensitivity (up to 100-200 Hz) and their omnidirectional sensitivity. In a spherical detector, five modes at a time will be excited, which is equivalent to five independent detectors oriented in different ways. This offers the opportunity, in the case of detection, to obtain direct information about the polarization of the wave and the direction to the source. A prototype spherical detector is already in operation in Leiden, Netherlands (mini-GRAIL)[7] with 1 m diameter and mode frequency ∼3.2 kHz. Two other detectors of the same size are under construction in Italy (Sfera) and Brazil (Shenberg).

3.2 Beam detectors

3.2.1 Laser Interferometers

A laser interferometer is an alternative gravitational wave detector that offers the possibility of very high sensitivities over a broad frequency band. Originally, the idea was to construct a new type of resonant detector with much larger dimensions. As one can realize from the relations (3.4) and (3.5), the longer the resonant detector is the more sensitive it becomes. One could then try to measure the relative change in the distance of two well-separated masses by monitoring their separation via a laser beam that continuously bounces back and forth between them. (This technique is actually used in searching for gravitational waves by using the so-called Doppler tracking technique, where a distant interplanetary spacecraft is monitored from earth through a microwave tracking link; the earth and spacecraft act as free particles.) Soon, it was realized that it is much easier to use laser light to measure relative changes in the lengths of two perpendicular arms; see Figure 6. Gravitational waves that are propagating perpendicular to the plane of the interferometer will increase the length of one arm of the interferometer, and at the same time will shorten the other arm, and vice versa. This technique of monitoring the waves is based on Michelson interferometry. L-shaped interferometers are particularly suited to the detection of gravitational waves due to their quadrupolar nature.

Figure 6 shows a schematic design of a Michelson interferometer; the three masses M_0, M_1 and M_2 are freely suspended. Note that the resonant frequencies of these pendulums should be much smaller than the frequencies of the waves that we are supposed to detect since the pendulums are supposed to behave like free masses. Mirrors are attached to M_1 and M_2 and the mirror attached on mass M_0 splits the light (beam splitter) into two perpendicular directions. The light is reflected on the

Figure 6: A schematic design of a Michelson interferometer.

two corner mirrors and returns back to the beam splitter. The splitter now half-transmits and half-reflects each one of the beams. One part of each beam goes back to the laser, while the other parts are combined to reach the photodetector where the fringe pattern is monitored. If a gravitational wave slightly changes the lengths of the two arms, the fringe pattern will change, and so by monitoring the changes of the fringe pattern one can measure the changes in the arm lengths and consequently monitor the incoming gravitational radiation [1].

Let us consider an impinging gravitational wave with amplitude h and (+) polarization, propagating perpendicular to the plane of the detector. We will further assume that the frequency is much higher than the resonant frequency of the pendulums and the wavelength is much longer than the arm length of our detector. Such a wave will generate a change of $\Delta L \sim hL/2$ in the arm length along the x-direction and an opposite change in the arm length along the y-direction, according to equations (2.19) and (2.20). The total difference in length between the two arms will be

$$\frac{\Delta L}{L} \sim h. \tag{3.6}$$

For a gravitational wave with amplitude $h \sim 10^{-21}$ and detector arm- length 4 km (such as LIGO), this will induce a change in the arm-length of about $\Delta L \sim 10^{-16}$. In the general case, when a gravitational wave with arbitrary polarization impinges on the detector from a random direction, the above formula will be modified by some angular coefficients of order 1.

[1] In practice, things are arranged so that when there is no actual change in the arms, all light that returns on the beam splitter from the corner mirrors is sent back into the laser, and only if there is some motion of the masses there is an output at the photodetector

If the light bounces a few times between the mirrors before it is collected in the photodiode, the effective arm length of the detector is increased considerably, and the measured variations of the arm lengths will be increased accordingly. This is a quite efficient procedure for making the arm length longer. For example, a gravitational wave at a frequency of 100 Hz has a wavelength of 3000 Km, and if we assume 100 bounces of the laser beam in the arms of the detector, the effective arm-length of the detector is 100 times larger than the actual arm-length, but still this is 10 times smaller than the wavelength of the incoming wave. The optical cavity that is created between the mirrors of the detector is known as a Fabry-Perot cavity and is used in modern interferometers.

In the remainder of this paragraph we will focus on the Gaussian sources of noise and their expected influence on the sensitivity of laser intereferometers.

a. Photon shot noise. When a gravitational wave produces a change ΔL in the arm-length, the phase difference between the two light beams changes by an amount $\Delta\phi = 2b\Delta L/\tilde{\lambda}$, where $\tilde{\lambda}$ is the reduced wavelength of the laser light ($\sim 10^{-8}$cm) and b is the number of bounces of the light in each arm. It is expected that a detectable gravitational wave will produce a phase shift of the order of 10^{-9}rad. The precision of the measurements, though, is ultimately restricted by fluctuations in the fringe pattern due to fluctuations in the number of detected photons. The number of photons that reach the detector is proportional to the intensity of the laser beam and can be estimated via the relation $N = N_0 \sin^2(\Delta\phi/2)$, where N_0 is the number of photons that the laser supplies and N is the number of detected photons. Inversion of this equation leads to an estimation of the relative change of the arm lengths ΔL by measuring the number of the emerging photons N. However, there are statistical fluctuations in the population of photons, which are proportional to the square root of the number of photons. This implies an uncertainty in the measurement of the arm length

$$\delta(\Delta L) \sim \frac{\tilde{\lambda}}{2b\sqrt{N_0}} \tag{3.7}$$

Thus, the minimum gravitational wave amplitude that we can measure is

$$h_{\min} = \frac{\delta(\Delta L)}{L} = \frac{\Delta L}{L} \sim \frac{\tilde{\lambda}}{bLN_0^{1/2}} \sim \frac{1}{bL}\left(\frac{\hbar c\tilde{\lambda}}{\tau I_0}\right)^{1/2}, \tag{3.8}$$

where I_0 is the intensity of the laser light (∼5-10 W) and τ is the duration of the measurement. This limitation in the detector's sensitivity due to the photon counting uncertainty is known as *photon shot noise*. For a typical laser interferometer the photon shot noise is the dominant source of noise for frequencies above 200 Hz, while its power spectral density $S_n(f)$ for frequencies 100-200Hz is of the order of $\sim 3 \times 10^{-23}\sqrt{\text{Hz}}$.

b. Radiation pressure noise: According to formula (3.8), the sensitivity of a detector can be increased by increasing the intensity of the laser. However, a very powerful laser produces a large radiation pressure on the mirrors. Then an uncertainty in the measurement of the momentum deposited on the mirrors leads to a proportional uncertainty in the position of the mirrors or, equivalently, in the measured change in the arm-lengths. Then, the minimum detectable strain is limited by

$$h_{\min} \sim \frac{\tau}{m}\frac{b}{L}\left(\frac{\tau\hbar I_0}{c\tilde{\lambda}}\right)^{1/2}, \tag{3.9}$$

where m is the mass of the mirrors.

As we have seen, the photon shot noise decreases as the laser power increases, while the inverse is true for the noise due to radiation pressure fluctuations. If we try to minimize these two types of noise with respect to the laser power, we get a minimum detectable strain for the optimal power via the very simple relation

$$h_{\min} \sim \frac{1}{L}\left(\frac{\tau\hbar}{m}\right)^{1/2} \tag{3.10}$$

which for the LIGO detector (where the mass of the mirrors is ∼100 kg and the arm length is 4 km), for observation time of 1 ms, gives $h_{\min} \approx 10^{-23}$.

c. Quantum limit. An additional source of uncertainty in the measurements is set by Heisenberg's principle, which says that the knowledge of the position and the momentum of a body is restricted from the relation $\Delta x \cdot \Delta p \geq \hbar$. For an observation that lasts some time τ, the smallest measurable displacement of a mirror of mass m is ΔL; assuming that the momentum uncertainty is $\Delta p \approx m \cdot \Delta L/\tau$, we get a minimum detectable strain due to quantum uncertainties

$$h_{\min} = \frac{\Delta L}{L} \sim \frac{1}{L}\left(\frac{\tau\hbar}{m}\right)^{1/2}. \tag{3.11}$$

Surprisingly, this is identical to the optimal limit that we calculated earlier for the other two types of noise. The standard quantum limit does set a fundamental limit on the sensitivity of beam detectors. An interesting feature of the quantum limit is that it depends only on a single parameter, the mass of the mirrors.

d. Seismic noise. At frequencies below 60 Hz, the noise in the interferometers is dominated by seismic noise. This noise is due to geological activity of the earth, and human sources, e.g., traffic and explosions. The vibrations of the ground couple to the mirrors via the wire suspensions which support them. This effect is strongly suppressed by properly designed suspension systems. Still, seismic noise is very difficult to eliminate at frequencies below 5-10 Hz.

e. Residual gas-phase noise. The statistical fluctuations of the residual gas density induce a fluctuation of the refraction index and consequently of the monitored

phase shift. Hence, the residual gas pressure through which the laser beams travel should be extremely low. For this reason the laser beams are enclosed in pipes over their entire length. Inside the pipes a high vacuum of the order of 10^{-9} Torr guarantees elimination of this type of noise.

Prototype laser interferometric detectors have been already constructed in the USA, Germany and UK more than 15 years ago. These detectors have an arm length of a few tens of meters and they have achieved sensitivities of the order of $h \sim 10^{-19}$. The construction of the new generation of laser interferometric detectors is near completion and some of them have already collected data. It is expected that in a year or so they will be in full operation. The US project named LIGO [8](Laser Interferometer Gravitational Observatory) consists of two detectors with arm length of 4 Km, one in Hanford, Washington, and one in Livingston, Louisiana. The detector in Hanford includes, in the same vacuum system, a second detector with arm length of 2 km. The detectors are already in operation and they have recently achieved the designed sensitivity.

The Italian/French EGO (VIRGO) detector [9] of arm-length 3 km at Cascina near Pisa is designed to have better sensitivity at lower frequencies is about to commence operation next year. GEO600 is the German/British detector build in Hannover [10]. It has 600 m arm length and is collecting data in collaboration with LIGO [11]. The TAMA300[12] detector in Tokyo has arm length of 300 m and it was the first major interferometric detector in operation. There are already plans for improving the sensitivities of all the above detectors and the construction of new interferometers in the near future.

3.2.2 Space Detectors

Both bars and laser interferometers are high-frequency detectors, but there are a number of interesting gravitational waves sources which emit signals at lower frequencies. The seismic noise provides an insurmountable obstacle in any earth-based experiment and the only way to surpass this barrier is to fly a laser interferometer in space. LISA [13] (Laser Interferometer Space Antenna) is such a system. It has been proposed by European and American scientists and has been adopted by ESA (European Space Agency) as a cornerstone mission, while recently NASA joined the effort and the launch date is expected to be around 2011.

LISA will consist of three identical drag-free spacecrafts forming an equilateral triangle with one spacecrafts at each vertex (Figure 7). The distance between the two vertices (the arm length) is 5×10^6 km. The spacecrafts will be placed into the same heliocentric orbit as earth, but about 20^o behind earth. The equilateral triangle will be inclined at an angle of 60^o with respect to Earth's orbital plane. The three spacecraft would track each other optically by using laser beams. Because of the diffraction losses it is not feasible to reflect the beams back and forth as is done

Figure 7: A schematic design of the space interferometer LISA.

with LIGO. Instead, each spacecraft will have its own laser. The lasers will be phase locked to each other, achieving the same kind of phase coherence as LIGO does with mirrors. The configuration will function as three, partially independent and partially redundant, gravitational wave interferometers.

At frequencies $f \geq 10^{-3}$, LISA's noise is mainly due to photon shot noise. The sensitivity curve steepens at $f \sim 3 \times 10^{-2}$ Hz because at larger frequencies the gravitational waves period is shorter than the round-trip light travel time in each arm. For $f \leq \times 10^{-2}$ Hz, the noise is due to buffeting-induced random motions of the spacecraft, and cannot be removed by the drag- compensation system. LISA's sensitivity is roughly the same as that of LIGO, but at 10^5 times lower frequency. Since the gravitational waves energy flux scales as $F \sim f^2 h^2$, this corresponds to 10^{10} times better energy sensitivity than LIGO.

3.2.3 Satellite Tracking

The Doppler delay of communication signals between the earth-based stations and spacecraft underlies another type of gravitational wave detector. A radio signal of frequency f_0 is transmitted to a spacecraft and is coherently transported back to earth, where it is received and its frequency is measured with a highly stable clock (typically a hydrogen maser). The relative change $\Delta f/f_0$ as function of time is monitored. A gravitational wave propagating through the solar system causes small perturbations in $\Delta f/f_0$. The relative shift in the frequency of the signals is proportional to the amplitude of gravitational waves. With this technique, broad-band searches are possible in the millihertz frequency band, and thanks to the very stable atomic clocks, it is possible to achieve sensitivities of order $h_{\min} \sim 10^{-13} - 10^{-15}$. Noise sources that affect the sensitivity of Doppler tracking experiments can be divided into two broad

classes a) instrumental and b) related to propagation. At the high-frequency end of the band accessible to Doppler tracking, thermal noise dominates over all other noise sources, typically at about 0.1 Hz. Among all other sources of instrumental noise (transmitter and receiver, mechanical stability of the antenna, stability of the spacecraft etc), clock noise has been shown to be the most important instrumental source of frequency fluctuations. The propagation noise is due to fluctuations in the index of refraction of the troposphere, ionosphere, and the interplanetary solar plasma. Both NASA and ESA have performed such measurements and there is a continued effort in this direction.

3.2.4 Pulsar Timing

Pulsars are extremely stable clocks and by measuring irregularities in their pulses we expect to set upper limits on the background gravitational waves (see next section). If an observer monitors simultaneously two or more pulsars, the correlation of their signals could be used to detect gravitational waves. Since such observation requires timescales of the order of 1 year, this means that the waves have to be of extremely low frequencies.

4. ASTRONOMICAL SOURCES OF GRAV. WAVES

The new generation of gravitational wave detectors (LIGO, VIRGO) have very good chances of detecting gravitational waves, but until these expectations are fulfilled, we can only make educated guesses as to the possible astronomical sources of gravitational waves. The detectability of these sources depends on three parameters: their intrinsic gravitational wave luminosity, their event rate, and their distance from the Earth. The luminosity can be approximately estimated via the quadrupole formula that we discussed earlier. Even though there are certain restrictions in its applicability (weak field, slow motion), it provides a very good order-of-magnitude estimate for the expected gravitational wave flux on Earth. The rate, at which various events with high luminosity in gravitational waves take place is extrapolated from astronomical observations in the electromagnetic spectrum. Still, there might be a number of *gravitationally luminous sources*, for example binary black holes, for which we have no direct observations in the electromagnetic spectrum. Finally, the amplitude of gravitational wave signals decreases as one over the distance to the source. Thus, a signal from a supernova explosion might be clearly detectable if the event takes place in our galaxy (2-3 events per century), but it is highly unlikely to be detected if the supernova explosion occurs at far greater distances, of order 100 Mpc, where the event rate is high and at least a few events per day take place. All three factors have to be taken into account when discussing sources of gravitational waves.

It was mentioned earlier that the frequency of gravitational waves is proportional to the square root of the mean density of the emitting system; this is approximately

true for any gravitating system. For example, neutron stars usually have masses around $1.4M_\odot$ and radii in the order of 10 km; thus if we use these numbers in the relation $f \sim \sqrt{GM/R^3}$ we find that an oscillating neutron star will emit gravitational waves primarily at frequencies of 2-3 kHz. By analogy, a black-hole with a mass $\sim 100M_\odot$, will have a radius of ∼300 km and the natural oscillation frequency will be ~ 100 Hz. Finally, for a binary system Kepler's law (see section 2.5) provides a direct and accurate estimation of the frequency of the emitted gravitational waves. For two $1.4M_\odot$ neutron stars orbiting around each other at a distance of 160 Km, Keplers law predicts an orbital frequency of 50 Hz, which leads to an observed gravitational wave frequency of 100 Hz.

4.1 Radiation from gravitational collapse

Type II supernovae are associated with the core collapse of a massive star together with a shock-driven expansion of a luminous shell which leaves behind a rapidly rotating neutron star or a black hole, if the core has mass of $> 2-3M_\odot$. The typical signal from such an explosion is broadband and peaked at around 1 kHz. Detection of such a signal was the goal of detector development over the last three decades. However, we still know little about the efficiency with which this process produces gravitational waves. For example, an exactly spherical collapse will not produce any gravitational radiation at all. The key question is what is the kinetic energy of the nonspherical motions, since the gravitational wave amplitude is proportional to this [equation (2.30). After 30 years of theoretical and numerical attempts to simulate gravitational collapse, there is still no great progress in understanding the efficiency of this process in producing gravitational waves. For a conservative estimate of the energy in non-spherical motions during the collapse, relation (2.32) leads to events of an amplitude detectable in our galaxy, even by bar detectors. The next generation of laser interferometers would be able to detect such signals from Virgo cluster at a rate of a few events per month.

The main source for deviations from spherical or axial symmetry during the collapse is the angular momentum. During the contraction phase, the angular momentum is conserved, and the star spins up to rotational periods of the order of 1 msec. In this case, consequent processes with large luminosity might take place in this newly born neutron star. A number of instabilities, such as the so-called *bar mode instability* and the *r-mode instability*, may occur which radiate copious amounts of gravitational radiation immediately after the initial burst. Gravitational wave signals from these rotationally induced stellar instabilities are detectable from sources in our galaxy and are marginally detectable if the event takes place in the nearby cluster of about 2500 galaxies, the Virgo cluster, 15 Mpc away from the Earth. Additionally, there will be weaker but extremely useful signals due to subsequent oscillations of the neutron star; *f*, *p* and *w* modes are some of the main patterns of oscillations (normal modes) of the neutron star that observers might search for.

These modes have been studied in detail and once detected in the signal, they would provide a sensitive probe of the neutron star structure and its supranuclear equation of state. Detectors with high sensitivity in the kHz band will be needed in order to fully develop this so-called gravitational wave asteroseismology.

If the collapsing central core is unable to drive off its surrounding envelope, then the collapse continues and finally a black hole forms. In this case the instabilities and oscillations that we discussed above are absent and the newly formed black hole radiates away, within a few milliseconds, any deviations from axisymmetry and ends up as a rotating or *Kerr* black hole. The characteristic oscillations of black holes (normal modes) are well studied, and this unique ringing down of a black hole could be used as a direct probe of their existence. The frequency of the signal is inversely proportional to the black hole mass. For example, it has been stated earlier that a $100M_{\odot}$ black hole will oscillate at a frequency of ∼100 Hz (an ideal source for LIGO), while a supermassive one with mass $10^7 M_{\odot}$, which might be excited by an infalling star, will ring down at a frequency of 10^{-3} Hz (an ideal source for LISA). The analysis of such a signal should reveal directly the two parameters that characterize any (uncharged) black hole; namely its mass and angular momentum.

4.2 Radiation from binary systems

Binary systems are the best sources of gravitational waves because they emit copious amounts or gravitational radiation, and for a given system we know exactly what is the amplitude and frequency of the gravitational waves in terms of the masses of the two bodies and their separation (see section 2.5). If a binary system emits detectable gravitational radiation in the bandwidth of our detectors, we can easily identify the parameters of the system. According to the formulas of section 2.5, the observed frequency change will be

$$\dot{f} \sim f^{11/3} \mathcal{M}^{5/3}, \tag{4.1}$$

and the corresponding amplitude will be

$$h \sim \frac{\mathcal{M}^{5/3} f^{2/3}}{r} = \frac{\dot{f}}{r f^3}, \tag{4.2}$$

where $\mathcal{M}^{5/3} = \mu M^{2/3}$ is a combination of the total and reduced mass of the system, called *chirp mass*. Since both frequency f and its rate of change $\dot{f}$ are measurable quantities, we can immediately compute the chirp mass (from the first relation), thus obtaining a measure of the masses involved. The second relation provides a direct estimate of the distance of the source. These relations have been derived using the Newtonian theory to describe the orbit of the system and the quadrupole formula for the emission of gravitational waves. Post-Newtonian theory inclusion of the most important relativistic corrections in the description of the orbit can provide more accurate estimates of the individual masses of the components of the binary system.

When analyzing the data of periodic signals the *effective amplitude* is not the amplitude of the signal alone but $h_c = \sqrt{n} \cdot h$, where n is the number of cycles of the signal within the frequency range where the detector is sensitive. A system consisting of two typical neutron stars will be detectable by LIGO when the frequency of the gravitational waves is ∼10Hz until the final coalescence around 1000Hz. This process will last for about 15 min and the total number of observed cycles will be of the order of 10^4, which leads to an enhancement of the detectability by a factor 100. Binary neutron star systems and binary black hole systems with masses of the order of $50M_\odot$ are the primary sources for LIGO. Given the anticipated sensitivity of LIGO, binary black hole systems are the most promising sources and could be detected as far as 200 Mpc away. The event rate with the present estimated sensitivity of LIGO is probably a few events per year, but future improvement of detector sensitivity (the LIGO II phase) could lead to the detection of at least one event per month. Supermassive black hole systems of a few million solar masses are the primary source for LISA. These binary systems are rare, but due to the huge amount of energy released, they should be detectable from as far as the boundaries of the observable universe.

4.3 Radiation from spinning neutron stars

A perfectly axisymmetric rotating body does not emit any gravitational radiation. Neutron stars are axisymmetric configurations, but small deviations cannot be ruled out. Irregularities in the crust (perhaps imprinted at the time of crust formation), strains that have built up as the stars have spun down, off-axis magnetic fields, and/or accretion could distort the axisymmetry. A bump that might be created at the surface of a neutron star spinning with frequency f will produce gravitational waves at a frequency of $2f$ and such a neutron star will be a weak but continuous and almost monochromatic source of gravitational waves. The radiated energy comes at the expense of the rotational energy of the star, which leads to a spin down of the star. If gravitational wave emission contributes considerably to the observed spin down of pulsars, then we can estimate the amount of the emitted energy. The corresponding amplitude of gravitational waves from nearby pulsars (a few kpc away) is of the order of $h \sim 10^{-25} - 10^{-26}$, which is extremely small. If we accumulate data for sufficiently long time, e.g., 1 month, then the effective amplitude, which increases as the square root of the number of cycles, could easily go up to the order of $h_c \sim 10^{-22}$. We must admit that at present we are extremely ignorant of the degree of asymmetry in rotating neutron stars, and these estimates are probably very optimistic. On the other hand, if we do not observe gravitational radiation from a given pulsar we can place a constraint on the amount of non axisymmetry of the star.

4.4 Cosmological gravitational waves

One of the strongest pieces of evidence in favor of the Big Bang scenario is the 2.7 K cosmic microwave background radiation. This thermal radiation first bathed the

universe around 380,000 years after the Big Bang. By contrast, the gravitational radiation background anticipated by theorists was produced at Planck times, i.e., at 10^{-32} sec or earlier after the Big Bang. Such gravitational waves have travelled almost unimpeded through the universe since they were generated. The observation of cosmological gravitational waves will be one of the most important contributions of gravitational wave astronomy. These primordial gravitational waves will be, in a sense, another source of noise for our detectors and so they will have to be much stronger than any other internal detector noise in order to be detected. Otherwise, confidence in detecting such primordial gravitational waves could be gained by using a system of two detectors and cross-correlating their outputs. The two LIGO detectors are well placed for such a correlation.

Acknowledgements

I am grateful to I. D. Jones, Th. Apostolatos, J. Ruoff, N. Stergioulas and J. Agtzidis for useful suggestions which improved the original manuscript.

References

[1] *Fundamentals of Interferometric Gravitational Wave Detectors*, P.R. Saulson, World Scientific (1994).

[2] *The Detection of Gravitational Waves*, D.G. Blair, Cambridge University Press (1991).

[3] *Relativistic Gravitation and Gravitational Radiation*, Cambridge University Press, Editors J-A. Marck and J-P. Lasota (1997).

[4] *Gravitational Radiation*, K.S. Thorne, in S.W. Hawking and W. Israel (eds), 300 Years of Gravitation, (Cambridge University Press, Cambridge).

[5] *Gravitational Wave Physics*, K.D. Kokkotas, Encyclopedia of Physical Science and Technology, 3rd Edition, Volume 7 Academic Press, (2002)

[6] *An Overview of Gravitational-Wave Sources* C. Cutler and K.P. Thorne gr-qc/0204090 April 30th 2002.

[7] mini-GRAIL http://www.minigrail.nl

[8] LIGO http://www.ligo.caltech.edu

[9] VIRGO http://www.pi.infn.it/virgo/virgoHome.html

[10] GEO http://www.geo600.uni-hannover.de

[11] *Detector Description and Performance for the First Coincidence Observations between LIGO and GEO* The LIGO Scientific Collaboration, gr-qc/0308043, August 14th, 2003.

[12] TAMA http://tamago.mtk.nao.ac.jp

[13] LISA http://lisa.jpl.nasa.gov

Second VESF School on Gravitational Waves

Cascina, Italy

May 28th – June 1st 2007

An Introduction to General Relativity, Gravitational Waves and Detection Principles

Dr Martin Hendry

University of Glasgow, UK

martin@astro.gla.ac.uk

Overview

These lectures present an introduction to General Relativity (GR) and its implications for the basic design properties of gravitational wave detectors.

In Sections 1 to 5 we discuss the foundations of GR and its key physical idea – that gravity manifests itself as a curvature of spacetime. We introduce the mathematical machinery necessary to describe this curvature and then briefly discuss the formulation and solution of Einstein's field equations, which encode the relationship between the curvature of spacetime and its matter and energy content.
In Sections 6 to 9 we then investigate the characteristics of *non-stationary* spacetimes. Although for simplicity we restrict our discussion to the **weak-field approximation**, we nevertheless derive a number of key results for gravitational wave research:

- We show that the free-space solutions for the metric perturbations of a 'nearly flat' spacetime take the form of a wave equation, propagating at the speed of light. This encapsulates the central physical idea of General Relativity: that the instantaneous 'spooky action at a distance' of Newton's gravitational force is replaced in Einstein's theory by spacetime curvature. Moreover, changes in this curvature – the so-called 'ripples in spacetime' beloved of popular accounts of gravitational waves – propagate outward from their source at the speed of light in a vacuum.

- We consider carefully the mathematical characteristics of our metric perturbation, and show that a coordinate system (known as the **Transverse–Traceless gauge**) can be chosen in which its 16 components reduce to only 2 independent components (which correspond to 2 distinct **polarisation states**).

- We investigate the geodesic equations for nearby test particles in a 'nearly flat' spacetime, during the passage of a gravitational wave. We demonstrate the quadrupolar nature of the wave disturbance and the change in the proper

distance between test particles which it induces, and explore the implications of our results for the basic design principles of gravitational wave detectors.

- As a 'taster' for the lectures to follow, we derive some basic results on the generation of gravitational waves from a simple astrophysical source: a binary star system. In particular, we estimate the amplitude and frequency from e.g. a binary neutron star system, and demonstrate that any gravitational waves detected at the Earth from such an astrophysical source will be extremely weak.

These lectures are extracted, adapted and extended from a 20 lecture undergraduate course on General Relativity and a short graduate course on gravitational waves – both of which I have taught in recent years at the University of Glasgow. Anyone who wishes to may access the complete lecture notes for the undergraduate course via the following websites:

Part 1: Introduction to General Relativity.

`http://www.astro.gla.ac/users/martin/teaching/gr1/gr1_index.html`

Part 2: Applications of General Relativity.

`http://www.astro.gla.ac.uk/users/martin/teaching/gr2/gr2_index.html`

Both websites are password-protected, with username and password 'honours'.

Suggested Reading

Much of the material covered in these lectures can be found, in much greater depth, in Bernard Schutz' classic textbook '*A first Course in General Relativity*' (CUP; ISBN 0521277035). For an even more comprehensive treatment see also '*Gravitation*', by Misner, Thorne and Wheeler (Freeman; ISBN 0716703440) – although this book is not for the faint-hearted (quite literally, as it has a mass of several kg!)

Notation, Units and Conventions

In these lectures we will follow the convention adopted in Schutz' textbook and use a metric with signature $(-,+,+,+)$. Although we will generally present equations in component form, where appropriate we will write one forms, vectors and tensors in bold face (e.g. $\tilde{\mathbf{p}}$ and $\vec{\mathbf{V}}$ and $\mathbf{T}$ respectively). We will adopt throughout the standard Einstein summation convention that an upper and lower repeated index implies summation over that index. We will normally use Roman indices to denote spatial components (i.e. $i = 1, 2, 3$ etc) and Greek indices to denote 4-dimensional components (i.e. $\mu = 0, 1, 2, 3$ etc).

We will generally use commas to denote partial differentiation and semi-colons to denote covariant differentation, when expressions are given in component form. We will also occasionally follow the notation adopted in Schutz, and denote the covariant derivative of e.g. the vector field $\vec{\mathbf{V}}$ by $\nabla\vec{\mathbf{V}}$. Similarly, we will denote the covariant derivative of a vector $\vec{\mathbf{V}}$ along a curve with tangent vector $\vec{\mathbf{U}}$ by $\nabla_{\vec{\mathbf{U}}}\vec{\mathbf{V}}$

Finally, in most situations we will use so-called **geometrised units** in which $c = 1$ and $G = 1$. Thus we effectively measure time in units of **length** – specifically, the distance travelled by light in that time. Thus

$$1 \text{ second} \equiv 3 \times 10^8 \text{ m} .$$

Recall that the gravitational constant, G, in SI units is

$$G \simeq 6.67 \times 10^{-11}\,\mathrm{N\,m^2\,kg^{-2}}$$

but the Newton is a composite SI unit; i.e.

$$1\,\mathrm{N} = 1\,\mathrm{kg\,m\,s^{-2}}$$

so that

$$G \simeq 6.67 \times 10^{-11}\,\mathrm{m^3\,kg^{-1}\,s^{-2}}$$

Replacing our unit of time with the unit of length defined above, gives

$$G \simeq 7.41 \times 10^{-28}\,\mathrm{m\,kg^{-1}}$$

So by setting $G = 1$ we are effectively measuring *mass* also in units of length. It follows that, in these new units

$$1\ \mathrm{kg} = 7.41 \times 10^{-28}\ \mathrm{m}\ .$$

In summary, then, our geometrised units take the form

$$\begin{array}{lrcl} \text{Unit of length:} & 1\ \mathrm{m} & & \\ \text{Unit of time:} & 1\ \mathrm{m} & \equiv & 3.33 \times 10^{-9}\ \mathrm{s} \\ \text{Unit of mass:} & 1\ \mathrm{m} & \equiv & 1.34 \times 10^{27}\ \mathrm{kg} \end{array}$$

1 Foundations of General Relativity

General relativity (GR) is Albert Einstein's theory of gravitation, which he published in 1916, 11 years after publishing his Special Theory of Relativity. While the latter theory probably has had wider cultural impact (*everyone* has heard of "E equals mc squared"!) GR is widely recognised among physicists as Einstein's real masterwork: a truly remarkable achievement, greater than all the feats of his 'Annus Mirabilis' in 1905. It was described by Max Born as "the greatest feat of human thinking about nature, the most amazing combination of philosophical penetration, physical intuition and mathematical skill".

GR explains gravitation as a consequence of the curvature of spacetime, while in turn spacetime curvature is a consequence of the presence of matter. Spacetime curvature affects the movement of matter, which reciprocally determines the geometric properties and evolution of spacetime. We can sum this up neatly (slightly paraphrasing a quotation due to John Wheeler) as follows:

Spacetime tells matter how to move,
and matter tells spacetime how to curve

A useful metaphor for gravity as spacetime curvature is to visualise a stretched sheet of rubber, deformed by the presence of a massive body (see Figure 1).

To better understand how Einstein arrived at this remarkable theory we should first briefly consider his Special Theory of Relativity.

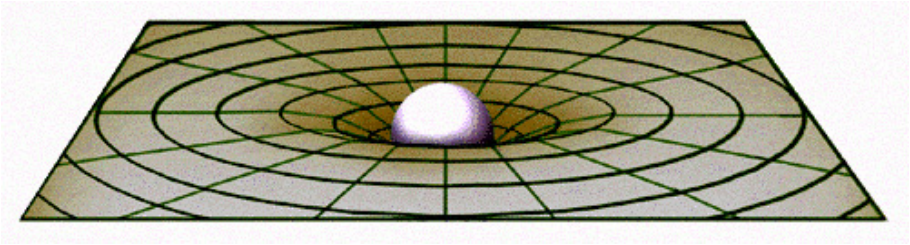


Figure 1: Familiar picture of spacetime as a stretched sheet of rubber, deformed by the presence of a massive body. A particle moving in the gravitational field of the central mass will follow a curved – rather than a straight line – trajectory as it moves across the surface of the sheet and approaches the central mass.

1.1 Special relativity

GR is a generalisation of special relativity (SR), in which Einstein set out to formulate the laws of physics in such a way that they be valid in all **inertial reference frames** – i.e. all frames in which Newton's first and second laws of motion hold – independently of their relative motion. At the heart of SR is the fundamental postulate of relativity: that the speed of light in the vacuum be the same for every inertial observer. Einstein showed that, as a consequence of this postulate, the Newtonian concept of a rigid framework of space and time against which physical phenomena were played out was no longer tenable: measurements of time and space cannot be absolute, but depend on the observer's motion and are related via the **Lorentz transformations**. Space and time as distinct entities gave way to a unified spacetime, and only the **spacetime interval** between events is independent of the observer's reference frame (referred to as **Lorentz frame**). For neighbouring events taking place at spacetime coordinates (t, x, y, z) and $(t + dt, x + dx, y + dy, z + dz)$, the interval is defined by

$$ds^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2 \tag{1}$$

Another inertial observer using a different coordinate system (t', x', y', z') (i.e. in a different Lorentz frame) will find the same value for the spacetime interval between the events, despite measuring them to have a different spatial and temporal

separation. Moreover:

- Intervals between neighbouring events with $ds^2 < 0$ are **timelike**; there exists a Lorentz frame ($\mathcal{A}$) in which the events occur at the same spatial coordinates, and in frame $\mathcal{A}$ the coordinate time separation of the events is equal to the **proper time** interval between them. Furthermore, the two events could lie on the **worldline** (i.e. the trajectory through spacetime) of a material particle.

- Intervals between neighbouring events with $ds^2 > 0$ are **spacelike**; there exists a Lorentz frame ($\mathcal{B}$) in which the events occur at the same coordinate time, and in frame $\mathcal{B}$ the spatial separation of the events is equal to the **proper distance** between them. The two events could *not* lie on the worldline of a material particle (since in frame $\mathcal{B}$ the particle would appear to be in two places at once!).

- Intervals between neighbouring events with $ds^2 = 0$ are **null** or **lightlike**; the two events could lie on the worldline of a photon

The geometry described by equation (1) differs markedly from that of 'flat' Euclidean space, and is known as Minkowski spacetime. We can re-write equation (1) more generally as

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu \tag{2}$$

where summation over the indices μ and ν is implied and both indices range over $0, 1, 2, 3$, corresponding to t, x, y, z. In equation (2) the $g_{\mu\nu}$ are components of the **metric** (we will provide a more formal definition of the metric shortly) which describes how intervals are measured in our spacetime. In general these components may be complicated functions of the spacetime coordinates but for Minkowski spacetime, in Cartesian coordinates and setting $c = 1$, the metric takes a very simple form

$$g_{\mu\nu} = \mathrm{diag}(-1, 1, 1, 1) \tag{3}$$

How does Newtonian gravity fit into this SR picture? We see immediately that the answer is "not well". Newtonian gravity is inherently non-relativistic since it describes the gravitational force between two masses as acting instantaneously and as depending on the distance separating the two masses. Different inertial observers would not agree about either point, and so would not agree about the force of gravity between the two masses. And what about *non-inertial* observers? The elegant geometry of Minkowski spacetime applied only to observers in uniform relative motion, not to accelerated motion. Yet Einstein sought a **fully covariant** theory, capable of describing the laws of physics (including gravity) in *any* coordinate system and for *any* relative motion. The key step towards achieving this came with Einstein's realisation that gravity and accleration are fundamentally equivalent – an idea enshrined in the **principle of equivalence**.

1.2 The equivalence principles

The principle of equivalence is often presented in two distinct forms: the **weak equivalence principle** and the **strong equivalence principle**.

1.2.1 The weak equivalence principle

In Newtonian physics the **inertial mass** of a body is a measure of its resistance to acceleration, and is the quantity that appears on the right hand side of the equation describing Newton's second law: "force equals mass times acceleration". The gravitational mass of a body, on the other hand, is the quantity that appears in Newton's law of universal gravitation. There is no *a priori* reason, in Newtonian physics, why these two masses should be equal and yet experimentally they are found to be identically so, to extremely high precision.

The **weak equivalence principle** (WEP) takes the equivalence of inertial and gravitational mass as axiomatic, stating that the inertial mass, m_I, and the gravi-

tational mass, m_G, of a body are indeed identically equal.

A profound result follows immediately from the WEP: a massive object 'freely falling' in a uniform gravitational field (e.g. a lift plummeting Earthwards after its cable has been snapped, or the interior of a spacecraft orbiting the Earth) will obey Newton's first and second laws of motion. In other words **the freely falling object inhabits an inertial frame in which all gravitational forces have disappeared**.

We call the reference frame inhabited by our freely-falling object a **local inertial frame** (LIF): the reference frame is only inertial over the region of spacetime for which the gravitational field is uniform. (The effective size of the LIF therefore depends on how rapidly the gravitational field varies as a function of position, and on how accurately we can measure the separation and velocity of freely-falling bodies).

The WEP has a number of important, and testable, physical consequences:

1. The empirically observed equality of gravitational and inertial mass is explained.

2. The acceleration of a test mass[1] in a gravitational field is entirely independent of its nature, mass and composition.

3. The path of a light ray will be bent by the gravitational field of a massive body.

4. A light ray emitted from the surface of a massive body will be *redshifted* (the effect is referred to as the **gravitational redshift**) when its wavelength is measured by a distant observer.

Experimental verification of the second point has a long and illustrious history (from Galileo to Apollo 15!). Experimental measurement of the gravitational redshift has

[1] i.e. an object with mass that is sufficiently small that it produces no measurable change in the gravitational field of the larger body towards which it is attracted.

been carried out from various astrophysical sources, as well as terrestrially via the Pound-Rebka experiment. (See e.g. Wikipedia for a clear and concise overview). We will briefly discuss gravitational light deflection later, in the context of the so-called *Schwarzschild solution* of Einstein's equations.

1.2.2 The strong equivalence principle

The strong equivalence principle (SEP) goes further and states that locally (i.e. in a LIF) *all* the laws of physics have their usual special relativistic form – except for gravity, which simply disappears. Moreover, the SEP states that there is *no* experiment that we can carry out to distinguish between a LIF which is freely-falling in a uniform gravitational field and an inertial frame which is in a region of the Universe far from any gravitating masses (and therefore well-described by the SR geometry of Minkowski spacetime).

Conversely, if we are inside a spaceship in our remote region of the cosmos free from gravity, we can *simulate* gravity by giving the rocket an acceleration; this acceleration will be indistinguishable from a uniform gravitational field with equal (but opposite) gravitational acceleration – gravity and acceleration are equivalent.

1.3 From special to general relativity

So the equivalence principles tell us that – provided our gravitational field is uniform (or can be reasonably approximated to be uniform) we can define a LIF: a coordinate system within which gravity has locally been 'transformed away', the laws of physics agree with SR and the metric of spacetime can be reduced to the simple Minkowski form of equation (3). Yet this is only the first step towards a fully covariant theory of gravity. In general gravitational fields are decidedly *not* uniform (the inertial frames defined by two freely-falling lifts in London and Sydney are very different) so we can *only* transform away their effects locally.

In order to be an effective theory of gravity, GR must provide us with the means to 'stitch together' LIFs across extended regions of spacetime, containing non-uniform gravitational fields. This stitching process turns the locally 'flat' geometry of Minkowski spacetime – which can be applied to each LIF – into the curved geometry that characterises the gravitational field of extended regions.

To describe spacetime curvature rigorously will require mathematical tools from the field of differential geometry: essentially extending the familiar description of physical quantities in terms of scalars and vectors to **tensors**. We will discuss tensors and their properties sparingly, in order not to get too bogged down in mathematics and too far removed from the physics of GR, but some discussion of tensors will be essential if we are to fully understand GR and its theoretical importance for the study of gravitational waves.

Before we introduce tensors to our mathematical toolbox, however, we can first gain some further physical insight into the relationship between gravity, acceleration and spacetime curvature by considering in a simplified way the trajectories of freely falling particles – paths known in GR as **geodesics**.

2 Introduction to Geodesic Deviation

2.1 Basic concepts of geodesics

According to Newton's laws the 'natural' trajectory of a particle which is not being acted upon by any external force is a straight line. In GR, since gravity manifests itself as spacetime curvature, these 'natural' straight line trajectories generalise to curved paths known as **geodesics**. These are defined physically as the trajectories followed by freely falling particles – i.e. particles which are not being acted upon by any *non-gravitational* external force.

Geodesics are defined *mathematically* as spacetime curves that parallel transport their own tangent vectors – concepts that we will explain in Section 3. For metric spaces (i.e. spaces on which a metric function can be defined) we can also define geodesics as extremal paths in the sense that – along the geodesic between two events E_1, and E_2, the elapsed proper time is an extremum, i.e.

$$\delta \int_{E_1}^{E_2} d\tau = 0 \tag{4}$$

Mathematically, the curvature of spacetime can be revealed by considering the deviation of neighbouring geodesics. The behaviour of geodesic deviation is represented qualitatively in Figure 2, which shows three 2-dimensional surfaces of different intrinsic curvature on which two ants are moving along neighbouring, initially parallel trajectories. On the leftmost surface, a flat piece of paper with zero intrinsic curvature, the separation of the ants remains constant as they move along their neighbouring geodesics. On the middle surface, a spherical tennis ball with positive intrinsic curvature, the ants' separation *decreases* with time – i.e. the geodesics move towards each other. On the rightmost surface, a 'saddle' shape with negative intrinsic curvature, the ants' separation *increases* with time – i.e. the geodesics move apart.

Specifically it is the **acceleration of the deviation between neighbouring**

geodesics which is a signature of spacetime curvature, or equivalently (as we would describe it in Newtonian physics) the presence of a non-uniform gravitational field. This latter point is important: geodesic deviation cannot distinguish between a *zero* gravitational field and a *uniform* gravitational field; in the latter case the acceleration of the geodesic deviation is also zero. Only for a non-uniform, or **tidal** gravitational field does the geodesic deviation accelerate. We will see later that it is precisely these tidal variations to which gravitational wave detectors are sensitive.

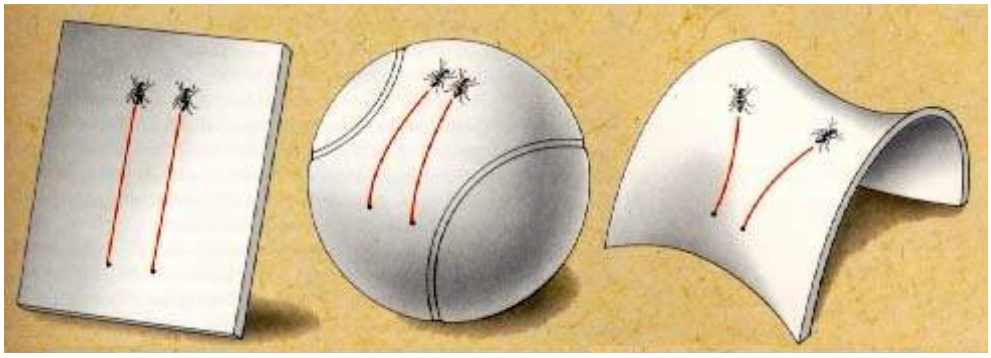


Figure 2: Geodesic deviation on surfaces of different intrinsic curvature. On a flat surface, with zero curvature, the separation of the ants remains constant – i.e. neighbouring geodesics remain parallel. On the surface of a sphere, with positive curvature, the separation of the ants decreases – i.e. neighbouring geodesics converge. On the surface of a saddle, with negative curvature, the separation of the ants increases – i.e. neighbouring geodesics diverge.

Before we consider a full GR description of the relationship between geodesic deviation and spacetime curvature, it is instructive to consider a rather simpler illustration: a Newtonian description of the behaviour of neighbouring free-falling particles in a non-uniform gravitational field.

2.2 Geodesic deviation in Newtonian gravity

Figure 3 is a (hugely exaggerated) cartoon illustration of two test particles that are initially suspended at the same height above the Earth's surface (assumed to be spherical) and are released from rest. According to Newtonian physics, the separation of these test particles will reduce as they freely fall towards the Earth, because they are falling in a non-uniform gravitational field. (The gravitational force on each particle is directed towards the centre of the Earth, which means that their acceleration vectors are not parallel).

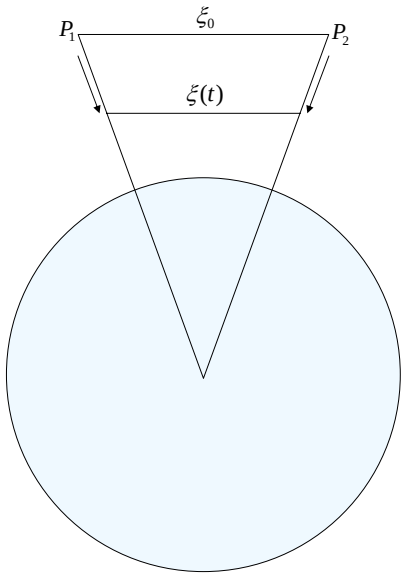


Figure 3: Cartoon illustrating how in Newtonian physics the separation of test particles will change in time if they are falling freely in a non-uniform gravitational field.

Suppose the initial separation of the test particles is ξ_0 and their distance from the centre of the Earth is r_0, while after some time t their separation is $\xi(t)$ and their distance from the centre of the Earth is $r(t)$. From similar triangles we can see that

$$\frac{\xi(t)}{r(t)} = \frac{\xi_0}{r_0} = k \tag{5}$$

where k is a constant. Taking derivatives with respect to time gives

$$\ddot{\xi} = k\ddot{r} = -\frac{kGM}{r^2} \tag{6}$$

where M is the mass of the Earth. Substituting for $k = \xi/r$ gives

$$\ddot{\xi} = -\frac{\xi}{r}\,\frac{GM}{r^2} = -\frac{GM\xi}{r^3} \tag{7}$$

If the test particles are released close to the Earth's surface then $r \approx R$, where R is the radius of the Earth, so $\ddot{\xi} = -GM\xi/R^3$. We can re-define our time coordinate and express this equation as

$$\frac{d^2\xi}{d(ct)^2} = -\frac{GM}{R^3c^2}\,\xi \tag{8}$$

Notice that in equation (8) the coefficient of ξ on the right hand side has dimensions [length]$^{-2}$. Evaluated at the Earth's surface this quantity equals about 2×10^{-23} m^{-2}.

2.3 Intrinsic curvature and the gravitational field

We can begin to understand the physical significance of equation (8) by making use of a 2-dimensional analogy. Suppose P_1 and P_2 are on the equator of a sphere of radius a (see Figure 4). Consider two geodesics – 'great circles' of constant longitude perpendicular to the equator, passing through P_1 and P_2, and separated by a distance ξ_0 at the equator. The arc distance along each geodesic is denoted by s and the separation of the geodesics at s is $\xi(s)$.

Evidently the geodesic separation is not constant as we change s and move towards the north pole N. We can write down the differential equation which governs the change in this geodesic separation. If the (small) difference in longitude between the two geodesics is $d\phi$ (in radians) then $\xi_0 = a\,d\phi$. At latitude θ (again, in radians) corresponding to arc length s, on the other hand, the geodesic separation is

$$\xi(s) = a\cos\theta\,d\phi = \xi_0\cos\theta = \xi_0\cos s/a \tag{9}$$

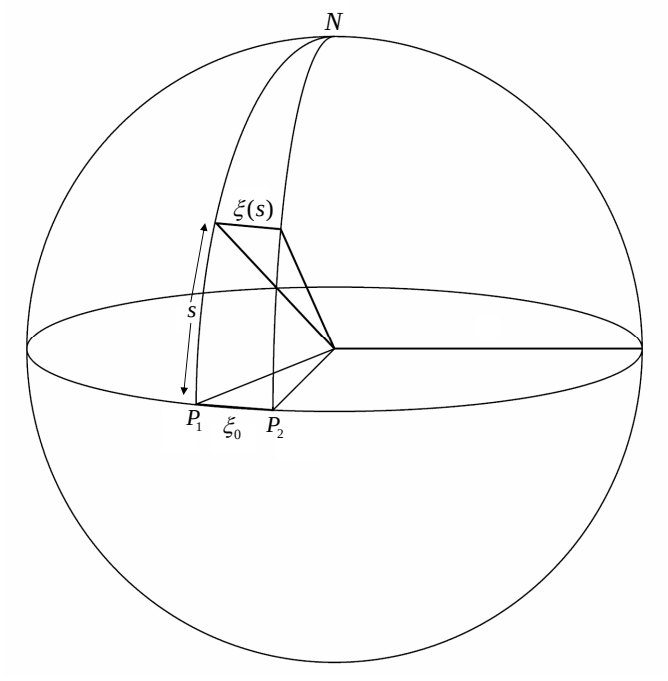


Figure 4: Illustration of the change in geodesic separation as we move along great circles of constant longitude on the surface of a sphere.

Differentiating $\xi(s)$ twice with respect to s yields

$$\frac{d^2\xi}{ds^2} = -\frac{1}{a^2}\xi \tag{10}$$

Comparing equations (8) and (9) we see that in some sense the quantity

$$\mathcal{R} = \left\{ \frac{GM}{R^3c^2} \right\}^{-\frac{1}{2}} \tag{11}$$

represents the *radius of curvature of spacetime* at the surface of the Earth. Evaluating this radius for the Earth we find that $\mathcal{R} \sim 2 \times 10^{11}$ m. The fact that this value is so much larger than the physical radius of the Earth tells us that spacetime is 'nearly' flat in the vicinity of the Earth – i.e. the Earth's gravitational field is rather weak. (By contrast, if we evaluate $\mathcal{R}$ for e.g. a white dwarf or neutron star then we see evidence that their gravitational fields are much stronger).

We will soon return to the subject of spacetime curvature and consider how it is expressed in GR. Before we can do that, however, we need to expand the mathematical armoury at our disposal.

3 A mathematical toolbox for GR

To deal rigorously with the geometrical properties of curved spacetimes we need to introduce a number of mathematical concepts and tools. We begin with the concept of a **manifold**.

3.1 Manifolds

A manifold is a continuous space which is locally flat. More generally we can regard a manifold as any set which can be continuously parametrised: the number of independent parameters is the *dimension* of the manifold, and the parameters themselves are the *coordinates* of the manifold. A *differentiable manifold* is one which is both continuous and differentiable. This means that we can define a scalar function (or *scalar field*) – ϕ, say – at each point of the manifold, and that ϕ is differentiable.

> **Mathematical Aside:** The formal mathematics of defining coordinates need not concern us in these lectures, but the interested reader can find useful discussions in any introductory textbook on differential geometry. Loosely speaking, it involves covering the points of the manifold by a collection of *open sets*, $\mathbf{U^i}$, each of which is mapped onto $\mathbf{R^n}$ by a one-to-one mapping, ϕ^i. The pair $(\mathbf{U^i}, \phi^i)$ is called a *chart*, and the collection of charts an *atlas*. One can think of each chart as defining a different coordinate system.

In GR we are concerned with a particular class of differentiable manifolds known as **Riemannian manifolds**. A Riemannian manifold is a differentiable manifold on which a *distance*, or *metric*, has been defined.

We can see why the mathematics of Riemannian manifolds are appropriate for GR. According to the WEP spacetime is locally flat and the interval between spacetime events is described by a metric, following equation (2). Moreover, any dynamical

theory of gravity will involve space and time derivatives; hence our description of spacetime should be differentiable. Riemannanian manifolds meet all three criteria.

3.2 Scalar functions on a manifold

One can define a function, f, on a manifold, $\mathbf{M}$. At any point, P, of the manifold the function takes a real value

$$f : \mathbf{M} \rightarrow \mathbf{R} \tag{12}$$

In a particular coordinate representation, P has coordinates $\{x^1, x^2, ..., x^n\}$. We may then write simply

$$f_P \quad = \quad f(x^1, x^2, ..., x^n) \tag{13}$$

Here f is called a *scalar* function; this means that its numerical value at each point of the manifold is the same real number, no matter which coordinate representation is used.

3.3 Vectors in curved spaces

The intuitive picture of a vector which we have learned in elementary maths and physics courses is based on the simple idea of an arrow representing a *displacement* between two points in space. Moreover a vector, $\vec{a}$, exists independently of our choice of coordinate system, but the *components* of $\vec{a}$ take different values in different coordinate systems, and we can define a transformation law for the components of the vector.

Consider, for example, the displacement vector, $\Delta\vec{x}$, with components Δx^μ and $\Delta x'^\mu$ in an unprimed and primed coordinate system respectively. Then, by the chain rule for differentiation

$$\Delta x'^\mu = \frac{\partial x'^\mu}{\partial x^\alpha} \Delta x^\alpha \tag{14}$$

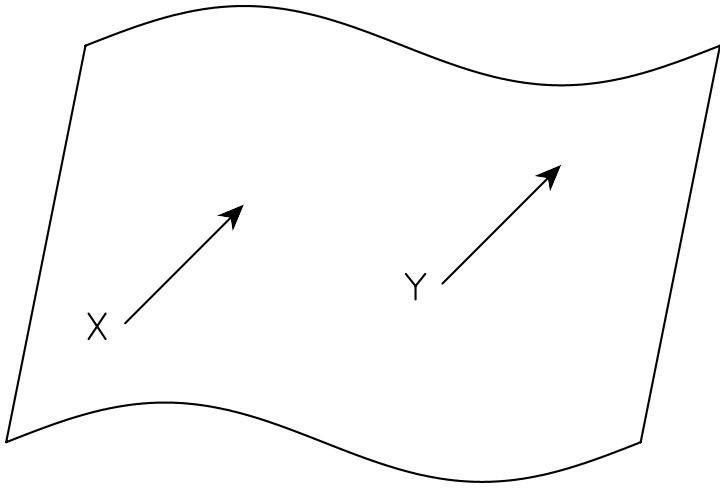


Figure 5: Vectors on a curved manifold. In general the components of $\Delta\vec{x}$ at X and of $\Delta\vec{y}$ at Y will be different, because they are defined at different points in the manifold.

Consider now two displacement vectors, $\Delta\vec{x}$ and $\Delta\vec{y}$. How can we decide if $\Delta\vec{x}$ and $\Delta\vec{y}$ are equal when – as shown in Figure 5 – they are defined at different points on our manifold? For vectors in flat space with Cartesian coordinates, for example, we can simply 'translate' $\Delta\vec{y}$ to X and compare the components of $\Delta\vec{y}$ with those of $\Delta\vec{x}$. This will **not** be valid for a general curved manifold, however, because the coefficients of the transformation law in equation (14) are in general functions of position. In other words, the transformation law between the primed and unprimed coordinate systems is in general *different* at different points of the manifold. Thus, it is *not* enough to define the components of a vector; we also need to specify the point of the manifold at which the vector (and its components) are defined.

The fact that the transformation law coefficients of equation (14) are in general functions of position also means that we have no 'universal' set of coordinate basis vectors on a curved manifold, as is the case for Euclidean space. There is, however, a means of defining a natural set of basis vectors for each point of the manifold

which allows us to develop a more general picture of what we mean by a vector – and one which is equally valid in a curved spacetime.

3.4 Tangent vectors

Suppose we have a scalar function, ϕ, defined at a point, P, of a Riemannian manifold, where P has coordinates $\{x^1, x^2, ..., x^n\}$ in some coordinate system. Since our manifold is differentiable we can evaluate the derivative of ϕ with respect to each of the coordinates, x^i, for $i = 1, ..., n$. In fact, since ϕ is completely arbitrary, we can think of the derivatives as a set of n 'operators', denoted by

$$\frac{\partial}{\partial x^i}$$

These operators can act on any scalar function, ϕ, and yield the rate of change of the function with respect to the x^i.

We can now define a **tangent vector** at point, P, as a linear operator of the form

$$a^\mu \frac{\partial}{\partial x^\mu} \equiv a^1 \frac{\partial}{\partial x^1} + a^2 \frac{\partial}{\partial x^2} + ... + a^n \frac{\partial}{\partial x^n} \tag{15}$$

(Note the use of the summation convention). This tangent vector operates on any function, ϕ, and essentially gives the rate of change of the function – or the *directional derivative* – in a direction which is defined by the numbers $(a^1, a^2, ..., a^n)$. We can define the addition of two tangent vectors in the obvious way

$$a^\mu \frac{\partial}{\partial x^\mu} + b^\mu \frac{\partial}{\partial x^\mu} = (a^\mu + b^\mu) \frac{\partial}{\partial x^\mu} \tag{16}$$

> **Mathematical Aside:** With this straightforward definition of addition, a little formal mathematics easily shows that the set of all tangent vectors form a *vector space*

Thus, the operator,

$$a^\mu \frac{\partial}{\partial x^\mu}$$

behaves like a vector, the components of which are $(a^1, a^2, ..., a^n)$. We can therefore write

$$\vec{a} = a^\mu \frac{\partial}{\partial x^\mu} \tag{17}$$

The n operators $\frac{\partial}{\partial x^\mu}$ can be thought of as forming a set of basis vectors, $\{\vec{e_\mu}\}$, spanning the vector space of tangent vectors at P.

What exactly do these basis vectors represent? We can find a simple geometrical picture for the $\vec{e_\mu}$ by first crystallising the notion of a *curve*, C, defined on our manifold. Our intuitive notion of a curve is simply of a connected series of points on the manifold; in the mathematical literature, however, we call this a *path*, and the term *curve* is instead reserved for the particular case of a path which has been *parametrised*.

Thus, a curve is a function which maps an interval of the real line into the manifold. Putting this more simply, a curve is a path with a real number (s, say) associated with each point of the path; we call s the parameter of the curve. Note also that once we choose a coordinate system each point on the curve has coordinates, $\{x^\mu\}$, which may also be expressed as functions of the parameter, s, i.e.

$$x^\mu = x^\mu(s) \quad \mu = 1, ..., n \tag{18}$$

Once we specify our coordinate system, we can consider a particular set of curves which use the *coordinates themselves* as their parameter. For example, point P with coordinates $\{x^1, x^2, ..., x^n\}$ lies on the n curves which we obtain by allowing only the value of x^i to vary along the i^{th} curve ($i = 1, ..., n$) and fixing all other coordinate values to be equal to their values at P. (To visualise a simple example, think of circles of equal latitude and longitude on the 2-sphere manifold). The basis vector, $\vec{e_i} \equiv \frac{\partial}{\partial x^i}$ can be thought of simply as the *tangent* to the i^{th} curve. This geometrical

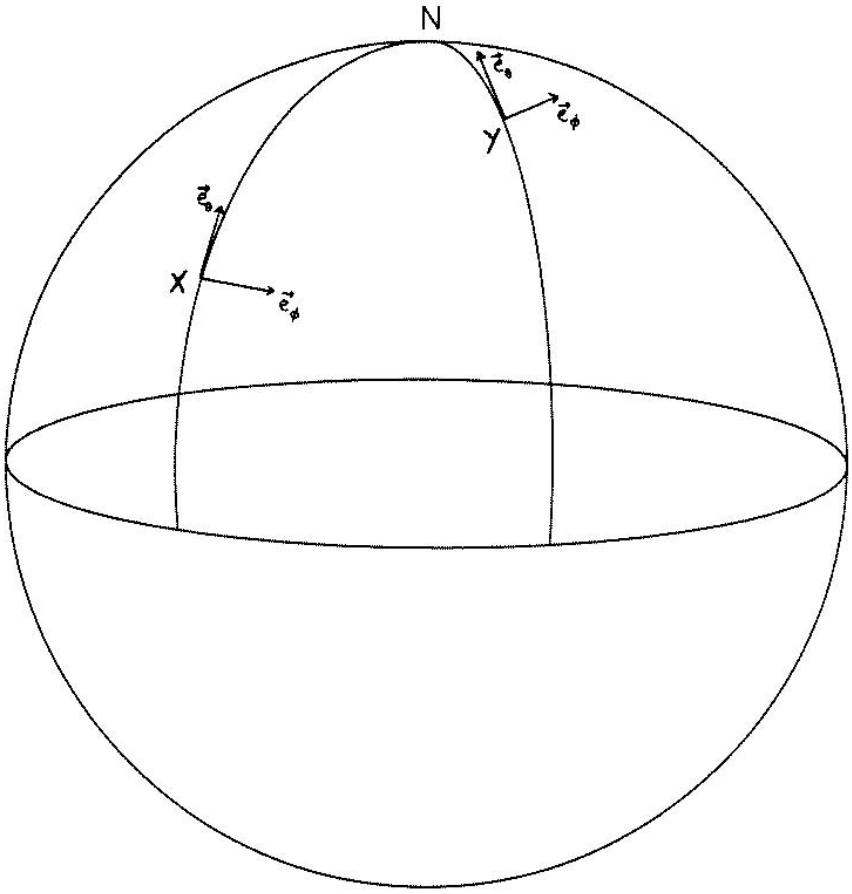


Figure 6: Illustration of the coordinate basis vectors defined at two points X and Y on the surface of a sphere. Note that the basis vectors $\vec{e_\phi}$ and $\vec{e_\theta}$ are *different* at points X and Y.

picture is illustrated in Figure 6, again for the straightforward example of the 2-sphere. Note that the basis vectors $\vec{e_\phi}$ and $\vec{e_\theta}$ are *different* at points X and Y of the manifold.

And what of a more general curve in the manifold? Here we simply connect the notion, introduced above, of a tangent vector as a directional derivative to our straightforward geometrical picture of a tangent to a curve. Figure 7 (adapted from Schutz) shows a curve, with parameter s, and with tangent vectors drawn at points with different parameter values. Suppose the coordinates of the points on the curve are $\{x^\mu(s)\}$, for $\mu = 1, ..., n$. Then the components, T^μ, of the tangent vector with

respect to the basis $\{\vec{e_\mu}\} \equiv \{\frac{\partial}{\partial x^\mu}\}$ are simply given by

$$T^\mu = \frac{dx^\mu}{ds} \tag{19}$$

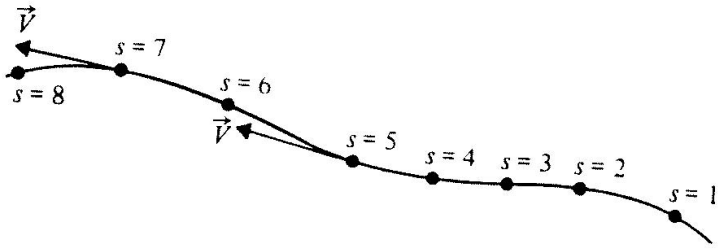


Figure 7: Schematic representation of a curve, parametrised by s, and showing tangent vectors drawn at $s = 5$ and $s = 7$.

To sum up, we can represent vectors as tangent vectors of curves in our manifold. Once we have specified our coordinate system, we can write down the components of a vector defined at any point of the manifold with respect to the natural basis generated by the derivative operators $\{\frac{\partial}{\partial x^\mu}\}$ at that point. A vector **field** can then be defined by assigning a tangent vector at *every* point of the manifold.

3.5 Transformation law for vectors

Suppose we change to a new coordinate system $\{x'^1, x'^2, ..., x'^n\}$. Our basis vectors are now

$$\vec{e'_\mu} \equiv \frac{\partial}{\partial x'^\mu}. \tag{20}$$

How do the components, $\{a^1, a^2, ..., a^n\}$, transform in our new coordinate system? To see how the law arises within the framework of our tangent vector description, let the vector $\vec{a}$ operate on an arbitrary scalar function, ϕ. Then

$$\vec{a}(\phi) = a^\nu \frac{\partial \phi}{\partial x^\nu} \tag{21}$$

By the chain rule for differentiation we may write this as

$$\vec{a}(\phi) = a^{\nu} \frac{\partial x'^{\mu}}{\partial x^{\nu}} \frac{\partial \phi}{\partial x'^{\mu}} \tag{22}$$

However, if we write $\vec{a}$ directly in terms of coordinate basis $\{\vec{e'_{\mu}}\} = \{\frac{\partial}{\partial x'^{\mu}}\}$, we have

$$\vec{a}(\phi) = a'^{\mu} \frac{\partial \phi}{\partial x'^{\mu}} \tag{23}$$

Comparing equation (22) with (23) we see that

$$a'^{\mu} = \frac{\partial x'^{\mu}}{\partial x^{\nu}} a^{\nu} \tag{24}$$

Thus the components of the tangent vector transform according to equation (24). We call this equation the transformation law for a **contravariant vector**, and say that the components of $\vec{a}$ transform **contravariantly**. (The term 'contravariant' is used to distinguish these vectors from another type of geometrical object – covariant vectors or 'covectors' – which we will meet in the next subsection. The more modern name for covariant vectors, however, is 'one-forms', and we will generally adopt that name in order to avoid this source of ambiguity). We denote the components of a contravariant vector using superscripts.

Equation (24) is the prototype transformation law for *any* contravariant vector. Any set of n components, A^{μ}, which can be evaluated in any coordinate system, and which transform according to the transformation law of equation (24), we call a contravariant vector.

3.6 Transformation law for one-forms

What is the relationship between the basis vectors $\vec{e'_{\mu}}$ and $\vec{e_{\mu}}$ in the primed and unprimed coordinate systems? From equation (20) we have

$$\vec{e'_{\mu}} = \frac{\partial x^{\nu}}{\partial x'^{\mu}} \vec{e_{\nu}} \tag{25}$$

Thus we see that the basis vectors do *not* transform in the same way as the components of a contravariant vector. This should not be too surprising, since the transformation of a basis and the transformation of components are different things: the former is the expression of *new* vectors in terms of *old* vectors; the latter is the expression of the *same* vector in terms of a new basis.

In fact, the form of the transformation in equation (25) is the same as the transformation law for another type of geometrical object, which we call a **covariant vector**, **covector**, or (in more modern literature) a **one-form**. Any set of n components, A_μ, which can be evaluated in any coordinate system, is said to be a one-form if the components transform according to the equation

$$A'_\mu = \frac{\partial x^\nu}{\partial x'^\mu} A_\nu \tag{26}$$

One can simply regard equation (26) as *defining* a one-form. Many modern textbooks on differential geometry, however, begin by defining a one-form as a linear mapping which acts on a vector to give a real number. Starting from this definition one can then arrive at equation (26).

One-forms are usually denoted by a tilde above a symbol, just as vectors are denoted by an arrow above a symbol. Thus e.g. $\tilde{p}(\vec{a})$ is a real number.

3.7 Transformation law for tensors

Having defined what we mean by vectors and one-forms, in terms of how their components transform under a general coordinate transformation, we can now extend our definition to the more general class of geometrical object which we call *tensors*.

A tensor of type (l, m), defined on an n dimensional manifold, is a linear operator which maps l one-forms and m (contravariant) vectors into a real number (i.e.

scalar). Such a tensor has a total of n^{l+m} components.

The transformation law for a general (l, m) tensor follows from its linearity, and from the transformation laws for a vector and one-form, in order that the scalar quantity obtained when the tensor operates on l one-forms and m vectors is independent of one's choice of coordinate system. We can write this general transformation law as follows

$$A'^{u_1\, u_2 \ldots u_l}_{r_1\, r_2 \ldots r_m} = \frac{\partial x'^{u_1}}{\partial x^{t_1}} \ldots \frac{\partial x'^{u_l}}{\partial x^{t_l}} \frac{\partial x^{q_1}}{\partial x'^{r_1}} \ldots \frac{\partial x^{q_m}}{\partial x'^{r_m}} A^{t_1\, t_2 \ldots t_l}_{q_1\, q_2 \ldots q_m} \tag{27}$$

This rather intimidating equation appears much more straightforward for some specific cases. First note that a contravariant vector is in fact a $(1, 0)$ tensor (since it operates on a one-form to give a scalar). Similarly a one-form is a $(0, 1)$ tensor (and more trivially a *scalar* is a $(0, 0)$ tensor).

A $(2, 0)$ tensor, say T^{ij}, is called a *contravariant* tensor of rank 2 and transforms according to the transformation law

$$T'^{ij} = \frac{\partial x'^i}{\partial x^k} \frac{\partial x'^j}{\partial x^l} T^{kl} \tag{28}$$

A $(0, 2)$ tensor, say B_{ij}, is called a *covariant* tensor of rank 2, and transforms according to the law

$$B'_{ij} = \frac{\partial x^k}{\partial x'^i} \frac{\partial x^l}{\partial x'^j} B_{kl} \tag{29}$$

An important example of a $(0, 2)$ tensor is the *metric tensor*, $g_{\mu\nu}$, which we already met in Section 1. This is a symmetric tensor, i.e.

$$g_{\mu\nu} = g_{\nu\mu} \quad \text{for all } \mu, \nu \tag{30}$$

We can now see that the form of equation (2) makes sense: the small displacements dx^μ and dx^ν transform as contravariant vectors; the metric tensor operates on these

vectors to give a scalar (the interval) which is invariant.

A tensor which has both upper and lower indices, which means that it has both contravariant and covariant terms in its transformation law, is known as a *mixed tensor*. The simplest example (after the trivial case of a $(0,0)$ tensor) is a $(1,1)$ tensor, D^i_j, say. Its transformation law is

$$D'^i_j = \frac{\partial x'^i}{\partial x^k}\frac{\partial x^l}{\partial x'^j}D'^k_l \tag{31}$$

An important example of a $(1,1)$ tensor is the Kronecker delta, δ^i_j, which has the property (in *any* coordinate system)

$$\delta^l_j = 1 \text{ when } j = l, \text{ and 0 otherwise} \tag{32}$$

3.8 Contraction of tensors

We can take the **inner product**, or **contraction** of a vector and one-form; i.e. we form the quantity A^iB_i (where, as usual, the summation convention is implied). This quantity is invariant in the sense that

$$A'^jB'_j = A^iB_i \tag{33}$$

We can generalise the operation of contraction to the case of any two tensors, and over an arbitrary number of indices, provided that an equal number of upper and lower indices are selected. In general, contraction over k indices will produce from a tensor of type (l,m) a new tensor of type $(l-k, m-k)$. For example, the contraction of the two tensors G^{ijk}_{lm} and R^s_{tu} over the indices i and t, j and u and l and s will give the $(1,1)$ tensor $G^{ijk}_{lm}R^l_{ij}$, where now only the indices k and m are free indices.

3.9 Raising and lowering indices

Given any contravariant vector A^μ it is possible to define, via the metric tensor, an associated one-form, which we denote as A_ν and which is defined by

$$A_\nu = g_{\mu\nu} A^\mu \tag{34}$$

This operation is often called *lowering the index*.

Similarly by using g^{ij} we can raise the index of a covariant quantity B_i to obtain a contravariant quantity B^i, viz.

$$B^i = g^{ij} B_j \tag{35}$$

The process of raising or lowering indices can be carried out with tensors of any rank and type. For example

$$D^{ijk\cdot\cdot}_{lm} = g_{lp} g_{mq} D^{ijkpq} \tag{36}$$

Some care must be taken in positioning the indices. The dots have been placed here to indicate the indices over which contraction has taken place, although in general we shall omit the dots and just write D^{ijk}_{lm}. Note that $D^{\cdot\cdot ijk}_{lm}$ defined by

$$D^{\cdot\cdot ijk}_{lm} = g_{lp} g_{mq} D^{pqijk} \tag{37}$$

is not the same as $D^{ijk\cdot\cdot}_{lm}$ unless D^{ijkpq} possesses some symmetry.

The *magnitude* of a vector with components A^μ is $g_{\mu\nu} A^\mu A^\nu$, which is of course invariant, since $g_{\mu\nu}$ is a $(0,2)$ tensor and A^μ and A^ν are both $(1,0)$ tensors. Notice that

$$g_{\mu\nu} A^\mu A^\nu = A_\mu A^\nu = g^{\mu\nu} A_\nu A_\nu \tag{38}$$

The quantity $g_{\mu\nu} A^\mu B^\nu$ is the scalar product of the two vectors.

3.10 Covariant differentiation and parallel transport

Any dynamical physical theory must deal in time varying quantities and, if this theory is also to be relativistic, spatially varying quantities too. Since GR is a covariant theory, we are confronted with the problem of constructing quantities that represent rates of change, but which can be defined in any coordinate system. In other words, we need to define a *derivative* which transforms covariantly under a general coordinate transformation: we call this a **covariant derivative**.

For any scalar function, say ϕ, defined on the manifold, the partial derivative

$$\phi_{,\nu} \equiv \frac{\partial \phi}{\partial x^\nu}$$

transforms as a $(0,1)$ tensor, i.e.

$$\phi'_{,\nu} \equiv \frac{\partial \phi'(x')}{\partial x'^\nu} = \frac{\partial x^\mu}{\partial x'^\nu}\frac{\partial \phi(x)}{\partial x^\mu} \equiv \frac{\partial x^\mu}{\partial x'^\nu}\phi_{,\mu} \tag{39}$$

On the other hand, partial derivatives of the components of a contravariant vector transform as:

$$A'^i_{,j} = \frac{\partial x'^i}{\partial x^k}\frac{\partial x^l}{\partial x'^j}A^k_{,l} + \frac{\partial x^l}{\partial x'^j}\frac{\partial^2 x'^i}{\partial x^l \partial x^k}A^k \tag{40}$$

The presence of the second term of equation (40),

$$\frac{\partial x^l}{\partial x'^j}\frac{\partial^2 x'^i}{\partial x^l \partial x^k}A^k,$$

is the reason why $A^i_{,j}$ does not transform as a tensor. The root of the problem is that computing the derivative involves subtracting vectors at two neighbouring points, at each of which the transformation law will in general be different. To overcome this problem we need to transport one of the vectors to the neighbouring point, so that they are subtracted at the *same* point in the manifold. This can be achieved by the

so-called **parallel transport** procedure, which transports the vector components along a curve in the manifold in a manner which preserves the angle between vectors.

Suppose we have a point with coordinates x^i while its neighbour has coordinates $x^i + dx^i$. It is easiest to consider first the parallel transport of the components of a *covariant* vector, or one-form, from $x^i + dx^i$ to x^i. Suppose that the change, δB_k, in the components will be a linear function of the original components, B_j, and of the displacement dx^i, so we can write

$$\delta B_k = -\Gamma^j_{ik} B_j dx^i \tag{41}$$

The coefficients Γ^j_{ik} are called the **Christoffel symbols** and they describe how the basis vectors at different points in the manifold change as one moves across the manifold, i.e.

$$\frac{\partial \vec{e}_\alpha}{\partial x^\beta} = \Gamma^\mu_{\alpha\beta}\ \vec{e}_\mu \tag{42}$$

It then follows that the covariant derivative of a one-form may be written as

$$B_{i;k} = B_{i,k} - \Gamma^j_{ik} B_j \tag{43}$$

i.e. $B_{i,k} - \Gamma^j_{ik} B_j$ transforms as a $(0,2)$ tensor. Given that the covariant and partial derivatives of a scalar are identical, it is then straightforward to show that the covariant derivative of a vector takes the form

$$A^i_{;k} = A^i_{,k} + \Gamma^i_{jk} A^j \tag{44}$$

which transforms as a $(1,1)$ tensor. The generalisation of equations (43) and (44) to tensors of arbitrary rank is straightforward.

It is also straightforward to show that the covariant derivative of the metric tensor is equal to zero, i.e. for all α, β, γ

$$g_{\alpha\beta;\gamma} = 0 \quad \text{and} \quad g^{\alpha\beta}_{;\gamma} = 0 \tag{45}$$

and this result can be used to obtain an expression for the Christoffel symbols in terms of the metric and its partial derivatives

$$\Gamma^i_{jk} = \frac{1}{2} g^{il}(g_{lj,k} + g_{lk,j} - g_{jk,l}) \tag{46}$$

3.11 The geodesic equation

As we discussed in Section 2, material particles not acted on by forces other than gravitational forces have worldlines that are **geodesics**. Similarly photons also follow geodesics. One can define a geodesic as an extremal path between two events, in the sense that the proper time along the path joining the two events is an extremum. Equivalently, one can define a geodesic as a **curve along which the tangent vector to the curve is parallel-transported**. In other words, if one parallel transports a tangent vector along a geodesic, it remains a tangent vector.

3.11.1 Geodesics of material particles

The worldline of a material particle may be written with the proper time, τ, as parameter along the worldline. The **four velocity** of the particle is the tangent vector to the worldline. One may show that the geodesic equation for the particle is

$$\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{\alpha\beta}\frac{dx^\alpha}{d\tau}\frac{dx^\beta}{d\tau} = 0 \tag{47}$$

3.11.2 Geodesics of photons

For photons, the proper time τ cannot be used to parametrise their worldline, since $d\tau$ for a photon is zero. We need to find some other way to parametrise the worldline of the photon (e.g. an angular coordinate along the trajectory) so that, with respect to this parameter (λ, say) the geodesic equation is satisfied, i.e.

$$\frac{d^2x^\mu}{d\lambda^2} + \Gamma^\mu_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} = 0 \tag{48}$$

The parameter λ is known as an **affine parameter**.

4 Spacetime Curvature in GR

4.1 The Riemann Christoffel tensor

In our simple example of Section 2 the curvature of our spherical surface depended on only a single parameter: the radius a of the sphere. More generally, the curvature of spacetime can be described by the **Riemann Christoffel tensor**, **R** (often also referred to simply as the **Riemann tensor**), which depends on the metric and its first and second order partial derivatives. The functional form of the Riemann Christoffel tensor can be derived in several different ways:

1. by parallel transporting of a vector around a closed loop in our manifold
2. by considering the commutator of the second order covariant derivative of a vector field
3. by computing the deviation of two neighbouring geodesics in our manifold

In view of our previous discussion we will focus on the third method for deriving **R**, although there are close mathematical similarities between all three methods. Also, the first method has a particularly simple pictorial representation. Figure 8 (adapted from Schutz) shows the result of parallel transporting a vector around a closed triangle. Panel (a) shows a flat surface of zero curvature; when the vector is parallel transported from A to B to C and then back to point A, the final vector is parallel to the original one. Panel (b) on the other hand shows a spherical surface of positive curvature; when we parallel transport a vector from A to B to C and back to A, the final vector is *not* parallel to the original one. We can express the net change in the components of the vector, after transport around the closed loop, in terms of the Riemann Christoffel tensor.

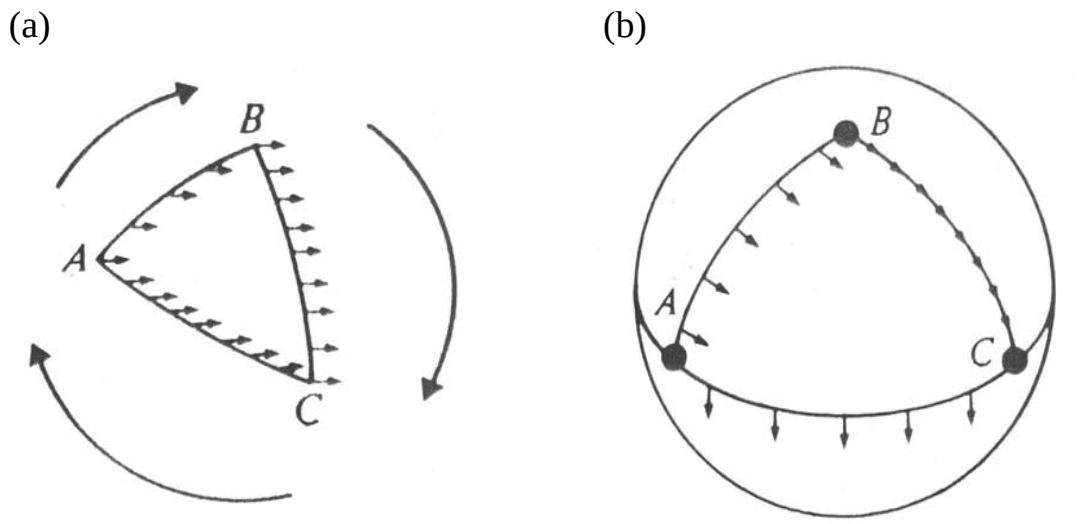


Figure 8: Parallel transport of a vector around a closed curve in a flat space (panel a) and a curved space (panel b).

4.2 Acceleration of the geodesic deviation

To see how we can derive the form of the Riemann Christoffel tensor via method 3, consider two test particles (labelled 1 and 2) moving along nearby geodesics. Let $\xi^\mu(\tau)$ denote the (infinitesimal) separation of the particles at proper time τ, so that

$$x_2^\mu(\tau) = x_1^\mu(\tau) + \xi^\mu(\tau) \tag{49}$$

Now the worldline of each particle is described by the geodesic equation, i.e.

$$\frac{d^2x_1^\mu}{d\tau^2} + \Gamma^\mu_{\alpha\beta}(x_1)\frac{dx_1^\alpha}{d\tau}\frac{dx_1^\beta}{d\tau} = 0 \tag{50}$$

and

$$\frac{d^2x_2^\mu}{d\tau^2} + \Gamma^\mu_{\alpha\beta}(x_2)\frac{dx_2^\alpha}{d\tau}\frac{dx_2^\beta}{d\tau} = 0 \tag{51}$$

Note also that, by Taylor expanding the Christoffel symbols at x_1 in terms of ξ, we may write

$$\Gamma^\mu_{\alpha\beta}(x_2) = \Gamma^\mu_{\alpha\beta}(x_1 + \xi) = \Gamma^\mu_{\alpha\beta}(x_1) + \Gamma^\mu_{\alpha\beta,\gamma}\xi^\gamma \tag{52}$$

Subtracting equation (50) from equation (51), substituting from equation (52) and keeping only terms up to first order in ξ yields the following equation for the acceleration of ξ^μ (dropping the subscript 1)

$$\frac{d^2\xi^\mu}{d\tau^2} + \Gamma^\mu_{\alpha\beta} v^\alpha \frac{d\xi^\beta}{d\tau} + \Gamma^\mu_{\alpha\beta} v^\beta \frac{d\xi^\alpha}{d\tau} + \Gamma^\mu_{\alpha\beta,\gamma} \xi^\gamma v^\alpha v^\beta = 0 \tag{53}$$

where we have used the fact that $v^\alpha \equiv dx^\alpha/d\tau$.

Equation (53) is not a tensor equation, since the Christoffel symbols and their derivatives do not transform as a tensor. We can develop the corresponding covariant expression, however, by taking *covariant* derivatives of the geodesic deviation. To this end, consider the covariant derivative of a general vector field $\vec{\mathbf{A}}$ along a geodesic with tangent vector $\vec{\mathbf{v}} \equiv d\vec{\mathbf{x}}/d\tau$. We write this covariant derivative as $\nabla_{\vec{\mathbf{v}}}\vec{\mathbf{A}}$, or in component form, introducing the covariant operator $D/D\tau$

$$\frac{DA^\mu}{D\tau} = v^\beta A^\mu_{;\beta} = \frac{dA^\mu}{d\tau} + \Gamma^\mu_{\alpha\beta} A^\alpha \frac{dx^\beta}{d\tau} \tag{54}$$

Now consider the second covariant derivative of the geodesic deviation, evaluated along the geodesic followed by test particle 1. From equation (54) in component form

$$\frac{D\xi^\mu}{D\tau} = \frac{d\xi^\mu}{d\tau} + \Gamma^\mu_{\alpha\beta} \xi^\alpha \frac{dx^\beta}{d\tau} \tag{55}$$

It then follows that

$$\frac{D^2\xi^\mu}{D\tau^2} = \frac{D}{D\tau}\left(\frac{D\xi^\mu}{D\tau}\right) = \frac{d}{d\tau}\left(\frac{D\xi^\mu}{D\tau}\right) + \Gamma^\mu_{\sigma\delta} \frac{D\xi^\sigma}{D\tau} v^\delta \tag{56}$$

Substituting again for $D\xi^\mu/D\tau$ we obtain

$$\frac{D^2\xi^\mu}{D\tau^2} = \frac{d}{d\tau}(\frac{d\xi^\mu}{d\tau} + \Gamma^\mu_{\alpha\beta}\xi^\alpha v^\beta) + \Gamma^\mu_{\sigma\delta}(\frac{d\xi^\sigma}{d\tau} + \Gamma^\sigma_{\alpha\beta}\ \xi^\alpha v^\beta)v^\delta \tag{57}$$

Now, applying the product rule for differentiation

$$\frac{d}{d\tau}(\Gamma^\mu_{\alpha\beta}\xi^\alpha v^\beta) = \Gamma^\mu_{\alpha\beta,\gamma}\frac{dx^\gamma}{d\tau}\xi^\alpha v^\beta + \Gamma^\mu_{\alpha\beta}\frac{d\xi^\alpha}{d\tau}v^\beta + \Gamma^\mu_{\alpha\beta}\xi^\alpha\frac{dv^\beta}{d\tau} \tag{58}$$

Since each particle's worldline is a geodesic we also know that

$$\frac{dv^\beta}{d\tau} = \frac{d^2x^\beta}{d\tau^2} = -\Gamma^\beta_{\sigma\delta}v^\sigma v^\delta \tag{59}$$

where again we have written $v^\beta = dx^\beta/d\tau$.

Substituting equations (58) and (59) into (57) and permuting some repeated indices, we obtain

$$\begin{aligned}\frac{D^2\xi^\mu}{D\tau^2} = \frac{d^2\xi^\mu}{d\tau^2} + \Gamma^\mu_{\alpha\beta,\gamma}v^\gamma\xi^\alpha v^\beta + \Gamma^\mu_{\alpha\beta}\frac{d\xi^\alpha}{d\tau}v^\beta + \Gamma^\mu_{\sigma\delta}\frac{d\xi^\sigma}{d\tau}v^\delta \\ +(\Gamma^\mu_{\beta\delta}\Gamma^\beta_{\alpha\sigma} - \Gamma^\mu_{\alpha\beta}\Gamma^\beta_{\alpha\delta})v^\sigma v^\delta\xi^\alpha\end{aligned} \tag{60}$$

However we can now use the result we obtained in equation (53) i.e.

$$\frac{d^2\xi^\mu}{d\tau^2} = -(\Gamma^\mu_{\alpha\beta}v^\alpha\frac{d\xi^\beta}{d\tau} + \Gamma^\mu_{\alpha\beta}v^\beta\frac{d\xi^\alpha}{d\tau} + \Gamma^\mu_{\alpha\beta,\gamma}\xi^\gamma v^\alpha v^\beta) \tag{61}$$

so that finally we obtain the compact expression

$$\frac{D^2\xi^\mu}{D\tau^2} = R^\mu{}_{\alpha\beta\gamma}\,v^\alpha\,v^\beta\,\xi^\gamma \tag{62}$$

where

$$R^\mu{}_{\alpha\beta\gamma} = \Gamma^\sigma_{\alpha\gamma}\Gamma^\mu_{\sigma\beta} - \Gamma^\sigma_{\alpha\beta}\Gamma^\mu_{\sigma\gamma} + \Gamma^\mu_{\alpha\gamma,\beta} - \Gamma^\mu_{\alpha\beta,\gamma} \tag{63}$$

One also frequently encounters the notation (see e.g. Schutz, Chapter 6)

$$\nabla_{\vec{\mathbf{v}}}\,\nabla_{\vec{\mathbf{v}}}\,\xi^\mu \;=\; R^\mu{}_{\alpha\beta\gamma}\,v^\alpha\,v^\beta\,\xi^\gamma \tag{64}$$

The $(1,3)$ tensor, $\mathbf{R}$, in equations (62) – (64) is the Riemann Christoffel tensor referred to at the beginning of this section. It is a tensor of rank 4, with 256 components – although the symmetries inherent in any astrophysical example greatly reduce the number of independent components.

If spacetime is flat then

$$R^{\mu}{}_{\alpha\beta\gamma} = 0 \tag{65}$$

i.e. all components of the Riemann Christoffel tensor are identically zero. We then see from equation (64) that the acceleration of the geodesic deviation is identically zero – i.e. the separation between neighbouring geodesics remains constant. Conversely, however, if the spacetime is curved then the geodesic separation changes along the worldline of neighbouring particles.

Equations (62) and (64) are of fundamental importance in GR. In a sense they are the properly covariant mathematical expression of the physical idea embodied in Wheeler's phrase *"spacetime tells matter how to move"*. Although we derived them by considering the geodesics of neighbouring material particles, we can equally apply them to determine how changes in the spacetime curvature will influence the geodesic deviation of photons. In that case in equation (62) we simply need to replace the proper time τ along the geodesic by another suitable affine parameter λ (say), and we also replace v^{α} by $dx^{\alpha}/d\lambda$. Otherwise, the form of equation (62) is unchanged.

The second part of Wheeler's statement – *"matter tells spacetime how to curve"* – has its mathematical embodiment in Einstein's equations, which relate the **Einstein tensor** (and thus, as we will see, implicitly the Riemann Christoffel tensor) to the **energy momentum tensor**. We consider Einstein's equations in the next section.

5 Einstein's equations

5.1 The energy momentum tensor

The energy momentum tensor (also known as the stress energy tensor) is the *source* of spacetime curvature. It describes the presence and motion of gravitating matter. In these lectures we will discuss the energy momentum tensor for the particular case of a **perfect fluid**, which is a mathematical idealisation but one which is a good approximate description of the gravitating matter in many astrophysical situations.

5.1.1 Perfect fluids

Many Newtonian gravitational problems can be considered simply as the interaction of a small number of point-like massive particles. Even in Newtonian theory, however, there are many contexts (e.g. the motion of stars in the Galaxy) where the number of gravitating 'particles' is too large to follow their individual trajectories. Instead we treat the system as a smooth continuum, or **fluid**, and describe its behaviour in terms of the locally averaged properties (e.g. the *density*, *velocity* or *temperature*) of the particles in each **fluid element** – by which we mean a small region of the fluid surrounding some point in the continuum within which the behaviour of the particles is fairly homogeneous.

This fluid description is also useful for many-particle systems in special relativity, although we must be careful about defining quantities such as density and velocity which are frame-dependent – i.e. we need to find a *covariant* description of the fluid (which, we will see, is why we require a *tensor* to describe the gravitating matter).

The simplest type of relativistic fluid is known as **dust**. To a physicist, a fluid element of dust means a collection of particles which are all at rest with respect to some Lorentz frame. Many textbooks (including Schutz) refer to this Lorentz frame as the **momentarily comoving rest frame** (MCRF) of the fluid element.

This name helps to reinforce the point that the fluid element as a whole may possess a bulk motion with respect to the rest of the fluid, and indeed this relative motion may not be uniform – i.e. the fluid element may be accelerating. At any moment, however, the instantaneous velocity of the fluid element allows us to define its MCRF, although the MCRF of neighbouring elements will in general be different at that instant, and the MCRF of the fluid element will also in general be different at different times. If the fluid element is dust, however, then at any instant in the MCRF of the fluid element the individual particles are taken to possess no random motions of their own.

Generally, however, the particles within a fluid element *will* have random motions, and these will give rise to **pressure** in the fluid (c.f. motions of the molecules in an ideal gas). A fluid element may also be able to exchange energy with its neighbours via **heat conduction**, and there may be **viscous forces** present between neighbouring fluid elements. When viscous forces exist they are directed parallel to the interface between neighbouring fluid elements, and result in a **shearing** of the fluid.

A relativistic fluid element is said to be a **perfect fluid** if, in its MCRF, the fluid element has *no* heat conduction or viscous forces. It follows from this definition that dust is the special case of a pressure-free perfect fluid.

5.1.2 Definition of the energy momentum tensor

We can define the energy momentum tensor, $\mathbf{T}$, in terms of its components in some coordinate system, $\{x^1, x^2, ..., x^n\}$, for each fluid element. Thus we define $T^{\alpha\beta}$ for a fluid element to be equal to the **flux of the α component of four momentum of all gravitating matter**[2] **across a surface of constant x^β**.

[2]By 'gravitating matter' we mean here all material particles, plus (from the equivalence of matter and energy) any electromagnetic fields and particle fields which may be present

Thus, the change, Δp^α, in the α component of the four momentum due to the flux through a surface element, ΔS_ν, at constant x^ν, is given by

$$\Delta p^\alpha = T^{\alpha\nu}\Delta S_\nu \tag{66}$$

(Note the use of the summation convention).

5.1.3 Components of T in the MCRF for dust

In this case the energy momentum tensor takes a very simple form. Since the particles in the fluid element are at rest, there is no momentum transfer. (For a general fluid, even if the particles are at rest there can be a flux of energy and momentum through heat conduction, but not for dust, which is a perfect fluid). Also there is no momentum flux, which means that $T^{ij} = 0$, $(i, j = 1, 2, 3)$. In fact the only non-zero component is $T^{00} = \rho$, the energy density of the fluid element.

5.1.4 Components of T in the MCRF for a general perfect fluid

This case is only slightly less straightforward than that of dust. Again the T^{00} component is equal to the energy density, ρ. Since there is no bulk motion of the fluid element and there is no heat conduction for a perfect fluid, the energy flux $T^{0i} = 0$ for $i = 1, 2, 3$. Moreover, from the symmetry of $\mathbf{T}$ we also have that the momentum density, $T^{i0} = 0$, for $i = 1, 2, 3$. For the spatial components, $T^{ij} = 0$ if $i \neq j$, since these terms correspond to viscous forces parallel to the interface between fluid elements and these forces are zero for a perfect fluid. Thus T^{ij} is a diagonal matrix. But T^{ij} must be diagonal in all reference frames – e.g. under all possible rotations. This is possible only if T^{ij} is a scalar multiple of the identity matrix, i.e. $T^{11} = T^{22} = T^{33}$.

Thus, T^{ii} is the flux of the i^{th} component of momentum in the x^i direction, perpendicular to the fluid element interface. Equivalently, it is the *force per unit area*,

perpendicular to the interface. This is just the **pressure**, P, exerted by the random motions of the particles in the fluid element. Hence we can write $\mathbf{T}$ as

$$\mathbf{T} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & P & 0 & 0 \\ 0 & 0 & P & 0 \\ 0 & 0 & 0 & P \end{pmatrix}. \tag{67}$$

5.1.5 Components of T in a general Lorentz frame

Extending our expression for $T^{\alpha\beta}$ from the MCRF to a general Lorentz frame is fairly straightforward, but the interested reader is referred e.g. to Schutz for the details and here we just state the result. If $\vec{u} = \{u^\alpha\}$ is the *four* velocity of a fluid element in some Lorentz frame, then

$$T^{\alpha\beta} = (\rho + P)u^\alpha u^\beta + P\eta^{\alpha\beta}, \tag{68}$$

where $\eta^{\alpha\beta}$ is the Minkowski metric of SR, as given by equation (3) but with raised indices.

Conservation of energy and momentum requires that

$$T^{\alpha\beta}_{,\beta} = 0. \tag{69}$$

i.e. the **divergence** of the energy momentum tensor is equal to zero.

5.1.6 Extending to GR

In Section 1 we introduced the strong principle of equivalence which stated that, in a LIF, all physical phenomena are in agreement with special relativity. In the light of our discussion of tensors, we can write down an immediate consequence of the strong principle of equivalence as follows

Any physical law which can be expressed as a tensor equation in SR has exactly the same form in a local inertial frame of a curved spacetime

This statement holds since, in the LIF, physics – and hence the form of physical laws – is indistinguishable from the physics of SR. This is a very important result because it allows us to generalise the form of physical laws involving tensors which are valid in SR to the case of GR, with semi-colons (denoting covariant derivatives) replacing commas (denoting partial derivatives) where appropriate.

How is this extension justified? From the principle of covariance a tensorial description of physical laws must be equally valid in any reference frame. Thus, if a tensor equation holds in one frame it must hold in any frame. In particular, a tensor equation derived in a LIF (i.e. assuming SR) remains valid in an arbitrary reference frame (i.e. assuming GR).

Hence, the energy momentum tensor for a perfect fluid in GR takes the form

$$T^{\mu\nu} = (\rho + P)u^\mu u^\nu + Pg^{\mu\nu}, \tag{70}$$

where $g^{\mu\nu}$ denotes the contravariant metric tensor for a general curved spacetime (which of course reduces locally to $\eta^{\mu\nu}$, according to the WEP).

We can extend to GR in this way the result of equation (69), on the conservation of energy and momentum. Thus, for a fluid element in a general curved spacetime

$$T^{\mu\nu}_{\;;\nu} = 0. \tag{71}$$

If this were *not* the case – i.e. if there existed some point, P, at which $T^{\mu\nu}_{\;;\nu} \neq 0$ – then we could construct a LIF at P in which all Christoffel symbols are zero. In this new frame covariant derivatives reduce to partial derivatives, implying that $T^{\mu\nu}_{\;,\nu} \neq 0$, which contradicts equation (69).

The general technique of using the principles of covariance and equivalence to extend the validity of tensor equations from SR to GR, usually by evaluating their components in the LIF where Christoffel symbols vanish, is a very powerful one and is commonly met in the relativity literature. It is sometimes referred to informally as the 'comma goes to semi colon rule'.

5.2 The Einstein tensor and Einstein's equations

We have seen that the Riemann Christoffel tensor, $R^{\mu}{}_{\alpha\beta\gamma}$, describes the curvature of spacetime. Given Wheeler's phrase *"matter tells spacetime how to curve"*, we expect the Riemann Christoffel tensor and energy momentum tensor to be related, and indeed they are via Einstein's equations. But we saw in Section 5.1 that $\mathbf{T}$ is a $(2,0)$ tensor. Thus, Einstein's equations must involve various *contractions* of the Riemann Christoffel tensor.

First we can contract the Riemann Christoffel tensor to form a $(0,2)$ tensor, which we call the **Ricci tensor** defined by

$$R_{\alpha\gamma} = R^{\mu}_{\alpha\mu\gamma} \tag{72}$$

i.e. contracting on the *second* of the lower indices. (N.B. some authors choose to define $R_{\alpha\gamma}$ as minus this value). We can also write the components of the Ricci tensor as

$$R_{\alpha\gamma} = g^{\sigma\delta} R_{\sigma\alpha\delta\gamma} \tag{73}$$

It is easy to show that $R_{\alpha\beta} = R_{\beta\alpha}$, i.e. the Ricci tensor is symmetric.

By further contracting the Ricci tensor with the contravariant components of the metric, one obtains the **curvature scalar**, viz:

$$R = g^{\alpha\beta} R_{\alpha\beta} \tag{74}$$

One may also use the metric to raise the indices of the Ricci tensor, and thus express

it in contravariant form, viz:

$$R^{\mu\nu} = g^{\mu\alpha} g^{\nu\beta} R_{\alpha\beta} \tag{75}$$

$R^{\mu\nu}$ is also symmetric.

Using the contravariant form of the Ricci tensor, we define the **Einstein tensor**, **G** as follows

$$G^{\mu\nu} = R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R \tag{76}$$

where R is the curvature scalar. Note that since $R^{\mu\nu}$ is symmetric, so too is $G^{\mu\nu}$.

The Einstein tensor is of crucial physical significance in GR, since it can be shown from the Bianchi identities (see Appendix 1) that

$$G^{\mu\nu}_{\ ;\nu} = 0 \tag{77}$$

Thus we have immediately **Einstein's equations** which state that

$$T^{\mu\nu}_{\ ;\nu} = G^{\mu\nu}_{\ ;\nu} \tag{78}$$

Einstein took the solution of these equations to be of the form

$$G^{\mu\nu} = kT^{\mu\nu} \tag{79}$$

where we can determine the constant k by requiring that we should recover the laws of Newtonian gravity and dynamics in the limit of a weak gravitational field and non-relativistic motion. In fact k turns out to equal $8\pi G/c^4$.

5.3 Solving Einstein's equations

Solving Einstein's equations is in general a highly non-trivial task. Provided that the metric tensor is known, or assumed, for the system in question, then we can directly compute the Christoffel symbols, Riemann Christoffel tensor and Einstein tensor, and use these to determine the geodesics of material particles and photons.

Given the Einstein tensor we can also compute the components of the energy momentum tensor, and thus determine the spatial and temporal dependence of physical characteristics such as the density and pressure of the system.

Proceeding in the other direction, however, is considerably more difficult. Given, or assuming, a form for the energy momentum tensor, Einstein's equations immediately yield the Einstein tensor, but from that point solving for the Riemann Christoffel tensor and the metric is in general intractable.

Certain exact analytic solutions exist (e.g. the Schwarzschild, Kerr and Robertson-Walker metrics) and these are applicable to a range of astrophysical problems. The Schwarzschild metric, for example, can be applied to derive some of the classical predictions of GR, including the advance of pericentre of planetary orbits and the gravitational deflection of light. (Appendix 2 derives the form of the Schwarzschild metric and briefly considers these classical GR predictions in more detail.)

However, as we mentioned on Page 2, the description of sources of gravitational waves requires a *non-stationary* metric (see below for a precise definition) for which we must resort to approximate methods if we are to make progress analytically.

Fortunately we will see that we can describe very well at least the *detection* of gravitational waves within the so-called **weak field approximation**, in which we assume that the gravitational wave results from deviations from the flat spacetime of Special Relativity which are small.

As we will see in the following sections, the weak field approximation greatly simplifies our analysis and permits gravitational waves detected at the Earth to be described as a linear perturbation to the Minkowski metric of Special Relativity.

6 Wave Equation for Gravitational Radiation

We show that the free-space solutions for the metric perturbations of a 'nearly flat' spacetime take the form of a wave equation, propagating at the speed of light.

6.1 Non-stationarity

The Schwarzschild solution for the spacetime exterior to a point mass is an example of a **static metric**, defined as a metric for which we can find a time coordinate, t, satisfying

1. all metric components are independent of t
2. the metric is unchanged if we apply the transformation $t \mapsto -t$

A metric which satisfies property (1) but not property (2) is known as **stationary**. An example is the metric of a spherically symmetric star which is *rotating*: reversing the time coordinate changes the sense of the rotation, even though one can find a coordinate system in which the metric components are all independent of time. The Kerr metric, which can be used to describe the exterior spacetime around a rotating black hole, is an example of a stationary metric.

In these lectures we explore some consequences of also relaxing the assumption of property (1), by considering spacetimes in which the metric components are time dependent – as can happen when the source of the gravitational field is varying. Such a metric is known as **non-stationary**.

6.2 Weak gravitational fields

6.2.1 'Nearly' flat spacetimes

Since spacetime is flat in the absence of a gravitational field, a weak gravitational field can be defined as one in which spacetime is 'nearly' flat. What we mean

by 'nearly' here is that we can find a coordinate system in which the metric has components

$$g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta} \tag{80}$$

where

$$\eta_{\alpha\beta} = \text{diag } (-1, 1, 1, 1) \tag{81}$$

is the Minkowski metric of Special Relativity, and $|h_{\alpha\beta}| << 1$ for all α and β.

A coordinate system which satisfies equations (80) and (81) is referred to as a **nearly Lorentz** coordinate system. Notice that we say that we can find *a* coordinate system satisfying these equations. It certainly does *not* follow that for *any* choice of coordinate system we can write the metric components of the nearly flat spacetime in the form of equations (80) and (81). Indeed, even if the spacetime is precisely Minkowski, we could adopt (somewhat unwisely perhaps!) a coordinate system in which the metric components were very far from the simple form of equation (81).

In some coordinate systems, therefore, the components may be enormously more complicated than in others. The secret to *solving* tensor equations in General Relativity is, often, to first choose a coordinate system in which the components are as simple as possible. In that sense, equations (80) and (81) represent a 'good' choice of coordinate system; just as equation (81) represents the simplest form we can find for the metric components in flat spacetime, so equation (80) represents the metric components of a nearly flat spacetime in their simplest possible form.

The coordinate system in which one may express the metric components of a nearly flat spacetime in the form of equations (80) and (81) is certainly not unique. If we have identified such a coordinate system then we can find an infinite family of others by carrying out particular coordinate transformations. We next consider two types of coordinate transformations which preserve the properties of equations (80) and (81). These are known respectively as **Background Lorentz transformations** and **Gauge transformations**.

6.2.2 Background Lorentz transformations

Suppose we are in the Minkowski spacetime of Special Relativity, and we define the inertial frame, S, with coordinates (t, x, y, z). Suppose we then transform to another inertial frame, S', corresponding to a **Lorentz boost** of speed v in the direction of the positive x-axis. Under the Lorentz transformation S' has coordinates given by, in matrix form

$$(t', x', y', z')^T = \begin{pmatrix} \gamma & -v\gamma & 0 & 0 \\ -v\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} (t, x, y, z)^T \tag{82}$$

where $\gamma = (1 - v^2)^{-1/2}$. (Remember that we are taking $c = 1$). We can write this in more compact notation as

$$x'^{\alpha} = \Lambda^{\alpha'}_{\beta} x^{\beta} \quad \equiv \quad \frac{\partial x'^{\alpha}}{\partial x^{\beta}} x^{\beta} \tag{83}$$

The Lorentz matrix has inverse, corresponding to a boost of speed v along the *negative* x-axis, given by

$$(t, x, y, z)^T = \begin{pmatrix} \gamma & v\gamma & 0 & 0 \\ v\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} (t', x', y', z')^T \tag{84}$$

or

$$x^{\alpha} = \Lambda^{\alpha}_{\beta'} x'^{\beta} \quad \equiv \quad \frac{\partial x^{\alpha}}{\partial x'^{\beta}} x'^{\beta} \tag{85}$$

Now suppose we are in a nearly flat spacetime in which we have identified nearly Lorentz coordinates (t, x, y, z) satisfying equations (80) and (81). We then transform to a new coordinate system (t', x', y', z') defined such that

$$x'^{\alpha} = \Lambda^{\alpha'}_{\beta} x^{\beta} \tag{86}$$

i.e. where the transformation matrix is identical in form to equation (82) for some constant v. In this new coordinate system the metric components take the form

$$g'_{\alpha\beta} = \Lambda^{\mu}_{\alpha'}\Lambda^{\nu}_{\beta'}g_{\mu\nu} = \frac{\partial x^{\mu}}{\partial x'^{\alpha}}\frac{\partial x^{\nu}}{\partial x'^{\beta}}g_{\mu\nu} \tag{87}$$

Substituting from equation (80) we can write this as

$$g'_{\alpha\beta} = \frac{\partial x^{\mu}}{\partial x'^{\alpha}}\frac{\partial x^{\nu}}{\partial x'^{\beta}}\eta_{\mu\nu} + \frac{\partial x^{\mu}}{\partial x'^{\alpha}}\frac{\partial x^{\nu}}{\partial x'^{\beta}}h_{\mu\nu} \tag{88}$$

Because of the particular form of the coordinate transformation in this case, it follows that

$$g'_{\alpha\beta} = \eta'_{\alpha\beta} + \frac{\partial x^{\mu}}{\partial x'^{\alpha}}\frac{\partial x^{\nu}}{\partial x'^{\beta}}h_{\mu\nu} = \eta_{\alpha\beta} + h'_{\alpha\beta} \tag{89}$$

(The last equation follows because the components of the Minkowski metric are the same in any Lorentz frame).

Thus, provided we consider **only** transformations of the form of equation (82) the components of $h_{\mu\nu}$ transform **as if** they are the components of a $(0,2)$ tensor defined on a background flat spacetime. Moreover, provided $v << 1$, then if $|h_{\alpha\beta}| << 1$ for all α and β, then $|h'_{\alpha\beta}| << 1$ also.

Hence, our original nearly Lorentz coordinate system remains nearly Lorentz in the new coordinate system. In other words, a spacetime which looks nearly flat to one observer still looks nearly flat to any other observer in uniform relative motion with respect to the first observer.

6.2.3 Gauge transformations

Suppose now we make a very small change in our coordinate system by applying a coordinate transformation of the form

$$x'^{\alpha} = x^{\alpha} + \xi^{\alpha}(x^{\beta}) \tag{90}$$

i.e. where the components ξ^{α} are functions of the coordinates $\{x^{\beta}\}$. It then follows that

$$\frac{\partial x'^{\alpha}}{\partial x^{\beta}} = \delta^{\alpha}_{\beta} + \xi^{\alpha}_{,\beta} \tag{91}$$

From equation (90) we can also write

$$x^{\alpha} = x'^{\alpha} - \xi^{\alpha}(x^{\beta}) \tag{92}$$

If we now demand that the ξ^{α} are small, in the sense that

$$|\xi^{\alpha}{}_{,\beta}| << 1 \quad \text{for all} \quad \alpha, \beta \tag{93}$$

then it follows by the chain rule that

$$\frac{\partial x^{\alpha}}{\partial x'^{\gamma}} = \delta^{\alpha}_{\gamma} - \frac{\partial x^{\beta}}{\partial x'^{\gamma}} \frac{\partial \xi^{\alpha}}{\partial x^{\beta}} \simeq \delta^{\alpha}_{\gamma} - \xi^{\alpha}_{,\gamma} \tag{94}$$

where we have neglected terms higher than first order in small quantities. We have also used the fact that the components of the Kronecker delta are the same in any coordinate system.

Suppose now that the unprimed coordinate system is nearly Lorentz – i.e. the metric components satisfy equations (80) and (81). What about the metric components in the primed coordinate system?

Since the metric is a tensor, we know that

$$g'_{\alpha\beta} = \frac{\partial x^{\mu}}{\partial x'^{\alpha}} \frac{\partial x^{\nu}}{\partial x'^{\beta}} g_{\mu\nu} \tag{95}$$

Substituting from equations (80) and (94) this becomes, to first order

$$g'_{\alpha\beta} = \left(\delta^{\mu}_{\alpha}\delta^{\nu}_{\beta} - \xi^{\mu}_{,\alpha}\delta^{\nu}_{\beta} - \xi^{\nu}_{,\beta}\delta^{\mu}_{\alpha}\right)\eta_{\mu\nu} + \delta^{\mu}_{\alpha}\delta^{\nu}_{\beta}h_{\mu\nu} \tag{96}$$

This further simplifies to

$$g'_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta} - \xi_{\alpha,\beta} - \xi_{\beta,\alpha} \tag{97}$$

Note that in equation (97) we have defined

$$\xi_{\alpha} = \eta_{\alpha\nu}\xi^{\nu} \tag{98}$$

i.e. we have used the Minkowski metric, rather than the full metric $g_{\alpha\nu}$ to lower the index on the vector components ξ^{ν}. This is permitted because we are working to first order, and both the metric perturbation $h_{\alpha\nu}$ and the components ξ^{ν} are small. We have also used the fact that all the partial derivatives of $\eta_{\alpha\nu}$ are zero.

Thus, equation (98) has the same form as equation (80) provided that we write

$$h'_{\alpha\beta} = h_{\alpha\beta} - \xi_{\alpha,\beta} - \xi_{\beta,\alpha} \tag{99}$$

Note that if $|\xi^{\alpha}{}_{,\beta}|$ are small, then so too are $|\xi_{\alpha,\beta}|$, and hence $h'_{\alpha\beta}$. Thus, our new primed coordinate system is *still* nearly Lorentz.

The above results tell us that – once we have identified a coordinate system which is nearly Lorentz – we can add an arbitrary small vector ξ^{α} to the coordinates x^{α} without altering the validity of our assumption that spacetime is nearly flat. We can, therefore, choose the components ξ^{α} to make Einstein's equations as simple as possible. We call this step choosing a **gauge** for the problem – a name which has resonance with a similar procedure in electromagnetism – and coordinate transformations of the type given by equation (99) are known as **gauge transformation**. We will consider below specific choices of gauge which are particularly useful.

6.3 Einstein's equations for a weak gravitational field

If we can work in a nearly Lorentz coordinate system for a nearly flat spacetime this simplifies Einstein's equations considerably, and will eventually lead us to spot that the deviations from the metric of Minkowski spacetime – the components $h_{\alpha\beta}$ in equation (80) – obey a **wave equation**.

Before we arrive at this key result, however, we have some algebraic work to do first. We begin by deriving an expression for the Riemann Christoffel tensor in a weak gravitational field.

6.3.1 Riemann Christoffel tensor for a weak gravitational field

In its fully covariant form the Riemann Christoffel tensor is given by

$$R_{\alpha\beta\gamma\delta} = g_{\alpha\mu} R^{\mu}_{\beta\gamma\delta} = g_{\alpha\mu} \left[\Gamma^{\sigma}_{\beta\delta} \Gamma^{\mu}_{\sigma\gamma} - \Gamma^{\sigma}_{\beta\gamma} \Gamma^{\mu}_{\sigma\delta} + \Gamma^{\mu}_{\beta\delta,\gamma} - \Gamma^{\mu}_{\beta\gamma,\delta} \right] \tag{100}$$

Recall from the previous section that, if we are considering Background Lorentz transformations – i.e. if we restrict our attention only to the class of coordinate transformations which obey equation (86) – then the metric perturbations $h_{\alpha\beta}$ transform as if they are the components of a $(0, 2)$ tensor defined on flat, Minkowski spacetime. In this case the Christoffel symbols of the first two bracketed terms on the right hand side of equation (100) are equal to zero. It is then straightforward to show that, to first order in small quantities, the Riemann Christoffel tensor reduces to

$$R_{\alpha\beta\gamma\delta} = \frac{1}{2} \left(h_{\alpha\delta,\beta\gamma} + h_{\beta\gamma,\alpha\delta} - h_{\alpha\gamma,\beta\delta} - h_{\beta\delta,\alpha\gamma} \right) \tag{101}$$

Moreover, it is also quite easy to show that equation (101) is invariant under gauge transformations – i.e. the components of the Riemann Christoffel tensor are independent of the choice of gauge. This follows from the form of equation (99) and the

fact that partial derivatives are commutative – i.e.

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$$

Hence, our choice of gauge will **not** fundamentally change our determination of the curvature of spacetime or the behaviour of neighbouring geodesics.

6.3.2 Einstein tensor for a weak gravitational field

From equations (100) and (101) we can contract the Riemann Christoffel tensor and thus obtain an expression for the Ricci tensor in linearised form. This can be shown (see Appendix 3) to take the form

$$R_{\mu\nu} = \frac{1}{2}\left(h^{\alpha}_{\mu\,,\nu\alpha} + h^{\alpha}_{\nu\,,\mu\alpha} - h_{\mu\nu,\alpha}{}^{,\alpha} - h_{,\mu\nu}\right) \tag{102}$$

where we have written

$$h \equiv h^{\alpha}_{\alpha} = \eta^{\alpha\beta} h_{\alpha\beta} \tag{103}$$

Recalling equation (81) we see that h is essentially the *trace* of the perturbation $h_{\alpha\beta}$. Note also that again we have raised the indices of the components $h_{\alpha\beta}$ using $\eta^{\alpha\beta}$ and not $g^{\alpha\beta}$. This is justified since $h_{\alpha\beta}$ behaves like a $(0,2)$ tensor defined on a flat spacetime, for which the metric is $\eta^{\alpha\beta}$. The derivation of equation (102) also uses the fact that all partial derivatives of $\eta^{\alpha\nu}$ are zero.

Note further that we have introduced the notation, generalising the definition of equation (98)

$$f^{\alpha} = \eta^{\alpha\nu} f_{\nu} \tag{104}$$

where f^{α} are the components of a vector. We can also extend this notation for raising and lowering indices to the components of more general geometrical objects, and to their partial derivatives. For example, in equation (102)

$$h_{\mu\nu,\alpha}{}^{,\alpha} \;=\; \eta^{\alpha\sigma}\left(h_{\mu\nu,\alpha}\right)_{,\sigma} \;=\; \eta^{\alpha\sigma}\, h_{\mu\nu,\alpha\sigma} \tag{105}$$

After a further contraction of the Ricci tensor, to obtain the curvature scalar, R, where

$$R = \eta^{\alpha\beta} R_{\alpha\beta} \tag{106}$$

and substitution into the equation

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu} R \tag{107}$$

we obtain, after further algebraic manipulation (see Appendix 3 for the details) an expression for the Einstein tensor, $G_{\mu\nu}$, in linearised, fully covariant form

$$G_{\mu\nu} = \frac{1}{2}\left[h_{\mu\alpha,\nu}{}^{,\alpha} + h_{\nu\alpha,\mu}{}^{,\alpha} - h_{\mu\nu,\alpha}{}^{,\alpha} - h_{,\mu\nu} - \eta_{\mu\nu}\left(h_{\alpha\beta}{}^{,\alpha\beta} - h_{,\beta}{}^{,\beta}\right)\right] \tag{108}$$

This rather messy expression for the Einstein tensor can be simplified a little by introducing a modified form for the metric perturbation defined by

$$\overline{h}_{\mu\nu} \equiv h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu} h \tag{109}$$

after which (see Appendix 3) equation (108) becomes

$$G_{\mu\nu} = -\frac{1}{2}\left[\overline{h}_{\mu\nu,\alpha}{}^{,\alpha} + \eta_{\mu\nu}\overline{h}_{\alpha\beta}{}^{,\alpha\beta} - \overline{h}_{\mu\alpha,\nu}{}^{,\alpha} - \overline{h}_{\nu\alpha,\mu}{}^{,\alpha}\right] \tag{110}$$

6.3.3 Linearised Einstein equations

Having ploughed our way through all of the above algebra, we can now write down Einstein's equations in their linearised, fully covariant form for a weak gravitational field, in terms of the (re-scaled) metric perturbations $\overline{h}_{\mu\nu}$. Since

$$G_{\mu\nu} = 8\pi T_{\mu\nu} \tag{111}$$

it follows that

$$-\overline{h}_{\mu\nu,\alpha}{}^{,\alpha} - \eta_{\mu\nu}\overline{h}_{\alpha\beta}{}^{,\alpha\beta} + \overline{h}_{\mu\alpha,\nu}{}^{,\alpha} + \overline{h}_{\nu\alpha,\mu}{}^{,\alpha} = 16\pi T_{\mu\nu} \tag{112}$$

Can we simplify equation (112) any further? Fortunately the answer is 'yes'.

We saw in Sections 6.2.3 and 6.3.1 that we can carry out a gauge transformation in a nearly Lorentz coordinate system and the new coordinate system is still nearly Lorentz, with (to first order) identical curvature. It would be useful, therefore, to find a gauge transformation which eliminated the last three terms on the left hand side of equation (112). In Appendix 4 we show that a transformation with this property always exists, and in fact is equivalent to finding a coordinate system in which

$$\overline{h}^{\mu\alpha}{}_{,\alpha} = 0 \tag{113}$$

We call this gauge transformation the **Lorentz gauge**, and it plays an important role in simplifying Einstein's equations for a weak gravitational field:

- Suppose we begin with arbitrary metric perturbation components $h_{\mu\nu}^{(\mathrm{old})}$ (defined on a background Minkowski spacetime).
- We transform $h_{\mu\nu}^{(\mathrm{old})}$ to the Lorentz gauge by finding vector components ξ^μ which satisfy the **Lorentz gauge condition**, explained in detail in Appendix 4. The new metric perturbation components $h_{\mu\nu}^{(\mathrm{LG})}$, in the Lorentz gauge, are given by

$$h_{\mu\nu}^{(\mathrm{LG})} \;=\; h_{\mu\nu}^{(\mathrm{old})} \;-\; \xi_{\mu,\nu} \;-\; \xi_{\nu,\mu} \tag{114}$$

- We convert $h_{\mu\nu}^{(\mathrm{LG})}$ to $\overline{h}_{\mu\nu}^{(\mathrm{LG})}$ using equation (109), i.e.

$$\overline{h}_{\mu\nu}^{(\mathrm{LG})} \;=\; h_{\mu\nu}^{(\mathrm{LG})} \;-\; \frac{1}{2}\eta_{\mu\nu}\, h^{(\mathrm{LG})} \tag{115}$$

- Provided the ξ^μ satisfy the Lorentz gauge condition, then the $\overline{h}^{(\mathrm{LG})}_{\mu\nu}$ components will satisfy equation (113). This in turn means that – when our metric perturbation is expressed in terms of $\overline{h}^{(\mathrm{LG})}_{\mu\nu}$ – the last three terms on the left hand side of equation (112) are all zero.

Thus, in the Lorentz gauge, the linearised Einstein field equations reduce to the somewhat simpler form (dropping the label '(LG)' for clarity)

$$-\overline{h}_{\mu\nu,\alpha}{}^{,\alpha} = 16\pi T_{\mu\nu} \tag{116}$$

6.3.4 Solution to Einstein's equations in free space

In free space we can take the energy momentum tensor to be identically zero. The free space solutions of equation (116) are, therefore, solutions of the equation

$$\overline{h}_{\mu\nu,\alpha}{}^{,\alpha} = 0 \tag{117}$$

or, using equation (105)

$$\overline{h}_{\mu\nu,\alpha}{}^{,\alpha} \quad \equiv \quad \eta^{\alpha\alpha}\overline{h}_{\mu\nu,\alpha\alpha} \tag{118}$$

In fact, when we write out equation (118) explicitly, it takes the form

$$\left(-\frac{\partial^2}{\partial t^2} \; + \; \nabla^2 \right) \overline{h}_{\mu\nu} \quad = \quad 0 \tag{119}$$

which is often also written as

$$\Box\, \overline{h}_{\mu\nu} \quad = \quad 0 \tag{120}$$

where the operator $\Box$ is known as the *D'Alembertian*.

Remembering that we are taking $c = 1$, if instead we write

$$\eta^{00} = -\frac{1}{c^2} \tag{121}$$

then equation (119) can be re-written as

$$\left(-\frac{\partial^2}{\partial t^2} \; + \; c^2 \, \nabla^2 \right) \overline{h}_{\mu\nu} \quad = \quad 0 \tag{122}$$

This is a key result. Equation (122) has the mathematical form of a wave equation, propagating with speed c. Thus, we have shown that the metric perturbations – the 'ripples' in spacetime produced by disturbing the metric – propagate at the speed of light as waves in free space.

7 The Transverse – Traceless Gauge

We show that a coordinate system can be chosen in which the 16 components of a linear metric perturbation reduce to only 2 independent components.

We now explore further the properties of solutions to equation (119). The simplest solutions are **plane waves**

$$\overline{h}_{\mu\nu} = \mathrm{Re}\left[A_{\mu\nu} \, \exp\left(i k_\alpha x^\alpha \right) \right] \tag{123}$$

where 'Re' denotes the real part, and the constant components $A_{\mu\nu}$ and k_α are known as the **wave amplitude** and **wave vector** respectively. (Note that, as they appear in equation (123), the k_α are the components of a one-form. However, since we are considering the weak field limit of a background Minkowski spacetime, converting between covariant and contravariant components is very straightforward).

Equation (123) may appear to restrict the metric perturbations to a particular mathematical form, but *any* $\overline{h}_{\mu\nu}$ can be Fourier-expanded as a superposition of plane waves.

The wave amplitude and wave vector components are not arbitrary. Firstly, $A_{\mu\nu}$ is symmetric, since $\overline{h}_{\mu\nu}$ is symmetric. This immediately reduces the number of independent components from 16 to 10. Next, given that

$$\overline{h}_{\mu\nu,\alpha}{}^{,\alpha} = \eta^{\alpha\sigma}\,\overline{h}_{\mu\nu,\alpha\sigma} = 0 \tag{124}$$

it is easy to show that

$$k_\alpha\,k^\alpha = 0 \tag{125}$$

i.e. the wave vector is a **null vector**.

Thus, equation (123) describes a plane wave of frequency

$$\omega = k^t = \left(k_x^2 + k_y^2 + k_z^2\right)^{1/2} \tag{126}$$

propagating in direction $(1/k^t)\,(k_x, k_y, k_z)$.

Also, it follows from the Lorentz gauge condition

$$\overline{h}^{\mu\alpha}{}_{,\alpha} = 0 \tag{127}$$

that

$$\left(\overline{h}^{\alpha}_{\mu}\right)_{,\alpha} = 0 \tag{128}$$

from which it then follows that

$$A_{\mu\alpha}\,k^\alpha = 0 \tag{129}$$

i.e. the wave amplitude components must be orthogonal to the wave vector $\mathbf{k}$.

Equation (129) is, in fact, four linear equations – one for each value of the free coordinate index μ. This means that we have sufficient freedom to fix the values of four components of $A_{\mu\nu}$, thus reducing from 10 to 6 the number of independent components $A_{\mu\nu}$.

Can we restrict the components of the wave amplitude further still? The answer is again 'yes', since still we have some additional freedom remaining in our choice of gauge transformation.

Note that the transformation defined by equation (134) does **not** determine ξ^μ uniquely. We saw in Appendix 3 that the Lorentz gauge condition requires that the ξ^μ satisfy

$$\left(-\frac{\partial^2}{\partial t^2} \; + \; \nabla^2 \right) \xi^\mu \quad = \quad \overline{h}^{(\mathrm{old})\mu\nu}_{\quad ,\nu} \tag{130}$$

However, to any set of components ξ^μ which satisfy equation (130), we could add the components ψ^μ to define a new transformation

$$x'^\mu \quad \mapsto \quad x^\mu + \zeta^\mu \quad = \quad x^\mu + \xi^\mu + \psi^\mu \tag{131}$$

and provided the ψ^μ satisfy

$$\left(-\frac{\partial^2}{\partial t^2} \; + \; \nabla^2 \right) \psi^\mu \quad = \quad 0 \tag{132}$$

then ζ^μ will *still* satisfy

$$\left(-\frac{\partial^2}{\partial t^2} \; + \; \nabla^2 \right) \zeta^\mu \quad = \quad \overline{h}^{(\mathrm{old})\mu\nu}_{\quad ,\nu} \tag{133}$$

so that the modified metric perturbation

$$h^{(\mathrm{TT})}_{\mu\nu} \; = \; h^{(\mathrm{old})}_{\mu\nu} \; - \; \zeta_{\mu,\nu} \; - \; \zeta_{\nu,\mu} \tag{134}$$

still express the Einstein tensor in the simplified, Lorentz gauge form of equation (122).

The label '(TT)' stands for **Transverse Traceless**, and the gauge transformation ψ^μ (or equivalently ζ^μ) defines the **Transverse Traceless gauge**. The reason

for this name, and the importance of the Transverse Traceless gauge for describing gravitational waves will become clear shortly.

Equation (132) gives us four additional equations with which we can adjust the components of our gauge transformation, in order to choose a coordinate system which makes $\overline{h}_{\mu\nu}$ – and hence $A_{\mu\nu}$ – as simple as possible. In fact, it can be shown that the freedom we retain in our choice of ψ^{μ}, while still satisfying the Lorentz gauge conditions, allows us to restrict further $A_{\mu\nu}$ by fixing the values of four more of its components, thus reducing it to having only two independent components.

Specifically, if **U** is some arbitrarily chosen four vector with components U^{β}, then we have sufficient freedom in choosing the components of ψ to ensure that the wave amplitude tensor satisfies

$$A_{\alpha\beta}\, U^{\beta} = 0 \tag{135}$$

Moreover, we can also choose the components of ψ so that

$$A^{\mu}_{\mu} = \eta^{\mu\nu}\, A_{\mu\nu} = 0 \tag{136}$$

i.e. we can set the trace of **A** to be equal to zero. (This is the origin of the 'Traceless' part of the name Transverse Traceless gauge).

To fix our ideas and to see explicitly the form of $A_{\alpha\beta}$ which emerges from our adoption of the Transverse Traceless gauge, consider a test particle experiencing the passage of a gravitational wave in a nearly flat region of spacetime. Suppose we now transform to the background Lorentz frame in which the test particle is at rest – i.e. its four-velocity U^{β} has components $(1, 0, 0, 0)$, which we may also write as

$$U^{\beta} = \delta^{\beta}_{t} \tag{137}$$

Equations (135) and (137) then imply that

$$A_{\alpha t} = 0 \quad \text{for all} \quad \alpha \tag{138}$$

Next suppose we orient our spatial coordinate axes so that the wave is travelling in the positive z-direction, i.e.

$$k^t = \omega\,, \quad k^x = k^y = 0\,, \quad k^z = \omega \tag{139}$$

and

$$k_t = -\omega\,, \quad k_x = k_y = 0\,, \quad k_z = \omega \tag{140}$$

It then follows from equation (129) that

$$A_{\alpha z} = 0 \quad \text{for all } \alpha \tag{141}$$

i.e. there is no component of the metric perturbation in the direction of propagation of the wave. This explains the origin of the 'Transverse' part of the Transverse Traceless gauge; in this gauge the metric perturbation is entirely transverse to the direction of propagation of the gravitational wave.

To summarise, in the Transverse Traceless gauge equation (123) simplifies to become

$$\overline{h}^{(\mathrm{TT})}_{\mu\nu} = A^{(\mathrm{TT})}_{\mu\nu} \cos\left[\omega(t-z)\right] \tag{142}$$

Equations (138) and (141), combined with the symmetry of $A_{\mu\nu}$, imply that the only non-zero components of $A_{\mu\nu}$ are A_{xx}, A_{yy} and $A_{xy} = A_{yx}$. Moreover, the traceless condition, equation (136), implies that $A_{xx} = -A_{yy}$. Hence, the components of $A_{\mu\nu}$ in the Transverse Traceless gauge are

$$A^{(\mathrm{TT})}_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & A^{(\mathrm{TT})}_{xx} & A^{(\mathrm{TT})}_{xy} & 0 \\ 0 & A^{(\mathrm{TT})}_{xy} & -A^{(\mathrm{TT})}_{xx} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{143}$$

It then follows trivially that

$$\overline{h}^{(\mathrm{TT})}_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \overline{h}^{(\mathrm{TT})}_{xx} & \overline{h}^{(\mathrm{TT})}_{xy} & 0 \\ 0 & \overline{h}^{(\mathrm{TT})}_{xy} & -\overline{h}^{(\mathrm{TT})}_{xx} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{144}$$

where

$$\overline{h}^{(\mathrm{TT})}_{xx} = A^{(\mathrm{TT})}_{xx} \cos\left[\omega(t-z)\right] \tag{145}$$

and

$$\overline{h}^{(\mathrm{TT})}_{xy} = A^{(\mathrm{TT})}_{xy} \cos\left[\omega(t-z)\right] \tag{146}$$

Finally, we should note that in the Transverse Traceless gauge it is trivial to relate the components $\overline{h}_{\mu\nu}$ to the original metric perturbation $h_{\mu\nu}$. Following equation (103) and substituting from equation (109) we define

$$\overline{h} = \eta^{\alpha\beta}\,\overline{h}_{\alpha\beta} = \eta^{\alpha\beta}\left(h_{\alpha\beta} - \frac{1}{2}\eta_{\alpha\beta}\,h\right) = h - 2h = -h \tag{147}$$

Clearly, then, for the particular case of the Transverse Traceless gauge, both $\overline{h}$ and h are identically zero. (We should not be surprised by this, because of how we constructed h and $\overline{h}$ in the first place). It then follows trivially that for all $\alpha\,,\ \beta$

$$\overline{h}^{(\mathrm{TT})}_{\alpha\beta} = h^{(\mathrm{TT})}_{\alpha\beta} \tag{148}$$

8 Effect of Gravitational Waves on Free Particles

We investigate the geodesic equations for the trajectories of material particles and photons in a nearly flat spacetime, during to the passage of a gravitational wave.

There is a danger that the mathematical details of the previous section might appear rather abstract, and detached from practical issues relating to the design and operation of gravitational wave detectors. However, nothing could be further from the truth. The simplifications introduced by the adoption of the Transverse Traceless gauge have important practical consequences for detector design, as we will now illustrate.

8.1 Proper distance between test particles

We saw from equations (144) – (146) that the amplitude of the metric perturbation is described by just two independent constants, A_{xx} and A_{xy}. We can understand the physical significance of these constants by examining the effect of the gravitational wave on a free particle, in an initially wave-free region of spacetime.

We choose a background Lorentz frame in which the particle is initially at rest – i.e. the initial four-velocity of the particle is given by equation (137) – and we set up our coordinate system according to the Transverse Traceless Lorentz gauge.

The free particle's trajectory satisfies the geodesic equation

$$\frac{dU^\beta}{d\tau} + \Gamma^\beta_{\mu\nu} U^\mu U^\nu = 0 \tag{149}$$

where τ is the proper time. The particle is initially at rest, i.e. initially $U^\beta = \delta^\beta_t$.

Thus, the initial acceleration of the particle is

$$\left(\frac{dU^\beta}{d\tau}\right)_0 = -\Gamma^\beta_{tt} = -\frac{1}{2}\eta^{\alpha\beta}\left(h_{\alpha t,t} + h_{t\alpha,t} - h_{tt,\alpha}\right) \tag{150}$$

However, from equation (138)

$$A_{\alpha t} = 0 \quad \Rightarrow \quad \overline{h}_{\alpha t} = 0 \tag{151}$$

Also, recall that $h = \overline{h} = 0$. Therefore it follows that

$$h_{\alpha t} = 0 \quad \text{for all } \alpha \tag{152}$$

which in turn implies that

$$\left(\frac{dU^{\beta}}{d\tau}\right)_0 = 0 \tag{153}$$

Hence a free particle, initially at rest, will remain at rest indefinitely. However, 'being at rest' in this context simply means that the *coordinates* of the particle do *not* change. This is simply a consequence of our judicious choice of coordinate system, via the adoption of the Transverse Traceless Lorentz gauge. As the gravitational wave passes, this coordinate system adjusts itself to the ripples in the spacetime, so that any particles remain 'attached' to their initial coordinate positions. Coordinates are merely frame-dependent labels, however, and do not directly convey any invariant geometrical information about the spacetime.

Suppose instead we consider the **proper distance** between two nearby particles, both initially at rest, in this coordinate system: one at the origin and the other at spatial coordinates $x = \epsilon\,,\ y = z = 0$. The proper distance between the particles is then given by

$$\Delta\ell = \int \left| g_{\alpha\beta} dx^{\alpha} dx^{\beta} \right|^{1/2} \tag{154}$$

i.e.

$$\Delta\ell = \int_0^{\epsilon} |g_{xx}|^{1/2} \quad \simeq \quad \sqrt{g_{xx}(x=0)}\ \epsilon \tag{155}$$

Now

$$g_{xx}(x=0) = \eta_{xx} + h^{(\mathrm{TT})}_{xx}(x=0) \tag{156}$$

so

$$\Delta\ell \simeq \left[1 + \frac{1}{2}h^{(\mathrm{TT})}_{xx}(x=0)\right]\epsilon \tag{157}$$

Since $h^{(\mathrm{TT})}_{xx}(x=0)$ in general is **not** constant, it follows that the proper distance between the particles will change as the gravitational wave passes. It is essentially this change in the proper distance between test particles which gravitational wave detectors attempt to measure.

8.2 Geodesic deviation of test particles

We can study the behaviour of test particles more formally using the idea of **geodesic deviation**, first introduced in Section 2. We define the vector ξ^α which connects the two particles introduced above. Then, for a weak gravitational field, from equation (62) and taking proper time approximately equal to coordinate time

$$\frac{\partial^2\xi^\alpha}{\partial t^2} = R^\alpha_{\mu\nu\beta}\,U^\mu\,U^\nu\,\xi^\beta \tag{158}$$

where U^μ are the components of the four-velocity of the two particles. Since the particles are initially at rest, then

$$U^\mu \quad = \quad (1,0,0,0)^T \tag{159}$$

and

$$\xi^\beta \quad = \quad (0,\epsilon,0,0)^T \tag{160}$$

Equation (158) then simplifies to

$$\frac{\partial^2\xi^\alpha}{\partial t^2} = \epsilon R^\alpha_{ttx} = -\epsilon R^\alpha_{txt} \tag{161}$$

Substituting from equation (101) for a weak gravitational field, we can write down the relevant components of the Riemann Christoffel tensor in terms of the non-zero components of the metric perturbation (remembering always that we are working in the Transverse Traceless gauge)

$$R^x_{txt} = \eta^{xx} R_{xtxt} = -\frac{1}{2} h^{(\mathrm{TT})}_{xx,tt} \tag{162}$$

$$R^y_{txt} = \eta^{yy} R_{ytxt} = -\frac{1}{2} h^{(\mathrm{TT})}_{xy,tt} \tag{163}$$

Hence, two particles initially separated by ϵ in the x-direction, have a geodesic deviation vector which obeys the differential equations[3]

$$\frac{\partial^2}{\partial t^2}\,\xi^x \;=\; \frac{1}{2}\epsilon \frac{\partial^2}{\partial t^2} h^{(\mathrm{TT})}_{xx} \tag{164}$$

and

$$\frac{\partial^2}{\partial t^2}\,\xi^y \;=\; \frac{1}{2}\epsilon \frac{\partial^2}{\partial t^2} h^{(\mathrm{TT})}_{xy} \tag{165}$$

Similarly, it is straightforward to show that two particles initially separated by ϵ in the y-direction, have a geodesic deviation vector which obeys the differential equations

$$\frac{\partial^2}{\partial t^2}\,\xi^x \;=\; \frac{1}{2}\epsilon \frac{\partial^2}{\partial t^2} h^{(\mathrm{TT})}_{xy} \tag{166}$$

and

$$\frac{\partial^2}{\partial t^2}\,\xi^y \;=\; -\frac{1}{2}\epsilon \frac{\partial^2}{\partial t^2} h^{(\mathrm{TT})}_{xx} \tag{167}$$

[3]We are being a little sloppy in our notation here, as we have defined ξ as a $(1,0)$ tensor and $\mathbf{h}$ as a $(0,2)$ tensor. However, since we are in a background Lorentz spacetime, the x and y components of vectors and one-forms are identical.

8.3 Ring of test particles: polarisation of gravitational waves

We can further generalise equations (164) – (167) to consider the geodesic deviation of two particles – one at the origin and the other initially at coordinates $x = \epsilon \cos\theta$, $y = \epsilon \sin\theta$ and $z = 0$, i.e. in the x-y plane – as a gravitational wave propagates in the z-direction. We can show that ξ^x and ξ^y obey the differential equations

$$\frac{\partial^2}{\partial t^2}\,\xi^x \;=\; \frac{1}{2}\epsilon\cos\theta\frac{\partial^2}{\partial t^2}h^{(\mathrm{TT})}_{xx} \;+\; \frac{1}{2}\epsilon\sin\theta\frac{\partial^2}{\partial t^2}h^{(\mathrm{TT})}_{xy} \tag{168}$$

and

$$\frac{\partial^2}{\partial t^2}\,\xi^y \;=\; \frac{1}{2}\epsilon\cos\theta\frac{\partial^2}{\partial t^2}h^{(\mathrm{TT})}_{xy} \;-\; \frac{1}{2}\epsilon\sin\theta\frac{\partial^2}{\partial t^2}h^{(\mathrm{TT})}_{xx} \tag{169}$$

Substituting from equations (146) and (148), we can identify the solution

$$\xi^x \;=\; \epsilon\cos\theta + \frac{1}{2}\epsilon\cos\theta\,A^{(\mathrm{TT})}_{xx}\cos\omega t \;+\; \frac{1}{2}\epsilon\sin\theta\,A^{(\mathrm{TT})}_{xy}\cos\omega t \tag{170}$$

and

$$\xi^y \;=\; \epsilon\sin\theta + \frac{1}{2}\epsilon\cos\theta\,A^{(\mathrm{TT})}_{xy}\cos\omega t \;-\; \frac{1}{2}\epsilon\sin\theta\,A^{(\mathrm{TT})}_{xx}\cos\omega t \tag{171}$$

Suppose we now vary θ between 0 and 2π, so that we are considering an initially circular ring of test particles in the x-y plane, initially equidistant from the origin. Figure 9 shows the effect of the passage of a plane gravitational wave, propagating along the z-axis, on this ring of test particles.

The upper panel shows the case where the metric perturbation has $A^{(\mathrm{TT})}_{xx} \neq 0$ and $A^{(\mathrm{TT})}_{xy} = 0$. In this case the solutions for ξ^x and ξ^y reduce to

$$\xi^x \;=\; \epsilon\cos\theta\left(\,1 \;+\; \frac{1}{2}A^{(\mathrm{TT})}_{xx}\cos\omega t\,\right) \tag{172}$$

and

$$\xi^y \;=\; \epsilon\sin\theta\left(\,1 \;-\; \frac{1}{2}A^{(\mathrm{TT})}_{xx}\cos\omega t\,\right) \tag{173}$$

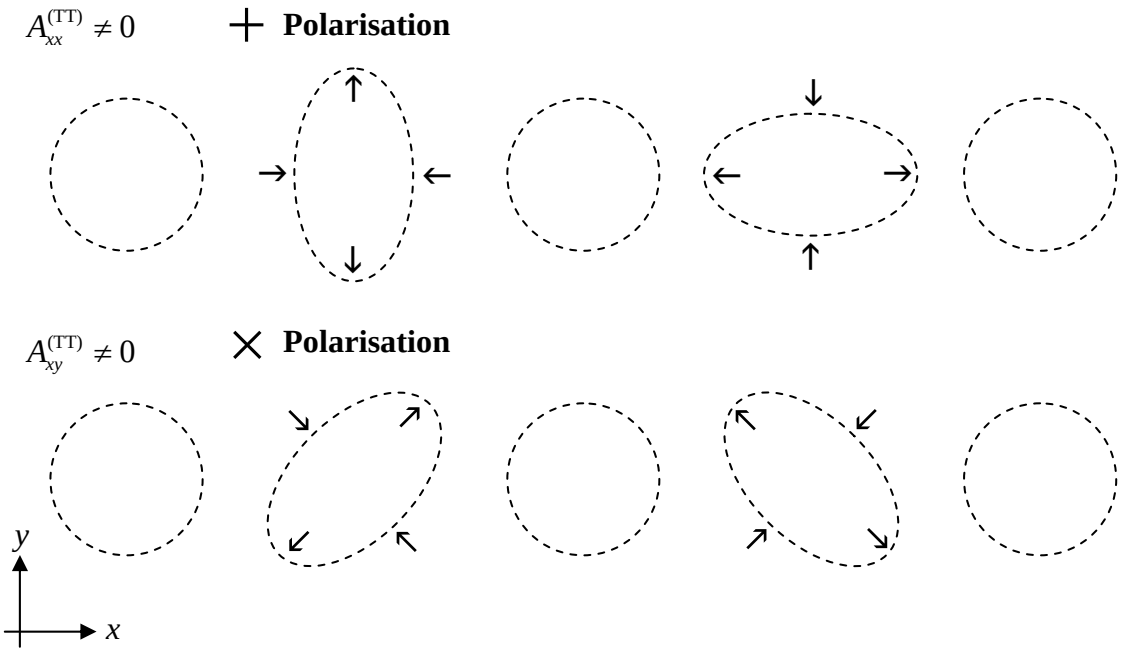


Figure 9: Cartoon illustrating the effect of a gravitational wave on a ring of test particles. The upper panel shows a wave for which $A_{xx}^{(\mathrm{TT})} \neq 0$ and $A_{xy}^{(\mathrm{TT})} = 0$, which we denote as the '+' polarisation. The lower panel shows a wave for which $A_{xy}^{(\mathrm{TT})} \neq 0$ and $A_{xx}^{(\mathrm{TT})} = 0$, which we denote as the '$\times$' polarisation.

Each of the five rings across the upper panel of Figure 9 corresponds to a different phase (i.e. different value of ωt) in the oscillation of the wave: the first, third and fifth phases shown are all odd multiples of $\pi/2$, so that the $\cos \omega t$ terms in equations (172) and (173) vanish. The second and fourth rings, on the other hand, correspond to a phase of π and 2π respectively. At phase π we can see from equations (172) and (173) that the effect of the wave will be to move test particles on the x-axis inwards – i.e. the gravitational wave *reduces* their proper distance from the centre of the ring – while test particles on the y-axis are moved outwards – i.e. the gravitational wave *increases* their proper distance from the centre of the ring. At phase 2π, on the

other hand, the wave will produce an opposite effect, increasing the proper distance from the ring centre of particles on the x-axis and reducing the proper distance of particles on the y-axis.

The lower panel of Figure 9 shows the contrasting case where the metric perturbation has $A^{(\mathrm{TT})}_{xy} \neq 0$ and $A^{(\mathrm{TT})}_{xx} = 0$. Again, the ring of test particles is shown for five different phases in the oscillation of the gravitational wave: $\pi/2$, π, $3\pi/2$, 2π and $5\pi/2$ respectively. In this case the solutions for ξ^x and ξ^y reduce to

$$\xi^x \;=\; \epsilon\cos\theta \;+\; \frac{1}{2}\epsilon\sin\theta\, A^{(\mathrm{TT})}_{xy}\cos\omega t \tag{174}$$

and

$$\xi^y \;=\; \epsilon\sin\theta \;+\; \frac{1}{2}\epsilon\cos\theta\, A^{(\mathrm{TT})}_{xy}\cos\omega t \tag{175}$$

To understand the relationship between these solutions and those for $A^{(\mathrm{TT})}_{xx} \neq 0$, we define new coordinate axes x' and y' by rotating the x and y axes through an angle of $-\pi/4$, so that

$$x' = \frac{1}{\sqrt{2}}\,(x - y) \tag{176}$$

and

$$y' = \frac{1}{\sqrt{2}}\,(x + y) \tag{177}$$

If we write the solutions for $A^{(\mathrm{TT})}_{xx} \neq 0$ in terms of the new coordinates x' and y', after some algebra we find that

$$\xi'^x \;=\; \epsilon\cos\left(\theta + \frac{\pi}{4}\right) \;+\; \frac{1}{2}\epsilon\sin\left(\theta + \frac{\pi}{4}\right)\, A^{(\mathrm{TT})}_{xy}\cos\omega t \tag{178}$$

and

$$\xi'^y \;=\; \epsilon\sin\left(\theta + \frac{\pi}{4}\right) \;+\; \frac{1}{2}\epsilon\cos\left(\theta + \frac{\pi}{4}\right)\, A^{(\mathrm{TT})}_{xy}\cos\omega t \tag{179}$$

Comparing equations (178) and (179) with equations (174) and (175) we see that our solutions with $A_{xy}^{(\mathrm{TT})} \neq 0$ are identical to the solutions with $A_{xx}^{(\mathrm{TT})} \neq 0$ apart from the rotation of $\pi/4$ – as can be seen from the lower panel of Figure 9.

We note some important features of the results of this section.

- The two solutions, for $A_{xx}^{(\mathrm{TT})} \neq 0$ and $A_{xy}^{(\mathrm{TT})} \neq 0$ represent two independent gravitational wave **polarisation states**, and these states are usually denoted by '+' and '×' respectively. In general any gravitational wave propagating along the z-axis can be expressed as a linear combination of the '+' and '×' polarisations, i.e. we can write the wave as

$$\mathbf{h} \;=\; a\,\mathbf{e}_+ \;+\; b\,\mathbf{e}_\times \tag{180}$$

  where a and b are scalar constants and the *polarisation tensors* $\mathbf{e}_+$ and $\mathbf{e}_\times$ are

$$\mathbf{e}_+ \;=\; \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{181}$$

  and

$$\mathbf{e}_\times \;=\; \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{182}$$

- We can see from the panels in Figure 9 that the distortion produced by a gravitational wave is **quadrupolar**. This is a direct consequence of the fact that gravitational waves are produced by changes in the curvature of spacetime, the signature of which is acceleration of the deviation between neighbouring geodesics. Recall our comment in Section 2.1 that geodesic deviation cannot

distinguish between a zero gravitational field and a uniform gravitational field: only for a non-uniform, **tidal** gravitational field does the geodesic deviation accelerate. Such tidal variations are quadrupolar in nature. In Section 9.1 below we discuss briefly a useful analogy with electromagnetic radiation that helps to explain why gravitational radiation is (at lowest order) quadrupolar.

- We can also see from Figure 9 that, at any instant, a gravitational wave is invariant under a rotation of 180° about its direction of propagation (in this case, the z-axis). By contrast, an electromagnetic wave is invariant under a rotation of 360°, and a *neutrino* wave is invariant under a rotation of 720°. This behaviour can be understood in terms of the *spin states* of the corresponding **gauge bosons**: the particles associated with the quantum mechanical versions of these waves.

  In general, the classical radiation field of a particle of spin, S, is invariant under a rotation of $360°/S$. Moreover, a radiation field of spin S has precisely two independent polarisation states, which are inclined to each other at an angle of $90°/S$. Thus, for an electromagnetic wave, corresponding to a photon of spin $S = 1$, the independent polarisation modes are inclined at 90° to each other.

  We can, therefore, deduce from the inclination of the gravitational wave polarisation states, that the **graviton** (which is, as yet undiscovered, since we do not yet have a fully developed theory of quantum gravity!) must be a spin $S = 2$ particle. The fact that electromagnetic waves correspond to a spin $S = 1$ field and gravitational waves correspond to a spin $S = 2$ field is also intimately connected to their mathematical description in terms of geometrical objects: spin $S = 1$ fields are **vector fields**, which is why we require only a vector description for the electromagnetic field; spin $S = 2$ fields, on the other hand, are **tensor fields**, which is why we required to introduce tensors

to describe the properties of the gravitational field.

8.4 The design of gravitational wave detectors: basic considerations

As we stated in Section 8.1, a change in the proper separation of test particles during the passage of a gravitational wave is the physical quantity which gravitational wave detectors are designed to measure. As will be discussed in detail in the subsequent lectures, in most of the gravitational wave detectors currently operational or planned for the future (e.g. LIGO, GEO600, VIRGO, LISA) these changes in the proper separation are monitored via measurement of the *light travel time* of a laser beam travelling back and forth along the arms of a Michelson Interferometer. Differences in the light travel time along perpendicular arms will produce interference fringes in the laser output of the interferometer.

We illustrate this in Figure 10, which shows in cartoon form the basic design of a Michelson interferometer gravitational wave detector. Here we suppose that gravitational wave with the '+' polarisation is propagating along the z-axis. Laser light of wavelength λ enters the apparatus at A, and is split into two perpendicular beams which are bounced off test mass mirrors M_1 and M_2 at the end of the arms (each of proper length L in the absence of any gravitational wave). The two beams are then re-combined and exit the system at B.

The three panels of Figure 10 denote three different phases of the wave as it passes through the system. In the left panel the wave causes no change in the proper length of the arms; in the middle panel the horizontal arm is shortened by ΔL while the vertical arm is lengthened by the same proper distance. In the right hand panel

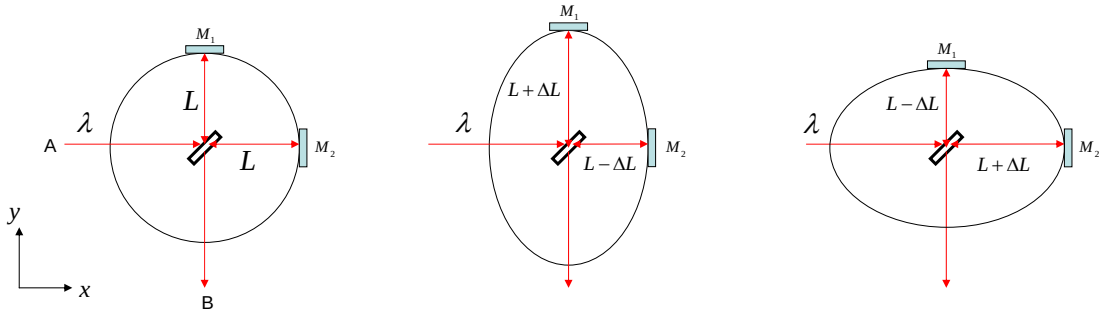


Figure 10: Cartoon illustration of the basic design features of a Michelson interferometer gravitational wave detector. See text for details.

we see the opposite: the horizontal and vertical arms are lengthened and shortened respectively by ΔL.

How are the physical dimensions of the interferometer related to the amplitude of the gravitational wave? Consider, for example, a gravitational wave $\mathbf{h} = h\mathbf{e}_+$ propagating along the z-axis. If we place two test masses along the x-axis, initially separated by proper distance L, we can see from equation (172) that the minimum and maximum proper distance between the test masses, as the gravitational wave passes, is $L - h/2$ and $L + h/2$ respectively. Thus, the fractional change $\Delta L/L$ in the proper separation of the test masses satisfies

$$\frac{\Delta L}{L} = \frac{h}{2} \tag{183}$$

Of course in general the arms of a gravitational wave detector will *not* be optimally aligned with the polarisation and direction of propagation of an incoming wave. Figure 11 sketches the orientation of one axis of a gravitational wave detector with respect to an incoming wave propagating along the z-axis. The detector axis is defined by standard spherical polar angles θ and ϕ. If the incoming wave has '+'

polarisation, i.e. $\mathbf{h} = h\,\mathbf{e}_+$, then the detector 'sees' an effective amplitude of

$$h_+ \;=\; h\,\sin^2\theta\,\cos 2\phi \tag{184}$$

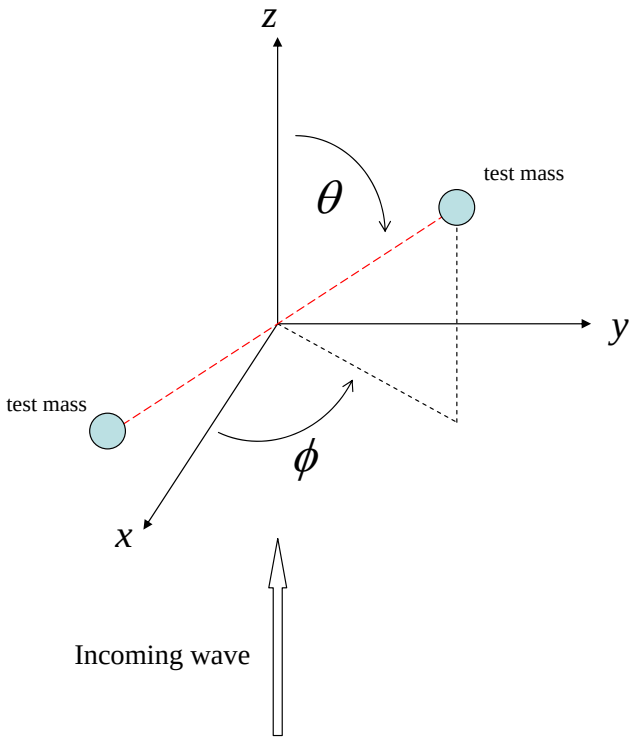


Figure 11: Cartoon illustration showing the relative orientation of a detector arm and the direction of propagation of a gravitational wave.

So we see that the wave produces a maximum response in the detector arm if $\theta = \pi/2$ and $\phi = 0$, and produces a *null* response for $\theta = 0$ or $\phi = \pi/4$. This makes sense when we consider Figure 11; we already commented previously that a metric perturbation produces no disturbance along its direction of propagation.

If, on the other hand, the incoming wave has the '$\times'$ polarisation, then in this case the detector 'sees' an effective amplitude of

$$h_\times \;=\; h\,\sin^2\theta\,\sin 2\phi \tag{185}$$

Now the wave produces a maximum response for $\theta = \pi/2$ and $\phi = \pi/4$, while the response is null for $\theta = 0$ or $\phi = 0$.

How large do we expect h to be? We have developed our understanding of the physics and mathematics of gravitational waves within the framework of the weak field approximation, which requires that $h << 1$. But how small is small? Much of the remainder of this School will focus on the enormous technical challenge which the detection of gravitational waves presents because h is indeed a tiny quantity: unless we are extraordinarily lucky, then from even the most cataclysmic astrophysical sources we expect h to be no larger than one part in 10^{20}.

To end these lectures, therefore, and as a pre-cursor to the more detailed treatments which will follow in later lectures, in the final section below we briefly consider the expected magnitude and character of gravitational waves from astrophysical sources.

9 The Production of Gravitational Waves

We discuss qualitatively the reasons why gravitational radiation is quadrupolar to lowest order, and we estimate the amplitude of the gravitational wave signal from a binary neutron star system.

9.1 The quadrupolar nature of gravitational waves

We can understand something important about the nature of gravitational radiation by drawing analogies with the formulae that describe electromagnetic radiation. This approach is crude at best since the electromagnetic field is a vector field while the gravitational field is a tensor field, but it is good enough for our present purposes. Essentially, we will take familiar electromagnetic radiation formulae and simply replace the terms which involve the Coulomb force by their gravitational analogues from Newtonian theory.

9.1.1 Electric and magnetic dipoles

In electromagnetic theory, the dominant form of radiation from a moving charge or charges is **electric dipole radiation**. For a single particle (e.g. an electron) of charge, e, with acceleration, $\mathbf{a}$, and dipole moment changing as $\ddot{\mathbf{d}} = e\,\ddot{\mathbf{x}} = e\,\mathbf{a}$, the power output, or luminosity, is given by

$$L_{\text{electric dipole}} \propto e^2\,\mathbf{a}^2 \tag{186}$$

For a general distribution of charges, with net dipole moment, $\mathbf{d}$, the luminosity is

$$L_{\text{electric dipole}} \propto e^2\,\dddot{\mathbf{d}}^2 \tag{187}$$

The next strongest types of electromagnetic radiation are **magnetic dipole** and **electric quadrupole radiation**. For a general distribution of charges, the luminosity arising from magnetic dipole radiation is proportional to the second time

derivative of the magnetic dipole moment, i.e.

$$L_{\text{magnetic dipole}} \propto \ddot{\mu} \tag{188}$$

where μ is given by a sum (or integral) over a distribution of charges:-

$$\mu \;=\; \sum_{q_i} \text{(position of } q_i) \times \text{(current due to } q_i) \tag{189}$$

9.1.2 Gravitational analogues

The gravitational analogue of the electric dipole moment is the **mass dipole moment**, **d**, summed over a distribution of particles, $\{A_i\}$

$$\mathbf{d} \;=\; \sum_{A_i} m_i \mathbf{x}_i \tag{190}$$

where m_i is the rest mass and $\mathbf{x}_i$ is the position of particle A_i.

By analogy with equation (188), the luminosity of gravitational 'mass dipole' radiation should be proportional to the second time derivative of **d** . However, the *first* time derivative of **d** is

$$\dot{\mathbf{d}} \;=\; \sum_{A_i} m_i \dot{\mathbf{x}}_i \quad \equiv \quad \mathbf{p} \tag{191}$$

where p is the **total linear momentum** of the system. Since the total momentum is conserved, it follows that the gravitational 'mass dipole' luminosity is zero – i.e **there can be no mass dipole radiation from any source**.

Similarly, the gravitational analogue of the magnetic dipole moment is

$$\mu \;=\; \sum_{A_i} (\mathbf{x}_i) \times (m_i \mathbf{v}_i) \quad \equiv \quad \mathbf{J} \tag{192}$$

where **J** is the **total angular momentum** of the system. Since the total angular momentum is *also* conserved, again it follows that the gravitational analogue of magnetic dipole radiation must have zero luminosity. Hence **there can be no dipole radiation of any sort from a gravitational source**.

The simplest form of gravitational radiation which has non-zero luminosity is, therefore, quadrupolar. We do not consider the mathematical details of quadrupolar radiation any further, save to point out that it can be shown that the quadrupole from a **spherically symmetric mass distribution** is identically zero. This suggests an important result: that, at least up to quadrupole order, **metric perturbations which are spherically symmetric do not produce gravitational radiation**. Thus, if e.g. the collapse of a massive star is spherically symmetric, it will generate *no* gravitational waves.

In fact, it is possible to prove that this result is *also* true for higher order radiation (e.g. octupole etc.), although the proof is very technical and is not discussed further. Interested readers are referred to Chapters 9 and 10 of Schutz '*A First Course in General Relativity*'.

9.2 Example: a binary neutron star system

Finally, we consider the example of the gravitational wave signature of a particular astrophysical system: a binary neutron star.

In general it can be shown (see, e.g. Schutz' textbook) that in the so-called *slow motion approximation* for a weak metric perturbation $h_{\mu\nu} << 1$ then for a source at distance r

$$h_{\mu\nu} = \frac{2G}{c^4 r}\ddot{I}_{\mu\nu} \tag{193}$$

where $I_{\mu\nu}$ is the **reduced quadrupole moment** defined as

$$I_{\mu\nu} = \int \rho(\vec{r}) \left(x_\mu x_\nu - \frac{1}{3}\delta_{\mu\nu} r^2 \right) dV \tag{194}$$

Consider a binary neutron star system consisting of two stars both of Schwarzschild mass M, in a circular orbit of coordinate radius R and orbital frequency f. For

simplicity we define our coordinate system so that the orbital plane of the pulsars lies in the $x-y$ plane, and at coordinate time $t=0$ the two pulsars lie along the $x-$axis. Substituting into equation (194)[4] it is then straightforward to show that

$$I_{xx} = 2MR^2\left[\cos^2(2\pi ft) - \frac{1}{3}\right] \tag{195}$$

$$I_{yy} = 2MR^2\left[\sin^2(2\pi ft) - \frac{1}{3}\right] \tag{196}$$

$$I_{xy} = I_{yx} = 2MR^2\left[\cos(2\pi ft)\sin(2\pi ft)\right] \tag{197}$$

From equations (193) and (195) – (197) it then follows that

$$h_{xx} = -h_{yy} = h\cos\left(4\pi ft\right) \tag{198}$$

and

$$h_{xy} = h_{yx} = -h\sin\left(4\pi ft\right) \tag{199}$$

where the amplitude term h is given by

$$h = \frac{32\pi^2 GMR^2 f^2}{c^4 r} \tag{200}$$

We see from equations (198) and (199) that the binary system emits gravitational waves at twice the orbital frequency of the neutron stars.

[4]taking the mass density distribution to be a sum of dirac delta functions – i.e. treating the pulsars as point masses

How large is h for a typical source? Suppose we take M equal to the Chandrasekhar mass, $M \sim 1.4 M_{\rm solar} = 2.78 \times 10^{30}$kg. We can then evaluate the constants in equation (200) and express h in more convenient units as

$$h = 2.3 \times 10^{-28} \ \frac{R^2[\mathrm{km}] f^2[\mathrm{Hz}]}{r[\mathrm{Mpc}]} \tag{201}$$

If we take $R = 20$km, say, $f = 1000$Hz (which is approximately the orbital frequency that Newtonian gravity would predict) and $r = 15$Mpc (corresponding to a binary system in e.g. the Virgo cluster), then we find that $h \sim 6 \times 10^{-21}$.

Thus we see that the signal produced by a typical gravitational wave source places extreme demands upon detector technology. In the lectures which follow we will explore how these technological challenges are being met and overcome.

Appendix 1: Proof that $G^{\mu\nu}_{\ ;\nu} = 0$

The key to proving this result is the Bianchi identities, which state that

$$R_{\alpha\beta\mu\nu;\lambda} + R_{\alpha\beta\lambda\mu;\nu} + R_{\alpha\beta\nu\lambda;\mu} = 0$$

Hence, our first step will be to prove that the Bianchi identities are true. The left hand side of the above equation is a $(0,5)$ tensor, so if we can show that its components are zero at an arbitrary point on the manifold in a given coordinate system, then they must be zero in all coordinate systems. We choose a LIF at arbitrary point, P and set up a coordinate system in which locally the Christoffel symbols are identically zero. From the definition of the Riemann Christoffel tensor

$$R^{\mu}_{\ \beta\gamma\delta} = \Gamma^{\sigma}_{\beta\gamma}\Gamma^{\mu}_{\sigma\delta} - \Gamma^{\sigma}_{\beta\delta}\Gamma^{\mu}_{\sigma\gamma} + \Gamma^{\mu}_{\beta\gamma,\delta} - \Gamma^{\mu}_{\beta\delta,\gamma}$$

Now, since $g_{\alpha\mu;\lambda} = 0$, we can write

$$R_{\alpha\beta\gamma\delta;\lambda} = (g_{\alpha\mu}R^{\mu}_{\ \beta\gamma\delta})_{;\lambda} = g_{\alpha\mu}R^{\mu}_{\ \beta\gamma\delta;\lambda}$$

Now all the Christoffel symbols are zero at P in our chosen coordinate system, so from the above two equations (changing covariant back to partial derivatives, which are interchangeable at P)

$$R_{\alpha\beta\gamma\delta;\lambda} = g_{\alpha\mu}(\Gamma^{\mu}_{\beta\gamma,\delta\lambda} - \Gamma^{\mu}_{\beta\delta,\gamma\lambda})$$

Substituting for the Christoffel symbols using

$$\Gamma^{\sigma}_{\alpha\beta} = g^{\sigma\nu}(g_{\nu\alpha,\beta} + g_{\nu\beta,\alpha} - g_{\alpha\beta,\nu})$$

and also using the fact that in our LIF $g_{\alpha\beta,\gamma} = g^{\alpha\beta}_{\ \ ,\gamma} = 0$ one obtains

$$R_{\alpha\beta\gamma\delta;\lambda} = \frac{1}{2}(g_{\alpha\gamma,\beta\delta\lambda} + g_{\alpha\delta,\beta\gamma\lambda} - g_{\beta\gamma,\alpha\delta\lambda} - g_{\beta\delta,\alpha\gamma\lambda})$$

Writing out the corresponding expressions for the other two terms on the left hand side of the Bianchi identities and adding gives the required result, i.e.

$$R_{\alpha\beta\gamma\delta;\lambda} + R_{\alpha\beta\lambda\gamma;\delta} + R_{\alpha\beta\delta\lambda;\gamma} = 0$$

in geodesic coordinates at P. Hence the Bianchi identities must hold in all coordinate systems because of their tensorial nature.

Now, from the symmetry properties of the Riemann Christoffel tensor, we have

$$R_{\alpha\beta\lambda\mu;\nu} = -R_{\alpha\beta\mu\lambda;\nu}$$

so that the Bianchi identities can be re-written as

$$R_{\alpha\beta\mu\nu;\lambda} - R_{\alpha\beta\mu\lambda;\nu} + R_{\alpha\beta\nu\lambda;\mu} = 0$$

Contracting this equation with $g^{\alpha\mu}$ we obtain

$$g^{\alpha\mu}R_{\alpha\beta\mu\nu;\lambda} - g^{\alpha\mu}R_{\alpha\beta\mu\lambda;\nu} + g^{\alpha\mu}R_{\alpha\beta\nu\lambda;\mu} = 0$$

Using the product rule for covariant differentiation and the fact that $g^{\alpha\mu}_{\ ;\beta} = 0$, we obtain

$$(g^{\alpha\mu}R_{\alpha\beta\mu\nu})_{;\lambda} - (g^{\alpha\mu}R_{\alpha\beta\mu\lambda})_{;\nu} + (g^{\alpha\mu}R_{\alpha\beta\nu\lambda})_{;\mu} = 0$$

This further simplifies to

$$R_{\beta\nu;\lambda} - R_{\beta\lambda;\nu} + R^{\mu}_{\ \beta\nu\lambda;\mu} = 0$$

We can further contract this equation by multiplying by $g^{\beta\nu}$, using the product rule and the fact that $g^{\alpha\mu}_{\ ;\beta} = 0$, to obtain (after contracting the first two terms and writing the third in terms of the fully covariant form of the Riemann Christoffel tensor)

$$R_{;\lambda} - R^{\nu}_{\lambda;\nu} + g^{\beta\nu}g^{\alpha\mu}R_{\alpha\beta\nu\lambda;\mu} = 0$$

Using the fact that

$$R_{\alpha\beta\nu\lambda} = -R_{\beta\alpha\nu\lambda}$$

the third term on the left hand side can therefore be re-written as

$$g^{\beta\nu}g^{\alpha\mu}R_{\alpha\beta\nu\lambda;\mu} = -(g^{\beta\nu}g^{\alpha\mu}R_{\beta\alpha\nu\lambda})_{;\mu} = -(g^{\alpha\mu}R_{\alpha\lambda})_{;\mu} = -R^{\mu}_{\ \lambda;\mu} \equiv -R^{\nu}_{\ \lambda;\nu}$$

Thus, we have

$$R_{;\lambda} - 2R^{\nu}_{\ \lambda;\nu} = 0$$

Multiplying each term by $-\frac{1}{2}g^{\mu\lambda}$ we obtain

$$g^{\mu\lambda}R^{\nu}_{\ \lambda;\nu} - \frac{1}{2}g^{\mu\lambda}R_{;\lambda} = 0$$

Once again using the product rule and the fact that the first covariant derivatives of the metric vanish this yields

$$(g^{\mu\lambda}R^{\nu}{}_{\lambda})_{;\nu} - (\frac{1}{2}g^{\mu\nu}R)_{;\nu} = 0$$

which of course is equivalent to

$$G^{\mu\nu}_{\ ;\nu} = 0$$

Appendix 2: The Schwarzschild solution

The Schwarzschild metric is a solution of Einstein's equations which describes the space-time exterior to a spherically symmetric, static mass, M.

A2.1: Derivation of the Schwarzschild metric

In most astrophysical situations, we can work with the metric tensor in **orthogonal** form. This means that one can identify a coordinate system in which the components, $g_{\alpha\beta}$, of the metric tensor satisfy

$$g_{\alpha\beta} = 0 \quad \text{for all } \alpha \neq \beta$$

We will derive the Schwarzschild metric in orthogonal form, so we begin by stating some general results which are valid for orthogonal metrics.

We can show that the Christoffel symbols take a simple form for an orthogonal metric

$$\begin{aligned}
\Gamma^{\lambda}_{\mu\nu} &= 0 \quad \text{for } \lambda,\ \mu,\ \nu \text{ all different} \\
\Gamma^{\lambda}_{\lambda\mu} &= \Gamma^{\lambda}_{\mu\lambda} = g_{\lambda\lambda,\mu}/2g_{\lambda\lambda} \\
\Gamma^{\lambda}_{\mu\mu} &= -g_{\mu\mu,\lambda}/2g_{\lambda\lambda} \\
\Gamma^{\lambda}_{\lambda\lambda} &= g_{\lambda\lambda,\lambda}/2g_{\lambda\lambda}
\end{aligned}$$

(Note: the summation convention does not apply in these equations).

For affine parameter, p, the geodesic equation can be re-written for an orthogonal metric in the form

$$\frac{d}{dp}\left(g_{\lambda\lambda}\frac{dx^\lambda}{dp}\right) - \frac{1}{2}\frac{\partial g_{\mu\mu}}{\partial x^\lambda}\left(\frac{dx^\mu}{dp}\right)^2 = 0$$

The Schwarzschild solution is also an example of a spherically symmetric static spacetime (which we henceforth refer to as S^4 for short). We can show that the interval for S^4 takes the general form

$$ds^2 = -e^\nu dt^2 + e^\lambda dr^2 + r^2\left(d\theta^2 + \sin^2\theta\, d\phi^2\right)$$

which, of course, has an orthogonal metric. Note that we have introduced the functions $\nu(r)$ and $\lambda(r)$ in place of g_{tt} and g_{rr}. Since the exponential function is strictly positive for all r, this replacement is legitimate provided that $g_{tt} < 0$ and $g_{rr} > 0$ for all points in our spacetime.

The Christoffel symbols for S^4 are given by

$$\begin{aligned}
\Gamma^t_{rt} = \Gamma^t_{tr} &= \tfrac{1}{2}\nu' & \Gamma^\theta_{r\theta} = \Gamma^\theta_{\theta r} &= \tfrac{1}{r}\\
\Gamma^r_{tt} &= \tfrac{1}{2}\nu' e^{\nu-\lambda} & \Gamma^\theta_{\phi\phi} &= -\sin\theta\cos\theta\\
\Gamma^r_{rr} &= \tfrac{1}{2}\lambda' & \Gamma^\phi_{r\phi} = \Gamma^\phi_{\phi r} &= \tfrac{1}{r}\\
\Gamma^r_{\theta\theta} &= -re^{-\lambda} & \Gamma^\phi_{\theta\phi} = \Gamma^\phi_{\phi\theta} &= \cot\theta\\
\Gamma^r_{\phi\phi} &= -re^{-\lambda}\sin^2\theta
\end{aligned}$$

All others zero

Further, we can write the Ricci tensor for S^4 as

$$R_{\lambda\nu} = \Gamma^\tau_{\lambda\nu}\Gamma^\sigma_{\tau\sigma} - \Gamma^\tau_{\lambda\sigma}\Gamma^\sigma_{\tau\nu} + \Gamma^\sigma_{\lambda\nu,\sigma} - \Gamma^\sigma_{\lambda\sigma,\nu}$$

i.e.

$$R_{tt} = \frac{1}{2}e^{\nu-\lambda}\left(\nu'' + \frac{1}{2}\nu'^2 - \frac{1}{2}\nu'\lambda' + \frac{2}{r}\nu'\right)$$

$$R_{rr} = -\frac{1}{2}\left(\nu'' + \frac{1}{2}\nu'^2 - \frac{1}{2}\nu'\lambda' - \frac{2}{r}\lambda'\right)$$

$$R_{\theta\theta} = 1 - e^{-\lambda}\left[1 + \frac{r}{2}\left(\nu' - \lambda'\right)\right]$$

$$R_{\phi\phi} = R_{\theta\theta}\sin^2\theta$$

and all other terms of the Ricci tensor are identically zero.

We are now in a position to derive the **Schwarzschild solution**. If the star is in an isolated region of space, then we can assume that all components of the Ricci tensor exterior to the star are identically zero. Thus

$$e^{\lambda-\nu}R_{tt} + R_{rr} = \frac{\nu' + \lambda'}{r} = 0$$

which in turn implies that

$$\nu + \lambda = \text{constant}$$

At large distances from the star we want the Schwarzschild metric to reduce to SR. Thus, as

$$r \to \infty \ , \quad \nu \to 0 \quad \text{and} \quad \lambda \to 0$$

which implies that

$$\nu + \lambda = 0$$

so that

$$e^{\nu} = e^{-\lambda}$$

This allows us to eliminate ν, giving

$$e^{-\lambda}\left(1 - \lambda' r\right) = 1$$

i.e.

$$\frac{d}{dr}\left(re^{-\lambda}\right) = 1$$

which we can integrate to give

$$e^{\nu} = e^{-\lambda} = 1 + \frac{\alpha}{r}$$

where α is a constant.

To evaluate α, suppose we release a material 'test' particle from rest. Thus, initially

$$\frac{dx^j}{d\tau} = 0 \quad \text{for } j = 1, 2, 3$$

where τ is **proper time**, and

$$\frac{dx^0}{d\tau} \equiv \frac{dt}{d\tau} \neq 0$$

Using the fact that

$$g_{\alpha\beta}\frac{dx^\alpha}{d\tau}\frac{dx^\beta}{d\tau} = -1$$

and after some reduction we see that

$$\frac{dt}{d\tau} = e^{-\nu/2}$$

We now make use of the geodesic differential equation for radial coordinate, r. At the instant when the particle is released this reduces to

$$\frac{d^2r}{d\tau^2} + \Gamma^r_{tt}\left(\frac{dt}{d\tau}\right)^2 = 0$$

After some further substitution we obtain finally

$$\frac{d^2r}{d\tau^2} = \frac{\alpha}{2r^2}$$

In the limit of a weak gravitational field this result must reduce to the prediction of Newtonian gravity, which is

$$\frac{d^2r}{dt^2} = -\frac{GM}{r^2}$$

where M is the mass of the star. If we adopt convenient units such that the gravitational constant, $G = 1$, this means that

$$\alpha = -2M$$

Finally, then, we can write down the invariant interval for the Schwarzschild metric

$$ds^2 = -\left(1 - \frac{2M}{r}\right)dt^2 + \frac{dr^2}{\left(1 - \frac{2M}{r}\right)} + r^2 d\theta^2 + r^2\sin^2\theta d\phi^2$$

A2.2: Geodesics for the Schwarzschild metric

The geodesics for a material 'test' particle in the Schwarzschild metric satisfy, with the proper time, τ, as affine parameter:

$$\frac{d}{d\tau}\left(g_{\lambda\nu}\frac{dx^\nu}{d\tau}\right) - \frac{1}{2}\frac{\partial g_{\mu\nu}}{\partial x^\lambda}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} = 0$$

Notice that the Schwarzschild metric coefficients are independent of both t and ϕ; hence if we set $\lambda = 0$ and $\lambda = 3$ then the second term on the left hand side of the geodesic equation vanishes. Given also that the Schwarzschild metric is *orthogonal*, it follows that

$$\frac{d}{d\tau}\left(g_{tt}\frac{dt}{d\tau}\right) = \frac{d}{d\tau}\left(g_{\phi\phi}\frac{d\phi}{d\tau}\right) = 0$$

Integrating this gives us

$$g_{tt}\frac{dt}{d\tau} = \text{ constant}$$

and

$$g_{\phi\phi}\frac{d\phi}{d\tau} = \text{ constant}$$

The geodesic equation for θ (i.e. $\lambda = 2$) is then

$$\frac{d}{d\tau}\left(r^2\frac{d\theta}{d\tau}\right) - \frac{1}{2}\frac{\partial}{\partial\theta}\left(r^2\sin^2\theta\right)\left[\frac{d\phi}{d\tau}\right]^2 = 0$$

which reduces to

$$r^2\frac{d^2\theta}{d\tau^2} + 2r\frac{dr}{d\tau}\frac{d\theta}{d\tau} - r^2\sin\theta\cos\theta\left(\frac{d\phi}{d\tau}\right)^2 = 0$$

This equation has a particular solution $\theta = \pi/2$; adopting this solution is equivalent to choosing the plane of the orbit of our material particle (e.g. a planet) to lie in the equatorial plane of our coordinate system. Making use of $\theta = \pi/2$ to simplify our solution, it follows that

$$\frac{dt}{d\tau} = \frac{k}{1 - \frac{2M}{r}}$$

and

$$\frac{d\phi}{d\tau} = \frac{h}{r^2}$$

where h and k are constants.

We can now obtain the geodesic differential equation for r, using

$$-1 = g_{tt}\left(\frac{dt}{d\tau}\right)^2 + g_{rr}\left(\frac{dr}{d\tau}\right)^2 + g_{\phi\phi}\left(\frac{d\phi}{d\tau}\right)^2$$

which in turn reduces to

$$\left(\frac{dr}{d\tau}\right)^2 = k^2 - 1 - \frac{h^2}{r^2} + \frac{2M}{r}\left(1 + \frac{h^2}{r^2}\right)$$

A2.3: The advance of pericentre in GR

Changing the dependent variable from r to u and the independent variable from τ to ϕ, our radial geodesic equation reduces to

$$h^2\left(\frac{du}{d\phi}\right)^2 = \left(k^2-1\right) - h^2u^2 + 2Mu\left(1+h^2u^2\right)$$

Differentiating, and cancelling the common factor of $\mathrm{d}u/\mathrm{d}\phi$ gives

$$\frac{d^2u}{d\phi^2} = -u + \frac{M}{h^2} + 3Mu^2$$

The effect of GR is to add the extra term $3Mu^2$ on the right hand side. For typical planetary orbits in the Solar System this extra term is tiny compared with the second term; e.g. for the Earth's orbit the ratio

$$\frac{3Mu^2}{M/h^2} \simeq 3\times 10^{-8}$$

Hence, because the extra GR term is very small anyway, we can obtain a very good approximation to the solution by replacing u in the u^2 term on the right hand side by the solution to the *Newtonian* version of this equation. Doing this we obtain

$$\frac{d^2u}{d\phi^2} = -u + \frac{M}{h^2} + 3\frac{M^3}{h^4}\left(1+2e\cos\phi+e^2\cos^2\phi\right)$$

We can write u as the sum of a 'Newtonian' and 'GR' part, so that u_{GR} describes the correction to the Newtonian orbit. Subtracting off the Newtonian solution gives

$$\frac{d^2u_{\mathrm{GR}}}{d\phi^2} = -u_{\mathrm{GR}} + 3\frac{M^3}{h^4}\left(1+2e\cos\phi+e^2\cos^2\phi\right)$$

which we can rewrite as

$$\frac{d^2u_{\mathrm{GR}}}{d\phi^2} + u_{\mathrm{GR}} = 3\frac{M^3}{h^4}\left(1+\frac{e^2}{2}+2e\cos\phi+\frac{e^2}{2}\cos 2\phi\right)$$

The right hand side of this equation takes the form $A + B\cos\phi + C\cos 2\phi$, where A, B and C are constants. It is easy to verify that particular integrals for each of these terms are, respectively

$$u_{\mathrm{GR}} = A$$

$$u_{\rm GR} = \frac{1}{2} B\phi \sin\phi$$
$$u_{\rm GR} = -\frac{1}{3} C \cos 2\phi$$

and the correction to the Newtonian orbit is given by the sum of these three particular integrals. Since each of the constants, A, B and C is of order the tiny constant M^3/h^4, we see that the first and third terms add to the Newtonian solution respectively a completely negligible constant and an equally negligible constant plus a tiny "wiggle".

The second term, on the other hand, is of a different form. Although the constant, B, is negligibly small, the presence of the ϕ means that this term produces a continually increasing – and thus ultimately non-negligible – effect. In fact

$$u = \frac{M}{h^2}\left(1 + e\cos\phi + \frac{3M^2}{h^2} e\phi \sin\phi\right)$$

Now, given that $3M^2/h^2$ is very small, and then using the approximations $\cos\beta \simeq 1$ and $\sin\beta \simeq \beta$ for small angle β, and the cosine addition formula, we can re-cast this last equation as

$$u = \frac{M}{h^2}\left[1 + e\cos\left(1 - \frac{3M^2}{h^2}\right)\phi\right]$$

Comparing this solution with its Newtonian analogue, we see that again the solution is elliptical in form and that u (and hence r) is a periodic function of ϕ. Notice, however, that the period, P, is given by

$$P = \frac{2\pi}{1 - 3M^2/h^2} > 2\pi$$

This means that the values of r trace out an approximate ellipse, but do not begin to repeat until *after* the radius vector has made a complete revolution. In other words the orbit can be regarded as an ellipse that 'precesses', so that the pericentre line advances each orbit by an amount, Δ, given by

$$\Delta = 2\pi\left(1 - \frac{3M^2}{h^2}\right)^{-1} - 2\pi \simeq \frac{6\pi M^2}{h^2} = \frac{6\pi M}{a(1-e^2)}$$

If we apply this equation to the orbit of Mercury, we obtain a perihelion advance which builds up to about 43 seconds of arc per century.

A2.4: Gravitational light deflection in GR

The geodesics for a photon in the Schwarzschild metric may be derived in a similar manner to Appendix 2.2, but we now must introduce a new affine parameter, λ (say), since the proper time for a photon is zero.[5]

For the 't' and 'ϕ' geodesic equations it is straightforward to see that we again obtain equations of the form

$$\frac{dt}{d\lambda} = \frac{k}{1-2M/r}$$
$$\frac{d\phi}{d\lambda} = \frac{h}{r^2}$$

since, for the 'θ' equation, we can again spot the particular solution $\theta = \pi/2$. It obviously then also follows that

$$\frac{d\theta}{d\lambda} = 0$$

We can then obtain

$$\left(\frac{dr}{d\lambda}\right)^2 = k^2 - \frac{h^2}{r^2} + \frac{2Mh^2}{r^3}$$

We now proceed as in Appendix 2.3, replacing the dependent variable, r by $u = 1/r$, and the independent variable λ by ϕ. This leads to

$$\frac{d^2u}{d\phi^2} + u = 3Mu^2$$

If we ignore the term on the right hand side we can see that a particular integral is

$$u = \frac{\cos\phi}{r_{\min}}$$

Following the same approach as in Appendix 2.3, we can obtain a very good approximation to the solution by replacing u on the right hand side by the solution to the corresponding Newtonian equation[6]. This gives the equation

$$\frac{d^2u}{d\phi^2} + u = \frac{3M}{r_{\min}^2}\cos^2\phi = \frac{3M}{2r_{\min}^2}\left(1+\cos 2\phi\right)$$

[5] In fact, the initial choice of affine parameter will not be important, since we will determine the trajectory of the photon with the coordinate ϕ as the independent variable

[6] If we regard photons (as, indeed, modern physics holds true) as particles with zero rest mass, then formally they should be 'immune' to Newton's gravitational force. If, on the other hand, we regard photons as having a negligible but non-zero mass then – even within a purely Newtonian

It is straightforward to verify that a particular integral of this approximation is

$$u = \frac{3M}{2r_{\rm min}^2}\left(1 - \frac{1}{3}\cos 2\phi\right)$$

from which it follows that the general solution is

$$u = \frac{\cos\phi}{r_{\rm min}} + \frac{3M}{2r_{\rm min}^2}\left(1 - \frac{1}{3}\cos 2\phi\right)$$

We can rewrite this, for e.g. the outgoing photon trajectory, as

$$u = \frac{\cos\left(\frac{\pi}{2} + \frac{\Delta\phi}{2}\right)}{r_{\rm min}} + \frac{3M}{2r_{\rm min}^2}\left[1 - \frac{1}{3}\cos\left(\pi + \Delta\phi\right)\right]$$

or

$$u = -\frac{\sin(\Delta\phi/2)}{r_{\rm min}} + \frac{3M}{2r_{\rm min}^2}\left[1 + \frac{1}{3}\cos\Delta\phi\right]$$

which further simplifies, since $\Delta\phi << 1$, to

$$u = -\frac{\Delta\phi}{2r_{\rm min}} + \frac{2M}{r_{\rm min}^2}$$

Setting $u = 0$ (i.e. $r \to \infty$) this finally gives us the General Relativistic result

$$\Delta\phi = \frac{4M}{r_{\rm min}} \equiv \frac{4GM}{c^2 r_{\rm min}} = \frac{2R_{\rm S}}{r_{\rm min}}$$

This is exactly twice the deflection angle predicted by a Newtonian treatment. If we take $r_{\rm min}$ to be the radius of the Sun (which would correspond to a light ray grazing the limb of the Sun from a background star observed during a total solar eclipse) then we find that

$$\Delta\phi = \frac{4 \times 1.5 \times 10^3}{6.95 \times 10^8} = 8.62 \times 10^{-6} \text{ radians} \quad = \quad 1.77 \text{ arcsec}$$

framework – we can calculate the predicted deflection angle as light passes close to a massive object. In fact, this calculation was first carried out in 1801 by Söldner. This is the Newtonian solution we are referring to here.

Appendix 3: Einstein's tensor in the weak field approximation

A3.1: The linearised Riemann Christoffel tensor

In Minkowski spacetime the Christoffel symbols are all identically zero. This reduces the Riemann Christoffel tensor to

$$R_{\alpha\beta\gamma\delta} \;=\; g_{\alpha\mu}\, R^{\mu}_{\beta\gamma\delta} \;=\; g_{\alpha\mu}\Gamma^{\mu}_{\beta\delta,\gamma} \;-\; g_{\alpha\mu}\Gamma^{\mu}_{\beta\gamma,\delta}$$

Substituting for the Christoffel symbols in terms of the metric and its derivatives

$$R_{\alpha\beta\gamma\delta} \;=\; g_{\alpha\mu}\,\frac{g^{\mu\sigma}}{2}(g_{\sigma\beta,\delta\gamma}+g_{\sigma\delta,\beta\gamma}-g_{\beta\delta,\sigma\gamma}) \;-\; g_{\alpha\mu}\,\frac{g^{\mu\sigma}}{2}(g_{\sigma\beta,\gamma\delta}+g_{\sigma\gamma,\beta\delta}-g_{\beta\gamma,\sigma\delta})$$

This reduces to

$$R_{\alpha\beta\gamma\delta} \;=\; \frac{1}{2}(g_{\alpha\delta,\beta\gamma}+g_{\beta\gamma,\alpha\delta}-g_{\alpha\gamma,\beta\delta}-g_{\beta\delta,\alpha\gamma})$$

If $g_{\alpha\beta}=\eta_{\alpha\beta}+h_{\alpha\beta}$ then

$$R_{\alpha\beta\gamma\delta} \;=\; \frac{1}{2}(h_{\alpha\delta,\beta\gamma}+h_{\beta\gamma,\alpha\delta}-h_{\alpha\gamma,\beta\delta}-h_{\beta\delta,\alpha\gamma})$$

A3.2: The linearised Ricci tensor and curvature scalar

Contracting the Riemann Christoffel tensor it then follows that

$$R_{\mu\nu}=R^{\sigma}_{\mu\sigma\nu}=\Gamma^{\sigma}_{\mu\nu,\sigma}-\Gamma^{\sigma}_{\mu\sigma,\nu}$$

To first order this reduces to

$$R_{\mu\nu}=\frac{1}{2}\eta^{\sigma\alpha}(h_{\alpha\nu,\mu\sigma}+h_{\mu\sigma,\alpha\nu}-h_{\mu\nu,\alpha\sigma}-h_{\alpha\sigma,\mu\nu})$$

Since the partial derivatives of $\eta^{\sigma\alpha}$ are zero, we can write this as

$$R_{\mu\nu}=\frac{1}{2}\left[(\eta^{\sigma\alpha}\,h_{\alpha\nu})_{,\mu\sigma}+(\eta^{\sigma\alpha}\,h_{\mu\sigma})_{,\alpha\nu}-\eta^{\sigma\alpha}h_{\mu\nu,\alpha}{}_{,\sigma}-(\eta^{\sigma\alpha}\,h_{\alpha\sigma})_{,\mu\nu}\right]$$

which further reduces to

$$R_{\mu\nu}=\frac{1}{2}\left[(h^{\sigma}_{\nu})_{,\mu\sigma}+\left(h^{\alpha}_{\mu}\right)_{,\nu\alpha}-h_{\mu\nu,\alpha}{}^{,\alpha}-h_{,\mu\nu}\right]$$

Thus the curvature scalar R is given by

$$R=\eta^{\alpha\beta}R_{\alpha\beta}=\frac{1}{2}\eta^{\alpha\beta}\left[\left(h^{\sigma}_{\beta}\right)_{,\alpha\sigma}+(h^{\sigma}_{\alpha})_{,\beta\sigma}-h_{\alpha\beta,\sigma}{}^{,\sigma}-h_{,\alpha\beta}\right]$$

A3.3: The linearised Einstein tensor

Combining our results for the Ricci tensor and curvature scalar we find

$$\begin{aligned} G_{\mu\nu} &= \frac{1}{2}\left[(h^{\sigma}_{\nu})_{,\mu\sigma} + \left(h^{\alpha}_{\mu}\right)_{,\nu\alpha} - h_{\mu\nu,\alpha}{}^{,\alpha} - h_{,\mu\nu}\right] \\ &\quad -\frac{1}{4}\eta_{\mu\nu}\,\eta^{\alpha\beta}\left[\left(h^{\sigma}_{\beta}\right)_{,\alpha\sigma} + (h^{\sigma}_{\alpha})_{,\beta\sigma} - h_{\alpha\beta,\sigma}{}^{,\sigma} - h_{,\alpha\beta}\right] \end{aligned}$$

To see that this expression reduces to equation (108) of Section 6.3.2 it is easiest to work backwards from that equation. We have, from equation (108)

$$\begin{aligned} \text{LHS} &= \frac{1}{2}\left[h_{\mu\alpha,\nu}{}^{,\alpha} + h_{\nu\alpha,\mu}{}^{,\alpha} - h_{\mu\nu,\alpha}{}^{,\alpha} - h_{,\mu\nu} - \eta_{\mu\nu}\left(h_{\alpha\beta}{}^{,\alpha\beta} - h_{,\beta}{}^{,\beta}\right)\right] \\ &= \frac{1}{2}\left[\eta^{\alpha\sigma}\,h_{\mu\alpha,\nu\sigma} + \eta^{\alpha\sigma}\,h_{\nu\alpha,\mu\sigma} - \eta^{\alpha\sigma}\,h_{\mu\nu,\alpha\sigma} - h_{,\mu\nu} - \eta_{\mu\nu}\eta^{\alpha\gamma}\eta^{\beta\sigma}\,h_{\alpha\beta,\gamma\sigma} + \eta_{\mu\nu}\eta^{\beta\alpha}\,h_{,\beta\alpha}\right] \\ &= \frac{1}{2}\left[(\eta^{\alpha\sigma}h_{\mu\alpha})_{,\nu\sigma} + (\eta^{\alpha\sigma}h_{\nu\alpha})_{,\mu\sigma} - h_{\mu\nu,\sigma}{}^{,\sigma} - h_{,\mu\nu} - \eta_{\mu\nu}\eta^{\alpha\gamma}\left(\eta^{\beta\sigma}h_{\alpha\beta}\right)_{,\gamma\sigma} + \eta_{\mu\nu}\eta^{\alpha\beta}h_{,\alpha\beta}\right] \\ &= \frac{1}{2}\left[(h^{\sigma}_{\nu})_{,\mu\sigma} + \left(h^{\sigma}_{\mu}\right)_{,\nu\sigma} - h_{\mu\nu,\sigma}{}^{,\sigma} - h_{,\mu\nu} - \eta_{\mu\nu}\eta^{\alpha\gamma}\,(h^{\sigma}_{\alpha})_{,\gamma\sigma} + \eta_{\mu\nu}\eta^{\alpha\beta}h_{,\alpha\beta}\right] \end{aligned}$$

The first four bracketed terms match the first four terms in the expression for the Einstein tensor given above. If we now consider the remaining four terms in the above expression for the Einstein tensor, then since $\eta^{\alpha\beta} = \eta^{\beta\alpha}$

$$-\frac{1}{4}\eta_{\mu\nu}\eta^{\alpha\beta}\left[\left(h^{\sigma}_{\beta}\right)_{,\alpha\sigma} + (h^{\sigma}_{\alpha})_{,\beta\sigma}\right] = -\frac{1}{2}\eta_{\mu\nu}\eta^{\alpha\beta}\,(h^{\sigma}_{\alpha})_{,\beta\sigma}$$

Also

$$\begin{aligned} \frac{1}{4}\eta_{\mu\nu}\eta^{\alpha\beta}\left[h_{\alpha\beta,\sigma}{}^{,\sigma} + h_{,\alpha\beta}\right] &= \frac{1}{4}\eta_{\mu\nu}\eta^{\alpha\beta}\left[\eta^{\sigma\gamma}h_{\alpha\beta,\sigma\gamma} + h_{,\alpha\beta}\right] \\ &= \frac{1}{4}\eta_{\mu\nu}\eta^{\sigma\gamma}\left(\eta^{\alpha\beta}h_{\alpha\beta}\right)_{,\sigma\gamma} + \frac{1}{4}\eta_{\mu\nu}\eta^{\alpha\beta}h_{,\alpha\beta} \\ &= \frac{1}{2}\eta_{\mu\nu}\eta^{\alpha\beta}h_{,\alpha\beta} \end{aligned}$$

Thus, apart from some permutation of repeated indices, we see that the remaining four terms of our expression for the Einstein tensor match exactly the final two terms of equation (108). This establishes that equation (108) is indeed the correct expression for $G_{\mu\nu}$.

A3.4: Linearised Einstein tensor in barred form

Equation (110) is also easiest to establish in reverse – i.e. we start with equation (110) and show that its terms can be rewritten in a manner that reduces to equation (108). Consider in turn each of the four bracketed terms on the right hand side of equation (110).

$$\overline{h}_{\mu\nu,\alpha}{}^{,\alpha} \;=\; \eta^{\alpha\sigma}\overline{h}_{\mu\nu,\alpha\sigma} \;=\; \eta^{\alpha\sigma}\left[h_{\mu\nu,\alpha\sigma} - \frac{1}{2}\eta_{\nu\mu}\,h_{,\alpha\sigma}\right]$$

$$\eta_{\mu\nu}\,\overline{h}_{\alpha\beta}{}^{,\alpha\beta} \;=\; \eta_{\mu\nu}\,\eta^{\alpha\sigma}\,\eta^{\beta\sigma}\,\overline{h}_{\alpha\beta,\gamma\sigma} \;=\; \eta_{\mu\nu}\,\eta^{\alpha\sigma}\,\eta^{\beta\sigma}\left[h_{\alpha\beta,\gamma\sigma} - \frac{1}{2}\eta_{\alpha\beta}\,h_{,\gamma\sigma}\right]$$

$$\overline{h}_{\mu\alpha,\nu}{}^{,\alpha} \;=\; \eta^{\alpha\sigma}\,\overline{h}_{\mu\alpha,\nu\sigma} \;=\eta^{\alpha\sigma}\left[h_{\mu\alpha,\nu\sigma} - \frac{1}{2}\eta_{\mu\alpha}\,h_{,\nu\sigma}\right]$$

$$\overline{h}_{\nu\alpha,\mu}{}^{,\alpha} \;=\; \eta^{\alpha\sigma}\,\overline{h}_{\nu\alpha,\mu\sigma} \;=\eta^{\alpha\sigma}\left[h_{\nu\alpha,\mu\sigma} - \frac{1}{2}\eta_{\nu\alpha}\,h_{,\mu\sigma}\right]$$

Hence we can write equation (110) as

$$\begin{aligned}
G_{\mu\nu} \;=\;& -\frac{1}{2}\eta^{\alpha\sigma}\,h_{\mu\nu,\alpha\sigma} + \frac{1}{4}\eta^{\alpha\sigma}\,\eta_{\mu\nu}\,h_{,\alpha\sigma} \\
& -\frac{1}{2}\eta_{\mu\nu}\,\eta^{\alpha\gamma}\,\eta^{\beta\sigma}\,h_{\alpha\beta,\gamma\sigma} \;+\; \frac{1}{4}\eta_{\mu\nu}\,\eta^{\alpha\gamma}\,\eta^{\beta\sigma}\,\eta\alpha\beta\,h_{,\gamma\sigma} \\
& +\frac{1}{2}\eta^{\alpha\sigma}\,h_{\mu\alpha,\nu\sigma} \;-\; \frac{1}{4}\eta^{\alpha\sigma}\,\eta_{\mu\alpha}\,h_{,\nu\sigma} \\
& +\frac{1}{2}\eta^{\alpha\sigma}\,h_{\nu\alpha,\mu\sigma} \;-\; \frac{1}{4}\eta^{\alpha\sigma}\,\eta_{\nu\alpha}\,h_{,\mu\sigma} \\
\\
=\;& -\frac{1}{2}\eta^{\alpha\sigma}\,h_{\mu\nu,\alpha\sigma} \;-\; \frac{1}{2}\eta_{\mu\nu}\,\eta^{\alpha\gamma}\,\eta^{\beta\sigma}\,h_{\alpha\beta,\gamma\sigma} \;+\; \frac{1}{2}\eta^{\alpha\sigma}\,h_{\mu\alpha,\nu\sigma} \\
& +\frac{1}{2}\eta^{\alpha\sigma}\,h_{\nu\alpha,\mu\sigma} \;+\; \frac{1}{2}\eta^{\alpha\sigma}\,\eta_{\mu\nu}\,h_{,\alpha\sigma} \;-\; \frac{1}{2}h_{,\mu\nu}
\end{aligned}$$

Comparing with our expression for $G_{\mu\nu}$ in Appendix A2.3, we see that – after changing some repeated indices and using the fact that the Minkowski metric is symmetric – the above expression is identical. This establishes that equation (110) is indeed the correct expression for the $G_{\mu\nu}$ in terms of $\overline{h}_{\mu\nu}$.

Appendix 4: The Lorentz Gauge Condition

First we show that, if $\overline{h}^{\mu\alpha}{}_{,\alpha} = 0$, the final three terms on the right hand side of equation (110) vanish. Consider the bracketed terms in turn

$$\begin{aligned}
\overline{h}_{\alpha\beta}{}^{,\alpha\beta} &= \left(\eta_{\alpha\gamma}\,\eta_{\beta\sigma}\,\overline{h}^{\gamma\sigma}\right)^{,\alpha\beta} \\
&= \left(\eta_{\alpha\gamma}\,\eta_{\beta\sigma}\,\eta^{\alpha\tau}\,\eta^{\beta\epsilon}\,\overline{h}^{\gamma\sigma}{}_{,\tau\epsilon}\right) \\
&= \delta^{\tau}_{\gamma}\,\delta^{\epsilon}_{\sigma}\,\overline{h}^{\gamma\sigma}{}_{,\tau\epsilon} \\
&= \overline{h}^{\gamma\sigma}{}_{,\gamma\sigma} \quad = \quad \left(\overline{h}^{\gamma\sigma}{}_{,\sigma}\right)_{,\gamma} \quad = \quad 0
\end{aligned}$$

$$\begin{aligned}
\overline{h}_{\mu\alpha,\nu}{}^{,\alpha} &= \eta_{\mu\gamma}\,\eta_{\alpha\sigma}\,\left(\overline{h}^{\gamma\sigma}{}_{,\nu}\right)^{,\alpha} \\
&= \eta_{\mu\gamma}\,\eta_{\alpha\sigma}\,\eta^{\alpha\tau}\,\overline{h}^{\gamma\sigma}{}_{,\nu\tau} \\
&= \eta_{\mu\gamma}\,\delta^{\tau}_{\sigma}\,\overline{h}^{\gamma\sigma}{}_{,\nu\tau} \\
&= \eta_{\mu\gamma}\,\overline{h}^{\gamma\sigma}{}_{,\nu\sigma} \quad = \quad \eta_{\mu\gamma}\left(\overline{h}^{\gamma\sigma}{}_{,\sigma}\right)_{,\nu} \quad = \quad 0
\end{aligned}$$

$$\begin{aligned}
\overline{h}_{\nu\alpha,\mu}{}^{,\alpha} &= \eta_{\nu\gamma}\,\eta_{\alpha\sigma}\,\left(\overline{h}^{\gamma\sigma}{}_{,\mu}\right)^{,\alpha} \\
&= \eta_{\nu\gamma}\,\eta_{\alpha\gamma}\,\eta^{\alpha\tau}\,\overline{h}^{\gamma\sigma}{}_{,\mu\tau} \\
&\quad \eta_{\nu\gamma}\,\delta^{\tau}_{\sigma}\,\overline{h}^{\gamma\sigma}{}_{,\mu\tau} \\
&\quad \eta_{\nu\gamma}\,\overline{h}^{\gamma\sigma}{}_{,\mu\sigma} \quad = \quad \eta_{\nu\gamma}\left(\overline{h}^{\gamma\sigma}{}_{,\sigma}\right)_{,\mu} \quad = \quad 0
\end{aligned}$$

So we see that the final three terms do indeed equal zero provided $\overline{h}^{\mu\alpha}{}_{,\alpha} = 0$.

Finally we establish the equation which must be solved in order that the Lorentz gauge condition $\overline{h}^{\mu\alpha}{}_{,\alpha} = 0$ is satisfied.

Suppose we begin with arbitrary metric perturbation components $h^{(\mathrm{old})}_{\mu\nu} \neq 0$. We define

$$h^{(\mathrm{LG})}_{\mu\nu} \;=\; h^{(\mathrm{old})}_{\mu\nu} \;-\; \xi_{\mu,\nu} \;-\; \xi_{\nu,\mu}$$

where the components ξ^μ are to be determined. We can also define

$$\begin{aligned}
\overline{h}^{(\mathrm{LG})}_{\mu\nu} &= h^{(\mathrm{LG})}_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}\,\eta^{\alpha\beta}\,h^{(\mathrm{LG})}_{\alpha\beta} \\
&= h^{(\mathrm{old})}_{\mu\nu} - \xi_{\mu,\nu} - \xi_{\nu,\mu} - \frac{1}{2}\eta_{\mu\nu}\left[\eta^{\alpha\beta}\left(h^{(\mathrm{old})}_{\alpha\beta} - \xi_{\alpha,\beta} - \xi_{\beta,\alpha}\right)\right] \\
&= h^{(\mathrm{old})}_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h^{(\mathrm{old})} - \xi_{\mu,\nu} - \xi_{\nu,\mu} + \eta_{\mu\nu}\,\eta^{\alpha\beta}\,\xi_{\alpha,\beta} \\
&= \overline{h}^{(\mathrm{old})}_{\mu\nu} - \xi_{\mu,\nu} - \xi_{\nu,\mu} + \eta_{\mu\nu}\,\xi^{\beta}{}_{,\beta}
\end{aligned}$$

Now

$$\begin{aligned}
\overline{h}^{(\mathrm{LG})\,\mu\nu} &= \eta^{\mu\alpha}\,\eta^{\nu\beta}\left[\overline{h}^{(\mathrm{old})}_{\alpha\beta} - \xi_{\alpha,\beta} - \xi_{\beta,\alpha} + \eta_{\alpha\beta}\,\xi^{\sigma}{}_{,\sigma}\right] \\
&= \overline{h}^{(\mathrm{old})\,\mu\nu} - \eta^{\mu\alpha}\,\eta^{\nu\beta}\,\xi_{\alpha,\beta} - \eta^{\mu\alpha}\,\eta^{\nu\beta}\,\xi_{\beta,\alpha} + \eta^{\mu\nu}\,\xi^{\sigma}{}_{,\sigma}
\end{aligned}$$

And

$$\overline{h}^{(\mathrm{LG})\,\mu\nu}{}_{,\nu} = \overline{h}^{(\mathrm{old})\,\mu\nu}{}_{,\nu} - \eta^{\nu\beta}\,\xi^{\mu}{}_{,\beta\nu} - \eta^{\mu\alpha}\,\xi^{s}igma_{,\sigma\alpha} + \eta^{\mu\nu}\,\xi^{\sigma}{}_{,\sigma\nu}$$

i.e.

$$\overline{h}^{(\mathrm{LG})\,\mu\nu}{}_{,\nu} = \overline{h}^{(\mathrm{old})\,\mu\nu}{}_{,\nu} - \eta^{\nu\beta}\,\xi^{\mu}{}_{,\nu\beta}$$

So we can ensure that $\overline{h}^{(\mathrm{LG})\,\mu\nu}{}_{,\nu} = 0$ provided we can find gauge components ξ^μ satisfying

$$\overline{h}^{(\mathrm{old})\,\mu\nu}{}_{,\nu} = \eta^{\nu\beta}\,\xi^{\mu}{}_{,\nu\beta} = \left(-\frac{\partial^2}{\partial t^2} + \nabla^2\right)\xi^\mu$$

We can always solve this equation for well-behaved metrics using standard methods for finding particular solutions of second order partial differential equations.

After Peer Review, Accepted for Publication in the ***Proceedings of the Space, Propulsion and Energy Sciences International Forum (SPESIF),*** 24-27 February, 2009, Edited by Glen Robertson.

Input Power Requirements for High-Frequency Gravitational Wave Generators

Robert M L Baker, Jr.

GravWave® LLC and Transportation Sciences Corporation, 8123 Tuscany Avenue, Playa del Rey, CA 90293, USA
Phone: 310-823-4143, Fax: 310-821-1694, DrRobertBaker@GravWave.com

Abstract: An analysis is accomplished of the input power requirement of High-Frequency Gravitational Wave (HFGW) generators. Several techniques are explored using both off-the-shelf and advanced-nanotechnology generator elements. It is concluded that proof-of-concept test, involving *N* off-the-shelf array elements could be of meter to kilometer length and require 25 MW or less power if array elements are in a staggered arrangement. The power and size of an operational nanotechnology HFGW generator or transmitter device can be greatly reduced by the focusing effect of N^2 radiator pairs. Utilization of conventional piezoelectric Film Bulk Acoustic Resonators (FBARs), tailored and scaled for HFGW generation, could provide the initial commercial generation means. The use of the new infrared-energized ring concept of Woods and the use of a double helix array proposed by Baker may even further reduce the power and size requirements of the device to <<20 W and mm in length and width.

Keywords: High-Frequency Gravitational Waves, HFGW, gravitational wave generator, laboratory GW generation
PACS: 95.55.Ym, 04.30.-w, 04.30.Tv, 04.80.Nn

INTRODUCTION

It has been demonstrated theoretically that high frequency relic gravitational waves (HFRGWs) in the high-frequency band of 10 GHz exist having amplitudes on the order of 10^{-30} m/m and that may be detectable by the proposed ultra-high sensitivity Li-Baker Chinese HFGW detector (Li et al., 2008; Baker, Stephenson and Li , 2008). In particular, the recent papers by Leonard Grishchuk (1977a; 1977b; 2007 and Grishchuk and Soloklin, 1991), confirm that relic high-frequency gravitational waves or HFGWs certainly exist and that the origin of these HFGWs is the cosmic background associated with the big bang. The interested reader is invited to visit the site for the 2007 Invitational 2nd HFGW Workshop at: www.earthtech.org/hfgw2/ in order to view these papers. If such a detector has been built and satisfactorily tested by sensing HFRGWs, then the laboratory generation of High-Frequency Gravitational Waves (HFGWs) can be considered. The figure of merit for such gravitational wave (GW) generation (Baker, Stephenson and Li, 2008) is that the amplitude of the GW is proportional to the distance between gravitational-wave radiators (e. g., the orbital major axis of a binary black hole pair), the force change (e. g., the orbital centrifugal-force change), the GW frequency and the number of in-phase element pairs, *N*, involved in the system squared. (Romero and Dehnen, 1981; Dehnen and Romero, 2003, page 6, Eq. (2.24)). As an example, of low-frequency GW generation, for PSR 1913+16, the distance between the GW radiators (binary neutron stars on orbit) is $4.1\text{x}10^{9}$ m, the change in (centrifugal) force is $1.16\text{x}10^{33}$ N, the frequency is 7.2x10-5 s^{-1} or $7.2\text{x}10^{-5}$ Hz and, since there is but one element (one neutron-star pair), it is multiplied by unity. This yields a GW power of $10.1\text{x}10^{24}$ W (very close to the value as obtained using conventional relativistic analyses (Baker, 2000a)). If we assume that the orbital plane of the star pair is normal to the direction to the Earth and that PSR 1913+16 is $9.5\text{x}10^{15}$ m from the Earth, then the GW flux at the Earth is $2.26\text{x}10^{-8}$ Wm^{-2}. The coalescing binary black holes are much closer to each other of course; however their force changes are billions of times greater and their frequencies momentarily higher than PSR 1913+16. This yields larger GW fluxes that may be detected by the

advanced Laser Interferometer Gravitational Observatory (LIGO) and by the proposed Laser Interferometer Space Antenna. (LISA). In the laboratory the aforementioned gravitational force change could not even approach those of the celestial sources so alternative means need to be investigated thus one must turn to other means of GW generation. Analyses have shown (Baker, 2006) that only electromagnetic, rather than gravitational forces can be utilized in the laboratory and that HFGWs can be generated there. Interferometer-based GW detectors as LIGO and LISA are not sensitive to GW frequencies above at most 2 kHz and, therefore, are inadequate for HFGW detection.

LABORATORY GENERATION OF HFGWs OVERVIEW

A number of devices for the laboratory generation of HFGWs have been proposed including the GASER (a gravitational-wave LASER first proposed by Halpren and Laurent (1964), some 40 years ago) discussed by Giorgio Fontana (2003); as well as an actual LASER generator of HFGWs as discussed by Baker, Li and Li (2006). A rather practical laboratory HFGW generator is one utilizing off-the-shelf components such as magnetron energized piezoelectric crystals or Film Bulk Acoustic Resonators or FBARs (Woods and Baker, 2005; Baker, Woods and Li, 2006). Another, more exotic HFGW generator involves the use of nuclear forces (Fontana and Baker, 2006; Fontana and Binder, 2009). The figure of merit is given explicitly by Baker, Stephenson and Li (2008). This figure of merit can be extended by considering other effects. Since in the laboratory the force change could not even approach those of the celestial sources. it would seem that the magnitude of any laboratory generated GWs could be best increased (1) by utilizing electromagnetic forces rather than gravitational, (2) by increasing the distance between the gravitational radiators, (3) by increasing the GW frequency and especially (4) by developing a large number of in-phase system elements. This last effect enters as the square of the number of elements, N , as proved using General Relativity analyses by Dehnen and Romero–Borja's analyses (Romero and Dehnen, 1981; Dehnen and Romero, 2003). Such N^2 dependence also may be the key to successful laboratory generation of GWs, especially High-Frequency Gravitational Waves (HFGWs). In that regard, recent proposal by Woods (Woods and Baker, 2009; Baker and Black, 2009)) propose the use of infrared-energized atomic nuclei, electrons and or molecules, which have a very large N, contained in a stack of waveguide rings (Patents Pending). In what follows we will briefly consider several aspects of laboratory HFGW generation.

The distance between GW radiators may be proportional to the GW wavelength in that it may have a useful GW-generation limit that is less than or equal to a GW wavelength. The wavelength is inversely proportional to the GW frequency. Thus given some value for the proportional constant, say unity or the distance between radiators equal to one GW wavelength, the GW frequency cancels out. As already noted it is important to take advantage of square of the number of in phase elements for useful laboratory HFGW generation. If we slice the elements in one dimension (the dimension along the axis of HFGW generation) in order to increase the number of elements, then the change in force per element will be inversely proportional to the number of elements. For example, if the elements are sliced into one hundred separate pieces, then each piece will have one hundredth of the force of the un-sliced element. Essentially, "f = ma" and it is assumed that the acceleration of the element was the same after the split as before. This result also follows Eq. (8) of Baker, Stephenson and Li (2008a), and if there were 100 splits of an FBAR, then there would be one-hundredth the GW flux resulting from each, but 100 more of them so the net effect according to the N^2 rule would be $(100)^2/100$ or a one-hundred fold increase in HFGW flux The frequency of the split elements may be a higher value -- but the attendant increase in GW power (proportional to the square of the higher frequency) and the decrease in power due to a smaller distance between tracks (assuming that the distance between tracks is one GW wavelength, which would be smaller) would cancel and there would be no net effect on HFGW amplitude. It is concluded, therefore, that the amplitude of the generated HFGWs is proportional to the number of in phase elements, N (not the square). A large number of elements for a given HFGW-generator length can be best realized by reducing the size of the individual elements to submicroscopic size (as discussed in U. S. Patent Number 6,784,591 (Baker 2000b). Let us consider the 1.8×10^8 cell-phone film bulk acoustic resonators or FBARs, 10,000 Microwave-Magnetron, proof-of-concept laboratory HFGW generator. Assuming a 10 μm distance or margin between the 100 μm on a side conventional FBARs, the overall length of the laboratory generator will be 110×10^{-6}m x 1.8×10^8 elements = 19.8 km. The same result, of course, as that found by Baker, Stephenson and Li (2008). It will have a total HFGW power of 0.066 W and for a distance out from the last in-line, in-phase FBAR element of one HFGW wavelength (6.1 cm) it will have a flux of 3.53 Wm^{-2}, yielding a HFGW amplitude there of $A = 4.9 \times 10^{-28}$ m/m. This result differs from the result of Baker, Stephenson and Li

(2008), since they took the distance out as 1.5 HFGW wavelengths (9 cm) not one wavelength or 6.1 cm. Although the frequencies may be different one can extrapolate approximately from the results of Dehnen and Romero–Borja's (2003) analyses in which the angle of the needle-like radiation pattern is inversely proportional to the square root of the product of the distance between the radiators (the width between FBAR bands or tracks) and *N*. The distance for the system discussed here is 6.1 cm and for Dehnen and Romero–Borja's, 2003 system 0.00001 m, for a factor of 6,100 and *N* differs by $1.8x10^{8}/5x10^{7}= 3.6$ for a product of $2.2x10^{4}$ and the inverse of the square root is $6.7x10^{-3}$. Using the result from Dehnen and Romero (2003) (Eq. (4.51)) of a needle half angle of 1.7 degrees we would extrapolate to 0.0115 degrees or very approximately $2x10^{-4}$ radians.

Since we are no longer constrained to the use of rudimentary off-the-shelf components as we were for the proof-of-concept apparatus, we can manipulate the submicroscopic elements. First, we will stagger the FBARs into two bands or tracks of 100 rows each or 110x100μm = 1.1 cm wide bands of FBARs placed a wavelength or 6.1 cm apart. We will stagger the rows, as shown schematically in Fig. 1, by displacing adjacent rows in the bands by 1.1 μm. Thus the overall length will be reduced to 198 m. Second, we can slice the 100 μm length of each FBAR element, along the direction of travel of the HFGW build up, into one-hundred 1 μm wide slices (exhibiting 0.1 μm margins). The power of each of these small slices will be 10^{7} W (10,000 off-the-shelf kilowatt Magnetron energizers) with their power spread over $1.8x10^{8}x100$ FBAR slices or 0.56 mW for each slice. If we assume a 1 mA current, then the voltage across the 1 μm wide slices is 0.56 V for a 560,000 V/m field strength, which should not arc. The staggered row displacements are now reduced to 11 nm. The overall length will be reduced to about 198 cm. Concentrating the 10 MW power to each of these 1.1 cm wide bands may prove to be difficult. Thus, as an example, we will replace the continuous-wave Magnetrons by a pulsed microwave source having one-microsecond-long pulses one second apart. The required average power for each FBAR band will now be 10 W. As a practical nanotechnology limit, we could reduce the slice width by two orders of magnitude to 10 nm. This would also require that the row displacements would be 110 pm (we are now into atomic if not sub-atomic dimensional changes). The overall length could be reduced to about 2 cm or the amplitude of the HFGWs could be increased to $A = 4.9x10^{-26}$. In this latter case the average energizing microwave power applied to each band would need to be increased to 1 kW. A preferred compromise in this apparent nano-technology limit might be to reduce the HFGWs generator's length to about 20 cm and increase the HFGW amplitude A to $4x10^{-27}$ m/m.

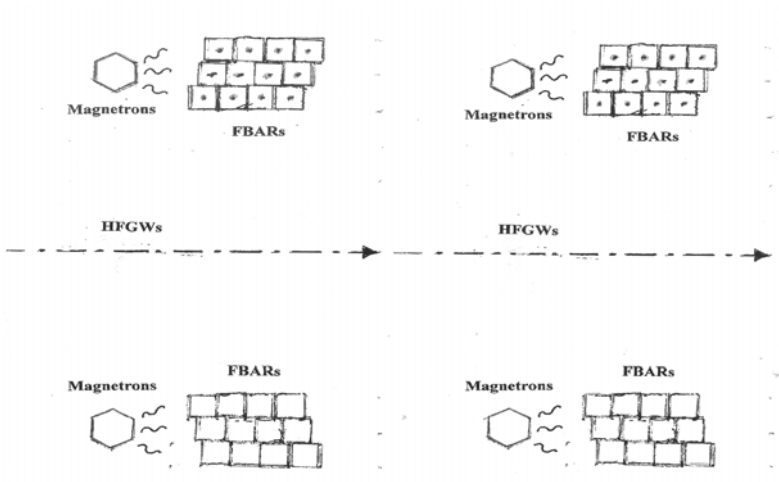


FIGURE 1. Illustration of the Parallel Staggered Tracks of FBARs

LABORATORY-GENERATION HFGW ELEMENT OVERVIEW

The complimentary approach to optimizing a practical HFGW generator is to increase the force produced by each element without increasing the required power, i.e., increasing element efficiency. This was initially done using

the modern light-weight piezoelectric FBARs rather than the heavy 10-gram crystals considered by Dehnen and Romero-Borja that were of 1981 vintage. Their paper on HFGW generation, utilizing conventional general-relativity analyses of a piezoelectric-crystal HFGW generator (Romero and Dehnen, 1981; Dehnen and Romero, 2003), agrees to within half a percent with the approach of Baker, Stephenson and Li (2008), as shown in Baker (2007). Special designs of FBAR-like elements for optimum force-generation efficiency will improve the HFGW generator performance beyond that for the usual cell-phone FBAR designs. Also staggered columns of split FBARs as discussed by Baker (2009) could be placed on double-helix ribbons (Baker and Black, 2009) with the Magnetron energizers conveniently situated on their axes. Another approach to element design is to utilize lasers whose targets are the force-generating elements. This HFGW generator means is initially presented in Baker, Li and Li (2006). Utilization of myriads of nano-scale lasers could generate high-frequency HFGW pulses.

The use of submicroscopic particles for HFGW generation was suggested by Baker (2000b) and was utilized as the basis of an infrared (IR) energized generator by Woods and Baker (2009). There is significant promise for the IR-generated HFGWs (Patents Pending). If you have a standing wave in a waveguide ring and excite it properly, then you have a GW source at its center. Please see Fig. 2 from Woods and Baker (2009). The easiest configuration to analyze requires that two counter-propagating traveling waves be excited inside this waveguide, thus producing a standing optical wave inside the guide. Regions of the guide separated by $\lambda/2$ will oscillate in anti-phase but the resultant HFGW will be in phase, since the produced GW is at doubled frequency. Therefore, this configuration is equivalent to the ring of discrete acoustic resonators (small masses) proposed previously by Woods and Baker (2005) for terrestrial HFGW production. Because the active material vibrates in phase and in opposite pairs and has circular symmetry, all the generated GW will combine in phase at the center of the torus. The GW flux produced at its center is proportional to the N submicroscopic particle pairs in the ring. There is no N^2 build up but there is a N build up. If you have a stack of n rings, which are excited in sequence at light speed as a generated, growing GW passes by, then you have a n^2 build up in GW flux. The IR wavelength is about 2.5×10^{-6}m and the IR waveguide has a cross-sectional area radius of $\lambda/4$ in order for it to be a monomode (lowest order mode) so that the phase doesn't change across the waveguide. Thus the cross-sectional area of each IR ring is $\pi\ (2.5\times10^{-6}\text{m}\ /4)^2 = 1.23\text{x}10^{-12}\text{m}^2$ and its diameter is 1.25×10^{-6} m. The volume of each 100-m radius toroidal ring is $2\pi\ (100)(\ 1.23\text{x}10^{-12}) = 7.7\text{x}10^{-12}\ \text{m}^3$. Woods and Baker (2009) divide the mass density of pentane by its molecular mass and that gives the density of jerkable masses = $6.3\times10^{28}\ \text{m}^{-3}$. Thus the number of masses in a 100-m radius circular wave guide $2N = (6.3\times10^{28})(\ 7.7\text{x}10^{-12}) = 4.85\text{x}10^{17}$ submicroscopic "particles" or potentially jerkable masses. According to Table 1 of Woods and Baker (2009) for pentane $P_i = 4.62\text{x}10^{-16}$ W. Thus the flux for all of the mass pairs in a single ring from Eq. (8) is $(1/2)(2N)(0.01146)\ P_i = 1.29\ \text{Wm}^{-2}$.

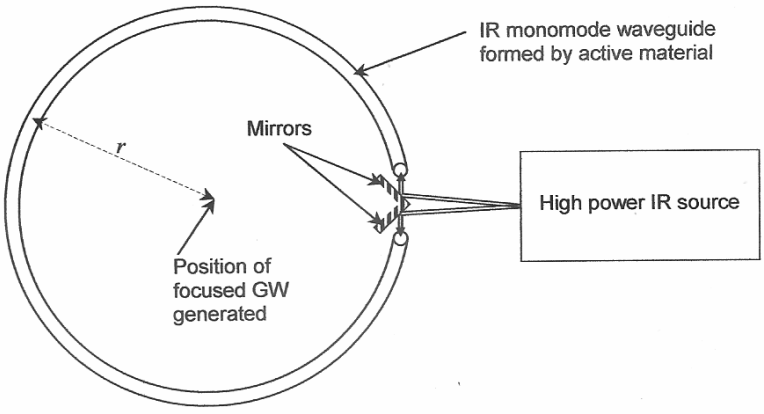


FIGURE 2. Circular Resonator Geometry Using Infra-Red Excitation.

Baker and Black (2009) reduce the ring radius to one meter, but set up 100 rings, concentrically (side by side in the same plane) with an average radius of the one meter. The reduced radius drops the P_i by $(100)^2$ to $4.62\text{x}10^{-20}$, but because of the 100 concentric rings the $N = 4.85\text{x}10^{17}/2$ remains the same. Thus the flux for a single "plate" of concentric rings is $1.29\text{x}10^{-4}\ \text{Wm}^{-2}$. We now stack some 10^6 of these plates on top of one another. Thus a 1.25-m high stack. In this case $n = 10^6$ and we can apply the n^2 law. Thus a HFGW total flux of $1.29\text{x}10^{8}\ \text{Wm}^{-2}$ will be

generated by the stack. Of course (as R. C. Woods has pointed out, Woods and Baker (2009)) we need to be careful how much power is fed to each ring. One possible arrangement is to feed the output of one ring to the input of the next. The problem here is that the source won't have a long enough coherence length, even if the attenuation of the IR doesn't kill the power after a ring or two. To avoid this, from one source the available energizing power, on the order of a MW but possibly reducible to less than 20 W, could be divided equally between all the rings and fed to them up the stack at the speed of light. Thus there is some uncertainty in the IR-ring generator and the FBAR generators would be the best for the proof-of-concept tests and the first practical applications. As computed by Baker and Black (2009) for the stack of rings, the amplitude of the laboratory-generated HFGWs is $A = 1.21 \times 10^{-28}$. Clearly there are a number of opportunities to enhance HFGW generation performance, utilizing special element designs, either by reducing the generator size or increasing the generated HFGW amplitude or both.

.

CONCLUSION

After proving out the Li-Baker detector on relic HFGWs, proof-of-concept tests (involving only off-the-shelf HFGW-array elements) can be accomplished with lengths on the order of 2 km depending upon placement of Magnetrons, FBARs and power lines. The length could be greatly reduced by a factor of 10 to 1000 (due to the N^2 effect for HFGW generated flux) if each track consisted of several close (110μm) parallel -multiple, staggered rows of off-the-shelf FBARs as shown schematically in Fig. 1. (Power proportional to N elements and beam area inversely proportional to N since the pencil beam is narrowed as N increases as shown in Baker and Black (2009)). Such a design would also allow for the power of the Magnetron beam (properly focused) to be more completely absorbed by the FBARs. The power requirement for the 20,000 off-the-shelf Magnetrons on the two linear tracks (not staggered) for the proof-of-concept tests would be about 25 MW, so that a power substation of that size would be required. Power and size of an operational nanotechnology HFGW generator or transmitter utilizing advanced FBARs can be greatly reduced by the focusing effect of N^2 radiator pairs. The use of the IR-rings concept of Woods and Baker (2009) may even further reduce the power and size requirements of the device as a future development. The use of nanotechnology could conceivably result in power and HFGW-generator size of less than 20 watts and only millimeters in length and width.

ACKNOWLEDGMENTS

The research and preparation was sponsored by GravWave® LLC and the GravWave division of Transportation Sciences Corporation.

REFERENCES

Baker, R. M L, Jr., "Gravitational Wave Generator," *United States Patent number 6,417,597B*, filed July 14, 2000a, granted July 19, 2002.

Baker R. M L, Jr., "Gravitational Wave Generator Utilizing Submicroscopic Energizable Elements," *United States Patent Number 6,784,591***,** filed December 27, 2000b, granted August 31, 2004.

Baker R. M L, Jr., "Layperson's Description of High-Frequency Gravitational Waves or "HFGWs," Lecture given at the ***Science and New Technology Special Interest Group*** at ***The California Club,*** Los Angeles, California, USA, October 17, 2006, available at: http://www.drrobertbaker.com/docs/Layperson's%20Description%20of%20HFGW.pdf

Baker R. M L, Jr., "Informal Comparative Analysis of Dehnen and Romero-Borja's HFGW-Generation Calculations with those of Baker, Stephenson and Li," pp. 16-21, October 17, 2007, http://www.gravwave.com/docs/Analysis%20of%20Lab%20HFGWs.pdf, accessed September 13, 2008.

Baker R. M L, Jr., Li, F. Y. and Li, R. "Ultra-High-Intensity Lasers for Gravitational Wave Generation and Detection" in the proceedings of *Space Technology and Applications International Forum (STAIF-2006)*, edited by M.S. El-Genk, American Institute of Physics Conference Proceedings, Melville NY **813**, pp. 1249-1258, 2006. available at http://www.drrobertbaker.com/docs/AIP;%20HFGW%20Laser%20Generator.pdf.

Baker,. R. M L Jr., R. Clive Woods and Fangyu Li, "Piezoelectric-Crystal-Resonator High-Frequency Gravitational Wave Generation and Synchro-Resonance Detection," in the proceedings of *Space Technology and Applications International Forum (STAIF-2006)*, edited by M.S. El-Genk, American Institute of Physics Conference Proceedings, Melville NY **813**, 2006. pp. 1280-1300, 2006.

Baker, R. M L, Jr., Stephenson, G. V. and Li . F. Y., "Analyses of the Frequency and Intensity of Laboratory Generated HFGWs," in the proceedings of *Space Technology and Applications International Forum (STAIF-2008)*, edited by M.S. El-Genk, American Institute of Physics Conference Proceedings, Melville, NY **969**, pp. 1036-1044, 2008a http://www.gravwave.com/docs/Analysis%20of%20Lab%20HFGWs.pdf

Baker R. M L, Jr., Stephenson G.V. and Li F. Y., "Proposed Ultra-High Sensitivity HFGW Detector," Proc. AIP Space Technology and Applications Int. Forum, Albuquerque, New Mexico, U.S.A. **969** 1045-1054, 2008b; available at http://www.gravwave.com/docs/Proposed%20Ultra-High%20Sensitivity%20HFGW%20Detector%2005-15-08.pdf

Baker R. M L, Jr., and C. S. Black, C. S., "Radiation Pattern for a Multiple-Element HFGW Generator," After Peer Review, Accepted for Publication in the *Proceedings of the Space, Propulsion and Energy Sciences International Forum (SPESIF*)**,** 24-27 February, Edited by Glen Robertson. (Paper 035), 2009.

Dehnen, H. and Romero, F. R. "Generation of GHz – THz High-Frequency Gravitational Waves in the laboratory," paper HFGW-03-102, *Gravitational-Wave Conference*, The MITRE Corporation, May 6-9, pp. 22 ff, 2003

Fontana, G. and Baker, R. M L, Jr., "The high-temperature superconductor (HTSC) gravitational laser (GASER)," paper HFGW-03-107, *Gravitational-Wave Conference*, The MITRE Corporation, May 6-9, 2003, available at http://earthtech.org/hfgw2/papers/Talk%203%20The%20HTSC%20GASER.pdf.

Fontana, G. and Baker, R. M L, J, "Generation of Gravitational Waves with Nuclear Reactions," in the proceedings of *Space Technology and Applications International Forum (STAIF-2006)*, edited by M.S. El-Genk, American Institute of Physics Conference Proceedings, Melville NY **813**, pp. 1352-1358, 2006. http://www.drrobertbaker.com/docs/AIP;%20HFGW%20Nuclear%20Generator.pdf

Fontana, G. and Binder, B., "Electromagnetic to Gravitational wave Conversion via Nuclear Holonomy," After Peer Review, Accepted for Publication in the *Proceedings of the Space, Propulsion and Energy Sciences International Forum (SPESIF*)**,** 24-27 February, Edited by Glen Robertson. (Paper 015), 2009.

Grishchuk L.P., "Gravitational Waves in the Cosmos and the Laboratory," Sov. Phys. Usp. **20** 319-334, 1977a (original Usp. Fiz. Nauk **121** 629-656, 1977)

Grishchuk L.P., "Graviton Creation in the Early Universe," Ann. N.Y. Acad. Sci. **302** 439-444, 1977b

Grishchuk L.P., "High-Frequency Relic Gravitational Waves, Their Detection and New Approaches," Proc. 2nd HFGW Workshop, IASA Austin, Texas, U.S.A., 2007; available at http://earthtech.org/hfgw2/

Grishchuk L.P. and Solokhin M., "Spectra of Relic Gravitons and the Early History of the Hubble Parameter," Phys. Rev. **D43** 2566-2571, 1991

Halpren, L. and Laurent, B.), "On the gravitational radiation of a microscopic system," *IL NUOVO CIMENTO*, Volume XXXIIIR, Number 3, pp. 728- 751, 1964

Li F. Y., Baker, R. M L, Fang Z., Stephenson G.V. and Chen, Z., "Perturbative Photon Fluxes Generated by High-Frequency Gravitational Waves and Their Physical Effects," *European Phys. J. **C,*** Published Online: 30 July 2008; available at http://www.gravwave.com/docs/Li-Baker%206-22-08.pdf

Romero B. F. and. Dehnen, H. "Generation of gravitational radiation in the laboratory," *Z .Naturforsch*, Volume **36a**, pp. 948-955, 1981

Woods, R. C. and Baker, R. M L, Jr., "Gravitational Wave Generation and Detection Using Acoustic Resonators and Coupled Resonance Chambers," in the proceedings of *Space Technology and Applications International Forum (STAIF-2005)*, edited by M.S. El-Genk, American Institute of Physics Conference Proceedings, Melville, NY **746**, 1298, 2005

Woods, R. C. and Baker, R. M L, Jr., "Generalized Generators of Very-High-Frequency Gravitational Waves Including Ring and Helix Devices," the ***Proceedings of the Space, Propulsion and Energy Sciences International Forum (SPESIF***), 24-27 February, Edited by Glen Robertson. (Paper 001), 2009.

P139 - Relativity
Problem Set 12: GRAVITATIONAL WAVES 2003 April

G. F. SMOOT
Department of Physics,
University of California, Berkeley, USA 94720

1 GRAVITATIONAL PLANE WAVES

1.1 Introduction to Gravitational Waves

By analogy to electromagnetism we expect that there should be gravitational waves – moving disturbances in the metric. These should have the following properties:
(a) Gravitational Waves should exist and travel with speed c which is the only Lorentz invariant velocity and is the relation between space and time.
(b) Every from of matter and energy with which we are familiar has the same gravitational effect – positive attraction. i.e. like having all one sign electric charge Thus no dipole radiation only quadrupole and higher order.
(c) Gravitational Waves are very weak. Consider two neutron stars of one solar mass ($M_{NS} = \mathrm{M}_\odot$) and radius $r_{NS} = 6$ km orbiting each other essentially in contact. (This is for example, clearly tidal forces will have a very disruptive effect on the two neutron stars.) The characteristic acceleration at the neutron stars is

$$a = \frac{2G\mathrm{M}_\odot}{(2r_{NS})^2} \simeq 9 \times 10^{11} m/s^2 \simeq 10^{11} g$$

The orbital period for half a rotation (symmetric pattern with identical neutron stars and thus the period of the graviational radiation) is $T/2 = 2\pi/\sqrt{r^3/G\mathrm{M}_\odot} \simeq 10^{-4}$ sec, so that the frequency of the gravitational radiation is about 10 kHz.

Consider a test mass at the Earth one parsec away (1 pc = 3.26 light years). This distance is the distance to the nearest known stars, excepting the Sun, and the test mass will be in the next section our detector. The acceleration $A_\oplus$ from the gravitaional radiation, i.e. proportional to $1/r$ rather than $1/r^2$, is

$$A = \frac{G\,(2m_\odot)\,a}{c^2 r} \simeq 0.09 m/s^2 \simeq 10^{-2} g$$

(d) Detection of gravitational waves is difficult. Consider a resonant frequency pendulum tuned to the gravitational waves. Using the formula for a gravitational pendulum $\omega = \sqrt{g/\ell}$, one finds that to get $f = 10^4$ Hz requires $\ell = 20\mathring{A}$. (24.8A)

However, the wavelength of the gravitational wave is $\lambda_{GW} = c/f \simeq 3x10^5 km/s/10^4/s = 30$ km. Thus a quarter-wavelength long detector would need to

be about 7.5 km. Thus the single pendulum is not a good match to the gravitational wave.

The velocity produced by the gravitational wave on the test mass on Earth is of order $V = A/\omega = A/(2\pi f) \simeq 3 \times 10^{-6}$ m/s. This is a very low velocity and difficult to dectect even with the Mössbauer effect. The amplitude of oscillation is $A/(2\pi f)^2 \simeq 0.5\mathring{A}$.

(e) The generation of gravitational waves is difficult. Just as it is difficult to detect gravitational waves; it is difficult to generate them. We will find that for astrophysical systems the dimensionaless strain h is roughly

$$h \sim \frac{r_{Schwarschild}}{r}\left(\frac{v}{c}\right)^2$$

The power in gravitational waves is proportional to the square of the internal power of the radiating object. The constant of proportionality is equivalent to dividing by a very large power.

$$P = \frac{P_{internal}^2}{P_{cGW}}, \quad \text{where} \quad P_{cGW} = \frac{c^5}{G} = 3.63 \ \times 10^{59}\text{erg/s} = 2 \times 10^5 \text{M}_\odot/s$$

1.2 Properties of Gravitational Waves

There are many similarities between gravitation and electromagnetism. In particular they have similar Poisson Equations relating the field to its sources. For electromagnetism

$$\Box^2 A^\nu = 4\pi J^\nu, \text{ where} \Box^2 A^\nu \equiv \partial_\mu \partial^\mu A^\nu - \partial^\nu \partial_\mu A^\mu = \frac{1}{c^2}\frac{\partial^2 A^\nu}{\partial t^2} - \frac{\partial^2 A^\nu}{\partial x^2} - \frac{\partial^2 A^\nu}{\partial y^2} - \frac{\partial^2 A^\nu}{\partial z^2} \quad (1)$$

where J^ν is the 4-D electric current and the last equality holds for the gauge condition $\partial_\mu A^\mu = 0$ and the usual rectilinear coordinates. For gravity one has (in the weak field limit)

$$\Box^2 \phi^{\mu\nu} = -16\pi G T^{\mu\nu} \quad \text{or} \quad \partial_\lambda \partial^\lambda (h^{\mu\nu} - \frac{1}{2}\eta^{\mu\nu} h) = -16\pi G T^{\mu\nu} \quad (2)$$

where $\kappa = \sqrt{16\pi G} \cong 2.04 \times 10^{-24}$ s (g-cm)$^{-1/2}$ and $T^{\mu\nu}$ the stress-energy tensor. Note the relation between the linearized part of the metric $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ where $\eta_{\mu\nu}$ is the Minkowski metric tensor and the field variable $\phi^{\mu\nu}$

$$\phi^{\mu\nu} \equiv h^{\mu\nu} - \frac{1}{2}\eta^{\mu\nu} h \quad \rightarrow \quad h^{\mu\nu} = \phi^{\mu\nu} - \frac{1}{2}\eta^{\mu\nu}\phi \quad (3)$$

Such a field is convenient for the gauge condition

$$\partial_\mu \phi^{\mu\nu} = 0 \quad (4)$$

It is no surprise that both Maxwell's and the linearized Einstein equations should have radiation solutions and that both solutions waves move with the speed

of light ($c = 1$ in our notation and is left out of the equation above). It is clear that adding any solution of the form

$$\Box^2 \phi^{\mu\nu} = 0, \quad \text{or, if} \quad \partial_\mu \phi^{\mu\nu} = 0, \quad \partial_\lambda \partial^\lambda \phi^{\mu\nu} = 0 \tag{5}$$

will work.

Try the plane-wave solution

$$\phi^{\mu\nu} = \epsilon^{\mu\nu} cos k_\alpha x^\alpha \tag{6}$$

where $\epsilon^{\mu\nu}$ is a constant tensor, and k_α is a constant vector in the direction of propagation, which are named the polarization tensor and the wavevector, respectively.

(a) Show that this is a solution to equation 5, if

$$k^\alpha k_\alpha = 0, \quad \text{and} \quad \epsilon^{\mu\nu} k^\mu = 0 \tag{7}$$

meaning that $\omega = ck$ where $k = \sqrt{k_x^2 + k_y^2 + k_z^2} = |\mathbf{k}|$ and thus the speed of the gravity wave is the speed of light. Note that this means in the weak field limit there is no dispersion. (The c for your information later. Ignore it for this exercise.)

Now without loss of generality, we can restrict ourself to a plane wave moving in the z-direction $k^\alpha = (\omega, 0, 0, \omega)$ so that

$$\phi^{\mu\nu} = \epsilon^{\mu\nu} cos(\omega t - \omega z) = \epsilon^{\mu\nu} cos(kt - kz) \tag{8}$$

(b) You can either show or take my word for it that equations 7 and the requirement that $h^{\mu\nu}$ and $\phi^{\mu\nu}$ are symmetric means that there only only six possible independent polarizations. Note that there are 16 components to the tensor and symmetry reduces that to 10 independent components. The two relations of equation 7 cut the 10 to 6. (You heard that these correspond to the ± 0, ± 1, and ± 2 helicity modes in class.)

(c) Two of the independent polarization solutions can be

$$\epsilon_\oplus^{\mu\nu} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{9}$$

and

$$\epsilon_\otimes^{\mu\nu} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{10}$$

Show that these are traceless and transverse to the direction of propagation. The claim is that these two are the only one of their set of six that can actually carry energy and momentum. The others can be transformed away by gauge transformations.

We can return to general plane waves

$$\begin{aligned} \phi_{\oplus}^{\mu\nu} &= A_{\oplus}\epsilon_{\oplus}^{\mu\nu} cos k_\alpha x^\alpha \\ \phi_{\otimes}^{\mu\nu} &= A_{\otimes}\epsilon_{\otimes}^{\mu\nu} cos k_\alpha x^\alpha \end{aligned}$$

with energy-momentum tensors

$$\begin{aligned} t_{\oplus}^{\mu\nu} &= A_{\oplus}^2 k^\mu k^\nu sin^2 k_\alpha x^\alpha \\ t_{\otimes}^{\mu\nu} &= A_{\otimes}^2 k^\mu k^\nu sin^2 k_\alpha x^\alpha \end{aligned}$$

The average flux in the direction of propagation is

$$t = \frac{1}{2c}(A_{\oplus}^2 + A_{\otimes}^2)\omega^2 = \rho_{GW}/c \tag{11}$$

and the energy density is related to the flux by c. The stress-energy tensor for waves propagating in the z-direction has for its non-zero components

$$T^{00} = \frac{1}{c}T^{0z} = \frac{1}{c}T^{zz} = c^2 < A_{\oplus}^2 + A_{\otimes}^2 >= \frac{1}{16\pi}\frac{c^2}{G} < h_+^2 + h_\times^2 > \tag{12}$$

1.3 Relative Motion

In class I showed that such gravitational plane waves caused two test particles to oscillate relative to each other. I used the weak field geodesic equation

$$\frac{dp_\mu}{d\tau} = \frac{1}{2}m h_{\alpha\beta,\mu}u^\alpha u^\beta \tag{13}$$

or for a massive particle

$$\frac{du_\mu}{d\tau} = \frac{1}{2}h_{\alpha\beta,\mu}u^\alpha u^\beta \tag{14}$$

to show that a particle at rest $u^\alpha = (1,0,0,0)$ experiences no acceleration and remains at rest. However, two particles that were separated by a distance d in a direction transverse to the direction of propagation and aligned with the $\oplus$ polarization experience a periodic variation in their position to first order

$$\delta l \cong d \times (\frac{1}{2}A_{\oplus} cos\omega t) \tag{15}$$

If a typical strong ($A_{\oplus} \simeq 10^{-18}$) gravity plane wave passes through the earth-moon system ($d \cong 3.84 \times 10^{10}$ cm), what is the variation in the earth-moon distance? If laser ranging to the moon is accurate to 10 cm, can the motion be readily detected?

If the relative distance to a space craft at a distance of 5 astronomical units (7.5×10^{13} cm) can be measured to 1 mm by doppler tracking, can one detect this wave? How accurately must one measure the separation to be able to observe such a wave? Could it be done with laser interferometry?

1.4 Effective tidal Force

A gravitational plane wave causes and apparent deformation of objects. We discussed the appearance of a tube-shaped distribution of matter lying along the direction of propagation of the plane wave. For the two linear polarizations $\oplus$ and $\otimes$ the tube appears alternately squashed in the directions of the polarization axis. Specifically for a plane wave propagating along the z-axis, the $\oplus$ polarization has at a fixed time (for the right phase)

$$\begin{aligned} x(z) &= r(1 - \frac{1}{2}A_{\oplus}\cos\omega z) \\ y(z) &= r(1 - \frac{1}{2}A_{\oplus}\sin\omega z) \end{aligned}$$

The $\otimes$ polarization would look the same but rotated 45°.

For the circular polarizations

$$\begin{aligned} \epsilon^{\mu\nu}_{RH} &= (\epsilon^{\mu\nu}_{\oplus} - i\epsilon^{\mu\nu}_{\otimes})e^{ik_\alpha x^\alpha} \\ \epsilon^{\mu\nu}_{LH} &= (\epsilon^{\mu\nu}_{\oplus} + i\epsilon^{\mu\nu}_{\otimes})e^{ik_\alpha x^\alpha} \end{aligned}$$

the shape is a twisted squashed tube. Waves with the first circular polarization $\epsilon^{\mu\nu}_{RH}$ have positive helicity meaning the angular momentum is parallel to the direction of energy flux. Waves with the second circular polarization $\epsilon^{\mu\nu}_{LH}$ have negative helicity meaning the angular momentum is antiparallel to the direction of energy flux. The angular momentum in the wave is just $2/\omega$ times the energy in the wave.

What is the effective tidal force on the particles in a tubular shape for a gravity wave of amplitude $A_{\oplus} \simeq 10^{-18}$? Hint: Use the equivalence principle and calculate the effective differential acceleration.

$$\frac{d^2x}{dt^2} = -\frac{1}{2}A_{\oplus}x_0\frac{d^2}{dt^2}cos\omega t \tag{16}$$

You can get a similar equation for y. Show that the radial component of the tidal force is

$$f_r = m\ \frac{1}{2}A_{\oplus}r_0\ \omega^2 cos\omega t\ cos2\phi \tag{17}$$

where $r_0^2 = x_0^2 + y_0^2$ is the radial coordinate and ϕ is the azimuthal angle to the polarization axis.

2 Generation of Gravitational Waves

We continue briefly with our analogy to electromagnetism. It is appropriate to note that the relative strength of gravity to electromagnetism for a simple effect such as and atomic transition of energy order $E \sim 1$ eV is $GE^2/e^2 \approx 3 \times 10^{-54}$. Thus we cannot expect to find gravity waves being significant in atomic or individual

particle interactions but will be significantly only as a bulk effect for a large aggregate collection. This (equivalent to the small value of κ) justifies our weak field approximation. This keeps the waves linear and with no dispersion.

The general solution to Poisson equations (1) and (2) are

$$\Box^2 A^\nu = 4\pi J^\nu, \qquad A^\nu(\mathbf{x},t) = \frac{1}{c}\int d^3x' \int dt' \frac{J^\nu(\mathbf{x}',t')}{|\mathbf{x}-\mathbf{x}'|}\delta(t'+|\mathbf{x}-\mathbf{x}'|/c-t) \tag{18}$$

integration over t' replaces $J^\nu(\mathbf{x}',t')$ by $J^\nu(\mathbf{x}',t-|\mathbf{x}-\mathbf{x}'|/c-t)$ and similarly

$$\Box^2 \phi^{\mu\nu} = -16\pi G T^{\mu\nu}, \quad \phi^{\mu\nu} = -\frac{16\pi G}{4\pi}\int d^3x' \frac{T^{\mu\nu}(\mathbf{x}',t-|\mathbf{x}-\mathbf{x}'|/c-t)}{|\mathbf{x}-\mathbf{x}'|} \tag{19}$$

(These were derived from the relation $\nabla^2(1/|\mathbf{x}-\mathbf{x}'|) = -4\pi\delta(\mathbf{x}-\mathbf{x}').$)

In considering radiation one usually divides the situation into various zones when the size of the system radiating is smaller than a wavelength.

Zone	*Conditions*
Near or Static	$d << r << \lambda$
Intermediate	$d << r \sim \lambda$
Far or Radiation Zone	$d << \lambda << r$

(20)

where d is the characteristic size of the radiating system, λ is the wavelength, and r is the distance to the observer.

In the far field region the field scales as $1/r$ and the field is perpendicular to $\mathbf{r}$ as can be seen from the following argument. In the far field zone $kr >> 1$ so that $|\mathbf{x}-\mathbf{x}'| \cong r - \hat{n}\cdot\vec{\mathbf{x}}'$ so that

$$\lim_{kr\to\infty} \vec{\mathbf{A}} \to \frac{e^{ikr}}{cr}\int e^{-ik\hat{n}\cdot\mathbf{x}'}\ \vec{\mathbf{J}}\ d^3x' \cong \frac{e^{ikr}}{cr}\sum_n \frac{(-ik)^n}{n!}\int (\hat{n}\cdot\vec{\mathbf{x}}')^n\ \vec{\mathbf{J}}\ d^3x' \tag{21}$$

where the second part comes from expanding $exp(-ik\hat{n}\cdot\vec{\mathbf{x}}')$ in a Taylor series since we have assumed that $kd << 1$ which implies that $k|x| << 1$.

We expect that quadrupole radiation is the lowest order radiation that gravity can produce since all gravitational charges carry the same sign and the amplitude of the field will be proportional to k^3 and the energy radiated proportional to k^6.

Aside: The general solution

$$\vec{\mathbf{A}} = \frac{4\pi ik}{c}\sum_{l,m} h_l^{(1)}(kr)Y_{lm}(\theta,\phi)\int j_l(kr')Y_{lm}^*(\theta',\phi')\ \vec{\mathbf{J}}\ d^3x' \tag{22}$$

where $h_l^{(1)}(kr)$ are the Hankel functions, $Y_{lm}(\theta,\phi)$ are the spherical harmonics, and $j_l(kr)$ are the Bessel functions.

Making similar approximation for the distance $|\mathbf{x}-\vec{\mathbf{x}}'| \cong r - \hat{n}\cdot\vec{\mathbf{x}}'$ for the gravitational case, one has

$$\phi^{\mu\nu} \cong \frac{16\pi G}{4\pi r}\int d^3x' T^{\mu\nu}(t-r,\mathbf{x}') \tag{23}$$

In the linear approximation, $T^{\mu\nu}$ satisfies the gauge condition

$$\partial_\nu T^{\mu\nu} = 0 \tag{24}$$

which implies

$$\frac{\partial}{\partial t}T^{k0} = -\partial_l T^{kl} \quad \text{and} \quad \frac{\partial}{\partial t}T^{00} = -\partial_l T^{0l}. \tag{25}$$

We can use the first relationship to get

$$\int T^{kl} d^3x = \frac{1}{2}\frac{\partial}{\partial t}\int (T^{k0}x^l + T^{l0}x^k) d^3x \tag{26}$$

and the second to get

$$\int (T^{k0}x^l + T^{l0}x^k) d^3x = \frac{\partial}{\partial t}\int T^{00}x^k x^l d^3x \tag{27}$$

This can be proven by integrating the right hand side by parts. Combined these give us

$$\int T^{kl} d^3x = \frac{1}{2}\frac{\partial^2}{\partial t^2}\int T^{00}x^k x^l d^3x. \tag{28}$$

This indicates that the integral over the stress-energy can be expressed in terms of derivatives of T^{00}. This is simply conservation of energy and momentum.

Putting this back into equation (23) we have

$$\phi^{\mu\nu} \cong -\frac{16\pi G}{4\pi r}\int T^{\mu\nu}(t-r, \mathbf{x}') = -\frac{16\pi G}{8\pi r}\frac{\partial^2}{\partial t^2}\int T^{00}(t-r, \mathbf{x}') d^3x' \tag{29}$$

We will make two other approximations in addition to the weak field and the far zone in our treatment; (1) the gravitational effects are relatively small so that we can replace T^{00} by ρ and (2) we will neglect special relativistic corrections ($v/c << 1$). Note that these approximations mean that we will be in the weak field limit in that the dimensionless strain

$$h \sim \frac{r_{schwarzschild}}{r}\frac{v^2}{c^2} \tag{30}$$

Thus for non-relativistic matter one has

$$\phi^{kl} \cong -\frac{16\pi G}{8\pi r}\frac{\partial^2}{\partial t^2}\int \rho(t-r, \mathbf{x}') d^3x' \tag{31}$$

This integral can be expressed in terms of the quadrupole moment tensor.

$$Q^{kl} = \int (x^k x^l - \frac{1}{3}r^2\delta^l_k)\rho(t-r, \mathbf{x}') d^3x \tag{32}$$

so that

$$\phi^{kl} \cong -\frac{16\pi G}{8\pi r}\{\frac{\partial^2}{\partial t^2}Q^{kl} + \delta^l_k\frac{\partial^2}{\partial t^2}\int r'^3\rho(t-r, \mathbf{x}') d^3x'\} \tag{33}$$

For the calculation of energy flux, the term proportional to δ^l_k can be omitted because that part of the wave carries no energy as it has the wrong polarization.

$$\phi^{kl} \cong -\frac{16\pi G}{8\pi r}\frac{\partial^2}{\partial t^2}Q^{kl} = -\frac{16\pi G}{8\pi r}\ddot{Q}^{kl} \tag{34}$$

This is the lowest order term for gravitational radiation and is clearly quadrupolar as asserted earlier. There can be higher order terms; however, in general one expects them to be significantly smaller than the quadrupole contribution.

It can be shown that the energy radiated per unit solid angle is

$$-\frac{d^2E}{dtd\Omega} = \frac{dPower}{d\Omega} = (\frac{16\pi G}{8\pi r})^2[\frac{1}{2}\dddot{Q}^{kl}\dddot{Q}^{kl} - \dddot{Q}^{kl}\dddot{Q}^{km}n^l n^m + \frac{1}{4}\dddot{Q}^{kl}\dddot{Q}^{mr}n^k n^l n^m n^r] \tag{35}$$

This angular distribution is quite complicated except in very symmetric cases. However, we can integrate over all solid angles to find the total power radiated

$$-\frac{dE}{dt} = \frac{G}{5c^5}\dddot{Q}^{kl}\dddot{Q}^{kl} \tag{36}$$

Now we have a general set of equations we can use to treat the case of various cosmologically and astrophysically interesting objects. We expect gravitational radiation to fall into one of three classes:

Sources of Gravitational Radiation

Type	Source	Characteristic Frequency
Periodic	orbiting and oscillating objects	$10^{-7} - 10^2$ Hz and $10^2 - 10^5$ Hz
Burst	Collisions, Supernova & infall/collapse	$10^{-1} - 10^4$ Hz
Stochastic	CGR, Inflation, Phase Transitions	all

We treat these in the next subsections.

2.1 Order of Magnitude Estimates

We estimated the size of the dimensionless strain produced by a mass

$$h \sim \frac{R_{schwarzschild}}{R}\frac{v^2}{c^2} \tag{37}$$

Now we can estimate the quadrupole moment and its third derivative to estimate the power loss.

$$\dddot{Q}^{lk} \sim \frac{GMR^2}{t_c^3} \sim \frac{GMv^3}{R} \tag{38}$$

where M, R, t_c, and v are the characteristic mass, size, time scale, and velocity of the source. Using our formula for the power radiated

$$-\frac{dE}{dt} = \frac{G}{5c^5}\dddot{Q}^{kl}\dddot{Q}^{kl} \sim \frac{G}{c^5}(\frac{M}{R})^2 v^6 \sim L_{cGW}(\frac{R_{Schwarzschild}}{R})^2(\frac{v}{c})^6 \tag{39}$$

where

$$L_{cGW} \equiv \frac{c^5}{G} = 3.63 \times 10^{59} \text{ erg s}^{-1} = 2.03 \times 10^5 \text{M}_\odot c^2 \text{s}^{-1} \tag{40}$$

If a steel rod of mass 100 tons (10^8 grams) and length 20 meters rotates at its breakup speed (30 radians/sec), what is the approximate radiated power in gravitational waves?

Show that the power in gravity waves for a system with internal power in quadrupole motion $L_{internal} \sim Mv^2/t_c$ is roughly

$$L_{GW} = \frac{L_{internal}^2}{L_{cGW}} \tag{41}$$

Since astrophysical systems are generally gravitationally bound, the virial theorem tells us

$$\text{Kinetic Energy} \sim \frac{MR^2}{t_c^2} \sim |\text{Potential Energy}| \sim \frac{GM^2}{R}. \tag{42}$$

So the characteristic time scale is

$$t_c \sim \sqrt{\frac{R^3}{GM}} \tag{43}$$

so that

$$L_{GW} \sim L_{cGW}(\frac{R_{Schwarzschild}}{R})^5 \tag{44}$$

and the gravitational wave energy radiated by a nonspherical self-gravitating system in a characteristic time is roughly

$$\Delta E \sim L_{GW} t_c \sim Mc^2(\frac{R_{Schwarzschild}}{R})^{7/2}. \tag{45}$$

Thus the gravitational wave emission efficiency is then

$$\Delta E \equiv \epsilon Mc^2, \quad \epsilon \sim (\frac{R_{Schwarzschild}}{R})^{7/2}. \tag{46}$$

Now we can estimate the order of magnitude for dimensionless strain of astrophysical objects at distance r

$$h \sim \epsilon^{2/7}\frac{R_{schwarzschild}}{r} \sim 3 \times 10^{-18}(\frac{\epsilon}{0.1})^{2/7}\frac{(M/\text{M}_\odot)}{(r/10kpc)}. \tag{47}$$

(10 kpc is roughly the distance to our Galactic center.)

3 Periodic Gravitational Waves

We consider here gravitational waves produced by sources whose density changes periodically.

3.1 Emission by a Vibrating Quadrupole

According to the equations above any system of masses with a time-dependent quadrupole moment whose third derivative $\dddot{Q}^{kl} \neq 0$ is not zero will radiate power in gravitational waves. Consider two equal masses connected by a spring - a linear oscillating quadrupole. This is a simplification of problems of interest but will give a scale for estimating situations of interest. We can use the result that a system of spherical masses has the same quadrupole moment as a system of point masses located at the centers of the spherical masses. We align the masses on the z-axis a distance b from the origin with oscillation amplitude a so that their positions are:

$$z = \pm(b + asin\omega t), \quad \text{and} \quad z^2 = b^2 + 2absin\omega t + a^2 sin^2\omega t \tag{48}$$

One can easily show that the quadrupole moment is

$$Q^{kl} = \frac{1}{3}\begin{bmatrix} -2mz^2 & 0 & 0 \\ 0 & -2mz^2 & 0 \\ 0 & 0 & 4mz^2 \end{bmatrix} = \frac{1}{3}(1+2\frac{a}{b}sin\omega t+(\frac{a}{b})^2 sin^2\omega t)\begin{bmatrix} -2mb^2 & 0 & 0 \\ 0 & -2mb^2 & 0 \\ 0 & 0 & 4mb^2 \end{bmatrix} \tag{49}$$

Thus to first order in a/b (so that we only get radiation at frequency ω) this gives

$$\phi^{kl} \cong \frac{16\pi G}{8\pi r}\frac{2a}{3b}\omega^2 sin\omega(t-r)Q(a=0)^{kl} \tag{50}$$

One can use the formula above to work out the angular distribution which comes to a simple pattern.

$$-\frac{d^2E}{dtd\Omega} = (\frac{16\pi G}{8\pi})^2[2mab\omega^3 cos\omega(t-r)]^2 \sin^4\theta \tag{51}$$

where θ is the usual polar angle. Note that this is similar to EM dipole radiation which depends on $\sin^2\theta$. The total emitted power is

$$-\frac{dE}{dt} = \frac{G}{5c^5}\dddot{Q}^{kl}\dddot{Q}^{kl} = \frac{32G}{15c^5}[mab\omega^3 \cos\omega(t-r)]^2 \tag{52}$$

and averaging over a cycle yields

$$-\frac{dE}{dt} = \frac{G}{5c^5}\dddot{Q}^{kl}\dddot{Q}^{kl} = \frac{16G}{15c^5}[mab]^2\omega^6. \tag{53}$$

The damping rate γ_{GW} of the oscillator due to gravity waves is defined as

$$\gamma_{GW} \equiv -\frac{1}{E}\frac{dE}{dt}. \tag{54}$$

Since each mass has kinetic energy of $\frac{1}{2}m\omega^2 a^2$, the damping rate for this oscillator is

$$\gamma_{GW} \equiv -\frac{1}{E}\frac{dE}{dt} = \frac{32G}{15c^5}mb^2\omega^4 \tag{55}$$

and the time for the energy to be damped to $1/e$ is $1/\gamma_{GW}$. Now this is a simple case but it allows us to make order of magnitude estimates for some more realistic astrophysical systems.

3.1.1 Oscillating Neutron Star

A neutron star could be formed asymmetrically or it might be impacted by infalling material such as an asteroid. Suppose that the neutron star has a mass of $m_{ns} = 0.7\mathrm{M}_\odot$, and radius $R_{ns} = 10$ km and that it is initially deformed on one axis so that $\delta R/R \sim 0.1$. It will oscillate on that axis as a nearly incompressible fluid. These parameters give the neutron star an approximate density of $\rho \sim 10^{14}$ gm/cm^{-3}. (Note, I have not checked to see that these are fully consistent with our formula - the Oppenheimer-Volkoff equation or the stiffer equation of state. If I or one of you have time, then we can fix it in the solution hand out.)

How much power is radiated in gravitational waves and what is the damping time? Is this likely to be the most efficient damping mechanism?

Hint: It is easy to show that the oscillation frequency is roughly $\omega \sim \sqrt{G\rho} \sim 3 \times 10^3$ Hz for an incompressible fluid. Use the relation that $v = \sqrt{\partial P/\partial\rho} \sim \sqrt{P/\rho}$ and $P = Gm^2/R^4$ and that $\omega \approx v/R$. Define the vibration energy as

$$E_{vib} = \frac{1}{2}m_{ns}(\delta R)^2\omega^2 \sim (\frac{\delta R}{R})^2 \times 10^{53} \text{ ergs} \tag{56}$$

3.1.2 White Dwarf, Sun, & Earth Oscillations

Gravitational radiation emitted by quadrupole vibrations might be important in the case of novas. Nova outbursts occur on white-dwarf stars in binary systems, where accretion from a companion star gradually accumulates nuclear fuel on the white dwarf surface until the fuel reaches a critical mass and explodes. Among other things, the explosion initiates vibrations in the body of the white dwarf, at characteristic frequencies of 0.01 to 1 Hz. The energy released in a nova explosion is typically 10^{45} ergs, of which as much as 10% goes into the vibrational motion of the star and is subsequently radiated in the form of gravitational waves.

What if the neutron star considered above was a white dwarf instead? How much power would be radiated from a 10% deformation and what would be the

characteristic decay time through the emission of gravitational radiation? Does this follow the order of magnitude scaling laws from earlier and the numbers in the paragraph above? Are gravity waves likely to be important to the dynamics of solar oscillations and earth oscillations? Make a table of system (neutron star, white dwarf, sun, earth) of the form

Table 1
Vibration GW Damping

Object	Mass	R	R_{BH}/R	GW Power	Damping Time
Neutron Star	0.7 $M_\odot$	10 km			
White Dwarf	0.7 $M_\odot$	10^4 km			
Sun ($M_\odot$, $R_\odot$)	1.99×10^{33} g	6.96×10^{10} cm	4.48×10^{-6}		
Earth	5.98×10^{27} g	6.38×10^{8} cm	1.4×10^{-8}		

Note that one might consider using the excitation of vibration modes in the earth as a gravitational wave detector. The confusion background comes from earthquakes, volcanos and mankind.

3.2 Emission by a Rotating Quadrupole

Another periodic source of gravitational waves, which occurs very commonly in astrophysics, is the rotating quadrupole consisting of two masses in orbits around their common center of mass.

As the damping effects of gravitational waves circularizes orbits and it is simpler to calculate, we will consider circular orbits. Since the quadrupole moment repeats when the masses interchange sides in the orbit, the frequency of the emitted gravitational waves is twice the orbital frequency. The angular distribution will be relatively complicated and we skip over it but note that in the direction perpendicular to the orbital plane the gravitational radiation will be circularly polarized while in the plane it will be linearly polarized as in the case of the vibrating quadrupole.

The radiated power of the system is

$$-\frac{dE}{dt} = \frac{32G}{5c^5}\left[\frac{m_1 m_2}{m_1+m_2}\right]^2 R^4\omega^6 = \frac{32G}{5c^5}\mu^2 R^4\omega^6 \tag{57}$$

where R is the separation between the masses m_1 and m_2 and ω is the orbital frequency.

For an astrophysical system, such as a binary star system with circular orbits, gravity holds the masses together and Kepler's third law gives a relationship between the separation R and the orbital period and thus angular frequency:

$$\omega^2 = \frac{G(m_1+m_2)}{R^3} \quad \text{or} \quad P = 0.545\left(\frac{R}{1000\text{ km}}\right)^{3/2}\left(\frac{M_\odot}{m_1+m_2}\right)^{1/2}\text{ sec} \tag{58}$$

so that the radiated power is

$$-\frac{dE}{dt} = \frac{32G^4}{5c^5R^5}(m_1m_2)^2(m_1+m_2). \tag{59}$$

As the binary loses energy by gravitational waves, the distance between the masses decreases at a rate given by

$$\frac{dR}{dt} = -\frac{64G^3}{5c^5R^3}m_1m_2(m_1+m_2) \tag{60}$$

at the same time the orbital frequency increases at a rate given by

$$\frac{d\omega}{dt} = -\frac{3\omega}{2R}\frac{dR}{dt} = \frac{96}{5}[\frac{G(m_1+m_2)}{c^2R^3}]^{3/2}\frac{Gm_1m_2}{c^2R} = \frac{96}{5}\frac{G}{c^5}\omega^3\frac{Gm_1m_2}{R} \tag{61}$$

where we have used the relationship that

$$E = -\frac{Gm_1m_2}{2R} \;\rightarrow\; \frac{dE}{dt} = \frac{Gm_1m_2}{2R^2}\frac{dR}{dt} \tag{62}$$

One can rearrange this to find the change in period $P_b = 2\pi/\omega$ (frequency) of the binary pair.

$$\frac{1}{P_b}\frac{dP_b}{dt} = \frac{3}{2}\frac{1}{R}\frac{dR}{dt} = -\frac{3}{2}\frac{1}{E}\frac{dE}{dt} = -\frac{96}{5}\frac{G^3(m_1+m_2)^2\mu}{c^5R^4} \tag{63}$$

$$= \frac{1}{2.61\times10^7\mathrm{yr}}(\frac{m_1+m_2}{\mathrm{M}_\odot})^{2/3}\frac{\mu}{\mathrm{M}_\odot}(\frac{1\mathrm{hr}}{P_b})^{8/3} \tag{64}$$

Assuming that the expressions above remain valid as $R\rightarrow0$, show that the time $\Delta t_{R\rightarrow0}$ until $R\rightarrow0$

$$\Delta t_{R\rightarrow0} = \frac{5}{256}\frac{c^5}{G^3}\frac{R_{now}^4}{(m_1+m_2)m_1m_2} \cong 1.84(\frac{R}{100\ \mathrm{km}})^4(\frac{m_1+m_2}{\mathrm{M}_\odot})^{-2}(\frac{\mathrm{M}_\odot}{\mu}) \tag{65}$$

and that

$$\frac{\Delta t_{R\rightarrow0}}{P_b} \sim 10^5(\frac{P}{1s})^{5/3} \quad \text{or} \quad \frac{1}{P_b}\frac{dP_b}{dt} \sim (10^5(\frac{P}{1s})^{5/3})^{-1} \tag{66}$$

Now we are in a position to calculate the gravitational radiation power, orbital radius change, and relaxation time for some astrophysical systems.

3.2.1 Jupiter Power

As by far the largest component of the solar system quadrupole, estimate the power emitted by Jupiter in gravitational radiation and the rate of change of its orbit (in cm/year). (Useful numbers for this problem: orbit radius 5.2 AU = 7.8×10^8 km, period of orbit 11.9 years implying $\omega = 1.68\times10^{-8}$ sec^{-1}, mass = 1.9×10^{30} g $\approx 10^{-3}\mathrm{M}_\odot$.

While you are at it, you might want to calculate the rate of change of the earth's orbital radius due to gravity wave emission and compare it to that of other effects you calculated in problem set 5.

3.3 SuperMassive Black Holes

Our current view is that in the center of most galaxies there are supermassive ($10^5 M_\odot$ to $10^8 M_\odot$) black holes. Occassionaly, smaller black holes or supermassive blach holes from the merging of galaxies find themselves in (co-)orbit around the central supermassive black hole. This happens at first because the black holes are so massive and thus sink towards the central of the potential. Later the gravitational radiation they emit causes them to sprial into closer and closer orbits until they eventually merge. One can show that the quadrupole radiation from two inspiraling black holes has the strain of

$$h_\times = 3.6 \times 10^{-22} \frac{1Gpc}{r} \left(\frac{M_1}{10M_\odot}\right) \left(\frac{M_2}{10^6 M_\odot}\right)^{2/3} \left(\frac{f}{0.01Hz}\right)^{2/3}$$

The frequency determines the orbit separation. The time at frequency f until inspiral is

$$T = 1.41 \times 10^6 sec \left(\frac{0.01Hz}{f}\right)^{8/3} \left(\frac{10M_\odot}{M_1}\right) \left(\frac{10^6 M_\odot}{M_2}\right)^{2/3}$$

$$h_{pred} = 3.4\times 10^{-22} \left(\frac{f}{10^{-2}Hz}\right)^{2/3} \left(\frac{M_{BH}}{10^4 M_\odot}\right)^{5/3} \frac{x}{(1+x)^{1/3}} \frac{1}{1-\sqrt{1+z}} \qquad f \sim 10^4 \left(\frac{M_\odot}{M_{BH}}\right) \frac{Hz}{1+z}$$

What is the strain expected at 3 Gpc from two supermassive black holes each with mass about $10^5 M_\odot$ with a frequency $f = 0.01Hz$? Roughly how long before they inspiral to a frequency of $f = 0.03Hz$? Make a rough plot of strain versus frequency with a few characteristic times called out -e.g. 1 year, 1 month, 1 week, 1 day, etc. Use the characteristic times for the proposed LISA (laser interferometer space array) of arond $10^{-4}\ Hz \leq f \leq 10^{-2}\ Hz$.

3.3.1 Rapidly Rotating Deformed Neutron Star

For a slightly deformed, homogeneous neutron star with moment of inertia $I = 2MR^2/5$, mass M, radius R, rotation period P_{ns}, and ellipticity, ϵ, the gravitational wave luminosity is

$$L_{GW} = -\frac{dE}{dt} = \frac{32G}{5c^5} I^2 \epsilon^2 (\frac{2\pi}{P})^6 \tag{67}$$

$$\approx 10^{38}\text{erg s}^{-1} (\frac{I}{4\times 10^{44}\text{g cm}^2})^2 (\frac{P_{ns}}{0.033\text{sec}})^{-6} (\frac{\epsilon}{10^{-3}})^2 \tag{68}$$

and at a time t after its birth since it will slow down

$$L_{GW} \approx 10^{45}\text{erg s}^{-1} (\frac{I}{4\times 10^{44}})^{1/2} (\frac{10^{-3}}{\epsilon}) (\frac{10^6\text{sec}}{t+10^4\text{sec}})^{3/2} \tag{69}$$

3.3.2 Close Binary Stars

Table 2
Interesting Binary Sources of GW

Object	Masses $M_\odot$	Distance (pc)	Frequency 10^{-6} Hz	GW Power ergs/s	Damping Time years	Strain at Earth (10^{-22}
Eclipsing Binaries						
ι Boo	1.0, 0.5	11.7	86			51
μ Sco	12, 12	109	16			210
V Pup	16.5, 9.7	520	16			46
Cataclysmic Binaries						
Am CVn	1.0, 0.041	100	1900			5
WZ Sge	1.5, 0.12	75	410			8
SS Cyg	0.97, 0.83	30	84			30
Binary X-ray Sources and/or Pulsars						
Cyg X-1	30, 6	2500	4.1			4
PSR1913+16	1.4, 1.4	5000	70*			0.12
PSR1534+12	1.4, 1.4	0.5	54*			
PSR2127+11C	1.4, 1.3	10.6	70*			

* Pulsars with elliptical orbits emit significantly at harmonics.
The next section will deal with these three binary pulsars.

There is a whole class of interesting binary star pairs that can be calculated approximately. The most spectacular binary star pair known at the moment is AM CVn (AM Canem Venticorum). Am CVn consists of a blue white dwarf and a low-mass white dwarf in an exceptionally small orbit around each other. Their period is 17.5 minutes! Do you know which star has the bigger radius? (Table 2 tells us that their masses are about 1 and 0.041 $M_\odot$, respectively.) What do you estimate for the ratio of the radii of the lower to higher mass star? Fill out the following table. (See formula in next section to help with the numerical factors.)

3.3.3 Binary Pulsars

There are three known binary pulsars and these and some of their parameters are listed in table 3.

Table 3
Galactic Binary Pulsars

Parameter	PSR 1534+12	PSR 1913+16	PSR 2127+11C
Reference	Wolszczan et al. 1991	Hulse & Taylor 1974	Prince et al. 1991
Flux (400 MHz)	36 mJy	5 mJy	0.6 mJy
Distance	1.1 ± 0.2 kpc	7.3 kpc	10.6 kpc (in M15)
Period P	10.098 hr	7.7519 hr	8.047 hr
eccentricity	0.273677	0.61713	0.68141
$M_1 + M_2$	2.6784 $M_\odot$	2.82837 $M_\odot$	2.71 $M_\odot$
$q = M_1(pulsar)/M_2$	0.97 ± 0.03	1.04	1.0 ± 0.2
$P/(2\dot{P})$	2.5×10^8 yr	1.1×10^8 yr	1.0×10^8 yr
$t_{merge-theory}$	27.3×10^8 yr	3.01×10^8 yr	2.20×10^8 yr
t_{meas}	30×10^8 yr	4.1×10^8 yr	3.20×10^8 yr

Table adapted from Phinney 1991 Ap J 380 L17.

As one can see the eccentricity is significant for these systems. This leads to significant emission at the orbital frequency (as opposed to twice) and in harmonics. The modified energy loss equation according to Peters and Mathews is

$$-\frac{dE}{dt} = L_{GW} = \frac{32G^4}{5c^5a^5}\mu^2(m_1+m_2)^3\ f(\epsilon) \tag{70}$$

$$\approx 3.0 \times 10^{33}\text{erg s}^{-1}(\frac{\mu}{\text{M}_\odot})^2(\frac{m_1+m_2}{\text{M}_\odot})^{4/3}(\frac{P_b}{1\text{hour}})^{-10/3}\ f(\epsilon) \tag{71}$$

where

$$f(\epsilon) = \frac{1 + \frac{73}{24}\epsilon^2 + \frac{37}{96}\epsilon^4}{(1-\epsilon^2)^{7/2}} \tag{72}$$

and a is the semi-major axis.

If the two masses are in an elliptical orbit with eccentricity ϵ then the energy emission in gravity waves rate (to a sign) is

$$L_{GW} = \frac{dE}{dt} = \frac{dE}{dt}|_{\epsilon=0}\ f(\epsilon) \tag{73}$$

and the angular momentum loss rate is

$$\frac{dJ}{dt} = \frac{dJ}{dt}|_{\epsilon=0}\ g(\epsilon) = -\frac{32}{5}\frac{G^{7/2}\mu^2(m_1+m_2)^{3/2}}{a^{7/2}}g(\epsilon) \tag{74}$$

where

$$g(\epsilon) = \frac{1 + \frac{7}{8}\epsilon^2}{(1-\epsilon^2)^2} \tag{75}$$

where these are averaged over the orbit.

Using Kepler's laws we find

$$\frac{1}{P_b}\frac{dP_b}{dt} = \frac{1}{P_b}\frac{dP_b}{dt}|_{\epsilon=0}\ f(\epsilon) = -\frac{96}{5}\frac{G^3(m_1+m_2)\mu}{c^5 a^4}\ f(\epsilon) \tag{76}$$

For an elliptic orbit

$$\epsilon^2 = 1 + \frac{2EJ^2}{G^2\mu^3(m_1+m_2)^2} \tag{77}$$

One can use this relationship to derive an equation for $d\epsilon/dt$ and find that it is negative, so that gravitational radiation tends to circularize an elliptical orbit.

$$\frac{d\epsilon}{dt} = -\frac{304}{15}\frac{G^3\mu\ (m_1+m_2)^2\epsilon}{a^4\ (1-\epsilon^2)^{5/2}}\left(1+\frac{121}{304}\epsilon^2\right) \tag{78}$$

and the semi-major axis a decreases at the rate

$$\frac{da}{dt} = \frac{dR}{dt}|_{\epsilon=0} f(\epsilon) = -\frac{64}{5}\frac{G^3\mu\ (m_1+m_2)^2}{a^3} f(\epsilon) \tag{79}$$

In the Newtonian regime, if we orient the polarization axes along the major and minor axes of the projection of the orbital plane, then the dimensionless strain is

$$h_+ = 2(1+cos^2 i)\ \frac{\mu}{R}[\pi(m_1+m_2)f]^{2/3} cos(2\pi \int \frac{df}{dt}dt\ t) \tag{80}$$

$$h_\times = \pm 4cos^2 i\ \frac{\mu}{R}[\pi(m_1+m_2)f]^{2/3} sin(2\pi \int \frac{df}{dt}dt\ t) \tag{81}$$

3.4 Death Spiral

As the orbital pair loses energy through emitting gravitational waves the orbit becomes more circular, the major axis decreases in size, and the orbital frequency increases. The binary pair will begin a gradual spiral towards each other. The inward motion is slow at first but increases rapidly as the orbit gets smaller. AS the distance gets fairly small the rate gets large and results in a crescendo of gravitational radiation and a death spiral that ends in their coalescence. During the death spiral phase the gravitational radiation frequency will move higher and higher - rapidly in an up glissando. This is a reasonable term since one can easily calculate that the maximum frequency for a binary neutron star is about 1 kHz. It may be up to about 10 times that for optimal black holes.

Let us consider the steps leading up to the end. We will trust our results to the point that Newtonian gravitation is a reasonable approximation. We can use the circular orbit parameters.

One can integrate the rate of change of the orbit radius to find

$$\frac{1}{4}(R^4 - R_0^4) = -\frac{64G}{5c^5} m_1 m_2 (m_1 + m_2)(t - t_0) \tag{82}$$

where the t_0 indicates the starting time of the calculation.

Earlier we derived a formula for the time remaining $\Delta t_{R\to 0}$ until the radius would go to zero, if the equations remained valid. The time remaining went as

$$\Delta t_{R\to 0} = \frac{5}{256}\frac{c^5}{G^3}\frac{R_{now}^4}{(m_1+m_2)^2\mu} = [1 - (\frac{r}{r_0})^4]\frac{r_0}{4}\frac{1}{[-dR/dt|_{t_0}} \tag{83}$$

Thus it is clear that the system spends most of its time at the large distance and much less at small separations.

It is easy to show that the orbital frequency is

$$f_{orbit} = \frac{\omega}{2\pi} = \frac{1}{P_b} = \frac{1}{2\pi}[\frac{5}{256}\frac{c^5}{G^{5/3}\mu(m_1+m_2)^{2/3}}\frac{1}{t_{R\to 0} - t}]^{3/8} \tag{84}$$

where $t_{R\to 0} = t_0 + \Delta t_{R\to 0}$. The gravitational wave frequency will be twice f_{orbit}.

Consider two neutron stars with masses 1.4 $M_\odot$orbiting at 100 km. The intial radiated power is about 6×10^{51} erg/s. The rate of decrease of the orbit radius is $-dR/dt|_{t_0} \sim 7 \times 10^6$cm/s = 70km/s. Clearly to get from 100 km to 20 km is going to take on the order of 1 second. The formula gives 0.36 seconds. Starting from 130 km separation would take about 1 second and 200 km would take about 6 seconds.

At 200 km the orbital frequency is about 50 Hz and the orbital velocity is about 0.1c. By 130 km these are up to 93 Hz and 0.13c and at 100 km this becomes 138 Hz and 0.289c. At 20 km the frequency is 1550 Hz and the orbital speed is 0.32c. Clearly a careful calculation will have to take into account special relativistic effects but our numbers should be accurate to better than 1% at 200 km and better than 10%. We would also have to take into account the spin-orbit coupling due to the frame-dragging of the rotating stars and orbit. This mostly causes a beat in the emitted radiation.

We can estimate the power radiated in gravitational waves in those last few moments.

$$\Delta E = \frac{G}{2} m_1 m_2 [\frac{1}{r_0} - \frac{1}{r}] \tag{85}$$

where the factor of two comes from the gravitational energy splitting between gravitational wave and kinetic energy. With $r_0 = 200$ km and $r = 20$ km one has

$$-\Delta E \simeq 10^{53} \text{ erg} \approx 0.04 M_\odot \tag{86}$$

which means that about 1.5% of the rest mass was radiated. Most of this energy is released in the last states of the coalescence. An extra pulse will be emitted when the neutrons stars collide and the resulting neutron star or black hole

oscillations/vibrations is likely to produce even more. We consider this in the next section.

In the May 1995 Scientific American article by Tvsi Piran that I handed out in class, the author claims that two neutron stars 700 km apart take 15 minutes before collision and and their orbital period shrinks from a fifth of a second to a few milliseconds. Check these numbers and see what mass system one needs for this to be true and are these numbers consistent with the quoted parameters for the three know binary pulsars?

4 Gravitational Wave Bursts

Now we move from periodic (or quasiperiodic) case to the burst case. Any system with a quadrupole whose third derivative is non-zero will radiate.

4.1 Little hits very Big

For illustration we will first consider a small particle falling towards a very large mass. This is just so that we can neglect the motion of the large mass and its radiation. That should make the effects clearer to understand. In the next section we treat this in the fashion of reduced mass. The energy loss can be found by evaluating the quadrupole radiation loss formula and using the quadrupole matrix for a single mass at a distance z on the z axis we found for in the vibrating quadrupole section above.

$$Q^{kl} = \begin{bmatrix} -\frac{1}{3}mz^2 & 0 & 0 \\ 0 & -\frac{1}{3}mz^2 & 0 \\ 0 & 0 & \frac{2}{3}mz^2 \end{bmatrix} \tag{87}$$

Hence,

$$-\frac{dE}{dt} = \frac{G}{5c^5}\dddot{Q}^{kl}\dddot{Q}^{kl} = \frac{2Gm^2}{15c^5}(6\dot{z}\ddot{z} + 2z\dddot{z})^2 \tag{88}$$

Now we need to evaluate the various derivatives of z. Here we consider the case that the small test mass m is falling from rest at a great distance (i.e. infinity) as a result of gravitational attraction to the large stationary mass M. Later you can consider how to treat a more general case of interactions. Conservation of energy gives

$$\frac{1}{2}m\dot{z}^2 = \frac{GmM}{|z|} \tag{89}$$

and Newtonian law of universal interaction and the equivalence principle gives

$$\ddot{z} = \frac{GM}{z^2} \tag{90}$$

Taking the derivative with respect to time

$$\dddot{z} = -\frac{2GM}{z^3}\dot{z} = -\frac{2GM}{|z|^{7/2}}(2GM)^{1/2} \tag{91}$$

Now we can estimate the energy loss by integrating dE along the particle trajectory.

$$\Delta E = -\int_{\infty}^{R} \frac{1}{|z|^{9/2}} 2Gm^2 \; 15c^5 (2GM)^{5/2} dz = -\frac{1}{R^{7/2}} \frac{4Gm^2}{105c^5} (2GM)^{5/2} \tag{92}$$

Clearly the closer the test particle gets the more radiation energy it emits. The question is where to cut off? Clearly, the Schwarzschild radius of the large mass is a hard cut off since none of the gravitational waves would escape.

$$\Delta E_{max} = -\frac{2}{105} mc^2 (\frac{m}{M}) \cong -0.019 mc^2 (\frac{m}{M}) \tag{93}$$

In the next section we will see that this may be as much as roughly a factor of two over more detailed calculations but does neglect the radiation from the oscillations of the system as the masses coalesce.

This is a moderately efficient gravitational wave generator. Consider a one solar mass object falling into a 10 $M_{\odot}$ black hole.

$$\epsilon = \frac{\Delta E}{mc^2} = 0.019 (\frac{m}{M}) = 0.002 \tag{94}$$

In the next section we handle the reduced mass case and get essentially same answer.

4.2 Collision of Two Objects

Consider two point masses m_1 and m_2 falling towards each other under the influence of gravity from rest at infinity. For this case we will assume that the collision is head on and that they move on the x axis. The center of mass is $x = 0$.

$$m_1 x_1 = -m_2 x_2 = \mu x \tag{95}$$

where

$$x \equiv x_1 - x_2, \qquad \mu \equiv \frac{m_1 m_2}{M}, \qquad M = m_1 + m_2 \tag{96}$$

Since

$$m_1 x_1^2 + m_2 x_2^2 = \mu x^2, \tag{97}$$

$$Q_{xx} = \frac{2}{3} \mu x^2, \qquad Q_{yy} = Q_{zz} = -\frac{1}{3} \mu x^2 \tag{98}$$

The equation of motion is

$$\ddot{x} = -\frac{GM}{x^2} \tag{99}$$

and

$$\frac{\dot{x}^2}{2} = \frac{GM}{x}, \tag{100}$$

yielding

$$\ddot{Q}_{xx} = -\frac{4}{3} G\mu M \frac{\dot{x}}{x^2}. \tag{101}$$

Using the formula for radiated power

$$-\frac{dE}{dt} = \frac{G}{5c^5}\left\langle \dddot{Q}_{xx}^2 + \dddot{Q}_{yy}^2 + \dddot{Q}_{zz}^2 \right\rangle = \frac{8}{15}\frac{G}{c^5}\mu^2 M^2 \left\langle \frac{\dot{x}^2}{x^4} \right\rangle. \tag{102}$$

This integral will diverge as $x \to 0$. However, we know that this non-relativistic treatment will fail as we get to the Schwarzschild radius as the gravitational radiation will not escape. If we cut the integral off at $x_{min} = 2GM/c^2$

$$\Delta E = \frac{2}{105}\frac{\mu^2 c^2}{M}. \tag{103}$$

Now we should expect that this is an overestimate as the closer one gets to the Schwarzschild radius the less gravitational wave energy escapes but we also ignored other relativistic effects. The fall of a test particle into Schwarzschild black hole of mass M gives

$$\Delta E = 0.0104\frac{\mu^2 c^2}{M}. \tag{104}$$

This is roughly half of our semi-classical approach. Epply and Smart have calculated the radiation form the head-on collision of two equal-mass Schwarzschild black holes to find

$$\Delta E = 0.001 M c^2 \tag{105}$$

however, there is uncertainty which is at the level of a factor of 2 so it might be in reasonable agreement with the formula above. We can presume that there is more radiation if the collision is not head-on or if there is significant angular momentum in the system.

How would you do this calculation for two particles colliding if in addition to the gravitational attraction, they had significant peculiar velocity? For example, gravitational bremsstrahlung between an electron and proton or two protons in a hot plasma.

The frequency spectrum from a burst of gravitational waves will clearly not be periodic or quasiperiodic. Instead it will be much like EM bremsstrahlung radiation. Modes will be excited from low frequencies right on up to the inverse of the fastest time in collision.

$$\Delta E = \int \frac{dE}{dt} dt = \frac{G}{5c^5}\int \dddot{Q}^{kl}\dddot{Q}^{kl} dt \tag{106}$$

The total energy radiated during the collision is

$$\Delta E = \int \frac{dE}{dt} dt = \frac{G}{5c^5}\int_{-\infty}^{\infty} \dddot{Q}^{kl}\dddot{Q}^{kl} dt \tag{107}$$

$$= \frac{1}{10\pi}\frac{G}{c^5}\int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} d\omega' \tilde{Q}^{kl}(\omega)\tilde{Q}^{kl}(\omega') e^{i(\omega-\omega')t} \tag{108}$$

where $\tilde{Q}^{kl}(\omega)$ is the fourier transform of $\ddot{Q}^{kl}$:

$$\tilde{Q}^{kl}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \ddot{Q}^{kl} e^{i\omega t} dt, \quad \text{and} \quad \ddot{Q}^{kl} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{Q}^{kl}(\omega) e^{-i\omega t} dt, \tag{109}$$

Parseval's theorem gives

$$\frac{dE}{d\omega} = \frac{2}{5} \frac{G}{c^5} |\tilde{Q}^{kl}(\omega)|^2 \tag{110}$$

For this case of two particles coming together under the influence of gravity in the low frequency regime

$$\frac{dE}{d\omega} \rightarrow \frac{2^{11/3}, (\frac{1}{3})^2}{5\pi 3^{7/3}} \frac{G}{c^5} (GM\omega)^{4/3} \mu^2 \quad \text{as} \quad \omega \rightarrow 0 \tag{111}$$

The high frequency cut off corresponds to the time scale at the Schwarzschild horizon. The figure handed out in class shows a representative spectrum of this type.

4.3 GW from Non-Spherical Collapse

One promising source of gravitational radiation is the non-spherical collapse of stellar cores (e.g. supernova). Detailed calculations are clearly very complicated. We consider a highly idealized symmetric case to get an order of magnitude estimate of the effect. Consider the Newtonian collapse of a homogeneous spheroid or ellipsoid.

$$-\frac{dE}{dt} = \frac{2}{375} \frac{GM^2}{c^5} \left\langle (\dddot{a}^2 - \dddot{c}^2)^2 \right\rangle \tag{112}$$

where a and c are the major and minor semi-axes. This collapsing configuration will give out radiation in proportion to the fourth power of the angular momentum.

$$\Delta E \sim (\frac{J}{GM^2/c})^4 Mc^2 \tag{113}$$

so with a lot of angular momentum the efficiency can be significant.

Estimate the efficiency $\epsilon = \Delta E/Mc^2$ in the formation of a remnant with the angular momentum of the Crab pulsar $J \sim 2 \times 10^{47}$ erg s^{-1}and for the pulsar PSR 1937+214 with $J \sim 4 \times 10^{48}$ erg s^{-1}.

5 Chaotic Gravitational Waves

The first set of chaotic gravitational waves that we encountered were the potential Cosmic Gravity Wave Thermal Background which you estimated would have an effective temperature of about 0.9 K. In a moment we will see how this could have come about; however, we should first consider how these might have been erased (stretched to very long wavelengths) and replaced by a new chaotic field by cosmic inflation.

5.1 Gravitons from Inflation

We saw in a previous lecture that inflation, in addition to producing density (scalar) fluctuations, would produce a random, chaotic field of gravitational (tensor) fluctuations. While these are outside the event (particle) horizon they are essentially frozen fluctuations in curvature

$$\delta ln(a) = \frac{\delta\rho}{\rho + P} \tag{114}$$

which Bardeen showed was a gauge invariant quantity. To first order we expect that these curvature fluctuations will be produced with roughly equal amplitude and thus potential fluctuations with dimensionless strain of order

$$(A_{\oplus,\otimes})_k = (h_{+,\times})_k \approx \frac{2}{\sqrt{\pi}} \frac{H_{inflation}}{M_{Planck}} \tag{115}$$

These fluctuations started their existence as part of the universe's wave function and as individual gravitons.

Once the universe is sufficiently old these curvature fluctuations were free to propagate as gravitational waves. That happens when a given graviton mode re-enters the horizon and the tensor metric fluctuations then propagate as gravitons. They also begin to redshift away due to the expansion of the universe.

The energy density of a chaotic field of gravitational waves has an energy density

$$\rho_{graviton} = < \dot{A}_{\oplus}^2 + \dot{A}_{\otimes}^2 > = \frac{\dot{h}_+^2 + \dot{h}_\times^2}{16\pi G} \tag{116}$$

or in power spectrum

$$\rho_{graviton} = \frac{1}{16\pi G} \int [(h_+)_k^2 + (h_\times)_k^2] k dk. \tag{117}$$

This leads us to the estimate of energy density power spectrum

$$k \frac{d\rho_{graviton}}{dk} = \frac{k^2[(h_+)_k^2 + (h_\times)_k^2]}{16\pi G} \tag{118}$$

Thus for the radiation-dominated phase of the universe the change in the energy density at each wavelength is proportional to a^{-2}. .....argument of why as energy per mode/wavelength is originally $\omega^2 A^2$ and it is redshifted away for a longer time inversely proportional to ω, one finds that for modes entering the horizon during radiation dominated phase all have equal energy density per mode (logarithmic interval) in λ .... Thus there is over the range of wavelengths $10^{10} to 10^{26}$ cm an essentially flat energy density or a dimensionless strain falling as $\propto \omega^{-2} \propto \lambda^2$. For longer wavelengths the mode entered during the matter-dominated phase and the energy density rises proportional to λ until one reaches the horizon. thus the

dimensionless strain grows proportional to the wavelength $\propto \omega^{-1} \propto \lambda$. Outside the horizon the dimensionless strain is (constant) equal to the value set by the expansion rate of inflation divided by the Planck mass (to the factor $2/\sqrt{\pi}$. See Figure xx. for an example of the gravitational wave spectrum expected for a sample model of inflation.

We can estimate the present (after expansion) strain of inflation epoch gravitational waves as a function of wavelength

$$h_0 \sim \Omega_\gamma^{1/2} \frac{\lambda_0 H_0}{c} (\frac{H_{inf}}{m_{pl} c^2})^2 \tag{119}$$

Thus a detector operating at a wavelength of $\lambda_0 = 10^3$ km, $H_0 \lambda_0 / c \sim 10^{-20}$. The density in radiation $\Omega_\gamma \sim 10^{-4}$, so $h_0 \sim 10^{-20} (H_{inf}/m_{pl}c^2)^2$.

5.2 Gravitational Waves from Thermal Collisions

We know that gravitational radiation is only emitted when particles actually undergo accelerations. In thermal plasmas, which are common in astrophysics and cosmology, there are thermal collisions - usually coloumb and elastic scattering - frequently. The energy per unit frequency interval emitted as gravitational waves in a collision of particles is

$$\frac{dE}{d\omega} = \frac{G}{2\pi} \sum_{in,out} (-1)^{in/out} m_n m_n \frac{1+\beta_{nm}^2}{(1-\beta_{nm}^2)^{1/2}} ln(\frac{1+\beta_{nm}}{1-\beta_{nm}}) \tag{120}$$

where β_{nm} is the relative velocity between particles n and m. For non-relativistic two-body scattering this reduces to

$$\frac{dE}{d\omega} = \frac{8G}{5\pi} \ \mu^2 \ v^4 \ sin^2\theta \tag{121}$$

where μ is the usual reduced mass, v is the relative velocity, and θ is the scattering angle in the center-of-mass frame.

The gravitational radiation produced by the collisions in a plasma or gas can be obtained by summing the radiated energies per collision, provided the collisions are incoherent (i.e. there is more time between collisions than it takes to radiate so that they do not interfere with each other coherently). This condition is that that we are considering radiation with $\omega_{GW} >> \omega_c$ where ω_c is the collision frequency per particle. In the incoherent regime

$$\frac{dE}{d\omega} = \frac{8G}{5\pi} \ \sum_{a,b} \mu_{ab}^2 n_a n_b \left\langle v_{ab}^2 \int \ \frac{d\sigma_{ab}}{d\Omega} \ sin^2\theta \ d\Omega \right\rangle \tag{122}$$

where n_a and n_b are the number densities of particles of type a and b, $d\sigma_{ab}/d\Omega$ is the center-of-mass-system differential scattering cross-section. The sum runs over all paris of particles and the average is taken over all collisions.

5.2.1 GW from a Plasma

As an example we calculate the gravitational radiation emitted by Coulomb scattering in a plasma. The Rutherford scattering we can retcross-section is

$$\frac{d\sigma_{ab}}{d\Omega} = \frac{q_a^2 q_b^2}{4v_{ab}^4 \mu_{ab}^2 sin^4(\theta/2)} \tag{123}$$

The integral over θ must be cut off at a minimum angle determined by the Debye screening of the Coulomb force at large impact parameters;

$$\int \frac{d\sigma_{ab}}{d\Omega}\ sin^2\theta\ d\Omega = \frac{4\pi q_a^2 q_b^2 ln(1/\theta_{min})}{v_{ab}^4 \mu_{ab}^2} \tag{124}$$

To average over collisions we must find $< v_{ab} >$, as one power is left. For a Maxwell-Boltzmann distribution

$$\langle v_{ab} \rangle = 2(\frac{2kT}{\pi \mu_{ab}})^{1/2} \tag{125}$$

Typically $ln(1/\theta_{min}) \sim 10$. For a plasma of completely ionized hydrogen we must take into account electron-proton and electron-electron collisions. (Why don't we worry about proton-proton and electron-helium plus proton-helium collisions?)

$$\frac{dP}{d\omega} = \frac{64 G n_e^2 e^4}{5c^5}\ (\frac{2kT}{\pi \mu_{ab}})^{1/2}\ (1+\sqrt{2})\ ln(1/\theta_{min}) \tag{126}$$

The electron collision frequency can be estimated as

$$\omega_c \approx \frac{e^4 n_e < v >}{(kT)^2} \approx \frac{e^4 n_e}{(kT)^{3/2} m_e^{1/2}}. \tag{127}$$

5.2.2 GW from the Sun & Stars

Now let us apply this to the hydrogen plasma in the solar or a stellar core. Within a volume of roughly 2×10^{31} cc this plasma has $T \simeq 10^7$ K, $n_e \simeq 3 \times 10^{25}$ cc, and $ln(1/\theta_{min}) \sim 4$. The collision frequency is roughly 10^{15} s^{-1}, which is three orders of magnitude less than the thermal frequency $\omega_{thermal} = kT/\hbar \approx 10^{18}$ s^{-1}, so that an estimate of the total power produced in gravitational radiation can be found by multiplying by $VkT/\hbar$. This gives about 10^8 watts or 100 megawatts.

5.2.3 GW from the Early Universe Plasma

We could do this same calculation for the early universe. What would be power in gravitational radiation produced by thermal collisions of the plasma at the time of nucleosynthesis? (Take a time of 1 sec and $kT = 1$ MeV.) We neglected photons in the plasma calculation above, should we take them into account here? How far back do we have to go in the universe before there is good thermal interaction between the gravitons and other constituents of the universe?

5.2.4 GW from the ElectroWeak Phase Transition

Extra gravitational radiation could be produced during the ElectroWeak Phase Transition (as well as other phase transitions) over the simple plasma conditions due to the traveling domain wall of the phase transition as well as their collisions with each other. ...

5.2.5 GW from Cosmic Strings

Due to the high tension equal to the mass per unit length μ in cosmic strings, they both try to straighten out at the speed of light as well as have oscillations and traveling waves. These lead to the copious production of gravity waves by strings.

We can make a rough order of magnitude estimate of the gravitational radiation from cosmic strings. Most radiation will come from oscillating loops which are formed when a string crosses itself. A loop naturally forms. Loops surviving to epoch given by time t in the radiation era will have a typical size $l \sim G\mu t$ and will produce gravitational waves of frequency $\omega \sim 1/l$ and energy density

$$\rho_{GW-s} = (\dot{M}t)n(l) \simeq \ \mu \ ln(l) \tag{128}$$

where $n(l)$ is the density of loops larger than l. During the radiation dominated era

$$n(l) \sim t^{-3/2} l^{-3/2} \tag{129}$$

and thus

$$\rho_{GW-s} \simeq \frac{(G\mu)^{1/2}}{Gt^2} \tag{130}$$

Since the energy density in relativistic radiations scales as $a^{-4} \propto t^{-2}$ and is essentially the critical density during the radiation dominated error, we have $\rho_\gamma \approx 1/(30Gt^2)$,

$$\Omega_{GW-s} \simeq 30(G\mu)^{1/2}\Omega_\gamma = 6 \times 10^{-7} h^{-2} (G\mu/10^{-6})^{1/2} \tag{131}$$

A more precise calculation (Brandenberger et al. 1986) gives a slightly lower estimate

$$\Omega_{GW-s} = 4 \times 10^{-8} h^{-2} \Omega_0^{-1} (G\mu/10^{-6})^{1/2} \tag{132}$$

Calculate the limit that is set by precision pulsar timing, on the energy density of gravity waves with frequencies of 1/4 yr^{-1}. Five kpc away is PSR 1937+21 discovered by Don Backer which has had a steady 1.6 ms pulse for four years. The time between the main pulse and an interpulse is given as 744.9 plus or minus 1.3 microseconds over those four years. It is part of a group of pulsars whose relative timing is good to about 3 microseconds. Does this come close to the expected signal from GUT cosmic strings? Do more years of observation help?

Hint: Calculate the strain needed to change the timing by 1 microsecond (distance = 5 kpc). Next calculate the energy density for gravitational waves with wave length of 4 light years and compare that to ρ_c to get Ω_{GW}.

6 Anisotropies from Gravitational Waves

What is the typical CBR anisotropy expected from a long wavelength gravity wave in which we are immersed? (Calculate a numerical value that sets the scale for the mean expected rms fluctuation if the energy scale of inflation at current horizon crossing is the GUT scale, first for the quadrupolar term due to the gravity wave in which we currently find ourselves and then for the chaotic field filling the universe. What would we expect for the apparent random velocity dispersion of clusters of galaxies relative to each other due to this chaotic field of gravity waves?

We expect the frequency shift for entering photons to be given the difference in the dimensionless strain at the point of reception and the point of emission.

$$\frac{\Delta T}{T} = z(\theta,\phi) = \frac{1}{2}(h_r - h_e)\ (1 - cos\theta)\ cos2\phi \tag{133}$$

where h_r and h_e are the dimensionless strain at the receiver and emitter respectively and theta is the angle between the line of sight and the direction of propagation of the plane wave $(\hat{\mathbf{n}} \cdot \vec{\mathbf{k}})$ and ϕ is the azimuthal angle around the direction of motion relative to the polarization of the wave.

For a random chaotic field we can estimate the resulting mean square amplitude of temperature or velocity variations as

$$\left\langle z^2 \right\rangle = \left\langle \frac{\Delta T}{T} \right\rangle = \frac{1}{3\pi^3}\ \rho_{GW}\ \omega^{-2} \tag{134}$$

where 5/8 of this is produced by the quadrupole.

Consider the random chaotic field produced by inflation which will have $< A >= 0$ but will have an rms value

$$\left\langle (\frac{\Delta T}{T})^2 \right\rangle = \frac{1}{4}\left\langle \omega h H_0^{-1} \right\rangle^2 \simeq \frac{\omega^2 h^2}{4H_0^2} \approx \frac{8\pi G \rho_{GW}}{H_0^2} \simeq 3\Omega_{GW} \tag{135}$$

THE GENERATION OF GRAVITATIONAL WAVES:

A REVIEW OF COMPUTATIONAL TECNIQUES *†

Kip S. Thorne

California Institute of Technology

Pasadena, California

TABLE OF CONTENTS

1. Introduction
   1.1. Nature of these lectures
   1.2. Regions of spacetime around a source
   1.3. The gravitational-wave field
   1.4. Formalisms for calculating wave generation

*Sections 1-6 of these lectures were presented at the International School of Cosmology and Gravitation, Erice, Sicily, 13-25 March 1975. Section 7, which is included here for completeness, was presented at the Second Latin-American Symposium on Relativity and Gravitation, Caracas, Venezuela, December 8-13, 1975. These are the lecture notes passed out at those conferences, except for a modest (and inadequate!) amount of updating as of January 1977.

†Supported in part by the National Aeronautics and Space Administration (NGR 05-002-256) and the National Science Foundation (AST 75-01398 A01). The updating and finalizing of the manuscript was carried out at Cornell University with partial support from the National Science Foundation (AST 75-21153).

2. Weak-Field Formalisms
2.1. Classification of weak-field formalisms
2.2. Special Relativity: A formalism of order (1,0)
2.3. Linearized Theory: A formalism of order (1,1)
2.3.1. A valid application of Linearized Theory
2.3.2. Invalid applications of Linearized Theory
2.4. Post-Linear Theory: A formalism of order (2,1)
2.5. Post-Linear Wave Generation: A formalism of order (2,2)
2.5.1. Sample applications
2.6. The slow-motion limit of the Post-Linear Formalism: Newtonian Theory [Order (2,1)] and Quadrupole-Moment Formalism [Order (2,2)]
2.6.1. Sample applications
2.7. The Post-Newtonian Wave-Generation Formalism [Order (3,3)]
2.7.1. Sample applications

3. Multipole-Moment Analysis of the Radiation Field
3.1. Gravitational-wave field
3.2. Symmetric,trace-free tensors and spherical harmonics
3.3. Multipole expansion of gravitational-wave field
3.4. Multipole expansion of energy, momentum, and angular momentum in waves
3.5. Discussion of the multipole formalism

4. Slow-Motion Formalisms
4.1. Overview
4.2. Multipole moments of a stationary source
4.2.1. ACMC-N coordinate system
4.2.2. Multipole moments
4.2.3. Example: the Kerr metric
4.3. The slow-motion, multipole-moment formalism for strong field sources
4.3.1. A sample application
4.4. Weak-field limit of the slow-motion formalism
4.5. Post-Newtonian multipole formalism

5. Perturbation Formalisms

6. Wave Generation by Sources with Strong Internal Fields and Fast, Large-Amplitude Motions

7. Propagation of Waves to Earth
7.1. The geometric-optics formalism
7.2. Example of propagation

8. Concluding Remarks

1. INTRODUCTION

1.1 Nature of these lectures

In these lectures I shall review techniques for calculating gravitational-wave generation. My emphasis will be on the techniques themselves, on their realms of validity, and on typical applications of them. Most derivations will be omitted or sketched only briefly, but I shall give references to places where the derivations can be found.

I shall presume that the reader is familiar with general relativity at the level of "track one" of Misner, Thorne, and Wheeler (1973) -- cited henceforth as "MTW." My sign conventions and notation will be the same as in MTW.

1.2 Regions of Spacetime Around a Source

One can characterize a source of gravitational waves, semi-quantitatively, by the following length scales:

$$L \equiv \text{"Size of source"} \equiv \left\{ \begin{array}{l} \text{radius of region inside which} \\ \text{the stress-energy } T^{\alpha\beta} \text{ is contained} \end{array} \right\},$$

$$2M \equiv \text{"Gravitational radius of source"} \equiv \left\{ \begin{array}{l} \text{2 x mass of source in} \\ \text{units where } G = c = 1 \end{array} \right\},$$

$$\lambda\!\!\!^{-} \equiv \text{"reduced wavelength of waves"} \equiv \left\{ \begin{array}{l} 1/2\pi \text{ x characteristic wave-} \\ \text{length of gravitational} \\ \text{waves emitted by source} \end{array} \right\}, \quad (1.2.1)$$

$$\left. \begin{array}{l} r_I \equiv \text{"inner radius of local wave zone"} \\ r_0 \equiv \text{"outer radius of local wave zone"} \end{array} \right\} \text{ (see below).}$$

Corresponding to these length scales, one can divide space around a source into the following regions (See Fig. 1):

Source:	$r \equiv$ radius $\lesssim L$
Strong-field region:	$r \lesssim 10\,M$ if $10\,M \gtrsim L$ typically does not exist if $L \gg 10\,M$
Weak-field near zone:	$L < r$, $10\,M \ll r$, $r \ll \lambda\!\!\!^{-}$ (1.2.2)
Wave generation region:	$r \lesssim r_I$ (includes source, strong-field region, and weak-field near zone)
Local wave zone:	$r_I \lesssim r \lesssim r_0$
Distant wave zone:	$r_0 \lesssim r$.

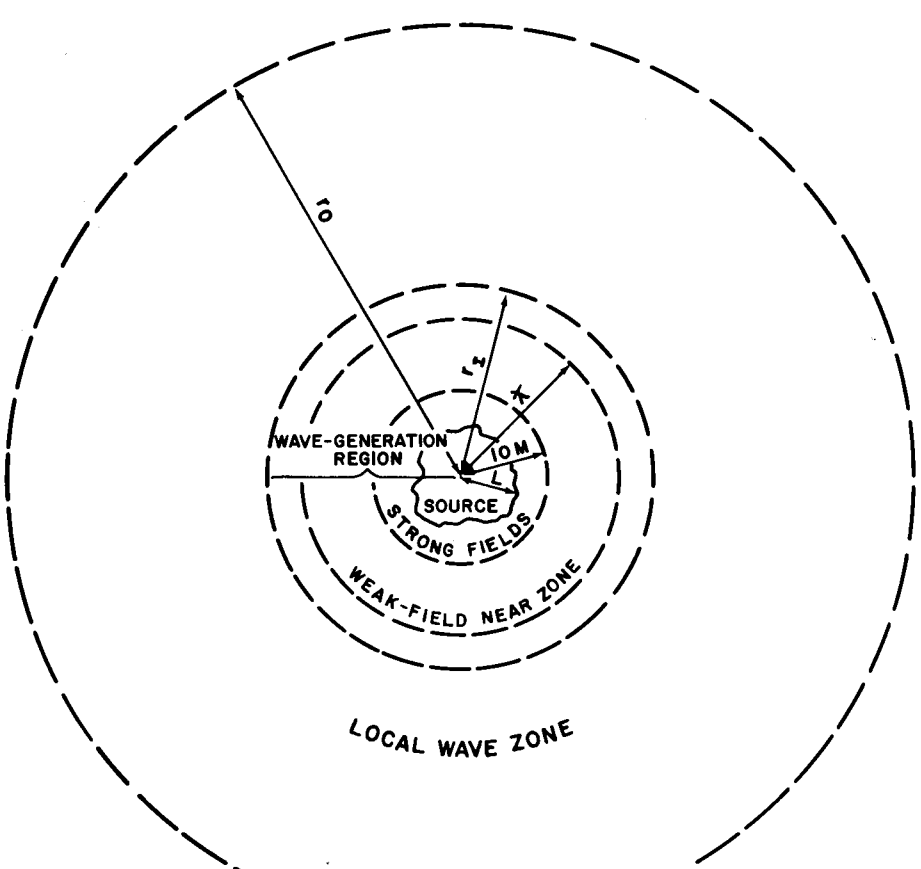


Figure 1. Regions of spacetime surrounding a source.

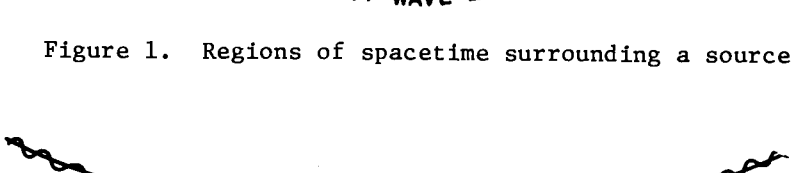


Figure 2. The local wave zone (schematic diagram). The smooth curve depicts the curvature of the background spacetime on which the waves (rippled curve) propagate. Near the source one cannot necessarily split the spacetime geometry into background plus waves; but in the local wave zone one can.

The "local wave zone" is the region in which (i) the source's waves are weak, outgoing ripples on a background spacetime; and (ii) the effects of the background curvature on the wave propagation are totally negligible (see Fig. 2).

The inner edge of the local wave zone r_I is the location at which one or more of the following effects becomes important: (i) the waves cease to be waves and become near-zone field, i.e., r becomes $\lesssim \lambda\!\!\!^{-}$; (ii) the gravitational pull of the source produces a significant redshift, i.e., r becomes $\backsim$ 2M = (Schwarzschild radius of source); (iii) the background curvature produced by the source distorts the wave fronts and backscatters the waves, i.e., $(r^3/M)^{1/2}$ becomes $\lesssim \lambda\!\!\!^{-}$; (iv) the outer limits of the source itself are encountered, i.e., r becomes $\lesssim$ L = (size of source). Thus, the inner edge of the local wave zone is given by

$$r_I = \alpha \times \text{Maximum } \{\lambda\!\!\!^{-},\ 2M,\ (M\lambda\!\!\!^{-}{}^2)^{1/3},\ L\}\ , \qquad (1.2.3)$$

$\alpha \equiv$ (some suitable number large compared to unity).

The outer edge of the local wave zone r_0 is the location at which one or more of the following effects becomes important: (i) a significant phase shift has been produced by the "M/r" gravitational field of the source, i.e., $(M/\lambda\!\!\!^{-}) \cdot \ell n(r/r_I)$ is no longer $<< \pi$; (ii) the background curvature due to nearby masses or due to the external universe perturbs the propagation of the waves, i.e., r is no longer $<< R_B$ = (background radius of curvature). Thus, the outer edge of the local wave zone is given by

$$r_0 = \text{Minimum } \{r_I \cdot \exp(\lambda\!\!\!^{-}/\beta M),\ R_B/\gamma\}, \qquad (1.2.4)$$

$\beta,\gamma \equiv$ (some suitable numbers large compared to unity).

Of course, we require that our large numbers α,β,γ, be adjusted so that the thickness of the local wave zone is very large compared to the reduced wavelength,

$$r_0 - r_I >> \lambda\!\!\!^{-}. \qquad (1.2.5)$$

Throughout these lectures I shall confine attention to sources which possess a local wave zone -- and I shall call such sources "isolated." It seems likely that every source of gravitational waves in the Universe today is "isolated." However, in the very early Universe the background curvature, $1/R_B^2$, was so large that sources might not have been isolated.

In complex situations the location of the local wave zone might not be obvious. Consider, for example, a neutron star passing very near a super-massive black hole. The tidal pull of the hole sets

the neutron star into oscillation, and the star's oscillations produce gravitational waves (Mashoon 1973; Turner 1977). If the hole is large enough, or if the star is far enough from it, there may exist a local wave zone around the star which does not also enclose the entire hole. Of greater interest - because more radiation will be produced - is the case where the star is very near the hole and the hole is small enough ($M_h \lesssim 100 M_\odot$) to produce large-amplitude oscillations, and perhaps even disrupt the star. In this case, before the waves can escape the influence of the star, they get perturbed by the background curvature of the hole. One must then consider the entire star-hole system as the source, and construct a local wave zone that surrounds them both.

The local wave zone acts as a buffer between the wave-generation region $r < r_I$ and the distant wave zone $r > r_0$. The existence of this buffer enables one to separate cleanly the theory of wave generation (applicable for $r < r_0$; treated in §§ 2-6 of these lectures) from the theory of wave propagation (applicable for $r > r_I$; treated in § 7 of these lectures).

1.3 The Gravitational-Wave Field

In the local wave zone where gravity is weak we shall use coordinate systems $(t,x,y,z) \equiv (x^0,x^1,x^2,x^3)$ which are centered on the source and are very nearly Minkowskiian; and we shall sometimes introduce the corresponding spherical coordinates (t,r,θ,ϕ) with

$$x = r\sin\theta\cos\phi\ ,\quad y = r\sin\theta\sin\phi\ ,\quad z = r\cos\theta. \qquad (1.3.1)$$

The components of the metric then differ only slightly from Minkowskii form

$$g_{\alpha\beta} \equiv \eta_{\alpha\beta} + h_{\alpha\beta},\quad \eta_{\alpha\beta} = \mathrm{diag}(-1,1,1,1),\quad |h_{\alpha\beta}| \ll 1. \qquad (1.3.2)$$

Throughout these lectures it will be adequate, in the local wave zone, to treat $h_{\alpha\beta}$ as a linearized field residing in flat space-time. If one knows $h_{\alpha\beta}$ in any such coordinate system (i.e. in any "gauge"), one can compute from it the "gravitational wave field" (§35.4 of MTW)

$$h^{TT}_{jk} \equiv (P_{ja}h_{ab}P_{bk} - 1/2\ P_{jk}P_{ab}h_{ab}). \qquad (1.3.3)$$

Here Latin indices run from 1 to 3; repeated Latin indices must be summed even if they are both "down"; and "TT" means "transverse, traceless part". The projection tensor used in this computation is

$$P_{jk} \equiv \delta_{jk} - n_j n_k\ ;\ n_j \equiv x^j/r. \qquad (1.3.4)$$

The task of wave-generation theory (§§ 2-6) is to determine h^{TT}_{jk} in the local wave zone. Once that has been done, wave-propagation theory (§ 7) can be used to calculate h^{TT}_{jk} in the distant wave zone.

1.4. Formalisms for Calculating Wave Generation

To calculate the generation of gravitational waves one must solve simultaneously the Einstein field equations and the equations of motion of the source. These equations are so complicated that one cannot hope to solve them exactly in any realistic situation. Therefore, one must resort to approximation schemes.

Several different approximation schemes have been devised for handling different types of sources. Most of these approximation schemes have been written in the form of "plug-in-and-grind" formalisms (i.e., computational algorithms). In these lectures I shall describe the following formalisms:

Weak-field formalisms (§2). These formalisms are valid for sources with weak internal gravitational fields ($L \gg 2M$; sources without any strong-field region; cf. Fig. 1). Section 2.1 will present a classification of weak-field formalisms and their realms of validity; and §§2.2-2.7 will present specific weak-field formalisms, each with its own realm of validity. A catalogue of the weak-field formalisms is given in Table 1.

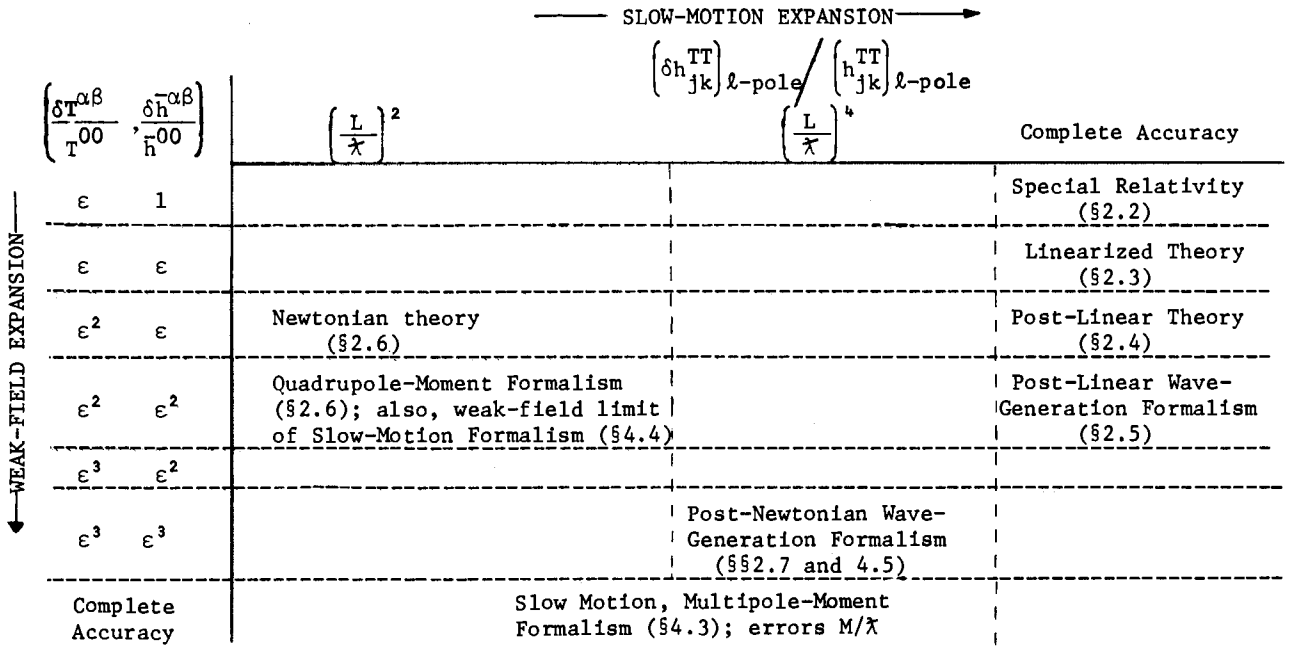
Multipole analysis of the radiation field (§3). The gravitational waves from any source can be resolved into multipole fields, and that multipole resolution yields simple formulas for the energy, linear momentum, and angular momentum carried off by the waves. Section 3 presents such a multipole analysis, restricted to the local wave zone. That analysis can be used fruitfully in conjunction with the wave-generation formulas of §§2,4, and 5.

Slow-motion formalisms (§4). These formalisms are valid for sources with slow internal motions ($L/\lambda\!\!\!^{-} \ll 1$). They make use of near-zone multipole-moment expansions, and consequently they tie in tightly with the multipole analysis of the radiation field (§3). The slow-motion formalisms are catalogued in Table 1 along with the weak-field formalisms.

Small-perturbation formalisms (§5). These formalisms are applicable to sources consisting of small dynamical perturbations of a nonradiative "background" system (e.g., small particles falling into black holes, and small-amplitude pulsations of stars).

TABLE 1

A CATALOGUE OF WEAK-FIELD FORMALISMS AND SLOW-MOTION FORMALISMS



Notation:

δ	means "errors in."
$\bar{h}^{\alpha\beta}$	is the gravitational field (including both nonradiative and radiative parts)
h^{TT}_{jk}	is the radiation field.
ε	is the magnitude of the source's internal gravitational fields.
L	is the source's linear size.
M	is the source's mass.
$\lambda\!\!\!^{-}$	is the reduced wavelength of the gravitational waves.

Formalisms for studying systems with strong internal fields and fast, large-amplitude internal motions (§6). Unfortunately, no analytic formalisms now exist for studying such systems. However one can study their evolution and the waves they emit by numerical solution of the Einstein field equations on a large computer. The necessary numerical techniques are now under development.

2. WEAK-FIELD FORMALISMS

2.1 Classification of Weak-Field Formalisms

Weak-field formalisms are applicable to systems with weak internal gravitational fields.

To compute a characteristic dimensionless strength ε of the internal field of a source, one can analyze the source as though spacetime were flat, using globally Minkowskii coordinates in which the center-of-mass is at rest. In these coordinates one can compute ε from the retarded integral

$$\varepsilon = \underset{\substack{\text{"relevant" values}\\ \text{of the field point}\\ (t,x^j)}}{\text{Maximum over all}} \left\{ \int \frac{T^{00}(t-|\underset{\sim}{x}-\underset{\sim}{x}'|,\ \underset{\sim}{x}')}{|\underset{\sim}{x}-\underset{\sim}{x}'|}\, d^3x' \right\} . \qquad (2.1.1)$$

Here T^{00} is the time-time component of the stress-energy tensor, and the "relevant" field points are those points at which one portion of the source interacts with fields produced by other portions of the source. One need not know ε with high precision; errors as large as a factor 3, say, are perfectly allowable.

For a source consisting of a single coherent body (e.g., a pulsating star) with mass M and linear size L,

$$\varepsilon \backsim M/L. \qquad (2.1.2a)$$

For a source consisting of several lumps, each with mass m and size ℓ, separated by distance $b \gg \ell$,

$\varepsilon \sim m/b$ if one is interested only in waves generated by relative motions of the lumps; (2.1.2b)

$\varepsilon \sim m/\ell$ if one is interested in waves generated by internal motions of the lumps. (2.1.2c)

Weak-field formalisms are valid only for sources that have $\varepsilon \ll 1$.

All weak-field formalisms have roughly the same structure: They utilize a single coordinate system $\{x^\alpha\}$ that covers the entire wave-generation region and local wave zone ($r \lesssim r_0$) and that is as nearly globally Minkowskiian as possible. In this coordinate system they define a "gravitational field" $\bar{h}^{\alpha\beta}$ (second-rank symmetric tensor) by

$$\bar{h}^{\alpha\beta} \equiv -[(-g)^{1/2} g^{\alpha\beta} - \eta^{\alpha\beta}], \tag{2.1.3}$$

where $g^{\alpha\beta}$ are the contravariant components of the metric tensor, where $g \equiv \det||g_{\alpha\beta}||$, and where $\eta^{\alpha\beta}$ are the components of the Minkowskii metric tensor [diag (-1,1,1,1)]. The formalisms then consist of field equations by which the stress-energy tensor $T^{\alpha\beta}$ (associated with matter and non-gravitational fields) generates the gravitational field $\bar{h}^{\alpha\beta}$, and equations of motion by which the gravitational field and internal stresses regulate the time evolution of the stress-energy tensor.

Thorne and Kovacs (1975; cited henceforth as TK) have devised a classification scheme for weak-field wave-generation formalisms--a scheme resembling that by which Havas and Goldberg (1962) classify "point-mass equation-of-motion formalisms." This scheme characterizes wave-generation formalisms by two integers n_T and n_h. These "order indices" tell one the magnitude of the errors made by the formalism--i.e., the amount by which the formalism's predictions should differ from those of exact general relativity theory:[1]

$$|(\text{errors in } T^{\mu\nu})/T^{00}| \sim \varepsilon^{n_T}, \tag{2.1.4a}$$

$$|(\text{errors in } \bar{h}^{\mu\nu})/\bar{h}^{00}| \sim \varepsilon^{n_h}. \tag{2.1.4b}$$

For example, a formalism of order $(n_T, n_h) = (1,1)$ makes fractional errors of order ε in both the stress-energy tensor and the gravitational field, while a formalism of order (2,1) makes fractional errors ε^2 in $T^{\mu\nu}$ and ε in $\bar{h}^{\mu\nu}$.

Errors in $\bar{h}^{\mu\nu}$, when fed into the equation of motion, produce errors in $T^{\mu\nu}$; and similarly, errors in $T^{\mu\nu}$, when fed into the field equations, produce errors in $\bar{h}^{\mu\nu}$. This feeding process places constraints on the order indices (n_T, n_h) of any self-consistent, weak-field wave-generation formalism:

$$n_h = n_T \quad \text{or} \quad n_h = n_T - 1. \tag{2.1.5}$$

[1] Note that all of the $|T^{\mu\nu}|$ are $\lesssim T^{00}$, and consequently all of the $|\bar{h}^{\mu\nu}|$ are $\lesssim \bar{h}^{00}$. This fact dictates the form of equations (2.1.4).

Thus, every weak-field formalism has order indices of the form (n,n-1) or (n,n) for some integer n.

TK devise a scheme for constructing formalisms of any desired order. They also show that any formalism of order (n,n-1) can readily be "strengthened" by augmenting onto it a higher-accuracy field-generation equation. The resulting, augmented formalism will have order (n,n).

In these lectures I shall not describe the general analysis of TK; I shall merely present examples of weak-field formalisms of various orders, and describe the augmentation process which improves their accuracy.

2.2 Special Relativity: A Formalism of Order (1,0)

Special Relativity is characterized by the equations of motion

$$T^{\mu\nu}{}_{,\nu} = 0, \tag{2.2.1}$$

which are oblivious of all gravitational effects. The largest of the individual terms that occur in these equations are of order

$$T^{00}/\ell, \tag{2.2.2}$$

where ℓ is a characteristic length scale inside the source. By contrast, the gravitational forces ignored by these equations of motion are

$$\Gamma^{\mu}{}_{\alpha\nu}T^{\alpha\nu}+\Gamma^{\nu}{}_{\alpha\nu}T^{\mu\alpha}\sim(\bar{h}^{00}/\ell)T^{00}\sim\varepsilon(T^{00}/\ell). \tag{2.2.3}$$

Comparison of the forces ignored (eq. 2.2.3) with the terms included (eq. 2.2.2) shows that special relativity makes fractional errors of order ε in the evolution of the stress-energy--and hence fractional errors of order ε in the stress-energy tensor itself. Evidently, the stress-energy tensor has order index $n_T = 1$.

The gravitational field, by contrast, has order index $n_h=0$, since special relativity is totally oblivious of gravity and therefore makes fractional errors $\delta\bar{h}^{\mu\nu}/\bar{h}^{00}\sim 1 = \varepsilon^0$.

Conclusion: Special relativity (eq. 2.2.1) is a weak-field formalism of order (1,0).

2.3. Linearized Theory: A Formalism of Order (1,1)

One can obtain Linearized Theory from Special Relativity very easily: One leaves unchanged the stress-energy tensor and its equations of motion

$$T^{\mu\nu}{}_{,\nu} = 0, \tag{2.3.1a}$$

but one postulates that this stress-energy generates a gravitational field $\bar{h}^{\mu\nu}$ by the relation

$$\bar{h}^{\mu\nu}(x) = 16\pi \int G(x,x') T^{\mu\nu}(x') d^4x'. \tag{2.3.1b}$$

Here d^4x' is the special relativistic volume element

$$d^4x' \equiv dx^{0'} dx^{1'} dx^{2'} dx^{3'}, \tag{2.3.2}$$

and $G(x,x')$ is the flat-space propagator (Green's function)

$$G(x,x') = \frac{1}{4\pi} \delta_{ret}[1/2(x^\alpha - x^{\alpha'})(x^\beta - x^{\beta'})\eta_{\alpha\beta}], \tag{2.3.3a}$$

$$\delta_{ret} \equiv \begin{cases} \text{Dirac delta function for } x^\alpha \text{ in future of } x^{\alpha'} \\ 0 \qquad\qquad\qquad\qquad \text{for } x^\alpha \text{ in past of } x^{\alpha'}. \end{cases} \tag{2.3.3b}$$

[Notice: (1) in the arguments of $\bar{h}^{\mu\nu}$, G, and $T^{\mu\nu}$ we omit the indices of the coordinates x^α and $x^{\alpha'}$; also (2) by integrating over time, $x^{0'}$, in eq. (2.3.1b) one obtains the expression

$$\bar{h}^{\mu\nu}(x^0, x^j) = 4 \int \frac{T^{\mu\nu}(x^{0'} = x^0 - |\underset{\sim}{x} - \underset{\sim}{x}'|, x^{j'})}{|\underset{\sim}{x} - \underset{\sim}{x}'|} d^3x', \tag{2.3.4}$$

which is familiar from elementary treatises on Linearized Theory; e.g., chapter 18 of MTW.]

The linearized gravitational field (2.3.1b) is a fairly good approximation to the exact general relativistic gravitational field, $\bar{h}^{\alpha\beta} \equiv -[(-g)^{1/2} g^{\alpha\beta} - \eta^{\alpha\beta}]$. It makes fractional errors of order ε. Hence, Linearized Theory has a gravitational order index $n_h = 1$, which is one order "better" than Special Relativity; but its stress-energy order index is the same as that of Special Relativity, $n_T = 1$. The total order of Linearized Theory is $(n_T, n_h) = (1,1)$.

To within the accuracy of Linearized Theory the metric perturbation $h_{\mu\nu}$ is the trace-reversal of $\bar{h}_{\mu\nu}$:

$$h_{\mu\nu} = \bar{h}_{\mu\nu} - 1/2\eta_{\mu\nu} \bar{h}; \quad \bar{h} \equiv \bar{h}_\alpha{}^\alpha. \tag{2.3.5a}$$

(Indices in Linearized Theory are raised and lowered with $\eta_{\mu\nu}$.) Since $h_{\mu\nu}$ and $\bar{h}_{\mu\nu}$ differ only by a trace, the gravitational-wave field in the local wave zone is (cf. eq. 1.3.3)

$$h^{TT}_{jk} = \bar{h}^{TT}_{jk} = P_{ja} P_{kb} \bar{h}_{ab} - 1/2 P_{jk} (P_{ab} \bar{h}_{ab}). \tag{2.3.5b}$$

The fractional errors in this gravitational-wave field are

$$(\text{fractional errors in } \bar{h}^{TT}_{jk}) \equiv \left| \frac{\delta \bar{h}^{TT}_{jk}}{\bar{h}^{TT}_{jk}} \right| = \left| \frac{\delta \bar{h}^{TT}_{jk}}{\bar{h}^{00}} \; \frac{\bar{h}^{00}}{\bar{h}^{TT}_{jk}} \right| \tag{2.3.6}$$

$$\sim \frac{\varepsilon}{|\bar{h}^{TT}_{jk}/\bar{h}^{00}|} \sim \frac{\varepsilon}{|\bar{h}^{TT}_{jk}/Mr^{-1}|} .$$

Here $\bar{h}^{00} \backsim$ (Newtonian potential) $\backsim M/r$ is the largest of the components of the gravitational field; cf. equation (2.1.4b) and footnote 1.

The "rules" for using Linearized Theory to calculate gravitational-wave generation are as follows: (1) Express the special relativistic stress-energy tensor $T^{\mu\nu}$ in terms of the non-gravitational variables of the specific system being analyzed. (2) Solve the special-relativistic equations of motion (2.3.1a) to determine the evolution of the stress-energy tensor. (3) Evaluate the integral (2.3.1b) to determine the first-order gravitational field $\bar{h}^{\mu\nu}$ in the local wave zone. (4) Project out the gravitational-wave field $\bar{h}^{TT}_{jk}$ using equation (2.3.5b). (5) Check, using equation (2.3.6), that the errors in the wave field are acceptably small.

This set of rules shows very clearly the sense in which "Linearized Theory is the theory of order (1,1) obtained by augmenting onto Special Relativity [theory of order (1,0)] field generation equations"; cf. next to last paragraph of § 2.1. In particular, when working in Linearized Theory one at first (Rules 1 and 2) pretends that one is in Special Relativity. Only when one starts evaluating the radiation field (Rules 3 and 4) and its errors (Rule 5) does one depart from Special Relativity.

Recently Press (1977) has transformed the gravitational-wave field (2.3.4) and (2.3.5) of Linearized Theory into the form

$$\bar{h}^{TT}_{jk} = \frac{2}{r}\left\{\frac{d^2}{dt^2}\int [T_{00}+2T_{0p}n_p+T_{pq}n_pn_q]_{ret} x_{j'} x_{k'} \, d^3x'\right\}^{TT}. \tag{2.3.7a}$$

Here "ret" means "evaluated at the retarded time"

$$[T_{00}]_{ret} \equiv T_{00}(t-|\underset{\sim}{x}-\underset{\sim}{x}'|, \underset{\sim}{x}') \tag{2.3.7b}$$

and n_p is the unit radial vector pointing from the source toward the distant observer,

$$n_p \equiv x^p/r \ . \tag{2.3.7c}$$

Press's expression (2.3.7a) is particularly useful for systems in which $|T_{0j}| \ll T_{00}$ and $|T_{jk}| \ll T_{00}$, since then it involves only the second moment of the retarded energy distribution, T_{00ret}. Similar expressions are encountered in the "quadrupole-moment formalism" (eqs. 2.6.4 and 2.6.5 below). However there one requires that the source be confined deep within its own near zone ("slow-motion assumption"), whereas Press's expression requires no such constraint.

Press's expression, appropriately modified, has wider validity than just Linearized Theory: Whenever one has a formalism in which the local-wave-zone field $\bar{h}^{\mu\nu}$ can be written as:

$$\bar{h}^{\mu\nu}(x^0,\underset{\sim}{x})=\frac{4}{r}\int \tau^{\mu\nu}{}_{ret}\, d^3x', \text{ with } \tau^{\mu\nu}{}_{,\nu}=0. \tag{2.3.8}$$

then (as Press emphasizes) expression (2.3.7a) is valid with $T^{\mu\nu}$ replaced by $\tau^{\mu\nu}$ --whatever that animal may be. The weak-field formalisms classified by TK (§2.1 above) all have this property; cf. §II of TK.

2.3.1. A Valid Application of Linearized Theory

Linearized Theory is applicable whenever one can ignore self-gravitational forces inside the source -- i.e., whenever fractional errors

$$\begin{aligned} &|\delta T^{\mu\nu}/T^{00}| \sim \varepsilon \sim (\text{internal gravitational potential}) \\ &|\delta \bar{h}^{TT}_{ij}/\bar{h}^{TT}_{ij}| \sim \frac{\varepsilon}{|\bar{h}^{TT}_{ij}/\bar{h}^{00}|} \end{aligned} \tag{2.3.9}$$

are acceptable. The following is an example of a valid application of Linearized Theory:

A steel bar of mass $M \sim 500$ tons, length $L \sim 20$ meters, and diameter $D \sim 2$ meters generates gravitational waves by rotating end-over-end with angular velocity $\omega \sim 28$ radians per second [Einstein (1918); Eddington (1922); §36.3 of MTW]. (Faster rotation would tear the bar apart.) For such a bar the internal gravitational field is

$$\varepsilon \sim \frac{M}{D} \sim \left(\frac{5\times10^{8}\,g}{2\times10^{2}\,cm}\right)\times\left(\frac{0.7\times10^{-28}\,cm}{g}\right) \sim 2\times10^{-22}. \qquad (2.3.10)$$

The radiation field and the gravitational potential $\bar{h}^{00}$, as calculated from Linearized Theory, turn out to be

$$|\bar{h}^{TT}_{ij}| \sim (ML^2\omega^2/r)\cos[2\omega(t-r) + \text{phase}] \qquad (2.3.11)$$

$$\bar{h}^{00} = 4M/r; \qquad (2.3.12)$$

and the fractional errors in the radiation field (eq. 2.3.6) are thus

$$\left|\frac{\delta\bar{h}^{TT}_{jk}}{\bar{h}^{TT}_{jk}}\right| \sim \frac{\varepsilon}{|\bar{h}^{TT}_{jk}/\bar{h}^{00}|} \sim \frac{\varepsilon}{(\omega L)^2} \qquad (2.3.13)$$

$$\sim \frac{10^{-22}}{[(2\times10^{3}\,cm)(28\,sec^{-1})(3\times10^{10}\,cm/sec)^{-1}]^2} \sim 10^{-11}$$

2.3.2. Invalid Applications of Linearized Theory

For most astrophysical systems internal gravity is important, and Linearized Theory is thus invalid. Examples are as follows:

(1) Gravitational waves from nonradial pulsations of a star. Let the star have mass M and radius R. Then its mean density ρ and its internal gravitational field strength ε are

$$\rho \sim M/R^3, \qquad \varepsilon \sim M/R; \qquad (2.3.14)$$

and its mean pressure [calculated from hydrostatic equilibrium, $dp/dr \sim \rho M/R^2$] is

$$p \sim \rho(M/R) \sim \varepsilon\rho\ . \qquad (2.3.15)$$

Since Linearized Theory makes fractional errors $|\delta T^{\mu\nu}/T^{00}| \sim \varepsilon$, its fractional errors in the pressure are

$$\frac{\delta p}{p} \sim \left|\frac{\delta T^{ij}}{T^{00}}\ \frac{T^{00}}{T^{ij}}\right| \sim \left|\varepsilon\,\frac{\rho}{p}\right| \sim 1. \qquad (2.3.16)$$

Thus, Linearized Theory makes unacceptable errors in the star's internal pressure forces, as well as in its internal gravity.

One should not be surprised to learn that its errors in the gravitational waves due to stellar pulsation are also unacceptably large

$$\left| \frac{\delta \bar{h}^{TT}_{jk}}{\bar{h}^{TT}_{jk}} \right| \gtrsim 1. \qquad (2.3.17)$$

(2) Gravitational waves from a near-encounter between two fast-moving stars. Linearized theory predicts no gravitational interaction between the stars. Therefore, according to Linearized Theory, each star proceeds undisturbed along its straight path through flat space, oblivious of the other star. As a result, no gravitational waves are produced -- an obviously incorrect prediction.

2.4 Post-Linear Theory: A Formalism of Order (2,1)

When analyzing the internal structure and motion of systems with significant self gravity, one needs a formalism which makes fractional errors $|\delta T^{\mu\nu}/T^{00}| \lesssim \varepsilon^2$ ---i.e., which has a stress-energy order index $n_T \geq 2$. The least accurate and least complex such formalism is Post-Linear Theory. It has order $(n_T, n_h) = (2,1)$.

For a derivation of Post-Linear Theory from general relativity theory, see Thorne and Kovács (1975; "TK").

In Post-Linear Theory one describes gravity by a gravitational field ${}_1\bar{h}^{\mu\nu}$, which is a good approximation to the general relativistic field (eq. 2.1.3):

$${}_1\bar{h}^{\mu\nu} = -[\sqrt{-g}\, g^{\mu\nu} - \eta^{\mu\nu}][1 + O(\varepsilon)]. \qquad (2.4.1)$$

Post-Linear Theory, like Special Relativity and Linearized Theory, utilizes the language of flat spacetime. For example, one raises and lowers indices on the gravitational field by means of the flat Lorentz metric $\eta_{\mu\nu}$; and one performs "trace reversals" in the familiar manner of Linearized Theory:

$$ {}_1h^{\mu\nu} = {}_1\bar{h}^{\mu\nu} - 1/2\eta^{\mu\nu}\, {}_1\bar{h}\ ; \quad {}_1\bar{h}^{\mu\nu} = {}_1h^{\mu\nu} - 1/2\eta^{\mu\nu}\, {}_1h; $$

$$ {}_1\bar{h} = -{}_1h = {}_1\bar{h}^{\alpha}_{\alpha} = -{}_1h^{\alpha}_{\alpha} \qquad (2.4.2)$$

(cf. chapter 18 of MTW). The metric of general relativity is approximated by

$$g_{\mu\nu} = \eta_{\mu\nu} + {}_1h_{\mu\nu}\,[1 + O(\varepsilon)]. \qquad (2.4.3)$$

The stress-energy tensor in Post-Linear Theory will typically be expressed in terms of the electromagnetic field $F_{\mu\nu}$, the matter variables (density, pressure, velocity, viscosity,...), and the gravitational field ${}_1\bar{h}^{\mu\nu}$. It must be constructed from these variables in accordance with all the rules of general relativity, to within fractional errors of $O(\epsilon^2)$.

The equations of motion of Post-Linear Theory are formally the same as those of general relativity

$$T^{\mu\nu}{}_{,\nu} + {}_1\Gamma^{\mu}{}_{\alpha\nu}\, T^{\alpha\nu} + {}_1\Gamma^{\nu}{}_{\alpha\nu}\, T^{\mu\alpha} = 0; \qquad (2.4.4a)$$

however, the Christoffel symbols ${}_1\Gamma^{\mu}{}_{\alpha\beta}$ are given by the "first-order approximation"

$${}_1\Gamma^{\mu}{}_{\alpha\beta} = \eta^{\mu\nu}\, {}_1\Gamma_{\nu\alpha\beta} \; ; \; {}_1\Gamma_{\nu\alpha\beta} = ({}_1h_{\nu\alpha,\beta} + {}_1h_{\nu\beta,\alpha} - {}_1h_{\alpha\beta,\nu}). \qquad (2.4.4b)$$

The field equations of Post-Linear Theory are the same as those of Linearized Theory

$${}_1\bar{h}^{\mu\nu,\alpha}{}_{\alpha} = -16\pi\, T^{\mu\nu}; \qquad (2.4.4c)$$

and their solution can be written, using the flat-space propagator (eq. 2.3.3), as

$${}_1\bar{h}^{\mu\nu}(x) = 16\pi \int G(x,x')T^{\mu\nu}(x')d^4x'. \qquad (2.4.5)$$

Notice that Post-Linear Theory differs from Linearized Theory in one crucial, but simple way: It allows the gravitational field ${}_1\bar{h}^{\mu\nu}$ to "push the matter around". [Christoffel symbols have been inserted into the equation of motion; compare eqs. (2.3.1a) and (2.4.4a).] This difference is crucial for astrophysics. It allows Post-Linear Theory to treat accurately the structure and evolution of stars, planets, planetary systems, star clusters, and near stellar encounters -- unless the stars are highly compact (i.e., unless they are neutron stars or black holes). Linearized Theory makes enormous errors on all such systems.

2.5. Post-Linear Wave Generation: A Formalism of Order (2,2)

Despite its fine ability to analyze the internal dynamics of astrophysical systems, Post-Linear Theory does a bad job of predicting their gravitational-wave generation. For wave generation it has no better accuracy than Linearized Theory.

Fortunately, there is a simple way to improve its accuracy. One need only append to it a gravitational field ${}_2\bar{h}^{\mu\nu}$, which is

more accurate than ${}_1\bar{h}^{\mu\nu}$:

$$ {}_2\bar{h}^{\mu\nu} = -[\sqrt{-g}\, g^{\mu\nu} - \eta^{\mu\nu}][1 + 0(\varepsilon^2)]. \qquad (2.5.1) $$

By appending ${}_2\bar{h}^{\mu\nu}$ onto Post-Linear Theory, one boosts its accuracy from order (2,1) to order (2,2). The resulting formalism, called the "Post-Linear Wave-Generation Formalism," bears the same relationship to Post-Linear Theory as Linearized Theory does to Special Relativity.

TK have derived a formula for ${}_2\bar{h}^{\mu\nu}$ in terms of ${}_1\bar{h}^{\mu\nu}$ and $T^{\mu\nu}$. Their formula involves the flat-space propagator

$$ G(x,x') = \frac{1}{4\pi}\, \delta_{ret}[1/2(x^\alpha - x^{\alpha'})(x^\beta - x^{\beta'})\eta_{\alpha\beta}] \qquad (2.5.2a) $$

[eq. (2.3.3)], and also its derivative with respect to the argument of the delta function

$$ G'(x,x') = \frac{1}{4\pi}\, \delta'_{ret}[1/2\,(x^\alpha - x^{\alpha'})(x^\beta - x^{\beta'})\eta_{\alpha\beta}]; \qquad (2.5.2b) $$

$$ \delta'_{ret}(z) \equiv (d/dz)\, \delta_{ret}(z). $$

The TK formula splits ${}_2\bar{h}^{\mu\nu}$ into five pieces

$$ {}_2\bar{h}^{\mu\nu} = {}_2\bar{h}^{\mu\nu}_D + {}_2\bar{h}^{\mu\nu}_F + {}_2\bar{h}^{\mu\nu}_{TL} + {}_2\bar{h}^{\mu\nu}_{TR} + {}_2\bar{h}^{\mu\nu}_W. \qquad (2.5.3) $$

Each piece has a particular physical interpretation. However, one must be careful not to take that interpretation too seriously -- for reasons discussed below. The five pieces are as follows:

Direct field, ${}_2\bar{h}^{\mu\nu}_D$. This field is produced directly by the Post-Linear stress-energy tensor $T^{\mu\nu}$; and it propagates from the source point $x^{\alpha'}$ to the field point x^α by means of the flat-space propagator. In other words, it propagates along the flat-space light cone with parallel-propagation of components and with a "1/r" fall-off of amplitude. The formula for this direct field is

$$ {}_2\bar{h}^{\mu\nu}_D(x) = 16\pi \int G(x,x')T^{\mu\nu}(x')[1 - {}_1\bar{h}(x')]d^4x'. \qquad (2.5.4) $$

This part of ${}_2\bar{h}^{\mu\nu}$ is $\sim M/r$, whereas the other four parts are typically $\lesssim M/r$.

Focussing field, ${}_2\bar{h}^{\mu\nu}_F$. When the gravitational field generated at $x^{\alpha'}$ propagates through regions of nonzero Ricci curvature -- i.e., through matter --, it gets focussed. This focussing increases

the amplitude of ${}_2\bar{h}^{\mu\nu}$, without changing its directionality (i.e., without changing the relative magnitude of its components). The amount of focussing between a source point x^α and a field point x^α is described by the "focussing function"

$$\alpha(x,x') \equiv 1/2\ X^\alpha X^\beta \int_0^1 {}_1R_{\alpha\beta}(x^{\mu'} + \lambda X^\mu)\lambda(1-\lambda)d\lambda, \qquad (2.5.5a)$$

$$X^\alpha \equiv x^\alpha - x^{\alpha'}.$$

Here ${}_1R_{\alpha\beta}(x^{\mu'} + \lambda X^\mu)$ is the first-order Ricci tensor, (calculated from $g_{\alpha\beta} = \eta_{\alpha\beta} + {}_1h_{\alpha\beta}$), evaluated at the event $x^{\mu'} + \lambda X^\mu$, which lies a fraction λ of the way along the straight line between source point and field point. A formula for ${}_1R_{\alpha\beta}$ is

$${}_1R_{\alpha\beta} = -1/2\ {}_1h_{\alpha\beta,\rho}{}^{,\rho} = 8\pi(T_{\alpha\beta} - 1/2\eta_{\alpha\beta}T). \qquad (2.5.5b)$$

For intuition into the focussing function, see Figure 3. In terms of the focussing function α, the flat-space propagator G, and the stress-energy tensor $T^{\mu\nu}$, the focussing field is

$${}_2\bar{h}_F^{\mu\nu}(x) = 16\pi \int \alpha(x,x')G(x,x')T^{\mu\nu}(x')d^4x'. \qquad (2.5.5c)$$

Tail Field, ${}_2\bar{h}_{TL}^{\mu\nu}$. Consider the gravitational field ${}_2\bar{h}^{\mu\nu}$ generated at an event $x^{\alpha'}$. It has a "wave front" that propagates outward, initially spherically and initially along the future light cone of $x^{\alpha'}$. However, focussing produces dimples in the wave front; and dimpling, when analyzed from the viewpoint of Huygens' principle, produces waves that radiate outward from the dimpled region in all directions. (See Figure 4.) The result is a "tail" of the wave field. Let a wave originating at $x^{\alpha'}$ arrive at an event $x^{\alpha''}$, with a dimple in its wave front due to focussing. The amplitude $\beta(x'', x')$ for this dimple to produce a tail is given by

$$\beta(x'',x') = X^{\alpha''} X^{\beta''} \int_0^1 {}_1R_{\alpha\beta}(x^{\mu'} + \lambda X^{\mu''})\lambda^2 d\lambda, \qquad (2.5.6a)$$

$$X^{\alpha''} \equiv x^{\alpha''} - x^{\alpha'}. \qquad (2.5.6b)$$

This amplitude is called the "tail-generating function." The tail which it generates is given by

$$\bar{h}_{TL}^{\mu\nu}(x) = -16\pi \iint_{x'\varepsilon I^-(x)} G(x,x'')\beta(x'',x')G'(x'',x')T^{\mu\nu}(x')d^4x''d^4x'. \qquad (2.5.6c)$$

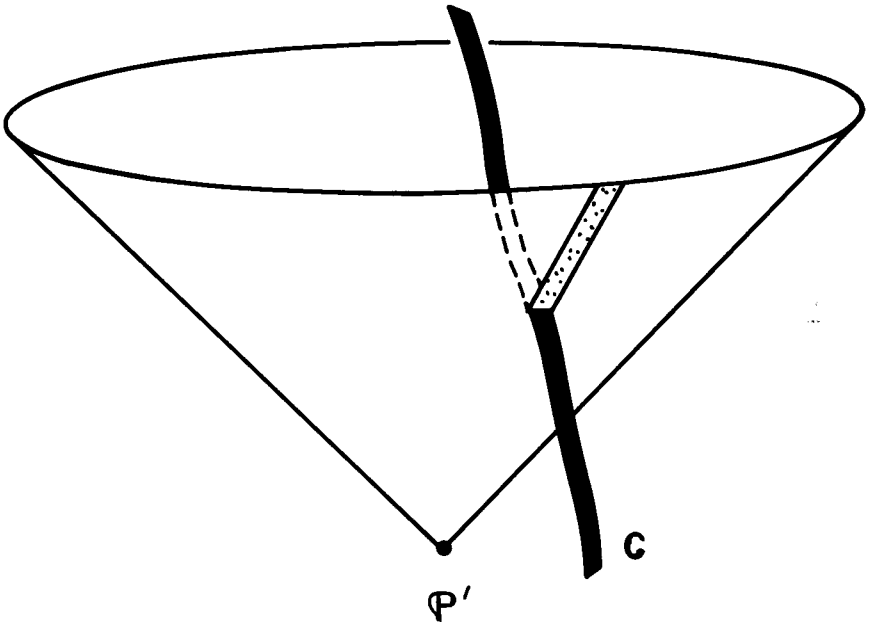


Figure 3. The region of focussing for a gravitational field generated at $\mathcal{P}'$, which propagates through spacetime containing a lump of matter (world tube $\mathcal{C}$). Focussing occurs, and α and $\bar{h}^{\mu\nu}_{\mathrm{F}}$ are nonzero, in the stippled region--i.e., in the region containing rays that have passed through the lump.

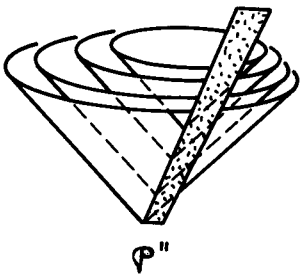


Figure 4. The tail of the gravitational field described in Figure 3. In the region of focussing (stippled region) the tail-generating function β is non-zero. Thus, each point $\mathcal{P}''$ in the stippled region is a source of tail -- and the tail propagates outward from each such $\mathcal{P}''$ along its future light cone.

Here the expression $x' \in I^-(x)$ means that the integration must extend over source points $x^{\alpha'}$ that lie inside but not on the past (flat-space) light cone of x^α.

Transition Field, $_2\bar{h}^{\mu\nu}_{TR}$. The gravitational field $_2\bar{h}^{\mu\nu}$ generated at an event $x^{\alpha'}$ does not really propagate along light cones of the (fictitious) flat space of Post-Linear Theory. Rather, it propagates along light cones of the slightly curved metric $g_{\mu\nu} = \eta_{\mu\nu} + {}_1h_{\mu\nu}$. The result is a time delay relative to flat-space propagation -- a delay measured, for example, in the "Shapiro radar time-delay experiment" (Shapiro 1964; §40.4 of MTW). The direct field (eq. 2.5.4) fails to take this time delay into account. Therefore, one must correct it for the effects of the time delay. The "transition field" $_2\bar{h}^{\mu\nu}_{TR}$ does that correcting. The correction is embodied in a "time-delay function"

$$\gamma(x,x') = \tfrac{1}{2}X^\alpha X^\beta \int_0^1 {}_1h_{\alpha\beta}(x^{\mu'} + \lambda X^\mu)d\lambda, \qquad (2.5.7a)$$

$$X^\mu \equiv x^\mu - x^{\mu'}. \qquad (2.5.7b)$$

For two events x and x' separated by a distance ℓ in the laboratory frame, this time-delay function is

$$\gamma(x,x') = \ell\Delta t_s, \qquad (2.5.7c)$$

where Δt_s is the "Shapiro time delay" (difference between curved-space and flat-space propagation times). In terms of this time-delay function, the transition field is given by[2]

$$\bar{h}^{\mu\nu}_{TR}(x) = 16\pi \int \gamma(x,x')G'(x,x')T^{\mu\nu}(x')d^4x'. \qquad (2.5.7d)$$

The name "transition field" is taken from electromagnetic theory: When a charged particle moves, with uniform velocity, through a medium of variable index of refraction, it radiates. The radiation (called "electromagnetic transition radiation") is caused by variations in the speed of propagation of the particle's Coulomb field.

[2]The expression for $\bar{h}^{\mu\nu}_{TR}$ given by TK differs slightly from eq.(2.5.7d): TK have to "truncate the external time delay" because they do not restrict attention to the local wave zone, wher the external delay is negligible.

The "gravitational transition field" (2.5.7d) has a similar origin: It is caused by variations in the "flat-space" speed of propagation of the particle's direct ("Coulomb") gravitational field.

Whump field, $_2\bar{h}^{\mu\nu}_W$. The gravitational field $_2\bar{h}^{\mu\nu}$ is generated not only by the material stress-energy tensor $T^{\mu\nu}$, but also by "gravitational stresses." For example, when two stars go flying past each other at high speed, their relative gravitational potential energy builds up quickly, then dies out quickly (i.e., goes "whump"), and in doing so it produces a burst of second-order gravitational field. TK call this field the "whump field", not only in the case of stellar fly-by, but also in general. The "gravitational stresses" $W^{\mu\nu}$ which generate it are a certain sum of products of first derivatives of $_1\bar{h}^{\alpha\beta}$; in particular,

$$W^{\mu\nu} \equiv {}_1t^{\mu\nu}_{LL} + (16\pi)^{-1}\, {}_1\bar{h}^{\mu\rho}{}_{,\sigma}\, {}_1\bar{h}^{\nu\sigma}{}_{,\rho} \tag{2.5.8a}$$

where $_1t^{\mu\nu}$ is the "Landau-Lifshitz pseudotensor," accurate up to fractional errors $\sim \varepsilon$:

$$\begin{aligned} {}_1t^{\alpha\beta}_{LL} = (16\pi)^{-1} \Big\{ &\tfrac{1}{2}\eta^{\alpha\beta}\, {}_1\bar{h}_{\lambda\nu,\rho}\, {}_1\bar{h}^{\lambda\rho,\nu} + {}_1\bar{h}^{\alpha}{}_{\lambda,\nu}\, {}_1\bar{h}^{\beta\lambda,\nu} \\ &-({}_1\bar{h}_{\nu\rho}{}^{,\alpha}\, {}_1\bar{h}^{\beta\nu,\rho} + {}_1\bar{h}_{\nu\rho}{}^{,\beta}\, {}_1\bar{h}^{\alpha\nu,\rho}) + \tfrac{1}{2}\, {}_1\bar{h}_{\lambda\nu}{}^{,\alpha}\, {}_1\bar{h}^{\lambda\nu,\beta} \\ &-\tfrac{1}{4}\eta^{\alpha\beta}\, {}_1\bar{h}_{\lambda\nu,\rho}\, {}_1\bar{h}^{\lambda\nu,\rho} - \tfrac{1}{4}\, {}_1\bar{h}^{,\alpha}\, {}_1\bar{h}^{,\beta} + \tfrac{1}{8}\eta^{\alpha\beta}\, {}_1\bar{h}_{,\lambda}\, {}_1\bar{h}^{,\lambda} \Big\}. \end{aligned} \tag{2.5.8b}$$

In terms of these stresses, the whump field is

$$ {}_2\bar{h}^{\mu\nu}_W(x) = 16\pi \int G(x,x')W^{\mu\nu}(x')\,d^4x'. \tag{2.5.8c}$$

The gravitational waves emitted by a Post-Linear source are described by the "tranverse, traceless part" of $_2\bar{h}^{\mu\nu}$, evaluated in the local wave zone

$$h^{TT}_{jk} = \bar{h}^{TT}_{jk} = P_{ja}\, {}_2\bar{h}_{ab}\, P_{bk} - \tfrac{1}{2} P_{jk}(P_{ab}\, {}_2\bar{h}_{ab}). \tag{2.5.9}$$

Here P_{jk} is the usual transverse projection tensor(eq. 1.3.4).

In the local wave zone $_2\bar{h}^{\mu\nu}$ satisfies the flat-space Einstein field equations $_2\bar{h}^{\mu\nu}{}_{,\alpha\beta}\eta^{\alpha\beta} = 0$ and the flat-space "Lorentz gauge condition" $_2\bar{h}^{\mu\nu}{}_{,\nu} = 0$, except for fractional errors of $O(\varepsilon^2)$. As a consequence, the gravitational-wave field h^{TT}_{jk}

transforms, under Lorentz transformations, in the usual manner for a field of spin two and zero rest mass.

Unfortunately, each of the five individual pieces of ${}_2\bar{h}^{\mu\nu}$, by itself, fails to satisfy the gauge condition. As a consequence, the transverse-traceless part of each piece (e.g., $[{}_2\bar{h}^{jk}_{TL}]^{TT}$) fails to transform as a spin-two, zero-rest-mass field. This fact prevents one from attributing a rigorous physical significance to each of the five pieces by itself. The physical interpretations given above are therefore somewhat heuristic. For further discussion, see TK, and also §VIc of Kovács and Thorne (1977a).

Crowley and Thorne (1977; their eq. 37) have derived several potentially useful formulas for the sum of the focussing, tail, and transition fields. An example is:

$$ {}_2\bar{h}^{\mu\nu}_F(x) + {}_2\bar{h}^{\mu\nu}_{TL}(x) + {}_2\bar{h}^{\mu\nu}_{TR}(x) = $$

$$ = -\frac{\partial}{\partial x^\alpha}\int G(x,x')\, {}_1\bar{h}^{\alpha\beta}(x')\frac{\partial}{\partial x^\beta}\, {}_1\bar{h}^{\mu\nu}(x')d^4x'. \qquad (2.5.10) $$

Crowley and Thorne also examine the issue of when one can combine the Post-Linear Wave-Generation Formalism with a point-particle description of gravitating bodies. The answer is non-trivial: Some sets of post-linear formulas are compatible with point-particle descriptions of matter; others are not.

The accuracy of the Post-Linear Wave-Generation Formalism is a slightly delicate issue. Until one looks closely, one expects

$$ (\text{fractional errors in } {}_2\bar{h}^{\mu\nu}) \sim \varepsilon^2; \qquad (2.5.11a) $$

$$ (\text{fractional errors in } {}_2\bar{h}^{TT}_{jk}) \sim \frac{\varepsilon^2}{|\bar{h}^{TT}_{jk}/\bar{h}^{00}|}\ . \qquad (2.5.11b) $$

However, the formalism as expounded above relies on two approximations which can sometimes produce larger errors than (2.5.11): First, it assumes that focussing is small -- i.e., that the focussing function satisfies

[3] See, e.g., §III.B of Eardley, Lee, and Lightman (1974), with the change of notation $\underset{\sim}{e}_{\hat{z}} \to \underset{\sim}{n} = \underset{\sim}{e}_{\hat{r}}$ = (unit radial vector), $\underset{\sim}{e}_{\hat{x}} \to \underset{\sim}{e}_{\hat{\theta}}$ = (unit vector in θ direction), $\underset{\sim}{e}_{\hat{y}} \to \underset{\sim}{e}_{\hat{\varphi}}$ = (unit vector in φ direction), and with specialization to the spin-2 case, $\Psi_2 = \Psi_3 = \Phi_{22} = 0$; and also Thorne (1977d).

$$|\alpha(x,x')| << 1 \text{ for all } x' \text{ inside source and } x \text{ in local wave zone.} \qquad (2.5.12)$$

Second, it assumes that the Shapiro time delays inside the source are short compared to the characteristic timescale for the source's motion -- i.e., that Δt_s (eq. 2.5.7c) and the reduced wavelength of the radiation, $\lambda\!\!\!^{-}$, satisfy

$$\Delta t_s \ll \lambda\!\!\!^{-} . \qquad (2.5.13)$$

These assumptions could be dropped, but only at the cost of making the formalism somewhat more complicated than it already is.

For realistic sources the "short-Shapiro-time-delay" assumption (2.5.13) is generally satisfied. However, one can readily imagine interesting weak-field sources ($\varepsilon << 1$) which violate the "small-focussing assumption." For example, for two stars of mass m and size ℓ separated by a distance $b >> \ell$, the focussing function for rays originating in one star and passing through the other is

$$\alpha \sim (b/\ell)(m/\ell) \sim 10^{-6}\ (b/\ell)(m/M_\odot)(\ell/R_\odot)^{-1} . \qquad (2.5.14)$$

This can be $\gtrsim 1$ for sufficiently large separations b.

The small-focussing and short-Shapiro-time-delay assumption force one to modify the error estimates (2.5.11). The correct error estimates are

$$(\text{fractional errors in } {}_2\bar{h}^{\mu\nu}) \sim \text{maximum } (\varepsilon^2, \varepsilon\alpha, \varepsilon\Delta t_s/\lambda\!\!\!^{-}), \qquad (2.5.15a)$$

$$(\text{fractional errors in } {}_2\bar{h}^{TT}_{jk}) \sim$$

$$\sim \frac{\text{maximum } (\varepsilon^2,\ \varepsilon\alpha,\ \varepsilon\Delta t_s/\lambda\!\!\!^{-})}{|\bar{h}^{TT}_{jk}/\bar{h}^{00}|} . \qquad (2.5.15b)$$

The "rules" for using the Post-Linear Wave-Generation Formalism are as follows: (1) Express the Post-Linear stress-energy tensor $T^{\mu\nu}$ in terms of the nongravitational variables of the specific system being analyzed, and in terms of the first-order gravitational field ${}_1\bar{h}^{\mu\nu}$. Do so in the manner of general relativity, up to fractional errors of $O(\varepsilon^2)$. (2) Solve the coupled Post-Linear equations of motion (2.4.4a) and Post-Linear gravitational field equations (2.4.4c) for the evolution of the source and its

first-order gravitational field ${}_1\bar{h}^{\mu\nu}$. (3) Evaluate the integrals (2.5.3)-(2.5.8) for field points in the local wave zone to determine th five pieces of the second-order gravitational field ${}_2\bar{h}^{\mu\nu}$. Project out the gravitational-wave field $\bar{h}^{TT}_{jk}$ in the local wave zone using equation (1.3.3) (5) Check, using equation (2.5.15b), that the errors in the wave field are acceptably small.

2.5.1. Sample Applications

The Post-Linear Wave-Generation Formalism can be applied, with good accuracy, to most astrophysical systems. For example, it can be applied to the two systems of §2.3.2, in which Linearized Theory was a failure: pulsations of a noncompact star, and near-encounters between fast-moving stars.

Actually, for the stellar pulsation problem -- and for any other problem characterized by slow internal motions (speeds much less than that of light) -- one need not resort to the Post-Linear Formalism in its full complexity. Rather, one can apply a slow-motion variant of the formalism (§2.6 below).

However, for systems with fast motion and weak but significant internal gravity, the Post-Linear Formalism is the simplest formalism that will do the full job. An example is "gravitational Bremsstrahlung radiation" produced in a near-encounter between nearly Newtonian stars ($M/R \ll 1$), which fly past each other with a relative velocity $v \gtrsim 0.1$ x (speed of light). This example is treated in detail by Kovács and Thorne (1977a,b). [Limiting cases and special features of the Bremsstrahlung problem are actually amenable to other techniques, which I do not review in these lectures. These include the Feynman-diagram method (originally designed for quantum gravity problems, but also applicable to classical problems; see Feynman 1963); the method of virtual quanta (Matzner and Nutku 1974); the method of Green's functions for the linearized Schwarzschild metric (a special case of the perturbation methods of §5 below; Peters 1970); and a new colliding-plane-waves tecnique (D'Eath 1977)].

2.6. The Slow-Motion Limit of the Post-Linear Formalism: Newtonian Theory [Order (2,1)] and Quadrupole-Moment Formalism [Order (2,2)]

Most astrophysical systems have not only weak internal gravity, $\varepsilon \ll 1$, but also slow internal motions and weak internal stresses. For such systems the Post-Linear Formalism of §2.4, which has order (2,1), reduces to the Newtonian Theory of Gravity; and the Post-Linear Wave-Generation Formalism of §2.5, which has order (2,2), reduces to the Quadrupole-Moment Wave-Generation Formalism. I shall briefly summarize these formalisms and their realms of validity. For details of the slow-motion transition from the Post-Linear Formalism to these formalisms, see TK. For derivations of these formalisms assuming slow motion from the beginning see, e.g., § 104 of Landau and Lifshitz (1962) or chapter 36 of MTW.

Realm of validity: One can characterize a slow-motion, weak-field system by the following parameters:

$M \equiv$ (mass of system)

$L \equiv$ (size of system)

$\lambda\!\!\!^{-} \equiv$ (characteristic timescale of system) = (reduced wavelength of radiation)

$v \equiv (|T^{0j}|/T^{00})_{max}$ = (maximum internal velocity) (2.6.1)

$S \equiv (|T^{ij}|/T^{00})_{max}$ = (maximum of stress/density)

$\varepsilon \equiv$ (magnitude of internal gravity) $\sim M/L$.

For a static system $\lambda\!\!\!^{-}$ is not zero; rather, it is the characteristic timescale for perturbations against which one hopes the system is stable (e.g., for small-amplitude pulsations, if the system is a static star).

Newtonian Gravitation Theory and the Quadrupole-Moment Formalism require for their validity the following conditions:

Weak gravity: $\varepsilon \ll 1$ (2.6.2a)

Slow motion: $L/\lambda\!\!\!^{-} \ll 1$, which implies $v \ll 1$ (2.6.2b)

Small Stresses: $S^2 \ll 1$, (2.6.2c)

System not near marginal stability:
$$\begin{cases} \omega^2 \equiv (1/\lambda\!\!\!^{-})^2 \gg (M/L)(M/L^3) \\ \omega^2 \equiv (1/\lambda\!\!\!^{-})^2 \gg S^2 (S/L)^2 \end{cases} \quad (2.6.2d)$$

[If the last condition is violated, then post-Newtonian gravitational forces may affect significantly the stability of the system and its motion; cf. Chandrasekhar (1964), or Boxes 24.2 and 26.2 of MTW.]

Newtonian Theory Summarized: Newtonian Theory describes gravity by a scalar gravitational potential U, and describes matter by its mass density ρ, its velocity v_j, and the stress t_{jk} measured in its rest frame. The equations of the theory are: (1) conservation of mass

$$\rho_{,t} + (\rho v_j)_{,j} = 0; \tag{2.6.3a}$$

(2) "Euler" equations of motion

$$\rho(v_{j,t}+v_{j,k}v_k) = \rho U_{,j} - t_{jk,k}; \tag{2.6.3b}$$

(3) gravitational field equation

$$U_{,jj} = -4\pi\rho, \tag{2.6.3c}$$

which has the solution

$$U(\underset{\sim}{x},t) = \int \frac{\rho(\underset{\sim}{x}',t)}{\left|\underset{\sim}{x} - \underset{\sim}{x}'\right|} \, d^3x'. \tag{2.6.3d}$$

For further details see, e.g., §39.7 of MTW.

Quadrupole-Moment Formalism Summarized: To calculate the gravitational waves from a nearly Newtonian system, one can proceed as follows: (1) Specify, by an equation of state, or viscosity, or some other method, the dependence of the stresses t_{jk} on ρ, v_j, and other variables of the system. (2) Solve the Newtonian equations (2.6.3) to determine the structure and evolution of the system. (3) Calculate the gravitational-wave amplitude h^{TT}_{jk} in the local wave zone using the following formula:

$$h^{TT}_{jk} = \frac{2}{r}\left[\ddot{\mathcal{I}}_{jk}(t-r)\right]^{TT}. \tag{2.6.4}$$

Here $\mathcal{I}_{jk}(t)$ is the "reduced quadrupole moment" of the source at time t

$$\mathcal{I}_{jk}(t) = \int \rho(t,\underset{\sim}{x}') \left[x^{j'} x^{k'} - \frac{1}{3}\delta^{jk}(\underset{\sim}{x}')^2\right] d^3x'; \tag{2.6.5}$$

$\ddot{\mathcal{I}}_{jk}$ is the second time derivative of $\mathcal{I}_{jk}$; r is distance from the source to the field point; and "TT" denotes the transverse-

traceless projection process of equation (1.3.3). (4) Verify that the "conditions of validity" (2.6.2) are satisfied.

2.6.1. Sample Applications

The quadrupole-moment formalism has been used more widely in gravitational-wave calculations than any other wave-generation formalism. Examples of its application are: the waves emitted by quadrupole pulsations of a star [see, e.g., Wheeler (1966)]; the waves emitted by orbital motion of a binary star system [see, e.g., Peters and Mathews (1963)]; the gravitational bremsstrahlung radiation from low-velocity near-encounters of two stars [low velocity limit of the problem described in §2.5.1; see, e.g., Ruffini and Wheeler (1971)]; and the "tidal gravitational radiation" produced when a Newtonian star is tidally deformed by passage near a black hole or other star [Mashoon (1973), Turner (1977)].

2.7 The Post-Newtonian Wave-Generation Formalism [Order (3,3)]

For systems with slow, but not extremely slow motions (e.g., $v \sim 0.3$), and weak, but not extremely weak gravity (e.g., $\varepsilon \backsim 0.1$), one needs a gravitational-wave formalism of higher accuracy than the Quadrupole-Moment Formalism. More specifically, one needs a formalism that takes account of Post-Newtonian effects in the structure and evolution of the system.

Epstein and Wagoner (1975) have devised such a formalism. Like all Post-Newtonian formalisms, it is valid only for gravitationally bound, or nearly gravitationally bound systems -- i.e., for systems with

$$v^2 \lesssim \varepsilon, \qquad S^2 \lesssim \varepsilon \tag{2.7.1a}$$

[cf. eq. (2.6.1)]; and it requires moderately weak internal gravity

$$\varepsilon \lesssim 0.1. \tag{2.7.1b}$$

The Epstein-Wagoner "Post-Newtonian Wave-Generation Formalism" has the following structure: (1) First one calculates the structure and evolution of the source using a Post-Newtonian Formalism of order $(n_T, n_h) = (3,2)$ -- e.g., the perfect-fluid Post-Newtonian formalism of Chandrasekhar (1965). (2) Then one evaluates the radiation field in the local wave zone using formulas to be given below (§4.5). [These formulas boost the order of the formalism up to $(n_T, n_h) = (3,3)$.] (3) Finally one verifies that the emitting system satisfies the Post-Newtonian validity conditions (2.7.1).

I delay presenting the Epstein-Wagoner equations until after I have treated multipole-moment formalisms, because Epstein and Wagoner utilize multipole moments in a nontrivial fashion.

2.7.1. Sample Applications

The Epstein-Wagoner formalism is being used in two different ways by the Stanford relativity group: (i) They use it to study deviations from the quadrupole-moment formalism for systems with moderately strong internal gravity or moderately high velocities or both [e.g. for moderately relativistic binary systems, and for stellar flybies at moderately high velocity (gravitational bremsstrahling); see Wagoner and Will (1976)]. (ii) They use it to seek insight into highly relativistic situations [e.g. waves from the collapse of a rotating star to form a black hole; Epstein (1976), Wagoner (1977)], where one can hope for, but is not guaranteed, more reliable results than from the quadrupole-moment formalism.

3. MULTIPOLE-MOMENT ANALYSIS OF THE RADIATION FIELD

3.1. Gravitational-Wave Field

I now wish to turn attention from weak-field formalisms to slow-motion formalisms. But before doing so, I must digress into a topic which is needed as an underpinning for the slow-motion discussion. This topic is the multipole structure of the gravitational-wave field for any isolated source. A number of researchers (not including me) have made major contributions to this subject; see end of §3.3 for references. Recently, I have taken the various contributions, have exhibited the relationships between some of their notations, and have tried to consolidate them into a single formalism in which (to me) the notation looks optimal. The following review of the subject is based on that work (Thorne 1977a--which I shall cite henceforth as TV, since it is Paper V in a series).

Consider the gravitational-wave field h^{TT}_{jk} in the local wave zone of an isolated source. The local wave zone is so constructed [eqs. (1.2.2)-(1.2.5)] that in it one can regard the background on which the waves propagate as completely flat. Consequently, the waves have the radially-propagating, flat-space form

$$h^{TT}_{jk} = r^{-1} A_{jk}(t - r,\theta,\phi) \qquad (3.1.1)$$

where A_{jk} is a symmetric, transverse, traceless amplitude that depends on retarded time $u = t - r$ and on angular location (θ,ϕ) around the source. The fact that A^{TT}_{jk} is symmetric,

transverse, and traceless is embodied in the algebraic relations

$$\text{Symmetry:} \quad A_{jk} = A_{kj}, \tag{3.1.2a}$$

$$\text{Transverse:} \quad A_{jk}n_k = 0, \text{where } \underset{\sim}{n} \equiv \underset{\sim}{x}/r = \text{unit radial vector}, \tag{3.1.2b}$$

$$\text{Traceless:} \quad A_{jj} = 0. \tag{3.1.2c}$$

It is often useful to expand the wave amplitude A_{jk} in spherical harmonics, and to identify the coefficients of the expansion as the multipole moments of the radiation field. In these lectures I shall follow the notation of Sachs (1961), Pirani (1965), MTW (Chap. 36), and TV (Thorne 1977a), which uses symmetric, trace-free tensors to represent spherical harmonics. The formalism of symmetric, trace-free tensors is described briefly in the next section.

3.2. Symmetric, Trace-Free Tensors and Spherical Harmonics

Let $\underline{A}$ be a symmetric, constant (position-independent) tensor of rank ℓ defined in 3-dimensional, flat space. We shall denote its components in a Cartesian coordinate system by $A_{k_1\ldots k_\ell}$, and the components of its completely symmetric part by $A_{(k_1\ldots k_\ell)}$:

$$A_{(k_1\ldots k_\ell)} \equiv \frac{1}{\ell!} \sum_{\pi} A_{k_{\pi(1)}\ldots k_{\pi(\ell)}}. \tag{3.2.1}$$

Here the summation goes over all $\ell!$ permutations, π, of $1\ldots\ell$. We shall denote the trace-free, symmetric part of $\underline{A}$ by the corresponding capital script letter, $\mathcal{A}$:

$$\mathcal{A}_{k_1\ldots k_\ell} = \sum_{n=0}^{[\ell/2]} (-1)^n \frac{\ell!(2\ell-2n-1)!!}{(\ell-2n)!(2\ell-n)!!(2n)!!} \times \tag{3.2.2}$$

$$\times \delta_{(k_1k_2} \cdots \delta_{k_{2n-1}k_{2n}} S_{k_{2n+1}\ldots k_\ell)j_1j_1\cdots j_nj_n}.$$

Here δ_{ab} = (Kronecker delta) are the components of the metric, $[\ell/2]$ means the largest integer less than or equal to $\ell/2$, $S_{k_1\ldots k_\ell} \equiv A_{(k_1\ldots k_\ell)}$, and the "double factorial" $(2n)!!$ is defined by (3.2.8) below.

Consider the set of all symmetric, trace-free tensors of rank ℓ ("STF-ℓ tensors"). The STF-ℓ tensors generate an irreducible representation of the rotation group, of weight ℓ. Hence, there

exists a one-to-one mapping between them and the spherical harmonics of order ℓ. To exhibit that mapping we introduce the unit radial vector $\underset{\sim}{n}$ of our Euclidean space, and we express its Cartesian components in terms of spherical coordinates, θ and ϕ :

$$\underset{\sim}{n}=\underset{\sim}{x}/|\underset{\sim}{x}|, \quad n_1+in_2=\sin\theta\, e^{i\phi}, \quad n_3=\cos\theta. \tag{3.2.3}$$

From the usual expression for the spherical harmonic $Y^{\ell m}(\theta,\phi)$ in terms of $\cos\theta$ and $\sin\theta\, e^{i\phi}$, it is easy to show [see pp. 289-290 of Pirani (1965) or see TV] that

$$Y^{\ell m}(\theta,\phi)=\mathcal{Y}^{\ell m}_{k_1\ldots k_\ell}\, n_{k_1}\ldots n_{k_\ell}. \tag{3.2.4}$$

Here $\mathcal{Y}^{\ell m}_{k_1\ldots k_\ell}$ is the following STF-ℓ tensor

$$\mathcal{Y}^{\ell m}_{k_1\ldots k_\ell} \equiv C^{\ell m} \sum_{j=0}^{[(\ell-m)/2]} a^{\ell m j}(\delta^1_{k_1}+i\delta^2_{k_1})\ldots(\delta^1_{k_m}+i\delta^2_{k_m})\cdot$$

$$\cdot\,\delta^3_{k_{m+1}}\ldots\delta^3_{k_{\ell-2j}}(\delta^{a_1}_{k_{\ell-2j+1}}\delta^{a_1}_{k_{\ell-2j+2}})\ldots(\delta^{a_j}_{k_{\ell-1}}\delta^{a_j}_{k_\ell}) \text{ for } m\geq 0, \tag{3.2.5}$$

$$\mathcal{Y}^{\ell m}_{k_1\ldots k_\ell} \equiv (-1)^m(\mathcal{Y}^{\ell\,-m}_{k_1\ldots k_\ell})^* \quad \text{for} \quad m<0.$$

$$C^{\ell m}\equiv(-1)^m\left[\frac{2\ell+1}{4\pi}\,\frac{(\ell-m)!}{(\ell+m)!}\right]^{1/2}, \quad a^{\ell m j}\equiv\frac{(-1)^j(2\ell-2j)!}{2^\ell j!(\ell-j)!(\ell-m-2j)!}.$$

Here $[(\ell-m)/2]$ means the largest integer less than or equal to $(\ell-m)/2$, and * means complex conjugate.

Henceforth, to simplify notation, when we encounter a sequence of many (say ℓ) indices on a tensor we shall denote it as follows

$$S_{A_\ell} \equiv S_{a_1\ldots a_\ell}. \tag{3.2.6a}$$

Similarly we shall abbreviate the tensor product of ℓ unit radial vectors by

$$N_{B_\ell} \equiv n_{b_1}\ldots n_{b_\ell}. \tag{3.2.6b}$$

In this abbreviated notation equation (3.2.4) says

$$Y^{\ell m}(\theta,\phi) = \mathcal{Y}^{\ell m}_{K_\ell} N_{K_\ell} . \qquad (3.2.4')$$

Thus, capital subscript letters denote sequences of lower-case subscript indices; and the number of indices in a sequence is denoted by a subscript to the capital subscript.

The tensors $\mathcal{Y}^{\ell m}_{K_\ell}$ with $-\ell \le m \le +\ell$ serve two roles: First, they generate the spherical harmonics of order ℓ (eq. 3.2.4'). Second, they form a basis for the $(2\ell+1)$-dimensional vector space of STF-ℓ tensors; i.e., any STF-ℓ tensor $\mathcal{F}$ can be expanded as

$$\mathcal{F}_{K_\ell} = \sum_{m=-\ell}^{\ell} F^{\ell m} \mathcal{Y}^{\ell m}_{K_\ell} . \qquad (3.2.7a)$$

The tensor components $\mathcal{F}_{K_\ell}$ are real if and only if $F^{\ell -m} = (-1)^m F^{\ell m}$. The expansion (3.2.7a) can be inverted as follows (see, e.g., TV)

$$F^{\ell m} = 4\pi \frac{\ell!}{(2\ell+1)!!} \mathcal{F}_{K_\ell} \mathcal{Y}^{\ell m}_{K_\ell} \qquad (3.2.7b)$$

where the "double factorial" is defined by

$$n!! \equiv n(n-2)(n-4)\ldots(2 \text{ or } 1). \qquad (3.2.8)$$

In practical calculations one can use spherical harmonics and STF-ℓ tensors interchangeably: Consider any sphere centered on the coordinate origin, and on that sphere consider any scalar function $f(\theta,\phi)$. One can expand $f(\theta,\phi)$ in spherical harmonics with complex-number expansion coefficients

$$f(\theta,\phi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} F^{\ell m} Y^{\ell m}(\theta,\phi); \qquad (3.2.9a)$$

alternatively, one can expand it in powers of the unit radial vector $\underset{\sim}{n}$, with coefficients that are STF-ℓ tensors

$$f(\theta,\phi) = \sum_{\ell=0}^{\infty} \mathcal{F}_{K_\ell} N_{K_\ell} . \qquad (3.2.9b)$$

The expansion coefficients of the two schemes are related by equations (3.2.7a,b).

Similar expansions can be made for vector fields and tensor fields. The resulting formalism, which relates vector and tensor spherical harmonics to STF-ℓ tensors, is presented in TV.

3.3. Multipole Expansion of Gravitational-Wave Field

The vacuum Einstein field equations permit the gravitational-wave amplitude A_{ij} (eq. 3.1.1) to be any symmetric, transverse, traceless function of retarded time $t-r$ and angle. One can show [see, e.g., TV] that this fact allows A_{ij} to be expanded in a multipole series of the form

$$h_{ij}^{TT} = r^{-1}A_{ij}$$

$$= \left[\sum_{\ell=2}^{\infty} (4/\ell!) r^{-1} \, {}^{(\ell)}\mathcal{I}_{ijA_{\ell-2}}(t-r)\, N_{A_{\ell-2}} \right. \qquad (3.3.1)$$

$$\left. + \sum_{\ell=2}^{\infty} [8\ell/(\ell+1)!] r^{-1} \varepsilon_{pq(i} \, {}^{(\ell)}\mathcal{S}_{j)pA_{\ell-2}}(t-r) n_q N_{A_{\ell-2}} \right]^{TT}.$$

Here the normalizations (factors of 4, 8, ℓ, etc.) have been chosen to make equations (4.4.3) look simple; ε_{abc} is the Levi-Civita tensor; parentheses around tensor indices denote symmetrization; ${}^{(\ell)}\mathcal{I}_{A_\ell}(t-r)$ and ${}^{(\ell)}\mathcal{S}_{A_\ell}(t-r)$ are the trace-free, symmetric coefficients of the expansion evaluated at retarded time $t-r$; a prefix superscript in parentheses, e.g., ${}^{(k)}\mathcal{S}(u)$, means that the quality is to be differentiated k times with respect to its argument u

$${}^{(k)}\mathcal{S}(u) \equiv d^k\mathcal{S}/du^k; \qquad (3.3.2)$$

and TT means that the transverse, traceless part is to be taken (eq. 1.3.3).

The coefficients ${}^{(\ell)}\mathcal{I}_{A_\ell}$ and ${}^{(\ell)}\mathcal{S}_{A_\ell}$ of the spherical harmonic expansions--integrated ℓ times (superscript ℓ removed)--are called the "mass" and "current" moments of the radiation field:

$$\mathcal{I}_{ab} \equiv \text{(mass quadrupole moment)}, \quad \mathcal{S}_{ab} \equiv \text{(current quadrupole moment)},$$

$$\mathcal{I}_{abc} \equiv \text{(mass octupole moment)}, \quad \mathcal{S}_{abc} \equiv \text{(current octupole moment)}, \qquad (3.3.3)$$

$\mathcal{I}_{A_\ell} \equiv$ (mass ℓ-pole moment), $\mathcal{S}_{A_\ell} \equiv$ (current ℓ-pole moment).

They are trace-free, symmetric tensors which depend on retarded time.

Notice that the mass quadrupole part of the radiation field (3.3.1) has a form familiar from the "Quadrupole Moment Formalism"

$$(h^{TT}_{jk})_{\text{mass quadrupole}} = 2r^{-1}\,\ddot{\mathcal{I}}^{TT}_{jk} \qquad (3.3.4)$$

[cf. eq. (2.6.4)] . Also notice that the mass-multipole moments produce a radiation field with "electric-type parity" (also called "even-type parity"), $\pi = (-1)^\ell$; while the current-multipole moments produce a radiation field with "magnetic-type" ("odd-type") parity, $\pi = (-1)^{\ell+1}$

If the radiation field is known, one can project out its multipole moments using the following integrals over a sphere of constant (t-r):

$$^{(\ell)}\mathcal{I}_{B_\ell} = \left[\begin{array}{l}\text{Symmetric, trace-}\\ \text{free part of}\end{array}\right]\left\{\frac{\ell(\ell-1)(2\ell+1)!!}{2(\ell+1)(\ell+2)}\,\frac{r}{4\pi}\int h^{TT}_{b_1b_2}\,n_{b_3}\cdots n_{b_\ell}\,d\Omega\right\}, \qquad (3.3.5a)$$

$$^{(\ell)}\mathcal{S}_{B_\ell} = \left[\begin{array}{l}\text{Symmetric, trace-}\\ \text{free part of}\end{array}\right]\left\{\frac{(\ell-1)(2\ell+1)!!}{4(\ell+2)}\,\frac{r}{4\pi}\int \varepsilon_{b_1jk}\,n_j\,h^{TT}_{kb_2}\,n_{b_3}\cdots n_{b_\ell}\,d\Omega\right\} \qquad (3.3.5b)$$

Multipole expansions of a generic radiation field have been given in various notations by Sachs (1961), Pirani (1965), Mathews (1962), Janis and Newman (1965), Bonnor and Rotenberg (1966), and Campbell and Morgan (1971). I like the above STF notation -- which is due to Sachs (1961), and Pirani (1965)-- because it ties in nicely with the theory of slow-motion sources (see below) and with the standard form of the quadrupole-moment formalism. Elsewhere (TV) I exhibit the relationship between the above STF expansions and expositions that use other conventions for spherical harmonics.

3.4. Multipole Expansion of Energy, Momentum, and Angular Momentum in Waves

The energy and linear momentum carried off by the radiation field (3.3.1) are most easily evaluated using the Isaacson (1968) stress-energy tensor for gravitational waves

$$T^{GW}_{\alpha\beta} = (1/32\pi)\left\langle h^{TT}_{jk,\alpha}\,h^{TT}_{jk,\beta}\right\rangle \qquad (3.4.1)$$

(cf. MTW, §§35.7 and 35.15). Here the brackets, $\langle\ \rangle$, denote an average over several wavelengths. The power radiated into a unit solid angle about the radial, $\underset{\sim}{n}$, direction is

$$(dE/d\Omega dt) \equiv -r^2 n_j T^{GW}_{j0} = +r^2\ T^{GW}_{00}\ ; \tag{3.4.2}$$

and the total power radiated is the integral of this power over a sphere lying in the local wave zone:

$$dE/dt = \int (dE/d\Omega dt) d\Omega. \tag{3.4.3}$$

By inserting the gravitational-wave field (3.3.1) into equations (3.4.1)-(3.4.3) and integrating one obtains (see TV)

$$\frac{dE}{dt} = \sum_{\ell=2}^{\infty} \frac{(\ell+1)(\ell+2)}{(\ell-1)\ell} \frac{1}{\ell!(2\ell+1)!!} \left\langle {}^{(\ell+1)}\mathcal{I}_{A_\ell}\ {}^{(\ell+1)}\mathcal{I}_{A_\ell} \right\rangle \tag{3.4.4}$$

$$+ \sum_{\ell=2}^{\infty} \frac{4\ell(\ell+2)}{(\ell-1)} \frac{1}{(\ell+1)!(2\ell+1)!!} \left\langle {}^{(\ell+1)}\mathcal{S}_{A_\ell}\ {}^{(\ell+1)}\mathcal{S}_{A_\ell} \right\rangle .$$

The waves carry linear momentum out radially--and, as with any locally plane-fronted radiation field, the magnitude of their momentum flux is the same as that of their energy flux:

$$(dP_j/d\Omega dt) \equiv r^2\ n_k\ T^{GW}_{kj} = r^2\ n_j T^{GW}_{00} = n_j\ (dE/d\Omega dt) \tag{3.4.5}$$

(cf. eq. 35.77j of MTW). The total linear momentum carried off by the waves,

$$dP_j/dt = \int (dP_j/d\Omega dt) d\Omega, \tag{3.4.6}$$

can be evaluated by inserting the wave field (3.3.1) into the above equations and integrating:

$$\frac{dP_j}{dt} = \sum_{\ell=2}^{\infty} \left\{ \frac{2(\ell+2)(\ell+3)}{\ell(\ell+1)!(2\ell+3)!!} \left\langle {}^{(\ell+2)}\mathcal{I}_{jA_\ell}\ {}^{(\ell+1)}\mathcal{I}_{A_\ell} \right\rangle + \right.$$

$$+ \frac{8(\ell+3)}{(\ell+1)!(2\ell+3)!!} \left\langle {}^{(\ell+2)}\mathcal{S}_{jA_\ell}\ {}^{(\ell+1)}\mathcal{S}_{A_\ell} \right\rangle + \tag{3.4.7}$$

$$\left. + \frac{8(\ell+2)}{(\ell-1)(\ell+1)!(2\ell+1)!!} \left\langle \varepsilon_{jpq}\ {}^{(\ell+1)}\mathcal{I}_{pA_{\ell-1}}\ {}^{(\ell+1)}\mathcal{S}_{qA_{\ell-1}} \right\rangle \right\}.$$

Note, that, whereas the energy radiated (eq. 3.4.4) involves a beating of each multipole moment against itself, the linear momentum radiated (eq. 3.4.7) involves a beating of "adjacent" multipole moments.

The dependence of the wave amplitude, h^{TT}_{ij}, on angle causes its wave fronts to be not quite precisely spherical--and thereby enables the waves to carry off angular momentum. One might hope that the angular momentum loss could be calculated by integrating $\epsilon_{jab}\, x_a\, T^{GW}_{b0}$ over a sphere surrounding the source. Unfortunately, such a procedure fails--for this reason: the averaging process that underlies equation (3.4.1) for $T^{GW}_{\alpha\beta}$ treats as zero the "tiny corrections" which die out as $1/r^3$. However, it is precisely the $1/r^3$ part that carries off the angular momentum. I was vaguely aware of this fact when writing the relevant sections of MTW, but was not sufficiently certain to spell it out explicitly. Subsequently Bryce DeWitt (1971) derived a simple, correct expression for the flux of angular momentum:

$$\frac{dS_j}{dt} = \frac{1}{16\pi}\int \epsilon_{jpq}\, x_p \left\langle \left(h^{TT}_{qa}\, {}^{(1)}h^{TT}_{ab}\right)_{,b} - \frac{1}{2} h^{TT}_{ab,q}\, {}^{(1)}h^{TT}_{ab} \right\rangle r^2\, d\Omega. \tag{3.4.8}$$

A word of interpretation is needed. This equation is correct only if the integral is evaluated in the asymptotic rest frame of the source. (Similarly for all previous formulae in this section.) As the source's linear momentum changes (eq. 3.4.7), its asymptotic rest frame gradually changes; and one must gradually change the reference frame in which one evaluates (3.4.8) (and all previous integrals).

Return to equation (3.4.8). By inserting the wave field (3.3.1) and integrating, one obtains

$$\frac{dS_j}{dt} = \sum_{\ell=2}^{\infty} \frac{(\ell+1)(\ell+2)}{(\ell-1)\ell!\,(2\ell+1)!!} \left\langle \epsilon_{jpq}\, {}^{(\ell)}\mathcal{I}_{pA_{\ell-1}}\, {}^{(\ell+1)}\mathcal{I}_{qA_{\ell-1}} \right\rangle \tag{3.4.9}$$

$$+\sum_{\ell=2}^{\infty} \frac{4\ell^2(\ell+2)}{(\ell-1)(\ell+1)!(2\ell+1)!!} \left\langle \epsilon_{jpq}\, {}^{(\ell)}\mathcal{S}_{pA_{\ell-1}}\, {}^{(\ell+1)}\mathcal{S}_{qA_{\ell-1}} \right\rangle.$$

Notice that this angular momentum radiated, like the energy radiated (eq. 3.4.4), involves a beating of each multipole with itself.

3.5. Discussion of the Multipole Formalism

Of what use are the above formulas? I view them as tools to to be used in studying the generation of gravitational waves from explicit sources. Given a source, one identifies the local wave zone. Using any technique one can dream up (and this is the tough part of the analysis!), one calculates the time-changing multipole moments of the source. One then plugs into the above formulas to get the radiation field and the rate it carries off energy, momentum, and intrinsic angular momentum.

As an alternative application, one can calculate the radiation field of a given source (the tough task; above formulas not necessarily useful); one can use equations (3.3.5) to resolve it into multipole pieces; and one can then use the other formulas above to read off the energy, momentum, and angular momentum radiated.

Once the radiation field h_{ij}^{TT} is known in the local wave zone, one can propagate it on outwards to Earth using the propagation equation of the geometric-optics formalism (§7 below). For typical situations the wave form (3.3.1) will remain highly accurate all the way to Earth, except for uninteresting phase shifts caused by the waves' self energy and by various masses present in the Universe, and except for changes in the weakest of the multipole fields caused by nonlinear beating together of stronger multipole fields.

4. SLOW-MOTION FORMALISMS

4.1. Overview

Turn attention now from weak-field systems (§2) and arbitrary systems (§3) to slow-motion systems--i.e., systems for which

$$\lambda\!\!\!^{-} \equiv (\text{reduced wavelength of radiation}) >> L \equiv (\text{size of source}). \quad (4.1.1)$$

Thorne (1977b--cited henceforth as TVI) has derived a slow-motion multipole-moment formalism for calculating gravitational-wave generation by such systems. This slow-motion formalism generalizes the quadrupole-moment formalism of §2.6 to include all multipole moments, and to encompass sources with arbitrarily strong internal gravity; cf. Table 1.

The slow-motion formalism takes, as its starting point, a multipole analysis of the external gravitational field of a stationary source. That stationary multipole analysis is described in §4.2; and the slow-motion wave-generation formalism built on it is described in §4.3. Then, in §4.4 the weak-field limit of

the slow-motion formalism is described. Throughout § 4 we shall use subscripted coordinate indices: $x_\alpha \equiv x^\alpha$.

4.2. Multipole Moments of a Stationary Source

Consider a stationary system residing in asymptotically flat spacetime and surrounded by vacuum. It is well known[4] that the vacuum geometry of spacetime surrounding such a stationary system is uniquely determined by two families of time-independent multipole moments: the mass moments, and the current moments. The object of this section is to give a definition of these moments (i.e., a set of conventions for them) that will tie in simply with radiation theory. More specifically, if the stationary source is set into very slow motion (§4.3 below), its moments as defined here (§4.2) will be identical to those which characterize the emitted radiation (§3 above).

4.2.1. ACMC-N Coordinate Systems

As a tool in defining the moments of a stationary source, we introduce a special class of coordinate systems: A coordinate system will be called "asymptotically Cartesian and mass centered to order N" ("ACMC-N") if and only if its metric coefficients have the following form. Expand them in inverse powers of "radius"

$$r \equiv [(x_1)^2 + (x_2)^2 + (x_3)^2]^{1/2} \quad . \tag{4.2.1}$$

The coefficients must be time-independent (so the Killing vector is $\partial/\partial x_0$); the leading term must be the Minkowskii metric; and the remaining terms must have the form

$$g_{\alpha\beta} = \eta_{\alpha\beta} + \frac{A^0_{\alpha\beta}\,(\underset{\sim}{n})}{r} + \frac{A^1_{\alpha\beta}\,(\underset{\sim}{n})}{r^2} + \ldots + \frac{A^{N+1}_{\alpha\beta}(\underset{\sim}{n})}{r^{N+2}} + [\text{terms that die out faster than } 1/r^{N+2}], \tag{4.2.2}$$

where

$$\underset{\sim}{n} = \underset{\sim}{x}/r = (x_j/r)\,(\partial/\partial x_j). \tag{4.2.3}$$

Expand each of the coefficients $A^P_{\alpha\beta}\,(\underset{\sim}{n})$ in spherical harmonics.

[4] See, e.g., van der Burg (1968), Geroch (1970), Clarke and Sciama (1971).

$A^0_{\alpha\beta}$ and A^1_{00} must involve only harmonics of order 0.

A^1_{0j} and A^1_{jk} must involve only harmonics of order 0 and 1.

(4.2.4)

$A^2_{\alpha\beta}$ must involve only harmonics of order 0,1, and 2.

$A^3_{\alpha\beta}$ must involve only harmonics of order 0,1,2, and 3

.

.

.

(4.2.4 con'd)

$A^N_{\alpha\beta}$ must involve only harmonics of order 0,1,...,N.

Note the absence of a dipole in A^1_{00}; that is what makes the coordinates "mass centered".

That such a coordinate system always exists is proved in TVI by showing that "deDonder coordinates," with appropriate specialization of the gauge, are ACMC-∞.

4.2.2. Multipole Moments

Given an ACMC-N coordinate system, we define the multipole moments of order $\ell \leq N+1$ in the following manner: The ℓ-order part of the $1/r^{\ell+1}$ piece of g_{00} is the mass ℓ-pole moment; and the ℓ-order, "odd-type parity" part of the $1/r^{\ell+1}$ piece of g_{0j} is the current ℓ-pole moment. More specifically,

$$A^{\ell}_{00} = [2(2\ell-1)!!/\ell!]\, \mathcal{I}_{A_\ell}\, N_{A_\ell} + (\text{parts of order} < \ell) \qquad (4.2.5a)$$

$$A^{\ell}_{0j} = [-4\ell(2\ell-1)!!/(\ell+1)!]\, \varepsilon_{jpq}\, \mathcal{S}_{pA_{\ell-1}}\, n_q N_{A_{\ell-1}} \qquad (4.2.5b)$$

$$+ (\text{"even-type" parts of order } \ell) + (\text{parts of order} < \ell)$$

(4.2.5b)

where $\mathcal{I}_{A_\ell}$ is the ℓ-pole mass moment, and $\mathcal{S}_{A_\ell}$ is the ℓ-pole current moment.

In TVI the self-consistency of this definition is proved; i.e., it is shown that every ACMC-N coordinate system gives the same result for the moments of order $\ell \leq N + 1$.

This definition of multipole moments is far less elegant than definitions given by other researchers.[4] However, the other definitions lead to moments which fail to tie in nicely with radiation theory. My moments must be nonlinear, algebraic functions of the moments defined elsewhere,[4] but I have not attempted to exhibit that relationship explicitly.

The above definition of multipole moments is embodied in the following formula for the metric coefficient of an ACMC-N coordinate system:

$$g_{00} = -1 + \frac{2\mathcal{I}}{r} + \frac{(0\text{-pole})}{r^2} + \frac{1}{r^3}\left[3\mathcal{I}_{ab}n_a n_b + (1\text{ pole}) + (0\text{ pole})\right.$$

$$+\ldots+\frac{1}{r^{\ell+1}}\left[\frac{2(2\ell-1)!!}{\ell!}\mathcal{I}_{A_\ell}N_{A_\ell} + (\ell\text{-1 pole})+\ldots+(0\text{ pole})\right]$$

$$+\ldots+\frac{1}{r^{N+1}}\left[\frac{2(2N-1)!!}{N!}\mathcal{I}_{A_N}N_{A_N} + (N\text{-1 pole})+\ldots+(0\text{ pole})\right]$$

$$\left.+\frac{1}{r^{N+2}}\left[\frac{2(2N+1)!!}{(N+1)!}\mathcal{I}_{A_{N+1}}N_{A_{N+1}} + (\text{poles with } \ell \neq N+1)\right]\right.$$

$$+\left[\text{terms that die out faster than } 1/r^{N+2}\right]. \qquad (4.2.6a)$$

$$g_{0j} = -\frac{1}{r^2}\left[2\varepsilon_{jpq}\mathcal{S}_p n_q + \left(\begin{array}{l}1\text{ pole part with}\\ \text{parity } \pi = -\end{array}\right) + (0\text{ pole})\right]$$

$$-\ldots-\frac{1}{r^{\ell+1}}\left[\frac{4\ell(2\ell-1)!!}{(\ell+1)!}\varepsilon_{jpq}\mathcal{S}_{pA_{\ell-1}}n_q N_{A_{\ell-1}} +\right.$$

$$\left.+\left(\begin{array}{l}\ell\text{ pole part with}\\ \text{parity } \pi = (-1)^\ell\end{array}\right) + (\ell\text{-1 pole}) + \ldots + (0\text{ pole})\right]$$

$$-\ldots-\frac{1}{r^{N+1}}\left[\frac{4N(2N-1)!!}{(N+1)!}\varepsilon_{jpq}\mathcal{S}_{pA_{N-1}}n_q N_{A_{N-1}} +\right.$$

$$\left.+\left(\begin{array}{l}N\text{ pole part with}\\ \text{parity } \pi = (-1)^N\end{array}\right) + (N\text{-1 pole}) + \ldots + (0\text{ pole})\right] +$$

$$- \frac{1}{r^{N+2}} \left[\frac{4(N+1)(2N+1)!!}{(N+2)!} \varepsilon_{jpq} \mathcal{S}_{pA_N} n_q N_{A_N} + \left\{ \begin{matrix} \text{N+1 pole part with} \\ \text{parity } \pi = (-1)^{N+1} \end{matrix} \right\} \right.$$

$$\left. + \text{(parts with } \ell \neq N+1) \right]$$

$$+ [\text{terms that die out faster than } 1/r^{N+2}] \quad . \qquad (4.2.6b)$$

$$g_{jk} = \delta_{jk} + \frac{\text{(0 pole)}}{r} + \frac{1}{r^2} [\text{(1 pole)} + \text{(0 pole)}] + \ldots$$

$$+ \frac{1}{r^{\ell+1}} [\text{(ℓ pole)} + \ldots + \text{(0 pole)}] + \frac{1}{r^{N+1}} [\text{(N pole)}$$

$$+ \ldots + \text{(0 pole)}] + \frac{1}{r^{N+2}} \left[\begin{matrix} \text{any angular} \\ \text{dependence} \end{matrix} \right]$$

$$+ [\text{terms that die out faster than } 1/r^{N+2}]. \qquad (4.2.6c)$$

4.2.3. Example: The Kerr Metric

As an example, consider the Kerr metric (external field of a stationary black hole). In Boyer-Lindquist coordinates, made quasi-Cartesian by defining

$$x_1 = r \sin\theta \cos\phi, \quad x_2 = r \sin\theta \sin\phi, \quad x_3 = r \cos\theta,$$

the nonzero metric coefficients are (cf. page 877 of MTW)

$$g_{00} = - \frac{r^2 + a^2 \cos^2\theta - 2Mr}{r^2 + a^2 \cos^2\theta} = -1 + \frac{2M}{r} - \frac{2Ma^2 \cos^2\theta}{r^3} + 0\left(\frac{1}{r^5}\right),$$

$$g_{0j} = - \frac{2\varepsilon_{jpq} J_p n_q}{r^2 + a^2 \cos^2\theta} = - \frac{2\varepsilon_{jpq} J_p n_q}{r^2} + 0\left(\frac{1}{r^4}\right), \qquad (4.2.7)$$

$$g_{jk} = \delta_{jk} + \frac{2M}{r} n_j n_k + \frac{1}{r^2} [\text{(0 pole)} + \text{(2 pole)}] + 0\left(\frac{1}{r^3}\right),$$

where

$$\underset{\sim}{J} \equiv Ma(\partial/\partial x_3). \qquad (4.2.8)$$

The presence of a quadrupole $1/r^2$ term in g_{jk} prevents this coordinate system from being ACMC-1; it is only ACMC-0. Hence, from the metric coefficients we can read off only the moments of order 0 and 1:

$\mathcal{I}$ ≡ (mass monopole moment; i.e., total mass-energy) = M;

$\mathcal{I}_a$ ≡ (mass dipole moment) = 0 [guaranteed to vanish because coordinates are mass centered];

$\underset{\sim}{\mathcal{S}}$ ≡ (current dipole moment; i.e., total angular momentum)

$$= Ma(\partial/\partial x_3) \quad . \qquad (4.2.9)$$

In order to deduce the higher-order moments one must perform a coordinate transformation that gets rid of the offending quadrupole term in g_{jk}. Such a transformation is exhibited in TVI.

4.3. The Slow-Motion, Multipole-Moment Formalism For Strong-Field Sources

Turn attention now from stationary sources to the slow-motion, multipole-moment formalism.

Consider a specific (but arbitrary) slow-motion, isolated system. If the system were not changing at all ($\lambda = \infty$), the metric everywhere outside the source would be stationary -- i.e., it would be describable by the stationary multipole-moment formalism of § 4.2. We can pass from that stationary metric to the true metric by gradually turning on the time dependence of the system's multipole moments -- i.e., by gradually letting λ decrease from infinity. Mathematically this is achieved by expanding the metric simultaneously in the two small dimensionless parameters R/r and R/λ, where R is the length scale that characterizes the weak-field, near-zone metric:

$$R \equiv \underset{\ell=1,2,\ldots}{\text{Maximum}} \left\{ \left| \frac{\mathcal{I}_{A_\ell}}{M} \right|^{1/\ell}, \left| \frac{\mathcal{S}_{A_\ell}}{M} \right|^{1/\ell} \right\} \qquad (4.3.1)$$

See TVI for details. For typical (not highly spherical) sources, R will be approximately equal to the size of the source L.

At zero order in R/λ, the R/r expansion of the metric must be formally identical to the general stationary expansion of § 4.2. But the time dependence of that "zero-order solution" will

generate corrections or order $(R/\lambda\!\!\!^{-})$, $(R/\lambda\!\!\!^{-})^2$, ... In TVI it is shown that those corrections are determined uniquely by the (time dependent) multipole moments, plus the demand that the near-zone metric match onto outoing waves. TVI also performs the match onto outgoing waves in the local wave zone, with the following result: The matching requires that the moments $\mathcal{I}_{A_\ell}$, $\mathcal{S}_{A_\ell}$ that appear in the zero-order [i.e., $(R/\lambda\!\!\!^{-})^0$] near-zone solutions [eqs. (4.2.6)] be identical to the moments that appear in the radiation field [eq. (3.3.1)] -- except for differences that are typically negligible:

$$|\delta\mathcal{I}_{A_\ell}| \lesssim MR^\ell(M/\lambda\!\!\!^{-}),\quad |\delta\mathcal{S}_{A_\ell}| \lesssim MR^\ell(M/\lambda\!\!\!^{-}). \qquad (4.3.2a)$$

For the most strongly radiating moment $\mathcal{M}_{A_\ell}$ (usually the mass quadrupole moment $\mathcal{I}_{A_2}$) the disagreement is

$$|\delta\mathcal{M}_{A_\ell}| \lesssim |\mathcal{M}_{A_\ell}|(M/\lambda\!\!\!^{-}). \qquad (4.3.2b)$$

This result leads to the following slow-motion formalism for calculating the generation of gravitational waves:
(1) Analyze the structure and evolution of the system in any convenient coordinates and by any fairly accurate approximation scheme. (2) From that analysis obtain an approximation to the external gravitational field which, at any instant, satisfies (to some degree of accuracy) the time-independent, vacuum Einstein field equations. (3) By transforming that external field to an ACMC coordinate system, read off its dominant multipole moments (the moments with the largest values of ${}^{(\ell)}\mathcal{I}_{A_\ell}$ or ${}^{(\ell)}\mathcal{S}_{A_\ell}$). (4) Plug those dominant moments into the gravitational-wave formulae of §3.

One attractive approximation scheme for use in steps (1) and (2) is the "instantaneous-gravity" approximation. In this scheme one sets to zero all time derivatives of the metric (but) not of the matter variables) when solving the Einstein field equations. This has the effect of removing all dynamical freedom from the gravitational field and making gravitational interactions within the source instantaneous rather than retarded. One automatically obtains an external gravitational field which satisfies the time-independent vacuum field equations; and, unless one has made a foolish choice of coordinates, the moments which one computes from that external field should contain errors no larger than $L/\lambda\!\!\!^{-}$. The radiation field computed from those moments will then have fractional errors $L/\lambda\!\!\!^{-}$ and $M/\lambda\!\!\!^{-}$. (Here L is the size of the source).

Typically the ℓ-pole mass and current moments will have magnitudes

$$\mathcal{I}_{A_\ell} \sim ML^\ell \ , \quad \mathcal{S}_{A_\ell} \sim M(L/\lambda\!\!\!^{-})L^\ell \ . \tag{4.3.3}$$

In this case their contributions to the radiation field [eq.(3.3.1)] will be

$$(h^{TT}_{jk})_{\text{mass } \ell\text{-pole}} \sim (M/r)\ (L/\lambda\!\!\!^{-})^\ell, \tag{4.3.4a}$$

$$(h^{TT}_{jk})_{\text{current } \ell\text{-pole}} \sim (M/r)\ (L/\lambda\!\!\!^{-})^{\ell+1} \ ; \tag{4.3.4b}$$

Since $L/\lambda\!\!\!^{-} \ll 1$ (slow-motion assumption), mass quadrupole radiation is usually far larger than the other multipoles. The current quadrupole field and mass octupole field are normally a factor $L/\lambda\!\!\!^{-}$ smaller. It is this fact which allows one to compute only the lowest few multipole moments when applying the slow-motion formalism.

In special cases (e.g., torsional oscillations of a neutron star) the mass quadrupole moment will vanish, or will be far smaller than its normal value. Then the current quadrupole or mass octupole radiation may dominate -- unless they, too, are abnormally small, allowing higher moments to make themselves felt.

Notice that, if one is interested in the linear momentum radiated by the source (eq. 3.4.7), one must compute not just the lowest significant multipole moment, but also the moments "adjacent" to it. Linear momentum is carried off only through the interference of adjacent multipole fields with each other.

4.3.1. A Sample Application

As a typical application of the slow-motion, multipole-moment formalism, consider a slowly rotating neutron star which is not quite axially symmetric. Ipser (1971) has formulated a general relativistic analysis of the interior of such a star, using the Regge-Wheeler (1957) formalism for small, strong-field deviations from spherical symmetry. Ipser's analysis shows how internal stresses, supported by the crystal structure of the star's mantle, maintain the star's deformation. It also gives formulas for the star's near-zone multipole moments in terms of the star's internal structure. By inserting those near-zone multipole moments into the slow-motion wave-generation formalism, one obtains the gravitational-wave field $\bar{h}^{TT}_{jk}$ produced by the star's rotation, and also the energy and angular momentum radiated. (Ipser computed the gravitational-wave field from his near-zone multipole moments by brute force, and discovered the then surprising result that it had the same form as in weak-field, slow-motion theory. That discovery was the original motivation for my constructing the above formalism.)

4.4 Weak-Field Limit of the Slow-Motion Formalism

For a slow-motion source ($L/\lambda\bar{}<<1$) with weak internal fields ($\varepsilon << 1$) and weak internal stresses ($S^2 << 1$), one can use Newtonian theory to analyze the interior region; cf. §2.6. Using Newtonian theory, one can express the multipole moments

$$\mathcal{I}_{A_\ell} \text{ and } \mathcal{S}_{A_\ell}$$

in terms of volume integrals over the source [see, eg., TV for detailed proof].

The multipole volume integrals must be performed in a mass-centered Cartesian coordinate system--i.e., in Cartesian coordinates with

$$\mathcal{I}_j(t) \equiv \int \rho(\underset{\sim}{x},t)\, x_j\, d^3x = 0. \tag{4.4.1}$$

[Note that such a coordinate system is automatically the rest frame of the source, since the time-derivative of equation (4.4.1) can be put in the form

$$\begin{aligned} 0 = d\,\mathcal{I}_j(t)/dt &= \int [\partial\rho(\underset{\sim}{x},t)/\partial t]\, x_j\, d^3x \\ &= \int \rho(\underset{\sim}{x},t)\, v_j d^3x = \text{(momentum of source)}. \end{aligned} \tag{4.4.2}$$

The third equality follows from the equation of mass conservation (2.6.3a) and an integration by parts.]

In a mass-centered Cartesian coordinate system, the volume integrals for $\mathcal{I}_{A_\ell}$ and $\mathcal{S}_{A_\ell}$ are

$$\mathcal{I}_{A_\ell} = \text{Symmetric trace-free part of } I_{A_\ell}, \tag{4.4.3a}$$

$$I_{A_\ell} = \begin{pmatrix}\ell\text{'th moment of}\\ \text{mass distribution}\end{pmatrix} = \int \rho x_{a_1} \cdots x_{a_\ell}\, d^3x; \tag{4.4.3b}$$

$$\mathcal{S}_{A_\ell} = \text{Symmetric trace-free part of } S_{A_\ell}, \tag{4.4.4b}$$

$$S_{A_\ell} = \begin{pmatrix}\ell\text{-1 moment of angular}\\ \text{momentum distribution}\end{pmatrix} \tag{4.4.4b}$$

$$= \int (\varepsilon_{a_1 jk} x_j \rho\, v_k)\, x_{a_2} \cdots x_{a_\ell}\, d^3x.$$

The "symmetric, trace-free part" can be computed from equations (3.2.1) and (3.2.2).

The volume integrals (4.4.3) and (4.4.4) are not precisely equal to the exact near-zone multipole moments of the source, of course. They contain errors of post-Newtonian order (cf. eq. 2.6.1):

$$(\delta \mathcal{I}_{A_\ell}) \text{ due to weak-field assumption} \sim ML^\ell[\varepsilon+(L/\lambda\!\!\!^{-})^2+S^2], \tag{4.4.5a}$$

$$(\delta \mathcal{S}_{A_\ell}) \text{ due to weak-field assumption} \sim M(L/\lambda\!\!\!^{-})L^\ell[\varepsilon+(L/\lambda\!\!\!^{-})^2+S^2]. \tag{4.4.5b}$$

These errors are in addition to the amounts (eqs. 4.3.2) by which the wave-zone multipole moments fail to agree with the near-zone multipole moments.

The weak-field limit of the slow-motion wave-generation formalism can be summarized by the following set of rules: (1) Use the Newtonian theory of gravity to analyze the structure and evolution of the source. (2) Calculate the lowest few multipole moments by evaluating expressions (4.4.3) and (4.4.4) in the mass-centered, Cartesian coordinates. (3) Insert those multipole moments into the local-wave-zone formulas of §3. Those formulas will then describe the lowest few multipoles of the radiation field. (4) Check that the radiation field is larger than the errors inherent (a) in the matching of near zone onto local wave zone (eqs. 4.3.2. with R = L), and (b) in the weak-field assumption

$$(\delta h^{TT}_{jk})_{\text{mass } \ell\text{-pole, errors}} \sim (M/r)\ (L/\lambda\!\!\!^{-})^\ell[\varepsilon+(L/\lambda\!\!\!^{-})^2+S^2], \tag{4.4.6a}$$

$$(\delta h^{TT}_{jk})_{\text{current } \ell\text{-pole, errors}} \sim (M/r)(L/\lambda\!\!\!^{-})^{\ell+1}[\varepsilon+(L/\lambda\!\!\!^{-})^2+S^2]. \tag{4.4.6b}$$

This weak-field, slow-motion formalism dates back to Einstein (1918), for the mass quadrupole part. The mass octupole and current quadrupole parts were first derived (so far as I know) by Papapetrou (1962, 1971); and the full formalism (in different notations from this) was first derived by Mathews (1962); see also Campbell and Morgan (1971). The earliest versions of the formalism assumed gravity so weak that one had to use Linearized Theory rather than Newtonian Theory in analyzing the source ("no self-gravity"). However, it was soon realized that a modest amount of self-gravity causes no problems in the formalism.

4.5 Post-Newtonian Multipole Formalism

We now return to the Epstein-Wagoner (1975) Post-Newtonian Wave-Generation Formalism, which was described qualitatively in §2.7. In this formalism Epstein and Wagoner find it most convenient to analyze the structure and motion of their source using a different Post-Newtonian gauge from that of Chandrasekhar (1965). Like Chandrasekhar, they describe the matter of their source by perfect-fluid thermodynamic variables measured in its local rest frame

$$\rho_o = \text{mass density}, \quad P = \text{pressure}, \quad \Pi = \text{specific internal energy}; \tag{4.5.1a}$$

and by the coordinate velocity of the fluid

$$v_j \equiv \frac{dx^j}{dt} . \tag{4.5.1b}$$

The internal gravity of the source they describe by the potentials U, V_j, Ψ, and χ which satisfy

$$U_{,jj} = -4\pi\rho_o, \quad V_{j,kk} = -4\pi\rho_o v_j,$$

$$\Psi_{,jj} = -4\pi\rho_o(v^2+U+\tfrac{1}{2}\Pi + \tfrac{3}{2}P/\rho_o), \quad \chi_{,jj} = -2U, \tag{4.5.2}$$

and which are related to the post-Newtonian metric by

$$g_{00} = -1+2U-2U^2+4\Psi-\chi_{,00}+O(\varepsilon^6) \tag{4.5.3a}$$

$$g_{0j} = -4V_j+O(\varepsilon^5) \tag{4.5.3b}$$

$$g_{jk} = 2U\delta_{jk}+O(\varepsilon^4) \tag{4.5.3c}$$

[In the Post-Newtonian formalism one assumes $\varepsilon \sim S^2 \sim (L/\lambda\!\!\!^{-})^2$.] The equations governing the evolution of the source in the Epstein-Wagoner gauge are the equation of mass conservation

$$[\rho_o(1+\tfrac{1}{2}\underset{\sim}{v}^2+3U)]_{,0}+[\rho_o(1+\tfrac{1}{2}\underset{\sim}{v}^2+3U)v_j]_{,j}=0, \tag{4.5.4a}$$

the equation of state

$$P=P(\rho_o,\Pi), \tag{4.5.4b}$$

the adiabatic equation of energy conservation

$$\rho_o\, d\Pi/dt + Pv_{j,j}=0, \quad d/dt \equiv \partial/\partial t + v_j\, \partial/\partial x_j, \tag{4.5.4c}$$

and the Euler equations of motion

$$\rho_o dv_i/dt - \rho_o U_{,i} + P_{,i} - (\Pi + \underset{\sim}{v}^2 + 4U + P/\rho_o)P_{,i}$$

$$+ (3\rho_o U_{,0} + P_{,0} + 4\rho_o v_j U_{,j})v_i + (4U - \underset{\sim}{v}^2)\rho_o U_{,i} - 4\rho_o V_{i,0}$$

$$+ 4\rho_o v_k (V_{k,i} - V_{i,k}) + \rho_o (\tfrac{1}{2}\chi_{,00i} - 2\Psi_{,i}) = 0 .$$

After one has computed the structure and evolution of the source using these equations, one can then compute the multipole moments which govern the radiation field. For typical sources -- to which Epstein and Wagoner restrict their attention -- the multipole components of the radiation field have the magnitudes (4.3.4). The mass quadrupole dominates with magnitude $(M/r)(L/\lambda\!\!\!^{-})^2$ and the post-Newtonian formalism is able to compute it with a fractional error $\sim \varepsilon^2 \sim (L/\lambda\!\!\!^{-})^4$. If one wishes to compute the other multipole contributions to similar absolute accuracy, then one must have fractional errors no larger than the following in the various multipole moments (cf. eq. 4.3.4):

$$\mathcal{I}_{A_2} : (L/\lambda\!\!\!^{-})^4, \quad \mathcal{I}_{A_3} : (L/\lambda\!\!\!^{-})^3, \quad \mathcal{I}_{A_4} : (L/\lambda\!\!\!^{-})^2, \quad \mathcal{I}_{A_5} : (L/\lambda\!\!\!^{-}),$$

$$\mathcal{S}_{A_2} : (L/\lambda\!\!\!^{-})^3 \quad \mathcal{S}_{A_3} : (L/\lambda\!\!\!^{-})^2 \quad \mathcal{S}_{A_4} : (L/\lambda\!\!\!^{-}). \qquad (4.5.5)$$

Up to this accuracy the multipole moments can be expressed in the following form:

$$(\mathcal{I}_{jk},\ \mathcal{I}_{ijk},\ \mathcal{S}_{jk}) = \text{Symmetric trace-free parts of } (I_{jk},\ I_{ijk},\ S_{jk}),$$

$$I_{jk} = \int\left[\tau_{00}x_j x_k + \frac{11}{21}\tau_{jk}r^2 + \frac{4}{21}\tau_{pp}x_j x_k - \frac{4}{7}x_p\tau_{pj}x_k\right]d^3x,$$

$$I_{ijk} = \int\left[\tau_{00}x_i x_j x_k + x_i\tau_{jk}r^2 + \frac{1}{3}\tau_{pp}x_i x_j x_k - x_p\tau_{pi}x_j x_k\right]d^3x, \qquad (4.5.6)$$

$$S_{ij} = \int\left[x_i\varepsilon_{jpq}x_p\tau^0{}_q - \frac{3}{28}\varepsilon_{ipq}x_p r^2\partial_t\tau_{qj} + \frac{1}{28}x_i\varepsilon_{jpq}x_p\partial_t\tau_{qs}x_s\right]d^3x,$$

$(\mathcal{S}_{ijk},\ \mathcal{I}_{ijk\ell})$ equal the Newtonian expressions (4.4.3), (4.4.4).

Here, the "effective energy density". τ_{00}, "effective momentum density" $\tau^0{}_j$, and "effective stress" τ_{jk} are

$$\tau_{00} = \rho_o(1 + \Pi + \underset{\sim}{v}^2 + 4U) - \frac{3}{8\pi}U_{,j}U_{,j} \qquad (4.5.7a)$$

$$\tau^0{}_j = \rho_o(1+\Pi+v^2+4U)v_j + Pv_j + \frac{3}{4\pi}U_{,0}U_{,j} + \frac{1}{2\pi}U_{,k}V_{k,j}$$

$$- \frac{1}{2\pi}V_k U_{,kj} - 2\rho_o V_j , \qquad (4.5.7b)$$

$$\tau_{jk} = \rho_o v_j v_k - \frac{1}{4\pi}U_{,j}U_{,k} - \frac{1}{2\pi}UU_{,jk}$$

$$+ \delta_{jk}(P + \frac{3}{8\pi}U_{,i}U_{,i} - 2\rho_o U). \qquad (4.5.7c)$$

Epstein and Wagoner use a different notation than ours for the multipole moments. The above formulas for the moments (eqs. 4.5.6) are derived in TV.

5. PERTURBATION FORMALISMS

Thus far I have described two large classes of wave-generation formalisms: weak-field formalisms (§2) and slow-motion formalisms (§4). Now I turn attention to a third large class: Perturbation formalisms.

The fundamental assumption underlying all perturbation formalisms is this: that the entire wave-generation region of space-time can be treated as a small perturbation which radiates, superimposed on a nonradiative but strongly curved "background" Examples are: (1) Small-amplitude pulsations of fully relativistic stars. Here the background is an unperturbed, equilibrium stellar model; and the perturbation is the pulsation. (2)Slow rotation of a slightly nonspherical neutron star (pulsar). Here the background is a nonrotating, spherical star; and the perturbation is both the deformation and the rotation. (3) Motion of a small object in the gravitational field of a black hole. Here the background is the Kerr metric of the black hole; and the small perturbation is the stress-energy tensor of the object, plus the gravitational field it produces.

There are a variety of different perturbation formalisms, each designed to handle a specific type of problem. Some applications make use of several formalisms combined together--and many applications combine a perturbation formalism for the wave-generation region with a multipole analysis of the radiation field (i.e., with a variant of the formalism described in §3).

In a recent review article (§II.C of Thorne 1977c) I have described most of the perturbation formalisms with which I am familiar. Rather than repeat that material here, I simply refer the reader to it.

6. WAVE GENERATION BY SOURCES WITH STRONG INTERNAL FIELDS AND FAST, LARGE-AMPLITUDE MOTIONS

The strongest sources of gravitational waves in the Universe should be sources with strong internal fields ($\varepsilon \sim 1$) and fast ($v \sim L/\lambda\!\!\!^{-} \sim 1$), large-amplitude internal motions. Examples are the highly nonspherical collapse of a star to form a black hole (or naked singularity, or whatever it does form), and a collision between two black holes. Unfortunately, such sources have eluded all efforts at analytic analysis.[5] There exists no formalism today by which one can calculate the waves they generate. Obviously, an accurate analysis of such systems is the most important and most difficult task lying ahead of us in the theory of gravitational-wave generation.

Fortunately great progress has been made on this task recently by Bryce DeWitt, Larry Smarr, Kenneth Eppley, and others (see Smarr 1977 for a review). Abandoning all hope of a truly analytic analysis, they turn to massive electronic computers as their key tool. Their method is elegant numerical solution of the full, nonlinear Einstein field equations. By now they have encountered and surmounted a number of serious numerical problems. The resulting numerical methods are nearly good enough to give reliable results for strong-field, high-speed, large-amplitude sources--and we can expect true reliability within another year or two. This, when it is achieved, will be very useful in planning gravitational-wave-detection efforts.

[5] An exception is the special situation of a collision between two black holes with relative velocity very nearly the speed of light [$\gamma=(1-v^2)^{-1/2} \gg 1$], for which D'Eath (1977) has formulated a remarkable "colliding-plane-wave" approximation that yields the dominant features of this radiation.

7. PROPAGATION OF WAVES TO EARTH

7.1. The Geometric-Optics Formalism

Once the gravitational-wave field h^{TT}_{jk} is known in the local wave zone, one can then propagate it out through the surrounding Universe to Earth, using the vacuum Einstein field equations. In nearly all situations one can use the geometric-optics approximation to the Einstein equations (e.g. Exercise 35.15 of MTW). In this section I describe the geometric optics formalism in a language different from, but equivalent to MTW.

This formalism, which is valid for $r \gg r_I$ (wave zone), describes spacetime by a background metric $g^{(B)}_{\mu\nu}$ through which the waves propagate. The waves are described by a gravitational-wave field $\psi_{\mu\nu}$ which reduces to

$$\psi_{jk} = h^{TT}_{jk}\,,\quad \psi_{j0} = 0,\quad \psi_{00} = 0 \quad \text{in local wave zone and in asymptotic rest frame of source.} \tag{7.1.1}$$

In an appropriate coordinate system (gauge) the full metric of spacetime is

$$g_{\mu\nu} = g^{(B)}_{\mu\nu} + \psi_{\mu\nu} + 0\,(\psi^2). \tag{7.1.2}$$

The waves are distinguished from the background by the very small length scale λ on which they vary

$$\lambda \equiv \begin{bmatrix}\text{length scale of}\\ \text{variations in } \psi_{\mu\nu}\end{bmatrix} \ll \mathcal{R}_B \equiv \begin{bmatrix}\text{radius of curvature}\\ \text{of background spacetime}\end{bmatrix}. \tag{7.1.3}$$

The geometric optics formalism remains valid so long as the propagating waves do not encounter regions of extremely strong curvature--i.e. regions where $\mathcal{R}_B \lesssim \lambda$.

In applying the geometric optics formalism to a specific problem, one proceeds as follows: (i) In the local wave zone and in the asymptotic rest frame of the source one describes the background metric by the line element

$$ds^2 = -dt^2 + dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) + O(M/r)dx^\alpha dx^\beta, \tag{7.1.4}$$

with the source located at the origin. (ii) One constructs null geodesics of $g^{(B)}_{\mu\nu}$ extending radially away from the source. Near the source each geodesic has the form

$$t - r = \tau_e = \text{const.},\quad \theta = \text{const.},\quad \phi = \text{const.} \quad \text{in local wave zone.} \tag{7.1.5}$$

These geodesics are called "rays." The gravitational waves propagate along them. (iii) Each ray is labeled by the proper time ("emission time" or "retarded time") τ_e at which it intersects the source; and by this ray labeling τ_e becomes a scalar field extending through all spacetime. (iv) Each ray has an affine parameter ζ, a world line $\mathcal{P}(\zeta)$ with coordinates $x^\alpha(\zeta)$, and a tangent vector ("propagation vector")

$$\vec{k} = d/d\zeta, \qquad k^\alpha = dx^\alpha/d\zeta. \tag{7.1.6}$$

Of course, since $\mathcal{P}(\zeta)$ is a null geodesic, $\vec{k}$ must satisfy

$$\vec{k}^2 \equiv k_\alpha k^\alpha = 0, \qquad k_{\alpha|\beta} k^\beta = 0. \tag{7.1.7}$$

Here and below a slash denotes covariant derivative with respect to the background metric. (v) One normalizes the affine parameter of each ray so that near the source

$$\zeta = r = \left\{\begin{array}{l}\text{proper distance from source}\\ \text{as measured in asymptotic}\\ \text{rest frame; eq.(7.1.4)}\end{array}\right\} \quad \text{in local wave zone.} \tag{7.1.8a}$$

As a result

$$k^0 = 1, \qquad \underset{\sim}{k} = \underset{\sim}{n} = \left\{\begin{array}{l}\text{unit radial vector}\\ \text{pointing away from}\\ \text{source}\end{array}\right\} \quad \begin{array}{l}\text{in source's asymptotic}\\ \text{rest frame and local}\\ \text{wave zone.}\end{array} \tag{7.1.8b}$$

(vi) From the above definitions and constructions one can show that throughout spacetime the gradient of the retarded time is equal to the propagation vector, except for sign

$$\vec{k} = -\vec{\nabla}\tau_e; \qquad k_\alpha = -\tau_{e,\alpha}. \tag{7.1.9}$$

(vii) In the local wave zone one imposes the starting conditions (7.1.1) on the gravitational-wave field $\psi_{\alpha\beta}$. Note that because h^{TT}_{jk} is transverse and traceless, $\psi_{\alpha\beta}$ is initially trace-free and orthogonal to the propagation vector

$$\psi \equiv \psi^\alpha_\alpha \equiv \psi_{\alpha\beta}\, g^{\alpha\beta}_{(B)} = 0, \quad \psi_{\alpha\beta} k^\beta = 0. \tag{7.1.10}$$

These properties are preserved as the waves propagate (cf. eq. 7.1.12 below). (viii) Initially, and after it has propagated, $\psi_{\alpha\beta}$ is a rapidly varying function of retarded time τ_e and in addition is a slowly varying function of location along surfaces of constant τ_e:

$$\psi_{\alpha\beta} = \psi_{\alpha\beta}\ (\tau_e; x^0, x^1, x^2, x^3), \tag{7.1.11a}$$

$$\dot{\psi}_{\alpha\beta} \equiv \frac{\partial \psi_{\alpha\beta}}{\partial \tau_e} = O\left[\frac{\psi_{\alpha\beta}}{\lambda}\right], \qquad \ddot{\psi}_{\alpha\beta} \equiv \frac{\partial^2 \psi_{\alpha\beta}}{\partial \tau_e^2} = O\left[\frac{\psi_{\alpha\beta}}{\lambda^2}\right] \tag{7.1.11b}$$

$$\left\{\frac{\partial \psi_{\alpha\beta}}{\partial x^\mu}\right\}_{\tau_e} = O\left[\frac{\psi_{\alpha\beta}}{\zeta} + \frac{\psi_{\alpha\beta}}{\mathcal{R}_B}\right], \tag{7.1.11c}$$

$$\psi_{\alpha\beta|\gamma} = \dot{\psi}_{\alpha\beta} k_\gamma + O\left[\frac{\psi_{\alpha\beta}}{\zeta} + \frac{\psi_{\alpha\beta}}{\mathcal{R}_B}\right]. \tag{7.1.11d}$$

Here the slash ($\psi_{\alpha\beta|\gamma}$) denotes covariant derivative with respect to the background metric, and O[X] means "of order X." (ix) $\psi_{\alpha\beta}$ propagates along the rays in accordance with the propagation equation

$$\psi_{\alpha\beta|\mu} k^\mu = -\frac{1}{2} k^\mu{}_{|\mu} \psi_{\alpha\beta}. \tag{7.1.12}$$

(x) One uses equation (7.1.12) to propagate the initial, local-wave-zone field out through the Universe to Earth.

Because, in an appropriate gauge, $\psi_{\alpha\beta}$ is the metric perturbation associated with the waves (eq. 7.1.2), one can use the usual formulas (Chapters 35 and 37 of MTW) to calculate from $\psi_{\alpha\beta}$ whatever properties of the waves one wishes. For example, one can use equations (35.62a,e)of MTW to compute the contribution of the waves to the Riemann curvature tensor of spacetime. By virtue of equations (7.1.11) that contribution is

$$R^{(GW)}_{\alpha\beta\gamma\delta} = \frac{1}{2}\left(\ddot{\psi}_{\alpha\delta} k_\beta k_\gamma + \ddot{\psi}_{\beta\gamma} k_\alpha k_\delta - \ddot{\psi}_{\beta\delta} k_\alpha k_\gamma - \ddot{\psi}_{\alpha\gamma} k_\beta k_\delta\right) + \text{fractional errors of } O[\lambda/\zeta + \lambda/\mathcal{R}_B]. \tag{7.1.13}$$

Similarly, one can use equation (35.70) of MTW to compute the Isaacson (1968) stress-energy tensor associated with the waves. It reduces to

$$T^{(GW)}_{\alpha\beta} = \frac{1}{32\pi} \langle \dot{\psi}_{\mu\nu} \dot{\psi}^{\mu\nu} \rangle k_\alpha k_\beta, \tag{7.1.14}$$

where $\langle\ \rangle$ means"average over several wavelengths of the waves."

The gravitational-wave field is not fully measurable. Only that part which contributes to the Riemann curvature tensor is measurable; the remainder can be changed at will by gauge transformations (infinitesimal coordinate transformations). The form of the generic gauge transformation which preserves all the above equations is

$$\psi_{\mu\nu}^{NEW} = \psi_{\mu\nu}^{OLD} + \Phi_\mu k_\nu + \Phi_\nu k_\mu \ , \qquad (7.1.15a)$$

where Φ_μ is a vector field which has the same variability properties as $\psi_{\alpha\beta}$ [eqs. 7.1.11], which is orthogonal to the propagation vector

$$\Phi_\mu k^\mu = 0, \qquad (7.1.15b)$$

and which satisfies the propagation equation

$$\Phi_{\mu|\alpha} k^\alpha = -\frac{1}{2} k^\alpha \Phi_{\mu|\alpha} \ . \qquad (7.1.15c)$$

An arbitrary observer with 4-velocity $\vec{u}$, at an arbitrary event in spacetime, may find it convenient to make $\psi_{\mu\nu}$ "spatial, transverse, and traceless" in his own rest frame. He can accomplish this by a gauge transformation of the above form with

$$\Phi_\mu = -\frac{u^\alpha \psi_{\alpha\mu}}{u^\gamma k_\gamma} + \frac{u^\alpha \psi_{\alpha\beta} u^\beta}{2(u^\gamma k_\gamma)^2} k_\mu \ . \qquad (7.1.16a)$$

The resulting ("NEW"; "Transverse Traceless") gravitational-wave field $\psi_{\mu\nu}^{TT}$ is related to the original ("OLD") one $\psi_{\mu\nu}$ by

$$\psi_{xx}^{TT} = \psi_{xx} = -\psi_{yy} = -\psi_{yy}^{TT}; \quad \psi_{xy}^{TT} = \psi_{xy} = \psi_{yx}^{TT} \ ; \quad \text{all other } \psi_{\alpha\beta}^{TT} = 0$$

in local rest frame of observer, with Minkowskii coordinates so oriented that $\vec{k} = k^0(\vec{e}_0 + \vec{e}_z)$. (7.1.16b)

In other words, this gauge transformation simply throws away all parts of $\psi_{\alpha\beta}$ except those that are purely spatial and are transverse to the propagation direction; and in the process it preserves the tracelessness of $\psi_{\alpha\beta}$.

7.2 Example of Propagation

As an example of the above formalism, consider a gravitational wave emitted by the gravitational collapse of a 10^6 solar-mass star at a time when the universe was only a few million years old.

Idealize the background metric, through which the waves propagate toward Earth, as a perfectly smooth Friedmann metric (MTW chapter 29)

$$ds^2 = g^{(B)}_{\mu\nu} dx^\mu dx^\nu = a^2[-d\eta^2 + d\chi^2 + \Sigma^2(d\theta^2 + \sin^2\theta d\phi^2)] \quad (7.2.1a)$$

$$a = a(\eta), \quad \Sigma = \begin{cases} \sin\chi & \text{if universe is "closed"} \\ \chi & \text{if universe is "flat"} \\ \sinh\chi & \text{if universe is "open"} \end{cases} ; \quad (7.2.1b)$$

and at each event in the universe introduce the orthonormal frame ("local proper rest frame of universe")

$$\vec{e}_{\hat{\eta}} = \frac{1}{a}\frac{\partial}{\partial\eta}, \quad \vec{e}_{\hat{\chi}} = \frac{1}{a}\frac{\partial}{\partial\chi}, \quad \vec{e}_{\hat{\theta}} = \frac{1}{a\Sigma}\frac{\partial}{\partial\theta}, \quad e_{\hat{\phi}} = \frac{1}{a\Sigma\sin\theta}\frac{\partial}{\partial\phi} . \quad (7.2.1c)$$

Place the supermassive star (source) at the origin of the Friedmann spatial coordinate system, $\chi_s = 0$; and let it emit its gravitational waves during an interval of time $\eta_s \le \eta \le \eta_s + \Delta\eta$. The duration of the burst [$\Delta\tau_e = a(\eta_s)\Delta\eta$ = (a few seconds)] will be very short compared to the age of the universe at the time of emission; and hence "a" will change negligibly during the emission

$$a_s \equiv a(\eta_s) >>> (da/d\eta)_s \, \Delta\eta . \quad (7.2.2)$$

The rays, along which the waves propagate, are null geodesics emanating from the star's world line ($\chi_s = 0$, $\eta_s \le \eta \le \eta_s + \Delta\eta$). By solving the geodesic equation in the Friedmann metric one obtains the following equations for the ray originating at ($\chi = 0$, $\eta = \eta_s + \tau_e/a_s$) and propagating in the (θ_e, ϕ_e) direction:

$$\chi = \eta - (\eta_s + \tau_e/a_s), \quad \theta = \theta_e, \quad \phi = \phi_e , \quad (7.2.3a)$$

$$k^\eta = k^\chi = \frac{d\eta}{d\zeta} = \frac{d\chi}{d\zeta} = \frac{a_s}{a^2}, \quad k^\theta = k^\phi = 0. \quad (7.2.3b)$$

Note that the "physical components" of the propagation vector (components in the local proper rest frame of the universe [7.2.1c]) are

$$k^{\hat{\theta}} = k^{\hat{\phi}} = 0, \quad k^{\hat{\eta}} = k^{\hat{\chi}} = (a_s/a) \text{ at general event} \quad (7.2.4)$$

$$= 1 \text{ at source.}$$

The source's retarded time, expressed as a function of Friedmann coordinates, is

$$\tau_e = a_s\,(\eta - \eta_s - \chi); \tag{7.2.5}$$

cf. equation (7.2.3a). The gradient of τ_e is the negative of the propagation vector, as required by equation (7.1.9).

In the asymptotic rest frame of the source ($\chi \ll 1$; $\eta_s \lesssim \eta \lesssim \eta_s + \Delta\eta$) the Minkowskii radial and time coordinates are related to Friedmann coordinates and to retarded time by

$$r = a_s\chi, \quad t = a_s(\eta-\eta_s); \quad \tau_e = t-r; \tag{7.2.6a}$$

and the line element is

$$ds^2 = -dt^2 + dr^2 + r^2\,(d\theta^2 + \sin^2\theta\, d\phi^2) \tag{7.2.6b}$$

$$+ O(M/r)\, dx^\alpha\, dx^\beta \leftarrow \text{contributions from gravity of source}$$

$$+ O(r^2/a^2 + t\, d\,\ell n\, a/dt) dx^\alpha dx^\beta \leftarrow \text{cosmological corrections.}$$

The gravitational waves, expressed in a gauge that is "TT" with respect to this asymptotic rest frame, have as their only nonzero components

$$\psi_{\hat\theta\hat\theta} = -\psi_{\hat\phi\hat\phi} \equiv \frac{1}{r}\, A_+(\tau_e,\theta,\phi) \qquad \text{in local wave zone.} \tag{7.2.7}$$

$$\psi_{\hat\theta\hat\phi} = \psi_{\hat\phi\hat\theta} \equiv \frac{1}{r}\, A_x(\tau_e,\theta,\phi)$$

The quantities A_+ and A_x are amplitudes for the two orthogonal polarization states "+" and "x". They are rapidly varying functions of retarded time $\tau_e = t-r$, and slowly varying functions of θ and ϕ :

$$\dot A \equiv \frac{\partial A}{\partial \tau_e} \sim \frac{A}{\lambda\!\!\!^{-}}, \qquad \frac{1}{r}\frac{\partial A}{\partial\theta} \sim \frac{A}{r}, \qquad \frac{1}{r\sin\theta}\frac{\partial A}{\partial\phi} \sim \frac{A}{r} \tag{7.2.8}$$

[cf. eq. (7.1.11) and note that in the asymptotic rest frame ζ = (affine parameter) = r].

The gravitational-wave field (7.2.7) propagates out into the Friedmann universe by means of the geometric-optics propagation equation (7.1.12). By solving this equation one discovers that

$$a\Sigma\psi_{\hat{\alpha}\hat{\beta}} = \text{constant along each ray,} \tag{7.2.9}$$

where Σ is the cosmological circumference function defined in equation (7.2.1b). Consequently (cf. eqs. [7.2.1b], [7.2.6a], [7.2.7], and [7.2.5]) the gravitational-wave field at an arbitrary event (η,χ,θ,ϕ) in spacetime is

$$\psi_{\hat{\theta}\hat{\theta}} = -\psi_{\hat{\phi}\hat{\phi}} = \frac{1}{a\Sigma} A_+[a_s(\eta-\eta_s-\chi),\theta,\phi] \tag{7.2.10a}$$

$$\psi_{\hat{\theta}\hat{\phi}} = \psi_{\hat{\phi}\hat{\theta}} = \frac{1}{a\Sigma} A_x[a_s(\eta-\eta_s-\chi),\theta,\phi]. \tag{7.2.10b}$$

Because the source is far from Earth, when these waves reach Earth they look plane waves. An observer on Earth can interpret them in terms of a local Minkowskii coordinate system with basis vectors

$$\frac{\partial}{\partial t} \equiv \vec{e}_0 = \vec{e}_{\hat{\eta}}, \quad \frac{\partial}{\partial z} \equiv \vec{e}_z = e_{\hat{\chi}}, \quad \frac{\partial}{\partial x} \equiv \vec{e}_x = \vec{e}_{\hat{\theta}}, \quad \frac{\partial}{\partial y} \equiv \vec{e}_y = \vec{e}_{\hat{\phi}}. \tag{7.2.11a}$$

If the location of Earth today is $(\eta_o, \chi_o, \theta_o, \phi_o)$, then the observer's Minkowskii coordinates and the global Friedmann coordinates near Earth are related by

$$t=a_o(\eta-\eta_o),\ z=a_o(\chi-\chi_o),\ x=a_o\Sigma_o(\theta-\theta_o),\ y=a_o\Sigma_o\sin\theta_o(\phi-\phi_o). \tag{7.2.11b}$$

The observer on Earth describes the universe's cosmological structure in terms of a Hubble expansion rate H_o and a deceleration parameter q_o; and he describes the source of the waves as having a cosmological redshift Z_s, which is related to the expansion factor by

$$1 + Z_s = a_o/a_s. \tag{7.2.12}$$

The equations of Friedmann cosmology permit one to express $a_o\Sigma_o \equiv$ (circumference of a circle that passes through Earth and is centered on the source at redshift Z_s)$/2\pi$ in terms of H_o, q_o, and Z_s:

$$a_o\Sigma_o \equiv R = \frac{H_o^{-1}}{q_o^2(1+Z_s)}\left[1-q_o+q_oZ_s-(1-q_o)(2q_oZ_s+1)^{1/2}\right]; \tag{7.2.13}$$

cf. eq. (29.33) of MTW. By combining equations (7.2.10) -(7.2.13) one obtains the following expressions for the gravitational-wave field that sweeps past Earth:

$$\psi_{xx} = -\psi_{yy} = \frac{1}{R}A_+\left[\frac{t-z}{1+Z_s} + \text{const.}, \theta_o,\phi_o\right], \tag{7.2.14a}$$

$$\psi_{xy} = \psi_{yx} = \frac{1}{R}A_x\left[\frac{t-z}{1+Z_s} + \text{const.},\theta_o,\phi_o\right], \tag{7.2.14b}$$

all other $\psi_{\mu\nu}$ vanish.

Notice that the time dependence, $A[(t-z)/(1+Z_s)+\text{const.}, \theta_o,\phi_o]$, of the waves as they sweep past Earth is identical to the time dependence of the emitted waves as measured by the source, $A[\tau_e,\theta_o,\phi_o]$, except for a redshift of $1+Z_s$ -- the same redshift as one sees in electromagnetic spectral lines. Notice also that for very large redshifts, $Z_s >>1$, the amplitude of the waves is independent of redshift:

$$\psi \sim \frac{A}{R} = \frac{A}{H_oq_o} \quad \text{for} \quad Z_s >> 1 \quad \text{and} \quad Z_s >> 1/q_o. \tag{7.2.15}$$

8. CONCLUDING REMARKS

Although it may appear from these lectures that the theory of gravitational-wave generation is a highly sophisticated and complex subject, one should not let this blind one to its gross inadequacies.

The strongest sources of gravitational waves in the universe--and the most promising sources for ultimate detection -- are those with strong internal gravity and fast large-amplitude internal motions. For them the only reliable technique of analysis is massive computer calculations (§6). All the fancy analytic tools of these lectures are helpless in the face of such sources!

I am indebted to Joseph Weber for his patience, and to Alessandra Exposito and Jim Isenberg for valuable assistance in the preparation of the typed manuscript.

REFERENCES

Bonnor, W.B., and Rotenberg, M.A. 1966, Proc. Roy. Soc. A, 289, 247.

Campbell, W.B. and Morgan, T. 1971, Physica, 53, 264.

Chandrasekhar, S. 1964, Astrophys. J., 140, 417.

Chandrasekhar, S. 1965, Astrophys. J., 142, 1488.

Clarke, C.J.S. and Sciama, D.W. 1971, Gen. Rel. Grav., 2, 331.

Crowley, R.J. and Thorne, K.S., 1977, Astrophys. J., in press; also available as Caltech Orange-Aid Preprint 450.

D'Eath, P. 1977, paper submitted for publication in Phys. Rev.

DeWitt, B.S. 1971, Lectures on General Relativity-Stanford, Fall 1971 (unpublished manuscript).

Eardley, D.M., Lee, D.L., and Lightman, A.P. 1974, Phys. Rev. D, 10, 3308.

Eddington, A.S. 1922, Proc. Roy. Soc. A, 102, 268.

Einstein, A. 1918, Berlin Sitzungsberichte, p. 154.

Epstein, R. 1976, Ph.D. Thesis, Stanford University.

Epstein, R. and Wagoner, R.V., 1975, Astrophys.J., 197, 717.

Feynman, R.P., 1963, Acta Physica Polonica, 24, 697.

Geroch, R. 1970, J. Math. Phys., 11, 2580.

Havas, P. and Goldberg, J.N. 1962, Phys. Rev., 128, 398.

Ipser, J.R. 1971, Astrophys. J., 166, 175.

Isaacson, R.A. 1968, Phys. Rev., 166, 1272.

Janis, A.I. and Newman, E.T. 1965, J. Math. Phys., 6, 902.

Kovács, S.J., Jr. and Thorne, K.S. 1977a, paper submitted to Astrophys. J: also available as Caltech Orange Aid Preprint

Kovács, S.J., Jr. and Thorne, K.S. 1977b, paper in preparation.

Landau, L.D. and Lifshitz, E.M. 1962, The Classical Theory of Fields (Oxford: Pergamon Press).

Mashoon, B. 1973, Astrophys. J., 181, L65.

Mathews, J. 1962, J. Soc. Indust. Appl. Math, 10, 768.

Matzner, R.A. and Nutku, Y. 1974, Proc. Roy. Soc. A, 336, 285.

Misner, C.W., Thorne, K.S., and Wheeler, J.A. 1973, Gravitation (San Francisco, W.H. Freeman and Co.); cited in text as MTW.

Papapetrou, A. 1962, Comptes Rendues, 255, 1578.

Papapetrou, A. 1971, Ann. Inst. Henri Poincare, 14, 79.

Peters, P.C. 1970, Phys. Rev. D, 1, 1559.

Peters, P.C. and Mathews, J. 1963, Phys. Rev., 131, 435.

Pirani, F.A.E. 1965 in A. Trautman, F.A.E. Pirani, and H. Bondi, Lectures on General Relativity (Englewood Cliffs, N.J.: Prentice-Hall).

Press, W.H., 1977, paper in press.

Regge T. and Wheeler, J.A. 1957, Phys. Rev., 108, 1063.

Ruffini, R. and Wheeler, J.A. 1971, in Proceedings of the Conference on Space Physics (Paris: ESRO), p. 45.

Sachs, R. 1961, Proc. Roy. Soc. A, 264, 309, esp. Appendix B.

Shapiro, I.I. 1964, Phys. Rev. Lett., 13, 789.

Smarr, L. 1977, in Proceedings of the Eighth Texas Symposium on Relativistic Astrophysics, to be published in Proceedings of the New York Academy of Sciences.

Thorne, K.S. 1977a, paper to be submitted for publication. Cited in text as TV.

Thorne,, K.S. 1977b, paper to be submitted for publication. Cited in text as TVI.

Thorne, K.S., 1977c , in N. Lebovitz ed., Proceedings of the (Chandrasekhar) Symposium on Theoretical Principles in Astrophysics and Relativity (Chicago: University of Chicago Press).

Thorne, K.S. 1977d, paper submitted to Astrophys. J.

Thorne, K.S. and Kovács, S.J., Jr. 1975, Astrophys. J., 200, 245. Cited in text as TK.

Turner, M. 1977, paper submitted for publication; preprint, Stanford University.

van der Burg, M.G.J. 1968, Proc. Roy. Soc. A, 303, 37.

Wagoner R.V. and Will, C.M. 1976, Astrophys. J., 210, 764.

Wagoner, R.V. 1977, in Proceedings of the Academia Nazionale dei Lincei International Symposium on Experimental Gravitation, Pavia, Italy, September 1976; B. Bertotti ed., in press.

Wheeler, J.A. 1966, Ann. Rev. Astron. Astrophys., 4, 423.

International Journal of Astronomy and Astrophysics, 2013, 3, 39-50
http://dx.doi.org/10.4236/ijaa.2013.32A007 Published Online June 2013 (http://www.scirp.org/journal/ijaa)

Gravity Caused by TEM Waves Operating on Dipoles in Atoms

Hans W. Giertz
Uppsa Research, Gnesta, Sweden
Email: hans@miklagaard.com

Received February 24, 2013; revised March 25, 2013; accepted April 3, 2013

Copyright © 2013 Hans W Giertz. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT

The study displays the existence of a gravitational singularity in the universe generating synchronized and extremely low frequency plane TEM (transverse electromagnetic) waves. It is proposed that atomic intrinsic electromagnetic fields create resonance with these plane TEM waves, causing atoms to receive and to re-emit synchronized plane TEM waves. The energy flow of synchronized plane TEM waves, travelling in opposite directions between e.g. two atoms, creates mutual force of attraction, *i.e.* gravity. Consequently, gravity is not an intrinsic atomic feature; however, the result of passive atoms exposed to electromagnetic energy. The study describes how plane TEM waves emitted by the gravitational singularity were measured. The study also displays how gravity from the earth, moon, sun and the gravitational singularity was measured and how gravity was simulated using an electronic device. The present electromagnetic law of gravity is compared with Newtonian geometric law of gravity.

Keywords: Gravity; Gravitation; Poynting Vector; Elementary Particle; Dipole; TEM Waves

1. Introduction

Newton made the geometric description of gravity. Despite elapsed time the theoretical description of gravity remains unsolved.

It's known that the atom contains charge, static electric- and magnetic dipole moment. These fields decrease with the square of the distance and can consequently not create gravity. It seems unlikely that the atom generates alternating electromagnetic fields by itself, and hence electromagnetic energy is often ruled out as the origin of gravity. The prevailing track is to explain gravity with string theory [1,2].

The present study builds on the observation that gravity comprises plane TEM (transverse electromagnetic) waves or energy. The theory of plane TEM waves is well known and hence, it is possible to describe gravity entirely by the laws of electromagnetism. All theory is found in university grade textbooks [3,4]. TEM waves can also be explained by its photon according to the wave-particle duality [5,6]. In the present paper the gravitational energy is described by the plane TEM wave.

The study builds on the observation that the universe contains one generator emitting plane TEM waves. All emitted plane TEM waves are quasi synchronized, facilitated by its extremely low frequency and origin. This generator or gravitational singularity can consist of circulating charge and is called the Gravity Source in the present study.

Electric, **E**, and magnetic, **B**, field vectors are linked in plane TEM waves and in this case they are almost static. An atomic elementary entity that contains a static electric field, $\mathbf{E}_i$, merged with a static magnetic field, $\mathbf{B}_i$, creates resonance with this type of plane TEM waves; it receives plane TEM waves and re-emits them without changing content. It is called MQ (mass quantum) in the present study.

In order to illustrate the simplicity of gravity a grossly simplified universe is described. This simplified universe contains a Gravity Source and only two MQs, MQ_1 and MQ_2. The Gravity Source emits plane TEM waves, TEM_1, to MQ_1. MQ_1 creates resonance, receives and re-emits TEM_1 uniformly in space and where a part is received by MQ_2. MQ_2 creates resonance, receives and re-emits TEM_1 uniformly in space and where a part is received by the Gravity Source. Simultaneously the Gravity Source emits plane TEM waves, TEM_2, to MQ_2, MQ_2 re-emits a part to MQ_1 and MQ_1 re-emits a part to the Gravity Source. Consequently, this simplified universe contains circulating, synchronized plane TEM waves travelling in opposite directions between the Gravity Source, MQ_1 and

MQ_2. According to the laws of electromagnetism, synchronized plane TEM waves travelling in opposite directions create mutual force of attraction, *i.e.* gravity. This model can be scaled up to any number of MQs, e.g. to today's universe.

In the present study the hypothesis is that gravity is electromagnetic energy generated centrally in the universe, distributed within the universe, received and re-emitted by fully passive atoms or MQs (mass quanta) in atoms.

This theory was verified in Section 2. Synchronized, plane TEM waves travelling in opposite direction were mapped on the laws of electromagnetism. It resulted in an equation describing the relation between the number of involved MQs, their distance and the mutual force of attraction. This relationship can be called the electromagnetic law of gravity and which was compared with the Newtonian geometric law of gravity.

Section 3 describes an electronic device that simulated the MQ. This device was used in Section 4 in order to empirically verify the theory in Section 2. Measurements were made according to a method described in [7]. Section 3 also describes an electronic device that contained a static magnetic field merged to an alternating electric field and where the frequency of this field was adjusted to the frequency of plane TEM waves emitted by the Gravity Source. It facilitated measurement of the Gravity Source frequency.

Section 5 covers discussion and where the measured (gravity) TEM wave's characteristics are mapped on the atom's nucleus, traditionally described by the Standard Model [8,9]. It results in a better understanding of the atom. The position of the earth's rotational axis relative to the Galactic Centre [10,11] and the Gravity Source is discussed based on data obtained from WMAP and the Planck mission [12], indicating the huge influence of the Gravity Source on the entire universe.

Conclusions are made in Section 6.

Aim of the study: the aim of the study is to present and to verify a theoretical model of gravity.

2. Theoretical Model of Gravity

Maxwell's equations $\operatorname{curl}\mathbf{H} = \boldsymbol{i} + \partial\mathbf{D}/\partial t$ and $\operatorname{curl}\mathbf{E} = -\partial\mathbf{B}/\partial t$ result in plane TEM (transverse electromagnetic) waves [3]. This is the most common electromagnetic energy in the universe resulting in e.g. radio waves, light and X-rays depending on frequency.

The behavior and characteristics of plane TEM waves is independent of its frequency. The field vectors **E** and **B** or $\mathbf{H}(\mathbf{B} = \mu_0\mathbf{H})$ are linked and perpendicular. The field vector amplitudes relate as $\mathrm{E} = c\mathrm{B}$, where c is the speed of light. The electric and magnetic field vector amplitudes E and B decrease with the distance from the source as 1/r. The Poynting vector **S** describes the TEM energy flow density [4]

$$\mathbf{S} = \mathbf{E} \times \mathbf{H} \tag{1}$$

The energy stored in the magnetic field is just equal to that in the magnetic field. The direction of energy flow is reversed for a wave travelling in the opposite direction because the phase of **E** and **H** is reversed.

Assume that one source, Source 1, radiates plane TEM_1 waves, described by its Poynting vector $\mathbf{S}_1(r)$ at the distance r from Source 1. Source 2 is positioned at the distance r. Source 2 is switched on whereby it radiates plane TEM_2 waves uniformly. In the direction towards Source 1 plane TEM_1 and TEM_2 waves interact because of the force between their synchronized field vectors. This also implies that there is energy transfer from TEM_1 to TEM_2 when TEM_2 propagates towards Source 1. In equilibrium there is energy transfer in both directions, *i.e.* from TEM_1 to TEM_2 vice versa. Note that **E** and **H** are always perpendicular and hence $\mathbf{S} = \mathbf{E} \times \mathbf{H}$ can be replaced by $\mathrm{S} = \mathrm{E} \cdot \mathrm{H}$. Equilibrium is defined by the amount of energy that Source 1 is able to deliver at Source 2, e.g. at the distance r, $\mathrm{S}_1(r)$, and the amount of energy that Source 2 is able to deliver at Source 1, *i.e.* $\mathrm{S}_2(r)$. This results in back reaction force $\mathrm{F}(r)$ proportional to $\mathrm{S}_1(r)$ on Source 1 and $\mathrm{S}_2(r)$ on Source 2, *i.e.* radiation reaction [4], and where γ is a constant. This creates mutual force of attraction $\mathrm{F}(r)$,

$$\mathrm{F}(r) = \gamma \cdot \mathrm{S}_1(r) \cdot \mathrm{S}_2(r) \tag{2}$$

It is now assumed that each source is an MQ (mass quantum, defined below) and where its Poynting vector is $\delta\mathrm{S}(r)$ and where r denotes the distance from the source. **E** and **H** decrease linearly with the distance r implying that $\delta\mathrm{S}(r)$ decreases with the square of the distance; $\delta\mathrm{S}(r) = \delta\mathrm{S}(0)/r^2$. The mutual force of attraction $\delta\mathrm{F}(r)$ between two MQs at distance r is

$$\delta\mathrm{F}(r) = \gamma \cdot \delta\mathrm{S} \cdot \delta\mathrm{S}/r^2. \tag{3}$$

It is now assumed that one MQ acts on p collocated MQs, belonging to the set **P** of all MQs at Source 2, and where each force can be described by $\delta\mathrm{F}(r)$. These MQs are synchronized resulting in the force

$$\sum_{p\in P} \delta\mathrm{F}(r) = \gamma \cdot \delta\mathrm{S} \cdot \sum_{p\in P} \delta\mathrm{S}\,/r^2 = \gamma \cdot (\delta\mathrm{S} \cdot \delta\mathrm{S}) \cdot p/r^2 \tag{4}$$

It is now assumed that n collocated MQs, belonging to the set **N** of all MQs at Source 1, act on p collocated MQs, belonging to the set **P** of all MQs at Source 2, resulting in the total force $\mathrm{F}(r)$,

$$\mathrm{F}(r) = \gamma \cdot \sum_{n\in N} \delta\mathrm{S} \cdot \sum_{p\in P} \delta\mathrm{S}/r^2 = \gamma \cdot (\delta\mathrm{S} \cdot \delta\mathrm{S}) \cdot n \cdot p/r^2 \tag{5}$$

$\delta\mathrm{S} \cdot \delta\mathrm{S}$ is a constant described by the intrinsic characteristics of the MQ, and hence Equation (5) is simpli-

Copyright © 2013 SciRes.

fied into

$$\mathrm{F}(r) = \xi \cdot n \cdot p / r^2 \qquad (6)$$

where ξ is a constant.

In the present study Equation (6) is called the law of gravity between two clusters, at distance r containing n respectively p MQs. Note that the cluster can be few MQs, an atom, a mass or a planet.

The implication of Equation (3) is that an atomic MQ generates gravity quantum and consequently, the present theory applies to quantum gravity.

It is proposed that elementary entities or more precisely mass quanta MQs, in the atomic nucleus, contain intrinsic, combined electric and magnetic fields or electromagnetic dipoles. Every MQ creates resonance and receives Gravity Source plane TEM waves. This results in emitted TEM waves. Two MQs and their emitted TEM waves create mutual force of attraction δF(r). This is illustrated with gravity between the earth and the sun. The earth contains n mass quanta (MQs) creating a force $n \cdot \delta \mathrm{F}(r)$ on every mass quanta (MQ) in the sun. The sun contains p mass quanta (MQs), thus the total force of attraction is $n \cdot p \cdot \delta \mathrm{F}(r)$ and that is equal to Equation (6). The earth's mass m_1 is proportional to the number of MQs on earth, *i.e.* $m_1 \sim n$, and the sun's mass $m_2 \sim p$. Equation (6) is then approximately equal to the Newtonian geometric law of gravity

$$\mathrm{F}_{\mathrm{Gravity}} = \xi \cdot n \cdot p / r^2 \approx \mathrm{F}_{\mathrm{Newton}} = \mathrm{G} \cdot m_1 \cdot m_2 / r^2 \qquad (7)$$

Note that F_G, G and the *constant* decreases with the distance to the Gravity Source.

3. Materials and Methods

All experiments were performed using one or many identical electronic devices called MQ. The MQ created resonance with plane TEM waves; it received and emitted plane TEM waves. Each MQ consisted of two neodymium magnets, with diameter 22 mm, spaced by a thin (0.05 mm) plastic foil. The south pole of one magnet attracted the north pole of the other. This created an internal magnetic field $\mathbf{B}_i$. One magnet was connected to the positive pole of a voltage source U. The other magnet was connected to the negative pole, **Figure 1**. This created an internal electric field $\mathbf{E}_i$, e.g. 10^6 V/m at $U = 50$ V. $\mathbf{B}_i$ and $\mathbf{E}_i$ were merged within the thin space between the magnets and could be described as electromagnetic dipoles. Measurements were performed with the neodymium magnets connected to a 9 V battery, forming a portable MQ, detached from external influence (e.g. AC, ground). This MQ was used in Experiments 5-10.

Experiments 1-4 and 11 were performed with the magnets connected to a low frequency signal generator with output voltage 0.5 mV RMS. Hence, the static electric field was replaced by an alternating electric field. The MQ created resonance at the signal generator frequency; it received and re-emitted plane TEM waves having this frequency. This device, also called MQ, was used in order to measure the frequencies of plane TEM waves originating from the Gravity Source.

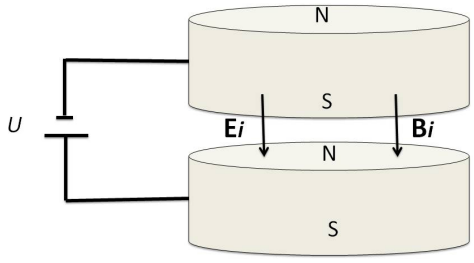


Figure 1. Generating TEM waves. Two neodymium magnets, with a diameter 22 mm, were spaced by a thin (0.05 mm) plastic foil with the south pole of one magnet attached to the north pole of the other. This created an internal magnetic field $\mathbf{B}_i$. One magnet was connected to the positive pole of a voltage source U, and the other magnet to the negative pole. This created an internal electric field $\mathbf{E}_i$, combined with $\mathbf{B}_i$. This device, called MQ received and re-emitted plane TEM waves.

Measurements of plane TEM wave field vector amplitudes were made as described in [7]. That report describes in detail how the position and amplitude of electric and magnetic field vectors were measured using a charge meter. The charge meter is described in detail. Measurements in the present study were made almost identically to those described in [7]. In summary the charge meter contained a probe with charge density ρ. The probe was moved, with constant speed, through the TEM wave electric field vectors **E**. This resulted in an electric body force $\rho\mathbf{E}$ on the probe charge density ρ. This current pulse was amplified and displayed. In this case **E** was field vectors, implying that div**E** was large and resulted in a distinct current pulse. This probe was also moved, with constant speed, through the TEM magnetic field vectors **B**, which resulted in a magnetic body force $\mathbf{J} \times \mathbf{B}$ on current **J** in the probe. In this case **B** was field vectors, implying that div**B** was large and resulted in a distinct current pulse. The measurement method had one important feature. The field vector's physical length was proportional to its amplitude and in the present study between 0.1 - 10 m. It was easy to measure and determine the physical length of the field vectors since div**E** and div**B** were large. Hence, this was an accurate measure of the relative amplitude. Consequently, all results are presented as relative amplitude.

TEM waves propagating in two directions, from Source 1 to Source 2 and from Source 2 to Source 1, were measured in the following way. A permanent magnet with magnetic field $\mathbf{B}_\mathbf{M}$ (1 μT) was inserted within the TEM wave's path, 1 m in front of Source 1. The propa-

Copyright © 2013 SciRes.

gating TEM wave consisted of electric **E** and magnetic **B** field vectors. The magnetic field $\mathbf{B_M}$ created a force on the positive magnetic field vectors $\mathbf{B}^+$ and on the negative magnetic field vectors $\mathbf{B}^-$. According to [3] it made the TEM wave's positive magnetic field vectors $\mathbf{B}^+$ divert the distance d_1 to one side and the negative magnetic field vectors $\mathbf{B}^-$ divert the distance d_2 to the opposite side. TEM waves propagating in one direction (e.g. from Source 1 to Source 2) were influenced so that their field vectors were diverted in the horizontal plane, see **Figure 2**, the electric field vectors **E** where oriented in the horizontal plane and the magnetic field vectors **B** were oriented in the vertical plane. TEM waves propagating in the opposite direction (e.g. from Source 2 to Source 1) were influenced so that their field vectors were diverted in the vertical plane, see **Figure 3**. The electric field vectors **E** where oriented in the vertical plane and the magnetic field vectors **B** were oriented in the horizontal plane. This enabled measurement of the two directions separately. Consequently, TEM waves propagating in opposite directions between two sources were rotated 90 degrees relative to each other. The field vectors propagated in smooth bows distanced by the magnetic field $\mathbf{B_M}$.

Measurements were performed 50 km south of Stockholm, Sweden.

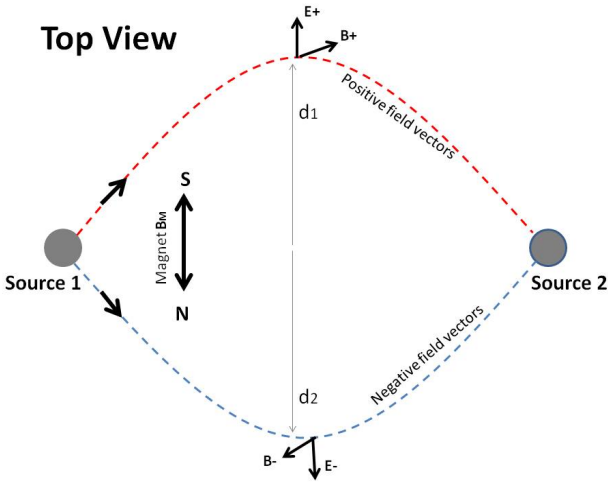


Figure 2. Measuring TEM waves propagating from Source 1. Two MQs, as described in Figure 1, Source 1 and Source 2, were positioned at 10 m distance. Source 1 was connected to a 9 V battery and Source 2 was also connected to a 9 V battery. This resulted in TEM waves propagating between the two sources. A permanent magnet $\mathbf{B_M}$ was inserted within the flow of TEM waves, 1 m in front of Source 1. TEM waves propagating from Source 1 to Source 2 were diverted in the horizontal plane. This resulted in that positive field vectors $\mathbf{E}^+$ and $\mathbf{B}^+$ were diverted the distance d_1 in the horizontal plane. The negative field vectors $\mathbf{E}^-$ and $\mathbf{B}^-$ were diverted the distance d_2 to the opposite side in the horizontal plane. This enabled separate measurement of TEM waves propagating from Source 1 to Source 2, respectively from Source 2 to Source 1.

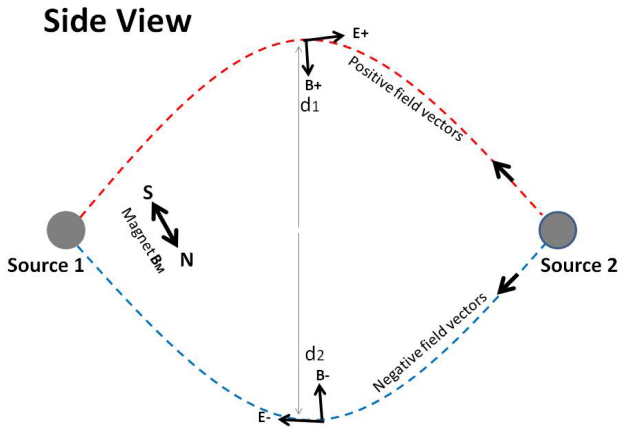


Figure 3. Measuring TEM waves propagating from Source 2. Two MQs, as described in Figure 1, Source 1 and Source 2, were positioned at 10 m distance. Source 1 was connected to a 9 V battery and Source 2 was also connected to a 9 V battery. This resulted in TEM waves propagating between the two sources. A permanent magnet $\mathbf{B_M}$ was inserted within the flow of TEM waves, 1 m in front of Source 1. TEM waves propagating from Source 2 to Source 1 were diverted in the vertical plane. This resulted in that positive field vectors $\mathbf{E}^+$ and $\mathbf{B}^+$ were diverted the distance d_1 in the vertical plane. The negative field vectors $\mathbf{E}^-$ and $\mathbf{B}^-$ were diverted the distance d_2 to the opposite side in the vertical plane. This enabled separate measurement of TEM waves propagating from Source 2 to Source 1, respectively from Source 1 to Source 2.

Analysis: the experiments were performed blindly, randomly and repeated 3 times.

4. Results

Experiments 1-4 were made with the MQ connected to the signal generator. The purpose with these experiments was to verify plane TEM wave frequency and its origin. Initially the signal generator frequency was increased from 1 to 100 Hz in increments of 0.1 Hz. TEM wave amplitude close to the MQ was measured as a function of frequency. It was observed that the MQ received and emitted plane TEM waves from the object being measured at two distinct frequencies; 69.9 Hz and 91.9 Hz. The signal generator was adjusted to these two frequencies in Experiments 1-4. It was also observed that the MQ always received and emitted plane TEM waves from the earth (*i.e.* ground). Hence, the MQ did not generate TEM waves; its only function was to create resonance at the plane TEM wave frequency.

Experiment 1. The purpose with this experiment was to measure the frequency of TEM waves received from the Gravity Source. The method was to create resonance (*i.e.* low impedance) in an MQ, connected to a signal generator, and to measure the amplitude of transmitted and absorbed TEM waves as function of frequency. The purpose was also to measure the direction of TEM waves

Copyright © 2013 SciRes.

received from and transmitted to the Gravity Source. One MQ, connected to a signal generator, was positioned on a table and this was called Source 1. The signal generator frequency was 69.9 Hz and then changed to 91.9 Hz. The generator frequency was carefully adjusted to maximum TEM wave field vector amplitude. 1‰ deviation in frequency resulted in approximately 50% change in TEM wave field vector amplitude, which illustrates the demand for accuracy. The MQ was activated at the frequency 69.9 Hz for 48 hours in order for the experiment to stabilize. It was observed that the field vector amplitude increased approximately 10 times during 48 hours. It also continued to increase after 48 hours by approximately 10% daily. It produced plane TEM_1 waves propagating from Source 1 towards north (earth's rotational axis). It followed a diurnal, counter-clockwise circle with approximately 20 degrees cone angel. Its peak altitude was approximately 75 degrees and occurred at noon (in Sweden) and the direction was towards the geographical north +/− 5 degrees. Plane TEM_2 waves propagated from a source in the universe towards Source 1, aligned along and merged with TEM_1. It was possible to measure TEM_1 and TEM_2 separately by inserting a magnet, with magnetic field $\mathbf{B_M}$ 1 m in front of Source 1, as explained in **Figure 2** and in **Figure 3**. The procedure is also explained in Experiment 10. Measurements displayed electric field vectors $\mathbf{E}^+$ and $\mathbf{E}^-$ and magnetic field vectors $\mathbf{B}^+$ and $\mathbf{B}^-$ perpendicular to the direction from Source 1 to Source 2 and from Source 2 to Source 1. In each direction the field vectors $\mathbf{E}^+$ and $\mathbf{B}^+$ respectively $\mathbf{E}^-$ and $\mathbf{B}^-$ were perpendicular, *i.e.* separated 90 degrees as in any plane TEM wave, e.g. radio wave [3]. **E** and **B** in one direction were rotated 90 degrees relative to **E** and **B** in the opposite direction, as described in **Figures 2** and **3**. In each direction the amplitudes were identical, *i.e.* $\mathbf{E}^+ = \mathbf{E}^-$ and $\mathbf{B}^+ = \mathbf{B}$; however, they varied between the two directions. Positioning an electric or magnetic field close to the field vectors changed their positions, confirming their nature and polarity. The method is described in [7]. The experiment was repeated, using generator frequency 91.9 Hz, displaying similar results. It is proposed that Source 2 was the Gravity Source generating plane TEM waves with the frequency 69.9 Hz and 91.9 Hz. The experiment displayed that the MQ, Source 1, received TEM_2 waves from the Gravity Source and that it produced TEM_1 waves which propagated in the direction towards the Gravity Source.

Experiment 2. The purpose with this experiment was to measure the frequency of TEM waves received from and transmitted to the sun. The method was to create resonance (*i.e.* low impedance) in an MQ, connected to a signal generator, and to measure the amplitude of transmitted and absorbed TEM waves as function of frequency. The purpose was also to measure the direction of TEM waves received from and transmitted to the sun. Experiment 1 was repeated; however, in this case TEM_1 and TEM_2 waves to and from the sun were measured at 69.9 Hz and then at 91.9 Hz. Measurements displayed results similar to Experiment 1, and where the only differences were the direction of the TEM waves and their amplitudes. TEM_1 and TEM_2 waves were always directed to and from the sun during daytime. The amplitude of plane TEM waves to/from the sun were less than 5% of the plane TEM waves to/from the Gravity Source, indicating the huge influence of the Gravity Source.

Experiment 3. The purpose with this experiment was to measure the frequency of TEM waves received from and transmitted to the moon. The method was to create resonance (*i.e.* low impedance) in an MQ, connected to a signal generator, and to measure the amplitude of transmitted and absorbed TEM waves as function of frequency. The purpose was also to measure the direction of TEM waves received from and transmitted to the moon. Experiment 1 was repeated; however, in this case plane TEM_1 and TEM_2 waves to and from the moon were measured at 69.9 Hz and then at 91.9 Hz. Measurements displayed results similar to Experiments 1 and 2, and where the only differences were the direction of the TEM waves and their amplitudes.

Experiment 4. The purpose with this experiment was to measure the frequency of TEM waves received from and transmitted to a mass. The method was to create resonance (*i.e.* low impedance) in an MQ, connected to a signal generator, and to measure the amplitude of transmitted and absorbed TEM waves as function of frequency. The purpose was also to measure the direction of TEM waves received from and transmitted to a mass. Experiment 1 was repeated; however, in this case plane TEM_1 and TEM_2 waves to and from a nearby object (e.g. 10 kg iron) were measured at 69.9 Hz and then at 91.9 Hz. Measurements displayed results similar to experiments 1, 2 and 3. The amplitude of $\mathbf{E}^+ = \mathbf{E}^-$ and $\mathbf{B}^+ = \mathbf{B}^-$ were proportional to the object's mass and inverse proportional to the objects distance r to the MQ, Source 1.

Experiments 5-7 aimed at verifying the theoretical model, presented in Section 2. One MQ or many collocated MQs were positioned at Source 1. Source 2 was located at the distance r = 10 m and one or many MQs were positioned at Source 2. All MQs were connected to a 9 V battery each. It produced plane TEM_1 waves propagating from Source 1 to Source 2 and plane TEM_2 waves propagating from Source 2 to Source 1. Measurements displayed electric field vectors $\mathbf{E}^+$ and $\mathbf{E}^-$ and magnetic field vectors $\mathbf{B}^+$ and $\mathbf{B}^-$ perpendicular to the direction from Source 1 to Source 2 and from Source 2 to Source 1. In each direction the field vectors $\mathbf{E}^+$ and $\mathbf{B}^+$ respectively $\mathbf{E}^-$ and $\mathbf{B}^-$ were perpendicular, *i.e.* separated 90 degrees as in any plane TEM wave, e.g. radio wave [3]. In each

Copyright © 2013 SciRes.

direction the amplitudes were identical, *i.e.* $\mathbf{E}^+ = \mathbf{E}^-$ and $\mathbf{B}^+ = \mathbf{B}^-$. It was also observed that Source 1 and Source 2 received and emitted plane TEM waves from the Gravity Source and from the earth (*i.e.* ground). Hence, the energy in TEM_1 and TEM_2 originated from these sources. Consequently, the MQs did not generate TEM waves; their only function was to create resonance with low frequency plane TEM waves and to re-emit them.

Experiment 5. The purpose of this experiment was to validate the theoretical model Equations (2) and (6); $F(r) = \gamma \cdot S_1(r) \cdot S_2(r) = \xi \cdot n \cdot p / r^2$. $S_1(r)$ and $S_2(r)$ were measured as a function of n and p. The purpose was also to simulate variation of n by varying the voltage U on one MQ. One MQ was positioned at Source 1 and one MQ at Source 2. The amplitude of the field vectors E^+ were measured in both directions (*i.e.* TEM_1 from Source 1 to Source 2 and TEM_2 from Source 2 to Source 1).

Then the number n of MQs at Source 1 was increased to 4 in steps of 1. Subsequently the number p of MQs at Source 2 was increased to 4 in steps of 1. Measurements displayed that the amplitude of the electric and magnetic field vectors $\mathbf{E}^+$, $\mathbf{E}^-$, $\mathbf{B}^+$ and $\mathbf{B}^-$ were proportional to $n \cdot p$ in both directions. Subsequently the MQ at Source 1 was connected to a variable voltage source and the voltage U on Source 1 was varied from 0 to 40 V in steps of 5 V. Measurements displayed that the amplitude of the electric and magnetic field vectors $\mathbf{E}^+$, $\mathbf{E}^-$, $\mathbf{B}^+$ and $\mathbf{B}^-$ were proportional to U in both directions, see **Figure 4**.

Experiment 6. The purpose of this experiment was to validate the theoretical model Equations (2) and (6); $F(r) = \gamma \cdot S_1(r) \cdot S_2(r) = \xi \cdot n \cdot p / r^2$. $S_1(r)$ and $S_2(r)$ were measured as a function of r. One MQ was positioned at Source 1 and one MQ at Source 2. The distance r between Source 1 and Source 2 was varied, starting with r = 1 m, then increased to 10 m in steps of 1 m. The amplitude of the field vectors $\mathbf{E}^+$, $\mathbf{E}^-$, $\mathbf{B}^+$ and $\mathbf{B}^-$ were measured in both directions as a function of the distance r. Measurements displayed that the amplitudes of the electric and magnetic field vectors $\mathbf{E}^+$, $\mathbf{E}^-$, $\mathbf{B}^+$ and $\mathbf{B}^-$ were inverse proportional to the distance r in both directions, see **Figure 5**.

Experiment 7. The purpose of this experiment was to verify that there was a transfer of energy from $S_1(r)$ to $S_2(r)$ vice versa according to Equation (2). One MQ was positioned at Source 1 and one MQ at Source 2 and the distance between them was 10 m. The amplitude of the field vectors $\mathbf{E}^+$, $\mathbf{E}^-$, $\mathbf{B}^+$ and $\mathbf{B}^-$ were measured at different distances from Source 1 along the straight direction from Source 1 to Source 2, and measured in both directions. Measurements displayed that the amplitude of the sum of electric and magnetic field vectors $\mathbf{E}^+$, $\mathbf{E}^-$, $\mathbf{B}^+$ and $\mathbf{B}^-$ from both directions (*i.e.* $TEM_1 + TEM_2$) were constant at every position between Source 1 and Source 2. $\mathbf{E}^+$, $\mathbf{B}^+$, $\mathbf{E}^-$ and $\mathbf{B}^-$ were spaced 90 degrees in both directions, as in any plane TEM wave.

In the following Experiments 8 and 9 an MQ, connected to a 9 V battery, was positioned at Source 1. Source 2 was a mass in Experiment 8 and the sun in Experiment 9.

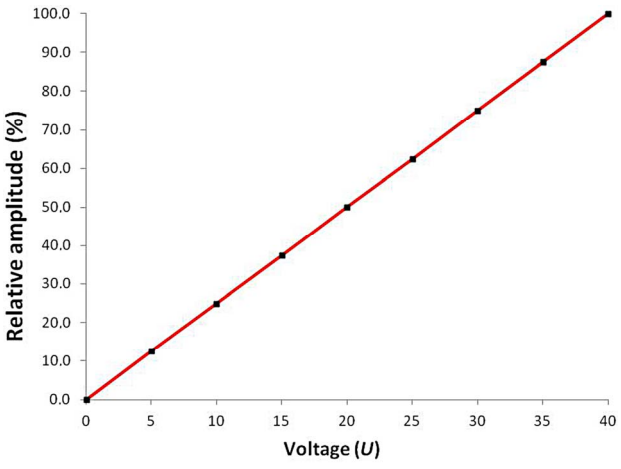


Figure 4. Field vector amplitude as function of voltage on one MQ. Two MQs, as described in Figure 1, Source 1 and Source 2, were positioned at 10 m distance. Source 2 was connected to a 9 V battery. The voltage *U* on Source 1 was varied from 0 - 40 V in steps of 5 V. The amplitude of the plane TEM_1 and TEM_2 wave's electric and magnetic field vectors ($\mathbf{E}^+$, $\mathbf{B}^+$, $\mathbf{E}^-$ and $\mathbf{B}^-$) was measured as a function of the voltage *U*. The amplitudes of all four field vectors were proportional to *U* and $\mathbf{E}^+ = \mathbf{E}^-$ and $\mathbf{B}^+ = \mathbf{B}^-$ in both directions. $\mathbf{E}^+$ amplitude in TEM_1 as function of *U* is displayed in Figure 4.

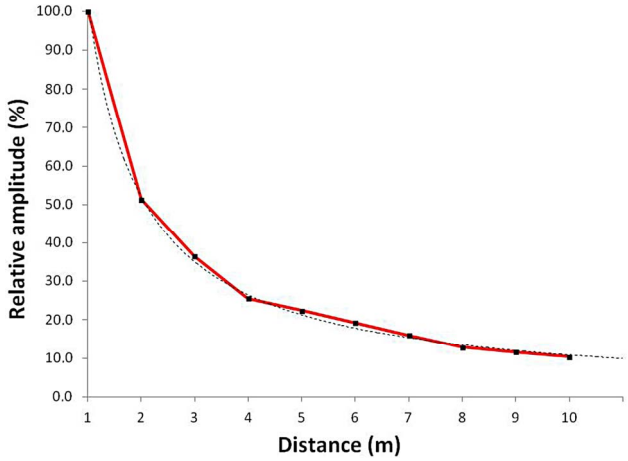


Figure 5. Field vector amplitude as a function of distance. Two MQs, described in Figure 1, were positioned at the distance *r*, which was initially 1 m and then increased to 10 m in steps of 1 m. The amplitude of TEM_1 and TEM_2 wave's electric and magnetic field vectors ($\mathbf{E}^+$, $\mathbf{B}^+$, $\mathbf{E}^-$ and $\mathbf{B}^-$) was measured as a function of the distance *r* in both directions. The amplitudes of all field vectors were proportional to1/r and $\mathbf{E}^+ = \mathbf{E}^-$ and $\mathbf{B}^+ = \mathbf{B}^-$ in both directions. $\mathbf{E}^+$ amplitude in TEM_1 as function of *r* is displayed in Figure 5. The dotted line represents the theoretical decline 1/*r*.

Copyright © 2013 SciRes.

Experiment 8. The purpose of this experiment was to validate the theoretical model Equation (2), Equations (6) and (7); $F(r) = \gamma \cdot S_1(r) \cdot S_2(r) = \xi \cdot n \cdot p / r^2 \approx F_{Newton} = G \cdot m_1 \cdot m_2 / r^2$. $S_1(r)$ and $S_2(r)$ were measured as a function of the mass m_2. A mass, m_2, (Source 2) which consisted of iron was positioned 5 m from the MQ at Source 1. The mass m_2 was increased from 0 to 10 kg in increments of 1 kg. Measurements displayed plane TEM_1 waves propagating from Source 1 to the mass m_2 and plane TEM_2 waves propagating from the mass m_2 to Source 1. The amplitude of the field vectors $\mathbf{E}^+$, $\mathbf{E}^-$, $\mathbf{B}^+$ and $\mathbf{B}^-$ were measured, in both directions, as a function of the mass m_2. Measurements displayed that the amplitude of the electric and magnetic field vectors $\mathbf{E}^+$, $\mathbf{E}^-$, $\mathbf{B}^+$ and $\mathbf{B}^-$ were proportional to the mass m_2 in both directions, see **Figure 6**. It was also observed that the mass m_2 received and emitted plane TEM waves to and from the Gravity Source and the earth (*i.e.* ground).

Experiment 9. The purpose with this experiment was to show that Equation (2), $F(r) = \gamma \cdot S_1(r) \cdot S_2(r)$, was only valid at steady state and to demonstrate how this steady state was achieved. Measurements were performed outside Stockholm, Sweden during the daytime. Source 1, consisting of one MQ, was positioned on a table. Measurement started by activating Source 1 (*i.e.* the MQ was connected to a 9 V battery). Measurements were performed 2 m in front of Source 1 and roughly in the direction of the sun. Measurements displayed plane TEM_1 waves propagating from Source 1 in direction towards the sun (Source 2) and plane TEM_2 waves propagating in the opposite direction. The amplitude of the field vectors $\mathbf{E}^+$, $\mathbf{E}^-$, $\mathbf{B}^+$ and $\mathbf{B}^-$ were measured as a function of time (both TEM_1 and TEM_2). Measurements displayed that the amplitude of TEM_1 electric and magnetic field vectors $\mathbf{E}^+$, $\mathbf{E}^-$, $\mathbf{B}^+$ and $\mathbf{B}^-$ increased approximately six times the initial amplitude within 16 minutes. Then they remained stable displaying no change in amplitude, see **Figure 7**.

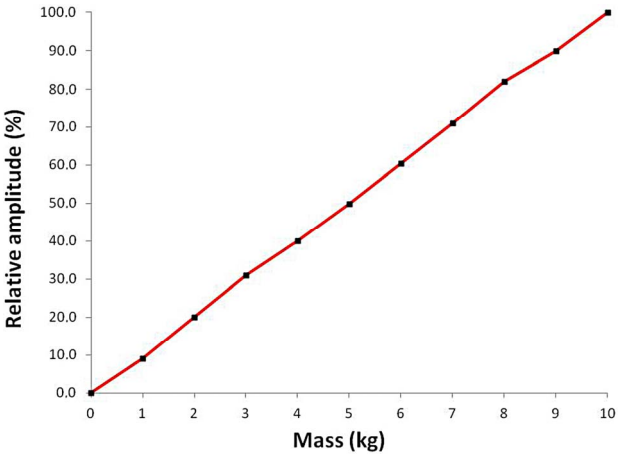


Figure 6. Field vector amplitude as a function of mass. One MQ (Source 1) was positioned 5 m from a mass which consisted of iron (Source 2). The mass was changed from 0 - 10 kg in steps of 1 kg. The amplitude of TEM_1 and TEM_2 wave's electric and magnetic field vectors ($\mathbf{E}^+$, $\mathbf{B}^+$, $\mathbf{E}^-$ and $\mathbf{B}^-$) was measured as a function of mass *m*. The amplitudes of all four field vectors were proportional to *m* in both directions and $\mathbf{E}^+ = \mathbf{E}^-$ and $\mathbf{B}^+ = \mathbf{B}^-$. $\mathbf{E}^+$ amplitude in TEM_1 as function of the mass *m* is displayed in Figure 6.

Measurement was also performed on plane TEM waves to and from the Gravity Source (after 48 hours). TEM wave amplitude of the Gravity Source was more than 20 times larger than the TEM wave amplitude towards the sun, indicating its large influence. Note that TEM wave amplitude from the Gravity Source continued to increase after 48 hours by approximately 10% daily. Hence, it is reasonable to assume that the influence from the Gravity Source was very large.

Experiment 10. The purpose of this experiment was to introduce a method which facilitated separate measurement of TEM waves travelling in opposite direction, as described in **Figures 2** and **3**. Experiments 1-9 were enhanced in the following way; a static magnetic field $\mathbf{B}_M$ (*i.e.* a 1 μT magnet) was inserted at the centre of the plane TEM_1 and TEM_2 waves, 1 m from Source 1, see **Figures 2** and **3**. It was observed that in Experiment 1,

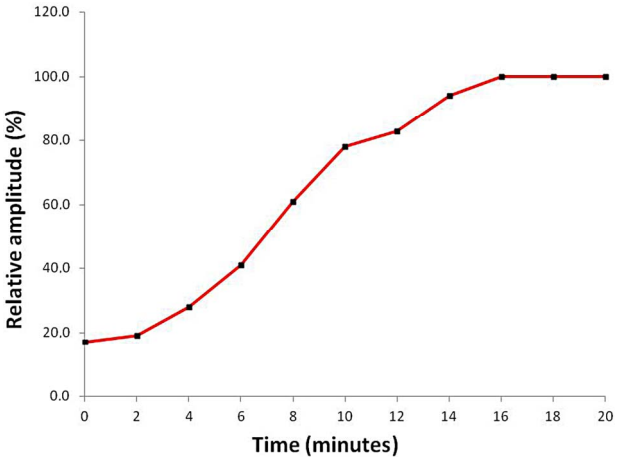


Figure 7. Field vector amplitude towards the sun as a function of time. One MQ, Source 1, was positioned on a table. The sun constituted Source 2. Measurements displayed plane TEM_1 waves propagating from Source 1 in direction towards the sun and plane TEM_2 waves propagating in direction from the sun to the MQ, Source 1. The amplitude of the plane TEM_1 and TEM_2 waves electric and magnetic field vectors ($\mathbf{E}^+$, $\mathbf{B}^+$, $\mathbf{E}^-$ and $\mathbf{B}^-$) was measured as a function of time. The amplitudes of field vectors were $\mathbf{E}^+ = \mathbf{E}^-$ and $\mathbf{B}^+ = \mathbf{B}^-$ at all times. After 16 minutes the amplitude was constant during an 8 hour measured period. The direction of plane TEM_1 waves and plane TEM_2 waves followed the position of the sun. After 16 minutes the length of TEM_1 field vectors were approximately 2 m at $U = 9$ V, indicating that the electric and magnetic field strength was significant and thus the field divergence divE was significant and easy to measure.

Copyright © 2013 SciRes.

TEM_1 was diverted much further than TEM_2, which provided a way to measure the plane TEM waves to and from the Gravity Source separately and to make a relative comparison between TEM_1 and TEM_2 amplitudes. It was also observed that in Experiment 9, TEM_1 was diverted much further than TEM_2, during the initial 8 minutes. It facilitated a way to measure the plane TEM waves to and from the sun separately and to make a relative comparison between TEM_1 and TEM_2 amplitudes.

Experiment 11. The purpose with this experiment was to measure extremely low frequency TEM waves received from the Gravity Source. The method was to create resonance (*i.e.* low impedance) in an MQ, connected to a signal generator, and to measure the amplitude of the absorbed TEM wave. One MQ, connected to a signal generator, was positioned on a table and this was called Source 1. The signal generator frequency was changed from 0.1 Hz to 1 Hz in steps of 0.01 Hz. The generator voltage was 0.5 mV RMS. It was observed that the MQ created resonance at 0.74 Hz and received plane TEM_2 waves from the direction of the Gravity Source. It was also observed that the MQ radiated 0.74 Hz plane TEM_1 waves to every type of matter, Source 2, in its vicinity. The amplitude of TEM_1 was proportional to the mass of Source 2. Furthermore, the amplitude of TEM_1, at Source 1, decreased linearly with the distance to Source 2. It was also observed that the amplitude of TEM_1 along the straight line between Source 1 and Source 2 decreased linearly with the distance from Source 1. It was not possible to observe any TEM_2 waves propagating from matter, Source 2, to the MQ, Source 1. It was concluded that the Gravity Source radiated plane TEM waves at 0.74 Hz and that this energy was absorbed by matter and probably re-emitted uniformly. The amplitude of the 0.74 Hz TEM wave, absorbed by matter, was significantly higher than the amplitude of 69.9 Hz and 91.9 Hz TEM waves propagating to the same matter (approximately 10 times larger).

It was observed in Experiments 1-11 that vertically propagating plane TEM waves (e.g. to/from the earth) resulted in that magnetic field vectors were oriented north to south caused by influence from the geomagnetic field. Otherwise the magnetic field vectors were oriented vertically, also the result of influence from the geomagnetic field. Consequently, the plane TEM waves were polarized which explains the efficiency of the measurement method used in Experiments 1-11.

5. Discussion

Experiments 2 and 9 displayed that gravity between the electronic device MQ (**Figure 1**) and the Gravity Source was much larger than gravity between the electronic device MQ and the sun. From Section 4 and 5 follows that gravity between the electronic device MQ and the Gravity Source was significant. It is proposed that the MQ within the atom (atomic MQ) behaves similar to the electronic device MQ. Consequently, the atomic MQ was exposed to two different gravity components, gravity between the atomic MQ and all other atomic MQs in the universe and gravity between the atomic MQ and the Gravity Source. It is proposed that the former is called gravitational gravity and the latter is called inertial gravity. The inertial gravity results in that the universe rotates (spirals) around the Gravity Source. Exposing the MQ to force accelerates the MQ relative to all other MQs and the Gravity Source. It can be illustrated with an MQ travelling through space, far away from planets and stars. It is primarily influenced by TEM waves from and to the Gravity Source. Its reference frame is the Gravity Source or more precisely TEM energy at MQ proximity. Its energy state relative to its reference frame (*i.e.* TEM energy generated by the Gravity Source) is changed when the MQ is accelerated and hence $\delta \mathrm{F} = \eta \cdot \mathrm{MQ} \cdot a$. a is the acceleration created by the quantified force δF on the inertial mass quantum $\eta \cdot \mathrm{MQ}$. A mass on earth experiences two reference frames. The first is the earth and its gravity and the second is the Gravity Source and its gravity. Hence, this mass m can be described as inertial mass, $m_i = \eta \cdot \sum \mathrm{MQ}$ and gravitational mass, $m_g = \beta \cdot \sum \mathrm{MQ}$ according to Equations (6) and (7). η and β are constants. This explains why, according to classical mechanics, the gravitational mass is proportional to the inertial mass. It is proposed that mass does not exist. Mass is only the observation of electromagnetic energy (*i.e.* TEM waves originating from the Gravity Source) operating on electric and magnetic fields, *i.e.* $\sum \mathrm{MQ}$ in atoms In that case Einstein's mass-energy equivalence may relate to the MQ, $\mathrm{E} = mc^2 = \zeta \cdot \sum \mathrm{MQ}$, where ζ is a constant.

The atom is in traditional physics described by its particles, quarks, gluon, graviton, baryon, proton, electron etc. and by its forces, the strong interaction, the weak interaction, electromagnetism and gravity, described in the Standard Model [8,9]. The present paper describes the atomic MQ (mass quantum) as an atomic entity creating gravity and inertia. The MQ consists of a static electric field in parallel with a magnetic field.

In its simplest form the MQ may consist of only two charges, a positive and a negative elementary charge forming an electric dipole. This dipole absorbs plane TEM waves from the Gravity Source with the frequency 0.74 Hz and this induces an electric and magnetic field with the frequency 0.74 Hz. Hence, this dipole consists of a static electric field and an almost static magnetic field. Consequently, the dipole can be regarded as an electromagnetic dipole which also creates resonance with plane TEM waves (gravity waves) at 69.9 Hz and 91.9 Hz. This dipole is called MQ in the present paper and it

Copyright © 2013 SciRes.

is the atom's elementary entity and also the atom's mass quantum.

The dipole (MQ) has electric and magnetic fields, electric and magnetic dipole moments, it produces gravity (mass) and it produces inertia (inertial mass). The force on one dipole (MQ) in the atom from all other dipoles (MQs) in the atom is entirely electromagnetic, *i.e.* caused by electromagnetic fields. The fields operating on the dipole consist of the vector sum of electric fields (Coulomb field) from all dipole's charges, the vector sum of electric dipole moments from all dipoles, the vector sum of induced electromagnetic fields (at 0.74 Hz) from all dipoles and the vector sum of induced electromagnetic dipole moments (at 0.74 Hz) from all dipoles.

The (induced) magnetic force (at 0.74 Hz) dominates within the atom's nucleus. It is proposed that this magnetic force represents what is commonly called gluon and the strong interaction. The electric force (Coulomb force) is strong within a dipole and between adjacent dipoles. Hence, the electric force can contribute to the strong interaction. However, within the atom's nucleus the electric force is normally smaller than the magnetic force. The force from electric and magnetic dipole moments is probably small. It is proposed that the electric force and the dipole moments normally represent what is commonly called the weak interaction and electromagnetism. The conclusion is that the atom can be described by configurations of MQs (dipoles) and elementary charges and by electromagnetic fields creating two forces; the magnetic force and the electric force. In addition the atomic MQs create resonance with gravity waves (TEM waves) and re-emit them. The flow of gravity waves between two MQs creates mutual force of attraction, *i.e.* gravity. Thus gravity is an electromagnetic force.

Positioning two MQs (dipoles) side-by-side; however, in opposite directions, results in strong magnetic and electric forces which fuses the MQs. MQs can be added forming a matrix. Such configuration can be described as a quark and various configurations can be described as various types of quarks.

Amalgamated quarks result in baryons, e.g. neutron. A proton is created when a positive elementary charge is added to one MQ in a neutron. An electron is created when a negative elementary charge is added to one MQ in a small configuration of MQs (e.g. a dipole with one positive charge and two negative charges). The sum of these configurations of MQs is the atom.

Consequently, the atom may consist of only positive and negative elementary charges configured into MQ, electron, quark, neutron, proton etc. Hence, the universe may consist of only positive and negative elementary charges and different configurations of these charges. All forces in the universe are electromagnetic forces induced by the position and motion (*i.e.* velocity and acceleration) of these elementary charges and their configurations, according to the laws of electromagnetism. The Gravity Source may consist of rotating charges and its generated TEM waves are the main provider and carrier [6] of energy within the universe. This model enables description of phenomena within the universe (e.g. dark energy, dark matter, expanding space, black hole, gravitational lensing and redshift) based on the laws of electromagnetism.

Equation (2) $\mathrm{F}(r)=\gamma\cdot\mathrm{S}_1(r)\cdot\mathrm{S}_2(r)$ describes the mutual force of attraction when the positions of the two sources are static relative to each other. When one source rotates relative to the other a force perpendicular to r is introduced and this force and the labor performed are minimized when the source's rotational axis is directed towards the other source. Experiment 1 displayed that the direction towards the Gravity Source was to the north, aligned to the earth's rotational axis; however, with a small misalignment. It is known that the Milky Way galaxy and its Galactic Centre contains the astronomical radio source Sagittarius A*. It has also been proposed that the Galactic Centre contains a super massive black hole. The earth's rotational axis is oriented towards the Galactic Centre; however, with a small misalignment [10,11]. Data from the WMAP mission display that distribution of energy in universe is anisotropic and may be aligned relative to the earth's rotational axis [12]. This anisotropic distribution of energy has recently (21 March 2013) been confirmed by data released from the Planck mission. Perhaps the earth's rotational axis and the Milky Way are oriented towards a concentration of mass in the universe and this concentration of mass is the Gravity Source. The universe expands and perhaps the Milky Way and the earth moves in direction from the Gravity Source, *i.e.* in the direction of the earth's geographical south? Compiling the present findings with known galaxy filaments may provide improved understanding of the universe.

It is proposed that there are three quanta. The positive and the negative elementary charges are the Charge Quantum. The atomic dipole is the Mass Quantum. The gravity wave and photon is the carrier of energy. The photon consists of electromagnetic energy superpositioned on gravity waves [6]. The gravity wave is the Energy Quantum. All three can be included in (modified) quantum physics. The fundament of modified quantum physics may comprise Charge Quantum, Energy Quantum and Mass Quantum operating on the laws of electromagnetism.

6. Conclusions

The study proposes that gravity is created by plane TEM (transverse electromagnetic) waves, which is the solution to Maxwell's equations $\operatorname{curl}\mathbf{H}=\boldsymbol{i}+\partial\mathbf{D}/\partial t$ and

Copyright © 2013 SciRes.

$\mathrm{curl}\,\mathbf{E} = -\partial\mathbf{B}/\partial t$ [3]. The gravitational force or mutual force of attraction between two synchronized plane TEM waves travelling in opposite direction is proportional to the scalar product of their Poynting vectors [4] according to Equation (2).

It is proposed that a gravitational singularity called the Gravity Source emitted plane TEM waves to elementary entities, or more precisely, to mass quanta, MQs, within the atom's nucleus. It is proposed that the MQ contained intrinsic, merged electric and magnetic fields or electromagnetic dipoles or one electric dipole. These almost static merged fields created resonance with the plane TEM wave's almost static field vectors. The received plane TEM waves were re-emitted as plane TEM waves and distributed to, in principle, every MQ in the universe and where the amplitude decreased linearly with the distance to the MQ. The content of these plane TEM waves remained unchanged. Simultaneously every MQ in the universe received plane TEM waves from, in principle, every other MQ in the universe. The plane TEM wave amplitude was inverse proportional to the distance between the MQs. Energy emitted by the Gravity Source was partly re-emitted to the Gravity Source. Hence, every MQ was also exposed to gravity from the Gravity Source.

The mutual force of attraction between two MQs is described by Equation (3). The mutual force of attraction between two masses (*i.e.* the sum of contributions from many MQs) is described by Equation (6), which is approximately equal to the Newtonian geometric law of gravity according to Equation (7).

The plane TEM wave frequency was measured in Experiments 1-4 using a device, MQ, which created resonance at a predetermined frequency; it received and re-emitted plane TEM waves at one particular frequency. It consisted of an internal alternating electric field $\mathbf{E}_i$ with frequency 0.74 Hz, 69.9 Hz or 91.9 Hz, merged with an internal static magnetic field $\mathbf{B}_i$.

Experiment 1. This experiment confirmed the existence of the Gravity Source generating plane TEM waves with the frequency 69.9 Hz and 91.9 Hz. The direction towards the Gravity Source was to the north and followed a diurnal, counter-clockwise circle with approximately 20 degrees cone angle. It was assumed that the diurnal change of direction was caused by the rotation of the earth and where the earth's axis was misaligned 10 degrees relative to the direction towards the Gravity Source.

Experiment 2. This experiment displayed that plane TEM wave's amplitude to/from the Gravity Source was at least twenty times the plane TEM wave's amplitude to/from the sun, confirming its large influence.

Experiments 2-4. These experiments confirmed that the sun, the moon, the earth and an object (e.g. iron) positioned in close proximity to the MQ, produced and received synchronized and mutually coupled plane TEM waves with frequency 69.9 Hz and 91.9 Hz. It indicated that these objects produced 69.9 Hz and 91.9 Hz plane TEM waves and that these plane TEM waves were synchronized to the Gravity Source's plane TEM waves.

Experiment 5. The purpose of this experiment was to validate the theoretical model Equation (2) and Equation (6); $\mathrm{F}(r) = \gamma \cdot \mathrm{S}_1(r) \cdot \mathrm{S}_2(r) = \xi \cdot n \cdot p/r^2$. $\mathrm{S}_1(r)$ and $\mathrm{S}_2(r)$ were measured as a function of n and p. n respectively p MQs were positioned at a distance r. Results displayed that $\mathbf{E}^+$, $\mathbf{B}^+$, $\mathbf{E}^-$ and $\mathbf{B}^-$ amplitudes were proportional to $n \cdot p$ confirming Equations (2) and (6) and also Equations (3)-(5). This was simulated by varying the voltage U on Source 1, resulting in proportional change in field vector amplitude and within the measurement accuracy as displayed in **Figure 4**.

Experiment 6. The purpose of this experiment was to validate the theoretical model Equation (2) and Equation (6); $\mathrm{F}(r) = \gamma \cdot \mathrm{S}_1(r) \cdot \mathrm{S}_2(r) = \xi \cdot n \cdot p/r^2$. $\mathrm{S}_1(r)$ and $\mathrm{S}_2(r)$ were measured as a function of r. In this experiment the distance r between MQs was varied, displaying that field vector amplitudes $\mathbf{E}^+$, $\mathbf{B}^+$, $\mathbf{E}^-$ and $\mathbf{B}^-$ were proportional to $1/r$ confirming Equations (2)-(6). Measured data was in good agreement with the theoretical decline $1/r$, displayed in **Figure 5**. (Note that field vectors decreased with $1/r$ while energy decreased as $1/r^2$).

Experiment 7. Plane TEM waves propagating between two MQs displayed that the sum of TEM_1 and TEM_2 field vector amplitudes were constant all the way from one MQ to the other MQ. It indicated the mutual coupling between synchronized plane TEM_1 and TEM_2 waves propagating in opposite direction. It is proposed that there was a transfer of energy between the two TEM waves. This transfer of energy and the back reaction force that the plane TEM waves created on its source created mutual force of attraction or gravity according to Equation (2). Consequently, Experiment 7 supports the theoretical model.

Experiment 8. The purpose of this experiment was to validate the theoretical model Equation (2), Equation (6) and Equation (7); $\mathrm{F}(r) = \gamma \cdot \mathrm{S}_1(r) \cdot \mathrm{S}_2(r) = \xi \cdot n \cdot p/r^2 \approx \mathrm{F}_{\mathrm{Newton}} = \mathrm{G} \cdot m_1 \cdot m_2/r^2$. $\mathrm{S}_1(r)$ and $\mathrm{S}_2(r)$ were measured as a function of the mass m_2.This experiment proved that an object (mass) produced gravity proportional to its mass m, confirming Equations (6) and (7). This experiment showed that a mass produced plane TEM waves identical the Gravity Source, the sun, the moon and also the electronic device used in Experiments 1-7. This indicates that the mass, the sun and the moon contained similar type of mechanism of intrinsic nature and similar to the function of the electronic device called MQ. It is proposed that the mass, the sun and the moon contained an elementary entity or mass quantum which created resonance and

Copyright © 2013 SciRes.

re-emitted gravity TEM waves originating from the Gravity Source. It is proposed that the mass quantum consisted of an intrinsic electric field or dipole merged with a magnetic field or dipole.

Experiment 9. The purpose with this experiment was to show that Equation (2), $F(r)=\gamma\cdot S_1(r)\cdot S_2(r)$, was only valid at steady state and to demonstrate how this steady state was achieved. It took approximately 16 minutes to stabilize the TEM_1 waves propagating from Source 1 (the MQ) to Source 2 (the sun) and the TEM_2 waves propagating from Source 2 to Source 1. The field vectors reached their maximum amplitude after approximately 16 minutes. The round trip to the sun was 16.6 minutes at the speed of light. It can be explained as follows. The MQ at Source 1 was switched on at time t = 0 minutes. After 8.3 minutes (t = 8.3) TEM_1 waves reached the sun creating the force $F(r)=\gamma\cdot S_1(r)\cdot S_2(r)$ at the sun. Then it took another 8.3 minutes (t = 16.6) before the influenced TEM_2 waves reached Source 1. The two synchronized TEM_1 and TEM_2 waves created mutual force of attraction between Source 1 and Source 2, and full gravity (at t = 16.6). The experiment confirmed the mutual exchange of energy between the two sources according to Equation (2) and also that plane TEM waves (off course) propagated with the speed of light.

Experiment 10 displayed that it was possible to separate TEM_1 waves from TEM_2 waves and that it was possible to measure the two directions separately.

Experiment 11 displayed that investigated matter contained some type of mechanism that absorbed plane TEM waves at 0.74 Hz, originating from the Gravity Source. Hence, TEM waves at 0.74 Hz were absorbed by matter and its atoms. Consequently, the absorbed TEM wave energy must induce electric and magnetic fields with the frequency 0.74 Hz within atoms.

The conclusion is that the universe contained a gravitational singularity called the Gravity Source. It generated plane TEM waves with the frequency 69.9 Hz and 91.9 Hz. Atomic mass quanta, MQ, consisting of intrinsic, merged static electric and magnetic fields or dipoles created resonance with these plane TEM waves. They were then re-emitted to other atomic MQs as plane TEM waves with the frequency 69.9 Hz and 91.9 Hz. Atomic MQs in the universe were synchronized to the Gravity Source which resulted in that received and emitted plane TEM waves were synchronized, causing exchange of energy and mutual force of attraction, which is called gravity. The atom consisted of many MQs. This resulted in an equation called the electromagnetic law of gravity, Equation (6). It was compared with the Newtonian geometric law of gravity, Equation (7). In addition the Gravity Source radiated plane TEM waves at 0.74 Hz, which was absorbed by matter and the absorbed plane TEM wave amplitude was linear to the mass, indicating an intrinsic atomic mechanism.

The following offers a possible explanation to the origin of the Gravity Source. It is proposed that the Gravity Source consisted of a kernel of charge rotating at three speeds (*i.e.* T = 1/0.74 s, T = 1/69.9 s and T = 1/91.9 s), constituting a gigantic electromagnetic generator. It emitted photons and their plane TEM waves with frequency 0.74 Hz, 69.9 Hz and 91.9 Hz.

Experiments 1-11 can also be summarized as follows. Gravity consisted of extremely low frequency plane TEM waves. Using state of the art measurement technique would require unrealistically long antennas. The method was to create plane TEM wave spatial divergence, which was created with the electronic device called MQ and polarization created by the geomagnetic field. The plane TEM wave spatial divergence created TEM wave field vector divergences div**E** and div**B** which were measured with a charge meter. This innovative measurement technique enabled unbundling of gravity.

It is proposed that the present study unifies gravity, described by Equation (6), with electromagnetic theory and quantum physics. Hence, this electromagnetic theory of gravity supplements Newtonian geometric theory of gravity, relativistic metric theory of gravity, Newton's second law of motion and Einstein's mass-energy equivalence.

REFERENCES

[1] J. H. Schwarz, "From Superstrings to M Theory," *Physics Reports*, Vol. 315, 1999, pp. 107-121. doi:10.1016/S0370-1573(99)00016-2

[2] O. Aharony and E. Silverstein. "Supercritical stability, transitions, and (pseudo)tachyons," *Physical Review D*, Vol. 75, No. 4, 2007. doi:10.1103/PhysRevD.75.046003

[3] B. I. Bleaney and B. Bleaney, "Electricity and Magnetism," Oxford University Press, London, 1965.

[4] D. B. Melrose and R. C. McPhedran, "Electromagnetic Processes in Dispersive Media," Cambridge University Press, Cambridge, 1991.

[5] S. S. Afshar, E. Flores, K. F. McDonald and E. Knoesel, "Paradox in Wave-Particle Duality," *Foundations of Physics*, Vol. 37, No. 2, 2007, pp. 295-305. doi:10.1007/s10701-006-9102-8

[6] H. W. Giertz, "Photons Are EM Energy Superpositioned on TEM Waves," *Open Journal of Microphysics*, in press, 2013.

[7] H. W. Giertz, "Extremely Low Frequency Electromagnetic Energy in the Air," *Journal of Atmospheric and Solar-Terrestrial Physics*, Vol. 72, No. 9-10, 2010, pp. 767-773. doi:10.1016/j.jastp.2010.03.022

[8] J. F. Donoghue, E. Golowich and B. R. Holstein, "Dynamics of the Standard Model," Cambridge University Press, Cambridge, 1994.

[9] S. Braibant, G. Giacomelli and M. Spurio, "Particles and

Copyright © 2013 SciRes.

Fundamental Interactions: An Introduction to Particle Physics," Springer Verlag, Heidelberg, 2009.

[10] D. C. Backer and R. A. Sramek, "Proper Motion of the Compact, Nonthermal Radio Source in the Galactic Center, Sagittarius A*," *The Astrophysical Journal*, Vol. 524 No. 2, 1999, pp. 805-815. doi:10.1086/307857

[11] S. S. Doeleman, *et al*., "Event-Horizon-Scale Structure in the Supermassive Black Hole Candidate at the Galactic Centre," *Nature*, Vol. 455, No. 7209, 2008, pp. 78-80. doi:10.1038/nature07245

[12] G. Hinshaw, *et al*., "Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Temperature Analysis," *The Astrophysical Journal Supplement Series*, Vol. 170, No. 2, 2007, pp. 288-334. doi:10.1086/513698

Copyright © 2013 SciRes.

Squeezed vacuum from a monolithic optical parametric oscillator

G. Breitenbach, T. Müller, S. F. Pereira, J.-Ph. Poizat,* S. Schiller, and J. Mlynek

Fakultät für Physik, Universität Konstanz, D-78434 Konstanz, Germany

Received April 10, 1995; revised manuscript received June 28, 1995

We have developed a stable, reliable source of continuous-wave squeezed vacuum at 1064 nm, using a subthreshold monolithic lithium niobate optical parametric oscillator. Quantum-noise reduction below the vacuum noise level in a bandwidth of 30 MHz with a maximum of 5.5 ± 0.2 dB at 2 MHz is demonstrated. The reconstruction of the Wigner function of a continuous-wave, strongly squeezed-vacuum state is obtained.
© 1995 Optical Society of America

1. INTRODUCTION

Squeezed states of light have been successfully demonstrated in several different schemes, for both phase and amplitude quadratures of the electromagnetic field.[1] Among these schemes, the optical parametric oscillator (OPO) operating both below and above threshold has been verified to be a very good squeezing generator. In addition to the demonstration of a high degree of squeezing,[2–4] these devices have been used in several applications such as sub-shot-noise interferometry and spectroscopy, noise reduction in optical amplification, and back-action-evading measurements.[5–8] It has also been suggested that the solution of many challenging problems in quantum mechanics, such as quantum nondemolition measurements, noiseless amplification, and Schrödinger cats, can be pursued with OPO's.[9–11] For most of these experiments and proposals, OPO's with good stability and low loss are essential.

To realize a compact, stable, low-loss system for the generation of squeezed states of light, we employ a monolithic resonator. The monolithic resonator is a simple device that consists of a nonlinear crystal coated on both faces with reflecting dielectric coatings. Such a design ensures good mechanical stability and small losses. Furthermore, because of the short length of the cavity, the linewidth of such a resonator is broad compared with those at discrete cavities, which implies a broad bandwidth in the noise reduction. The performance of monolithic resonators has been successfully tested for second-harmonic generation and frequency downconversion and for the generation of nonclassical light such as twin beams and bright squeezing in second-harmonic generation.[12–14]

In this paper we combine these techniques to build a compact continuous-wave source of squeezed vacuum at 1064 nm. Our research is partially motivated by the possible application of such a squeezed-light source in planned gravitational wave interferometers and ultrasensitive polarimeters, which will operate at this wavelength.[15,16] In addition to the measurement of broadband squeezing, we applied the technique of optical homodyne tomography[17,18] to reconstruct, for the first time to our knowledge, the Wigner function of a strongly squeezed state in the continuous-wave regime.

The paper is organized as follows. In Section 2 we describe the experimental setup and characterize the OPO. In Section 3 we present the measurements of squeezing and the reconstruction of the Wigner function. Conclusions and future prospects are given in Section 4.

2. EXPERIMENTAL SETUP AND CHARACTERIZATION OF THE OPO

A general overview of the experimental setup is shown in Fig. 1. A frequency-doubled diode-pumped Nd:YAG laser (Lightwave Electronics Model 122) pumps a monolithic degenerate type-I OPO whose output is analyzed in a balanced homodyne detector.

The OPO is a 7.5-mm-long standing-wave resonator consisting of a $LiNbO_3$ crystal doped with 5% MgO. The end faces of the crystal are polished spherically, with 10-mm radii of curvature. One end of the crystal is coated with a high reflector at both 1064 and 532 nm; the other side is the output coupler with transmission $T = 2.1\%$ at 1064 nm and high transmission at 532 nm. With this configuration, the nonresonant pump double passes the resonator to enhance the nonlinear coupling, so that the threshold is reduced. The faces perpendicular to the crystal c axis are coated with gold for electro-optic modulation. The resonator is embedded in an aluminum oven whose temperature is actively controlled to a few millikelvins. The measured finesse of the OPO is 265, which, considering the free spectral range of 9 GHz, gives a linewidth of $2\Gamma = 34$ MHz (FWHM).

Crucial for the suitability of a resonator for the generation of strong squeezing is its escape efficiency $\rho = T/(T + A)$, where T is the transmission of the output coupler and A is the total internal round-trip loss of the resonator. The value of ρ is equal to the maximum achievable degree of squeezing. For a monolithic resonator, it is not easy to measure the propagation losses A and the output coupler transmission T separately. A

0740-3224/95/112304-06$06.00 © 1995 Optical Society of America

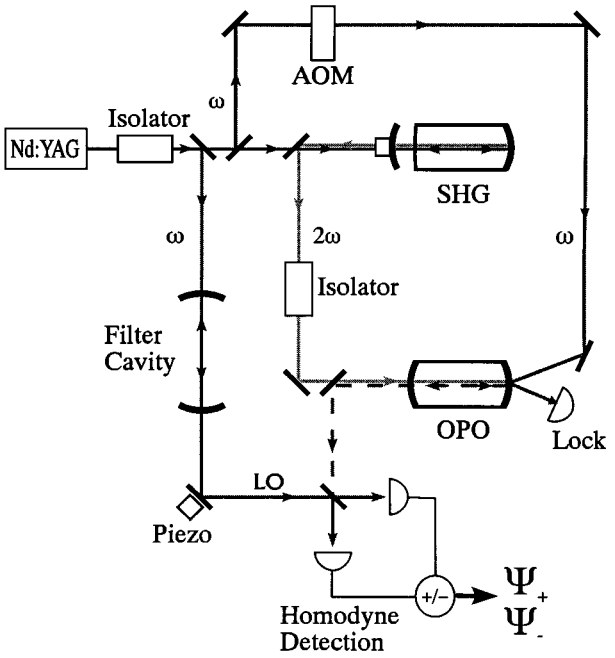


Fig. 1. Experimental scheme for the generation of squeezed vacuum with a monolithic $LiNbO_3$ resonator. AOM, acousto-optic modulator; SHG, second-harmonic generator; LO, local oscillator; piezo, piezoelectric mirror actuator.

direct way to determine the escape efficiency consists of measuring the conversion efficiencies for resonant second-harmonic generation or for parametric downconversion. For both processes the maximum conversion efficiency is equal to ρ. In previous experiments we measured maximum conversion efficiencies of 82% and 84% for frequency doubling and downconversion, respectively, at pump powers near 110 mW.[13,19] Correcting these values for the imperfect mode-matching results in an escape efficiency ρ of 88%. Assuming perfect detection efficiency, the best degree of squeezing possible with this resonator is 9.5 dB.

A basic requirement for our system is temporal stability of the squeezing. This condition can be achieved by active stabilization, whereby the laser frequency is locked to the OPO cavity resonance and the frequency-doubling cavity is locked to the laser frequency. For the stabilization of the laser frequency a probe beam at 1064 nm is frequency shifted by 180 MHz with an acousto-optic modulator and is injected through the high-reflector end of the monolithic standing-wave cavity into a higher-order transverse mode, such that the laser frequency (the subharmonic of the OPO's pump frequency) coincides with the resonator's TEM_{00} mode. Thus degenerate operation of the OPO is ensured. An error signal for the feedback to the laser is generated by a modified Pound–Drever technique,[20] in which the OPO crystal is modulated at 14.5 MHz, and the ensuing amplitude modulation of the probe wave is measured in reflection by a rf photodetector (see Fig. 1).

In fact, the high frequency stability of the laser and the OPO and the large OPO linewidth allow us to keep the laser frequency in resonance manually for several minutes. Most measurements shown here were therefore done without active stabilization of the laser.

The frequency-doubling cavity is independently locked to the laser frequency with a similarly generated error signal controlling the cavity length by means of a piezoelectric mirror actuator.

The squeezed vacuum at the output of the OPO is combined with a strong local oscillator in a balanced homodyne detector.[21] As photodetectors we used passivated Epitaxx ETX500 InGaAs photodiodes with the covers removed. We measured the quantum efficiency to be 97%, using a Laser Instrumentation thermopile for calibration. The local oscillator is split off from the laser source and is filtered by a high-finesse cavity to reduce the excess amplitude noise of the diode-pumped Nd:YAG laser at frequencies of a few megahertz. The filter cavity is 50 cm long, with a bandwidth of 140 kHz. Figure 2 shows the spectral densities $\Psi_{\pm}$ of the photocurrents $i_{\pm}$ from the plus and the minus junctions of the homodyne detector for the beams before and after the filter cavity. A comparison of the traces shows that the excess noise of the laser is strongly suppressed by the filter cavity, so that the local oscillator is shot-noise limited at frequencies ≥ 1 MHz. The suppression of a strong coherent modulation applied to the local oscillator is better than 20 dB. The shot-noise level, determined by comparison of Ψ_+ and Ψ_- when the open port of the beam splitter has vacuum input, is accurate to 0.3 dB for a wide range of frequencies. When the balancing of power in the two detectors is optimized for a particular frequency, the accuracy is of the order of 0.2 dB. With a local oscillator power of 2 mW the shot-noise level is 14 dB above the electronic noise level of the detectors at lower frequencies and 5 dB for frequencies above 24 MHz.

The alignment as well as the mode matching of the pump into the OPO were optimized by operation of the OPO above threshold, as described in Ref. 19. Further optimization of the pump alignment and the OPO temperature was done by measurement of the parametric amplification of a small signal at 1064 nm injected into the high-reflector port of the cavity. With pump powers near the threshold power of 28 mW, parametric gain as high as 170 was observed.

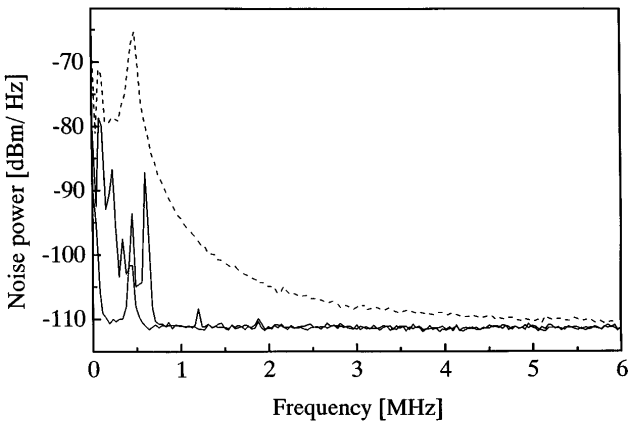


Fig. 2. Spectral density of the amplitude noise of the diode-pumped Nd:YAG laser as a function of frequency. Shown are Ψ_+ before the filter cavity (top, dashed trace) and Ψ_+ after the filter cavity (middle, solid trace). The bottom trace shows Ψ_- for both cases. The peak at 500 kHz is due to the relaxation oscillation of the laser. The second peak at 600 kHz in the trace of Ψ_+ after the filter cavity is the frequency modulation used for locking the cavity. The resolution bandwidth was 10 kHz.

3. MEASUREMENTS OF SQUEEZING

A. Squeezing Spectrum

The squeezed vacuum from the OPO is detected by a measurement of the photocurrent fluctuations i_- from the balanced homodyne detector. The resulting current is sent to a spectrum analyzer, where the spectral density Ψ_- is recorded. The maximum noise enhancement S_+ (antisqueezing) and reduction S_- (squeezing) correspond to the fluctuations of the quadrature components $\hat{x} = (\hat{a} + \hat{a}^\dagger)$ and $\hat{p} = (\hat{a} - \hat{a}^\dagger)/i$ of the electric-field operator $\hat{E}$. For a pump power P and an analysis frequency Ω, the spectra are given by[2,22]

$$S_\pm(\Omega, P) = \pm \frac{4\sqrt{P/P_{\text{th}}}}{(\Omega/\Gamma)^2 + (1 \mp \sqrt{P/P_{\text{th}}})^2}, \tag{1}$$

where $\Gamma = 17$ MHz is the cavity linewidth (HWHM) and $P_{\text{th}} = 28$ mW is the threshold power of the OPO. The measured spectral densities of the photocurrents are related to the spectra by

$$\Psi_\pm(\Omega, P) = \Psi_0[1 + \xi\rho S_\pm(\Omega, P)], \tag{2}$$

where Ψ_0 is the spectral density for a vacuum-state input, ρ is the escape efficiency of the resonator, and ξ is the detection efficiency; the last includes propagation losses after the OPO, homodyne efficiency, and detector quantum efficiency.

Because we used broadband detectors that were calibrated between 1 and 30 MHz, we were able to record the whole spectrum of squeezing and antisqueezing by keeping the phase of the local oscillator constant at either minimum or maximum noise level. After subtracting the electronic noise and normalizing the squeezing and the antisqueezing to the shot-noise level, we obtained the traces shown in Fig. 3. The observation of squeezing at still higher frequencies was limited only by the bandwidth of the photodetectors employed in our detection scheme. The theoretical spectra are in good agreement with the experimental data, as shown by the curves presented in Fig. 3, where we used the total efficiency $\delta\xi$ and P/P_{th} as fit parameters.

The measurement of squeezing at a fixed frequency $\Omega = 2$ MHz is shown in Fig. 4, where trace (i) is the squeezing as a function of the local oscillator phase, trace (ii) is the shot noise, and trace (iii) is the squeezing when the phase of the local oscillator is fixed manually at the minimum noise level. By averaging the noise of trace (iii), we determine the squeezing level to be 5.5 ± 0.2 dB below the vacuum noise level. For the measurement shown, the pump power was approximately 3/4 the threshold power. Further increasing the pump power led to a higher gain, but additional noise degraded the squeezing. We believe that this effect is due to pump fluctuations, inasmuch as the pump was not shot-noise limited at 2 MHz. From measurements at lower pump powers we find the total efficiency of detection $\xi = 92\%$, which agrees within 2% with the following measured individual efficiencies: 98.5% homodyne, 98% propagation after the OPO, and 97% detector quantum efficiency. Once we corrected for these values, the inferred squeezing amounted to 7.9 dB outside the resonator.

B. Squeezing Ellipse and Wigner Function

The usual way to represent a squeezed state of the electromagnetic field in the phase space defined by its quadrature components x and p is a depiction of an uncertainty region with the shape of an ellipse. The principal axes of this ellipse correspond to the standard deviation of the fluctuations in the (possibly rotated) squeezed and the antisqueezed quadrature components. The ellipse represents a contour of the Wigner function $W(x, p)$ of that state.

The Wigner function is a quasi-probability distribution in phase space, differing from a classical distribution mainly by the fact that it may be negative and by the fact that it cannot be measured directly, inasmuch as it is a joint distribution of the eigenvalues of two noncommuting observables.[23] What is measurable, though, for any pair of rotated quadrature components $\hat{x}_\theta = \hat{x}\cos(\theta) + \hat{p}\sin(\theta)$, $\hat{p}_\theta = -\hat{x}\sin(\theta) + \hat{p}\cos(\theta)$, are the marginal distributions $P_\theta(x_\theta) = \int_{-\infty}^{\infty} W[x_\theta \cos(\theta) -$

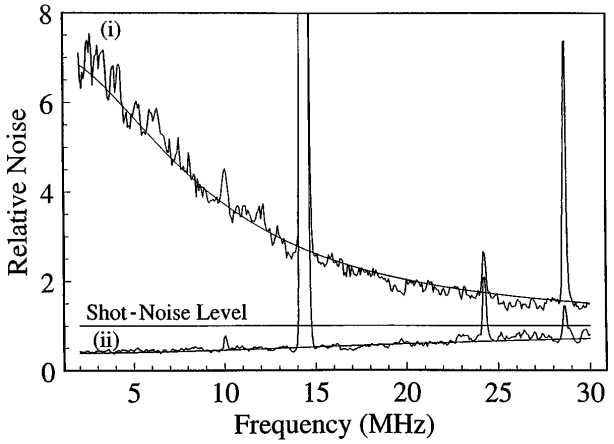


Fig. 3. Spectral density Ψ_- as a function of the analysis frequency Ω, with trace (i) being the antisqueezing spectrum and trace (ii) the squeezing spectrum. The theoretical expectations are denoted by the solid curves superimposed upon traces (i) and (ii). The peak at 14.5 MHz and its harmonic at 29 MHz come from the modulation of the doubling cavity used for locking. The resolution bandwidth was 100 kHz; the video bandwidth, 100 Hz.

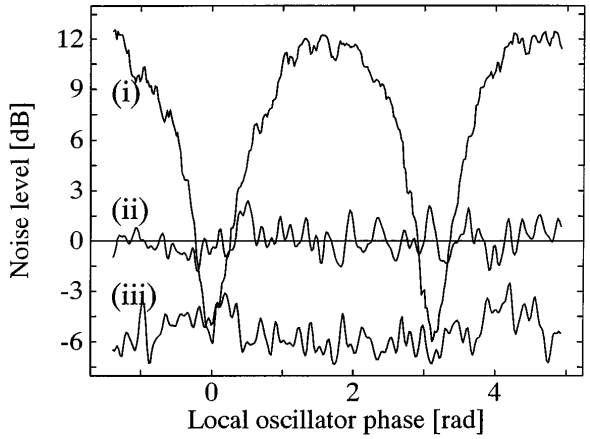


Fig. 4. Squeezed vacuum from the monolithic OPO. In trace (i) the phase of the local oscillator is being scanned; in trace (iii) the phase is fixed manually for minimum noise. One obtains the shot-noise level by averaging trace (ii). The resolution bandwidth was 100 kHz; the video bandwidth, 1 kHz.

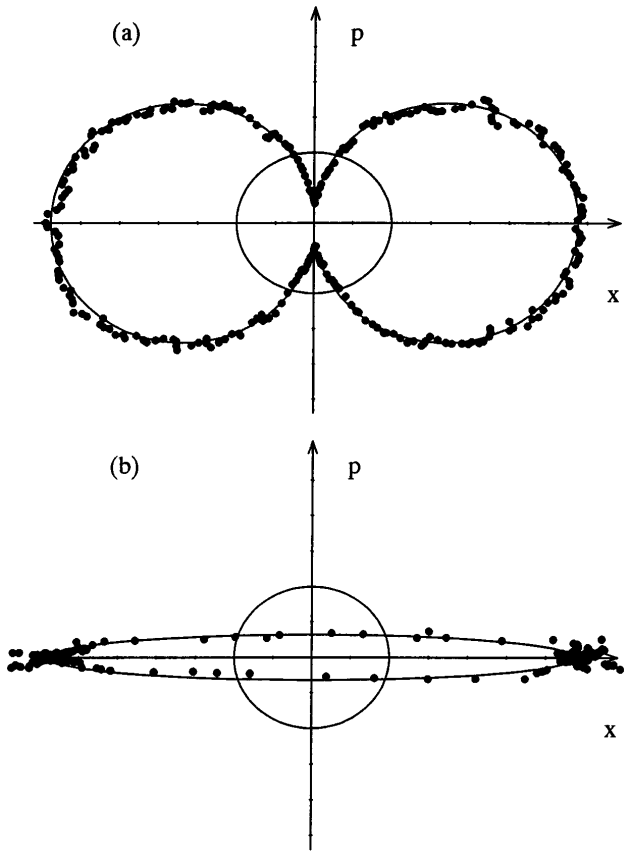


Fig. 5. (a) Standard deviations of the electric-field amplitudes of a squeezed-vacuum state in phase space. The circle represents the unsqueezed vacuum state. (b) The same data transformed into a squeezing ellipse. The points correspond to measured values corrected for total efficiency, and the curves represent the theoretical expectations.

$p_\theta \sin(\theta),\ x_\theta \sin(\theta) + p_\theta \cos(\theta)]\mathrm{d}p_\theta$.[17] These marginal distributions are classical probability distributions. For the case of a squeezed state, they are Gaussian. The recently established technique of optical homodyne tomography permits the reconstruction of the Wigner function from these marginal distributions by use of the inverse Radon transform.[17,18] Because the Wigner function is related to the density matrix through a Fourier transform, it is possible to determine completely the quantum-mechanical state of a system.

In the following discussion, we first show how to visualize the uncertainty region of the quantum-mechanical state of the light field in our experiment based on a measurement of the variances, and then we describe the reconstruction of its entire Wigner function.

The square roots of the experimental values of the variances (measured as in Fig. 4) correspond, when corrected for escape and detection efficiency and the electronic noise floor, to the standard deviation of the electric-field operator $\hat{E}$ in phase space as a function of the phase angle θ. Its equation is given by $\Delta\hat{E}(\theta) = [(\Delta\hat{x})^2 \cos^2(\theta) + (\Delta\hat{p})^2 \sin^2(\theta)]^{1/2}$, which is the equation of a lemniscate. The phase-space representation of our experimental data is shown in Fig. 5(a). The circle represents the unsqueezed vacuum state. The principal axes of the lemniscate, $a = \Delta\hat{x}$ and $b = \Delta\hat{p}$, correspond to maximum noise suppression or enhancement in the respective quadrature component.

The variances of the E field can be thought of as being projections of the uncertainty region in phase space onto the direction of observation determined by the phase angle θ. Geometrically, this means that the E-field variances form the pedal curve of the uncertainty region.[24] Therefore the uncertainty region and thus the contour of the Wigner function are calculated with the transformation

$$[r(\theta)\cos(\theta),\ r(\theta)\sin(\theta)] \Rightarrow \left[\frac{a^2}{r(\theta)}\cos(\theta),\ \frac{b^2}{r(\theta)}\sin(\theta)\right]. \tag{3}$$

The result for a 4-dB squeezed-vacuum state is shown in Fig. 5(b). The area of the ellipse and the circle are approximately equal, as is expected for a minimum uncertainty state.

Although this transformation of the values of the variances to the contour of the Wigner function is correct for a minimum uncertainty state, it is not correct for arbitrary states of the light field. The information contained in the variances is in general not sufficient to infer the geometrical shape of the uncertainty region.

We gain more information about the state of the light field by recording not just the variance but rather the whole distribution $P_\theta(x_\theta)$ of the electric-field fluctuations.

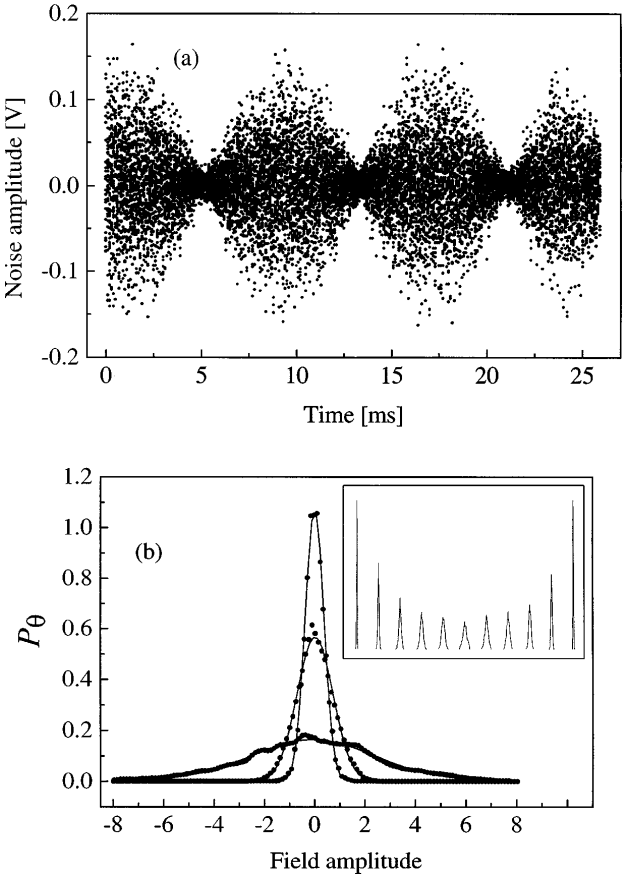


Fig. 6. (a) Noise of the photocurrent i_- of the squeezed vacuum from the OPO at 2 MHz as a function of the local oscillator phase. The detection bandwidth is 100 kHz. (b) Distributions of the photocurrent fluctuations for maximum squeezing, shot noise, and maximum antisqueezing. The curves are fits of Gaussian distributions. The inset shows the distributions for eleven different phase angles from 0 to π.

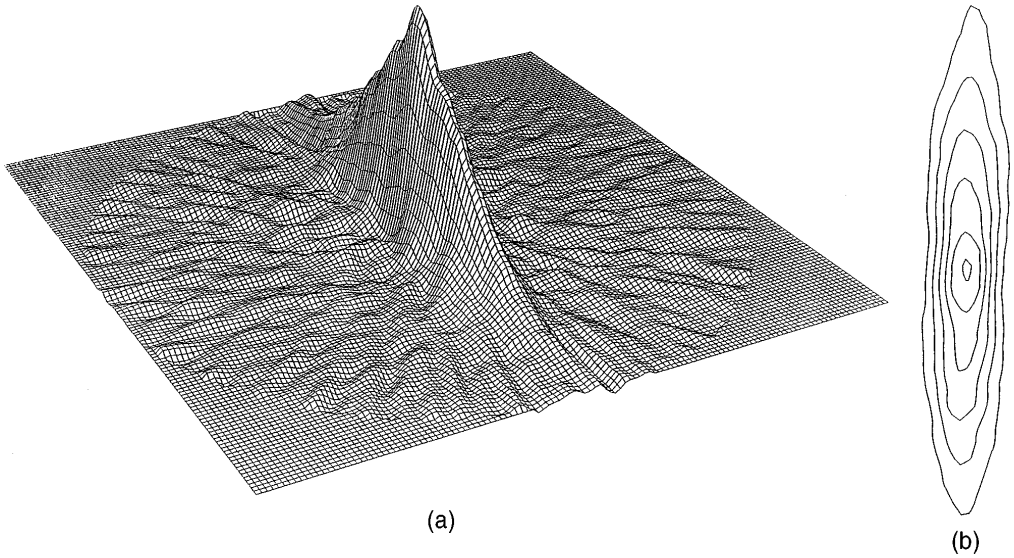


Fig. 7. (a) Reconstruction of the Wigner function for the squeezed vacuum. (b) Contour plot of the Wigner function. The ellipticity of the contours is slightly smaller than in Fig. 5 because no corrections for the detection efficiency were made.

As mentioned above, this procedure allows us to employ the method of optical homodyne tomography to determine the Wigner function completely. The inverse Radon transform corresponds in two dimensions to the step of backprojecting the variances, carried out above.

For a squeezed-vacuum mode of frequency Ω relative to the carrier optical frequency, the Wigner function is given by

$$W_{\mathrm{sq}}(x,\,p) = (1/2\pi)\exp\left(-\frac{1}{2}\{[1+\rho S_{-}(\Omega)]x^2 + [1+\rho S_{+}(\Omega)]p^2\}\right), \tag{4}$$

where S_{+} and S_{-} are defined by Eq. (1). Thus there is a spectrum of Wigner functions that differ in the ratio of their half-widths. For our reconstruction we chose the field mode at $\Omega = 2$ MHz.

Following the procedure described in Ref. 18, we first obtained the distributions of the electric-field amplitude at several different phase angles in phase space. The intermediate-frequency output of the spectrum analyzer at 2 MHz with a resolution bandwidth of 100 kHz, was recorded by a Nicolet 400 digitizing oscilloscope, as the local oscillator phase was slowly scanned. In this way it is possible to observe the variance of the electric field directly on the spectrum analyzer while accumulating the amplitude data. A typical trace of the resulting photocurrent fluctuations is shown in Fig. 6(a). We divided this time trace of 161,000 points into 79 sections corresponding to 79 different phase angles. For each phase angle interval the time-dependent amplitude noise was sorted into 128 bins. Thus histograms of 128 bins at 79 different phase angles θ were created, each having the shape of a Gaussian [Fig. 6(b)]. These histograms correspond, when normalized, to the Gaussian distributions of the amplitude noise at different phase angles.

Finally, the collection of histograms was transformed with the inverse Radon transform. The main part of the inversion algorithm consists of a convolution of the measured histogram values with a given filter function (filtered backprojection). A quadratic regularization was used to approximate this filter function. The Wigner function and its contour lines are shown in Fig. 7. Note that no corrections for detection and escape efficiency were made. For the chosen parameters, the numerical error of this reconstruction process is estimated to be less than 2%, which is consistent with a comparison of the ratio of the two half-widths of the Wigner-function sections at $x = 0$ and $p = 0$ and the directly observed squeezing and antisqueezing of 5.5 and 11 dB for this measurement.

4. CONCLUSIONS

In conclusion, we have used a monolithic continuous-wave OPO pumped below threshold to generate squeezed vacuum at 1064 nm. With this device 5.5 dB of squeezing was observed at a frequency of 2 MHz, and the whole squeezing spectrum from 2 to 30 MHz was recorded. Furthermore, the Wigner function of the squeezed-vacuum state was reconstructed. With the same device we have generated a beam of bright squeezed light by the process of parametric deamplification. 4-dB noise reduction in the amplitude noise of a 0.25-mW beam was achieved.[25] The level of detected squeezing was limited mainly by the finite detection efficiency and the noise of the pump laser. We plan a further investigation of the effect of the latter in the future.

The reconstruction of the Wigner function of a squeezed state represents a first step in a complete experimental characterization of a squeezed state. Because of the high degree of squeezing and its good stability, the experimental determination of the photon statistics and density-matrix elements in Fock-state representation appears to be possible.[26]

We believe that, beyond the demonstration of the performance of monolithic resonators as very stable, compact squeezing sources with a high degree of squeezing, they can be useful in several applications of squeezed light. The broad bandwidth makes it suitable for spectroscopic investigations, and its simplicity should make

proposed experiments that require more than one source of squeezed light more accessible.

ACKNOWLEDGMENTS

We are indebted to R. Paschotta for essential help at the beginning of the experiment, especially for the design of the frequency doubler. We thank M. G. Raymer, A. Faridani, and D. MacAlister for providing us with the original source code for the inverse Radon transform, and U. Janicke and U. Leonhardt for several improvements regarding the numerical inversion. Furthermore, we gratefully acknowledge several enthusiastic discussions with A. G. White and S. Kohler. S. F. Pereira was supported by the Humboldt Foundation. J.-Ph. Poizat was supported by the European Community (EC) network "Non-classical Light" and thanks the Konstanz group for its hospitality during his stay. Financial support for this research was provided by the Deutsche Forschungsgemeinschaft, by the EC network Nonclassical Light, and by EC Basic Research Project 6934-QUINTEC.

S. Schiller's e-mail address is stephan.schiller@uni-konstanz.de.

*Permanent address, Institut d'Optique, B. P. 147, 91403 Orsay, France.

REFERENCES

1. See J. Opt. Soc. Am. B **4** (1987); J. Mod. Opt. **34** (1987) (special issues on squeezed states of light); Appl. Phys. B **55,** (1992) (special issue on quantum-noise reduction in optical systems).
2. L.-A. Wu, H. J. Kimble, J. L. Hall, and H. Wu, "Generation of squeezed states by parametric down conversion," Phys. Rev. Lett. **57,** 691 (1986).
3. E. S. Polzik, J. Carri, and H. J. Kimble, "Atomic spectroscopy with squeezed light for sensitivity beyond the vacuum-state limit," Appl. Phys. B **55,** 279 (1992).
4. J. Mertz, T. Debuisschert, A. Heidmann, C. Fabre, and E. Giacobino, "Improvements in the observed intensity correlation of optical parametric oscillator twin beams," Opt. Lett. **16,** 1234 (1991).
5. M. Xiao, L.-A. Wu, and H. J. Kimble, "Precision measurement beyond the shot-noise limit," Phys. Rev. Lett. **59,** 278 (1987); P. Grangier, R. E. Slusher, B. Yurke, and A. La Porta, "Squeezed-light-enhanced polarization interferometer," Phys. Rev. Lett. **59,** 2153 (1987).
6. E. S. Polzik, J. Carri, and H. J. Kimble, "Spectroscopy with squeezed light," Phys. Rev. Lett. **68,** 3020 (1992).
7. Z. Y. Ou, S. F. Pereira, and H. J. Kimble, "Quantum noise reduction in optical amplification," Phys. Rev. Lett. **70,** 3229 (1993).
8. S. F. Pereira, Z. Y. Ou, and H. J. Kimble, "Back-action evading measurements for quantum nondemolition detection and quantum optical tapping," Phys. Rev. Lett. **72,** 214 (1994)
9. B. Yurke, "Optical back-action-evading amplifiers," J. Opt. Soc. Am. B **2,** 732 (1985); P. Smith, M. Collett, and D. F. Walls, "Quantum nondemolition measurement with an optical parametric amplifier," Opt. Commun. **102,** 1051 (1993).
10. I. E. Protsenko, L. A. Lugiato, and C. Fabre, "Spectral analysis of the degenerate optical parametric oscillator as a noiseless amplifier," Phys. Rev. A **50,** 1627 (1994).
11. L. Krippner, W. J. Munro, and M. D. Reid, "Transient macroscopic quantum superposition states in degenerate parametric oscillation: calculations in the large quantum limit using the positive *P*-representation," Phys. Rev. A **50,** 4330 (1994).
12. W. J. Kozlovsky, C. D. Nabors, and R. L. Byer, "Efficient second harmonic generation of a diode-laser-pumped Nd:YAG laser using monolithic $MgO:LiNbO_3$ external resonant cavities," IEEE J. Quantum Electron. **24,** 913 (1988); W. J. Kozlovsky, C. D. Nabors, R. C. Eckardt, and R. L. Byer, "Monolithic $MgO:LiNbO_3$ doubly resonant optical parametric oscillator pumped by a frequency-doubled diode-laser-pumped Nd:YAG laser," Opt. Lett. **14,** 66 (1989).
13. R. Paschotta, K. Fiedler, P. Kürz, R. Henking, S. Schiller, and J. Mlynek, "82% Efficient continuous-wave frequency doubling of 1.06 μm with a monolithic $MgO:LiNbO_3$ resonator," Opt. Lett. **19,** 1325 (1994).
14. C. D. Nabors and R. M. Shelby, "Two-color squeezing and sub-shot noise signal recovery in doubly resonant optical parametric amplifier," Phys. Rev. A **42,** 556 (1990); A. Sizmann, R. J. Horowicz, G. Wagner, and G. Leuchs, "Observation of amplitude squeezing of the up-converted mode in second harmonic generation," Opt. Commun. **80,** 138 (1990); P. Kürz, R. Paschotta, K. Fiedler, and J. Mlynek, "Bright squeezed light by second-harmonic generation in a monolithic resonator," Europhys. Lett. **24,** 449 (1993); R. Paschotta, M. Collett, P. Kürz, K. Fiedler, H.-A. Bachor, and J. Mlynek, "Bright squeezed light from a singly-resonant frequency doubler," Phys. Rev. Lett. **72,** 3807 (1994).
15. A. Giazotto, "Wide-band measurement of gravitational waves: the VIRGO project," Nuovo Cimento C **6,** 955 (1992); J. Hough, G. P. Newton, N. A. Robertson, H. Ward, B. F. Schutz, K. Danzmann, A. Rüdiger, R. Schilling, and W. Winkler, "Proposal for a 600 m laser interferometric gravitational wave antenna—GEO 600," Internal Rep. 190 (Max Planck Institut für Quantenoptik, Munich, 1994).
16. D. Bakalov, G. Cantatore, G. Carugno, S. Carusotto, P. Favaron, F. Della Valle, I. Gabrielli, U. Gastaldi, E. Iacopini, P. Micossi, E. Milotti, R. Onofrio, R. Pengo, F. Perrone, G. Petrucci, E. Polacco, C. Rizzo, G. Ruoso, E. Zavattini, and G. Zavattini, "PV-LAS: vacuum birefringence and production and detection of nearly massless, weakly coupled particles by optical techniques," Nucl. Phys. B Proc. Suppl. **35,** 180 (1994).
17. K. Vogel and H. Risken, "Determination of quasiprobability distributions in terms of probability distributions for the rotated quadrature phase," Phys. Rev. A **40,** 2847 (1989).
18. D. T. Smithey, M. Beck, M. G. Raymer, and A. Faridani, "Measurement of the Wigner distribution and the density matrix of a light mode using optical homodyne tomography: application to squeezed states and the vacuum," Phys. Rev. Lett. **70,** 1244 (1993).
19. G. Breitenbach, S. Schiller, and J. Mlynek, "81% Conversion efficiency in frequency-stable continuous-wave parametric oscillation," J. Opt. Soc. Am. B **12,** 2095 (1995).
20. R. Drever, J. Hall, F. Kowalski, J. Hough, G. Ford, A. Munley, and H. Ward, "Laser phase and frequency stabilization using an optical resonator," Appl. Phys. B **31,** 97 (1983).
21. H. P. Yuen and V. W. S. Chen, "Noise in homodyne and heterodyne detection," Opt. Lett. **18,** 177 (1983); B. Schumaker, "Noise in homodyne detection," Opt. Lett. **19,** 189 (1984).
22. M. J. Collett and D. F. Walls, "Squeezing spectra for nonlinear optical systems," Phys. Rev. A **32,** 2887 (1985).
23. E. P. Wigner, "On the quantum correction for thermodynamic equilibrium," Phys. Rev. **40,** 749 (1932).
24. R. Loudon, "Graphical representation of squeezed-state variances," Opt. Commun. **70,** 109 (1989).
25. S. Schiller, G. Breitenbach, S. F. Pereira, R. Paschotta, A. G. White, and J. Mlynek, "Generation of continuous-wave bright squeezed light," in *Laser Frequency Stabilization and Noise Reduction*, Y. Shevy, ed., Proc. Soc. Photo-Opt. Instrum. Eng. **2378,** 91 (1995).
26. U. Leonhardt, Department of Physics, Humboldt-Universität, 12484 Berlin, Germany (personal communication, 1995); M. G. Raymer, Department of Physics, University of Oregon, Eugene, Oregon 97403 (personal communication, 1995); G. M. D'Ariano, C. Macchiavello, and M. G. A. Paris, "Detection of the density matrix through optical homodyne tomography without filtered back projection," Phys. Rev. A **50,** 4298 (1994).

Journal of Space Exploration, Mehta Press

Foreword to the sixth (special) issue, September 2014

Guest editor: Prof. Jochem Hauser,
Hpcc-Space GmbH, Hamburg, Germany

Festschrift in Honor of the 80th Birthday of

Dr. Williamson Berry

ret., Head of Propulsion and Aerothermodynamics Division, European Space research and Technology Centre European Space Agency Noordwijk, The Netherlands

This special Issue of the *Journal of Space Exploration* is dedicated to my friend, colleague, and (former) boss at ESA, the eminent

Dr. Williamson (William, Bill) Berry

on the occasion of his 80th birthday (recent picture). As head of the Propulsion and Aerothermodynamics Division at ESTEC for more than twenty-six years, Bill, as one of the leading chemical propulsion experts in Europe, exerted a major influence on European aerospace industry to further his goal in making European rocketry competitive with U.S. and Russian industry. When Bill left ESA in July 1999, he could look back on the Ariane success story, to which ESA in general, and he

personally had contributed a major part. Space activities in Europe had changed from humble beginnings into commercial business. In particular, Bill played an active role in the creation of an autonomous, world-class capability in spacecraft propulsion and aerothermodynamics for Europe. However, Bill exerted also a major influence on my career, when he hired me (a physicist) as the head of the aerothermodynamics section in 1988 in his division. Among many other things, we worked together in the early 1990s on the Hermes space-plane, the European version of the U.S. space shuttle. Though Bill was responsible for the European propulsion activities, which means that his view had to be on technically realizable propulsion systems, he was well aware of the severe limits of current propulsion and had a keen vision on the physics of advanced propulsion. I still remember our discussions on *field propulsion*, a term already coined by W. Corliss in 1960, i.e., propulsion without propellant based on hitherto unknown long-range physical fields, which became again a major topic in NASA's *Breakthrough Propulsion Physics* program from 1996-2001.

Because of his outstanding personality and open-mindedness William Berry assembled a long list of friends among his engineering colleagues during his long career in ESA, but I am proud (as a non-engineer) to consider myself friend of Bill who also was a wonderful boss during my time at ESA. In addition to his skillful work in the world of European propulsion, William Berry always saw himself as a contributing citizen both to the larger field of science and humanity, which included, of course, animal world.

When I was asked by the editor in chief of JSE, Prof. T. Musha, to choose a topic of my own for the sixth issue of JSE, I suggested the theme

Spaceflight Perspectives from Novel Concepts of
Spacetime, Gravitation and Symmetries.

Numerous colleagues, both in the academic departments and in the engineering community whom I contacted, approved this title. I am convinced that William Berry would have done the same, but, for obvious reasons, I could not contact him. A dominant theme in Bill's work was the quest to find the ultimate space propulsion technology. Of course, he was well aware that such a technology has to come from novel physics, and not from the refinement of engineering.

It therefore feels very appropriate to me to dedicate this special issue of JSE as Festschrift to William Berry.

At the European Space Agency (ESA) we experienced first-hand the immense technical difficulties and extreme cost to place relatively small payloads into low earth orbit. Space propulsion is still dealing with the technologies developed in the 50s and 60s of the last century, and the vision portrayed by Werner von Braun in his famous article in Collier's magazine in 1952, entitled *Man on the Moon*, did not become a reality. The shuttle era came to an end in 2012, and NASA will not fly a next generation space vehicle any time soon. The problem of propulsion is not the engineering, but it is linked to the underlying physics that remains unchanged since the days of ancient Chinese rockets. It is the physical principle of classical momentum conservation, which stands in the way of producing an efficient, and effective propulsion system. It is the basic physics itself that prevents progress.

Consequently, all contributions to this Festschrift belong to the (speculative) physics of advanced propulsion, energy generation, and gravitation, and report on the progress made since 2001, when the NASA Breakthrough Propulsion Physics program ended. When I contacted the potential contributors, I offered them free choice of topic, but mentioned that all papers will be reviewed, and none of the editors would interfere with the verdict of the reviewers. I feel that the final collection of papers displays an exciting variety, just as William Berry does wish to see!

Recently, the scientific community has been shaken by numerous scandals, where authors published exciting results that eventually had to be retracted, because data had been deliberately manipulated (remember, for instance, the infamous hockey stick curve in climate science, or the large number of articles with falsified measurements in the field of bioscience from authors already considered for a noble prize). Moreover, journals were found to publish articles of questionable quality, i.e., using an insufficient review process. As a consequence, all papers presented in this special issue were thoroughly reviewed. Finally, seven out of eight papers were accepted, three papers were completely rewritten and four papers saw minor modifications. The review process for the paper of my long-time co-author Walter Dröscher and for my own paper was handled by the editor in chief of JSE. I am indebted to these referees, both in Europe and the U.S., who do their work in anonymity, for their patient, but relentless efforts to improve style, readability, clarity, and content of the papers.

All of the authors have been in the field long enough to be widely known from their publications, but of course, there is no guarantee that the novel ideas discussed in these papers actually reflect physical reality. If they do, a revolution will occur both in science and technology, but it may well be that our ideas turn out to be wrong. If, in hindsight, some or all of the ideas presented here, appear incorrect, the reader should be assured that, at least, we were acting as honest fools, which is scientifically legitimate as was stated by the well known physicist Richard P. Feynman.

The last paper, accessible for the intelligent layman, by C. Daigle, summarizes from a design perspective the ensuing revolutionary technological aspects that could follow from this novel physics.

In this special issue we are presenting to the scientific community the state of theoretical and experimental research concerning the latest results in the emerging field of physics for novel gravity-like fields that might represent a new paradigm shift regarding the very nature of gravitation, including the evaluation of recent experiments that might have measured extreme gravitomagnetic fields outside general relativity.

Therefore, the motivation to further exploring the mysterious nature of gravitation is understandable, aiming beyond Newtonian (Einsteinian) gravity. Gravitation has maintained the interest of researchers at every stage in the history of physics, and it became Einstein's quest to unify gravitation with the other forces since 1916. Hence,

it should be no surprise that new theoretical attempts along with experimental work are presented in this special issue to continue where Einstein was forced to leave off.

The quantization of the gravitational field has been unsuccessful, despite great efforts in this direction. The problem may be that the number of fundamental forces is not known, in other words, there is a belief that only four forces exist (strong, weak, electromagnetic, and gravitational force). Perhaps gravity is of a more subtle nature than Newtonian gravity, and an interaction between gravity and electromagnetism might exist?

New gravitational experiments have been published since 2006, and geometrical theories from the 1950s have gained prominence (for instance, by Finzi, Heim, Wheeler). They have been extended and combined with concepts of modern physics, trying to explain novel experimental results for extreme gravitomagnetic and gravity-like fields. For instance, in his monograph on *Quantum Field Theory*, M. Kaku presents a calculation of the Coleman-Weinberg potential that might be employed to calculate the coupling strength for extreme gravitomagnetic fields. More recently, as pointed out by A. Zee in *Quantum Field Theory in a Nutshell*, gravity might be the square of two spin 1 fields (it should be noted that particles of spin 1 can be described by Yang-Mills fields), an idea that also might be applicable in the explanation of the experiments on extreme gravitomagnetic fields. These and other exciting ideas are presented to the reader, shedding new light on the nature of gravity as well as the number and type of fundamental forces that should exist in Nature.

As Einstein felt, the most important objective of any theory is to comprise as few and basic elements as possible without contradicting physical experience in conjunction with practical applications. For example, a relationship between the physical phenomena of electromagnetism and gravitation might exist at cryogenic temperatures. Any novel theory must be verifiable by laboratory experiments or astronomical observations.

Performing gravitational experiments is not an easy endeavor since highly sensitive devices have to be produced and utilized at cryogenic temperatures, often at liquid Helium temperature. Even if experimental findings or theories eventually cannot be verified, one should not denounce the serious experimenter or theorist for failure, since the history of science has shown that every step forward is a complicated venture, needless to say that all programs for novel theoretical models initially contain many unclear points. But this is true even for established theories. Gravitational experiments are notoriously difficult as can be seen from the fact that the physics of gravitational wave astronomy, despite the early efforts of J. Weber starting out in 1969, is still not an established fact.

As was pointed out by theoretical physicist Richard P. Feynman, there occur numerous strange phenomena in the complex situations of solid-state physics. He prophetically foresaw an enormous number of technical applications that could arise from such physics. Might it be possible that a combination of low temperature and

solid state physics could lead to strange phenomena and, this is the most important point, to a large number of technical applications, but this time in the field of gravitational engineering?

Finally, in order for science to progress, both theorists and experimenters have to be willing to leave the save (funding issues) and (sometimes) sacred ground of established physical theory and dare to propose speculative novel ideas. If a blind alley is met, the courage to reverse one's direction of research is required. In particular, the speaking silence of the Large Hadron Collider at Geneva is at odds with all of the so-called advanced physical theories beyond the standard model of particle physics. Instead, if ideas of novel gravitational fields at cryogenic temperatures turn out to be true, the new scientific age of gravitational engineering might have begun.

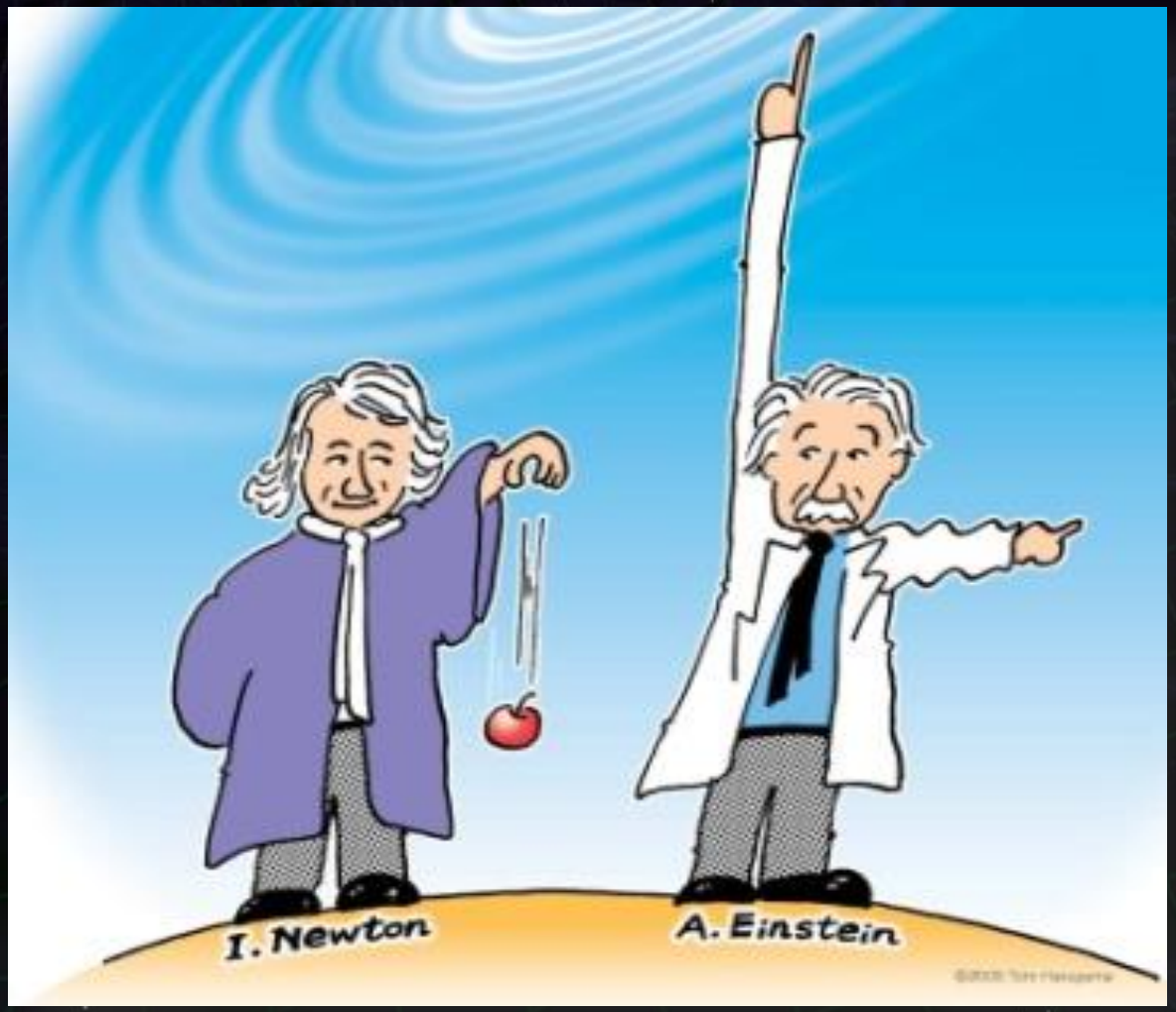
Whether or not this special issue of JSE, which can be considered a sequel of the recent book on *Gravity-Superconductor Interactions: Theory and Experiment, G. Modanese, G. A. Roberson (eds.), Bentham Publishers 2012,* is revealing novel scientific facts, or even stands for a paradigm shift in science, remains undecided at present. New experiments, hopefully as suggested in this special issue, need to be carried out in order to prove the existence of gravity-like fields outside Newtonian gravity. Nevertheless, I am convinced that the articles presented here contain a large amount of exciting material that will be of substantial benefit to those readers willing to leave the trodden path of science and to accept the challenge.

Subtle is the Lord…Albert Einstein

Jochem H. Hauser

Professor (em) of HPC, Ostfalia University of Applied Sciences,
Campus Suderburg, Germany
and
Scientific Director,
High Performance Computing and Communications for Space GmbH,
Hamburg, Germany

Gravity and Gravitational-Wave Physics



Masaki Ando
(National Astronomical Observatory of Japan)

Global COE Symposium 'Development of Emergent New Fields' (Feb. 13, 2013, Kyoto University, Kyoto)

Self Introduction

- Was a Special Associate Professor of this GCOE program for 3.5 years (2009.1 – 2012.5)

  → Special Research Unit for
  'Gravity and Gravitational-Wave Physics'
  (Belonging to 'Tentaikaku' astrophysics theory group)

- From June 2012, an Associate Professor of NAOJ (National Astronomical Observatory of Japan)

  Gravitational-Wave Project office

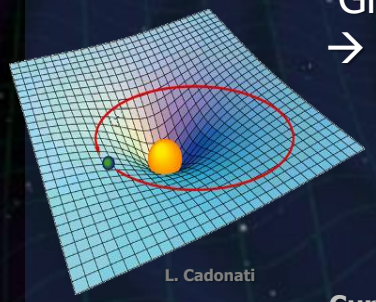
  - Host for TAMA300 GW detector
  - Co-host for KAGRA GW antenna

  ~20 members (1 Prof. to come, 1 Assoc. Prof., 5 Research Associates, 3PDs, 3 Engineers, 2 Secretaries, and 4 Grad. Students)

Gravity and Gravitational waves

General Relativity

Gravity
→ Nature of space-time



Einstein equation

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}$$

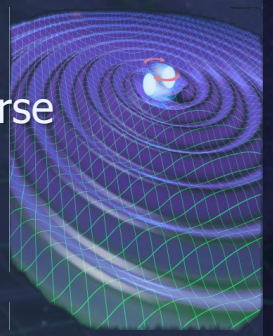
Curvature Of space time

Mass (Energy Momentum)

AC component

Gravitational wave

GW astronomy
New probe for the Universe
Violent phenomena,
Early universe, …
Test of general relativity



DC component

Gravity

Direct Probe of space-time
Prospects on problems
in modern physics
Fundamental Physics
ISL, EP

Obs.

Precise Measurement

Displacements of
Macroscopic objects
Laser interferometer
Quantum optics, …
Fundamental noises
Thermal fluctuation, …

Measurement

KAGRA（かぐら）

LCGT

2nd generation GW detector in Japan

Obs. Start ~2017 → Direct detection of GW

Large-scale Detector

Baseline length: 3km

High-power Interferometer

Cryogenic interferometer

Mirror temperature: 20K

Underground site

Kamioka mine,

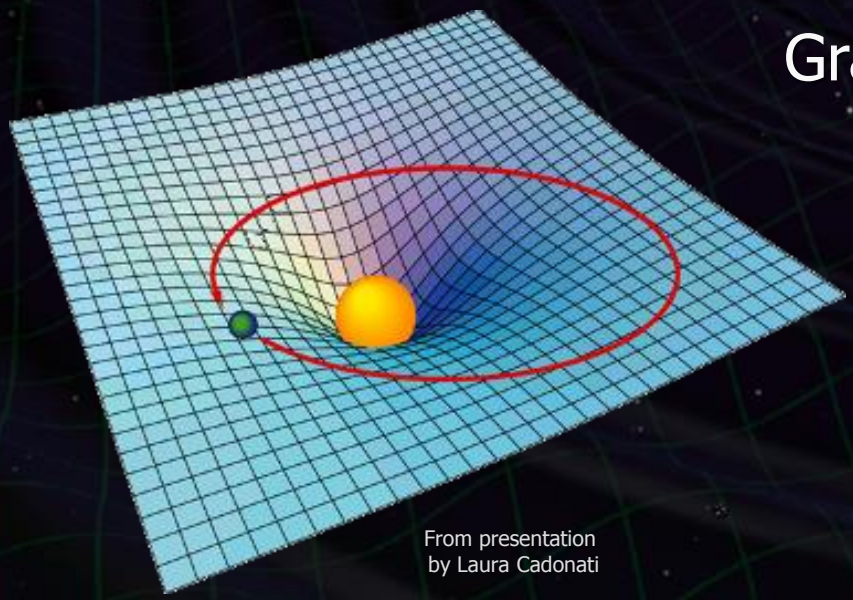
1000m underground

Outline

- Gravitational-Wave Astronomy
- Overview KAGRA GW Antenna
- Current Status of KAGRA
- Summary

Gravitational Wave Astronomy

Gravitational Wave



From presentation
by Laura Cadonati

General Relativity

Gravity : Curvature of space-time

Acceleration of Mass
→ Fluctuations in space-time
→ Propagates as
'Ripples in space-time'

"Mass tells space-time how to curve,
and space-time tells mass how to move."
John Archibald Wheeler

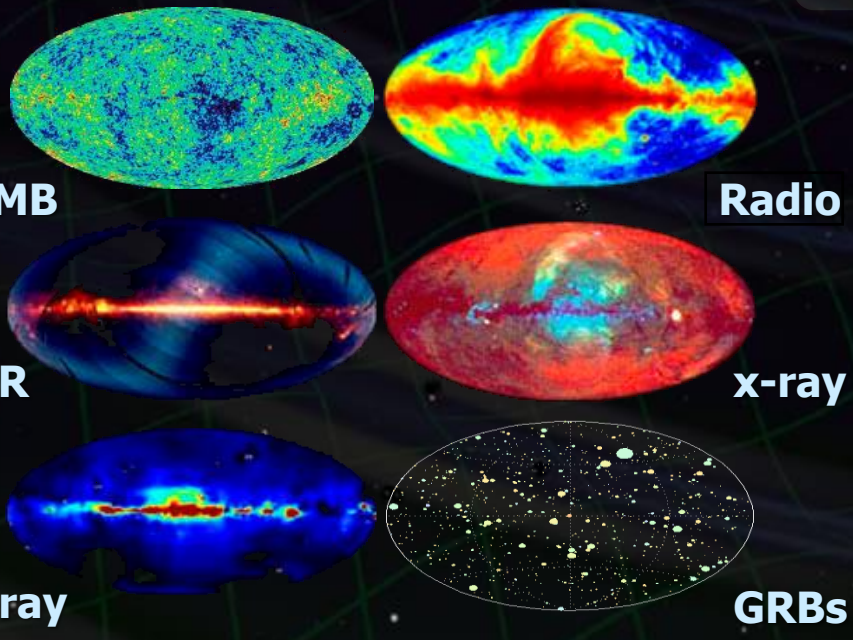
Gravitational Waves

Gravitational-Wave Astronomy

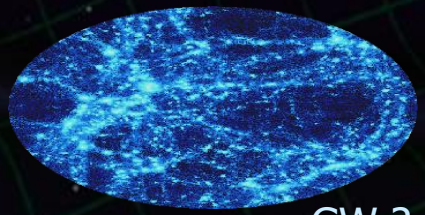
Reveal the universe by Gravitational Waves.

Nature of GWs
- Radiated by accelerated masses
- Strong transmissivity

New probe to the universe
- Complementary with EMWs.
- Unique sciences
  - Early universe before CMB era
  - High-energy phenomena



+



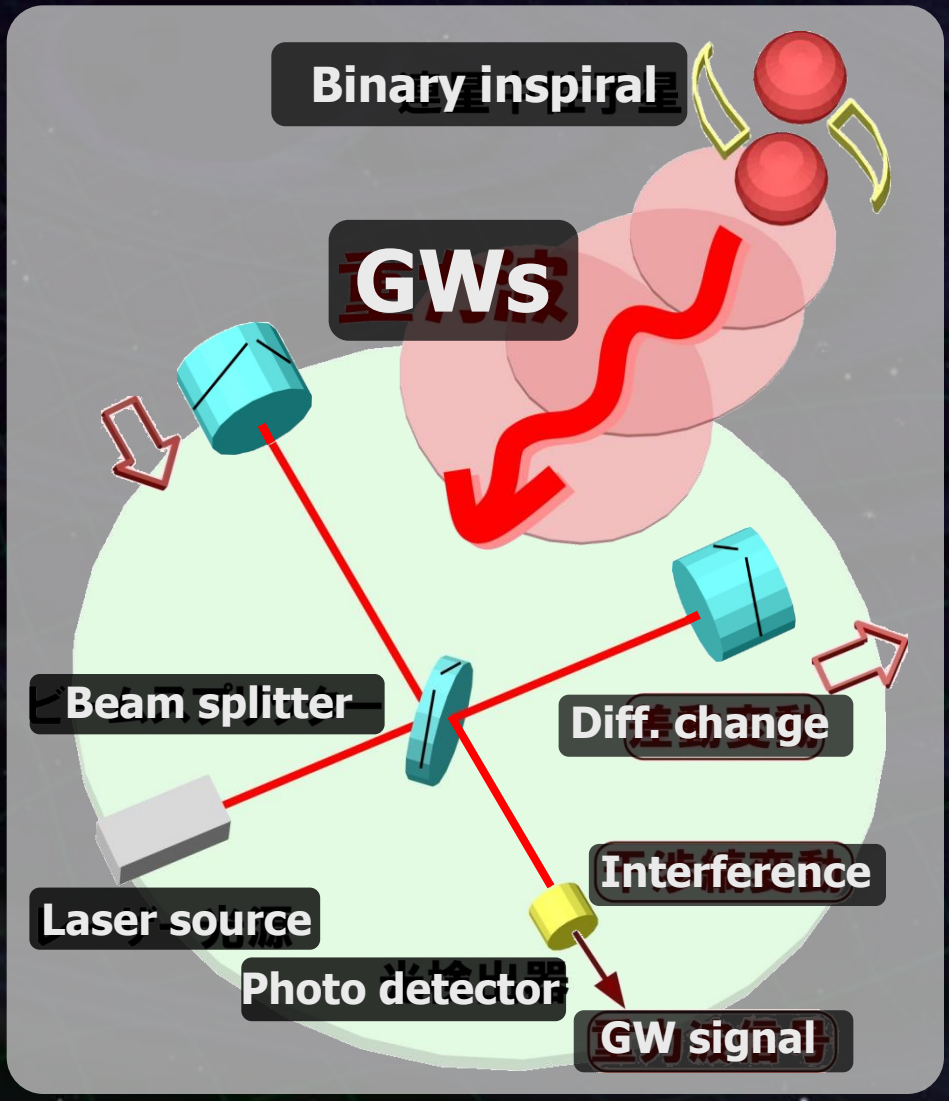
Laser Interferometric Detector

Laser Interferometer

(Michelson interferometer)

When GW comes...

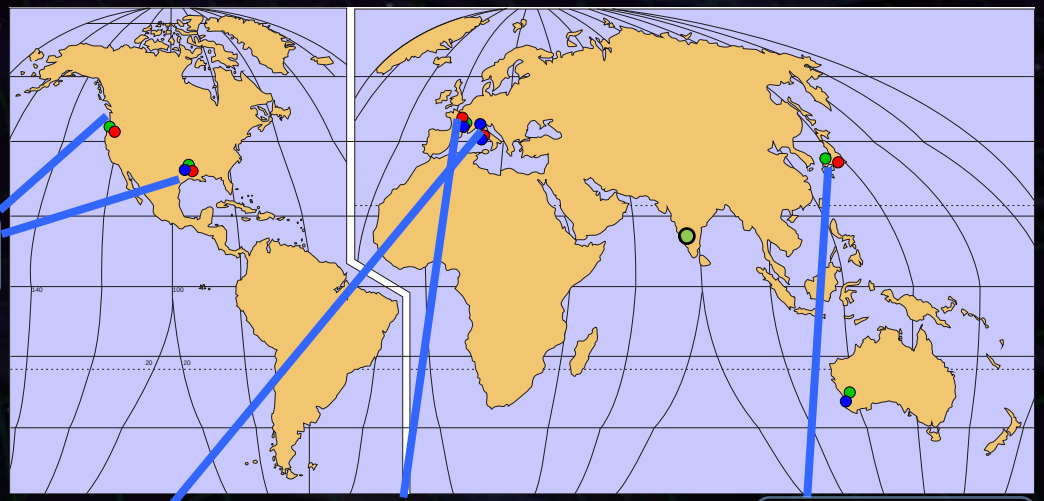
Differential length (strain) changes in two arms
↓
Detected at photo detector



2nd-Generation GW Antennas

International observation network
for GW astronomy will be on-line in ~5 years.
>200Mpc range →Event rate ~10 events/year

aLIGO (USA)
4km x 2 (or3)

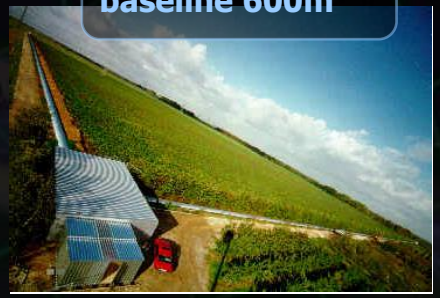


LIGO-Austreria

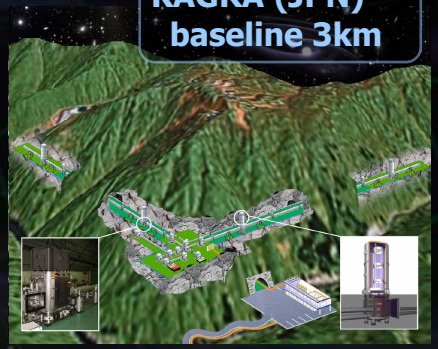
LIGO-India
in proposal

indigo

GEO-HF (GER-UK)
baseline 600m



KAGRA (JPN)
baseline 3km



Adv.VIRGO (ITA-FRA)
baseline 3km

Target of Ground-based Detectors

Terrestrial Detectors – Obs. Band ~10Hz – 1kHz

⇨ Compact and high-energy astronomical phenomena

Neutron Star

Black Hole

Early Universe

Pulsar

Supernova

Binary merger

Stochastic Background

EMRI

Soft Gamma-Ray Repeater

Star Oscill ation

Long GRB

Short GRB

Quasi-normal Oscillation

EM Waves
Neutrino
High-energy CR

Follow-up observation
Numerical Relativity

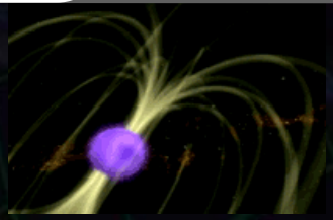
Understandings on High-E phenomena

EoS of high-density Matter → Nuclear Phys.

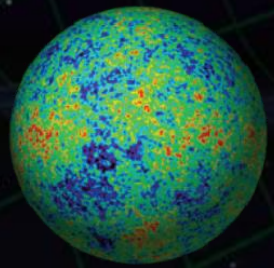
Test of Relativity Physics in Strong Gravity

GW targets and data analysis

Signal duration

Waveform	Short (bursts)	Long (stationary)
Known	Binary merger → Chirp wave, Ringdown wave	Pulsar, LMXB → Continuous
Unknown	Stellar core collapse → burst wave	 Soft gamma-ray repeater

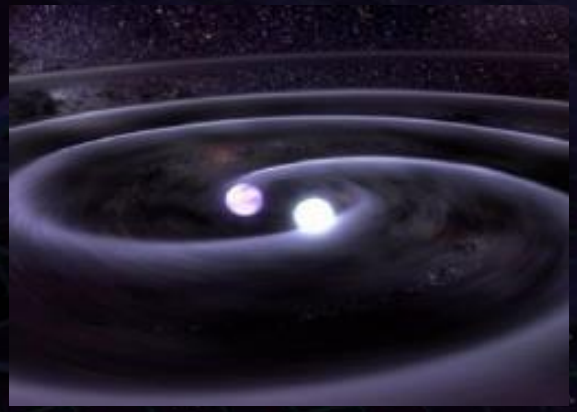
Gamma-ray bursts



Stochastic background → Random wave

Neutron-star inspiral

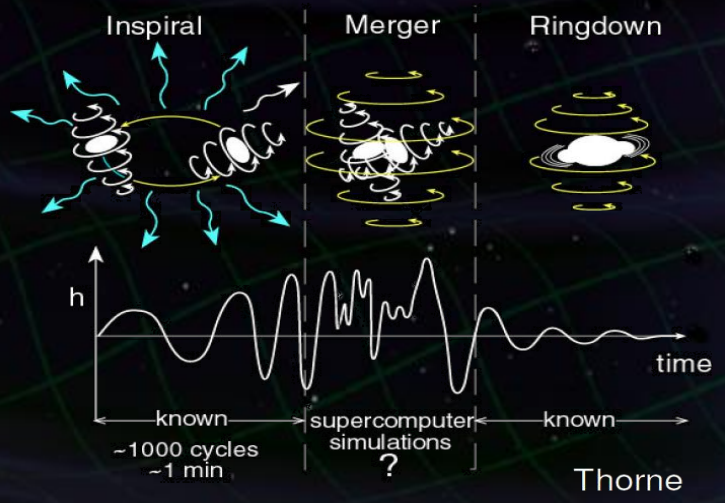
Primary target :
Inspiral and merger of NS binary



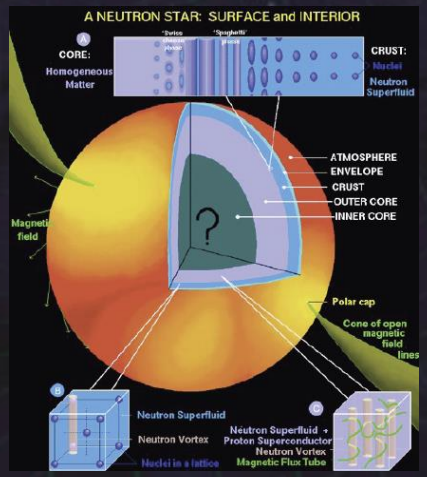
- Quantitative estimation of event rate from pulsar observations.
- Precise waveform is predicted.
  → Sophisticated analysis method using an optimal filter.

⇩

Promising for first detection



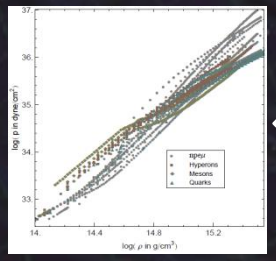
EoS of Neutron Stars



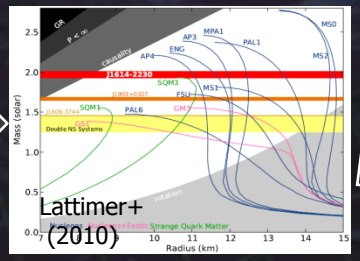
Neutron-star EoS

(Density $2 \sim 15\rho_0$)

Density-Pressure



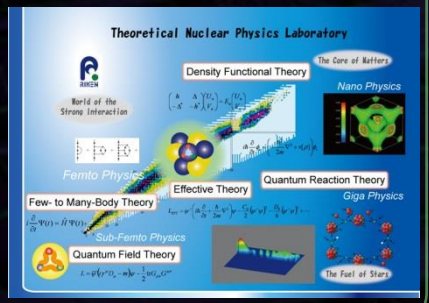
Radius-Mass of NS



Lattimer+ (2010)

Nuclear Physics

Natural Laboratory for high-density nuclear physics



理研Web pageより

High-energy Phenomena

→ Astrophysics, Frontier Physics

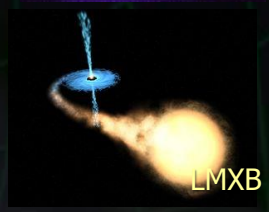
Radio Pulsar

Supernova

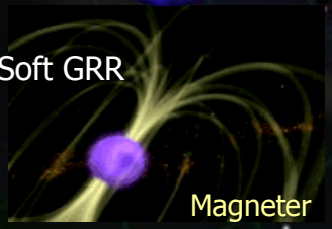
Short GRB

Binary Merger

Short GRB

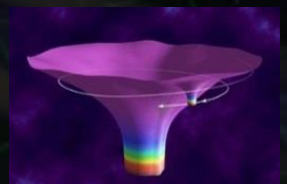


LMXB



Soft GRR

Magneter



Black-hole formation

Relativity and Cosmology

Compact Binary

→ Standard Siren

Physics in Strong Gravity

Cosmological Parameters

Messenger and Read, arXiv:1107.5725

Numerical Simulation

Hotokezaka+, PRD (2011)

Equal-mass NS merger

HMNS

$M > M_{\rm max}$ Supported by centrifugal force

Type Ⅰ

Collapse to BH

Type Ⅱ

Lifetime < 5ms

Short lived

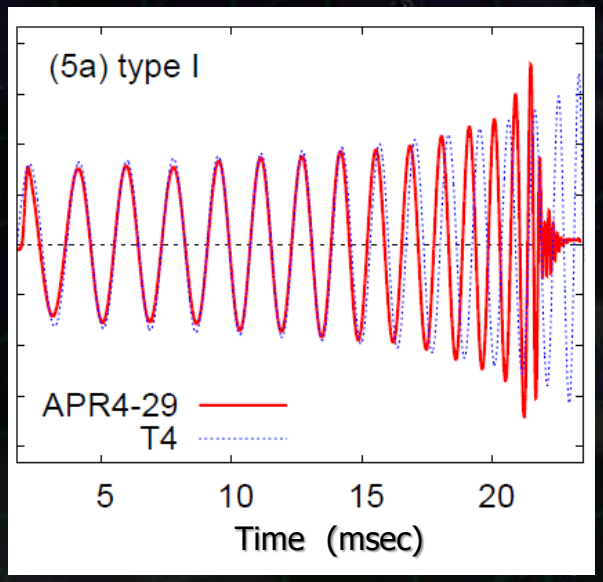
Collapse to BH

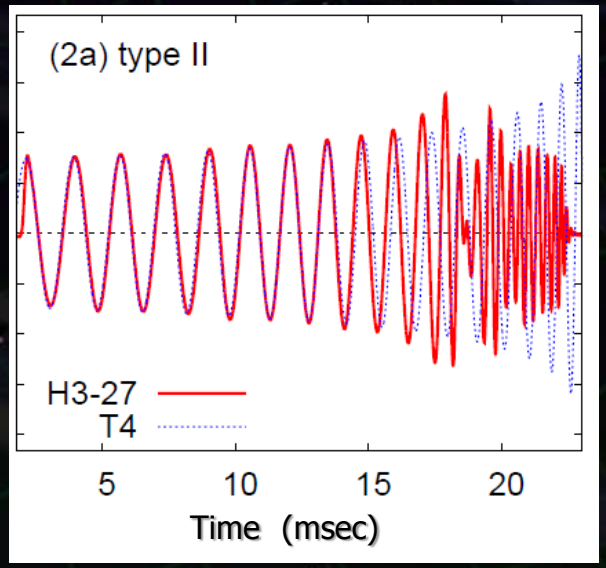
Type Ⅲ

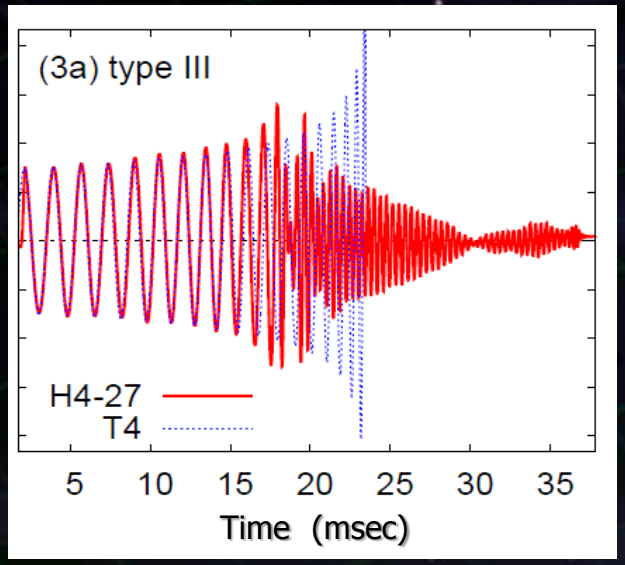
Lifetime > 5ms

Long lived

Collapse to BH







KAGRA Project

KAGRA（かぐら）

LCGT

Large-scale Cryogenic Gravitational-wave Telescope

2nd generation GW detector in Japan

Large-scale Detector

Baseline length: 3km

High-power Interferometer

Cryogenic interferometer

Mirror temperature: 20K

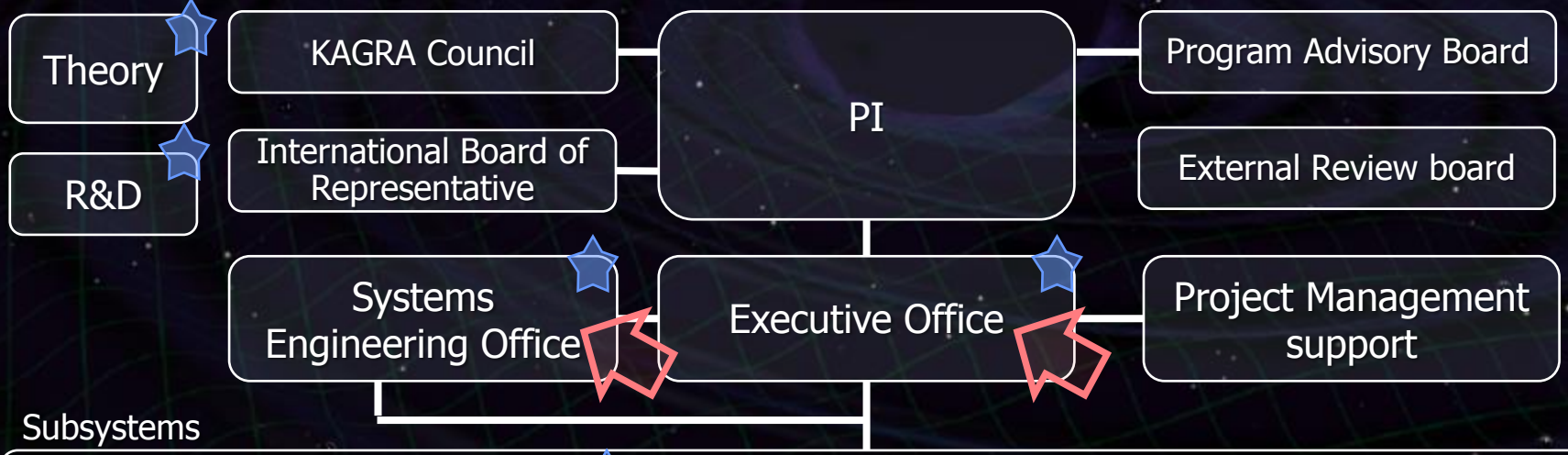
Underground site

Kamioka mine,

1000m underground

Organization of KAGRA

~150 Collaborators (Host : ICRR, Co-hosts: NAOJ, KEK)

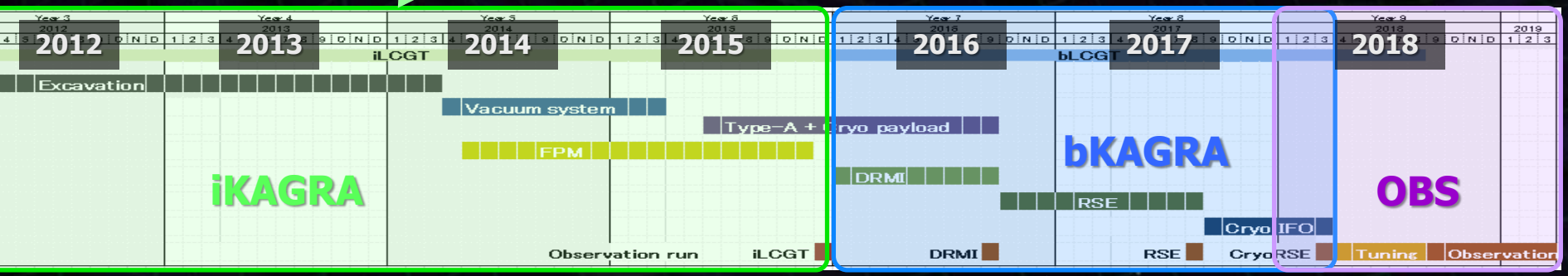


Subsystems

Tunnel (TUN)	Facility (FCL)	Vacuum (VAC)	Cryogenics (CRY)	Vibration Isolation (VIS)
Mirror (MIR)	Laser (LAS)	Main Interferometer (MIF)	Input-output Optics (IOO)	Auxiliary Optics (AOS)
Analog Electronics (AEL)	Digital System (DGS)	Data Management (DMG)	Data Analysis (DAS)	Geophysics Interferometer (GIF)

KAGRA Schedule

- **iKAGRA** (2010.10 – 2015.12)
  3-km FPM interferometer
  - Baseline 3km room temp.
  - Operation of total system with simplified IFO and VIS.



- **bKAGRA** (2016.1 – 2018.3)
  Operation with full config.
  - Final IFO+VIS configuration
  - Cryogenic operation.

Schedule and Budget

	FY2010	FY2011	FY2012	FY2013	FY2014	FY2015	FY2016	FY2017
Budget	'Leading-edge Research Infrastructure' program (~98M$) for iKAGRA			'Specially Promoted Research' program (~5M$) for detector upgrade				
	Budget from MEXT (~33M$) for excavation			Budget from MEXT (~20M$) for detector upgrade				

KAGRA configuration

iKAGRA

- KAGRA facility
- 3km simple room-temp. interferometer

Upgrade

bKAGRA

- Cryogenic mirrors
- Full-power RSE configuration

Purpose

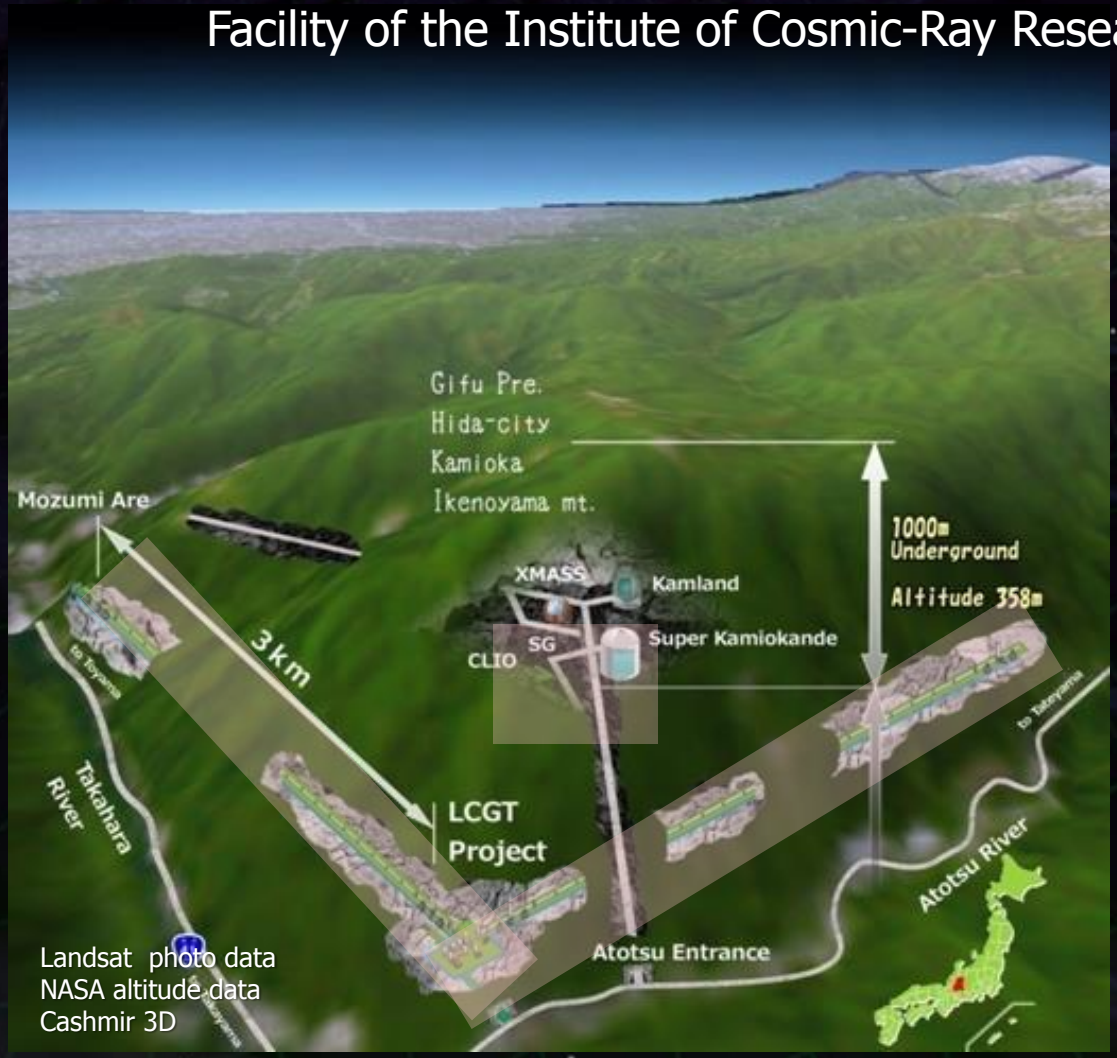
Preparation of infrastructure

GW detection and astronomy

KAGRA Site

Underground site at Kamioka, Gifu prefecture

Facility of the Institute of Cosmic-Ray Research (ICRR), Univ. of Tokyo.



Neutrino
- Super Kamiokande, Kamland

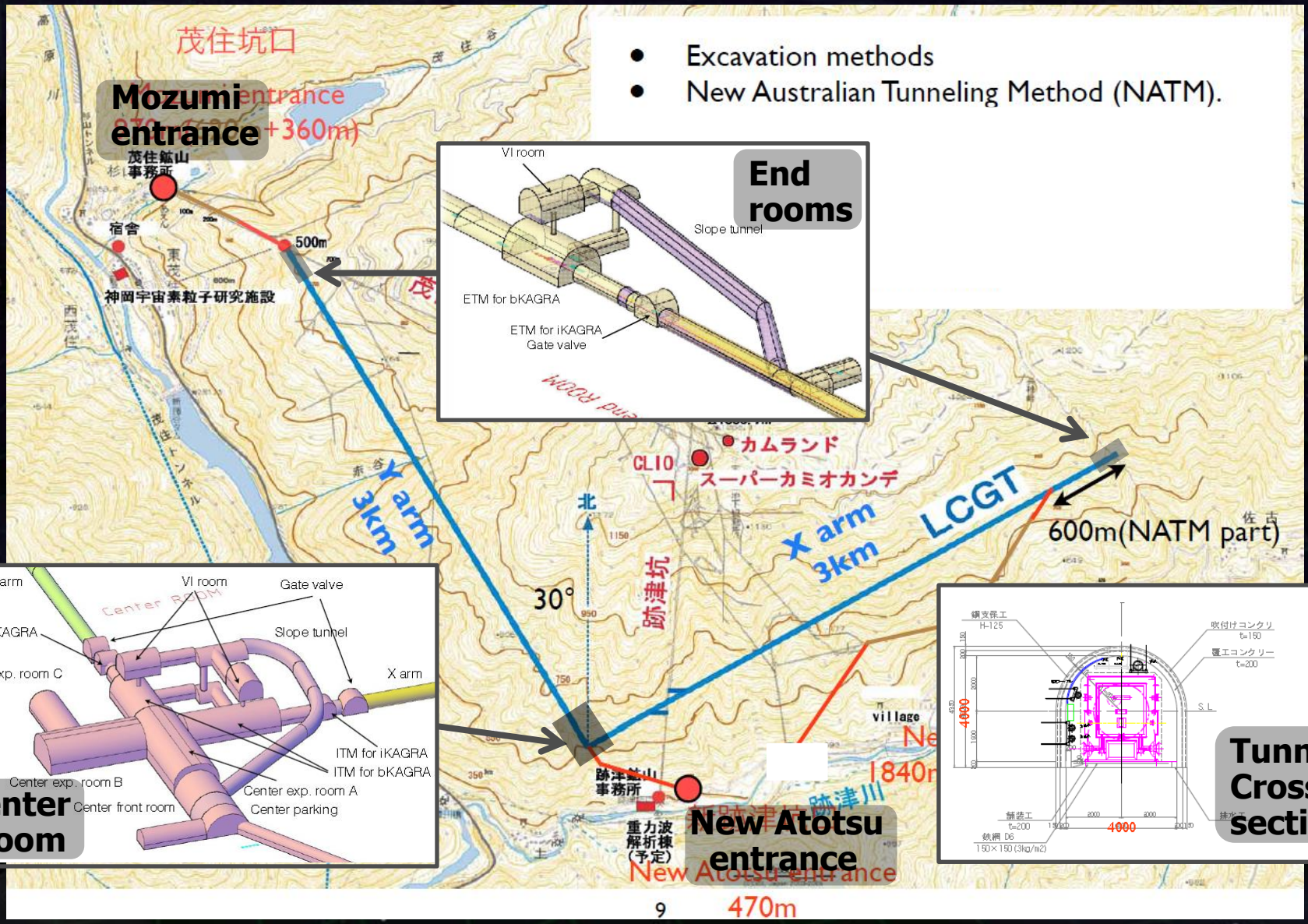
Dark matter
- XMASS

Gravitational wave
- CLIO, KAGRA

Geophysics
- Strain meter

- 220km away from Tokyo
- 1000m underground from the top of the mountain. (Near Super Kamiokande)
- 360m altitude
- Hard rock of Hida gneiss (5 [km/sec] sound speed)

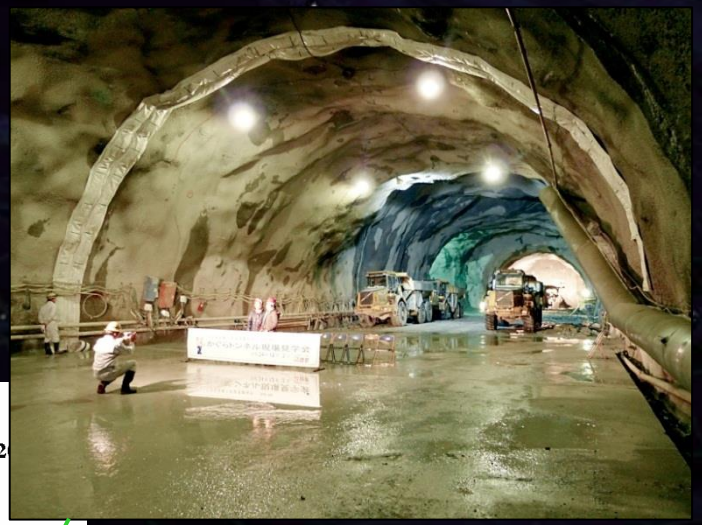
Tunnel Design

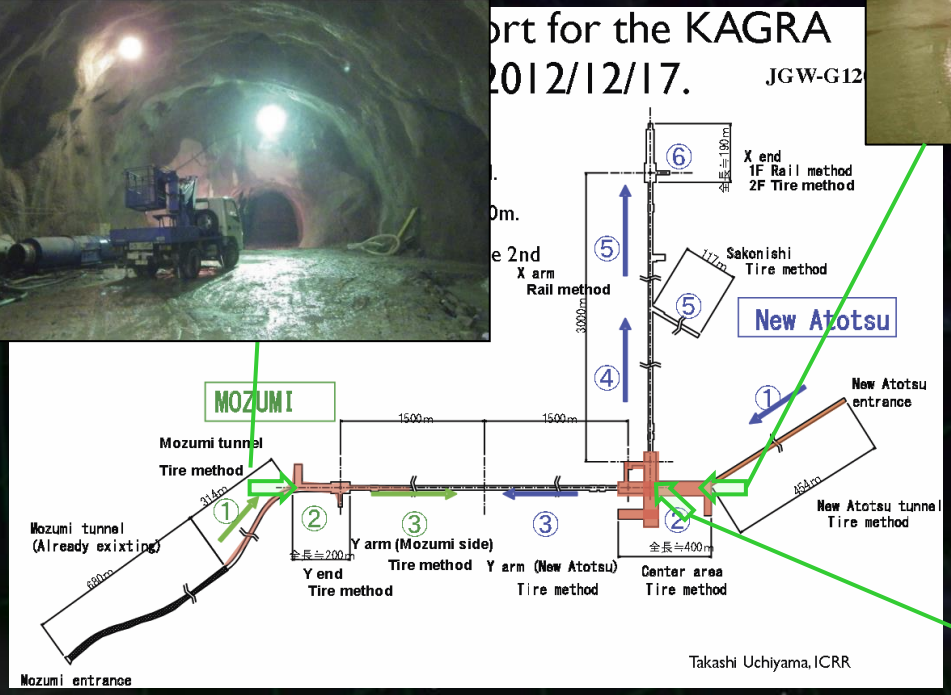


Status of Tunnel Excavation

Mozumi :
Y-arm tunnel
(~1000m)

New Atotsu：Center room





Surface Facility at Kamioka

Rent and remodel a public building (140m^2) for free.

→ On-site office and laboratory for GW group.

Aug. 29, 2012
Announcement for local people
→ Open as office in Nov.

KAGRA Vacuum duct

- 12m, Φ800mm ducts for 3km x 2 arms.
  → ~90% of 478 ducts have been delivered.

Press to form a duct

Bellows for each duct

Presentation
By Y.Saito (KEK)

Transportation to Kamioka

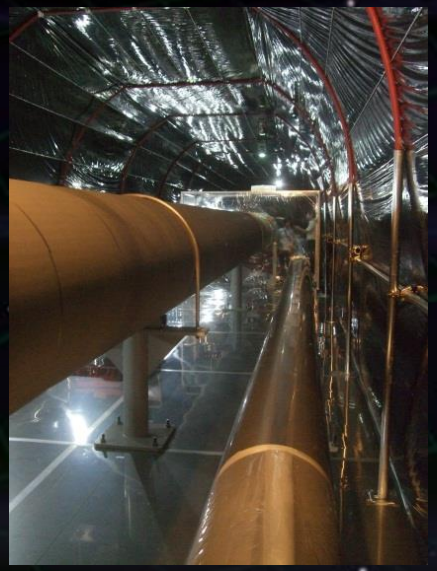
Baking at MIRAPRO Co.
Noda/MESCO, Kamioka

Test at MIRAPRO Co. Noda

Installation Test Facility

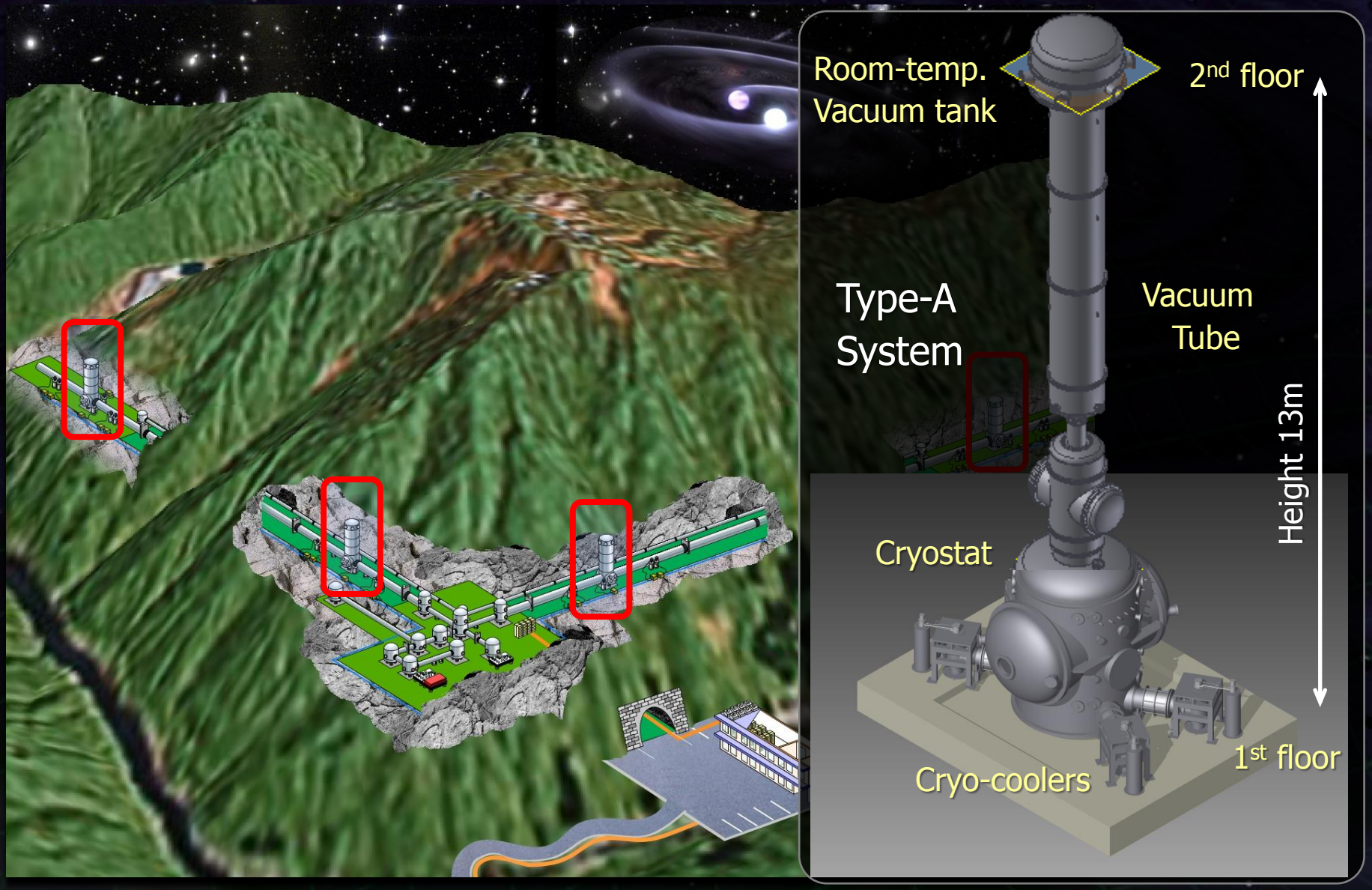
KAGRA tunnel simulator for installation test

(MIRAPRO, Noda factory)



June 28, 2012, Photo by Kamiizumi and Iwasaki (ICRR)

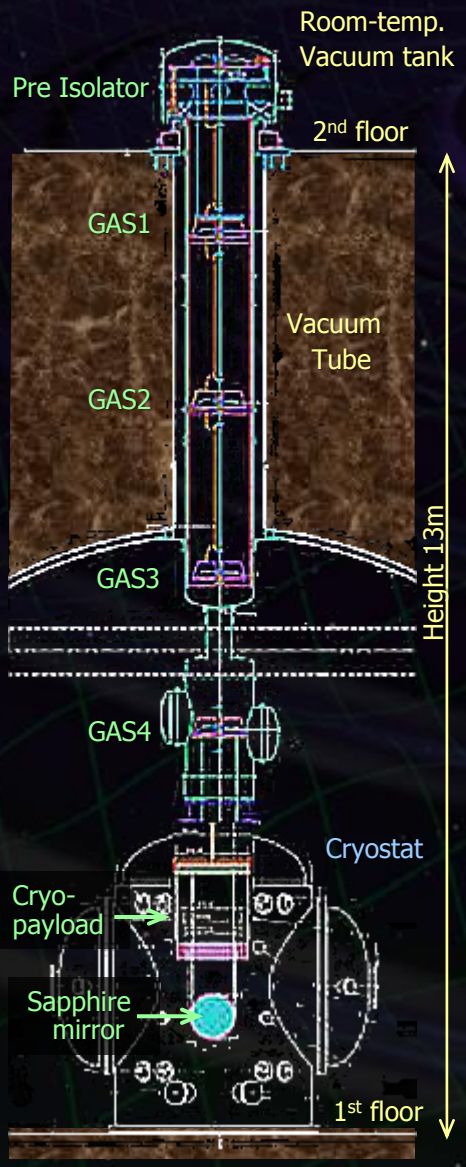
Cryogenic Isolator

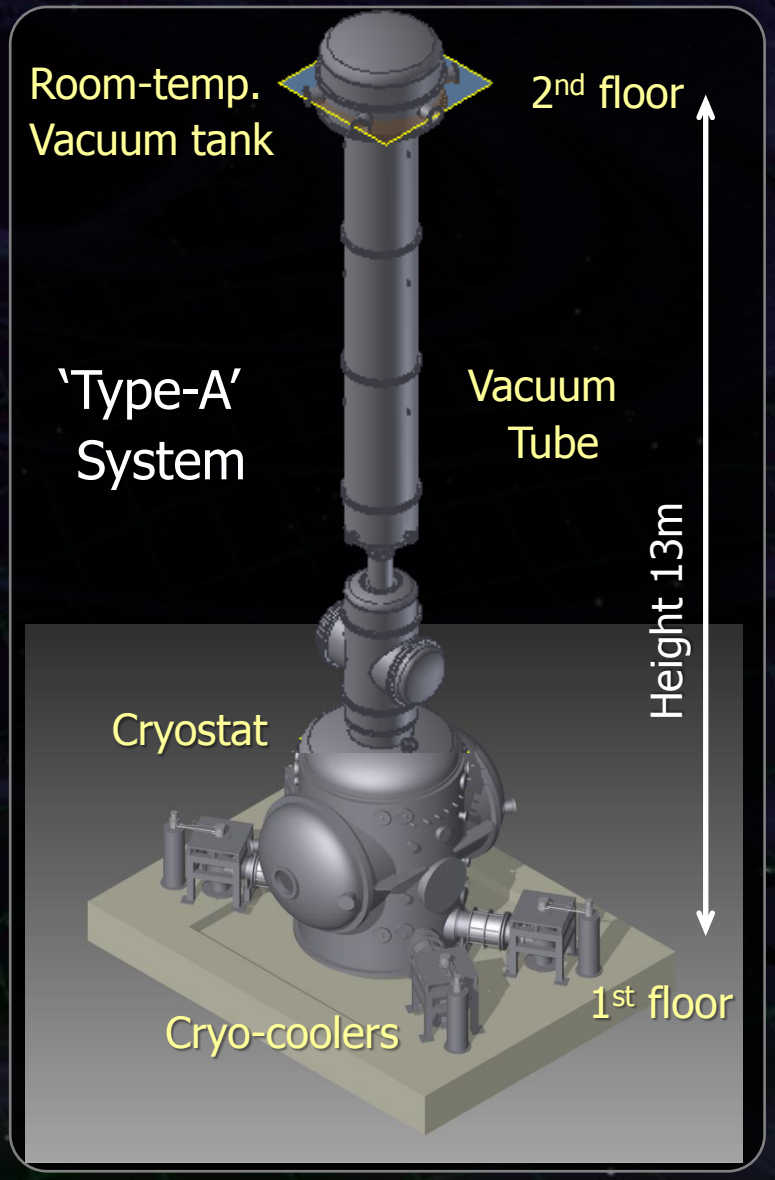


Cryogenic Mirror Isolator

Room-temp.
Filter chain

Cryogenic
Payload





Cryostat Construction

Cryostat #1 in preparation for installation of radiation shield.

Cryostat #2 in leak test.

3rd and 4th cryostats under construction

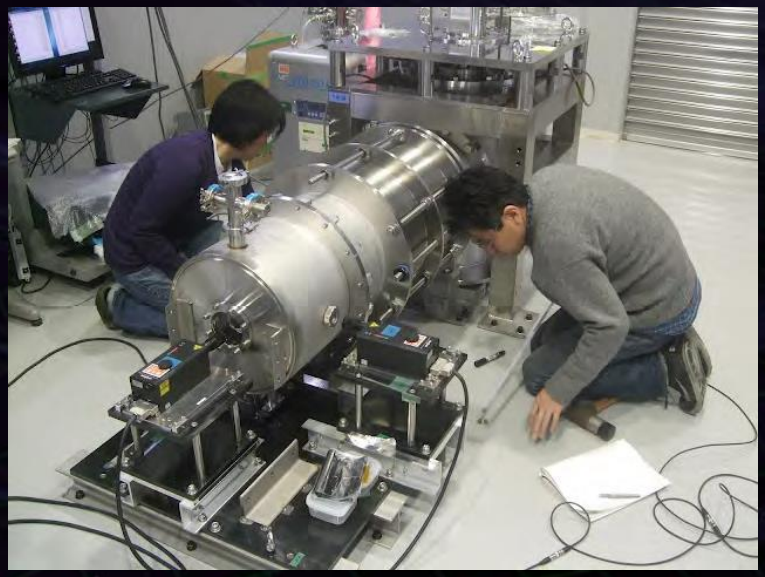
Radiation shield

Toshiba Keihin Factory (Oct 31, 2012)

Cryo-cooler Construction

Cryo-cooler units at ICRR (Kashiwa)

Vibration measurement



Storage at ICRR (Akeno)

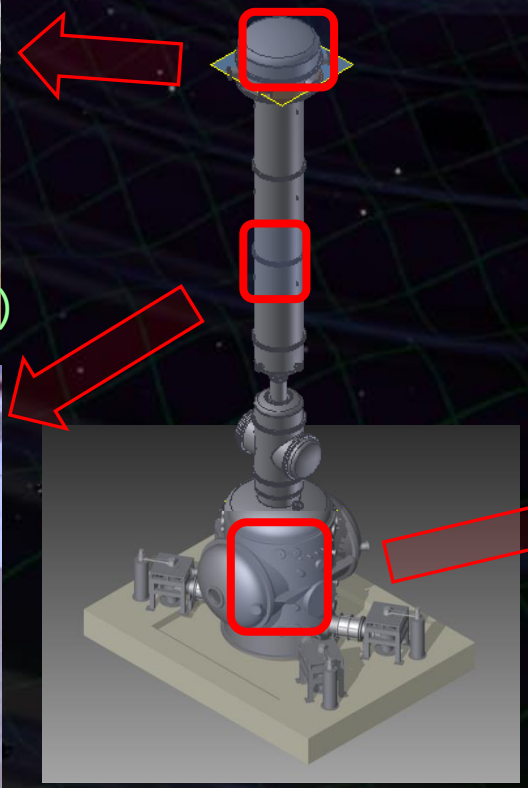
T.Suzuki at External Review (April 2012)

Sapphire Mirror Isolator

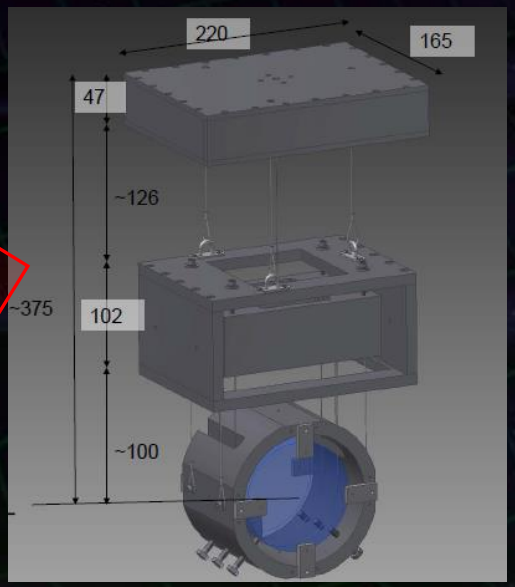
Pre Isolator prototype at Lucca(Pisa)

GAS filter prototype

'Type-A' system



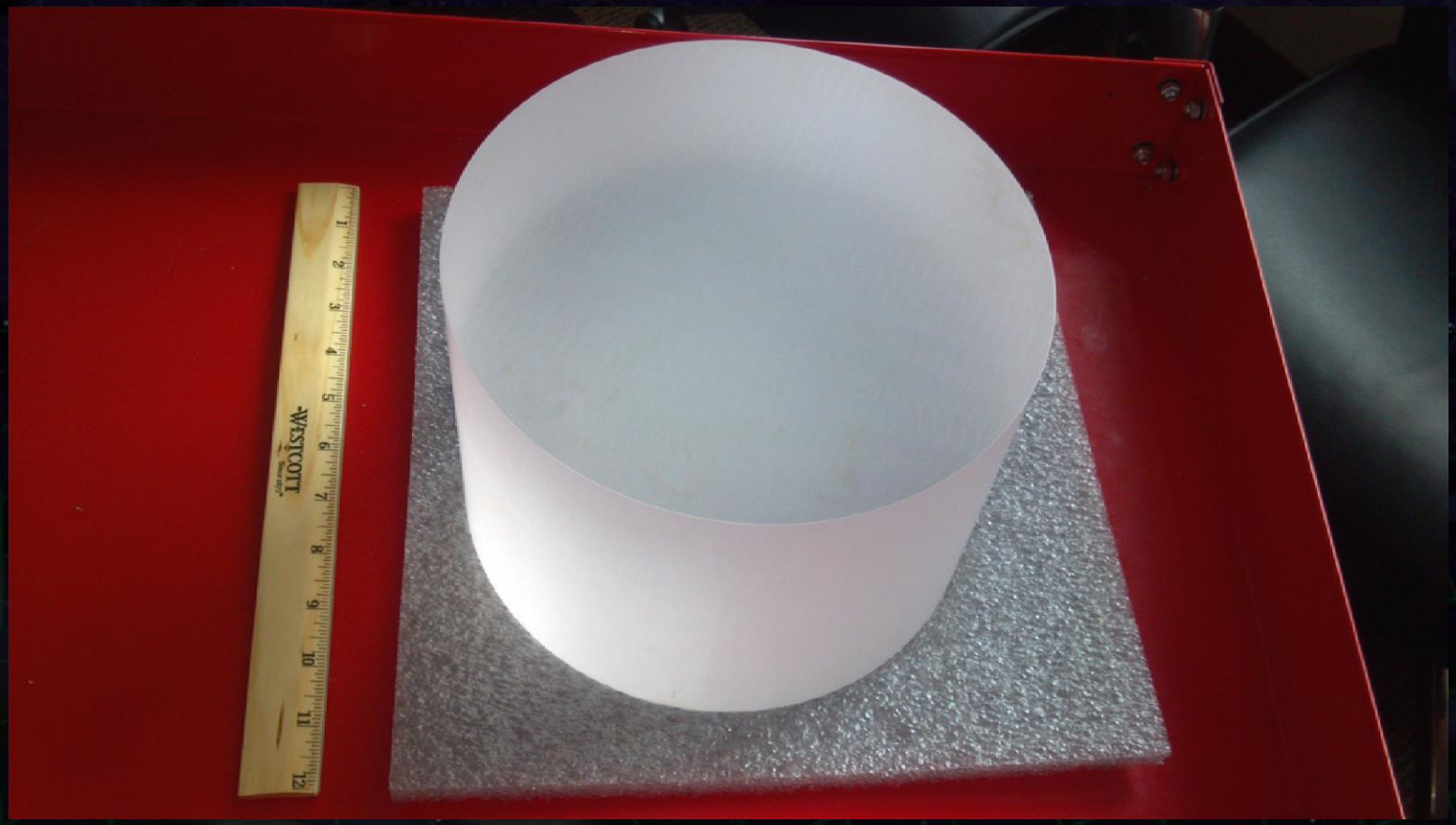
Cryogenic payload
½-scale prototype



Isolator Prototypes

Pre-Isolator Prototype test (Kashiwa)

Sapphire Mirror



2 Sapphire substrates were delivered
(Φ220mm, t 150mm, c-axis)

Summary

Summary

KAGRA : Under Construction

- Sufficient sensitivity for direct GW detection
- Form global network as one of the 2nd-gen. detectors

⇨ Aim to detect GW, and to open new astronomy

- KAGRA will demonstrate 3rd generation detector techniques: cryogenics and underground

Status

- Technology based on TAMA and CLIO experiences
- Tunnel and facilities are becoming real.
- Prototype developments : SAS, Cryostat, Control Sys.

Experiences in GCOE

For me…

- Enjoyed life in Kyoto.
- Research experiences in Kyoto was vary fruitful.
  - 'Tentaikaku' astrophysics theory group:
    → Interactions with theorists
  - Quantum Optics group.
    → Precise atomic spectroscopy for fundamental phys.

⇩

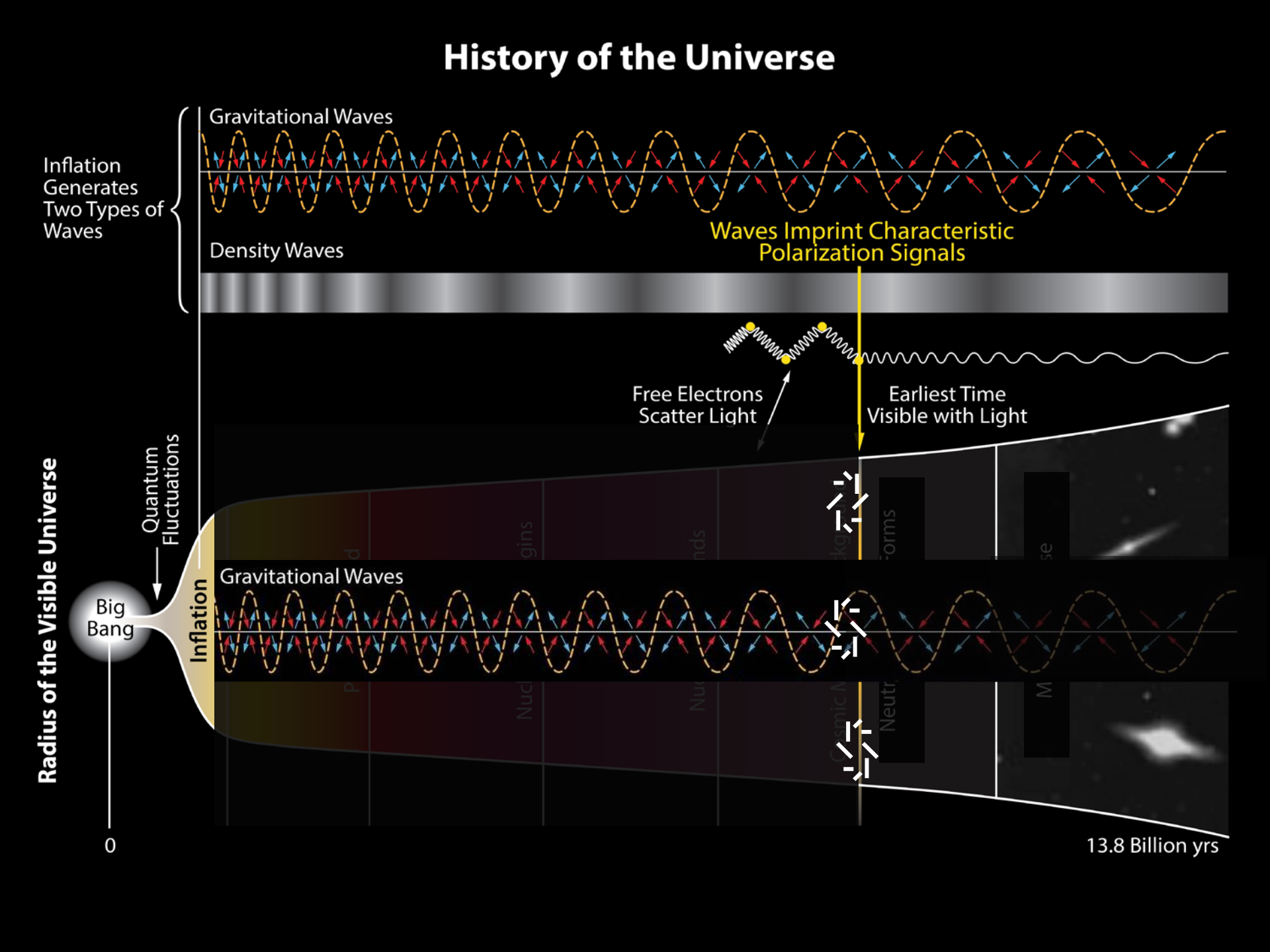
Opportunities to tackle on new fields

Thank You Very Much!!!

End

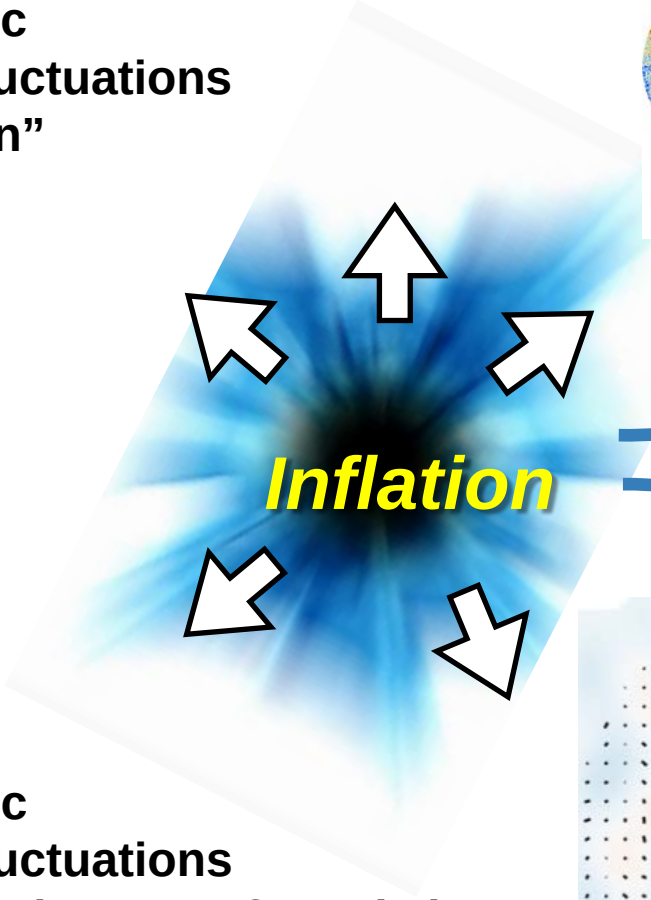
Inflationary cosmology with BICEPs

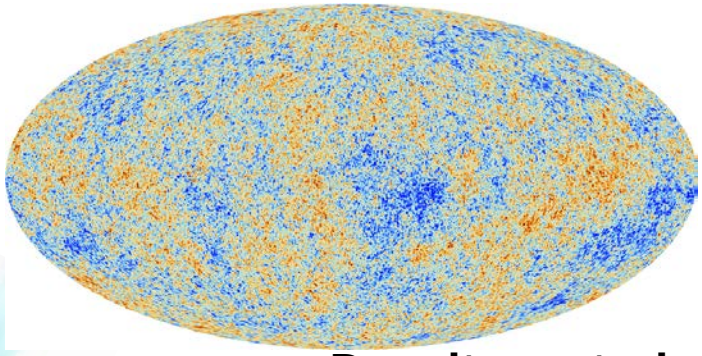
Chao-Lin Kuo
Stanford University
SLAC National Accelerator Lab


History of the Universe
Inflation Generates Two Types of Waves
Gravitational Waves
Density Waves
Waves Imprint Characteristic Polarization Signals
Free Electrons Scatter Light
Earliest Time Visible with Light
Quantum Fluctuations
Big Bang
Inflation
Gravitational Waves
Radius of the Visible Universe
0
13.8 Billion yrs

Density perturbations and gravitational waves

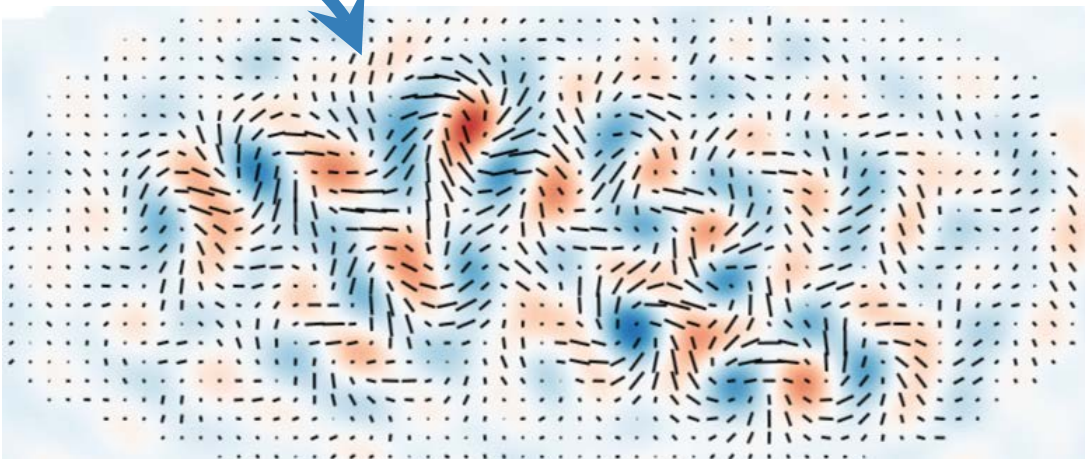
Sub-atomic vacuum fluctuations of "inflaton"





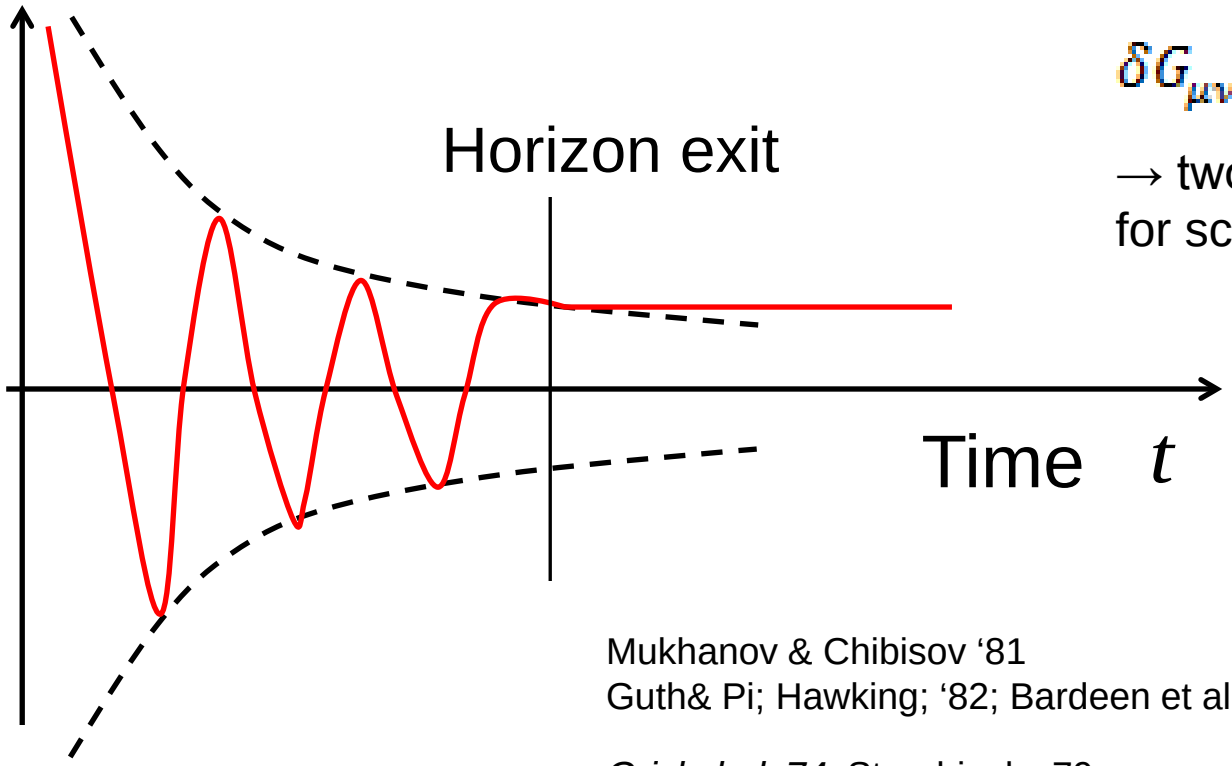
Density perturbations *studied* by Planck, WMAP, SPT, etc.

Gravitational waves



Sub-atomic vacuum fluctuations of *graviton* (quanta of gravity)

Generation of ***scalar/tensor*** perturbations



$$\delta G_{\mu\nu} = 8\pi G \delta T_{\mu\nu}$$

→ two linear wave equations for scalar /tensor

Mukhanov & Chibisov ‘81
Guth& Pi; Hawking; ‘82; Bardeen et al., ’83, Sasaki ‘83

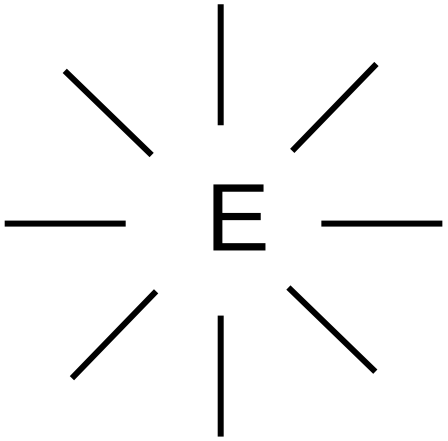
Grishchuk 74; Starobinsky 79
Rubakov et al, 82; Frabri & Pollock , 82

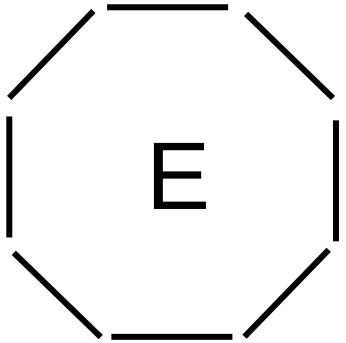
Quantum fluctuations in the vacuum state of the *inflaton*/*graviton* fixes the r.m.s of the linear solutions

Inflationary B-modes, known as the "Holy Grail" of cosmology

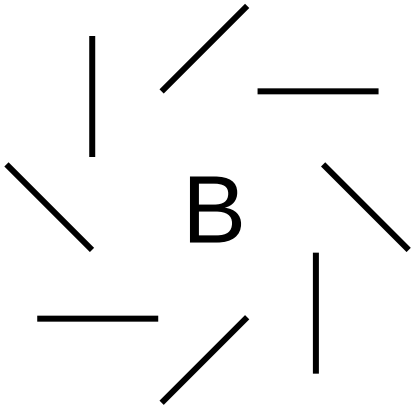
- Started out as *graviton* vacuum fluctuations
- Energy scale of inflation ~ expansion rate ~ GW amplitude
- Alternative models generate no GW
- *Field range and "UV" completeness*

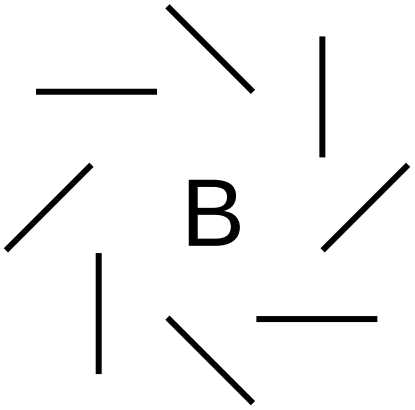
Only gravitational waves can generate B-modes

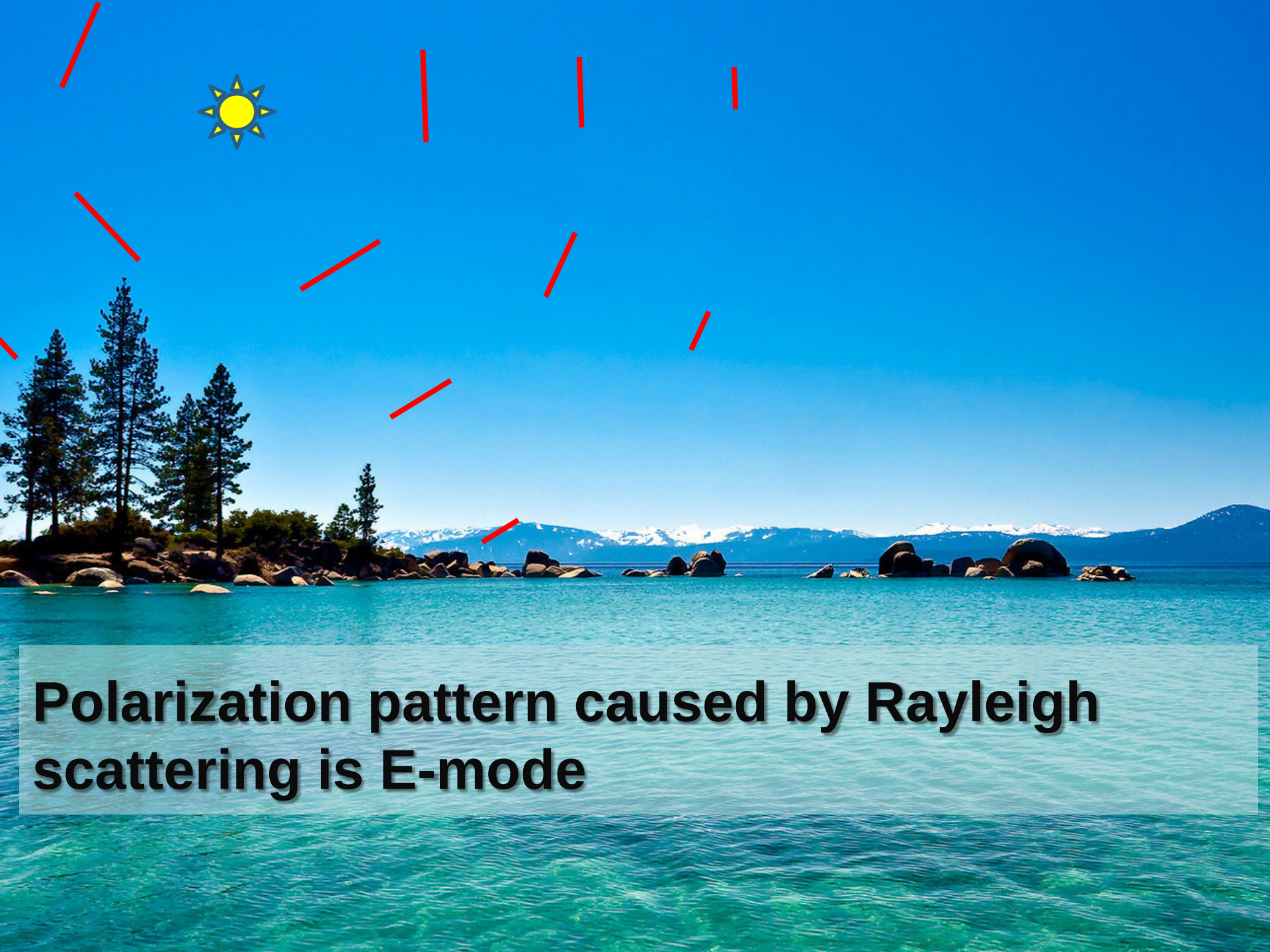


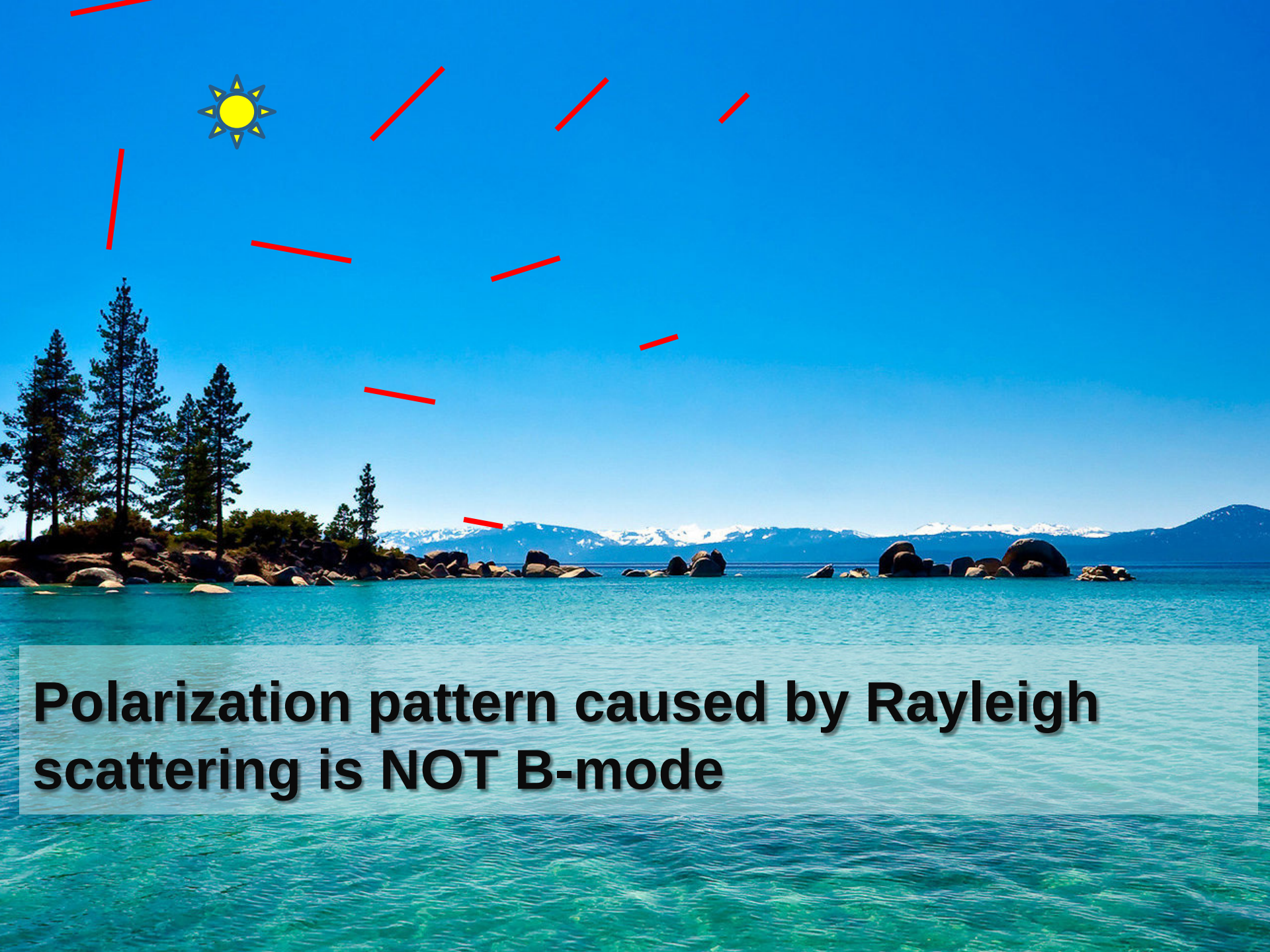


Seljak & Zaldarriaga ‘97
Kamionkowski, Kosowsky, Stebbins ‘97



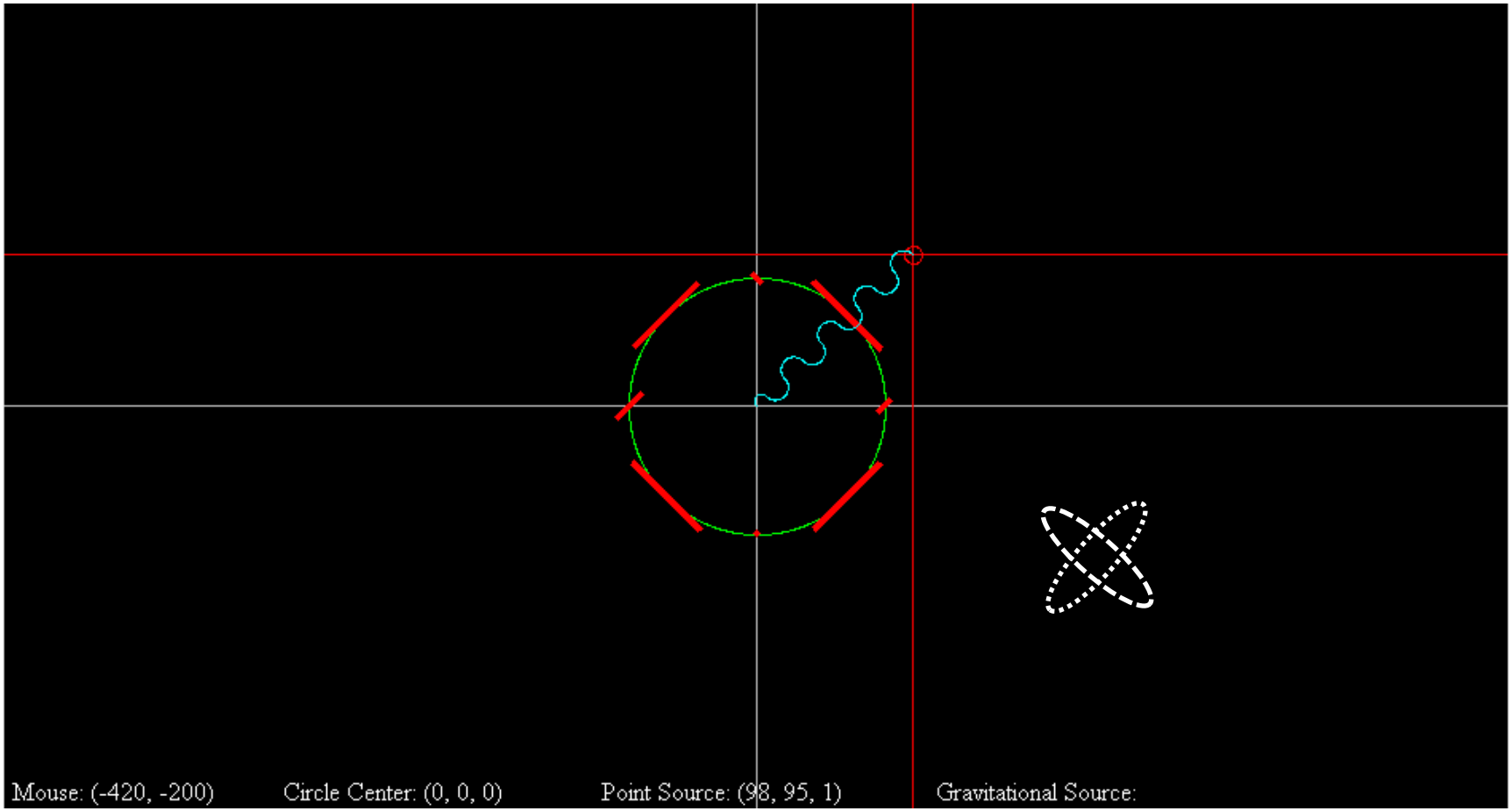



Polarization pattern caused by Rayleigh scattering is E-mode

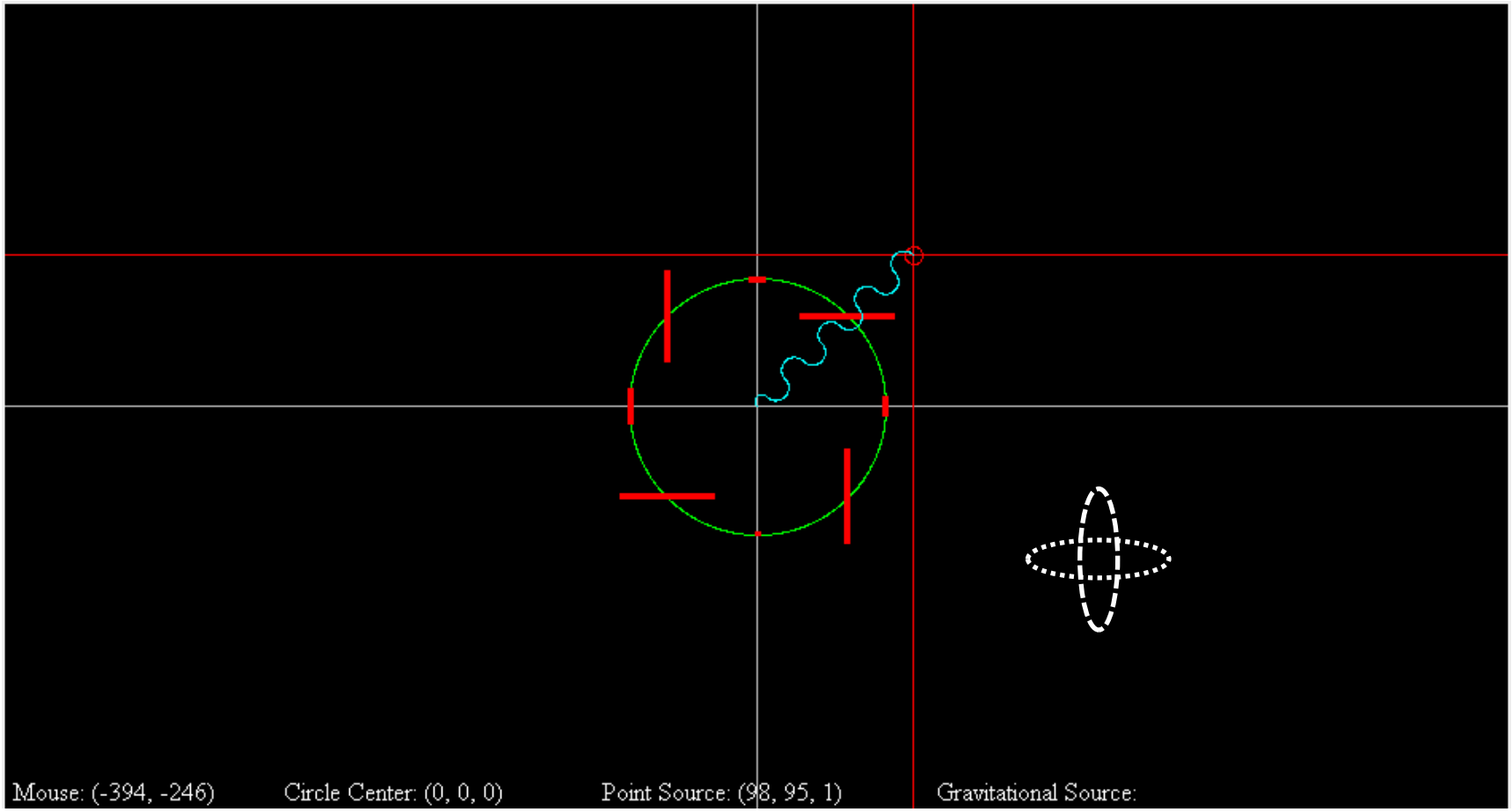


Polarization pattern caused by Rayleigh scattering is NOT B-mode

Gravitational waves generate *E*-mode polarization



Gravitational waves generate *B*-mode polarization



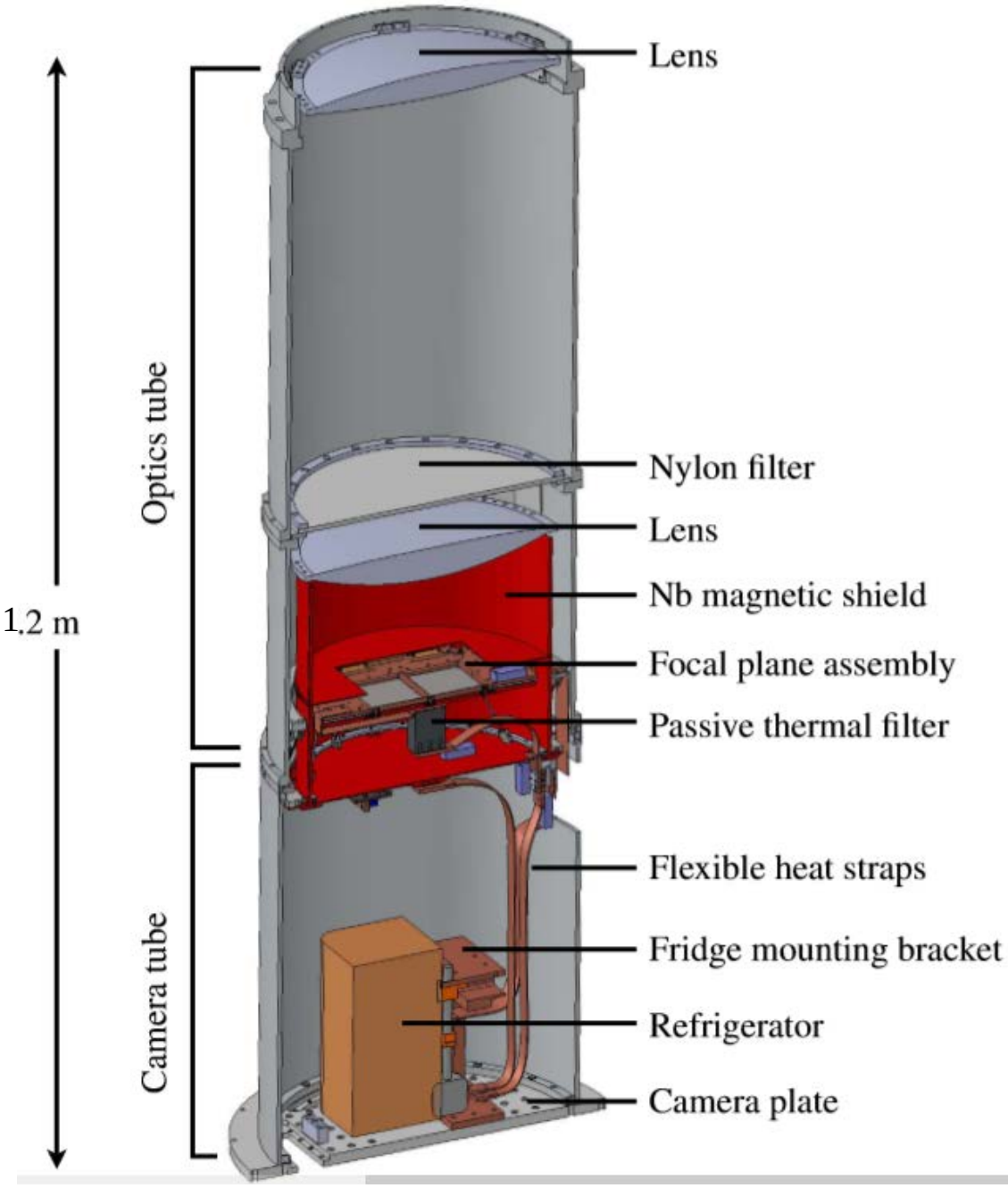
!!!

The polarization pattern is unique, but small

Vertical / Horizontal
differ by
1 part in 30,000,000

South Pole is the Mecca of CMB research (BICEP1, BICEP2, Keck Array, BICEP3)

- High, dry, cold, low water vapor in the atmosphere
- Stable climate for continuous 6 months
- Great logistical support (NSF-Office of Polar Program+ Lockheed)



BICEP/Keck series
BICEP1/2/3
Keck Array

microwave (95/150 GHz)
Superconducting sensors
Low temperature physics (0.25K)

Lithographic detectors
High packing density
Mass production

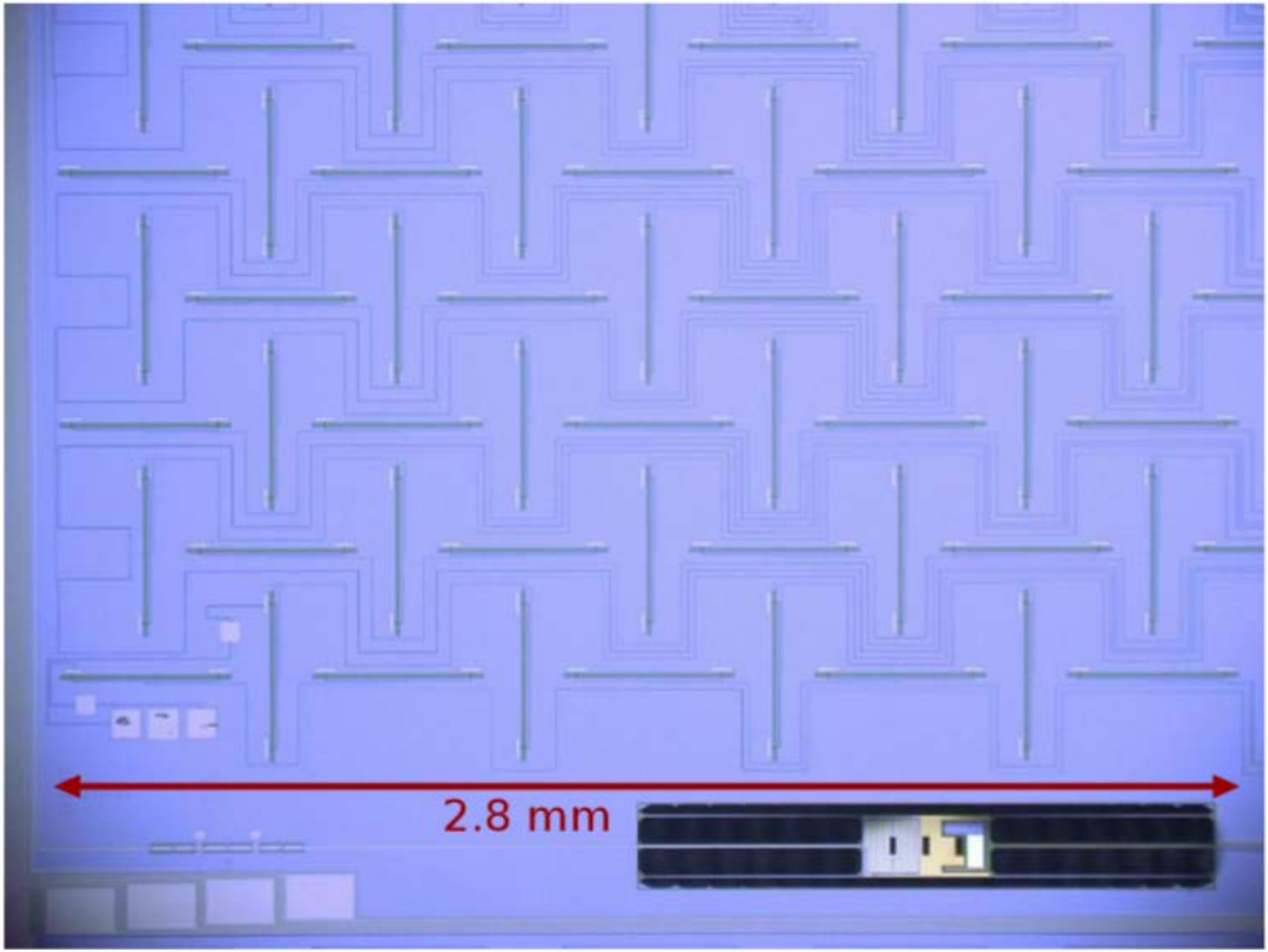
On going program + R&D for the future

BICEP1: 2006, 2007, 2008

BICEP2: 2010, 2011, 2012

Keck Array: 2011, 2012, 2013, …

BICEP3: 2015…

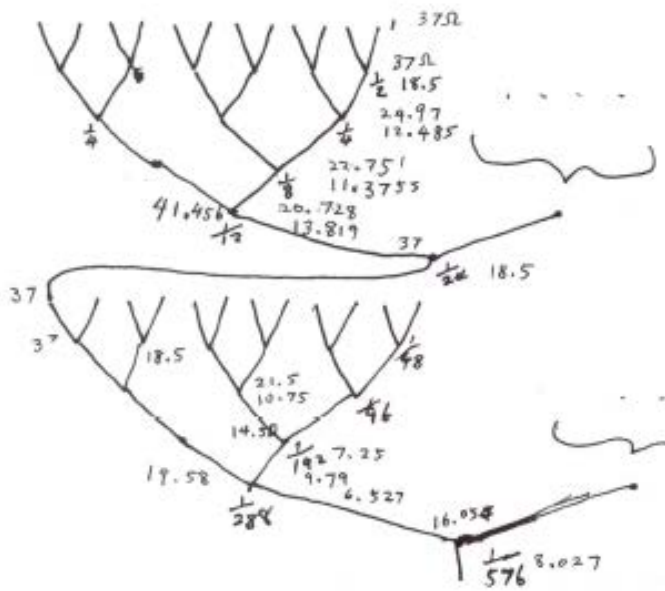

2.8 mm

9/21/06

Dual pol H-antenna

Peter "3"

impedance = 37 + 15j Ω



170 μm	
500	
500 × 2	
500 × 2	
500 × 6	
500 × 12	
170	
300	
300 × 2	
300 × 2	
300 × 6	

CLK notebook
a design was found
after 3 years
9/21/2006

~clkuo/polarimeter/feednetwork/taper/

Maximum reflectivity ~ 1% inband taper_sll_H_antenna_peter3

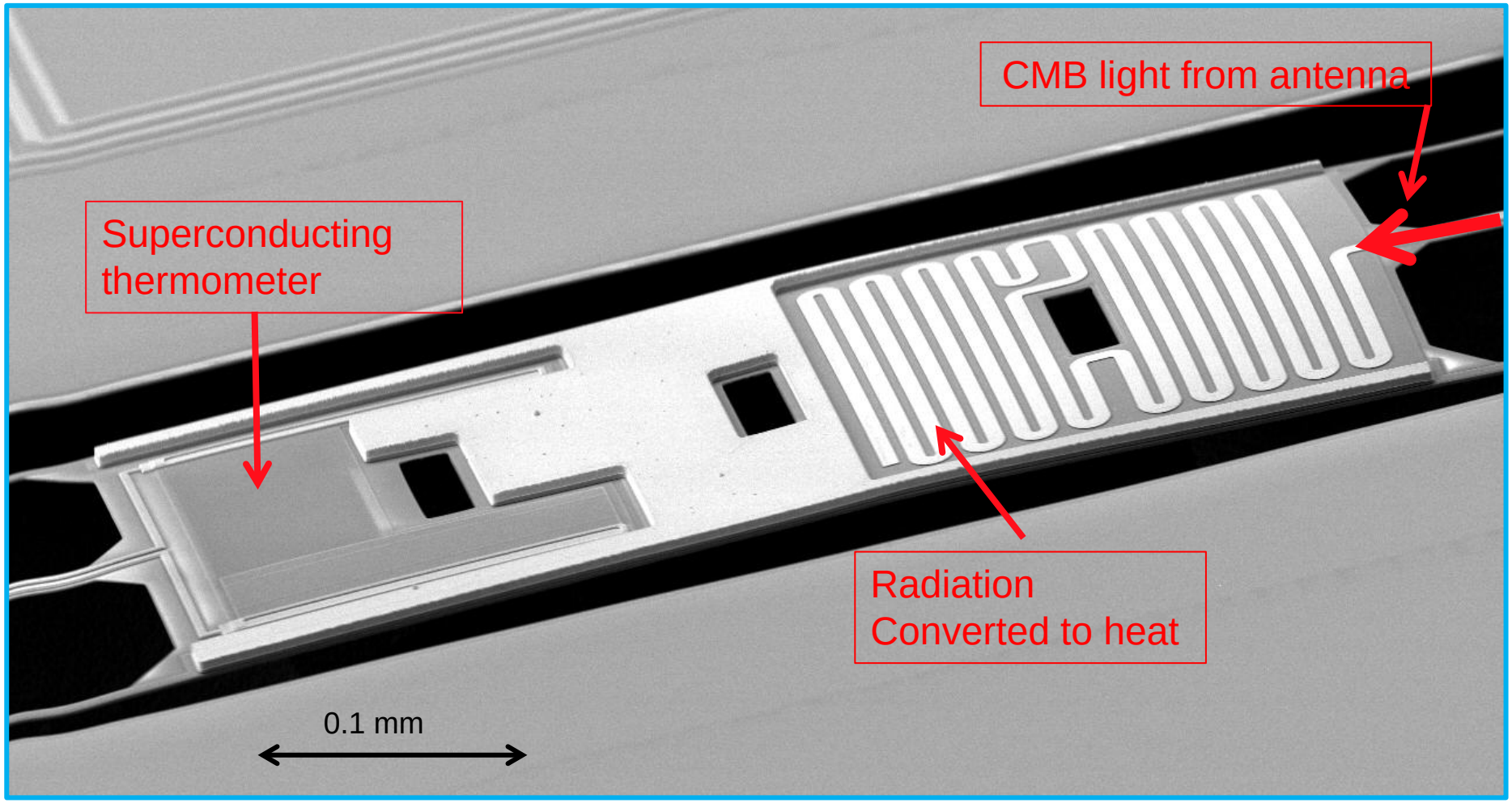
2nd level | ε = 4.5, thickness = 3000 Å | supermix | 1st level

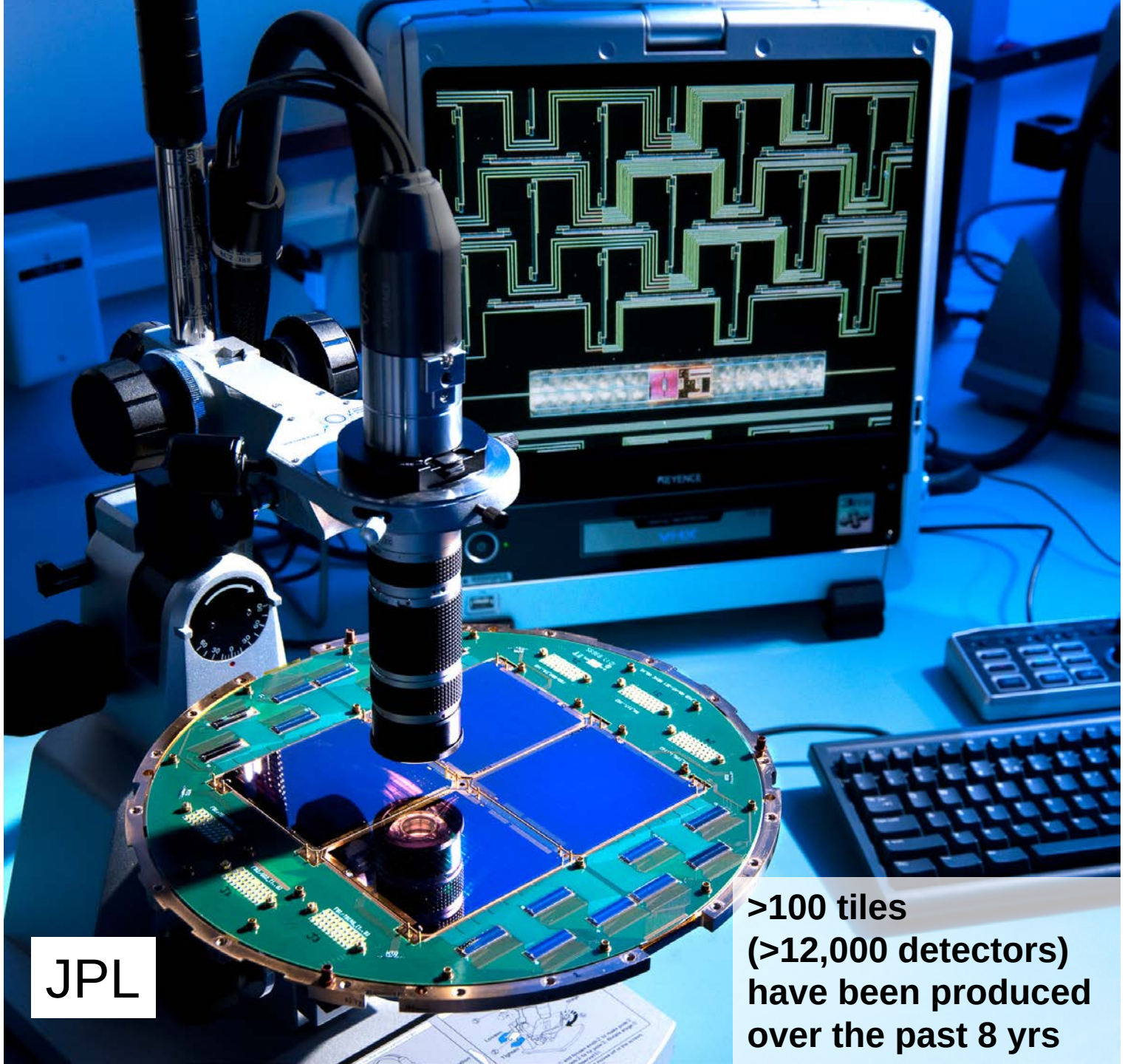
Z(Ω)	W(μm)
37	0.931
18.5	2.53
21.5	2.083
10.75	4.87
14.5	3.42
7.25	7.605
9.79	5.42

Z (Ω)	W (μm)
24.97	1.706
12.485	4.09
22.751	1.933
11.3755	4.562
20.728	2.185
13.819	3.625
41.456	0.765

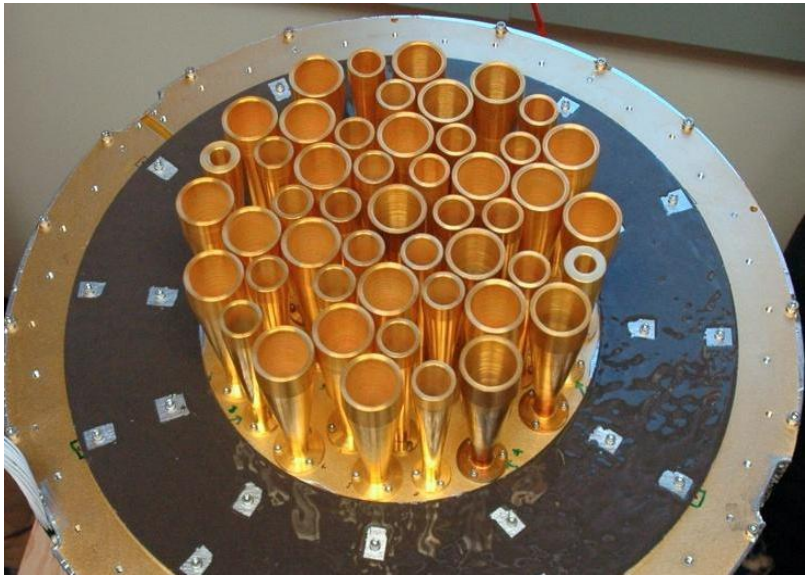
Detecting the CMB radiation

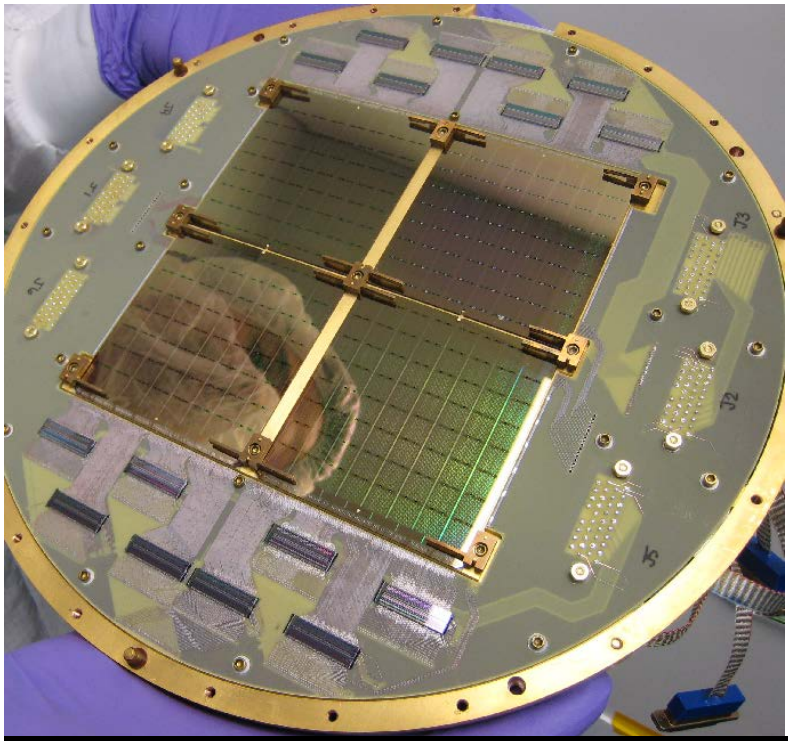
BICEP2 Detector: Transition-Edge Superconductor




JPL
>100 tiles
(>12,000 detectors)
have been produced
over the past 8 yrs

3 BICEP2 year = 30 BICEP1 years!

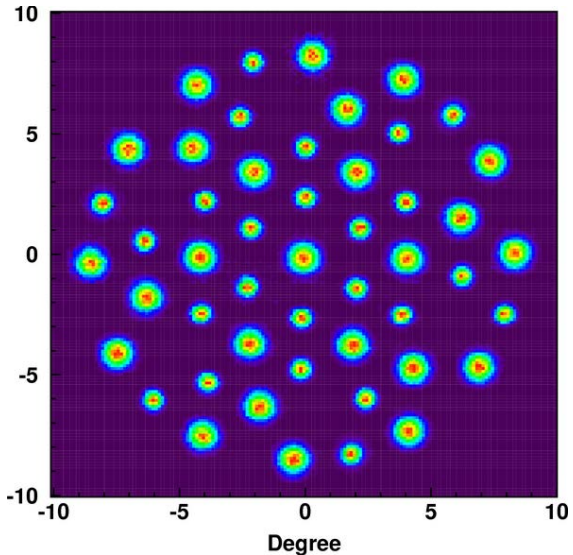


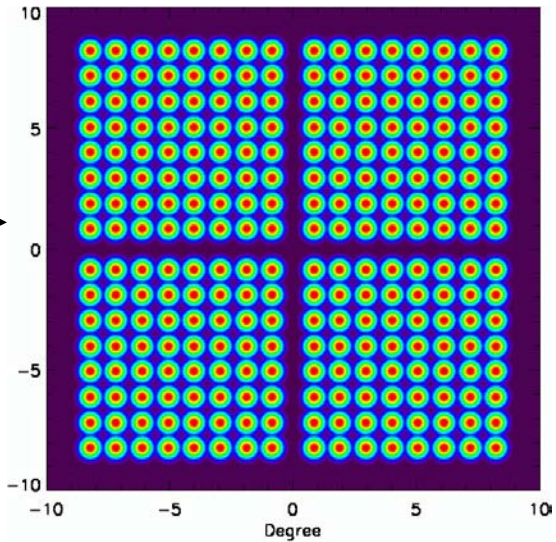


JPL : antenna-coupled TES arrays

BICEP1

48
150 GHz
detectors





BICEP2

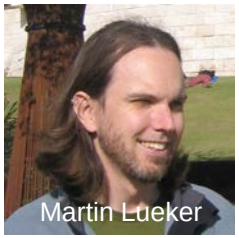
512
150 GHz
detectors

VE RI TAS
HARVARD
CALIFORNIA INSTITUTE OF TECHNOLOGY
1891
cea
UCSD
JPL
NIST
CARDIFF UNIVERSITY
UNIVERSITY OF TORONTO
UBC

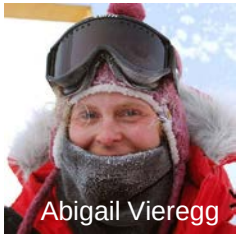
BICEP2 Postdocs

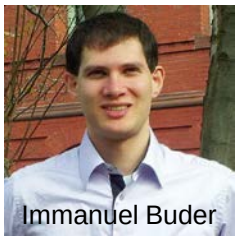
Colin Bischoff

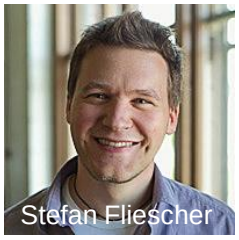
Jeff Filippini


Martin Lueker

Walt Ogburn


Abigail Vieregg


Immanuel Buder


Stefan Fliescher

Roger O'Brient

Angiola Orlando

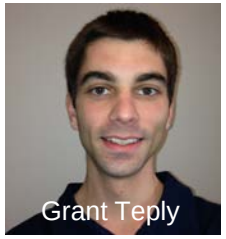
Zak Staniszewski

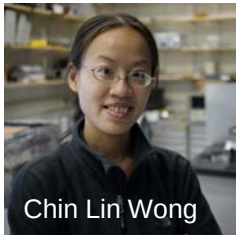
BICEP2 Graduate Students

Randol Aikin

Justus Brevik

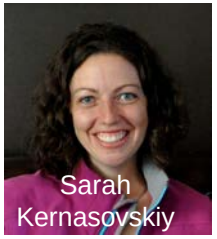
Chris Sheehy


Grant Teply


Chin Lin Wong

Kirit Karkare

Jon Kaufman


Sarah Kernasovskiy

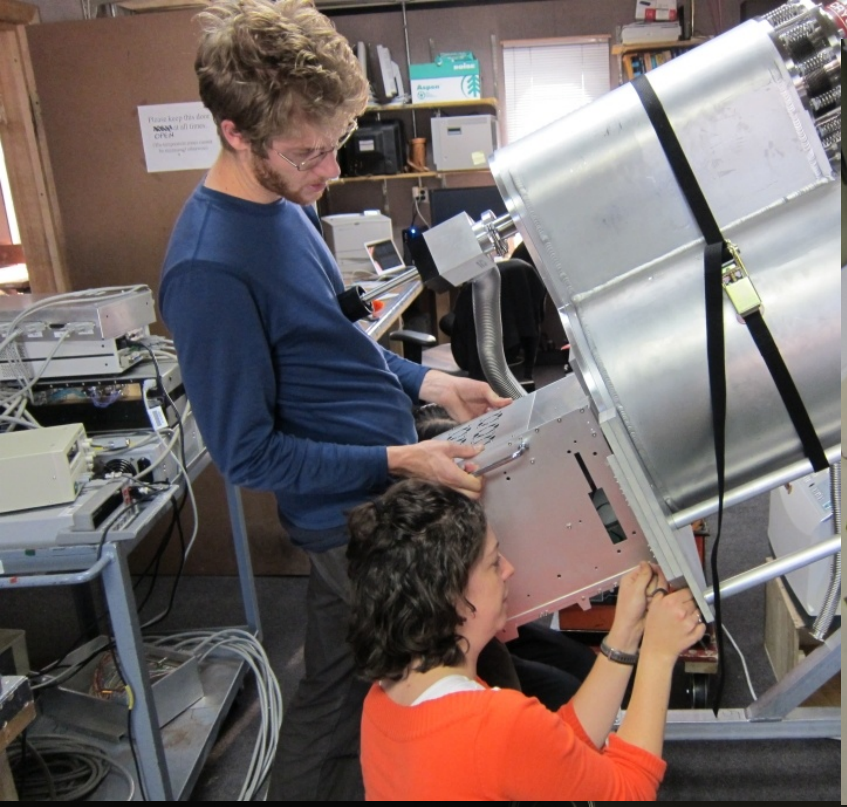
Jamie Tolan

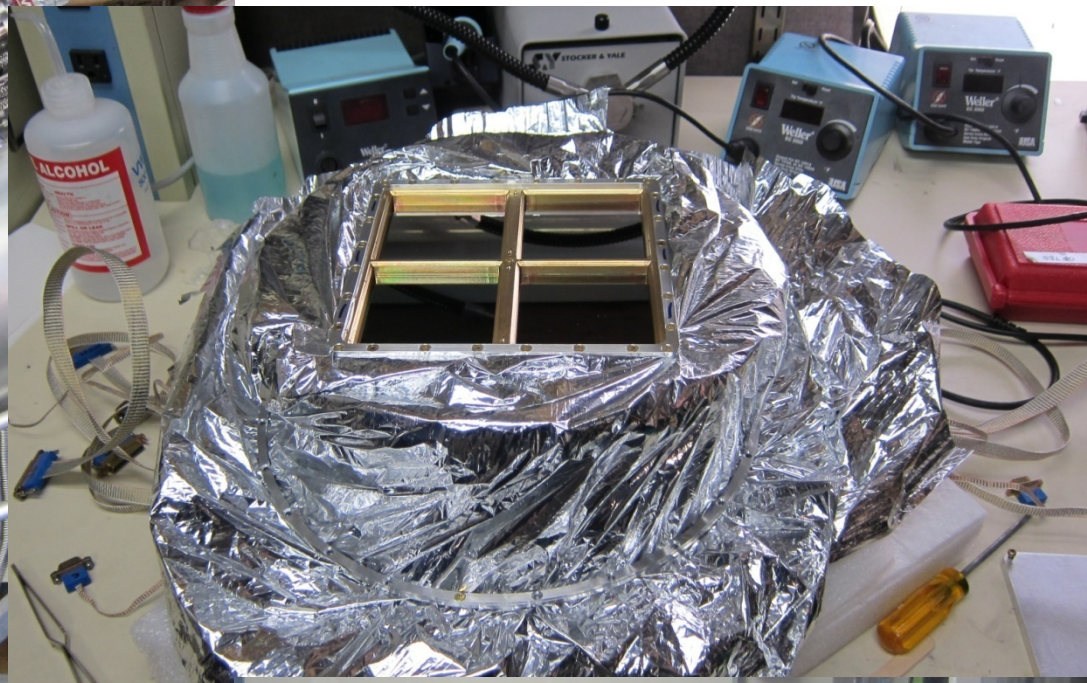
BICEP2 Winterovers

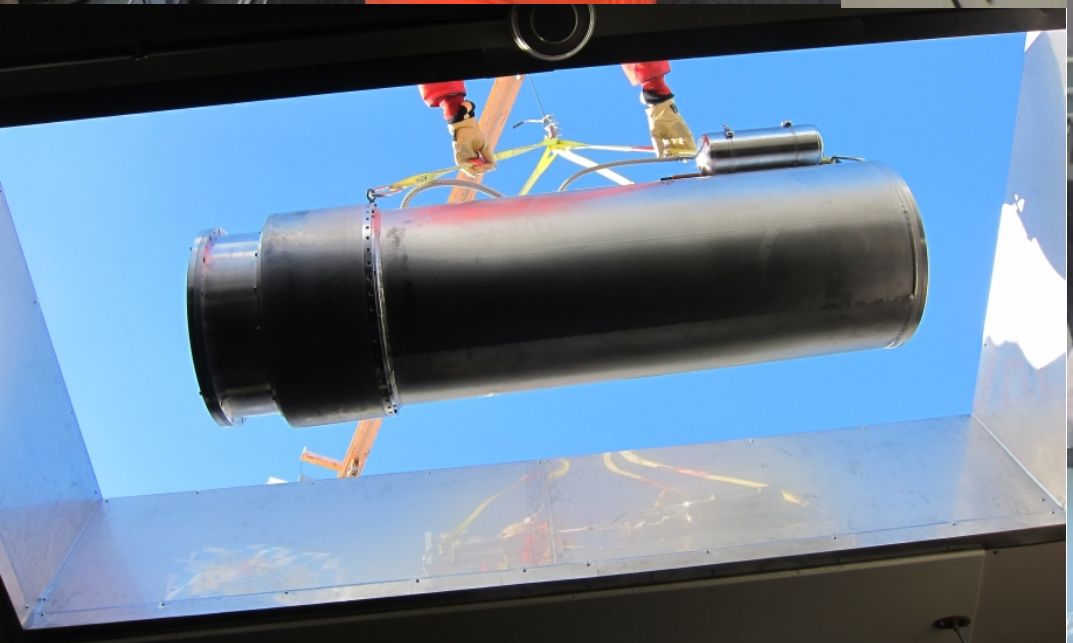
Steffen Richter 2010

Steffen Richter 2011

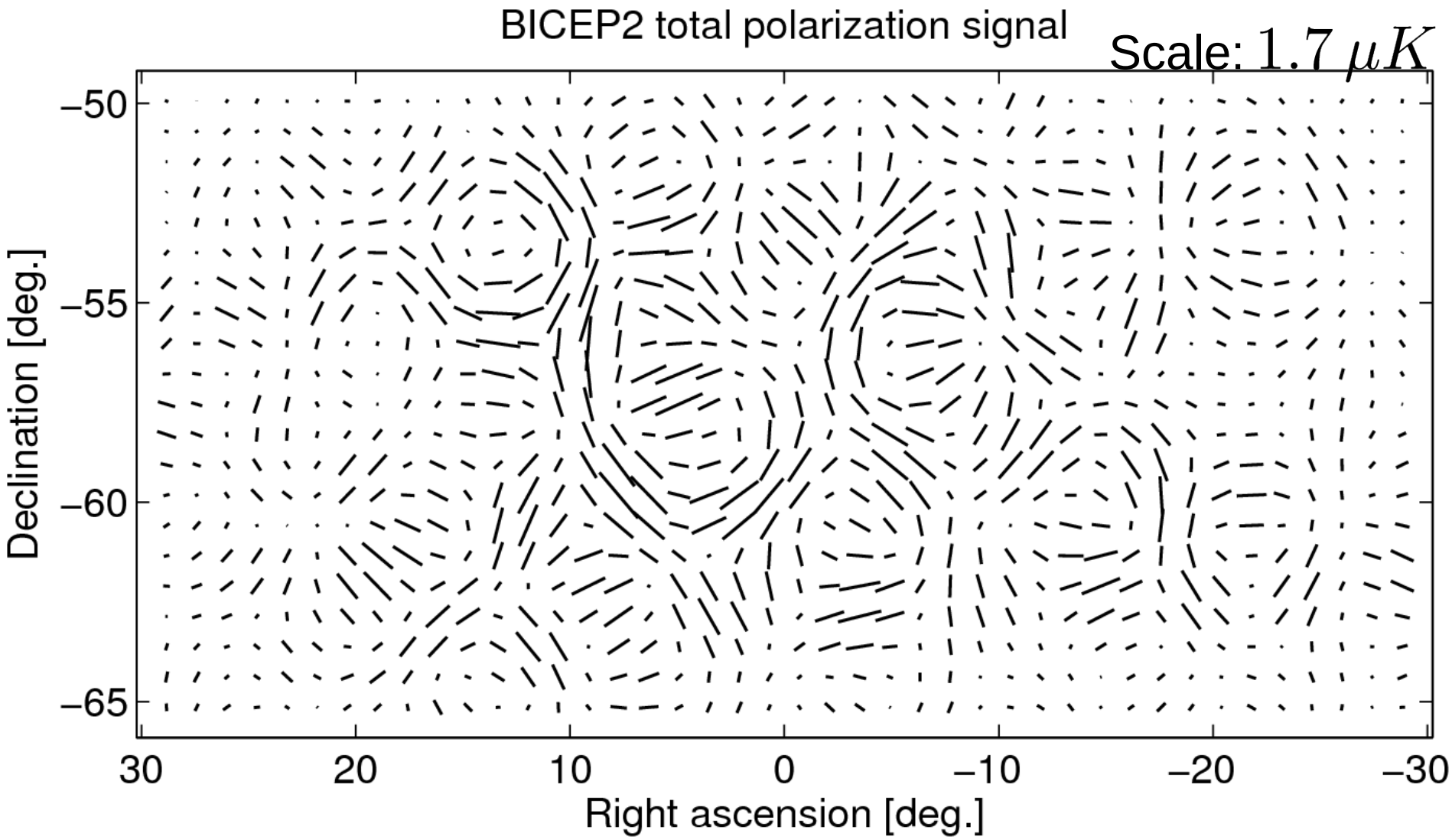
Steffen Richter 2012



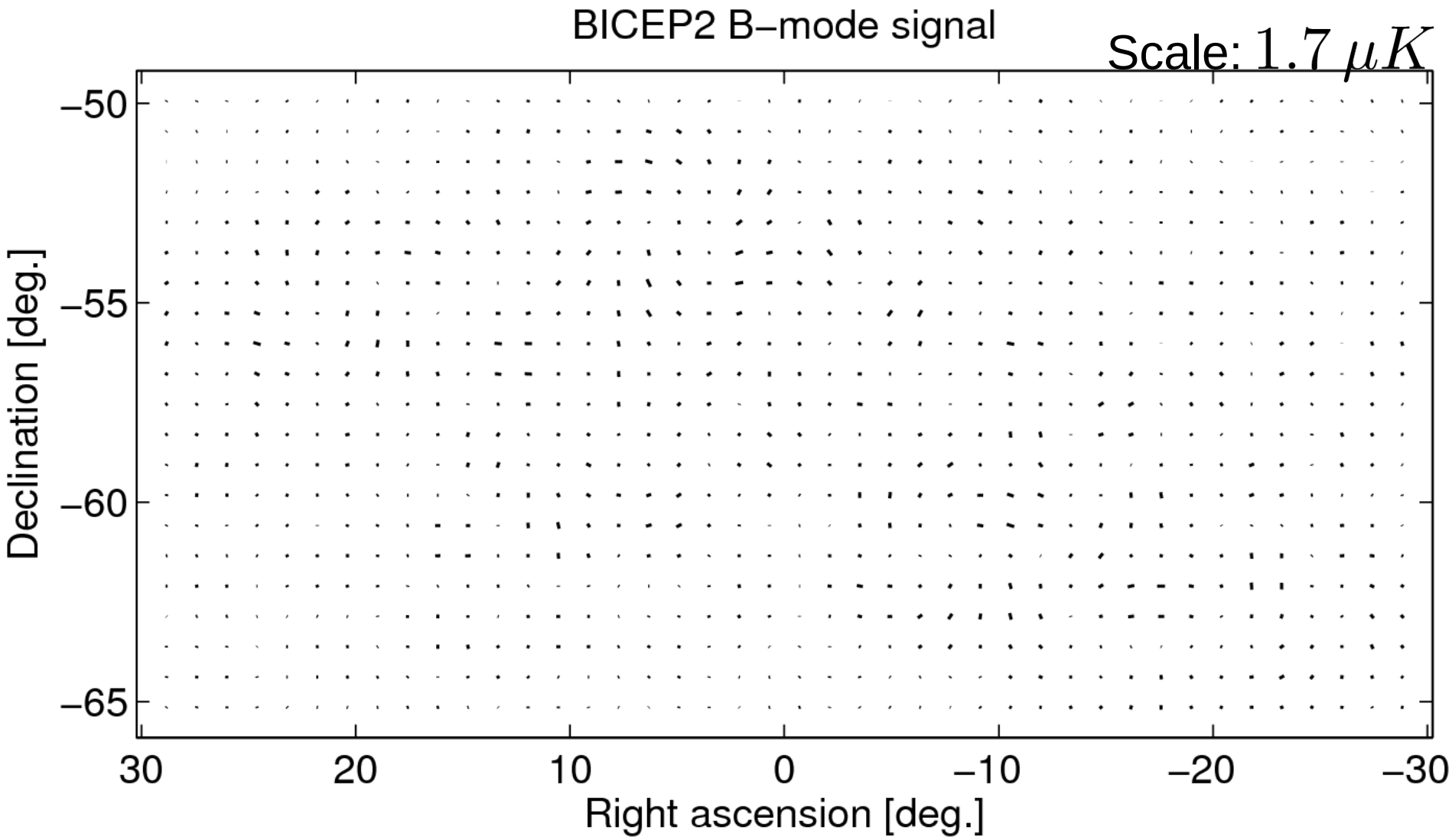

ALCOHOL
Weller
Weller



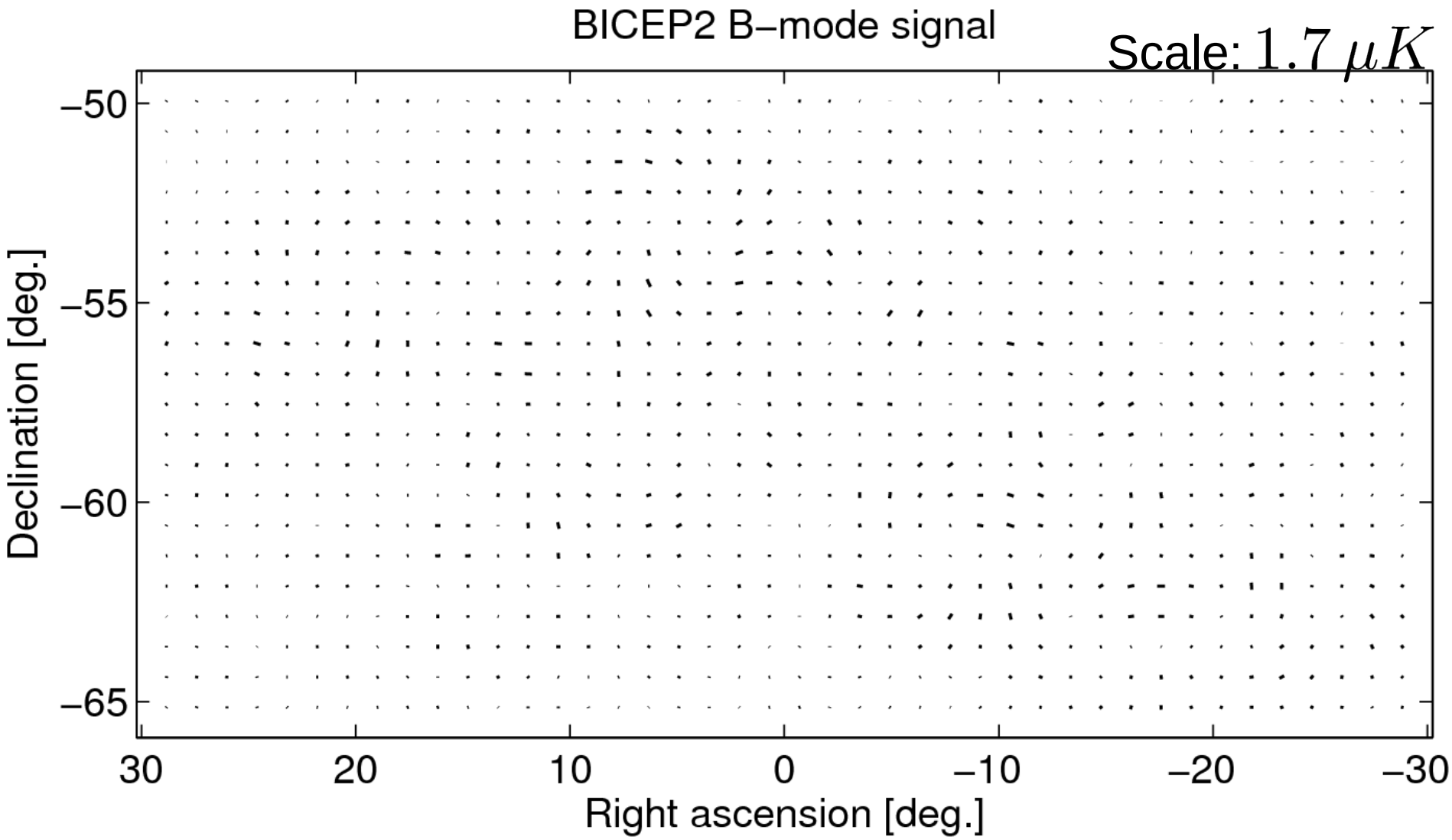
Total polarization



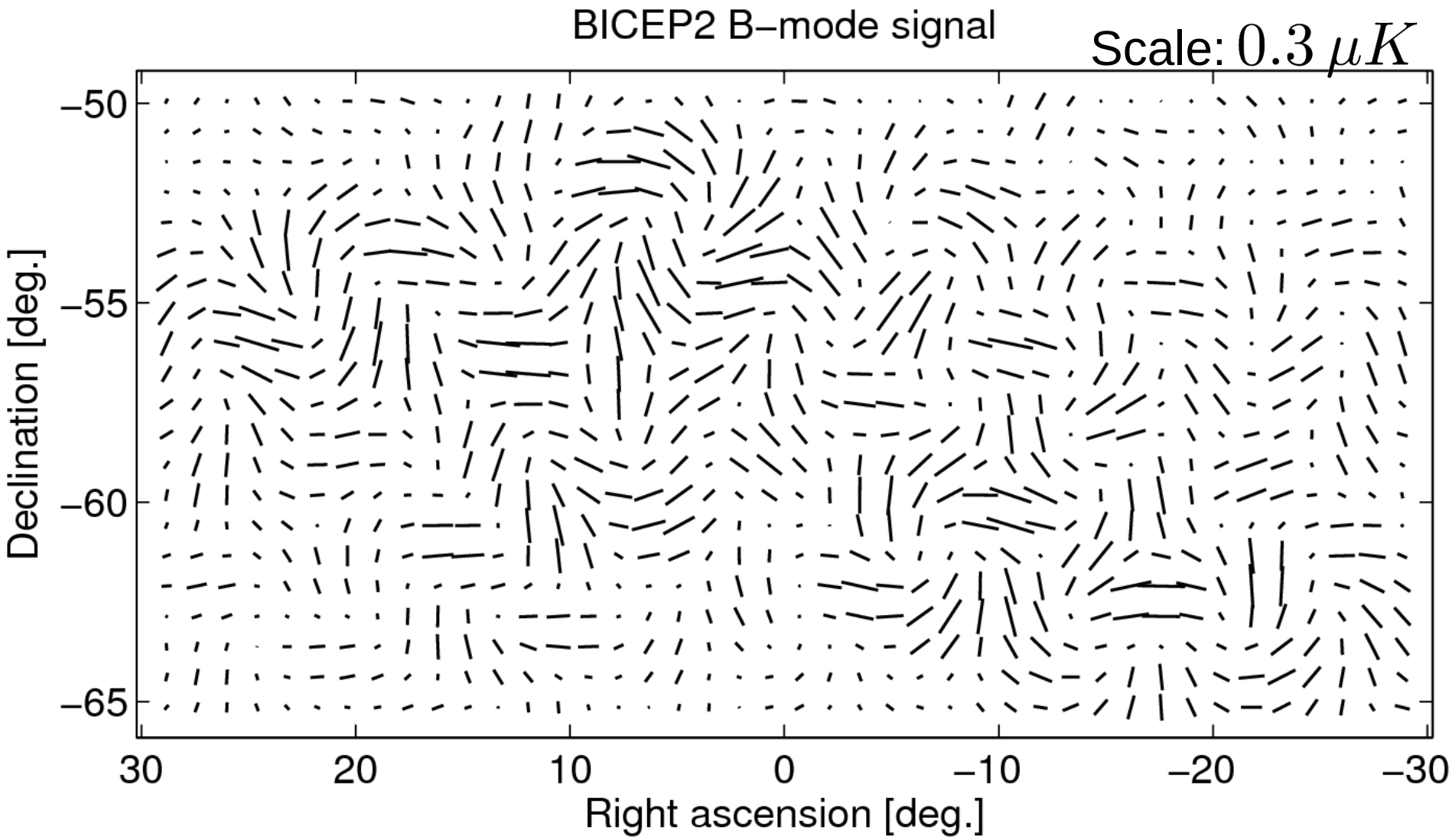
B-mode contribution



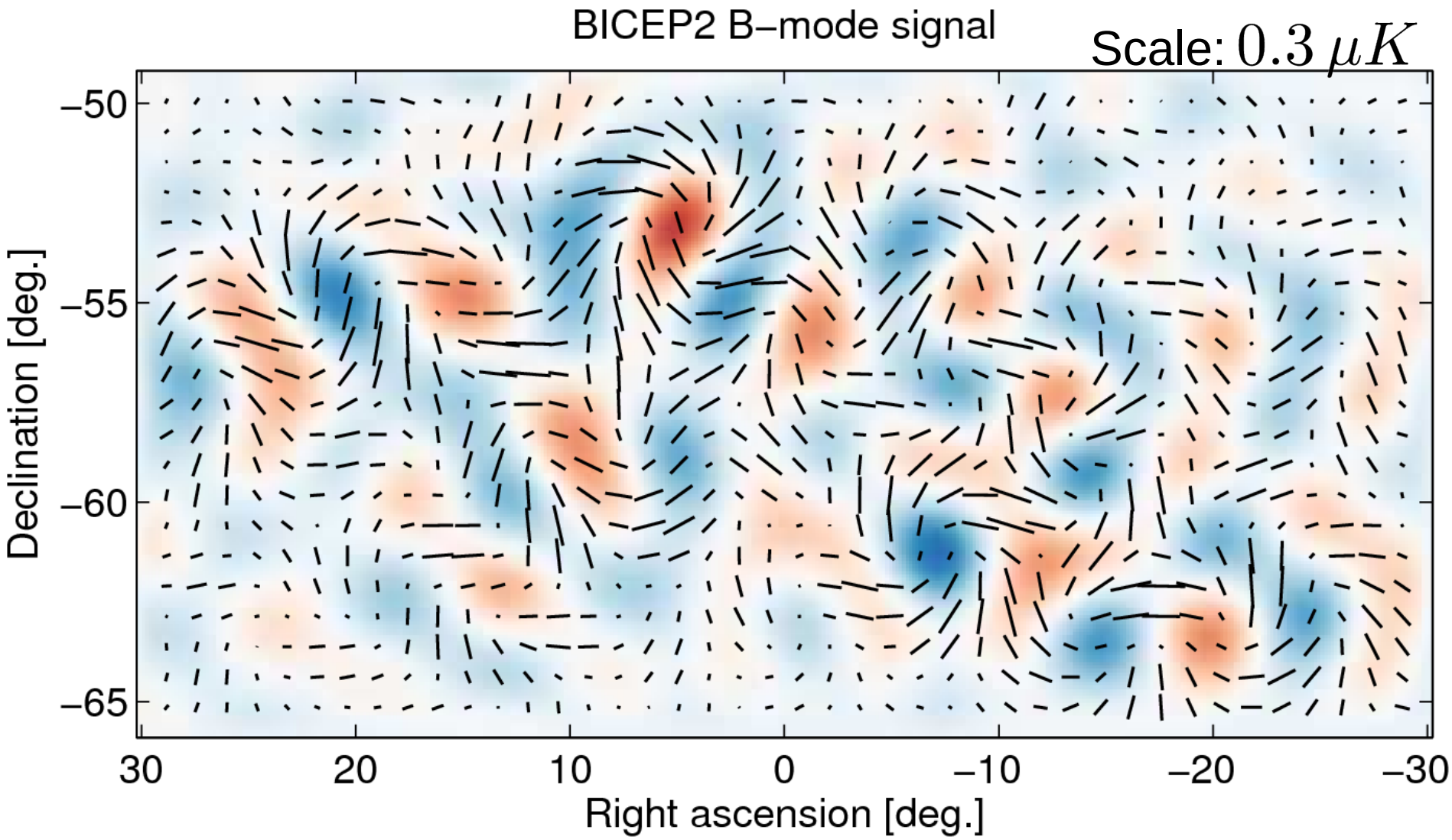
B-mode contribution



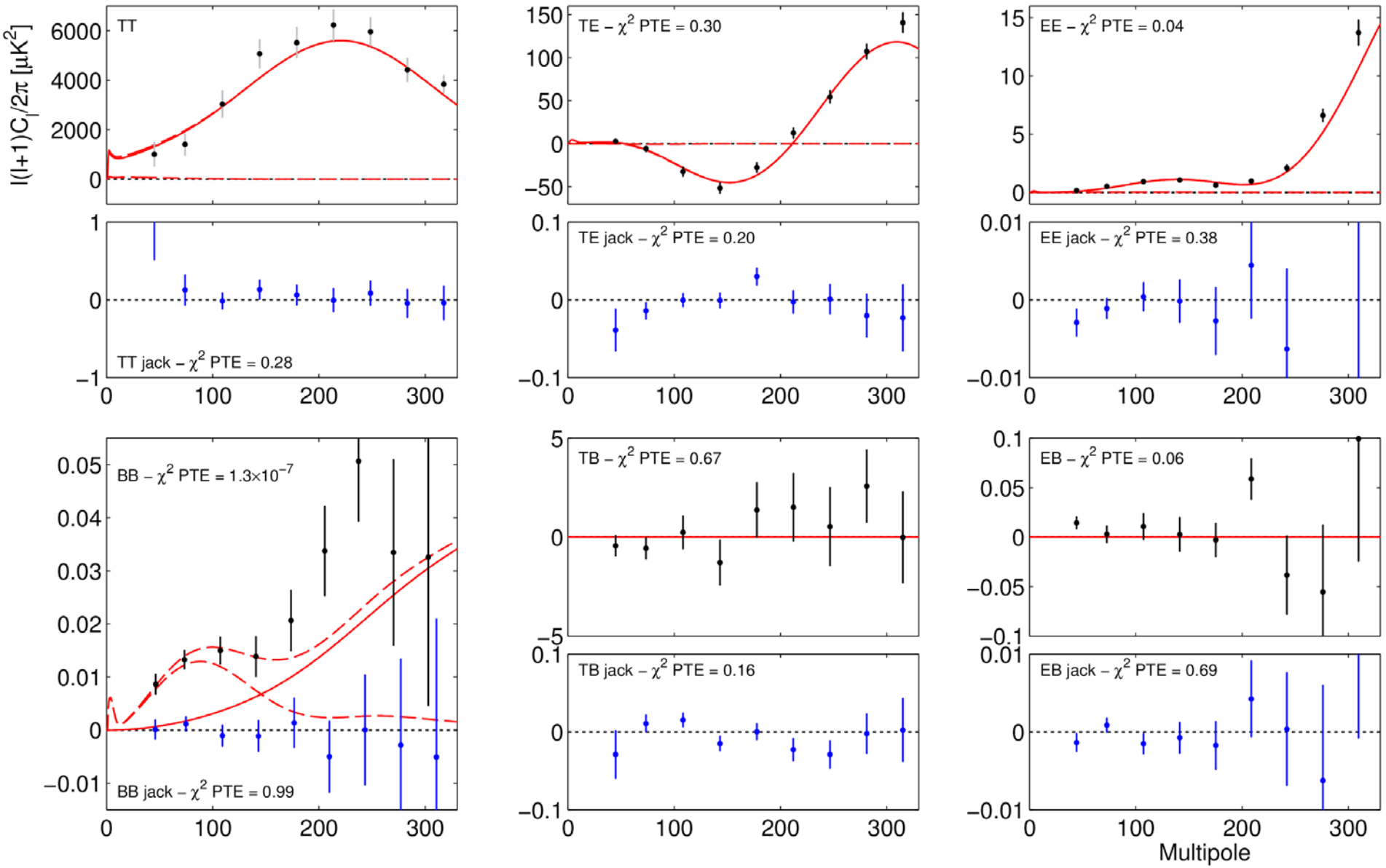
B-mode contribution



B-mode contribution

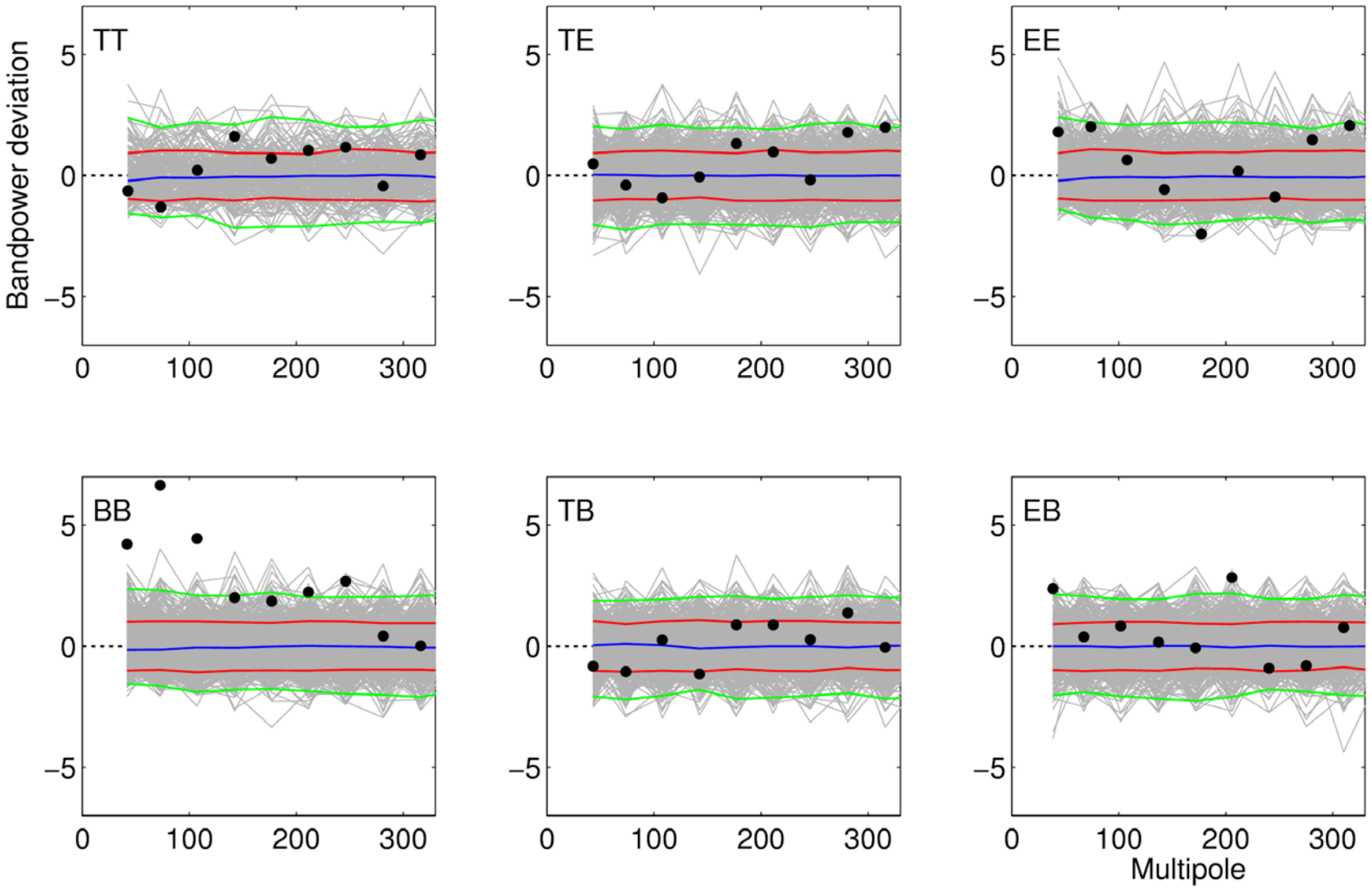


Temperature and Polarization Spectra



The Bicep2 Collaboration

Bandpower Deviations



Bandpower deviations from mean of lensed-ΛCDM+noise simulations and normalized by the std of those sims

- real data
- lensed-ΛCDM + noise sims
- ± 1σ
- ± 2σ

Check Systematics: Jackknifes

TABLE 1
JACKKNIFE PTE VALUES FROM χ^2 AND χ (SUM-OF-DEVIATION) TESTS

Jackknife	Bandpowers 1–5 χ^2	Bandpowers 1–9 χ^2	Bandpowers 1–5 χ	Bandpowers 1–9 χ
Deck jackknife				
EE	0.046	0.030	0.164	0.299
BB	0.774	0.329	0.240	0.082
EB	0.337	0.643	0.204	0.267
Scan Dir jackknife				
EE	0.483	0.762	0.978	0.938
BB	0.531	0.573	0.896	0.551
EB	0.898	0.806	0.725	0.890
Tag Split jackknife				
EE	0.541	0.377	0.916	0.938
BB	0.902	0.992	0.449	0.585
EB	0.477	0.689	0.856	0.615
Tile jackknife				
EE	0.004	0.010	0.000	0.002
BB	0.794	0.752	0.565	0.331
EB	0.172	0.419	0.962	0.790
Phase jackknife				
EE	0.673	0.409	0.126	0.339
BB	0.591	0.739	0.842	0.944
EB	0.529	0.577	0.840	0.659
Mux Col jackknife				
EE	0.812	0.587	0.196	0.204
BB	0.826	0.972	0.293	0.283
EB	0.866	0.968	0.876	0.697
Alt Deck jackknife				
EE	0.004	0.004	0.070	0.236
BB	0.397	0.176	0.381	0.086
EB	0.150	0.060	0.170	0.291
Mux Row jackknife				
EE	0.052	0.178	0.653	0.739
BB	0.345	0.361	0.032	0.008
EB	0.529	0.226	0.024	0.048
Tile/Deck jackknife				
EE	0.048	0.088	0.144	0.132
BB	0.908	0.840	0.629	0.269
EB	0.050	0.154	0.591	0.591
Focal Plane inner/outer jackknife				
EE	0.230	0.597	0.022	0.090
BB	0.216	0.531	0.046	0.092
EB	0.036	0.042	0.850	0.838
Tile top/bottom jackknife				
EE	0.289	0.347	0.459	0.599
BB	0.293	0.236	0.154	0.028
EB	0.545	0.683	0.902	0.932
Tile inner/outer jackknife				
EE	0.727	0.533	0.128	0.485
BB	0.255	0.086	0.421	0.036
EB	0.465	0.737	0.208	0.168
Moon jackknife				
EE	0.499	0.689	0.481	0.679
BB	0.144	0.287	0.898	0.858
EB	0.289	0.359	0.531	0.307
A/B offset best/worst				
EE	0.317	0.311	0.868	0.709
BB	0.114	0.064	0.307	0.094
EB	0.589	0.872	0.599	0.790

Splits the 4 boresight rotations

Amplifies differential pointing in comparison to fully added data. Important check of deprojection. See later slides.

Splits by time

Checks for contamination on long (“Tag Split”) and short (“Scan Dir”) timescales. Short timescales probe detector transfer functions.

Splits by channel selection

Checks for contamination in channel subgroups, divided by focal plane location, tile location, and readout electronics grouping

Splits by possible external contamination

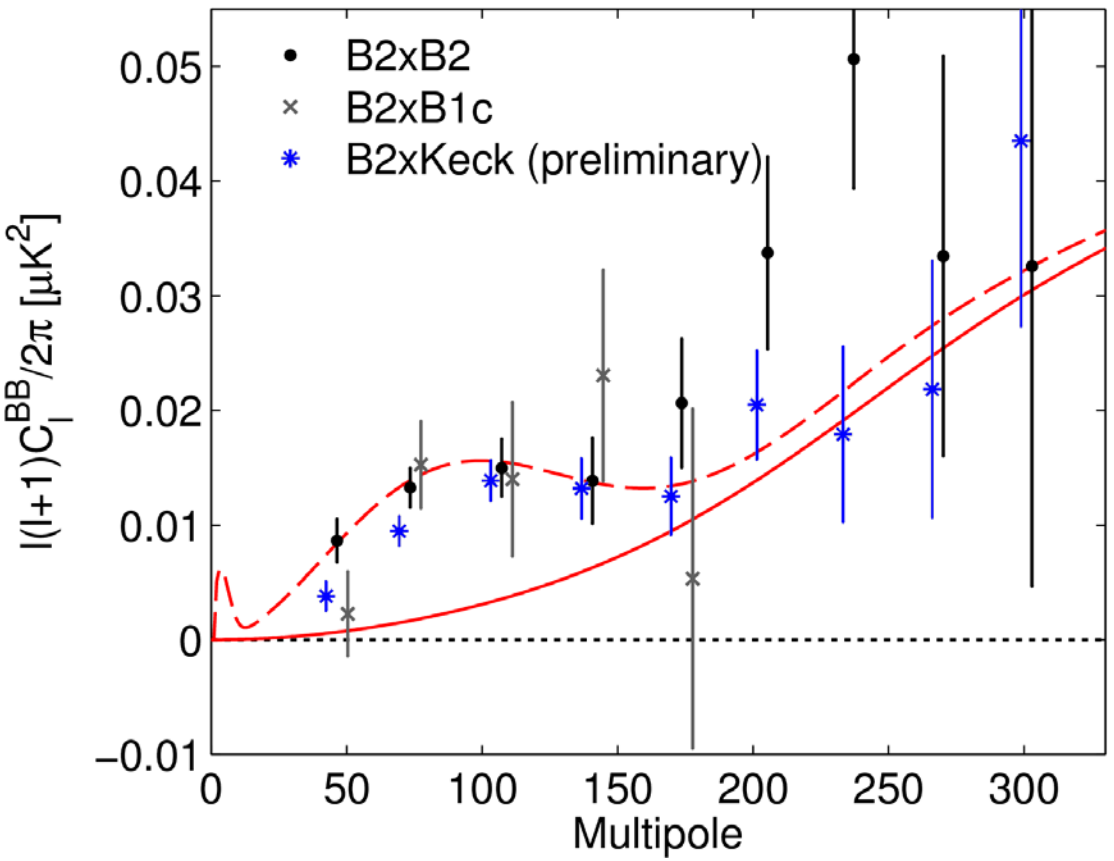
Checks for contamination from ground-fixed signals, such as polarized sky or magnetic fields, or the moon

Splits to check intrinsic detector properties

Checks for contamination from detectors with best/worst differential pointing. “Tile/dk” divides the data by the orientation of the detector on the sky.

Additional Cross Spectra

Form cross spectrum between BICEP2 and BICEP1 combined (100 + 150 GHz):

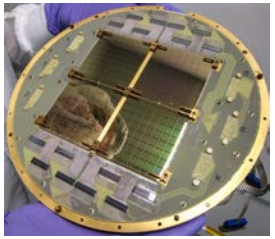


BICEP2 auto spectrum compatible with B2xB1c cross spectrum

~3σ evidence of excess power in the cross spectrum

Additionally form cross spectrum with 2 years of data from *Keck Array*, the successor to BICEP2

Excess power is also evident in the B2xKeck cross spectrum



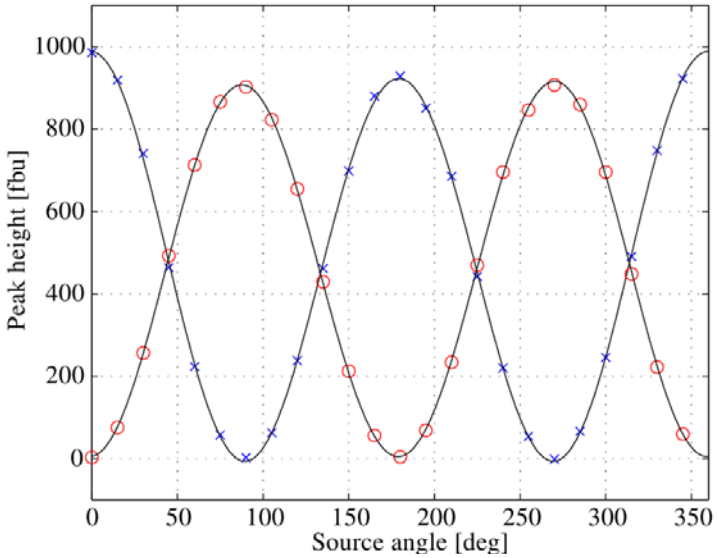
Cross spectra:
Powerful additional evidence against a systematic origin of the apparent signal

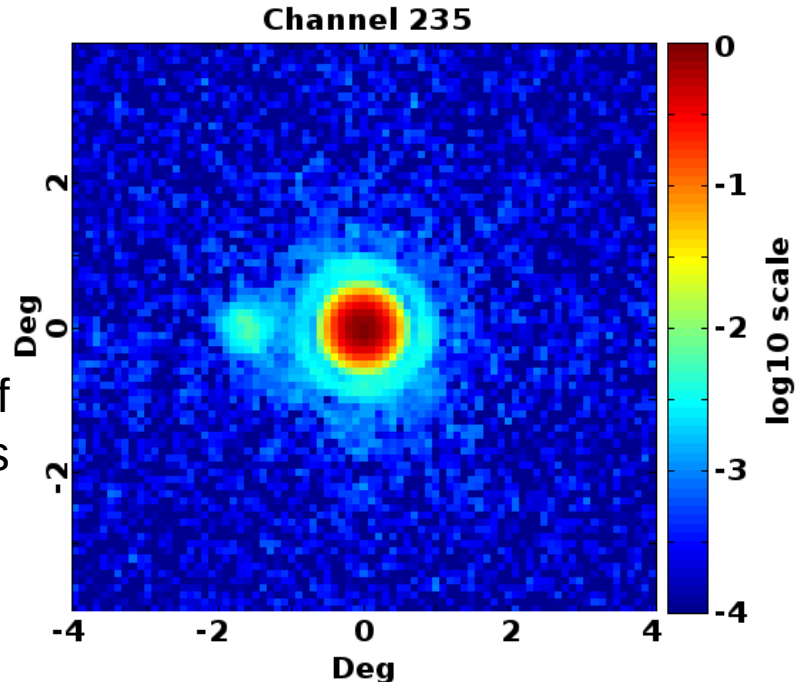
Calibration Measurements

For instance...

Far field beam mapping

Detector Polarization Calibration

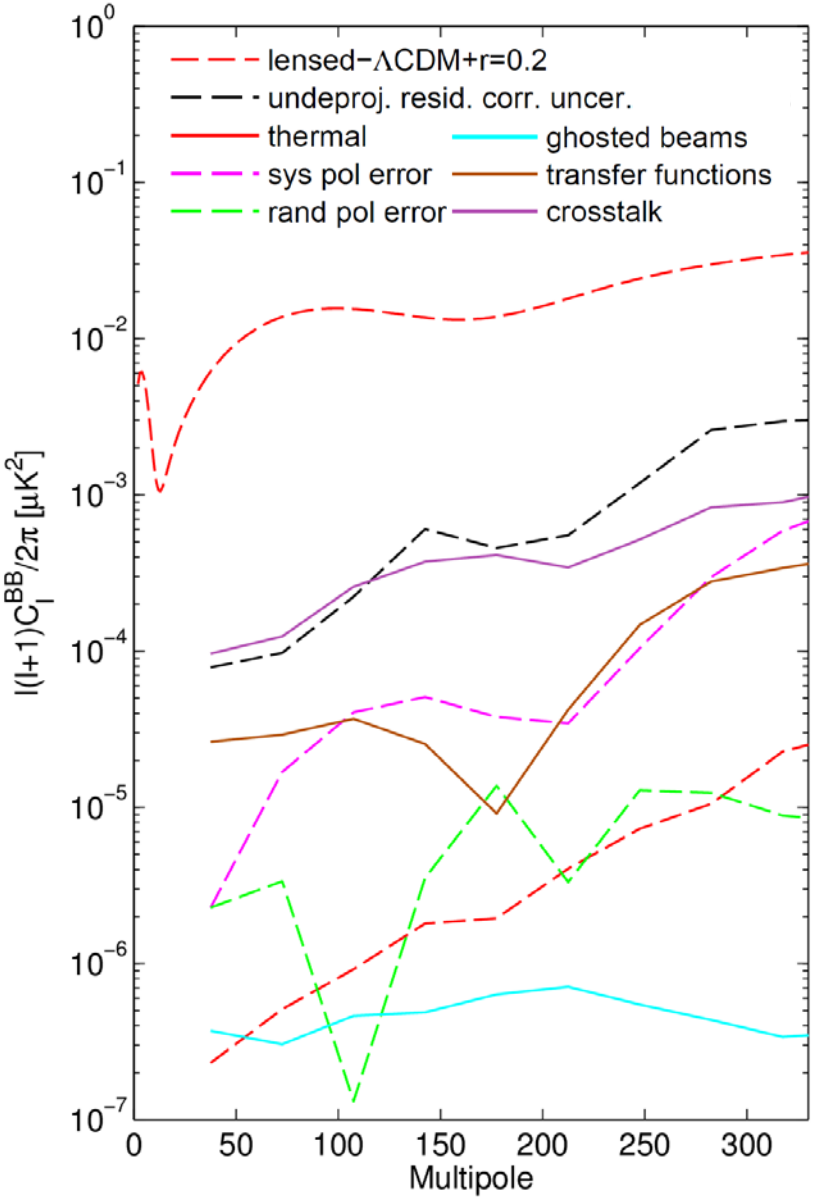




Hi-Fi beam maps of individual detectors

Detailed description in companion Instrument Paper

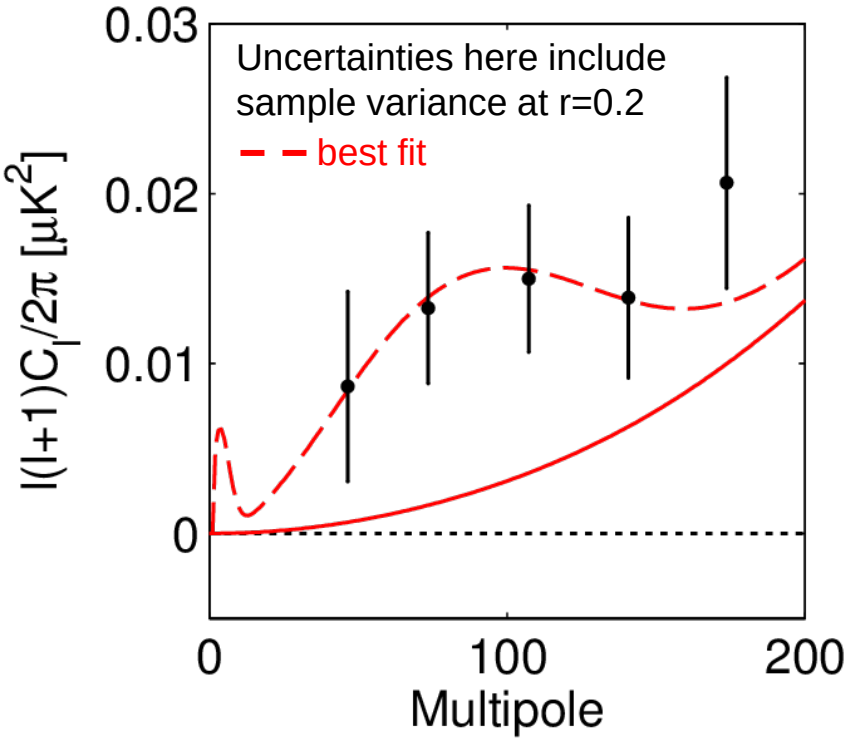
Systematics beyond Beam imperfections



All systematic effects that we could imagine were investigated!

We find with high confidence that the apparent signal *cannot be explained* by instrumental systematics!

Constraint on Tensor-to-scalar Ratio r



Substantial excess power in the region where the inflationary gravitational wave signal is expected to peak

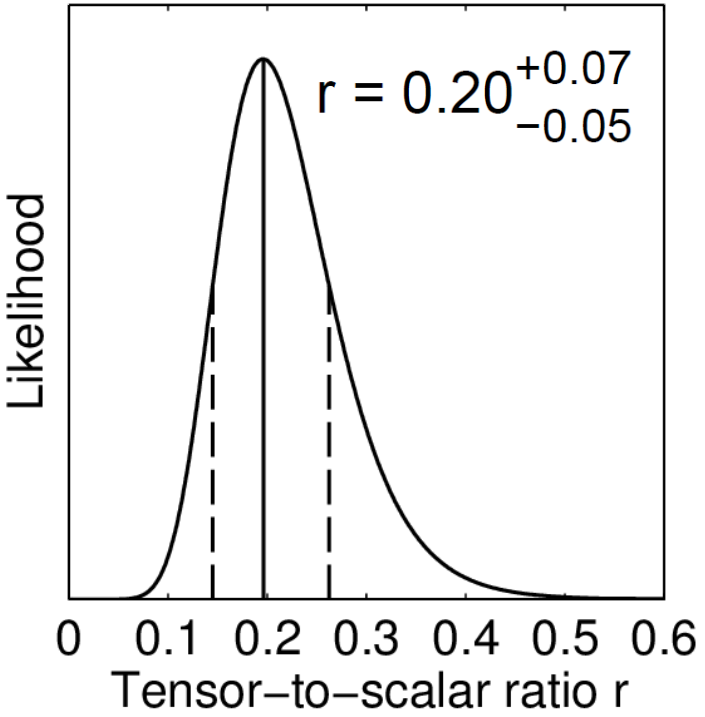
Find the most likely value of the tensor-to-scalar ratio r

Apply "direct likelihood" method, uses:

→ lensed-ΛCDM + noise simulations

→ weighted version of the 5 bandpowers

→ B-mode sims scaled to various levels of r (n_T=0)



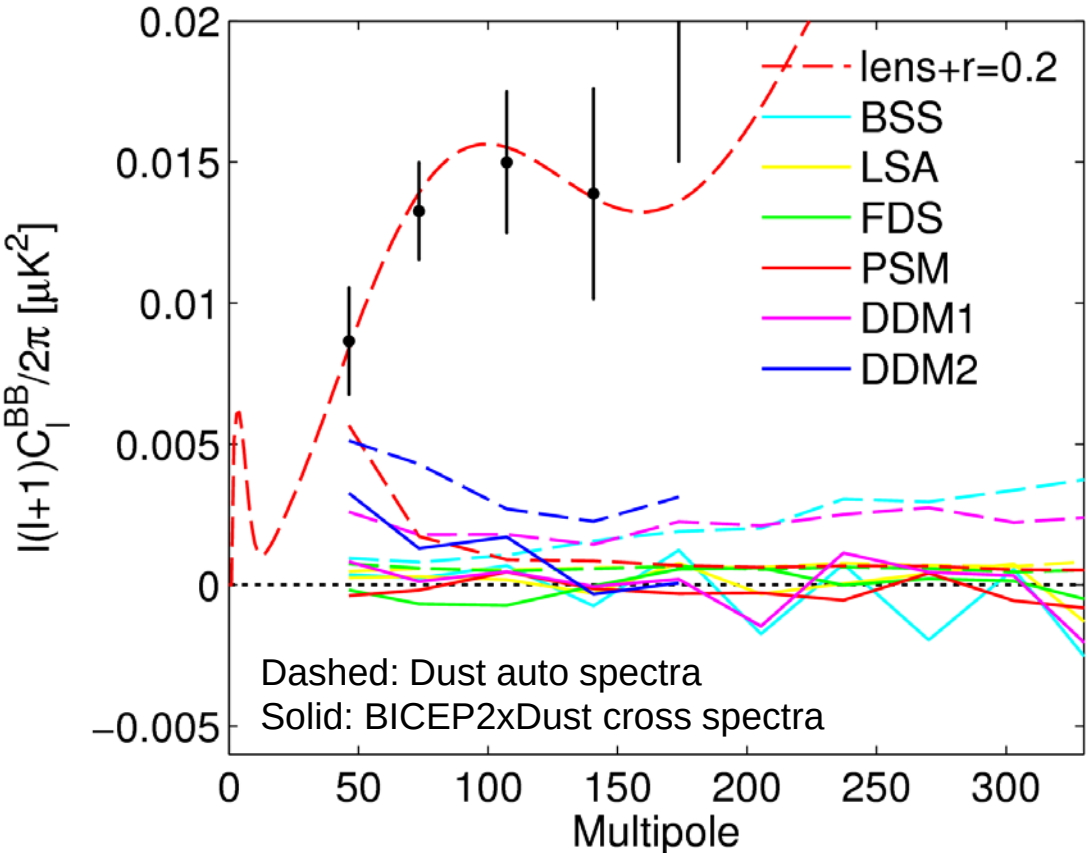
Within this simplistic model we find:

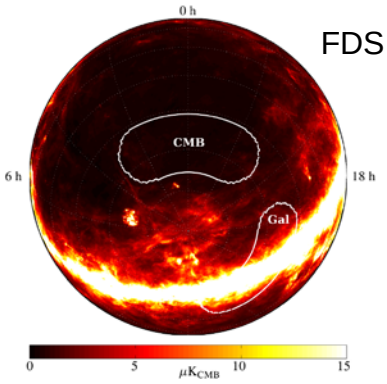
r = 0.2 with uncertainties dominated by sample variance

PTE of fit to data: 0.9
→ model is perfectly acceptable fit to the data

r=0 ruled out at 7.0σ

Polarized Dust Foreground Projections





FDS Model

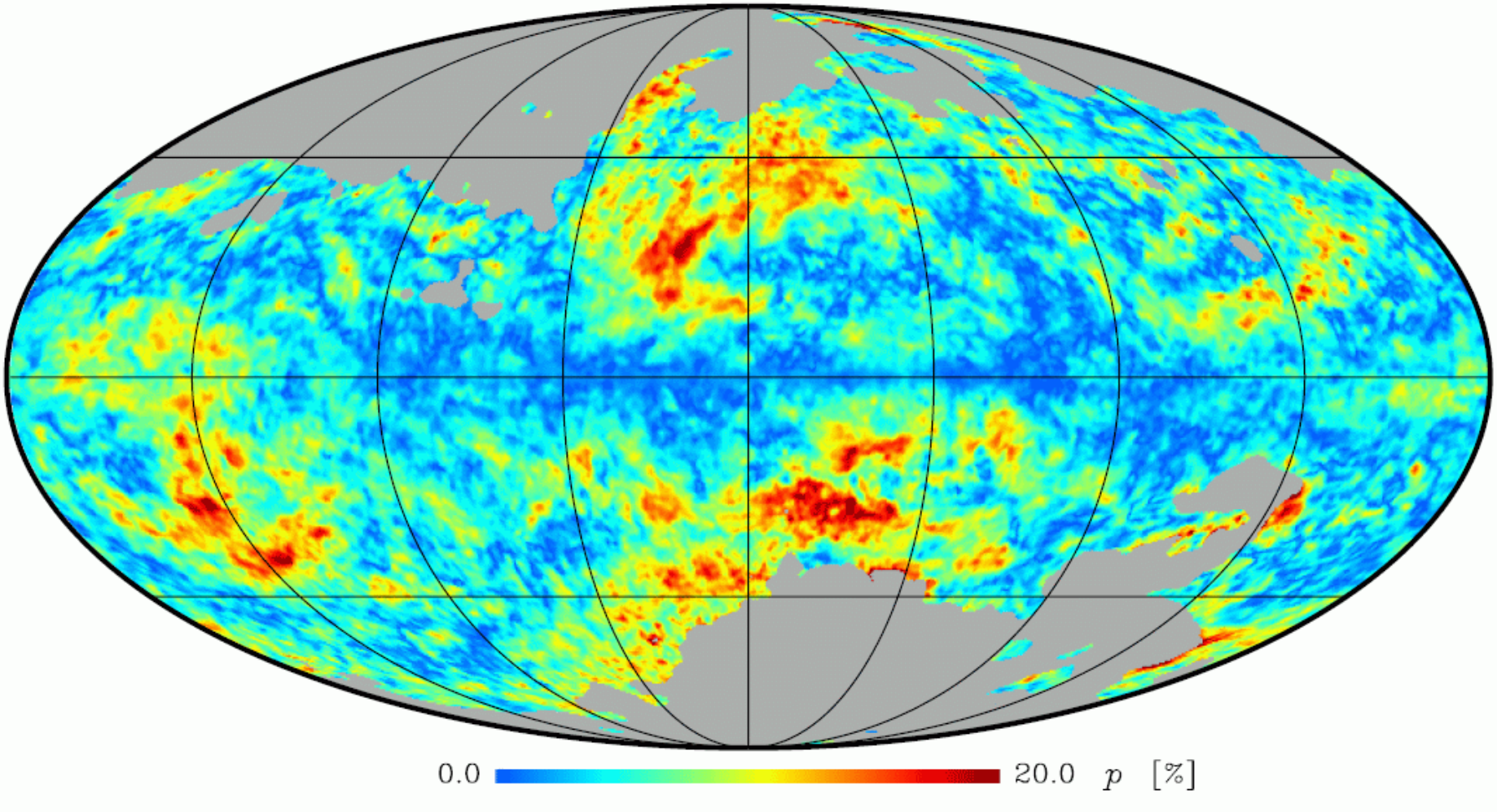
The BICEP2 region is chosen to have extremely low foreground emission.

Use various models of polarized dust emission to estimate foregrounds.

All dust auto spectra well below observed signal level.

Cross spectra consistent with zero.

Polarization fraction from *Planck XIX* (May 6)



The Bicep2 Collaboration

Polarization fraction from *Planck XIX* (May 6)

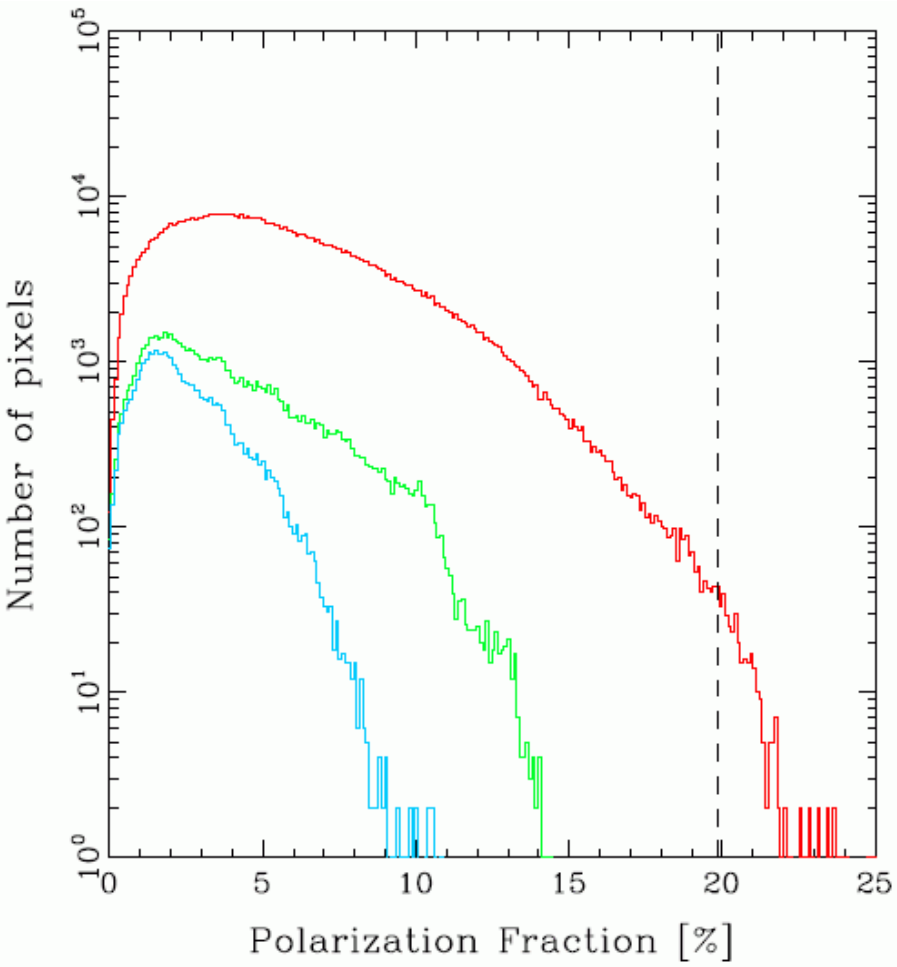
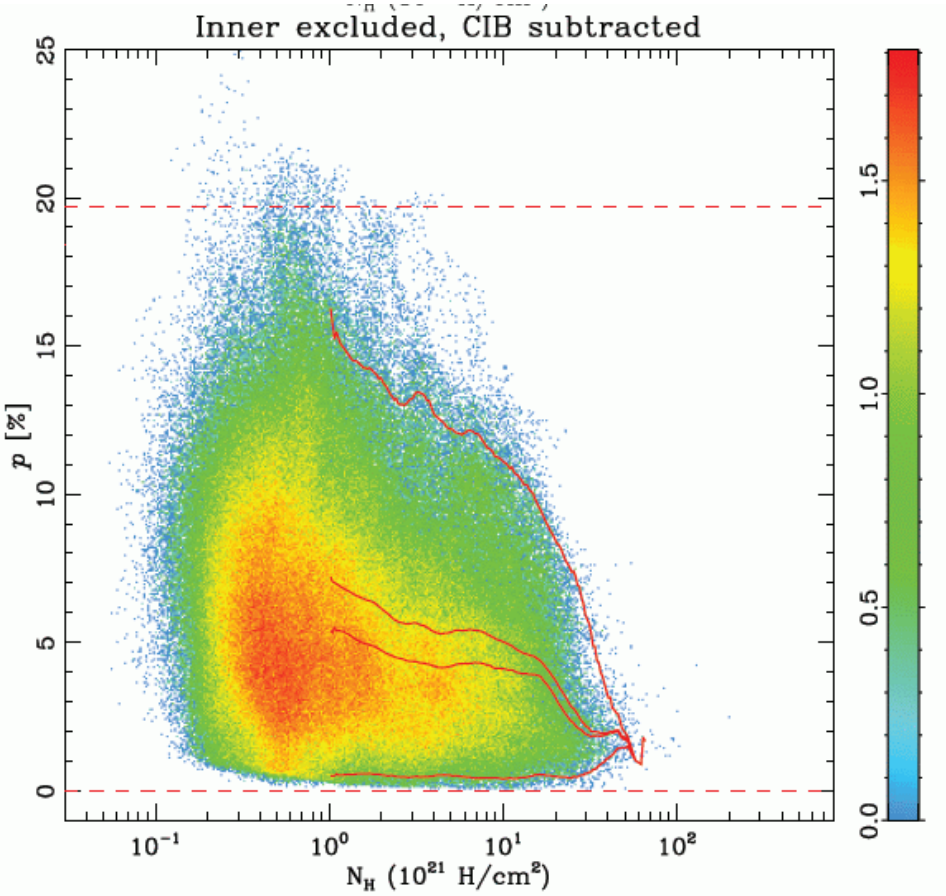


Fig. 7. Histograms of the observed polarization fraction at 1° resolution for the whole sky shown in Fig. 1 (*red*), the Galactic plane within $|b_{II}| < 5°$ (*green*) and the inner Galactic plane within $|b_{II}| < 5°$ and $|\ell_{II}| < 90°$ (*blue*).

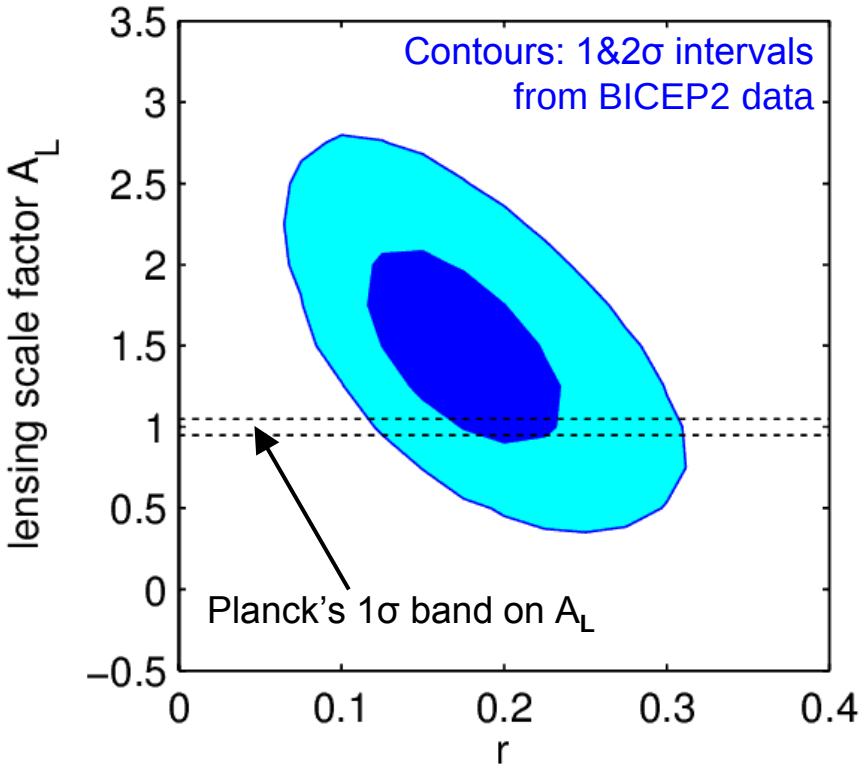


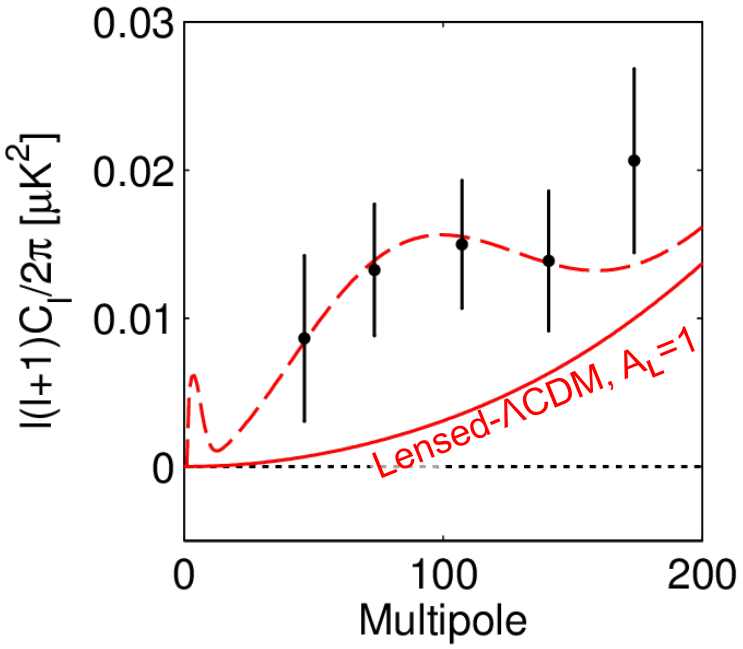
The Bicep2 Collaboration

Joint Constraint on r and Lensing Scale Factor

Lensing deflects CMB photons, slightly mixing the dominant E-modes into B-modes -- dominant at high multipoles

Planck data constrain the amplitude of the lensing effect to $A_L = 0.99 \pm 0.05$.





In the joint constraint on r and A_L we find:

BICEP2 data is perfectly compatible with a lensing amplitude of A = 1.

Marginalizing over r, we detect lensing B-modes at 2.7σ

Compatibility with Indirect Limits on r

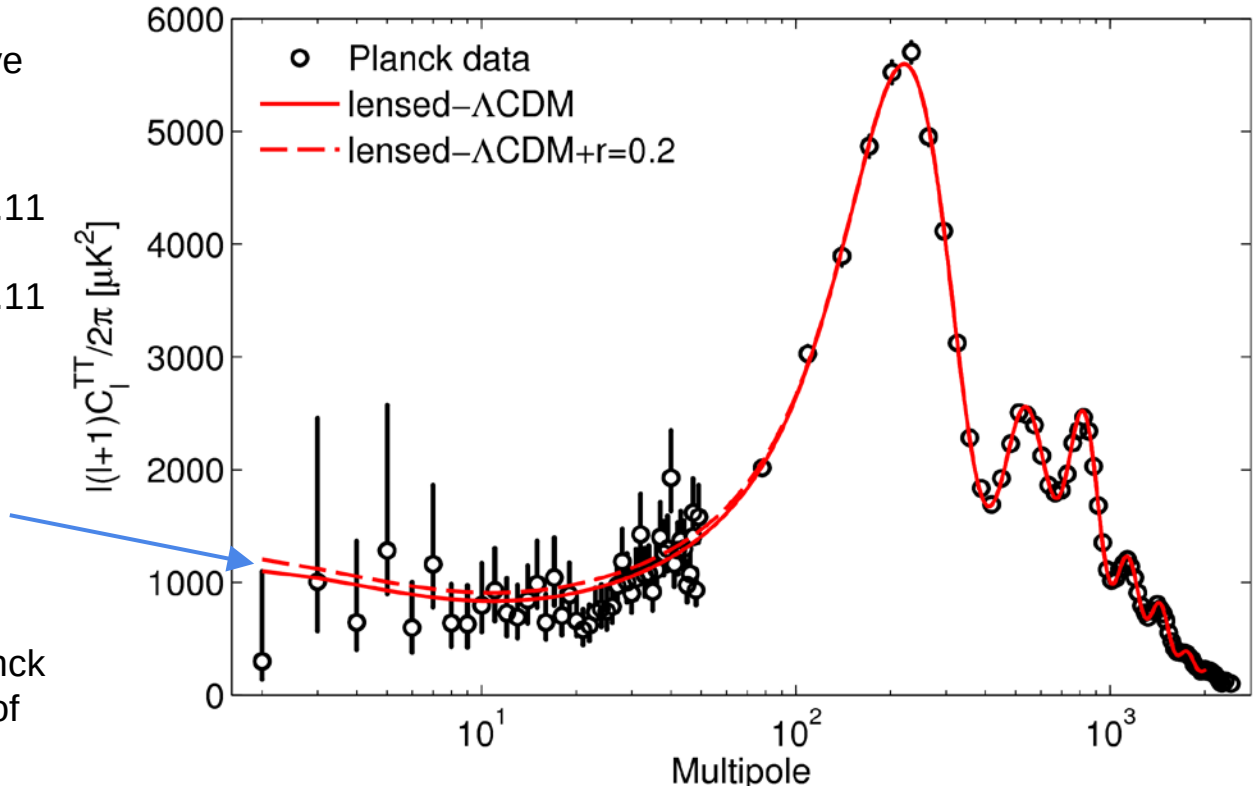
Using temperature data over a wide range of angular scales limits on r have been set:

SPT+WMAP+BAO+H_0 : $r < 0.11$

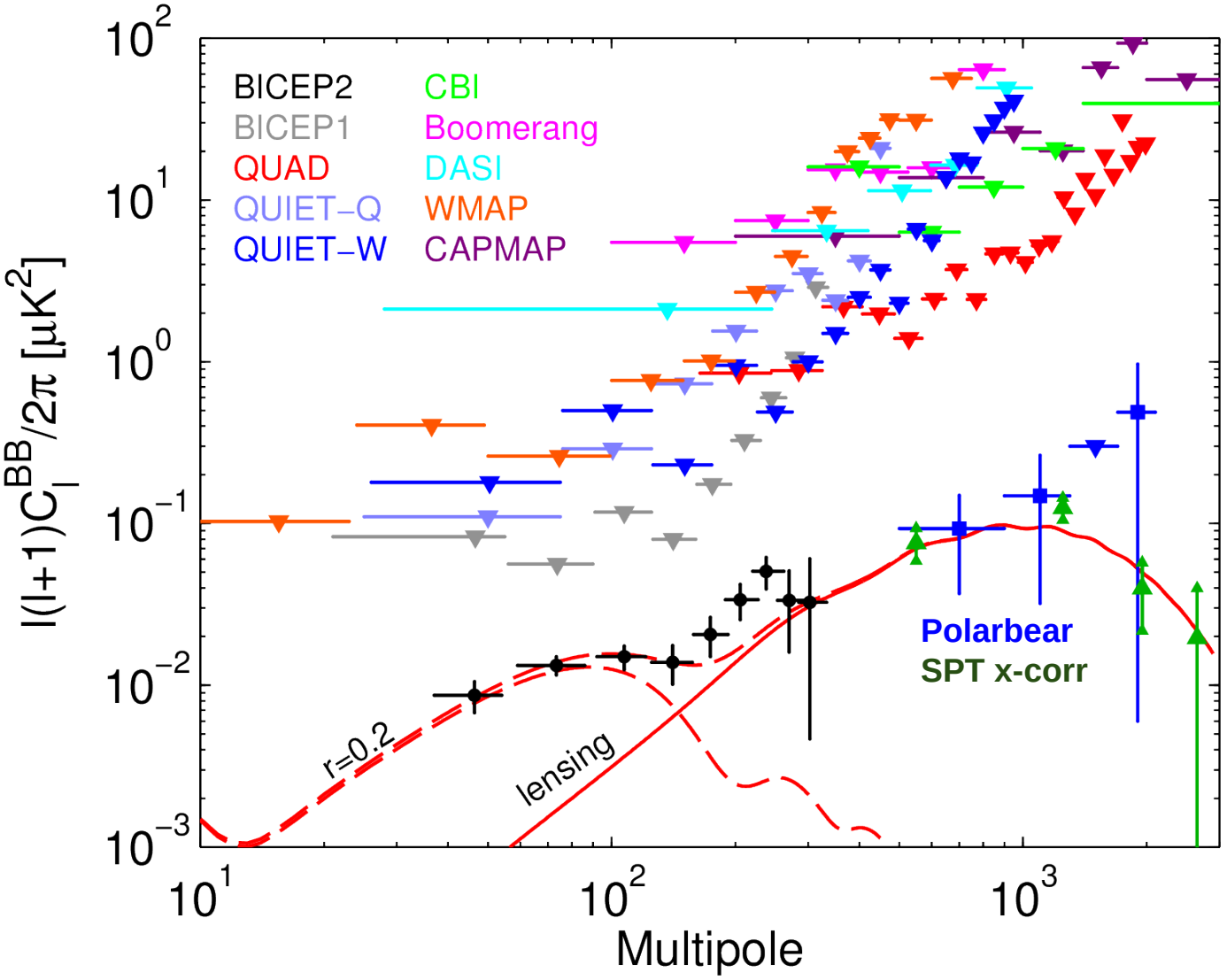
Planck+SPT+ACT+$WMAP_{pol}$: $r < 0.11$

r=0.2 makes a small change to the temperature spectrum.

(In this plot r=0.2 simply added to Planck best fit model with no re-optimization of other parameters)



BICEP2 and upper limits from other experiments:



Future

- We really want to confirm the signal (Keck-100, BICEP3, SPTpol, Planck, …)
- Increase the frequency and sky coverage
- Test the scale invariance of B-modes (n_T)
  - All sky survey
  - delensing with large aperture telescopes
- Test the Gaussianity of B-modes
- Search for the (tensor) reionization bump
- Can we fit tensor into TT spectrum ?

Prospects

BICEP1: 2006, 2007, 2008

BICEP2: 2010, 2011, 2012

Keck Array: 2011, 2012, 2013,

2014 (576 100GHz detectors)…

BICEP3: 2015…

Prospects

BICEP1: 2006, 2007, 2008

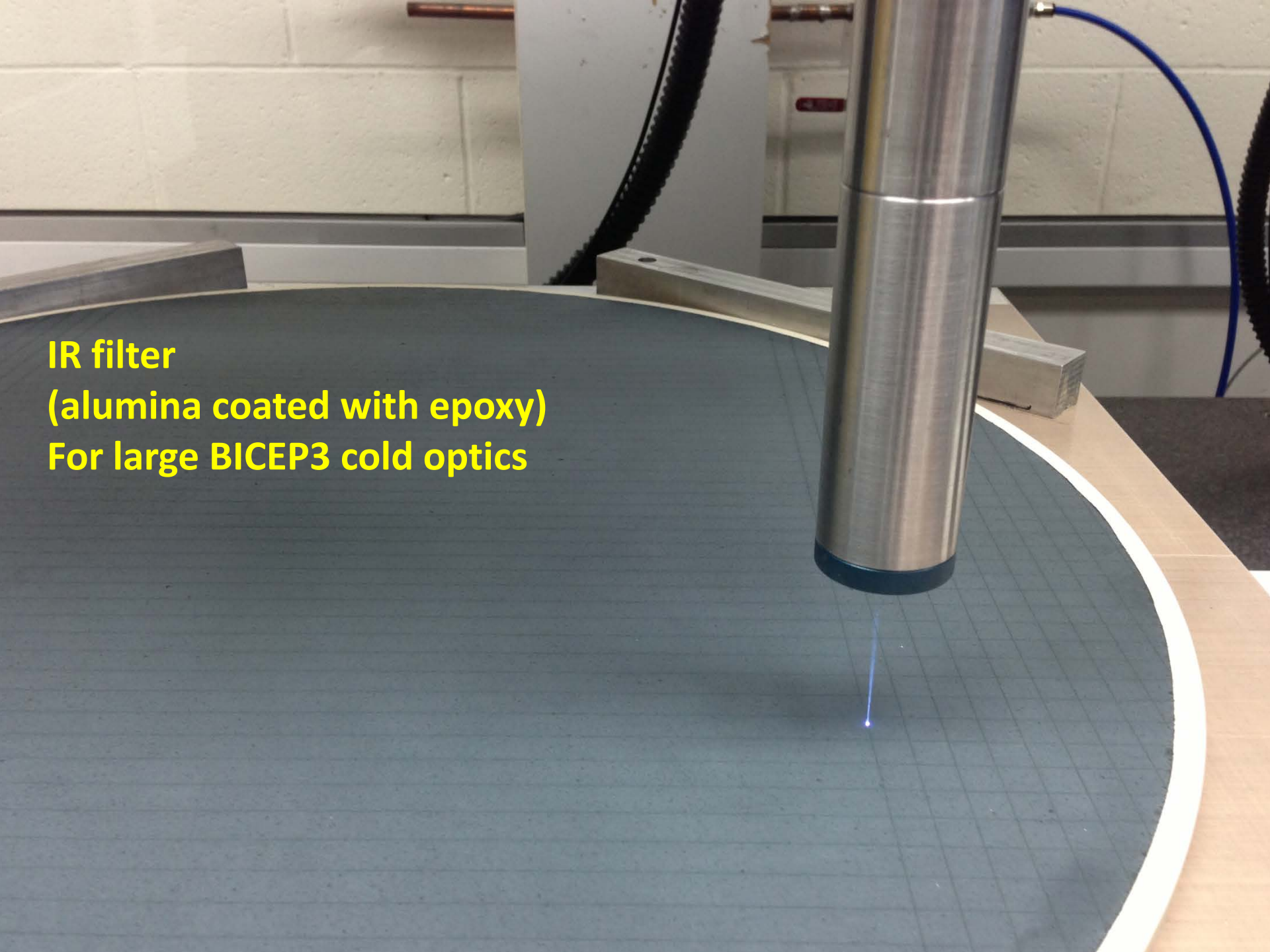
BICEP2: 2010, 2011, 2012

Keck Array: 2011, 2012, 2013,

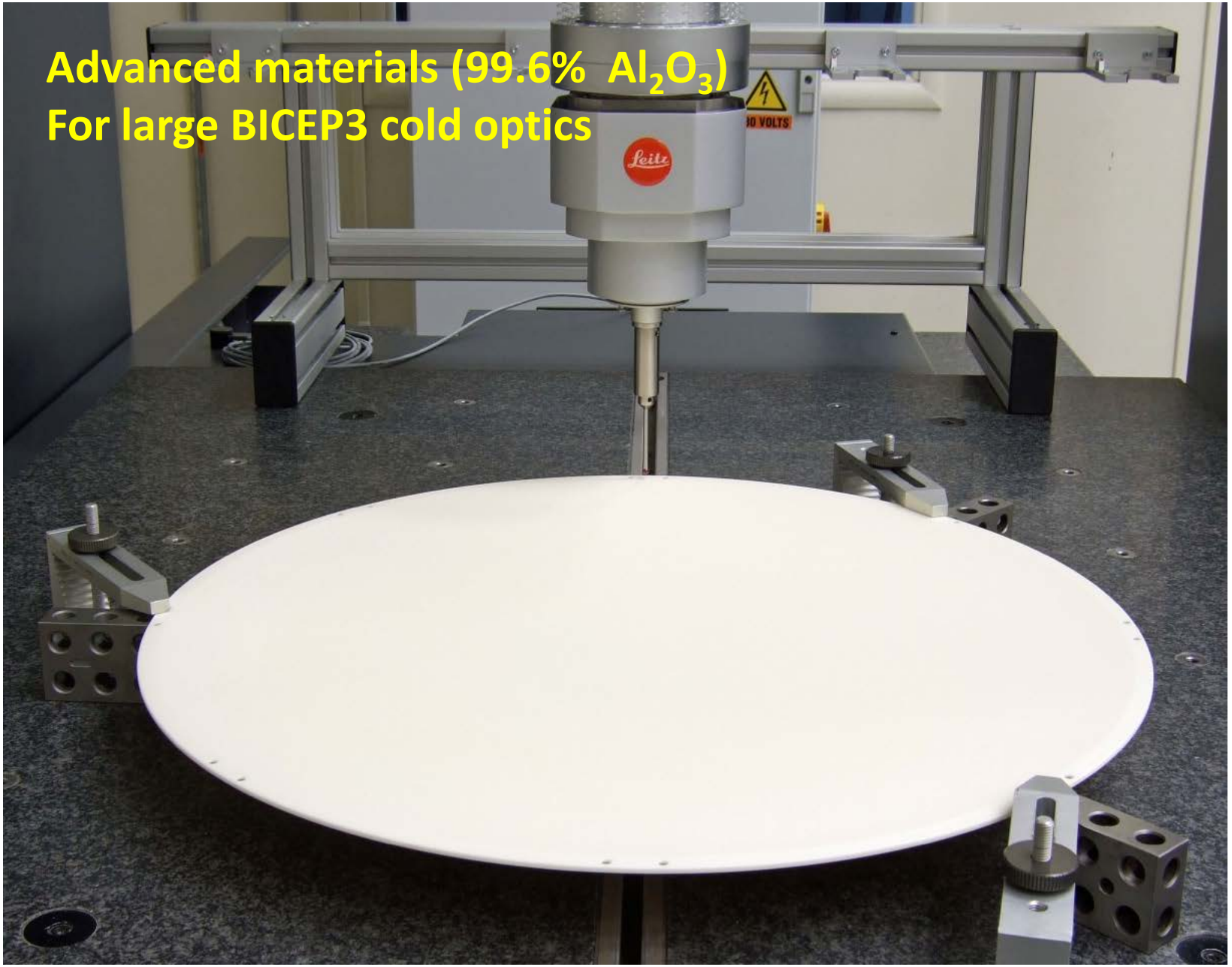
2014 (576 100GHz detectors)…

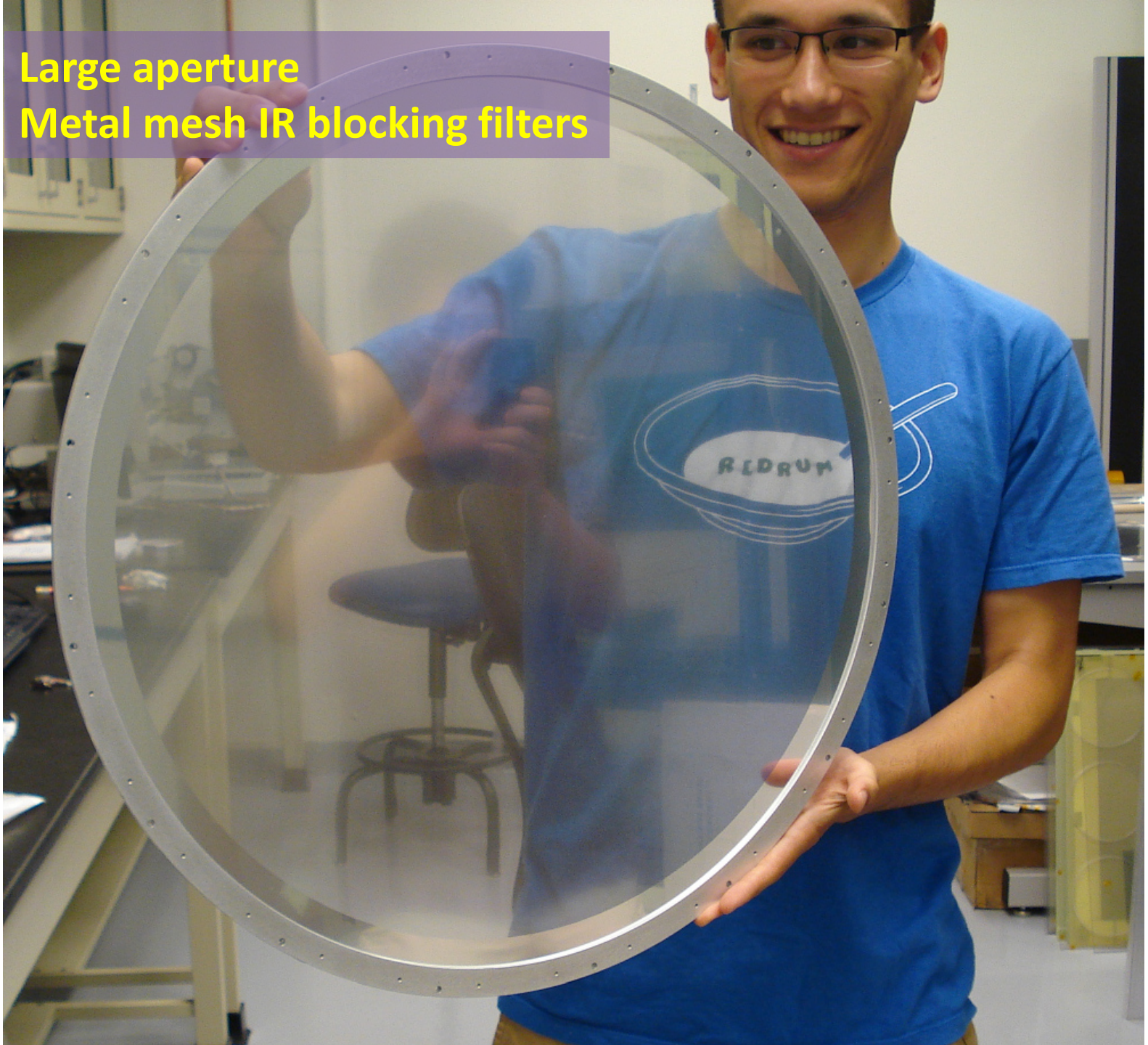
BICEP3: 2015 –

(another 2560 100GHz detectors)


IR filter
(alumina coated with epoxy)
For large BICEP3 cold optics

Epoxy-based
AR-coating
On curved lens


Advanced materials (99.6% Al_2O_3)
For large BICEP3 cold optics

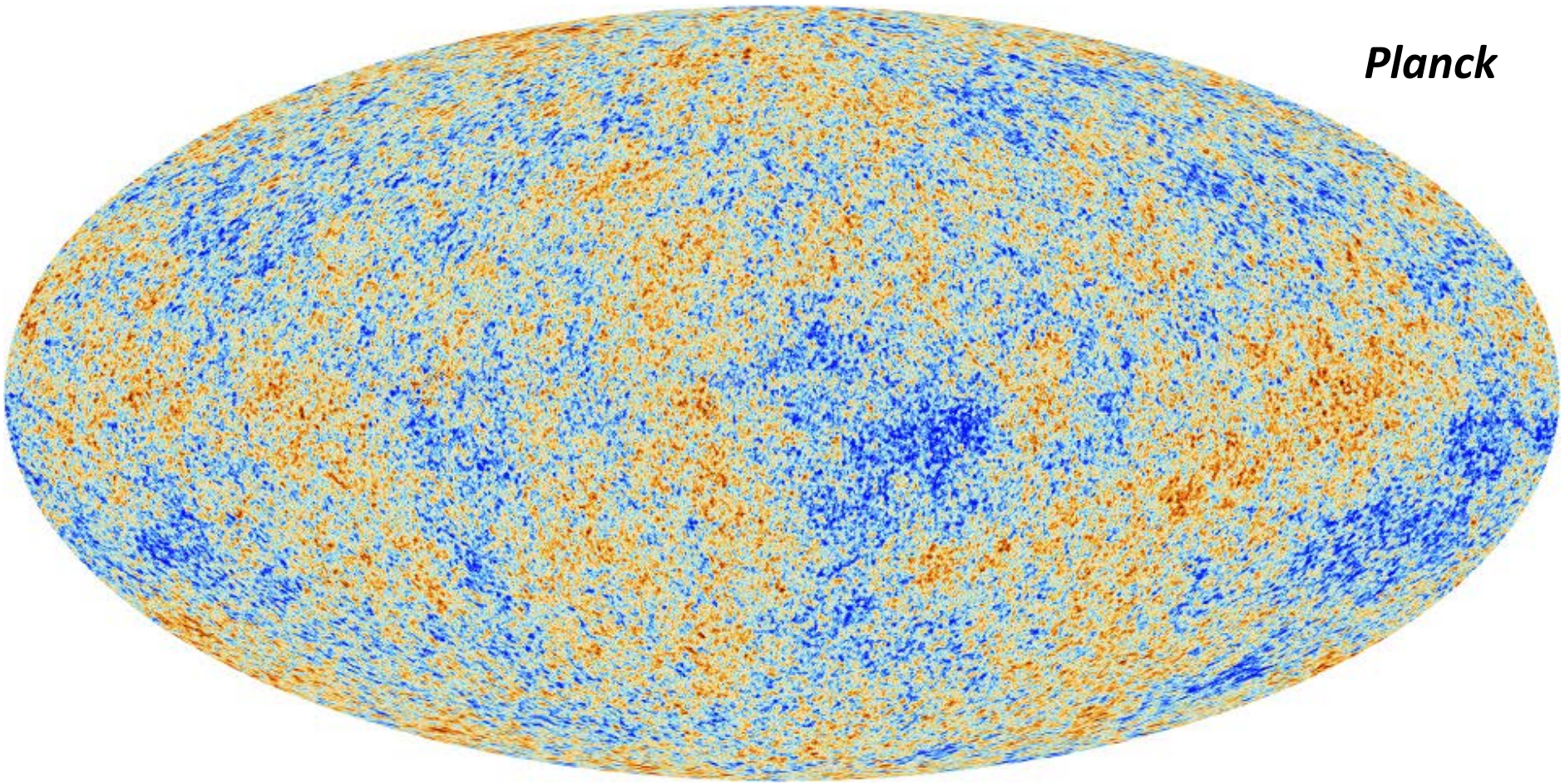

Large aperture
Metal mesh IR blocking filters

CMB after

Planck, ACT, SPT, POLARBEAR, BICEP3?

Three fundamental *maps* of our Universe

Map # 1
(scalar perturbations)



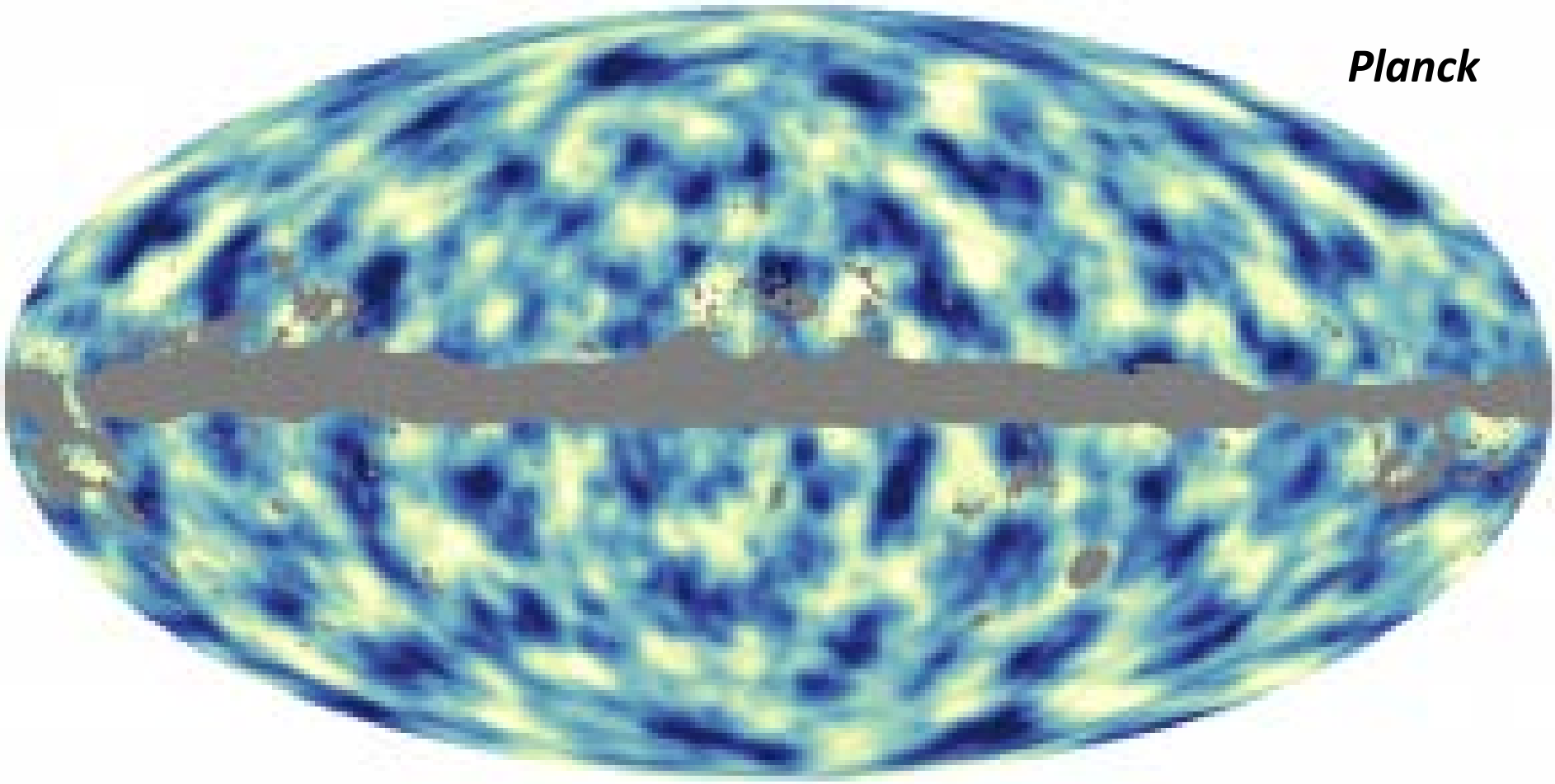
Planck

Three fundamental *maps* of our Universe

Map # 2
(tensor perturbations)

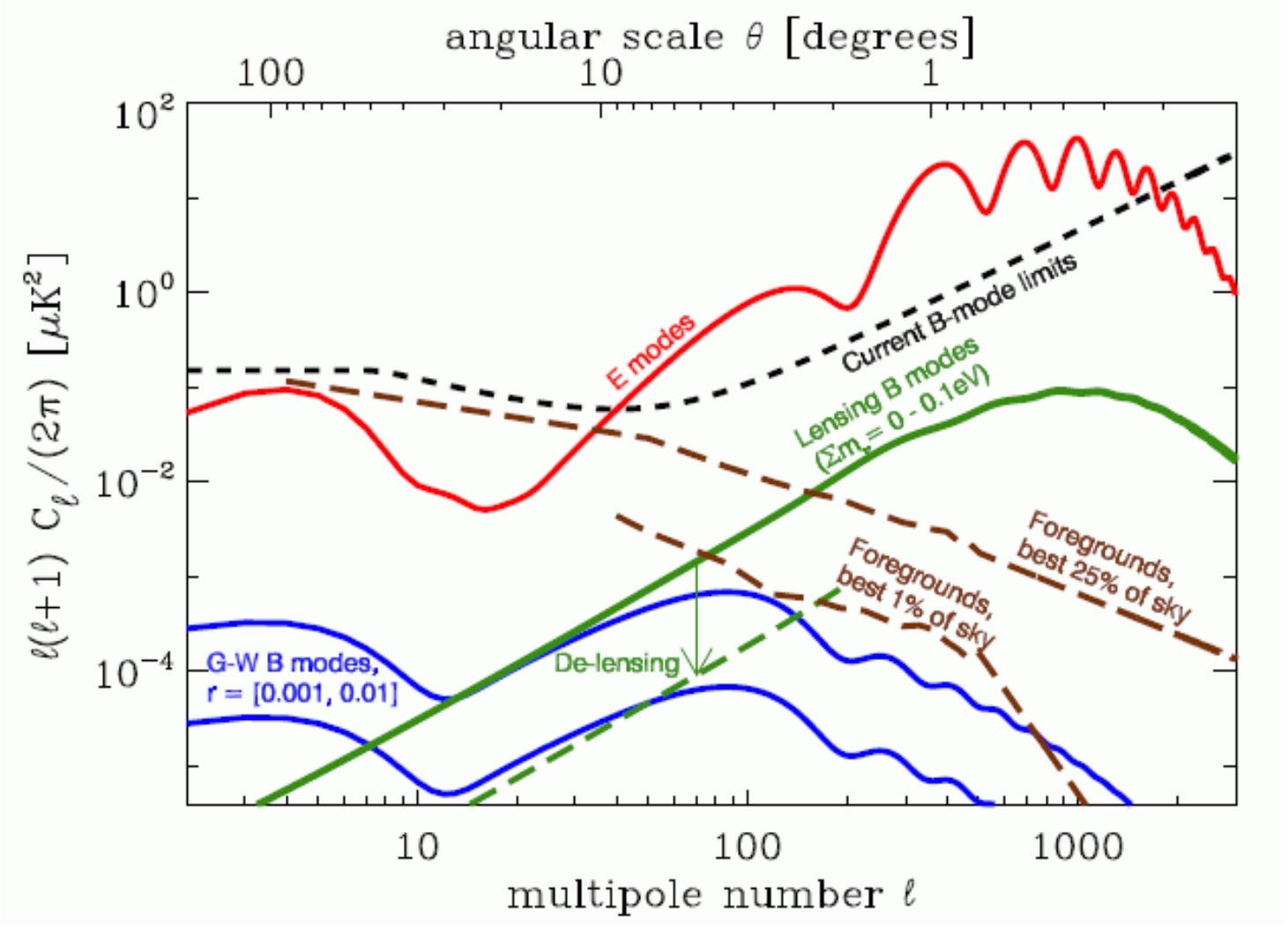
Three fundamental *maps* of our Universe

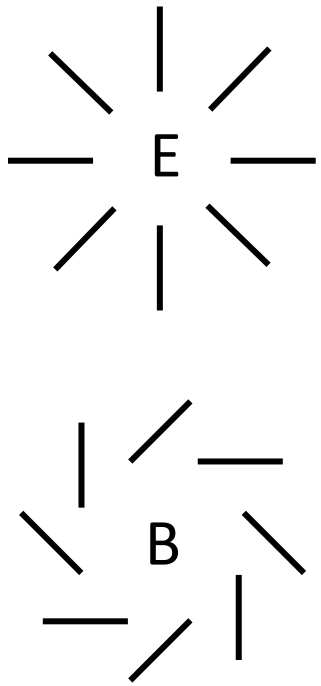
Map # 3
(lensing deflection field)



Planck

CMB B-mode polarization





Snowmass Process (HEP community summer study)

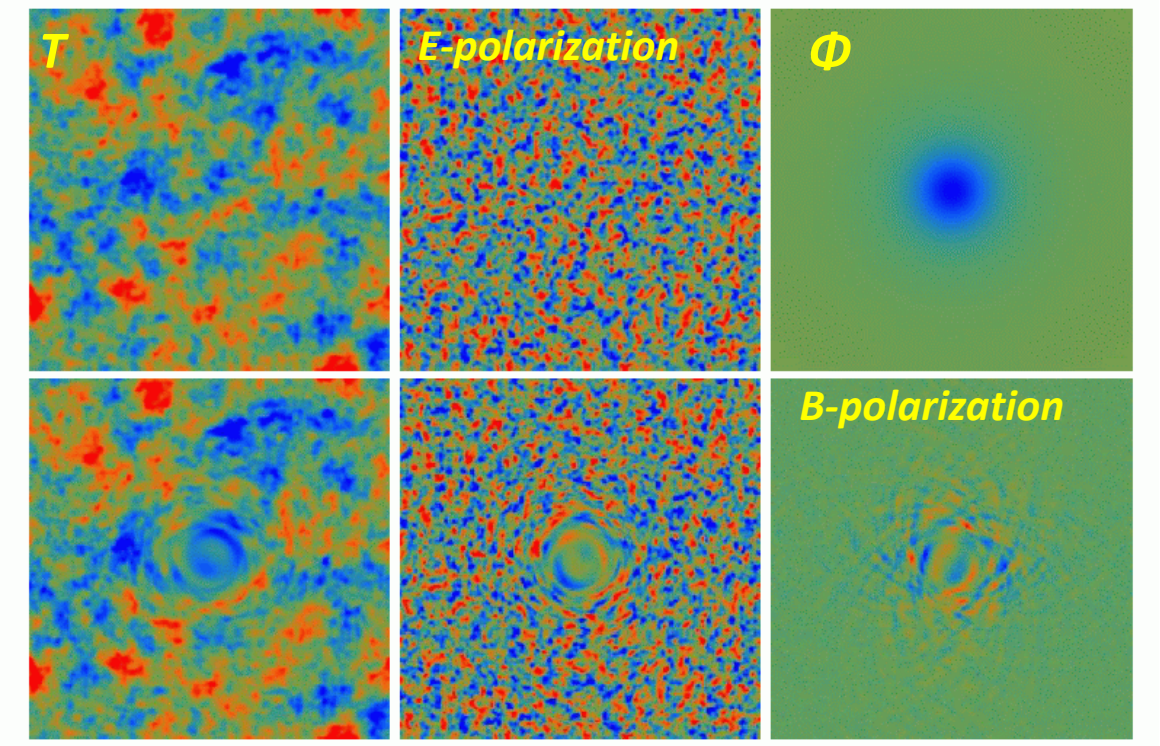
- **P5 = US Particle Physics Projects Prioritization Panel**

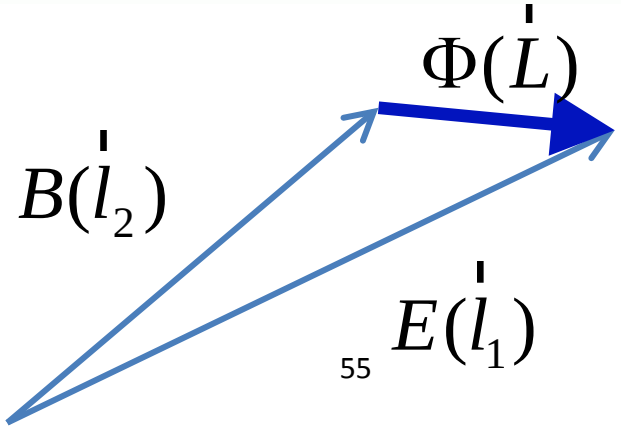
May 23, 2014:

"Recommendation 18: Support CMB experiments as part of the core particle physics program. The multidisciplinary nature of the science warrants continued multiagency support."

Budget for CMB-S4 project is included under all three budget scenarios considered.

Lensing reconstruction

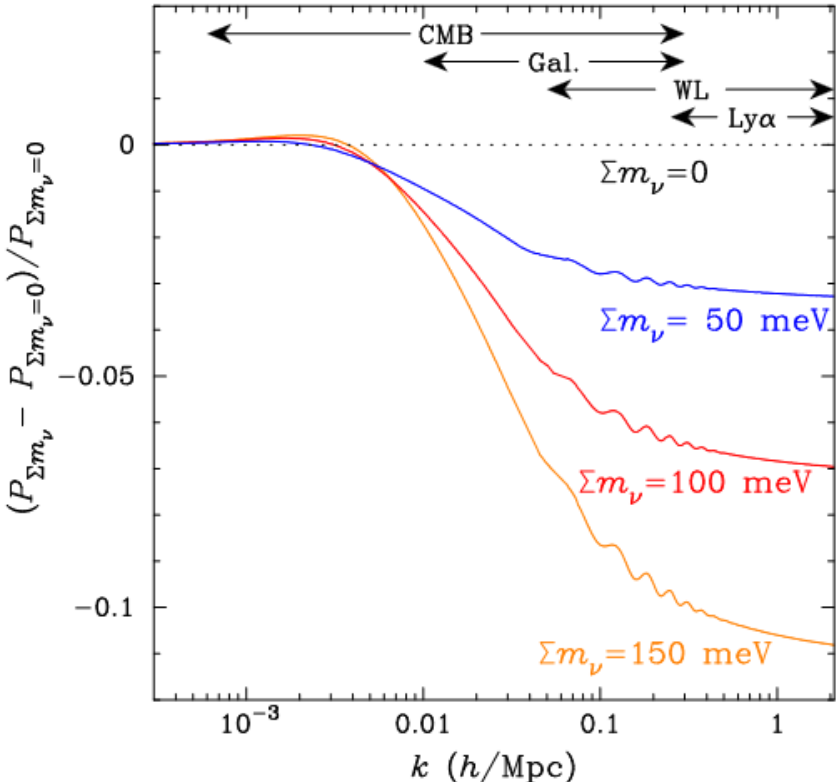


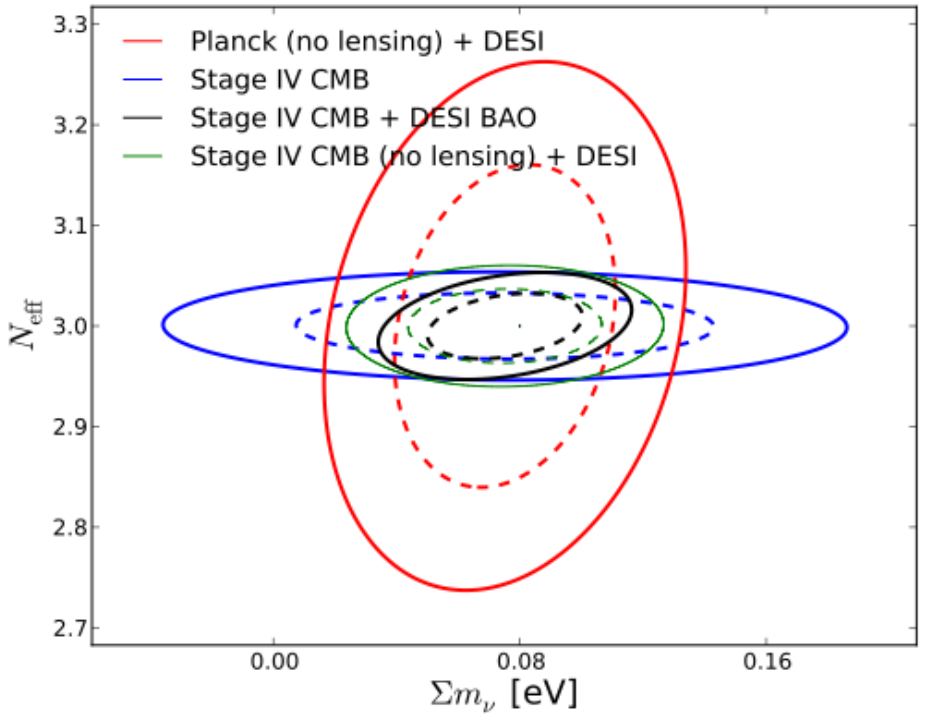


CMB as a probe of neutrinos

Arcminute-measurements of CMB polarization is a probe of cosmic neutrinos

- Number of relativistic species N_{eff} (CMB "damping tail" in polarization)
- Sum of neutrino masses ($\sum m_\nu$) suppresses structures (lensing *B*-modes)





ENCYCLOPEDIA

A Guide to Designing Future Ground-based CMB Experiments

W. L. K. Wu,[1,2,*] J. Errard,[3,4] C. Dvorkin,[5] C. L. Kuo,[1,2] A. T. Lee,[4,6] P. McDonald,[6] A. Slosar,[7] and O. Zahn[8]

[1] *Department of Physics, Stanford University, 382 Via Pueblo Mall, Stanford, CA 94305, USA*
[2] *Kavli Institute for Particle Astrophysics and Cosmology, SLAC, 2575 Sand Hill Road, M/S 29, Menlo Park, CA 94025, USA*
[3] *Computational Cosmology Center - Lawrence Berkeley National Lab, 1 Cyclotron Road, Berkeley, CA 94720, USA*
[4] *Department of Physics, University of California, Berkeley, CA 94720, USA*
[5] *Institute for Advanced Study, School of Natural Sciences, Einstein Drive, Princeton, NJ 08540, USA*
[6] *Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA*
[7] *Brookhaven National Laboratory, Upton, NY 11973, USA*
[8] *Berkeley Center for Cosmological Physics, Department of Physics, University of California, and Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*

(Dated: February 18, 2014)

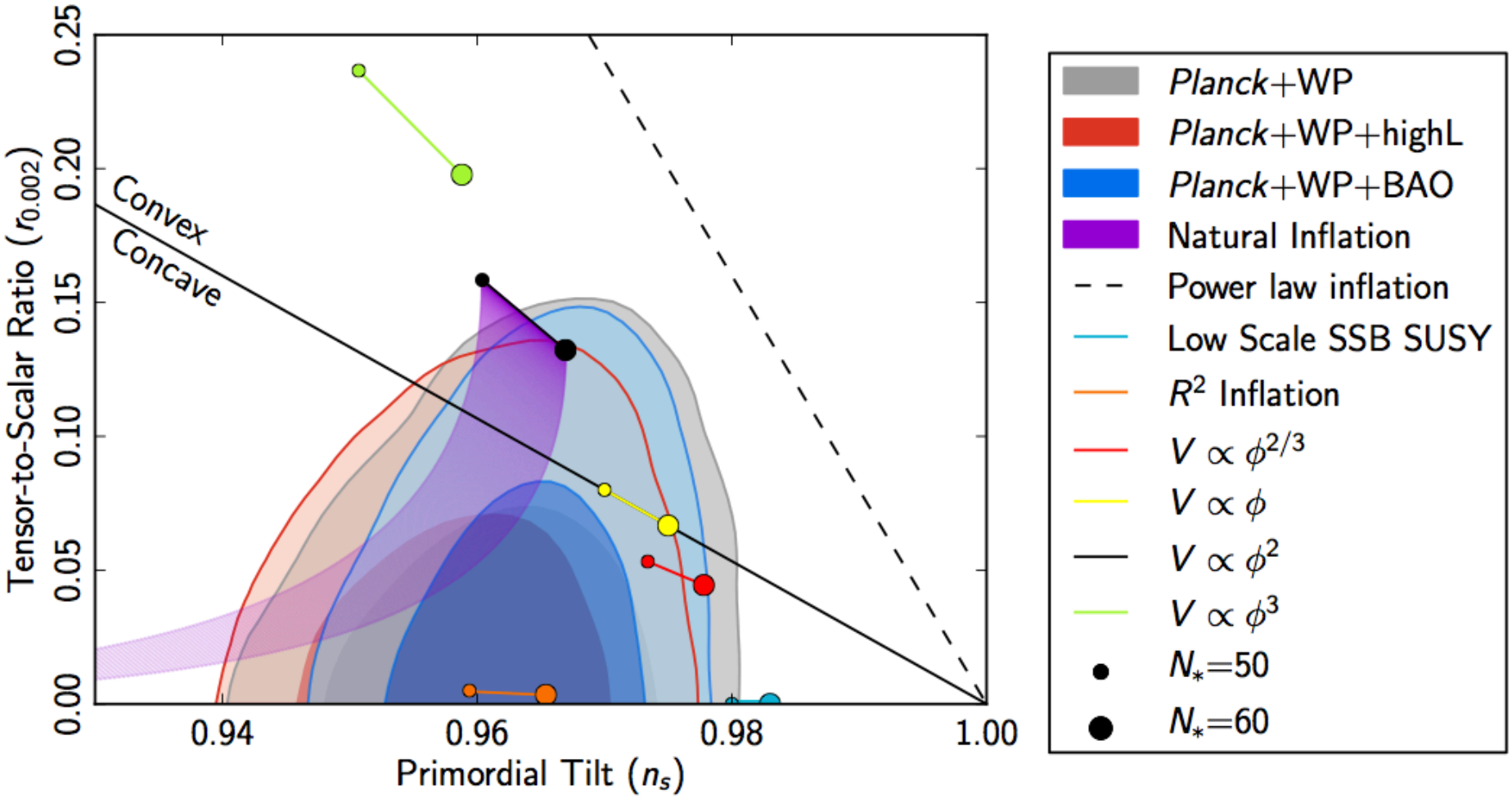
In this follow-up work to the High Energy Physics Community Summer Study 2013 (HEP CSS 2013, a.k.a. SNOWMASS), we explore the scientific capabilities of a future Stage-IV Cosmic Microwave Background polarization experiment (CMB-S4) under various assumptions on detector count, resolution, and sky coverage. We use the Fisher matrix technique to calculate the expected uncertainties in cosmological parameters in $\nu\Lambda$CDM that are especially relevant to the physics of fundamental interactions, including neutrino masses, effective number of relativistic species, dark-energy equation of state, dark-matter annihilation, and inflationary parameters. To further chart the landscape of future cosmology probes, we include forecasted results from the Baryon Acoustic Oscillation (BAO) signal as measured by DESI to constrain parameters that would benefit from low redshift information. We find the following best 1-σ constraints: $\sigma(M_\nu) = 15$ meV, $\sigma(N_{\rm eff}) = 0.0156$, Dark energy Figure of Merit $= 303$, $\sigma(p_{ann}) = 0.00588 \times 3 \times 10^{-26}$ cm^3/s/GeV, $\sigma(\Omega_K) = 0.00074$, $\sigma(n_s) = 0.00110$, $\sigma(\alpha_s) = 0.00145$, and $\sigma(r) = 0.00009$. We also detail the dependences of the parameter constraints on detector count, resolution, and sky coverage.

Inflation: spectral index

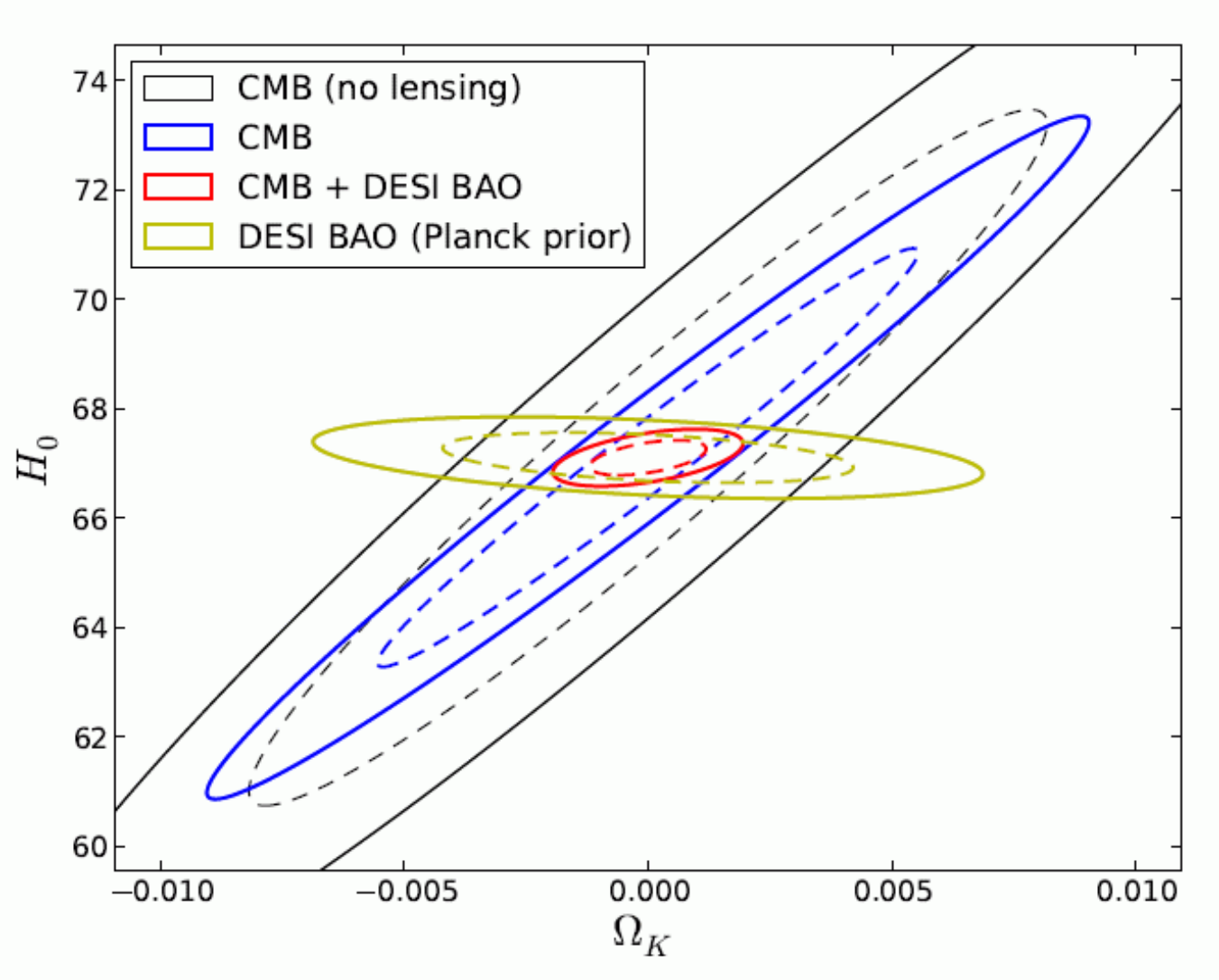
	CMB				CMB+BAO			
	$1'$	$2'$	$3'$	$4'$	$1'$	$2'$	$3'$	$4'$
10^4 detectors								
$f_{sky} = 0.25$	2.91	2.94	2.98	3.04	2.19	2.23	2.29	2.36
$f_{sky} = 0.50$	2.11	2.13	2.16	2.21	1.64	1.67	1.71	1.75
$f_{sky} = 0.75$	1.76	1.77	1.80	1.83	1.39	1.42	1.45	1.48
10^5 detectors								
$f_{sky} = 0.25$	2.66	2.73	2.80	2.86	1.93	1.98	2.04	2.12
$f_{sky} = 0.50$	1.94	1.97	2.01	2.06	1.44	1.47	1.51	1.56
$f_{sky} = 0.75$	1.60	1.63	1.66	1.70	1.22	1.24	1.28	1.32
10^6 detectors								
$f_{sky} = 0.25$	2.38	2.48	2.62	2.73	1.70	1.76	1.83	1.92
$f_{sky} = 0.50$	1.75	1.81	1.90	1.96	1.28	1.32	1.37	1.43
$f_{sky} = 0.75$	1.45	1.51	1.57	1.61	1.10	1.12	1.16	1.20

Table V: n_s 1-σ constraints in units of 10^{-3} from CMB and from CMB+BAO. "CMB" includes lensing.

Inflation: spectral index



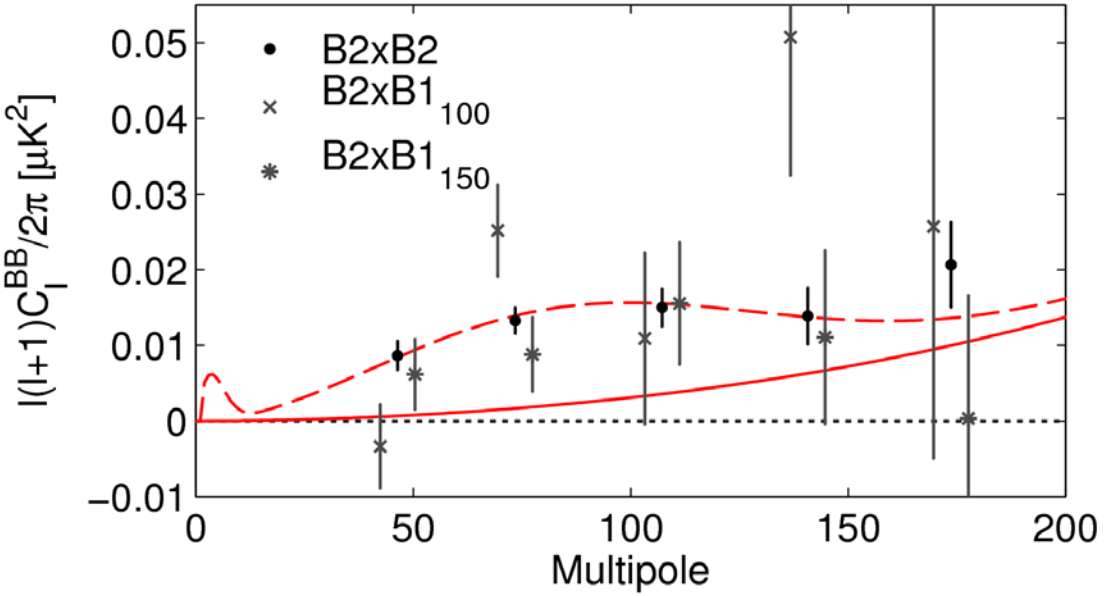
Inflation: Mean spatial curvature



1σ ~
few x 10^{-3}

Thank you !
Keith Vanderlinde

Spectral Index of the B-mode Signal

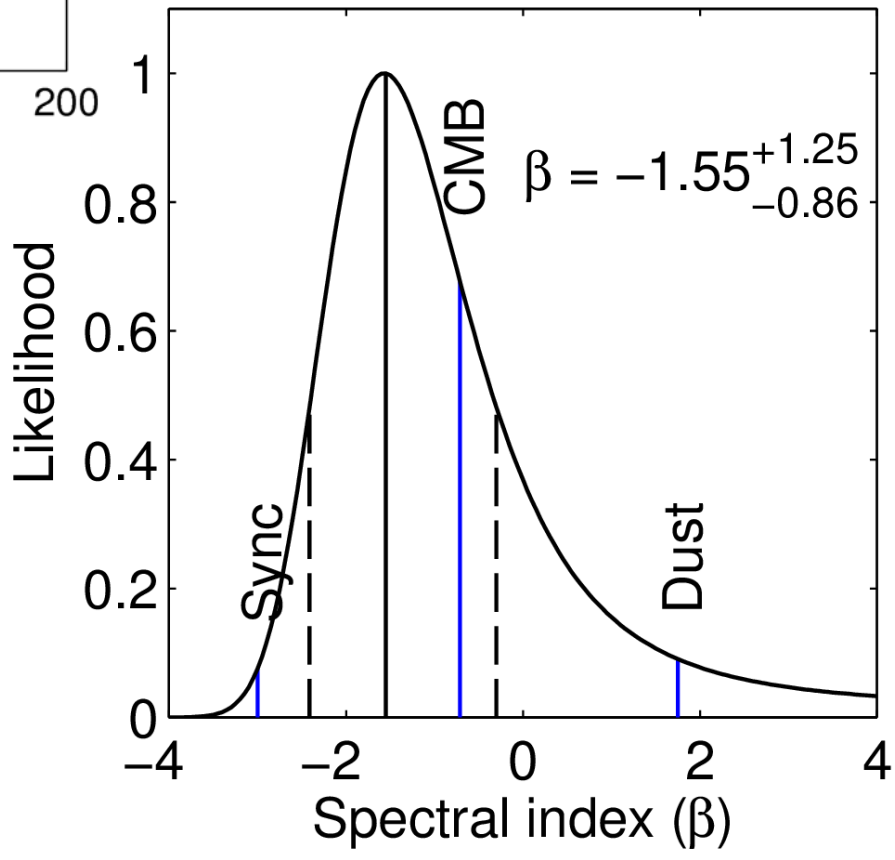


Likelihood ratio test: consistent with CMB spectrum, disfavor pure dust/sync at **2.2/2.3σ**

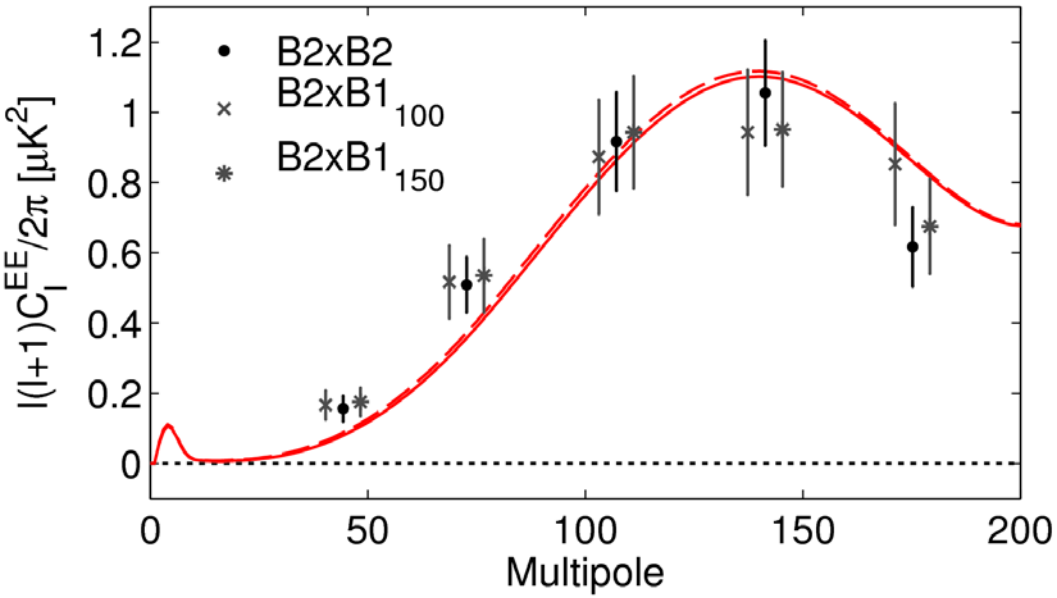
Comparison of B2 auto with $B2_{150}$ x $B1_{100}$ constrains signal frequency dependence, independent of foreground projections

If **dust**, expect little cross-correlation

If **synchrotron**, expect cross higher than auto



Spectral Index of the E-mode Signal

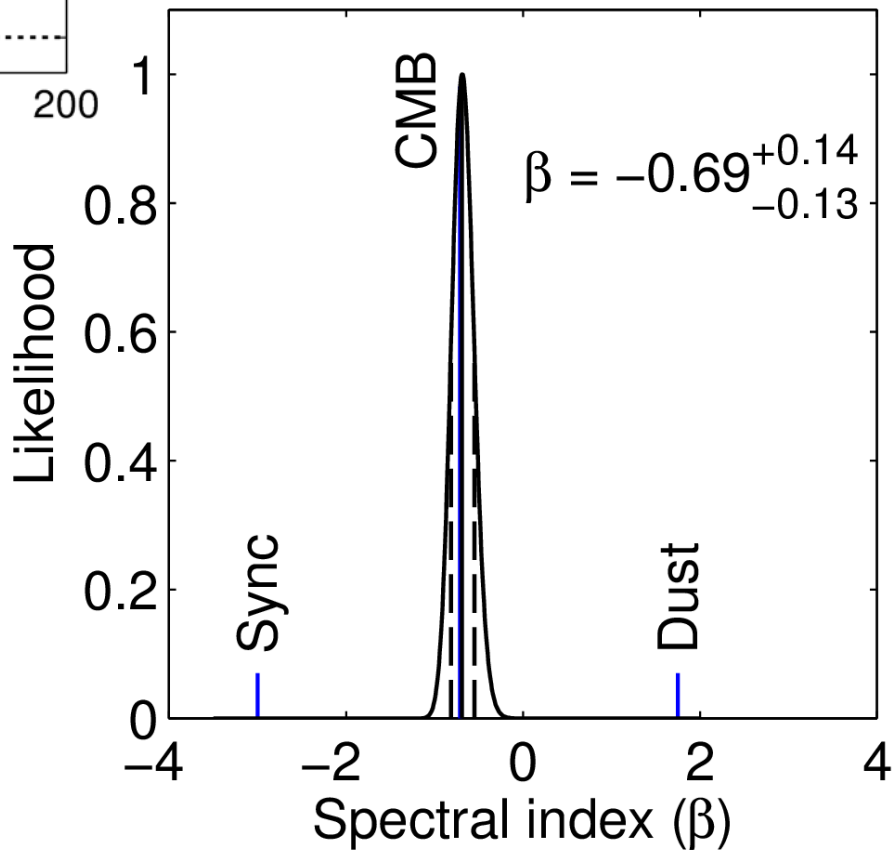


Likelihood ratio test: consistent with CMB spectrum, disfavor pure dust/sync at **11/30σ**

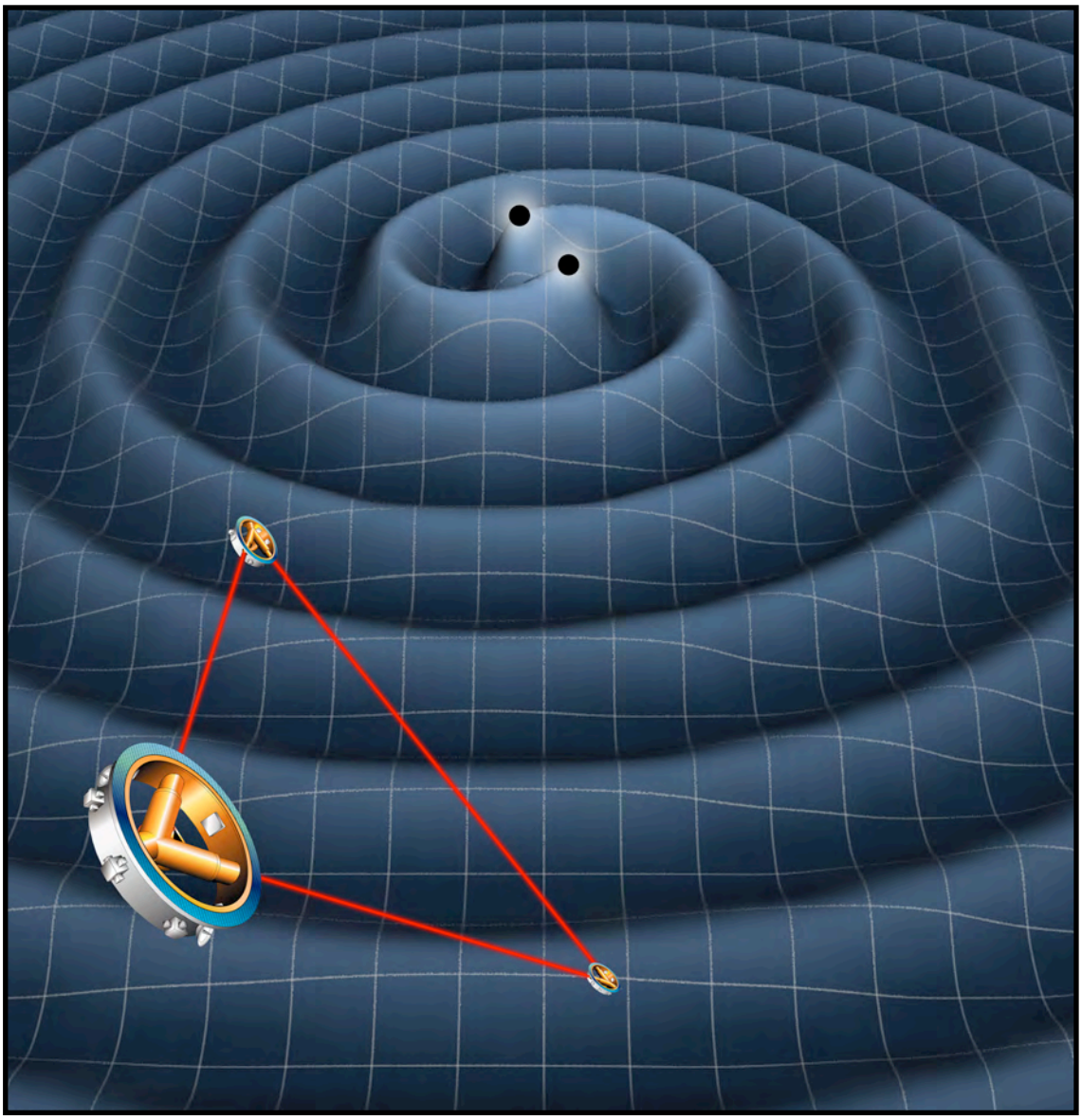
Comparison of B2 auto with $B2_{150}$ x $B1_{100}$ constrains signal frequency dependence, independent of foreground projections

If **dust**, expect little cross-correlation

If **synchrotron**, expect cross higher than auto



LISA: Probing the Universe with Gravitational Waves



31 May 2007

(Revised: 31 March 2009)

LISA-LIST-RP-436
Version 1.2
10 March 2009

This is a revision of a document originally prepared for the National Research Council *Beyond Einstein Program Assessment Committee* (BEPAC) in 2007.

Prepared by:

Lead author: Tom Prince (Caltech/JPL)

Contributors and editors: John Baker (GSFC), Pete Bender (Univ of Colorado), Pierre Binetruy (APC - Paris), Joan Centrella (GSFC), Teviet Creighton (JPL), Jeff Crowder (JPL) , Curt Cutler (JPL), Karsten Danzman (Univ. of Hannover and the Albert Einstein Institute), Steve Drasco (JPL) , Lee S. Finn (Penn State), Craig Hogan (Univ. of Washington), Piero Madua (Univ of Santa Cruz), Cole Miller (Univ of Maryland), Milos Miloslavljevic (Univ. of Texas, Austin), Gijs Nelemans (Radboud University Nijmegen), Sterl Phinney (Caltech), Bonny Schumaker (JPL), Bernard Schutz (Albert Einstein Institute), Michele Vallisneri (JPL), Marta Volonteri (Univ. of Michigan) and Karen Willacy (JPL).

Thanks to the many other members and friends of the LISA science community who provided valuable input and suggestions.

Table of Contents

Executive Summary 1

1. Gravitational Waves: An Overview 9
2. LISA Mission Overview 19
3. Black Hole Astrophysics: Supermassive Black Holes in Galactic Nuclei 27
4. Black Hole Physics: Confronting General Relativity with Precision Measurements of Strong Gravity 43
5. Precision Cosmometry and Cosmology 57
6. Ultra-compact Binaries 66
7. New Physics and the Early Universe 81
8. LISA and the Key Questions of Astronomy and Physics 88

Appendices

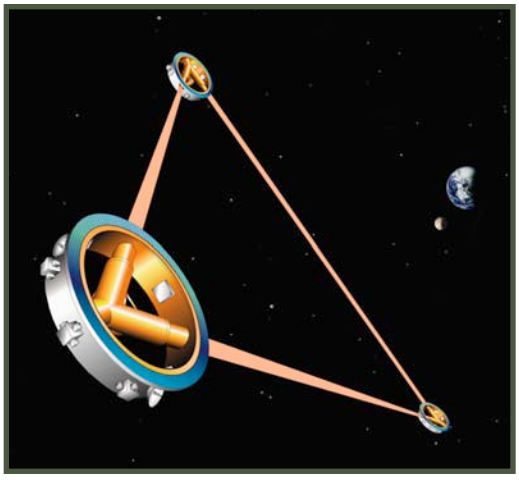
A1. LISA Science Objectives 92
A2. Acronyms 94
A3. References 96

Executive Summary

The New Science of Gravitational Waves

Einstein's theory of spacetime and gravity, general relativity, predicts that motions of mass produce propagating vibrations that travel through spacetime at the speed of light. These gravitational waves (as the vibrations are called) are produced abundantly in the Universe and permeate all of space[1]. Measuring them will add an altogether new way to explore what is happening in the Universe: rather than studying the propagation and transformation of conventional particles and fields in spacetime, as all science has done up to now, we will sense vibrations of the fabric of spacetime itself. Studying this new form of energy will convey rich new information about the behavior, structure, and history of the physical universe, and about physics itself. When gravitational waves become observable they will provide a new and uniquely powerful probe of the extremes of spacetime, from the Big Bang to black holes, to address the deep questions that have emerged in Einstein's vision of the cosmos: what powered the Big Bang, what happens to space and time in black holes and what is the mysterious dark energy accelerating the expansion of the Universe?

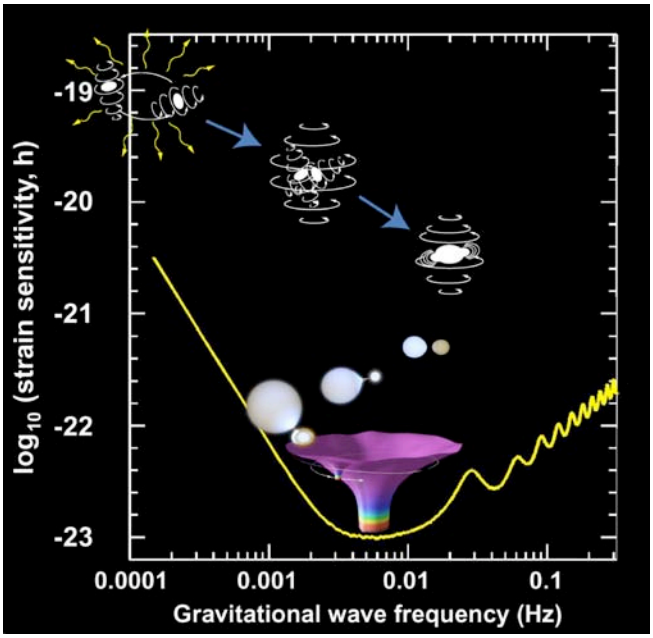
LISA is a space mission designed to measure gravitational radiation over a broad band at low frequencies, from about 0.03 to 100 millihertz, a band where the Universe is richly populated in strong sources of gravitational waves. It measures signals from a wide range of different sources: massive black holes merging in galaxies at all distances; massive black holes consuming smaller compact objects; known binary compact stars and stellar remnants; members of known populations of more distant binaries; and probably other sources, possibly including relics of the extremely early Big Bang, which are as yet unknown.



[1] Gravitational waves have never been directly detected. Nevertheless, the existence of gravitational waves is in little doubt as their effects have been measured precisely, if indirectly. Any theory of gravity consistent with special relativity will exhibit gravitational waves, and the predictions of general relativity should be quantitatively reliable for LISA because the long standing best evidence for of gravitational waves is the orbital decay of the Hulse-Taylor binary pulsar, which radiates at frequencies only marginally below LISA's operating band. Moreover, LISA is sensitive to any time-dependent tidal gravitational force, even to those of Newtonian theory produced by, say, the Sun. Therefore LISA will be able to detect the gravitational waves predicted by any reasonable theory of gravity.

These strong signals convey detailed information addressing a wide range of science: the history of galaxies and black holes in the Universe; general relativity itself and the behavior of spacetime; precision measurements of the Universe as a whole; the physics of dense matter and stellar remnants; and possibly new physics associated with events in the early Universe or relics predicted in string theory.

In the same way that electromagnetic radiation accompanies acceleration of electric charges, gravitational radiation accompanies quadrupolar acceleration of any kind of mass or energy. Quadrupolar motion in a system with mass M and size R at a distance D typically perturbs spacetime with a dimensionless metric-strain amplitude (fractional variation in proper spatial separations) of about $h \approx \left(GM/(Rc^2)\right)^2 (R/D)$. LISA senses this by monitoring the changes in the distances between inertial proof masses. LISA uses precision laser interferometry across a vast distance of space to compare separations among proof masses that are protected by the spacecraft from non-gravitational disturbances. Over a separation of $L \approx 5$ million kilometers, LISA's ability to measure variations of $\Delta L \approx 0.05$ picometers (rms over one year observation) corresponds to a strain amplitude sensitivity of about $h \approx \Delta L/L \approx 10^{-23}$. LISA *coherently* measures spacetime strain variations, including frequency, phase, and polarization, all of which reflect large-scale properties of the systems that produce them and are therefore direct traces of the motions of distant matter.



The LISA strain sensitivity curve showing approximate frequency range and strain amplitude of three categories of sources: massive black hole inspiral, ringdown, and merger (uppermost source); ultra-compact Galactic binaries (middle source); capture of stellar mass compact object by massive black holes (lower source).

LISA is an astronomical observatory of unprecedented versatility and range. Its all-sky field of view ensures that it misses nothing. Its coherent mode of observing allows it to resolve and distinguish overlapping signals and locate them on the sky. Its dynamic range of 10^5 or more in amplitude (10^{10} in energy) allows it to study sources within the Galaxy and out to the edge of the Universe. Finally, LISA's wide frequency band (more than three decades) allows it to study similar sources of widely different masses and cosmological redshifts. Because gravitational waves penetrate all regions of time and space with almost no attenuation, LISA can sense waves from the densest regions of matter, the earliest stages of the Big Bang, and the most extreme warpings of spacetime near black holes.

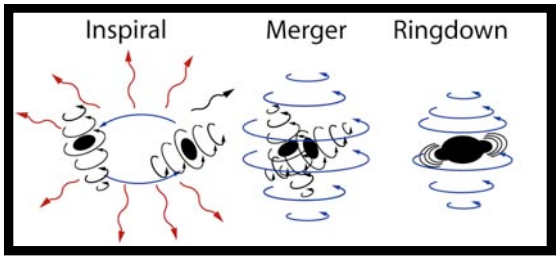
Survey of LISA Science

We now provide a brief survey of the key scientific measurements that LISA will perform. These measurements address the basic scientific goals of the LISA mission, which are captured formally in the LISA Science Objectives listed in the first appendix. The scientific background for the LISA science measurements and objectives is discussed extensively in the various sections of this document.

LISA records the inspirals and mergers of binary black holes, the most powerful transformations of energy in the Universe, allowing precision measurements of systems composed only of pure dynamical spacetime.

The strongest gravitational waves are generated by systems with the largest gravitational fields GM/R, hence large masses and small sizes. The strongest of all are generated by interactions of black holes, dense knots of pure spacetime energy with $GM/Rc^2 \approx 1$. At LISA frequencies the strongest sources are massive black hole binaries with about 10^4 to 10^7 times the mass of the sun. Two black holes orbit each other, spiral together as they lose energy by radiation, and finally merge. The waves from these events — many cycles over a long inspiral, climaxing in a brief series of powerful waves during a violent merger, and a final ringdown to a quiescent single black hole — record dynamical general relativity not only in its purest form but also in its most violent, nonlinear behavior: a maximally warped vacuum spacetime interacting with itself.

The black hole binaries start with wide orbits at low frequencies. As they lose energy their frequency increases and their radiation strengthens. A typical source enters the LISA band a year or more before the final merger so many orbits are recorded, encoding details of the system properties and behavior, position on the sky, and absolute distance. The coherent phase and polarization information obtained over LISA's solar-orbit baseline (and variable inclination) can often pinpoint where a source is in the sky to better than a degree. In the last hours or minutes the signal-to-noise ratio grows very high, often into the hundreds to thousands depending on distance. At its peak luminosity, around the moment of merger, a black hole binary is the most extreme transformation of mass-energy of any kind in the Universe, radiating a power $\approx 10^{-3}c^5/G$ (or ~10^{49} watts), in a few wave cycles, for a time of about 100 GM/c^3. This peak radiated power is about 1000 times more than all the stars in the visible Universe. The merger throes of a million solar mass binary black hole merger last about 500 seconds. Massive black hole binary inspiral and merger

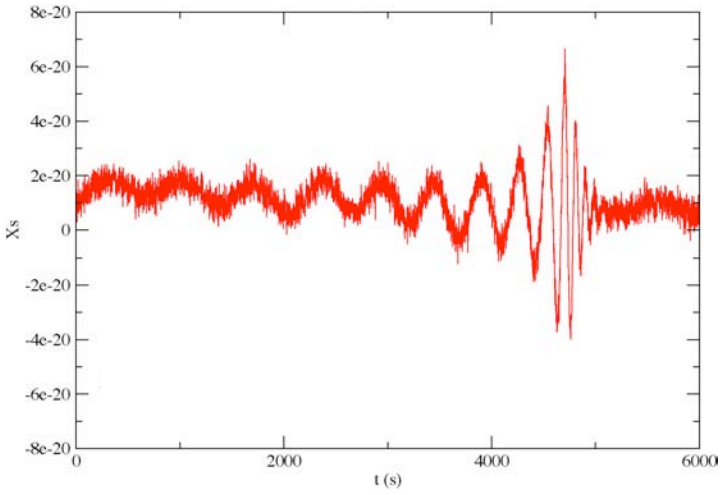


events are such powerful radiators that LISA can detect them anywhere, out to the largest redshifts where galaxies might exist.

The detailed study of waveforms from black hole binaries provides a rich testbed for general relativity. Recent breakthroughs now allow numerical computation of Einstein's field equations throughout the entire inspiral and merger event, yielding a detailed map of the predicted gravitational waveform that will be the first detailed test of dynamical, strong-field general relativity. Waveforms coherently correlated over many orbits (from ≈ 10 to ≈ 1000 depending on mass and redshift) recorded in the LISA signal stream, and detection of events with a signal to noise of a thousand or more, allow precise tests of the theory as well as precise measurements of system parameters to a precision of order 10^{-2} to 10^{-3}. Comparison with the computed details of the inspiral and merger waveform will provide a powerful test of the binary black hole model assumed for these systems.

LISA will map isolated black holes with high precision, verifying whether they are the stationary "no hair" spacetime configurations described by the Kerr metric, completely specified by four numbers: the mass and three components of spin.

In general relativity the final isolated spinning black hole is described mathematically as a particular, precisely specified spacetime shape called a Kerr metric, that depends only on the physics of gravity and not at all on the history or environment of the black hole. Comparison of the ringdown waveform with theory can verify that the final black hole which arises from a merger is indeed described by the Kerr solution, and satisfies the "no hair" theorem of general relativity that states that an isolated, stationary black hole is completely specified by its mass, charge and angular momentum. The LISA signals during the merger phase are so strong that the signal-to-noise ratio is often greater than 100 even in one oscillation cycle: signal waveforms are visible on an oscilloscope type display of raw data even to the naked eye, so even if general relativity were to be wrong at the levels allowed by our existing tests (*e.g.* the double binary pulsar J07037 -3039) we would be able to use LISA data to make sense of what is happening.



Gravitational wave signal for the final few orbits, plunge, merger and ringdown of a massive black hole binary. The signal is the sum of the gravitational waveform and simulated LISA noise. Note that the waveform stands up well above the noise and is visible in fine detail. A merger event such as that shown here can have a total signal to noise of 200 even at a redshift, z = 15.

A small compact object orbiting a spinning massive black hole.

LISA also uses a second type of source to explore the spacetime near a massive black hole. Driven by chance encounters, a much smaller mass compact objects — such as a degenerate dwarf, neutron star or stellar-mass black hole — sometimes finds itself captured by the massive black hole, after which it orbits many times until it finally plunges into the horizon and disappears. The gravitational waves from these extreme mass ratio inspirals (EMRIs) encode a detailed map of a relatively unperturbed massive black hole, predicted to be a pure Kerr knot of highly curved, spinning spacetime. About 10^5 wave cycles are measured for each source, emitted from orbital paths exploring deep into different parts of the relativistic region near the massive black hole. The specific mass quadrupole and higher moments predicted by the Kerr solution are measured with a precision of about 10^{-4}, and precision tests of small variations about the equilibrium Kerr solution — the small amounts of "hair" added by the perturbing object — are measured at the one percent level. Gravitational waves from these events map in exquisite detail the cleanest and most accurately predicted structures in all of astrophysics, whose mathematical elegance Chandrasekhar once likened to that of atoms.

LISA directly observes how massive black holes form, grow, and interact over the entire history of galaxy formation.

Optical, radio and x-ray astronomy have produced abundant evidence that nearly all galaxies have massive black holes in their central nuclei (and indeed that some recently merged galaxies even have two black holes). These nuclear black holes have a profound effect on galaxy formation; the influence of black hole powered jets on the intergalactic gas out of which galaxies form is in some cases directly observed. There is a circumstantial case, but no direct evidence, that the formation of this population of black holes was associated with a multistage process of binary inspiral and merger, together with accretion. LISA will obtain direct and conclusive evidence and study details of this process via gravitational radiation.

In standard concordance cosmology, the first massive black holes naturally arise from the very first, supermassive stars. In this scenario, black hole binaries begin to form from a high redshift, $z \approx 20$, when galaxies start to assemble by a series of (hundreds to thousands of) hierarchical mergers of smaller protogalaxies. When two galaxies merge into one, their central black holes sink to the center of the new galaxy, find each other, inspiral and merge. There are so many galaxies forming in the Universe observed by LISA that mergers happen quite frequently: estimates based on standard galaxy formation theory suggest that if black holes indeed grew by hierarchical merging, LISA detects a merger event about once or twice every week on average, from a wide range of redshifts extending back to massive binaries in early protogalaxies at $z \approx 15$. At any given time, in addition to the actual mergers, these models predict

that LISA observes inspiral signals from hundreds of binaries in the final years before their merger. LISA digs directly and intimately into the detailed evolution of galactic nuclei: the large sample of binaries provides a direct record of the whole history of galaxy formation in the observable Universe, and of the processes that grew their central black holes and shaped their nuclei.

In addition to mergers of massive black holes, LISA will also observe the inspiral of stellar mass black holes into the massive black holes in the centers of normal galaxies. These are the extreme mass ratio inspiral events (EMRIs) mentioned earlier. The parameters measured from extreme mass ratio events yield a census of isolated massive black hole spins and masses in many galaxies today, a revealing relic of black hole history. The local universe also produces observable inspirals of less compact stars and stellar remnants that probe the rich astrophysics near the massive central black holes as they consume piecemeal the various stellar populations in their vicinity.

LISA measures precise, gravitationally-calibrated absolute luminosity distances to high redshift, with the potential of contributing uniquely to measurement of the Hubble constant and dark energy.

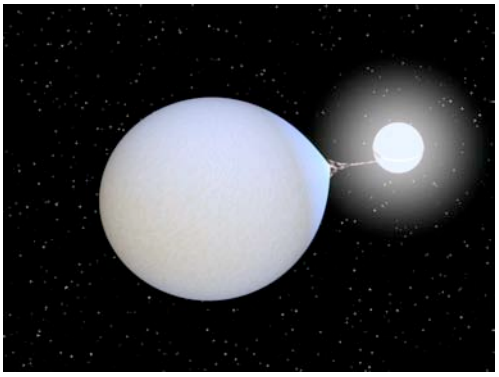
Because the inspiral leading up to the merger is a clean, pure vacuum spacetime system of two black holes, properties of the radiation can be computed exactly in general relativity, so that the black hole masses, spins, orientations and even the exact distance can be reconstructed from LISA data. (Roughly speaking, the final wave cycle period tells the final absolute Schwarzschild radius, and the ratio of that length to the distance is the metric strain, h.) These inspiral distances are both individually precise and absolutely calibrated, using only pure gravitational physics, and they cover a wide range of redshift. In the absence of lensing effects the absolute physical luminosity distance to a single LISA inspiral event is typically estimated from the waveform alone with on the order of one percent precision, and in some cases with as good as 0.1% precision. If identification of the host galaxy[2] allows an independent redshift determination, the redshift-distance relation is also measured with high precision. Black hole binaries thus represent a unique and independent new capability for precision cosmology that complements other techniques. Even a small number of sources at moderate redshift calibrates the distance scale and Hubble constant an order of magnitude better than any current method — a powerful constraint on dark energy models in combination with microwave background data. The expected large sample of high redshift inspiral events may lead to measurements of dark energy

[2] LISA's waveform fitting can often pinpoint the direction of a source to much better than a degree, and the distance estimate also narrows the redshift range considerably; nevertheless there may be many thousands of galaxies in the LISA "error box" for a given source. Models suggest that the host may be identified from a telltale nuclear starburst associated with the merger, or from variability associated with the disrupted disks around the merging holes, but galaxy nuclei are too little understood to make a firm prediction. Study of LISA electromagnetic counterparts may provide an exploratory bonanza for wide field synoptic imaging and spectroscopy across the electromagnetic spectrum, but it is also possible that identification of hosts will prove elusive.

parameters comparable in precision to other methods, but with independent calibration and completely different systematic errors. The main source of error, especially at high redshift, is the noise induced by cosmic weak gravitational lensing along the line of sight, but in a statistical sample this is controllable, and indeed provides unique new information about the nature and clustering of dark matter over time.

LISA studies in detail thousands of compact binary stars in the Galaxy, providing a new window into matter at the extreme endpoints of stellar evolution.

In addition to mergers and meals of distant black holes, LISA detects many lower mass binary systems in our galaxy, mostly very compact remnants of normal stars, called white dwarfs. Very soon after turning on, LISA will quickly detect a handful of nearby binary compact stars already studied and named. These "verification binaries" provide sources with known positions and periods ensuring particular, predictable LISA signals. Signals are also certain to appear from populations in our galaxy of numerous and various remnants, including white dwarfs and neutron stars, which are known to exist from some that emit electromagnetically. Simple extrapolation of known nearby samples to the whole Galaxy predicts that LISA will detect thousands of binaries. The most compact binaries (those at high frequency) will be measured in detail as individual sources from across the Galaxy, while at lower frequencies only the nearby ones will be individually distinguished; millions of others from across the Galaxy will blend together into a confusion background. LISA provides distances and detailed orbital and mass parameters for hundreds of the most compact binaries, a rich trove of information for detailed mapping and reconstruction of the history of stars in our galaxy, and a source of information about tidal and other non-gravitational influences on orbits associated with the internal physics of the compact remnants themselves. LISA may also detect at high frequencies the background signal from compact binaries in all the other galaxies.



LISA may find entirely new phenomena of nature not detected using light or other particles.

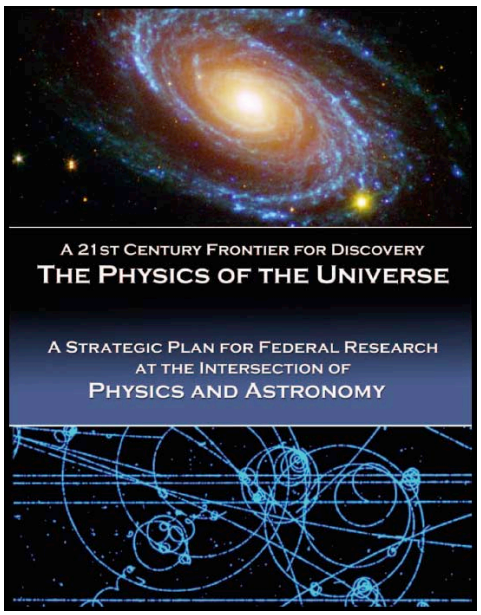
Given that all forms of mass and energy couple to gravity, other sources of gravitational waves may exist that are not known from extrapolating current electromagnetic observations. LISA's frequency band can indeed be extrapolated to very high redshift where we do not yet have any direct observations, and to a regime where LISA itself will be our first information of any kind about the nonlinear behavior and motion of matter. For example, the LISA frequency band in the relativistic early Universe corresponds to horizon scales at the Terascale frontier, where phase transitions of new forces of nature or extra dimensions of space may have caused catastrophic, explosive bubble growth and efficient gravitational wave production. LISA is capable of detecting a stochastic background from such events from about 100 GeV to about

1000 TeV, if gravitational waves in the LISA band were produced with an overall efficiency more than about 10^{-7}, a typical estimate from a moderately strong relativistic first-order phase transition. This corresponds to times about 3×10^{-18} to 3×10^{-10} seconds after the start of the Big Bang, a period not directly accessible with any other technique. Reaching much further still beyond the range of any particle accelerator, LISA also deeply probes possible new forms of energy such as cosmic superstrings, relics of the early Universe predicted in some versions of string theory, that are invisible in all ways except by the gravitational waves they emit. In principle, their signature could provide direct evidence for new ideas unifying all forms of mass and energy, and possibly even spacetime itself.

LISA and the Key Questions of Astronomy and Physics

LISA addresses forcefully and directly a broad cross-section of the research priorities and key questions raised by recent astronomy and physics decadal, community, agency and White House reports. These include National Research Council reports such as *Astronomy and Astrophysics in the New Millenium* and *Connecting Quarks with the Cosmos*, as well as the National Science and Technology Council report *A 21st Century Frontier of Discovery: The Physics of the Universe*.

LISA approaches many deep questions set forth in these documents in a new and unique way not possible with any other space- or ground-based observatory. Indeed, LISA has the potential to transform much of physics and astronomy: it senses the remote Universe in an entirely new way, and explores many new kinds of phenomena that can be explored in no other way. Its discoveries are likely to expand the scope of both astronomy and physics significantly and to reshape the science questions of the future. By directly sensing the dynamical activity of spacetime, LISA feels the beating heart of Einstein's cosmos. The new science of gravitational waves provides the most direct way to advance the goals of NASA's Beyond Einstein program.



1. Gravitational Waves: An Overview

Most everything we know about the Universe we have learned from *light*: since ancient times, electromagnetic waves have been messengers from the cosmos to our eyes, and later to our telescopes and our antennas. More recently, we have begun to parse the messages of more exotic carriers: like the elusive neutrinos from the Sun and beyond. We are now ready to add an altogether new modality to science: sensing vibrations of the very fabric of spacetime. *Gravitational waves* will add a many-voiced soundtrack to the rich imagery of the cosmos (see Hogan 2006).

In Einstein's 1915 theory of general relativity, the geometry of spacetime is not a passive setting for the dynamics of matter and energy, but an equally dynamic player. Matter and energy cause spacetime curvature, which in its turn guides the free fall of matter and energy. Remarkably, spacetime can support curvature without any matter: *black holes*, the densest masses in the Universe, are objects of pure spacetime wrapped around itself; *gravitational waves* are self-sustaining, undulatory excitations of spacetime, carrying energy and traveling at the speed of light. Unlike electromagnetic radiation (but much like neutrinos) gravitational waves interact very weakly with matter, and can penetrate anything without losing intensity. This makes them powerful probes of faraway regions and extreme conditions, but it also makes them very hard to detect. Only recently has technology advanced to the point of

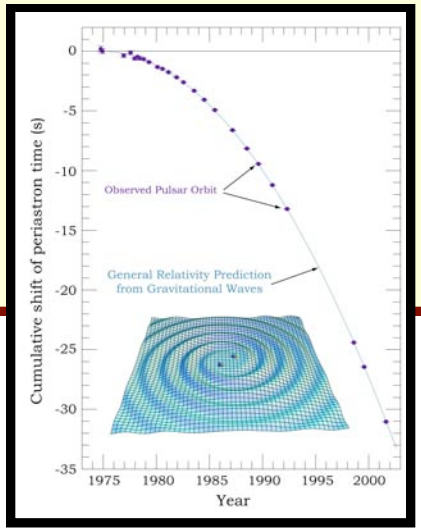
In 1967, the first radio pulsar was discovered by Jocelyn Bell and Anthony Hewish (for which Hewish received the Nobel Prize in 1974). Pulsars were quickly identified as neutron stars, the incredibly compressed remnants of the supernova explosions of stars.

In 1974, Russel Hulse and Joseph Taylor discovered the first binary pulsar, PSR 1913+16. This system consists of two neutron stars (of which only one emits as a pulsar) with an orbital period of eight hours. The general theory of relativity predicts that the orbiting stars stir spacetime around them, losing energy by emitting gravitational waves and therefore spiraling closer together.

In 1993, Taylor and Hulse received the Nobel prize for showing that the orbital period of PSR 1913+16 is decreasing at exactly the rate predicted by Einstein's "quadrupole formula" for the emission of gravitational waves in binaries.

In 2004 an international team of astronomers announced the discovery of PSR J0737-3039A/B, a *double* binary pulsar of two neutron stars, both of which emit detectable radiation. This system is a unique general-relativistic laboratory: timing studies of its emissions over 2.5 years have already improved on some of the most stringent tests on general relativity from other binary pulsars and from Solar-system dynamics (Kramer *et al.* 2006).

The gravitational waves from neutron binary stars await direct detection by LISA and by ground-based detectors.



Precession of periastron in PSR 1913+16 as measured in 27 years of radio data and as predicted by general relativity.

building apparatus sensitive enough to measure the minute effects of gravitational waves on matter.

Gravitational waves will reveal the most violent events in the Universe, the collision and coalescence of two black holes. In the final minutes before the merger, the power radiated in gravitational waves reaches 10^{49} W, a thousand times more luminous than all the stars in all the galaxies in the visible Universe put together. These mergers will allow us to test how well Einstein's equations work in such extreme conditions, offering us insight into the strongest and most violently dynamical spacetimes Nature has produced since the Big Bang.

Gravitational waves produce tiny oscillations in the distance between *freely-floating* masses isolated from all forces other than gravity; these oscillations can be monitored using laser interferometry. This principle is implemented in the ground-based Laser Interferometer Gravity-wave Observatory (LIGO), funded by NSF. LIGO and similar observatories worldwide will likely make the first detection of gravitational waves at relatively high frequencies, between 1 and 1000 Hz. The *Beyond Einstein* Great Observatory LISA will be sensitive in a broad band at much lower frequencies, between 10^{-4} and 0.1 Hz. It will detect entirely different sources, in great numbers, and with exquisite precision.

With LISA we will observe the coalescence of the supermassive black holes at the centers of merging galaxies; the radiation from thousands of individual ultra-compact binaries in our galaxy; the infall of small black holes, neutrons stars, and white dwarfs into the supermassive black holes at galactic centers. These sources will provide rich astrophysical information about the evolution of galaxies and stars, and they will serve as unique laboratories for precision measurements that may lead to new insights into fundamental physics and cosmology. In addition to these well-known astrophysical sources, LISA will search for gravitational waves from the very early Universe, as well as Big Bang remnants such as cosmic superstrings.

What are Gravitational Waves?

Electromagnetic (EM) waves are self-sustaining oscillations of the electric and magnetic fields, propagating through spacetime. By contrast, gravitational waves (GWs) are oscillations of spacetime itself (see Thorne 1987 and Flanagan & Hughes 2005 for reviews). Einstein predicted GWs shortly after developing his theory of general relativity (GR), but the first experimental verification of their existence had to wait over 60 years, until the binary pulsar observations by Hulse and Taylor starting in 1974 (see sidebar).

According to GR, GWs propagate at the speed of light, acting *tidally* by stretching and squeezing any extended distribution of matter or energy through which they pass. This warping action is transverse to the direction of wave propagation. GWs contain two dynamical degrees of freedom, which can be identified in the "+" (plus) and " $\times$" (cross) polarizations, corresponding to the axes associated with the stretching and squeezing. For instance, as depicted in Figure 1-1, a pure "+" polarization squeezes along the x axis and stretches along the y axis, and then vice-versa one half-cycle later.

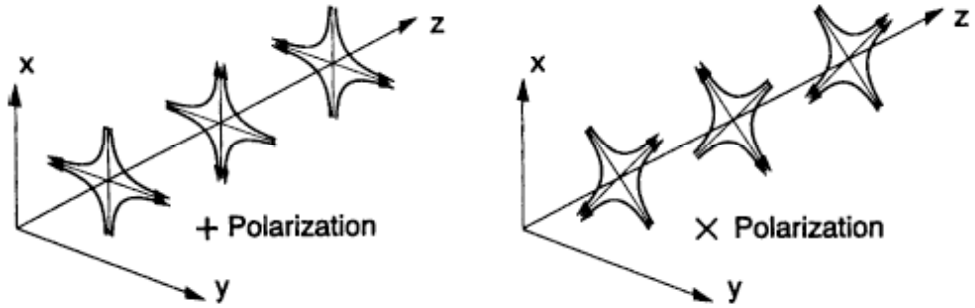
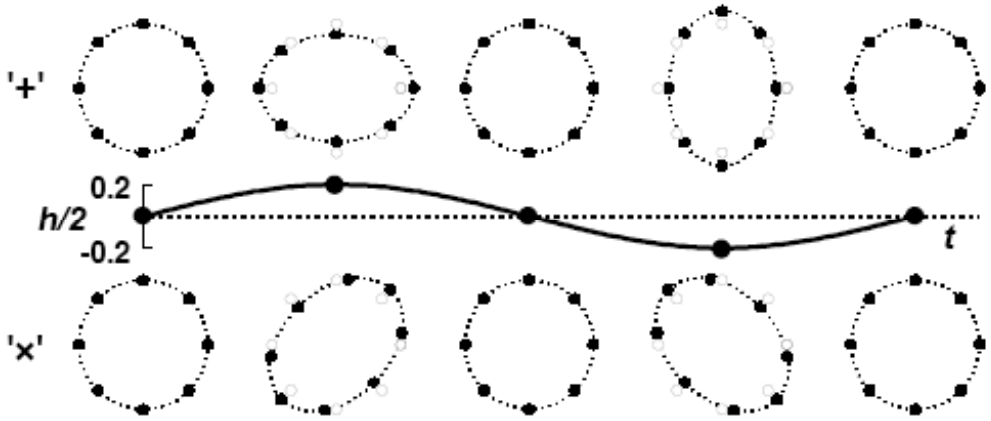


Figure 1-1: The effect of linearly polarized gravitational waves is to alternately stretch and squeeze the intervening matter and energy in perpendicular directions, as visualized here by the motions of a set of freely-floating *test* particles (*i.e.*, bodies small enough that their own gravitational field is negligible).



Just as EM waves are generated by accelerated charges, GWs are generated by accelerated masses. Because of charge conservation, an oscillating charge dipole is the lowest-order time-dependent distribution that can produce EM waves; because of mass and momentum (*i.e.*, mass dipole) conservation, a variable mass *quadrupole* is needed to produce GWs. (Technically it is the second time derivative of the transverse-traceless part of the quadrupole moment that generates GWs.)

EM waves arise from the interactions of atoms, nuclei, or other particles within astrophysical sources. EM waves are typically generated in numerous individual emitting volumes, much smaller than the astrophysical object of interest, so the wavelength of radiation is also much smaller than the object. For this reason, EM waves permit us to image the object if it is close enough or big enough. But the short wavelength has a disadvantage: we typically receive an incoherent superposition of radiation from many independent regions in the source, and if the source is not close enough to resolve then it is often a difficult and uncertain job to model the emission process well enough to go from the information we get about many different wavelength-scale regions up to the much larger scale of the entire astrophysical system.

By contrast, GWs are generated by the bulk mass distribution of the objects, so the wavelength of GW radiation is typically comparable to or larger than the size of the entire emitting region. (For instance, for two black holes orbiting each other and losing energy by gravitational radiation, the wavelength of the GW radiation is 10 or 20 times the radius of the orbit.) Thus, GW observations do not generally allow imaging, and the extraction of information from waveforms proceeds with audio-like methods such as time–frequency analysis. Because GW radiation is emitted coherently from the entirety of the astrophysical object, it provides *direct* information about the object's large-scale structure. Moreover, GW observations allow us to extract information from the *phase* of the wave as well as its amplitude or intensity. The phase evolution often carries more information about the detailed dynamics of the emitter than the amplitude does.

An astronomical observatory for GWs has very different characteristics from observatories of EM radiation. It will help to understand the range and versatility of LISA's observations as described in this document if we make some of these differences explicit.

First, LISA has a large intrinsic dynamic range: it could in principle measure accurately signals over an amplitude range of 10^5 or an energy range of 10^{10}. This is because it measures tiny changes in separations between the proof masses and therefore always operates in a linear regime. LISA is designed to study sources well below its mean noise level (extracted by matched filtering) up to the strongest expected sources in the Universe.

Second, LISA has a very large frequency range, spanning four decades, limited by its size and by the difficulty of isolating the proof masses at low frequencies. This means that, unlike optical, ultraviolet, or infrared observatories, LISA is less likely to miss distant sources because they are cosmologically redshifted to lower frequencies: it will be able to study homogeneously populations of objects out to the highest redshifts.

Third, LISA has all-sky acceptance of signals; it sweeps three different quadrupolar antenna patterns across the sky as it orbits the Sun, so that its sensitivity for all but the shortest transient sources is fairly isotropic. Unlike any imaging EM observatory, LISA is not pointed, so it does not miss any signals if they are above its noise level. This is particularly important because it allows LISA to detect strong transient events like black hole coalescences without having to point at the source; the strongest events in the Universe are necessarily transient, because they radiate far too much energy to be sustained in a steady state. Although LISA is an all-sky detector, it can nevertheless reconstruct event positions through its data analysis, by using phase modulation (Doppler effects) and amplitude modulation, which are available to it because it makes coherent observations. LISA can also separate thousands of simultaneously superimposed signals because it uses phase information to resolve them, so its all-sky acceptance does not lead to confusion except where there are really huge numbers of sources.

Finally, because GWs interact very weakly with matter, LISA is not troubled by absorption, scattering, or obscuration in any of its observations. The best illustration of this is in its search for a cosmological background of radiation from inflation: it can in principle see right back to the end of the inflationary epoch, through all stages of decoupling, symmetry breaking, and particle creation. GWs will give us our deepest views of the interiors of very dense environments, our

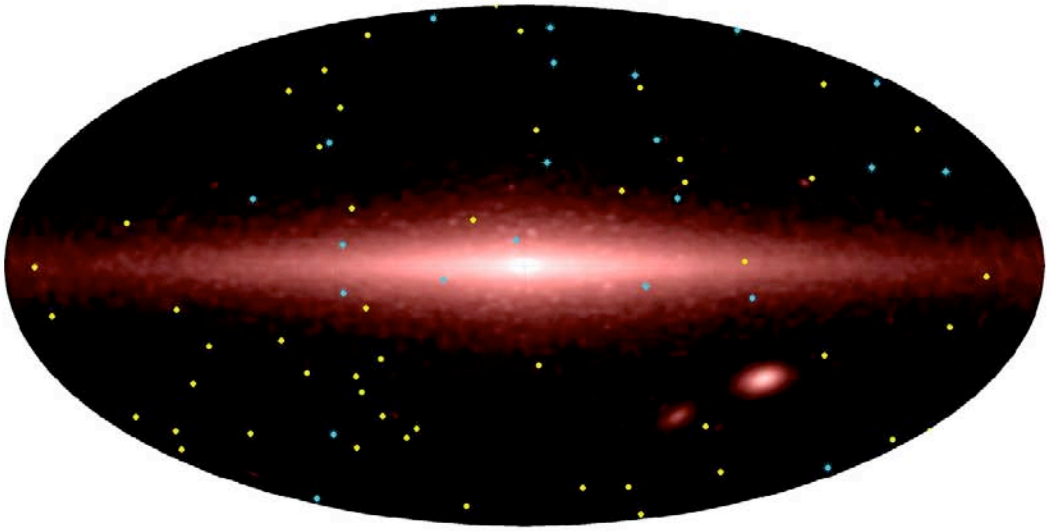


Figure 1-2: Simulation of the gravitational wave sky in the LISA band. The plane of the Galaxy is clearly visible as a horizontal band of emission from the population of compact white dwarf binaries. Also visible are the Magellanic clouds. Massive black hole mergers and capture of compact objects by massive black holes, which are temporary events, are indicated by yellow

only direct information about black holes, leading to their unequivocal identification and our first possibility to observe any possible structures in the electrically uncharged dark matter in the Universe.

We also note that the GW observable is the amplitude of the wave (or *wave strain*) h, the fractional amount of the stretching and squeezing discussed above. The amplitude h falls off only as $1/r$. A consequence is that strong sources of GWs (such as the binary inspirals of massive black holes, which LISA will see with amplitude signal-to-noise ratios ~ 1000 at a redshift $z = 1$), can essentially be detected out to arbitrarily large redshifts.

The distinctive characteristics of GWs ensure that they will provide a unique new channel to study the Universe, complementing information gathered over decades from EM channels, and probing previously inaccessible dense and dark regions of the Universe. The potential for discovery and surprise is great.

The Gravitational-Wave Universe in the LISA band

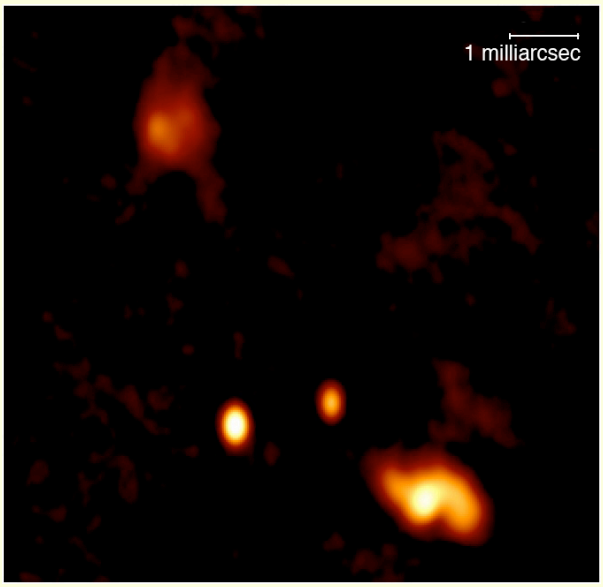
Although GWs have not been detected yet, we know enough about the contents of the Universe to make reasonably accurate guesses about some of the GW sources that LISA will observe (see Hughes 2003 & 2006 for reviews). As discussed earlier, in the same way that accelerated electric charges generate EM radiation, accelerated mass and energy of any kind generate gravitational radiation. The periodic motion of a system of mass M and size R creates GWs at a distance D with a strain amplitude of about $h \approx \left(GM/(Rc^2)\right)^2 (R/D)$, with frequency determined by the frequency of the motion. The shapes and strengths of the observed waves give us details about the structure and behavior of the system that produced them.

The strongest waves are generated by the systems with the largest gravitational fields GM/R, which correspond to large masses and small sizes. The strongest of all are generated by the interactions of black holes, which have $GM/Rc^2 \approx 1$. The lightest black holes (remnants of single stars, with about ten times the mass of the Sun) emit at the highest frequencies, in the 100-Hz band accessible to ground-based detectors.

By contrast, the strongest sources in the far lower LISA band (between 0.1 to 100 mHz) are the *super*massive black holes (SMBHs)

A supermassive black hole binary with parsec-scale orbital separation

Massive black hole binaries with parsec-scale orbital separation are expected to be quite common in the Universe. Since however the corresponding angular separations are very small, it is hard to image the binary components separately, and until recently the only known massive black-hole binaries had kiloparsec-scale separation. Recently, radio imaging with the Very Long Baseline Array has revealed a supermassive black-hole binary (0402+379) with an angular separation of just 7 milliarcseconds, corresponding to a projected spatial separation of 7.3 parsecs (Rodriguez *et al.* 2006). The binary sits in an elliptical galaxy, and its total mass is estimated to be about 10^8 Solar masses. The radio map of this amazing system is shown in the figure and consists of two compact cores associated with the two supermassive black holes, as well as two extended radio lobes. The radio lobes are estimated to be only about 1000 years old, indicating very recent activity associated with one of the two black holes. While this particular system is not necessarily expected to coalesce in a Hubble time because of its large mass, it demonstrates the existence of binary black holes resembling the progenitors of the coalescing systems that will be observed by LISA. There are so many galaxies in the observable volume that mergers are estimated to occur about once or twice a week, and signals from multiple binaries will be observed at any given time.



15 GHz VLBA image of 0402+379 (courtesy of Greg Taylor)

at the centers of galaxies; these are the remnants of the process of galaxy formation, with about 10^4 to 10^7 times the mass of the Sun. Optical, radio, and X-ray astronomy have produced abundant evidence that nearly all galaxies have massive black holes in their central nucleus, and that some of them even have two (see sidebar).

Mergers of SMBHs happen frequently: galaxies are continually forming, in a hierarchical fashion, from the mergers of smaller galaxies, and whenever two galaxies merge their central black holes sink to the center and find each other. SMBH mergers are so powerful that LISA can see them out to a wide range of redshifts, extending back to the first protogalaxies at $z \approx 15$. Estimates from standard galaxy formation theory suggest that LISA will detect SMBH mergers about once or twice every week (Volonteri 2006), but predictions are very uncertain at the high-redshift end, which sits beyond the reach of EM observations. LISA will lift the veil of these cosmic "dark ages", providing a direct record of the history of galaxy formation and central black hole growth in the observable Universe.

Smaller galactic objects can also be captured (and eventually consumed) by the central black hole. Compact objects such as degenerate dwarfs, neutron stars, and black holes will sometimes be driven by chance encounters into a close orbit around the massive black hole: a dance of death that they will repeat many times until they finally plunge into the black hole event horizon. The GWs from these extreme–mass-ratio inspirals (EMRIs) encode a detailed map of spacetime geometry around the SMBH. The history and environment of the black hole leave no imprint on this geometry, which is a very pure and beautiful solution (the *Kerr metric*) of the equations of general relativity. Thus, EMRI signals will test Einstein's theory by probing the most accurately predicted structures in all of astrophysics, whose mathematical elegance Chandrasekhar once likened to that of atoms.

In addition to black holes, many other known systems in our Universe can produce GWs in LISA's frequency band. Soon after it is turned on, LISA will quickly detect a handful of nearby binary stars, which have known periods and positions (and even assigned names), and which will appear in the LISA data with predictable, distinctive signatures.

A numerous Galactic population of *undiscovered* degenerate-dwarf binaries will be observed all across the LISA band; we know that such objects exists from EM observations in our vicinity, but LISA will detect thousands of individual binaries throughout the Galaxy. At low frequencies millions binaries more from across the Galaxy will blend together into a confusion background in the LISA data which will nevertheless teach us about the statistics of their population. At higher frequencies, the binaries have more powerful signals, and are farther apart in frequency space, allowing LISA to characterize each individually. At high frequencies, LISA may also detect the background signal from the degenerate binaries in other distant galaxies, allowing us to place constraints on cosmic star formation rates.

Given that all forms of matter and energy couple to gravity, it seems likely that the Universe will treat LISA to yet other GW sources that we cannot anticipate on the basis of our EM observations. This is especially true for observations at very high redshifts, where LISA may give us the very first clues to the unknown conditions of matter and energy in the very early Universe.

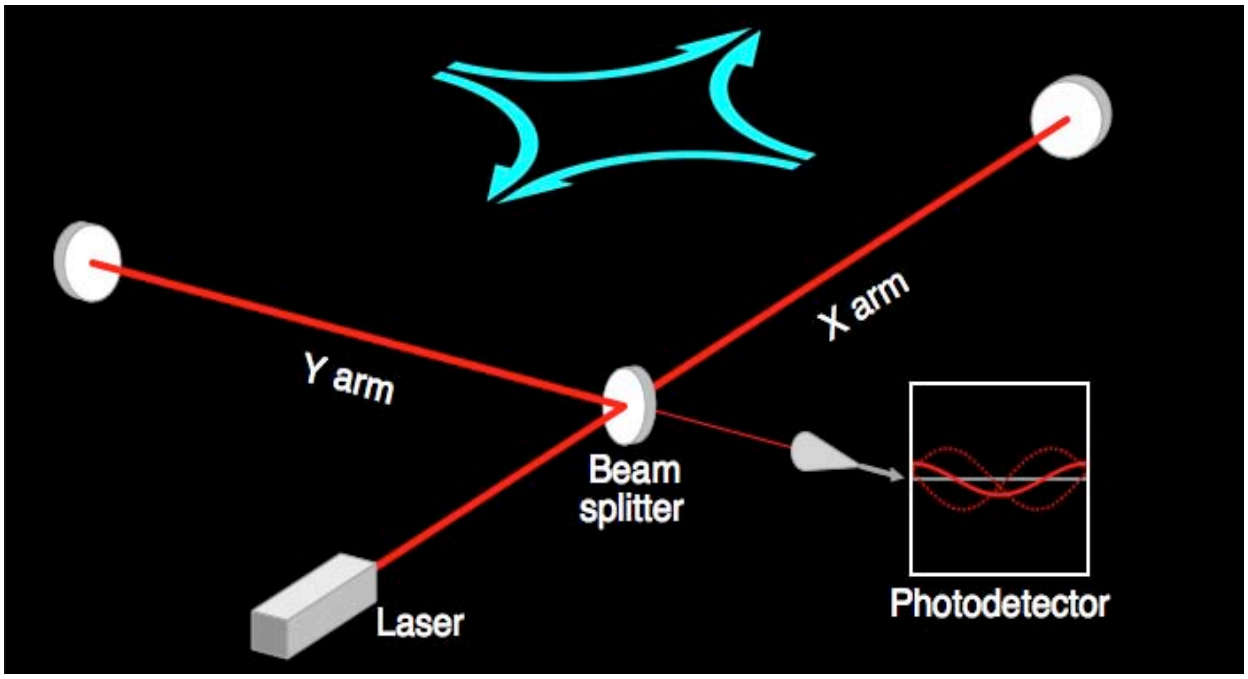


Figure 1-3: Idealized Michelson interferometer with laser, beam splitter, end mirrors, and photodetector at the exit port. the blue arrows represent a perpendicularly propagating polarized gravitational wave.

In the relativistic early Universe, the LISA frequency band corresponds to the *Terascale frontier*, where the phase transitions of exotic fields or extra spatial dimensions may have caused catastrophic and explosive bubble growth, with efficient GW production. LISA will also probe *superstrings*, relics of the early Universe predicted in some versions of string theory. These exotic structures, which are completely invisible except for the GWs they emit, could produce strong, distinctive LISA signatures; they could provide direct evidence that all forms of matter and energy, and possibly even spacetime itself, are ultimately made of quantum strings.

How are Gravitational Waves Detected?

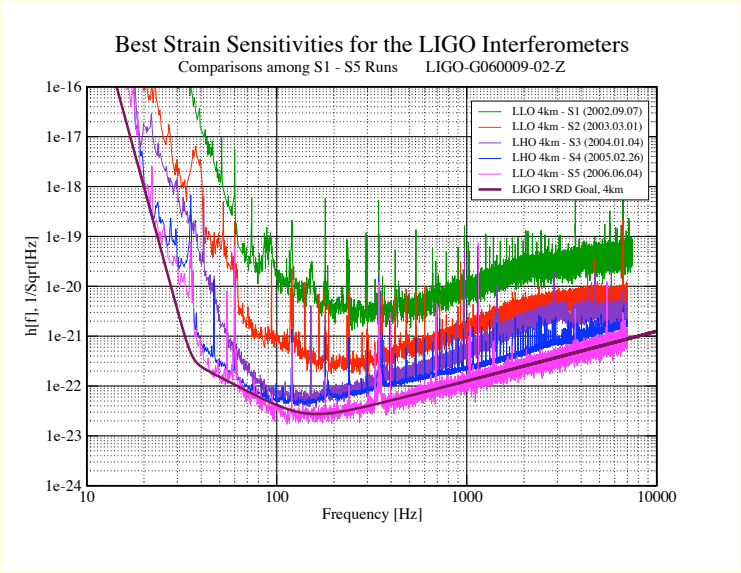
Einstein's great epiphany was that gravity is the manifestation of the curvature of spacetime, the background for all the interactions of matter and energy. Freely-falling *test* bodies (small compared to spacetime curvature and undisturbed by other forces) thread spacetime along *geodesics*, the straightest paths possible through this curved arena. Nearby, approximately parallel geodesics are pushed together and pulled apart by spacetime curvature. Indeed, GWs are waves of spacetime curvature, experienced by test bodies as an oscillating change in their relative distance.

To understand how this principle is used to detect GWs, it is useful to visualize an idealized Michelson laser interferometer (see Figure 1-3) whose components are floating freely in space, at

LIGO reaches its design sensitivity

The absolute accuracies required by GW observations may at first seem unachievable. The LIGO design sensitivity (black line) implies measurements of the change in distance between interferometer mirrors that are accurate to better than one part in 10^{22} in the relevant frequency band. Yet, as shown in this figure by the progression of experimental noise curves for the LIGO Hanford 4 km interferometer (colored lines), the dedication and ingenuity of countless scientists and engineers has made these measurements possible.

Because of LISA's much longer arm length, the metrology accuracy required locally at each spacecraft is much less exacting than the LIGO requirement, by a factor of about one half million .



rest with each other, and far removed from any gravitating bodies. The power measured by a photodetector at the exit port of the interferometer is a simple function of the phase difference of the two light beams that are divided at the beamsplitter, propagated across the two arms, and recombined at the exit port. Incoming GWs (consider for simplicity a plane GW, propagating perpendicularly to the plane of the interferometer, with "+" polarization aligned with the two arms) alternately increase the distance experienced by light traveling along one arm and decrease the distance along the other, creating oscillations in the power measured at the exit port.

The currently operating ground-based GW detectors such as LIGO (Abramovici *et al.* 1992) implement much more sophisticated versions of this setup (for instance, in LIGO light is made to bounce several times along cavities in each arm, enhancing the phase-accumulation effect). The free fall of interferometer components is approximated, along the "sensitive" axes, by suspending them from pendulums that are carefully isolated from the environment.

The sensitivity and frequency range of ground-based interferometers is set by their size (with armlength ranging from a few hundred meters to 4 km for LIGO), and by various noise sources, including photon shot noise in the phase-difference measurement, and thermal noise in the suspensions and optical components. To detect an amplitude of $h \approx 10^{-22}$ LIGO must measure displacements of the order of 10^{-19} m, an extremely challenging target. At frequencies below 10 Hz, seismic noise from terrestrial sources is especially challenging to the mechanical isolation systems. Below about 1 Hz, another terrestrial noise begins to be larger than expected astronomical signals: gravity-gradient noise from time-dependent changes in the local Newtonian gravitational field. No GW detector can be isolated from these tidal gravitational fields, so that detection in the LISA frequency band requires a detector far from the Earth.

LISA brings the Michelson measurement concept to the grander scale and much quieter environment of space. While ground-based detectors are naturally limited by the Earth's curvature to km armlengths, LISA can be more than a million times larger. Longer arms magnify the effect of GWs, so LISA can achieve 10^{-22} strain accuracy by measuring displacements of the order of fractions of a picometer, a million times larger than those LIGO has to cope with. When this is combined with the absence of seismic and gravity-gradient noise, LISA can achieve remarkable sensitivity to the lower-frequency GWs emitted by larger, more massive systems. Thus, while advanced ground-based detectors may observe neutron stars or stellar-mass black holes coalescing out to distances of hundreds of megaparsecs, LISA will be sensitive to supermassive black hole coalescences out to redshifts greater than 10, allowing it to plumb the earliest stages of galaxy formation.

While LISA can be thought of as a "Michelson interferometer in space", its actual implementation is quite different from a conventional Michelson interferometer. The experimental approach is closer to that of spacecraft Doppler tracking, in which the observed quantity is the frequency change in the signal from a distant spacecraft. In LISA, each spacecraft sends a beam of laser light to each of the other two distant spacecraft and in turn receives a beam from each of the two spacecraft. The received laser light is coherently combined at a photodetector with the light from an onboard reference laser and the frequency difference is recorded as a beat signal. The beat signals recorded at each of the three spacecraft are delayed in time and recombined in a technique called "Time Delay Interferometry" (Tinto 2005), which essentially creates three virtual Michelson interferometers whose output signals represent the basic LISA science data stream.

2. LISA Mission Overview

The LISA mission uses three identical spacecraft whose positions mark the vertices of an equilateral triangle five million km on a side, in orbit around the Sun. LISA can be thought of as a giant Michelson interferometer in space, with a third arm that provides independent information on the two gravitational wave (GW) polarizations as well as redundancy. The spacecraft separation — the interferometer arm length — sets the range of GW frequencies LISA can observe (from about 0.1 mHz to above 0.1 Hz). This range was chosen to reveal some of the most interesting sources – mergers of supermassive black holes, ultra-compact binaries, and the inspirals of stellar-mass black holes into supermassive black holes. The center of the LISA triangle traces an orbit in the ecliptic plane, 1 AU from the Sun and 20° behind Earth, and the plane of the triangle is inclined at 60° to the ecliptic (see Figures 2-1 and 2-2). The natural free-fall orbits of the three spacecraft around the Sun maintain this triangular formation throughout the year, with the triangle appearing to rotate about its center once per year.

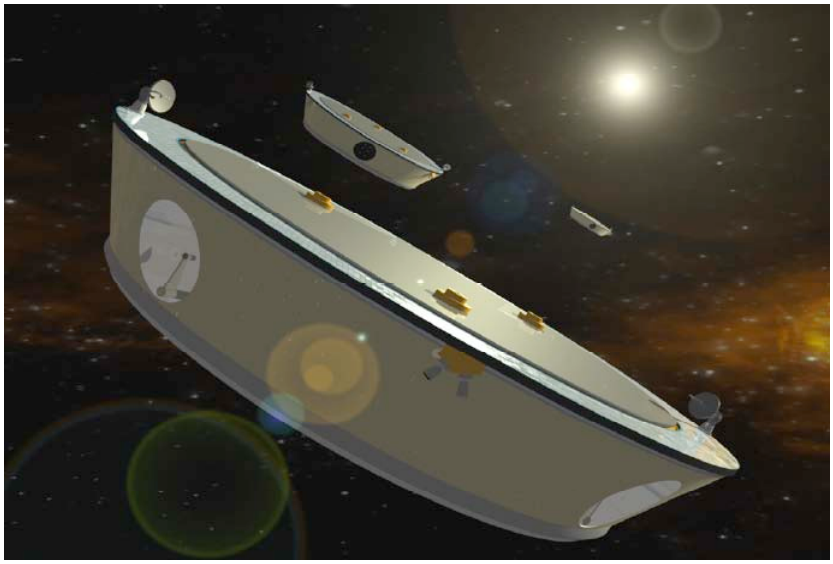


Figure 2-1: The LISA constellation of three spacecraft separated from one another by five million km (0.03 AU), with the Sun visible in the distance at 1 AU.

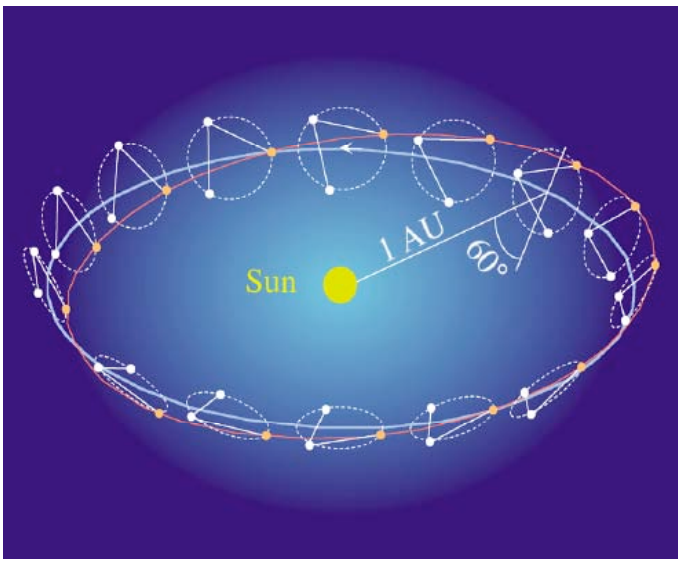


Figure 2-2: The LISA constellation's heliocentric 1-AU–radius orbit, with the plane of the triangle inclined at 60° to the ecliptic. The triangle appears to rotate once around its center in the course of a year's revolution around the Sun.

Rather than a physical Michelson interferometer, the actual implementation used for LISA resembles the technique known as spacecraft Doppler tracking, but realized with infrared laser light instead of radio waves. The laser light going out from one spacecraft to the other corners is not reflected back directly,

because diffraction losses over such long distances would be too great. Instead, in analogy with an RF-transponder scheme, the laser on the distant spacecraft is phase-locked to the incoming light and transmitted back at full intensity. When the transponded laser light arrives back at the original spacecraft, it is superposed with a portion of the original laser beam, which serves as the local oscillator in a standard heterodyne detection scheme. This relative phase measurement gives information about the length of that interferometer arm, modulo an integer number of wavelengths of the laser light. The difference between the phase measurements for the two arms gives information about the relative changes in the two arms — the gravitational-wave signal.

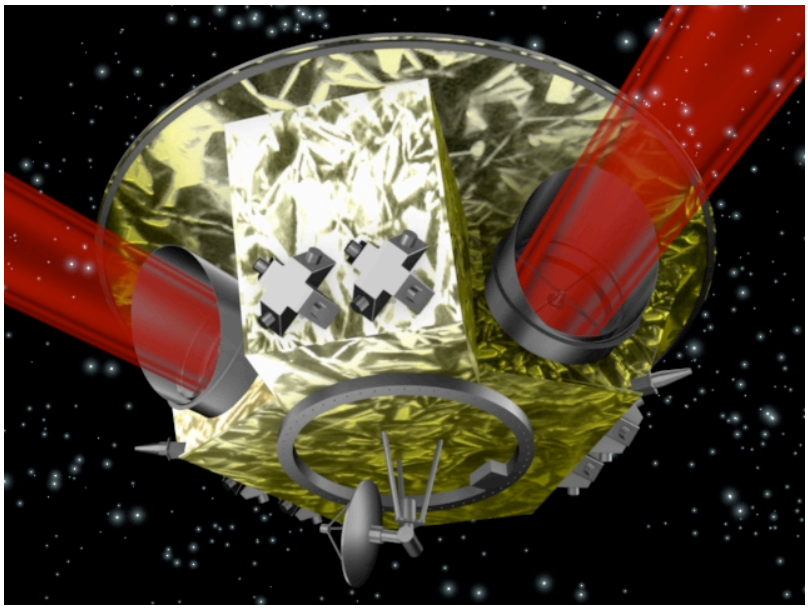


Figure 2-3: View of the exterior of one LISA spacecraft, showing the two telescope assemblies, two of the six thruster packages, and the telemetry antenna.

A two-arm interferometer can be prone to phase errors due to laser frequency fluctuations. If the arms were exactly equal in length, then laser frequency fluctuations would cancel perfectly in the arm length difference measurement. Unfortunately, annual variations in the LISA spacecraft orbits prevent perfect cancellation of laser noise. To minimize the measurement error from laser phase noise the lasers are frequency stabilized — first to an optical cavity, and then to the 5-million-km interferometer arm (Sheard *et al.* 2003, Herz 2005). Any residual laser frequency noise in the LISA measurements will be removed by post-processing on the ground using a technique called Time Delay Interferometry (TDI; Tinto *et al*. 2003, Tinto & Dhurandar 2005).

Each spacecraft contains a pair of optical assemblies oriented at roughly 60° to each other (see Figure 2-3). Each assembly is pointing toward a similar one on the corresponding distant spacecraft, to form a (non-orthogonal) Michelson interferometer. Through a 40-cm aperture telescope on each assembly, a laser beam from a 1.064-μm Nd:YAG master laser and 1-W Yb fiber amplifier is transmitted to the corresponding remote spacecraft. The same telescope is used to collect the very weak incoming beam (around 100 pW) from the distant spacecraft and direct it to a sensitive photodetector, where it is combined with a local-oscillator beam derived from the original local laser light. At the heart of each assembly is a vacuum enclosure containing a free-flying polished platinum-gold cube, 4 cm in size — the "proof mass" that serves as an inertial reference for the local optical assembly (see Figure 2-4). A passing gravitational wave will produce a relative strain in this large "optical truss," causing an increase in the optical path

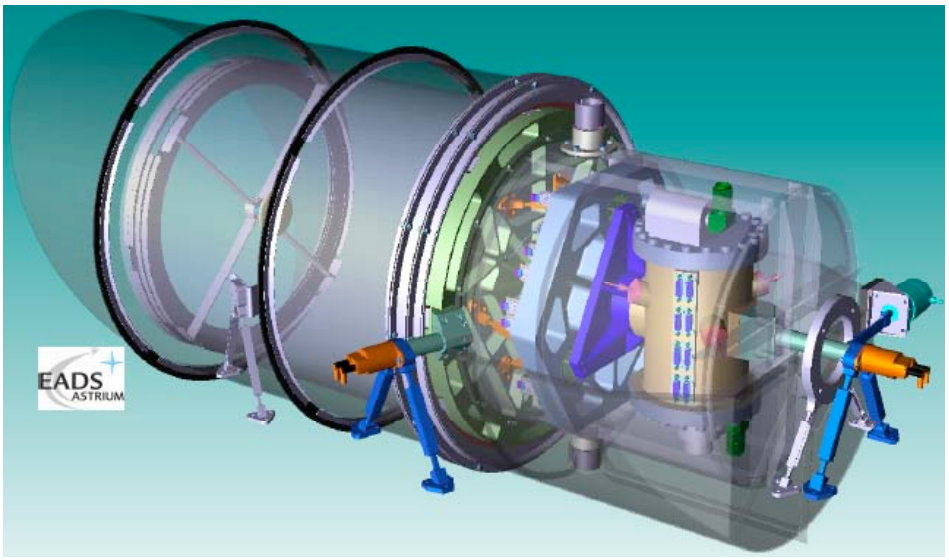


Figure 2-4: Optical assembly containing (from left to right) the telescope assembly, the optical bench, and the proof mass vacuum enclosure.

length between the proof masses forming one arm while causing a decrease for the other arm. These length changes are measured interferometrically with sub-Angstrom precision. In this way, LISA will be sensitive enough to detect gravitational-wave induced strains of amplitude $h = \Delta l / l \sim 10^{-23}$ in one year of observation, with a signal-to-noise ratio of 5.

The spacecraft surrounding each pair of optical assemblies serves primarily to shield the proof masses from the adverse effects of solar radiation pressure fluctuations; the spacecraft positions do not enter directly into the measurements. Nevertheless, in order to minimize disturbances to the proof masses from fluctuating forces in their vicinities, each spacecraft must be kept moderately centered around the proof masses (to about 10 nm/$\sqrt{\text{Hz}}$ in the measurement band). This is achieved by a “drag-free" control system based on small electric thrusters and displacement sensors. Specifically, both capacitive and optical sensors are used to measure the displacements and rotations of the proof masses relative to the spacecraft. These offset signals are then fed back to control micro-Newton thrusters, which force the spacecraft to follow its proof masses. The thrusters are also used to control the attitude of the spacecraft relative to the incoming optical wavefronts, using signals derived from quadrant photodiodes. Capacitive actuation is used to adjust the positions or orientations of the proof masses when needed.

Each of the three LISA spacecraft has a launch mass allocation of about 480 kg (including margin) including the payload, propulsion module, and the spacecraft adapter. The propulsion modules use a chemical propulsion system to transfer each spacecraft from the Earth orbit to the final position in interplanetary orbit. All three spacecraft can be launched together by a single Atlas V (531). Each spacecraft carries a small, steerable antenna used for transmitting the science and engineering data, in the Ka-band to the NASA Deep Space Network. The nominal mission lifetime is five years. Additional information about the LISA mission can be found in Hammesfahr (2001) and Cramer *et al.* (2003).

LISA Capabilities

Sensitivity

The sensitivity of LISA is shown in Figure 2-5 together with a comparable sensitivity curve for the future ground-based Advanced LIGO. Sensitivity over a logarithmic frequency interval is shown here in terms of the dimensionless strain $\sqrt{f}\sqrt{S_{\rm h}(f)}$ (where $\sqrt{S_{\rm h}(f)}$ is the 1-σ level of strain noise spectral amplitude), in order to facilitate comparison with the much higher measurement frequencies of LIGO. The LISA sensitivity curve divides into three regions: a low-frequency region where proof-mass acceleration noise dominates, a mid-frequency region where shot noise and optical-path measurement errors dominate, and a high-frequency region where the sensitivity curve rises as the wavelength of the GW becomes shorter than the LISA arm length. Additionally, a diffuse background of unresolved galactic binaries is expected to contribute to the measured strain level in the frequency range from 0.1 - 1 mHz; this component is indicated in Figure 2-5.

Figure 2-5 illustrates also the different astrophysical sources that LISA will study, contrasted with those studied by ground-based interferometers such as LIGO. In general, a ground-based

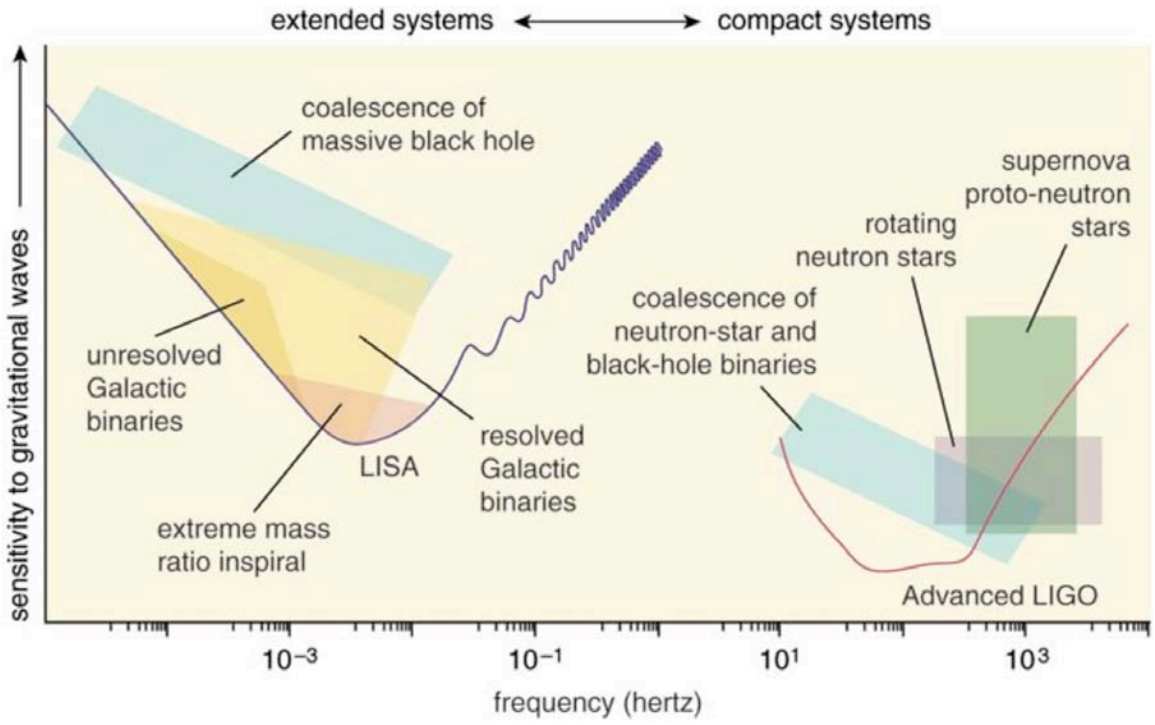


Figure 2-5: Strain sensitivities predicted for LISA and Advanced LIGO. The curve is plotted in units of ($\sqrt{fS_{\rm h}(f)}$), where $S_{\rm h}(f)$ is the LISA noise power spectrum with $\sqrt{S_h(f)}$ representing the 1-σ noise level. The units of ($\sqrt{fS_{\rm h}(f)}$) are the natural units for plotting sensitivity per logarithmic frequency interval.

interferometer is limited to frequencies above about 10 Hz because of seismic noise, corresponding to wavelengths less than about 30,000 km. The sources it studies will therefore generally be very compact sources, *e.g.*, with orbital separations comparable to the wavelength. Such sources include neutron-star binaries and stellar-mass black hole binaries just before merger. In contrast, LISA will see most of its sources in the frequency range below about 10 mHz, corresponding to wavelengths greater than about 30 million km. This is the characteristic size of systems involving massive black holes, as well as compact white-dwarf binaries.

Sky Position Determination

LISA can determine a source's position in the sky using techniques that are similar to those used in the radio frequency domain, namely AM and FM modulation. For higher frequencies, above about 1 mHz, LISA will use a technique similar to the localization of pulsars *via* pulsar timing, *i.e.*, observing the phase or Doppler shifts of the GW signal as a function of the position of LISA in its orbit about the Sun. At low frequencies the wavelengths are too long to do positioning using frequency modulation. Consequently, the position is determined from the modulation of the amplitude of the GW signal as the LISA constellation changes aspect in its orbit about the Sun. Both of these methods can provide sub-degree location accuracy for strong GW sources.

Distance Determination

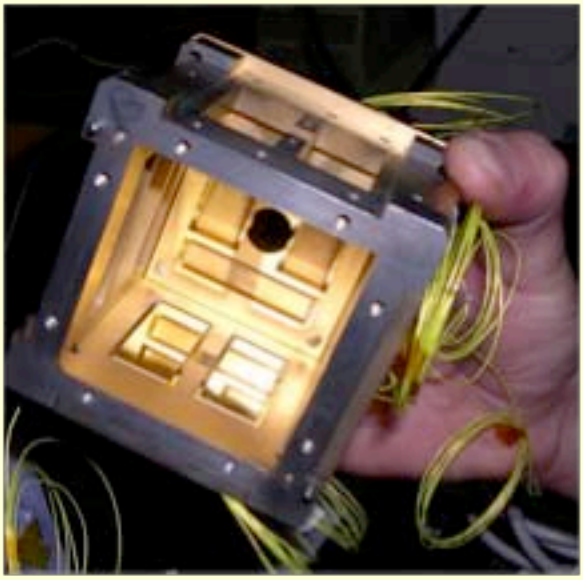
A unique feature of gravitational wave measurements of astrophysical systems is the capability of determining precise distances for many GW sources. A coalescing binary has a characteristic frequency of emission which is twice the orbital frequency and which increases in time due to the loss of energy to GW emission. The rate of increase of the frequency, *df/dt*, depends on the so-called "chirp mass", $M_{\rm chirp} = (1+z)(M_1 M_2)^{3/5}(M_1+M_2)^{1/5}$, where z is the redshift and M_1 and M_2 are the masses of the two members of the binary. The amplitude of the GW signal at the detector depends both on the chirp mass and on the distance. By determining the chirp mass from the rate of increase of the frequency the signal amplitude can then be used to determine the distance. LISA will typically be able to measure the distance to a coalescing massive black hole binary to a few-percent accuracy. In contrast, the redshift z is not independently determined from GW observations. Dimensionful system parameters such as mass and spin enter the system's dynamical evolution as timescales. For distant sources, these timescales acquire a redshift factor $(1 + z)$; as a consequence, the inferred system parameters likewise acquire factors of $(1 + z)$. Redshift is always degenerate with a measured system's intrinsic parameters; it must be determined by associating a merging system with a host galaxy. We thus have the situation that for electromagnetic (EM) observations distance determination is typically difficult, but redshift determination reasonable straightforward, while the opposite is true for GW observations. The excellent distance determination capability of LISA is extremely

LISA Pathfinder: Flight Test of Key LISA Technologies

The European Space Agency is preparing a flight validation of the LISA drag-free and interferometry systems on the LISA Pathfinder (LPF) mission, planned for 2011. LPF will provide validation and risk reduction for key LISA technologies by carrying two drag-free proof masses and an interferometer to measure their relative displacements using heterodyne techniques as planned for LISA. Essentially, LPF shrinks one LISA arm down to fit inside a single spacecraft; it thereby loses its sensitivity to gravitational waves but tests the LISA sensitivity to relative proof-mass displacements.

The picometer displacement sensitivity required of LISA's interferometer is not particularly challenging: it is about one-million times more relaxed than LIGO currently achieves on the ground, and it is only a thousand times more stringent than Albert Michelson achieved in his original interferometer 125 years ago. LPF will demonstrate that this can be achieved in space, but the primary reason for launching LPF is to validate the drag-free system. LPF will demonstrate a proof-mass assembly that can survive launch vibration and still provide the required degree of isolation from external disturbances in space. LPF will also carry two independent microthruster hardware and control systems, one supplied by ESA and the other by NASA. It will provide an opportunity to compare the stability, reliability, and precision of these key components of the drag-free system.

The LPF design, incorporating any lessons learned from the mission, will form the basis of the LISA drag-free and interferometry systems. ESA plans to give LISA the final go-ahead for Phase C/D once the results of LPF have been evaluated.



The LISA pathfinder engineering model of the drag-free sensor housing, showing the capacitive sensors.

important for determining the number density of BH mergers versus epoch, for possible H_0 and dark energy measurements, and for determining the distribution of ultra-compact binaries in the Galaxy.

Science Operations

LISA will deliver a rich and diverse scientific data set providing numerous opportunities for guest investigations associated with LISA observations. Opportunities will include: complementary EM observations during flight and pre- and post-flight, development of analysis techniques for LISA data, analysis of the LISA data for detection and characterization of sources of gravitational waves, and use of LISA generated source lists, waveforms, catalogues for physics and astrophysics research investigations and theoretical investigations. These are discussed further in a later section.

The starting point for most science analysis of LISA data will be the construction of three science data channels, the so-called "Time-Delay Interferometry (TDI) variables". A full discussion of the TDI data is beyond the scope of this

document and further information can be found in Tinto & Dhurandar (2005). Basically, the three TDI data channels can be thought of as containing the information about the "+" and "×" gravitational wave polarizations, as well as a third channel of information, the "Symmetrized Sagnac" channel, which is relatively insensitive to gravitational waves and can be used as a monitor of LISA instrumental noise, very important for LISA measurements of the diffuse GW background.

LISA data are expected to contain the signals from many sources simultaneously. Well-developed techniques exist to detect and characterize multiple sources. Thus, in addition to the calibrated TDI channels, a data product will be produced that will be a "model" for the LISA data providing at any time the best estimate for the source content of the LISA data, including waveform estimates, effectively providing three source-subtracted LISA data channels for use by researchers interested in searching for new sources.

In addition to the actual scientific data channels, LISA will provide "alerts" of specific events, in particular predictions of merger times and locations of massive black hole binaries as well as catalogues and ephemerides for detected sources: extra-galactic black hole binaries, ultra-compact galactic binaries, inspiral of stellar-mass compact objects into massive black holes in external galaxies, *etc*. These data products will provide a wealth of opportunities for follow-up ground- and space-based astronomical facilities as well as opportunities for theoretical research.

Status of LISA Data Analysis

There is already a robust community effort in development and demonstration of LISA data analysis including publication of over 250 papers. LISA has an active program of data simulation and "mock data challenges" to refine the understanding of the science capabilities of LISA (Arnaud *et al.* 2006). Techniques have been developed for detection and parameter estimation for all classes of LISA sources: massive binary black hole coalescence, galactic white-dwarf binaries, extreme mass ratio inspirals, and stochastic background. These techniques have a rich heritage, not only from ground-based gravitational wave detectors, but also from related analysis fields such as sonar, radar, seismology, radio astronomy, and voice recognition. Activity is underway to perform detailed simulations and demonstrations for the various sources and analysis techniques.

LISA mission quick facts

LISA Pathfinder: The joint ESA-NASA LISA Pathfinder/ST-7 mission will demonstrate the core technologies needed for both LISA's interferometers and for the 'drag-free' spacecraft control that protects LISA's proof masses from disturbance.

LISA: The Laser Interferometer Space Antenna is a joint mission of NASA and ESA. Specifications below are per the baseline of design of 2006 but are subject to change.

Launch and cruise: A single Atlas-V rocket lifts all three spacecraft, each with its own propulsion module. They reach their final solar orbits after a 13-month cruise.

Orbits: Each drag-free spacecraft is in an independent heliocentric orbit, trailing Earth by about 20 degrees.

Station-keeping: There is no station-keeping. The orbits are chosen to keep the spacecraft at the corners of a nearly equilateral triangle throughout the mission life.

Inter-spacecraft separation: About 5 million kilometers (0.03 Earth-Sun distance, or 17 light-seconds).

Spacecraft:
Size: 2.7-m diameter, 0.7-m thick
Mass: 575 kg
Orientation: 60 degrees to the Sun, constant.
Power: 820W per spacecraft, from solar cells.
Attitude and drag-free control: six micro-Newton thrusters, 4 to 30 μN each
Telemetry: X-band, DSN, under 10 kbps

Payload (per spacecraft):
Lasers (two): 1-W diode-pumped 1064-nm Nd:YAG lasers, frequency-stabilized to an onboard reference cavity and also to the inter-spacecraft arms.

Proof masses (two): 2-kg Au-Pt cubical proof masses. Electrostatic actuation perpendicular to the measurement axes. Charge control by UV illumination.

Telescopes (two): 40-cm diameter, f/12, used both to transmit and receive.

Measurement and error budget:

Measurement: Optical heterodyne: The received laser beam (about 10^{-10} W) is mixed with about 1mW of local laser light on a quadrant photodiode. Measurements from all three pairs of spacecraft are combined using time-delay interferometry to synthesis two independent Michelson interferometers and a third independent Sagnac interferometer.

Optical-path errors: Caused by detector shot noise, pointing jitter etc., total not to exceed 2 x 10^{-11} m/√(Hz) over the LISA bandwidth.

Acceleration errors: Caused by unmeasured residual forces on proof masses. Total not to exceed 3 x 10^{-15} $(m/s^2)/\sqrt{(Hz)}$ from 0.1 to 1 mHz. Divide by $(2\pi f)^2$ for resulting displacement error.

Strain sensitivity: Primary sensitivity to gravitational waves of frequency 0.03 mHz – 0.1 Hz.

Source localization in space: Below 1 mHz or lifetime < months: degrees
Above 10 mHz, lifetime > months: arcminutes

3. Black Hole Astrophysics: Supermassive Black Holes in Galactic Nuclei

There is now abundant evidence that nearly all galaxies have massive black holes in their central nuclei. These nuclear black holes can have profound effects on galaxy formation, releasing huge amounts of energy into the galaxy *via* accretion-powered jets. The formation of this population of massive black holes is thought to be associated with a multistage process of binary inspiral and merger, together with accretion. LISA will detect the merger events directly, thus tracing the growth and merger history of massive black holes and their host galaxies. LISA will search for a population of seed black holes at early epochs and use precise measurements of black-hole spins to help determine the relative importance of black-hole growth mechanisms. LISA will also probe the rich astrophysics in the nuclei of normal galaxies by observing the inspiral of compact objects into the massive black holes in their centers.

Key science questions

- When did the massive black holes in galactic nuclei form, and how did they grow?
- What fraction of proto-galaxies contained massive black holes in their cores, as a function of redshift?
- What are the mass and spin distributions of the supermassive black holes in galactic nuclei?
- What is the role of black hole mergers in early hierarchical structure assembly?
- What are dynamics of stars near massive black holes in galactic nuclei?

Evidence for supermassive black holes in galactic nuclei

Supermassive black holes accreting gas in galactic nuclei were first proposed in the 1960s (Salpeter 1964; Zel'dovich & Novikov 1964) to explain the enormous luminosities of the newly discovered quasars. Refinements of this idea have become the generally accepted explanation for the electromagnetic and kinetic emissions from all active galactic nuclei (AGN; see, *e.g.*, Krolik 1999).

The disks of gas around accreting black holes (mass $M_\bullet$) in active galaxies are inferred to have luminosities approaching the Eddington limit (at which radiation pressure on the Thomson cross section σ_T of electrons balances the gravitational attraction on them and the protons from which they were stripped):

$$L_{\rm Edd} = 4GcM_\bullet m_p/\sigma_T = 10^{46}\ {\rm erg\ s^{-1}}(M_\bullet/10^8 M_\odot)\ .$$

Thus, the black holes in AGN, producing radiation from accreted rest-mass with efficiency $\varepsilon \equiv 0.1\varepsilon_{0.1}$ must be increasing in mass by accretion with an e-folding time known as the Salpeter time,

$$t_S = \varepsilon M_\bullet c^2/L_{\rm Edd} = 4\times 10^7 \varepsilon_{0.1}\ {\rm yr}\ .$$

Since t_s is about one percent of the age of the Universe at redshift $z = 1$, at most about one percent of black holes can be radiating at near the Eddington limit; the rest must be quiescent. This crude estimate is consistent with observations: the deepest images from the Hubble Space Telescope (HST) show nearly 10^6 galaxies per square degree. Only about one percent have the variable nuclei characteristic of AGN (Cohen *et al.* 2006). In its longest exposures, the Chandra X-ray observatory also detects about 10^4 active galactic nuclei in each square degree of sky (Brandt & Hasinger 2005). We might expect, then, that the deep Universe is populated by billions of invisible black holes, representing 99% of the total population. Gravitational wave observatories such as LISA can reveal these hidden objects.

As Lynden-Bell pointed out, another consequence of the ~ 1% "duty cycle" is that a large fraction of local galaxies must have been quasars in their youth, and must today harbor relic black holes in their nuclei (Lynden-Bell 1969). A 1% duty cycle could mean that 1% of galaxies host a black hole that is active for most of its life or, at the other extreme, that most galaxies host black holes that are active only 1% of the time. The latter interpretation has recently been confirmed spectacularly by the discovery that in the centers of almost all bright nearby galaxies (including our own Milky Way) the velocities of stars and gas begin to rise in the Keplerian fashion expected if there were a central point mass dominating the central potential (Ferrarese & Ford 2005; see Figure 3-1).

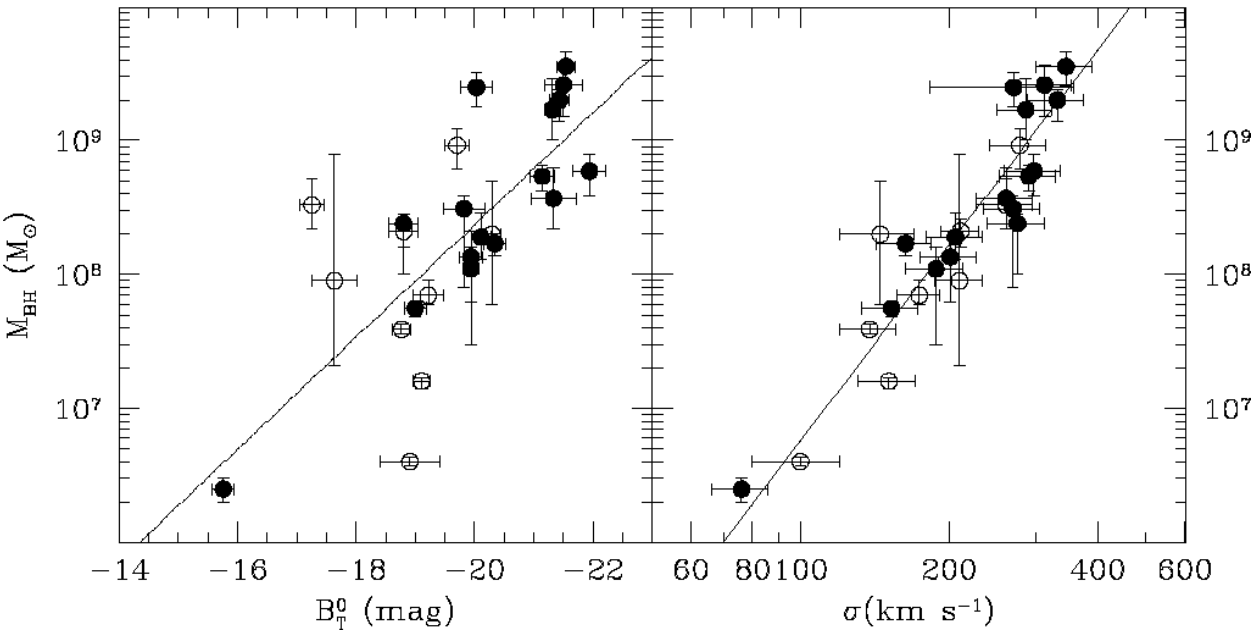


Figure 3-1: The correlation between black hole mass *M*• and the luminosity of the host galaxy's stellar bulge (left) and between *M*• and the host galaxy's bulge velocity dispersion σ (right) for all detections in galaxies near enough that current instruments resolve the region in which the black hole dominates the dynamics. Filled symbols show elliptical galaxies; open symbols show spiral and lenticular galaxies (from Ferrarese & Ford 2005; see also Gebhardt *et al.* 2000).

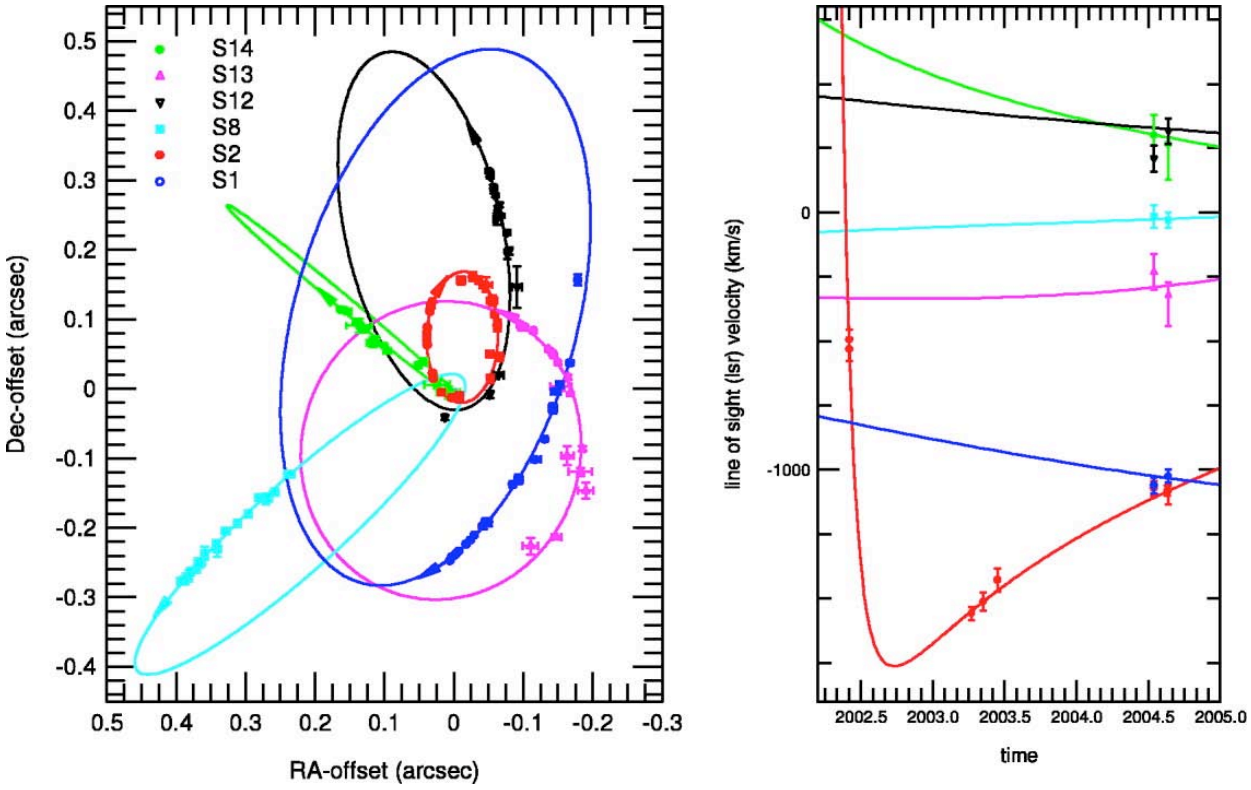


Figure 3-2: The orbits on the sky, and radial velocity measurements of six stars close to the central point mass in the Milky Way. At their peribothrons, stars S2 and S14 pass within 5×10^{-4} pc = 100 AU = 1300 Schwarzschild radii of the central black hole, of mass $4 \times 10^6 M_\odot$ (from Eisenhauer *et al.* 2005).

Hubble Space Telescope press releases consistently call these central masses black holes, and we follow that practice here. But one of the most important questions LISA will address is whether these masses are actually the Kerr black holes of Einstein's relativity. In almost all cases, current measurements cannot probe regions closer than $\sim 10^5$ times the Schwarzschild radius of the inferred black hole, so other astrophysical models (*e.g.*, dense clusters of stellar-mass black holes) are not conclusively ruled out. But in the notable case of our own Milky Way, one can measure the orbits of stars passing as close as 1300 Schwarzschild radii from the $4 \times 10^6 M_\odot$ central mass (Ghez *et al.* 2005; Eisenhauer *et al.* 2005), and one can rule out even contrived astrophysical alternatives to black holes (Maoz 1998),

Massive or Supermassive?

The terms "massive black hole" and "supermassive black hole" often are used almost interchangeably to refer to the massive objects in the centers of galaxies. In this document we generally will use the following definitions:

Supermassive Black Hole (SMBH): $10^7 M_\odot < M_{BH}$

Massive Black Hole (MBH): $10^4 M_\odot < M_{BH} < 10^7 M_\odot$

Intermediate Mass Black Hole (IMBH): $10^2 M_\odot < M_{BH} < 10^4 M_\odot$

though not exotic models invoking new physics such as soliton stars with radii of less than a few Schwarzschild radii. See Figure 3-2.

The black hole mass estimates derived from the kinematics of stars and gas in the nuclei of nearby representative galaxies also allow one to estimate the space density of local black holes (Figure 3-1). These estimates give

$$\rho_\bullet = 2\text{-}5 \times 10^5 M_\odot \ \mathrm{Mpc}^{-3}$$

(for $H_0 = 70$ km s^{-1} Mpc^{-1}; Aller & Richstone 2002; Marconi *et al.* 2004). The uncertainty in $\rho_\bullet$ results principally from uncertainties and dispersion in the correlations between black hole mass and galaxy luminosity or between black hole mass and velocity dispersion (*c.f.* Lauer *et al.* 2006, Tundo *et al.* 2006).

Figure 3-3: Snapshots from a simulation of a collision of two spiral galaxies, similar to the Milky Way, containing $10^5 M_\odot$ seed black holes at their centers. The images show only the gas in the galaxies. Color indicates temperature and brightness gas density. The collision drives both star formation in the galaxy and gas accretion onto the black holes, causing them to merge. The resulting quasar expels most of the gas from the galaxy, leaving a gas-poor galaxy containing a $\sim 10^8 M_\odot$ black hole (from Di Matteo *et al*. 2005).

One can compare $\rho_\bullet$ to the total increase in the mass density of black holes derived from the total radiation density emitted by AGN, an argument suggested by Soltan in 1982. If the total radiation density emitted by AGN were produced by accretion onto black holes with an efficiency ε of converting accreted rest-mass to radiation, then it must have increased the mass-density of supermassive black holes by

$$\Delta\rho_\bullet \approx 3.5 \times 10^5 \varepsilon_{0.1}^{-1} M_\odot \ \mathrm{Mpc}^{-3}$$

where $\varepsilon_{0.1} \equiv \varepsilon/0.1$ (Marconi *et al*. 2004; Soltan 1982). Most of this radiation comes from AGN with redshifts between 0.5 and 3. There is probably a 50% uncertainty in $\Delta\rho_\bullet$ due to uncertainties in the bolometric corrections and in corrections for obscured AGN and faint high-redshift AGN.

The striking correlation shown

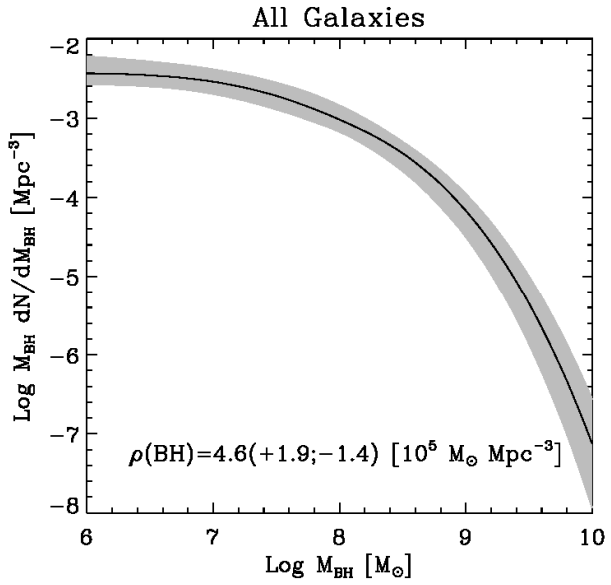


Figure 3-4: The local space density of black holes in galactic nuclei as a function of black hole mass, inferred from measurements of the kinematics of stars and gas in the nuclei of nearby galaxies (from Marconi *et al*. 2004).

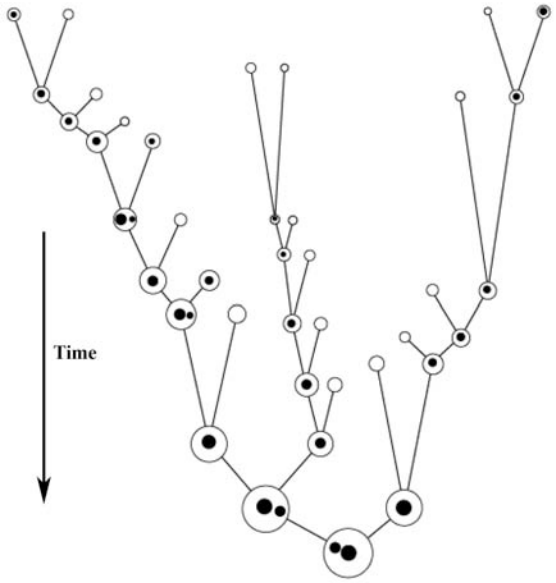


Figure 3-5: A cartoon of the merger-tree history of the assembly of a galaxy and its central black hole in cold-dark matter cosmology. Time increases from top to bottom, and the horizontal axis is a spatial direction. In this case the final galaxy is assembled from the merger of twenty smaller galaxies, containing a total of four seed black holes, and results in four mergers of binary black holes.

in Figure 3-1 between the mass of a galaxy's nuclear black hole and the galaxy's stellar mass and velocity dispersion (and hence the depth of its gravitational potential well) has convinced astronomers that throughout cosmic history, the growths of galaxies and their nuclear black holes have been tightly coupled (Gebhardt *et al.* 2000; Tremaine *et al.* 2002). Events in the galaxy decide the rates of the black hole's accretion, stellar capture, and merger; and jets and radiation from the black hole's accretion disk can remove gas from the galaxy and its surroundings (Figure 3-3).

The estimates for the growth in black hole mass, $\Delta\rho_\bullet$, can be compared to the estimate for the local density of black holes. Most of the contribution to the local black hole mass density comes from black holes with masses between 10^8 and 10^9 $M_\odot$ (Figure 3-4). Similarly, the main contribution to $\Delta\rho_\bullet$ comes from black holes in the same mass range at z = 2 - 3 (Merloni 2004). Since these billion-solar-mass black holes have grown in mass mainly through accretion, by a few orders of magnitude from z = 3 to now, the high-redshift Universe must have been dominated by black holes of a smaller variety, $10^6 - 10^7 M_\odot$. Were the smaller black holes quietly growing by mergers, while the supermassive ones were growing by accretion? Here lies a great potential for discovery by LISA, which will detect mergers of mass 10^5 - 10^7 $M_\odot$ with high signal-to-noise out to redshift $z \sim 20$ or higher.

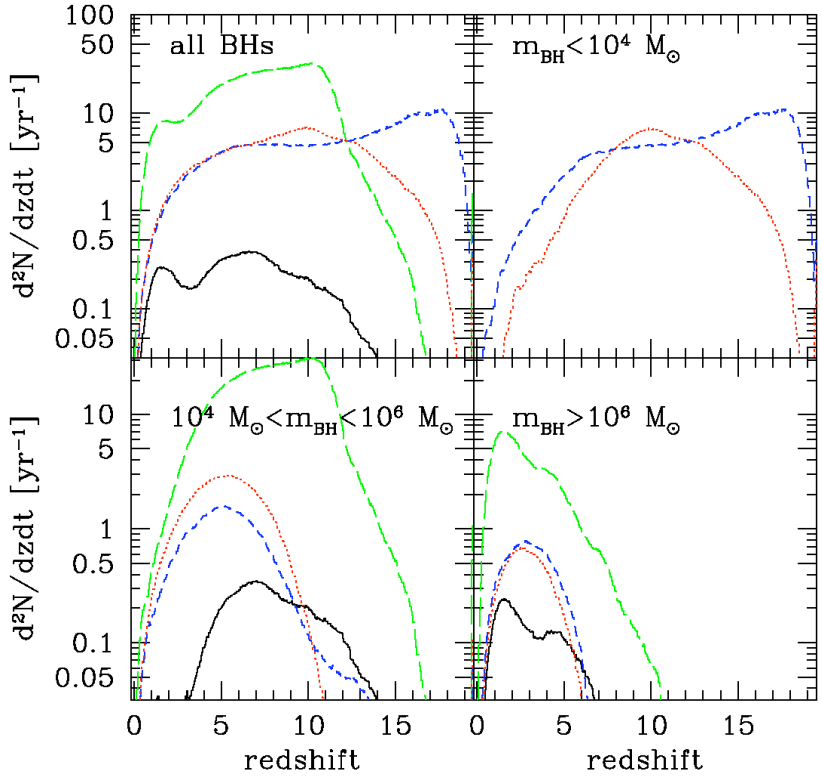


Figure 3-6: Predicted rates of massive black hole binary coalescences observable by LISA. Red, blue curves: scenarios with 150 $M_\odot$ black hole seeds resulting from the collapse of the first stars formed at $z > 20$, with black hole growth occurring by gas accretion during galaxy mergers. The black hole mergers are assumed to be driven initially by spherical stellar dynamics alone (red) or to be prompt (blue, *e.g.*, driven by gas drag). Black, green curves: scenarios with heavy seed black holes ~ 10^4-10^5 $M_\odot$. The green curve assumes the heavy seeds were produced in all dark halos with masses > 10^8 $M_\odot$ by $z = 15$. These merge to produce today's black holes (Koushiappas *et al.* 2004) The black curve assumes a much smaller number of similarly substantial seeds only in dark halos of > 10^{10} $M_\odot$, and requires that today's black holes grew mainly by gas accretion (from Volonteri 2006).

Growth and merger history of massive black holes

Expected merger rates

There is a simple argument bounding the number of mergers of supermassive black holes that LISA is likely to see. Hubble Space Telescope observes more than 10^{10} galaxies. Most bright local galaxies contain central supermassive black holes. Fossil evidence for mergers among local galaxies implies that about 70% of these have undergone a merger during the 0.8×10^{10} years since redshift $z = 1$ (Toomre 1977; Bell *et al.* 2006). Therefore, the galaxy merger rate at $z < 1$ must be close to one per year, and if the black holes in merging galaxies merge in turn, the merger rate of supermassive black holes should also be at least one per year.

Observations show that our Universe is dominated by cold dark matter, and that its initial spectrum of perturbations was such that the first objects to collapse under their self-gravity were tiny systems the size of dwarf galaxies. These then fell into each other to create larger ones (see Figure 3-5). Present-day galaxies like our Milky Way grew by the merger of more than 1000 subunits, which started forming already at redshift higher than 20. If each of these subunits initially contained a seed black hole of $10^4 M_\odot$, the merger rate seen by LISA could be as high as one thousand per year!

These elementary arguments suggest strongly that the merger rate of binary black holes that LISA will see lies in the range 1 - 1000 yr^{-1} (see Figure 3-6). The actual rate is proportional to: (1) the fraction of proto-galactic fragments that contain seed black holes massive enough for LISA to detect their mergers, multiplied by (2) the fraction of galaxy mergers that lead to black hole mergers. Our theoretical understanding of these fractions is limited and neither is well constrained observationally. LISA's measurements offer our best hope of determining them.

Seed black holes

The number densities and masses of the seed black holes largely determine their merger history. The similarity between $\rho_\bullet$ (the current best estimate for the mass density in black holes) and $\Delta\rho_\bullet$ (the density of rest mass that must have been accreted in order to produce the observed AGN) suggests that accretion does play a large part in the building of supermassive black holes. However, the estimates are also consistent with a significant fraction of the present black hole mass density being already present as black hole "seeds" by z = 10 - 20 (Koushiappas *et al.* 2004). In this scenario only a few of the seeds have grown through accretion to a billion solar masses, and these make up most of the mass we detect in supermassive black holes today and explain most of the accretion luminosity. The majority of seeds did not grow to high mass by accretion, so their presence and evolution is best traced by observing their mergers directly through gravitational radiation.

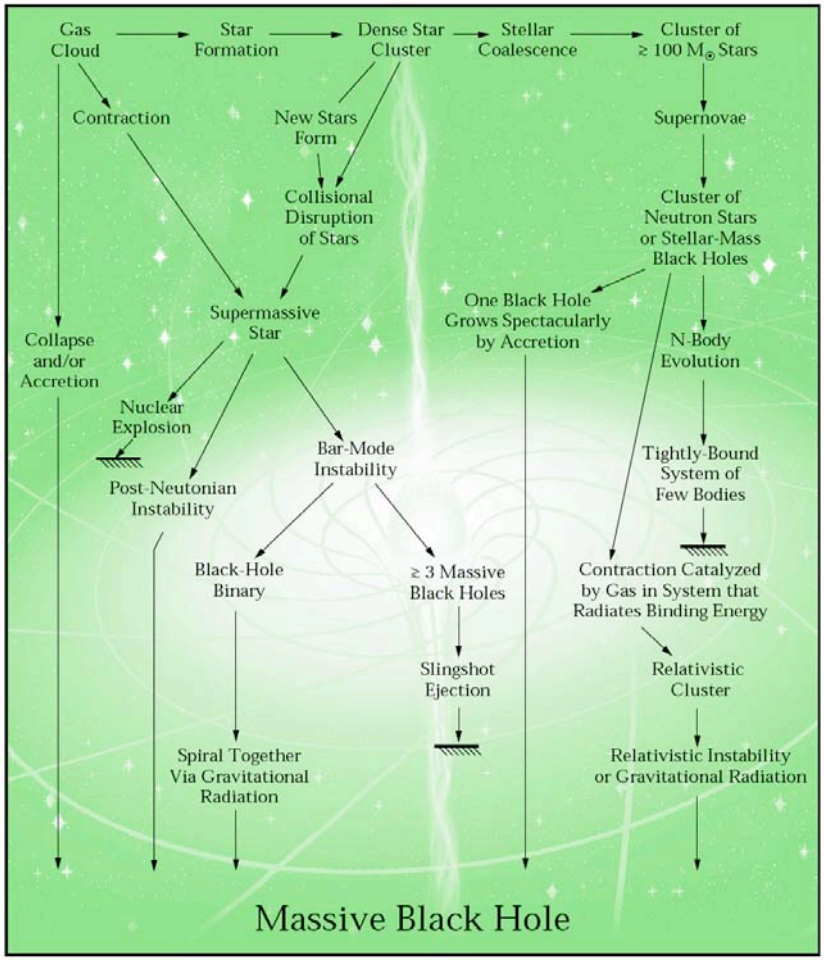


Figure 3-7: Scenarios for the formation of a massive black hole out of dense gas in the center of galaxies. From http://www.slac.stanford.edu/pubs/beamline/31/1/31-1-trimble.pdf (after Rees 1978).

Broadly, there are two

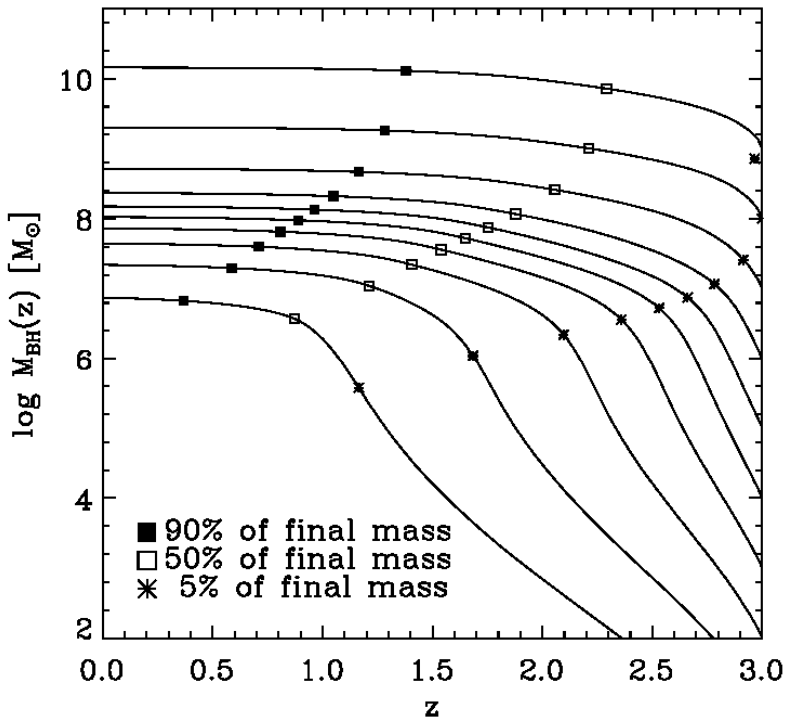


Figure 3-8: One possible growth history of black holes in active galaxies due to gas accretion, as a function of redshift and black hole mass, inferred from an X-ray selected sample of active galactic nuclei, assuming near-Eddington accretion (from Marconi *et al.* 2004).

currently popular scenarios for these "seeds". The first supposes that they are the ~100 $M_\odot$ black hole remnants of the first stars to form from primordial gas in the Universe (so-called Population III stars: Abel *et al*. 2002; Bromm *et al.* 2002; Bromm & Larson 2004; Madau & Rees 2001).

The second scenario is motivated in part by the existence of quasars at z = 5 - 6 (when the Universe was not even one billion years old) so luminous as to require black holes $> 10^9 M_\odot$ even at that early time. This second scenario supposes that much more massive seeds (10^4-$10^5 M_\odot$) could grow either by direct collapse of radiation-dominated massive star-like objects (Rees 1978; Bond *et al.* 1984), or by rapid accretion at super-Eddington rates (Begelman *et al.* 2006). See Figures 3-7 and 3-8.

Understanding how black hole seeds formed is a challenging observational task. There may be a population of "mini-quasars" at high redshift powered by accretion onto seed black holes, but if the mass of the seeds is below $\sim 10^6 M_\odot$ their flux, diluted by the large distance light has to travel from the formation redshifts ($z \sim 10$ - 30), is too weak for single sources to be detected electromagnetically. There is however evidence for a nearby population of low-mass AGN at $z <$ 0.2 (Barth *et al*. 2005). Seeds of mass $\sim 10^4 M_\odot$ can nevertheless be identified at high redshift during their mergers by observing their emission of gravitational radiation with LISA. Mergers may in fact play a much more important role in the higher-redshift Universe, because at $z > 5$ the merger rate is not limited by the Soltan-type argument that constrains the number of mergers at z < 3. Furthermore, constraints derived from the X-ray background actually limit the overall growth by accretion for black holes at $z > 6$.

The role of massive black holes in the evolution of early cosmic structure.

Active galactic nuclei powered by supermassive black holes keep the Universe ionized at $z < 4$, structure the intergalactic medium (IGM), and probably regulate star formation in their host galaxies. Intermediate mass black holes accreting gas from the surrounding medium may shine as "miniquasars" at redshifts as high as $z \sim 20$, with dramatic effects on the thermodynamics of the IGM (Kuhlen & Madau 2005). There are significant uncertainties about this key period in structure formation, from the fate of first stars (Abel, Bryan, & Norman 2002; Bromm *et al*. 2002; Bromm & Larson 2004) to the growth of $M > 10^9\ M_\odot$ black holes from seed black hole precursors in redshift $z > 6$ quasars (Madau & Rees 2001; Bromm & Loeb 2003) to the role of black holes in the reheating and reionization of the Universe (Madau *et al.* 2004) to the establishment at early epochs of the currently observed tight correlation between black hole mass and galactic bulge properties (Ferrarese & Merritt 2000; Gebhardt *et al.* 2000; Tremaine *et al*. 2002; Ferrarese & Ford 2005).

Models of hierarchical structure assembly predict that galaxy-sized dark matter halos will start to be common at redshifts $z \sim 10$ - 20 (see, *e.g.,* Mo & White (2002) for a pedagogical summary). This is also the epoch in which stars and galaxies first form in abundance, hence it is the beginning of the nonlinear phase of the Universe. However, electromagnetic observations of the properties and interactions of these early galaxies will be extremely challenging, because the first ones will be under-luminous dwarfs and because the surface brightness decreases as $(1 + z)^{-4}$.

In contrast, LISA observations of massive black hole coalescences from this epoch will be uniquely powerful in probing the halo mergers required in current models. The central black holes in many early halos are expected to have masses $M \sim 10^4$ - $10^6\ M_\odot$, corresponding to the redshifted mass range $10^5\ M_\odot < M(1+z) < 10^7\ M_\odot$ to which LISA will be most sensitive. The best current models predict tens of detected mergers per year at redshifts $z > 10$ (*e.g.,* Sesana *et al.* 2004). This is enough to characterize the redshift and mass distributions and discriminate between classes of structure formation models, in a crucial realm that is otherwise inaccessible to observation.

What happens when black holes collide and merge?

When two galaxies collide, the gravitational interactions between their stars and dark matter merge them into a ball of stars in $\sim 10^8$ yr. If both galaxies contain nuclear black holes, the black holes orbit within the ball of stars, and lose energy by deflecting the orbits of stars that pass near them. Within several orbital periods, the two black holes eject all the stars in their vicinity, and the rate of subsequent shrinking of their orbits is controlled by whatever stars and gas dribble in from larger radii, plus gravitational radiation (Merritt & Milosavljevic 2005). For black holes with mass $< 10^7 M_\odot$ (the ones to which LISA is sensitive), which live in galaxies with bulges smaller than the Milky Way's, gravitational radiation alone will merge the black holes in less than

10^{10} yr (Yu 2002). In more massive galaxies, the larger pair of black holes can eject all their surrounding stars before the pair is close enough for gravitational radiation to merge them. Then, without a further dribble of stars or gas, the black holes could be left orbiting each other for longer than the age of the Universe, at orbital separations of 0.1 - 1 pc and orbital periods of 10^2 - 10^4 yr. However, it appears that many galaxies have sufficiently flattened or tri-axial bulges of stars, or enough gas that the continuing supply of material for the pair to eject continues to cause the pair's orbit to shrink. The black holes are eventually driven close enough together that gravitational radiation takes over and merges them on a fairly short timescale. Observational evidence, too, favors prompt mergers even among the massive black holes. If the close pairs were long lived, a third merger would frequently lead to a third black hole interacting with the pair, leading to ejection of all three black holes from the nucleus of the galaxy, contrary to observation (Haehnelt & Kauffmann 2002). Furthermore, the most compelling explanation for X-shaped and double-double radio galaxies is that they result from reorientation of the black hole spin following a merger (Zier 2006 and references therein).

Black Hole Spins

No hair and no naked singularities

In general relativity, astrophysical black holes are predicted to be completely described by exactly two parameters: mass and *spin* (vacuum relativity also allows an electric charge, but this is quickly shorted out to gravitationally insignificant levels in the real world). This is the famous theorem that "black holes have no hair " (Thorne 1995). The total spin angular momentum S is usually specified in terms of the angular momentum per unit mass $a = S/M$ or the dimensionless $a_\bullet = S/(GM^2/c) = S/M^2 = a/M$ (the last two equalities being true only in relativists' units where $G = c = 1$). For a specified mass M, a black hole described by relativity cannot have $a_* > 1$ without showing a naked *singularity i.e.* one uncloaked by an event horizon (and this would generally be forbidden by the Cosmic Censorship conjecture). The simple non-rotating Schwarzschild black hole has $a_* = 0$.

Predicted spins from formation and growth

So what values $0 < a_* < 1$ do astronomers now expect for astrophysical black holes? Rotating gas clouds and stars have $a_* > 1$, so black holes that form from their collapse have high spin ($a_* = 0.75$), and leave much of their angular momentum in a residual disk that could accrete and raise a_* still more (Shibata & Shapiro 2002; Gammie *et al.* 2004). However, as discussed above in the context of seed black holes, most of the black holes in galactic nuclei have masses well above likely seed masses. Doubling the mass of a black hole is also enough to change a_* by of

order unity, so the values of a_* probably depend more on the growth history than on the original spins of the seeds (Hughes & Blandford 2003, Gammie *et al.* 2004, Volonteri *et al.* 2005). One possible evolutionary path involves rapid disk accretion of interstellar gas onto the black hole. Evidently this was the dominant path for the rare quasar black holes, which are too large to be probed by LISA. Another possible path involves a combination of gas accretion with many generations of black hole mergers. A third possibility for the growth of black holes is the random accretion of small packets of less dense material, such as material from tidally disrupted individual stars.

Merging binary black holes of comparable masses also have orbital angular momentum much larger than the maximum allowed for the merged black hole. However, recent numerical relativity simulations of the merger of equal mass, non-spinning black holes show that enough angular momentum is radiated to avoid a naked singularity, and the final steady-state merged black hole has $a_* = 0.7$ (Pretorius 2005). Simulations of the mergers of equal-mass rapidly rotating Kerr holes suggest it is unlikely to create a maximally rotating hole in this way; a maximum final value of $a/M = 0.89$ was produced from two holes with individual spin parameters $S/M^2 = 0.757$ aligned with the orbit (Campanelli *et al.* 2006).

By contrast, growing a black hole by accreting small companions (*e.g.*, stars or stellar mass black holes of mass M_c) which fall in on isotropically distributed orbits causes a secular decrease $a_\bullet$. If $M \gg M_c$, the angular momentum is essentially determined by a random walk, and $a_* \sim (M_c/M)^{1/2} \sim 10^{-3}$. This mechanism of black hole growth is most plausible for lower luminosity AGN and black holes of $< 10^7 M_\odot$ (Milosavljevic *et al.* 2006). In nuclear clusters dense and massive enough to supply enough stars to grow black holes of 10^8-$10^9 M_\odot$, stars collide (Figure 3-7; Rauch 1999). The subsequent evolution is hard to model, and there is no compelling evidence for the existence of such massive clusters.

Accreting unmagnetized gas in a thin disk with a steady direction of angular momentum drives a black hole to $a_* = 0.998$ after the accretion has increased its mass by a factor of 2.5 (Thorne 1974). Magnetized accretion disks are less effective at spinning up black

Table 3-1 Black hole spin characteristics for different growth scenarios (a-e), and representative efficiencies in converting accreted mass into radiation. Here efficiencies are calculated assuming no torque inside the innermost circular orbit, and corrected for the radiation emitted by the disk and swallowed by the hole (Thorne 1974).

black hole spin a_*	thin disk radiation efficiency (corrected for capture by hole) $\epsilon = L_{\rm disk}/(\dot{M}c^2)$
0^a	0.057
0.7^b	0.133
0.9^c	0.151
0.998^d	0.308
1^e	0.400

[a] result of isotropic accretion of small bodies.
[b] result of collapse or equal mass merger.
[c] approx equilibrium spin in magnetised disk accretion.
[d] equilibrium spin in unmagnetised disk accretion.
[e] maximal rotation before naked singularity appears.

holes, since they also lose angular momentum electromagnetically through disk winds, torques in the region where the gas begins to plunge into the black hole, and through magnetically-mediated extraction directly from the black hole. Simulations suggest that these effects may limit the final spins to $a_* \sim 0.9$, in contrast to the $a_* = 0.998$ of an unmagnetized thin disk (Krolik *et al.* 2005; Hawley & Krolik 2006). These results are summarized Table 3-1.

Although massive black holes could form by any combination of these processes, numerical astrophysics and relativity have show that these different mechanisms can produce very different distributions of black hole spins, skewing the distribution to high or low spins or spreading it evenly over a wide range of spins. No other astronomical observable can provide such a direct probe into the history of a key structural constituent of the Universe, and LISA will be able to measure such spins to better than 1% accuracy.

Merging massive black holes: Open questions and LISA's answers

There is strong evidence from electromagnetic observations that the many diverse phenomena of active galactic nuclei, ranging from giant radio jets to quasars, ultra-luminous infrared galaxies, the X-ray background and gamma-rays from blazars are all caused by accretion onto black holes. There is also strong evidence that galaxies grow by mergers, and that those mergers cause inflow of gas into the nuclear regions, exciting starbursts and feeding the central black holes. The winds and radiation from the stars and accreting black holes in turn expel the remaining gas, and determine the structure of the galaxy (Begelman 2003; Murray *et al*. 2005; Hopkins *et al*. 2006). While this evidence from electromagnetic observations is compelling, it is incomplete. Many fundamental questions about black holes and the galaxies they inhabit remain, which LISA will address:

- Except in a very small number of cases, the masses of the black holes in both quiescent and active galaxies are uncertain by factors of at least a few (due to the difficulty of constraining observationally the complications of stellar and gas dynamics). This prevents us from improving our understanding of accretion disk structure and dynamics. LISA observations of merging MBHs will reveal both their masses to typically less than 1% (Lang & Hughes 2006).

- Today, almost nothing is known observationally about the high-redshift merger tree that led to the galaxies we see today (see Figure 3-7). LISA observations of merging MBHs represent one of our best chances of observing this complex history. LISA can determine the binary's luminosity distance, typically to within a several percent. Assuming the standard concordance cosmology, this distance gives us the binary's redshift. Thus one can deduce the MBH merger rate as a function of z, and so trace the early history of galaxy mergers and the build-up of MBH masses over time (at least for those of solar masses). Today, the fraction of low-luminosity galaxies with black holes at $< 10^6 M_\odot$ is unknown at all redshifts.

- The spins of both quiescent and active black holes are poorly known. The distribution of spins will be strongly diagnostic of the mechanism of black hole growth (see discussion above on predicted spins). And the spins themselves are vital to many models of electromagnetic

phenomena: jet formation, jet twists, accretion efficiency, tidal disruption events. LISA observations of MBH mergers will reveal the spin parameters a_* of the merging BHs to within 1% (Lang & Hughes 2006). EMRI detections (see below) should typically reveal the MBH spin a_* to within ~ 0.01% (Barack & Cutler 2004).

Direct gravitational wave observations of these objects, which we now assume to be black holes, whether through their mergers or their capture of compact stellar-mass companions, should resolve these questions with a precision and certainty unachievable by other modes of observation.

Stellar captures and the dynamics of galactic nuclei

Some of the most exciting astrophysics with LISA will come from observing the gravitational waves produced by inspirals of stellar-mass compact objects into massive black holes. There is compelling indirect evidence that these inspirals are common throughout the Universe and occur about once per 10 million years in galaxies like the Milky Way (Freitag 2003, Hopman & Alexander 2006). Because the mass ratio for these binaries is typically ~ 10^{-5}, these sources are commonly referred to as extreme mass ratio inspirals, or EMRIs.

White dwarfs, neutron stars, and stellar-mass black holes all share the property that they reach the last stable orbit around the massive black hole (MBH) before they are tidally disrupted; hence all three types of compact stars can in principle lead to observable EMRI signals. However, black holes (BHs), being more massive, are expected to dominate the observed rate for LISA, for two reasons: mass segregation tends to concentrate the heavier compact stars nearer the MBH, and BH inspirals have higher signal-to-noise, and so can be seen within a much larger volume.

There are currently about 20 confirmed stellar-mass black holes, with estimated masses in the range ~ 5 - 15 $M_\odot$. However, it is generally believed that there are tens of millions of them in our galaxy, since they are expected to be produced in the deaths of massive (greater than ~ 20 - 25 $M_\odot$) stars (Woosley *et al.* 2002). The mechanism of mass segregation insures that the fraction of stars that are black holes is much higher very near the central MBH than at a random spot in the galaxy. In the Milky Way, near Sgr A*, it is estimated that there are several thousand stellar-mass BHs in the innermost parsec (Miralda-Escudé & Gould 2000; Freitag *et al*. 2006) and that stellar-mass BHs dominate the total stellar mass within 0.01 parsec.

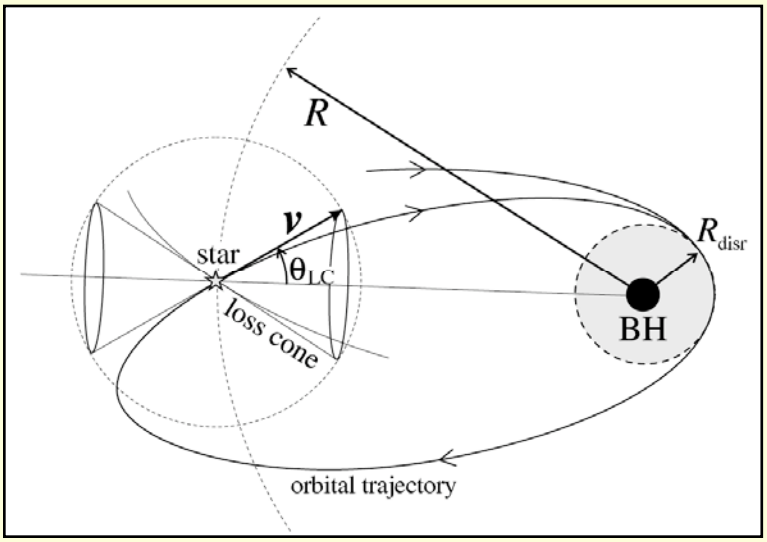
The oldest and best understood mechanism for creating EMRIs is that two-body encounters between stars in a dense stellar cluster near the MBH result in diffusion in each one's specific angular momentum, J. For non-rotating (Schwarzschild) MBHs, if J is reduced below about $4M$ (where M is the MBH mass), then the star plunges directly into the MBH. More interesting for LISA are the compact stars that diffuse into very high-eccentricity orbits, but that do not plunge directly. If the changes in J bring the value down to roughly $6M$, substantial energy starts to be lost due to gravitational radiation. The apocenter of the orbit shrinks as energy is radiated, and it

is possible for the apocenter to shrink enough that the compact star essentially ceases to interact with the other stars near the center of the galaxy, and gravitational radiation reaction completely drives the final evolution of the orbit.

Unfortunately, other scenarios can occur also. If J random-walks away from near zero before the BH-MBH system becomes substantially more tightly bound, then the losses due to gravitational radiation become negligible. Or, if J decreases to about $4M$ before much energy is radiated away, the black hole will plunge into the MBH rather than spiral in gradually. The gravitational wave signal then would not last long enough to produce a detectable signal-to-noise ratio. For a MBH of ~ 3 - 4 million solar masses (such as the one at the center of our Milky Way), gradual inspirals – the observable LISA sources – arise from the stars that are scattered towards the MBH from distances within a few times 0.01 parsec (Hopman & Alexander 2006). For such gradual inspirals, the last years of inspiral will typically be observable by LISA. Compact stars that are scattered into low-J orbits start out with orbital eccentricity very close to unity ($1 - e \sim 10^{-4}$), and their orbits remain moderately eccentric until the final plunge; median final eccentricities might be $e \sim 0.2$.

Mechanisms for Creating Extreme Mass Ratio Inspiral (EMRI) Events

Three different mechanisms for the production of EMRIs have been explored in the literature. The oldest and best-understood mechanism is the diffusion of stars in angular-momentum space, due to two-body scattering. Compact stars in the inner ~ 0.01 pc will sometimes diffuse onto very high eccentricity orbits, such that gravitational radiation will then shrink the orbit's semi-major axis and eventually drive the compact star into the MBH. This diffusion of stars into the "loss cone" is illustrated in the figure below.



Important physical effects setting the overall rate for this mechanism are mass segregation, which concentrates heavy, ~ 10 solar mass BHs close to the MBH, and resonant relaxation (the stars in the inner cusp are on nearly periodic orbits around the MBH, so the same two stars will interact with each other repeatedly), which increases the rate of diffusion (Hopman & Alexander 2006). The other two proposed mechanisms for producing EMRIs are tidal disruption of binaries that pass close to the MBH (Miller *et al.* 2005) and the creation of massive stars (and their rapid evolution into compact objects) in accretion disks surrounding MBHs (Levin 2006). It is not clear which mechanism will actually dominate the rate; indeed, all three could be roughly comparable.

It is estimated that most LISA detections of EMRIs will come from ~ 10 $M_\odot$ BHs spiraling into MBHs with masses in the range ~ 10^6 - $10^{6.5}$ $M_\odot$. The space density of MBHs in this mass range is ~ $1.7 \times 10^{-3}\,\mathrm{Mpc}^{-3}$, and LISA can detect such sources out to $z \sim 1$, corresponding to a co-moving volume of ~ 200 Gpc^3. The two-body scattering mechanism discussed above leads

to a rate (of observable inspirals) in Milky Way-type galaxies of $\sim 2.5 \times 10^{-7}\,\mathrm{yr}^{-1}$ (Hopman & Alexander 2006), which then implies a LISA detection rate ~ 85/yr (Gair *et al.* 2004). (The above estimate ignores "edge effects" related to the fact that inspirals ending in the first few months after LISA turns on will generally not be observable, since not enough signal-to-noise will have built up. Also EMRIs with unfavorable orientations will not be observable to $z = 1$, so a more conservative estimate would be ~ 50/yr.) The uncertainty in this rate is perhaps a factor of ~ 20.

In addition to the two-body scattering mechanism, other proposed channels for EMRIs are tidal disruption of binaries that pass close to the MBH (Miller *et al.* 2005) and creation of massive stars (and their rapid evolution into black holes) in the accretion disks surrounding MBHs (Levin 2006). Either of these channels could lead to a rate of the same order of magnitude as quoted above for two-body scattering, and this multiplicity of channels gives added confidence that at least one of them will produce a significant detection rate for LISA. Given a reasonable sample of detections, it should also be possible to infer the relative contributions from the different EMRI channels listed above. That is because two-body scattering leads to EMRIs that are moderately eccentric and have arbitrary inclination with respect to the MBH spin. In contrast, tidal disruptions of binaries lead to EMRIs that also have arbitrary inclination, but whose eccentricities are very close to zero; this is because tidal disruption results in orbits with (initially) much larger pericenter than for two body scattering, giving more time for radiation reaction to circularize the orbit before it becomes visible to LISA. Disk formation of EMRIs leads to sources with zero eccentricity and zero inclination (*i.e.*, orbital angular momentum parallel to the MBH's spin angular momentum). The orbital inclination and eccentricity can both be measured with very high accuracy from the LISA data (Barack & Cutler 2004).

During the last years before merger, the motion is highly relativistic, with speeds typically a third the speed of light. The period for precession of periapsis is comparable with the period for radial motion, and the orbital plane precesses just a few times less rapidly if the MBH has large spin. Thus the motion is quite complex, and correspondingly, the number of templates needed to do a brute force search for the signal is extremely high. Indeed, a brute force search for the signal over a year or more of data will not be possible, even with the computer power available in 2016.

Fortunately, Gair *et al.* (2004) were able to show that the search problem can be overcome by using shorter time intervals for the initial searches, and then combining the results incoherently. For three-week initial intervals, the computational resources required are reduced to reasonable levels. The main penalty of this approach is that the signal-to-noise ratio required for a successful search is increased from roughly 15 to 30-35. (That is, the source strength must be such that the combined, matched-filtering signal-to-noise using two independent LISA channels is 30-35.) Whether other search methods can be found that will avoid this increase in the required signal-to-noise ratio is now under investigation.

Recently, much more detailed simulations of EMRI events and of the stellar and compact object populations in galactic nuclei have been carried out by several different groups (Alexander & Hopman 2003; Freitag 2004; Hopman & Alexander 2005, 2006). As noted, there are still sub-

stantial uncertainties in the rates for EMRI events, because current observations simply do not impose strong constraints on models of the stellar swarms in galactic nuclei. The event rates for black hole EMRIs as a function of MBH mass will give detailed information on conditions in the galactic nuclei that probably cannot be obtained in other ways. The estimated event rates for white dwarf and neutron star EMRIs are considerably lower, but such events also are likely to contribute to clarifying conditions in the galactic nuclei.

Barack & Cutler (2004) showed that EMRI observations will permit measurement of the MBH spin parameter a_* with extremely high accuracy: $\Delta a_* \sim 10^{-4}$. Hence LISA is likely to provide a databank of hundreds or more black hole spins in the low-redshift ($z < 1$) Universe. As noted above, this spin distribution is highly sensitive to the evolutionary paths of the MBHs (Hughes & Blandford 2003, Gammie *et al*. 2004, Volonteri *et al.* 2005). No other astronomical observable provides such a direct probe into the history of MBHs, which chart the history of galaxies themselves.

LISA Science Objectives and Investigations relevant to this section (see Appendix 1)

1. Understand the formation of massive black holes

1.1. Search for a population of seed black holes at early epochs.

1.2. Search for remnants of the first (Pop III) stars through observation of intermediate-mass black hole captures, both at formation and at later epochs.

2. Trace the growth and merger history of massive black holes and their host galaxies

2.1. Determine the relative importance of different black hole growth mechanisms as a function of redshift.

2.2. Determine the merger history of 10^4 to 3×10^5 $M_\odot$ black holes before the era of the earliest known quasars ($z \sim 6$).

2.3. Determine the merger history of 3×10^5 to 10^7 $M_\odot$ black holes at later epochs ($z < 6$).

2.4. Determine the masses and spins of low-z massive black holes (MBH) using observations of extreme mass ratio inspiral (EMRI) events ($z \lesssim 1$).

2.5. Enable the search for electromagnetic counterparts to MBH mergers in order to determine the host galaxy.

3. Explore stellar populations and dynamics in galactic nuclei

3.1. Characterize the immediate environment of MBH in $z < 1$ galactic nuclei from EMRI capture signals.

3.2. Study intermediate-mass black holes from their capture signals.

3.3. Improve our understanding of stars and gas in the vicinity of Galactic black holes using coordinated gravitational and electromagnetic observations.

4. Black Hole Physics: Confronting General Relativity with Precision Measurements of Strong Gravity

Setting the stage

Key Science Questions

- Is the strong field gravity correctly described by GR?
- Are the massive dark central objects in galaxies really black holes?

General relativity (GR) is a theory of gravity, one of the fundamental forces of nature, in which gravitational fields are manifested as curvature of spacetime. GR was put forth by Albert Einstein nearly one hundred years ago as a remarkable marriage of physical insight and mathematical beauty. In its first experimental triumph, GR provided an elegant solution to an outstanding physical conundrum, the missing 43 seconds of arc per century in the precession of Mercury's perihelion that could not be accounted for by Newtonian gravity. The advances in our ability to measure such small effects during the latter part of the 19^{th} century provided this precision probe of the gravitational field near the Sun, and a confrontation with Newtonian gravity. Today, in the early years of the 21^{st} century, LISA provides the means to confront GR with experimental measurements at levels of precision and gravitational field strengths unimaginable a hundred years ago.

GR has no adjustable (or free) parameters and makes solid, specific predictions. While any test can therefore potentially be fatal, any failure of GR can point the way to new physics. Confronting GR with experimental measurements, particularly in strong gravitational fields, is an essential enterprise. And despite its great successes, we know that GR cannot be the final word on gravity, since it is a classical theory and so must break down at the Planck scale. As yet there is no complete, quantum theory of gravity, and gravitation is not yet unified with the other fundamental fields

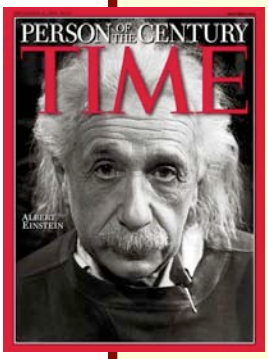
While so far GR has passed all the tests to which we have subjected it (Will 2006), most of these tests have been in the weak-field regime, which we can define by using the parameter $\varepsilon \sim v^2/c^2 \sim GM/(Rc^2)$. Here v is the typical velocity of the bodies, M their mass, and R their typical separation. In weak gravitational fields, $\varepsilon \ll 1$. For many astrophysical situations, such as Solar System dynamics, a fully general relativistic treatment is beyond our capabilities, but it is both practical and perfectly adequate to describe the dynamics using post-Newtonian (PN) equations, which are derived from the full general relativistic equations by systematically expanding them as a power series in ε. For the tests of GR that have been carried out in our Solar System, second-order corrections are of order $\varepsilon \sim 10^{-16}$, and so to-date it has been sufficient to model Solar System dynamics using first-order post-Newtonian equations. Solar

System tests have been completely consistent with general relativity, to this order of approximation.

Binary pulsars, which are essentially very stable and accurate clocks with typical orbital velocities $v/c \sim 10^{-3}$, can be seen as Nature's gift to relativists (Lorimer 2005). They have been excellent laboratories for precision tests of general relativity in gravitational fields considerably stronger than those prevailing in our Solar System. Current observations of several binary pulsars are perfectly consistent with GR predictions, as calculated through first post-Newtonian order. Observations of the first binary pulsar to be discovered, PSR B1913+16, also provided the first astrophysical evidence for gravitational radiation (earning the Nobel Prize for the binary's discoverers, Hulse and Taylor). Loss of energy due to gravitational-wave (radiation reaction) emission causes the binary's orbit to slowly shrink; $\dot{P}$ (an $O(\varepsilon^{2.5})$ effect) agrees with that predicted by GR to within 0.2%, which is within the error bars (Weisberg & Taylor 2004). The double pulsar system, PSR J0737-3039 A and B, discovered three years ago, should in the future provide even better tests of GR, for several reasons: it is the most highly relativistic of the known pulsar binaries; its distance and acceleration (relative to us) are much smaller; and, since both neutron stars are pulsars, it was possible to determine the mass ratio almost immediately from Newtonian-level dynamics, rather than having to infer it from relativistic effects. The last has made possible several additional GR tests that were not available previously (Kramer, *et al.* 2006). The orbital period derivative has already been measured (and is consistent with GR at the 1.4% level), and second-order post-Newtonian effects are expected to be measurable within several years (Kramer *et al.* 2006). (Why is radiation reaction – an $O(\varepsilon^{2.5})$ effect – more measurable than the $O(\varepsilon^{2})$ post-Newtonian effects? Because radiation reaction drains energy from the system and so causes inspiral, its effect on the orbital phase grows quadratically with time. By contrast, the first- and second-order post-Newtonian equations are conservative, and so their effect on the orbit grows only linearly in time. For observation times T of order $P/\varepsilon^{-0.5}$, the non-conservative $O(\varepsilon^{2.5})$ effect is more measurable than the second-order post-Newtonian corrections.)

LISA observations of coalescing massive black hole (MBH) binaries, or of stellar-mass compact objects spiraling into massive black holes, will allow us to confront GR with precision measurements in a strong-field/high-velocity regime not accessed by Solar System test or binary pulsar measurements. LISA will observe physics taking place just outside the event horizon, where typical velocities are $v/c \sim 1$. The merger of comparable-mass MBH binaries at cosmological distances produces an enormously powerful burst of gravitational radiation, which LISA will be able to measure with amplitude signal-to-noise as high as several thousand. In the months prior to merger, LISA will detect the gravitational waveform due to the binary's inspiral and from that inspiral waveform, the masses and spins of the two MBHs can be determined to high accuracy. Given these physical parameters, numerical relativity will be able to predict the exact shape of the burst waveform, and this can be compared directly to the observed burst – providing an ideal test of the theory. Stellar-mass compact objects spiraling into MBHs will provide a qualitatively different sort of test, but an equally exquisite one. The compact object

travels on a near-geodesic of the spacetime of the MBH, and as it spirals in, it effectively maps out the spacetime surrounding the MBH. For these extreme-mass-ratio inspirals (EMRIs), LISA will typically observe of order 10^5 cycles of inspiral waveform, all of which are emitted as the compact object spirals from 10 horizon radii down to a few horizon radii. Encoded in these waves is an extremely high precision map of the MBH's spacetime metric. Better opportunities than these for confronting GR with actual strong-field observations could hardly be hoped for. (One caveat, however, is that LISA observations of black hole binaries cannot discriminate between general relativity and scalar-tensor theories of gravity. The reason is that black holes do not support scalar fields; *i.e.*, they have no scalar "hair". Even after LISA flies, the best limits on scalar-tensor theories will come from Solar System and binary pulsar measurements; see Esposito-Farese 2004)



Did Einstein have the last word on gravity?

At the dawn of the 20^{th} century, Newton's theory of gravity was astonishingly successful. It provided a compelling and beautiful explanation of gravitational effects throughout the known cosmos. Newtonian gravity successfully predicted the existence of a previously unknown planet in the solar system. The theory passed every known experimental test, except one. Precision measurements of the precession of the perihelion of Mercury's orbit about the Sun revealed a small discrepancy of 43 seconds of arc per century. Newtonian gravity was unable to account for this anomalous precession, which was seen only in the orbital motion of the planet closest to the Sun – the planet where the Sun's gravity is strongest. GRs first triumph was to naturally account for this anomalous 43 seconds of arc per century.

At the dawn of the 21^{st} century, Einstein's theory of gravity is astonishingly successful. It also provides a compelling and beautiful explanation of gravitational effects through the cosmos as we know it today. GR successfully predicted phenomena such as gravitational lensing, black holes, and gravitational waves, and provides a natural framework for the expansion of the Universe. GR has passed every experimental test to date. But there is no well established quantum theory of gravity, and gravity is not yet unified with the other fundamental fields.

LISA provides unparalleled opportunities for making precision measurements in the regime of very strong and dynamical gravitational fields. Will such measurements unveil any flaws in the otherwise solid and successful edifice of GR? Even on macroscopic scales Newton did not have the last word on gravity.... did Einstein?

Now, this strong-field regime will quite likely be observed by ground-based gravitational-wave detectors several years before LISA flies (*e.g.*, the Advanced LIGO detectors should come online $\sim$ 2014 and LIGO is expected by that point to have observed BH mergers where the components are of roughly comparable mass. However, even the brightest BH mergers that LIGO will likely observe will still have amplitude signal-to-noise ratio ~100 times smaller than the brightest MBH mergers that LISA will observe. The precision with which LISA can measure the merger and ringdown waveforms will correspondingly be ~ 100 times better than for ground-based detectors. Similarly for extreme mass ratio inspirals: while the ground-based detectors may detect binaries with mass ratios $\sim 10^{-2}$ (*e.g.*, a neutron star spiraling into a 100 $M_\odot$ BH), in observations lasting $\sim 10^2 - 10^3$ cycles, the precision with which the spacetime can be mapped in such cases is at least two orders of magnitude worse than what is achievable with LISA's EMRI sources. Thus LISA will test our understanding of gravity in the most extreme conditions of strong and dynamical fields, and with a precision that is two orders of magnitude better than attainable from the ground.

GR has been extraordinarily fruitful in correctly predicting new physics, including the gravitational bending of light (or gravitational lensing), the gravitational redshift, black holes and gravitational waves. GR also provided the overall framework for modern cosmology, including the expansion of the Universe. Even if GR is fully correct (on lengthscales of astrophysical interest), LISA may reveal more new physics from strong-field GR. The famous singularity theorems of Penrose and Hawking assert that sufficiently compact objects must collapse, resulting in some spacetime singularity, but it is only a conjecture that the singularity is generically clothed by a black hole's event horizon. Might LISA reveal naked singularities or some other object formed of strongly warped spacetime? Could some central objects in galactic nuclei represent some other form of matter, such as massive boson or soliton stars? Since our understanding today of the nonlinear, strong gravity regime of GR is quite limited, LISA's "tests of GR" could reveal new objects that are unexpected, but perfectly consistent with GR.

The inspiral, merger, and ringdown of MBH binaries

LISA's strongest sources are expected to be coalescing MBH binaries where the components have roughly comparable masses, $0.1 \lesssim M_2/M_1 < 1$. The coalescence waveforms will be visible by eye in the data stream, standing up well above the noise, as illustrated in Figure 4-1.

As depicted in Figure 4-2, the coalescence can be described in three stages: inspiral, merger, and ringdown (Flanagan & Hughes 1998), all of which will typically be observable by LISA. The inspiral stage is a relatively slow, adiabatic process in which the BHs spiral together on quasi-circular orbits. The BHs have wide enough separations that they can be treated as point particles within the PN approximation; consequently, this stage can be computed analytically, with high-order PN expansions. The inspiral is followed by the dynamical

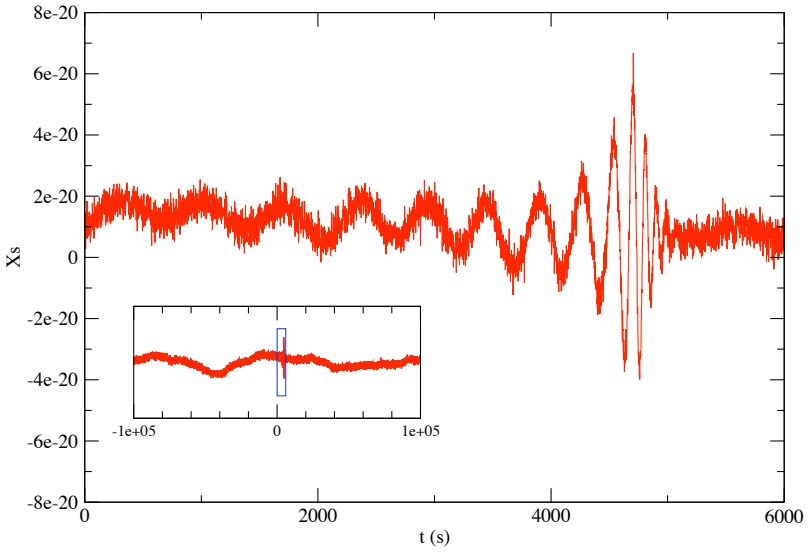


Figure 4-1 Gravitational wave signal for the final few orbits, plunge, merger and ringdown of an MBH-MBH binary. Here both MBHs have mass $10^5 M_\odot$ and are not spinning, the binary is at z=15, and is seen face-on. The signal is the sum of the gravitational waveform and simulated LISA noise. Note that even at z=15, the waveform stands up well above the noise and is visible in fine detail. The inset shows a longer stretch of data, containing the merger waveform. The large-amplitude, low-frequency "wiggles" are due to LISA's acceleration noise, which rises at lower frequencies.

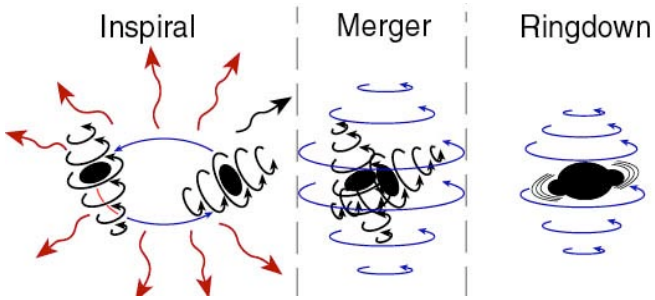


Figure 4-2 Schematic drawing of the three stages of the final coalescence of a comparable mass black hole binary, with spinning holes.

merger, in which the BHs leave their quasi-circular orbits and plunge together, forming a highly distorted remnant BH. Here, the BH velocities approach $v/c \sim 1$, the PN approximation breaks down, and the system can only be analyzed using numerical relativity simulations of the full Einstein equations. The distorted remnant settles into a stationary Kerr BH as it "rings down" by emitting gravitational radiation. While numerical relativity is required to tell us the initial state of the distorted BH, the evolution of that distortion — its "ringing down"— can be interpreted in terms of BH perturbation theory. At the end of the ringdown the final black hole is left in a quiescent state, with no residual structure besides its Kerr geometry. Its spacetime metric is then determined fully by its mass and spin, as required by the BH "no-hair" theorem.

For equal-mass MBH binaries with total mass in the range $3 \times 10^5 < M(1+z) < 3 \times 10^7 M_\odot$ where z is the cosmological redshift, the three stages all have comparable signal-to-noise; *i.e.*, within an order of magnitude of each other (see Figure 6 of Flanagan & Hughes 1998). From a typical LISA observation of the inspiral part of the signal, it should be possible to determine the physical parameters of the binary to extremely high accuracy. Using these parameters, numerical relativity should be able to predict precisely the merger and ringdown waves. The merger and ringdown waveforms will typically stand out well above the noise (by a factor $\sim 10^2 - 10^3$ after the noise has been band-passed to remove the large noise contribution from very low- and high-frequencies), so an extremely clean comparison will be possible between the observed waveforms and the predictions of general relativity.

We will now discuss the three stages of binary coalescence in greater detail, taking them in turn.

The inspiral stage: inferring the binary's masses and spins

With orbital velocities typically in the range $v/c \sim 1$, most of the inspiral stage can be well described using high-order post-Newtonian equations. The inspiral waveform is a chirp: a sinusoid that increases in frequency and amplitude as the BHs spiral together. The part of the inspiral stage that is observable to LISA lasts of order months to years. (More precisely, when the gravitational-wave frequency sweeps past 10^{-4} Hz, the time remaining until merger is

approximately $1.7 \times 10^8 s(0.25/\eta)[M(1+z)/2\times 10^5 M_\odot]^{-5/3}$ where $M \equiv M_1 + M_2$ is the total mass of the binary and $\eta \equiv M_1 M_2/M^2$ is the symmetric mass ratio. LISA will observe the last $10^4 - 10^5$ GW cycles from the inspiral. Since the inspiral signal is quite well understood theoretically, matched filtering can be used to dig these inspiral waveforms out of the noise, starting more than a year before the final merger, when the waves do not yet stick up above the noise as in Figure 4-1. And because the inspiral waveforms are strong, long-lived, and well-understood, it should be possible from the inspiral alone to determine the system parameters to amazingly high accuracy. Both masses should be determined to within fractional error $\sim 10^{-3}$, and their spins should be measured to within ~ 0.1 - 1% (Lang & Hughes 2006).

These are the size of random errors due to instrumental noise and confusion noise from white-dwarf (WD) binaries. Although waveforms available today are completely adequate for detection of waves, there will be some systematic errors in measured parameters due to the fact that the inspiral "template" waveforms, necessary for parameter estimation, will not be perfectly known theoretically. Highly accurate numerical solutions for tens of thousands of orbits are well beyond our current abilities, and will likely remain so for some time. However, great advances have been made in computing the inspiral to high order in post-Newtonian theory. For non-spinning bodies on nearly circular orbits, the post-Newtonian dynamics have been completely solved up through order $O(\varepsilon^{3.5})$ and the radiation reaction effects have been solved up through $O(\varepsilon^6)$: three-and-a-half orders beyond the order where radiation reaction effects first enters (Nissanke & Blanchet 2005). The influence of BH spin on the waveform phase has already been calculated to 2.5 orders beyond the lowest-order radiation reaction terms (Blanchet, Buonanno & Faye 2006). And for binaries with very small mass ratio, the post-Newtonian expansion is known to still much higher order. Additionally, it should also be possible to use numerical solutions to the late stages of inspiral to check and refine proposed methods for accelerating the convergence of PN expansion, such as Pade approximants. Therefore it seems very likely that well before LISA flies, theorists will be able to reduce to a very low level any systematic parameter estimation errors due to inaccurate theoretical inspiral waveforms.

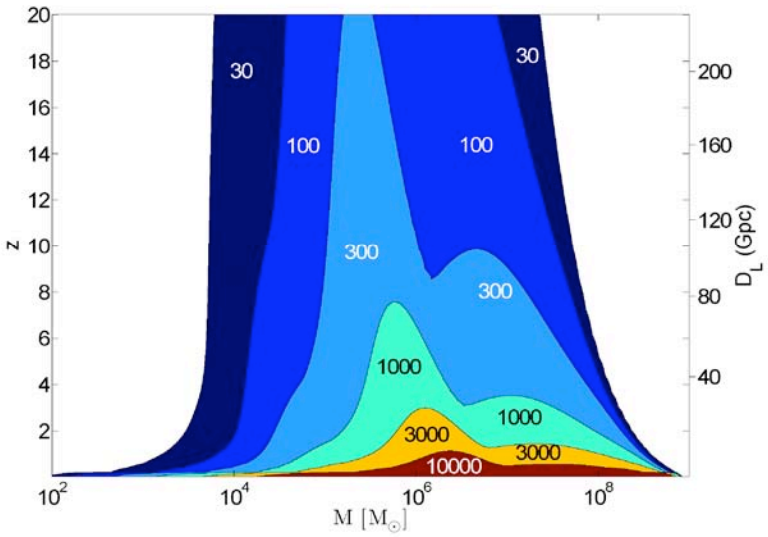
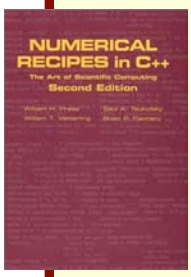


Figure 4-3 Plot of total signal-to-noise for all three stages of merger, for equal-mass binaries, as a function their total mass and redshift.

The merger stage: Spectacular bursts

The inspiral is followed by a dynamical merger that produces a burst of gravitational waves. This is a brief event, comprising only a few cycles if the holes spin slowly but more if fast, and lasting $\sim$ 1 hr $M/(10^6 M_\odot)$. However, it is fantastically energetic: during merger the gravitational-wave luminosity is $L_{GW} \sim 10^{23} L_\odot$, in that time producing more power than all the stars in the observable Universe. The final merger of MBH-MBH binaries occurs in the very strong-field, highly non-linear and highly dynamical regime of GR and is the strongest gravitational wave source that LISA is clearly predicted to see. LISA will be able to see the merger of two $10^4 M_\odot$ BHs out to redshift $z \sim 20$, and for mergers of $\sim 10^6 M_\odot$ BHs at $z \sim 1$ the LISA signal-to-noise will be in the thousands; see Figure 4-3. As mentioned above, LISA observations of the inspiral should give the masses and spins of the MBH components to $\sim$0.1%. With these in hand, numerical relativity will make a very specific prediction for the merger and ringdown radiation from the system. Comparison with the waveform that LISA actually observes will allow us to confront the predictions of GR with an ultra-high precision measurement in the fully nonlinear and dynamical regime of strong gravity.

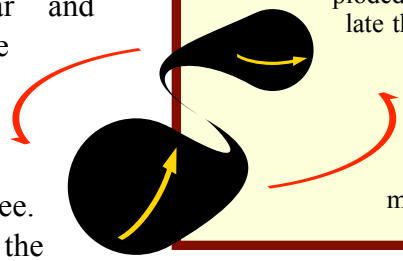


Numerical Relativity: Meeting the Theory Challenge

Recent dramatic advances in numerical techniques for simulating BH mergers and computing their waveforms have boosted opportunities to confronting GR with LISA observations. The importance – and difficulty – of calculating the gravitational waveforms from black hole mergers had been recognized by the AANM, Q2C, and Gravitational Physics decadal reports as a major theory challenge in this field. In the words of the Q2C report (page 118):

"Nearly as difficult as building these (gravitational wave) observatories, however, is the task of computing the gravitational waveforms that are expected when two black holes merge. This is a major challenge in computational general relativity and one that will stretch computational hardware and software to the limits. However, a bonus is that the waveforms will be quite unique to general relativity, and if they are reproduced observationally, scientists will have performed a highly sensitive test of gravity in the strong-field regime."

Numerical relativists have attempted to calculate gravitational waveforms from black hole mergers for over 30 years. The first full orbit of an equal mass, non-spinning black hole binary was finally achieved in late 2003. Roughly 1½ years later the field exploded, as several groups developed the means to simulate the final few orbits, plunge, merger and ringdown using dramatically different methodologies and achieving the same resulting gravitational waveforms. The first simulations with non-equal masses, and with spins, followed quickly. Numerical relativity is currently in an era of much activity and rapid progress, on a broad front.



To carry out this test, we must be able to compute the burst from merger and ringdown by solving Einstein's equations numerically. Such numerical simulations should start with the final few orbits of the black holes and then proceed into their plunge and merger, signaled by the formation of a single horizon, and continue through the ringdown of the remnant black hole. This has proved to be a very difficult undertaking, with attempts going back more than 30 years. Until 2005, the simulation codes were beset with many difficulties, including instabilities that caused the codes to crash before the evolution reached even one orbital period. Fortunately, a

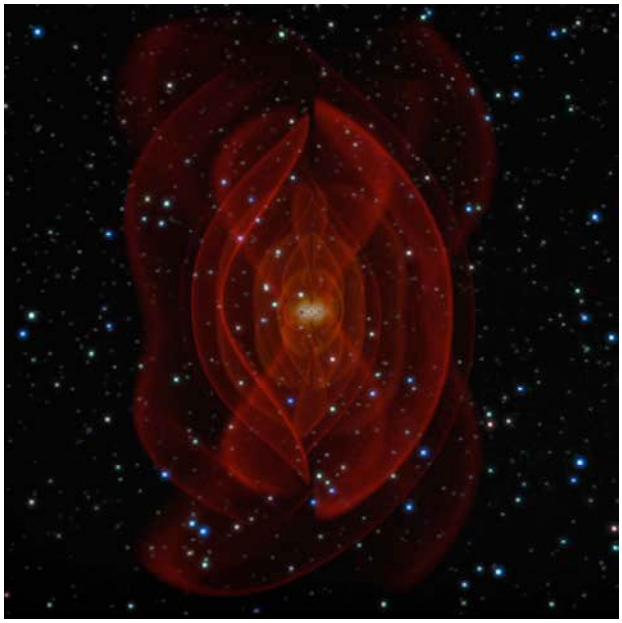


Figure 4-4 Gravitational waves from a numerical simulation of a binary black hole merger. The black holes are still separated in this frame and the gravitational waves are shown as the reddish contours.

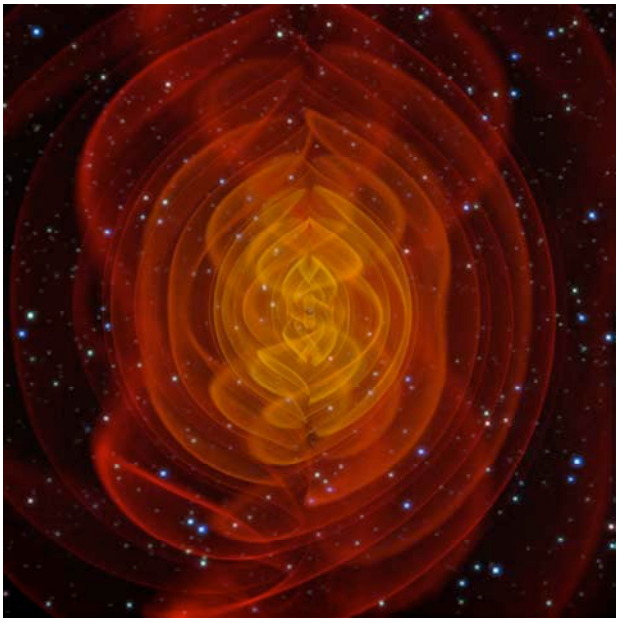


Figure 4-5 Here the black holes have merged and the resulting gravitational waves have increased in intensity, shown by the yellow and orange contours.

series of remarkable breakthroughs has occurred recently; see sidebar. With the development of new methods for evolving black holes on numerical grids, several groups can now evolve black hole binaries for several orbits and then through plunge, merger, and ringdown. So far, three groups have computed gravitational waveforms from the merger of two equal-mass, non-spinning BHs, and they all get the same simple shape shown in Figure 4-1 (Baker, Campanelli & Pretorius 2006). In addition, there is consensus that the merger of two equal-mass Schwarzschild BHs produces a final remnant BH with spin $J \sim 0.7GM/c$. Figures 4-4 and 4-5 are visualizations of the merger and emitted waves; Figure 4-1 is the waveform with superposed LISA noise.

Simulations of binaries in which the black holes have equal masses and spins, and non-equal masses but no spins, have also been carried out. Numerical relativity is now in an era of rapid progress, and it seems reasonable to expect that BH mergers with unequal masses and generic spins will be amenable to numerical solution well before LISA flies. In short, recent theoretical breakthroughs in numerical relativity expand the science that LISA can do, making LISA especially timely.

The ringdown stage: BH spectroscopy

Though numerical relativity waveforms from colliding holes naturally include the ringdown waves, those waves are well understood analytically. GR predicts are that every "excited" black

hole settles down to a stationary state, characterized entirely by its mass, spin and angular momentum[3], through the emission of gravitational radiation. These ringdown waves consist of a superposed set of BH quasi-normal modes (QNMs), waves with exponentially damped sinusoidal time dependence, plus a far weaker "tail" that decreases as (1/time)[6]. The modes are strongly damped as the mode energy is radiated away to infinity, so the final ringdown stage will be brief, lasting a few cycles

The quasinormal modes (QNMs) of Kerr black holes can be solved for using perturbation theory. That is, the spacetime metric is written as the Kerr metric plus a small perturbation, and Einstein's equations are expanded to first-order in that perturbation. The solutions can be decomposed into a sum of eigenmodes with complex eigenfrequencies. It was discovered in the 1970s that the partial differential equations for the eigenmodes/frequencies of Kerr BHs can be fully separated (*i.e.*, reduced to ordinary differential equations), so they are known to essentially arbitrary accuracy. While there are a countable infinity of these modes (corresponding to the angular order and overtone of the perturbation from the stationary state), the lowest order modes are the most readily excited and the most weakly damped; in practice only a few modes are likely to be observed. The frequencies and damping times of these ringdown QNMs are completely determined by the mass and angular momentum of the final, merged MBH.

We note that if two different QNMs are detected in a ringdown, then the ringdown radiation itself provides a test of strong-field GR and the hypothesis that the central massive objects in galactic nuclei are really Kerr BHs. The reason is that from two modes one measures four parameters (the frequencies and damping times of both modes), which must all be consistent with SAME mass and spin values (Dreyer 2003). Thus, in the same way that we can identify chemical elements through their spectroscopic fingerprint, so we can uniquely identify a BH (*i.e.*, determine its mass and spin) from the spectrum of its ringdown radiation. On the other hand, if the observed radiation arises from a different source (*e.g.*, a boson star), or if general relativity does not correctly describe gravity in the extremes of strong fields and dynamical spacetimes, the spectrum would very likely be inconsistent with that predicted by general relativity for black holes.

Extreme-Mass-Ratio Inspirals: Precision Probes of Kerr Spacetime

Introduction

Observational evidence for the existence of MBHs at the centers of galaxies is currently based on modeling the gravitational potentials of these objects using the motions of stars and gas, and comparing the results with those expected if the central object were a black hole. The best case today comes from stellar motions near the center of our galaxy, which reveal the presence of a compact dark object of mass $M \sim 4 \times 10^6 M_\odot$; stellar orbits show the central mass to be point-

[3] And electric charge, though for astrophysical black holes the charge has a negligible effect on the spacetime structure.

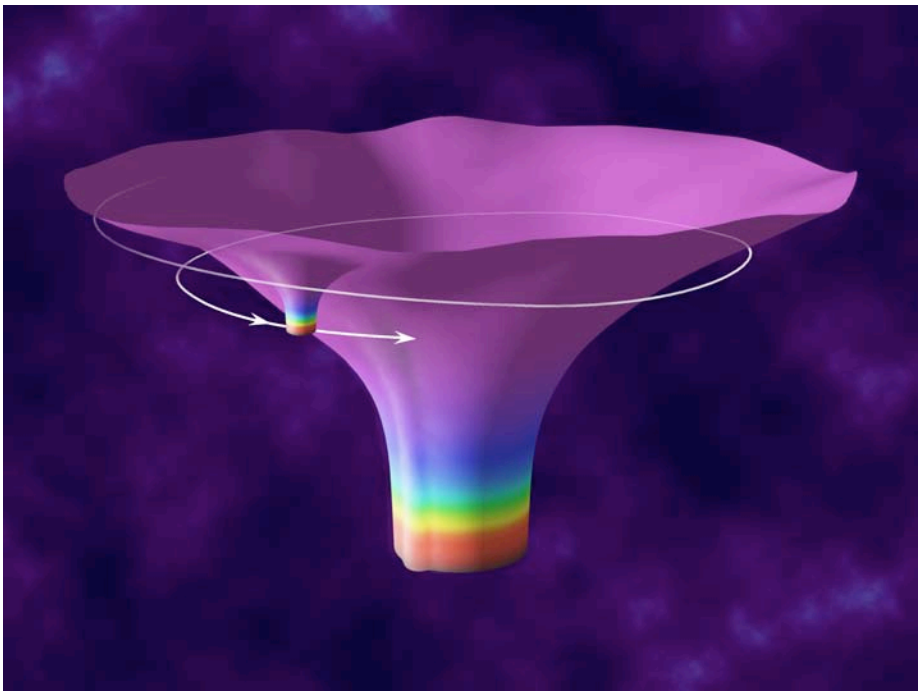


Figure 4-6 Embedding diagram of an EMRI, with the smaller black hole orbiting in the spacetime of the larger black hole. The colors depict the slowing of time (the "lapse" function) as one nears the horizons and the shape depicts the geometry of space in the orbital plane.

like down to a scale of $\sim$ 100 AU. LISA will map the structure of MBH spacetimes on length scales $\sim 10^4$ times smaller - the size of the horizon.

An inspiraling binary with one body much less massive than the other is referred to as an extreme-mass-ratio inspiral (EMRI). In the LISA context, the larger body is a $\sim 10^5 - 10^7 M_\odot$ MBH and the smaller one a stellar-mass compact object (white dwarf, neutron star, or stellar-mass BH) or perhaps an intermediate-mass BH, with mass $\sim 10^3 M_\odot$. Thus LISA EMRIs will have mass ratios $10^{-2} \lesssim m_2/m_1 \lesssim 10^{-7}$. While white dwarfs, neutron stars and BHs should all reach the event horizon of a MBH before being tidally shredded, the best current estimates are that inspirals of $\sim 10\ M_\odot$ BHs will dominate the LISA detection rate (both because mass segregation concentrates them more strongly near the MBH, and because, being more massive, they produce a stronger signal and so can be seen to greater distance). Therefore in this section we concentrate on the $\sim\ 10\ M_\odot$ BH case, shown in Figure 4-6, (but with the small hole enlarged in size for easy visualization).

Even the closest EMRI will yield a signal too weak to be immediately visible above the LISA noise, so matched filtering will be necessary to dig EMRI signals out of that noise. Fortunately, the matched filtering signal-to-noise scales like the square root of the number of observed GW cycles, and LISA will observe each EMRI for a timescale of order years, or equivalently for of order 10^5 cycles, so the matched filtering signal-to-noise will be a factor $\sim$ 300 higher than the instantaneous signal-to-noise. Unfortunately the number of independent year-long templates is so vast that straightforward matched-filtering, using a grid of templates densely covering the entire parameter space, will not be feasible computationally, and one will likely have to resort to filtering that is suboptimal (compared to straightforward Wiener filtering), but more computationally efficient. The best current estimate is that EMRIs with (total, matched-filtering) signal-to-noise > 30-35 will be detectable with such methods (Gair *et al.* 2004). Then a

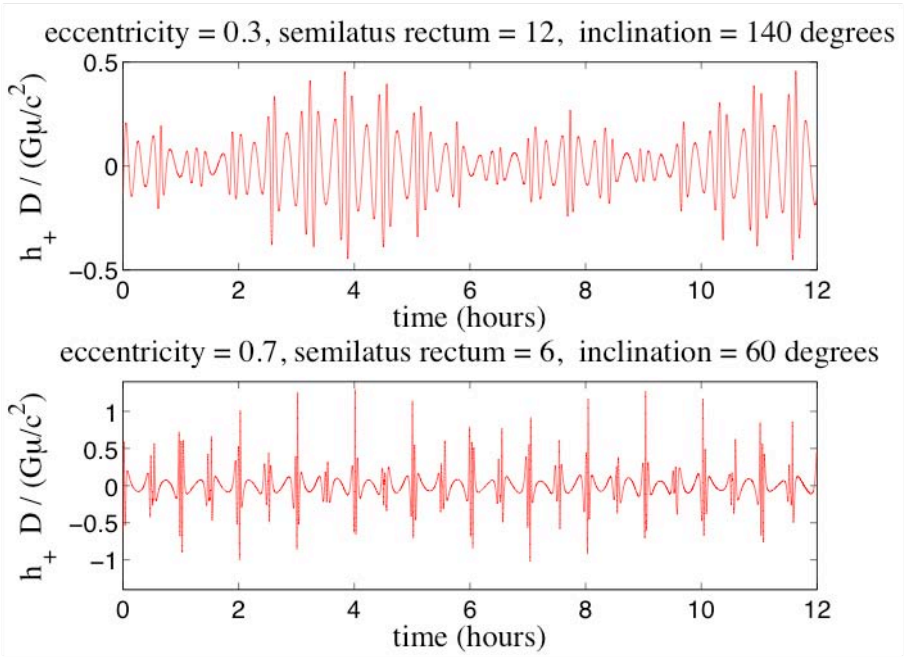


Figure 4-7 Segments of generic EMRI waveforms (Drasco & Hughes 2006); These are the plus-polarized waves produced by a test mass μ orbiting a $10^6 M_\odot$ black hole that is spinning at 90% of the maximal rate allowed by general relativity, a distance D from the observer. The two panels correspond to different configurations of the test mass' orbit, characterized by the shown parameters. The top panel assumes a slightly eccentric and inclined retrograde orbit modestly far from the horizon. The bottom panel assumes a highly eccentric and inclined prograde orbit much closer to the horizon. The amplitude modulation visible in the top panel is mostly due to Lense-Thirring precession of the orbital plane. The bottom panel's more eccentric orbit produces sharp spikes at each pericenter passage.

$10\,M_\odot$ BH spiraling into an MBH will be observable by LISA out to $z\sim1$. Given the current best estimate for the rate of BH inspirals in the Milky Way ($\sim 2.5\times10^{-7}$/yr; Hopman 2006), and extrapolating this rate to the rest of the Universe, leads to a LISA detection rate of ~ 50 - 100/yr, with the strongest sources having (matched-filtering) signal-to-noise > 100 (Gair *et al.* 2004).

The EMRIs are expected to be very clean astrophysical systems (except perhaps in the few percent of galaxies containing active galactic nuclei, where interactions with the accretion disk could possibly affect the dynamics). Over timescales of order a day, the orbits of the smaller body are essentially geodesics in the spacetime of the MBH. On longer timescales, the loss of energy and angular momentum due to gravitational-wave emission causes the smaller body to spiral in; *i.e.*, the geodesic's "constants" of integration change slowly over time. For LISAs observation time (of order years), the orbits are highly relativistic (radius < 10 Schwarzchild radii) and display extreme forms of perihelion precession and precession of the orbital plane due to dragging of inertial frames by the holes spin. Figure 4-7 shows two sample waveforms, corresponding to short stretches of time.

To correctly model the orbit over timescales longer than ~1 day, one must include gravitational radiation reaction, which causes the small body to lose energy and angular momentum. The natural approach to this problem is to take advantage of the tiny mass ratio m_2/m_1 and use perturbation theory; *i.e.,* treat the small body as producing a metric perturbation on the spacetime of the MBH. It is simplest to model the small body as a point particle, so one can ignore its internal structure [as was justified by the work of Mino *et al.* (1997)], but the cost of this simplification is that the perturbed metric diverges at the location of the particle, and so

the radiation reaction force must be regularized. A general (and quite beautiful) regularization prescription was given independently by Mino *et al.* (1997) and Quinn & Wald (1997), but implementing their prescription numerically has been a challenge. How to do so is at once a fascinating question in pure general relativity and a practical requirement for deriving the maximum science possible from LISA. It is therefore a very attractive problem for a relativist to work on, and a highly talented community of ~ 25 relativists is now heavily engaged in this subfield. Recently, Mino (2005) has suggested a relatively simple and practical solution to this problem, based on calculating the average effect of radiation reaction over many orbits rather than the instantaneous self-force. It is not yet clear whether Mino's proposal will yield waveforms that are sufficiently accurate for LISA's purposes, but, if not, there is also a well-developed alternative: Barack & Ori (2003) have developed a mode-sum numerical procedure for regularizing the self-force, and its effectiveness has been demonstrated for circular orbits in Schwarzschild, where the results can be checked using simpler methods (Barack & Lousto 2005). Therefore it seems very likely that theorists will provide extremely accurate EMRI template waveforms well before LISA flies.

Given the very large number of EMRI GW cycles that accumulate over a year of LISA observations (~ 100,000), a fit of the observed gravitational waves to theoretically calculated templates will be very sensitive to small changes in the templates physical parameters. As mentioned above, this sensitivity makes the search computationally challenging, but it allows extremely accurate determination of the source's parameters, once an EMRI signal is identified. For example, (assuming that GR is correct and the central massive object is a Kerr BH), LISA should be able to determine the mass and spin of the MBH to fractional accuracy $\sim 10^{-5} - 10^{-3}$ (Barack & Cutler 2004).

EMRI tests of the Kerr-ness of the central massive object in galactic nuclei

This level of precision suggests that we can do more — use the EMRIs as a highly precise observational test of the "Kerr-ness" of the central massive object. That is, if we do NOT assume that the larger object is a black hole, we can use gravitational waves from an EMRI to map the spacetime of that object. The spacetime outside a stationary axisymmetric object is fully determined by mass moments M_l and current multipole moments S_l. The mass moments tell us about the distribution of mass and energy in the source; they are analogous to electric charge moments in classical electricity and magnetism. The current multipole moments about the *motions* of mass and energy; they are analogous to magnetic moments. Since these moments fully characterize the spacetime, the orbits of the smaller object, and the gravitational waves it emits, are therefore determined by the spacetime multipolar structure. By observing these gravitational waves with LISA we can characterize the spacetime of the central object.

Indeed, Ryan (1995) showed that for inspiraling trajectories that are slightly eccentric and slightly non-equatorial, *in principle* all the multipole moments are redundantly encoded in the emitted gravitational waves, through the time-evolution of the orbit's three fundamental frequencies. Basically, these are the fundamental frequencies associated with the r, θ, and ϕ motions

(Drasco & Hughes 2004), or, equivalently, the radial frequency and the two precession frequencies. Extracting the moments from EMRI waves is analogous to geodesy, in which the distribution of the earth's mass is determined by studying the orbits of satellites. Black hole geodesy – also known as bothrodesy – is very powerful because Kerr black holes have a very special multipolar structure. A Kerr black hole with mass M and spin parameter a (in units with G=c=1) has multipole moments given by $M_l + iS_l = M(ia)^l$. Thus, $M_0 = M$, $S_l = aM$, $M_2 = -a^2 M$, and similarly for all other multipole moments; they are all completely determined by the first two moments, the black hole mass and spin. This is nothing more than the "no hair" theorem for black holes: a black hole's properties are entirely determined by its mass and spin.

How accurately can M_2 be extracted from an observed EMRI signal, independently of M and S_l? It was recently estimated that M_2 could be measured to within $\Delta M_2 \sim 10^{-2} - 10^{-4} M^3$ (Barack & Cutler, in preparation), along with $\Delta M/M$ and $\Delta S_l/M^2$ to $10^{-5} - 10^{-3}$. Thus the standard picture for the central massive object — a Kerr BH described by GR — makes a definite prediction regarding M_2, which can be checked to very high precision using EMRI data. Any inconsistency with the Kerr relation could signal a failure of GR, or the discover of a new type of compact object, or some surpisingly strong perturbation from other material or object.

Other tests of the Kerr-ness of the central massive object have also been proposed. Kesden *et al.* (2005) have argued persuasively that one can use EMRI waves to distinguish definitively between a central MBH and a boson star. In the BH case the GW signal "shuts off" shortly after the inspiraling body reaches the last stable orbit (and then plunges through the event horizon), while for a massive boson star, the signal gets prolonged, and its frequency derivative changes sign, as the body enters the boson star and spirals toward its center.

The above tests take the standard model of the central object (a Kerr BH described by GR) and either 1) compare it with a different model (*e.g.*, a boson star) or 2) embed that model in a larger one with extra free parameters (*e.g.*, Kerr, but with arbitrary quadrupole moment), and ask whether the best fit to the data is consistent with Kerr values for the extra parameters. These are all essentially comparison tests. There seems to be no unique or optimum way of constructing such tests —basically because today there is no compelling alternative to general relativity to compare against it. Nevertheless, tests of this sort would be very useful in either cementing confidence in the standard picture, or homing in on discrepancies.

But precisely because there is no compelling alternative to GR today, it is important to also consider other ways of "confronting GR" with EMRI data, which do not involve comparison with (perhaps somewhat *ad hoc*) alternatives. For instance, one could break an observed EMRI signal into, say, four consecutive pieces (each piece with signal-to-noise $\sim 100/\sqrt{4} = 50$), and estimate the binary's physical parameters from each piece separately. It is clear that very precise estimates of both BH masses and the MBH spin could be obtained from each segment, and all four segments must yield the same parameters, within the error bars. It is clear that inter-segment comparison would yield a quite precise test of the null hypothesis that the central massive object is indeed a Kerr BH. Other tests of a similar nature could easily be formulated.

Bounding the mass of the graviton

From LISA observations, it should be possible to improve significantly on the current Solar System bound on the graviton mass: $m_g < 4 \times 10^{-22}$ eV. Two different methods for bounding have been proposed. The first method is based on LISA observations of WD binaries. A large fraction of the gravitationally strongest WD binaries will be eclipsing, so optical observations will tell us the instant when, as seen from Earth, one body is (almost) directly behind the other. This is also the instant, when the (square of) the gravitational-wave strain field goes through a minimum, assuming that the gravitons and photons propagate at the same speed. However if gravitons and photons travel at different speeds, due to a non-zero graviton mass m_g, there will be a phase delay between the optical and gravitational signals. By bounding this phase delay, one bounds m_g. A single binary should be sufficient to set a limit of $m_g < 10^{-22}$ eV. With N such sources, the bound should decrease like $N^{-1/2}$, and N should be order 100, so a bound of $m_g < 10^{-23}$ eV could be achieved this way (Cooray & Seto 2004). (Note the bound from any binary is independent of its distance, D; doubling D would double the delay for any m_g but the error bar on the phase delay also doubles, since the signal-to-noise is halved.)

The second method is based on measuring the phase evolution of MBH inspiral signals. This phase evolution is predicted rather precisely by high-order post-Newtonian calculations. However if $m_g > 0$, the higher-frequency gravitational waves emitted late in the inspiral travel faster than the lower-frequency waves emitted earlier, and to some extent "catch up" with them. Since MBH inspirals are much stronger GW emitters than WD binaries, this leads to a correspondingly stronger bound: $m_g < 10^{-25}$ eV (Berti, Cardoso & Will 2006).

LISA Science Objectives and Investigations relevant to this section

(see Appendix 1)

5. Confront General Relativity with observations

5.1. Detect gravitational waves directly and measure their properties precisely.

5.2. Test whether the central massive objects in galactic nuclei are the black holes of general relativity.

5.3. Make precision tests of dynamical strong-field gravity in massive black hole binaries.

5. Precision Cosmometry and Cosmology

Precision cosmometry is defined as the art of making precision measurements of the world or the Universe. The measurement of distance forms the foundation for much of astronomy and cosmology. An example is the Hubble constant which has had a history of refinement spanning decades. LISA has the potential to make fundamental contributions to precision cosmometry with its ability to provide gravitationally calibrated distances to sources with a 3% accuracy or better.

Using precision cosmometry for cosmology

LISA measurements of waveforms from black hole binary merger inspirals yield gravitationally calibrated, absolute distances to high redshift. The individual raw absolute precision for a single event ranges from typically about 0.2% to 0.4% at a redshift $z = 1$ to 1% to 3% at $z \approx 3$ to 5. The absolute physical calibration, high per-event precision, and large redshift range all represent new and unique capabilities. A redshift-distance relation with this approach requires identification of the host galaxy to obtain a redshift independently, and additional errors are added by weak lensing noise at high z. Even a small number of events at $z < 1$ may allow LISA to measure the Hubble constant to about better than 1% accuracy, and together with events at higher redshifts may probe global curvature and cosmic dark energy with a precision comparable with other methods. The technique complements other methods; their combination provides unique information about the new physics of dark energy, and new tests of concordance cosmology.

Precision cosmology characterizes the structure and behavior of the Universe as a whole: its global curvature, its expansion with time, and the behavior of perturbations. The global curvature of space is a relic of the earliest observable moments of inflation and carries information about the initial conditions of the Universe; cosmic expansion history tests models of the new physics of dark energy; and cosmological perturbations test the dynamical predictions of general relativity on the largest scales. More than simply mapping our Universe, precision cosmology explores in detail the behavior of space, time, matter and energy at the opposite extremes to black holes: the lowest density, the largest scales, and the earliest times.

For the most powerful tests we seek not only high precision, but also a variety of different cosmometric techniques that measure global spacetime in different ways (see "Cosmology Primer" box). Precision measurements of cosmic microwave background (CMB) anisotropies (from COBE, balloon- and ground-based experiments, WMAP [Spergel *et al.* 2006], and in the future, Planck) set the highest standard of quality: CMB now reliably determines certain combinations of cosmological parameters with precision at a level of a few percent. Combining other types of measurements with the CMB data breaks degeneracies in fundamental quantities, increases reliability by controlling systematic errors, probes recent expansion where dark energy dominates, and allows deeper questions to be asked: for example, whether dark energy varies with time or reflects a pathology in the theory of gravity on large scales (rather than a new form of energy).

Key science questions

- What is the nature of dark energy?
- What is the global geometry of the Universe?
- What is the Hubble constant?
- Can redshifts be determined from electromagnetic observations by identifying the host galaxies of black hole merger events.

A cosmology primer

Precision cosmology aims to measure the global properties of our spacetime: the geometry (and statistical fluctuations) of the temporally evoloving, nearly spatially uniform 4-dimensional manifold that best represents all of the Universe that we can see. This geometry contains information about the curvature of 3-space, thought to be a relic of how the Universe began, and about the character of the average mass-energy content of the Universe over time.

Abundant evidence indicates that the Universe on the largest scales is nearly homogeneous and isotropic: on a large scale average, it is statistically the same everywhere and in all directions. Such a smooth, symmetric universe can be visualized as a uniform 3D space (which may have a uniform positive, flat, or negative spatial scalar curvature k/a^2, as specified by a curvature parameter k=1,0,-1), that uniformly changes size with time according to a cosmic scale factor $a(t)$. Referred to the present, the size is inversely proportional to a dimensionless ratio (1+z) where z is called the redshift. Points "at rest" in this space appear to recede at a speed cz proportional to their distance with physical rate of recession per unit separation given by the Hubble parameter $H(t) = \dot{a}/a$.

Einstein's equations relate the evolution of a to the mass-energy density and curvature:

$$\left(\frac{\dot{a}}{a}\right)^2 \equiv H^2 = \frac{8\pi G}{3}\sum_i \rho_i - \frac{kc^2}{a^2}$$

where the mass-energy is divided here into components i with density, ρ_i. As the Universe expands, the matter and energy spread out. Furthermore, if a component has a significant pressure (momentum flux) p_i, the work done during the expansion changes its internal energy. Local energy conservation gives:

$$\dot{\rho}_i = -3H(1+w_i)\rho_i,$$

where $w_i \equiv p_i/\rho_i c^2$ is the equation of state parameter for each component. For discrete components with constant w_i, the evolution of the scale factor can be written

$$H(a)^2 = H_0^2 \sum_i \Omega_i \left(\frac{a_0}{a}\right)^{3(1+w_i)}$$

where H_0 denotes the present-day Hubble constant, and the Ω_i represent the present-day densities of mass-energy components in units of the so-called critical density,

$$\rho_c = (3/8\pi G)H_0^2.$$

The history of the expansion of the Universe reflects both its geometry and composition. The sum includes a term Ω_k with w_k=-1/3 representing the dynamical effect of spatial curvature, components describing radiation (with w_k=1/3), and nonrelativistic (baryonic and cold dark) matter (with w_m=0). Of particular interest is the value of w_{DE} (or more generally, $w_{DE}(t)$) which reflects the physical character of mysterious cosmic dark energy. Current data suggest that the Universe is dominated by dark energy ($\Omega_{DE} \sim 0.74$), and w_{DE} is constrained at about the ten percent level by relative supernova distances (Astier *et al.* 2006), w_{DE} = -1.023 ± 0.090(random) ± 0.054(systematic). Einstein's cosmological constant Λ, corresponding to the gravitation of a uniform physical vacuum, has w_Λ=-1, with no t dependence. Precise data are capable of discovering whether the dark energy is consistent with an Einstein Λ or other models of $w_{DE}(t)$, estimating the other parameters including Ω_k, and testing whether general relativity is an accurate description of global behavior. Some techniques (such as relative distance distances from Type Ia supernovae or other standard candles, or gravitational lenses with time delays) measure the global spacetime metric directly via photon propagation. Others (such as weak gravitational lensing by intergalactic dark matter perturbations, or surveys of high redshift galaxy clusters) also measure the response of the smooth metric to mass density perturbations via the growth of structure. A few techniques offer absolute physical calibration of distances; these include Sunyaev-Zeldovich mapping of plasma in galaxy clusters, and, more precisely, mapping of the galaxy correlations imprinted by acoustic oscillations of baryons in the baryon/radiation plasma before recombination. The most precise data at present, the anisotropy spectrum of the microwave background, constrains most directly the behavior of the plasma/dark matter system at high redshift, as well as global light propagation and is most powerful when combined with one or more other techniques that probe the nearby Universe.

LISA will use gravitational wave propagation to obtain absolute, gravitationally calibrated luminosity distances, a combination of attributes not shared by any other technique. Even a small number of absolute distances measures H_0, a measurement of central importance. LISA/BHB constraints are similar to the absolute calibration provided by Baryon Acoustic Oscillation technique, but provide a pure-physics calibration not reliant on cosmological models; by combining these two techniques, new constraints can be placed on "dark radiation" such as new types of relativistic particles.

Improved precision in measurements of other quantities, such as absolute and relative distances, growth of structure and fluctuation power spectrum have thus emerged as a top priority of cosmological research. Over the next decade several large programs are being carried forward with this goal (NRC Committee on the Physics of the Universe 2003, Albrecht *et al.* 2006). Each of the proposed techniques has complementary strengths, weaknesses, sources of systematic errors and physical and astronomical assumptions, and thus it is prudent to pursue all of them.

Unique features of black hole binary distances compared with other techniques

LISA's distance measurements to black hole binary (BHB) mergers will independently and precisely measure absolute distances in an entirely new way. Among all the techniques used for cosmological measurement, BHB is unique in several ways, being:

- Physically calibrated assuming only general relativity (no astronomical assumptions about system configuration or environment);
- Absolutely calibrated (a true physical distance in laboratory units based only on gravity — not a distance ratio to an astronomically defined reference);
- Of very high intrinsic precision (ranging from 0.2% to 3% for a single event at $1 < z < 5$); and
- Useful over a very wide range of redshift (detectable BHB

A special challenge is calibration of the large-scale cosmos to absolute (ultimately, laboratory) standards of length or time. Such measurements allow globally-measured quantities (such as CMB angles and galaxy redshifts) to be connected to locally-measured quantities, such as the temperature of the cosmic microwave background, cosmic chronometers, and element abundances. Traditionally this absolute calibration employs a cosmic distance ladder, using direct geometrical parallax measurements of nearby stars to calibrate indirect measures for larger distances, in a series of steps extending to cosmological scales. Other absolute calibrators are now becoming competitive but present major challenges in systematic reliability and precision, and require a variety of assumptions — again, requiring multiple approaches for a robust result (see box on “Absolute Distance Calibration”).

LISA will add a unique and complementary new tool: absolutely calibrated distances determined by measuring the waves generated by massive binary black hole inspiral. Previous sections provide detailed discussion of these waves; for our purpose here, the main point to note is that measurement of these inspiral waves makes it possible to directly determine the luminosity distance to a source with a relative precision that can be as good as 0.1%. The main weakness of this tool relative to other techniques is the still-uncertain number and redshift distribution of events that can provide useful distances, as discussed below. For several applications, the need to identify a host galaxy and corresponding redshift may limit the applicability of this technique. On the other hand the intrinsic precision may be higher than any other technique, possibly in some respects even better than the CMB, and it brings an absolute physical calibration based on gravity alone, unlike any other technique. Even with a small number of events, the unique features of black hole binary inspirals — their reliable absolute calibration, inherent precision, and large range — introduce a new capability that promises to make all other precision measurements more robust and informative.

Absolute cosmometry with black hole binaries

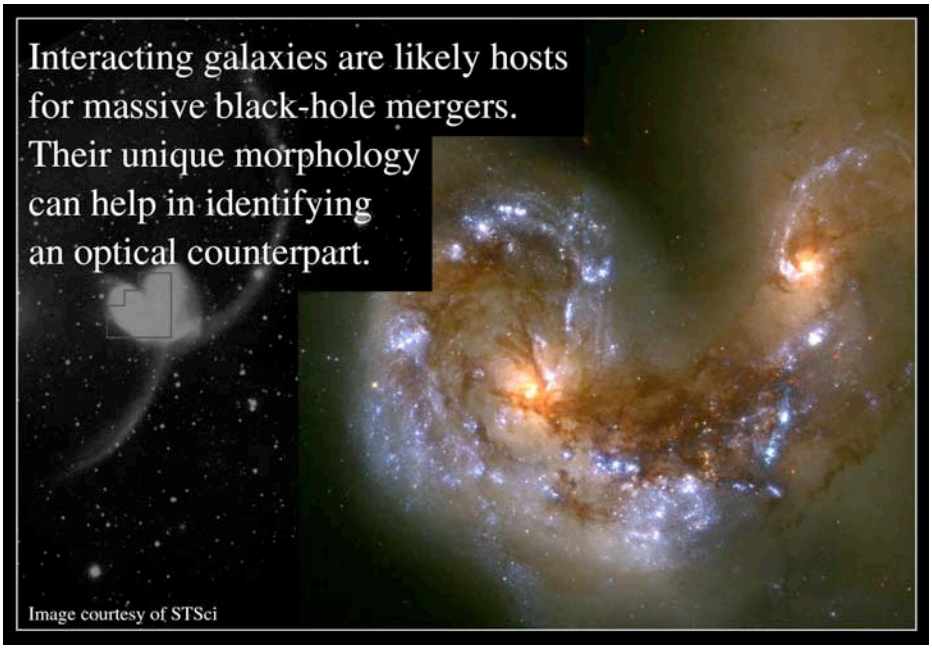
The principle of estimating distances from measured waveforms is elegantly simple (Schutz 1986): the chirping time τ of an inspiral/merger event, together with its orbital frequency ω and strain h, gives an absolute luminosity distance $D \approx c/\omega^2 \tau h$, with a numerical factor depending on details of the configuration that are precisely determined by the measured waveform. (As explained in previous sections, roughly speaking, the directly measured redshifted chirp mass tells the redshifted final absolute Schwarzschild radius, and the ratio of that length to the luminosity distance is the metric strain, h.)

LISA is synergistic with deep multiband imaging, synoptic monitoring, and deep multiobject spectroscopy. LISA data and electromagnetic data will rely on each other and will be used iteratively to create a more powerful probe than either on its own, both of precision cosmology and of the behavior of galactic stars and gas surrounding massive black holes. LISA waveforms provide a final level of precision better than any other technique but require supporting electromagnetic data to get there.

However, as discussed in detail in Section 2, *LISA Mission Overview*, LISA measurements cannot independently determine the redshift of a source. In gravitational wave measurements, the source's intrinsic frequency and chirp time are always measured in combination with cosmic redshift: $\omega = \omega_{int}/(1+z)$, $\tau = (1+z)\tau_{int}$. The redshift is always degenerate with the source's intrinsic parameters, and cannot be determined from the GW data alone. An independent measurement of redshift is therefore needed. This may be accomplished by getting the optical redshift to the host galaxy, for instance by identifying an electromagnetic radiation counterpart to the event.

If a cosmological model is assumed, then a redshift can be inferred self-consistently by requiring agreement between the estimated distance and that predicted from the cosmology. Thus to the extent that cosmological parameters are roughly known, only a certain narrow range of redshift is allowed for the host. Once a host is identified, the precise direction allows a still more precise estimate of distance from LISA. Thus LISA data will be used iteratively and synergistically with electromagnetic data where these are available, to achieve maximal precision.

Identification of the host galaxy to a merger event can occur in many different ways. For an unusually nearby ($z < 1$) event with large black hole mass and signal-to-noise ratio, the LISA error box in angle and redshift might contain fewer than a thousand candidate galaxies, and the host might be easily identified from large scale optical morphology (such as tidal tails) as the site of a recent merger event. More typically, the error box contains of order ten thousand galaxies, and the merger event may not be associated with a major visible disruption; in many cases the redshift is also very high, making optical identification difficult. On the other hand, models suggest that the same merger that creates a black hole binary often sends interstellar gas raining into the new galaxy nucleus, and the same gas that helps the two holes dissipate energy and momentum as they sink towards each other also forms a bright infrared nuclear starburst. For



large binary mass ratios, and total binary masses at the upper end of the range of expected LISA BHB events (more than about $10^6 M_\odot$), the starburst is expected to still be active when the merger event occurs (Dotti *et al.* 2006). Such starbursts will be visible to very high redshift with JWST, and observations in the infrared penetrate deeply even into highly obscured nuclei.

An even more distinctive signature to identify the host galaxy may come from responses of the material near the black holes' horizons to the merger event, which can lead to observable time-variable X-ray/UV emission from the galactic nucleus, modulated by the evolution of the potential as the inspiral progresses. (In the case of obscured nuclei, the time variations may appear in reprocessed infrared emission.) For most LISA-observable holes (less than a few million $M_\odot$), gas accretion disks are evacuated during the merger process but then reestablish emission within a few years, causing an "afterglow" — a newly observable X-ray source (Milosavljevic & Phinney 2005). For higher mass holes, the interaction of accretion discs as the holes approach each other can lead to observable, variable X-ray precursors to the merger event (Dotti *et al.* 2006). If the two holes have disks and/or black-hole-spin-powered jet emission, the disrupted disks and jets can show nonthermal signatures that may appear from radio to gamma rays. Recent evidence of low power AGN (presumably from low mass holes) suggest that such evidence may not be at all rare or unusual. Only a tiny fraction (much smaller than $\sim 10^{-10}$) of the variable gravitational energy of the system needs to appear as a variable electromagnetic signal to be clearly visible.

At redshifts significantly less than unity, the relevant inspiral events consist of a compact stellar mass or intermediate mass black hole (IMBH) captured by a massive black hole in a galactic nuclei. These events will be extremely large mass ratio inspirals (see section on EMRIs). The signal-to-noise ratio and per-event precision are not as high for these as for binaries of comparable mass, but are still possibly good enough (a few percent per event, possibly for hundreds of events) for precision measurement of H_0. Even if electromagnetic counterparts for EMRIs are not found, adequate redshift calibration for a large sample might be obtained statistically from galaxy surveys in the LISA error boxes.

Absolute Distance Calibration

One of the most important sources of error in constraining global parameters at present is the determination of the absolute cosmic distance scale at low redshift, the Hubble constant H_0. This is not well constrained by CMB measurements, which provide an absolute ruler at high redshift, the sound horizon size at recombination r_S=147.8 ± 2.6 Mpc. The CMB data confine models to a narrow degeneracy line, well approximated by Ω_k = -0.3040 + 0.4067 Ω_Λ (Spergel *et al.* 2006) but do not distinguish well between points along this line. A precise measure of absolute distance, even at low redshift, narrows the allowed region to a small interval on this line. With the HST Key Project estimates of H_0, WMAP yields Ω_k =-0.003 ± 0.015, with the error dominated by H_0.

The best distance ladder estimate, the HST Key Project value for the Hubble constant H_0 = 72 ± 8 km/s/Mpc (Freedman *et al.* 2001), still has an uncertainty of more than ten percent. The possibility of systematic errors in this value cannot be completely ruled out, as perhaps suggested by other different HST estimates (Sandage *et al.* 2006) [H_0 = 62.3 ± 1.3 (random) ± 5.0 (systematic) km/s/Mpc] and by recent evidence of a low value ≈ 61 km/s/Mpc from independent Cepheid calibration via eclipsing binaries (Bonanos *et al.* 2006). Other absolute calibration techniques include gravitational lens time delays, and Sunyaev-Zeldovich observations of hot gas in galaxy clusters. So far the systematic errors in these techniques have prevented reliable precision at better than the ten percent level.

A new promising technique now emerging uses very long baseline radio observations of proper motions of distant megamasers in disk galaxies; this method is essentially geometrical, with some assumptions about orbit configuration and motions. Currently the formal error from this method [H_0 = 74 ± 3 (random) ± 6 (systematic) km/s/Mpc based on one object still tied to Cepheids rather than distant Hubble flow (Macri *et al.* 2006)] is comparable to the Key Project error. As more objects are observed in the coming decade, the megamaser technique may achieve a precision at the few percent level at sufficiently large distances that Cepheids or other secondary calibrators will be unnecessary.

The Baryon Acoustic Oscillation (BAO) technique also provides an absolute calibration, based on the known physics of the dark matter/baryon-radiation plasma system around recombination. Currently it offers precision at the 4% level (Eisenstein *et al.* 2005), with prospects for improvement from a larger sample size. This calibration depends on precise understanding of the matter/radiation energy ratio, largely constrained by CMB.

LISA's calibration is based only on the physics of gravitational waves, without reference to cosmological contents, detailed properties of astronomical systems, such as history or environment, or even Standard Model physics. (All of the relevant parameters of the inspiralling black holes can be measured from the waveform and ultimately connect to laboratory measurements of Newton's G and speed of light c). For a single object, the main source of error is noise from gravitational lensing, which can change the measured h and therefore inferred distance by an amount that is unknown for each individual object. This effect is partially correctable using a statistical sample and in any case is small at $z < 1$.

By measuring the inspiral waves, particularly in the case where binaries are of comparable mass, LISA will pinpoint many events to better than a degree (in some cases, much better). LISA will also be able to predict the time of merger well in advance (often, hours to weeks), allowing searches for precursors and afterglows, as well as emission associated with the moment of the merger itself. In general, finding and studying electromagnetic signatures from an active inspiral/merger nucleus requires deep imaging over a wide (~ degree) field, in multiple wavebands. The time variable signatures require sampling on a timescale as long as intervals of years, down to a time resolution scale as short as the ~ 1000 second orbital period. Fortunately, such capability is being created in many bands, including large infrared arrays on large aperture telescopes, extremely capable optical systems such as LSST (which will image galaxies to $z > 1$, in a field 4 degrees across, in less than a minute), extremely wide angle radio interferometry

(being demonstrated in systems such as the Allen Telescope Array), and new capabilities in space from the infrared (JWST) to gamma rays (GLAST).

LISA electromagnetic counterpart sources may provide a rich field of study, with many identifiable hosts and counterparts. LISA sources have possible electromagnetic counterparts over a wide variety of wavebands and timescales: potentially, an exploratory bonanza providing access to new phenomena over a huge range of scales. On the other hand given our ignorance about the processes in galactic nuclei associated with massive black holes and their mergers, it is also possible that the observable electromagnetic signatures may be very rare, in which case precision cosmology with these sources will be impractical.

Another important source of uncertainty is the rate of events, as discussed in Section 3, "Black hole astrophysics". Standard galaxy formation theory suggests that the first black holes formed from the first massive stars, in the first baryonic collapses at redshifts of order 20. Subsequent hierarchical clustering of halos leads to mergers of holes on successively larger scales, eventually forming the population of nuclear massive black holes found today in galactic nuclei. Detailed models of this process (Sesana *et al.* 2004, Volonteri 2006) predict that LISA will see many dozens of massive black hole merger events per year, spread over a range of masses from 10^4 to 10^7 solar masses and a range of redshifts from 1 to 20. On the other hand a more conservative model for the formation of the massive black holes, based only on observed populations of massive black holes, and allowing them to grow as much as possible via accretion rather than mergers, can produce rates an order of magnitude smaller. The total rate and redshift distribution makes a big difference in LISA's capability for precision cosmology. Similarly, the rate of EMRIs is highly uncertain (covering a similar range of possible rates), because of required extrapolation to nuclear stellar populations in galaxies where we have very limited information at present.

The most important practical limit on precision of distance measurements at high redshift comes from gravitational lensing, which can magnify and brighten an object, or demagnify it relative to the mean for a given cosmology, and thus bias its inferred distance. Because high precision is the goal, even small-amplitude modulation by weak lensing is a concern. This effect must be controlled at high redshift by using a statistical sample, as is done for example with supernova distance indicators. The overall precision thus depends on the BHB redshift distribution and event rates.

Impact of LISA/BHB distances: Examples

Since the BHB technique yields independent and physically calibrated absolute distances it complements other techniques of precision cosmology, many of which yield relative distances only, and all of which use different assumptions with radically different systematic errors and biases from BHBs.

The potential contribution of BHB to precision cosmology has not yet been evaluated as thoroughly as other tools, and is subject to the above uncertainties, some of which will not be

resolved until LISA flies. Nevertheless it is useful to cite several benchmark examples of potential impact:

- In a homogeneous universe, the BHB technique with realistic LISA capabilities and source parameters yields absolute distances with a raw per-event precision from about 0.2% to 0.4% at $z = 1$, to about 1% to 2% $z = 3$ to $z = 5$ (Holz & Hughes 2005), depending on the source properties; in some cases, even 0.1% precision is possible, as shown in Figure 5 -1 (Lang & Hughes 2006).

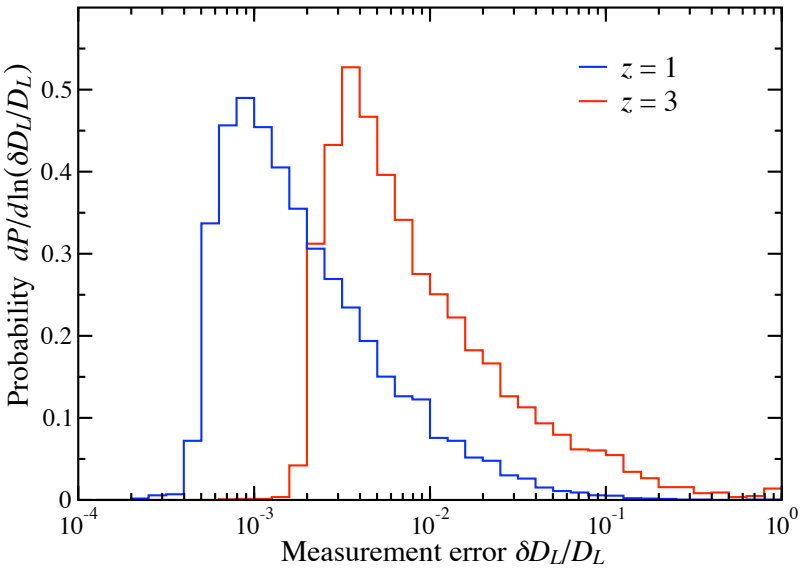


Figure 5-1: Simulated distribution of errors in BHB luminosity distances D_L at $z = 1$ (blue) and $z = 3$ (red), with black hole masses of 10^5 and $6 \times 10^5 M_\odot$, assuming that an electromagnetic counterpart allows precise sky position determination (Holz & Hughes 2005). Even at $z = 3$ most measurement errors are less than 2%.

- Since most lines of sight to $z < 1$ are nearly empty of dark matter, lensing errors are relatively small. A few massive inspiral events, or a large number of EMRI events, at redshift less than or of order unity may lead to a reliable measurement of the Hubble constant, H_0 to better than 1%, an order of magnitude improvement over current techniques. Absolute gravitational calibration adds unique information even if other techniques attain comparable formal precision with different assumptions.

- Calibration of the absolute distance scale, in combination with CMB measurements alone, and a definite scaling law for the dark energy *w(a)*, allows a determination of *w* with high precision (Hu 2005, Eisenstein & White 2004). With Planck quality CMB data alone, calibration of the Hubble constant at 2% precision achieves *w* accuracy of better than 3% (Olling 2007). Similarly, a one percent constraint on absolute distance, combined with the CMB data, yields $\sim 10^{-3}$ error on global curvature Ω_k (Knox 2006); in this respect the constraints are similar to those obtainable in the future from Baryon Acoustic Oscillations (Eisenstein *et al.* 2005). With both techniques (that is, LISA and BAO together), a new tight constraint can be derived on the density of any invisible relativistic species or "dark radiation" that affects the BAO calibration.

- LISA/BHB distances are also useful at moderate *z* (of order 2 to 3) if the BHB sample is large enough for fitting to a lensing amplification model. Even with conservative assumptions about lensing noise, a sample of 100 LISA/BHB events provides comparable precision to a sample of ≈ 3000 SNIa distances at comparable redshift (Dalal *et al.* 2006).

• The considerable number of events from higher redshift, out to $z \sim 20$, will provide measurements of black hole mass mergers as a function of distance and measurements of integrated weak lensing along multiple lines of sight to high redshift. The galaxy counterparts of these sources will be infrared galaxies, many of them in very early stages of assembly, and many of them potentially observable with JWST. LISA's observations will complement the revolutionary JWST views of early structure and galaxy formation, adding detailed information about the early growth of nuclear regions, massive black holes, and dark matter density fluctuations at high redshift.

LISA Science Objectives and Investigations relevant to this section (see Appendix 1)

6. Probe new physics and cosmology with gravitational waves

6.1. Study cosmic expansion history, geometry and dark energy using precise gravitationally calibrated distances and redshifts measured from electromagnetic counterparts or determined statistically.

6. Ultra-Compact Binaries

Summary

Close binary stars are double stars in which two compact objects such as white dwarfs and neutron stars, orbit each other in a short period. Binaries with orbital periods below a few hours emit gravitational radiation in the LISA band and are generally called ultra-compact binaries. They have relatively weak gravitational wave signals in comparison to massive black hole binaries, but are numerous in the Galaxy and even the Solar neighborhood. The prospects for LISA in this area of astrophysics are truly spectacular. Several thousand systems are expected to be detected individually, with their parameters determined to high precision, while the combined signals of the millions of compact binaries in the LISA band will form a well detectable background signal. This should be compared to less than 50 ultra-compact binaries known today. The shear number of detections will allow study of entire populations of binaries in such detail that not only the properties of this particular population can be determined, but through comparison with models, the formation of these binaries and thus many preceding phases in binary evolution can be strongly constrained. This has a strong bearing on our understanding of many high-energy phenomena in the Universe such as supernova explosions, gamma-ray bursts and X-ray sources as they share parts of the evolution history of the binaries detectable by LISA.

The number of sources in the LISA band also ensures that many of them can be detected at high signal-to-noise ratio (often larger than 100), allowing for detailed study of individual binaries. For many the

Key science questions

- Is general relativity the correct theory of gravitation?
- Is there a large population of ultra-compact binaries in the Galaxy?
- How did compact binaries form and what is the outcome of a common-envelope phase?
- What is the nature of the fundamental physical interactions in compact binaries?
- How are the compact binaries distributed in the Galaxy and what does that tell us about the formation and evolution of the Galaxy?

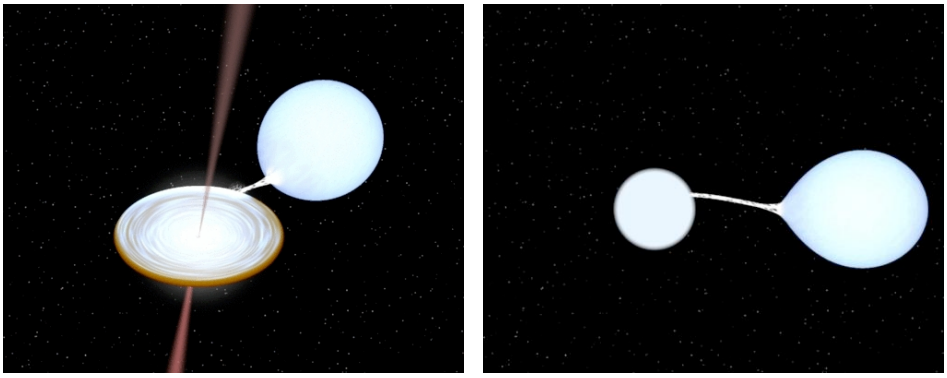


Figure 6-1: Impression of two types of ultra-compact binaries: (left) an ultra-compact X-ray binary in which a neutron star accretes from a compact companion; (right) an AM CVn binary in which a white dwarf accretes from a compact companion.

frequency and phase evolution can be studied. This will enable the study of the physics of tides and mass transfer in unprecedented detail. The extreme conditions of short orbital periods, strong gravitational fields and high mass-transfer rates are unique in astrophysics.

The LISA measurements will provide very different information compared to what can be deduced from electromagnetic detections. In particular LISA's capability to determine distances as well as the fact that the gravitational wave signals are unaffected by interstellar dust provide huge advantages over other detection techniques. Compared to GAIA, LISA will detect a quite different population, with most distance determinations right in the Galactic center — where most ultra-compact binaries are — rather than between the Sun and the Galactic center. The distance determinations will make it possible to map the distribution of many compact binaries in the Galaxy, providing a new method to study Galactic structure. On the other hand, many of the detectable binaries can be studied electromagnetically thus providing a strong complementarity if both techniques are combined. Dedicated complementary observing programs as well as public data releases will ensure simultaneous and follow-up electromagnetic observations.

The importance of electromagnetic observations

The LISA measurements will provide complementary information to what can be deduced from electromagnetic detections. In particular its capability to determine distances as well as the fact that the gravitational wave signals are unaffected by interstellar dust provide huge advantages over other detection techniques. On the other side, the positional accuracy of LISA is quite limited compared to electromagnetic detectors. This highlights the importance of coordinated gravitational wave and electromagnetic observations, a unique possibility when studying gravitational waves at low frequencies. Although the majority of the binaries discovered by LISA will not likely be detected with electromagnetic detectors due to the dust absorption, the shear number of expected gravitational wave detections will ensure that there are still many tens or hundreds for which it is possible to do complementary electromagnetic observations. This will yield not only their position in the sky — thus putting strong priors on the gravitational wave parameter estimation — but also provide additional information, ranging from temperature and age estimates if only broadband colors can be measured, to individual masses and even radii if the system is eclipsing or spectra can be taken.

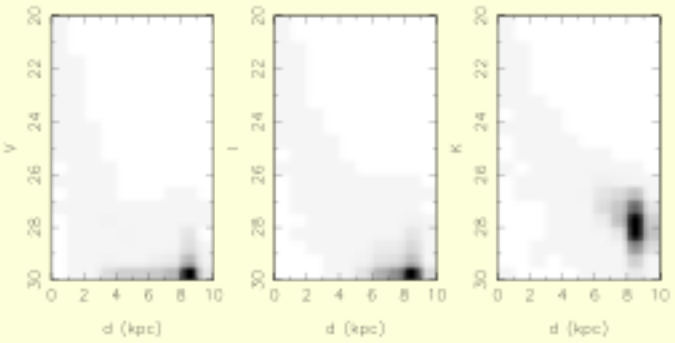


Figure 6-2: Optical and infrared magnitudes of the expected double white dwarf binaries detected by LISA.

A number of guaranteed detectable sources are known to date. Some of these can be used to verify instrument performance by looking for a gravitational signal at twice the orbital period and comparing the signal to expectations. In addition, once LISA has detected several nearby binaries and determined their sky position they can be observed optically thus providing an additional quantitative check on instrument sensitivity. Among the currently identified sources are the two most exotic but controversial binaries known, possibly with ultra-short orbital

Tutorial: Ultra-compact binaries

At the end of the 1960s, binary stars with periods of less than one hour were discovered, and astrophysicists recognized that the two stars are so close together that an ordinary star like the Sun could not fit in their orbits. The close proximity of the objects suggests that both components are remnants: white dwarfs, helium stars, neutron stars, or even black holes that form after stars exhaust their nuclear fuel. The components of ultra-compact binaries are stars that have evolved beyond hydrogen-core burning.

A useful distinction can be made between systems in which the two components are far apart (detached systems) and systems in which the components are so close together that mass is being transferred from one star to the other (interacting systems). Examples of detached systems are double neutron stars and double white dwarfs. Their existence in the LISA band is inferred from longer period systems that are known and are evolving towards shorter periods, such as the double pulsar PSR J0737 -3039 and the double white dwarf WD 0957-666.

Currently some 30 interacting ultra-compact binaries are known. They naturally divide in two classes, based on their observed characteristics (Figure 6-1). In ultra-compact X-ray binaries, gas falls onto a neutron star whose potential well is so deep that the gas heats up to millions of Kelvin and produces abundant X-rays. The second class comprises the AM CVn stars, so named because the first of them was a variable star discovered in the constellation Canes Venatici (hunting dogs). In AM CVn systems gas falls onto a white dwarf, which has a much shallower potential well than a neutron star. The gas is heated to only about 100 000 K, and most radiation is emitted at optical and UV wavelengths in the range of 100–600 nm. The number of known systems is increasing regularly due to new discoveries or period determinations of X-ray sources that turn out to be ultra-compact X-ray binaries.

periods. As LISA will really open a new window to the Universe, arguably the most exciting prospect for LISA is the detection of many additional completely unknown and highly exotic sources in the Galaxy.

LISA as a workhorse: thousands of new binaries

Ultra-compact binaries (see Tutorial Box: Ultra-compact binaries) will completely dominate the number of source detections by LISA. Current estimates suggest the number of resolved compact binaries that will be detected by LISA will be counted in thousands, if not ten-thousands. These systems will be dominated by the shortest period systems and most will have periods less than about 20 min (Figure 6-3). If we compare that with the fact that less than 50 ultra-compact binaries[4] are known today — less than a handful with periods shorter than 20 min — it is clear that the knowledge of this population will be revolutionized by LISA. As these systems are relatively short lived and faint, there is no hope to detect these system with any other means than via gravitational radiation. Their detection will allow us to test different models for the common-envelope phase: the dominant uncertainty in our understanding of binary evolution and many high-energy phenomena. The internal statistical accuracy delivered by the shear number of detected sources will ensure that the common-envelope phase will be put to the most critical test that can be expected in the midterm future of astrophysics.

[4] http://www.astro.ru.nł-nelemans/dokuwiki

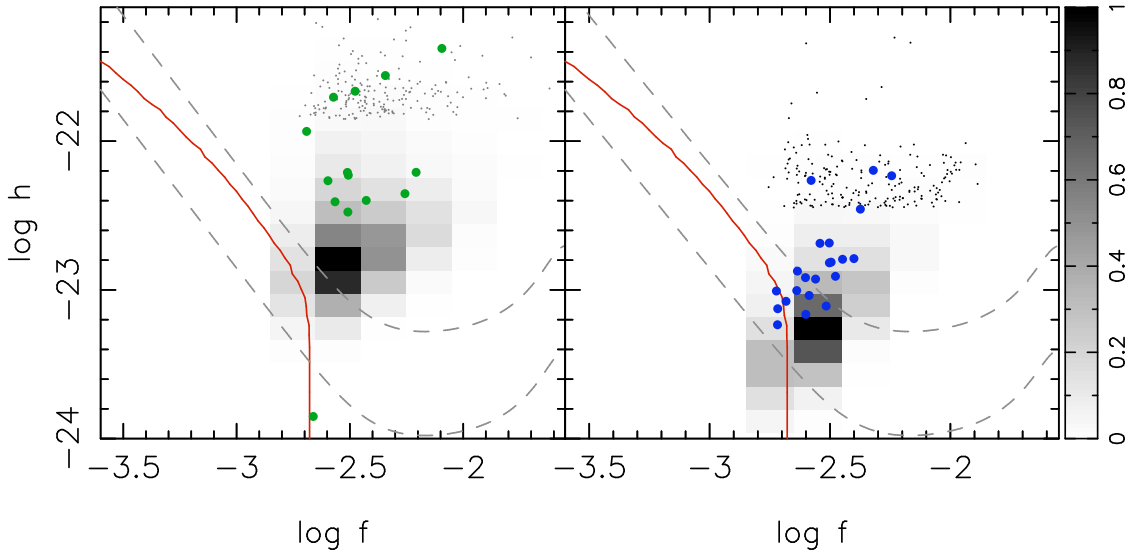


Figure 6-3: GWR wave amplitude *h* as function of the GWR frequency *f* (Hz) for the Galactic binaries that are expected to be detectable by LISA. The left panel shows the (10658) double white dwarf systems as the grey shade, with the 200 strongest sources as points, to increase their visibility. The right panel shows the (9831) resolved AM CVn systems that are expected, again showing the 200 strongest sources as points. Over plotted with the large symbols are the neutron star binaries in the left panel and the ultra-compact X-ray sources in the right panel. The average double white dwarf background is plotted as the solid line, while the dashed curves show the LISA sensitivity for a integration time of 1 year giving S/N of 1 and 5 respectively. Based on Nelemans *et al*. (2004).

Key question: the outcome of the common envelope phase

Less than half of the stars in the Universe are single, the majority being part of a binary, a triple or a higher order system. Of the binaries on the order of half are formed with a sufficiently small orbital separation that during the evolution of the components into giants or super giants the stars will interact. Especially for low-mass stars the majority of interactions are unstable and will lead to runaway mass transfer. Based on the observed short orbital periods of binaries that have passed this stage it is argued that somehow the companion of the giant ends up inside the giant's outer layers. During that common envelope phase, friction reduces the velocity of the companion, leading to orbital shrinkage and transfer of angular momentum from the orbit into the envelope of the giant. Along with angular momentum, orbital energy is deposited in the envelope, whose matter is then unbound from the giant's core, leading to a very compact binary consisting of the core of the giant and the original companion (Paczynski 1976).

Virtually all compact binaries and most of the systems giving rise to high-energy phenomena (such as X-ray binaries, relativistic binary pulsars and possibly gamma-ray bursts) have experienced at least one common-envelope phase. Given the importance of this phase in high-energy astrophysics our understanding of the physics and our ability to predict the outcome of

Tutorial: Evolution of a star

Based on the fundamental set of structure equations — balancing gravity with pressure and accounting for energy generation and transport — it has been possible to construct stellar models that compare very well to the observed stars in the local Universe. Stars first stay on the main sequence, when they are burning hydrogen in their high pressure and temperature cores just as our Sun is currently doing. The nuclear burning, however, cannot go on forever. In time, no hydrogen remains in the core where the nuclear burning takes place. Once the hydrogen is gone, the pressure in the core cannot be sustained; gravity wins the battle of forces, and so the core contracts. Meanwhile, the outer layers of the star expand, typically by an impressive factor of 10 to 100, and the star becomes a red giant. Stars less massive than about 8 times the mass of the Sun lose their outer layers into space due to a stellar wind. All that remains is a compact degenerate core, about the size of Earth but typically 300 000 times more massive, forever cooling and visible as a small, initially hot star called a white dwarf.

More massive stars will continue nuclear fusion and their core will be burnt into heavier and heavier elements until it is made of iron. But neither fusion nor fission can produce energy from iron, so the iron core has no way of counterbalancing gravity and it collapses. The extreme pressure forces the electrons and protons to combine into a degenerate gas of neutrons that, if the core is not too heavy, generates enough additional pressure to halt the contraction. The resulting neutron star has a radius of about 10 km and a mass of 3×10^{30} kg. If the core is massive enough, even the pressure of the degenerate neutrons cannot offset gravity, and the core collapses into a black hole. Neutron stars are formed from the cores of stars with masses between about 10 and 25 solar masses; black holes, as far as we know, are created from stars with masses greater than about 25 solar masses. During the formation of either a neutron star or a black hole, the outer layers of the star are blown into space in a violent explosion called a supernova.

the common-envelope phase are alarmingly poor. Theoretical progress to understand the phase from first physical principles is slow (*e.g.* Taam & Sandquist 2000) and observational tests have led to questioning of the applicability of the standard formalism described above. Comparison of the parameters of the thousands of binaries detected by LISA with model predictions will provide a direct test of the different proposed outcomes of the common-envelope phase and our understanding of the preceding binary evolution in general.

Formation of ultra-compact binaries in globular clusters

Of special interest are globular clusters, dense assemblies of up to a million stars that represent the oldest stellar populations in the Galaxy. The angular resolution that can be achieved with LISA is such that globular clusters can be resolved, so that the cluster sources can be distinguished from the Galactic disc sources (Figure 6-4). Globular clusters have a strong overabundance of bright X-ray sources probably due to dynamical interactions. In particular it seems that the number of ultra-compact X-ray binaries is severely enlarged compared to the other Galactic populations. It is not clear if ultra-compact binaries with white dwarf components are overproduced as well. LISA will directly test this by determining the number of ultra-compact binaries in globular clusters. This will provide unique information, since the details of how these interactions lead to the formation of ultra-compact X-ray binaries are poorly understood (Verbunt & Lewin 2006).

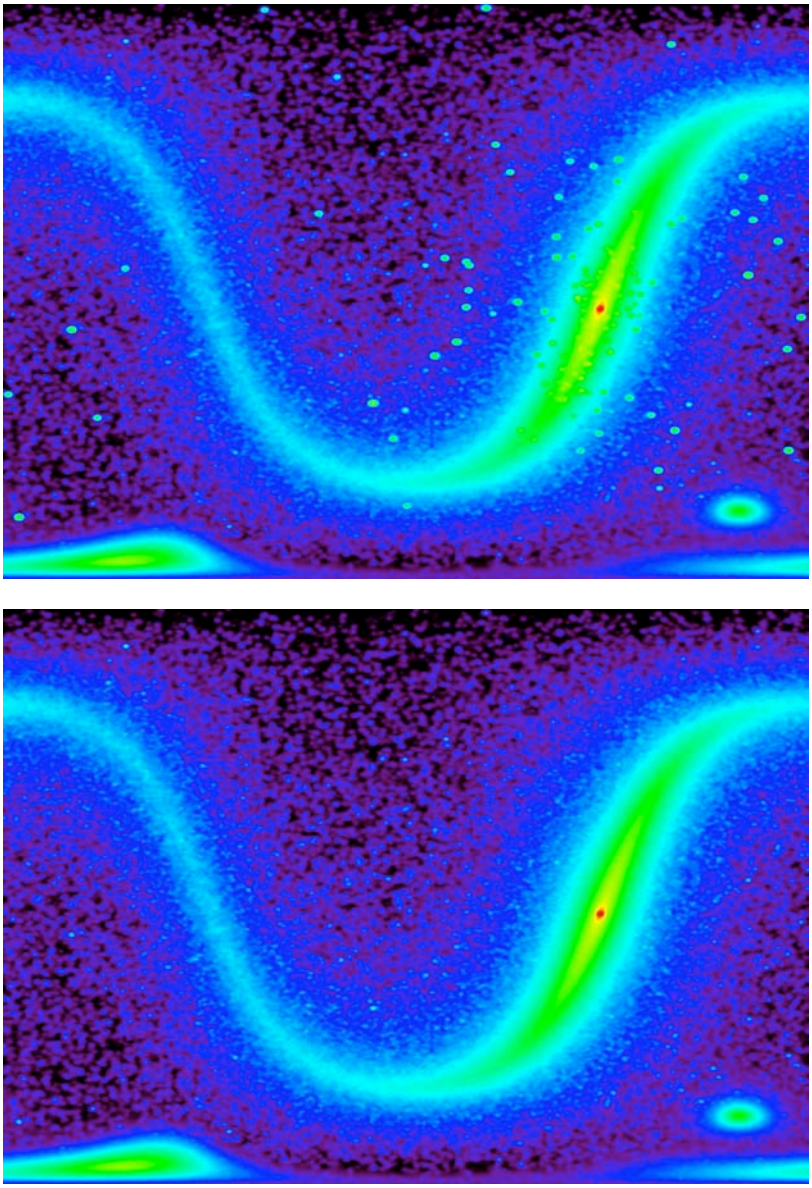


Figure 6-4: Sky brightness in GWR as seen through LISA eyes showing the Galactic disc and bulge, globular clusters, LMC and SMC. In the top plot the number of ultra-compact binaries in globular clusters is assumed to be enhanced compared to the field making the globular clusters stand out, while in the bottom plot the number is not enhanced.

The foreground of Galactic gravitational waves

At frequencies below a few mHz the number of sources in the Galaxy is so large that only a small percentage, the brightest (and closest) sources, will be individually detected. The vast majority will form an unresolved foreground signal in the detector. (We distinguish the Galactic "foreground" from the diffuse extragalactic "background".) This foreground is often described as an additional noise component which is misleading for two reasons. The first is that there is a lot of astrophysical information in the foreground. The overall level of the foreground is a measure of the total number of ultra-compact binaries, which is very valuable information given the current uncertainty levels in the normalization of the population models. The spectral shape of the foreground also contains information about the homogeneity of the sample, as simple models of a steady state with one type of binary predict a very distinct shape. In addition, due to the concentration of sources in the Galactic center the foreground is strongly modulated during the year, with periods in which the foreground is more than a factor 2 lower than during other periods (Figure 6-5). The cyclo-stationary character of the foreground allows for further tests of the population models and thus the uncertainties in binary evolution theory *e.g.* Edlund *et al.* (2005)

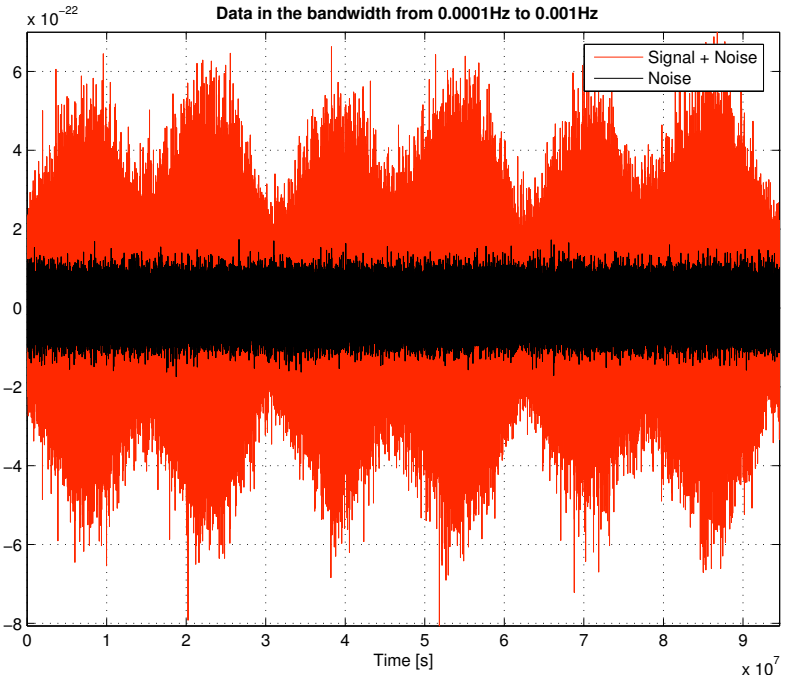


Figure 6-5: Level of the Galactic gravitational wave background as function of time, showing the yearly change. From Edlund *et al.* (2005)

The extragalactic ultra-compact binary background

The masses of white dwarfs and their minimal separation set a maximum to the signal strength of an individual binary, which puts the distance of individual binaries detectable by LISA to less than 100 kpc. This means the LMC and SMC are just within reach, but no individual ultra-compact binaries can be detected further away. However, the combined extragalactic signal — dominated by the signals from sources around redshift of 1 — is expected to be just around the limit of the sensitivity of LISA at a few mHz (Farmer & Phinney 2003). So if the resolved binaries in the Galaxy can be effectively removed, (non) detection of this extragalactic compact-binary background will provide insight in the cosmic star formation provided that the Galactic binaries have greatly reduced the uncertainties in the binary evolution that also hinder the extragalactic background predictions.

Verifying LISA and testing GR

The first and arguably most fundamental test that LISA will perform is direct measurement of gravitational radiation, although this might well have happened earlier using ground-based detectors. However, LISA is able to test a much simpler and cleaner system than the ground-based detectors by measuring the monochromatic signal of a detached binary. The brightest expected sources have signal-to-noise ratios larger than 100 and will allow detailed tests of the monochromatic and higher order terms. For the most compact high-frequency sources (above a few mHz) the first and second derivative of the frequency also will be measured. In addition, the most exciting possibility and expectation is the unexpected. LISA will likely discover objects and phenomena that are completely unknown and might shed new light on our understanding of the Universe.

Verification binaries

A subset of the known ultra-compact binaries have been recognized as ideal verification sources, as they should be detected in a few weeks to months and thus can be used to verify the performance of the instrument (Stroeer & Vecchio 2006). The reliability of the verification binaries has been improved recently by determinations of distances and systems parameters thus providing predictions of the expected signals with well defined error bars (Figure 6-6). Their expected monochromatic nature within the LISA mission time prevents astrophysical effects (see next section) hampering their detection.

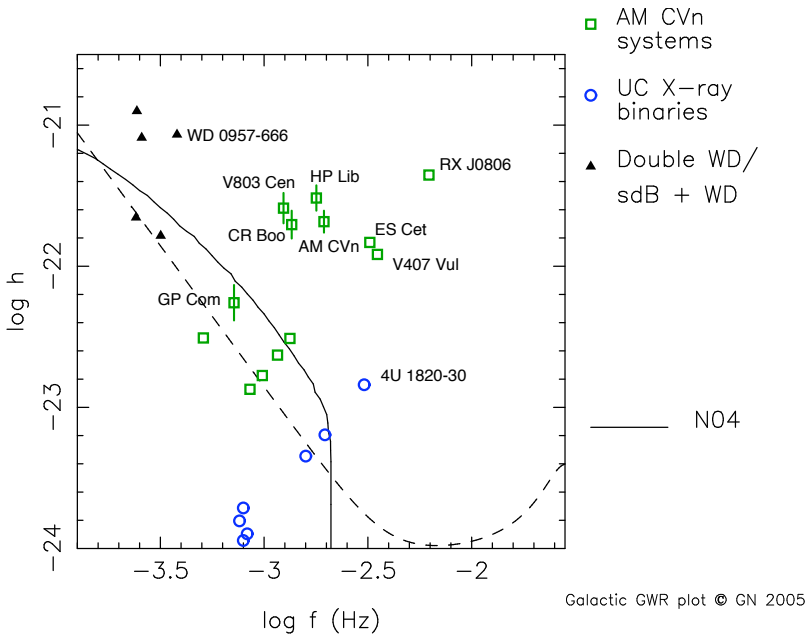


Figure 6-6: Gravitational wave strain versus frequency for the verification binaries. The instrument sensitivity and the average Galactic background are also plotted. Based on Roelofs *et al.* (2006)

The LISA measurements will immediately test the determination of the system parameters and, more importantly, will provide the definitive answers the debate about the nature of the two shortest period binaries RX J0806.3+1527 and V407 Vul. These have received special attention in the past few years because they show repeating signals with periods of 5.4 and 9.5 minutes. In some interpretations, those times are the remarkably small orbital periods of binaries in which two white dwarfs are separated by about a quarter of the Earth Moon distance. A good deal of debate and uncertainty attends the two systems, and several competing theories purport to explain them, ranging from a detached pair of magnetic white dwarfs in which their X-ray emission is produced by induction, to direct accretion models in which the observed periods are not orbital periods at all (see Nelemans 2006). No model is strongly favored, and we might have to wait until LISA is launched before the two objects are understood.

Outlook: expected developments in the next decade

There are a number of initiatives that will improve our knowledge of ultra-compact binaries in the next decade, before LISA will fly. One of the major contributions to the increase in the number of known AM CVn systems in the last years has been the Sloan Digital Sky Survey (SDSS), in which 6 new systems have been found. This number will certainly go up with the

ongoing extension of the SDSS to lower Galactic latitudes (the SEGUE survey) and the European Galactic Plane Surveys, two surveys that are particularly well designed for finding compact binaries. However, all SDSS systems have relatively long orbital periods (longer than about 30 minutes). Two surveys capable of finding AM CVn stars with periods less than 30 minutes are underway or will start soon: the Rapid Time Survey (RATS[5]) and OmegaWhite[6].

For the future there are initiatives to find more ultra-compact X-ray binaries both through the continued monitoring of the sky to search for X-ray transients with RXTE and other satellites, as well as through dedicated X-ray and optical surveys of the Galactic Bulge that are currently in development.

With these developments the number of verification sources available for LISA will be several tens allowing detailed tests of the performance of the instrument.

Studying the astrophysics of compact binaries using LISA

Although the effect of gravitational radiation on the orbit will dominate the evolution of the binaries detected by LISA, additional physical processes will cause strong deviations from the simple point mass approximation. The two most important interactions that occur are tides — when at least one of the stars in a binary system is not in co-rotation with the orbital motion or when the orbit is eccentric — and mass exchange. Because many binaries will be easily detected, these interactions do not hamper their detection, but instead will allow for tests of the physics underlying these deviations. Gravitational wave measurements will provide a completely complementary approach to many aspects of astrophysics because they are optimal for the study of short period systems in contrast to the current bias towards bright electromagnetic systems and events.

Physics of tidal interaction

LISA measurements of individual short period binaries will give a wealth of information on the physics of tides and mass transfer (stability). For detached systems with little or no interaction, the frequency evolution is well determined. The strain amplitude h, the frequency f and its derivatives are connected by:

$$h \propto M^{5/3} f^{2/3} d^{-1} \quad (1)$$

$$\dot{f} \propto M^{5/3} f^{11/3} \quad (2)$$

$$\ddot{f} = \frac{11}{3}\frac{\dot{f}}{f} \quad (3)$$

[5] http://www.mssl.ucl.ac.uk/~gtbr/rats.html

[6] http://www.astro.ru.nłomegawhite/

where M is the chirp mass and d is the distance. Thus measurement of h, f, $\dot{f}$ provides chirp mass and distance. Additional measurement of $\ddot{f}$ gives a direct test of the dominance of gravitational wave radiation in the frequency evolution (third equation). Tidal interaction between white dwarfs in detached systems before the onset of mass transfer will give rise to distinct deviations of the frequency evolution as compared to systems with no or little tidal interaction. The strength of the tidal interaction is virtually completely unknown, with estimates ranging over many orders of magnitude (Marsh *et al.* 2004). Knowledge of the strength of the tides is not only important for the understanding of the physics of tides in general and the physics of white dwarf interiors, but has important consequences for the tidal heating (and possibly observability) of LISA sources and the stability of mass transfer between white dwarfs (Racine *et al.* 2006).

Physics of mass-transfer stability

Detached ultra-compact binaries will evolve to shorter and shorter periods due to the angular momentum loss through gravitational wave radiation. At sufficiently short orbital period (a few minutes) one of the stars becomes larger than its Roche lobe — the equipotential surface that crosses the minimum of the potential between the two stars — and material "leaks" out of the potential well of one star onto the other star. Depending on the difference between the change of the radius of this star and the Roche lobe upon mass transfer, there may be positive or negative feedback, leading to either limited, stable mass transfer, or a runaway mass-transfer instability.

For double white dwarfs and white dwarf - neutron star binaries the stability of the ensuing mass transfer has important consequences, both for the number of expected mass-transferring LISA sources as well as a number of open astrophysical questions. The stable systems will form detectable interacting binaries (AM CVn systems or ultra-compact X-ray binaries). LISA will detect a number of double white dwarfs and AM CVn systems that are so close to the onset of mass transfer that the stability of the mass transfer can be tested directly by comparing the two numbers. In addition, LISA will detect several ultra-compact X-ray binaries at the very early stages of mass transfer, providing a test of the mass transfer stability in these systems as well.

For AM CVn systems the dominant uncertainty in the mass-transfer stability is again the tidal interaction between the two white dwarfs. Most likely the mass transfer will proceed via the direct impact configuration: due to the proximity of the two stars the mass transfer stream lands directly on the surface of the accreting white dwarf, rather than wrapping around the accreting stars, interacting with itself to form a flat accretion disk in the plane of the orbit (Webbink 1984). It turns out that the stability of the mass transfer depends critically on the tidal interaction between the two white dwarfs (Marsh *et al.* 2004). In absence of any tidal interaction there will be additional angular momentum loss from the orbit due to the transfer of angular momentum from the orbit to the accreting star which will consequently spin-up. This is different from the cases where the accretion is via a disc for which most of the angular momentum generally is stored in the disc and eventually via very efficient tidal interaction put back into the orbit.

However, efficient tidal coupling between the accreting star and the companion might return the angular momentum back to the orbit as well (Racine *et al.* 2006, d'Souza *et al.* 2006; see Figure 6-7), thus reducing the magnitude of the spin-up.

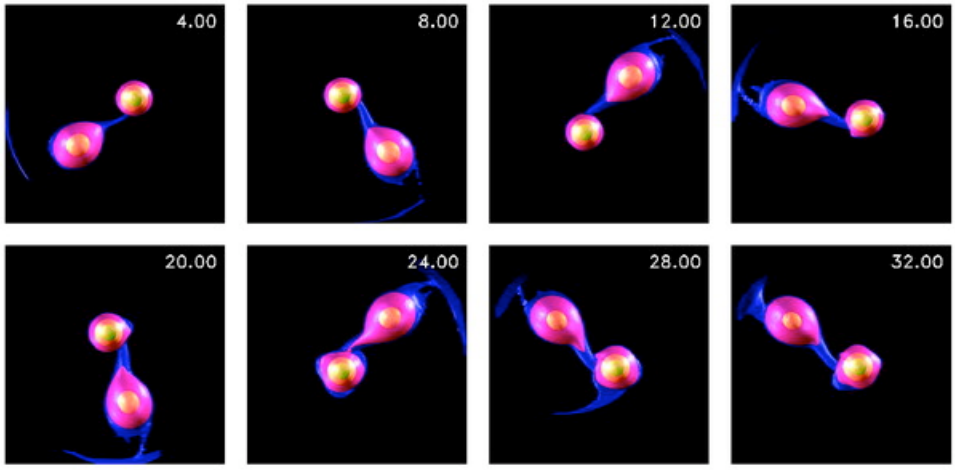


Figure 6-7: Numerical simulation of the stability of mass transfer in a binary with mass ratio 0.5, in which the mass transfer is stable (d'Souza *et al.* 2006).

The difference between efficient and inefficient tidal coupling is rather dramatic: the fraction of double white dwarfs estimated to survive the onset of mass transfer can drop from about 20 percent to 0.2 percent (Nelemans *et al.* 2001) depending on assumptions about the tidal coupling. This difference is very easily measurable with LISA.

For ultra-compact X-ray binaries, the stability issue is completely different. At the onset, the mass transfer is orders of magnitude above the Eddington limit for a neutron star (the mass transfer rate at which the potential energy liberated in the accretion can couple to the infalling gas to blow it away). For normal stars and white dwarfs this would likely lead to complete merger of the system, but the enormous amount of energy liberated when matter is falling into the very deep potential well of a neutron star allows matter to be dumped on it at rates up to a thousand times the Eddington limit. This allows the formation of ultra-compact X-ray binaries from white dwarf-neutron star pairs, if the white dwarf has a low mass (see Yungelson *et al.* 2002). LISA will unambiguously test this prediction, by detecting a several tens of ultra-compact X-ray binaries with periods between 5 and 20 minutes (Figure 6-3).

Double white dwarf mergers

The 80 to 99.8 percent of the double white dwarfs that experience run-away mass transfer and merger likely give rise to quite spectacular phenomena. Although it is not expected that LISA will be able to witness the actual merger of a double white dwarf it will certainly detect the shortest period binary known, expected at a period of about 3 minutes and give an extremely good estimate of their merger rate. Mergers of double white dwarfs have been proposed as progenitors of single subdwarf O and B stars, R Corona Borealis stars and maybe all massive white dwarfs. The merger of a sufficiently massive double white dwarf has been proposed as the trigger for type Ia supernova events, the standard candles that play an important role in our

understanding of the expansion of the Universe and the existence of dark energy. Alternatively, if the merger does not lead to an explosion a (rapidly spinning) neutron star will be formed. This is one possible way to form isolated millisecond radio pulsars as well as magnetars, which have been proposed as sites for short gamma-ray bursts *e.g.* Levan *et al.* (2006). LISA will be able to put strong constraints on these hypotheses. By measuring (chirp) masses and coalescence times LISA will directly determine the merger rate for double white dwarfs with different masses which can then be compared with the rates and population of their possible descendants determined by other means.

Neutron star and black hole binaries

In addition to the thousands of white dwarf binaries, several tens of neutron star binaries will be detected by LISA, if the current observational and theoretical estimates of the formation rate of neutron star binaries turn out to be reliable (*e.g.* Nelemans *et al.* 2001). For all these populations, the sample will be complete at the shortest periods as the systems can be seen throughout the Galaxy and thus the sample will be independent of selection effects such as those present in radio pulsar surveys which currently hamper our knowledge about this population in the Galaxy. In addition, by the time LISA will fly, Advanced LIGO will likely have detected a number of double neutron star mergers from far away galaxies, so these measurements together will test our ability to extrapolate our population models from our own galaxy to the rest of the Universe.

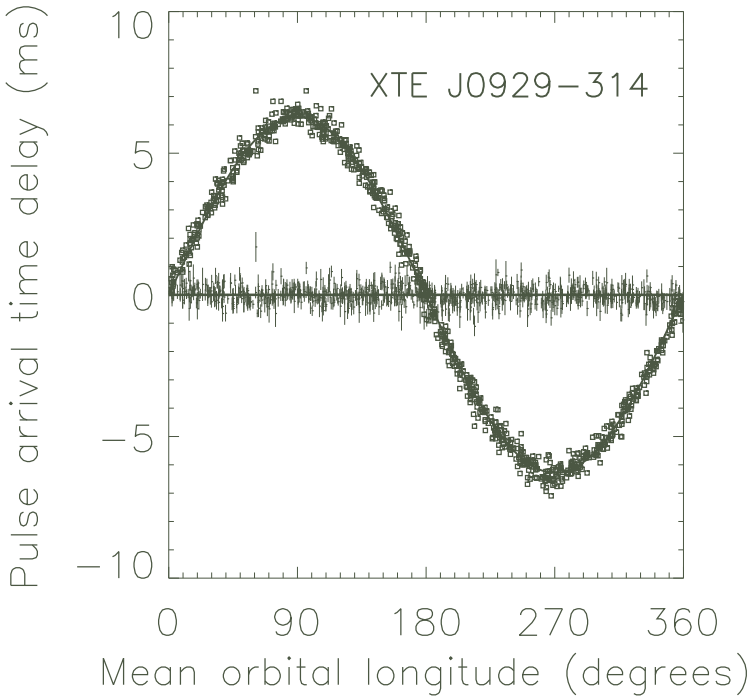


Figure 6-8: Imprint of the 40 min orbital period on the arrival times of the X-ray pulsations in the ultra-compact X-ray binary XTE J0929-314. From Galloway *et al.* (2002)

Another opportunity for LISA to test a long-standing hypothesis concerning neutron stars is provided as LISA will not be biased towards detecting transient systems that suddenly show up as bright X-ray sources. A very special situation might arise in the

millisecond X-ray pulsars in ultra-compact X-ray binaries. In recent years, observations of X-ray pulsations from many ultra-compact X-ray binaries have enabled astrophysicists to determine the rotation rate of the neutron star using the NASA mission RXTE (Figure 6-8). As had been expected on theoretical grounds, the neutron stars are spinning rapidly several hundred times per second due to the angular momentum gained from infalling matter. The measurements give credence to the idea that rapidly spinning neutron stars observed as millisecond radio pulsars are descendants of accreting neutron stars in binary systems (*e.g.* Bhattacharya & van den Heuvel 1991). However, scientists have yet to establish the exact role of ultra-compact binaries in the formation of these pulsars. The distribution of spin periods discovered in X-ray binaries suggests additional neutron star angular momentum loss on top of the plasma physics interaction between the accretion and magnetic field of the spinning neutron stars (Chakrabarty *et al*. 2003). One exciting possibility is that this additional angular momentum loss is due to strong gravitational wave emission (Bildstein 1998). In that case ultra-compact X-ray binaries might be the only sources that could be studied simultaneously with LISA and Advanced LIGO, with LISA detecting the orbital period and LIGO detecting the neutron star period.

If they exist in sufficient quantities, LISA will easily pick up black hole binaries in the Galaxy. There is a lot of uncertainty about the number of ultra-compact stellar mass binary black holes in the Galaxy *e.g.* Belczynski *et al*. (2002) but if they exist they will show up in the LISA band and indeed this is the only way such systems may ever be detected.

New studies of galactic structure with LISA

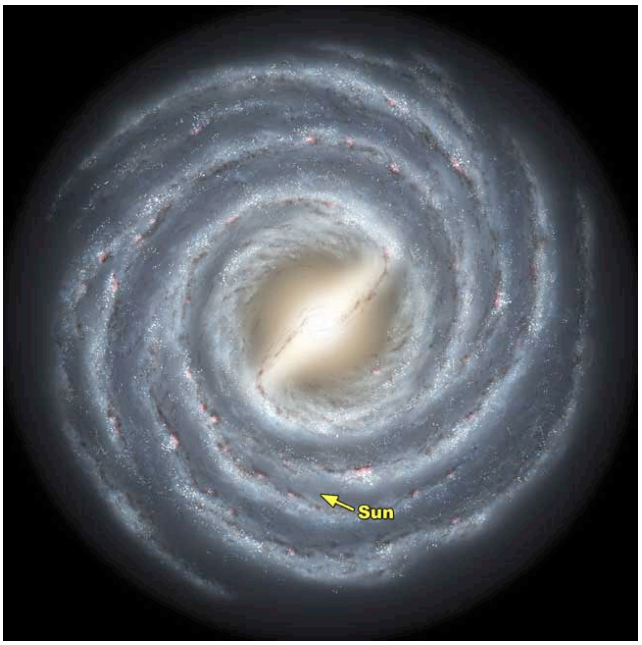
One of the major advantages of LISA is that it will determine distances for more than 100 compact binaries by measuring their $\dot{f}$ (Equation 2). The exact ability of LISA to determine distances depends critically on the mission lifetime, as larger life times lead to more accurate $\dot{f}$ measurements.

The directional dependence of the Galactic background as well as the directional accuracy for the resolved systems allow for a statistical assessment of the contributions of the different Galactic components such as the Galactic bulge (with its bar), the thin and thick disc and especially the Galactic halo.

Galactic halo

The halo sources will of course be distributed completely differently from the other Galactic components. The LISA directional sensitivity will immediately pick up any strong halo population. The halo population might be enhanced compared to the disc as the formation and evolution of binaries in the halo may have been quite different. Such old and metal poor population can locally only be studied in globular clusters, where the formation and evolution of

binaries is generally completely altered by dynamical effects. There is currently a very interesting suggestion. Two of the 18 known AM CVn systems seem to be most consistent with belonging to the halo. They have very low metal abundances and anomalous velocities. If true this implies a very large number of AM CVn stars in the halo, maybe as many as in the rest of the Galaxy. LISA would certainly pick up the short period systems in this population.



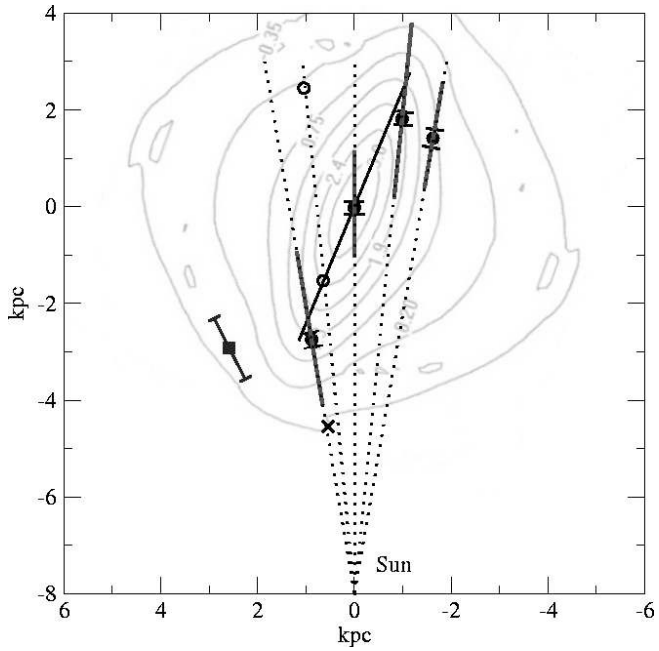


Figure 6-9: Different models for the Galactic bar: above the Spitzer GLIMPSE model, below the model of Bizants and Gerhardt with measured distances to red clump stars (from Babusiaux & Gilmore 2006). Note the difference in the angle of the bar.

Galactic center and bar

The Galactic center is one of the most interesting areas of the Galaxy, with a central super-massive black hole surrounded by a dense assembly of stars with intriguing properties. Dynamical effects, in particular mass segregation, will lead to many interactions close to the central black hole so that wide binaries will become tighter or will be disrupted (for a review see Alexander 2005). This likely leads to an increase in the number of ultra-compact binaries as well as the possibility of extreme mass-ratio inspirals (see other sections in this volume). However, current observational techniques, fundamental observational limits due to the extreme faintness of these sources as well as severe limitations to theoretical calculations of these processes result in a virtual absence of any constraints on these populations and this will likely remain the situation until LISA will fly.

Another major question about the central region of the Galaxy is the size and orientation of a bar (see Hadamache *et al.* 2006, Gerhard 2002, also Figure 6-9). Because the ultra-compact binaries are expected to closely follow the mass distribution in the Galaxy, direct

measurement of distances and directions to hundreds of ultra-compact binaries in the Bulge will put a constraint on the bar, independent of the interpretation of star counts.

Galactic disc scale height

The level and shape of the double white dwarf background will provide information on the scale height of the ultra-compact binary population (Benacquista & Holley-Bockelmann 2006) in the disc of the Galaxy. For many of the resolved sources the LISA measurements will also provide an estimate of the orbital inclination which for the first time will give hints on the dynamics of the formation of binaries from interstellar clouds because in a statistical way the angular momentum vectors of the binaries can be compared to the overall angular momentum of the Galaxy.

LISA Science Objectives and Investigations relevant to this section

(see Appendix 1)

4. Survey compact stellar-mass binaries and study the structure of the Galaxy

4.1. Elucidate the formation and evolution of Galactic stellar-mass binaries; constrain the diffuse extragalactic foreground.

4.2. Determine the spatial distribution of stellar mass binaries in the Milky Way and environs.

4.3. Improve our understanding of white dwarfs, their masses, and their interactions in binaries and enable combined gravitational and electromagnetic observations.

7. New Physics and the Early Universe

Signatures of motions in the early Universe

LISA can detect a broad band stochastic background caused by a first-order phase transition with critical temperature from about 100 GeV to 1000 TeV. Proposed phase transitions include electroweak or supersymmetry breaking, dynamically active submillimeter extra dimensions, and some types of inflationary reheating. LISA may also detect a background, and possibly bursts, from otherwise unobservable cosmic superstrings.

Gravitational waves penetrate all of cosmic history, and LISA explores scales, epochs, and new physical effects not accessible in any other way (Figure 7-1). A detectable gravitational wave background in the LISA band is predicted by a number of new physical ideas for early cosmological evolution (Maggiore 2000, Hogan 2006a). Two important mechanisms for generating stochastic backgrounds are relativistic phase transitions and cosmic strings:

- Many types of new physics predict first-order phase transitions leading to bulk motions from cataclysmic bubble nucleation, cavitation, collisions, and turbulence (Figure 7-2). The cosmic expansion rate at a temperature of about 1TeV, corresponding to an apparent horizon size of about $c/H = ca/\dot{a} \approx 1$ mm at that time, is redshifted now to a frequency

$$f_0 = \dot{a}(t) \approx 10^{-4}\text{Hz}\,[H(t) \times 1\text{mm}/c]^{1/2} \approx 10^{-4}\text{Hz}\,(T/1\text{TeV})\,.$$

Thus, LISA's frequency band of about 0.1 to 100 millihertz today corresponds to the horizon at and beyond the Terascale frontier of fundamental physics. This allows LISA to probe bulk motions at times about 3×10^{-18} to 3×10^{-10} seconds after the Big Bang, a period not directly accessible with any other technique. Taking a typical broad spectrum into account, LISA has the sensitivity to detect cosmological backgrounds caused by new phase transitions from 0.1 to 1000 TeV, if more than a modest fraction $\approx 10^{-7}$ of the energy density is converted to gravitational radiation in LISA's band (Figure 7-3).

- Fundamental string theory, the subject of intense theoretical study as a unified framework for all particles and forces of nature, also predicts the possibility of new fundamental objects called cosmic superstrings, stretched to astronomical size by the cosmic expansion, that lose energy principally through gravitational radiation with a very broad and uniquely identifiable

Key science questions

- Is there a first-order phase transition at or beyond TeV energies?
- Are there extra dimensions at the submillimeter scale?
- Do stable superstrings exist, and can they be blown up to form cosmic strings?
- What was the quantum state of the Universe at or before the Big Bang?
- How did inflation end?
- Were there violent events in the early Universe that left no relic trace in conventional particles and fields?

spectrum. LISA will be our most sensitive probe for these objects by many orders of magnitude and so offers the possibility of detecting direct evidence of fundamental strings.

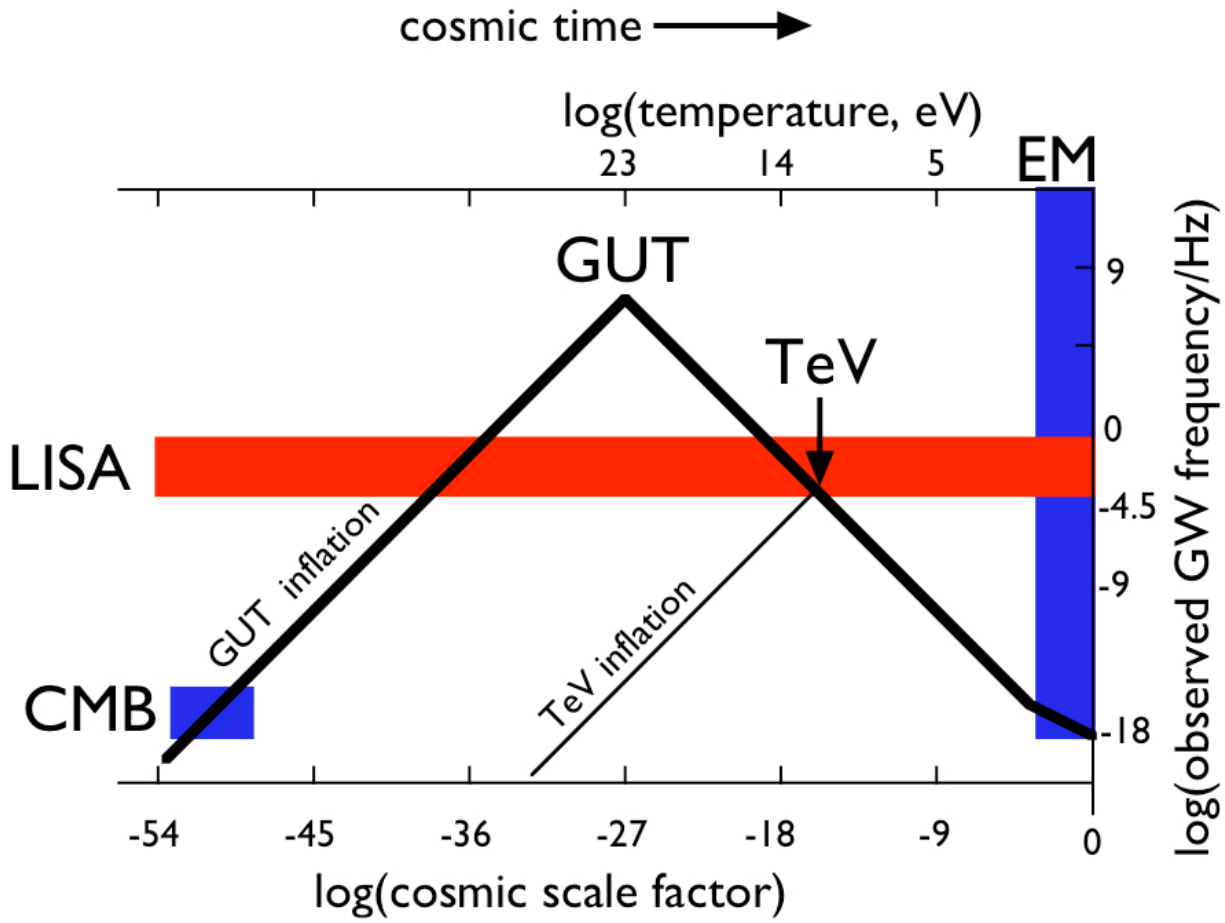


Figure 7-1: The history of mass motions the Universe, in units that show the unique penetration power and discovery potential of gravitational waves in cosmology. The observed (redshifted) frequency of wave-generating motions is shown as a function of cosmic scale factor *a*, with the present epoch at the right. The red bar shows the range of cosmic history accessed by LISA. Blue regions are accessible to electromagnetic (EM) observations; the bar to the right shows the Universe since recombination (some sources are also seen by LISA, where the two bars overlap), and the box at lower left shows the relic fluctuations imprinted in the CMB at inflation. The heavy solid line shows the redshifted Hubble rate or horizon scale, for a standard Grand Unified Theory (GUT) inflationary cosmology; the peak represents the epoch of reheating. Regions above this line correspond to motions smaller (*i.e.*, at higher frequencies) than the apparent horizon. LISA probes all processes within the the horizon up to about 1000 TeV, as well as a regime of cosmic inflation different from CMB observations.

First-Order Cosmological Phase Transitions: Bulk Motion from Bubble Nucleation, Cavitation, Collisions, Turbulence

Abundant evidence suggests that the physical vacuum was not always in its current state, but once had a significantly higher free energy. This idea is fundamental and general: it underlies symmetry breaking in theories such as the Standard Model and its supersymmetric extensions, and cosmological models including almost all versions of inflation. Common to all these schemes is the feature that a cold, nearly uniform free energy contained in the original ("false") vacuum is liberated in a phase transition to a final (or "true") vacuum, and eventually converted into thermal energy of radiation and hot plasma.

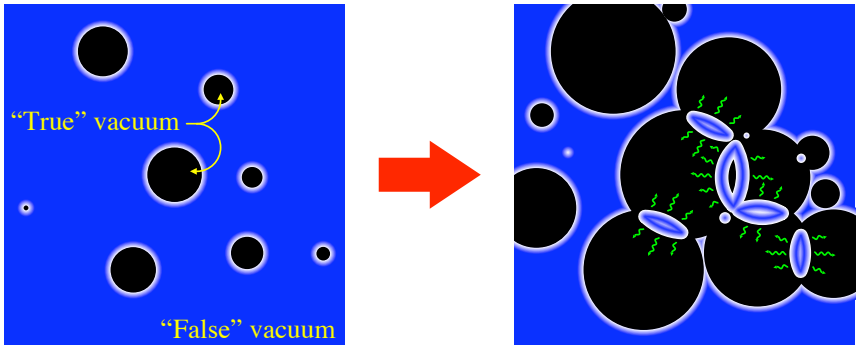


Figure 7-2: First-order phase transition between a "false" and a "true" vacuum. When expanding bubbles of the new phase collide, the latent heat of the old phase is converted into relativistic matter flows, which can produce gravitational radiation.

In many theories the conversion between vacuum states corresponds to a first-order phase transition. In an expanding universe this leads to a cataclysmic process. After supercooling below the critical temperature for the transition, a thermal or quantum jump across an energy barrier leads to the formation of bubbles of the new phase at widely separated nucleation sites. The bubbles rapidly expand and collide (Figure 7-2). The internal energy is thus converted to organized flows of mass-energy, whose bulk kinetic energy eventually dissipates via turbulence and finally thermalizes. The initial bubble collision and subsequent turbulent cascade lead to relativistic flows and acceleration of matter that radiate gravitational waves on a scale not far below the horizon scale (Witten 1984, Hogan 1986, Kosowsky *et al.* 2002, Dolgov *et al.* 2002).

Dynamics of Warped Submillimeter Extra Dimensions

A "theory of everything" based on quantum superstrings requires many, as yet invisible extra dimensions for mathematical consistency. The sizes of the dimensions, their shapes, and how they are stabilized are unknown. If they exist, gravity can penetrate into them so they must be small or highly "warped" — with sizes or radii of curvature below the submillimeter scale limits set by direct laboratory tests of the gravitational inverse-square law. (The scales probed by Standard Model particles and fields are much smaller than this, but fields other than gravity might be confined to a 3-dimensional subspace or "brane" living in a larger dimensional space.)

LISA, the LHC and the electroweak phase transition

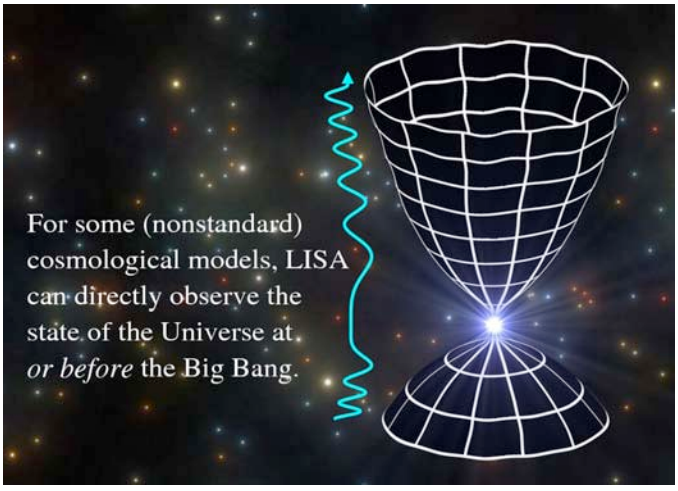
A moderately strong first order TeV-scale electroweak phase transition leads to the production of a detectable background of gravitational waves through bubble collisions and turbulence (Hogan 1986, Grojean & Servant 2006). In the Standard Model and its minimal supersymmetric extension, the phase transition associated with electroweak symmetry breaking by a Higgs field is now expected to be second order. However, other extensions of the Standard Model predict a first order phase transition, and electroweak baryogenesis generally requires it. In the next few years CERN's Large Hadron Collider will probe details of the Higgs sector and the nature of phase transitions in the Terascale region.

Since the Hubble length at the Terascale is about a millimeter, the current threshold where possible new effects of extra dimensions might appear happens to be about the same in the laboratory gravity, particle/field, and cosmological realms: that is, at present, laboratory gravity experiments, accelerator physics, and LISA cosmology converge on the same new regime in very different ways. It is even possible that new properties of gravity on this scale are related to cosmic dark energy (whose energy density is about $(0.1\text{mm})^{-4}$ in particle units).

The dynamics associated with the stabilization of extra dimensions at a certain size or warp radius might introduce a source of free internal energy released coherently on a "mesoscopic" (submillimeter to nanometer) scale, leading to a detectable background (Hogan 2000, Randall & Servant 2006). If the extra dimensions are much smaller than the Hubble length when the stabilization occurs, the behavior of the extra dimensions is nearly equivalent to scalar field behavior as viewed in conventional 3-dimensional space, with effects similar to the phase transitions just discussed (Figure 7-3). Brane condensation also introduces a new kind of mechanism for generating gravitational waves: motion and curvature of our Standard Model brane in the extra dimensions. LISA's high frequency limit at 1000 TeV corresponds to direct probes of extra dimensions as small as 10^{-6} mm.

Terascale Inflationary Reheating

Inflation represents an extraordinarily coherent behavior of an energetic scalar field that is nearly uniform across the observable Universe. After inflation, the internal potential energy of this field is converted into a thermal mix of relativistic particles, in a process known as "reheating". The reheating temperature might be as cool as 1 TeV, especially in some braneworld models where the Planck scale is itself not far above the Terascale.



There is no reason to assume a quiet,

orderly reheating process; the decay of the inflaton energy may be violently unstable. In many scenarios, the conversion begins with macroscopically coherent but inhomogeneous motions that eventually cascade to microscopic scales. Quantum coherent processes such as "preheating" transform the energy into coherent classical motions which, like the phase transitions discussed above, generate backgrounds of the order of 10^{-3} of the thermal plasma density (Felder & Kofman 2006, Easther & Lim 2006). As with those transitions, the characteristic frequency of the background matches the LISA band if the final reheating occurred at 0.1 to 1000 TeV.

Exotic Inflationary Quantum Vacuum Fluctuations

The amplification of quantum vacuum fluctuations during inflation leads to a background of primordial gravitational waves. An optimistic estimate of this background in the case of conventional inflation limits these to less than about 10^{-10} of the cosmic microwave background energy density, far below LISA's sensitivity; in many inflation models it is much less (Chongchitnan & Efstathiou 2006). However, some unconventional versions of inflation, particularly pre-Big-Bang or bouncing brane scenarios, predict possibly detectable backgrounds in the LISA band (see for example Buonanno 2003, Buonanno *et al.* 1997). Although some key parameters remain unknown, which limits the predictive power of these models, they are significantly constrained by gravitational wave backgrounds.

Backgrounds and Bursts from Cosmic Strings

String theory is the leading candidate for a theory unifying all of physics: both the quantum fields of the Standard Model, and the spacetime dynamics of general relativity. Many versions of string theory predict the cosmological formation of cosmic superstrings (Polchinski 2005): quasi-stable strings that form after inflation and are stretched to enormous length by the cosmic expansion. In equivalent field-theory language, cosmic strings arise from certain types of phase transitions, and stable relics of the high-energy phase persist in the form of one-dimensional strings that resemble flux tubes or trapped vortex lines.

Observational evidence for string theory

The discovery of an identifiable superstring background (and possibly, but less likely, superstring bursts) would be the first direct evidence that all particles and fields are indeed actually made of strings, and would help shape the rich mathematical insights of string theory into a model of the real world.

The primordial network of strings spawns isolated, oscillating loops that ultimately radiate almost all of their energy into gravitational waves. Their gravitational waves are mainly governed by a dimensionless parameter $G\mu/c^4$, where μ is the energy per unit length, or tension, in the string. Current limits on gravitational wave backgrounds from pulsar timing already

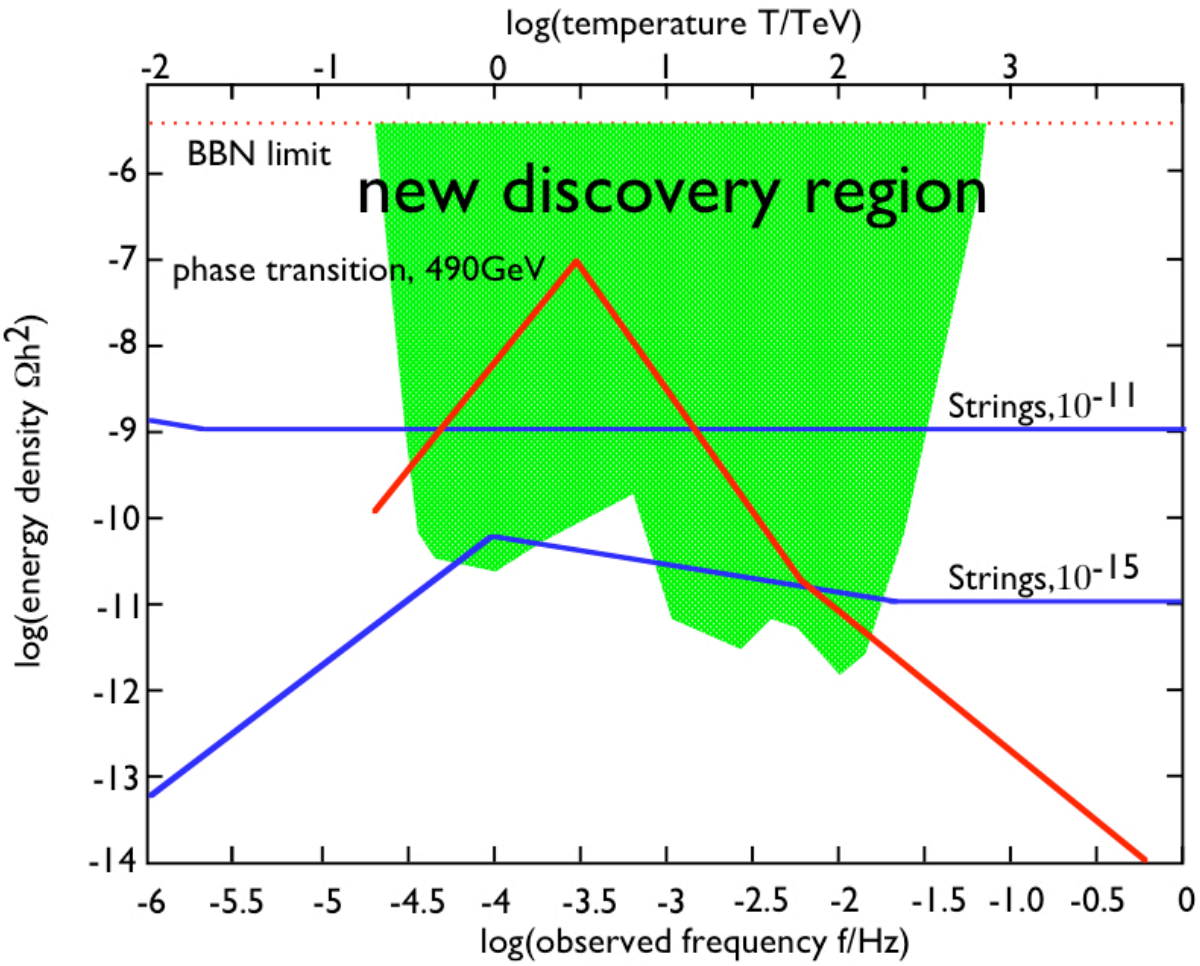


Figure 7-3: Gravitational wave background broad-band energy density is shown in units of the critical density for $h_0 = 1$. The new discovery region shown for LISA assumes a broad band limit using Sagnac calibration to distinguish instrument noise from isotropic gravitational wave backgrounds (Hogan & Bender 2001, updated with optimistic estimate for confusion from unresolved extragalactic binaries from Farmer & Phinney 2003). LISA's sensitivity extends about seven orders of magnitude below the energy density of thermal radiation, and six orders of magnitude below current limits from Big Bang Nucleosynthesis. The background spectrum from cosmic superstrings (Hogan 2006b) is shown for two values of string tension, $G\mu/c^4 = 10^{-11}$ and 10^{-15}. A model spectrum is shown from a first order phase transition of warped extra dimensions (Randall & Servant 2006), which is typical of strongly first order transition spectra; the spectrum moves to the right in the case of a higher transition critical temperature, down and right for a smaller nucleation scale or a smaller latent heat. Top axis is labeled by the temperature where the horizon size maps to the corresponding observed frequency.

suggest that if cosmic strings exist, they must be so light that they have no observable effects except from their gravitational radiation.

Figure 7-3 includes predicted stochastic background spectra (Hogan 2006b) from strings for two values of $G\mu/c^4$ spanning a range of scenarios motivated by brane world inflation. (These estimates are plotted for a "large loop" scenario where newly formed loops are about 0.1 of the horizon size.) The spectrum from cosmic strings is distinguishably different from that of phase

transitions or any other predicted source: it has nearly constant energy per logarithmic frequency interval over many decades at high frequencies, including the range where LISA is likely to observe it, and falls off at low frequencies since large string loops are rare and radiate slowly. LISA's sensitivity to $G\mu/c^4$ is many orders of magnitude deeper than the best possible future sensitivity from pulsar timing (which works at nanohertz frequencies, where light strings radiate very weakly and are buried in confusion noise).

If the strings are not too much lighter than $G\mu/c^4 \approx 10^{-10}$, occasional distinctive bursts might be seen from loops that happen to beam gravitational waves in our direction from "cusp catastrophes", where a momentary event produces a sharply-bent bit of string moving at nearly the speed of light (Damour & Vilenkin 2005, Siemens *et al.* 2006). These rare events, if they are intense enough to stand out above the background, are recognizable from their universal waveform, which derives just from the geometry of the cusps. Although individual burst events, if detected, give the clearest signature of a string source, the first detectable sign of a superstring loop population is likely their integrated stochastic background as shown in Figure 7-3 (Hogan 2006b).

LISA Science Objectives and Investigations relevant to this section

(see Appendix 1)

6. Probe new physics and cosmology with gravitational waves

6.2. Measure the or set bounds on the spectrum of cosmological GW backgrounds.

6.3. Search for burst events from cosmic string cusps.

8. LISA and the Key Questions of Astronomy and Physics

LISA addresses forcefully and directly many of the research priorities and big questions raised by recent astronomy and physics decadal and community reports such *Astronomy and Astrophysics in the New Millennium*, and *Connecting Quarks with the Cosmos*. It might not be an exaggeration to say that LISA will completely transform much of physics and astronomy. Since LISA will study the universe in an entirely new way, its discoveries are also likely to expand the scope of both sciences significantly and reshape the questions of the future.

LISA science addresses a very broad cross-section of some of the most important and exciting science priorities identified by major national and international reviews and studies. For instance, the roadmap for the NASA Beyond Einstein program identified five major science objectives and thirteen research focus areas (NASA SEU 2002). LISA, as a Great Observatory in the Beyond Einstein program, addresses all five science objectives, with particularly strong emphasis on four of the five. Furthermore, LISA addresses eight of the thirteen identified research focus areas, many in a way that is unique and impossible to accomplish using conventional space- and ground-based observatories. The Beyond Einstein objectives and focus areas for which LISA has

Beyond Einstein science objectives addressed by LISA

Science Objective 1: Find out what powered the Big Bang

Research Focus Area 1. Search for gravitational waves from inflation and phase transitions in the Big Bang.

Research Focus Area 2. Determine the size, shape, and energy content of the Universe.

Science Objective 2: Observe how black holes manipulate space, time, and matter

Research Focus Area 3. Perform a census of black holes throughout the Universe.

Research Focus Area 4. Determine how black holes are formed and how they evolve.

Research Focus Area 5. Test Einstein's theory of gravity and map space-time near the event horizons of black holes and throughout the Universe.

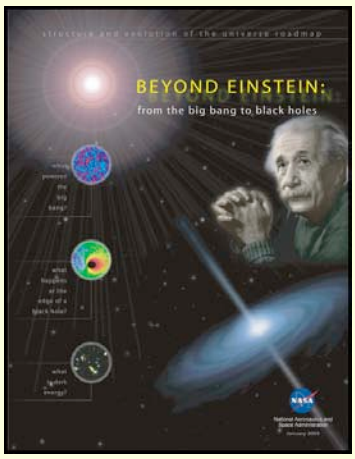
Research Focus Area 6. Observe stars and gas plunging into black holes.

Science Objective 3: Identify the mysterious dark energy pulling the Universe apart.

Research Focus Area 7. Determine the cosmic evolution of the dark energy pulling the Universe apart.

Science Objective 5: Understand the development of structure in the Universe.

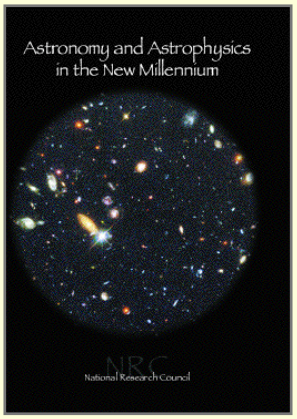
Research Focus Area 12: Discover how the interplay of baryons, dark matter, and gravity shapes galaxies and systems of galaxies.

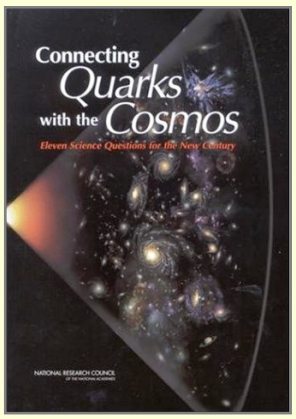


particularly powerful capabilities are given in the accompanying box.

In the US, LISA has been endorsed as a high priority mission in several influential reports. In particular, National Research Council (NRC) has issued several reports emphasizing the opportunities for exciting new science using gravitational waves. In 1999 the NRC report,

Major science questions of the AANM and Q2C reports addressed in unique ways by LISA

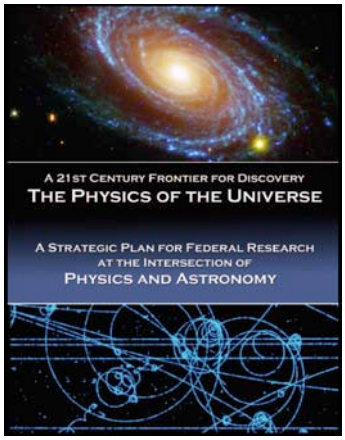




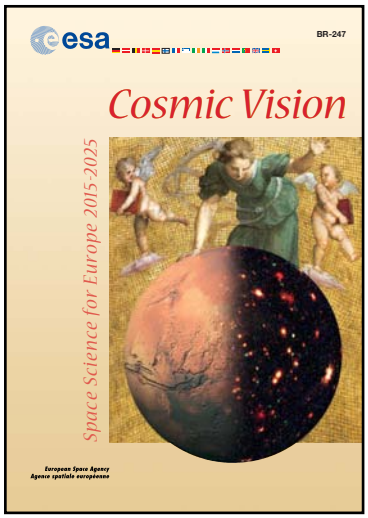
Report	Major Science Question	Relevant LISA Science
AANM & Q2C	How did the universe begin, evolve into today's structures, what is its destiny?	Gravitational waves are the only radiation that can be used to study the earliest beginnings of the Universe
AANM	How do galaxies form and evolve?	Massive black holes and mergers play a fundamental role in galaxy evolution
AANM	How do stars form and evolve?	Ultra-compact binaries represent the extreme endpoints of stellar binary evolution
Q2C	Did Einstein have the last word on gravity?	Black hole mergers allow precision tests of dynamic and static strong field gravity
Q2C	What is the nature of Dark Energy?	Precision distance estimates using black hole binaries offer a potentially powerful capability
Q2C	Are there additional spacetime dimensions?	Gravity is one of our best probes of the existence of extra dimensions

Gravitational Physics: Exploring the Structure of Space and Time (NRC 1999), described 12 science opportunities, several of which require LISA observations to be realized. A short time later, the 2000 NRC Decadal Survey, *Astronomy and Astrophysics in the new Millennium* (AANM: NRC 2001), stated that "LISA is unique among the recommended new initiatives in

that it is designed to detect the gravitational radiation predicted by Einstein's theory of general relativity. The direct measurement of gravitational radiation from astrophysical sources will open a new window onto the universe and enable investigations of the physics of strong gravitational fields." Likewise, the 2003 NRC report *Connecting Quarks with the Cosmos* (Q2C: NRC 2003) endorsed LISA and characterized LISA as having "great potential to address questions that lie at the boundary between physics and astronomy". The science in the AANM and Q2C reports was summarized in a set of 16 fundamental questions, 5 from the AANM report and 11 in the Q2C report. Many of these very important questions are addressed in a new and unique way by LISA science, as indicated in the accompanying box.



Following the AANM, Beyond Einstein, and Q2C reports, the National Science and Technology Council (NSTC) issued a major report, A 21st Century Frontier of Discovery: *The Physics of the Universe - A Strategic Plan for Federal Research at the Intersection of Physics and Astronomy*. This report endorsed the Beyond Einstein program as a key component of the strategy for federal research investment in physics and astronomy and stated that the execution of the LISA mission is "necessary to open up this powerful new window on the universe and create the new field of gravitational wave astronomy." Most recently, the NRC undertook a *Review of Progress in Astronomy and Astrophysics Toward the Decadal Vision*, issuing a letter report (NRC 2005). This report reviewed post-AANM astronomical discoveries and highlighted in particular progress in three broad areas: the universe and the nature of matter and energy, our place in the cosmos, and the formation and evolution of black holes. LISA is recognized as a facility-class mission, together with Con-X, that will "provide a broad and flexible science return across all of astrophysics, as have HST, CGRO, Chandra and Spitzer before them."



In Europe, LISA and LISA Pathfinder are integral components of ESA's Cosmic Vision Scientific Programme announced in 2005 (ESA 2005). The ESA space science program has traditionally been organized in long-term strategic plans. The Horizon 2000 plan was brought forward in 1985, followed by Horizon 2000 Plus in 1995. All of the missions in these programs have been realized subsequently, showing a remarkable stability in strategic planning. LISA first appeared in Horizon 2000 Plus, as a Cornerstone mission in Fundamental Physics. The other Cornerstone missions in this program are Herschel and Gaia in Astronomy and Bepi Colombo in Solar System Science, all to be launched soon, to follow Rosetta, which is flying already. LISA has subsequently been reconfirmed and it is now taken for granted in the Cosmic Vision Programme: "Despite some delays and descoping owing to budgetary constraints, the promises of Horizon 2000 will be broadly fulfilled when the astronomical missions Herschel and Planck set off into space in 2007. The second step in this

decadal series is Horizon 2000 Plus, including highly promising missions such as Gaia, BepiColombo, JWST, LISA and Solar Orbiter. Another challenge is LISA (Laser Interferometer Space Antenna), a joint ESA/NASA project which, by searching for gravitational waves, will open a new window on the Universe. On LISA Pathfinder (2011), ESA will test European and American contributions to the amazing technology required for this project."

1. ESA Cosmic Vision science objectives

2. ***Science Question 1: What are the conditions for planet formation and the emergence of life?***
   3. Research Topic 1.1. From gas and dust to stars and planets.
   4. Research Topic 1.2. From exo-planets to biomarkers.
   5. Research Topic 1.3. Life and habitability in the solar system.
6. ***Science Question 2: How does the solar system work?***
   7. Research Topic 2.1. From the sun to the edge of the solar system.
   8. Research Topic 2.2. The giant planets and their environments.
   9. Research Topic 2.3. Asteroids and other small bodies.
10. ***Science Question 3: What are the fundamental physical laws of the universe?***
    11. Research Topic 3.1. Explore the limits of contemporary physics.
    12. Research Topic 3.2. The gravitational wave Universe.
    13. Research Topic 3.3. Matter under extreme conditions.
14. ***Science Question 4: How did the Universe originate and what is it made of?***
    15. Research Topic 4.1. The early Universe.
    16. Research Topic 4.2. The Universe taking shape.
    17. Research Topic 4.3. The evolving violent Universe.
    18.

Cosmic Vision is constructed around four general Science Questions with three research topics each. While LISA as a Cornerstone mission in Fundamental Physics clearly belongs to Research Topic 3.2, it addresses all Research Topics in both Science Question 3 and Science Question 4.

Several of the ESA member states have roadmaps that make statements on space science and LISA in particular. In the German "Denkschrift Astronomie" (Deutsche Forschungsgemeinshaft 2003), LISA is a Golden Bullet and rated top priority. In the UK, LISA and LISA Pathfinder are contained in the Particle Physics and Astronomy Research Council (PPARC) Strategic Plan and have just been confirmed in the 2006 Delivery Plan (PPARC 2006). In the European Community's ApPEC Astroparticle Physics in Europe Roadmap (ApPEC 2006), active support for LISA is also foreseen as a top priority.

Appendix 1: LISA science objectives and scientific investigations

1. Understand the formation of massive black holes

1.1. Search for a population of seed black holes at early epochs.

1.2. Search for remnants of the first (Pop III) stars through observation of intermediate-mass black hole captures, both at formation and at later epochs.

2. Trace the growth and merger history of massive black holes and their host galaxies

2.1. Determine the relative importance of different black hole growth mechanisms as a function of redshift.

2.2. Determine the merger history of 10^4 to 3×10^5 $M_\odot$ black holes before the era of the earliest known quasars ($z \sim 6$).

2.3. Determine the merger history of 3×10^5 to 10^7 $M_\odot$ black holes at later epochs ($z < 6$).

2.4. Determine the masses and spins of low-z massive black holes (MBH) using observations of extreme mass ratio inspiral (EMRI) events ($z \lesssim 1$).

2.5. Enable the search for electromagnetic counterparts to MBH mergers in order to determine the host galaxy.

3. Explore stellar populations and dynamics in galactic nuclei

3.1. Characterize the immediate environment of MBH in $z \lesssim 1$ galactic nuclei from EMRI capture signals.

3.2. Study intermediate-mass black holes from their capture signals.

3.3. Improve our understanding of stars and gas in the vicinity of extragalactic black holes using coordinated gravitational and electromagnetic observations.

4. Survey compact stellar-mass binaries and study the structure of the Galaxy

4.1. Elucidate the formation and evolution of Galactic stellar-mass binaries; constrain the diffuse extragalactic foreground.

4.2. Determine the spatial distribution of stellar mass binaries in the Milky Way and environs.

4.3. Improve our understanding of white dwarfs, their masses, and their interactions in binaries and enable combined gravitational and electromagnetic observations.

5. **Confront General Relativity with observations**

5.1. Detect gravitational waves directly and measure their properties precisely.

5.2. Test whether the central massive objects in galactic nuclei are the black holes of general relativity.

5.3. Make precision tests of dynamical strong-field gravity in MBH binaries.

6. Probe new physics and cosmology with gravitational waves

6.1. Study cosmic expansion history, geometry and dark energy using precise gravitationally calibrated distances in cases and redshifts measured from electromagnetic counterparts or determined statistically.

6.2. Measure or set bounds on the spectrum of cosmological GW backgrounds.

6.3. Search for burst events from cosmic string cusps.

7. Search for unforeseen sources of gravitational waves

Appendix 2: Acronyms

AANM	Astronomy and Astrophysics in the New Millennium, a NRC report
AGN	active galactic nuclei
AM	amplitude modulation
AU	astronomical unit (150 million km)
BAO	baryon acoustic oscillation
BBN	Big Bang nucleosynthesis
BH	black hole
BHB	black hole binary
CMB	cosmic microwave background
COBE	Cosmic Background Explorer
CGRO	Compton Gamma Ray Observatory
EM	electromagnetic
EMRI	extreme mass ratio inspiral
ESA	European Space Agency
FM	Frequency Modulation
GAIA	Global Astrometric Interferometer for Astrophysics
GLAST	Gamma ray Large Area Space Telescope
GLIMPSE	Galactic Legacy Infrared Mid-Plane Survey Extraordinaire
GR	general relativity
GUT	grand unified theory
GW	gravitational wave
GWR	gravitational wave radiation
HST	Hubble Space Telescope
IMBH	intermediate mass black hole
JWST	James Webb Space Telescope
LHC	Large Hadron Collider
LIGO	Laser Interferometer Gravitational wave Observatory
LISA	Laser Interferometer Space Antenna
LMC	Large Magellanic Cloud
LPF	LISA Pathfinder
LSS	Large Scale Structure
MBH	massive black hole
NASA	National Aeronautics and Space Administration
NRC	National Research Council
NSF	National Science Foundation
NSTC	National Science and Technology Council
PN	post-Newtonian
Q2C	Connecting Quarks with the Cosmos, a NRC report

QNM	quasi-normal mode
RATS	Rapid Time Survey
RF	radio frequency
RMS	root mean square
RXTE	Rossi X-ray Timing Explorer
SDSS	Sloan Digital Sky Survey
SEGUE	Sloan Extension for Galactic Understanding and Exploration
SEU	Structure and Evolution of the Universe
SMBH	supermassive black hole
SMC	Small Magellanic Cloud
SNAP	Supernova Acceleration Probe
SNR	signal-to-noise ratio
TDI	time delay interferometry
UV	ultra-violet
WD	white dwarf
WMAP	Wilkinson Microwave Anisotropy Probe

Appendix 3: References

Executive Summary

NASA SEU 2002 *Beyond Einstein: from the big bang to black holes*, Structure and Evolution of the Universe Roadmap Team, NP-2002-10-510-GSFC

National Research Council (NRC) 1999 *Gravitational Physics: Exploring the Structure of Space and Time*, National Academy Press

National Research Council (NRC) 2001 *Astronomy and Astrophysics in the New Millennium,* National Academy Press

National Research Council (NRC) 2003 *Connecting Quarks with the Cosmos: Eleven Science Questions for the New Century,* National Academy Press

National Research Council (NRC) 2005 *Review of Progress in Astronomy and Astrophysics Toward the Decadal Vision: Letter Report*, National Research Council (http://www.nap.edu/catalog/11230.html)

National Science and Technology Council (NSTC) *A 21st Century Frontier of Discovery: The Physics of the Universe - A Strategic Plan for Federal Research at the Intersection of Physics and Astronomy*

1. Gravitational Waves: An Overview

Abramovici A, *et al.* 1992 "LIGO: The Laser Interferometer Gravitational-Wave Observatory" *Science* **256**, 325–333

Flanagan E, Hughes S 2005 "The basics of gravitational wave theory" *New J Phys* **7**, 204

Hogan C 2006 "The sounds of spacetime" *American Scientist* **94**, 534–541

Hughes S 2006 "A brief survey of LISA sources and science" gr-qc/0609028

Hughes S 2003 "Listening to the universe with gravitational-wave astronomy" *Annals of Physics* **303**, 142–178

Kramer M, *et al.* 2006 "Tests of general relativity from timing of the double pulsar" *Science* **314,** 97–102

Rodriguez C, *et al.* 2006 "A compact supermassive binary black hole system*" ApJ* **646**, 49–60

Thorne K 1987 "Gravitational Radiation" in Three Hundred Years of Gravitation eds. Hawking S, Israel W (Cambridge: Cambridge University Press) p330–458

Schutz B 1999 "Gravitational wave astronomy" *Class Quant Grav* **16**, A131–A156

Tinto M, Dhurandar S 2005 "Time-Delay Interferometry" *Living Reviews in Relativity* **8**, 4 (http://www.livingreviews.org/lrr-2005-4).

Volonteri M 2006 “Supermassive black hole mergers and cosmological structure formation” in Proceedings of the Sixth LISA Symposium, Eds Merkowitz S, Livas J, *AIP Conf Series* **873**, 61–69, astro-ph/0609741

2. LISA Mission Overview

Arnaud K, *et al.* 2006 “The mock LISA data challenges: An overview” in Proceedings of the Sixth LISA Symposium, Eds Merkowitz S, Livas J, *AIP Conf Series* **873**, 619–624, gr-qc/0609105/

Cramer B, *et al.* 2003, “Laser Interferometer Space Antenna: Technology Readiness & Implementation Plan” TRIP Report, NASA, available at http://www.srl.caltech.edu/lisa/documents.html.

Hammesfahr A 2001 “LISA mission study overview” *Class Quant Grav* **18**, 4045–4051.

Herz M 2005 “Active laser frequency stabilization and resolution enhancement of interferometers for the measurement of gravitational waves in space” *Optical Engineering* **44**, 090505.

Sheard B, *et al.* 2003 “Laser frequency stabilization by locking to a LISA arm” *Physics Letters A* **320**, 9-21.

Tinto M, Dhurandar S 2005 “Time-Delay Interferometry” *Living Reviews in Relativity* **8**, 4 (http://www.livingreviews.org/lrr-2005-4).

Tinto M, *et al.* 2003 “Implementation of time-delay interferometry for LISA”, *PRD* **67**, 122003

3. Black Hole Astrophysics: Supermassive Black Holes in Galactic Nuclei

Abel T, *et al.* 2002 “The Formation of the First Star in the Universe” *Science* **295**, 93-98

Alexander T, Hopman C 2003 “Orbital inspiral into a massive black hole in a galactic center” *ApJL* **590**, L29

Aller M, Richstone D 2002 “The cosmic density of massive black holes from galaxy velocity dispersions” *AJ* **124**, 3035–3041

Barack L, Cutler C 2004 “Confusion noise from LISA capture sources” *Phys Rev D* **69**, 082005

Barth A, *et al.* 2005 “Dwarf Seyfert 1 nuclei and the low-mass end of the M_{BH}–σ relation” *ApJ* **619**, L151–L154

Begelman M, *et al.* 2006 “Formation of supermassive black holes by direct collapse in pregalactic haloes” *MNRAS* **370**, 289–298

Begelman M 2003 “AGN Feedback Mechanisms” in Carnegie Observatories Astrophysics Series, Vol 1: Coevolution of Black Holes and Galaxies, Cambridge Univ Press astro-ph/0303040

Bell E, *et al*. 2006 “The merger rate of massive galaxies” *ApJ* **652**, 270–276

Bond J, *et al.* 1984 “The evolution and fate of very massive objects” *ApJ* **280**, 825–847

Brandt W, Hasinger G 2005 "Deep extragalactic X-ray surveys" *ARAA* **43**, 827–859

Bromm V, Larson L 2004 "The first stars" *Annual Reviews Astronomy Astrophysics* **42**, 79–118

Bromm V, Loeb A 2003 "Formation of the first supermassive black holes" *ApJ* **596**, 34–46

Bromm V, *et al.* 2002 "The formation of the first stars. I. The primordial star-forming cloud" *ApJ* **564**, 23–51

Campanelli M, *et al.* 2006 "Spinning-black-hole binaries: The orbital hang up" *Phys Rev D* **74**, 041501

Cohen S, *et al.* 2006 "Clues to active galactic nuclei growth from optically variable objects in the Hubble Ultra Deep Field" *ApJ* **639**, 731–739

DiMatteo T, *et al.* 2005 "Energy input from quasars regulates the growth and activity of black holes and their host galaxies" *Nature* **433**, 604–607, Original of Figure 7 available at http://web.phys.cmu.edu/~simtiziana/BHGrow/

Eisenhauer F, *et al.* 2005 "SINFONI in the Galactic Center: Young stars and infrared flares in the central light month" *ApJ* **628**, 246–259

Ferrarese L, Ford H 2005 "Supermassive black holes in galactic nuclei: Past, present and future research" *Space Science Reviews* **116**, 523–624

Ferrarese L, Merritt D 2000 "A Fundamental Relation between Supermassive Black Holes and Their Host Galaxies" *ApJ* **539,** L9–L12

Freitag M 2004 "Captures of stars by a massive black hole: Investigations in numerical stellar dynamics" in The Astrophysics of Gravitational Wave Sources, Eds Centrella J, *AIP Conf Series*, **686**, 109–114 (also astro-ph/0306064).

Freitag M 2003 "Gravitational waves from stars orbiting the Sagittarius A* black hole" *ApJ* **583**, L21–L24

Freitag M, *et al.* 2006 "Stellar remnants in galactic nuclei: Mass segregation" *ApJ* **649**, 91–117

Gair J, *et al.* 2004 "Event rate estimates for LISA extreme mass ratio capture sources" *Class Quant Grav* **21**, S1595–S1606

Gammie C, *et al.* 2004 "Black hole spin evolution" *ApJ* **602**, 312–319

Gebhardt K, *et al.* 2000 "A relationship between nuclear black hole mass and galaxy velocity dispersion" *ApJ* **539**, L13–L16

Ghez A, *et al.* 2005 "Stellar orbits around the Galactic Center black hole" *ApJ* **620**, 744–757

Haehnelt M, Kauffmann G 2002 "Multiple supermassive black holes in galactic bulges" *MNRAS* **336**, L61–L64

Hawley J, Krolik J 2006 "Magnetically driven jets in the Kerr Metric" *ApJ* **641**, 103–116

Hopkins P, *et al.* 2006 "A unified, merger-driven model for the origin of starbursts, quasars, the cosmic X-ray background, supermassive black holes and galaxy spheroids" *ApJ Suppl* **163**, 1–49

Hopman C, Alexander T 2006 "The effect of mass segregation on gravitational wave sources near massive black holes" *ApJ* **645**, L133–L136

Hopman C, Alexander T 2005 "The orbital statistics of stellar inspiral and relaxation near a massive black hole: Characterizing gravitational wave sources" *ApJ* **629**, 362--372

Hughes S, Blandford R 2003 "Black hole mass and spin coevolution by mergers" *ApJ* **585**, L101–104

Koushiappas S, *et al.* 2004 "Massive black hole seeds from low angular momentum material" *MNRAS* **354**, 292–304

Krolik J 1999 Active Galactic Nuclei, Princeton Univ Press

Krolik J, *et al.* 2005 "Magnetically driven accretion flows in the Kerr metric. IV. Dynamical properties of the inner disk" *ApJ* **622**, 1008–1023

Kuhlen M, Madau P 2005 "The first miniquasar" *MNRAS* **363***,* 1069-1082

Lang R, Hughes S 2006 "Measuring coalescing massive binary black holes with gravitational waves: The impact of spin-induced precession" *Phys Rev D* **74**, 122001

Lauer T, *et al.* 2006 "The masses of nuclear black holes in luminous elliptical galaxies and implications for the space density of the most massive black holes" astro-ph/0606739

Levin 2006 "Starbursts near supermassive black holes: Young stars in the Galactic Center and gravitational waves in the LISA band" *MNRAS* **374,** 515–524

Lynden-Bell D 1969 "Galactic nuclei as collapsed old quasars" *Nature* **223**, 690–694

Madau P, Rees M 2001 "Massive black holes as population III remnants" *ApJ* **551**, L27–L30

Madau P, *et al.* 2004 "Early reionization by miniquasars" *ApJ* **604**, 484–494

Maoz E 1998 "Dynamical constraints on alternatives to supermassive black holes in galactic nuclei" *ApJ* **494**, L181–L184

Marconi A, *et al.* 2004 "Local supermassive black holes, relics of active galactic nuclei and the X-ray background" *MNRAS* **351**, 169–185

Merloni A 2004 "The anti-hierarchical growth of supermassive black holes" *MNRAS* **353**, 1035–1047

Merritt D, Milosavljevic M. 2005 "Massive black hole binary evolution" *Living Reviews in Relativity* 8, http://relativity.livingreviews.org/Articles/lrr-2005-8

Miller MC, *et al.* 2005, "Binary encounters with supermassive black holes: Zero-eccentricity LISA events" *ApJL* **631**, L117-L120

Milosavljevic M, *et al.* 2006 "Contribution of stellar tidal disruptions to the X-ray luminosity function of active galaxies" astro-ph/0602289

Miralda-Escudé J, Gould A 2000 "A cluster of black holes at the galactic center" *ApJ* **545**, 847–853

Mo H, White S 2002 "The abundance and clustering of dark haloes in the standard ΛCDM cosmogony" *MNRAS* **336,** 112–118

Murray N, *et al*. 2005 "On the maximum luminosity of galaxies and their central black holes: Feedback from momentum-driven winds" *ApJ* **618**, 569–585

Pretorius F 2005 “Evolution of binary black hole spacetimes” *Phys Rev Lett* **95**, 121101

Rauch K 1999 “Collisional stellar dynamics around massive black holes in active galactic nuclei" *ApJ* **514**, 725–745

Rees M 1978 “Quasars” *The Observatory* **98**, 210–223

Salpeter E 1964 “Accretion of interstellar matter by massive objects” *ApJ* **140**, 796–800

Sesana A, *et al.* 2004 “Low frequency gravitational radiation from coalescing massive black hole binaries in hierarchical cosmologies” *ApJ* **611**, 623–632

Shibata M, Shapiro S 2002 “Collapse of a rotating supermassive star to a supermassive black hole: Analytic determination of the black hole mass and spin” *ApJ* **577**, 904–908

Soltan A 1982 “Masses of quasars” *MNRAS* **200**, 115–122

Thorne K 1995, in Black Holes and Time Warps: Einstein’s Outrageous Legacy, 273

Thorne K 1974 “Disk-accretion onto a black hole. II. Evolution of the hole” *ApJ* **191**, 507–520

Toomre A 1977 “Mergers and some consequences” in Evolution of Galaxies and Stellar Populations, Eds. Tinsley B, Larson R; Yale Univ Press, 401–426

Tremaine S, *et al.* 2002 “The slope of the black hole mass versus velocity dispersion correlation” *ApJ* **574**, 740–753

Tundo E, *et al.* 2006 “Scatter in the $M_\bullet - \sigma$ and $M_\bullet - L$ correlations and consequences for the inferred black hole mass function” astro-ph/0609297

Volonteri M 2006 “Supermassive black hole mergers and cosmological structure formation” in Proceedings of the Sixth LISA Symposium, Eds Merkowitz S, Livas J, *AIP Conf Series* **873**, 61–69, astro-ph/0609741

Volonteri M, *et al.* 2005 “The distribution and cosmic evolution of massive black hole spins” *ApJ* **620**, 69–77

Woosley S, *et al.* 2002 “The evolution and explosion of massive stars” *Rev Mod Phys* **74**, 1015–1071

Yu Q 2002 “Evolution of massive binary black holes” *MNRAS* **331**, 935–958

Zel'dovich Ya, Novikov I 1964 *Dokl. Akad. Nauk. SSSR* **158**, 811

Zier C 2006 “Merging of a massive black hole binary II” astro-ph/0610457

4. Black Hole Physics: Confronting General Relativity with Precision Measurements of Strong Gravity

Baker J, *et al.* 2007 “Comparisons of binary black hole merger waveforms”, gr-qc/0701016

Baker J, *et al.* 2006 “Binary black hole late inspiral: Simulations for gravitational wave observations” gr-qc/0612117

Barack L, Lousto C 2005 “Perturbations of Schwarzschild black holes in the Lorenz gauge: Formulation and numerical implementation” *Phys Rev D* **72**, 104026

Barack L, Cutler C 2006 "Using LISA EMRI sources to test off-Kerr deviations in the geometry of massive black holes" gr-qc/0612029

Barack L, Cutler C 2004 "LISA capture sources: Approximate waveforms, signal-to-noise ratios, and parameter estimation accuracy" *Phys Rev D* **69**, 082005

Barack L, Ori A 2003 "Gravitational self-force on a particle orbiting a Kerr black hole" *Phys Rev Lett* **90** 111101

Berti E, *et al.* 2006 "On gravitational-wave spectroscopy of massive black holes with the space interferometer LISA" *Phys Rev D* **73**, 064030

Blanchet L, *et al.* 2006 "Higher-order spin effects in the dynamics of compact binaries II. Radiation field" *Phys Rev D* **74**, 104034

Centrella J 2006 "The final merger of comparable mass binary black holes" in Proceedings of the Sixth LISA Symposium, Eds Merkowitz S, Livas J, *AIP Conf Series* **873,** 70–81, gr-qc/0609172

Cooray A, Seto N 2004 "Graviton Mass from Close White Dwarf Binaries Detectable with LISA" *Phys Rev D* **69**, 103502

Drasco S, Hughes S 2006 "Gravitational wave snapshots of generic extreme mass ratio inspirals" *Phys Rev D* **73**, 024027

Drasco S, Hughes S 2004 "Rotating black hole orbit functionals in the frequency domain" *Phys Rev D* **69**, 044015

Dreyer O, *et al.* 2004 "Black hole spectroscopy: Testing general relativity through gravitational wave observations" *Class Quant Grav* **21,** 787–804

Esposito-Farese G (2004) "Tests of scalar tensor gravity" in Phi in the Sky: The Quest for Cosmological Scalar Fields, Eds. Martins C, *et al*., *AIP Conf Series* **736,** 35–52

Flanagan E, Hughes S 1998 "Measuring gravitational waves from binary black hole coalescences: I. Signal to noise for inspiral, merger, and ringdown" *Phys Rev D* **57**, 4535–4565

Gair R, *et al.* 2004 "Event rate estimates for LISA extreme mass ratio capture sources" *Class Quant Grav* **21**, S1595–S1606

Hopman C 2006 "Astrophysics of extreme mass ratio inspiral sources" astro-ph/0608460

Kesden M, *et al.* 2005 " Gravitational-wave signature of an inspiral into a supermassive horizonless object" *Phys Rev D* **71**, 044015

Kramer M, *et al.* 2006 "Tests of general relativity from timing the double pulsar" *Science Express*, Sept 14, 2006, astro-ph/0609417

Lang R, Hughes S 2006 "Measuring coalescing massive binary black holes with gravitational waves: The impact of spin-induced precession" *Phys Rev D* **74**, 122001

Lorimer D 2005 "Binary and millisecond pulsars", *Living Rev Relativity* **8**, 7 (www.livingreviews.org/lrr-2005-7)

Mino Y 2005 "Self-force in the radiation reaction formula – adiabatic approximation of a metric perturbation and an orbit" *Prog Theor Phys* **113**, 733–761

Mino Y, *et al.* 1997 "Gravitational radiation reaction to a particle motion" *Phys Rev D* **55**,

3457–3476

Nissanke S, Blanchet L 2005 "Gravitational radiation reaction in the equations of motion of compact binaries to 3.5 post-Newtonian order" *Class Quant Grav* **22** 1007–1032

Quinn T, Wald R 1997 "Axiomatic approach to electromagnetic and gravitational radiation reaction of particles in curved spacetime" *Phys Rev D* **56** 3381–3394

Ryan F 1995 "Gravitational waves from the inspiral of a compact object into a massive, axisymmetric body with arbitrary multipole moments" *Phys Rev D* **52**, 5707–5718

Weisberg J, Taylor J 2004 "Relativistic binary pulsar B1913+16: Thirty years of observations and analysis" in Proc Aspen Conference, ASP Conf Series Binary Radio Pulsars Eds Rasio F, Stairs I astro-ph/0407149

Will C 2006 "The confrontation between general relativity and experiment" *Living Rev Relativity* **9**, 3 (www.livingreviews.org/lrr-2006-3)

5. Precision Cosmometry and Cosmology

Albrecht A, *et al*. 2006 "Report of the dark energy task force" astro-ph/0609591

Astier P, *et al*. 2006 "The Supernova Legacy Survey: measurement of Ω_M, Ω_Λ and w from the first year data set" *A&A* **447**, 31–48

Bonanos A, *et al.* 2006 "The first direct distance determination to a detached eclipsing binary in M33" *ApJ* **652** 313–322

Dalal N, *et al.* 2006 "Short GRB and binary black hole standard sirens as a probe of dark energy" *Phys Rev D* **74**, 063006

Dotti M, *et al.* 2006 "On the search of electromagnetic cosmological counterparts to coalescences of massive black hole binaries" *MNRAS* **372**, 869–875

Eisenstein D, White M 2004 "Theoretical uncertainty in baryon oscillations" *Phys Rev D* **70**, 103523

Eisenstein D, *et al*. 2005 "Detection of the baryon acoustic peak in the large-scale correlation function of the SDSS luminous red galaxies" *ApJ* **633**, 560–574

Freedman W, *et al.* 2001 "Final results from the Hubble Space Telescope key project to measure the Hubble constant" *ApJ* **553**, 47–72

Holz D, Hughes S 2005 "Using gravitational-wave standard sirens" *ApJ* **629**, 15–22

Hu W 2005 "Dark energy probes in light of the CMB" in Observing Dark Energy, Eds Wolff S, Lauer T, ASP Conf Ser **399**, 215

Knox L 2006 "On precision measurement of the mean curvature" *Phys Rev D* **73,** 023503

Lang R, Hughes S 2006 "Measuring coalescing massive binary black holes with gravitational waves: The impact of spin-induced precession" *Phys Rev D* **74**, 122001

Macri L, *et al.* 2006 "A New Cepheid distance to the maser-host galaxy NGC 4258 and its implications for the Hubble constant" *ApJ* **652** 1133–1149

Milosavljevic M, Phinney E 2005 “The afterglow of massive black hole coalescence” *ApJ* **622**, L93–L96

National Research Council (NRC) 2003 *Connecting Quarks with the Cosmos: Eleven Science Questions for the New Century,* National Academy Press

Olling R 2007 “Accurate extra-galactic distances and dark energy: Anchoring the distance scale with rotational parallaxes” astro-ph/0607607

Sandage A, *et al.* 2006 “The Hubble constant: A summary of the host program for the luminosity calibration of Type Ia supernovae by means of Cepheids” *ApJ* **653,** 843–860

Schutz B 1986 “Determining the Hubble constant from gravitational wave observations” *Nature* **323**, 310–311

Sesana A, *et al.* 2004 “Low-frequency gravitational radiation from coalescing massive black hole binaries in hierarchical cosmologies” *ApJ* **611**, 623–632

Spergel D, *et al.* 2006 “Wilkinson microwave anisotropy probe (WMAP) three year results: Implications for cosmology” astro-ph/0603449

Volonteri M 2006 “Supermassive black hole mergers and cosmological structure formation” in Proceedings of the Sixth LISA Symposium, Eds Merkowitz S, Livas J, *AIP Conf Series* **873**, 61–69, astro-ph/0609741

6. Ultra-Compact Binaries

Alexander T 2005 “Stellar processes near the massive black hole in the Galactic center” *Phys Rep* **419,** 65–142

Babusiaux C, Gilmore G 2005 “The structure of the Galactic bar” *MNRAS* **358**, 1309–1319

Belczynski K, *et al.* 2002 “A comprehensive study of binary compact objects as gravitational wave sources: Evolutionary channels, rates, and physical properties” *ApJ* **572**, 407–431

Benacquista M, Holley-Bockelmann K 2006 “Consequences of disk scale height on LISA confusion noise from close white dwarf binaries” *ApJ* **645**, 589–596

Bhattacharya D, van den Heuvel E 1991 “Formation and evolution of binary and millisecond radio pulsars” *Physics Rep* **203**, 1–124

Bildsten L 1998 “Gravitational radiation and rotation of accreting neutron stars” *ApJ* **501**, L89–L93

Chakrabarty D, *et al.* 2003 “Nuclear-powered millisecond pulsars and the maximum spin frequency of neutron stars” *Nature* **424**, 42–44

d'Souza M, *et al.* 2006 "Numerical simulations of the onset and stability of dynamical mass transfer in binaries" *ApJ* **643**, 381-401

Edlund J, *et al.* 2005 “White-dwarf white-dwarf galactic background in the LISA data” *Phys Rev D*, **71**, 122003

Farmer A, Phinney E 2003 “The gravitational wave background from cosmological compact binaries” *MNRAS* **346**, 1197–1214

Galloway D, *et al.* 2002 “Discovery of a high-latitude accreting millisecond pulsar in an ultracompact binary” *ApJ* **576**, L137–L140

Gerhard O 2002 “The galactic bar” In ASP Conf. Ser. 273: The Dynamics, Structure & History of Galaxies: A Workshop in Honour of Professor Ken Freeman Eds Da Costa G, Jerjen H, 73–84

Hamadache C, *et al.* 2006 “Galactic bulge microlensing optical depth from EROS-2” *A&A* **454**, 185

Levan A, *et al.* 2006 “Short gamma-ray bursts in old populations: magnetars from white dwarf-white dwarf mergers” *MNRAS* **368**, L1–L5

Marsh T, *et al.* 2004 “Mass transfer between double white dwarfs” *MNRAS* **350**, 113–128

Nelemans G 2006 “Ultracompact binary stars” *Physics Today* **59**, 26

Nelemans G, *et al.* 2004 “Short-period AM CVn systems as optical, X-ray and gravitational-wave sources” *MNRAS* **349**, 181–192

Nelemans G, *et al.* 2001 “Population synthesis for double white dwarfs. II Semi-detached systems; AM Cvn binaries” *A&A* **368**, 939–949

Paczynski B 1976 “Common envelope binaries” In Structure and Evolution of Close Binary Systems, Eds Eggleton P, Mitton S, Whelan J; Kluwer, 75–80

Racine E, *et al.* 2006 “Non-dissipative tidal synchronization in accreting binary white dwarf systems” astro-ph/0610692.

Roelofs G, *et al.* 2006 “HST/FGS parallaxes of AM CVn stars and astrophysical consequences” *ApJ*, submitted

Stroeer A, Vecchio A 2006 “The LISA verification binaries” Class Quant Grav 23 S809–S817

Taam R, Sandquist E 2000 "Common envelope evolution of massive binary stars" *Annual Rev Astron Astrophys* **38**, 113–141

Verbunt F, Lewin W 2006 “Globular cluster X-ray sources” In Compact Stellar X-ray Sources Eds Lewin W, van der Klis M, *Cambridge Astrophysics Series* **39**; Cambridge University Press, 341–379

Webbink R 1984 “Double white dwarfs as progenitors of R Coronae Borealis stars and Type I supernovae” *ApJ* **277**, 355–360

Yungelson L, *et al*. 2002 “On the formation of neon-enriched donor stars in ultracompact X-ray binaries” *A&A* **388**, 546–551

7. New Physics and the Early Universe

Brustein R, *et al.* 1995 “Relic gravitational waves from string cosmology” *Phys Lett B* **361**, 45–51

Buonanno A 2003 “Gravitational waves from the early Universe” gr-qc/0303085

Buonanno A, *et al.* 1997 "Spectrum of relic gravitational waves in string cosmology" *Phys Rev D* **55**, 3330

Chongchitnan S, Efstathiou G 2006 "Prospects for direct detection of primordial gravitational waves" *Phys Rev D* **73**, 083511

Damour T, Vilenkin A 2005 "Gravitational radiation from cosmic (super)strings: Bursts, stochastic background, and observational wind" *Phys Rev D* **71**, 063510

Dolgov A, *et al*. 2002 "Relic backgrounds of gravitational waves from cosmic turbulence" *Phys Rev D* **66**, 103505

Easther R, Lim E 2006 "Stochastic gravitational wave production after inflation" *J Cosmol Astropart Phys* **0604**, 010

Farmer A, Phinney E 2003 "The gravitational wave background from cosmological compact binaries" *MNRAS* **346**, 1197–1214

Felder G, Kofman L 2006 "Nonlinear inflation fragmentation after preheating" hep-ph/0606256

Grojean C, Servant G 2006 "Gravitational waves from phase transitions at the electroweak scale and beyond" hep-ph/0607107

Hogan C 2006a "Gravitational wave sources from new physics" in Proceedings of the Sixth LISA Symposium, Eds Merkowitz S, Livas J, *AIP Conf Series* **873**, 30–40, astro-ph/0608567

Hogan C 2006b "Gravitational waves from light cosmic strings: Backgrounds and bursts with large loops" *Phys Rev D* **74**, 043526

Hogan C 2000 "Gravitational waves from mesoscopic dynamics of the extra dimensions" *Phys Rev Lett* **85,** 2044–2047

Hogan C 1986 "Gravitational radiation from cosmological phase transitions" *MNRAS* **218**, 629–636

Hogan C, Bender P 2001 "Estimating stochastic gravitational wave backgrounds with the Sagnac calibration" *Phys Rev D* **64**, 062002

Khlebnikov S, Tkachev I 1997 "Relic gravitational waves produced after preheating**"** *Phys Rev D* **56**, 653–660

Kosowsky A, *et al.* 2002 "Gravitational radiation from cosmological turbulence" *Phys Rev D* **66**, 024030

Maggiore M 2000 "Gravitational wave experiments and early universe cosmology" *Physics Rep* **331**, 283–367

Polchinski J 2005 "Cosmic superstrings revisited" *Int J Mod Phys A* **20**, 3413

Randall L, Servant G 2006 "Gravitational waves from warped spacetime" hep-ph/0607158

Siemens X, *et al.* 2006 "Gravitational wave bursts from cosmic (super)strings: Quantitative analysis and constraints" *Phys Rev D* **73**, 105001

Witten E 1984 "Cosmic separation of phases" *Phys Rev D* **30**, 272–285

8. LISA and the Key Questions of Astronomy and Physics

ApPEC 2006 "Astroparticle Physics in Europe Roadmap", (https://ptweb.desy.de/appec/roadmap.html)

Deutsche Forschungsgemeinshaft 2003, "Denkschrift Astronomie (Status and Perspectiven der Astronomie in Deuschland 2003-2016)", Wiley-VCH (ISBN 3-527-27220-8)

ESA 2005 "Cosmic Vision: Space Science for Europe 2015-2025", *ESA Publications Division* **BR-247 (**http://sci.esa.int/science-e/www/object/index.cfm?fobjectid=38542#) (see also http://www.esa.int/esaSC/SEMA7J2IU7E_index_0.html).

NASA SEU 2002 *Beyond Einstein: from the Big Bang to black holes*, Structure and Evolution of the Universe Roadmap Team, NP-2002-10-510-GSFC

National Research Council (NRC) 1999 *Gravitational Physics: Exploring the Structure of Space and Time*, National Academy Press

National Research Council (NRC) 2001 *Astronomy and Astrophysics in the New Millennium,* National Academy Press (AANM)

National Research Council (NRC) 2003 *Connecting Quarks with the Cosmos: Eleven Science Questions for the New Century,* National Academy Press (Q2C)

National Research Council (NRC) 2005 *Review of Progress in Astronomy and Astrophysics Toward the Decadal Vision: Letter Report*, National Research Council (http://www.nap.edu/catalog/11230.html)

National Science and Technology Council (NSTC) *A 21st Century Frontier of Discovery: The Physics of the Universe - A Strategic Plan for Federal Research at the Intersection of Physics and Astronomy*

PPARC 2006 "Delivery Plan",(http://www.pparc.ac.uk/ap/PPARC_Delivery_Plan06.pdf)

Benefits of Artificially Generated Gravity Gradients for Interferometric Gravitational-Wave Detectors

L. Matone[1], P. Raffai[2], S. Márka[1], R. Grossman[1], P. Kalmus[1], Z. Márka[1], J. Rollins[1], V. Sannibale[3]

[1] Columbia University, Department of Physics, New York, NY 10027, USA
[2] Eötvös University, Institute of Physics, 1117 Budapest, Hungary
[3] California Institute of Technology, LIGO Laboratory, Pasadena, CA 91125, USA

Abstract. We present an approach to experimentally evaluate gravity gradient noise, a potentially limiting noise source in advanced interferometric gravitational wave (GW) detectors. In addition, the method can be used to provide sub-percent calibration in phase and amplitude of modern interferometric GW detectors. Knowledge of calibration to such certainties shall enhance the scientific output of the instruments in case of an eventual detection of GWs. The method relies on a rotating symmetrical two-body mass, a Dynamic gravity Field Generator (DFG). The placement of the DFG in the proximity of one of the interferometer's suspended test masses generates a change in the local gravitational field detectable with current interferometric GW detectors.

1. Introduction

Dynamic gravity fields generated by rotating masses have been used previously in several experimental tests; however, their exploitation in conjunction with interferometric gravitational detectors has not been addressed until now. Forward and Miller [1] in 1967 developed a gravity field generator that allowed them to calibrate an orbiter sensor capable of measuring the lunar mass distribution. A similar technique was used by Weber *et al.* [2, 3] to calibrate a GW bar detector, where a volume of matter was acoustically stressed at 1660 Hz and the resulting noise excess in the detector was found to be consistent with theory. At the University of Tokyo, in the 1980s, a series of experiments were conducted to test the law of gravitation up to a distance of 10 m [4, 5, 6, 7, 8]. In these studies, the coupling between the dynamic field, generated by a rotating mass, and the quadrupole moment of a mechanical oscillator antenna was measured confirming the gravitational law within experimental uncertainties [7, 8]. In the 1990s, the gravitational wave group at the University of Rome developed and carried out experiments [9, 10] on the cryogenic GW bar detector, EXPLORER, at CERN. A device, with quadrupole moment of $\mathcal{M}_2 = 6.65 \times 10^{-2}$ kg m^2 and rotating in the frequency range of $450 - 470$ Hz, was developed to calibrate the antenna and was also used to confirm existing upper limits to Yukawa-like gravitational potential violations at laboratory scale.

The increased sensitivity and bandwidth of modern interferometric gravitational wave detectors warrants a new investigation into and opens exciting new possibilities for application of advanced gravity field generators in GW research. Presently interferometric gravitational wave detectors are reaching their design sensitivity enabling us to probe for gravitational radiation from sources well beyond the Local Group of galaxies. The response of these detectors to GW radiation is usually evaluated by direct injection of possible waveforms with known amplitude via magnetic actuators, also used for active control of the test masses' (essentially the interferometer mirrors) displacement. In addition, displacement in the test mass position can be induced by local gravity fields produced by a Dynamic gravity Field Generator (DFG). A DFG is essentially a symmetric rotating object with a significant quadrupole moment. When it is placed in the proximity of one of the interferometer mirrors, the induced change due to the device's quadrupole moment can be measured by the GW detectors such as the Laser Interferometer Gravitational Wave Observatory (LIGO) [11, 12], the VIRGO experiment [13], the 300 m Laser Interferometer Gravitational Wave Antenna (TAMA300) [14] and the GEO600 interferometer [15]. Future detectors, such as Advanced LIGO (AdLIGO) [16], offer higher sensitivity.

Several authors (see for example [17, 18, 19]) pointed out that gravity gradient (or Newtonian) noise, generated by density fluctuations in the Earth and the atmosphere, can be a potentially limiting noise source in advanced interferometric GW detectors. Motion of massive bodies (e.g. due to human activity) in the vicinity of the interferometer test masses also alters the local gravitational field, mainly at low

frequencies [20, 21, 17]. Gravity gradient noise manifests itself as an induced motion of the interferometer mirrors due to the fluctuation of the local gravity field. The DFGs described here can be used to modulate the local gravitational field around the test mass (TM) at a precise frequency and phase on a well-controlled manner and thereby directly validate/evaluate the expected noise generation and coupling mechanisms to complex structures.

In addition, DFGs have the potential to provide sub-percent amplitude and phase calibration of interferometric GW detectors. In the case of LIGO currently there are two calibration methods in use. The first one uses the interferometer TM's coil-magnet actuator to calibrate the gravitational wave channel (see for example [22] and [23]) while the second method uses the radiation pressure exerted on the TM by an independent laser source (see for example [24], [25], [26] and recently [27]). A DFG provides an alternative and independent sub-percent calibration, significantly improving the current accuracy of several percents (see e.g. [28]).

In this work we describe a hypothetical two-body DFG coupled to an ideal interferometric gravitational wave detector. The induced displacement on the suspended TM is dominated by the quadrupole moment of the DFG mass distribution in the case of a symmetric device. Any undesired system asymmetry will contribute to the dipole moment and can be measured and accounted for directly. We asses the application of such devices for the calibration of interferometric GW detectors as well their possible usage in gravity gradient noise studies that will eventually limit the performance of long baseline detectors at low frequencies.

Additionally, two DFGs in a null experiment setup can be used to explore violations to Newton's $1/r^2$ law well beyond the current limits. We investigated this possibility in detail for the LIGO, Advanced LIGO and Virgo detectors via numerical simulations. This is the subject of a separate publication. [29]

2. Newtonian field dynamics from a two-mass DFG

In analytical derivations, throughout this paper we will treat the suspended interferometer TM and the masses of the two-body DFG as point masses for simplicity. First we calculate the acceleration, along the laser beam axis, the mass is subjected to from a DFG configuration shown in fig.(1). Masses m_1 and m_2 are separated by a distance r_1 and r_2, respectively, from the center of rotation and are rotating at a frequency of $f_0 = \omega_0/(2\pi)$. The center of mass of mirror M and the DFG's center of rotation are separated by a distance d, where $d > r_{1,2}$.

Assuming that the distance between the DFG's i-th mass and the mirror is $h_{\rm i}$, the Newtonian potential at the mirror's center of mass is

$$V^{\rm c} = \sum_{i=1}^{2} V_{\rm i}^{\rm c} = -GM \sum_{i=1}^{2} \frac{m_{\rm i}}{h_{\rm i}}. \tag{1}$$

Introducing the variables $R_1 = r_1/d$, and $R_2 = -r_2/d$, $h_{\rm i}$, being a function of time can

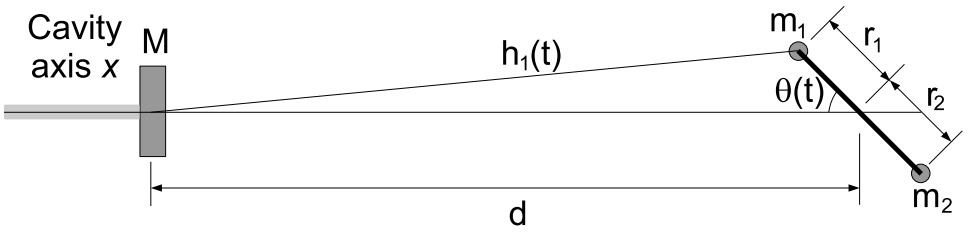


Figure 1. Schematic of an ideal symmetric two-mass DFG. The system consists of two masses, m_1 and m_2, separated by a distance of r_1 and r_2 from the center of rotation. The center of rotation is placed a distance d away from the mirror's center of mass. The system rotates at a frequency of $f_0 = \omega_0/(2\pi)$ where $\theta(t) = \omega_0 t$. The x axis denotes the interferometer's optical axis and only accelerations along this axis are considered.

be written as

$$h_{\rm i}(t) = d\,\sqrt{1 + R_{\rm i}^2 - 2R_{\rm i}\cos\theta(t)} \tag{2}$$

where $\theta(t) = \omega_0 t$ (see fig.(1)). The magnitude of the TM's induced acceleration along the laser beam axis is

$$a^{\rm c} = \frac{1}{M}\left|\frac{\partial V^{\rm c}}{\partial d}\right| = \frac{G}{d^2}\sum_{i=1}^{2} m_{\rm i} B_{\rm i}(R_{\rm i},\theta). \tag{3}$$

Here $B_{\rm i}(R_{\rm i},\theta)$ is a geometrical factor

$$B_{\rm i}(R_{\rm i},\theta) = \frac{1 - R_{\rm i}\cos\theta}{(1 - 2R_{\rm i}\cos\theta + R_{\rm i}^2)^{3/2}}. \tag{4}$$

For the case of a much smaller lever arm $r_{\rm i}$ than the distance d ($R_{\rm i} \ll 1$) we can expand $V^{\rm c}$ thereby expressing the induced acceleration $a^{\rm c}$ in terms of the n-th multipole moment $\mathcal{M}_{\rm n}$ of the DFG's mass distribution

$$a^{\rm c} = \frac{G}{d^2}\sum_{n=0}^{\infty}\frac{n+1}{d^n}\cdot\mathcal{M}_{\rm n}\cdot P_{\rm n}(\cos\theta) \tag{5}$$

where

$$\mathcal{M}_{\rm n} = m_1 r_1^n + (-1)^n m_2 r_2^n \tag{6}$$

and $P_{\rm n}(\cos\theta)$ is the Legendre polynomial of n-th order.

We remark that the DFG's dipole moment, as well as the higher-order odd moments, contribute only to the odd harmonic terms, whereas the quadrupole moment and the higher-order even terms, contribute only to the even harmonic terms. In the case of an ideally symmetric DFG, all odd moments vanish and the induced displacement is dominated by the quadrupole moment $\mathcal{M}_2$ at twice the rotation frequency.

2.1. Induced Displacement from the Newtonian Potential

The suspended TM can be considered as a free body for frequencies well above the eigenfrequencies of the suspension which typically lie around 1 Hz [30]. Neglecting the

time-independent term, double integrating eq.(5) with respect to time and considering only the dominant terms in the first few harmonics, the TM's displacement along the laser beam axis, x, can be written as

$$x(t) \simeq \frac{G}{(d\ \omega_0)^2} \times \left[2 \cdot \frac{\mathcal{M}_1}{d} \cdot \cos\omega_0 t + \frac{9}{16} \cdot \frac{\mathcal{M}_2}{d^2} \cdot \cos 2\omega_0 t + \frac{5}{18} \cdot \frac{\mathcal{M}_3}{d^3} \cdot \cos 3\omega_0 t \right] \tag{7}$$

In the case of a symmetric two-mass DFG, the dipole and the octopole contribution vanishes and the quadrupole moment $\mathcal{M}_2$ dominates. For initial LIGO throughout the paper we will consider the case of $m_1 = m_2 = 1.5$ kg, $r_1 = r_2 = 0.25$ m (equivalent to a quadrupole moment of $\mathcal{M}_2 = 0.1875$ kg m^2), with a rotation frequency of $f_0 = \omega_0/(2\pi) = 51$ Hz and a distance of $d = 2.5$ m. The resulting RMS displacement change $x_{\rm rms}$ at twice the rotation frequency is 1.24×10^{-18} m and scales according to

$$x_{\rm rms} \simeq 1.24 \times 10^{-18}\ {\rm m} \times \left(\frac{\mathcal{M}_2}{0.1875\ {\rm kg\ m^2}} \right) \times \left(\frac{51\ {\rm Hz}}{f_0} \right)^2 \times \left(\frac{2.5\ {\rm m}}{d} \right)^4 \tag{8}$$

Fig.(2) shows the design sensitivities for initial LIGO, AdLIGO and VIRGO also including LIGO's nominal displacement sensitivity for the beginning of the fifth science run [31] (S5) (data taking began in November 2005 and it is scheduled to collect one year of integrated data). The LIGO detectors' displacement sensitivity at 102 Hz is $\sim 2 \times 10^{-19}$ m$/\sqrt{\rm Hz}$ (see gray curve in fig.(2)).

The signal-to-noise ratio (SNR), defined as the ratio of the RMS signal to the displacement noise spectrum density integrated for a time T, gives a measure of how much a given stimulus is above background. For the above mentioned device, in the case of LIGO during S5 at 102 Hz (that is twice the above mentioned rotation frequency), and for an integration time of 1 s, we obtain an SNR of 6. In general terms, for an arbitrary noise floor $\tilde{n}$, and integration time T, the SNR scales as

$$SNR \simeq 6 \times \left(\frac{2 \times 10^{-19}\ {\rm m}/\sqrt{\rm Hz}}{\tilde{n}} \right) \times \left(\frac{T}{1\ {\rm s}} \right)^{1/2} \times \left(\frac{x_{\rm rms}}{1.24 \times 10^{-18}\ {\rm m}} \right) \tag{9}$$

where $x_{\rm rms}$ is shown in eq.(8). At the present sensitivity level of LIGO it is possible to sense such a dynamically changing gravity field from the DFG in question using a relatively small integration time.

Fig.(3) shows the SNR for different detectors as a function of twice the rotational frequency with an integration time of half an hour. Using once again the example cited above (DFG of quadrupole moment $\mathcal{M}_2 = 0.1875$ kg m^2), the top portion of the figure shows that for rotational frequencies ranging between 10 Hz and 500 Hz a distance of 2.5 m can be used for the initial LIGO detectors.

The center part of fig.(3) shows the SNR for the VIRGO detector. Due to the detector's sensitivity at low frequencies, low rotational frequencies, as low as ~ 10 Hz,

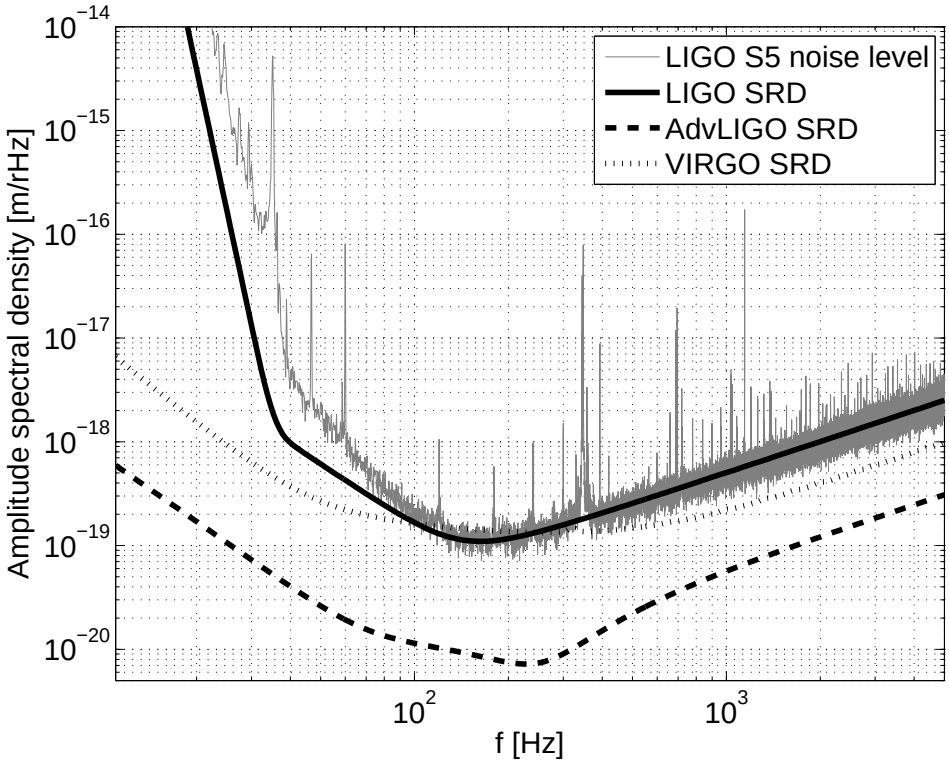


Figure 2. The nominal displacement sensitivity of LIGO (gray trace) at the beginning of the fifth science run (S5) together with its design curve (black); the design sensitivity (SRD) for Advanced LIGO (dashed) and the European VIRGO detector (dotted) are also shown.

could be used. The bottom portion of fig.(3) shows the response from the Advanced LIGO interferometer.

We would like to point out, that the choice of the rotational frequency as well as the DFG placement affects the level of spurious couplings. This is addressed in the sections to come.

2.2. A hypothetical DFG design

In fig.(4) we show a hypothetical DFG design based on the parameters discussed in this section. It consists of an Aircraft Grade (6Al/6V/2Sn) Titanium disc 60 cm in diameter and 10 cm in height. The disc has two cylindrical slots, 50 cm apart, which can hold different materials. The choice of materials was motivated by the desire to maximize density difference and strength while still keeping the material cost within the bounds of reason. We use Tungsten cylinders, 3.6 cm in diameter, as an example in the following sections. Practical details, such as the expansion and stress factors of the DFG under prolonged operating conditions must be modeled and simulated by finite element analysis methods, then subsequently measured and taken into account. These studies are beyond the scope of this paper.

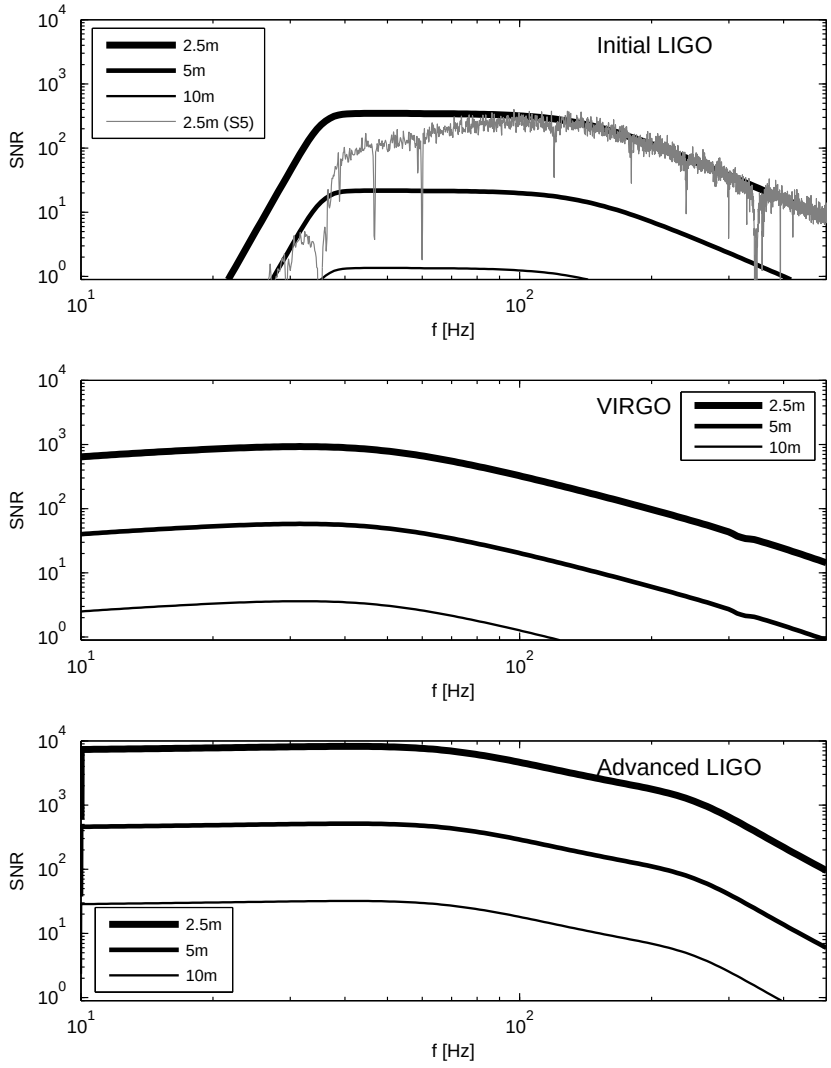


Figure 3. The Signal-to-Noise Ratio due to a DFG of quadrupole moment $\mathcal{M}_2 = 0.1875$ kg m^2 and 1/2 hour of integration time. Top: initial LIGO with the DFG positioned 2.5, 5 and 10m away from the TM; Middle: VIRGO for positions of 2.5, 5, and 10m; Bottom: Advanced LIGO with distances to the TM of 2.5, 5 and 10m.

2.3. Gravity Gradient Noise Studies with DFGs

A DFG in the proximity of the test mass of the interferometer can be used to experimentally investigate and model the coupling between the varying gravity field and the complex suspension system of the test mass in many fundamental configurations. The artificial gravity gradient field generated by a DFG not only couples to the test mass but also into all stages of the multistage suspension system and gives rise to possible second order effects. By varying the placement and rotation frequency of the

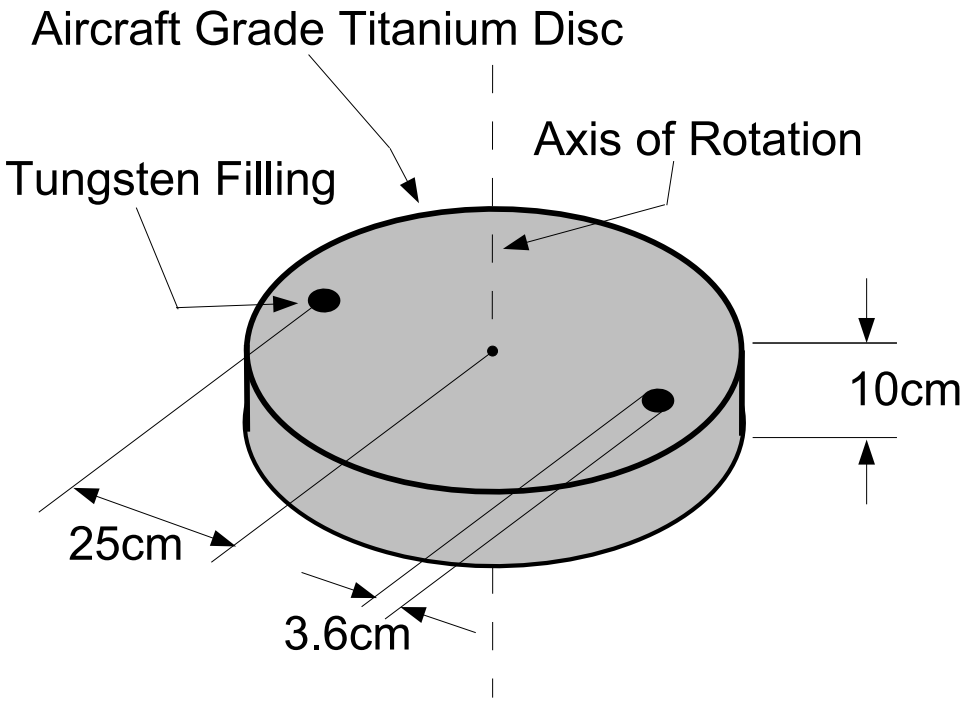


Figure 4. Sketch of a hypothetical DFG. The DFG consists of an Aircraft Grade (6Al/6V/2Sn) Titanium disc 50 cm in diameter and 10 cm in height. It holds two Tungsten cylinders at 20 cm from the rotation axis. The radius of the Tungsten cylinders is 1.8 cm.

DFG, this artificial gravity field can simulate conceivable gravity gradient noise sources specific to the local environment of the interferometric detector in question. With a DFG the dependence of the TM's displacement on the orientation of the gravity gradient noise source can be mapped: the DFG can be installed at different distances from the TM in the axis of the laser beam as well as placed off-axis and out of the plane of the interferometer. This is especially important since gravity gradient noise couples to the system from each direction on different ways thus potentially introduces problems into the detection chain via hard to track second order effects and possible nonlinear couplings. Additional advantage of artificially generated dynamic gravity gradients is that the frequency dependence of the interferometer's response to Newtonian noise sources could be mapped out in detail, which is especially important for the low frequency region. The results might eventually be used in generating approaches for mitigating the effect of local gravity gradients in future detectors at low frequencies besides providing accurate information about the nature of this noise source.

2.4. Calibration of an Interferometric GW Detector using a DFG

In this section we address the level of precision we must achieve when using the DFG as a calibration tool. While the present calibration accuracy of 2-10% in amplitude and phase [28] (depending on frequency range) might seem adequate for upper limit

and event rate studies, it will be important to know the calibration of the detector to a higher accuracy when the collaborations enter the "detection era". In the context of a detected signal via a global network of interferometric GW detectors, where the waveform is recoverable, amplitude and phase calibration known to a higher precision shall be beneficial. Coherent network methods will perform better, pointing accuracy will increase, source distance information can be recovered and used with a higher accuracy. With an increase in calibration certainty the precision of waveform and polarization recovery is expected to improve, which in turn allows for better scientific output.

With the DFG method, the achievable calibration accuracy would be limited by the uncertainty in the gravitational constant, G, at the subpercent level. To estimate the calibration uncertainty we first consider the TM displacement $x_{\rm rms}$ induced by the DFG due to its quadrupole moment $\mathcal{M}_2$ is given in eq.(8). In statistical terms (assuming a large number of DFGs identical within practical tolerances), the relative uncertainties in the measurement of the gravitational constant ($\delta G/G$), in the mass ($\delta m/m$), arm length ($\delta r/r$), rotation frequency ($\delta f_0/f_0$) and distance from the TM center of mass ($\delta d/d$) add in quadrature leading to a relative uncertainty on the induced displacement ($\delta x/x$) and is approximately described by

$$\left(\frac{\delta x}{x}\right)^2 \simeq \left(\frac{\delta G}{G}\right)^2 + \left(\frac{\delta m}{m}\right)^2 + 4\left(\frac{\delta r}{r}\right)^2 + \\ + 4\left(\frac{\delta f_0}{f_0}\right)^2 + 16\left(\frac{\delta d}{d}\right)^2 \tag{10}$$

Our goal is to achieve sub-percent precision in amplitude calibration, therefore we need to keep the relative uncertainties of every DFG parameter well below $\simeq 0.1\%$.

The currently accepted value of the gravitational constant, G, is $((6.6742\pm0.0010)\times 10^{-11}\mathrm{m}^3\mathrm{kg}^{-1}\mathrm{s}^{-2})$. This means, that there is a $\simeq 0.015\%$ contribution to the relative uncertainty on the induced displacement just by taking into consideration the precision of previous G measurements. G contributes as the leading term in limiting the precision of amplitude calibration if the uncertainties related to manufacturing and/or measurement of the other parameters contributing to each of the other four terms in equation (10) is below 0.015%. Thus we require

$$\begin{aligned} \frac{\delta m}{m} &\leq 1.5\times 10^{-4} \\ \frac{\delta r}{r} &\leq 7.5\times 10^{-5} \\ \frac{\delta f_0}{f_0} &\leq 7.5\times 10^{-5} \\ \frac{\delta d}{d} &\leq 3.75\times 10^{-5} \end{aligned} \tag{11}$$

These levels of uncertainties adding up in quadrature yield 0.035% uncertainty in ($\delta x/x$), more than adequate for a sub-percent amplitude calibration.

Considering the DFG described in sec.(2.1), with $m_1 = m_2 = 1.5$ kg, $r_1 = r_2 = 0.25$ m, rotation frequency of $f_0 = 51$ Hz and distance from the TM of $d = 2.5$ m, the required uncertainties in (11) translate to

$$\begin{aligned} \delta m &= 2.25 \times 10^{-4} kg \\ \delta r &= 1.9 \times 10^{-5} m \\ \delta f_0 &= 3.8 \times 10^{-3} Hz \\ \delta d &= 9.4 \times 10^{-5} m \end{aligned} \tag{12}$$

Most off the shelf balances can be used to measure the DFG masses, while the ultimate precision of mass determination, δm, is ~ 50 μg [32] with a state of the art mass comparator. Uncertainties on δr are determined by machining precision and can be kept within ~ 1 μm [33].

The uncertainty in the rotational frequency f_0 can be addressed by using a precision optical encoder to provide pulses which can be used to phase lock the DFG to an atomic clock or GPS. In this case, the uncertainty is limited by the encoder itself or the servo system. For a rotation period of $1/f_0 = 20$ ms and an off-the-shelf 16-bit optical encoder providing a square pulse train at 3.2 MHz ($\simeq 300$ ns/pulse), the relative position of the square wave rising edge with respect to the atomic clock signal can be determined within $\delta t \simeq 10$ ns. This allows for a high precision of $\delta f_0/f_0 \sim 10^{-6}$.

Distance d could change somewhat over time when a DFG is used as a calibration device. The thermal variations in the TM and DFG housings are kept within fractions of a degree and should not play a significant role. The tidal-compensation system, a servo-mechanism acting on the position of the TM to compensate for earth-moon and earth-sun tidal effects, displaces the TM locally with peak to peak excursions of the order of ~ 300 μm (see [34] and [35]). This kind of excursion can be taken into account during the calibration.

Distance d can be directly measured via laser based range finding (i.e. Light Detection and Ranging, LIDAR) technologies, which can provide better than $\delta d \simeq 1$ μm uncertainty in lab environments [36].

When direct distance measurement between the TM and the DFG is not possible, an alternative method for finding d can be adopted. The $2\omega_0$ component can be measured as a function of d by varying the DFG's position by a well known amount and using a χ^2 minimization procedure to estimate the effective distance d. For simplicity, let the distance vary linearly

$$d(t) = d_0 + v\ t \tag{13}$$

where v is the DFG's pivot velocity along the beam axis. Following eq.(8), the uncalibrated interferometer response R_{IFO} to the DFG's stimulus can be described as

$$R_{IFO} = \frac{K}{(d_0 + v\ t)^4} \tag{14}$$

where K and d_0 are free parameters. A linear sweep of the pivot's position would provide an estimate of d_0 while any residual would provide information on any $d(t)$ component that could potentially be significant. The uncertainty in d_0 will be statistical in nature and eventually will be limited to the systematic uncertainty of the other parameters, such as the dipole moment and the rotation frequency. In this case

$$\frac{\delta d}{d} \sim \frac{\delta r}{r} \sim 10^{-5} \tag{15}$$

There are also other uncertainties that need to be addressed for realistic measurements, most of them are second order in nature. For example, an accurate alignment of the DFG is necessary: the effective arm length $\tilde{r}$ is altered if the plane of rotation of the DFG is not aligned with the plane of the interferometer. Restricting this change to 19 μm (same as the uncertainty in r) restrains the leveling of the DFG to 0.7°, which is achievable with commercial optical positioning methods.

Placing the DFG out of line with the Fabry-Perot arm introduces another second-order source of error. This error creates a distance $\tilde{d}$ which differs from d. Requiring their relative change $(\tilde{d} - d)/d$ to be of the order of 10^{-5} sets an alignment requirement to the cavity with an order of 1 cm, which is easily achievable with optical positioning.

The phase of the rotating DFG can be measured by phase locking the DFG to an atomic clock or GPS. The phase uncertainty (in the first order) is therefore due to $\delta t/t$. Taking the previously mentioned $\delta f_0/f_0 \sim 10^{-6}$ value as a basis, the precision of phase calibration can even be better than $\simeq 0.01\%$.

The quoted accuracy of calibration for the LIGO detector for recent science runs [28] is at the 6-10% level and valid for a broad range of frequencies and for the entire length of the science run. The inherent accuracy of the calibration method itself is at the order of 1-2% [37]. Using DFG as calibration tool this can be pushed down to the subpercent level for amplitude calibration and an even better accuracy can be achieved for phase calibration.

To take full advantage of this proposed calibration method for interferometric GW detectors, we envision a DFG positioned at around 2-3 m from each end mirror of the two arms of the interferometer. The rotation frequencies can be chosen such that subpercent level calibration could be provided for the most sensitive region of the detector response. The employment of two separate DFGs, rotating at slightly different frequencies, would allow the calibration of the two interferometer arms separately in a spectrally similar region. Additionally, with longer integration times higher order harmonics become detectable, thus the device can be used for calibration of interferometer response of frequency regions at points separated by the DFG's rotation frequency.

3. Mitigation of Spurious Couplings from the DFG's Motor

In interferometric GW detectors, using DFGs as a calibration tool means that the new device will be put in close proximity (e.g. 2.5 meters) of the test mass for a prolonged period of time, while the GW detector itself is in a continuous data taking mode. Thus,

it is necessary that spurious coupling of the DFG to the suspended mirror be negligible, as detailed in this chapter. The only acceptable effect on the GW data should be the fine and easily filterable lines at the multiples of the rotational frequency of the DFG. Of most concern is the electro-magnetic coupling via the motor driving the system, the acoustic coupling via the local interferometer optical sensors and the seismic vibrations induced by an unbalanced DFG.

3.1. Electro-Magnetic Coupling

There are two ways the motor's electro-magnetic field could couple to the test mass. One coupling is the interaction of the motor's electro-magnetic (EM) field with the interferometer electronics residing next to the DFG. The other way is through the coupling of the DFG's EM field with the coil-magnet system needed to drive the TM in position.

With the proper Electromagnetic Interference (EMI) shielding in place and using DC permanent magnet servo motors the parasitic emission can be mitigated. The DFG could be equipped with a non-integer gear ratio to completely separate the EM harmonics from the Newtonian signal since the induced displacement appears at harmonics of the rotation frequency of the DFG and not of the motor.

It is also possible to completely eliminate the mechanical coupling via an Eddy Current Motor, which also simplifies the DFG, balancing and bearing design.

3.2. Acoustic coupling

For the LIGO interferometers, acoustic signals near the detector could potentially couple directly to the gravitational wave channel. A possible coupling mechanism could consist of an acoustic stimulus exciting the beam position on an optical sensor. If the sensor in question is used to feedback on TM positions, the acoustic excitation finds its way into the detector. This effect is mitigated by installing the DFG in its own vacuum envelope.

3.3. Seismic coupling

The most obvious concern is the TM coupling through the ground. For an ideally symmetric DFG, as described in earlier sections, the dipole moment vanishes and so does its contribution to the Newtonian field. Any asymmetry in the system creates a non-null dipole moment at the rotation frequency, introducing ground vibration. In this section we use a simple model to estimate this cross-coupling for the initial LIGO case.

For an asymmetric DFG, the device's center of rotation will be subjected to a sinusoidal force F' at the rotation frequency ω_0 whose RMS value along the beam axis can be written as

$$F'_{\rm rms} = \frac{1}{\sqrt{2}}\,\omega_0^2\,\mathcal{M}_1 \tag{16}$$

where $\mathcal{M}_1$ is the dipole moment of the DFG. The displacement $\delta x_{\rm react}$ of the reaction mass due to the asymmetry, to first order approximation, is

$$\delta x_{\rm react} = \frac{m\, r}{\sqrt{2}\, M_{\rm react}}\ (\epsilon_r + \epsilon_m) \tag{17}$$

where $\epsilon_r = \delta r/r$ and $\epsilon_m = \delta m/m$. The TM displacement can then be expressed as

$$\delta x_{\rm rms} = \delta x_{\rm react}\ {\rm R(f)} \tag{18}$$

where R(f) is the attenuation factor provided by LIGO's seismic isolation stage and suspension.

To estimate the motion of the cement slab beneath both the DFG and the TM we select achievable uncertainty requirements of 11. For a plausible reaction mass $M_{\rm react}$ of 100 tons (assuming a concrete slab $10m \times 10m \times 0.5m$) its mass displacement is

$$\delta x_{\rm react} = 6 \times 10^{-10}{\rm m} \tag{19}$$

LIGO's stack[38] reduces this displacement down by a factor of $\sim 10^6$ at 51 Hz while the suspension stage [30] brings it down by an another factor of $\sim (51\ {\rm Hz}/0.74\ {\rm Hz})^2 = 4500$. This results in a TM displacement of

$$\delta x_{\rm rms} = 1.3 \times 10^{-19} m \tag{20}$$

which is below the noise floor of LIGO and is a negligible spurious coupling.

The above estimated effect of seismic coupling can be further reduced by attaching the rotating DFG to a light slab with very small reaction mass $M_{\rm react}$. The seismic signal of a high-precision seismometer coupled to the slab, resulting from system asymmetries, can be substantially reduced by iterative adjustment of the balancing of the DFG. Attaching this balanced DFG to a heavy slab with higher $M_{\rm react}$ will reduce $\delta x_{\rm react}$ to be well below the ambient seismic field. The reduction factor is given by the ratio of the reaction mass of the light slab to the reaction mass of the heavy slab. This can lead to a TM RMS displacement even orders of magnitude smaller than as given in eq.(20).

4. Safety

Significant kinetic energy (i.e. tens of kJs) is stored in the DFG once it rotates and crucial safety considerations must be addressed. There are two major points of failure management to be concerned with. (a.) The vacuum chamber of the DFGs must be made strong enough to withstand the damage of an accidentally disintegrating disk. This is the standard solution for high speed gyroscopes. (b.) For added security, the gap between the inner wall of the vacuum chamber and the outer edge of the rotating disk must be kept relatively *small*. In the event of an incident where the DFG's material starts to yield or its angular acceleration is uncontrolled the disk will expand radially touching the sidewall and slowly stop, preempting a catastrophic failure. These conditions can be met using Finite Element Analysis (FEA) aided design, in-house destructive testing of sacrificial parts and relying only on X-ray rated base materials.

5. Conclusion

These initial feasibility studies of simple DFGs indicate that they are capable of dynamically changing the local gravitational field by an amount detectable by current interferometric gravitational wave detectors.

The DFGs can be designed, manufactured, tuned and characterized to be symmetric and safe enough to eliminate concerns about vibrations and spurious couplings, once positioned in the proximity of one of the suspended TMs.

The generated gravity gradient signal is proportional to the DFG's quadrupole moment with its signature appearing at twice the rotation frequency. At the present detector sensitivity level of LIGO, systematic uncertainties due to the DFGs can be well below the 0.1% level in amplitude with insignificant timing uncertainties. This apparatus provides a detector-independent calibration technique that can significantly surpass the achievable precision of other existing calibration methods.

The DFG also offers a unique and distinctive way to generate a differential arm length displacement for gravitational wave detectors. Apart from calibration objectives, it could also be used to validate the expected noise generation and coupling mechanism of Newtonian noise, possibly a limiting factor in advanced gravitational wave detectors.

There are many details that need attention when designing and manufacturing a practical device. Finite element analysis of the DFGs and subsequent experimental studies are necessary to completely understand the stresses the DFG is subjected to. The DFGs will be enclosed in a separate vacuum chamber. A prototype design and test will be necessary to balance the disk and test vibration control. Other mostly practical problems, such as safety, can also be solved as was shown in past applications/experiments that have used rapidly rotating instruments.

6. Acknowledgements

The authors gratefully acknowledge the support of the United States National Science Foundation for the construction and operation of the LIGO Laboratory and the Particle Physics and Astronomy Research Council of the United Kingdom, the Max-Planck-Society and the State of Niedersachsen/Germany for support of the construction and operation of the GEO600 detector. The authors also gratefully acknowledge the support of the research by these agencies and by the Australian Research Council, the Natural Sciences and Engineering Research Council of Canada, the Council of Scientific and Industrial Research of India, the Department of Science and Technology of India, the Spanish Ministerio de Educacion y Ciencia, The National Aeronautics and Space Administration, the John Simon Guggenheim Foundation, the Alexander von Humboldt Foundation, the Leverhulme Trust, the David and Lucile Packard Foundation, the Research Corporation, and the Alfred P. Sloan Foundation.

The LIGO Observatories were constructed by the California Institute of Technology and Massachusetts Institute of Technology with funding from the National Science

Foundation under cooperative agreement PHY-9210038. The LIGO Laboratory operates under cooperative agreement PHY-0107417.

The authors are also grateful for the support of the United States National Science Foundation under cooperative agreement PHY-04-57528 and Columbia University in the City of New York. We greatly appreciate the support of LIGO collaboration. We are indebted to many of our colleagues for frequent and fruitful discussion. In particular we'd like to thank G.Giordano, R.Adhikari, V.Sandberg, M.Landry, P.Sutton, P.Shawhan, R.DeSalvo, H.Yamamoto, Y. Aso and C.Matone for their valuable comments on the manuscript.

This paper has been assigned LIGO Document Number LIGO-P060056-00-Z.

7. References

[1] Forward, R. L. and Miller, L. R. *J. Appl. Phys.*, 38:512–+, 1967.
[2] J. Sinsky and J. Weber. New Source for Dynamical Gravitational Fields. *Physical Review Letters*, 18:795–797, May 1967.
[3] J. A. Sinsky. Generation and Detection of Dynamic Newtonian Gravitational Fields at 1660 cps. *Physical Review*, 167:1145–1151, March 1968.
[4] H. Hirakawa, K. Tsubono, and K. Oide. Dynamical test of the law of gravitation. *Nature*, 283:184–+, January 1980.
[5] K. Oide, K. Tsubono and H. Hirakawa. *Jap. J. Appl. Phys.*, 19:L123, 1980.
[6] T. Suzuki, K. Tsubono and K. Kuroda. *Jap. J. Appl. Phys.*, 20:L498, 1981.
[7] Y. Ogawa, K. Tsubono, and H. Hirakawa. Experimental test of the law of gravitation. *Phys. Rev. D.*, 26:729–734, August 1982.
[8] K. Kuroda and H. Hirakawa. Experimental test of the law of gravitation. *Phys. Rev. D.*, 32:342–346, July 1985.
[9] Astone P. et al. *Z. Phys. C*, 50:21–+, 1991.
[10] Astone P. et al. Experimental study of the dynamic Newtonian field with a cryogenic gravitational wave antenna. *European Physical Journal C*, 5:651–664, 1998.
[11] Abbott B. et al. Detector description and performance for the first coincidence observations between LIGO and GEO. *Nuclear Instruments and Methods in Physics Research A*, 517:154–179, January 2004.
[12] D. Sigg and the LIGO Science Collaboration. Status of the LIGO detectors. *Classical and Quantum Gravity*, 23:51–+, April 2006.
[13] Acernese, F. et al. *Classical and Quantum Gravity*, 23:S63, 2006.
[14] Takahashi, R. et al. *Classical and Quantum Gravity*, 21:S403, 2004.
[15] Lück, H. et al. *Classical and Quantum Gravity*, 23:S71, 2006.
[16] http://www.ligo.caltech.edu/advLIGO.
[17] P. R. Saulson. Terrestrial gravitational noise on a gravitational wave antenna. *Phys. Rev. D.*, 30:732–736, August 1984.
[18] S. A. Hughes and K. S. Thorne. Seismic gravity-gradient noise in interferometric gravitational-wave detectors. *Phys. Rev. D.*, 58(12):122002–+, December 1998.
[19] Beccaria M. et al. Relevance of Newtonian seismic noise for the VIRGO interferometer sensitivity. *Class. Quantum Grav.*, 15:3339–3362, 1998.
[20] R. Spero. Prospects for Ground Based Detectors of Low Frequency Gravitional Radiation. *AIP Conf. Proc.*, 96:347–350, 1983. in Science Underground.
[21] T. Suzuki and H. Hirakawa. Local Field and Gravitational Radiation Experiments. *J. Phys. Soc. Japan*, 48(2):685–686, 1980.

[22] L. Matone et al. Input Test Mass (ITM) Absolute Calibrations: Fringe Counting, Fringe Fitting, and Sign Toggling Methods. 2002. LIGO Internal Document T020141.
[23] R. Adhikari et al. Calibration of the LIGO detectors for the First LIGO Scientific Run. 2003. LIGO Internal Document T030097.
[24] D. Sigg. Strain Calibration in LIGO. 1997. LIGO Internal Document T970101.
[25] J. Bruursema. Calibration of the LIGO Interferometer Using the Recoil of Photons. 2003. LIGO Internal Document T030266.
[26] E. Goetz. Commissioning of the Photon Calibrators. 2004. LIGO Internal Document T040196.
[27] P. Kalmus. Commissioning of the Photon Calibrators for S5. 2005. LIGO Internal Document T050221.
[28] G. Gonzalez et al. Calibration of the LIGO detectors for S3. 2005. LIGO Internal Document T050059.
[29] P. Raffai, L. Matone, I. Bartos, Z. Marka, S. Marka. Tests of non-Newtonian Gravity Using Dynamic Gravity-Gradients in Interferometric Gravitational-Wave Detectors. *Phys. Rev. D.*, to be submitted.
[30] J. Hazel S. Kawamura and F. Raab. Suspension Preliminary Design. 1996. LIGO Internal Document T960074.
[31] LIGO collaboration. Displacement Sensitivities for the LIGO Interferometers - Early Performance for S5. 2006. LIGO Internal Document G060011.
[32] Sartorius CC10000U-L, http://sartorius.balances.com/sartorius/mass-comparators.html.
[33] in the case of Titanium, see for example http://www.cranfieldprecision.com/.
[34] E. Morganson. Developing an Earth-Tides Model for LIGO Interferometers. 1999. LIGO Internal Document T990181.
[35] L. Matone. Slow Ground Motion Observed by LIGO. 2000. LIGO Internal Document G000325.
[36] K. Minoshima and H. Matsunoto. *Conference on Lasers and Electro-Optics Europe High-precision distance measurement using the frequency comb of an ultrashort pulse laser*, page 440, 2005.
[37] personal conversation, M. Landry.
[38] J.Giaime et al. *Rev. Sci. Instr.*, 67:208, 1996.

From the *Proceedings of the Gravitational-Wave Conference*, edited by P. Murad and R. Baker, The MITRE Corporation, Mclean, Virginia, May 6-9, 2003, Paper HFGW-03-117.

Generation of High-Frequency Gravitational Waves (HFGW) by Means of an Array of Micro- and Nano-Devices

by
Robert M. L. Baker, Jr.†

ABSTRACT

The process by which High-Frequency Gravitational Waves (HFGW) are generated by means of the time rate of change of the acceleration of a mass or masses, termed a "jerk" or a "shake," is developed. Arrays of micro- and nano-devices, termed energizing and energizable elements, are utilized to generate a train of coherent gravitational waves. As the waves progress along the axis of such devices they are reinforced by the energizable elements, under the control of a computer logic system in order to be modulated for applications such as communication. Starting with a theoretical non-rotating, but ratcheting or jerking rim or ring, linear devices, such as a stack of ratcheting rims or rings, evolve. These devices emulate a rotating rim. But the changing centrifugal force vector of a rotating rim or ring, which is tangent to rim and represents a jerk, is replaced by the electromagnetically, reciprocally jerked energizable elements of a non-rotating rim that do not involve large g loads. Two specific devices are described to illustrate the concept: the first such device involves a barrel or stacks of concentric rims whose surface or edges are covered with an array of ultra-small micromagnets (energizable elements) surrounded by a sheath of ultra-small microcoils (energizing elements); it generates approximately 380 kW of HFGW. The second is much smaller, 18 mm in length, and through use of a superconducting lens delivers a flux of 6.3×10^{-7} watts per square meter at a 7 km distant receiver – about that of a handheld radio transceiver. Other HFGW generators utilizing nanowire lattices for a high-frequency nanomechanical resonators and parallel current-carrying plates are also investigated and the question of null GW generation for symmetrical systems is considered.

The problem, which all of the devices discussed in this paper solve, is to cause a system of masses, which could be mini-magnets, micro-devices, nano-devices, individual molecules, submicroscopic particles, or individual electrons (as in a superconductor), under computer logic control, to move in concert with a jerk in order to *build up* (generate) HFGW with either planar or cylindrical wave propagation. Such jerking masses produce a very long sequence of HFGW pulses having significant average power and ability to carry information without generating incapacitating heat, causing disruptive g loads, or producing overpowering EM radiation.

Nomenclature

A	area
a	semi-major axis of a two-body orbit
a	acceleration
B	magnetic flux density
c	speed of light or, alternatively, the electron mobility speed

---

† Senior Consultant, GRAVWAVE® LLC and Transportation Sciences Corporation, 8123 Tuscany Avenue Playa del Rey, California 90293, USA. Telephone: (310) 823-4143. E-mail: robert.baker.jr@comcast.net.

Copyright 2003 by R. M. L. Baker, Jr. Published by The MITRE Corporation with permission.

$D_{\alpha\beta}$	quadrupole moment-of-inertia tensor
d	diameter or distance between conductors
E	energy
e	eccentricity of a two-body orbit
F	force per unit length
f	force
$\mathbf{f_{cf}}$	centrifugal-force vector
G	universal gravitational constant
h	strain in the space-time continuum , ? *l*/*l* caused by GW passage
I	moment of inertia
i	current
l	length
M	mean anomaly for a two-body orbit
m	mass of an object on orbit in characteristic units or of any object in kilograms
m	sum of the masses of a pair of binary stars or mass of a rod in kilograms
N	noise or index of GW refraction or number of pulses in an activation time interval
n	mean motion for a two-body orbit or number of objects or elements or number of rim or ring sections in a stack
n	number of coil turns or number of concentric rims or rings
P	the magnitude of the power of a gravitational-radiation source
p	parameter or semilatus rectum = $a(1-e^2)$
q	charge or periastron distance, a(1-e)
R	resistance or range
r	radial distance to an object on orbit; alternately, the effective radius of gyration
r	radius of a magnetic core, piston or barrel or stack of rims
S	GW flux
s	distance or displacement
t	time
t`	spinning-rod time
V	volume or speed
v	true anomaly of a two-body orbit
v	velocity or frequency
x	axis of orthogonal coordinate system
y	axis of orthogonal coordinate system
z	axis of rotation orthogonal to x and y axis
α	attenuation or diffraction angle
Δ	small increment
Δf_{cfx}	incremental x component of centrifugal force
Δf_{cfy}	incremental y component of centrifugal force
Δt	time increment
d	fraction of a linear-motor, GW generator's barrel radius that is an energizing-element sheath and/or energizable-element core
δl	thickness of a given rim or ring
δm	differential mass
δt	differential time or activation time interval
?	the central angle of a rotating rod
κ_{I3dot}	coefficient (constant or function) of the kernel in the d^3I/dt^3 formulation of the quadrupole
λ	wavelength
μ	$= m_1 + m_2$ = sum of masses on a two-body orbit in characteristic units
μ_0	permeability of free space
ν	frequency
σ	absorption cross section
τ	characteristic time; for heliocentric unit systems = 5.022×10^5 seconds
ω	angular rotational rate

Subscripts

1	refers to mass one
2	refers to mass two
a	current in one wire
b	current in adjacent wire
cf	centrifugal
d	diffraction
GW	gravitational wave
l	longitudinal
p	phase
r	radial
t	tangential
x	x component
y	y component

1. INTRODUCTION

The general concept of the devices discussed in this paper is to simulate scientifically acceptable generation of gravitational waves (GW) like those that are produced by energizable celestial systems such as rotating binary stars, star-black-hole collisions, star explosions, star collapse, binary black holes, spinning black holes, etc. through the use of smaller macro- and micro-, terrestrial or laboratory energizable systems. Such terrestrial systems generate well over 40 orders of magnitude more force intensity by virtue of their use of non-gravitational forces (nuclear or electromagnetic compared to gravitational) than a typical celestial system and well over 12 orders of magnitude greater frequency (THz or QHz {10^{15} Hertz; the term Quadrahertz, QHz, is preferred over Petahertz or PHz} and higher compared to kHz or very small fractions of a Hz) than a typical celestial system. Terrestrial energizable systems produce significant and useful GW according to the various designs of the devices to be described, even though they are orders of magnitude smaller and less massive than the extraterrestrial celestial systems. In the various designs of these devices, large numbers of small energizable elements are energized in sequence or in concert, by energizing or stimulating elements, to emulate the motion of a much larger and extended body in order to enhance the generation of GW.

The specific concept, which will be expanded upon, requires applying a long series of rapid "jerks" or "shakes" or third-time-derivative motion to a mass or series of masses, using relatively strong magnetic, electric, or nuclear forces. The devices described in the present paper will be shown to generate significant *High-Frequency Gravitational Waves* or **HFGW** without disruptive g loads. The effect will be measurable in the laboratory since it affects or warps the *spacetime* geodesic over very small distances (due to high frequency and short GW wavelength) and thereby will produce HFGW detectable by utilizing detectors described in this Conference. If the energizable elements are uncharged, then there may be little or no attendant electromagnetic (EM) radiation.

In order to illustrate the concept, a circular rim, which does not rotate, but ratchets or jerks, is described. This rim is then evolved into a practical gravitational-wave generator. The system of masses described in this paper (and subject to jerks) can be small or mini magnets micro- or nano-devices, molecules, sub-microscopic particles, electrons (e.g., in a superconductor; about 10^{20} per cubic centimeter), etc. The misconception that the laboratory generation of GW is not feasible is fed by the example of a spinning rod given in most introductory textbooks. Such a rod utilizes the change in the centrifugal force vector to generate GW and is torn apart well before any significant GW is generated. The devices discussed herein are completely different and utilize electromagnetic forces and reciprocating, not rotational, motion in order to generate GW.

2. JERK FORMULATION OF THE QUADRUPOLE EQUATION

There is no new Physics here, simply a different approach or formulation of the conventional equations utilized to estimate GW power in order to render engineering applications more apparent. I will employ the standard quadrupole equation, which was originally formulated by Einstein in 1918, to compute the HFGW power. I will formulate that basic quadrupole approximation in terms of a change in force, ?f, over a short time interval, ?t, which is defined as a "***jerk***." The derivation of this basic jerk equation will be accomplished by two separate analysis paths: one starting with the third derivative of the moment of inertia formulation of the quadrupole equation and the other starting with the spinning rod (or binary orbit) formulation of the quadrupole equation. The resulting jerk equation will be numerically checked against the known result for the binary star pair PSR 1913 + 16.

2.1 *Derivation from Third Time Derivative of the Moment of Inertia*

As is well known and noted specifically in a letter (dated January 19, 2000) to me from Dr. Geoff Burdge, Deputy Director for Technology and Systems of the *National Security Agency*: "Because of symmetry, the quadrupole moment can be related to a principal moment of inertia, I, of a three-dimensional tensor of the system and … can be approximated by

$$-dE/dt \approx -G/5c^5 (d^3I/dt^3)^2 = - 5.5x10^{-54} (d^3I/dt^3)^2." \quad (1A)$$

or from Eq. (110.16), p. 355 of Landau and Lifshitz [5] :

$$P = | - dE/dt | = (G/45c^5)(d^3D_{\alpha\beta}/dt^3)^2 \text{ [watts]} \quad (1B)$$

or

$$P = 1.76x10^{-52} (d^3I/dt^3)^2 \text{ [watts]}. \quad (1C)$$

This is Einstein's ***quadrupole equation***.

In Eq. (1A), k in Burdge's notation is G (not, however, the Einstein tensor) and the units in Eq. (1C) are in the MKS system [watts] not the cgs. In order to introduce the jerk concept let us consider the hypothetical example of a rim that, like the ratchet wheel of a mechanical watch, need not be uniformly rotating or, in fact, not rotating at all. In this case, for a collection of masses, which are small permanent magnets, along the rim,

$$I = \delta m\, r^2 \quad [kg\text{-}m^2], \quad (2)$$

where

δm = mass of an individual magnetic sites around the rim [kg], and

r = the distance from a pivot out to any single δm on the rim [m] (or more exactly, the *radius of gyration* of the rim). Thus

$$d^3I/dt^3 = \delta m d^3 (r^2)/dt^3 = 2r\delta m(d^3r/dt^3) + \ldots \quad (3A)$$

Approximately, by delta differentiation,

$$2r\{dm(d^3r/dt^3)\} \sim 2r\{dm?(d^2/dt^2)/?t\} \quad (3B)$$

and, by noting that by Newton's second law of motion,

$$f_r = dm(d^2r/dt^2), \quad (4A)$$

we have, again by delta differentiation,

$$dm?(d^2r/dt^2) = ?f_r \quad (4B)$$

where f_r = radial force on δm and $?f_r$ is the rapid increase in f_r over time ?t (**the jerk**). The third derivative of I is, therefore, approximated by

$$d^3I/dt^3 \cong 2r\,\Delta f_r/\Delta t, \quad (5)$$

in which Δf_r is the nearly instantaneous **increase** in the force on magnetic (or other energizable element) sites, δm, caused by the magnetic field of current-carrying coils (or other energizing elements) when they are turned on and off or pulsed by transistors or ultra-fast switches resulting in a jerk.

Let us now visualize a stack of such rims; each one composed of a circle of small permanent magnets that are surrounded by a close-by ring of coils (please see FIG. (3B)). In this regard, the coils adjacent to the periphery of each rim are sequenced (at the local GW speed, say the speed of light) along the stack of rims from one rim to the next in order to generate or build up the train of coherent HFGW as they move through the stack of rims (energizable magnetic sites). In order **not** to build up acceleration the jerks are reciprocating; but (arguably) due to the square in the kernel of the quadrupole equation, the GW radiates in both directions along the axis of the circular rims (through their centers) no matter which direction the peripheral magnetic masses are jerked. In summary, by substituting Eq. (5) into Eq. (1C),

$$P = 1.76x10^{-52} (2r\Delta f_r/\Delta t)^2 \text{ [watts]}, \quad (6)$$

which is the **jerk formulation of the quadrupole equation.**

2.2 *Derivation from a Spinning Rod*

An alternative derivation of Eq. (6) is as follows: From Eq. (1), p. 90 of Joseph Weber [1] one has for Einstein's formulation of the gravitational-wave (GW) radiated power of a rod spinning about an axis through its midpoint having a moment of inertia, I [kg-m^2], and an angular rate, ω [radians/s] (also please see, for example, pp. 979 and 980 of Misner, Thorne, and Wheeler [2], in which I in the kernel of the quadrupole equation also takes on its classical-physics meaning of an ordinary moment of inertia):

$$P = 32GI^2\,\omega^6/5c^5 = G(I\omega^3)^2/5(c/2)^5 \text{ [watts]} \quad (7)$$

or , with $I = r^2m$ (r being the radius of gyration of the rod)

$$P = 1.76x10^{-52}(I\omega^3)^2 = 1.76x10^{-52}(r\{rm\omega^2\}\omega)^2 \quad \text{[watts]} \quad (8)$$

where $\{rm\omega^2\}$ can be associated with the magnitude of the rod's centrifugal-force vector, $\mathbf{f_{cf}}$. Equation (8) is the *same equation as that given for two bodies on a circular orbit* on p. 356 of Landau and Lifshitz [5] ($I=\mu r^2$ in their notation) where $\omega = n$, the orbital mean motion [radians/s].

The $\mathbf{f_{cf}}$ vector reverses every half period at **twice the angular rate of the rod** (and a $\mathbf{f_{cf}}$ component's magnitude completes one complete period in half the rod's period). Thus the GW frequency is 2(ω/2p), where ? is in [radians/s]. The change in the centrifugal-force vector itself (which I call a "jerk" when divided by a time interval) is a differential vector at right angles to the $\mathbf{f_{cf}}$ vector and directed tangentially along the arc that the dumbbell or rod moves through. The differential change in, for example, the x-component of the change in centrifugal force, $?f_{cfx}$, is f_{cfx} ?? and the change in the y-component is $?f_{cfy}$ is f_{cfy} ?? , where ? is the central angle of the rotating rod in radians. By delta differentiation of $f_{cf}^2 = f_{cfx}^2 + f_{cfy}^2$,

$$f_{cf}?\ f_{cf} = f_{cfx}?\ f_{cfx} + f_{cfy}?\ f_{cfy} \qquad (9A)$$

and when one associates the components $?f_{cfx,y}$ with $f_{cfx,y}$?? and, after dividing by ?t` (t` being spinning-rod time), and noting that ??/?t` =? ,

$$f_{cf}?\ f_{cf} / ?t` = (f_{cfx}^2 + f_{cfy}^2)?\ . \qquad (9B)$$

Thus $?\ f_{cf} / ?t` = f_{cf}?$ ; but ?t` = ½?t since the period of the GW is half the period of the rod, so that

$$2?\ f_{cf} / ?t = f_{cf}? , \qquad (9C)$$

but $f_{cf} = \{rm?^2\}$ so

$$2?f_{cf}/?t = \{rm?^2\}? \qquad (9D)$$

and substituting Eq. (9D) into Eq. (8) yields

$$P = 1.76x10^{-52}(2r\Delta f_{cf}/\Delta t)^2, \qquad (10)$$

where $(2r\Delta f_{cf}/\Delta t)^2$ is the kernel of the quadrupole approximation equation and $\Delta f_{cf}/\Delta t$ is, again, the **jerk.** Equation (10) is identical to Eq. (6), but arrived at by a different analysis path.

Equation (6), like Eqs. (1), (7), (8) and (10), are approximations for GW power and may **only hold accurately** for $r << \lambda_{GW}$ and for speeds of the GW generator components far less than the speed of light, c. Please see, for example, Pais [3], p. 280 and Thorne [4], p. 357. (On the other hand, Leonid P. Grishchuk at this Conference suggested that the requirement that $r << \lambda_{GW}$ may not be a stringent one.)

2.3. *Validation Based on Orbit of PSR 1913+16*

As a numerical validation of Eq. (10), that is a validation of the use of a jerk to estimate gravitational-wave power, let us utilize the approach for computing the gravitational-radiation power of the pulsar **PSR 1913+16** observed by Hulse and Taylor [6] to demonstrate the existence of GW.

2.3.1 *Orbital Elements*

Since the observation of the binary pulsar PSR 1913+16 (identifies right ascension of 19 degrees 13 minutes and declination of 16 degrees North) represents the only *experimental confirmation* of gravitational waves, insight into the jerk approach can be found in the analyses of such a double-star system. Thus please bear with the following rather laborious arithmetic.

The pair of PSR 1913+16 stars will coalesce in $3.5x10^8$ years due to GW radiation and produce a rather continuous GW until that time. It is the pair's coalescing that **exactly** agrees with GW-generation theory (utilizing orbital mechanics) that indirectly confirms the existence of GW. According to J. H. Taylor, Jr. [6], the period of their mutual rotation is 7.75 hours (or $\mathbf{2.79x10^4}$ **[s]**), periastron is 1.1 solar radii (one solar radius is $6.965x10^8$ [m]), and apastron is 4.8 solar radii. It's radius of gyration is essentially the semi-major axis a = (1.1 + 4.8)/2 = 2.95 solar radii = $(2.95)(6.965x10^8)$ = $\mathbf{2.05x10^9}$ **[m]**. Each star exhibits a mass of about 1.4 solar masses (one solar mass is $1.987x10^{30}$ [kg]) so that together their mass is $m = m_1+m_2 = (2)(1.4)(1.987x10^{30})$ = $\mathbf{5.56x10^{30}}$ **[kg]**. According to a perusal of binary-star catalogs by John Mosley of the *Griffith Observatory*, the binary pulsar PSR 1913+16 is at a distance from

our Sun of 23,300 light years (one light year is $9.5x10^{15}$ [m]). If there was complete diffraction, then the reference area over which the GW would spread at the Sun's distance would be a sphere having an area of $(4\pi)(2.33x10^4 \text{ x } 9.5x10^{15})^2 = 6.2x10^{41}$ [m^2].

2.3.2 *Gravitational-Wave Power—the Quadrupole Equation*

In the case of a binary star pair such as PSR 1913+16, the magnitude of the GW power, P, is computed from the quadrupole equation, which for two masses on orbit about one another is given, for example, by an equation on p. 356 of L. D. Landau and E. M. Lifshitz [5] or Peters and Mathews [7]. The time-constant factor in the equation for P is

$$8G^4 m_1^2 m_2^2 \mu/(15c^5). \qquad (11)$$

The *time-variable factor* in P is a function of the true anomaly, v, and orbital eccentricity, e, as given in [5] :

$$(1+e\cos v)^4([1+\{e/12\}\cos v]^2+e^2\sin^2 v)/(a[1-e^2])^5. \qquad (12)$$

In conventional astrodynamic/celestial-mechanics notation (see Samuel Herrick [8]) this factor (i.e., Eq. (12)) is

$$p/r^6+(dr/d\tau)^2/12\mu r^4, \qquad (13)$$

where p is the orbital "parameter" or semilatus rectum (= $a\{1 - e^2\}$) in [AU], r is the radial distance between the two masses [AU], τ is the characteristic time measured in $k_s$days or in units of $5.022x10^6$ [s] for a heliocentric-unit system (utilized by Taylor [6] and others for PSR 1913+16), μ is the sum of the two masses, i. e., $\mu = m_1 + m_2$ [solar masses], and as usual G = $6.67423x10^{-11}$ [m^3/kg-s], and c is the speed of light = $3x10^8$ [m/s]. Note that one AU (astronomical unit) = $1.496x10^{11}$ [m].

The GW power radiated, P, which causes a perturbation in the semi-major axis, a, (and an attendant secular decrease in the orbital period) is obtained by integrating the time-variable factor, Eq. (13), over the orbital period using the mean anomaly, M, as independent variable, which is directly proportional to the time (that is, M = n [t-T], where n is the mean motion {ω in Landau and Lifshitz's [5] notation, p. 357} n= 2p/Period = $2p/2.79x10^4$ = **$2.25x10^{-4}$ [1/s]**, and T is the time of periastron passage).

2.3.3 *Accuracy of the Results*

The value of the average GW power, P, is computed from observations that define the eccentricity (based primarily upon Doppler-shift determination of the range rate at periastron and apastron), semi-major axis, and orbital orientation angles of PSR 1913+16. The error in the computed value of P is related to the observational error leading to the determination of the orbital elements as well as the determination of the masses of the pair of neutron stars. For example, a 0.1 percent change in the measurement of range rate at periastron results in a 0.28 percent change in GW power, P, and a 0.1 percent change in the mass of the stars results in a 0.33 percent change in GW power.

2.3.4 *Centrifugal Force and Acceleration*

The x and y average delta **centrifugal force** component(s), $\Delta f_{cfx,y}$ (which will later be utilized to validate the fundamental jerk equation numerically) are both

$$man^2 = (5.56x10^{30})(2.05x10^9)(2.25x10^{-4})^2 = \mathbf{5.77x10^{32}\ [N]} \qquad (14)$$

divided by *m* yields the average centrifugal acceleration = 103.78 [m/s^2] = 10.6 [g's]. At periastron, r = q = a(1-e) = $(2.05x10^9)(1-0.641)$ = $7.36x10^8$ [m] (with e = 0.641), the centrifugal acceleration is $q(dv/d\tau)^2$ where $dv/d\tau = \sqrt{(\mu p)}/r^2$ (please see Baker [9] , p. 13). In this latter case μ = 2.8 [solar masses], a = 2.95 [solar radii] = $(2.95)(6.965x10^8$ [m/solar radii])/$1.496x10^{11}$ [m/AU] = 0.01373 [AU], p = $a\{1-e^2\}$ = $0.01373\{1- 0.4109\}$ = 0.00809 [AU], and q = r = $7.36x10^8$ [m]/$1.496x10^{11}$[m/AU] = 0.00495 [AU]. After inserting these numbers I have $dv/d\tau$ = $(\sqrt{[2.8x0.00809]}/[0.00495]^2)/5.022x10^6$ [s/k_sday] = $1. 223x10^{-3}$ [radians/s]. Thus the centrifugal **acceleration** at periastron of the star pair is $q(dv/dt)^2$ = $(7.36x10^8$ [m] $)(1.223x10^{-3}$ [radians/s])^2 = $1.101x10^3$ [m/s^2] = **112 [g's]** – apparently *still within the weak-field approximation of Einstein's GW equations.*

2.3.5 *Comparison of Results*

From Eq. (14) I computed that each of the components of force change, $\Delta f_{cfx,y}$ = $5.77x10^{32}$ [N] (multiplied by two since the centrifugal force reverses its direction each half period) and Δt = (1/2)(7.75hrx60minx60sec) = $1.395x10^{4}$ [s] for the half period. Thus using the jerk approach:

$$P = 1.76x10^{-52}\{(2r\Delta f_{cfx}/\Delta t)^2 + (2r\Delta f_{cfy}/\Delta t)^2\} =$$

$$1.76x10^{-52}(2x2.05x10^{9}x5.77x10^{32}/1.395x10^{4})^2 x2$$

$$= \mathbf{10.1x10^{24}} \text{ [watts]} \quad (15)$$

versus the result of **$9.296x10^{24}$** [watts] using Landau and Lifshitz's more exact two-body-orbit formulation given by Eqs. (1.1) and (1.2) of Baker [10] integrated using the mean anomaly not the true anomaly as independent variable. The most stunning closeness of the agreement is, of course, fortuitous since due to orbital eccentricity there is little symmetry among the $\Delta f_{cfx,y}$ components around the orbit and there are small errors inherent in the approximations of Eqs. (3A) and (3B) and, of course Eq. (5) leading to Eq. (10). Nevertheless, since the results for GW power are so close, orbital-mechanic formulation compared to the utilization of a jerk, **the correctness of the jerk formulation is well demonstrated!**

3. COMPUTATION OF HFGW POWER

There are some very sophisticated and exact computer simulations of the generation of gravitational waves (please see, for example, S. F. Ashby, *et al* [11]). The quadrupole approximation utilized herein by me and, for example, by Romero and Dehnen [12] and others at this Conference is probably less exact. On the other hand, the computer simulations are less relevant to the devices involved in the generation and detection of HFGW in the laboratory. These computer simulations describe GW generation by strong-field astrophysical phenomena (e.g., neutron stars, black holes, etc.), coupled spacetime and general relativistic hydrodynamic equations, and are usually restricted to gravitational forces ; not non-gravitational forces involved in laboratory HFGW generation. I will first discuss the meaning of the term quadrupole

3.1 *Meaning of Quadrupole*

The basic physical process for generating a gravitational wave is the third (or higher) time derivative of the motion of a mass, termed a "jerk" or "shake" or ?f/?t, that is, ?f is an increase in force, f, on the mass carried out over a small time interval, ?t. As noted in Baker [10], that physical process produces a gravitational wave with a **power** given by, for example, the quadrupole approximation (as originally derived by Einstein) or it could be determined directly from the special and general relativity equations (using a computer-implemented numerical integration as, for example, discussed in. Ashby, *et al* [11]). As noted in Conference paper HFGW-03-101, the quadrupole is the lowest-order solution to the GW propagation problem That is, the quadrupole itself is **not** the physical process at all, but only one means of establishing the power of the generated gravitational waves – the lowest-order solution.

Other algorithms, often most complicated, can define other GW properties such as direction, polarization, constructive/ destructive interference, etc. This situation is similar to Newton's Laws, which govern the physical process of planetary motion. The effect of that motion can be computed using, for example, the two-body approximation, or it could be determined directly from the equations of motion described by Newton's Laws, using a computer- implemented numerical integration.

The two-body approximation itself is **not** the physical law at all, but only one means of describing the resultant motion – a "lowest-order solution." In the case of a nuclear-reaction-generated gravitational wave, in which a nuclear particle is ejected from a nucleus, it is like a small rocket, or in the case of electrons shaken in a resonance cavity, plasma beam, superconductor, etc., there is a third time derivative of the motion of the nucleus in the first case or electron mass in the second case, or a jerk, which produces gravitational waves whose power can be estimated, for example, by the quadrupole approximation. Thus when I mention a "quadrupole-produced gravitational wave" I'm really implying the fundamental

physical concept of the jerk and not the computational means for establishing the power of the gravitational wave.

3.2 *Harmonic Motion.*

As far as a harmonic motion of a mass or a pair of masses is concerned (harmonic oscillator), gravitational waves are generated. Just as in the case of a pendulum, the usual descriptor of harmonic motion, there exists a third time derivative of the pendulum bob. It is the jerk of that bob that produces the gravitational wave, which can be estimated using a quadrupole approximation or computed exactly by means of a rather complicated solution of the equations of special and general relativity.

4. LABORATORY MICRO- AND NANO-SCALE HFGW GENERATION DEVICES

In this section I will describe an Individual Independently Programmable Coil System or ***IIPCS*** (U. S. Patent No. 6,160,336), miniature integrated circuits, which provide for the emulation of a device that is much more extensive than the individual energizable elements (e.g., the small permanent magnets), and I will summarize what all of the HFGW-generation devices accomplish.

4.1 *Individual Independently Programmable Coil System*

I will now discuss HFGW generation devices that utilize, for example, microchip and nanotechnology in order to generate HFGW in the laboratory. For the very large number of ultra-small, sub-millimeter coil elements utilized in some of the devices discussed, a miniaturized integrated circuit can be utilized (please see, for example, the coil turn of *Al* utilized by Y. Acremann, *et al* [13]). They will be embedded in or imprinted on a silicon chip, organic material, or in connection with polymer-based or superconductor devices. They will consist of multiple layers with appropriate sequencing time delays to ensure near simultaneity of the magnetic fields interaction as the direct-current train of approximately one-picosecond or shorter pulses simultaneously traverse each coil set on the chip levels. The timing sequence could be integrated in the chip with the ultra-fast switches or transistors or through other semi-conductor devices. The myriad of these small coils in a three-dimensional array (please see FIG. (1)) act on the field of a small magnet to produce the jerk.

Since the jerk is generated by an electromagnetic process, there could be significant EM radiation generated that could reduce the efficiency of the device. It should, however, be emphasized that **it is <u>not</u> the magnetic field that generates the HFGW**, but rather the mass of the magnets (or other energizable elements) that are jerked that generates the GW. The magnetic material exhibits magnetic sites (perhaps on a molecular level or ferromagnetic atoms) that, of course, include electrons; but in this case (as opposed to a superconductor HFGW generator) it is not the electron mass being jerked that produce the GW, but rather the actual magnet's mass.

4.2 *Miniaturized Integrated Circuits*

A preferred design (U. S. Patent No. 6,417,597) utilizes conventional computer chips or wafers of a computer logic system, containing IIPCS circuit elements (U. S. Patent No. 6,160,336). These circuits are about 18 micrometers or less apart and include a synchronizing clock, input/output ports, and sub-millimeter coils on 50 to 100 micrometer centers. The chips are about 6 mm to 9 mm square and are obtained from silicon wafers. These chips are sewn into a circuit-board roll with an approximately 25-micrometer-diameter gold thread. Several layers of this roll (for example, 25) are connected in a fixed location or band adjacent to the moving or non-moving (jerking or non-jerking) spindle's rim and they form the IIPCS in the spindle rim's magnetic field. (The rolls of chips just mentioned are routinely fabricated by French-owned *Oberthur Card Systems* {a plant exists at Rancho Dominguez, California}, French-based *Gemplus*, *Schlumberger* {Paris and New York}, and California-based *Frost & Sullivan*.)

4.2.1 *Coil Sets*

In the proposed miniaturized integrated circuit devices, as exhibited in FIG. (1), there will be a very large number of ultra-small, sub-

millimeter or microscopic coil sets or elements, 56, embedded in or imprinted upon a silicon chip, 57, in multiple layers. Ultra-fast micro-switches or transistors of the IIPCS, 58, will launch a long a series of current pulses, 59, of approximately nanosecond to picosecond or less duration moving at the electron's mobility speed (approximately light speed, c) that will be timed to reach the individual coil sets or elements almost simultaneously (with the same rise time as discussed in Y. Acremann *et al* [13]). These pulses can travel along several individual conductors, as in FIG. (1), or along one single conductor per line, as in FIG. (2), and thereby interact with the magnetic field of a nearby magnet on the rim, 60, in concert.

This interaction will result in a third-time-derivative motion or jerk of the uncharged magnetic masses on the rim to generate a train of gravitational waves. **The effect is exactly the same as a rotating or ratcheting rim with the change in centrifugal force (jerks) replaced by the reciprocating jerks of the magnets attached to the rim.** The ultra-fast switches are preferably semiconductor-based, such as the semiconductor optical amplifier (SOA), a semiconductor nonlinear interferometer such as a nonlinear *Sagnac* interferometer on a phosphide semiconductor chip, etc. (please see, for example, D. Cotter, *et al* [14], pp. 1523-1528).

4.2.2 *Pulse Duration*

The pulse duration will be such as to completely energize any given coil set as it passes through it in order to produce a magnetic field interaction. As has been emphasized, the interaction will result in a third-time-derivative circumferential motion or jerk of a cylindrical stack of rims, 63, shown in FIG. (3A) (figure 2 of US Patent 6,417,597) and generate a long GW train of successive GW pulses having axis, 29. This stack or barrel is surrounded by and immediately adjacent to a sheath of IIPCS-controlled coil sets, 64. The cross–section of the barrel or an individual rim of the stack is shown in FIG. (3B) (figure 8A of US Patent 6,160,336). The coils (myriads of them represented by a single coil) 26, interact with the tiny rim magnets, 24; produce jerks along axes, 27, which emulate a ratcheting rim, 15. In the case of a design with the current-pulse train on a single conductor interconnecting a line of coil sets, FIG. (2), there will be a build up of impulses to full value as the current-impulse train progresses down the line of coil sets. Use of a single conductor wire for each line of coil sets reduces the resistive power loss. In each line of coils set in series 61 there will be time delays, 62, between coil sets to ensure simultaneity of the current pulses reaching any given coil set. The myriad of miniature coil sets (incased in layers of chips) will energize (jerk) each tiny rim magnet.

4.2.3 *Parallel-Conductor Stacks*

In FIG. (4), ultra-fast switches or transistors of the IIPCS, 58, will launch a long series of direct-current pulses acting in either direction, 59, of approximately nanosecond or picosecond or less duration moving at the electron's mobility speed along individual conductors or single interconnecting computer wires in order to produce current pulses, 59, acting in concert to generate modulated jerks and resulting HFGW (GHz to THz and higher frequencies) with axis, 29 (or perpendicular to that axis). The current pulses will be timed to reach parallel-plate conductors, 66, which may have different masses or may have ballast, 67, attached and/or carry different current, and/or have different modulus of elasticity and/or are constructed differently in their mountings for the purpose of exhibiting asymmetrical mass displacements, jerks or "hammer blows." The asymmetry is required in order to avoid the null situation to be discussed in Section 9. (Such a concept of utilizing the force between parallel current-carrying conductors is similar to the nanowire or nanoplate devices to be described in Section 8.)

4.3 *Emulation of a Much More Extensive Body*

As a GW front passes by the energizable, e. g., in the case of parallel-plate, elements (schematically shown in FIGS. (5A) and (5B) as 80, 84, 86, and 88) or individual members of a stack of (jerking) rims as shown in FIG. (3A), they are energized in sequence thereby increasing the wave's amplitude. In FIG. (5B) such an effect is schematically illustrated as GW 83 ,85, 87, and 89 build up to accumulate the GW, 82, wave front shown also in FIG. (5A).

Thus a linear device having a much longer effective length (or radius of gyration), r, or a cylindrical stack of rims having a much larger mass than any single rim is emulated. Again, this is all subject to experimental verification. It is to be emphasized that **any unwanted EM radiation can be screened out.**

In FIG. (6), ultra-fast switches or transistors of the IIPCS, 58, will launch a long series of current pulses, of approximately picosecond duration along individual conductors or single, interconnecting conductor wires that will be timed to reach individual, sub-millimeter micro- or nano-electromechanical elements, or piezoelectric crystals, etc. 56. This will be in sequence to reinforce and cause a *build up* of the amplitude of a **coherent** GW beam (as in FIG. (5B)) having axis, 29. The ensemble of electromechanical elements (including other kinds of energizable elements such as nanomachines) will also be embedded in or imprinted on a silicon chip in multiple layers. FIGS. (7A) and (7B) (figures 7A and 7B of US Patent 6,160,336) exhibit the ultra-fast, micro-switches (1.1d,u to 1.4 d,u and 1.5 *l*,r to 1.8 *l*,r) set so no current flows through the coils and ultra-fast, micro-switches (2.1 d,u to 2.4 d,u and 2.5 *l*,r to 2.8 *l*,r) set so that the current flows through the coils right to left. Other switch setting can reverse this current direction.

4.4 *Summary*

The problem, which all of the devices discussed in this paper solve, is to cause a system of masses, which could be mini-magnets, micro-devices (e.g., small plates), nano-devices (e.g., nanowires), individual molecules, submicroscopic particles, or individual electrons (as in a superconductor) to move in concert with a jerk in order to *build up* (generate) HFGW with either planar or cylindrical wave propagation. Such jerking masses produce a very long sequence of HFGW pulses having significant average power without generating incapacitating heat, causing disruptive g loads, or producing overpowering EM radiation.

As I have emphasized, the problem is solved in several alternative ways by utilizing an array of energizable elements (e. g., rim-magnets, coils, parallel plates, piezoelectric crystals, dielectrics, capacitors, nanomachines, high-temperature superconductors, electrons, nuclear particles, laser beams, etc.) to be activated by energizing elements (e.g., coils, submicroscopic particles, laser beams, etc.) under comp uter-logic-system control.

As already noted, these energizable elements are activated or energized in the correct sequence with correct timing by the IIPCS computer (computer-controlled logic system) to accumulate a GW (moving at local GW speed in the energizable mass, which may or may not be near to the vacuum light speed) as the GW front moves in the mass or collection of masses. Essentially, the IIPCS causes the entire mass or collection of masses, or rims, or molecules to jerk effectively in unison or in step with the GW wave front and generate coherent HFGW. That is, the jerk will progress in step with the GW front and build the GW amplitude up – somewhat similar to a cyclotron pulsing a charged particle as it circles around in its magnetic field, or, possibly, like a traveling-wave amplifier and similar to the coherent GW generation suggested by Romero and Dehnen [12]. Energizable elements (that jerk when energized) are energized in sequence as the GW front passes. As has been seen, these elements taken together emulate a much larger, more extensive mass. **That is, the entire mass "appears" to the GW (as it passes) to be a single larger mass (e.g., a solid massive cylindrical flywheel) being jerked cohesively.** Experiments suggested at this Conference would not only shed light on such HFGW characteristics, but also, as suggested by Y. Acremann, *et al* [13] in their discussion of the processional motion of the magnetization vector "… forms the basis for realistic models of magnetization dynamics in a largely unexplored but technologically increasingly relevant (picosecond) time scale."

5. MAGNETIC FIELD BUILD UP AND HEAT LOSS

I will commence the analysis using the theoretical example of the ratcheting or jerking rim and then evolve the device into both a stack of rims and into a linear form. It should be recognized again that a rotating rim could generate GW, but in order to generate significant and continuous GW its rotational rate would need to be so large that the rim would be torn apart! Thus a rotating rim has been replaced by a ratcheting or jerking rim and that rim will be replaced by a stack of rims composed of

individual jerking rim elements and finally by a linear motor.

5.1 *Magnetic Field Build Up*

Although of little concern in most applications, the length of time to "build-up" the magnetic field of the coils is important here as it is in the experimental work of Y. Acremann [13]. The electrons must complete sufficient coil turns (moving at the current pulse speed – about $2.3x10^8$ [m/s] or 77% of light speed) in approximately a picosecond to "launch" most of the magnetic field that produces the impulsive force (like a "hammer blow") or jerk when it interacts with the static magnetic field of permanent or electromagnets as effectively "carried around" by the ratcheting rim. Thus, they must be very tightly wound with each coil "set" having a total length of less than 0.3 mm (0.0003[m] or 300 micrometers). The coils are tightly wound in order to react rapidly enough as the current pulses move through the coils at electron migration (for the ultra fast switch semiconductors) or current pulse speed, about 0.77c according to some experiments by Spring [19] who found current speeds from 0.66c to 0.9c. Note the electrons themselves, like water molecules in an ocean wave, do not move with the wave and have drift speeds on the order of only one m/s. If each of the ultra-small, sub-millimeter coil sets consist of two coils or turns, as exhibited in FIGS. (7A) and (7B), then their diameters are on the order of $d = 0.3/2\ \pi = 0.05$[mm] = 50 [μm] or less. (Note that the single-turn coil of *Al,* utilized by Y. Acremann, *et al* [13] was about 6 [μm] in diameter.) The coil wire could be made of gold having about a 0.015-mm or 15-micrometer diameter. The resistance for such wire at room temperature is about 135 [ohms/m] -- high-temperature superconducting (HTSC) material would be useful here. As will be seen from Section 5.4.2, in the spin-up jerk mode the IIPCS will need to build up **0.26 [Tesla]** flux density, at the appropriate polarity and interval, e. g., every 0.044 [m] for magnets of that rim spacing (the requirements for the spin-down jerk mode are essentially the same, but reversed). The magnetic flux density, B, is given by

$$B = \mu_o n i/l \text{ [Tesla]} \qquad (16)$$

where $\mu_o = 4\pi x10^{-7}$ (permeability of free space), n is the number of coil turns, i is the current through the coils [amps], and l is the length of the coil conductors [m]. The double coil sets will be placed on 50 to 100 [micrometer] centers, so that there will be about $2x100x100 = 2x10^4$ coil turns on each square-centimeter level of the stack of 25 coil levels or layers. With $l = 0.044$ [m], $n = 25x2x10^4 = 5x10^5$, $i = 9.1x10^3/5x10^5 = 0.018$ [amps] or 18 milliamperes, , $ni = 9.1x10^3$ [amp turns], so that Eq. (6) yields **B = 0.26 [Tesla].**

5.2 *Heat Loss*

The total length of 15-micrometer-diameter gold wire across any given layer or level is 100(rows) x100 (coil and jumper/time-delays)x(600 micrometers) = 6 [m]. For the 25 layers or levels there will 150 [m] of wire with a resistance of 150[m] x135[ohms/m] = $2.025x10^4$ [ohms]. Since on average every other pulse interval across a conductor wire will carry no current (resulting of course, in a lower average GW power), the heat loss per centimeter of chip stack or semi-conductor layers is

$$(1/2)i^2R = \mathbf{3.28\ [watts]}. \qquad (17)$$

This heat loss can be reduced by 32% by using 25-micrometer-diameter wires for the time-delay jumpers, but high-temperature superconductors (HTSCs) for this purpose are contemplated. In addition there may be some energy loss or resistance occasioned by EM radiation generated during the GW-generation process **– reduced or eliminated since the jerked masses are uncharged**. Such an EM energy loss can be reduced by the design of the energizing coil elements and controlling the direction of current pulses by the IIPCS. Concerns of the influence of magneto resistance (MR) of both the conductors and the semiconductor circuits and the dynamics of the impulsive magnetic field buildup should be addressed during experiments as would be the aforementioned EM radiation, which could significantly reduce the efficiency of the HFGW generator. Note that alternating currents are **not** utilized (only direct-current, positive pulses) in order not to drive the electrons to the conductor's skin and thereby increase resistance. This is probably not a problem if superconductors are utilized and their utilization is contemplated in most of the HFGW devices described in this paper.

5.3 *Direction of the Generated HFGW*

The HFGW is expected to progress both ways along the axis of the rims (or cylindrical stack of independent rims) since there is a square associated with the kernel of the quadrupole equation—so there is no preferred direction along the axis of the cylinder. Such a concept will be subject to experimental verification and the propagation direction may be dependent upon whether or not there is circular polarization or cross "+" polarization or a combination. Between the rims HTSC lenses can be inserted to concentrate the HFGW along the axis of the rings. As has been emphasized a large number of ultra-fast switches, preferably semiconductor based (exhibited in FIGS. (7A) and 7(B)), would be activated simultaneously along the progressing gravitational wave front in one of the two directions by the coil-control computer, with communication lines of nearly equal length to all switches. The coil-control computer logic system of the IIPCS could also activate coil sets inside and outside the rim, not just outside as shown in FIGS.(3A) and (3B) and concentric "layers" of rims and coils could be utilized.

The coils, which are closely adjacent to magnets or magnetic sites in the rim or rims, are to be sequentially activated in order to generate coherent HFGW in one direction (non-coherent HFGW will propagate in the other direction). Depending upon the proximity of the coils and the duration of the current pulses, there may be currents induced in one coil juxtaposed to another. *Any induced currents will produce deterministic reverberations for which the computer logic system of the IIPCS that can be programmed to account for*. In any event, the reverberations would subside as the current-produced magnetic pulses either collapse or clear the ensemble of chip layers at light speed.

5.4 *Linear-Motor, Linear-Jerk GW*

The foregoing discussion of a ratcheting rim is included first in order to bridge the gap between the rather conventional celestial GW mechanisms involving circular motion and the terrestrial laboratory cylindrical stack of "rotating" or jerking rims. Another laboratory device is similar to that proposed by Romero and Dehnen [12] and involves a linear motor device involving linear motion. (Yet another variant of the ratcheting rim is a ratcheting rod shown in FIG. (8), which could be oriented at various angles.) In the following subsections I discuss the linear-motor concept, estimate the coil-magnet force (utilized for all the devices discussed in this paper), and by means of a numerical example, determine the material acceleration.

5.4.1 *Concept*

The linear-motor design of the HFGW-generation device, sometimes referred to as a *linear induction motor* or LIM, is visualized to involve a single sector of the ratcheting rim with the impulsive forces, ?f, being longitudinal and tangential to the ring of energizable elements, $\Delta f_{l,}$ rather than radial. If the rim magnets and adjacent coils were peeled off from the rim and laid out flat, then the result would be a linear motor. Please see FIGS. (9A) – (9D) for a schematic of the progression of such a "peeling." In a very hypothetical case , a 2000-meter-radius exemplar device once peeled would be 2pr = 2px2000 = 6283 [m] in length and, since for this linear mass distribution $I = (1/3)mr^2$ and $d^3I/dt^3 \cong (2/3)r\,\Delta f_l\,/\Delta t$, the effective radius or radius of gyration is 6283/3 = 2094 [m] ~ 2000 [m] (a measure of the mass distribution) and I assume that it is a tube 3 [m] in diameter. One-centimeter-wide chip rolls would be placed longitudinally along the sides of central, cylindrical, permanent- (or electro-) magnetic tube, core, piston, or barrel, 63, consisting of an array of magnetic energizable element sites, 57, as shown in FIG. (3A) if one replaces the stack of jerking rims by a solid cylinder of magnets surrounded by a sheath of coils. The thin (approximately one cm thick) band of Alnico 5 permanent magnets could be replaced by far stronger electromagnets that face outward as in FIG. (3B).

5.4.2 *Estimate of Force*

In general, permanent magnets exhibit irregular magnetic fields and associated forces. As a rule of thumb a band of juxtaposed 1.75-inch-long (0.044[m]) magnets will lift a weight in excess of 30 pounds per 1.75 inches or {30x12"/1.75" = 206 pounds)/(2.2 pounds per kilogram)} {3.28 feet per meter} = 307 [kg/m]

x{ 9.8 Newtons per kilogram weight}or produce about **3000 [N/m]** of longitudinal force, f, per meter. Each 1.75-inch permanent magnet has a flux density, B, of about 2,600 gauss or **0.26 [Tesla]** developed every 4.4 [cm]. This matches the magnetic flux density of the juxtaposed coils from Eq. (16). Thus since each meter-long, square-centimeter segment of the roll would produce about **3000 [N/m]** of longitudinal force, f_l, per meter and all together they form a sheath of sub-millimeter coils (energizing elements) surrounding this central magnetic core, tube, piston, or barrel of a stack of rings of juxtaposed magnets. I will be extrapolating these numbers to micro- or nano-magnets so it is important to establish a specific magnetic force per unit volume, which of course would be much larger if electromagnets replace the permanent magnets and HTSCs were introduced. The impulsive force per unit volume for the meter long square-centimeter cross-section magnet and closely adjacent coil combination, is $?f_l/?V = 3000[N/m]/(0.01[m])^2 = \mathbf{3x10^7\ [N/m^3]}$. As mentioned already, rings of such uncharged elements, whose planes are parallel to the passing GW crest, will be energized or jerked tangentially to the rings as the GW crest passes and add to its amplitude so as to generate coherent HFGW. Note that in this case the motion of the magnetic mass is asymmetrical (either "in" or "out") so that there is a quadrupole moment without GW cancellation. (In this regard, please see Section 9.) Pinto and Rotoli [15] (p. 567) indicate that "... the quadrupole formula is only valid provided a suitable surface integral (vanishes), which is the case for a series of point sources" such as the energizable elements of the subject device and that of Romero and Dehnen [12].

5.4.3 *Numerical Example*

As a numerical example for the linear-motor design, there would be about one roll or 25-layer strip of chips spaced around and adjacent to the cylindrical barrel of the linear motor (64, shown in FIG. (3A)) in a longitudinal direction (parallel to the barrel axis) every two centimeters forming the sheath. There would be πx3[m]x100[cm/m]/2[cm] = 471 strips around the barrel's circumference, each one having a length of 2px1000 = 6283 [m] so that

$$\Delta f_l=(471)(6283[m])(3000[N/m]) = \mathbf{9x10^9\ [N]} \quad (18)$$

and with κ_{mr3dot} = 32 (the theoretical quadrupole-approximation value to be established experimentally since r may not be less than λ_{GW}),

$$P= 1.76x10^{-52}\ (?\ x6283x9x10^9/10^{-12})^2 = 0.25\ [watts]. \quad (19)$$

Thus, with the reference area being the two 3 [m] diameter "barrels" or "pipes" or "tubes" or cylindrical ends (GW propagating in both directions so the area is doubled) with a thickness of one centimeter, area = 2 (3p) (0.01) = 0.19 [m^2], the generated HFGW flux is about 0.25/0.19 = 1.3 [watts/m^2] near the hypothetical device. The average HFGW flux or signal would be about 1 [watt/m^2]. As a point of reference I again compare our terrestrial HFGW generator to celestial Low-Frequency Gravitational Wave or LFGW generation. Thus, for the sake of argument (although admittedly, it is like comparing "apples to oranges") the 1 [watt/m^2] is compared to $\mathbf{4x10^{-16}\ [watts/m^2]}$ maximum signal from a 500 mega parsec [Mpc] distant, 1000 black-hole (BH) radius semimajor-axis binary black hole (BBH) osculating orbit and $\mathbf{5x10^{-5}\ [watts/m^2]}$ from a 6-BH radius osculating orbit ([10], pp. 19 and 27) – or over **ten-thousand times stronger** than the LFGW signal from a 6 BH-radii BBH osculating orbit just before merger**!**

For cylindrical GW, in case the barrel magnets participated in harmonic oscillation (each end's uncharged magnetic sites moved in and out harmonically relative to the other), the reference area would be (6283)(3π) = $6x10^4$ [m^2] and the GW flux would be $0.25/6x10^4 = 4x10^{-6}$ [watts/m^2]. Although the jerked masses are uncharged, the high-frequency electromagnetic fields may generate significant EM radiation that will be studied in any experimental effort.

5.4.4 *Material Accelerations*

The acceleration, *a*, caused by the jerk is obtained by multiplying the activation time, dt, by the time rate of change of acceleration. The time rate of change of acceleration is $da/dt \sim ?a/?t$ and by Newton's second law $?f = m?a$ so $?a = ?f/m$ and

$$da/dt \sim \Delta f/m\Delta t. \quad (20)$$

From Eq. (18) ?f = $9x10^9$ [N], m (mass) per

meter for the cylinder of magnets is equal to (3.8[kg/m] of magnet strips)(471 strips per meter of cylinder)(6283 [m] cylinder length) = 1.12 x 10^7 [kg], and ?t = 10^{-12} [s]. Thus

$$da/dt = 9x10^9/(1.12x10^7)(10^{-12}) = 8x10^{14}\ [m/s^3] \quad (21)$$

Let us suppose that the total activation time for jerking a magnet in one direction, dt, is three picosecond pulse lengths or $3x10^{-12}$ [s]. Therefore, a = (d*a*/dt)dt = $(8x10^{14})(3x10^{-12})$ ~ $2.4x10^3$ [m/s^2] or 245 g's and far less than, say, the acceleration experienced by a bullet in the barrel of a gun. Thus the magnet would not disintegrate. The acceleration is also probably within the weak-field limit of Einstein's equations since for PSR 1913 + 16 the acceleration at periastron is over 100 g's (Section 2.3.4) and probably much grater acceleration is encountered for a spinning neutron star for which the Einstein equations presumably hold.

In the *extreme case* of 100 picoseconds of continuous jerk (in the same direction), dt = 10^{-10} [s], the speed would build up to

$$ds/dt = (da/dt)\delta t^2/2 = (8x10^{16})x10^{-20}/2 = 0.0004\ [m/s] \quad (22)$$

and the displacement of the magnetic mass (composed of many magnetic surface sites, 57, of the linear motor, piston, or barrel shown in FIG. (3A)) is

$$s = (da/dt)\delta t^3/6 = (8x10^{16}/6)x10^{-30} = 1.3x10^{-16}\ [m]. \quad (23)$$

Again there could be considerable "motion" of the magnetic mass, but even in the most unlikely case of an extremely long series of jerks in the same direction (100 pulses), it goes a very small distance before the IIPCS reverses the built-up acceleration, speed, and displacement and *the stresses in the material of the device would be minimal.*

In general, for permanent-magnet and coil combinations of all of the devices discussed in this paper, which exhibit an N-pulse long activation time, the acceleration is

$$da/dt = \{(?f \text{ per meter})/(\text{mass per meter})\}N \quad (24)$$

with dt = N ?t. We have calculated that ?f per meter = 3000 [N/m}, that mass per meter = 3.8 [kg/m], and with N = 3 (the ?t's cancel out), d*a*/dt = (3000/3.8)x3 = 2368 [m/s^2] or 245 g's as before. **Since the ?t (and, therefore, the frequency) cancel out, the material (magnet) acceleration is only dependent upon the ratio of force to mass of the magnets and the number of current pulses, N, in the total time that the magnet is activated (jerked) in one direction before the IIPCS reverses the jerk.**

6. HIGH-INTENSITY HFGW GENERATOR

A high-intensity HFGW-generation would necessarily involve a much shorter pulse duration, e.g., ten attoseconds or 10^{-17} [s] (100 PHz or QHz frequencies). In this regard, Raymond Lewis told us during this Conference (5/7/03) that "... quantum jump might greatly reduce ?t to, say, 10^{-18} [s]...." I will configure the generator as before as a barrel composed of a stack of individual and separate rims whose edges are again covered with a juxtaposed array of ultra-small micromagnets (energizable elements) surrounded by a sheath of ultra-small microcoils (energizing elements). The device is again represented by the schematic drawn in FIG. (3A) whose cross section or individual separate rims are shown in FIG. (3B) except that there are multiple concentric rims or rings around the cylinder axis. If the HFGW spreads out from one rim to the next, then a thin HTSC lens (please see paper HFGW-03-120) can be inserted between the rims in order to concentrate the HFGW down the axis of the stack as its intensity is built up. At the end of the stack of rims or rings there is a final HTSC lens that concentrates the HFGW on a focal plane.

The power of the device is given by a variant of Eq. (6). Each rim or ring has magnets on its periphery that are energized by an adjacent shell of coils (using the IIPCS for timing). As already noted there are many concentric rims or rings at, say, two-centimeter intervals along the axis of the rim or ring stacks to allow for a one-centimeter thick (maximum thickness) HTSC lens to be sandwiched in between the one-centimeter thick, d*l*, ring sets. I will set the overall length of the generator to be *l* = 500 [m] so that there are n = (*l*/d*l*)/2 = 25,000 rims (or sets of concentric rings) along the axis of the cylinder (FIG. (3A)). I utilize the force per meter that has been calculated for the (rather weak) permanent magnets of F = 3000 [N/m] and set the radius of the outermost rim to be r_0 = 10 [m],

so that the rims or rings in any cross section number $n = (r_0/dl)/2 = 500$. The power of the HFGW generator is given by (half goes the opposite way)

$$P = \tfrac{1}{2}x1.76x10^{-52}\sum_{k=1,\,step2}^{k=500}\{2r_k\,Fn n(2p r_k)/?t\}^2\ [watts] \qquad (25)$$

where $r_k = 2k\,dl$ [m] and $?t = 10^{-17}$ [s]. The numerical result is about **380 kW**.

At the focus of the last or end lens, the size of the diffraction pattern on the focal plane defines the maximum HFGW flux. The diameter of the diffraction pattern is $1.22?_{GW}$, where $?_{GW} = c\,?t$ and for the ten-attosecond (100Qz), $?_{GW} = 3x10^{-9}$ [m]. Ideally all of the HFGW power is concentrated in this diffraction pattern, which has an area of $p(1.22x3x10^{-9}/2)^2 = 1.05x10^{-17}$ [m^2]. So the flux is **$3.7x10^{22}$ [watts/m^2]**. Such a hypothetical and ideal flux density compares favorably with the ultra-intense laser pulses of 10^{23} [watts/m^2], which produce proton energies of "... up to 58 MeV..." [18].

7. MINITURIZED HFGW GENERATOR

For the purpose of having a specific numerical example of a miniaturized HFGW generator, suppose that the dimensions of the transmitter or miniaturized device involve an energizable-element stack of tiny rims (e.g., rings or circles of microscopic magnets) that is 6 [mm] thick surrounding a 3 [mm] radius energizing-element core (e.g., microscopic coils). The diameter of the device is 2(3+6) = 18 [mm]. The radius of gyration would be 6 [mm]. Let us also suppose that the device is 18 [mm] in length. At the receiver, which I assume to be 7 [km] away, I will introduce a d = 18 [mm] diameter superconducting lens to gather and focus the HFGW in order to concentrate or amplify the signal at the receiver. I will consider that $?f_l/?V$ can be increased ten fold by increased magnetic efficiencies due to the use of superconducting electromagnets (rather than rather weak permanent magnets) to $3x10^8$ [N/m^3]. I will also consider a reduction in pulse time to one 100 attoseconds or $?t = 10^{-16}$ [s]. The longitudinal-force pulse, $?f_l$ = (Volume) x($?f_l/?V$) = $(p[(9x10^{-3})^2 - (3x10^{-3})^2]\,[0.018])\,x(3x10^8) = (4.07x10^{-6})(3x10^8) = 1.22x10^3$[N]. Thus from Eq. (6) I find (with half the GW, the non-coherent half, going in the opposite direction)

$$P = \tfrac{1}{2}x\ 1.76x10^{-52}\{(2)(0.006)(1.22x10^3)/10^{-16}\}^2 = \mathbf{1.89x10^{-18}}\ [watts]. \qquad (26)$$

This power from the forward, "coherent-radiation" end is distributed over an area defined by the diffraction pattern at a distance of 7 [km] or range, $R = 7x10^3$ [m]. The diffraction angle, a_d, at the apex of a cone of HF GW is given by (please see Conference paper HFGW -03-120) $a_d \sim ?_{GW}$ /device-diameter = $c?t/(0.018) = (3x10^8)(10^{-16})/(0.018) = 1.67x10^{-5}$ [radians].

The area of the conical spread of the HF GW is

$$a = p(a_d\,R/2)^2 = p(1.67x10^{-5}x7x10^3/2)^2 = \mathbf{1.07x10^{-2}}\ [m^2]. \qquad (27)$$

The 18 [mm] diameter lens, which concentrates the HF GW at the receiver, has a grasp or GW gathering power, or amplification of $(d/?_{GW})^2 = \{(0.018)/(3x10^8)(10^{-16})\}^2 =$ **$3.6x10^9$** . Putting it all together the signal at the receiver is $\{(1.89x10^{-18})/(1.07x10^{-2})\}\{3.6x10^9\} =$ **$6.3x10^{-7}$ [watts/m^2]**, which is about an order of magnitude larger than a ten-watt isotropic EM transmitter at a 7 [km] distance ($1.6x10^{-8}$ [watts/m^2], p.42 of [10]).

Note that *the HFGW signal at the receiver is inversely proportional to the **sixth power** of the system's pulse length, ?t,* (including the lens at the receiver). The foregoing is a bit of a simplification since, like the discussion of the high-intensity design in Section 6, one would turn to a concentric, cylindrical-layer construction – not to a simple sheath and single rim configuration. Thus the energizing elements (e.g., coils) and energizable elements (e.g., magnetic sites on the rims) would be close enough for the GW waves (of wavelength $c?t = (3x10^8)\,(10^{-16}) = 3x10^{-8}$ [m] or 30 nanometers – probably much smaller in a superconductor) marching down the cylinder coherently, to build up with an electron migration distance of only $(2.38x10^8)(10^{-16}) = 23.8$ nanometers [nm].

8. NANOMECHANICAL RESONATOR

A recent paper by Melosh, *et al* [17] raises the possibility of utilizing high-

frequency nanomechanical resonators for generating HFGW. They indicate some 10^{11} junctions (presumably where resonators could be located) per square centimeter of wafer or chip. The resonance frequency of the wires is "... expected to range from tens of MHz to GHz." At the GHz frequency, $?t \sim 10^{-9}$ [s]. The ?f is more difficult to estimate. Let us assume that the force applied to the resonators is like the force between two parallel current-carrying conductors mentioned in Section 4.2.3 and exhibited schematically in FIG. (4). In this case the classical equation for the force (impulse) is

$$?f = \mu_0 l i_a i_b/2pd, \quad (28)$$

where $\mu_0 = 4px10^{-7}$, l is the length of the wire section, d is the distance between wires, and i_a and i_b are the currents flowing through any adjacent pair of wires. The wires can be as small as 20 [nm] in diameter for a wire cross section of about $3x10^{-16}$[m^2]. If one were able to conduct a TAmp/m^2 or 10^{12} Amps/m^2 (a most difficult problem especially if superconductors are contemplated), then the current in the two wires would be three tenths of a milliamp or $3x10^{-4}$ [A]. I will assume that at the resonator the wire section, l, subject to the force and the distance, d, between wires are about equal. Thus Eq. (28) yields

$$?f = (4px10^{-7})\,(3x10^{-4})^2/(2p) = 3.6x10^{-21}[N] \quad (29)$$

and if I had about a three-meter stack of these (assumed 1 [mm] thick) wafers or chips (so that the radius of gyration for a computer-logic system control to build up a coherent GW along the stack, would be r = 1 [m]), and with 10^{11} resonators per wafer, Eq. (6) yields

$$P= 3x10^3 x1.76x10^{-52} x\{2(3.6x10^{-21})(10^{11})/10^{-9}\}^2$$

$$= 6.8x10^{-50} \quad [watts]. \quad (30)$$

Clearly this is **not** a viable HFGW generator. Even if one could impress a QHz frequency on the nanowires one could only expect a 10^{12} improvement – not enough.

A better way to implement a system like this would be to utilize an array of parallel plates that are, for example, one centimeter on a side and one millimeter thick. A two-meter square plate would hold about ten thousand of them plus associated circuitry and, again, there could be a three-meter stack for a one-meter radius of gyration for the build up of a coherent beam. In this case the area of the "wire" would be (0.01) (0.1) = 10^{-3} [m^2]. Assuming only 10^5 [A/m^2], a 100 [A] current would pass through each pair (probably superconducting to avoid heat). Let us suppose that I can separate the plates by 100 [nm] or 10^{-7} [m] and that for the centimeter plates l = 0.01 [m]. Thus evaluation of Eq. (28) yields

$$?f = (4px10^{-7})(0.01)(100)^2/2px10^{-7}=200 \text{ [N]} \quad (31)$$

with QHz current pulses $?t = 10^{-15}$ [s]. There are 10^4 of the energizable elements on each plate as the GW wave passes by and if the plates are one-half centimeter apart (therefore 3[m]/0.05 = 600 of them in the stack) , so that Eq. (6) yields

$$P= 600x1.76x10^{-52}\{2x10^4 x200/10^{-15}\}^2 = 1.7x10^{-6} \text{ [watts]}. \quad (32)$$

All of this is very hypothetical, but much closer to a realistic HFGW generator.

9. SYMMETRY AND NULL GW GENERATION

There exist some situations in which a jerk exists in a system of masses, but there is no attendant GW generation. One often refers to these as *symmetrical situations.* A situation in which a system is so symmetrical that one can think of the GW as canceling out and becoming null. Of course there are reactive jerks when a star collides with a black hole resulting in some acceleration of both bodies, or one neutron star orbits another with action and reaction on both bodies and in both cases GW is generated. Also GW would be expected to be generated with the reactive jerks of coils and magnets, motion of resonate cavity walls, Cooper pairs in a superconductor, electrons in a dielectric, etc. But there are situations, such as an isotropic explosion of a star, in which there is a jerk due to an expanding shell of gas and **no** GW is generated.

This situation warrants some attention. Consider FIG. (10A) in which there is shown the cross section of an exploding shell of gas having diameter, d. I consider two opposite small incremental masses of the shell, $?m_1$ and $?m_2$, which are jerked to the top and bottom of the

figure. Let us suppose that they are jerked so that each mass generates two oppositely moving GWs, normal to the direction of the jerk in a plane perpendicular to the plane of the figure. I show as $?_{GW}$. The usual assumption for the efficacy of the quadrupole approximation for estimating GW power is that $?_{GW}$ »d, which is the case for most celestial LFGW generators. Thus I have the situation shown in FIG. (10B) and the GW cancel out and become null.

10. CONCLUSIONS

The jerk formulation of the quadrupole approximation was derived in two alternative ways for the HFGW power of a laboratory HFGW generation device. The formulation was numerically validated by analysis of a well-known GW-generating binary pulsar, PSR 1913+16. By means of numerical examples, it was shown that the resistive heat loss and device component acceleration are well within tolerable limits.

Micro- and nano-scale HFGW generator components have been described in connection with a computer logic system to facilitate the generation of coherent HFGW. A device consisting of a cylindrical stack or jerking rims was also described in connection with two specific designs: a high-intensity HFGW generator and a miniaturized HFGW generator. Moreover, the outputs of HFGW fluxfrom these devices range from $6.3x10^{-7}$ [watts/m^2] to $3.7x10^{22}$ [watts/m^2]. In either case the fluxes are greatly increased by increasing the frequency of the jerks to the QHz or higher and by including a HTSC lens to concentrate the HFGW. Linear motor HFGW-generator designs were configured and studied. The situation in which component symmetry prohibits the generation of GW due to destructive GW interference was examined. A nanomechanical resonator concept is analyzed, but not found to be especially efficient. And finally, the major conclusion is that **laboratory HFGW generation devices are feasible, practical, and warrant experimental investigation**.

REFERENCES

[1] Joseph Weber (1964), "Gravitational Waves," in *Gravitation and Relativity*, Chapter 5, pp. 90-105, W. A. Benjamin, Inc., New York.

[2] Charles W. Misner, Kip Thorne, and John Archibald Wheeler (1973), *Gravitation*, W. H. Freeman and Company, New York.

[3] Abraham Pais (1982), *Subtle is the Lord … The Science and the Life of Albert Einstein*, Oxford University Press, pp. 38, 242, 280, and 384.

[4] K. S. Thorne (1987), "Gravitational radiation," Chapter 9 of *300 Years of Gravitation*, Cambridge Press.

[5] L. D. Landau and E. M. Lifshitz (1975), *The Classical Theory of Fields*, Fourth Revised English Edition, Pergamon Press, pp. 348, 349, 355-357.

[6] H. Taylor, Jr. (1994), "Binary pulsars and relativistic gravity," *Reviews of Modern Physics*, Volume 66, Number 3, July, 1994 pp. 711-719.

[7] P. C. Peters and J. Mathews (1963), "Gravitational Radiation from Point Masses in a Keplerian Orbit," *Physical Review*, Volume 131, pp. 435-440.

[8] Samuel Herrick (1971), *Astrodynamics Volume 1*, Van Nostrand Reinhold, pp. 60 and 61.

[9] Robert M. L. Baker, Jr. (1967), *Astrodynamics, Application and Advanced Topics*, Academic Press, New York.

[10] Robert M. L. Baker, Jr. (2000), "Preliminary tests of fundamental concepts associated with gravitational-wave spacecraft propulsion," *American Institute of Aeronautics and Astronautics: Space 2000 Conference and Exposition*, Paper Number 2000-5250, September 20, August 21, 2001, Revision. (Please see Internet site at: http://drrobertbaker.com/RevisedAIAAPaper.htm .)

[11] F. Ashby, Ian Foster, James M. Lattimer, Norman, Manish Parashar, Paul Saylor, Schutz, Edward Seidel, Wai-Mo Suen, F.

D. Swesty, and Clifford M. Will (2000), "A Multipurpose code for 3-D relativistic astrophysics and gravitational wave astronomy: application to coalescing neutron star binaries," *Final Report for NASA CAN NCCS5-153*, October 15, 30 pages

[12] F. Romero B. and H. Dehnen (1981), "Generation of gravitational radiation in the laboratory," *Z. Naturforsch*, Volume 36a, pp.948-955. And: Heinz Dehnen and Fernando Romero-Borja (2003), "Generation of GHz - THz band, high-frequency gravitational waves in the laboratory," *Paper HFGW-03-102 Gravitational-Wave Conference, The MITRE Corporation*, May 6-9.

[13] Y. Acremann, *et al*, (2000), "Imaging processional mf otion of the magnetization vf ector," *Science*, Volume 290, October 20, pp. 492-495.

[14] D. Cotter, *et al* (1999), "Non-linear optics for high-speed digital information processing," *Science*, Volume 286, November 19, pp. 1523-1528.

[15] I. M. Pinto and G. Rotoli (1988), "Laboratory generation of gravitational waves," *Proceedings of the 8th Italian Conference on General Relativity and Gravitational Physics*, Cavlese (Trento), August 30 – September 3, World Scientific-Singapore, pp. 560-573.

[16] Ning Li and Douglas G. Torr (1992), "Gravitational effects on the magnetic attenuation of super conductors," *Physical Review B*, Volume 46, Number 9, p. 5491. (HFGW refraction).

[17] Nicholas A. Melosh, Akram Boukai, Frederic Diana, Brian Gerardot, Antonio Badolato, Pierre M. Petroff, and James R. Heath (2003), "Ultrahigh-density nanowire lattices and circuits," *Science*, Volume 300, April 4, pp.112-115.

[18] K. W. D. Ledingham, P. McKenna, and R. P. Singhal (2003), "Applications for nuclear phenomena generated by ultra-intense lasers," *Science*, Volume 300, May 16, p.1107.

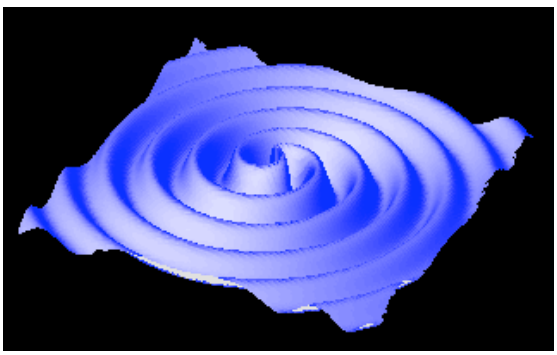
[19] Spring, L. (1954), "Discoveries and Observations," item number 68., http://www.teslatech.info/ttstore/articles/spring/discover.htm, accessed October 9, 2002.

Le onde gravitazionali

NAUTILUS

CHE COSA SONO LE ONDE GRAVITAZIONALI?

Perturbazioni del campo gravitazionale che si propagano alla velocità della luce, previste dalla **Teoria della Relatività Generale** formulata da **Albert Einstein**.
Einstein rivoluzionò i concetti di **spazio** e di **tempo** ipotizzando che uno stesso fenomeno, osservato da sistemi di riferimento diversi, potesse apparire diverso, al punto che anche il tempo potesse scorrere in modo diverso.
Al sistema di riferimento che descrive un fenomeno si deve quindi aggiungere, alle tre coordinate spaziali, una coordinata temporale passando da tre a quattro dimensioni **(spazio-tempo).**
La teoria della relatività generale, che descrive i fenomeni connessi con la gravitazione, prevede che la forma dello spazio-tempo dipenda dalla distribuzione della materia in esso presente: la massa di una stella incurva la trama dello spazio-tempo come se questo fosse un telo invisibile.
Quando tale distribuzione viene modificata, ad esempio perché una massa si sposta molto rapidamente, la forma dello spazio-tempo cambia di conseguenza. Questa variazione non è istantanea in tutto l'Universo ma si propaga, dal luogo dove ha avuto origine, alla velocità della luce. La propagazione della variazione della forma dello spazio-tempo prende il nome di onda gravitazionale e possiamo assimilarla all'increspatura sulla superficie di un lago dopo il lancio di un sasso.



QUALI SONO LE SORGENTI DI ONDE GRAVITAZIONALI?

Qualunque massa in movimento genera onde gravitazionali ma il segnale emesso è talmente debole che per avere la speranza di osservarlo è necessario avere come sorgenti grandi masse in movimento, come i buchi neri, o eventi catastrofici, come l'esplosione di supernove.

SI POSSONO RIVELARE LE ONDE GRAVITAZIONALI?

Le onde gravitazionali interagiscono debolmente con la materia. Per fare un esempio un'onda gravitazionale che attraversasse il Sole perderebbe soltanto una parte su 10^{16} della sua energia. La rivelazione di tali onde pone dunque problemi fisici e tecnologici d'eccezione.

A COSA SERVE RIVELARE LE ONDE GRAVITAZIONALI?

La loro rivelazione potrebbe essere una scoperta scientifica fondamentale per lo studio del nostro universo.
E' grazie alla loro caratteristica di interagire debolmente con la materia che, viaggiando praticamente indisturbate nello spazio, attraversano qualunque corpo celeste si trovi lungo il loro cammino.
Ciò permette di poterle utilizzare come sonde per fare osservazioni di zone del nostro universo al momento nascoste dalle stelle o dalla materia interstellare, come ad esempio gli interni delle galassie, fornendone una radiografia.
Questa conoscenza permetterebbe di studiare la distribuzione di materia nell'universo e contribuirebbe agli studi sulla sua origine ed evoluzione. Si è infatti ipotizzato che negli istanti immediatamente successivi al Big Bang sia stata emessa una rilevante quantità di onde gravitazionali. La rivelazione di queste onde di origine cosmologica fornirebbe una fotografia dell'universo nei suoi primissimi momenti di vita.
La debolezza, unita alla bassa interazione con la materia, non ha consentito, finora, una rivelazione diretta delle onde gravitazionali.

LNF-INFN Divulgazione e Pubbliche Relazioni – Silvia Miozzi

COME SI RIVELANO LE ONDE GRAVITAZIONALI?

Le onde gravitazionali possono essere rivelate utilizzando **rivelatori risonanti** o interferometrici.
Un rivelatore risonante funziona come un'antenna: il passaggio dell'onda gravitazionale provoca una variazione nelle dimensioni del rivelatore nel piano perpendicolare alla direzione di propagazione dell'onda.
Un rivelatore interferometrico misura la variazione della distanza relativa fra due masse libere, indotta dal passaggio dell'onda nello spazio che le separa.
In entrambi i tipi di rivelatori le variazioni indotte dall'onda sono inferiori alle dimensioni di un atomo.

CHE COS'E' NAUTILUS?

E' il rivelatore di onde gravitazionali risonante ultracriogenico istallato ai Laboratori Nazionali di Frascati dell'INFN dal 1992.
NAUTILUS fa parte di una rete internazionali di rivelatori che include ALLEGRO (Louisiana, USA), AURIGA (INFN, Legnaro), EXPLORER (CERN, Ginevra).

COME E' FATTO NAUTILUS?

NAUTILUS è una barra cilindrica di alluminio lunga 3 m con diametro di 0,6 m e del peso di 2770 Kg.
Il segnale da raccogliere è talmente debole che per poter essere rivelato è necessario eliminare tutti i disturbi, sia quelli dovuti alla stessa agitazione termica del materiale di cui è fatta l'antenna, sia tutti quelli provenienti dall'esterno (traffico, eventi microsismici, anche il semplice calpestio).
Per risolvere il primo problema l'antenna deve essere mantenuta ad una temperatura molto bassa. **NAUTILUS** può essere raffreddata fino a 0,1 gradi sopra zero assoluto ed in queste condizioni è l'oggetto di grandi dimensioni più freddo all'interno dell'Universo.
L'isolamento dell'antenna dalle perturbazioni esterne richiede poi l'uso di sospensioni e ammortizzatori di grande efficacia.
Sotto l'effetto dell'onda gravitazionale l'antenna entra in vibrazione. Le vibrazioni vengono convertite prima in un segnale elettrico e poi in uno magnetico che viene successivamente misurato da un dispositivo detto SQUID (Superconducting Quantum Interference Device). Misurando variazioni di lunghezza pari a 10^{-18} m, **NAUTILUS** è tra gli oggetti più sensibili finora mai costruiti.

ESISTONO LE ONDE GRAVITAZIONALI?

In diversi laboratori sparsi per il mondo si stanno eseguendo esperimenti sulla rivelazione di onde gravitazionali utilizzando sia antenne risonanti, sia interferometri. Nessuno di questi apparati è ancora riuscito ad individuare un segnale inequivocabilmente interpretabile come il passaggio di un'onda gravitazionale.
Effetti indiretti delle onde gravitazionali sono stati già osservati: per esempio la perdita di energia di un sistema binario di stelle che, ruotando molto rapidamente l'una attorno all'altra, irraggiano nello spazio onde gravitazionali.
I successi della relatività generale in altri campi ci rendono fiduciosi del fatto che tali onde esistano.

LIGO AND THE DETECTION OF GRAVITATIONAL WAVES

Large detectors on opposite sides of the country are about to start monitoring the cosmos for the gravitational waves that general relativity tells us should be emanating from catastrophic astrophysical events.

Barry C. Barish and Rainer Weiss

The idea of gravitational waves was already implicit in the 1905 special theory of relativity, with its finite limiting speed for information transfer. The explicit formulation for gravitational waves in general relativity was put forward by Einstein[1,2] in 1916 and 1918. He showed that the acceleration of masses generates time-dependent gravitational fields that propagate away from their sources at the speed of light as warpages of spacetime. Such a propagating warpage is called a gravitational wave.

The best empirical evidence we have of the existence of gravitational radiation is indirect. It comes from the 1974 discovery and beautiful observations, by Russell Hulse and Joseph Taylor,[3] of the first binary pulsar ever found. (See PHYSICS TODAY, December 1993, page 17.) Exploiting the clockwork pulsar signal from the neutron star, they were able to monitor the orbital period of the binary star system with exquisite precision and confirm that it was indeed gradually speeding up at just the rate predicted for the general-relativistic emission of gravitational waves.

The *direct* detection of gravitational waves will mark the opening of a new window on the near and far reaches of the cosmos. For physics, its most important promise is the direct observation of gravitation in highly relativistic settings, so that one can test general relativity in the strong-field limit, where it is not merely a small correction to Newtonian gravity. (See the companion article in this issue by Clifford Will, on page 38.) In that limit, the strong curvature of the spacetime geometry should show us fundamentally new physics.

By the time they reach us, the gravitational waves are, of course, only very weak perturbations on our local flat space. But they will provide information about the strong-field regions where they began. The detection of the waves will also allow us to determine the wave properties of the gravitational radiation—for example, their propagation velocity and polarization states.

For the astrophysicist, the observation of gravitational waves will provide a new and very different view of the universe. These waves arise from motions of large aggregates of matter, rather than from the particulate sources that produce electromagnetic waves. Because gravitational waves are not scattered as they propagate between source and observer, they should provide information about what's happening in the innermost and densest regions of the astrophysical source. Probing the universe in this very different way, gravitational radiation is likely to bring us exciting surprises and unanticipated new astrophysics.

A new generation of detectors based on suspended mass interferometry promises to attain the requisite sensitivity for observing gravitational waves. (See figure 1.) These new detectors are the fruit of a quarter-century of worldwide technology development, design, and construction. The US effort, called LIGO (Laser Interferometer Gravitational-Wave Observatory), is a joint Caltech–MIT project, supported by the National Science Foundation. LIGO is a pair of L-shaped laser interferometers: one in Hanford, Washington, the other, some 3000 km away, in Livingston, Louisiana. (See figure 2.) Each evacuated interferometer arm is 4 km long.

The LIGO facilities at both sites have now been completed, and detector installation is under way. Following a two-year commissioning program, we expect the first sensitive searches for astrophysical gravitational waves to begin in 2002. This initial search, sensitive to changes (strains) as small as a part in 10^{21} in the lengths of the interferometer arms, will be the first attempt to detect gravitational waves at a sensitivity that reaches plausible estimates for astrophysical source strengths. It will mark a 100- to 1000-fold improvement over previous searches—both in sensitivity and bandwidth.

The two LIGO interferometers will operate in coincidence, so as to filter out local noise. In fact, to provide additional coincidence surety, a third independent interferometer, half as long as the other two, will share the vacuum system of the full-size interferometer at Hanford. Also, one determines the direction and polarization of a gravitational wave by measuring arrival-time differences between geographically dispersed detectors. At the Hanford and Livingston support facilities, efforts will continue on the development of improved and special-purpose detectors of increased search and follow-up sensitivity.

Gravitational radiation

The gravitational wave plays a role in gravitation similar to that of the electromagnetic wave in electricity and magnetism. But because mass, unlike charge, comes in only one sign and the momentum of a free system must be conserved, the lowest-order source of gravitational radiation is quadrupolar.

The radiation field causes a strain in space itself, transverse to the propagation direction. The strain pattern contracts space along one transverse dimension while expanding it along the orthogonal direction in the transverse plane. One way of imagining this distortion of space

BARRY BARISH *is a professor of physics at the California Institute of Technology in Pasadena and director of the LIGO laboratory.* RAINER WEISS *is a professor of physics at the Massachusetts Institute of Technology in Cambridge and the LIGO project's integration scientist.*

© 1999 American Institute of Physics, S-0031-9228-9910-040-1

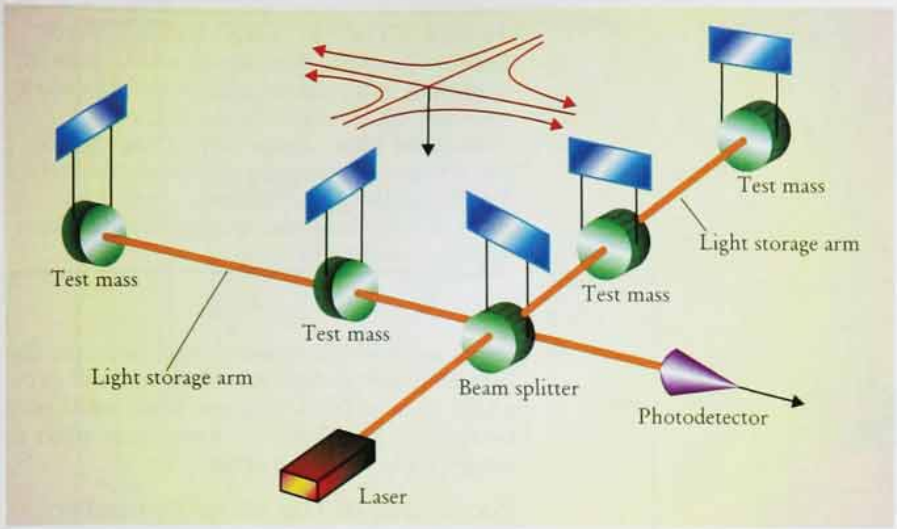


FIGURE 1. THE LIGO GRAVITATIONAL-WAVE DETECTORS are equal-arm Michelson laser interferometers whose hanging mirrors serve as the gravitational test masses. An incident gravitational wave, indicated in red by the stress pattern coming down from above, stretches one interferometer arm and compresses the other, causing a difference between light travel times in the two orthogonal arms. This time difference is manifested in the interference pattern when the two laser beams recombine on the way to the photodetector, which can measure phase shifts to a few ten-billionths of an interference fringe.

is to look at the weave in a piece of cloth when it's pulled along one dimension. The little squares of the weave distort in just this way. Furthermore the strain is quite uniform, so that the relative motion of points in the cloth depends linearly on their separation.

It is this linear increase in the relative motion that provides the motivation for LIGO's 4 km interferometer arms. Such ambitious length is intended to provide adequate sensitivity to passing gravitational waves in the face of inevitable local perturbations. The tensor character of gravity (the putative graviton is a spin-2 particle) means that the push/pull pattern of the strain field for a plane gravitational wave has two orthogonal polarizations. If a candidate signal really is a gravitational wave, and not just noise, the half-length auxiliary interferometer at the Hanford site should see a coincident displacement half as large as that experienced by its full-length neighbor.

One can also think of the gravitational wave as producing a tidal force field such that the relative force between masses grows as their separation. The field will pull masses together along one transverse direction while pushing them apart along the orthogonal direction. How one chooses to view the wave–matter interaction is really a matter of taste. But one must be careful to maintain consistency in one's view and not mix the geometric and tidal representations.

The strength of the gravitational field is expressed by the dimensionless strain

$$h = \left(\frac{GM}{Rc^2}\right)\left(\frac{v^2}{c^2}\right)$$

where the factor in the first pair of brackets is the Newtonian potential due to a source mass M at a distance R, divided by the square of the speed of light. On the surface of the Earth, that comes to 10^{-8}, a very weak field. But, at the surface of a neutron star, it can be as large as 10^{-1}. And at the horizon of a black hole, it is close to unity—the ultrarelativistic limit of a strong gravitational field. The factor in the second bracket pair—an estimate of the system's kinetic energy in asymmetric motion relative to its rest energy—is a measure of the strength of the relativistic dynamics.

This expression gives us an immediate estimate of the scale of the strains we might encounter. Consider, for example, a solar-mass source at a Galactic distance, moving at about 10% of the speed of light. In such a case, the strain we would observe halfway across the Galaxy would not exceed a part in 10^{18}. This simple estimate explains why we have to take such heroic measures to detect the strains. Even over the 4 km span of a LIGO arm, the relative displacement of two objects would be only a few times 10^{-13} cm, just about the size of an atomic nucleus! It's even worse than that. As discussed below, plausible sources typically lead to strains of only 10^{-21}, corresponding to LIGO displacements a thousand times smaller than the width of a nucleus.

Candidate sources

There is a large range of processes in the universe that should produce detectible gravitational waves.[4] Terrestrial interferometers like LIGO will search in the frequency range from 10 Hz to 10 kHz for characteristic signals from a variety of astrophysical sources for which one might hope to discern the signatures of gravitational radiation over the background noise. (See figure 3.)

▷ **Chirp signals.** The terminal spiraling of a star into a "compact" binary partner (a neutron star or a black hole) will produce radiation that increases in amplitude and frequency as the two move toward final coalescence. This chirp signal can be well characterized, depending on parameters such as the mass, separation, and orbital eccentricity. That makes it possible to formulate efficient detection templates.

▷ **Burst signals.** The collapse of a supernova may produce gravitational radiation. Type II supernova collapses can generate strong gravitational radiation, *if* the core collapse departs sufficiently from axial symmetry. Estimates suggest that detection might be possible for such collapses as far out as the Virgo Cluster of galaxies, some 50 million

Initial parameters for the LIGO detectors

Arm length	4000 m
Arm cavity storage time	880 μs
Laser type and wavelength	Nd:YAG, $\lambda = 1064$ nm
Input power at recycling cavity	6 W
Power recycling gain	30
Mirror mass	10.7 kg
Mirror diameter	25 cm
Mirror loss	$< 1\times10^{-4}$
Mirror internal Q	1×10^{6}
Cavity input mirror transmission	3×10^{-2}
Pendulum Q (structure damping)	1×10^{5}
Pendulum period (single)	1 s
Seismic isolation system	–110 dB at 100 Hz

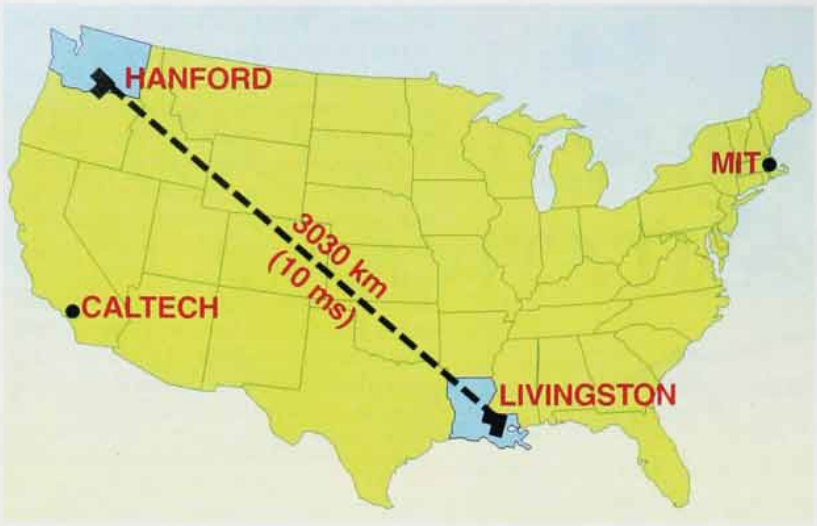


FIGURE 2. THE TWO LIGO SITES, 3030 km apart, in Hanford, Washington and Livingston, Louisiana, will work in coincidence. The recent photo of the Hanford site shows the two orthogonal 4 km vacuum pipes going off into the distance. The vacuum system also houses a smaller auxiliary interferometer, with 2 km arms to help distinguish true gravitational wave signals from noise.

light-years away. That would yield type II supernova observation rates of one or more per year. Another possible burst source accessible to LIGO is the brief, burplike oscillation of a black hole's event horizon just after it swallows a star. The detection of supernova or postprandial black hole events will require coincident observation of burst signals in several geographically dispersed interferometers.

▷ **Periodic signals.** Radiation from the nonaxisymmetric motion of a neutron star, or of the nuclear fluid on its surface might produce periodic signals in the detectors. Happily, for many known pulsars the frequency of such periodic signals lies within LIGO's sensitivity band. The searches for periodic gravitational signals from identified neutron stars will be facilitated by the fact that one can track the system continually over very many cycles, taking account of the gradual slowing of the pulsar's spin and the Doppler shifts and amplitude variation due to the Earth's diurnal and annual motions. We expect to perform general sky searches as well as targeted searches of known pulsars.

▷ **Stochastic signals.** Signals from gravitational waves emitted in the first instants of the early universe—as far back as the Planck epoch at 10^{-43} seconds—can be detected by way of correlations of background signals from two or more detectors. Some models of the early universe predict detectable signals. Such relic gravitational radiation would provide us with an exciting new cosmological probe.

The initial parameters for the LIGO interferometers have been chosen to provide a sensitivity with a reasonable chance for detecting gravitational waves. (See the table on page 45.) The anticipated rates for the various sources, however, are burdened with large uncertainties. As future advances in detector sensitivity increase the distance over which one can find sources, the rate at which events are observed will grow *as the cube* of LIGO's reach. That lends particularly high priority to a vigorous effort to improve the system's sensitivity.

Basic idea of the interferometer

A Michelson interferometer operating between freely suspended masses is ideally suited to detect the antisymmetric compression and distension of space induced by gravitational waves.[5] Figure 1 is a schematic drawing of the LIGO equal-arm Michelson interferometer. The two interferometer arms, each 4 km long in the full-length detectors, have identical light-storage times. Light sent from the laser light source to the beam splitter is divided evenly between the two arms.

Having traversed the arms, the light is reflected back to the splitter by mirrors at their far ends. On the return journeys to the photodetector, the roles of reflection and transmission in the splitter are interchanged for the two beams and, furthermore, the phase of the reflected beam is inverted by 180°. Therefore the recombined beams heading toward the photodetector interfere destructively, while the beams heading back to the laser source interfere constructively. If the interferometer arms are of precisely equal length, the photodetector ideally sees no light, all of it having been diverted, by perfect interference, back to its source.

One would get this kind of perfect interference if the beam geometry provides a single phase over the propagating wavefront. An idealized uniphase plane wave has this property, as does the Gaussian wavefront in the lowest-order spatial mode of a laser. Then, provided the arms are equal in length (or their length difference is a multiple of half the wavelength of the monochromatic beam), the photodetector sees no light at all. The destructive interference over the entire beam wavefront is complete.

If, in the absence of any disturbance, the interferometer is carefully balanced so that no light appears at the photodetector, a sufficiently strong gravitational wave passing though the interferometer can disturb this balance and cause light to fall on the detector. That, in essence, is how LIGO will sense gravitational waves. To obtain the required sensitivity, we have made the arms 4 km long, and we have included two refinements:

▷ First, the intensity change at the photodetector due to a gravitational wave depends on the interaction time of the wave with the light in the arms. The longer this interaction time—up to half the period of the gravitational wave—the larger is the resulting optical phase shift and the consequent change of the light intensity at the photodetector. To gain further interaction time, beyond what one gets simply from the 4 km arm length, the initial

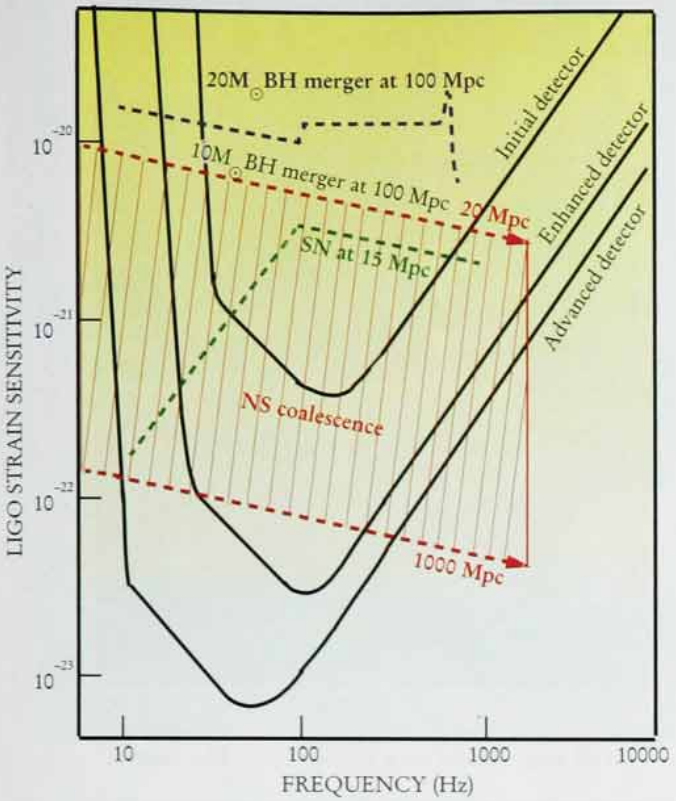


FIGURE 3. RMS STRAIN SENSITIVITY LIMITS as a function of signal frequency, for three LIGO generations indicated by the U-shaped black curves, are compared with signal estimates for various astrophysical sources. The "enhanced" detector is anticipated for 2006, and the "advanced" detector four years later. The shaded red region indicates the strain signal expected from the coalescence of two neutron stars at distances from 20 to 1000 megaparsecs (1 Mpc = 3×10^6 light-years), and from the merger of two 10 $M_\odot$ black holes at least 100 Mpc away. The larger and more structured signal expected from the merger of two 20 $M_\odot$ black holes at 100 Mpc is indicated by the purple dashed curve. The green dashed curve indicates the signal expected from an asymmetric supernova 15 Mpc away. One expects a few events per year within the red parallelogram.

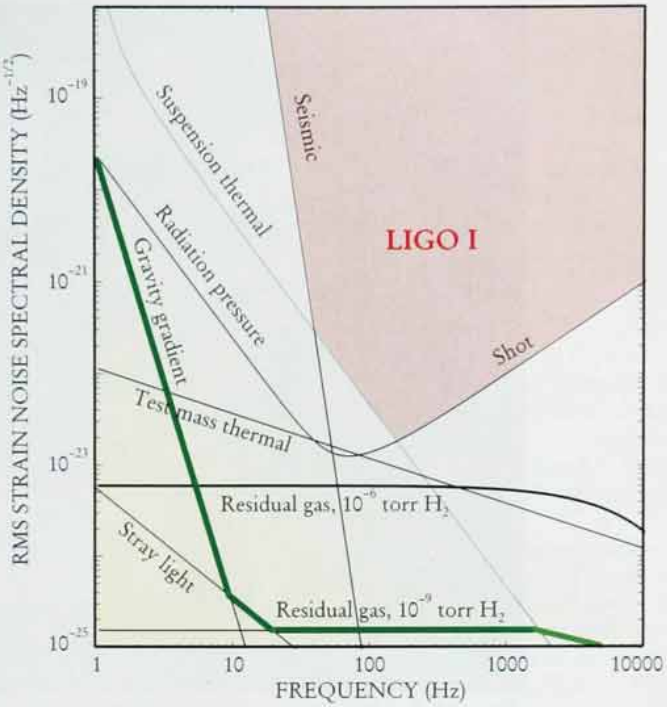


FIGURE 4. LIMITING NOISE SOURCES for the initial LIGO interferometers are shown on a plot of frequency against spectral noise density. The vertical axis denotes the RMS strain noise in 1 Hz of bandwidth. The noise increases as the square root of bandwidth, so that the noise in a typical 100 Hz window would be 10 times that shown on the axis. At the lowest frequencies, sensitivity is limited by geophysical and man-made seismic noise; at intermediate frequencies by thermal noise; and at the highest frequencies by the shot noise of photon statistics. The green line represents the minimum noise at the present LIGO facilities, irrespective of eventual detector upgrades.

LIGO interferometers will also fold the optical beams within the arms by means of optical cavities. This trick results in a light-storage time of about 1 millisecond. That's about 50 times longer than a simple straight transit through a 4 km arm.

▷ A second refinement increases the interfering light intensity by making the entire interferometer a resonant optical storage cavity. Most of the light interferometrically diverted from the photodetector direction—when the arms are unstrained—returns toward the light source. That makes it possible to achieve a significant gain by placing another mirror between the laser and the beam splitter. By properly choosing this extra mirror's position and making its transmission equal to the optical losses inside the interferometer, one can match the losses so that no light at all is reflected back to the laser. This is equivalent to increasing the laser power by about a factor of 30, without adversely affecting the frequency response of the interferometer to a gravitational wave.

Sensitivity limits

The success of the detector will ultimately depend on how well we can control the noise in the measurement of the exceedingly small strains we have been discussing. That has been the prime technological challenge in this field for the past several decades, and it is the central focus of our development of the technology for LIGO. The noise we have to contend with is broadly divided into sensing noise, random force noise and, ultimately, quantum noise. Sensing noise involves the various phenomena that limit our ability to sense and register the small motions in question. Random force noise, on the other hand, results from disturbances that cause small motions of the suspended masses. Eventually one confronts the ultimate quantum noise limit. This orderly classification presumes that one is careful enough in the design and execution of the experiment to reach the fundamental limits. The quantum limit will not be an issue for the first or second generation of LIGO detectors. So we do not address it in this article. There is, however, important ongoing work that seeks to understand the quantum noise limit and develop techniques to circumvent it in measuring the strain.

In order to approach the fundamental limits, we have made extensive use of two concepts in experimental physics promoted by Robert Dicke (1916–97) of Princeton University. The first is the technique of modulating the signal to be detected at frequencies far above the $1/f$ noise due to the drift and gain experienced by all instruments. For example, we measure the optical phase to determine the motion of an interference fringe at radio frequency rather than near DC.

A second concept is to apply feedback to physical vari-

FIGURE 5. LIGO'S SEISMIC ISOLATION SYSTEM consists of four layers of masses and springs. Each of the coil springs seen here is made by lining the inside of a straight metal tube with rubbery damping material and then filling the lined tube with a line of metal slugs strung on a rubber core. The tube is then coiled and sealed.

ables in the experiment in order to control and damp large excursions at low frequencies. The variable is measured by way of the control signal required to hold it stationary. A good example is the position of the interferometer mirrors. At low frequencies, we maintain the interferometer fringe at a fixed phase by holding the mirrors at fixed positions with coil/magnet actuators.

Sensing noise

Our ability to determine the relative motions of the mirrors at the ends of the arms interferometrically is limited by the smallest change in optical phase that we can measure. The light emitted by a conventional laser is in a coherent state in which the photon occupation number n obeys a Poisson distribution with variance

$$\Delta n = \sqrt{n} = \sqrt{\dot{n}\tau}$$

where $\dot{n}$ is the rate at which photons encounter the beam splitter and τ is the integration or observation time. Because the phase and photon occupation number are conjugate variables obeying an uncertainty relation, one gets

$$\Delta\phi = 1/\sqrt{\dot{n}\tau}$$

for the variance in the interferometric measurement of the relative phase of the recombining beams at the photodetector. We expect the optical-phase variance in the initial LIGO detector to be about 3×10^{-10} radians, corresponding to a strain variance in a 10-millisecond measurement of about 2×10^{-22}. That would be the fundamental Poisson limit. It is sometimes called the shot-noise limit, because it can also be derived from the statistics of photon counting in the photodetection. This shot noise determines LIGO's sensitivity limit for frequencies above 300 Hz. (See figure 4.)

Before one reaches this limit, however, one has to deal with a host of practical problems, such as laser frequency fluctuations, laser amplitude noise, and stabilization of the beam geometry. We must also reduce additional sensing noise terms that can occur in the beam propagation—for example, scattering by residual gas molecules and scattering off the vacuum tube walls driven by seismic and acoustical noise. We limit these effects by using baffling and low-scatter optics in the evacuated beam tubes. But even if one controls these noise terms and achieves the fundamental Poisson noise limit, one cannot easily reduce the noise any further by simply increasing the laser power to get more photons. That remedy raises problems of optical heating in the mirrors and coatings and, finally, radiation pressure fluctuations.

Random force noise

At lower frequencies, the sensitivity limit is set by how well the motions of our test masses—the hanging mirrors—are controlled. At the lowest frequencies (about 10–100 Hz), the largest disturbances come from "seismic noise"—the motion of the Earth's surface driven by wind, water flow, and human activities, as well as by low-level earthquakes. At intermediate frequencies (100 Hz), the principal culprit is thermal noise—that is, Brownian motion driven by thermal excitations. Less important for the initial LIGO interferometers, but increasingly significant as the detectors are upgraded, will be fluctuations in the Newtonian gravitational forces on the mirrors resulting from density fluctuations in the ground and the atmosphere and, ultimately, the radiation pressure fluctuations.

In general, these random forces are not correlated at the different mirrors, and they are independent of the length of the interferometer arms. By contrast, displacements due to gravitational waves grow linearly with the arm length. That's our principal motivation for going to the expense and trouble of having 4 km arms.

The LIGO suspended mirrors, which serve as the test masses, are isolated from motions of the Earth by cascaded stages of vibration isolation. The first level of isolation, consisting of four stages of springs and masses, reduces the seismic motion a millionfold at frequencies around 100

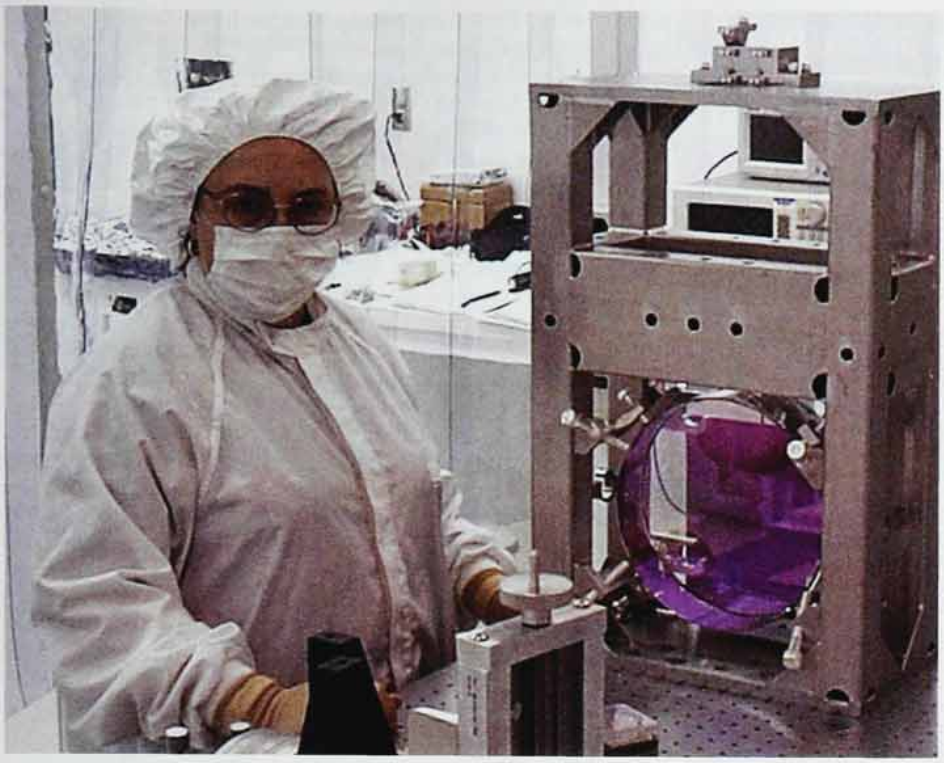


FIGURE 6. A LIGO MIRROR, 25 cm in diameter and 10 cm thick, is made of ultrapure fused silica. Its purple coating is a highly reflective multilayer stack of dielectric materials. Absorption and scattering losses must not exceed a few parts per million.

Hz, and progressively more at higher frequencies. (See figure 5.) This isolation works much like the suspension in a car. The final stage of the isolation system is the hanging mirror itself. Each test mass is, in effect, a pendulum suspended by flexures. The pendulum provides another stage of vibration isolation. But, more important, it also serves to reduce the influence of thermal noise.

Mechanical thermal noise enters the system by exciting the pendulum, causing the test mass to move, and by exciting acoustic waves that disturb the mirror surface. The acoustic noise can be represented as a superposition of the motions induced in the normal modes of the mass. The strength of the perturbation is estimated by taking the overlap of the acoustic-mode shape with the optical wavefront. The equilibrium thermal excitation of each normal mode at temperature T is $kT/2$, yielding significant motion at the principal resonant frequencies. Therefore we choose these frequencies to be outside LIGO's detection band for gravitational waves.

The thermal noise is a more fundamental and difficult problem than the seismic noise. Our primary techniques for reducing the thermal noise are to cool specific modes and to design systems with low dissipation. The seismic noise, by contrast, is motion relative to the inertial frame. So one can use the inertial frame as a reference to reduce the driving accelerations.

Detection strategies and confidence

In developing LIGO's search techniques, statistical tests, and detection criteria, we seek to minimize false observations. Within the statistics associated with the instrument noise, a viable gravitational wave signal from a distant astrophysical source must appear in the data streams of all three LIGO interferometers in the US, and of any other detectors in a worldwide network of comparably sensitive instruments.

For specific astrophysical searches, we will require signals consistent with calculated expectations of how the frequency varies with time. For the terminal in-spiraling of a binary system with a neutron star, for example, one can calculate the waveform as a function of the system parameters. So we can compare a candidate chirp signal over thousands of cycles, as it crosses LIGO's sensitivity band, with detailed templates of calculated waveforms.

Futhermore, the geographically dispersed detectors will have to exhibit consistent waveforms in proper coincidence. There will also be anticoincidence vetoes to weed out environmental effects. The hardest problem in a burst search is the elimination of false signals associated with non-Gaussian noise in the individual interferometers. By requiring multiple-detector coincidence, we can reduce the rate of such spurious events to less than one per decade.

Periodic sources will have to satisfy a very special set of criteria. The observed signal must exhibit amplitude modulation and Doppler frequency modulation consistent with the effects of the Earth's rotation and revolution around the Sun.

A stochastic background of gravitational waves can be detected by searching for a common "noise" in a set of interferometers. The detection requires the cross-correlation of two or more interferometers. In the LIGO geographic configuration, the cross correlation will be made between the Washington and Louisiana interferometers, with some penalty in bandwidth due to the large separation. We will also be able to correlate the two interferometers at the Washington site, assuming that their independence is not overly compromised by correlated perturbations at the same location.

Plans for the future

At first, LIGO will carry out a broadband search, because we do not know what kinds of astrophysical or cosmological sources we are most likely to see first. The LIGO facilities have been designed for a lifetime of 30 years, during which time, we expect, there will be a continuing and active program of detector development. The facilities can accommodate detectors operating at the quantum limit of a 1 ton mass and at the Newtonian limits imposed by the terrestrial environment. The vacuum and optical systems have been designed so as not to compromise eventual operation at these ultimate limits. It should be possible eventually to operate improved LIGO detectors that are several hundred times more sensitive than what we will start with next year.

Our initial detector design is a compromise between performance and technical risk. It incorporates some educated guesses as to what directions we should take to arrive at a reasonable probability for finding gravitational waves. It is a broadband system with modest optical power in the interferometer arms and a low-risk vibration isola-

science serving society

POSTDOCTORAL POSITION IN EXPERIMENTAL PHYSICS/CHEMISTRY

The Chemical Science and Technology (CST) Division is seeking qualified postdoctoral candidates to perform research involving the optical and magnetic trapping of selected radioactive atoms. The successful applicant will join a multidisciplinary research team composed of atomic physicists, nuclear physicists, and nuclear chemists. Our research is concentrated on fundamental electroweak interaction measurements (beta-decay asymmetry and beta-neutrino correlations of polarized atoms) and on the sympathetic cooling and Fermi degeneracy studies of cold atoms. The ultra-sensitive detection and quantitative isotopic ratio measurements of selected species is also under development. Applicants must have a strong background in experimental physics and/or chemistry with experience in using single-frequency lasers and in performing high-precision atomic spectroscopic measurements. Expertise in laser cooling, trapping, and fluorescent detection of atoms is highly desirable. For additional technical information, please contact **Dr. David J. Vieira** at **vieira@lanl.gov.**

A Ph.D. completed within the last three years or soon to be completed is required. Further details about the Postdoctoral Program may be found at: http://www.hr.lanl.gov/postdoc/. For consideration, submit a resume and publications list along with a cover letter outlining current research interests, including the names of three people who will provide references to postdoc-jobs@lanl.gov (no attachments, please!) OR submit two copies to: **Postdoc Program Office, AIP-PD995065, MS P290, Los Alamos National Laboratory, Los Alamos, NM 87545**. NOTE: Reference **#AIP-PD995065** in the e-mail Subject line (or the address) and cover letter.

AA/EOE

Los Alamos
NATIONAL LABORATORY
Operated by the University of California for the Department of Energy

UCI
University of California, Irvine

University of California, Irvine Major Initiative in the Physics and Chemistry of Materials

As part of a new major multidisciplinary initiative in materials physics and chemistry, the School of Physical Sciences at the University of California, Irvine, invites applications for at least five new faculty openings at all levels. Applications are encouraged from candidates in all areas of materials physics and chemistry, including but not limited to the areas of biomaterials, nanomaterials, electronic, photonic and magnetic materials, polymers, ceramics, atomic scale manipulation, novel applications of materials to technology, and theoretical work related to these topics. At the Assistant professor level we are seeking outstanding Ph.D.s who will establish vigorous research programs in areas that will complement existing strong programs in both the Department of Physics and Astronomy and the Department of Chemistry. The ability to effectively teach physics or chemistry at the undergraduate and graduate levels is required. Applicants for Assistant Professor positions should send their curriculum vitae, a list of publications and a research proposal to **Dean Ronald Stern, Materials Search Committee, School of Physical Sciences, University of California, Irvine, California, 92697-4675.** Applicants should also arrange to have three letters of recommendation submitted on their behalf. Candidates for the senior faculty openings should submit their curriculum vitae including a short description of their present research activities, and a list of potential references to the address above. To ensure full consideration during the 1999/2000 academic year, applications and all supporting materials should be received by **November 15, 1999.**

The University of California is an Equal Opportunity Employer committed to excellence through diversity.

tion system. The mirror suspensions have been well tested in prototype interferometers.

We expect to make improvements in the LIGO interferometers following the first scientific data run, which is scheduled to end in 2004. These improvements will include a new suspension system, provided by the collaborating GEO project, to further reduce the thermal noise. We may also, at that point, change to sapphire test masses. We also expect that significant improvement in the seismic isolation of the test masses will extend LIGO's sensitive observation band down to 10Hz.

We plan to reduce the sensing noise by going to a new interferometer configuration and by applying higher-power lasers in conjunction with improved optical materials and techniques to handle the higher power.

We expect that LIGO's sensitivity at 100 Hz will be improved by about a factor of 15, and that the overall high-sensitivity band will be expanded significantly to both lower and higher frequencies. That should expand the cosmic volume LIGO can search at a given sensitivity—and hence the discovery rate—by a factor close to 3000.

In the longer run, greater changes in the detector might use still newer interferometer configurations to drive the system to the ultimate limits dictated by quantum fluctuations and fluctuations in terrestrial gravity. It will be particularly interesting to improve LIGO's sensitivity for detecting periodic sources and possibly even a stochastic background of primordial gravitational waves. Searching for this speculative primordial background at high frequencies, where stochastic noise is tolerable, can be accomplished by using interferometers that greatly reduce the phase noise of the interference fringes at the cost of reduced bandwidth.

The scientific collaboration

As we enter LIGO's commissioning phase, we have expanded the scientific community's involvement by creating the LIGO Scientific Collaboration. It presently consists of about 30 research groups comprising more than 200 physicists and astrophysicists. We expect the collaboration to continue to grow and become the scientific center of LIGO as it develops over the next decade.

It is, of course, difficult to predict how LIGO will really evolve. But we believe we have set out on a course that has bright prospects for the early detection of gravitational waves. We plan a flexible approach toward improvements that will either let us follow up sources that have been detected or, if we find nothing at first, undertake more sensitive searches.

There are plenty of opportunities for new technical ideas and search methods. We look forward to developing an international collaboration with other gravitational-wave detectors to form a world-wide network. After LIGO's first data run, we plan to interleave subsequent searches with a series of detector upgrades that promise to lead to ever-enhanced sensitivity, making the direct detection of gravitational waves a reality within the next decade.

References

1. A. Einstein, Reports of the Physical–Mathematical Session of the Royal Prussian Academy of Sciences (1916), p. 688.
2. A. Einstein, Reports of the Physical–Mathematical Session of the Royal Prussian Academy of Sciences (1918), p. 154.
3. R. A. Hulse, J. H. Taylor, Astrophys. J. **195**, L51 (1975).
4. K. S. Thorne, in *300 Years of Gravitation*, S. Hawking, W. Israel, eds., Cambridge U. P., Cambridge, England (1987), chap. 9.
5. P. R. Saulson, *Fundamentals of Interferometric Gravitational Wave Detectors*, World Scientific, Singapore (1994). ■

Numerical Modeling of Gravity Wave Generation by Deep Tropical Convection

TODD P. LANE

Department of Mathematics and Statistics, Monash University, Clayton, Australia and Cooperative Research Centre for Southern Hemisphere Meteorology, Clayton, Australia

MICHAEL J. REEDER

Department of Mathematics and Statistics, Monash University, Clayton, Australia

TERRY L. CLARK

National Center for Atmospheric Research, Boulder, Colorado*

(Manuscript received 8 March 1999, in final form 21 July 2000)

ABSTRACT

Although convective clouds are known to generate internal gravity waves, the mechanisms responsible are not well understood. The present study seeks to clarify the dynamics of wave generation using a high-resolution numerical model of deep convection over the Tiwi Islands, Australia. The numerical calculations presented explicitly resolve both the mesoscale convective cloud cluster and the gravity waves generated. As the convective clouds evolve, they excite gravity waves, which are prominent features of the model solutions in both the troposphere and stratosphere. The source location is variable in time and space but is related to the development of individual convective cells. The largest amplitude gravity waves are generated when the cloud tops reach the upper troposphere.

A new analysis technique is introduced in which the nonlinear terms in the governing equations are taken as the forcing for linear gravity waves. The analysis shows that in the present calculation, neither the shear nor the diabatic heating are the dominant forcing terms. Instead, the wave source is most easily understood when viewed in a frame of reference moving with the wind at the level of neutral buoyancy, whereupon the source may be described as a vertically oriented, oscillating convective updraft. This description is consistent with the properties of the modeled stratospheric waves.

1. Introduction

Vertically propagating gravity waves are accompanied by a vertical flux of horizontal momentum, with horizontal momentum transferred to the mean flow whenever the waves are transient or experience attenuation (by dissipation, saturation, wave breaking, or absorption at a critical level). The force exerted by such waves plays an important role in shaping the mean flow in the middle atmosphere, and for this reason, nearly all general circulation models (GCMs) include, albeit crudely, parameterizations of gravity wave drag. However, most researchers have focused on orographically generated gravity waves even though waves generated by nonorographic sources are thought to contribute significantly to the total momentum budget of the atmosphere (Fritts 1993). Such nonorographic sources include midlatitude frontal systems (e.g., Reeder and Griffiths 1996), upper-level jets (e.g., Guest et al. 2000), and dry boundary layer convection (e.g., Clark et al. 1986; Kuettner et al. 1987; Hauf and Clark 1989). Deep tropical convection is another wave source and is the focus of the present paper.

It is likely that deep tropical convection is an important source of gravity waves. According to Gray (1973), precipitating convective clusters occupy about 15%–20% of the tropics between 5° and 25° latitude, and numerous observational studies link gravity wave activity[1] to tropical convection (e.g., Larsen et al. 1982;

* The National Center for Atmospheric Research is sponsored by the National Science Foundation.

Corresponding author address: Todd Lane, NCAR, P.O. Box 3000, Boulder, CO 80307-3000.
E-mail: lane@ucar.edu

[1] "Gravity wave activity" is an imprecise term commonly used to describe observed wave-like perturbations in thermodynamic and dynamical variables. However, it is usually not possible to establish conclusively whether or not the perturbations are due to gravity waves.

© 2001 American Meteorological Society

Pfister et al. 1993; Sato et al. 1995; Karoly et al. 1996). Moreover, idealized two-dimensional modeling shows that tropical squall lines are strong gravity wave radiators (Fovell et al. 1992; Alexander et al. 1995; Alexander and Holton 1997). Among other things, gravity waves generated by deep tropical convection are thought to be important in driving the quasi-biennial oscillation (e.g., Dunkerton 1997).

As well as contributing to the drag in the middle atmosphere, convectively generated gravity waves modify the region around the convective cloud, communicating the effect of latent heating within the cloud to the environment (e.g., Bretherton and Smolarkiewicz 1989; Mapes 1993; Shutts and Gray 1994). Mapes (1993) showed that a heating profile typical of a mesoscale convective system generates gravity waves that displace low-level parcels upward. He argued that these displacements reduce the convective inhibition (CIN), thereby increasing the possibility of further convection. In this way, convectively generated gravity waves may be important in organizing convection and initiating new cells (e.g., Lin et al. 1998).

Although the gravity waves generated by convective clouds are thought to be important, it is not possible to observe the wave field in sufficient detail in both time and space to determine how much these waves contribute to the global momentum budget. Current observational techniques make simplifying assumptions about the wave field and hence introduce uncertainty in the deduced wave parameters, especially the momentum flux. Furthermore, convective systems are complicated, and it is difficult to accurately observe their evolution and internal structures. At present, explicit cloud-scale modeling is the only way to describe the cloud and wave field in sufficient detail to explore the dynamics of the wave source and quantify the drag exerted by the waves on the middle atmosphere.

The approach adopted in the present study is to simulate deep tropical convection using a high resolution three-dimensional numerical model and to use the numerical solution to investigate the generation of gravity waves by the convection. The calculations presented are initialized using an idealized sounding motivated by the conditions observed during the recent Maritime Continent Thunderstorm Experiment (MCTEX), and follow the evolution of a convective cloud cluster over the Tiwi Islands just north of Darwin, Australia. The convection is initiated by sea-breeze convergence, and the numerical model explicitly resolves the process by which the waves are generated.

MCTEX was an international field experiment held during November and December 1995 focusing on the deep multicellular thunderstorms that form regularly over Bathurst and Melville Islands, which together comprise the Tiwi Islands [see Keenan et al. (1994) for details on MCTEX]. These thunderstorms are known locally as ''Hectors,'' and they develop during the months leading up to the summer monsoon and during monsoon breaks; a climatology of these storms and a description of their evolution is presented by Keenan et al. (1990) [see also the numerical modeling study by Golding (1993) and Crook (1997) and the recent review by Reeder and Smith (1998)]. Convection generally develops in the early afternoon along a line of sea-breeze convergence. Thereafter, the convection is initiated and controlled by the evaporatively driven cold pools. Later on, individual storms merge and may form squall-like structures that move off the islands. Keenan et al. (1994) found that the vertical velocity and radar structure have more similarity to continental convection than to oceanic convection. In particular, maximum updraft velocities are nearly an order of magnitude larger than those typical of oceanic convection. The environment in which the storms develop typically has moderate convective available potential energy (CAPE) (approximately 1500 J kg^{-1}), low to moderate shear (ranging from 0.25 to 2.5 $\times$ 10^{-3} s^{-1}), and high low-level moisture (approximately 20 g kg^{-1}) (Keenan et al. 1990).

A detailed understanding of how convective clouds generate gravity waves is a prerequisite to parameterizing their effects in GCMs. At present, there are at least three different hypotheses in the literature. Perhaps the most common hypothesis is that gravity waves are forced by temporal variations in the diabatic heating within the cloud (e.g., Salby and Garcia 1987; Bretherton 1988; Lin et al. 1998; Pandya and Alexander 1999). This idea forms the basis of a gravity wave parameterization developed recently by Chun and Baik (1998). A second possibility is the so-called *obstacle effect.* It is argued that if the cloud updraft develops in shear, as is usually the case, then the cloud will partially block the flow, producing a cloud-relative flow across the top of the cloud (e.g., Clark et al. 1986; Hauf and Clark 1989). With this idea in mind, it is sometimes said that clouds behave like *moving mountains* (e.g., Pfister et al. 1993). A third hypothesis, first proposed by Pierce and Coroniti (1966), was explored by Fovell et al. (1992). These authors argued that when individual updrafts within a convective system strike the tropopause, they decelerate quickly, subsequently oscillating about their level of neutral buoyancy. Fovell et al. described the individual clouds as *mechanical oscillators.*

The central aim of the present study is to clarify how convective clouds generate gravity waves. The present paper presents an analysis of the wave source. Evaluation of these source terms shows that the dominant terms are related to the idea of a mechanical oscillator of Fovell et al. (1992). The nonlinear analysis developed here is motivated by Lighthill's theory for the generation of sound (Lighthill 1952, 1954). In his theory, Lighthill treated sound waves as linear disturbances forced by (nonlinear) unstratified turbulent motion and introduced a radiation tensor to characterize the wave source. We show that a similar radiation tensor can be formulated, which describes the source of linear motions forced by convective clouds, including the generation of gravity

waves. The ideas behind Lighthill's radiation tensor have been adapted previously in a number of problems dealing with the generation of gravity waves by internal circulations. For example, Stein (1967) used a similar approach to investigate the generation of gravity waves in an isothermal (solar) atmosphere, and Ford (1994) recently applied the method to the emission of inertia-gravity waves by vortical flows in a rotating shallow water system. Reeder and Griffiths (1996) adapted Lighthill's approach to study the generation of gravity waves by fronts and jets.

All previous numerical studies on gravity wave generation by deep tropical convection modeled intense two-dimensional squall lines (Fovell et al. 1992, Alexander et al. 1995; Alexander and Holton 1997). Such systems are highly organized and appear to be strong gravity wave radiators. The present paper explores the generation of gravity waves in three-dimensional cloud systems, such as those that develop over the Tiwi Islands. Such systems are only weakly organized when compared with squall lines, but they are much more common in the Tropics. The convective clouds modeled here are, in essence, highly transient point gravity wave sources, whereas the squall lines modeled previously are long-lived line sources.

The paper is organized as follows. Section 2 briefly summarizes the numerical model used and outlines the numerical experiments considered. Section 3 discusses the evolution of the convective system and the structure of the gravity waves generated. The vertical flux of horizontal momentum in the stratosphere is calculated, and the structure and evolution of a particular convective cell is examined in detail. In section 4, a forced, linear wave equation is formulated, and a *source tensor* is introduced. The modeled convection is analyzed in terms of wave forcing, and the dynamics of wave generation are clarified. In section 5, the wave response in the stratosphere is shown to be consistent with the forcing mechanisms identified in section 4. Finally, our conclusions are summarized in section 6.

2. The numerical model

a. Model description

The numerical model used was originally developed by Clark (1977). In brief, the model is three-dimensional, nonhydrostatic, and anelastic. It is formulated in terms of terrain following height coordinates, and the finite difference approximations are second order in both space and time. The model includes parameterizations of the boundary layer physics, warm rain processes, ice microphysics (Koenig and Murray 1976), and subgrid scale turbulence (Lilly 1962; Smagorinsky 1963). For simplicity, the calculations presented here assume an f-plane approximation and neglect internal absorption and scattering of solar radiation. Over land, the calculations assume that the Bowen ratio is 1 and the albedo is 0.2.

An important feature of the numerical model is its capacity for two-way interactive grid nesting [see Clark and Farley (1984) for details]. Numerical solutions calculated on lower-resolution ''coarse'' grids are interpolated to higher-resolution ''fine'' grids using quadratic or cubic polynomials, depending on the variable being interpolated. These interpolated fields provide initial and boundary conditions for the higher-resolution domains. To complete the two-way interaction, the fine grid numerical solutions are averaged onto the coarse grid during the numerical integration. The averaging is defined in such a way as to conserve mass. Accordingly, the coarse grid solutions are affected by the fine grid solutions. Moreover, grid refinement and grid stretching are used in the vertical direction [see Clark and Hall (1996) for details].

All numerical experiments described are initialized with a single thermodynamic sounding taken at 0400 Local Standard Time (LST) 27 November 1995 during MCTEX (LST = UTC + 9.5 h). The radiosonde sounding on this day reached a height of 24 km. Above this height, isothermal and dry initial conditions are assumed. Figures 1a and 1b show the initial profiles of water vapor mixing ratio and potential temperature. The tropopause is located at a height of 16 km, and the moisture profile is conditionally unstable. Due to the high surface moisture, the lifting condensation level (LCL) is only 700 m above the surface. The level of free convection (LFC) ranges from 1.5 to 3.0 km, depending on the origin of the lifted parcel; likewise the level of neutral buoyancy (LNB) ranges from 10 to 14 km.

The intent of the study is not to simulate the details of the convection observed on 27 November but instead, to use this case to investigate how deep tropical convection excites gravity waves. With this in mind, an idealized wind profile was chosen for the present study. The wind profile retains the basic features of the observations, but its idealized form simplifies the analysis by reducing the number of active source terms, hence clarifying the source mechanism of the waves (see sections 4 and 5). Such an idealization allows the modeled stratospheric waves to be related unambiguously to certain properties of the wind profile. The initial wind profile is easterly at all heights and is shown in Fig. 1c. Above the boundary layer, the flow is 5 m s^{-1} throughout most of the troposphere. Just below the tropopause, the flow increases linearly through a 3-km deep layer to 10 m s^{-1} in the stratosphere. The main difference between the idealized wind profile and that taken during MCTEX is that the observed profile has a low-level easterly jet. This jet probably plays an important part in organizing the convection and determining the way in which the convective line propagates. These effects are neglected here.

Numerical cloud models commonly include a local-

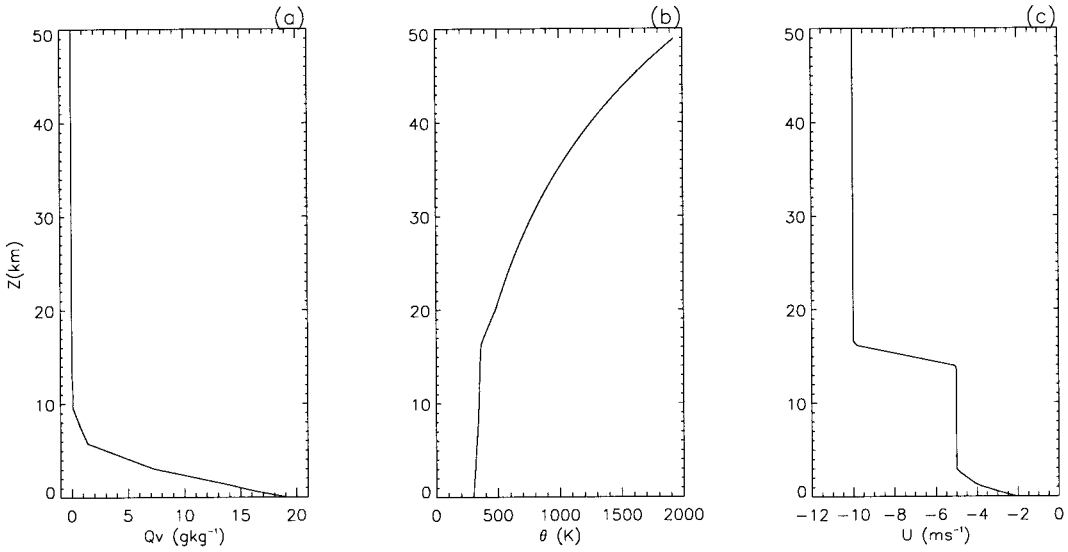


FIG. 1. Background fields used to initialize all model integrations. (a) Water vapor mixing ratio, Q_v (g kg^{-1}); (b) potential temperature, θ (K); and (c) zonal wind, U (m s^{-1}).

ized low-level buoyancy perturbation (often called a warm bubble) as part of the initial conditions. The warm bubble subsequently rises and initiates convection. However, the present study is concerned principally with the waves generated by deep convection, and therefore, it is critical not to introduce unrealistic gravity waves in the initial conditions. The gravity waves generated in the model must be unequivocally due to the clouds and not to an artifact of the initialization. For this reason, we chose to initiate the convection by explicitly modeling the island sea breeze similar to that observed during MCTEX, where deep convection develops once the circulation becomes strong enough to overcome the low-level CIN.

b. Experimental design

This paper describes the results of three numerical experiments. The primary difference between the experiments is the resolution and size of the model domains.

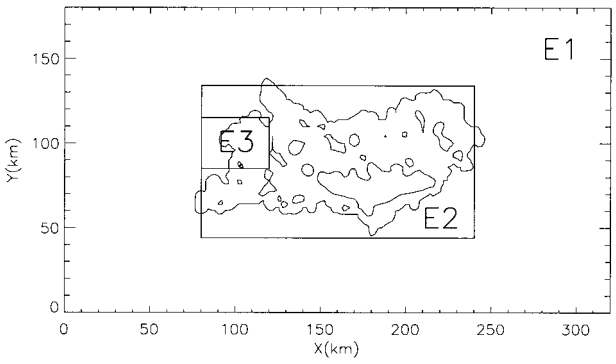


FIG. 2. Outline of model domains. Shown are domains initialized in experiments E1 to E3 (denoted Domains 1 to 3 in text). Also shown are topographic contours of 0 and 50 m.

Experiment E1 is the lowest resolution experiment considered here. It is designed to test the sensitivity of the modeled clouds, and the waves they generate, to changes in spatial resolution. Furthermore, the reduced computational cost of E1 allows the momentum flux to be calculated over a longer time interval than is possible with the other higher-resolution numerical experiments. Experiment E1 uses only one domain (Domain 1) surrounding the Tiwi Islands and is centered at (11.55°S, 130.56°E). Figure 2 shows the size and extent of the model domains relative to the Tiwi Islands. For simplicity, all surrounding topography not directly associated with the Tiwi Islands has been neglected. Domain 1 has 2-km horizontal resolution and extends vertically to 50 km. The vertical grid is stretched so that the grid spacing varies from 50 m at the surface to 800 m in the stratosphere (Fig. 3). The uppermost 10 km of the vertical grid includes a Rayleigh-friction sponge to absorb vertically propagating disturbances without reflection. The vertical grid is designed to resolve three important components of the flow field: the boundary layer convergence, the convective mixing in the troposphere, and the gravity waves in the stratosphere. Experiment E1 is initialized at 0700 LST and is integrated for 15 h until 2200 LST.

Experiment E2 is designed to resolve both the convective clouds and excited gravity waves. In this experiment, a second domain (Domain 2) is nested within Domain 1 (see Fig. 2), and the two domains interact in the way described earlier in section 2a. Domain 2 uses 1-km horizontal resolution with a vertical grid that extends to 19 km (see Fig. 3). Although both horizontal and vertical resolution remain unchanged in the upper

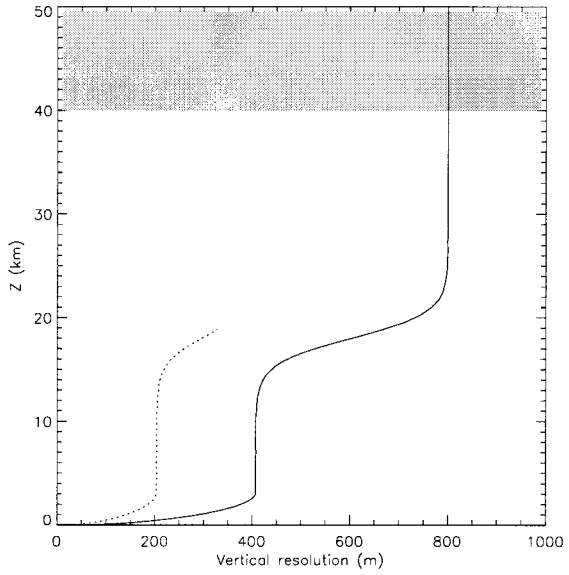


FIG. 3. Vertical grid spacings for Domain 1 (solid) and Domains 2 and 3 (dashed). The region of the sponge absorber is shaded.

stratosphere, the convective clouds, which are the source of the gravity waves, are better-resolved in E2 than in E1. This experiment begins at 1100 LST, taking its initial conditions from experiment E1, and is integrated until 1330 LST.

Experiment E3 uses three interactive grids to focus on one particular convective cell identified in E2. The purpose of E3 is to examine in detail how the cloud generates gravity waves. In this experiment, a third domain (Domain 3), with horizontal resolution of 0.5 km, is nested within Domain 2 (see Fig. 2) taking its initial conditions from E2 at 1230 LST. Experiment E3 is integrated until 1330 LST. The vertical grid remains the same as Domain 2 of E2. Domain 2 is, in turn, nested within Domain 1.

3. Numerical experiments

After the initialization of experiment E1, the islands are slowly heated, and a sea-breeze circulation develops. The dominant circulation appears as two zonally oriented convergence lines, which propagate inland in opposing directions from the northern and southern coastlines. By 1100 LST, the two convergence lines have penetrated inland approximately 20 km in the meridional direction, and at this time, the maximum meridional wind speed near the surface is approximately 5 m s^{-1}. The depth of the sea breeze is approximately 0.5 km, with weaker return flow aloft. The distance between the leading edges of the two sea breezes is approximately 15–20 km. Prior to collision of the sea breezes (approximately 1200 LST), the convergence causes sufficient vertical displacement of parcels to overcome the CIN, and the convection deepens rapidly.

a. Evolution of the convective clouds

Most of the modeled convection forms over the Tiwi Islands. Accordingly, our analysis of the convection focuses on experiment E2, which has an intermediate resolution domain covering the islands (Fig. 2). Prior to the initialization of experiment E2 at 1100 LST, shallow boundary layer convection is scattered over the islands, and the convection remains relatively shallow until about 1200 LST. After this time, the convection develops rapidly, and by 1300 LST, it reaches its maximum intensity. Thereafter, the convection decays while being advected by the easterly background flow.

Figure 4 shows the total cloud water (liquid water plus ice) in a horizontal cross section 10 km above mean sea level and depicts the temporal variation in the spatial position of deep convective cells. Note that those cells through which the cross section passes have nearly reached their LNB. At 1200 LST, a number of deep cells are located over the western half of the islands. Half an hour later, some of these cells arrange themselves in a north–south line, and by 1300 LST, the cells merge to form a convective line approximately 40 km long. The convective line continues to move westward, with a large cell developing at its northern end. In addition to the convective line, there are a number of other deep cells scattered across the islands.

A zonal cross section through $y = 100$ km illustrates the vertical structure of the convection and, in particular, the large cell at the northern end of the convective line (Fig. 5). At 1200 LST, the zonal cross section of vertical velocity and total moisture (water vapor, cloud water, and ice) shows a number of cumulus cells over the islands (Figs. 5a and 5b). These cells have cloud bases around 4 km and cloud tops between 9 and 12 km. Their updraft velocities are relatively small, and the cells are scattered with little organization.

By 1300 LST, there are a number of convective cells over the islands in various stages of development (Figs. 5c and 5d). A very deep cell is evident at approximately $(x, y) = (100, 100)$ km. This is the large cell at the northern end of the convective line (Fig. 4c), which will be examined in detail in section 3c. It is rapidly developing and has strong updrafts with vertical velocities on the order of 20–30 m s^{-1}. There are also two smaller cells upstream of the large cell. The updraft velocities in these two cells are less than 1 m s^{-1}, and they do not appear to be developing further. All cells in Figs. 5c and 5d are separated by approximately 35 km, and for this cross section, they are located at approximately $x = 100, 135, 170$ km, with another growing cell evident in the boundary layer at $x = 205$ km. It is unclear whether this spacing is caused by resonance between the clouds and the generated gravity waves (cf. Clark et al. 1986) or whether it is simply the natural scale set by the convection alone.

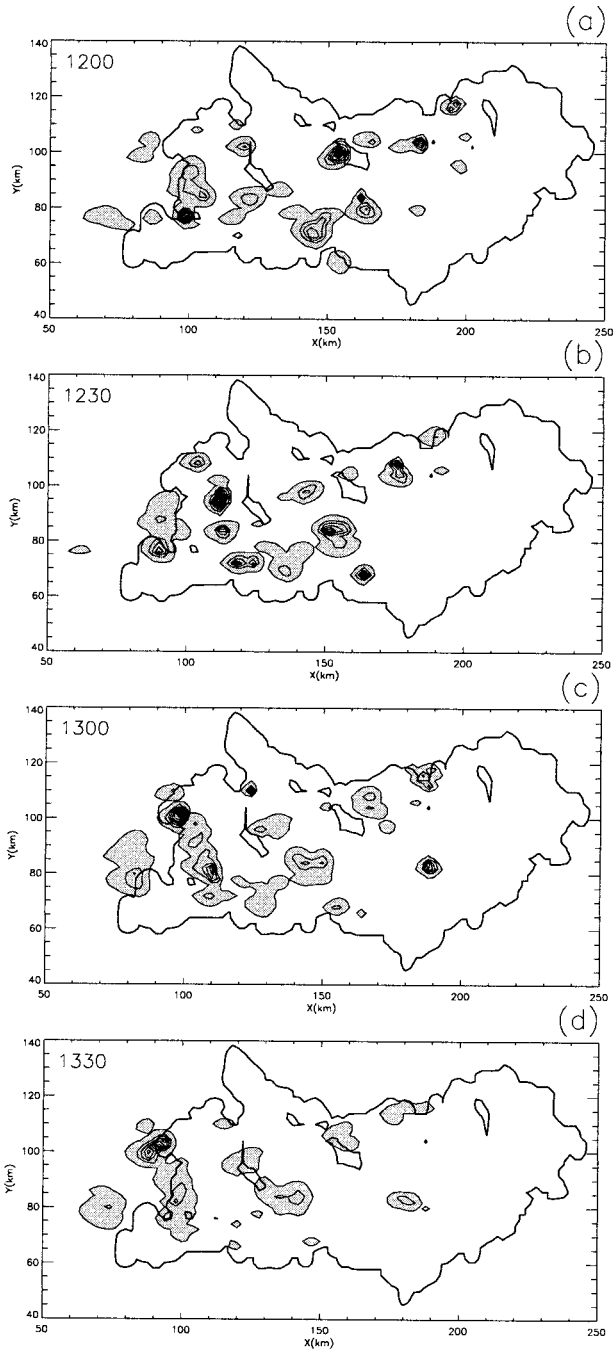


FIG. 4. Horizontal cross section of total cloud water (ice and liquid water) through $z = 10$ km of Domain 1 for experiment E2. (a) 1200 LST, (b) 1230 LST, (c) 1300 LST, and (d) 1330 LST. Contour intervals of 1 g kg^{-1} with the minimum level equal to 0.1 g kg^{-1}. Also shown is the zero topographic contour (thick).

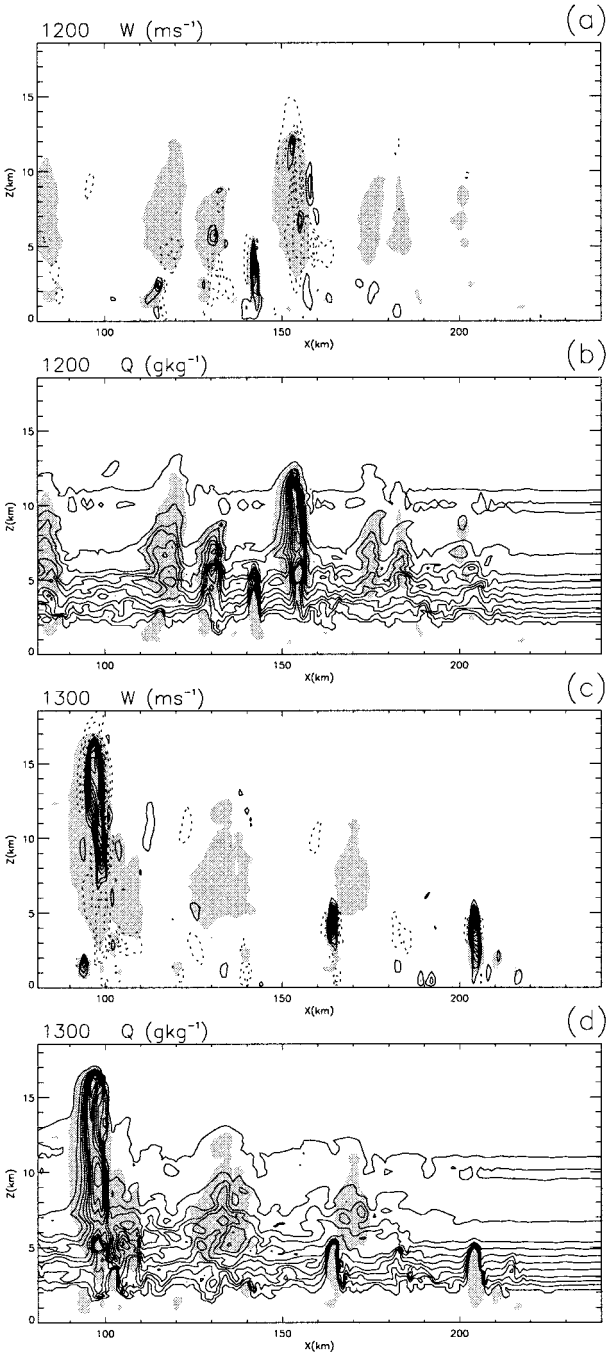


FIG. 5. Zonal cross section through $y = 100$ km for Domain 2 experiment E2. (a) Vertical velocity, w (1 m s^{-1} interval) at 1200 LST; (b) total moisture, Q (1 g kg^{-1} interval) at 1200 LST. (c) and (d) are the same as (a) and (b) except at 1300 LST. The shaded area in all plots is the 0.1 g kg^{-1} total cloud water contour.

b. Sensitivity to spatial resolution

The sensitivity of the modeled convection to changes in the resolution can be determined by comparing the numerical solutions from Domain 1 of experiments E1 and E2. The numerical solutions on Domain 1 from E2 should better represent the convection due to the two-way interactive nesting of the intermediate resolution domain (Domain 2) within it. A horizontal cross section of the cloudy air in the plane $z = 10$ km in E1 at 1300 LST (Fig. 6) can be compared directly to Fig. 4c to

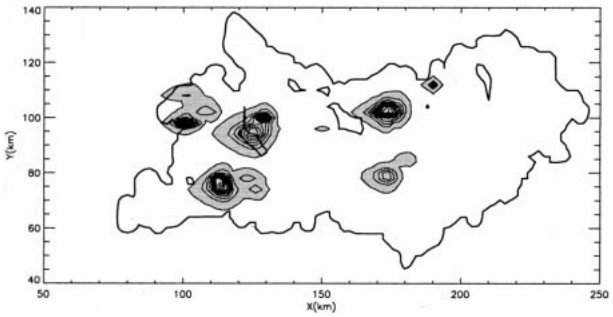


FIG. 6. Horizontal cross section of total cloud water (ice and liquid water) through $z = 10$ km of Domain 1 for experiment E1; for 1300 LST. Contour intervals of 1 g kg^{-1} with the minimum level equal to 0.1 g kg^{-1}. Also shown is the zero topographic contour (thick).

determine the sensitivity of the convection to spatial resolution. The comparison shows that the number and position of deep convective cells is sensitive to the spatial resolution, although the size and position of the largest cells are similar. Experiment E2 better resolves the boundary layer convergence and consequently more accurately represents those processes that initiate convection. Golding (1993) concluded that 4-km horizontal grid spacing was sufficient to qualitatively resolve these convective clusters over the Tiwi Islands. A similar conclusion was reached by Weisman et al. (1997), who suggested that 4-km horizontal grid spacing is required to resolve the mesoscale structures of squall lines. Although 4 km may be the upper limit for resolving the gross features of such convective clusters, it is not sufficient to resolve internal cloud structures such as convective updrafts, which may have spatial scales of only a few kilometers. It will be shown in section 4 that the individual updrafts play central roles in generating gravity waves and must be resolved when investigating the dynamics of the wave source.

c. Generated gravity waves

Three horizontal cross sections of the vertical velocity half an hour apart and at a height of 20 km are shown in Fig. 7. The cross sections lie in the stratosphere well above the top of the clouds, and velocity perturbations are produced by vertically propagating gravity waves. The wave fronts are almost circular and appear to be radiating outward (and upward) from highly localized sources. At 1230 LST, there is a strong, localized wave source at approximately $(x, y) = (90, 75)$ km (Fig. 7a). Half an hour later, another source develops at approximately $(x, y) = (100, 100)$ km (Fig. 7b), and by 1330 LST, this source dominates the wave generation (Fig. 7c). This source is advected downstream by the background flow, and at 1330 LST, is located at approximately $(x, y) = (90, 100)$ km. This source corresponds to the large cell at the northern end of the convective line discussed in section 3a. Throughout the integration, the wave sources remain highly localized, although their number and position are highly variable. These wave sources are strongly correlated with deep individual convective cells. Furthermore, the temporal variability suggests that clouds do not generate waves throughout their entire evolution, but instead, gravity waves are generated during a particular part of clouds' life cycles.

At a height of 20 km, the wave field is incoherent. According to linear wave theory, the vertical component of the group velocity is a function of the horizontal and vertical wavelength. Consequently, a wave packet will disperse into its constituent waves as it propagates upward. For this reason, the wave field at 40 km appears more coherent than at lower levels (Fig. 8). Crook et al. (1991) noted this effect for gravity waves in the lower troposphere, and Prusa et al. (1996) found that waves generated by a localized source at the tropopause were dominated by one wavelength after they had propagated vertically into the middle atmosphere. The increased coherence of the waves underscores the highly localized and variable nature of the source. The wave signature, and hence the apparent source of the waves, lags the corresponding signature at 20 km due to the time taken for the waves to propagate vertically.

Recently, Dewan et al. (1998) analyzed gravity waves in the satellite observations from the Midcourse Space Experiment. The density and temperature fluctuations, which were measured in the upper stratosphere (at approximately 40 km), showed wave fronts arranged in concentric circles centered over a large thunderstorm. These observations are remarkably similar to the cross sections shown in Fig. 8 and lend support to the numerical solutions presented here.

We consider now zonal and meridional cross sections of the vertical velocity at 1300 LST. The two cross sections intersect at $(x, y) = (100, 100)$ km (Figs. 9a and 9b). The zonal cross section (Fig. 9a) shows a wave packet centered over the group of convective cells and an upstream bias to the waves. In contrast, the gravity wave field in the meridional cross section (Fig. 9b) is approximately symmetric about the source. The amplitude of the waves is small, approximately 0.3 m s^{-1}, and their wavelength spectra (in the stratosphere) have peaks in the horizontal (both zonal and meridional) and vertical at approximately 17 and 5 km, respectively. These will be determined more accurately later in this section.

The zonal velocity and the buoyancy are shown in Figs. 9c and 9d. Upstream from the source, the horizontal velocity exhibits wave-like perturbations similar to those in the vertical velocity, although the wave signature in zonal velocity of the downstream propagating waves is more pronounced than the upstream propagating waves. The reason is that the phase lines of the downstream propagating waves are more horizontal. As the phase lines become more horizontal, the relative amplitude of the horizontal velocity increases and vertical velocity decreases. This is important because an analysis restricted to vertical velocity signatures may

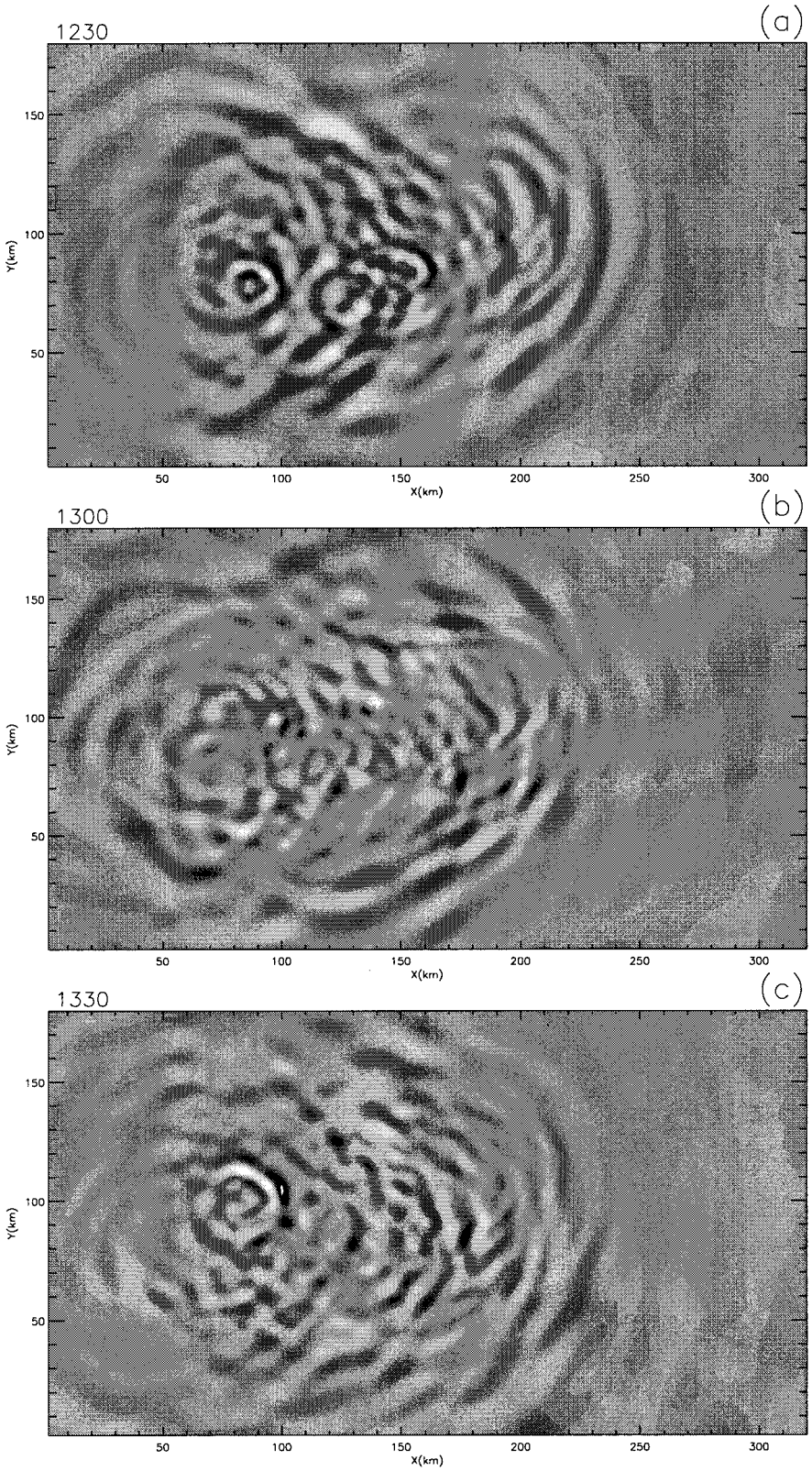


FIG. 7. Horizontal cross section of vertical velocity of Domain 1 experiment E2. (a) 1230 LST, (b) 1300 LST, and (c) 1330 LST for $z = 20$ km.

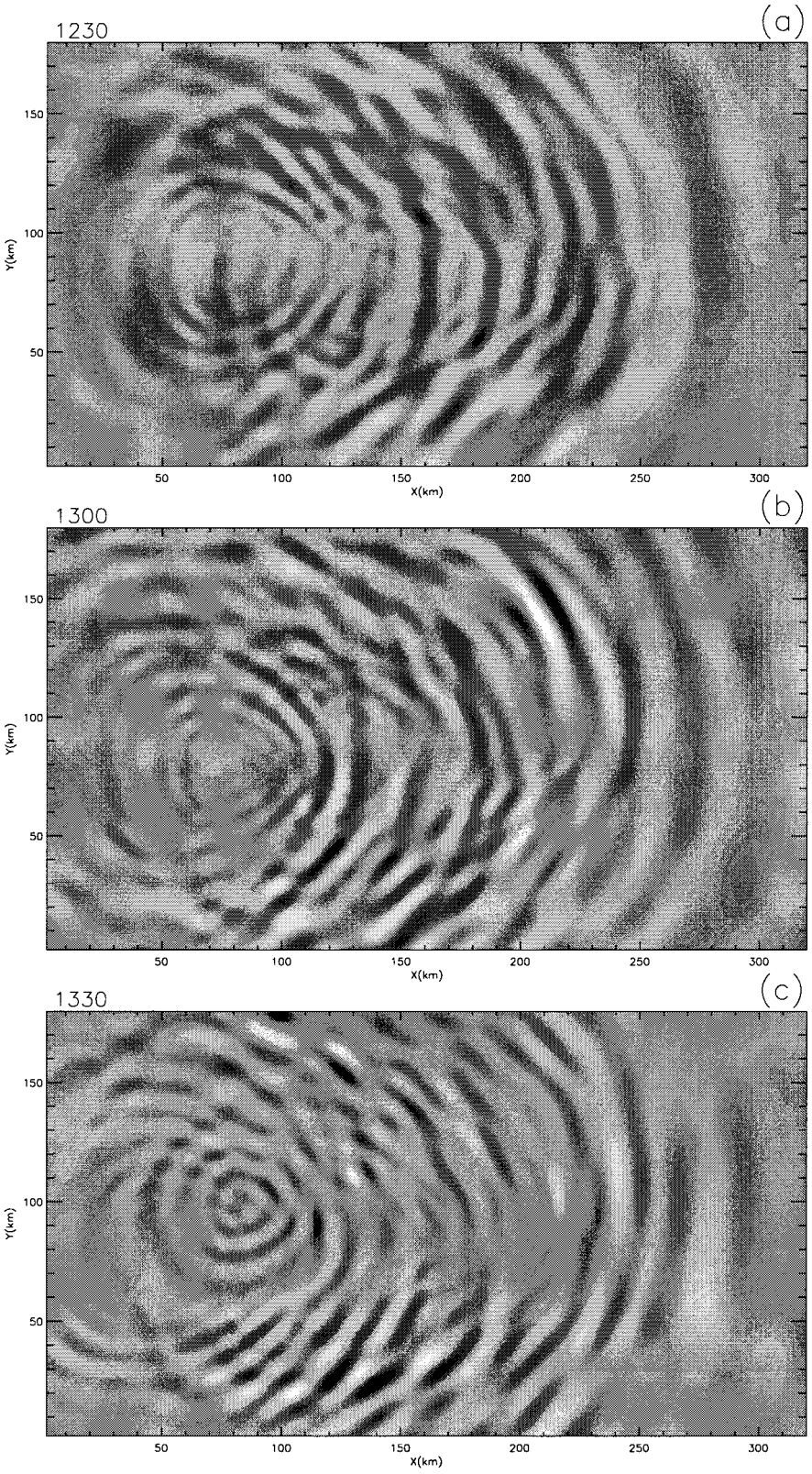


FIG. 8. Same as Fig. 7 except for $z = 40$ km.

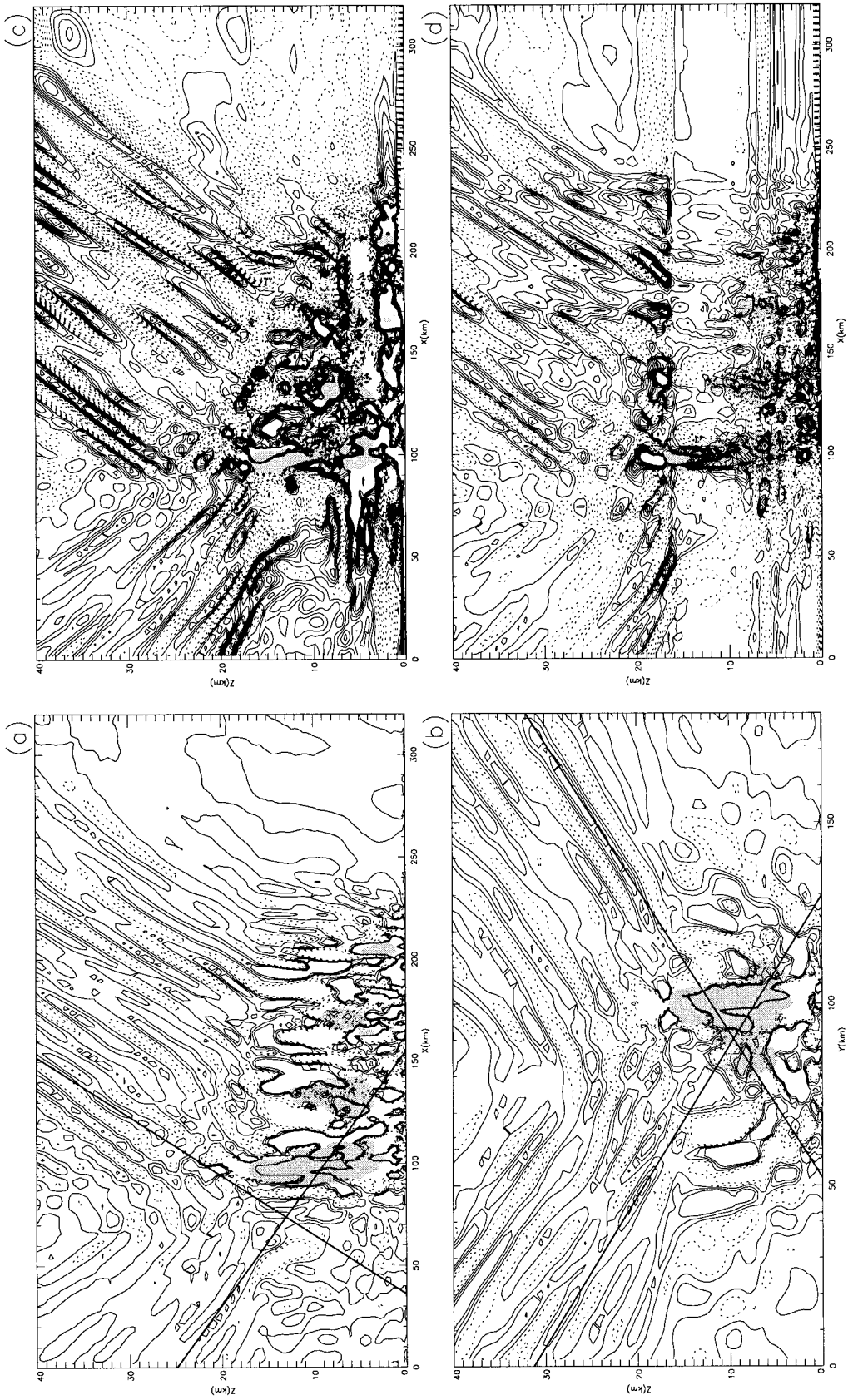


FIG. 9. (a) Zonal cross section of vertical velocity through $y = 100$ km for Domain 1 experiment E2. (b) Meridional cross section of vertical velocity through $x = 100$ km for Domain 1 experiment E2. (c) and (d) show perturbation zonal velocity and dry buoyancy, respectively, for the same zonal cross section as (a). All plots are valid at 1300 LST and feature the 0.1 g kg^{-1} total cloud water contour (shaded). Negative contours are dashed with velocity contours having a 0.1 m s^{-1} interval and buoyancy contours having a 5×10^{-3} m s^{-2} interval. (a) and (b) also have solid lines representing the angles of some phase lines for later analysis. Note that (b) has a different horizontal scale from (a), (c), and (d).

miss the downstream propagating waves. As buoyancy perturbations reflect parcel displacements, the wave signatures in the buoyancy field are in quadrature with the vertical velocity. These ideas are discussed further in section 5a.

Power spectra of the vertical velocity for Domain 1 E2 are calculated to determine more accurately the dominant horizontal and vertical wavelengths in the stratosphere. Two-dimensional m–k (zonal) and m–l (meridional) spectra[2] are shown in Fig. 10. These spectra are calculated using the zonal and meridional cross sections of vertical velocity shown in Fig. 9 for heights of $20 < z < 40$ km. These spectra are smoothed over three adjacent wavenumber bins, squared, and then normalized. Note also that in these spectra, $m < 0$, which corresponds to negative vertical phase speed, and therefore $k > 0$ represents westerly propagating waves and $k < 0$ represents easterly propagating waves. Both of these spectra (Figs. 10a and 10b) have maximum power at a horizontal wavenumber of approximately 3.7×10^{-4} m^{-1} (wavelength of 17 km) and vertical wavenumber of approximately 1.3×10^{-3} m^{-1} (wavelength of approximately 5 km). The zonal spectrum is asymmetric, about $k = 0$, due to the stratospheric flow. Upstream propagating waves have smaller wavenumbers and larger wave signatures. This is due to wave refraction and will be discussed in more detail later in the section. The meridional spectrum is approximately symmetric in terms of the location of the vertical wavenumber maximum, except the southerly propagating waves ($l > 0$) appear to have more power. This is probably due to the focus of the gravity waves in the upper stratosphere being south of the center of the domain, thus there are more southerly propagating waves in the domain. (Note that the position of the focus of the waves in the upper stratosphere lags the position of the source in the troposphere due to the finite vertical group velocity of the waves.)

The zonal cross section of vertical velocity (Fig. 9a) shows that the gravity waves are refracted strongly as they propagate from the troposphere to the stratosphere. In the troposphere, the phase lines are oriented almost vertically, implying that the waves are trapped or evanescent or that the vertical wavelength is very long. One should bear in mind that in the troposphere, it is difficult, and probably not dynamically meaningful, to distinguish convective updrafts from gravity waves. In the stratosphere, the phase lines are oriented at an angle to the vertical, implying that the waves are propagating vertically. This refraction of the waves at the tropopause is due to a change in static stability and will be discussed in more detail in section 5a. The mechanism that selects the horizontal wavelength of the gravity waves is discussed in section 5b. The convective cells in the troposphere appear to be separated by a distance equal to twice the dominant horizontal wavelength (approximately 17 km) in the stratosphere (Fig. 9a), and there is a signature of this in the power spectrum (Fig. 10a). However, as noted earlier in section 3a, it remains unclear at present what role (if any) the gravity waves play in organizing the convection. This is a topic for future research.

d. Momentum fluxes

The vertical flux of horizontal momentum (or simply momentum flux) is an important quantity when considering how gravity waves affect the large-scale circulation. The horizontally averaged (zonal) momentum flux is defined as $(\overline{\rho}uw)$, where $\overline{\rho}$ is the background density; u and w are the zonal and vertical perturbations from the background flow ($U(z)$, 0, 0), respectively; and $\overline{(\ )}$ represents a simple horizontal average. Quantifying the momentum flux due to convectively generated gravity waves is crucial for the development and verification of GCM parameterizations.

The evolution of the horizontally averaged (zonal) momentum flux for Domain 1 from experiment E1 is shown in Fig. 11. The momentum flux is largest in the troposphere, and its evolution strongly reflects the life cycle of the convection over the island. Prior to approximately 1000 LST, the momentum flux is associated with the well-mixed boundary layer. After this time, the tropospheric momentum flux is associated mainly with transport by the convective clouds. Between approximately 1000 and 1200 LST, the convection deepens slowly, reaching a height of about 9 km. At 1200 LST, the sea-breeze convergence is sufficiently strong to overcome the CIN, and the convection deepens rapidly to the tropopause. The convection remains deep until approximately 1400 LST, after which it begins to decay. By 2200 LST, vertical mixing is confined to heights below the middle troposphere.

In the stratosphere, the momentum flux is associated principally with vertically propagating gravity waves. Early in the model run, the stratospheric momentum flux is produced by small-amplitude gravity waves generated by the sea breeze. However, after about 1200 LST, the convection deepens rapidly, and the stratospheric momentum flux increases sharply, reaching a maximum at approximately 1400 LST. The upper-stratospheric momentum flux lags that in the lower stratosphere because it takes time for the waves to reach these heights. The stratospheric momentum flux decreases during the afternoon to nearly zero by the end of the model integration at 2200 LST.

The model calculation employs open lateral boundaries, allowing waves to propagate out of the model domain without significant reflection. Some waves propagate through the lateral boundaries of the model and therefore do not contribute to the momentum flux shown

[2] Note that k, l, and m are the zonal, meridional, and vertical wavenumbers, respectively.

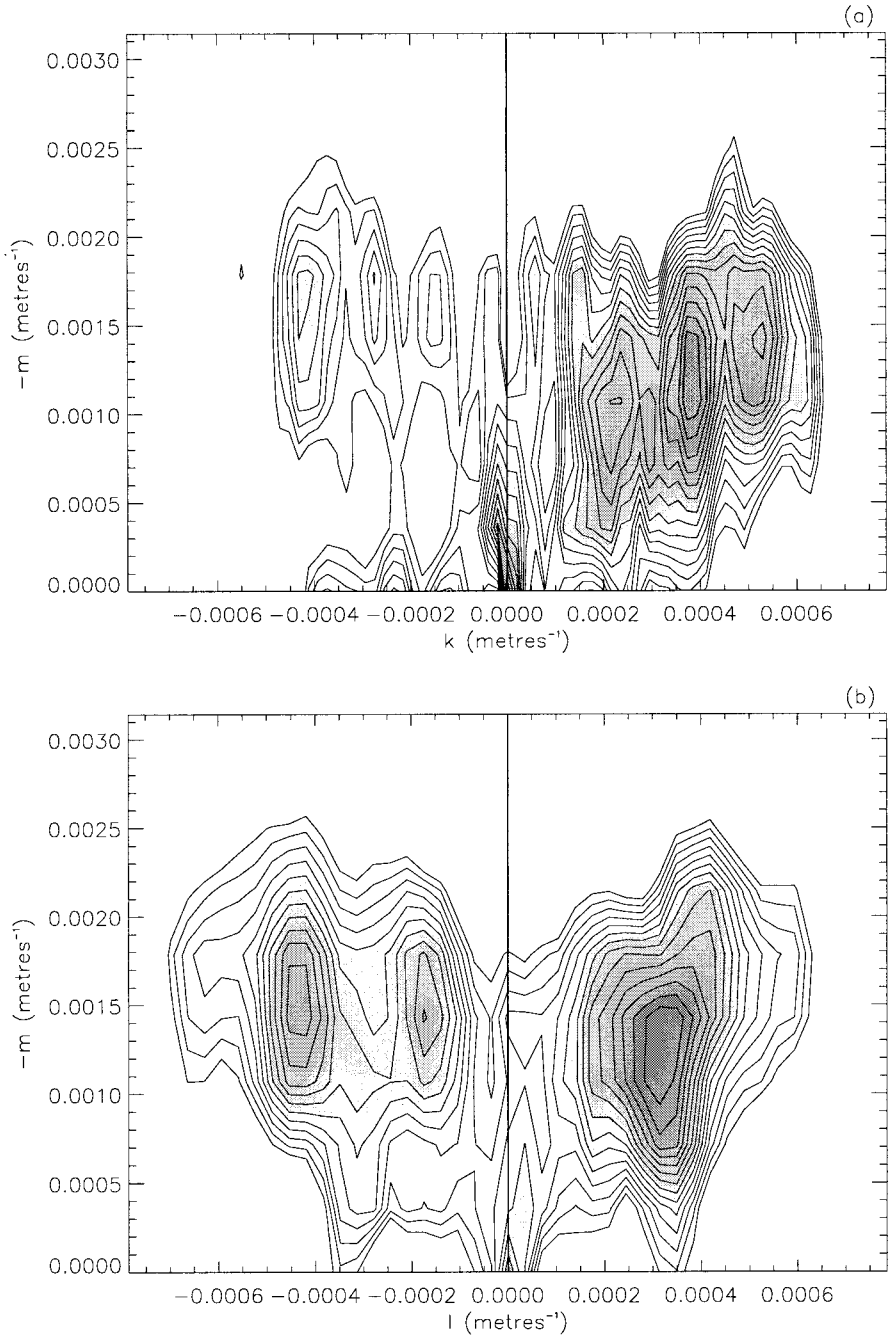


FIG. 10. Normalized two-dimensional power spectra of vertical velocity at 1300 LST. (a) *m–k* (zonal) spectrum calculated at $y = 100$ km and $20 < z < 40$ km. (b) *m–l* (meridional) spectrum calculated at $x = 100$ km and $20 < z < 40$ km. To aid interpretation, the spectra are normalized with logarithmic contours, and only the largest 1.5 orders of magnitude are shown ($10^{-1.5}$–1 in 15 contours).

in Fig. 11. Hence, the calculated momentum flux should be considered a lower bound on the magnitude of the total momentum flux. Also, such calculations of the momentum flux are sensitive to the size of the region over which the average is taken, particularly if the wave source is not in the center of the averaging region. Figure 12 shows the meridional average of the momentum flux for Domain 1 in experiment E2 at 1300 LST. The pattern of momentum flux is highly variable in space because the individual convective clouds, which generate the

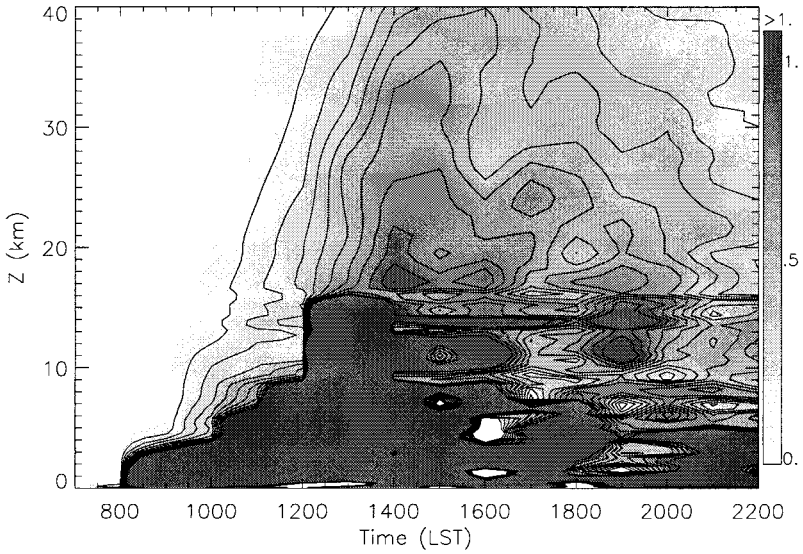


FIG. 11. Temporal evolution of the horizontally averaged (zonal) momentum flux for experiment E1 ($\times 10^{-3}$ Nm^{-2}); contour interval is 1×10^{-4} Nm^{-2}, with the zero contour omitted.

waves, are highly transient. Upstream from the wave source, the momentum flux is generally positive, while downstream from the source, it is generally negative.

The dominant wave-generating cloud is not in the center of the domain at 1300 LST (Fig. 12). Instead, the main wave source is located at approximately $x =$ 100 km. The relative contribution of the upstream and downstream propagating waves to the momentum flux can be estimated by averaging the momentum flux over two different regions. The first region is between $x =$ 0 km and $x = 100$ km and principally comprises those waves propagating downstream from the dominant source. The second region lies between $x = 100$ km and $x = 200$ km and mainly includes those waves propagating upstream from the dominant source. These averages and their sum are shown in Fig. 13a. The upstream component of the momentum flux is largest, and the downstream component acts to reduce the total momentum flux only in the lower stratosphere. This is a rather crude method of separating the upstream com-

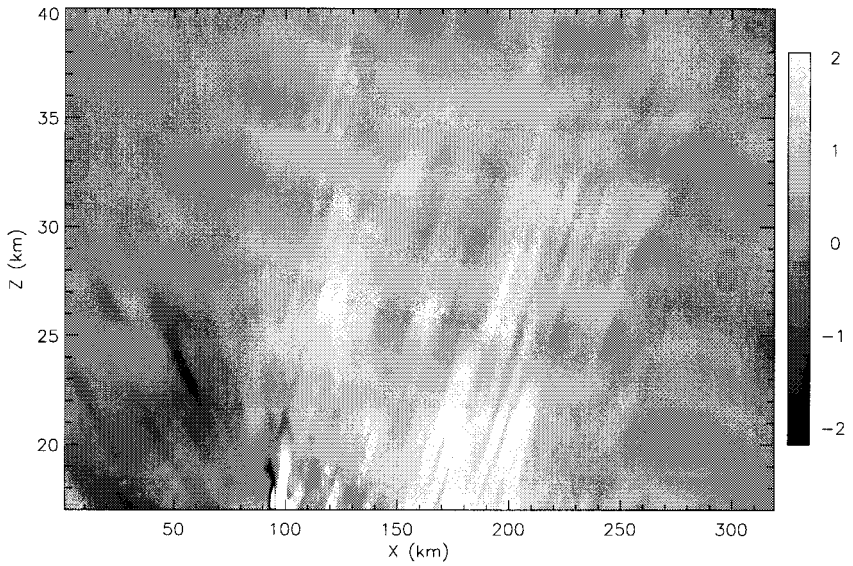


FIG. 12. Meridional average of the momentum flux for Domain 1 experiment E2 at 1300 LST ($\times 10^{-3}$ Nm^{-2}).

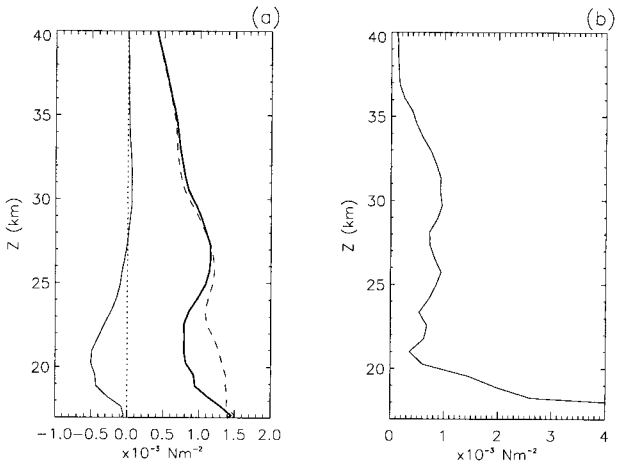


FIG. 13. Profiles of the momentum flux for experiment E2 at 1300 LST. (a) The horizontal average is calculated over $0 \leq x \leq 100$ km (solid), $100 \leq x \leq 200$ km (dashed), and $0 \leq x \leq 200$ km (bold) for $0 \leq y \leq 180$ km. (b) The horizontal average is calculated in the stratosphere over the region of Domain 3 in E3.

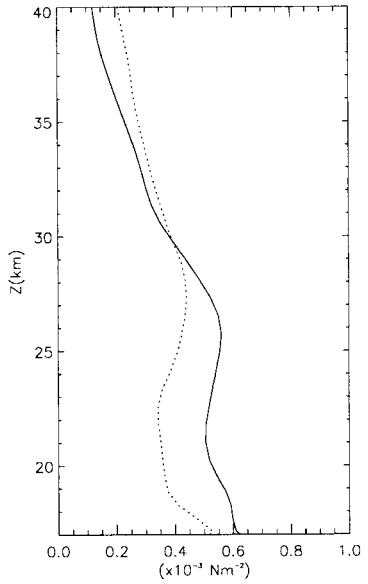


FIG. 14. Vertical profile of the horizontally averaged (zonal) momentum flux at 1300 LST for Domain 1 experiment E1 (solid) and experiment E2 (dashed).

ponents of the momentum flux, in particular, considering that the wave source is not a true point source but rather a cluster of sources.

The momentum flux generated by an individual cloud is estimated by taking a horizontal average for E2 in the region above Domain 3 of E3 (described in more detail in the next section). This average is calculated at 1300 LST and is shown in Fig. 13b. At this time, the convective cloud chosen is the strongest wave source in the model domain, and this strength is reflected in the large momentum flux in the lower stratosphere. Such a localized average of the momentum flux can be compared to localized observations of gravity wave momentum fluxes taken close to convective clouds (for example from aircraft or radiosondes). However, this profile has little value in terms of the momentum budget of the middle atmosphere because it does not incorporate the reduction in momentum flux due to radial propagation of the waves. As time continues, the magnitude of the momentum flux reduces as the waves disperse and radiate outward. Eventually, the pattern of momentum flux resembles one of the pulses in Fig. 12.

Experiment E1 has been used to illustrate the temporal evolution of the momentum flux. However, as shown earlier, smaller-scale convective structures are sensitive to the numerical resolution. To determine the sensitivity of the stratospheric momentum flux to tropospheric resolution, a profile of momentum flux in the stratosphere from E1 at 1300 LST is compared to a profile at the same time from E2 (Fig. 14). Although the two profiles have similar numerical values and similar vertical gradients, they show important differences. These differences can be explained by the slightly different positions and stages of development of the convective cells in the two experiments. For example, any change in the position of the clouds will change how much momentum is transported through the lateral boundaries by the waves, which in turn, will change the horizontally averaged momentum flux. Unfortunately, the computational cost of these large numerical calculations made it impracticable to explore resolution sensitivity further.

e. High-resolution convection and the wave-generation mechanism

In section 3c, experiment E2 was used to identify a dominant wave source at approximately $(x, y) = (100, 100)$ km. In order to better resolve the structure of this source, a higher-resolution domain surrounding this region is initialized at 1230 LST. This domain is the basis of experiment E3. However, when higher-resolution domains are initialized from lower-resolution outer domains, the modeled atmosphere takes some time to adjust. Typically, this adjustment time is on the order of the buoyancy period, $2\pi/N$, which is approximately 10 min in the troposphere. The maximum convective updraft at initialization is about 30 m s^{-1}, and therefore in 10 min, a parcel moving at this velocity would traverse the depth of the troposphere. Thus, by about 1240 LST, the flow has adjusted to the finer computational grid.

Zonal cross sections of vertical velocity (Figs. 15a–d), potential temperature (Figs. 16a–d), and total moisture (Figs. 17a–d) depict the evolution of the large cell at the northern end of the convective line, discussed earlier in section 3a. This cell is the strongest wave-producing cloud in the model runs. At 1240 LST, there is a strong updraft at approximately $x = 100$ km, with maximum vertical velocity of approximately 40 m s^{-1} (Fig 15a).

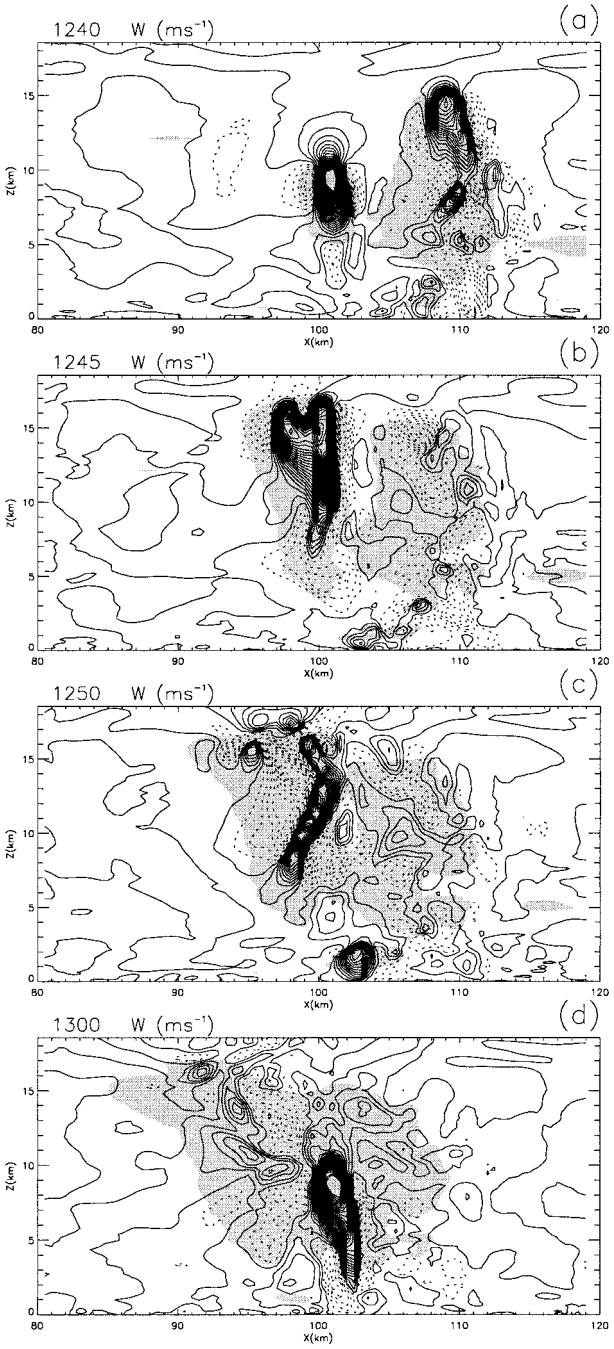


FIG. 15. Zonal cross section of vertical velocity, w, through y = 98 km of Domain 3 experiment E3 for (a) 1240 LST, (b) 1245 LST, (c) 1250 LST, and (d) 1300 LST. Contour intervals are 1 m s^{-1} with negative values dashed. Shaded is the 0.1 g kg^{-1} total cloud water contour.

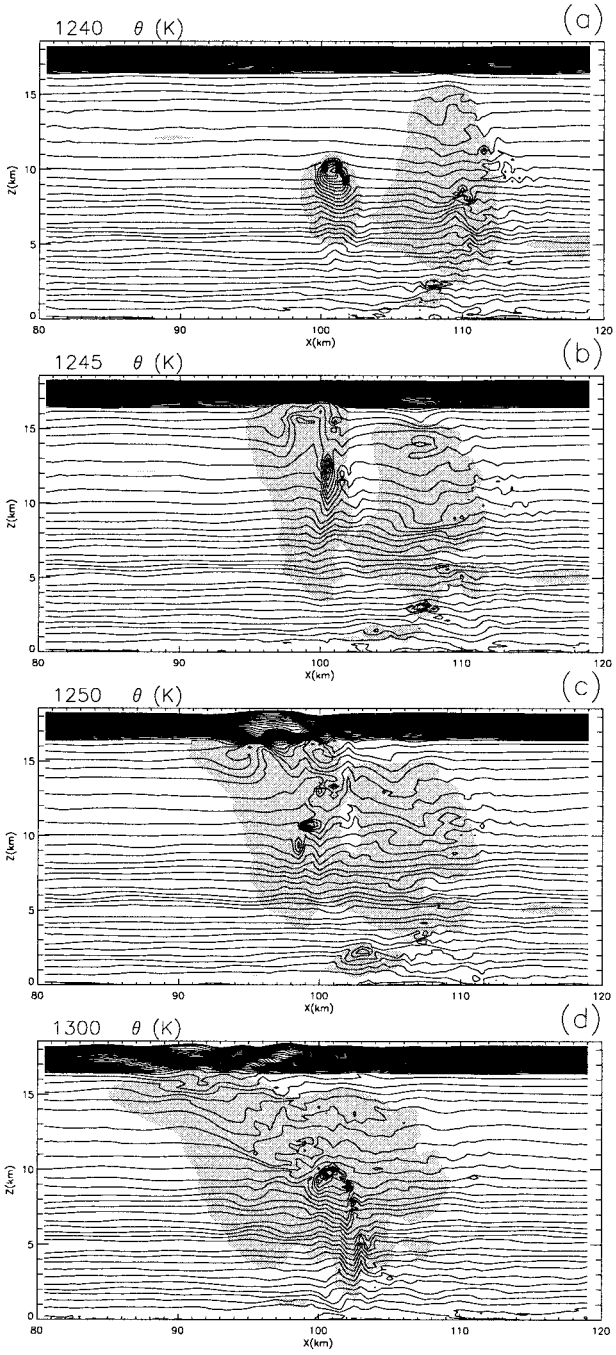


FIG. 16. Same as Fig. 15 except potential temperature, θ, with 2 K intervals.

At this time, the updraft has not mixed moisture above z = 12 km (Fig. 17a). However, the cell is rapidly developing, and there is a well-defined unstable bubble of potentially warm air at the top of the cloud (Fig. 16a). Therefore, further vertical mixing can be expected. Weak subsidence on the edges of the updraft indicates that stable air is being displaced laterally and downward around the vertically moving column of air. There is another deep cell to the east of the updraft at approximately x = 110 km. This cell is in its later stages of development and will not be considered in this analysis.

The updraft has overshot its LNB (Fig. 16b) at 1245 LST and has mixed air vertically to the tropopause. The updraft has penetrated the stratosphere, with cloudy air

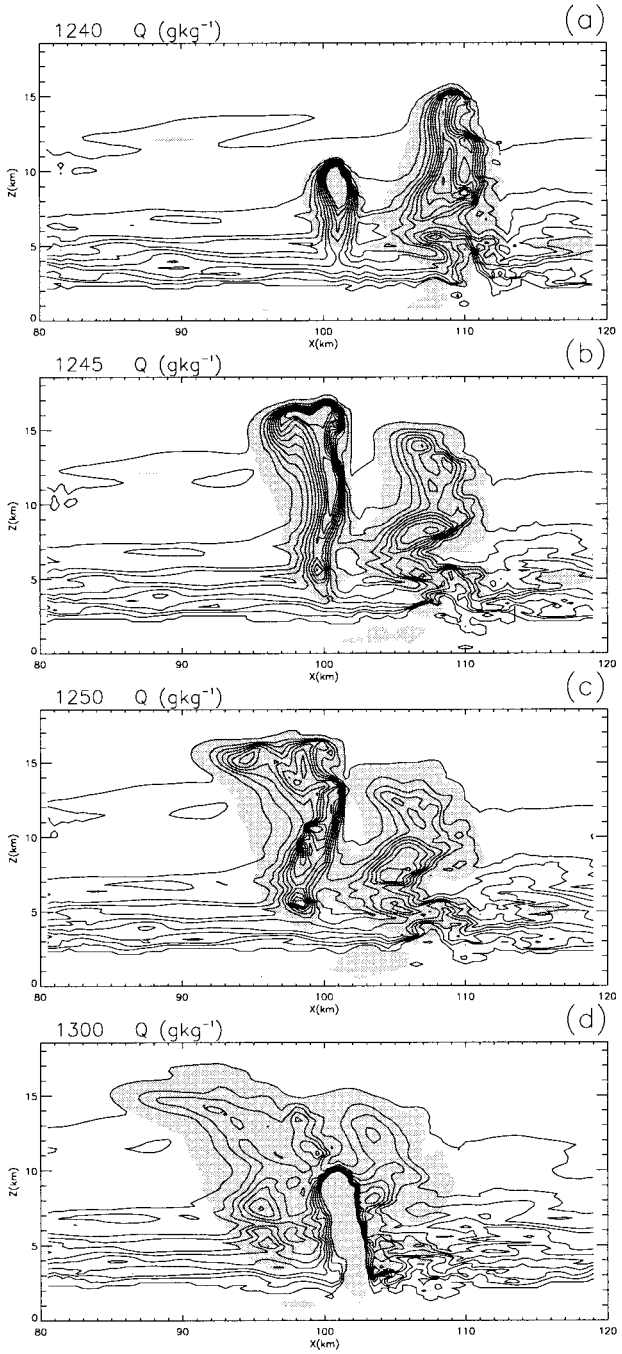


FIG. 17. Same as Fig. 15 except total moisture, Q, with 1 g kg^{-1} intervals.

evident in the lowest 1 km (Fig. 17b). To some extent, the tropopause acts like a lid, with stable air above the updraft being forced laterally and downward at 3–4 m s^{-1} by the upward moving mass of air. The displaced air was neutrally buoyant just below the tropopause, and after being displaced downward, has become positively buoyant. There exists an upper-tropospheric bulge in the cloudy air, indicating that cloudy air is forced outward while being decelerated. The maximum vertical velocity of the updraft is approximately 16 m s^{-1} (Fig. 15b), and Fig. 16b shows that the LNB of this air mass is approximately $z = 13$ km. Consequently, the net force on the upward-moving air above $z = 13$ km will be downward. In fact, a small portion of this updraft has already reversed its direction at $(x, z) = (99, 16)$ km.

At 1250 LST, the upward moving air of the main updraft has reversed due to over shooting its LNB, and a large downdraft exists (Fig. 15c) in the same region as the strong updraft at 1245 LST. This downdraft has a maximum value of approximately 11 m s^{-1} near the LNB of this air ($z = 13$ km). The downward displaced air of Fig. 15b has now reversed direction forming two weak updrafts (approximately 4 m s^{-1}), with the maximum value located at approximately $z = 15$ km. Located laterally adjacent to these two weak updrafts are two downdrafts of equivalent strength and spatial scale.

At 1300 LST, the dominant motions observed during the evolution of the convective cells are no longer present. The convective cells now exist as a large mass of cloudy air with a pronounced anvil. This large mass begins to merge with adjacent masses of cloudy air while being advected by the easterly background flow.

During the development of the updraft, two separate modes of oscillation are evident. First, there is the very strong oscillatory motion of the updraft about its LNB (approximately $z = 13$ km). This oscillation has approximate vertical orientation and has a vertical velocity amplitude of 10–15 m s^{-1}. Second, an oscillation occurs when stable air is displaced by the rising updraft at the tropopause. This displaced air oscillates about its original height (approximately $z = 15$ km) with a vertical velocity amplitude of 4 m s^{-1}. This secondary oscillation is of a much smaller spatial scale than the main updraft and appears to be generating external wave modes of equivalent amplitude, which propagate horizontally away from the updraft. Furthermore, the main oscillation has generated a vertical velocity component directly above it in the stratosphere. This region of vertical velocity appears to be the source of stratospheric gravity waves.

4. Analysis of the gravity wave source

The previous section focused on the structure and evolution of the convection and the stratospheric wave field and showed that the individual convective cells generated waves. The present section seeks to answer the question: how do the convective cells generate waves? Answering this question is central to understanding how convection modifies its environment, and is a prerequisite to developing accurate and physically realistic parameterizations of convectively generated waves.

a. The source tensor

The following analysis is motivated by Lighthill's (1952) theory for the aerodynamic generation of sound and by Ford's (1994) adaptation of Lighthill's theory to inertia-gravity waves in rotating shallow water. In linear gravity wave theory, all nonlinear terms in the governing equations are neglected. This assumption is generally valid for small-amplitude waves far from their source but may not be satisfied close to their source. Indeed, as we shall see, these nonlinear terms can be interpreted as the source of linear motions, including gravity waves. In the remainder of this section, these ideas are applied to convective clouds, which are of course highly nonlinear systems.

We begin by separating the velocity components into a vertically varying background flow plus a (finite amplitude) perturbation,

$$\mathbf{u} = (U(z) + u, V(z) + v, w),$$

and rewriting the governing equations with all nonlinear terms on the right-hand side. The resulting expressions are

$$D_t u + U_z w + \phi_x = F_u, \tag{1}$$

$$D_t v + V_z w + \phi_y = F_v, \tag{2}$$

$$D_t w + \phi_z - b = F_w, \quad \text{and} \tag{3}$$

$$D_t b + wN^2(z) = \dot{b} + F_b = F_B, \tag{4}$$

where

$$F_u = -\left[(u^2)_x + (uv)_y + \frac{1}{\overline{\rho}}(\overline{\rho}uw)_z\right],$$

$$F_v = -\left[(uv)_x + (v^2)_y + \frac{1}{\overline{\rho}}(\overline{\rho}vw)_z\right],$$

$$F_w = -\left[(uw)_x + (vw)_y + \frac{1}{\overline{\rho}}(\overline{\rho}w^2)_z\right],$$

$$F_b = -\left[(ub)_x + (vb)_y + \frac{1}{\overline{\rho}}(\overline{\rho}wb)_z + \frac{wb}{g}N^2(z)\right].$$

Here, u, v, and w are the perturbation velocity components in the three coordinate directions x, y, z, respectively. Note that these perturbations need not be small-amplitude. The variable b is the (dry) buoyancy defined as $b = g(\theta - \overline{\theta})/\overline{\theta}$, where θ is the potential temperature and $\overline{\theta}$ is its background profile. The variable N^2 is the Brunt–Väisälä frequency, defined as $(g/\overline{\theta})(d\overline{\theta}/dz)$, and $\phi = p/\overline{\rho}$, where p is the perturbation pressure and $\overline{\rho}$ is the background density. The term D_t is the linear operator representing the material derivative following the background flow and is defined as $D_t = \partial_t + U(z)\partial_x + V(z)\partial_y$. The term $\dot{b}$ is the total rate of change of (dry) buoyancy due to diabatic processes. The variables F_u, F_v, F_w, and F_b are written in flux form and physically represent the advection of the perturbation velocity and perturbation buoyancy by the perturbation wind fields. The term F_B is simply the sum of diabatic and nonlinear buoyancy terms. The turbulence closure terms in the governing equations have been neglected for simplicity, since in all cases presented later, these terms are at least one order of magnitude smaller than the nonlinear forcing terms. Similarly, coriolis effects are small and have been neglected in the analysis.

The individual flux terms, which comprise the nonlinear right-hand side of Eqs. (1)–(4) (e.g., $(u^2)_x$, $(uv)_y$, and $(1/\overline{\rho})(\overline{\rho}uw)_z$ combine to form F_u), are all equally important. This is probably due to the isotropic nature of deep convection. However, when the convection is most intense, the terms that involve vertical velocity are generally larger in magnitude than those that do not. The one exception is the term $(wb/g)N^2$ from F_b, which is approximately one order of magnitude smaller than the other terms that comprise F_b.

The continuity equation is

$$u_x + v_y = -\frac{1}{\overline{\rho}}(\overline{\rho}w)_z = -w_z + \frac{1}{H_s}w, \tag{5}$$

where $H_s(z) = -\overline{\rho}(\partial\overline{\rho}/\partial z)^{-1}$ is the density scale height. Combining Eqs. (1)–(5) gives a single, forced linear wave equation for w,

$$D_{tt}\left[\nabla^2 w - \left(\frac{w}{H_s}\right)_z\right] - D_t\left[\mathbf{U}_z \cdot \boldsymbol{\nabla}\left(\frac{w}{H_s}\right) + \mathbf{U}_{zz} \cdot \boldsymbol{\nabla} w\right] + N^2\nabla_h^2 w = \mathcal{F}. \tag{6}$$

Note that the homogeneous form of Eq. (6) (in the limit as $H_s \to \infty$) is identical to Eq. (8) from Bretherton (1966). The source function $\mathcal{F}$ can be concisely written by a simple change in notation. Let

$$(x, y, z) = (x_1, x_2, x_3),$$

$$(u, v, w) = (u_1, u_2, u_3),$$

$$(F_u, F_v, F_w) = (F_1, F_2, F_3),$$

$$(U, V, 0) = (U_1, U_2, U_3).$$

Then

$$\mathcal{F} = \frac{\partial^2 T_{ij}}{\partial x_i \partial x_j}, \tag{7}$$

where

$$T_{ij} = \delta_{ij}(D_t F_3 + F_B) - \delta_{i3}\left(D_t F_j + \frac{dU_j}{dz}F_i + \delta_{ij}F_B\right) + \frac{dU_j}{dz}F_i. \tag{8}$$

Alternatively, the components of the tensor can be written in matrix form as

$$T_{ij} = \begin{bmatrix} (D_t F_w + F_B + U_z F_u) & \frac{1}{2}(U_z F_v + V_z F_u) & -\frac{1}{2} D_t F_u \\ \frac{1}{2}(U_z F_v + V_z F_u) & (D_t F_w + F_B + V_z F_u) & -\frac{1}{2} D_t F_v \\ -\frac{1}{2} D_t F_u & -\frac{1}{2} D_t F_v & 0 \end{bmatrix}.$$

We will call T_{ij} the source tensor. It is second order, symmetric, and analogous to Lighthill's (1952) radiation tensor. Equations (6), (7), and (8) imply that the generation of linear motions, including gravity waves, can be associated with the source tensor. In general, the perturbation velocities and the buoyancy will be relatively large within a convective cloud, and consequently, the source tensor (and hence $\mathcal{F}$) will be important within the cloud. However, away from convective clouds, the source tensor will be small, provided the waves remain small amplitude, in which case the forcing is negligible. Thus, the forcing will appear as a compact source confined to the convective clouds. However, it is important to bear in mind that the source function will force a range of linear circulations as well as gravity waves.

The background flow is unidirectional in the model calculations. Consequently, the source function and source tensor reduce to (in Cartesian notation)

$$\mathcal{F} = \nabla_h^2(D_t F_w + F_B) + (U_z F_u)_{xx} + (U_z F_v)_{xy} - (D_t F_u)_{xz} - (D_t F_v)_{yz}, \tag{9}$$

and

$$T_{ij} = \begin{bmatrix} (D_t F_w + F_B + U_z F_u) & \frac{1}{2} U_z F_v & -\frac{1}{2} D_t F_u \\ \frac{1}{2} U_z F_v & (D_t F_w + F_B) & -\frac{1}{2} D_t F_v \\ -\frac{1}{2} D_t F_u & -\frac{1}{2} D_t F_v & 0 \end{bmatrix}, \tag{10}$$

respectively. Equation (10) is used to determine the relative importance of shear, convective updrafts and diabatic heating in generating gravity waves. Note that F_b and F_B are calculated directly from the perturbation fields, whereas $\dot{b}$ is defined as a residual.

b. Source analysis of the modeled convection

Experiment E3 is a high-resolution numerical integration that focuses on a particular deep convective cell. The convective cell chosen is the strongest wave radiator in the coarser-resolution experiments E1 and E2. A deep updraft is generated during the evolution of the cell and has the appearance of an unstable bubble of moist air (Figs. 15–17). The updraft overshoots its LNB and subsequently decelerates. This particular updraft is crucial in the generation of the gravity waves observed in the calculations. Using the source function, we now analyze this convective cell using the data from Domain 3 in experiment E3.

The source function, $\mathcal{F}$, can be used to identify the region in which the waves are generated. The value is calculated using second-order differences in time and space, and zonal cross sections through $y = 98$ km at 1240, 1245, 1250, and 1300 LST are shown in Figs. 18a–d. Note that $\mathcal{F}$ is relatively noisy because its calculation requires fourth derivatives of model data. At 1240 LST (Fig. 18a), the developing updraft has a distinct signature in $\mathcal{F}$, with a maximum at a height of approximately 10 km. Once the updraft has overshot the LNB (Fig. 18b) and the direction of the updraft has been reversed (Fig. 18c), $\mathcal{F}$ is nonzero between heights of approximately 8 and 16 km (the tropopause). However, $\mathcal{F}$ is consistently largest in the region of 10 km $< z <$ 13 km, which is the region of the LNB. Perhaps the most important point is that although the updraft reaches the tropopause and displaces the isentropes in the lower stratosphere (Figs. 16b and 16c), the magnitude of $\mathcal{F}$ is a maximum well below the tropopause. After the oscillating updraft has dissipated, the dominant feature of $\mathcal{F}$ is another developing updraft (Fig. 18d). Given that $\mathcal{F}$ has a consistent maximum at the LNB of the oscillating updraft, it is likely that this oscillation is the dominant wave producing mechanism. As shown later in section 5a, this idea is consistent with the structure of the waves generated.

The source function [Eq. (9)] can be used to determine the principal physical processes responsible for the gravity waves. The terms that comprise Eq. (9) are grouped according to the physical process they represent, that is, nonlinear advection, diabatic heating, and environmental wind shear. Whereupon, the source function is

$$\mathcal{F} = \underbrace{\nabla_h^2(D_t F_w + F_b) - (D_t F_u)_{xz} - (D_t F_v)_{yz}}_{\text{advection terms}} + \underbrace{\nabla_h^2(\dot{b})}_{\text{heating}} + \underbrace{(U_z F_u)_{xx} + (U_z F_v)_{xy}}_{\text{shear}}.$$

These three terms will be referred to as $\mathcal{F}_a$, $\mathcal{F}_h$, and $\mathcal{F}_s$,

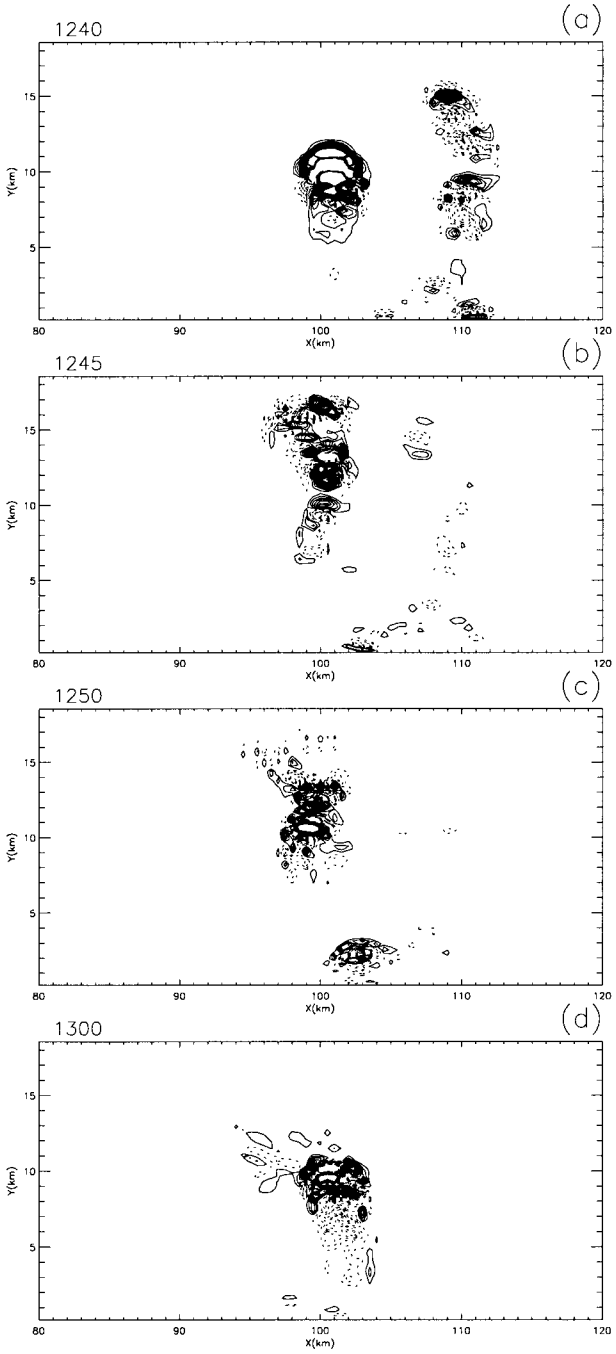


FIG. 18. Zonal cross section through $y = 98$ km of the source function for (a) 1240 LST, (b) 1245 LST, (c) 1250 LST, and (d) 1300 LST. Contour interval is 2×10^{-9} m^{-1} s^{-3}, with negative values dashed, and the zero contour omitted.

respectively. A zonal cross section of these terms through $y = 98$ km at 1245 LST is shown in Figs. 19a–c. Recall that at 1245 LST, the bulk of the updraft is still ascending even though it has overshot its LNB.

Consider the advective source term, $\mathcal{F}_a$, and the diabatic heating source term, $\mathcal{F}_h$ (Figs. 19a and 19b). At this time, the greater contribution to the total forcing (Fig. 18b) comes from the nonlinear advection, with the magnitude of $\mathcal{F}_h$ being approximately one tenth of $\mathcal{F}_a$. Furthermore, $\mathcal{F}_h$ is restricted to the middle troposphere, whereas the maximum of $\mathcal{F}_a$ is in the upper troposphere. This is an important result, as the majority of previous studies dealing with convectively generated waves have assumed a priori that diabatic heating is the dominant wave-producing mechanism (see, e.g., Bretherton 1988; Lin et al. 1998; Chun and Baik 1998). While the cloud ultimately owes its existence to diabatic heating, our results indicate that the production of buoyancy by diabatic heating may not be the main process generating the gravity waves in the numerical model.

One needs to be cautious when making deductions about the relative importance of the forcing terms without explicitly solving the forced wave equation [Eq. (6)]. Even though $\mathcal{F}_h$ is much smaller in amplitude than $\mathcal{F}_a$, it has a more coherent structure and may be an efficient wave generator. Nonetheless, the relative size of $\mathcal{F}_a$ strongly suggests that it cannot be neglected.

Consider now the contribution of the environmental wind shear terms, $\mathcal{F}_s$ (Fig. 19c), to the total source function. The contribution from shear is negligible in this case, with $\mathcal{F}_s$ being at least an order of magnitude smaller than all other terms. Although $\mathcal{F}_s$ is nonzero in Fig. 19c, the contour interval is one fifth of that used for Figs. 19a and 19b. While the convective updraft penetrates the shear layer, the contribution of this mechanism to the source function is negligible. Upon neglecting the effects of shear, the source tensor reduces to

$$T_{ij} = \begin{bmatrix} (D_t F_w + F_B) & 0 & -\frac{1}{2} D_t F_u \\ 0 & (D_t F_w + F_B) & -\frac{1}{2} D_t F_v \\ -\frac{1}{2} D_t F_u & -\frac{1}{2} D_t F_v & 0 \end{bmatrix}. \quad (11)$$

The minor role played by the background wind shear in generating the waves is consistent with the fields in Figs. 7–9. In particular, the horizontal wavelength is the same in both the zonal and meridional cross sections even though there is a shear layer only in the background zonal wind. Nevertheless, it will be shown in the next section that the wind shear plays an important role in changing the propagation characteristics of the waves.

5. Stratospheric wave response to tropospheric oscillations

Section 4 examined the *source* of the gravity waves in the model. In particular, analysis of the source tensor suggested that the waves are probably generated as air parcels in the cloud updraft oscillate about their LNB.

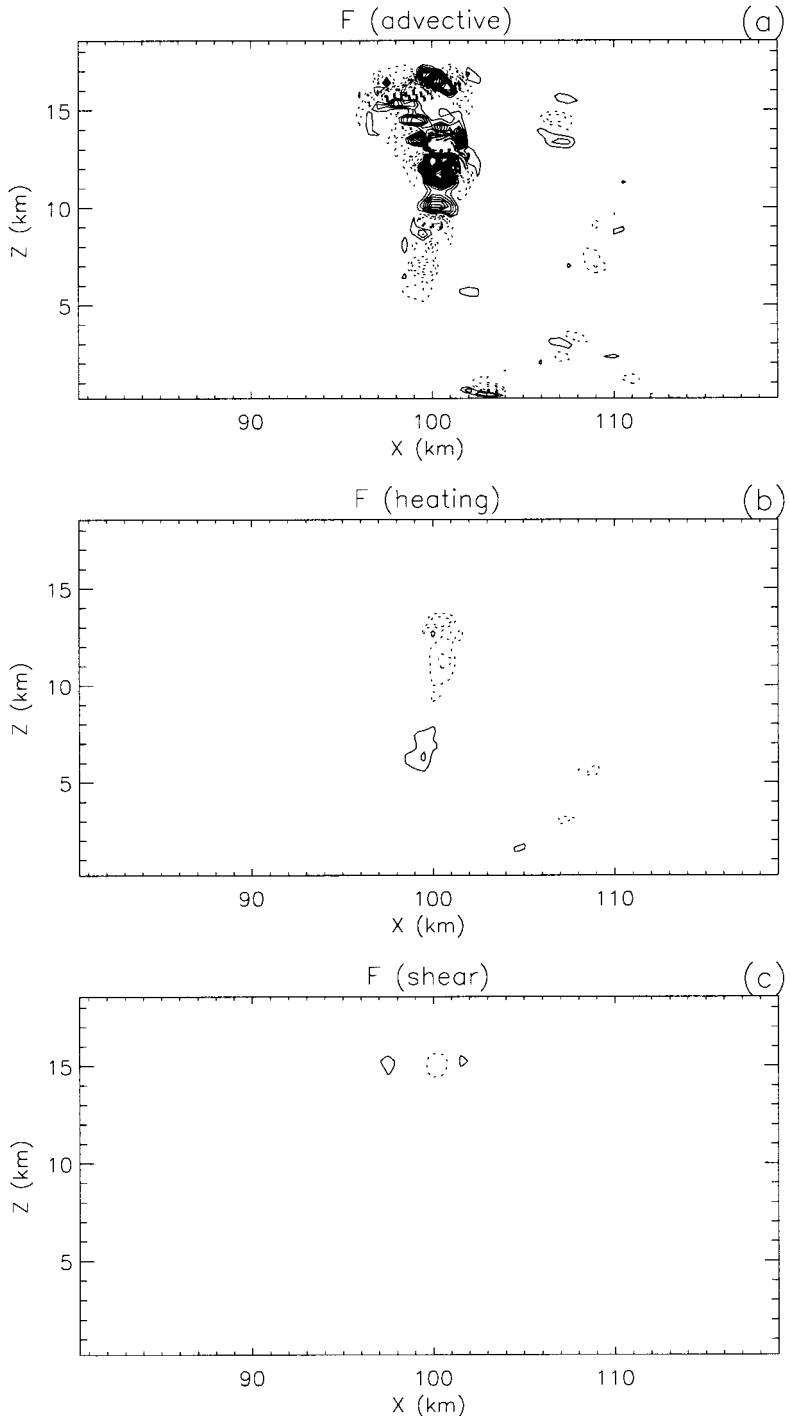


FIG. 19. Zonal cross section through $y = 98$ km at 1245 LST of (a) $\mathcal{F}_a$, (b) $\mathcal{F}_h$, and (c) $\mathcal{F}_s$. Contour interval is 2×10^{-9} m^{-1} s^{-3} in (a) and (b) and 0.4×10^{-9} m^{-1} s^{-3} in (c). Negative values are dashed, and the zero contour is omitted.

This is essentially the mechanism proposed by Pierce and Coroniti (1966). In this section, we examine the *response* of the atmosphere to this kind of forcing, and we show that the wave field is consistent with the hypothesized forcing.

a. Refraction effects

In the case examined, the gravity waves propagate radially away from the source and therefore have no preferred direction of propagation. This suggests that

vertical wind shear, which is only nonzero in the zonal direction, does not play a significant role in the generation of the gravity waves. The analysis of the source tensor has lent support to this. Also, as the wave fronts are circular, in the zonal plane through the wave source, the local meridional wavenumber l is zero, and therefore the analysis using the dispersion relation can be completed in two dimensions. Likewise, in the meridional plane through the source, the local zonal wavenumber k is zero. With these points in mind, the vertical wind shear is neglected in this section, and the two-dimensional dispersion relation is written as

$$N^2 \cos^2(\alpha) = k^2c^2 = (\omega \pm kU)^2, \qquad (12)$$

where α is the angle of phase lines from the vertical, k is the horizontal wavenumber, U is the background flow speed, c is the horizontal phase speed, and ω is the frequency of the wave in the chosen frame of reference. The parameters k and c are conserved along ray paths providing the background flow, U, is steady and horizontally homogeneous. Note that substituting l for k and $V = 0$ for U into Eq. (12) yields the dispersion relation for the meridional direction.

Let U_{LNB} be the background zonal flow at the LNB, and consider a frame of reference translating at this speed. In this frame of reference, the convective updraft is vertical and stationary, and it oscillates with a frequency equal to the background static stability in the troposphere, N_t. The tropospheric waves generated by the source will be evanescent, and consequently, their phase lines will be vertical ($\alpha = 0$). Hence, the real part of the vertical wavenumber m will be zero.

In the stratosphere, however, the dispersion relation is

$$N_s^2 \cos^2(\alpha) = k^2c^2 = (\omega \pm k\tilde{U})^2,$$

where $\omega = N_t$, $\tilde{U} = U - U_{\mathrm{LNB}}$ is the flow in the stratosphere in the frame of reference moving with the LNB, and N_s is the stratospheric Brunt–Väisälä frequency. This implies that

$$\cos(\alpha) = \frac{N_t \pm \tilde{U}k}{N_s}. \qquad (13)$$

The minus sign represents waves propagating to east, and the plus sign represents waves propagating to the west. In the case of zero LNB relative flow ($\tilde{U} = 0$), Eq. (13) reduces to

$$\cos(\alpha) = \frac{N_t}{N_s}. \qquad (14)$$

Using a two-dimensional numerical model, Fovell et al. (1992) examined the generation of gravity waves by a squall line. They concluded that the individual cells that comprise the squall line were responsible for generating the waves, and they discussed the dynamical process in terms of a horizontally propagating mechanical oscillator. Fovell et al. (1992) related the frequency of the oscillator to the lifetime of the individual convective cells. In all other respects, their analysis is very similar to that described above, the principle differences being that ours is performed in the frame of reference moving with the wind at the LNB, and that the frequency of the oscillation in this frame of reference is set by the tropospheric Brunt–Väisälä frequency. Performing the analysis in the frame of reference of the LNB exposes most clearly the simple process responsible for the pattern of gravity waves produced in the stratosphere by the model. The increased stratification above the tropopause refracts the waves, tilting phase lines away from the source, while the flow relative to the LNB refracts the phase lines in the direction of the background flow.

Equations (13) and (14) can be used to predict the slope in the stratosphere of the lines of constant phase. In the model calculations, the LNB lies between 10 and 14 km, depending on the level from which an air parcel is lifted. This is below the shear layer, and the velocity at the LNB is equal to $(-5, 0)$ m s^{-1}. Above the shear layer the velocity is equal to $(-10, 0)$ m s^{-1}. The upper-tropospheric Brunt–Väisälä frequency, N_t, is equal to 0.007 s^{-1}, and the stratospheric Brunt–Väisälä frequency, N_s, is equal to 0.028 s^{-1}. Note that N_t and N_s correspond to periods of 15.0 min and 3.7 min, respectively. In the absence of rotation, the meridional direction can be treated independently of the zonal direction. As there is no background flow in the meridional direction, the slope of the phase lines in a meridional cross section is given by Eq. (14). In the zonal direction, Eq. (13) describes the slope of the phase lines in the stratosphere. The two branches represent upstream propagating waves (negative branch) and downstream propagating waves (positive branch). Figure 20 shows α as a function of wavelength. In the absence of LNB relative flow, the gravity wave pattern is symmetric about the source, and the angle at which the phase lines slope is independent of wavelength. In the case of nonzero LNB relative flow, the phase lines are rotated in the direction of the flow by an angle that is dependent on the horizontal wavelength of the waves.

Dark lines parallel to the lines of constant phase in the stratosphere are marked in Figs. 9a and 9b. In the case with nonzero LNB relative flow (Fig. 9a), the angle from the vertical of the upstream oriented phase lines is 71° ± 1°, and the angle of downstream oriented phase lines is 81° ± 1°. In the case of zero LNB relative flow (Fig. 9b), the waves are symmetric about the source, and the phase lines make an angle of 76° ± 1° to the vertical. These angles compare very well to the solutions of Eqs. (13) and (14). The angle of the phase lines for zero LNB relative flow is within the uncertainty of measurement. For the case with nonzero LNB relative flow, the gravity waves generated by the model are rotated by 5° in the direction of the relative flow. According to Fig. 20, a rotation of 5° corresponds to a horizontal wavelength of approximately 15 km. Given that the spectral peak in the horizontal wavelength is at ap-

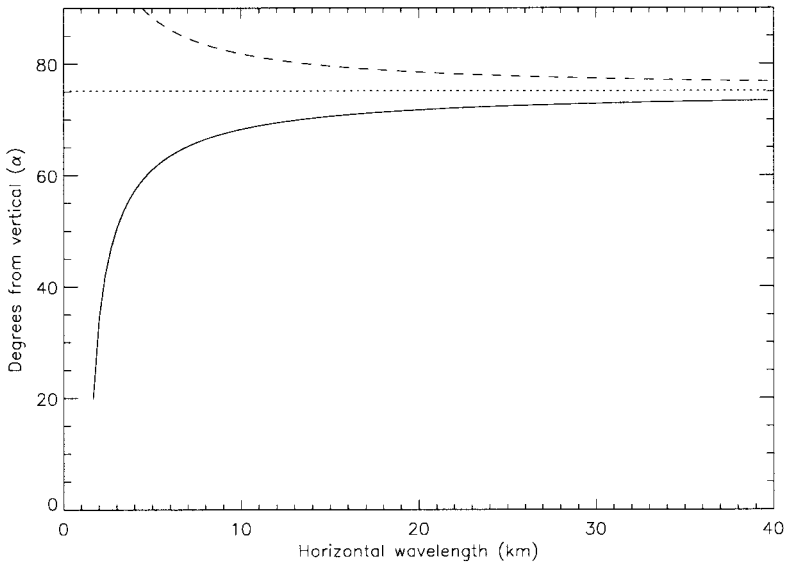


FIG. 20. Angle of the lines of constant phase in the stratosphere from the vertical (α) as a function of horizontal wavelength. The dotted line represents no stratospheric flow with respect to the tropospheric source. Shown also are the upstream (solid) and downstream (dashed) propagating solutions for a stratospheric flow of -5 m s^{-1} relative to the tropospheric source.

proximately 17 km, the theoretically predicted properties of the waves agree well with those of the modeled waves. Such close agreement lends strong support to the hypothesis that the gravity wave source is a vertically oriented oscillator moving at the velocity of the wind at the LNB. Although a spectrum of waves will be generated, the above argument describes the dominant waves.

The phase speed of the gravity waves as seen in the frame of reference of the LNB is $c = \pm N_t/k$ in the zonal direction or $c = \pm N_t/l$ in the meridional direction. Using $N_t = 0.007$ s^{-1} and $k, l = 2\pi/17$ km^{-1}, the phase speed relative to the LNB is 18.9 m s^{-1}. The ground-based phase speed is equal to $c = U_{\text{LNB}} \pm N_t/k$, where the plus sign is for waves propagating in the positive x direction (east), and the minus sign is for waves propagating in the negative x direction (west). It follows that the upstream propagating waves have a ground-based phase speed of 13.9 m s^{-1}, and the downstream propagating waves have a ground-based phase speed of 23.9 m s^{-1}. Moreover, the ground-based frequencies of the upstream and downstream propagating waves are 5.14×10^{-3} s^{-1} and 8.83×10^{-3} s^{-1}, respectively. These frequencies correspond to periods of 20.3 minutes and 11.9 minutes respectively. Of course, in the meridional direction, the ground-based and LNB reference frames are identical, and therefore, the phase speed and frequency remain unchanged.

In the cloud cluster considered here, there are a number of convective clouds with different cloud-top heights. The depth of each cloud, and hence the height of its LNB, depends on a number of factors, including entrainment, intensity, and the initial local forcing. The depth of the convection is essentially irrelevant here because the Brunt–Väisälä frequency and the background wind velocity are constant in the troposphere. Therefore, N_t and U_{LNB} are the same for each cloud, and thus, each cloud contributes to the same part of the ground-based frequency spectrum. However, in an atmosphere with tropospheric wind shear, clouds with different depths would possess different values of U_{LNB}, and each cloud would generate waves with different ground-based frequencies. This would broaden the ground-based frequency spectrum. The effect of tropospheric shear on the gravity wave generation will be considered in a future study.

The momentum flux comprises contributions from upstream propagating waves (positive) and downstream propagating waves (negative). The momentum flux for each branch varies like $\overline{\rho} A_u A_w$, where A_u and A_w are the amplitudes of the perturbation horizontal and vertical velocity, respectively. The size of A_u and A_w changes with the angle of the phase lines. (This is discussed in more detail below.) Therefore, even though shear does not appear to be generating the dominant waves, it changes the slope of upstream and downstream propagating waves. Consequently, the shear produces a zonal asymmetry in the waves and, hence, a nonzero (zonal) momentum flux. However, the downstream propagating waves still make a contribution to the (zonal) momentum flux, although it is smaller than the contribution of the upstream propagating waves. Figure 13a shows that the downstream propagating waves reduce the magnitude of the momentum flux in the lower stratosphere.

Often the wave signatures in vertical velocity or potential temperature are used to analyze gravity waves.

Consider Figs. 9a and 9d. In these figures, the upstream propagating waves appear to dominate the wave field. However, upstream propagating waves are not preferentially excited. As shown earlier, a shear layer refracts the gravity waves. This refraction affects the signature of the waves in vertical velocity, potential temperature, and perturbation horizontal velocity. Using continuity, it can be shown that (in two dimensions) the amplitudes of the horizontal velocity perturbations, A_u, and the vertical velocity, A_w, are related by the expression

$$A_u = -\frac{m}{k}A_w. \tag{15}$$

It follows that when the phase lines of gravity waves tilt to the horizontal (vertical), A_u increases (decreases) with respect to A_w. However, A_w is not conserved following packet trajectories.[3] The wave action, $\tilde{A}$, [first defined by Bretherton (1966)] is conserved along packet trajectories and (in two dimensions) is defined as

$$\tilde{A} = \frac{E_0}{(\omega - Uk)},$$

where E_0 is the wave energy per unit volume defined by

$$E_0 = \frac{1}{2}\overline{\rho}\left(u^2 + w^2 + \frac{b^2}{N^2}\right).$$

It can be shown that the amplitude of w varies as

$$A_w^2 \sim (\omega - Uk)^3\frac{\tilde{A}}{\overline{\rho}N^2},$$

where $\tilde{A}$ depends only on ω and k [see Bretherton (1966) for details]. It is also possible to show that if the background flow is steady and horizontally uniform, then the wave frequency (ω) and horizontal wavenumber (k) are conserved following packet trajectories. Note also that

$$(\omega - Uk)^2 = \frac{N^2k^2}{k^2 + m^2}.$$

Thus in the stratosphere (where k, N, and ω are all constant), the amplitude of the vertical velocity signature will decrease as the magnitude of the vertical wavenumber increases (i.e., as the phase lines are refracted to the horizontal). Therefore, even though the wave action remains constant, a refraction of the phase lines to the horizontal (vertical) will reduce (increase) the vertical velocity signature. Therefore, without taking into account these changes in wave signature due to wave refraction, analyses may be biased toward upstream propagating waves.

[3] A packet trajectory is the trajectory defined by the group velocity of a wave packet.

b. Horizontal wavelength selection

In the previous section, it was shown that the tropospheric gravity waves satisfy the assumptions that $\omega = N_t$ in the frame of reference of the LNB and that $m = 0$. However, under these conditions, the dispersion relation is automatically satisfied for all horizontal wavelengths. This prompts the following question: what determines the horizontal wavelength? The analysis assumes that the frequency, vertical wavenumber, and horizontal wavenumber are all real, which is true for a freely oscillating wave mode with infinite spatial extent. However, the convective updraft, which generates the waves, is damped and has limited spatial extent. In this section, the convective updraft is described as a spatially evanescent damped oscillation, and a simple analysis of the dispersion relation is used to explain how the horizontal wavelength is selected.

Rearranging the dispersion relation [Eq. (12)] for the troposphere, the horizontal wavenumber can be written as a function of the frequency and vertical wavenumber

$$k^2 = m^2 \Bigg/ \left(\frac{N_t^2}{\omega^2} - 1\right), \tag{16}$$

where k, m, and ω may be complex. As before, we assume that the real parts of ω and m are N_t and 0, respectively. Therefore, k, m, and ω have the form

$$k = k_R + ik_I,$$

$$m = 0 + im_I, \quad \text{and}$$

$$\omega = N_t - i\omega_I.$$

Hence, the vertical velocity can be written as

$$w = A(k)\exp(-k_Ix - m_Iz - \omega_It)\exp[i(k_Rx - N_tt)].$$

The horizontal, vertical, and temporal e-folding scales over which the convective updraft decays are $1/k_I$, $1/m_I$, and $1/\omega_I$, respectively. Physically, $1/k_I$ is determined by the width of the convective updraft, $1/m_I$ is related to the distance air parcels overshoot their LNB, while $1/\omega_I$ is presumably related to the turbulent mixing. Of course, the convective system will produce mesoscale variations in the background flow, which will affect the trapping and propagation characteristics of the gravity waves. Such effects are difficult to quantify.

From Fig. 15b, the e-folding decay scale in the vertical is estimated to be approximately 4 km, whereupon k can be calculated as a function of ω_I using Eq. (16). The horizontal wavelength, $2\pi/k_R$, and the horizontal e-folding scale, $1/k_I$, are plotted as a function of ω_I in Fig. 21a. Similarly, the horizontal wavelength can be considered as a function of the e-folding horizontal scale; this is plotted in Fig. 21b. If the convective updraft can be described as a spatially evanescent damped oscillator, then Eq. (16) predicts a single (real) horizontal wavelength, explaining why the stratospheric wave field is almost monochromatic. This mode ''leaks'' from the

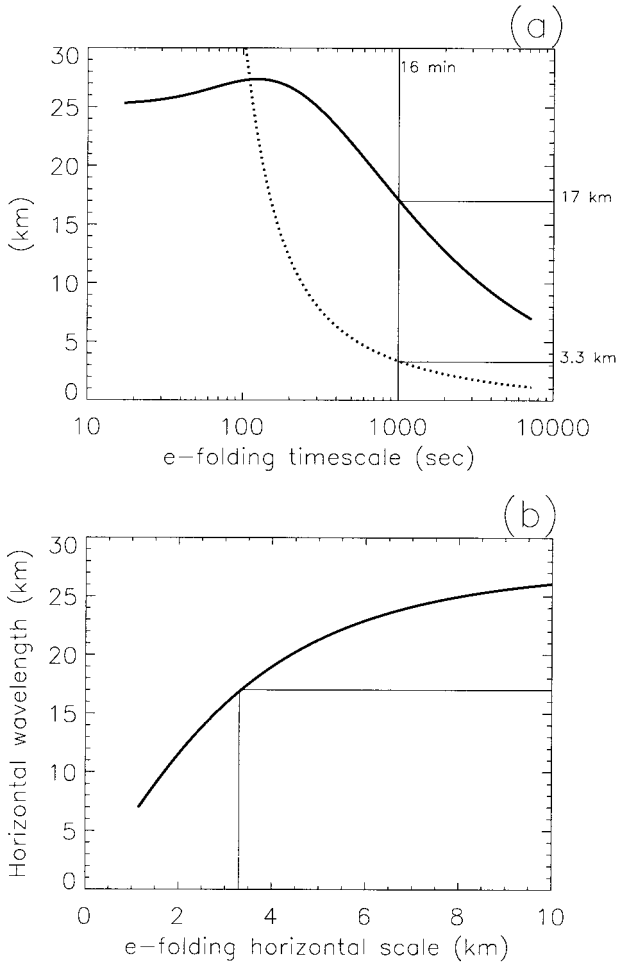


FIG. 21. (a) The horizontal wavelength $2\pi/k_R$ (solid) and *e*-folding horizontal scale $1/k_I$ (dashed) as a function of the *e*-folding timescale $1/\omega_I$. (b) The horizontal wavelength $2\pi/k_R$ as a function of the *e*-folding horizontal scale $1/k_I$. Both plots show the intersection of the curves with the dominant horizontal wavelength from the model calculations (17 km).

troposphere into the stratosphere where it propagates vertically.

In the model calculations, the dominant horizontal wavelength in the stratosphere is approximately 17 km. Assuming that the *e*-folding decay scale in the vertical is 4 km, Fig. 21 shows that the updraft must have horizontal and temporal *e*-folding decay scales of approximately 3.3 km and 16 min, respectively. These values are consistent with those found in the model solution.

6. Summary and conclusions

Convective clouds are known to generate gravity waves. These waves are thought to be important in shaping the mean flow in the middle atmosphere and in modifying the environment of the cloud. The central aim of the present paper has been to clarify how clouds generate such waves.

In this study, a high-resolution numerical model has been used to simulate the evolution and structure of a convective cloud cluster over the Tiwi Islands, and the gravity waves generated by the cluster have been examined in detail. Three numerical runs were presented. The principal differences between each run was the horizontal resolution, which was 2, 1, or 0.5 km. Each run was initialized with an idealization of a particular radiosonde sounding taken over the Tiwi Islands during MCTEX. Island thunderstorms, such as these, are triggered by sea-breeze convergence, and because of their regularity and simple forcing mechanism, they are ideal for studying convectively generated waves.

In the model runs, a cluster of deep convective clouds developed over the western half of the Tiwi Islands. These clouds merged to form a north–south line approximately 40 km long moving westward with the background flow. A number of large convective cells developed in each model run, although in the calculation with coarsest resolution, smaller-scale convective structures were not well resolved. The position of the convective cells was also sensitive to spatial resolution.

As the convection evolved, gravity waves formed part of the numerical solution in both the troposphere and stratosphere. However, the largest amplitude gravity waves developed when the convection deepened and reached the upper troposphere. The gravity wave source was highly localized and highly variable in both space and time. The position of the source was directly related to the position of developing convective cells. The amplitude of the gravity waves in the stratosphere was small. For example, the vertical velocity fluctuations were about 0.3 m s^{-1}. The waves were remarkably monochromatic and had a horizontal wavelength of about 17 km. This wavelength was shown to be set by the width of the convective updraft, the degree of turbulent mixing, and the distance by which rising air parcels overshot their LNB.

The stratospheric momentum flux associated with the vertically propagating waves was greatest shortly after the convective clouds reached the upper troposphere. The waves and their source were highly transient, producing pulses of momentum flux in the stratosphere. Consequently, the momentum flux was sensitive to the area over which it was averaged.

The *source tensor* and *source function* were introduced for the anelastic equations of motion with arbitrary background shear but no background rotation. The source function was used to clarify the source of the gravity waves generated by the modeled convection. Although there was some uncertainty in the conclusions, the results were consistent with the hypothesis that the oscillation of the convective updrafts about their LNB was the dominant mechanism forcing the gravity waves.

Although the production of buoyancy by latent heating is central to the development of convective clouds, it has been shown that in our model, it is not the dominant term forcing the linear wave equation. Rather, non-

linear advection was the dominant term in the source function. This suggests that diabatic heating may not be the primary mechanism forcing the gravity waves. In contrast, most previous work on convectively generated gravity waves has assumed, a priori, that the transient diabatic heating generates the waves (e.g., Bretherton 1988; Lin et al. 1998; Chun and Baik 1998).

Shear was also found to be unimportant in generating gravity waves in our model runs. In particular, the gravity waves showed no preferential direction of propagation, with the dominant horizontal wavelength being the same in all directions from the source. We note that if the waves were generated by the so-called *obstacle effect* (e.g., Clark et al. 1986) or a *moving mountain* effect (e.g., Pfister et al. 1993), upstream propagating gravity waves would have been dominant, and the shear terms in the source tensor would have been important contributors.

It has been shown that treating the wave source as a vertically oriented oscillator moving with the LNB is consistent with the slope of the phase lines in the stratosphere. In the frame of reference moving with wind at the LNB, the frequency of the gravity waves was equal to the Brunt–Väisälä frequency in the troposphere. It was also shown that although shear does not generate the waves, it refracts the phase lines in the direction of the flow relative to the LNB. This refraction creates an asymmetry between the upstream and downstream propagating waves, thereby producing a nonzero momentum flux. The refraction also reduces the wave signature in the vertical velocity of the downstream propagating waves, thereby giving the appearance that upstream waves dominate the stratospheric wave field. Despite this, the downstream propagating waves make an important contribution to the momentum flux in the lower stratosphere.

The results of this paper emphasize that convective clouds generate gravity waves when the convective updraft rapidly decelerates. With this in mind, one might ask the following: how does the role of shear change if a shear layer is centered on the LNB? Collocating the tropopause and the LNB might represent one such case of maximizing shear and buoyancy effects. The analysis presented can also be used to better understand how squall lines, boundary layer thermals, or jet streaks excite gravity waves. These are topics of current and planned research.

Acknowledgments. We would like to thank Andrew Crook for his help in setting up the Tiwi Island simulation. We are grateful to Joan Alexander, Dale Durran, Roger Smith, and two anonymous reviewers whose comments helped improve the manuscript.

REFERENCES

Alexander, M. J., and J. R. Holton, 1997: A model study of zonal forcing in the equatorial stratosphere by convectively induced gravity waves. *J. Atmos. Sci.,* **54,** 408–419.

——, ——, and D. R. Durran, 1995: The gravity wave response above deep convection in a squall line simulation. *J. Atmos. Sci.,* **52,** 2212–2226.

Bretherton, C. S., 1988: Group velocity and the linear response of stratified fluids to internal heat or mass sources. *J. Atmos. Sci.,* **45,** 81–93.

——, and P. K. Smolarkiewicz, 1989: Gravity waves, compensating subsidence and detrainment around cumulus clouds. *J. Atmos. Sci.,* **46,** 740–759.

Bretherton, F. P., 1966: The propagation of groups of internal gravity waves in a shear flow. *Quart. J. Roy. Meteor. Soc.,* **92,** 466–480.

Chun, H.-Y., and J.-J. Baik, 1998: Momentum flux by thermally induced internal gravity waves and its approximation for large scale models. *J. Atmos. Sci.,* **55,** 3299–3310.

Clark, T. L., 1977: A small-scale dynamic model using a terrain-following coordinate transformation. *J. Comput. Phys.,* **24,** 186–215.

——, and R. D. Farley, 1984: Severe downslope windstorm calculations in two and three spatial dimensions using anelastic interactive grid nesting: A possible mechanism for gustiness. *J. Atmos. Sci.,* **41,** 329–350.

——, and W. D. Hall, 1996: On the design of smooth, conservative vertical grids for interactive grid nesting with stretching. *J. Appl. Meteor.,* **35,** 1040–1046.

——, T. Hauf, and J. P. Kuettner, 1986: Convectively forced internal gravity waves: Results from two-dimensional numerical experiments. *Quart. J. Roy. Meteor. Soc.,* **112,** 899–925.

Crook, N. A., 1997: Simulation of convective storms over the Tiwi Islands and comparison with observations from MCTEX. Preprints, *22d Conf. on Hurricane and Tropical Meteorology,* Fort Collins, CO, Amer. Meteor. Soc., 376–377.

——, T. L. Clark, and M. W. Moncrieff, 1991: A numerical study of the Denver cyclone. Part II: Interaction with the convective boundary layer. *J. Atmos. Sci.,* **48,** 2109–2126.

Dewan, E. M., and Coauthors, 1998: MSX satellite observations of thunderstorm-generated gravity waves in mid-wave infrared images of the upper stratosphere. *Geophys. Res. Lett.,* **25,** 939–946.

Dunkerton, T., 1997: The role of gravity waves in the quasi-biennial oscillation. *J. Geophys. Res.,* **102,** 26 053–26 076.

Ford, R., 1994: Gravity wave radiation from vortex trains in rotating shallow water. *J. Fluid Mech.,* **281,** 81–118.

Fovell, R., D. Durran, and J. R. Holton, 1992: Numerical simulations of convectively generated stratospheric gravity waves. *J. Atmos. Sci.,* **49,** 1427–1442.

Fritts, D. C., 1993: Gravity wave sources, source variability and lower middle atmosphere effects. *Coupling Processes in the Lower and Middle Atmosphere,* E. V. Thrane, T. A. Bix, and D. C. Fritts, Eds., Kluwer Academic, 191–208.

Golding, B. W., 1993: A numerical investigation of tropical island thunderstorms. *Mon. Wea. Rev.,* **121,** 1417–1433.

Gray, W. M., 1973: Cumulus convection and larger scale circulations. I. Broadscale and mesoscale considerations. *Mon. Wea. Rev.,* **101,** 839–855.

Guest, F., M. J. Reeder, C. J. Marks, and D. J. Karoly, 2000: Observations of stratospheric inertia-gravity waves over Macquarie Island. *J. Atmos. Sci.,* **57,** 737–752.

Hauf, T., and T. L. Clark, 1989: Three-dimensional numerical experiments on convectively forced internal gravity waves. *Quart. J. Roy. Meteor. Soc.,* **115,** 309–333.

Karoly, D. J., G. L. Roff, and M. J. Reeder, 1996: Gravity wave activity associated with tropical convection and detected in TOGA COARE sounding data. *Geophys. Res. Lett.,* **23,** 1153–1174.

Keenan, T. D., B. R. Morton, X. S. Zhang, and K. Nyguen, 1990: Some characteristics of thunderstorms over Bathurst and Melville Islands near Darwin, Australia. *Quart. J. Roy. Meteor. Soc.,* **116,** 1153–1172.

——, and Coauthors, 1994: Science plan—Maritime Continent Thun-

derstorm Experiment. BMRC Research Rep. 44, 61 pp. [Available from BMRC, GPO Box 1289K, Melbourne, VIC 3001, Australia.]

Koenig, L. R., and F. W. Murray, 1976: Ice-bearing cumulus cloud evolution: Numerical simulation and general comparison against observations. *J. Appl. Meteor.,* **15,** 747–762.

Kuettner, J. P., P. A. Hildebrand, and T. L. Clark, 1987: Convection waves: Observations of gravity wave systems over convectively active boundary layers. *Quart. J. Roy. Meteor. Soc.,* **113,** 445–467.

Larsen, M. F., W. E. Swartz, and R. F. Woodman, 1982: Gravity-wave generation by thunderstorm observed with a vertically-pointing 430 MHz radar. *Geophys. Res. Lett.,* **9,** 571–574.

Lighthill, M. J., 1952: On sound generated aerodynamically. I. General theory. *Proc. Roy. Soc. London,* **211A,** 564–587.

——, 1954: On sound generated aerodynamically. II. Turbulence as a source of sound. *Proc. Roy. Soc. London,* **222A,** 1–32.

Lilly, D. K., 1962: On the numerical simulation of buoyant convection. *Tellus,* **14,** 145–172.

Lin, Y.-L., R. L. Deal, and M. S. Kulie, 1998: Mechanisms of cell regeneration, development, and propagation within a two-dimensional multicell storm. *J. Atmos. Sci.,* **55,** 1867–1886.

Mapes, B. E., 1993: Gregarious tropical convection. *J. Atmos. Sci.,* **50,** 2026–2037.

Pandya, R. E., and M. J. Alexander, 1999: Linear stratospheric gravity waves above convective thermal forcing. *J. Atmos. Sci.,* **56,** 2434–2446.

Pfister, L., S. Scott, and M. Loewenstein, 1993: Mesoscale disturbances in the tropical stratosphere excited by convection: Observations and effects on the stratospheric momentum budget. *J. Atmos. Sci.,* **50,** 1058–1075.

Pierce, A. D., and S. C. Coroniti, 1966: A mechanism for the generation of acoustic-gravity waves during thunderstorm formation. *Nature,* **210,** 1209–1210.

Prusa, J. M., P. K. Smolarkiewicz, and R. R. Garcia, 1996: Propagation and breaking at high altitudes of gravity waves excited by tropospheric forcing. *J. Atmos. Sci.,* **53,** 2186–2216.

Reeder, M. J.,, and R. K. Smith, 1998: Mesoscale meteorology. *Meteorology of the Southern Hemisphere,* D. J. Karoly and D. G. Vincent, Eds., Amer. Meteor. Soc., 201–241.

——, and M. Griffiths, 1996: Stratospheric inertia-gravity waves generated in a numerical model of frontogenesis. II. Wave sources, generation mechanisms and momentum fluxes. *Quart. J. Roy. Meteor. Soc.,* **122,** 1175–1195.

Salby, M., and R. Garcia, 1987: Transient response to localised episodic heating in the tropics. Part I: Excitation and short-time near field behaviour. *J. Atmos. Sci.,* **44,** 458–498.

Sato, K., H. Hashiguchi, and S. Fukao, 1995: Gravity waves and turbulence associated with cumulus convection observed with the UHF/VHF clear-air doppler radars. *J. Geophys. Res.,* **100,** 7111–7119.

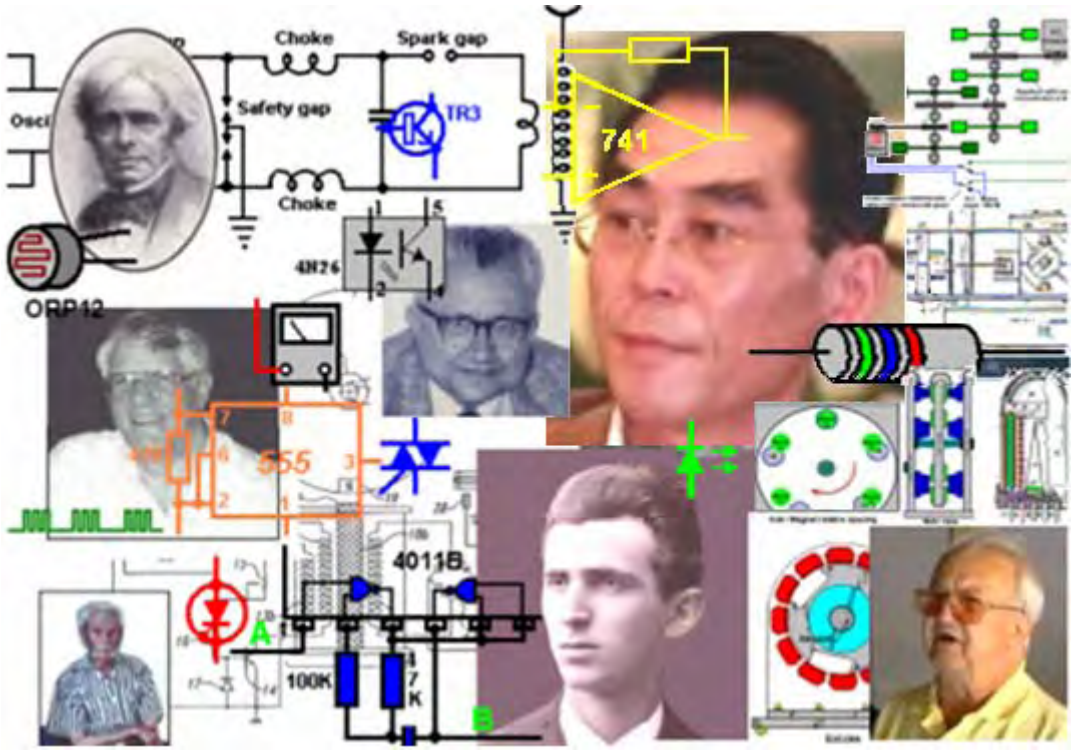
Shutts, G. J., and M. E. B. Gray, 1994: A numerical modelling study of the geostrophic adjustment process following deep convection. *Quart. J. Roy. Meteor. Soc.,* **120,** 1145–1178.

Smagorinsky, J., 1963: General circulation experiments with the primitive equations. I. The basic experiment. *Mon. Wea. Rev.,* **91,** 99–164.

Stein, R. F., 1967: Generation of acoustic and gravity waves by turbulence in an isothermal stratified atmosphere. *Sol. Phys.,* **2,** 385–432.

Weisman, M. L., W. C. Skamarock, and J. B. Klemp, 1997: The resolution dependence of explicitly modeled convective systems. *Mon. Wea. Rev.,* **125,** 527–548.

Practical Guide to 'Free-Energy' Devices



Author: Patrick J. Kelly

Version: 26.6 Release date: 5 Nov 2014

This eBook has more than 400 bookmarks to speed up your access to any section of interest. The contents of this eBook is only about 5% of the information on the http:www.free-energy-info.tuks.nl web site

Preface

Here is a small amount of background information in order that you can understand the nature of this "***Practical Guide to Free-Energy Devices***".

I am just an ordinary person who became interested in "free-energy" as a result of a television programme entitled 'It Runs on Water' shown in the 1980s by a UK television company called 'Channel 4'. At the present time that programme can be seen at https://www.youtube.com/watch?v=xHISTsiR9qc. From my point of view, the content of this documentary seemed to be rather unsatisfactory as it suggested quite a number of very interesting things but gave no real hard and fast specifics for the viewer to follow up on to investigate the subject further. However, it had the enormous benefit of making me aware that there was such a thing as "free-energy".

My attempts to find out more were not very successful. I bought paper copies of several of Stan Meyer's HHO gas patents from the Patent Office in 1986 but while they were interesting, they did not provide much in the way of additional information. Searching on the internet at that time did not produce much more in the way of practical information. Things have changed dramatically since then and there has been an enormous increase in available information. But, even today, it is relatively difficult to find much in the way of direct, useful and practical information on free-energy systems and techniques. Much of the information consists of chatty, lightweight articles describing people, events and inventions in vague, broad outline terms which are almost completely lacking in specifics.

These articles have the style of saying "There is a new invention called a 'bus' which is used to carry passengers from place to place. We saw one the other day, it was painted green and blue and looked most attractive. It is driven by Joe Bloggs who wears an engaging smile and a hand-knitted sweater. Joe says that even his children could drive a bus as it is so easy to do. Joe expects to retire in six months time as he is going to take up gold prospecting." While I'm sure that an article like that is interesting, the sort of description which I would want would be: "There is a new invention called a 'bus' which is used to carry passengers from place to place. We saw one the other day, and were very impressed as it has seats for some forty-five people. It has bodywork made of pressed aluminium, a wheel at each corner of its considerable 40' x 10' structure, a five litre diesel engine made by the Bosworth Engineering Company of Newtown, and has power-assisted steering, hydraulic brakes and ……".

There are also many articles, scientific papers and books, some of which, quite frankly, I am not able to understand as the authors think mathematically and express themselves in equations (where they frequently do not define the terms which they use in their equations, making them effectively meaningless). I do not think in mathematical equations, so I do not share in this much higher level of thinking and analysis, though I do have some of these papers on my web site for the benefit of visitors who do have the ability to understand them easily.

After a long period of searching and investigating I was beginning to gather enough information to be fairly confident of what was being done, what had already been achieved, and some of the possible background reasons for the effects which were being observed. Early in 2005 I decided that as I had encountered so much difficulty and had to put in so much effort to find out the basics of "free-energy" that it could be helpful to others if I shared what I had found out. So I wrote the first edition of this presentation and created a simple web site to make it available to others. Of course, this body of information is not static – on the contrary, it is very fast-moving. Consequently, this information digest is updated and refined typically many times each year. The present form of presentation is the third style of layout which has been used as the volume of material has increased.

It should be stressed that this information is what I have discovered as part of my interest in the subject and is mainly a reporting on what is being said by other people. I have not built and proved every device described – to do that would take many lifetimes, so please understand that this is just an attempt to aid your own investigation. While it can be proved that some device works as described, through independent replication and verification, the reverse is not true. If someone were to build a device and fail to get it to work as described, then the most that can honestly be said is that an unsuccessful attempt was made to replicate it. It does not, of course, show that the original device did not operate exactly as described, just that the (possibly inept) attempt at replication, was not successful. In some instances, you will see that I have expressed the opinion that the device is not viable, or, as in the case of the 'Nitro Cell' that I do think that it does work, but as many people have tried to build it and failed

to get the results described, that it can't be recommended as an investigation project. However, as soon as I said that, a local man announced that he had made two Nitro Cells and attached them to his two Ford Transit trucks and that resulted in saving him an estimated £500 over the course of two years.

I do not suggest that this set of information covers every possible device, nor that my description is by any means the complete and definitive statement of everything to be known on the subject. The old saying applies here: "If you think you know all the answers, then you just haven't heard all the questions!" So, this material is just an introduction to the subject and not an encyclopaedia of every known device.

I should like to thank the very large number of people who have most kindly given me their permission to reproduce details of some of their work, providing photographs, checking what I have written, suggesting additions, etc. Also those kind people who have given me permission to reproduce their own works directly on my web sites or in my documents. There seems to be a common thread of concern among many people that shows as a desire to share this information freely rather than to try to make money from selling it, and I thank these people for their generosity.

Many people hold "conspiracy theory" views and believe that there is a concerted effort to suppress this information, and more especially, to prevent free-energy devices reaching the market. Personally, I think that while that is certainly true, much of this opposition may be just the normal reaction of vested commercial interests. If you were making a profit of literally millions per hour, would you welcome the introduction of a system which would eventually cut your income to zero? If not, then how much would you be willing to pay someone to make sure that the present system is never changed – a million? A billion? While this opposition is definitely there and people who stand to lose money and/or power through change will continue to oppose this knowledge, and to a much greater extent, the introduction of any commercial free-energy device, almost the entire focus of the information is on devices – what they do, how they are made and how they may operate when they draw additional energy from the local environment.

Let me stress again, that this set of information is not by any means the final word on the subject, but just an introduction to the subject by a single person who makes no claims to knowing all the answers. Enjoy your research – I hope you are successful in every respect.

Patrick Kelly

April 2008

A Practical Guide to 'Free-Energy' Devices

Contents

Index 1

Introduction . . . . . I -1

Chapter 1: Magnet Power
The ShenHe Wang permanent magnet motor-generator . . . . . 1 - 1
The Ecklin-Brown generator . . . . . 1 - 4
The Howard Johnson magnet motor . . . . . 1 - 6
The 'Carousel' permanent magnet motor . . . . . 1 - 7
The Robert Tracy permanent magnet motor . . . . . 1 - 9
The Ben Teal electromagnet motor . . . . . 1 - 9
The Jines permanent magnet motor . . . . . 1 - 11
The Stephen Kundel permanent magnet motor . . . . . 1 - 12
Charles Flynn's permanent magnet motor . . . . . 1 - 15
Steorn's Magnet Motor . . . . . 1 - 24
George Soukup's Magnet Motor . . . . . 1 - 26
Dietmar Hohl's V-Magnet Motor and Jes Ascanius' implementation of it . . . . . 1 - 29
Constructing a simple permanent magnet motor . . . . . 1 - 30
Muammer Yildiz's Magnet Motor . . . . . 1 - 35
Donald Kelly's Magnet Motor . . . . . 1 - 54
Mike Brady's Perendev Magnet Motor . . . . . 1 - 55
Magnetic shielding from Pasi Mäkilä . . . . . 1 - 55
The Permanent Magnet Motor of Victor Diduck . . . . . 1 - 56

Chapter 2: Moving Pulsed Systems
The Adams Motor/Generator . . . . . 2 - 1
Lidmotor's Low-voltage Rotor . . . . . 2 - 12
The Phemax Self-Powered motor/generator . . . . . 2 - 13
The Kromrey No-Drag Electrical Generator . . . . . 2 - 16
The Teruo Kawai Motor . . . . . 2 - 25
James Hardy's Water-jet Generator . . . . . 2 - 27
Georges Mourier's COP=10,000 Motor/Generator Patent . . . . . 2 - 29
The RotoVerter . . . . . 2 - 39
Coil-shorting techniques . . . . . 2 - 44
Raoul Hatem's Energy System . . . . . 2 - 48
Lawrence Tseung's COP = 3.3 Pulsed Rotor . . . . . 2 - 50
Art Porter's Electromagnetic Motor System . . . . . 2 - 56
Raising DC Motor Efficiency . . . . . 2 - 57
Direct Coupled Motors . . . . . 2 - 67

Chapter 3: Motionless Pulsed Systems
Charles Flynn's devices . . . . . 3 - 1
Lawrence Tseung's magnetic frame . . . . . 3 - 1
Thane Heins' dual toroidal magnetic frame . . . . . 3 - 3
The High-power Motionless Generator of Clemente Figuera . . . . . 3 - 6
The Zero back-EMF Coils of Alexkor . . . . . 3 - 26
The Self-Powered Generators of Barbosa and Leal . . . . . 3 - 27
The Ultra-simple Device of Lorrie Matchett . . . . . 3 - 60
The Motionless Generator of Theodore Annis and Patrick Eberly . . . . . 3 - 66
The Richard Willis COP>1 Motionless Electrical Generator . . . . . 3 -75
The 'silverhealtheu' bar generator . . . . . 3 - 77
The 1982 patent of Heinrich Kunel . . . . . 3 - 77
Valeri Ivanov's Motionless Generator . . . . . 3 - 84
The Motionless Generators of Kelichiro Asaoka . . . . . 3 - 85
Stephan W. Leben's device . . . . . 3 - 99
Floyd Sweet's VTA . . . . . 3 - 101

The magnetic generator of Ashley Gray . . . 3 - 102
Pavel Imris' Optical Generator . . . 3 - 106
The Michel Meyer and Yves Mace Isotopic Generator . . . 3 - 109
The Colman / Seddon-Gilliespie Generator . . . 3 - 110
Don Smith's Magnetic Resonance System . . . 3 - 112
Belkired Mohamed's Don Smith analysis . . . 3 - 112
Ming Cao's Developments . . . 3 - 140
Russian Developments . . . 3 - 145
Chinese Developer 'Salty Citrus' . . . 3 - 148
Another Russian Development . . . 3 - 157
Using two Tesla Coils back-to-back . . . 3 - 157
Joseph Boyd's power transformer . . . 3 - 160
The 'Gegene' Magnetic Plate Arrangement . . . 3 - 173
Tariel Kapanadze's Self-powered Devices . . . 3 - 176
An analysis of the Kapanadze devices . . . 3 - 181
The Cold Electricity Coil of 'UFOpolitics' . . . 3 - 182
Stanley Meyer's Magnetic Particles Power Generator . . . 3 - 191
The E-Stress Power Generator . . . 3 - 203

Chapter 4: Gravity-Powered Systems

Lawrence Tseung's Gravity Impact wheel . . . 4 - 1
The Chas Campbell System . . . 4 - 4
Jacob Bitsadze's Pulley Ratio Information . . . 4 - 6
The self-powered generator of José Luis García del Castillo . . . 4 - 7
The Wilson self-powered DC generator . . . 4 - 11
The Bedini Pulsed Flywheel . . . 4 - 12
James Hardy's Water-jet Generator . . . 4 - 13
The Two-year magnetic Pendulum . . . 4 - 14
Jerzy Zbikowski's Chain Drive system . . . 4 - 17
Gravitational Effects . . . 4 - 18
Mikhail Dmitriev's nudged weights gravity wheel . . . 4 - 20
The Dale Simpson Gravity Wheel . . . 4 - 22
The Veljko Milkovic Pendulum / Lever system . . . 4 - 25
Amr Al-Hossary's Hinged-PlateDesign . . . 4 - 25
The Murilo Luciano Gravity Chain . . . 4 - 27
Other Gravity Designs . . . 4 - 35
James Kwok's "Hidro" Self-Powered Generator . . . 4 - 35
The Ribero Buoyancy Patent . . . 4 - 48
The 30-kilowatt Gravity-Powered Motor . . . 4 - 50
Bhaskara's Gravity Wheel . . . 4 - 51
William Skinners' torque multiplier . . . 4 - 53

Chapter 5: Energy-Tapping Pulsed Systems

Frank Prentice's horizontal wire system . . . 5 - 1
Dave Lawton's Water Fuel Cell . . . 5 - 3
John Bedini's Battery Pulse-Charger . . . 5 - 4
The Tesla Switch . . . 5 - 5
A 3-Battery Switching Circuit . . . 5 - 11
The Self-powered Free-Energy Generators of Carlos Benitez . . . 5 - 20
Bozidar Lisac's 1-battery charger patent . . . 5 - 37
Bob Boyce's Toroid . . . 5 - 46
Don Smith's devices . . . 5 - 52
Tariel Kapanadze's system . . . 5 - 53
The insights of Vladimir Utkin . . . 5 - 53
Walter Ford's high-powered crystal set . . . 5 - 54
Lawrence Tseung's "FLEET" toroid . . . 5 – 56
The LED and CFL lighting circuits of Jeanna and Lidmotor . . . 5 – 58
Advanced Joule Thief circuits . . . 5 - 66
LaserSaber's Joule Thief lighting circuits . . . 5 - 80
The Ed Gray / Marvin Cole Power System . . . 5 - 83
Tesla's Experiments . . . 5 - 85
The Alberto Molina-Martinez Generator . . . 5 - 91
The Hubbard Self-powered Generator . . . 5 - 92
The Joseph Cater Self-powered Generator . . . 5 - 94

Floyd Sweet's VTA . . . . . 5 - 102
Rosemary Ainslie's COP=17 heater . . . . . 5 - 104
Joseph H Cater's Self-sustaining generator . . . . . 5 - 105
Dr Oleg Gritskevitch's COP=100 1.5 MW generator . . . . . 5 - 106

Chapter 6: Battery-Charging Pulsed Systems

Bedini Pulsing . . . . . 6 - 2
Roger Andrews' switching system . . . . . 6 - 2
Ron Knight's professional battery advice . . . . . 6 - 5
Ron Pugh's Charger. . . . . 6 - 5
Ossie Callanan's charging system . . . . . 6 - 18
The Self-charging system . . . . . 6 - 24
The Relay Coil Battery Charger. . . . . 6 - 25
The Re-wired Fan Charger . . . . . 6 - 26
The Automotive Coil Charger . . . . . 6 - 28
The Self-charging Motor . . . . . 6 - 30
The Alexkor multiple battery charger. . . . . 6 - 31
Howerd Halay's cold electricity conditioning . . . . . 6 - 39
The Ron Cole One-Battery Charger . . . . . 6 - 40
The Tesla Switch. . . . . 6 - 42
The self-charging motor . . . . . 6 - 43
The 'UFOpolitics' motor . . . . . 6 - 46
Some battery charging suggestions . . . . . 6 - 46

Chapter 7: Aerial Systems and Electrostatic Generators

Nikola Tesla's Aerial Plate Power System . . . . . 7 - 1
The Jes Ascanius replication of Tesla's Aerial Plate System . . . . . 7 - 8
The printed circuit board of Dragan Kljajic . . . . . 7 - 12
Thomas Henry Moray's System. . . . . 7 - 14
Herman Plauson's System. . . . . 7 - 31
Roy Meyer's System . . . . . 7 - 33
Raymond Phillips' RF System. . . . . 7 - 34
The 'Alexkor' Aerial System . . . . . 7 - 38
The 'TREC' Aerial System of Lawrence Rayburn . . . . . 7 - 39
The Patent Application of Harold Deyo . . . . . 7 - 41
The Electrostatic Generator of Charles Goldie . . . . . 7 - 51
The Electrostatic Generator of D. Le May. . . . . 7 - 55
The Electrostatic Generator of Onezime Breaux . . . . . 7 - 68

Chapter 8: Fuel-less Motors

The Bob Neal Motor. . . . . 8 - 1
The Scott Robertson air compressor. . . . . 8 - 4
The Leroy Rogers Motor . . . . . 8 - 6
The Vortex Tube . . . . . 8 - 21
The Eber Van Valkinburg Motor . . . . . 8 - 23
The Clem Motor . . . . . 8 - 23
The Papp Motor . . . . . 8 - 25
The Robert Britt Motor . . . . . 8 - 28
The Michael Eskeli Turbine . . . . . 8 - 28
James Hardy's Water-Jet Generator. . . . . 8 - 47
The Cahill & Scott Generator . . . . . 8 - 48

Chapter 9: Passive Systems

Hans Cohler's device . . . . . 9 - 1
Thomas Trawoeger's pyramid. . . . . 9 - 3
Peter Grandics' pyramid . . . . . 9 - 20
Les Brown's pyramid . . . . . 9 - 22
Joseph H. Cater's explanation of how a pyramid actually works . . . . . 9 - 29
Pier Luigi Ighina's "ERIM" and "Elios" passive devices. . . . . 9 - 31
The Joe Cell . . . . . 9 - 35
Co-axial Cable Electrets . . . . . 9 - 67

Chapter 10: Vehicle Systems

Booster design . . . 10 - 1
Simple DC boosters . . . 10 - 2
The “Smack's” booster . . . 10 - 17
The “Hotsabi” booster . . . 10 - 18
The Zach West booster . . . 10 - 18
Bill Williams’ “DuPlex” booster . . . 10 - 27
The ‘Hogg’ Electrolyser . . . 10 - 28
Advanced Boosters . . . 10 - 34
The Bob Boyce DC electrolyser . . . 10 - 34
Pulsed Water-Splitters . . . 10 - 35
Dave Lawton's Meyer Replication . . . 10 - 35
Dr Cramton's water-splitter . . . 10 - 54
Bob Boyce's Toroidal water-splitter . . . 10 - 60
John Bedini Water-Splitting . . . 10 - 61
Dave Lawton’s Auto-Tune Circuit . . . 10 - 62
Running an Electrical Generator on Water Alone . . . 10 - 63
Running an Unmodified Generator on Water Alone . . . 10 - 88
Stan Meyer's Water Injection System . . . 10 - 92
Peter Lindemann running a lawnmower on water alone . . . 10 - 97
Peter Lowrie’s High-Current Electrolyser System . . . 10 - 98
Water use in engines . . . 10 - 99
Ted Ewert’s Vortex Tube . . . 10 - 102
Cam timing Issues . . . 10 - 107
Robert Krupa’s ‘FireStorm’ Spark Plug . . . 10 - 108
Plasma Ignition . . . 10 - 109
Roger Maynard’s Water Vapour Injection System . . . 10 - 110
“Fuelsavers” streamlining . . . 10 - 112
The Ram Implosion Wing . . . 10 - 112
High mpg Carburettors . . . 10 - 112
Allen Caggiano’s High-mpg carburettor . . . 10 - 114
A Fuel Vapourisation System . . . 10 - 118
Vortex Fuel Reforming (GEET) . . . 10 - 119
Electric Power . . . 10 - 121
The Weird Nature of Water and Advanced Electrolysis . . . 10 - 123
The Electrolyser of Charles Garrett . . . 10 - 129
The Electrolyser of Archie Blue . . . 10 - 130
The HHO cell of Paul Zigouras . . . 10 - 132

Chapter 11: Other Devices and Theories

The Tesla Generators . . . 11 - 1
The Aspden Device . . . 11 - 3
Paulo & Alexandra Correa . . . 11 - 15
Professor Konstantin Meyl . . . 11 - 16
Tesla’s MHD unit . . . 11 - 16
The Unified Field Theory . . . 11 - 18
Maurice Cotterell’s Explanation of Gravity . . . 11 - 22
John R. R. Searle . . . 11 - 32
The Gravity Wave Detector . . . 11 - 33
The Butch Lafonte Motor / Generator . . . 11 - 35
The Joseph Newman’s COP = 8 Device . . . 11 - 40
Daniel Cook’s Induction Coil . . . 11 - 51
Michael Eskeli’s Fuel-less Heater . . . 11 - 53
Karl Schappeller’s Free-energy Device . . . 11 - 60
Condensation-Induced Water Hammer . . . 11 - 75
William Hyde's COP = 10 Electrostatic Power Generator . . . 11 - 76
The Suppression of Knowledge . . . 11 - 85

Chapter 12: Electronics Tutorial

Voltage . . . 12 - 1
Resistance . . . 12 - 2
Semiconductors – Transistors . . . 12 - 10
Diodes . . . 12 - 19
Alternating Current . . . 12 - 21
Coils (Inductors) . . . 12 - 21
The Ben Teal Motor . . . 12 - 23

Transformers . . . 12 - 21
Rectification and Power Supplies . . . 12 - 26
Multivibrators . . . 12 - 29
The Bistable . . . 12 - 29
The Monostable . . . 12 - 30
The Astable . . . 12 - 31
Inverters . . . 12 - 32
Digital Logic and Truth Tables . . . 12 - 32
The NAND gate and Gating . . . 12 - 34
The Latch . . . 12 - 37
The NE555 Timer Chip . . . 12 - 41
The 741 Op-amp Chip . . . 12 - 56
The SCR (Thyristor) . . . 12 - 59
The Triac . . . 12 - 59
A simple, versatile, Schmitt inverter signal generator . . . 12 - 62
The CD4022BC Divide-By-Eight Chip . . . 12 - 62
The cascaded CD4017B Divide-By-N circuitry . . . 12 - 64
The PIC chip . . . 12 - 66
Capacitors . . . 12 - 71
AC Circuitry, inductors and resonance . . . 12 - 73
Choosing unspecified component values . . . 12 - 78
Prototype Construction . . . 12 - 81
Test Equipment . . . 12 - 85
Power Supply Unit . . . 12 - 88
The Oscilloscope . . . 12 - 90
Measuring magnetic field strength . . . 12 - 91
The Weird Stuff . . . 12 - 92

Chapter 13: Doubtful Devices
Paul Baumann's "Thestakia" . . . 13 - 1
The Homopolar or "N-Machine" . . . 13 - 4
The Borderland Science Homopolar AC Generator . . . 13 - 5
Tesla's Homopolar Generator . . . 13 - 7
The "Romag" and "Mini-Romag" Generators . . . 13 - 8
Cold Fusion . . . 13 - 11
Moller's Atomic Hydrogen Generator . . . 13 - 11
Muammer Yildiz's "Ocean Star" Electrical Generator . . . 13 - 13
Jesse McQueen . . . 13 - 20
The Nitro Cell ("D18") . . . 13 - 23
The HydroStar and HydroGen . . . 13 - 42
Hydrogen from Aluminium . . . 13 - 46
Francois Cornish . . . 13 - 47
Ultrasonic water-splitting . . . 13 - 48
The MEG . . . 13 - 49
Dave Lawton's assymetric MEG variation . . . 13 - 51
The Hitch-Hiker's MEG description . . . 13 - 52
The Devices of Hans Coler . . . 13 - 62

Chapter 14: Renewable Energy Devices
Heaters . . . 14 - 1
The Wood-gas Stove . . . 14 - 1
Mr Teslonian's wood-gas stove / electrical generator / fuel producer . . . 14 - 4
The solid-fuel electricity stove . . . 14 - 5
Henry Paine's HHO patent . . . 14 - 7
Sang Nam Kim's HHO heaters . . . 14 - 7
The H-Cat HHO heater from Justin Church . . . 14 - 10
Eugene Frenette . . . 14 - 11
Eugene Perkins . . . 14 - 12
The Peter Davey Heater . . . 14 - 15
The Jean-Christophe Dumas Heater . . . 14 - 17
The re-wired halogen heater . . . 14 - 20
Home-build Wind Generator . . . 14 - 21
William McDavid's Wind-Power or Water-Power Generator . . . 14 - 22
Frank Herbert's Wind Generator . . . 14 - 23
Mead and Holmes Power System . . . 14 - 25

Solar Ovens . . . . . 14 - 26
Solar Water Pasteurisation. . . . . 14 - 34
Solar Water Stills. . . . . 14 - 52
Sonic Water Pumps . . . . . 14 - 60
The Ram Pump . . . . . 14 - 62
Discouraging mosquitoes and small flies.. . . . . 14 - 64
Wave Power . . . . . 14 - 65
Cooling Using Heat . . . . . 14 - 69
Solar Panels . . . . . 14 - 71
Low-cost Cooling Systems. . . . . 14 - 100

Chapter 15: The Present Situation
Why free-energy devices are not in your local shop . . . . . 15 - 1

Chapter 16: Questions and Answers
What to Build . . . . . 16 - 1
Questions on Electrolysis . . . . . 16 - 7
Questions on Magnet Motors . . . . . 16 - 17
Questions on Don Smith's Devices. . . . . 16 - 19
Questions on General Topics . . . . . 16 - 22

Appendix
US and UK Wire sizes and capacities . . . . . A -1
Frank Fecera's permanent magnet motor patent . . . . . A - 2
Howard Johnson's permanent magnet motor patent . . . . . A - 46
Harold Ewing's Carousel permanent magnet-generator patent. . . . . A - 56
The Pavel Imris minimal power lighting system patent . . . . . A - 74
The Colman/Seddon-Gillespie 70-year battery patent . . . . . A - 83
The Jon Sok An Lenz-less electrical generator patent . . . . . A - 87
The Molina Martinez self-powered electrical generator patent . . . . . A - 100
Michael Ognyanov's solid-state electrical generator patent . . . . . A - 113
Edwin Gray's electric motor patent . . . . . A - 119
Edwin Gray's electric power supply patent . . . . . A - 142
The Adams-Aspden electrical motor-generator patent . . . . . A - 148
William Barbat's self-powered electrical generator patent . . . . . A - 164
John Reardon's AC generator patent . . . . . A - 196
Geoffrey Spence's self-powered electrical generator patent . . . . . A - 213
Robert Alexander's COP = 2.93 electrical generator patent. . . . . A - 228
Shigeaki Hayasaka's electrical generator patent . . . . . A - 236
Larry Jamison's electrical generator patent application . . . . . A - 252
Teruo Kawai's COP>1 electric motor patent . . . . . A - 259
Joseph Newman's Energy Generator patent. . . . . A - 280
Philip Brody's very high output ceramic solar devices patent . . . . . A - 298
Charles Flynn's controlled magnetic devices patent . . . . . A - 338
The **M**otionless **E**lectromagnetic **G**enerator patent . . . . . A - 411
Dan Davidson's acoustic-magnetic electrical generator patent. . . . . A - 427
John Bedini's battery-charging patent . . . . . A - 434
John Bedini's motor-generator patent . . . . . A - 445
John Bedini's pulse-charging system patent . . . . . A - 456
Richard Weir and Carl Nelson's battery replacement patent . . . . . A - 470
Hermann Plauson's aerial power systems patent . . . . . A - 485
Roy Meyers' Electricity-producing device patent . . . . . A - 524
Paulo and Alexandra Correa's free-electricity patent . . . . . A - 532
Paulo and Alexandra Correa's energy conversion patent . . . . . A - 561
The Mead-Nachamkin ZPE to electricity conversion patent . . . . . A - 603
Stanley Meyer's Water Fuel patent 4,936,961 . . . . . A - 620
Stanley Meyer's hydrogen injection system for vehicles patent 4,389,981 . . . . . A - 627
Stanley Meyer's hydrogen gas burner patent 4,421,474. . . . . A - 637
Stanley Meyer's hydrogen generation and enhancement patent 5,149,407 . . . . . A - 642
Stanley Meyer's water fuel generator patent CA 2,067,735 . . . . . A - 659
Stanley Meyer's WFC control circuitry patent WO 92/07861 . . . . . A - 670

Stephen Meyer’s water-splitting patent application 2005/0246059 . . . . . A - 680
Henry Puharich’s water-splitting patent 4,392,230 . . . . . A - 689
Shigeta Hasebe’s spiral electrolyser patent . . . . . A - 719
Stephen Chambers’ hydroxy generator patent (Xogen Power Inc.) . . . . . A - 725
Charles Garrett’s water carburettor patent . . . . . A - 740
Archie Blue’s electrolyser patent . . . . . A - 748
Ruggero Santilli’s plasma arc electrolysis patent . . . . . A - 754
Chak Chang’s low-voltage low-temperature plasma patent application. . . . . A - 764
Juan Aguero’s water-engine patent application . . . . . A - 823
Stephen Horvath’s water-powered car patent . . . . . A - 830
Christopher Eccles’ water-splitting cell patent . . . . . A - 857
Spiro Spiros’ COP>1 electrolyser patent . . . . . A - 864
Henry Paine’s hydroxy gas conversion patent . . . . . A - 901
Boris Volfson’s gravity drive patent. . . . . A - 904
Charles Pogue’s first high-mpg carburettor patent . . . . . A - 919
Charles Pogue’s second high-mpg carburettor patent . . . . . A - 927
Charles Pogue’s third high-mpg carburettor patent . . . . . A - 932
Ivor Newberry’s high-mpg carburettor patent. . . . . A - 940
Robert Shelton’s high-mpg carburettor patent . . . . . A - 945
Harold Schwartz’s high-mpg carburettor patent . . . . . A - 949
Oliver Tucker’s high-mpg carburettor patent . . . . . A - 952
Thomas Ogle’s high-mpg carburettor patent . . . . . A - 955
Stephen Kundel’s permanent magnet motor . . . . . A - 968
Charles Flynn’s permanent magnet motor . . . . . A - 992
Claude Mead and William Holmes’ wind power storage system . . . . . A - 1021
The Richard Willis COP>>1 electrical generator . . . . . A - 1028
Graham Gunderson's pulsed toroid generator . . . . . A - 1038
Mark McKay's investigation into Edwin Gray's technology . . . . . A - 1046
Mike Brady’s “Perendev” magnet motor patent . . . . . A - 1107
Donald Kelly’s magnet motor patent . . . . . A - 1117
Bob Neal’s compressor patent . . . . . A - 1125
Leroy Rogers’ compressed air engine patent . . . . . A - 1132
Eber Van Valkinburg’s compressed air/oil engine patent . . . . . A - 1148
Josef Papp’s inert-gas engine patent . . . . . A - 1153
Robert Britt’s Inert-gas engine patent . . . . . A - 1186
Floyd Sweet’s “VTA” Electrical Generator . . . . . A - 1196
Meguer Kalfaian’s Electrical Generator . . . . . A - 1210
Theodore Annis and Patrick Eberly’s Motionless Generator . . . . . A - 1220
William McDavid’s Wind-Power or Water-power Generator. . . . . A - 1232
The Self-powered Generators of Barbosa and Leal . . . . . A - 1252
Web links to Scientific Papers and Videos . . . . . A - 1283

Alphabetical Index of Devices

AC generator patent, John Reardon . . . A - 196
Acoustic water pumps: Bellocq, Dickinson and Benson . . . 14 - 61
Advanced Electrolysis, Moray B. King . . . 10 - 120
Aerial power systems patent, Hermann Plauson . . . A - 485
Aerial system, Frank Prentice . . . 5 - 1
Aerial system, Jes Ascanius . . . 7 - 9
Aerial system, Lawrence Rayburn . . . 7 - 39
Aerial system, Nikola Tesla . . . 7 - 1
Aerial system, Raymond Phillips . . . 7 - 34
Aerial system, Thomas Henry Moray . . . 7 - 14
Air Vortex Turbine, Ted Ewert . . . 10 - 89
Aspden Effect, Harold Aspden . . . 5 - 66
Asymmetrical Motionless Generator, Dave Lawton . . . 13 - 51
Atomic Hydrogen generator, William Lyne . . . 13 - 11
Automotive Relay battery pulser, Imhotep . . . 6 - 28
Battery-charging patent, John Bedini . . . A - 434
Battery-pulser, Alexkor . . . 6 - 31
Battery-pulser, Bozidar Lisac . . . 5 - 37
Battery-pulser, John Bedini . . . 6 - 2
Battery-pulser, Ossie Callanan . . . 6 - 18
Battery-pulser, Ron Pugh . . . 6 - 5
Battery replacement patent, Richard Weir and Carl Nelson . . . A - 470
Battery technology, Ronald Knight . . . 6 - 5
Blood electrifier, Bob Beck . . . 11 - 88
Boosters, Various . . . 10 - 1
Britt Engine . . . 8 - 28
Buoyancy patent, Renato Ribera . . . 4 - 48
Capacitor Battery Pulser, Ron Cole . . . 6 - 40
Caravan Power System, Claude Mead and William Holmes . . . 14 - 21
Car relay pulse charger, Imhotep . . . 6 - 28
Chain Drive, Jerzy Zbikowski . . . 4 - 17
Clem engine, Richard Clem . . . 8 - 23
Co-axial Cable Electrets . . . 9 - 67
Cold Electricity Coil of ‘UFOpolitics’ . . . 3 - 182
Colloidal Silver, Bob Beck . . . 11 - 88
Compressed-air engine, Bob Neal . . . 8 - 1
Compressed-air engine, Leroy Rogers . . . 8 - 6
Compressed-air tank, Scott Robertson . . . 8 - 4
Compressed-air/oil engine, Eber Van Valkenburg . . . 8 - 23
Controlled magnetic devices patent, Charles Flynn . . . A - 338
Cooling systems for low-power use . . . 14 - 96
COP = 2.93 electrical generator patent, Robert Alexander . . . A - 228
COP>1 electric motor patent, Teruo Kawai . . . A - 259
COP>1 electrolyser patent, Spiro Spiros . . . A - 864
Davey water heater, Peter Davey . . . 14 - 15
Direct Coupled Motors . . . 2 - 67
Drinking Water Systems . . . 14 - 48
Electrets . . . 9 - 68
Electrical energy from air, Nikola Tesla . . . 11 - 1
Electrical free-energy generation, Harold Aspden . . . 11 - 3
Electrical generator, Alfred Hubbard . . . 5 - 90
Electrical generator, Arthur Cahill & John Scott . . . 8 - 48
Electrical generator, Joseph Cater . . . 5 - 103
Electrical generator, Alberto Molina-Martinez . . . 5 - 89
Electrical generator, Clemente Figuera . . . 3 - 6
Electrical generator, Borderland Science . . . 13 - 8
Electrical generator, Ecklin-Brown . . . 1 - 4
Electrical generator, Dr Oleg Gritskevitch . . . 5 - 104
Electrical generator patent application, Larry Jamison . . . A - 252
Electrical generator, Meyer-Mace . . . 3 - 103
Electrical generator, Raoul Hatem . . . 2 - 48
Electrical generator, Raymond Kromrey . . . 2 - 16
Electrical generator, Richard Willis . . . 3 - 75

Electrical generator, self-powered, Mr Wilson . . . . . 4 - 7
Electrical Generator, ‘silverhealtheu’. . . . . 3 - 77
Electrical generator patent, Shigeaki Hayasaka . . . . . A - 236
Electrical Generator, Stan Meyer . . . . . 3 - 191
Electrical Generator, Tariel Kapanadze . . . . . 3 - 176
Electrical motor-generator patent, Adams-Aspden . . . . . A - 148
Electricity-producing device patent, Roy Meyers . . . . . A - 524
Electricity-producing solid-fuel stove . . . . . 14 - 5
Electric motor, Ben Teal . . . . . 1 - 9
Electric motor patent, Edwin Gray . . . . . A - 119
Electrical motor, Georges Mourier . . . . . 2 - 29
Electrical motor, Edwin Gray . . . . . 5 - 81
Electric motor, Robert Adams . . . . . 2 - 1
Electric motor, Teruo Kawai . . . . . 2 - 25
Electric power supply patent, Edwin Gray . . . . . A - 142
Electrolyser, Archie Blue . . . . . 10 - 128
Electrolyser, Bob Boyce . . . . . 5 - 46
Electrolyser, Bob Boyce . . . . . 10 - 33
Electrolyser, Bob Boyce . . . . . 10 - 59
Electrolyser, Charles Garrett . . . . . 10 - 126
Electrolyser, Paul Zigouras . . . . . 10 - 129
Electrolyser, Peter Lowrie . . . . . 10 - 95
Electrolyser, Selwyn Hughes . . . . . 10 - 95
Electrolyser, Zach West . . . . . 10 - 18
Electrolyser patent, Archie Blue. . . . . A - 748
Electrostatic Power Generator, Onezime Breaux . . . . . 7 - 68
Electrostatic Power Generator, Charles Goldie . . . . . 7 - 51
Electrostatic Power Generator, William Hyde . . . . . 11 - 76
Electrostatic Power Generator, D. Le May . . . . . 7 - 55
Elios device, Pier Luigi Ighina . . . . . 9 - 34
Energy-conversion patent, Paulo and Alexandra Correa . . . . . A - 561
Energy Generator patent, Joseph Newman . . . . . A - 280
ERIM device, Pier Luigi Ighina . . . . . 9 - 31
E-Stress Power Generator, harisingh. . . . . 3 - 203
Fan Battery Pulser, Imhotep . . . . . 6 - 26
FLEET power toroid . . . . . 5 - 56
Francois Cornish hydrogen generator . . . . . 13 - 47
Free-electricity patent, Paulo and Alexandra Correa . . . . . A - 532
Free-energy device, Karl Schappeller . . . . . 11 - 60
Fuelsavers . . . . . 10 - 109
Fuji camera circuit adaption . . . . . 5 - 62
GEET fuel re-former system, Paul Pantone . . . . . 10 - 107
Generator running on water alone . . . . . 10 - 63
Gravity-chain device, Murilo Luciano . . . . . 4 - 27
Gravity device, Amr Al-Hossary. . . . . 4 - 26
Gravity drive patent, Boris Volfson . . . . . A - 904
Gravity-tapping generator, Chas Campbell . . . . . 4 - 4
Gravity Wave Detector, Dave Lawton . . . . . 11 - 33
Gravity wheel, Bhaskara. . . . . 4 - 51
Gravity wheel, Dale Simpson . . . . . 4 - 22
Gravity wheel, Lawrence Tseung . . . . . 4 - 1
Gravity wheel, Mikhail Dmitriev . . . . . 4 - 20
Heater, Fuel-less, Michael Eskeli . . . . . 11 - 53
Heater, COP=17 by Rosemary Ainslie . . . . . 5 - 102
Heaters . . . . . 14 - 1
Heater, Magnetic. . . . . 16 - 39
H-Cat HHO heater, Justin Church . . . . . 14 - 10
Hidro Generator, James Kwok . . . . . 4 - 35
High-mpg carburettor patent, Charles Pogue . . . . . A - 919
High-mpg carburettor patent, Charles Pogue . . . . . A - 927
High-mpg carburettor patent, Charles Pogue . . . . . A - 932
High-mpg carburettor patent, Harold Schwartz . . . . . A - 949
High-mpg carburettor patent, Ivor Newberry . . . . . A - 940
High-mpg carburettor patent, Oliver Tucker . . . . . A - 952
High-mpg carburettor patent, Robert Shelton . . . . . A - 945

High-mpg carburettor patent, Thomas Ogle . . . . . A - 955
Hinged-Plate Gravity Device, Amr Al-Hossary . . . . . 4 - 25
Hogg electrolyser . . . . . 10 - 28
Homopolar generator, Michael Faraday . . . . . 13 - 4
Hotsabi booster, "Hotsabi" . . . . . 10 - 18
Hydraulic Ram Pump . . . . . 14 - 58
Hydrogen from Aluminium . . . . . 13 - 46
Hydrogen gas burner patent 4,421,474, Stanley Meyer . . . . . A - 637
Hydrogen generation and enhancement patent 5,149,407, Stanley Meyer . . . . . A - 642
Hydrogen injection system for vehicles patent 4,389,981, Stanley Meyer . . . . . A - 627
HydroStar and HydroGen devices . . . . . 13 - 42
Hydroxy gas conversion patent, Henry Paine . . . . . A - 901
Hydroxy generator patent, Stephen Chambers . . . . . A - 725
Hydroxy-burning heaters . . . . . 14 - 7
Induction Coil generator, Daniel Cook . . . . . 11 - 51
Inert-gas engine, Josef Papp . . . . . 8 - 25
Inert-gas engine, Robert Britt . . . . . 8 - 28
Inverter system, Jesse McQueen . . . . . 13 - 20
Joe Cell, Joe Nobel . . . . . 9 - 35
Joseph Newman Motor, Joseph Newman . . . . . 11 - 40
Joule Ringer circuits, Lasersaber . . . . . 5 - 80
Joule Thief advanced circuits . . . . . 5 - 66
Joule Thief lighting circuits . . . . . 5 - 56
Lead-out energy, Lawrence Tseung . . . . . 4 - 1
Lawnmower running on water alone . . . . . 10 - 94
Lenz-less electrical generator patent, Jon Sok An . . . . . A - 87
Long-life battery, Colman / Seddon-Gillespie . . . . . 3 - 110
Low-voltage low-temperature plasma patent application, Chak Chang . . . . . A - 764
Magnetic shielding from Pasi Mäkilä . . . . . 1 - 55
Magnet motor-generator, permanent 5kW, Shen He Wang . . . . . 1 - 1
Magnet Motor, Art Porter . . . . . 2 - 57
Magnet motor, Carousel . . . . . 1 - 7
Magnet-generator patent, Carousel . . . . . A - 56
Magnet motor patent, Charles Flynn . . . . . 1 - 15
Magnet Motor, Dietmar Hohl . . . . . 1 - 29
Magnet Motor, Jes Ascanius . . . . . 1 - 29
Magnet motor patent, Frank Fecera . . . . . A - 2
Magnet Motor, George Soukup . . . . . 1 - 26
Magnet motor, Howard Johnson . . . . . 1 - 6
Magnet motor patent, Howard Johnson . . . . . A - 46
Magnet motor, John Jines . . . . . 1 - 11
Magnet motor, Donald Kelly . . . . . 1 - 54
Magnet motor, Perendev (Mike Brady) . . . . . 1 - 55
Magnet motor, Robert Tracy . . . . . 1 - 9
Magnet motor, simple . . . . . 1 - 30
Magnet motor, Steorn . . . . . 1 - 24
Magnet motor, Stephen Kundel . . . . . 1 - 12
Magnet motor, Victor Diduck . . . . . 1 - 56
Magnetic Conversion, Don Smith . . . . . 3 - 112
Magnetic Conversion, Don Smith . . . . . 5 - 52
Magnetic flux, Charles Flynn . . . . . 3 - 1
Magnetic frame, Lawrence Tseung . . . . . 3 - 1
Magnetic frame, Thane Heins . . . . . 3 - 3
Magnetic Plate, The ‘Gegene’ . . . . . 3 - 173
Magnetohydrodynamic drive, Nikola Tesla . . . . . 11 - 17
MEG, Tom Beardon et al. . . . . 13 - 49
Minimal power lighting system patent, Pavel Imris . . . . . A - 74
Motionless Generators of Kelichiro Asaoka . . . . . 3 - 84
Motionless Generator of Dietmar Kunel . . . . . 3 - 77
Motionless Generator of Valeri Ivanov . . . . . 3 - 84
Motionless **E**lectromagnetic **G**enerator patent . . . . . A - 411
Motor-generator, Butch Lafonte . . . . . 11 - 35
Motor-generator patent, John Bedini . . . . . A - 445
Motor-generator patent, Georges Mourier . . . . . 2 - 29
N-machine generator, Michael Faraday . . . . . 13 - 2

Neal Compressed-air Engine . . . . . 8 - 1
Nitro Cell . . . . . 13 - 22
Ocean-Star generator, Muammer Yaldiz . . . . . 13 - 13
One-battery switch, Ron Cole . . . . . 6 - 40
Optical Amplifier, Pavel Imris . . . . . 3 - 106
Pancake Coil, Nikola Tesla . . . . . 5 - 66
Pendulum/lever device, Veljko Milkovic . . . . . 4 - 25
Phase-Lock-Loop water-splitting circuit . . . . . 10 - 62
Plasma arc electrolysis patent, Ruggero Santilli . . . . . A - 754
Power System for Caravans, Mead and Holmes . . . . . 14 - 21
Power tube, Edwin Gray . . . . . 5 - 81
Pulse-charging system patent, John Bedini . . . . . A - 456
Pulsed flywheel generator, Jacob Bitsadze . . . . . 4 - 6
Pulsed flywheel generator, Chas Campbell . . . . . 4 - 4
Pulsed flywheel generator, John Bedini . . . . . 4 - 12
Pulsed flywheel generator, Jim Watson . . . . . 4 - 11
Pulsed flywheel generator, Lawrence Tseung . . . . . 2 - 50
Pyramid, James Brock . . . . . 9 - 15
Pyramid, Joseph H. Cater . . . . . 9 - 29
Pyramid, Les Brown . . . . . 9 - 22
Pyramid, Paulo and Alexandra Correa . . . . . 11 - 15
Pyramid, Peter Grandics . . . . . 9 - 21
Pyramid, Thomas Trawoeger . . . . . 9 - 3
Radium power generation, Nikola Tesla . . . . . 11 - 2
RF aerial to DC power, Raymond Phillips . . . . . 7 - 34
Raising DC motor efficiency . . . . . 2 - 57
Ram Implosion Wing, Robert Patterson . . . . . 10 - 109
Ram Pump . . . . . 14 - 58
Refrigeration through heating, Albert Einstein . . . . . 14 - 66
Romag generator, Magnetic Energy . . . . . 13 - 8
RotoVerter, Hector Torres . . . . . 2 - 39
Screened transformer, Joseph Boyd . . . . . 3 - 160
Searle Effect Device, John R. R. Searle . . . . . 11 - 32
Self-powered water-jet electrical generator, James Hardy . . . . . 2 - 27
Self-powered water-jet electrical generator, James Hardy . . . . . 4 - 13
Self-powered water-jet electrical generator, James Hardy . . . . . 8 - 47
Self-powered water-jet electrical generator, James Hardy . . . . . 2 - 27
Self-powered electrical generator, Nilson Barbosa and Cleriston Leal . . . . . 3 - 27
Self-powered electrical generator, Carlos Benitez . . . . . 5 - 20
Self-powered electrical generator, Joseph H Cater . . . . . 5 - 103
Self-powered electrical generator, José Luis García del Castillo . . . . . 4 - 7
Self-powered electrical generator patent, Molina Martinez . . . . . A - 100
Self-powered electrical generator patent, Tariel Kapanadze . . . . . 3 - 176
Self-powered electrical generator, Stephan W. Leben . . . . . 3 - 99
Self-powered electrical generator patent, William Barbat . . . . . A - 164
Seventy-year battery patent, Colman/Seddon-Gillespie . . . . . A - 83
Smack's Booster, Eletrik . . . . . 10 - 17
Solar ovens, BYU . . . . . 14 - 22
Solar panels . . . . . 14 - 67
Solar Still Water-makers . . . . . 14 - 48
Spiral electrolyser patent, Shigeta Hasebe . . . . . A - 719
Static electricity, Lorrie Matchett . . . . . 3 - 61
Stromerzeuger, Hans Coler . . . . . 13 - 62
Stromerzeuger, Hans Coler . . . . . 9 - 1
Tesla Coil, Correas . . . . . 11 - 15
Tesla Coils back-to-back, Nick Giannopoulous . . . . . 3 - 157
Tesla Switch, Nikola Tesla . . . . . 5 - 6
Tesla Switch, Nikola Tesla . . . . . 6 - 42
Testatika, Paul Baumann . . . . . 13 - 1
The Thirty Kilowatt Gravity-Powered Motor . . . . . 4 - 50
Toroidal Power Unit, Bob Boyce . . . . . 5 - 46
Torque multiplier, William Skinner . . . . . 4 - 53
Turbine, Michael Eskeli . . . . . 8 - 28
Ultrasonic hydrogen generator . . . . . 13 - 48
Very high output ceramic solar devices patent, Philip Brody . . . . . A - 298

Vortex Fuel Reformers . . . . . 10 - 116
Vortex tube . . . . . 8 - 21
VTA, Floyd Sweet . . . . . 3 - 102
VTA, Floyd Sweet . . . . . 5 - 102
VTA, Floyd Sweet . . . . . A - 1196
Waste spark handling . . . . . 10 - 70
Water Acoustic Pump, Toribio Bellocq . . . . . 14 - 55
Water Acoustic Pump, Richard Dickinson . . . . . 14 - 57
Water Acoustic Pump, Arthur Bentley . . . . . 14 - 57
Water carburettor patent, Charles Garrett . . . . . A - 740
Water clusters, Moray B. King . . . . . A - 740
Water-engine patent application, Juan Aguero . . . . . A - 823
Water Fuel Cell, Dr Scott Cramton . . . . . 10 - 53
Water Fuel Cell, Dave Lawton . . . . . 5 - 4
Water Fuel Cell, Dave Lawton . . . . . 10 - 34
Water fuel generator patent CA 2,067,735, Stanley Meyer . . . . . A - 659
Water Fuel patent 4,936,961, Stanley Meyer . . . . . A - 620
Water-jet self-powered generator, James Hardy . . . . . 2 - 27
Water-jet self-powered generator, James Hardy . . . . . 8 - 47
Water-maker, Calice Courneya . . . . . 14 - 56
Water-maker, Elmer Grimes . . . . . 14 - 56
Water-makers . . . . . 14 - 48
Water Heater, Peter Davey . . . . . 14 - 15
Water-powered car patent, Stephen Horvath . . . . . A - 830
Water-Powered Generator, William McDavid . . . . . 14 - 18
Water-Powered Generator, William McDavid . . . . . A - 1233
Water-splitting cell patent, Christopher Eccles . . . . . A - 857
Water-splitting patent application 2005/0246059, Stephen Meyer . . . . . A - 680
Water-splitting patent 4,392,230, Henry Puharich . . . . . A - 689
Water vapour injection systems . . . . . 10 - 107
Water injection systems, Stan Meyer . . . . . 10 - 89
WFC control circuitry patent WO 92/07861, Stanley Meyer . . . . . A - 670
Wind Generator, Dan Bartmann and Dan Fink . . . . . 14 - 17
Wind Generator, Frank Herbert . . . . . 14 - 19
Wind-Powered Generator, William McDavid . . . . . 14 - 18
Wind-Powered Generator, William McDavid . . . . . A - 1233
Wind Power Storage System, Claude Mead and William Holmes . . . . . A - 1021
Wire sizes and capacities . . . . . A -1
ZPE to electricity conversion patent, Mead-Nachamkin . . . . . A - 603

Alphabetical Index of People

Adams, Robert . . . . . 2 - 1
Adams, Robert . . . . . A – 148
Aguero, Juan . . . . . A - 823
Alexander, Robert . . . . . A - 228
Alexkor . . . . . 6 - 31
Alexkor . . . . . 3 - 26
Al-Hossary, Amr . . . . . 4 - 26
Ainslie, Rosemary . . . . . 5 - 102
An, Jon Sok . . . . . A - 87
Andrews, Roger . . . . . 6 - 3
Annis, Theodore . . . . . 3 - 29
Asaoka, Kelichiro . . . . . 3 - 85
Ascanius, Jes . . . . . 11 - 3
Aspden, Harold . . . . . 11 - 3
Aspden, Harold . . . . . A - 148
Aum, Johnny . . . . . 5 - 70
Barbat, William . . . . . A - 164
Barbosa, Nilson . . . . . 3 - 27
Bartmann, Dan . . . . . 14 - 17
Baumann, Paul . . . . . 13 - 1
Beardon, Tom . . . . . 13 - 49
Bearden, Tom . . . . . A - 411
Beck, Bob . . . . . 11 - 88
Beck, Bob . . . . . 12 - 44
Bedini, John . . . . . 4 - 12
Bedini, John . . . . . 5 - 4
Bedini, John . . . . . 6 - 2
Bedini, John . . . . . A - 434
Bedini, John . . . . . A - 445
Bedini, John . . . . . A - 456
Belkired, Mohamed . . . . . 3 - 112
Bellocq, Toribio . . . . . 14 - 56
Benitez, Carlos . . . . . 5 - 20
Bentley, Arthur . . . . . 14 - 56
Bhaskara . . . . . 4 - 51
Bitsadze, Jacob . . . . . 4 - 6
Borderland Science . . . . . 13 - 6
Blue, Archie . . . . . 10 - 127
Blue, Archie . . . . . A - 748
Boyce, Bob . . . . . 5 - 46
Boyce, Bob . . . . . 10 - 33
Boyd, Joseph . . . . . 3 – 160
Brady, Mike . . . . . 1 - 55
Brady, Mike . . . . . A - 1107
Breaux, Onezime . . . . . 7 - 68
Brinkley, William . . . . . 13 - 46
Britt, Robert . . . . . 8 - 28
Britt, Robert . . . . . A - 1186
Brock, James . . . . . 9 - 16
Brody, Philip . . . . . A - 298
Brown, Les . . . . . 9 - 23
Bundock, Jim . . . . . 13 - 40
Byzehr, Joseph . . . . . 4 - 6
Caggiano, Allen . . . . . 10 - 111
Cahill, Arthur . . . . . 8 - 34
Callanan, Ossie . . . . . 6 - 18
Campbell, Chas . . . . . 4 - 4
Cao, Ming . . . . . 3 - 140
Cater, Joseph H. . . . . 5 - 92
Cater, Joseph H. . . . . 5 - 105
Cater, Joseph H. . . . . 9 - 29
Cater, Joseph H. . . . . 11 - 85
Chambers, Stephen . . . . . A - 725

Chang, Chak . . . . . A - 764
Church, Justin . . . . . 14 - 10
Clem, Richard . . . . . 8 - 23
Coe, Graham . . . . . 9 - 14
Cole, Marvin . . . . . 5 - 81
Cole, Ron . . . . . 6 - 40
Coler, Hans . . . . . 13 - 62
Coler, Hans . . . . . 9 - 1
Colman, Harold . . . . . 3 - 110
Colman, Harold . . . . . A - 83
Cook, Daniel . . . . . 11 - 51
Cornish, Francois . . . . . 13 - 45
Correa, Paulo and Alexandra . . . . . 11 - 15
Correa, Paulo and Alexandra . . . . . A - 561
Cotterell, Maurice . . . . . 11 - 22
Cramton, Dr Scott . . . . . 10 - 53
Davey, Peter . . . . . 14 - 15
Davidson, Dan . . . . . A - 427
Davson, Cryil . . . . . 11 - 60
Deyo, Harold . . . . . 7 - 41
Diamanti, Gabriele . . . . . 14 - 51
Dickinson, Richard . . . . . 14 - 57
Diduck, Victor . . . . . 1 - 56
Dmitriev, Mikhail . . . . . 4 - 20
Drbal, Karel . . . . . 9 - 7
Eberly, Patrick . . . . . 3 - 66
Eccles, Christopher . . . . . A - 857
Ecklin, John W. . . . . 1 - 4
Einstein, Albert . . . . . 14 - 65
Electrodyne Corporation . . . . . 5 - 5
EnergeticForum . . . . . 10 - 92
Eskeli, Michael . . . . . 11 - 53
Eskeli, Michael . . . . . 8 - 28
EVGRAY, Yahoo forum . . . . . 2 - 33
Ewert, Ted . . . . . 10 - 99
Ewing, Harold . . . . . A - 56
Faraday, Michael . . . . . 13 - 4
Fecera, Frank . . . . . A - 2
Figuera, Clemente . . . . . 3 - 6
Fink, Dan . . . . . 14 - 17
Flynn, Charles . . . . . 1 - 15
Flynn, Charles . . . . . 3 - 1
Flynn, Charles . . . . . A - 338
Flynn, Charles . . . . . A - 992
Ford, Walter . . . . . 5 - 54
Frenette, Eugene . . . . . 14 - 11
Garrett, Charles . . . . . 10 - 126
Gadgetmall . . . . . 5 - 62
Garrett, Charles . . . . . A - 740
Giannopoulos, Nikanor . . . . . 3 - 157
Goldie, Charles . . . . . 7 - 51
Grandics Peter . . . . . 9 – 21
Gray, Ashley . . . . . 3 - 103
Gray, Edwin . . . . . 5 - 84
Gray, Edwin . . . . . A - 119
Gray, Edwin . . . . . A - 142
Gray, Edwin . . . . . A - 1046
Grimes Elmer . . . . . 14 - 55
Gritskevitch, Dr Oleg . . . . . 5 - 104
Gunderson, Graham . . . . . A - 1038
Hardy, James . . . . . 2 - 27
Hardy, James . . . . . 4 - 14
Hardy, James . . . . . 8 - 47
Hasebe, Shigeta . . . . . A - 719
Hatem, Raoul . . . . . 2 - 48

Hayasaka, Shigeaki . . . . . A - 236
Heath, Brian. . . . . 6 - 26
Heins, Thane . . . . . 3 - 3
Herbert, Frank . . . . . 14 - 18
Hohl, Dietmar . . . . . 1 - 29
Holdgate, Ed . . . . . 10 - 33
Holmes, William . . . . . 14 - 30
Holmes, William . . . . . A - 1021
Horvath, Stephen . . . . . A - 830
Hubbard, Alfred . . . . . 5 - 92
Hyde, William . . . . . 11 - 76
Hydrogen Garage . . . . . 10 - 34
Ighina, Pier Luigi . . . . . 9 - 31
Imhotep . . . . . 6 - 25
Imhotep . . . . . 6 - 28
Imris, Pavel . . . . . 3 - 105
Imris, Pavel . . . . . A - 74
Ivanov, Valeri . . . . . 3 - 84
Jamison, Larry . . . . . A - 252
Jeanna . . . . . 5 - 58
Jines, John . . . . . 1 - 11
Johnson, Howard . . . . . 1 - 5
Johnson, Howard . . . . . A - 46
José Luis García del Castillo . . . . . 4 - 7
Kalfaian, Meguer . . . . . A - 1210
Kapanadze, Tariel . . . . . 3 - 176
Kapanadze, Tariel . . . . . 5 - 53
Kawai, Teruo . . . . . 2 - 25
Kawai, Teruo . . . . . A - 259
Kelly, D. A. . . . . 13 - 1
Kelly, D. A. . . . . 1 - 54
Kelly, D. A. . . . . A - 1117
Kenny, James . . . . . 13 - 49
Kim, Sang Nam. . . . . 14 - 7
King, Moray B. . . . . 7 - 20
King, Moray B. . . . . 10 - 122
Kljajic, Dragan . . . . . 7 - 12
Knight, Ronald . . . . . 6 - 5
Kromrey, Raymond . . . . . 2 - 16
Krupa, Robert. . . . . 10 - 105
Kundel, Stephen . . . . . 1 - 12
Kunel, Dietmar . . . . . 3 - 77
Kwok, James . . . . . 4 - 35
Lafonte, Butch . . . . . 11 - 35
‘Lasersaber’ . . . . . 5 - 80
Lawton, Dave . . . . . 13 - 51
Lawton, Dave . . . . . 5 - 4
Lawton, Dave . . . . . 10 - 34
Lawton, Dave . . . . . 11 - 33
Leal, Cleriston .. . . . . 3 - 27
Leben, Stephan W. . . . . 3 - 99
Le May, D.. . . . . 7 - 55
‘Lidmotor’ . . . . . 2 - 12
‘Lidmotor’ . . . . . 5 - 58
Lindemann, Peter . . . . . 2 - 57
Lindemann, Peter . . . . . 10 - 94
Lisac, Bozidar. . . . . 5 - 37
Lowrie, Peter . . . . . 10 - 95
Luciano, Murilo . . . . . 4 - 27
Lyne, William . . . . . 13 - 7
Mace, Yves . . . . . 3 - 109
Mäkilä, Pasi . . . . . 1 - 55
Martinez, Molina . . . . . A - 100
Matchett, Lorrie. . . . . 3 - 61
Maynard, Roger . . . . . 10 - 107

McFreey, William . . . 3 - 181
McKay, Mark . . . A - 1046
McQueen, Jesse . . . 13 - 20
Mead, Claude . . . 14 - 21
Mead, Claude . . . A - 1021
Mead, Franklin . . . A - 603
Meyer, Michel . . . 3 - 109
Meyer, Stanley . . . 3 - 191
Meyer, Stanley . . . 10 - 84
Meyer, Stanley . . . A - 620
Meyer, Stanley . . . A - 627
Meyer, Stanley . . . A - 637
Meyer, Stanley . . . A - 642
Meyer, Stanley . . . A - 659
Meyer, Stanley . . . A - 670
Meyer, Stanley . . . A - 680
Meyers, Roy . . . A - 524
Milkovic, Veljko . . . 4 - 25
Molina-Martinez, Alberto . . . 5 - 86
Moller, Nikolas . . . 13 - 11
Moore, Kenneth . . . 13 - 49
Moore, Dr. Terry . . . 11 - 30
Moray, Thomas Henry . . . 7 - 14
Mourier, Georges . . . 2 - 29
Nachamkin, Jack . . . A - 603
Naudin, Jean-Louis . . . 13 - 8
Neal, Bob . . . 8 - 1
Neal, Bob . . . A - 1126
Nelson, Carl . . . A - 470
Newberry, Ivor . . . A - 940
Newman, Joseph . . . 11 - 40
Newman, Joseph . . . A - 280
Nobel, Joe . . . 9 - 35
Ogle, Thomas . . . A - 955
Ognyanov, Michael . . . A - 113
Paine, Henry . . . 14 - 7
Paine, Henry . . . A - 901
Papp, Josef . . . 8 - 25
Papp, Josef . . . A - 1153
Patrick, Stephen . . . 13 - 49
Patterson, Robert . . . 10 - 111
Perkins, Eugene . . . 14 - 12
Phillips, Raymond Snr . . . 7 - 34
Plauston, Hermann . . . 7 - 31
Plauston, Hermann . . . A - 485
Pogue, Charles . . . A - 919
Pogue, Charles . . . A - 927
Pogue, Charles . . . A - 932
Porter, Art . . . 2 - 56
Prentice, Frank . . . 5 - 1
Pugh, Ron . . . 6 - 6
Puharich, Henry . . . A - 689
Quirey, David . . . A - 689
Rayburn, Lawrence . . . 7 - 39
Reardon, John . . . A - 196
Ribero, Renato . . . 4 - 48
Robertson, Scott . . . 8 - 4
Rogers, Leroy . . . 8 - 6
Rogers, Leroy . . . A - 1132
Rothman Technologies . . . 13 - 44
Santilli, Ruggero . . . A - 754
‘Salty Citrus’ . . . 3 - 148
Schappeller, Karl . . . 11 - 60
Schwartz, Harold . . . A - 949
Scott, John . . . 8 - 48

Searle, John R. R. . . . . 11 - 32
Seddon-Gillespie, Ronald . . . . . 3 - 110
Seddon-Gillespie, Ronald . . . . . A - 83
Shelton, Robert. . . . . A - 945
Silverhealtheu . . . . . 3 - 77
Simpson, Dale . . . . . 4 - 22
Skinner, William . . . . . 4 - 53
Smith, Donald Lee . . . . . 3 - 112
Smith, Donald Lee . . . . . 5 - 52
Soukup, George . . . . . 1 - 26
Spence, Geoffrey . . . . . A - 213
Spiros, Spiro . . . . . A - 864
Steorn . . . . . 1 - 24
Stevens, Peter . . . . . 9 - 14
Sweet, Floyd . . . . . 3 - 102
Sweet, Floyd . . . . . 5 - 102
Sweet, Floyd . . . . . A - 1196
Szilard, Leo . . . . . 14 - 61
Teal, Ben . . . . . 4 - 5
Tesla, Nikola . . . . . 5 - 6
Tesla, Nikola . . . . . 5 - 76
Tesla, Nikola . . . . . 7 - 1
Tesla, Nikola . . . . . 11 - 1
Tesla, Nikola . . . . . 11 - 17
Teslonian, Mr . . . . . 14 - 4
Tewari, Paramahamsa . . . . . 13 - 7
TheGuru2You. . . . . 3 - 99
The Mad Scientist . . . . . 5 - 70
Torres, Hector . . . . . 2 - 39
Tracy, Robert . . . . . 1 - 9
Trawoeger, Thomas . . . . . 9 - 3
Tseung, Lawrence . . . . . 2 - 50
Tseung, Lawrence . . . . . 3 - 1
Tseung, Lawrence . . . . . 4 - 1
Tseung, Lawrence . . . . . 5 - 56
Tucker, Oliver. . . . . A - 952
UFOpolitics . . . . . 2 - 57
UFOpolitics . . . . . 3 - 182
UFOpolitics . . . . . 6 - 46
Utkin, Vladimir . . . . . 5 - 53
Valera, Cherepanov . . . . . 5 - 51
Van Valkenburg, Eber. . . . . 8 - 23
Van Valkenburg, Eber. . . . . A - 1148
Volfson, Boris . . . . . A - 904
Wang, ShenHe . . . . . 1 - 1
Watson, Jim . . . . . 4 - 12
Weir, Richard . . . . . A - 470
West, Zach . . . . . 10 - 18
Willis, Richard . . . . . 3 - 75
Willis, Richard . . . . . A - 1028
Wilson, Mr . . . . . 4 - 11
Yildiz, Muammer . . . . . 13 - 13
Yildiz, Muammer . . . . . 1 - 35
Zbikowski, Jerzy . . . . . 4 - 17
Zigouras, Paul . . . . . 10 - 129

A Practical Guide to 'Free-Energy' Devices

Overview

This eBook contains most of what I have learned about this subject after researching it for a number of years. I am not trying to sell you anything, nor am I trying to convince you of anything. When I started looking into this subject, there was very little useful information and any that was around was buried deep in incomprehensible patents and documents. My purpose here is to make it easier for you to locate and understand some of the relevant material now available. What you believe is up to yourself and is none of my business. Let me stress that almost all of the devices discussed in the following pages, are devices which I have not personally built and tested. It would take several lifetimes to do that and it would not be in any way a practical option. Consequently, although I believe everything said is fully accurate and correct, you should treat everything as being "hearsay" or opinion.

Some time ago, it was commonly believed that the world was flat and rested on the backs of four elephants and that when earthquakes shook the ground, it was the elephants getting restless. If you want to believe that, you are fully at liberty to do so, however, you can count me out as I don't believe that.

The Wright brothers were told that it was impossible for aeroplanes to fly because they were heavier than air. That was a commonly believed view. The Wright brothers watched birds flying and since, without question, birds are considerably heavier than air, it was clear that the commonly held view was plain wrong. Working from that realisation, they developed aeroplanes which flew perfectly well.

The years passed, and the technology started by the Wright brothers and their careful scientific measurements and well-reasoned theory, advanced to become the "science" of aeronautics. This science was used extensively to design and build very successful aircraft and "aeronautics" gained the aura of being a "law".

Unfortunately, somebody applied aeronautic calculations to the flight of bumblebees and discovered that according to aeronautics, bumblebees couldn't possibly fly as their wings could not generate enough lift to get them off the ground. This was a problem, as it was perfectly possible to watch bees flying in a very competent manner. So, the "laws" of aeronautics said that bees can't fly, but bees actually **do** fly.

Does that mean that the laws of aeronautics were no use? Certainly not - those "laws" had been used for years and proved their worth by producing excellent aircraft. What it did show was that the "laws" of aeronautics did not yet cover every case and needed to be extended to cover the way that bees fly, which is through lift generated by turbulent airflow.

It is **very** important to realise that what are described as scientific "laws" are just the best working theories at the present time and it is virtually certain that those "laws" will have to be upgraded and extended as further scientific observations are made and further facts discovered. Let's hope those four elephants don't get restless before we have a chance to learn a bit more!

Introduction

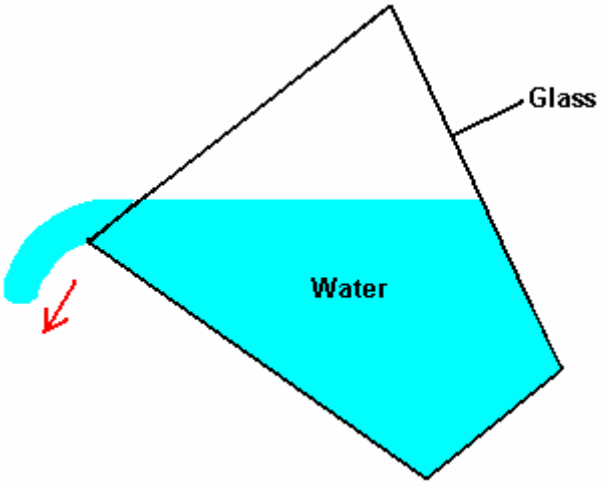
It should be stressed at this point, that this material is intended to provide you with information and only that. If you should decide, on the basis of what you read here, to build some device or other, you do so solely and entirely at your own risk and on your own responsibility. For example, if you build something in a heavy box and then drop it on your toe, then that is completely your own responsibility (you should learn to be more careful) and nobody other than yourself is in any way liable for your injury, or any loss of income caused while your toe is recovering. Let me amplify that by stating that I do not warrant that any device or system described in this document works as described, or in any other way, nor do I claim that any of the following information is useful in any way or that any device described is useful in any way or for any purpose whatsoever. Also, let me stress that I am not encouraging you to actually construct any device described here, and the fact that very detailed construction details are provided, must not be interpreted as my encouraging you to physically construct any device described in this document. You are welcome to consider this a work of fiction if you choose to do so.

I apologise if this presentation seems very elementary, but the intention is to make each description as simple as possible so that everybody can understand it, including people whose native language is not English. If you are not familiar with the basic principles of electronics, then please read the simple step-by-step electronics tutorial in Chapter 12 which is intended to help complete beginners in the subject.

At this point in time - the early years of the twenty-first century - we have reached the point where we need to realise that some of the “laws” of science do not cover every case, and while they have been very useful in the past, they do need to be extended to cover some cases which have been left out until now.

For example, suppose a bank robber broke into a bank and stole all of the cash there. How much could he take? Answer: “every coin and every note”. The limit is the sum total of all cash in the building. This is what the “Law” of Conservation of Energy is all about. What it says is very simple – you can’t take out any more than there is there in the beginning. That seems pretty straightforward, doesn’t it?

As another example, consider a glass tumbler filled completely with water. Using common sense, tell me, how much water can be poured out of the glass? For the purposes of this illustration, please take it that temperature, pressure, gravity, etc. all remain constant for the duration of the experiment.

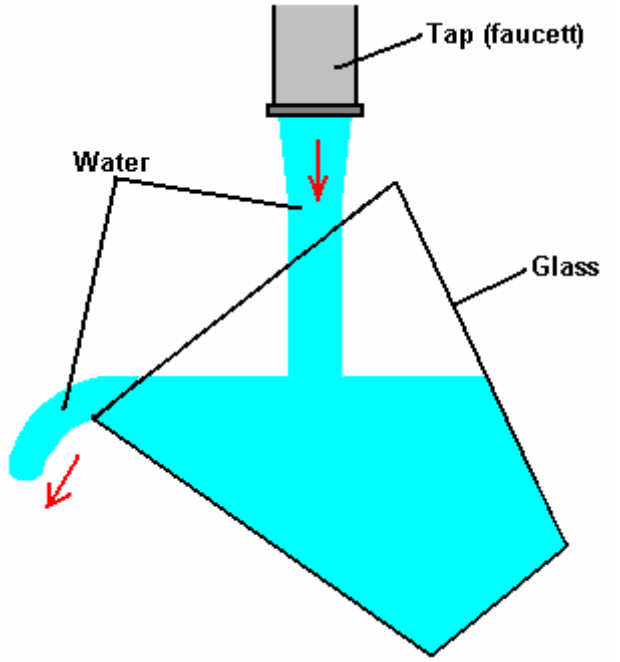


The answer is: “the exact volume contained inside the tumbler”. Agreed. This is what present day science says. To be strictly accurate, you will never be able to pour all of the water out as a small amount will remain, wetting the inside of the glass. Another way of putting this is to say that the “efficiency” of the pouring operation is not 100%. This is typical of life in general, where very few, if any, actions are 100% efficient.

So, are we agreed with current scientific thinking then – the maximum amount of water which can pour out of the tumbler is the total volume inside the tumbler? This seems simple and straightforward, doesn’t it? Science thinks so, and insists that this is the end of the story, and nothing else is possible. This arrangement is called a “closed system” as the only things being considered are the glass, the water and gravity.

Well, unfortunately for current scientific thinking, this is not the only possible situation and “closed systems” are almost unknown in the real world. Mostly, assumptions are made that the effects of anything else around will cancel out and add up to a net zero effect. This is a very convenient theory, but unfortunately it has no basis in reality.

Let’s fill our glass with water again and begin to pour it out again, but this time we position it underneath a source of flowing water:



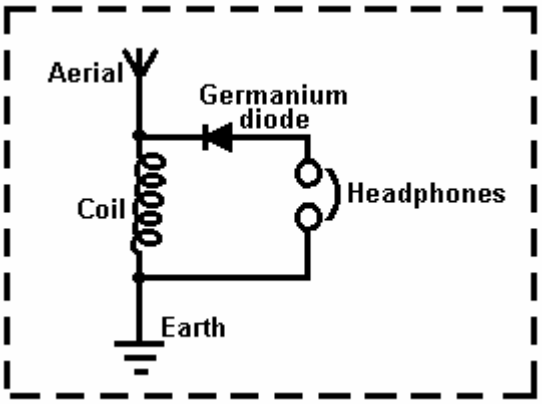
So, now, how much water can be poured out of the tumbler? Answer: “millions of times the volume of the tumbler”. But hang on a moment, haven’t we just said that the absolute limit of water poured from the tumbler has to be the volume inside the tumbler? Yes, that’s exactly what we said, and that is what current science teaching says. The bottom line here is that what current science says does in fact hold true for most of the time, but there are cases where the basic assumption of it being a “closed system” is just not true.

One popular misconception is that you can’t get more energy out of a system than you put into it. That is wrong, because the sentence was worded carefully. Let me say it again and this time, emphasise the key words: “**you** can’t get more energy out of a system than **you** put into it”. If that were true, then it would be impossible to sail a yacht all the way around the world without burning any fuel, and that has been done many times and none of the driving energy came from the crews. If it were true, then a grain mill driven by a waterwheel would not be able to produce flour as the miller certainly does not push the millstones around himself. If that were true, then nobody would build windmills, or construct solar panels, or tidal power stations.

What the statement should say is “more energy can’t be taken out of a system than is put into it or is already in it” and that is a very different statement. When sailing a yacht, the wind provides the driving force which makes the trip possible. Notice that, it is the environment providing the power and not the sailors. The wind arrived without them having to do anything about it, and a lot less than 100% of the wind energy reaching the yacht actually becomes forward thrust, contributing to the voyage. A good deal of the energy arriving at the yacht ends up stretching the rigging, creating a wake, producing noise, pushing the helmsman, etc. etc. This idea of no more energy coming out of a system than goes into it, is called “The Law of Conservation of Energy” and it is perfectly right, in spite of the fact that it gets people confused.

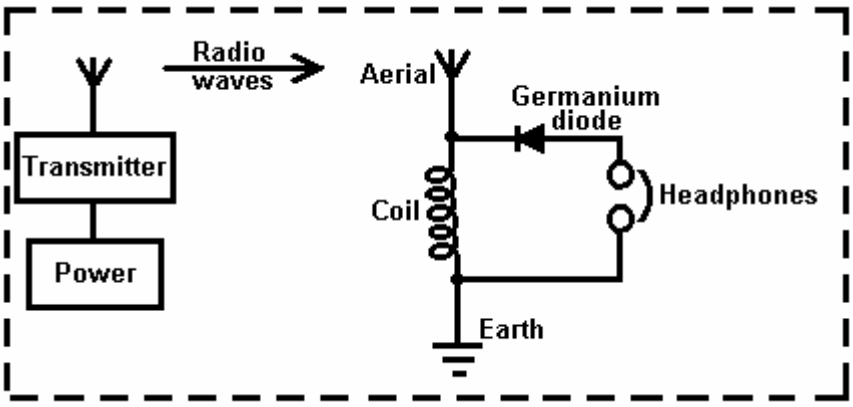
“Free-Energy Devices” or “Zero-Point Energy Devices” are the names applied to systems which appear to produce a higher output power than their input power. There is a strong tendency for people to state that such a system is not possible since it contravenes the Law of Conservation of Energy. It doesn’t. If it did, and any such system was shown to work, then the “Law” would have to be modified to include the newly observed fact. No such change is necessary, it merely depends on your point of view.

For example, consider a crystal set radio receiver:



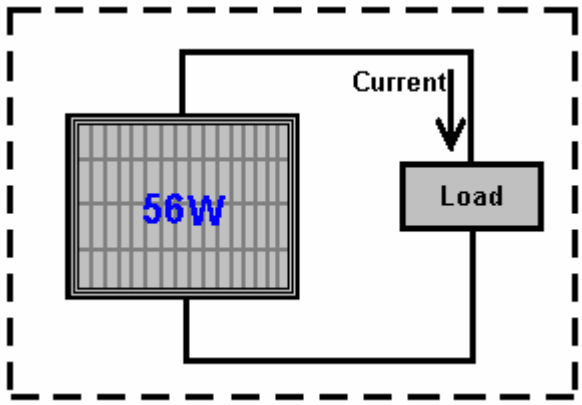
Looking at this in isolation, we appear to have a free-energy system which contradicts the Law of Conservation of Energy. It doesn't, of course, but if you do not view the whole picture, you see a device which has only passive components and yet which (when the coil is of the correct size) causes the headphones to generate vibrations which reproduce recognisable speech and music. This looks like a system which has no energy input and yet which produces an energy output. Considered in isolation, this would be a serious problem for the Law of Conservation of Energy, but when examined from a common sense point of view, it is no problem at all.

The whole picture is:

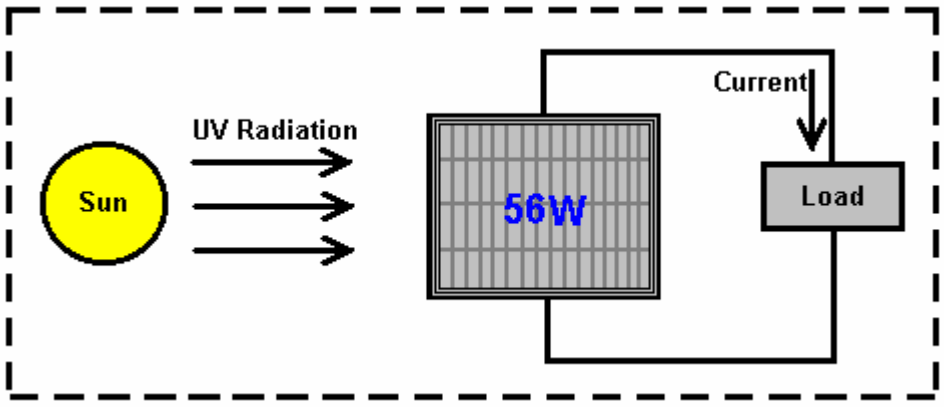


Power is supplied to a nearby transmitter which generates radio waves which in turn, induce a small voltage in the aerial of the crystal set, which in turn, powers the headphones. The power in the headphones is far, far less than the power taken to drive the transmitter. There is most definitely, no conflict with the Law of Conservation of Energy. However, there is a quantity called the "**C**oefficient **O**f **P**erformance" or "**COP**" for short. This is defined as the amount of power coming out of a system, divided by the amount of power that **the operator** has to put into that system to make it work. In the example above, while the **efficiency** of the crystal set radio is well below 100%, the **COP** is greater than 1. This is because the owner of the crystal radio set does not have to supply any power at all to make it work, and yet it outputs power in the form of sound. As the input power from the user, needed to make it work is zero, and the COP value is calculated by dividing the output power by this zero input power, the COP is actually infinity. **Efficiency** and **COP** are two different things. Efficiency can never exceed 100% and almost never gets anywhere near 100% due to the losses suffered by any practical system.

As another example, consider an electrical solar panel:



Again, viewed in isolation, this looks like (and actually **is**) a Free-Energy device if it is set up out of doors in daylight, as current is supplied to the load (radio, battery, fan, pump, or whatever) without the user providing any input power. Again, Power Out with no Power In. Try it in darkness and you find a different result because the whole picture is:



The energy which powers the solar panel comes from the sun.. Only some 17% of the energy reaching the solar panel is converted to electrical current. This is most definitely not a contravention of the Law of Conservation of Energy. This needs to be explained in greater detail. The Law of Conservation of Energy applies to closed systems, and only to closed systems. If there is energy coming in from the environment, then the Law of Conservation of Energy just does not apply, unless you take into account the energy entering the system from outside.

People sometimes speak of “over-unity” when talking about the efficiency of a system. From the point of efficiency, there is no such thing as “over-unity” as that would mean that more power was coming out of the system than the amount of power entering the system. Our trusty bank robber mentioned above would have to take out of the bank vault, more money than was actually in it, and that is a physical impossibility. There are always some losses in all practical systems, so the efficiency is always less than 100% of the power entering the system. In other words, the efficiency of any practical system is always under unity.

However, it is perfectly possible to have a system which has a greater power output than the power input which **we** have to put into it to make it work. Take the solar panel mentioned above. It has a terribly low efficiency of about 17%, **but**, we don’t have to supply it with any power to make it work. Consequently, when it is in sunlight, it’s Coefficient Of Performance (“COP”) is it’s output power (say, 50 watts) divided by the input power needed to make it work (zero watts) which is infinity. So, our humble, well-known solar panel has terrible efficiency of 17% but at the same time it has a COP of infinity.

It is now generally accepted that “Dark Matter” and “Dark Energy” form more than 80% of our universe. There is nothing sinister about the adjective “Dark” as in this context, it merely means that we cannot see it. There are many useful things which we utilise, which we can’t see, for example, radio waves, TV signals, magnetism, gravity, x-rays, etc. etc.

The actual situation is, that we are sitting in a vast field of energy which we can't see. This is the equivalent of the situation for the crystal set shown above, except that the energy field we are in is very, very much more powerful than the radio waves from a radio transmitter. The problem is, how to tap the energy which is freely available all around us, and get it to do useful work for us. It can definitely be done, but it is not easy to do.

Some people think that we will never be able to access this energy. Not very long ago, it was widely believed that nobody could ride a bicycle faster than 15 miles per hour because the wind pressure on the face of the rider would suffocate him. Today, many people cycle much faster than this without suffocating - why? - because the original negative opinion was wrong.

Not very long ago, it was thought that metal aircraft would never be able to fly because metal is so much heavier than air. Today, aircraft weighing hundreds of tons fly on a daily basis. Why? - because the original negative opinion was not correct.

It is probably worth while, at this point, to explain the basics of Zero-Point Energy. The experts in Quantum Mechanics refer to how the universe operates as "Quantum Foam". Every cubic centimetre of "empty" space is seething with energy, so much in fact, that if it were converted using Oliver Heaviside's equation (made famous by Albert Einstein) $E = mC^2$ (that is Energy = Mass multiplied by a very big number), then it would produce as much matter as can be seen by the most powerful telescope. There is actually nothing "empty" about space. So why can't we see anything there? Well, you can't actually see energy. All right then, why can't you measure the energy there? Well, two reasons actually, firstly, we have never managed to design an instrument which can measure this energy, and secondly, the energy is changing direction incredibly rapidly, billions and billions and billions of times each second.

There is so much energy there, that particles of matter just pop into existence and then pop back out again. Half of these particles have a positive charge and half of them have a negative charge, and as they are evenly spread out in three-dimensional space, the overall average voltage is zero. So, if the voltage is zero, what use is that as a source of energy? The answer to that is "none" if you leave it in it's natural state. However, it is possible to change the random nature of this energy and convert it into a source of unlimited, everlasting power which can be used for all of the things we use mains electricity for today - powering motors, lights, heaters, fans, pumps, ... you name it, the power is there for the taking.

So, how do you alter the natural state of the energy in our environment? Actually, quite easily. All that is needed is a positive charge and a negative charge, reasonably near each other. A battery will do the trick, as will a generator, as will an aerial and earth, as will an electrostatic device like a Wimshurst machine. When you generate a **Plus** and a **Minus**, the quantum foam is affected. Now, instead of entirely random plus and minus charged particles appearing everywhere, the **Plus** which you created gets surrounded by a sphere of minus charge particles popping into existence all around it. Also, the **Minus** which you created, gets surrounded by a spherical-shaped cloud of plus-charge particles popping into existence all around it. The technical term for this situation is "**broken symmetry**" which is just a fancy way of saying that the charge distribution of the quantum foam is no longer evenly distributed or "symmetrical". In passing, the fancy technical name for your **Plus** and **Minus** near each other, is a "**dipole**" which is just a techno-babble way of saying "two poles: a plus and a minus" - isn't jargon wonderful?

So, just to get it straight in your mind, when you make a battery, the chemical action inside the battery creates a Plus terminal and a Minus terminal. Those poles actually distort the universe around your battery, and causes vast streams of energy to radiate out in every direction from each pole of the battery. Why doesn't the battery run down? Because the energy is flowing from the environment and not from the battery. If you were taught basic physics or electrical theory, you will probably have been told that the battery used to power any circuit, supplies a stream of electrons which flows around the circuit. Sorry Chief - it just ain't like that at all. What really happens is that the battery forms a "dipole" which nudges the local environment into an unbalanced state which pours out energy in every direction, and some of that energy from the environment flows around the circuit attached to the battery. The energy does **not** come from the battery.

Well then, why does the battery run down, if no energy is being drawn from it to power the circuit? Ah, that is the really silly thing that we do. We create a closed-loop circuit (because that's what we have always done) where the current flows around the circuit, reaches the other battery terminal and immediately destroys the battery's "dipole". Everything stops dead in it's tracks. The environment becomes symmetrical again, the massive amount of readily available free-energy just disappears and you are back to where you started from. **But**, do not despair, our trusty battery immediately creates the Plus and Minus terminals again and the process starts all over again. This happens so rapidly that we don't see the breaks in the operation of the circuit and it is the continual recreation of the dipole which causes the battery to run down and lose it's power. Let me say it again, the battery

does **not** supply the current that powers the circuit, it never has and it never will - the current flows into the circuit from the surrounding environment.

What we really need, is a method of pulling off the power flowing in from the environment, without continually destroying the dipole which pushes the environment into supplying the power. That is the tricky bit, but it has been done. If you can do that, then you tap into an unlimited stream of inexhaustible energy, with no need to provide any input energy to keep the flow of energy going. In passing, if you want to check out the details of all of this, Lee and Yang were awarded the Nobel Prize for Physics in 1957 for this theory which was proved by experiment in that same year. This eBook includes circuits and devices which manage to tap this energy successfully.

Today, many people have managed to tap this energy but very few commercial devices are readily available for home use. The reason for this is human rather than technical. More than 10,000 Americans have produced devices or ideas for devices but none have reached commercial production due to opposition from influential people who do not want such devices freely available. One technique is to classify a device as "essential to US National Security". If that is done, then the developer is prevented from speaking to anyone about the device, even if he has a patent. He cannot produce or sell the device even though he invented it. Consequently, you will find many patents for perfectly workable devices if you were to put in the time and effort to locate them, though most of these patents never see the light of day, having been taken for their own use, by the people issuing these bogus "National Security" classifications.

If you feel that this opposition to free-energy and related technology is a figment of my imagination and that the people who state that more than 40,000 free-energy device patents have already been suppressed, then please consider this extract from a 2006 reminder to Patent Office staff in America to single out all patents which have to do with free-energy and any related subjects and take those patent applications to their supervisor to be dealt with differently to all other patent applications:

B. Subject matter of special interest in TC 2800

1. Perpetual motion machines; classes 310 and 290
2. Anti-gravity devices
3. Room temperature superconductivity; class 310
4. Free energy – Tachyons, etc.
5. Gain-Assisted Superluminal Light Propagation (faster than the speed of light); class 702, 359
6. Other matters that violate the general laws of physics; classes 73, 290.
7. Applications containing claims to subject matter which, if issued, would generate unfavorable publicity for the USPTO, class 84, 702.
8. Reexamination proceedings involving patents in litigation and:
   The court decision/verdict is subject to review by the Supreme Court
   The court decision includes high monetary awards
   The technology and companies involved would likely generate high publicity

Here "USPTO" is the United States Patent and Trademark Office, which is a privately owned commercial company run to make money for it's owners.

The purpose of this eBook is to present the facts about some of these devices and more importantly, where possible, explain the background details of why and how systems of that type function. As has been said before, it is not the aim of this book to convince you of anything, just to present you with some of the facts which are not that easy to find, so that you can make up your own mind on the subject.

The science taught in schools, colleges and universities at this time, is well out of date and in serious need of being brought up to date. This has not happened for some time now as people who make massive financial profits have made it their business to prevent any significant advance for many years now. However, the internet and free sharing of information through it, is making things very difficult for them. What is it that they don't want you to know? Well, how about the fact that you don't have to burn a fuel to get power? Shocking, isn't it !! Does it sound a bit mad to you? Well, stick around and start doing some thinking.

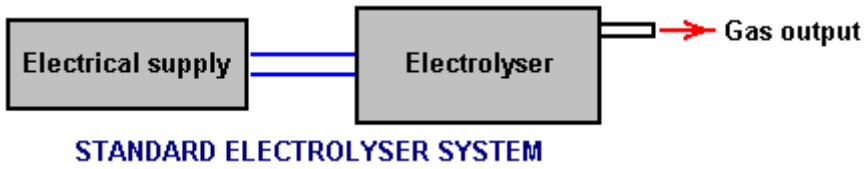
Suppose you were to cover a boat with lots of solar panels which were used to charge a large bank of batteries inside the boat. And if those batteries were used to operate electric motors turning propellers which drive the boat along. If it is sunny weather, how far could you go? As far as the boat can travel while the sun is up and if the battery bank is large, probably most of the night as well. At sun-up on the next day, you can continue your journey. Oceans have been crossed doing this. How much fuel is burned to power the boat? None !! Absolutely none at all. And yet, it is a fixed idea that you have to burn a fuel to get power.

Yes, certainly, you can get power from the chemical reaction of burning a fuel - after all, we pour fuel into the tanks of vehicles "to make them go" and we burn oil in the central heating systems of buildings. But the big question is: "Do we have to?" and the answer is "No". So why do we do it? Because there is no alternative at present. Why is there no alternative at present? Because the people making incredibly large financial profits from selling this fuel, have seen to it that no alternative is available. We have been the suckers in this con trick for decades now, and it is time for us to snap out of it. Let's have a look at some of the basic facts:

Let me start by presenting some of the facts about electrolysis. The electrolysis of water is performed by passing an electric current through the water, causing it to break up into hydrogen gas and oxygen gas. This process was examined in minute detail by Michael Faraday who determined the most energy efficient possible conditions for electrolysis of water. Faraday determined the amount of electric current needed to break the water apart, and his findings are accepted as a scientific standard for the process.

We now bump into a problem which scientists are desperate to ignore or deny, as they have the mistaken idea that it contradicts the Law of Conservation of Energy – which, of course, it doesn't. The problem is an electrolyser design by Bob Boyce of America which appears to have an efficiency twelve times greater than Faraday's maximum possible gas production. This is a terrible heresy in the scientific arena and it gets the average "by the book" scientist very up-tight and flustered. There is no need for this worry. The Law of Conservation of Energy remains intact and Faraday's results are not challenged. However, an explanation is called for.

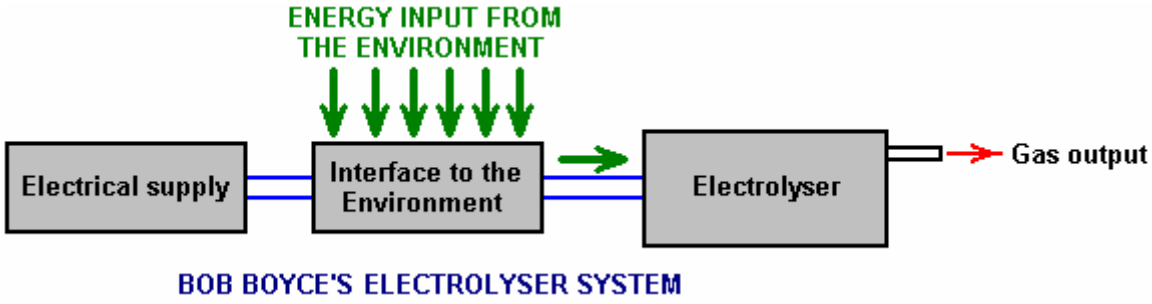
To start with, let me show the arrangement for a standard electrolyser system:



STANDARD ELECTROLYSER SYSTEM

Here, current is supplied to the electrolyser by the electrical supply. The current flow causes breakdown of the water contained in the electrolyser, resulting in the amount of gas predicted by Faraday (or less if the electrolyser is not well designed and accurately built).

Bob Boyce, who is an exceptionally intelligent, perceptive and able man, has developed a system which performs the electrolysis of water using power drawn from the environment. To a quick glance, Bob's design looks pretty much like a high-grade electrolyser (which it is) but it is a good deal more than that. The practical construction and operational details of Bob's design are shown in http://www.free-energy-info.tuks.nl/D9.pdf, but for here, let us just consider the operation of his system in very broad outline:



BOB BOYCE'S ELECTROLYSER SYSTEM

The very important distinction here is that the power flowing into the electrolyser and causing the water to break down and produce the gas output, is coming almost exclusively from the environment and **not** from the electrical supply. The main function of Bob's electrical supply is to power the device which draws energy in from the environment. Consequently, if you assume that the current supplied by the electrical supply is the whole of the power driving the electrolyser, then you have a real problem, because, when properly built and finely tuned, Bob's electrolyser produces up to 1,200% of Faraday's maximum efficiency production rate.

This is an illusion. Yes, the electrical input is exactly as measured. Yes, the gas output is exactly as measured. Yes, the gas output is twelve times the Faraday maximum. But Faraday's work and the Law of Conservation of Energy are not challenged in any way **because** the electrical current measured is used primarily to power the interface to the environment and nearly all of the energy used in the electrolysis process flows in from the local environment and is not measured. What we can reasonably deduce is that the energy inflow from the environment is probably about twelve times the amount of power drawn from the electrical supply.

At this point in time, we do not have any equipment which can measure this environmental energy. We are in the same position as people were with electrical current five hundred years ago – there was just no equipment around which could be used to make the measurement. That, of course, does not means that electrical current did not exist at that time, just that we had not developed any equipment capable of performing measurement of that current. Today, we know that this environmental energy exists because we can see the effects it causes such as running Bob's electrolyser, charging batteries, etc. but we can't measure it directly because it vibrates at right-angles to the direction that electrical current vibrates in. Electrical current is said to vibrate "transversely" while this zero-point energy vibrates "longitudinally", and so has no effect on instruments which respond transversely such as ammeters, voltmeters, etc.

Bob Boyce's 101-plate electrolyser produces anything up to 100 litres of gas per minute, and that rate of production is able to power internal combustion engines of low capacity. The vehicle alternator is perfectly capable of powering Bob's system, so the result is a vehicle which appears to run with water as the only fuel. This is not the case, nor is it correct to say that the engine is powered by the gas produced. Yes, it does utilise that gas when running, but the power running the vehicle is coming directly from the environment as an inexhaustible supply. In the same way, a steam engine does not run on water. Yes, it does utilise water in the process, but the power that runs a steam engine comes from burning the coal and not from the water.

The Basics of "Free-Energy":

This beginner's introduction presumes that you have never heard of free-energy before and would like an outline sketch of what it is all about, so let's begin at the beginning.

We tend to have the impression that people who lived a long time ago were not as clever as we are - after all, we have television, computers, mobile phones, games consoles, aeroplanes, …. But, and it is a big "but", the reason why they did not have those things is because science had not advanced far enough for those things to become possible. That did **not** mean that the people who lived before us were any less clever than we are.

For example, could you, personally, come up with an accurate calculation of the circumference of the Earth? This has to be without prior knowledge, no satellites, no astronomical information, no calculators, no computers and no experts to guide you. Eratosthenes did by observing the shadows in two wells some 800 kilometres apart. When was that? More than two thousand years ago.

You have probably heard of the geometry of Pythagoras who lived hundreds of years before Eratosthenes, and that geometry is still used in remote areas to lay out the foundations for new buildings. You have probably heard of Archimedes who worked out why things float. He lived more than two thousand years ago. So, how do those people stack up against you and me? Were they stupid people?

This is quite an important point because it demonstrates that the body of scientific information enables many things which were not thought possible in earlier times. This effect is not restricted to centuries ago. Take the year 1900. My father was a youngster then, so it is not all that long ago. It would be another three years before Orville and Wilbur Wright made their first 'heavier-than-air' flight, so there no aircraft around in 1900. There were no radio stations and most definitely, no television stations, nor would you have found a telephone inside a house. The only serious forms of information were books and periodicals or teaching establishments which relied on the knowledge of the teachers. There were no cars and the fastest form of transport for the average person was on a galloping horse.

Today, it is difficult to grasp what things were like not all that long ago, but come closer in time and look back just fifty years. Then, people researching in scientific fields had to design and build their own instruments before they ever got to experimenting in their chosen fields of knowledge. They were instrument makers, glass-blowers, metal workers, etc. as well as being scientific researchers. Nowadays there are measuring instruments of all

kinds for sale ready-made. We have silicon semiconductors which they didn't have, integrated circuits, computers, etc. etc.

The important point here is the fact that advances in scientific theory have made possible many things which would have been considered quite ridiculous notions in my father's time. However, we need to stop thinking as if we already know everything there is to know and that nothing which we think of as "impossible !!" could ever happen. Let me try to illustrate this by remarking on just a few things which as recently as the year 1900 would have marked you out as a "lunatic crank", things which we take for granted today because, and only because, we are now familiar with the science behind each of these things.

Certainties in the year 1900

A metal aeroplane weighing 350 tons couldn't possibly fly - everyone knows that !!

You couldn't possibly watch someone who is a thousand miles away - talk sense !!

No ! Of course you can't speak to somebody who lives in a different country unless you visit them !

The fastest way to travel is on a galloping horse.

A machine could never beat a man at chess - be realistic !

Today, we know that these things are not just possible, but we take them for granted. We have a mobile phone in our pocket and could easily use it to talk to friends in other countries almost anywhere in the world. It would seem very strange if we could not do that any more.

We each have a television and can watch, say, a golf tournament taking place at the other side of the world. We watch in real time, seeing the result of each stroke almost as soon as the golfer does himself. Even suggesting that such a thing was possible might have got you burnt at the stake for witchcraft, not all that long ago, but not having television would seem a very strange situation for us today.

If we see a 350 ton metal Boeing 747 aircraft flying past, we would not think it to be strange in any way, let alone think it was "impossible". It is routine, casual travel at 500 mph, a speed which would have been considered to be a fantasy when my father was young. The fact that the aircraft is so heavy, is of no concern to us as we know that it will fly, and does so, routinely, every day of the year.

We take for granted, a computer which can do a million things in one second. Today, we have lost the understanding of how big "a million" is, and we know that most people are likely to lose a game of chess if they play against a computer, even a cheap chess computer.

What we need to understand is that our present scientific knowledge is far from being comprehensive and there is still a very large amount to be learned, and that things which the average person today would consider "impossible" are quite liable to be casually routine day-to-day devices in just a few years time. This is **not** because we are stupid but instead it is because our current science still has a long way to go.

The objective of this website (http://www.free-energy-info.tuks.nl) is to explain some of the things which current science is not teaching at the present time. Ideally, we want a device which will power our homes and cars without the need to burn a fuel of any kind. Before you get the idea that this is some new and wild idea, please remember that windmills have been pumping water, milling grain, lifting heavy loads and generating electricity for a very long time now. Water wheels have been doing similar work for a very long time and both of these devices are fuel-less.

The energy which powers windmills and water wheels comes to us via our Sun which heats air and water, causing wind and rain, feeding power to our devices. The energy flows in from our local environment, costs us nothing and will keep on coming whether we make use of it or not.

Most of the pictures of wind generators and water wheels which you will see, show devices which would take a large amount of money to set up. The title of this eBook is "The Practical Guide to Free-Energy Devices" and the word "practical" is intended to indicate that most of the things spoken about are things which you, personally, have a reasonable chance of constructing for yourself if you decide to do so. However, while in chapter 14 there are instructions for building your own wind-powered electrical generator from scratch, pumping water uphill without using a fuel and utilising wave power at low cost, these things are subject to the weather. So, because of this, the main subject is the next generation of commercial devices, devices which do not need a fuel in order to function and power our homes and vehicles, devices which operate no matter what the weather is doing.

Perhaps I should remark at this point, that the commercial introduction of this new wave of hi-tech devices is being actively opposed by people who will lose a very large stream of revenue when it does eventually happen, as it most certainly will. For example, Shell BP which is a typical oil company, makes about US $3,000,000 profit per hour, every hour of every day of every year, and there are dozens of oil companies. The government makes even more than that out of the operation, with 85% of the sale price of oil in the UK being government tax. No matter what they say, (and they both do like to talk "green" in order to gain popularity), neither would ever for a single moment, consider allowing the introduction of fuel-less power devices, and they have the financial muscle to oppose this new technology at every possible level.

For example, some years ago Cal-Tech in the USA spent millions proving that on board fuel reformers for vehicles would give us all better fuel economy and cleaner air. They did long-term testing on buses and cars to provide proof. They teamed up with the very large auto-parts supplier Arvin Meritor to put these new devices in production vehicles. Then "One Equity Partners" bought out Arvin Meritor's division that did all the final work to get fuel reformers put into all new vehicles. They created a new company, EMCON Technologies, and that company dropped the fuel reformer from their product line, not because it did not work, but because it **did** work. This is not "conspiracy theory" but a matter of public record.

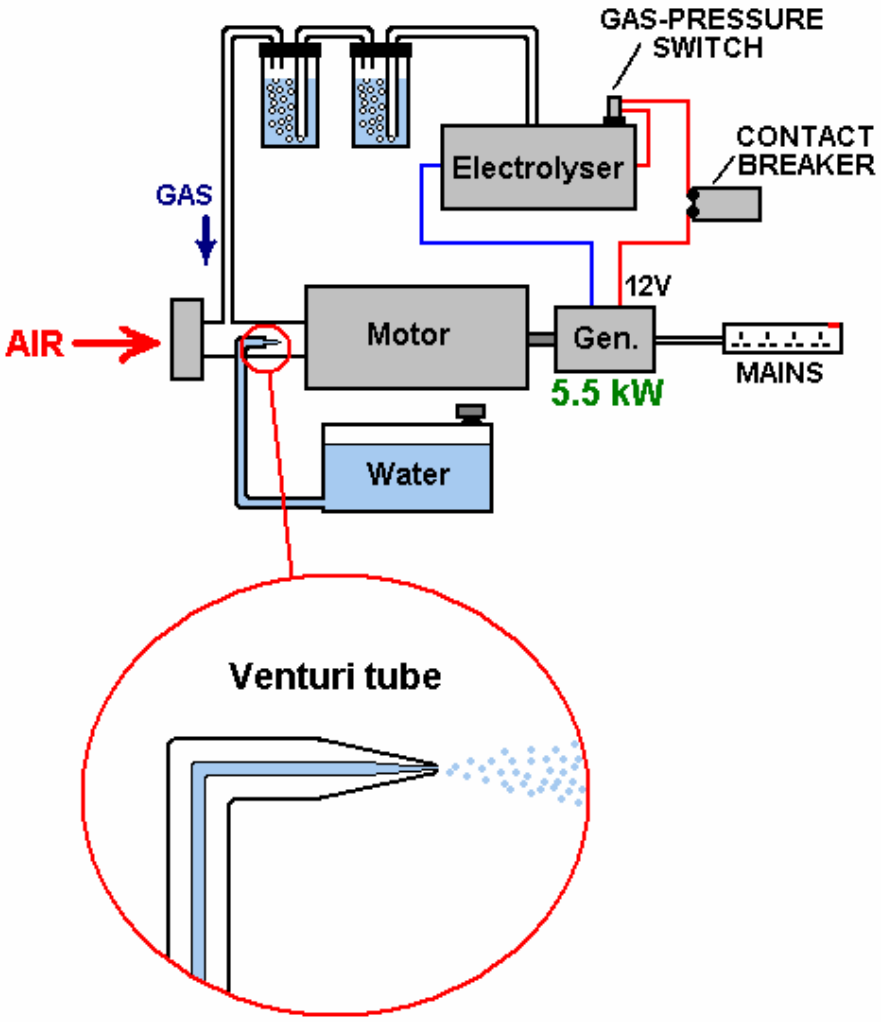
Some years ago, Stanley Meyer, a very talented man living in America, found a very energy-efficient way of breaking water into a mixture of hydrogen gas and oxygen gas. He pushed on further and found that a vehicle engine could be run on quite a small amount of this "HHO" gas if it was mixed with air, water droplets and some of the exhaust gas coming from the engine. He got funding to allow him to start manufacturing retro-fit kits which would allow any car to run on water alone and not use any fossil fuel at all. You can imagine how popular that

would have been with the oil companies and the government. Just after getting his funding, Stan was eating a meal at a restaurant when he jumped up, said "I've been poisoned !", rushed out into the car park and died on the spot. If Stan was mistaken, and he died of 'natural causes', then it was remarkably convenient timing for the oil companies and the government, and his retro-fit kits were never manufactured.

Even though Stan left behind many patents on the subject, until recently nobody managed to replicate his very low-power electrolyser, then Dave Lawton in Wales achieved the feat and many people have since replicated it by following Dave's instructions. More difficult still is getting an engine to run on no fossil fuel as Stan did, but recently, three men in the UK achieved just that by getting a standard petrol-engined electrical generator to run with water as the only fuel. Interestingly, this is not something which they want to pursue as they have other areas which appeal more to them. Consequently, they have no objections to sharing the practical information on what they did and the details are in chapter 10.

In very brief outline, they took a standard 5.5 kilowatt generator and delayed the spark timing, suppressed the 'waste' spark and fed the engine a mix of air, water droplets and just a small amount of HHO gas (which they measured at a flow rate of just three litres per minute). They test-loaded the generator with four kilowatts of electrical equipment to confirm that it worked well under load, and then moved on to a larger engine. This is the general style of generator which they used:

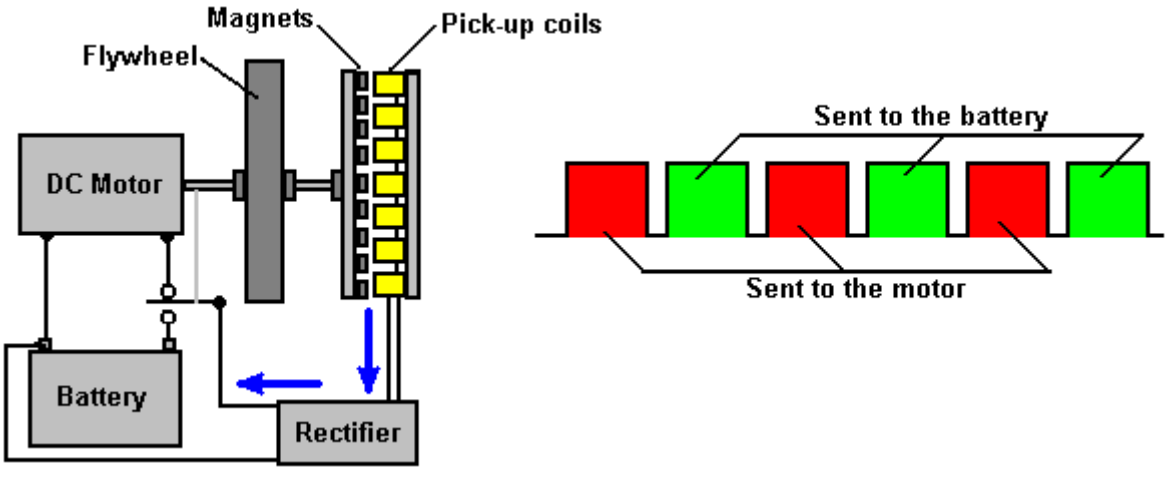
And their arrangement for running it without petrol, is shown in outline here, the full details being in chapter 10, including how to make your own high-performance electrolyser:



Conventional science says that it can prove mathematically that it is quite impossible to do this. However, the calculation is massively flawed in that is not based on what is actually happening and worse still, it makes initial assumptions which are just plain wrong. Even if we were not aware of these calculations, the fact that it has been done is quite enough to show that the current engineering theory is out of date and needs to be upgraded.

In passing, it might be remarked that an isolated, almost self-sufficient commune in Australia has been supplying their electrical needs by running ordinary electrical generators on water as the only (apparent) fuel for many years now.

However, let us now consider a device built by John Bedini, another talented man in America. He built a battery-powered motor with a flywheel on the shaft of the motor. This, of course, does not sound like startling stuff, but the crunch is that this motor ran in his workshop for more than three years, keeping it's battery fully charged during that time - now that **is** startling. The arrangement is like this:

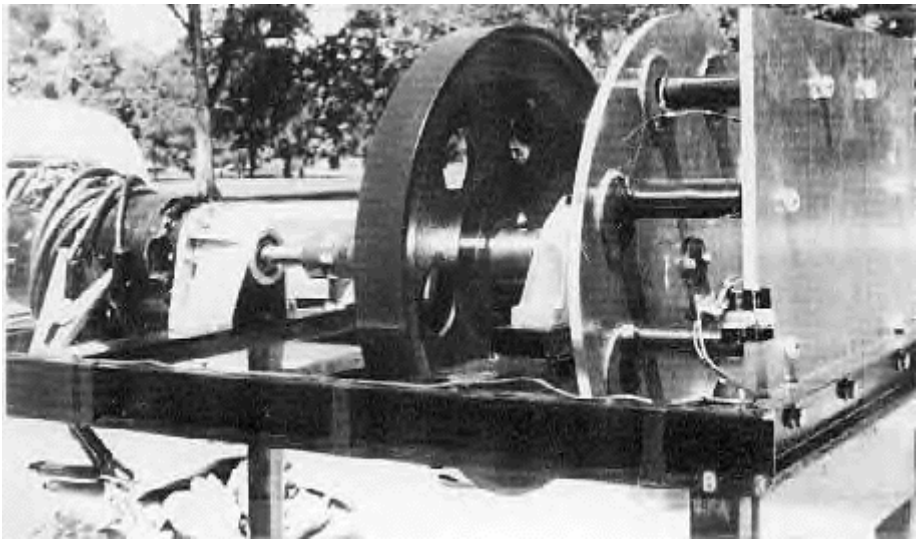


What makes this arrangement different from a standard set-up is that the battery powered motor is not connected directly to the battery but instead is fed with a rapid series of DC pulses. This has two effects. Firstly, that method of driving a motor is very efficient electrically speaking and secondly, when a flywheel is driven with a series of pulses, it picks up additional energy from the local environment.

One other unusual feature is the way that the motor shaft spins a disc with permanent magnets mounted on it. These sweep past a matching set of coils attached to a stationary board, forming an ordinary electrical generator and the resulting electrical power which is generated is converted to DC current and fed back to the driving battery, charging it and maintaining its voltage.

Standard theory says that a system like this has to be less than 100% efficient because the DC motor is less than 100% efficient (true) and the battery is only about 50% efficient (true). Therefore, the conclusion is that the system cannot possibly work (false). What is not understood by conventional science is that the pulsed flywheel draws in additional energy from the local environment, showing that conventional science theory is inadequate and out of date and needs to be upgraded, after all, this is not a ‘closed system’.

An American called Jim Watson built a much larger version of John's system, a version which was twenty feet (6 meters) long. Jim's version not only powered itself, but generated 12 kilowatts of excess electrical power. That extra 12 kilowatts of power must be a considerable embarrassment for conventional science and so they will either ignore it, or deny that it ever existed, in spite of the fact that it was demonstrated at a public seminar. This is what Jim's device looked like:



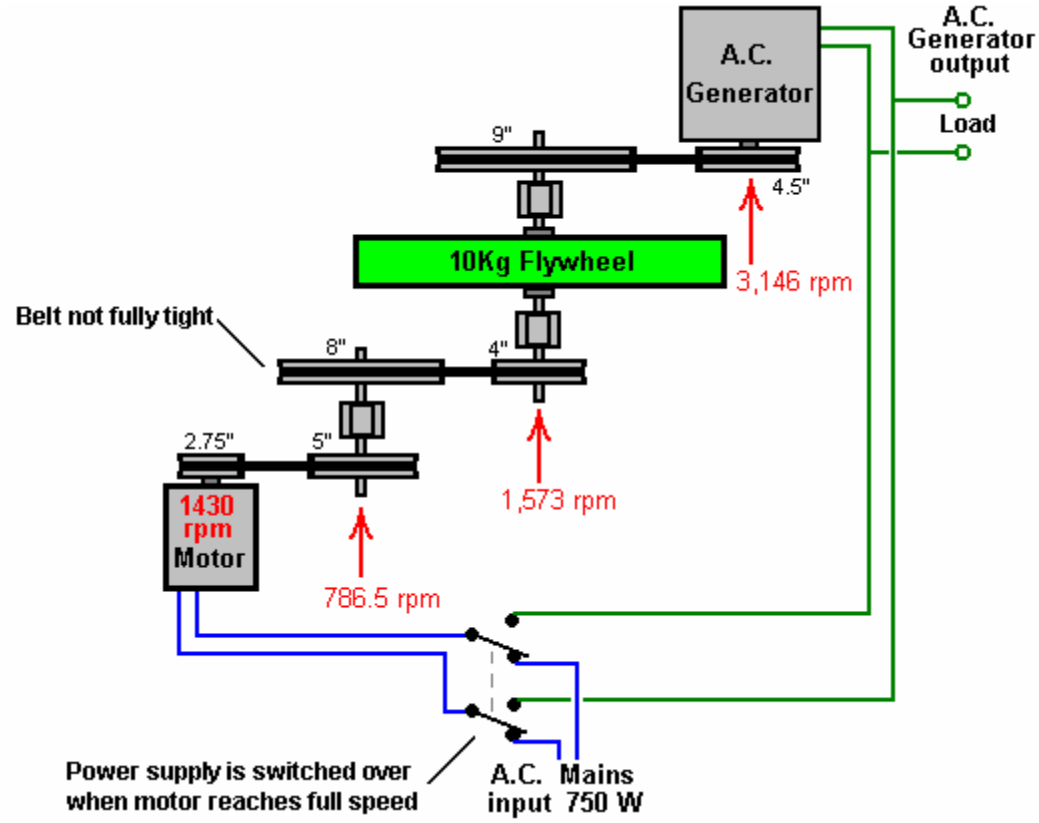
Working quite independently, an Australian called Chas Campbell, discovered the same effect. He found that if he used an AC motor plugged into the mains to drive a flywheel which in turn drove an ordinary generator, that it was possible to get a greater power output from the generator than the amount needed to drive the motor.

Chas used his motor to drive a series of shafts, one of which has a heavy flywheel mounted on it, like this:

The final shaft drives a standard electrical generator and Chas found that he could power electrical equipment from that generator, electrical mains equipment which required greater current that his mains-driven motor did.

Chas then took it one stage further and when the system was running at full speed, he switched his mains motor over from the wall socket to his own generator. The system continued to run, powering itself and driving other equipment as well.

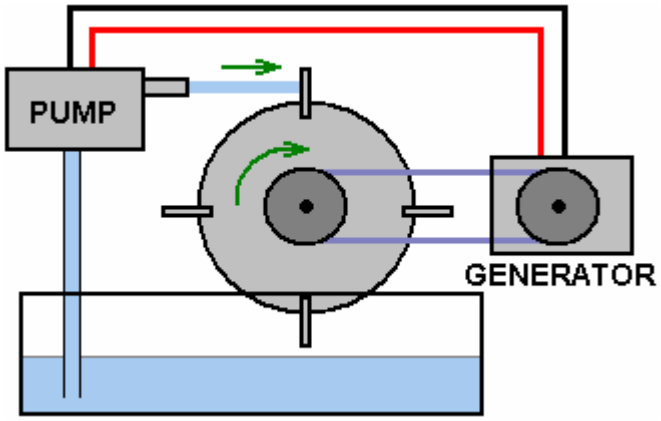
Conventional science says that this is impossible, which just goes to show that conventional science is out of date and needs to be upgraded to cover system like this where excess energy is flowing in from the local environment. Here is a diagram of how Chas Campbell's system is set up:



James Hardy has put a video on the web, showing a variation of this same principle. In his case, the flywheel is very light and has simple paddles attached around the rim of the wheel:

He then aims a powerful jet of water from a high-powered water pump, directly at the paddles, driving the wheel round with a rapid series of pulses. The shaft, on which the wheel is mounted, drives a standard electrical generator which lights an ordinary light bulb:

The really interesting part comes next, because he then unplugs the electrical supply to the water pump and switches it over to the generator which the wheel is driving. The result is that the pump powers itself and provides excess electricity which can be used to power other electrical equipment. The arrangement is like this:



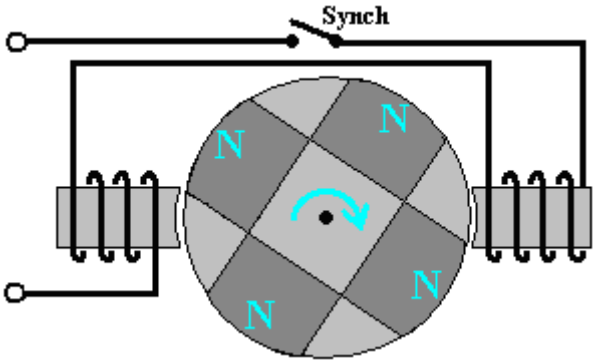
Once again, conventional science says that this is impossible, which in turn, demonstrates that conventional science is out of date and needs to be expanded to include these observed facts.

Permanent Magnets provide continuous power. This is because the poles of the magnet form a dipole, unbalancing the zero-point energy field near the magnet, and causing a continuous flow of energy which we call “lines of magnetic force”. This should be obvious as a magnet can support its own weight on the vertical face of a

refrigerator, for years on end. Conventional science says that permanent magnets can't be used as a source of power. However, the reality is that conventional science just doesn't know the techniques necessary for extracting that power. The lines of magnetic flow around any magnet are symmetrical and in order for a magnet to provide a useful directional force, it is necessary to arrange magnets in such a way that their combined magnetic field is no longer symmetrical. Doing that is not easy, but there are many alternative methods. Magnets are attracted to iron and that principle along with several other techniques have been successfully use by the New Zealander, **Robert Adams** who produced a motor which is, typically, at least 800% efficient. This, of course, is impossible according to conventional science. Robert was told that if he shared the information, he would be killed. He decided that at his age, being killed was not a major thing, so he went ahead and published all the details.

Present day motors driven by electricity are always less than 100% efficient as they are deliberately wound in a symmetrical way in order to make them inefficient. The Adams motor looks like a motor driven by electrical pulses, but it is **not**. The motor power comes from the permanent magnets mounted on the rotor and not from an electrical pulse applied to the electromagnets attached to the stator. The magnets are attracted to the metal cores of the stationary electromagnets. This provides the driving power of the motor. The electromagnets are then powered just enough to overcome the backwards drag of the magnets when they have just passed by the cores of the electromagnets.

The system works like this:



1. The magnets are attracted to the iron cores of the electromagnets, rotating the drive shaft and powering the motor.

2. The moving magnets generate electrical power in the windings of the electromagnets and this power is used to charge the driving battery.

3. When the permanent magnets reach the electromagnets, a small amount of electrical power is fed to the windings of the electromagnets in order to overcome any backward pull hindering the rotation of the drive shaft.

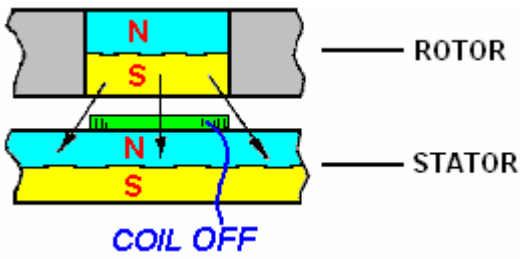
4. When that power supplied to the electromagnets is cut off, the Back EMF pulse is captured and used to charge the driving battery.

5. Although not shown in the diagram above, there are normally additional pick-up coils mounted round the rotor and if they are connected briefly at the right moment, they generate extra current and when they are switched off, their resulting reversed magnetic field also boosts the rotor on it's way, and that can raise the COP over 1000. One replication using this technique has an electrical input of 27 watts and a 32 kilowatt output.

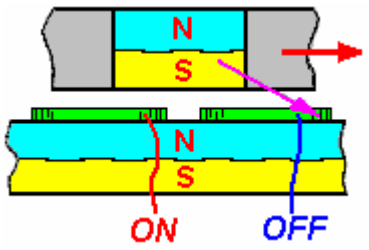
When operated in this way, the Adams Motor has a power output far in excess of the input power needed to make it run. The design confuses conventional science because conventional science refuses to accept the concept of energy flow into the motor, from the local environment. This is all the more strange, considering that windmills, water wheels, hydro-electric schemes, solar panels, wave-power systems, tidal power systems and geothermal energy systems are accepted and considered perfectly normal, in spite of the fact that they all operate on energy flowing in from the local environment. It is difficult to avoid the conclusion that vested interests are working hard to prevent conventional science accepting the fact that free-energy is all around us and there for the taking. Perhaps it is the case that they want us to go on paying for fuel to burn to "make" energy to power our homes and vehicles.

Another example of magnet power being used in the design of a powerful motor comes from **Charles Flynn**. He uses a similar method of electrical screening to prevent magnetic drag hindering the drive shaft rotation. Instead

of using electromagnets, Charles uses permanent magnets on both the rotor and the stator, and a flat coil of wire to create the blocking fields:

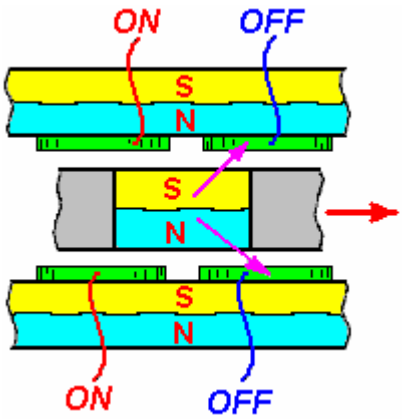


When the coil does not have current flowing through it, it does not produce a magnetic field and the South pole of the rotor magnet is attracted equally forwards and backwards by the North pole of the stator magnet. If there are two coils as shown below, and one is powered and the other is not powered, the backward pull is cancelled out and the forward pull causes the rotor to move forwards:

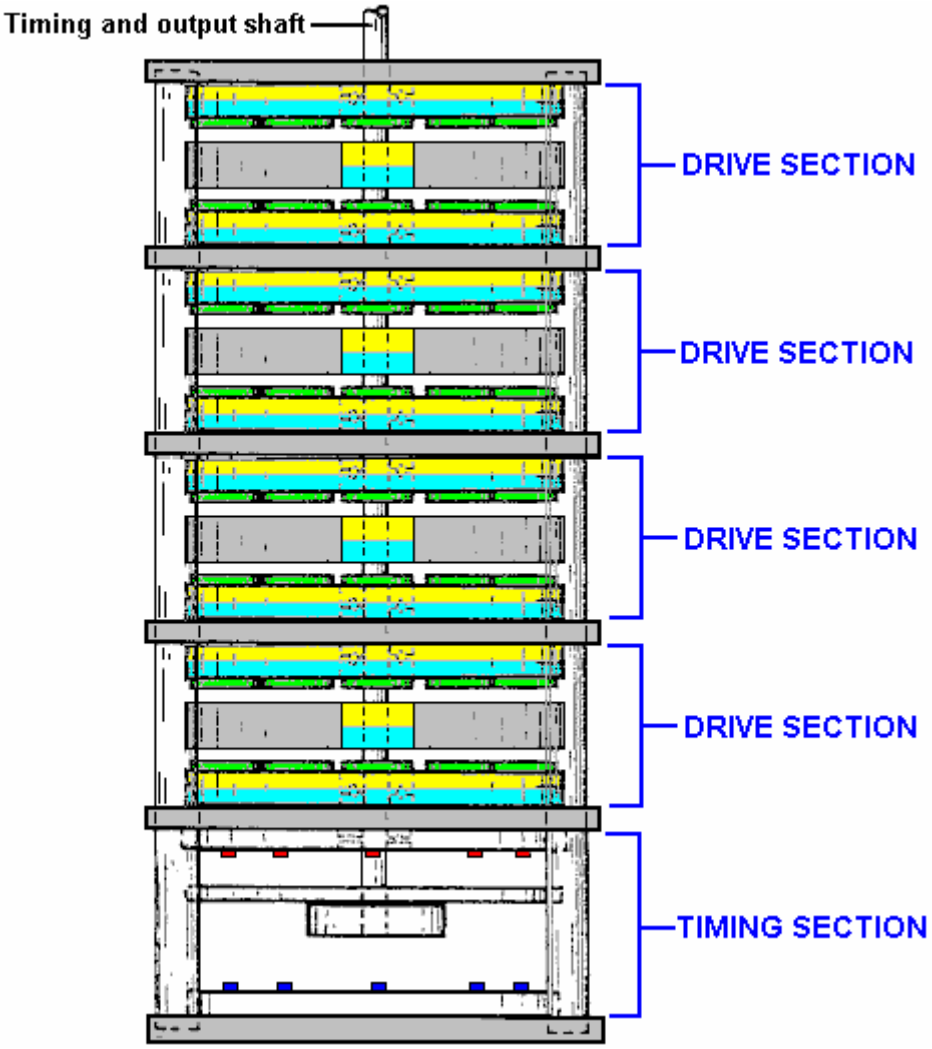


Conventional science takes a quick glance at this arrangement and proclaims that the motor efficiency has to be less than 100% because of the large electrical pulse needed to make the shaft turn. This just demonstrates a complete lack of understanding of how the motor operates. There is no "large electrical pulse" because the motor is **not** driven by electrical pulses, but instead it is driven by the attraction of many pairs of magnets, and only a very small electrical pulse is applied to cancel the backward drag as the magnets move past. To put this in context, the powerful prototype motor built by Charles ran at 20,000 rpm and the power for the coils was supplied by an ordinary 9-volt "dry-cell" battery quite incapable of supplying heavy currents.

The motor is easily made more powerful by using a stator magnet on both side of the rotor magnet, as shown here:



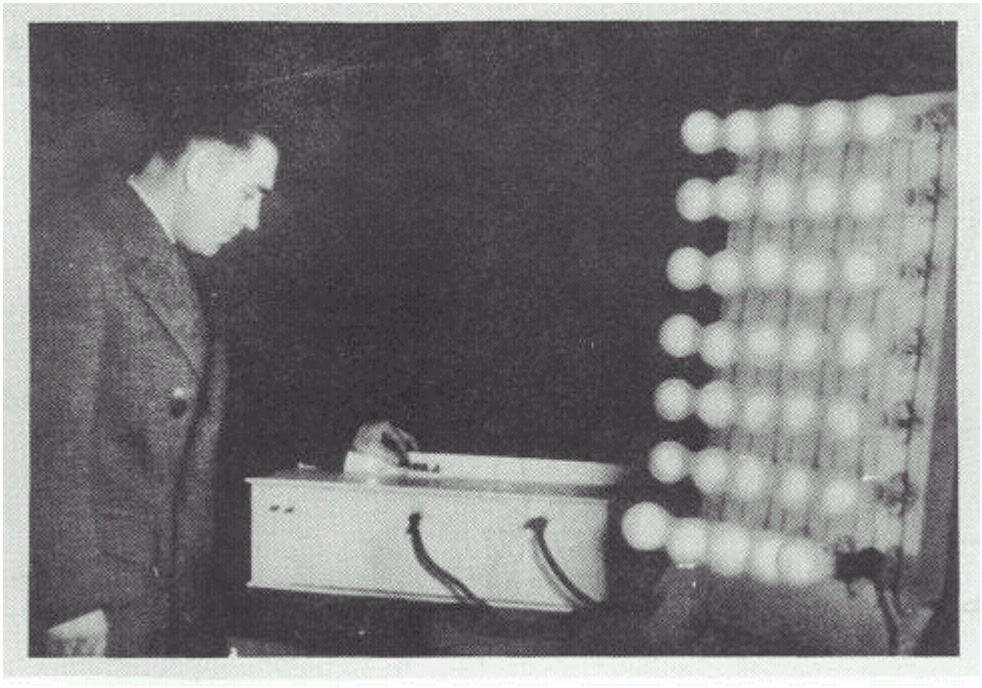
There is no real limit to the power of this motor as layer after layer of magnets can be mounted on a single drive shaft as shown here:



The electrical pulses to the screening coils can be synchronised by the light from Light-Emitting Diodes mounted in the timing section, shining through holes in a timing disc attached to the drive shaft of the motor. The light falling on light-dependant resistors on the other side of the disc, provide the switching for the coil-powering electricity.

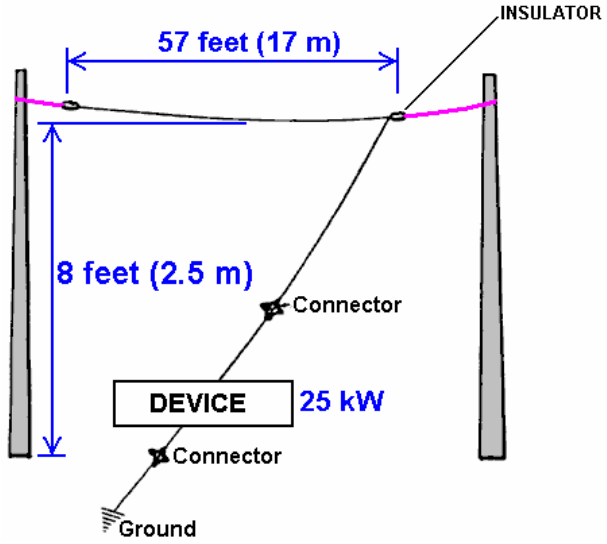
Aerial systems. We are surrounded by so much energy that a simple aerial and earth connection can draw in very large amounts of electrical power from the local environment.

Thomas Henry Moray ran frequent public demonstrations during which he lit banks of light bulbs to show that useful amounts of energy could be drawn from the environment:



Moray's device could produce output powers up to fifty kilowatts and it had no moving parts, just a simple aerial and an earth. In spite of the frequent demonstrations, some people would not believe that this was not a hoax, so Moray invited them to choose a place and he would demonstrate the power available at any location they wanted.

They drove fifty miles out into the countryside and picked a really isolated spot away from all power lines and the very few commercial radio stations in the area. They set up a very simple aerial estimated by one observer to be just fifty seven feet long and only seven or eight feet off the ground at its lowest point:



The earth connection was an eight-foot length of gas pipe which was hammered into the ground. There is no significance in the earth connection being a gas pipe, as that was used just because it was to hand at the time. The bank of lights being powered by Moray's device, grew brighter as the gas pipe was driven further and further into the ground, providing a better and better earth connection. Moray then demonstrated that when the aerial was disconnected, the lights went out. When the aerial was connected again, the lights were lit again. He then disconnected the earth wire and the lights went out and stayed out until the earth wire was connected again. The sceptics were completely convinced by the demonstration (which is most unusual for sceptics as sceptics often refuse to accept anything which contradicts their current beliefs).

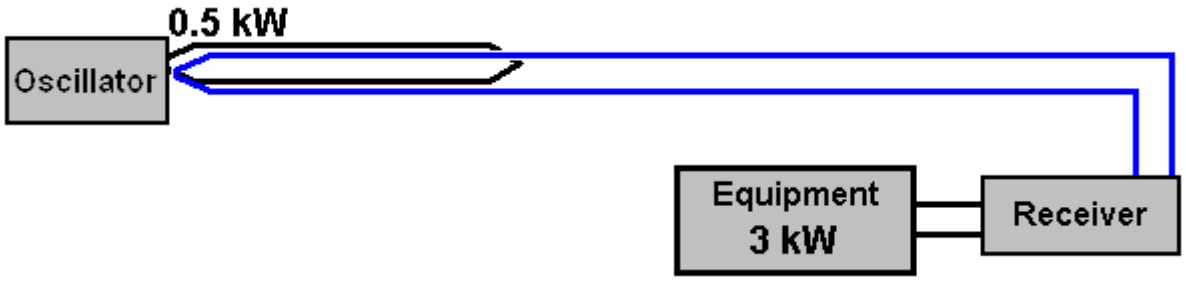
Moray's is one of several excellent and very successful devices which I can't tell you exactly how to replicate (because the details were never revealed and Moray was intimidated into silence) but the important point here is that a 57-foot aerial raised just 8-feet from the ground can provide kilowatts of electrical power at any location, if you know how to do it.

Moray's demonstrations were highly unpopular with some people and he was shot at in his car. He put bullet-proof glass in his car, so they came into his laboratory and shot at him there. They succeeded in intimidating him into stopping his demonstrations or publishing the exact details of how to replicate his aerial power system.

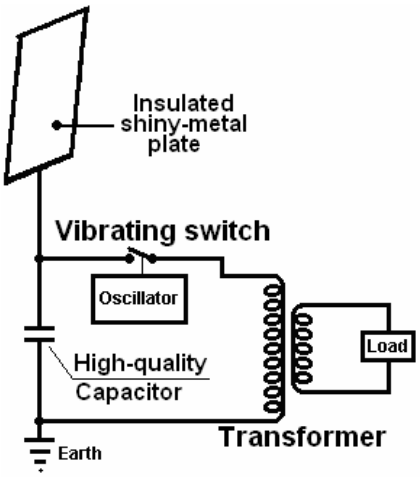
Lawrence Rayburn has recently developed an aerial system with one part raised thirty feet above the ground. He powers his farm with it and has measured more than 10 kilowatts being drawn from it.

Hermann Plauson has a patent which reads more like a tutorial on how to extract useful power from an aerial. He describes installations where one producing 100 kilowatts of excess power he calls a "small" system.

Frank Prentice has a patent on an 'aerial' system where he drives a wire loop alongside a long length of wire mounted just seven or eight inches (200 mm) above the ground. His input power is 500 watts and the power drawn from the system is 3,000 watts, giving an excess of 2.5 kilowatts (COP=6):

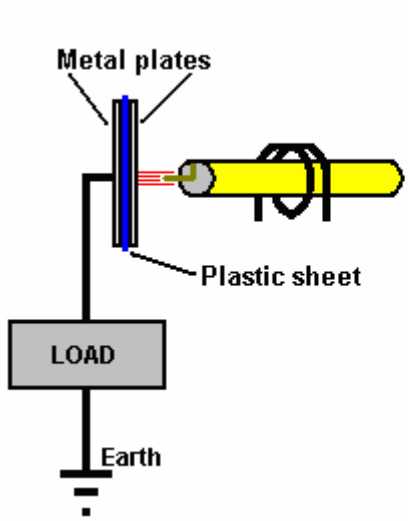


Nikola Tesla, probably the most famous person in the free-energy field, has a patent on an aerial system which uses a shiny metal plate with insulated faces as the main component of his aerial. As is common in this field, a high-quality capacitor is used to store the energy initially and then that power is pulsed through a step-down transformer which lowers the voltage and raises the current available, as shown here:

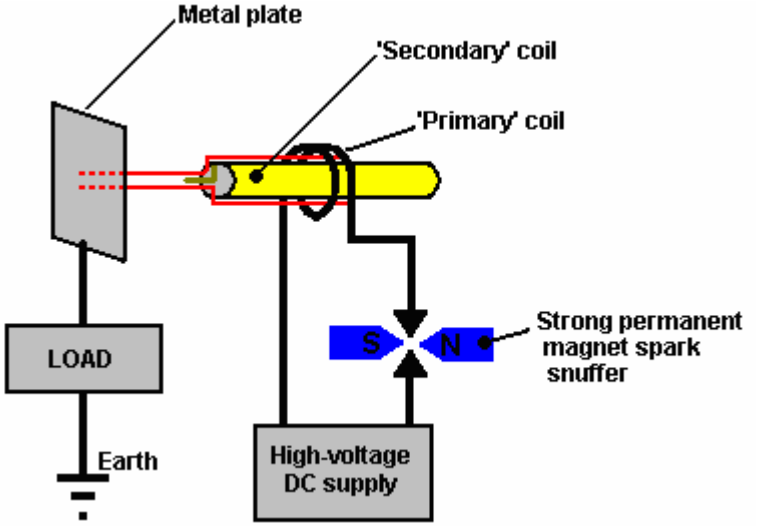


Instead of using an aerial, it is possible to use a **Tesla Coil** which produces very high currents **if** the primary winding is placed in the middle of the secondary winding and not at one end which is the usual configuration. With one method, Tesla directs the output on to a single metal plate and powers a load between the plate and the earth.

Don Smith demonstrates this in a video currently on YouTube. He uses a capacitor made from two metal plates with a sheet of plastic between them, instead of Tesla's insulated single plate. The load is powered between the capacitor and earth. The video shows Don using a 28-watt hand-held Tesla Coil and producing what looks like several kilowatts of power in the earth line.



Don Smith



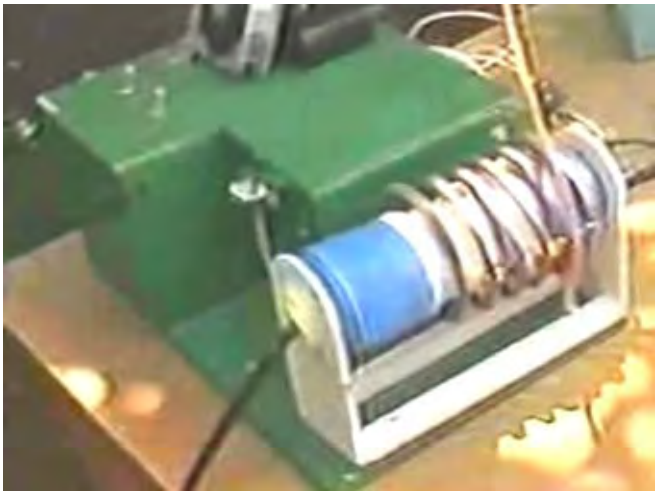
Nikola Tesla

Don points out that the output power is proportional to the **square** of the voltage and the **square** of the frequency: So if you double the frequency and double the voltage there will be 16 times as much output power.

Tariel Kapanadze demonstrates this in a web video of his interview for Turkish TV. It shows him making an earth connection by burying an old car radiator, and then lighting a row of light bulbs from a Tesla Coil style fuel-less device. While the commentary is not in English, the video is very informative. You will notice that this is a substantial power output coming from a device built with a very basic style of construction where bare wires are twisted together to form an electrical connection.

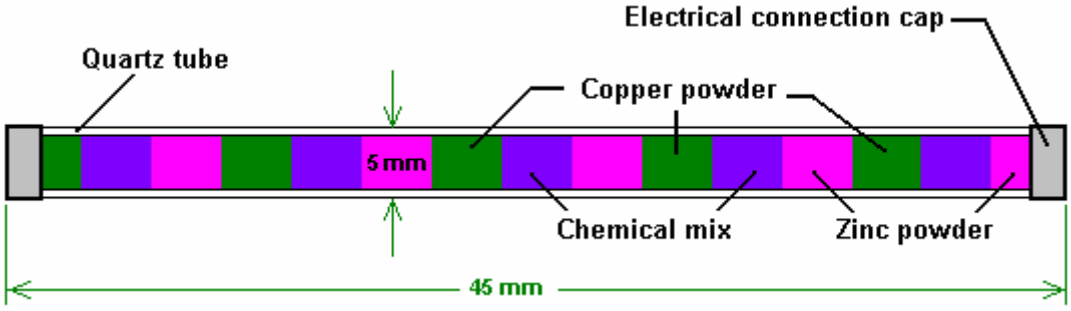
When the starting battery is removed, the equipment is held in the air to show that it is self-contained and self-powered. This is another confirmation that free-energy is all around us and ready to be taken by anyone who knows how. Tariel is seen here lighting a row of five light bulbs hanging from a broom handle placed across the backs of two chairs - not exactly a high-tech, high-cost form of construction this!

This is a picture of his circuit housing, spark gap and output transformer:



However, I can't provide you with the exact details as Tariel has never revealed how he does it and there is every indication that he never will. He says that if he told how it worked, then "you would laugh as it is so simple".

The **Colman / Seddon-Gillespie** 70-year battery. A quite different approach to getting fuel-less power was taken by Colman and Seddon-Gillespie who developed a tiny tube of harmless chemicals - copper, zinc and cadmium:

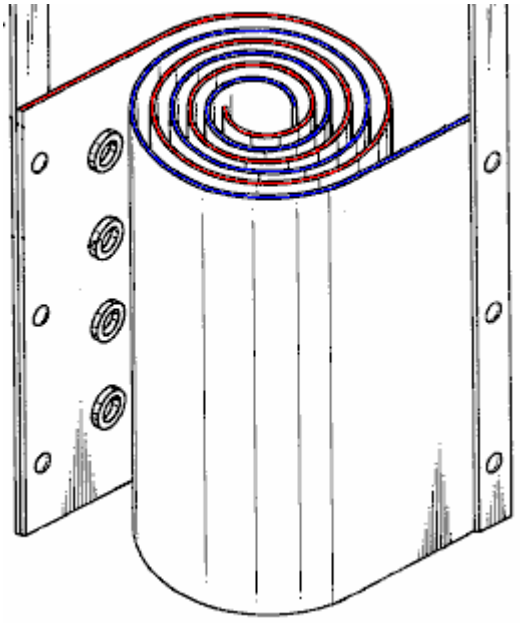


They found that if his tube was subjected to a few seconds of high-frequency electromagnetic radiation, then it became radioactive for about one hour. During that time, a kilowatt of electrical power could be drawn from this tiny tube. Near the end of the hour, another burst of electromagnetic waves keeps the tube radioactive and maintains the output current. Lead shielding is used to make this a safe device. They have a patent on this device. The expected working life of one of these tubes is estimated as being seventy years.

Electrolysis. Michael Faraday did a really excellent job of investigating how much energy was required to change water from it's liquid state into a mixture of hydrogen gas and oxygen gas. Conventional science has latched on to this information and refuses to believe that it is not the last possible word on electrolysis.

This is akin to saying that the fastest a man can propel himself over the ground is by running, and refusing to accept the fact that there might be a later invention of a bicycle which would allow a much faster human-powered speed over the ground.

This is maintained in spite of the fact that a patent has been awarded to **Shigeta Hasebe** for a different style of electrolysis, using magnets and spiral electrodes like this:



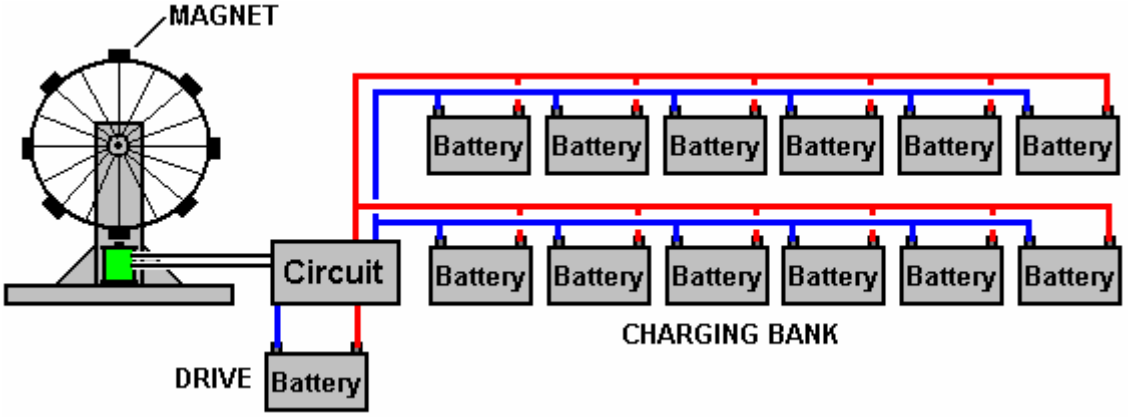
In his patent, Shigeta indicates his disappointment that his laboratory tests only showed an efficiency of ten times greater than that of Faraday while his calculations showed that he could be getting twenty times the Faraday result. The different method, along with the use of powerful magnets at the top and bottom of his electrode pairs, bypassed the limits which Faraday had established by changing the working conditions.

Stanley Meyer of the USA discovered a method of splitting water into it's gas form, using very little power. Stan's work has been replicated by Dave Lawton and many other people. For example, Dr Scott Cramton has produced the "HHO" gas mix produced by the electrolysis of water, at a rate of 6 litres per minute with a power input of just 36 watts (12 volts at 3 amps). This is dramatically better than Faraday thought was possible and it allows power production through recombining the hydroxy gas to give water again, as the power produced is well above the amount of power needed to split the water in the first place. It should be remarked in passing, that most of the power produced when HHO is recombined into water, does not come from the hydrogen (even though in its HHO form it is typically four time more energetic than hydrogen gas), but from charged water clusters which are generated during the electrolysis process.

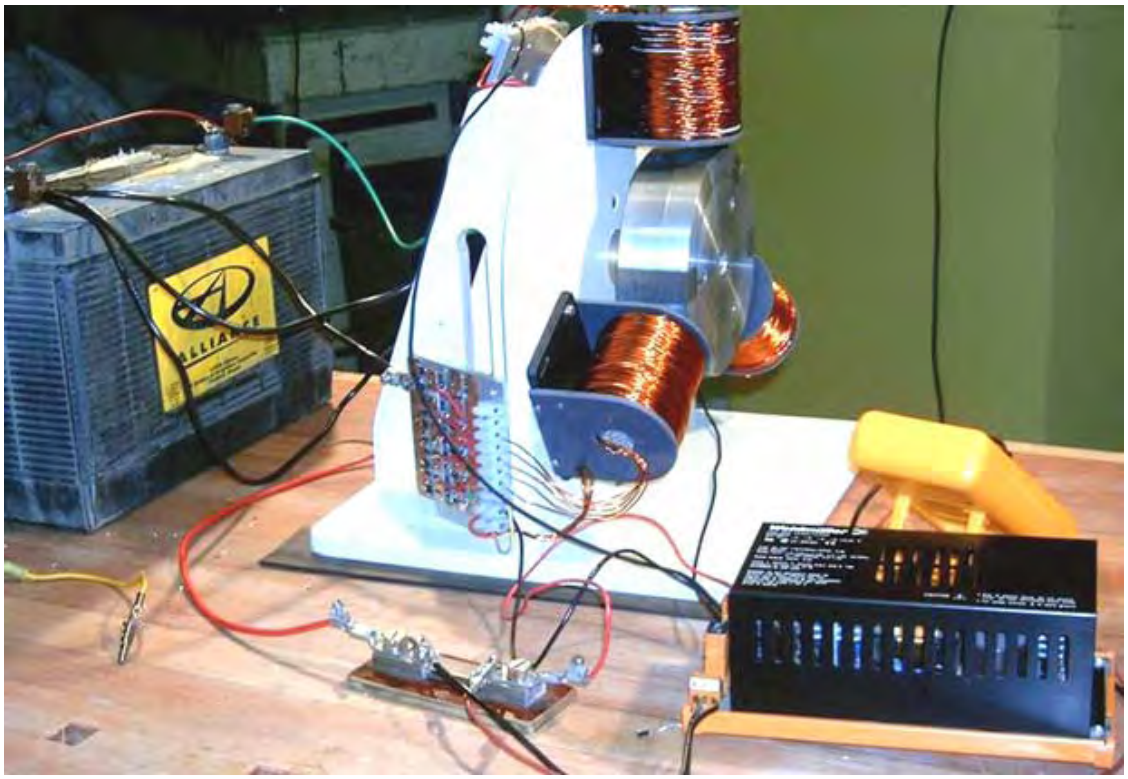
John Bedini of the USA has patented a system for the rapid charging of batteries with a pulsed waveform. Using banks of batteries tends to be very expensive, very space-consuming and replacement batteries are needed at frequent intervals, giving the user a disposal problem and additional cost. Batteries have the serious restriction that they get damaged and their life shortened if the rate of discharge is less than 20 hours. So a 100 Amp-Hour battery can only manage a 5 amp current (60 watts) if it is not to be damaged.

John Bedini's spike-generating system can charge several batteries at the same time. The snag is that you can't use the batteries to power equipment while they are being charged, so you need two sets of batteries. The system is easy to make and use. The best performance that I have come across is where there is thirteen times more power output than the power input.

There are several variations on John's pulser. The most common is a bicycle wheel with ferrite permanent magnets attached to the rim:



As the wheel spins, the approaching magnet generates a voltage in one winding of an electromagnet. This triggers a circuit which powers a second winding of the electromagnet. This pulse pushes the magnet away, keeping the wheel spinning. When the power to the coil is cut off, the resulting "Back EMF" voltage spike is fed to the batteries being charged. If the spike is sharp enough, it can cause an inflow of additional energy from the local environment. Interestingly, the rate at which the wheel rotates is directly proportional to the amount of charge in the batteries being charged. Here is a picture of **Ron Pugh**'s high-quality construction of a Bedini pulse charger:



Conclusion:
The term "Free-Energy" generally means a method of drawing power from the local environment, without the need to burn a fuel. There are many different successful methods for doing this and these methods span many countries and many years.

The amount of power which can be collected can be very high and the few kilowatts needed to power a household is most definitely within the reach of most of the devices mentioned.

****

In this brief introduction, not much detail has been given about the devices mentioned and only a small selection of devices has been covered. Much more detail is available in the various chapters of this eBook

The 'bottom line' is that energy can definitely be drawn from the local environment in sufficient quantities to supply all of our needs. For whatever reason, conventional science appears determined not to accept this basic fact and denies it at every opportunity. It seems likely that vested financial interests are the root cause of this refusal to accept the facts. The true scientific method is to upgrade scientific theory in the light of observed fact and new discoveries, but the true scientific method is not being followed at the present time.

Therefore, I invite you to examine the facts, read the information in this eBook and the additional information on the website http://www.free-energy-info.tuks.nl/ and make up your own mind on the subject. Please note that this is not a fixed body of information and this eBook normally gets a significant upgrade on a regular basis Consequently, I suggest that you download a new copy say, once per month in order to stay up to date with what is happening. Good luck with your research.

Patrick Kelly
http://www.free-energy-info.com

Chapter 1: Magnet Power

Note: If you are not at all familiar with basic electronics, you might find it easier to follow parts of this chapter if you read chapter 12 first.

One thing which we are told, is that permanent magnets can't do any work. Oh yes, magnets can support themselves against the pull of gravity when they stick on your refrigerator, but, we are told, they can't do any work. Really?

What exactly is a permanent magnet? Well, if you take a piece of suitable material like 'mild' steel, put it inside a coil of wire and drive a strong electrical current through the coil, then that converts the steel into a permanent magnet. What length of time does the current need to be in the coil to make the magnet? Less than one hundredth of a second. How long can the resulting magnet support its own weight against gravity? Years and years. Does that not strike you as strange? See how long you can support your own body weight against gravity before you get tired. Years and years? No. Months, then? No. Days, even? No.

Well if you can't do it, how come the magnet can? Are you suggesting that a single pulse for a minute fraction of a second can pump enough energy into the piece of steel to power it for years? That doesn't seem very logical, does it? So, how does the magnet do it?

The answer is that the magnet does not actually exert any power at all. In the same way that a solar panel does not put any effort into producing electricity, the power of a magnet flows from the environment and not from the magnet. The electrical pulse which creates the magnet, aligns the atoms inside the steel and creates a magnetic "dipole" which has the same effect that the electrical "dipole" of a battery does. It polarises the quantum environment surrounding it and causes great streams of energy flow around itself. One of the attributes of this energy flow is what we call "magnetism" and that allows the magnet to stick to the door of your refrigerator and defy gravity for years on end.

Unlike the battery, we do not put it in a position where it immediately destroys its own dipole, so as a result, energy flows around the magnet, pretty much indefinitely. We are told that permanent magnets can't be used to do useful work. That is not true.

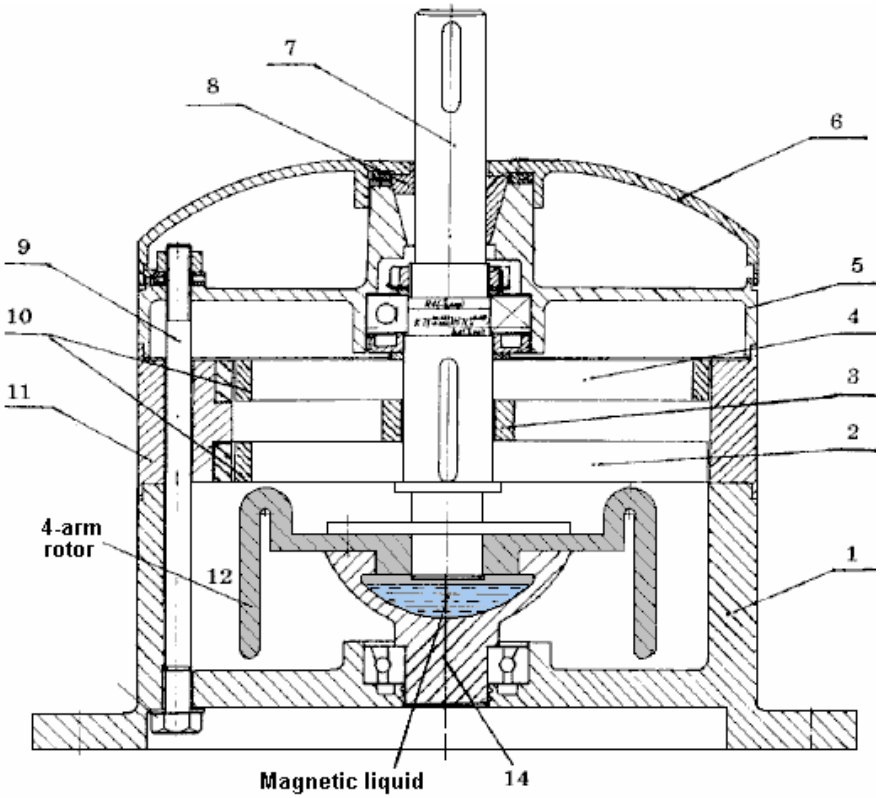
ShenHe Wang's Permanent Magnet Motor.
This is a picture of a Chinese man, ShenHe Wang, who has designed and built an electrical generator of five kilowatt capacity. This generator is powered by permanent magnets and so uses no fuel to run. It uses magnetic particles suspended in a liquid. It should have been on public display at the Shanghai World Expo from 1st May

2010 to 31st October 2010 but the Chinese government stepped in and would not allow it. Instead, they would only allow him show a wristwatch-size version which demonstrated that the design worked but which would be of no practical use in power generation:

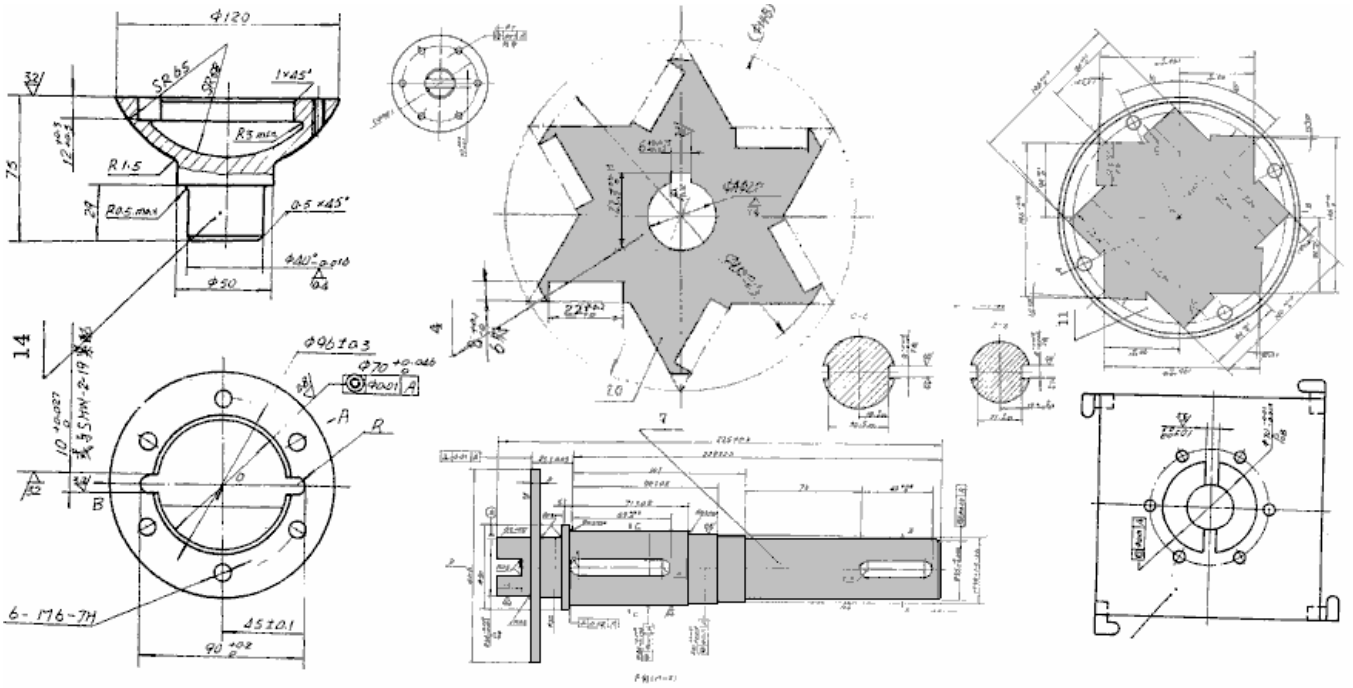
Most inventors don't seem to realise it, but almost every government is opposed to members of the public getting hold of any serious free-energy device (although they are happy to use these devices themselves). Their objective is to dominate and control ordinary people and a major factor in that is to control the supply and cost of power. A second method used everywhere is to control money, and without noticing it, governments manage to take away about 78% of people's income, mainly by concealed methods, indirect taxes, charges, fees, … If you want to know more about it, then visit www.yourstrawman.com but please understand that the reason why free-energy devices are not for sale in your local shop has to do with political control and vested financial interests and has nothing whatsoever to do with the technology. All technological problems have been solved, literally thousands of times, but the benefits have been suppressed by those in power.

Two of Mr Wang's 5 kilowatt generators successfully completed the Chinese government's mandatory six-month "Reliability and Safety" testing programme in April 2008. One large Chinese consortium has started buying up coal-fired electricity generating stations in China in order to refurbish them with pollution-free large versions of Wang's generator. Some information on the construction of the Wang motor is available here: http://www.free-energy-info.tuks.nl//Wang.pdf.

The motor consists of a rotor which has four arms and which sits in a shallow bowl of liquid which has a colloidal suspension of magnetic particles in it:



There is a patent on the motor but it is not in English and what it reveals is not a major amount.



It was Mr Wang's intention to give his motor design to every country in the world and invite them to make it for themselves. This very generous attitude does not take into account the many vested financial interests in each country, not the least of which is the government of that country, which will oppose the introduction of any device which taps into free-energy and which, consequently, would destroy their continuous streams of income. It is even possible that you would not be allowed to go to China, buy one and bring it back with you for use at home.

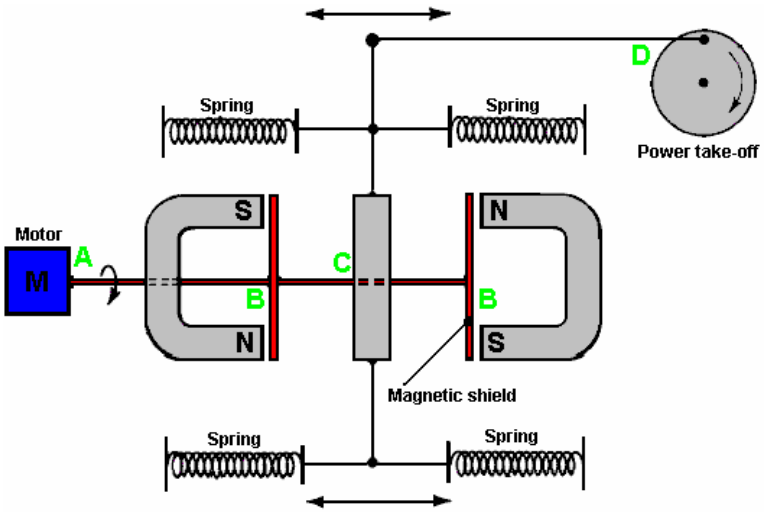
It is not easy to arrange permanent magnets in a pattern which can provide a continuous force in a single direction, as there tends to be a point where the forces of attraction and repulsion balance and produce a position in which the rotor settles down and sticks. There are various ways to avoid this happening. It is possible to modify the magnetic field by diverting it through a soft iron component.

There are many other designs of permanent magnet motor, but before showing some of them, it is probably worth discussing what useful work can be performed by the rotating shaft of a permanent magnet motor. With a home-built permanent magnet motor, where cheap components have been used and the quality of workmanship may not be all that great (though that is most definitely not the case with some home construction), the shaft power may not be very high. Generating electrical power is a common goal, and that can be achieved by causing permanent magnets to pass by coils of wire. The closer to the wire coils, the greater the power generated in those coils. Unfortunately, doing this creates magnetic drag and that drag increases with the amount of electrical current being drawn from the coils.

There are ways to reduce this drag on the shaft rotation. One way is to use an Ecklin-Brown style of electrical generator, where the shaft rotation does not move magnets past coils, but instead, moves a magnetic screen which alternatively blocks and restores a magnetic path through the generating coils. A commercially available material called “mu-metal” is particularly good as magnetic shield material and a piece shaped like a plus sign is used in the Ecklin-Brown generator.

John Ecklin’s Magnetic-Shielding Generator.

John W. Ecklin was granted US Patent Number 3,879,622 on 29th March 1974. The patent is for a magnet/electric motor generator which produces an output greater than the input necessary to run it. There are two styles of operation. The main illustration for the first is:

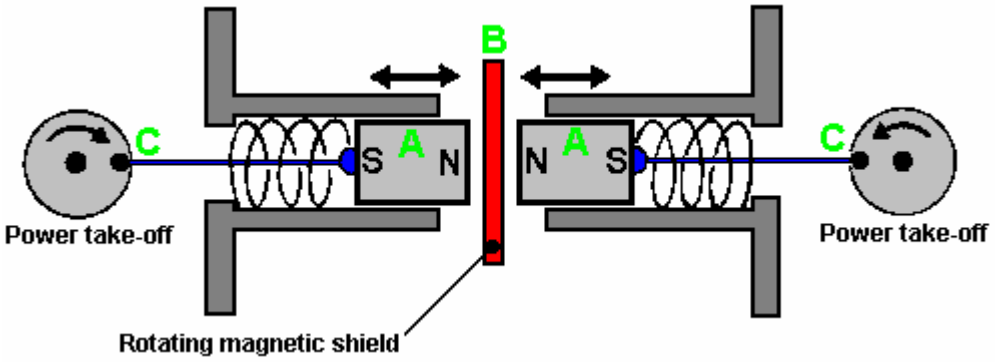


Here, the (clever) idea is to use a small low-power motor to rotate a magnetic shield to mask the pull of two magnets. This causes a fluctuating magnet field which is used to rotate a generator drive.

In the diagram above, the motor at point ‘A’ rotates the shaft and shielding strips at point ‘B”. These rectangular mu-metal strips form a very conductive path for the magnetic lines of force when they are lined up with the ends of the magnets and they effectively shut off the magnet pull in the area of point ‘C’. At point ‘C’, the spring-loaded traveller is pulled to the left when the right-hand magnet is shielded and the left hand magnet is not shielded. When the motor shaft rotates further, the traveller is pulled to the right when the left-hand magnet is shielded and the right hand magnet is not shielded. This oscillation is passed by mechanical linkage to point ‘D’ where it is used to rotate a shaft used to power a generator.

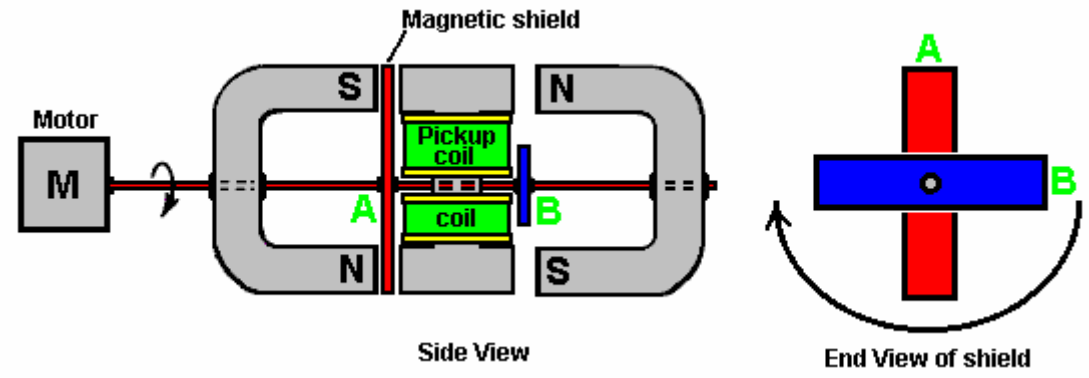
As the effort needed to rotate the magnetic shield is relatively low, it is claimed that the output exceeds the input and so can be used to power the motor which rotates the magnetic shield.

The second method for exploiting the idea is shown in the patent as:



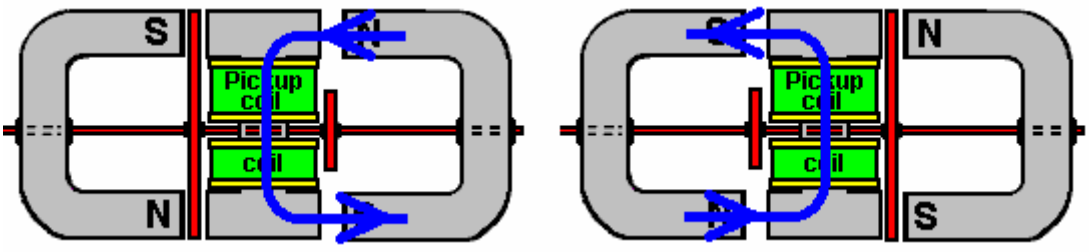
Here, the same shielding idea is utilised to produce a reciprocating movement which is then converted to two rotary motions to drive two generators. The pair of magnets 'A' are placed in a housing and pressed towards each other by two springs. When the springs are fully extended, they are just clear of the magnetic shield 'B'. When a small electric motor (not shown in the diagram) moves the magnetic shield out of the way, the two magnets are strongly repelled from each other as their North poles are close together. This compresses the springs and through the linkages at 'C' they turn two shafts to generate output power.

A modification of this idea is the **Ecklin-Brown Generator**. In this arrangement, the movable magnetic shielding arrangement provides a direct electrical output rather than a mechanical movement:



Here, the same motor and rotating magnetic shield arrangement is used, but the magnetic lines of force are blocked from flowing through a central I-piece. This I-piece is made of laminated iron slivers and has a pickup coil or coils wound around it.

The device operates as follows:

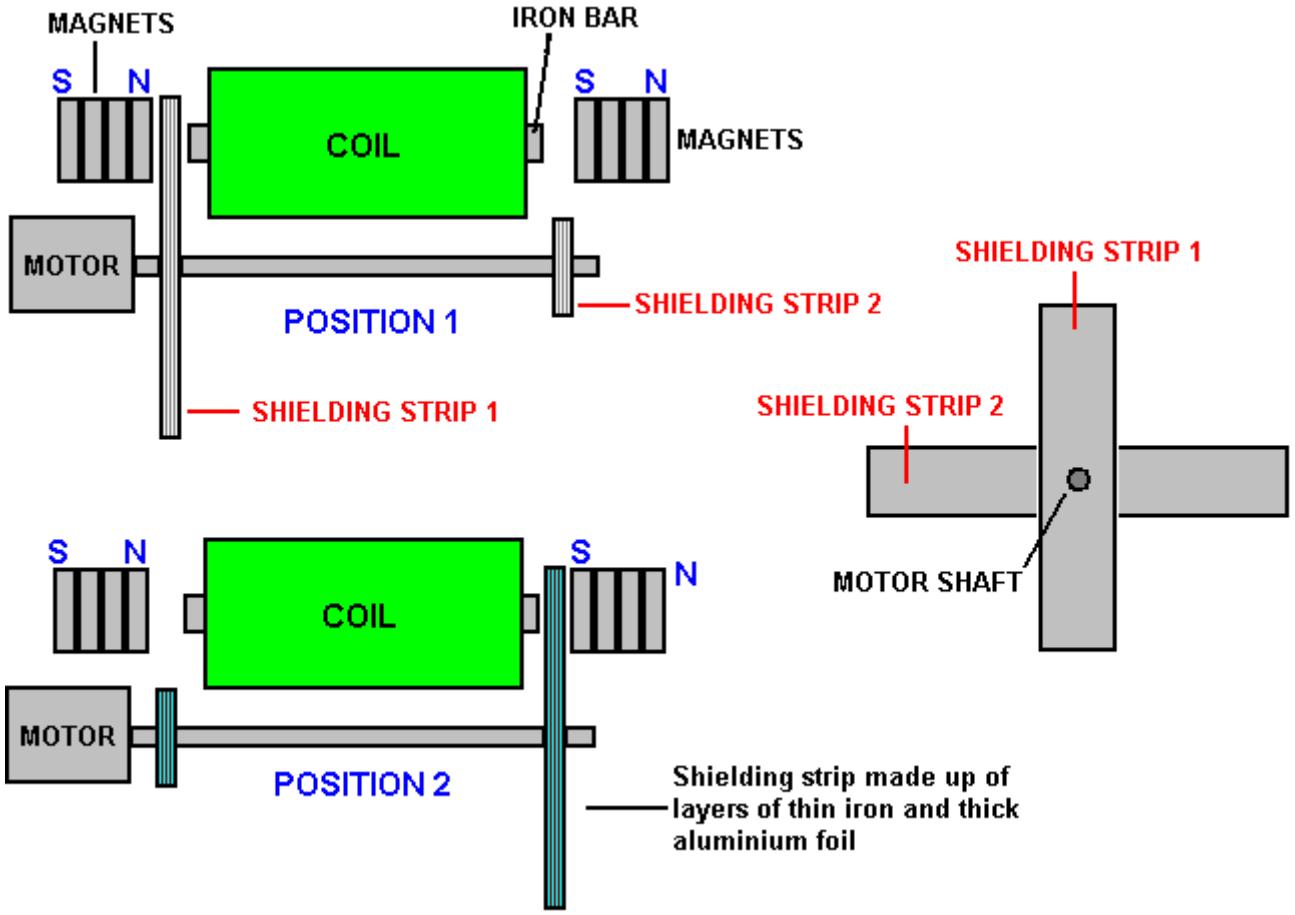


In the position shown on the left, the magnetic lines of force flow **downwards** through the pickup coils. When the motor shaft has rotated a further ninety degrees, the situation on the right occurs and there, the magnetic lines of force flow **upwards** through the pickup coils. This is shown by the blue arrows in the diagram. This reversal of magnetic flux takes place four times for every rotation of the motor shaft.

While the Ecklin-Brown design assumes that an electric motor is used to rotate the mu-metal shield, there does not seem to be any reason why the rotation should not be done with a permanent magnet motor.

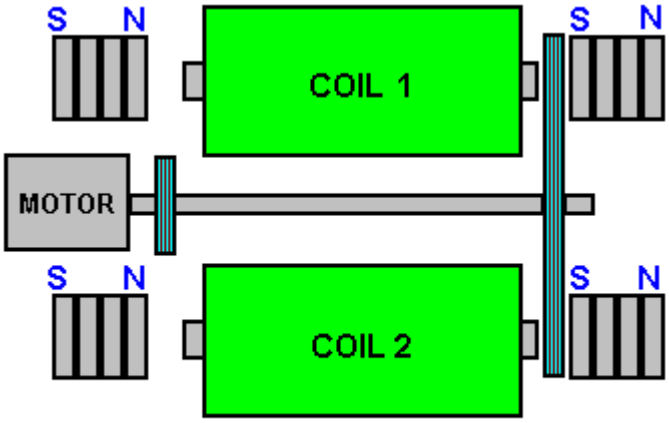
Toroidal shapes are clearly important in many devices which pull in additional energy from the environment. However, the Ecklin-Brown generator looks a little complicated for home construction, the principle can be used with a much more simple style of construction where the cores of the output coils are straight bars of suitable material such as ‘soft’ iron or perhaps the more readily available masonry anchors:

If using the masonry anchors, be sure to cut the conical end off as it alters the magnetic effect in an undesirable way. Using a hand hacksaw and a vise, cutting the end off is a very easy thing to do and that allows an ordinary helical coil to be wound either directly on the shaft or on a simple bobbin which slides on to the shaft. With any such coil, the voltage produced increases as the number of turns in the coil increases. The maximum current draw depends on the thickness of the wire as the thicker the wire, the greater the current which it can carry without overheating.

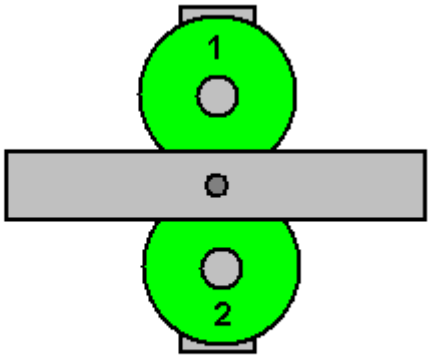


We can use an ordinary magnet or set of magnets at each end of the straight core to cause a strong magnetic field to flow through the core of our coil. As the motor spins the two screening arms they pass alternately between the magnet at one end of the core and then the magnet at the other end of the core, creating a fluctuating magnetic field passing through the coil.

The drawing shows just one output coil, but there could be two coils:

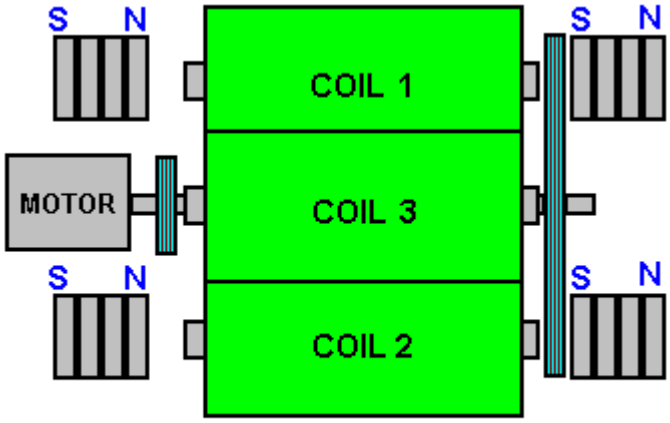


SIDE VIEW

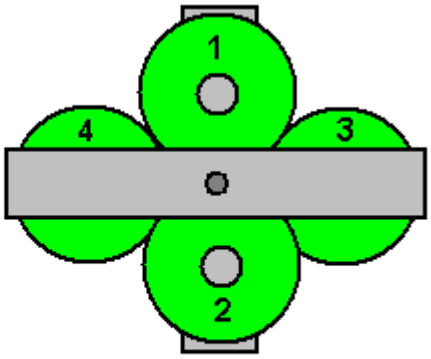


END VIEW

Or there could be four coils:



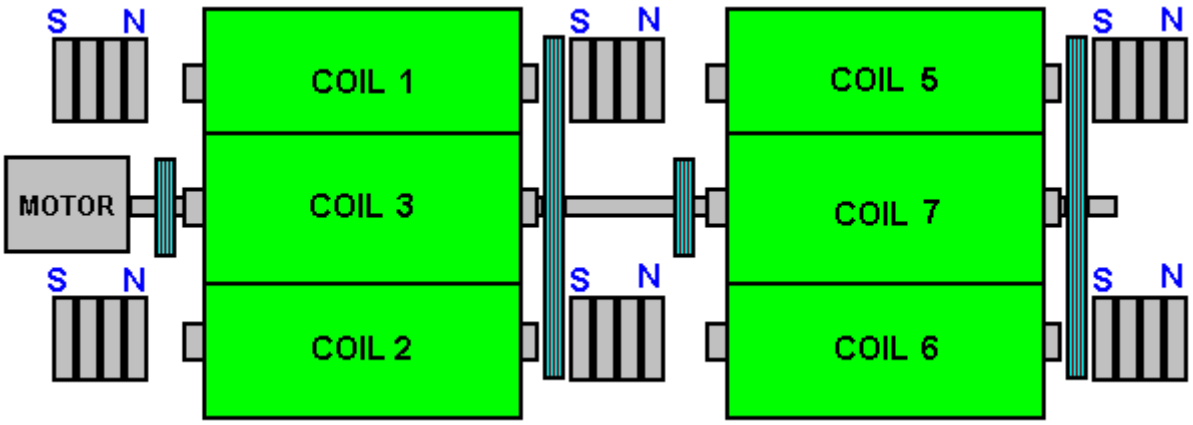
SIDE VIEW



END VIEW

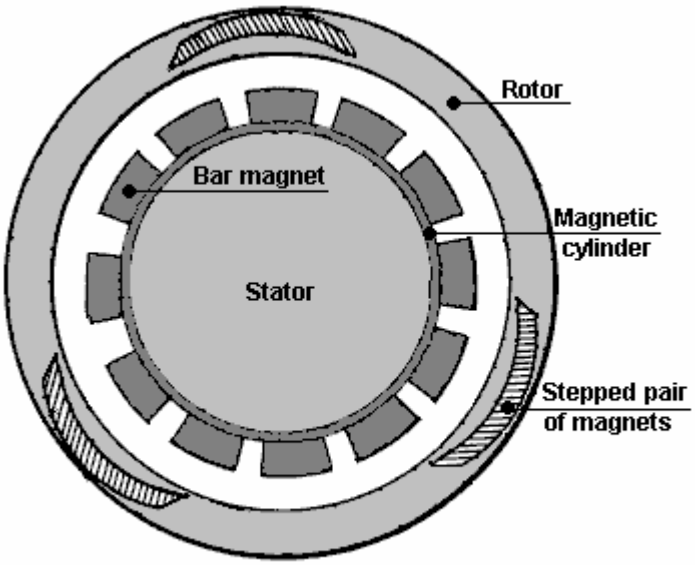
The coils can be connected in parallel to increase the output current, or they can be connected in series (in a chain configuration) to increase the output voltage. While the drawings show the shields connected directly to the motor drive shaft (a short length of plastic sleeving from a piece of wire would probably be used to help with alignment of the motor shaft and the shielding axle) there is no reason why the shielding should not be on a separate axle mounted in bearings and driven by a belt and pulley wheel arrangement.

With a separate shielding axle, allows a long, stiff axle to be used and that allows there to be additional coils and magnets. The result could be like this:



Howard Johnson's Permanent Magnet Motor.

Returning to permanent magnet motors themselves, one of the top names in this field is Howard Johnson. Howard built, demonstrated and gained US patent 4,151,431 on 24th April 1979, from a highly sceptical patent office for, his design of a permanent magnet motor. He used powerful but very expensive Cobalt/Samarium magnets to increase the power output and demonstrated the motor principles for the Spring 1980 edition of *Science and Mechanics* magazine. His motor configuration is shown here:

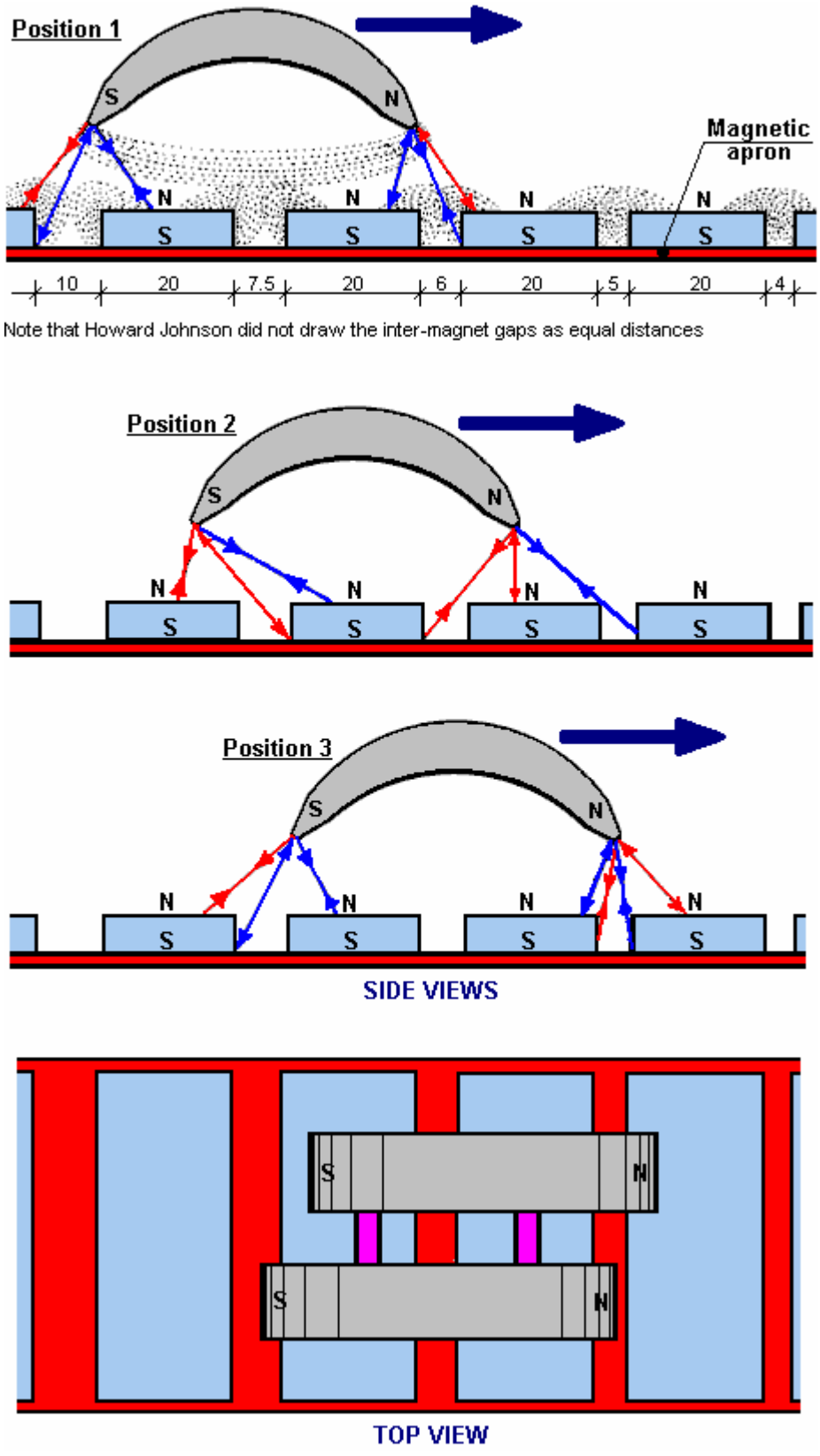


Note that the gaps between the magnets are not a constant width.

The point that he makes is that the magnetic flux of his motor is always unbalanced, thus producing a continuous rotational drive. The rotor magnets are joined in stepped pairs, connected by a non-magnetic yoke. The stator magnets are placed on a mu-metal apron cylinder. Mu-metal is very highly conductive to magnetic flux (and is expensive). The patent states that the armature magnet is 3.125" (79.4 mm) long and the stator magnets are 1" (25.4 mm) wide, 0.25" (6 mm) deep and 4" (100 mm) long. It also states that the rotor magnet pairs are **not** set at 120 degrees apart but are staggered slightly to smooth out the magnetic forces on the rotor. It also states that the air gap between the magnets of the rotor and the stator are a compromise in that the greater the gap, the smoother the running but the lower the power. So, a gap is chosen to give the greatest power at an acceptable level of vibration.

Howard considers permanent magnets to be room-temperature superconductors. Presumably, he sees magnetic material as having electron spin directions in random directions so that their net magnetic field is near zero until the electron spins are aligned by the magnetising process which then creates an overall net permanent magnetic field, maintained by the superconductive electrical flow.

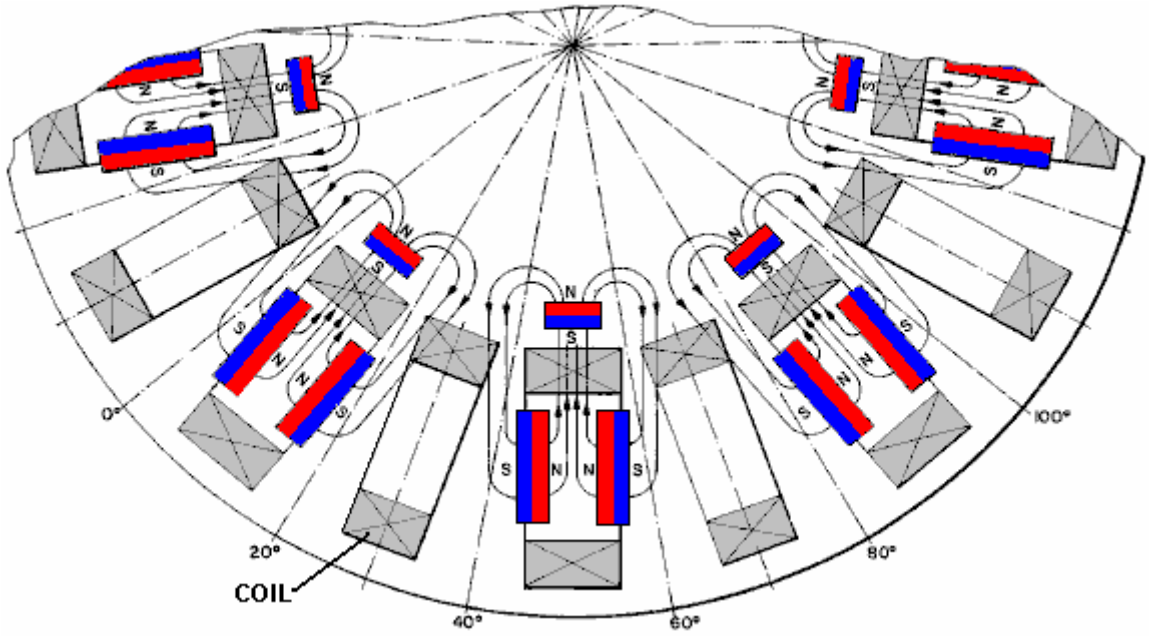
The magnet arrangement is shown here, with the inter-magnet gaps assessed from the drawing in Howard's patent:



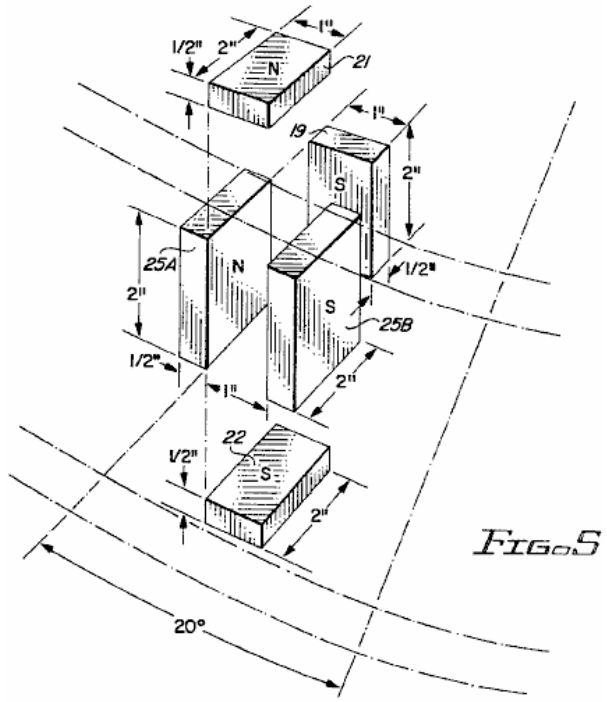
A magazine article on this can be seen at http://newebmasters.com/freeenergy/sm-pg48.html.

The "Carousel" Permanent Magnet Motor/Generator.

US Patent 5,625,241, included in the Appendix, presents the specific details of a simple electrical generator powered by permanent magnets alone. This generator can also be used as a motor. The construction is not particularly complicated:



It uses an arrangement where permanent magnets are associated with every second coil set around the rotor. Operation is self-powered and the magnet arrangement is clearly defined:



And the physical arrangement of the device is not particularly complicated:

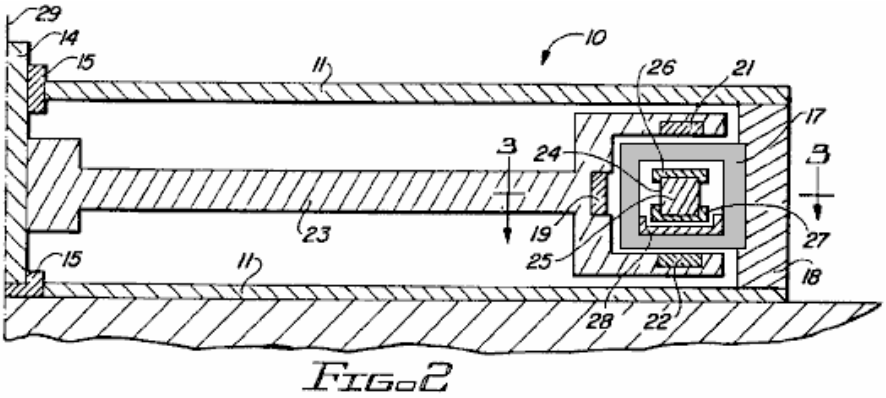
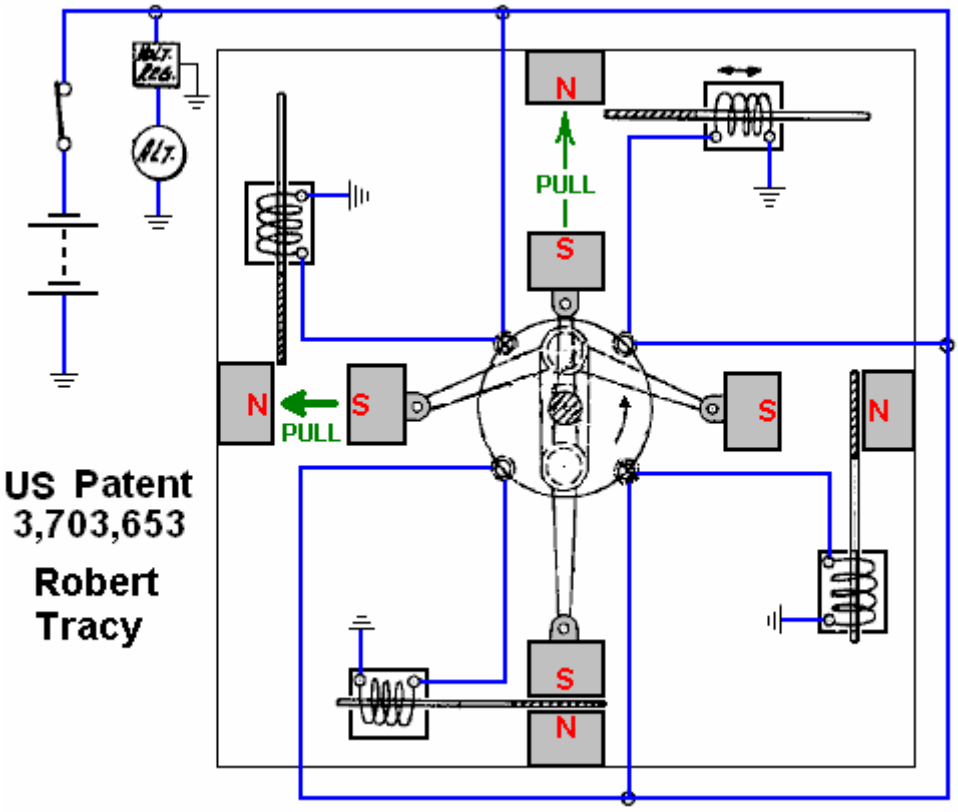


FIG. 2

This is a patent which is definitely worth reading and considering, especially since it is not a complicated presentation on the part of the authors, Harold Ewing, Russell Chapman and David Porter. This seemingly very effective generator appears to be overlooked at the present time. It seems quite clear that permanent magnet motors are a wholly viable option for the home constructor and they are capable of substantial power outputs over long periods, however, it should be noted that motors using magnets alone are notoriously difficult to get operational and while it can be done, motors which use moving shielding or pulsed electrical shielding are much more viable for the first-time constructor – motors such as the Charles Flynn motor or the Stephen Kundel motor.

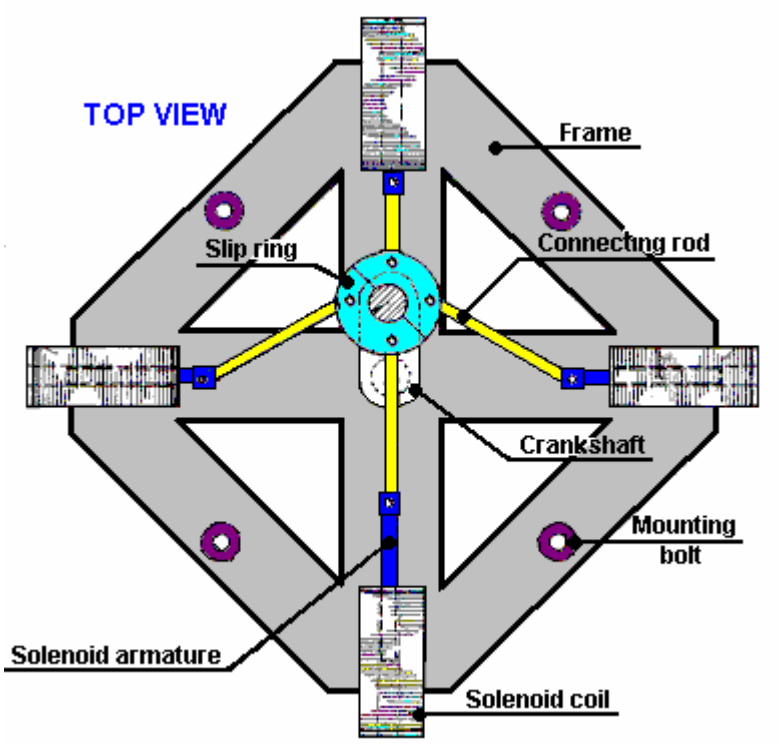
Robert Tracy's Permanent Magnet Motor.

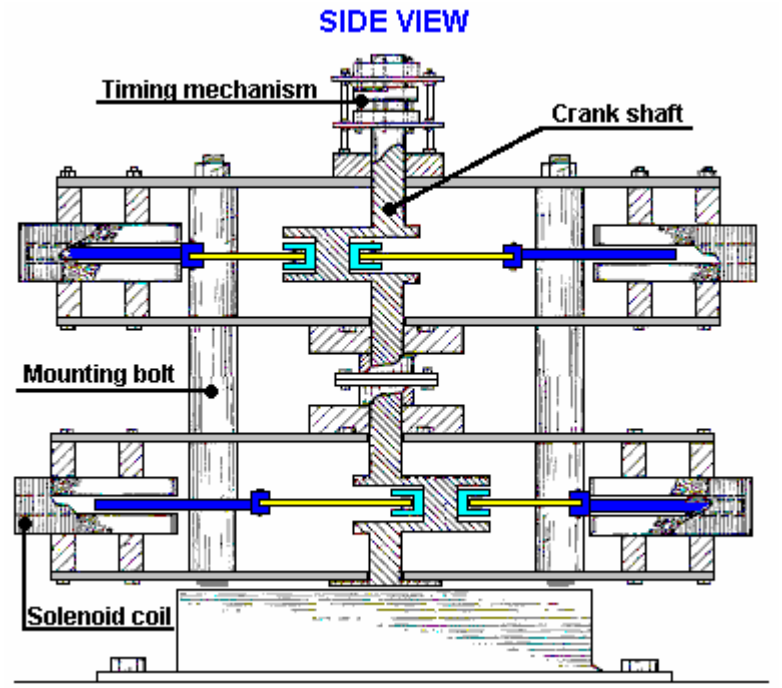
Some people have opted for permanent magnet motors where the field is shielded at the appropriate moment by a moving component of the motor. Robert Tracy was awarded US Patent Number 3,703,653 on 21st November 1972 for a "Reciprocating Motor with Motion Conversion Means". His device uses magnetic shields placed between pairs of permanent magnets at the appropriate point in the rotation of the motor shaft:



Ben Teal's Electromagnet Motor.

Motors of this kind are capable of considerable power output. The very simple motor, originally built by Ben Teal using wood as the main construction material, was awarded US Patent Number 4,093,880 in June 1978. He found that, using his hands, he could not stop the motor shaft turning in spite of it being such a very simple motor design:



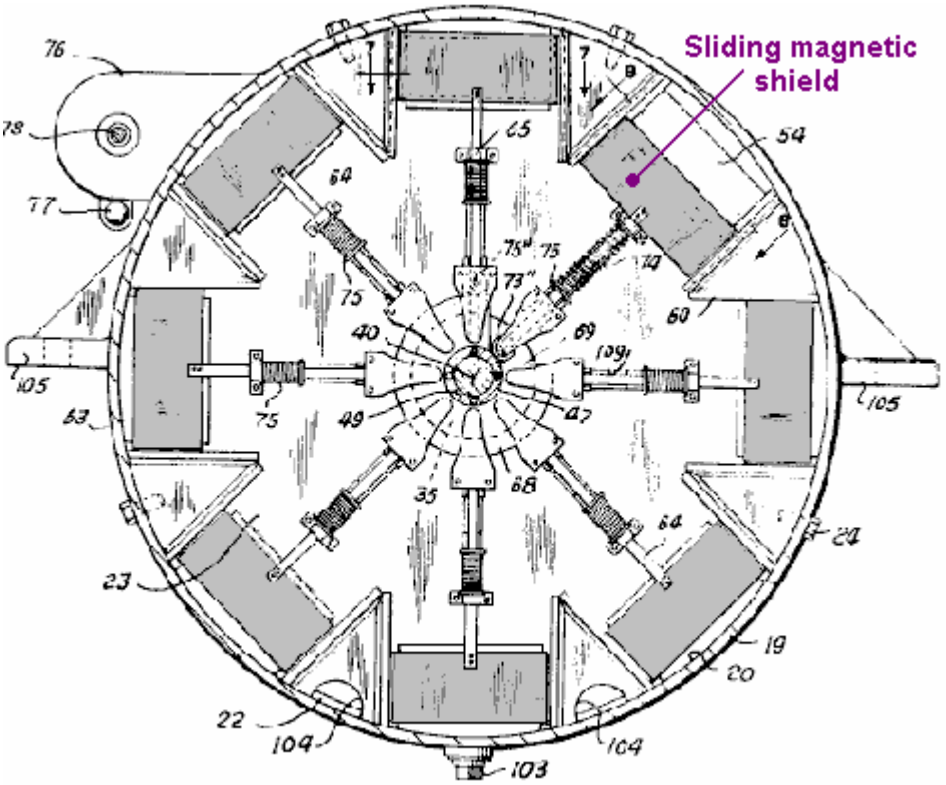


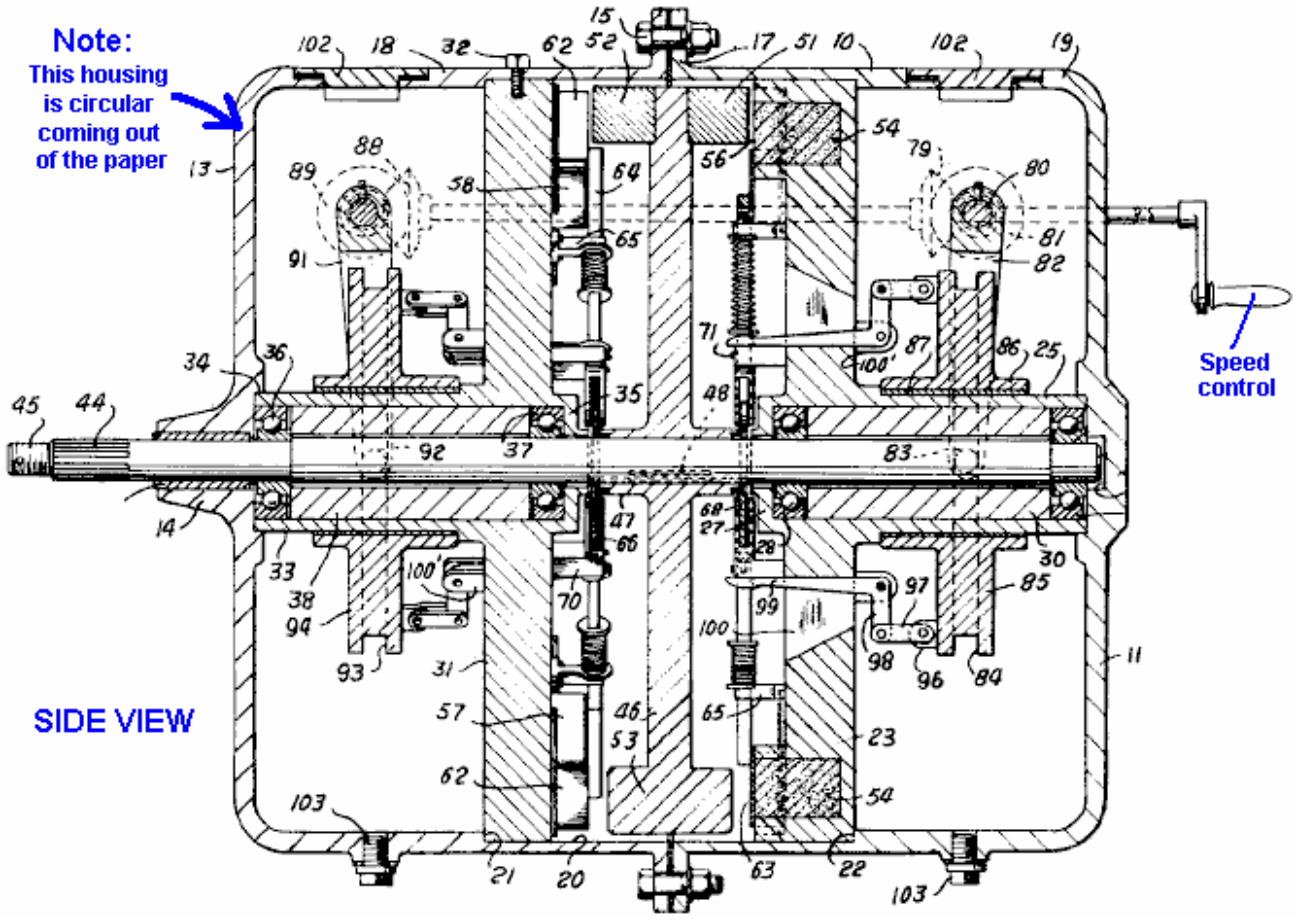
The motor operation is as simple as possible with just four switches made from springy metal, pushed by a cam on the rotor shaft. Each switch just powers it's electromagnet when it needs to pull and disconnects it when the pull is completed. The resulting motor is very powerful and very simple. Additional power can be had by just stacking one or more additional layers on top of each other. The above diagram shows two layers stacked on top of one another. Only one set of four switches and one cam is needed no matter how many layers are used, as the solenoids vertically above each other are wired together in parallel as they pull at the same time.

The power delivered by the Teal motor is an indication of the potential power of a permanent magnet motor which operates in a rather similar way by moving magnetic shields to get a reciprocating movement. Placing a resistor and capacitor across each switch contact both suppresses sparks and feeds current back to the battery when the contact opens, and this extends the battery life considerably.

The Jines Permanent Magnet Motor.

James E. Jines and James W. Jines were awarded US Patent 3,469,130 on 23rd September 1969 “Means for Shielding and Unshielding Permanent Magnets and Magnetic Motors Utilising the Same” and which is in the Appendix. This magnet motor design uses selective shielding of the drive magnets to produce a continuous force in one direction. It also has a mechanical arrangement to progressively adjust the shielding to adjust the power of the motor.

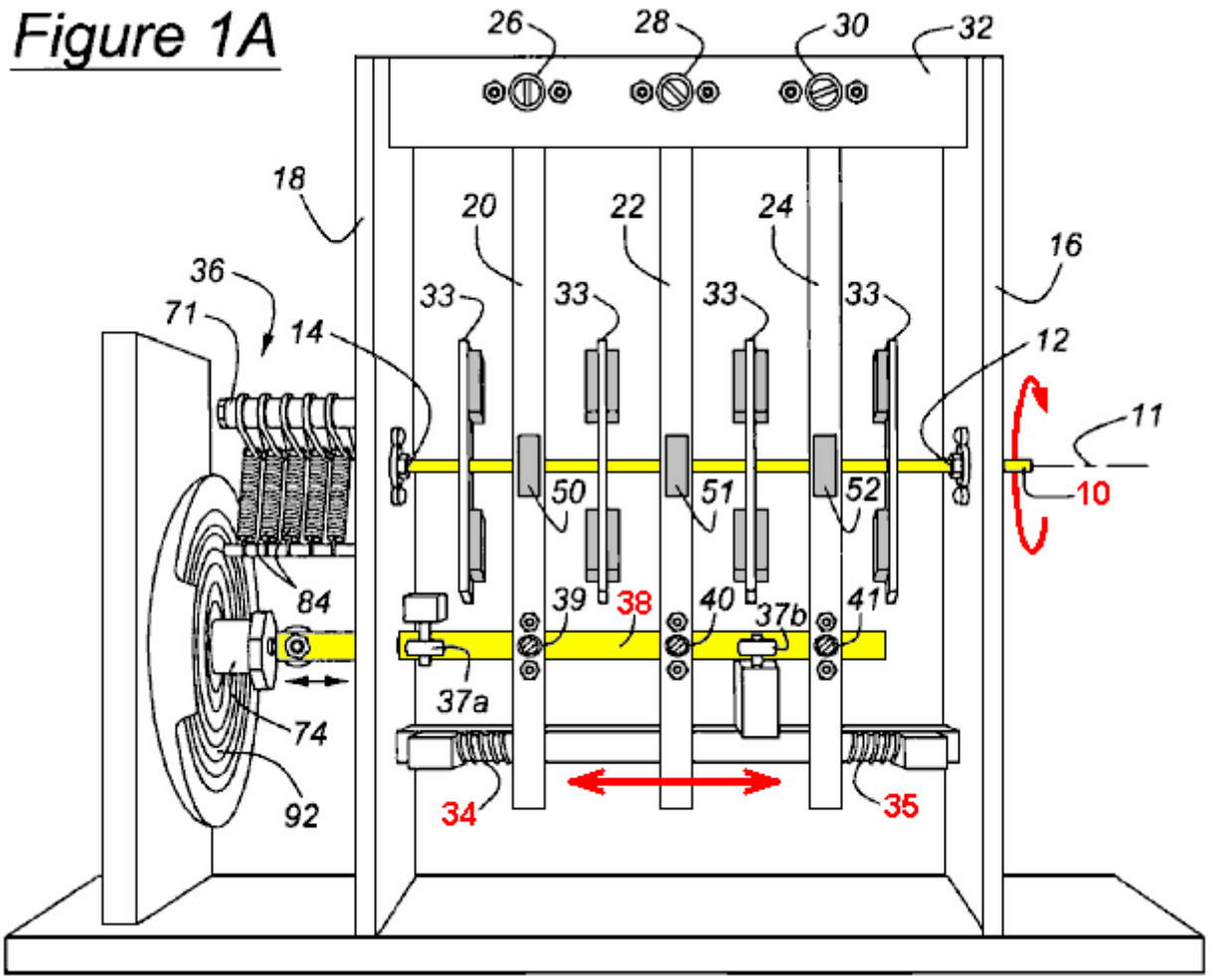




This is a very interesting design of magnetic motor, especially since it does not call for any materials which are not readily available from many suppliers. It also has the advantage of not needing any form of exact adjustment or balancing of magnetic forces to make it operate.

Stephen Kundel's Permanent Magnet Motor.

Stephen Kundel's motor design is shown in full detail in his patent which is shown on page A - 968 of the Appendix. It uses a simple oscillating motion to position the "stator" magnets so that they provide a continuous rotational force on the output shaft:



Here, the yellow arm marked **38**, rocks to the right and left, pushed by a solenoid coil **74**. There is no obvious reason why this rocking motion could not be achieved by a mechanical linkage connected to the rotating output shaft **10**. The three arms **20**, **22** and **24**, being pivoted at their upper points, are pushed into a central position by the springs **34** and **35**. The magnets **50**, **51** and **52**, are moved by these arms, causing a continuous rotation of the output drive shaft **10**. The movement of these magnets avoids the position where the magnets reach a point of equilibrium and lock into a single position.

Figure 2

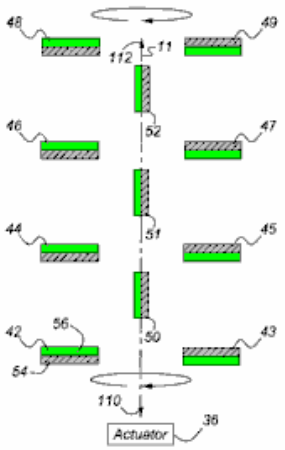
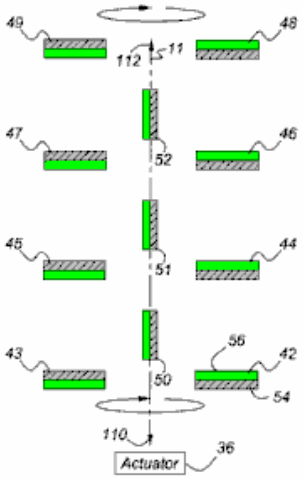


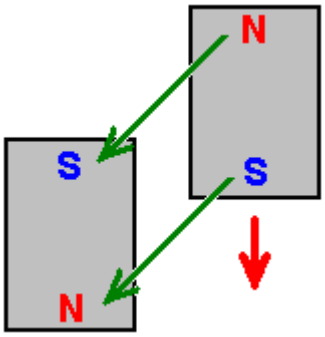
Figure 3



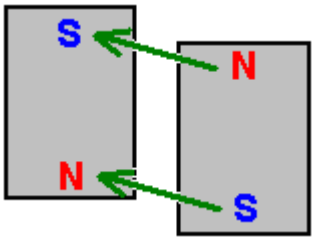
Figures 2 and 3 show the position of the magnets, with the Figure 3 position showing a point in the output shaft rotation which is 180 degrees (half a turn) further on than the position shown in Figure 2.

Some other, more powerful magnet arrangements which can be used with this design are shown in the full patent in the Appendix.

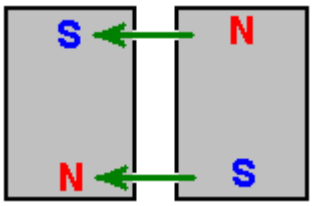
This design does not seem to appeal to many constructors in spite of the fact that it must be one of the easiest magnet motors to set up and make work. The output power level can be as big as you want as additional layers of magnets can be added. The operation is very simple and it can, perhaps, be seen more easily if just one lever arm is considered. The lever arm has just two working positions. In one position it acts on one set of rotor magnets and in the second position it acts on a second set of rotor magnets. So, we will look at each set in turn. If there are two magnets near each other, one fixed in position and the other free to move like this:



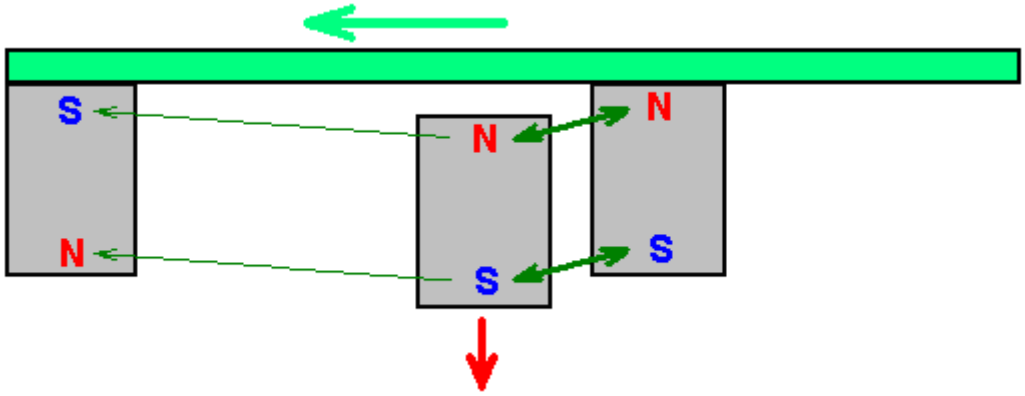
The magnets have a strong attraction to each other because of the North and South poles attracting each other. However, as the two South poles repel each other, the movement of the approaching magnet is not directly along the green arrows shown but initially is in the direction shown by the red arrow. This situation continues with the moving magnet approaching he fixed magnet and the pull between them getting stronger all the time. But, the situation changes immediately the moving magnet reaches it's closest point to the fixed magnet. Momentum starts to carry it past, but at that point the direction of the pull between the magnets starts to oppose the onward movement of the moving magnet:



If the fixed magnet remains in that position, then the moving magnet will oscillate briefly and come to a halt directly opposite the fixed magnet like this:



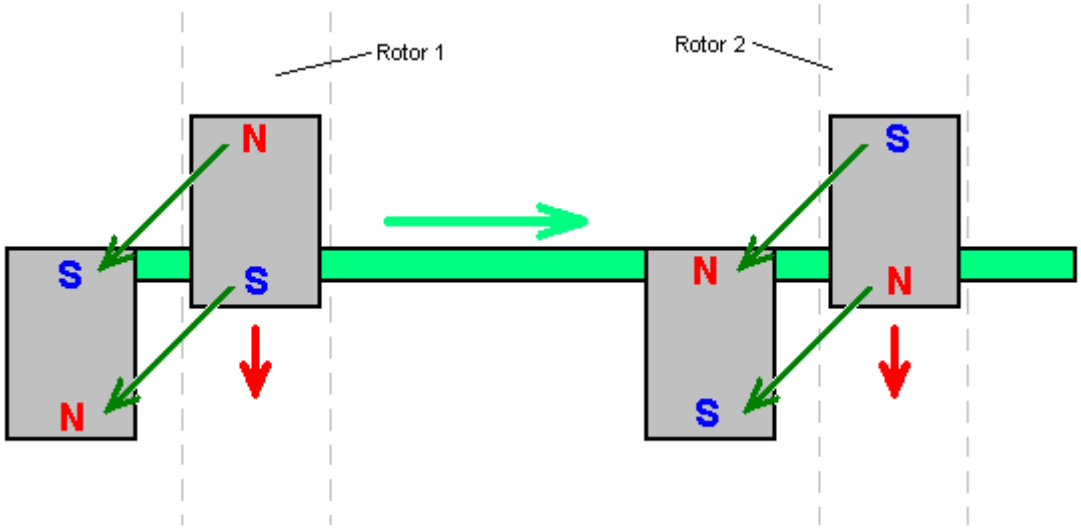
The attraction forces between the two magnets is now wholly horizontal and there is no force on the movable magnet to cause it to move. This is simple stuff, understood by anyone who has examined permanent magnets in order to see what they do. Stephen Kundel is well aware of this, and so he moves the "fixed" magnet rapidly out of the way before the reverse-direction pull slows the moving magnet down. He moves the magnet sideways and slides another one into position like this:



The new magnet is now much closer to the moving magnet and so has a much greater influence on it. The poles of the new magnet match the poles of the moving magnet which causes them to push apart very strongly, driving the moving magnet onwards in the direction it was moving in. The moving magnet moves very quickly and so gets out of the range of the fixed magnets quite quickly, at which point, the "fixed" magnets of the stator are moved back into their original position where they act in the same way on the next moving magnet attached to the rotor.

This very simple operation only requires a small force to move the stator magnets sideways between their two positions, while the force between the stator magnets and the rotor magnets can be high, producing considerable rotational power to the axle on which the rotor discs are attached.

The efficiency of the system is further boosted because when the stator magnets are in the first position shown, the second "fixed" magnet is not sitting idle but instead, it acts on the magnet of the next rotor disc:



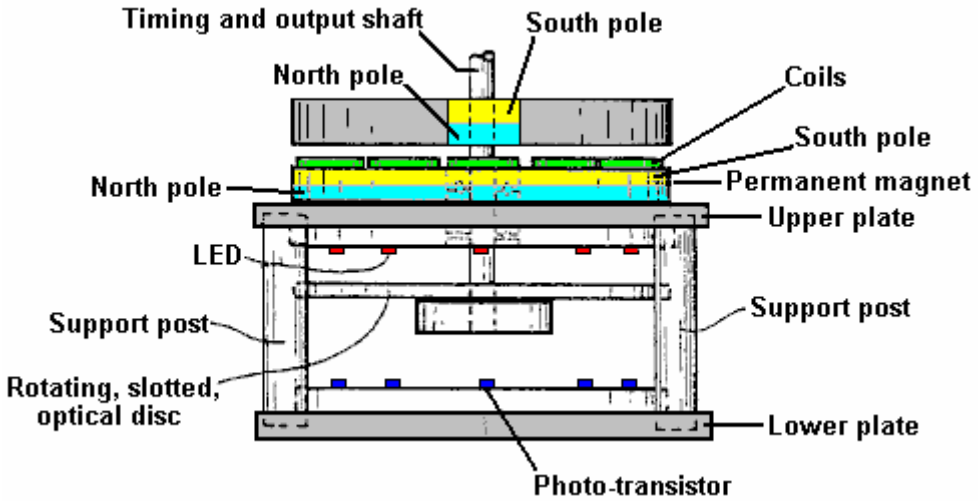
For this, the magnets attached to Rotor disc 2 have to be positioned so that their poles are the reverse of those attached to Rotor disc 1. Stephen uses a loudspeaker to wobble the horizontal bar on which the stator magnets are mounted, backwards and forwards as a loudspeaker has that mechanism already built into it. Don Kelly's

permanent magnet motor also uses this very simple idea of moving the stator magnets out of the way at the appropriate moment.

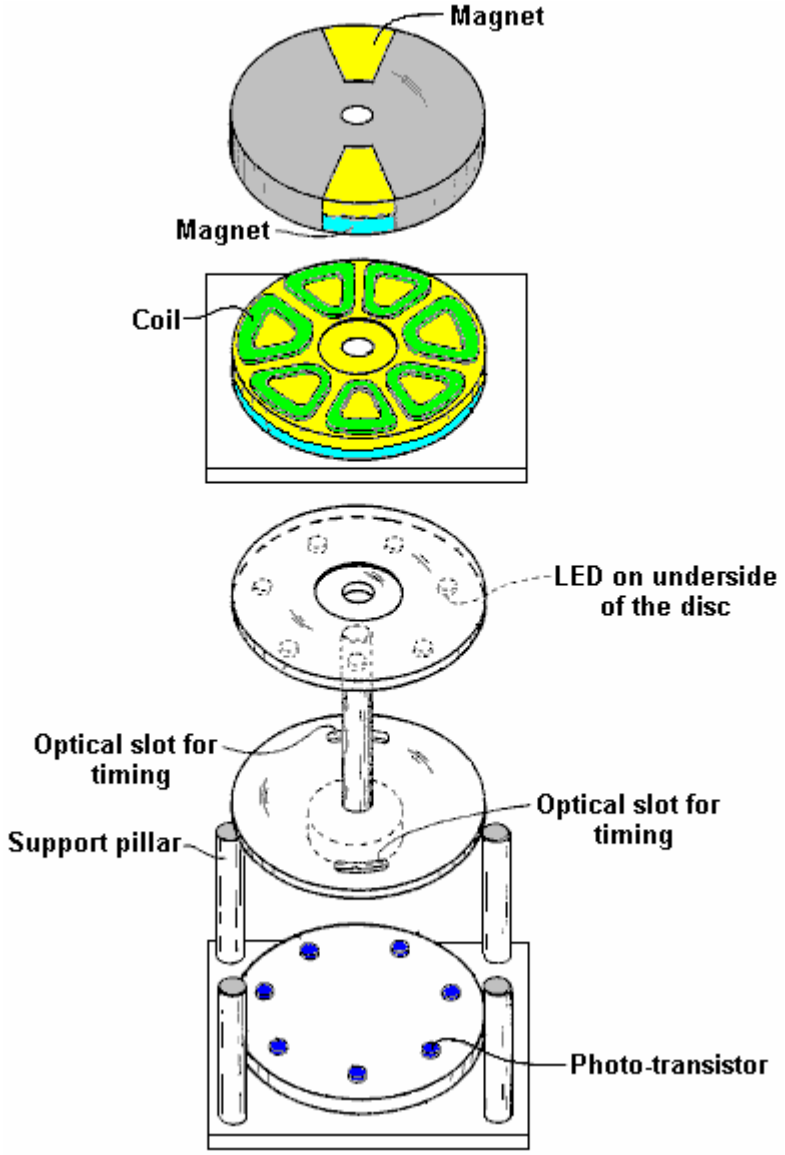
Charles “Joe” Flynn’s Permanent Magnet Motor.

Patent US 5,455,474 dated 3rd October 1995 and shown in full in the Appendix, gives details of this interesting design. It says: “This invention relates to a method of producing useful energy with magnets as the driving force and represents an important improvement over known constructions and it is one which is simpler to construct, can be made to be self starting, is easier to adjust, and is less likely to get out of adjustment. The present construction is also relatively easy to control, is relatively stable and produces an amazing amount of output energy considering the source of driving energy that is used. The present construction makes use of permanent magnets as the source of driving energy but shows a novel means of controlling the magnetic interaction or coupling between the magnet members and in a manner which is relatively rugged, produces a substantial amount of output energy and torque, and in a device capable of being used to generate substantial amounts of energy.”

The patent describes more than one motor. The first one is like this when seen from the side:

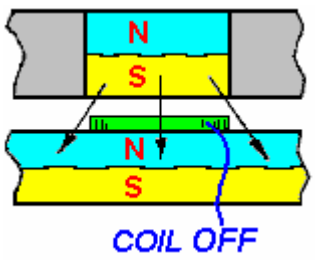


An exploded view, shows the different parts clearly:

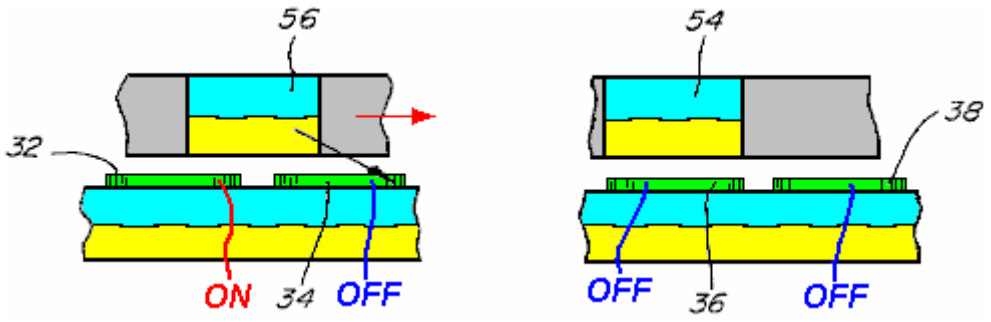


This construction is relatively simple and yet the operation is powerful. The power is provided by three magnets, shown shaded in blue and yellow. The lower magnet is in the form of a disc with the poles arranged on the large, circular, flat faces. This is the stator magnet which does not move. Positioned above it is a disc made of non-magnetic material (shaded in grey) and which has two magnets embedded in it. This disc is the rotor and is attached to the central vertical shaft.

Normally, the rotor would not rotate, but between the two discs there is a ring of seven coils which are used to modify the magnetic fields and produce powerful rotation. The powering up of these coils is very simple and it is arranged by shining a beam of Ultra Violet light from one of the Light-Emitting Diodes through a slot in an optical-timing disc attached to the rotating shaft. The LEDs and the photo-transistors are aligned with the centres of the seven coils. The position and width of the slot controls which photo-transistor gets switched on and for how long it remains powered up. This is a very neat and compact arrangement. The really interesting part of the design is how the coils modify the magnetic fields to produce the output power of the device. The orientation of the magnet poles can be swapped over, provided that this is done for all three magnets.



Shown here is the situation when one of the rotor magnets has rotated to where it is above one of the coils which is not yet powered up. The South pole of the rotor magnet is attracted to the North pole which is the entire upper face of the stator magnet as shown by the three arrows. If a voltage is applied to the coil, then this magnetic coupling is disrupted and altered. If any torque is developed as a result of the coil being powered up, then it will be developed to either side of the energised coil. If the coil is not powered up, then there will be full attraction between the magnets and no rotational force will be produced. You will notice that there are two rotating magnets (an even number) and seven coils (an odd number) so when one of the rotor magnets is above a coil, then the other isn't. This staggering of the two positions is essential for generating smooth, continuous rotational torque and self-starting without any need to rotate the shaft manually.

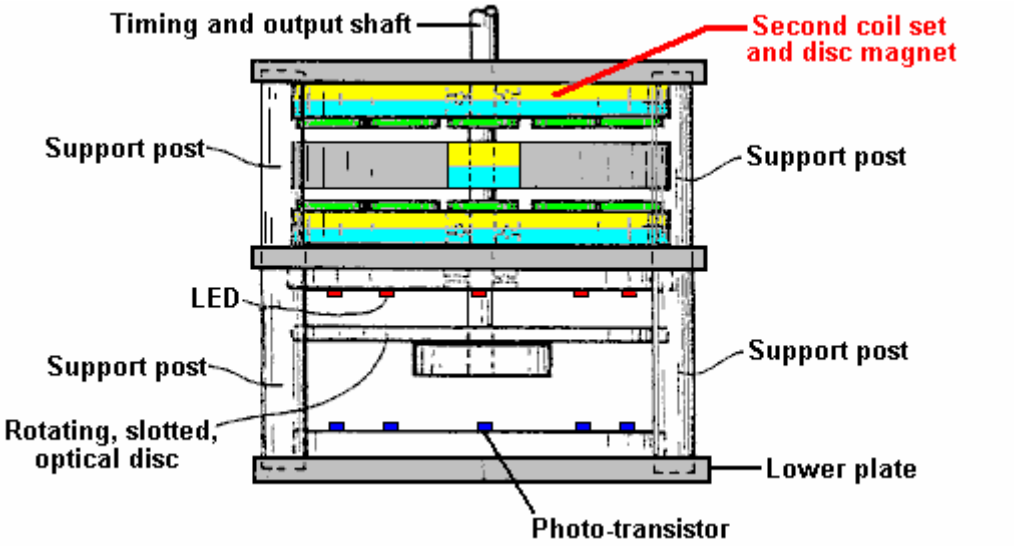


The diagram above shows a piece from both sides of the rotor disc, to explain the operation of the coils. On the left, magnet **56** overlaps coil **32** and coil **34**. Coil **32** is powered up and this breaks the magnetic link on the left hand side of magnet **56**. But, coil **34** is not powered up, so the attraction between magnet **56** and the disc magnet under the coils remains. Even though this attraction is at a downward angle, it creates a push on the rotor, driving it towards the right as shown by the red arrow.

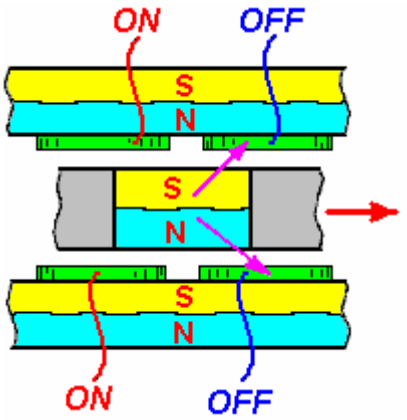
While this is happening, the situation around the other side of the rotor disc, is shown on the right. Here, magnet **54** is above coil **36** and that coil is not powered up, so there is no resulting drive in either direction - just a downward pull on the rotor magnet, towards the stator magnet below it. The adjacent coil **38** is also not powered up and so has no effect on the rotation. This method of operation is **very** close to that of the motor design of Robert Adams described in the next chapter. It is important to understand that this method of operation is nothing like that of the John Bedini pulsers where the rotation of a disc is caused by the electrical pulse applied to a coil creating a repulsion thrust to a rotor magnet. Instead, here, the coil acts as a magnetic shield, being provided with the minimum possible power to do its job. The coil is, in effect, a shield which has no moving parts, and so is a very clever mechanism for overcoming the tendency for the rotor magnets to lock on to the stator magnets and preventing rotation.

At any moment, six of the seven coils in this design are inactive, so in effect, just one coil is powered. This is not a major current drain. It is important to understand that the power of this motor is provided by the permanent magnets pulling towards each other. Each of the two magnets applies a horizontal pull on the rotor every seventh of a turn, that is, every 51.1 degrees in the rotation. As the coils are an uneven number, the rotor gets a magnetic pull every 25.5 degrees in the rotation, first from one rotor magnet and then from the other rotor magnet.

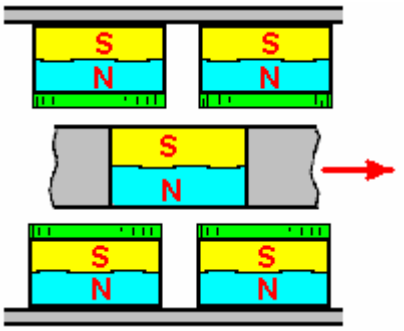
It follows then, that the power of the motor can be increased by adding more magnets. The first step in this search for additional power is to add a second disc magnet and coils on the other side of the rotor, so that there is a second pull on the magnet. This has the added advantage that it balances the downwards pull of the first disc magnet with an upward pull, giving an enhanced and balanced horizontal thrust as shown here:



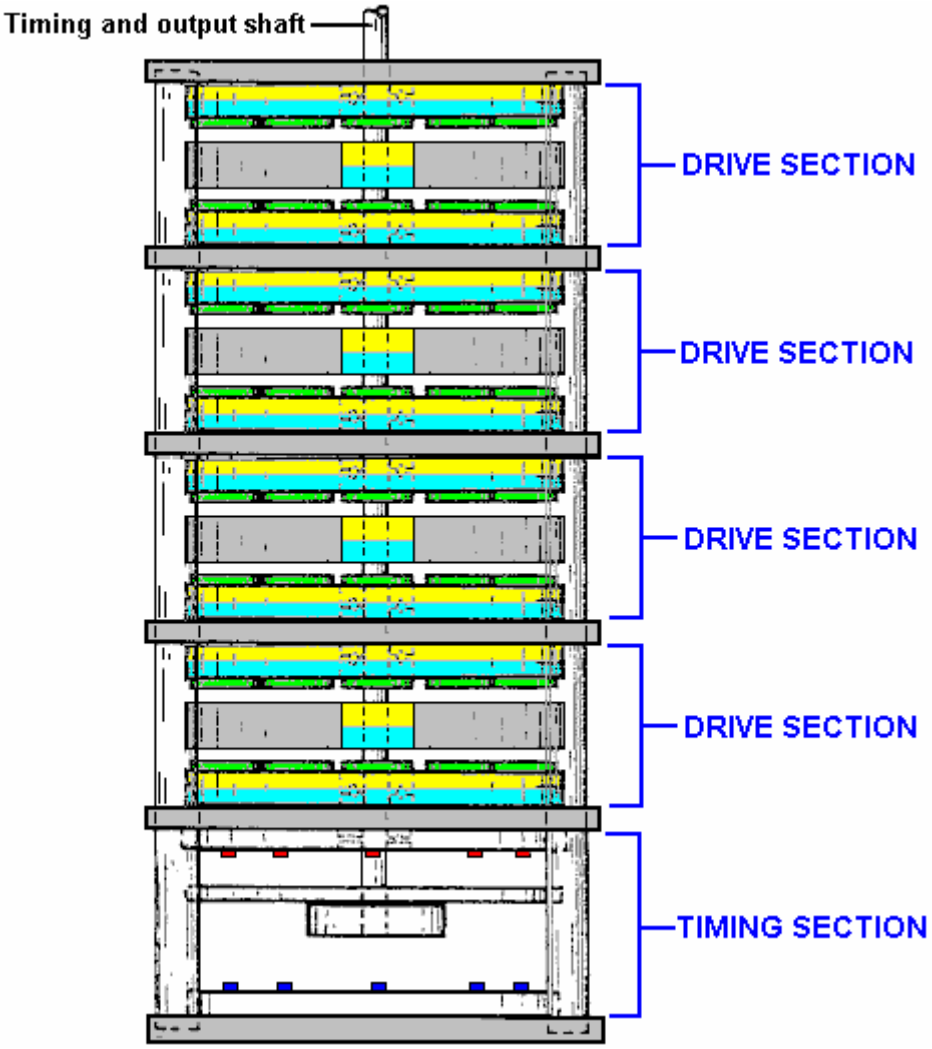
The coil switching with the additional layer of coils is shown here:



This produces a larger horizontal thrust. While this design goes for optimum performance, I suggest that a much more simple form of construction with a ring of standard circular neodymium magnets could be used instead of one large disc magnet, and ordinary circular coils placed on top of the circular magnets, and this allows large diameter rotors to be constructed, the larger diameter giving greater output shaft power:

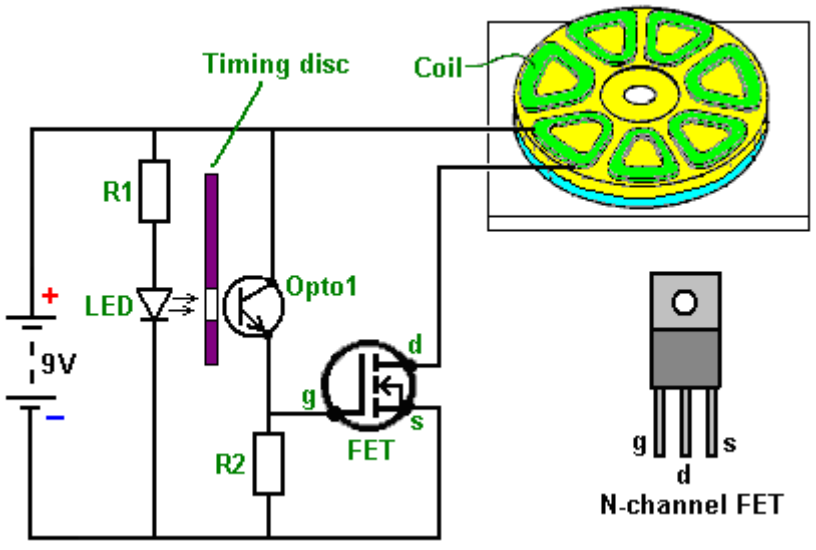


To increase the power of the output shaft further again, additional sets of magnets and coils can be added as shown here:



It should be remembered that the timing section shown above could be replaced by a NE555 timer circuit which generates a steady stream of On / Off pulses. When those pulses are fed to the coils, the motor rotates, slaving itself to the pulse rate. This gives an immediate speed control for the motor as well as avoiding the need for the precise positioning of the slotted disc which allows the LEDs to shine directly on to the phototransistors at the appropriate instant. If that approach is taken, then the timing section shown above would be omitted.

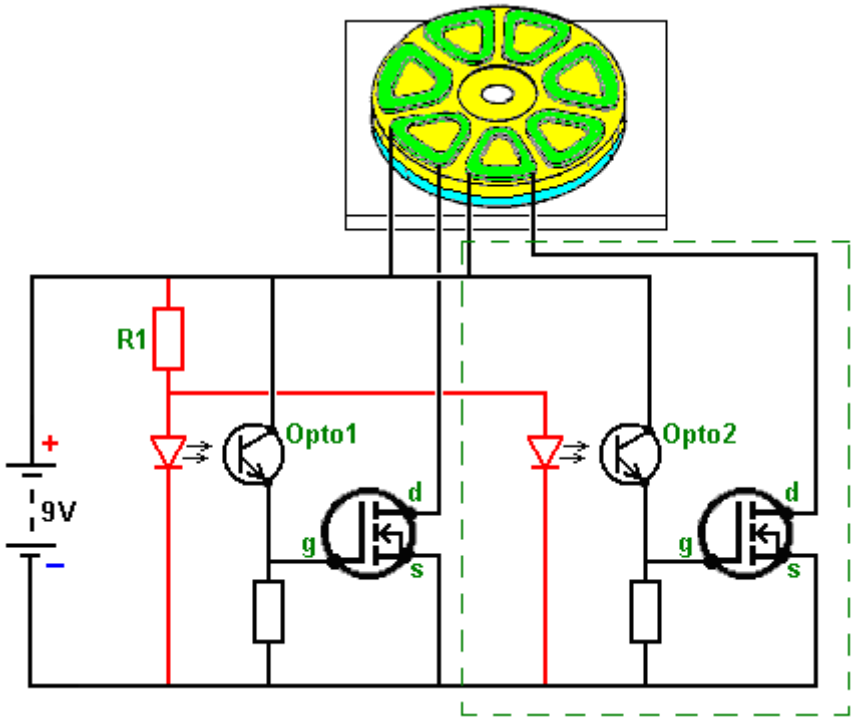
The circuitry that Charles specifies for powering the coils to block the magnetic fields of the permanent magnets uses N-channel MOSFETs and is very simple. Here is his circuit for driving one of the coils:



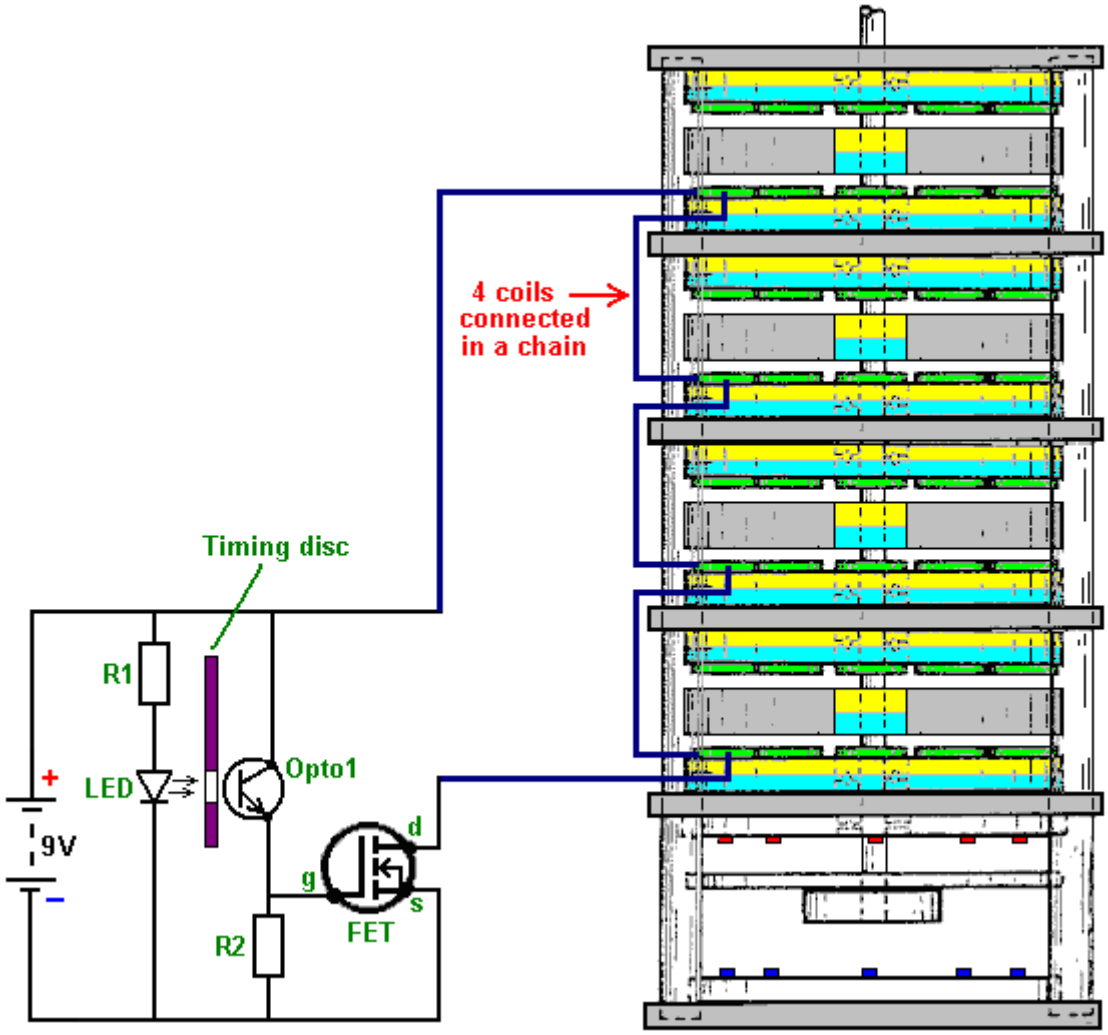
Just five components are used. The current through the coil is controlled by a transistor. In this case it is a Field-Effect Transistor usually called a "FET". The most common type of FET is used, namely an "N-channel" FET which is the rough equivalent to an NPN transistor as described in Chapter 12. A FET of this type is switched off when the voltage on it's "gate" (marked "g" in the diagram) is 2.5 volts or lower. It is switched on when the voltage on it's gate is 4.5 volts or more.

In this circuit we want the FET to switch on when the motor's timing disc is in the right position and be off at all other times. This is arranged by shining the light from a Light-Emitting Diode or "LED" through a hole in the timing disc which rotates with the shaft of the motor. When the hole is opposite the LED for the coil which is to be powered up, light shines through the hole and on to a light-sensitive device, Charles has opted to use a Light-Sensitive transistor, but a light-dependent resistor such as an ORP12 could be used instead. When the light shines on the "Opto1" device in the circuit diagram, it's resistance falls dramatically, raising the voltage on the gate of the FET and switching it on. When the timing disc hole moves past the LED, the light is cut off and the FET gate voltage drops down, switching the FET off. This arrangement causes the coil of the motor to be switched on and off at just the right time to give a powerful rotation of the motor shaft. In the circuit, the resistor "R1" is there to make sure that the current flowing through the LED is not excessive. The resistor "R2" has a low value compared to the resistance of "Opto1" when no light falls on it, and this holds the gate voltage of the FET down to a low value, making sure that the FET is completely off.

As you can see, this is basically a very simple circuit. However, as one of these circuits is used for each coil (or each pair of coils if there is an even number of coils in this slice of the motor), the circuit in the patent looks quite complicated. It is actually very simple. The resistor "R1" is used to limit the current flow through all of the LEDs used and not just one LED. You could, of course, use one resistor for each LED if you wanted to. The circuit for powering two coils (and not showing the timing disc) looks like this:



The section inside the green dashed line being the identical circuit for the second coil. This addition to the circuit is made for each coil, at which point, the motor is ready to run. If, as would be normal, several layers of magnets are being used, then the coils positioned above each other can be connected in a chain like this:

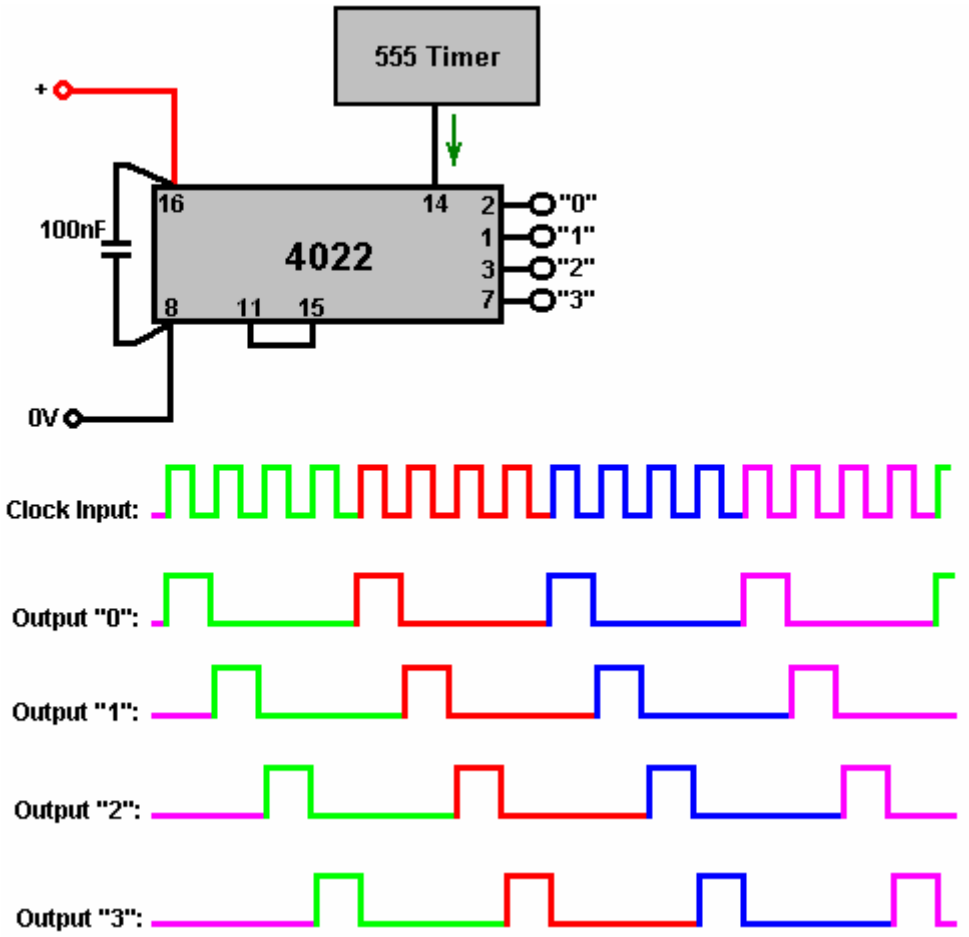


Connecting several coils "in series" (in a chain) like this, reduces the number of electronic components needed and it makes sure that the pulses to each of these coils is at exactly the same instant. Alternatively, it is possible to wire these coils across each other "in parallel", the choice is generally dictated by the resistance of the coils. The patent drawing shown above seems to indicate that there is a big gap between the LEDs and the optical devices. This is probably not the case as most people would choose to keep the gap between the LED and the light-dependent device as small as possible, mounting them so that they are just clear of the timing disc on each side of it.

In this patent, Charles Flynn remarks that this magnet motor can be used for almost any purpose where a motor or engine drive is required and where the amount of energy available or required to produce the driving force may vary little to nil. Charles has produced motors of this type which are capable of rotating at very high speed - 20,000 rpm and with substantial torque. Lesser speeds can also be produced, and the motor can be made to be self-starting. Because of the low power required to operate the device, Charles has been able to operate the motor using just a nine volt, off-the-shelf dry battery.

One application which seems most appropriate for this motor design is the Frenette heater shown in Chapter 14. Using this motor to drive the discs inside the heater drum would produce a heater which appears to be driven by just a nine-volt battery. However, while that is the appearance, the reality is that the power of this motor comes from the permanent magnets and **not** from the battery. The battery current is only used to prevent the backward pull of the magnets and it is **not** used to drive the motor.

While the use of a timing disc is a very satisfactory arrangement, it is also possible to use electronic circuitry instead of the mechanical timing disc, the opto devices and the LEDs. What is needed here is a device which produces a series of voltage pulses which can be used to drive the gate voltage of each FET from below 2.5 volts to over 4.5 volts. It looks as if the well-known 555 timer chip would be suited to this task and it would certainly run off the nine-volt battery. However, we have more than one set of coils which need to be run. For example, if we have say, four sets of coils to drive by powering up four different FET transistors one after the other, then we could use a "Divide-by-Eight" chip, like the 4022 chip. This chip can be set to divide by any number from two to eight. All that is needed to select the number to divide by, is one connection between two of the pins on the chip.



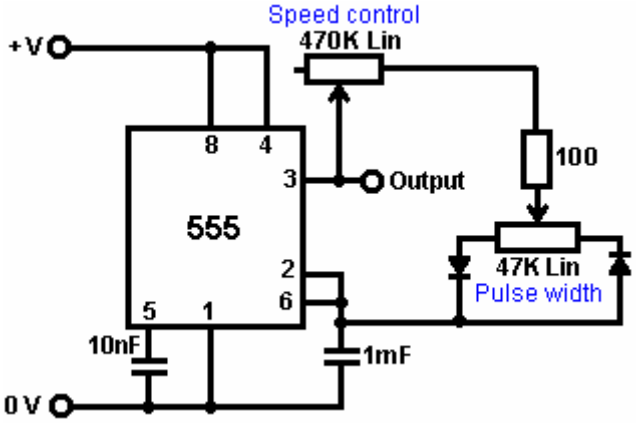
The output voltage on the pins marked "1", "2", "3" and "4" goes high one after the other as shown in the diagram above. So, each of these output pins would be connected to the FET gates in that order and the FETs would get switched on in that same order.

With the 4022 chip, the connections for the rate of division are as follows:

For ‘Divide by 7’ operation, connect pin 10 to pin 15
For ‘Divide by 6’ operation, connect pin 5 to pin 15
For ‘Divide by 5’ operation, connect pin 4 to pin 15
For ‘Divide by 4’ operation, connect pin 11 to pin 15
For ‘Divide by 3’ operation, connect pin 7 to pin 15
For ‘Divide by 2’ operation, connect pin 3 to pin 15

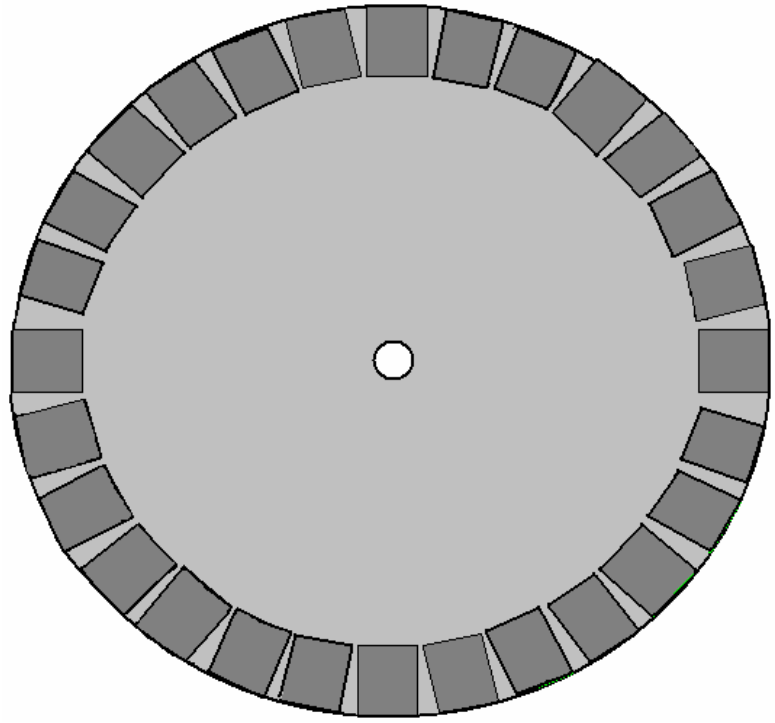
When using a circuit like this, the pulse rate from the 555 chip is set to a very low value like half a second, so that the motor shaft can get started. Once it gets moving, the pulse rate is gradually increased to speed the motor up. One advantage of this method is that it allows speed control, and if the motor was being used to power a Frenette heater, then the speed control would also act as a temperature control for the heater.

A possible 555 chip circuit might be:



As this allows the speed to be controlled and when the required speed is reached, the pulse width can then be adjusted to give the minimum current draw to maintain that speed. There are, of course, many other suitable circuits which could be used instead of this one and Chapter 12 will fill you in on some of them as well as explaining how circuits work and how to build them.

If it so happens that it is difficult to find suitable circular magnets with the poles on opposing faces, then I suggest that it should be possible to use standard rectangular magnets throughout and rectangular coils as shown here:



And while this arrangement is not as magnetically efficient as a circular magnet, it does have the convenience of allowing the construction of a rotor of any chosen size. Ideally, unlike the stator shown above, there should be an odd number of magnets, or failing that, an odd number of coils. Alternatively, the rotor could have an odd number of magnets so as to allow self-starting. But, it should be noted that if the motor is to be driven by an electronic pulsing system, then it is very much more simple to have an even number of magnets on the stator and start the motor moving by hand. This is because with an odd number of stator magnets, the opto sensors are not exactly opposite each other and so do not fire together. With an even number of stator magnets, the coils which are 180 degrees apart can be wired together as they fire at exactly the same time. With the slotted optical timing disc, the slots are exactly opposite each other and match the width of the rotor magnets, but the coils (nearly) opposite each other are not powered on and off at exactly the same time, although their powered arcs are likely to overlap for part of their operation. This could be catered for electronically by using a monostable delay for the coil on the opposite side of the disc.

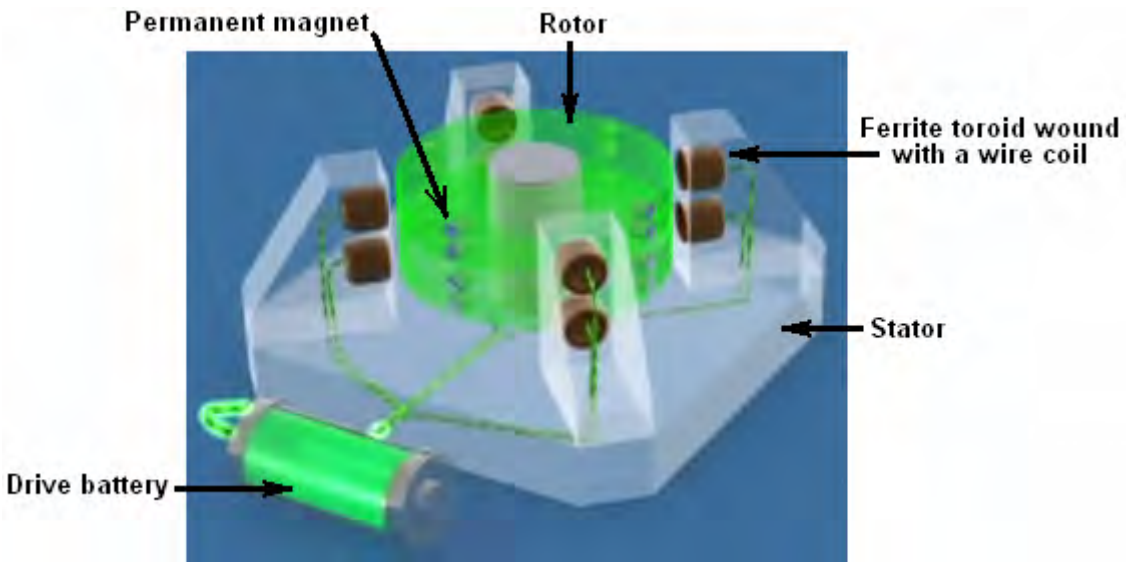
The objective of each coil is to just, and only just, cancel out the magnetic field of the permanent magnet underneath it. The magnetic field produced by the coil depends on the current flowing in the coil, the number of turns in the coil and the area of the coil. The current flowing depends on the diameter of the wire and the voltage applied to it. It is probably necessary to mount just one magnet on the stator and experiment with the coil until

your current drive and coil allow the rotor to spin freely. Whatever the coil result is, should be ok for all of the magnets even though they are likely to vary in strength a bit.

Steorn's Magnetic Devices.

The Irish company Steorn have produced a system which is almost identical to the Charles Flynn magnet motor just described. They call their device "Orbo" and its operation is pretty much the same. The advance made by Steorn is that they have devised a very clever magnetic masking system using ferrite toroids wound with a copper wire coil. This is a slick method of switching magnetic attraction on and off. When the coil carries a sufficient current it generates a circular magnetic field spiralling around the toroid and not going outside the toroid. This field does not have an attraction for outside magnets. It makes no difference if the direction of the current flow through the coil is reversed as the resulting magnetic field just spins around the toroid in the opposite direction and performs exactly the same magnetic blocking of the ferrite ring which forms the toroid. If no current flows, then the copper wire does not block off the influence of the ferrite ring and the permanent magnets on the rotor are strongly attracted to it, causing the rotor to spin.

On their web site www.steorn.com, Steorn illustrate their design like this:



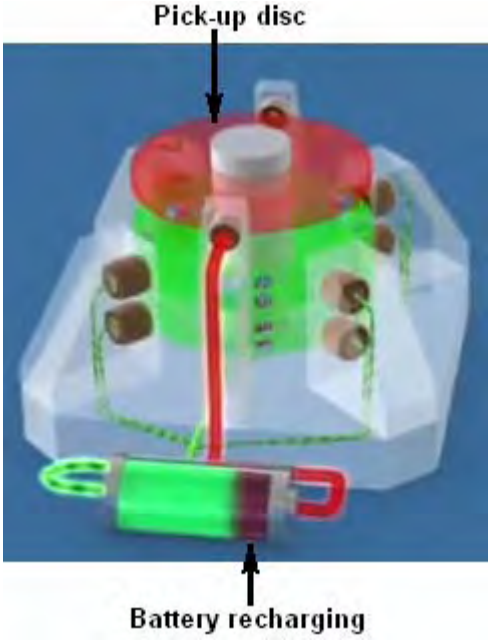
In this implementation, eight ferrite rings are mounted on the stator in four locations ninety degrees apart. These are wound with copper wire coils which can be powered by a battery, via a timing mechanism. The rotor has embedded in it, eight pairs of small permanent magnets, also spaced ninety degrees apart.

In exactly the same way as the Adams motor described in chapter 2, the current through the coils is set to the minimum level which allows the rotor to spin freely. The timing mechanism is then switched in and the motor and the rotor given a spin. The rotor magnets are strongly attracted to their corresponding ferrite rings mounted on the stator posts and this accelerates the rotor.

If no current is passed through the coils, then the rotor will oscillate backwards and forwards for a short time before coming to rest with the magnets as close to the ferrite rings as possible. To prevent this happening, the timing circuit senses when the magnets reach the ferrite rings, and passes that minimum current through the coils, trapping the rings inside a magnetic field which has no effect on the rotor magnets. The momentum of the rotor causes it to spin on past the stator rings to a position where the magnets are closer to the next rings than they are to the ones which they have just passed, at which point, the current is cut off and the magnetic attraction to the ferrite rings returns. This is identical to one mode of operation of the Adams motor.

The next step is also identical to that of the Adams motor, namely, to add on some pick-up coils to convert some of the rotating magnetic energy into electrical energy, either to recharge the driving battery or to power other equipment, or both.

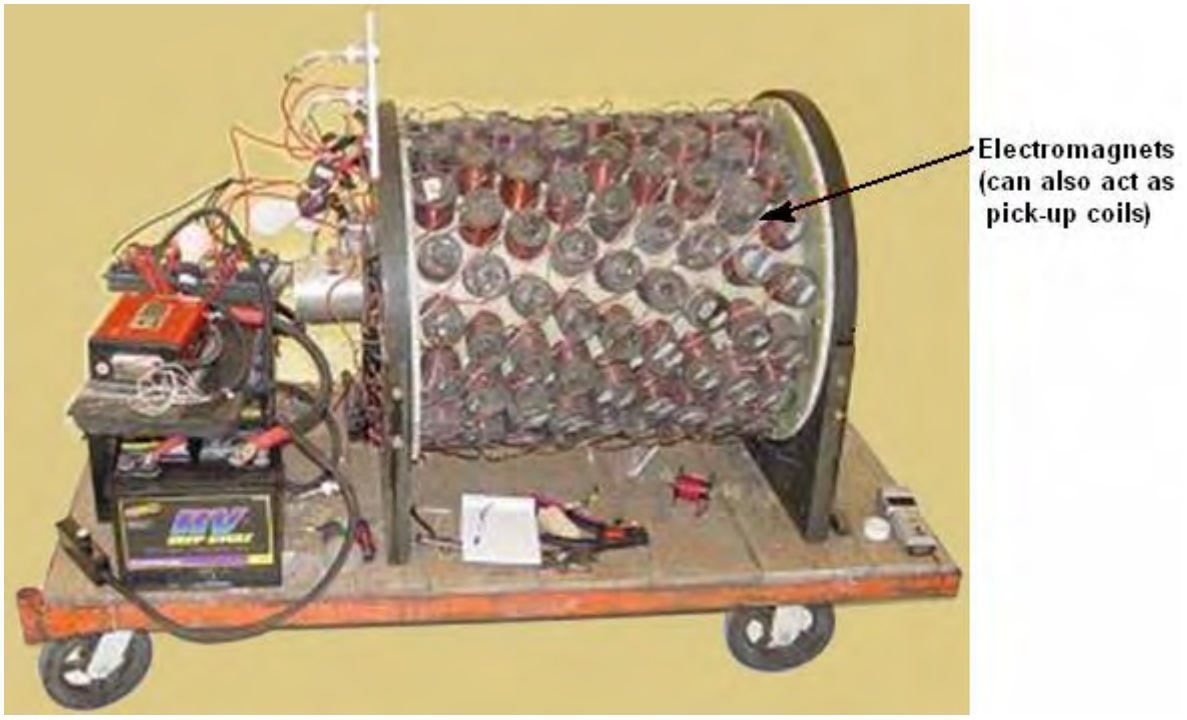
Steorn's arrangement for doing this is to add an additional disc, containing permanent magnets, to the rotor and positioning wire coils opposite those magnets as is normal for a generator. Steorn choose to show the resulting energy charging up the battery again:



Video presentations on this style of motor/generator are at:
http://www.youtube.com/watch?v=AXamGLyRkt8&NR=1
http://www.youtube.com/watch?v=rg3rLqYMzN4&feature=related and
http://jnaudin.free.fr/steorn/indexen.htm

We tend to think of this style of magnet-powered motor as being low-power. This is probably because it is often the case that the demonstration proof-of-principle implementations shown are minor devices. These motors can be very powerful and the one shown here, designed and built by Mr Sung of China has an output power of 20 kilowatts or twenty-seven horsepower:

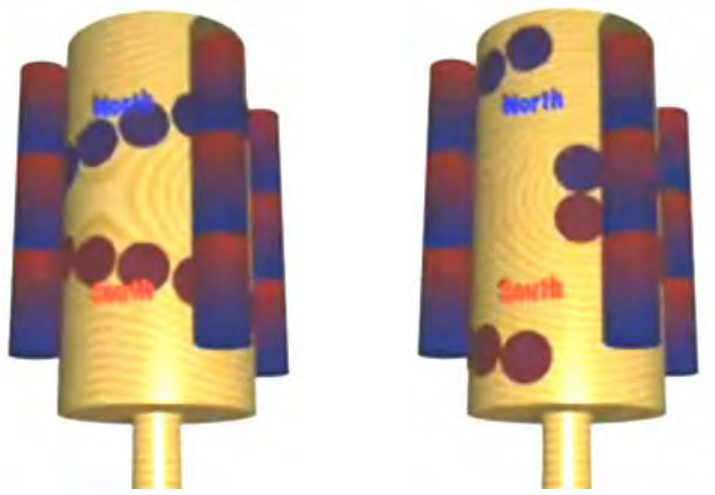
And another design which has a larger diameter and about 144 magnets has a reported output of 225 horsepower:



You will notice that each ring of magnets is positioned further around the rim of the cylinder providing powerful pulses from 64 magnets every 22.5 degrees of rotation, so it is little wonder that the motor has considerable shaft power. Some of the coils can be switched to collect power if the working conditions do not need the full shaft output power, charging the drive battery. The rotating inner cylinder has permanent magnets mounted on it.

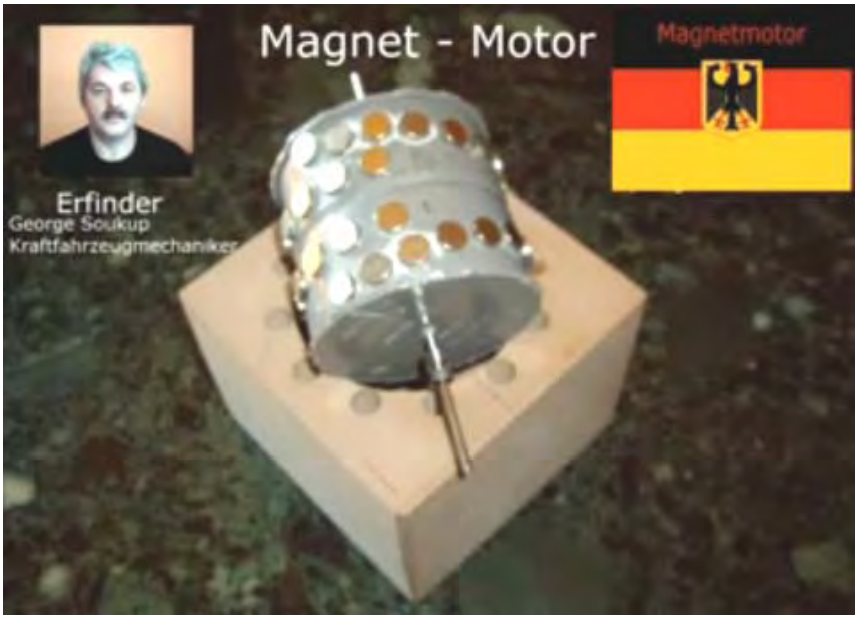
George Soukup's Permanent Magnet Motor.

There used to be a video on the web, showing a magnet motor built on the "V" style of magnet placement which has two sets of permanent magnets spaced like this:



This style of magnet arrangement (North magnets shown in blue and South in red) has a locking point where the switch from wide spacing to narrow spacing occurs and this causes the rotation to stop there.

The implementation shown in this video has the V magnets spaced rather more widely apart as shown here:



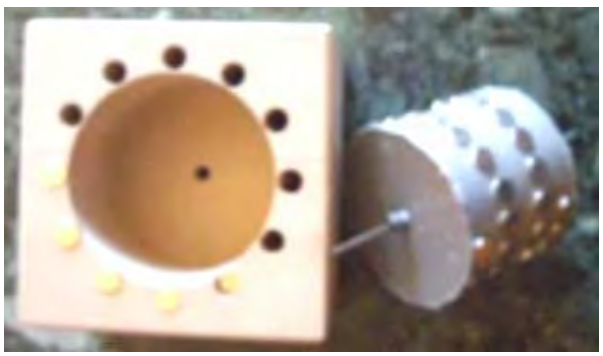
The taper is much less pronounced with an inner gap some four times greater than the gap to the outer ring. It also appears that the last inner magnet has a greater gap around the drum than the remaining ring of magnets.

The housing is very simple looking, with an evenly spaced ring of twelve holes to take long magnets with alternating North and South magnetised areas along their length. You will notice from the photographs, that George has cavities to take up to twelve stacks of stator magnets, although he only uses any five of them for his demonstrations.

The housing has considerable clearance for the drum and magnets. The rear shaft bearing is just set into the back of the housing:

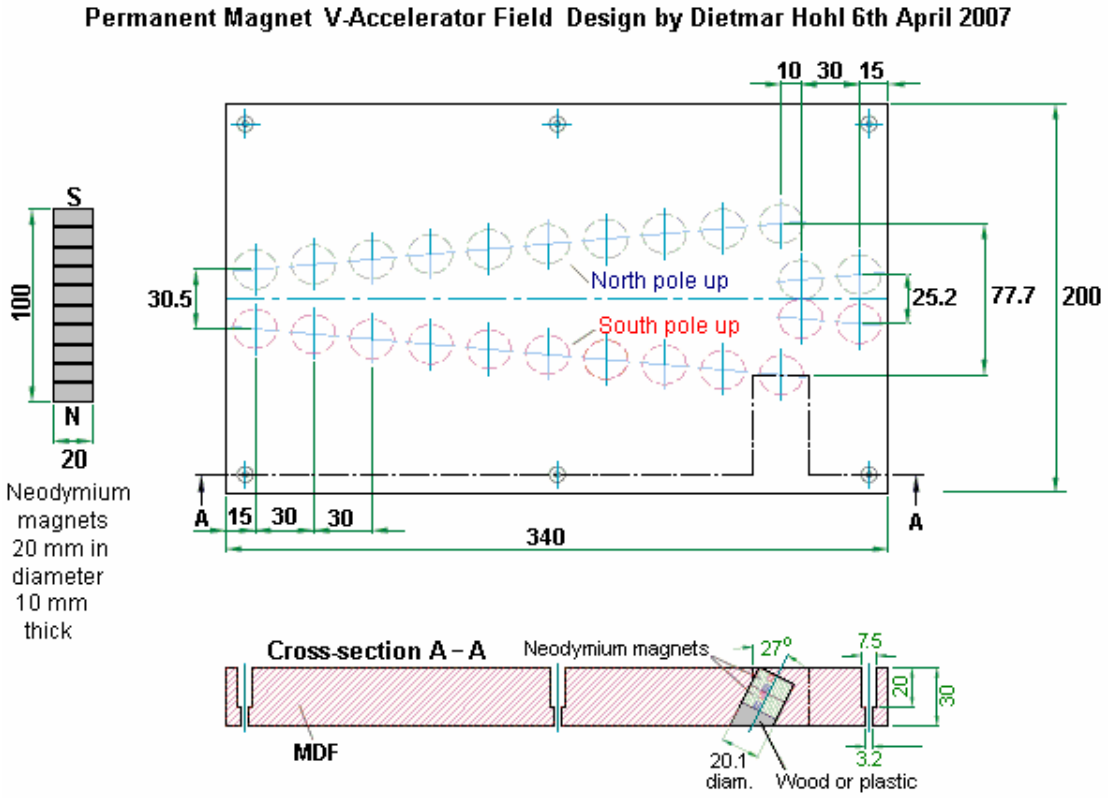
The front has two sheets of acrylic, one to hold the insert magnets in place and one to provide the shaft's front bearing support:

As there is no commentary with the video it is a little difficult to pick up all of the details, but it seems that positioning stator magnets allows the motor to overcome the normal sticking point of the typical V-motor arrangement. The video shows various arrangements including the non-symmetrical grouping shown here where four or five consecutive magnets are used and the remaining slots left empty:



Dietmar Hohl's Permanent Magnet Motor

If you would like to make a simple motor of this type, then the information provided by Dietmar Hohl, passed to me by Jes Ascanius of Denmark, shows you how. He uses 20 mm diameter round neodymium magnets 10 mm thick, stacked in pairs in the stator of this layout:

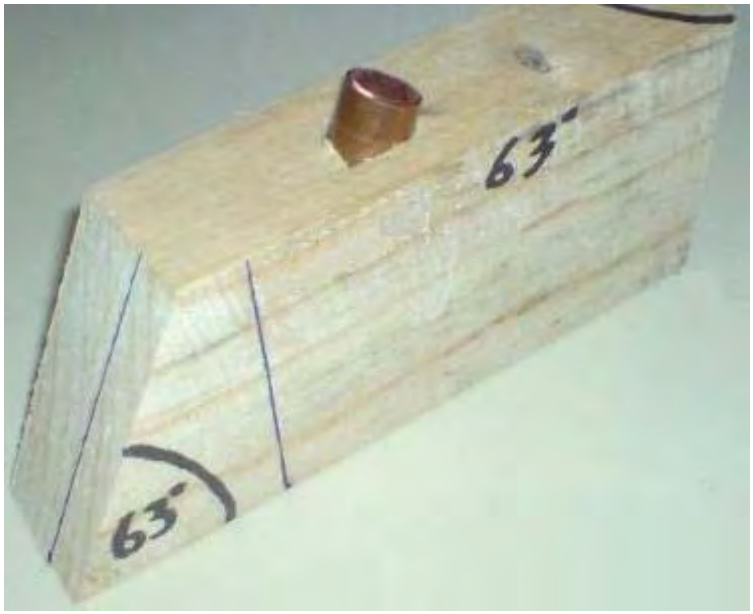


This shows a magnetic gate arrangement built on a flat piece of Medium-Density Fibreboard 30 mm thick. The holes drilled in it are 20.1 mm in diameter and positioned so as to take two of the 10 mm thick magnets stacked together. The holes are drilled at an angle of 63 degrees to the horizontal or 27 degrees to the vertical, whichever way you prefer to think of it. On one side of the board, the inserted magnets have their North poles facing upwards, while on the other side of the board, the magnets are inserted with their South poles facing upwards. Dietmar shows six holes to take bolts or screws to fasten the piece of MDF to a larger board or table. Those do not form any part of the magnetic system and can be omitted. A video of one version of it in action can be found at http://www.free-energy-info.tuks.nl//Vtrack.mpg.

The gate operates by causing a stack of ten of the magnets to roll along the V-shaped track and pass smoothly across the junction with the next set of V-positioned magnets. There can be as many of these V-sets as you want and the magnet stack will still keep rolling. This is one of the few magnetic gate designs which adapts to drum operation as a motor rotor.

The magnets are positioned at an angle in order to use the magnetic fields at the edge of the magnets. They are stacked in pairs in order to increase their power. The power of the motor depends on the strength of the magnets, how close the stator magnet stacks are to the VF-track magnets and the number of stacks of stator magnets. If you decide to construct one of these motors, then it is suggested that you make things easier for yourself by keeping the curvature low, using three or four of the Vs. With Dietmar's dimensions, a 2-V drum would be 216.5 mm (8.5") in diameter, a 3-V drum would have a 325 mm (12.8") diameter and a 4-V drum a diameter of 433 mm (17") and those dimensions include the 30 mm (1 3/16") strip which holds the magnets, so the inner drum diameters are 30 mm less in each case.

When making the motor drum, it is possible to use a flexible material to hold the magnets. This allows the strip to be laid out flat while the holes are drilled, and then attached to the outside of a rigid drum with a 60 mm lesser diameter than the ones mentioned above. Jes Acanius of Denmark shows how a jig can be made to make drilling the holes easier:



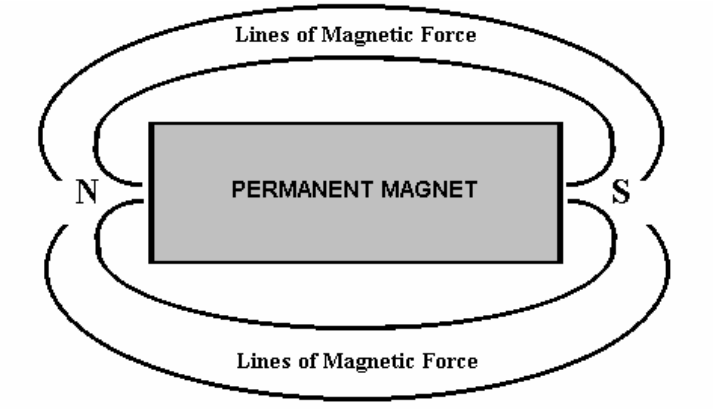
This one has had a length of copper pipe inserted at the correct angle, in order to direct the drill bit at the exact angle required. This motor has been successfully replicated by Jes Ascanius of Denmark who used 10 mm magnets which were to hand, and again with square magnets which were pushed into round holes and not even angled in this proof-of-concept implementation which only took one hour to build using scrap material to hand, and which did work:

With Dietmar's design using angles magnet pairs, the number of magnets needed is quite high. For a single V, there are 58 magnets. For a 2-V version, 106 magnets. For a 3-V version, 154 magnets and for a 4-V version, 202 magnets if there is only one stack of stator magnets, so ten extra magnets need to be added to the count for each additional ten-magnet stack of stator magnets. The motor power is likely to increase as the diameter increases as the lever arm that the magnet has to turn the drum, increases – double the diameter to (almost) double the power.

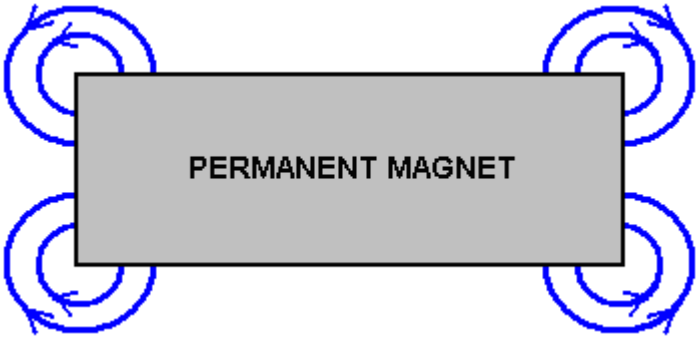
Simple Permanent Magnet Motors

It is very difficult to use the power of permanent magnets to make a motor powered by them alone. The Dietmar Hohl design shown above is one of the very few which can readily be made and tested at home. The problem is that almost all magnets have a symmetrical magnetic field, while what is needed for a magnet-powered motor is an asymmetrical magnetic field. Consequently, magnets have to be combined in ways which distort their normal field shape. You will notice that in the Hohl motor, the drive magnets are angled and that is an important feature of using magnets in motors.

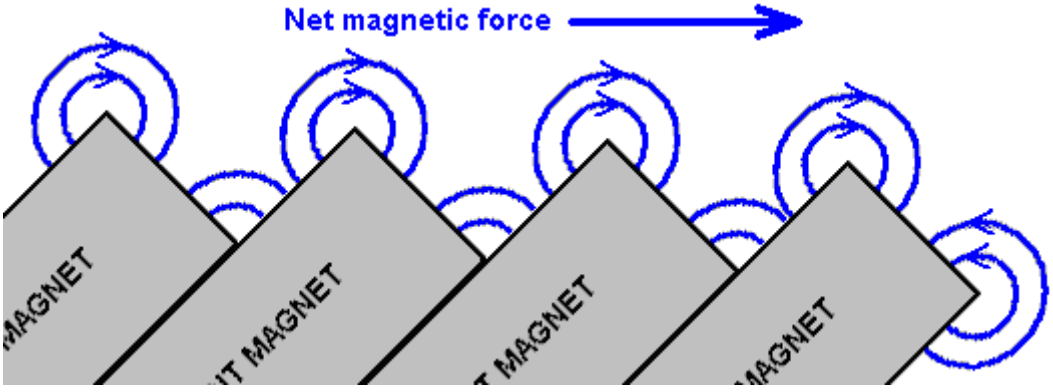
Schools currently teach that the field surrounding a bar magnet is like this:



This is deduced by scattering iron filings on a sheet of paper held near the magnet. Unfortunately, that is not a correct deduction as the iron filings distort the magnetic field by their presence, each becoming a miniature magnet in it's own right. More careful measurement shows that the field actually produced by a bar magnet is like this:

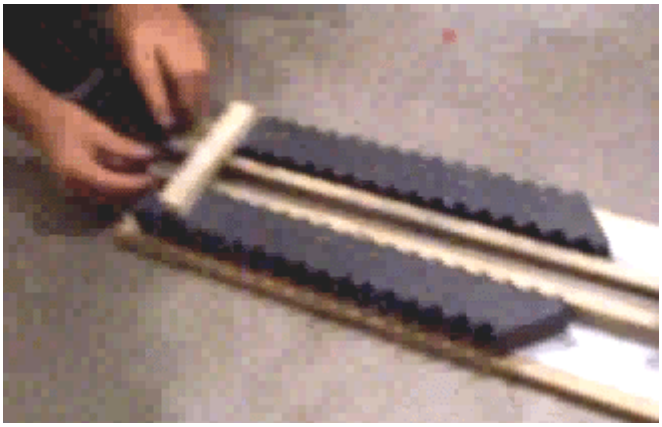


There are many lines of force, although these diagrams show only two of them. The important factor is that there is a rotating field at each corner of a typical bar magnet. It follows then that if a row of magnets is placed at a an angle, then there will be a resulting net field in a single direction. For example, if the magnets are rotated forty five degrees clockwise, then the result would be like this:



With this arrangement, the opposing corners of the magnets as shown here, are lower down and so there should be a net magnetic force pushing to the right just above the set of magnets. However, the situation is not as simple and straightforward as you might imagine. The additional lines of magnetic force which have not been shown in the diagram above, act further out from the magnets and they interact, creating a complex composite magnetic field. It is frequently found that after four or five magnets that a short gap needs to be left before the line of magnets is continued on.

Two boys; Anthony and Andreas, have used this magnet arrangement to create a magnetic track and they have a lot of fun, sending a magnet sliding between two of these rows of angled magnets. Initially, they used the cheaper ceramic magnets and got a very satisfactory movement when using a neodymium magnet as the moving component:



You will notice that they have managed a row of 18 ceramic magnets on each side of their track and the results which they are getting are very good. They have three videos on the web at the present time:

https://www.youtube.com/watch?v=Vo2-Qb3fUYs

https://www.youtube.com/watch?v=VeXrFfw4RSU

https://www.youtube.com/watch?v=VTbFfEEE_qU

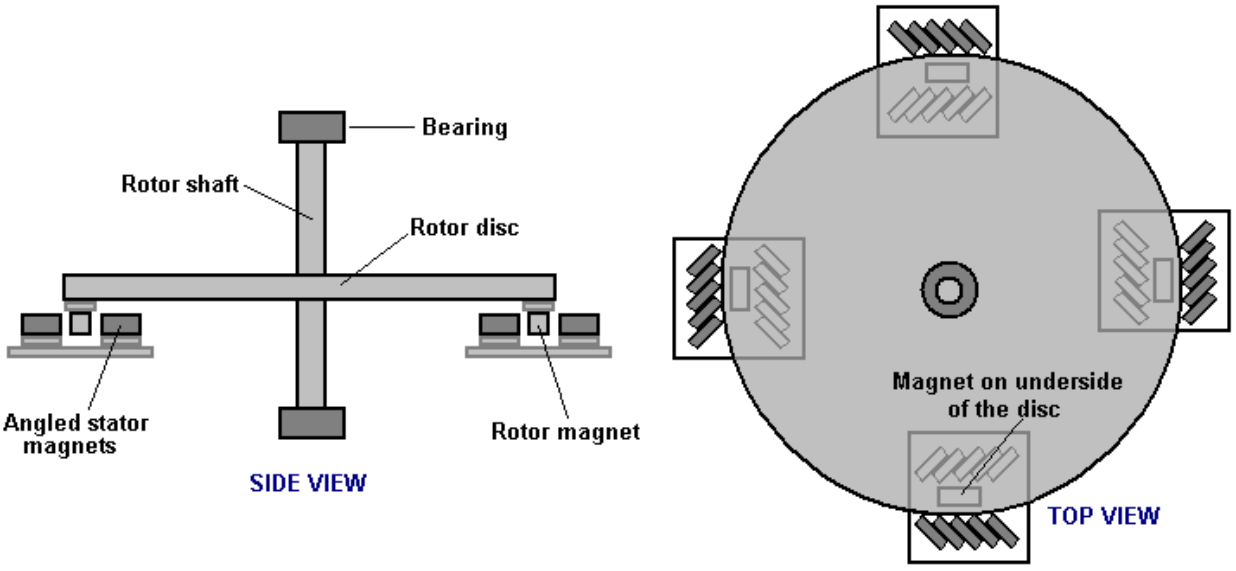
The moving magnet is made up of four 12 mm x 12 mm x 12 mm (or half-inch by half inch by half inch) neodymium magnets attached North - South - North - South - North - South - North - South:

They have not disclosed all of the details of what they are using (accidentally rather than by intention). The ceramic stator magnets are 48 mm x 20 mm x 10 mm with the poles on each of the main faces. They position each magnet with it's North pole facing towards the track and they angle the magnets at 45 degrees. There is a 15 mm gap between the stator magnets and the moving magnets on both sides of the track. Wooden strips direct the moving magnets.

Neodymium magnets have very different characteristics to those of ceramic magnets (and that is not just strength of the magnetic field). It is not unusual for experimenters to find that devices will work well with one type of magnet but not with the other type. Here the developers have also tried using two sets of five angled neodymium magnets on each side of their track and the result was a more powerful thrust on their moving magnets.

The magnets are held in place in this picture, by wooden dowels driven into the base plank. They used these in order to avoid any magnet-fastening material which could alter the magnetic field.

The next step would be for them to power a motor using their magnetic track technique. However, this has been tried many times and the conclusion is that it is **VERY** hard to change a straight magnetic track into one which forms a complete circle. Therefore, I would suggest the following arrangement:

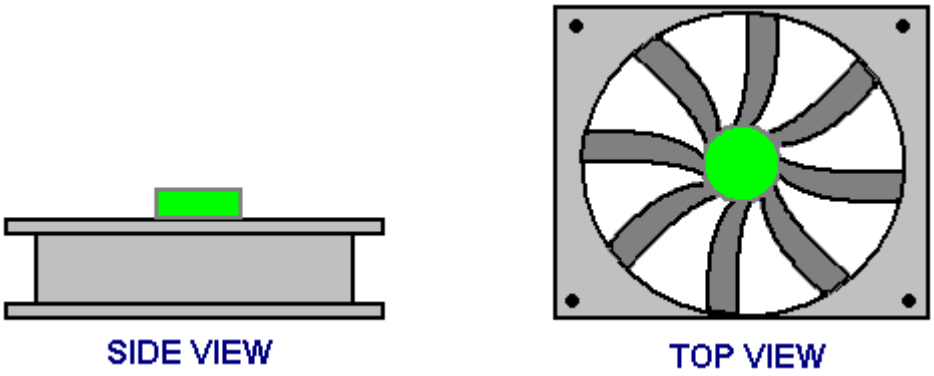


Here, a simple disc rotor has four magnets (of the type used to move down the magnetic track) attached to the underside of the disc and positioned so that they move through four short sets of four, or at the outside, five angled stator magnets as the disc spins. It does not matter if the rotor shaft is horizontal or vertical. If the disc spins well, then sets of two air-core pick-up coils can be positioned between each of the stator magnet arrays so that electricity is generated as the rotor magnets pass by overhead. If a constructor decides to attach two rotor discs to the one rotor shaft, then the two rotors should be positioned so that the rotor shaft gets pushed every 45 degrees of rotation rather than every 90 degrees as shown here. This style of motor is definitely within the scope of the average person to build should they be inclined to do so.

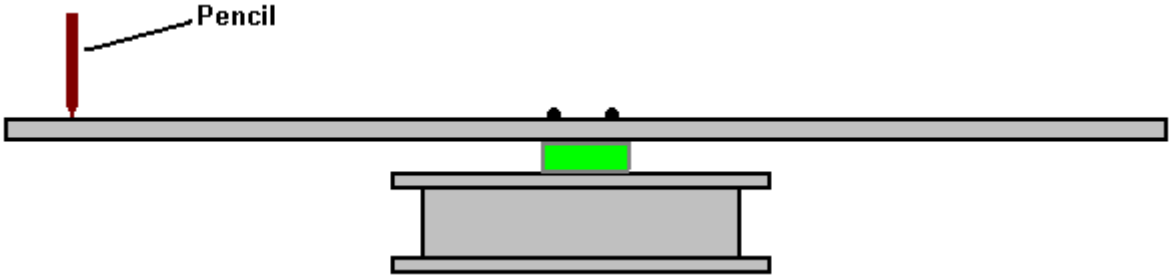
I have been asked to say how I personally would go about constructing a prototype of this nature. As I have very limited constructional skills, I would do it like this:

For the bearing, I would pick a computer cooling fan, as these have very good bearings and if one is not to hand inside an old, obsolete computer, then they can be bought very, very cheaply. The diameter of the fan is not important. These fans generally look something like this:

As the part of the fan which spins round does not normally project above the stationary frame, a spacing disc of wood or plastic is needed to provide the clearance. The disc is glued to the centre of the fan using perhaps, Impact Evostick, epoxy resin or super glue. It would then look like this:



A square of wood can then be screwed to the spacer, like this:



And as I am hopeless at creating good-quality mechanical devices, I would then hold a pencil very steadily against a support and give the wood a spin, so that the pencil draws a perfect circle exactly centred on the bearing of the fan. Then, marking the wood and the spacer so that there is no doubt as to which way round the wood is attached to the spacer, I would unscrew the wood and cut around the pencil line very carefully, smoothing the edges of the disc gently with fine sandpaper. Screwing the disc back in place, a spin should confirm that the edge of the disc stays steadily in place with no wavering of the edge. Actually, if the disc is not perfect, that is not a major problem as it is the rotor magnets which need to be positioned accurately, and for that, another pencil line can be produced by spinning the disc when the desired position has been determined.

Permanent magnets vary enormously in size and strength, so when magnets are purchased, it is a matter of testing them using a track of the type used by Anthony and Andreas. The stator magnets are angled at about 45 degrees to the track and with just four on each side, it is a case of finding the spacing between the two sets of angled magnets which pushes the stator magnets furthest along the track.

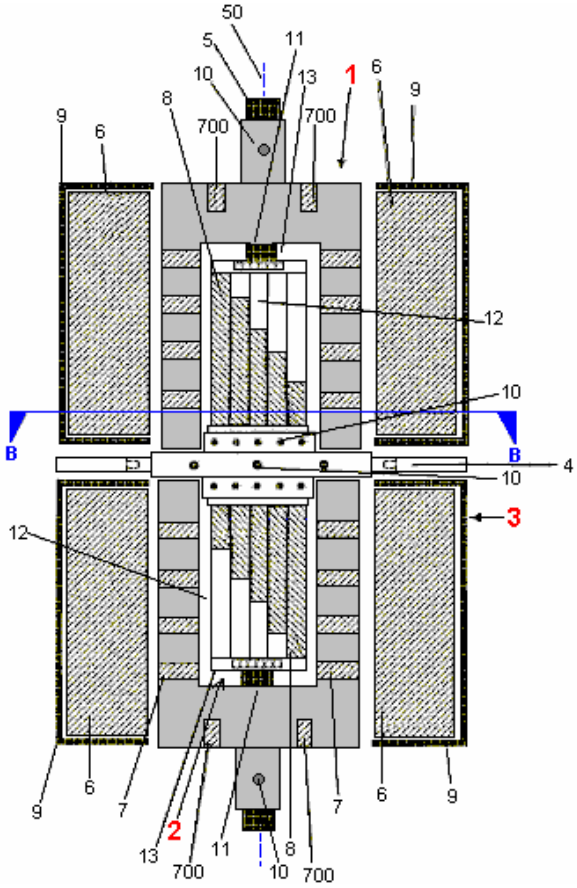
Muammer Yildiz's Permanent Magnet Motor.

Muammer Yildiz has developed a powerful permanent magnet motor, patented it, and demonstrated it to the staff and students of a Dutch university. During the demonstration, the mechanical power output was estimated at 250 watts and immediately after the demonstration, the motor was completely taken apart to show that there were no hidden power sources. There is a video showing this demonstration, located at:
http://pesn.com/2010/04/22/9501639_Yildiz_demonstrates_magnet_motor_at_Delft_University/

Please note that this is an attempted translation of the German language text of his patent and so, the accuracy of the content is not absolutely certain although it is likely to be reasonably accurate.

Patent EP 2,153,515 **17th February 2010** **Inventor: Muammer Yildiz**

DEVICE HAVING AN ARRANGEMENT OF MAGNETS



ABSTRACT

The device has a rotating axial drive shaft **5** supported so that it rotates inside a stator **2**, which is surrounded by an outer stator **3**. The rotor is firmly connected to the drive shaft. The outer stator has dipole magnets **6** which are positioned on the inner surface of a circular cylinder **9**. These outer magnets are evenly spaced around the surface of the surrounding cylinder.

DESCRIPTION

This invention is a device for generating an alternating magnetic field that interacts with a stationary magnetic field. The interaction of a stationary magnetic field with an alternating magnetic field has been used for some time, for example in brushless DC motors and in magnetic levitation.

One object of this invention is to provide an improved device for generating an alternating magnetic field that interacts with a stationary magnetic field. This is achieved as described in Claim 1, by the special arrangement of the dipole magnets of the inner stator, the rotor and the outer stator which creates a magnetic effect which keeps the rotor floating freely between the inner stator and the outer stator, and this acts as a magnetic bearing.

Surprisingly, it has been shown that the special layout of the dipole magnets of the inner stator, the rotor and the outer stator during rotation of the rotor, generates an alternating magnetic field is which allows a largely loss-free movement of the rotor as it spins between the inner stator and the outer stator. This very useful effect can be used for a variety of technical applications, for example, a particularly low-friction bearing is preferred for supporting a shaft which has to rotate at high speed.

In the following description, when mathematical terms, especially geometric terms, are used - terms such as “parallel”, “perpendicular”, “plane”, “cylinder”, “angle”, etc. as is typical when producing technical drawings, but it must be understood that these things are never achieved in practice, due to the manufacturing tolerances of the components. It is therefore important to realise that this description refers to the ideal situation, which will never be achieved. Therefore, the reader needs to understand that generally accepted tolerances will be involved in practice.

The output shaft spins around one axis, called the “shaft axis”. The shaft itself is preferably constructed as a straight cylinder of circular cross-section.

In a preferred embodiment of this invention, the magnets project slightly out of the inner stator. This is also the case for both the rotor and the outer stator. A partial overlap of two magnets is achieved when a plane perpendicular to the shaft axis, passes through both of the two magnets and the two magnets are considered to overlap if this situation occurs.

A partial overlap of three magnets occurs when a plane perpendicular to the shaft axis runs through each of the three magnets. The degree of overlapping does not affect the description and the amount of overlap of any two of the three magnets can be anything from 1% to 100%, where the magnets overlap completely.

In a particularly preferred embodiment of the invention, the magnets of the inner stator and the rotor are able to align completely. In addition to this, the outer stator is constructed so that it can be rotated around the shaft axis so that the contact ratio between the magnets of the rotor and the magnets of the outer stator can be adjusted to give any degree of overlap from 0% to 100%.

Three imaginary cylinders are produced. One by the magnets of the inner stator, a second by the rotor magnets as they spin around the shaft axis and the third is created by the magnets of the outer stator. The axes of these three cylinders is the same as the shaft axis.

Ideally, the rotor will have the shape of a drum or a cup, that is, a hollow cylinder with a circular cross-section or a piece of pipe whose one end face is covered by circular disk. In the centre of the disc, the rotor has a hole through which the shaft passes. The disc can also have a collar which is used to clamp the rotor to the shaft by means of a bolt passing through the drive shaft or by grub screws tapped into the collar. Whichever method is used, the rotor magnet assembly is connected securely to the drive shaft. The use of a clamping screw has the advantage of allowing the rotor to be taken apart for maintenance or repair. The hollow cylinder section of the rotor, is arranged so that there is a small air gap between it and both the inner and outer stators.

The hollow rotor cylinder has two, or more, permanent magnets mounted on it. These are equally spaced around the circumference of the rotor cylinder and positioned so as to be parallel to the drive shaft axis. The outer stator is cylindrical in shape and surrounds the rotor, leaving a small air gap between them and it’s axis is aligned with the drive shaft axis. Ideally, the magnets mounted on the inside of the outer stator cylinder, are aligned with the drive shaft axis and their pole faces are at right angles to the shaft axis. That is, a line drawn through the North and South pole faces of these magnets will point at the drive shaft, and so one pole face will face the rotor.

It is also possible for the magnets of the outer stator to be rod-shaped and to form a complete ring around the inner face of the outer stator cylinder. If this is done, then the magnetic rings need to be separated from each other by non-magnetic spacers and the whole length of the outer stator will be covered with these magnetic rings and spacers. In this case, the inner and outer stators are mounted in a fixed relationship to each other by means of brackets or other mounting methods.

Ideally, the rotor is held in position by the magnetic fields of the two stators and “floats free” between them. This is the preferred method. However, it is possible for the drive shaft to run the entire length of the device and to be supported in roller bearings.

One possible construction is to have both of the stators made in two separate parts. These need to be exactly symmetrical relative to the drive shaft axis. The outer stator pieces can also be arranged to be capable of rotational adjustment relative to the inner stator which always has a fixed position. Another option with this particular arrangement is to have the distance of the outer stator components adjustable, so that the air gap between the rotor and the outer stator magnets can be manually adjusted.

An angle “alpha” is defined as the angle between the magnetic axis of a magnet of the inner stator and a tangent to the circumference of the inner stator at that point. An angle “beta” is defined as the angle between the magnetic axis of a rotor magnet and a tangent to the rotor circumference at that point. An angle “gamma” is defined as the angle between the magnetic axis of a magnet of the outer stator and a tangent to the circumference of the outer stator at that point. In a preferred embodiment of this invention, each of these angles

is between 14 degrees and 90 degrees.

It is a particular advantage if the permanent magnets of both the inner and outer stator have a either a rectangular or trapezoidal cross-section when seen as being cut by a plane perpendicular to the shaft axis. It is also particularly advantageous if the rotor magnets have a circular cross-section when viewed as being cut by that plane perpendicular to the shaft axis. Other, non-symmetrical magnet cross-sections are possible, such as trapezoidal, triangular, or irregularly shaped cross sections.

It is possible for all of the magnets of the inner stator to have identical shapes. Similarly, it is possible for all of the magnets of the outer stator to have identical shapes. It is also possible for all of the rotor magnets to have the same shape. However, the positioning of the magnetic North and South poles of the various magnets will not be identically position as will be seen from the following detailed description.

The magnets of the inner stator, the rotor and the outer stator have a magnetic orientation which causes them to repel each other at every angular position of the rotor. For example, the magnets of the inner stator can have their North poles facing outwards and in that case, the magnets on the rotor will have their North poles facing inwards towards the inner stator. Similarly, the magnets of the outer stator would then have their South poles facing inwards in order to repel the (outer) South poles of the rotor magnets.

Further features, details and advantages of the invention will be apparent from the following description of an embodiment of the invention and the associated drawings as shown here:

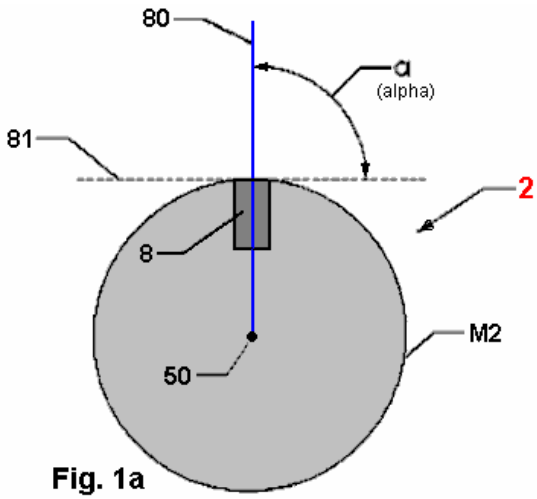


Fig. 1a

Fig.1 is a schematic representation of the device.

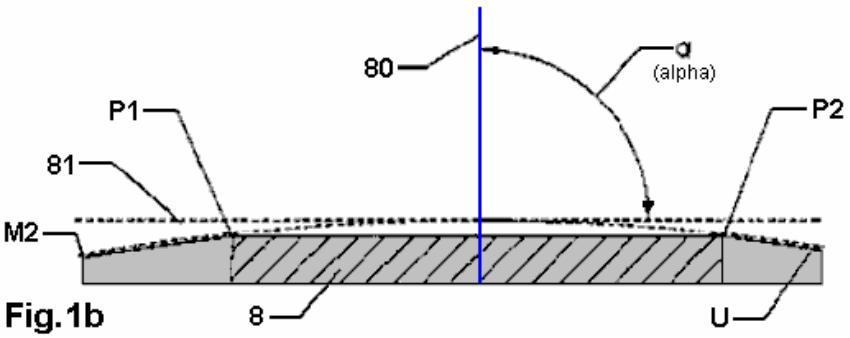


Fig.1b

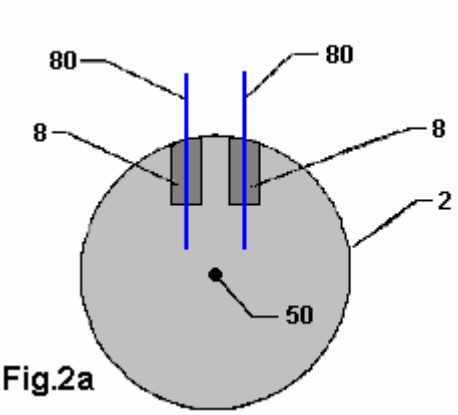


Fig.2a

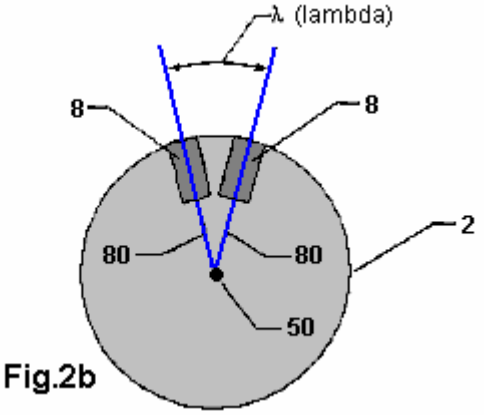


Fig.2b

Fig.2a is an oblique view of the inner stator without magnets and **Fig.2b** is a view of the inner stator at right angles to the shaft axis.

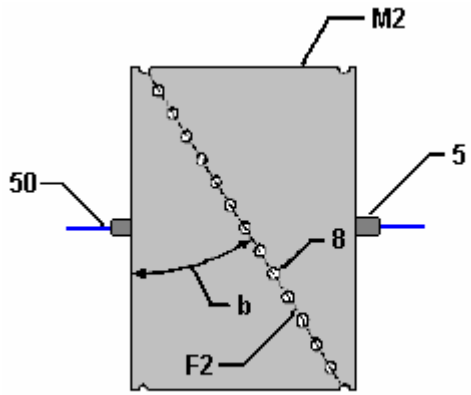


Fig.3a

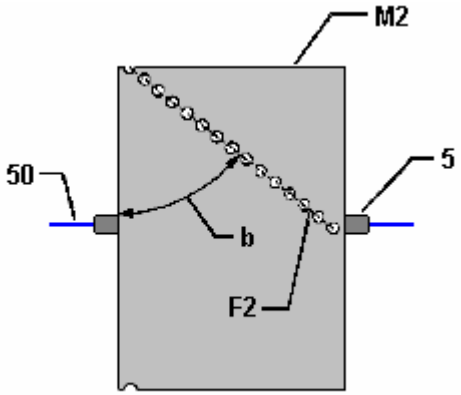


Fig.3b

Fig.3 Shows a magnet arrangement for the inner stator

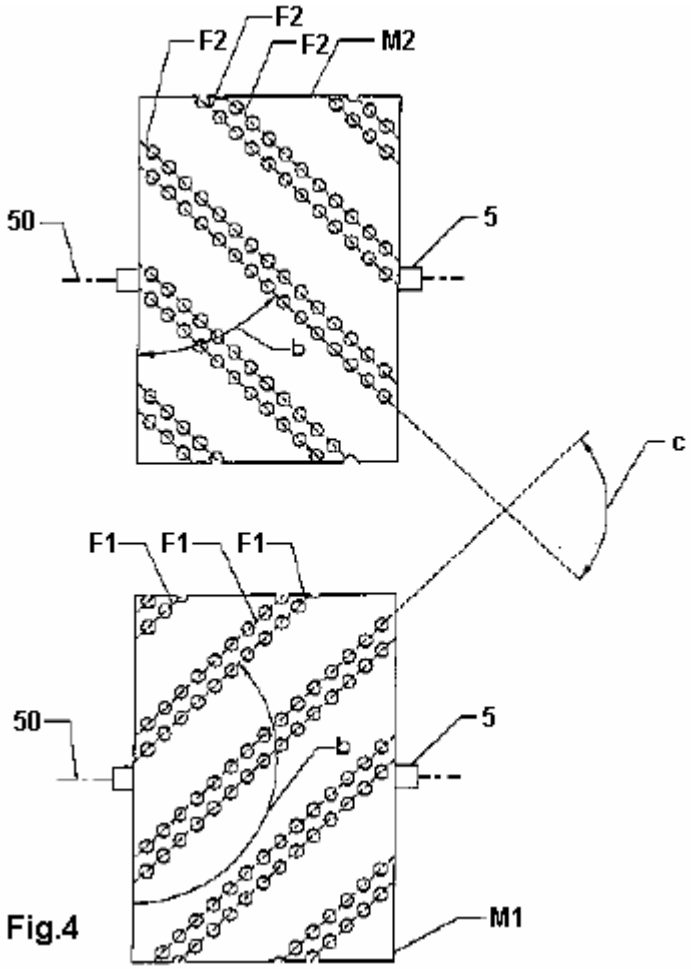


Fig.4

Fig.4 a section through the inner stator, along the line **A--A** indicated in **Fig.12b**

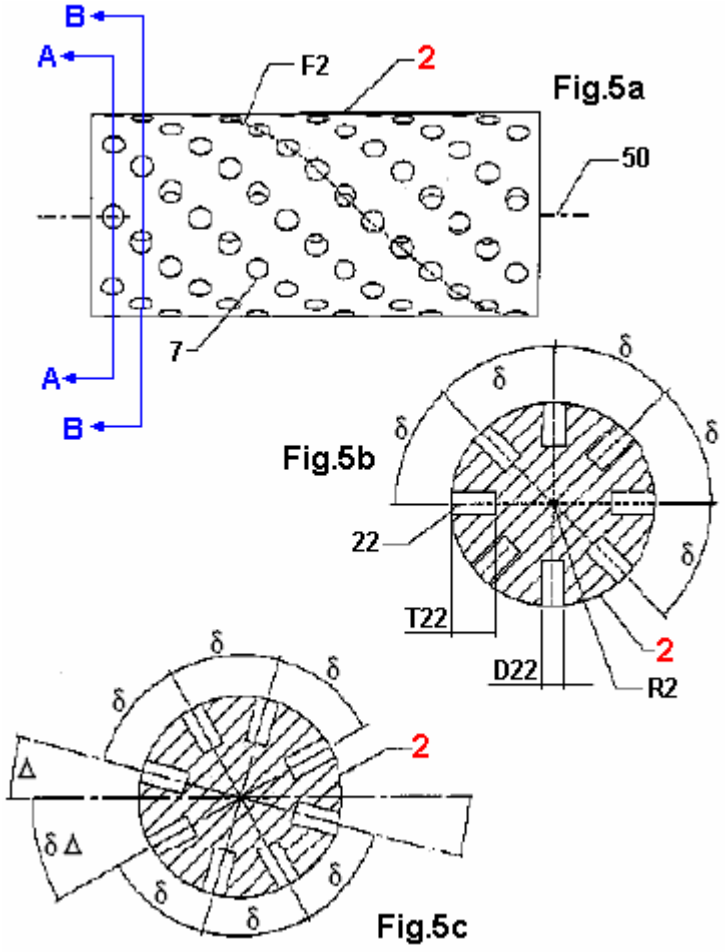
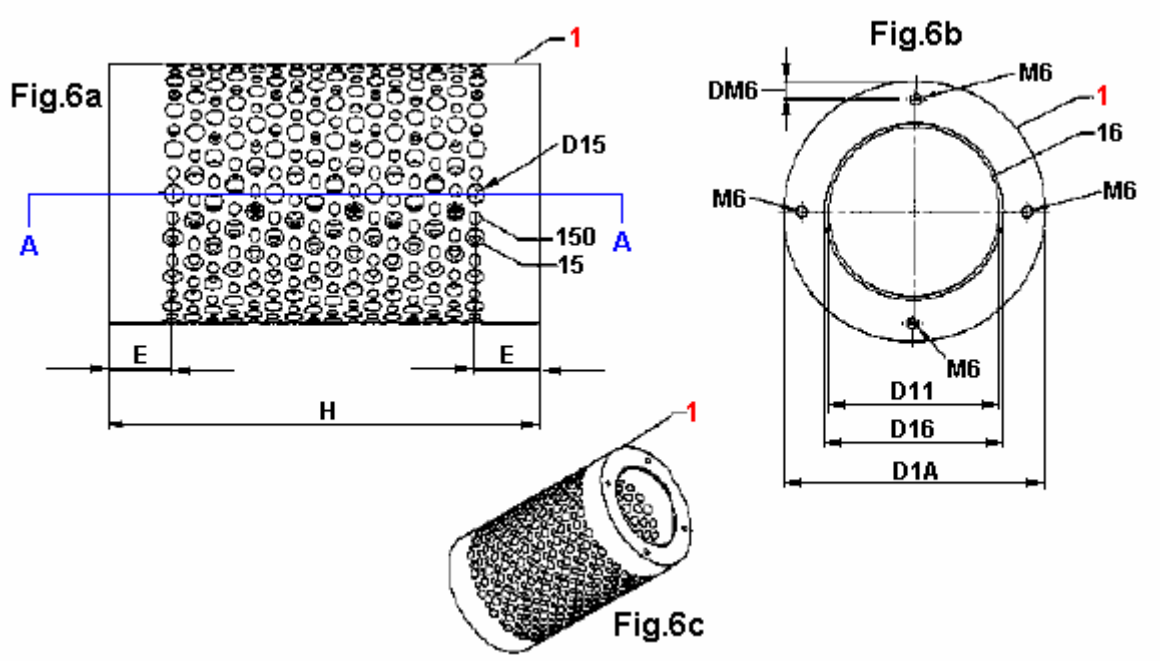


Fig.5a is a view of the fastening device perpendicular to the shaft axis and **Fig.5b** is a view of the fastening device in the direction of the shaft axis



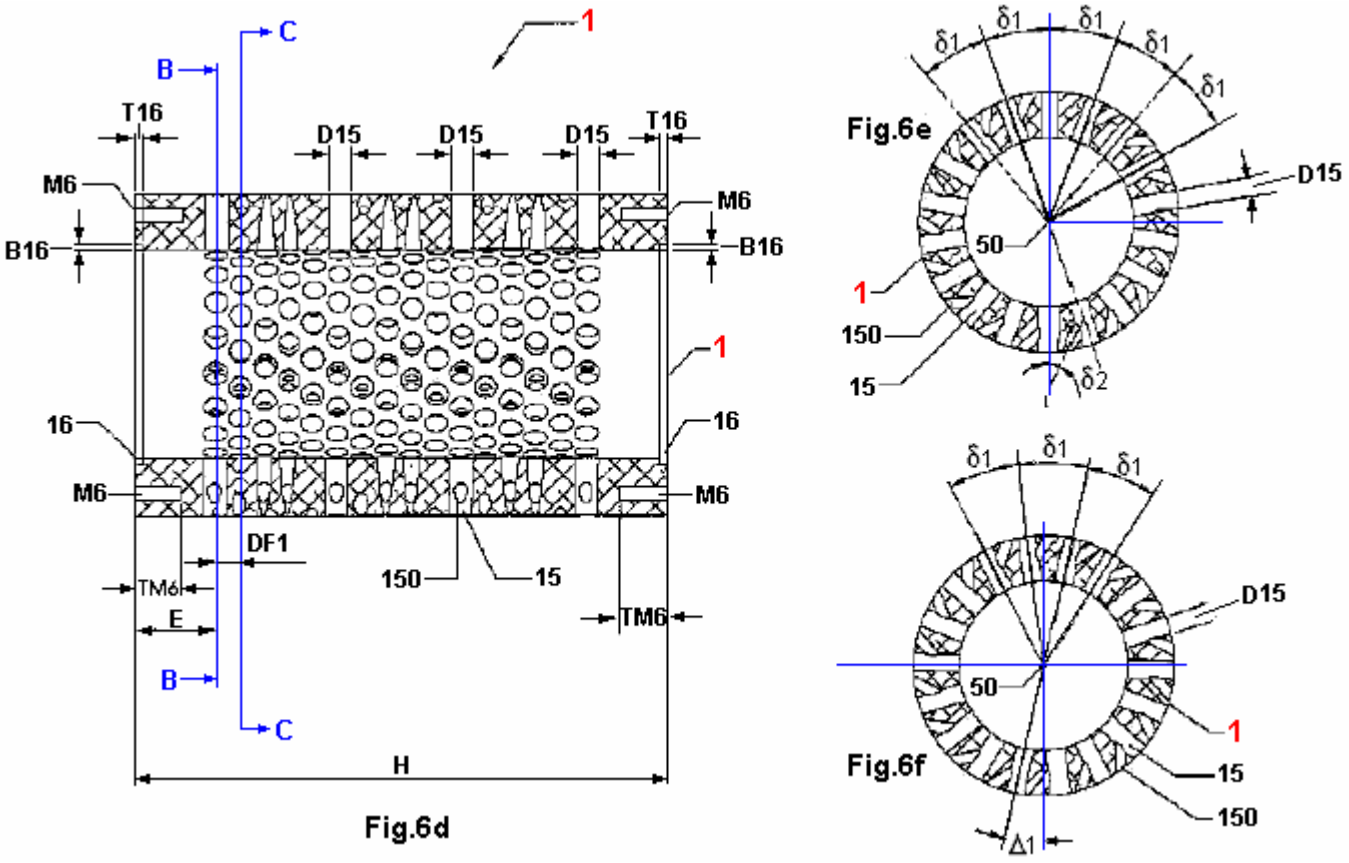


Fig.6 is a perspective view of the rotor

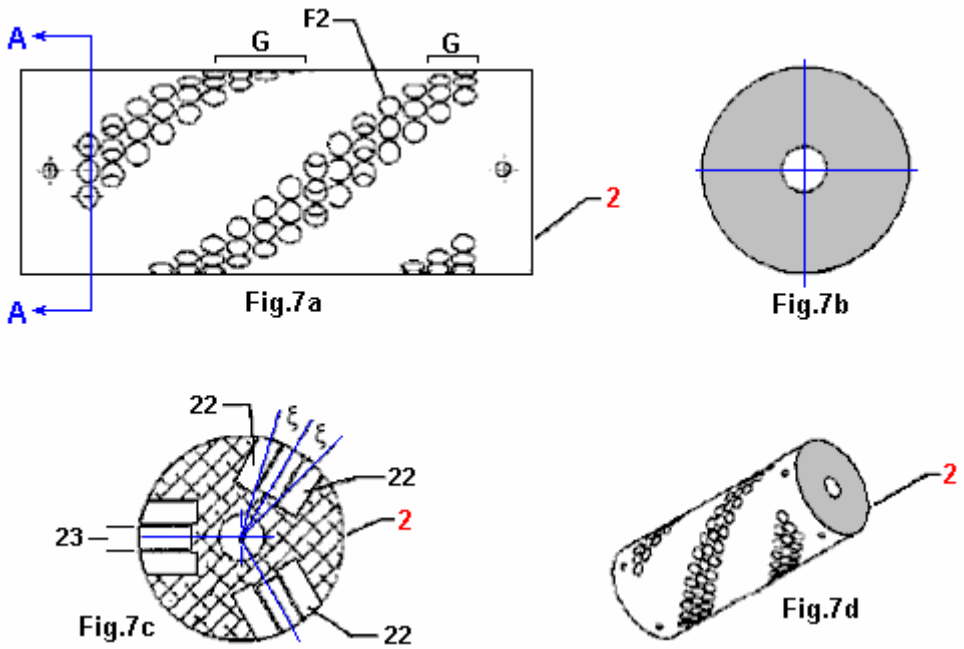


Fig.7a is a schematic view of the inner stator and rotor. **Fig.7b** is a diagram of possible angle of the magnetic axis of the magnets in the rotor;

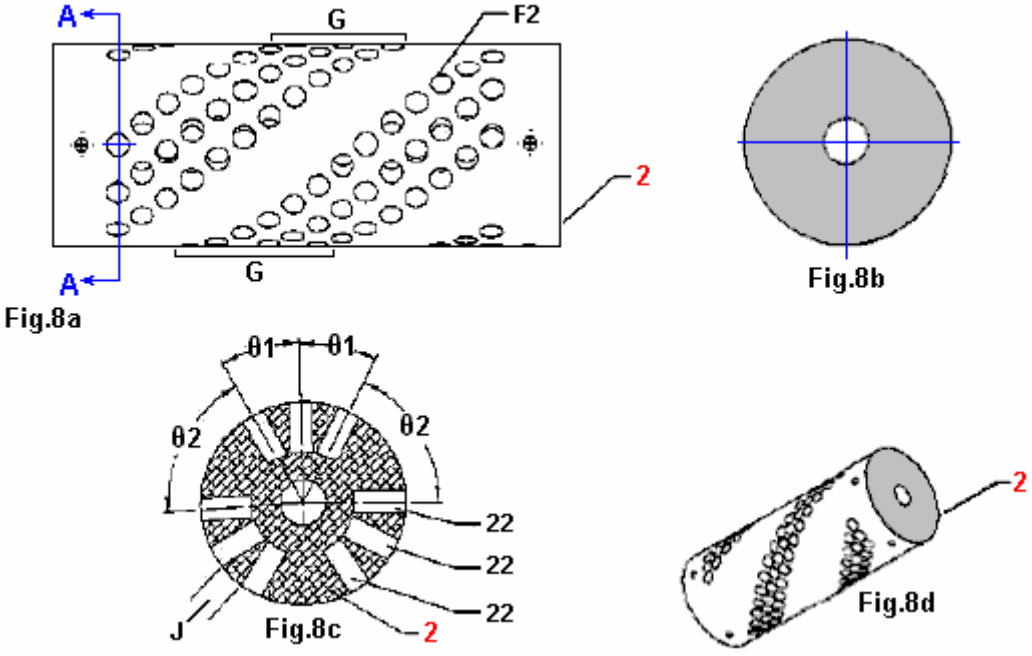


Fig.8a shows the magnetic arrangement of the rotor, along the direction **X--Y** indicated in **Fig.16**. **Fig.8b** is a detailed view of the rotor shown in **Fig.8a**.

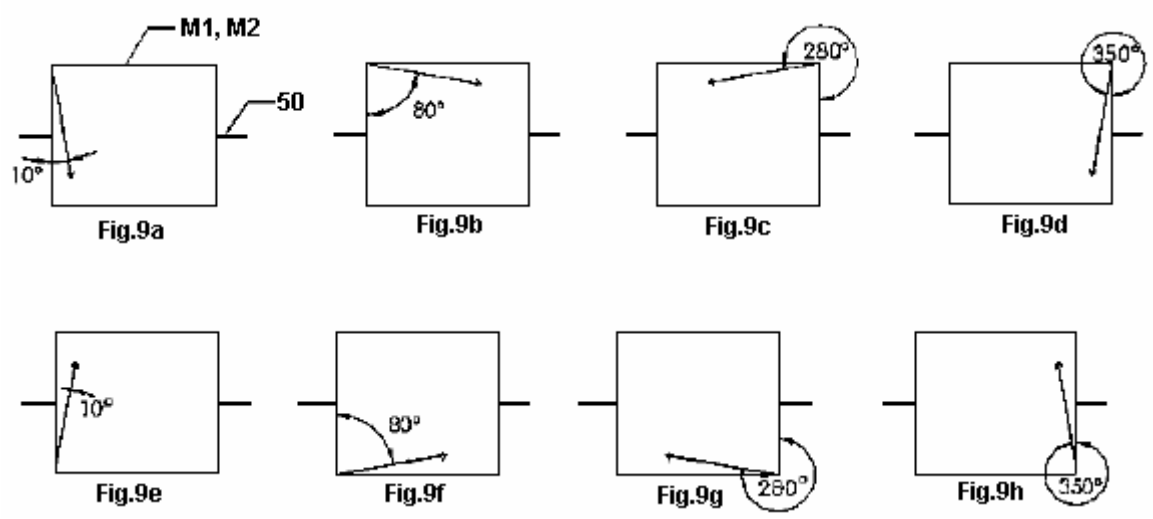


Fig.9a to 9h show the angles of sets of magnets installed in the rotor when viewed from the side. These are shown in greater detail later in this description.

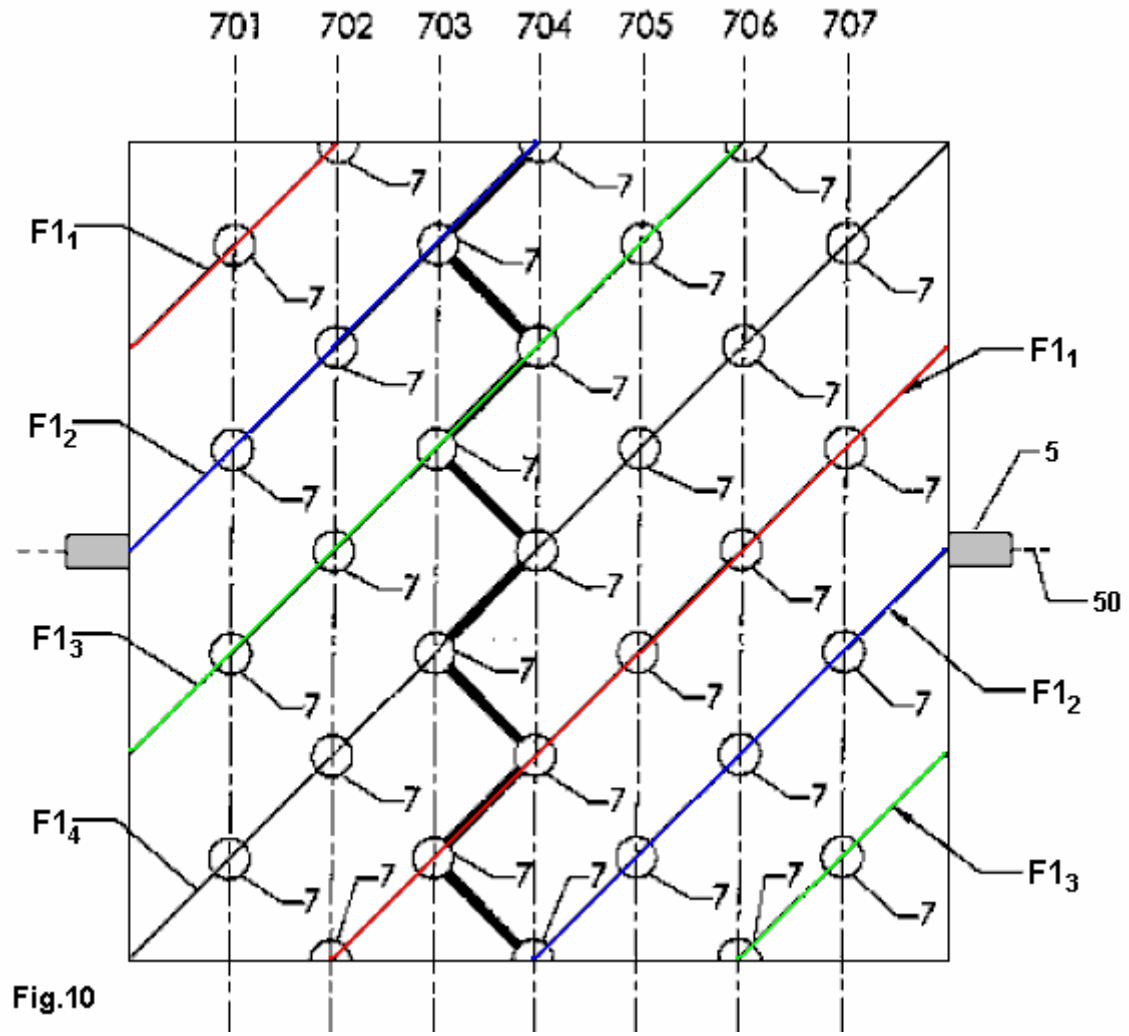


Fig.10

Fig.10 shows the positions of magnet strings embedded in the rotor. These are given in more detail later on.

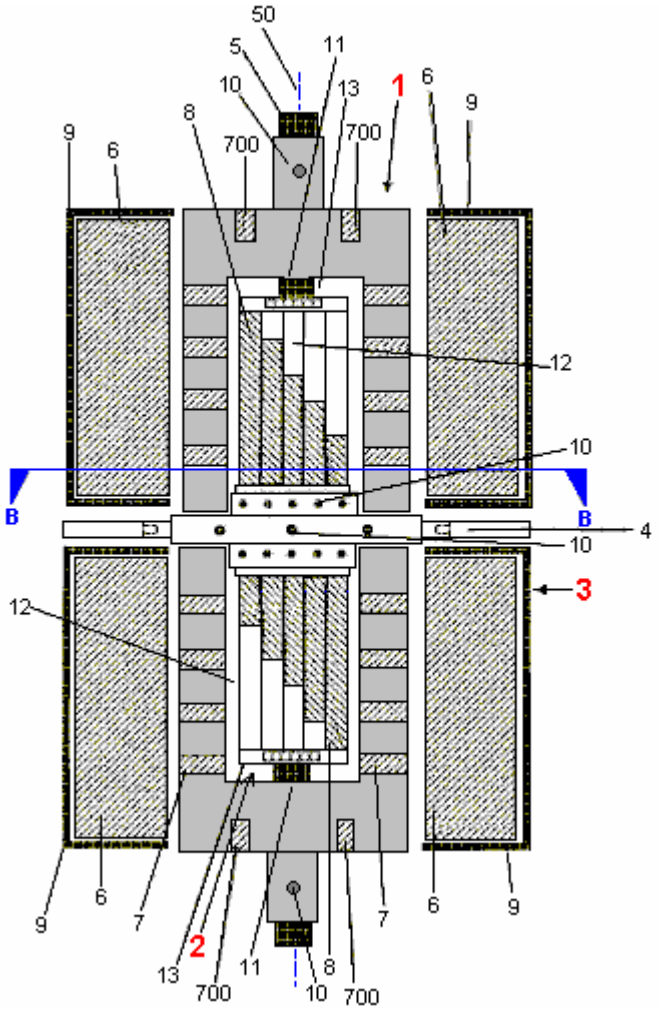


Fig.11 shows the arrangement of magnets on both stators and the rotor, shown as a section along the shaft axis.

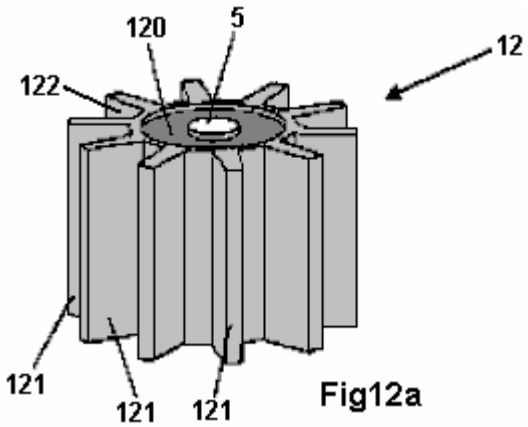


Fig.12a shows the arrangement of cylinder and fins on the rotor before the rotor magnets are installed in the spaces between the fins.

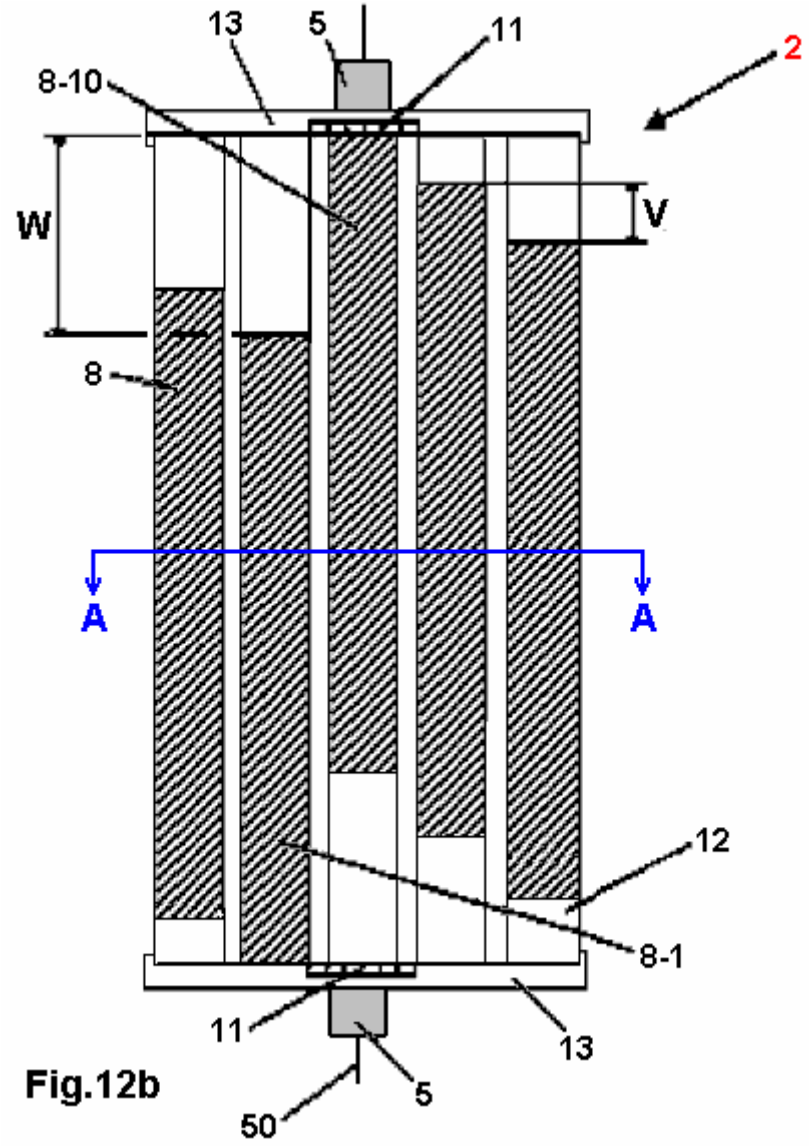


Fig.12b

Fig.12b shows the arrangement of the magnets of the rotor, as seen in a view at right angles to the longitudinal axis of the rotor.

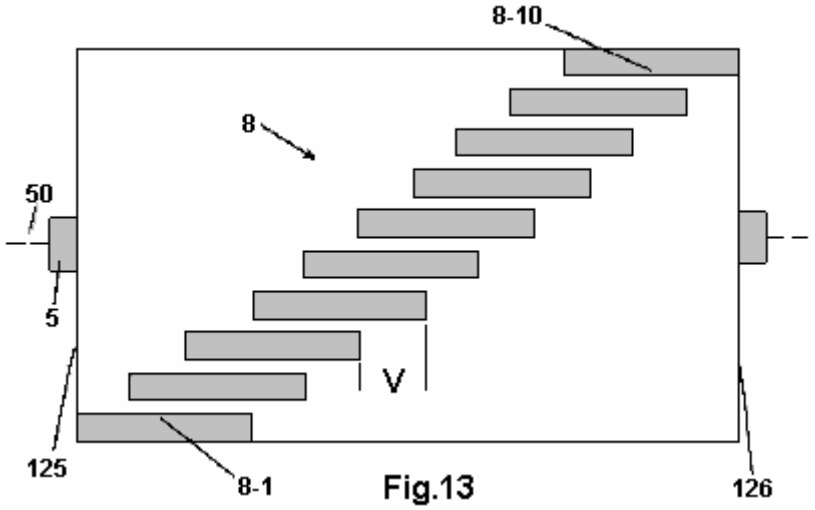


Fig.13

Fig.13 shows the stepped positioning of the magnets of the rotor. This view shows the surface of the rotor and it's shaft, opened out and laid flat. That is, the rectangle show here is actually the whole of the cylindrical

surface of the rotor. In this view, the fins between the magnets are not show in order to emphasise the stepping of the magnets relative to each other.

DETAILED DESCRIPTION

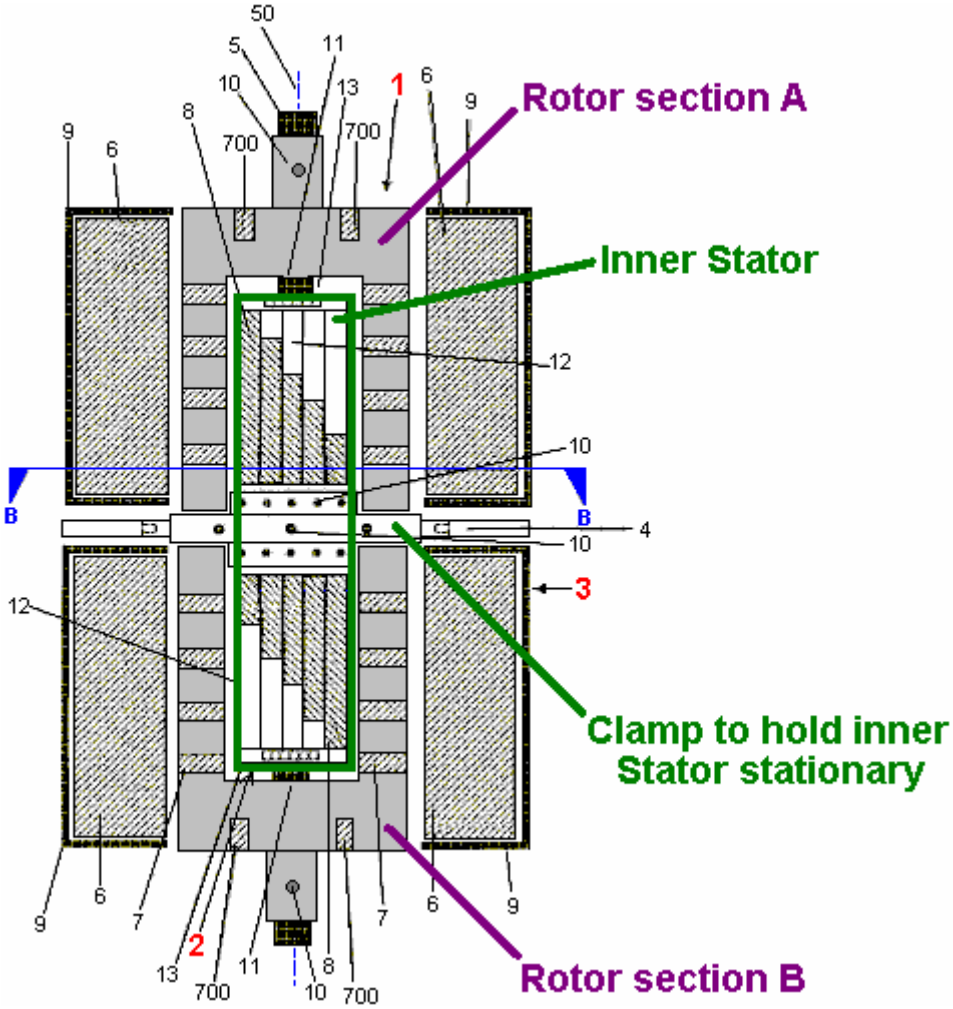


Fig.1 shows a schematic representation of the device having an inner stator **2**, a rotor **1** and an outer stator **3**, which are arranged coaxially around the shaft axis **50** of a pivoting rod-shaped shaft **5**. The cylindrical inner stator **2** has at each end, an end cap **13** which is in the form of a circular disc with a ball-race bearing **11** mounted on it. The bearing **11**, maintains the position of the inner stator **2** relative to shaft **5**. The drive shaft **5** is normally made from a non-magnetic material such as plastic, (not steel) and typically, has a diameter of 10 mm to 40 mm and a length of 100 mm to 400 mm.

The inner stator **2** has a core **12** with magnets **8** mounted on it's outer surface. The inner stator **2** is held stationary by a mounting device **4**, which is secured in position in a mechanical housing (not shown), and is held firmly fixed in this way.

The rotor **1** consists of two mirror-image rotor drums, each with a pipe section and a circular disc section which is clamped rigidly to drive shaft **5** by means of grub screws **10**. Each of the rotor drums has magnets **7** mounted on it. These magnets **7**, are positioned in five distinct places and they have one magnetic pole facing towards the shaft and the other pole facing radially outwards.

The rotor drums are positioned so that there is a cylindrical air gap between them and the inner stator **2**. This air gap is usually of the order of 3mm to 50 mm. Although the two halves of the rotor are separated by the clamping mechanism **4** which prevents the inner stator from rotating, the rotor halves are positioned so that the magnets within them are balanced and so there is no irregular force generated when shaft **5** is spun at high speed. At the ends of the rotor drums there are magnets **700** as the objective of this design is to have the rotor suspended magnetically.

The outer stator **3** is composed of two separate half cylinders **9**. Each of these cylinders **9**, contains magnets **6** mounted on it's inner face. Although each section of the outer stator consists of a hollow cylinder, the outer ends of the stator housing form a complete disc which surrounds the drive shaft **5** and forming a complete enclosure rather than leaving the device open at the ends. There is an air gap between the faces of the magnets mounted on the inner surface of the cylindrical frame **9** and the faces of the magnets mounted on the rotor. These sets of magnets face each other and the air gap between them is also typically 3 mm to 50 mm. The magnets on each of the stators are parallel to the shaft axis **50**. The outer stators is constructed so that it can be moved relative to the inner stator, thus altering their magnetic overlap. This alteration can be made by moving the outer stator when the motor is actually running.

The magnets designated **6**, **7**, and **8**, are dipole magnets and in a preferred embodiment, these are permanent magnets, for example, consisting of SmCo (samarian cobalt) and/or NdFeB (neodymium/iron/boron). It is also possible for one or more of these magnets to be an electromagnet. The magnetic flux density of the magnets **6**, **7**, and **8** is preferably in a range from 0.4 to 1.4 Tesla.

The frame is preferably made from a non-magnetic material such as aluminium with a wall thickness from 2 mm to 10 mm.

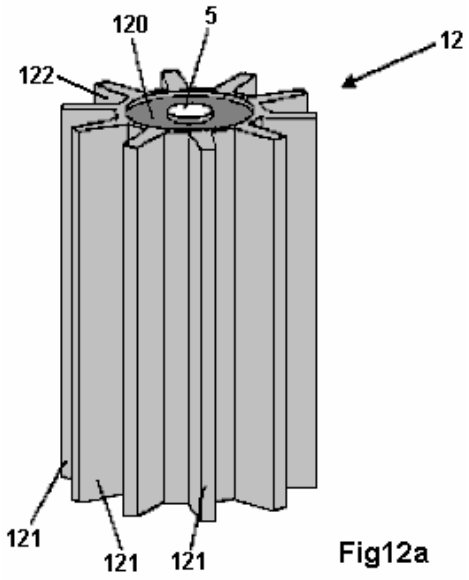


Fig12a

Fig.12a shows an inner stator frame made from a non-magnetic material (such as aluminium or copper). The frame **12** has a circular cylinder **120** which has attached to it's outer surface, radial ribs **121**. Each of these ribs extends along the central axis of the cylinder **120** along the full length of the cylinder, that is, from its' base to the top surface. The ribs are distributed uniformly over the cylinder circumference, forming grooves **122**. Cylinder **120** has a central hole along it's axis for shaft **5** to run through. Both of the end surfaces of cylinder **120** are recessed to accommodate one of the ball bearings **11**. The diameter of the stator core 12 is typically 50 mm to 500 mm with a length of 100 mm to 300 mm. The width of the ribs **121** is generally not more than 100 mm and is usually about 20% of the length of the ribs **121**.

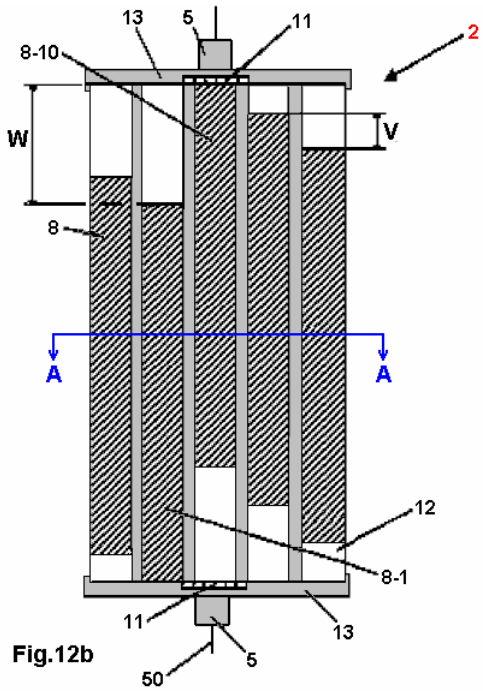


Fig.12b

Fig.12b shows a schematic representation of the inner stator **2**. The inner stator **2** is composed of the inner stator frame **12**, the magnets **8** and the end caps **13**. The magnets **8** are of equal length but their length is less than the length of the stator core **12**. These magnets form the outer surface of the stator. They are seated in the grooves **122** and held in position by the ribs **121**. The first magnet **8-1** is inserted flush with the end cap **13**. The other magnets 8 each have an axial offset **V** along the shaft axis **50** arranged so that there is an even stepping of the magnets with the final magnet **8-10** butting up against the second end plate **13**. The axial offset V is the total overall gap **W** divided by (n - 1), where n is the number of magnets and so, **V** varies with the number of magnets used. In a typical arrangement, **V** is 5% of the length of the magnets **8**.

The end caps **13** have a diameter of 50 mm to 500 mm and a thickness of 5 mm to 20 mm. A typical length for the magnets **8** is 100 mm. The magnet dimensions are arranged so that when they are positioned in the grooves **122**, the inner stator **2** has a substantially uniform outer surface.

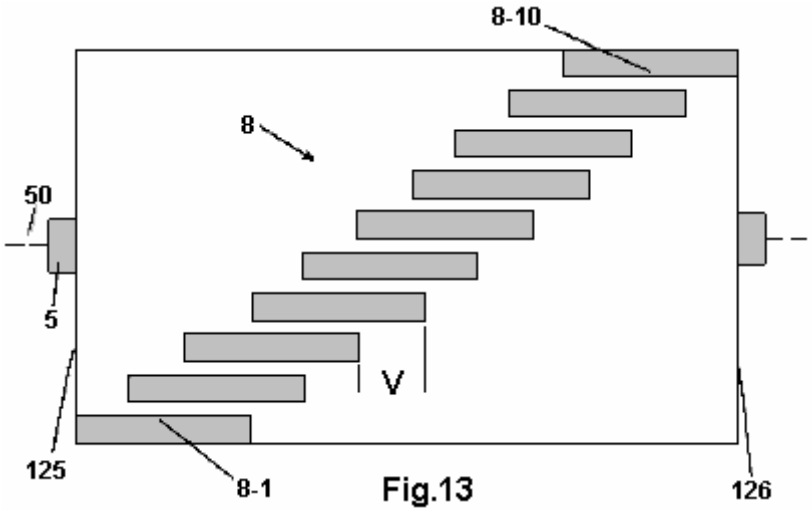


Fig.13

Fig.13 shows an opened-out view of the outer surface of the inner stator **2**. Here, ten magnets **8** are arranged with even spacing. The under side of the magnets taper in the direction of the shaft axis **50** and so they have a lesser width near the centre of the stator than they do at the outside surface. The first magnet **8-1** is positioned

with its end face aligned with the base **125** of the inner stator core **12**. The remaining nine magnets (8-2 to 8-10) are each offset by the amount **V** with the last magnet **8-10** reaching the top surface of the inner stator core **126**.

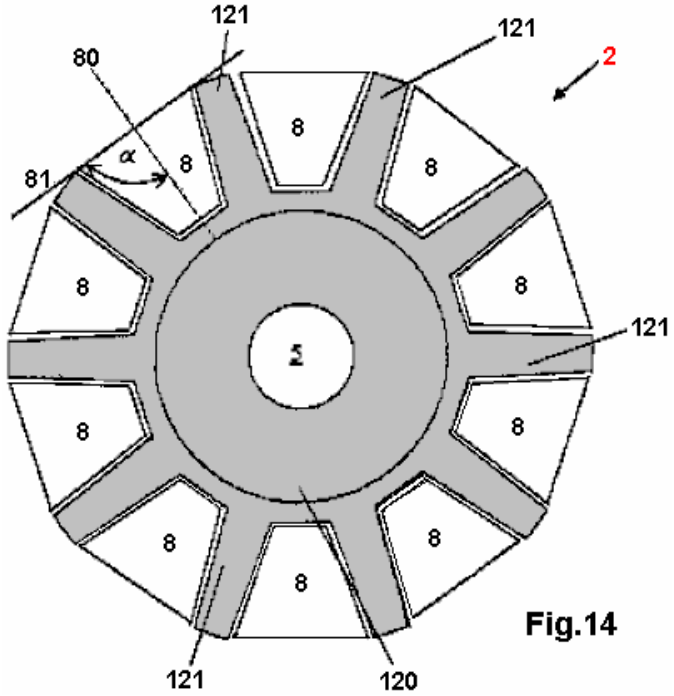


Fig.14

Fig.14 shows a cross-section through the inner stator **2** along the plane **A--A** of **Fig.12b**. The inner stator **2** has a hollow cylinder **120**, through which the central axis of the shaft **5** passes. Running along the outer surface of the cylinder are the ribs **121**. The hollow cylinder **120** typically has a diameter of 100 mm and a length of 170 mm. In the gaps formed between the ribs **121** the magnets **8** are placed. When seen in the plane **A--A** these magnets have a trapezoidal cross-section. These magnets have two magnetic poles and the magnets are positioned so that the magnetic axis **80** which runs through the two poles is radial within the section plane **A--A**. An angle **α** [alpha] formed at the intersection of the magnetic dipole axis **80** of a magnet **8** and the tangent **81** to the ribs **121** can have a value between 14 degrees and 90 degrees. In the case shown in **Fig.14** the angle alpha is 90 degrees.

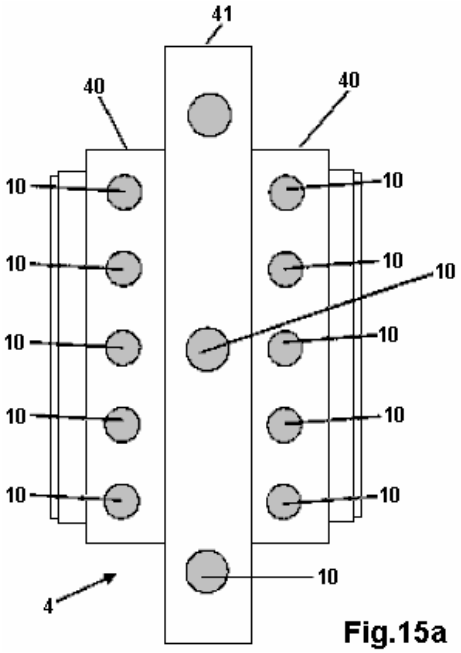


Fig.15a

Fig.15a shows the fastening device **4** in a view perpendicular to the shaft axis **50**. The fastening device **4** has an inner hollow cylinder **40** with a smaller radius and an outer fixing ring plate **41** with larger radius. The inner hollow cylinder **40** and the outer ring fastening plate **41** are connected together. The hollow cylinder **40** is used for receiving and fixing the inner stator **2** by means of screws **10**. The fastening ring **41** is part of a mechanical housing (not shown) for holding the device firmly positioned.

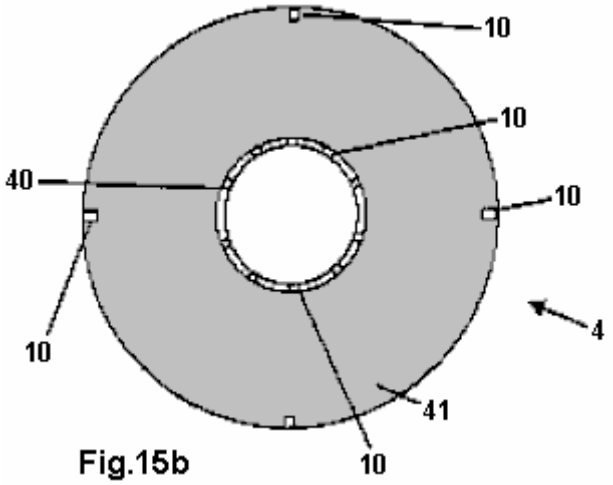


Fig.15b

Fig.15b shows the fastening device **4** in a view in the direction of the shaft axis **50**. The mounting ring plate **41** has at it's periphery, four screws **10** for attachment to the mechanical housing of the hollow cylinder **40** which has on its circumference, a number of screws **10** for fixing the inner stator in place.

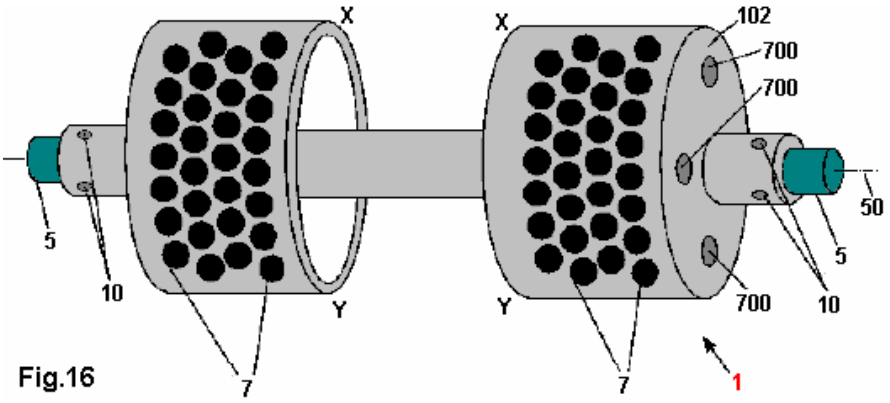


Fig.16

Fig.16 is a view of the rotor **1**, which is clamped to shaft **5** by means of the screws **10**. The rotor **1** consists of two separate drums attached to a central hollow shaft. Mounted in its outer surface are a series of magnets **7** sunk into circular holes. The rotor itself is constructed using a non-magnetic material such as aluminium or copper. The distance between the two rotor drums is 15 mm and they have an outer diameter of 165 mm, a height of 70 mm and a wall thickness of 26 mm. Each rotor drum has a top surface annular disk **102**, into which two or more magnets **700** are sunk. These are positioned uniformly around the circumference of the disc as shown in the diagram. The magnetic dipole axis of magnets **700** is parallel to the shaft axis **50**.

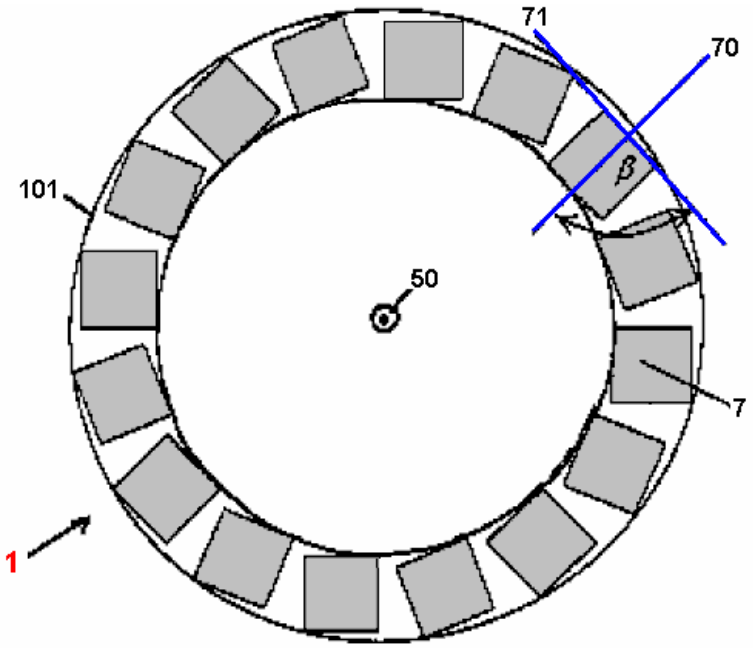


Fig.17a is a schematic representation of the possible orientations of the rotor magnets **7** when seen as viewed looking parallel to the shaft axis **50**. The magnetic dipole axis **70** of rotor magnets seven is in a plane which is radial to the shaft axis **50**. The angle **β** [beta] between the magnetic dipole axis **70** and the tangent **71** breaks through the outer periphery of the hollow cylinder **101** of the rotor **1** and this angle can have values between 14 degrees and 90 degrees.

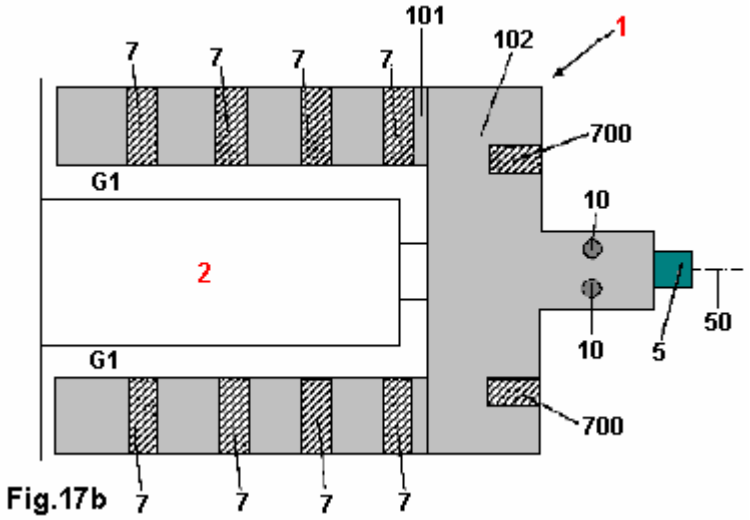


Fig.17b

Fig17b is a schematic view of one rotor drum and part of the inner stator **2**, where the view is perpendicular to the shaft axis **50**. The rotor **1** is clamped to the shaft **5** by the screws **10** and held rigidly in position. The shaft **5** passes through a ball bearing inset into the inner stator **2** and so can rotate freely relative to the inner stator. The rotor has two drum, or bell-shaped, sections which surround the inner stator. The rotor **1** has a hollow cylindrical section **101**, which extends away from the top surface **102**. Since the inner stator is fixed and prevented from rotation by it's anchoring device (component **4** in **Fig.1**), the rotor spins the hollow cylinder **101** around it. The hollow cylinder **101** of rotor **1** is separated from the inner stator **2** by an annular air gap **G1**. The hollow cylinder **101** of rotor **1** has magnets **7** sunk into holes in it. The top surface **102** of the rotor **1** also has holes in it and these are used to install the magnets **700** in it.

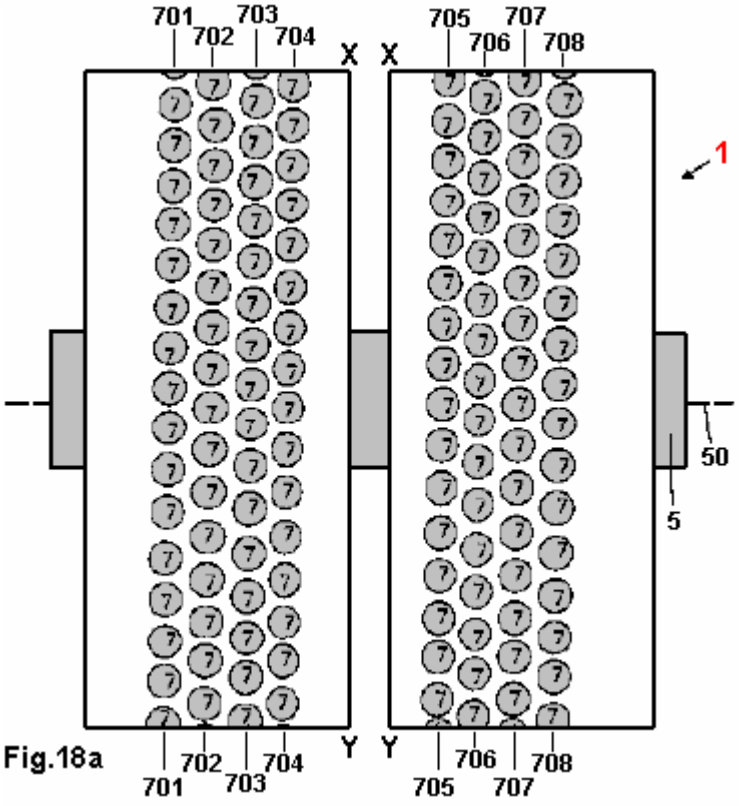


Fig.18a

Fig.18a shows the outer surfaces of the two halves of the rotor drum **1** laid out flat instead of curved into a circle in the **X--Y** plane shown in **Fig.16**. This surface is perpendicular to the shaft axis **50** and rows of magnets **7** are positioned in rows **701 to 708**. Each of these rows is slightly offset in relation to the row beside it, resulting in a zig-zag layout of the magnets **7**.

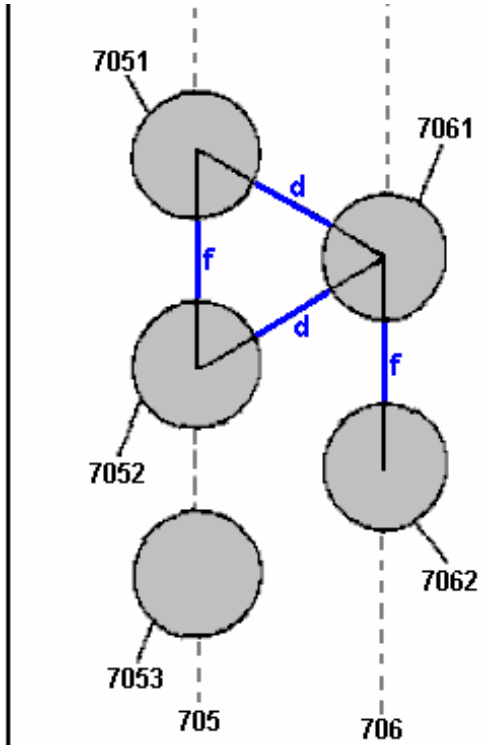


Fig.18b shows, in enlarged detail, the positioning of the magnets **7** shown in **Fig.18a**. The centres of the magnets **7** in the rows **705** and **706** have a constant separation **f** between their edges. The distance between any two adjacent rows, say, **705** and **706**, is chosen so that the arrangement is as shown in **Fig.18b** with constant magnetic separation of length **d** between the edges of the magnets in adjacent rows. For example, the magnets **7051** and **7052** are exactly the same distance apart as magnets **7061** and **7062** the adjacent row **706**. Also, the centres of the three magnets **7051, 7052** and **7061** form an isosceles triangle. This relationship holds for all of the

magnets in all seven series **701** to **708**. Although the magnets **7** are shown in the diagrams as being circular, they could well be other shapes such as square or hexagonal.

The length **d** ranges from about 3 mm to 50 mm. A distance which is particularly preferred, is 5 mm. The distance **f** ranges from about 10 mm to 70 mm.

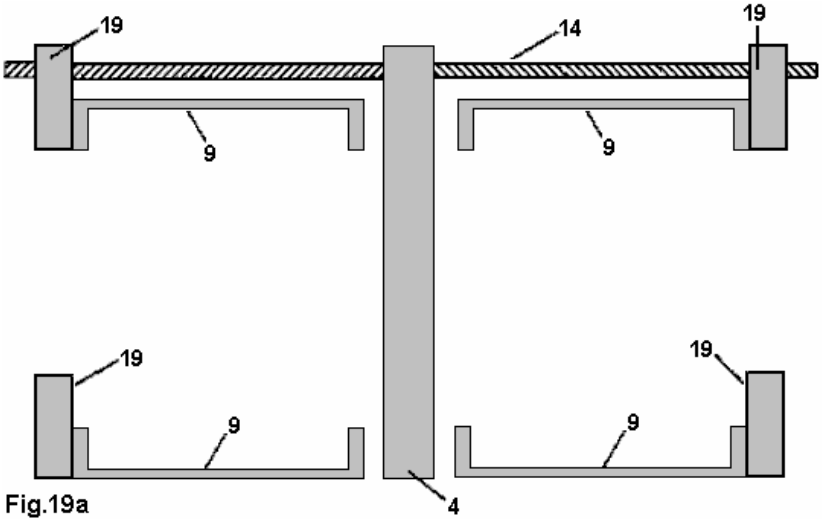


Fig.19a

Fig.19a shows a longitudinal section through the mechanical housing for the device, i.e. a section parallel to the shaft axis **50**. The mechanical housing includes the support piece **4** for clamping the inner stator **2** to prevent it from rotating, the mount **19** for guiding the movable halves of the outer stator **3**, and a rotating threaded rod **14** which can move both halves of the outer stator **3** relative to the rotor and/or the inner stator **2**. The gear shaft **14** has two threaded sections with threads which run in opposite directions (right-hand and left-hand threads). The rotation of this shaft causes the two halves of the outer stator housing to move in a symmetrical manner in opposite directions, inwards or outwards. The guide devices **19** are mounted on the gear shaft **14** and so they only move in one plane. The outer cylindrical sections **9** which house the outer stator **3** are firmly attached to the end caps **19**. Typically, this mechanical housing has a height of 400 to 600 mm, a width of 400 mm and a depth of 530 mm.

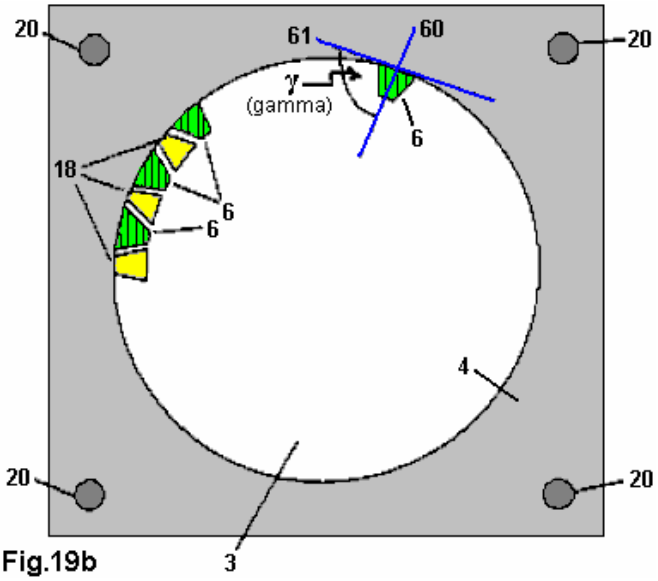


Fig.19b

Fig.19b is a section through the outer stator **3**, the section plane is perpendicular to the shaft axis **50**. The outer stator **3** has arranged in it, a ring of non-magnetic fasteners **18**, between which magnets **6** are secured. For reasons of clarity, only some of the magnets **6** are shown although these magnets are mounted on the entire circumference of the outer stator **3**. The size of the magnets **6** and the non-magnetic fasteners **18** is chosen so that they form a hollow cylinder whose central axis is in the direction of the shaft axis **50**. The magnetic dipole axis **60** of the magnets **6** are perpendicular to the shaft axis **50**. An angle **γ** [gamma] between the magnetic dipole axis **60** and a tangent **61** to the outer periphery of the hollow cylindrical outer stator **3** is between 14 degrees and 90 degrees. The outer stator **3** is connected to the mounting block **4**, which includes the mounting columns **20**.

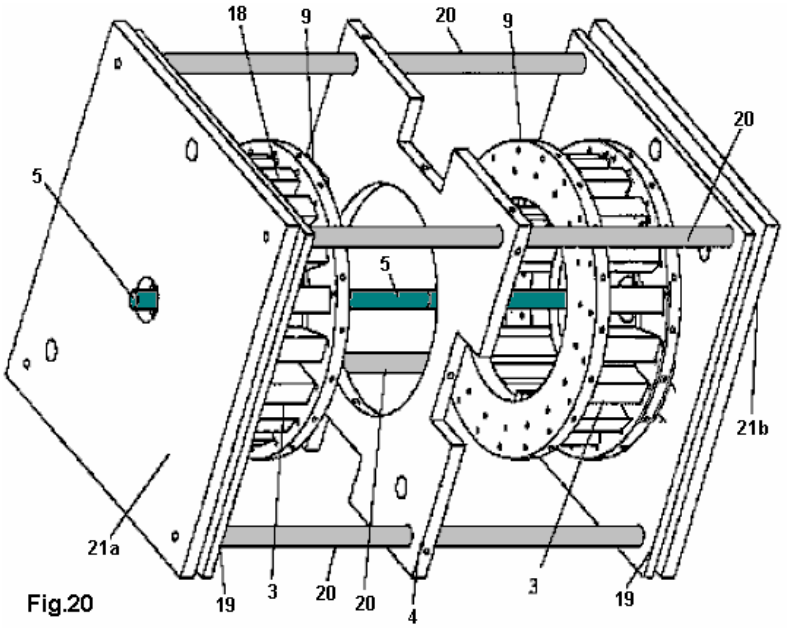
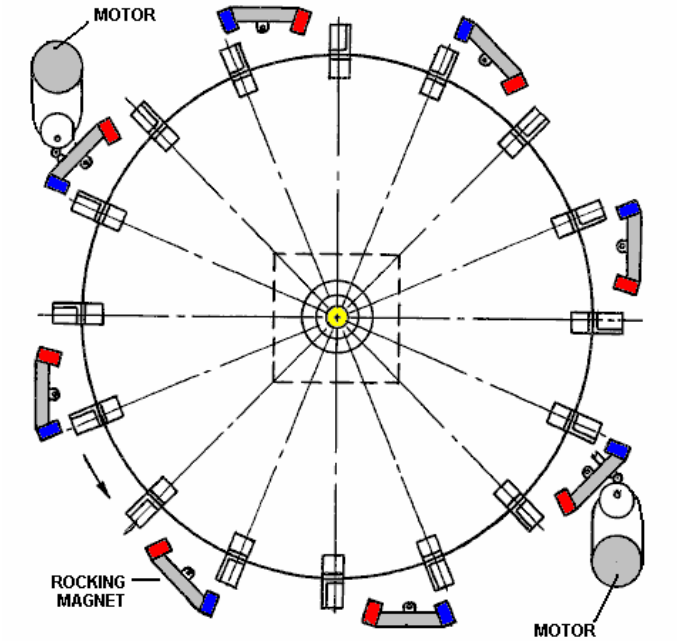


Fig.20

Fig.20 is a perspective view of the mechanical housing for the device. Additional practical details are available in the patent.

Donald Kelly's Permanent Magnet Motor.

In 1979, Mr Kelly was granted a patent on a permanent magnet motor design. He comments that apart from it being very difficult to generate sufficient power to mechanically move the stator magnets slightly to achieve continuous rotation, the resulting rate of revolutions is very low. For those reasons, he has opted to move the stator magnets slightly using small DC motors. His design is included here as it is a concept which is relatively easy to understand. The overall idea is not unlike that of Stephen Kundel who rocks the stator magnets with a solenoid, as shown earlier in this chapter. The objective here is to use a small electrical current to generate a powerful rotation far greater than would be possible from the electrical current itself, and so, produce what is in effect, a power multiplication through the use of permanent magnets. A slightly reworded copy of his patent is shown in the Appendix.

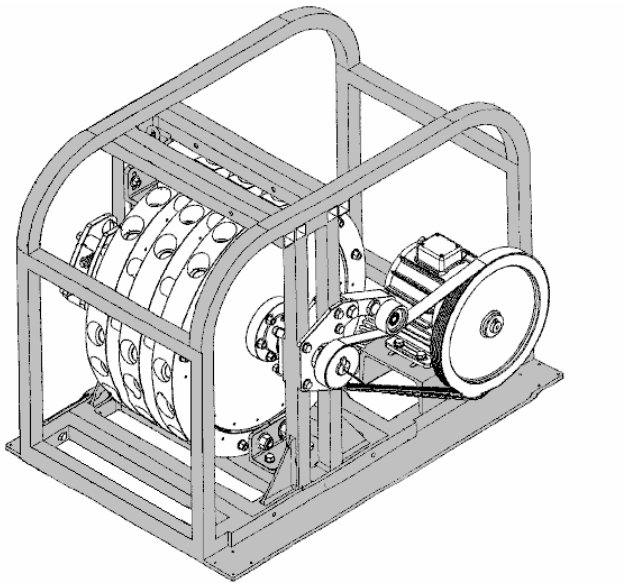


The operation is a simple strategy. Eight sets of magnets are mounted on rocker arms. These have two main positions. In the first position, the rocker magnets attract the magnets mounted on the rotor. When the rotor moves because of this attraction and reaches a point where there is about to be a backward drag on the rotor, the position of the rocker arms is altered so that the first set of rocker magnets are moved out of the way to a position where they have little effect due to their increased distance from the rotor magnets. This rocker movement also

moves magnets of the opposite polarity which push the rotor magnets on their way. In this design, the attraction and the push are applied to different sets of magnets. If the attraction is on magnets 1, 3, 5, etc. then the push is on magnets 2,4,6, etc. But, in spite of this, the pull and push are applied to every rotor magnet as it passes. The power needed to operate the electric motors is minimal as the power of the motor is provided by the magnets. Instead of two tiny motors, it would be possible to operate the rocker arms using small solenoids and if the motor is used to power an electrical generator, then the design could be made self-powered by using some of the electrical output to provide the necessary input power. The sketch above shows just one layer of the motor, but there can be as many layers as you like, each driving the single output shaft, and increasing it's power with every layer.

Mike Brady's "Perendev" Magnet Motor.

One of the most widely known permanent magnet motors is the "Perendev" motor, which catches the imagination of most people. It is said that dozens of these motors have been made and sold as motor/generators with an output of not less than 100 kilowatts. As far as I am aware, this has not been confirmed, nor have there been independent tests made on the motor other than a brief test by Sterling Allan. However, let me stress again that it is **very** difficult to get any permanent-magnet-only motor operating and it is much easier to start with one like the Adams motor shown in Chapter 2, or the Charles Flynn motor shown earlier in this chapter. Please notice as well, that the magnets used in this design are non-standard magnets and so will be difficult to get and probably very expensive because of that and specialised magnetic shielding is used.

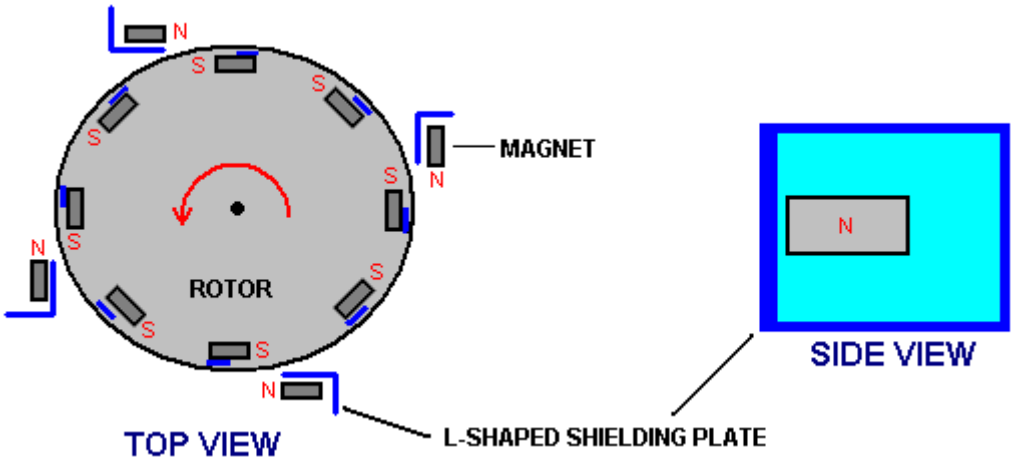


Mike's Patent Application WO 2006/045333 A1 dated 4th May 2006 is shown in the Appendix. In mid 2010, Mike had so much difficulty in getting his design into commercial production that his financial backers are most unhappy with the situation, and if Mike is having difficulty in replicating it (as did Howard Johnson with his magnet motor) then a newcomer to this field would be well advised to stick with magnet motors which use movement of the stator magnets, such as Don Kelly, Stephen Kundel and others, or magnet motors using mechanical or electrical shielding such as the Charles Flynn motor, the Robert Tracy motor, or the Jines motor.

Magnetic shielding from Pasi Mäkilä

A method of blocking a magnetic field using simple materials, comes form Pasi Mäkilä of Finland. His video showing this is at https://www.youtube.com/watch?v=14ayyu9PVSI and he concentrates on placing shielding around a cylindrical magnet:

However, when used as general shielding, a series of flat steel and aluminium layers can be used and while Pasi uses aluminium sheet 1.5 mm thick and zinc-plated steel which is 1 mm thick he suggests using thinner sheets. He suggests using four layers of steel with a sheet of aluminium between the steel sheets and perhaps one or more layers of aluminium on the outside. Pasi's main aim is to share this arrangement to allow people to make permanent magnet motors. One arrangement which may well be worth trying out is to use the shielding to block the backward drag of rotor magnets passing stator magnets, perhaps like this:



With this arrangement, the South poles of the rotor magnets are attracted to the exposed North poles of the stator magnets, causing the rotor to rotate. As soon as a rotor South passes the stator North pole, the stator shielding blocks the reverse pull which would normally slow the rotor down.

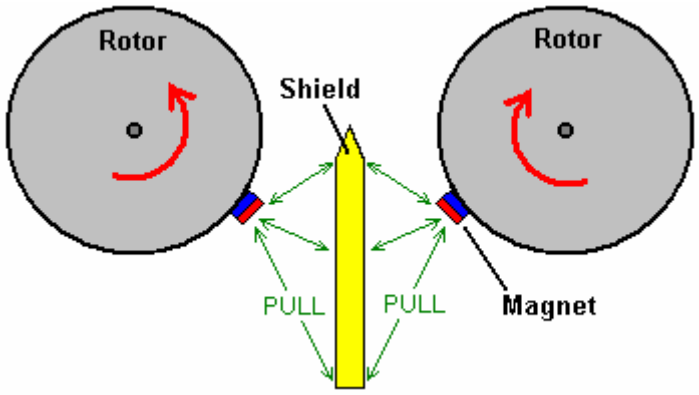
We then have the repulsion of the North pole of the stator magnet and the North pole of the rotor magnet. To block that, a short length of shielding is placed beside the north pole end of the rotor magnet. It would probably be an advantage to run the stator north pole shielding over the top and underside of the rotor magnet to cause major magnetic blocking.

This magnetic motor design is just a suggestion and has not yet been built and tested.

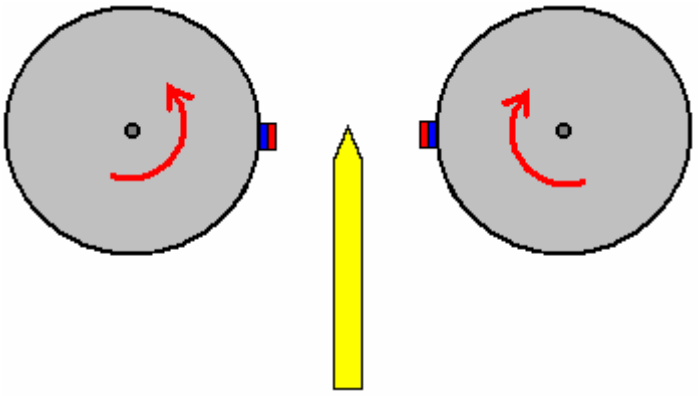
The Twin Rotor Suggestion

When you are considering shielding magnets using iron or steel, you need to remember that fridge magnets stick to refrigerators because the refrigerators are made of steel. This demonstrates the fact that there is an attraction between magnets and iron or steel. Consequently, if a magnet is shielded with steel so that it blocks the whole of the magnetic field of the magnet, a second magnet will be attracted to that metal shielding material. At http://www.youtube.com/watch?v=vUcWn1x3Tss there is, at the present time, a video by "magneticveil" where he proposes the use of this feature of simple shielding in the construction of a magnet motor.

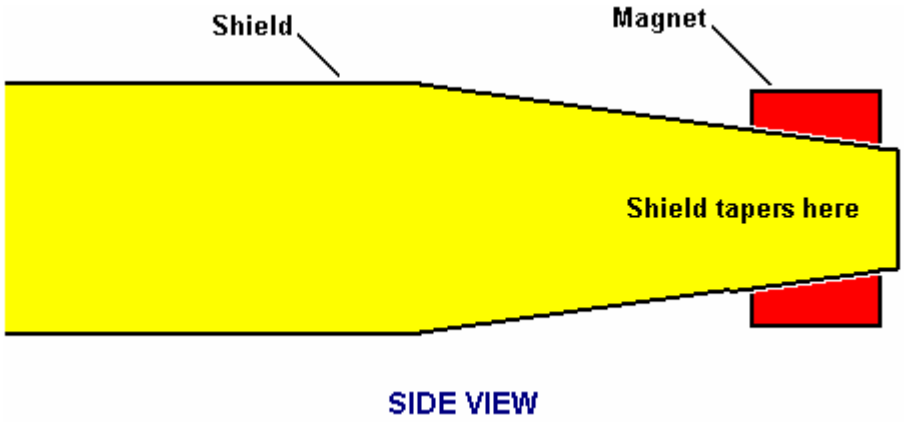
He suggests using two rotors geared together. The rotors have magnets on them, but for the purposes of explanation, just one pair of magnets are shown here:



Each magnet is attracted to the metal shield material between the rotors. This causes the rotors to rotate in the direction shown by the red arrows. The magnets are drawn to the nearest point to the shield which they can reach as shown here:

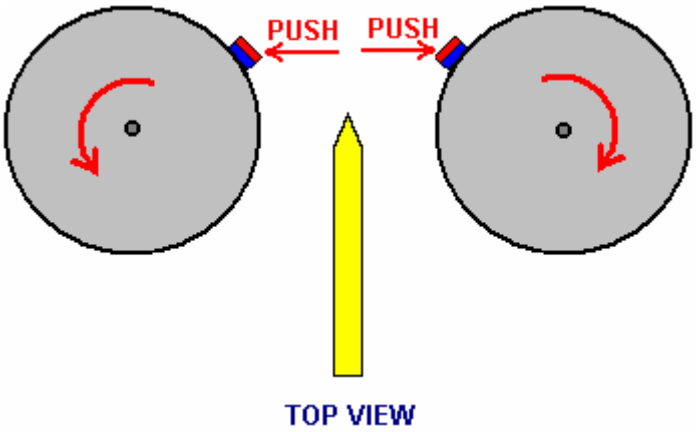


At this point you would expect the rotors to stop moving and lock into a stationary position. However, the interesting idea is to adjust the shape of the shield like this:



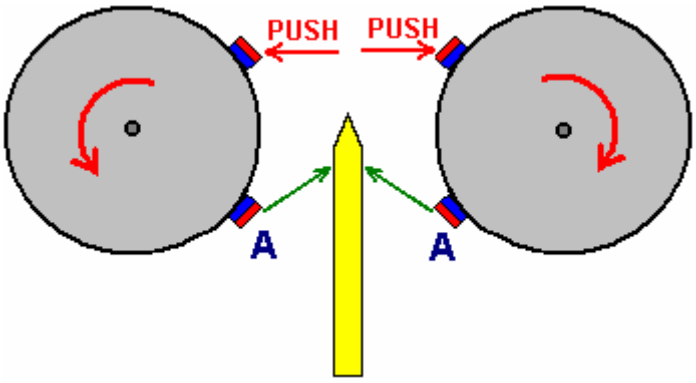
At the end of the shield, its width is reduced and tapered so that the magnetic field from the magnet behind it exactly matches the attraction of the magnet on the near side of the shield. This has the effect of giving a completely neutral zone at the tip of the shield, with neither an attraction or a repulsion in that region. The degree of tapering depends on the strength of the magnets, the thickness and material of the shield and the spacing between the magnets and the shield, and it needs to be discovered by experiment.

This neutral zone stops there being a major pull between the magnets and the shield, and so momentum carries the rotors on past the end of the shield. This produces a situation like this:



Here, the magnets have moved past the shield and are repelling each other strongly. They are beyond the axles of the rotors, so the repelling force produces a turning effect on each rotor. This is the situation with just one pair

of magnets, but each rotor will have many magnets on it. This produces an additional turning effect. Consider just one other pair of magnets, in the same position as our first diagram:



The pull between the magnets “A” and the shield, adds to the rotation caused by the push between the unshielded magnets. This arrangement of magnets and shield should allow continuous rotation of both rotors and the motor can be stopped by removing the shield.

It should be noted that this arrangement uses magnets in repulsion mode. That is, the outward-facing poles of the magnets on both rotors are the same. There have been reports of permanent magnet motors where the magnets were in repulsion mode, and while these motors ran well, it was found that after about three months, the magnets lost their magnetisation. If at all possible, magnets should be used in their attraction mode. This is not possible in the above twin-rotor arrangement, so if one is being constructed, it might be a good idea to arrange the physical construction in such a way that the rotor magnets can easily be removed. This allows remagnetisation of the magnets, or alternatively, their replacement if very cheap types are used.

The Permanent Magnet Motor of Victor Diduck

In US patent application number US2007/0296284 of 27th December 2007, there is shown a convincing design for a powerful permanent magnet motor. Here is one of the embodiments from that patent – one which looks reasonably easy to build.

Magnetic Motor

Abstract:
A magnetic motor having a magnetic drive assembly magnetically coupled to a magnetic slave assembly. The drive assembly has at least one drive magnet. In one embodiment the drive magnet is mounted on a cowling. In another embodiment the drive magnet is mounted on a drive wheel. The slave assembly has at least one slave wheel mounted on a slave shaft. At least one slave magnet is mounted on the slave wheel. In one embodiment slave magnets are mounted in grooves running diagonally across the face of the slave wheel. In another embodiment the slave magnets are mounted in notches cut into the slave wheel. The drive magnet is magnetically coupled to the slave magnet with the poles arranged in a like-faces-like orientation. The gap between the drive magnet and slave magnet can be adjusted in order to optimise the magnetic coupling the between them. The slave wheel and its slave shaft are caused to rotate by the magnetic coupling between the drive magnet and the slave magnet. The slave shaft can be coupled to an output device such as an electric generator.

BACKGROUND OF THE INVENTION
There have been a number of attempts to perfect magnetic motors; for instance, U.S. Pat. No. 4,151,431 issued to Howard Johnson. However, in most such devices no working models have been achieved. In order to make a permanent magnet motor operate it is necessary to accomplish a switching function equivalent to that accomplished in electric motors by brushes, commutators, alternating current, or other means. In permanent magnet motors magnetic leakage must be shielded so as to reduce energy lost as eddy-current energy. A proper combination of materials, geometry, and magnetic concentration is required in order to be able to construct a magnetic motor that can operate continuously.

SUMMARY OF THE INVENTION
A magnetic motor is provided comprising a magnetic drive assembly magnetically coupled to a magnetic slave assembly. The magnetic slave assembly includes a rotatable slave shaft upon which is mounted at least one rotatable slave wheel. Upon the slave wheel is mounted at least one slave magnet. The magnetic drive assembly

includes at least one drive magnet which is magnetically coupled to the slave magnet in a like-faces-like orientation. As a result of the magnetic coupling between the drive magnet and the slave magnet, magnetic forces produced between the coupled drive magnet and slave magnet drive the rotatable slave wheel, making it rotate and therefore causing the slave shaft to rotate. The slave shaft is coupled to an output device such as the armature of an electric generator.

The slave assembly is coupled to a frame. The slave wheels are fixed to the shaft so that the wheels rotate together. Each slave wheel has embedded in its surface a plurality of slave magnets set in indentations cut into the slave wheel. One pole of each slave magnet is exposed and facing outwards from the surface of the slave wheel, and the other magnet pole faces the slave wheel. Either the north pole or the south pole of the slave magnets may face outward, as long each magnet has the same pole facing outwards.

In one embodiment the indentations in the slave wheels for receiving the slave magnets form spaced apart, parallel grooves running from one side of the surface of the slave wheel to the other for receiving the slave magnets. The angle of each groove across the surface of the slave wheel is preferably about 35 degrees with respect to horizontal. The direction of orientation of the grooves of the other of the slave wheels is also about 35 degrees off of the horizontal, but in the opposite direction to that of the first wheel.

In another embodiment the indentations in the slave wheels for receiving the slave magnets are notches cut into the slave wheel at measured and equal intervals along the edges of the wheel, intervals of 45 degrees being preferred.

In this “cowling” embodiment of the invention, the magnetic drive assembly comprises a pair of non-magnetic cowlings surrounding and substantially enclosing each of the slave wheels. Each pair of cowlings forms a semi-circular surface having a diameter slightly larger than the diameter of its respective slave wheel. The concave curvature of the cowlings faces the slave wheels. Mounted on the convex surface of the cowlings are a plurality of permanent drive magnets. The drive magnets are mounted so that they present to the slave magnets the same pole as the slave magnets present to the drive magnets; i.e., like-faces-like: north-to-north or south-to-south. Neither the cowlings nor their drive magnets rotate.

In the various embodiments, the gap between the drive magnets and the slave magnets is adjustable.

BRIEF DESCRIPTION OF THE DRAWINGS

Further features and advantages of the invention will be apparent from the following detailed description taken in conjunction with the accompanying drawings, where:

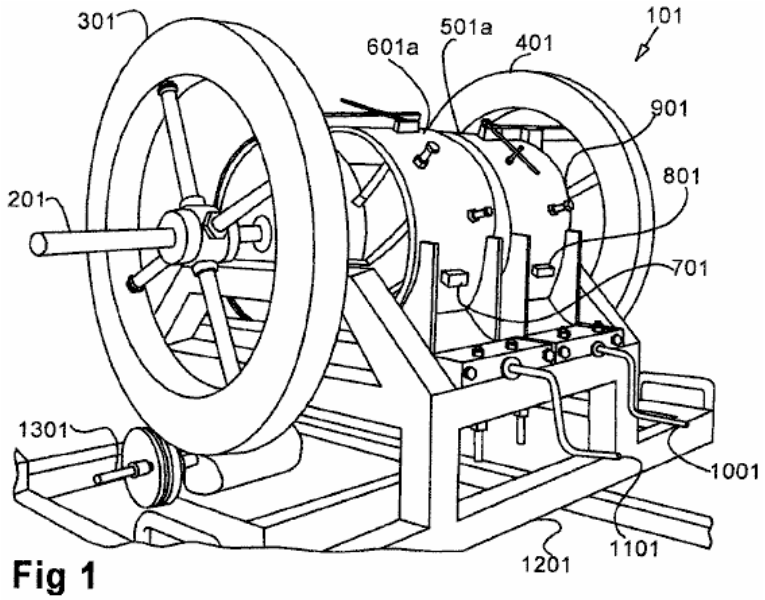


Fig 1

Fig.1 is a perspective view of the cowling embodiment of the magnetic motor with fly wheels attached.

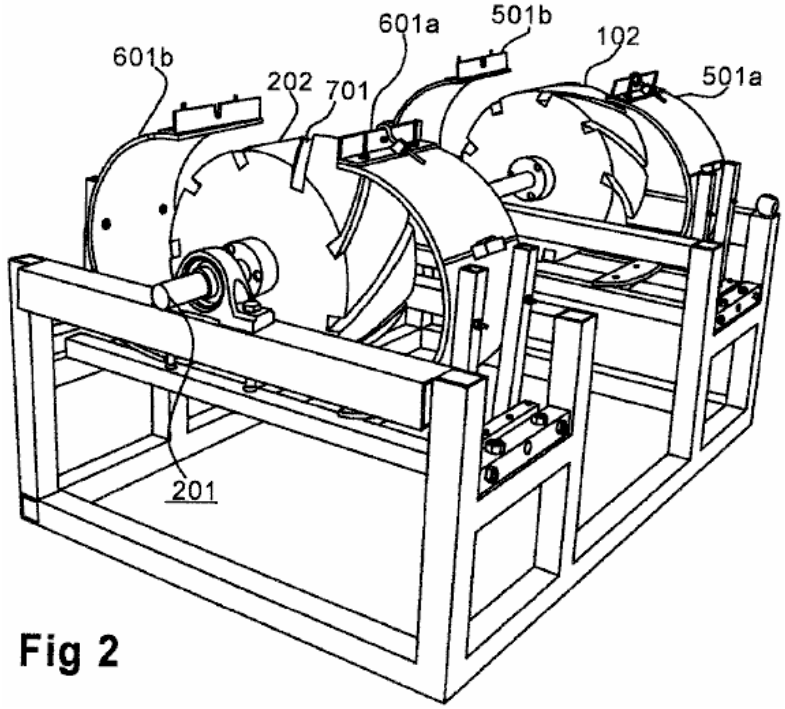


Fig 2

Fig.2 is partially disassembled perspective view of the cowling embodiment of the magnetic motor.

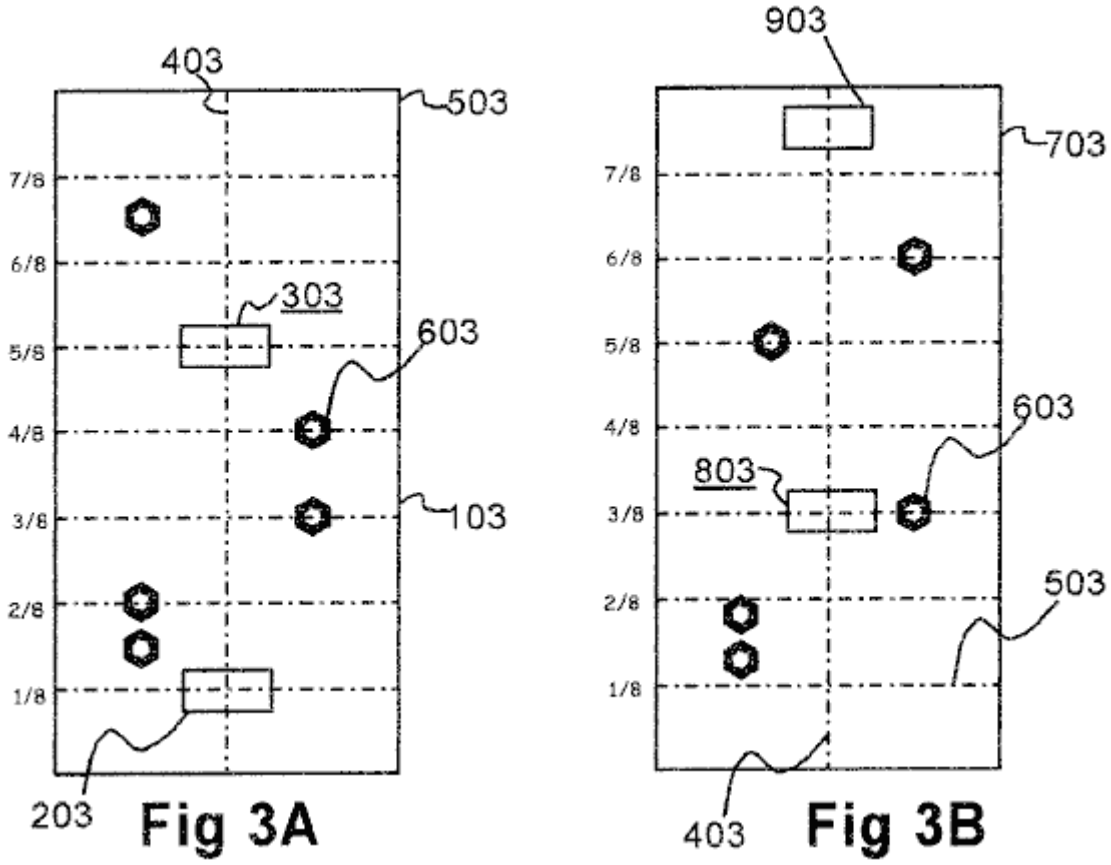


Fig 3A Fig 3B

Fig.3 is a diagram of the magnet placement on the cowling.

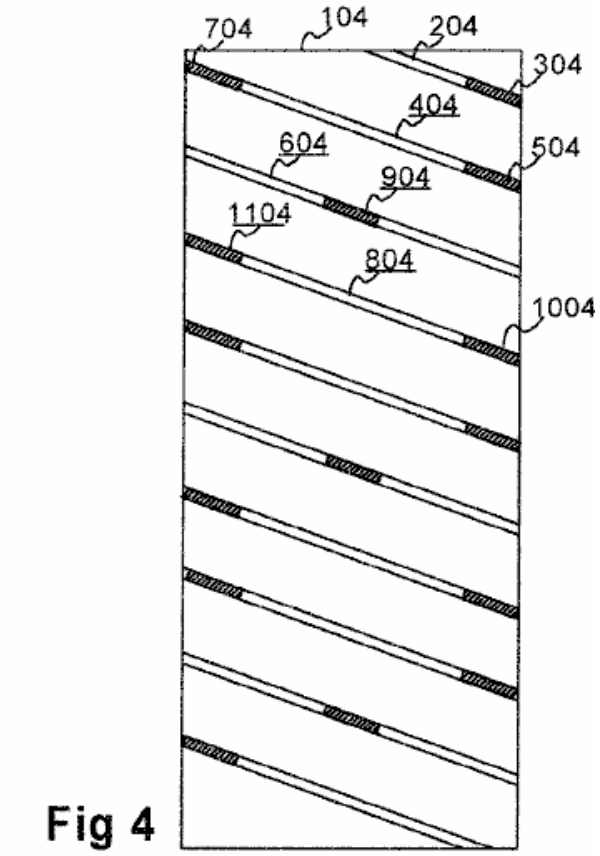


Fig 4

Fig.4 is a schematic diagram of one slave wheel of the cowlings embodiment showing the position of the permanent magnets.

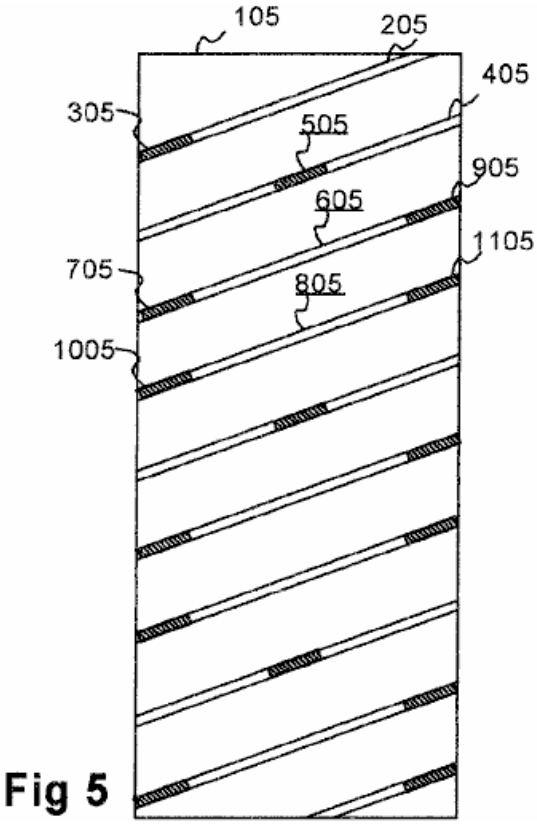


Fig 5

Fig.5 is a schematic diagram of another slave wheel of the cowlings embodiment showing the position of the permanent magnets.

DETAILED DESCRIPTION OF THE INVENTION

In the various embodiments of the invention there is generally provided a magnetic drive assembly and a magnetic slave assembly, with a magnetic field coupling the drive assembly to the slave assembly such that when the drive assembly rotates it causes the slave assembly to rotate. The coupling is entirely magnetic, where no chains, gears, pulleys, worm drives or other physical couplers are required.

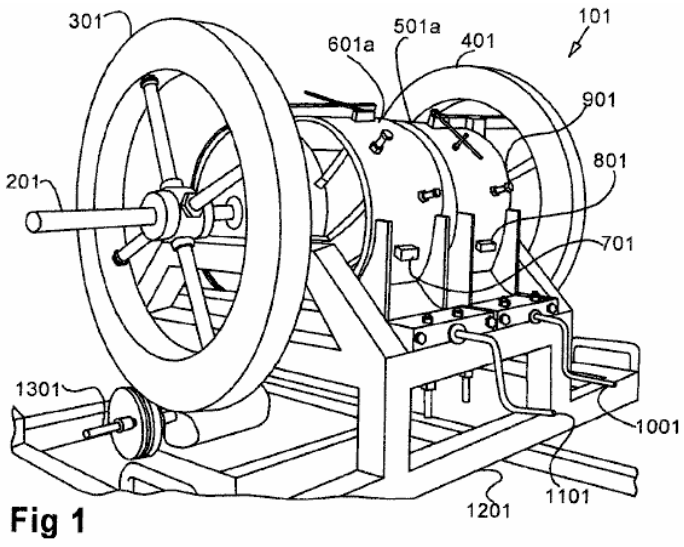


Fig 1

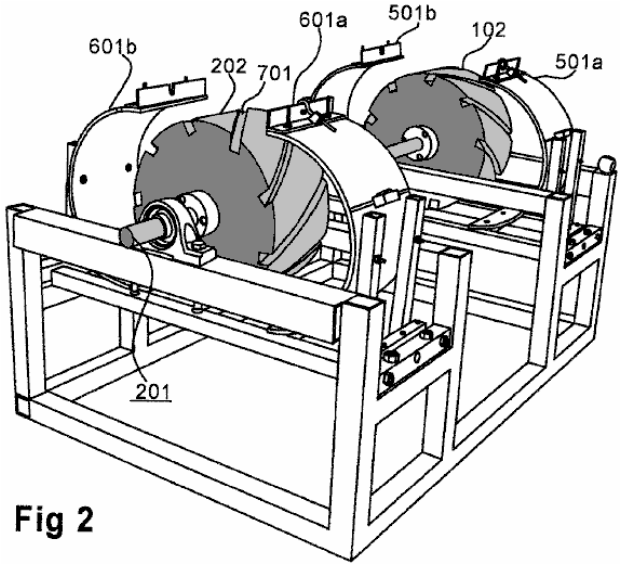


Fig 2

Fig.1 and **Fig.2** show a first embodiment of the invention, referred to herein as the "cowling" embodiment. In this embodiment the magnetic slave assembly of the magnetic motor **101** comprises two solid non-magnetic slave wheels **102** and **202**, most clearly seen in **Fig.2**. The slave wheels are mounted on a slave shaft **201**. **Fig.1** shows an embodiment in which optional fly-wheels **301**, **401** are mounted on slave shaft **201**. The fly-wheels may be conveniently mounted at or near the ends of the slave shaft. A device **1301** for generating electric current is provided that is directly coupled to the slave shaft, or indirectly coupled through a fly-wheel, as shown in **Fig.1**, or though some other element of the magnetic slave assembly.

Except for elements noted herein, the invention is constructed of a non-magnetic material. Pheotic plastic or ceramic materials are currently preferred for the slave wheels and drive wheels, but a wide variety of non-magnetic materials is acceptable so long as the material does not create or exacerbate eddy currents. The diameter of the slave wheels in currently operating models is approximately 10 inches (250 mm), and the width approximately 5 inches (125 mm). The optimum dimensions of the slave wheels will be determined by the specific application of the invention.

As seen in **Fig.2**, each slave wheel has a plurality of grooves running from one side to the other. One such groove is designated **701**. The grooves in one wheel are oriented at an angle of about 35 degrees to the slave wheel edge, while the grooves of the second wheel are oriented at about 35 degrees to the opposite edge, as can be seen clearly in **Fig.2**.

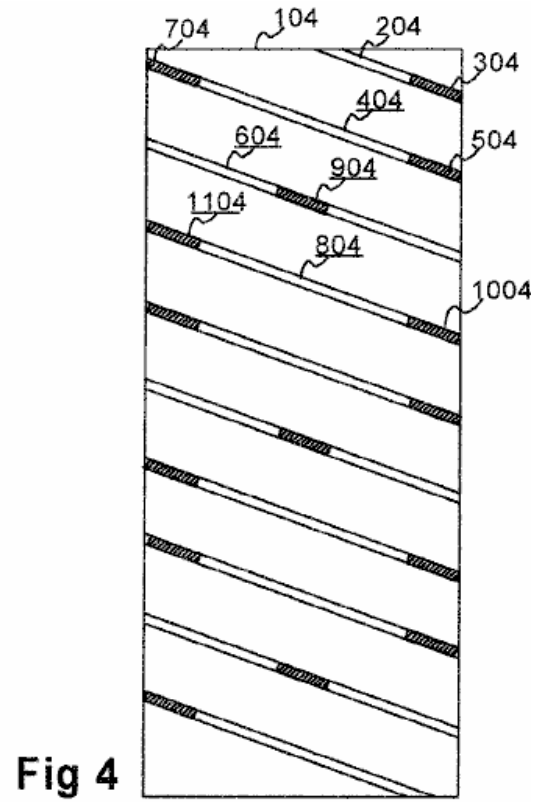


Fig 4

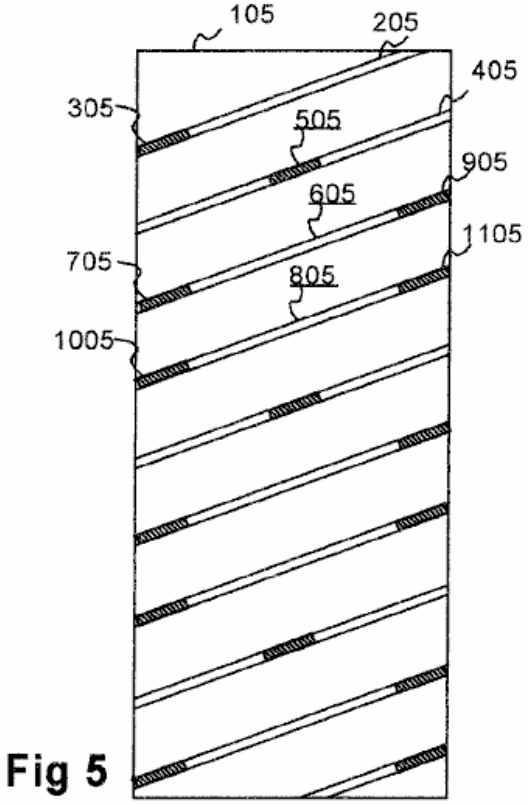


Fig 5

Fig.4 and **Fig.5** demonstrate the orientation of the grooves and the placement of the slave magnets. The rectangles **104** and **105** represent the surfaces of the slave wheels as if they were laid out flat. The grooves in slave wheel **104** slope downwards from left to right at an angle of about 35 degrees to the horizontal. The grooves in slave wheel **501** slope upward from left to right at an angle of about 35 degrees to the horizontal. In **Fig.4**, grooves **204**, **404**, **604**, and **804** are representative of the grooves in one slave wheel. Grooves **205**, **405**, **605**, and **805** of the slave wheel represented in **Fig.5** are representative of grooves in the other slave wheel.
Slave magnets are fitted into the grooves. In **Fig.4**, representative slave magnets are **304**, **504**, **704**, **904**, **1004**, and **1104**. The preferred position of the slave magnets is that two adjacent grooves have magnets positioned at their ends as shown with **304**, **504**, and **704** in grooves **204** and **404**. The next groove **604** has a single slave magnet **904** centrally placed. This pattern of two grooves with end magnets and the third with a central magnet is repeated. The preferred embodiment has a total of 9 grooves and 15 slave magnets per slave wheel. **Fig.5** shows that the same pattern is used in the second slave wheel, for instance in the manner in which slave magnets **305**, **505**, **706**, **905**, **1005**, and **1105** are positioned in grooves **205**, **405**, **605**, and **805**.

In the preferred embodiment, the north pole of each slave magnet faces outwards from the groove; however, having the south pole facing outwards produces equally satisfactory results. The magnets can be glued into place or otherwise firmly fixed so they do not shift. The attractive forces these magnets produce if opposite poles are allowed to make magnetic contact requires approximately 1200 ft. lbs. to overcome. Slave and drive magnets are permanent magnets and have the same pole facing outwards, producing repulsive forces on the order of a measured 38 gauss.

The magnetic drive assembly of the “cowling” embodiment comprises paired clam-shell cowlings **601*a***, **601*b*** and **501*a*, 501*b***, best seen in **Fig.2**, which shows the cowlings in an open position, exposing the slave wheels. **Fig.1** shows the cowlings in the closed position, in which the invention operates. Crank handles **1001**, **1101** operate worm-drives to provide for opening and closing the cowlings in order to adjust the gap between the cowlings and the slave wheels, and, hence, the gap between the drive magnets and the slave magnets.

Fig.1 also shows drive magnets **701**, **801** placed on the outer surface of cowlings **501*a*** and **601*a*** respectively. A plurality of ferro-magnetic bolts **901** penetrate the clam-shell cowling through threaded holes. These bolts modify the magnetic field and eliminate dead spots. The placing of the drive magnets and bolts is discussed below.

From **Fig.1** it can be seen that the combined curvature of the paired clam-shell cowlings results in them nearly surrounding their respective slave wheel when in the closed position. That is, each member of a cowling pair surrounds somewhat less than 180 degrees of the slave wheel's circumference so that when juxtaposed in the closed position, together they surround nearly 360 degrees of the slave wheel circumference.

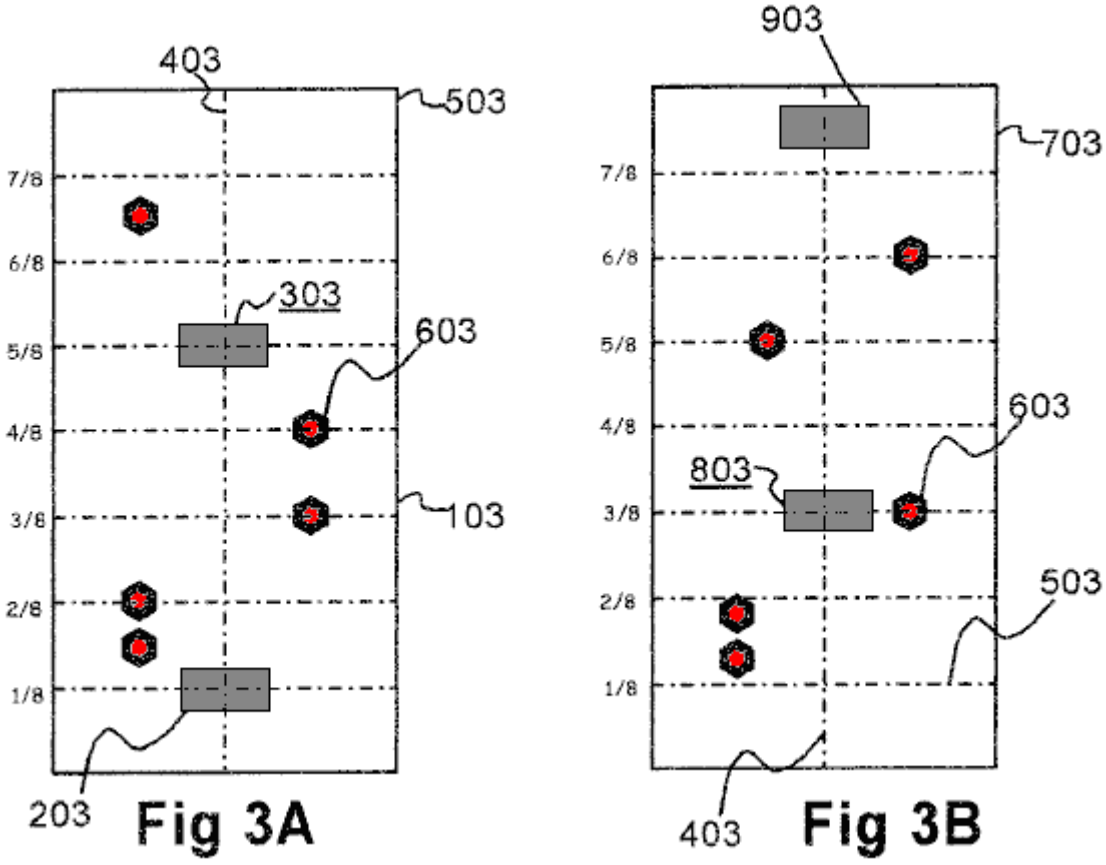


Fig.3A and **Fig.3B** represent a pattern for mounting the slave magnets on the outside, or convex, surface of one pair of cowlings. The figure represents the cowling-halves **103**, **703** as if they were laid flat. Guide lines are provided in the figure to indicate the longitudinal bisecting lines **403** and horizontal lines **503** dividing each cowling into eighths.

With respect to the cowling-half shown in **Fig.3A,** two permanent drive magnets, **203**, **303** are glued to the outside surface of the cowling on line **403** bisecting the cowling longitudinally. One drive magnet **203** is placed approximately one eighth of the way from one end. The second drive magnet **303** is placed three eighths of the way from the opposite end. Ferro-magnetic bolts **603** are inserted in the cowling through threaded holes. The purpose of the bolts is to modify the magnetic field to eliminate dead spots.

With respect to the cowling-half shown in **Fig.3B**, drive magnet **803** is placed three eighths of the way from one end, and drive magnet **903** is placed one eighth (one sixteenth ?) of the way from the other end. Again, ferro-magnetic bolts **603** are provided for eliminating dead spots in the magnetic field.

The diameter across each slave wheel is approximately 10 inches (250 mm). Measured from the bottom of groove **404** the diameter is 9 inches(225 mm). Consequently, the arc length from the bottom of one groove to the bottom of an adjacent groove is π inches (i.e., 3.14 inches or 80 mm).

The drive magnets are glued or otherwise firmly fixed to the outer or concave surfaces of the cowlings. Assuming that the slave magnets have been mounted in the grooves of the slave wheels with the north pole facing outwards, the north pole of each drive magnet is fixed against the cowling surface so that like poles face one another. As the cowlings are moved toward the slave wheels by turning the cranks **1101**, **1001** the drive magnets repel the slave magnets, causing the slave wheels to rotate.

Adjustment of the spacing between the cowlings and the slave wheels by means of cranks **1101**, **1001** adjusts the strength of the interaction of the fields of the drive magnets and slave magnets and, hence, the torque on the slave wheels.

As shown in **Fig.1**, fly-wheels **301**, **401** can optionally be mounted on the slave shaft. The preferred position is at or near the end of the shaft.

Slave shaft **201** thus turns as a result of the magnetic force from the cowlings being applied to the slave wheels. This shaft can be coupled to an output such as the armature of a generator **1301**, either directly or through a flywheel, as shown. Alternatively, the magnetic motor could itself drive a hydraulic pump of a transmission,

thereby reducing the number transmission components and the overall complexity of transmissions. Many different applications for this motor become obvious once it is realised that by using very strong permanent drive magnets useful power can be generated.

It is possible to vary the dimensions of the slave wheels. Presently, the preferred diameter is approximately 10 inches and a width of 5 inches. The motor can operate with the slave shaft **201** vertical or horizontal. While aluminium is a suitable material for the motor, the use of a hard plastic or ceramic materials have also been used with success. Pheotic plastic is presently preferred.

By using two slave wheels rather than just one, any dead spots in one wheel will be compensated for by the other wheel. The upper limit of the number of slave wheels is not yet known. The lower limit is one.

Patrick Kelly
www.free-energy-info.tuks.nl
www.free-energy-info.com
www.free-energy-info.co.uk

Chapter 2: Moving Pulsed Systems

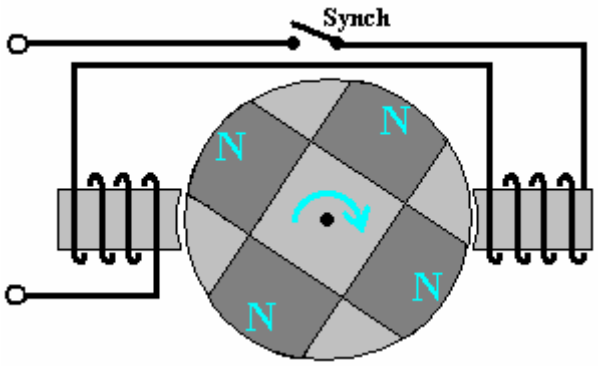
Note: If you are not at all familiar with basic electronics, you might find it easier to understand this chapter if you read chapter 12 first.

There are three categories of pulsed system and we will consider each in turn. These are drive-pulsed systems, energy-tapping pulsed systems and gravity free-energy pulsing systems. Here we will look at systems where an electrical pulse is used to cause the device to operate by creating a temporary magnetic field caused by electric current flowing through a coil or “electromagnet” as it is often called. Many of these systems are rather subtle in the way that they operate. One very well-known example of this is

The Motor/Generator of Robert Adams.

The late Robert Adams, an electrical engineer of New Zealand designed and built several varieties of electric motor using permanent magnets on the rotor and pulsed electromagnets on the frame of the motor (called the "stator" because it does not move). He found that if they were configured correctly, then the output from his motors exceeded their input power by a large margin (800%).

The diagram of his motor intended to show the basic operating principle is shown here:



If a motor is built like this, then it will most certainly work but it will never reach 100% efficiency let alone exceeding the 100% mark. It is only with a specific configuration which is hardly ever publicised that high performance figures can be achieved. While Robert has shown several different configurations, in order to avoid confusion I will describe and explain just one of them. I am indebted to several of Robert's friends and colleagues for the following information and I should like to express my thanks to them for their help and support in bringing you this information.

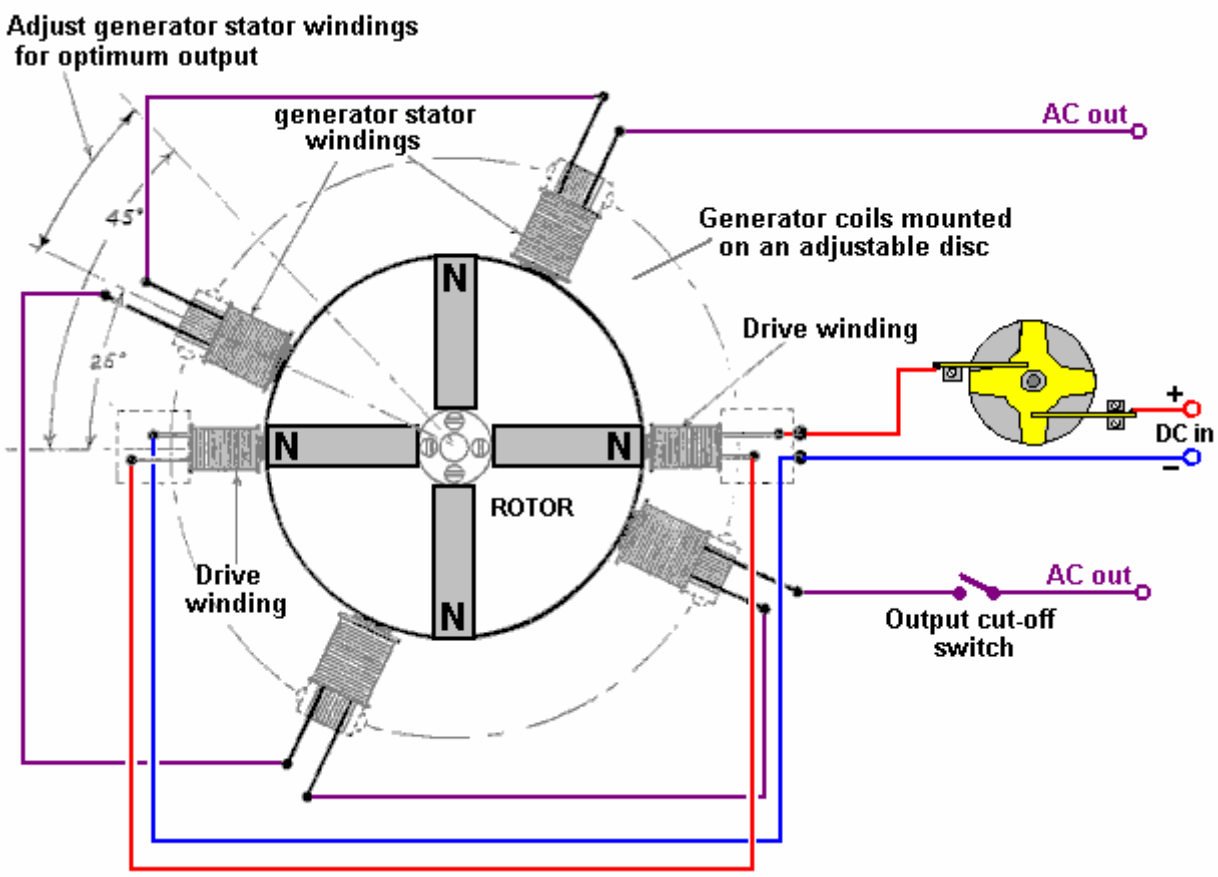
First and foremost, high performance can only be achieved with the clever use of power collection coils. These coils need to be positioned accurately and their power collection restricted to just a very short arc of operation by connecting them to, and disconnecting them from, the output circuit at just the right instant so that the back EMF generated when the current draw stops, actually contributes to the drive of the rotor, speeding it on it's way and raising the overall efficiency of the motor/generator as a whole.

Next, the shape of the magnets used is important as the length to width proportion of the magnet alters the pattern of it's magnetic fields. In direct opposition to the diagram shown above, the magnets need to be much longer than their width (or in the case of cylindrical magnets, much longer than their diameter).

Further, a good deal of experimentation has shown that the size and shape of the electromagnets and pick-up coils has a major influence on the performance. The cross-sectional area of the core of the pick-up coils should be four times that of the cross-sectional area of the permanent magnets in the rotor. The reverse is true for the cores of the drive coils as their cores should have a cross-sectional area of just one quarter of the rotor magnet cross-sectional area.

Another point which is almost never mentioned is the fact that big circuit gains will not be achieved unless the drive voltage is high. The minimum should be 48 volts but the higher the voltage, the greater the energy gain, so voltages in the 120 volts (rectified US mains voltage) to 230 volts (rectified mains voltage elsewhere) should be considered. Neodymium magnets are not recommended for drive voltages under 120 volts.

This is one of Robert's test circuits:

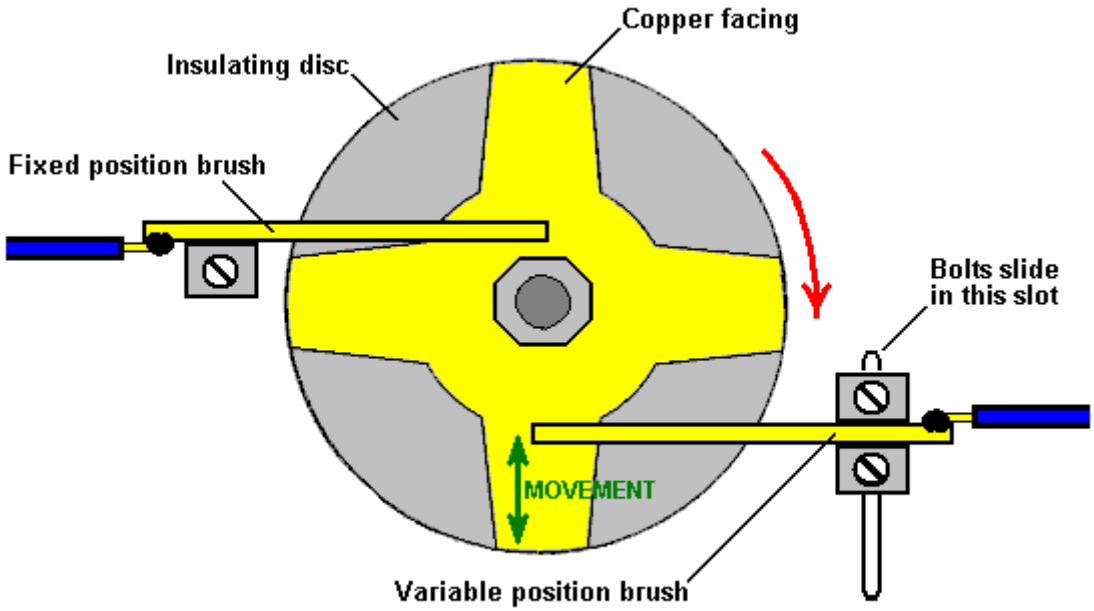


Notice that the cores of the "generator" pick-up coils are very much wider than the cores of the drive coils. Also notice the proportions of the magnets where the length is much greater than the width or diameter. The four generator windings are mounted on a single disc allowing them to be moved through an angle to find the optimum operating position before being locked in position and the two drive coils are mounted separately and held clear of the disc. Notice also that the power pick-up coils are much wider compared to their length than the drive coils are. This is a practical feature which is explained in greater detail later.

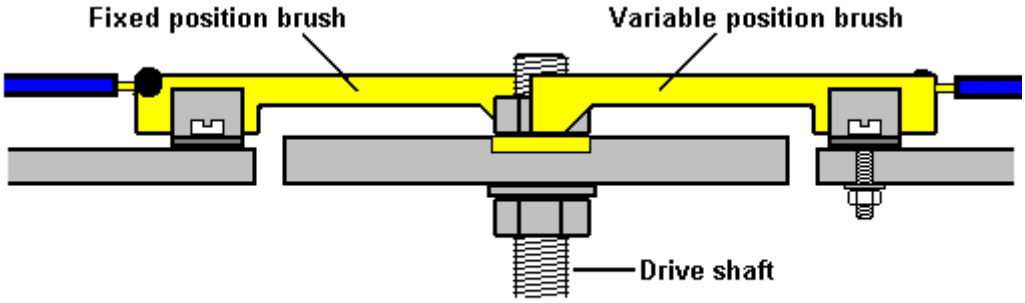
The DC input is shown passing through Robert's custom-made contactor switch which is mounted directly on the shaft of the motor/generator. This is a mechanical switch which allows an adjustable On / Off ratio, which is known as the "Mark/Space Ratio" or, if the "On" period is of particular interest, the "Duty Cycle". Robert Adams indicates that when the motor is running and has been adjusted to it's optimum performance, then the Mark/Space ratio should be adjusted to minimise the On period and ideally get it down to about 25% so that for three quarters of the time, the input power is actually switched off. There are various ways of achieving this switching while still having a very sharp turn on and turn off of the power.

Robert considered mechanical switching of the drive current to be a very good option although he was not opposed to using the contact to power a transistor to do the actual switching and so reduce the current through the mechanical contacts by a major factor. His reasons for his preference for mechanical switching are that it gives very sharp switching, needs no electrical power to make it operate and it allows current to flow in both

directions. The current flow in two directions is important because Robert produced various ways of getting the motor to feed current back into the driving battery, allowing it to drive the motor for long periods without lowering its voltage hardly at all. His preferred method of switching is shown here:



TIMING GEAR - TOP VIEW



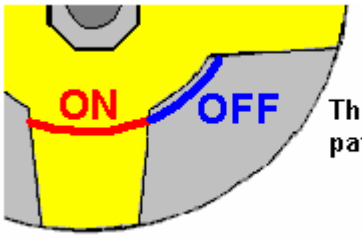
TIMING GEAR - SIDE VIEW

This switching gear operates as follows: The timing disk is bolted securely to the drive shaft of the motor and its position is set so that the electrical switch-on occurs when the rotor magnet is exactly aligned with the drive coil core. Adjustment of that timing is done by loosening the locking nut, rotating the disc very slightly and clamping the disc in position again. A spring washer is used to keep the assembly tight when the device is running. The disc has a star-shaped piece of copper sheet set into its surface and two silver-tipped, copper arm "brushes" slide across the surface of the copper star.

One of these two brushes is fixed in position and slides across the copper star near the drive shaft, making a permanent electrical connection to it. The second brush slides alternatively on the non-conducting surface of the disc and then over the conducting arm of the copper. The second brush is mounted so that its position can be adjusted and, because the copper arms taper, that alters the ratio of the "On" time to the "Off" time. The actual switching is achieved by current flowing through the first brush, through the copper arm and then through the second brush. The brush arms shown in the diagram above rely on the springiness of the copper arm to make a good brush-to-copper electrical connection. It might be preferred to use a rigid brush arm, pivot it and use a spring to ensure a very good contact between the brush and the copper star at all times.

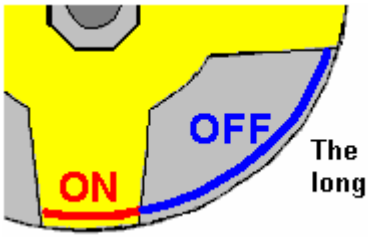
The adjustment of the On to Off time, or "Mark/Space Ratio" or "Duty Cycle" as the technical people describe it, could perhaps do with some description. If the moveable brush is positioned near the centre of the disc, then,

because of the tapering of the copper arms, the part of the non-conducting disc that it slides over is shorter and the part of the conducting copper arm with which it connects is longer, as the two sliding paths are about the same length, the current is on for about the same length as it is off, giving a Mark/Space ratio of about 50% as shown here:



The On path is about the same length as the Off path and so the Mark/Space ratio is about 50%

If, instead, the moveable brush is positioned near the outside edge of the disc, then because of the tapering of the copper arm, the On path is shorter and the non-conducting Off path is very much longer, being about three times as long as the On path, giving a Mark/Space ratio of about 25%. As the moveable brush can be positioned anywhere between these two extremes, the Mark/Space ratio can be set to any value from 25% to 50%.

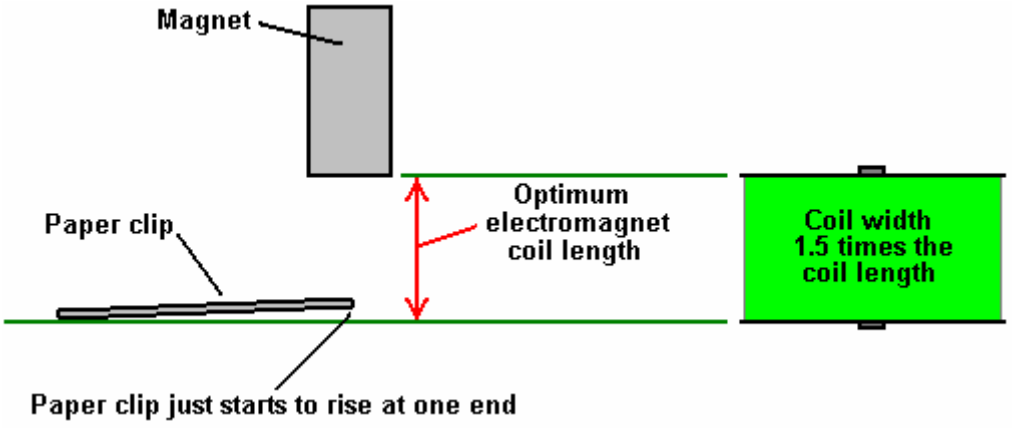


The On path is shorter and the Off path is much longer giving a Mark/Space ratio of about 25%

The two brushes can be on the same side of the drive shaft or on opposite sides as shown. One important feature is that the brushes touch in a position where the disc surface is always moving directly away from the brush mounting, causing any drag to be directly along the arm and giving no sideways loading on the brush. The diameter of the device is usually one inch (25 mm) or less.

You will also notice that the output is switched although the diagram does not give any indication of how or when that switching takes place. You will notice that the diagram has angles marked on it for the optimum positioning of the pick-up coils, well, an Adams Motor builder with a forum ID of "Maimariati" who achieved a Coefficient Of Performance of 1,223, found that the optimum switching for his motor is On at 42 degrees and Off at 44.7 degrees. That tiny 2.7 degree part of the rotor turn gives a substantial power output and cutting the output current off at that point causes the back EMF of the coils to give the rotor a substantial additional boost on its way. His input power is 27.6 watts and the output power is 33.78 kilowatts

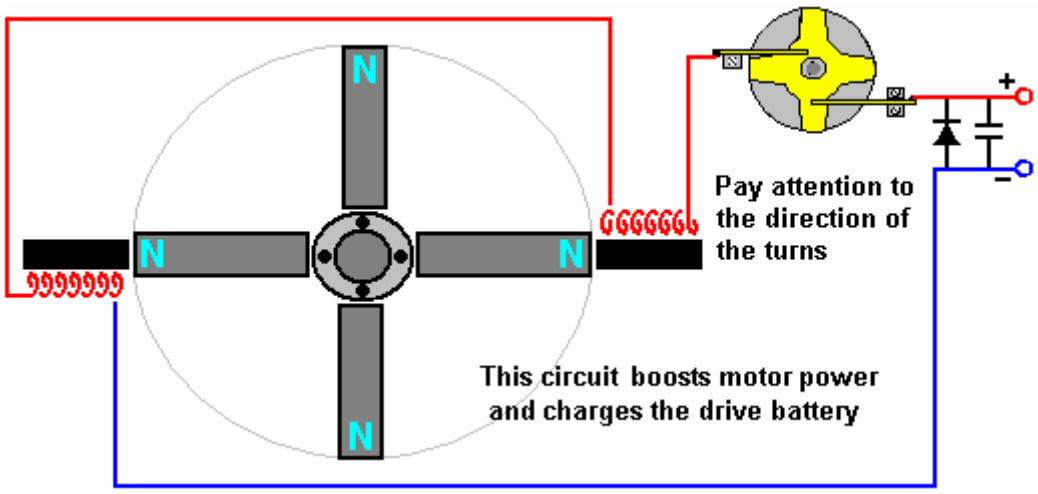
Now for some practical details. It is suggested that a good length for the power pick-up coils can be determined by using the “paper clip test”. This is done by taking one of the permanent magnets used in the rotor, and measuring the distance at which that magnet just begins to lift one end of a 32 mm (1.25 inch) paper clip off the table. The optimum length of each coil from end to end is exactly the same as the distance at which the paper clip starts to lift.



The core material used in the electromagnets can be of various different types including advanced materials and alloys such as ‘Somalloy’ or 'Metglas'. The power pick-up coil proportions are important as an electromagnet becomes less and less effective as its length increases, and eventually, the part furthest from the active end can actually be a hindrance to the effective operation. A good coil shape is one which you would not expect, with the coil width being, perhaps 50% greater than the coil length:

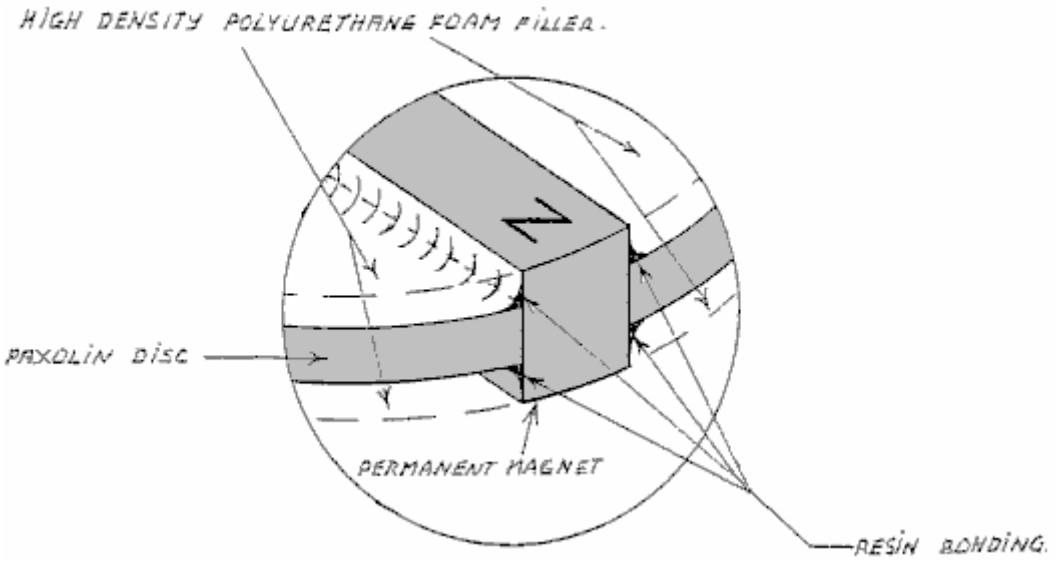
Contrary to what you would expect, the device draws in energy from the local environment better if the end of the pick-up coil farthest from the rotor is left unaffected by any other part of the device and the same applies to the magnet facing it. That is, the coil should have the rotor at one end and nothing at the other end, that is, no second rotor behind the coil. The speed at which the voltage is applied to, and removed from, the coils is very important. With very sharp voltage rises and falls, additional energy is drawn from the surrounding environmental energy field. If using transistor switching, then the IRF3205 FET has been found to be very good and a suitable driver for the FET is the MC34151.

If using a Hall-effect semiconductor to synchronise the timing, say the UGN3503U which is very reliable, then the life of the Hall-effect device is much improved if it is provided with a 470 ohm resistor between it and the positive supply line, and a similar 470 ohm resistor between it and the negative line. These resistors in series with the Hall-effect device effectively “float” it and protect it from supply-line spikes".



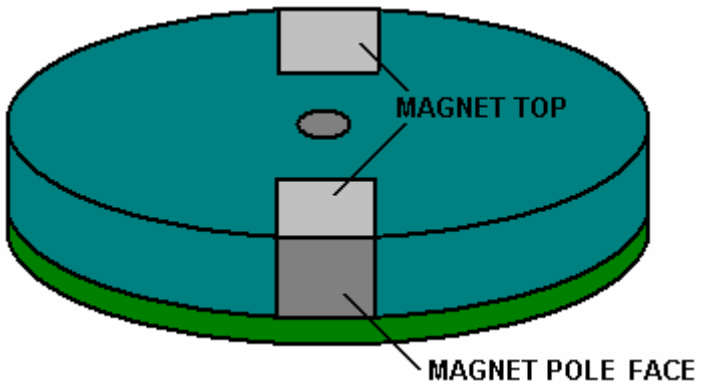
Here, two electromagnets are driven by the battery via Robert's 4-arm commutator which is mounted on the rotor shaft. Some of the recommendations given by Robert are the opposite of what you would expect. For example, he says that a single rotor construction tends to be more electrically efficient that one where several rotors are mounted on a single shaft. Robert is against the use of reed switches and he recommends making one of his commutators.

At one stage, Robert recommended the use of standard transformer shims for constructing the cores of the electromagnets. This has the advantage that matching bobbins for holding the coil windings are readily available and can still be used for pick-up coils. Later on, Robert swung towards the use of solid cores from the old PO Series 3000 telephone relays and eventually said that electromagnet cores should be solid iron.

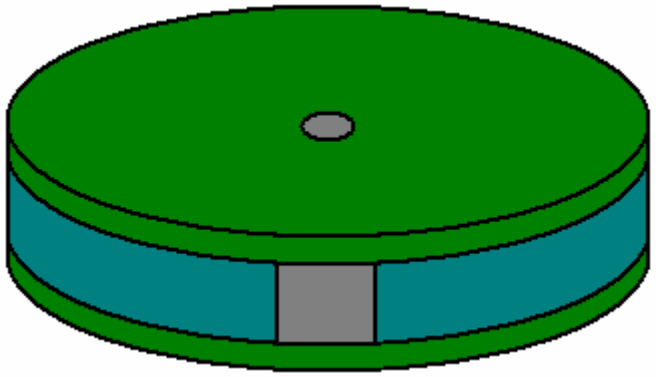


The diagrams presented by Robert show the magnets located on the rim of the rotor and pointing outwards. If this is done, then it is essential that the magnets in the rotor are firmly attached on at least five of their six faces and the possibility of using a ring of non magnetic material such as duct tape around the outside should be considered. That style of construction also lends itself to streamlining the rotor by having a completely solid construction, although it might be remarked that the motor would run better and more quietly if it were enclosed in a box which had the air pumped out of it. If that is done, then there will be no air resistance and because sound can't pass through a vacuum, quieter operation is bound to result.

While this may sound a bit complicated, there is no reason why it should be. All that is needed is two discs and one central disc which is the thickness of the magnets, with slots cut in it, the exact size of the magnets. The assembly starts with the lower disc, magnets and central disc. These are glued together, probably with epoxy resin, and that holds the magnets securely on four faces as shown here:



Here, the magnets are attached on the lower face, the right and left faces, and the unused pole face, and when the upper disc is attached, the upper faces are also secured and there is the minimum of air turbulence when the rotor spins:

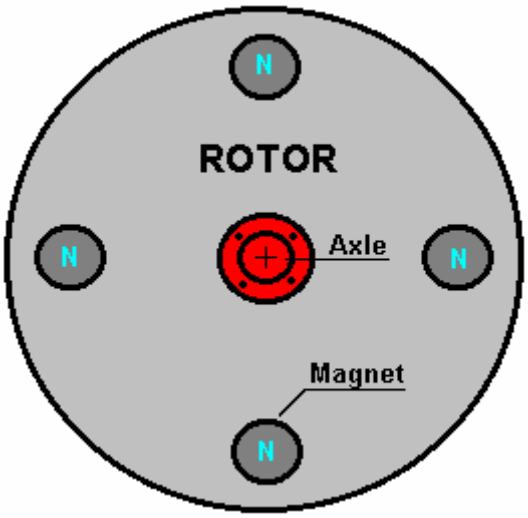


There is a "sweet spot" for the positioning of the power pick-up coils and it will usually be found that this is two or three millimetres away from the rotor. If that is the case, then there will be room for an outer band of duct tape on the rim of the rotor to provide additional protection against the failure of the magnet attachment method.

High-power versions of the motor/generator need to be enclosed in a metal box which is earthed as they are quite capable of generating a substantial amount of high frequency waves which can damage equipment such as oscilloscopes and create TV reception interference. There would probably be an improvement in performance as well as a reduction in sound if the box was airtight and had the air pumped out of it. If that is done, then there will be no air resistance as the rotor spins and since sound does not pass through a vacuum, quieter operation is possible.

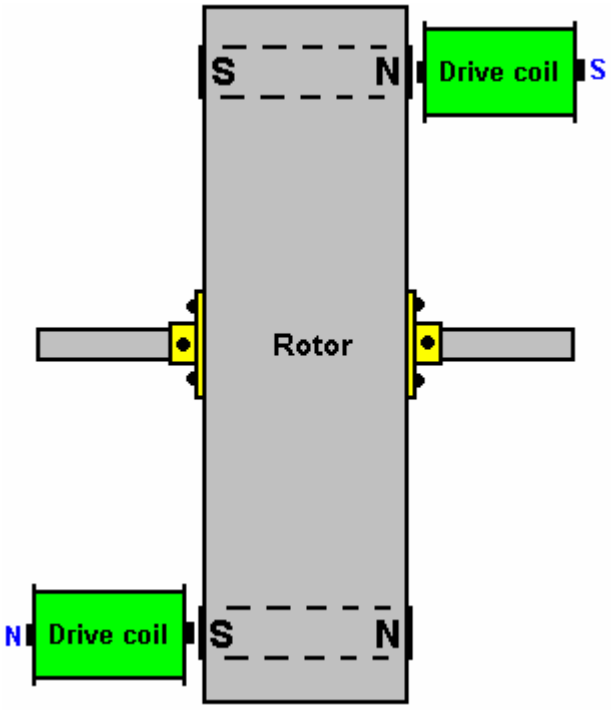
Experienced rotor builders do not like the radial magnets style of construction because of the stresses on the magnet attachments if high rotational speeds are reached. It should not need to be said, but it is obviously a major requirement to keep your hands well away from the rotor when the motor is running as it is perfectly possible to be injured by the high-speed movement if you are careless. Please remember that this presentation must not be considered to be a recommendation that you build or use any device of this nature and it must be stressed that this text, in common with the entire contents of this eBook, is intended to be for information purposes only and no representations or warranties are implied by this presentation. Should you decide to construct, test or use any device, then you do so entirely at your own risk and no liability attaches to anybody else if you sustain any kind of injury or property damage as a result of your own actions.

Because of the mechanical stresses caused during rotation, some experienced constructors feel that the magnets should be embedded in the rotor as shown here where they are kept well clear of the rim of a rotor which is made from a tough material. This is so that the outer strip of the material prevents the magnets breaking loose and becoming dangerous high-speed projectiles, which at best would destroy the electromagnets and at worst could injure someone quite badly:



It needs to be remembered that the proportions of the magnets are for the magnet length to be more than the diameter, so in cases like this where circular magnet faces are to be used, the magnets will be cylindrical and the rotor needs to have a significant thickness, which will depend on the magnets which are available locally. The magnets should be a tight push-fit in their holes and securely glued in place.

Robert Adams has used this construction style as well. However, if an arrangement like this is used, then there will be a substantial sideways pull on the rotor as it reaches the electromagnet core, tending to pull the magnets out of the rotor.



It is important that the rotor should be perfectly balanced and have the minimum amount of bearing friction possible. This calls for precision construction and either roller or ball bearings. The construction style shown above has the distinct advantage that it has an open end to both the magnet and the coils and this is believed to facilitate the inflow of environmental energy into the device.

It may be my ignorance showing here, but I have a problem with this version. The difficulty as I see it is that the magnet/core pull and the subsequent drive thrust when the coil is powered, form a "turning couple" as they both try to rotate the axle in the same direction. This places a substantial loading on the axle bearings, usually amplified by the radius of the rotor being greater than the distance from the rotor to the axle bearings. This load will be in the tens of kilograms range and will be applied and reversed perhaps forty times per second. To me, that appears like a vibration load and is directly opposed to the "perfectly balanced" rotor operation being sought. The radial magnet arrangement generally shown by Robert Adams does not have any of this kind of loading at all because the coils are exactly opposite each other and their loads cancel each other out exactly. The choice is, of course, up to the builder and his assessment of the advantages and disadvantages of the different styles of construction.

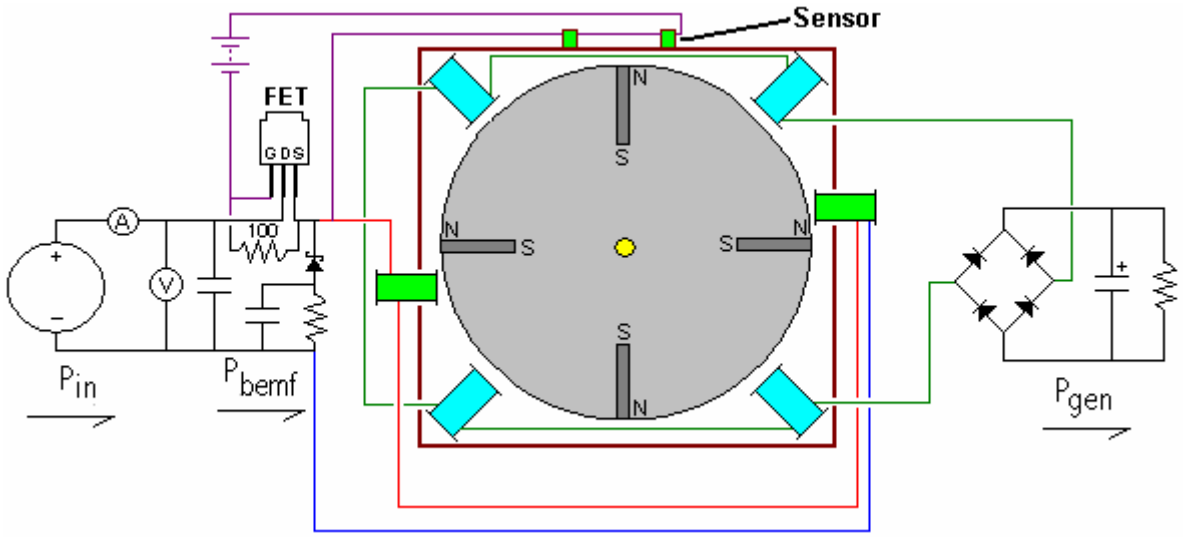
When getting ball-race bearings for an application like this, please be aware that "closed" bearings such as these are not suitable as supplied:

This is because this type of bearing is usually packed with dense grease which completely destroys its free motion, making it worse as a bearing than a simple hole-and-shaft arrangement. However, in spite of this, the

closed or "sealed" bearing is popular as the magnets tend to attract dirt and dust and if the device is not enclosed in a steel box as is necessary for the high power versions, then having the seal is considered to be an advantage. The way to deal with the grease packing is to soak the bearing in an isopropyal solvent cleaner to remove the manufacturer's grease, and then, when it has dried out, lubricate the bearing with two drops of a high quality thin oil. If it is intended to house the motor/generator in an earthed, sealed steel box then an alternative type of bearing which might be suitable is an open design like this:

especially if the air is removed from the box. Some constructors prefer to use ceramic bearings which are supposed to be immune to dirt. One supplier is http://www.bocabearings.com/main1.aspx?p=docs&id=16 but as with everything else, these choices have to be made by the builder and will be influenced by his opinions.

I'm not sure where it came from, but here is a circuit diagram showing a transistor drive and the return of the back EMF of the drive coils to the driving power supply. Using this method, about 95% of the drive current can be returned, lowering the current draw enormously:



The diode feeding the power back to the supply is a Shotky type because of it's high-speed operation. It needs to be able to handle the peak pulse power and so should be one of the more robust types. What this circuit does not have is the very important switching on the output coils circuit. Another strange item is the way that the FET sensor is arranged with two sensors rather than one and with an additional battery. While it must be admitted that the current draw of the FET gate should be very low, there still does not seem to be much reason to have a second power supply. One other peculiarity in this diagram is the positioning of the drive coils. With them offset as shown, it has the effect of them being at an angle relative to the rotor magnets. It is not at all clear if this is an advanced operating technique or just poor drawing - I am inclined to assume the latter although I have no evidence for this other than the circuit design and the low quality of the original drawing which had to be improved considerably to arrive at the diagram shown above.

The coil generator output should be fed into a capacitor before being passed to whatever equipment is to be powered by the device. This is because the energy is being drawn from the local environment and is not conventional energy. Storing it in a capacitor converts it to a more normal version of electrical power, a feature which has also been mentioned by Don Smith and by John Bedini although their devices are quite different in operation.

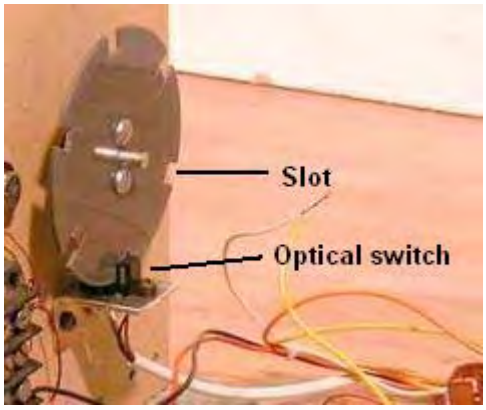
The DC resistance of the coil windings is an important factor. The overall resistance should be either 36 ohms or 72 ohms for a complete set of coils, whether they are drive coils or power pick-up coils. Coils can be wired in parallel or in series or in series/parallel. So, for 72 ohms with four coils, the DC resistance of each coil could be 18 ohms for series-connected, 288 ohms for parallel connected, or 72 ohms for connection in series/parallel where two pairs of coils in series are then wired in parallel.

To help with assessing the wire diameter and length which you could use, here is a table of some of the common sizes in both American Wire Gage and Standard Wire Gauge:

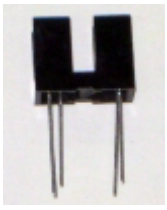
AWG	Dia mm	SWG	Dia mm	Max Amps	Ohms / 100 m
11	2.30	13	2.34	12	0.47
12	2.05	14	2.03	9.3	0.67
13	1.83	15	1.83	7.4	0.85
14	1.63	16	1.63	5.9	1.07
15	1.45	17	1.42	4.7	1.35
16	1.29	18	1.219	3.7	1.48
18	1.024	19	1.016	2.3	2.04
19	0.912	20	0.914	1.8	2.6
20	0.812	21	0.813	1.5	3.5
21	0.723	22	0.711	1.2	4.3
22	0.644	23	0.610	0.92	5.6
23	0.573	24	0.559	0.729	7.0
24	0.511	25	0.508	0.577	8.7
25	0.455	26	0.457	0.457	10.5
26	0.405	27	0.417	0.361	13.0
27	0.361	28	0.376	0.288	15.5
28	0.321	30	0.315	0.226	22.1
29	0.286	32	0.274	0.182	29.2
30	0.255	33	0.254	0.142	34.7
31	0.226	34	0.234	0.113	40.2
32	0.203	36	0.193	0.091	58.9
33	0.180	37	0.173	0.072	76.7
34	0.160	38	0.152	0.056	94.5
35	0.142	39	0.132	0.044	121.2

So far, we have not discussed the generation of the timing pulses. A popular choice for a timing system is to use a slotted disc mounted on the rotor axle and sensing the slots with an "optical" switch. The "optical" part of the switch is usually performed by UV transmission and reception and as ultra violet is not visible to the human eye, describing the switching mechanism as "optical" is not really correct. The actual sensing mechanism is very simple as commercial devices are readily available for performing the task. The sensor housing contains both a UV LED to create the transmission beam, and a UV dependent resistor to detect that transmitted beam.

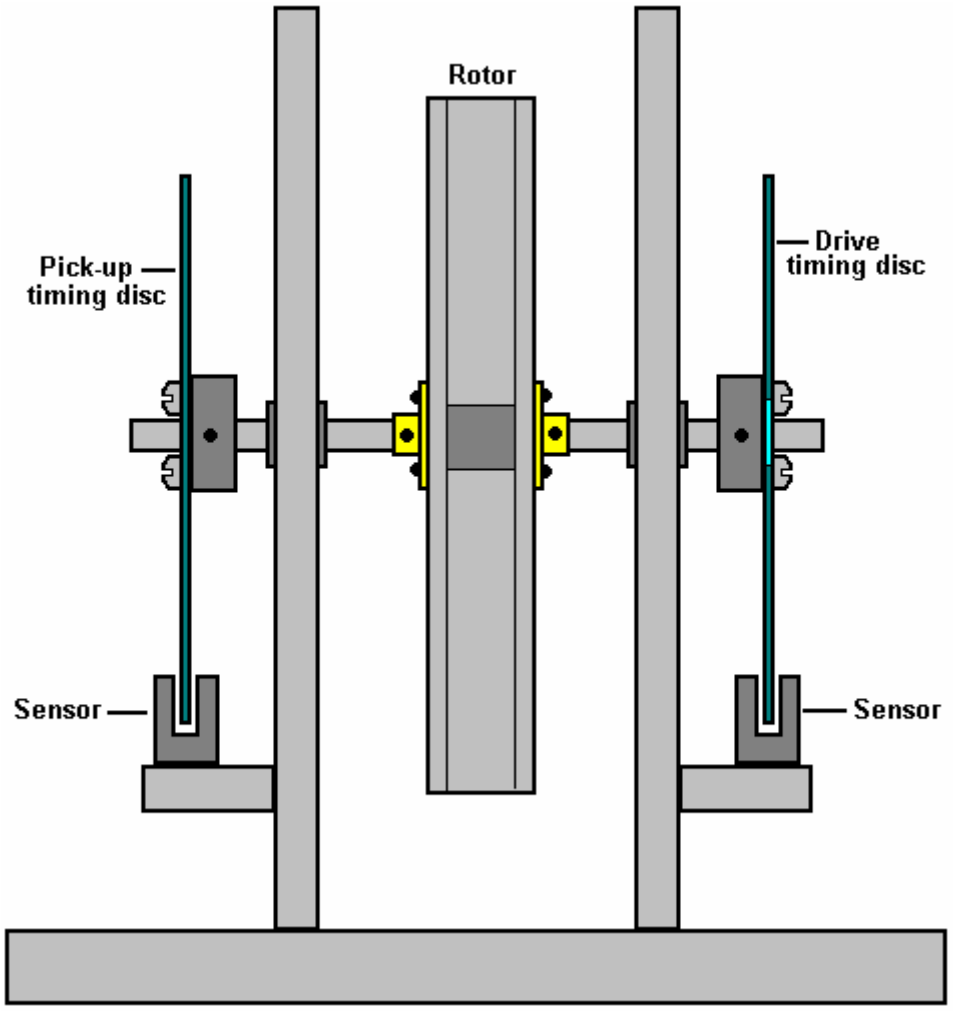
Here is an example of a neatly constructed timing mechanism made by Ron Pugh for his six-magnet rotor assembly:



and the switch/sensor:



This device happens to be one which is supplied by www.bayareaamusements.com under their product code number : OP-5490-14327-00. As the slotted disc rotates, one of the slots comes opposite the sensor and allows the UV beam to pass through to the sensor. That lowers the resistance of the sensor device and that change is then used to trigger the drive pulse for whatever length of time the slot leaves the sensor clear. You will notice the balanced attachment method used by Ron to avoid having an unbalanced rotor assembly. There can be two timing discs, one for the drive pulses and one for switching the power pickup coils in and out of the circuit. The slots in the power pick-up timing disk will be very narrow as the switch-on period is only about 2.7 degrees. For a six-inch diameter disc where 360 degrees represents a circumference length of 18.85 inches (478.78 mm) a 2.7 degree slot would be only 9/64 inch (3.6 mm) wide. The arrangement for an axial magnet rotor set-up could be like this:



So to recap, the things which are necessary for getting an Adams Motor output into the serious bracket are:

1. A performance of COP>1 can only be achieved if there are power pick-up coils.

2. The rotor magnets need to be longer than they are wide in order to ensure the correct magnetic field shape and the rotor must be perfectly balanced and have bearings as low-friction as possible.

3. The face area of the rotor magnets needs to be four times that of the drive coil cores and one quarter the area of the core of the power pick-up coils. This means that if they are circular, then the drive coil core diameter needs to be half the diameter of the magnet and the magnet diameter needs to be half the diameter of the power pick-up core. For example, if a circular rotor magnet is 10 mm across, then the drive core should be 5 mm across and the pick-up core 20 mm across.

4. The drive voltage needs to be a minimum of 48 volts and preferably, a good deal higher than that.

5. Do not use neodymium magnets if the drive voltage is less than 120 volts.

6. The drive coils should not be pulsed until they are exactly aligned with the rotor magnets even though this does not give the fastest rotor speed.

7. Each complete set of coils should have a DC resistance of either 36 ohms or 72 ohms and definitely 72 ohms if the drive voltage is 120 volts or higher.

8. Collect the output power in large capacitors before using it to power equipment.

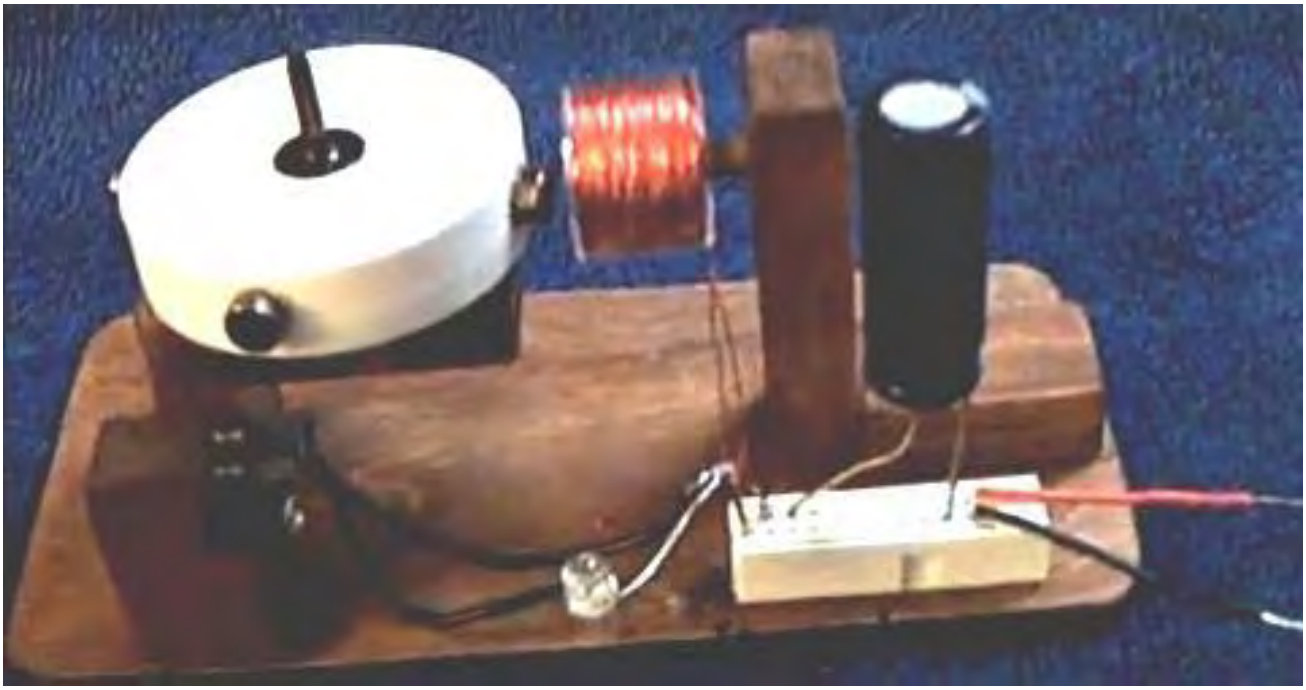
It may also be possible to boost the output power further, by using the Coil-Shorting technique shown in the section of this chapter on the RotoVerter.

If you want the original drawings and some explanation on the operation of the motor, then two publications from the late Robert Adams can be bought from www.nexusmagazine.com where the prices are quoted in Australian dollars, making the books look much more expensive than they actually are.

http://www.totallyamped.net/adams/index.html is a really impressive collection of well-informed practical material on building and using an Adams motor with details of sensors and how they work, core materials and their performances and how to locate the "sweet spot" - very highly recommended web site.

Lidmotor's Low-voltage Rotor

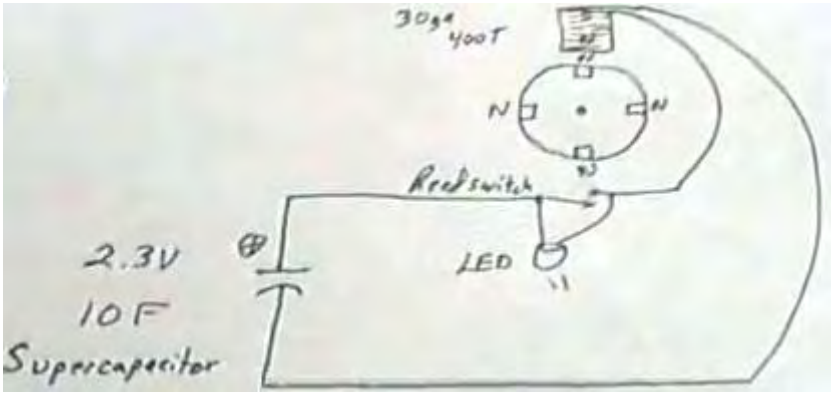
One very experienced developer whose YouTube ID is "Lidmotor" (because he makes motors from the lids of jars) has a short video at https://www.youtube.com/watch?v=SjWCprVXer8 showing a very simple rotor design with one of his jar lids mounted on a single bearing:



The white lid has four magnets attached to it spaced out evenly around the lid at ninety-degree intervals. Facing them is an air-core coil mounted on a non-magnetic support dowel and wound with 400 turns of '30-gauge' wire on a plastic spool. As 'Lidmotor' is American, the '30 gauge' wire is likely to be American Wire Gauge #30 with a diameter of 0.255 mm as opposed to the European Standard Wire Gauge size which has a 0.315 mm diameter.

An air-core coil has no effect on the passing rotor magnets IF it is not carrying current. An output coil will cause drag on the rotor if current is being drawn from the coil, and so timed output switching as used by Robert Adams would be needed to not just overcome the drag, but to push the rotor on its way as well.

There are two very important features of this rotor drive design. One is the fact that a supercap (10-Farad, 2.3V) is used to drive the rotor and when supplied with a charge of only 0.5V to 1.0V, can spin the rotor for up to thirty minutes. That very long time is likely to be a feature of the second important item which is that he has placed an LED across the reed switch used to power the coil. When the reed switch opens, a back-EMF voltage spike occurs and the LED feeds that voltage pulse back into the 100% efficient supercap, recovering most of the current used to drive the rotor. This is the same method as used by Robert Adams in his motor designs. Lidmotor presents his circuit like this:



While the circuit shows the magnet operating the reed switch as being 180 degrees around from the coil, the photograph indicates that the switching magnet is one 90 degrees away. Any of the other magnets can be used. The reed switch operation is adjusted to get the best performance. This is done by moving the switch backwards and forwards along the moving path of the magnet to make the switching occur earlier or later. The objective is to push the rotor magnet on its way by pulsing the coil very briefly just after the rotor magnet has passed the centre of the coil. The length of time that the reed switch is closed can be adjusted by moving the switch closer to the magnet for a longer switch-closed time, or further away for a shorter switch closure. It is also possible to alter the closed time by positioning the switch across the path of the magnet travel or parallel to it.

If you are not familiar with a reed switch, it is just a glass tube, filled with an inert gas, and with two overlapping metal strips inside the tube:

Reed switch

The external magnetic field magnetises the strips and they spring together due to magnetic attraction and spring apart again when the magnetic field moves away. These switches come in various sizes and the smallest version tends to be unreliable and has a very low maximum current capacity. The larger versions are much more robust.

Lidmotor's circuit is very simple and very effective, even though the rotor will have minimal weight and no significant drag. One wonders if adding a second coil and a diode feeding the supercapacitor, if the system could not become self-running.

Phemax's Inertial-Propulsion and Electrical Generation System.

Phemax Technologies, Inc. have developed what they describe as their 'sustainable transportation system' which is based on their inertial-propulsion generation wheel, which uses what they call '3-D Coraxial Hybrid Induction' (CHI), in which 'Coraxial' refers to their "combined radial-axial" technique where drive is provided by a radial pulsed magnet system and power extraction is achieved using an axially-orientated magnet/coil system.

Vice President and Chief Technology Officer, Michael Hsueh and Sales Manager, Sabrina Li

Taiwan inventor, Tajen (David) Chi, says that his company, "Phemax Technologies Inc." will be demonstrating their technology

in September 2010 at an expo, after which they will allow interested and qualified parties to come to their premises for licensing discussions.

Their device can take the place of the battery component of an electric vehicle or it can be used to supplement and recharge a battery bank in the vehicle. At the September exhibition, they will allow visitors to ride their 2 kW test car. David intends to put their test data on the Internet. With a rotational input of 1 NM at 500 rpm, a single inertial-propulsion generation wheel can produce 1 KW of axial generative power.

David also said that a 150W output from his motor actually produces between 180 and 200 watts of mechanical output, while a 1500W output produces a mechanical output of 1800 to 2000 watts (96 Volts at 20 Amps), as measured by a watt meter, speed meter, torque meter, and oscilloscopes. He says that these measurements were made by Michael Hseuh, Vice President and Chief Technology Officer.

Ceramic-based and carbon-based ultra-capacitors are used to return some of the output power to the input in order to keep the machine running continuously without the need for a power supply when the engine is running and providing it's output. When driving a vehicle powered by this device, the mechanical driving torque is provided via a continuously-variable transmission.

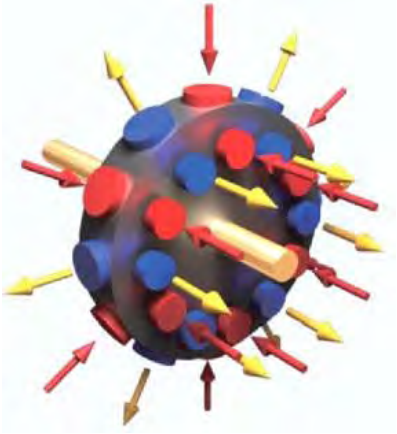
At this time, Phemax Technologies, Inc. have a 150-Watt and a 2 kW prototype which they plan to demonstrate in September 2010. He says that they generally run these prototypes for eight hours each day in their lab. The central principle of the Phemax Technologies Inc. process is what they call "CORAXIA" which stands for "combined radial-axial" hybrid induction, abbreviated to "CHI", which is David's surname. They also call this the "3-D Coaxial Hybrid Induction" system. They state that their electromagnetic 3-D arrangement enables the wheel to have both radial flux propulsion and axial flux generation with just a single rotor and two separate stators. As the radial electric drive rotates, the mechanical energy due to the inertia of the wheel and the ground power transmission mechanism enables the wheels of a vehicle to generate electricity as they rotate.

There are two videos showing the device generating electricity, located here:
http://www.youtube.com/watch?v=W_IzhpZxxcQ and
http://www.youtube.com/watch?v=O8frdR-fnO here.

Another application is where the electrical output from the device is being used to produce a hydrogen/oxygen gas mix from water using an underwater plasma arc:

David says he is self-taught in this area. He has three patents: Taiwan patent M352472, M372891 and M382655 (which are not in English). From his search of the patent index, the Internet and YouTube, David said he has not yet seen a patent or experimental set-up similar to what they have developed.

This motor/generator is unusual in that it uses pulsed rotor drive on magnets mounted on the circumference of the rotor and simultaneously, it picks up electrical energy from a coil/magnet arrangement mounted on the side faces of the rotor as shown here:



The basic drive/generator unit can be replicated on a single shaft to give increased power, still without increased frictional losses between the drive and the power generation:

It is unusual to see this technique used as it can be difficult to avoid interaction between the different magnetic fields. However, David has been entirely successful in doing this and the arrangement has no transmission losses between the drive motor and the generator as these are integral components of the system.

This system is capable of driving a self-powered air-conditioning unit and a 5 kW proof-of-concept prototype is shown here:

This device is also capable of driving electric lighting and with a 5 kilowatt output self-powered output, it can power most household needs. A typical washing machine has a 2.25 kilowatt maximum electrical draw as does the average tumble-dryer. Most fan-heaters are 3 kilowatts or less when on full output.

If a wind-power generator is already in place, then using the mechanical power of the generator to spin the axle of one of these CHI generators results in a very high-power electricity generation system.

Contact: Tajen (David) Chi, Taiwan email: chitajen@gmail.com

Inquiries to: Phemax Technologies Inc., Michelle Chen, Manager of Marketing and Business Development.
Tel: 886-2-2371-5622 e-mail: michelle@phemax.com

Raymond Kromrey's Electrical Generator.

Where the objective is to produce electricity from a rotating magnetic field, there has always been a search for some method of either reducing, or eliminating altogether, the drag on the rotor when electric current is drawn from the generator. One design which claims to have very limited drag caused by current draw is the Kromrey design. The main characteristics of this design are said to be:

1. It has almost constant electrical power output even when the rotor speed is altered by as much as 35%.

2. It can continue to operate with it's electrical output short-circuited, without heating the rotor or causing a braking effect.

3. The production efficiency (electrical output divided by the driving force) is high.

4. The frequency of it's AC output power can be adjusted to that required by the equipment which it powers.

5. The rotor can be spun at any rate from 800 rpm to 1,600 rpm.

6. The simple construction allows manufacturing costs to be about 30% less than other generators.

7. This generator is recommended for supplying power at or above the 1 kilowatt level.

Here is the patent for this device:

Patent US 3,374,376 **19th March 1968** **Inventor: Raymond Kromrey**

ELECTRIC GENERATOR

My present invention relates to an electric generator which converts magnetic energy into electric energy using two components which can rotate relative to each other, i.e. a stator and a rotor, one having electromagnets or permanent magnets which induce a voltage in a winding which forms part of an output circuit mounted on the other component.

Conventional generators of this type use a winding which whose conductors form loops in different axial planes so that opposite parts of each loop pass through the field of each pole pair, twice per revolution. If the loops are open circuit, then no current flows in the winding and no reaction torque is developed, leaving the rotor free to turn at the maximum speed of its driving unit. As soon as the output winding is connected across a load or is short-circuited, the resulting current flow tends to retard the motion of the rotor to an extent which depends on the intensity of the current and this makes it necessary to include compensating speed-regulating devices if it is necessary to maintain a reasonably constant output voltage. Also, the variable reaction torque subjects the rotor and its transmission to considerable mechanical stresses and possible damage.

It is therefore the general object of this invention to provide an electric generator which has none of the above disadvantages. Another object is to provide a generator whose rotor speed varies very little in speed between open circuit operation and current delivery operation. Another objective is to provide a generator whose output voltage is not greatly affected by fluctuations in its rotor speed.

I have found that these objectives can be achieved by rotating an elongated ferromagnetic element, such as a bar-shaped soft-iron armature, and a pair of pole pieces which create an air gap containing a magnetic field. Each of the outer extremities of the armature carries a winding, ideally, these windings are connected in series, and these coils form part of a power output circuit used to drive a load. As the armature rotates relative to the air gap, the magnetic circuit is intermittently completed and the armature experiences periodic remagnetisations with successive reversals of polarity.

When the output circuit is open, the mechanical energy applied to the rotor (less a small amount needed to overcome the friction of the rotating shaft) is absorbed by the work of magnetisation, which in turn, is dissipated as heat. In actual practice however, the resulting rise in temperature of the armature is hardly noticeable, particularly if the armature is part of the continuously air-cooled rotor assembly. When the output circuit is closed, part of this work is converted into electrical energy as the current flow through the winding opposes the magnetising action of the field and increases the apparent magnetic reluctance of the armature, and so the speed of the generator remains substantially unchanged if the output circuit is open or closed.

As the armature approaches its position of alignment with the gap, the constant magnetic field tends to accelerate the rotation of the armature, aiding the applied driving force. After the armature passes through the gap there is a retarding effect. When the rotor picks up speed, the flywheel effect of its mass overcomes these fluctuations in the applied torque and a smooth rotation is experienced.

In a practical embodiment of this invention, the magnetic flux path includes two axially spaced magnetic fields traversing the rotor axis and substantially at right angles to it. These fields are generated by respective pole pairs co-operating with two axially spaced armatures of the type already described. It is convenient to arrange these two armatures so that they lie in a common axial plane and similarly, the two field-producing pole pairs also lie in a single plane. The armatures should be laminated to minimise eddy currents, so they are made of highly permeable (typically, soft-iron) foils whose principle dimension is perpendicular to the rotor axis. The foils can be held together by rivets or any other suitable method.

If the ferromagnetic elements are part of the rotor, then the output circuit will include the usual current-collecting means, such as slip-rings or commutator segments, depending on whether AC or DC current output is desired. The source of coercive force in the stator includes, advantageously, a pair of oppositely positioned, yoke-shaped magnets of the permanent or electrically energised type, whose extremities constitute the pole pieces mentioned above. If electromagnets are used in the magnetic circuit, then they may be energised by an external source or by direct current from the output circuit of the generator itself.

I have found that the terminal voltage of the output circuit does not vary proportionately to the rotor speed as might be expected, but instead, it drops at a considerably slower rate with decreasing rotor speed. So, in a particular tested unit, this voltage fell to only about half its original value when the rotor speed was dropped to one

third. This non-linear relationship between terminal voltage and driving rate produces a substantially constant load current and therefore, electric output over a wide speed range, at least under certain load conditions, inasmuch as the inductive reactance of the winding is proportional to frequency (and consequently, to rotor speed) so as to drop off more rapidly than the terminal voltage, in the event of a speed reduction, with a resulting improvement in the power factor of the load circuit.

If the magnetic circuit contains only a single pole pair per air gap, the flux induced in the rotating armature will change its direction twice per revolution so that each revolution produces one complete cycle of 360 electrical degrees. In general, the number of electrical degrees per revolution will equal 360 times the number of pole pairs, it being apparent that this number ought to be odd since with even numbers it would not be possible to have poles alternating in polarity along the path of the armature and at the same time to have the North and South poles of each pair at diametrically opposite locations. In any case, it is important to dimension the curved facing faces of the pole pairs in such a manner so as to avoid allowing the armature to bridge between adjoining poles, so it is necessary to make the sum of the arcs spanned by these faces (in the plane of rotation) equal to considerably less than 360 degrees electrical.

The invention will now be described in more detail, reference being made to the accompanying drawings in which:

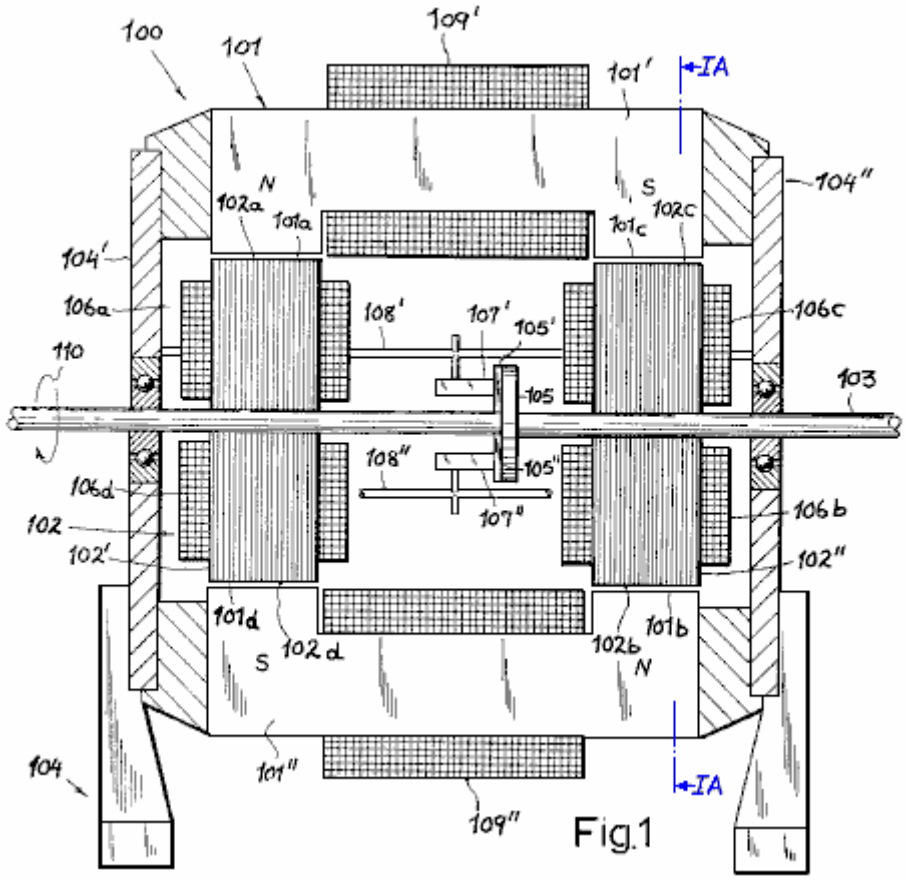


Fig.1

Fig.1 and **Fig1A**. illustrate a first embodiment of my invention, shown in axial section and in a cross-sectional view taken on line IA - IA of Fig.1 respectively.

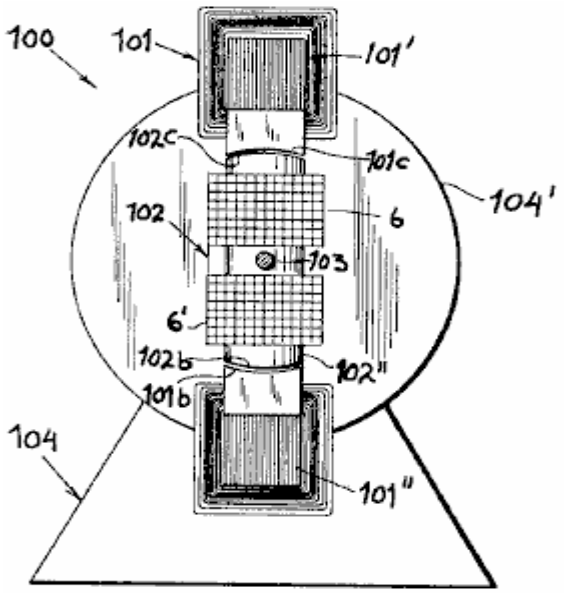


Fig.1A

Fig.2 and **Fig.3** are perspective views illustrating two other embodiments.

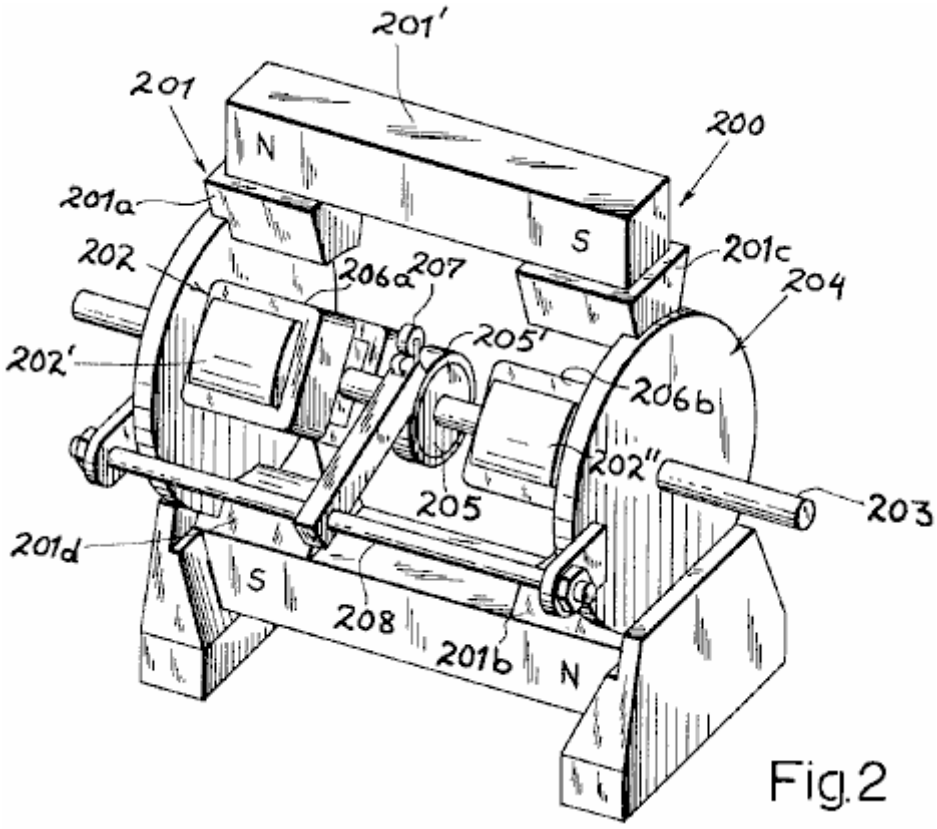


Fig.2

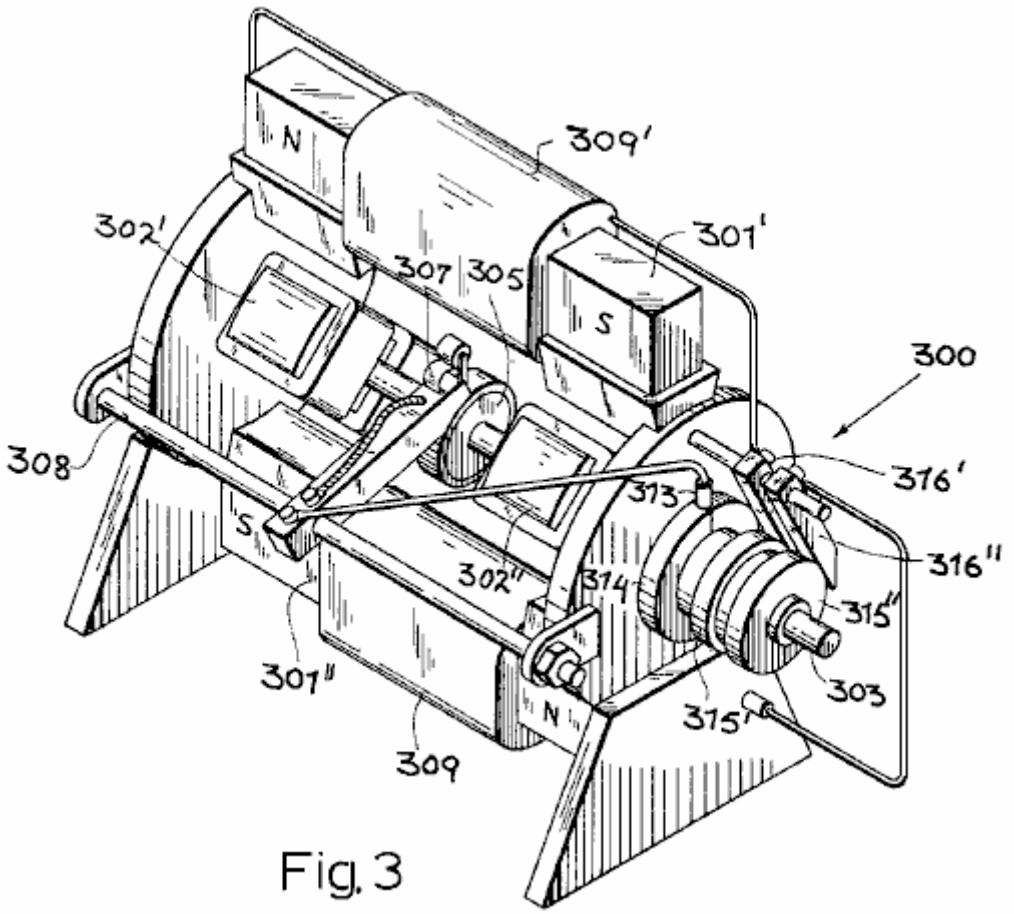


Fig. 3

Fig.4 and **Fig.5** illustrate diagrammatically, two output circuit arrangements, one for a DC output and one for an AC output.

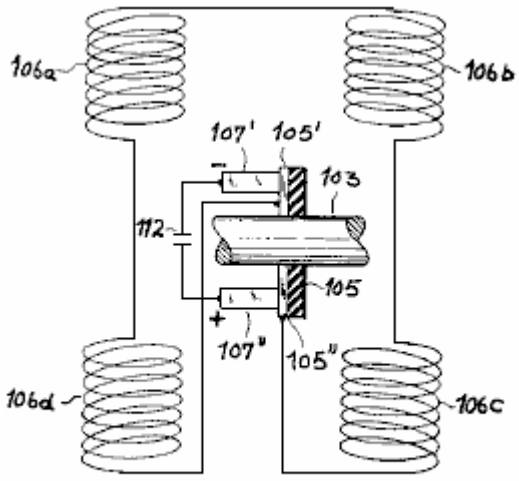


Fig. 4

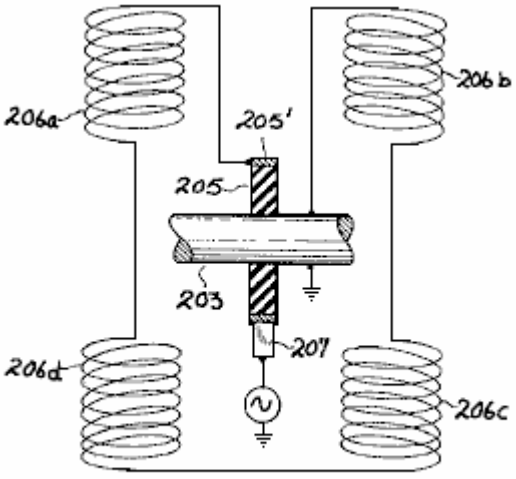


Fig. 5

Fig.6 is a somewhat diagrammatic illustration of an arrangement for comparing the outputs of a conventional generator and a generator according to this invention.

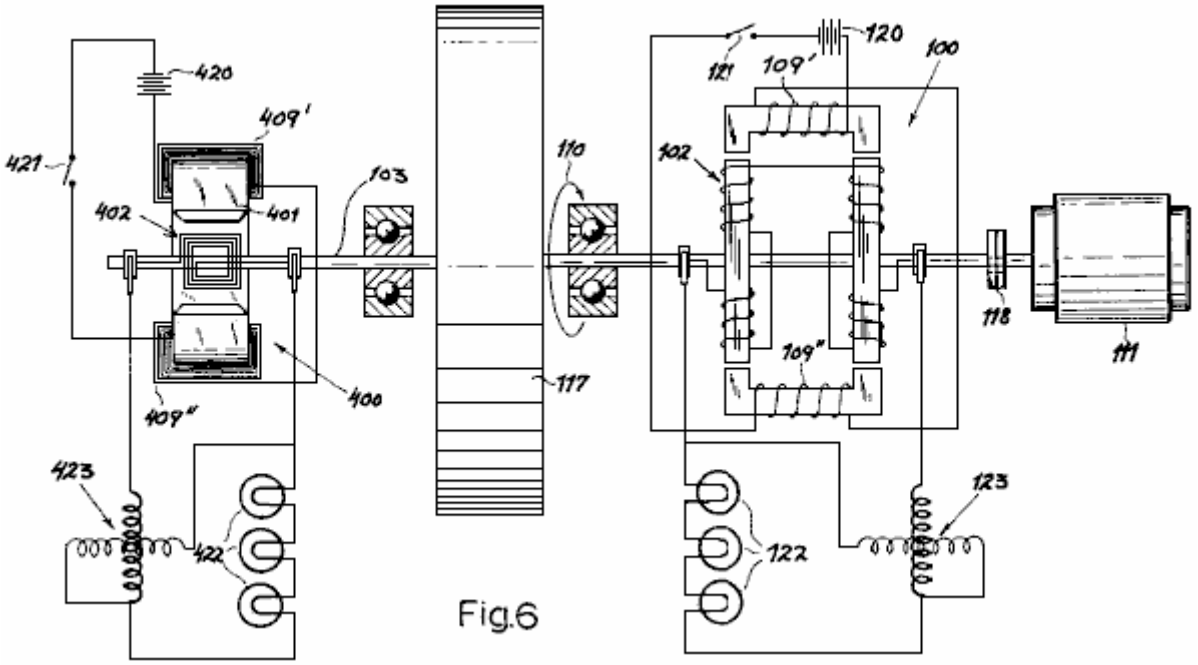


Fig.6

The generator **100** shown in **Fig.1** and **Fig.1A** comprises a stator **101** and a rotor **102** which has a pair of laminated armatures **102'** and **102"**, carried on a shaft **103** which is free to rotate in bearings mounted in the end plates **104'** and **104"**, of a generator housing **104** which is made from non-magnetic material (e.g. aluminium) which is rigidly attached to the stator.

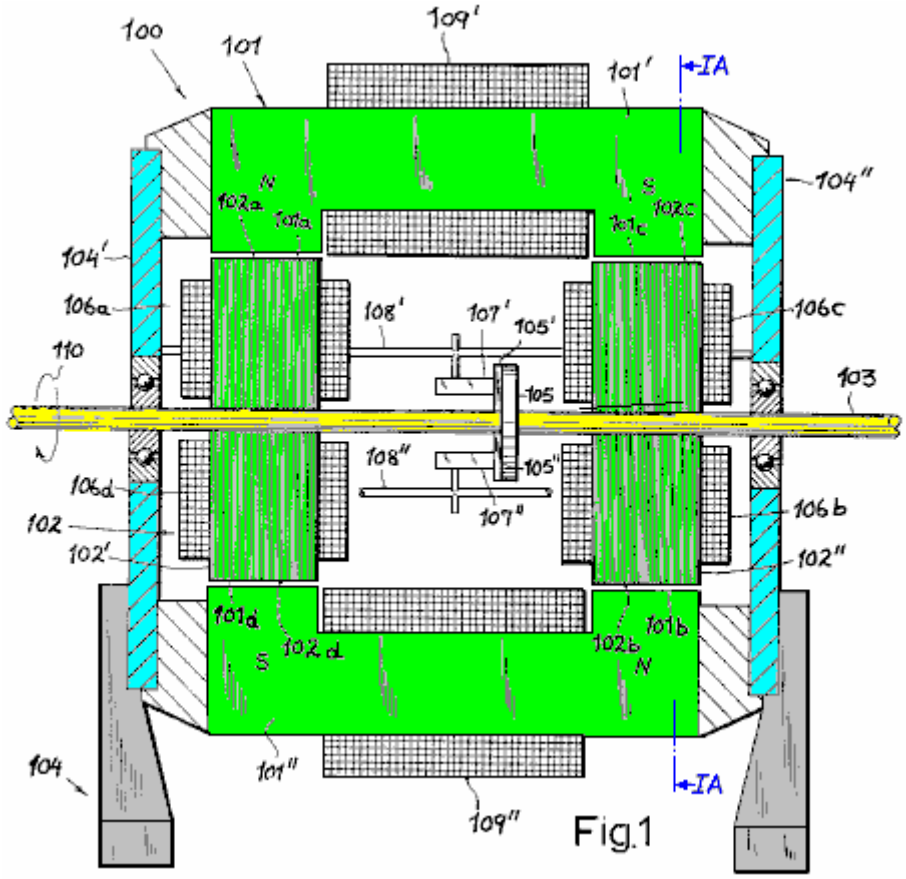


Fig.1

Shaft **103** is coupled to a source of driving power indicated diagrammatically by an arrow **110**. The stator **101** includes a pair of yoke-shaped laminated electromagnets **101'** and **101"** whose extremities form two pairs of co-planar pole pieces, designated respectively **101a, 101b** (North magnetic pole) and **101c, 101d** (South magnetic pole). The pole pieces have concave faces, facing towards the complimentary convex faces **102a, 102d** of armature **102'** and **102b, 102c** of armature **102"**. These faces whose concavities are all centred on the axis of

shaft **103**, extend over arcs of approximately 20^{O} to 25^{O} each in the plane of rotation (**Fig.1A**) so that the sum of these arcs adds up to about 90^{O} geometrically and electrically.

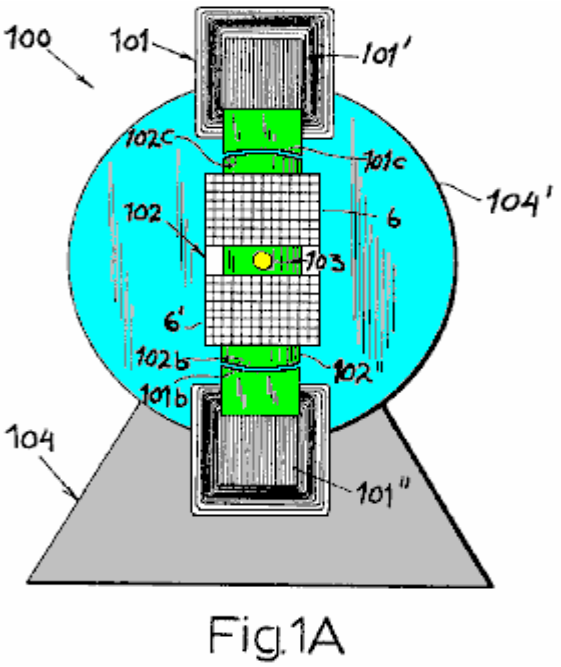


Fig.1A

The stator magnets **101', 101"** are surrounded by energising windings **109', 109"** which are connected across a suitable source of constant direct current (not shown). Similar windings, each composed of two series-connected coils **106a, 106d** and **106b, 106c**, surround the rotor armatures **102'** and **102"**, respectively. These coils form part of an output circuit which further includes a pair of brushes **107', 107"** which are carried by arms **108', 108"** on housing **104** with mutual insulation brushes **107', 107"** co-operate with a pair of commuter segments **105', 105"** (see also **Fig.4**) which are supported by a disc of insulating material **105**, mounted on shaft **103**.

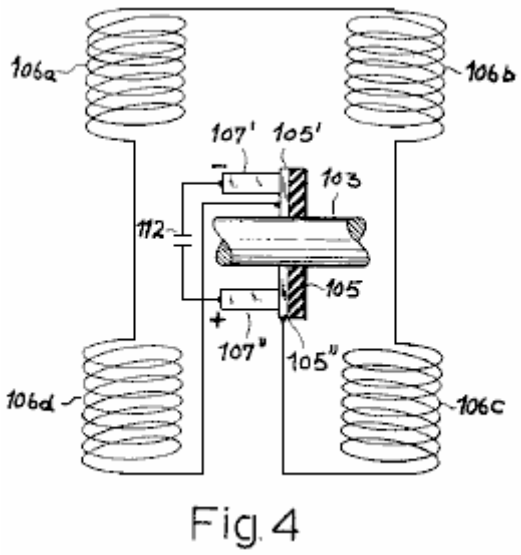


Fig.4

By virtue of the series-connection of coils **106a-106d** between the segments **105'** and **105"**, as illustrated in **Fig.4**, the alternating voltage induced in these coils gives rise to a rectified output voltage at brushes **107'** and **107"**. The unidirectional current delivered by these brushes to a load (not shown) may be smoothed by conventional means, represented by capacitor **112** in **Fig.4**.

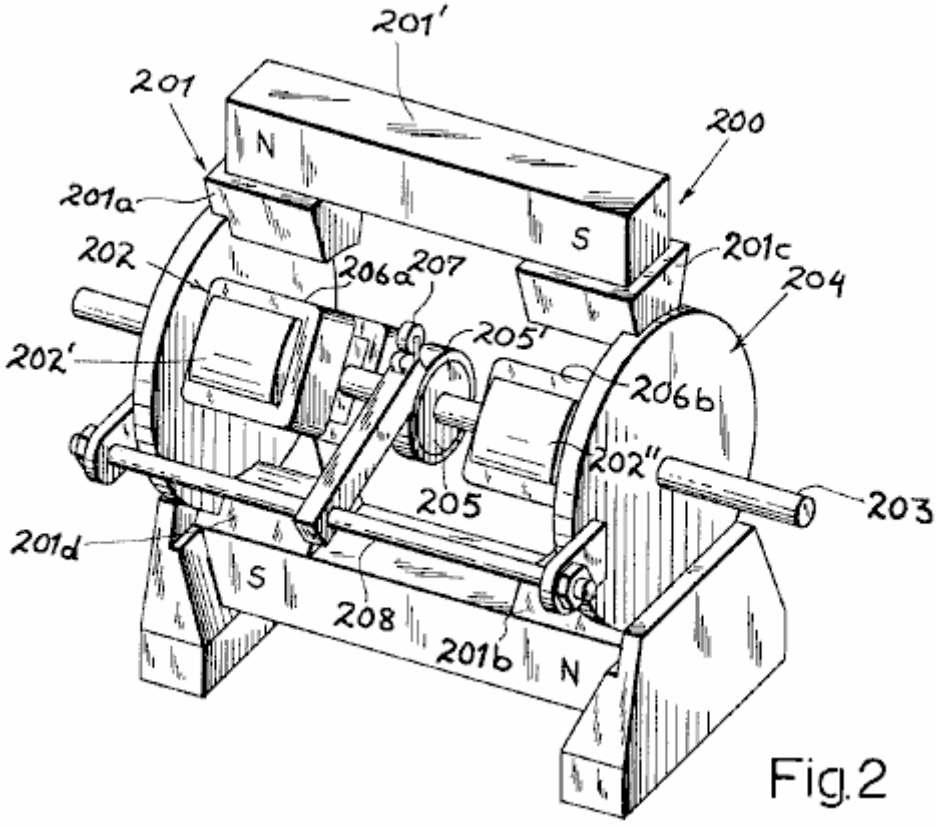


Fig.2

Fig.2, shows a modified generator **200**, whose housing **204**, supports a stator **201** essentially consisting of two permanent bar magnets **201'** and **201"**, extending parallel to the drive shaft **203** (on opposite side of it), each of these magnets being rigid and each having a pair of sole shoes **201a, 201c** and **201b, 201d** respectively. Rotor **202** is a pair of laminated armatures **202'** and **202"**,similar to those of the previous embodiment, whose output coils **206a, 206b, 206c** and **206d** are serially connected between a slip-ring **205'**, supported on shaft **203** through the intermediary of an insulating disc **205**, and another terminal here represented by the grounded shaft **203** itself. Slip-ring **205'** is contacted by brush **207** on holder **208**, the output of this brush being an alternating current of a frequency determined by the rotor speed.

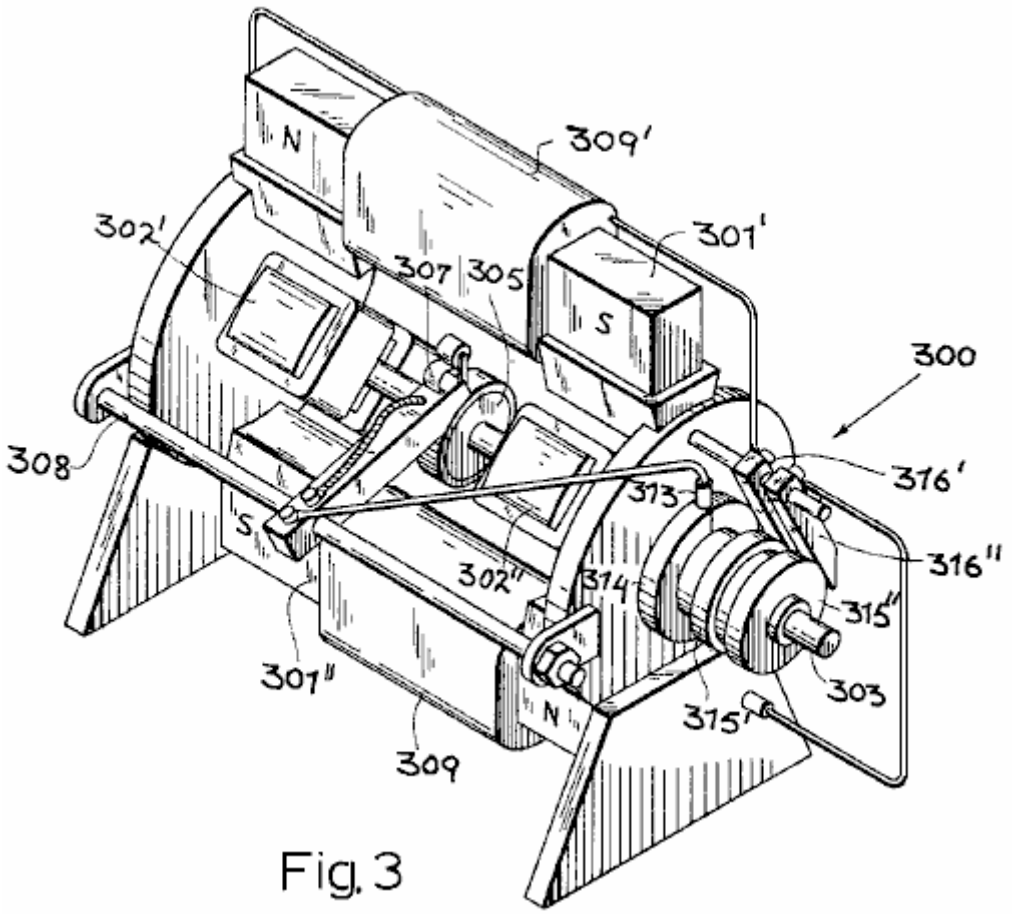


Fig. 3

Fig.3 shows a generator **300** which is basically similar to the generator **100** shown in **Fig.1** and **Fig.1A**. It's shaft **303** carries a pair of laminated soft-iron armatures **302', 302"** which can rotate in the air gaps of a pair of electromagnets **301', 301"** which have windings **309'** and **309"**. The commutator **305** again co-operates with a pair of brushes **307**, only one of which is visible in **Fig.3**. This brush, carried on an arm **308**, is electrically connected to a brush **313** which engages with a slip-ring **314** positioned on an extremity of shaft **303** which also carries two further slip-rings **315', 315"** which are in conductive contact with ring **314** but are insulated from the shaft. Two further brushes **316', 316"** contact the rings **315', 315"** and respectively are connected to windings **309'** and **309"**. The other ends of these windings are connected to an analogous system of brushes and slip-rings on the extremity of the opposite shaft, and arranged so that the two commutator brushes are effectively bridged across the windings **309'** and **309"** in parallel. Therefore, in this embodiment, the stator magnets are energised from the generator output itself, it being understood that the magnets **301'** and **301"** (made, for example, of steel rather than soft iron) will have a residual coercive force sufficient to induce an initial output voltage. Naturally, the circuits leading from the brushes **307** to the windings **309', 309"** may include filtering as described in connection with **Fig.4**.

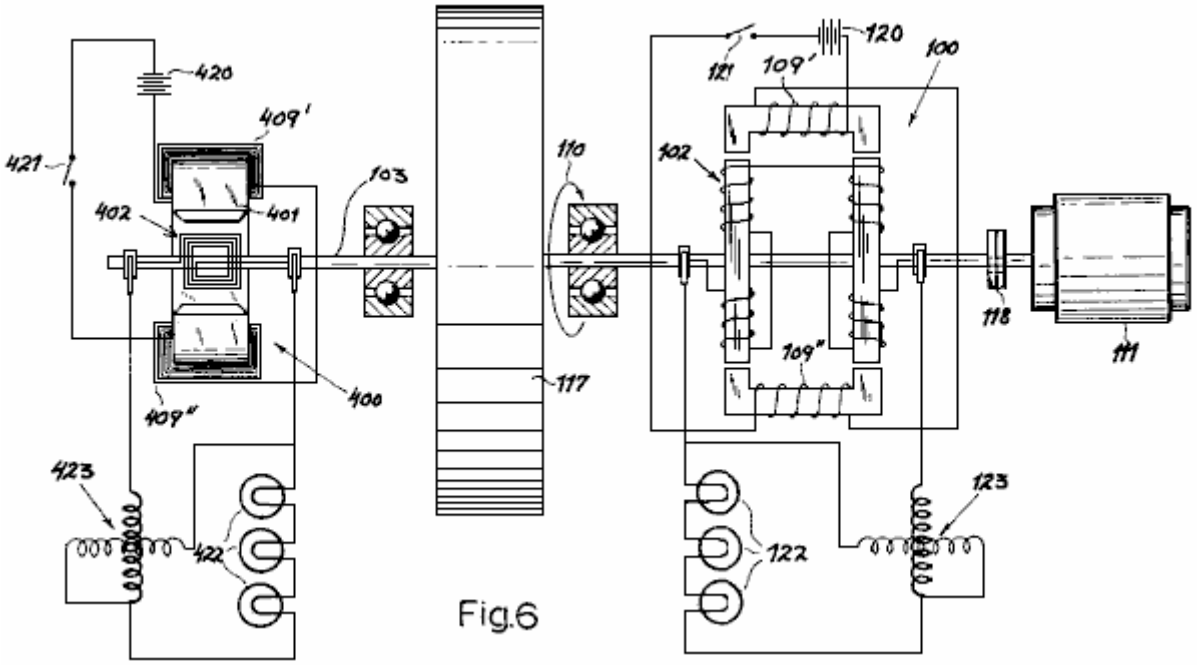


Fig.6

Fig.6 shows a test circuit designed to compare the outputs of a generator of this design, such as the unit **100** of **Fig.1** and **Fig.1A**, with a conventional generator **400** of the type having a looped armature **402** which rotates in the gap of a stator magnet **401** which is fitted with energising windings **409', 409"**. The two generators are interconnected by a common shaft **103** which carries a flywheel **117**. This shaft is coupled through a clutch **118** to a drive motor **111** which drives the rotors **402** and **102** of both generators in unison, as indicated by arrow **110**. Two batteries **120** and **420**, in series with switches **121** and **421**, represent the method of supplying direct current to the stator windings **109', 109"** and **409', 409"** of the two generators.

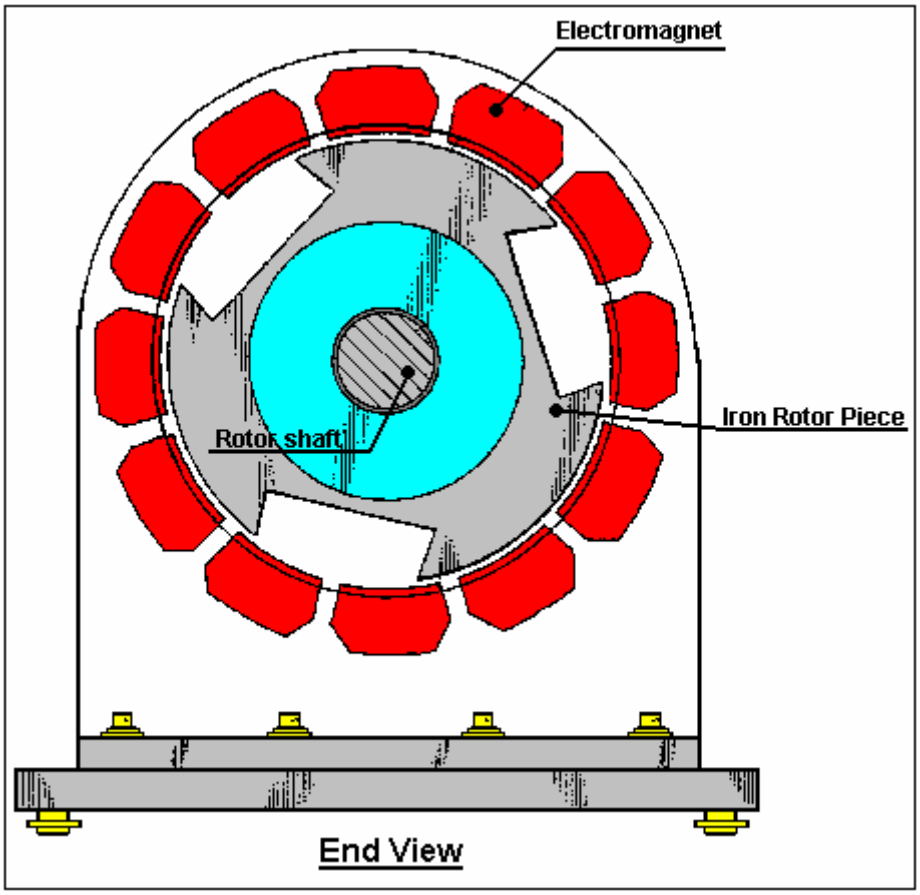
The rectified output of generator **100** is delivered to a load **122**, shown here as three incandescent lamps connected in series, and with a combined consumption of 500 watts. Generator **400**, provides current into an identical load **422**. Two watt meters **123** and **423** have their voltage and current windings connected respectively in shunt and in series with their associated loads **122** and **422**, to measure the electric power delivered by each generator.

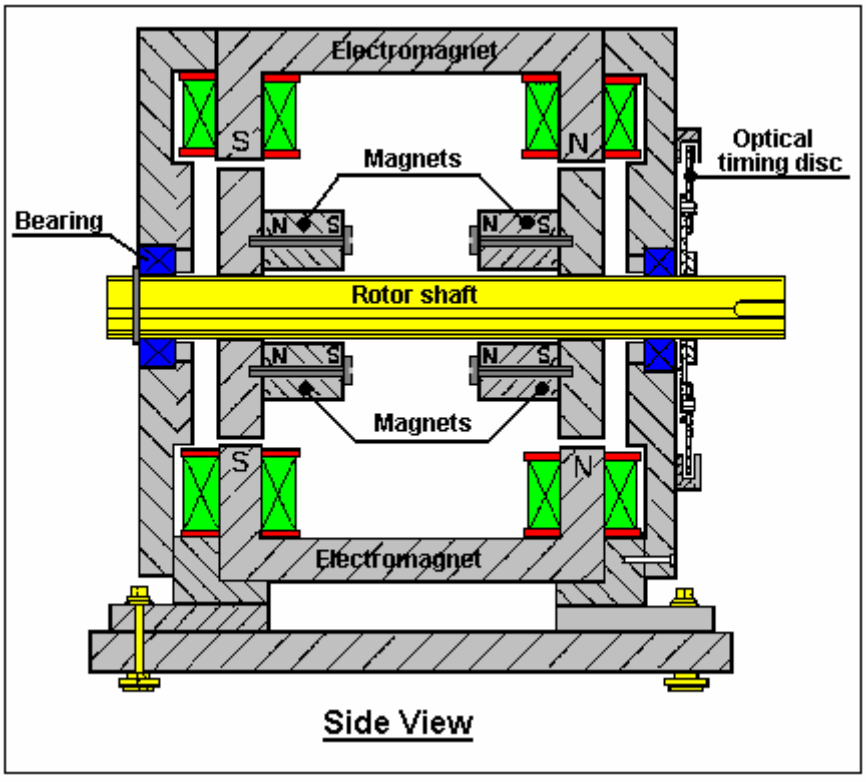
When clutch **118** is engaged, shaft **113** with it's flywheel **117** is brought to an initial driving speed of 1,200 rpm. at which point, the switch **421** in the energising circuit of the conventional generator **400**, is closed. The lamps **422** light immediately and the corresponding wattmeter **423** shows an initial output of 500 watts. However, this output drops immediately as the flywheel **117** is decelerated by the braking effect of the magnetic field on armature **402**.

Next, the procedure is repeated but with switch **421** open and switch **121** closed. This energises generator **100** and the lamps **122** light up, wattmeter **123** showing an output of 500 watts, which remains constant for an indefinite period of time , there being no appreciable deceleration of flywheel **117**. When the clutch **118** is released and the rotor speed gradually decreases, the output of generator **100** is still substantially 500 watts at a speed of 900 rpm. and remains as high as 360 watts when the speed dropped further to 600 rpm. In a similar test with a generator of the permanent magnet type, such as the one shown at **200** in **Fig.2**, a substantially constant output was observed over a range of 1600 to 640 rpm.

Teruo Kawai's COP=1.6 Magnetic Motor.

In July 1995, a patent was granted to Teruo Kawai for an electric motor. In the patent, Teruo states that a measured electrical input 19.55 watts produced an output of 62.16 watts, and that is a COP of 3.18. The main sections of that patent are included in the Appendix.





In this motor, a series of electromagnets are placed in a ring to form the active stator. The rotor shaft has two iron discs mounted on it. These discs have permanent magnets bolted to them and they have wide slots cut in them to alter their magnetic effect. The electromagnets are pulsed with the pulsing controlled via an optical disc arrangement mounted on the shaft. The result is a very efficient electric motor whose output has been measured as being in excess of its input.

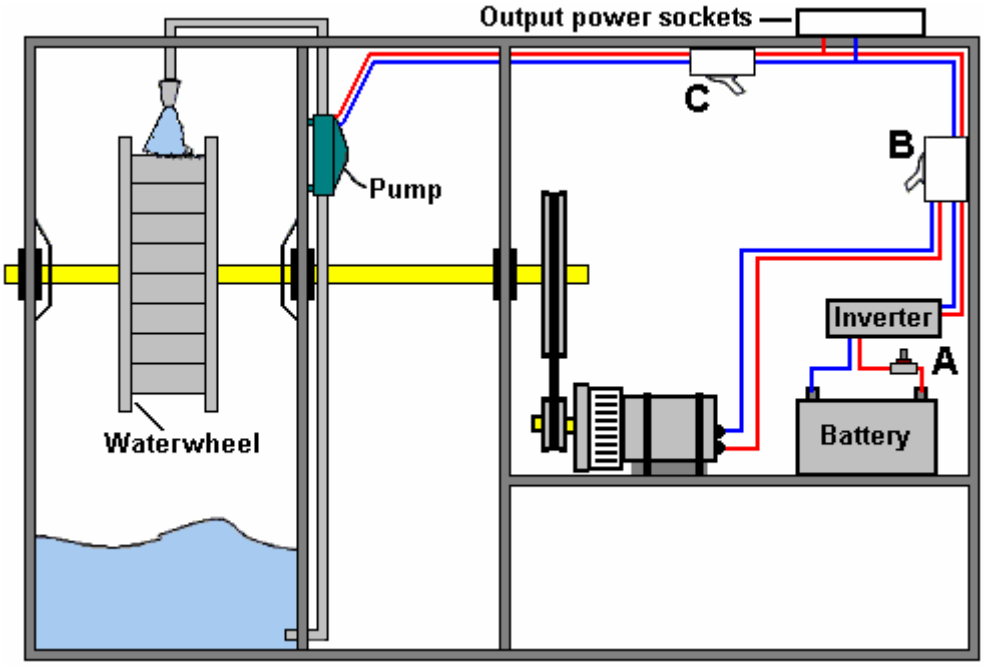
James Hardy's Self-Powered Water-Pump Generator.
There is a video on Google which shows an interesting self-powered water-pump driven, electrical generator at:
http://www.youtube.com/watch?v=IGpXA6qhH_Q

This is a very simple device where the jet of water from the pump is directed at a simple water-wheel which in turn, spins an electrical alternator, powering both the pump and an electric light bulb, demonstrating free-energy.

Initially, the generator is got up to speed, driven by the mains electrical supply. Then, when it is running normally, the mains connection is removed and the motor/generator sustains itself and is also able to power at least one light bulb. The generator output is normal mains current from a standard off-the-shelf alternator.

James has Patent Application US 2007/0018461 A1 published in 2007 on his design. In that application he points out that a major advantage of his design is the low noise level produced when the generator is running. In the video and the pictures above, the demonstration has the housing opened up in order to show how the generator system works, but during normal use, the compartments are completely sealed.

In his document, James shows the overall system like this:



The housing is divided into three separate compartments. The first compartment has a strong axle shaft running through it, supported on ball or roller bearings – possibly ceramic for this environment. The bearings are protected by being covered by splash guards which keep the water (or other liquid) off them. A waterwheel of almost any type is mounted on the shaft and a high-capacity water pump directs a stream of liquid on to the waterwheel, striking the paddles at right angles in order to provide the maximum impact.

This first compartment is sealed in order to contain all of the liquid inside it and the bottom is effectively a sump for the liquid. A pipe located near the bottom of the compartment feeds the liquid to the pump which is located in the second compartment. The pump boosts the liquid through a nozzle, directing it at the waterwheel. While almost any nozzle will work, it is usual to choose one which produces a concentrated jet of liquid in order to generate the largest possible impact. One would expect that the larger the diameter of the waterwheel, the more powerful the system would be. However, that is not necessarily the case as other factors such as the overall weight of the rotating members might affect the performance. Experimentation should show the most effective combination for any given pump.

The rotating shaft is given a third bearing supported by the side of the final compartment. The shaft then has a large diameter belt pulley mounted on it, the belt driving a much smaller pulley mounted on the shaft of the generator. This raises the rate at which the generator shaft is rotated. If the pump operates on AC mains voltage, then the generator will be one which generates mains voltage AC. If the pump operates on, say, 12 volts, then the generator will be one which generates 12 volts DC. The diagram above, shows the arrangement for a mains voltage system as that is probably the most convenient. If a 12-volt system is chosen, then the inverter can be omitted.

The generator is started by pressing the 'normally open' press-button switch marked "**A**" in the diagram. This passes the battery power through to the 1-kilowatt inverter which then generates AC mains voltage. The switch marked "**B**" is a "changeover" switch, and for starting, it is set so that it passes the AC power through switch "**A**" to the pump. This causes the pump to turn on and direct a powerful jet of liquid at the waterwheel, forcing it around and so powering the generator. When the generator gets up to full speed, switch "**B**" is flipped over, disconnecting the inverter and feeding the generator power through to the pump, keeping it running and supplying additional power to the output power sockets mounted on top of the housing. The press-button switch is released, disconnecting the battery which is no longer needed. Switch "**C**" is an ordinary On/Off mains switch which is needed if you want to turn the generator off.

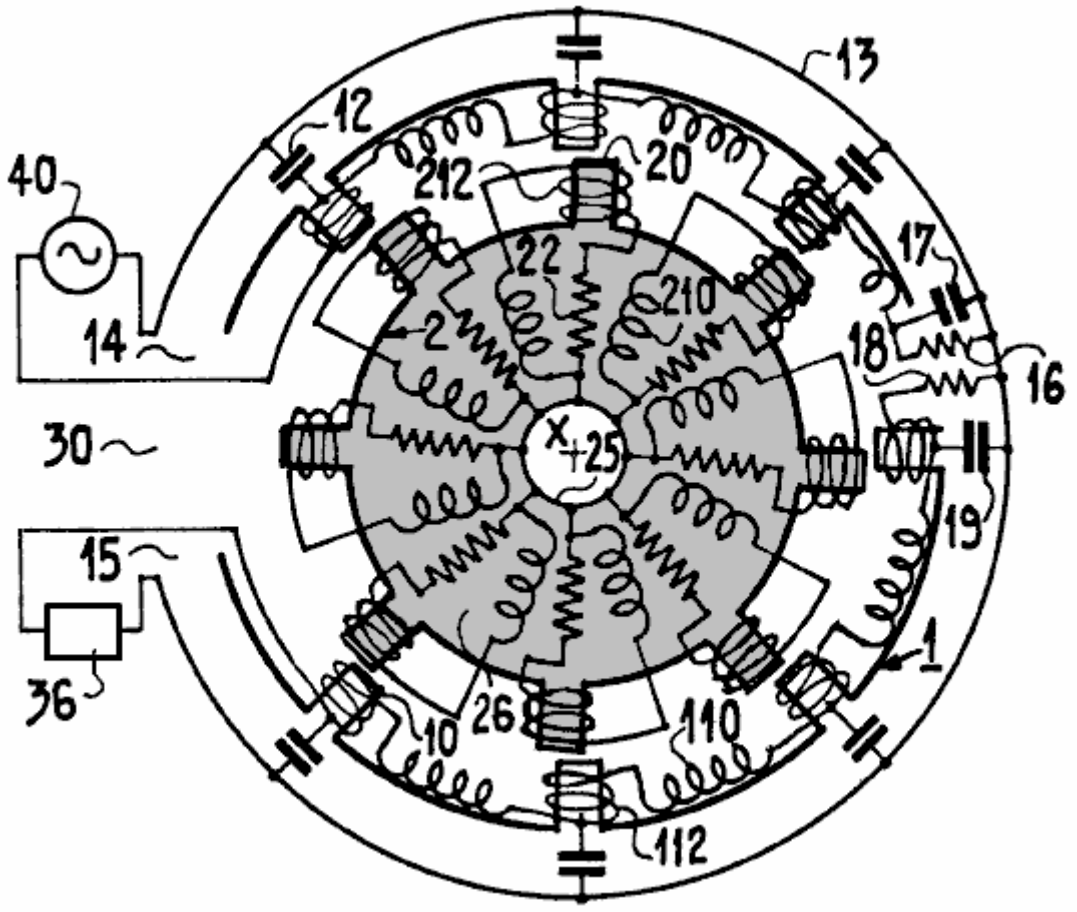
A major advantage of this generator system is that the main components can be bought ready-made and so only very simple constructional skills and readily available materials are needed. Another advantage is that what is happening can be seen. If the pump is not working, then it is a simple task to discover why. If the generator is not spinning, then you can see that and sort the problem. Every component is simple and straightforward.

James suggests that a suitable pump is the 10,000 gallons per hour "Torpedo Pump" from Cal Pump, web site: http://www.calpumpstore.com/products/productdetail/part_number=T10000/416.0.1.1:

Georges Mourier's COP=10,000 Motor/Generator Patent.
This patent from Frenchman Georges Mourier is quite remarkable in that it states clearly in the patent that it has a Coefficient Of Performance of 10,000 with an input power of just 10 watts producing an output of 100 kilowatts, which, considering the massive opposition of the US Patent Office to any claim of output power being greater than the input needed to produce that output, is little short of amazing. This patent contains a fair amount of mathematics where Georges goes about showing why there is a power amplification. This can be ignored by most people as the point to concentrate on is how to build a power-amplifying motor/generator.

US Patent 4,189,654 19th February 1980 Inventor: Georges Mourier

Electrical machine operating as a generator or as an amplifier



Abstract
The invention relates to an electrical machine. The stator **1** is a delay line, in the form of a low-pass filter in the example, formed by inductance **11** and capacitors **12** connected between these inductances and the common conductor **13**. The rotor **2** comprises dissipating elements **22** incorporated in circuits **26**, separate in the example,

and having a common point **25**. It is put in movement by a motor. The machine operates as a high-gain amplifier having a wide band of high-frequency signals applied to the input **14** of the stator, separated from the output **15** by the decoupling zone **30**. High powers are obtainable. Application to installations for testing vibration of industrial equipment and to high-power long wave radio transmission.

Description

The invention relates to an electrical machine capable of operating as a generator and as an amplifier. The machine comprises a fixed part, or stator, in which moves a moving part designated hereinafter by the term “rotor”, by analogy with the case of machines of the prior art in which the movement in question is a movement of rotation, although this movement may be other than a rotation and in particular a rectilinear translation in the case of the invention.

The stator consists of a line having two conductors which have two input terminals and two output terminals; it’s rotor comprises resistive elements under conditions which will be described in detail later.

In operation, a wave is propagated between the input terminals and the output terminals in question in this line. Electrical machines are known from U.S. Patent 3,875,484, in which the stator comprises inductances and capacitances incorporated in a transmission line, as in the machines of this invention, along which there is propagated, in operation, an electric wave, but contrary to the case of the invention, this line has only one pair of terminals to which those of the alternating current source are connected. The application of the voltage of this source between these terminals causes the rotation of the rotor of the machine which, as it is operating as a motor, does not have an output. Owing to the structure, a brief indication of which has been given above, the machine of this invention is intended, on the contrary, to operate as a generator or as an amplifier; it has an output constituted by the other pair of terminals of the stator, the rotor being driven by an exterior motor.

A better understanding of the invention will be had from the ensuing description with reference to the accompanying Figures which represent:

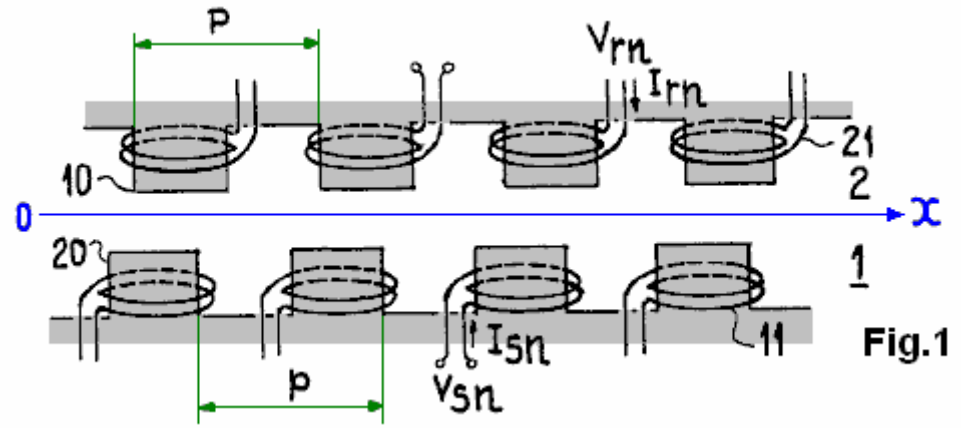
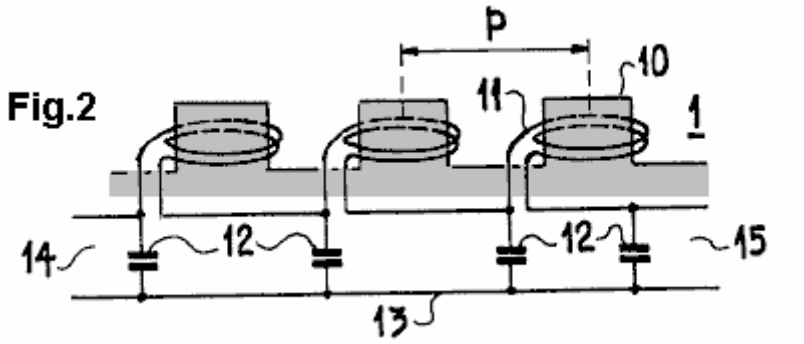


FIG.1, diagrammatically, the stator assembly and the rotor of a machine to which the invention is applied;



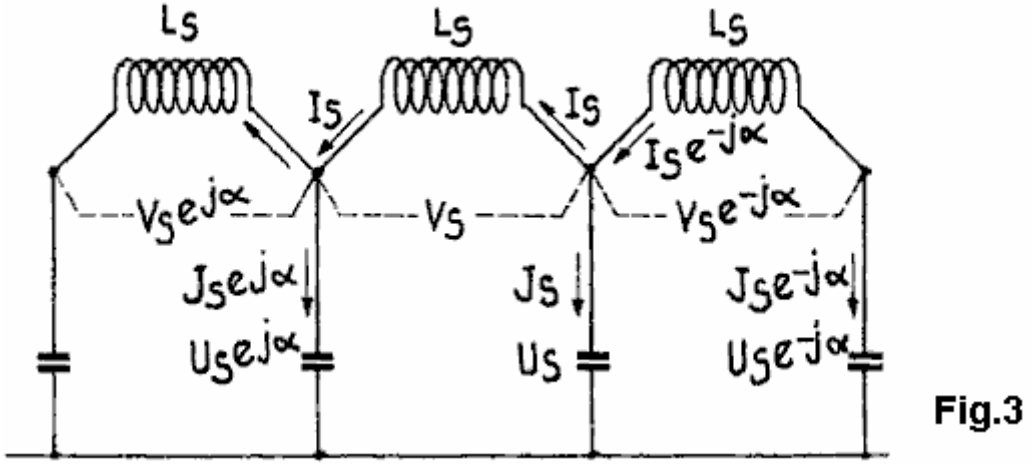


Fig.3

Fig.2 and **Fig.3**, diagrammatically, an embodiment of an electrical circuit of the stator of a machine of the invention and the corresponding diagram;

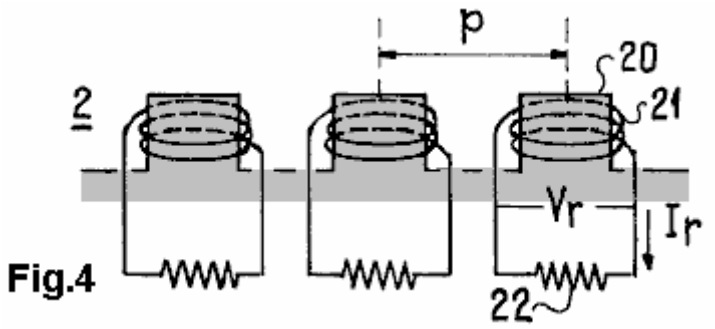


Fig.4

Fig.4, diagrammatically, a rotor structure of the machines of the invention;

Fig.5

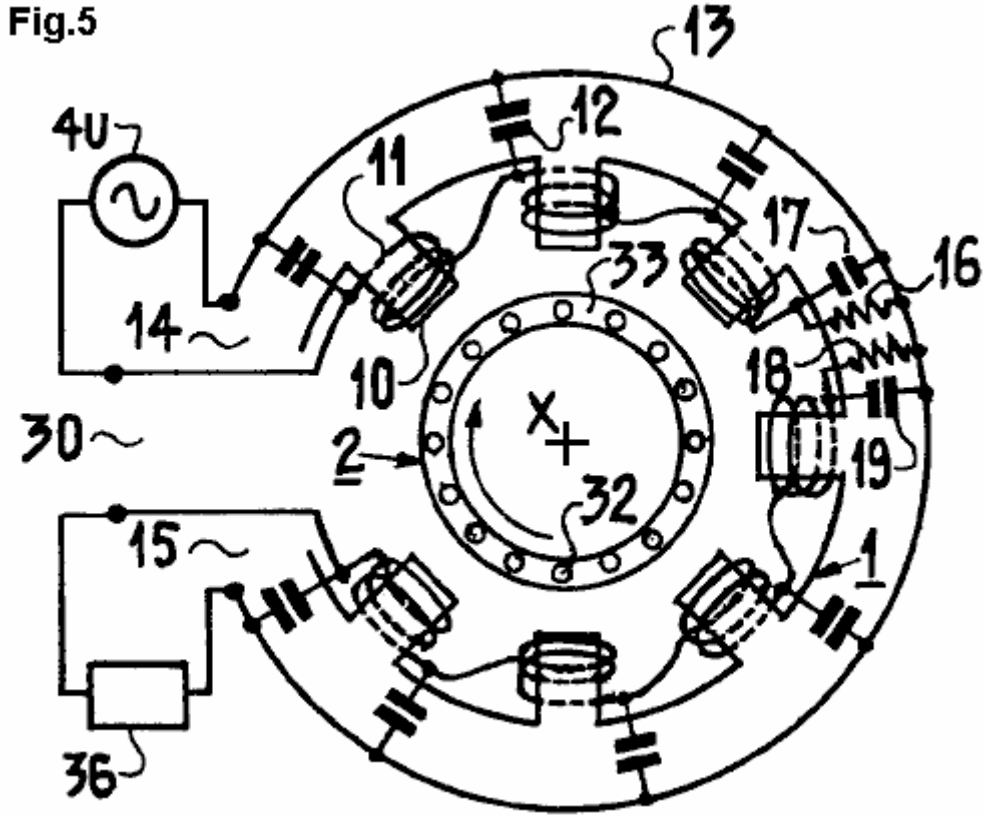


Fig.5, a diagrammatic view of a variant embodiment of the machine of the invention having a stator according to the design of **Fig.2**;

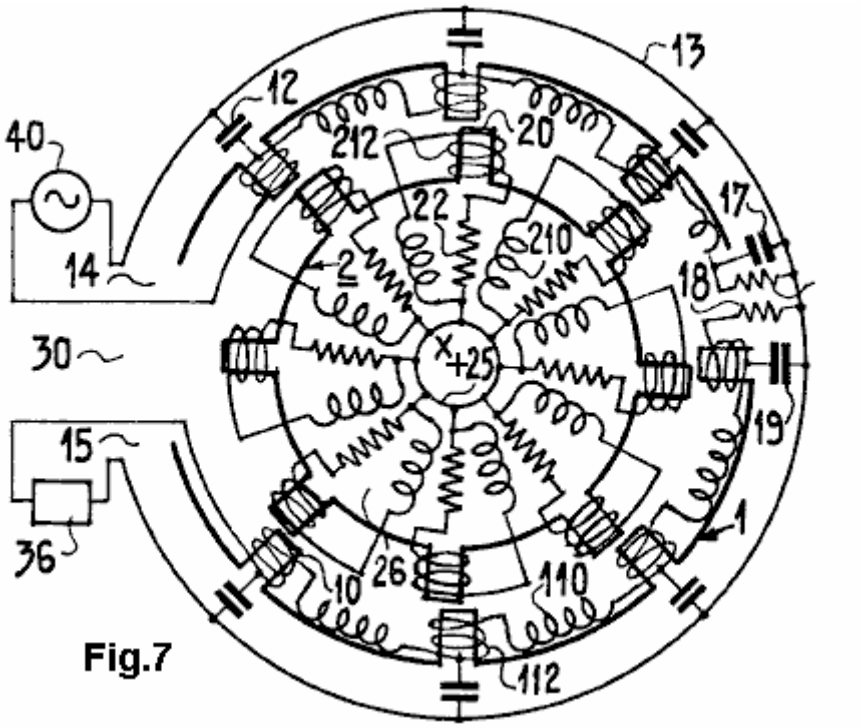


Fig.7

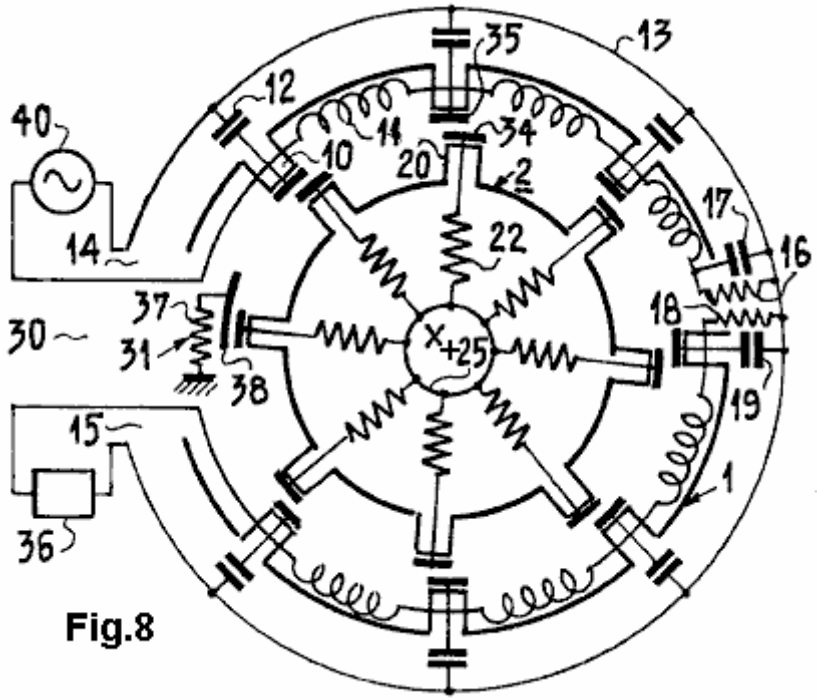


Fig.8

Fig.7 and **Fig.8**, diagrammatic views of two variant embodiments of the machine of the invention operating as amplifiers;

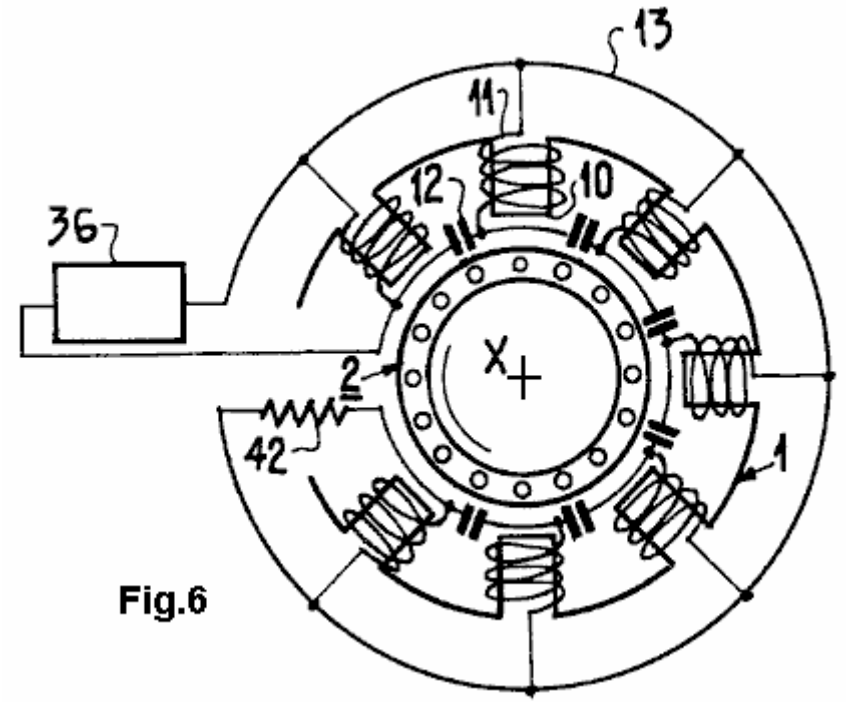


Fig.6

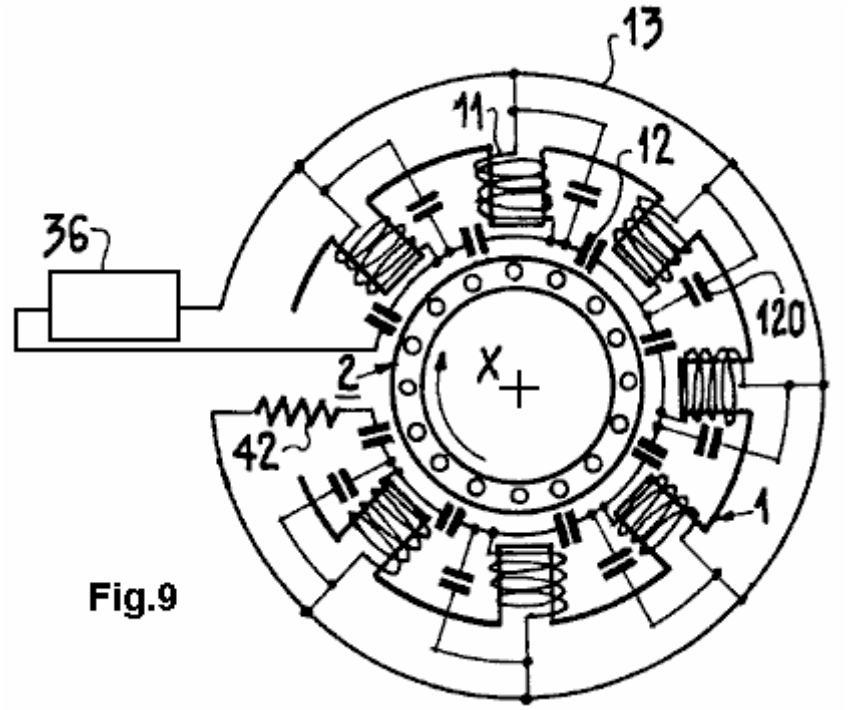


Fig.9

Fig.6 and **Fig.9**, two modifications of the machine of the invention operating as generators.

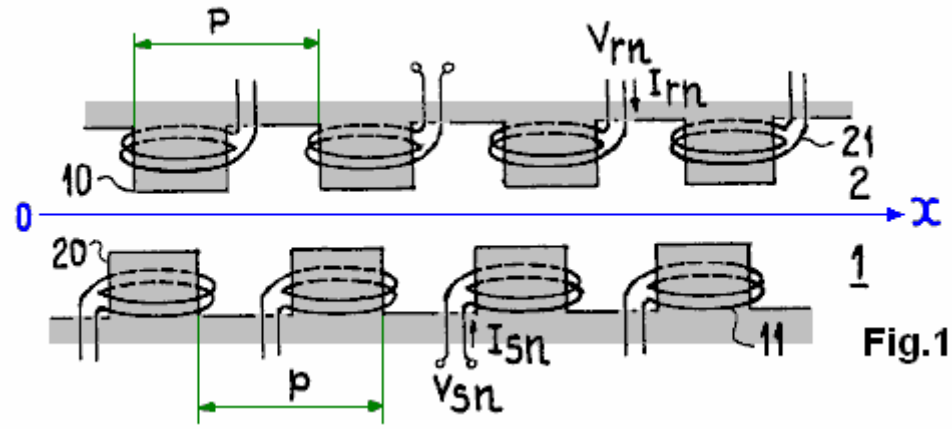


Fig.1

Fig.1 shows diagrammatically, a machine structure to which the invention is applied, in which the stator **1** and the rotor **2** comprise poles aligned in the direction o → x along which the rotor movement occurs. These poles are marked as **10** and **20** and are repeated with the same horizontal spacing or pitch **p** in both the rotor and the stator. References **11** and **21** are the inductive windings through which the coupling occurs between a pole of the rotor and a pole of the stator when they align in the course of their movement.

V_{sn} and V_{rn} are the instantaneous values of the voltages at the ends of these windings, and I_{sn} and I_{rn} are the currents in those coils at the moment of this alignment, n being the number assigned to the pole in each part of the machine. In order to simplify the notations, this index number 'n' will be omitted when mentioning the voltages and currents in question. φ designates the magnetic flux and M the coefficient of mutual inductances of this coupling. There are then obtained the following known equations:

$$\phi_s = L_s I_s + M I_r \quad \ldots\ldots\ldots\ldots (1)$$
$$\phi_r = M I_s + L_r I_r \quad \ldots\ldots\ldots\ldots (2)$$
$$V_s = j\omega L_s I_s + j\omega M I_r \quad \ldots\ldots\ldots\ldots (3)$$
$$V_r = j\omega' M I_s + j\omega' L_r I_r \quad \ldots\ldots\ldots\ldots (4)$$

where $\mathbf{L_s}$ and $\mathbf{L_r}$ are the values of inductances **11** and **21** and $\boldsymbol{\varphi_s}$ and $\boldsymbol{\varphi_r}$ the magnetic flux in these windings.

In these equations, the quantities **ω** and **ω'** are the angular frequencies of the currents in the stator and the rotor respectively; **ω** = 2πf, where **f** is the corresponding frequency. For equal phase differences between two successive poles of the stator and rotor, the pulses **ω** and **ω'** in the relation:

$$\omega' = \omega - \beta u \quad \ldots\ldots (5)$$

where **β** is in accordance with the usual definition the constant of propagation of the magnetic field wave along the stator, and **u** is the speed at which the rotor moves in front of the stator in the direction of the reference axis o – x in the case of the machine of the embodiment of **Fig.1**.

$$\beta = \frac{2\pi}{\Lambda} = \frac{\omega}{v_\phi}$$

if **Λ** and **vφ** respectively are the wavelength and the phase velocity in the stator.

In the embodiment of the machine of this invention having a stator in the form of a low-pass filter transmission line with an input **14** and an output **15**, and the inductances **11**, charge-accumulating elements **12** mounted between these inductances, and a common conductor **13** as shown in **Fig.2** and **Fig.3**, there is added to the foregoing equations the following equation:

$$\frac{V_s}{I_s} = -\frac{2}{jC_s\omega}(1 - \cos\alpha) \quad \ldots\ldots\ldots\ldots (6)$$

where **α** is equal to $\boldsymbol{\beta_p}$ and $\mathbf{C_s}$ is the value of the capacitors **12** (**Fig.2**) included in the construction of the filter, it being observed that there is obtained with the notations of **Fig.3**:

$$I_s = \frac{-J_s}{1 - \exp(-j\alpha)} \quad \text{and} \quad V_s = U_s(1 - \exp(j\alpha)), \text{ with } U_s = -j\frac{J_s}{C_s\omega}$$

There is also added:

$$V_r + R_r I_r = 0 \quad \ldots\ldots\ldots\ldots (7)$$

the equation for a rotor constructed, as shown in **Fig.4**, by separate resistive circuits each of which comprises, in addition to the foregoing, coupling inductance **21**, a resistor **22** of value R_r.

By eliminating the magnitudes of flux, voltage and current between the homogeneous equations (1), (2), (3), (4), (6) and (7) and by replacing **ω'** with it's expression taken from equation (5), there is obtained an equation in **β** whose roots have an imaginary part, whence it results that the variation with respect to time of the wave propagated along the stator in the direction **o --- x** of **Fig.1**, **whose amplitude** is proportional, in accordance with the conventional notation, to **exp j(ωt-βx)**, **undergoes an amplification in the course of this propagation**.

The equation in question is:

$$(\beta - \frac{\omega}{u} + \frac{j}{u\frac{L_r}{R_r}})(\beta^2 \quad \beta^2 1) = j\frac{\beta^2 1}{u\frac{L_r}{R_r}} \quad \frac{\frac{M^2}{L_r L_s}}{1 - \frac{M^2}{L_r L_s}} \quad \ldots\ldots (8)$$

$$\text{with} \quad \beta^2 1 = \omega^2 C_s L_s \frac{1}{p^2}(1 - \frac{M^2}{L_4 L_s})$$

and assuming that **α** is small, that is to say, that the machine has a large number of phases in the electro-technical sense of the word, the roots of which will be calculated by their relative difference **z** from a reference value chosen to be equal to **ω / u**, which is but little different, under the conditions of operation of the machines of the invention, from the quantity β_1 defined above; there is obtained:

$$z = \frac{\beta - \beta_1}{\beta_1} \text{ and } \beta = \beta_1(1+z) \quad \ldots\ldots\ldots (9)$$

The foregoing ratio **ω/u** is none other than the value of **β** in synchronous machines, in which **ω'** is null (equation 5). In the machines of this invention, **ω'** is non-zero: these machines operate in the asynchronous manner.

The equation in **z** is written:

$$z^2 + 2j\gamma z - j\gamma = 0 \quad \ldots\ldots\ldots (8')$$

in the particular case of a coupling coefficient

$$\mu = \frac{M^2}{L_r L_s} \; de \; \frac{1}{2}$$

It has two roots having an imaginary part. **The machine is capable of amplifying a signal applied to the input of the stator**. The order of magnitude of this amplification is shown below as an example. The expression of these roots is:

$$z\pm = -j\gamma \pm j\sqrt{\gamma^2 - j\gamma} \quad \ldots\ldots\ldots (10)$$

with:

$$z\pm = \frac{R_r}{L_r} \quad \frac{1}{2u\beta_1'}$$

to which the following numerical values correspond for a certain number of values of **γ**. The amplified wave corresponds to the root **Z+**.

	0.01	0.03	0.1	0.3	1
Z+	0.0703 + 0.610	0.120 + 0.094j	0.212 + 0.135j	0.334 + 0.149j	0.455 + 0.098j

The gain "**g**" in power per unit length of the stator is, according to equation (9) is:

$$g\ \mathrm{db/m} = 8.7\beta_1\text{-(imaginary part of z)}$$

In the given example, this gain reaches it's maximum value for the wave corresponding to **z+**, for **γ** between 0.1 and 0.3. The imaginary part of **z** is then in the neighbourhood of 0.15, which gives for **g**, in respect of a stator whose length is 5 wavelengths (x = 5), about 40 decibels. It will be observed that the corresponding propagation constant **β** differs only slightly in absolute value from the constant **ω/u** corresponding to synchronous operation. The machine of this invention operates under conditions which differ only slightly from synchronous operating conditions.

The wave passing through the stator is amplified in the manner of an electromagnetic wave which is propagated along the delay line of wave propagating tubes employed in hyper frequency described, among others, by J. R. Pierce in "Travelling Wave Tubes", Van Nostrand Co, 1950. The amplification occurs, as for these tubes, in a wide band around the central frequency.

In the foregoing, the rotor has been represented, for ease of description, with a structure having poles (reference **20** in the Figures). With the scope of the invention the rotor could be in the form of the well-known squirrel-cage structure of asynchronous machines. The above calculations remain valid in giving to the notion of resistance and to the notion of inductance the signification they have in this case.

The foregoing properties have been established by means of simplifying hypotheses which permit a clearer presentation. These properties apply generally to all machines whose stator and rotor would have the indicated electrical structure, even in the case of an operation which would differ slightly from these hypotheses in the neighbourhood of synchronous frequency. Moreover, the case was considered of a coupling by inductance between the rotor and the stator with a mutual coefficient **M**. Within the scope of the invention, this coupling could also be of an electrostatic nature between the conductors of the stator and rotor; similar equations would be obtained by substituting for the coefficient **M** the coefficient of electrostatic influence between the conductors in question, for the quantity $\mathbf{L_r}$ that, **Cr**, of the capacitance of each circuit of the rotor and, after interversion between $\mathbf{L_s}$ and $\mathbf{C_s}$. In this case $\mathbf{V_r}$ (equation 7) represents the voltage drop at the terminals of the capacitance of the rotor. The machine of this invention is also capable of operating as a generator, as will be seen below. A few embodiments of the structure of the machine of this invention will be given here.

Fig.5

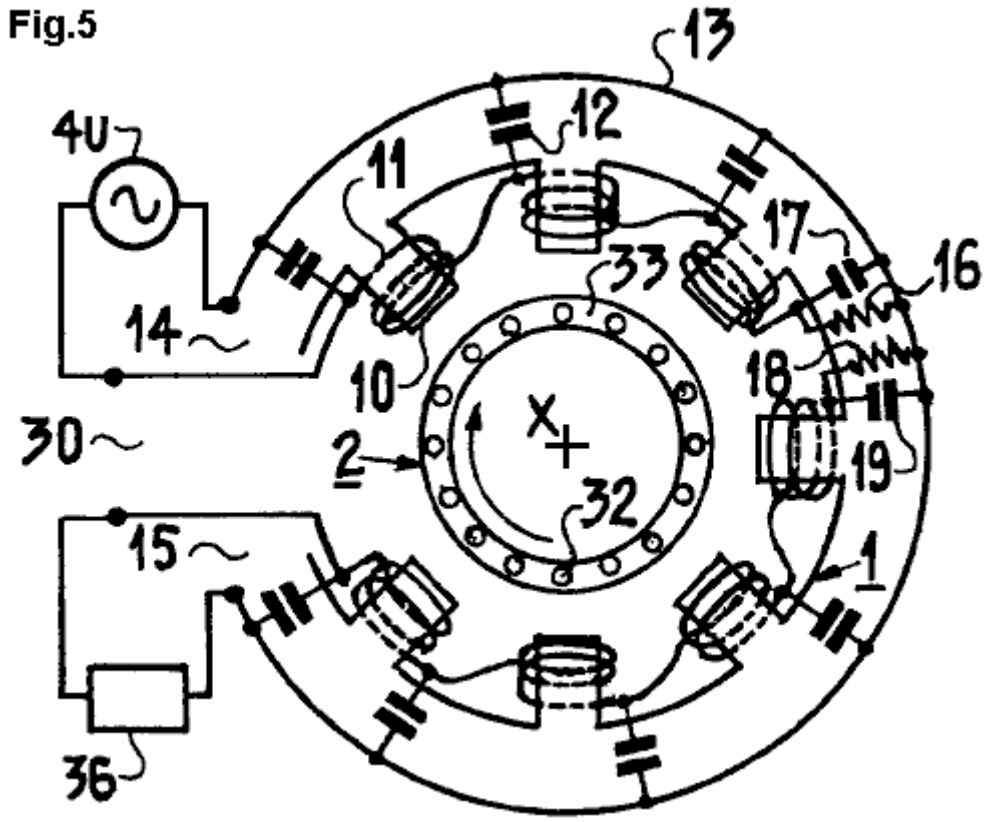


Fig.5 shows a first modification of this structure comprising a stator in accordance with the low-pass filter diagram of **Fig.2**. The rotor is a squirrel-cage, the reference numerals **32** and **33** are the bars and the end faces. In operation, it is driven by a rotor (not shown) which rotates it in the direction of the arrow about axis **X** which is common to the rotor and the stator of the machine. A source of alternating voltage **40** is applied to input **14** of the

stator, and applied to the output **15** is a load **36**, the impedance of which is equal to the characteristic impedance of the line of which the stator is part. In this arrangement of revolution about the axis **X**, a decoupling zone **30** separates the input and output of the stator. Further, in this zone, and in order to avoid any risk of coupling between input and output of the stator by the circuits of the rotor, there is provided any damping device considered necessary, an embodiment of which is given below.

The different elements of the filter constituting the stator are damped by the resistors **16**, and **18** which are connected as shown in the Figure between the windings **11** and the conductor **13** common to the terminals of the capacitors **17** and **19**.

Such a machine operates as an amplifier of the signal applied to the input of the stator with a gain which is of the order of 40 db in the numerical example given above. Such machines may be used as supply sources for high-power vibrators for the testing of industrial equipment of all kinds. They have the advantage over presently-known installations of this type of avoiding the steep leading edges and the high frequencies which result in their spectrum. Output powers of 100 kilowatts may be obtained with 10 watts applied at the input, with frequencies ranging up to 50 kHz. The same machine may be used as a high-power amplifier in radio broadcasting.

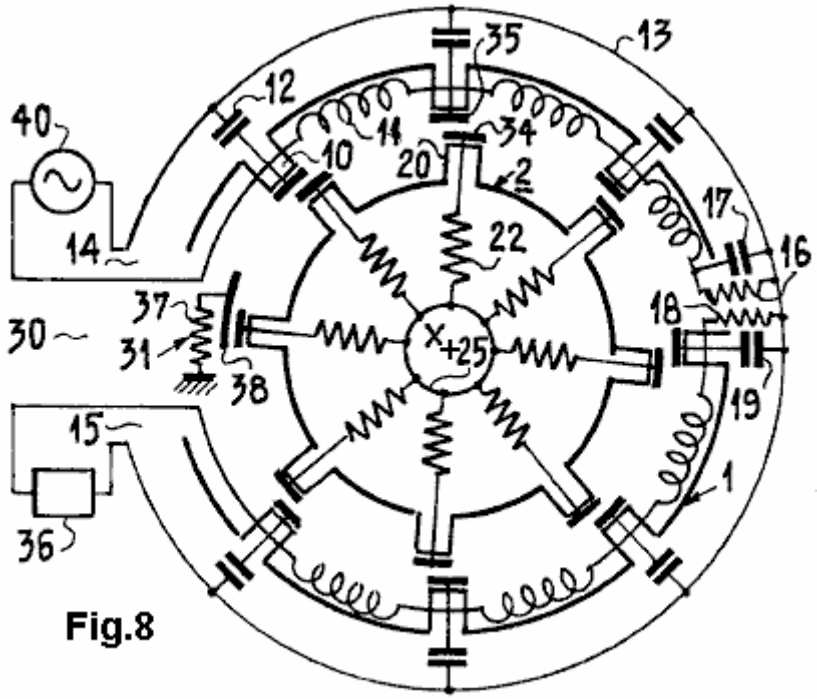


Fig.8

Fig.7 and **Fig.8** give two other modifications of the arrangement of the circuits of the rotor in which the same reference numbers indicate the same components as in the preceding Figures. In **Fig.8**, the coupling between the rotor and stator is electrostatic. It occurs between conductors **34** and **35** when they face each other in the course of the rotation of the rotor.

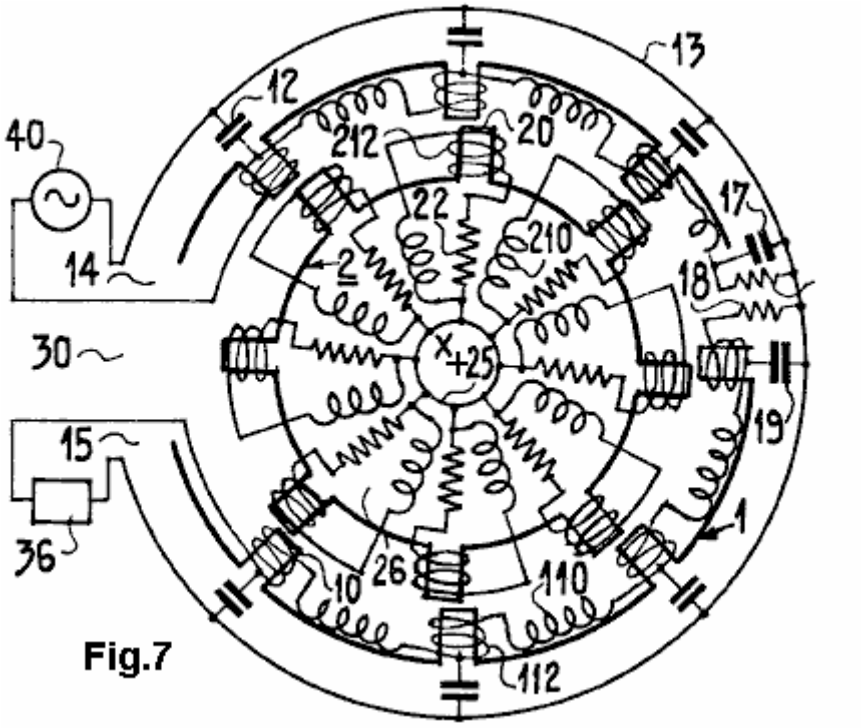


Fig.7

In **Fig.7, 110** and **112** designate the two constituent parts of the self-inductances associated with each one of the poles **10** of the stator, and **210** and **212** the parts of the self-inductances of each pole **20** of the rotor. In both **Fig.7** and **Fig.8**, **25** is a common conductor and **26** indicates all of the elements associated with each pole of the rotor. In the modification of **Fig.8**, an example is given of the construction of the damping device **31** in the zone **30**. A rail **38** is earthed through resistor **37** and through a contact (not shown) of the circuits of the rotor out along this rail **38** when they pass through the zone **30**.

The same machine is capable of operating as a generator (**Fig.6** and **Fig.9**).

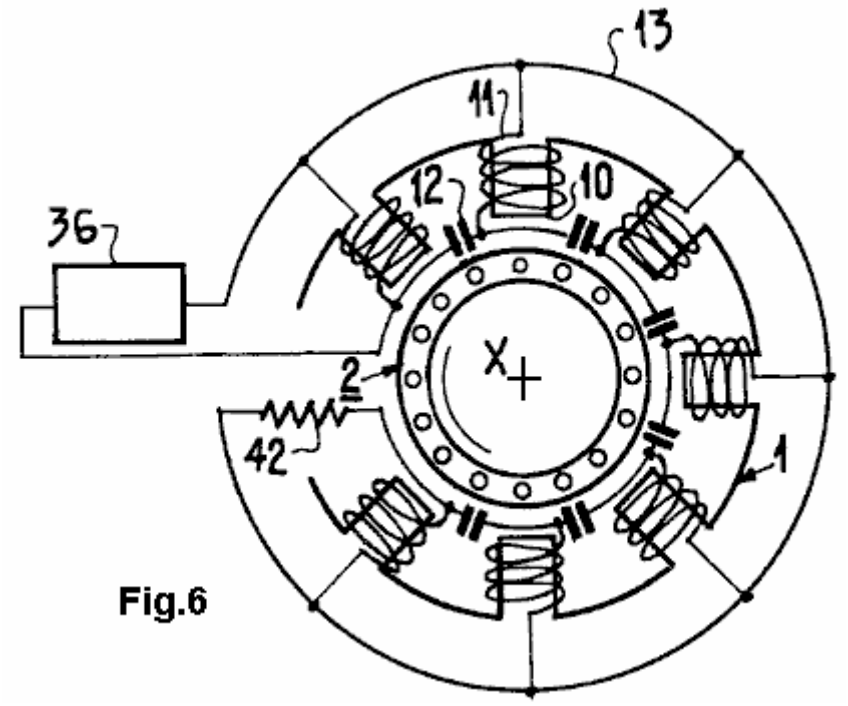


Fig.6

Fig.6 represents a machine whose stator is connected as a high-pass filter: each pole winding is connected at one end to the common conductor **13** and at the other end to two capacitors **12**, as shown above. The filter thus constructed has a phase velocity which varies very rapidly with the frequency and in the direction opposed to that of the flux of electrical energy (inverse wave). The rotor has a velocity rather close to that of the phase velocity whereas the energy flows back to the load **36** in the opposite direction. The upstream side (for the energy) of the filter terminates on the characteristic impedance **42** of the filter. A generator is obtained in this way having a

frequency which is determined within a wide band by the speed of rotation of the rotor. In the example shown in the Figure, the rotor has a squirrel-cage structure.

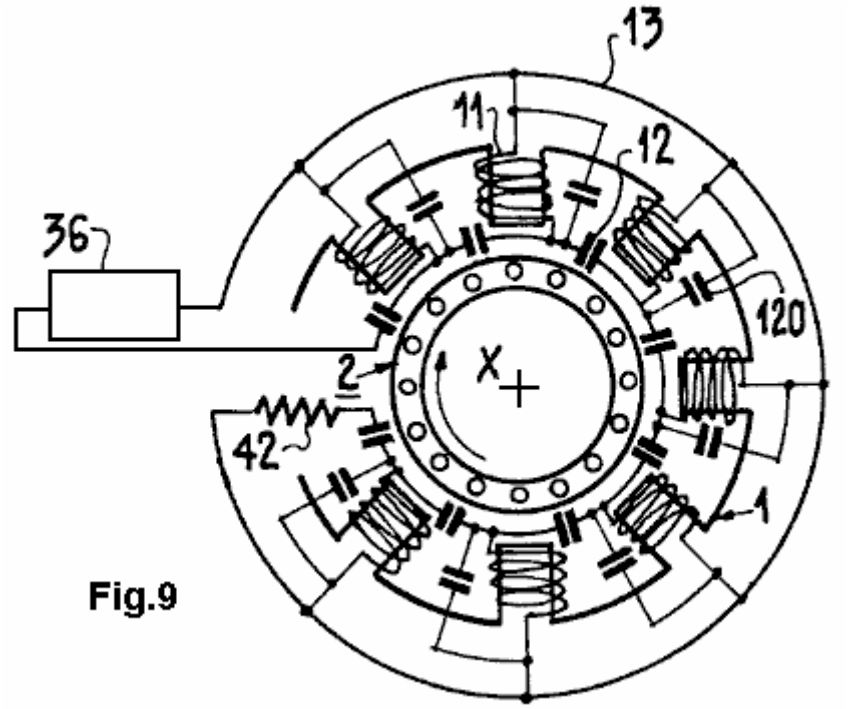


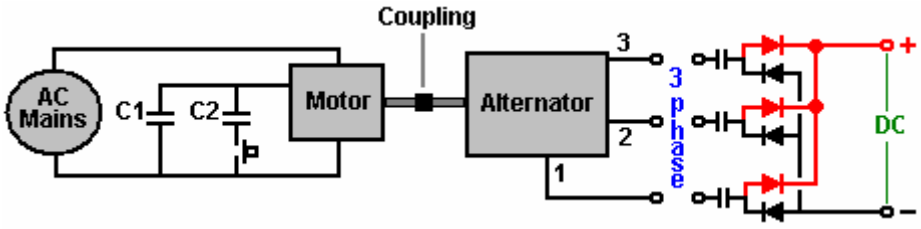
Fig.9

The machine of **Fig.9** differs from that of **Fig.6** by the addition of supplementary capacitors **120** arranged in parallel with the inductances **11**. The circuit thus obtained is an inverse wave band-pass filter. The width of its band-pass is determined by the respective values of the capacitors **12** and **120**. The phase velocity remains rapidly variable as a function of the frequency. The machine operates as a generator, the frequency of which, depends only very slightly on the velocity of rotation of the rotor.

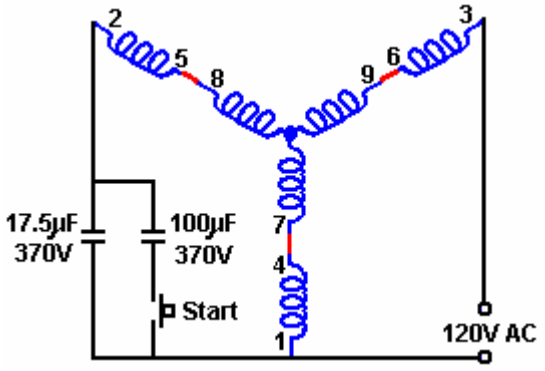
The "RotoVerter" Power Amplification System.

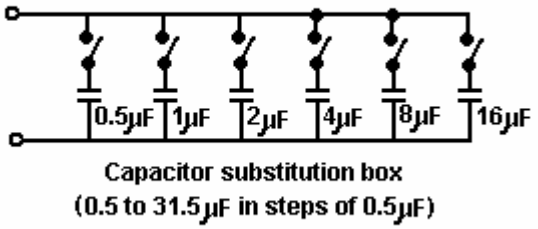
Not all pulsed-drive systems use permanent magnets as part of their drive mechanism. For example, the RotoVerter, designed by Hector D Peres Torres of Puerto Rico, and which has been reproduced by several independent researchers, producing at least 10 times more output power than the input power, uses standard three-phase electric motors instead of magnets.

This system has been reproduced by several independent researchers and it produces a substantial power gain when driving devices which need an electrical motor to operate. At this time, the web site: http://panacea-bocaf.org/rotoverter.htm has considerable details on how to construct the device as do the http://www.scribd.com/doc/2965018/HighEfficiencyForElectricMotors and the http://www.scribd.com/doc/26347817/RV-Energy-Saving-X documents. The outline details are as follows:



The output device is an alternator which is driven by a three-phase mains-powered, 3 HP to 7.5 HP motor (both of these devices can be standard 'asynchronous squirrel-cage' motors). The drive motor is operated in a highly non-standard manner. It is a 240V motor with six windings as shown below. These windings are connected in series to make an arrangement which should require 480 volts to drive it, but instead, it is fed with 120 volts of single-phase AC. The input voltage for the motor, should always be a quarter of its rated operational voltage. A virtual third phase is created by using a capacitor which creates a 90-degree phase-shift between the applied voltage and the current.





Capacitor substitution box
(0.5 to 31.5µF in steps of 0.5µF)

The objective is to tune the motor windings to give resonant operation. A start-up capacitor is connected into the circuit using the press-button switch shown, to get the motor up to speed, at which point the switch is released, allowing the motor to run with a much smaller capacitor in place. Although the running capacitor is shown as a fixed value, in practice, that capacitor needs to be adjusted while the motor is running, to give resonant operation. For this, a bank of capacitors is usually constructed, each capacitor having its own ON/OFF switch, so that different combinations of switch closures give a wide range of different overall values of capacitance. With the six capacitors shown above, any value from 0.5 microfarad to 31.5 microfarad can be rapidly switched to find the correct resonant value. These values allow combined values of 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, ...by selecting the appropriate switches to be ON or OFF. Should you need a value greater than this, then wire a 32 microfarad capacitor in place and connect the substitution box across it to test higher values step by step to find the optimum value of capacitor to use. The capacitors need to be powerful, oil-filled units with a high voltage rating - in other words, large, heavy and expensive. The power being handled in one of these systems is large and setting one up is not without a certain degree of physical danger. These systems have been set to be self-powered but this is not recommended, presumably because of the possibility of runaway with the output power building up rapidly and boosting the input power until the motor burns out.

The Yahoo EVGRAY Group at http://groups.yahoo.com/group/EVGRAY has a large number of members many of whom are very willing to offer advice and assistance. A unique jargon has built up on this forum, where the motor is not called a motor but is referred to as a “Prime Mover” or “PM” for short, which can cause confusion as “PM” usually stands for “Permanent Magnet”. RotoVerter is abbreviated to “RV” while “DCPMRV” stands for “Direct Current Permanent Magnet RotoVerter” and “trafo” is a non-standard abbreviation for “transformer”. Some of the postings in this Group may be difficult to understand due to their highly technical nature and the extensive use of abbreviations, but help is always available there.

To move to some more practical construction details for this system. The motor (and alternator) considered to be the best for this application is the “Baldor EM3770T” 7.5 horsepower unit. The specification number is 07H002X790, and it is a 230/460 volts 60Hz 3-phase, 19/9.5 amp, 1770 rpm, power factor 81, device.

The Baldor web site is www.baldor.com and the following details should be considered carefully before trying any adaption of an expensive motor. The following constructional photographs are presented here by kind permission of Ashweth of the EVGRAY Group.

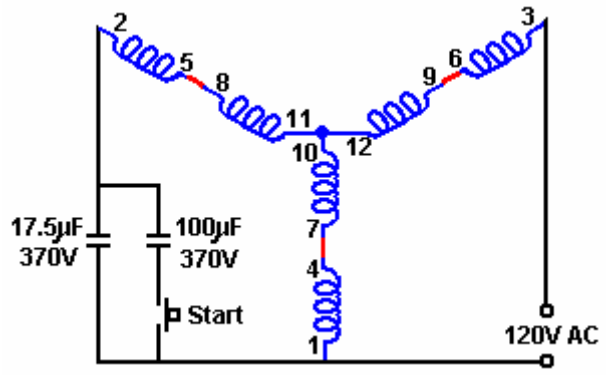
The end plate of the drive motor needs to be removed and the rotor lifted out. Considerable care is needed when doing this as the rotor is heavy and it must **not** be dragged across the stator windings as doing that would damage them.

The second end-plate is then removed and placed on the opposite end of the stator housing:

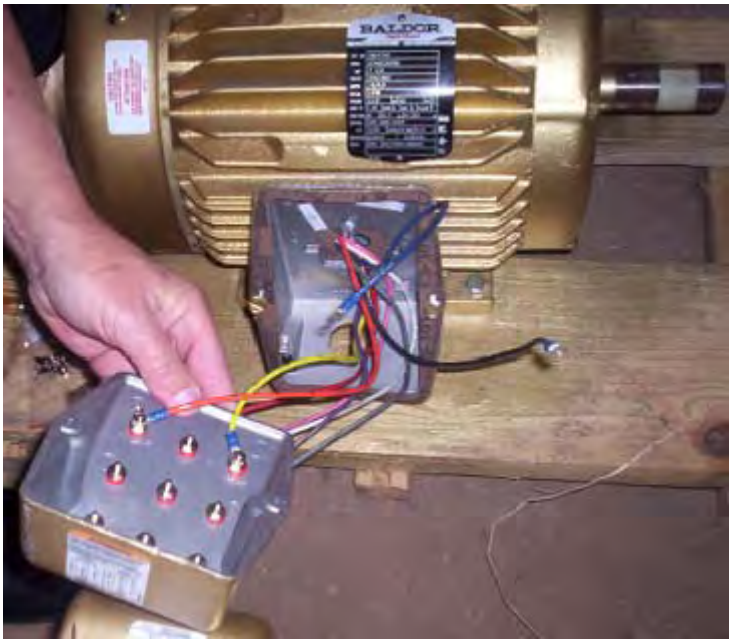
The fan is removed as it is not needed and just causes unnecessary drag, and the rotor is inserted the opposite way round to the way it was removed. That is, the housing is now the other way round relative to the rotor, since the rotor has been turned through 180 degrees before being replaced. The same part of the shaft of the rotor passes through the same end plate as before as the end plates have also been swapped over. The end plates are bolted in position and the rotor shaft spun to confirm that it still rotates as freely as before.

To reduce friction to an absolute minimum, the motor bearings need to be cleaned to an exceptional level. There are various ways of doing this. One of the best is to use a carburettor cleaner spray from your local car accessories shop. Spray inside the bearings to wash out all of the packed grease. The spray evaporates if left for a few minutes. Repeat this until the shaft spins perfectly, then put one (and only one) drop of light oil on each bearing and do not use WD40 as it leaves a residue film. The result should be a shaft which spins absolutely perfectly.

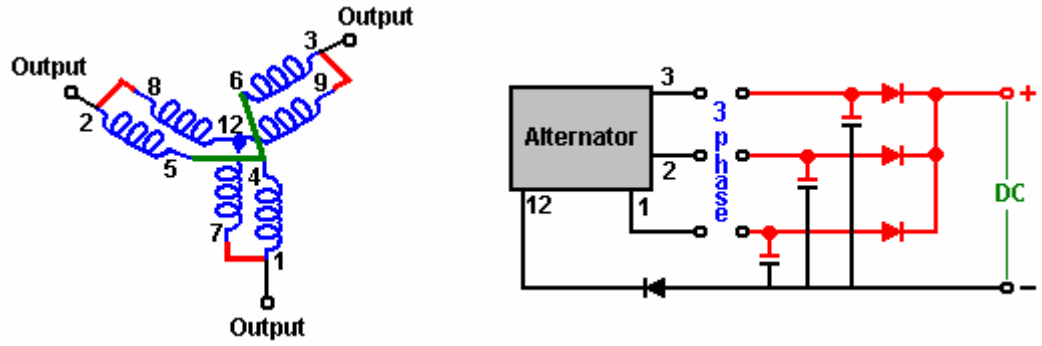
The next step is to connect the windings of the two units. The motor (the “Prime Mover”) is wired for 480 volt operation. This is done by connecting winding terminals 4 to 7, 5 to 8 and 6 to 9 as shown below. The diagram shows 120 volts AC as being the power supply. This is because the RotoVerter design makes the motor operate at a much lower input than the motor designers intended. If this motor were operated in the standard way, a 480 volt 3-phase supply would be connected to terminals 1, 2 and 3 and there would be no capacitors in the circuit.



It is suggested that the jumpering of the motor windings is more neatly done by removing the junction box cover and drilling through it to carry the connections outside to external connectors, jumpered neatly to show clearly how the connections have been made for each unit, and to allow easy alterations should it be decided to change the jumpering for any reason.



The same is done for the unit which is to be used as the alternator. To increase the allowable current draw, the unit windings are connected to give the lower voltage with the windings connected in parallel as shown below with terminals 4,5 and 6 strapped together, 1 connected to 7, 2 connected to 8 and 3 connected to 9. This gives a three-phase output on terminals 1, 2 and 3. This can be used as a 3-phase AC output or as three single-phase AC outputs, or as a DC output by wiring it as shown here:



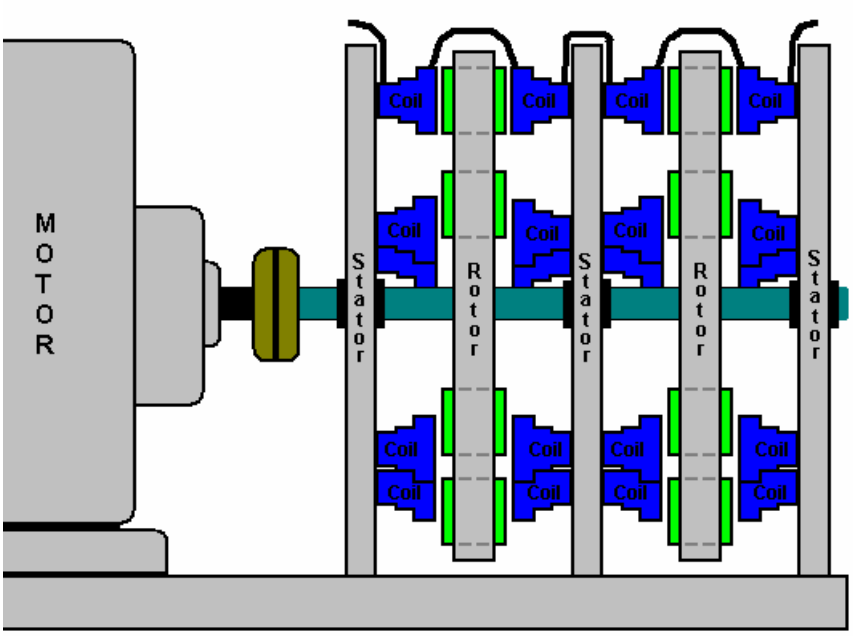
The motor and the alternator are then mounted securely in exact alignment and coupled together. The switching of the direction of the housing on the drive motor allows all of the jumpering to be on the same side of the two units when they are coupled together, facing each other:

The input drive may be from an inverter driven from a battery charged via a solar panel. The system how needs to be 'tuned' and tested. This involves finding the best 'starting' capacitor which will be switched into the circuit for a few seconds at start-up, and the best 'running' capacitor.

To summarise: This device takes a low-power 110 Volt AC input and produces a much higher-power electrical output which can be used for powering much greater loads than the input could power. The output power is much higher than the input power. This is free-energy under whatever name you like to apply to it. One advantage which should be stressed, is that very little in the way of construction is needed, and off-the-shelf motors are used. Also, no knowledge of electronics is needed, which makes this one of the easiest to construct free-energy devices available at the present time. One slight disadvantage is that the tuning of the "Prime Mover" motor depends on its loading and most loads have different levels of power requirement from time to time. A 220 Volt AC motor can also be used if that is the local supply voltage.

If an alternator is being driven by the RotoVerter motor (the "Prime Mover") but although the shaft is being rotated rapidly there is no output voltage, then it is likely that the alternator has been sitting around unused for a long time and has lost the magnetic properties which it needs at start-up. To fix this, connect each of the three output windings, one at a time, across a car battery for about five seconds to develop some magnetism and the alternator will then work. This is a one-off thing only needed after a long period of inactivity.

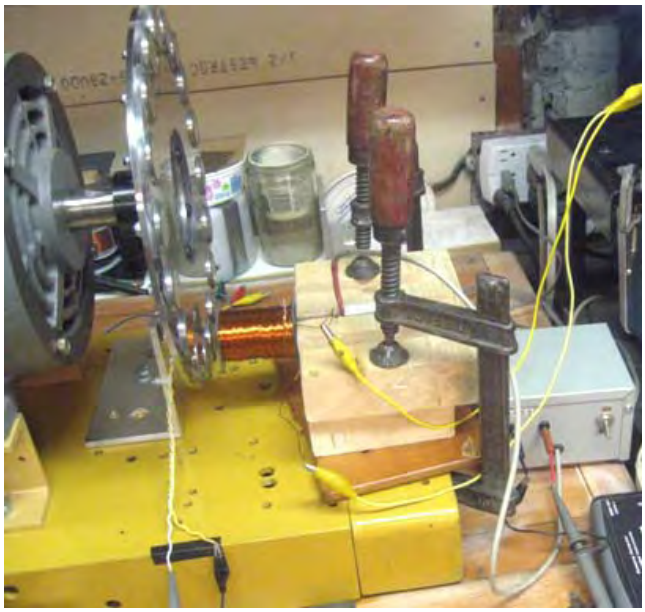
It is not essential to construct the RotoVerter exactly as shown above, although that is the most common form of construction. The Muller Motor mentioned earlier, can have a 35 kilowatt output when precision-constructed as Bill Muller did. One option therefore, is to use one Baldor motor jumpered as the "Prime Mover" drive motor and have it drive one or more Muller Motor style rotors to generate the output power:



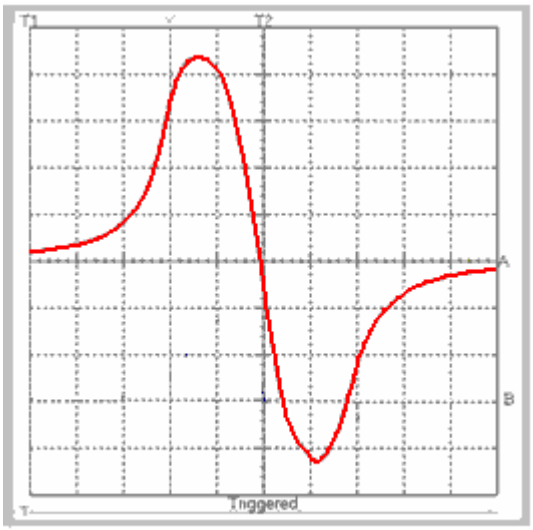
Power Boosting Through Coil Short-Circuiting.

The RotoVerter output and the Muller Motor/Generator output (and possibly, the Adams Motor output) can be increased very substantially by a technique developed by "Kone" the moderator of the EVGRAY Yahoo forum already mentioned. The technique is to place a dead short-circuit across each output coil, just as the magnetic field of that coil reaches a maximum. This is done five times in rapid succession and can boost the output power by an estimated factor of 100 times.

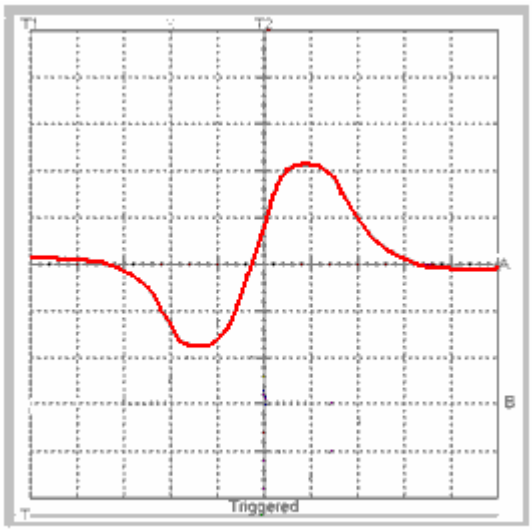
On the surface, it appears complete madness to place a short-circuit across the very output which you are generating as the whole point of the exercise. However, it is not as mad as it seems. At the peak point, the coil itself contains a large amount of energy and when a short-circuit is placed across it, the result is quite unusual. The effect of short-circuiting an iron-cored coil has been demonstrated by Ron Pugh of Canada with a bench test using this equipment:



Here, a device for measuring magnetic fields is sensing the magnetic field of the coil as the rotor magnets move past the coil. The measurement is done with the coil operating normally and then again, with the coil short-circuited. The results are shown in these oscilloscope displays:

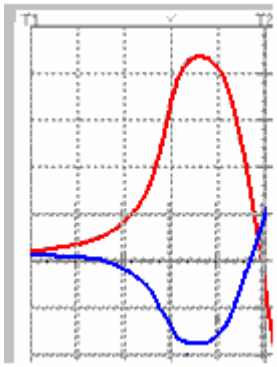


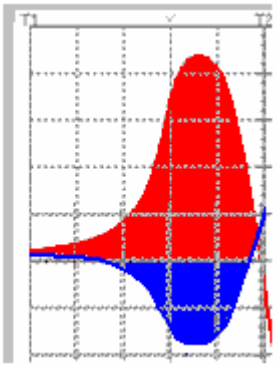
OPEN COIL



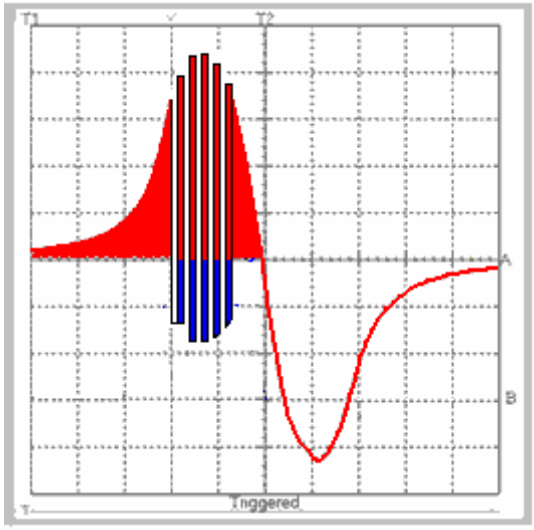
SHORT–CIRCUITED COIL

Quite surprisingly, the magnetic field is reversed by the short-circuit. If we consider just the first half of the cycle:





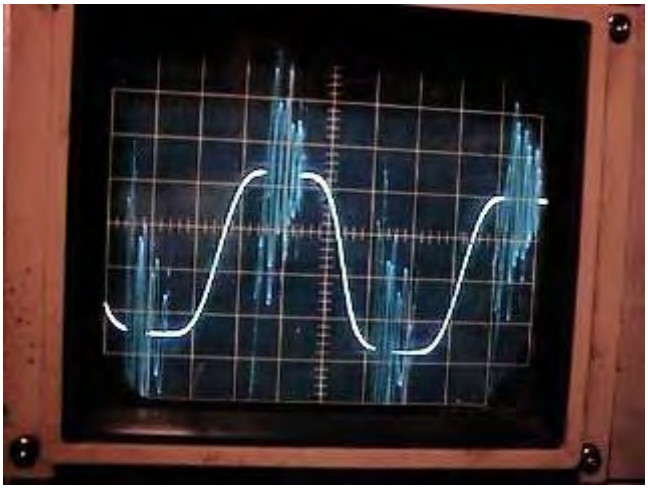
You will notice that when the open coil has a very strong positive position (relative to a North magnetic pole), the short-circuited coil has a strong reading of the opposite polarity. Therefore, at any given instance near the peak, there is the potential for a major magnetic reversal if the coil were to be switched from the red area into the blue area and back again. You might imagine that if the coil short-circuiting were performed very rapidly, that there would be a result like this:



However, this is not really possible with an iron-cored coil as it is not able to reverse it's magnetisation rapidly enough to produce this effect. Coils with iron cores might get up to 3,000 reversals per second although 1,000

would probably be a more realistic figure. For higher frequencies, a ferrite core could be used and for yet higher frequencies again, an iron-dust epoxy-encapsulated core is needed. For unlimited frequencies, an air-core coil is used.

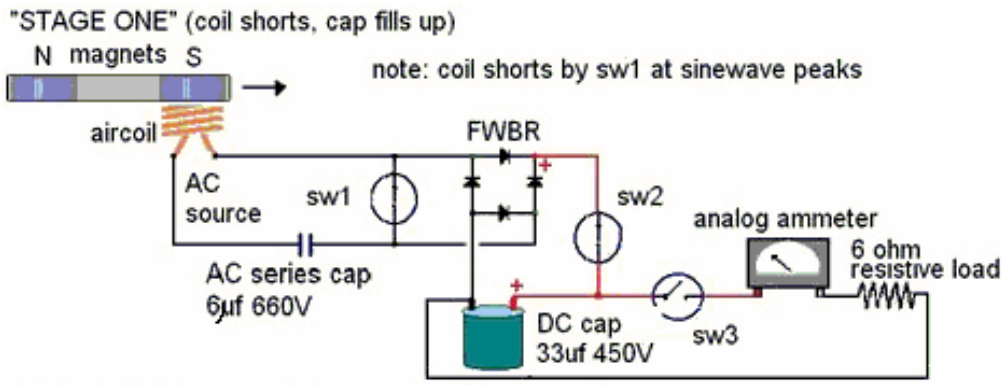
In the RotoVerter / Muller Motor example where Kone has demonstrated major energy gains, the arrangement is different to Ron Pugh's bench-test example. Firstly, the magnets on the rotor present two poles to the coil as they pass by, giving a full, even sine wave output. Secondly, Kone uses an air-core coil and he has the fast switching implemented to take advantage of that coil type:



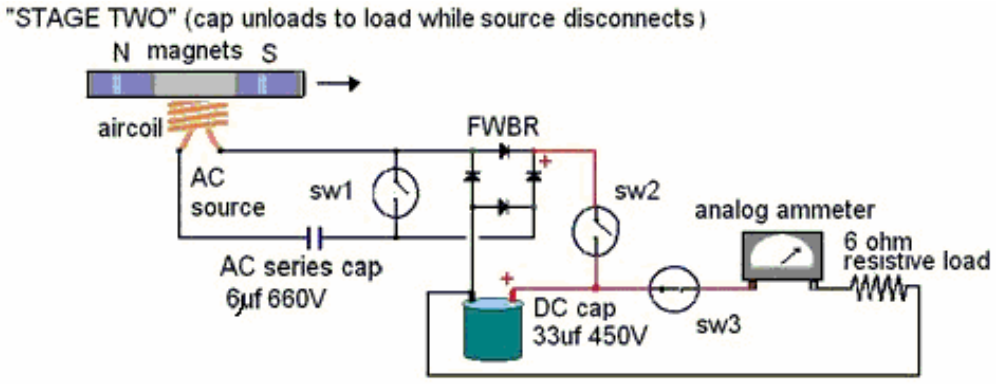
This screen shot shows exactly the same effect with the oscillating pulses plunging down to exactly where the negative trace would be at that time during each of the short-circuited periods. The display is showing a twenty-volt peak-to-peak waveform with each complete sine wave cycle taking 2 milliseconds.

Instead of the original gradual magnetic rise to a single peak, there are now five very sharp magnetic reversals, each of which are substantially larger that the original peak. It is the change of magnetic flux in the pick-up coil which produces the output power, so it can be seen that with this additional switching, a massive increase in flux change has been produced in the output coil. This increase is both a larger magnetic swing and a much greater rate of change of the flux, and as the rotor is spinning at some 1,800 rpm and has many magnets in it, the overall magnetic power increases by a major factor. Please note that in the following diagrams produced by Kone, the rotor magnet has a South pole which reaches the pick-up coil first, followed by a North pole passing the coil. This produces an excellent sine wave output in the coil.

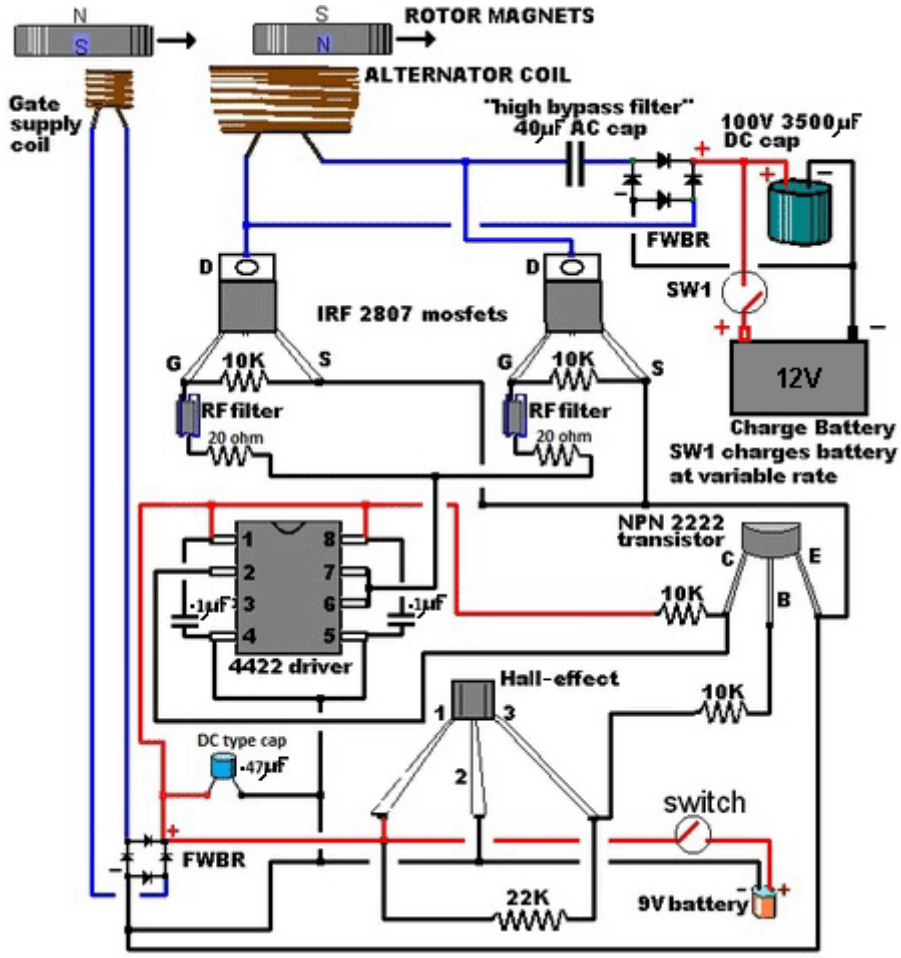
The additional switching is performed by a mechanical brush and contact system and the power collection circuit used for the short-circuited coil is:



and when the coil is not short-circuited, the circuit is:



Doug Konzen has been developing this circuitry and generously sharing his results freely. His web site is at http://sites.google.com/site/alternativeworldenergy/shorting-coils-circuits and his most recent practical circuit is:



There is a video on the web where coil-shorting is used on the output from a replication of John Bedini's Window Motor. John's Window Motor is a cylinder with magnets mounted in it, spun inside a large coil by a pulsed motor:

The motor can be a self-runner, but for the coil-shorting demonstration, it was just spun briefly by hand, producing voltage pulses of about 16 volts. When the coil-shorting is switched on, those pulses rise to about 440 volts even

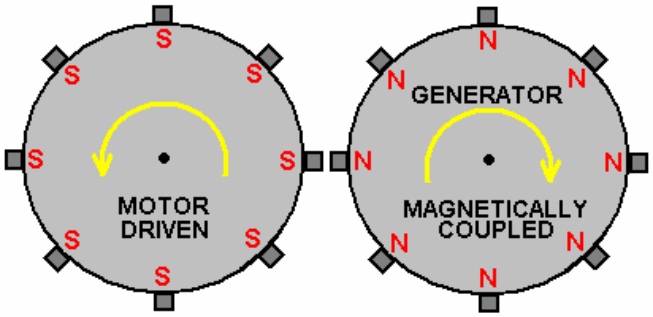
though the coil shorting was not the optimum five times at peak (which would probably have raised the voltage pulses to about 1,600 volts). The oscilloscope displays of the test shown on video are:

I'm told that this coil-shorting technique is used in industry but is consider to be a 'trade secret'. Details of John Bedini's Window Motor can be found at http://www.fight-4-truth.com/Schematics.html.

Raoul Hatem's Magnetic Coupling System.

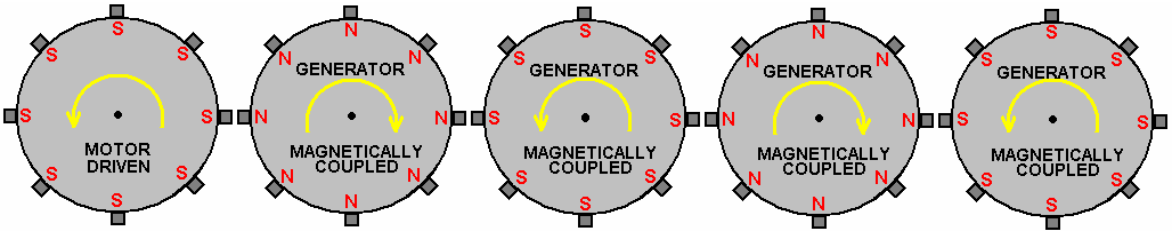
Generally, the RotoVerter has very low input power when not loaded and about a 90% energy reduction when under load. The ideal situation is where there is a constant load as the tuning of the RotoVerter does depend to some degree on the load. However, the RotoVerter performance can be increased very substantially by using the techniques introduced by Raoul Hatem in 1955, which conventional science will not accept because according to current theory, any such energy gain has to be "impossible" and so, cannot happen no matter what evidence there is:

Raoul Hatem's heretical statement is that using spinning magnets draws in energy from the environment, allowing a system to have COP>1 (ever heard of John Searle's spinning magnet systems?). His method is to use a motor (whether RotoVerter or not) to spin a heavy rotor disc with 36 powerful rare-earth magnets mounted on it. Then, using an identical heavy disc with magnets mounted on a generator to give a magnetic coupling between the motor and the generator can give not just coupling, but an energy gain as well:

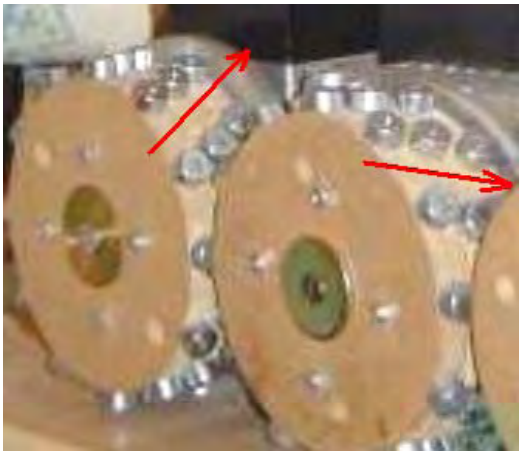


The heavy rotor provides some flywheel effect which helps with the operation of the system. Even with one motor as shown above, there is an energy gain as demonstrated in a recent video demonstration of the effect at http://www.dailymotion.com/video/xi9s9b_moteur-magnetique-de-leon-raoul-hatem_webcam#.UaGyVTcr6Bo where a simple system produces 144 watts of excess power. However, the really big gains are got when several generators are driven by just the one motor. In passing, it may be remarked that there are two separate energy gain systems operating here. Firstly, the rotating magnetic field acts directly on the excess electrons in the local environment, drawing them into the system just as the fluctuating magnetic field of the secondary winding of any transformer does. Secondly, the rotors are receiving a rapid stream of drive pulses, and as Chas Campbell has demonstrated, that draws in excess energy from the gravitational field.

Anyway, you will notice that the powerful magnets used have their North poles outwards on one rotor while the adjacent rotor has the South poles outwards. The very strong attraction between these opposite poles cause the generator disc to rotate in step with the motor disc. This process allows many generators to be driven by just the one motor as shown here and in the photograph above:



For ease of drawing, the diagram above shows only eight magnets per rotor disc, but you will notice in the photograph (and in the video) that there are three stepped rows of magnets on each rotor:

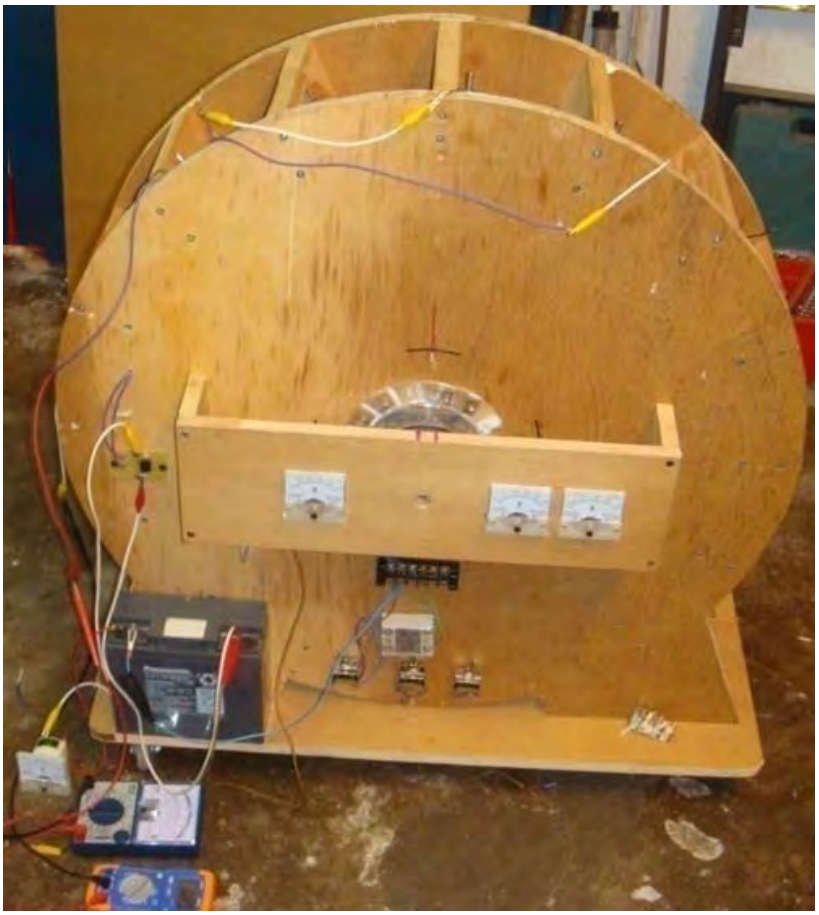


You will also notice that direction of the stepping is reversed on every second rotor disc in order for the magnets to match each other in position as they rotate in opposite directions.

While this type of arrangement gives a major increase in the output power compared to the input power, the situation can be further enhanced by retrieving some of the input power by means of electronic circuitry, and both Phil Wood and David Kousoulides have kindly shared their methods for doing this.

Lawrence Tseung's COP=3.3 Pulsed-Flywheel.

Lawrence has been presenting his theory of lead-out energy which indicates that excess energy is drawn from the environment when there is an impact. The method of producing this effect which he has followed is to create an unbalanced wheel and demonstrate that excess energy is produced. It should be stressed that energy is never created or destroyed and so, when he measures more energy in his device than the energy which he uses to power it, energy is not being created but is instead, being drawn in from the local environment. Lawrence has recently demonstrated a prototype to members of the public:



This simple device was demonstrated to have 3.3 times as much output power as the input power needed to make it operate. This is an early prototype which was demonstrated in October 2009 and Lawrence and his helpers are working on to produce more advanced models which have kilowatts of excess electrical power.

Mr Tseung remarks: "The **Lee-Tseung Lead-Out Energy Theory** was first disclosed to the world on 20th December 2004 at Tai Po, in Hong Kong. The Lead-Out Energy Theory basically says that one can lead-out (or bring-in) Energy from the surrounding environment into a Lead-Out Energy Machine. The total Input energy is equal to the sum of the Supplied Energy plus the Lead-Out Energy. For example, if the supplied energy is 100 units and the lead-out energy is 50 units, the device's total Input Energy will be 150 units. This means that the Output Energy can be more than the Supplied Energy of 100 units provided by the person using the device.

If we ignore the small loss of energy caused by less than 100% efficiency of the device itself, then the Output Energy will be the whole of the 150 units. If we use 50 of the output energy units and feed back 100 of the output units as the Supplied Energy, then that Supplied Energy can again lead-out another 50 units of excess output Energy for us to use. Thus a Lead-Out Energy Machine can continuously lead-out pollution-free, virtually inexhaustible and readily available energy for us to use. We do not need to burn any fossil fuel or pollute our

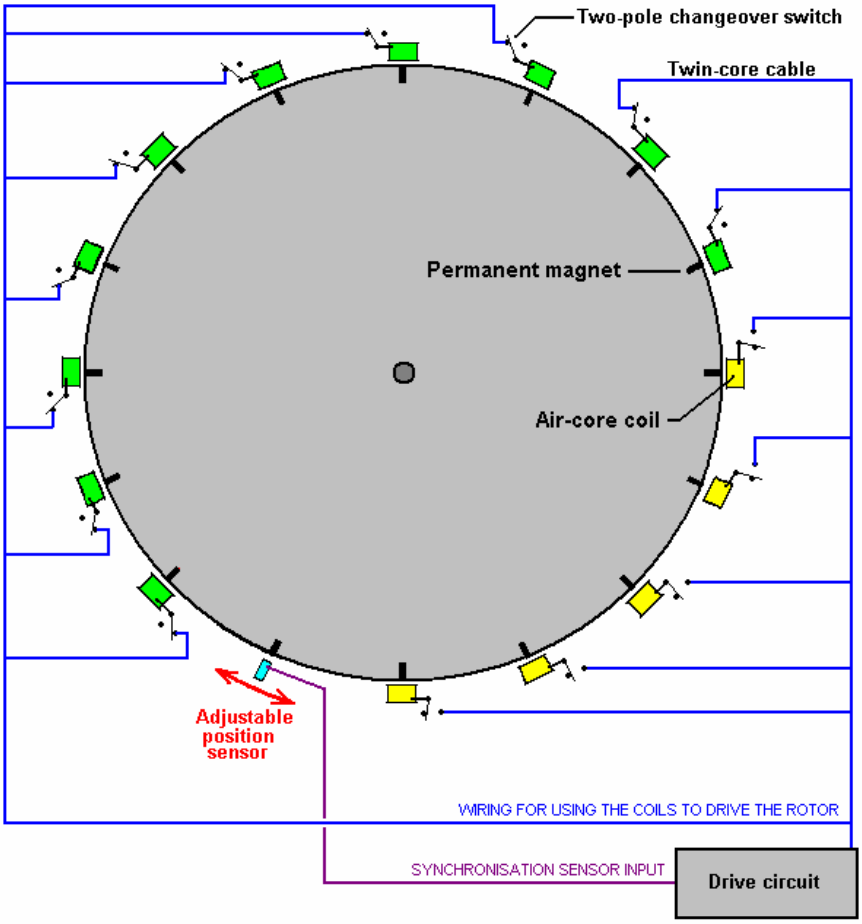
environment. The two examples of Lead-Out energy which we access are Gravitational and Electron-Motion energy.

The Lead-Out Energy theory does not violate the Law of Conservation of Energy. The Law of Conservation of Energy has been used as a roadblock for the so called “Overunity” devices. The patent offices and the scientific establishment routinely dismiss an invention as belonging to the impossible “perpetual-motion machine” category if the inventor cannot identify the energy source of his invention.

We got the help of Mr. Tong Po Chi to produce a 60 cm diameter Lead-Out Energy machine in October 2009. The Output Energy of that device is greater than the Input Energy by a factor of 3 times. These results are confirmed by voltmeters and ammeters measuring the Input and Output energies.

The Tong wheel has been shown at two Open Shows in Hong Kong (Inno Carnival 2009 and Inno Design Tech Expo) in November and December 2009. Over 25,000 people have seen it. The Better Hong Kong Radio Show has video recorded it, the discussions being conducted in Chinese. At this time, the Tong wheel is at the Radio Studio available for experts to view and examine with their own instruments."

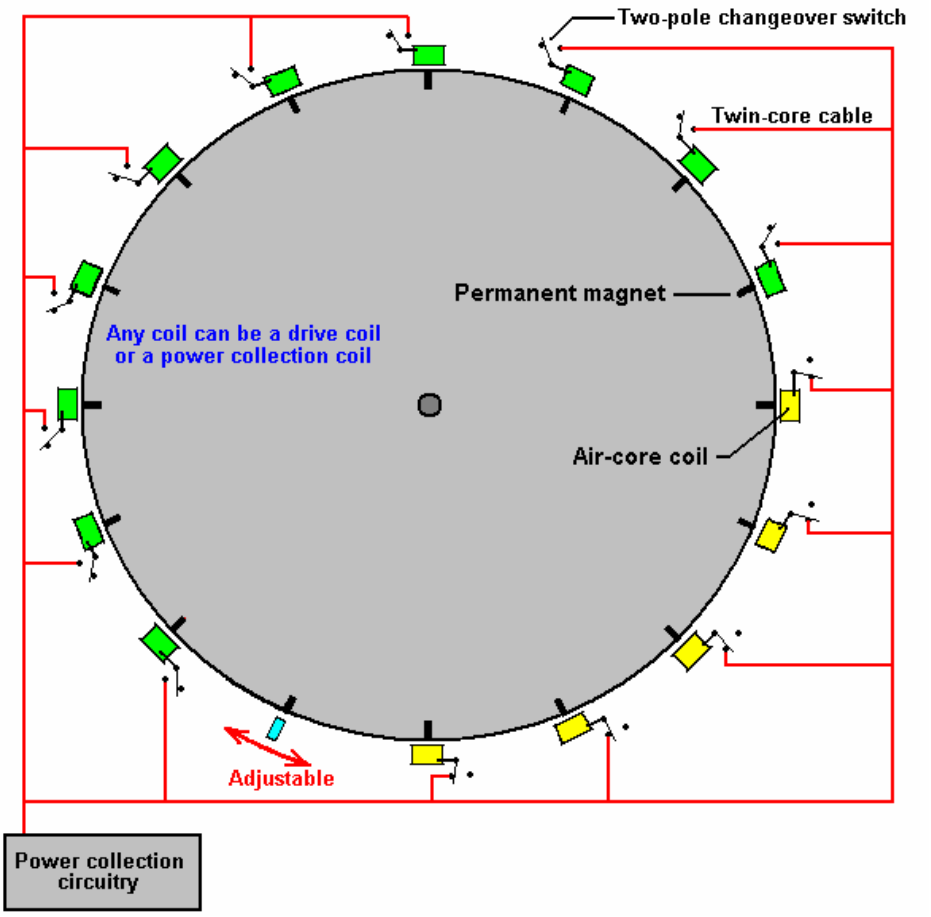
The Tong wheel has a diameter of 600 mm and this large size is considered to be important. It has 16 permanent magnets mounted on its rim and 15 air-core coils mounted around it on the stator. There is one position sensor. The coils can be switched to act as drive coils or as energy collection coils:



With this arrangement, if the positions the switches as shown for ten of the fifteen coils shown here, then they act as drive coils. The sensor is adjusted so that the drive circuit delivers a brief energising pulse to those coils just after the magnets have passed their exact alignment position with the coils. This causes them to generate a magnetic field which repels the magnets, thrusting the rotor around.

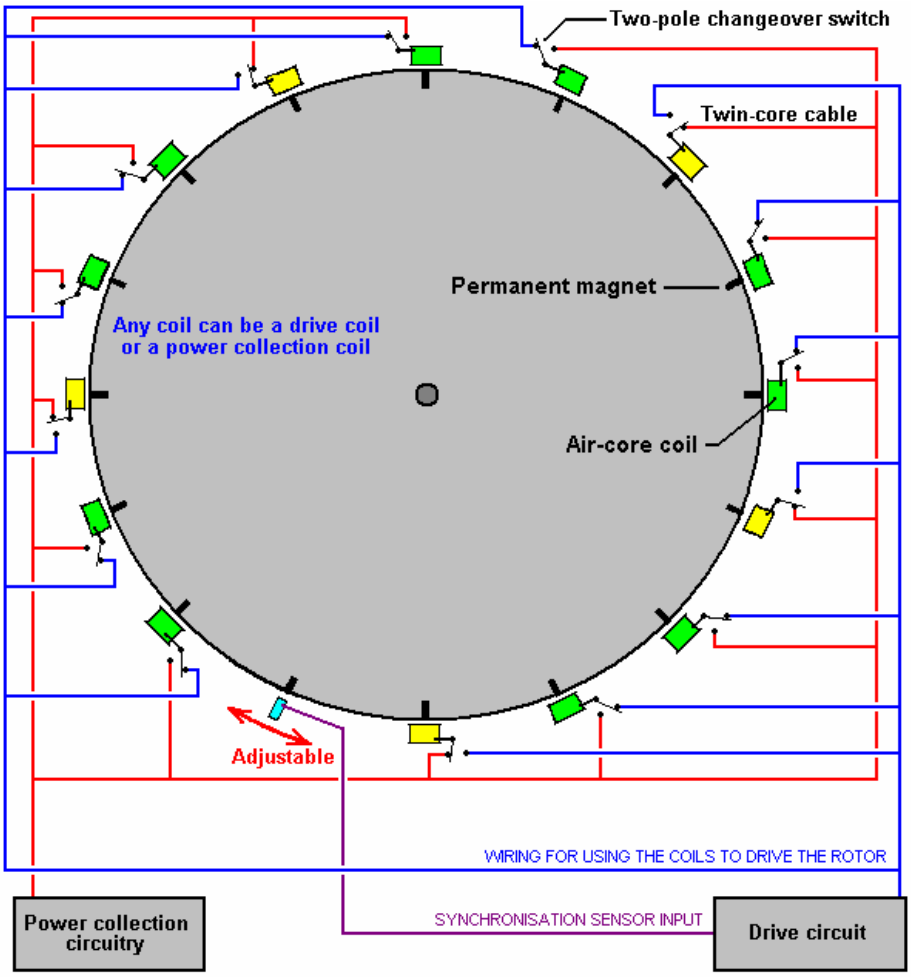
The pulse is very brief, so very little power is needed to accomplish this pulsing. As mentioned before, any number of coils can be switched to provide this driving force. With this particular wheel construction by Mr Tong, the best number has been found to be ten drive coils.

The power pick-up is achieved by gathering the electricity generated in some of the coils as the magnets move past them:



In this particular arrangement, five of the coils gather energy while ten provide the drive. For the sake of simplicity, the diagram shows the five collection coils adjacent to each other and while that would work, the wheel is better balanced if the drive coils are evenly spaced out around the rim. For that reason, this switching would actually be selected to give five sets of two drive coils followed by one pick-up coil as that gives a perfectly balanced thrust on the wheel.

The two diagrams above are shown separately in order to make it clear how the drive switching and the power pick-up switching are arranged. The full design arrangement and the balanced switching are shown in the following diagram which indicates how the full design is implemented on this particular implementation of the wheel design. The sensor can be a coil feeding a semiconductor switching circuit, or it can be a magnetic semiconductor called a Hall-effect device which can also feed a semiconductor circuit. An alternative would be a reed switch which is a simple mechanical switch encased in an inert gas inside a tiny glass envelope. Suitable switching circuits are described and explained in chapter 12 of this eBook.

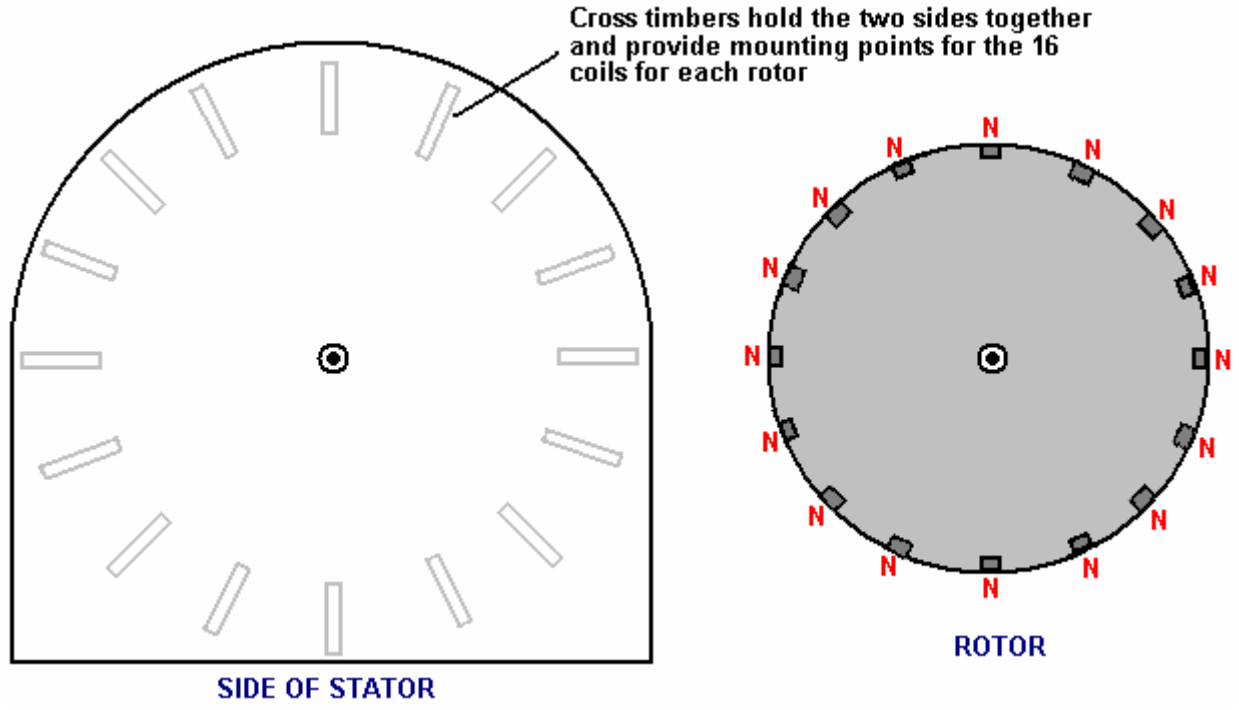


Mr Tseung remarks that the large wheel size is due to the fact that the Pulse Force takes time to impart the impulse to the wheel and lead-out energy from the environment into the system. If you want to see this actual wheel, you can email Dr. Alexandra Yuan at ayuan@hkstar.com to make an appointment. The Tong wheel is located at the Better Hong Kong Radio Studio in Causeway Bay, Hong Kong. Just say that you want to see the Lead-Out Energy Machine. The demonstration can be in English or in Chinese. Ideally, there should be a group of at least six visitors with one or more being a qualified engineer or scientist, and you are welcome to bring your own cameras and/or test equipment. It is planned to produce a version which has a 300 watt output, and another with a 5 kilowatt output. Educational kits are also planned.

If you decide to replicate this particular design, then to raise the output power level you might consider putting another set of coils around the wheel and either using them as fifteen additional energy pick-up coils or alternatively, pulsing the wheel twice as often. Adding one or more additional rotor discs to the same rotating shaft is also an option and that has the advantage of increasing the rotor weight and improving the effect of the impulses on the rotor.

The diameter of the wire used to wind the coils is a design choice which has a wide scope. The thicker the wire, the greater the current and the larger the impulse given to the wheel. The coils are normally connected in parallel as shown in the diagrams.

Because of the way magnetic field strength drops off with the square of the distance, it is generally considered good design practice to make the coils one and a half times as wide as they are deep, as indicated in the diagrams above, but this is not a critical factor. This design is, of course, a version of the Adams motor described at the start of this chapter. Although motors of this kind can be built in many different ways, the construction used by Mr Tong has some distinct advantages, so here is a little more detail on how I understand the construction to be carried out.

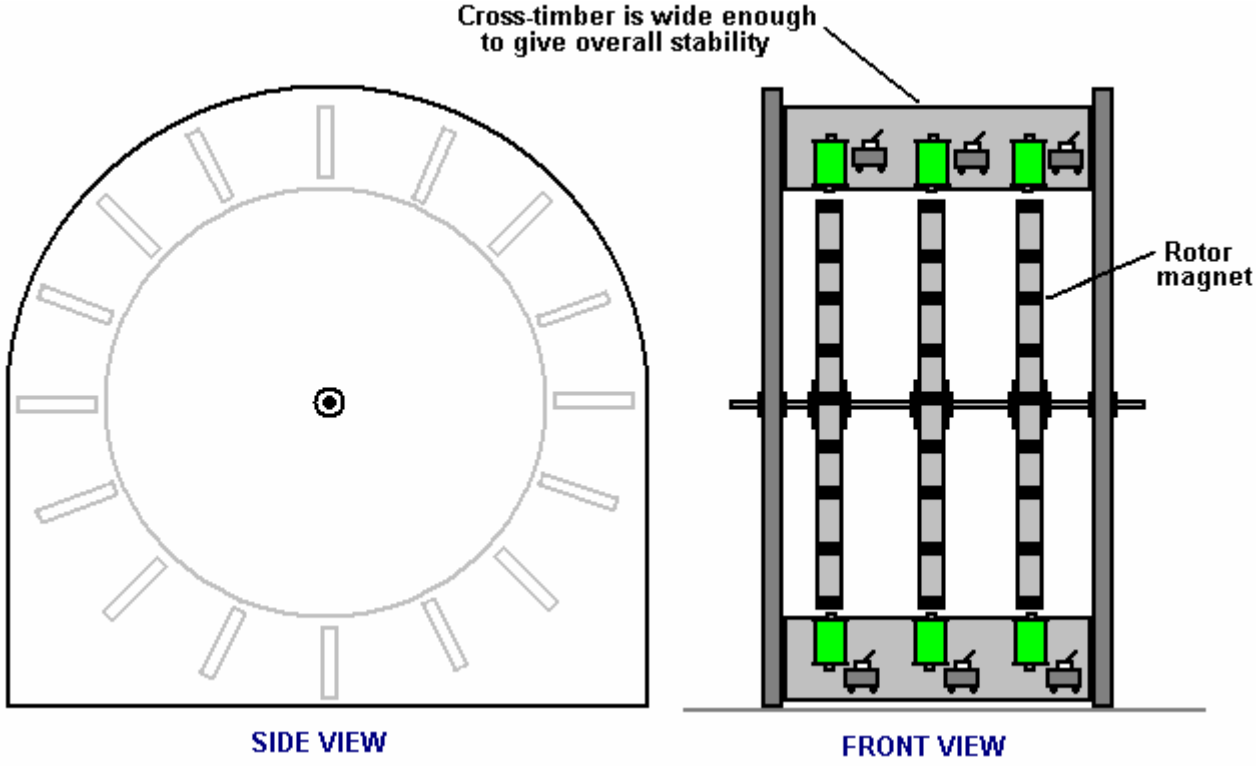


There are two side pieces which are attached together by sixteen cross timbers, each of which are held in place by two screws at each end. This produces a rigid structure while the construction method is as simple as is possible, using readily available materials which are worked with the most basic of hand tools. The construction also allows the motor to be taken apart completely without any difficulty, transported as a “flat-pack” package and then assembled at a new location. It also facilitates people who want to see the motor taken apart after a demonstration in order to assure themselves that there is no hidden power source.

Each of the cross timbers provide a secure mounting platform for an electromagnet and it’s associated switch. In the implementation by Mr Tong, there appears to be just the one rotor, configured as shown above with sixteen permanent magnets mounted in it’s rim. The magnetic poles of these magnets are all orientated in the same direction. That is to say, the magnetic poles facing outwards are all either South or all North poles. It is not critical whether the outward facing poles are North or South as Robert Adams used both arrangements with great success, but having said that, most people prefer to have the North poles facing outwards.

Robert has always said that one rotor was enough, but his techniques were so sophisticated that he was able to extract kilowatts of excess power from a single small rotor. For us, just starting to experiment and test a motor of this type, it seems sensible to stick with what Mr Tong has experienced success. However, this build by Mr Tong is not his final motor but just one in a series of continuously improved motors.

The following diagram shows an arrangement which has three rotors attached to a single shaft and while you may choose to construct this with just one rotor, if the cross timbers are long enough, then one or two extra rotors can be added in very easily at a later date.



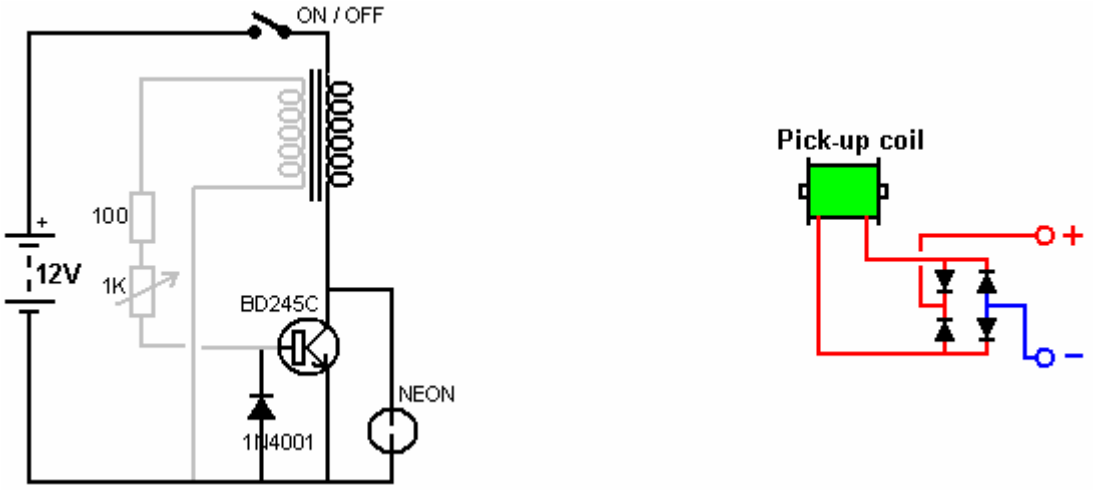
Here, just two of the cross timbers are shown. The electromagnet coils used by Mr Tong are air-core as that type have the least effect on the passing magnets. However, electromagnets with cores tend to be much more power for any given current flowing through them. In theory, the core should be made of lengths of insulated iron wire as that would reduce power loss through eddy currents flowing in the core, but Robert actually recommends solid metal cores, and as he was the most experienced person in this field, paying attention to what he said seems sensible.

The core material needs to be a metal which magnetises easily and powerfully, but which does not retain any of its magnetism when the current stops flowing. Not many metals have those characteristics and soft iron is usually recommended. Nowadays, soft iron is not always readily available and so a convenient alternative is the central bolt of a masonry anchor which has excellent properties:

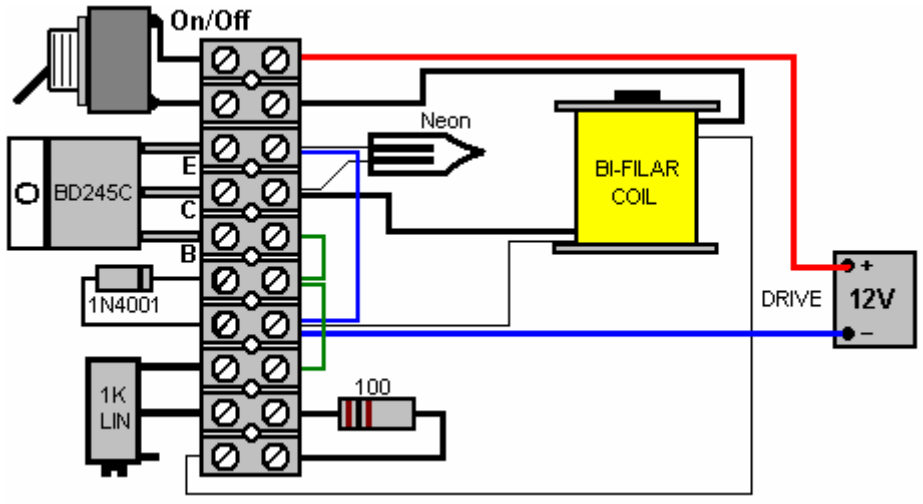
The shaft of the bolt can be cut quite easily with a hacksaw, but be sure to remove (or file down) the head of the bolt as the increase in diameter has a marked effect on the magnetic properties of the electromagnet core if it is left in place. The bolt shown above is a M16 x 147 mm masonry anchor bolt with a bolt diameter of 10 mm. Some makes of dry-ink felt white-board markers have a rigid body which fits the 10 mm bolt exactly and provide an excellent tube for constructing an electromagnet bobbin.

With a core in the electromagnets, the rotor gets additional rotating power. Initially, the magnets on the rotor are attracted to the electromagnet cores, giving the rotor a turning force which does not require any current to be supplied. When the rotor magnets are at their closest point to the electromagnet cores, the windings are powered up briefly and that gives the rotor magnets a strong push away, causing the rotor to spin.

There are many different designs of simple drive circuits and it is probably worth trying out different types to see which works best with your particular build of motor. In the same way, there are many kinds of collection circuits for taking off some of the excess power generated. The most simple of these is just a diode bridge, perhaps feeding a battery and charging it up for use at a later time. If you get sophisticated with the collection circuit and just take power off for a very short period of time at the correct moment, the cutting off of the current draw, causes a back-EMF magnetic pulse in the collection electromagnet which causes it to give the rotor an extra drive push – both current collection and rotor drive in one combined package.



Here are two of the most simple circuits possible, one for drive and one for power collection. The drive circuit transistor is switched on by a voltage generated in the grey coil by a rotor magnet passing by. The transistor then feeds a large current pulse to the black coil, driving the rotor on its way. The neon and the diode are there to protect the transistor and a physical layout for this circuit might be:

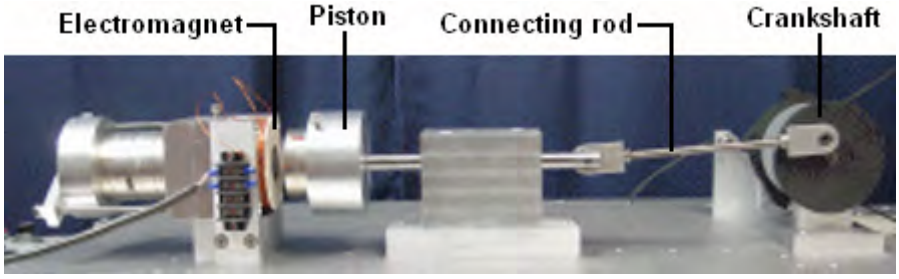


The 1K variable resistor is adjusted to give the best performance and the On/Off switch is optional. More advanced circuits, such as the one on page 2 - 9 can also be tried and the performance compared. Generally speaking, I would expect a three-rotor version to give a better performance than a single rotor implementation, but experimentation is needed.

Art Porter's Magnetic System.

Art uses an electromagnet with a ring magnet mounted on the core. When the coil is pulsed so as to augment the field of the permanent magnet, Art says that his prototype produces 2.9 times the field strength of the permanent magnet on it's own. When the coil is pulsed in the direction which opposes the field of the permanent magnet, the resulting magnetic field is zero.

This is a very serious change of magnetic field which can be used in different applications. One which Art has implemented is using the arrangement to power a crankshaft motor. Art shows several different motor builds including this one:



With this arrangement, Art states that 95% of the motor power comes from the permanent magnet. Art's website is at http://www.gap-power.com/index.html and he has a very interesting, lengthy video showing all of the details at http://www.gap-power.com/videos/Full%20Length%20Video.wmv.

In the video, Art attempts to apply Ohm's Law in an attempt to analyse the operation and is puzzled when the oscilloscope readings do not match his Ohm's Law calculations. He thinks that there is a contradiction between the oscilloscope and Ohm's Law, which he repeatedly stresses is a law of Nature, and he concludes that one of them has to be wrong. In actual fact, neither are wrong because Ohm's Law only applies to DC current flow in resistive circuits, and Art is not using steady DC current or a resistive load.

Applying short DC pulses to a substantial coil of wire is the equivalent of applying AC to that inductor. Ohm's Law does not apply due to the inductance of the coil. There is a Power Factor involved and Back EMF voltage pulses, so the oscilloscope readings are what is needed to calculate the input and output powers.

This arrangement is almost identical to that used in the Charles Flynn magnet motor described in chapter 1, and very close to the techniques used by Robert Adams in a properly tuned Adams Motor as shown at the start of this chapter. In my opinion, the effect which Art is exploiting would best be used if a large pick-up coil is placed against the end of the electromagnet core and the drive coil pulsed at the coil's (high) resonant frequency as that will minimise the input power and maximise the output power. Stacking those units in a bank could well produce a very sizeable excess electrical output. Our thanks are due to Art and his colleagues for sharing their research work freely for others to replicate and progress further.

DC Motor Efficiency

Commercially available DC motors are deliberately designed and manufactured to have extremely poor performance. In my opinion, the reason for this is that a properly designed electric motor could easily do away with the need for using internal combustion engines in vehicles and that would not suit the oil companies or their owners, the New World Order cartels. Worse still, electric motors with COP>1 open the way to self-powered free-energy systems and that would never do !!

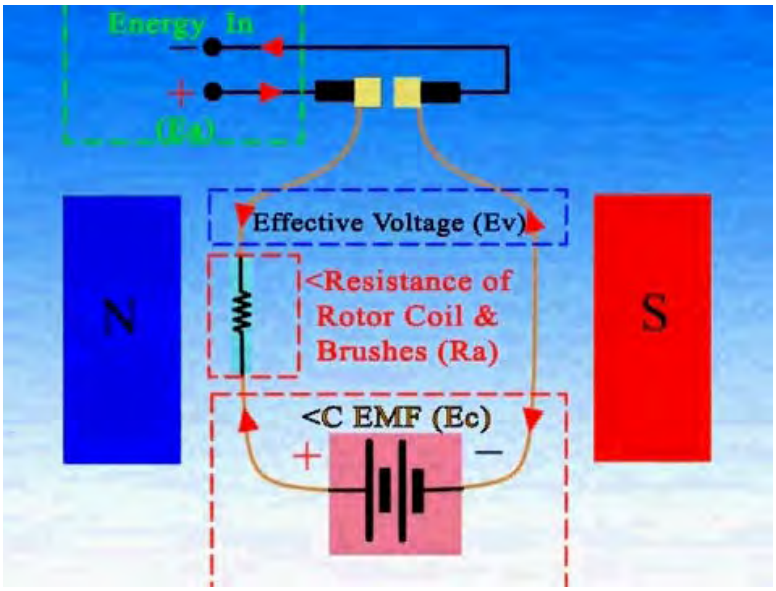
The video: http://torrentdb.in/95d95cff2bbC8039ed9/Peter-Lindemann-Electric-Motor-Secrets.torrent is available on the web and I strongly recommend that you watch all of it. It presents the basic facts very nicely. In brief outline, present day motors act both as a motor and as a generator of electrical power, but they are deliberately wound so that the power generation is used to oppose the input power and so produce a completely crippled output. During World War II, a German Engineer re-wired a standard electric motor and made it self-powered, that is, it ran and produced mechanical output power without the need for any input power once it had been started. That shows the potential of a properly constructed electric motor with the same size and general structure of any commercial electric motor.

Presumably, he did that by adding extra brushes and using some of the windings in generator mode with their output powering the drive windings which were arranged asymmetrically. There was also one other man who achieved self-powered re-wiring of a motor, but neither of those men made their information public knowledge.

The con job which has been run on us for many decades now is to wind the motor in such as way that the magnetic fields inside the motor oppose each other. When a current is passed through a coil of wire, it stores energy in that coil, and when the current flow is cut off, that energy needs to flow back out of the coil and it will do so in the reverse direction. This is sometimes called back-EMF ("Electro-Motive Force") although many people are not happy with that description. However, no matter what you call it, there is energy stored in the coil and that energy can be used to do useful work. But, the motor manufacturers choose to wind the motor so that instead of extracting that useful power, they use it to oppose a major part of the input power, creating a weak motor which heats up due to the wasted energy.

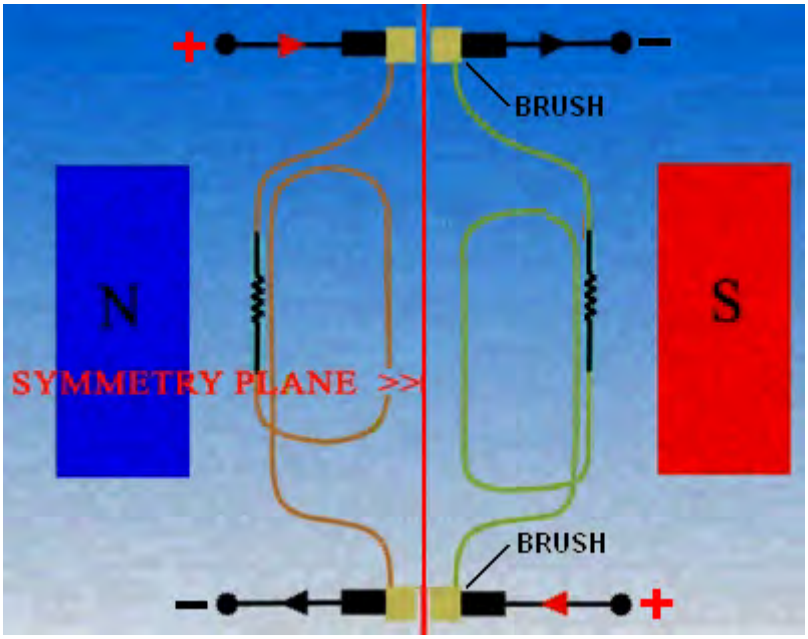
Contributor “UFOpolitics” points out that a deliberate mis-design of electric motors has, for the last 130 years, been presented to us as the only way to make and operate such motors. He states that because the windings are arranged in a symmetrical way, that a braking effect is produced which reduces the output power of the motor by anything from 50% to 90%. That is, a properly wound motor would have anything from twice to ten times the output power for the same input power. This mis-design guarantees that present day motors are always less than 100% efficient and always heat up when run. This mis-design is caused by using symmetrical windings in the motor.

Standard motor wiring is quite different and the killer effect is caused by having two windings which face each other, powered simultaneously with currents flowing in opposite directions. This causes a complete conflict between the magnetic fields and that destroys the efficiency of the motor: A very experienced experimenter has started a forum thread on the energetic-forum, both to explain this and to show new and more advanced construction methods and to answer questions and encourage replications and further developments. The forum is at: http://www.energeticforum.com/renewable-energy/11885-my-asymmetric-electrodynamic-machines.html and is definitely worth visiting, especially if you are good with mechanical devices. The experimenter uses the forum ID of “UFOpolitics” and he has produced an animated video in an attempt to explain the basic problems with present day DC electric motors: http://www.youtube.com/watch?v=Mj4rV0AoI-Q&feature=channel&list=UL. He points out that a problem winding in the standard DC motor looks like this:



The input current for any winding is fed in through a single pair of brush contacts. The generated electrical power “Ec” is not extracted and is forced to oppose the input energy “Ea”, leaving only a fraction of the input power to actually run the motor. It is likely that a motor of this type will only operate at 25% of it’s potential efficiency.

“UFOpolitics” has produced and demonstrated a simple way of overcoming this problem while using the existing motor housing, magnets and brush contacts. He does this by extracting the generated electrical power as a useful output and so preventing that useful power being used against the motor’s operation. To implement this, he adds one additional pair of brushes and re-winds the motor coils like this:



Here, one pair of brushes is at the top and one pair at the bottom of the ‘armature’ (the bit that rotates inside the motor housing and provides the mechanical power output). The coils are rewound to form a series of separate vertical coils, connecting to one brush terminal at the top and one brush terminal at the bottom as shown above.

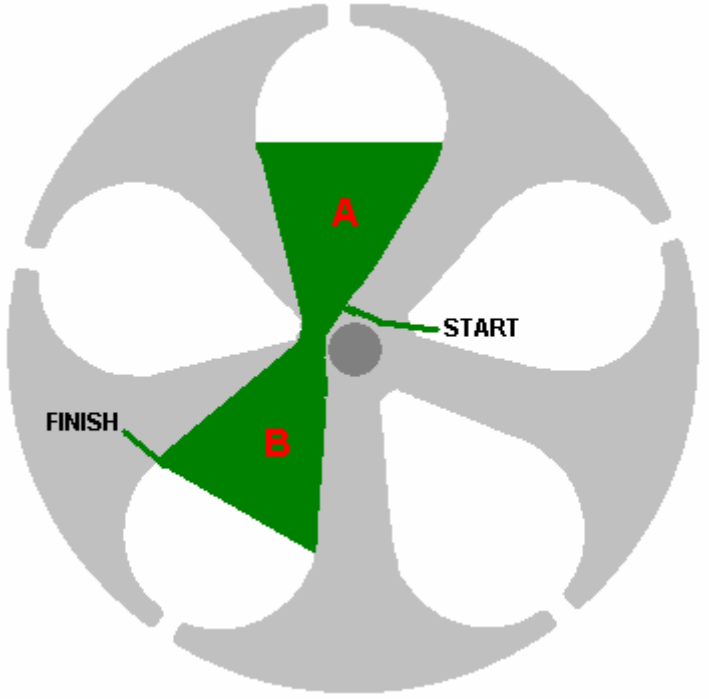
The input power is between the terminals on the left and flows through the coil shown in brown. The current flow generates a magnetic field, causing rotation because of the Permanent magnets marked “N” (for a magnet which has it’s North pole facing the coils) and “S” (for a magnet which has it’s South pole facing the coils). The black zig-zag line represents the resistance to current flow of the wire and brush contacts.

The coil shown in green on the right represents that same coil at a later moment when it has been disconnected from the power supply and rotated until it reaches that position, at which point, the energy stored in it is taken off as a useful output via the right hand pair of brushes. However, this is just an explanatory diagram and it does not show the very important fact that the discharging coil **must not** directly face a driving coil, because if it does, then the energy discharge would create a magnetic field which would interfere with the magnetic field of the driving coil and create a major problem.

Right, to say that again, any one coil is powered on the left hand side to drive the armature around and provide the output shaft with turning power (“torque”). Then that rotation disconnects that coil from the input power, leaving the coil charged with energy which has nowhere to go. That charged coil continues round until it hits the second set of brushes, which allow it to discharge through a load and do useful work.

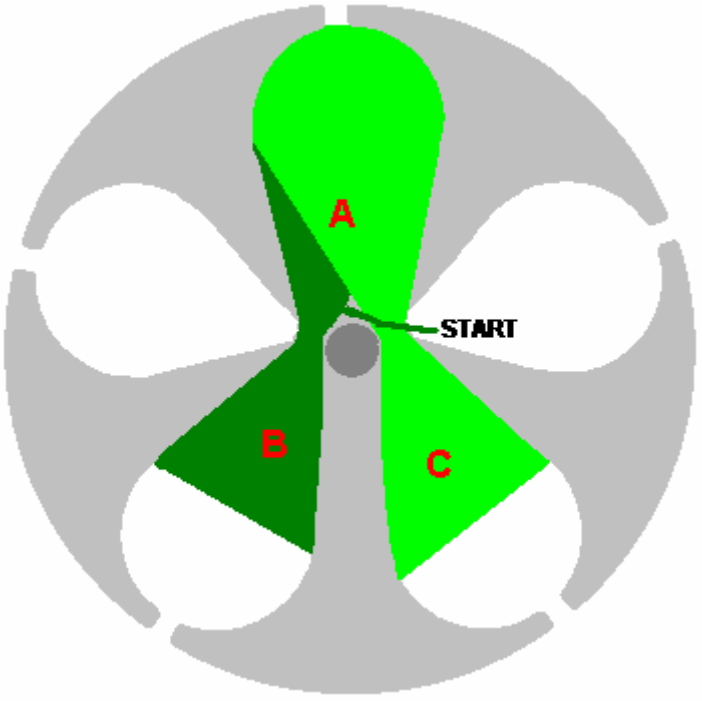
The really clever part of the adaption of the motor is best seen from above the vertical rotor. If, for example, you were to take a five-pole DC motor apart and remove the windings, the shaft and armature body might look like this:

When making an asymmetrical wound rotor, the windings go like this.

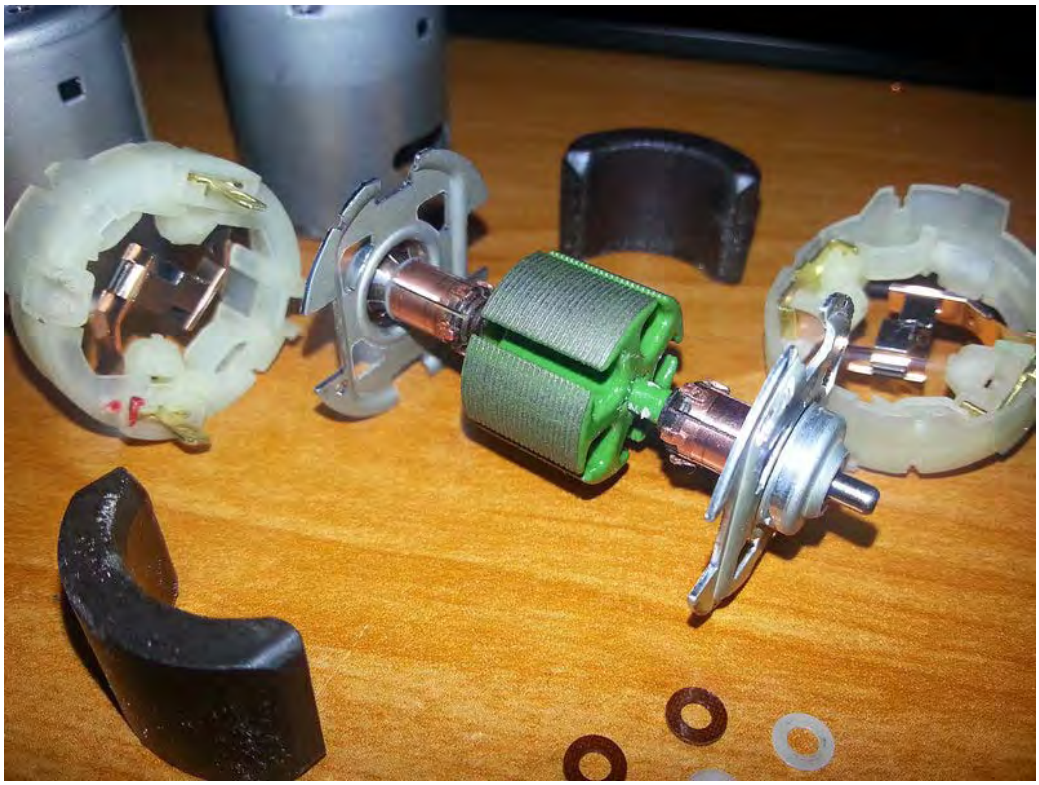


The start of the wire is secured at the top and then fed downwards through the opening “A” and back up through the opening “B”. For the small Radio Shack motor, this winding would be 25 turns of # 30 AWG wire (described as radio Shack ‘red’ wire, with a copper wire diameter of 0.255 mm). If you are re-winding a motor armature, please understand that each wire turn needs to be pulled tight in order to make a tight, solid and robust coil which will not vibrate unduly when the armature is spinning.

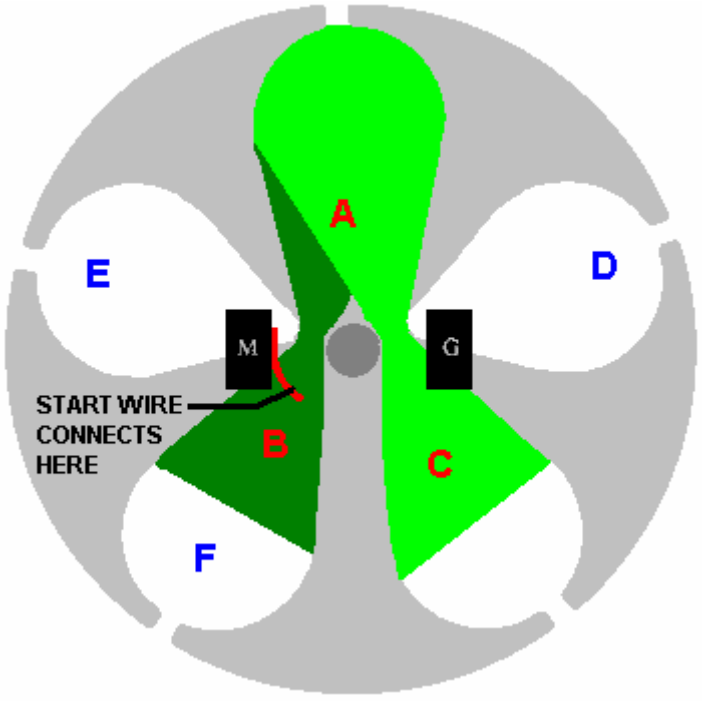
The end of the wire marked “FINISH” is **not** cut, but is taken down through opening “A” and this time, up through opening “C”. For clarity, these continuing turns are shown in a different colour, but please realise that it is the same single strand of wire being used throughout:



The final wire turn goes down through opening “A” and finishes at the other end of the body of the armature. In these views, the wire runs down into the paper, each turn forming a cylinder. This view may give you a better visual picture of what the coils are wound on:



The next step is to connect the START and FINISH wire ends of this V-shaped double coil to the “commutator” slip rings which allow current to be passed through the coil at just the right moment. Seen again from one end of the armature, the connections are like this:



The commutator slip rings are connected further up on the drive shaft and the start of the winding wire (shown previously in dark green) is connected to the top commutator sector in the position shown here. The finishing end of the wire is connected to the corresponding commutator sector at the far end of the shaft – that is, the sector directly in line with the upper sector just connected to the start of the wire.

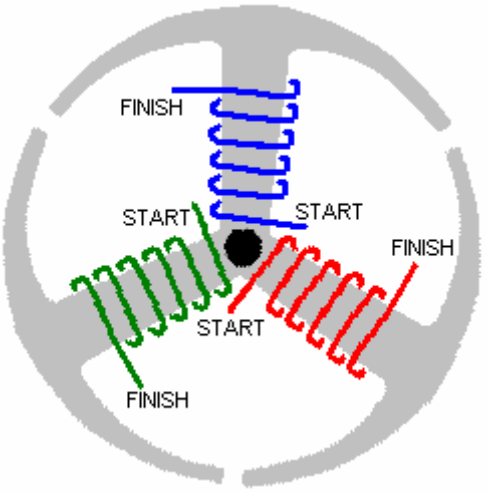
This completes the first of five identical V-shaped coils. The next coil is wound in the same way. The armature is rotated one sector counter-clockwise so that sector “D” replaces “A” at the top and the next coil is wound with the wire starting at the top and going down through opening “D” and up through opening “E”, repeating the same number of turns, and then, without cutting the wire the next set of wires are wound going down through opening “D” and back up through opening “F”. The start of the wire is then connected to the commutator sector which spans between openings “A” and “E” and the end connected to the corresponding commutator sector at the other end of the shaft..

For each of the remaining three windings, the shaft is rotated one position counter-clockwise and the same winding and connecting procedure carried out. When completed, no matter which opening is placed at the top of the view along the shaft, the windings and commutator sector for the wire connections will be identical.

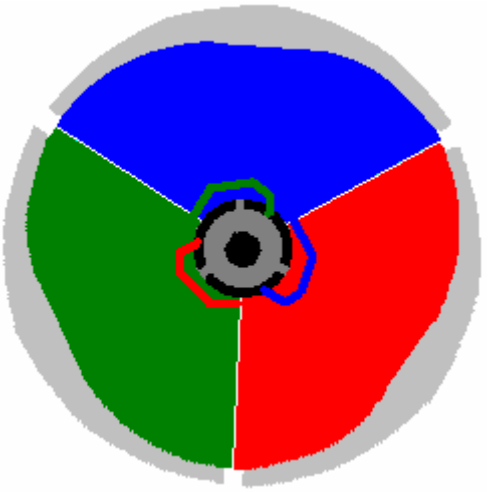
3-Pole Motors

The winding arrangement is slightly different for motors which have three poles (or multiples of three poles such as 6, 9, 12, etc poles). For the very simple 3-pole motors, the armature looks like this:

and with this style of armature, the winds are around the three arms, like this:



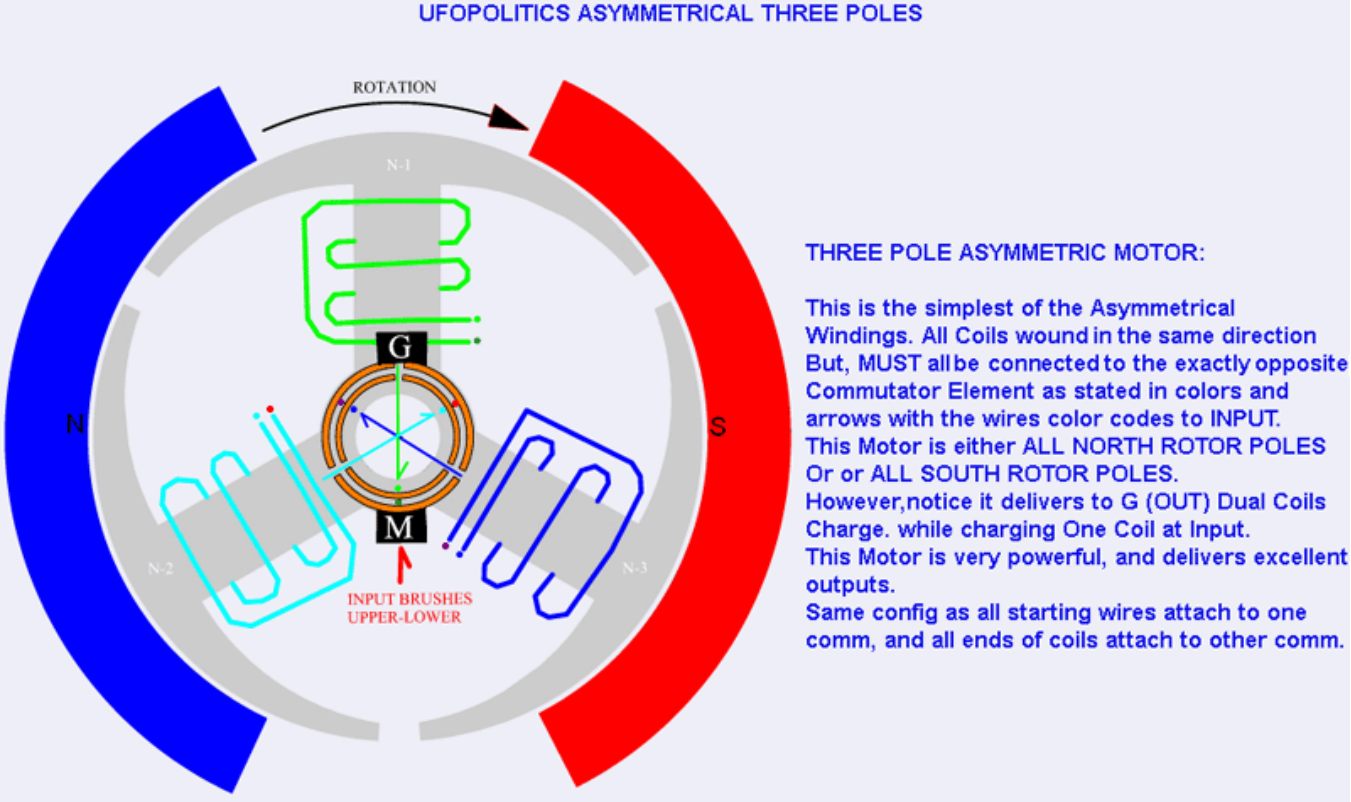
And as before, the commutator sectors at the top are duplicated at the bottom, allowing separate input and output circuits for each of the three coils. The windings have many turns, filling the available space and each winding is connected to the slip ring sector directly opposite it, like this:



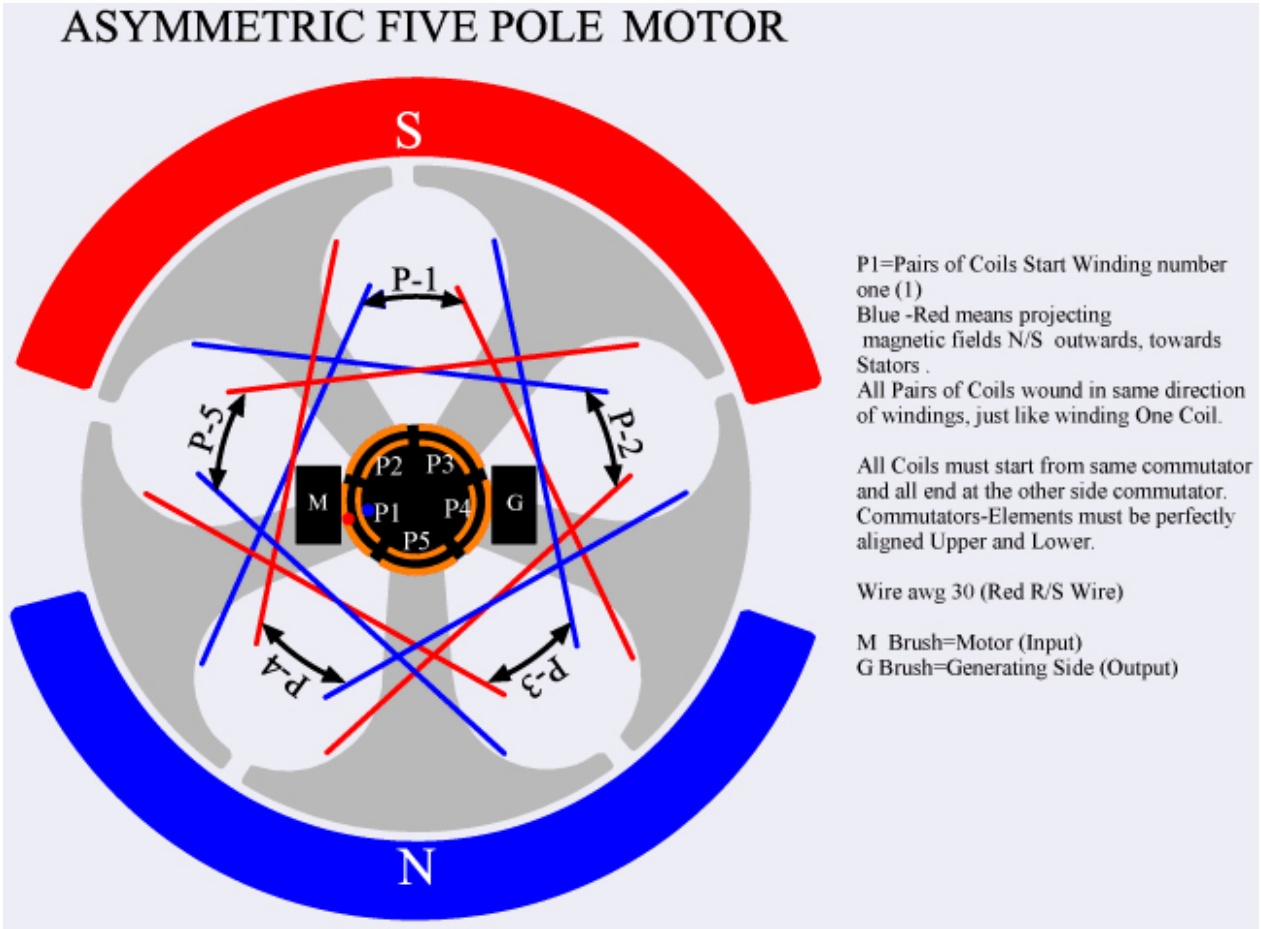
The Start of each winding is connected to the commutator slip ring sector at the top of the armature and the Finish is connected to the slip ring sector directly below it, that is, the sector which is at the same angle as the top one where the Start of the wire is connected. This allows the brushes which press against the slip ring sectors to connect to both ends of each coil in turn as the armature rotates. Three pole motors are particularly powerful and motors with six poles can be re-wound with pairs of adjacent sectors amalgamated to give three larger sectors. Nine pole motors can have three adjacent sectors wound as a single coil to provide the same effect as a three pole motor, and twelve pole motors can have four adjacent sectors wound as a single coil.

The positioning of the brushes is important. With the three-pole and five-pole arrangements, the brushes are aligned with the gaps between the magnets which surround the armature. However, the re-wound motor can be 'tuned' for improved torque and reduced drive current by adapting the motor housing to allow some adjustment of the position of the brush and commutator slip rings relative to the coils. This adjustment need only be slight as the angular movement of the brushes will be small. It is, of course, essential that the upper and lower adjusted positions move by exactly the same angular amount so that every upper commutator slip ring sector remains exactly above it's corresponding lower slip ring sector. In other words, the commutator slip ring sector at the top and bottom of each coil, must be exactly aligned vertically so that the electrical connections are made and broken at exactly the same instant

The commutator and brush arrangement are shown here in UFOpolitics' diagrams:



The Commutator brush marked “G” (for “Generator”) takes away the energy stored in each coil and passes it to an electrical load. The Commutator brush marked “M” (for “Motor”) feeds energy into the coil from the battery which is driving the motor. The red and blue stripes surrounding the armature are two permanent magnets. The magnet shown in red has it’s South pole facing the armature and the magnet shown in blue has it’s North pole facing the armature. This creates a magnetic field flowing horizontally across the armature. The five-pole arrangement is like this:



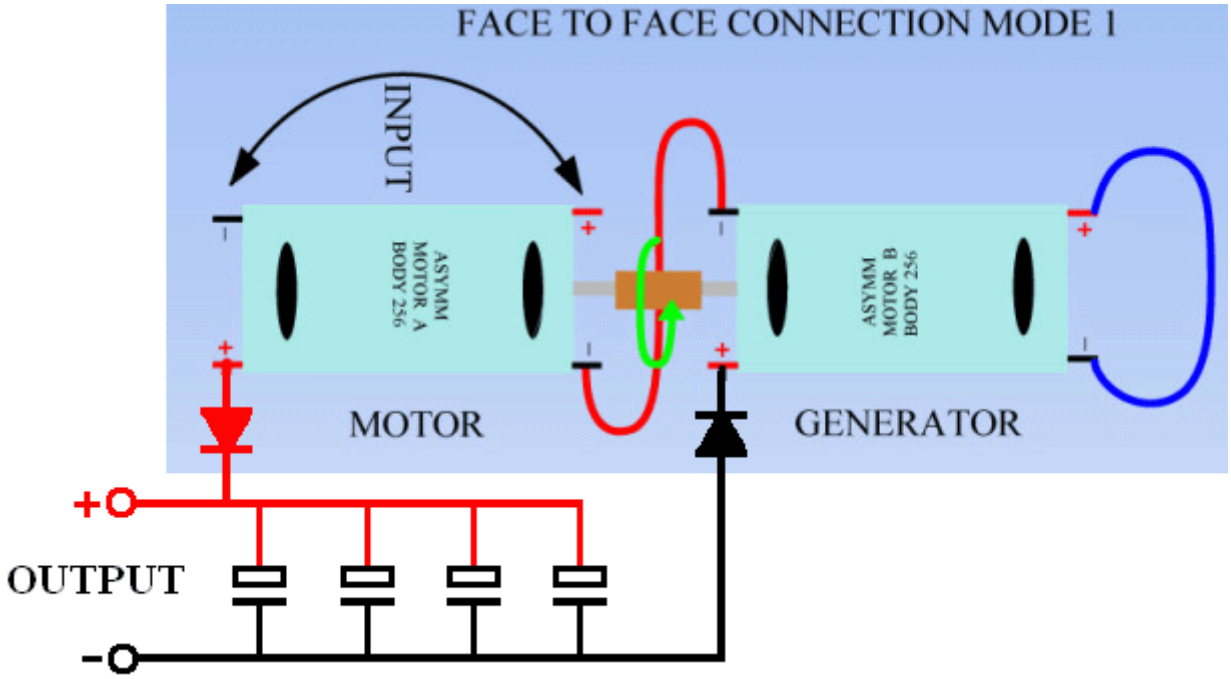
Here, the designation “R/S” stands for “Radio Shack” which is a chain of stores in America. In the forum, that is sometimes changed to “RS” and should not be confused with the large electronics outlet “Radio Spares” whose trademark is “RS”. UFOpolitics has suggested that the cheap 5-pole DC motor available from Radio Shack should be used by experimenters to become familiar with re-winding DC motor coils. Being a cheap product, those motors do not have a particularly high build quality, but they are suitable motors for experiments. Forum members share the details of how they dealt with adapting these and other motors.

I have to admit that motor windings and operation tend to confuse me and I sometimes find it difficult to understand what ‘UFOpolitics’ means when he talks about different winding strategies. However, it seems reasonably clear at this early stage of forum development, that his objective is to produce two things:

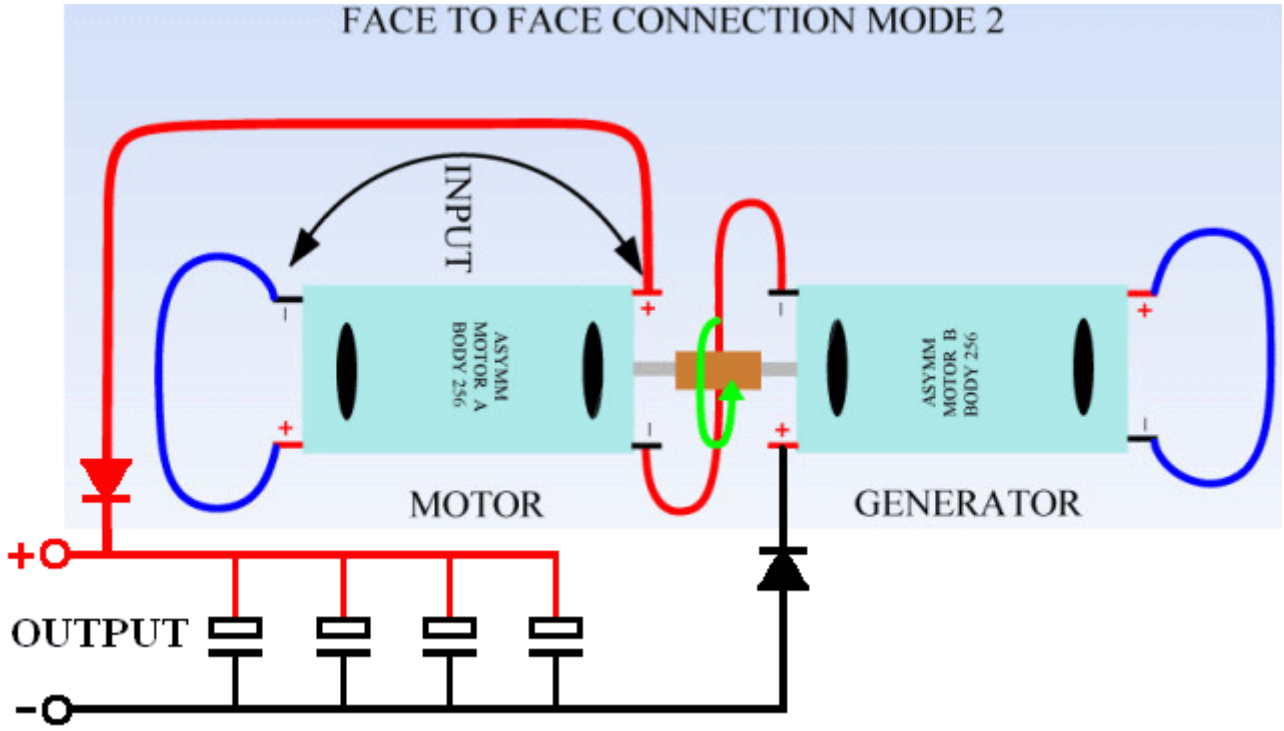
1. A very powerful electric motor which can be used in serious forms of road transport as well as for other practical applications, and

2. A powerful motor/generator combination which can produce useful generated electrical power.

While ‘UFOpolitics’ is very patiently going through many of the possible variations on how a DC motor can be wound and connected, and showing various forum members where they have failed to get some of their windings positioned correctly, he has also shown some of the best ways of connecting a re-wound motor used as a driver or “Prime Mover” as some people like to call it, and a re-wound motor which is to be used as an electrical generator. He shows two important ways for making a very effective Motor/Generator combination, as shown here:



It needs to be realised that these arrangements are not conventional arrangements and that the re-wound motors operate in a different way to motors bought ‘off the shelf’. For this reason, it is necessary to isolate the electrical output to prevent current flowing through the load from affecting the operation of the Motor/Generator combination. This can be done by placing a diode in each of the output lines and charging a capacitor bank which is then used to feed whatever load is to powered. If my understanding is correct, then feeding any cold electricity produced into a capacitor causes the current to become conventional hot electricity. It is not clear if that action is part of this arrangement although the circuitry shown should be used. This is the second version:



'UFOpolitics' comments on these arrangements as follows: As we excite the input of the Motor, the Generator will start producing energy and that additional energy will flow through the Motor Output side because they are connected in series here. Two rectifiers must be connected at both output terminals, Positive and Negative, to avoid back flow from closing the circuit through the load.

As the Motor accelerates, the Generator boosts the energy flow which then runs through the Motor augmenting the Output Fields and when the output is loaded then an 'Engagement' of both Machines occurs as they start to compensate each other through their output flows. It should be understood that the Output should be Capacitor Banked in a dedicated Reservoir.

When designing a Generator for a specific, existing Asymmetric Motor machine, it must be understood that Generator Interactions should be considered to run as "Counter Rotation" to the Motor Machine's originally conceived rotation (which is easily done by just moving brush-lines passing stator bisector angles to the opposite of those needed for a Motor, or alternatively, setting the timing backwards). This will definitively enhance the assisted rotation of both Machines when connected together in this Face-to-Face mode.

As I do not find the forum comments easy to understand, I recommend that you visit the forum and read the posts as you may well understand the conversations easier to follow than I do.

On the forum, 'Sanskara316' states "I have re-wound a small 3-volt 3-pole motor. I used an almost dead, 6-volt sealed lead-acid battery to power the motor. This battery just sits at around 4 volts and if given a load, even a small LED, it's voltage drops to 1 volt. The re-wound motor started very slowly - barely spinning, then after a minute or two it started to spin faster, and I noticed that the voltage on the battery was slowly climbing. I connected a small LED flashlight to the generating side and it lit up. Now the battery voltage under load is around 2+ volts. It's been running for an hour now and the machine squeals a lot. It is conditioning the battery and the meter cannot be showing what really is happening. The motor draws 300 ma?? – That's not possible as the battery just doesn't have that power". To which 'UFOpolitics' remarks: "Well I am glad you have witnessed some of the 'Effects'... these re-wound motors do recondition batteries...remember, Radiant Energy is taking over the Machine...so Radiant Energy comes out through the Input also… which is the reason why we get high Volts-Amps reading on a meter ...these motors use very small amounts of current and volts. Inside the motor, every coil is being 'Self-Electromagnetically-Pulsed' because they auto-disconnect from the power source, then the next coil in the sequence is assisted by the first coil when it has rotated to it's next position, and so on. The commutator switching has become a 'Self-Oscillator' for every independently-energised coil."

Another forum member 'prochiro' says: "I have also replicated the battery-charging events that 'Sanskara316' indicated. I started with a 12-volt 4 Amp-Hour battery which I had been using with another circuit two weeks ago and had not recharged it after using it for hours. It was sitting at 12.40 volts. I took my best-running re-wound motor, plugged it in direct and ran it. The battery voltage dropped to 12.24 volts and stayed at that level for 30 seconds. The battery voltage then started to rise 1/100 of a volt per minute. When it was at 12.27 volts, I disconnected the motor (the total run time was less than 5 minutes). I then let it rest for five minutes. At the end

of the five minutes, the battery voltage had risen to 12.43 volts and is still at that voltage now. Just think what a larger motor would do on a big battery bank. Everybody needs to document this test as it proves what 'UFOpolitics' said."

New DC motors, and particularly cheap motors, will have brushes which do not mate cleanly with the commutator slip ring sectors and so, when the modification has been made, running the motor for some time allows the brushes to wear in and that raises the efficiency of the electrical connections which in turn, improves the performance of the motor. If you wish to build and test one of these motors, then you can find help and support in the forum with your questions answered and numerous videos and photographs from different experimenters to help you.

Direct Coupled Motors

At http://www.youtube.com/watch?v=sj0ovfLvFTQ&feature=share Raoul Melendres shows a very interesting arrangement of two motors:

Probably intended to make you think more than anything else, this circuit appears to have the right hand motor, which is less than 100% efficient, being driven by the left hand motor which is also less than 100% efficient. The output appears to be arranged to provide DC to a transistor astable multivibrator which drives the secondary of a mains transformer, which is also less than 100% efficient. The much higher output voltage from the primary of the transformer is full-wave rectified and then smoothed using a capacitor and fed to the input of the driving motor, which presumably, drags the low-power output down to around six volts. This voltage supposedly drives the motor and has some current left over to charge a cell-phone battery. From a theoretical point of view, this looks impossible but the interesting question is "is it impossible?". Personally, I would have expected the pulley wheel on the drive motor to be larger than the pulley wheel on the generator.

Patrick Kelly
http://www.free-energy-info.tuks.nl
http://www.free-energy-info.com
http://www.free-energy-info.co.uk
http://www.free-energy-devices.com

Chapter 3: Motionless Pulsed Systems

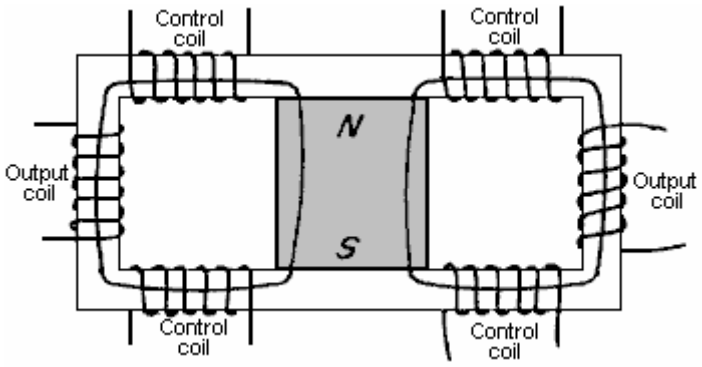
Note: If you are not at all familiar with basic electronics, you might find it easier to understand this chapter if you read chapter 12 first.

The pulsed devices mentioned so far have had moving parts but rotating or fluctuating magnetic fields can be created without moving parts. An example of this is Graham Gunderson's solid-state electric generator shown in US Patent Application 2006/0163971 A1 of 27th July 2006 which is shown on page A-1038 of the appendix. Another example is:

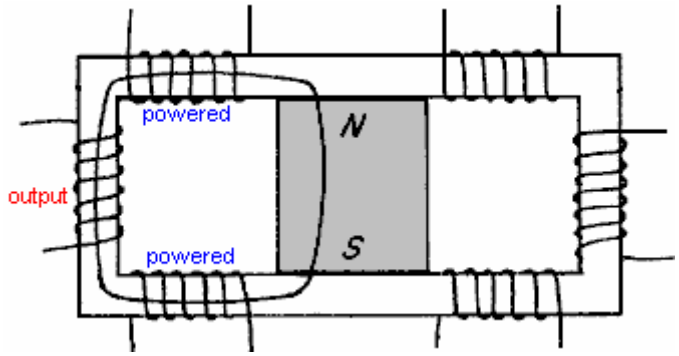
Charles Flynn's Magnetic Frame.

Another device of this type comes from Charles Flynn. The technique of applying magnetic variations to the magnetic flux produced by a permanent magnet is covered in detail in the patents of Charles Flynn which are included in the Appendix. In his patent he shows techniques for producing linear motion, reciprocal motion, circular motion and power conversion, and he gives a considerable amount of description and explanation on each, his main patent containing a hundred illustrations. Taking one application at random:

He states that a substantial enhancement of magnetic flux can be obtained from the use of an arrangement like this:



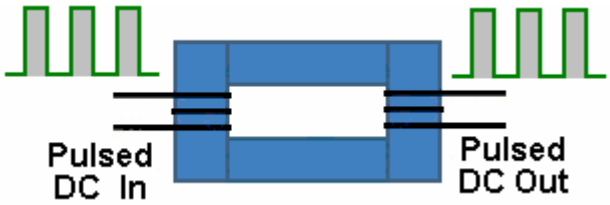
Here, a laminated soft iron frame has a powerful permanent magnet positioned in it's centre and six coils are wound in the positions shown. The magnetic flux from the permanent magnet flows around both sides of the frame.



The full patent details of this system from Charles Flynn are in the Appendix, starting at page A - 338.

Lawrence Tseung's Magnetic Frame.
Lawrence Tseung has recently produced a subtle design using very similar principles. He takes a magnetic frame of similar style and inserts a permanent magnet in one of the arms of the frame. He then applies sharp DC pulses to a coils wound on one side of the frame and draws off energy from a coil wound on the other side of the frame.

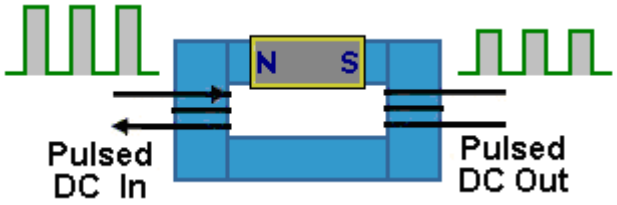
He shows three separate operating modes for the devices as follows:



(1) No Permanent Magnet, No Lead-Out Energy, Maximum COP = 1

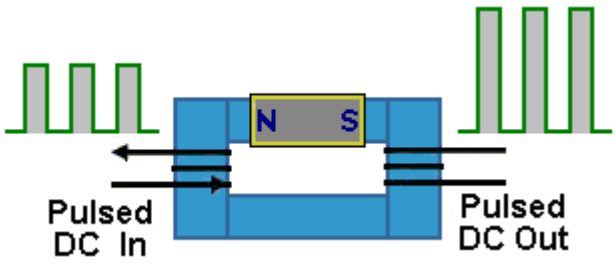
Lawrence comments on three possible arrangements. The first on shown above is the standard commercial transformer arrangement where there is a frame made from insulated iron shims in order to cut down the "eddy" currents which otherwise would circulate around inside the frame at right angles to the useful magnetic pulsing which links the two coils on the opposite sides of the frame. As is very widely known, this type of arrangement never has an output power greater than the input power.

However, that arrangement can be varied in several different ways. Lawrence has chosen to remove a section of the frame and replace it with a permanent magnet as shown in the diagram below. This alters the situation very considerably as the permanent magnet causes a continuous circulation of magnetic flux around the frame before any alternating voltage is applied to the input coil. If the pulsing input power is applied in the wrong direction as shown here, where the input pulses generate magnetic flux which opposes the magnetic flux already flowing in the frame from the permanent magnet, then the output is actually **lower** than it would have been without the permanent magnet.



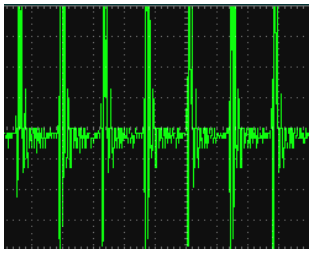
(2) Permanent Magnet Opposes Magnetic Flux: COP < 1

However, if the input coil is pulsed so that the current flowing in the coil produces a magnetic field which reinforces the magnetic field of the permanent magnet then it is possible for the output power to exceed the input power. The "Coefficient of Performance" or "COP" of the device is the amount of output power divided by the amount of input power which the user has to put in to make the device operate. In this instance the COP value can be greater than one:



(3) Permanent Magnet Enhances Magnetic Flux: COP > 1

As it upsets some purists, perhaps it should be mentioned that while a square wave input signal is applied to the input of each of the above illustrations, the output will not be a square wave although it is shown that way for clarity. Instead, the input and output coils convert the square wave to a low-quality sine wave which only becomes a pure sine wave when the pulse frequency exactly matches the resonant frequency of the output winding. The oscilloscope display shown here is a typical output power waveform which has nearly 390,000 of these pulses per second.



There is a limit to this as the amount of magnetic flux which any particular frame can carry is determined by the material from which it is made. Iron is the most common material for frames of this type and it has a very definite saturation point. If the permanent magnet is so strong that it causes saturation of the frame material before the input pulsing is applied, then there can't be any effect at all from positive DC pulsing as shown. This is just

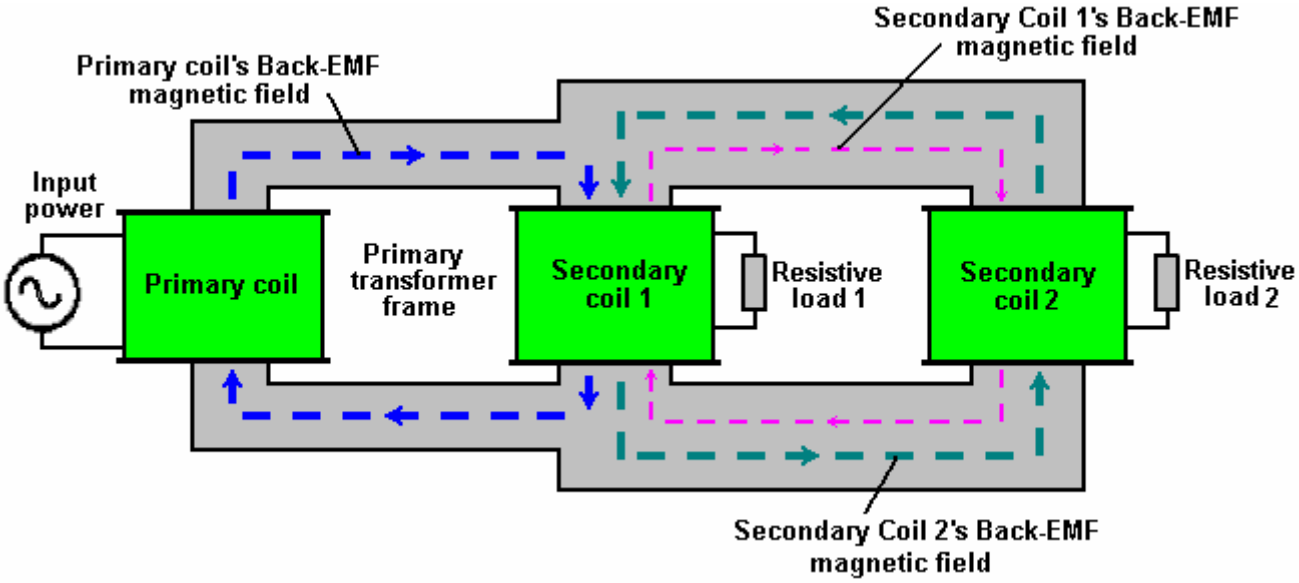
common sense but it makes it clear that the magnet chosen must not be too strong for the size of the frame, and why that should be.

As an example of this, one of the people replicating Lawrence's design found that he did not get any power gain at all and so he asked Lawrence for advice. Lawrence advised him to omit the magnet and see what happened. He did this and immediately got the standard output, showing that both his input arrangement and his output measuring system both worked perfectly well. It then dawned on him that the stack of three magnets which he was using in the frame were just too strong, so he reduced the stack to just two magnets and immediately got a performance of COP = 1.5 (50% more power output than the input power).

The Transformers of Thane Heins.

Thane has developed, tested and patented a transformer arrangement where the output power of his prototype is thirty times greater than the input power. He achieves this by using a figure-of-eight double toroid transformer core. His Canadian patent CA2594905 is titled "Bi-Toroid Transformer" and dated 18th January 2009. The abstract says: The invention provides a means of increasing transformer efficiency above 100%. The transformer consists of a single primary coil and two secondary coils. The two secondary coils are set on a secondary toroidal core which is designed to be maintained at a lower magnetic resistance than the primary toroidal core throughout the entire operating range of the transformer. Thus, when the transformer secondary delivers current to a load, the resulting Back-EMF is not allowed to flow back to the primary due to the higher magnetic resistance of that flux path, instead, the secondary coil's Back-EMF follows the path of least magnetic resistance into the adjacent secondary coil.

You will notice that in the following diagram, the secondary transformer frame on the right is much larger than the primary transformer frame on the left. This larger size produces a lower magnetic resistance or "reluctance" as it is known technically. This seems like a minor point but in fact it is not, as you will see from the test results.



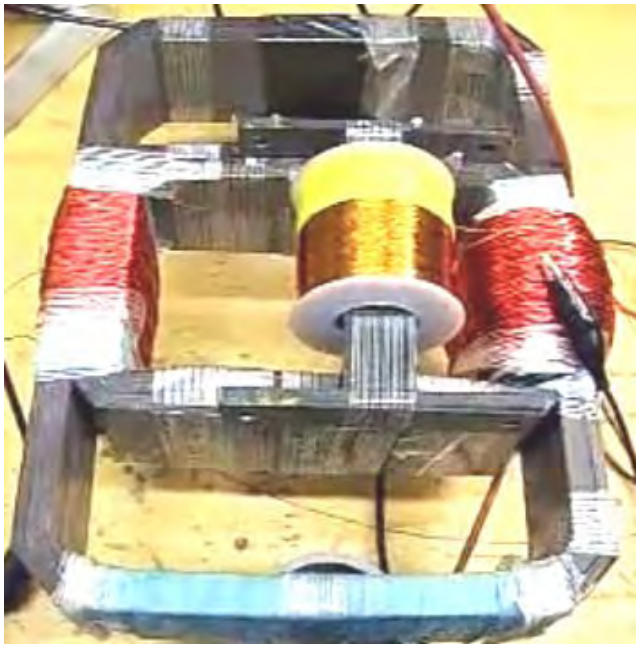
In a conventional transformer, the power flowing in the primary winding induces power in the secondary winding. When the power in the secondary winding is drawn off to do useful work, a Back-EMF magnetic flux results and that opposes the original magnetic flux, requiring additional input power to sustain the operation.

In this transformer, that opposing magnetic flow is diverted through a larger magnetic frame which has a much lower resistance to magnetic flow and which, as a result, bleeds off the problem flux, sending it through secondary coil 2 in the diagram above. This pretty much isolates the input power from any opposition, resulting in a massive improvement in the operation efficiency.

In the patent document, Thane quotes a prototype test which had a primary coil winding with 2.5 ohms resistance, carrying 0.29 watts of power. The secondary coil 1 had a winding with 2.9 ohms resistance, receiving 0.18 watts of power. The Resistive load 1 was 180 ohms, receiving 11.25 watts of power. The secondary coil 2 had a winding with 2.5 ohms resistance, and received 0.06 watts of power. Resistive load 2 was 1 ohm, receiving 0.02 watts of power. Overall, the input power was 0.29 watts and the output power 11.51 watts, which is a COP of

39.6 and while the document does not mention it directly, the primary coil should be driven at it's resonant frequency.

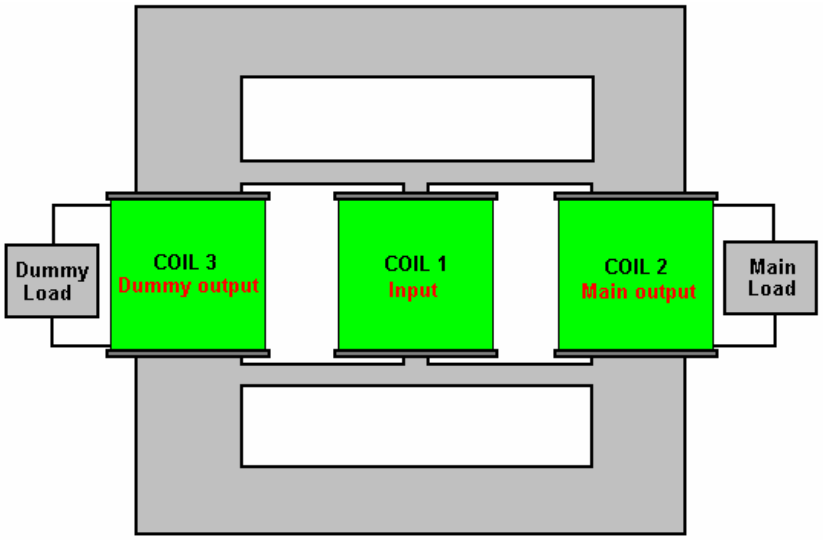
A variation of this arrangement is to attach an outer toroid to the existing bi-toroid arrangement, like this:



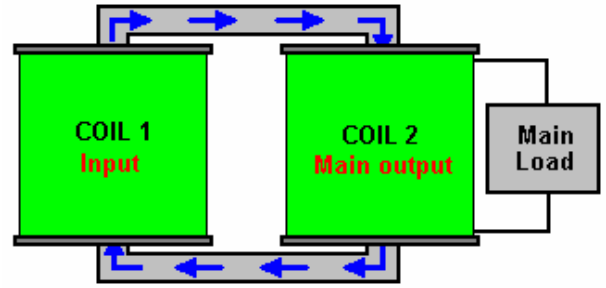
This prototype, as you can see, is fairly simple construction, and yet, given an input power of 106.9 milliwatts, it produces an output power of 403.3 milliwatts, which is 3.77 times greater.

This is something which needs to be considered carefully. Conventional science say that "there is no such thing as a free meal" and with any transformer, you will get less electrical power out of it than you put into it. Well, this simple looking construction demonstrates that this is not the case, which shows that some of the dogmatic statements made by present day scientists are completely wrong.

This version of Thane's transformer is made like this:

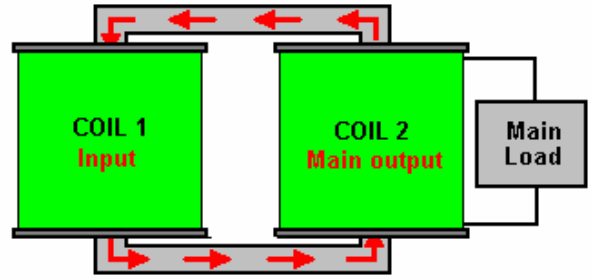


The way that off-the-shelf transformers work at the moment is like this:



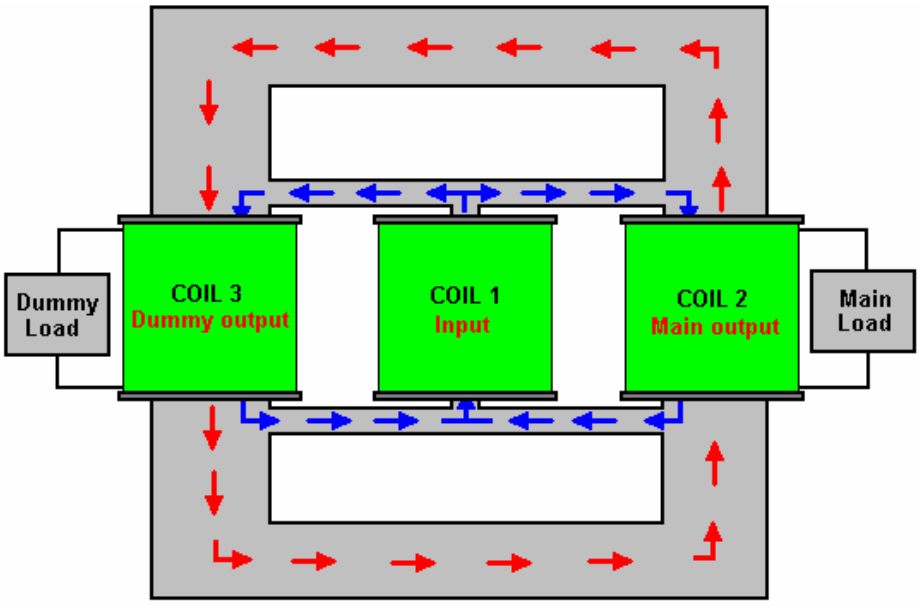
When a pulse of input power is delivered to Coil 1 (called the "Primary winding"), it creates a magnetic wave which passes around the frame or "yoke" of the transformer, passing though Coil 2 (called the "Secondary winding") and back to Coil1 again as shown by the blue arrows. This magnetic pulse generates an electrical output in Coil 2, which flows through the electrical load (lighting, heating, charging, video, or whatever) providing it with the power which it needs to operate.

This is all well and good but the catch is that the pulse in Coil 2 also generates a magnetic pulse, and unfortunately, it runs in the opposite direction, opposing the operation of Coil 1 and causing it to have to boost it's input power in order to overcome this backward magnetic flow:



This is what makes current scientific "experts" say that the electrical efficiency of a transformer will always be less than 100%.

Thane has overcome that limitation by the simple and elegant technique of diverting that backward pulse of magnetism and channelling it through an additional magnetic path of lower resistance to magnetic flow through it. The path is arranged so that Coil 1 has no option but to send it's power through the frame as before, but the return pulse takes a much easier path which does not lead back to Coil 1 at all. This boosts the performance way past the 100% mark, and 2,300% has been achieved quite readily (COP=23). The additional path is like this:



Not shown in this diagram are the reverse pulses from Coil 3. These follow the easier outside path, opposing the unwanted back pulse from coil 2. The overall effect is that from Coil 1's point of view, the tiresome back pulses

from Coil2 have suddenly disappeared, leaving Coil 1 to get on with the job of providing power without any hindrance.

This simple and elegant modification of the humble transformer, converts it into a free-energy device which boosts the power used to drive it and outputs much greater power. Congratulations are due to Thane for this technique.

At the present time there are videos showing how this transformer works:
http://www.youtube.com/watch?v=5KfwiXJ8apk and
http://www.youtube.com/watch?v=GcAYhM0LX9A&playnext=1&feature=sub&list=TLJjwIlLxS9jQ.

The High-power Motionless Generator of Clemente Figuera

Clemente Figuera of the Canary Islands died in 1908. He was a highly respected individual, an Engineer and University Professor. He was awarded several patents and was known to Nikola Tesla. Figuera's design is very simple in outline. He has avoided the performance-killing Lenz Law magnetic feedback by splitting a transformer into three parts. Two parts form the primary winding and are shown on the left and on the right. The third part is the secondary winding which is located in the centre. Because of the splitting of the primary into two parts, Lenz's Law has been abolished for this design, allowing a spectacular performance where the current drawn from the secondary winding has no effect on the current flowing in the two halves of the primary winding. There is also, no back-EMF as current flows continuously in both halves of the primary winding. The very clever method used by Clemente makes the strength of the current in the two halves of the primary to oscillate with one side repeatedly having first much more current and then far less current than the other half. This generates alternating current in the secondary, current which can be drawn off and used for useful work, powering lights, heaters, motors, etc. The following information comes from a man who wishes to remain anonymous. On 30th October 2012, he made the following comments about his repair to a Figuera patent which was missing some of the content. He says:

CLEMENTE FIGUERA AND HIS INFINITE ENERGY MACHINE

I heard of Clemente Figuera for the first time from one of the Tesla articles. In 1902 the Daily Mail announced that Mr. Figueras (with an "s"), a Forestry Engineer in the Canary Islands, and for many years Professor of Physics at St. Augustine's College, Las Palmas, had invented a generator which required no fuel. The newspaper article says that "He claims to have invented a generator which can collect the electric fluid, to be able to store it and apply it to infinite purposes, for instance, in connection with shops, railways and manufacturers. He will not give the key to his invention, but declares that the only extraordinary point about it is that it has taken so long to discover a simple scientific fact. Señor Figueras has constructed a rough apparatus by which, in spite of it's small size and it's defects, he obtains 550 volts, which he utilises in his own house for lighting purposes and for driving a 20 horse-power motor. Señor Figueras is shortly coming to London, not with models or sketches, but with a working apparatus. His inventions comprise a generator, a motor, and a sort of governor or regulator, and the whole apparatus is so simple that a child could work it." [Taken from "Perpetual Motion – A History of an Obsession"].

I was in one of the forums when someone mentioned Clemente Figuera and provided some links to documents referring to his work [1]. In one of the documents, I found what looks to be the only page showing sketches from one of his patents. After restoring the faint lines which show the wire connections, I was very surprised to see the similarities between the embodiment of Mr. Figuera's drawing and one of my own for over-unity transformers.

I was very eager to read any information about Mr Figuera's work and the operation of his 'Infinite Energy Machine'. It looks very suspicious that the pages describing the most important part of the machine have been 'lost'. I then decided to just figure this machine out for myself.

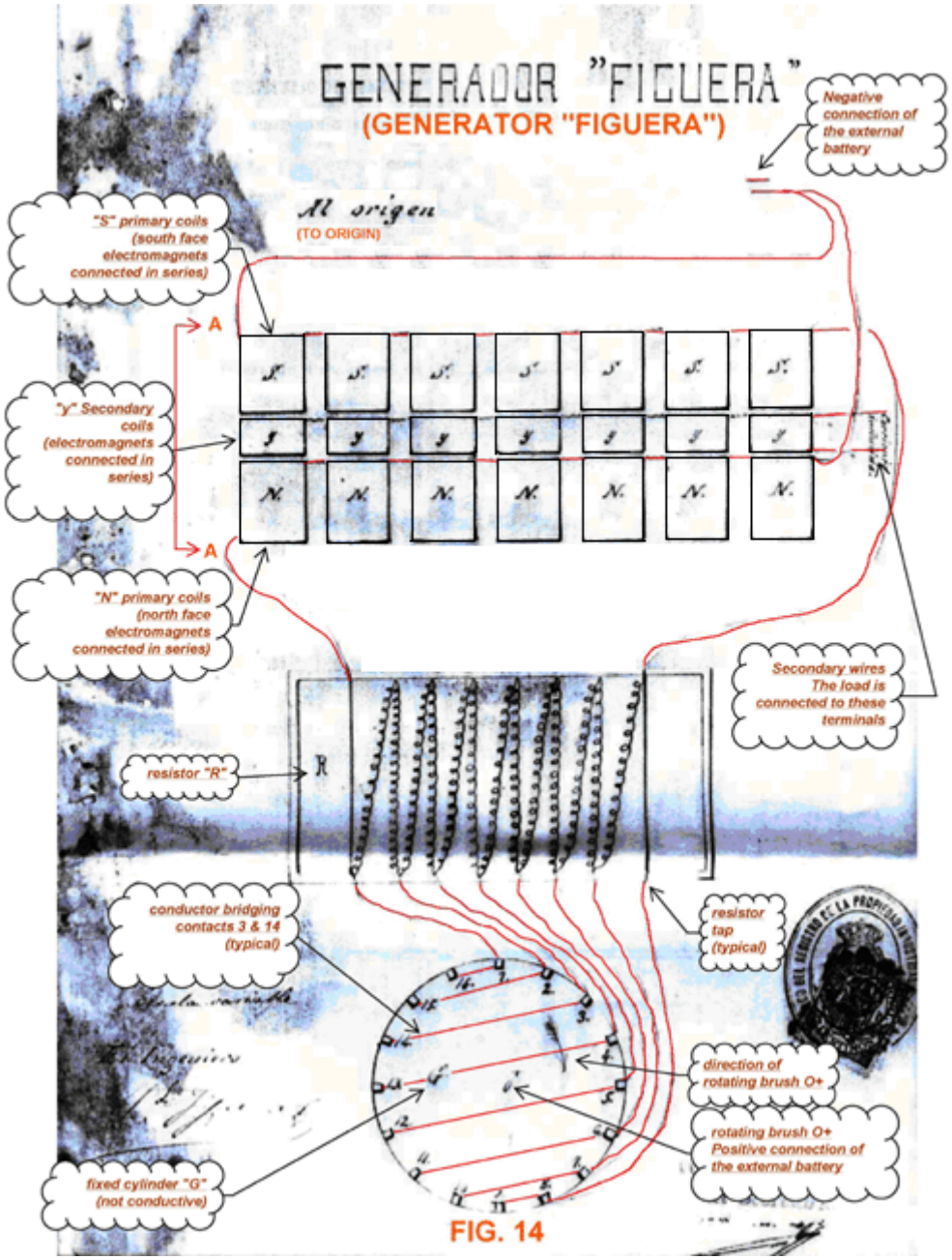

GENERADOR "FIGUERA"
(GENERATOR "FIGUERA")
Negative connection of the external battery
Al origen
(TO ORIGIN)
"S" primary coils (south face electromagnets connected in series)
A
"y" Secondary coils (electromagnets connected in series)
A
"N" primary coils (north face electromagnets connected in series)
Secondary wires The load is connected to these terminals
resistor "R"
R
conductor bridging contacts 3 & 14 (typical)
resistor tap (typical)
direction of rotating brush O+
rotating brush O+ Positive connection of the external battery
fixed cylinder "G" (not conductive)

FIG. 14

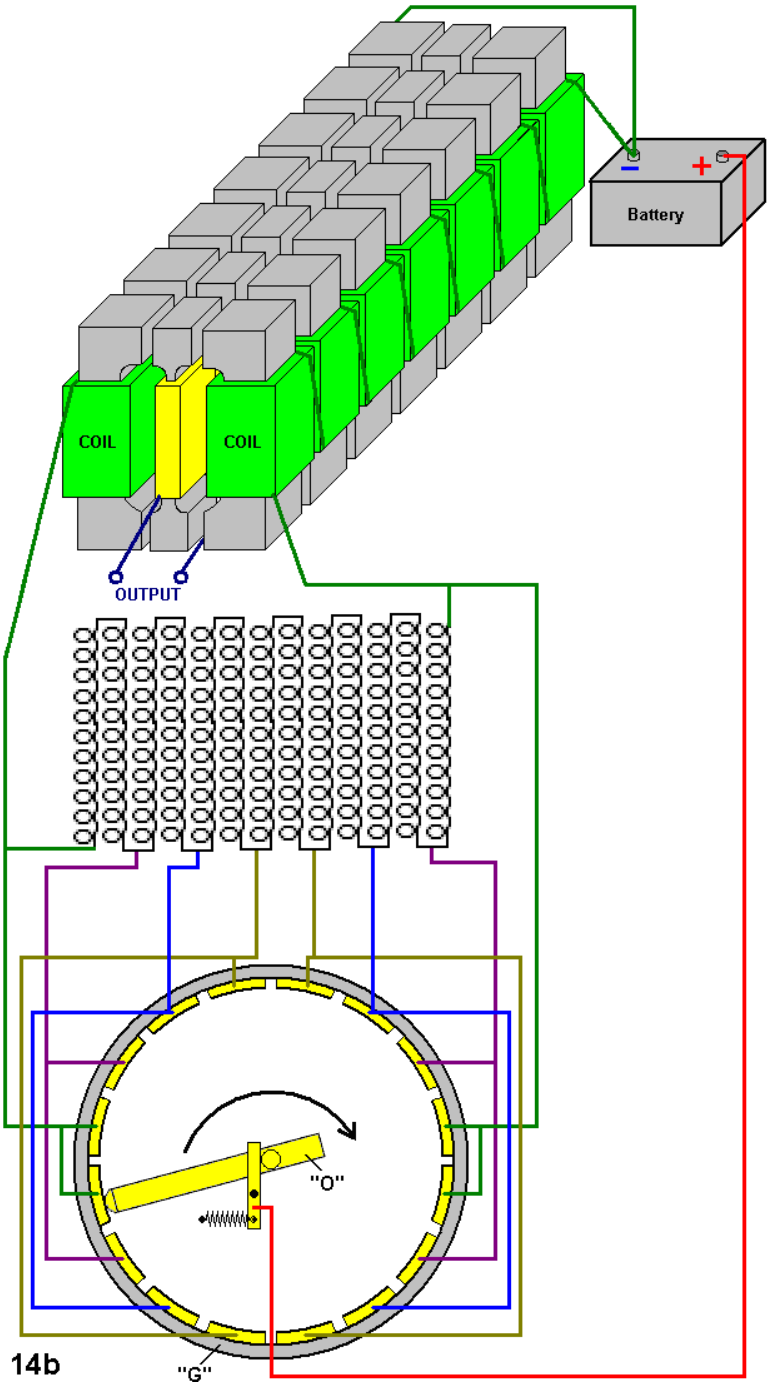


Fig. 14b

Please note that the rotating contact brush needs to be a “Make Before Break” type. That is, it needs to bridge across the gap between adjacent stator contact strips so that there is no sparking due to the current flow being interrupted.

According to Mr. Figuera, an over-unity transformer can be built without using permanent magnets, and based on a very simple concept. Figuera’s generator consists of three rows of electromagnets, where each row is connected in series. The rows of “S” and “N” electromagnets function as the primary of the transformer, while the row of “y” electromagnets, located in the centre, functions as the secondary. The “S” and “N” stand for South and North poles, respectively. The apparatus includes a resistor “R” having multiple taps connected to a type of distributor formed by a cylinder “G” and brush “O”. The brush “O” rotates inside the cylinder “G” changing the connection to the resistor taps. When the brush “O” rotates around the eight taps, it generates two stepped half-cycle sine waves which are 90° out of phase with each other. I suggest that Fig.15 is the wiring diagram as originally disclosed by Mr. Figuera in his patents. The most significant component of the system is the arrangement of the electromagnets shown in section A-A of figure 14. Keep in mind that each electromagnet shown in figure 15 corresponds to a row of seven electromagnets connected in series as shown in figure 14. In addition, I recommend that when building this apparatus, at least for the first implementation, that you try to replicate all of the details of the device shown in the patent. For example, figure 14 shows the top area of the “S” and “N” electromagnets being approximately equal to twice the top area of the “y” electromagnets.

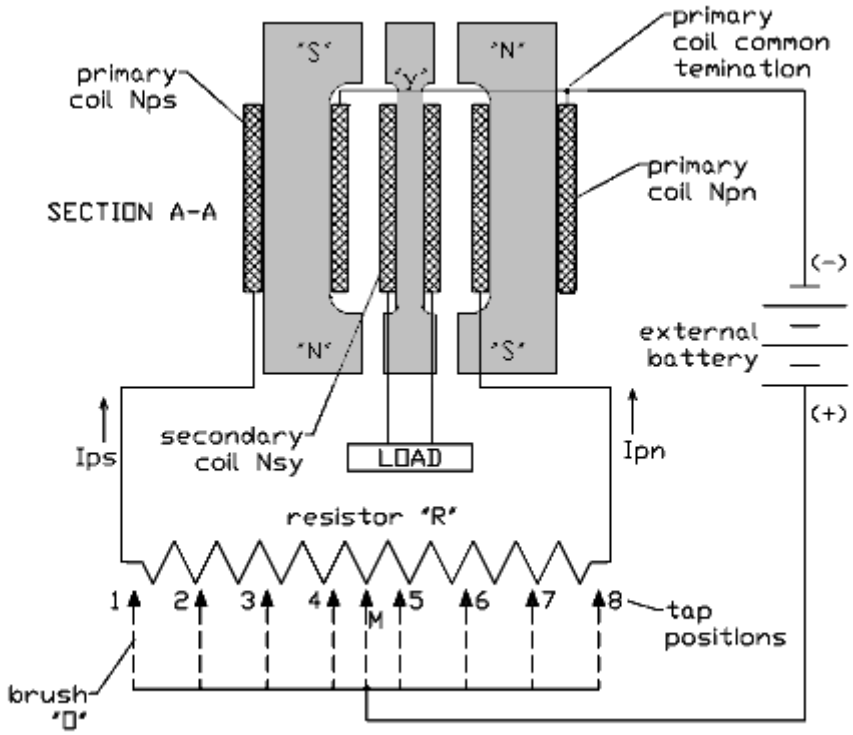


FIG. 15 **SIDE VIEW**

Even though Mr. Figuera used stepped sinusoidal currents **Ips** and **Ipn**, I consider the resistor shown in Fig.15, to be a linear variable resistor having infinite 'taps' and the voltage and current generated to be pure half-cycle sine waves which are 90° out of phase. The coils of the "S" and "N" electromagnets are connected together and attached to the negative terminal of the battery. The other ends of both electromagnets are connected to both ends of the resistor "R". The sliding contact "O" is connected to the positive terminal of the battery and is rotated continuously making electrical connections repeatedly from left to right and then back from right to left across the multi-tap resistor "R". The position of the sliding contact "O", determines the magnitude of the DC currents **Ips** and **Ipn** passing through the primary coils "S" and "N". For instance, when the brush is in position 1, the "S" coils receive the full voltage of the battery, producing the maximum current **Ips** and maximum magnetic field **Bps**, while at the same time, the current **Ipn** and magnetic field **Bpn** of the "N" coils are at their minimum values because they are now connected to the battery through the maximum value of the resistor "R". Figure 21 shows the voltage, current, and magnetic field waveforms flowing through these coils. The voltage induced in the secondary coils "y" is a sinusoidal alternating voltage. The secondary voltage should be zero when the magnitudes of the currents Ips and Ipn are equal. At this point, the magnetic fields **Bps** and **Bpn** induce two voltages of the same magnitude and opposite polarity.

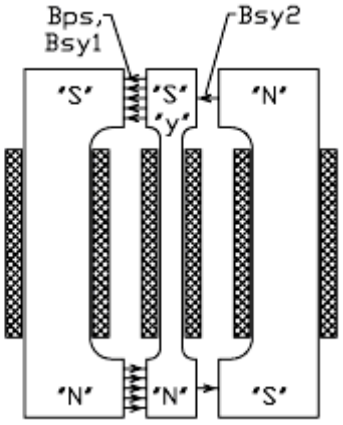


FIG. 16 POSITION 1

The magnetic interaction of the "S", "N", and "y" electromagnets is shown in Fig.16 to Fig.20. Figure 16 illustrates the situation when the brush "O" is at position 1. Here, the current **Ips** and magnetic field **Bps** are at their maximum, while the current **Ipn** and magnetic field **Bpn** are at their minimum values. When the secondary current **Isy** starts flowing, the "y" coils generate a magnetic field **Bsy** which opposes the magnetic field **Bps** in accordance with Lenz's law. As a consequence, a South pole is created at the top of the "y" electromagnet and a

North pole at the bottom. Because magnets of the same polarity repel and opposite polarities attract, it is likely that some of the induced magnetic field **Bsy2** is diverted through the iron core of the "N" electromagnet, which represents a lower reluctance path. And, if the induced magnetic field **Bsy** can be rerouted so as to avoid opposing the magnetic field **Bps** which generates it, then, it might be possible to have an over-unity transformer.

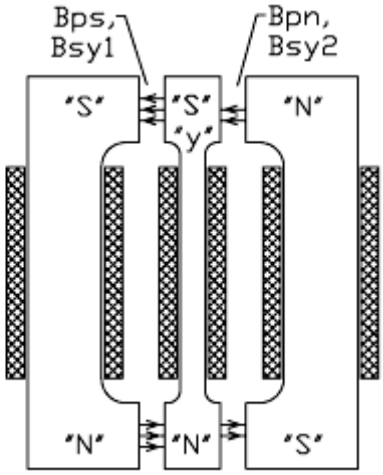


FIG. 17 POSITION 3

Fig.17 illustrates the situation when the sliding contact "O" is at position 3. The primary current **Ips** and the primary magnetic field **Bps** are decreasing in magnitude while the magnitude of the primary current **Ipn** and magnetic field **Bpn** are both increasing. The primary current **Ips** (and **Bps**) is still larger than primary current **Ipn** (and **Bpn**). As shown in the figure, part of the induced magnetic field **Bsy2** is still coupled to the "N" electromagnets.

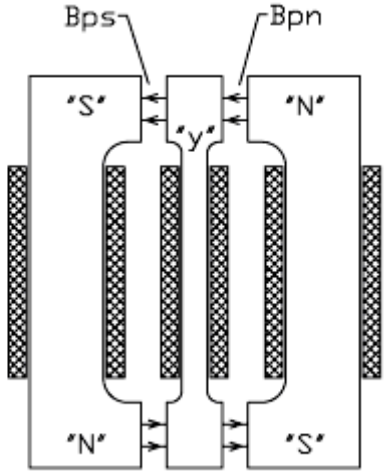


FIG. 18 POSITION M

Fig.18 illustrates the scenario when the brush is at position M. This position is exactly at the centre of the resistor "R" and both currents **Ips** and **Ipn** are of equal magnitudes, and as a result, the magnetic fields **Bps** and **Bpn** are also equal. The net voltage Vsy, current **Isy**, and magnetic field **Bsy** induced in the secondary coils "y" are all zero.

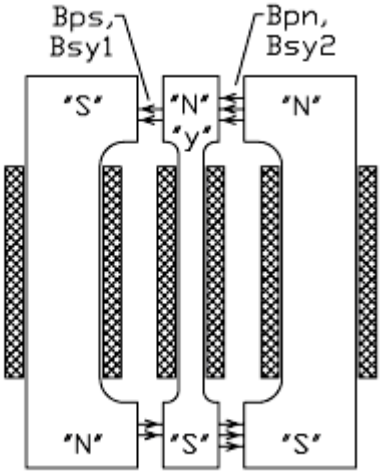


FIG. 19 POSITION 6

Figure 19 shows the situation when sliding contact "O" is at position 6. The primary current **Ips** and the primary magnetic field **Bps** are still decreasing in magnitude while the magnitude of the primary current **Ipn** and the magnetic field **Bpn** are increasing. The primary current **Ips** (and **Bps**) is now of lower magnitude than primary current **Ipn** (and **Bpn**). Because the magnetic field **Bpn** of the "N" electromagnets is stronger than the magnetic field **Bps** of the "S" electromagnets, the polarity of the induced voltage Vsy, current **Isy**, and magnetic field **Bsy** are reversed in accordance with Lenz's law. In this situation, the secondary electromagnets "y" present the north poles at the top and the south poles at the bottom making the "y" and "N" electromagnets to repel and the "y" and "S" to attract. Because of the now higher reluctance of the "N" electromagnets and lower reluctance of the "S" electromagnets, it is expected that part of the induced magnetic field **Bsy** will couple with the "S" electromagnets, and therefore, the effect of Lenz's law is minimised.

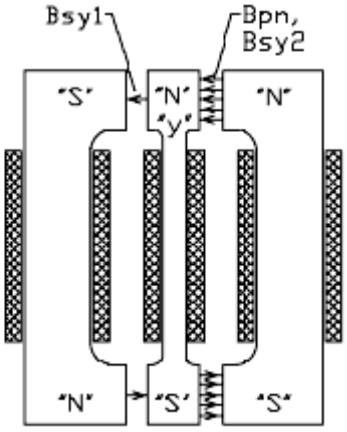


FIG. 20 POSITION 8

Fig.20 illustrates the situation when the brush "O" is at position 8. The primary current **Ipn** and the magnetic field **Bpn** are at their maximum values. The induced secondary voltage Vsy, current **Isy**, and magnetic field **Bsy** are also maximum and of opposite polarities to those which they had at position 1. Again, part of the induced secondary magnetic field **Bsy** is attracted by the "S" electromagnet mitigating the effect of Lenz's law.

References:

[1] http://orbo.es/
http://www.bibliotecapleyades.net/tesla/esp_tesla_27.htm
http://www.alpoma.net/tecob/?page_id=8258

Our thanks are due to the anonymous contributor who produced the above information on the work of Clemente Figuera whom I had never heard of before. Recently, the forum member 'hanlon1492' of the overunity.com forum http://www.overunity.com/12794/re-inventing-the-wheel-part1-clemente_figuera-the-infinite-energy-machine/#.UXu9gzcQHqU has shared a translation of Figuera's complete 1908 patent, lodged just days before he died, and it is reproduced here with thanks to 'hanlon1492' for his work and for freely sharing the results:

PATENT by CLEMENTE FIGUERA (year 1908) No. 44267 (Spain)

Ministry of Development General Board of Agriculture, Industry and Commerce. Patents of Invention. Expired. Dossier number 44267. Instruction at the request of D. Clemente Figuera. Representative Mr. Buforn. Presented in the register of the Ministry in the 31st October 1908, at 11:55 received in the negotiated in the 2nd November 1908.

ELECTRICAL GENERATOR "FIGUERA"

BACKGROUND

If we rotate a closed circuit inside a spinning magnetic field, with the closed circuit positioned at right angles to the lines of magnetic force, a current will be induced in the closed circuit for as long as there is movement, and the sign of that induced current will depend on the direction in which the closed circuit moves.

This is the basis of all magnetic machines and electric dynamos from the original, invented by Pixii, in France and later modified and improved by Clarke to reach the design of the current dynamos of today.

The principle on which this theory is based, has the unavoidable need for the movement of either the induction circuit or the magnetic circuit, and so, these machines are considered to be a transformer of mechanical work into electricity.

PRINCIPLE OF THE INVENTION

Considering carefully what happens in a dynamo in motion, we see that the coil turns of the induction circuit approach and move away from the magnetic centres of the magnets or electromagnets, and those turns, while spinning, pass through sections of the magnetic field of different magnetic strengths, because, while the maximum magnetic strength is in the centre of the core of each electromagnet, this action weakens as the induction coil moves away from the centre of the electromagnet, only to increase again when it is approaching the centre of another electromagnet with opposite sign to the first one.

Because we all know that the effects seen when a closed circuit approaches and moves away from a magnetic centre are the same as when the circuit is motionless and the magnetic field increased and decreased in intensity, since any variation of the magnetic flow traversing a circuit produces an induced electrical current. Then, consideration was given to the possibility of building a machine which would work, based, not on the principle of movement as current dynamos do, but based on the principle of increasing and decreasing the strength of the magnetic field, or the strength of the electrical current which produces it.

The voltage from the total current of the current dynamos is the sum of all of the induced currents generated in every turn of the induction coils. Therefore it does not matter if these induced currents were generated by rotating the induction coils, or by varying the magnetic flux which passes through them. In the first case, a greater amount of mechanical work is required than the amount of electricity generated, while in the second case, the force needed to produce the variation of magnetic flux is so insignificant that it can easily be taken from the output generated by the machine.

Up to the present day, no machine based on this principle has been constructed for the production of large electrical currents, and which among other advantages, has overcome the necessity for motion and so, the energy needed to produce it.

In order to attain the production of large industrial electrical currents, using the principle that electrical current can be provided by just changing the flow of magnetic flux through an induction circuit, the above disclosure should be sufficient, however, as this operating principle needs to embodied in a practical machine, there is a need to describe it in order to fully disclose how to carry out a practical application of this principle.

This principle is not new since it is just a consequence of the laws of induction stated by Faraday in the year 1831: what it is new and claimed in this patent, is the application of this principle to a machine which produces large industrial electrical currents and which, up to now, has only been obtained by transforming mechanical work into electricity.

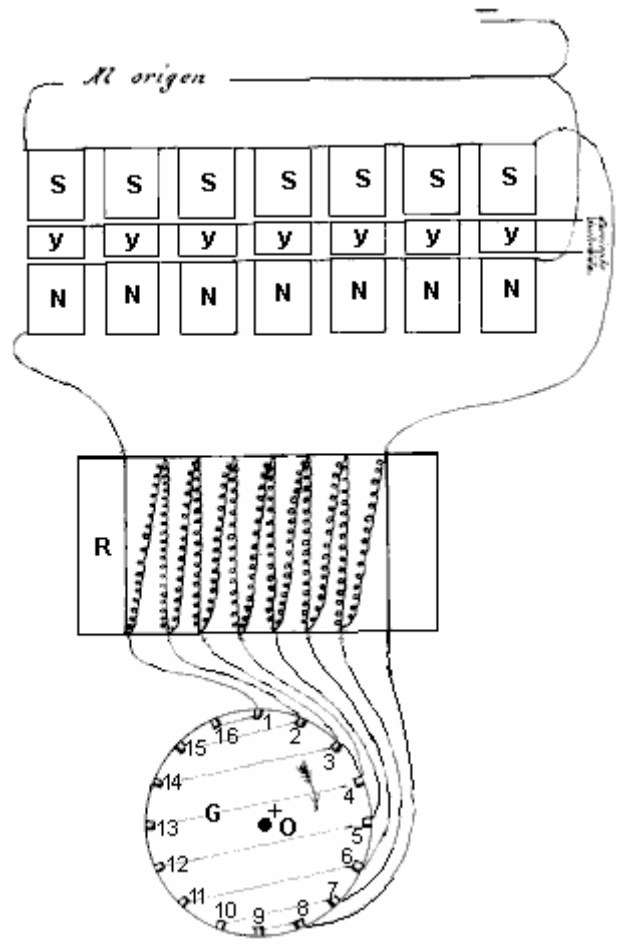
We will therefore, provide a description of a machine based on the above principle disclosed in this patent; but it must be understood, and what is sought is the patent for the application of this principle, that all machines built based on this principle, will be included in the scope of this patent, whatever the form and way that has been used to make the application.

DESCRIPTION OF GENERATOR OF VARIABLE EXCITATION "FIGUERA"

The machine is comprised of a fixed inductor circuit, consisting of several electromagnets with soft iron cores enhancing induction in the induction circuit, which is also fixed in position and motionless, and which is composed of several coils, accurately positioned. As neither of the two circuits spin, there is no need to make them round, nor leave any space between one and the other.

Here what it is constantly changing is the intensity of the excitatory current which drives the electromagnets and this is accomplished using a resistance, through which circulates a operating current, which is taken from one power source and passed through one or more electromagnets, thus magnetising one or more electromagnets. When the current is higher, the magnetisation of the electromagnets is increased, and when it is lower, the magnetisation is decreased. Thus, varying the intensity of the current, varies the magnetic field which crosses through the induction circuit.

To assist in understanding this idea, it is convenient to refer to the attached drawing which is no more than a sketch intended to assist in understanding the operation of the machine built to implement the principle outlined above.



Suppose that electromagnets are represented by rectangles marked '**N**' and '**S**'. Located between their poles is an induction circuit represented by the line of small rectangles marked '**y**'. A resistor 'R', drawn here in a simple form to help understanding of the entire system. Shown as '**+**' and '**-**', is the excitation power, drawn from an external source. As can be seen in the drawing, the different sections of this resistor connect with the commutator bars embedded in a stationary cylinder of insulating material. A sliding-contact brush '**O**', which always connects with more than one contact, rotates, carrying the excitation current. One of the ends of the resistor is connected to the electromagnets **N**, and the other end of the resistor is connected to the electromagnets **S**. Half of the terminals of the resistor go to half of the commutator bars of the cylinder. The other half of these commutator bars are connected directly to the first set of commutator bars.

The operation of the machine is as follows: the brush **O** rotates inside cylinder **G** and is always in contact with two of the commutator bars. When the brush is touching contact **1** the current, flowing from the external source passes through the brush and powers electromagnets **N** to their maximum level of magnetisation, but the current passing through electromagnets **S** is insufficient to magnetise them because the whole of resistance **R** is too great to allow sufficient current to magnetise them. Therefore, the **N** electromagnets are fully powered while the **S**

electromagnets are not sufficiently powered to be magnetised.

When the brush connects with contact **2**, the whole of the current will not flow through electromagnets **N** because it has to pass through part of the resistor. Consequently, some current will pass through the electromagnets **S** because it has to overcome less resistance than in the previous case. This same reasoning applies to the case when brush **O** connects with each of the different contacts around the first semicircle. Then the brush **O** starts to connect with the commutator contacts in the other half, each of which are directly connected to their corresponding commutator contacts in the first half. In short, the resistor has the function of a current-splitter, powering either one set of electromagnets or the other set of electromagnets repeatedly. It can be seen that electromagnet sets **N** and **S** operate in a complementary manner, because while the first set is being progressively powered up, the other set is being progressively powered down. This sequence is repeated continuously causing an orderly a constant variation of the magnetic fields passing through the induction circuit. This action can be maintained by just the simple rotation of a brush or group of brushes which rotate in a circle inside cylinder **G** driven by a small electric motor.

As indicated by the drawing the current, once it has flowed through the electromagnets, returns to the power source where it originated. A small part of the output current from this device can be used to provide the 'external' excitation power mentioned above, thus making the machine self-exciting and to provide the current to operate the small motor which moves the brush causing the switching. Once started with an external power source, that external power source can be removed and the machine will continue to work indefinitely without any external power source.

This invention is really new, very daring, and above all, has huge technical and industrial consequences in all areas. This patent was not applied for until a working machine based on these principles had been built, thus proving the concept to be sound and practical.

ADVANTAGES OF THE ELECTRICAL GENERATOR "FIGUERA"

1. The completely free production of DC or AC electric current of any voltage which can be used for:
   a. Providing a driving force.
   b. Production of light.
   c. Production of heat.
   d. All other existing uses of electricity.

2. There is no need whatsoever for a driving force of any kind or chemical reactions or fuel consumption.

3. Needs little or no lubrication.

4. Is so simple that it can be easily operated by anyone.

5. Does not produce smoke, noise, or vibration when operating.

6. Indefinite operational life.

7. Has a wide range of uses: home management and industrial.

8. Easy construction.

9. Cheap to produce and market

NOTE

A 20-year patent is requested for a "NEW GENERATOR OF ELECTRICITY, so-called "FIGUERA" of variable excitation, designed to produce electrical currents for industrial applications without using either driving force, nor chemical reactions. The machine is essentially characterised by two series of electromagnets which form the inductor circuit, between whose poles induction coils are placed. Both the induction and inductor circuits remain motionless and yet are able to produce a current induced by the constant variation of the intensity of the magnetic field forcing the excitatory current (coming at first from any external source) to pass through a rotating brush which, in its rotation movement, connects with the commutator bars or contacts of a ring distributor or cylinder whose contacts are connected to a resistor whose value varies from a maximum to a minimum and vice versa, according with the commutator bars of the cylinder which operates, and for that reason the resistance is connected to the electromagnets N by one of its side, and the electromagnets S at the other side, in such a way that the excitatory current will be magnetising successively with more or less strength, the first electromagnets, while, simultaneously decreasing or increasing the magnetisation in the second set, determining these variations

in intensity of the magnetic field, the production of the current in the induced, current that we can use for any work for the most part, and of which only one small fraction is derived for the actuation of a small electrical motor which rotates the brush, and another fraction goes to the continuous excitation of the electromagnets, and, therefore, converting the machine to become self-exciting, being able to remove the external power which was used initially to excite the electromagnets. Once the machinery is in motion, no new force is required and the machine will continue in operation indefinitely.

All in accordance with the described and detailed in this report and as represented in the drawings which are attached.

Barcelona, the 30th of October, 1908. Signed: Constantino de Buforn.

*****

There are some practical points which have not been included so far and which need to be mentioned. The Figuera patent shows the electromagnets as just rectangles, and while C-shaped electromagnet cores have been indicated and discussed, there is a distinct possibility that the electromagnet cores are just I-shaped or even a short cylinder which is several times wider than it is tall. These more simple shapes could make it very much easier to construct, although the C-shaped core need only be three straight sections placed together.

It is essential to construct each of the cores of the electromagnets from iron and only iron. While a laminated core does minimise eddy currents, in this application, a laminated core has a major negative magnetic effect (something which is not generally known).

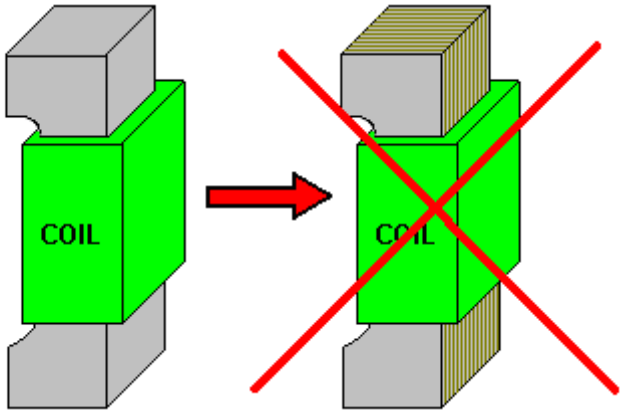


Fig. 22

I have to agree wholeheartedly with the anonymous contributor when he recommends that any attempted replications stay as close to the arrangement shown in the patent drawing, and have seven separate sets of three electromagnets. However, for subsequent experiments, a somewhat easier construction with just one set of electromagnets might be tried, making the electromagnets equal in length to the seven separate units:

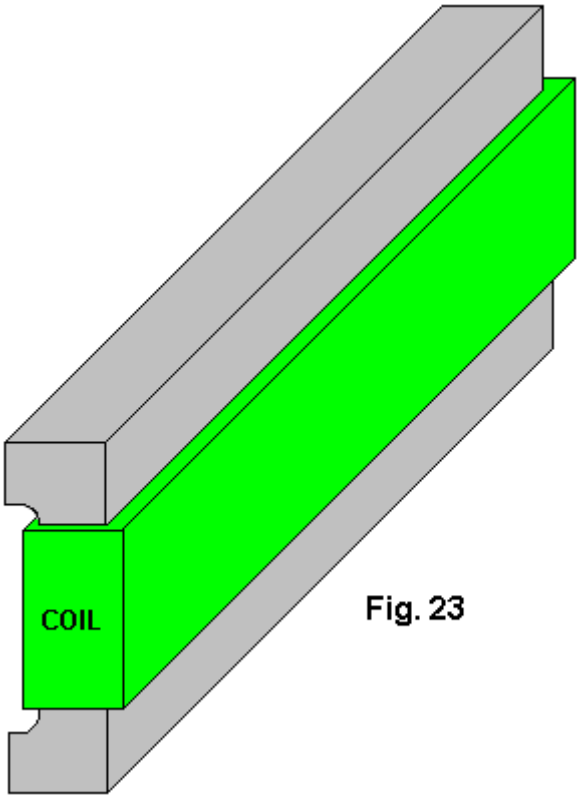


Fig. 23

This arrangement has advantages if the design is taken on into manufacturing as less construction is needed.

Figure 15 shows two electromagnets connected at the top to the battery Minus and at the bottom to the battery Plus. But, one is marked with a North pole at the top and the other with a South pole at the top, so perhaps some explanation would be helpful. If the coils are connected that way, then one will have to be wound in a clockwise ("CW") direction and the other in a counter-clockwise ("CCW") direction:

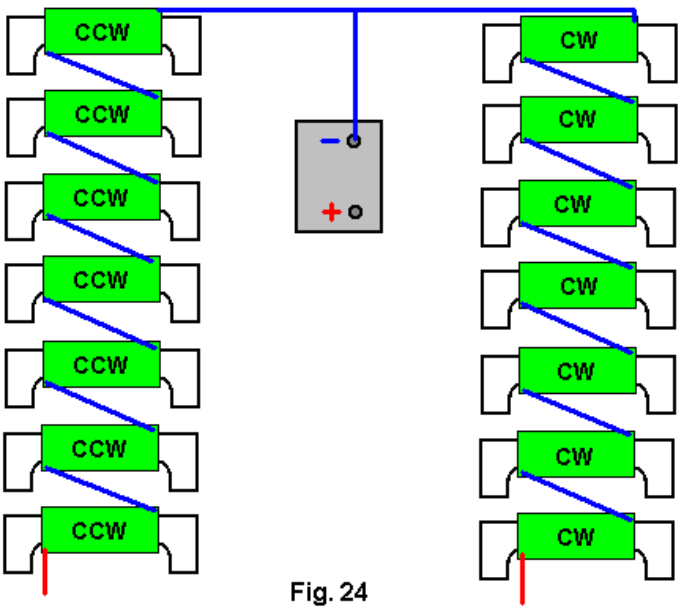


Fig. 24

Or the alternative is to have all of the electromagnets wound in the same way, and adjust the connections:

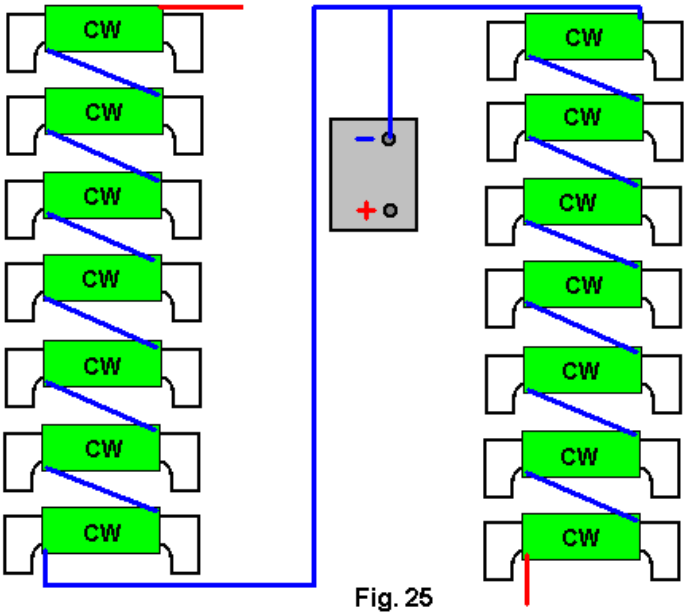


Fig. 25

The Figuera design was implemented more than a hundred years ago, and so Clemente did not have any semiconductors available to him, and so he used a motor-driven commutator arrangement to produce the electrical switching which he needed.

While I am in no way opposed to mechanical switching, especially where prototypes are concerned, there has to be an advantage in using solid-state switching, and while I am by no means an expert in that field, the following suggestions might be useful for experienced circuit builders.

In spite of the wire-wound resistor bank having only eight connection points, the switching has to have sixteen outputs due to the backwards and forwards switching sequence which is used. A solid-state 16-way switching module can be constructed from two CD4017 Divide-by-Ten integrated circuits like this:

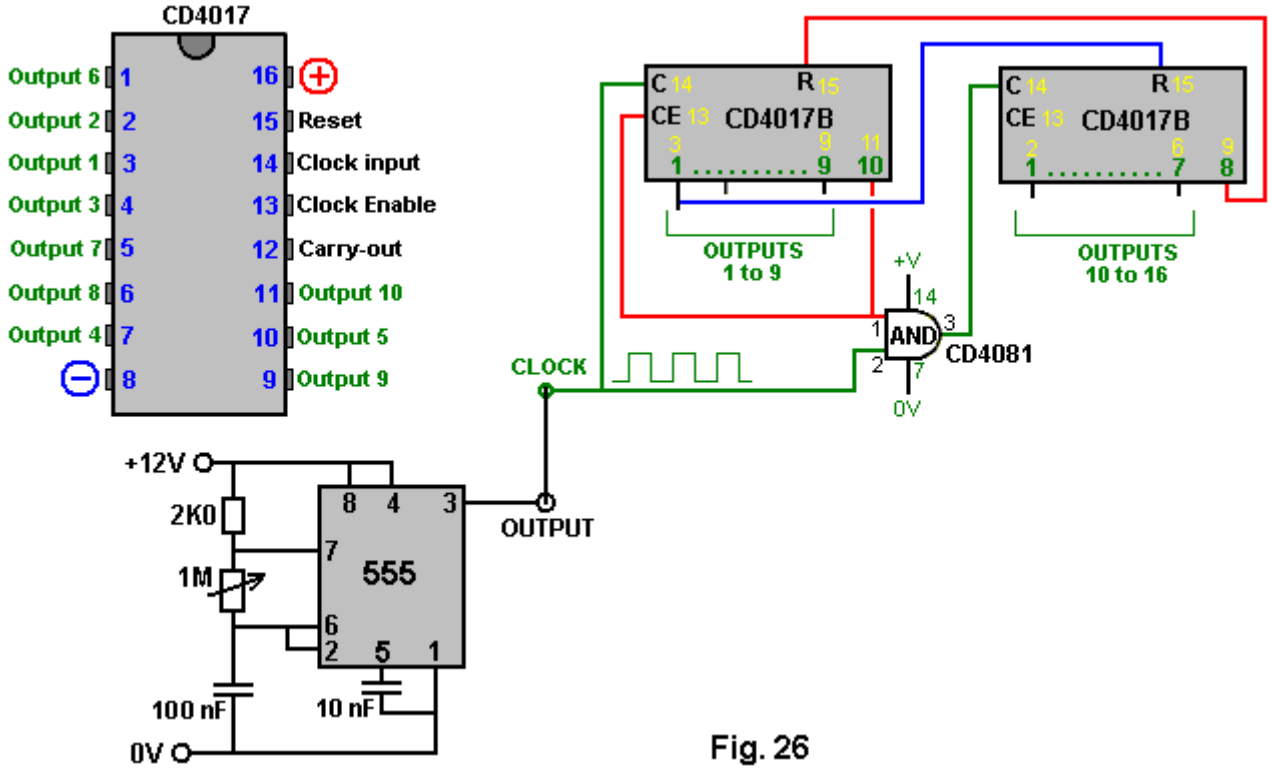


Fig. 26

This arrangement gives sixteen outputs in sequence, so two outputs need to be connected together in order to match the mechanical switching which Clemente used. Presumably, it would not be advisable to connect two outputs directly together, and so an isolation diode (say, a 1N4148 type) would be required on each output.

Please note that the pin connections shown here have been revised as it appears that in this circuit, the output pin 3 of the second 4017 chip does not function as expected:

Output Number	Chip and Pin Nos	Paired with Output	Resistor Connection Point
1	Chip 1 Pin 3	16 (Chip 2 pin 6)	1
2	Chip 1 Pin 2	15 (Chip 2 pin5)	2
3	Chip 1 Pin 4	14 (Chip 2 pin1)	3
4	Chip 1 Pin 7	13 (Chip 2 pin 10)	4
5	Chip 1 Pin 10	12 (Chip 2 pin7)	5
6	Chip 1 Pin 1	11 (Chip 2 pin 4)	6
7	Chip 1 Pin 5	10 (Chip 2 pin 2)	7
8	Chip 1 Pin 6	9 (Chip 1 pin 9)	8
9	Chip 1 Pin 9		
10	Chip 2 Pin 2		
11	Chip 2 Pin 4		
12	Chip 2 Pin 7		
13	Chip 2 Pin 10		
14	Chip 2 Pin 1		
15	Chip 2 Pin 5		
16	Chip 2 Pin 6		

Eight power transistors can be used to energise each resistor connection point in the sequence required. As mechanical switching was used by Clemente, it really did not matter which way round the battery connections were made. We can match his switching exactly by using PNP power transistors (or possibly, P-channel FETs) which would make the arrangement like this (with just two of the eight connections being shown):

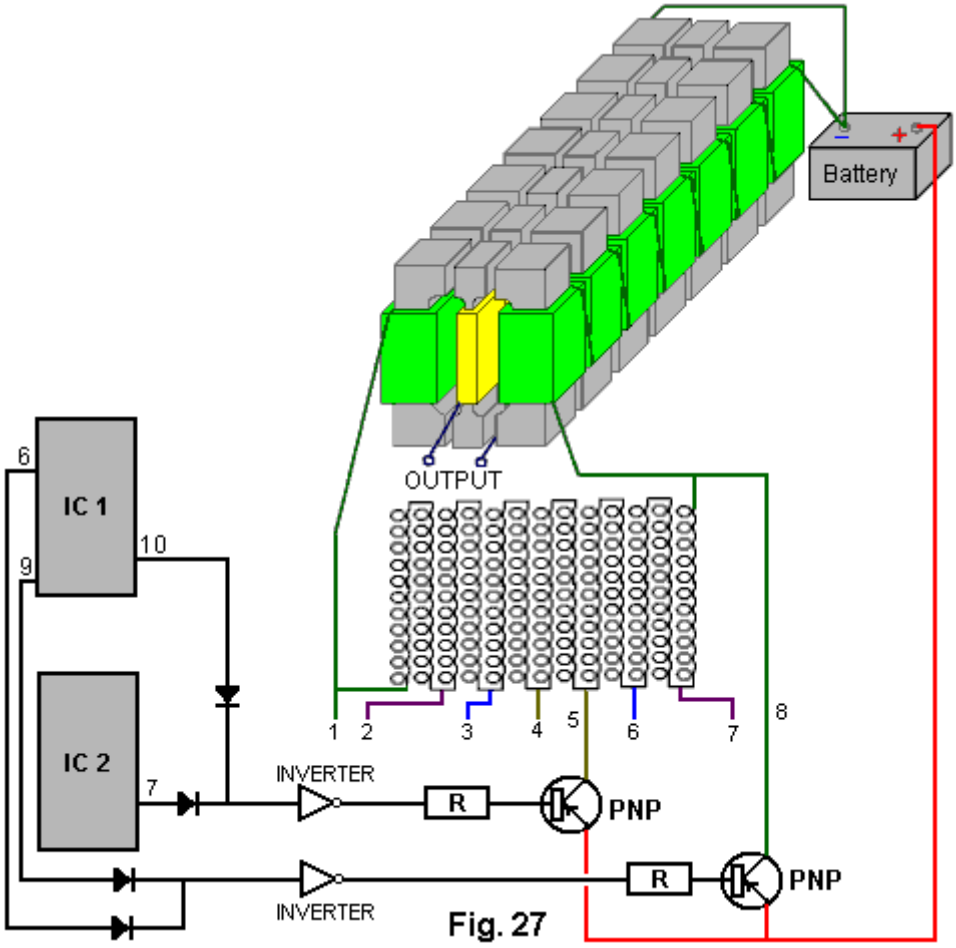


Fig. 27

Or you can reverse the battery for the easier NPN option:

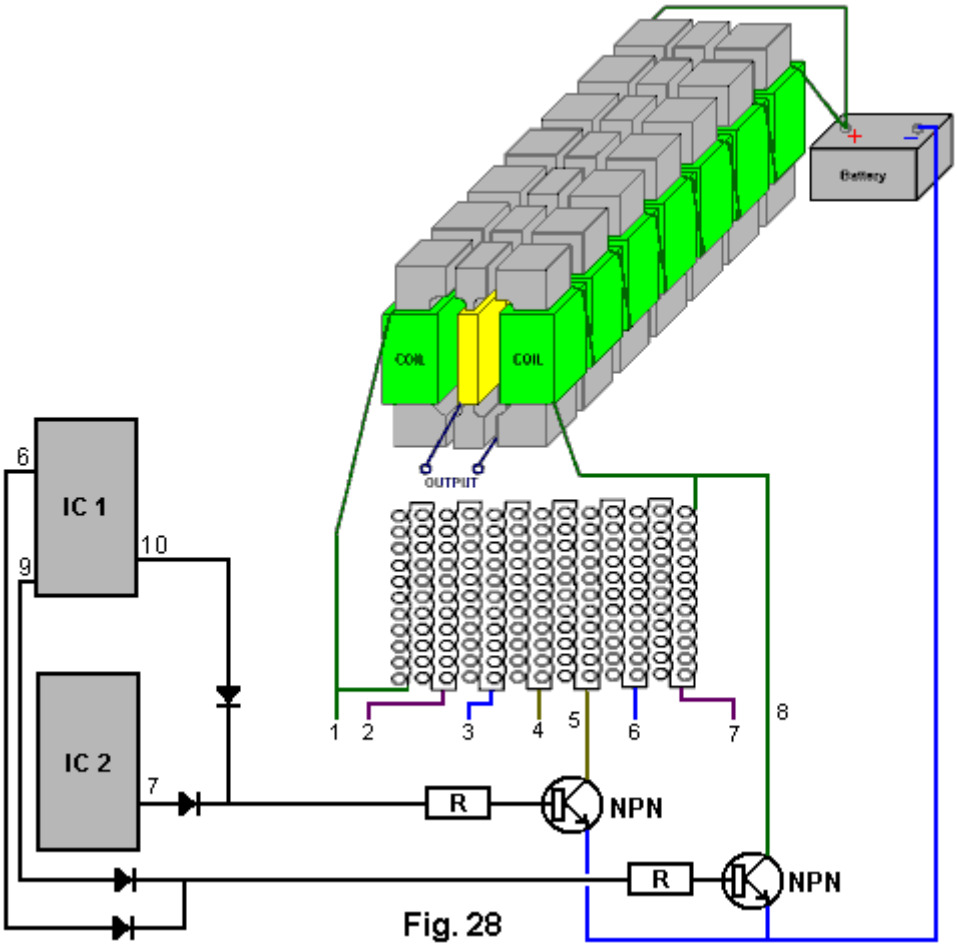


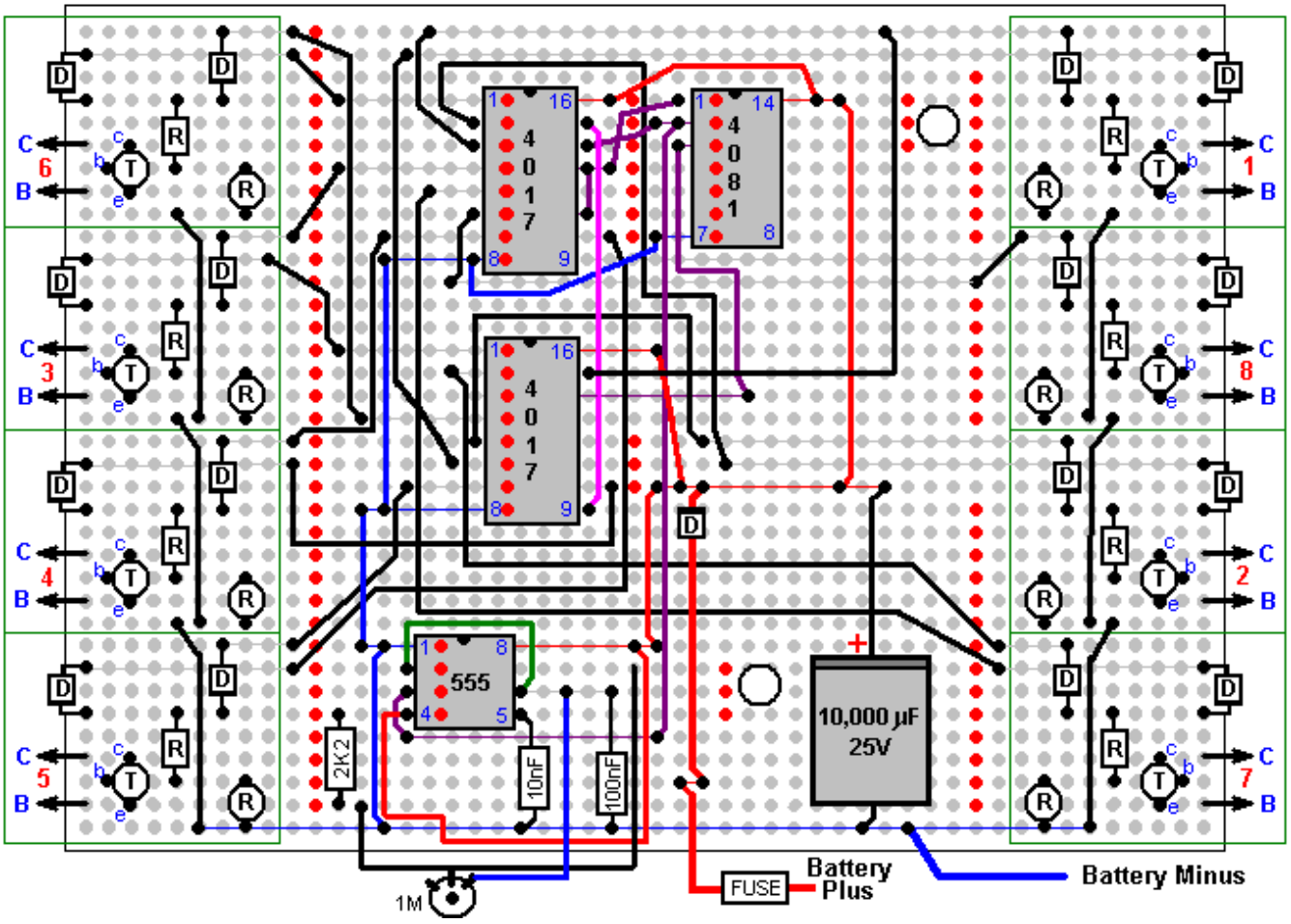
Fig. 28

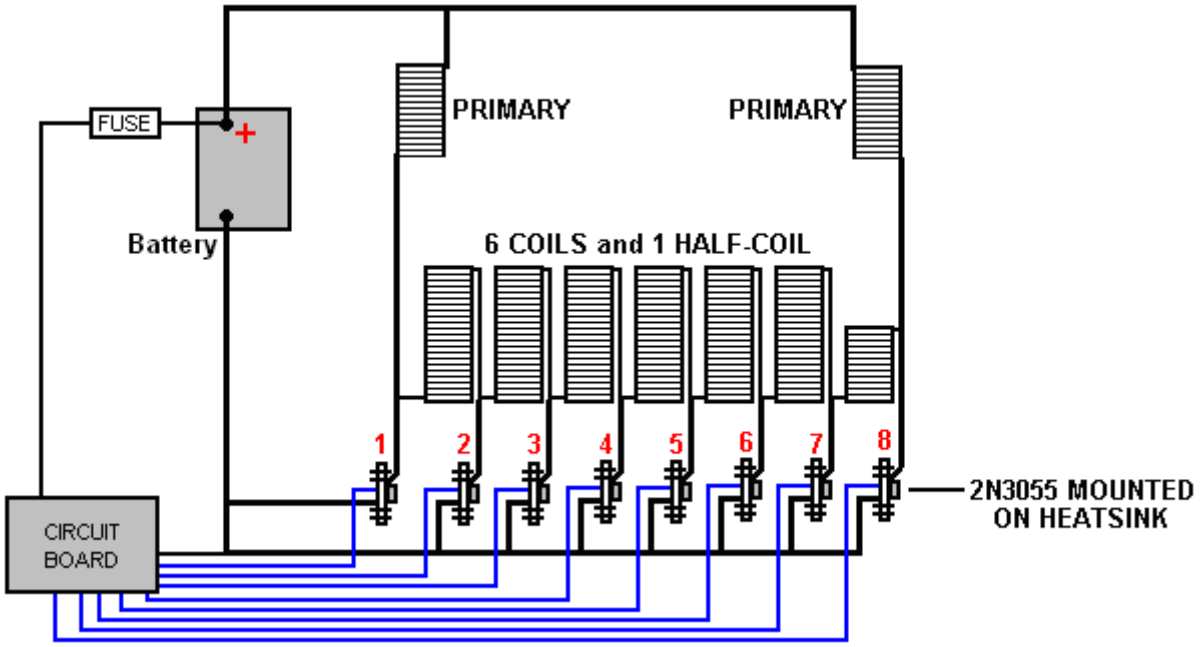
I have been asked by an electronics novice to show a possible construction form for this sort of circuit. I am not particularly good at that sort of thing, but here are a couple of diagrams of a non-optimised layout for a standard size of common stripboard:

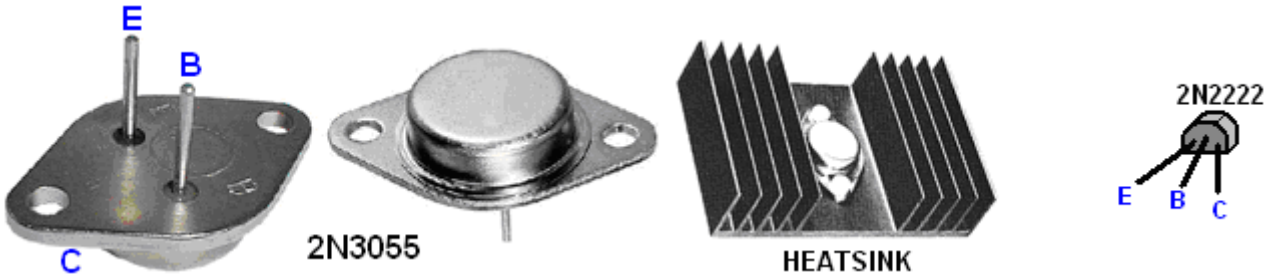
T = 2N2222 Transistor D = 1N4001 Diode R = 4.7K Resistor

C = Connection to 2N3055 Collector B = Connection to 2N3055 Base

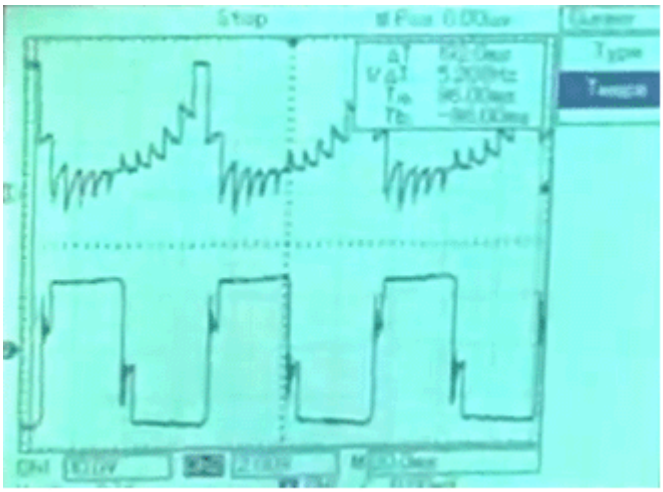
● = Break in the copper strip on the underside of the board



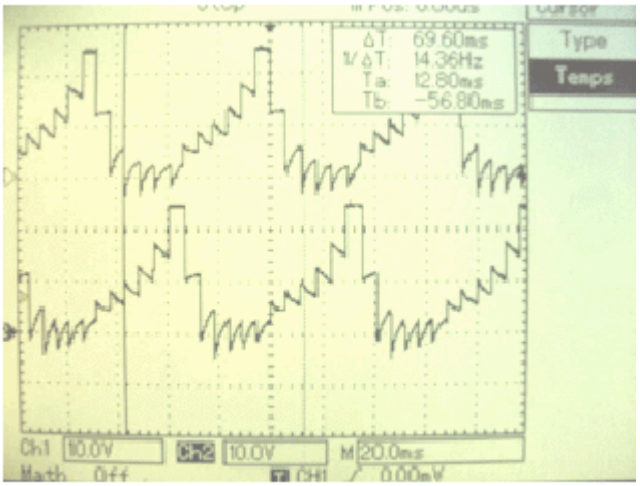




Experienced experimenter 'Woopy' has posted a video of a quick experiment to test the working principle of this Figuera design. It is at http://www.youtube.com/watch?v=HlOGEnKpO-w&feature=g-u-u and in it, he short-circuits the secondary winding, showing that the input power is totally unaffected by the current draw from the secondary.
He shows some very interesting oscilloscope shots:



One primary and the secondary output



Both primaries, showing their phases

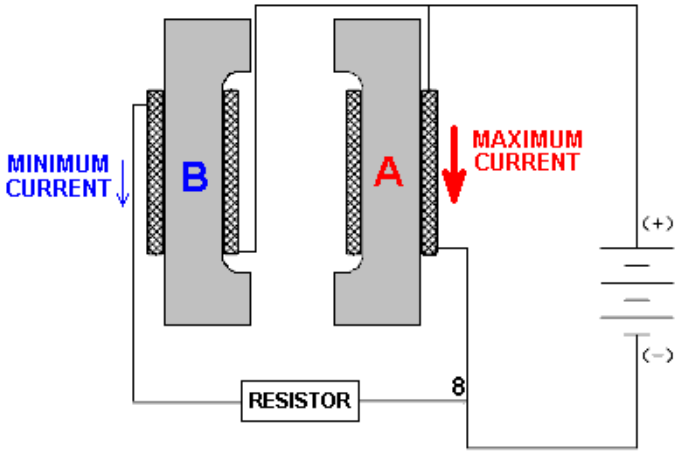
The first screen shot surprises me as it shows clearly that the output is actually an excellent square wave while I would have expected it to be a sine wave as it is coming from a coil which has inductance. The second shot shows very clearly, how the two banks of primary electromagnets operate out of phase with each other thanks to Woopy's mechanical 6-way switching arrangement. It is reported that Mr Figuera ran a 20-horsepower motor with his prototype and if that motor were fully loaded, then that is 15 kilowatts of power, easily enough to power a household.

Please bear in mind that if the electromagnets are made from iron, whether laminated or not, that iron restricts the frequency, probably to 500 Hz or less, and so it is necessary to keep the frequency that low if using a solid-state circuit to drive the transformer. For 60 Hz output with mechanical switching, requires the motor to run at 3,600 rpm which is fairly fast although definitely achievable. Also, the output power will be limited by the current handling capacity of the wire in the secondary winding. The first page of the Appendix shows the current capacities for the standard AWG and swg wire sizes.

Because this Figuera design is so important, being low-voltage, high power and not needing tuning I have recently been asked to explain it in greater detail and suggest some component values for people starting to experiment with it. I am not an electronics expert, and so my suggestions need to be taken as just that, namely, suggestions for a possible starting point for experimentation.

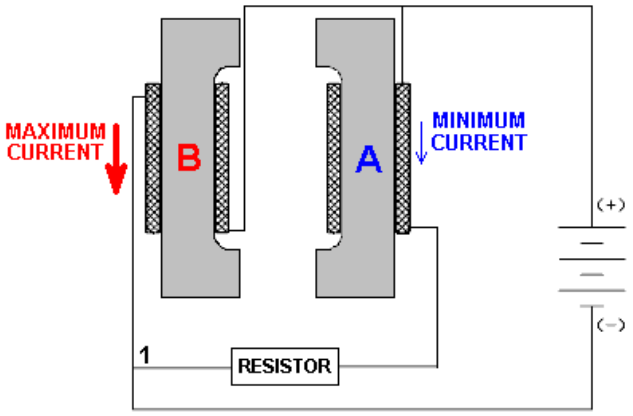
The first point is that the two halves of the primary winding of the transformer become electromagnets when current flows through their windings. The strength of an electromagnet increases as the current flow increases. Large current: strong magnet. Small current: weak magnet.

Clemente Figuera's circuit is arranged so that the current through the windings is made to vary so that when one magnet is strong, the other one is weak. It works like this:



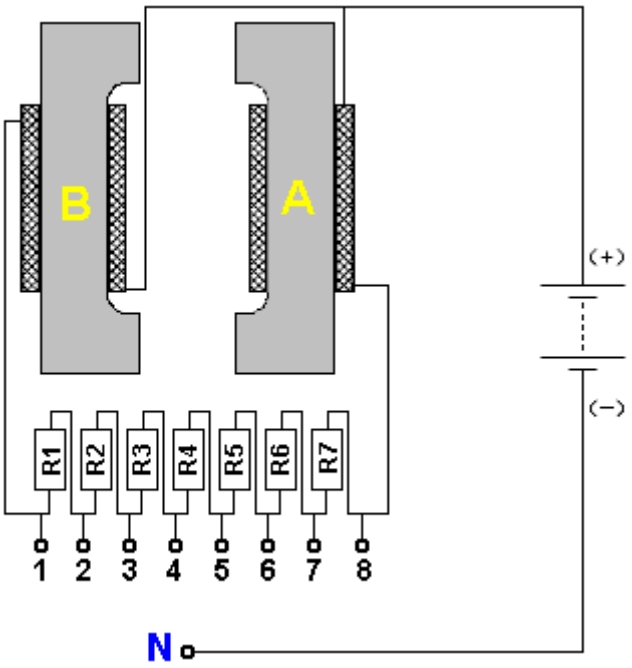
When the mechanical (or transistor) switching connects the battery to point ‘8’ in the previous diagrams, we get the situation shown above. Current from the battery flows directly through the right-hand electromagnet “A”, making it the strongest magnet that it can be at that battery voltage. The electromagnet “B” on the left gets current flow from the battery all right, but that current is reduced because it has to flow through the resistor.

When the switching changes and the battery is connected to point “1” in the previous diagrams, we get this arrangement:



Here, electromagnet “B” is free of the resistor and gets it’s maximum possible current, making it the strongest magnet which it can be at that battery voltage, while electromagnet “A” has it’s current reduced by the resistor getting in the way, making it the weakest magnet it can be when the system is running.

If we switched between these two positions, we would get a square wave style of operation, but Clemente did not do that. Instead, he split the resistor into seven parts (if Fig.14 is drawn correctly, one part having only half the resistance of the other parts). This makes the arrangement like this:



When the battery negative “N” is connected to point “2”, then the current flow through electromagnet “B” is hindered by resistor R1, but the current flow through electromagnet “A” is hindered by resistors R2 and R3 and R4 and R5 and R6 and R7, which together, have a far higher resistance than R1 on its own. This makes the current flow through electromagnet “B” far greater than the current flow through electromagnet “A”.

When the battery negative “N” is connected to point “3”, then the current flow through electromagnet “B” is hindered by resistor R1 and resistor R2, but the current flow through electromagnet “A” is hindered by resistors R3 and R4 and R5 and R6 and R7, which together, have a far higher resistance than resistors R1 and R2. This makes the current flow through electromagnet “B” still greater than the current flow through electromagnet “A”.

When the battery negative “N” is connected to point “4”, then the current flow through electromagnet “B” is hindered by resistors R1, R2 and R3, and the current flow through electromagnet “A” is hindered by resistors R4, R5, R6 and R7, which together, have a higher resistance than resistors R1, R2 and R3. This makes the current flow through electromagnet “B” somewhat greater than the current flow through electromagnet “A” (nearly a balanced flow as resistor R7 is only half the value of each of the other resistors.

When the battery negative “N” is connected to point “5”, then the current flow through electromagnet “B” is hindered by resistors R1, R2, R3 and R4, while the current flow through electromagnet “A” is hindered by resistors R5, R6 and R7, which together, now have a lower resistance than resistors R1, R2, R3 and R4. This makes the current flow through electromagnet “B” somewhat less than the current flow through electromagnet “A”.

When the battery negative “N” is connected to point “6”, then the current flow through electromagnet “B” is hindered by resistors R1, R2, R3, R4 and R5, while the current flow through electromagnet “A” is hindered by resistors R6 and R7, which together, now have a much lower resistance than resistors R1, R2, R3, R4 and R5. This makes the current flow through electromagnet “B” much less than the current flow through electromagnet “A”.

When the battery negative “N” is connected to point “7”, then the current flow through electromagnet “B” is hindered by resistors R1, R2, R3, R4, R5 and R6, while the current flow through electromagnet “A” is hindered by resistor R7, which has a very much lower resistance than resistors R1, R2, R3, R4, R5 and R6 together. This makes the current flow through electromagnet “B” very much less than the current flow through electromagnet “A”.

Clemente has arranged the battery switching sequence to be to points 1, 2, 3, 4, 5, 6, 7, 8, 8, 7, 6, 5, 4, 3, 2, 1, repeating over and over again. This makes the connections to points 1 and 8 to be twice as long compared to the connection times for the intermediate points, giving a sine-wave shape rather than a sawtooth shape.

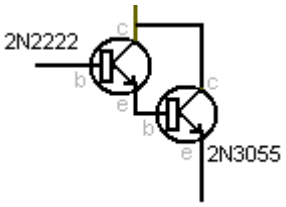
There is current flow through both electromagnets at all times. The current flow is never broken although, as you can see, the intensity of the current flow varies all the time with each electromagnet getting stronger than the other one repeatedly.

The mechanical switching used by Clemente will work perfectly well, although there will be motor noise and wear on the switch contacts. A solid state version will be silent, more reliable and much longer lasting. There are many

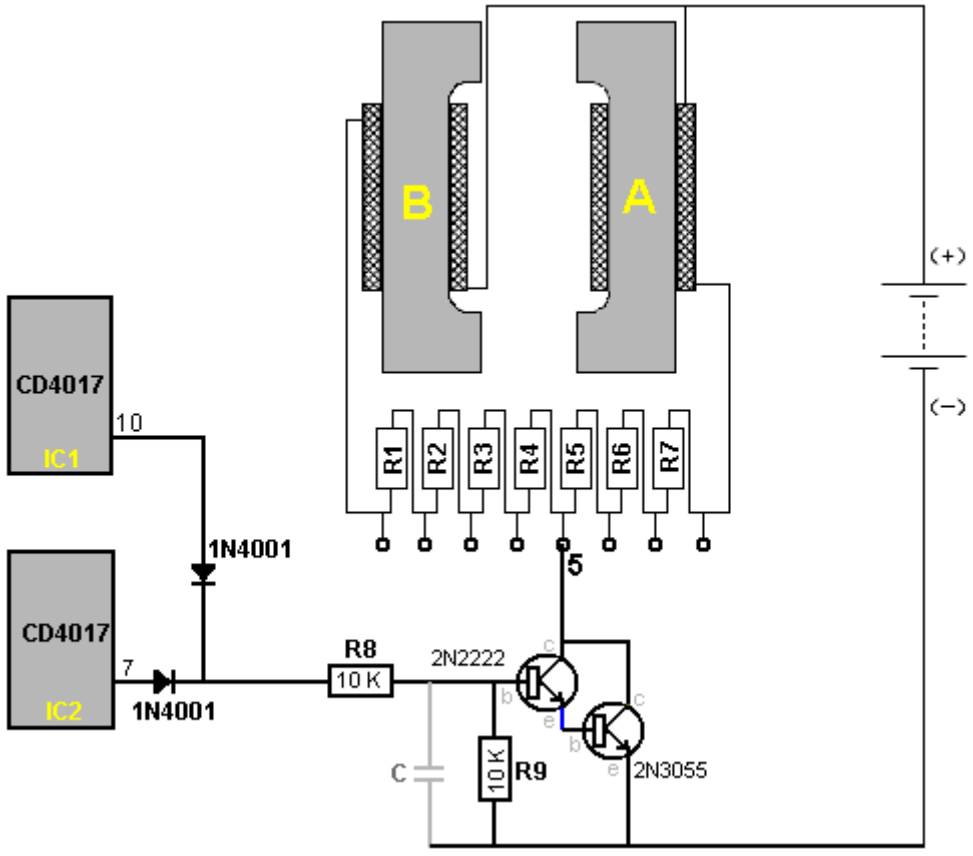
different way to build most electronic circuits and each builder will have his own favourite way of constructing the circuit. This Figuera circuit does not specify the battery voltage and so some people will want to use a twelve volt battery. As many FET transistors need as much as ten volts in order to switch on properly, a twelve volt supply is probably a little low for them, and so I suggest using the older bipolar transistors.

As the transistor has to carry the current which passes through the electromagnets, it needs to be able to handle considerable current flow. The very common 2N3055 transistor can do that (as can many other suitable transistors). The switching rate is very, very slow for a transistor and so speed is not an issue. The voltage is very low, and so that is not an issue either and so the 2N3055 transistor is definitely a possible choice.

In common with most high-power transistors, the current gain is low being between 20 and 30 typically. That means that to switch it on properly, a current of one twentieth of the switched current has to be fed into the base of the transistor. That base current is too high to be convenient, so we can raise the transistor gain to around 6000 by adding in a low-power transistor such as the 2N2222 transistor. The two transistors are connected together in a configuration called a 'Darlington Pair' which looks like this:



In this arrangement, the two Collectors are connected together, while the Emitter of the 2N2222 transistor feeds into the Base of the 2N3055 power transistor. With a high gain of six thousand or so for our transistor pair, we need to limit the current flowing through their combined Base-to-Emitter junction, and so we introduce a current limiting resistor R8 in the following circuit suggestion:



The 10K resistor value shown would limit the transistor current to about nine amps, while a 4.7K resistor would allow around eighteen amps. Each transistor pair is only on for one eighth of the time, but the 2N3055 transistors need to be mounted on a heat-sink. If a single metal plate is used as a heat-sink for all eight 2N3055 transistors, then mica washers (available from the supplier of the transistors) must be used between each transistor and the plate because the Collector of each 2N3055 transistor is it's metal case and in this circuit, the Collectors do not connect to a common point. The mica washers pass heat but not electricity. Separate heat-sinks can, of course, be used.

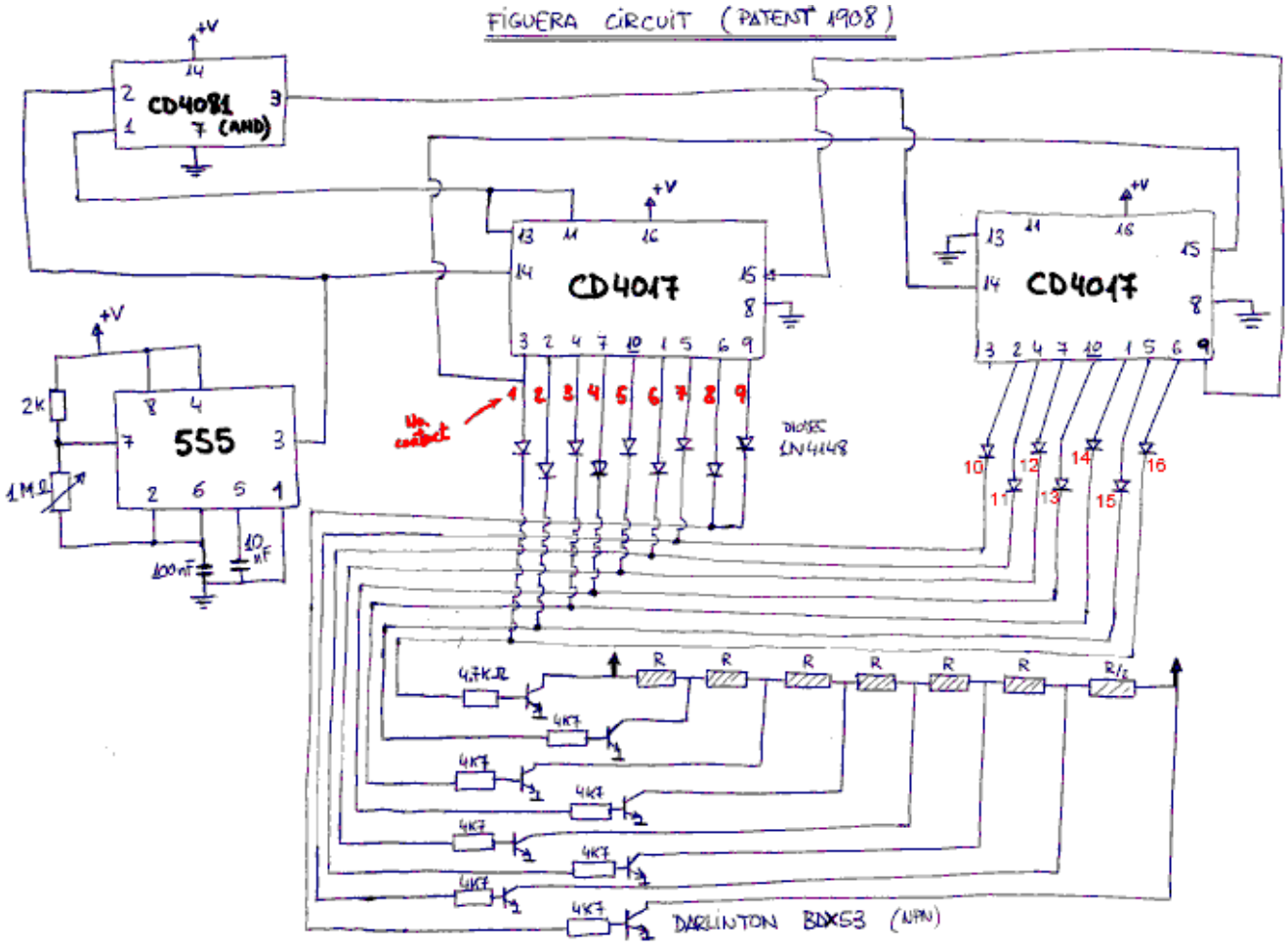
The capacitor “C” in the above circuit diagram will probably not be needed. The switching needs to maintain a constant current flow through both electromagnets. I would expect the 4017 chip switching to be fast enough to allow this to happen. If that proves not to be the case, then a small capacitor (probably 100nF or less) can delay the switch-off of the transistors just long enough to allow the next transistor in the sequence to be switched on to provide the required ‘Make-Before-Break’ switching.

As indicated in the table above, the 4017 pins which feed the transistor pairs through the 1N4148 (or similar) diodes are:

IC1 pin 3 and IC2 pin 6 for resistor connection point 1.
IC1 pin 2 and IC2 pin 5 for resistor connection point 2.
IC1 pin 4 and IC2 pin 1 for resistor connection point 3.
IC1 pin 7 and IC2 pin 10 for resistor connection point 4.
IC1 pin 10 and IC2 pin 7 for resistor connection point 5.
IC1 pin 1 and IC2 pin 4 for resistor connection point 6.
IC1 pin 5 and IC2 pin 2 for resistor connection point 7.
IC1 pin 6 and IC1 pin 9 for resistor connection point 8.

This Figuera design is very attractive as it uses only simple, readily available materials, low voltage and does not require difficult tuning. It also has the potential to be self-powered if part of the output is used to provide a voltage-stabilised power supply for the input power and the remaining output power can be kilowatts if the wire diameters chosen can carry that much current. Chapter 12 explains electronic circuitry in more detail.

A contributor who wishes to remain anonymous does not like the circuit arrangement shown above and prefers this circuit which he has built and tested:

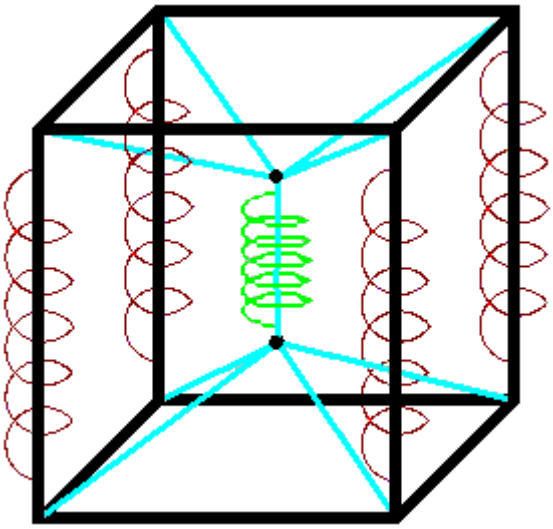


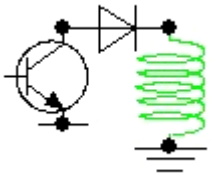
The NPN Darlington transistor BDX53 in this circuit is not available everywhere around the world and if that is the situation in your area, then using a 2N2222 transistor feeding a 2N3055 (or TIP3055) transistor as shown originally will work just the same.

The Alexkor Zero-Back-EMF Coils

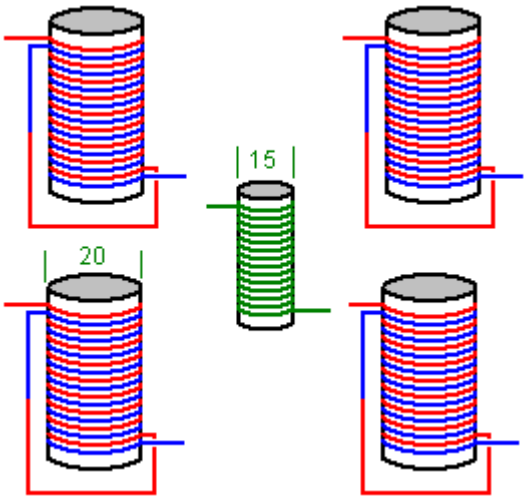
Alex in Russia who has shared several of his motionless pulse-charging systems for batteries, now shares his design which does not appear to have any back-EMF effect on the primary coil. If that is the case, then any increase in output current draw does not have a corresponding increase in the current flowing through the primary coil. That is completely different to the way in which a conventional transformer operates.

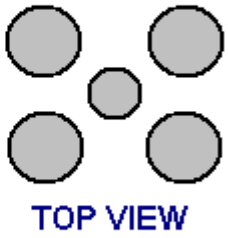
The arrangement is somewhat like the Transmitter / Receiver arrangement of Don Smith and while it looks to be a simple arrangement, is isn't. Alex draws his coil configuration like this:



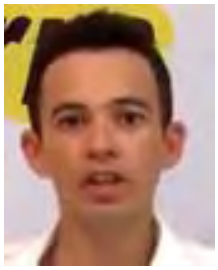


Here, his chosen form of construction is a frame of twelve lengths of 20 mm diameter plastic pipes – four at the top, four at the bottom and four verticals. Each pipe is filled with ferrite powder and there is an output coil wound on each of the four vertical pipes. Suspended in the centre is the primary coil which is 15 mm in diameter. All five coils are wound using 0.5 mm diameter enamelled copper wire (swg 25 or AWG #24). While Alex's drawing shows a single strand of wire, the actual arrangement for the four output coils is that they are wound as a single layer bi-filar coil:





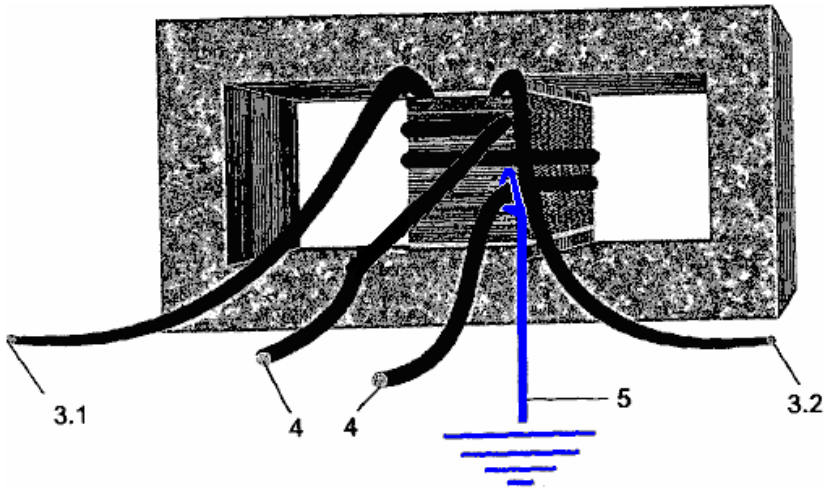
For this, the output coils are wound with two strands of wire side by side, in a single layer along the length of the plastic pipe. Then, the start of one wire is connected to the end of the other wire. As the coils are filled with ferrite, they can operate at high frequency, when the 15 mm primary coil is fed with either DC pulses or an AC sine wave. Each output coil can provide a separate output or the output coils can be connected in series to give a higher voltage or connected in parallel to give a higher output current.



The Self-Powered Generators of Barbosa and Leal

In July 2013, two Brazilian men, Nilson Barbosa and Cleriston Leal, published a series of patents which appear to be very significant. Their patent WO 2013/104042 published on 18th July 2013, is entitled "Electromagnetic device for Capturing Electrons from the Ground to Generate Electricity" and has some very interesting features. It describes a simple device which they describe as an "electron trap". Their patents are written in Portuguese and an attempted translation of three of them is included at the end of the Appendix.

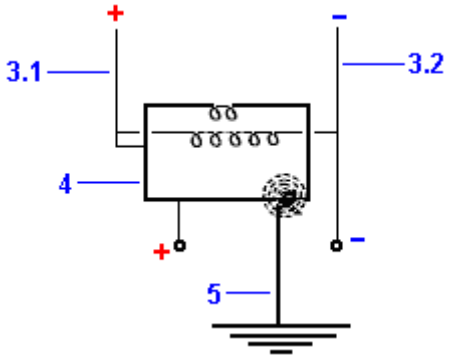
An unusual feature of this design is the fact that it has a continuous conductive loop, in which it is claimed, current flows continuously, even without the need for an applied voltage. Instead, it is the magnetic fields of electromagnets which keep the current flowing. They state that an insignificant amount of input power produces a substantial power output, and they consider a COP of 100 to be about the minimum performance which can be expected from the design. That is a 1 watt input for a 100 watt output. One version of the electron trap looks like this:



The inventors describe their device like this: "this electromagnetic-field-generating device, powered by a power source, produces an electromagnetic field which induces an electric current in a closed conductive circuit, creating an interaction between the magnetic poles of the equipment and the magnetic poles of the earth - through both electromagnetic attraction and repulsion. An endless supply of electrons is drawn from the earth into the conductive closed loop, which is connected to the ground through a conductive interconnected grid. The attracted electrons add to the current already flowing in the conductive closed loop, making power available for driving high-power loads, although the device itself is supplied with only a small amount of power."

One very interesting feature is that the continuous-loop coil formed by wire **4** in the diagram above, is literally, only two turns of wire. The power-gaining mechanism, amazingly, is the earth wire (shown in blue) which is merely wrapped around wire **4** and not directly connected to it as the electron-transfer link is by induction. With this arrangement, the current circulating in the closed loop wire **4**, attracts more electrons from the ground, flowing through the wrapped connection of wire **5,** into wire **4**, augmenting the current flow there by a major amount. Wire **3** can have an alternating voltage applied to it in order to get alternating current in wire **4**, but please understand that the current flowing in wire **4** is **not** the result of the current in wire **3**. If the current in wire **3** is DC, then the current in wire **4** will be DC as this is **not** a conventional transformer, but instead, it is an electron trap, operating in an entirely different way.

The electron trap can be connected in an AC circuit of this type:



Here, the earth wire **5** is wrapped around the continuous loop wire **4**, feeding it additional electrons captured from the ground. The ends of wire **4** are connected together to form the loop, and that connection also forms the positive side of the output (where a DC output is being produced). The magnetic field produced by the current flowing in wire **3**, acts on the electron flow coming from the earth, but as it does not provide any of the electric power flowing in wire loop **4**, the current flowing in wire **3** can be tiny, without affecting the power output.

In their patent WO 2013/104043, also of 18th July 2013, they show several different ways of connecting their electron trap in a useful circuit. For example, like this:

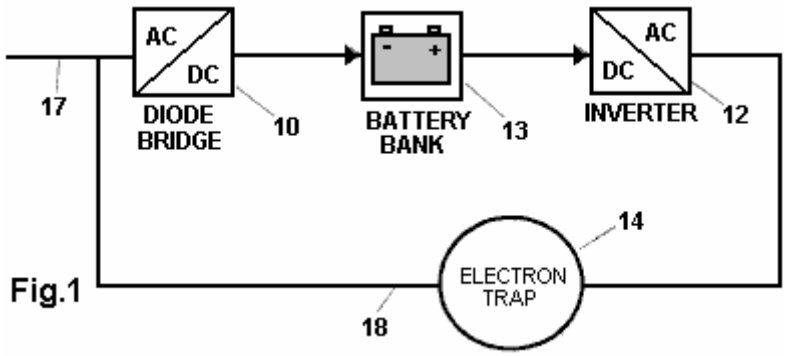
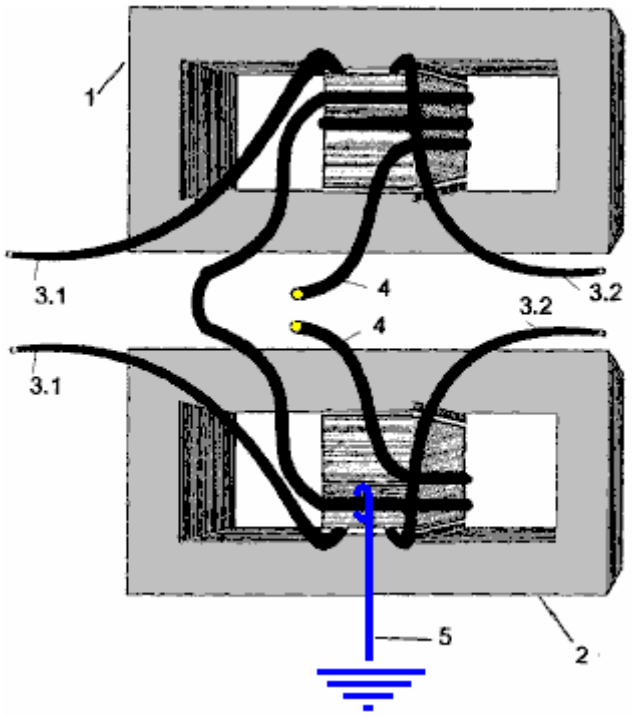


Fig.1

Here, the battery **13**, is used to power an ordinary inverter **12**, which produces a high alternating voltage, in this case, at very low power. That voltage is applied to the wire 3.1 to 3.2 of the electron trap, creating an oscillating magnetic field, which creates an oscillating inflow of electrons into the closed loop wire (4), which creates an amplified electrical output at the same frequency – typically 50 Hz or 60 Hz as those are the common mains frequencies. That amplified power output from the electron trap **14**, is passed along wire **18** to an ordinary diode bridge **10**, and the pulsing DC from the bridge is smoothed and used to replace the battery input to inverter **12**. The battery is now switched out of the circuit and, as well as making the overall circuit self-powered, the power coming from the electron trap is used to recharge the battery if it needs recharging (and/or, perhaps, to charge the batteries of an electric car). Because the electron trap needs almost no input power at all, the input power to the inverter is very small, and so a good deal of additional AC power can be drawn off through cable **17**, and used to drive powerful electrical loads, with no electrical power being needed from the battery. Being self-powered, the COP value for the circuit is infinity.

Just as there are several different ways of using an electron trap in a circuit, there are several ways of constructing and connecting an electron trap. While it is possible to arrange the components so that the power output is 2-phase or 3-phase, here we will just deal with the ordinary, household, single-phase power supply.

The first variation is to use more than one frame. Two frames can be connected like this:



This is the actual drawing from the patent and it presents a slight problem in that it is not physically possible to implement the number **4** wire in the way shown. Each frame will have two complete turns wound on it, although the drawing does not show this. Because of the inaccuracy of the drawing, I am not able to say if the coil turns on frame **2**, are in the same direction as those on frame **1**. There are four possible ways of winding these 2-turn coils when interconnecting them, so perhaps experimentation can be used to determine which method works best.

With this two-frame arrangement, there is just the one earth wire **5**, as before, again, it is wrapped around wire **4** rather than being physically connected to it. The continuous wire loop **4** has two ends as before, but there are now two **3.1** wire ends and two **3.2** wire ends. The Portuguese translation programs produce highly questionable results for this area of the patent, but I gather that the inventors intend the two 3.1 ends to be connected together and the two 3.2 ends to be connected together, and then the joined ends are treated exactly as before, effectively putting the two windings in parallel.

One disadvantage of this design is that it is not portable due to the earth connection. Barbosa and Leal deal with this problem in their patent WO 2013/104041 of the same date where they show a method of constructing an electron trap which collects excess electrons from the air. If you feel that there are no excess electrons in the air, then consider the fact that all of the aerial designs in chapter seven all extract and use those electrons. Also, consider the amount of electricity in a lightning strike, where much of the electrical energy comes from the air, and remember that world wide, there are between 100 and 200 lightning strikes every second.

The free-electrons-in-the-air electron trap is somewhat more complicated than the earth-wire electron trap, with four pairs of coils (**3** and **4**) being mounted inside two aluminium hemispheres (**1**):

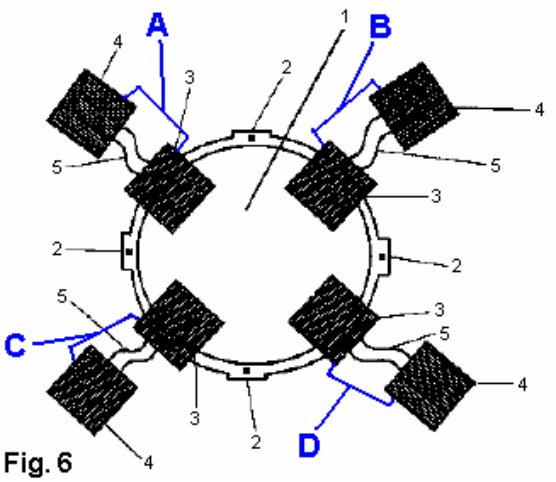


Fig. 6

The methods for using the air-electrons trap are the same as those for the earth-wire electron trap.

An earth-wire video demonstration is here: http://www.youtube.com/watch?v=iRSP7h73u-Q with 22 watts producing 6 kilowatts.

An attempted translation of the three Barbosa/Leal patents is here:

WO Patent 2013/104043 **18th July 2013** **Inventors: Nilson Barbosa and Cleriston Leal**

ELECTRIC ENERGY GENERATION SYSTEM WITH FEEDBACK

Note: These three patents are in Portuguese and what is shown here is a low-quality attempt at translation into English using a translation program. The originals can be downloaded free from: http://worldwide.espacenet.com/singleLineSearch?locale=en_EP.

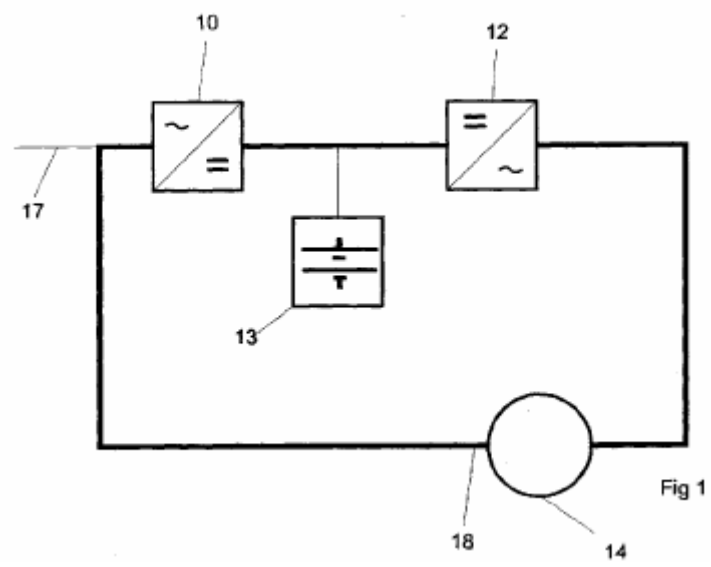


Fig 1

Abstract:
The present invention relates to electric energy generation equipment comprising a basic circuit formed by a rectifier (**10**), for example, an AC/DC converter connected in series to an inverter (**12**), for example, a DC/AC converter, and a bank of batteries (**13**) connected in series between the rectifier (**10**) and the inverter (**12**). An electron-capturing element (**14**), which can be either a free space electron-capturing element or, alternatively, an earth electron-capturing element, is connected in series to the basic circuit formed by the rectifier (**10**), the inverter (**12**) and the battery assembly (**13**). The bank of batteries (**13**) powers the basic circuit because it is connected to the system. Consequently, the inverter (**12**) converts direct current into alternating current and supplies this

current to the electron-capturing element (**14**). After receiving the electric current from the inverter (**12**), the electron-capturing element (**14**) starts capturing electrons from the alternating current and powering the rectifier (**10**), which converts the alternating current into a direct current in order to recharge the bank of batteries (**13**) and power the inverter (**12**) which powers the electron-capturing element, closing the feedback loop, and also providing electric energy for consumption by external loads.

WIPO Patent Application WO/2013/104043 Filing Date: 01/11/2013
Application Number: BR2013/000016 Publication Date: 07/18/2013

Assignee: EVOLUÇÕES ENERGIA LTDA (Rua Santa Tereza 1427-B Centro - Imperatriz -MA, CEP -470 - Maranhão, 65900, BR)

SELF-POWERED ELECTRICITY GENERATOR.

Technical field

The present invention relates to a device for generating electricity, in particular self-powered equipment for generating electricity.

Description of the Related Art

There are many methods for generating electricity using electromagnetism, but all of these are electromechanical devices using magnets and have limited generating capacity and an ecological impact which makes them unsuited to large scale projects.

Objectives of the Invention

The aim of this invention is the sustainable generation of electricity, using a generator which is able to produce large amounts of electricity from an extremely low input current, which initially is supplied by a bank of batteries, but subsequently is supplied by the output from the generator which is also able to power external loads.

The above objective, and other objectives, are achieved by the present invention through the use of a typical Uninterruptible Power Supply circuit comprising of an AC/DC rectifier feeding a battery bank which powers a DC/AC inverter, which is connected to a device to trap electrons from space (as described in Brazilian patent application No. BR1020120008378 of 13th January 2012) or alternatively, a device which extracts electrons from the Earth (as described in Brazilian patent application No. BR1020120008386 of 13th January 2012), which then passes the extracted electrons to the AC/DC rectifier, charging the battery bank, thus closing the loop as well as providing electricity to power external loads.

The self-powered system for generating electricity from the present invention can be fixed or mobile. It is fixed when using electron capture from the earth due to the ground connection, or mobile when using electron capture from space.

The self-powered electricity generating system of this invention may be configured in several different ways, each using the same inventive concept but using different arrangements of components. Different versions include single-phase, two-phase or three-phase versions, producing outputs of any power and voltage.

Brief Description of the Drawings

The present invention will now be described with the aid of drawings, but this patent is not limited to the versions and details shown in these drawings, although they show additional details and advantages of the present invention.

The drawings:

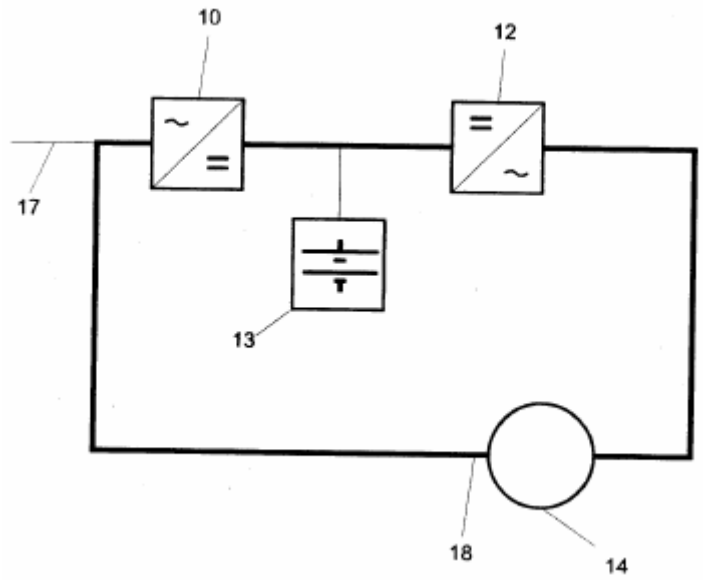


Figure 1 - shows a basic circuit system for self-powered electricity generation of the present invention

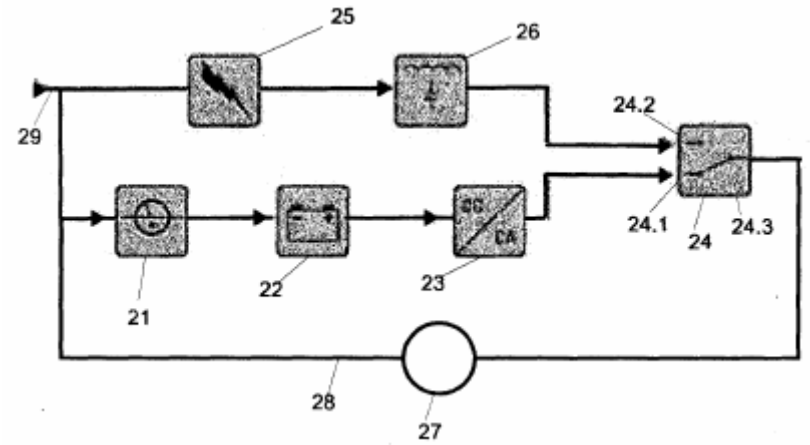


Figure 2 - shows a first embodiment of the constructive system for self-powered electricity generation of the present invention;

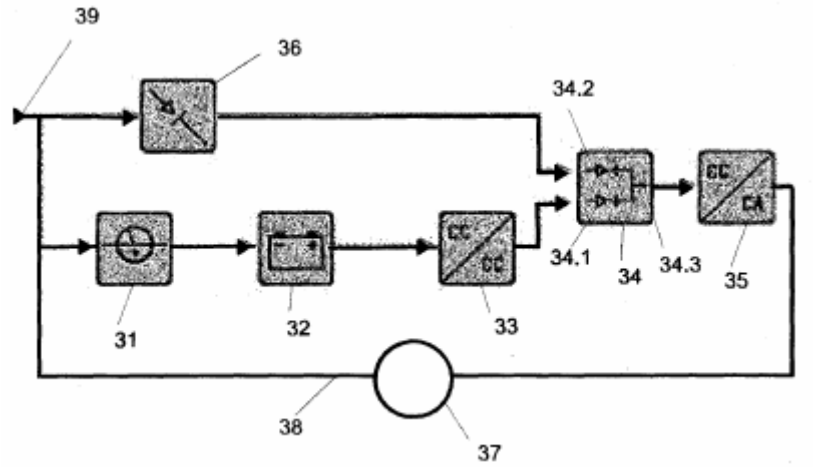


Figure 3 - shows a second embodiment of the self-powered system for generating electricity of the present invention;

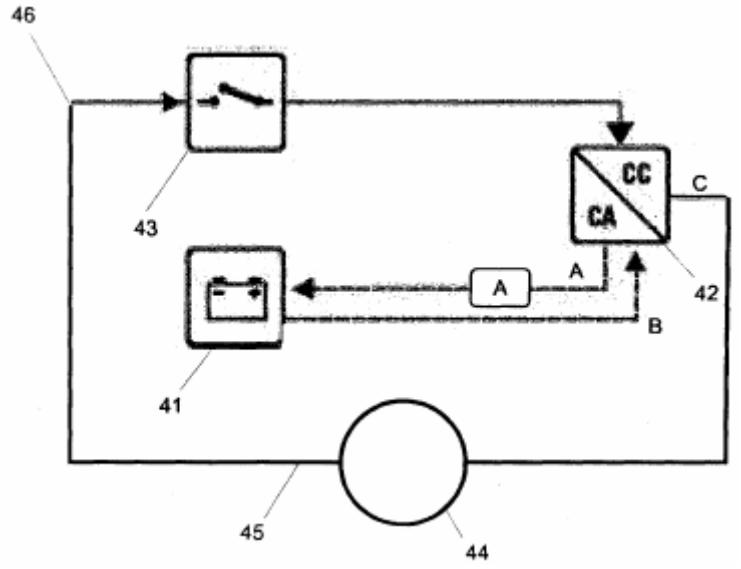


Figure 4 - shows a third embodiment of the self-powered system for generating electricity of the present invention;

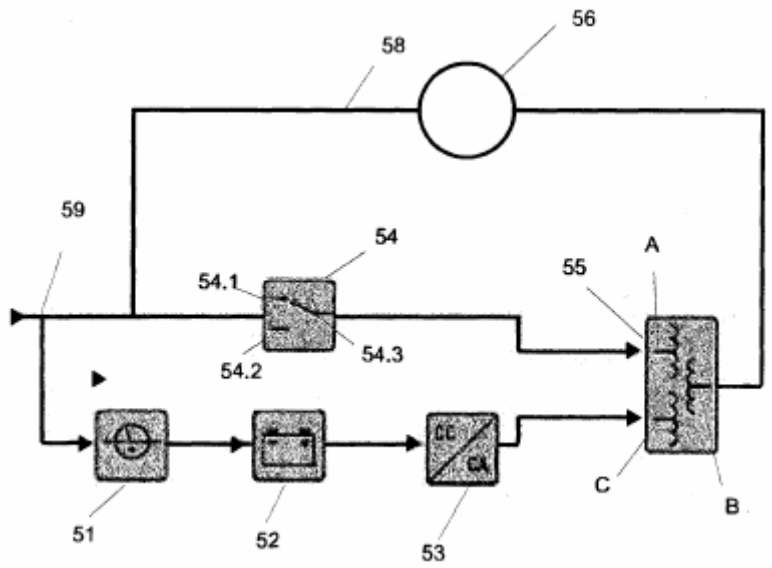


Figure 5 - shows a fourth embodiment of the self-powered system for generating electricity of the present invention;

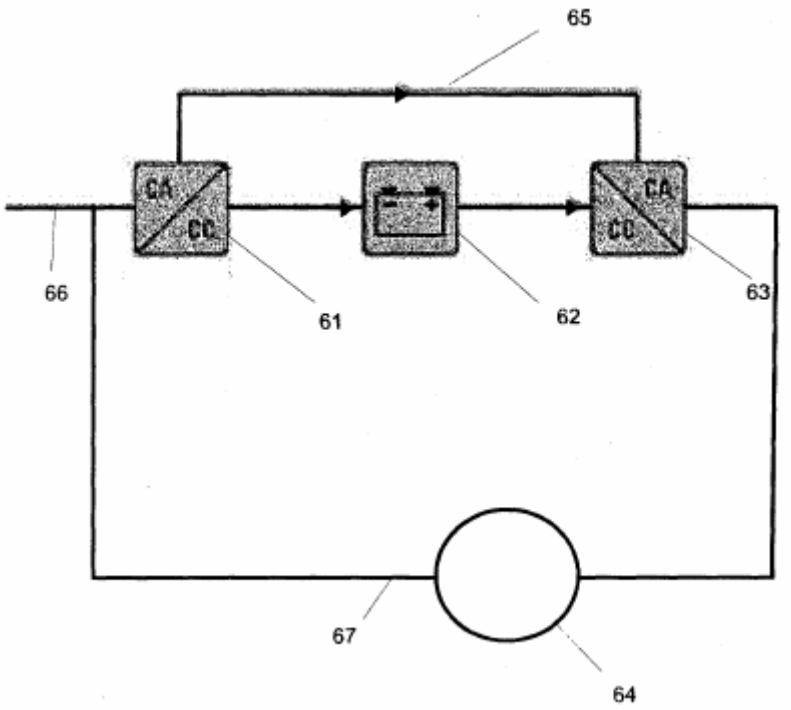


Figure 6 - shows a fifth embodiment of the self-powered system for generating electricity of the present invention;

Detailed description of the Invention:

There are different ways of closing the self-feeding cycle depending on the circuit configuration chosen. Some of these arrangements are shown in **Figures 2** to **6**, wherein the main circuitry continues to oscillate, continuously generating instant electricity.

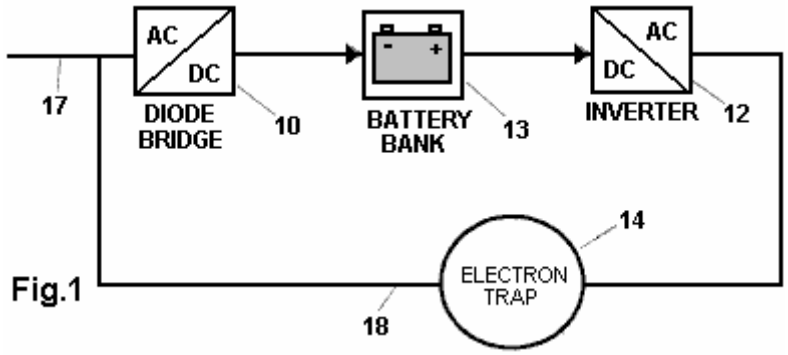


Fig.1

As shown in **Fig.1**, the self-powered system for generating electricity comprises a basic circuit consisting of a rectifier (AC/DC converter) **10** which is connected in series to an inverter (DC/AC) **12**. A bank of batteries **13** is connected between the rectifier **10** and the inverter **12**. The output from the DC/AC inverter **12**, connects to an electron-trap **14** which can extract electrons from space (as described in Brazilian patent application No. BR1020120008378 of 13th January 2012) or alternatively, extracts electrons from the Earth (as described in Brazilian patent application No. BR1020120008386 of 13th January 2012).

When connected, the battery bank **13** provides power to the DC/AC inverter **12** which converts the direct current into alternating current and provides current to the electron-trap **14**. The output of the electron trap **14** is passed through wire **18**, to the AC/DC bridge rectifier **10**, which keeps the battery bank charged as well as powering the DC/AC inverter **12**. Additional power is passed to external equipment through wire **17**.

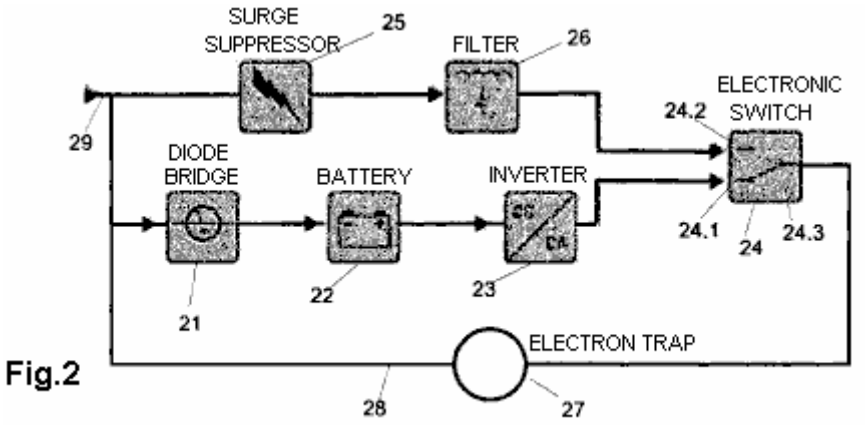


Fig.2

Fig.2, shows another embodiment of the system of this self-powered electric power generation equipment. It comprises a typical Uninterruptible Power Supply circuit of a battery charger (AC/DC converter) **21** connected to a drive device (a DC/AC inverter) **23** and between them, a battery bank **22** forming the basic circuit. Additional devices are an electron-trap **27** which may collect free electrons from space (as defined in Brazilian patent application No. BR1020120008378 of 13th January 2012) or, alternatively, collects electrons from the Earth (as described in Brazilian patent application No. BR1020120008386 of 13th January 2012). The 3-phase electronic switch **24** normally connects **24.1** to **24.3** connecting the electron trap **27** to inverter **23**. Connected in parallel is the surge suppressor **25**, which, when activated, via filter **26**, causes switch **24** to disconnect the **24.3** to **24.1** link and instead, connect **24.3** to **24.2**.

An alternative arrangement for use in emergency situations, is to use the system no longer self-powered. For this, the system is comprised of a power input from an external power source, directly to the interconnection point **29** to provide power to surge suppressor **25**, which provides power to feed the power output point **28** in order to power external loads. When the electron-trap **27** is turned off, the electronic transfer switch **24** reverts to its default position which connects point **24.1** to point **24.3** causing the circuit to function, once again, in its self-feeding mode. As soon as the electron sensor **27** provides sufficient power to the over-voltage sensor **25**, it operates the transfer switch **24** through filter **26**, ending the self-feeding phase and supplying energy directly to the power output point **28**, in order to feed external loads.

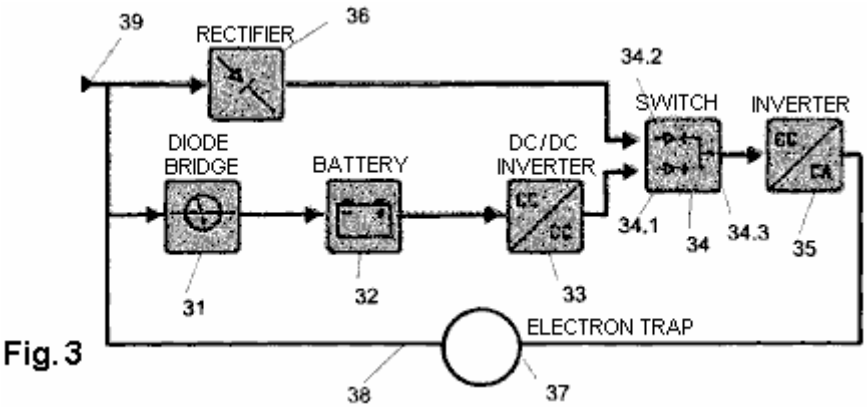


Fig. 3

Fig.3 shows another embodiment of the self-powered system for generating electricity, comprising a device which includes the basic circuit of a typical Uninterruptible Power Supply, consisting of a battery charger (AC/DC converter) **31** connected to a drive device (inverter DC/AC) **35** and attached to them, a battery bank **32**. This basic circuit together with other devices is connected to an electron-trap **37** for collecting free electrons from surrounding space or, alternatively, an Earth-connected electron trap **37**. We have then, a bank of batteries **32** connected to the DC/DC converter **33**, which is connected to the phase transfer switch **34 / 34.1** which is connected to point **34.3**, which connects to the inverter **35**, and so, the electron-trap **37**.

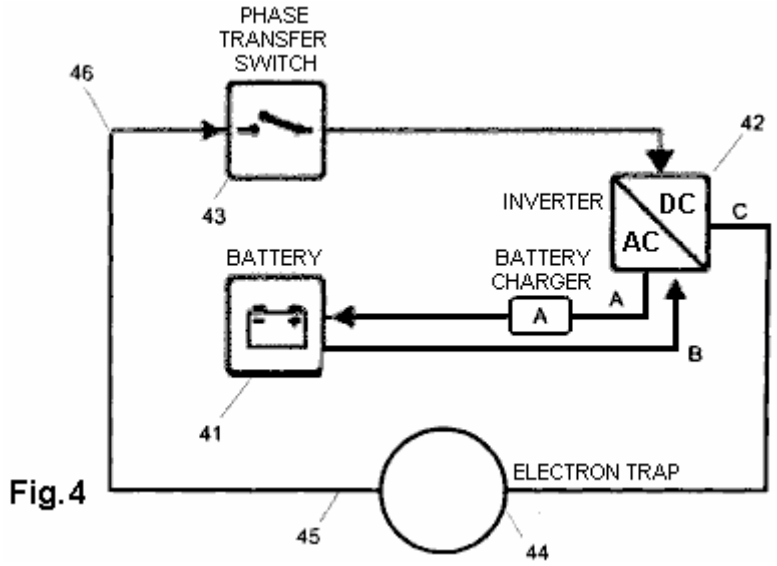


Fig. 4

Fig.4 shows another embodiment of the system for self-powered electricity generation which is comprised of a basic circuit of a typical uninterruptible power supply, consisting of a battery charger (AC/DC converter) **A** connected to an inverter (DC/AC) **42** and attached to them, battery bank **41**, and this basic circuit together with other devices are connected to a free space electron-capture device **44** or an earth-connection electron-trap **44**. Comprising thus, a battery charger **A** connected to a battery bank **41**, which is connected in series with inverter **42** at point **B** which is in series with point **C** of inverter **42** which is in series with the electron sensor **44**, which is in series with the phase transfer switch **43** via the three-phase load output connection point **45**. The phase transfer switch **43** is in series with the inverter **42**, which is connected in series the (AC/DC converter) battery charger **A** feeding the battery bank **41**.

An alternative construction for use in emergency situations, in which the system ceases to be self-powered, the system may include power input from an external power source, via the interconnection point **46**, thus providing electricity output **45**, to power external loads. The battery bank **41** provides power to the inverter **42** which converts the direct current into alternating current and feeds the electron trap **44**. The phase transfer switch closes when the batteries need recharging.

Sensor **44** captures electrons, producing alternating current, which feeds the phase transfer switch **43** with alternating current input power. The phase transfer switch **43** feeds the inverter **42** which charges the batteries, closing the self-powering loop which provides power at the output **45**, feeding both the power input and any external loads.

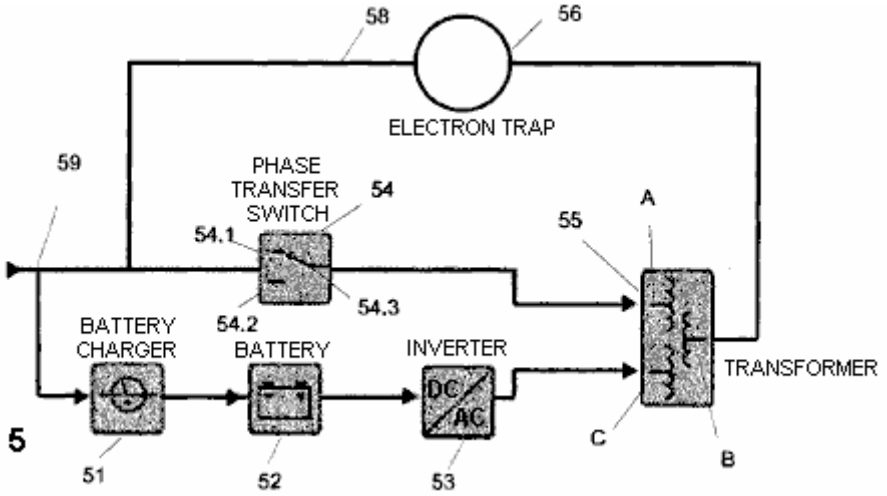


Fig. 5

Fig.5 shows another embodiment of the system for self-powered electric power generation equipment comprising a circuit which includes a typical uninterruptible power supply comprising a battery charger (AC/DC converter) **51** connected to a DC/AC inverter **53** and attached to them, a battery bank **52**. This basic circuit together with other devices are connected to a space free-electron capture device **56** (as defined in Brazilian patent application No. BR1020120008378 of 13/1/12) or, alternatively, an earthed free-electron collector **56** (as defined in Brazilian

patent application No. BR1020120008386 of 13/1/12). This then comprises a battery charger **51** which is connected in series with a battery bank **52**, which is connected in series with the inverter **53**, which is connected in series with the transformer **55** at its point **C**, which is in series with its point **B** which is in series with the electron collector **56**, which is in series with the battery charger **51** which is connected to the load exit point **58**, which is also the circuit entry point **59**, which is in series with the phase transfer switch **54** section **54.1**, which is connected to terminal **54.3**, which is in series with point **A** of the transformer **55** which exits at point **B**. Points **A** and **54.3** as well as the parallel points **54.1** and **54.2**, are all parallel to the battery charger **51**, the battery bank **52**, the inverter **53** and to point **C** of the transformer **55**.

An alternative construction for use in emergency situations, in which the system ceases to be self-powered, the system may include an external power input point **59**, allowing phase transfer switch **54** to provide power output **58**, to feed external loads. Battery bank **52** provides power to the inverter **53**, which converts the direct current into alternating current, feeding point **C** of the transformer, which comes out at points **B** and **A** of the transformer **55**. Point **B** of the transformer feeds the electron-trap **56** producing alternating current which feeds the battery charger **51**, recharging the battery bank **52**.

The battery charger **51** is connected in parallel with the transfer switch **54** via connection points **54.1** and **54.3**, feeding point **A** of the transformer, which comes out at point **B**. Point **A** of the transformer and the switch transfer points **54.3** and **54.1** are in parallel to the battery charger **51**, the battery **52**, the inverter **53** and point **C** of the transformer **55**.

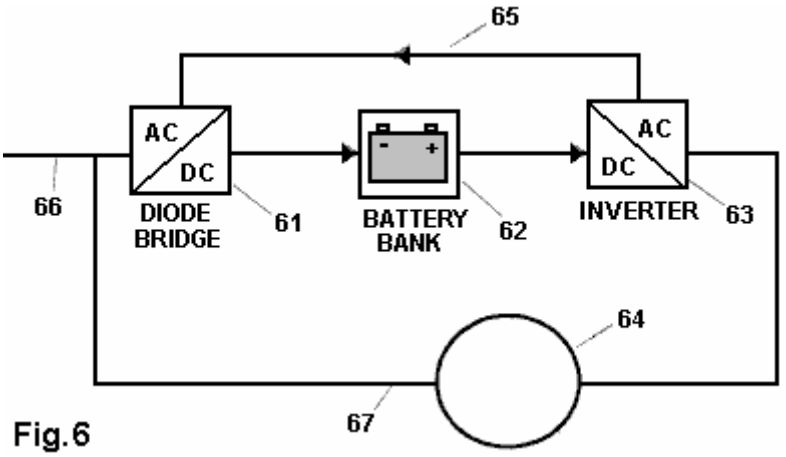


Fig.6

Fig.6 shows another embodiment where a rectifier **61** is connected to an inverter **63** and a battery bank **62**, and to a space free-electron trap **64** or alternatively, an earth electron trap **64** comprising thus, a delta (AC/DC) converter **61**, which is connected in series to a battery bank **62**, which is connected in series with the (DC/AC) inverter **63**, which is in series with the electron collector **64** which is connected in series with the delta converter (AC/DC) **61** whose AC part is in series with the alternating AC current inverter **63** via a connecting wire **65** which is in parallel with the DC part of the delta converter **61** with the battery bank **62** and the DC part of inverter **63**. An alternative construction for use in emergency situations, in which the system ceases to be self-powered, the system may comprise a power input from an external power source, via the interconnection point **66** connected to the delta converter **61**, the output **67** supplying power, to the external loads.

Battery bank **62** provides power to the inverter **63**, which converts the direct current into alternating current, powering the free-electron collector **64**. The captured electrons from collector **64** form an alternating current which feeds the delta converter **61** via an output power load wire **67**.

The alternating part of the three-phase delta converter **61** is fed with alternating current from inverter **63** via connecting wire **65**, which is connected in parallel to the continuous DC delta converter **61**, which feeds the battery bank **62** and with the continuous portion the inverter **63**, closing the cycle of self-feeding and supplying power at the output **67**, which is the output power point.

Having described examples of preferred embodiments, it should be understood that the scope of the present invention encompasses other possible forms of construction, using the electron collectors connected to a basic circuit of a typical uninterruptible power supply of energy, known as a UPS, comprising a rectifier device (an AC/DC converter) **10**, connected to one inverter (DC/AC converter) **12**, and attached between them, an energy storage device (typically, a battery bank).

A very important part of the above patent is the device described as a “collector of free-electrons”, either from the earth or from space. We have to go to the patent applications mentioned above to find the details of these designs:

Application Number: BR2013/000015, Publication Date: 07/18/2013, Filing Date: 01/11/2013

Assignee: EVOLUÇÖES ENERGIA LTDA (Rua Santa Tereza 1427-B Centro - Imperatriz, MA- CEP -470 - Maranhäo, 65900, BR)

ELECTROMAGNETIC ELECTRON TRAP FOR ELECTRIC POWER GENERATION

Technical Field

The present invention refers to electromagnetic equipment for electric power generation or alternatively for thermal power generation. More specifically equipment capable of producing abundant electricity and thermal energy from a tiny amount of input electrical energy

Description of the Related Art

According to Lenz's law, any induced current has a direction such that the magnetic field it generates opposes the change in magnetic flux which produced it. Mathematically, Lenz's Law is expressed by the negative sign (-) that appears in the formula of Faraday's Law, as follows.

The magnitude of the induced emf (ε) in a conducting loop is equal to the rate of change of magnetic flux (Φ_B) with time:

$$\varepsilon = -\frac{d\Phi_B}{dt} \qquad \text{Equation 1}$$

As an example of application of Faraday's Law, we can calculate the electromotive force induced in a rectangular loop that moves in or out, with constant speed, a region of uniform magnetic field. The magnetic field flux through the surface limited by the loop is given by:

$$\phi = xLB \qquad \text{Equation 2}$$

and its variation in time:

$$\frac{\Delta\phi}{\Delta t} = \left(\frac{\Delta x}{\Delta t}\right)LB = vLB \qquad \text{Equation 3}$$

So:

$$\varepsilon = vLB \qquad \text{Equation 4}$$

and if the coil has a resistance (R) and the induced current:

$$i = \frac{\varepsilon}{R} = \frac{vLB}{R} \qquad \text{Equation 5}$$

A conductor traversed by an electric current immersed in a magnetic field undergoes the action of a force given by:

$$F = I\vec{L} \times B \qquad \text{Equation 6}$$

Thus, the effect of the current induced in the loop appears as forces Ff, and F - FM. The first two cancel each other out and the third is cancelled by an external force P_{EXT} needed to maintain the constant speed loop.

As the force FM must oppose the force F_{EXT}, current (i) induced in the loop by varying the magnetic flux must have the meaning indicated in **Fig.3**. This fact is a particular example of Lenz's Law.

Considering the experimental activities discussed with Faraday's law, when a magnet approaches a coil, the induced current in the coil has a direction as shown in **Fig.1**. This generates a magnetic field whose north pole is

facing the north pole of the magnet, that is, the field generated by the induced current opposes the motion of the magnet.

When the magnet is moved away from the coil, the current induced in the coil has a direction opposite to that shown in **Fig.1**, thereby generating a magnetic field whose south pole is facing the north pole of the magnet. The two poles attract each other, that is, the field generated by the induced current opposes the movement of the magnet away from the coil. This behaviour is present in all current power generators, and known as 'engine brake' is highly undesirable as it increases the resistance and so, the energy loss.

When two electromagnetic coils are placed facing each other, as shown in **Fig.2**, there is no current in either of them. At the instant of power-up of one of the coils, the current in the coil, generates an induced current in the second coil. When powered up, the current in the coil goes from zero to its maximum value, and then remains constant.

Thus, when the current is changing, the magnetic field generated by it, (whose north pole faces the second coil) is also changing and so the magnetic flux of this field through the second coil is also changing. Then there is a current induced in the second coil whose sense is such that the magnetic field it generates tends to decrease the flow mentioned above, that is, its north pole confronts the north pole of the first field coil.

When the power switch is opened, the current in the first coil drops from its maximum value to zero, and correspondingly its magnetic field decreases. The flux of the magnetic field in the second coil also decreases, and the induced current now flows in the opposite direction. This current flow direction produces an enhancing magnetic field, that is, it has a south pole facing the north pole of the field of the first coil.

Thus, there is a realisation of the principle of conservation of energy, expressed by Lenz's law, wherein any induced current has an effect which opposes the cause that produced it. Assuming that the induced current acts to favour the variation of the magnetic flux that produced the magnetic field of the coil, it would have a south pole facing the north pole of the approaching magnet, causing the magnet to be attracted towards the coil.

If the magnet were then released, it would experience an acceleration toward the coil, increasing the intensity of the induced current and thus create an enhanced magnetic field. This field, in turn, would attract the magnet with increasing force, and so on, with a continuing increase in the kinetic energy of the magnet.

If energy were to be withdrawn from the magnet-coil system at the same rate at which the kinetic energy of the magnet increases, then there would be an endless supply of energy. So it would be a perpetually operating motor, which would violate the principle of conservation of energy. Therefore, it can be concluded that current generators feature a large energy loss during the generation of electricity.

Objectives of the Invention

An objective of the present invention is to contribute to the generation of sustainable energy, proposing an electromagnetic machine capable of producing abundant electricity from an extremely low input of electrical energy.

The above objective and other objectives are achieved by the present invention by a device comprised of at least one electromagnetic field-generating device (without a core or with at least one core) powered by an electrical power source (without a core or with at least one core) having their coils, or sets of coils, wound on at least one common conductive member in a closed circuit which itself has a polarised voltage which is connected to at least one conductive interconnection element which is connected to a grounding grid, these interconnections creating a new technical effect, namely, the appearance of an electric current which keeps circulating in a closed conductive loop, and which can therefore be used to power external loads.

The device which is the object of the present invention operates as follows: the electromagnetic field generating device, powered by a power source, produces an electromagnetic field which induces an electric current in a closed conductive circuit, creating an interaction between the magnetic poles of the equipment and the magnetic poles of the earth - through both electromagnetic attraction and repulsion. An endless supply of electrons is drawn from the earth into the conductive closed loop, which is connected to the ground through a conductive interconnected grid. Attracted electrons add to the current already flowing in the conductive closed loop, making power available for driving high-power loads, although the device itself is only supplied with a small amount of power. Thus, advantageously, the device which is the object of the present invention, acts as a trap for electrons from the earth and this allows the generation of electricity.

Advantageously, the present electromagnetic equipment generates either electricity or thermal energy, providing access to this new source of energy is through an electromagnetic field. The interconnections of the components

of the electron-trap of the present invention, cause an advantageous new technical effect, namely, the appearance of an electric current which keeps circling in the conductive closed circuit, with or without voltage being applied and even without a load being connected to the loop - provided that the electron-trap is connected.

The proposed sensor can also be used to generate thermal power, depending on the form in which you want to use the effect of the flow of electrical current produced in this electromagnetic equipment.

For the generation of thermal energy in amounts proportional to the power of the electron-trap, through the movement of electrons in the conductive closed loop itself, the resistance should be increased by increasing the number of turns around the cores in the conductive element of the closed circuit, and in that instance, the coils of the electromagnetic field generating device, will then be made of heat-insulated electrical circuit components, bearing in mind the required temperature which is to be produced. The thermal energy generated by the electron-trap can be used in any application from domestic to industrial applications.

This technology can also be used for various technical purposes in electric machines. By "electrical machines", it should be understood to include: static electrical machines, transformers, ballasts, rotating electrical machines, synchronous machines, dual power supply machines, current rectifiers in synchronous cascade, external pole machines, synchronous current machines alternating current machines and/or direct current machines, electronic equipment and electrical resistances. The capture of electrons can provide single-phase, two-phase or three-phase supplies, operating at low, medium or high voltage.

The capture of electrons by induction, does not impact on the environment. The fact is that we use as the capturing force, only a negligible amount of electricity relative to the current captured by the sensor. The relationship between power input and the quantity of electricity generated by the electron-trap is at least 1 to 100, that is, for each 1 watt provided to the sensor, there is at least 100 watts of power available for external loads. This relationship, however, is not limited, as it depends on the mounting of the electron-trap and the objectives of the circuit, and so, the generated power can be greater than 100 times the input power.

Another advantage of the earthed electron- trap proposed in the present invention is that the electron-trap can transport electrons from point "A" to point "B" without a voltage drop across the closed-loop conductive element - if it is biased with a voltage - regardless of the distance between the points depending on the strength and quantity of the electromagnetic field generating devices. It is also possible to transport electrons when the conductive element in a closed circuit is itself not polarised. Thus, the electric current is transported without voltage, just by the magnetic field formed between the device and the generator of the electromagnetic field.

Brief description of the Drawings

The present invention will now be described with the aid of drawings, but the design is not limited to the implementations shown in these drawings, although they show other details and advantages of the present invention.

The figures show:

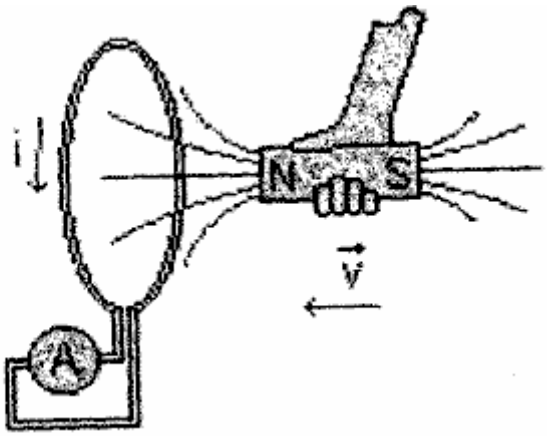


Fig.1

Fig.1 - illustrates Faraday's law.

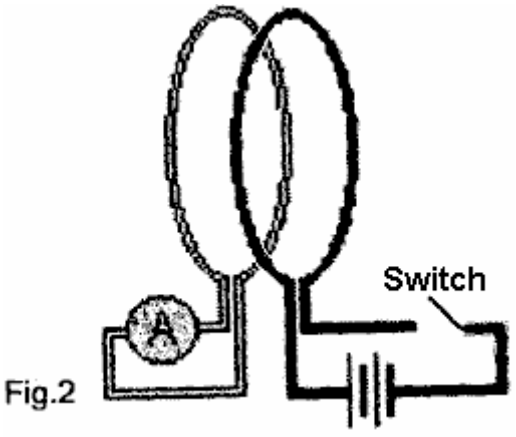


Fig.2

Fig.2 – is a representation of Faraday's law.

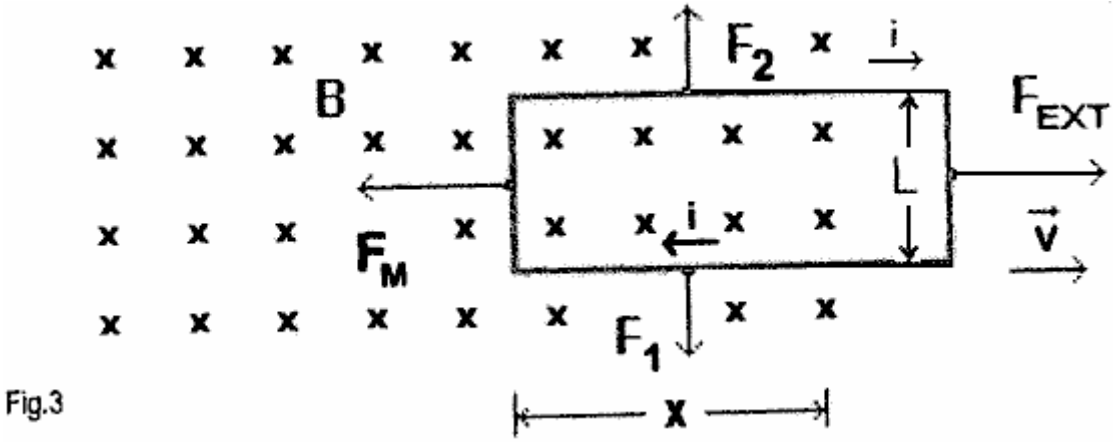


Fig.3

Fig. 3 – is a representation of Faraday's law.

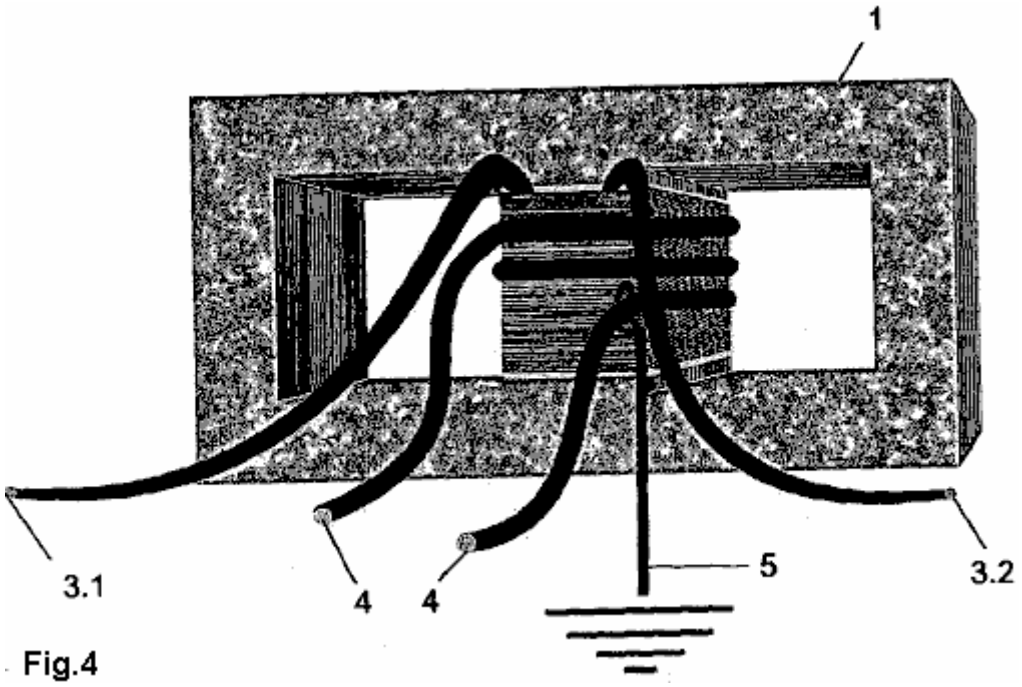


Fig.4

Fig. 4 - is a perspective view of an electron-trap with a single phase coil.

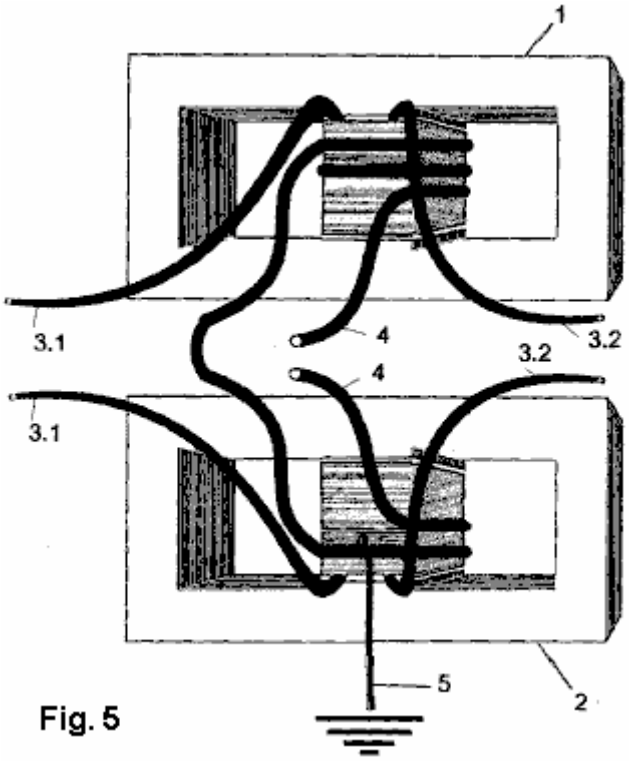


Fig. 5

Fig.5 – is a perspective view of a single-phase electron trap with two coils.

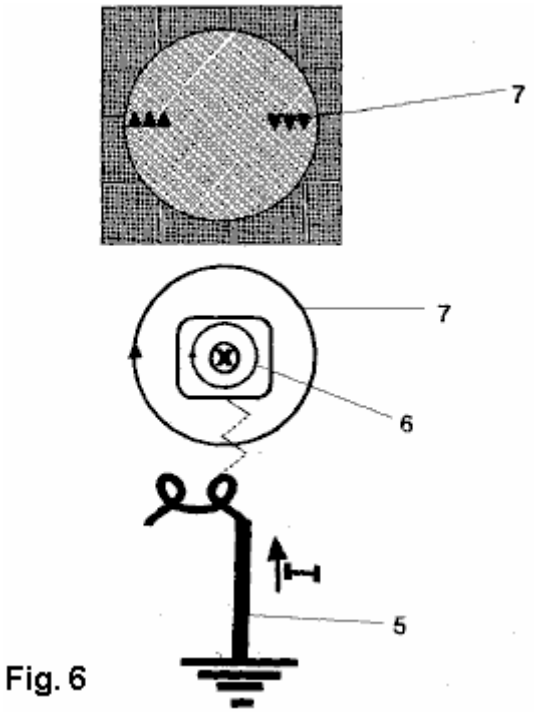


Fig. 6

Fig.6 - is a representation of the effect of electromagnetic flux in the coils around the cores of the electron trap.

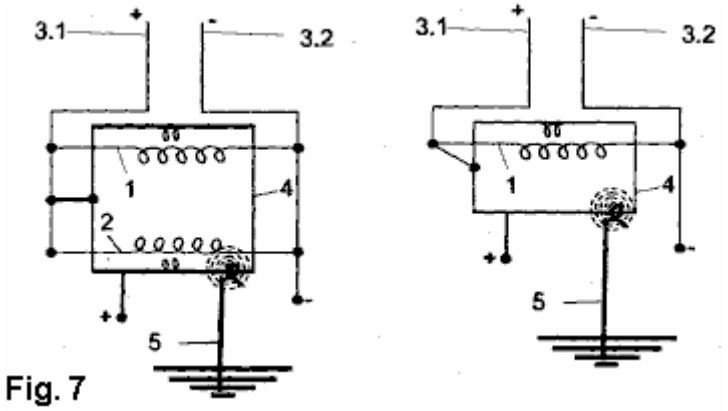


Fig. 7

Fig.7 - is a representation of an electrical circuit with two coils of the link/coil conductor polarised.

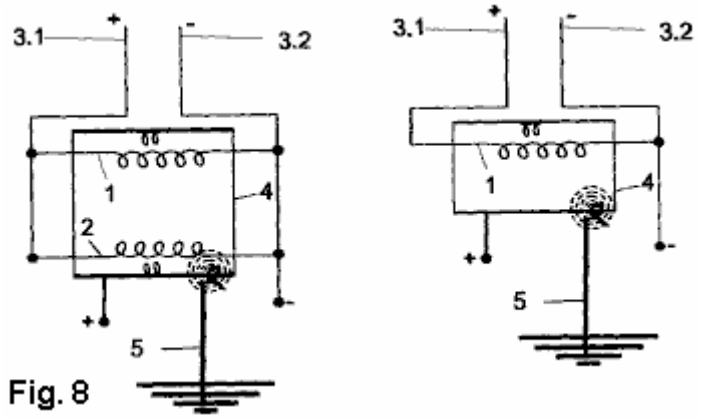


Fig. 8

Fig.8 - is a representation of an electrical circuit with two coils of the link/coil conductor not polarised.

Detailed Description of the Drawings

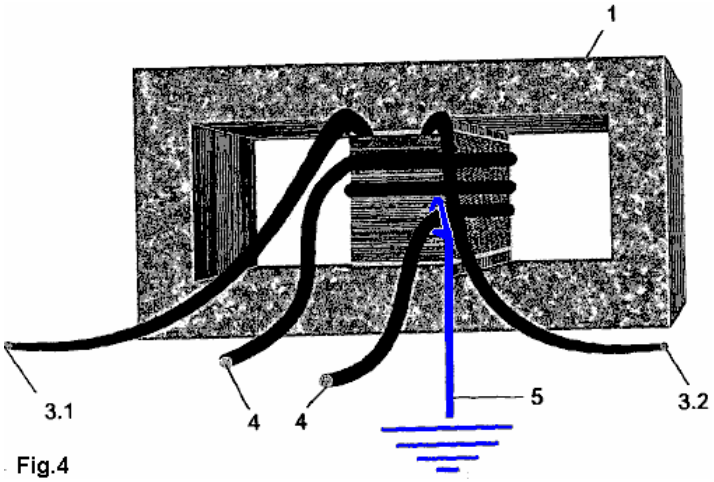


Fig.4

Fig.4 shows one of several types of electron-trap proposed by the present invention, where the electron-trap is single-phase and consists of at least one electromagnetic field-generating device with at least one set of coils, in this case it happens to be an electromagnetic type coil with one common magnetic core, but it could alternatively have any number of windings of any kind and shape. However, the electron-trap proposed by the present invention can be constructed with a different type of electromagnetic field generating device, such as an electromagnetic inductor or magnet of any type or shape, or any combination of them, and in unlimited numbers for each phase of the electron trap.

When winding these coils, for example, coil **4-4**, each coil must have at least one complete turn, preferably two turns if the objective is to generate electricity, and preferably four turns if the objective is provide thermal energy. The number of turns in the coils wound around the common core, is directly related to the amount of current to be generated.

At least one conductive interconnection element, in this case the driving member **5** - which can be copper or any other suitable conductive, material whether insulated or not insulated, connects or loop-links wire **4** to the ground grid. The connection between the conductor **5** and wire **4** is by electromagnetic induction. Winding **4** is also the power supply for the loads which are to be powered by the captured electrons.

Also in **Fig.4**, the power wires **3.1** and **3.2** (live phase and neutral) have an input from an external power coil **1** which can be energised from any external source of electricity such as a power grid. The trapped electrons can be configured to supply DC or AC current. Thus, if the coil **1** power source is alternating electrical current - AC, then the electron-trap provides alternating electrical current. If the power source is continuous electrical current - DC, then the electron-trap provides continuous electrical current - DC. The electrical supply provided by the trapped electrons can be single-phase, two-phase or three-phase, and at low, medium or high voltage.

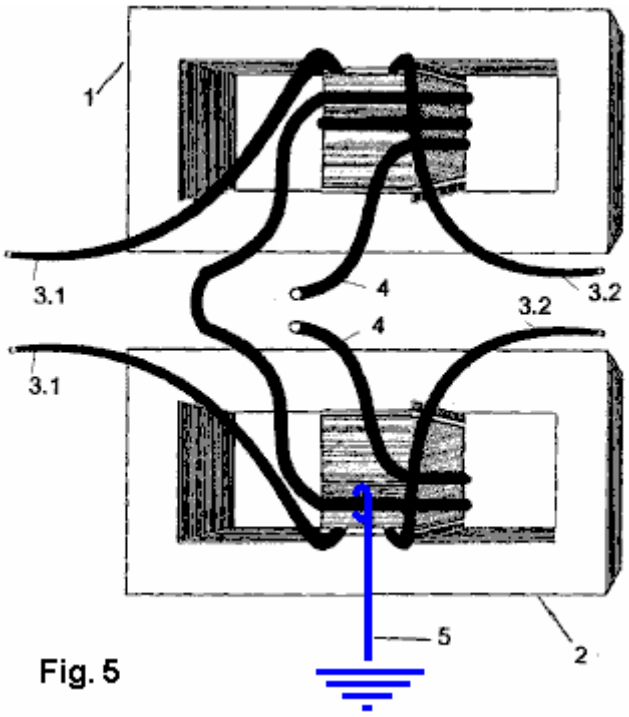


Fig. 5

Fig.5 shows an electron-trap with two single-core phase coils: **1** and **2**, although these coils may be of any type and shape. However, the electron-trap proposed by the present invention can be constructed with other types of electromagnetic field generating device, with at least one electromagnetic inductor or electromagnet which can be of any type and shape, with any combination of them, and in unlimited quantities in each phase of the electron-trap.

The coils on frames **1** and **2** may have other shapes, but they must each have at least one complete turn, particularly in coil **4**. The number of turns in this winding are directly related to the amount of current to be generated. This coil also makes the interconnection between the coils **1** and **2** forming the link between their two cores.

At least one conductive interconnection element, in this case the driving member **5** - which can be copper or any other suitable conductive, material whether insulated or not insulated, connects or loop-links wire **4** to the ground grid. The connection between the conductor **5** and wire **4** is by electromagnetic induction.

In electron-traps which have numerous sets of coils **1** and **2**, the ends of all of the power-supply conductors **3.1** can all be connected to each other, and all of the **3.2** conductor ends may be connected together. Thus, all of the coils **1** and **2** can be fed exactly the same voltage. The power to energise coils **1** and **2** can be provided from any external source of supply of electricity such as a power grid.

In electron-traps which have numerous coils **1** and **2**, a single coil winding **4** connects the cores of all of the coils **1** and **2**.

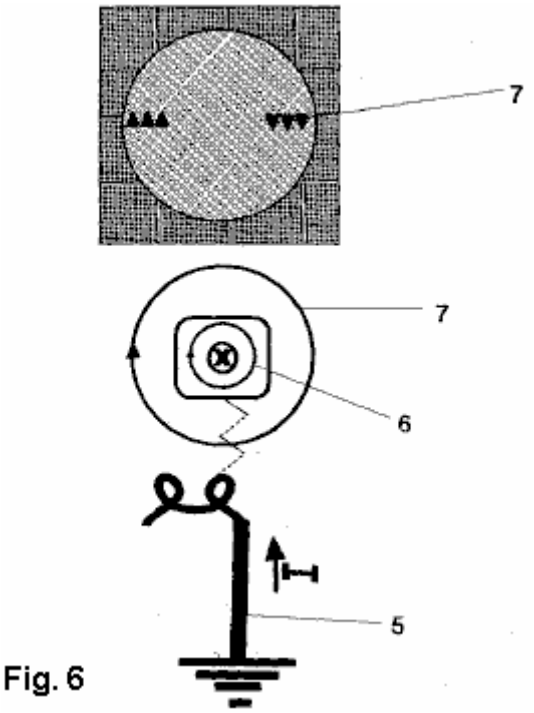


Fig. 6

The diagram shown in **Fig.6**, illustrates the magnetic induction **6** around the core "**X**" of the coil **1**. This induction causes electrical current flow in the conductor coil link **7/4**, attracting electrons from the earth, through the conductive member **5**, to the magnetic field of the electron-trap, where those electrons are added to the current generated by induction in the link coil **4** conductor loop circulating between north and south magnetic poles.

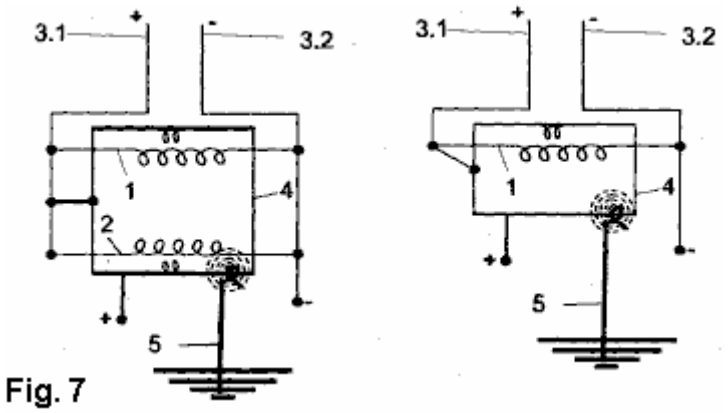


Fig. 7

Fig.7 shows how the connections should be made in one version of the electrical circuit of the electron-trap proposed in this invention. The diagram shows the electrical circuit of an electron-trap where the link/coil driver **4** is polarised with a voltage. This is one form of construction for an electron-trap which has two coils **1** and **2**, where a link/coil loop conductor **4** is biased with a voltage, that is, there is a link connecting the coil conductors **4** of a power supply **3.1** or **3.2**, whatever the stage.

In this way, earth electron-traps, by adopting this circuit, that is, with the link/conductor loop **4** and polarised voltage on coils **1** and **2**, besides being used as a power source for external loads, can also be used for thermal power generation.

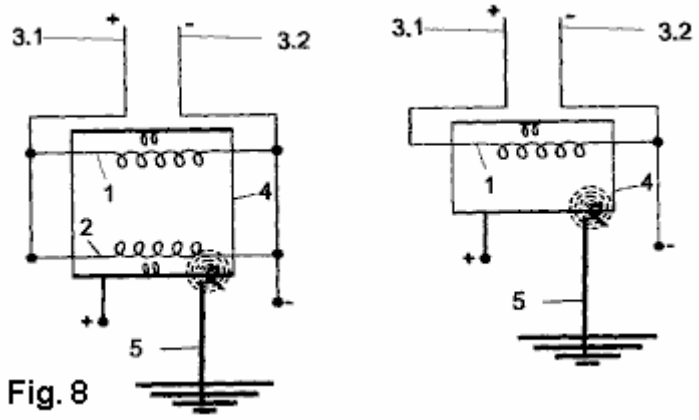


Fig. 8

Fig.8 shows how connections should be made in another electric circuit electron-trap proposed in this invention. the circuit illustrates a circuit of an electron-trap with a non-polarised link / coil driver **4**. This is one form of construction of the electron-trap where a link / coil conductor **4** of the spiral conductor coils **1** and **2** is not polarised, that is, there is no such link connecting conductor / conductor coil conductors **4** of a coil **3.1** or **3.2**.

Thus, earth electron-traps adopting this circuit, that is, with the link coil not polarised, the current flows without there being voltage in the link/coil conductor **4** joining the first and second coils by electromagnetic induction. They can also be used for generating thermal energy.

The structure of the circuit - in the open or closed coils **1** and **2**, and always in the closed link / loop lead **4** - makes it possible to generate current by induction and electron capture by electromagnetism on the link conductor **4** - where current is generated and stays in motion with or without voltage, as the coils **1** and **2** are being fed. Thus, the present invention provides a new concept for electrical energy generation, since it is obtained from an electric current circling without consumption and even without an output load being attached to it.

Additionally, because the induced electrical current flows regardless of the voltage present, it can be used as a current stabiliser for electrical networks whether they be single-phase, two-phase or three-phase, with low, medium or high voltage.

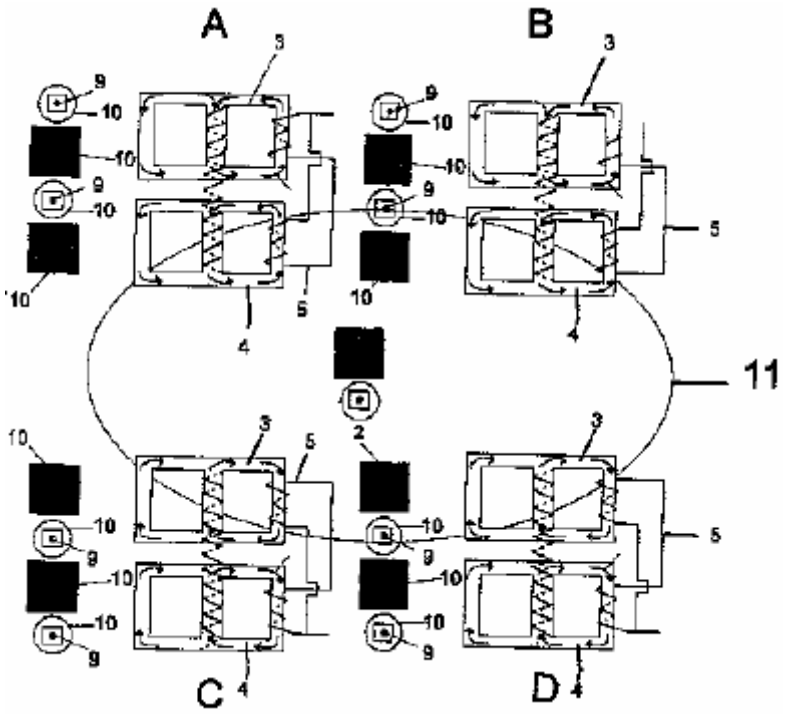
ELECTROMAGNETIC DEVICE FOR CAPTURING FREE SPACE ELECTRONS TO GENERATE ELECTRICITY

Application Number: BR2013/000014

Publication Date: 07/18/2013

Filing Date: 01/11/2013

Assignee: EVOLUÇÕES ENERGIA LTDA (Rua Santa Tereza 1427-B Centro - Imperatriz -, MA - CEP -470 - Maranhão, 65900, BR)



Abstract:

The invention relates to a device that comprises at least three sets (**A, B, C, D**) of at least one device for generating an electromagnetic field (**3**) and (**4**), powered by an electricity source (without a core or with at least one core) the cores thereof or any extension thereof, preferably the windings or sets of windings thereof, being surrounded by at least a single conductive element forming a polarised and energised closed-circuit (**5**), the sets of electromagnetic-field generating devices (**3**) and (**4**) being linked together by their opposing poles to encourage the interaction of their electromagnetic fields, which ideally, are located between two hollow metal hemispheres (**1**) so as to concentrate and enhance the electromagnetic fields, these interconnections causing, as a novel technical effect, the emergence of an electrical current that circulates, with or without voltage, in the conductive element forming a closed-circuit (**5**) - even if no load is connected.

Description:

"ELECTROMAGNETIC EQUIPMENT FOR FREE ELECTRON-CAPTURE FROM SPACE, FOR ELECTRICITY GENERATION".

Technical Field

The present invention relates to electromagnetic equipment for electrical power generation and/or thermal power generation. More specifically, equipment capable of producing abundant electricity and thermal energy from a tiny input of electrical energy.

Description of the Related Art

According to Lenz's law, any induced current has a direction such that the magnetic field it generates opposes the change in magnetic flux that produced it. Mathematically, Lenz's Law is expressed by the negative sign (-) that appears in the formula of Faraday's Law, as follows.

The magnitude of the induced emf (ε) in a conducting loop is equal to the rate of change of magnetic flux (Φ_B) with time:

$$\varepsilon = -\frac{d\Phi_B}{dt} \qquad \text{Equation 1}$$

As an example of application of Faraday's Law, we can calculate the electromotive force induced in a rectangular loop that moves in or out, with constant speed, a region of uniform magnetic field. The magnetic field flux through the surface limited by the loop is given by:

$$\phi = xLB \qquad \text{Equation 2}$$

and its variation in time:

$$\frac{\Delta\phi}{\Delta t} = \left(\frac{\Delta x}{\Delta t}\right)LB = vLB \qquad \text{Equation 3}$$

So:

$$\varepsilon = vLB \qquad \text{Equation 4}$$

and if the coil has a resistance (R) and the induced current:

$$i = \frac{\varepsilon}{R} = \frac{vLB}{R} \qquad \text{Equation 5}$$

A conductor traversed by an electric current immersed in a magnetic field undergoes the action of a force given by:

$$F = IL \times B \qquad \text{Equation 6}$$

Thus, the effect of the current induced in the loop appears as forces Ff, and F - FM. The first two cancel each other out and the third is cancelled by an external force P_{EXT} needed to maintain the constant speed loop.

As the force FM must oppose the force F_{EXT}, current (i) induced in the loop by varying the magnetic flux must have the meaning indicated in **Fig.1**. This fact is a particular example of Lenz's Law.

Considering the experimental activities discussed with Faraday's law, when a magnet approaches a coil, the induced current in the coil has a direction as shown in **Fig.2**. This generates a magnetic field whose north pole is facing the north pole of the magnet, that is, the field generated by the induced current opposes the motion of the magnet.

When the magnet is moved away from the coil, the current induced in the coil has a direction opposite to that shown in **Fig.2**, thereby generating a magnetic field whose south pole is facing the north pole of the magnet. The two poles attract each other, that is, the field generated by the induced current opposes the movement of the magnet away from the coil. This behaviour is present in all current power generators, and known as 'engine brake' is highly undesirable as it increases the resistance and so, the energy loss.

When two electromagnetic coils are placed facing each other, there is no current in either of them. At the instant of power up one of the coils, the current in the coil, generates an induced current in the second coil. When powered up, the current in the coil goes from zero to its maximum value, and then remains constant.

Thus, when the current is changing, the magnetic field generated by it, (whose north pole faces the second coil) is also changing and so the magnetic flux of this field through the second coil is also changing. Then there is a current induced in the second coil whose sense is such that the magnetic field it generates tends to decrease the flow mentioned above, that is, its north pole confronts the north pole of the first field coil.

When the power switch is opened, the current in the first coil drops from its maximum value to zero, and correspondingly its magnetic field decreases. The flux of the magnetic field in the second coil also decreases,

and the induced current now flows in the opposite direction. This current flow direction produces an enhancing magnetic field, that is, it has a south pole facing the north pole of the field of the first coil.

Thus, there is a realisation of the principle of conservation of energy, expressed by Lenz's law, wherein any induced current has an effect which opposes the cause that produced it. Assuming that the induced current acts to favour the variation of the magnetic flux that produced the magnetic field of the coil, it would have a south pole facing the north pole of the approaching magnet, causing the magnet to be attracted towards the coil.

If the magnet were then released, it would experience an acceleration toward the coil, increasing the intensity of the induced current and thus create an enhanced magnetic field. This field, in turn, would attract the magnet with increasing force, and so on, with a continuing increase in the kinetic energy of the magnet.

If energy were to be withdrawn from the magnet-coil system at the same rate at which the kinetic energy of the magnet increases, then there would be an endless supply of energy. So it would be a perpetually operating motor, which would violate the principle of conservation of energy. Therefore, it can be concluded that current generators feature a large energy loss during the generation of electricity.

Objectives of the Invention

The present invention aims to contribute to the generation of sustainable energy, proposing electromagnetic equipment capable of producing abundant electricity from an extremely low input of electrical energy.

The above objective and other objectives are achieved in the present invention by a device comprising at least three sets of at least one electromagnetic field generating device (without a core or with at least one core) powered by an electrical power source, having their cores or any extension of them, with their coils or sets of coils, wound on at least one common conductive member in a closed circuit which is polarised by a voltage source, and these sets of electromagnetic field generating devices are arranged with their poles in confrontation, to promote the interaction of electromagnetic fields, and, preferably, positioned between two hollow metallic hemispheres, in order to focus and enhance their electromagnetic fields - these interactions cause a new technical effect - the emergence of an electric current which keeps flowing in a closed loop, with or without voltage being applied to that closed loop, current which is capable of powering external loads - even if no load is attached to it.

The device which is the object of the present invention operates as follows: Sets of electromagnetic field generating devices to be powered by an electrical power source, produce an electromagnetic field which induces an electric current in a closed conductive circuit, creating an interaction between the magnetic poles, and through repeated electromagnetic attraction and repulsion, provides an endless supply of electrons to the conductive closed loop itself.

The electrons attracted by this technique, augment the current flowing in the closed conductive loop, which provides the current to power external loads of high power, in spite of the fact that the device itself is supplied with only a small level of power. Thus, advantageously, the device which is disclosed in the present invention forms a trap for electrons from space, resulting in the generation of electricity. The interconnections of the components of the electron-trap cause, a new technical effect, namely, the appearance of an electric current which keeps circling in a closed circuit, even without any voltage being applied to the closed circuit, and even without a load being connected to it. The present electromagnetic equipment generates electricity or thermal energy, providing access to this new source of energy through the use of an electromagnetic field.

The proposed sensor can also be used for the generation of thermal energy depending on the form of circuit which is to be used, resulting from the flow of electric current produced by this electromagnetic equipment.

This field generates a flow of electric current induced by electromagnetic coils, which appears in the linking interconnecting devices generating electromagnetic fields with electromagnets, inductors or magnets. This chain operates in a manner favourable to the variation of the magnetic flux produced by the magnetic field in the electron-trap. Thus, it creates a north pole and a south pole, providing an endless supply of electric current without resistance between the links which interconnect the devices which are generating the electromagnetic fields. So, induced electric current is generated with or without voltage in the interconnection links of electromagnetic field-generating devices, depending on the connection method of the electrical circuit of the electron-trap.

The free-electrons collected by the space electron-trap can form alternating current (AC) or direct current (DC). The ratio of input power to output power is 1 to 100, that is, the generated power can be 100 times greater than the input power when there is at least one link / coil driver between the coils and the inductors or electromagnets. This relationship, however, is not limited to a factor of 100, as it depends on the shape of the electron-trap and its objective.

Another advantage of the free space electron-trap of the present invention is that, with thermal insulation of the components in the electric circuit, it is possible to produce thermal energy at low, medium or high temperature, through the movement of the electrons in the conductors, coils and/or electromagnets. The temperature generated is linked directly to the number of turns in the coils.

Thermal power generation performed by the sensor can be used for boiling and/or evaporation of liquids to be used in other types of energy generation, for example, replacing the use of coal and natural gas.

Another advantage of the proposed electron-trap of the present invention is that the electron-trap can transport electrons from one point "A" to a point "B", without a voltage drop in the link - if it is polarised - regardless of the distance between the points, depending on the strength and quantity of the electromagnetic field-generating devices. It is also possible to transport the electrons when the link devices generating the electromagnetic field are not polarised. In this way, the electric current is conveyed without voltage but only by the magnetic field formed between the coils. This methodology can be used in various fields.

Because of its simple construction, the electron-trap is a simple device which is compact, and performs low-cost power generation which can be used in all types of machinery, equipment and devices of all kinds, and many areas of application which require electricity in order to operate. The electron-trap can have single-phase, two-phase or three-phase output, and can generate electric current at low, medium or high voltage.

Brief description of the Drawings

The present invention will now be described with the aid of drawings, but the design is not limited to the implementations shown in these drawings, although they show other details and advantages of the present invention.

The figures show:

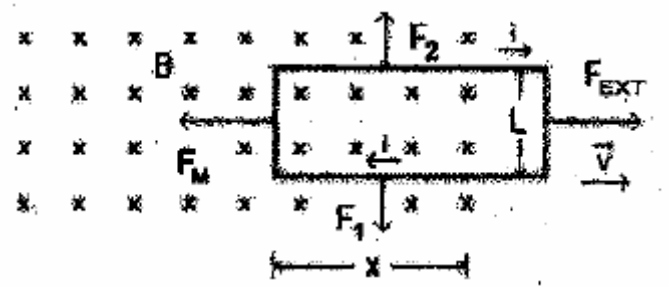


Fig.1 - illustrates Faraday's law.

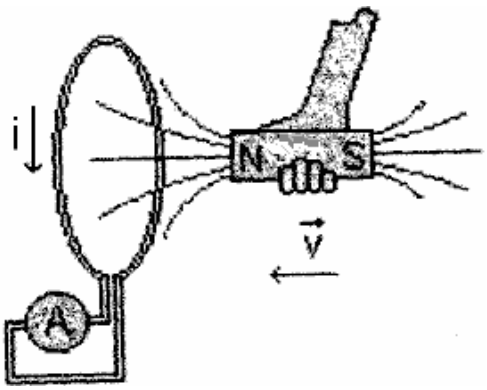


Fig.2 illustrates Faraday's law where a magnet approaches a coil of just one turn..

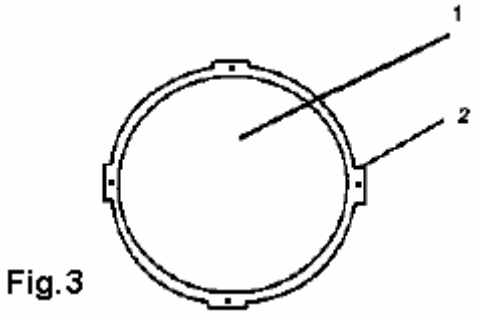


Fig.3

Fig.3 is a view of one metallic hemisphere seen from above.

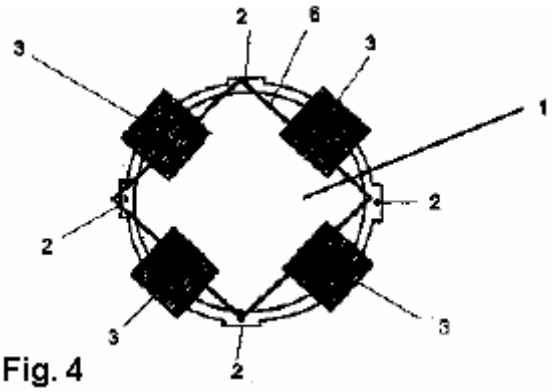


Fig. 4

Fig.4 is a bottom view of the hemisphere with the coils in place.

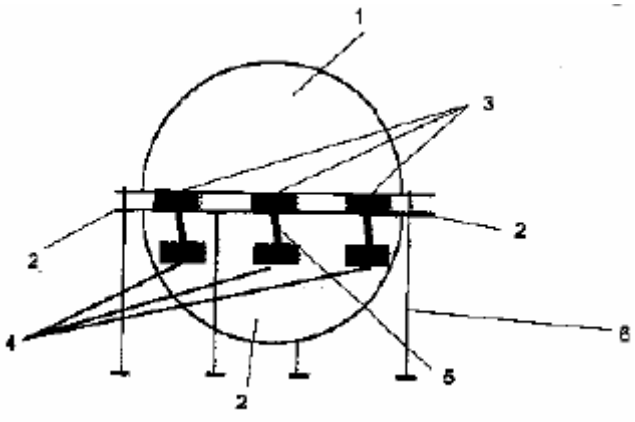


Fig.5 is a side view of the free-space electron-trap.

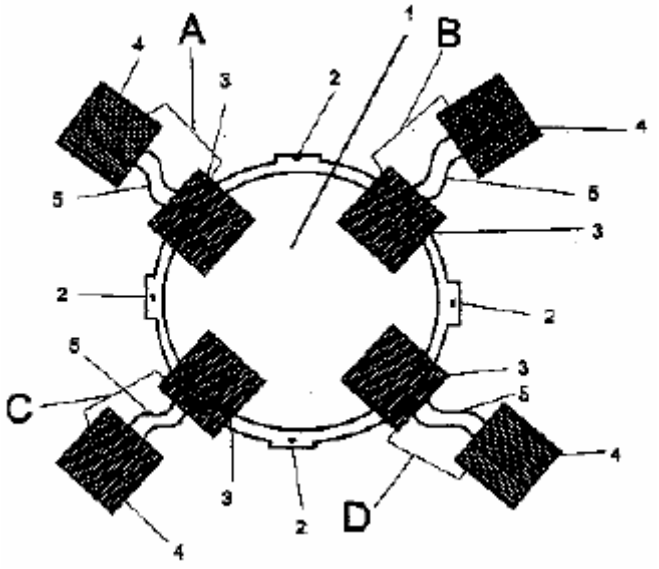


Fig.6 is an underside view of the space electron-trap, with its coils and electromagnets

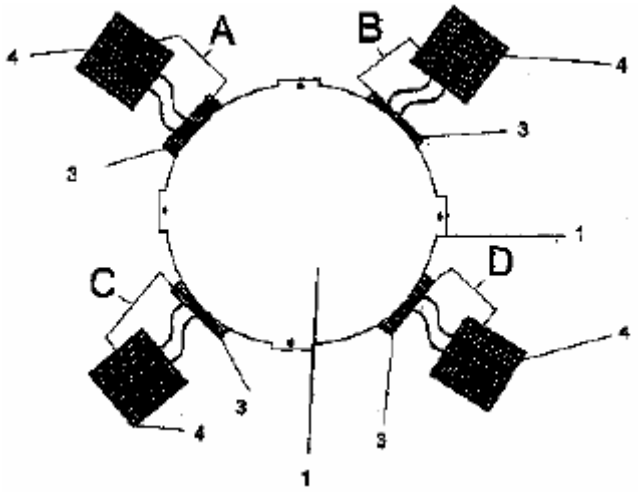


Fig.7 a view from above of the space electron-trap with its coils and electromagnets.

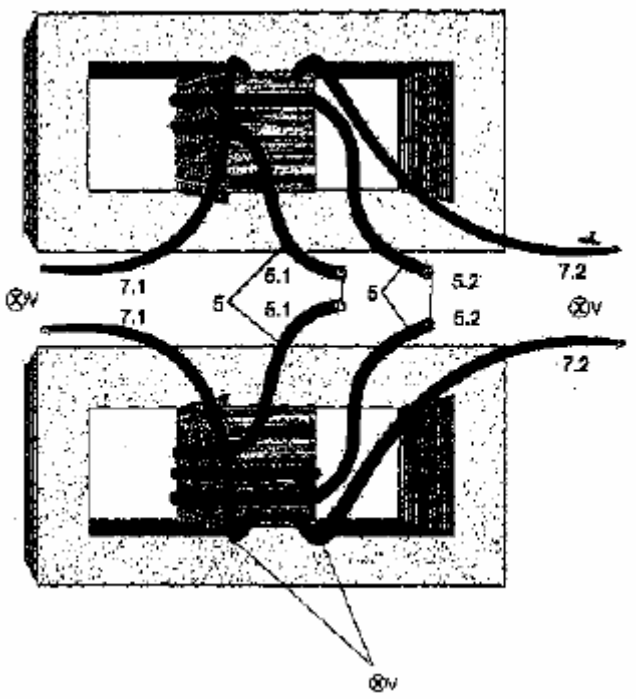


Fig.8 is a perspective view of an electron-trap with its coils.

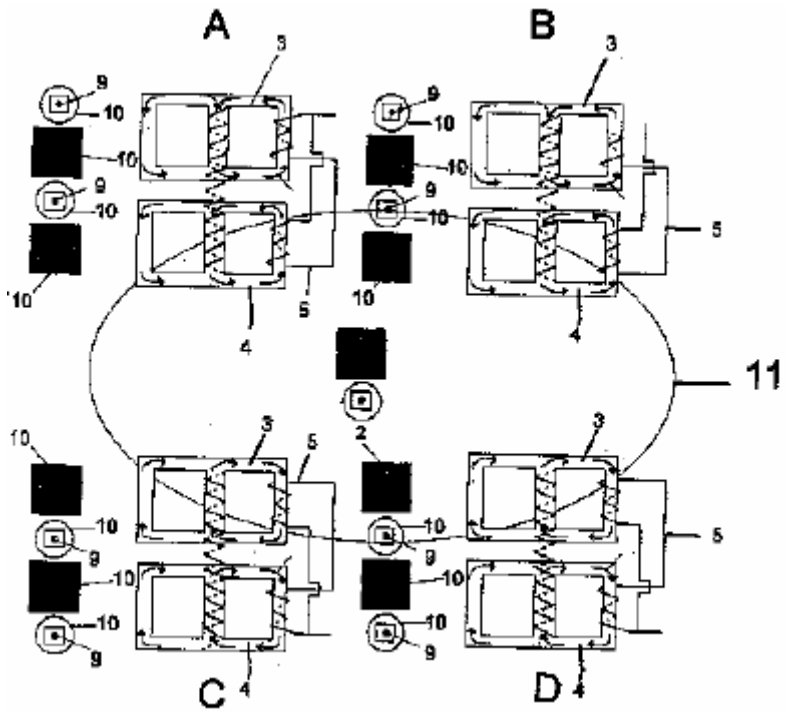


Fig.9 shows the circuit diagram of the device, indicating the effect of electromagnetic field.

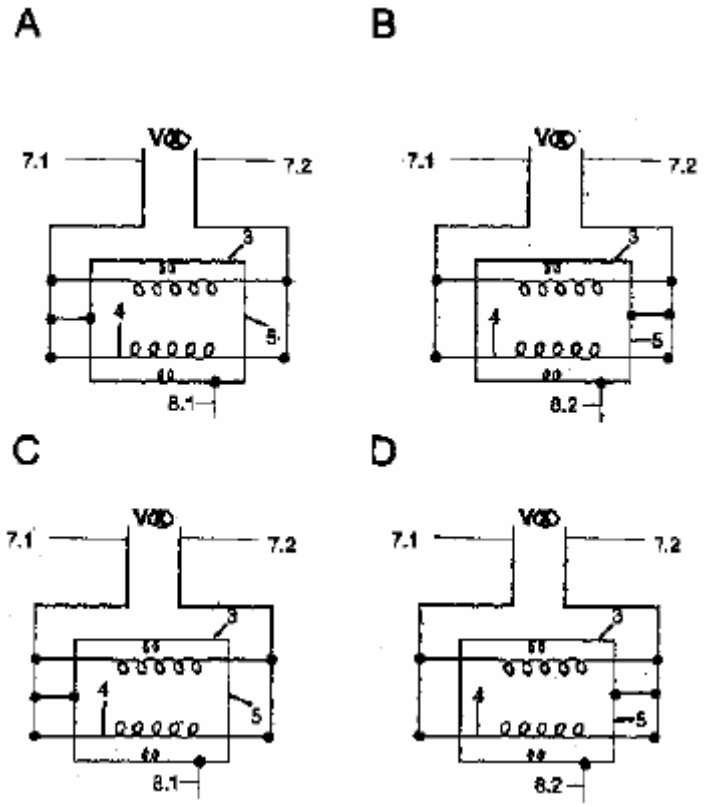


Fig.10 - shows the circuit diagram of the connection of the inductor coils in sets (**A, B, C** and **D**).

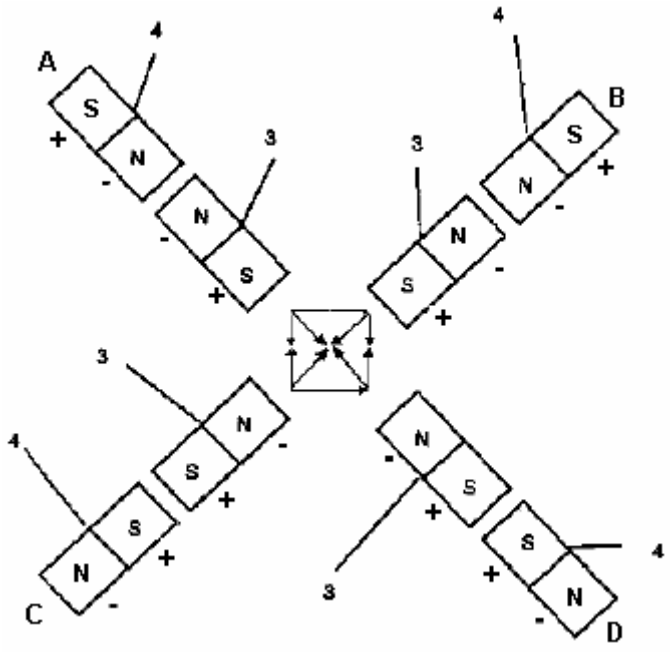


Fig.11 - is an electromagnetic diagram representation of north and south poles of the sets of coils (**A, B, C** and **D**).

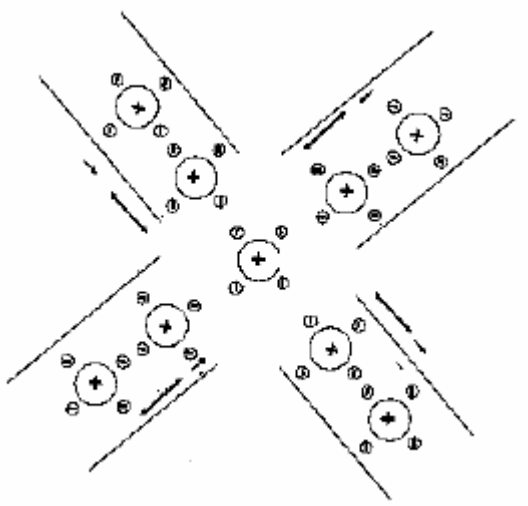


Fig.12 is a representation of the electrons being attracted and repelled by the device.

Detailed Description of the Drawings

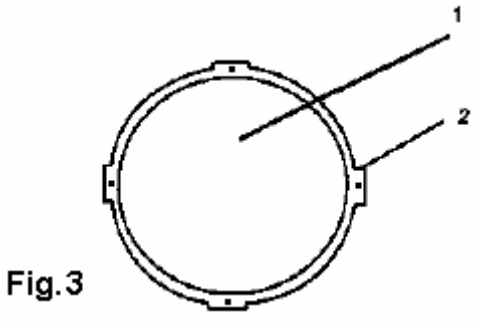


Fig.3 is a top view of one of the two hollow metallic hemispheres **1** which is part of the electron trap of free space proposed in this invention. Hemisphere **1** is preferably made from, but not limited to, aluminium, and it has mounting tabs **2**.

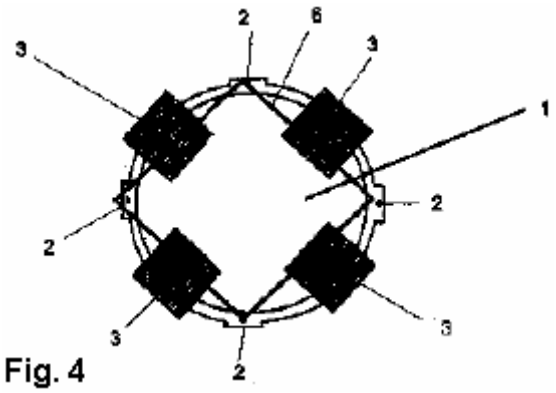


Fig.4 is a bottom view of metallic hemisphere **1**. It has four electromagnetic field generating devices **3**, positioned around the hemisphere and fixed to support **6** which is attached to hemisphere **1** by mounting tabs **2**.

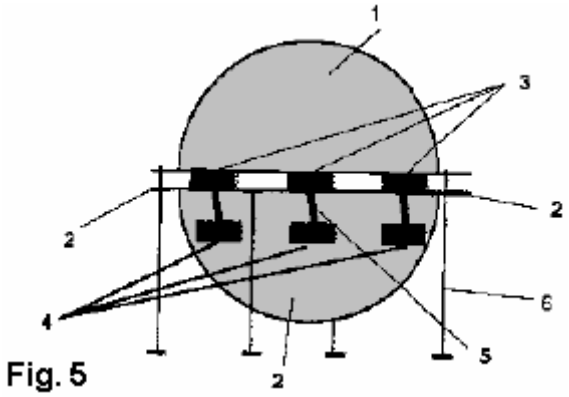


Fig. 5

Fig.5 is a side view of the free space electron-trap. It shows the two metallic hemispheres **1** and **2** (which form an imperfect sphere), and three of the coils **3** which are attached to the mounting tabs **2** and three inductors **4** which form the closed circuit itself, and which are attached by conductors **5**, and support member **6** on which are mounted coils **3** and their components.

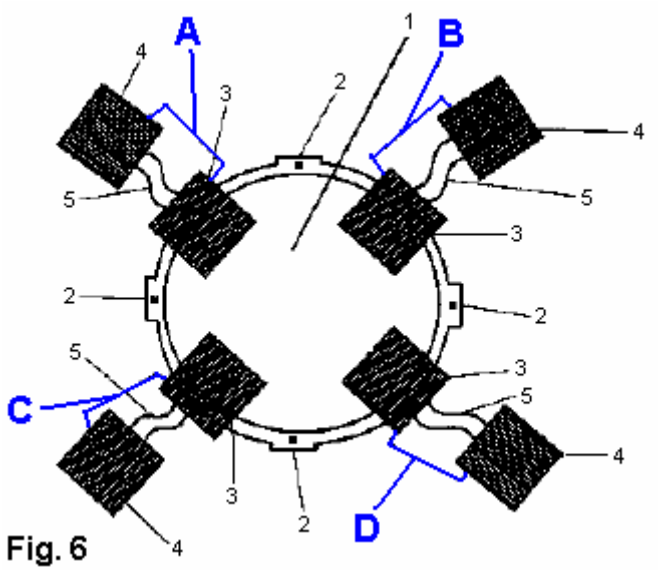


Fig. 6

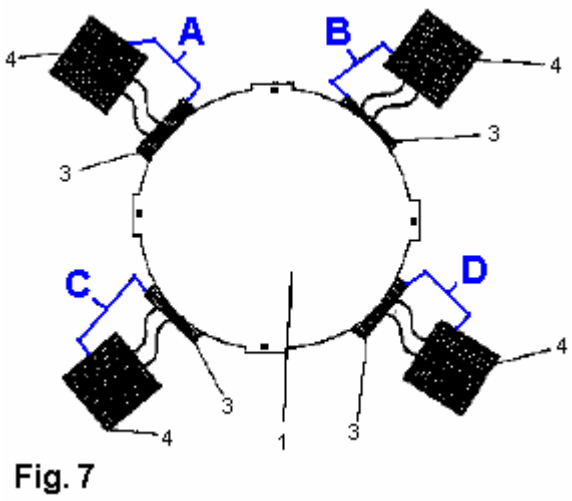


Fig. 7

Fig.6 and **Fig.7** show the top and bottom views of the metallic hemisphere **1** which accommodates four coils **3** attached to the holder **6** (not shown) which is secured to the hemisphere **1** by its mounting tabs **2**. **Fig.6** also shows the inductors or electromagnets **4** their corresponding coils **3** and their interconnecting conductors **5**. Each coil **3** and its linked inductor **4** forms a set. In **Figures 6** and **7** there are four such sets, marked **A**, **B**, **C** and **D**. The coils **3**, connected by their links **5**, each have at least one turn, and if the objective is to generate electricity, then preferably two turns, and if the objective is thermal energy, then four turns. The coils **3** may have various different shapes. The number of turns in the coil **3** are directly related to the amount of current to be generated, and the connecting links **5** may be either a single conductor or more than one conductor, the cross-sectional area of conductor **5** being selected to carry the current which is to be generated.

In sets **A, B, C** and **D**, the link conductors **5** have at least one turn around coils **3**. This winding is connected to the respective electromagnets **4** of each set (**A, B, C** and **D**) as shown in **Figures 6** and **7**. Please note that the inductors and electromagnets **4** can be any type of inductor, and other types of coil may be used.

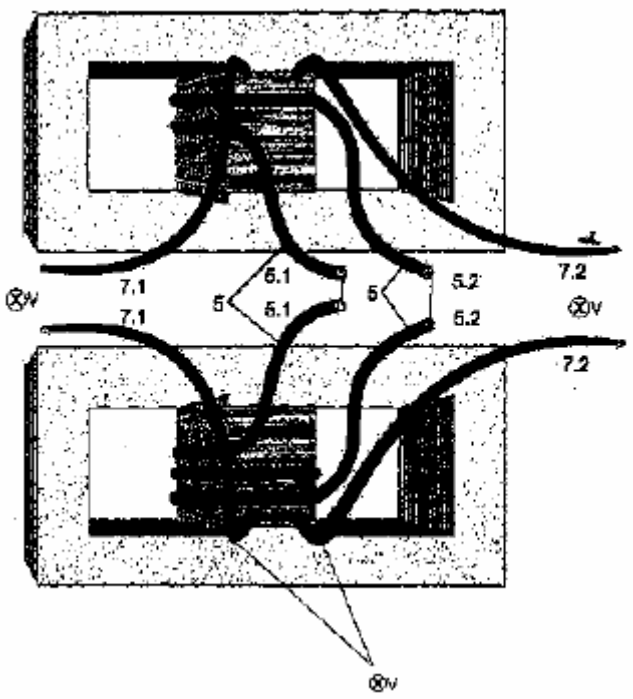


Fig.8 shows the inter-connecting coils **5** for each of the five sets **A, B, C** or **D** linking between coils **3** and **4** in each set. As shown in **Fig.6** and **Fig.7**, the link **5** makes the connection between coils **3** and **4**. This means that the wires marked **5.1** are all connected together, and the wires marked **5.2** are all connected together. Doing this, establishes the interconnection links **5** shown in the drawings. The power supply wires marked **7.1** are connected together as are the wires marked **7.2**. The wires marked **7.1** are connected to the live phase of the external power supply, while the other ends marked **7.2** are connected to the neutral of the external power supply.

In the space free-electron trap of the present invention, the coils **3** can be either single-phase, two-phase or three-phase. Also, the coils **3** may be powered by any voltage (V). The power coil **3** can be energised by any source of electrical energy such as a power grid. The electron-trap can be configured to produce alternating current or direct current. So, if the external power supply is alternating electrical current - AC, then the electron-trap provides an alternating electrical current output. If the power supply is DC, then the electron-trap provides an output of continuous electrical current - DC. The electron-trap can be configured for single-phase, two-phase or three-phase operation, with low, medium or high voltage outputs.

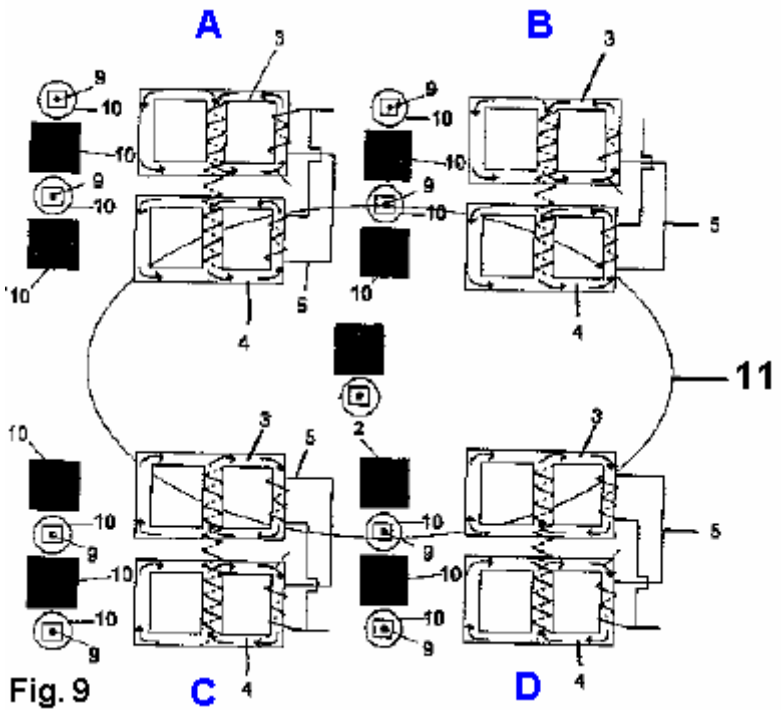


Fig.9 shows an electron-trap circuit diagram with four sets **A, B, C** and **D** of inductor coils **3** and **4**. Induction is produced around core **9** of the three sets of coils **A, B, C** and **D**. The effect of the interaction of the

electromagnetic fields **11** is shown. The induction via core **9**, causes the circulation of electric current in the links **5**, attracting the free electrons through the electromagnetic field of the trap. Then, the electrons join with the current generated by induction on link **5**, circulating between the magnetic poles north-south and south-north.

By way of example, the coils **3** are shown wound on a single phase column type core, but these can also be of any kind or shape. The electron-trap proposed by the present invention can be constructed with another type of electromagnetic field generating device which has at least one electromagnetic coil or magnet or electromagnetic inductor which can be of any kind or shape, or any combination of those, and with any number in each phase of the electron-trap.

The electron capture occurs through an electromagnetic field which is formed with the connection of coils **3** with the electromagnets or inductors **4** through the links **5** between the eight components.

This closure produces the displacement of the electrons in the coil **3** set (**A**) (for simplicity, referred to as coil **3A**), these electrons are attracted by the protons of coil **3D**, and are repelled by the electrons of the electromagnetic field of the coil **3D** itself. These coil **3D** electrons are attracted by the protons of the coil **3B**, and are repelled by the electrons of the electromagnetic field of coil **3B**. These electrons of coil **3B** are attracted by the protons of coil **3C**, and are repelled by the electrons of the electromagnetic field of the coil **3C** itself. Similarly, the **3C** coil electrons are attracted by protons of the **3A** coil, and are repelled by the electrons of the electromagnetic field of the coil **3A** itself. These coil **3A** electrons are attracted by the protons of the **3D** coil, and are repelled by the electrons of the electromagnetic field of the **3D** coil itself. Analogously, the coil **3D** electrons are attracted by the protons of the coil **3B**, and are repelled by the electrons of the electromagnetic field of the coil **3B** itself. These **3B** coil electrons are attracted by the protons of coil **3C**, and are repelled by the electrons of the coil itself induced **3C**, and then the coil **3C** electrons are attracted by protons of coil **3A**, and are repelled by the electrons of the electromagnetic field of the coil **3A** itself. That cycle continues as the sets of coils **A, B, C** and **D** are being fed by a voltage. These endless attractions and repulsions generate an electric current in the link coil **5**.

In the electron-trap, the voltage is stable. Regardless of the amount of current generated-which can be very high, the voltage will be the same in the electric circuit of the sensor, because the current moves through the attraction and repulsion of the electrons, regardless of voltage.

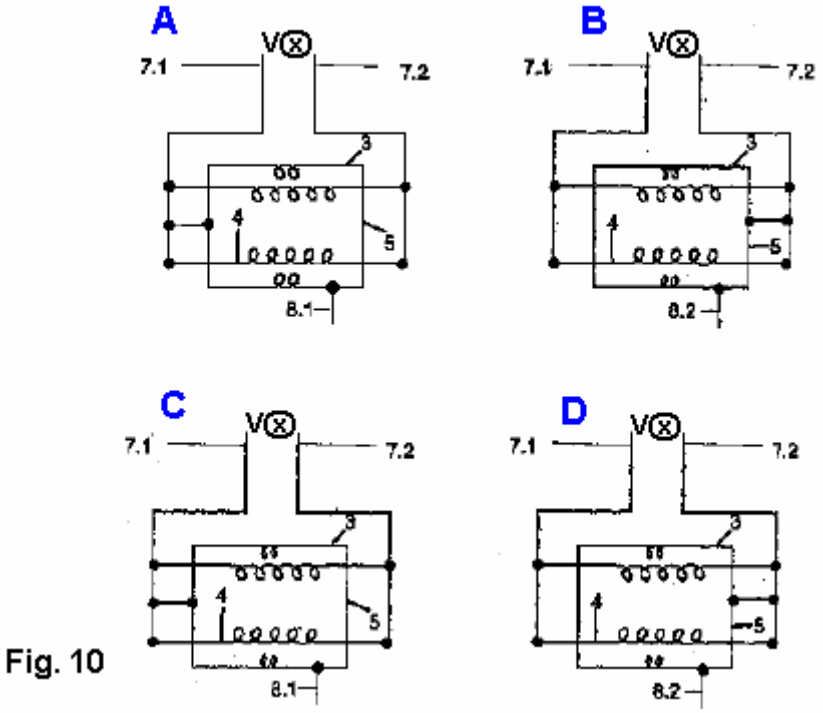


Fig. 10

Fig.10 illustrates a circuit diagram of the electrical connection between the coils **3** and **4** in sets **A, B, C** and **D**. It can be seen that the sets **A, B, C** and **D** are enclosed between the coils **3** and their associated inductors or electromagnets **4**. The supply conductors **7.1** and **7.2**, of sets **A, B, C** and **D** must be interconnected. When feeding power to the coils **3** and **4** the phase should be connected to **7.1** and the neutral to **7.2**.

The sets **A, B, C** and **D** after being fed with electric current, generate voltage through the attraction and repulsion of the electrons in the linking coil **5**, where there is at least one output load **8.1**, which should be connected joining sets **A** and **C**, and at least one load output **8.2**, which should be connected joining sets **B** and **D**. The output points **8.1** and **8.2** are the respective phases and neutral of power points **7.1** and **7.2**.

In this way, a single-phase electron-trap is created by two pairs of sets of coils/inductors **3** and **4**.

The **3/4** electromagnet coil set can be replaced by a **3/3** coil set, without any disadvantage to the electron-trap.

Sets **A, B, C** and **D**, are inserted into a hollow metal hemisphere **1** preferably constructed from - but not limited to - aluminium. The hemisphere **1**, whose function is to concentrate and maximise their electromagnetic fields, simulating an electron cloud, has a fixed support **6** connected to attachment tabs **2**, and to which the coils **3** are fixed.

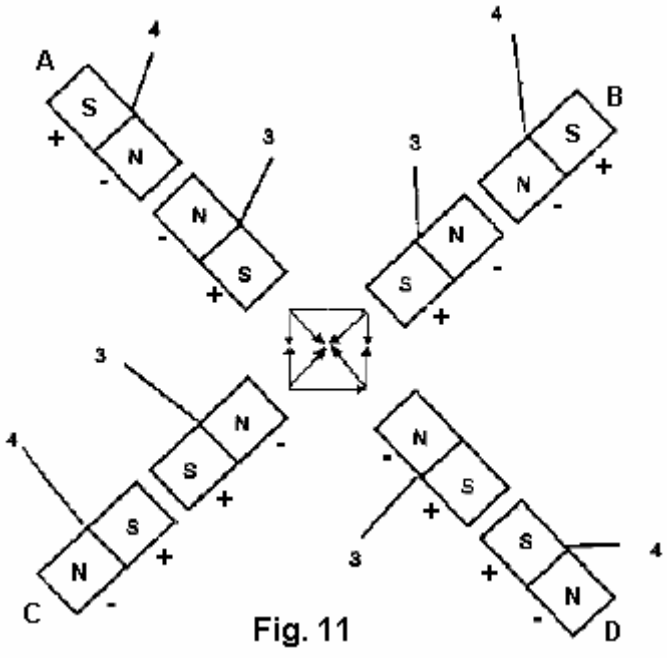


Fig. 11

Fig.11 is a diagram of the electromagnetic north and south poles of the inductor coils **3** and **4** of sets **A, B, C** and **D** of the electron-trap. The electromagnetic behaviour described for **Fig.9** is again demonstrated by the formation of the magnet assembly to the North Pole and South Pole being attracted and repelled by the lines of force of the magnet from the point "**A**" to point "**D**", point "**A**" to point "**B**", the point "**B**" to point "**C**", point "**C**" to point "**A**", and so on, as long as there is an electromagnetic field. The electromagnetic field of the space electron-trap provides that induced current in a direction similar to the variations of the magnetic flux that produced it. So, the magnetic field creates a north pole and a south pole in each of the sets **A, B, C** and **D**, as shown in **Fig.11**.

By feeding the coils **3** of the electron-trap with a desired voltage a magnetic field is generated in coils **3**, between the four sets **A, B, C** and **D**, which form a flow of electrons. This flow of electrons augments the electron flow which is circulating in the closed-loop link-coil **5**, thus implementing free electron capture from space. The electromagnetic field of the coil **3A** runs north to south, the electromagnetic field of the coil **3B** runs north to south, the electromagnetic field of the coil **3C** flows from south to north, and the electromagnetic field of the coil **3D** flows from south to north, as shown in **Fig.11**. It should be noted that the sets **A, B, C** and **D** can be formed by any combination of coil, magnet and electromagnet.

The south to north electromagnetic field induces current flow in the coil **3A**. The north to south electromagnetic field induces current flow in the coil **3B**. The north to south electromagnetic field induces current flow in the coil **3C** and The north to south electromagnetic field induces current flow in the coil **3D**. The induced current flow can have any power and it can be single-phase, two-phase or three-phase current.

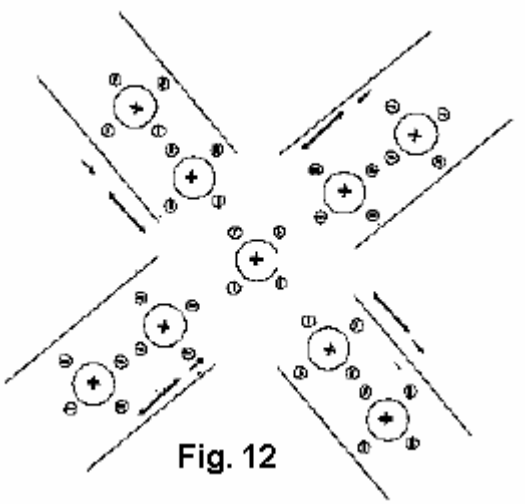


Fig. 12

Fig.12 shows the electrons being attracted and repelled by the induction coils **3** and **4**. Being repelled and attracted by electromagnetic induction, the electric current flows without resistance.

The electron-trap produces electromagnetic waves which can be used for various purposes, including signal transmission at any frequency and for any purpose. The capture is caused by these electromagnetic waves. The same physical effect can be achieved by the combination of the capture devices of other technologies, including electromechanical, electric, electronic, electromagnetic, or through the combination of a magnet or any other magnetised materials.

The space free electron-trap of the present invention is a renewable source of electrical power production and a new way of generating energy through the capture effect, generating flows of electrons, generating ordered movement of electrons - electric current - as shown in **Figures 9, 11,** and **12**. Electrons can move without any voltage difference in the continuous loop **5**. Alternatively, the loop may be biased with any chosen voltage.

The relevant Barbosa and Leal patents in Portuguese can be downloaded here:

http://www.free-energy-info.tuks.nl/Barbosa1.pdf

http://www.free-energy-info.tuks.nl/Barbosa2.pdf

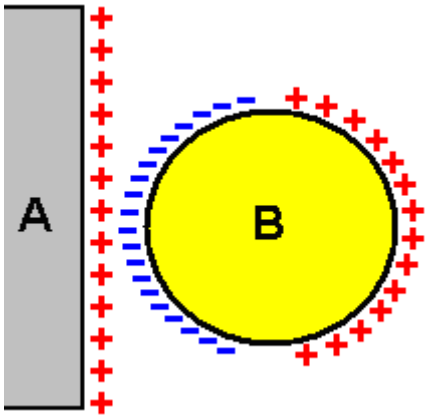
http://www.free-energy-info.tuks.nl/Barbosa3.pdf

Free energy from Lorrie Matchett

The style of operation used by Barbosa and Leal looks as if it is related to the developments of Lorrie Matchett. On 16th June 2008, Lorrie Matchett published his very simple design for a device which captures usable free-energy (video: http://youtu.be/eGD9o7D4To8). His device is based on a very simple and well-know principle of static electricity. This is a principle which is taught in schools all around the world but is generally considered to be of no importance as static electricity is considered too low-power to be of any use. I seriously doubt that anyone who has been struck by lightning would consider static electricity to be “low-power” and suggesting that to them is likely to expand your vocabulary with some words which are seldom heard.

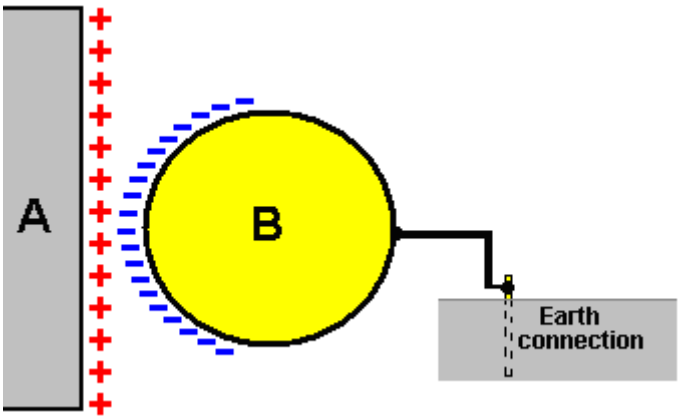
Important Note: the following details mention the use of mains voltages and so let me stress that this presentation is for information purposes only and must not be construed as being a recommendation that you construct or use any such device. Should you choose to ignore this and construct and use Lorrie Matchett’s device, then please be fully aware that you do so entirely at your own risk and nobody else is in any way responsible for the results of what you do.

The principle which is being used here is that an electrically charged object causes the migration of opposite charges on the surface of any object brought close to it. For example if a charged surface if brought close to a metal sphere, then this happens:



The ordinary metal sphere "**B**" which has no particular charge on it is very much affected by being close to a charged surface "**A**" and the closer it gets, the greater the effect. The surface of the sphere had an even distribution of positive and negative charges on its surface, giving it an overall charge of about zero, but the charged surface changes all that. The positive charges on surface "**A**" attract the negative charges on the surface of the sphere causing them to migrate towards surface "**A**". While the positive charges on surface "**A**" do repel the existing positive charges on the surface of the sphere, the migrated negative charges of the sphere itself have an even greater effect, causing the segregation of electrical charges shown above. The situation returns to normal if the sphere is moved away again.

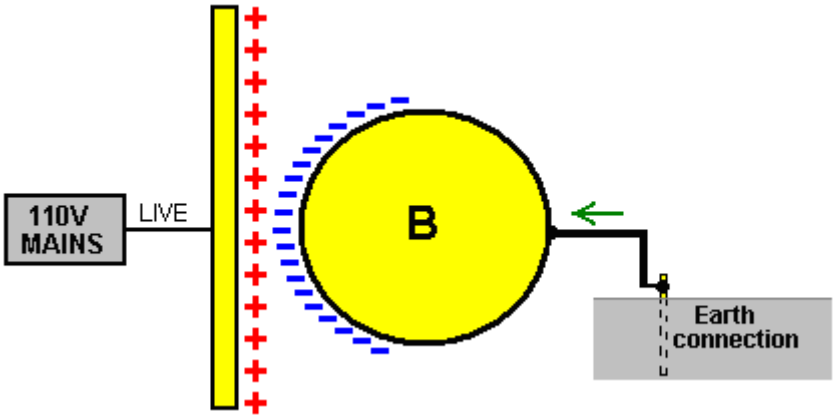
However, the situation changes considerably if the metal sphere "**B**" is connected to the ground:



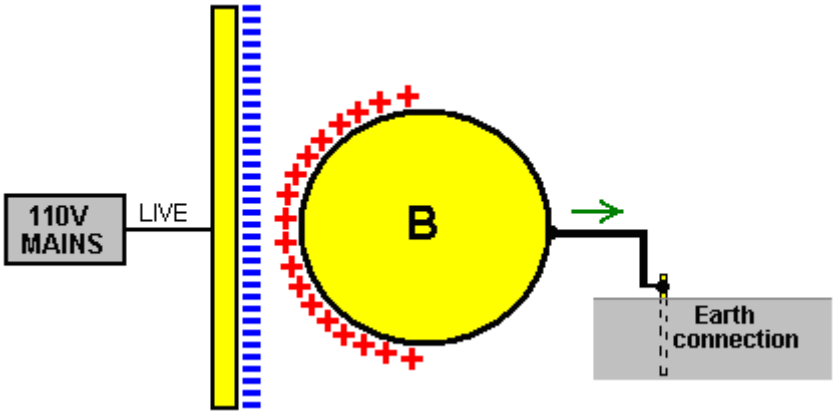
The movement of charges on the surface of the sphere is the same as before, but the Earth has millions of spare charges of both kinds and so, immediately siphons off the excess positive charges on the side of the sphere away from charged surface "**A**". You will notice that charged surface "**A**" is not directly involved in any way and no charge moves from "**A**" to "**B**".

The same effect is seen if the surface "**A**" is negatively charged (except for the fact that the sphere has positive charges rather than the negative charges shown above. The only current flow is along the wire connecting the sphere to the earth connection.

Lorrie Matchett uses this principle, and for the charged surface he connects one end of a brass rod to the 'Live' side of a 100V 60 Hz mains electricity supply. The other end of the brass rod is not connected to anything else. This produces this situation for 8.3 milliseconds:



And then for the following 8.3 milliseconds the mains reverses and you get this situation:

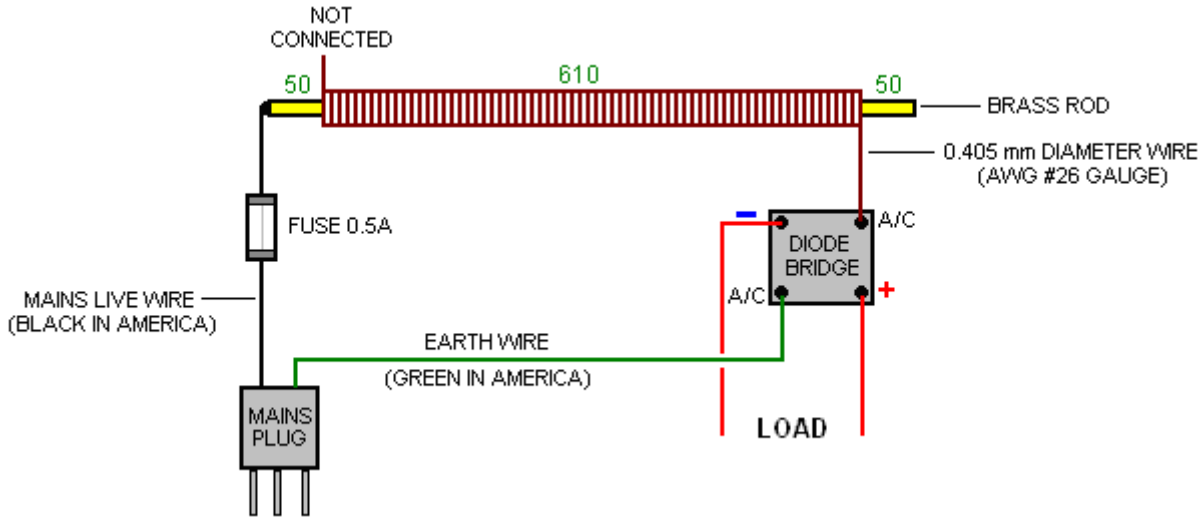


The result of this is that there is a backwards and forwards flow of static electricity along the earth connecting wire, a flow which reverses direction sixty times per second. This is not conventional electricity but is the same form of electricity which is collected by an aerial. Nikola Tesla's patents show many different ways of utilising this static electricity, as does Herman Plauson in his patent (www.free-energy-info.com/Chapter7.pdf). Thomas Henry Moray produced fifty kilowatts of continuous power from quite a small aerial. Paul Baumann of the Swiss commune produced several kilowatts from static electricity. Lorrie Matchett settles for just a few watts and he does it like this:

He connects the live wire of a 110V (RMS) AC mains supply to a brass rod 28-inches (710 mm) long and 3/16 inches (4.76 mm) in diameter. The rod is not directly connected to anything else and so does not form part of a closed loop circuit and so, no current flows from the mains. It must be stressed that the rod and connecting wire are potentially very dangerous and need to be insulated very carefully to ensure that touching them will not cause an electric shock. Please understand very clearly that as no current of any kind is drawn from the mains that this circuit is **not** "stealing electricity from the mains".

For convenience, and only for convenience, Lorrie uses the earthing system of the house mains supply by connecting a green earthing wire to the earth pin of his mains plug. It needs to be clearly understood that this has nothing directly to do with the mains supply and any good quality separate earth would be at least as good as the earthing point inside the mains plug. Effectively, there is only one mains connection.

Instead of using a metal sphere as shown in the illustrations above, Lorrie uses a coil of wire wound around the insulation layer on his brass rod, and he passes the alternating flow of static electricity, drawn from the earth, through a standard diode bridge as shown here:

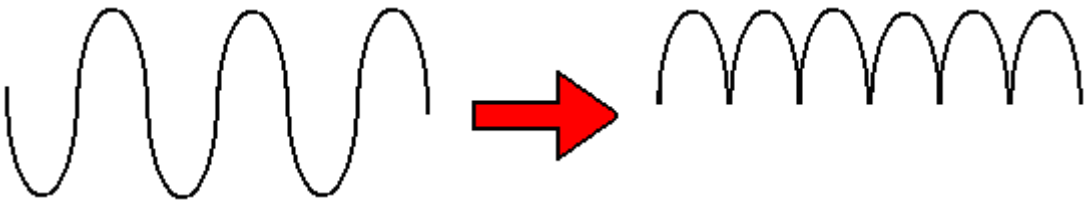


Lorrie covers the brass rod with insulation which is as thin as possible. He suggest heat-shrink tubing for the insulation and on top of it he winds 0.405 mm diameter, solid-core enamelled copper wire, covering a 24-inch (610 mm) length of the rod, placing the turns closely side by side and leaving 2-inches (50 mm) clear at each end of the rod. Thicker wire should not be used.

He also shows a 500 milliamp fuse in the mains supply line. I am not at all happy about that as that fuse can power five incandescent 100-watt mains bulbs connected in parallel, and do you really want that amount of power flowing through you if your insulation is not good enough and you touch it? If you use a fuse in that position I would suggest a 20 mm glass quick-blow 100 mA fuse (mainly because no smaller one is readily available). The fuse is not needed for the circuit and is there in an attempt to protect careless humans.

The coil wound on the insulated brass rod is only connected at one end and that end goes to one of the two "Alternating Current" tags on a 3A diode bridge. Lorrie does not specify the voltage rating for the diode bridge, but it needs to be a minimum of 170-volts if the mains is a 110V (RMS) type, and double that for a 220V (RMS) mains connection. I have no idea why he specifies a 3-amp rating, but the minimum bridge available locally at 3-amps which I would recommend is a 400V rated unit which is supplied at trivial cost.

We need to understand the effect of the diode bridge. It halves the available voltage and doubles the frequency as illustrated here:

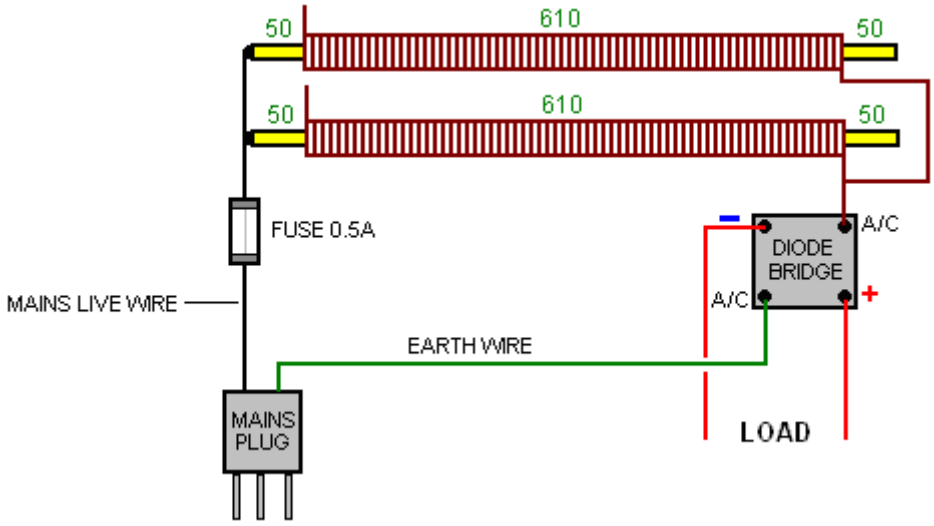


A 110V supply is supposed to swing from Minus 155V to Plus 155V and back again sixty times per second, which is an overall voltage swing of 310V. When passed through a diode bridge that changes to a voltage waveform which swings from Zero volts to Plus 154V and back again 120 times per second, which is an overall voltage swing of 154V which is an average or "RMS" voltage of 109V due to the sine wave shape.

In the rest of the world, the mains voltage is 220V (RMS) nominal, alternating fifty times per second and the Live mains wire is colour coded brown in the UK and the earth wire yellow/green stripes. In passing, the Neutral wire is white for the American 110V system and blue for the 220V system used in the UK.

This design has been brought to my attention by Jes Ascanius of Denmark who is a very able developer of all kinds of free-energy designs. He has replicated this design of Lorrie Matchett and confirms that it works. He has also taken the design further and shares some of the practical details which he has discovered through his own experimentation:

For greater power, additional rods can be used:



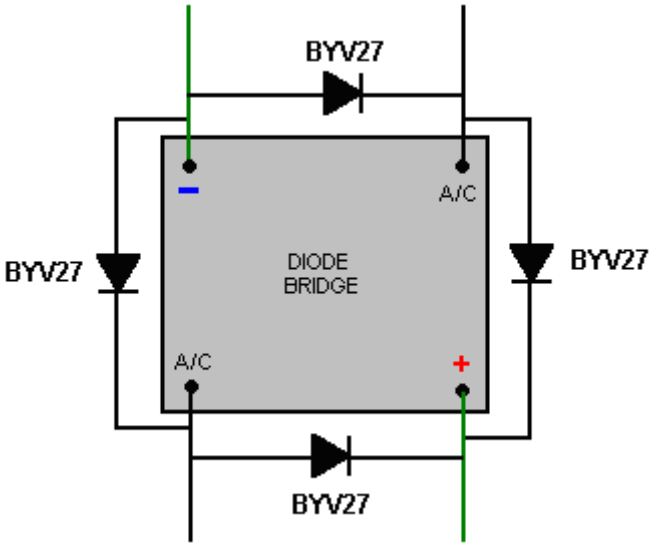
While brass is considered to be the best material for the rod, the diameter is not critical in any way and any size from 5 mm to 20 mm can be used and instead of a rod, a length of brass pipe should be quite suitable. It is also possible to use other materials for the rod but doing that reduces the output power available.

Jes has checked the output of his implementation with the mains fuse removed. The result was an output voltage of 2.6V picked up from the many 220V 50Hz signals generated by the mains wiring all around the place for lighting and sockets. When the fuse is inserted, the voltage rises immediately to 129V with two rods or 162V with five rods. When that voltage is loaded with a 7-watt LED lighting array, the voltage gets pulled down to 61V, but good lighting is being produced for zero current draw from the mains. I would expect that putting a reasonably large capacitor across the load, that the reservoir effect of the capacitor would improve the LED output. Jes has a video of this at https://www.youtube.com/watch?v=zeBqYb2QoAM&feature=player_embedded.

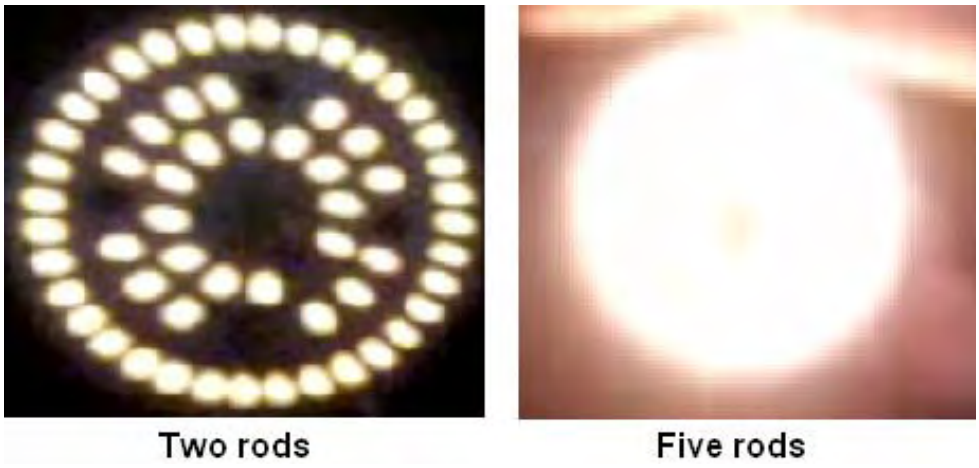
Jes initially used two long rods wound with coils:

And later, five rods. His AC ammeter is sensitive enough to show that due to inefficiencies caused by the tiny stray capacitance between the rods and the coils, there is a very slight current draw from the mains. The mains wattage is far less than the output wattage of the system.

An improvement implemented by Jes is adding four high-speed BYV27 diodes to the ordinary diode bridge like this:

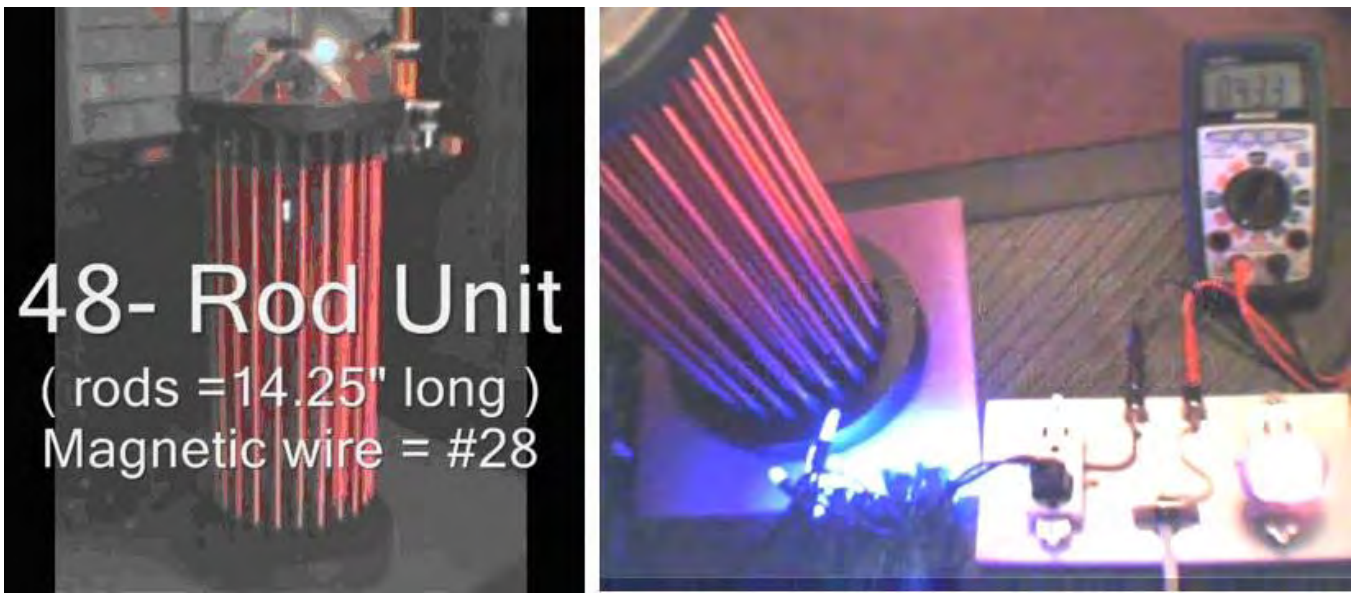


This has the effect of improving the action of the diode bridge and allows more power to extracted from each cycle of the energy flow. When using two brass rods, Jes gets his 5-watt LED array to light up like this:



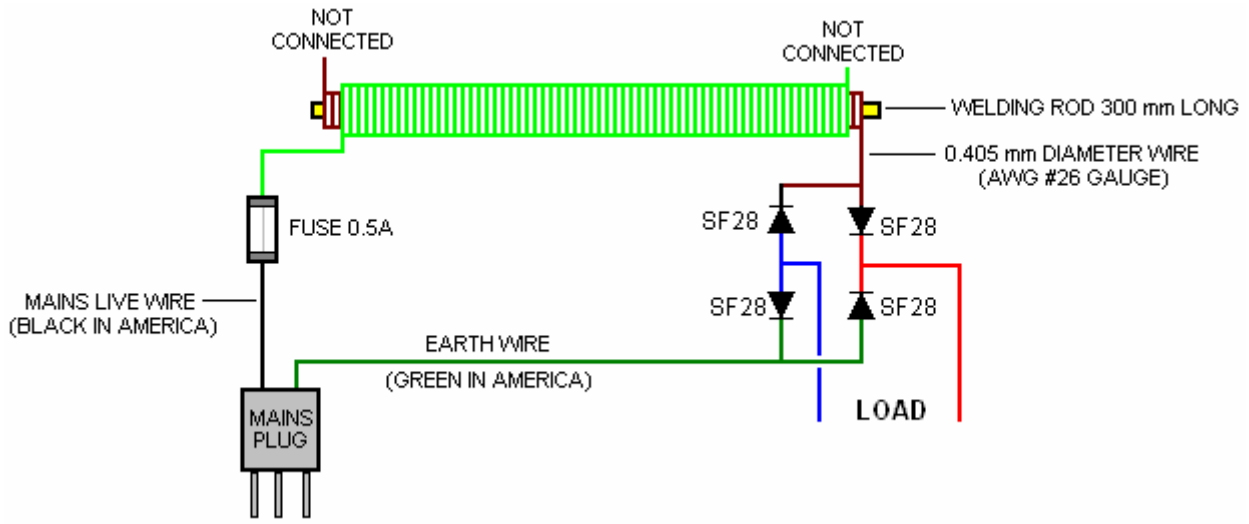
Two rods Five rods

Lorrie also extended his development to a remarkable 48 rods:



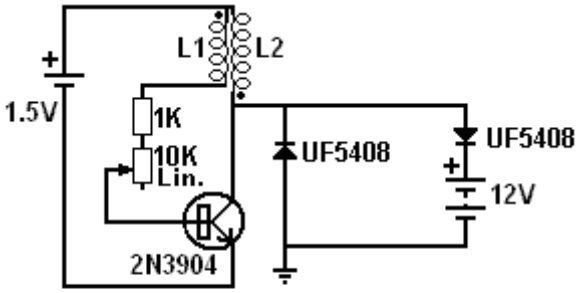
Video link: http://youtu.be/hJyZK6t9qcA The electrical output could be used to charge batteries. Adding extra turns does not increase the output voltage. If the number of turns in each coil matches the output load, then the output power will be greater.

Alexkor in Russia, who is expert in recharging batteries has experimented with this concept and he uses ten coils connected in parallel. He does not use brass, but instead uses the much shorter 300 mm long, 3 mm diameter welding rods with their chemical coating removed. Also, these rods are only used to raise the effectiveness of two separate coils wound on each rod. Each coil is 700 to 750 turns of 0.4 mm diameter wire and the connections are made to the coils and not the rods, as shown here for a single coil pair :



Alex isolates his set of 10 coil-pairs inside a short length of plastic piping:

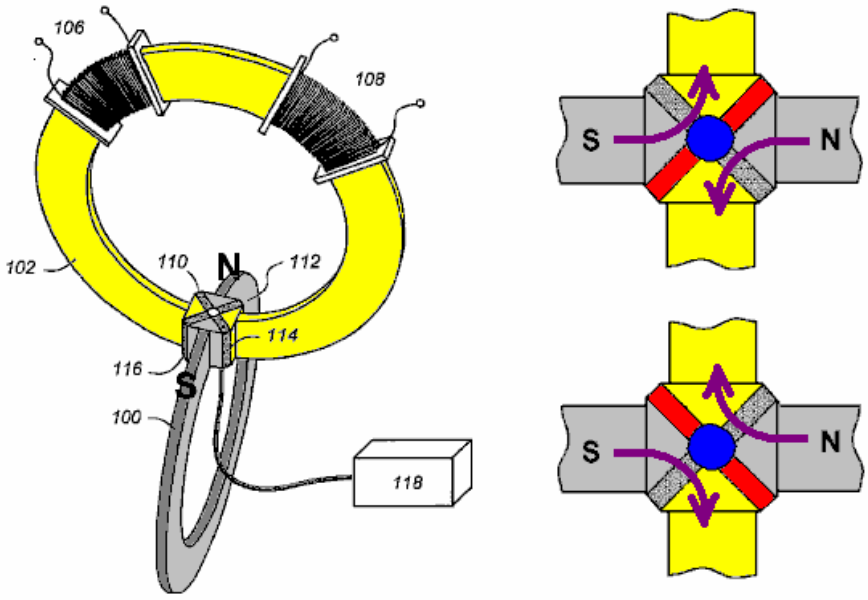
and uses them to power his battery-charging circuit:



Coils: 0.5 to 1.0mm diameter solid copper wire length: 1 to 2 metres bi-filar wound

The Solid-State Magnetostrictive System of Annis and Eberly.

Theodore Annis & Patrick Eberly have produced a variation on this multiple-magnetic-path method which is shown in their US Patent Application 20090096219. They have opted to use a motionless reluctance switch which is a solid-state device which can block magnetic flow when energised. They have arranged one of their devices like this:



The ring shown in grey is a magnet which connects to the ring shown in yellow through two diagonal ‘reluctance’ (magnetic flow) switches. The yellow ring can carry magnetic flux and the control box marked **118** switches the diagonal strips on and off in turn, causing the magnetic flux to reverse it’s direction through the yellow ring. The coils wound on the yellow ring pick up this reversing magnetic flux and pass it out as an electric current. While only one pair of rings are shown here, the design allows for as many rings as are needed to be connected together as shown here:

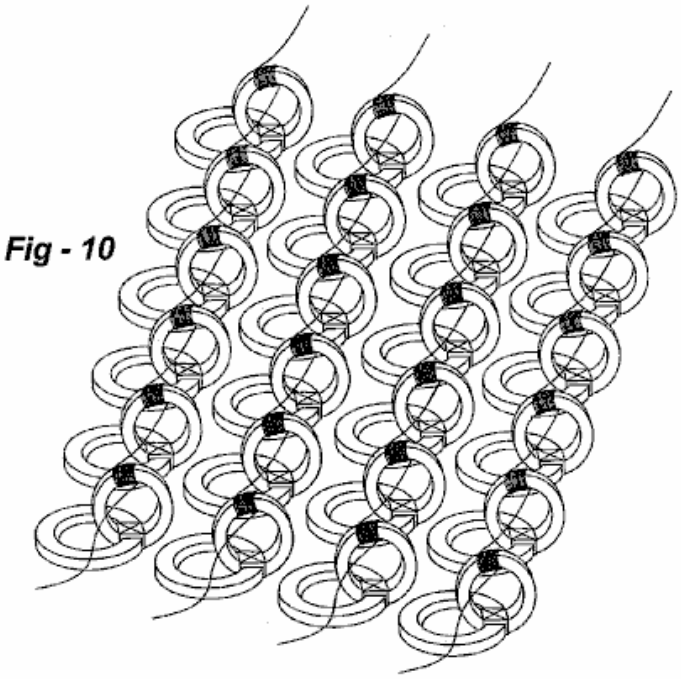


Fig - 10

The patent says: ”The currently preferred motionless reluctance switch is described by Toshiyuki Ueno & Toshiro Higuchi, in their paper entitled „Investigation of the Dynamic Properties of a Magnetic Flux Control Device composed of Laminations of Magnetostrictive Piezoelectric Materials” – University of Tokyo 2004. As shown in Fig.4, this switch is made of a laminate of a Giant Magnetostrictive Material **42**, a TbDyFe alloy, bonded on both sides to a Piezoelectric material **44**, **46** to which electricity is applied. The application of electricity causes the reluctance of the piezoelectric material to increase.

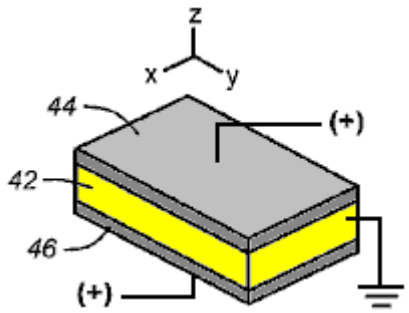


Fig - 4

This original patent application is included in the Appendix.

However, very interestingly, there is another, completely different patent application from Annis and Eberly, with the same publication date and the same number. It is not at all obvious to me how that could be, but here is the bulk of that other patent application (the original being in the Appendix).

ENERGY GENERATION APPARATUS AND METHODS BASED UPON MAGNETIC FLUX SWITCHING

Abstract

In an electrical energy generator, at least one permanent magnet generates flux and a magnetisable member forms the single flux path. An electrically conductive coil is wound around the magnetisable member, and a plurality of flux switches are operative to sequentially reverse the flux from the magnet through the member, thereby inducing electrical current in the coil. A “Figure of Eight” construction comprises two continuous loops of magnetisable material sharing a magnetisable member common to both loops. An alternative configuration uses stacked loops and a separate piece of material acting as the magnetisable member. One end of the magnet is coupled to one of the loops, with the other end being coupled to the other loop. Each loop further includes two flux switches operated in a 2 × 2 sequence to sequentially reverse the flux through the magnetisable member. A relatively small amount of electrical power is used to control the magnetic flux of a permanent magnet by switching the flux between alternate paths. The resulting power from the switched magnetic flux yields substantially more power than the power required for the input switching.

Description

FIELD OF THE INVENTION

This invention relates generally to energy generation and, in particular, to methods and apparatus wherein magnetic flux is switched through a flux path to produce electricity.

BACKGROUND OF THE INVENTION

Magnetic flux may exist in “free-space,” in materials that have the magnetic characteristics of free-space, and in materials with magnetically conductive characteristics. The degree of magnetic conduction in magnetically conductive materials is typically indicated with a B-H hysteresis curve, by a magnetisation curve, or both.

Permanent magnets may now be composed of materials which have a high coercively (Hc), a high magnetic flux density (Br), a high magneto motive force (mmf), a high maximum energy product (BHmax), with no significant deterioration of magnetic strength over time. An example is the NdFeB permanent magnet from VAC of Germany, which has an Hc of 1,079,000 Amperes/meter, a Br of 1.427 Tesla, an mmf ranging up to 575,000 Ampere-turns, and a BHmax of 392,000 Joules/meter^3.

According to Moskowitz, “Permanent Magnet Design and Application Handbook” 1995, page 52, magnetic flux may be thought of as flux lines which always leave and enter the surfaces of ferromagnetic materials at right angles, which never can make true right-angle turns, which travel only in straight or curved paths, which follow the shortest distance, and which follow the path of lowest reluctance (resistance to magneto motive force).

Free space presents a high reluctance path to magnetic flux. There are many materials which have magnetic characteristics similar to those of free space. There are other materials which offer a low or lower reluctance path for magnetic flux, and it is these materials that typically comprise a defined and controllable magnetic path.

High-performance magnetic materials for use as magnetic paths within a magnetic circuit are now available and are well suited for the (rapid) switching of magnetic flux with a minimum of eddy currents. Certain of these materials are highly non-linear and respond to a "small" applied magneto motive force (mmf) with a robust generation of magnetic flux (B) within the material. The magnetisation curves of such materials show a high relative permeability (ur) until the "knee of the curve" is reached, at which point ur decreases rapidly approaching unity as magnetic saturation (Bs) is reached.

Some of these non-linear, high-performance magnetic materials are referred to as "square" due to the shape of their B-H hysteresis curves. An example is the FINEMET® FT-3H nanocrystalline core material made by Hitachi of Japan. Other examples include Superperm49, Superperm80, SuperMalloy, SuperSquare80, Square50, and Supermendur, which are available from Magnetic Metals in the USA.

A "reluctance switch" is a device or means that can significantly increase or decrease (typically increase) the reluctance of a magnetic path. This is ideally done in a direct and rapid manner, while allowing a subsequent restoration to the previous (typically lower) reluctance, also in a direct and rapid manner. A reluctance switch typically has analogue characteristics. By way of contrast, an off/on electric switch typically has a digital characteristic, as there is no electricity "bleed-through." With the current state of the art, however, reluctance switches exhibit some magnetic flux bleed-through. Reluctance switches may be implemented mechanically, such as to cause keeper movement to create an air gap, or electrically by various other means.

One electrical reluctance switch implementation uses a control coil or coils wound around a magnetic path or a sub-member that affects the path. U.S. Navy publication, "Navy Electricity and Electronics Series, Module 8 - Introduction to Amplifiers" September 1998, page 3-64 to 3-66 describes how to modulate alternating current by changing the reluctance of the entire primary magnetic path by these means, one of which is used in a saturable-core reactor and the other in a magnetic amplifier. Flynn, U.S. Pat. No. 6,246,561; Patrick et al., U.S. Pat. No. 6,362,718; Pedersen, U.S. Pat. No. 6,946,938; Marshall, and US Patent Application 2005/01256702-A1 all disclose methods and apparatus that employ this type of reluctance switch for switching magnetic flux from a stationary permanent magnet or magnets for the purpose of generating electricity (and/or motive force).

Another electrical means of implementing a reluctance switch is the placement within the primary magnetic path of certain classes of materials that change (typically increase) their reluctance upon the application of electricity. Another electrical means of implementing a reluctance switch is to saturate a sub-region of a primary magnetic path by inserting conducting electrical wires into the material comprising the primary magnetic path. Such a technique is described by Konrad and Brudny in "An Improved Method for Virtual Air Gap Length Computation," in IEEE Transactions on Magnetics, Vol. 41, No. 10, October 2005.

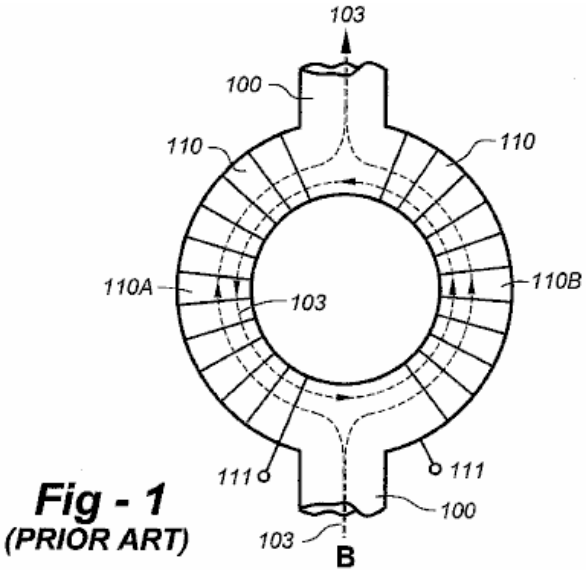


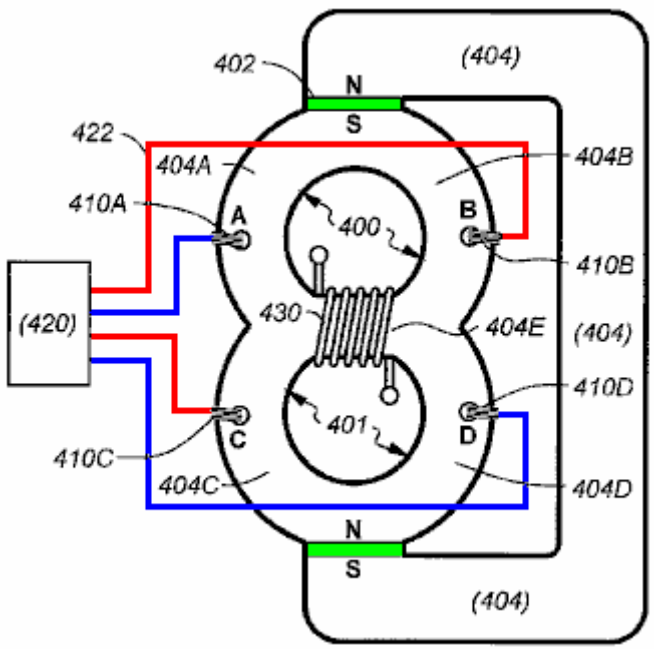
Fig - 1
(PRIOR ART)

Another electrical means of implementing a reluctance switch is described by Valeri Ivanov of Bulgaria on the web site www.inkomp-delta.com, shown in **Fig.1**. An electric toroid **110** is inserted into a primary magnetic path (**100**), so that the primary magnetic path is divided into two sub-paths **110A** and **110B**. A net magnetic flux reduction

effect in the primary magnetic path **100** results from the combination of the effects in the two sub-paths **110A** and **110B**, each of which results from different physics principles. In the first sub-path **110A**, the magnetic flux generated by applying electrical current to the windings **110** around toroidal path **110** opposes and subtracts from its portion of the magnetic flux **103** received from the primary magnetic path **100** yielding a reduced magnetic flux, which is also further reduced by a decrease in the sub-path **110A**'s relative permeability thereby increasing the reluctance of the sub-path. In the second sub-path **110B**, the magnetic flux generated by applying electrical current to the toroid windings **111** adds to its portion of the magnetic flux **103** received from primary magnetic path **100** yielding an increased net magnetic flux that approaches or exceeds the knee of the material's magnetisation curve thereby reducing its relative permeability and increasing its reluctance.

SUMMARY OF THE INVENTION

This invention is directed to methods and apparatus where magnetic flux is switched in both direction and intensity through a flux path to produce electricity. The apparatus broadly comprises at least one permanent magnet generating flux, a magnetisable member forming the flux path, an electrical conductor wound around the magnetisable member, and a plurality of flux switches operating to reverse sequentially, the flux from the magnet flowing through the member, thereby inducing electrical current in the coil.



The preferred embodiment includes first and second loops of magnetisable material. The first loop has four segments in order **A, 1, B, 2**, and the second loop has four segments in order **C, 3, D, 4**. The magnetisable member couples segments **2** and **4**, and the permanent magnet couples segments **1** and **3**, such that the flux from the magnet flows through segments **A, B, C, D** and the magnetisable member. Four magnetic flux switches are provided, each controlling the flux through a respective one of the segments **A, B, C, D**. A controller is operative to activate switches **A-D** and **B-C** in an alternating sequence, thereby reversing the flux through the segment and inducing electricity in the electrical conductor. The flux flowing through each segment **A, B, C, D** is substantially half of that flowing through the magnetisable member prior to switch activation.

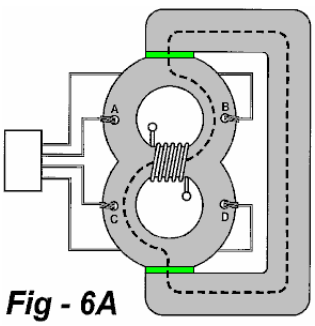


Fig - 6A

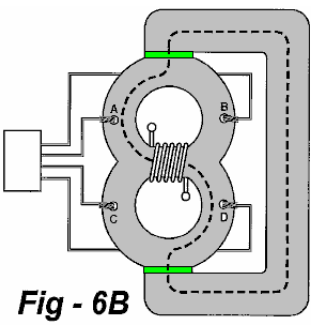


Fig - 6B

The loops and magnetisable member are preferably composed of a nanocrystalline material exhibiting a substantially square BH intrinsic curve. Each magnetic flux switch adds flux to the segment it controls, thereby magnetically saturating that segment when activated. To implement the switches, each segment may have an aperture formed through it and a coil of wire wound around a portion of that segment and through the aperture. The controller may be at least initially operative to drive the switch coils with electrical current spikes.

The first and second loops may be toroidal in shape, and the loops may be spaced apart from one another, with **A** opposing **C**, and **1** opposing **3**, and with **B** opposing **D** and **2** opposing **4**. The magnetisable member in this case is preferably a separate piece of material. Alternatively, the first and second loops may form a "Figure of eight" shape, with the two loops intersecting to form the magnetisable member.

The permanent magnet(s) and the material comprising the magnetic paths are preferably proportioned such that the material through the common segment is at or slightly below its maximum relative permeability before the electrically conducting output coil is energized. In the preferred embodiments, the power resulting from the switched magnetic flux yields substantially more power than the power required for the input switching.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig.1 is a drawing of a prior art reluctance switch in the form of an electrical toroid inserted into a primary magnetic path;

Fig.2 is a detail drawing of a reluctance switch according to the invention;

Fig.3A and **Fig.3B** are detail drawings showing the use of four reluctance switches according to the invention;

Fig.4 is a drawing which depicts a preferred embodiment of the invention;

Fig.5 is a detail drawing an alternative reluctance switch according to the invention implemented through split laminations;

Fig.6A and **Fig.6B** show the operation of an energy generator according to the invention;

Fig.7A is an exploded view of a preferred energy generator construction;

Fig.7B is a side view of the construction shown in **Fig.7A**;

Fig.8 is a simplified schematic diagram of components used to simulate the apparatus of the invention;

Fig.9A is a diagram that shows the current delivered to one pair of flux switches in the simulation;

Fig.9B is a diagram that shows the current delivered to the other set of flux switches in the simulation;

Fig.10 shows the output of the simulation shown here; and

Fig.11 is a block diagram of a controller applicable to the invention

DETAILED DESCRIPTION OF THE INVENTION

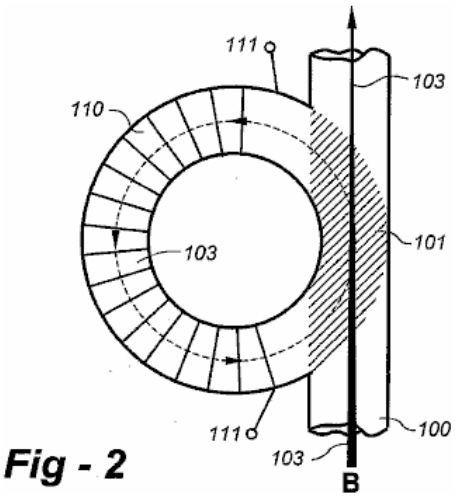


Fig - 2

Fig.2 is a detail drawing of a reluctance switch according to the invention. The reluctance switch includes the following components: a closed magnetic path **110** comprised of a high performance magnetic material (preferably a non-linear material exhibiting a "sharp knee" as saturation is approached), around which is wound a coil **111**. The closed magnetic path **110** shares a common segment **101** with a primary magnetic path **100**, in which magnetic flux **103** is induced by a permanent magnet (shown in subsequent drawings). Electric current is applied to windings **111** having a polarity and sufficient amperage so that the magnetic flux generated in the path

of switch **110** is additive to the magnetic flux **103** from the permanent magnet, such that the primary path **110** approaches or reaches magnetic saturation.

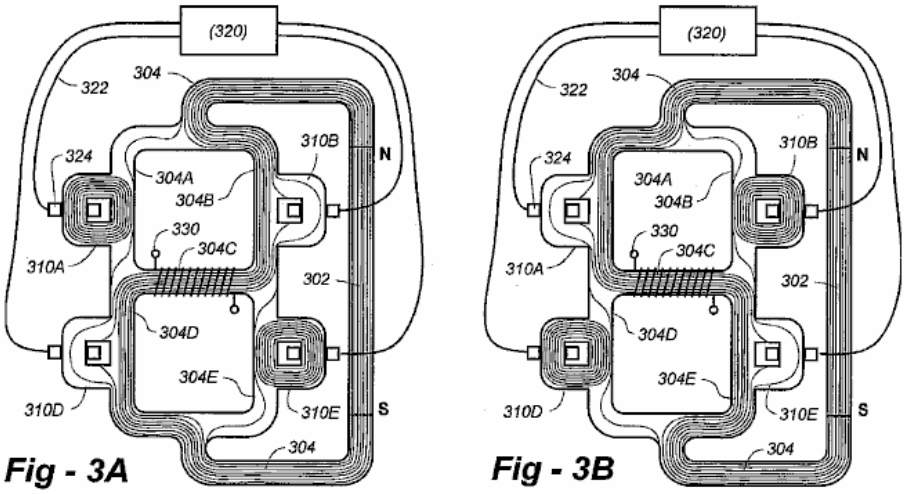


Fig.3A and **Fig.3B** are detail drawings of an apparatus which employs four reluctance switches according to the invention in a manner similar to that disclosed in U.S. patent application Ser. No. 11/735,746 entitled “Electricity Generating Apparatus Utilising a Single Magnetic Flux Path”, the entire content of which is incorporated herein by reference. In this and in all embodiments described herein, the geometry of the closed magnetic paths may be circular (toroidal), rectangular, or any other closed-path shapes . A primary path **304** carries the flux from permanent magnet **302** unidirectionally. Flux switch pairs **310 A/E** and **310 B/D** are activated in alternating fashion to reverse the flux in magnetisable member **304C**, thereby inducing electrical current in winding **330**. **Fig.3A** shows the flux flow in one direction, and **Fig.3B** shows it reversed.

In **Fig.3A**, switches **310A** and **310E** are activated by controller **320** in electrical communication with the windings on the switches such as through conductor **322** to winding **324**. The additional flux in switches **310A** and **310E** are additive with the flux that would otherwise be present in segments **304A** and **304E**, thereby saturating these paths, causing the flux through segment **304C** to be in the direction shown. In **Fig.3B**, switches **310B** and **310D** are activated, saturating segments **304B** and **304D**, and reversing the flow.

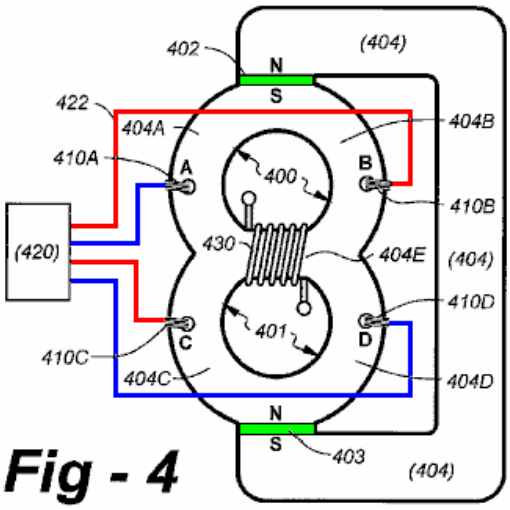


Fig.4 is a drawing that depicts an embodiment of the invention using circular toroids **400**, **401** and multiple permanent magnets **402**, **403** disposed in the primary path **404**. The two toroids **400**, **401** intersect, forming magnetisable member **404E**. A coil **430** is wound around the member **404E**, as shown.

The primary magnetic path **404** interconnects the upper end of loop **400** and the lower end of loop **401**. One of the magnets, **402**, couples one end of the primary magnetic path **404** to the first loop **400**, and another, **403**, couples the other end of the primary magnetic path **404** to the second loop **401**.

In this, and all of the embodiments described here, the permanent magnets are strong, rare-earth magnets, and multiple magnets of any length (thickness) may be used in each case. Further, in all embodiments, the loops, primary magnetic path and/or magnetisable member are preferably constructed from a high magnetic permeability

material such as the FINEMET FT-3H nanocrystalline soft magnetic material available from Hitachi. The invention is not limited in this regard, however, as alternative materials, including laminated materials, may be used.

The connections of the primary magnetic path **404** to the two loops **400**, **401** create four segments apart from the magnetisable member **404E**, the four segments including two opposing segments **A, B** in the first loop on either side of magnet **402**, and the two opposing segments **C, D** in the second loop on either side of magnet **403**.

Four magnetic flux switches are provided, each being operative to control the flux through a respective one of the four segments. A controller **420** is operative to activate the switches associated with segments **A** and **D**, and then **B** and **C**, alternately, thereby reversing the flux through the member **404E**, and so, inducing electrical current in coil **430**.

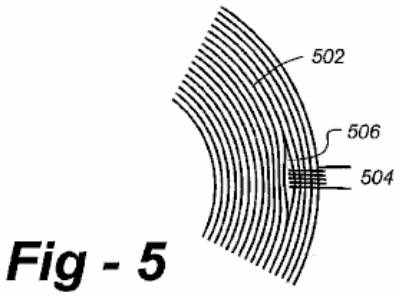


Fig - 5

Apertures may be formed through each of the four segments, with the switches being implemented by coils **410A** to **410D** which pass through the apertures and around an outer (or inner) portion of each segment. As shown in **Fig.5**, if the loops are fabricated with laminated material **502**, the laminations may be split at **506** to accommodate coil **504**. The percentage of the segment surrounded by the coil may vary in accordance with the material used, the waveforms presented to the coils, and other factors, with the goal being to magnetically saturate each segment through activation of the associated switch, thereby reversing the flux through path **404E**.

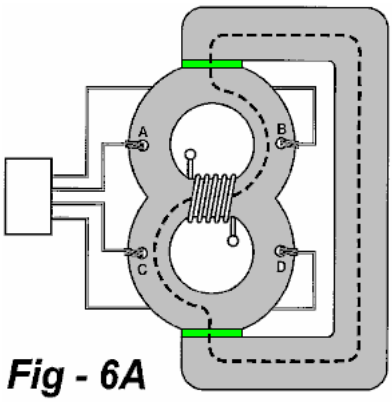


Fig - 6A

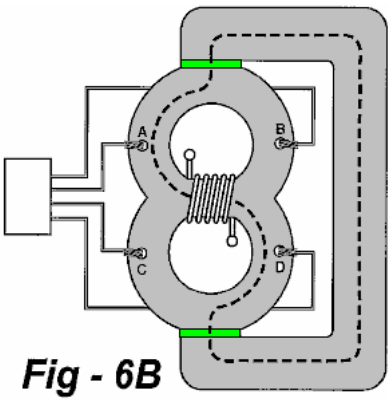


Fig - 6B

Fig.6A and **Fig.6B** show the operation of the apparatus of **Fig.4**. The primary path **404** carries the flux from permanent magnets **402** and **403** unidirectionally. Reluctance switches **410A** to **410D** are activated alternately to reverse the flux in segment **404E** which, in turn, induces electrical current in winding **430**. **Fig.6A** shows the flux flow in one direction, and **Fig.6B** shows it flowing in the opposite direction.

In **Fig.6A**, switches **410A** and **410D** are activated by controller **420** in electrical communication with the windings on the switches, such as through conductors **422** to switch **410B**. The flux provided by switches **410A** and **410D**, thereby saturating these paths, causing the flux through segment **404C** to be in the direction shown. In **Fig.6B**, switches **410B** and **410C** are activated, saturating segments **404B** and **404D**, thereby reversing the flux through path **404E**.

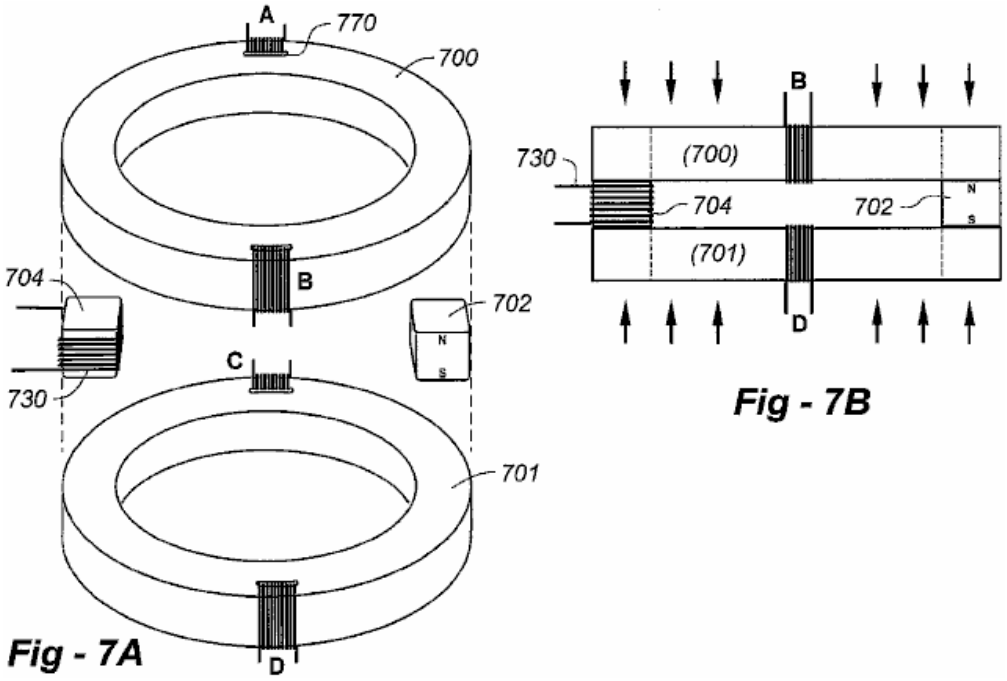


Fig.7A depicts a preferred construction of the apparatus depicted in **Fig.4**, **Fig.6A** and **Fig.6B**. Loops **400** and **401** are implemented as complete toroids **700**, **701**. This is important, since preferred high-performance magnetic materials are currently available in regular shapes of this kind. Note that, in this case, curved slots such as **770** are formed through the sides of each toroid to implement flux switches **A** to **D**. The magnetisable member in this embodiment is implemented with a block of material **704**, preferably the same high-performance magnetic material used to construct loops **400**, **401**. Permanent magnet **702**, shown at **702**, preferably has the same length as block **704**, enabling the various constituent parts to be held together with compression, shown in **Fig.7B**.

The following sections summarise some of the important characteristics of the preferred embodiments:
In terms of materials, the apparatus benefits from the use of nanocrystalline material with a “Square” BH intrinsic curve, a high Br (remanence) which is about 80% of its Bs (saturation), a low Hc (coercivity), and a fast magnetic response time to saturation. An example is FineMet FT-3H from Hitachi of Japan, which has a Br of 1.0 Tesla, a Bs (saturation) of 1.21 Tesla, a time to saturation (Bs) of 2 usec, and an Hc of −0.6 amp-turns/meter.

Modern permanent magnets are used with a square BH intrinsic curve, a Br in the range of 1.0 Tesla or more, and high Hc in the range of −800,000 amp-turns/meter or more. An example is the NdFeB magnet from the German company VAC, which has a Br of 1.427 Tesla and an Hc of −1,079,000 amp-turns/meter.

An important consideration is the matching of the magnet to the nanocrystalline material, both in Tesla rating and in cross-sectional area. The magnet's Br should be below the Bs of the nanocrystalline material. If the magnet is too “strong” for the nanocrystalline material, this may cause the nanocrystalline material to saturate at the area of contact with the magnet.

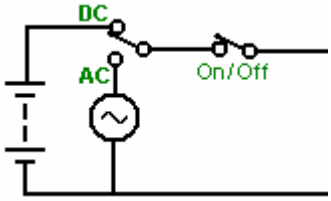
The current driving the reluctance switches in the prescribed 2 × 2 sequence should have a sharp rise in the leading edge (Tr) of each pulse with a pulse width (Pw) and Amperage value that are sustained until released at the end of the pulse width (Tf). The table below shows the effects of input current pulse rise times (Tr) on the output. There exists a narrow band of Tr, before which there is small power output, at which there are excellent power output and COPs in the range of 200 to 400 or greater, and after which there is no major increase in power output. The COP of this device without the coupling circuit is defined as “Output power/Drive Power” for the switches.

Tr	Output Power	Waveform Description
1.0E–4 secs	50 Watts	Spikes
7.5E–5	50 Watts+	Spikes with intermittent 30 Kilowatt square waves
5.0E–5	15 Kilowatts	Square waves after 3 cycles
1.0E–5	15 Kilowatts	Square waves after 1 cycle

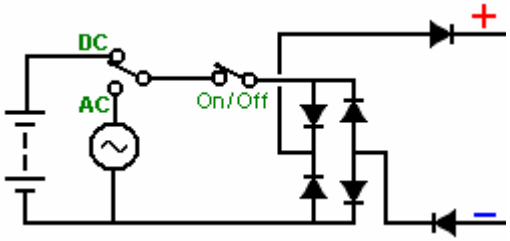
The Motionless Generator of Richard Willis.

On 28th May 2009 a European Patent application was filed by Richard Willis, entitled "Electrical Generator". During a TV interview, Richard stated that his design has COP=3,600. Available commercially from his Canadian company and sold under the name "Magnacoster", early in 2010 his advertised pricing was US $4,200 for a unit which has four separate 100 amp 12V outputs, giving a combined maximum output power of 4.8 kilowatts. A larger unit was priced at US $6,000 with four separate 24V outlets providing a 9 kilowatt combined output. The house-powering unit which is supplied with a 12 kilowatt inverter to provide mains AC power and which gets connected direct to the circuit-breaker box of the house, was priced at US $15,000. One particularly interesting statement made by Richard is that the output power is at a higher frequency than the input power. He suggests that the electrical signal bounces around inside the device, multiplying the power as it goes and giving the output higher voltage and higher current than the input. The design of the device is most interesting as it is very simple. It is shown in his patent application WO 2009065219, a somewhat reworded copy of which is included in the Appendix to this eBook. Richard's web site is http://www.vorktex.ca/page/235610203. However, while Richard's designs do indeed work, he appears to be experiencing problems with the output wiring melting due to the very high current, and more importantly, the generation of high levels of unwanted electromagnetic radiation. These problems appear to have prevented him from supplying any commercial units at this time.

The circuit is based on a pulsed coil and two magnets and it has a number of unusual features. The power supply is unusual:

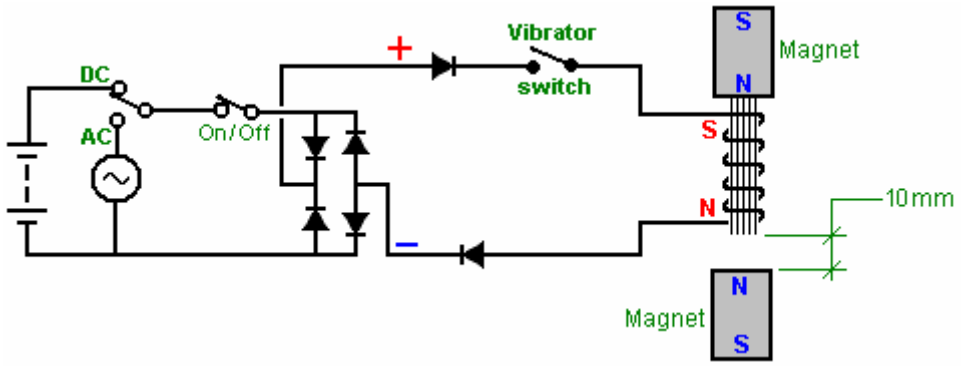


Richard arranges it like this so that either DC or AC can be used as the input power and so he follows that arrangement with a diode bridge, followed by two more diodes as shown here:



This is an interesting arrangement when the input is DC as it would be a more usual arrangement to have the diode bridge only in the AC input section and not included for the DC input where it just drops the input voltage and wastes electrical power unnecessarily. Still, that is the way it is shown in the patent, so that is the way it is shown here.

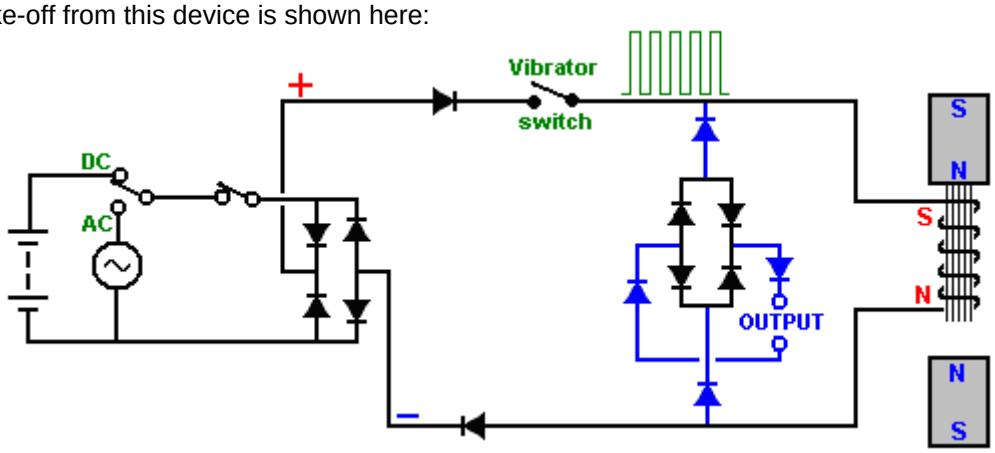
The input power supply is fed to an electromagnet but is converted into a pulsed supply by the use of an interrupter switch which may be mechanical or electronic:



As can be seen, the arrangement is particularly simple although it is an unusual configuration with the electromagnet core touching one of the permanent magnets and not the other. The magnet and electromagnet poles are important, with the permanent magnet North poles pointing towards the electromagnet and when the electromagnet is powered up, it's South pole is towards the North pole of the permanent magnet which it is touching. This means that when the electromagnet is powered up, it's magnetic field strengthens the magnetic field of that magnet.

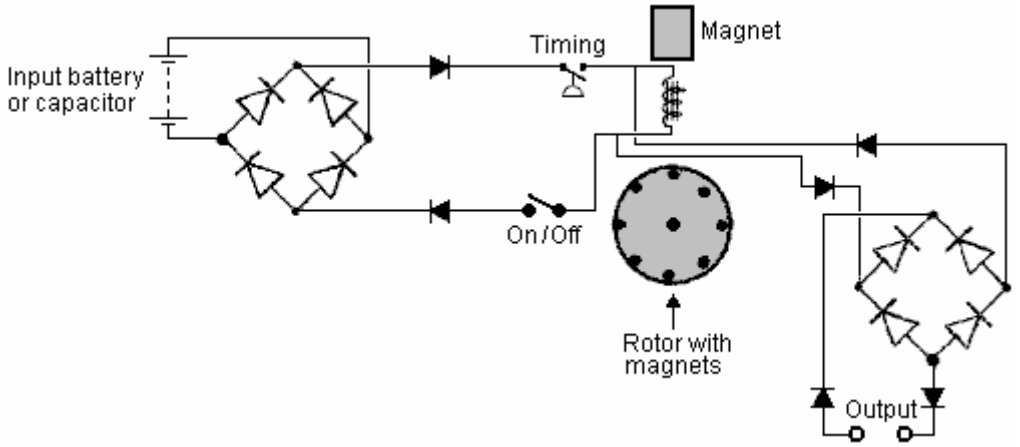
There is a one-centimetre gap at the other end of the electromagnet and it's North pole opposes the North pole of the second permanent magnet. With this arrangement, each electromagnet pulse has a major magnetic effect on the area between the two permanent magnets. In the diagram shown above, just a few turns of wire are shown on the electromagnet core. This is just for clarity and it does not mean that only a few turns should be used. The strength of the magnets, the electromagnet wire thickness and number of turns are related to each other and experimentation will be needed to determine the best combination.

The energy take-off from this device is shown here:



Richard states that the input power can be anywhere from under one volt to one million volts while the input current can be anything from under one amp to one million amps, so he clearly envisages a major range of constructions and components. The core material for the electromagnet is specified as ferrite, mumetal, permalloy, cobalt or any non-permeable metal material. It seems likely that iron filings embedded in epoxy resin is likely to be a suitable material as it can respond very rapidly to sharp pulses and it seems clear that in common with almost every other similar free-energy device, the rapidity of rise and fall of the power pulse is of major importance. Having said that, Richard states that the frequency of pulses in the output section is greater than the frequency of pulses applied to the input section. From this it seems likely that the device should be tuned so that the input pulses should be at a lower harmonic of the resonant frequency of the device. It is worth reading Richard's full description which is near the end of the Appendix.

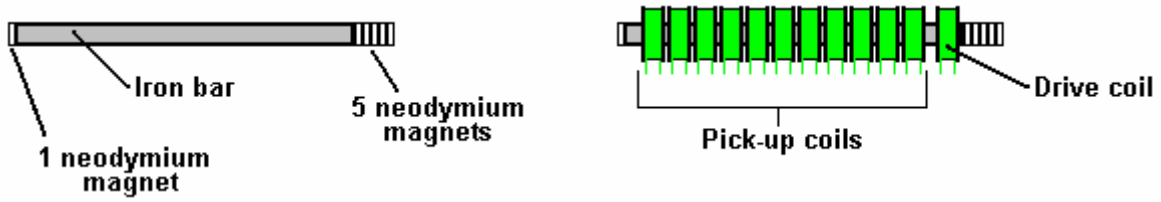
A second version of the circuit looks like a modification of the John Bedini pulsed rotor battery charging circuit with a rotor substituting for the second permanent magnet:



This enhances the operation of the Bedini device by providing an initial magnetic field in the coil.

The Generator Proposal of "Silverhealtheu".

One of the EVGRAY yahoo forum members whose ID is 'silverhealtheu' has described a simple device which appears to be not unlike the Richard Willis generator above.



The device consists of an iron bar one inch (25 mm) in diameter and one foot (300 mm) long. At one end, there is a stack of five neodymium magnets and at the opposite end, a single neodymium magnet. At the end with the five magnets, there is a coil of wire which is strongly pulsed by a drive circuit. Down the length of the bar, a series of pick-up coils are positioned. Each of these coils picks up the same level of power that is fed to the pulsing coil and the combined output is said to exceed the input power.

The Motionless Generators of Heinrich Kunel

While Richard Willis of Magnacoaster in Canada, has gone commercial with his generators, making each of the generator outputs multiples of 12V 100A, a lot of very interesting earlier information can be found in the 1982 patent of Heinrich Kunel. The patent describes four separate configurations of his basic design, a design which looks very much like that used by Richard Willis. Here is an attempted translation of the Kunel patent which is in German:

PATENT: DE3024814 28 January 1982 Inventor: Kunel, Heinrich

PROCEDURES AND DEVICES FOR ENERGY PRODUCTION

Application Number: DE19803024814 19800701
Priority Number(s): DE19803024814 19800701
IPC Classification: H02N11/00 EC Classification: H02K53/00

DESCRIPTION

The invention concerns procedures and devices for energy production, which convert the magnetic flux from permanent magnets, without the need for rotary, or any other form of motion, into temporally variable induced flux and produce large fluctuations of the induced magnetic flux into electric current, without the need for a thermal circuit, or torque or chemical process, and in such a manner that the electric current is amplified.

The energy problems of our time are sufficiently well-known in the professional world. The transformation of conventional primary energy forms into technically usable energy is relatively cost-intensive. Additionally, by doing this, precious and ever scarcer raw materials are being destroyed and pollutants are accumulated, pollutants which are able to cause the end of mankind.

In order to reduce or avoid these and other disadvantages during power production, it is suggested according to this invention, that for the production of electric current by induction, no torque is used, but instead, the magnetic flux from permanent magnets is converted into an induced flux with large rapid variations, e.g. into a rapidly changing or pulsating induced flux, which can generate an electric current.

In accordance with this invention, one proceeds from the basic arrangement where a permanent magnet is connected to an iron core via one or both of its poles, the core consisting of, for example, dynamo iron, pure iron or amorphous iron or any similar suitable material which has little or no core losses.

Following the procedure of this invention if, for example, one pole of a bar magnet touches such an iron core and is aligned with the longitudinal axis of that core, then both the magnet and the core act as a magnet.

During this initial magnetisation process of the core a magnetic flux flows in the core, which induces a current in a conducting circuit surrounding the core.

If, in addition to the permanent magnet, a coil is wound around the core, and arranged in such a way that a current flowing in that coil wholly or partially interrupts the magnetic flux flowing through the core, then this causes another modification of the magnetic flux passing through a coil wound around the core.

If this interruption of the flux flowing from the permanent magnet is large, and has the waveform of a Alternating Current, then a pulsating Direct Current is induced in the coil wound around the core.

With an AC input to the magnetic flux modifier coil wound on the core beside the magnet, the coil receives two current flow direction changes per AC cycle, and so the magnetic flux coming from the permanent magnet is interrupted once and released once during each input power cycle. In this way, a pulsating induced Direct Current is generated in the output coil caused by the pulsating magnetic flux in the core.

It was found that the induced flux from a permanent magnet achieves its full initial value of the magnetic flux density in the core also at the free end of the soft-iron magnetic core, even if there are several induction windings each with the same number of turns and the conductor cross section are appropriate as in a magnetic flux modification coil wound on the core, without changing the strength of the magnetic flux density or the remanence of the permanent magnet.

The permanent magnet is not demagnetised when used to provide the flux in the core, no energy is extracted from it, unlike an electromagnetic core, whose windings require more operating current than that produced as an output. With an electromagnetic core, as much input current is required, as is induced in the output winding, corresponding to the relationships of the well-known transformer. Therefore it is important to produce the induction flux by using a permanent magnet.

After the basic primary system, one can build e.g. energy linear generators or energy circle generators or other arising or suitable types and forms of energy generators, without the necessity of a rotor or a stator or any such mobile sections or providing torque in the generator.

The invention is so designed as to make only frequency control governed by electrical means so that the internal induction flux in the generator core is essentially caused by the magnetic field of the permanent magnet.

In the drawings constructional examples are schematically represented according to the invention:

Fig.1 shows a linear energy generator in longitudinal cross-section.

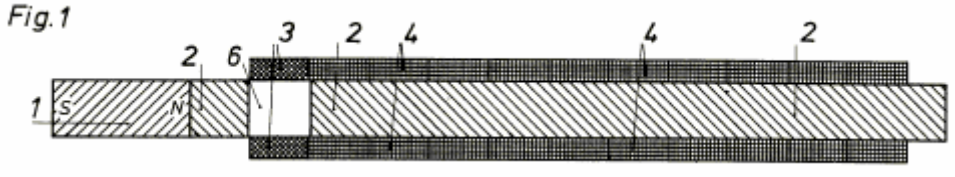


Fig.2 shows a linear energy generator in the instantaneous state of the transmission of the induction flux of the permanent magnet to the generator core and

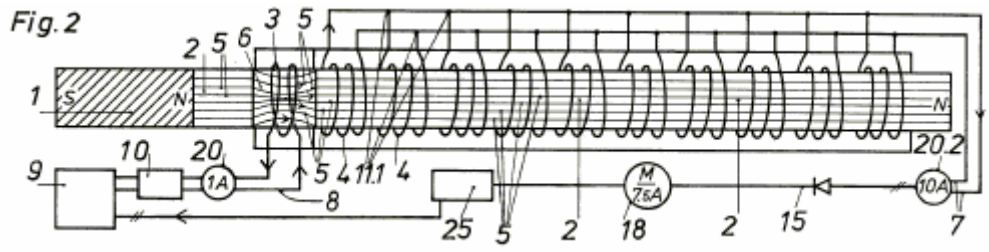


Fig.3 represents a linear energy generator at the moment of the interruption by the induced flux;

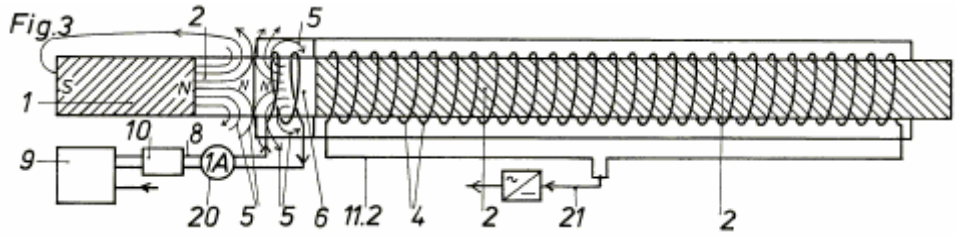


Fig.4 explains a static energy pulse generator with a closed magnetic circuit at the moment the transmission of the induction flux of the permanent magnet to the generator core block,

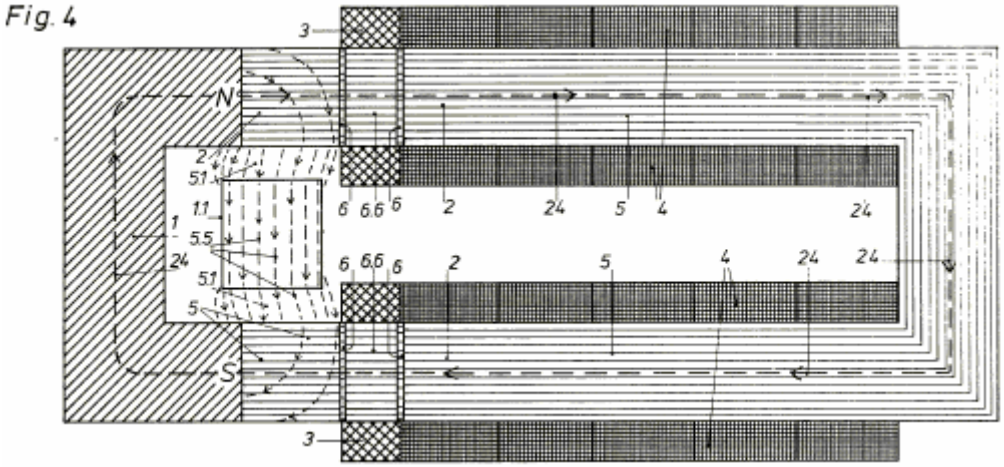


Fig.5 is a schematic diagram of the functional way in which the feedback process operates in a system according to the present invention,

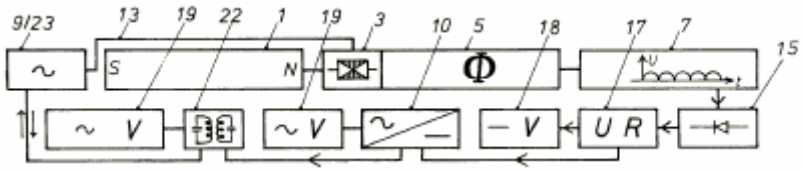


Fig. 5

Fig.6 shows a double linear energy generator with some of its sub-elements and

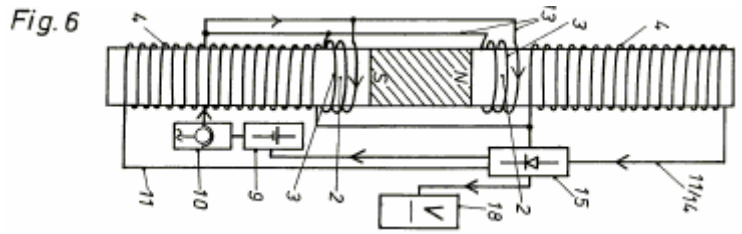
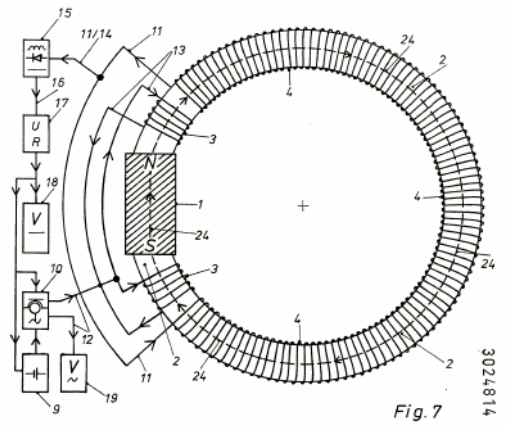
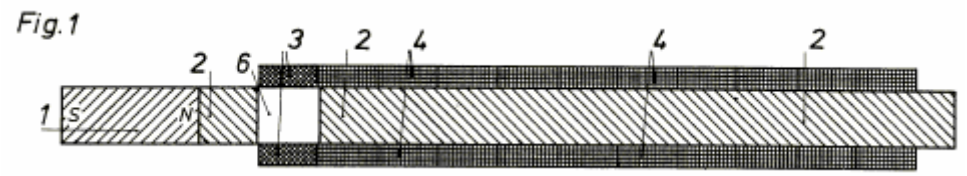


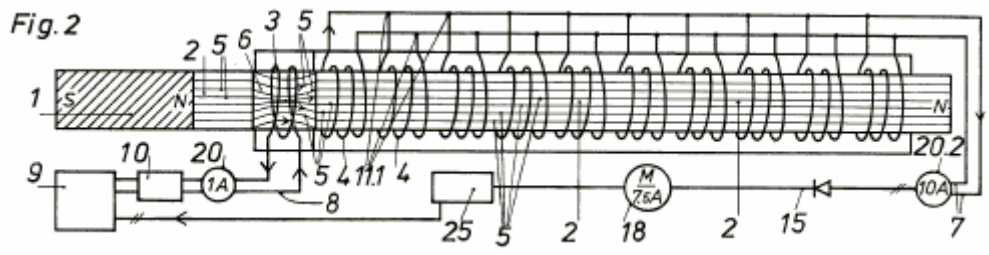
Fig.7 shows an energy generator, according to the present invention, with cyclic pulsing operation and with some of its sub-elements in and at the energy circuit.





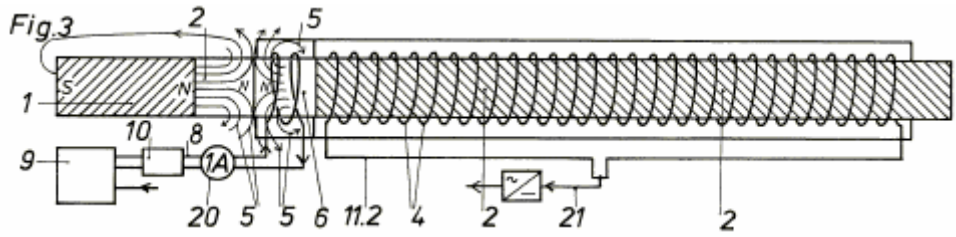
The linear energy generator shown in cross-section in **Fig.1**, consists of a permanent bar magnet **1** with a soft-iron magnetic generator core **2**, which can be a single piece, or, as shown here, divided into two sections. The magnetic flux modification coil is not directly attached to the permanent magnet **1**, so that the permanent magnet **1** is not subjected to the alternating fields produced by the magnetic flux modification coil **3**.

On the generator core **2**, following the magnetic flux modification coil **3**, there are several output coils **4**. An air gap **6** serves as a gate or circuit-breaker for the magnetic flux coming from the permanent magnet **1** and the magnetic induction flux for the induction windings **4**.



With this arrangement, an alternating current applied to the magnetic flux modification coil **3** is used to produce an alternating magnetic field in the air-gap **6**, so that, as shown clearly in **Fig.2**, with each phase of the alternating current the induced magnetic flux **5** is directed first to the core **2** and then against the permanent magnet **1**, as

shown in **Fig.3**, causing the magnetic flux **5** induced in core **2** by the permanent magnet **1**, is interrupted totally or partially, and thus experiences a modification which varies with time.



If an alternating current is supplied to the magnetic flux modification coil **3** e.g. with a frequency by 50 Hz, then the induced flux **5** in core **2** experiences one hundred modifications per second, inducing in the induction windings **4** a pulsating direct current **14** which has 50 positive maximum values per second.

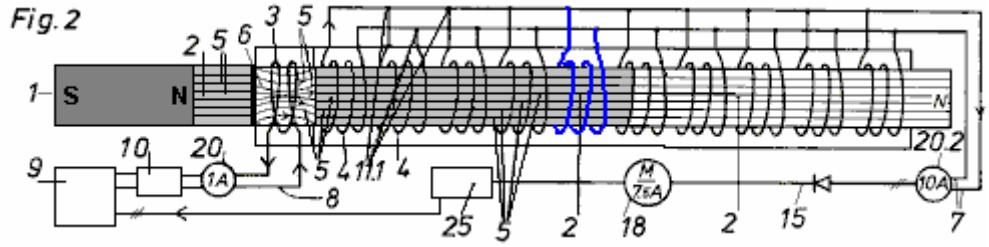
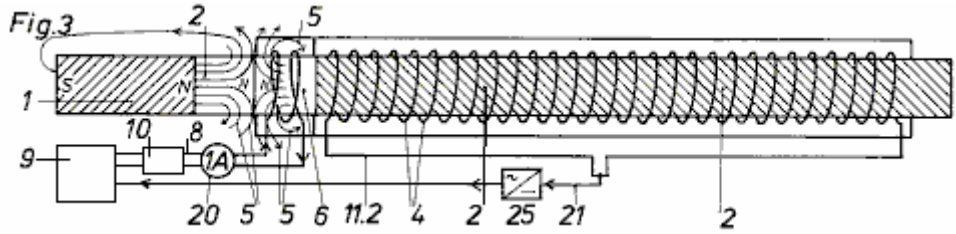


Fig.2 illustrates that on the generator core **2** several induction windings **4** are wound, **which correspond to the number of turns with same wire diameter as used in the magnetic flux modification coil 3**.

The permanent magnet **1** requires no electric current to produce its magnetic flux and nevertheless has at its North Pole end **N** of the core **2** the same magnetic saturation, as produced by the multiple number of turns of induction coil **4** of **Fig.2** or in a continuous coil **4** as shown in **Fig.3**, receives a multiple version of the input current which is needed for the excitation of the magnetic flux modification coil **3**.

From the power source **9** the energising current flows to the pulse generator **10**, the 1-amp ammeter **20** displays the intensity of the electric current. The induced current **7** or the pulsating direct current **14** is added via connections **11,1** and is measured by the 10-amp ammeter **20,2**. The rectifier **15** (not shown) produces pulsating Direct Current which is smoothed and supplied as the DC output **18**, through wire **21** to the battery charger **25**, which provides the input for the power source **9**.



The interruption of the magnetic flux by the direction change of the Alternating Current, necessary for the repeated modification of the induction flux **5**, is shown in **Fig.3**. When the induction flux **5** is interrupted, the wire coil(s) **11.2** is negative in that instant. Wire **21** provides the connection to the output power (not shown) and the input power source **9**.

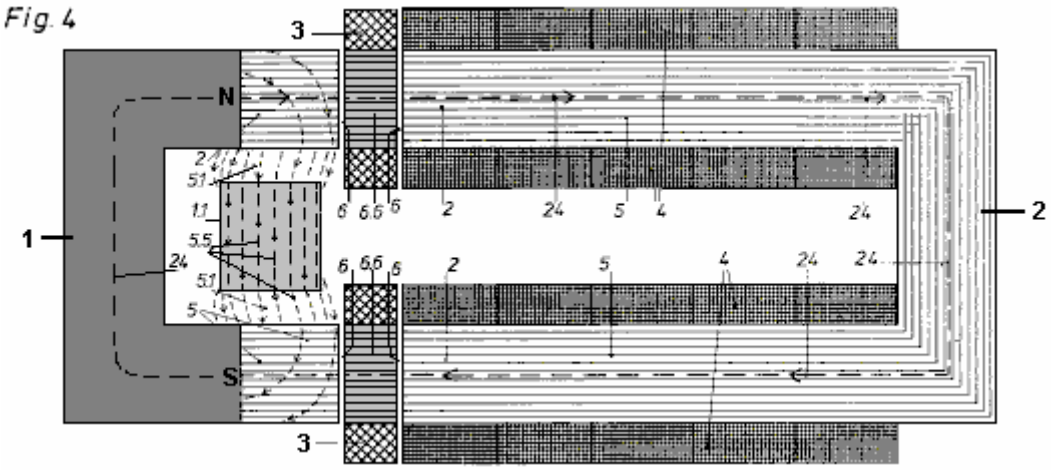


Fig.4 shows an embodiment of the above design which uses a U-shaped permanent magnet **1** and a U-shaped generator core **2** which has its two ends facing towards the poles of the permanent magnet **1**. Wound around two narrow air gaps **6** and its own narrow core, is a magnetic flux modification coil **6.6**.

Fig.4 shows the instantaneous state, when the magnetic flux **5** is transferred from the magnetic flux modification coil **3** into the generator core **2** forming a closed magnetic circuit **24**. The magnetic flux modification coil **3** here has a core **6.6**, which alternately makes and interrupts the passage of the magnetic flux **5** between two narrow air gaps **6** from the permanent magnet **1** to the generator core **2** which has the induction windings **4**, so that each pulse of the induction flux **5** induces a current in the induction windings **4**. Thus a pulsating Direct Current results which is several times larger than the input energising current.

When the direction of current flow in the soft-iron magnetic core **6.6** of the magnetic flux modification coil **3** is changed, then at that time, the magnetic flux **5** of the permanent magnet **1** flows over into the iron guidance block **1.1** and flows along the paths **5.5** and **5.1** to the South-pole of the permanent magnet **1** or to the equilibrium in the air gaps between the North-pole of iron guidance block **1.1** and the South-pole of the permanent magnet **1**. The dashed line **5.5** from North-pole to South-pole through the iron guidance block **1,1** depict the magnetic flow when it is prevented from travelling through the generator core **2**.

This steering of the magnetic flux **5** prevents leakage flux entering the generator core **2** and so the maximum RMS of the induced current is achieved, as the generator core **2** is without magnetic excitation.

PJK note: I have difficulty in accepting this as the excitation coil would appear to produce an equivalent magnetic flux into the U-shaped frame **2** as it blocks the permanent magnet flux passing across the air gaps between magnet **1** and electromagnets **3**, and to make things worse, magnetic flux flows about a thousand times more easily through soft iron than through air. However, we do know that modulating the flux from a permanent magnet with the magnetic field from a coil is very effective in producing COP>1 as has been shown by the independent replication of Lawrence Tseung's magnetic frame covered earlier in this chapter.

The following schematic diagram **Fig.5**, shows the sequence in the cyclic process e.g. in an energy generator according to **Fig.4**.

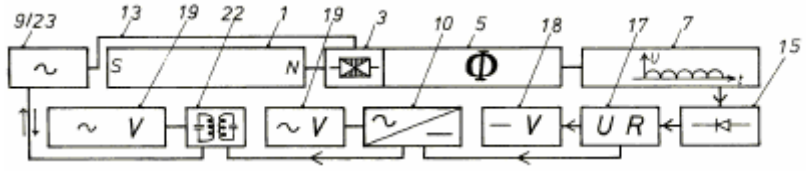


Fig. 5

The pulsating current from the power source **9** or Alternating Current **12** from the power-grid **23** flows through wire **13** to the magnetic flux modification coil **3** and produces a pulsating induction current **7** or pulsating direct current **14**, which is converted by the rectifier **15** to smooth Direct Current **16** which is passed to the voltage regulator **17**, and then on as Direct Current **16** now at the desired voltage, on to the DC output **18** and to the current transformer **10**, by which the received Alternating Current **12** is led to the AC output **19** and coupled by AC link **22** with the power-grid input **23**, whereby the AC output **19** can be supplied with current from the power-grid or from the electric output of the energy generator.

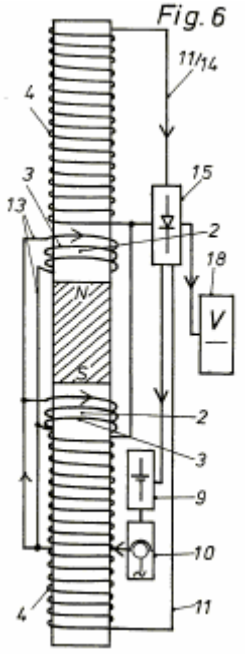
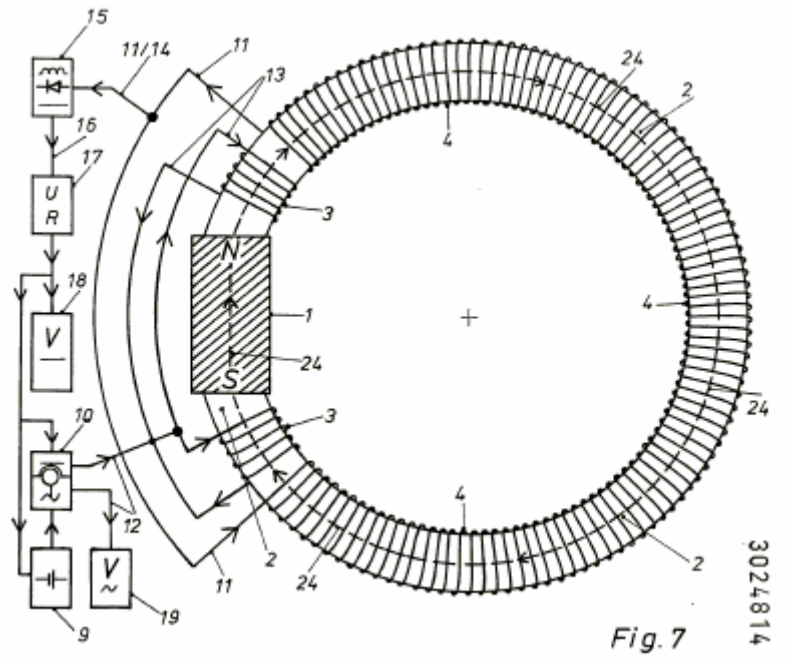


Fig.6 shows a double linear generator in accordance with the invention. On a preferably straight-line permanent magnet are fitted generator cores **2** of dynamo iron to its two poles. The magnetic flux modification coil **3** gets its working current from the power source **9** through a current transformer or from pulse generators **10** fed by the excitation circuit **13**. Through appropriate windings **4** an output of either Direct Current or Alternating Current can be obtained.

Pulsating Direct Current **14** from the induction current circuit **11** is smoothed and passed to the DC output **18** and to the input power source **9**.



A further version of the system in accordance with the invention is shown in **Fig.7**. This is also a motionless energy producer, although it has a cyclic arrangement and function. In this generator there are no mobile sections such as a rotor and the induction flux **5**, like the induction current **7**, develops from a cyclically pulsating induction flux.

The permanent magnet **1** is inserted into a circular generator core **2**. The magnetic flux modification coils **3** can be operated with pulsating Direct Current **14** or as here with Alternating Current **12**. For example, the Direct Current **16** from the power source **9** is led into a current transformer **10** converted into Alternating Current and fed into the exciter circuit **13**.

The magnetic flux modification coils are created so that the positive maximum value of the Alternating Current **12** opens and supports the natural flow of the permanent magnet flux **5** passing from the North-pole to the South-pole through the circular generator core **2** to form a closed magnetic circuit **24**.

If the magnetic flux modification coils **3** on both sides of the permanent magnet **1** carry the maximum negative value of the Alternating Current **12**, then the natural magnetic flux in the generator core **2** is constricted by induction flux moving in the opposing direction in the magnetic flux modification coils **3** and this interrupts the magnetic flux **5** either totally or partially.

In the case of temporally large modification of this sequence in the coil **4** a pulsating direct current **14** is caused, which is led by the induction circuit **11** to the electric rectifier **15**, in which the pulsating direct current **14** is reduced to a smooth DC current. The Direct Current **16** can be passed to the DC output **18** and the input power supply **9** and the current transformer **10** which supplies the AC output **19** and AC to the magnetic flux modification coils **3**.

If the induction flux **5** is restricted in the core **2** by the magnetic flux modification coils **3**, then an Alternating Current with a smaller negative maximum value is produced and the arithmetic average value during one period shifts down to zero.

According to this invention in a cyclic process an energy circle is created, with considerable energy surplus for the supply of the various outputs as well as for the maintenance of the operation of this system.

According to this invention, by avoiding the need for torque, the same induction effect is obtained by the magnetic flux modification coils and the use of permanent magnets in energy generation, as with the conventional generators which use the transformation of torque into electricity, whereby however, the energy value of the input torque is greater than the energy value of the electricity this generated.

It was found that from each pole of the permanent magnet to the two ends of a U-shaped generator core one of the magnetic flux modification or guidance coils with or without core for transmission of the magnetic induction flux has to be designated in such a manner that by alternating induction flux, which is caused by the permanent magnet e.g. in the rhythm of the phase change of an alternating current frequency of the energising current the generator core is constantly commutated, as the North-pole will be transferred alternating to one and the other open end of the core and the coils likewise leading to the core to s-pole of the permanent magnet close the reversible magnetic circuit in the core with each current pulse, which is caused by a permanent magnet.

In this way the induction flux in the core experiences its desired flux direction change and produces in the output windings of the generator an Alternating Current of the same frequency as those of the energising current, however with identical frequency to that of the input energising current.

Since the pulsating or reversible induction flux is caused by a permanent magnet, no electric current is necessary for its production also for the whole length of the generator core and its output windings, because the reversible magnetic excitation of the core takes place indirectly in each case or directly via a permanent magnet, whose remanence is changed by the magnetic excitation of the generator core in accordance with the invention.

The system according to invention for energy production and energy producer can be extremely efficient e.g. in high frequency operation with electronically controlled direct current pulse operation and it can output a multiple of the necessary input current and electricity be produced in this way, is produced without material being used up and without a thermal circuit or a torque being necessary.

If several of these generators are cascaded in graduated increasing size e.g. in a series where the second generator receives the full output from first and third generator receives the full output of the second, then with a power multiplier of 10 for each generator, the sixth generator in the chain will have a 1000 MW performance if there is a 1000 W power input to the first (and smallest) generator at the start of the series.

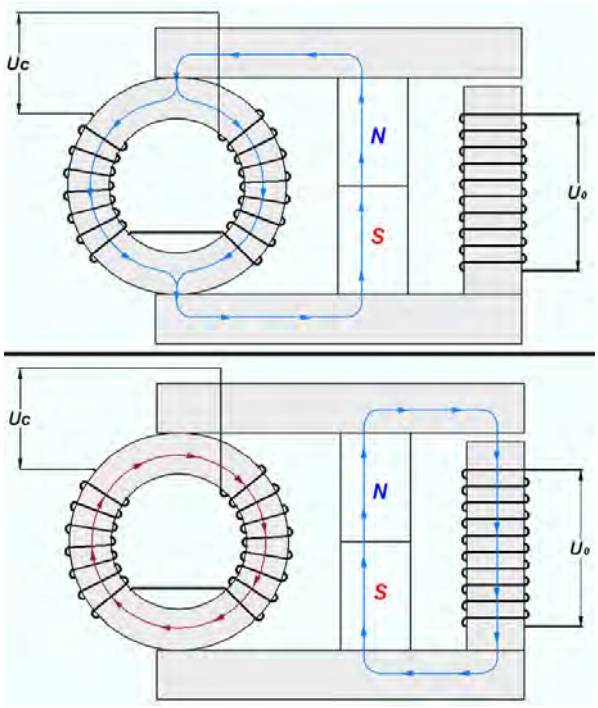
Thus it is possible to replace all well-known primary energies and procedures of the energy conversion with the systems and energy generators according to invention, for all future on economical into electricity because of their high costs, since these cannot operate in any way remotely as economically, as it is possible with the devices of this invention.

Valeri Ivanov's Motionless Generator.

There are other devices which have what appears to be a very important air-gap in a magnetic frame. One of these was displayed on a Bulgarian website and is on the web page which is located at http://www.inkomp-delta.com/page3.html, put up by Valeri Ivanov in 2007.

Valeri lives in Elin Pelin, Bulgaria and his motionless generator has COP=2.4 performance. Videos: http://www.youtube.com/watch?v=7IP-buFHKKU and http://www.youtube.com/watch?v=npFVaeSbk1Q are for his design, and it appears that he is about to go commercial: http://www.inkomp-delta.com/index.html in May 2014.

It is shown that an effective device can be constructed from a permanent magnet, a toroid and a laminated iron yoke. The arrangement is displayed like this:



When the input coil is pulsed with an input voltage, it causes a flux reversal in the frame around which the output coil is wound, generating an electrical output.

There is another forum related to this and the better known MEG of Tom Bearden's which can be found at http://tech.groups.yahoo.com/group/MEG_builders/message/1355 where that particular message states that Valeri's device can be made to work at frequencies as low as 50 Hz and can use standard laminated iron frame components and produces Coefficient Of Performance figures up to 5.4 (that is, the output power is more than five times the input power). A demo video is at http://inkomp-delta.com/page10.html but not in English. It may well be that in order to work well, the MEG needs a very narrow input coil with an air-gap on each side of it and the same may well apply to Lawrence Tseung's magnetic frame shown earlier in this chapter.

The Motionless Generators of Kelichiro Asaoka

Kelichiro Asaoka received US patent 5,926,083 some two years before the well-known MEG patent of Tom Bearden and his associates. Personally, I find it hard to understand how the MEG patent (in the Appendix) could have been awarded when the Asaoka patent was already in place. However, here is most of the content of the Asaoka patent:

US Patent 5,926,083 20 July 1999 Inventor: Kelichiro Asaoka

Static magnet dynamo for generating electromotive force based on changing flux density of an open magnetic path

ABSTRACT

A static magnet dynamo including at least one permanent magnet having different poles; a first core comprising a soft magnetic material and which couples the different poles of the permanent magnet to form a closed magnetic path; a second core comprising a soft magnetic material which couples to the closed magnetic path via a

paramagnetic material to form an open magnetic path; a magnetised coil wound around a portion of the first core where the closed magnetic path is formed; and an induction coil wound around a portion of the second core. A direction of a flux of the closed magnetic path is changed by applying an alternating voltage to the magnetised coil, generating an electromotive force in the induction coil by electromagnetic induction due changes in a flux of the open magnetic path induced by the change in direction of the flux of the closed magnetic path.

FIELD OF THE INVENTION

This invention relates to a dynamo which generates electromotive force by electromagnetic induction by changing the flux passing through an induction coil. More particularly, this invention relates to a static magnet dynamo that changes magnets that pass through an induction coil without turning the armature or electromagnet.

DISCUSSION OF BACKGROUND

Dynamos currently in practical use are so designed as to generate electromotive force by electromagnetic induction by changing the flux passing through an induction coil. Dynamos that generate power this way come in a wide variety, ranging from large models used in hydroelectric, thermal or atomic power plants to small models such as small dynamos with a diesel engine.

In all dynamo models mentioned above, the armature and electromagnet are turned, to change the flux passing through the induction coil, thus generating electromotive force in the induction coil by electromagnetic induction. For examples, the armature and electromagnet are turned by the torque of a water turbine in hydroelectric power generation, by the torque of the steam turbine in thermal and atomic power generation, and by the torque of the diesel engine in small dynamos.

Disadvantages:

Dynamos that generate electromotive force by electromagnetic induction as mentioned above are so designed that, regardless of the size of the dynamo, the armature and electromagnet are turned in order to change the flux passing the induction coil. These dynamos are disadvantageous in that the said turning of the armature and electromagnet generates vibrations and noise.

OBJECTS OF THE INVENTION

The purpose of this invention is to provide a static magnet dynamo devoid of any torque-giving means or other moving part to eliminate vibrations and noise, in order to resolve the various problems mentioned above.

To resolve the above problems, this invention is composed as described below.

The static magnet dynamo involved in this invention consists of at least one permanent magnet, a first core consisting of a soft magnetic material forming a closed magnetic path by coupling the different poles of the said permanent magnet, a second core consisting of a soft magnetic material forming an open magnetic path by being coupled to the closed magnetic path via a paramagnetic material, a magnetised coil wound around a part consisting of only the closed magnetic path of the first core, and an induction coil wound around the second core. The point of this invention is to generate electromotive force in the induction coil by electromagnetic induction by changing the direction of the flux of the closed magnetic path by applying an alternating voltage to the magnetised coil and by changing the flux of the open magnetic path induced by changes in the direction of the flux in the closed magnetic path.

Effects:

In the above configuration, the static magnet dynamo involved in this invention consists of a first core consisting of a permanent magnet and a closed magnetic path, a second core consisting of an open magnetic path via a paramagnetic material, a magnetised coil wound around the part consisting only of the closed magnetic path of the first core, and an induction coil wound around the second magnetic path. The dynamo is so designed as to generate electromotive force in the induction coil by electromagnetic force by changing the direction of the flux of the first core by applying an alternating voltage to the magnetised coil, and by changing the flux of the second core induced by changes in the direction of the flux of the first core.

This makes it possible to change the flux passing through the induction coil without a torque-giving means or other moving part and to generate electromotive force in the induction coil by electromagnetic induction, thus enabling power generation without causing vibrations or noise. This dynamo can also be downsized and made available at low prices.

Other characteristics and benefits of this invention will be made clear by the description given below with diagrams attached.

BRIEF DESCRIPTION OF THE DRAWINGS

A more complete appreciation of the invention and many of the attendant advantages thereof will be readily obtained as the same becomes better understood by reference to the following detailed description when considered in connection with the accompanying drawings, wherein:

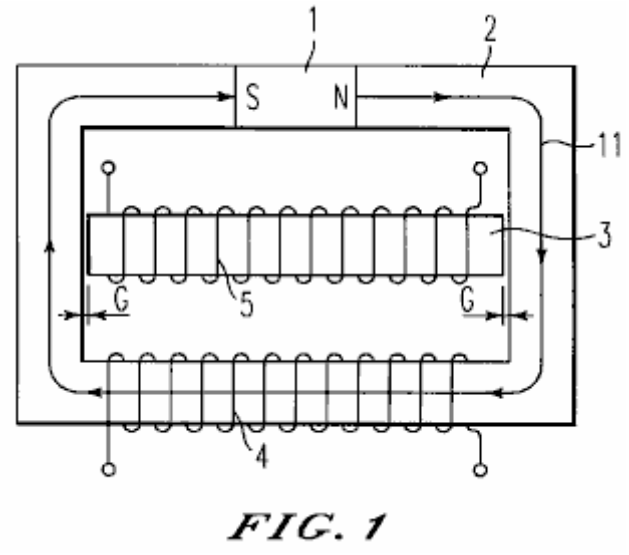


FIG. 1

Fig.1 represents a basic configuration of a static magnet dynamo with an open magnetic path involved in this invention.

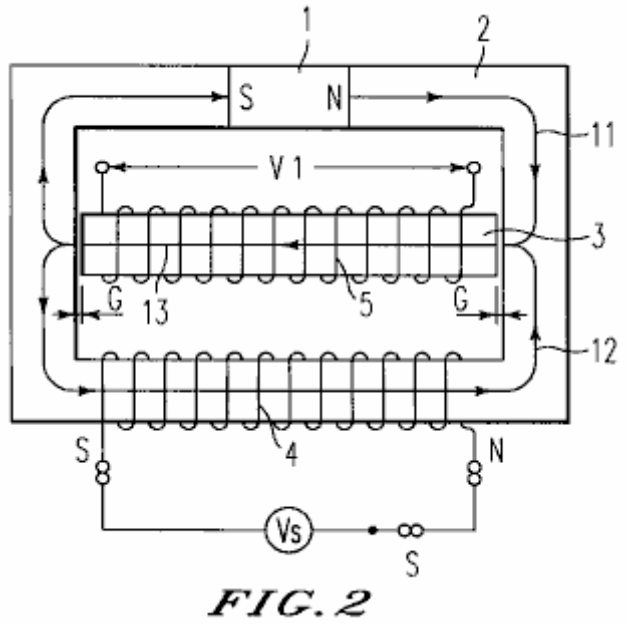


FIG. 2

Fig.2 represents how a flux in the direction opposed to that of a permanent magnet typically occurs in the magnetised coil.

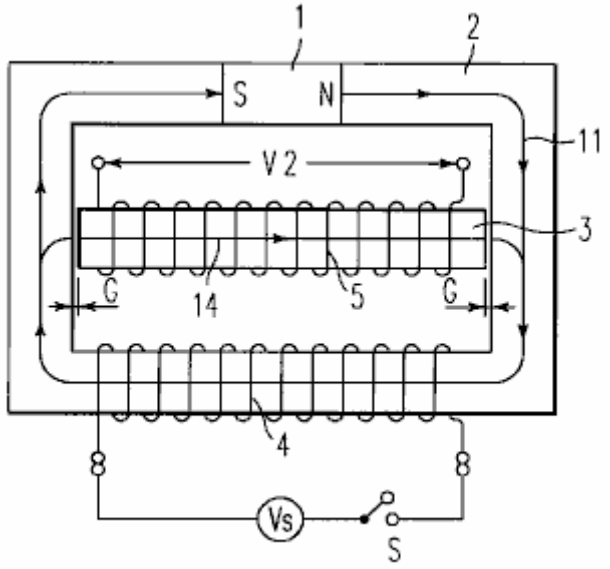


FIG. 3

Fig.3 represents how a flux in the direction opposed to that of a permanent magnet typically disappears from the magnetised coil.

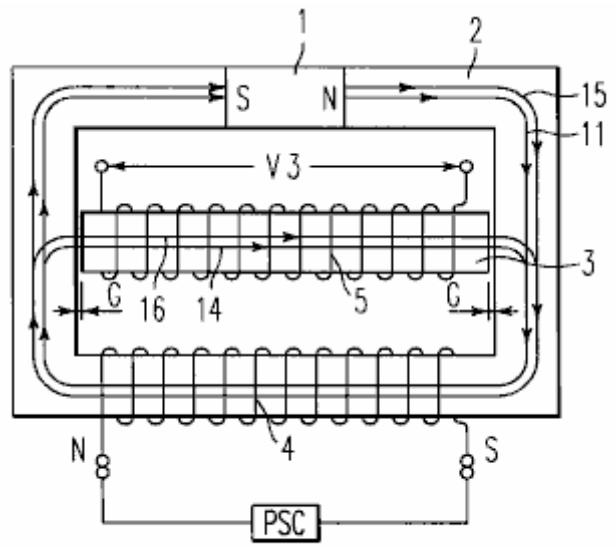


FIG. 4

Fig.4 represents how a flux in the same direction as that of the permanent magnet typically occurs in the magnetised coil.

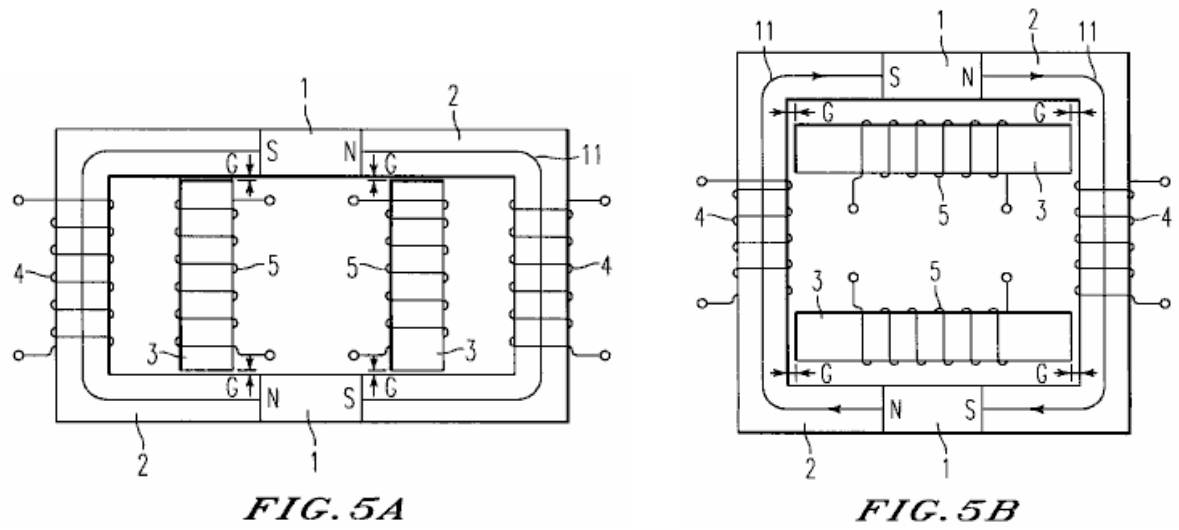


Fig.5 is a first embodiment of the static magnet dynamo involved in this invention.

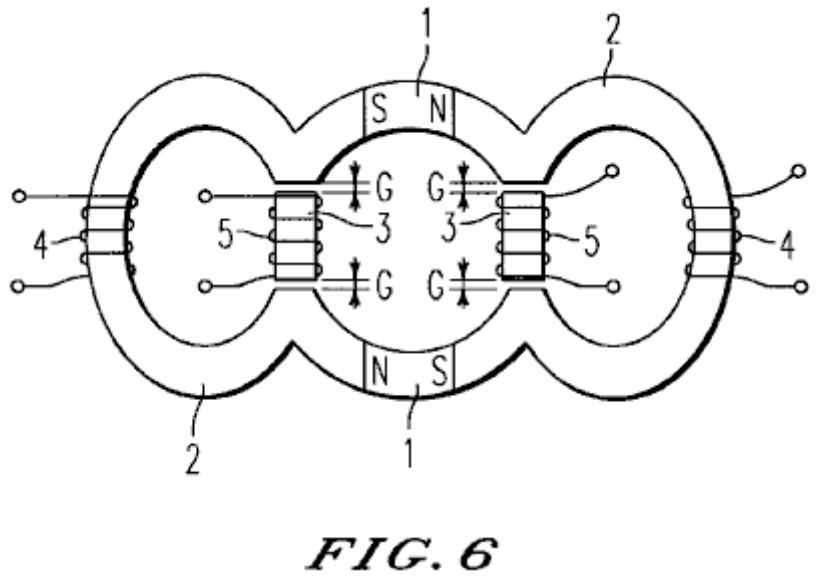


Fig.6 is a second embodiment of the static magnet dynamo involved in this invention.

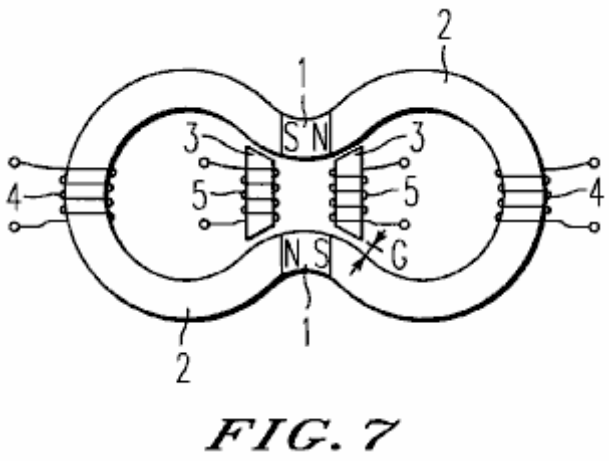


Fig.7 is a third embodiment of the static magnet dynamo involved in this invention.

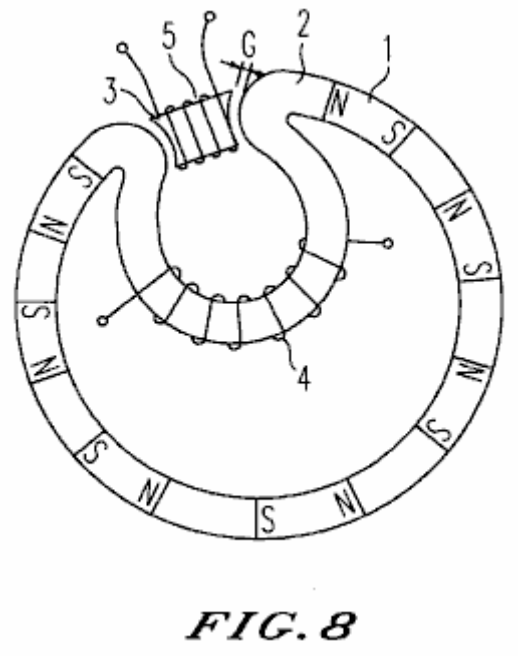


FIG. 8

Fig.8 is a fourth embodiment of the static magnet dynamo involved in this invention.

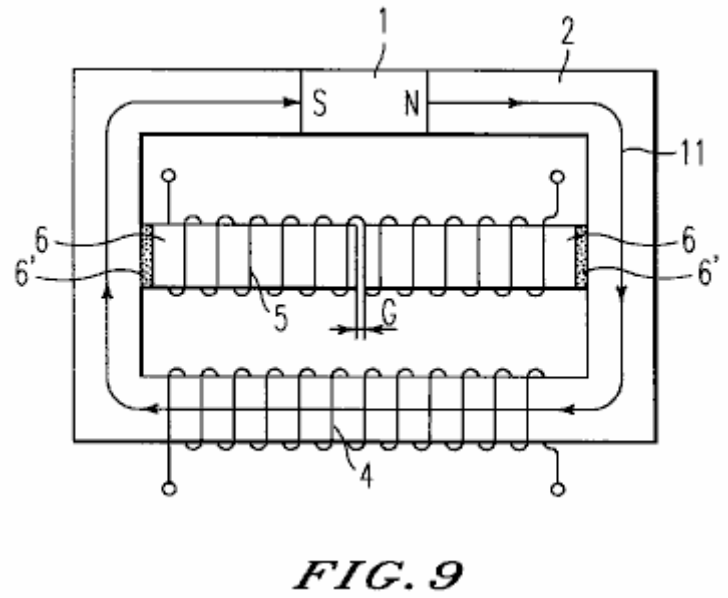


FIG. 9

FIG. 9 is a fifth embodiment with an open magnetic path.

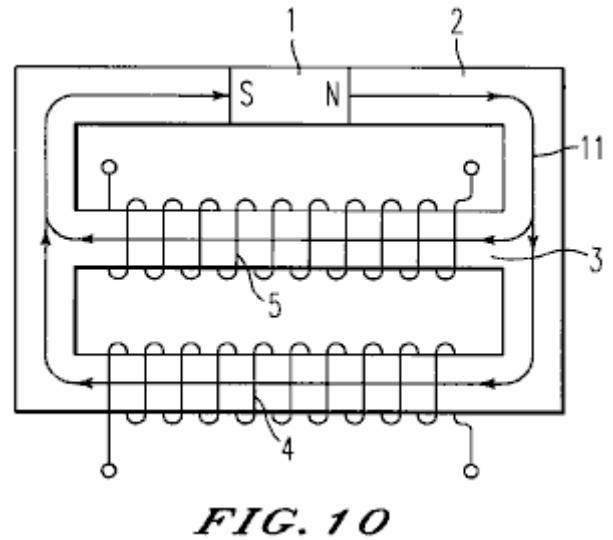


FIG. 10

Fig.10 is a basic configuration of a static magnet dynamo with a closed magnetic path involved in this invention.

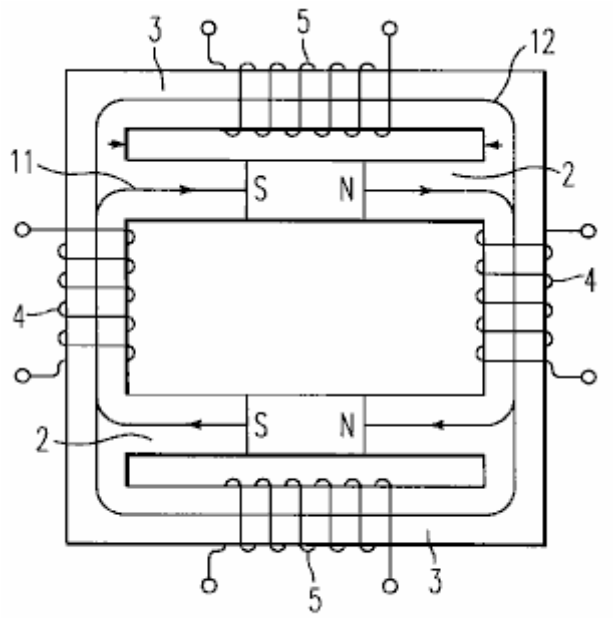


FIG. 11

Fig.11 is a first embodiment of the static magnet dynamo with a closed magnetic path involved in this invention.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

Referring now to the drawings, wherein like reference numerals designate identical or corresponding parts throughout the several views, and more particularly to **Fig.1** where there is illustrated a basic configuration of the static magnet dynamo with one permanent magnet. **Figs. 2, 3**, and **4** describe how the static magnet dynamo represented in **Fig.1** generates power.

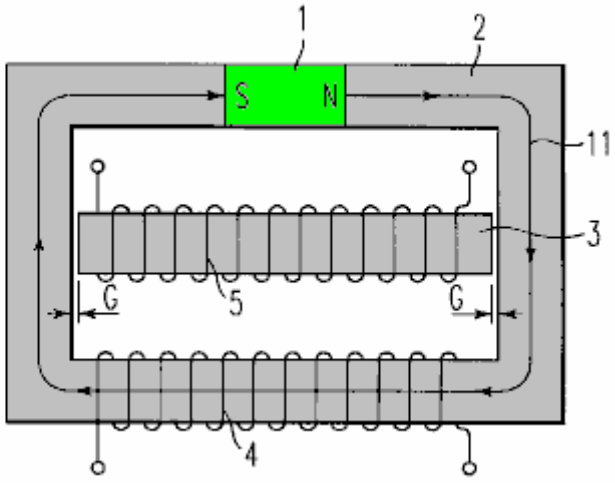


FIG. 1

As indicated in the figures, the first core **2** formed to couple the permanent magnet **1** and the different poles of the permanent magnet **1** in an annular manner, forms a closed magnetic path. This closed magnetic path is then equipped with a second core **3** via a paramagnetic material 10 μm to 5 mm thick. This results in the formation of an open magnetic path consisting of a permanent magnet **1**, part of a first core **2**, a paramagnetic material, and a second core **3**. The part consisting only of the closed magnetic path of the first core **2** is wound around with a magnetised coil **4**. The second core **3** is then wound around with an induction coil **5** designed to generate electromotive force by electromagnetic induction.

Here, the permanent magnet **1** is a magnet with a high residual flux density, a great coercive force, and a large maximum energy product for higher power generation efficiency. Typical materials used here are neodymium iron boride magnet ($Nd_2Fe_{14}B$), samarium cobalt magnet (Sm_2Co_{17}), or samarium iron nitride ($Sm_2Fe_{17}N_2$).

The first core **2** and the second core **3** are made of a soft magnetic material having a high permeability, with high initial, maximum, and other permeability levels, high residual flux density and saturation magnetisation, and small coercive force, thus making effective use of the flux of the magnetic path for power generation. Examples include Permalloy based alloys.

Applicable paramagnetic materials are those with a specific permeability comparable to that of a vacuum, such as air, copper, and aluminium. When air is used as a paramagnetic material, that is, when a gap **G** is secured between the first core **2** and the second core **3**, the second core **3** is retained with a solid paramagnetic material. The figures represent embodiments with a gap **G**, without a solid paramagnetic material designed to retain the second core **3**.

Following is a description of how a static magnet dynamo of the above configuration generates power. First, when no voltage is applied to the magnetised coil **4** of the static magnet dynamo, a first flux **11** is formed in the first core **2** in the direction going from the N pole to the S pole of the permanent magnet **1**. In this state, no flux has been formed in the second core 3 coupled via the gap **G**.

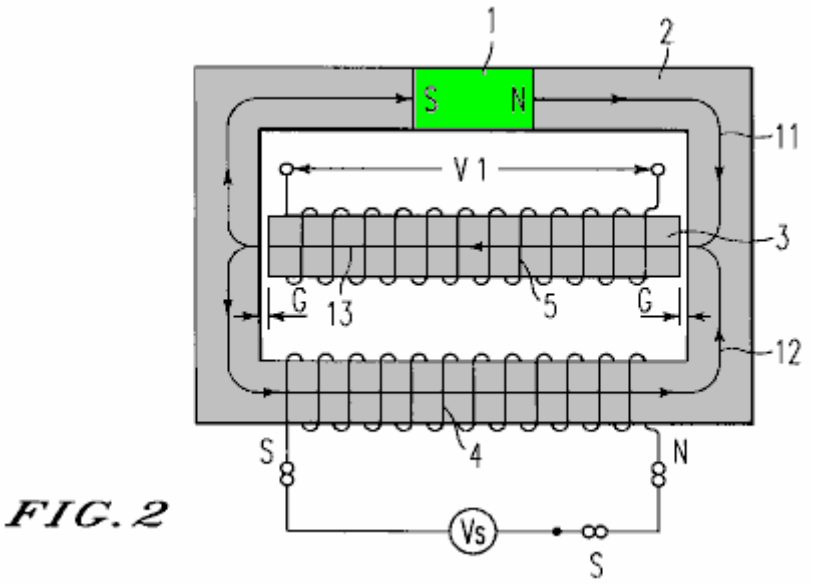


FIG. 2

A voltage can be applied to the magnetised coil **4** in three manners described below. In the first voltage application, as indicated in **Fig.2**, a DC voltage **VS** is applied to the magnetised coil **4** in the direction that the voltage repels the first flux **11** of the first core **2** generated by the permanent magnet **1**, and vice versa, that is, in such a manner that the second flux **12** occurs in the reverse direction of the first flux **11**. As a result, the first flux **11** repels the second flux **12** and vice versa, so that the flux more easily leaks from the closed magnetic path. The first flux **11** and the second flux **12**, which more easily leak from the closed magnetic path, jump across the gap **G** and enter the second core **3**, so that a third flux **13** is induced in the second core **3**. Furthermore, the induction of this third flux **13** changes the flux passing through the induction coil **5**, so that electromotive force **V1** occurs in the induction coil **5**, resulting in power being generated.

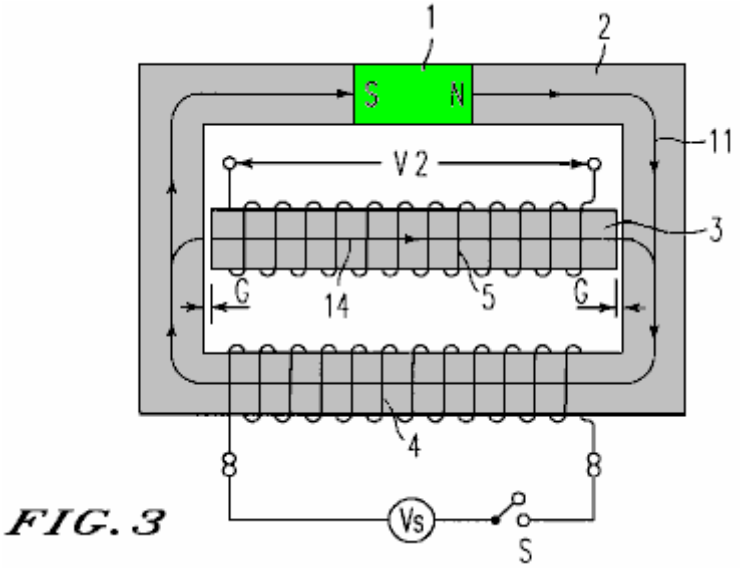


FIG. 3

Next, removing the DC voltage applied to the magnetised coil **4** prompts the first core **2** to try going back to a state where only the first flux **11** is formed as indicated in **Fig.1**. At that time, the second core **3** has a flux in the reverse direction of the third flux **13**, that is, the fourth flux **14** indicated in **Fig.3**, in order to kill the third flux **13**. Then, the induction of the fourth flux **14** changes the flux passing the induction coil **5**, so that electromotive force **V2** occurs in the induction coil **5**, resulting in power being generated.

Power generation in this first voltage application can be realised by a static magnet dynamo involved in this invention, a DC power supply to apply a DC voltage **VS** to the magnetised coil **4**, and a switching circuit that turns the DC power supply on and off. A contact-less switching circuit can be made if a semiconductor switching device, such as a thyristor, is available.

The second voltage application is the same as the first voltage application up to the point where the third flux **13** is induced in the second core **3** by applying a DC voltage **VS** to the magnetised coil **4** so as to generate the second flux **12** in the reverse direction of the first flux **11** and where the third flux **13** is induced to generate electromotive force **V1** in the induction coil **5**, thus generating power.

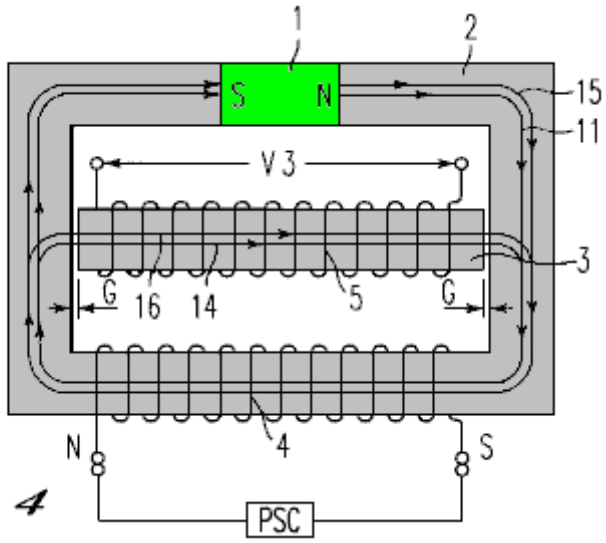


FIG. 4

Next, changing the polarity of the DC voltage applied to the magnetised coil **4** generates in the first core **2** the first flux **11** caused by the permanent magnet **1**, as well as the fifth flux **15** in the same direction as the first flux, caused by the magnetised coil **4**. Here, the first flux **11** is given the fifth flux **15**, so that the second core **3** is given the fourth flux **14** as indicated in **Fig.4**, as well as the sixth flux **16** in the same direction as the fourth flux **14**. Furthermore, inducing the fourth flux **14** and the sixth flux **16** changes the flux passing through the induction coil **5**, so that an electromotive force **V3** larger than the electromotive force **V2** is generated in the rotary coil to produce power.

This second voltage application requires a polarity switching circuit **PSC** that changes the polarity of **DO** voltage instead of a switching circuit that turns on and off the DC voltage applied to the magnetised coil **4** in the first voltage application. This polarity switching circuit can be made of a semiconductor switching device, similarly to the switching circuit in the first voltage application.

In the third voltage application, AC voltage **VS** is applied to the magnetised coil **4** instead of applying DC voltage to the magnetised coil **4** in the second voltage application with the polarity changed. The flux generated by applying AC voltage to the magnetised coil **4** becomes an alternating flux that alternates between the second flux **12** in **Fig.2** and the fifth flux **15** in. **Fig.4**. Then, the flux induced in the second core **3** is the third flux **13** in **Fig.2** when the second flux **12** is generated, and is the fourth flux **14** trying to kill the sixth flux **16** and the third flux **19** in **Fig.4** when the fifth flux **15** is generated. That is, the flux induced in the second core **3** naturally also becomes an alternating flux.

In power generation of this third voltage application, AC voltage is applied to the magnetised coil **4**, which overcomes the need for a switching circuit or polarity switching circuit **PSC**, which was needed in the first and the second voltage application, so that the device becomes simplified. Furthermore, the flux induced in the first core **2** and the second core **3** becomes an alternating flux induced by AC voltage, so that the dynamo functions also as a transformer having a gap **G** between the first core **2** and the second core **3**. It is therefore possible to increase further the electromotive force **V** generated by electromagnetic induction in the induction coil **5**.

Next, the power generation efficiency of a static magnet dynamo involved in this invention is described. The static magnet dynamo can be considered as a transformer if its permanent magnet **1** is removed and there is a gap **G**.

A transformer entails an eddy current loss **We** and hysteresis loss **Wh** of the core, and a loss **Wr** due to the electric resistance of the coil. These factors are in a relation formulated below.

Total loss **W1 = We + Wh + Wr** . . . . . .(1)

Let the input be **Win** and the output **Wo**, and the **Win** becomes equal to the total loss, so that the conversion efficiency of the transformer is

Eff = Wo / Win = Wo(We + Wh + Wr) < 1 . . . . . (2)

In reality, in **Fig.1**, the closed magnetic path consisting of the first core **2** contains a permanent magnet **1**. The flux of this permanent magnet **1** therefore contributes to power generation. Therefore, in **Fig.1**, let the input be **Win2** and the output **Wo2**, then

Wo2 = Wp + αWin2 . . . . . (3)

Where **Wp** represents power resulting from the flux of the permanent magnet **1** contributing to power generation, and **α** represents a conversion efficiency obtained when the device is considered as a transformer with a gap **G**.

Therefore, power generation efficiency is:

Eff = Wo2 / Win2 or .

Eff = (Wp / Win2) + α . . . . . (4)

Here, since **α** < 1, if **Wp / Win2** > 1, that is, if power obtained resulting from the flux of the permanent magnet **1** contributing to power generation is larger than dynamo power supplied to the magnetised coil **4**, power generation efficiency becomes no less than 1, so that the device can display its performance as a dynamo.

Thus, the inventor examined as described below how much the flux of the permanent magnet **1** contributes to the induction of the third flux **13** in **Fig.2**. First, the inventor provided static magnet dynamos of the basic configuration indicated in **Fig.1**, one with a permanent magnet **1** and another without a permanent magnet **1**. The inventor then compared the power levels needed to induce fluxes of equal flux densities to the second core **3** of each embodiment, that is, the power levels supplied to the magnetised coil **4**. As a result, an embodiment with a permanent magnet **1** required only a very low power level to be supplied to the magnetised coil **4**. It was observed that the power level required was no more than one fortieth of that of the embodiment without a permanent magnet **1**, depending on the test condition.

In a static magnet dynamo involved in this invention, therefore, **Win2** can be made sufficiently smaller than **Wp**, so that the inventor considers it possible to make **Wp / Win2** > 1.

Embodiment 1

Next, as the first embodiment, a static magnet dynamo system composed of two static magnet dynamos of the basic configuration is described based on **Fig.5**.

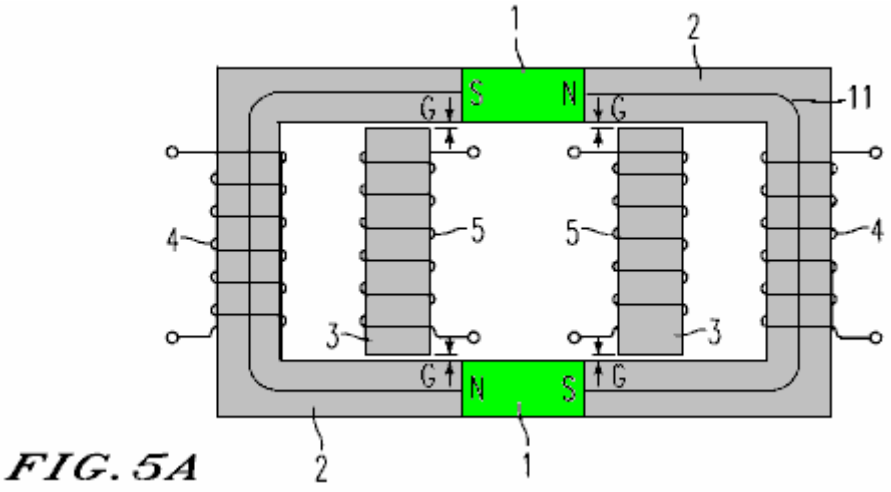


FIG. 5A

In **Fig.5A**, in a static magnet dynamo, a closed magnetic path is made of two permanent magnets **1** and two first cores **2** formed so as to couple the different poles of one permanent magnet **1** with the other permanent magnet **1** in an annular manner. This closed magnetic path is then equipped with a second core **3** via a gap **G**. This forms an open magnetic path consisting of a permanent magnet **1**, part of a first core **2**, a paramagnetic material, and a second core **3**.

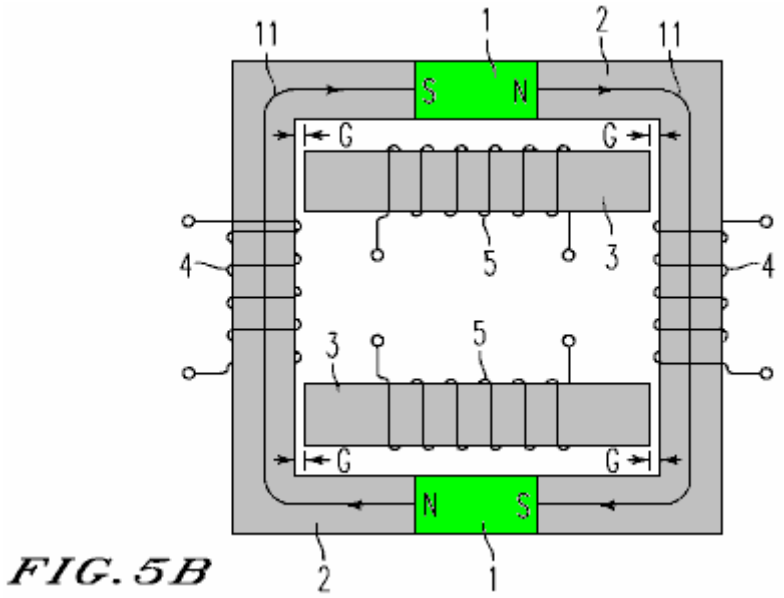


FIG. 5B

This open magnetic path can be arranged in two different ways. In one configuration, as indicated in **Fig.5A**, one open magnetic path can be made of two permanent magnets **1** and two second cores **3**. In the other configuration, as indicated in **Fig.5B**, one open magnetic path can be made of one permanent magnet **1** and another can be made of one first core **2**. The static magnet dynamos in **Fig.5A** and **Fig.5B** do not differ substantially in terms of effect results, except that their patterns forming such an open magnetic path differ.

The part forming only a closed magnetic path of each first core **2** is wound around with a magnetised coil **4**. Each second core **3** is then wound around with an induction coil **5** which generates electromotive force by electromagnetic induction.

This static magnet dynamo forms a first flux **11** in the first core **2** in the direction going from the N pole to the S pole of the permanent magnet **1**, with no voltage applied to the magnetised coil **4**. Furthermore, the action of this dynamo applying voltage to the magnetised coil **4** and generating electromotive force in the induction coil **5** by electromagnetic induction to generate power is similar to static magnet dynamos of the basic configuration. The static magnet dynamo with two permanent magnets **1** as mentioned above has well-balanced magnetic paths. Since the flux of the permanent magnets **1** can be effectively used, this embodiment achieves higher power generation efficiency than static magnet dynamos of the basic configuration.

The first embodiment is a static magnet dynamo system composed of two static magnet dynamos of the basic configuration. Similarly, a static magnet dynamo system can be made as a combination of three or more static magnet dynamos of the basic configuration (**Figs. 1-4**). In that case, similarly to the first embodiment, an open

magnetic path can be formed in two manners. One configuration is the formation of one open magnetic path by coupling all permanent magnets **1** with a second core **3**. The other is the formation of as many open magnetic paths as permanent magnets by coupling the N pole of each permanent magnet **1** to the S pole with a second core **3**.

Embodiment 2

Next, the second embodiment of the present invention is represented in **Fig.6**,

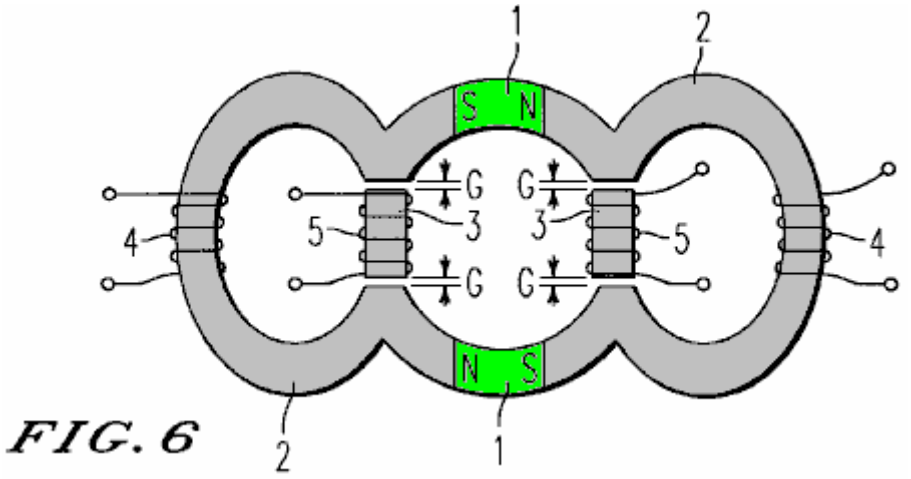


FIG. 6

the third embodiment in **Fig.7**,

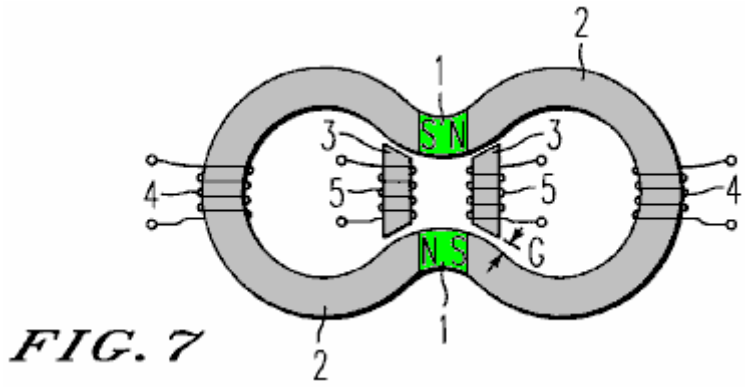


FIG. 7

and the fourth embodiment in **Fig.8**.

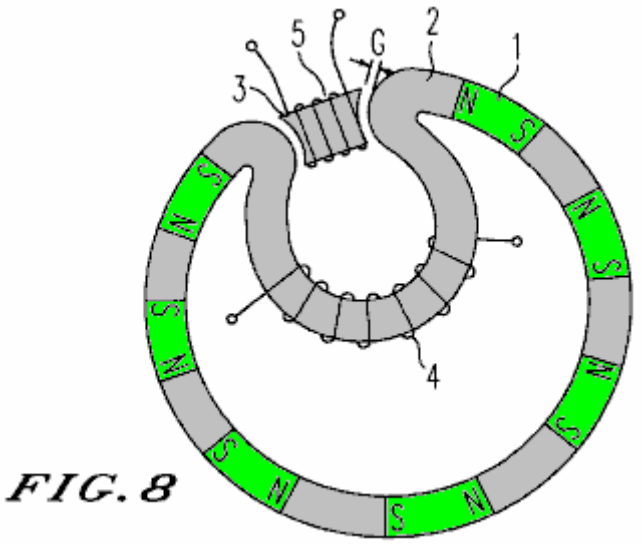


FIG. 8

In these embodiments, the action of applying voltage to the magnetised coil **4** and generating electromotive force in the induction coil **5** by electromagnetic induction is similar to that of a static magnet dynamo of the basic configuration (**Figs. 1-4**).

The second and the third embodiments represented in **Fig.6** and **Fig.7** have the same basic configuration as the first embodiment, except that the first core **2** in each embodiment is shaped quite differently.

In the second embodiment, the part opposed to the end of the second core **3** sticks out toward the end of the second core **3**. Thus, the leakage flux due to the repellence of the first flux **11** and the second flux **12** generated in the first core **2** jumps across the gap **G** and enters the second core **3** with greater ease.

Embodiment 3

The third embodiment is so designed that the part coupling the second core **3** is that part of the first core **2** which is nearest to the permanent magnet **1** and, to shorten the open magnetic path even further, the two permanent magnets **1** are close to each other. Since a flux tends to form a closed magnetic path with the shortest distance, the leakage flux due to the repellence of the first flux **11** and the second flux **12** generated in the first core **2** jumps across the gap **G** and enters the second core **3** with greater ease.

Embodiment 4

The fourth embodiment indicated in **Fig.8**, as opposed to a static magnet dynamo of the basic configuration, consists of a first loop where permanent magnets **1** with multiple closed magnetic paths are arranged circularly with the fluxes oriented in the same direction, and of a second loop which is wound around with a magnetised coil **4** and installed inside the first loop. Furthermore, the parts with their first cores **2** coupling the first loop to the second one stick out toward each other across a specified gap. The parts where this first core **2** stick out are coupled together with a second core **3** via a gap **G** to form an open magnetic path. This reinforces the flux of the permanent magnets **1** and makes it easier for the leakage flux due to the repellence of the first flux **11** and the second flux **12** generated in the first core **2** to jump across the gap **G** and enter the second core **3**.

Embodiment 5

The configuration of a static magnet dynamo involved in this invention has so far been described in terms of embodiments where an open magnetic path is connected to the first core **2** at both ends of the second core **3** via a paramagnetic material. However, this invention is not limited to these embodiments.

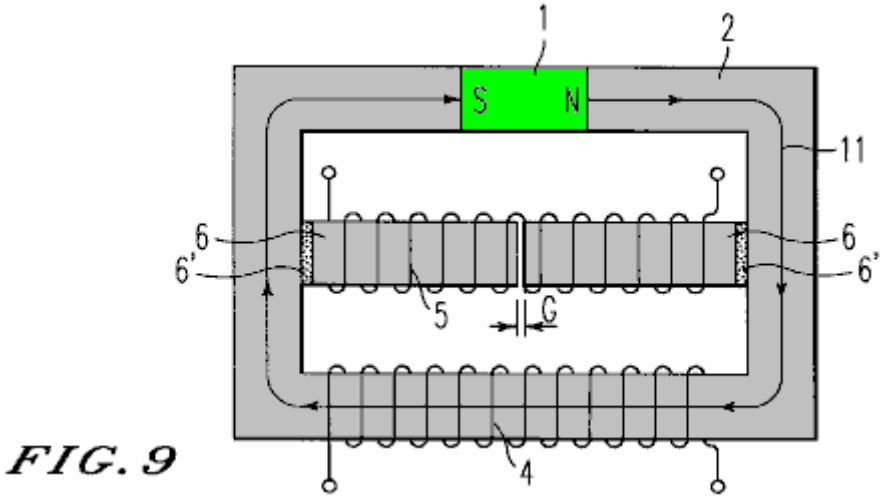


FIG. 9

That is, as indicated in **Fig.9**, the open magnetic path may be embodied by extending any two parts of the first core **2** in the direction that they approach each other, thus defining them as core extensions **6**, and coupling these core extensions **6** via a paramagnetic material **6'**. This embodiment can be applied to all embodiments mentioned above.

Embodiment 6

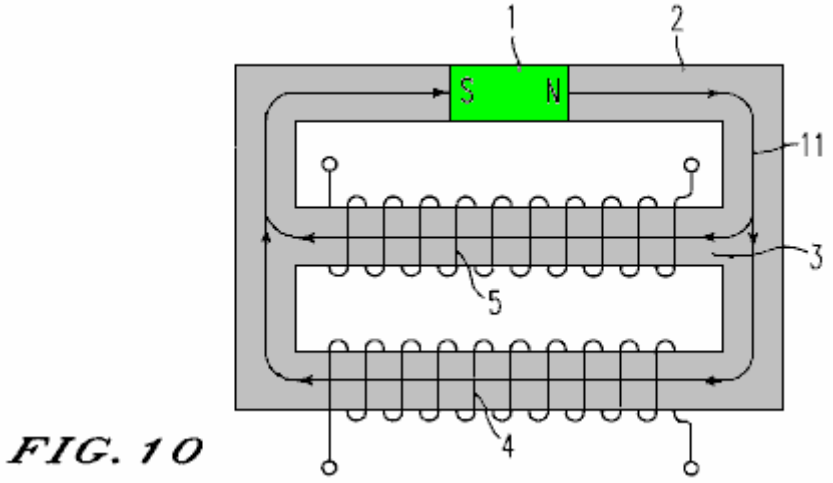


FIG. 10

As indicated in **Fig.10**, a closed magnetic path consists of a permanent magnet **1** and a first core **2** formed so as to couple the different poles of the said permanent magnet **1** in an annular manner. This closed magnetic path is then equipped with a second core **3** so that it comes magnetically in parallel with the permanent magnet **1**, so that a bypass closed magnetic path is composed of a permanent magnet **1**, part of a first core **2**, and a second core **3**.

The part consisting only of the closed magnetic path of the first core **2** is wound around with a magnetised coil **4**. The second core **3** is then wound around with an induction coil **5** designed to generate electromotive force by electromagnetic induction.

The action of a static magnet dynamo of the above configuration generating power is described below. First, when no voltage is applied to the magnetised coil **4** of a static magnet dynamo, the first core **2** forms a first flux **11** in the direction going from the N pole to the S pole of the permanent magnet **1**. In this state, a flux similar to that of the first core **2** is generated in the second core **3** as well.

Embodiment 7

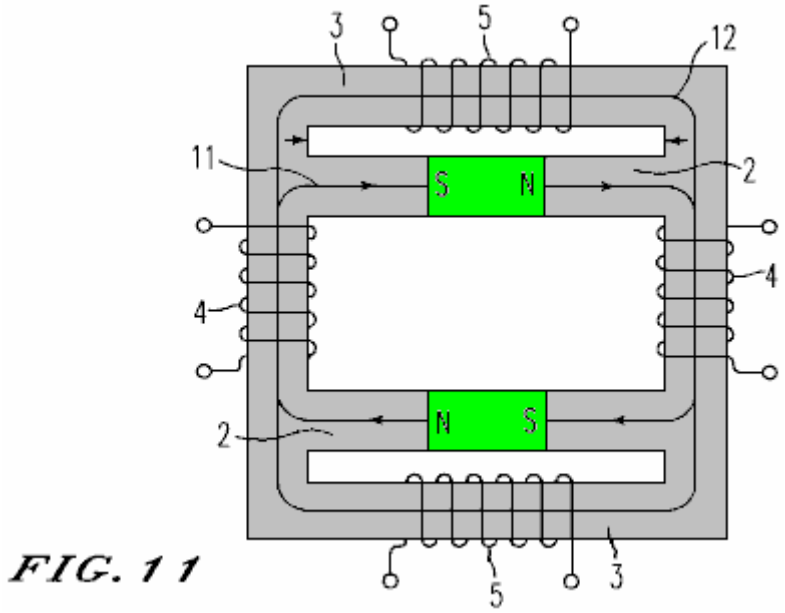


FIG. 11

The seventh embodiment is described below based on **Fig.11**, in terms of a static magnet dynamo system composed of two static magnet dynamos of the basic configuration and with the relative position of the permanent magnets changed.

In a static magnet dynamo, a closed magnetic path is composed of two permanent magnets **1** and two first cores **2** so designed as to couple the different poles of one of the permanent magnets **1** with the other permanent magnet **1** in an annular manner. This closed magnetic path is then equipped with a second core **3**. This results in the formation of a bypass closed magnetic path consisting of a permanent magnet **1**, part of a first core **2**, a paramagnetic material, and a second core **3**.

The parts where a closed magnetic path of each first core **2** alone is formed are wound around with a magnetised coil **4**. Each second core **3** is then wound around with an induction coil **5** designed to generate electromotive force by electromagnetic induction.

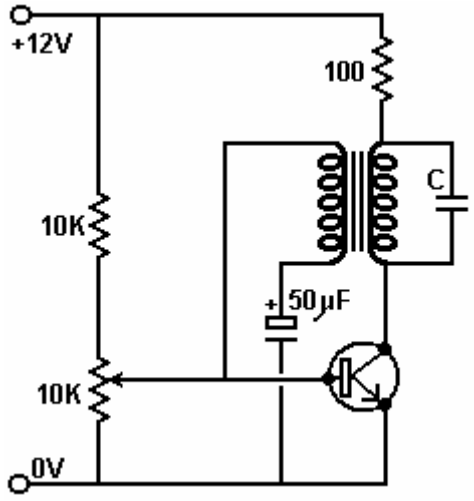
In this static magnet dynamo, where no voltage is applied to the magnetised coil **4**, a first flux **11** is formed in the first core **2** in the direction going from the N pole to the S pole of the permanent magnet **1**. The action of applying voltage to the magnetised coil **4** and generating electromotive force in the induction coil **5** by electromagnetic induction to generate power is similar to that of a static magnet dynamo of the basic configuration.

In the aforementioned static magnet dynamo incorporating two permanent magnets **1**, magnetic paths are arranged in a well-balanced manner. This makes it possible to make effective use of the flux of the permanent magnets **1**, so that power generation efficiency is higher than that of a static magnet dynamo of the basic configuration.

This invention has so far been described somewhat in detail in terms of the most favourable embodiments. Since it is clear that a wide variety of embodiments can be realised without opposing the philosophy and scope of this invention, this invention will not be limited to any particular embodiment, except for the limitations described in the attached claim.

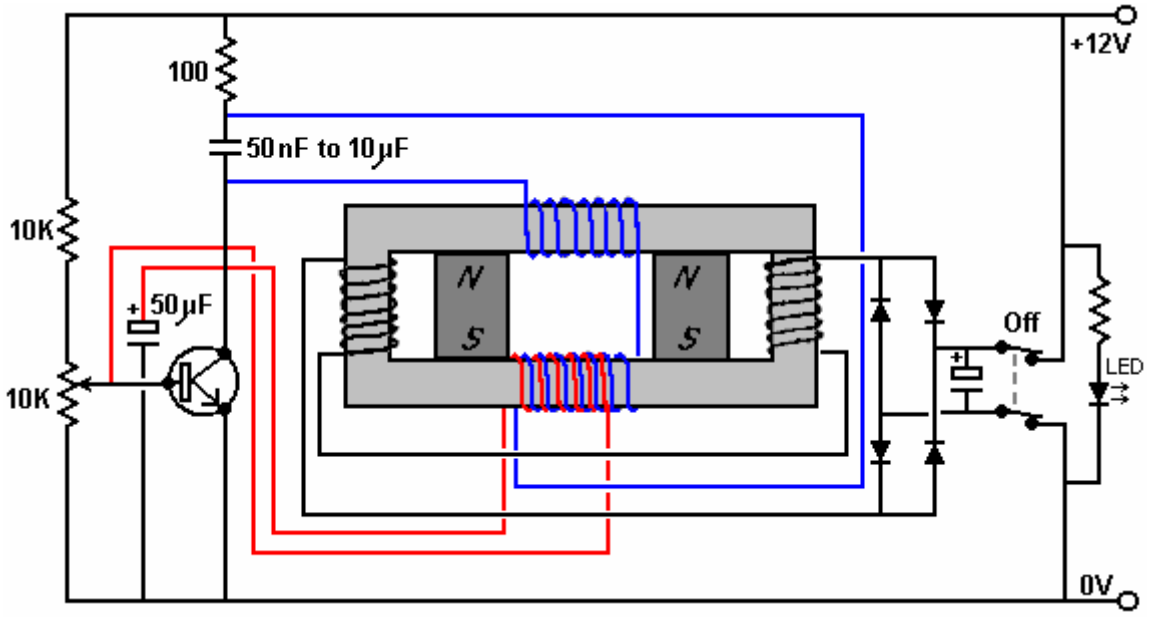
Stephan Leben's Circuits.

There is an interesting video posted on YouTube at http://www.youtube.com/watch?v=9zh_C3yvJH0 where Stephan W. Leben whose ID is "TheGuru2You" posts some really interesting information. He starts with a circuit produced by Alexander Meissner in 1913 and shown here:



Stephan states that he has built this circuit and can confirm that it is a self-resonating powering circuit. Once a twelve volt supply is connected to the input terminals, the transistor switches on powering the transformer which feeds repeating pulses to the base of the transistor, sustaining the oscillations. The rate of oscillation is governed by the capacitor marked "**C**" in the circuit diagram above and the coil across which it is connected.

Stephan suggests combining Alexander Meissner's circuit with Charles Flynn's magnetic amplification circuit. Here the transformer is switched to become the Charles Flynn oscillator winding plus a second winding placed alongside for magnetic coupling as shown here:



The transistor stage is self-oscillating as before, the transformer now being made up of the red and blue coil windings. This oscillation also oscillates the Flynn magnetic frame, producing an electrical output via the black coils at each end of the magnetic frame. This is, of course, an oscillating, or AC output, so the four diodes produce a full-wave rectified (pulsating) DC current which is smoothed by the capacitor connected to the diodes.

This circuit would be started by touching a 12 volt source very briefly to the output terminals on the right. An alternative would be to wave a permanent magnet close to the red and blue coils as that generates a voltage in the coils, quite sufficient to start the system oscillating and so, becoming self-sustaining. Stephan suggests using the piezo crystal from a lighter and connecting it to an extra coil to produce the necessary voltage spike when the coil is held close to the blue coil and the lighter mechanism clicked.

A surprising problem would be how to switch the device off since it runs itself. To manage this, Stephan suggests a two-pole On/Off switch to disconnect the output and prevent it supplying the input section of the circuit. To show whether or not the circuit is running, a Light-Emitting Diode ("LED") is connected across the output and the current flowing through it limited by a resistor of about 820 ohms.

Anyone wanting to try replicating this device will need to experiment with the number of turns in each coil and the wire diameter needed to carry the desired current. Stephan states that you need to have at least twice the weight of copper in the (black) output coils as there is in the (blue) input coils in order to allow the device produce excess power. The first page of the Appendix shows the current carrying capacity for each of the standard wire diameters commonly offered for sale. As this is a fairly recently released circuit, I am not aware of any replications of it at this time.

Floyd Sweet's "VTA" Generator.

Another device in the same category of permanent magnets in conjunction with oscillating coils, was produced by Floyd Sweet. The device was dubbed "Vacuum Triode Amplifier" or "VTA" by Tom Bearden.

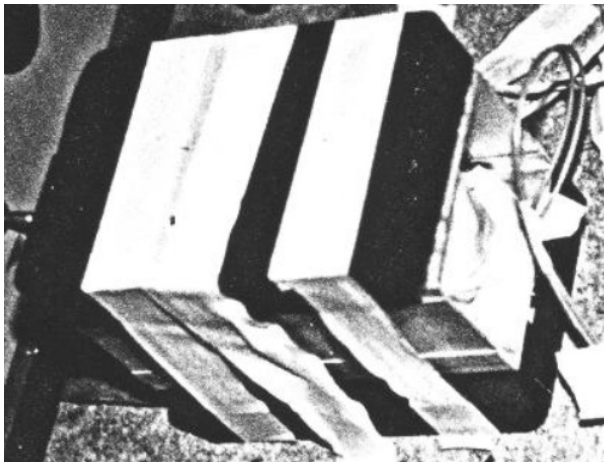
The device was capable of producing more than 500 watts of output power at 120 Volts, 60 Hz requiring less than one third of one milliwatt as input power. The output power can operate AC motors, lights, heaters and when rectified, DC motors.

Thanks is due to Horst Weyrich who has recently provided me with links to useful material which I had not seen before. This link: http://www.youtube.com/watch?v=0gM9natKIyY is to a video in which Floyd shows most of the magnet conditioning process.

Recently, some additional information on Floyd Sweet's device, has been released publicly by an associate of Floyd's who goes just by his first name of "Maurice" and who, having reached the age of seventy has decided that it is time to release this additional information. That information can be found in the Appendix. I am not aware of anybody who has succeeded in replicating Floyd's VTA, but here is as much information as I have at this time.

In the above video, Floyd talks about separating two of his conditioned permanent magnets with an "air gap" which is astonishing as he is putting them on opposite sides of a length of aluminium channel with thick walls and aluminium has a major damping effect on magnetic fields:

This arrangement which seems quite mad, is confirmed by a picture from Floyd's lab. As shown here:



This shows clearly that the open ends of the channel are not between the two magnets allowing an unrestricted magnetic field to flow between them, but instead, two channel thicknesses of aluminium are between the two magnets, obstructing the magnetic flow – quite remarkable !!

Floyd shows two coils being used to condition the magnets. The first is the large vertical coil shown here in front of Floyd:

The second coil is not seen as it is inside the vertical coil, sitting flat on the base, and consisting of an entire reel of AWG #17 (1.15 mm diameter) wire, something like this:

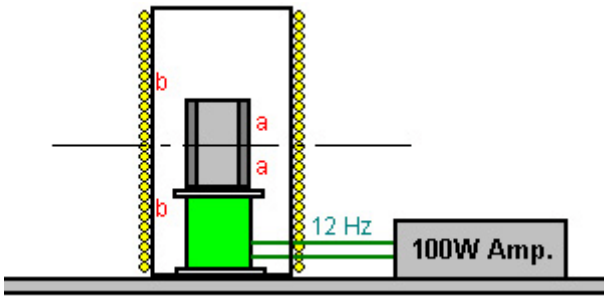
This coil operates effectively as an air-core solenoid, producing a strong axial magnetic field inside the larger coil which surrounds it. This inside coil is driven by a sine wave signal in the 10Hz to 15Hz range, boosted through a 100-watt audio amplifier which provides the current needed to impose the sine wave on this low impedance coil without distorting the waveform.

The first step is to determine the resonant frequency of each of the two permanent magnets to be used. The ferrite magnets used appear to be about 6 x 4 x 1 inch (150 x 100 x 25 mm). As they will eventually be used as a pair, one end of each is marked so that they can be aligned in the correct orientation after conditioning. That size of magnet appears to have a resonant frequency of about 12Hz, but each magnet will be slightly different.

The inside, low frequency coil is powered up at around 12Hz, the length of the magnet aligned with the Earth's magnetic field (that is, North/South), and placed on top of the vertical coil. An iron shim as used in transformer core construction is placed vertically on top of the magnet as an indication of resonating:

As shown in the video, the sine wave frequency feeding the AWG #17 coil is adjusted slowly to find the point at which the iron shim vibrates most strongly. That frequency is noted, and the same is done for the second magnet. It is not likely that the two resonant frequencies will be the same, and so the average frequency for the pair is used.

Next, the two magnets are placed in attraction mode, one on each side of the aluminium channel, with their marked ends at the same end of the channel. That is, the North pole face of one will touch the aluminium and the South pole face of the other will touch the aluminium. The two magnets and their separating aluminium channel are then placed inside the main coil and aligned so that the outer North pole faces southwards and the opposite external South pole faces Northwards. The large coil dimensions are arranged so that the centre of the magnets is at the centre of the outer coil:



While the inner coil continues to be powered up with a powerful sine wave, the outer coil is now fed a 60Hz stream of sharp voltage pulses. These are generated by charging up a 16,000 microfarad 250V capacitor and then discharging it through an SCR (Thyristor) connected to the outer coil. It is important that the spikes be sharp as they are imposing that frequency on the internal magnetic structure of the magnets. Presumably, if aiming for a European 240V version, then the outer coil would be pulsed at 50Hz rather than the American 60Hz and the capacitor would be a 450V rated type.

The working pictures of Floyd's prototype powering a load, seem to indicate that the one input coil and one output coil as described by Ashley Gray below, are the same as used by Floyd as we see wires coming out of the open end of the channel.

Horst wishes to emphasise that it is not easy to make a working replication of Floyd's design as the people at www.hyiq.org have tried to replicate it for a number of years now, without any success.

Ashley Gray of New Zealand.

In April 2014, I was sent some information about a colleague of Floyd Sweet's – Ashley Gray of Nelson, New Zealand. The version described by Ashley appears to be understandable.

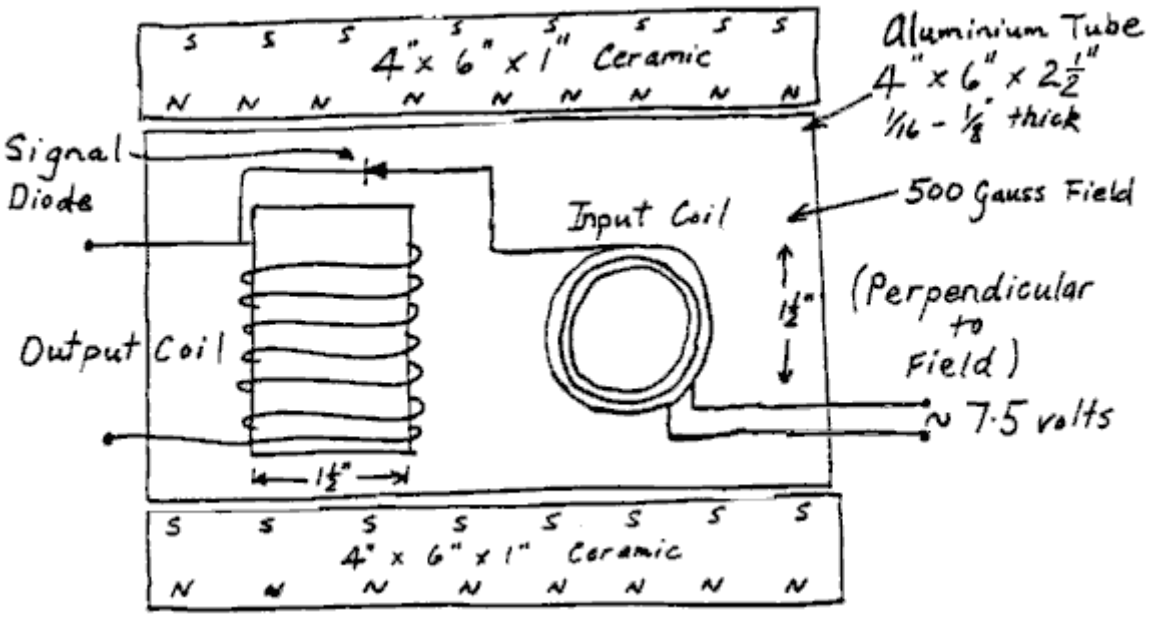
On 20th June 1994, Ashley says:

After a trip to America in 1985 when I was first introduced to Floyd Sweet, I was invited to go back and work with him. At that time he was being funded by Mark Goldes of the Aesop Institute, and Darryl Roberts was working as co-coordinator for the Institute in L.A. After working with Floyd for some time we left America for England. At that stage the "politics" had become difficult. While we were in England, we were contacted by Mark Goldes and told that Floyd had obtained some results which they would like us to verify for them.

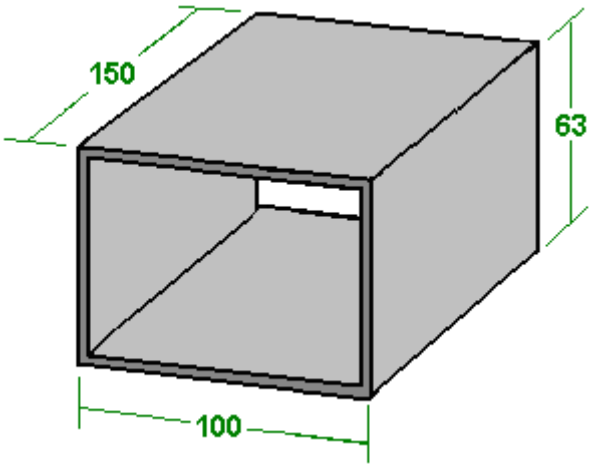
On our return to New Zealand, Darryl Roberts sent us the Lab Notes which he had recorded during the first tests of the "Space Quanta Modulator" and it's construction details. We were asked to repeat the experiments to verify the results. We constructed the device but were unable to get any results at that time. In the light of the new information which had been released, I conducted some further experiments and managed to obtain some interesting results without magnet 'conditioning' which, as far as I was aware, was not used in the original device.

The initial device which I constructed when in America, consisted of two 1-inch x 1-inch (25 x 25 mm) neodymium magnets mounted in a steel frame. There were two 'modulating' windings and one output winding. It was driven by a specially built sine-wave oscillator which was adjustable from 1 kHz to 2 kHz. We did not get any output or significant result from this device. Floyd felt that this was due to the high field strength of the neodymium magnets and the closed magnetic path. Floyd did not mention anything about magnet conditioning being necessary.

A second prototype was then built, using Barium Ferrite magnets size 6-inches x 4-inches x 1-inch (150 x 100 x 25 mm):

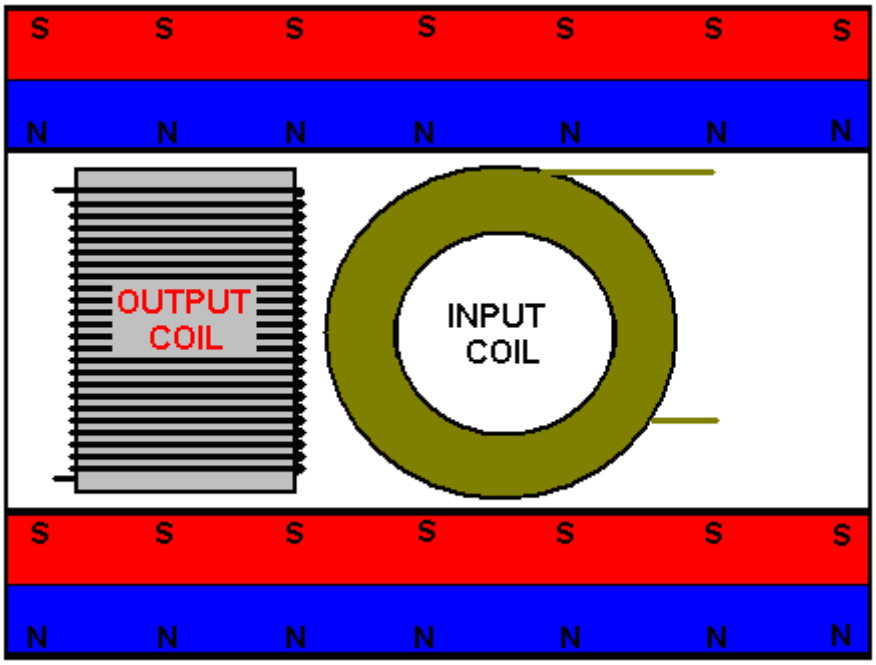


Ashley appears to be using magnets which are not 'conditioned'. Ashley has had what I consider to be very significant results from his prototype with an output power of 111 watts for an input of just 0.001 watts (a COP of 111,000). Ashley's successful build has an aluminium housing. People get the very mistaken idea that aluminium is not magnetic because magnets do not stick to it. The reality is that aluminium has a really major effect on magnetic fields and can be used as magnetic shielding if thick enough. Ashley's design uses two coils at right-angles to each other and that style of operation can be seen in other free-energy designs. Anyway, take a look at the version built by Ashley:



The aluminium enclosure is shallow. The ends are 4" x 2.5" which is 100 x 63 mm. The diameter of the input coil is 1.5 inches or 38 mm. Similarly, the output coil length has to be less than 63 mm.

The diagram below may give a slightly better idea of the dimensions involved in the construction. At the present time, magnets of that size are available for £14 each in the UK. They each have an 8 Kg pull and are very heavy.



The magnetic lines of force flow through the length of the output coil and through the width of the input coil. As you can see from the diagram, the unit is compact in spite of the very large magnets. The input needed is a good-quality sine wave. Ashley also says:

Details From Lab Notes of First Successful Tests

Original Test Setup:

A Signal Generator made by Wavetek, USA, was used to drive the input coil.

Input coil: 1.5" diameter 120 turns #20 gauge (0.812 mm diameter, overall resistance about 1 ohm)

Input = 7.5 volts at 3.l microamps = 23 micro watts

Output Coil: 1.5" diameter 12 turns #12 gauge (2.05 mm diameter)

Output = 10.4 volts sine wave at 1.84 amps = 19.15 watts at approx 400 Hz

Comments:

Frequency generally affected resistive 1.8 amp 20 watt bulb load proportionally - brightness increased with frequency, decreased with frequency except at certain points when it appeared inversely related, increasing as frequency decreased etc.

First Modifications:

The Signal generator was replaced with a purpose-built Sine Wave Oscillator of 9 volt output. The input coil was increased to 250 turns of #18 gauge (1.024mm) and the output coil was increased to 24 turns of #18 gauge(1.024 mm diameter) wire. Magnets, spacing, etc. all remained the same.

Input: 7.2 volts at 143 micro amps (0.001 watts)

Output: 24.2 volts at 4.6 amps = 111watts. Frequency 388 and 402 Hz

Comments:

By increasing the area of the wire exposed to/or occupying the fluctuating magnetic field the output was doubled. The exact proportions /ratios of the space filling volume of winding to output had not been determined at the time of writing. Magnet size seems to be less important than the volume of the windings, wire diameter, input voltage and current.

The current is only limited by the impedance of the wire which rises dramatically in the magnetic field to several hundred thousand ohms, while the impedance when outside the magnetic the field is only 2 or 3ohms @ 400 Hz. (250 turns #18 (1 mm diameter) wire).

The AC excitation current is only required to support I^2R losses as the magnetic field requires no additional power, as it is not loaded by the wires passing through the field.

The unit ran for 10 to 12 hours with no heating occurring but no longer duration tests were performed. The tests were witnessed by three people.

Technical Notes:

The quality of the oscillator is important - there should be no harmonic distortion i.e. it needs to be a pure sine wave.

The signal diode divides the current into the circuit, and being parallel - puts a small microamp current into the power coil as well as the excitation coil. This works with the magnets in such a way that there is produced a vector complementation.

When in close proximity to the magnets, the output load bulbs vibrate.

The Optical Generator of Pavel Imris.

Pavel was awarded a US patent in the 1970's. The patent is most interesting in that it describes a device which can have an output power which is more than nine times greater than the input power. He achieves this with a device which has two pointed electrodes enclosed in a quartz glass envelope which contains xenon gas under pressure (the higher the pressure, the greater the gain of the device) and a dielectric material.

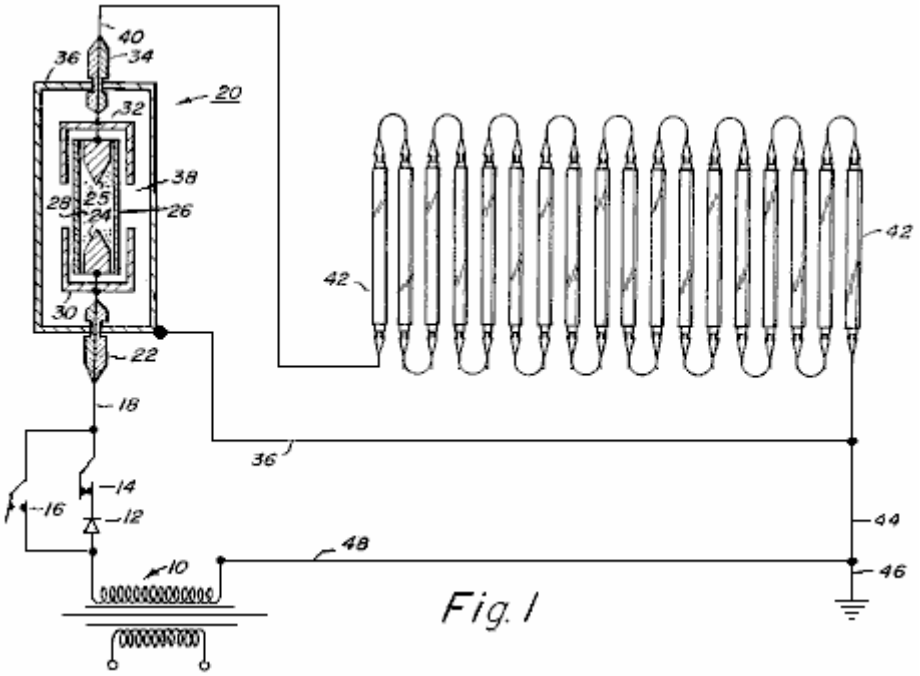


Fig. 1

Here, the power supply to one or more standard fluorescent lamps is passed through the device. This produces a power gain which can be spectacular when the gas pressure in the area marked '24' and '25' in the above diagram is high. The patent is included in this set of documents and it contains the following table of experimental measurements:

Table 1 shows the data to be obtained relating to the optical electrostatic generator. **Table 2** shows the lamp performance and efficiency for each of the tests shown in **Table 1**. The following is a description of the data in each of the columns of **Tables 1 and 2**.

Column	Description
B	**Gas used in discharge tube**
C	**Gas pressure in tube** (in torrs)
D	**Field strength across the tube** (measured in volts per cm. of length between the electrodes)
E	**Current density** (measured in microamps per sq. mm. of tube cross-sectional area)
F	**Current** (measured in amps)
G	**Power across the tube** (calculated in watts per cm. of length between the electrodes)
H	**Voltage per lamp** (measured in volts)
K	**Current** (measured in amps)
L	**Resistance** (calculated in ohms)
M	**Input power per lamp** (calculated in watts)
N	**Light output** (measured in lumens)

Table 1

		Optical	Generator	Section		
A	**B**	**C**	**D**	**E**	**F**	**G**
Test No.	**Type of discharge lamp**	**Pressure of Xenon**	**Field strength across lamp**	**Current density**	**Current**	**Power str. across lamp**
		(Torr)	**(V/cm)**	**(A/sq.mm)**	**(A)**	**(W/cm.)**
1	Mo elec	-	-	-	-	-
2	Xe	0.01	11.8	353	0.1818	2.14
3	Xe	0.10	19.6	353	0.1818	3.57
4	Xe	1.00	31.4	353	0.1818	5.72
5	Xe	10.00	47.2	353	0.1818	8.58
6	Xe	20.00	55.1	353	0.1818	10.02
7	Xe	30.00	62.9	353	0.1818	11.45
8	Xe	40.00	66.9	353	0.1818	12.16
9	Xe	60.00	70.8	353	0.1818	12.88
10	Xe	80.00	76.7	353	0.1818	13.95
11	Xe	100.00	78.7	353	0.1818	14.31
12	Xe	200.00	90.5	353	0.1818	16.46
13	Xe	300.00	100.4	353	0.1818	18.25
14	Xe	400.00	106.3	353	0.1818	19.32
15	Xe	500.00	110.2	353	0.1818	20.04
16	Xe	600.00	118.1	353	0.1818	21.47
17	Xe	700.00	120.0	353	0.1818	21.83
18	Xe	800.00	122.8	353	0.1818	22.33
19	Xe	900.00	125.9	353	0.1818	22.90
20	Xe	1,000.00	127.9	353	0.1818	23.26
21	Xe	2,000.00	149.6	353	0.1818	27.19
22	Xe	3,000.00	161.4	353	0.1818	29.35
23	Xe	4,000.00	173.2	353	0.1818	31.49
24	Xe	5,000.00	179.1	353	0.1818	32.56

Table 2

		Fluorescent	Lamp	Section	
A	**H**	**K**	**L**	**M**	**N**
Test No.	**Voltage**	**Current**	**Resistance**	**Input Energy**	**Light Output**
	(Volts)	**(Amps)**	**(Ohms)**	**(Watts)**	**(Lumen)**
1	220	0.1818	1,210	40.00	3,200
2	218	0.1818	1,199	39.63	3,200
3	215	0.1818	1,182	39.08	3,200
4	210	0.1818	1,155	38.17	3,200
5	200	0.1818	1,100	36.36	3,200
6	195	0.1818	1,072	35.45	3,200
7	190	0.1818	1,045	34.54	3,200
8	182	0.1818	1,001	33.08	3,200
9	175	0.1818	962	31.81	3,200
10	162	0.1818	891	29.45	3,200
11	155	0.1818	852	28.17	3,200
12	130	0.1818	715	23.63	3,200
13	112	0.1818	616	20.36	3,200
14	100	0.1818	550	18.18	3,200
15	85	0.1818	467	15.45	3,200
16	75	0.1818	412	13.63	3,200
17	67	0.1818	368	12.18	3,200
18	60	0.1818	330	10.90	3,200
19	53	0.1818	291	9.63	3,200
20	50	0.1818	275	9.09	3,200
21	23	0.1818	126	4.18	3,200
22	13	0.1818	71	2.35	3,200
23	8	0.1818	44	1.45	3,200
24	5	0.1818	27	0.90	3,200

The results from Test No. 24 where the gas pressure is a very high 5,000 Torr, show that the input power for each 40-watt standard fluorescent tubes is 0.9 watts for full lamp output. In other words, each lamp is working to its full specification on less than one fortieth of its rated input power. However, the power taken by the device in that test was 333.4 watts which with the 90 watts needed to run the 100 lamps, gives a total input electrical power of 423.4 watts instead of the 4,000 watts which would have been needed without the device. That is an output power of more than nine times the input power.

From the point of view of any individual lamp, without using this device, it requires 40 watts of electrical input power to give 8.8 watts of light output which is an efficiency of about 22% (the rest of the input power being converted to heat). In test 24, the input power per lamp is 0.9 watts for the 8.8 watts of light produced, which is a lamp efficiency of more than 900%. The lamp used to need 40 watts of input power to perform correctly. With this device in the circuit, each lamp only needs 0.9 watts of input power which is only 2.25% of the original power. Quite an impressive performance for so simple a device!

The Michel Meyer and Yves Mace Isotopic Generator.

There is a French patent application number FR 2,680,613 dated 19th August 1991 entitled “Activateur pour Mutation Isotopique” which provides some very interesting information. The system described is a self-contained solid-state energy converter which abstracts large amounts of energy from an ordinary iron bar. This is also shown in Michel’s Czechoslovakia Patent No.284,333

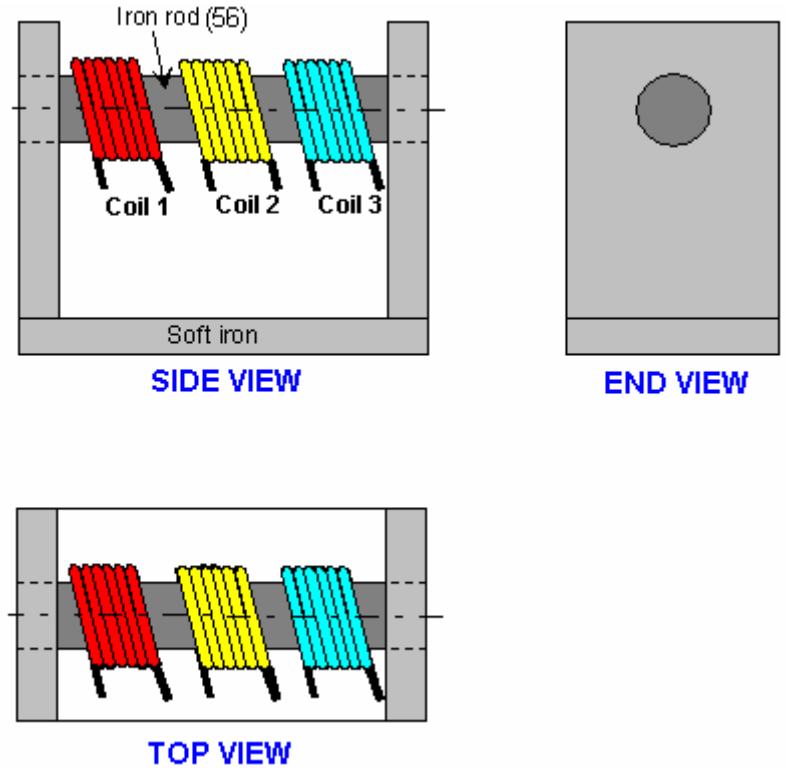
The inventors describes the technique as an “isotopic mutation effect” as it converts ordinary iron (isotope 56) to isotope 54 iron, releasing large amounts of electrical energy in the process. This excess energy can, they say, be used to drive inverters, motors or generators.

The description of the mechanism which is being used by the device is: “the present invention uses a physical phenomenon to which we draw attention and which we will call ‘Isotopic Change’. The physical principle applies to isotope 56 iron which contains 26 protons, 26 electrons and 30 neutrons, giving a total mass of 56.52 Mev, although its actual mass is 55.80 Mev. The difference between the total mass and the actual mass is therefore 0.72 Mev this which corresponds to an energy of cohesion per nucleon of 0.012857 Mev.

So, If one introduces an additional 105 ev of energy to the iron core isotope 56, that core isotope will have a cohesion energy level of 0.012962 Mev per nucleon corresponding to iron isotope 54. The instability created by this contribution of energy will transfer the isotope 56 iron to isotope 54 causing a release of 2 neutrons.

This process generates an excess energy of 20,000 ev since the iron isotope 54 is only 0.70 Mev while isotope 56 has 0.72 Mev. To bring about this iron isotope 56 conversion, we use the principle of Nuclear Magnetic Resonance."

The practical method for doing this is by using three coils of wire and a magnetic-path-closing support frame of iron as shown in this diagram:



In this arrangement,

Coil 1: Produces 0.5 Tesla when fed with DC, converting the iron bar into an electromagnet
Coil 2: Produces 10 milli-Tesla when fed with a 21 MHz AC sinewave signal
Coil 3: Is the output coil, providing 110, 220 or 380 volts AC at about 400 Hz depending on the number of turns in the coil

This simple and cheap system has the potential for producing substantial energy output for a very long time. The inventors claim that this device can be wired to be self-powered, while still powering external devices. Coil 1 turns the iron rod into an electromagnet with it's flux channelled in a loop by the iron yoke. Coil 2 then oscillates that magnetic field in resonance with the isotope 56 iron atoms in the rod, and this produces the isotope conversion and release of excess energy. Coil 3 is wound to produce a convenient output voltage.

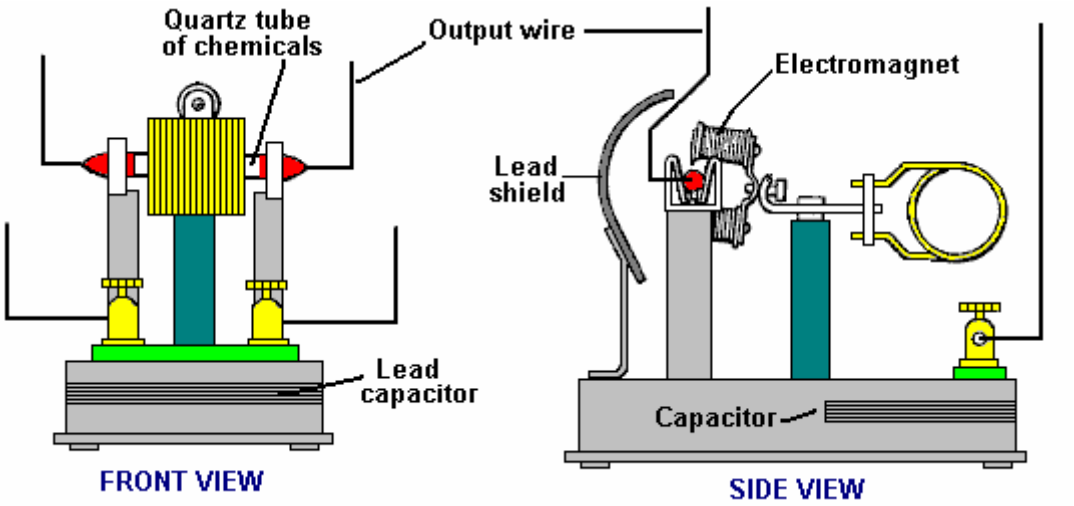
The Colman / Seddon-Gilliespie Generator.

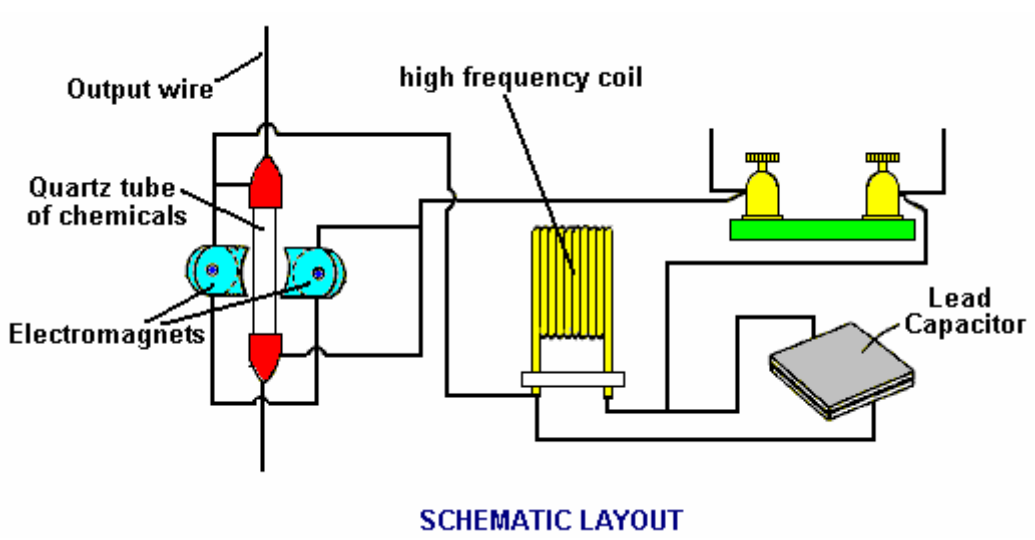
This device, patented by Harold Colman and Ronald Seddon-Gillespie on 5th December 1956, is quite remarkable. It is a tiny lightweight device which can produce electricity using a self-powered electromagnet and chemical salts. The working life of the device before needing refurbishment is estimated at some seventy years with an output of about one kilowatt.

The operation is controlled by a transmitter which bombards the chemical sample with 300 MHz radio waves. This produces radioactive emissions from the chemical mixture for a period of one hour maximum, so the transmitter needs to be run for fifteen to thirty seconds once every hour. The chemical mixture is shielded by a lead screen to prevent harmful radiation reaching the user. The patent, GB 763,062 is included in the Appendix.

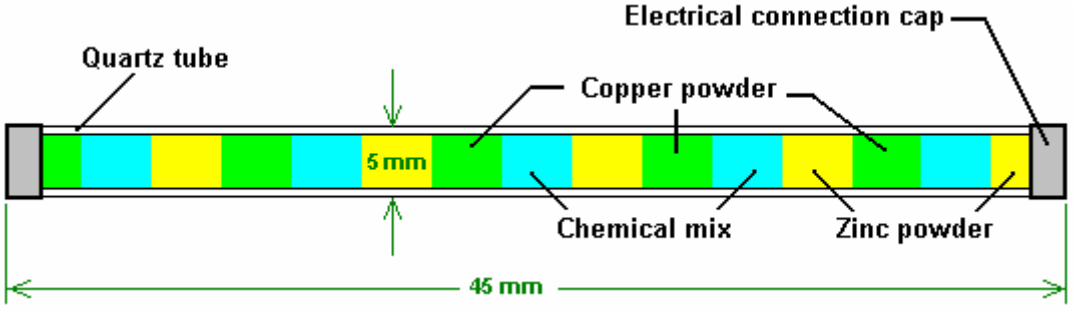
This generator unit includes a magnet, a tube containing a chemical mixture of elements whose nuclei becomes unstable as a result of bombardment by short waves so that the elements become radio-active and release electrical energy, the mixture being mounted between, and in contact with, a pair of different metals such as copper and zinc, and a capacitor mounted between those metals.

The mixture is preferably composed of the elements Cadmium, Phosphorus and Cobalt having Atomic Weights of 112, 31 and 59 respectively. The mixture, which may be of powdered form, is mounted in a tube of non-conducting, high heat resistivity material and is compressed between granulated zinc at one end of the tube and granulated copper at the other end, the ends of the tube being closed by brass caps and the tube being carried in a suitable cradle so that it is located between the poles of the magnet. The magnet is preferably an electro-magnet and is energised by the current produced by the unit. The transmitter unit which is used for activating the generator unit may be of any conventional type operating on ultra-shortwave and is preferably crystal controlled at the desired frequency.





The transmitter unit is of any suitable conventional type for producing ultra shortwaves and may be crystal controlled to ensure that it operates at the desired frequency with the necessity of tuning. The quartz tube containing the chemical mixture, works best if made up of a number of small cells in series. In other words, considering the cartridge from one end to the other, at one end and in contact with the brass cap, there would be a layer of powdered copper, then a layer of the chemical mixture, then a layer of powdered zinc, a layer of powdered copper, etc. with a layer of powdered zinc in contact with the brass cap at the other end of the cartridge. With a cartridge some forty five millimetres long and five millimetres diameter, some fourteen cells may be included.



The Devices of Don Smith.

One free-energy developer who had most impressive claims for his devices is the late Don Smith who produced many spectacular devices, generally with major power output. Don says that his understanding comes from the work of Nikola Tesla as recorded in Thomas C. Martin's book "The Inventions, Researches, and Writings of Nikola Tesla" ISBN 0-7873-0582-0 available from http://www.healthresearchbooks.com and various other book companies. This book can be downloaded from http://www.free-energy-info.tuks.nl/ as a pdf file, but a paper copy is much better quality and easier to work from.

Many experimenters have spent considerable time and effort in attempts to replicate the work which Don reports and while COP>1 has definitely been achieved, high power has not yet been reached. If you want more detail then it can be found here: http://www.free-energy-info.tuks.nl/DonSmith.pdf as a free download.

Mohamed assesses Don Smith's Devices

Late in 2014, Mohamed released the following document after two years of examining and testing Don Smith's main design. If preferred, his document can be downloaded as a separate pdf file:
http://www.free-energy-info.com/Mohamed.pdf

بسم الله الرحمن الرحيم

In the name of Allah, most gracious, most merciful

The Resonance Energy Device Explained

Preface

This presentation is a favour from Allah (God), the Thanks is to HIM even though it's the result of more than two years of deep thinking about the resonance energy device invented by Donald Lee Smith. I was interested in this device due to the huge amount of power which it is able to provide. The device has no moving parts and is small in size. This presentation is an attempt to explain two important sources of information about the resonance energy device; they are a document and a video:

The document is located here: http://www.free-energy-info.com/Smith.pdf

The video can be seen here: http://www.youtube.com/watch?v=cQkYAh8Qgb4

The information contained in the document is free and open to the public. I think that it is time for such information to become widely known - free energy is free because it's for everyone. Energy is everywhere in an unlimited quantity, ready to be taken with minimum effort, the resonance energy device described here is one of

the most attractive free energy devices, it depends on a multidimensional transformer which makes possible a harmonic exchange between the positive energy reality and the unseen negative sea of energy. This is possible using a spark-gap arrangement, the energetic symmetry between the positive and negative energy oceans make it preferable to use the term **ambient background energy** rather than the name **zero point energy**. The video pointed to above, is an excellent source that can be viewed in parallel with this work, and when considered together, the reader will understand both the video and this presentation.

At the start of the above video, the inventor explains the importance of the magnetic field in generating electricity. The magnetic energy upsets the background energy and that results in a separation of electrons between the two ends of a coil, this separation of electrons is a source of electric power.

The resonance energy device is based on a very important idea, namely, that magnetism and electricity are two sides of a single entity!

Matter and energy are two aspects of the same thing as Oliver Heaviside expressed in his famous energy equation $E = MC^2$. The electrical equivalent of this in our Resonance Energy Device is:

$$E = (\text{Volts} \times \text{Amperes}) \times (\text{Cycles per Second})^2$$

In this presentation, we will learn together how the device works but before any attempt to use this information, please pay attention to the following caution:

You need to read this document again and again to understand the power behind this device; this device is a power producer! You are near a power production factory but in a small size, this special coil only needs voltage and when we oscillate that voltage, the voltage turns into real current !!

This means that even a short-circuit isn't permitted in the collection system if you plan to use HV capacitors. This is not the place for amateurs, please do not attempt to implements or use the information shown here unless you are experienced and skilled. Neither the publisher nor the author makes any representations as for the completeness or the accuracy of the information contained here and disclaims any liability for damages or injuries resulting from your actions.

Algeria, August 2014
hopehope3012 (at) gmail (dot)com

Let's examine how a magnetic field generates a voltage in an ordinary coil and why our coils lack the energy key that Donald Smith has found!

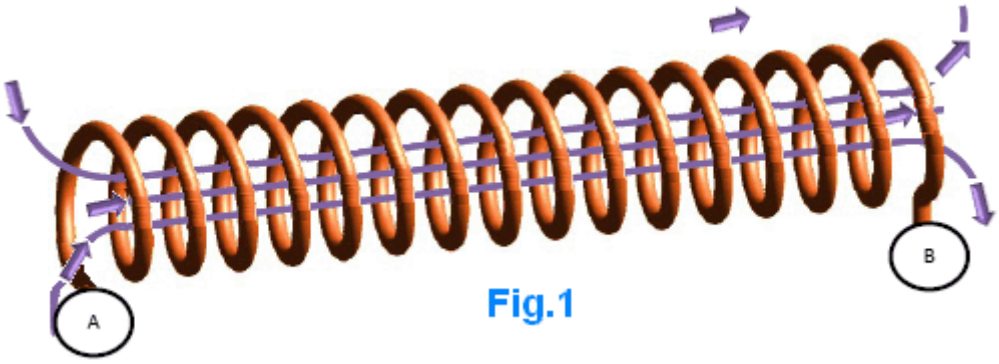


Fig.1

When a magnetic field penetrates a coil, the induced rotating electric field generates additive elementary voltages along the coil. Every turn in the coil receives the same voltage, we extract the power along the length of the coil,

but we miss the power available in the area inside the coil !! To understand this, let's examine the induced electric field:

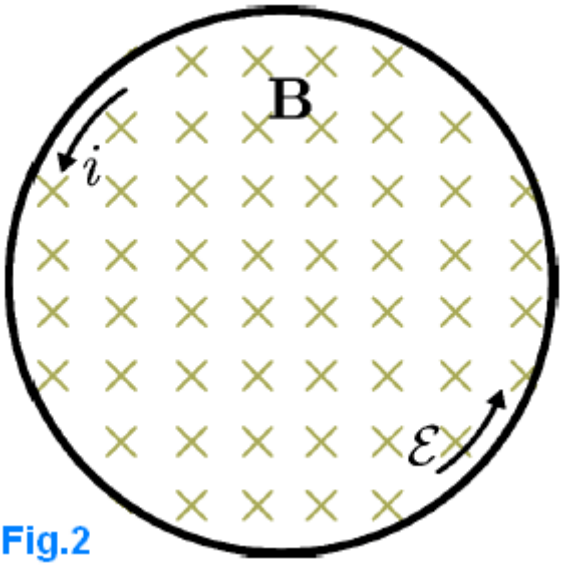


Fig.2

If we examine Fig.2, which represents a single turn of the coil shown in Fig.1, we see that the magnetic field penetrates all of the area inside that ring. This magnetic field will be transformed to a voltage because it will drive a rotating electric field, this field is able to induce a voltage and this induced voltage will cause the current to flow due to the different of voltage between the two points **A** and **B** in Fig.1.

The important, generally forgotten thing about the induced electric field, is its availability inside our coil as shown in Fig.3:

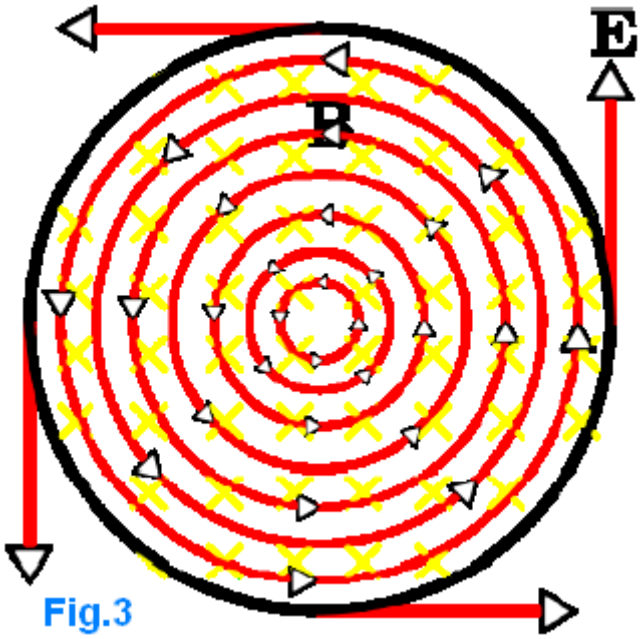


Fig.3

The induced electric field exists independent of the conducting loop. In other words, an induced electric field permeates all of the space within the region of the changing magnetic field, as indicated by the red field lines in Fig 3. What about this field? It is wasted power. It is wasted power at point **A**, as well as all of the distance between the two points **A** and **B**.

If we want to achieve power amplification we have to combine the magnetic field with the induced electric field in such a way as to conserve the non-conservative electric field! The induced rotating electric field will remain non-conservative but we could play with the induced voltage created by that field using a Tesla Bi-filar Coil ("TBC").

We need to extend the capacitive side of a Tesla bi-filar coil in order to benefit from the induced E field to a reasonable degree, if we try to use a flat TBC we will face the same problem in our solenoid coils where we concentrate between two points… The heart of the Resonance Energy device is based on an extended Tesla bi-filar coil (Fig.4)

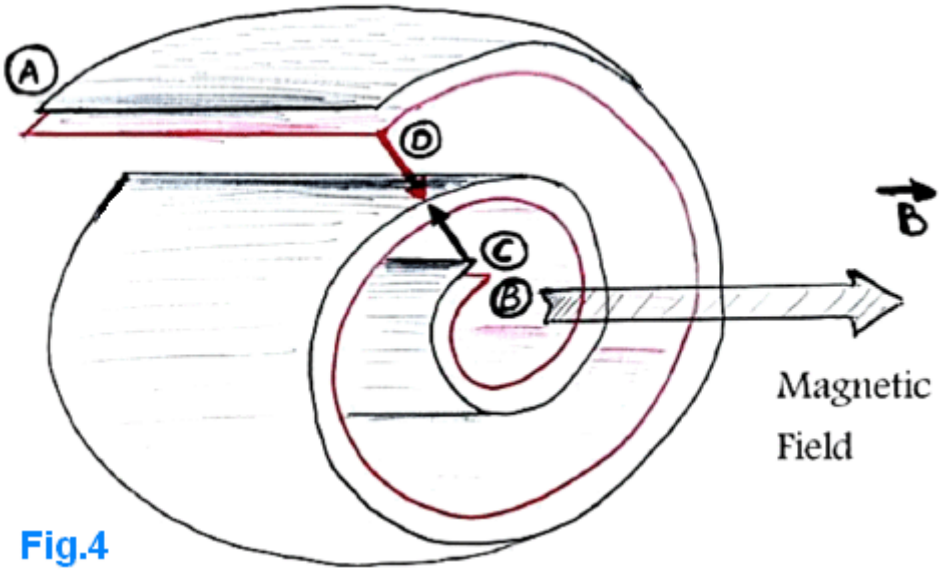


Fig.4

During my search for the specification of Tesla bi-filar coil I wasn't able to find a lot of information such as the resonance frequency of this special coil as if it has been forgotten, although it has been known since 1894!

I felt it's without advantage; fortunately at that time I successfully built a new type of asymmetrical capacitor that has four plates rather than three plates, I was able to replicate my source voltage, I gave it the name C1/C2 system, when I fed C1 with high voltage I was able to take the voltage from C2 without direct contact, the device is based on the electric field interaction between C1 and C2.

The C1/C2 system opened my mind to the possibility of charging a capacitor without the need of direct contact; with the device. I began thinking in a different way, wondering how I could combine the two side of electricity in just one device. The device shown in Fig.4 can be simplified as shown in Fig.5:

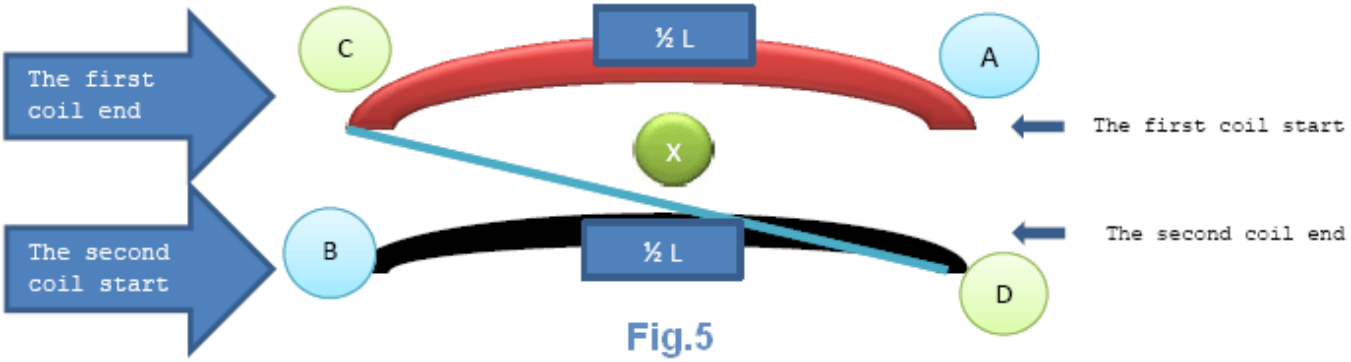


Fig.5

The magnetic side is the path from **A** to **B**, it is the path which the current will follow, and then the magnetic energy will be formed and concentrated inside our coil, the magnetic side is the hidden side of electricity because we can't conserve this state for a long time as we do in the capacitors. This side is directly related to the ambient background energy or zero point energy. (Later we will see why Donald Smith prefers the name of ambient background energy as opposed to ZPE)

To make things easier we could take the permanent magnet as an example, the permanent magnet appears to create the magnetic field all the time without requiring us to provide any visible power for which we have to pay.

The electric side of this amazing coil is the most difficult thing to understand. The capacitor in our extended Tesla bi-filar coil is just one point! It's the point "**X**" but from a static perspective, we say that it is between points **C** and

D. The path from **C** to **D** is the way to conserve the direction of current needed to build the coil and ensure the maximum voltage is generated between the two plates of the capacitor. This capacitor is a dynamic capacitor and not a static capacitor. In this capacitor the displacement current is absent... because it gathers the two polarities of electricity in one device. The magnetic side is the current side where the current starts from point **A** and flows to point **B**. The electric side is the point at which the maximum voltage between the two plates exists. I think that this is the most difficult point to understand in the Tesla Bi-filar Coil, because we see it as just a coil and not a coil with a built-in capacitor!

Back to 1894 and take a look at what Mr. Nikola Tesla said about his bi-filar coil:

> **My present invention has for its object to avoid the employment of condensers which are expensive, cumbersome and difficult to maintain in perfect condition, and to so construct the coils themselves as to accomplish the same ultimate object.**

Let us start by analysing how the extended Tesla bi-filar coil works. Imagine that the following drawing is a charged, extended Tesla coil with the positive and negative parts, shown as red and black plates respectively.

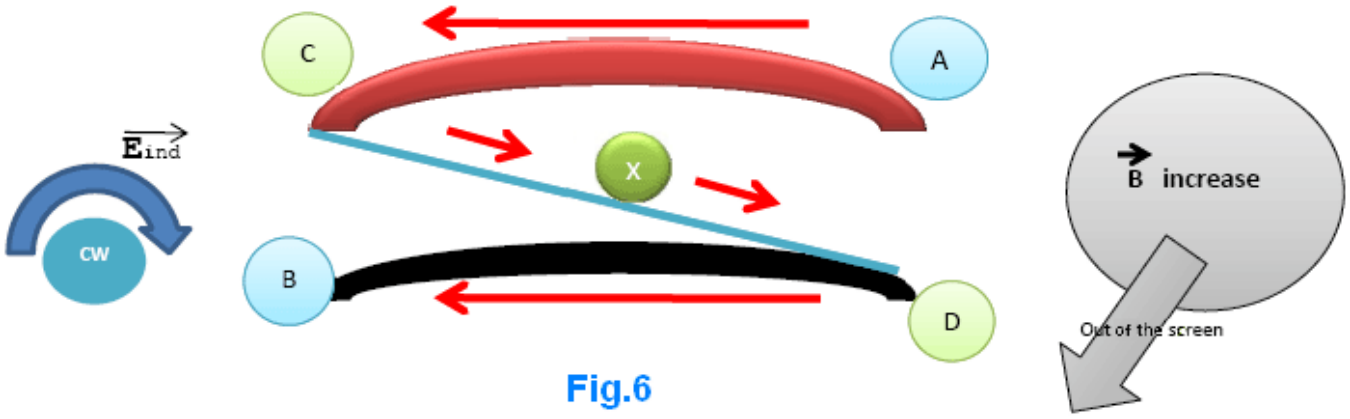


Fig.6

When we connect these two plates by their points **C** and **D**, an electric current will flow from point **A** to point **B** as shown by the red arrows. When the induced electric field finds a coil which is wound counter-clockwise ("CCW") – the red plate - versus another coil wound clockwise ("CW") – the black plate - this electric field will generate an induced voltage which will tend to recharge the capacitor again!

The rotating electric field will start to build an induced electric voltage from point **B** which will be 0V in this case because the induced electric fields rotate in a clockwise direction. When the magnetic field increases, the electric current will flow counter-clockwise and so the direction of the magnetic field will be out the screen as indicated in Fig.6.

In this case we achieve a cost-free voltage difference between adjacent plates. This situation attracts electrons which were not previously available, to become incorporated and produce a very large net gain in potential, this gain is real !!

The induced rotating electric field flows in closed loops but this design causes it to build a voltage difference between adjacent plates. This fact prevents the voltage difference between the adjacent plates from dropping and it makes the replication of electric power with radio frequency a practical operation.

Because of this, we can understand why resonance can't produce the power but it replicates the power with radio frequency and this is the secret of the huge power which this device is able to give with the equation which Don Smith gave us:

$$\text{Power in one second} = \tfrac{1}{2} C V^2 F^2$$

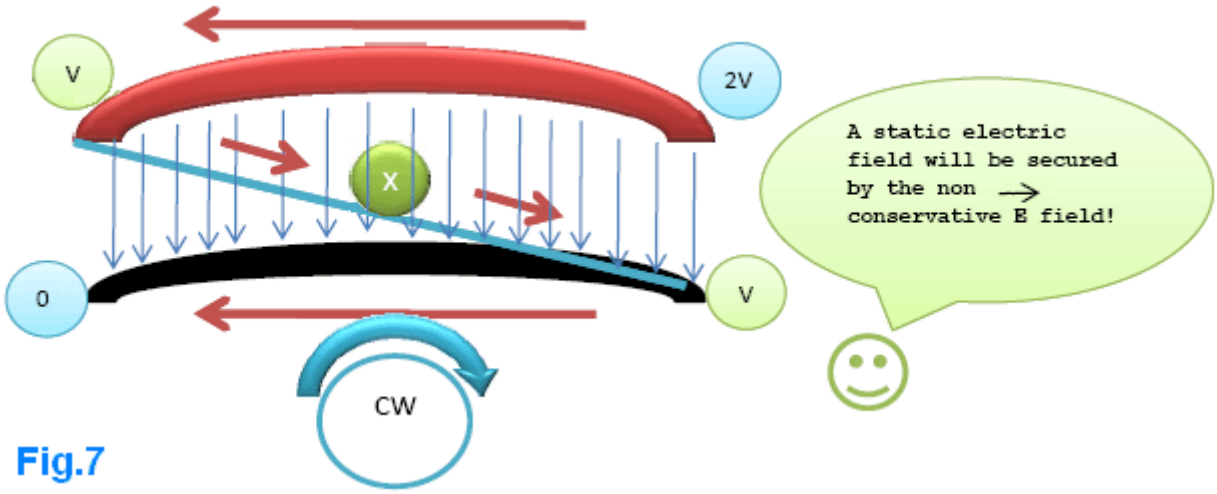


Fig.7

The induced rotating electric field will build induced elementary voltages; **B** is the starting point for that field, so it will have 0V but when we arrive at point **D** it will have V volts, which same voltage will be available at point **C**. The induced rotating electric field will continue it's rotation producing a value of 2V at point **A** !! (Fig.7).

A question arises: where is the capacitor if **C** and **D** both have the same voltage?

The capacitor in our extended Tesla bi-filar coil (TBC) is a dynamic capacitor, it will be formed only when the coil oscillates. When that happens, the induced electric field will give **C** and **D** the same value of voltage but every adjacent turn will receive the same difference of voltage needed for attracting ambient electrons which weren't originally in the system but now are incorporated in large numbers, providing the excess power. This capacitor appears due to the junction **CD**, and from a dynamic point of view, it is the point **X**.

This device treats the magnetism and electricity as if they are the same thing. When this happens both sides help the other side. The displacement current is absent here because it appears when we separate electricity and magnetism. When we do that, we fill the gap by introducing something isn't real. The displacement current does not exist !!

When you move a permanent magnet in the air, **an induced electric field will appear regardless of the presence of a coil.** The electricity is there, it comes from the background energy and returns to its background level, both electricity and magnetism have the same origin, the magnetism is the energy side (magnetic field), the electricity is the physical side (electrons).

The physical side appears when we place a conducting coil in an area where there is a changing magnetic field. The magnetic field causes the electrons to rotate CCW and produce voltage and rotate CW and produce current, voltage electrons appear to be more negative because they are more active, current electrons appear to be less negative. We seek the number of electrons separated by the changing magnetic field. The current (I) is the energy killer because it absorbs the power of the voltage electrons!

In this system, the magnetic field penetrates inside a special capacitor coil where both sides of **electric power** are found in a unified state. We discussed the phase when the capacitor discharges across itself, and now it is time to see how the coil will charge the capacitor with a reversed polarity.

It's important to review the resonance in a normal L/C circuit to understand the work of the extended TBC. Imagine that you have a charged capacitor, when connected to a coil a current starts to flow, when the current reaches its maximum value, this means that the capacitor was totally discharged, the magnetic field will have the greater value and start to decrease and induce a voltage which will charge the capacitor again with the opposite polarity.

The same will happen in our extended TBC, when the magnetic field reaches its maximum value it will then decrease and induce a voltage which will charge the capacitor again, the complexity here is to understand the positions of coil and capacitor.

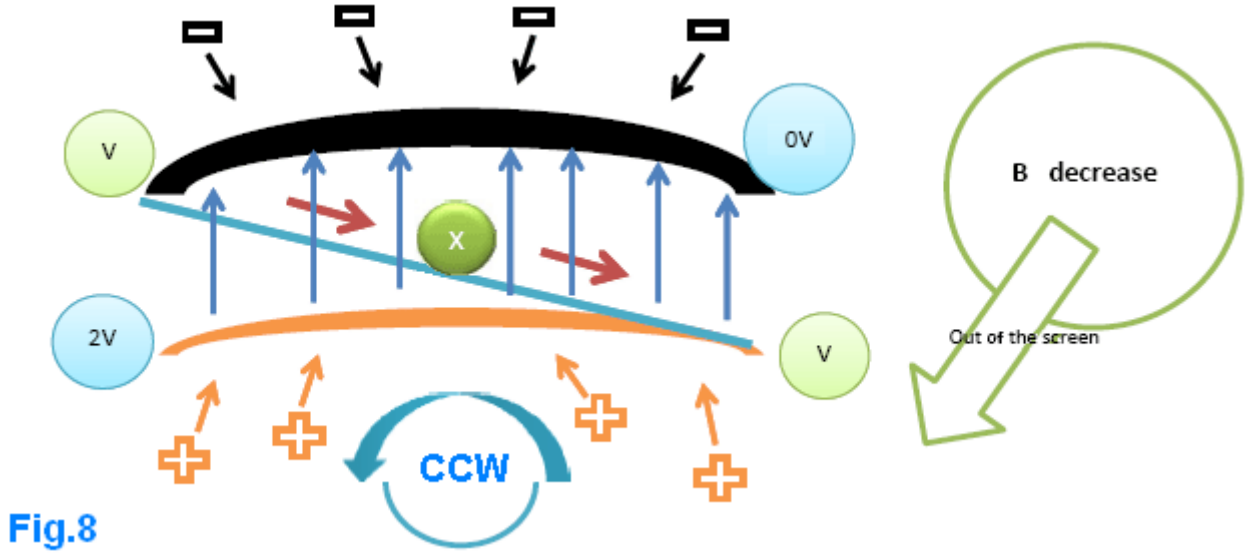


Fig.8

In a dynamic perspective the coil sees the capacitor in the **X** position, the coil situated between **A** and **B**, comparing this to what happens in a normal L/C circuit at this phase the coil has to drive a current **outside its region**, the outside region between **A** and **B** has to be position **X**. when the magnetic field induces a current to recharge the capacitor between **C** and **D**, the induced electric field will do this job because it will change the direction of rotation to be CCW (Fig.8), and this behaviour will tend to keep the current at a high value while the voltage increases! The position of the capacitor in this device has to be the position of a spark gap. A spark gap in this position will improve the performance **dramatically** and amplify the power gain.

The Resonance Energy device works with radio speeds above 20 kHz. The ambient background electrons are usually found in a dormant or inactive state and this makes the drawing-in of ambient electrons a difficult operation. For this reason, we need an earth connection to improve the performance of this system, later in the video, Don Smith talks about something new, the earth connection isn't needed because there is something called **air grounding**.

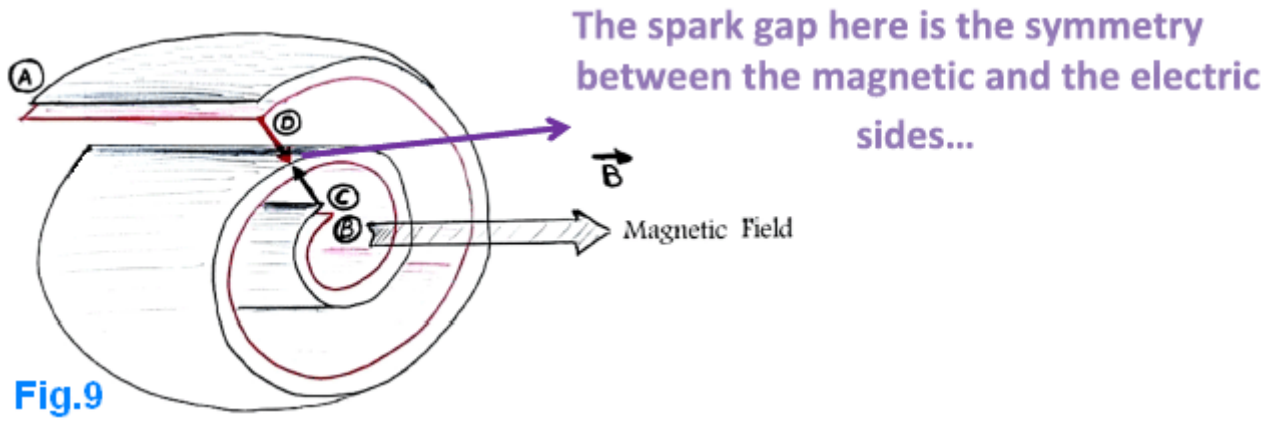


Fig.9

To activate air electrons effectively we need another source of power, actually this source of power exists as a negative energy but before we go further, we have to see the symmetry between the electric and magnetic sides. When we start from **A** to **C** we have **half coil-capacitor** while from **D** to **B** we have **capacitor-half coil**.

To understand how our extended TBC receives negative energy, it is better to see the way in which the device oscillates. There are two methods of feeding this coil:

The direct method

Here we have to know the resonant frequency of this coil or alternatively, we need to have the resonant frequency covered by our feeding circuit. This method is the best because we don't need lots of power, because when we achieve resonance our extended TBC will have a very high impedance, and so, a maximum voltage will be available between **A** and **B**, and fewer turns will be needed in the **reactant coil L2**.

The indirect method
Here we don't have to know the resonant frequency, if we feed this coil from **A** and **B** the capacitor will charge. The spark gap has a very high resistance and it will not fire until there is a voltage difference between points **C** and **D**, at that moment a spark will occur, and when that happens, the resistance of the spark gap drops from a very high value to very low value and that short-circuits any power coming from the power source until the natural coil resonance finishes!

When the capacitor is fully charged, the maximum voltage appears between points **A** and **B** which is the coil side, **when the spark occurs the capacitor will transform into a coil which has a capacitor built into it!**
This operation provides the natural resonance needed without causing any problems.

Caution!

When we feed this coil from point **A** to point **B** and then disconnect it from the power source, you would then expect to discharge it by short-circuiting point **C** to point **D** (usually from the location of spark gap). If you do this and then touch the coil **you will definitely receive a substantial high voltage shock!** Even if you see the spark, the spark will not discharge this capacitor immediately but it will cause the coil to oscillate again and recharge itself. This device isn't a simple capacitor because when you short-circuit the spark gap it will transform the electrical energy into electromagnetic energy capable of recharging the capacitor again.

If you try to short-circuit points **A** and **B** in an attempt to discharge the extended TBC, then these two points will **transform** into electric points (due to the symmetry). From a static perspective, the capacitor can take an **AB** position as well!

The solution is to short circuit **C** to **D** permanently and then short circuit **A** to **B** at the same time.

As I mentioned before, there is a need for extra energy in order to activate air electrons effectively. In reality, this is not exactly what will happen as the arrangement is complicated. The negative energy enters this device in an unusual way. The extended Tesla Bi-filar Coil will supply reactive energy in enormous amounts because there are common characteristics between reactive energy and negative energy. Reactive power is a kind of electrical power, measured in volts-amps-reactive ("vars") which cannot do work in its present form. For sinusoidal waveforms, the formula for reactive power is

$$\textbf{Reactive power} = \mathbf{V} \times \mathbf{I} \times \sin(\theta)$$

The reactor coil in our system is a shortened version of a normal parallel L/C circuit. The output energy received in the reactant coil has to be a reactive energy because of the presence a coil in parallel with a capacitor. The phase angle between current and voltage is 90 degrees, and so, the active energy in this case will be zero.

$$\textbf{Active power} = \mathbf{V} \times \mathbf{I} \times \cos(90 \text{ degrees}) = 0$$

The system acts like a negative resistor, it's a **non-dissipative** system because it's an energy absorbing system:

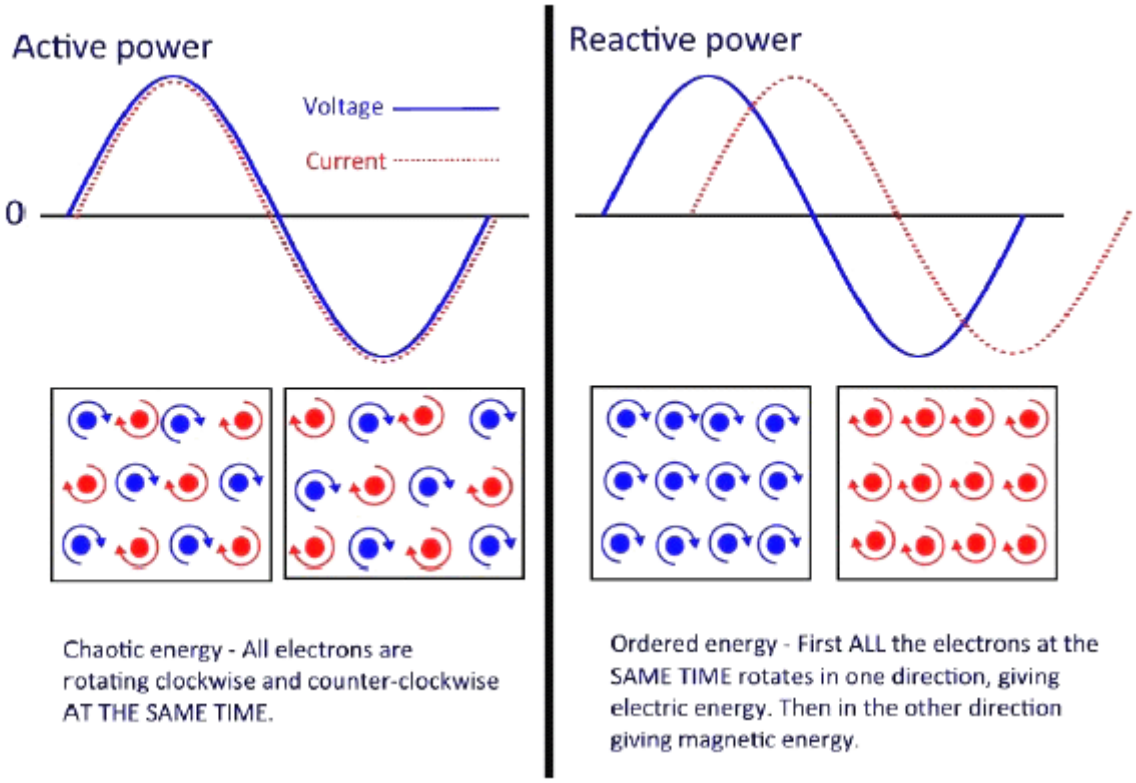


Fig.10

The reactive energy in this system is an advantage. The energy is ordered and so we can easily achieve **super-conductivity** at room temperature!

As shown in Fig.10, active energy is a **chaotic energy** and so it is not easy to get rid of resistance. The important point where we need to focus our attention is where the voltage has a maximum value the current will be totally absent. Negative energy is a kind of electric **power** obtained when sparking (on-off) from a high voltage direct current (Fig.11).

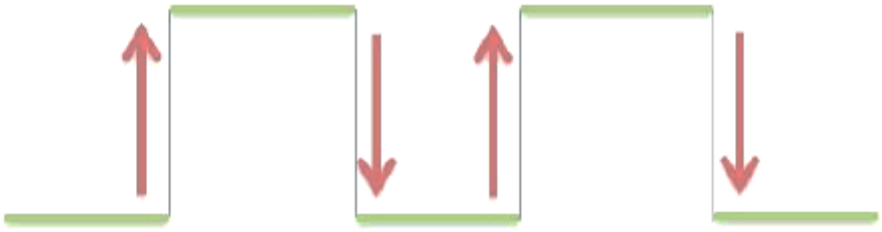


Fig.11

This is why we need a high-voltage diode in our power supply circuit. The resonance in a normal parallel L/C circuit does not require this.

If we connect a spark-gap between point **C** and point **D**, then the negative energy will be drawn into our system with the same speed as positive energy! The symmetry between the magnetic and electric sides in our positive energy reality will open the correct door for the undetectable negative energy reality!

Initially, when the capacitor starts to discharge, the current increases but the induced rotating electric field will tend to keep the voltage at a high value. The capacitor discharge through the spark-gap (which requires a large amount of voltage), the current flow does not start immediately. Initially, the current increases but the spark does not occur. This pushes the voltage up higher (behaviour which is known in parallel L/C circuits), then the current increases to a high value very quickly, while the voltage is drawn down to a level which cannot sustain the spark. When the capacitor is totally discharged the current flow through the spark-gap is at its maximum value. Consequently, the extended Tesla Bi-filar Coil produces a square wave rather than the expected sine-wave which is produced by an ordinary parallel LC circuit. The square wave produced when the spark occurs, contains waves of all frequencies and so, even if the time during the spark is short, there will still be thousands of oscillations in that time. I know that it is not easy to visualise this, but it is what actually happens.

This special genius design solves the most difficult problems in cold electricity, due to it's reversed behaviour compared to hot electricity. Cold electricity prefers to flow along materials which we consider to be insulators while hot electricity prefers to flow along materials which we consider to be conductors.

According Tom Bearden, with negative electricity the capacitor performs in the way that a coil does with hot electricity, and the coil will perform like a hot electricity capacitor (Fig.12).

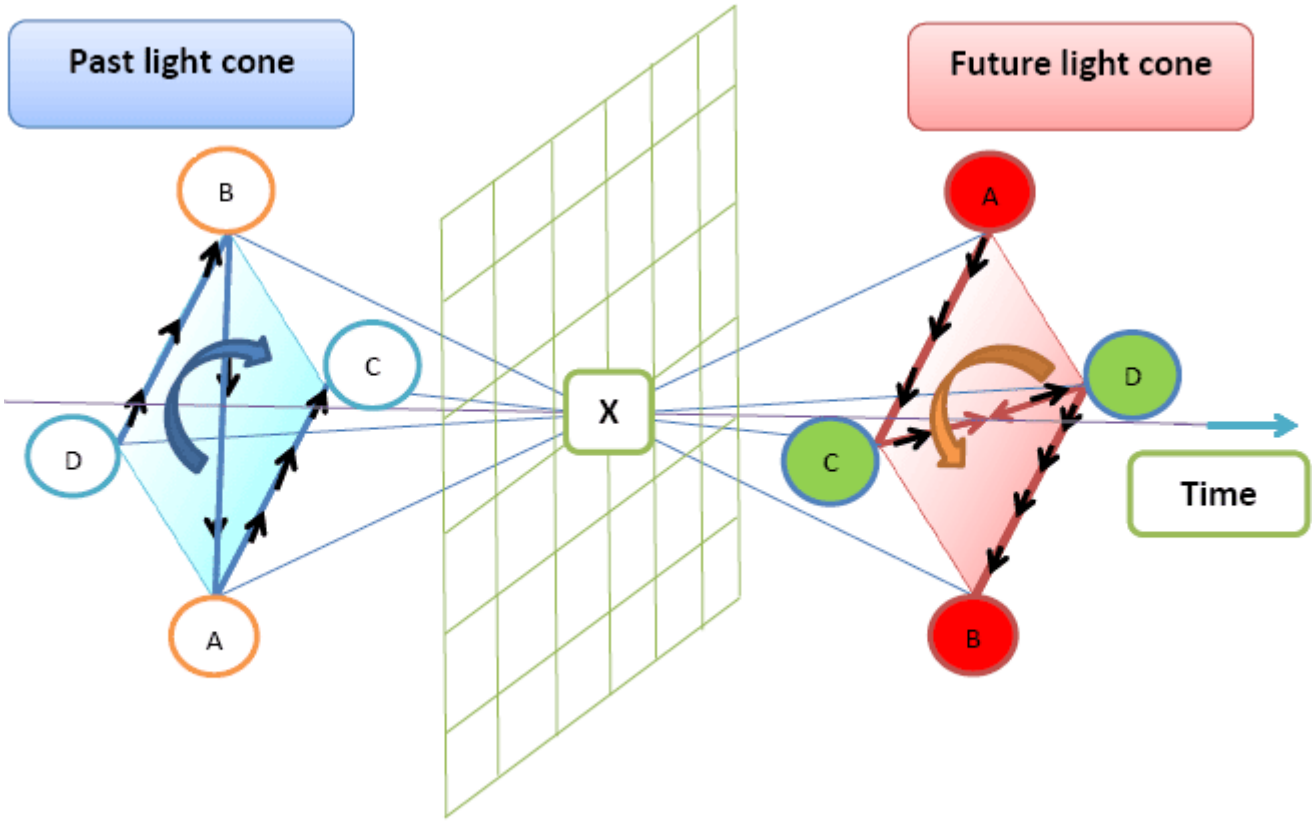


Fig.12 (the capacitor is discharging across itself)

the above illustration is an attempt to understand how cold electricity work but it's better to take a look at Floyd sweet explanation of his VTA device in the Appendix (page A-1209) :

it is important to note that so long as positive energy is present in a **positively-flowing time regime**, then unity and over-unity power gains are not possible. The summation of the losses due to resistance, impedance, friction, magnetic hysteresis, eddy currents and windage losses of rotating machinery will always reduce overall efficiency below unity for a closed system. The laws of conservation of energy always apply to all systems. However, **the induced motional E-field changes the system upon which those laws need to be applied**. Since the vacuum triode operates in **more than four dimensions** and **provides a link between the multi-dimensional reality of the quantum state and the Dirac Sea**, we are now dealing with an open-ended system and not the "closed system" within which all conservation and thermodynamic laws were developed. To achieve unity, the summation of all magnetic and ohmic losses must equal zero. To achieve this state, **negative energy and negative time need to be created**. When this is achieved, **all ohmic resistance becomes zero and all energy then flows along the outside of conductors in the form of a special space field**.

The above explanation describes the VTA device but also demonstrates the work of the Resonance Energy device of Don Smith !!

The correct model of capacitor is the extended Tesla Bi-filar Coil because it provides a link between the multi-dimensional reality of the quantum state and the Dirac Sea of negative energy. From a positive energy perspective, **AB** describes the coil while **CD** describes the capacitor, **but** the coil will transform into capacitor **AB** in the region of negative energy; and in the same area the capacitor will transform into the coil !!

How can we explain this physically!?

The energy equation and Dirac's equation call for both positive and negative energy. Thus they are symmetrical with respect to energy, as are the forces of physics positive repulsive forces increase positive energy, while negative attractive forces increase negative energy. According to the modern kinetic theory of mass-energy, negative energy would merely be a **vibration of charges at right angles to our ordinary dimensions in an "imaginary" direction**. (Fig.13).

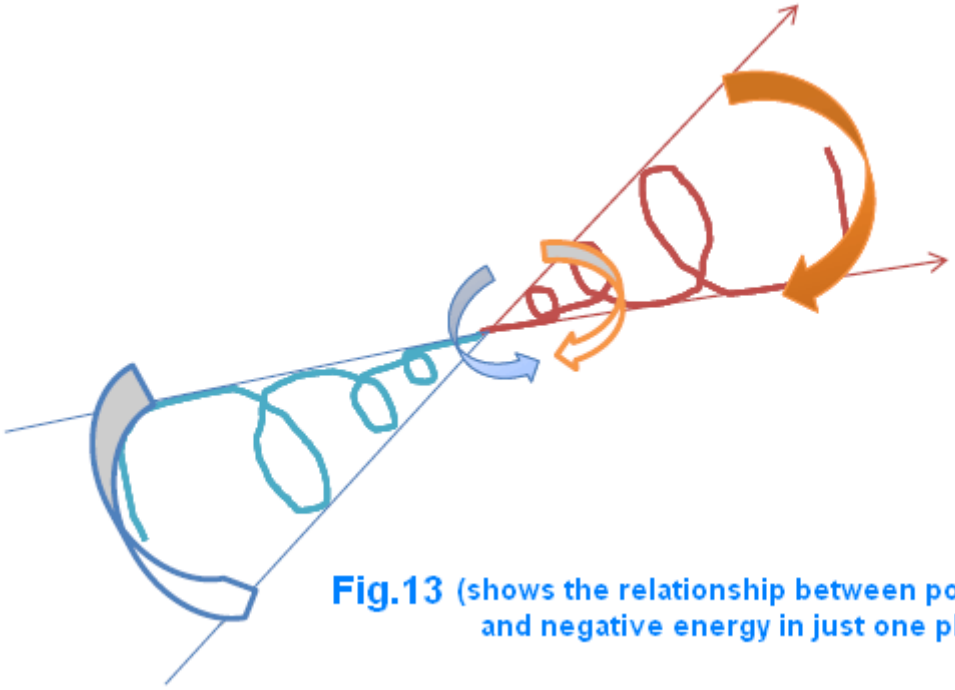


Fig.13 (shows the relationship between positive and negative energy in just one phase)

To understand Fig.13 correctly, we need to remember that we are confined by our positive energy reality; in our extended Tesla Bi-filar Coil we need the negative energy in our positive side, the spark gap in CD position is the flexible place for both positive and negative energy to be combined.

Tom Bearden has an important book entitled "Energy from the Vacuum". The following text is on page 236 where he explains the specification of a true negative resistor:

The true negative resistor is an open dissipative system ***a priori***, and equilibrium thermodynamics therefore does not apply. Instead, the thermodynamics of open systems far from equilibrium applies. The negative resistor freely receives energy from outside the system (from the environment), and "dissipates" it in interception and collection actions inside the system, to freely increase the available potential energy in the system.
In circuits, the main characteristic of a negative resistor is that the environment freely furnishes some excess energy to (i) power the load, and/or (ii) move the current back against the voltage, particularly when shunted across the back emf region of the source dipole. The *operator* does not have to furnish this excess energy dissipated to propel the current backwards or dissipated to power the load!

The true negative resistor in our system is the blue extended TBC where negative energy moves the current back against the voltage; this power will charge the coil electrically if it's acting as a capacitor!

The electric current in the negative energy region (past light cone) works in a reverse manner compared to the electric current in positive energy region (future light cone), the symmetry between the magnetic **energetic** side

and the **physical** electrical side will curve space for the negative energy power to enter our device through our **flexible** spark-gap zone which represent a one dimensional capacitor **X**, the negative energy which can be represented by attraction forces will find its way through the spark-gap to increase the electric energy through **AB**, the system will continue its divergence and the clockwise rotation of the negative energy current will increase the counter clockwise rotation of the hot electric current (potential energy), this tends to amplify the power between point **A** and point **B** which represents the voltage gained by this virtual current!

We are examining the first instant when the capacitor discharges across itself (transform into coil), the curved **AB** space will transform into one point in the negative energy sea which is the capacitor in our imaginary extended TBC.

Similarly, when the magnetic field collapses and charges the capacitor across **CD**, the magnetic field returns to its background level, the space-time continuum is reversed by the fields which are produced in the presence of excited coherent space flux. These quanta have been attracted from, and ultimately extracted from, the virtual vacuum, the infinitely inexhaustible Diac Sea (from Floyd Sweet papers in the Appendix)

Like charges have repulsion behaviour on the positive energy side, while they attract each other in the negative energy region. This information is essential for an understanding of the nature of the negative energy sea.

When the turn comes (second time period) for our capacitor to be charged again with opposite polarity, the system will diverge toward negative energy space to close the loop in that space !! The current passes from **C** to **D** to charge the capacitor but in the virtual dimension it starts from **D** and finishes in **C**. This power will charge the capacitor magnetically if it's a coil.

As you see there is positive real energy and undetectable negative real energy. I think that Don Smith preferred the name *ambient background energy* as opposed to *Zero Point Energy* because there are two regions from which we can take power, namely, over ambient background energy and below ambient background energy.

At this point, we can understand why cold electricity prefers insulators rather than conductors. This kind of power is capable of running in an imaginary dimension **parallel and reversed** relative to our ordinary, familiar dimension. But… according to Floyd Sweet; when run in parallel with positive energy however, cancellation (annihilation) of **opposing** power types occurs. This has been fully tested in the laboratory.

This applies to the chaotic positive energy flowing regime time when voltage electrons (ccw rotation) and current electrons (cw rotation) run together at the same time, our Reactive Energy resonance system works in harmony with negative energy, our previous study allows us to draw us some important conclusions:

For the first time period, we have (C discharging across L);
+ energy increase magnetic energy ---- I
- energy increase electric energy ---- V

For the second time period we have (L charging C);
+ energy increase electric energy ---- V
- energy increase magnetic energy ---- I

since magnetic energy is the current and electrical energy is the voltage and because they are out of phase (reactive energy), positive energy will work in harmony with negative energy and no cancellation will appear.

Our extended TBC is a very important device not only because it can supply unlimited electric power but it give us an exceptional opportunity to understand the way that energy flows inside our universe!

When the device oscillates it produces cold electricity and hot electricity, this means that the device is able to relate to both positive and negative energy. The flow of energy has two directions; from positive to negative and vice versa. Let's think about the physical way in which things work.

In this analysis I'm trying to explain some deep physical aspects about positive and negative energy. As illustrated in Fig.12 and Fig'13, it is important to consider energy flow in relation to time. These two pictures just represent my thoughts and understanding of the subject.

Cold electricity has the ability to produce an electric response when it interacts with metal surfaces. This can also be seen in the Casimir Effect where two non-magnetic metal plates, which are not carrying an electrostatic charge, are suspended very close to each other. The plates do not hang straight down but move towards each other.

Cold electricity has the ability to produce an electric response when contacting metal surfaces because it's able to ionise empty space. In our extended TBC, when the spark occurs, we are actually colliding the space-time field through one point (the spark-gap).

The space-time field is, in my opinion, the space where negative and positive energy exist together, they exist together but cancel each other out due to a constant ratio. If we take a mass with a gravity field around it, and we move the mass and create a mass current, a new field is also created. It is a different kind of gravity field with no source and no sink, when the speed of the mass increases, then the created gravity field increases also. If the mass reaches the speed of light, then this means that it has the value $E = mC^2$ as a positive energy. mC^2 is the maximum value exchange between positive and negative energy admissible by zero point fluctuation (ZPF) for that mass to exist the way it does in its time-space field, the mass has two options to reach the speed of light:

1. It will transform into exotic matter.
2. It will break the time-space structure.

The only places which provide these two conditions are black holes. Black holes exist in the centre of galaxies which provide the rotational energetic symmetry between the mass and gravity field - see Fig.14, Fig.15 and Fig.16.

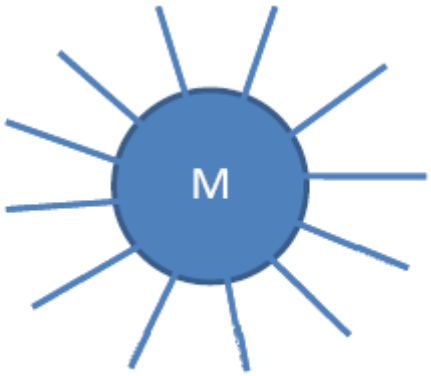


Fig.14 GRAVITATIONAL FIELD

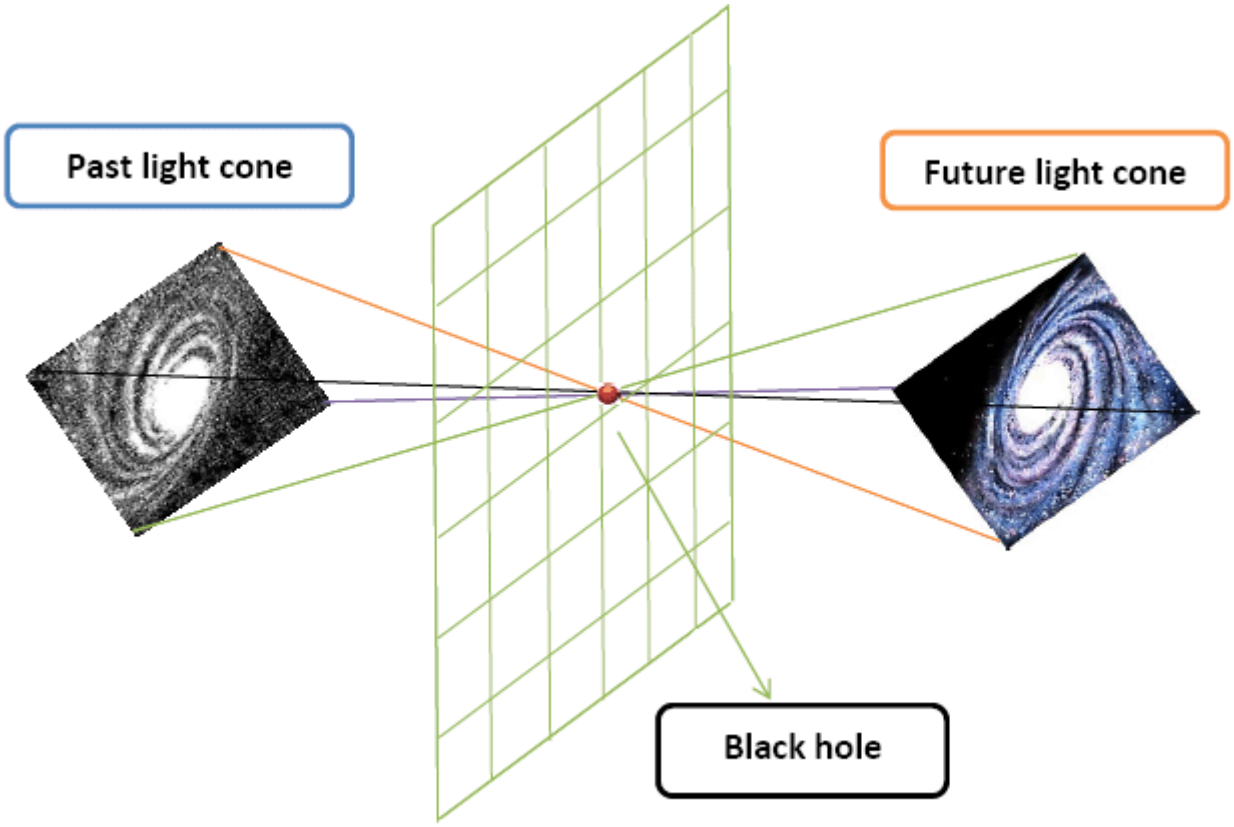


Fig.15 notice the similarity between the black hole and the spark gap

Fig.16 The massive spiral galaxy NGC 1365 has a huge black hole in it's heart, spinning at nearly the speed of light

For a spiral galaxy to maintain its shape with cosmic dimensions (the diameter of galaxy may exceed 100,000 light years) there is a need for a negative energy to be the underlying energy for space-time throughout all of that galaxy. That negative energy has to transport virtual particle instantly! The transformed physical matter (including space-time!) provided by black holes furnishes an excess of positive energy in the galaxy, providing stability and symmetry. Black holes are not a fracture in space-time but they are essential.

The above explanation will help to clear the way for a better understanding of the nature of electric energy. This explains why a sharp positive-going DC electric pulse interacts with negative energy to produce cold electricity which is an instant response from the negative energy sea. This response does not start from the spark-gap, but it ends in it!

The negative energy will rotate to finish in the spark gap, this will squeeze the space-time to provide excited coherent virtual particles which in turn produce electronic responses when contacting a metal surface. From my point of view, the electronic responses created in metal surfaces have a magnetic angular momentum. Cold electricity is able to charge a capacitor to a much higher voltage than the capacitor's voltage rating, even if the capacitor's rated voltage is low. The question which springs to mind is; do electric fields inside a capacitor charged with cold electricity really exist?

If the answer is yes, then why isn't the capacitor destroyed? In my opinion, it is because the electronic responses caused by cold electricity have magnetic angular momentum instead of electric field lines. I suggest that the presence of magnetic field lines between the positive and negative plates of a capacitor charged with cold electricity are actually as shown in Fig.17.

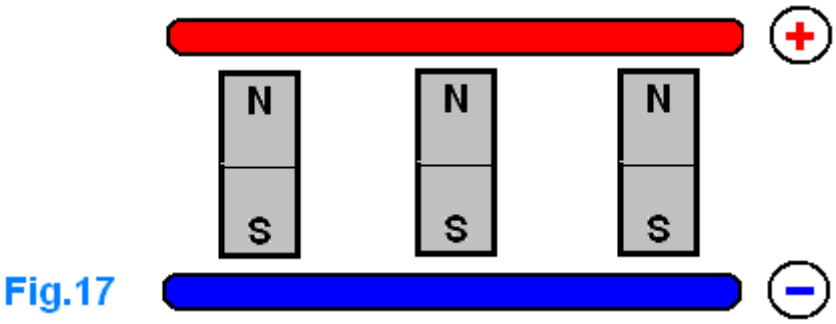


Fig.17

When the spark gap impacts on space-time (that is, when the spark occurs) the response from the negative energy sea looks as if it should neutralise the excititation created in the positive energy side (Fig.18). We can't detect the movement of negative energy, and so we only see the impact which it has in our positive energy reality.

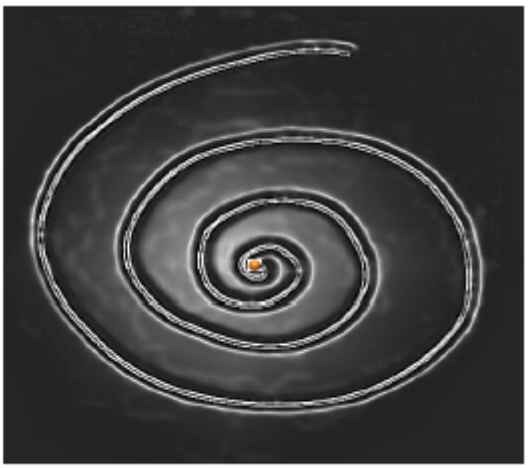


Fig.18

The small red ball in the above drawing, is the spark gap which is the door for negative energy to enter our positive energy reality; the negative energy sea will react both before and after the spark occurs.

Referring back to Fig.14, before the spark-gap fires the negative energy will rotate starting from the spark-gap to neutralise the excitation created in the positive energy side (Fig.19a) and when the spark-gap finishes firing, the negative energy will end in the position of the spark gap (Fig.19b).

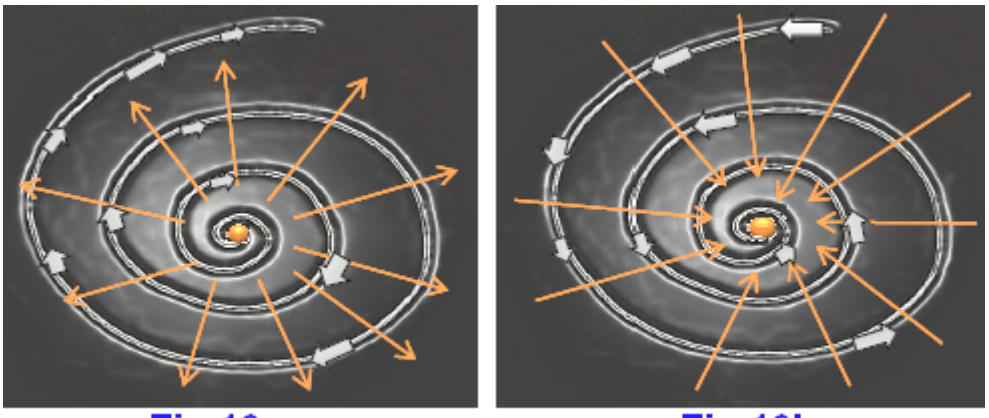


Fig.19a Fig.19b

The Bloch wall area in an ordinary permanent magnet, is the area of electron separation. Let's see how this happens in our extended Tesla Bi-filar Coil. During the first time period, when the capacitor starts discharging across itself to become a parallel L/C circuit, point **A** will provide a maximum voltage while point **B** is the maximum current. The current flow starts from point **A** and finishes at point **B**. The system is now producing magnetic energy and because of the magnetic field increase, the electrons will start from point **B** and flow to point **A** which causes clockwise rotation to neutralise the counter clockwise spin of the voltage electrons, and cold electricity will charge the coil electrically if it is acting as a capacitor, and it will push the current to go back against the voltage by providing a magnetic angular momentum (the clockwise rotation shown in Fig.19a) at point **X** the result is to turn back the voltage electrons, causing strong initial potential electrical energy which increases the electrical energy. Current in cold electricity is the equivalent of voltage in hot electricity. The Bloch wall is the place where negative energy interacts with our E-TBC, in other words when the spark-gap fires, the current will not start immediately because the negative energy will supply a virtual current by providing a CW rotation in the Bloch wall area **X**. This virtual current is a compensator of the real current but it will not absorb the power from the voltage electrons which increase the available potential electrical energy. All this happens before the real current increases to provide the magnetic energy.

Caution: Please be aware that high voltage capacitors have dielectric recovery which stores the electric field for a long time. High voltage capacitors need 5 minutes or more to discharge completely.

Energy is everywhere and in enormous amounts ready to be taken for free. When we do that, we do not reduce the available power because the universe is full of energy, the energy in our universe is the source, physical matter is the energy in a visible form and the energy is in invisible matter form.

The presence of the negative energy sea beside our positive energy reality, raises an important question, namely, why they are separated when they could be united? They are separated to let our universe exist in the way that it does. Negative energy serves our existence because it's designed to be the **under** background energy level for our **foreground** positive energy reality. Our existence is a thin part between those two energetic oceans. Negative energy is extremely active until the point when it appear to be nothing!

We now need to explain another important behaviour of our extended Tesla bifilar coil, namely, it's super conductivity at room temperature.

One of the puzzles in this device is its capability to equal the voltage with the current. The wire of the coil **AB** can receive ambient electrons because it's the surface for the capacitor **CD**; let's examine this surface in Fig.20:

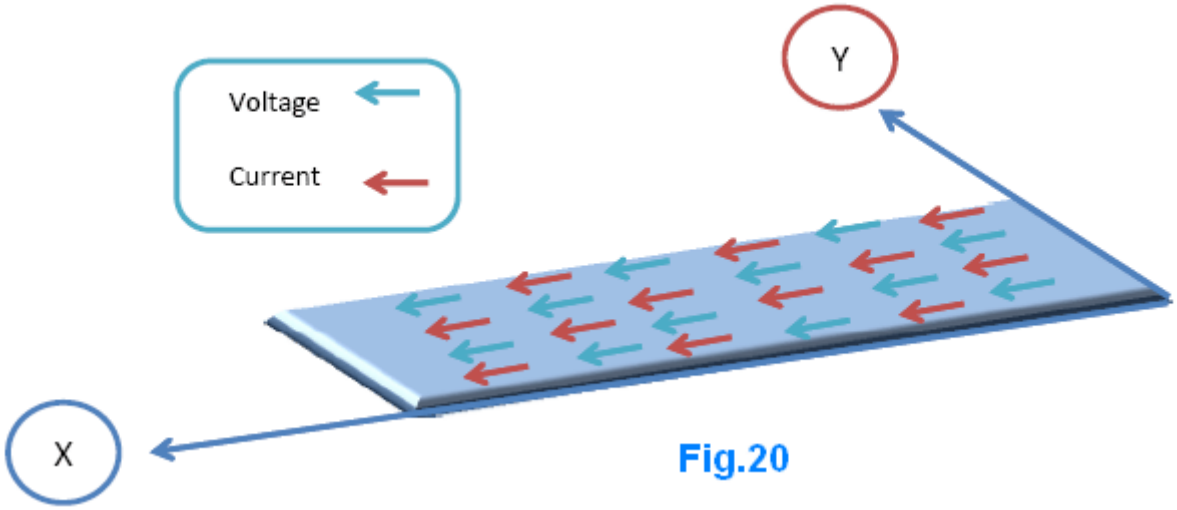


Fig.20

When the capacitor discharges across itself to become a parallel L/C circuit, the induced rotating electric field (with the help of negative energy) creates a difference of voltage between adjacent plates, this voltage according the Gauss law causes **new electrons** to be present in the system.

$$\oint \vec{E} \cdot d\vec{A} = \frac{Q_{enclosed}}{\varepsilon_0}$$

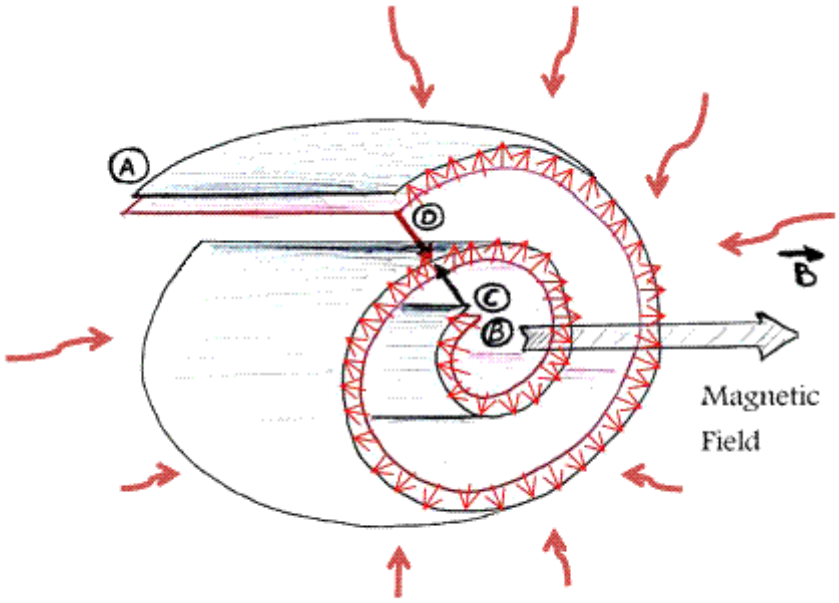


Fig.21 (ambient electrons can be taken from the other plate as one of them has to lose electrons)

When ambient electrons enter our system (Fig.21), they increase the power gain across the Y axis (Fig.20). When the current flows inside the extended TBC, parallel currents will be added while parallel voltage will be the same in the Y axis, while in the X axis serial voltage will be added while serial currents will be the same!!

On the Y-axis: sum (I) equals (V)
On the X-axis: sum (V) equals (I)

This system treats the voltages and currents in the same way, the voltages and the currents are physically equal.

When this happens the device squares the electromagnetic flux and becomes a near-unity system in every process which will replicate the electric power according to the working frequency. This is a near-unity system due to the super conductivity at room temperature where the electrons don't face the usual reduction encountered in an ordinary parallel L/C circuit.

A normal L/C circuit cannot produce the super conductivity at room temperature because the exchange between the electric power and the magnetic power must lower one of them in each process. In our extended TBC they are combined in such a way so as to amplify the power in every process, and so the total available electrical energy in each cycle is twice the power available in a charged capacitor which can be seen from the following relationship:

$$\textbf{Power} = 0.5 \times C \times V^2$$

(consider the similarity between the magnetic and electrical energy in a resonating parallel L/C circuit)

Here, I need to explain the importance of the reactive electric power in the Resonance Energy device, in an alternating electric system where the voltage and current go up and down at the same time (Fig.10). Only active, real power is transmitted and when there is a time shift between current and voltage both reactive and active power are transmitted. When this time shift is 90 degrees (π/2 degrees) the transmitted real power will be zero as discussed above. This does not mean that there is no power, but it does mean that **we cannot use this power in this alternating form,** we have to transform it into Direct Current so that both current and voltage are united.

Reactive power looks like a skipping rope (Fig.22):

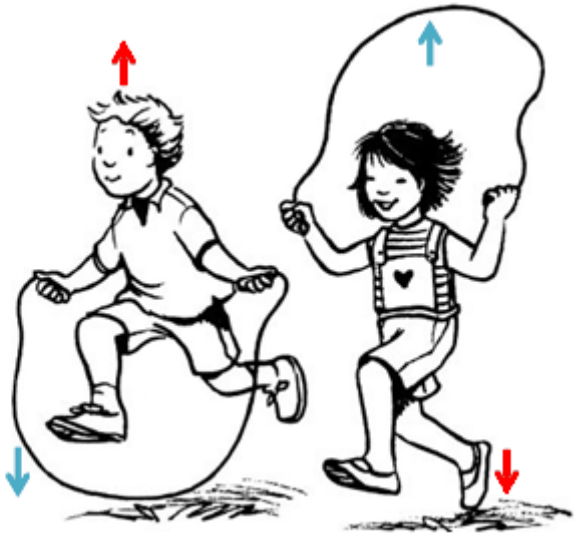


Fig.22

Imagine the **voltage** to be the rope and the children's bodies are the **current**. Reactive power looks like a skipping rope, the active power will not let the kids' bodies move correctly. Reactive power is an essential part of the Resonance Energy device, and a skipping rope is good example which shows how kids go up and down without any problem. This sort of movement exists in our device.

The separation between voltage and current in the Resonance Energy device is crucial for producing and cloning electric power at radio frequency speed. A proper method of collecting and converting the huge available electric energy is needed.

The example given in Fig.22 is important when planning to collect and convert the available electric power. If we simply use a step-down transformer it is highly likely that we will alter the current which will reduce the gained power. With reactive energy, when the voltage is high the current is low. A step-down transformer will lower the voltage but it can't amplify the current as expected! In a normal transformer we amplify the current depending on the available active power (V x I):

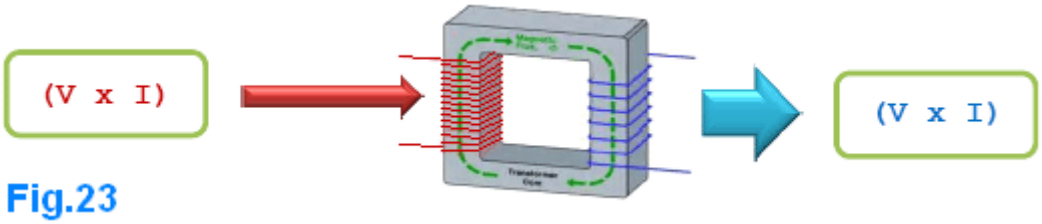


Fig.23

Physically (Fig.23) the **electromagnetic flux** inside the transformer has two components, the electrical component **V** and the magnetic component **I**, for successful transfer of electric power from the primary to the secondary, both of them are needed at the same time. In our case, when **V** is high the product (**V x I**) is low due to the time shift, even if the available power could achieve megawatts!

Another factor which we have to take into consideration, is the high-speed needed to replicate the power, using a transformer to lower the voltage imposes the need for a special transformer core which is able to respond at radio frequencies. These facts have to be taken seriously if we want to collect the available energy effectively.

Fig.24

Personally, I prefer to improve on the use of high-voltage diodes as shown in Fig.24. It is better to use a diode bridge constructed with **fast recovery** high-voltage diodes. Fast recovery diodes have the ability to return to their blocking state very quickly, making it possible for the other half-oscillation to be accumulated in the high-voltage capacitors, each cycle (**up** and **down** in Fig.25) will give power similar to the power available in a charged capacitor given by the following equation **½ C V^2**.

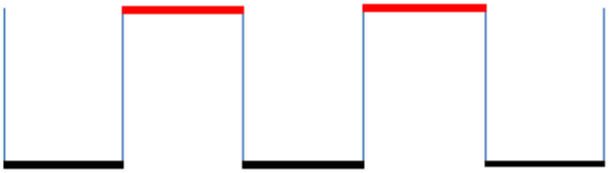


Fig.25

The amazing, energetic behaviour of the extended Tesla Bi-filar Coil make it totally different from an ordinary parallel L/C circuit. Our extended TBC gives twice the frequency of an equivalent parallel L/C circuit. This means that if you form the same inductance with the same capacitance in an ordinary parallel L/C circuit, then that will produce only half the frequency that the same combination produces with an extended Tesla bifilar coil form!

I have not been able to verify this because I don't have an oscilloscope or a frequency meter. That, of course, is no excuse for not thinking about how the device will act, so the following analysis is an attempt to imagine the energy equation given by Mr. Donald Smith as:

$$\textbf{Power in one second} = 0.5 \times C \times V^2 \times F^2$$

To simplify things, let us analyse only the voltage. When the capacitor discharges across itself to become a complete parallel LC circuit, at that instant, the magnetic field reaches it's maximum value. What makes this system different, is the induced rotating electric field. This field will instantly charge the capacitor with the opposite polarity before the induced current resulting from the collapsing magnetic field can do it. As we learned before, this is the key to energy amplification.

Resonance is the key for energy multiplication, our extended TBC acts like one device, so the passage from the positive cycle to the negative cycle takes no time. In other words, the device has the ability to change it's direction of charge instantly. The yellow zone in Fig.26 is absent (compared to a normal parallel L/C circuit), when the energy cycles up and down (Fig.25) the device gives twice the power available in the capacitive side of the extended TBC.

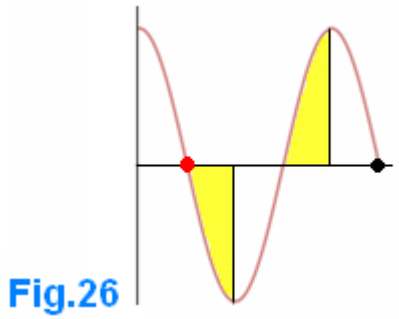


Fig.26

(the cycle in an ordinary L/C circuit is from the red point to the black point. In our device, that is only half of the time which explains why an E-TBC produces twice the frequency when compared to an ordinary parallel combination)

The power is each cycle will be given by:

$$\frac{1}{2} C V^2 \times 2 = C V^2 \ldots\ldots\ldots\ldots\ldots\ldots 1$$

Since the frequency will affect both the voltage and the current we will examine the equation number 1 as follows:

CV x V, the product CV is the quantity of charge available in a charged capacitor, if we divide it by 1 second this will give us the current since Q/T is the current in one second. Let us suppose that the frequency is 3 Hz.

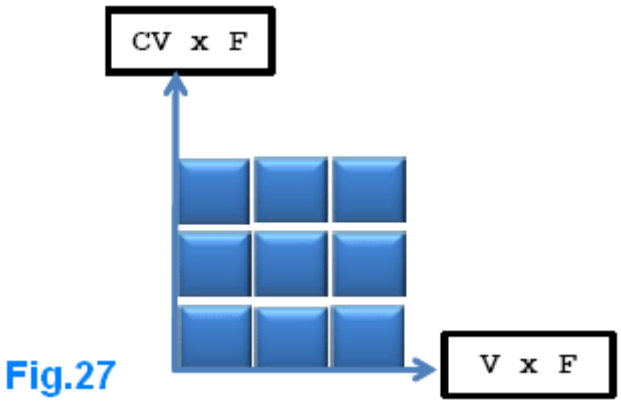


Fig.27

From Fig.27 we can see that the total available power is proportional to 9 which is the square of the given frequency. Each cycle has the power of C V^2, the number of cycles in one second gives us the frequency; the frequency will replicate CV which is the current and this will give us CV x F and replicate the voltage by the product V x F, this analysis is the best explanation of why the voltage equals the current in this system, because CV x F is the available current and V x F is the available voltage! This seems strange; how could the product V x F be the available voltage since the result is very high since we are working with radio frequencies above 20 KHz?

The example given in Fig.23 will help us understand this. The energy formula gives the available DC power when converted from its alternating state; the **electromagnetic** flux will be squared causing amplification of both **current** and **voltage**. In each second, the available power is **CV x F** x **V x F**, the power obtained depends on the number of turns in coil L2, and the limiting factor is the product **V x F** which is a very high number in practical systems. Below this factor, the current is very high being the product **CV x F**!! This explains why a megawatt sized unit can fit quite easily on a breakfast table and it explains why this device is able to give any required level of energy.

The energetic formula of our extended TBC can now be written as follows:

$$\textbf{Power in one second} = C \times V^2 \times F^2$$

This equation gives the available power in watts when converted into high-voltage Direct Current. When the device oscillates the power obtained is **pure reactive energy, Volt-Ampere-Reactive (VAR) is present while active electric power (W) is absent in this dynamic state**, Fig.28:

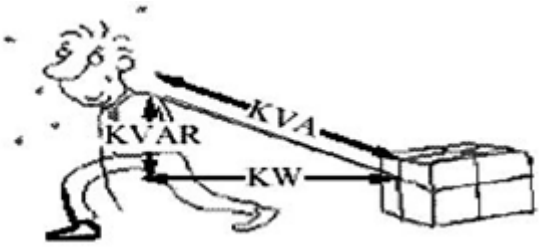


Fig.28

(this explains why reactive energy isn't capable of doing real work in it's alternating state - when KW=0 then KVAR has no effect)

In practical terms, the extended TBC is just a high-voltage capacitor which has the ability to let the current penetrate inside itself, so it has both magnetic and electric specifications.

Practical section

A free-energy device is something which is fascinating, being hasty in wanting to build and test one is common, but that is not good. **High-voltage with high current isn't a game!**, your first mistake may well be your last. If you decide to build this device in your house it is a good idea to use locks and keys and sticking a high-voltage warning symbol on the device is a sensible action.

I am not encouraging you to actually construct the device described here; the above theoretical information provided is the most important section. When the device is fully understood, then taking care when near it will be automatic. This device is a very special Tesla coil, when increasing the voltage in a normal transformer the current drops, **but here the current increases in the same manner as the voltage does!** This device has current equal to the voltage. Resonance will impact both the voltage and the current. The special geometrical design of the extended Tesla Bi-filar Coil, including the flexible position of the spark-gap, produces the needed rotational energetic symmetry between positive and negative energy. As we have already seen, the spark gap opens the door for a massive inflow of electrical energy to be present. **I personally got shocked indirectly from the L2 coil and I certify the risk of this device**.

For construction, the first thing we need is a source of high-voltage. The device can be fed using two different methods as described here. The first is the direct method where the high-voltage power source has the same frequency as the natural resonant frequency of the extended Tesla Bi-filar Coil. The second method is the indirect method; where there is no need to know the frequency of the reactor (active) coil.

The high-voltage power source is needed to feed the extended TBC which is the reactor coil (active coil), Fig.29 shows an easy to build oscillator:

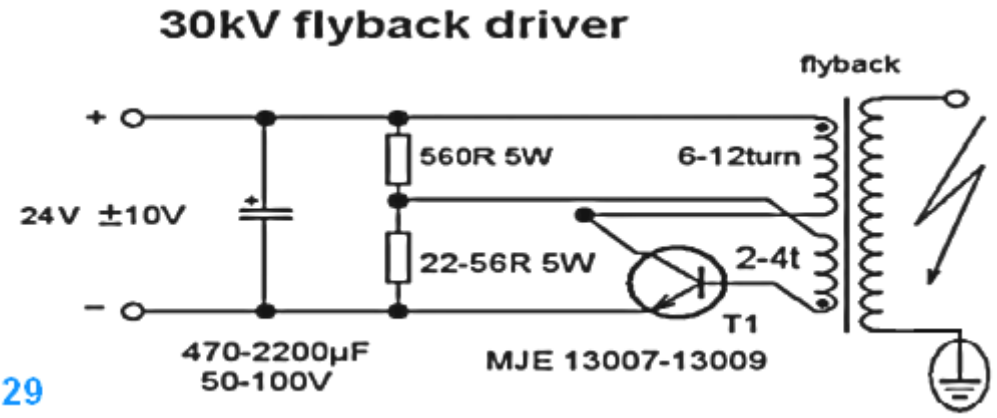


Fig.29

It's better to use a flyback transformer which has a high-voltage diode built into it. Flyback transformers are readily available and cheap. The above circuit diagram is for a flyback transformer, where a high-voltage power source enter our reactor coil via points **A** and **B** (Fig.30):

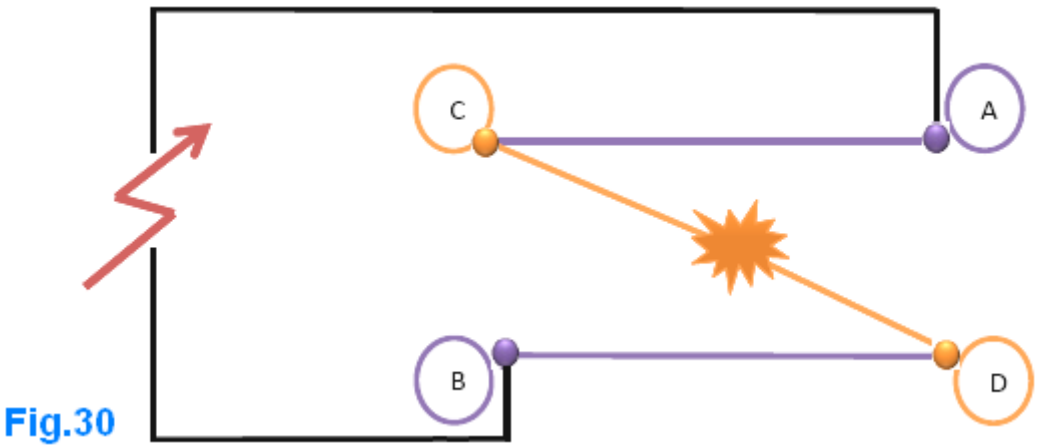


Fig.30

Initially, the capacitor charges up to the value needed by the spark-gap. When the voltage across the electrodes of the spark-gap reaches a high enough value, a spark occurs causing the resistance of the spark-gap to jump from a very high value to a very small value, short-circuiting any power coming from the power source until the natural resonance finishes. The capacitor transforms into a complete coil which has its capacitor built into it. The natural resonance of the TBC is assured when using this method, but it does have some disadvantages. The frequency produced by power supply/spark-gap combination has to be high enough to allow more power to be produced and this calls for a powerful power source. On the other hand, the voltage between the reactor coil **AB** will be **limited** by the distance between the electrodes of the spark-gap. This imposes the need for a large number of turns in the L2 coil.

The output current obtained is directly related to the available voltage between the capacitor plates which form the device between **A** and **B**. We need to bear in mind that the capacitor incorporated inside our extended TBC operates in a dynamic fashion where no displacement current exists.

The direct method of feeding the reactor coil with its own exact natural frequency is the best way to get the most available power, but this creates a real problem as it is not at all easy to find a high-voltage power supply adjustable over the range of frequencies which we want, especially frequencies above 200 KHz. We may require our extended TBC to work above 200 KHz, and for that, we may need impedance-matching capacitors (Fig.31).

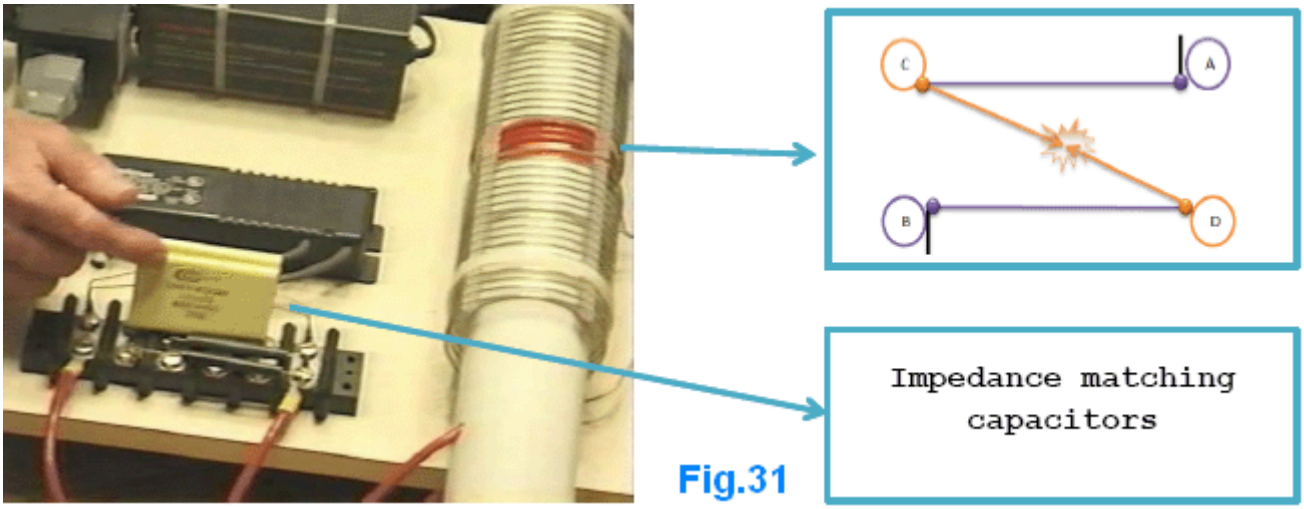


Fig.31

The two yellow capacitors seen above are for impedance matching since the working frequency in Don Smith's device was very high, requiring a neon-tube driver to supply it.

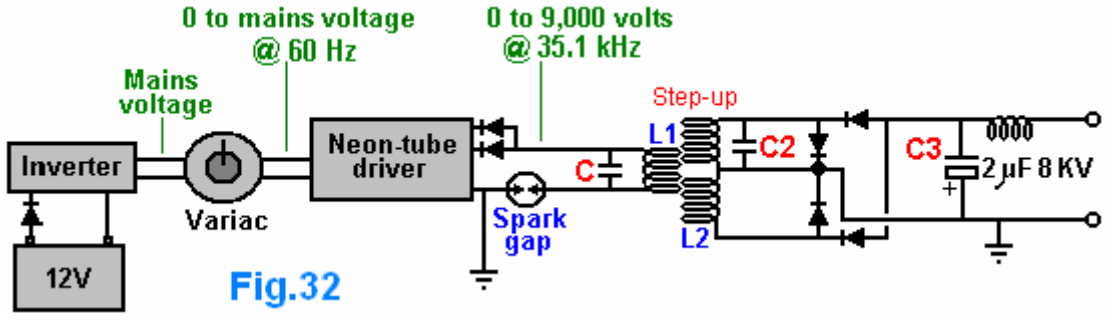


Fig.32

Impedance matching is simply the process of making one impedance look like another; in our situation it is necessary to match the load impedance to the source. For example, if the extended TBC resonates at 2.4 MHz, (this is the load), and the neon-tube driver operates at 35.1 KHz (this is the source) we need to add parallel capacitors to our extended TBC in order to make it resonate at 35.1 KHz.

In practice, you need to short-circuit the points **C** and **D** and measure the inductance of the E-TBC (L2 has to be in place for this measurement). After that, remove the short-circuit and measure the capacitance of the E-TBC. This gives you two values "C" and "L".

The resonance of the Extended –Tesla Bi-filar Coil is given by:

$F^2 = 1 / (\pi^2 \times LC)$ …… (a) The resonant frequency of an E-TBC is double

When you add a capacitor for impedance matching to the extended TBC, the resonance frequency will decrease with the following relationship

$F^2 = 1 / (4\pi^2 \times L(c+c^*))$ …… (b) where F is the frequency of the source

I believe that we can use the above equation to calculate the value of c* which needs to be added in order to achieve resonance. From equation (b) we could write

$$\frac{1}{F^2} = 4\pi^2 LC + 4\pi^2 LC^* \qquad \text{and so we have}$$

$$C^* = \frac{1}{4\pi^2 \quad L}\left(\frac{1}{F^2} - 4\pi^2 LC\right) \ldots\ldots\ldots\ldots\ldots\ldots \quad (c)$$

Using equation (c), we will be able to calculate the needed capacitor for impedance matching; the value obtained is in Farads, and that equation the frequency is in Hz and the inductance in Henries.

When you have the correct C* value and you power your device up (Fig.33), resonance will not be achieved instantly because the spark gap forms the full L/C parallel circuit only when it fires!!

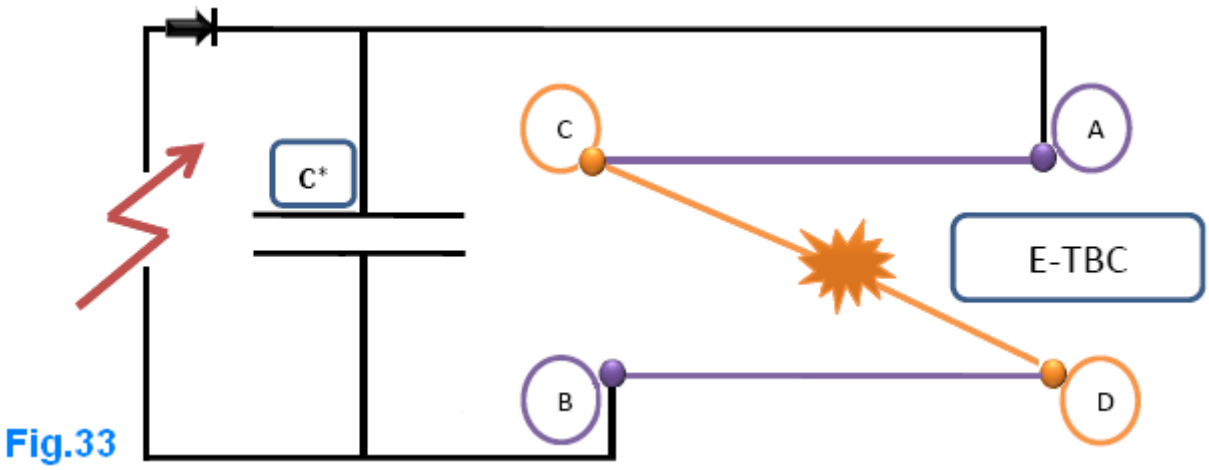


Fig.33

The first thing which happens is the charging of capacitor C*, after that the capacitor "C" of the extended TBC will be charged until it reaches the voltage needed to make the spark-gap fire. When that happens, the spark-gap has a very low resistance value, making the E-TBC fully formed. At this point, the incoming electric energy from the high-voltage power supply will find a load whose impedance matches it's resonant frequency. This, in turn, produces the maximum possible voltage across inductance L of the E-TBC. Further, the resulting electromagnetic field will increase cycle by cycle causing the device to resonate fully after a very short time.

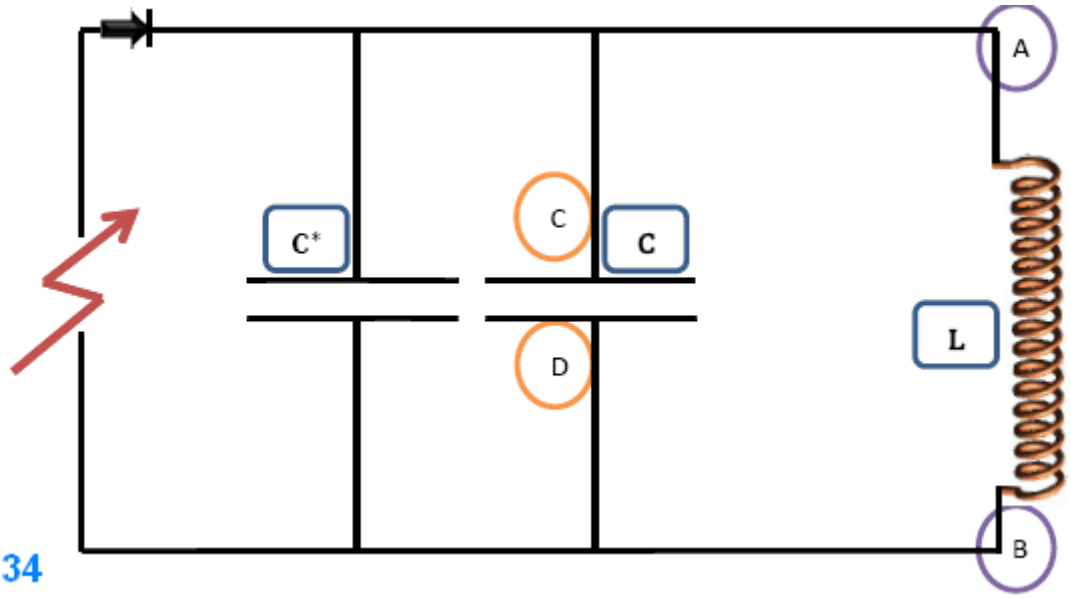


Fig.34

This is the equivalent circuit when the spark-gap fires

Constructing the extended TBC

The extended Tesla Bi-filar Coil is just a high-voltage capacitor which has a magnetic behaviour. So, it is both a capacitor and coil at the same time. Constructing this device is relatively simple. You need two lengths of aluminium foil, each 1.2 meters long (later I will explain the possibilities of modifying the characteristics of an extended TBC). Because it is a capacitor, you need 3 pieces of polyethylene sheet, each 1.3 meters long.

To construct a high voltage capacitor usually you need 2 pieces of polyethylene sheet but it's better to use 3 pieces since we are working with a high voltage (Fig.35), this will depend on your skills in constructing a high voltage capacitor.

You need welding rods to insure the electrical conductivity of the aluminium foils, Fig.36 indicates how to do that. Actually the best length and width of aluminium foil need some experimentation; you have to establish the positions of the magnetic and electrical points (Fig.37). **A** and **B** are the magnetic points (coil) while **C** and **D** are the electric points (capacitor).

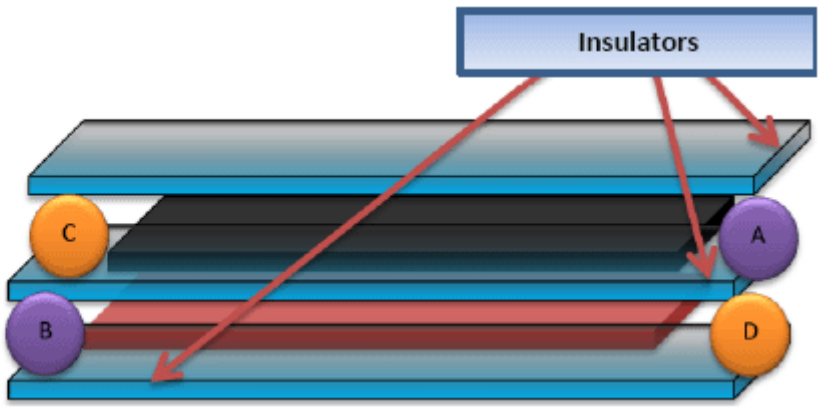


Fig.35

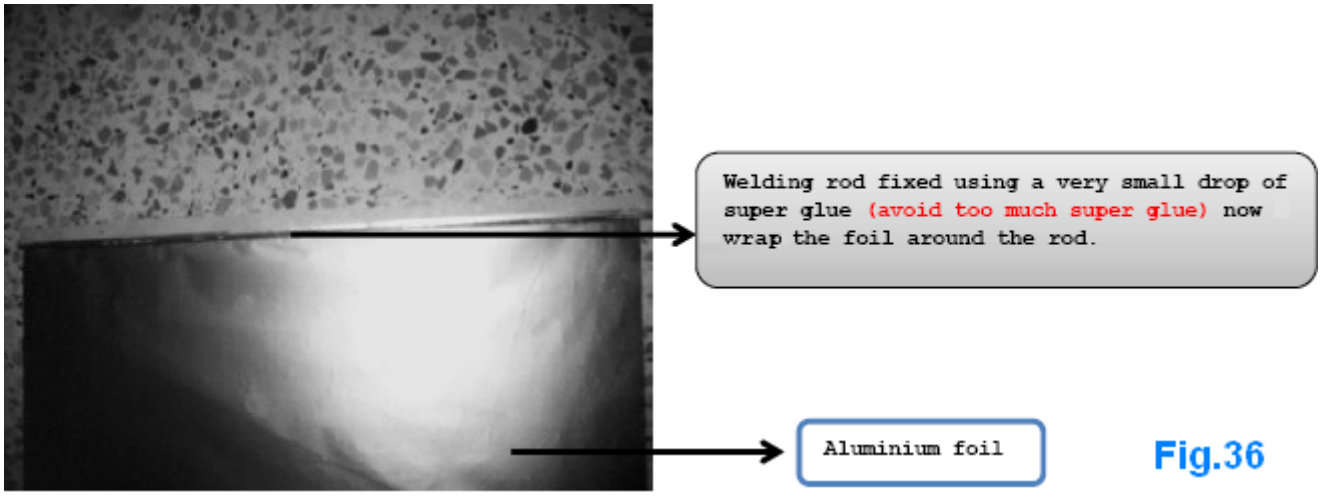


Fig.36

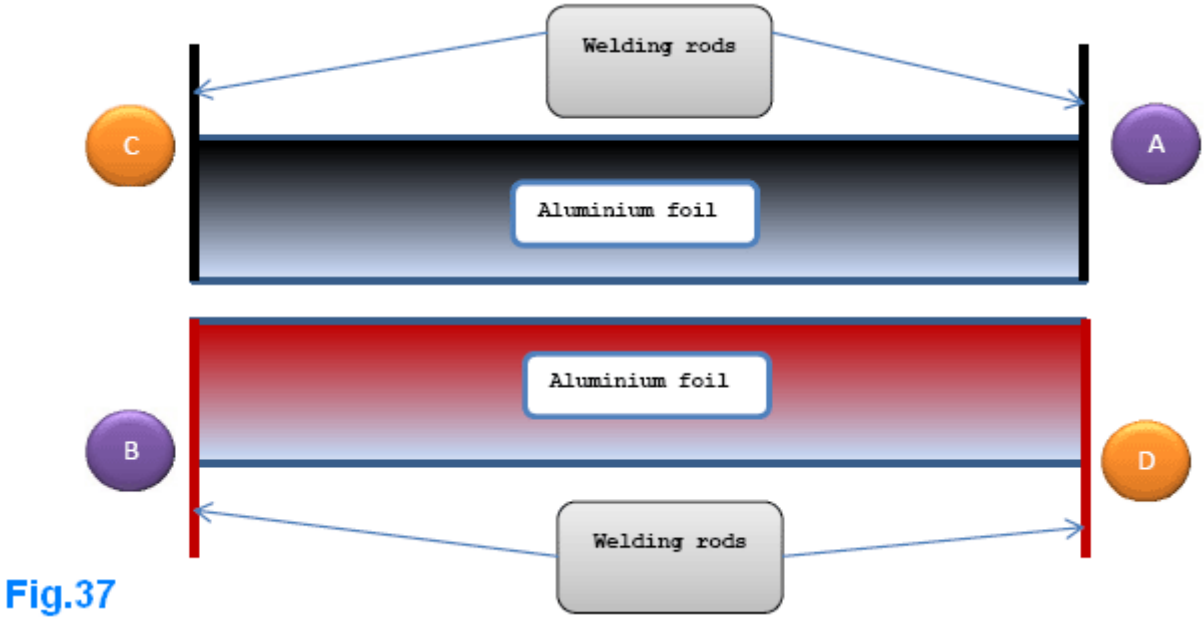


Fig.37

The E-TBC need a coil former for it to be wound on it (Figure 38)

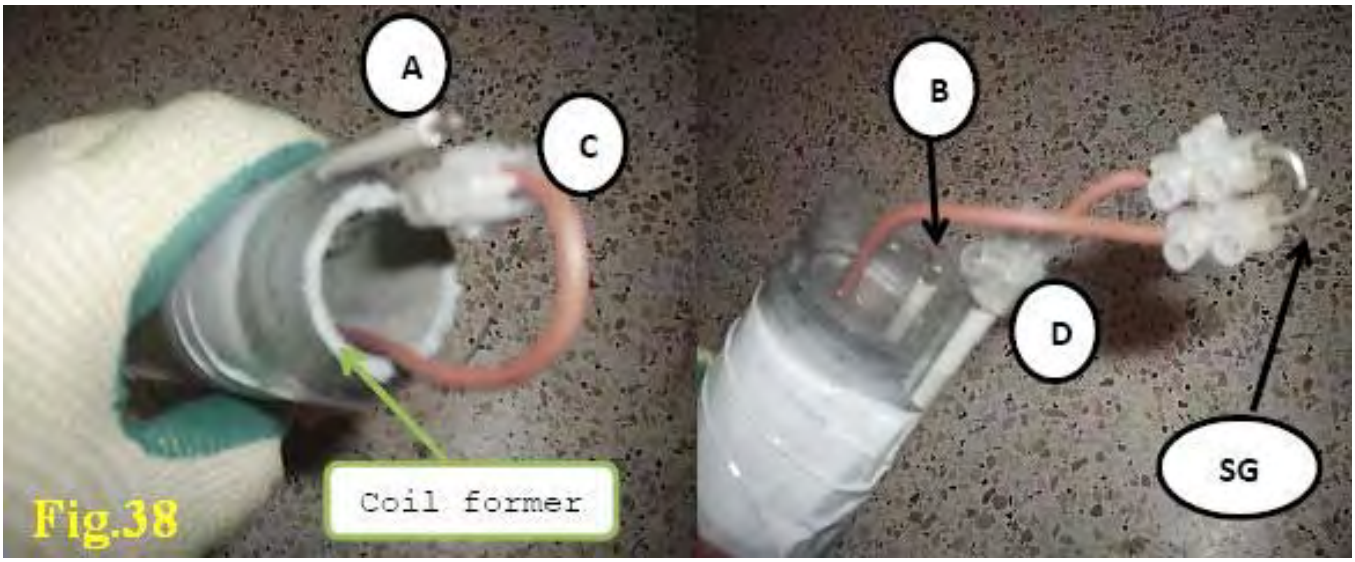


Fig.38

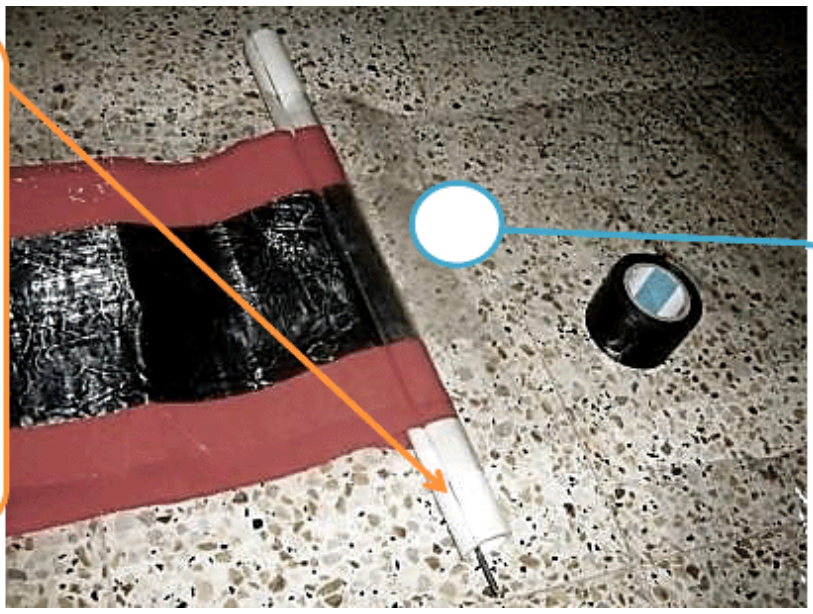
The following photo shows the dimensions used in the prototype (Fig.39)

The width of aluminum foil is 10 cm; the length of each capacitor plate is 120 cm

Fig.39

Internal arcing is a common problem with high-voltage capacitors (Fig.40)

*Notice the cutoff along the plastic tube; this will help in penetration the aluminum foils, the points **C** and **B** will be inside the tube but **A** and **D** will be outside, now firmly turn the plastic tube to form your **E-TBC.***



An internal arcing is a common problem in a high voltage capacitor; a good isolation is needed as well as a good insulator.

Fig.40

The best combination between the length and the width of an extended TBC has a major influence on the electric energy obtained at the L2 coil (the reactant coil). For example, the electric arcing experienced at the L2 coil using the E-TBC shown in Fig.41, was very weak, the length shown by the black arrow is very much greater than the width (orange arrow), this give a weak magnetic flux due to the small coil inductance, the coil inductance is very important because it will transform the replicated power into electromagnetic flux.

Fig.41

The reactant coil (L2):

The coil length is about 25cm as shown in Figure 42, the diameter is 6 cm, and the thickness of the wire is 1.18 mm (AWG #17 or swg 18) and the number of turns is about 200. Fig.43 shows some nice sparks from L2 coil:

The spark shown above is very strong but it can't burn through thin paper! This proves that the electric energy obtained is reactive, and so it can't do work as it is. Converting a high-voltage reactive electric power into Direct Current isn't easy to achieve. Working with a voltage over 10KV is really dangerous, in my device I had about 40KV of reactive power which to be converted.

Ideas:
To solve this problem let's think again about the energy equation of our extended TBC. The idea is to work with a step-down method rather than using the step-up technique.

The energy equation can be written like this: **Power in one second = CVF** x **VF**

CVF is the available current in one second since C is the capacitance value of the E-TBC, V is the voltage used, and F is the resonance frequency.

VF is the limit value when stepping up the voltage across the L2 coil, the gained electric power goes up proportional to the voltage value across L2, when achieving VF the total obtained power will be exactly: **$C V^2 F^2$** which is a very high power level. I know that this can be confusing, but this system has equal voltage and current. When stepping up the voltage, the current remains constant because it depends on the product **CVF**. We could increase the voltage by adding more turns to the L2 coil, when doing this the current is the same but the available electric energy will be given by: **Power in one second = CVF x V***

Where **V*** is the voltage across L2.

Using this new equation will help us a great deal in determining the voltage needed across the L2 coil in order to achieve the required electric power.

Example:
Imagine you have the following working conditions:

C = 10 nF
V = 30 KV
F = 100 KHz
You need a power of 30 KW, what L2 voltage is needed to achieve this power level?

Using the previous relation will give us:

$30{,}000 = CVF \times V^*$
$30{,}000 = 10 \times 10^{-9} \times 30 \times 10^3 \times 100 \times 10^3 \times V^*$
$30{,}000 = 30 \times V^* \Rightarrow V^* = 1{,}000$ volts

to obtain 30 KW you only need 1000 volts across the L2 coil. To achieve this power level you need direct resonance, feeding the E-TBC with a high-voltage only, without direct resonance, will not give this result since the voltage across the coil of the E-TBC will be limited by the electrode separation of the spark gap.

The voltage **V** across the E-TBC is very important here because the product CVF is the current obtained using the reactant coil (L2). The working frequency **F** is important too. Similar to an ordinary transformer, if we want to use the step-down method effectively, we have to think about using too many turns when winding the E-TBC. When designing an E-TBC, it is important to think about the length of the capacitor plates because the length between **B** and **D** will give the total value of the induced voltage between the capacitor plates which increase the electromagnetic flux (Figure 44).

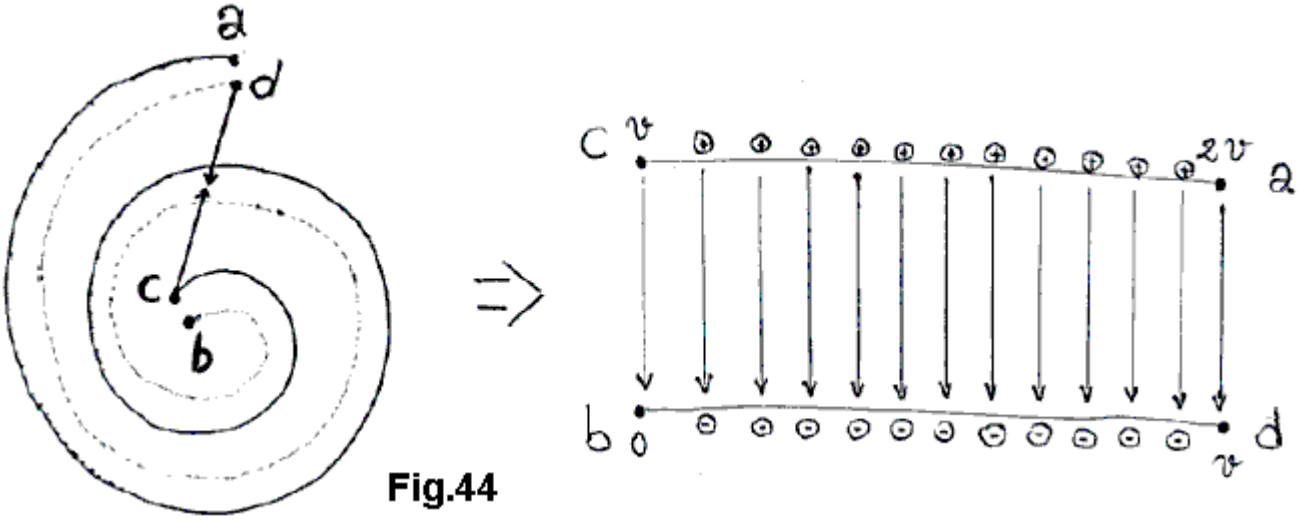


Fig.44

A simple diagram of the resonance energy device can be like the following drawing (Fig.45):

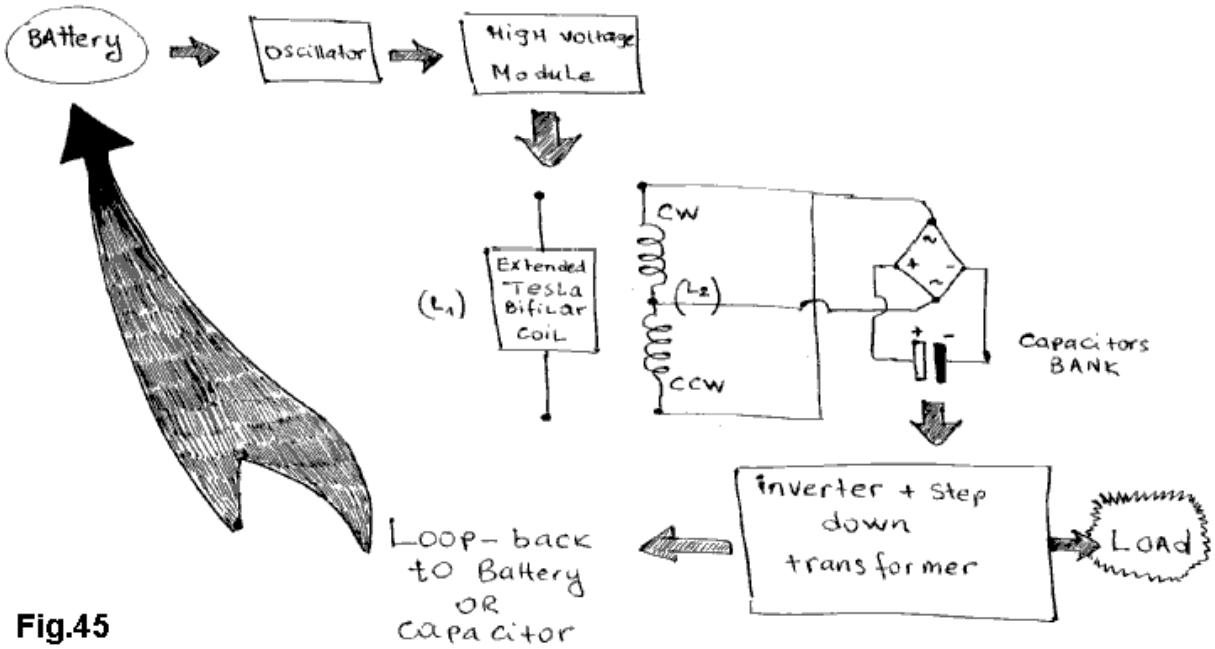
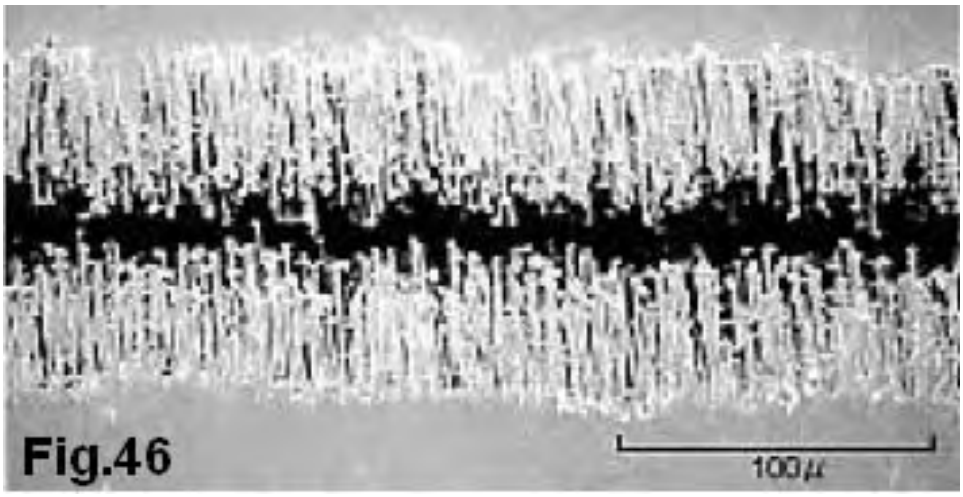


Fig.45

When converting the reactive electric power from the L2 coil to Direct Current and stepping down the energy obtained to the mains voltage and frequency (for example, 220V, 50Hz), the current will be boosted yet again.

Some applications may not need an inverter. An electric heater can be fed directly from the capacitors banks but we have to prevent the Alternating Current coming from L2 to enter the heater by using another high inductance coil.

Another idea about increasing the capacitive side of the extended TBC is to use etched aluminium foil to increase the surface area. The foil can be treated chemically using high-voltage. The result is shown in Fig.46:



Maybe this is the technique used by Donald Smith to avoid the need for an earth connection. I have already mentioned about using negative energy to avoid the earth connection requirement in the Resonance Energy device but I can't guarantee that this is the method used by Donald Smith.

Any questions, or suggestions are welcome via my e-mail: hopehope3012 (at) gmail (dot) com

Chinese Developer Ming Cao.
a free-energy developer in mainland China, comments on the designs of Don Smith, and Tariel Kapanadze. He says:

None of these things originate from me, they come from Tesla and God.

1. The most important issue, is resonance. Don Smith said that we should make the wire length of the primary coil to be one quarter of the wire length of the secondary coil in order that they will resonate together. My experiments show that this is not true. In a Tesla Coil, the primary coil and it's capacitor form a tank circuit which is an L/C circuit, which oscillate at it's own resonant frequency, and when it does that, it generates a longitudinal wave at that exact frequency. The frequency of this longitudinal wave is determined by the inductance of the primary coil combined with the capacitance of it's tank circuit capacitor, and not the wire length of the primary coil alone. The secondary coil with it's sphere at the top, together form an antenna, which transmits this longitudinal wave. The secondary coil and it's top sphere together form a quarter-wave resonant antenna for this longitudinal wave. They do not form an L/C circuit and that is why very few people have managed to replicate Don Smith's devices.

2. In the devices of Don Smith and Tariel Kapanadze, there is no sphere. We see a single coil as the secondary. This is no longer a quarter-wave antenna, but a half-wave antenna. The highest voltage shows up at the very centre of this coil, and zero voltage shows up at the two ends of the coil winding. These are where the energising coil and the pick up coil should be positioned.

3. The longitudinal wave which passes through the secondary coil is not a current at all, it is a signal running through it, so if we let the secondary to charge a capacitor, we will get nowhere. All we will get is hot electricity caused by the loose induction coupling. The arc at the top of a typical Tesla coil is lightning voltage, and no capacitor on earth can handle that voltage, so even a very high voltage capacitor will be over stressed and the arc will shock through it.

4. The speed of this wave is well defined. It depends on the total capacitance of the coil, and the sphere if there is one. In a typical Tesla coil, the bigger the sphere, the bigger the capacitance, and the lower the resonant frequency of the secondary coil. People are trying to explain it by L/C circuit theory, but this is not necessarily true. Increased capacitance will slow the wave down. If there is no sphere, as in the devices of Don Smith and Tariel Kapanadze, the total capacitance is quite small, and so, the speed of the wave should approach the value (π/2) x C, where C is the speed of light. This speed of the longitudinal wave is claimed by Tesla himself. I kind of verified this by experiment. I said "kind of", because in my experiment, I got a speed of (π/2) x C x (8/9). The wave is slowed down by the copper losses and the capacitance of the coil, mostly the capacitance, but it's definitely faster than the speed of light.

5. So, to tune the secondary, we should not use the speed of light at all, Don was playing a game with us here. Take Don's device as an example. If we position the primary coil at the centre of the secondary coil, then that middle point of the secondary should either be grounded or connected to a large metal sphere, and each half of the secondary coil should act as a half-wave antenna. Also, the pick-up coils should be located at the two far end terminals. The speed of the longitudinal wave along the secondary coil is unpredictable and so we can only predict a general speed range, we cannot tell whether it is already resonating by performing calculations. Like Nick Giannopoulo's arrangement (see below) and Tesla's patent diagram, there are two quarter wave coils, whose inner terminals are connected together and open to the air. Here 'open to the air' means that it is different from the other turns of the coil. The longitudinal wave is climbing the turns rather than passing along the wire. But at the end of each quarter wave coil, there is no other turn to climb any more, only a long wire for it to travel along. This straight long wire is open to the air and provides a capacitance for the whole device, and this additional capacitance will slow down the longitudinal wave which is passing through it, so that the resonant frequency for the combination of these two secondary coils will be lower. But if we eliminate the straight wire, and make it a single half-wave secondary coil, the longitudinal wave can keep climbing the turns, and there's no additional capacitance, so that the speed of the longitudinal wave will be very close to (π/2) x C, and the resonant frequency will be higher. We can use the same wire length and same diameter coil former to build different devices, which will work at completely different frequencies. So the resonant frequency is unpredictable and we need to find the exact frequency by equipment measurement, or it won't work. The only correct way of tuning the secondary is shown by Eric Dollard in his video of the 1980s, entitled "Eric Dollard Transverse and Longitudinal Wave" which at the present time can be found on YouTube at http://www.youtube.com/watch?v=6BnCUBKgnnc.

6. A pick-up coil is always necessary, and it should be positioned near the zero node of a standing wave. This is one of the only two ways of harness the longitudinal wave. This method is the dynamic way, the other way is the static method, which I believe was used by Ed Gray.

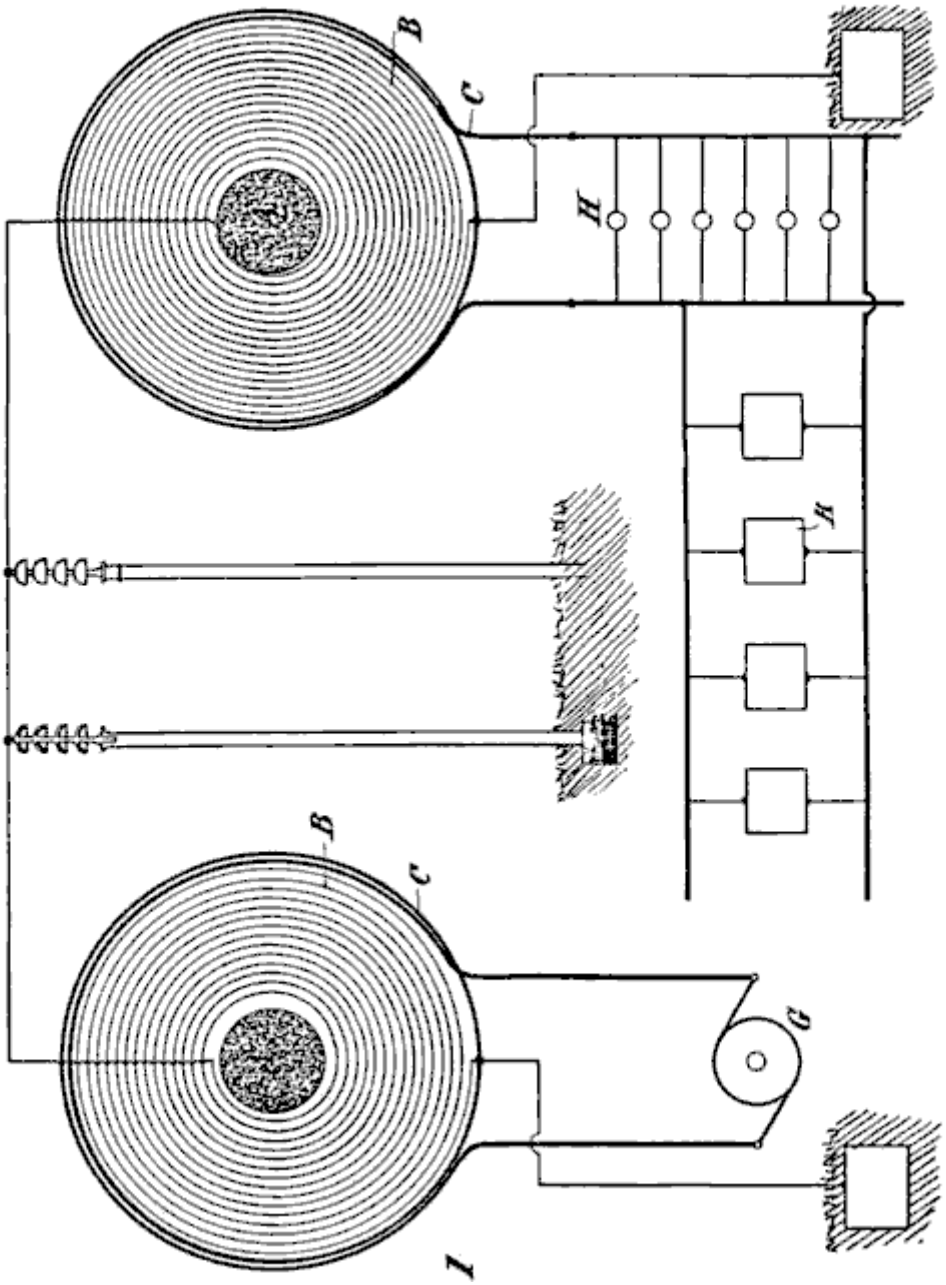
7. In Dr. Peter Lindemann's book and video, he says that Tesla is using unidirectional current. I have to disagree with this. When we charge a capacitor and discharge it through a spark gap, the discharge current "bounces" between the two plates of the capacitor, until the energy is all lost at the spark gap. This process repeats itself endlessly in a typical Tesla Coil. We can see this primary waveform with an oscilloscope and it is alternating current. Thousands of Tesla Coils work in this way and generate lightning. I am confident that this is how it operates.

8. It is not like Don Smith said, that doubling the voltage quadruples the output. It does look like that, but it is actually the current flowing through the primary doing the job. Of course we increase the current by increasing the breakdown voltage of the spark gap by widening the gap. But fundamentally, it is the current which is doing the job. Ed Gray's tube uses a short straight copper bar as the energising 'coil', but it's not a coil, it has little inductance to generate voltage, it only has high current passing through it to energise the longitudinal wave. Of course I haven't actually seen this process, it is a conclusion which is not fully based on experiment.

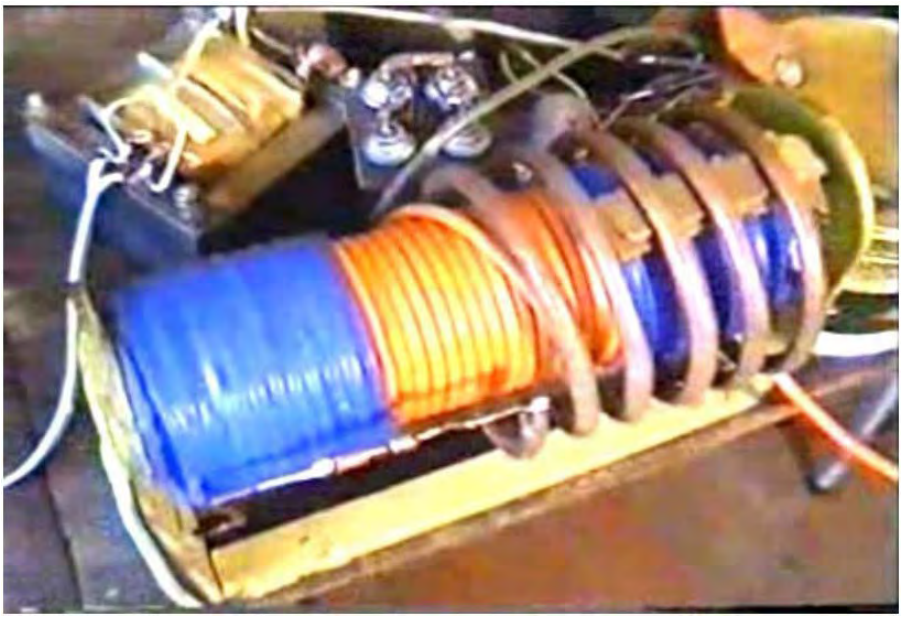
9. The larger the number of turns in the pick-up coil, the higher the output voltage will be. I still cannot understand how the pick-up process works, but it does pick up more energy.

I get all these by low voltage from a signal generator, as I haven't finished building a high voltage device yet, although I'm already working on it. But I think it's safe for me to believe that these results are solid and good enough to share.

Here is a image from Tesla's patent 593,138 Electrical Transformer.



We can see it's exactly the same as Nick Giannopoulos' setup, except that Tesla is using a generator in this diagram, I believe for simplicity. As long as the generator is generating the exact frequency of current, it will work fine. The secondary at the energising side is a quarter-wave coil, and at the pick-up side is another quarter-wave coil. The highest voltage is at the far end of these two secondary coils and their connecting wire, and zero voltage is at the very outside turn of each of the coils. Now if we change the spiral form coil to helical, it becomes Nick's set-up. And let's take this further, we can shorten the connecting wire until the two solenoid secondary coils actually become one big coil, then, when combined it is a half-wave coil, and the highest voltage is at the middle point of it. Now it becomes Don Smith's and Tariel Kapanadze's device, like this:



Because the energy is also coming back from the energising side, Kapanadze adds another pick-up coil right underneath the primary energising coil. This arrangement, I think, is very hard to replicate, because it is so very hard to tune, for several reasons:

1. The Secondary wire length is rather short, and the wave speed is very very close to (π/2) x C, so the frequency should be very high, at least 5-7MHz I would guess, or perhaps even higher.

2. The pick-up coil and the energising primary coil is too close to the centre point of the half wave secondary coil. Because the centre point is the point of highest voltage, if the input is a little high, there would be arc shock between the secondary to the energising coil and the pick-up coil, at lightning voltage levels, and so even the best isolation is useless. Also, the centre point is very very sensitive, any conductor close to it will add to the total capacitance of the coil and of course that will alter the half-wave resonant frequency. This adds more difficulty to the tuning adjustment. Besides, after all, people don't even know it's a half-wave coil if he doesn't tell us.

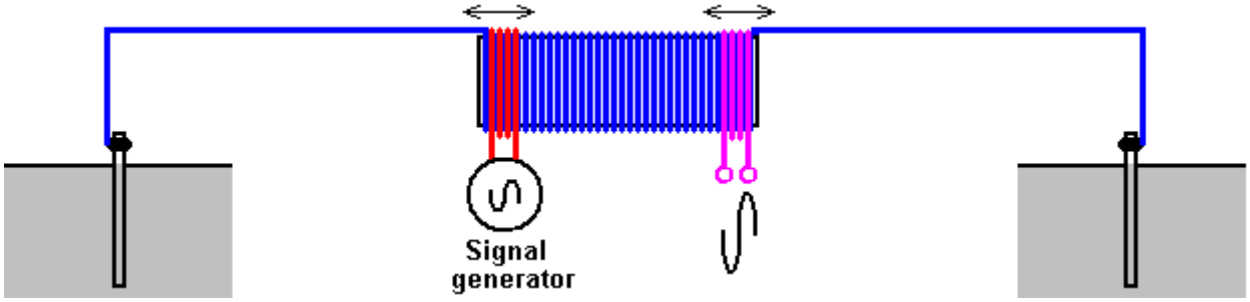
3. The coupling coefficient **K** is a little high, this will increase the hot transformer effect by inductive coupling, and that will not help at all.

Don Smith did indeed say something useful. He said that we can make the secondary coil a fixed size, and then slide the primary coil inside it. Well based on experimental results, this sliding process is altering the actual effective length of the secondary coil. In general, we should assess coil size by counting the turns from the turn right underneath the primary energising coil, to the turn right underneath the pick-up coil, this section is the actual secondary, and this section should be a half-wave resonate coil, the rest of the coil just sits there doing nothing.

But it's not that simple, the terminals of the secondary coil should connect to the earth or to a large sphere, or a typical Tesla Coil secondary with the same quarter-wave resonant frequency. Otherwise the signal will bounce backwards and forwards in the coil producing a mess, or generating an arc, and this is bad for performance, and this is why a solid ground connection is desirable. And this is the true meaning when Don says "slide the primary coil to do the fine tuning".

So, returning to the Kapanadze device, the energising coil covers a large area of the secondary coil, making the effective length of the secondary coil very much shorter, again, boosting the working frequency of the device even higher. For such a device, it is impossible to tune it without a 20Mhz signal generator, an oscilloscope and complete understanding of how a longitudinal wave behaves. For a start, I don't even know where to connect the oscilloscope probe or which terminal should connect to the ground, I'm so lucky to be able to watch Eric Dollard's old video, and I recommend everybody to watch that video, watch it over and over again, also many other educational videos from Eric. A lot of fundamental stuff about how a longitudinal wave behaves are explained there, it's like a treasure map covered in dust in a quiet corner of an open library.

Ming's video http://www.youtube.com/watch?v=1p41KLfOM2E&feature=youtu.be demonstrates what he is saying here. For the video he uses an input coil, a monitoring coil and a secondary coil, each end of which is earthed using separate earth connections:



Ming also remarks:

For the set-up in the video, the secondary coil is wound using 1mm diameter enamelled copper wire, 365 turns around a 160mm diameter PVC pipe. The total coil length is 39.5cm. The total wire length of the secondary is approximately 182m. The white material is several layers of insulating glue to prevent arcing between adjacent turns when working with high voltage. The primary coil and the pick-up coil are wound with audio cable which is more than 4 square millimetres in cross section. The primary coil has 2 strands, 2 turns. The pick-up coil has 4 strands and only one turn. I use this thick wire, because I am going to use these coils for my high voltage project.

For a low voltage experiment like shown in the video, it would be quite adequate to use ordinary copper wire of 1 square millimetre cross section (swg 18 or AWG 17). If the secondary wire length is reduced, then the resonant frequency will be higher, but the principle is the same.

If only low voltage is going to be used - perhaps just to study the nature of longitudinal waves, then the secondary coil can be made using very thin wire of 0.3 to 0.4mm diameter (swg 30 to swg 27) enamelled copper wire, which will cost much less. I made my coils with thick wire because I intend to continue using high voltages.

It's been a long time, but I've got some more understanding about harnessing radiant energy. I have made two additional videos: http://www.youtube.com/watch?v=WJUfj53geBo and http://www.youtube.com/watch?v=BdBjKVyKBZA In these two videos, I explain the method of converting Tesla's 'cold' electricity to normal 'hot' electricity by storing it in a capacitor. I strongly believe that, the method shown in the second video is exactly what Don is doing with his famous device, which has no pick-up coil, just a two-part secondary.

In the first video, I replace the pick-up coil with an uncoated sheet of copper, to show people that, this is not a transformer, and so, is not based on electromagnetic induction. The pick-up coil is fundamentally, a piece of metal which can be electrified by a longitudinal wave. I can remove the diode and the capacitor, and just let the copper sheet discharge to ground through a spark gap and two ordinary 200-watt incandescent light bulbs connected in series, the light bulbs are pretty bright although not fully lit, but pretty bright in spite of this being a non-resonant situation. They look like this:

The copper sheet is electrified, and it's charge is flowing to ground, and it is this very process which forms the current. So if we consider it as a transformer, and consider the pick-up coil as a inductor, and add a load to this "inductor" to form a closed loop, then we are going in the wrong direction.

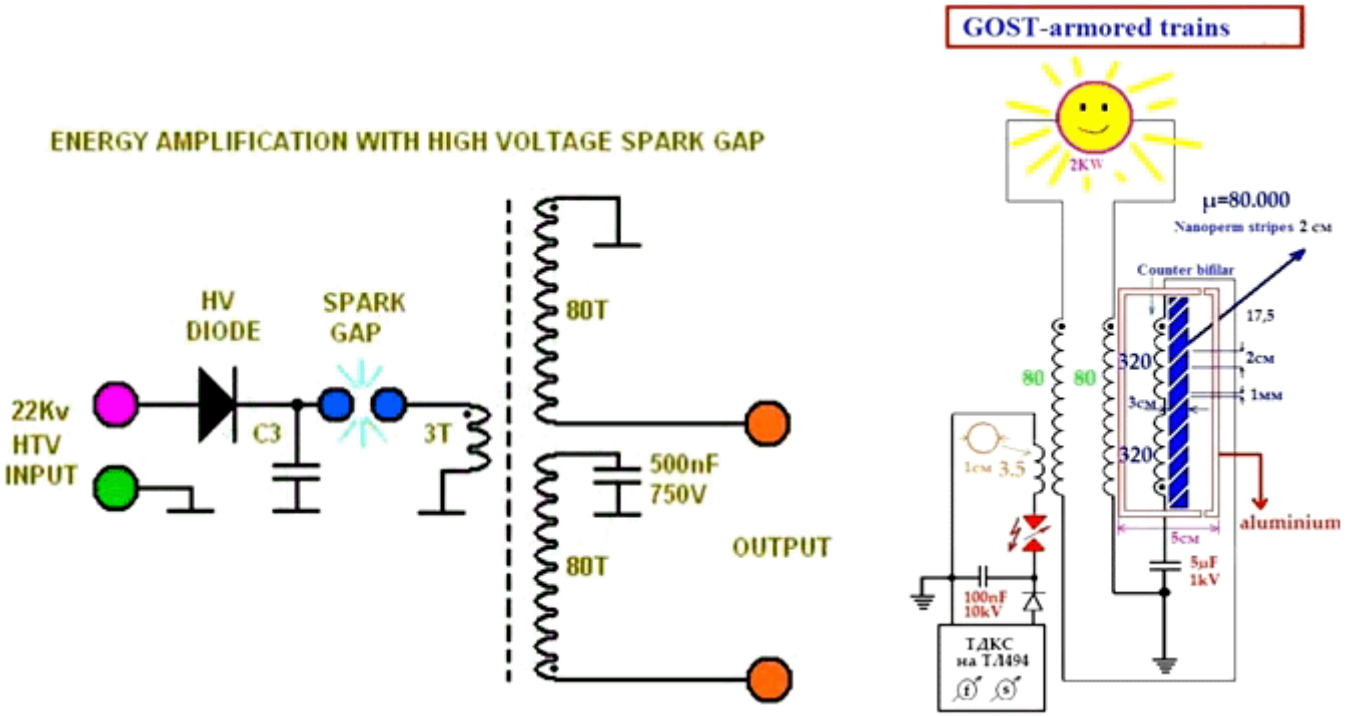
Then I re-read about Nick Giannopoulos' device, and I noticed that he said that the light coming from his bulb is blue and white. Following his circuit diagram, I believe that it is like this:

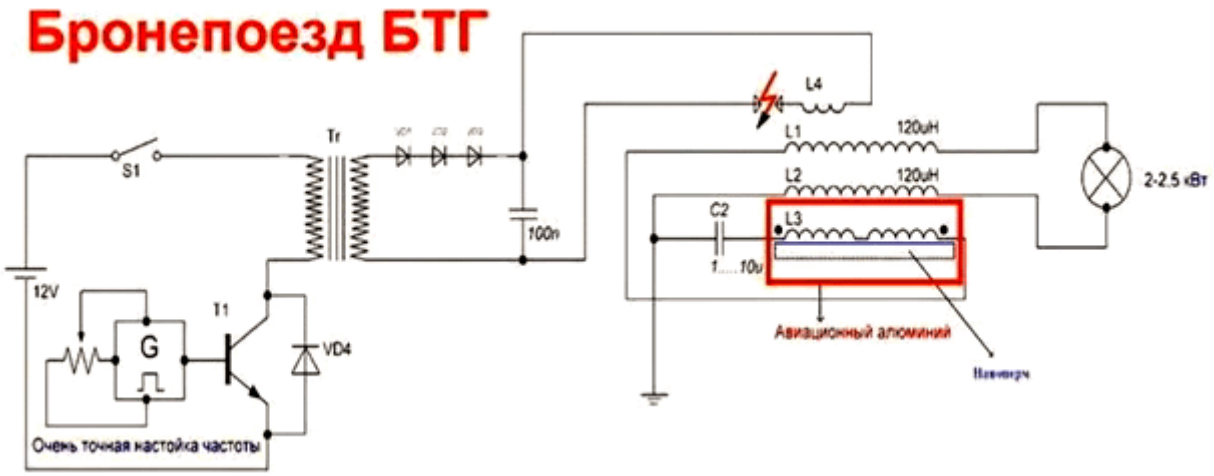
I get this kind of light when I attach the light bulb directly to the sheet copper without an earth connection or any other additional wire. Here at this stage, we have no 'hot' electricity. The blue-white light is caused by the high voltage of the metal, to which the bulb is attached. The high voltage is not caused by induction, it is purely static charge on the metal surface, caused by longitudinal wave electrification. If we use Tesla's specially made bulbs as shown in his lectures, we have his single-wire lighting system, and we will have a very bright light suitable for general-purpose lighting instead of this kind of blue-white light. Generally speaking, my bare copper sheet is the equivalent of Nick's pick-up coil plus his step-down transformer, which, of course, is not a transformer at all.

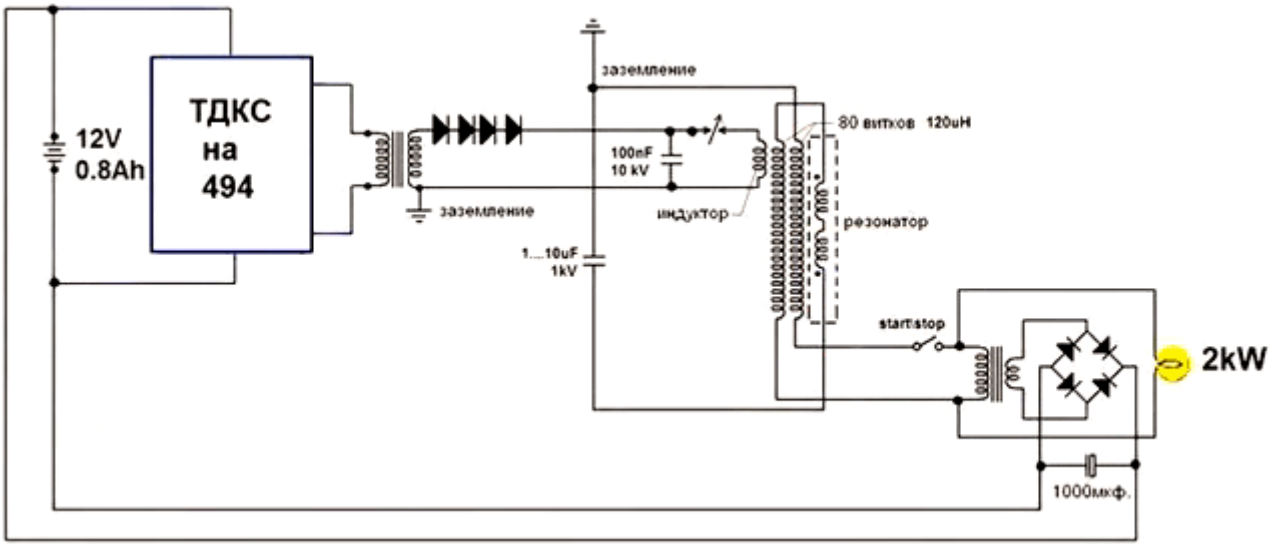
Note: As will be seen in the videos, Ming uses two separate earth connections. One is the earthing wire of his mains electricity and the other is a connection to his cold water pipes.

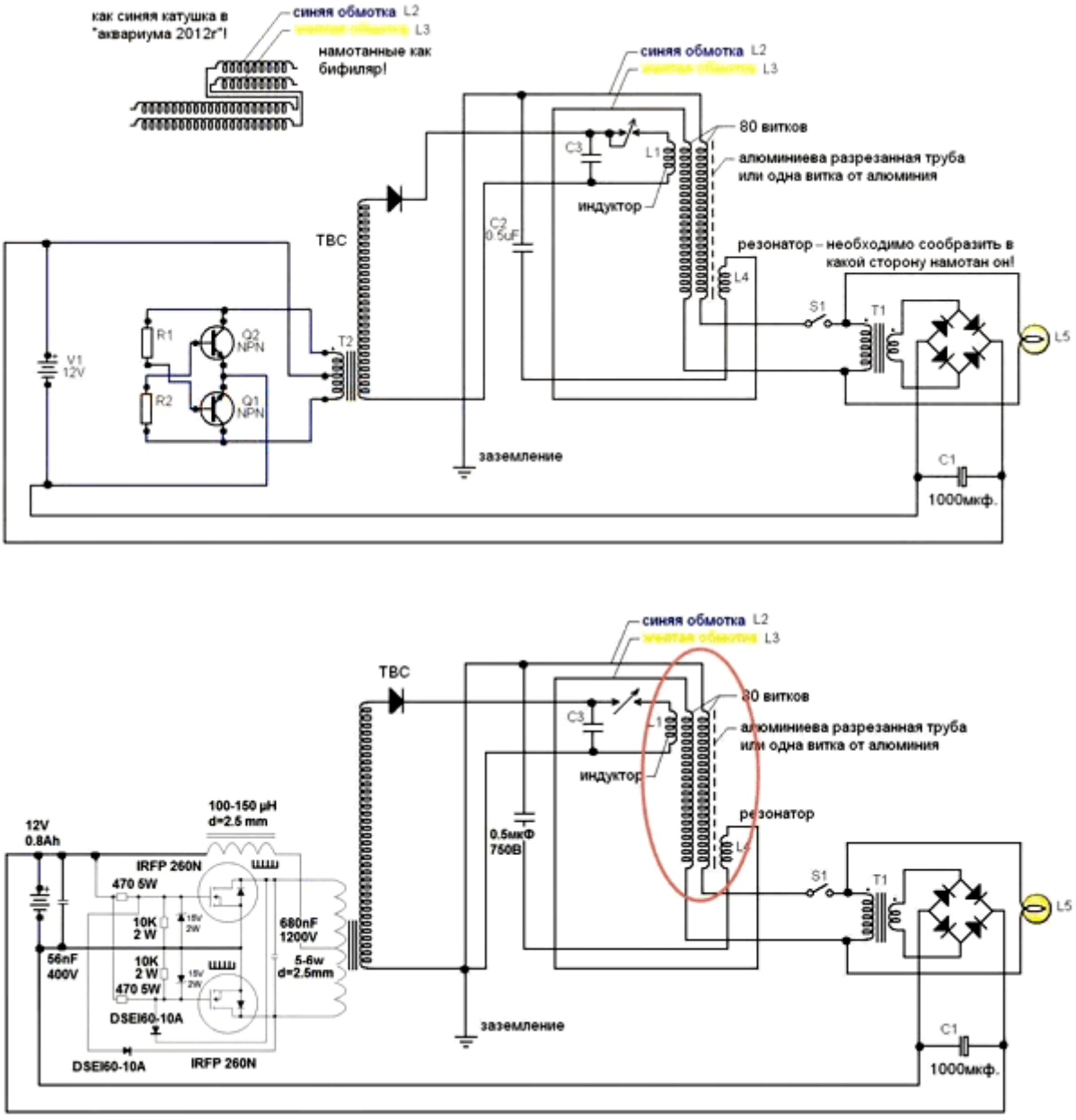
A Russian Developer

A Russian developer has lit a large light bulb with a self-powered Kapandze-style circuit:


ENERGY AMPLIFICATION WITH HIGH VOLTAGE SPARK GAP
HV DIODE
SPARK GAP
22Kv HTV INPUT
C3
3T
80T
500nF 750V
80T
OUTPUT
GOST-armored trains
2KW
μ=80.000
Nanoperm stripes 2 см
Counter bifilar
17,5
2см
1мм
3см
320
320
80
80
1см 3.5
5см
aluminium
5μF 1kV
100nF 10kV
ТДКС на TL494


Бронепоезд БТГ
S1
Tr
12V
G
T1
VD4
100n
L4
L1
120uH
L2
120uH
C2
L3
1....10u
2-2.5 кВт
Авиационный алюминий
Очень точная настойка частоты

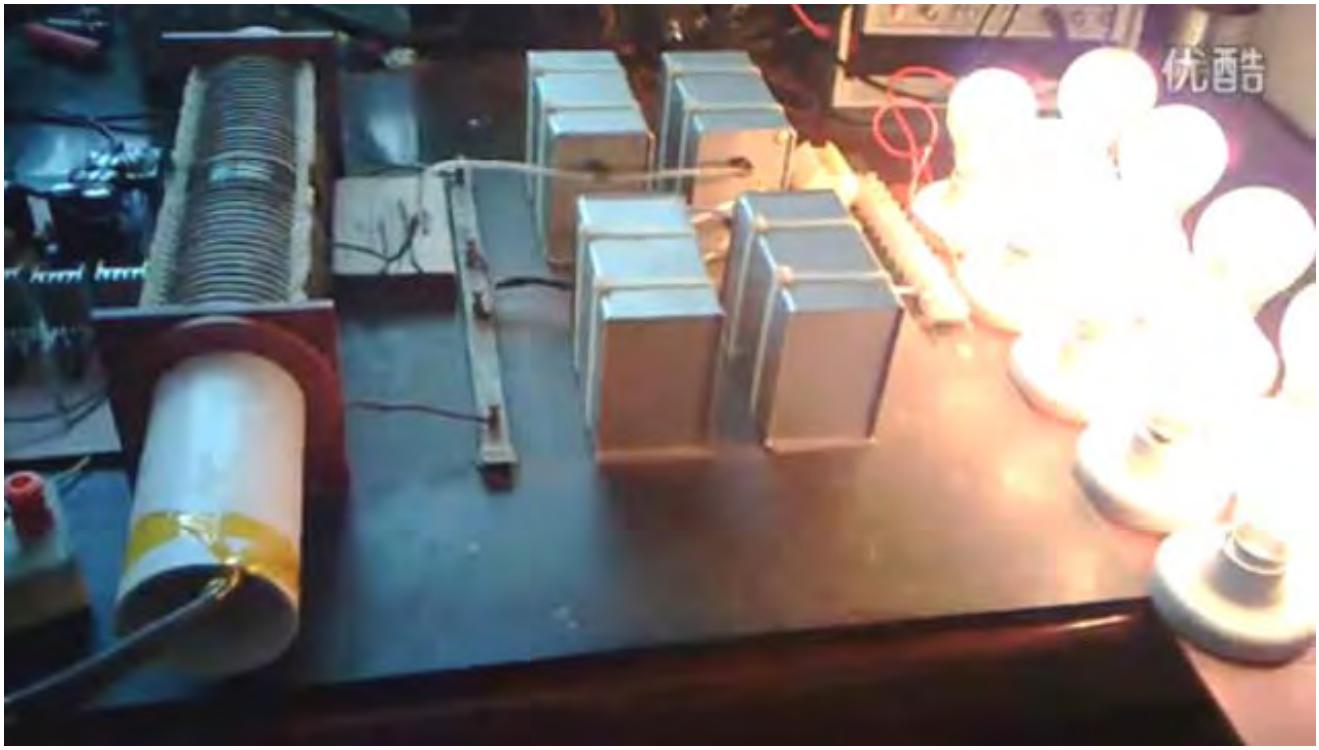

12V 0.8Ah
ТДКС на 494
заземление
заземление
100nF 10 kV
1....10uF 1kV
индуктор
80 витков 120uH
резонатор
start/stop
2kW
1000мкф.

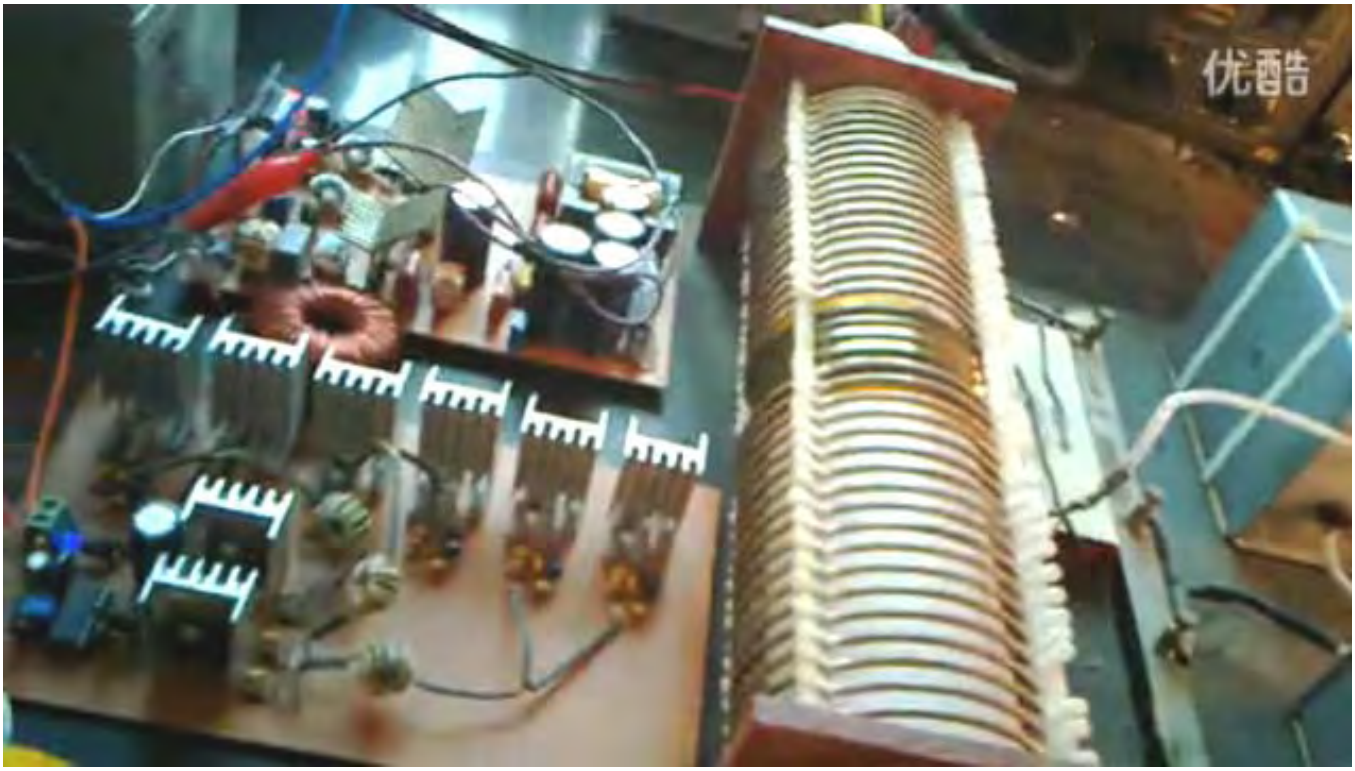


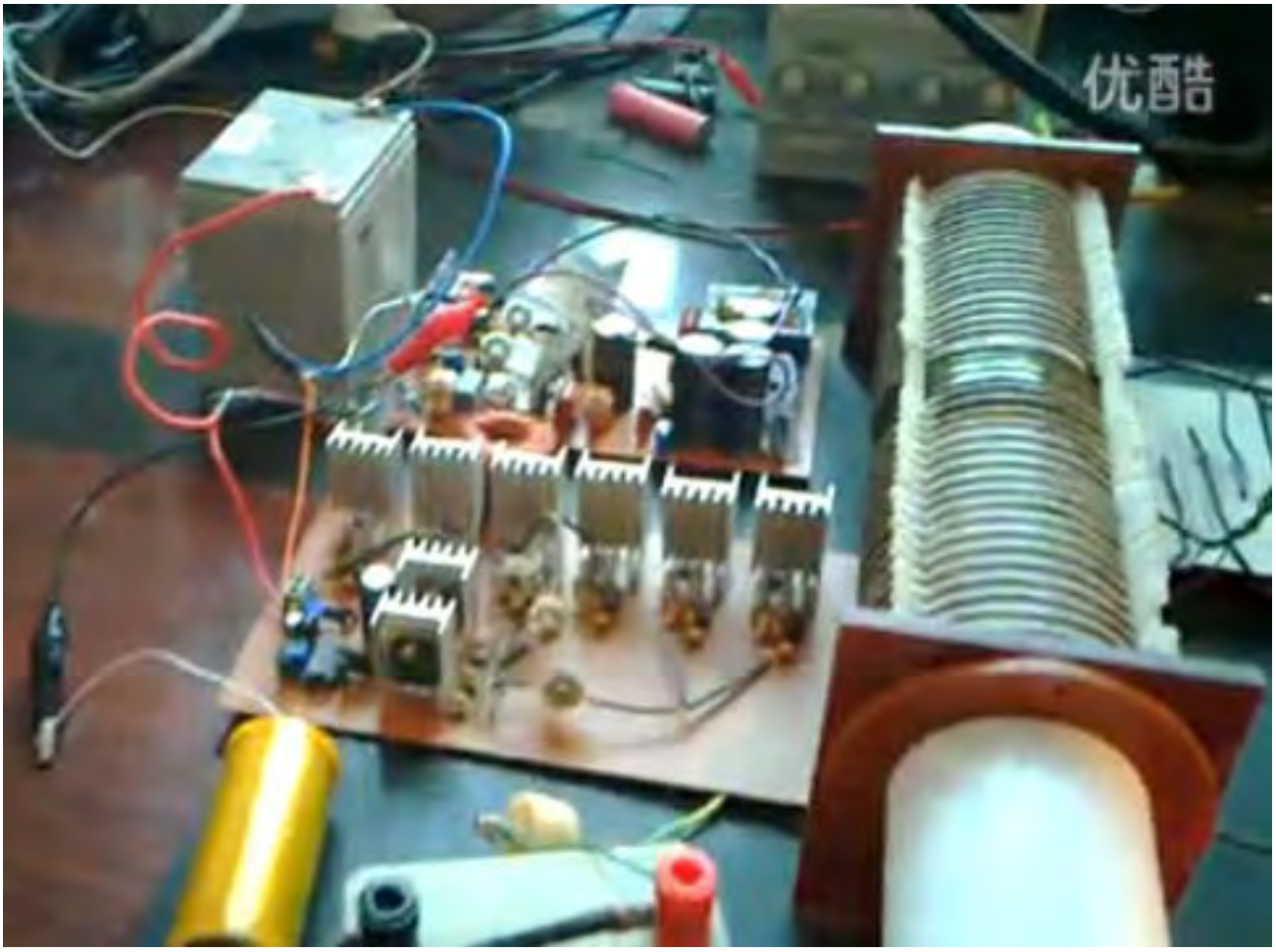
Video at http://www.youtube.com/watch?v=5nxKqfkkndw&feature=youtu.be shows self-powered bulb (needs an earth connection):

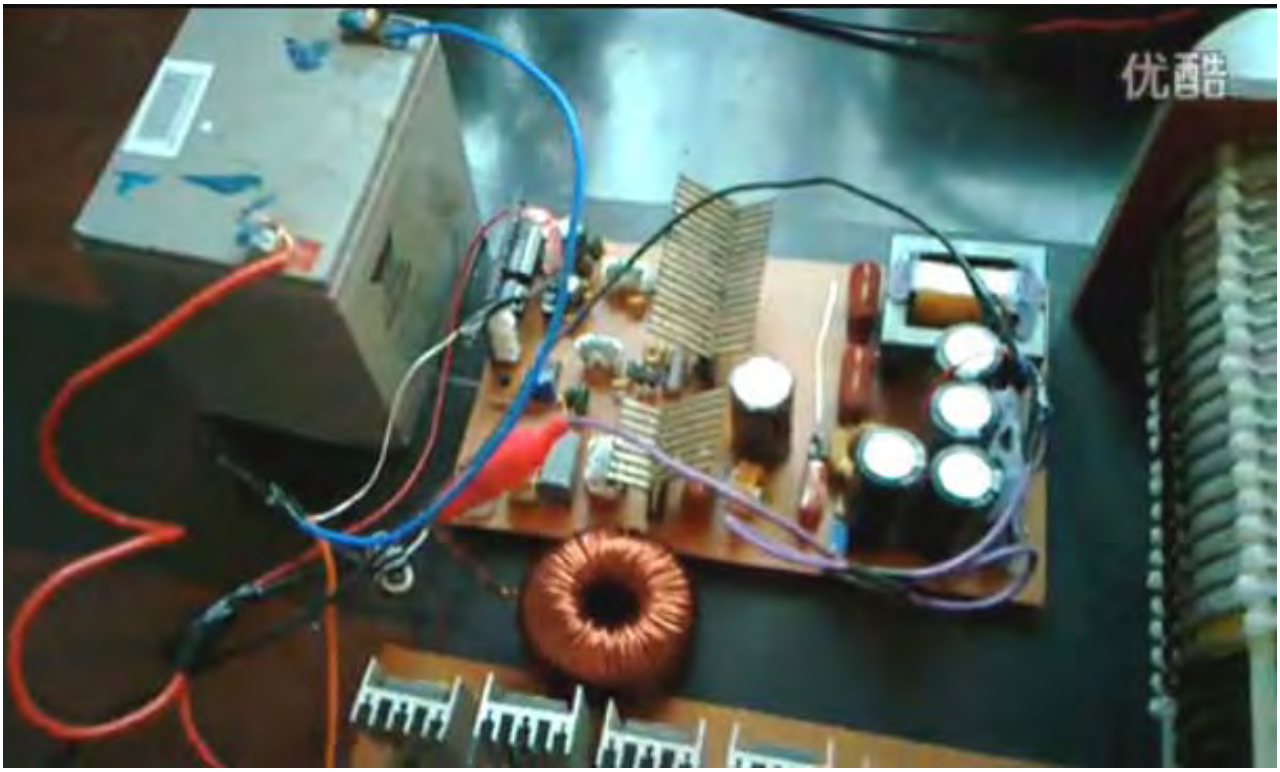
‘Salty Citrus’ Chinese Developer.
A Chinese developer whose forum ID is ‘Salty Citrus’, has replicated Don Smith’s main device very successfully. Using an input of 12V at 1A to 2A (24 watts) he is lighting ten 100-watt light bulbs to a high level of brightness. The Chinese language video relating to this can be seen at:
http://www.energysea.net/forum.php?mod=viewthread&tid=1350&extra=&page=1

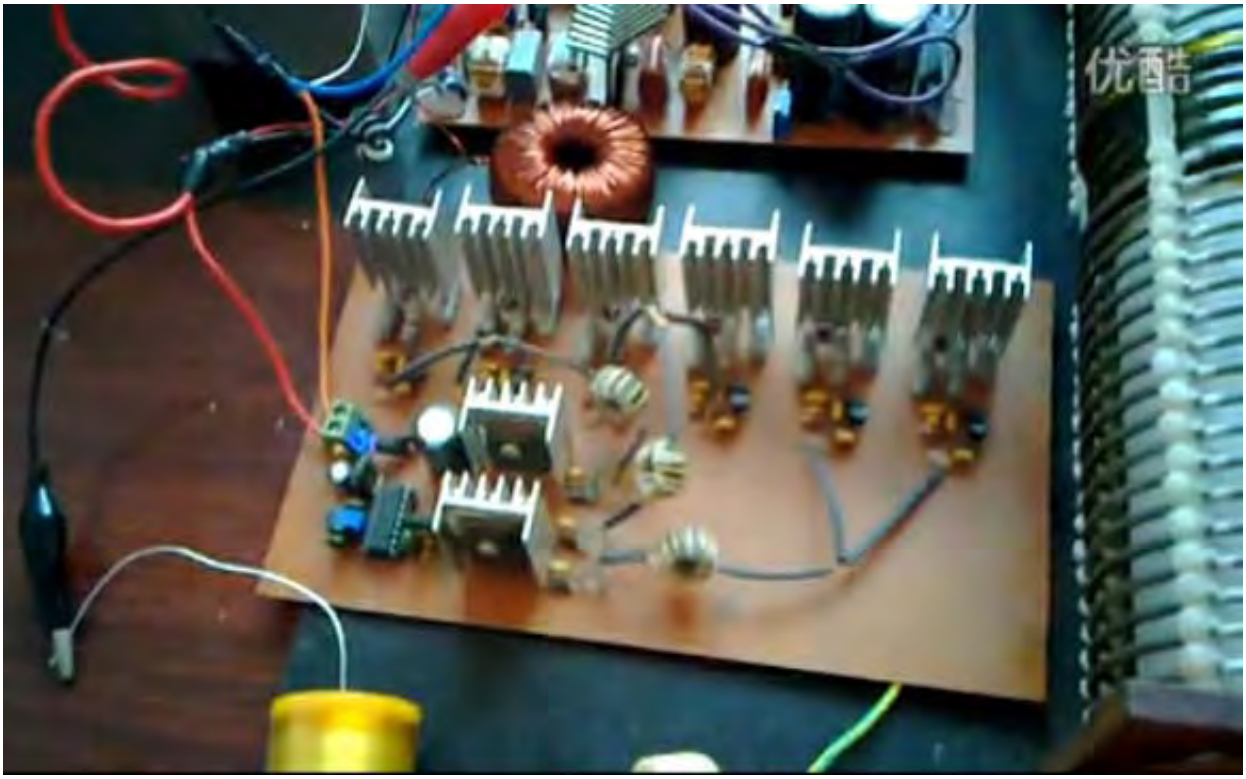
Here are some of the frames from that video:



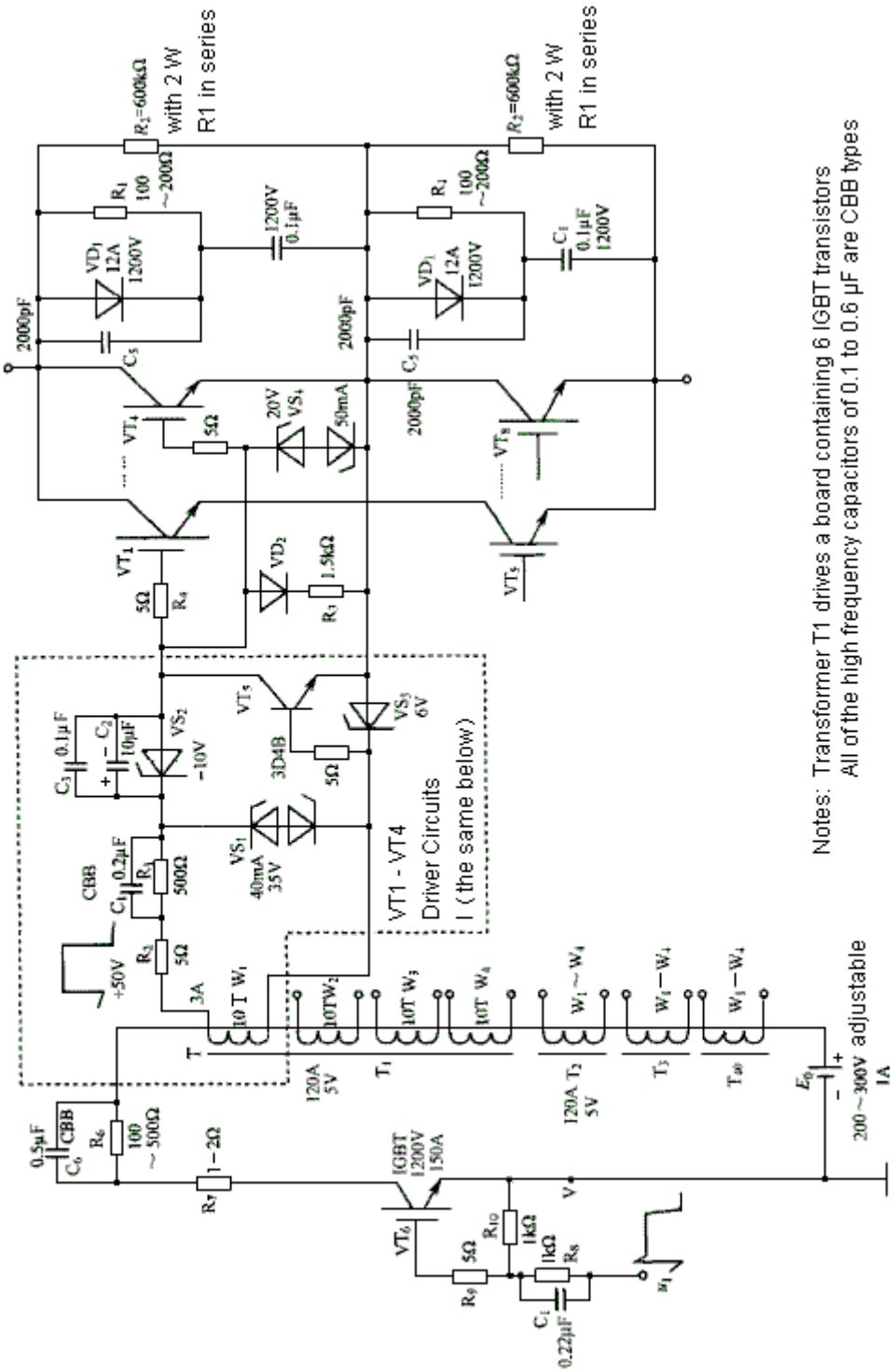

优酷告

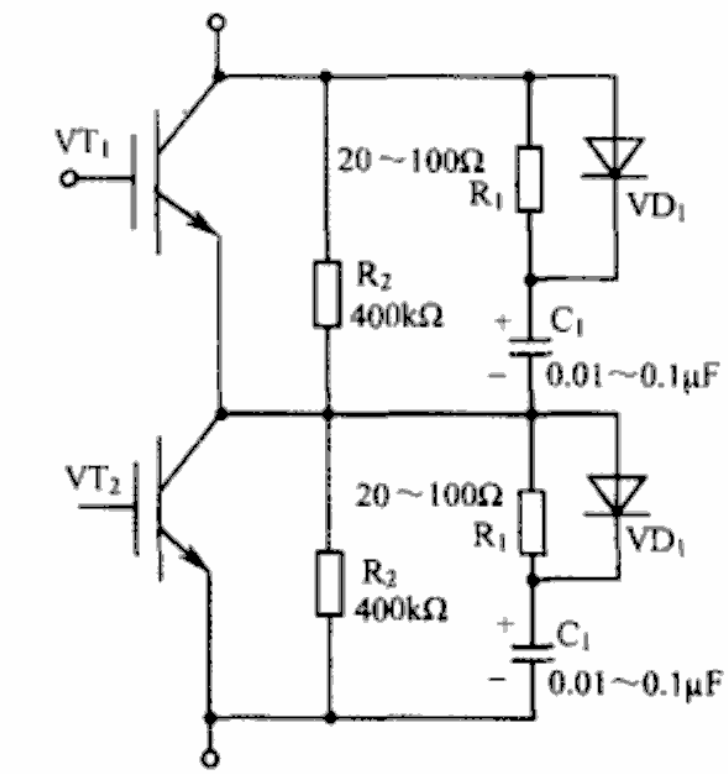

优酷告





The circuitry used is shown here:


R2=600kΩ with 2 W R1 in series
R1 100 ~200Ω
VD1 12A 1200V
1200V 0.1μF
2000pF
C5
VT4
5Ω
20V VS4
50mA
VT8
VT1
5Ω R4
VD2
1.5kΩ R3
VT5
3D4B
VS3 6V
C3 0.1μF
C2 10μF
VS2 -10V
CBB
C1 0.2μF
R2 500Ω
VS1 40mA 35V
R5 5Ω
+50V
3A
10 T W1
10TW2
10T W3
10T W4
W1 ~ W4
T
T1
120A 5V
T2
T3
T10
E0
200 ~ 300V adjustable 1A
VT1 - VT4 Driver Circuits I (the same below)
Notes: Transformer T1 drives a board containing 6 IGBT transistors
All of the high frequency capacitors of 0.1 to 0.6 μF are CBB types
0.5μF CBB C6
R6 100 ~ 500Ω
R7 1 – 2Ω
IGBT 1200V 150A
VT6
R9 5Ω
R10 1kΩ
R8 1kΩ
C4 0.22μF
V
W1



Subsequently, a forum post by a Mexican man says:

Hello ‘Salty Citrus’,

I love your video!!!! I can really appreciate the amount of work you and your group have spent to develop and perfect the Don Smith / Tesla free-energy device. Thank you for pursuing such a noble cause.

I am intrigued by your switching network using the CREE CMF20120. How did you wire the MOSFET's?
You used a UCC3825A Pulse-Width Modulator to clock the signal --> MOSFETS --> Gate Drive Transformers (x3) --> push-pull transistors --> CMF20120? Did you run the CMF20120 in series? Sorry about so many questions, but I am totally impressed by your ingenuity, and completely agree that your solid-state solution has undoubted benefits over Tesla's conventional spark gap.

I would be honoured if you could take the time to answer my questions. I would love to replicate your circuits.

I wish you the best of luck with your endeavours.

Sincerely,

‘Lost_bro’ (half a world away)

Re: ‘Lost_bro’

Thanks for the compliment. The success does credit to my team. Thanks to my team. Yes, the CMF20120 run in series in this solution. The voltage balancing between each MOSFET is critical as is the balancing between RC and DC voltage created by R.

Welcome to our forum for the exchange of information. China is an hospitable country. If you have any information or ideas, please don’t hesitate to share them with us. ‘Half a world away’ is not a great distance.

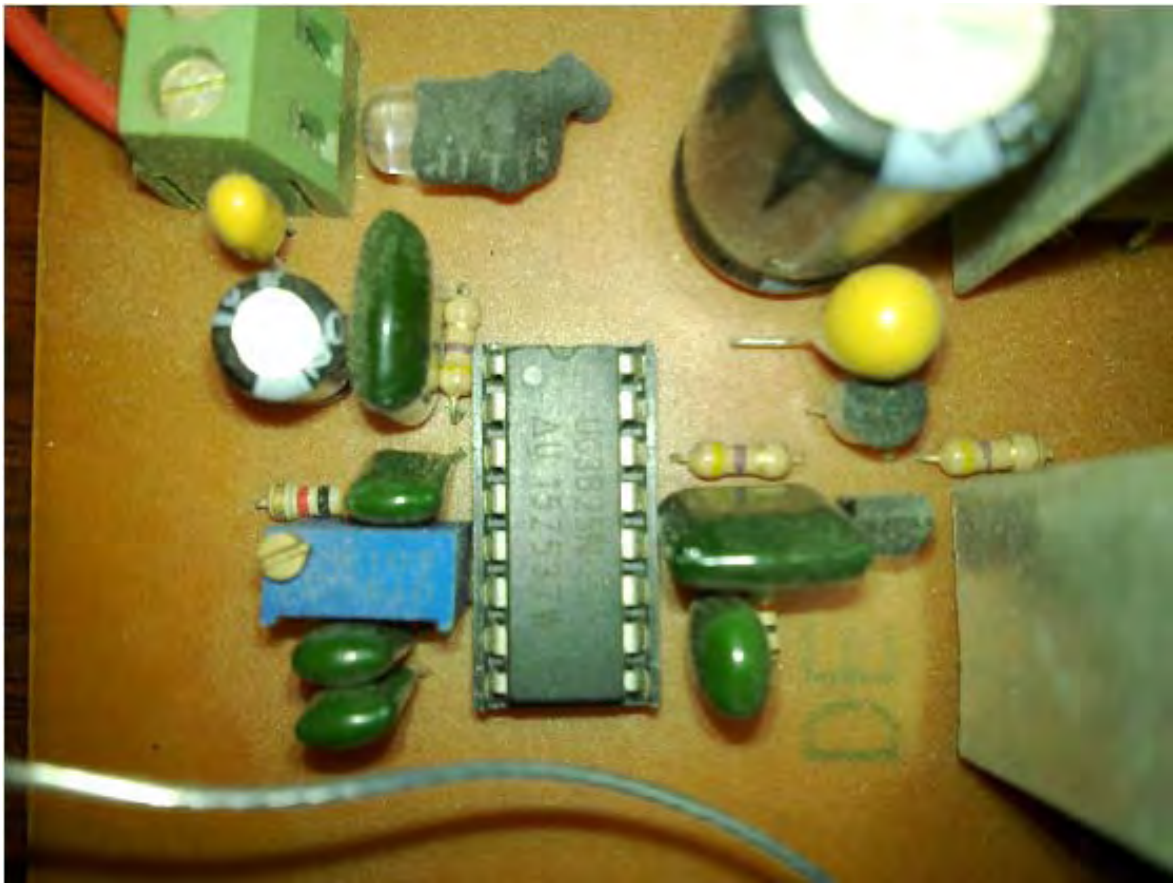
All the best,

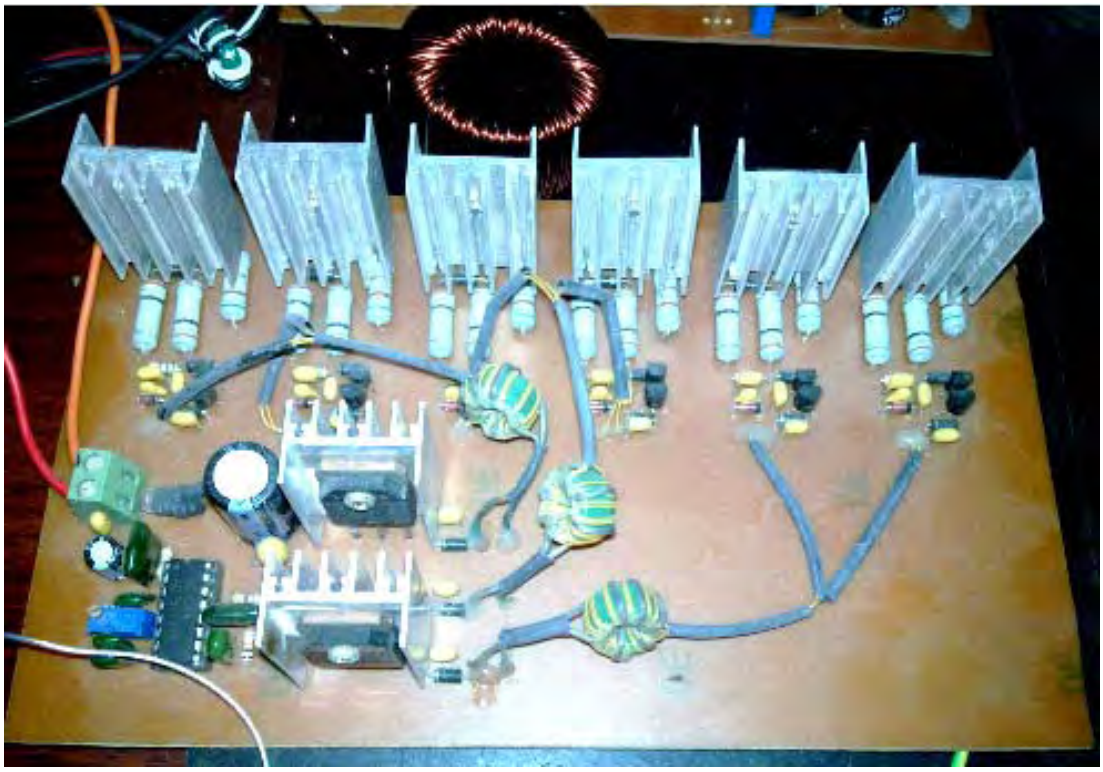
Sincerely

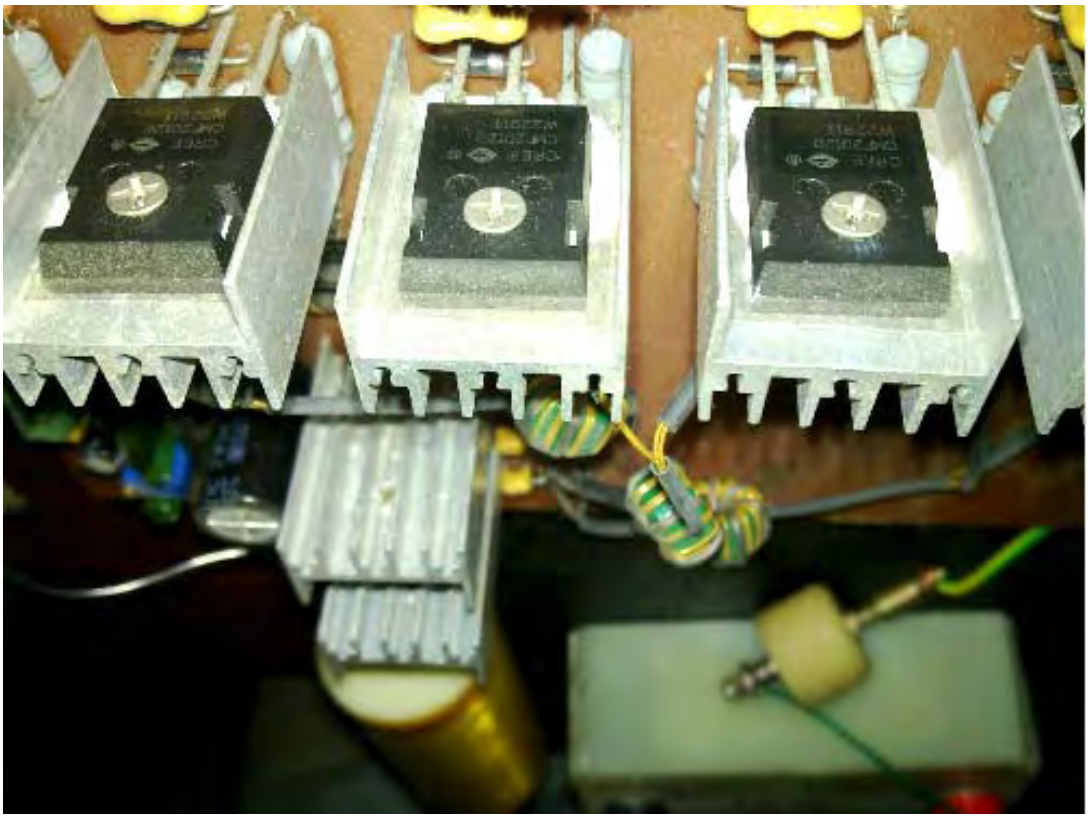
‘Salty Citrus’

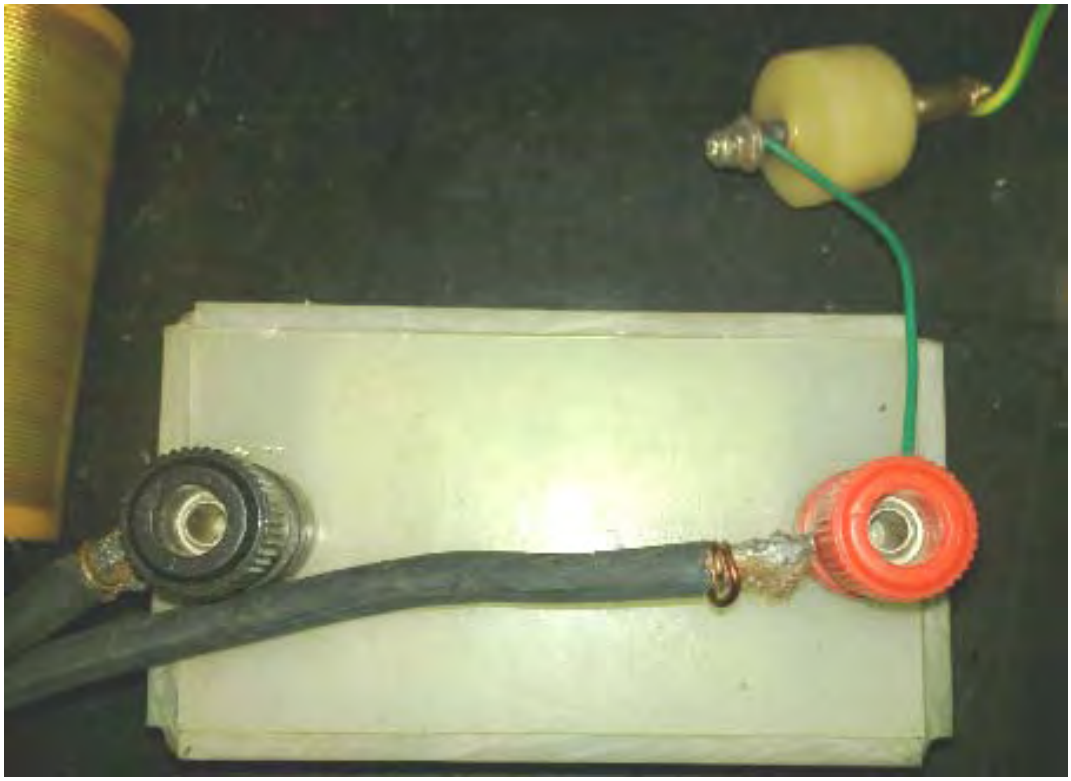
An earlier entry on the Chinese forum translates as:

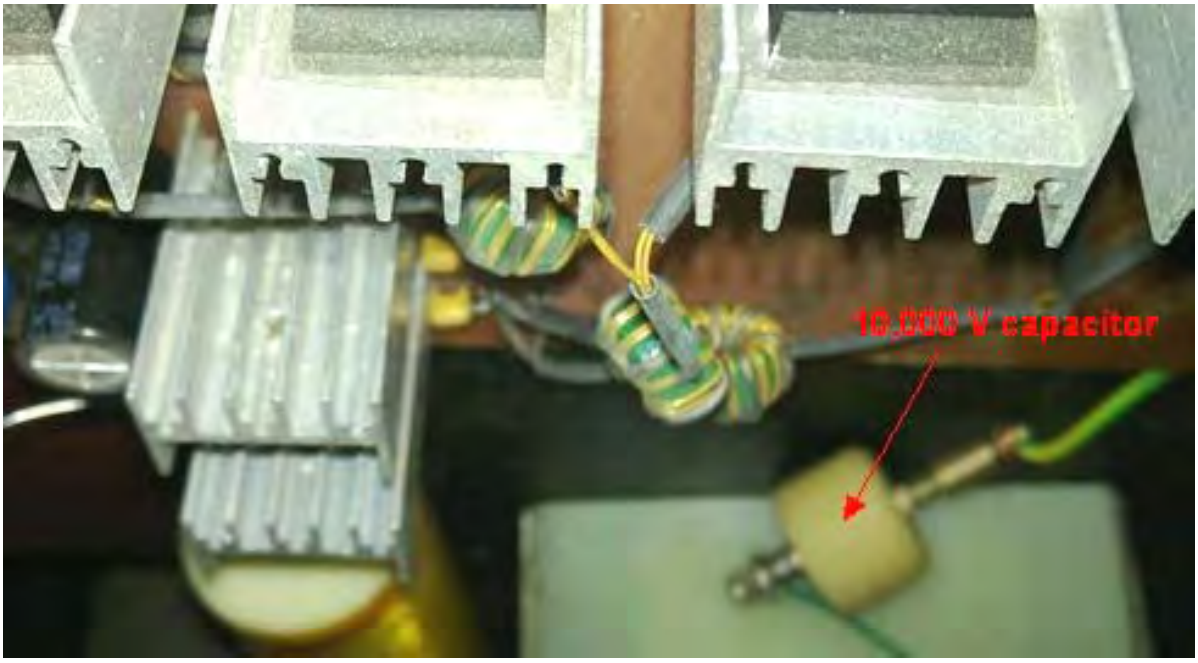
Here is an earlier build. It is simple and has no step-down section and so cannot be self-powered:










10,000 V capacitor

Each bulb is 100 watts. The first board has a 12-volt input and an adjustable output which can be varied from 500V to 1600V (any higher voltage would damage the four 450V 20 microfarad capacitors). In the video, the variable resistor is used to set the voltage level of the FBT after boost as the voltage step-up circuit can go up as high as 3,000 volts.

The L2 coil is wound in a single direction and has just one tap at the centre. The idea is from Tesla's Colorado Springs Notes, in which Tesla disclosed the best method for a resonant driver. The frequency used in this circuit is about 230 kHz.

Question: There is nothing to do with quarter-wavelength, but is there anything with the length of the L1 and L2 coils on quarter-wavelength?
Answer: I think that the phase is more important.

Question: Do you need a Phase-Locked Loop circuit with a certain phase difference?
Answer: Basically, I use a fixed frequency, I have tried a Phase-Locked Loop and the effect is the same.

Question: Do you use direct drive with the spark gap only being used to limit voltage?
Answer: You can use a vacuum tube to drive it.

Question: If you drive it directly, then the loading will be very big and the current will increase, whereas if you use a spark gap, then the spark will become smaller and the current will be steady.
Answer: If the load affects the input, then you cannot drive it even with a spark gap. If you trigger with a spark gap, then the load will not increase the input. The spark gap is just a switch.

Question: Is there any direct Lenz relationship between the load and the primary?
Answer: Once the phase has been adjusted, the primary has no adverse effect on the secondary.

Commenting on his circuitry, 'Salty Citrus' states:

The diode symbols with a tick indicate a Zener diode (or bidirectional TVS-Transient Voltage Suppressor or "varistor"). For example, in this circuit, they are used to suppress the Grid voltage of the MOSFET, to maintain the gate voltage within the range of +20V to -20V. The above circuit is just a description of the structure of the MOS series method. Specific components will be needed for your own requirements considering the MOSFETs being used in your construction.

The voltage E0 can be adjusted. The source can be made using a TL494 IC operating at 12V, or alternatively, an adjustable, voltage-stabilised inverter can be used. The voltage setting depends on the numbers of MOSFETs which are being used in series and the parameters of Grid voltage and the turns ratio of the isolation transformer. The circuit is arranged so that each MOSFET has its own separate isolation transformer, and all of the primary windings of those transformers are connected in series to form a single current path. The number of turns in the primary of each isolation transformer is exactly the same. To drive an IGBT(or MOSFET), VT6 provides a high-frequency pulse current to drive the Gates of the MOSFETs, so as to achieve consistent switching.

In my circuit, the frequency used is 220 kHz, for this frequency, I use six MOSFETs type CMF2012 (1200V, 37A, Resistance Drain-to-Source of just 80 milliohms). This MOSFET from CREE has excellent performance, but you have to design the drive circuit carefully, 2V to 22V for the Gate voltage will be best. I particularly stress that it is very important that MOSFETs operated in series, require voltage balancing and an accurate drive. Especially important is having synchronized drive signals and the rise and fall time of the drive signal should be as short as possible, so that the switching time difference between the MOSFETs will be short, and that improves the high frequency operation.

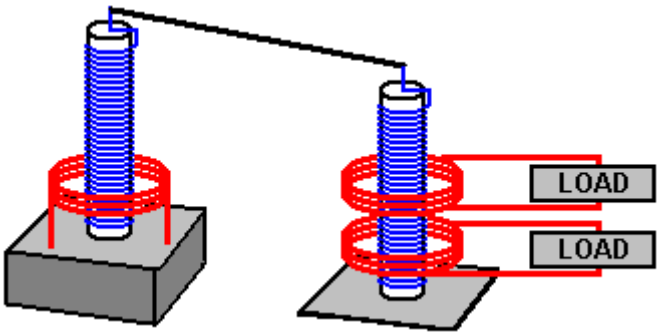
Another Russian Development

At http://www.youtube.com/watch?v=sRLFfs9I18Q there is a video which shows a self-powered motionless generator with an output of 105 watts:

Thanks are due to Wesley for his translation of the Russian soundtrack. No circuit details are available that this time, but the video has “Part 1” in the title and so further details may be provided at a later time.

Tesla Coils Back-to-Back

I have been told of one man who used his common sense and produced an impressive result. He used a Tesla Coil as the driving force, and then used a second Tesla Coil back-to-back with the first one, to step the high voltage back down again. Doing that, he was able to light a series of powerful light bulbs from the “L1” output coils. He also confirmed that doubling the voltage, quadrupled the power output, verifying what Don said. He also found that adding additional coils with bulbs to the output Tesla Coil, did not increase the input power at all, did not cause any of the existing light bulbs to shine any less brightly, and yet lit the additional bulbs. That would appear to be confirmation of Don’s statement that any number of magnetic copies of the original oscillating magnetic field of the first Tesla Coil, can provide a full-power electrical output without requiring any additional input power. I’m no expert, but my understanding of the arrangement is:



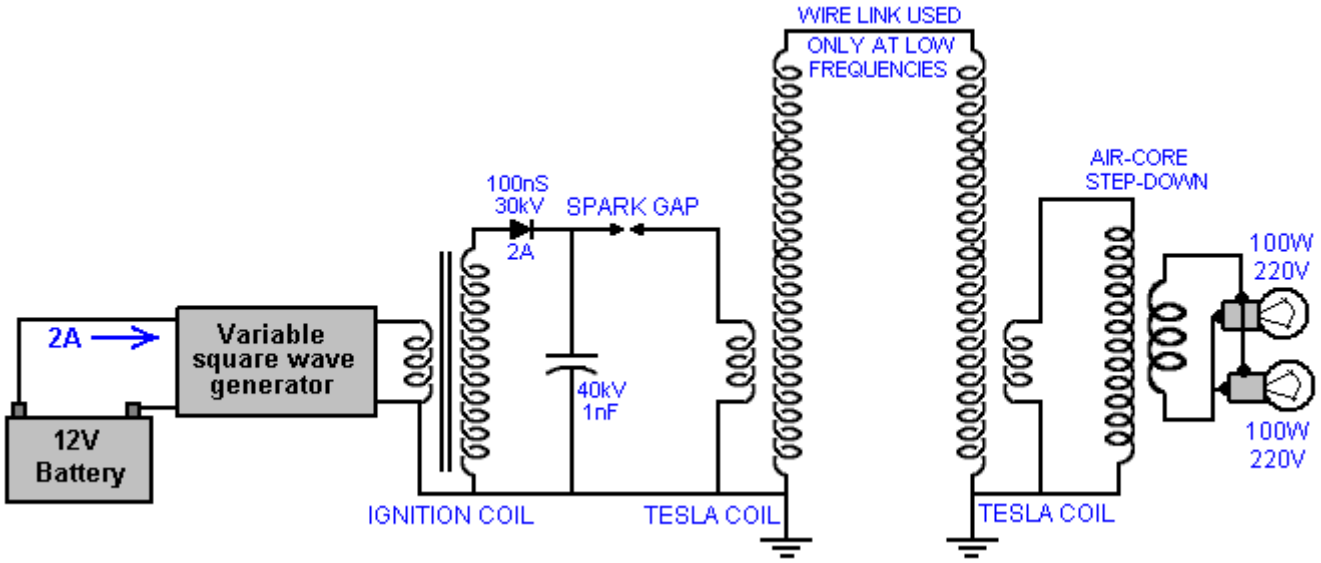
As the large diameter coil is exactly one quarter the length of the smaller diameter coil, there is an automatic resonance of both when the applied frequency is just right. As the first narrow coil is identical to the second narrow coil, they are also automatically resonant together. Again, as the large coils which feed the loads are exactly one quarter the wire length of the narrow coils, they also resonate at the common frequency and at that frequency, the input power is at its minimum while the output power is at its maximum. The spike at the top of each of the narrow coils is connected with a wire to channel the generated power from the first Tesla Coil to the second one.

This arrangement may seem too simple to be effective, but with Tesla technology "too simple" just does not apply. This can be seen clearly from the work of **Nikanor "Nick" Giannopoulos**. Before he ever learned anything about electronics, Nick read and understood Nikola Tesla's "Colorado Spring Notes" (http://www.free-energy-info.tuks.nl/TeslaCSN.pdf 60Mb) and this helped with his present level of understanding. Interestingly, and perhaps not surprisingly, Nick had difficulty with conventional electronics after becoming familiar with Tesla's technology.

Nick used a square wave signal generator adjustable from 50 kHz downwards and with a fully adjustable Mark/Space ratio. This was used to drive an oil-filled car ignition coil, which, as he points out is not a Tesla Coil in spite of the frequently held view that it is. Ignition coils only operate at low frequency due to the limitations of their core material. However, John Stone points out that certain coil designs, such as those for the Fiat 'Punto' car, are built in such a way that replacing the core with ferrite should be possible, and that would allow high frequency operation.

Anyway, Nick uses a standard car ignition coil at lower frequency and uses it to feed a spark gap like this which is constructed from two chipboard screws:

His circuit is:



Nick has had very impressive results from his circuit, although it is still very much a work in progress with more development and testing still to be done. The 24 watt input of 12V at 2A is producing two very brightly lit 220V light bulbs. This does not tell us very much about the actual output power as bulbs are notorious for lighting brightly at low power levels, especially if the frequency is high. But, a very important point is the quality of the light which is an unusual, blue-white colour, quite unlike the colour produced when connected to the 220V mains supply. This is generally a sign of the power being 'cold' electricity. While he has not yet had the opportunity to test it, Nick believes that the circuit as it stands now is quite capable of powering much higher loads, and considering the colour of the light, I would be inclined to agree with him, although anything like that has to be tested and proven before any solid conclusions can be drawn from what is already known about the performance. The circuit performance is much improved if two separate physical earth ground connections are made.

Please don't fall into the trap of thinking that because the sparks are occurring at less than 5 kHz, that the Tesla coils also operate at that frequency. If you strike a bell which vibrates at 400 Hz, does that mean that you have to hit it 400 times every second in order to hear it? Actually, no, you don't, and the same thing applies here where the resonant frequency of the Tesla coils is approximately 650 kHz. The primaries are wound on 100 mm diameter PVC pipe sections and 19 turns of 1.02 mm diameter enamelled copper wire is used for them (19 swg or #18 AWG). The secondary coils are wound on 70 mm diameter PVC pipe using 0.41 mm diameter enamelled copper wire (27 swg or #26 AWG) with a total length of four time the primary winding wire length. As you will see later on in this chapter, resonance in a coil involves a standing wave inside the wire. That standing wave is

created by the signal bouncing off the end of the wire and being reflected back. At frequencies other than the resonant frequency, this results in a constantly changing set of many different waves travelling in both directions and at different intensities (what could reasonably be described as a total mess). When the resonant frequency is fed to the coil, then all of that mess disappears and just one waveform remains, and at any point along the wire, that waveform appears to be stationary although, of course, it is not actually stationary, just the effect of the peaks always occurring at exactly the same spot and the troughs occurring at exactly the same spot, making successive waves look exactly the same as the previous one.

This feature has one very practical aspect, namely that if you run the same wire away from the coil turns to connect to whatever the next circuit component happens to be, then the wave inside the wire will not bounce back at the end of the coil turns but will continue on to the end of the wire before bouncing back. So, the connecting wire length has to be included when reckoning the wire length in the turns of the coil. On the other hand, if the wire in the coil turns is terminated at the ends of the coil and wire of a very different diameter is used for connecting to the next component in the circuit, then the signal inside the wire will bounce back from the sudden change in wire diameter and so the connecting wire length will not be part of the wire length in the turns of the coil. This is an important feature if you are aiming for an exact 4:1 wire length ratio (and 4:1 wire weight) between the Tesla Coil windings in order to impose an automatic resonance between the two windings.

It should be noted that PVC (especially non-white PVC) has a very restrictive effect on high frequency coils. At low frequencies, PVC is ok, but it drags down the coil performance as the frequency rises, lowering the "Q" (for "Quality") factor of the coil. Using acrylic instead of PVC overcomes this. Alternatively, coating the PVC with a high-voltage insulating material such as shellac or one of the proprietary coating agents, will improve matters considerably. The ideal, of course, is to have no former at all and have the coil standing unaided because of it's own strength.

The Screened Transformer of Joseph Boyd.

It is not at all clear if the Boyd power system should be located here or in chapter 7 which deals with aerials. Joseph talks about the way that radio circuits operate and why only very limited power appears to be the limit of radio receivers. He explains a method of extracting serious levels of power from a transmitting coil and a receiving coil, but the higher power levels require an input oscillator, and so, while an aerial and earth can be used for lower power levels, input power is needed for optimum performance. Here is part of Joseph's patent application:

US Patent Application 2008/0129397 5th June 2008 Joseph Boyd

ELECTROMAGNETIC ELECTRIC GENERATOR

Abstract:
An electrical generator that uses a high frequency oscillator in a tuned circuit, set to resonate with the transmitter coil of a full-length high frequency transformer unit, to generate electromagnetic energy, to transform this energy to electrical energy and to collect this energy.

1. Field of the Invention

The present invention is an electrical generator that uses a high frequency oscillator in a tuned circuit, set to resonate with the transmitter coil of a full-length high frequency transformer unit, to generate electromagnetic energy, to transform this energy to electrical energy and to collect this energy.

2. Description of the Related Art

If an oscillator circuit is properly connected to a tuned antenna so that it resonates, a current will flow between the aerial and the ground, and this produces the high frequency electromagnetic air waves and ground waves of our radios and other electronic equipment.

An oscillator of the same type used in electromagnetic wave transmission equipment is used to generate the electromagnetic energy used in this patent. These electromagnetic transmitters are well developed and are used world wide, and broadcast at frequencies that extend from the longest radio waves to the very short ones. Certain radios send their signals great distances, some even travelling around the world.

Although these high frequency electromagnetic energy waves are all around us, this energy has long been considered impossible to collect on a large scale due to the induction characteristics of the electromagnetic wave

as it passes a metallic object. As the wave goes by a wire tuned to resonate at the frequency of the wave, it induces an electrical charge in the wire, but to use this charge, we need another wire to close the circuit and let the charge flow. If we use another wire alongside the first wire and connected to it, the wave induces a charge in it exactly like in the first wire and no current will circulate in the two wires.

This problem of collecting the energy of the wave was solved by the invention of the half-length electromagnetic transformer, but the half-length electromagnetic transformer applies only to the means of collecting the atmospheric energy. The invention of the full-length electromagnetic transformer of this invention, however, allows us to combine the generation of the electromagnetic wave and the electrical converter into one compact unit.

BRIEF SUMMARY OF THE INVENTION

Basically, this unit uses an oscillating electric current to generate an electromagnetic wave, which releases a much greater electric current, and the total electrical energy derived in this manner, is over and above the amount of energy needed to operate the equipment.

It has long been assumed that the only energy involved in high frequency electromagnetic transmission is that supplied by the operator to drive his equipment. The actual energy of the electromagnetic wave is often over a hundred times greater than this since the amount of electromagnetic energy in the earth is practically unlimited, there appears to be no limit to the size of the electromagnetic generators, or to the size of the power plants based on this energy source. This energy is available, world wide, free for the taking.

This energy is related to the light waves and is probably a variation of the light waves, however, the radio type waves are longer than light waves and are vibrating at a lower frequency. Light waves are a source of high energy just for the taking, also. Anything that is brought up to a high heat, will give off light energy. A very small wire in a light bulb, when brought to a high heat will release a light ray of such power that it will go all the way to the moon. This is natural energy, produced by the speed of the earth through space. Using the math of Dynetics, the speed of the earth necessary to give any pound of earth the atomic energy of one pound of uranium, came out exactly the same as the speed of light (186,300 miles per second). The fact that the math came out exactly at this speed leaves little doubt that the speed of the earth through space is the speed of light, and that every pound of material on earth has the energy of one pound of uranium, due to this speed.

The electromagnetic energy in the low frequency range differs from other types of energy, in many ways, but of interest to us is that it is propagated by electrical currents, travels through the air like the light waves, and is detected and may be collected, when it induces an electric charge in a wire.

This is an ideal source of energy. The generators may be hand-held or large enough to replace the biggest power plants. They may be used to drive motor cycles, sleds, autos, trucks, trains, ships and planes. The fact that the output is in the form of electrical energy is, in itself, a great benefit, but the fact that the generating equipment is light and compact is a real plus for all types of mobile equipment.

It is possible that this invention will supply all of the electrical energy needed in the homes making the distribution lines unnecessary, and if used to drive automobiles, our dependence on oil will be a thing of the past.
This invention makes possible an abundance of energy, available to mankind any where in the world. Even the poorest nations will have an abundance of energy.

The oscillating equipment that generates the electromagnetic wave used in this invention includes an oscillator, of some type driving a tuned transmitter coil that resonates with a tuned collector coil in a full-length electromagnetic transformer. The induced current is collected in the collector coil and may be rectified and stored in a battery or used to do work. The oscillator circuit is an ordinary oscillator circuit, driven by a tube, crystal or even an electoral arc, and the tuning means and rectification set-up are standard.

The basic element that is novel to this invention is the full-length electromagnetic transformer unit that is made up of two or more metallic pipe like sheaths side by side. The sheaths are not connected together, electrically. Two or more coils are wound in the sheaths.

The transmitter coil uses an insulated wire, which is threaded up through one sheath, and down through another sheath a number of times forming a long flat continuous circuit of wire inside of the pipe like sheaths. And then the collector coil is threaded up through the sheaths, and wound the same way. The two coils may have a different number of turns. The coils are tuned to resonate at the oscillator frequency and an electromagnetic wave is generated in the transmission coil. The wave induces a charge in that part of the collector coil that is in the same sheath, and next to it, and if the wave is moving up in the sheath, the charges of all of the collector wires in that sheath are moving up, and if the wave is moving down, the charges of all of the collector wires are moving down. But the transmitter wave in one sheath does not induce a current in the wires of another sheath, nor does it

induce a current in a wire outside of that sheath. This allows a current induced in one sheath to freely circulate in the other sheaths or on an outside wire.

When the transmitter coil is brought up to resonance and broadcasts its electromagnetic wave inside of the sheaths, it benefits us in two ways. It keeps the wave inside of the sheaths and prevents it from spreading far and wide, and it concentrates the wave on that part of the collector coil in the same sheath.
Since each of the collector wires are of the same length, say ½ wave length of the electromagnetic wave, and since they are parallel and side by side, a resonant electromagnetic wave induces equal charges in all of the collector wires within the sheath. These induced currents are exactly in phase and are connected in series, so that the voltages add up to an amount proportional to the number of turns.

More than two sheaths may be used with the coils wound inside, or the ff11 length-transformer may be combined with the half-length transformer where the winding is part inside the sheath and part outside of the sheaths. The oscillator may be replaced by an aerial in cases where low power is needed. The inductance coil may be left entirely outside of the transformer, and the coupling made to the transmission coil by magnetic induction.
These and other objects, features and advantages of the present invention will become more apparent upon reading the following specification in conjunction with the accompanying drawing figure.

BRIEF DESCRIPTION OF THE FIGURES

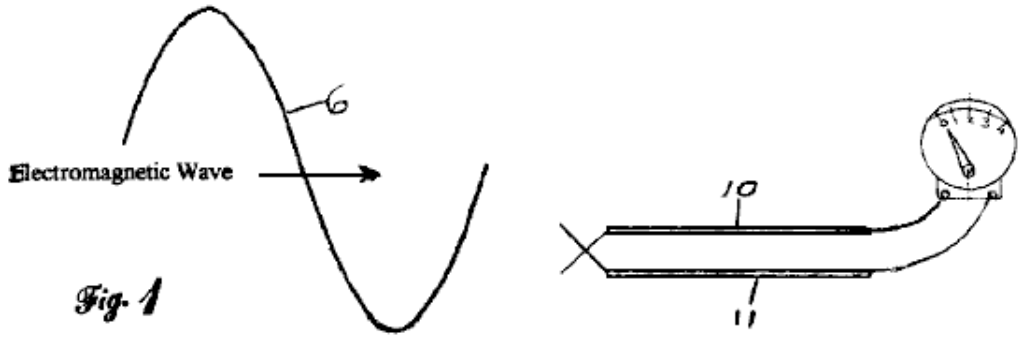


Fig.1 is a view of an electromagnetic wave, due to resonance, passing two side by side wires tuned to the same frequency. The wave generates equal alternating charges in each wire, and no current flows when the wires are connected.

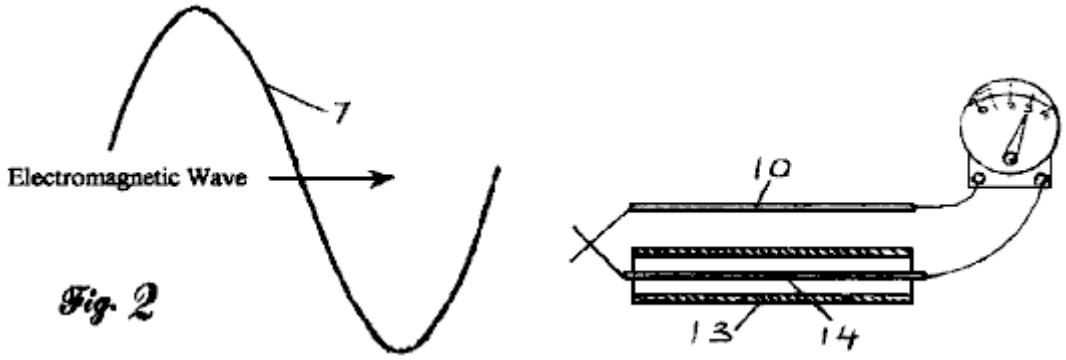


Fig.2 is a view of a resonating electromagnetic wave passing two wires, tuned to resonate, where one wire is surrounded by a metal sheath. The outer metal sheath stops the wave and prevents it from inducing a charge in the shielded wire. The Charge induced in the outside wire now flows freely through the shielded wire.

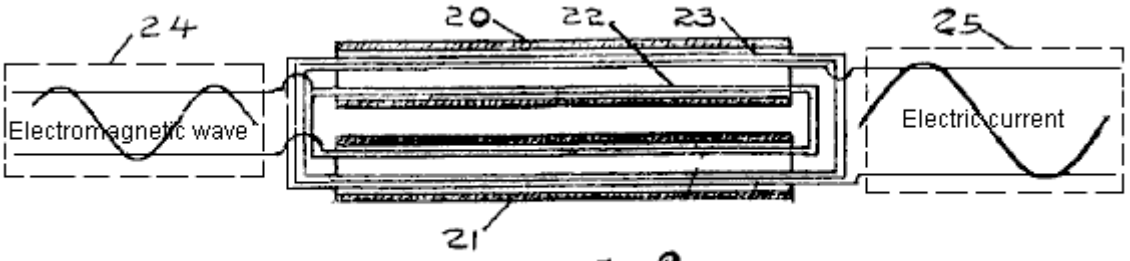


Fig. 3

Fig.3 is a view of a full length electromagnetic transformer made of two metallic pipe like sheaths, having two wires threaded up through the inner opening of one sheath and down through the inner opening of the other sheath, a number of times to form two continuous coils. When an electromagnetic wave is fed into the tuned transmitter coil, an alternating current is induced in the tuned collector coil.

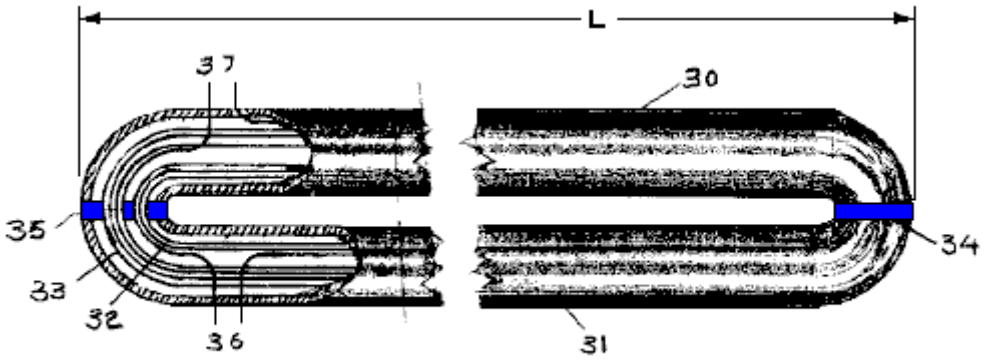


Fig. 4

Fig.4 is a view of a full-length electromagnetic transformer, having two sheaths, insulated from each other, and a cutaway portion shows the coils inside.

Fig. 5

Fig.5 is a view of a full-length electromagnetic transformer, having four sections insulated from each other.

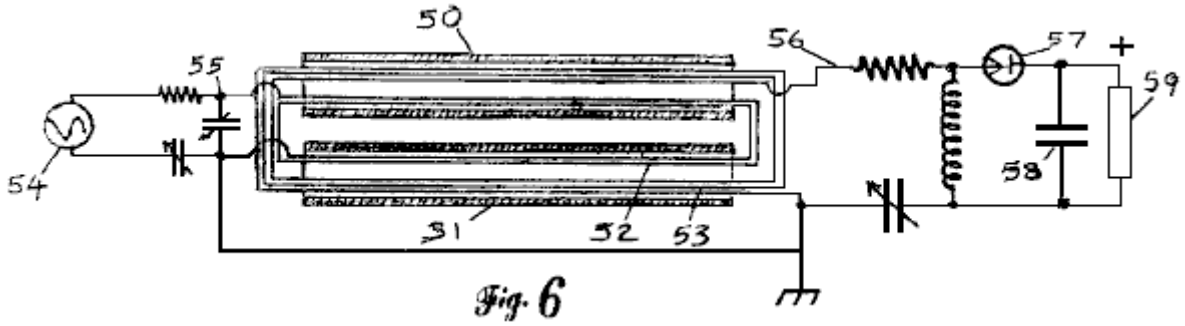


Fig. 6

Fig.6 is a section view of a full length electromagnetic transformer made of two metallic pipe like sheaths, where an oscillator circuit is fastened to a tuned transmitter coil, which induces a current in the tuned collector coil; and a half-wave rectifier circuit converts the high frequency current to DC current.

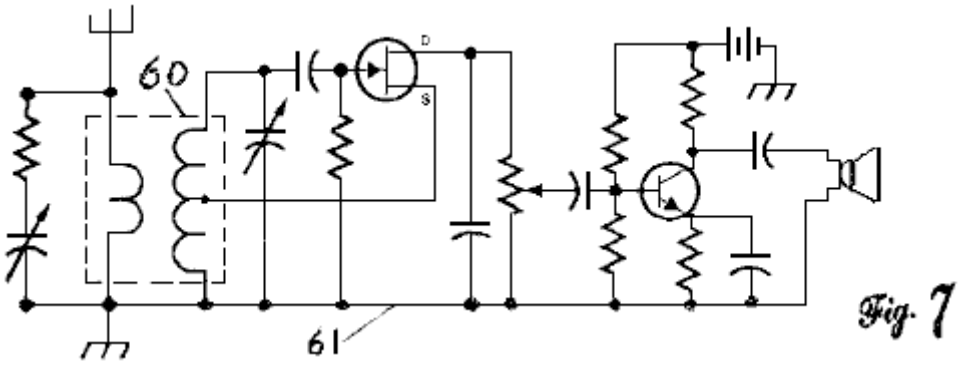


Fig. 7

Fig.7 shows a view of a full-length transformer, used as an inductance in a radio receiver or other similar electronic circuit **61**. In this case the transmitter coil is between the aerial and ground, and the collector coil acts as the radio frequency inductance.

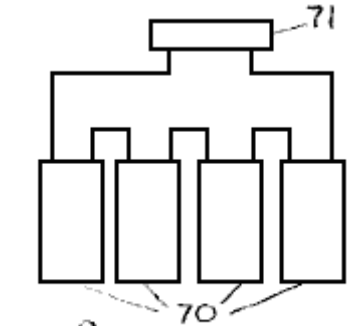


Fig. 8

Fig.8 shows a number of full-length transformers, connected in series. They all have rectifiers to convert the high frequency AC current to DC current and the voltages at the load add up.

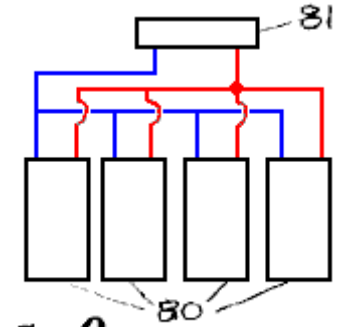


Fig. 9

Fig.9 shows a view of a full-length transformer, connected in parallel. In this case the currents at the load add up.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

When a circuit is made to resonate at some frequency, it generates an electromagnetic wave. These waves move at the speed of light and travel great distances, and this makes possible our radios, televisions and cell phones. These waves, in passing a wire suspended in the air, set up alternating charges in the wire, varying at the frequency of the wave. If we put an inductance, that is tuned to the frequency of the wave, between the wire and the ground the circuit will resonate and charges will flow back and forth on the wire. This allows us to pick out the wave we are interested in, amplify it and read its message. We have long known that the voltage of the wave we choose to receive is greatly amplified when we make our circuit resonate at that wave's frequency, but we have been unable to accumulate this additional energy from the wave because we are limited to what energy we can pick up by the equipment that we use.

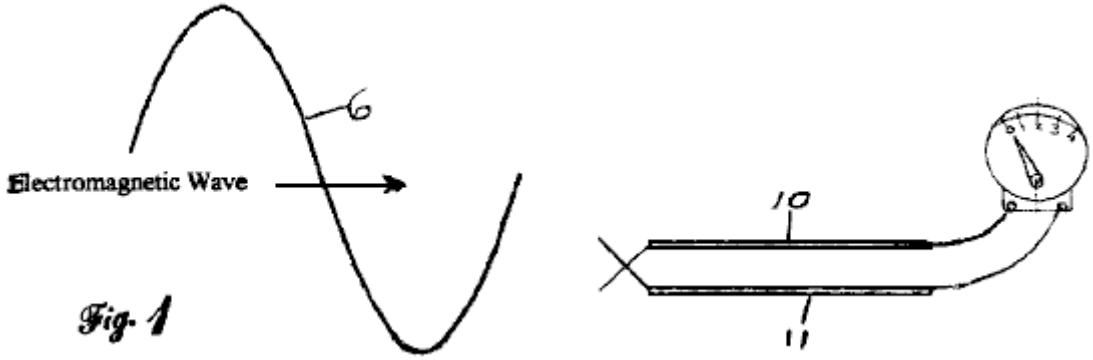


Fig. 1

Fig.1 shows why this is so. The passing electromagnetic wave **6** generates an alternating electric charge in the wire **10**, tuned to the wave frequency to close the circuit so that the induced charge can circulate, we have added a second wire **11**, but if we connect the ends of the wires together, the electromagnetic wave **6** induces a charge in the second wire exactly like the one in the first wire, and no current will flow. Any effort to circulate the charge induced in wire **10** is blocked by an equal opposing charge in **11**.

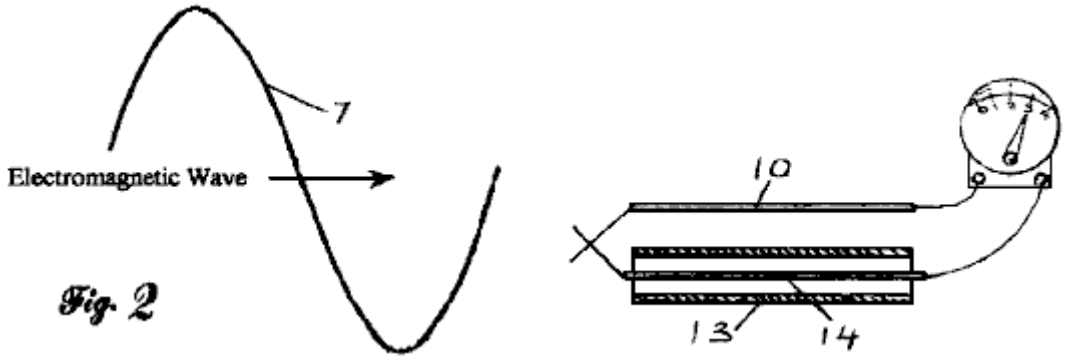


Fig. 2

In **Fig.2** we have replaced the wire **11** with a metallic, pipe like sheath **13** having an insulated wire **14** inside when an electromagnetic wave **7** passes, the wave is stopped at the sheath **14** and does not induce any charges in the inner wire **14**. This allows the inside wire **14** to conduct the charge induced in the wire **12**.

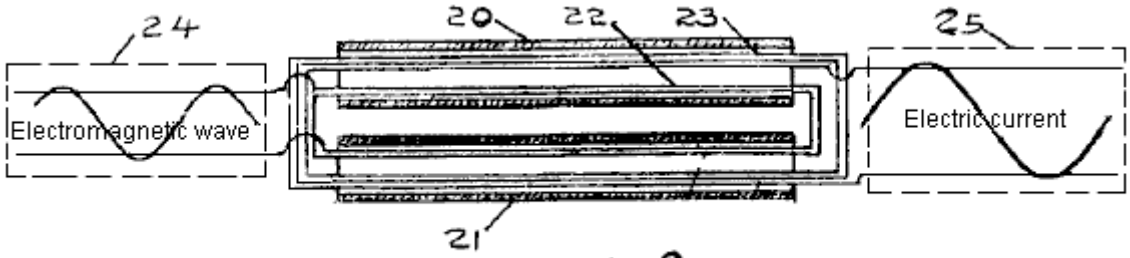


Fig. 3

In **Fig.3** we see two sheaths **20** and **21**, parallel and insulated from each other. Two separate coils **22** and **23** are wound inside the sheaths. The coil **22** is the transmitter coil and it is tuned to resonate at the frequency of the oscillator circuit **24**, and this generates an electromagnetic wave that induces a current in the collector coil **23**, which is adapted and used in the tuned collector circuit **25**.

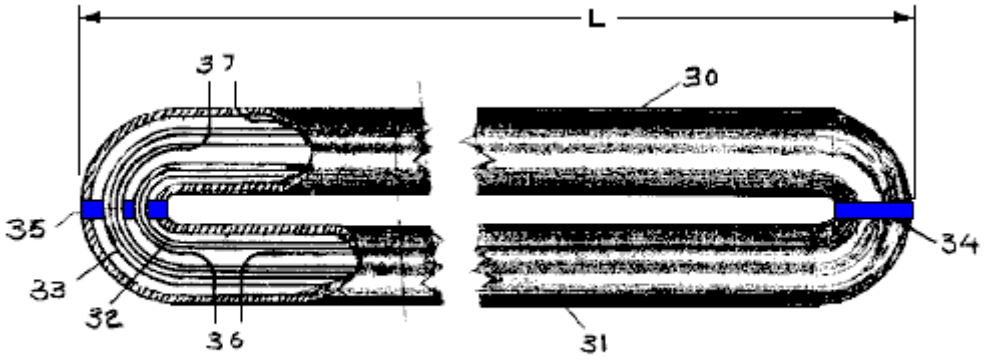


Fig. 4

Fig.4 is a view of a full-length electromagnetic electric generator using two separate sheaths **30** and **31**, where the electromagnetic wave is almost completely contained within the sheaths. The sheaths are parallel and insulated from each other by insulators **34** and **35**. The transmitter coil **32** and the collector coil **33** are shown in the cut-away, and are wound completely inside the sheaths. Additional coils may be added as needed. The transmitter coil **32** is tuned to resonate at the frequency of the oscillator circuit, which connects in at **36**. And this generates an electromagnetic wave that is contained within the sheaths and induces a current in the collector coil **33**, which is tuned to resonate at the wave frequency and is coupled to the collector circuit at the terminal **37**. The insulating separators **34** and **35** are necessary to prevent induced current from flowing in the sheaths **30** and **31**.

Fig. 5

A number of sheaths may be combined, as is shown in **Fig.5**. In this case four sheaths **40**, **41**, **42** and **43** are combined, so that the coils wound inside the sheaths are in series, and the sheaths are electrically separated by the insulators **44**, **45**, **46** and **47**. The tuned circuit for the transmitter coil is hooked up at connectors **48**, and the collector circuit connects to the collector coil at **49**.

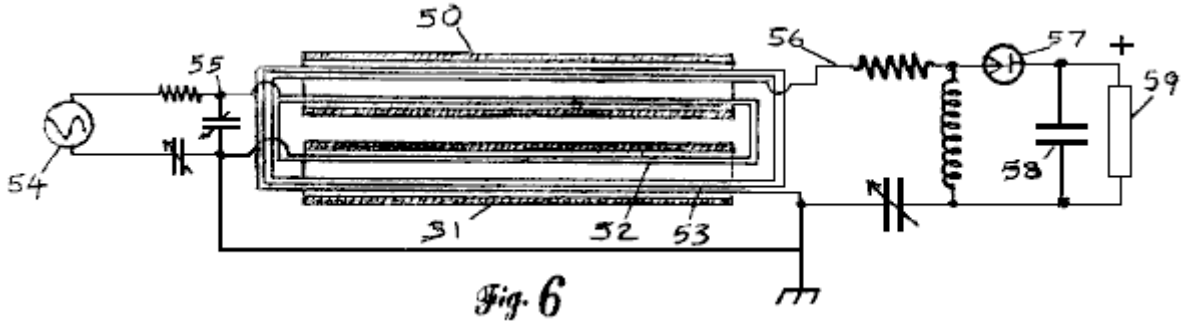


Fig. 6

A simple circuit for operating the full-length electromagnetic transformer is shown in **Fig.6**. The two sheaths **50** and **51** are shown with the transmitter coil **52** and collector coil **53**. The transmitter coil is connected to the tuned circuit **55**, which is driven by the oscillator **54**. The oscillator **54** is tuned to oscillate at some frequency and the transformer circuit **55** and collector circuit **56** are tuned to resonate with it. At resonance the transmitter coil **52** emits an electromagnetic wave which induces a current in the collector coil **53**. Shown here is a half-wave rectifier **57** and capacitor **58** connected to the load **59**. The rectifier is necessary on each collector unit, because we can add the direct currents of the separate units, but the alternating currents of the units might be out of phase and without the rectifier they would cancel out.

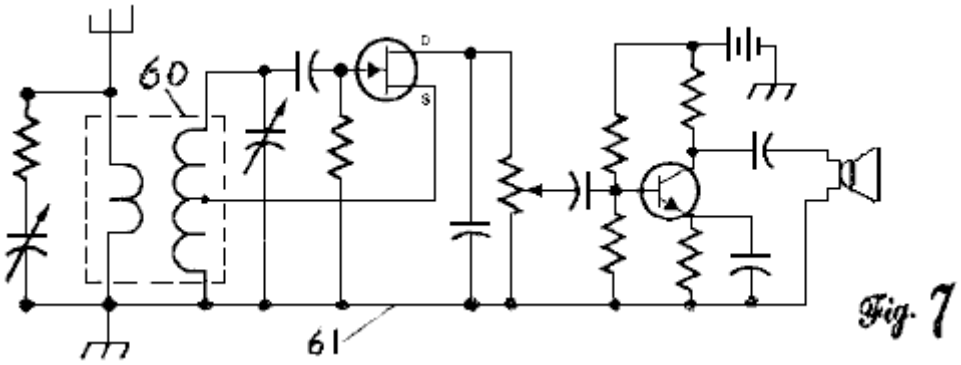


Fig. 7

Fig.7 shows the transmitter coil connected between an aerial and the ground, where the aerial circuit and collector circuit are tuned to resonate at a desired frequency to greatly boost the sensitivity and to amplify the signal. This type application will work equally well with transmitters.

Also, using the half-length transformer as an aerial and the full-length transformer as the radio frequency transformer works the same with both receivers and transmitters. The oscillator circuit can use a high frequency magnetic inductance and the full-length transformer unit to resonate with a half-length transformer which acts as an aerial. A unit such as this would greatly add to the power of a radar system.

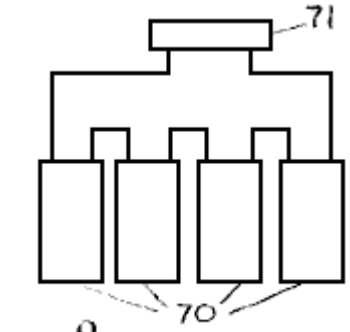


Fig. 8

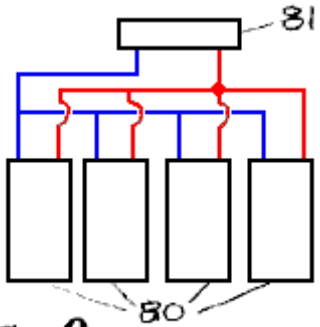


Fig. 9

The full-length units, when combined with rectifiers, may be connected in series as in **Fig.8**, or in parallel as in **Fig.9**

While the invention has been disclosed in it's preferred forms, it will be apparent to those skilled in the art that many modifications, additions, and deletions can be made therein without departing from the spirit and scope of the invention and its equivalents as set forth in the following claims.

Claims:

1. An electric generator that uses an oscillating circuit to resonate with inductances enclosed in a full-length electromagnetic transformer unit, comprising: two or more parallel, metallic, pipe-like sheaths, insulated from each other; an electromagnetic transmitter coil, which is wound inside the sheaths, using an insulated wire, threaded up through the inside opening of one of the sheaths and down on the inside of a second sheath, one or more times, to form a continuous coil; a means of tuning the transmitter coil to the frequency of the oscillating circuit, so that the transmitter coil resonates and generates an electromagnetic wave that is contained within the sheaths; a collector coil composed of an insulated wire, threaded up through the inside opening of one of the sheaths and down on the inside of a second sheath, one or more times, to form a continuous coil, wherein the electromagnetic wave of the transmitter coil induces an electric charge in the collector coil; and a means of tuning the collector coil to resonate at the frequency of the oscillating circuit; so that the electromagnetic wave induces a current in the collector coil, and since the coil is contained within the sheaths, the current is free to flow, and the high frequency alternating current, so generated, can be used elsewhere in an electronic circuit, or altered and used as a power source.

2. The full-length electromagnetic transformer unit of claim 1, where the collected current is rectified, and the DC current is stored in capacitors and used to do work.

3. The full-length electromagnetic transformer unit of claim 1, where the transmitter coil acts as a tuned

inductance, in a receiver, between the aerial and the ground, and a tuned collector coil resonates to increase the sensitivity and the amplitude of the signal.

4. The full-length electromagnetic transformer unit of claim 3, where the tuned transmitter coil resonates with an oscillator and an aerial to act as a transmitter for radios, televisions, cell phones, radars and computers.

5. The full-length electromagnetic transformer unit of claim 1, where the lengths of the sheaths are some multiple of the wave length of a particular electromagnetic frequency.

6. The full-length electromagnetic transformer unit of claim 5, where a number of the units are connected in parallel to increase the current.

7. The full-length electromagnetic transformer unit of claim 5, where a number of the units are connected in series to increase the voltage.

8. The full-length electromagnetic transformer unit of claim 5, where the parts are reduced to a size small enough to fit in, and supply power to, a cell phone, a lap-top computer, or other electrical appliance.

9. The full-length electromagnetic transformer unit of claim 1 further comprising three or more, metallic, parallel, side by side, pipe like sheaths, where the sheaths are not electrically connected, and have two or more coils wound inside the sheaths.

****

Up to the present time (June 2013), I have not heard of anyone attempting to replicate the design shown in this patent, and so it is just being put forward here in case somebody wants to try it. It appears to be a clever technique. The power gain is maximised by tuning the transmitting and receiving coils to the frequency of the oscillator, although in practice, it is highly likely that the oscillator frequency would be adjusted to the transmitter coil as it is so easy to adjust the frequency of an oscillator.

Boyd does not go into great detail about attaining resonance, and that is generally a major difficulty in any design which does not have automatic tuning. It needs to be borne in mind that the length of wire in each coil (and possibly its weight) is a key factor. Boyd talks about the coils possibly having the same number of turns, and that is fine **provided** that the coils are identical in size, that is, having the same shape when viewed from the top and the same depth of turns when viewed from the side, and exactly the same number of turns with each coil having exactly the same wire length. Resonance in a length of wire, whether it is laid out straight or wound into a coil, tends to confuse many people. Richard Quick's very clear explanation of resonance in any length of wire, in his US patent 7,973,296 of 5th July 2011 is very helpful. He says:

"Quarter-Wave" Resonance; Standing Electromagnetic Waves"

One of the two main types is electrical resonance is referred to here as quarter-wave resonance. This type of resonance depends almost entirely on the length of a wire element For reasons described below, if a segment or length of wire is one quarter as long as the "voltage waves" which are travelling through the wire, then a set of "reflected" waves will be added to the emitted waves, in a synchronised alignment which creates stronger "superimposed waves".

Accordingly, an understanding of the "quarter-wave" phenomenon will help a reader understand how a straightforward and easily-controlled factor (i.e., the length of a wire ribbon which will be used to form a spiral coil) can help create a "quarter-wave" resonant response, which will create the types of electromagnetic pulses and fields referred to as "standing waves".

The speed at which a voltage impulse is transmitted through a metal wire is extremely fast. It is essentially the same as the speed of light, which travels 300 million meters (186,000 miles) in a single second (that distance would circle the earth more than 7 times).

If wavelength (in meters) is multiplied by frequency (cycles per second), the result will be the speed of light, 300 million meters/second. Therefore, the wavelength of an "alternating current" (AC) voltage, at some particular frequency, will be the speed of light, divided by which frequency.

Therefore, using simple division, if an alternating voltage operates at a frequency of 1 megahertz (MHz), which is a million cycles per second, then the "wavelength" at that frequency will be 300 meters. If the frequency halves become 500 kilohertz, the wavelength becomes twice as long (600 meters); and, if the frequency were to increase to 2 megahertz, the wavelength drops to 150 meters.

It should be noted which the term "cycles" is what scientists call "a dimensionless unit", which drops out and becomes silent when other physical terms are multiplied or divided.

At AC frequencies of 10 kilohertz or greater, the common references to “alternating current” (AC) voltage begin using a different term, which is “radio-frequency” (RF) voltage. Accordingly, RF voltage is a form (or subset) of AC voltage, which operates at frequencies higher than 10 kilohertz. RF power generators are readily available, and are sold by numerous companies which can be easily located by an Internet search, using the term “RF power generator”. For example, Hotek Technologies Inc. (hotektech.com) sells two RF power generators, called the AG 1024 and AG 1012 models, which can provide output power at frequencies ranging from 20 kHz to 1 MHz; the 1012 model has a power output of 1000 watts, while the 1024 model has a power output of 2000 watts. The output frequency of any such RF power supply can be adjusted and “tuned” across the entire range of operating frequencies, merely by turning knobs or manipulating other controls in a power supply of this type.

In a wire having a fixed and unchanging length, the easiest way to create a “standing wave” is to adjust the RF frequency emitted by a power supply with an adjustable frequency, until the “tuned” frequency creates a wavelength which is 4 times as long as the wire. This principle is well-known to physicists, and it is commonly referred to as “quarter-wave” behaviour, since the length of the wire segment must be one quarter as long as the wavelength. Since it is important to this invention, the principles behind it are illustrated in a series of drawings provided in **Fig.1** to **Fig.4**, all of which are well-known prior art.

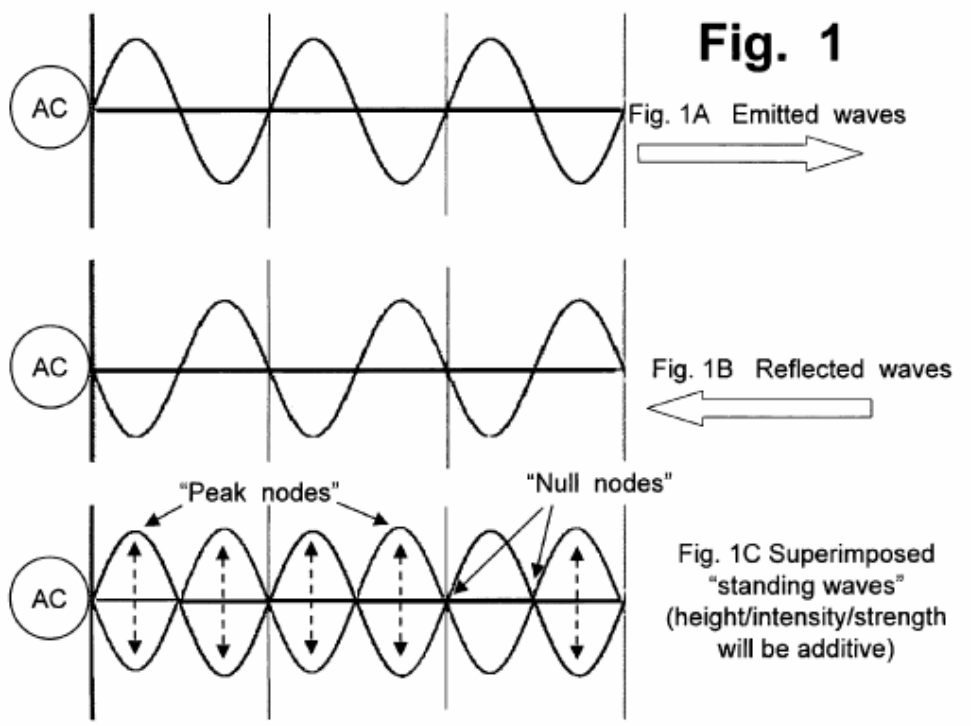


Fig.1A indicates an idealized wavelength of an alternating voltage, depicted by a sine wave which is being sent from an AC power supply (shown by a circle at the left end of a horizontal straight wire) into the “input” end of the wire. The voltage waves travel through the wire towards the right, as indicated by the block arrow in **Fig.1A**. When the waves reach the end of the wire, they cannot leave the wire (at least, not in a simplified and “ideal” system, which is being assumed and used here to explain the principle of how a simple straight wire can create a standing wave). Therefore, the voltage wave will effectively “bounce” or “reflect” back from the tip of the wire, and the “reflected wave” will begin travelling back through the wire, going in the opposite direction, as indicated by the left-pointing block arrow in **Fig.1B**.

Because of the laws of conservation of energy, the reflection and “return travel” of these types of waves, when they bounce off the tip of a wire, is actually quite good, and rather efficient, as discussed below, provided which the wire tip does not emit sparks, arc discharges, or other forms of “escaping” electrical energy.

Accordingly, **Fig.1A** depicts a set of “emitted waves” travelling towards the right, while **Fig.1B** depicts an idealised set of “reflected waves” travelling toward the left along the same wire.

Fig.1C illustrates what happens when both sets of waves (emitted and reflected) are superimposed on each other. Since the two sets of waves are travelling at exactly the same speed, and since they have exactly the same wavelength, they will create a “standing wave” pattern when they are added together. As can be visualised from **Fig.1C**, there will be a set of locations, along the length of the wire, which can be referred to as “peak nodes”, where the AC voltage reaches it’s maximum.

At a location halfway between a pair of adjacent “peak nodes”, there will be a spot which can be called a “null node”, a “zero node”, a trough or valley node, or similar terms. At each “null node” location, the AC voltage will

appear to be not fluctuating at all. Those are the sites, along the length of the wire, where each “positive” hump (created by a sine wave travelling toward the right) will be counter-balanced and offset by a “negative hump” with exactly the same height, travelling at an identical speed toward the left.

As a result, this type of response within a wire creates a “standing wave”. If the instantaneous voltage is measured at a “null node”, it would appear that nothing is happening, in terms of fluctuating voltage. Furthermore, the “null node” will not be moving, along the length of the wire; instead, it will appear to be standing still.

This can be demonstrated, in a coil, by using a “grounded lead” to test for voltages along the length of a coil. If a “grounded lead” coupled to a volt meter is used to touch the surfaces of a series of strands in a non-insulated coil (such as a coil made of thin copper tubing, wrapped around a plastic cylindrical shape, as used in the types of large transformers used by hobbyists to create “Tesla coils” which will emit large and visually impressive electrical arcs), the “test lead” will detect no apparent voltage at a null node, which will occur at some particular strand in the coil. At a different strand of the coil, the “test lead” will detect an alternating voltage which has twice the strength and intensity of the voltage being emitted by the power supply.

If voltage is measured at a “peak node”, the voltage will be doing something which can be called, using vernacular or laymen's terms, “the full-tilt boogie”. The AC voltage levels will be moving back and forth, between: (i) a very high and intense positive voltage, to (ii) an equally intense negative voltage. This is indicated by the “bubble” shapes shown along the wire in **Fig.1C**.

The “bubbles” which are shown in **Fig.1C** can help someone understand how standing waves are created, and how they act in a synchronised manner. However, which drawing fails to show another result which is very important in what actually happens in a standing wave. For purposes of description and analysis at this introductory level, the system can be assumed to be “ideal”, which implies a perfect “mirror-image” reflection of each wave from the right end of the wire. An “ideal” system also implies that no reflections occur at the left hand end of the wire where the power supply is located, and all “reflected” wave activity simply ceases. In real circuits and wires of this type, second and third order reflections do in fact occur, and they are used to further increase the strength and power output of these types of systems; however, those additional factors and “harmonics” should be ignored until after the basic principles of this type of system have been grasped and understood.

In an ideal system, when the reflected waves (which are travelling toward the left, in the wire segments illustrated in **Fig.1**) are “superimposed” on the emitted waves (travelling toward the right), the “peak” positive voltage which will be instantaneously reached, at the highest point of each “bubble” shown in **Fig.1C**, will occur when the positive peak of an emitted wave crosses a mirror-image positive peak of a reflected wave, travelling in the opposite direction. Accordingly, when those two “positive peak” values are added to each other, the instantaneous positive peak voltage which will occur, in the wire, will actually be twice as intense as the “positive peak” voltage being emitted by the AC power supply.

An instant later, at that exact point on that segment of wire, a negative peak voltage will be created, which will be the sum of (i) the negative peak voltage emitted by the power supply, and (ii) the negative peak voltage of a reflected wave also will pass through, travelling toward the left. At which instant, when those two negative peak voltages are added to each other, the instantaneous negative voltage which will occur, in the wire, will be twice as intense as the “negative peak” voltage being generated by the AC power supply.

A more accurate and representative visual depiction of a “standing wave” in a wire would actually show the heights of the peaks as being twice as tall as the peaks of the emitted voltage waves, and the reflected voltage waves. However, which depiction might confuse people, so it usually is not shown in drawings of “standing waves”.

Accordingly, the instantaneous response in the wire, at a location halfway between two “null nodes”, is doing something which can fairly and properly be called “the full-tilt double double boogie”. The “double double” phrase (note which it contains not just one but two “doubles”) was added to that phrase, for two reasons:

(i) To emphasise the fact that each and every voltage peak (maximum positive, and maximum negative) will be twice as strong, and twice as intense, as the maximum positive and negative peak voltages emitted by the power supply; and,

(ii) to point out that the frequency of the superimposed “bubbles”, shown in **Fig.1C**, is actually twice as fast as the frequency of the AC cycle which is emitted by the power supply, as discussed below.

The “twice the intensity” result is directly comparable to what an observer will see, if a large mirror is placed behind a light bulb in an otherwise dark room. The mirror effectively keeps the room dark, everywhere behind the mirror, so there is no “magical doubling” of the light in the room; which would violate the basic law of conservation of energy. Instead, what the mirror does is to shift light away from the backside of the mirror, and keep that light energy on the reflective side of the mirror. Anyone standing in front of the mirror will see two apparent light bulbs. Both of those light bulbs (the original bulb, and the reflected image) will have the same brightness (if the mirror is perfect). Therefore, the mirror will double the intensity of the light energy reaching the observer.

That same effect, in a circuit, will happen if the end of a wire acts like a mirror. If a wire does not have any components which will cause it to become an active “emission source” (which is the behaviour of transmission

antennas and certain other components), in a way which efficiently releases voltage-created energy into the atmosphere, then the basic rules which require conservation of energy will prevent that energy from simply disappearing and ceasing to exist. As a result, even if the end of a wire is not designed to be a perfect reflector, a large portion of the voltage wave will indeed reflect off the wire tip, and travel back through the same wire, in a "second pass".

To understand adequately, the type and amount of "wave reflection" which occurs at a wire tip, consider what happens if a light bulb is shining in a room which has shiny, glossy white paint on all the walls and ceilings; then, consider how it would look if the same light bulb were located in a room with all of the walls and ceilings painted "matt black". The total amount of light which would be available, to carry out a task such as reading a newspaper, clearly would be much greater in the white room, because light reflects off white paint, even though white paint does not even begin to approach the type of "reflection quality or clarity" which a mirror creates. The difference in what happens, when light intensity in a room painted matt black is compared to a room painted a glossy white, does not arise from the presence or absence of "reflection quality or clarity"; instead, it is governed by the laws of conservation of energy. When light shines on to a surface which is painted matt black, the light energy is absorbed by the paint, and it literally warms the paint up. In contrast to that, glossy white paint will not absorb light energy, so it reflects the light back out, for a "second pass" through the air which fills a room.

Because of the laws of conservation of energy, and without depending on any "quality of reflectance" characteristic of wire tips, electrical energy cannot simply disappear, when it reaches the end of a wire. Instead, there are only two things which can happen to that energy:

(i) the electrical energy can be emitted into the surroundings, such as by emitting sparks, arcs, or radio-frequency signals which will carry energy; or

(ii) if the energy is not emitted by the tip of the wire, then, by simple necessity and because of the basic law of conservation of energy, it must be reflected back into the wire, and it will be forced to travel back through the wire again.

If a wire has a long and tapered tip, then the reflected wave might become somewhat diffused, and it might lose some portion of the "clarity" of the wave. However, since wavelengths in the frequencies of interest here are hundreds of meters long, the type of tip created by a conventional wire cutter will not create any significant diffusion, in a reflected wave. And, unlike the white-painted walls of a room, there is not a large area which is available, at the tip of a wire, which can create scatter, spread, or diffusion. As a result, the tip of a wire will be a relatively efficient mirror-type reflector, when an AC voltage is "pumped" into one end of the wire.

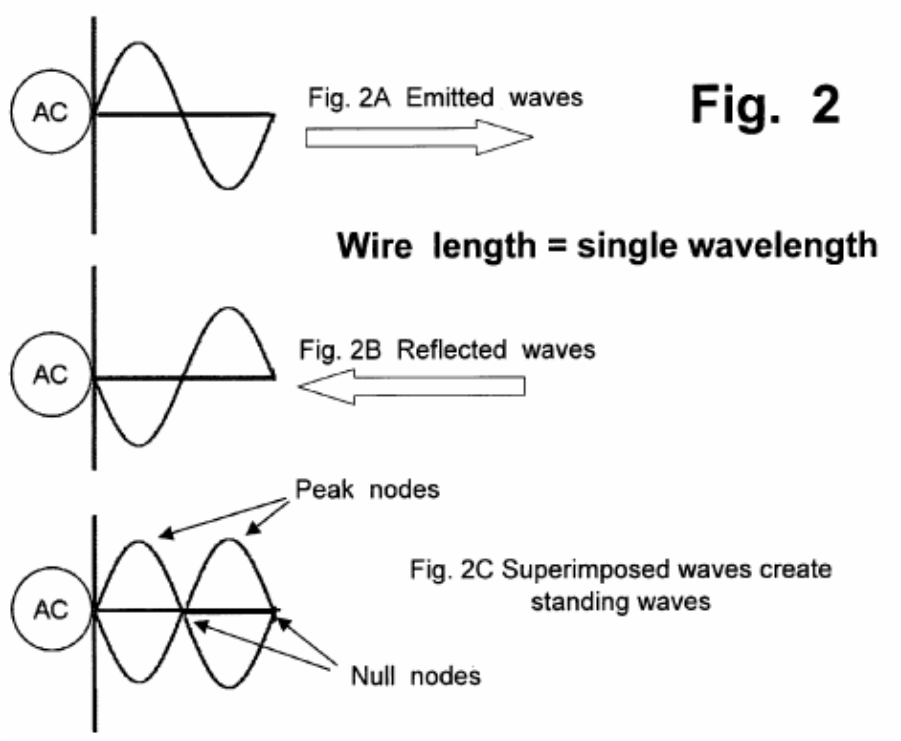
The second factor mentioned above, when the "double-double" boogie phrase was mentioned, relates to a doubling of the frequency of a standing wave. When a standing wave is created in a wire by reflection of an emitted AC voltage wave, the frequency of the standing wave is, quite literally, double the frequency of the emitted wave.

This can be seen, visually, by noting that in the emitted AC voltage, shown in **Fig.1A**, a single complete wavelength contains both a "positive hump" and a "negative hump". Accordingly, three complete sine waves, divided into three segments by the imaginary vertical lines, are shown in **Fig.1A**.

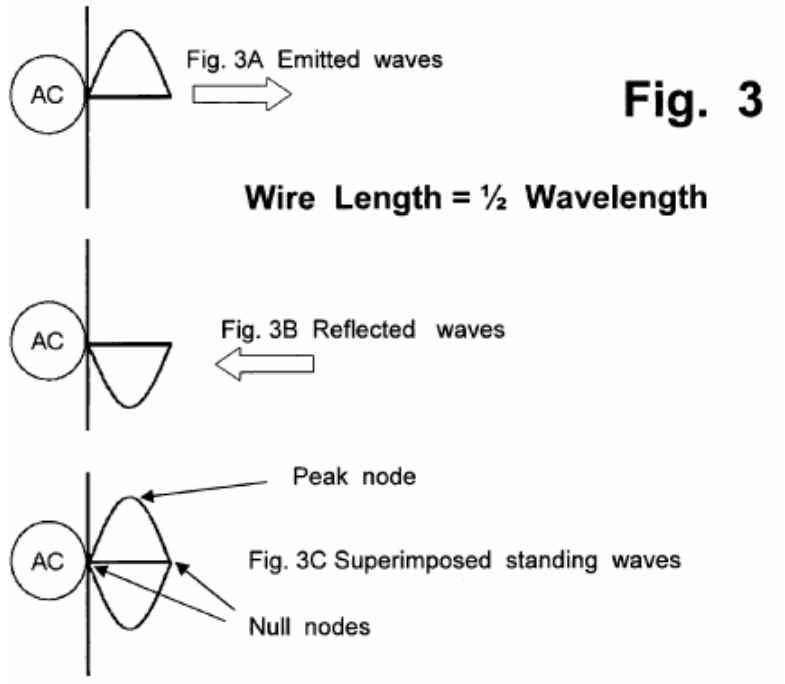
By contrast, each and every "bubble" shown in **Fig.1C** depicts a complete and total "wavelength", in a standing wave. Six of those standing wave "bubbles" fit into exactly the same length of wire which holds only 3 emitted wavelengths from the power supply.

The "frequency doubling" effect of standing waves is important, because AC systems can convey and release energy in a manner which increases, as the frequency of the AC voltage supply increases. To some extent, this is analogous to saying that, if a motor can be run at twice the speed (while still generating the same torque), then the work output of that motor can be twice as great, at the higher speed. That analogy is not entirely accurate, since work output from an electric device which uses AC power depends on "area of the curve" functions which occur when sine waves are involved. Nevertheless, as a general principle, if the frequency of the voltage peaks increases, then power output will also increase, in many types of electric circuit components.

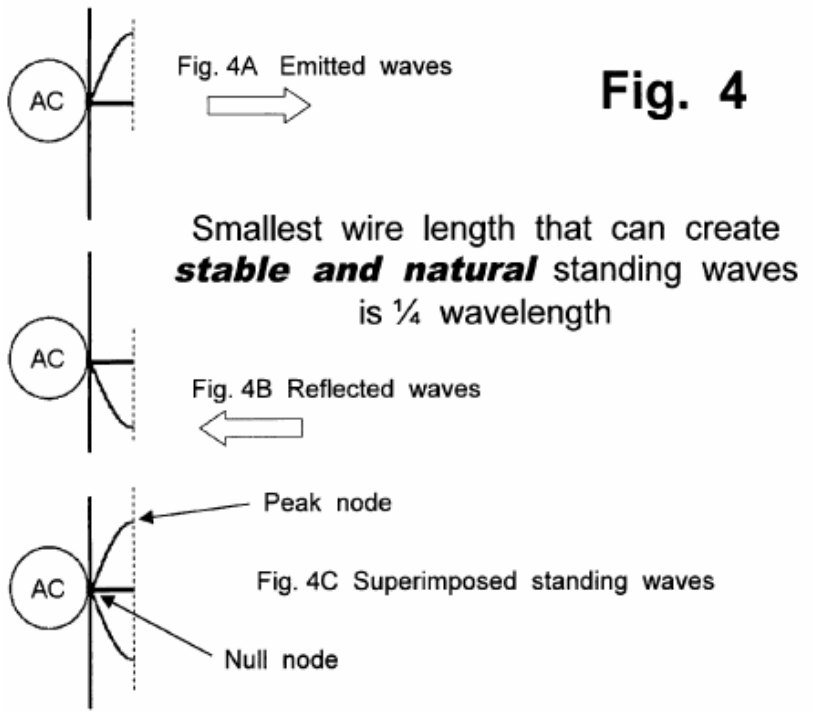
In the three panels of **Fig.1**, the wire length is three times as long as the wavelength of the voltage from the power supply. However, to create standing waves, a wire length does not need to be any particular multiple of the wavelength of an AC voltage. As can be seen by considering **Fig.1C**, the same types of "bubbles" would be created: (i) if the wire length were exactly twice as long as the wavelength; or, (ii) if the wire length were the same length as the wavelength.



Accordingly, **Fig.2** (which includes **Fig.2A** showing an emitted wave, **Fig.2B** showing a reflected wave, and **Fig.2C** showing the superimposed "bubbles") shows what happens in a wire segment which has a length which is equal to a single wavelength from an AC voltage at a fixed frequency. A resonant standing wave will be formed, with a frequency which is double the frequency of the input AC voltage. which same result will apply, in a wire having any length which is an exact (integer) multiple (such as 1x, 2x, 3x, etc.) of the wavelength of the AC voltage being pushed (or forced, driven, pumped, etc.) into the wire segment.



Moving to still shorter wires, the same principle also applies to any wire with a length equal to one half of an AC voltage wavelength. As shown in **Fig.3** (which includes **Fig. 3A** showing an emitted wave, **Fig. 3B** showing a reflected wave, and **Fig. 3C** showing the superimposed "bubbles"), if the wire length is one half of the wavelength, a natural and resonant standing wave will still form, with a frequency which is double the frequency of the input AC voltage.



Finally, moving to a still shorter wire, the same principle also applies to any wire which has a length equal to one quarter of an AC voltage wavelength, as shown in **Fig.4A**, **Fig.4B**, and **Fig.4C** Even though it does not stretch across or cover a complete "bubble", the standing wave shown in **Fig.4C** is nevertheless a stable, natural, and resonant "standing wave", with a frequency which is exactly twice the frequency of the input AC voltage.

It is possible to create partially stable and semi-resonant responses, using one eighth, one sixteenth, or shorter lengths of wire, by using additional devices which can remove electrical power from the system, or which can generate effects which are usually called "harmonics". However, those are not the types of natural and stable responses which can be created by a simple, basic system consisting of nothing more than: (i) a wire having a fixed length and a "reflective" tip; and (ii) an AC power source with a frequency which can be "tuned" until it creates a resonant response in any wire segment having a suitable length.

Therefore, since quarter-wave wire lengths are the shortest lengths which can create natural and stable standing waves, the conventional term which is commonly used, to describe what happens when a wire creates a resonant standing-wave response, is a "quarter-wave" response.

In some devices, telescoping components (or other elements which can alter the effective length of a wire-type element) can be used to alter the ability of the element to respond to a fixed wavelength. Many types of antennas use this approach, if they need to process signals which are being transmitted on fixed and known frequencies. However, those examples are not relevant to spiral coil reactors, which will use an approach which involves tuning and adjusting the frequency of the voltage which is being supplied to a reactor, until a resonant response is observed in coils with fixed and unchanging lengths.

It should also be noted that certain types of "tuning" elements (such as capacitors, which can have either fixed or adjustable capacitance levels) can also be coupled electrically to a wire, in a manner which "emulates" adding more length to that wire. This approach can be used to alter (or increase the range of) the frequencies to which a wire circuit will respond resonantly.

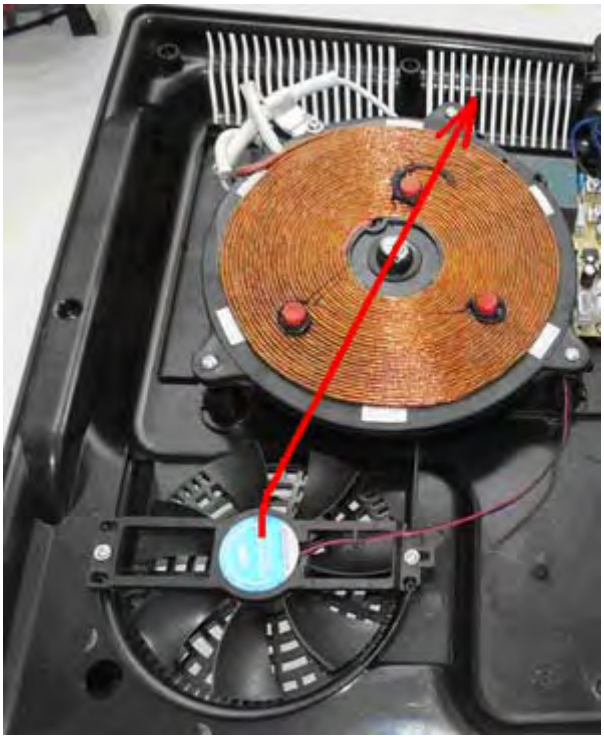
The 'Gegene' Magnetic Arrangement.

As we have seen from what Don Smith has said, a very effective method of gaining additional power is to make a high frequency magnetic transmitter as that allows several outputs to be taken from the transmitter without increasing the input power in any way. Recently, a clever idea for a simplified version of this has been shared on the web. As far as I am aware, this device was first presented by the Lithuanian 'FreeEnergyLT' whose website is at http://freeenergylt.narod2.ru/dynatron/

and the information then replicated and documented by J L Naudin on his website http://jnaudin.free.fr/gegene/indexen.htm and named ‘Gegene’ being short for ‘Great Efficiency Generator’. The clever idea is to use a commercial induction hot-plate as the transmitter. These have recently become available at low cost, this one:

Sold in the UK by Maplin, has power levels adjustable from 300 watts to 2000 watts, and costs only £30 delivered to your address. These devices operate by generating a powerful high frequency oscillating magnetic field which induces eddy currents in any magnetic material placed on the surface of the cooker. That is, cookware which is made of cast iron or steel (not stainless steel which is supposedly non-magnetic). The heating is very rapid and completely uniform across the item of cookware which is very helpful when cooking. The hot plate is controlled by sophisticated electronics which will not switch on unless there is an iron object on the plate and which varies the frequency and current in a way chosen by the designer.

The circuitry produces the magnetic field by pulsing current through a large, flat coil in the centre of the case as can be seen in this photograph of a typical induction plate with the case opened:



The brown coil gets hot, and so there are spacers on it to prevent the cool outer casing form picking up the heat of the coil. There is also a fan which draws air in from underneath the case and blows it across the coil in order to keep the heat down.

In order to use this magnetic transmitter, we need to place a suitable output coil on the plate, and power a load from the energy collected by that coil. This is a fairly recent idea and so there is still a good deal of experimentation going on, testing different coils, and various loads. It is generally agreed that the best load is a non-inductive load with halogen lamps and ordinary filament light bulbs being recommended. Halogen lamps are used in some low-cost commercial heaters, and they are very effective method of radiant heating. In his video at http://www.youtube.com/watch?v=LbAhUwHvJCE, Laurent powers seven separate 400-watt halogen lamps using a small 800-watt maximum plate which has a small 120 mm diameter transmitter coil:

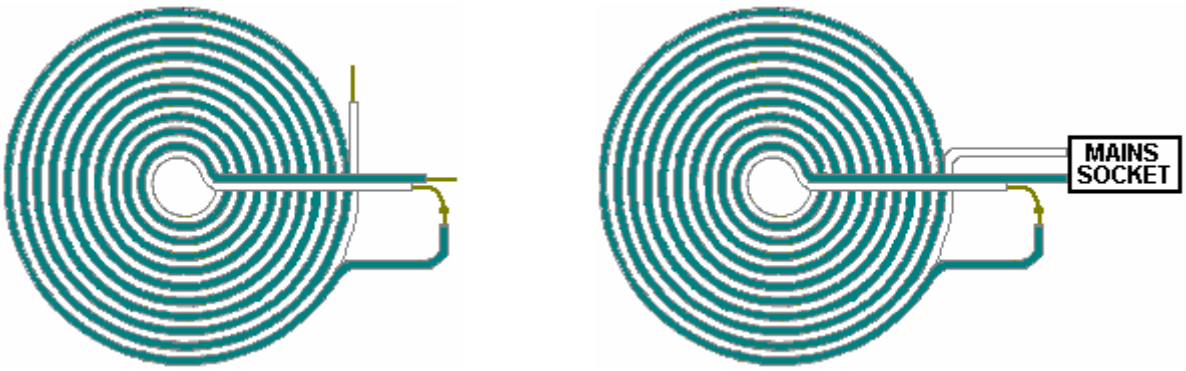
No particular power output is claimed by Laurent, but as you can see, the 2800 watts of halogen lamps are brightly lit while a wattmeter on the input to the plate reads just 758 watts. It seems to be fairly clear that there is a significant power gain with this arrangement. Then, Laurent places an additional coil on top of the first one and shows it lighting a 100-watt filament light bulb very brightly:

It is actually quite difficult to see the brightness of lamps shown in a video as the video camera automatically turns down the brightness of the recording. The important point here is that there is significant power output from a second coil, without there being any increase in the power input to the transmitter coil in the induction plate.

There are many different designs of electronics in commercial induction plates. Most will not start operating until a magnetic object is placed on top of the plate. If that is done, then the object needs to be removed very promptly as it heats up very rapidly. Fortunately, most plate designs keep operating as soon as the induction process is started and so there in no problem with removing the metal cookware (or whatever is used to start the process). Laurent's very small induction plate does not have that protection circuitry and so starts up as soon as it is switched on.

Jean-Louis Naudin uses a 2000-watt induction plate turned down to its 1000-watt setting. It has a 180 mm diameter pick-up coil. He says that for him, it is essential to have at least 1500-watts of load or else the induction plate will shut down with an error code indicating that no cookware is present.

The coils used are Tesla bi-filar pancake types, typically, attached to a thin sheet of MDF or plywood, say 2 mm thick, with superglue. Laurent's 120 mm coil has ten turns and Jean-Louis' 160 mm coil has sixteen turns, needing about 5 metres of twin-core wire, and Laurent's about 2.5 metres of wire. I suggest that the wire should be rated for mains voltage and have, perhaps 1 sq. mm cross-sectional area of copper wire in each conductor. A Tesla pancake coil is wound like this:



Please remember that this arrangement involves high voltages and so is not suitable for newcomers to electronics. This presentation is strictly for information purposes only and it is not a recommendation that you attempt to implement anything shown here, and if you choose to do so, then the responsibility is yours and yours alone.

Tariel Kapanadze's Self-Powered Generators

Tariel Kapanadze, like Don Smith, appears to have based his work on that of Nikola Tesla. There has been a video on the web, of one of his devices in operation, but it appears that the video has been removed. However, part of it can be seen here: http://www.youtube.com/watch?v=l3akywcvb9g The video commentary was not in English and so the information gathered from it is not as complete as it might be. However, in spite of that, a number of useful things can be learned from it. Unfortunately, Tariel refuses to share the details of his designs.

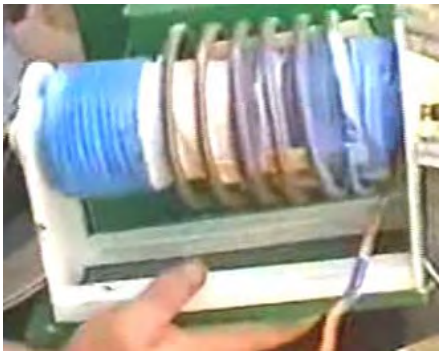
The video shows a demonstration being staged in a back garden, I believe, in Turkey. Strong sunshine was casting dense shadows which made video detail less than perfect. Essentially, Tariel demonstrated one of his builds of a Tesla-style free-energy device, powering both itself and a row of five light bulbs.

One of the most encouraging things about this video is that the construction and operation was of the most basic kind, with not the slightest suggestion of expensive laboratory work or anything high-precision. This is most definitely a backyard construction within the scope of any knowledgeable person.

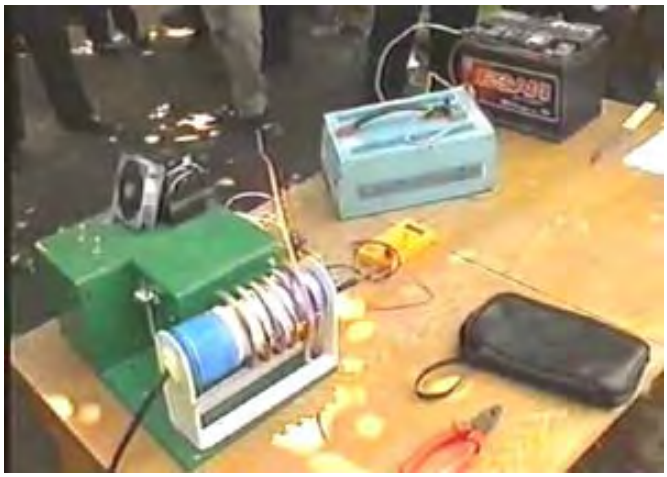
Electrical connections were made by twisting bare wires together:

and where necessary, tightening the twist with a pair of pliers:

This shows clearly that a high-power and very useful free-energy device can be made with the most simple of construction methods - no expensive connectors here, just a zero-cost twisted connection.



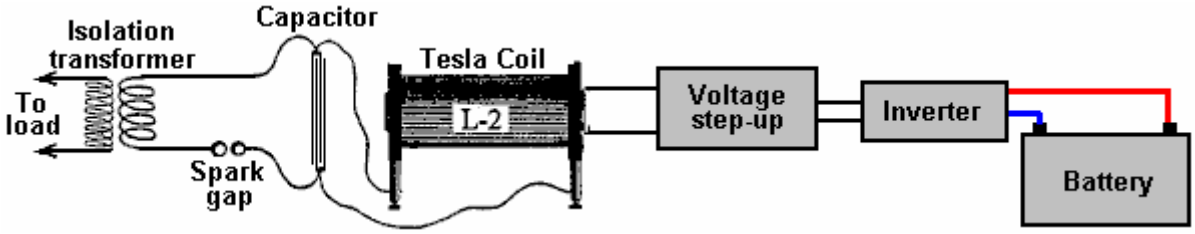
The device being displayed is a Tesla Coil powered, earth-connected system of the type already described. You will notice that the thick primary winding is not placed at one end of the central secondary winding but is much closer to the centre of the coil. Remember that Don Smith states that if the primary coil is placed centrally, then the amount of current which the coil can deliver is very large, in spite of the fact that most people think that a Tesla Coil can only produce trivial currents. Notice also that this Tesla Coil appears to be mounted on a cheap kitchen-roll holder. I have seen it said that Tariel makes a new device for each demonstration and takes it apart afterwards, so if that is correct, then it is likely that there is no great effort or expense involved in making one of these systems.



The main operational components are shown here, placed on one small table. There is a lead-acid battery (which is removed later in the demonstration), what appears to be an inverter to produce mains AC voltage from the battery, a high-voltage step-up system housed in a green box for safety reasons, a Tesla Coil, a spark gap mounted on the box and a fan-cooled component, probably a solid-state oscillator system driving the Tesla Coil. Not seen in this picture, is an item contained in a small box which might well be a high-voltage capacitor.

Two earth connections are organised. The first one is an old car radiator buried in the ground:

and the second is a bare wire wrapped around a garden tap's metal pipe and twisted tight as shown above. It is distinctly possible that the circuit is based on this circuit of Tesla's:



Perhaps, the battery powers the inverter which produces mains voltage, which is then stepped up to a high voltage level by the enclosed electronics. This then drives the Tesla Coil, producing both very high voltage and current with the capacitor storing the energy as a reservoir. The spark gap then pulses this energy, driving the primary winding of the isolation transformer which produces a lower voltage at substantial current (depending on the current-handling capacity of the transformer itself) powering the load, which in this case, is a row of light bulbs.

The load is a row of five light bulbs hung from a brush handle placed across the backs of two chairs:

As you can see, this is not exactly high-tech, high-cost construction here, with all of the materials being used for other things afterwards.

Initially, the battery is used to power the inverter and it is demonstrated that the current being drawn from the inverter is substantially less than the power entering the load. In conventional terms, this appears impossible, which is an indication that the conventional terms are out of date and need to be updated to include the observed facts from demonstrations such as this.

As the system is putting out a good deal more power than is required to drive it, might it not be possible to use part of the output power to provide the input power. This is often called "closing the loop" and it is demonstrated in this video as the next step.

First, the circuit is altered so that the input power connection to the inverter is taken from the output. Then the circuit is powered up using the battery as before. The battery is then disconnected and removed altogether, and the people helping with the demonstration pick up all of the active items and hold them up in the air so as to show that there are no hidden wires providing the extra power from some hidden source. The items on the table are not part of the circuit:

There is some additional information on Tariel including videos of some of his more powerful, newer designs at http://peswiki.com/index.php/Directory:Kapanadze_Free_Energy_Generator#Official_Website although it has to be said that there does not appear to be very much on him or his work available at this time.

In December 2009 an anonymous contributor e-mailed to say that Kapanadze returned to the ex-USSR republic of Georgia and that the video soundtrack is in the Georgian language and after the demonstration, the interview is in Russian. He has kindly translated the parts which relate to the device, as follows:

Question: What are you showing us today?
Answer: This is a device which draws energy from the environment. It draws 40 watts as it starts up, but then it can power itself and provide an output of 5 kilowatts. We don't know how much energy can be drawn from the environment, but in an earlier test, we drew 200 kilowatts of power.

Question: Is it possible to solve the energy problems of Georgia?
Answer: We consider that they have already been solved.

Question: Please tell us in simple terms, how your device works.
Answer: (1) Power is drawn from the battery to get the device running
(2) If we want, we can use part of the output power to drive a charger and charge the battery
(3) When the device is running, we can remove the battery and it then operates self-powered. This particular unit can deliver 5 kilowatts of power which is enough for a family. We can easily make a version which supplies 10 kilowatts. We don't know what the practical power limit is for a unit like this. With this particular device we have here, we do not draw more than 5 kilowatts as we don't want to burn out the components which we used in this build.

Question: Does your invention pick up current from mains wires?
Answer: The mains has nothing to do with this device. The energy produced comes directly from the environment.

Question: What do you call your device and do you dedicate it to anyone?
Answer: I would not dream of claiming this device to be my invention, I just found something which works. This is an invention of Nikola Tesla and all the credit is his. Tesla has done so much for mankind but today he is just forgotten. This device is his invention, his work.

Question: Why are you so sure that this is a design of Nikola Tesla's?
Answer: Because I worked from his invention - his design. I discovered how to get automatic resonance between the primary and secondary windings. The most important thing is to achieve resonance. Melnichenko came close to solving this problem. The government of Georgia refuses to take this invention seriously.

Question: You said that resonance must be maintained. Which parts resonate?

Answer: Here (pointing to the green box) and here (pointing to the Tesla Coil mounted on the top of the green box). The resonator is inside the green box and at present, it is secret until patented.

Question: How much would one of these units cost?
Answer: When mass produced, it would cost between 300 and 400 US dollars for a unit which has an output of 5 or 6 kilowatts.

Question: How much did it cost you to build this demonstration device?
Answer: About eight thousand (currency not specified). Parts had to be got in from twenty different places.

Question: Is this your house?
Answer: No, I rent this place because we have sold all that we have to make these devices. And, having done it, the government and many scientists say "We are not interested because a device like that is impossible and can't possibly exist!". I have not been allowed to make a presentation to them, but people who understand the Tesla Coil understand how this device works.

Kapanadze is an architect by profession and has not had any training in either physics or Electrical Engineering. The information on which this design was based was downloaded free from the internet.

One of the most important aspects of this video is the confirmation it gives for the work of Tesla and of Don Smith, in that it shows clearly, yet again, that large amounts of energy can be drawn from the local environment, without the need to burn a fuel. Another video: http://www.youtube.com/watch?v=gErefbcTz-U

People frequently ask for construction drawings or alternatively, outlets where they can buy one of his devices. Unfortunately, Tariel is not willing to share the details of his designs and so they will probably never be manufactured. Many people have tried to analyse and replicate his design.

Kapanadze Analysis by William McFreey

Shown below are two analysis papers on Tariel Kapanadze's designs. These papers by William McFreey are written primarily for physicist, and so they may not be easily understood by you if you are not familiar with scientific symbols and notation, so please feel free to move past this section..

While William McFreey's paper is highly technical, the designs disclosed in it have been prototyped and proved to be fully functional. However, it needs to be understood that there are several ways of extracting free-energy from our local environment. For example:

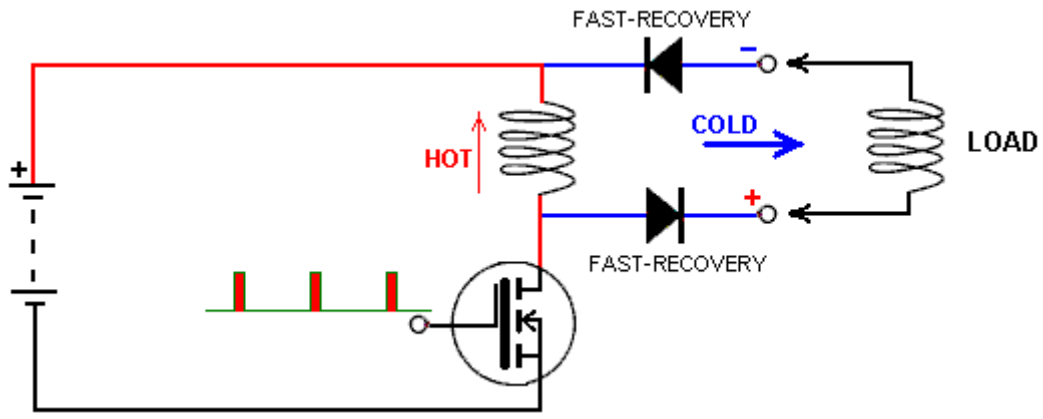
1. Thomas Henry Moray demonstrated that it is possible to draw kilowatts of energy directly from the local environment, using an aerial of relatively minor dimensions.
2. Stanley Meyer demonstrated that water can be separated into a hydrogen and oxygen gas mix using tiny power levels. This allows an ordinary petrol generator to be run self-powered on water alone.
3. Bob Boyce has produced a pulsed toroidal circuit which allows a battery to recharge itself.
4. Robert Adams produced a motor/generator design where permanent magnets generate an electrical output far in excess of the power needed to operate it.
5. Carlos Benitez designed a system which is self-powered and which provides kilowatts of excess power using just simple, standard electronic components.
6. Thane Heins has built and demonstrated simple asymmetrical transformers where the output power is more than thirty times greater than the input power.
7. Clemente Figuera produced a split-transformer design where the Lenz "Law" effect does not apply and so it produces far greater output power than input power.
8. John Bedini designed a pulsed flywheel motor/generator system which ran self-powered for years. Jim Watson built a large version which had many kilowatts of excess power.
9. Dr Oleg Gritskevitch built a self-powered motionless toroidal generator which produced 1.5 megawatts for two years.
10. James Hardy has demonstrated how the jet of water from a powerful water-pump can spin a generator fast enough to self-power the pump and supply additional electrical power for other equipment.
11. Mikhail Dmitriev has produced a gravity-powered generator which uses a small electrical motor to deflect weights on a rotor and that system produces kilowatts of excess power drawn from the gravitational field.

12. Lawrence Tseung has shown how a magnet embedded in an iron frame produces a greater output than the input power when it is powered with DC pulses.
13. Lawrence Tseung has also demonstrated that a heavy rotor powered by electromagnet pulses can have a substantially greater output power than the power needed to drive it.
14. Veljko Milkovic has demonstrated how combining a pendulum with a lever produces a system which has far greater mechanical output power than the mechanical power needed to operate it.
15. Richard Willis has shown how pulsing a magnetic material can produce kilowatts of excess power.
16. James Kwok has shown that introducing air into the bottom of a tank of water can generate tens of kilowatts of excess power through buoyancy.
17. Dietmar Hohl has published his design where permanent magnets cause permanent rotation of a cylinder without the need for any external source of power.
18. Mummar Yildiz has demonstrated a 300 watt permanent magnet only motor and then taken the motor completely apart to prove that there is no other source of power.
19. Lawrence Tseung has produced a variation on the Joule Thief circuit which has greater output power than the input power (something which I personally have built and can confirm).
20. Floyd Sweet produced a system where a tiny amount of power is used to wobble the magnetic poles of a specially conditioned magnet, producing thousands of times greater output power than the input power needed to power the system.

There are many other proven systems. The two papers from William McFreey can be freely downloaded from: http://www.free-energy-info.tuks.nl/McFreey.pdf.

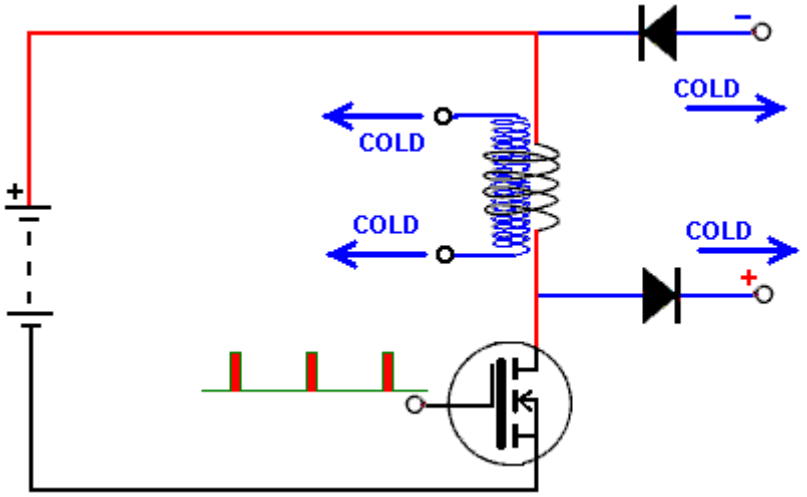
The Cold Electricity Coil of 'UFOpolitics'

A man who uses the forum ID of 'UFOpolitics' has been sharing his insights and experiences on various different forums, such as the one dealing directly with the production and use of cold electricity in solid-state circuits: http://www.energeticforum.com/renewable-energy/10529-my-motors-got-me-tap-into-radiant-energy-1.html His insights are unusual and very important. His basic statement is that if a coil is pulsed, using a circuit like this:



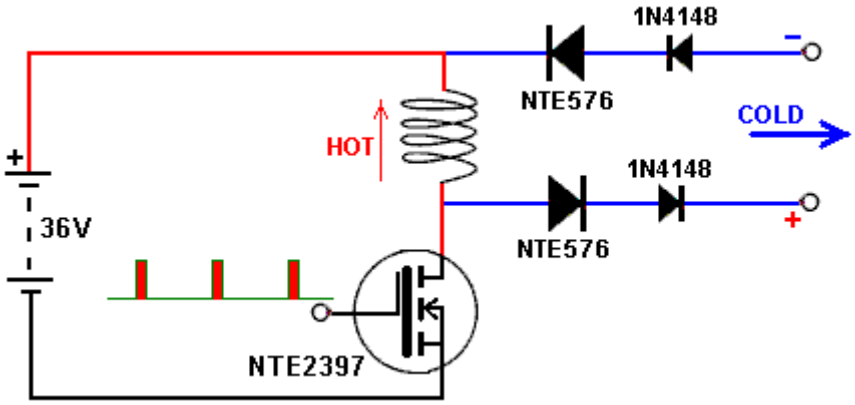
then conventional hot electricity pulses the coil when the transistor is switched ON, but if that current is switched OFF rapidly, then there is an inflow of cold electricity into the coil from the surrounding environment. That inflow of energy can be collected and diverted to power a load through the use of two high-speed diodes which can carry considerable current as the power inflow is substantial. The inflow of energy occurs when the transistor is switched OFF and so it is desirable to have the transistor switched off for most of the time, in other words, a low percentage Duty Cycle for the transistor. There **must** be a significant load on the cold electricity output. If there is not, then the cold electricity will flow back into the hot electricity section of the circuit and it may damage the transistors. Tom Bearden states that resistors boost cold electricity rather than hindering it's flow, so the load should be a coil, a DC motor with brushes or a fluorescent light bulb.

It has been observed that the incoming energy tends to flow inwards towards the centre of the coil, so an additional method of collecting this extra energy is to place a second coil inside the main coil, and wound in the same direction as it, like this:



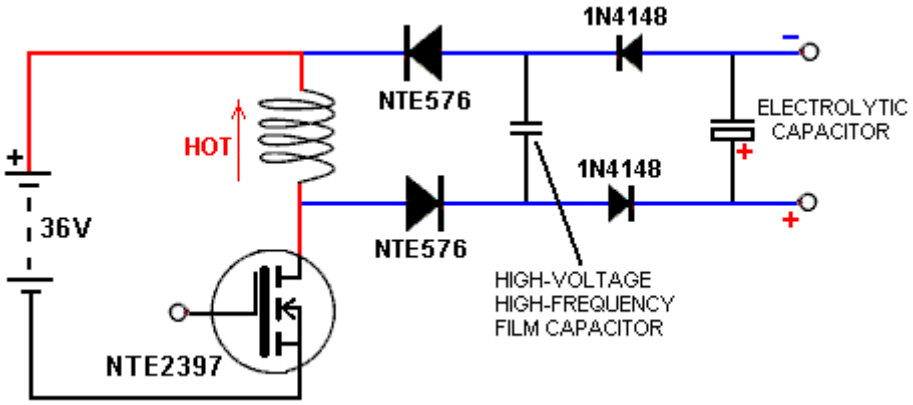
This provides two separate, independent cold electricity power outputs. Diodes are not needed for the inner 'secondary' coil. This inner coil is a pick-up coil and is not related in any way to the number of turns in the hot electricity pulsing coil. Instead, this coil collects inflowing cold electricity during the period when the pulsing coil is switched OFF. The hot electricity pulsing coil can be wound directly on top of the extra pick-up coil or the extra coil can be wound separately and placed inside the main coil spool.

Very surprisingly, it is recommended that the powerful high-speed diode used to channel the cold electricity out of the circuit, be followed by a small 1N4148 silicon epitaxial planar high-speed diode (75V 0.45A) as this is said to clean up the cold electricity output even more. It is important that the cold electricity has to encounter the more powerful silicon diodes before reaching the 1N4148 diodes, so the order of the diodes is very important, and should be as shown here:

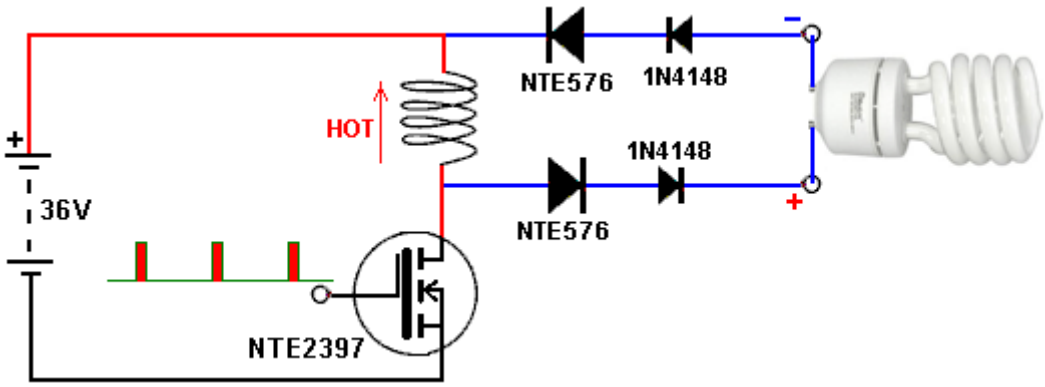


Alternative diodes for the NTE576 (6A, 35nS, 400V) are the NTE577 (5A, 70nS, 1000V) and the HFA16PB (16A, 19nS, 600V). The main requirement is high-speed operation, voltage rating of at least 400V and current rating of at least 5 amps..

There is one additional thing to be done with this circuit when a DC output is required and that is to apply filtering to the output. First, when the energy has passed through the NTE576 (or equivalent) power diodes, it encounters a high-frequency (low capacity) high quality film capacitor placed across the output in order to siphon off any high-frequency voltage ripple before it is passed through the small 1N4148 diodes and into a smoothing and storage electrolytic capacitor. Storing the cold electricity in the electrolytic capacitor converts it into conventional hot electricity.



While this circuit looks like something which you just switch on and it works, that is not the case as there is an essential start-up procedure where the signal applied to the transistor is started at just a few cycles per second and 50% duty cycle and that input is then adjusted carefully and slowly while monitoring the voltages and currents produced by the circuit. This is a seriously powerful system with the capability of producing a major power output.



It is very important that the circuit is not powered up without a suitable load on the cold electricity output. A suitable load is a self-ballasted 230-volt fluorescent light. It must be understood that just flipping the power switch to it's ON position is not sufficient to get an inflow of cold electricity. Instead, it is necessary to progress the start-up sequence carefully, and a fluorescent light is particularly helpful for doing this although a neon bulb is also a popular choice of temporary load, because these devices allow the current flow in the load to be assessed visually.

Before switch-on, the input oscillator is set to 50% duty cycle and minimum frequency. Then the frequency is raised very slowly, causing the lamp to start flashing. As the frequency is raised, the current drawn from the battery needs to be monitored as it is the current flowing through the transistor, and the current is kept down by lowering the duty cycle progressively. This process is continued carefully and if successful, the colour of the light produced will initially be purple or green before reaching continuous bright white light. Videos showing the light produced and the fact that it is not dangerous to life or affected by water can be seen at http://www.youtube.com/watch?v=W1KALMgFscg&list=UUdmFG5BeS0YnD2b5zasXXng&index=1&feature=plcp .

The driving force is a series of powerful magnetic pulses, and implementing the physical circuit to achieve that requires careful construction. The battery driving the circuit is a 36 volt combination of cells. The coil is wound as an air-core construction on a 2-inch (50 mm) diameter spool and the DC resistance is arranged to be about 1.4 or 1.5 ohms. This, in turn, requires a substantial drive from the transistor and so it is normal to connect six powerful output transistors in parallel in order to spread the current flow between them as well as dissipating the heat generated across several transistors bolted to a common heat-sink of generous area.

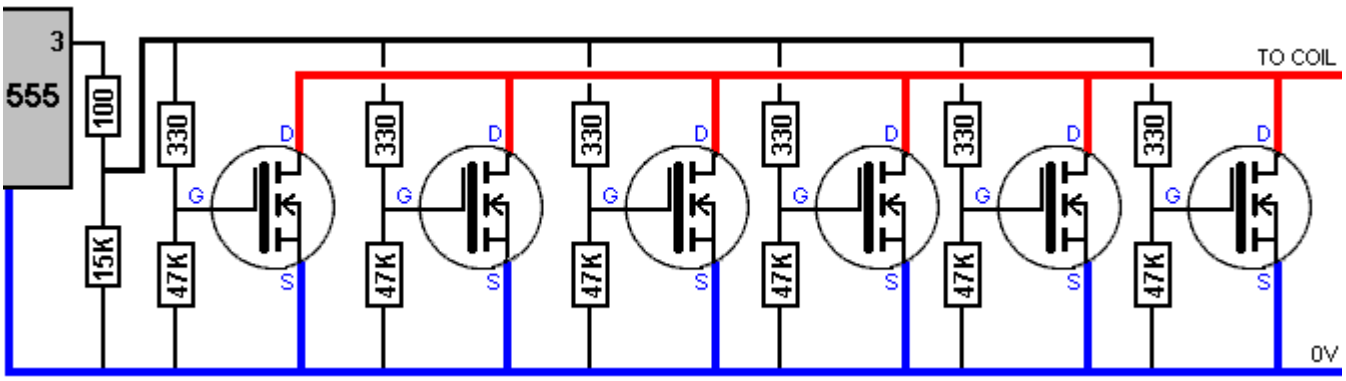
How the coil is wound is something to consider. The objective is to have a coil of about 1.5 ohm resistance and which has the maximum magnetic effect for the current passed through it. Copper wire has become very expensive and so it would be very costly to wind the coil with vast lengths of thick wire, not to mention the very large size and great weight which would be produced by doing that. The copper wire options in Europe are typically to work with half-kilogram reels of wire. The details of some of these are as follows:

Gauge (swg)	Wire length	Total ohms	Amps/strand	Strands	Total amps
14	17.5 m	0.09	9.3	None	-
16	27 m	0.22	5.9	None	-
18	48 m	0.71	3.7	None	-
20	85 m	2.23	1.8	2	3.6
22	140 m	6.07	1.2	4	4.8
24	225 m	15.81	0.73	11	8.0
26	340 m	35.70	0.45	24	11.0
28	500 m	77.50	0.29	52	15.0

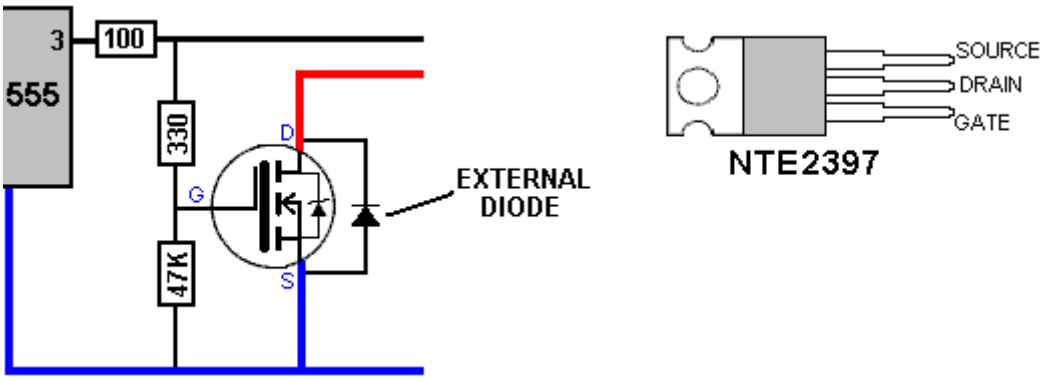
We can see from this that a 500 gram reel of 14 swg wire has a total resistance of just 0.09 ohms and so it would take sixteen reels (weighing 8 kilograms and costing a lot of money) to wind just a one-strand coil using that wire, producing a coil which could carry a current of 9.3 amps. As opposed to that, a single reel of 28 swg could provide 52 separate windings, which when connected in parallel, could carry 15 amps as well as costing and weighing far less. It would be tedious, but not impossible, to wind a 52-strand coil, so a more reasonable number of strands connected in parallel might be used. We are aiming at a DC resistance of about 1.45 ohms in any coil arrangement which we select.

The magnetic field produced by a single strand is generally less than the magnetic field produced by two strands carrying the same total current. So, if we were to pick 22 swg wire, then we could measure out four 133.5 metre lengths, join them at the start, and wind the four strands simultaneously, side-by-side to form a coil with a DC resistance of 1.45 ohms. It is important that the strands are exactly the same length so that they carry exactly the same current and no one strand gets overloaded with current due to it having a lower resistance than the other strands. It should be realised that as the maximum current which the wire can carry is 4.8 amps and the resistance is only 1.45 ohms, the maximum continuous DC voltage which can be sustained by the coil is only 7 volts, and so as a 36-volt battery is being used, we **must** adjust the frequency and duty cycle very carefully, especially since we are starting at very low frequencies. If the full battery voltage is applied continuously to the coil, then the coil will be destroyed.

Various members of the forum have suggested, built and tested different circuits for feeding a variable-frequency variable-duty-cycle drive signal to the output transistor. However, 'UFOpolitics' recommends a simple 555 timer circuit. If you are not familiar with electronic circuits, then read chapter 12 which explains them in some detail, including the 555 timer family of circuits. The point stressed by 'UFOpolitics' is that the output taken from pin 3 of the 555 chip passes first through a 100 ohm resistor and then, every transistor gets a separate feed via a two resistor voltage divider pair. The 47K Gate-to-Ground resistor is to ensure that the FET turns off properly. It may be possible to increase the value of these resistors but they should never be less than 47K.

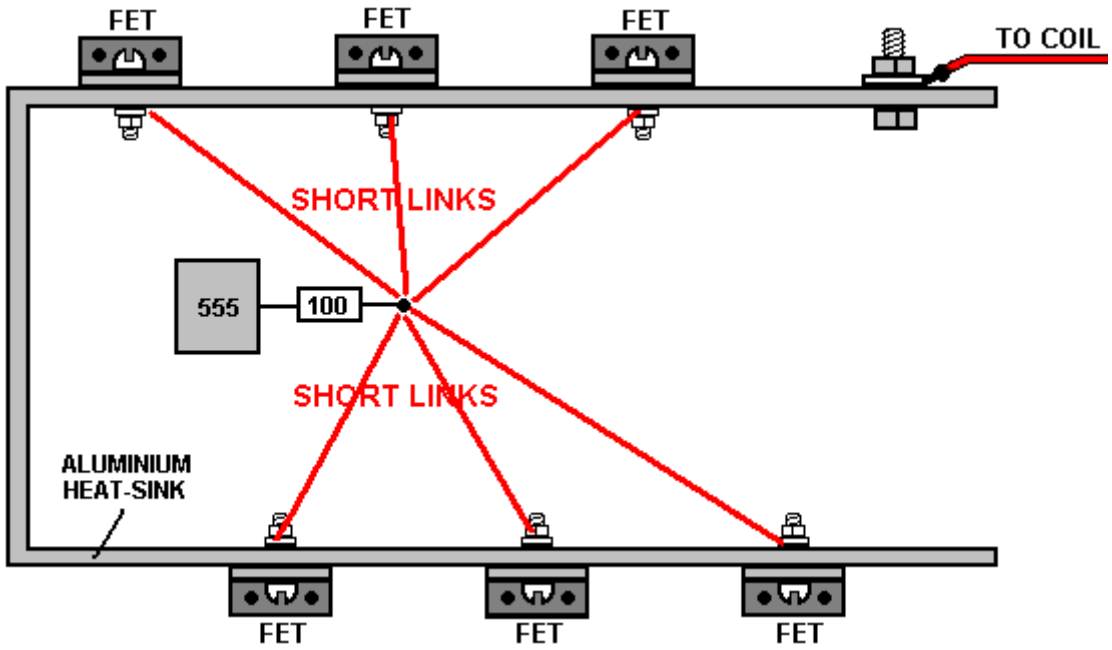


The thick lines in this diagram indicate heavy-duty wiring which can carry high currents without generating any real heat when doing so. It is also recommended that although the FET has an internal diode, an extra external high-speed diode (NTE576 or similar), be connected across each FET in order to boost the switching speed:



A FET has a gate capacitance of about 1 nF. The faster it can be charged / discharged the faster the FET will switch (and stay cool). What determines the speed of charge / discharge for the gate capacitance is the length of wire from driver to gate or gates is inductance (where one metre of wire produces 0.05µH). In addition to that, different lengths of Gate connection wire will create different switching delays and the different inductances can then initiate High Frequency oscillations with repetitive ON/OFF/ON/OFF switching actions. The result might be burned FETS and lack of cold electricity activities.

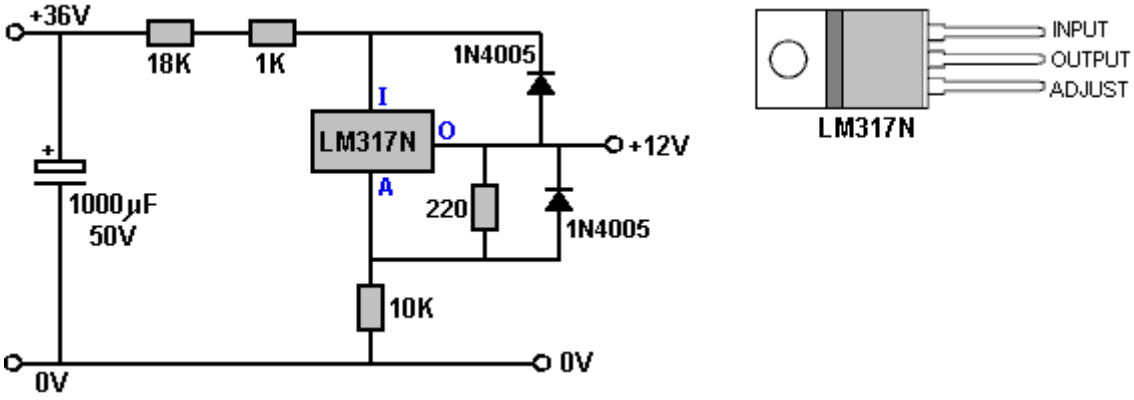
Another point made by 'UFOpolitics' is that the physical layout should have the connecting wires or tracks kept as short as possible and he suggests this layout:



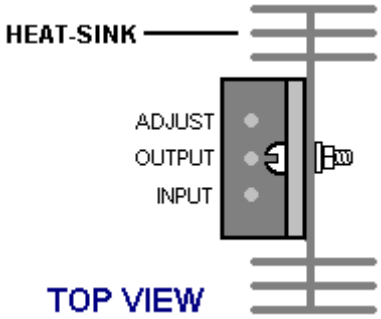
There are two things to note here. Firstly, the 100 ohm resistor coming from pin 3 of the 555 timer IC is positioned centrally between the six FET transistors mounted on the aluminium heat-sink, and that point is carried closer to each FET with a low-resistance conductor to give a good-quality link for the resistors feeding the Gate of each FET. Secondly, the heat-sink itself is also used to provide a low-resistance electrical connection to the coil which the FETs are driving. The connection to the heat-sink is via a nut and bolt clamping a solder tag firmly to a cleaned area of the heat-sink. Each FET is electrically connected to the heat sink through it's mounting tag which forms it's heat-sink connection as well as connecting to the Drain of the Transistor. However, if the aluminium heat sink is a black anodised type, then, apart from cleaning between each FET and the heat-sink contact area, it is worth running a thick wire also linking the central FET pins to the output wire connection point.

The transistors used in the prototype, and recommended for replications are the NTE2397. This is not a very common transistor in Europe at this time and so the popular IRF740 might perhaps be used as it appears to have all of the main characteristics of the NTE2397 transistor. 'UFOpolitics' suggests the 2SK2837 (500V, 20A, 80A pulsed), or the IRFP460 (500V, 0.27 Ohm, 20A and 80A pulsed).

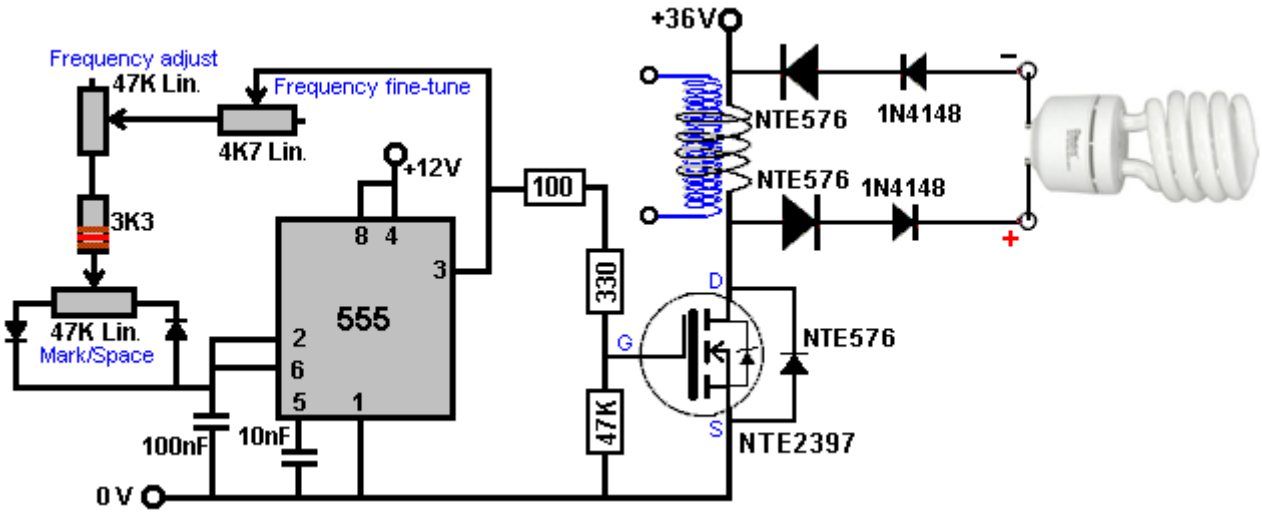
As the 555 timer has a maximum supply voltage of 15 volts, an LM317N voltage-stabiliser chip is used to create a 12-volt supply from the 36-volt battery (a 24V battery could be used):



The LM317N integrated circuit should be attached to a good heat sink as it is dropping off 24 of the 36 volts powering the circuit, and so, has to dissipate twice the power that the NE555 chip uses:



There are various pulsing circuits which have been used successfully with this system. 'UFOpolitics' considers the NE555 chip to be the most straightforward, so perhaps my suggestion for this arrangement might be a suitable choice:

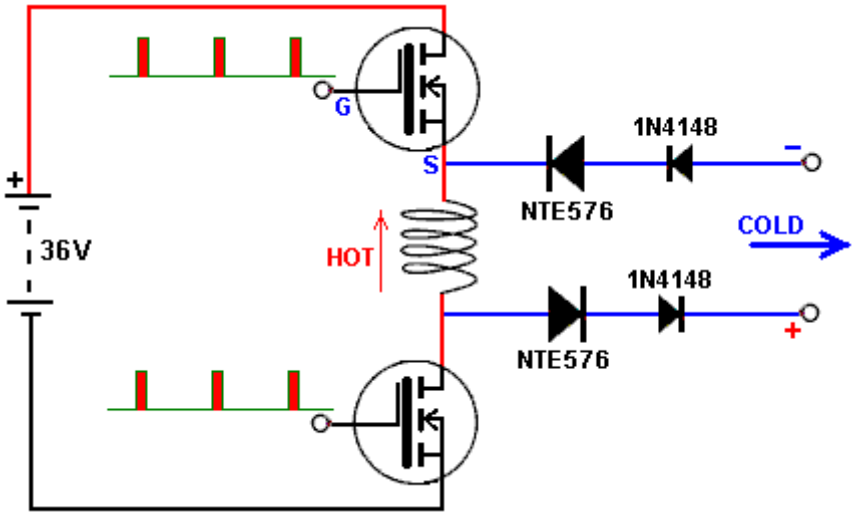


This gives fine control of the frequency and independent adjustment of the Mark/Space ratio or 'Duty Cycle' and it needs only three very cheap components other than the controls. If the expensive multi-turn high quality variable resistors are available, then the 4.7K 'fine-tune' variable resistor can be omitted as those variable resistors make the adjustments easier to control. The 'Lin.' In the diagram stands for 'Linear' which means that the resistance varies steadily at a constant rate as the shaft of the variable resistor is rotated.

In the ‘UFOpolitics’ circuit, it is important to turn the frequency down to it’s minimum value and set the Mark/Space ratio to 50%, before powering the circuit down. Otherwise it would be easy to power the circuit up with a much higher frequency than is advisable and so, causing damage to some of the circuit components.

There are ways to boost the performance over what has already been described. One way is to insert a stainless steel core inside the coil. Stainless steel is supposed to be non-magnetic but in practice, that is not always the case. However, ideally, this steel core is improved by altering it’s crystalline structure by heating it up and then quenching it by submerging it in cold water.

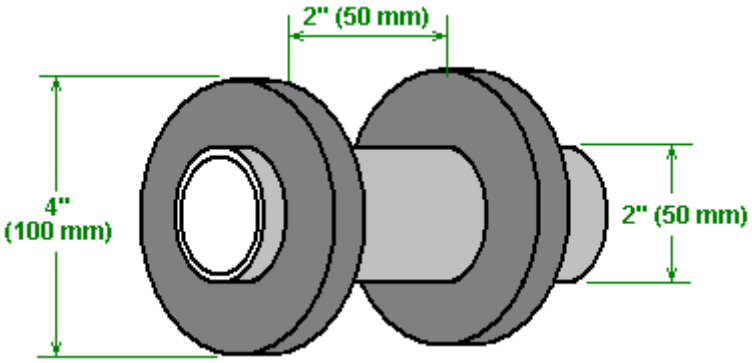
Another improvement is to isolate the coil better at switch-off through the use of a second transistor. Having a ‘switched-off’ transistor at each end of the coil certainly blocks the flow of hot electricity, but if Tom Bearden is correct, the resistance of the transistors in their OFF state will actually boost the flow of cold electricity as it reacts in the reverse way to how hot electricity reacts. The arrangement is like this:



While this looks like a very simple circuit to implement, that is not the case. The upper transistor is switched ON by the voltage difference between it’s Gate “G” and it‘s Source “S”. But, the voltage at it’s Source is not fixed but varies rapidly due to the changing current in the coil, and that does not help when solid and reliable switching of the upper transistor is needed. A P-channel FET could be used instead and that would have it’s Source connected to the fixed voltage of the Plus of the 36V battery. That would help the switching enormously, but there would still be timing issues between the two transistors switching ON and OFF at exactly the same time. Other circuits have been suggested for doing that type of switching, but in the early stages, ‘UFOpolitics’ recommends that things be kept as simple as possible, so using just one transistor is the best option.

Switching speed is an item of major importance, even to the extent that the reduction in the speed of switching caused by using more than one transistor in parallel has caused the suggestion to be made that it might actually be a better option to use just one FET since these high-performance FETs are capable of carrying the whole of the switching current, and it is mainly to lower the FET operating temperature that multiple FET use is suggested. Every extra FET used in parallel, slows the switching down. However, it should be realised that there is a somewhat greater risk of burning the FET out if just one is used.

The coil dimensions recommended are two-inch (50 mm) diameter and 2-inch length. The wound coil is likely to be about three-inches (75 mm) so making the flange diameter 4-inches (100 mm) is realistic:



The recommended material is fibreglass which has high heat-resisting properties as well as being easy to work, the personal choice of 'UFOpolitics' is Polyester Resin with Methyl Ethyl Kethol (MEK) Hardener. A suggested alternative is acrylic, which is not as heat resistant. Acrylic is excellent for high-frequency applications but this circuitry does not operate at high frequencies. Whatever spool material is chosen, it needs to be non-magnetic. When connected in the circuit, the start of the coil winding wire goes to the battery positive.

Here is another coil wound on acrylic tube and with all four diodes connected to the ends of the coil:

It should be understood that cold electricity provides almost unlimited power and it has uses which are not readily understood by many people.

'UFOpolitics' suggests that the hot electricity drive circuitry be tested initially using just a resistive load. If everything checks out correctly, then test with a lesser value resistor in series with the coil, and if that checks out satisfactorily, then testing cautiously with the coil on it's own.

Cold electricity can charge batteries rapidly and after a series of charge and discharge cycles, batteries become 'conditioned' to cold electricity and the experiences of Electrodyne Corp. staff show that large conditioned batteries which are fully discharged, can be recharged in under one minute. A member of the present forum has tried this with the 'UFOpolitics' circuit and he reports:

Yesterday a friend and I took 6 identical, old, 12V, 115Ah batteries and made two 36V banks. We set up bank "A" (the better three) to power the device to charge bank "B". Bank A was 37.00v at rest and Bank B was 34.94V.

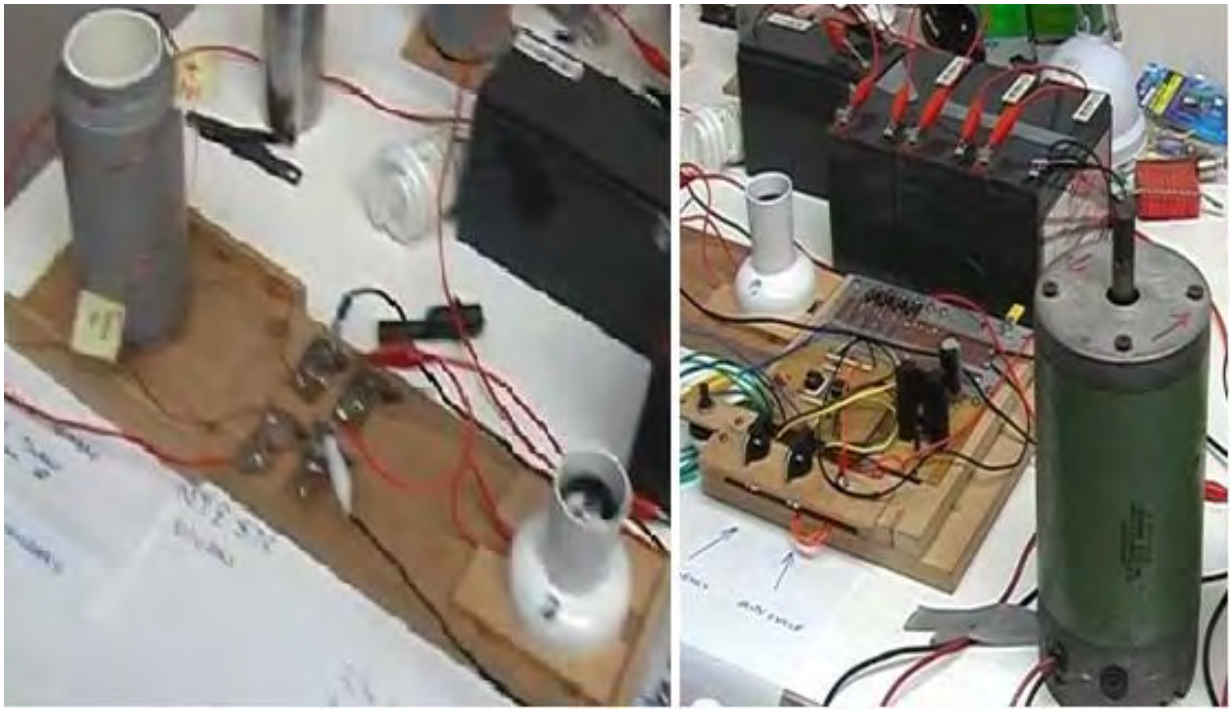
My lowest frequency was 133Hz (I need to change my cap and add another 100k pot with the one which is controlling the frequency) and the duty cycle was at 13%. We started at 2A draw on the Primary circuit.

As I raised the frequency, the batteries on charge jumped up to 38.4V then dropped evenly to 36.27V and started up again (at about 0.01V every 2 seconds). After two and a half hours, they were up to 39.94V. At this point we stopped the charging and let everything rest for 10 minutes. So far everything seems very normal for this kind of charging, except that the device appears to be very stable and powerful...pushing the batteries right on up continuously. The Primary battery voltage dropped initially to 36.20V and stayed there the whole time, then recovered to 36.98V during the 10 minute rest.

Then we switched battery banks A and B and charged the opposite way for about 20 minutes. We stopped and rested things again, swapped the banks back and started charging bank B again for another 20 minutes and stopped. After letting the batteries rest for a few hours in order to get truer readings, bank A was at 37.07V and bank B was at 38.32V. Both battery banks had gained power. These were not very good batteries, either. One of the bank B batteries was at 10.69V at the start. Another interesting note: The amp draw on the Primary dropped from 2A to 1.5A as the frequency was raised from 133Hz to about 550Hz.

This was with the very first use of cold electricity with these low-grade batteries and a major improvement can be expected after many additional charge/discharge cycles. This completely overcomes the factors which make a battery bank unsuitable for household power. If an entire battery bank can be recharged in just minutes, then it opens the way for serious household power using a battery bank.

Cold electricity can also run motors very powerfully. Forum member 'Netica' found that putting a capacitor across the motor terminals improved the running very substantially, giving impressive performance. His video of this is at http://www.youtube.com/watch?feature=player_detailpage&v=7uAYKhrPDPc and the motor, running off an air-core coil with no steel insert. His set-up looks like this:



It is also possible to submerge cold electricity circuits in water without causing any harm:

A video of this is located here: http://www.youtube.com/watch?v=W1KALMgFscg&feature=channel&list=UL including demonstrating the use of very powerful light bulbs. A general running demonstration is here: http://www.youtube.com/watch?v=yVzhKpEqUgc&feature=player_embedded .

Stanley Meyer's Electrical Particle Generator.
Stan, who is famous for his water-splitting and related automotive achievements, actually held about forty patents on a wide range of inventions. Here is one of his patents which circulates magnetic particles in a fluid, and while the fluid does move, none of the other components in the device move and a high level of constructional skills is not called for. This is a highly efficient generator of electricity.

This is a slightly re-worded excerpt from this Stan Meyer patent. Although it does not state it in the patent, Stan appears to make it understood that this system produces a significant power gain – something with Patent Offices find very difficult to accept.

Patent CA 1,213,671 4th February 1983 Inventor: Stanley A. Meyer

ELECTRICAL PARTICLE GENERATOR

ABSTRACT
An electrical particle generator comprising a non-magnetic pipe in a closed loop having a substantial amount of magnetised particles encapsulated inside it. A magnetic accelerator assembly is positioned on the pipe, which has an inductive primary winding and a low-voltage input to the winding. A secondary winding is positioned on the pipe opposite to the primary winding. Upon voltage being applied to the primary winding, the magnetised particles are passed through the magnetic accelerator assembly with increased velocity. These accelerated particles passing through the pipe, induce an electrical voltage/current potential as they pass through the secondary winding. The increased secondary voltage is utilised in an amplifier arrangement.

BACKGROUND AND PRIOR ART
The prior art teachings expound the fundamental principle tat a magnetic field passing through inductive windings will generate a voltage/current or enhance the voltage across it if the winding is a secondary winding.

It is also taught by the prior art, that a magnetic element in a primary inductive field will be attracted at one end of the coil and repelled at the other end. That is, a moving magnetic element will be accelerated in motion by the attraction and repulsion of the magnetic field of the primary inductive winding.

In the conventional step-up transfer, the voltage across the secondary is a function of the number of turns in the secondary relative to the number of turns in the primary winding. Other factors are the diameter of the wire and whether the core is air or a magnetic material.

SUMMARY OF THE INVENTION

The present invention utilises the basic principle of the particle accelerator and the principle of inducing a voltage in a secondary winding by passing a magnetic element through it.

The structure comprises a primary voltage inductive winding having a magnetic core, plus a low-voltage input. There is a secondary winding with a greater number of turns than the turns in the primary winding, plus an output for using the voltage induced in that winding.

The primary winding and core are positioned on one side of an endless, closed-loop, non-magnetic pipe. The secondary windings are positioned on the opposite side of the endless pipe. The pipe is filled with discrete magnetic particles, preferably of a gas, and each particle has a magnetic polarised charge placed on it.

Due to their magnetic polarisation charges, the particles will sustain some motion. As the particles approach the accelerator assembly, which is the primary coil, the magnetic field generated by the coil attracts the particles and accelerates them through the coil. As each particles passes through the coil, the repulsion end of the coil boosts the particle on it's way. This causes each particle to exit from the coil with an increased velocity.

As the magnetic particles pass through the secondary coil winding, they induce a voltage across the ends of that coil. Due to the larger number of turns, this induced voltage is much higher than the voltage across the primary coil.

The main objective of this invention is to provide an electrical generator which is capable of producing a voltage/current of much greater magnitude than has been possible previously. Another objective is to provide a generator which uses magnetic particles and a magnetic accelerator. Another object is to provide a generator which can control the amplitude of the output. Another objective is to provide a generator which can be used with DC, AC, pulsed or other configurations of waveforms. Another objective is to provide a generator which can be used in either a single-phase or a 3-phase electrical system. Another objective is to provide a generator for developing magnetised particles for use in an electrical particle generator. Another objective is to provide an electrical generator which uses readily available components to construct a simple embodiment of this invention.

BRIEF DESCRIPTION OF THE DRAWINGS

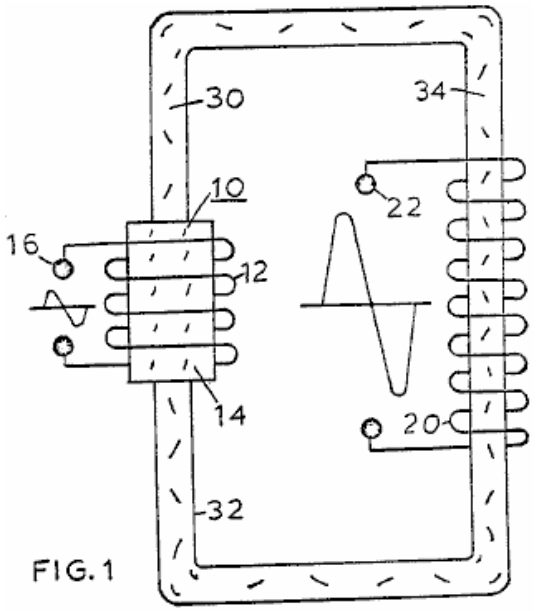


FIG. 1

Fig.1 is a simplified illustration of the principles of the invention, shown partially in cross-section and partially pictorially.

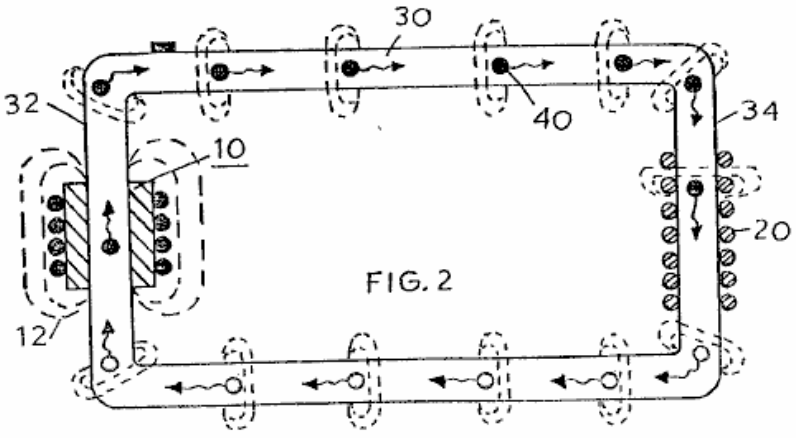


Fig.2 is an electrical schematic illustration of the embodiment shown in **Fig.1**.

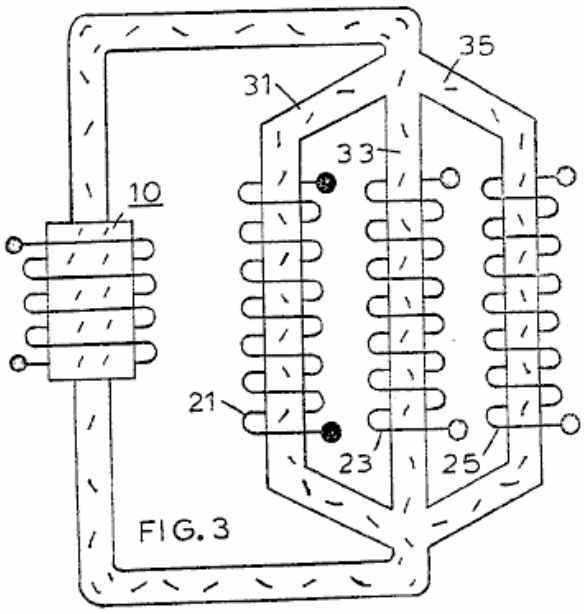


Fig.3 is an illustration similar to **Fig.2** but which is adaptable to 3-phase use.

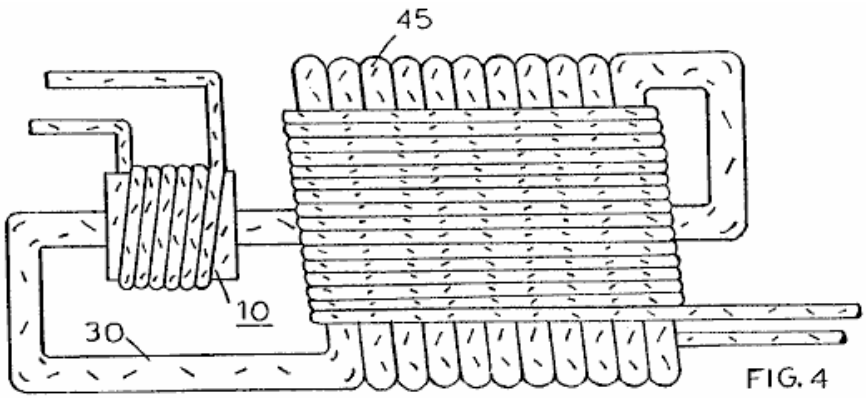


Fig.4 is a first alternative arrangement of a preferred implementation of the invention.

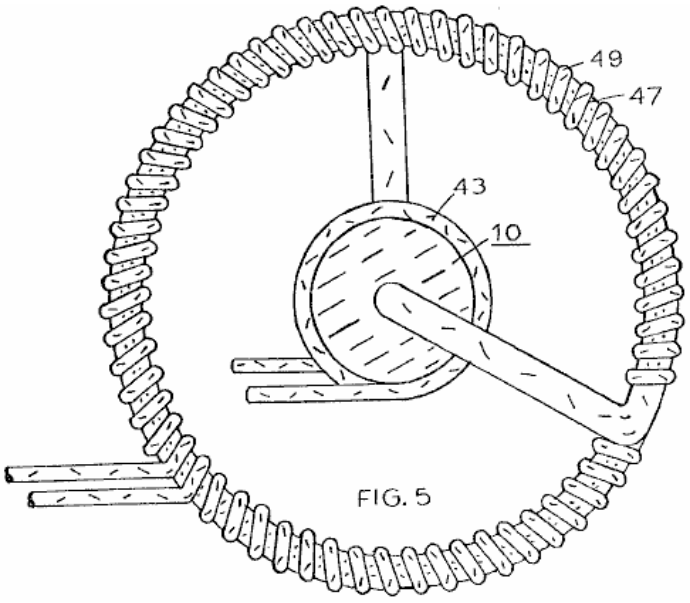


FIG. 5

Fig.5 is another alternative arrangement of a preferred embodiment of the invention.

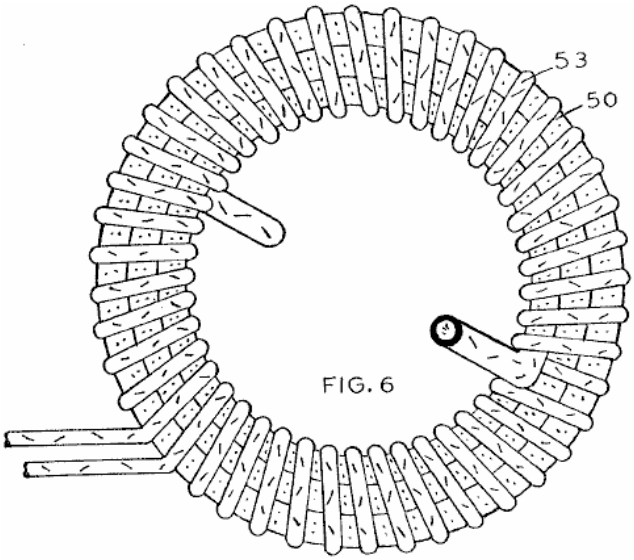


FIG. 6

Fig.6 is another alternative arrangement of a preferred embodiment of this invention.

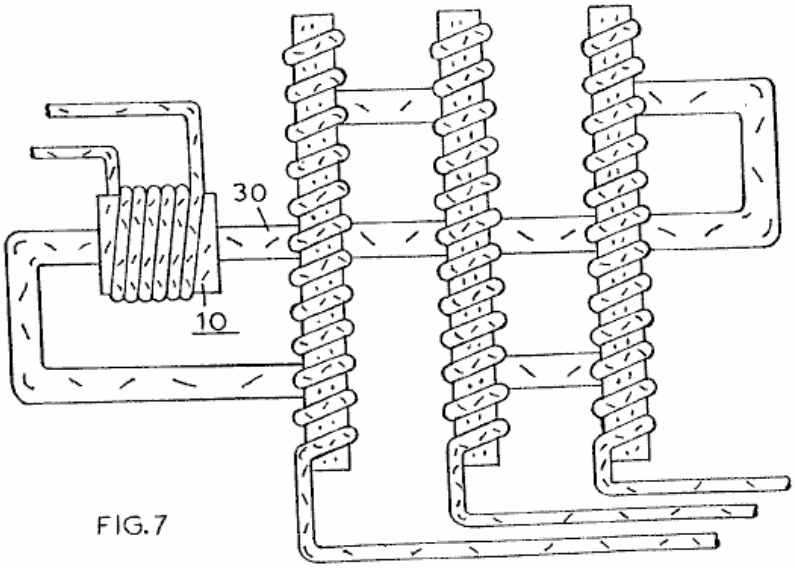


FIG. 7

Fig.7 is another alternative arrangement of a preferred embodiment of this invention.

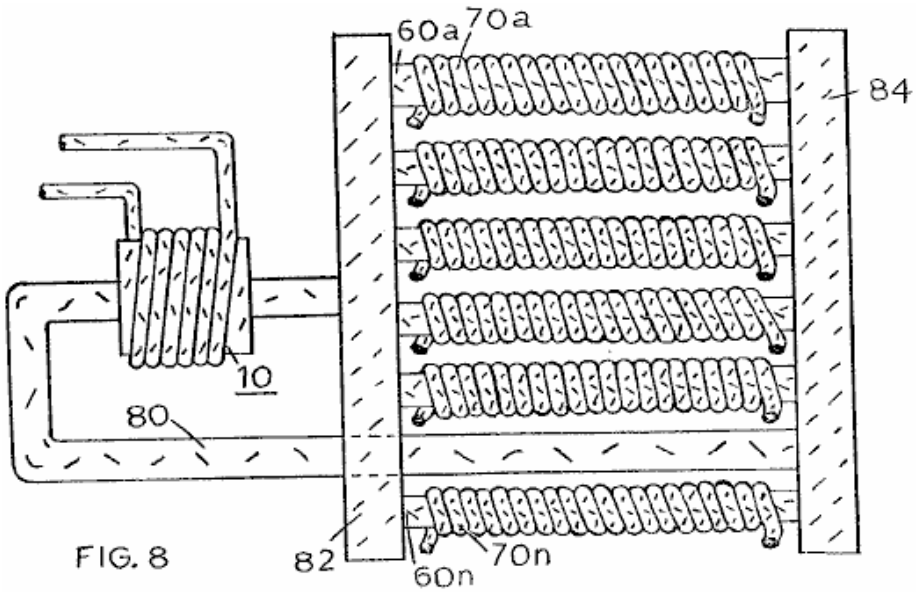


FIG. 8

Fig.8 is another alternative arrangement of a preferred embodiment of this invention.

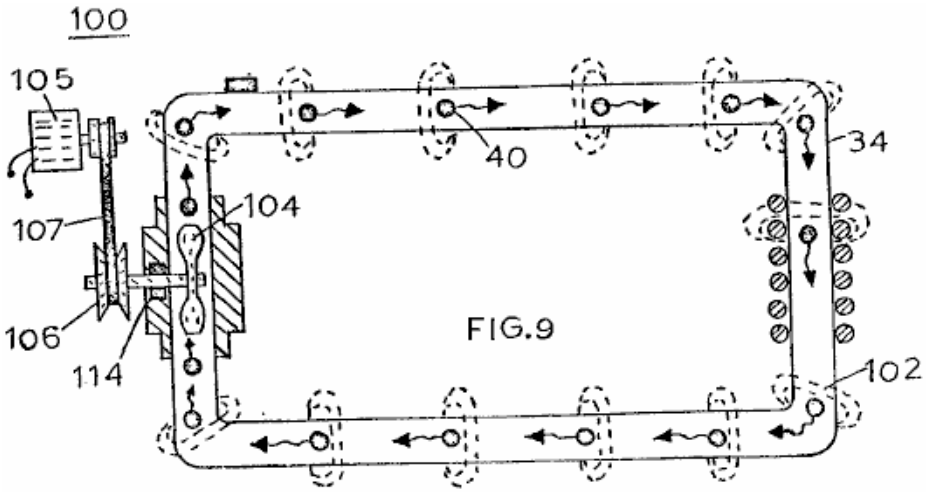


FIG.9

Fig.9 is an alternative arrangement for a magnetic drive particle accelerator assembly.

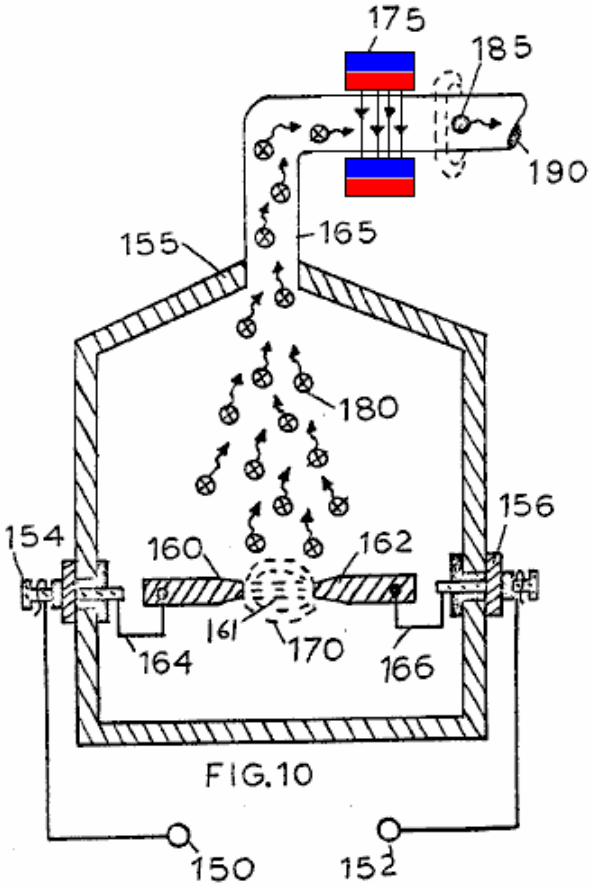


FIG.10

Fig.10 is an illustration of an alternative method of producing the magnetised particles used in this invention.

DETAILED DESCRIPTION

Fig.1 and **Fig.2** show the invention in it's most simplified schematic form:

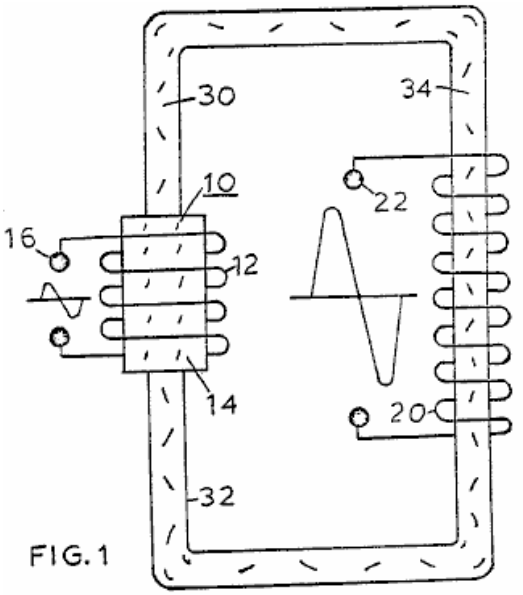


FIG.1

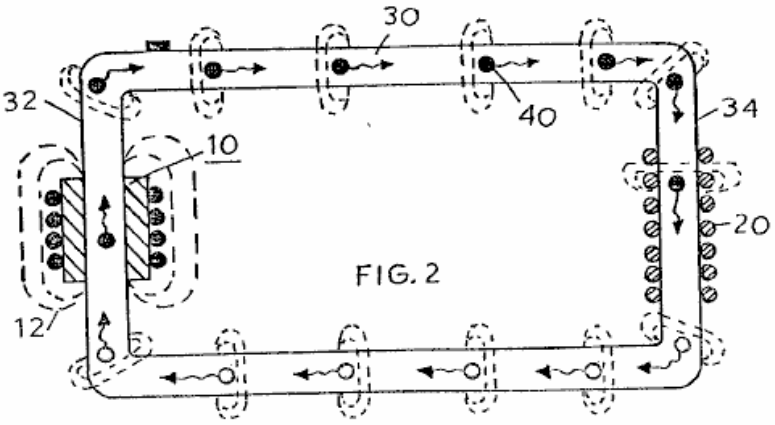


FIG. 2

It comprises a primary coil magnetic accelerator assembly **10**, a closed-loop non-magnetic pipe **30**, and a secondary winding **20**. The magnetic accelerator assembly is comprised of primary windings **12**, a magnetic core **14**, and voltage taps **16**. The primary windings are positioned around end **32** of the closed-loop pipe **30** which is made from non-magnetic tubing.

At the opposite end **34** of the closed-loop pipe **30**, are the secondary windings **20**. The end terminals **22** of the secondary winding **20**, allow the voltage generated in the winding to be used. Contained inside pipe **30**, there is a substantial number of magnetic particles **40** as shown in **Fig.2**. The particles **40** must be light enough to be freely mobile and so may be particles suspended in a fluid medium such as gas, liquid or light-weight movable solid particles. Of these options, the use of a gas is preferred. If solid particles are used as the transporting medium, then it may be desirable to remove all air from inside the pipe in order to reduce the resistance to the flowing particles. Each of the particles **40** is magnetised and the following description refers to one individual particle and not to the mass of particles as a whole.

The voltage applied to terminals **16** of primary winding **12**, is a low voltage, and it's magnitude may be used as an input signal control. By varying the input voltage, the accelerator will vary the speed of the circulating particles, which will, in turn, vary the magnitude of the voltage/current output of the secondary winding **20**. The output **22** of the secondary transformer winding **20**, is a high voltage/current output.

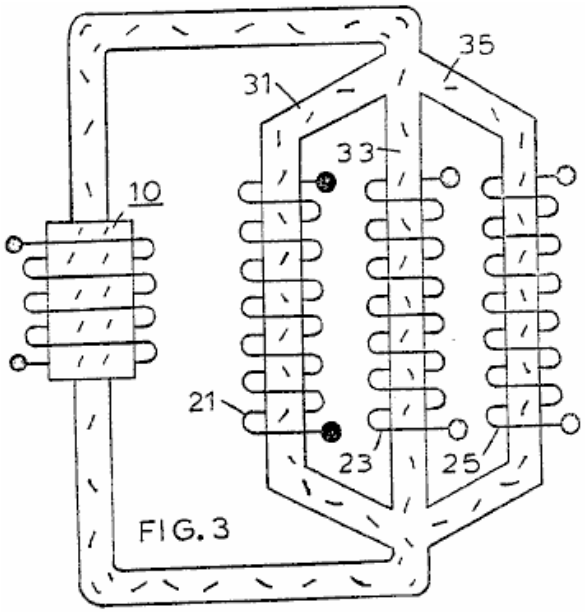


FIG. 3

It can be appreciated that the system shown in **Fig.1** and **Fig.2** where there is just one closed loop, provides a single-phase output in the secondary winding **20**. **Fig.3** shows a closed-loop arrangement with three parallel non-magnetic tubes **31**, **33** and **35**, each with it's own output winding **21**, **23** and **25**. Each of these three windings are a single-phase output, and as their three pipes share a common input junction and a common output junction, these three output windings provide a balanced 3-phase electrical system.

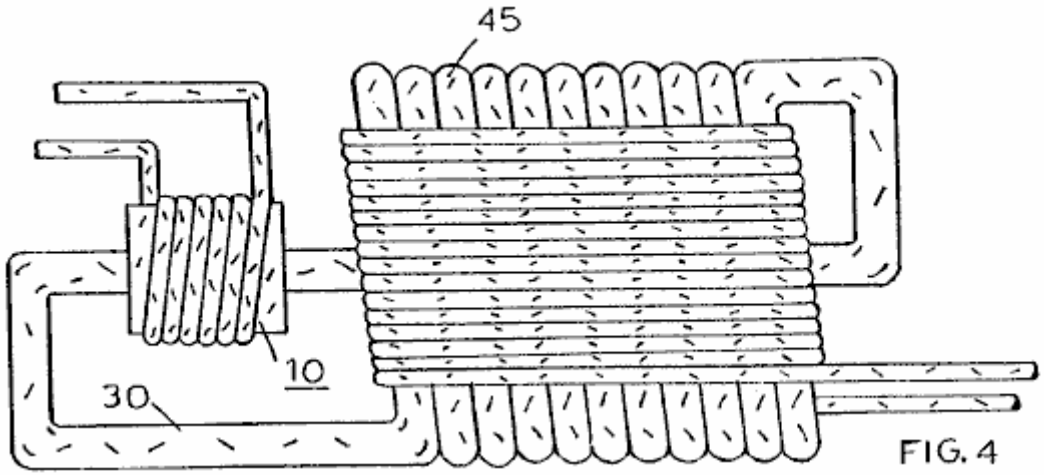


FIG. 4

Fig.4 shows an electrical power generator which operates exactly the same as those shown in **Fig.1** and **Fig.2**. Here, the arrangement is for use in an environment where there is a high moisture content. An insulating coating **45**, completely covers pipe **30** as well as all of the electrical windings. **Fig.4** also illustrates the fact that increasing the number of turns for any given wire diameter increases the voltage/current output of the device. In this physical configuration, both vertical and horizontal directions are used which allows a large-diameter pipe to be used with a substantial number of turns of heavy-gauge high-current wire.

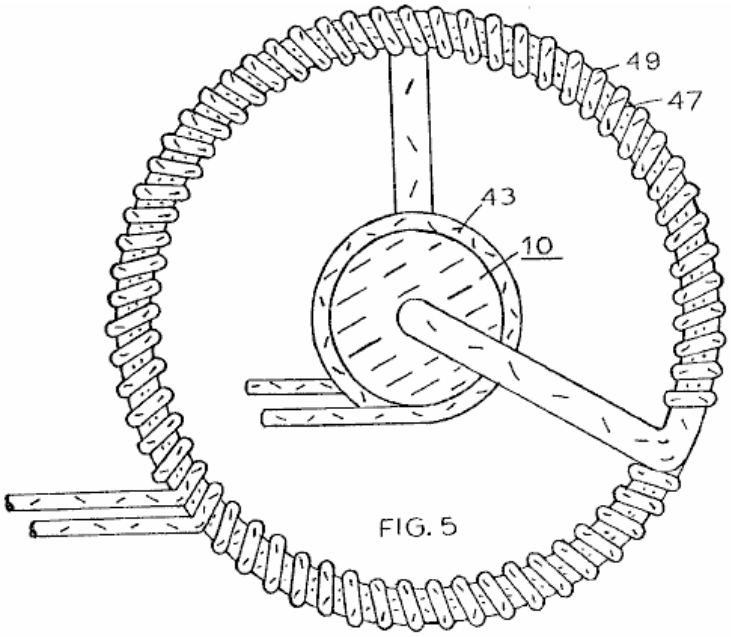


FIG. 5

Fig.5 shows a coil arrangement **49**, which uses the entire magnetic flux in the closed-loop tubing **47**. This is a co-axial arrangement with the primary winding **43** as a central core.

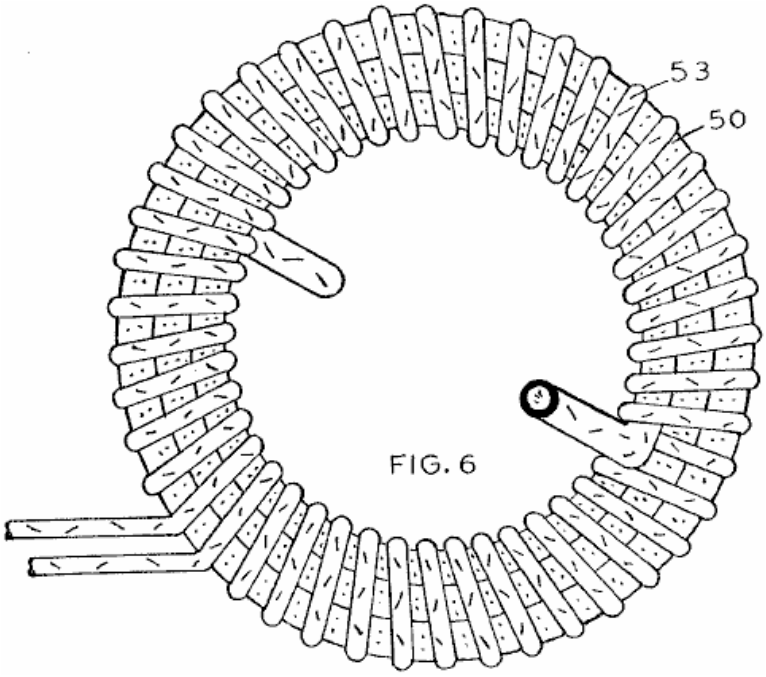


FIG. 6

Fig.6 illustrates a concentric spiral configuration of the tubing **50**, with the secondary windings **53** covering it completely.

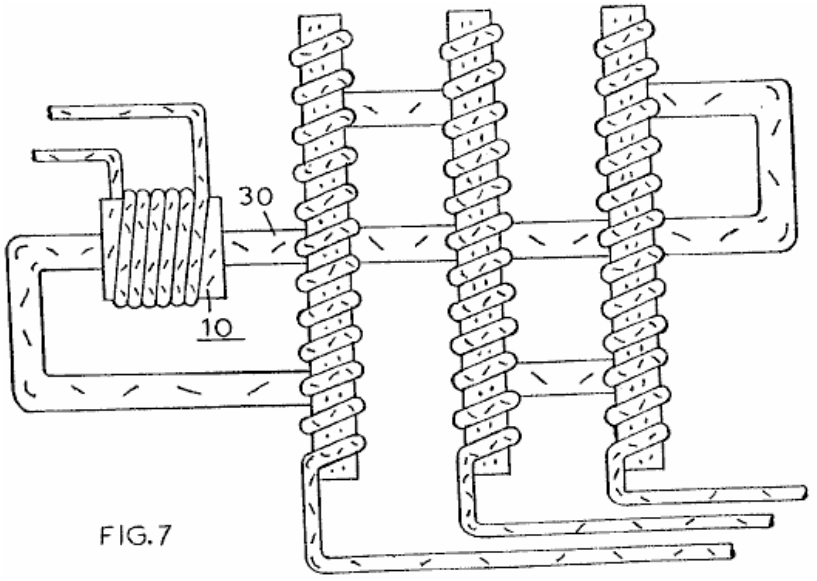


FIG. 7

Fig.7 shows an arrangement where the particle accelerator **10** is wound over the tubing **30** in much the same way as in **Fig.1** and **Fig.2**. However, in this arrangement, the tubing **30** is a continuous closed loop arranged in a series-parallel configuration where there are three secondary windings providing three separate outputs while the tubing **30** runs in series through those three windings.

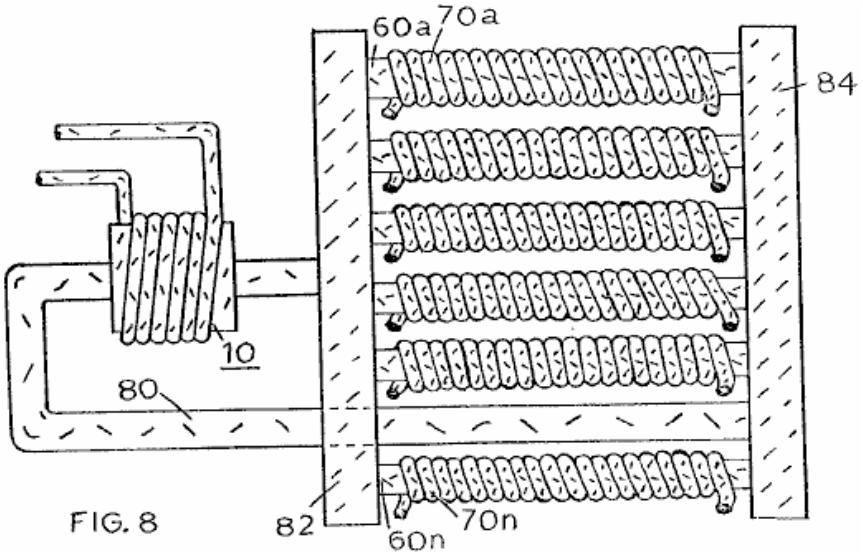


FIG. 8

Fig.8 shows a configuration which is the reverse of that shown in **Fig.7**. Here, there are several pick-up coils wound in series and unlike the earlier configurations, the tubing **80** is not continuous. In this arrangement, there is an input manifold **82**, and an output manifold **84**, and several separate tubes **60a, 60b, 60c, ….. 60n** interconnecting those two manifolds. Each of those separate tubes has it's own separate secondary coil **70a, 70b, 70c, ….. 70n** wound on it.

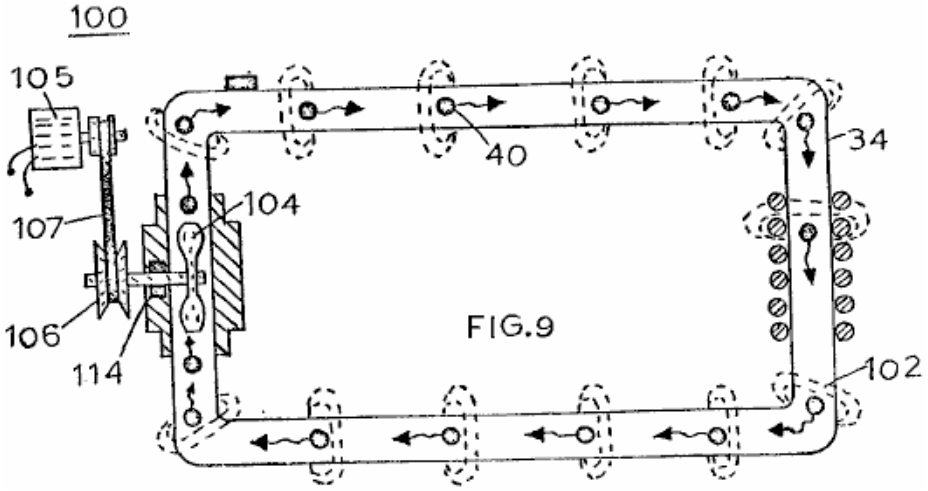


FIG.9

The magnetic particle accelerator **10**, can be different in design to that shown in **Fig.1**. **Fig.9** shows a mechanical particle accelerator **100**. In this arrangement, the magnetic particles **102** are permanently magnetised prior to being encapsulated in the non-magnetic pipe **110**. The particles 102 are accelerated by fan blade or pump **104** rotated by mechanical drive assembly **106**. The mechanical drive for assembly **106** may be a belt-drive pulley **112**, or similar device driven by an electric motor. A sealing bearing **114** keeps the particles **102** inside the pipe **110**.

It has been stated that the magnetic particles traversing the secondary coils, generate a voltage/current in them. It must be understood, however, that that the particles are actually traversing the magnetic field of those coils.

Also, the pipe **30** has been described as a non-magnetic pipe. There are certain non-magnetic pipes which would not work with this invention. Pipe **30** must be capable of passing magnetic lines of force.

A significant feature of each of the various embodiments already described, is the generation of the magnetic particles which are encapsulated within the tubing.

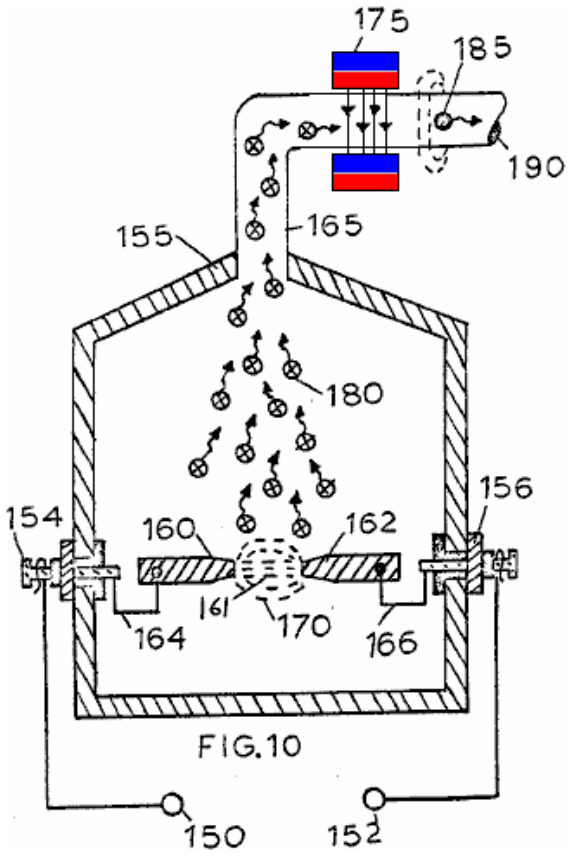


FIG.10

Fig.10 shows an apparatus for carrying out the process of vaporising material to produce suitable particles which are then magnetised by being subjected to a magnetic field. The chamber **155** is an evacuated chamber having

electrodes, made from magnetisable metal, **160** and **162**. A voltage is applied between terminals **150** and **152**, and this drives a current through terminals **154** and **156**, to spark-gap electrodes **160** and **162**, generating an arc which vaporises the tip material of the electrodes, producing particles **180**. These particles rise and enter tube **190**, passing through a magnetic field generator **175**. This gives each particle a magnetic charge and they continue on their way as magnetically-charged particles **185**, passing through port **190** to reach the electrical particle generator described above.

In the simplified embodiment shown in **Fig.1** and **Fig.2**, as well as the other preferred embodiments mentioned, it was indicated that a low voltage was applied to the particle accelerator **10**. Upon acceleration, a high voltage/current would be induced in the secondary pick-up coil **20**. A most significant advantage of the present invention is that the voltage amplification is not related to the shape of the waveform of the input voltage. Specifically, if the input is DC a DC voltage will be output. An AC input will produce an AC output. A pulsed voltage input will produce a pulsed voltage output and an input voltage of any other configuration will produce an output having that same configuration.

The Work of Russ Gries.

Russ Gries has produced a video presentation and analysis of the above Stan Meyer patent. http://www.youtube.com/watch?v=OnAmTmxBpAQ .

The very experienced Alex Petty is joining with Russ in working on replicating Stan's system and Alex's web site is at www.alexpetty.com. A discussion forum linked to this is at http://open-source-energy.org/forum/ and there is information at http://www.overunity.com/index.php?topic=5805.285 and high-resolution pictures can also be seen in Russ' video at http://www.youtube.com/watch?v=JOarpi6sDD4. Russ' own website is at http://rwgresearch.com/ and an additional video of the most recent developmental work being undertaken at: http://www.youtube.com/watch?v=adzVQRsS1KY&feature=youtu.be.

There are various important things which are commented on and Russ is to be commended for drawing attention to them. For the moment, please forget about HHO as that is a separate issue. As far as I can see, the patent does not claim that the device is COP>1 but instead that the device is a power transformer which potentially has a greater power output than conventional transformers since there is no Lenz Law reverse magnetic path from the output coil winding to affect the input power.

Having said that, Stan in his video points out ways to boost the power of the device, namely:

1. Increase the strength of the magnetic particles

2. Increase the speed of the magnetic particles

3. Lower the distance between the magnetic particles and the output winding.

The magnetic particles can be produced in various ways, but the most effective appears to be by filling the arcing chamber with argon gas and using iron, nickel or cobalt electrodes. The reason for this is that the electric arc does not only generate minute particles of the electrode material, but it also interacts with the argon, stripping off electrons and causing some of the metal particles to combine with the modified argon gas molecules to form a magnetic gas. That gas will always remain a magnetic gas due to the atomic bonding as it is not just minute particles of metal physically suspended in a gas due to their tiny size.

You will recall from chapter 1, that the very successful ShenHe Wang magnet motor/generator has a magnetic liquid as a key component. Here, Stan is producing a much lighter magnetic gas and the advantage of that lightness is that it can be boosted to very high speeds without any danger. The larger the number of modified argon molecules, the greater the magnetic effect when they pass through a coil of wire. The argon gas can be passed through the arc chamber over and over again so that a very high percentage of the gas is magnetic. Alternatively, if you are sophisticated in the design of the particle generator, you can arrange for the molecules which have become magnetic, to be pulled off into storage by a magnetic field.

Stan talks about pumping the magnetic gas through whatever pipe loop arrangement you decide to use, by a pump, but he promptly moves on to using a magnetic coil to boost the gas forward as the coil has no moving parts and so, no mechanical wear. This is only one reason. The main reason is that with magnetic acceleration, the gas speed can become very high indeed and in his video he talks about the speed of light. However, I personally do not believe that anything remotely like a speed that great could be achieved inside a pipe loop of small diameter. Nevertheless, speeds well in excess of what a mechanical pump can achieve are likely to be produced by magnetic acceleration.

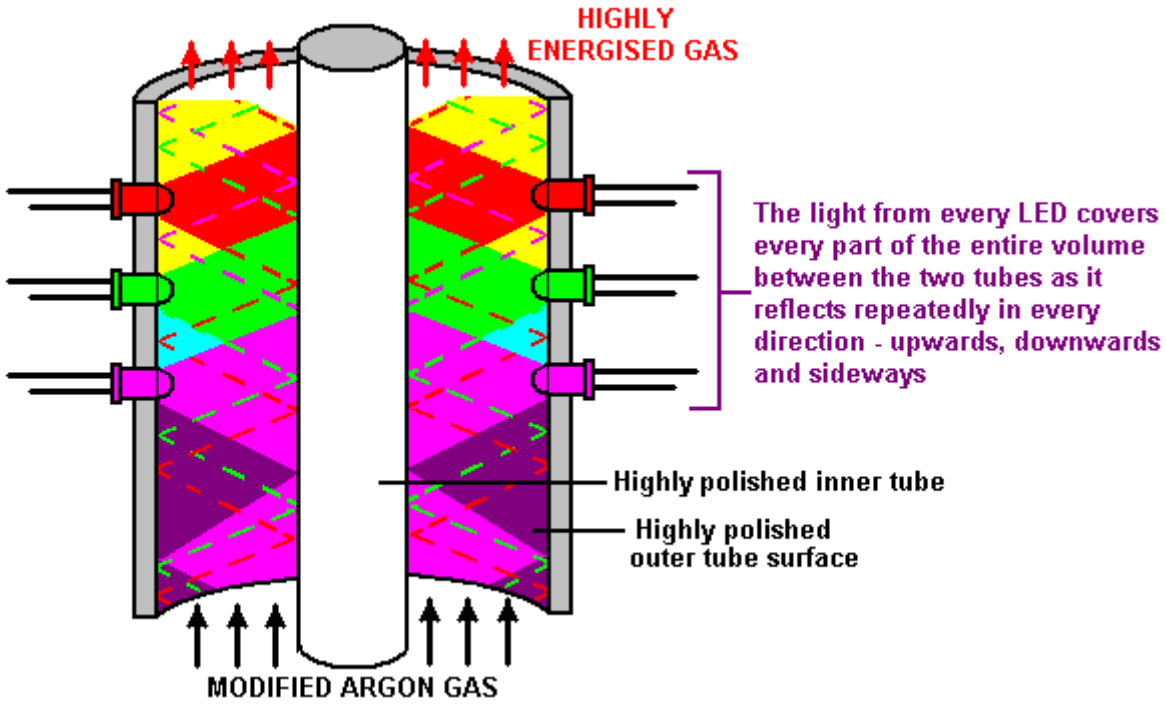
Russ, in his discussion, points out that on most of Stan's surviving prototypes, the coil which is used for the acceleration is constructed using several apparently separate coils, and he speculates that each coil section is powered sequentially, causing a rippling magnetic field. While that is definitely possible, I don't see that a style of coil powering would have any advantage as opposed to powering all of the coils continuously. However, if sequential powering is believed to be an advantage, then the 'Divide-by-N' circuitry of chapter 12 can be used to provide the sequential powering or any more complex sequence.

Stan then points out that the output voltage can be increased by increasing the number of turns on the output coil and/or having additional output coils. This is easily understood conventional electrics. But, he then goes on to point out that the output will also be increased if the electrons of the modified argon molecules are raised to a high orbital level. This places the electromagnetic electrons (as described in chapter 11) closer to the output coils and presumably also allows the gas to be accelerated to a greater speed by the driving magnetic field.

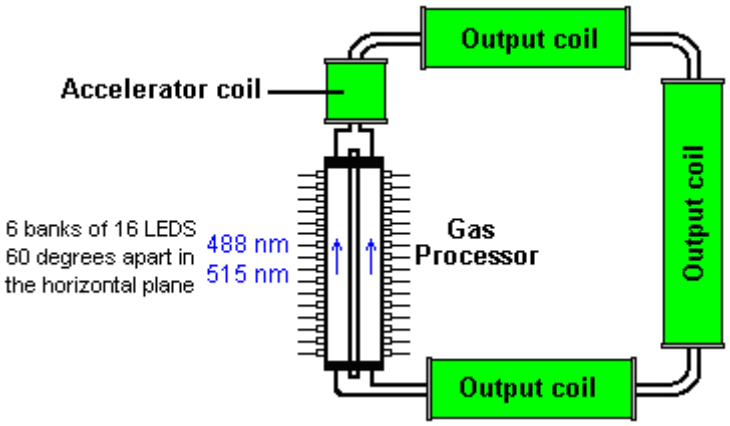
This power boosting of the gas is achieved using Stan's "Gas Processor" described in chapter 10. The Gas Processor pumps electromagnetic energy into the gas through the use of banks of Light-Emitting Diodes which produce light of the correct wavelength to add energy to that particular gas.

If you check on the internet for the wavelength of argon, you find conflicting information, with some sites saying that the wavelength is 1090 nanometres ("nm") and most others saying both 488 nm and 514.5 nm. Most LEDs produce a band of frequencies, so it would be a case of picking LEDs whose band of frequencies include the wanted wavelength.

The Gas Processor itself, consists of a central tube which is polished to a mirror finish on the outside, surrounded by a larger tube which is highly polished on the inside. The LED light is then bounced between these polished surfaces until it is absorbed by the gas which is passed through the gap between the two tubes. This is not easy to illustrate, but it might be shown like this:



In Stan's design, he uses six columns of sixteen LEDs, with each column of LEDs spaced out evenly around the outer tube. So, to boost the Magnetic Particle Generator to greater power levels, a Gas Processor is placed in the loop of tubing:



The Gas Processor normally has a coil mounted at each end and it may be convenient to use coils in those positions as accelerator coils. It may also be an advantage to apply a pulsed high-voltage between the inner and outer tubes of the Gas Processor. As it stands, this looks as if it has a high possibility of being a COP>1 electrical device.

The E-Stress Power Generation System

On the seventh of September 2013 a man with the ID of 'harisingh' published the following information. I have tried to contact him to ask his permission to publish his work, but without any success. I have no information on it other than what is shown here. What he says is:

The E-Stress Power Generating System is a most intriguing over-unity power generating project. It is extremely versatile while being relatively easy to construct and it is readily upgradable.

The Golden Key or Golden Mean is the principle which allows this device to draw power directly from the e-field's relative vacuum energy density using induction-less coils. What makes this design so unique can be seen from the following illustrations. What makes this dual induction-less coil so special is that it nullifies the effects of electrostatic induction using only a fraction of the power required to sustain the charge-displacement which creates the affecting E-field. In conventional electronic circuits, coils and capacitors are generally kept away from each other, but in this circuit, their interaction is the key to success !

The E-Stress Amplifier consists of three cylindrical capacitors and two induction-less coils with external circuitry designed to start-up and maintain the entire system and load. The interior and exterior capacitors, **CDI** and **CDE**, are charged up and kept charged by a voltage source **Vc** which is in the range of 50 to 90 volts DC. The charge on these capacitors remains for a long time, governed by the resistance of the dielectric, (provided that there are no accidental short circuits) so, the energy required to maintain the charge on these capacitors is minimal.

The third capacitor **Cr**, is sandwiched between the interior and exterior capacitors and is independent of **Vc**. When the other two capacitors, **CDI** and **CDE**, become charged, capacitor **Cr** becomes charged as well but at a slightly lower voltage due to dielectric voltage drop. This charging effect is a result of electrostatic induction. Separating the concentric capacitors are two special induction-less coils. When current passes through these coils, the electrostatic induction forces are temporarily neutralised, allowing the charged capacitor **Cr** to discharge and oscillate with an exterior power coil or transformer as shown in **Fig.7**. The induction-less coils **Lo** are pulsed with DC at the same rate as the natural resonant frequency of the tank circuit (**Cr-Lr**). Keeping the pulsing rate within 10% of the tank circuit's resonant frequency will maintain maximum output power. Because of the flexibility in the construction parameters, determining the resonant frequency and power capabilities will be difficult without advanced analysing equipment. So to overcome these difficulties, a variable-frequency oscillator can be used as it allows ready determination of the appropriate frequency range needed. A steady constant load will make this determination easier as well. When the maximum output power is reached, you can measure the resistance of the oscillator's variable resistor to aid in determining, and finally setting, oscillation at the natural resonant frequency of the tank circuit. An additional 1K ohm variable resistor can be connected in series with the 100K variable resistor shown, to allow easy "fine" tuning.

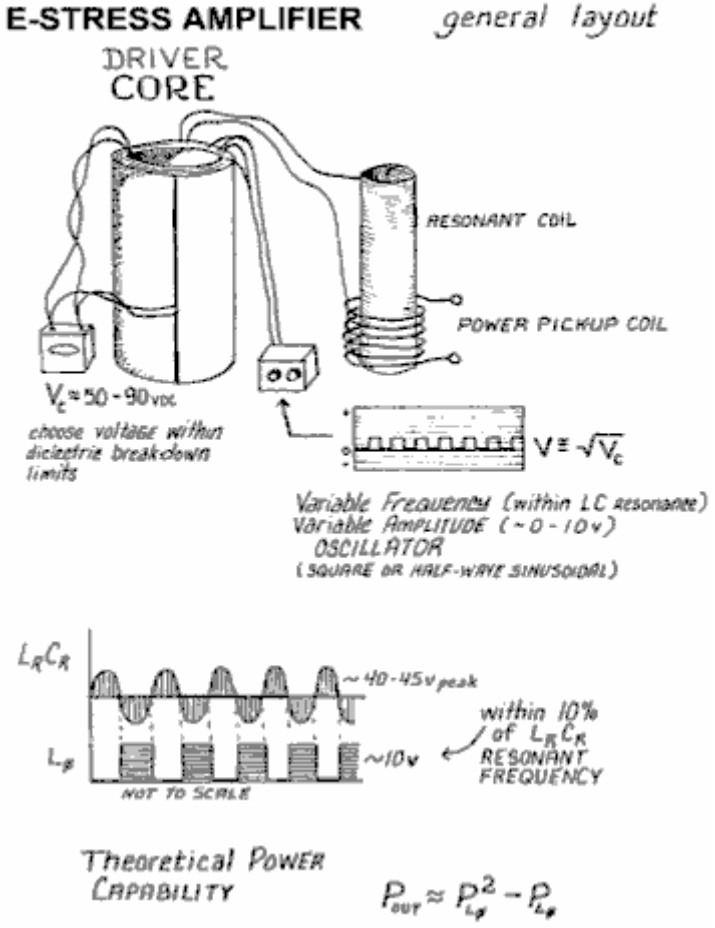


Fig.1 shows the general layout of the entire E-Stress Power System. This diagram illustrates the driver core and the variable frequency oscillator which determines the tank circuit resonance. **Fig.1** also shows the DC power source of around 50 to 90 volts, which initially charges the "core" capacitor **CDI** / **CDE**. Also shown in this diagram is the resonant coil and power output or "pickup" coil.

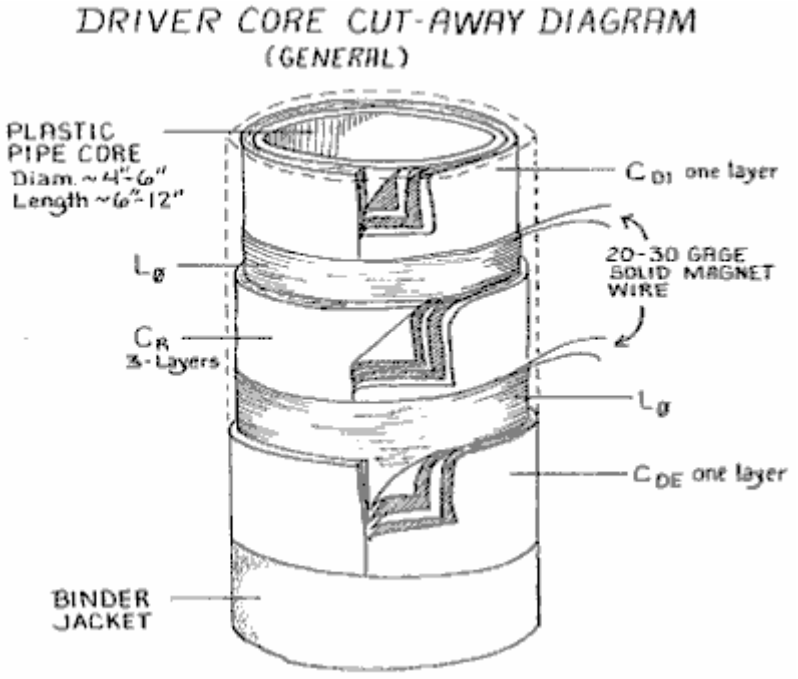


Fig.2 shows the "driver core" in a cutaway, perspective diagram. The PVC plastic "core" material can be a 6 inch (150 mm) diameter foot long (300 mm) piece of PVC pipe in this scaled-down model system. In this diagram, please note the inner and outer one-layer capacitor capacitors **CDE** and **CDI**. Also note that the middle capacitor

(**Cr**) is a three-layer capacitor made from heavy aluminium foil or stainless steel. The induction-less coils can be seen in two places on either side of capacitor **CR**. The induction-less coils are made from solid gauge #20 to #30 magnet wire (0.812 mm to 0.255 mm diameter enamelled copper wire).

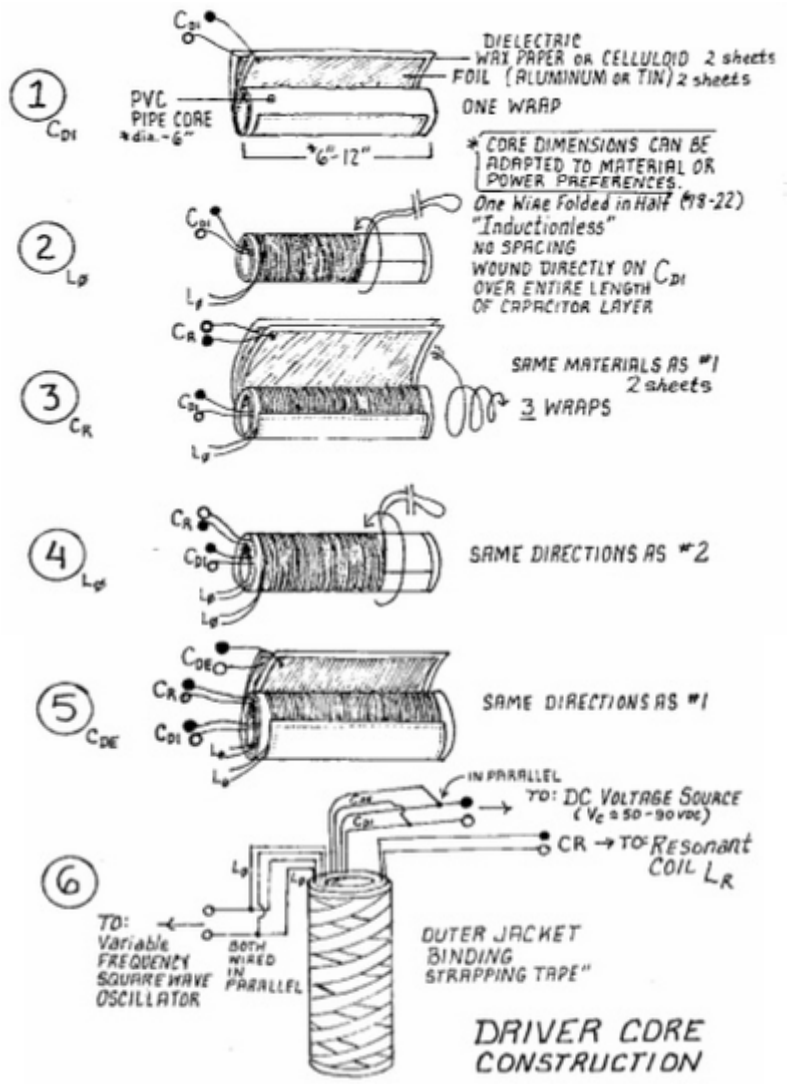


Fig.3 above, illustrates how the centre driver "core" is constructed. There are six basic steps to constructing the driver "core" assembly. First begin with the PVC pipe as shown and begin wrapping the first capacitor **CDI**. Step two shows the first induction-less coil **L1**. Note that for the induction-less coil, the wire is "folded" back and the two conductors are wound together as shown. This coil is single-wound with #18 to #22 gauge wire (1.024 mm to 0.644 mm diameter). Step three shows the centre capacitor **Cr**. This capacitor is created in the same manner as the previous capacitor except that it has three layers of turns. Step four illustrates the second induction-less coil which is wound in the same direction as the first coil. This second coil **Lo** is single-layer. Step five is the final capacitor **CDE**, which consists of a single wrap, and is wrapped in the same direction as the two previous capacitors. The final step, six, is to wrap the entire assembly with strapping tape to form an outer jacket when finished.

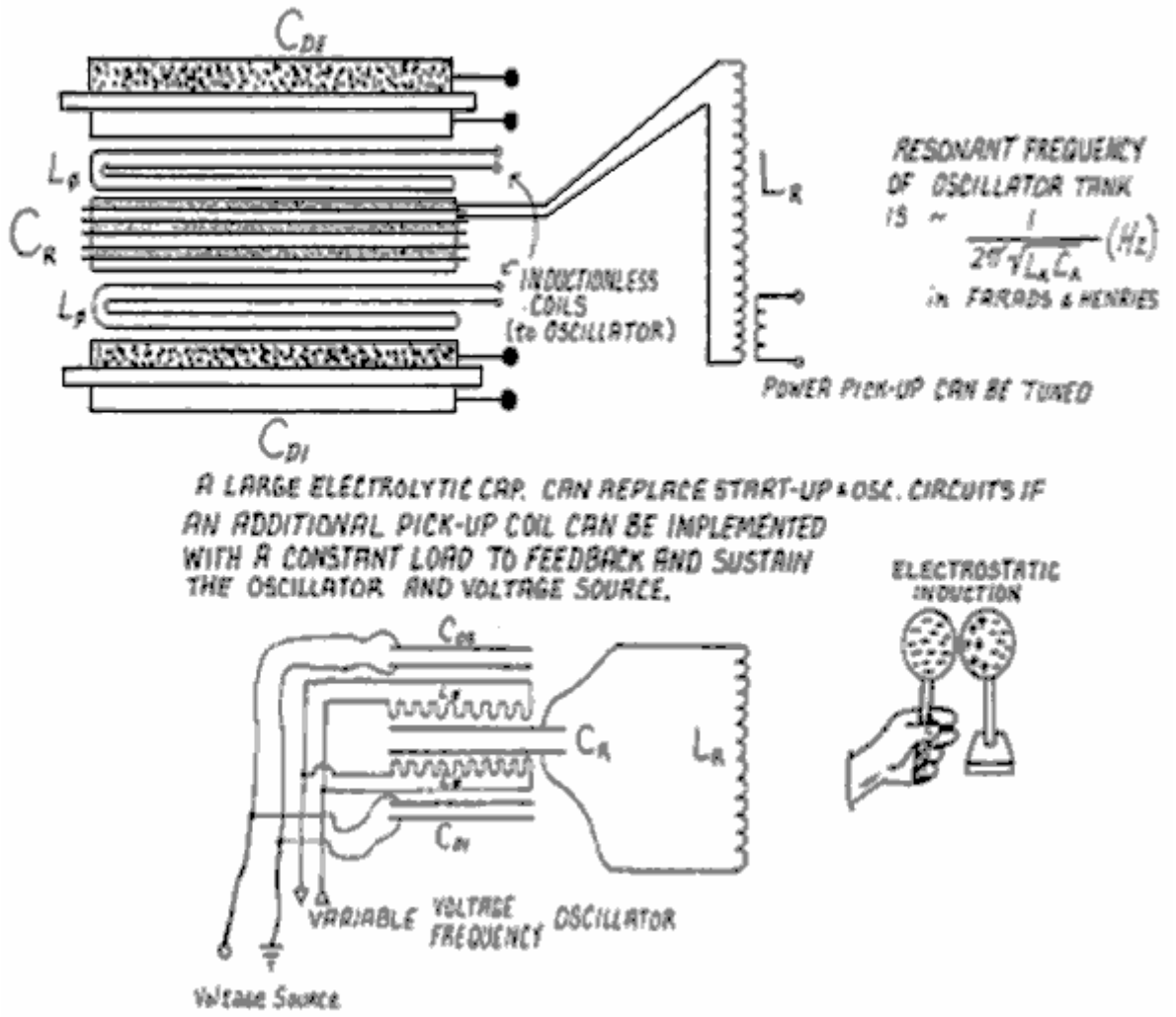


Fig.4 above, is an electrical diagram of the main capacitor "core" and the coil hookups of the E-Stress Power System. Note the parallel connections of the inner and outer capacitors **CEI** / **CDE**, which connect to the 90 volt DC power source. Also note the parallel connections of the induction-less coils which connect to the variable frequency oscillator circuit. The centre capacitor **Cr** is shown connected to the resonant power coil **Lr**.

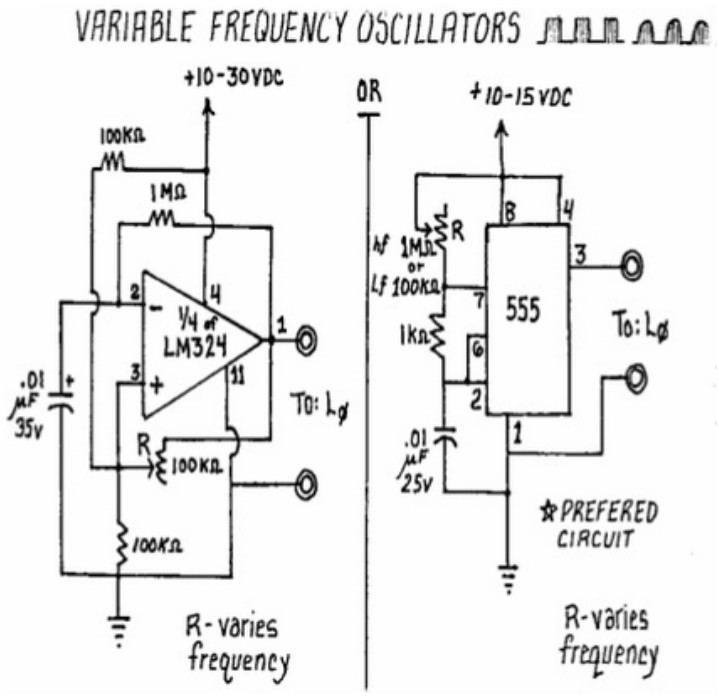


Fig.5 above, displays the two different types of variable frequency oscillators which are used to drive the induction-less coils. The first oscillator displayed consists of an LM324 Op-amp configured to produce feedback and thus oscillate. The second example oscillator consists of an LM555 timer IC. Either example oscillator can be used to drive the induction-less coils.

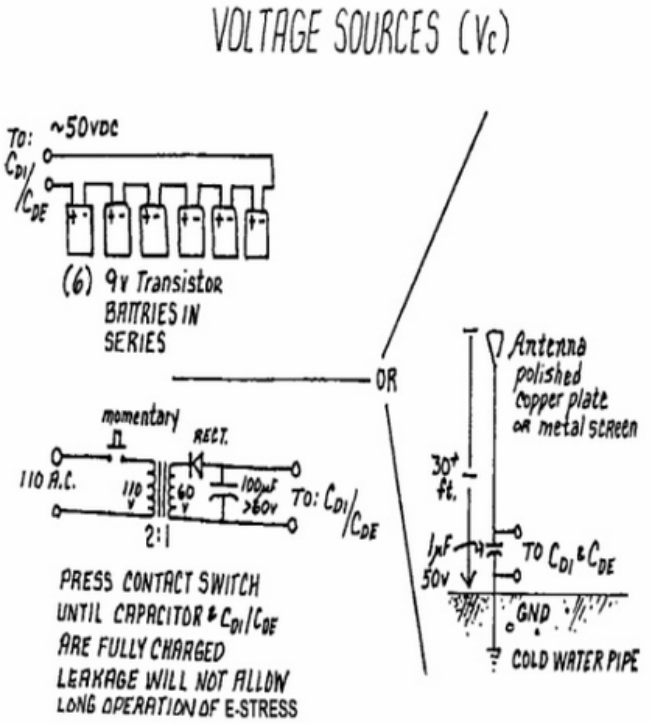


Fig.6 above, shows the DC voltage source **Vc** which is applied to the **CDE** / **CDI** capacitors which are connected in parallel to form the electrostatic field. The DC voltage source could be one of the three types of voltages sources as shown. A battery could be used, which consists of six nine volt batteries. You could also fabricate an AC to DC power source or you could elect to create your own antenna DC source as shown. The battery method affords a quick means to test the circuitry and is safe as well.

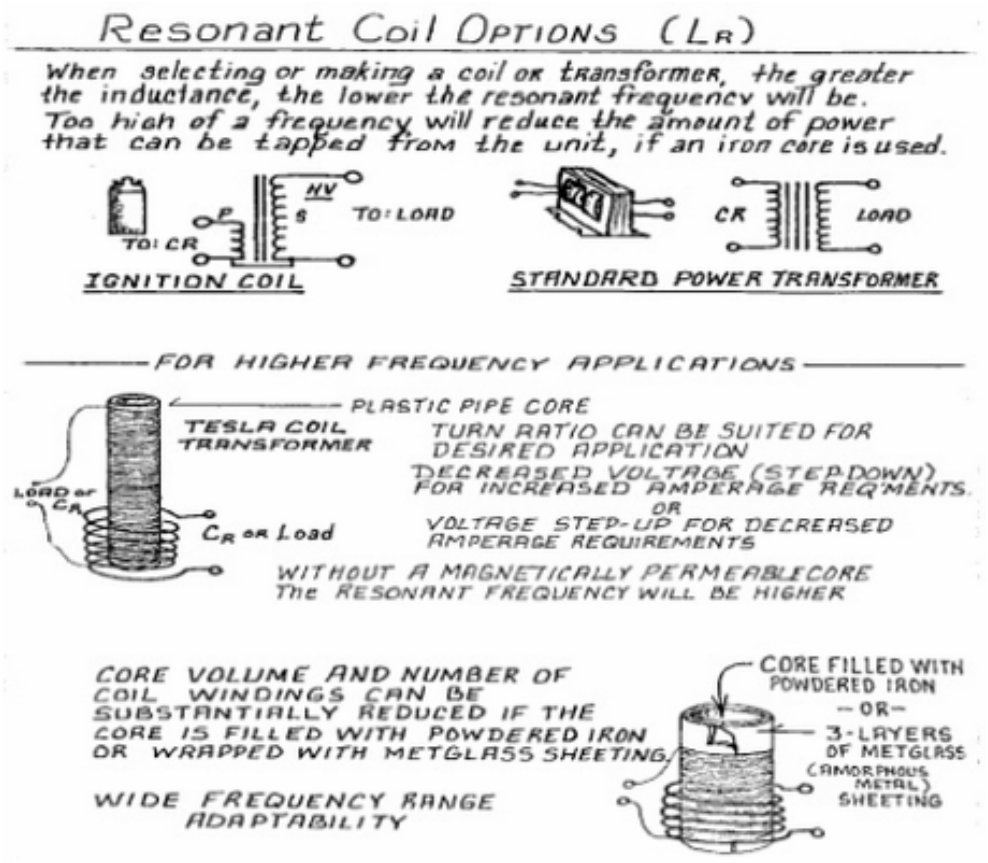


Fig.7 above, illustrates the options for the two different types of resonant coil **Lr**. Basically, there are two coil options: the standard iron core power transformer and the high frequency Tesla type coil such as the ignition type coil. You will need to determine the type of output you wish to have in the overall design. For conventional designs you would most likely construct the standard transformer shown at the bottom of the diagram, which consists of a core filled with powdered iron or Metglass.

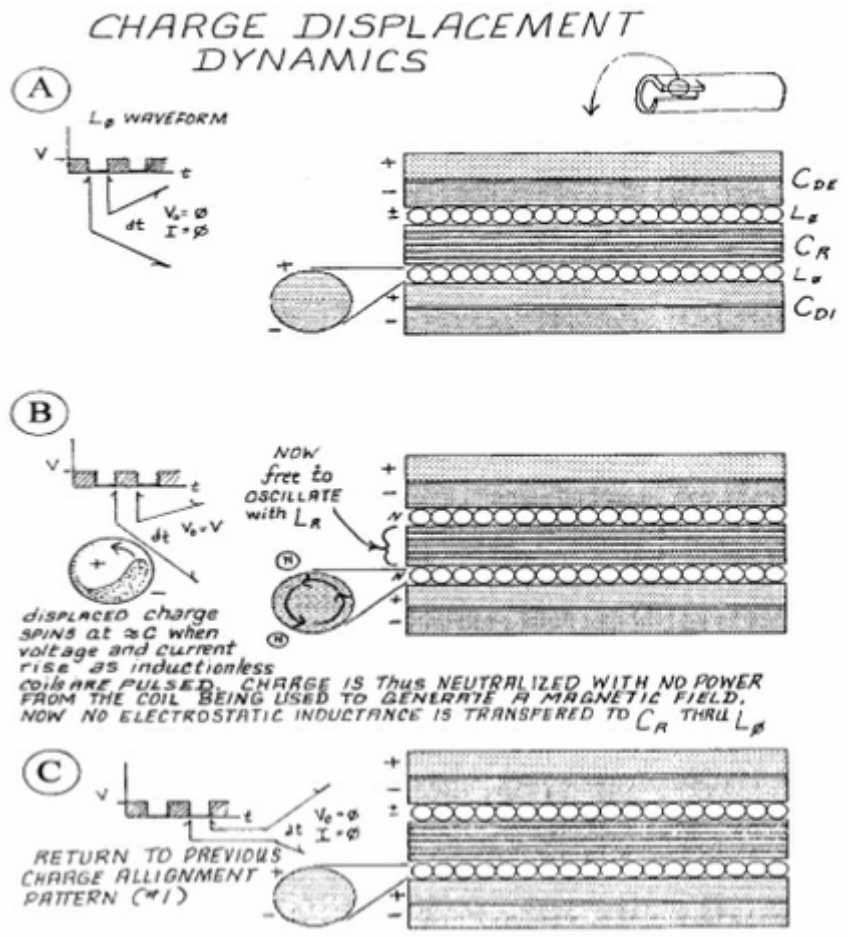


Fig.8 above, discusses the charge displacement dynamics involved in the E-Stress Amplifier power system. Diagram **A** illustrates the induction-less coil **Lo** waveform. Diagram **B** depicts the displaced charge spins when the voltage and current rise as the induction-less coils are pulsed, while diagram **C** shows the waveform during the return to the previous charge patterns, starting the cyclic over again.

Patrick Kelly
www.free-energy-info.tuks.nl
www.free-energy-info.com
www.free-energy-info.co.uk
www.free-energy-devices.com

Chapter 4: Gravity-Powered Systems

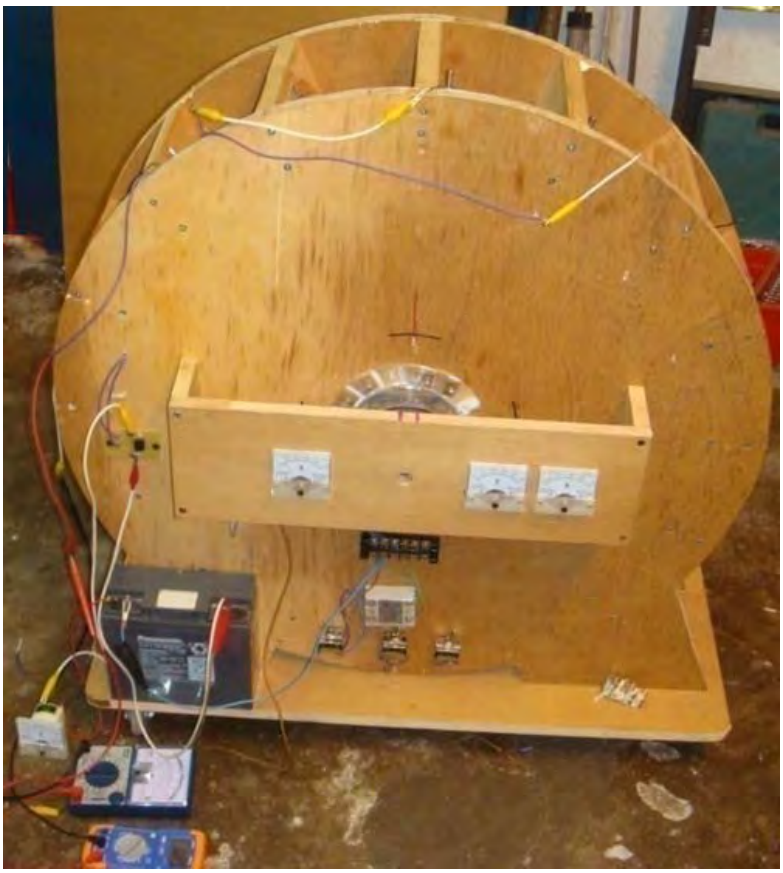
Note: If you are not at all familiar with basic electronics, you might find it easier to follow parts of this chapter if you read chapter 12 first.

Lawrence Tseung's COP=3.4 Pulsed-Flywheel Generator.

It is generally not realised that excess energy can be obtained from pulsing a flywheel or other gravitational device.

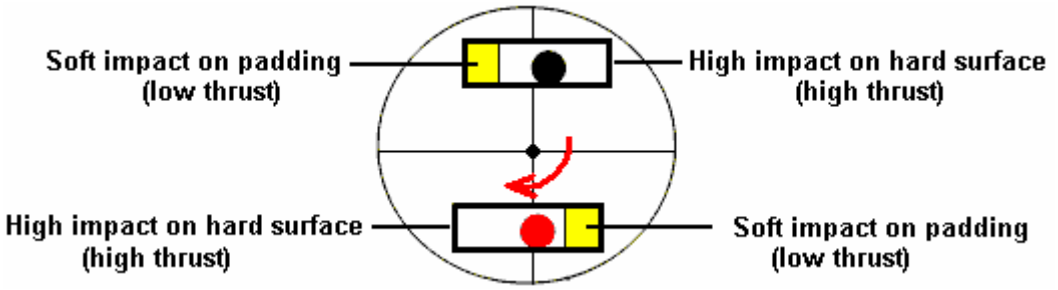
This fact has recently been stressed by Lawrence Tseung who refers to the extra energy obtained in this way as being "Lead-out" energy. This gravitational feature has been part of university Engineering courses for decades, where it has been taught that the loading stress on a bridge caused by a load rolling across the bridge is far less than the stress caused if that same load were suddenly dropped on to the bridge.

This impulse technology has been known for some time and it is demonstrated driving a canoe in the video at http://www.youtube.com/watch?v=aKWPht3fU-o but Lawrence points out the potential for using it as a method for gaining excess energy for practical use. In October 2009, Lawrence and his band of helpers ran public demonstrations of an early prototype electrical pulsing system which produces excess output energy of COP = 3.3, that is, with 3.3 times more output energy than the user has to put into it to make it work:



Video: http://www.youtube.com/watch?v=tiafQ9R-REs&feature=mfu_in_order&list=UL Lawrence is busy developing this device further as he intends to construct one with a output energy excess of several kilowatts.

Behind this device is Lawrence's "Lead-out" theory and for this he suggests a simple arrangement to demonstrate the principle. He presents the case of a rotor which has two substantial weights contained in two cylinders attached to the rotor:

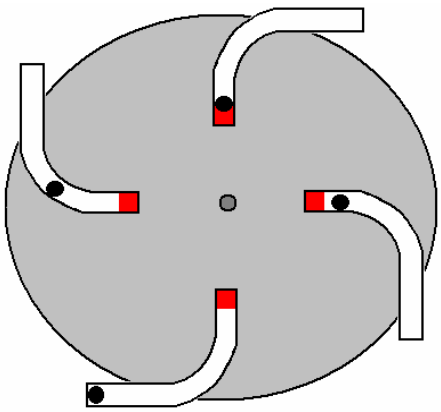


As the disc rotates, the ball falls down the length of the tube. At one end, the tube has a rigid cap which causes a significant impact when the ball hits it. The other end of the tube is padded and that cushions the impact which causes a net imbalance in the impacts and that maintains the rotation.

There is a prototype implementation on YouTube but the implementation is not adequate and the disc stops rotating after five minutes. The YouTube video slot is located at: http://www.youtube.com/watch?v=zykButGc22U&feature=related and there are two significant problems with that particular build. Firstly, the tube rotation is too slow to be effective and instead of the weight falling under gravity and accelerating to a good speed before the impact, the weight just rolls gently down a minor slope and does not make a major impact.

Secondly, the weights are far too small for the size of the wheel and there are only two weights providing impacts very widely spaced apart as the wheel rotates only slowly. One man made a ten-foot version and it rotated steadily for ten months after which time his wife insisted that it be taken apart as it was too noisy.

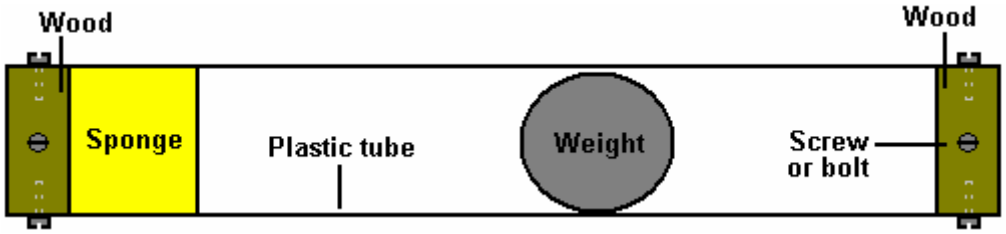
I would suggest some modifications to the wheel as Lawrence is far too busy with developing his COP>1 pulse implementation. Firstly, the movement of each weight should be delayed until the tube is much nearer the vertical. This can be achieved by curving part of the tube like this:



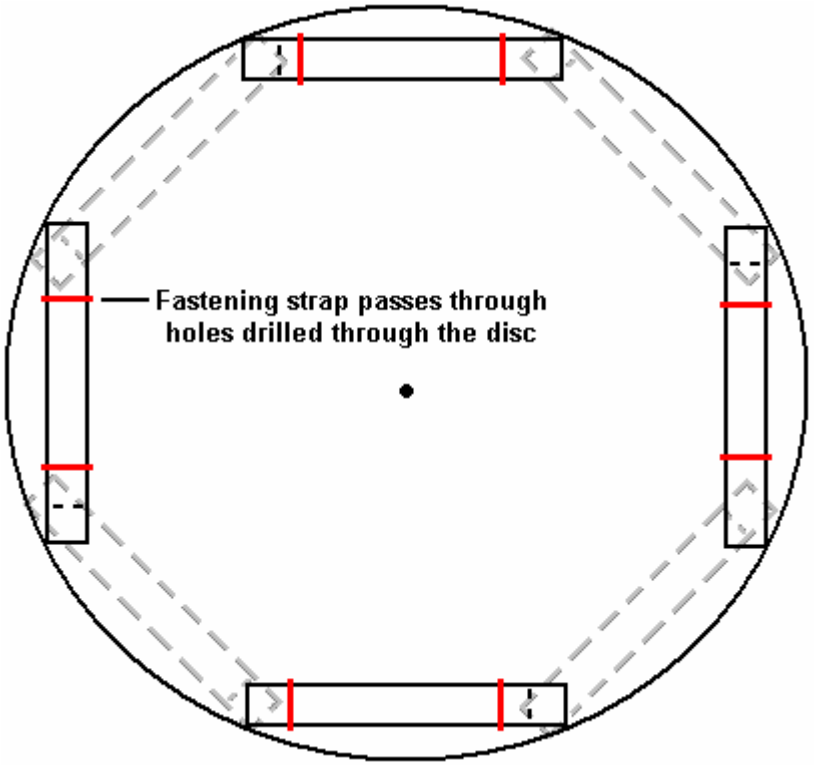
This way, the ball does not start rolling until the main part of the tube is near vertical. This allows a much greater acceleration and impact. The weighted ball should be much larger, say 2" (50 mm) in diameter and made of lead, in order to generate a significant thrust. Also, the cushioned ends of the tubes should be aligned with the pivot of the wheel so that any residual impact does not generate a turning force in the wrong direction. there is a negative turning effect due to the lever arm of the bottom weight. This turning force is only there for a small arc of rotation as the weight will roll inwards as soon as the tube section rises above the horizontal and as the tube then transitions into a circular curve, the movement inwards is gentle. It probably would be better if the tubes were angled slightly more in the clockwise direction, rather than exactly as shown in the diagram.

Secondly, there should be eight tubes on the disc, four on each side and one side staggered by 45 degrees so that there is a driving impact every 45 degrees instead of the 180 degrees of the version shown in the YouTube video. With that arrangement of four times as many impacts, each substantially greater, and no significant reverse impacts, the wheel has a much better chance of successful rotation without needing to be particularly large. The wheel itself should not be light as it acts as a flywheel and a pulsed flywheel has already been shown to produce excess power. The wheel bearings should be ball races and not the closed variety because those ones are packed with grease and have a serious resistance to rotation. Instead, the open-sided variety of ball bearing should be used as they rotate very freely.

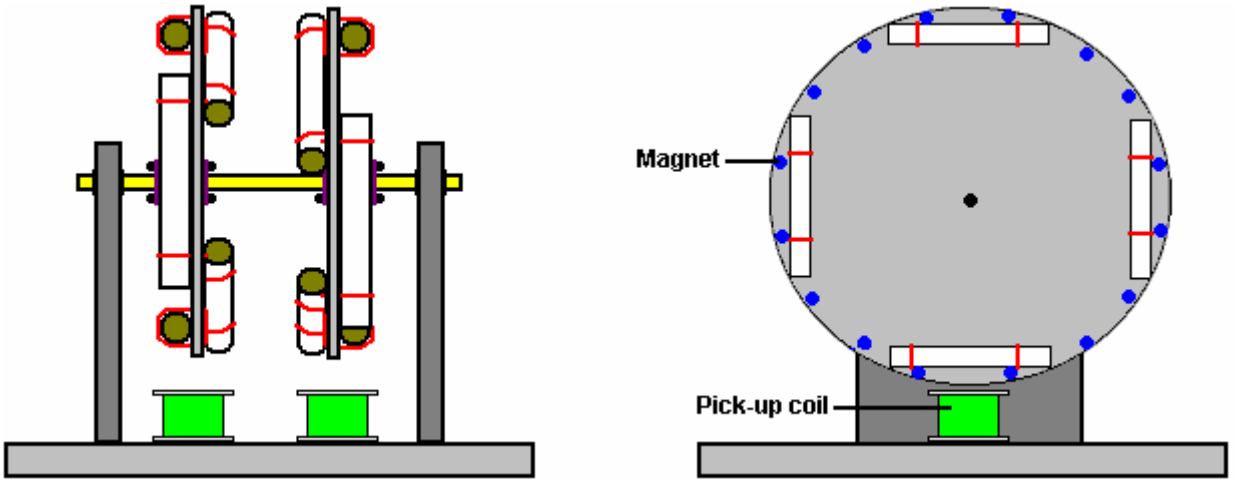
Using straight tubes for illustration, each tube could be like this:



Here, a wood disc is fitted to each end of a piece of plastic tube and held securely in place with screws or bolts which pass through small holes drilled in the plastic pipe and screw into the wooden disc. A piece of thick sponge is glued to the disc at one end and the heavy weight inside the tube is not a tight fit so that it can move very freely inside the tube. Four of these tubes are fitted to each side of each disc used in the device as shown here:



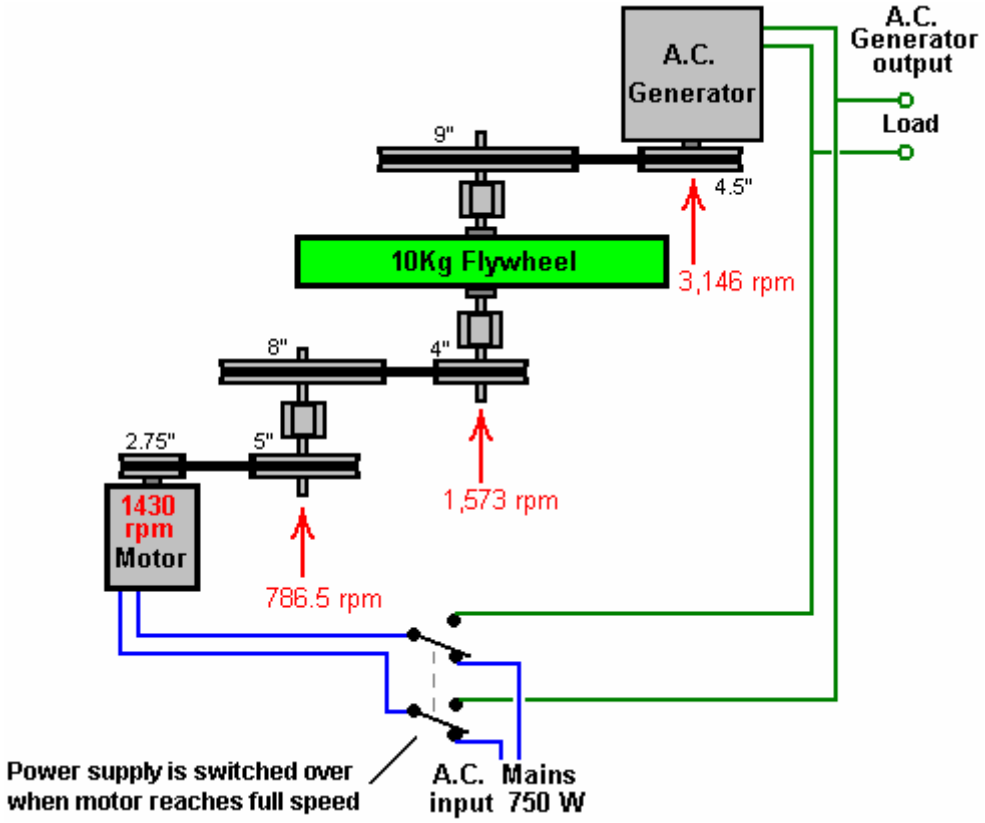
The four tubes attached to the back of the disc are 45 degrees away from the tubes mounted on the front of the disc. Each tube is attached securely in place with straps which pass through the disc and are secured on the far side. The tubes can also be glued in place to further strengthen the attachment. These eight tubes give an unbalanced impact for every 45 degrees of rotation. If two of these discs are attached to a common rotor shaft, then the second disc can be positioned 22.5 degrees around from the first one. That arrangement gives an unbalanced impact for every 22.5 degrees of rotation. If three discs were placed on a common rotor shaft and evenly positioned, then there would be an unbalanced impact every 15 degrees of rotation, which is 24 impacts per rotation. A two-disc arrangement might look like this:



If the rotor spins well, then it would be worth while attaching a series of magnets to the discs, being careful to keep each disc perfectly balanced. One or more air-core coils can then be used to determine if current can be drawn from the device without stopping the rotation. The coils should not have a magnetic core as that would cause a major drag on the rotation whether current was being drawn or not.

Chas Campbell's Flywheel System.

Recently, Mr. Chas Campbell of Australia demonstrated electrical power gain with a flywheel system which he developed:



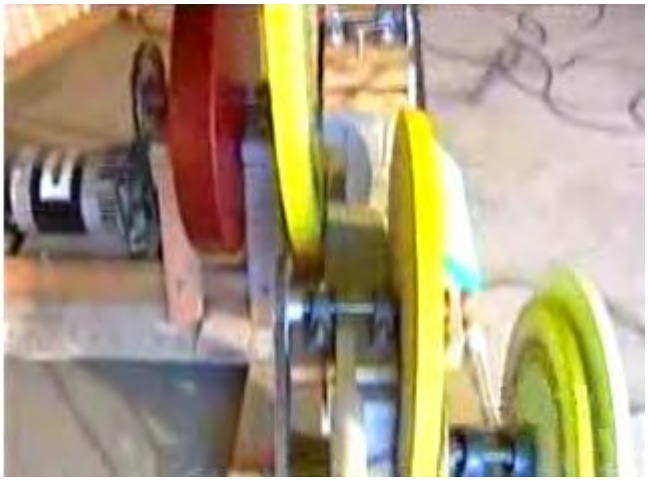
But what this diagram does not show, is that a couple of the drive belts are left with excessive slack. This causes a rapid series of jerks in the drive between the mains motor and the flywheel. These occur so rapidly that they do not appear noticeable when looking at the system operating. However, this stream of very short pulses in the drive chain, generates a considerable amount of excess energy drawn from the gravitational field. Chas has now confirmed the excess energy by getting the flywheel up to speed and then switching the drive motor input to the output generator. The result is a self-powered system capable of running extra loads.

Let me explain the overall system. A mains motor of 750 watt capacity (1 horsepower) is used to drive a series of belts and pulleys which form a gear-train which produces over twice the rotational speed at the shaft of an electrical generator. The intriguing thing about this system is that greater electrical power can be drawn from the output generator than appears to be drawn from the input drive to the motor. How can that be? Well, Mr Tseung's gravity theory explains that if a energy pulse is applied to a flywheel, then during the instant of that pulse, excess energy equal to **2mgr** is fed into the flywheel, where "**m**" is the mass (weight) of the flywheel, "**g**" is the gravitational constant and "**r**" is the radius of the centre of mass of the flywheel, that is, the distance from the axle to the point at which the weight of the wheel appears to act. If all of the flywheel weight is at the rim of the wheel, the "r" would be the radius of the wheel itself.

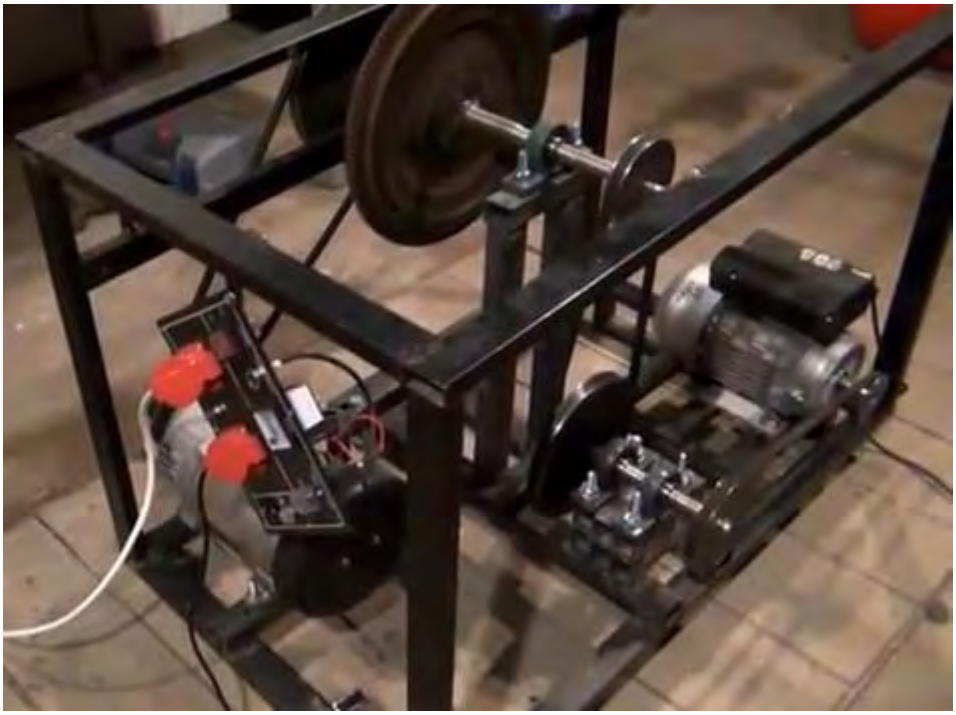
This means that if the flywheel (which is red in the following photographs) is driven smoothly at constant speed, then there is no energy gain. However, if the drive is not smooth, then excess energy is drawn from the gravitational field. That energy increases as the diameter of the flywheel increases. It also increases as the weight of the flywheel increases. It also increases if the flywheel weight is concentrated as far out towards the rim of the flywheel as is possible. It also increases, the faster the impulses are applied to the system.

However, Jacob Bitsadze points out that another mechanism comes into play even if all of the belts are correctly tensioned. The effect is caused by the perpetual inward acceleration of the material of the flywheel due to the fact that it rotates in a fixed position. He refers to it as being 'the rule of shoulder of Archimedes' which is not something with which I am familiar. The important point is that Chas Campbell's system is self-powered and can power other equipment.

Now take a look at the construction which Chas has used:

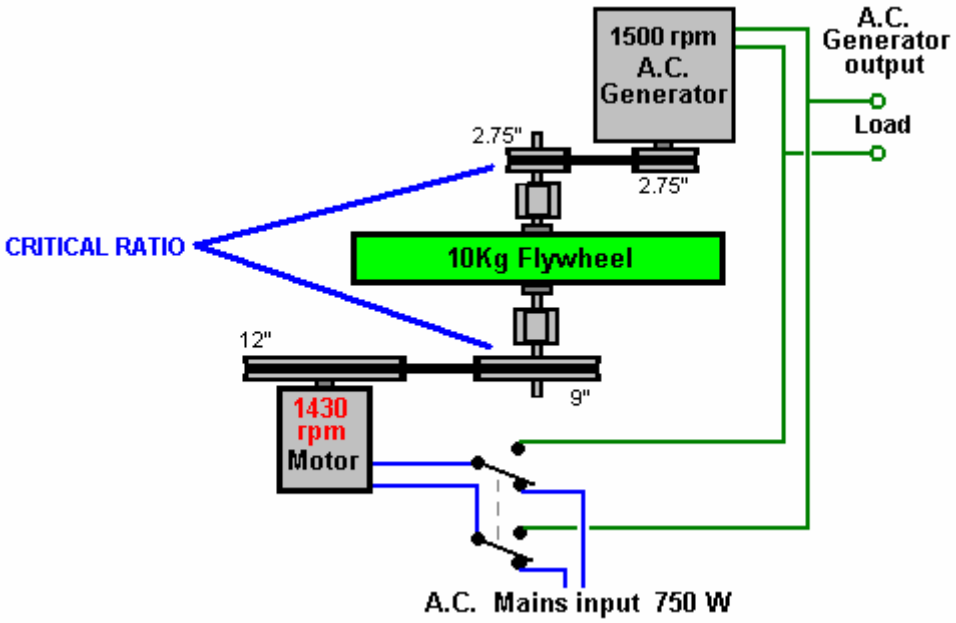


You notice that not only does he have a heavy flywheel of a fair size, but that there are three or four other large diameter discs mounted where they also rotate at the intermediate speeds of rotation. While these discs may well not have been placed there as flywheels, nevertheless, they do act as flywheels, and each one of them will be contributing to the free-energy gain of the system as a whole. A replication video with 750 watts input and 2340 watts output is here: http://www.youtube.com/watch?v=98aiISB2DNw and this implementation does not appear to have a heavy flywheel as you can see from this picture, although the largest pulley wheel looks as if it contains considerable weight:



Jacob Byzehr's Analysis.

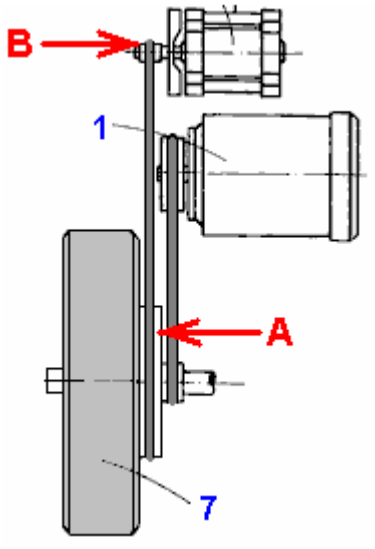
In 1998, Jacob lodged a patent application for a design of the type shown by Chas Campbell. Jacob has analysed the operation and he draws attention to a key design factor:



Jacob states that a very important feature for high performance with a system of this kind is the ratio of the diameters of the driving and take-off pulleys on the shaft which contains the flywheel, especially with systems where the flywheel rotates at high speed. The driving pulley needs to be three or four times larger than the power take-off pulley. Using Chas' 1430 rpm motor and a commonly available 1500 rpm generator, the 12:9 step-up to the shaft of the flywheel gives a satisfactory generator speed while providing a 3.27 ratio between the 9-inch diameter driving pulley and the 2.75" diameter power take-off pulley. If a generator which has been designed for wind-generator use and which has it's peak output power at just 600 rpm is used, then an even better pulley diameter ratio can be achieved.

The Self-powered Generator of José Luis García del Castillo

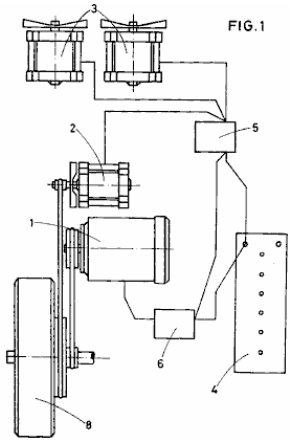
In 1998, Spanish patent ES 2,119,690 was granted to José Luis García del Castillo. I suspect that the auxiliary generators shown in the patent are only there to get the patent accepted by the patent examiner, rather than because they are actually needed. If that is correct, then the design is almost the same as Chas Campbell's design, although built in a more compact form:



As Jacob Byzehr points out, an energy gain is achieved through inertial acceleration caused by having the pulley wheel "A" attached to the flywheel shaft, larger than the pulley wheel "B" attached to the shaft of the generator. As drawn, there is a major difference in those diameters. Here is an attempted translation of the patent:

Patent: ES 2119690 Date: 1 Oct 1998 Inventor: José Luis García del Castillo

AUTONOMOUS ENERGY REGENERATION SYSTEM



Abstract

The system comprises an electric motor drive (1), a main generator (2), auxiliary generators (3), a battery (4), a charging regulator (5), and a speed regulator (6). The system is intended to generate its own operating power, and provide an extra supply for other purposes.

DESCRIPTION

Field of the invention

The present invention refers to a self-contained system of energy regeneration, which in addition has several advantages set out below.

Background of the invention

It has been known for many years, how to construct machines which can generate electric current. These are known by the generic name of "electric power generators", consisting of rotating machine that transforms

mechanical power into electrical power as a result of alternative action between a magnetic field and a moving conductor.

However, the various types of generator which make up the current state of the art, require the help of a motor, which transforms mechanical power into electrical energy, and that motor requires an independent power source which must be supplied continuously.

Thus, a system capable of generating its own power supply as well as providing an extra power supply for other purposes, is not known in the current state of the art.

Summary of the invention

The applicant for the present patent has designed an self-contained energy regeneration system, capable of producing its own operating energy in addition to generating a surplus which can be used in electrical networks using voltage converters required for any electrical installation, whether in homes, offices, warehouses etc.., with it is possible to reach places where it is difficult to install the power grid, allowing its use as an alternative source of energy other than solar or wind power.

Other applications would be in the automotive field, as a power source for motorcycles, cars, etc. by connecting the system to the propelling motor, and thus achieving the necessary motion of the vehicle.

Overall, the system is comprised of the following basic components:
1. An electric traction motor.
2. A main generator.
3. Various auxiliary generators.
4. A battery or accumulator.
5. A load and output-power controller.
6. A speed controller.

The electric drive motor supplies the necessary electromotive force needed for the system to operate, the generator supplies power to the system, charging the battery and providing direct power to the traction motor when needed, or if the battery is fully charged, then just to power the motor. It can go provide direct mechanical power by using pulleys and belts, gears or any other means.

The auxiliary generators are responsible for supplying backup power and can use propellers or be in the form of a turbine operating by the action of the wind or by gears attached to a flywheel placed in the traction motor.

The function of the battery is to provide the power needed to start the motor and in addition, to supply any extra power that the engine may need during operation. The battery is recharged by the main generator which is driven directly by the motor. The function of the charge controller is to prevent the battery becoming overcharged. The function of the speed controller is to control the speed of the drive motor.

The present invention offers the advantages described above, as well as others which will be understood from the example embodiment of the system described in detail below, to facilitate understanding of the features stated above, and introducing at the same time, various in addition to the present specification. It should be understood that the drawings are only by way of example and they do not limit the scope of the present invention in any respect, being just an example of one form of construction.

Brief description of the drawings

In the drawings:

Fig.1, is a diagram of the system as one example of a practical embodiment of the invention.

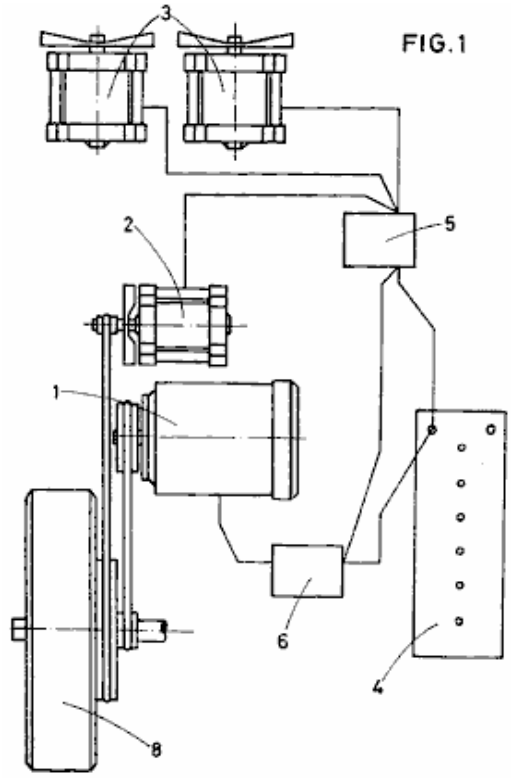
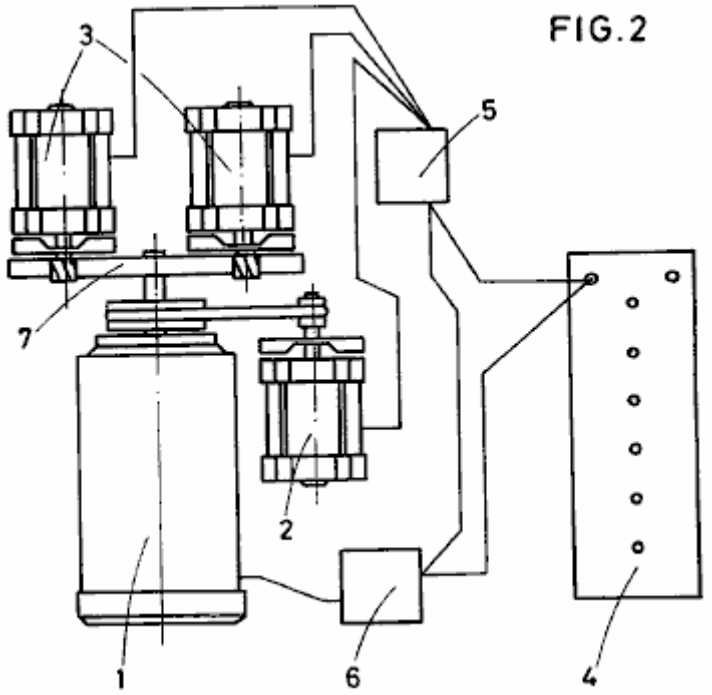
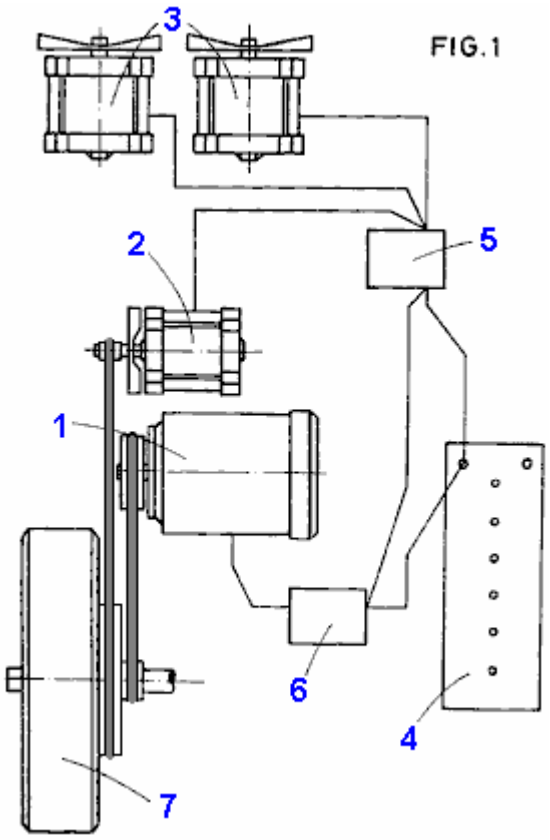


Fig.2 shows an alternative embodiment of this invention.



Description of a preferred embodiment of this invention

As shown in the drawing, the self-contained system of energy regeneration, in accordance with an embodiment of the present patent, comprises an electric traction motor (**1**), a main generator (**2**), several auxiliary generators (**3**), a battery or accumulator (**4**), a charge controller and power supply (**5**), and a motor speed controller (**6**).



The electric drive motor (**1**) provides the necessary electromotive force for operating the system, and its voltage and power are selected in accordance with whatever size of system you wish to construct.

The main generator (**2**) supplies power to the system, firstly, to recharge the battery (**4**), and secondly, direct power to the motor (**1**) if it requires it. When the battery (**4**) is fully charged, its charge is maintained by power from the motor, supplied by suitable pulleys or other method of transmission of mechanical power.

The speed of revolution of the generator must be arranged through choice of the gearing between the motor and the generator, so that when the motor is operating at its maximum speed and drawing its maximum current, that the generator is spinning fast enough to supply that current. The main generator (**2**) will be therefore connected electrically to the battery (**4**) and mechanically to the motor (**1**). The auxiliary generators (**3**), are in the form of a turbine, operating by wind action or by gears attached to a flywheel (**7**), driven by motor (**1**), as shown in **Fig.2**. These auxiliary generators (**3**) provide reserve power for the system.

The battery (**4)**, must have a capacity which exceeds the maximum power of the motor (**1**), and its role in the system is to provide the power needed to start the motor (**1**), and to supply any extra energy which the motor (**1**), may need during operation. The battery is recharged directly by the main generator (**2**) which is driven by the motor (**1**).

The charge controller and power distributor (**5**) is positioned between the main generator (**2**), the auxiliary generators (**3**) and the battery (**4**). Its job is to regulate the current draw from the battery (**4**), to prevent excessive current draw. It also distributes any surplus power as a direct feed to the generator (**2**) and the drive motor (**1**) when it needs additional current. The auxiliary generators (**3**) can either provide additional power to the motor (**1**), or their power output may be used for any other power needs.

The speed regulator (**6**), is intended to regulate the speed of the motor (**1**), this adjustment is gradual, and is adjusted to match the intended use of the system.

System applications can be many and diverse, noteworthy among which are uses in the automotive field, where it can be used as a means for propelling cars and other motor vehicles, with the motor (**1**) connected to a drive wheel which propels the vehicle. In these applications, the auxiliary generators (**3**), can be fitted with propellers or be turbine-shaped, so that the passing wind provides extra energy to the electrical system. With electric vehicles, only the main generator is connected to a drive wheel.

Other applications for this system are in the field of energy supply, i.e. use in electrical networks. The advantage of using this patented system is that the power supply is practically inexhaustible and clean, and subject only to component wear and tear during operation.

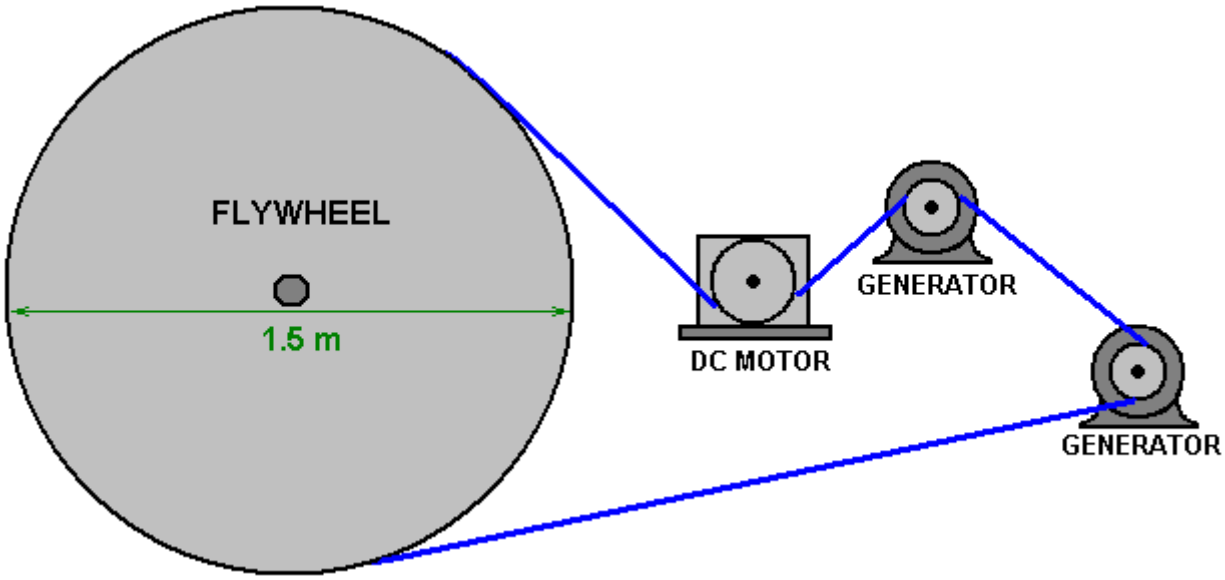
Its operation is as follows:
Battery (**4**) provides the energy needed to start the motor (**1**), and allows for the possibility of increased electrical input being needed at certain times during operation of the system. The battery (**4**) is electrically connected to the motor (**1**) through the speed controller (**6**), which is fed from the main generator (**2**) by the load distribution controller (**5**). The generator (**2**), is driven directly by the drive motor (**1**), and the transmission of motion from one component to another is through gears, belts and pulleys, or any other conventional means.

The auxiliary generators (**3**) are electrically connected to the charge controller (**5**), and the battery (**4**), and their movement by the wind generates their own energy, having turbine blades or propellers, or through being spun by the flywheel (**7**) which is connected directly to the motor (**1**). The energy produced by these auxiliary generators (**3**) may be used for charging other batteries for later use, or used directly to power other electrical equipment or electrical installation networks. Through voltage converters it is possible to convert the voltage produced by the system to a voltage which is suitable for use in other equipment.

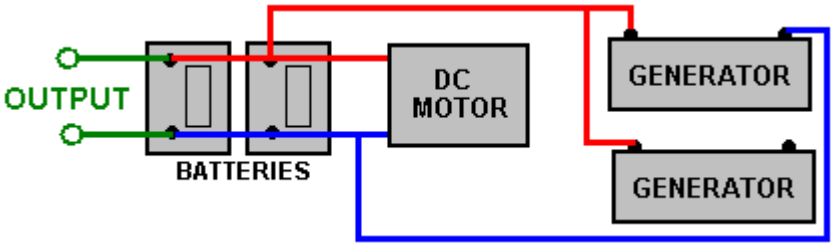
The Wilson Self-Powered DC Generator

Mr. Wilson of Texas built a self-powered generator system using an old table and some car parts. His construction was shaky, but in spite of that, it powered itself and other equipment. The table which he used was five feet (1.5 m) in diameter and 2-inches (50 mm) thick which means that it will have weighed at least 130 pounds or 60 Kilograms which is a substantial amount, well in excess of that used by Chas Campbell with his AC self-powered system. In this DC construction the system was driven by a standard, unmodified, off-the-shelf DC motor powered by two car batteries wired in parallel to give a larger current capacity. These batteries were kept charged up by two 'generators' from pre-1964 American cars (the closest available today are permanent magnet alternators). These generators also powered additional equipment and Mr Wilson pointed out that three or more generators could be run by the system, giving a substantial level of excess electrical power.

The machine has to be described as 'shaky' because he chose to convert the table top into a V-pulley belt drive flywheel by driving a series of nails into the edge of the wooden disc, with those nails angled to form a V shaped gap through which he ran a pulley belt. After three days of continuous running, those nails started to come out, causing him to power the system down. This unit was built around 1990, and if anyone decides to attempt a replication, then I suggest that the rim of the wooden disc is grooved to take the belt rather than relying on nails. The arrangement was like this:



There was also a belt-tensioning roller which is not shown in the diagram above which assumes that the flywheel has been grooved to take the drive belt. Schematically, the arrangement was like this:



Here, the additional output can be used directly for powering 12-volt equipment or an inverter can be used to provide mains voltage and frequency. A typical inverter looks like this:

The battery power is connected to one end using thick cables to carry the heavy current, and one or more mains sockets are provided at the other end of the case, along with an On/Off switch and power indicators. Inverters come in many sizes and power ratings, generally ranging from 150 watts to 3,000 watts (3 kW). The more expensive ones are specified as "True Sine-Wave Output" but very few present day items of equipment will not run well on the cheaper versions which do not produce a true sine-wave output.

Mr Wilson decided not to patent his design and instead wanted it to be open-source information for anybody to use freely. However, the Jesse McQueen patent shown in chapter 13 looks to be Mr Wilson's design although the flywheel does not appear to be mentioned there. It should be stressed that the generator output needs to be high and so permanent magnet types are considered to be essential for this application. The specialised motor (and consequently, generator) winding methods of 'UFOpolitics' shown in chapter 2, raise efficiencies by a factor of typically 300% or more, and so would raise the output of this system very substantially if they were applied to the motor, or the generators, or both.

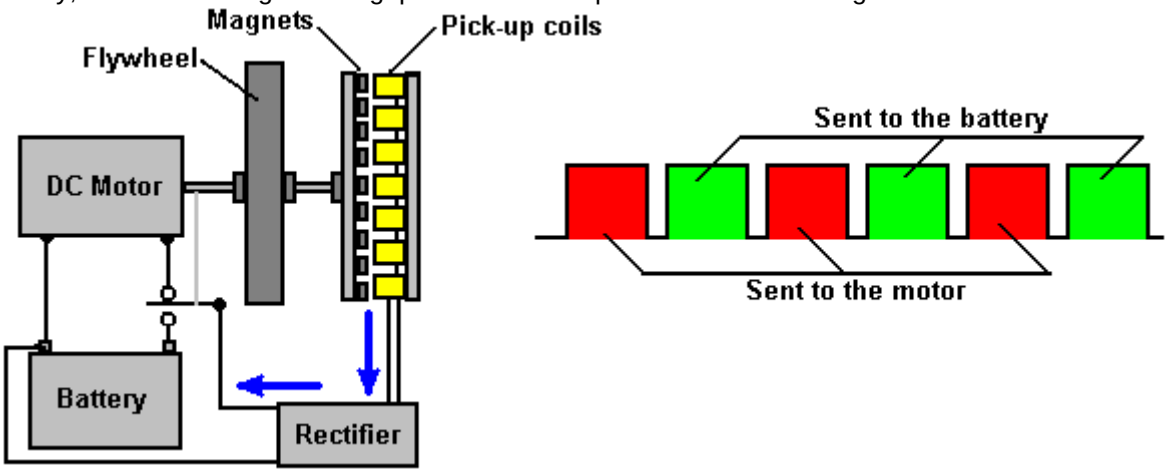
John Bedini's COP=8 Pulsed Flywheel.

The Chas Campbell system is not an isolated case. On page 19 of the book "Free Energy Generation - Circuits and Schematics" John Bedini shows a diagram of a motor/generator which he has had running for three years continuously while keeping it's own battery fully charged. At John's web site http://www.icehouse.net/john1/index11.html about half way down the page, there is a black and white picture of a very large construction version of this motor built by Jim Watson and which had an excess power output of many kilowatts, due to the very large size and weight of it's flywheel. The excess energy is drawn from the Earth's gravity field, and so, two factors are involved.

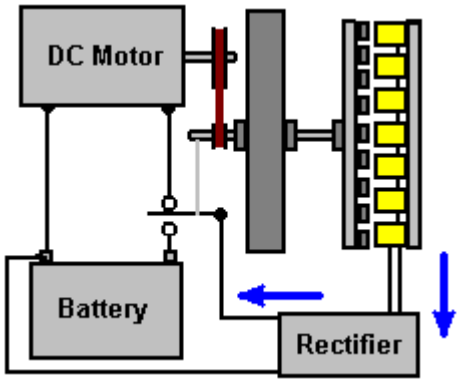
The first is the size, weight and speed of oration of the flywheel itself and the second is the effectiveness of the drive connection between the drive motor and the flywheel. In John's prototype, the flywheel is relatively small, limiting the output power and calling for careful manual tuning of the system, and John shows the motor being directly linked to the shaft of the flywheel, and if that is the case, then in my opinion, that limits the output power considerably as can be seen from the work of Jacob Byhehr earlier in this chapter.

The general strategy is that the motor spins the flywheel and the flywheel shaft spins a disc with permanent magnets mounted on it. The magnets have their South poles facing corresponding helically wound coils which are connected in series. As the magnets pass by the coils, a voltage is generated and current is then drawn from the coils and fed, first to the motor to power it and then secondly to the battery to keep it charged.
John shows his switching mechanism as a mechanical attachment on the flywheel shaft with a conducting sector of about 110 degrees of arc. This gives equal duration pulses being fed back to the motor and then to

the battery, with there being a short gap between each pulse and the following one:



While this is a simple concept, in my opinion, there is scope for improvement. As current is drawn from the output coils, the current flow creates magnetic effects which cause drag which opposes the rotation of the flywheel. This suggests that controlling the timing of the current draw with the switching arrangement used by Robert Adams (chapter 2) would convert that drag into a push which would help the flywheel on its way instead of impeding it. Probably of greater importance would be to gear the drive motor up as Jacob points out. Leaving the generator disc directly attached to the flywheel shaft, the motor might be geared up by, say, a 2:1 ratio:



John's mechanical switching has the advantage of always being synchronised with the flywheel but it has the disadvantage of wear and tear on the mechanical parts. An electronic circuit to replace the mechanical parts should not be difficult to arrange and if you want the feedback from the generator section to be synchronised with the flywheel (which actually does not seem to be in any way necessary), then an optical disc or a magnetic sensor can be used. This generator system of John's can have considerable excess output.

James Hardy's Self-Powered Water-Jet Generator.

As described in more detail in Chapter 2 and Chapter 8, there is a very simple device based on a high-power water pump. In this system, a small quantity of water is pumped around continuously, in the same general style as an ornamental fountain. The difference here is that a high speed jet of water is produced and directed at a turbine wheel. The turbine wheel can be of any type as indicated in the patent which James has been awarded for this design. In the video at present on the web, the water wheel is of very simple design and yet works well – it is shown here:

Small discs are attached to the wheel at widely spaced intervals around it's rim. The water jet hits these and applies an impulse to the wheel, driving it around, but also adding extra energy through those impulses.

The waterwheel is coupled to a standard electrical generator via pulleys and V-belts. The system is started using the mains supply and then when it is running at full speed, the electrical supply for the pump is switched over from the mains to the output of it's own generator. This is exactly the same as Chas Campbell does with his pulsed flywheel and both systems are capable of powering additional standard electrical equipment intended for mains use.

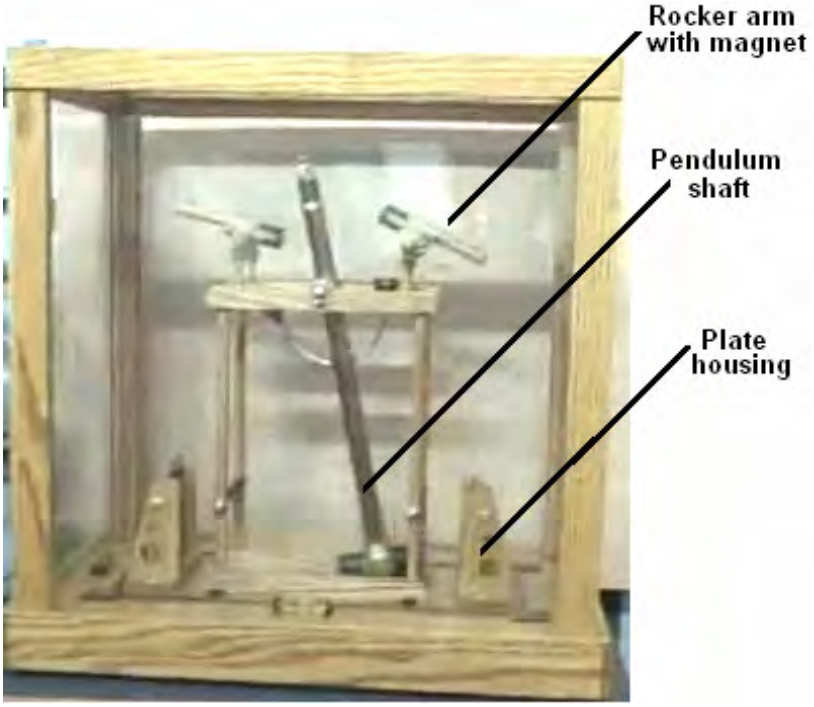
Chas Campbell's flywheel, John Bedini's flywheel and this water-jet generator all demonstrate very clearly that environmental energy is readily available for us to use any time we choose to do so. All that is necessary is for us to construct one of these devices.

The Permanent Magnet Pendulum.

At the present time, there is a short video clip on YouTube, showing a pendulum which has been running unaided for two years: http://www.youtube.com/watch?v=SZjNbjhxgt4 and which uses both gravity and magnetism to keep going. The device is installed in a case with transparent sides:

The pendulum itself looks rather like a sledgehammer due to it's rigid shaft and the additional magnets mounted on the weight. The above picture shows the pendulum at the end of it's swing to the right and the picture below, in it's extreme left hand swing position:

Which indicates the swing covers a fairly short distance. Mounted near the top of the pendulum, there are two pivoted arms which look quite like microphones, due to having large magnets mounted on their innermost ends:



The device operates like this: The pendulum swings to the right and as it does so, it raises a magnet attached to the pendulum shaft by a curved silver arm:

Presumably, the arm is curved to avoid the constructional complications at the pendulum pivot which would be caused by a straight mounting arm attached to the pendulum shaft. The rising magnet attached to the pendulum pushes the magnet end of the rocker arm upwards even though it does not come close to it.

The rocker arm is used to raise and lower a plate which has a magnet mounted in it. The raising and lowering is achieved by having two cords attached to the end of the rocker arm and their other ends attached to the two upper corners of the moving plate:

The plate slides in two slots in the support housing and the plate movement is relatively small:

The tipping up of the lever arm drops the plate down as the pendulum approaches the plate. This introduces a magnetic braking effect where some of the momentum of the pendulum weight is stored in the opposing magnetic fields of the pendulum magnets and the plate magnet. This brakes the pendulum movement and gives it a magnetic push on its opposite swing, sustaining it's swinging day after day after day.

This is a clever arrangement and the device on display has been built to a very high standard of construction. It does not appear to have any additional energy take off, but seems quite likely that air-core coils could be used along the swing path to generate electrical power. The arrangement appears so close to John Bedini's pendulum battery charger that it may well be possible to use a pendulum of this type to charge batteries just as John does.

While this looks like a very simple device, it is highly likely that it requires exact adjustment of the length of the lever arms, the magnetic gap sizes in relation to the strength of the magnets, etc. etc. Repeated small adjustments are probably needed to get the device operating smoothly and sustaining the pendulum swing. All in all though, it is a very interesting device.

Jerzy Zbikowshi's Mechanical Power Amplifier.

We come now to a device which I would love to describe as "impossible" but reluctantly, I can't actually do that. On the surface, this device has every appearance of being impossible, and yet it has been measured in a laboratory as being 147% efficient. Perhaps the laboratory measurements are wrong, however, there seems to be very little scope for measurement error as the device is so basically simple. My problem is that if the results are 100% genuine, which is distinctly possible, then a series of these arranged in a circle, each driving the next one, it would create a self-powered device and I can't explain where the driving power would come from. I can understand pretty much every other device in this eBook, but this one has me stumped. As I don't have any basis for claiming to be a genius, I am sharing the information here and I will let you decide if it can work as the patent claims that it does.

The patent in question is the very innocent looking US 7,780,559 entitled "Chain Transmission" which innocently states that it is a single-chain system for rotating a large gearwheel at the same rate as a smaller, driving gearwheel, and without question, that is exactly what it does. At this point, my Engineering training jumps in and says "sure, but the overall mechanical efficiency will be less than 100% and while the larger gearwheel does rotate at the same rate, it will do so far less powerfully, and you have exactly the same effect as driving the second shaft with a small gearwheel which has a large gearwheel bolted to it.

The only problem with this is that testing appears to show that this is not the case and in fact, (probably due to the larger lever arm of the larger gearwheel radius) the arrangement has an output power which was measured in the prototype as being 47% greater than the input power. OK, so how does it work?

In the diagram shown here, a small-diameter driving wheel marked "**1**" has exactly the same number of teeth as the much larger driven wheel marked "**2**". As they are linked by a chain, these two wheels rotate at exactly the same rate, that is, the revolutions per minute are exactly the same for each of those two wheels.

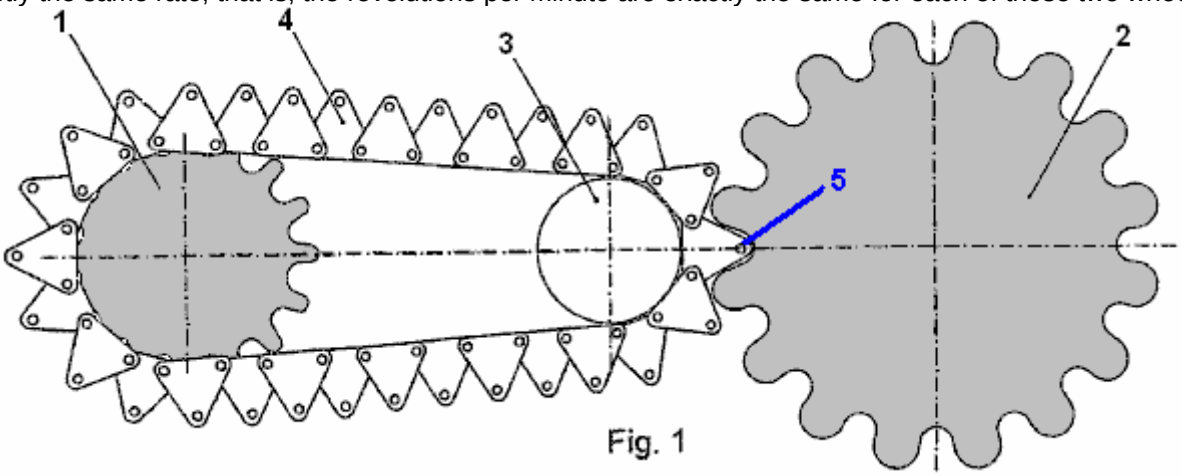


Fig. 1

The way that the chain manages to push the larger teeth of the large wheel is by having the driving roller "**5**" raised by a triangular link "4" so that it has the same rotational pitch as the teeth on the larger wheel.

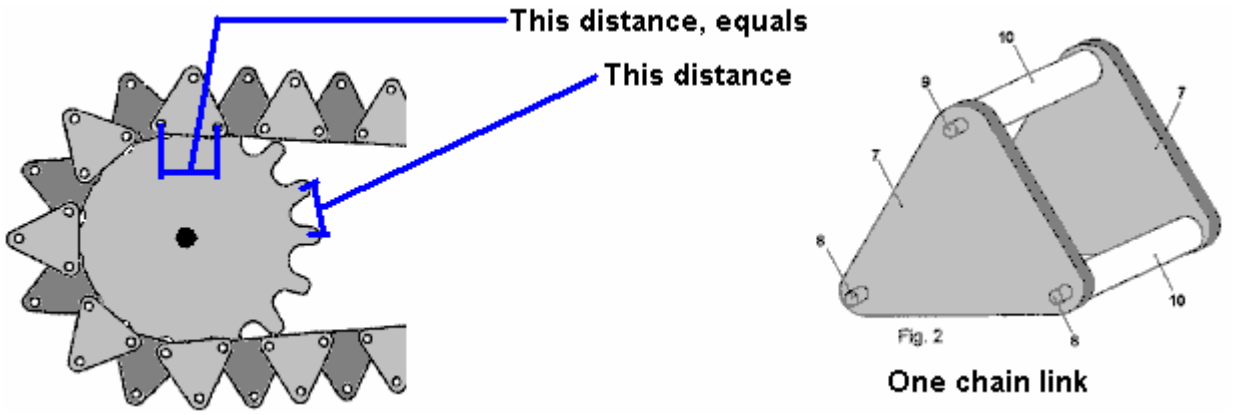
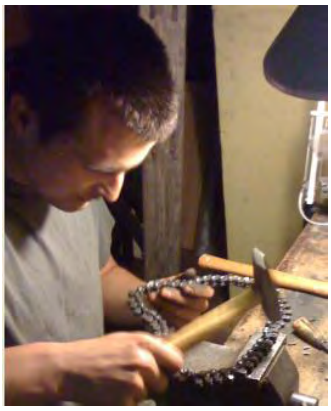


Fig. 2

One chain link



My immediate reaction to this is to say that as the triangular links in the drive chain have a somewhat narrower base than their height, that this will cause the driving roller "**5**" to have a less powerful drive than the driving wheel "1". But if the lab measurements made on the prototype are correct, then that increased level arm effect is not sufficient to overcome the gains caused by the increased radius of the larger wheel. The lab measurements were made at the certified laboratory of the Institute of Electrical Machines and Drives of the Technical University of Wroclaw, Poland. A video presentation in Polish can be seen at http://www.focus.pl/video/film/perpetuum-mobile/. It is difficult to see how this chain drive could be COP>1 but it has the advantage that anyone with good mechanical construction skills can test it without the need for any knowledge of electronics.

Gravitational Effects.

We are all familiar with the effects of gravity. If you drop something, it falls downwards. Engineers and scientists are usually of the opinion that useful work cannot be performed on a continuous basis from gravity, as, they point out, when a weight falls and converts it's "potential energy" into useful work, you then have to put in just as much work to raise the weight up again to its starting point. While this appears to be a sound analysis of the situation, it is not actually true.

Some people claim that a gravity-powered device is impossible because, they say that it would be a "perpetual motion" machine, and they say, perpetual motion is impossible. In actual fact, perpetual motion is not impossible as the argument on it being impossible is based on calculations which assume that the object in question is part of a "closed" system, while in reality, it is most unlikely that any system in the universe is

actually a “closed” system, since everything is immersed in a massive sea of energy called the “zero-point energy field”. But that aside, let us examine the actual situation.

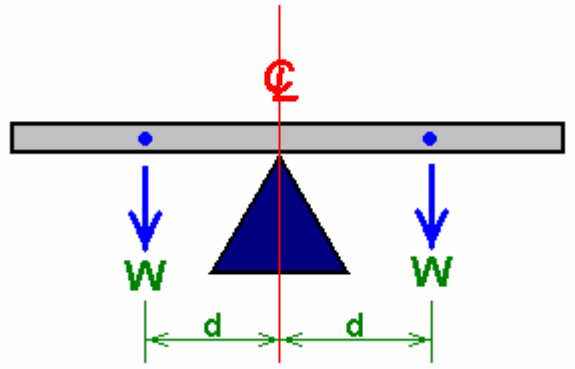
Johann Bessler made a fully working gravity wheel in 1712. A 300 pound (136 Kg) wheel which he demonstrated lifting a 70 pound weight through a distance of 80 feet, demonstrating an excess power of 5,600 foot-pounds. Considering the low level of technology at that time, there would appear to be very little scope for that demonstration to be a fake. If it were a fake, then the fake itself would have been a most impressive achievement.

However, Bessler acted in the same way as most inventors, and demanded that somebody would have to pay him a very large amount of money for the secret of how his gravity wheel worked. In common with the present day, there were no takers and Bessler took the details of his design to the grave with him. Not exactly an ideal situation for the rest of us.

However, the main argument against the possibility of a working gravity wheel is the idea that as gravity appears to exert a direct force in the direction of the earth, it therefore cannot be used to perform any useful work, especially since the efficiency of any device will be less than 100%.

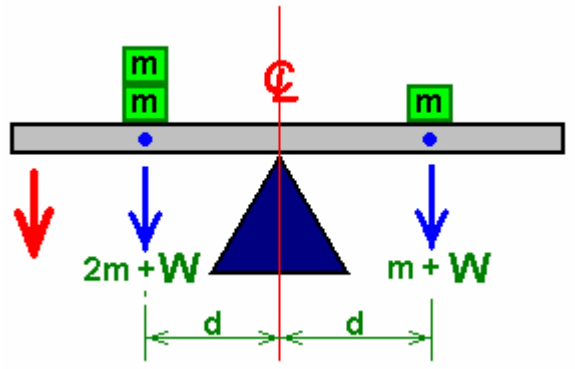
While it is certainly agreed that the efficiency of any wheel will be less than 100% as friction will definitely be a factor, it does not necessarily follow that a successful gravity wheel cannot be constructed. Let us apply a little common sense to the problem and see what results.

If we have a see-saw arrangement, where the device is exactly balanced, with the same length of a strong plank on each side of the pivot point, like this:



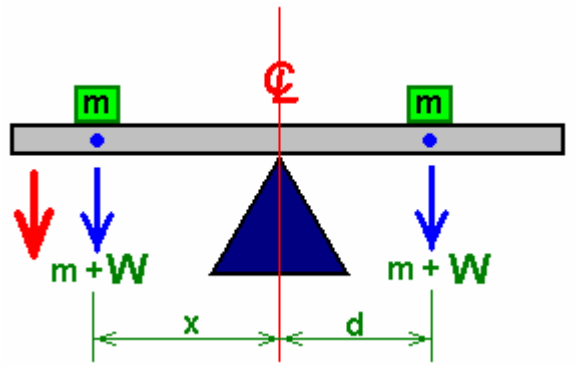
It balances because the weight of the plank (“**W**”) to the left of the support point tries to make the plank tip over in a counter-clockwise direction, while exactly the same weight (“**W**”) tries to tip it over in a clockwise direction. Both turning forces are d times W and as they match exactly, the plank does not move.

The turning force (d times W) is called the “torque”, and if we alter the arrangement by placing unequal weights on the plank, then the beam will tip over in the direction of the heavier side:



With this unequal loading, the beam will tip down on the left hand side, as indicated by the red arrow. This seems like a very simple thing, but it is a very important fact. Let me point out what happens here. As soon as the weight on one side of the pivot is bigger than the weight on the other side (both weights being an equal distance from the pivot point), then the heavy plank starts to move. Why does it move? Because gravity is pushing the weights downwards.

One other point is that the distance from the pivot point is also important. If the added weights “**m**” are equal but placed at different distances from the pivot point, then the plank will also tip over:



This is because the larger lever arm "**x**" makes the left hand weight "**m**" have more influence than the identical weight "**m**" on the right hand side.

Do you feel that these facts are just too simple for anyone to really bother with? Well, they form the basis of devices which can provide real power to do real work, with no need for electronics or batteries.

The following suggestions for practical systems are put forward for you to consider, and if you are interested enough test out. However, if you decide to attempt to build anything shown here, please understand that you do so entirely at your own risk. In simple terms, if you drop a heavy weight on your toe, while other people may well be sympathetic, nobody else is liable or responsible for your injury - you need to be more careful in the future ! Let me stress it again, this document is for information purposes only.

The Deflected-Weight Generator of Mikhail Dmitriev.

Mikhail is a Russian experimenter who has worked for many years developing and testing gravity-powered devices. His persistence has paid off and he has been very successful. His work is shown on Sterling Allan's web site http://peswiki.com/index.php/Directory:Mikhail_Dmitriev_Gravity_Wheel where there are videos and photographs of several of his prototypes. It is envisaged that large versions which generate 6 to 12 kilowatts of excess power will become available for purchase in 2011. Each of his various designs is based on the principle of having weights attached to a wheel and arranging for those weights to be offset outwards when falling and offset inwards when rising. Because of the different lever arms involved, that gives a force imbalance which causes the wheel to rotate continuously and if the weights are of a considerable size, then the rotation is powerful and can be used to generate electrical energy.

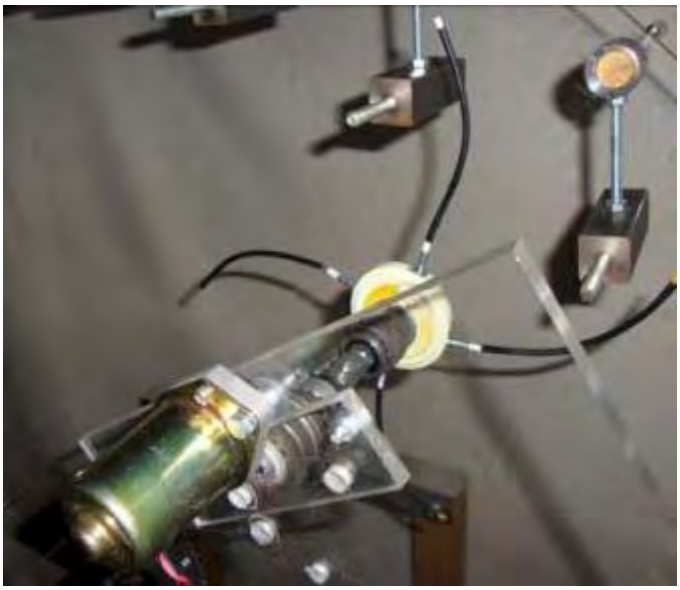
In order to arrange for the weights to be offset as the wheel goes around, each weight is suspended on a pivoted arm:

For the device to operate as required, that suspension arm needs to be moved to (say) the right when falling and be centred or deflected to the right when rising. Mikhail has chosen to use a small amount of electrical power to make this happen, because the energy provided by gravity in turning the wheel far outweighs the small electrical input needed to make the wheel rotate.

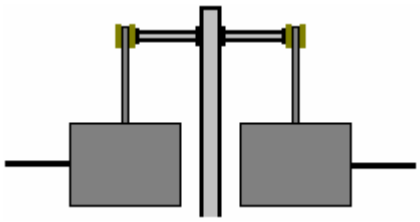
Several mechanisms for making this happen have been tested as you can see from Sterling's presentation. One method is to push the lever arms to the right with a simple rotating disc which has deflector arms attached to it:

After being given the sideways push, each weight stays off centre until it reaches the bottom of it's travel. Please remember that while the weights show here are tiny, a full-size working device will have weights which weight a total of perhaps 130 kilograms and the forces involved are then large. The picture above is a little difficult to make out as the rotating disc is transparent and the support for the rotating arms is also transparent. The horizontal metal arm is there to support the transparent panel on which the 'arms wheel' bearing is mounted.

An alternative method is to use a small motor which drives the arms directly as shown here:

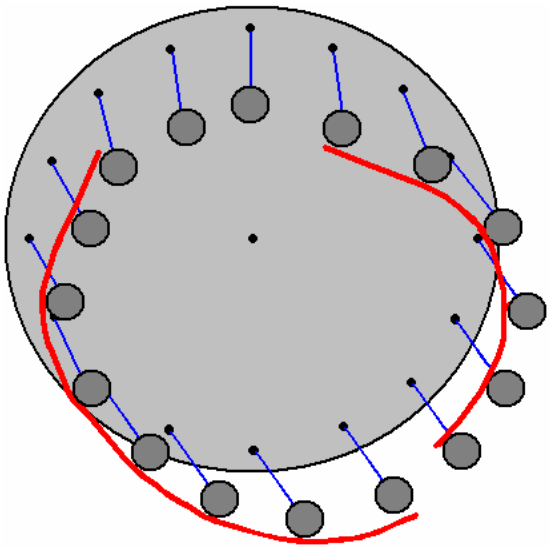


Each weight is held rigidly and so when the motor arm presses against it, the lever arm is pushed out sideways without the weight twisting away from the motor arm. These prototype weights are not heavy, but when a working unit is being built they will have considerable weight, so to get a well balanced arrangement, it might be advisable to have weights on both sides of the wheel so that there is no offset axial load placed on the shaft which supports the wheel:



Mikhail's arrangement works well when it relies on the swinging movement of the weights to keep them off centre during the time when they are falling and you can watch a video of that happening. However, it

makes one wonder if it would not be possible to arrange for this movement without the need for a motor, although using a motor is a very clever and sensible method of ensuring rotational power. Perhaps if two stationary deflectors were used, one to keep the weights out to the right when falling and one to keep them out to the right when rising, a viable system might be created. Perhaps something like this:

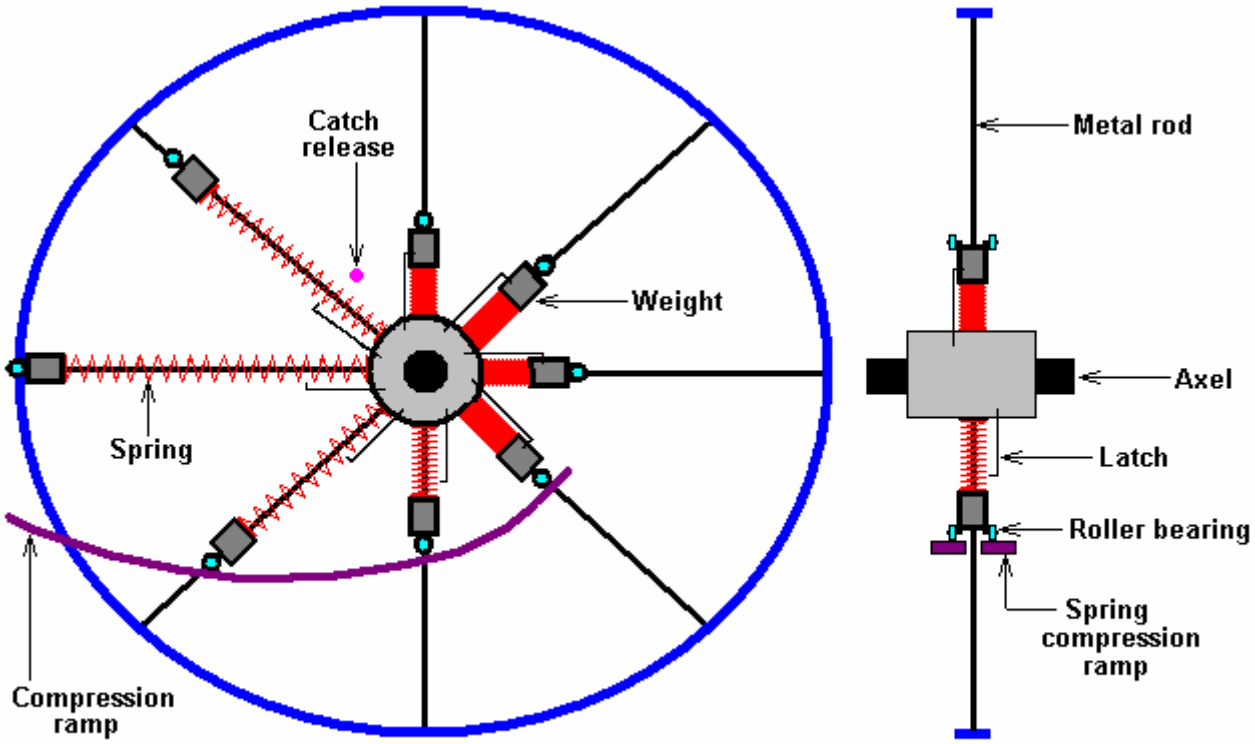


Admittedly, the deflector pieces would have a smoother shape than drawn here, but the principle is shown in spite of the poor quality of the diagram. Where heavy weights are involved, each could have a roller bearing pressing between the weight and the deflector shield in order to minimise friction as the weight slides past. Alternatively, the deflector shield could be faced with powerful magnets opposing similar magnets attached to the weights which would give a no-contact, very low friction movement. A fairly similar idea is part of the next entry from Dale Simpson.

Dale Simpson's Gravity Wheel Suggestion.

The design of gravity-operated machines is an area which has been of considerable interest to a number of people for quite some time now. The design shown here comes from Dale Simpson of the USA. It should be stressed that the following information is published as open-source, gifted to the world and so it cannot be patented by any individual or organisation. Dale's prototype wheel has a diameter of about five feet, utilising weights of a substantial value. The overall strategy is to create excess torque by having the weights slide along metal rods radiating from a central hub somewhat like the spokes of a cart wheel. The objective is to create an asymmetrical situation where the weights are closer to the hub when rising, than they are when falling.

The difficulty with designing a system of this type is to devise a successful and practical mechanism for moving the weights in towards the hub when they are near the lowest point in their elliptical path of movement. Dale's design uses a spring and a latch to assist control the movement of each weight. The key to any mechanical system of this type is the careful choice of components and the precise adjustment of the final mechanism to ensure that operation is exactly as intended. This is a common problem with many free-energy devices as careless replication attempts frequently result in failure, not because the design is at fault, but because the necessary level of skill and care in construction were not met by the person attempting the replication. Here is a sketch of Dale's design:



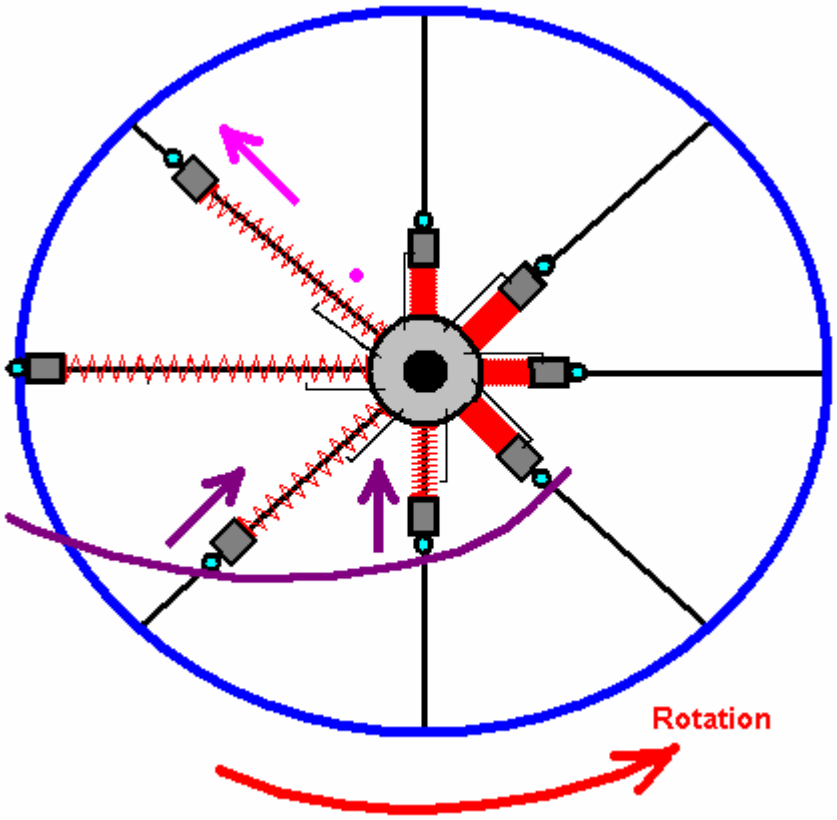
The wheel has an outer rim shown in blue and a central hub shown in grey. Metal spokes shown in black run out radially from the hub to the rim. Eight spokes are shown in this diagram as that number allows greater clarity, but a larger number would probably be beneficial when constructing a wheel of this type.

The wheel as shown, rotates in a counter-clockwise direction. Each weight, shown in dark grey, has a pair of low-friction roller bearings attached to it. There is also a spring, shown in red, between the weight and the hub. When a weight reaches the 8-o'clock position, the roller bearings contact a spring compression ramp, shown in purple. This ramp is formed of two parts, one on each side of the spokes, providing a rolling ramp for each of the two roller bearings. The ramp is formed in a curve which has a constant rate of approach towards the hub of the wheel.

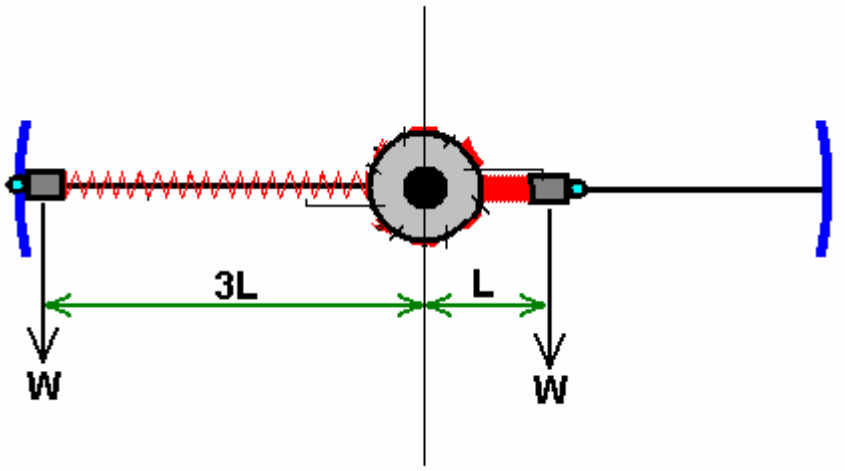
The ramp is positioned so that the spring is fully compressed when the weight has just passed the lowest point in its travel. When the spring is fully compressed, a latch holds it in that position. This holds the weight in close to the hub during its upward movement. The springs are not particularly powerful, and should be just strong enough to be able to push the weight back towards the rim of the wheel when the spoke is at forty five degrees above the horizontal. The "centrifugal force" caused by the rotation assists the spring move the weight outwards at this point. The push from the spring is initiated by the latch being tripped open by the latch release component shown in pink.

The weights have an inward motion towards the hub when they are pushed by the wheel's turning motion which forces the roller bearings upwards along the spring-compression ramp. They have an outward motion along the spokes when the catch holding the spring compressed is released at about the 11-o'clock position. The latch and the release mechanism are both mechanical - no electronics or electrical power supply is needed in this design.

These details are shown in the diagram below:



The question, of course is, will there be enough excess power to make the wheel rotate properly? The quality of construction is definitely a factor as things like the friction between the weights and their spokes needs to be very low. Let us consider the forces involved here:



Take any one weight for this calculation. Any excess rotational energy will be created by the difference between the forces attempting to turn the wheel in a clockwise direction and those forces trying to turn the wheel in a counter-clockwise direction. For the purpose of this discussion, let us assume that we have built the wheel so that the compressed-spring position is one third of the spring-uncompressed position.

As the weights are all of the same value “W”, the see-saw turning effect in a clockwise direction is the weight (“W”) multiplied by it’s distance from the centre of the axle (“L”). That is, W x L.

The turning effect in the counter clockwise direction is the weight (“W”) multiplied by it’s distance from the centre of the axle (“3W”). That is, W x 3 x L.

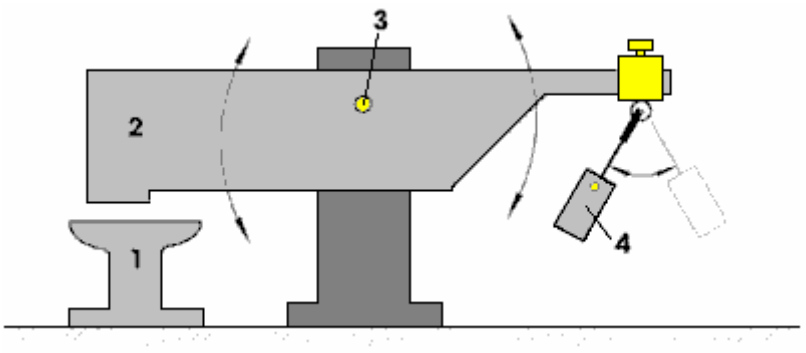
So, with WL pushing it clockwise, and 3WL pushing it counter-clockwise, there is a net force of (3WL - WL), i.e. a net force of 2WL driving the wheel in a counter-clockwise direction. If that force is able to push the weight in towards the hub, compressing the spring and operating the spring latch, then the wheel will be fully operational. There is actually, some additional turning power provided by the weights on the left hand side of the diagram, both above and below the horizontal, as they are a good deal further out from the axle than those with fully compressed and latched springs.

The only way of determining if this design will work correctly is to build one and test it. It would, of course, be possible to have several of these wheels mounted on a single axle shaft to increase the excess output power available from the drive shaft. This design idea has probably the lowest excess power level of all those in this document. The following designs are higher powered and not particularly difficult to construct.

Veljko Milkovic's Pendulum/Lever System.

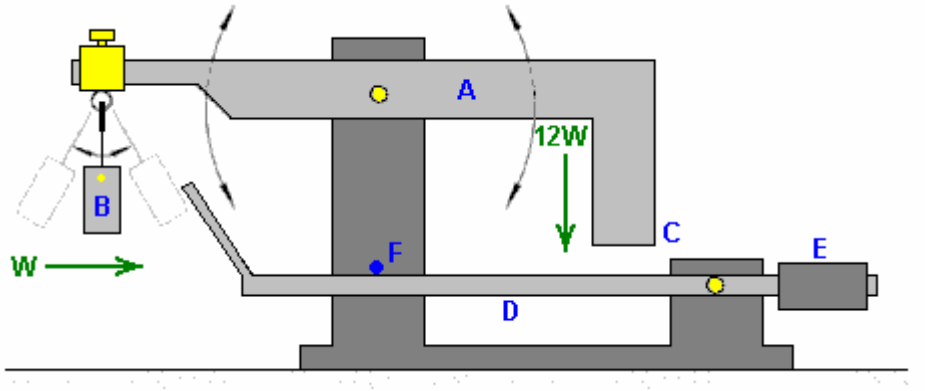
The concept that it is not possible to have excess power from a purely mechanical device is clearly wrong as has recently been shown by Veljko Milkovic at http://www.veljkomilkovic.com/OscilacijeEng.html where his two-stage pendulum/lever system shows a COP = 12 output of excess energy. COP stands for "Coefficient Of Performance" which is a quantity calculated by diving the output power by the input power **which the operator has to provide** to make the system work. Please note that we are talking about power levels and not efficiency. It is not possible to have a system efficiency greater than 100% and it is almost impossible to achieve that 100% level.

Here is Veljko's diagram of his very successful lever / pendulum system:



Here, the beam **2** is very much heavier than the pendulum weight **4**. But, when the pendulum is set swinging by a slight push, the beam **2** pounds down on anvil **1** with considerable force, certainly much greater force than was needed to make the pendulum swing.

As there is excess energy, there appears to be no reason why it should not be made self-sustaining by feeding back some of the excess energy to maintain the movement. A very simple modification to do this could be:

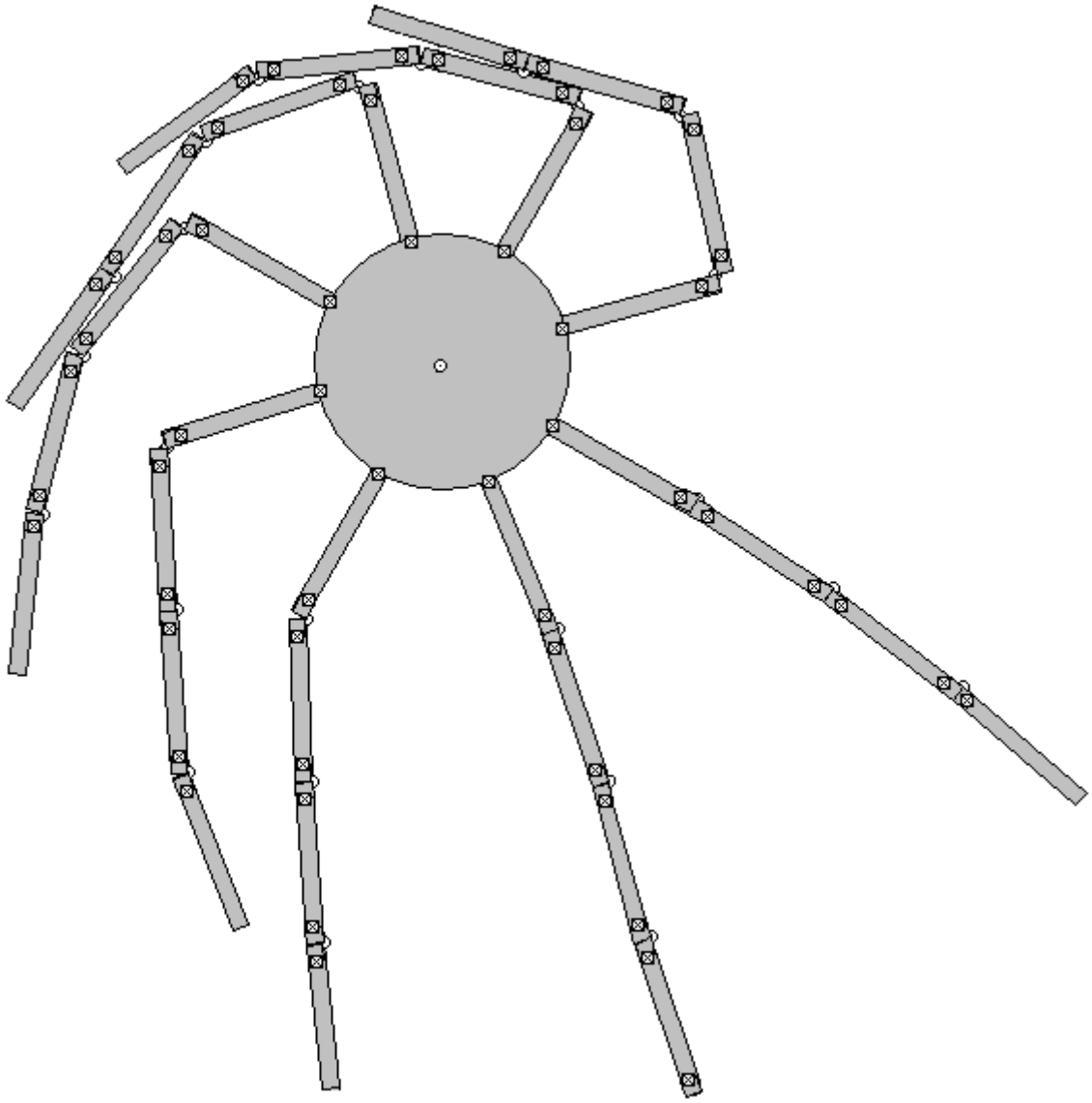


Here, the main beam **A**, is exactly balanced when weight **B** is hanging motionless in it's "at-rest" position. When weight **B** is set swinging, it causes beam **A** to oscillate, providing much greater power at point **C** due to the much greater mass of beam **A**. If an additional, lightweight beam **D** is provided and counterbalanced by weight **E**, so that it has a very light upward pressure on its movement stop **F**, then the operation should be self-sustaining.

For this, the positions are adjusted so that when point **C** moves to its lowest point, it just nudges beam **D** slightly downwards. At this moment in time, weight **B** is at its closest to point **C** and about to start swinging away to the left again. Beam **D** being nudged downwards causes its tip to push weight **B** just enough to maintain its swinging. If weight **B** has a mass of "**W**" then point **C** of beam **A** has a downward thrust of 12**W** on Veljko's working model. As the energy required to move beam **D** slightly is quite small, the majority of the 12**W** thrust remains for doing additional useful work such as operating a pump.

Amr Al-Hossary's Patent.

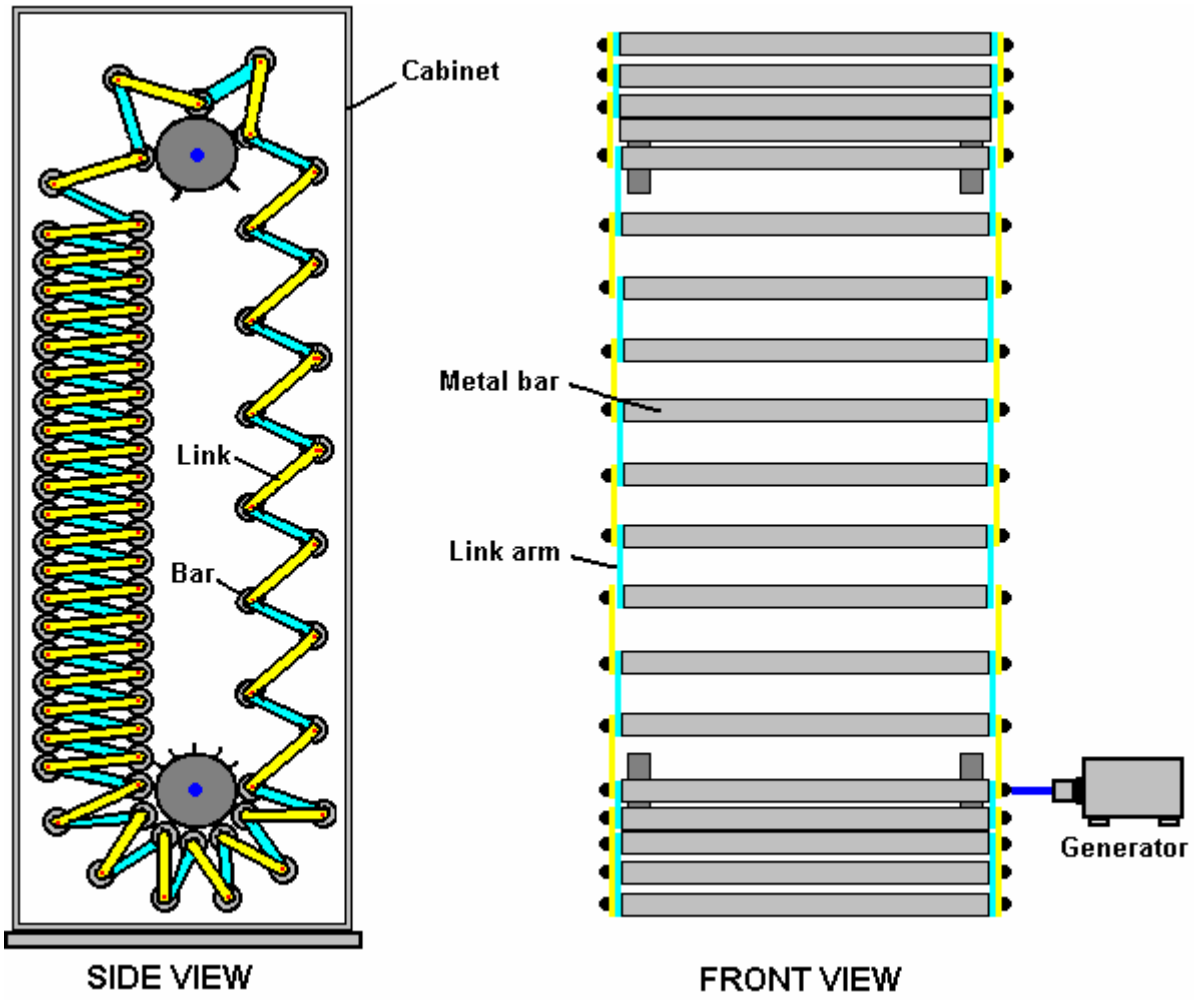
If that is the case, then the design should perhaps be modified along the lines of the patent application of Amr Al-Hossary where the hinged plates are attached directly to the rotor. That way, the plate lever arm definitely comes into play:



Here, the hinged arms or plates will also give an impact impulse when they open out to their full extent and that provides additional turning power. However, the imbalance between the two side is not a major amount and so this design is not likely to provide a large amount of torque for driving external loads.

Murilo Luciano's Gravity Chain Suggestion.

Murilo Luciano of Brazil, has devised a very clever, gravity-operated power device which he has named the "Avalanche-drive". Again, this design cannot be patented as Murilo has gifted it to the world as a royalty-free design which anybody can make. This device continuously places more weights on one side of a drive shaft to give an unbalanced arrangement. This is done by placing expandable links between the weights. The links operate in a scissors-like mode which open up when the weights are rising, and contract when the weights are falling:



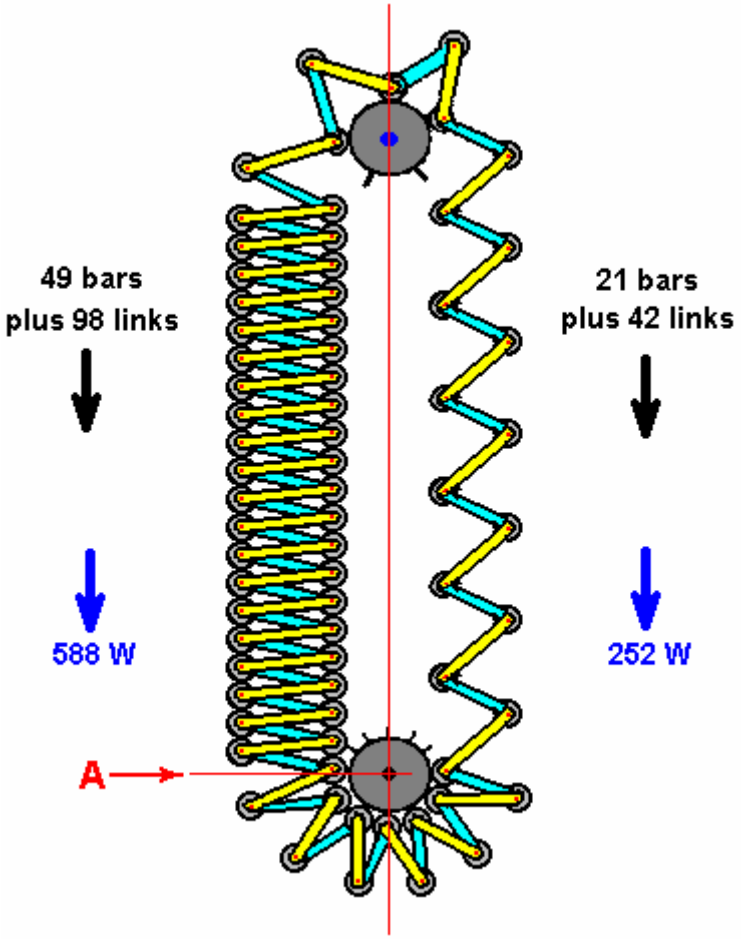
In the arrangement shown here, the weights are shown as steel bars. The design is scaleable in both height, width and the mass and number of weights. In the rough sketch above, the practical details of controlling the position of the bars and co-ordinating the rotation of the two support shafts are not shown in order to clarify the movement. In practice, the two shafts are linked with a pair of toothed sprockets and a chain. Two sets of vertical guides are also needed to control the position of the bars when they are in-between the four sprockets which connect them to the drive shafts, and as they go around the sprocket wheels.

In the sketch, there are 79 bar weights. This arrangement controls these so that there are always 21 on the rising side and 56 on the falling side (two being dead-centre). The resulting weight imbalance is substantial. If we take the situation where each of the linking bars weighs one tenth as much as one of the bar weights, then if we call the weight of one link "W", the rising side has 252 of these "W" units trying to turn the sprockets in a clockwise direction while 588 of the "W" units are trying to turn the sprockets in an counter-clockwise direction. This is a continuous imbalance of 336 of the "W" units in the counter-clockwise direction, and that is a substantial amount. If an arrangement can be implemented where the links open up fully, then the imbalance would be 558 of the "W" units (a 66% improvement) and the level arm difference would be substantial.

There is one other feature, which has not been taken into account in this calculation, and that is the lever arm at which these weights operate. On the falling side, the centre of the weights is further out from the axis

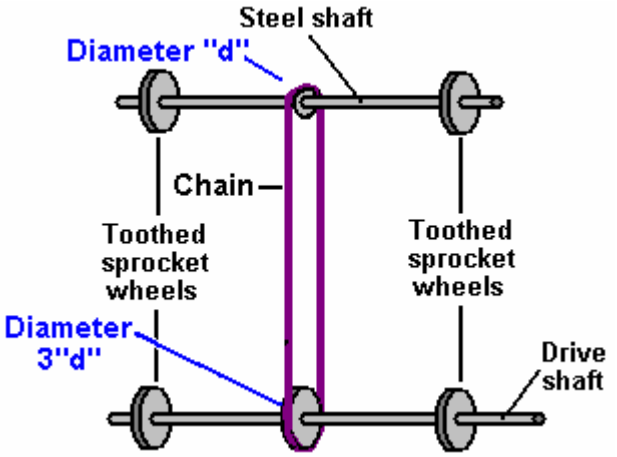
of the drive shafts because the link arms are nearly horizontal. On the rising side, the links are spread out over a lesser horizontal distance, so their centre is not as far out from their supporting sprocket. This difference in distance, increases the turning power of the output shafts. In the sketch above, an electrical generator is shown attached directly to one output shaft. That is to make the diagram easier to understand, as in practice, the generator link is likely to be a geared one so that the generator shaft spins much faster than the output shaft rotates. This is not certain as Murilo envisages that this device will operate so rapidly that some form of braking may be needed. The generator will provide braking, especially when supplying a heavy electrical load.

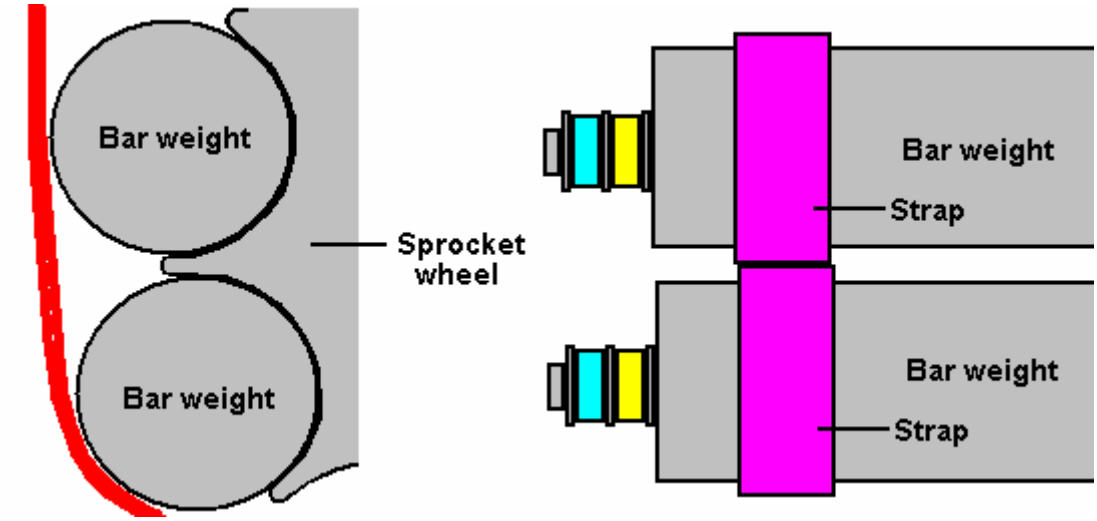
This diagram shows how the two side of the device have the unbalanced loading which causes a counter-clockwise rotation:



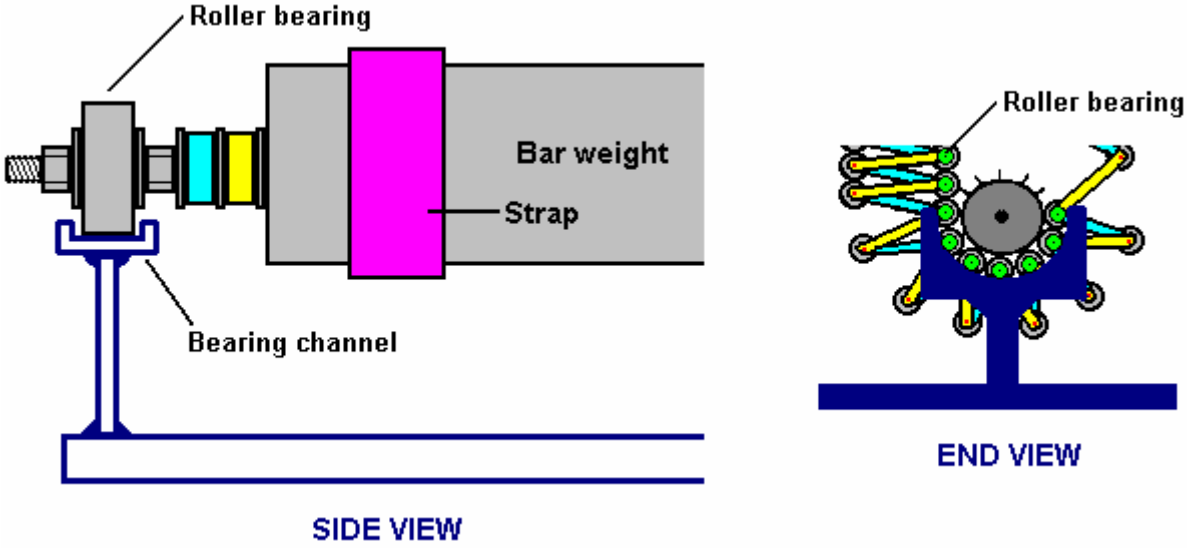
The diagrams shown above are intended to show the principles of how this device operates and so for clarity, the practical control mechanisms have not been shown. There are of course, many different ways of controlling the operation and ensuring that it works as required. One of the easiest building methods is to link the two shafts together using a chain and sprocket wheels. It is essential to have the same number of bar weights passing over the upper sprocket wheels as pass under the lower sprocket wheels. On the upper sprocket wheels, the bars are spread out, say, three times as far apart than they are on the lower sprocket wheels, so the upper sprockets need to rotate three times as fast as the lower ones. This is arranged by using a lower drive-chain sprocket wheel which has three times the diameter of the upper one.

The driving force provided by the weight imbalance of the two columns of rod weights needs to be applied to the lower sprocket wheels at point "**A**" in the diagram above. For this to happen, there has to be a mechanical connection between the stack of bar weights and the sprocket wheel. This can be done in different ways. In the above concept diagrams, this link has been shown as a sprocket tooth or alternatively, a simple pin projection from the sprocket wheel. This is not a good choice as it involves a considerable amount of machining and there would need to be some method to prevent the bar rotating slightly and getting out of alignment with the sprocket wheel. A much better option is to put spacers between the bar weights and have the sprocket teeth insert between the bars so that no bar slots are needed and accurate bar positioning is no longer essential. This arrangement is shown below:





The description up to here has not mentioned the most important practical aspects of the design. It is now time to consider the rising side of the device. To control the expanded section of the chain, and to ensure that it feeds correctly on to the upper sprocket wheels, the gap between successive bar weights must be controlled.



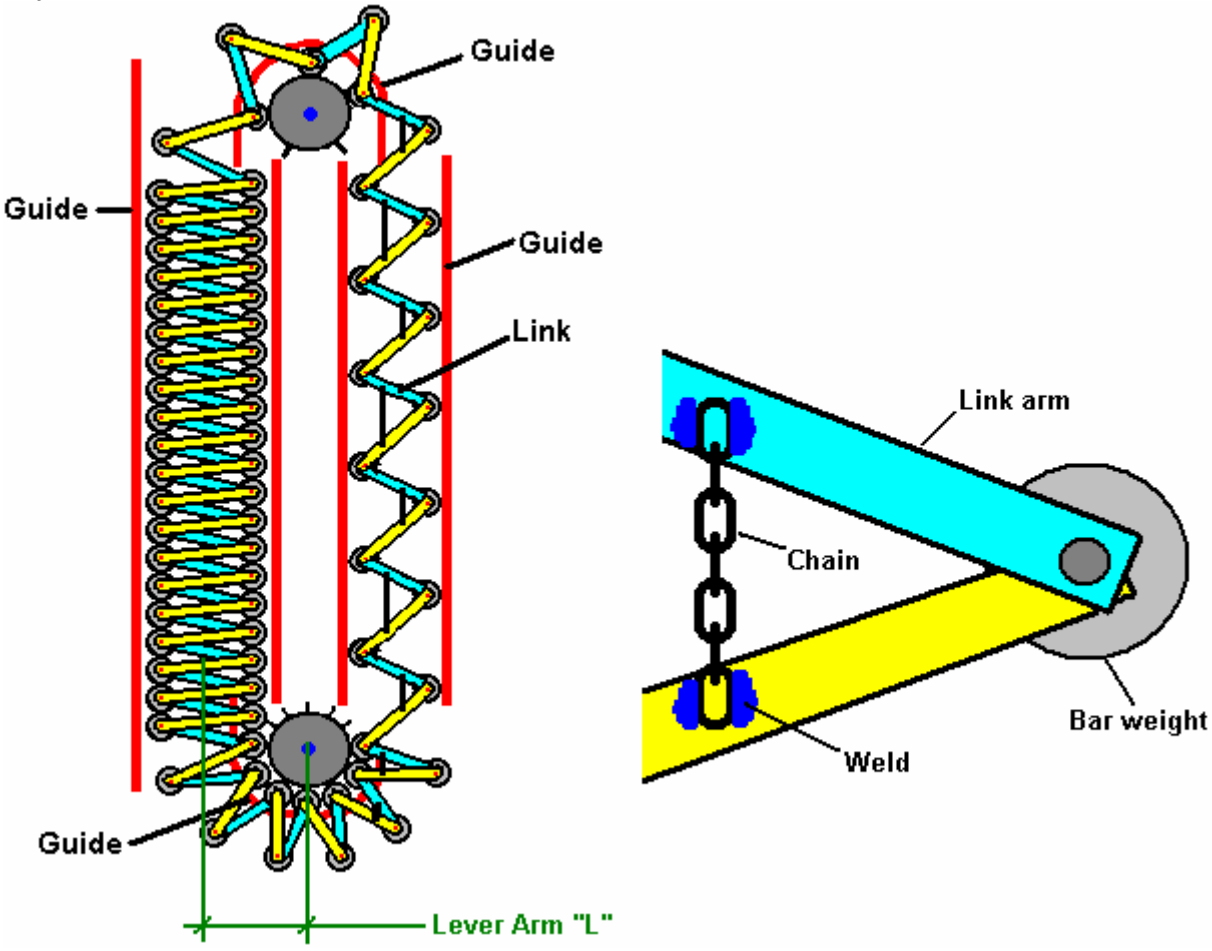
A guiding channel can be used, as shown here, and standard ball-bearings or roller-bearings can be attached to the ends of the weights by using threaded rod (or a bolt with the head inside the weight) and locking nuts.

In the example shown here, which is of course, just one option out of hundreds of different implementations, the bars on the rising side are three times as far apart as those on the falling side. This means that on the upper sprocket wheels, only every third tooth will connect with a bar weight. This is shown in the following diagram. However, if the linked weights were left to their own devices, then the rising side bars would hang down in one straight line. While that would be optimum for drive power, Murilo does not envisage that as a practical option, presumably due to the movement of the links as the bar weights move over their highest point. In my opinion, that arrangement is quite possible to implement reliably provided that the length of the links is selected to match the sprocket distance exactly, however, Murilo's method is shown here.

Murilo's method is to use additional restraining links between the weights. The objective here is to make sure that when the weights spread out on their upward journey, that they take up positions exactly three bar widths apart, and so feed correctly on to the teeth of the upper sprocket wheel. These links need to close up on the falling side and open up on the rising side. They could be fabricated from short lengths of chain or from slotted metal strips with a pin sliding along the slot.

Whichever method is chosen, it is important that the links stay clear of the bars and do not prevent the bars stacking closely together on the falling side as that would prevent them seating correctly on the teeth of the lower sprocket wheels. The easiest precision option for the home constructor is using chain, where two bar weights are positioned on the upper sprocket wheel to give the exact spacing, and the tensioned chain is welded in position, as shown below. Placing the chain inside a plastic tube causes it to take up an "A" shape standing outwards from the links when they move into their closed position. This keeps the chains from getting between the link bars. In addition, the chains are staggered from one pair of link bars to the next, as shown below, as an additional measure to keep the operation both reliable and quiet..

In the diagram below, only a few of these restraining links are shown in order to keep the diagram as simple as possible. It is not a good choice to make the upper bar sprocket wheels three times larger than the lower sprocket wheels as this would force both the rising and falling sections of chain out of the vertical, which in turn introduces friction against the guides. The central 1:3 gearing is needed to make sure that the chains on the rising side are fully stretched and the spacing of the bar weights matches the upper sprocket spacing exactly.

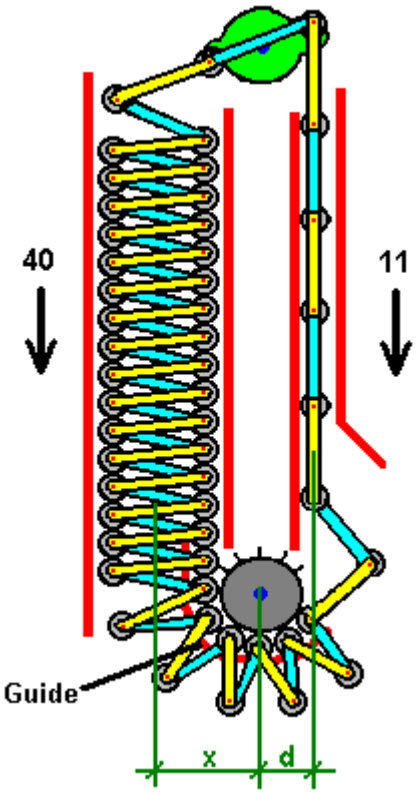


The diagrams have not shown the supporting framework which holds the axles in place and maintains the unit in a vertical position, as this framing is not specialised in any way, and there are many acceptable

variations. A sensible precaution is to enclose the device in an upright box cabinet to make sure that there is no chance of anything getting caught in the rapidly moving mechanism. This is an impressive design of Murilo's, who recommends that in the implementation shown above, that the links shown in blue are made 5% longer than those shown in yellow, as this improves the weight distribution and drive of the lower sprocket wheel..

A washing machine has a maximum power requirement of 2.25 kW and in the UK a suitable 3.5 kW alternator costs £225 and needs to be spun at 3,000 rpm for full output.

While the above description covers Murilo's main design, it is possible to advance the design further, raising its efficiency in the process as well as reducing the construction effort needed to build it. For this version, the main components remain the same, with the upper axle geared to the lower axle as before and the upper axle rotating faster than the lower one. The main difference is that on the rising side, the chain opens up completely. This does away with the need for the chain links, moves the rising weights much closer in and reduces the number of rising weights:



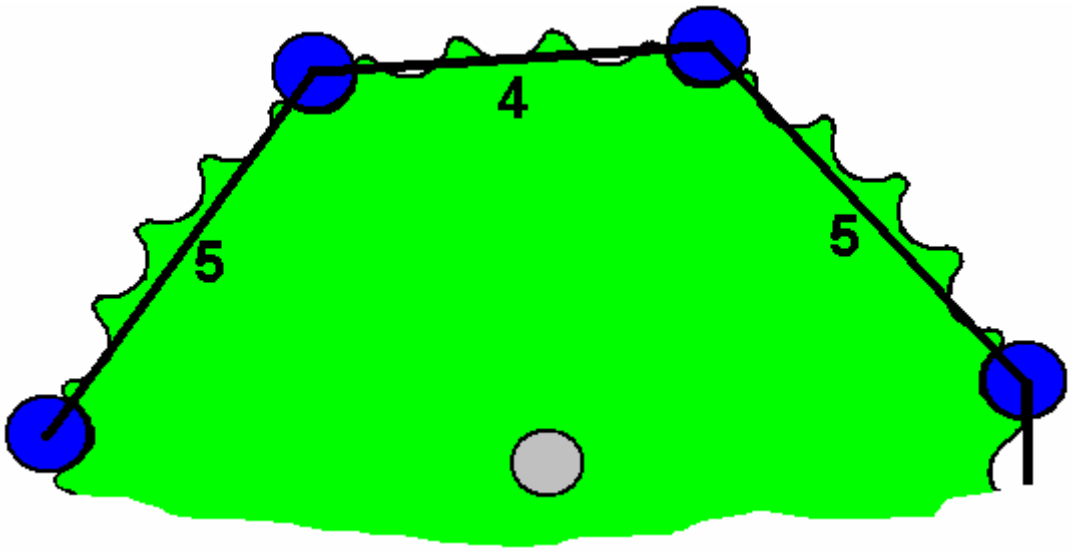
With a reduced number of weights in the diagram above, the weight imbalance is a very substantial 40:11 ratio with the massive advantage of a substantially reduced lever arm "**d**" which is much smaller than the lever arm "**x**" of the falling weights. This is a major imbalance, giving 40x pulling the axle in a counter-clockwise direction and only 11d opposing that movement.

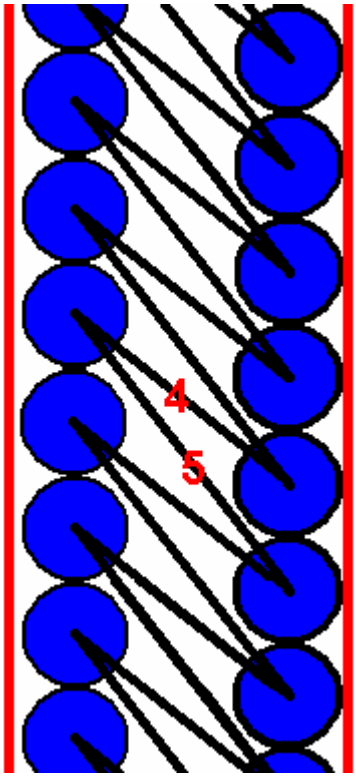
In the description so far, it has been assumed that all components will be made of metal. This is not necessarily the best choice. Firstly, metal moving against metal does make a noise, so guides made robustly of thick plastic or other similar material would be a good choice for the guides for the weights.

The weights themselves could equally well be made from strong plastic piping filled with sand, lead pellets, concrete or any other convenient heavy material. The pipes would then have strong end caps capable of holding the pivots for the links. The sprocket wheels themselves could well be made from thick plastic material which would give a quieter operation and which could be bolted to the power take-off shaft with a bolt placed right through the axle.

Most of the dimensions are not critical. Increasing the diameter of the lower sprocket wheel will increase the power of the output axle but will lower its speed. Adding more weights will increase both the output power and to a lesser degree, the speed, but will increase the overall size of the unit and its overall weight and cost. Making each weight heavier will raise the output power, or reduce the overall size if the weight is contained in fewer weights. Increasing the length of the links means fewer weights on the rising side but will require larger sprocket wheels.

It is not necessary to have all the links the same size. If the lengths are chosen carefully and the indentations in the upper sprocket wheel cover the entire circumference, then every second link can be one indentation shorter which tips the weights into a more compact and effective column on the falling side:





With this arrangement, the outer weights, shown here on the left, press down very firmly on the inside column of weights, making a compact group. If using plastic pipes with concrete then the hinge arrangement for the rods can be very simple, with a bolt set in the concrete as shown below.

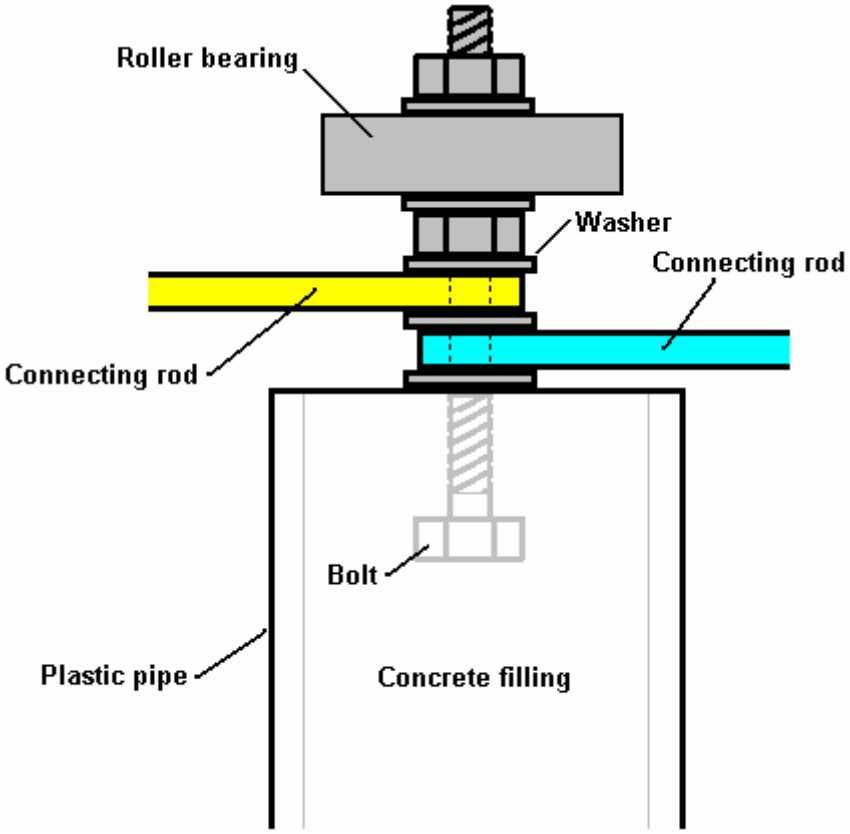
The rods, washers and bolt can be supported on a thin, rigid strip placed across the top of the pipe. When the concrete has gone solid, the strip is removed and the gap produced by its removal then allows free movement of the rods. If this technique is used, then the bar weights are cast in two steps, with a tightly fitting disc pushed part way up inside the pipe so that one end can be filled while the other end remains open and ready for the completion of the other end.

One advantage of using plastic pipes is that if the sprocket wheels are made from a tough high-density plastic material, such as is used for food chopping boards, and the weight guides are also made from tough

plastic, then there should be no metal-upon-metal noise produced during operation, if the bolt holes in the connecting rods are a good fit for the bolts used.

The concrete or mortar used as a filling can be made wet and pliable, since mechanical strength is not an issue here, and a filling with no voids in it is desirable. Even low quality concrete (caused by more water than absolutely necessary) would be more than adequate for this purpose.

The arrangement at the ends of a concrete-filled plastic pipe bar weight could be constructed like this:



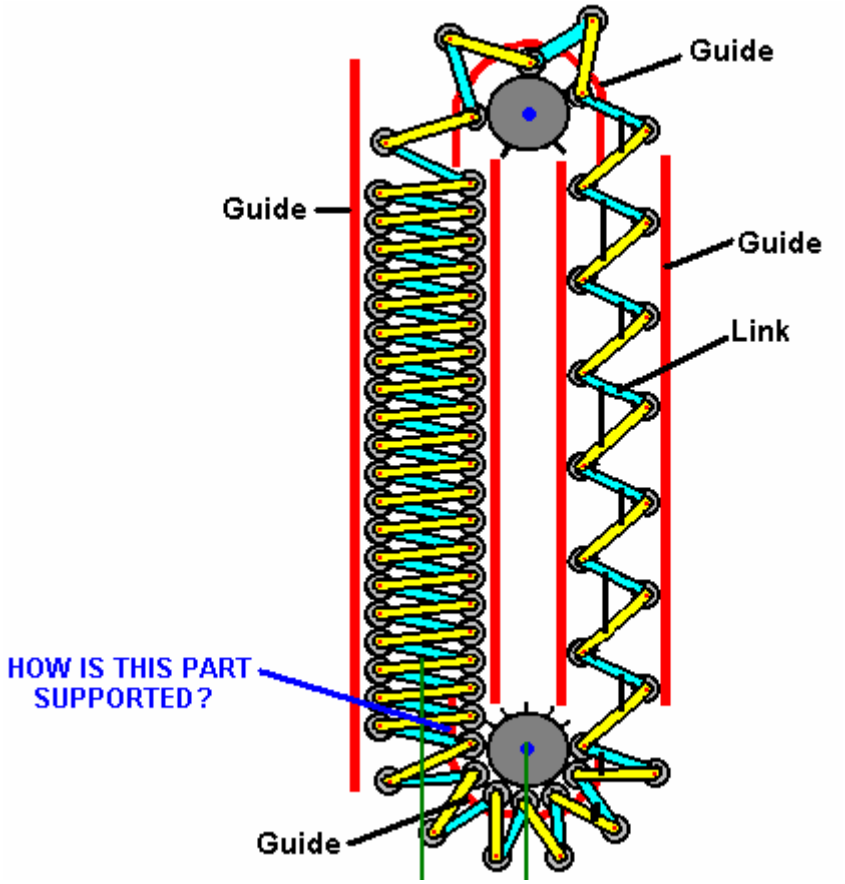
There is a very strong inclination when building a device to make it operate smoothly. Where excess energy is being drawn from the gravity field, the reverse is necessary, with a jerky operation being the optimum. Remember that the extra energy only occurs during the duration of the impulses causing the jerks. It follows then, that in an ideal situation, any device of this type should be driven by a rapid series of strong impulses. In practice, using a heavy flywheel or any similar component which has a high inertial mass, although a rapid series of sharp pulses is being applied to the component and jerky operation is not visible to the human eye, excess energy is still being “led-out” and made available to do useful work.

One other observation which may be of interest, and that it the feedback from builders of gravity wheels which says that the power output from a gravity wheel is greater if the axle is horizontal and the rotating wheel is aligned exactly with magnetic East-West.

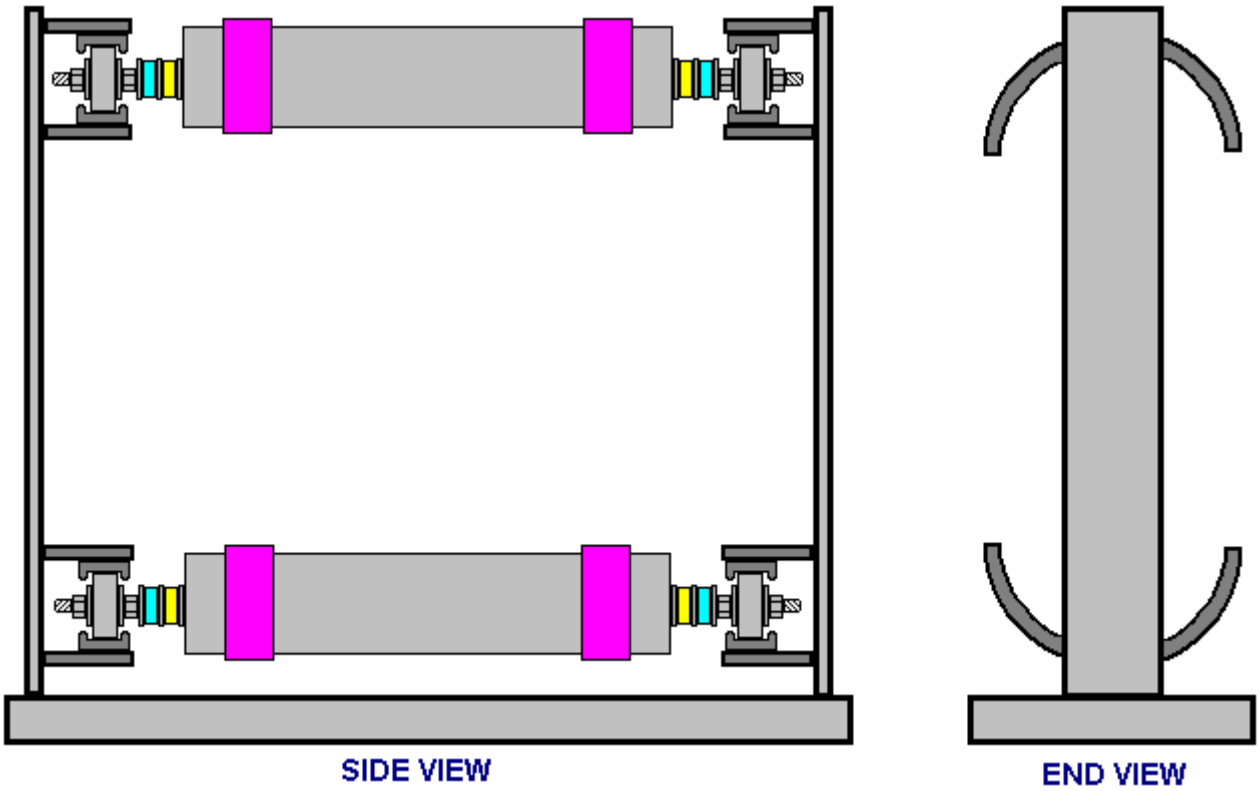
A Practical Construction Query

I have just been asked about the practical issues of mounting the guiding components for the weights. I must apologise for not making it clear that the diagrams in this description are intended to show the overall methods of operation, rather than being a direct construction arrangement. There will be several ways of constructing an implementation of each device. Here is one suggestion for a practical construction method for the gravity chain device.

The query was as follows:

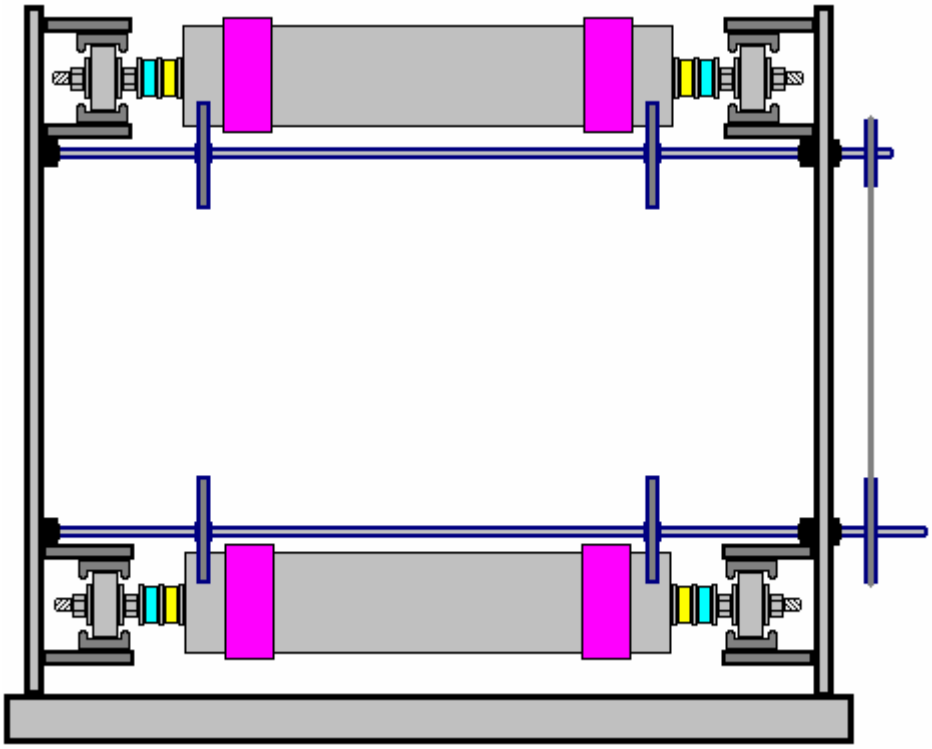


It is pointed out that the lower guide as shown, can't be supported from inside as the weights sweep through the area which would be used for that support. Also, it can't be supported from outside as the connecting rods have to move through the area where that support would be positioned. A solution has been suggested where the lower guide is supported by a strap from the upper guide, the strap running between the inner and outer weights. That is a solution which could work, but it introduces significant unnecessary friction. An alternative method is to place the guides outside the moving weights as shown here:



This method provides a low-friction channel for the roller-bearings to move along. This controls the position of the weights very accurately and the end walls also provide the supports for the axles which synchronise the positions of the weights and provide gearing between the axles if that is required. For clarity, just two of the many weights are shown and the overall proportions distorted so that the diagram will fit on the page.

With the axles, it might look like this:



Here, the axle shafts are geared together outside the end wall and either a chain or a belt drive used. The lower shaft allows a power take-off. The ratio of the diameters of the pulley wheels or sprocket wheels dictates the relative rates of rotation of the two shafts.

Other Gravity-Powered Designs.

Sterling Allan reports on **Bobby Amarasingam**'s design which has 12 kilowatts of excess power: http://pesn.com/2010/12/04/9501738_British_gravity_motor_generates_12_kilowatts/

Also reported by Sterling is the **Smith-Caggiano** gravity/momentum/centrifugal-force generator design. The report is at: http://www.peswiki.com/index.php/Directory:OBM-Global%27s_Angular_Force_Generator

Another of Sterling's reports is on the **Chalkalis Gravity Wheel** which can be seen at:
http://peswiki.com/index.php/OS:_F._M._Chalkalis_Gravity_Wheel

Buoyancy

While we are aware of buoyancy being used to convert wave power into electricity, we seem to neglect the idea of using the very powerful buoyancy forces (caused by gravity) as a direct tool at locations away from the sea. This is definitely a mistake because serious levels of power can be generated from such a system. One such system is:

The "Hidro" Self-Powered Generator of James Kwok.

This design demonstrates yet again, the practical nature of drawing large quantities of energy from the local environment. Commercial versions are offered in three standard sizes: 50 kilowatt, 250 kilowatt and 1 megawatt and licensing partners are being sought. This generator which James has designed can be seen at the Panacea-bocaf.org web site at http://panacea-bocaf.org/hidrofreeenergysystem.htm and on James' own web site at http://www.hidroonline.com/ both of which have video clips explaining how the design works. The method is based on different pressures at different depths of water, gravity, and on the buoyancy of air-

filled containers. The system does not rely on wind, weather, sunlight, fuel of any type, and it can operate all the time, day or night, without causing any kind of pollution or hazard. This particular design calls for a water-filled structure of some height, a source of compressed air and a pulley system, and without wishing to be in any way critical, it seems rather more complicated than it needs to be. If, unlike James, you have not done the mathematics for the system, you would assume that the amount of power generated by a system like this would be less than the amount of power needed to make it operate. However, that is definitely very far from reality as considerable excess power is gained through the natural forces of the local environment which make the system operate. Part of the patent application which James made is shown here:

US 2010/0307149 A1 **Date: 9th Dec. 2010** **Inventor: James Kwok**

HYDRODYNAMIC ENERGY GENERATION SYSTEM

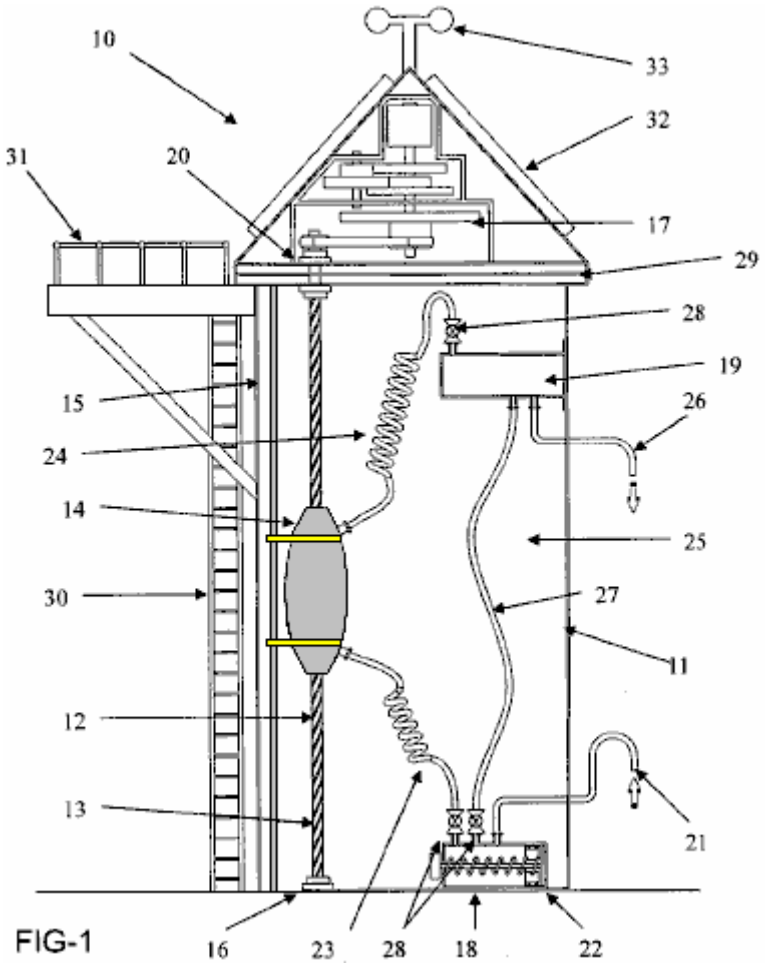


FIG-1

Fig.1 is a cross-sectional view of an embodiment of the energy generation system of the present invention. Here, the energy generation system **10** comprises a vessel **11** in the form of a water tank and a shaft **12** which can rotate about it's longitudinal axis. The shaft **12** is provided with a helical screw groove **13** and is connected at it's lower end to a bearing **16** which allows it to rotate freely about its longitudinal axis.

The upper end of the shaft is connected to a generator **17** which is a flywheel system. The rotational energy of shaft **12** may be transferred to the generator through a ratchet-cog system **20**. A buoyant inflatable capsule **14** is provided along with its guiding mechanism **15** which is in the form of a wire or pole to assist in the smooth vertical movement of buoy **14**.

There is a first air reservoir **18** located in a lower portion of the vessel **11** and a second air reservoir **19** located in an upper portion of the vessel **11**. The first reservoir **18** draws air from the atmosphere, in through air intake port **21**. Once the pressure in the first reservoir has reached a predetermined value, a piston **22** is actuated, forcing air through hose **23** into the buoyant capsule **14**, which, when inflated, begins to move upwards through water tank **11**, as the buoy **14** has become less dense than the fluid **25** (such as fresh

water or saltwater) in tank **11**. This in turn causes rotation of shaft **12**, and activation of the power generator **17**, thereby generating power.

When buoy **14** reaches the upper limit of its travel, the air in the buoy may be forced to flow through a second hose **24** and into the second air reservoir **19**. When air is removed from the buoy it moves downwards through vessel **11** under gravity and with the assistance of ballast (not shown). The downward movement of buoy **14,** causes rotation of the shaft **12**, which drives the generator **17**, thereby generating power.

Air stored in the second reservoir **19** may be vented to the atmosphere through a vent **26** if the pressure in the second reservoir **19** becomes too high. Alternatively, air may flow from the second reservoir **19** into the first reservoir **18** through a third hose **27** so that less air must be drawn into the first reservoir **18** when buoy **14** reaches the lower limit of its travel and must once again be inflated with air from the first reservoir **18**.

The hoses **23**, **24** and **27** are provided with non-return valves **28** to ensure that air will flow in only one direction through the system **10**. Vessel **11** may be provided with ventilation **29** as required and it may also be provided with access stairs **30** and an access platform **31** so that maintenance may be carried out as required. The system may also be provided with a solar energy collection device **32** to generate at least a portion of the energy required to drive piston **22** and the non-return valves **28**. Energy produced by the solar energy collection device **32** may also be used to power a light or beacon **33** to indicate the location of the system **10**.

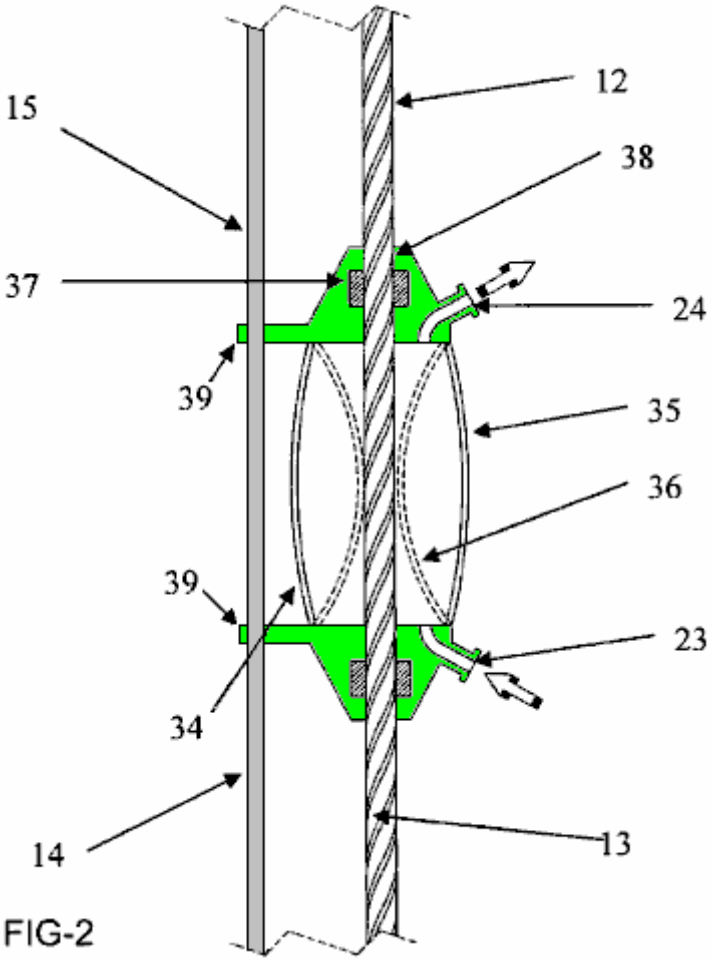


FIG-2

Fig.2 shows one arrangement for buoy **14** comprising an inflatable capsule **34**. This figure illustrates the shape of the walls of the inflatable capsule **34** when inflated **35** and when deflated **36**. Air passes into capsule **34** through hose **23** and exits from the capsule through hose **24**.

The buoy **14** also has a sleeve **37** attached to it. This sleeve has projections which engage with the helical groove **13** of shaft **12**, thereby causing rotation of the shaft when the buoy moves relative to shaft **12**. Sleeve **37** is provided with ballast **38**, such as stainless steel weights that assist in the downward movement of the buoy when it is deflated.

Buoy **14** is attached to a guide pole **15** and the buoy has a pair of arms **39** which slide on the guiding pole **15** and assist in the smooth vertical movement of the buoy.

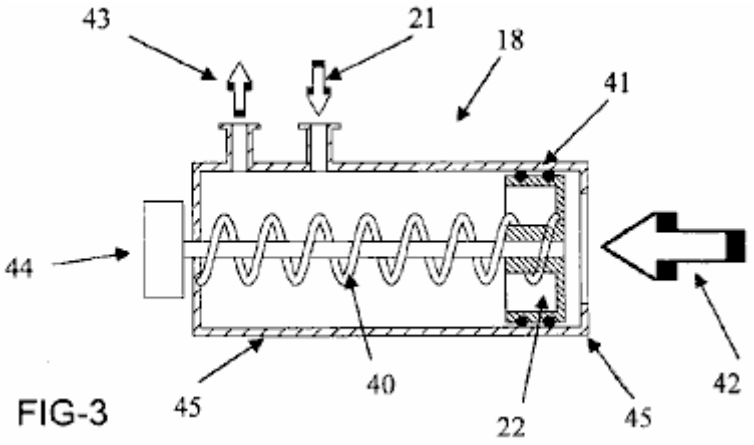


FIG-3

Fig.3 shows one version of the first air reservoir **18**. Air is drawn into reservoir **18** through air intake **21**. The reservoir includes a piston **22** associated with a spring **40**, the piston **22** being provided with seals **41** to prevent leakage of air.

When pressure, such as hydrostatic pressure, is applied in the direction of arrow **42**, the piston moves to the left of the reservoir **18** compressing spring **40** and forcing air out through outlet **43**. A motor **44** is provided to reverse the movement of the piston **22**. Reservoir **18** may be fixed to the floor of the vessel.

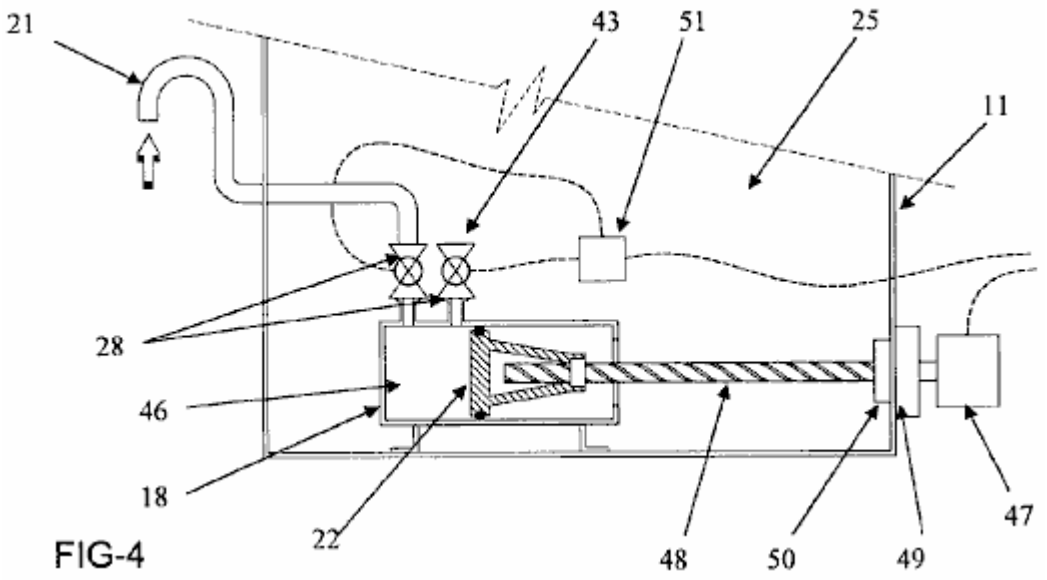


FIG-4

An alternative construction of the first air reservoir **18** is shown in **Fig.4**. In this embodiment, reservoir **18** is housed within a vessel **11** containing a fluid **25**. Air enters reservoir **18** through air intake **21** and is held in a chamber **46**. The reservoir has a piston **22** and the movement of the piston **22** towards the left of the reservoir **18** forces air in the chamber **46** out through air outlet **43**.

Piston **22** is driven by motor **47** which rotates the helically-grooved shaft **48**. The motor is linked to the shaft by a ratchet and cog mechanism **49**, which is provided with a spring loaded seal **50** on the inner surface of vessel **11**. An actuator **51,** may be used to control the opening and closing of non-return valves **28** as well as the actuation of motor **47**.

Fig.5 illustrates a cross-sectional view of an energy generation system according to one of the embodiments of the present invention:

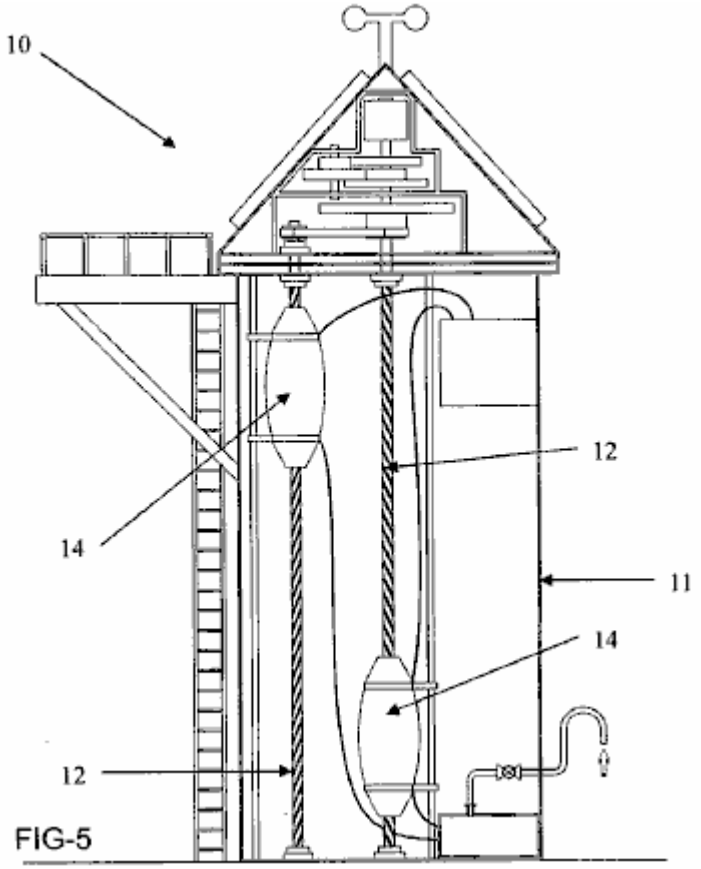


FIG-5

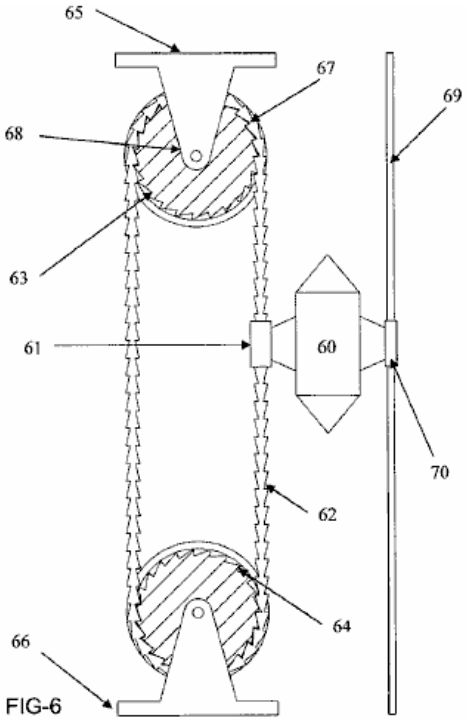


FIG-6

Fig.5 shows an embodiment where which a pair of buoys **14** are present. Each buoy is associated with its own shaft **12** and may move up and down inside vessel **11** independent of one another.

In **Fig.6**, an alternative embodiment of the present invention is illustrated, where the buoy **60** has a connecting method **61** in the form of a cylindrical sleeve through which a guide chain **62** passes. Chain **62** is provided in an endless loop and is located on an upper tracking device **63** and a lower tracking device **64**, both of which are pulleys. The upper pulley **63** may be fixed to an upper wall (not shown) of a vessel (not shown) via a bracket **65**, while the lower pulley **64** may be fixed to a lower wall (not shown) of a vessel (not shown) via a bracket **66**.

The connection mechanism **61** contains ratchets which engage with the links of the chain **62** when buoy **60** moves downwards. Thus, as buoy **60** moves downwards, chain **62** also moves, thereby causing both the upper and lower pulleys to rotate in a clockwise direction. The upper and lower **64** pulleys have a series of indentations **67** corresponding to the shape of the links of the chain **62**. In this way, the chain **62** sits in the indentations **67** and grips the tracking device (**63**, **64**), thereby ensuring that the tracking device (**63**, **64**) rotates.
In the embodiment of the invention illustrated in **Fig.6**, a work shaft **68** is associated with the upper pulley **63** such that rotation of the upper pulley results in rotation of the work shaft **68**. The work shaft **68** is located substantially perpendicular to the direction of travel of the buoy **60**. The work shaft drives a generator to produce power.

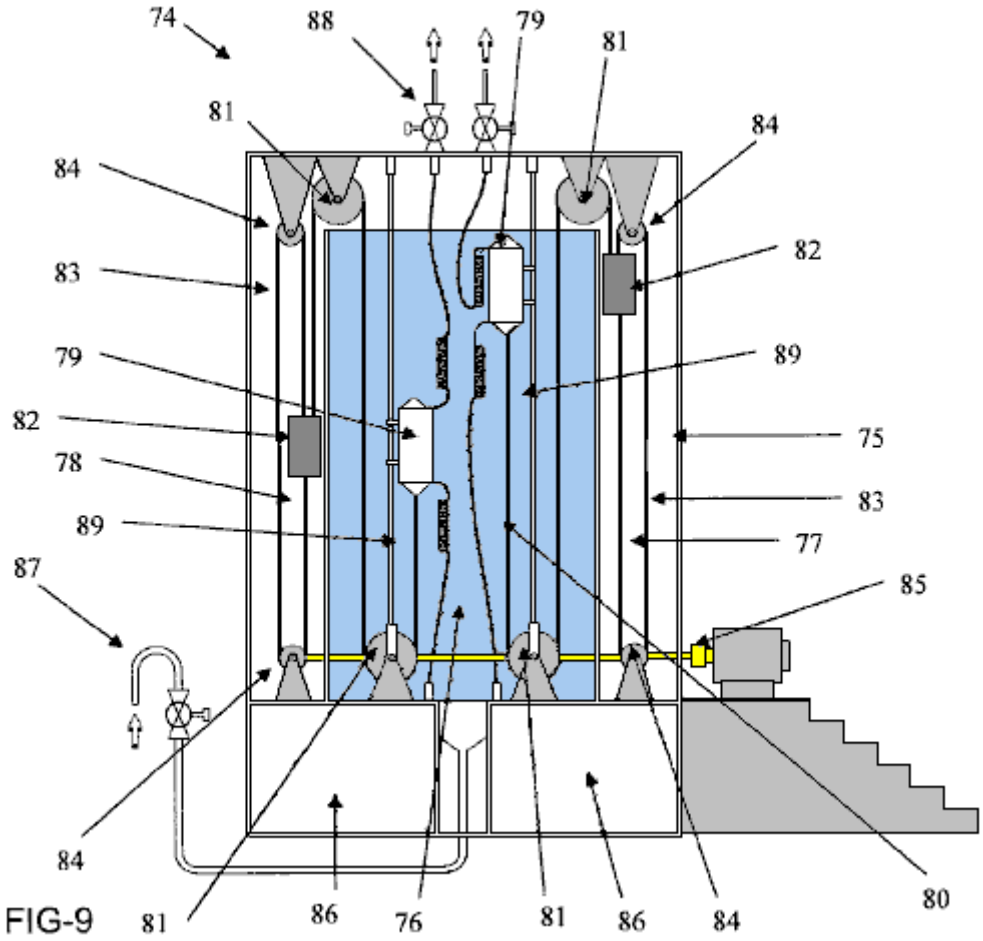


FIG-9

Fig.9 shows an alternative embodiment of this energy generation system **74**. The system is comprised of a vessel **75** having a fluid-filled "wet" compartment **76** and one or more "dry" compartments (in this case, a pair of dry compartments **77**, **78**) with no liquid in them. These dry compartments may be fabricated from any suitable material, such as, concrete, steel, fibreglass, plastic or any combination of materials.

The system also has a pair of buoys **79** each with a deflatable bladder-like construction. The buoys have guide rails **89** which ensure that the buoys move smoothly up and down inside the vessel **75**.

In this embodiment of the invention, air reservoirs **86** are located in the base of the vessel **75**. Air enters the reservoirs **86** through inlet **87**, while air exiting from the buoy **79** is vented through valves **88**. The vented air may either be expelled to the atmosphere or recycled to the reservoirs **86**.

Each of the buoys is designed to be connected to one end of a chain or rope **80**. A weight **82** is connected to the other end of the chain or rope **80**. The chain or rope **80** has a series of pulleys **81** such that when the buoy is inflated and filled with air, the buoyancy is greater than the weight **82** and so the buoy rises in the vessel.

When the buoy **79** is deflated, weight **82** is heavier than the buoyancy and so the buoy sinks in vessel **75**. In the embodiment illustrated here, the weights **82** are located in the dry compartments **77**,**78**. There are several reasons for this, including that, by locating the weights **82** in the dry compartments **77**,**78**, the velocity of the weights **82** in the downward direction is increased, and therefore an increase in the energy produced by the system **74** is experienced.

The weights **82** are associated with second ropes or chains **83**, such that vertical movement of the weights **82** results in the rotation of the second ropes or chains **83** around a pair of sprockets **84**. Rotational energy generated by the rotation of the second ropes or chains **83** is transferred to a power generation device **85** (such as a turbine or the like) in order to generate power (e.g. electrical power)

***

In spite of its mechanical complexity, the Hidro design is offered as a commercial generator with tens of kilowatts of excess power, indicating that buoyancy is a significant method of generating power, based on the fact that water is hundreds of times heavier than air. Due to its weight, movement in water is slow but can be very powerful. The helical groove method of converting the vertical movement of the floats into rotational power is used because of this as it has a very high ratio between shaft turns and movement along the shaft. This can be understood when you consider the fact that a complete revolution of the shaft is caused by the float moving up just one step to the next thread position directly above. The turns ratio for the complete float movement is determined by the angle of the groove cut into the drive shaft.

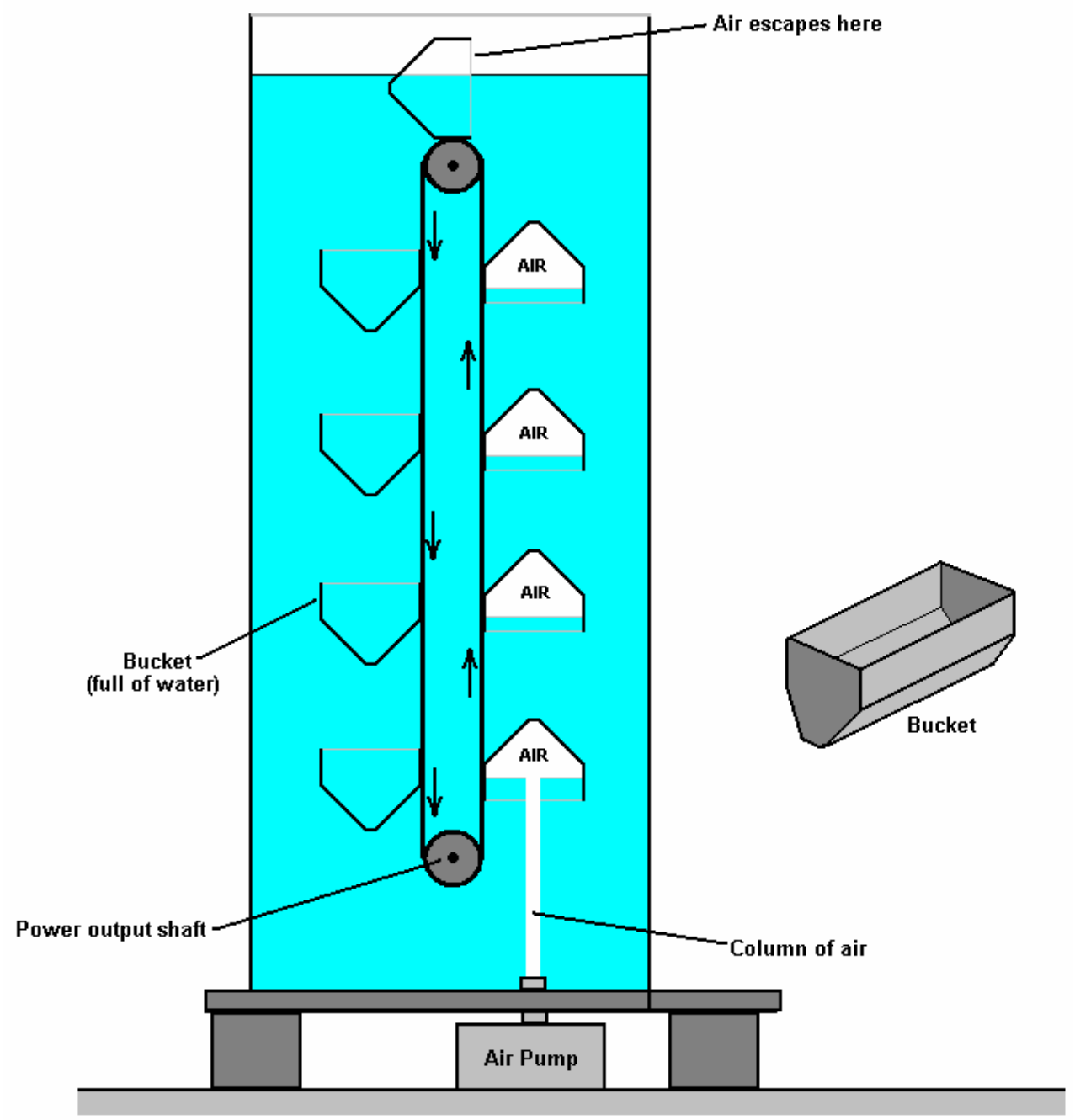
One other thing which needs to be considered for such a project is the weight of the overall structure when filled with water. The overall weight is liable to be many tons and so the footing underneath the generator needs to be very robust. Also, while compressed air is mentioned, giving the impression of cylinders of compressed air or gas, For continuous operation one would expect an air pump to be used. Whether or not an air pump is used, the diameter of the air hoses needs to be considered. Most people think that a gas can flow along a pipe or tube very easily. That is not the case. If you want to get a feel for the constriction caused by a pipe, then take a one metre length of 6 mm diameter plastic tube and try blowing through it. No significant amount of air will pass through the tube even if you blow very hard. The web site http://www.engineeringtoolbox.com/natural-gas-pipe-sizing-d_826.html shows this table:

Capacity of Pipe (MBH ≈ CFH)							
Pipe Size (inches)		Pipe Length (feet)					
Nominal	Inside diameter	10	20	40	80	150	300
1/2	0.622	120	85	60	42	31	22
3/4	0.824	272	192	136	96	70	50
1	1.049	547	387	273	193	141	100
1 1/4	1.380	1200	849	600	424	310	219
1 1/2	1.610	1860	1316	930	658	480	340
2	2.067	3759	2658	1880	1330	971	686
2 1/2	2.469	6169	4362	3084	2189	1593	1126
3	3.068	11225	7938	5613	3969	2898	2049
4	4.026	23479	16602	11740	8301	6062	4287
5	5.047	42945	30367	21473	15183	11088	7841
6	6.065	69671	49265	34836	24632	17989	12720
8	7.981	141832	100290	70916	50145	36621	25895

Notice the major difference in carrying capacity of any of these pipes with just the change from a 10-foot (3 metre) length to a modest 20-foot (6 Metre) length, and those lengths are the sort of lengths needed for many applications. Also, look at the figures for, say, the 0.5 inch (nominal) diameter pipe. With just a 10-foot length, it would take a full two minutes to pump just one cubic foot of air through it. It follows then, that pipes of considerably larger diameter are needed for a project like the ‘Hidro’.

It is possible to construct a much more simple version of the ‘Hidro’, perhaps like this:

A Simple Buoyancy Power Generator



A simple hydraulic, buoyancy-powered generator can be constructed, with two or more horizontal, rotating shafts submerged in water in such a way that they are effectively positioned one above the other. Each shaft has one, and preferably two or more sprocket wheels mounted on it. Each of these sprocket wheels engages with a continuous chain loop which also engages with the sprocket wheel which is positioned vertically above it. These vertical chain loops form a belt-style support for a series of identical buckets. On one side of the vertical belt the buckets have their open face upwards and on the other side the bucket openings are facing downwards. An air pump is positioned directly underneath the set of buckets which have the bucket openings facing downwards. The air pump generates an upward-moving stream of air which collects in the rising buckets, displacing the water filling the bucket. This results in a powerful upward thrust caused by the buoyancy of that bucket, and the thrust causes the bucket to move upwards, rotating both horizontal shafts and bringing another water-filled bucket into position above the air pump. A gearing system transfers the rotation torque thus produced, to a generator which produces electricity for general purpose uses.

This is a generator whose input shaft is rotated through buoyancy caused by air-filled containers submerged in a tank of water or some other suitable heavy liquid. Continuous, powerful rotation of the generator shaft is produced through the use of one or more conventional, commercially available air pumps. An air pump is

used to fill a series of containers which are open at one end and which are attached to what is effectively a belt arrangement created by two strong chain-link loops which mesh with sprocket wheels mounted on two shafts, either, or both of which can be utilised for the extraction of useful power, preferably for driving an electricity generator but not necessarily limited to that function as any powerful torque has many useful applications.

Objectives are to provide a power generation system which is very simple in form and which can be understood, operated and maintained by people with minimal training. Also, a system which uses components which are already readily available, thus avoiding significant manufacturing costs, and one which operates without the need for any kind of complex mechanism or high-precision equipment and which can operate with a wide range of commercially available products.

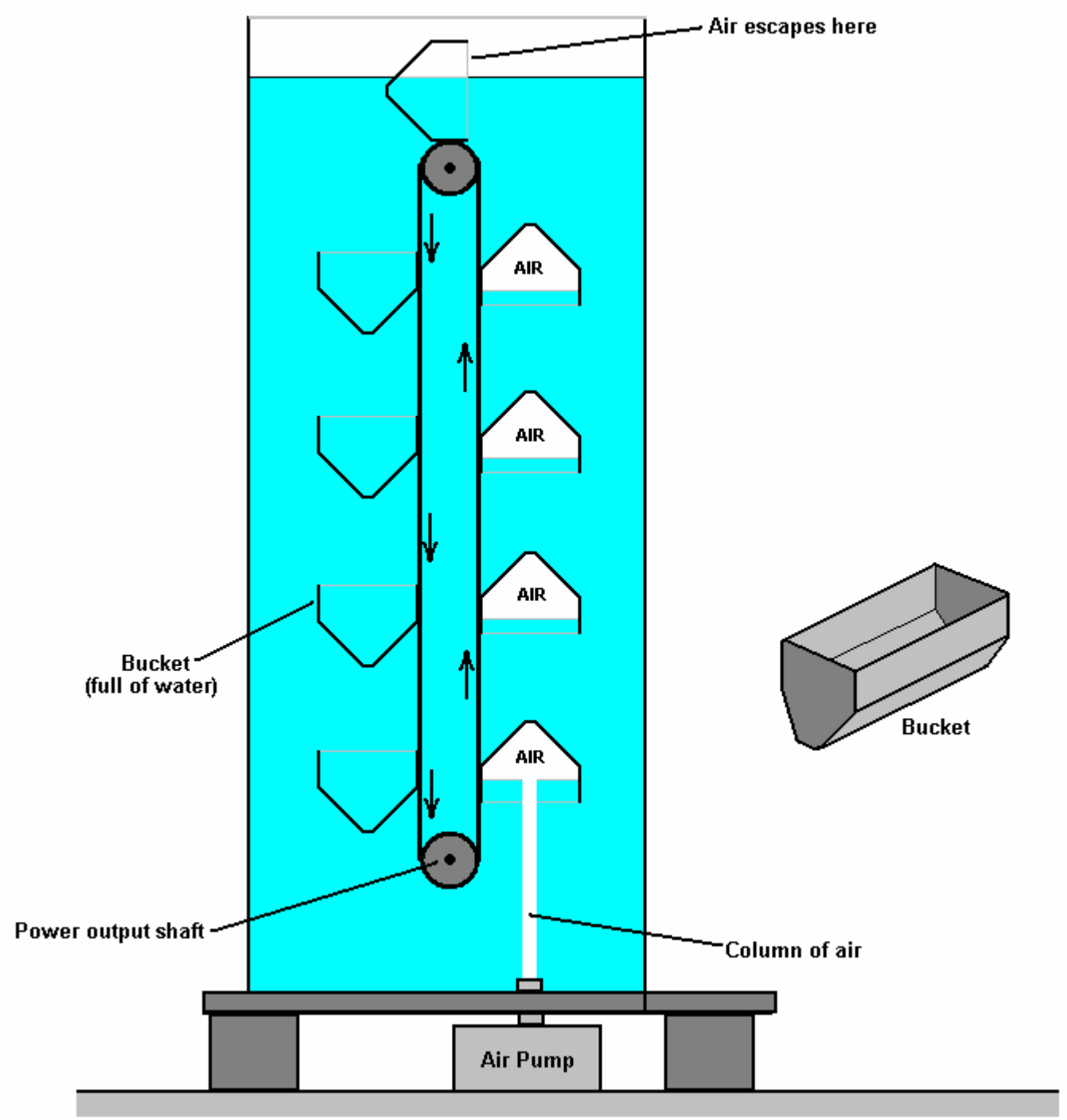


Fig.1, is a simplified partial schematic cross-sectional view showing the main components of the generator as seen from one end.

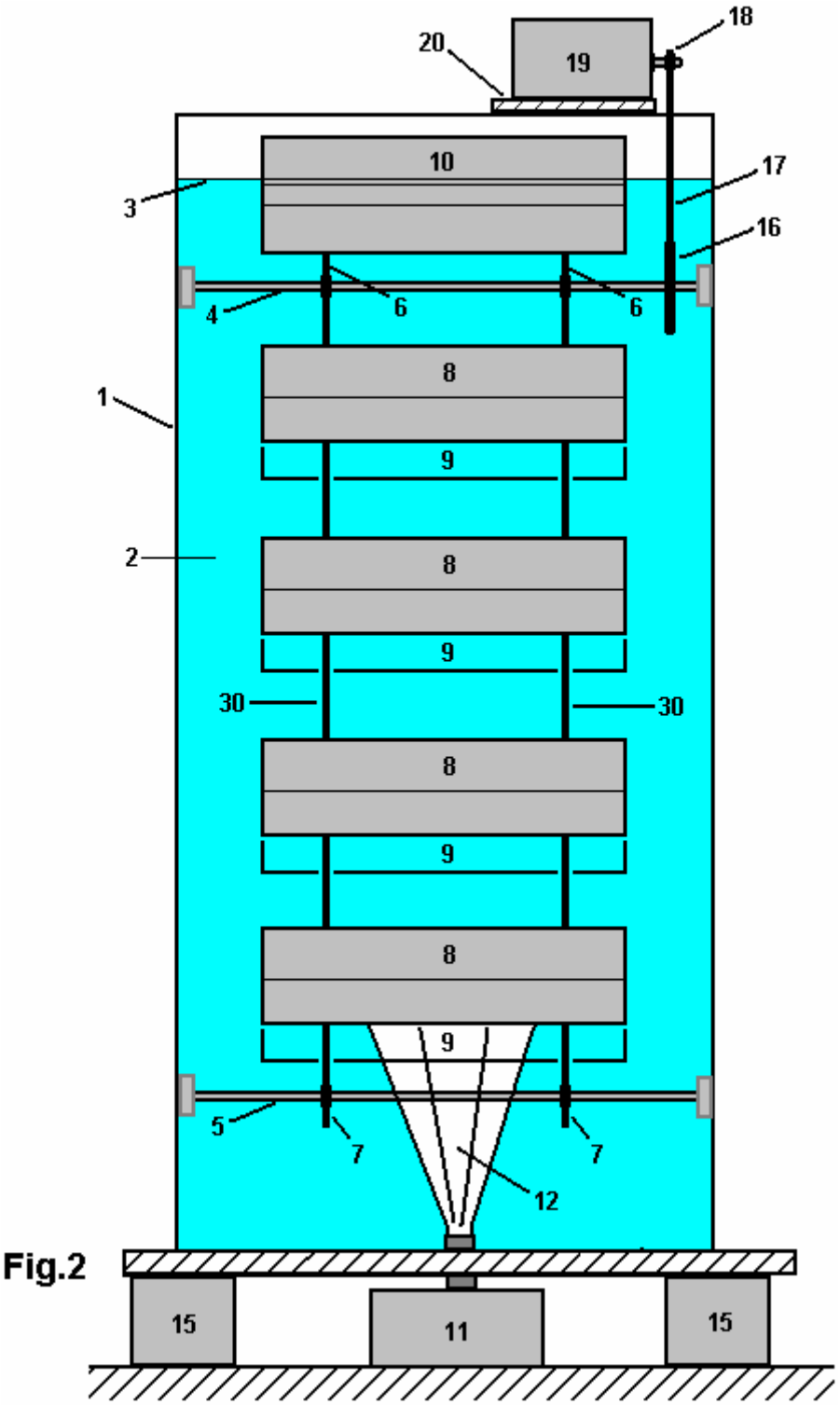


Fig.2

Fig.2, is a cross-sectional conceptual schematic view showing the front view of the generator in its most simple form.

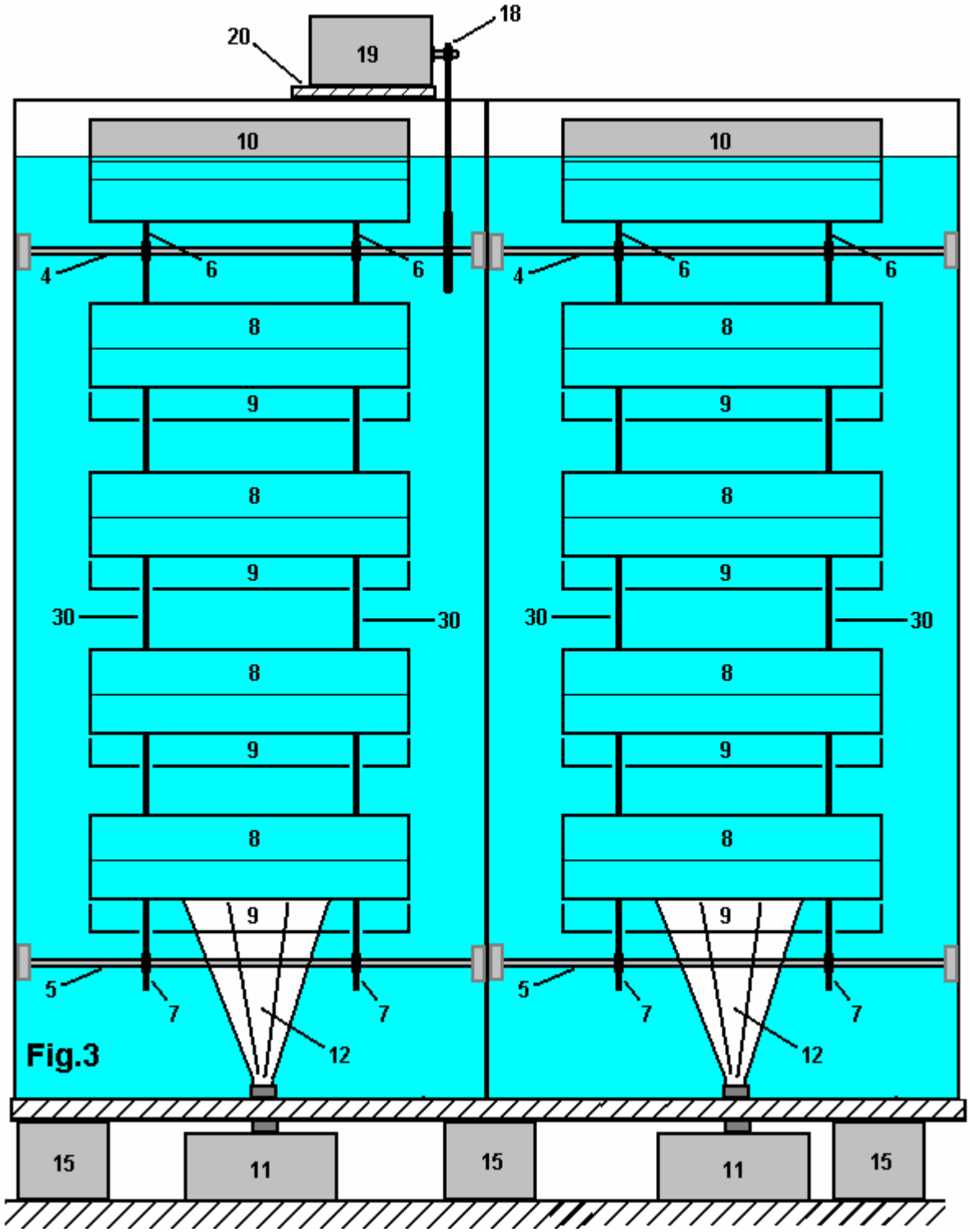


Fig.3, is a cross-sectional conceptual schematic view showing the front view of the generator where more than one set of buckets is used.

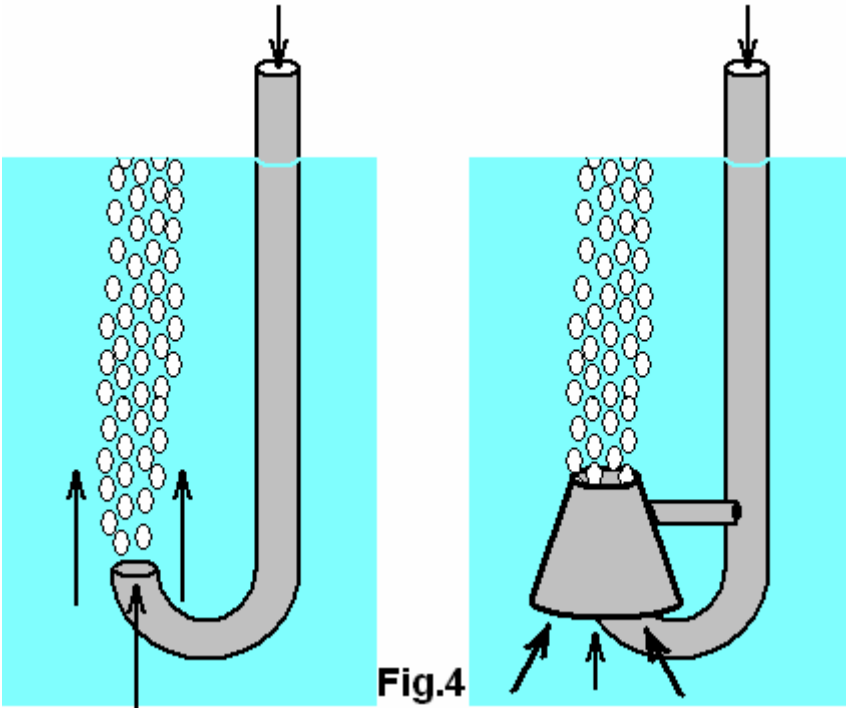


Fig.4

Fig.4, is a perspective conceptual view showing arrangements for a simplified air-feed system which operates from above the tank.

Fig.1, illustrates the overall concept of the generator in its most simple form where lightweight rigid buckets are used to capture the rising air from the air pump. In this Figure, a water tank **1**, holds water or other suitable liquid **2**. The surface of the liquid **3**, is indicated to illustrate the fact that a bucket **10**, which is in the process of turning over at the top of its orbital motion, is positioned so that one edge of the bucket is clear of the surface of the water, which allows the air which was trapped inside the bucket to escape into the atmosphere and the water fill the entire bucket causing only a very minor turbulence when doing so. This is a desirable, but not essential feature as the air trapped in any bucket will escape upwards as soon as the bucket starts its downward movement, positioning its open end upwards, although this causes unnecessary turbulence inside the tank. One possible bucket shape is shown in perspective view, but many different bucket shapes may be used, including flexible membrane types or alternatively, hinged-plate types which have very much reduced resistance to moving through the water when in their collapsed state during their downward movement.

The buckets **8, 9** and **10**, are attached to two strong chains **30**, which mesh with the upper sprocket wheel **6**, mounted on the upper axle **4**, and the lower sprocket wheel **7**, which is mounted on the lower axle **5**. Although it is not visible in **Fig.1**, there are two upper sprocket wheels **6**, two lower sprocket wheels **7**, and two chain loops **30,** although these can be seen in **Fig.2**.

The tank is supported on a robust plate **14**, which itself is supported by a series of pillars **15** which rest on a secure footing **16**, providing operating space underneath the tank for the installation and maintenance of the air pumping equipment. As fresh water weighs 1000 Kg per cubic metre, the weight of the operational generator system is substantial and so this must be allowed for when assessing the footing needed to support the tank and it's contents. While a thin-wall tank is shown in **Fig.1**, many different forms of tank may be utilised, including earth bank and plastic membrane styles, or resurfaced abandoned well shafts. The tank of **Fig.1** presumes that the lower axle **5** is taken out through the wall of tank **1**, using an arrangement similar to that used for the drive shafts which power the screws of ships and other power vessels. While an arrangement of that type provides a drive shaft which is conveniently close to the ground, the much more simple arrangement shown in **Fig.2** where the output power is taken off using the very simple chain and sprocket wheel method utilised for the bucket supports (chain **30,** and sprocket wheels **6** and **7**). In general, the more simple and straightforward any design is, the better it works in practice and the lower any maintenance costs become.

Referring again to **Fig.1**, when activated, air pump **11** produces a stream of air **12**, which flows rapidly upwards. This stream of air **12**, once established, does not have to push against the head of water as immediately above the nozzle of the pump is a rapidly rising column of air, sustained both by the exit velocity from pump **11** and the natural upward movement caused by the relative weights of water and air (as water is

several hundred times heavier than air). This column of air would normally flow straight upwards in calm water, but should it be found that turbulence in the water tends to push the rising air away from its vertical path, baffles can be placed around the pump and positioned so that the air stream is forced to stay within the same section of water taken up by the rising buckets.

The rising air enters the lowest of the rising buckets and collects in it, forcing the water out of the open bottom of the bucket. If the rising bucket is not completely filled with air before the next bucket moves between it and the air pump, the trapped air will expand as the bucket rises and the water pressure reduces due to the lesser depth. Any one bucket with a substantial amount of air in it will create a very significant upward force due to buoyancy, air being about one thousand times lighter than water.

Each bucket on the rising side adds to that upward force and consequently, the chains **30** need considerable strength. The weight of the buckets on each side of the chain match and so the main advantage of light buckets is to lower the inertial mass of the moving parts. Movement through the water is relatively slow but this is offset by gearing between the output drive shaft and the generator's input shaft. The power of the system can be increased by adding more buckets in the vertical chain, increasing the water depth accordingly. Other ways of increasing the power include increasing the volume inside each bucket and/or increasing the flow rate produced by the air pump or pumps used. Another simple method is shown in **Fig.3** and discussed below. An alternative to air pumps is to use tanks of a compressed, non-polluting gas, possibly air.

The buckets shown in the various Figures are rigid, very simple shapes, possibly made by a plastic moulding process in order to be cheap, strong, lightweight and permanently water-resistant. There are, of course, many possible variations on this including using rigid hinged plates sealed with a strong flexible membrane, allowing the buckets to fold and become streamlined on their downward path, and opening as soon as they turn to start their upward movement. There are many mechanisms which can provide this movement, but it is a matter of opinion as to whether or not the extreme simplicity of rigid buckets is worth sacrificing

Fig.2 shows a schematic layout of the generator when seen from the side. The same numbers apply to the components already seen in **Fig.1**. The arrangement seen in **Fig.2** is the most simple, basic, single bucket set. The near-side rising buckets **8** obscure the view of the far-side falling buckets **9** and only the lowest part of the falling buckets **9** can be seen in this view. **Fig.1** shows buckets which are some two and a half times longer than they are wide, but this, of course, is just one option among literally thousands of possible proportions. The size and shape of buckets is related to the performance and number of air pumps being used for any one set of buckets and that choice depends on what is available locally at a reasonable price. It would not be unusual for two or three air pumps to be used side by side along the length of the bucket **8** although **Fig.2** only shows a single pump.

Fig.2 also shows a simple method for power take off where a large diameter sprocket wheel **16** is mounted on the upper axle **4**, and driving a much smaller diameter sprocket wheel **18** which is mounted on the drive shaft of the electricity generator **19** which is mounted on plate **20** which is attached securely to the top of tank **1**.

Fig.3 shows one of the possible arrangements for increasing the system power without increasing the depth of water used. Here, the axles **4** and **5** extend far enough to allow another set of buckets to drive them, increasing the torque very substantially. While **Fig.3** shows one extra set of buckets, there is, of course, no reason why there should not be three or more sets of buckets side by side. It should be noted however, that the partitions shown between the bucket sets are not there just to reduce the water swirling but are needed to support the bearings which are essential for the extended axles, since without those, the diameter of the bars used for the axles would have to increase very markedly to avoid unwanted flexing along their length. While the second set of buckets has been shown aligned exactly with the first set, there is an advantage in offsetting them relative to each other so that the output torque is more even with buckets emptying and filling at different points in the bucket cycle.

Fig.4. shows a method for further simplification, where the air is pumped from above the water surface. It is a matter of concern to most people, that the pressure of the head of water above the air pump is a major obstacle to overcome and will be a continuous opposing force during operation of the generator. If air is being injected from underneath the tank, then initially, that pressure head has to be overcome. However, once the air flow is established, a vertical cigar-shaped area of water vortex is established by the rising air stream. This three-dimensional annular vortex negates the water head in the small area immediately above the air nozzle, and almost sucks the air out of the pump, after the initial introduction of the air has been accomplished.

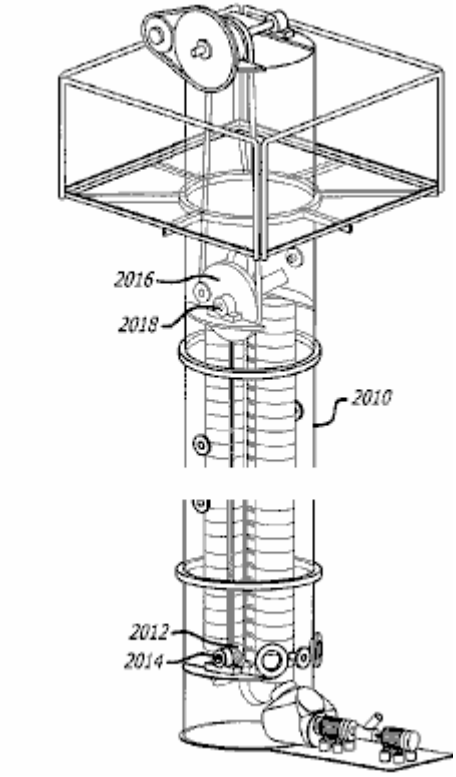
There is another way of achieving this desirable effect without ever having to pump against the total head of water, and that is to use a mobile air pipe as shown in **Fig.4**. Initially, the air pump is started and lowered a short distance into the water. The opposing head of water is not large and the water vortex can be established quite easily. The pipe is then lowered very slowly, so as to maintain the vortex at a progressively lower depth, where, in spite of the increased head of water, the pump does not have to overcome that head. When the pipe outlet reaches the operational depth, it is then rotated to bring it under the set of rising buckets. The major advantage of this arrangement is that the tank is as simple as possible, with no possibility of leaking, and so abandoned wells can be modified to become energy generators. Alternatively, an earth bank can be erected to form an above-ground tank, possibly sealed with a plastic membrane. This method also avoids needing to support the weight of tank and water above a work-area where the air pump or compressed-air cylinders are located and maintained. The creation of the water vortex can be assisted by the addition of a cowl around the pipe exit as shown in this figure, but that is an optional feature.

The Ribero Buoyancy Patent.

While internal combustion engine demonstrate that considerable power can be had from motion which moves backwards and forwards continuously, that sort of action is not very efficient as there is continuous reversal of the oscillating drive components. The floats in the (very successful) 'Hidro' design shown above. A different design is shown in the 2011 patent of Renato Bastos Ribero of Brazil. Here is an excerpt from that patent:

US 7,958,726 **14th June 2011** **Inventor: Renato Bastos Ribero**

Apparatus and associated methods to generate useable energy



Abstract:

The present disclosure relates to an apparatus and associated methods for generating energy by capturing and taking benefit of the energy generated by any quantity of air surfacing inside water. In exemplary embodiments, the apparatus comprises compressing a lower density gas in a liquid medium, allowing the gas to naturally rise to the surface of the liquid medium and then capturing the energy generated by the surfacing gas.

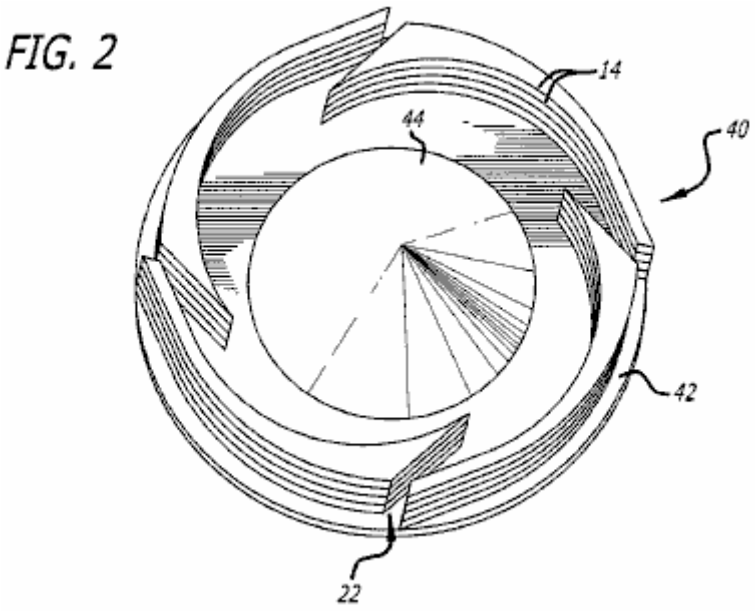
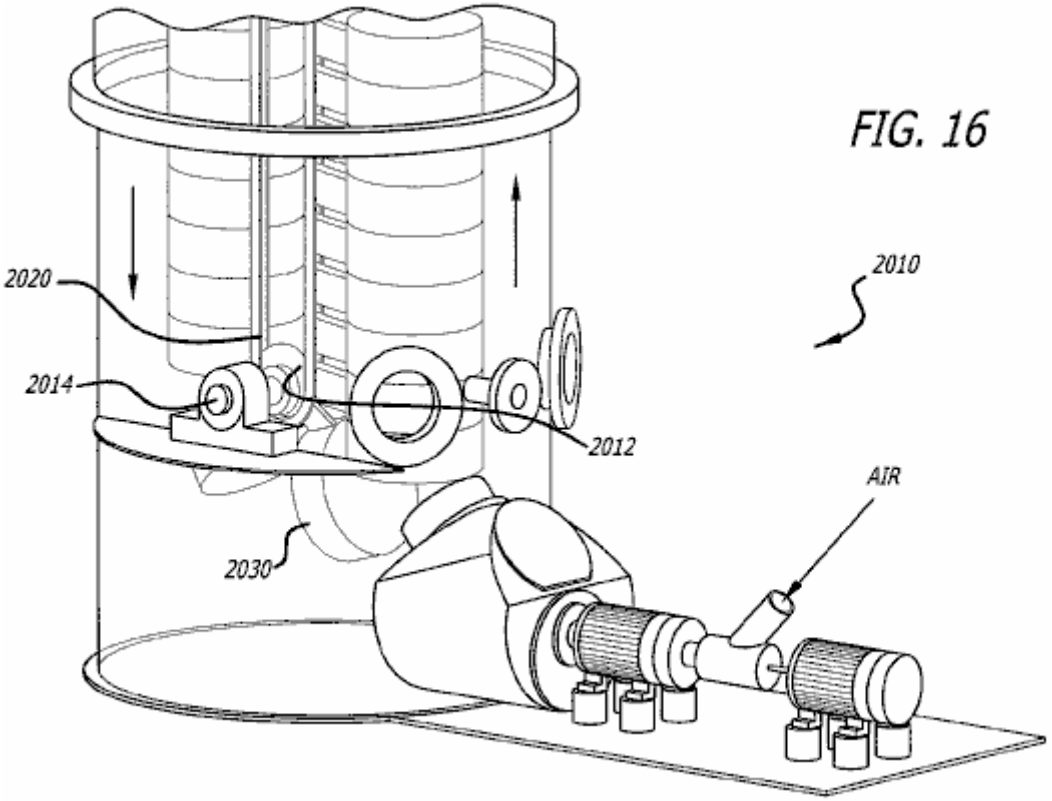


Fig.2 is a perspective upper view of a rotor disc for compressing a gas into a liquid medium. This disclosure is in two stages which, in this case, work together. The first stage consists of the creation of energy with the introduction of air at the bottom part of a water column. Once introduced, the air creates energy when moving towards the surface. The introduction of air into the water is the main point of this first part of the disclosure. A method was created in order to use a very small quantity of energy when doing this.

The teeth on the disc, when rotated in water, water to flow away from the area between the teeth, lowering the pressure there and allowing the easy introduction of air in that area. Without air introduction, water would not move away and the pressure would not be reduced. The purpose of the cone is to spread the incoming air.

The second section of the present disclosure refers to a system with the objective of capturing the air which is inserted at the bottom of a water column or tank, while moving towards the water surface.



As shown in **Fig.16**, buckets **2030** descend on a continuous chain **2020** which passes around a lower and an upper wheel or pulley. On reaching the lower pulley, the buckets turn around the bottom pulley **2014** and line up again on the ascending side of the chain. Immediately after turning around the bottom pulley, each bucket receives the rising airflow.

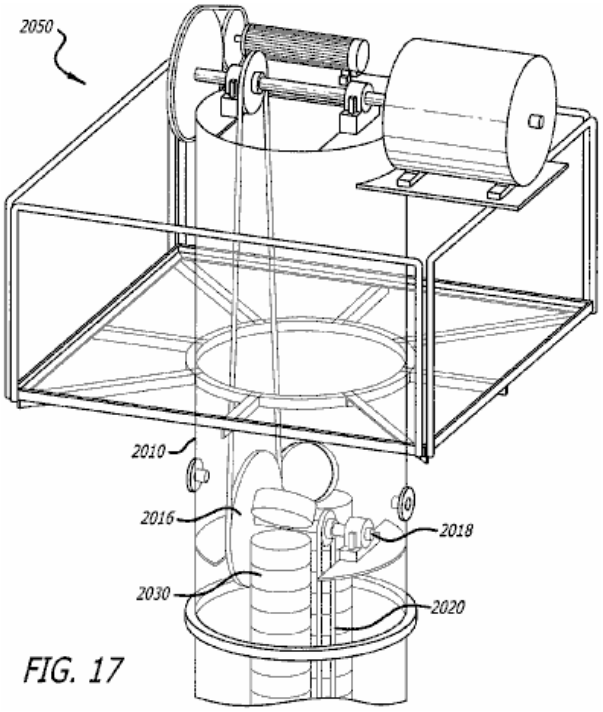


Fig.17 shows the arrangement at the top of the water tank **2010**. Buckets **2030** rising due to being filled with air, turn over as they pass around the top pulley **2018**, releasing the air inside them and start descending again towards the bottom pulley. The upper axle **2018** has an estimated rotational speed of 120 rpm.

In Fig.17 I (Ribero) am showing the transmission of this energy to an axle at the top of the water column where we have a generator **2050** requiring a rotation of 300 rpm plus an engine connected to another generator with rotation of 600 rpm. This part of Fig.17 is only illustrative to show that we shall generate energy at the primary axle at 120 rpm, or use any kind of transmission to more convenient rotational speeds.

***

I think that the words marked in red indicate that although this patent has been granted, the generator has never been built and is only an idea. Personally, I am highly dubious about the mechanisms which are supposed to give reduced water pressure at the air intake, as I don't think that they would work, or if they do, certainly not for the reasons stated. What he wants to do can certainly be done, but not in the way that he suggests. If the axles are rotating at the 120 rpm which he suggests, then that would allow less than one eighth of a second to fill each bucket and while the notion of reduced water turbulence through the buckets touching each other is attractive, I don't think that the method described is feasible.

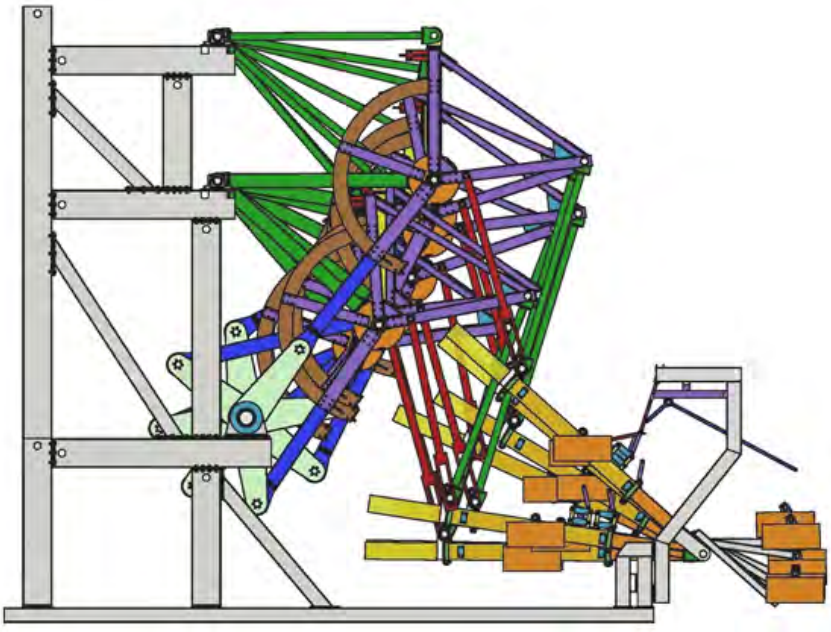
So, while we can be sure that buoyancy methods are perfectly capable of generating serious power, we need a better design than either of the two shown here as the Hidro appears to be very expensive to build.

The Thirty Kilowatt Motor

The recent website http://www.rarenergia.com.br/ shows a gravity-powered motor which is capable of driving a 30-kilowatt electrical generator.

This is certainly not a home-build project and the cost of construction and day-to-day maintenance cost makes this appear to be a very uneconomic project. However, one enormous advantage of the building of these two generators is that they show very clearly that free-energy is available and perfectly viable. The people building these two constructions point out very clearly that these are gravity-powered motors rather than just electrical generators. While these motors can drive electrical generators, it is emphasised that they

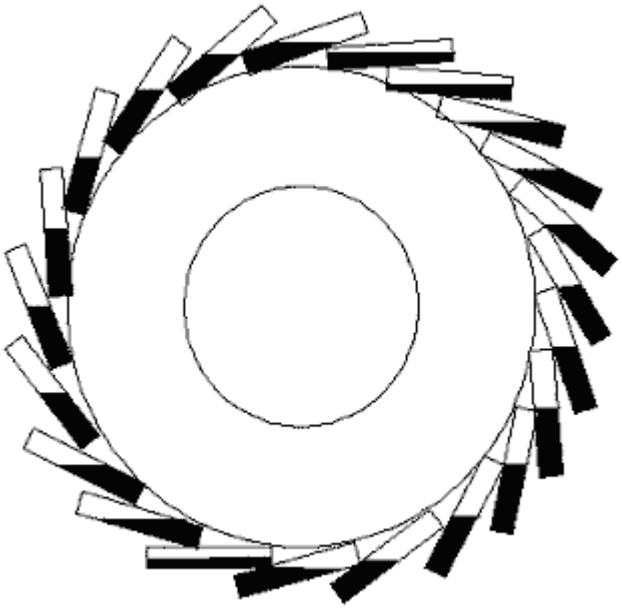
can also perform any task which needs a motor, such as pumping, drilling, etc. The size of these motors is substantial as can be seen from the following pictures:



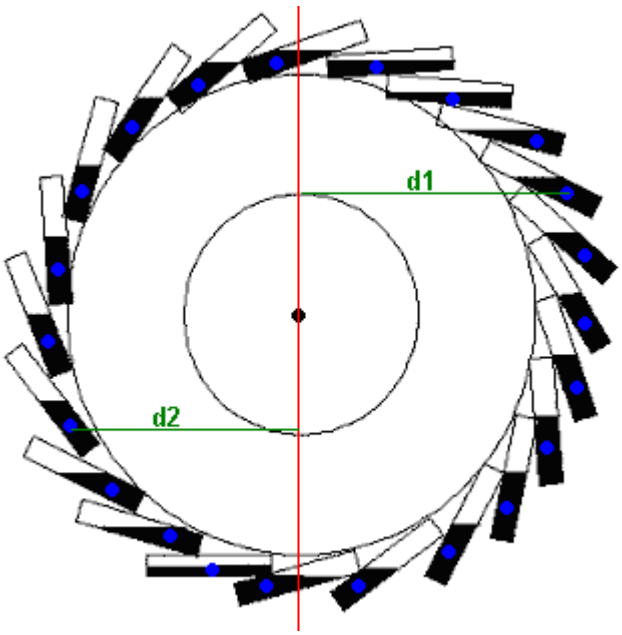
Bhaskara's Gravity Wheel

I have recently been made aware of the Baskara gravity wheel invented in 1150 in India. This is a most interesting idea which appears to have considerable potential. There is what appears to be a marginal replication at http://www.dailymotion.com/video/xygxsy_bhaskara-wheel-overbalanced-chain_tech. Any gravity-powered device needs to be large and heavy if it is to produce serious usable power. The tiny device

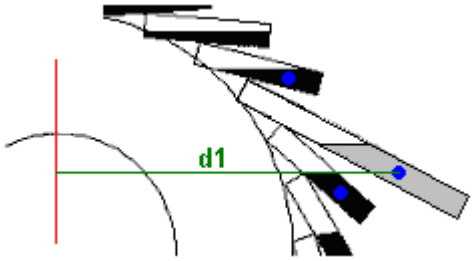
shown in the video is shown stationary which shows that the bearing friction is far too great for the weight of liquid involved. Properly built, it would be impossible to have the wheel stationary unless it were locked in place as the weight imbalance would start it rotating from any stationary position. The overall design of the wheel is generally shown like this:



The idea is that the liquid on the right has its Centre of Gravity further from the axle of the wheel, than that on the left. The “Centre of Gravity” of any object is the point at which all the weight of the object is considered to act. It is usually at the centre of an object which has a regular shape. In this case, the turning effect is created by very small differences in lever arm lengths:

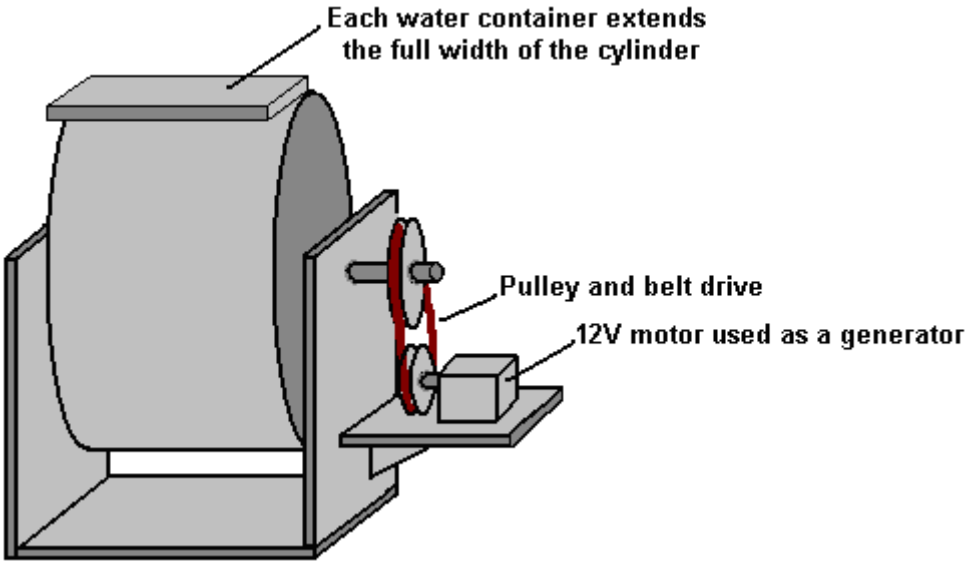


Here, the Centre of Gravity of each body of liquid is marked with a blue dot. The lever arms “d1” and “d2” show the distances for two tubes on opposite side of the axle. Although it is not obvious in this illustration, the length “d1” is actually longer than the length “d2”. The tubes are fixed in position and do not move relative to the wheel and so, the only moving part is the liquid and the rotating wheel. As shown here, the mechanism is not efficient as much of the potential power is not being accessed. The turning force of the liquid is created by the liquid flowing outwards, away from the axle. That flow is being severely limited by the short tube lengths. It would be much more realistic to use a much longer tube like this:



Just by lengthening the tube, there is a major increase in the movement of the liquid away from the axle. The increased tube length has no noticeable effect on the rising side of the wheel.

Most people think of the wheel as being small with a few small tubes attached to it. If actual power extraction is the objective, then the wheel becomes a cylinder and the “tubes” extend all the way across the cylinder. I would suggest that the cylinder diameter is made about one meter and the cylinder length about half a metre if you are going to carry it through a doorway, and much longer if not. In this arrangement, the “tubes” become shallow, flat containers, making the arrangement something like this:



If the liquid containers are made of, say, 3 mm MDF, then I would suggest that the sides are made using MDF of 6 mm thickness to make the container rigid as the weight of liquid will be quite high. The inside of the container is sealed if the construction material is not waterproof, then a coating of varnish, paint or one of the waterproofing sprays will prevent water absorption. The liquid is added before the top of the container is glued in place. Exactly the same amount of liquid should be poured into each container before sealing and a measuring jug should be used to make sure that each half-filled container matches the weight of each of the other containers.

A twelve volt DC motor can be used as a generator, driven by a belt and two pulleys, in order to get some useful electrical output from the device. Materials other than a liquid have been suggested as the heavier the material is, the more powerful the generator becomes. Mercury would be very good as it is very heavy, but as it is so poisonous, (not to mention expensive), it is not a realistic choice. Lead shot or steel ball bearings have been suggested but they would probably not flow across the containers well enough to be a good choice, and so a liquid is perhaps the best thing to use, water being an obvious choice.

The Torque Generator of William F. Skinner

In 1939, William Skinner of Miami in Florida, demonstrated his fifth generation generator powered by spinning weights. His demonstration can still be seen at http://www.britishpathe.com/video/gravity-power where he shows his design powering a twelve-foot lathe, a drill press and a power hacksaw, all simultaneously. The newsreel commentator states that the output power was “1200% of the input power” which is COP=12 but it is highly likely that he should have said “1200 times” rather than “1200%” because

he continues to state that using the design would allow a one-horsepower (746 watts) input power to power 3,500 homes. If it were COP=12 then each of those 3,500 homes would receive less than 2.6 watts, which is clearly wrong. At the much more likely COP=1200, each household would receive on average, 255 watts, which might be just possible in 1939 when few appliances were electric. Anyway, Skinner's impressive equipment could be driven by a single cotton thread drive band while powering his whole workshop. It looked like this:

This design has four nearly vertical shafts, each braced to give additional rigidity. These rotating shafts pass their rotating power to the mechanical output drive belt seen on the left. Each of these rotating shafts has a heavy weight in the form of a thick, short cylinder mounted high up near the top of the shaft and what is probably an even heavier weight in the form of a long narrower cylinder attached near the bottom of the shaft as seen just to the right of the output drive belt. These four identical sets of shafts with their pairs of weights spin two or three times per second and produce the whole of the output power.

As far as I am aware, Skinner never patented his design or disclosed how it worked. However, the operating principle is very simple indeed although it may take you a while to grasp how it works. You can check this out quite easily for yourself if you have access to an old-fashioned chair with four rigid legs like this:

Tilt the chair over so that it is balanced on one leg. You will notice that almost no effort is involved in keeping it in that position as all of the weight is supported by the floor through just one of the legs. Now, move the top of the chair by a very small amount and keep the top of the chair in that position. You will notice two things: first, very little effort was needed to move the top of the chair and second, the chair now swings around and becomes stationary on the same side that the top of the chair was moved.

Notice two other things: the chair swung around because of your moving the top slightly and **you** did not swing it around, and if the chair is heavy, the amount of energy in the swinging chair is very much greater that the amount of energy which you applied to the top of the chair.

If you were to keep moving the top of the chair in a tiny circle, then the chair will spin around continuously for however long that you choose to wobble the top of the chair. The amount of energy in the spinning chair is very much greater than the energy which you are expending to make the chair spin. So from where is that extra energy coming?

What is happening is that the chair swings round under gravity to reach the lowest possible point for it with the new position of the top of the chair. But, before it can get there, you move the top of the chair further around and so the chair has to swing further in order the reach the lowest point. But before it can get there, you move the top again ….. The chair keeps swinging round and round, pulled by gravity, for as long as you choose to keep moving the top. But, no matter how heavy the chair, very little effort is needed from you to cause the spinning.

Skinner had a mechanism at the top of each vertical drive shaft, and that mechanism kept moving the top of the shaft in a small circle while allowing the shaft to rotate freely at all times. That caused the very heavy weights attached to the shaft to keep spinning around, and he used that power of the heavy spinning weights to power his whole workshop. Moving the top of the shafts required so little power that he used a 93-watt electric motor and to show that he was not even using all the power of that small motor, he used a single cotton thread as a drive band to move the tops of the four power output shafts.

His mechanism looks complicated. This is partly due to the fact that there are four identical power shafts with their weights, mounted in the one compact frame and that makes the device look more complicated than it really is. It is also due to the fact that the system shown in the newsreel is William's fifth version of the device. It is likely that his earlier, much more simple versions worked well and encouraged him to build even fancier versions.

There are two forums where members of those forums are trying to work out exactly how his final version machine worked and then replicate the design for current day use as it is a neat system for accessing additional usable power. Those forums are at:

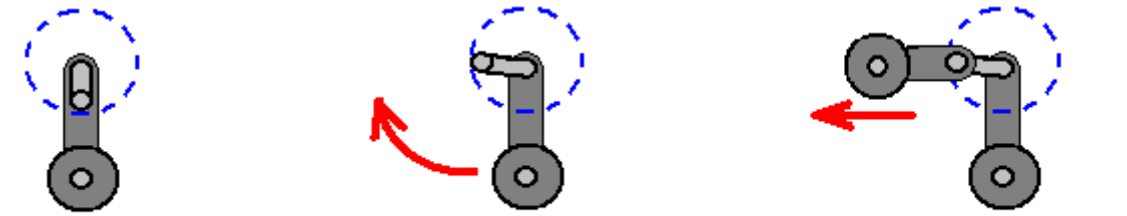
http://www.overunity.com/14655/1939-gravity-power-multiply-power-by-1200/#.U5y0gXaqmJA and

http://www.energeticforum.com/renewable-energy/17195-william-f-skinner-1939-gravity-power.html

It needs to be remembered however, that it is not actually necessary to replicate William's fifth version, but instead it would be quite enough to use the principle of the spinning chair to produce a simple mechanism where the input power is far less than the output power.

If we consider what is happening, then perhaps we can understand Skinner's complicated-looking arrangement. We can consider just one of the four axle shafts The large weight is spinning around in a circle and that motion is then used to power the output shaft. In order to reduce the effort needed to spin the weight, the axle shaft has been made thinner and four bracing rods have been used to brace the shaft in exactly the same way that sailing yacht masts are usually braced with "spreaders" to hold the bracing out from the mast and so give greater overall stiffness. So we can ignore those bracing bars as they have nothing to do with the actual operation of his design, but are merely his choice out of many different construction options.

Remember the spinning chair and consider what has to be done to spin Skinner's heavy weight. The top of the shaft has to be moved in a small circle. Looking down from the top the situation is like this:

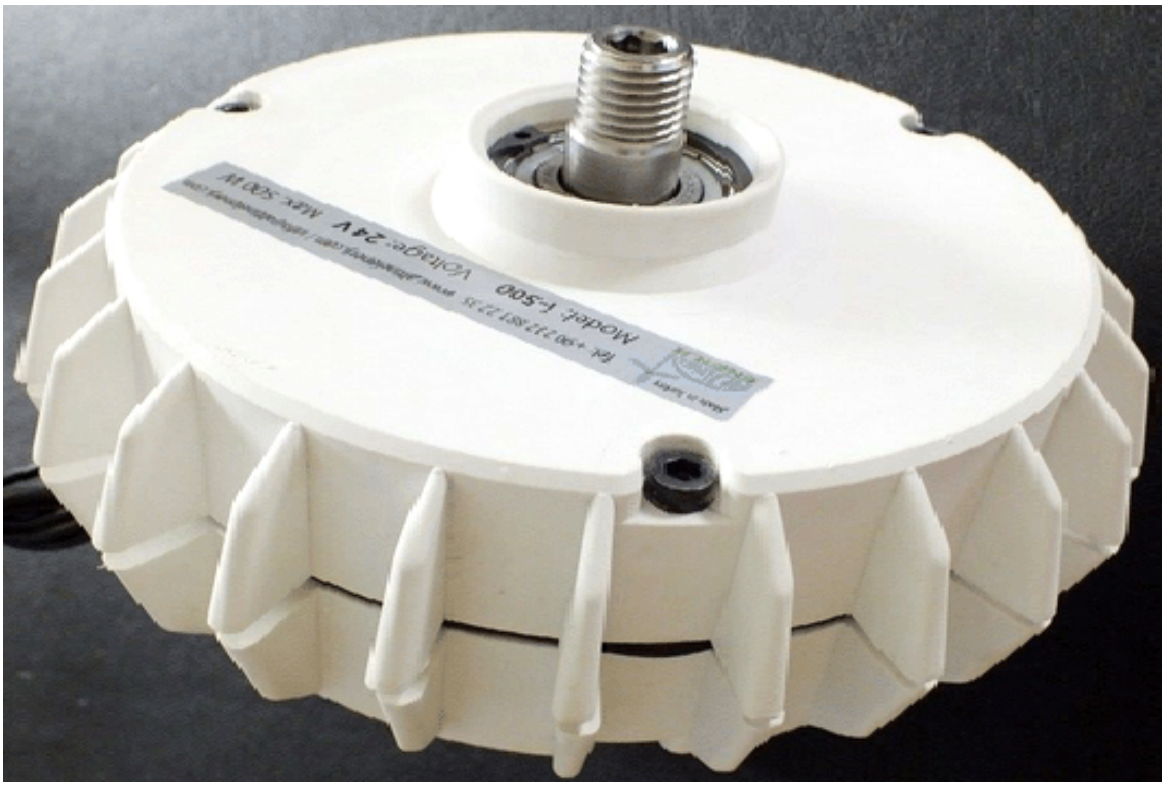


When the system is switched off, the weight attached to the bottom of the shaft comes to rest directly underneath the top of the shaft. When the system is started again, the first move is to shift the top of the axle shaft ninety degrees around. This is the start of the rotary movement and initially, the movement is slow as it takes the heavy weight some time to get moving. To reduce the effort of moving the top of the shaft ninety degrees ahead of the big lower weight, Skinner has added a weight at the top to assist the movement in that direction.

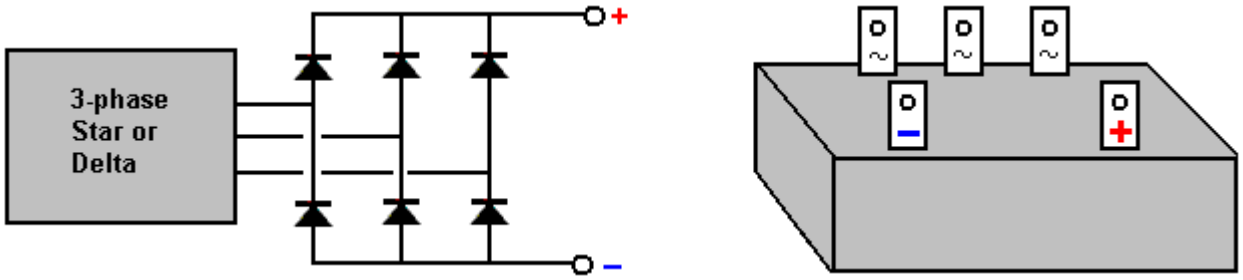
Skinner also took advantage of his very large workshop to use a belt-driven mechanism above the top of the shaft, in order to reduce the effort of moving the top of the axle shaft even further (to the level where it could be driven by a cotton thread). He used four separate identical shafts in his construction for two reasons: first, the overall output power is increased and second, any sideways forces stressing the mounting frame are matched on every side, which is helpful when you have heavy weights on a rotating arm as Skinner did.

As the output shafts appear to be rotating at about 150 rpm, Skinner opted to use a straight mechanical drive. Back in 1939, electrically-driven equipment was not as widespread as it is today, but nowadays we would probably prefer to have an electrical output rather than a mechanical drive although that mechanical drive could be used for driving pumps and other low-speed devices. So, we are faced with introducing some form of gearing which can raise that 150 rpm to the much higher level preferred by most alternators.

While it would be possible to use an ordinary 12-volt motor as a generator and produce a 12-volt electrical output, it is likely to be more convenient to use an off-the-shelf electrical generator, perhaps a very low-friction one like this which has been designed for wind-power operation and which has a 12V or 24V 3-phase output:

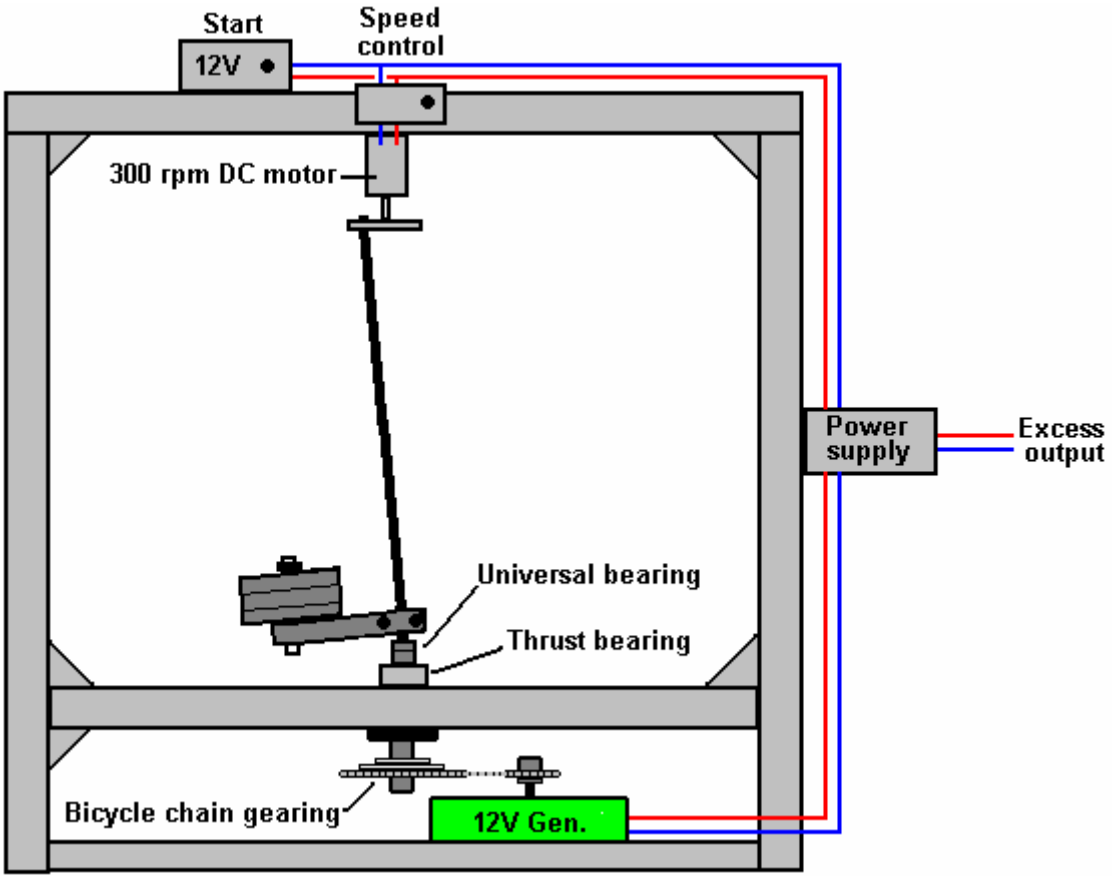


The fact that the output is 3-phase can sound a little daunting, but the conversion to DC is quite straightforward:

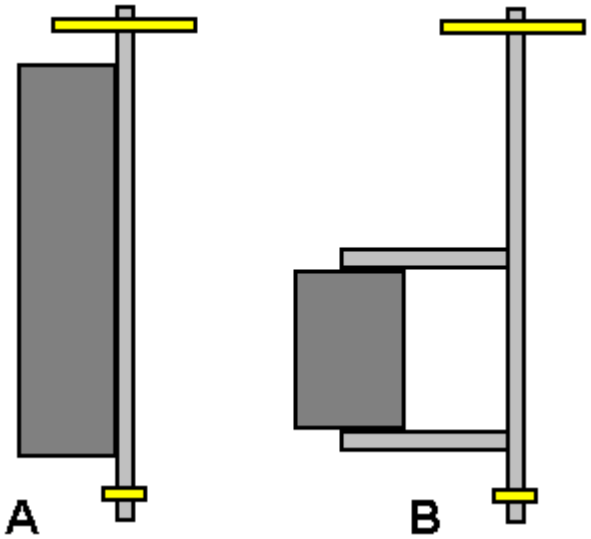


The output can be converted to DC with six ordinary diodes or a integrated diode arrangement can be used where there is a connecting tag for each of the three outputs and a separate tag for the DC Plus and for the DC Minus. The currents involved are quite high as 400 watts at 12-volts represents more than 33 amps and the peak output of 500 watts is a current of about 42 amps. For that reason, the 3-phase rectifier blocks are rated at 50 amps which sounds very high until you do the calculations and discover what the current is likely to be. It should also be borne in mind that the DC output wire has to carry that level of current on a continuous basis and so fairly robust wire is needed. If the voltage were 220V then the wire would be carrying more than 9 kilowatts at that current flow, and so the normal 13-amp mains wire is just not sufficient and instead, we need to use thick wire or more than one strand of wire for both the Plus and the Minus connections.

This particular generator is not expensive and can output 400 watts of electricity (33 amps) continuously. As the Skinner type appears to be spinning at 150 Hz, a gearing up of the output speed would allow greater output, so perhaps for a home-builder, the physical arrangement might be like this:



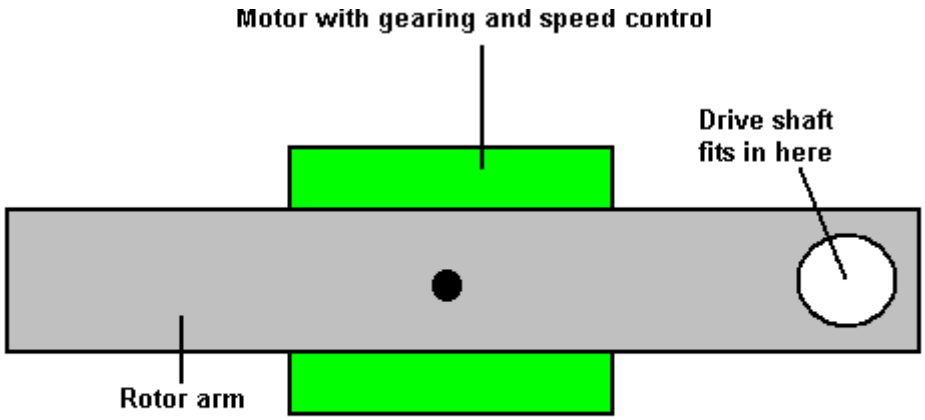
There are, of course, many different forms of construction which might be used, but with each of them, the question is, “how do you make the angled shaft rotate powerfully?”. If you can work out the complexities of Skinner’s fifth version shown in the newsreel, then that would certainly do the job. However, we would prefer a much more simple design and so we do not necessarily have to copy what Skinner did but instead we can just apply the principle which he demonstrated. One possible arrangement might be to imitate the chair experiment using a strong shaft with a weight attached to one side of it, perhaps like this:



Version “A” uses the weight to stiffen the shaft but doing that raises the centre of gravity of the combined shaft and weight which may not be convenient. Version “B” increases the torque for any given weight by moving the centre of gravity of the weight away from the centreline of the shaft by means of extension arms. As the shaft rotates at a constant rate, the load on the shaft will be essentially constant and there should not be any significant flexing of the shaft although it might bend and remain with that same bend during all of the time when it is spinning if the weight is very high relative to the stiffness of the shaft.

We do have to input some power to rotate the top of the drive shaft, but if we arrange things in any one of the hundreds of viable configurations, then the output power will be massively greater than our input power. An alternative arrangement which allows speed control (and so, output power control) is to take some of the generated output of electricity and use that to power an electrical drive which positions the top of the drive shaft.

There will be many different ways of achieving that movement. One method for doing this might be:



SEEN FROM UNDERNEATH

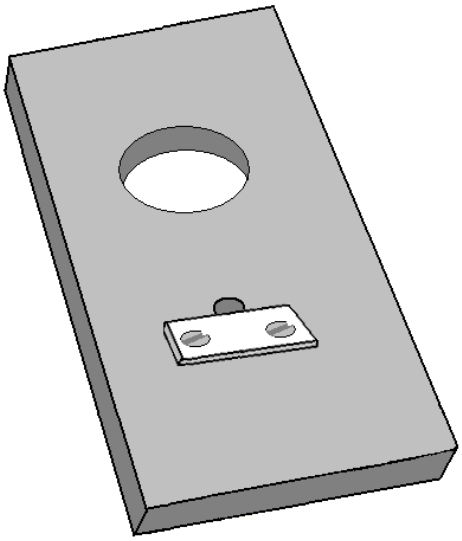
Here, the small electric motor shown in green is geared down and used to move the top of the drive shaft at whatever rate of revolution that we consider to be satisfactory, using a standard DC motor speed controller.

It should be noted that no matter what angle is chosen for the axle shaft, that is always a constant relative to the motor arm moving it round in the circle at the top of the shaft. This means that no roller bearing is needed as there is no relative movement and the shaft will automatically take up that fixed angle. The drive motor arm moving the top of the shaft will probably not be long, as Skinner appeared to be moving the top of his shafts by about 40 mm away from the centreline of the bottom pivot, making only one degree or so for the angle of the shaft on every side of the vertical.

It is, of course, not essential to convert the output power to electricity and instead it could be used in the same way that Skinner did, driving mechanical equipment such as water pumps for irrigation or extracting water from wells, milling operations for processing grain or for operating any form of workshop equipment. It is also not necessary to build the device anywhere near as large as Skinner did, and small versions could be used to power lighting systems, operate fans or cooling systems or for any other minor household requirements.

The power output from the machine can be increased by increasing the weight attached to the output shaft, or by increasing the length of the arm holding the weight, or by tilting the output shaft through a greater angle (which increases the input power needed, but probably not by much), or perhaps by scaling the whole thing up so that it is physically bigger. Skinner's design uses stiffening bracing on the output shaft, which suggests that the lighter the shaft is, the better the performance. Because of this, a prototype build might use a timber shaft of perhaps, 33 mm square as that is both light and very strong and rigid and it is a good shape for ensuring that there is no slipping of the arm which supports the weights. The top of the shaft is reduced slightly so that it has a circular cross-section. A 300 rpm motor rotates at a maximum of 5 turns per second and so is suitable for rotating the axle shaft. A suitable, low-cost motor of that type, looks like this:

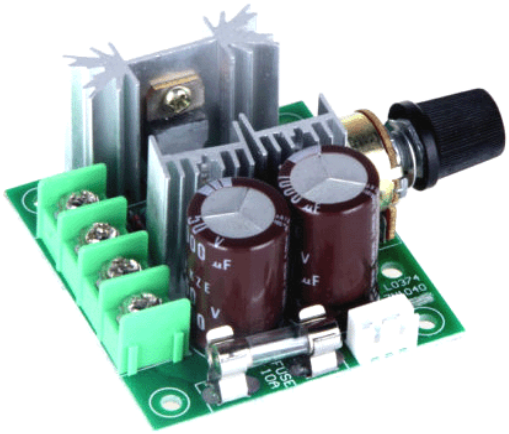
The motor needs to be linked to the shaft in a simple way which ensures that there will be no shaft slippage:



Perhaps cutting a suitable sized hole through a strip of material and using a strip of metal pressed into the flat face of the motor drive shaft (in addition to the hole being a tight push fit) would be adequate for this. A screwed collar or layer of epoxy resin holds the plate firmly to the motor as the plate is positioned below the motor and so gravity tends to pull the plate off the motor shaft at all times.

It would initially be assumed that a ball bearing or roller bearing would be needed in this motor arm, but that is not the case as the axle shaft does not rotate relative to the motor arm and while the axle shaft can be a loose fit in the hole, there is certainly no need for a bearing.

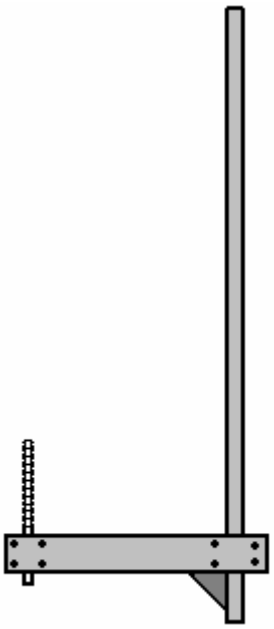
A commercial DC Motor Speed Controller can be used to bring the shaft rotation speed gradually up from a stationary start to the chosen rate of revolution:



Using a commercial module like this means that no electronics knowledge is needed to build a working generator of this type.

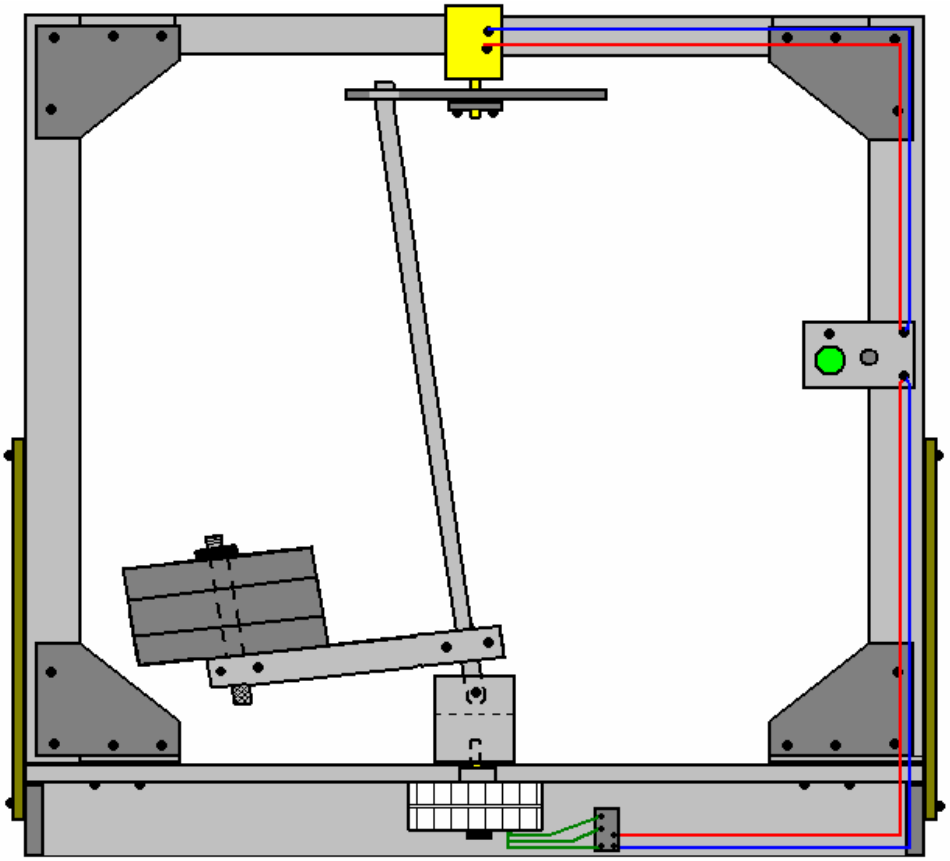
There are many options for providing the necessary weight which drives the generator. One possibility is to use a barbell shaft with as many weights as are required, that being a very simple alteration:

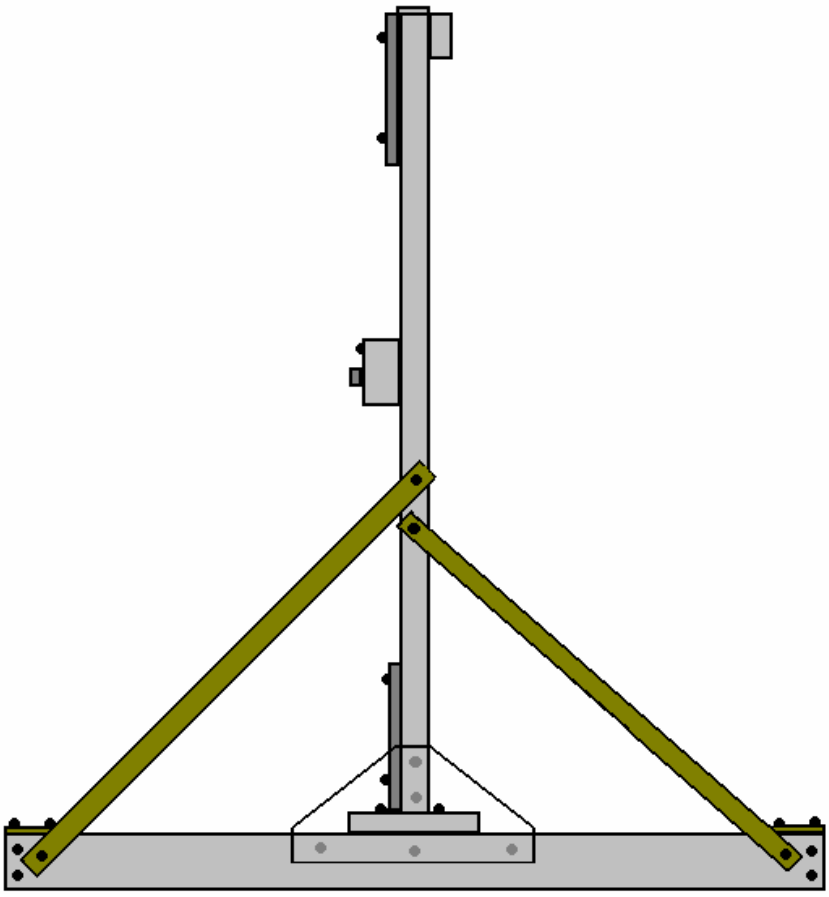
One of the hand grips can be cut and used directly as part of the mounting, perhaps like this:



This simple arrangement allows the weight discs to be added and secured in any combination desired. As dumbbells are supplied in pairs, there are four discs of each side which allows a wide range of weight options going up in jumps of just 1 Kg which is very convenient. If the axle shaft has a square cross section, there is no tendency for the lever arm to slide around the shaft

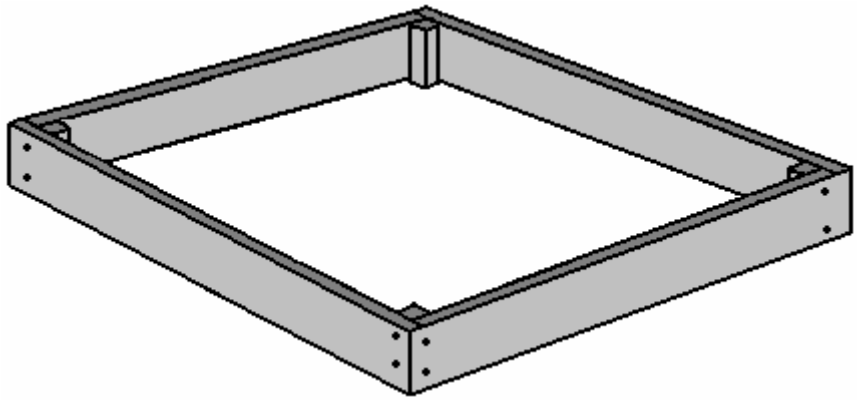
The following sketches are not to scale, but one form of construction might be:



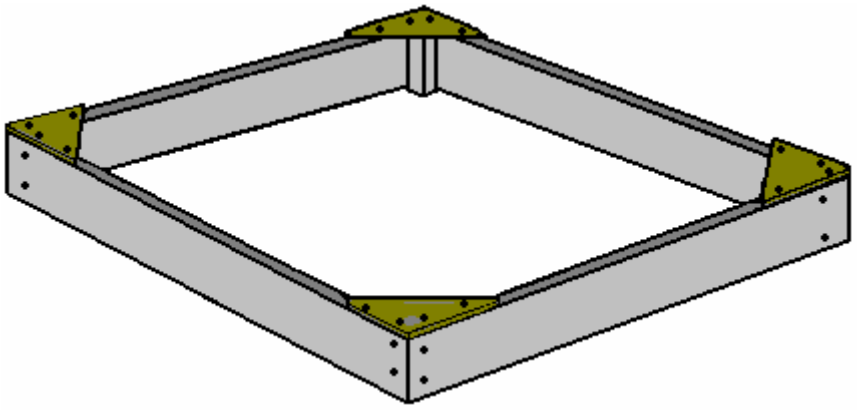


For this style of construction, four pieces of, perhaps, 70 x 18 mm Planed Square Edge timber are cut to perhaps 1050 mm and two 33 x 33 x 65 mm pieces epoxied and screwed to two of the pieces, 18 mm in from the ends:

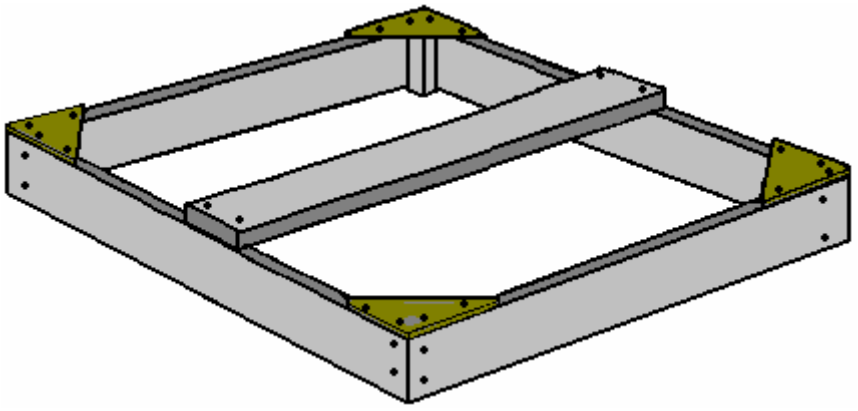
Then the four pieces are screwed together while resting on a flat surface:



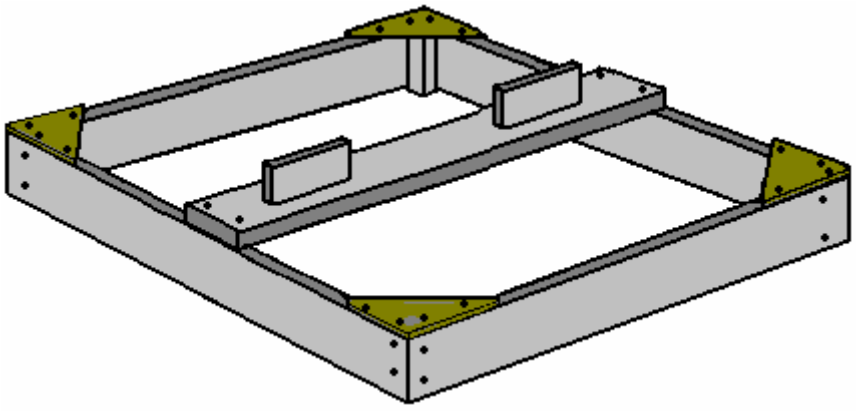
Then corner bracing triangles of MDF are screwed in place:



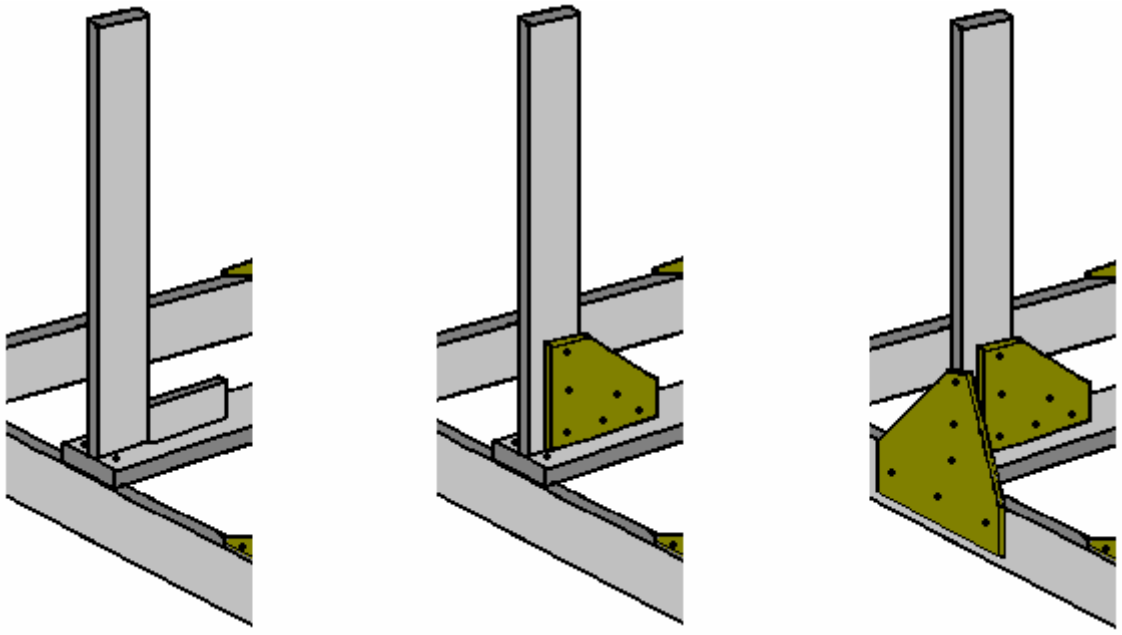
Then a 130 x 25 mm thick plank is attached across the width at the centre point and screwed in place:



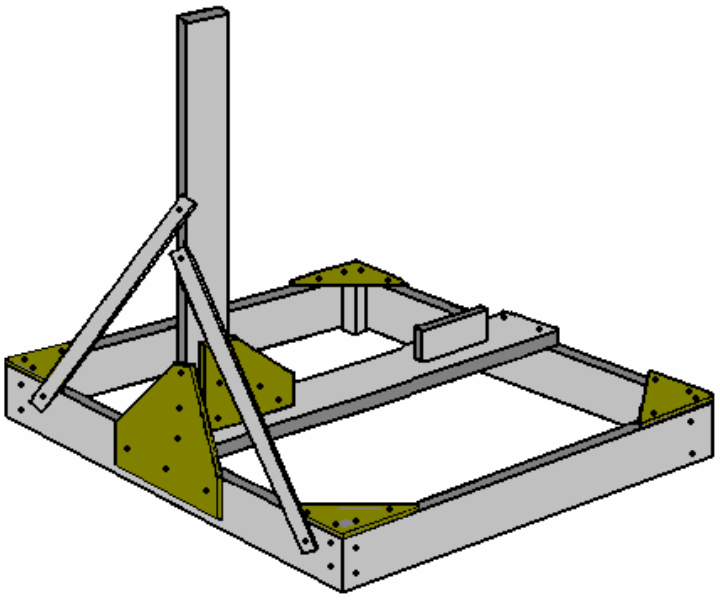
Next, two lengths of the 18 mm thick timbers about 180 mm long are epoxied and screwed to the centre of the 25 mm thick plank, leaving 70 mm clearance to the end of the plank:



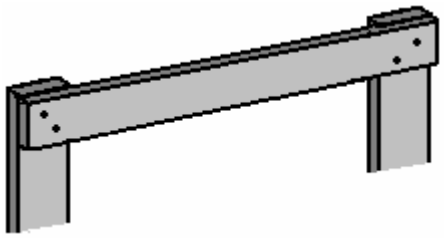
Two timber strips 1350 mm long, are cut and erected vertically, being attached by screws coming upwards through the 25 mm thick plank, and by MDF bracing triangles on one side and across the lower end of the verticals. If a spirit level is used to ensure that the vertical timber is actually vertical, then first, the four corners of the floor frame need to be weighted down to overcome any twisting and the floor frame confirmed to be actually horizontal before attaching the vertical timbers:



Each vertical needs to be braced on both sides with diagonal strip, either metal or timber:

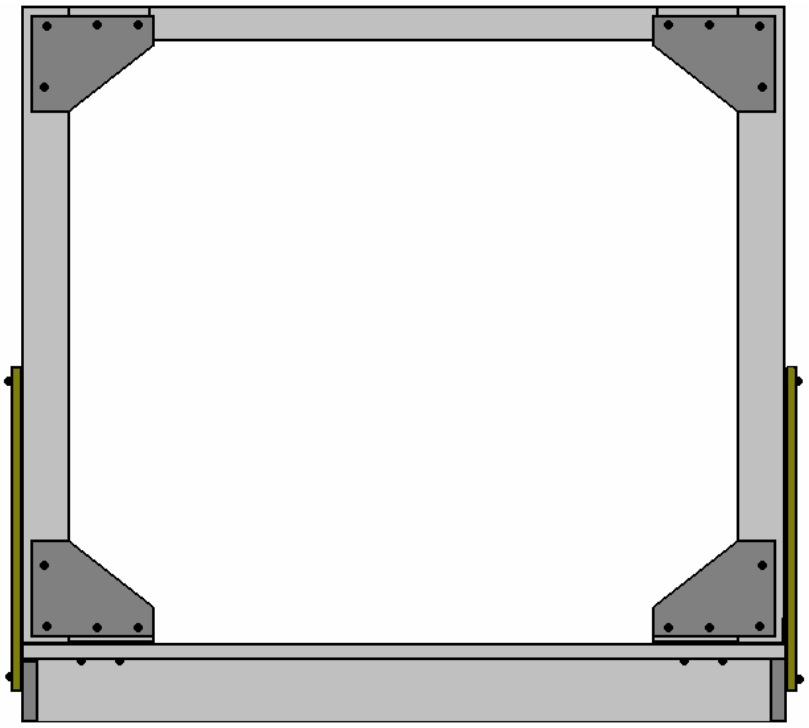


An 18 mm thick timber strip is screwed to the tops of the verticals. This deliberately positions the timber 18 mm off centre as the motor which rotates the top of the axle shaft has to be attached to the middle of this newest timber and that places the motor shaft very close to the central point of the base:

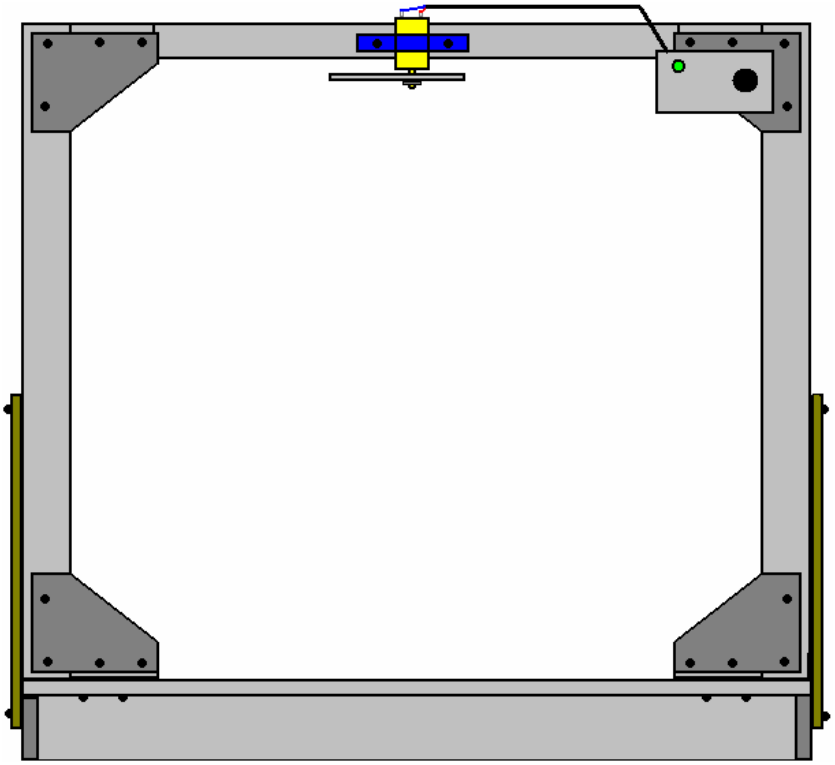


One slight disadvantage is that a packing piece is needed for the triangular MDF bracing pieces which increase the frame rigidity at the top:

At this stage, the construction will look like this:



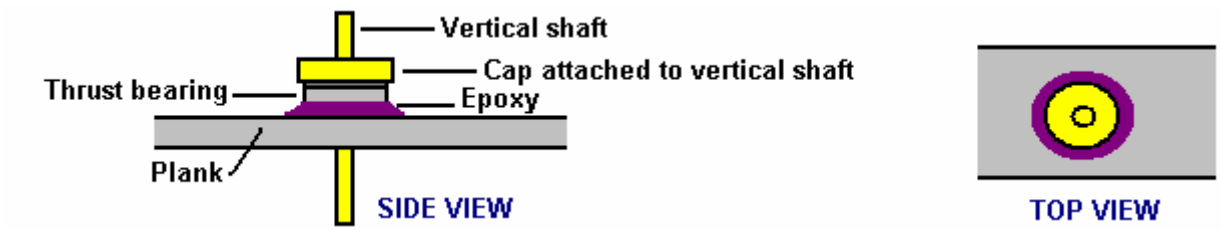
At this point, the 300 rpm motor with it's actuator arm and the speed-control box can be fitted. The motor is located centrally, and the control box can be positioned anywhere convenient. The control box is merely a 12-volt battery pack of 1.2V NiMh AA-size batteries connected through a push-to-make press button switch and the commercial DC Motor Speed Controller, to the 300 rpm motor. With this arrangement, the motor can be powered up by pressing the button and adjusting the speed slowly up from stationary, getting the rotor weight moving gradually faster and faster until its best operating speed is reached. When everything is in place, then the rectified output of the alternator is fed into the control box, so that the Start button can be released and the device becomes self-powered from part of the output power. The initial step looks like this:



It should be explained that, with the exception of the 25 mm thick plank, all of this construction is only loaded very lightly as rotating the top of the axle shaft does not take much power or effort at all. Almost all of the rotating weight is located at the bottom of the axle shaft and that weight rests on some form of bearing which rests in the middle of the 25 mm plank.

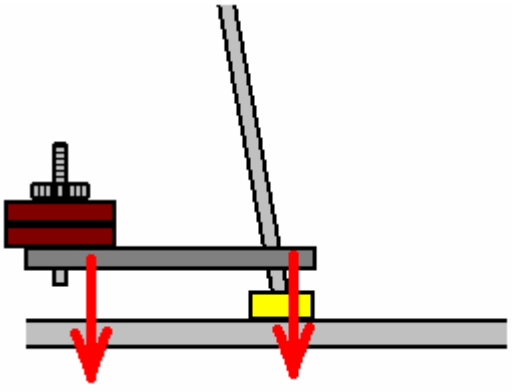
For a small version of the generator, such as this one, the rotating weight does not need to be all that great and so, the forces generated by the weight and its rotation about the bearing need not be a major thing. However, in spite of the fact that we are only dealing with limited forces which can be handled by simple components, people may be inclined to use a thrust bearing instead of allowing the weight to rest on the shaft of the alternator. A bearing of that kind may look like this:

Here, the base and inner ring do not move while the top outer ring revolves freely and can support a major load while it rotates. If we choose to use one of these, then an arrangement like this could be used:

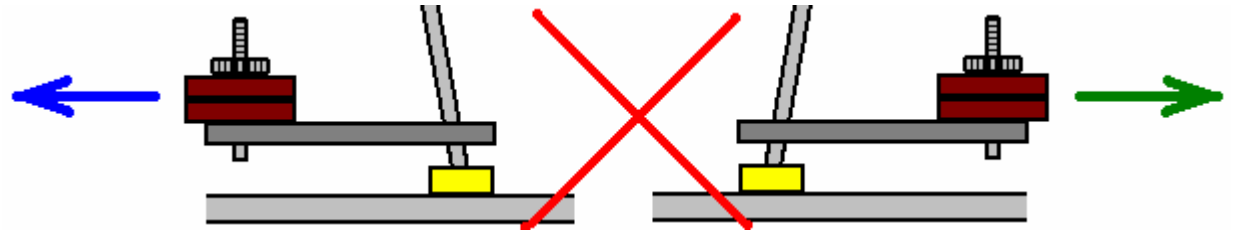


This combination has a cap (shown in yellow) with a central vertical shaft (yellow) attached to it, tightly encasing the upper ring of the bearing whose lower ring is securely attached to the 25 mm thick plank (grey) perhaps using epoxy resin (purple). This allows free rotation of the upper ring and vertical shaft while carrying significant loading. The power take-off in the arrangement shown is from the shaft projecting beneath the plank. Generally speaking, the electrical power output increases with increased speed of rotation, so gearing the alternator up so that it rotates much faster than the axle shaft is desirable and this arrangement may be convenient for that. If it is important to have the power take-off above the plank, then a strong bracket can be used to raise the bearing high enough above the plank to accomplish that.

There are two separate forces acting on the bearing. One is always downwards as the bearing supports the rotating weight:

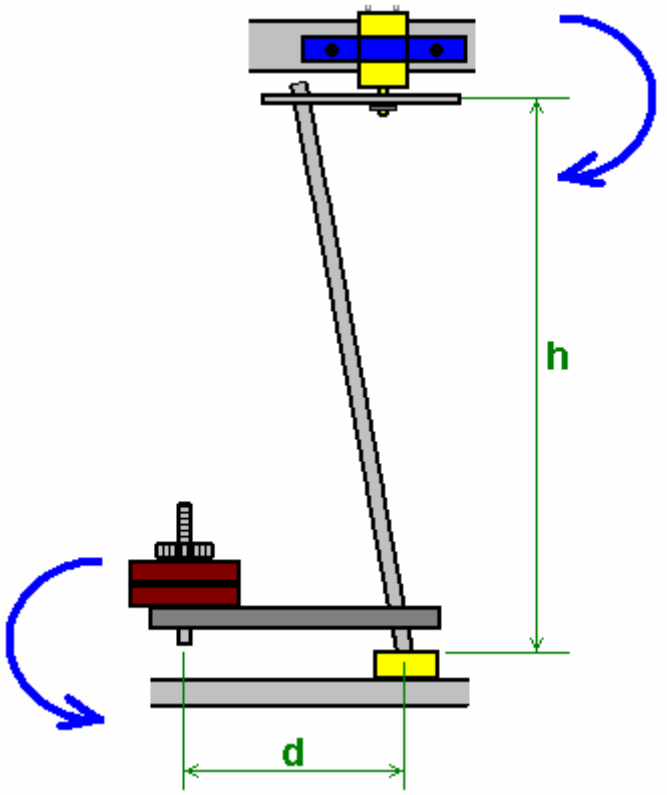


Then there is the sideways forces caused by the rotation of the (unbalanced) weight:



This sideways force is normally considered to be a major problem, however, in this instance, the weight is not being whirled around and trying to escape from the shaft in a horizontal direction, but instead, the weight is turning under gravity powered by its own weight, and the forces generated are quite different and in a different direction. Also, the rate of rotation is very small compared to the speeds which we automatically think about when considering an orbiting weight, typically, this rotation only being between 150 and 300 rpm.

As far as the loading on the axle drive motor is concerned, the situation is like this:

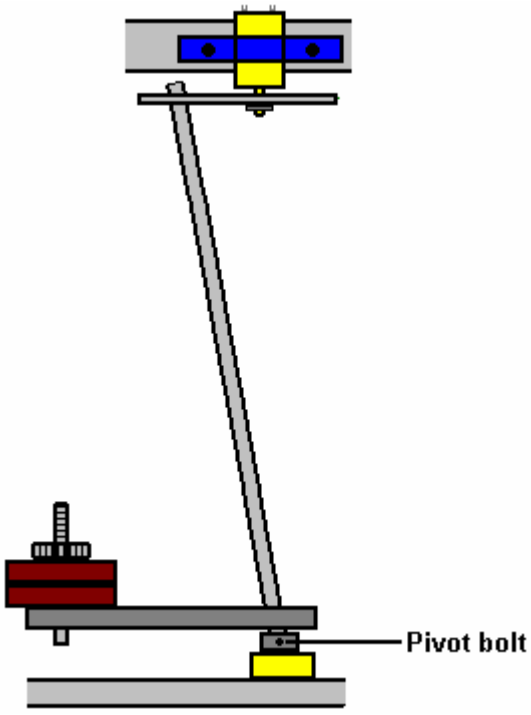


This is the position when at rest. The pull on the motor shaft at the top of the axle shaft is **W x d / h** where **W** is the weight at the end of arm **d**. The situation changes immediately the top of the axle shaft is rotated and the weight **W** starts to swing under the influence of gravity.

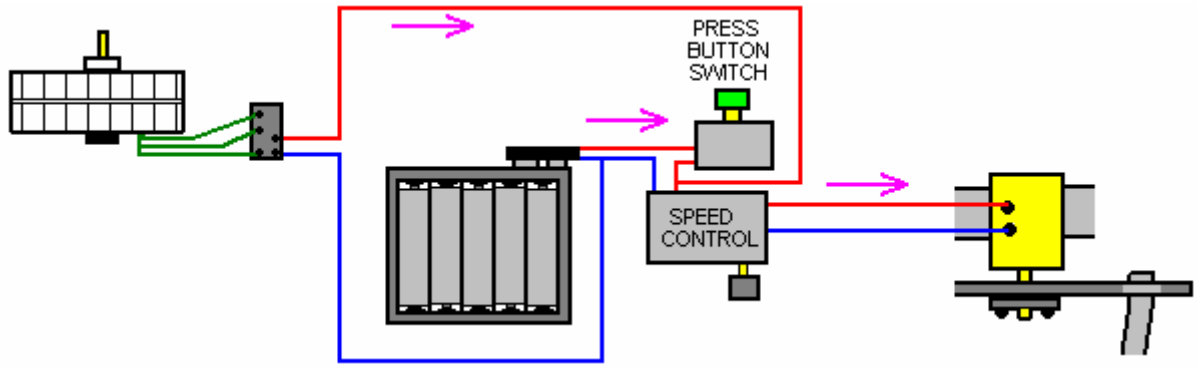
I am told that the axle shaft needs to be light. With small weights, a rigid wooden shaft is adequate and it does not flex under the loading. I am assured that the bottom of the axle shaft needs a universal joint and a major version of this generator where the weights are very high, that is certainly true as the shaft will flex if designed to its minimum specification, but under these much less stressed conditions, there will be no flexing of the shaft when it is pulled sideways and as the shaft angle is a constant, I do not believe that any such joint is necessary. However, many people will wish to include one. These bearings come in different forms, and one of them looks like this:

It must be remembered that if a joint like this is fitted, then it will not be in constant motion, that is, the joints will take up one particular position and will maintain that position during the whole of the time that the generator is in operation.

A compromise would be to provide a hinged movement in one plane by pivoting the axle shaft joint just above the thrust bearing:



The electrical connections are quite straightforward:



The 12-volt battery pack of 1.2V AA-size batteries is connected to the motor speed controller when the button of the press-button switch is held down. This powers the motor, and as the axle shaft speeds up progressively, the generator starts producing power which is always fed to the speed controller box. As soon as the generator gets up to speed the press button switch can be released and the system runs on power produced by the generator. Excess power will be drawn from the generator output, but those links are not shown in the diagram.

Patrick Kelly
www.free-energy-info.tuks.nl
www.free-energy-info.com
www.free-energy-info.co.uk
www.free-energy-devices.com

Chapter 5: Energy-Tapping Pulsed Systems

Note: If you are not at all familiar with basic electronics, you might find it easier to understand this chapter if you read chapter 12 first.

One very interesting feature of free-energy devices is that although various devices which appear to be completely different and have different apparent applications, the background operation is often the same. It is clear that a sharp positive going DC electric pulse interacts with the surrounding energy field, making large quantities of free-energy available for anyone who has the knowledge of how to gather and use that extra energy.

Let me stress again that "over-unity" is an impossibility. Over-unity suggests that more energy can be taken out of a system than the total energy which goes into the system. This is not possible as you can't have more than 100% of anything. However, there is another perfectly valid way of looking at the operation of any system, and that is to rate the output of the system relative to the amount of energy that the user has to put in to make it work. This is called the "Coefficient Of Performance" or "COP" for short. A COP = 1 is when all of the energy put in by the user is returned as useful output. A COP>1 is where more useful energy comes out of the device than the user has to put in. For example, a sailing boat in a good breeze transports people along without the need for the energy of movement to be supplied by the crew. The energy comes from the local environment and while the efficiency is low, the COP is greater than 1. What we are looking for here is not something to tap wind energy, wave energy, sunlight energy, river energy, thermal energy or whatever but instead we want something which can tap the invisible energy field which surrounds us all, namely the "zero-point energy" field or the "ambient background".

For this, let us look at pulsing circuits used by a wide range of people in a number of apparently quite different devices. An electrical "pulse" is a sudden voltage rise and fall with very sharply rising and falling voltages. However, pulses are seldom generated as isolated events when working with practical devices, so it is probably better to think of a train of pulses, or a "waveform" with very sharp rising and falling edges. These can be called oscillators or signal generators and are so commonplace that we tend not to give them a second thought, but the really important factors for using an oscillator for zero-point energy pick-up is the quality of the signal. Ideally, what is needed can be a perfect square wave with no overshoot, and the voltage level never going below zero volts, or a complex waveform, also with very sharp attack and decay times. These waveforms are a good deal more difficult to generate than you might imagine.

Even in these days of sophisticated solid-state electronic devices, the best method of creating a really sharp voltage pulse is still considered to be a spark gap, especially one which has the spark chopped off suddenly by the use of a strong magnetic field at right angles to the spark gap. For an example of this style of operation, consider the following device.

Frank Prentice's COP=6 Pulsed Aerial System.

Electrical Engineer Frank Wyatt Prentice of the USA invented what he described as an 'Electrical Power Accumulator' with an output power six times greater than the input power (COP = 6). He was granted a patent in 1923 which says:

My invention relates to improvements in ELECTRICAL POWER ACCUMULATORS and the like, wherein the earth, acting as rotor and the surrounding air as a stator, collects the energy thus generated by the earth rotating on it's axis, utilises it for power and other purposes.

In the development of my WIRELESS TRAIN CONTROL SYSTEM for railways, covered by my United States Letters Patent Number 843,550, I discovered that with an antenna consisting of one wire of suitable diameter supported on insulators, three to six inches above the ground and extending one half mile, more or less in length, the antenna being grounded at one end through a spark gap, and energised at the other end by a high frequency generator of 500 watts input and having a secondary frequency of 500,000 Hz, would produce in the antenna, an oscillatory frequency the same as that of the earth currents and thus electrical power from the surrounding media was accumulated along the length of the transmission antenna and with a closed oscillatory loop antenna 18 feet in length run parallel with the transmission antenna at a distance of approximately 20 feet, it was possible to obtain by tuning the loop antenna, sufficient power to light to full candle power a series bank of 50 sixty-watt carbon lamps. Lowering or raising the frequency of 500,000 Hz resulted in a diminishing of the amount of power received through the 18 foot antenna.

Likewise, raising the transmission antenna resulted in a proportionate decrease of power picked up on the receiving antennae and at 6 feet above the earth no power whatsoever was obtainable without a

change of voltage and frequency.

It is the objective of my generic invention to utilise the power generated by the earth, by the means described here and illustrated in the drawings. The two drawings show simple and preferred forms of this invention, but I wish it to be understood that no limitation is necessarily made as to the exact and precise circuits, shapes, positions and structural details shown here, and that changes, alterations and modifications may be made when desired within the scope of my invention and as specifically pointed out in the claims.

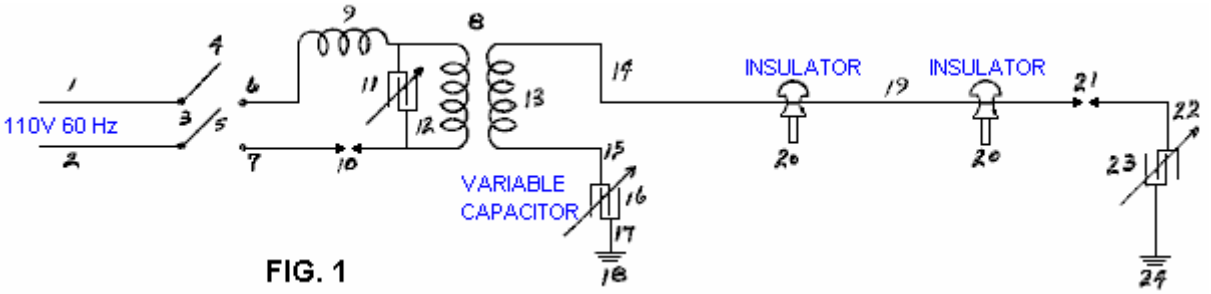


FIG. 1

Referring particularly to **Fig.1**, **1** and **2** are alternating current feed wires supplying 110 volts 60 cycles per second to a high-frequency generator. **3** is a switch with poles **4** and **5**, while **6** and **7** are the connections to the high-frequency transformer **8**, which is used to step-up the frequency to 500 kHz and the voltage to, say, 100 kV. **9** is an inductor, **10** is a spark gap, **11** is a variable capacitor, **12** is the primary winding and **13** the secondary winding of transformer **8**. The secondary winding is connected to ground through variable capacitor **16**, and wire **17**. Wire **14** connects transformer **8** to the main transmission antenna **19** which is supported along it's length on insulators **20**. Spark gap **21** is positioned between the main transmission antenna **19** and the ground **24**, passing through connecting wire **22** and variable capacitor **23**. The main transmission antenna **19**, can be any desired length.

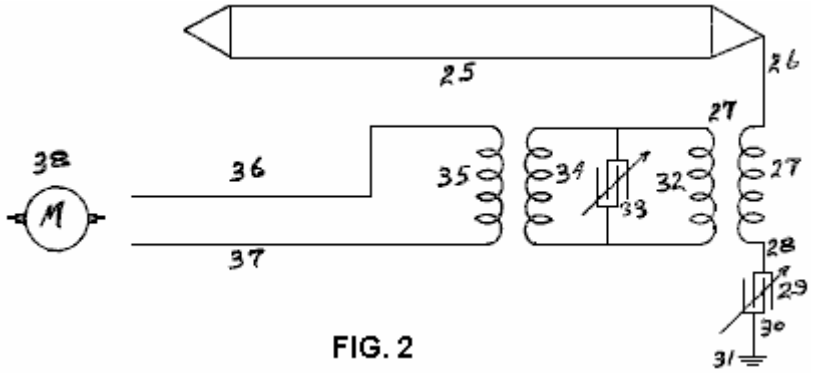
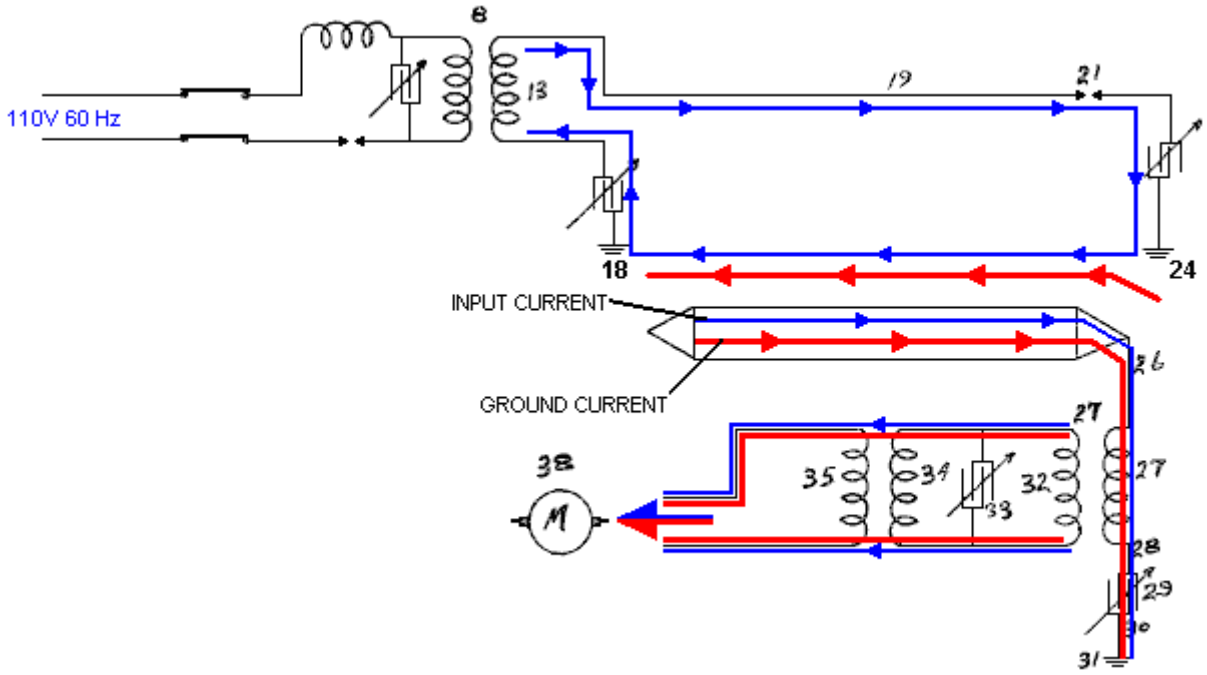


FIG. 2

In **Fig.2**, **25** is a closed oscillating loop antenna of any desired length. For greatest efficiency, it is run parallel with the main transmission antenna **19** of **Fig.1**. Wire **26** is connected to the secondary winding **27** of a step-down transformer which winding then goes to ground **31** through variable capacitor **29**. The primary winding **32** of the step-down transformer has variable capacitor **33** connected across it and it feeds directly into winding(s) **34** of frequency transformer(s) which supply current through winding(s) **35** to a motor "**M**" or other electrical load(s).

Having described the drawings, I will now describe the operation of my invention. Operate switch **3** to connect the input power. Adjust spark gap **10** and variable capacitor **11** so that 100,000 volts at a frequency of 500,000 cycles per second is delivered to step-up transformer **8** of **Fig.1**. Next, adjust spark gap **21** of the transmission antenna **19** so that all (voltage) peaks and nodes are eliminated in the transmission of the 100,000 volts along the antenna by the current surges across spark gap **21**. The high-frequency alternating current flowing through spark gap **21** passes through variable capacitor **23** to ground **24** and from there, back through the ground to earthing point **18**, through variable capacitor **16** and back to winding **13** of transformer **8**. As the 500,000 Hz

current is the same as the earth-generated currents and in tune with it, it naturally follows that accumulation of earth currents will amalgamate with those for transformer **8**, providing a reservoir of high-frequency currents to be drawn upon by a tuned circuit of that same 500 kHz frequency, such as that shown in **Fig.2**, where the antenna **25** is turned to receive a frequency of 500 kHz, which current then passes through transformer **27**, any frequency-adjusting transformer(s), and on to power the load(s) **38**.



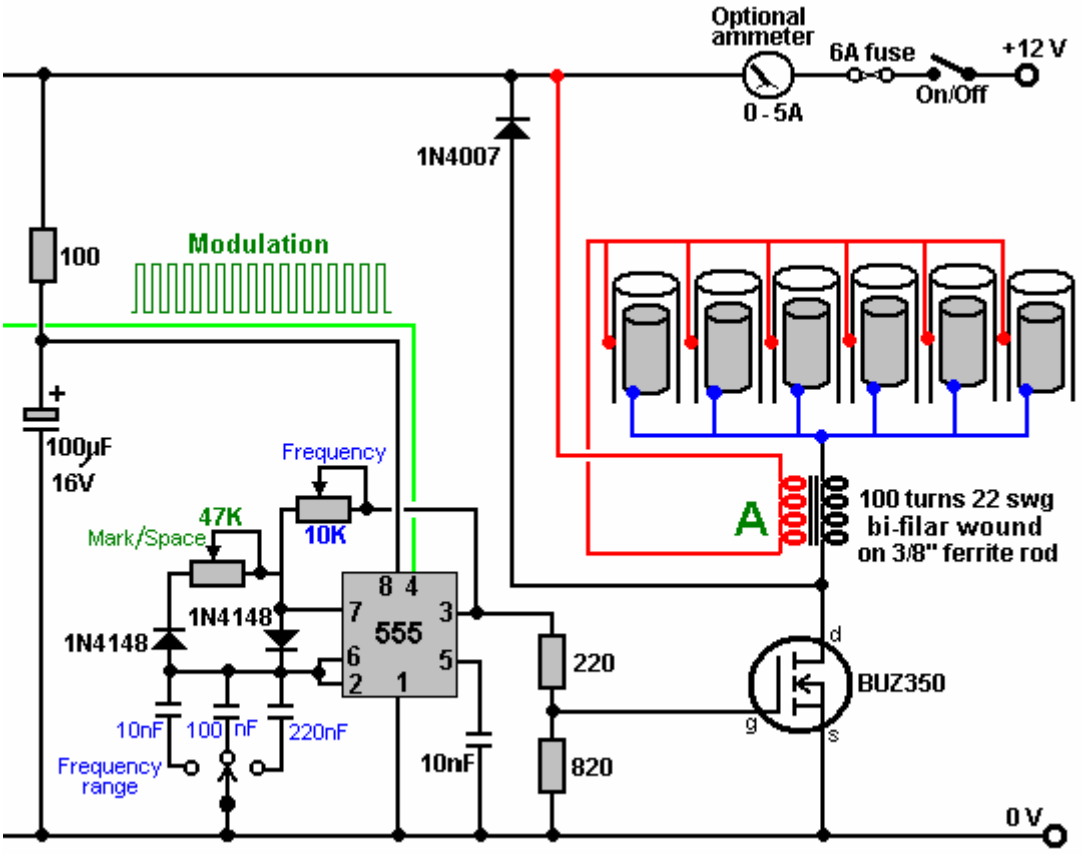
The return of current through the earth from transmission antenna **19**, is preferable to return through a wire as the ground return current picks up more earth currents than a wire does. I also prefer under certain conditions, to use a single antenna wire in place of the closed loop antenna shown in **Fig.2**. Under certain operational requirements, I have had improved performance by having the transmission antenna elevated and carried on poles many feet above the earth, and with that arrangement it is necessary to use a different voltage and frequency in order to accumulate earth currents.

This system of Frank's effectively applies very sharply pulsed DC pulses to a long length of wire supported in a horizontal position not far above the ground. The pulses are sharp due to both the spark gap on the primary side of the transformer, along with the spark-gap on the secondary (high voltage) side of the transformer. An input power of 500 watts gives a 3 kW power output from what appears to be an incredibly simple piece of equipment.

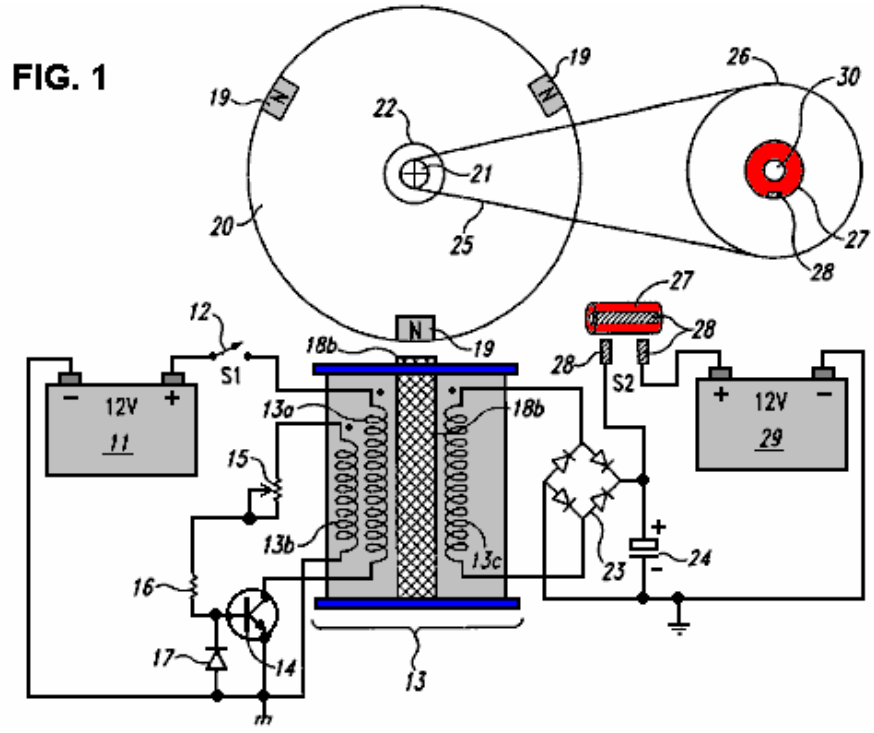
Dave Lawton's Solid-State Circuit.

A solid-state semiconductor circuit which has proved successful in producing pulses like this is shown as part of Dave Lawton's replication of Stan Meyer's Water Fuel Cell. Here, an ordinary NE555 timer chip generates a square wave which feeds a carefully chosen **F**ield-**E**ffect **T**ransistor the BUZ350 which drives a water-splitter cell via a combined pair of choke coils at point "A" in the diagram below.

Stan Meyer used a toroidal ferrite ring when he was winding these choke coils while Dave Lawton uses two straight ferrite bars, bridged top and bottom with thick iron strips. Chokes wound on straight ferrite rods have been found to work very well also. The effects are the same in all cases, with the waveform applied to the pipe electrodes being converted into very sharp, very short, high-voltage spikes. These spikes unbalance the local quantum environment causing vast flows of energy, a tiny percentage of which happens to flow into the circuit as additional power. The cell runs cold, and at low input current, quite unlike an ordinary electrolysis cell where the temperature rises noticeably and the input current needed is much higher.

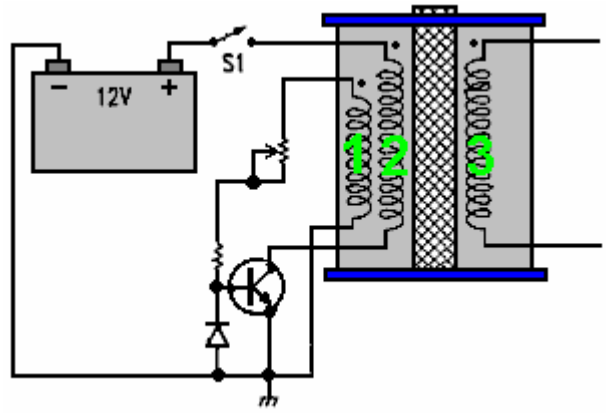


John Bedini's Battery-Charging Circuit.
John Bedini uses this same pulsing of a bi-filar wound coil to produce the same very short, very sharp voltage spikes which unbalance the local energy field, causing major flows of additional energy. The figure shown here is from his US patent 6,545,444.



John has produced and generously shared, many designs, all of which are basically similar and all using a 1:1 ratio bi-filar wound transformer. This one uses a free-running rotor with permanent magnets embedded in it's rim, to trigger sharp induced currents in the windings of the coil unit marked "**13b**" which switches the transistor on, powering winding "**13a**" which powers the rotor on its way. The pick-up coil "**13c**" collects additional energy from

the local environment, and in this particular circuit, feeds it into the capacitor. After a few turns of the rotor (dictated by the gear-down ratio to the second rotor), the charge in the capacitor is fed into a second "on-charge" battery.



The rotor is desirable but not essential as the coils marked **1** and **2** can self-oscillate, and there can be any number of windings shown as **3** in the diagram. Winding **3** produces very short, sharp, high-voltage spikes, which is the essential part of the design. If those sharp pulses are fed to a lead-acid battery (instead of to a capacitor as shown above), then an unusual effect is created which triggers a link between the battery and the immediate environment, causing the environment to charge the battery. This is an amazing discovery and because the voltage pulses are high-voltage courtesy of the 1:1 choke coils, the battery bank being charged can have any number of batteries and can be stacked as a 24-volt bank even though the driving battery is only 12 volts. Even more interesting is the fact that charging can continue for more than half an hour after the pulsing circuit is switched off.

It can be tricky to get one of these circuits tuned properly to work at peak performance, but when they are, they can have performances of COP>10. The major snag is that the charging mechanism does not allow a load to be driven from the battery bank while it is being charged. This means that for any continuous use, there has to be two battery banks, one on charge and one being used. A further major problem is that battery banks are just not suitable for serious household use. A washing machine draws up to 2.2 kilowatts and a wash cycle might be an hour long (two hours long if a "whites" wash and a "coloureds" wash are done one after the other which is not uncommon). During the winter, heating needs to be run at the same time as the washing machine, which could well double the load.

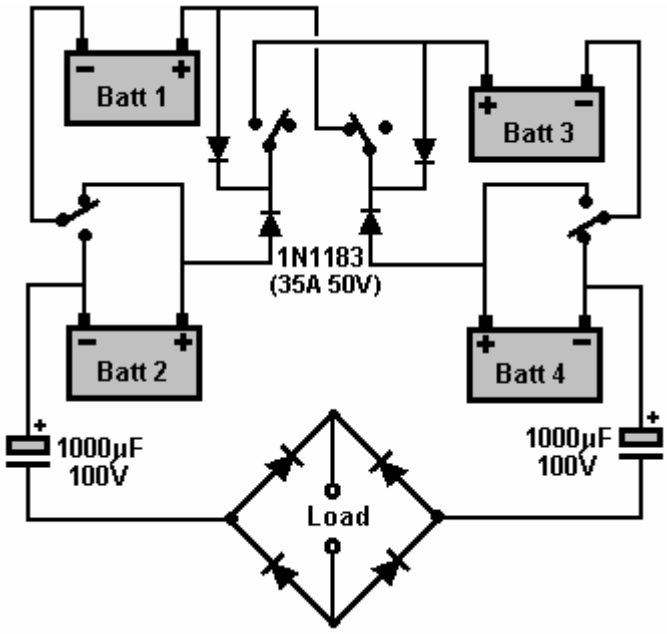
It is recommended that batteries are not loaded much beyond their "C20" rate, that is, one twentieth of their Amp-Hour nominal rating. Say that 85 Amp-Hour deep-cycle leisure batteries are being used, then the recommended draw rate from them is 85 Amps divided by 20, which is 4.25 amps. Let's push it and say we will risk drawing double that, and make it 8.5 amps. So, how many batteries would we need to supply our washing machine assuming that our inverter was 100% efficient? Well, 2,200 watts on a 12-volts system is 2,200 / 12 = 183 amps, so with each battery contributing 8.5 amps, we would need 183 / 8.5 = 22 large, heavy batteries. We would need twice that number if we were to treat them right, plus twice that again for household heating, say 110 batteries for an anyway realistic system. That sheer size of battery banks is not realistic for your average householder or person living in an apartment. Consequently, it appears that the Bedini pulse-charging systems are not practical for anything other than minor items of equipment.

However, the really important point here is the way that when these short pulses are applied to a lead-acid battery, a link is formed with the environment which causes large amounts of energy to flow into the circuit from outside. This is extra "free-energy". Interestingly, it is highly likely that if the pulses generated by Dave Lawton's water-splitter circuit shown above, were fed to a lead-acid battery, then the same battery-charging mechanism is likely to occur. Also, if a Bedini pulse-charging circuit were connected to a water-splitting cell like the Lawton cell, then it is highly probable that it would also drive that cell satisfactorily. Two apparently different applications, two apparently different circuits, but both producing sharp high-voltage pulses which draw extra free-energy from the immediate environment.

The Tesla Switch.

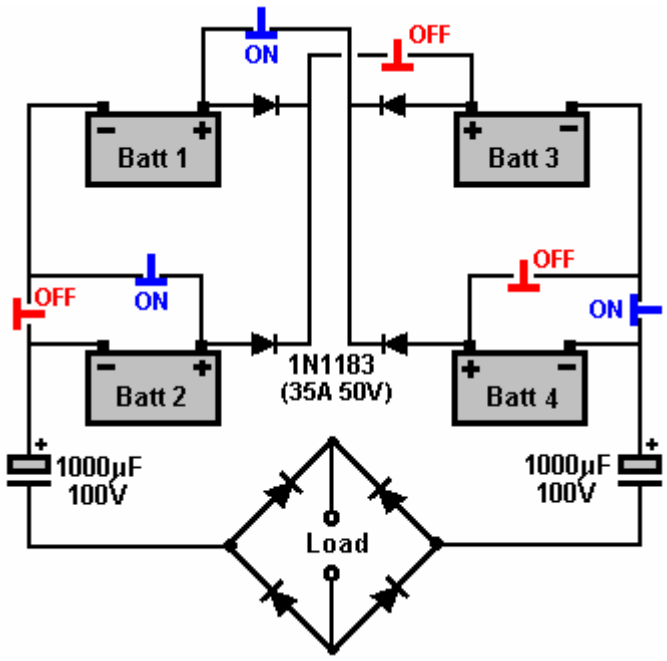
It doesn't stop there. Nikola Tesla introduced the world to Alternating Current ("AC") but later on he moved from AC to very short, sharp pulses of Direct Current ("DC"). He found that by adjusting the frequency and duration of these high-voltage pulses, that he could produce a whole range of effects drawn from the environment - heating, cooling, lighting, etc. The important point to note is that the pulses were drawing energy directly from the local environment. Leaving aside the advanced equipment which Tesla was using during those experiments and moving to Tesla's simple-looking 4-battery switch, we discover the same background operation of sharp voltage pulses drawing free-energy from the environment.

Consider the Electrodyne Corp. circuit (shown in "The Manual of Free-Energy Devices and Systems", 1986) tested by them for a period of three years:



Please note that when I shared this circuit diagram several years ago, someone persuaded me that the diodes were shown the wrong way round, and because of that, I have shown these diodes incorrectly. The diagram above is the one shown by the Electrodyne Corp. staff, and is correct.

As the switching used by this device was a mechanical device which has six switches where three are **ON** and three are **OFF** at any moment, the Electrodyne Corp. staff present the circuit diagram like this:



With switching like this:

It is recommended that this simple-looking circuit has an inductive load, preferably a motor, but consider the results of that very extended period of testing. If the switching rate and switching quality were of a sufficiently high standard, then the load could be powered indefinitely. The batteries used were ordinary lead-acid batteries, and after the three years of tests, the batteries appeared to be in perfect condition. Their tests revealed a number of very interesting things. If the circuit was switched off and the batteries discharged to a low level, then when the circuit was switched on again, the batteries returned to full charge in under one minute. No heating occurred in the batteries in spite of the massive charging rate. If the circuit was switched off and heavy current drawn from the batteries, then heat would be produced which is quite normal for battery discharging. The system operated lights, heaters, television sets, small motors and a 30-horsepower electric motor. If left undisturbed, with the circuit running, then each battery would charge up to nearly 36 volts with no apparent ill effects. Control circuitry was developed to prevent this over-charging. This, of course, is easy to do as all that is required is to place a relay across one battery and have it disconnect the circuit when the battery voltage reaches whatever voltage is considered to be a satisfactory maximum voltage.

These test results show spectacular battery charging and battery performance, quite outside the normal range associated with these ordinary lead-acid batteries. Are they being fed very short, very sharp pulses, like the previous two systems? It would look as if they were not, but one other very interesting piece of information coming from Electrodyne is that the circuit did not operate correctly if the switching rate was less than 100 Hz (that is 100 switchings in one second). The Electrodyne switching was done mechanically via three discs mounted on the shaft of a small motor.

One other detail reported by the Electrodyne testers, is that if the switching speed exceeded 800 times per second, that it was “dangerous” but unfortunately, they didn’t say why or how it was dangerous. It clearly was not a major problem with the batteries as they were reported to be in good shape after three years of testing, so definitely no exploding batteries there. It could well be as simple a thing that the voltage on each battery rose so high that it exceeded the voltage specifications of the circuit components, or the loads being powered, which is a distinct possibility. It is possible that at more than 800 pulses per second, the charging produced excessive cooling which was not good for the batteries.

It is generally accepted that for a circuit of this nature to work properly, the switching has to be very sudden and very effective. Most people have an immediate urge to use solid-state switching rather than the mechanical switching used by Electrodyne. A 'thyristor' or 'SCR' might be suitable for this, but the sharp switching of a PCP116 opto-isolator driving an IRF540 FET is impressive and a TC4420 FET-driver could substitute for the opto-isolator if preferred. It is possible that having a slight delay after the switches have turned On and Off, can prove very effective.

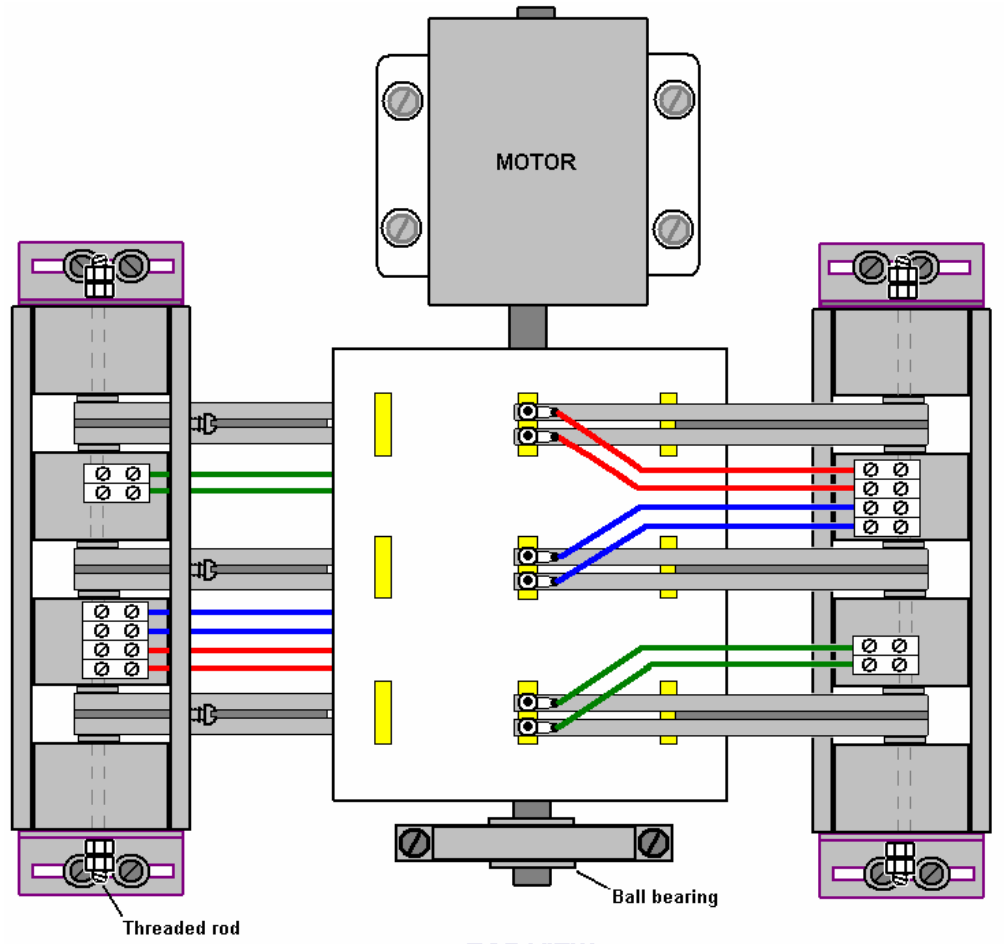
The Electrodyne Corp. staff used three identical discs mounted on the shaft of a motor as shown above. This allows the contact "brushes" to be located on opposite sides of the discs. There are, of course, many possible alternative constructions and I have been asked to show how I would choose to build this type of mechanical switching. The common idea of using mechanical relays is not very practical. Firstly, relays have trouble switching at the speeds suggested for this circuit. Secondly, with a contact life of say, two million and a switching speed of just 100 times per second, the relays would reach their projected lifespan after two weeks of operation, which is not a very practical option.

The objective is to have a simple construction which produces several switchings for each revolution of the motor, easy adjustment of the timing of two separate sets of three switches (one set being OFF when the other set is ON), a construction which can be taken apart and then assembled again without altering the timing, and an electrical connection method which is straightforward. Obviously, the construction needs to use components which are readily available locally, and ideally, only require simple hand tools for the construction.

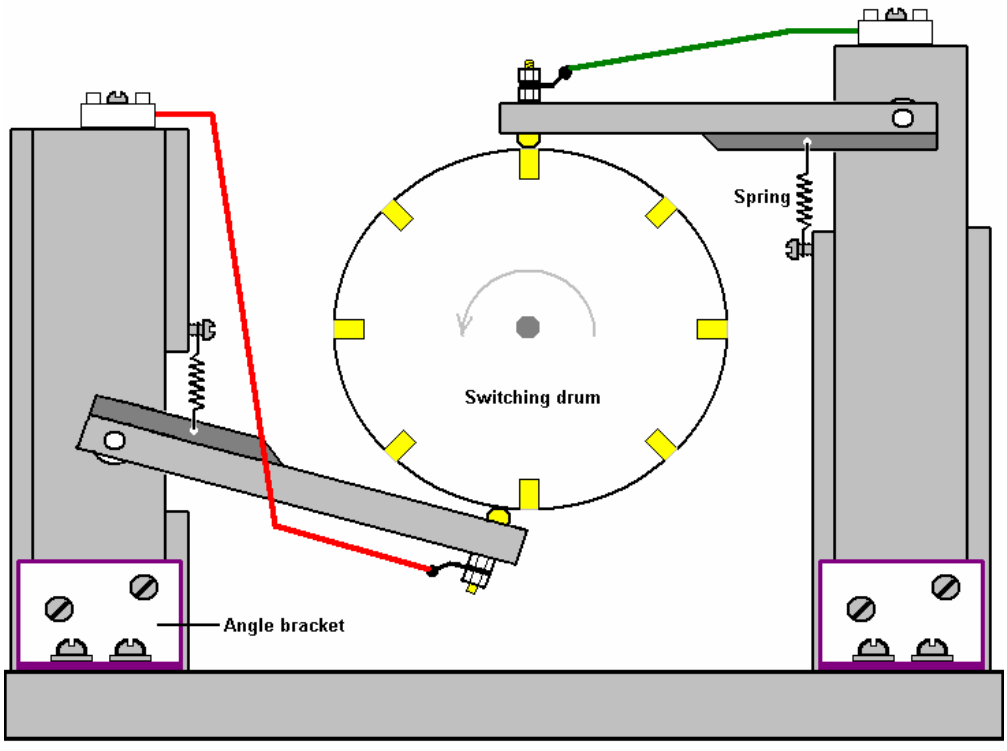
This suggested construction allows adjustment of the timing for both the start of the first set of switches and the start of the second set of switches. It should also be possible to introduce a short gap between the operation of these two sets of switches. This particular design is assuming a gap between each switching operation as that may be beneficial.

The switch contacts are rigid arms, pulled against the rotating drum by springs. The contacts touching the drum can be of various types and the ones shown are brass or copper cheese-head screws or bolts which are particularly convenient as they allow standard solder tags to be used to make the connections to the switch wires which then run across to ordinary electrical screw connectors, all of which can be accessed from above. I would suggest that four screw connectors should be used as a block as that allows them to be fastened in position with two screws which then stops them rotating when the wires are being tightened. There should not be any need for the conducting inserts in the switching cylinder to be particularly wide in the direction of rotation.

A practical construction method might be:



TOP VIEW



SIDE VIEW

The contact arms are shown as attached to each other in pairs. A lower level of construction accuracy can be allowed if they are all kept separate and a spring used for each arm rather than one spring for two arms as shown

in the drawing. I strongly recommend that the switching drum be solid and the brass or copper insets be a fair thickness and keyed securely into the drum. The surface of the inserts should be very gently eased into exact alignment with the surface of the drum, possibly by the very careful use of a small file or lathe if you are lucky enough to have access to one.

The pivots for all of the switching arms can be a length of threaded rod with lock nuts at each end. There should be almost no movement of the switching arms when the drum is spinning, so no enormous precision is needed for the holes in the switching arms, through which the threaded rod runs. Having said that, it must be understood that each switch in the set of three, must turn On and OFF at the same time, so the contacts on the spring-loaded arms must slide on to and off the conducting strips in the switching cylinder, at exactly the same time.

The drawing shows three conducting inserts at each of eight evenly-spaced positions around the circumference of the drum. The number around the drum is not critical although this suggestion gives eight switchings per revolution. If you choose to use a different number, you need to remember that the positioning of the arms underneath the drum will be different. You need to arrange it so that just after one set runs off its conducting strips that the other set slides on to it's conducting strips. Both sets of switches must not be ON at the same time as that short-circuits the batteries, which is probably not a good idea.

The timing adjustment is achieved by moving the supporting block slightly, by easing the four clamping screws, sliding the block and tightening the screws again. This, of course, is done when the drum is not rotating.

Each set of six switching arms needs to have all of the arms exactly the same length between the sliding contact (shown as a bolt head) and the pivot hole. Each of the conducting strips inset into the drum, need to be aligned exactly and be exactly the same width, otherwise the switching action will be ragged and not properly synchronised.

The supports for the switching arms can be either a single block with slots cut in it or the easier construction shown, where it is fabricated from several standard rectangular pieces and glued and/or screwed together.

The unequal amount of conducting strip compared to the non-conducting part means that there will be a timing gap between each pair of On/Off switchings. In spite of that, the battery switching will be a 50% Duty cycle as required. The switching sequence will then be: On / Off / Pause, On / Off / Pause, On / Off / Pause ….. and that may well be a desirable arrangement as having an inter-pulse delay can be very good for battery charging.

However, please don't imagine that the Tesla Switch described here is a 'plug-and-play' device which you can switch on and it will give you the sort of outputs mentioned above, as that is very much not the case. You need to see the Tesla Switch as being a long-term development project with high potential.

If you use the Tesla Switch circuit with manual switches and run each phase for many minutes before altering the switching, it can give up to four times better performance than running the load off the four batteries in parallel. That is not what the Tesla Switch is all about.

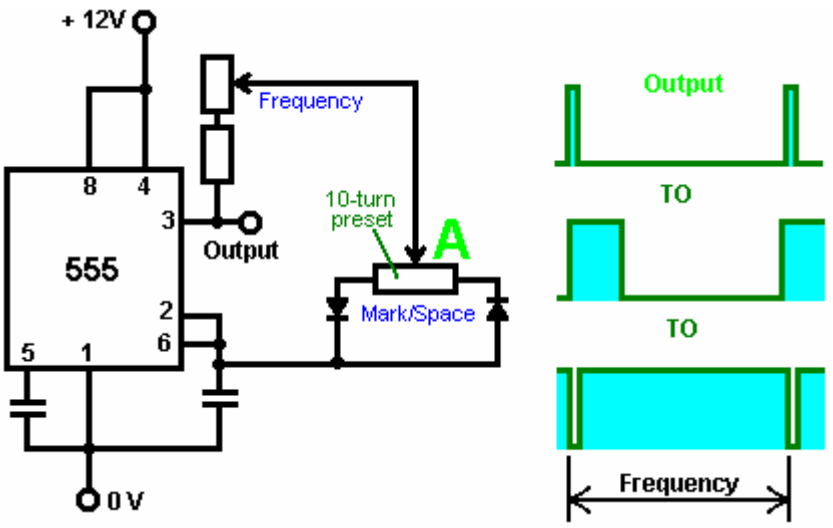
The Tesla Switch is one of the more difficult devices to get operational, in spite of the fact that it appeals to a large number of people. There are three possible modes of operation. If the diodes are turned the wrong way round so that they can feed current from each battery, then the operation will definitely be COP<1 but it will be a good deal better than operating without the switch circuit in place.

The second way has only been achieved by John Bedini as far as I am aware. This is where the circuitry is the same but the circuit components and connecting wires are adjusted very carefully to produce circuit resonance. When that happens, the circuit becomes self-powering although there is little or no extra power for other devices.
The third way was developed and tested over three years by staff of the Electrodyne Corporation in America. In this version, the diodes are reversed and they only feed sharp voltage spikes back to the batteries, through the diodes which supposedly don't allow current to flow in that direction. This is a very different form of operation where the operating power flows into the circuit from the local environment. The batteries need to be 'conditioned' through long periods of being operated this way as the 'cold electricity' used in the circuit is the opposite of the 'hot electricity' which the batteries have been using up to now. This long conditioning period is generally enough to make the average builder give up and believe that the circuit just doesn't work. Dave Lawton was faced with exactly the same type of problem when he attempted to replicate Stan Meyer's "Water Fuel Cell". It appeared 'dead' and produced nothing during a whole month of testing, and then it suddenly burst into life, producing large amounts of HHO gas mix for almost no electrical input. Without his exceptional patience, Dave would never have succeeded. I believe that the same applies to the Tesla Switch when wired correctly with the diodes blocking current flow from the batteries – it is likely to take long-term and patient testing before the system swings into life.

One experimenter who did not believe the diodes could possibly work that way round, tested the arrangement and discovered that in spite of the theory, in practice, the reverse-biased diodes actually pass very sharp voltage

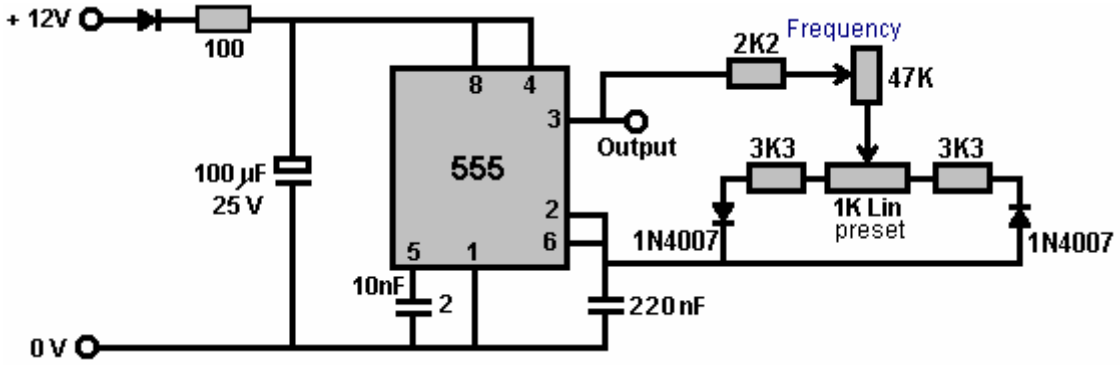
spikes to the batteries, so the effect may well be like a slick version of the various battery pulsing circuits shown in chapter 6.

Mechanical switching appears to work very well indeed, but if we decide to try using electronic circuitry, then we need to get an exact 50% Mark/Space ratio using a switching circuit, and so the following style of circuit might be used with a multi-turn preset resistor in position “A”:



Here, the frequency is not noticeably affected by adjustment through a very wide range of Mark/Space settings. The output from Pin 3 needs to drive a very sharp switching combination such as a TC4420 FET driver connected to IRF540 FETs.

Perhaps the circuit might be something like this:



This circuit allows the Mark-Space ratio to be adjusted without altering the frequency, and the frequency can be adjusted without affecting the Mark-Space setting in any way. In the Tesla Switch circuit, three switches need to be in their On position and the other three switches in their Off position, so we will arrange this by using the ordinary NE555 timer circuit shown above, with it's adjustable Mark-Space ratio (that is, variable On-to-Off ratio). We will use this circuit to drive six opto-isolators which will turn the six transistors On and Off in groups of three as required. To get the very high switching speed needed, PCP116 opto isolators should be used and although these are difficult to find, every effort should be made to get them as they enhance the switching speed.

Variable resistors come in a wide range of types. It is probably best to use a preset type as they are very easy to adjust and hold their settings very solidly. Also, when the correct setting is found, the component will be left on that position permanently. Some common types are:

where some can be adjusted from the top and others adjusted from the side. All of them can be mounted directly on the strip-board or printed circuit board used to construct the circuit.

However, the problem is to decide the direction of current flow and provide solid state components accordingly, as the Tesla Switch circuit almost certainly does not run with conventional electronic design. If you were to reverse the diodes shown in the first circuit diagram in this section, then the circuit will remain solidly COP<1 although some people have managed an operational improvement of 32 times over just using the batteries straight to power the load. With the diodes as shown in the first two diagrams in this section, the circuit operates by drawing in energy from the environment and that operates in an entirely different way in a circuit.

It is interesting to note that in the 1989 patent US 4,829,225 granted to Yury Podrazhansky and Phillip Popp, their evidence is that batteries charge much better and have a longer life if they are pulsed in a specific way. Their formula is that the battery should be given a powerful charging pulse lasting for a period of time between a quarter of a second and two seconds, the pulse being the Amp-Hour rating of the battery. That is, for an 85 AHr battery, the charging pulse would be 85 amps. That pulse is then followed by a discharging pulse of the same, or even greater current but only maintained for only 0.2% to 5% of the duration of the charging pulse. Those two pulses are then followed by a resting period before the pulsing is repeated. They quote the following examples of their experiences when using this method:

Battery:	**9V alkaline**	**1.25V NiCad 0.5 AHr**	**15V NiCad 2 AHr**	**12V lead-acid 40 AHr**
Charging current:	0.5 Amps	1.2 Amps	3.0 Amps	48 Amps
Charging period:	550 mS	700 mS	500 mS	850 mS
Discharge current:	6 Amps	6 Amps	14 Amps	85 Amps
Discharge period:	2 to 3 mS	2 mS	2 mS	3 mS
Rest period:	15 to 20 mS	7 to 10 mS	10 mS	15 mS
Charging range:	50% to 100%	20% to 100%	20% to 100%	20% to 100%
Total charging time:	12 to 15 mins	20 mins	35 to 40 mins	40 mins

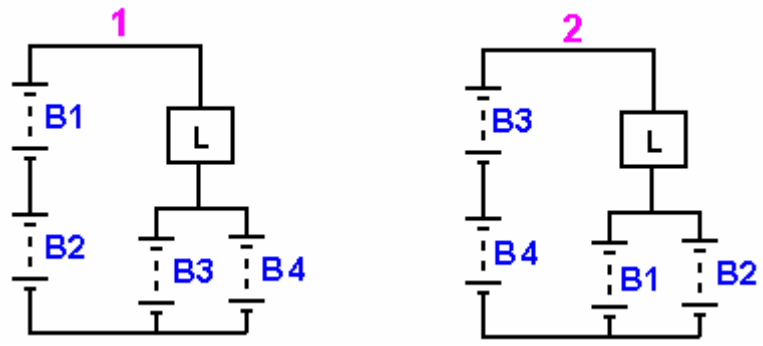
Interestingly, this appears to confirm the charging potential of the Tesla Switch style of operation, especially if there is a short rest period between the two sets of switch operations.

A Three Battery Switching System

Continuing the Tesla Switch style of operation, it is possible to get the same effect as the Tesla Switch circuit, using only three batteries (or three capacitors). Discussed almost a century ago by Carlos Benitez in his patents, and more recently described by John Bedini, just three batteries can be used if more complicated circuit switching is used. Carlos points out that there has to be an energy loss due to wires heating up and batteries not being 100% efficient. He overcomes these problems with some very clever circuitry which is covered in the following section. However, it is not at all certain that this is actually the case as experimentation indicates that it is possible for this kind of battery switching to maintain the battery charge levels far beyond the expected.

Here is an untested suggestion for how it might be possible to produce a portable, self-powered powerful light. There are many possible variations on this, and the following description is just intended as an indication of how a three battery switching system might be constructed. If you are not familiar with simple electronics, then I suggest that you study the basic electronics tutorial of chapter 12.

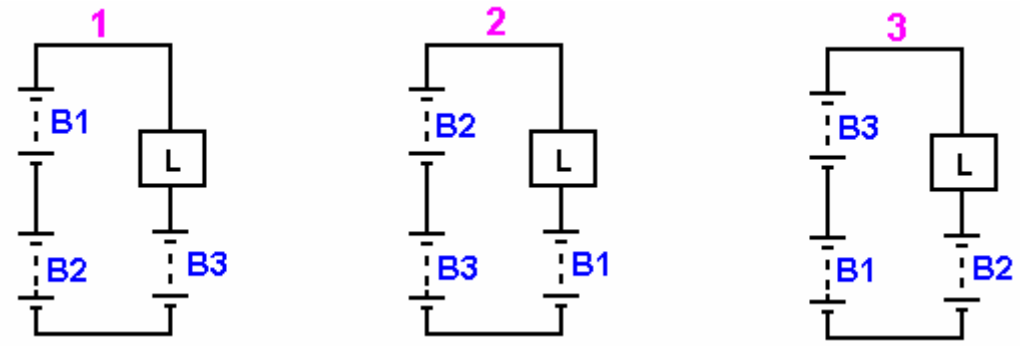
Battery charging can be accomplished in various different ways. Obviously, the more the electrical loading can be reduced, the lesser the need for recharging. Two methods for doing this involve passing the same electrical current repeatedly through the load, as shown here:



The switching for this arrangement can be implemented in various different ways, but essentially, in Stage 1, batteries ‘B1’ and ‘B2’ provide twice the voltage of batteries ‘B3 ‘ and ‘B4’, causing current to flow through the load ‘L’ and into batteries ‘B3’ and ‘B4’, driven by the voltage difference which is normally, the same as the voltage of any one of the batteries on its own. Each of the batteries ‘B3’ and ‘B4’ receive only half of the current supplied by batteries ‘B1’ and ‘B2’, and so, there is, not surprisingly, an energy loss. However, for half of the time, batteries ‘B3’ and ‘B4’ are receiving charging current instead of supplying current to the load.

In Stage 2, the batteries are swapped around and the process repeated with batteries ‘B3’ and ‘B4’ supplying current to the load and batteries ‘B1’ and ‘B2’. Tests have shown that with this arrangement, the load ‘L’ can be powered for longer than if all four batteries were connected in parallel and used to supply the load directly. With this system, each battery receives half of the load current for half of the time.

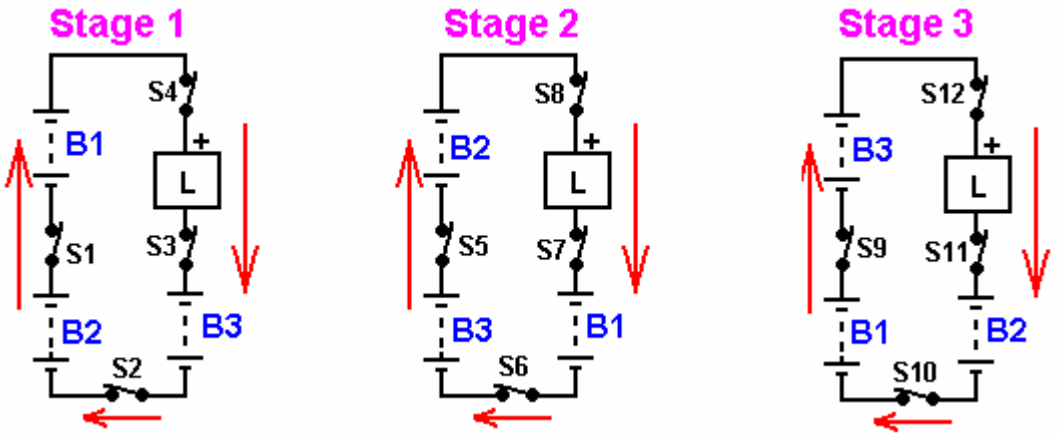
An alternative method which uses the same principle, but three batteries instead of four, and where each battery receives all of the load current for one third of the time, is like this:



Here, the batteries are switched around sequentially, with two of them in series causing current flow through the load ‘L’ and into the third battery. There is, of course, an overall energy loss, and so, additional energy from an external source needs to introduced to keep the load powered continuously. However, as with the four-battery system, the load ‘L’ can be kept powered longer by the batteries arranged like this than would occur if all three batteries were connected in parallel and used to power the load directly.

As before, the switching for a system of this kind can be implemented in various different ways. For long-term reliability, solid-state switching is preferred, and as NPN transistors are low-cost and readily available, they are shown here in one of the preferred configurations:

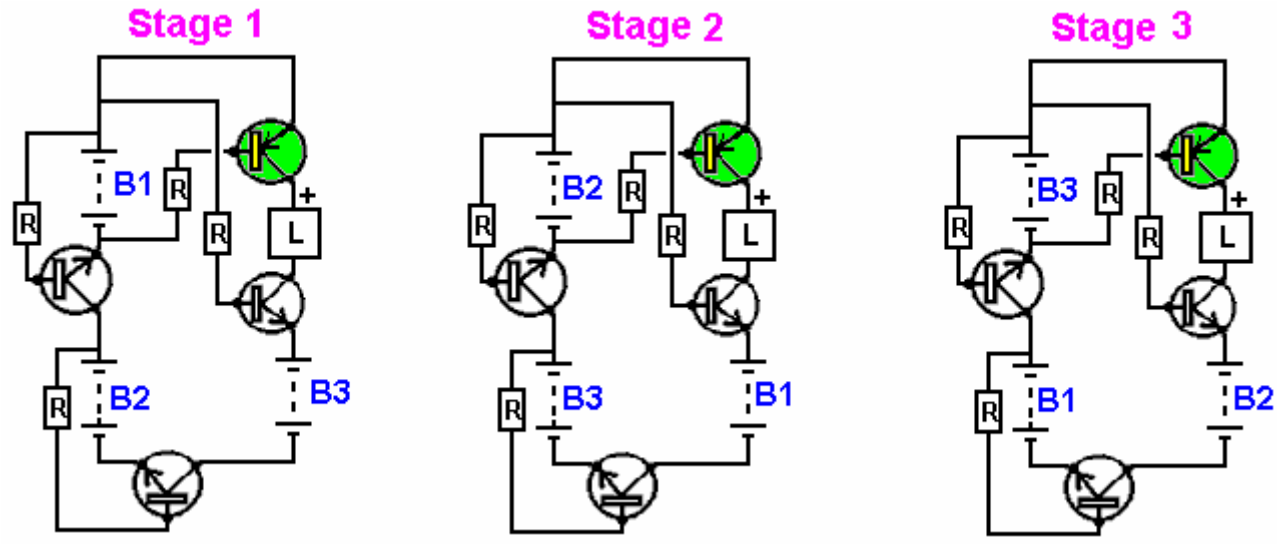
As each inter-battery connection is different for each of the three stages of operation of this circuit, it is necessary to have four switches for each stage. In order to establish the necessary details for connection of the transistors, as this circuit does not have the normal positive and negative rails, the (nominal) current flow directions need to be examined. These are shown here:



Obviously, the current flows from the higher series-connected voltage to the lower single battery voltage. The twelve virtual switches are numbered from 'S1' to 'S12' respectively, and if each represents an NPN transistor, then we also need to ensure that the current flow direction is correct for the transistor and to identify a higher voltage point which can be used to feed current into the base of each transistor. These details are listed here:

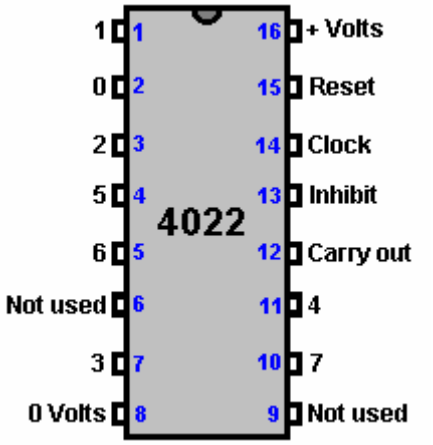
Stage	Switch	Type	Collector	Base resistor	Emitter
1	1	NPN	B2 Plus	B1 Plus	B1 Minus
	2	NPN	B3 Minus	B2 Plus	B2 Minus
	3	NPN	Load Minus	B1 Plus	B3 Plus
	4	PNP	Load Plus	B1 Minus	B1 Plus
2	5	NPN	B3 Plus	B2 Plus	B2 Minus
	6	NPN	B1Minus	B3 Plus	B3 Minus
	7	NPN	Load Minus	B2 Plus	B1 Plus
	8	PNP	Load Plus	B2 Minus	B2 Plus
3	9	NPN	B1 Plus	B3 Plus	B3 Minus
	10	NPN	B2Minus	B1 Plus	B1 Minus
	11	NPN	Load Minus	B3 Plus	B2 Plus
	12	PNP	Load Plus	B3 Minus	B3 Plus

The suggested switching arrangement therefore, looks like this:

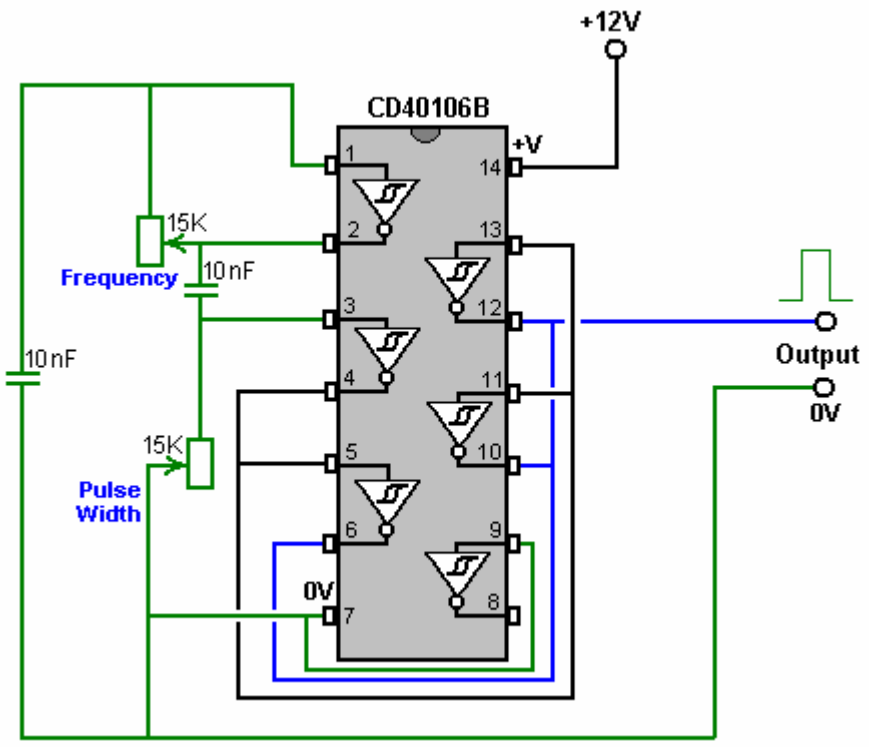


While the above diagram shows each stage with permanently connected base resistors, that is, of course, only to display the conceptual arrangement. Each resistor is passed through an opto-isolator and each set of four opto-isolators are driven by one of three separate outputs of equal duration. One possible arrangement for this could be as indicated below.

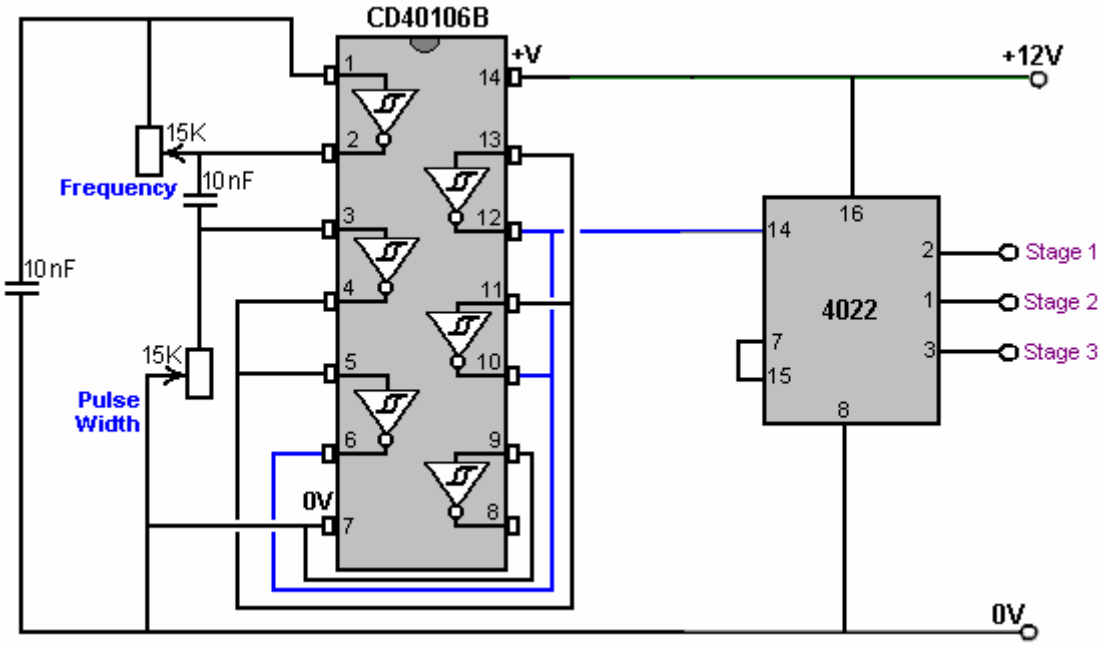
The CD4022 Divide-by-Eight chip can be arranged to divide by three rather than eight, by connecting its pin 7 to pin 15. The physical chip connections are:



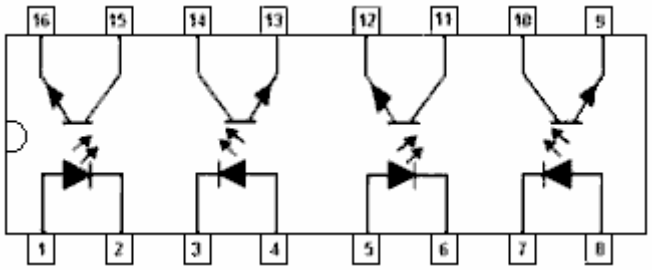
The chip needs a clock signal in order to function. There are many different ways of generating a clock signal, and the one shown here is very cheap, simple and has adjustable frequency and adjustable Mark/Space ratio, although, as the signal is to be used to trigger the action of a Divide-By-Three chip, there is no need for this clock signal to have a 50% Mark/Space ratio. The chip supply current is so tiny, that it really does not matter what the Mark/Space ratio is:



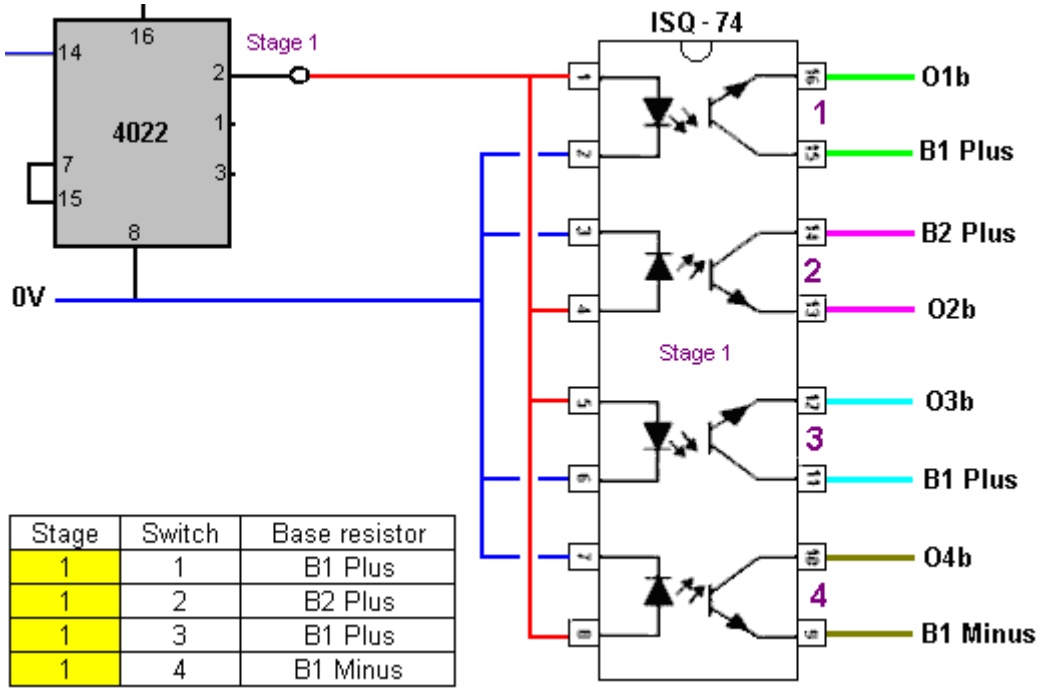
Using this circuit as the clock signal, the opto-isolator circuit could be:



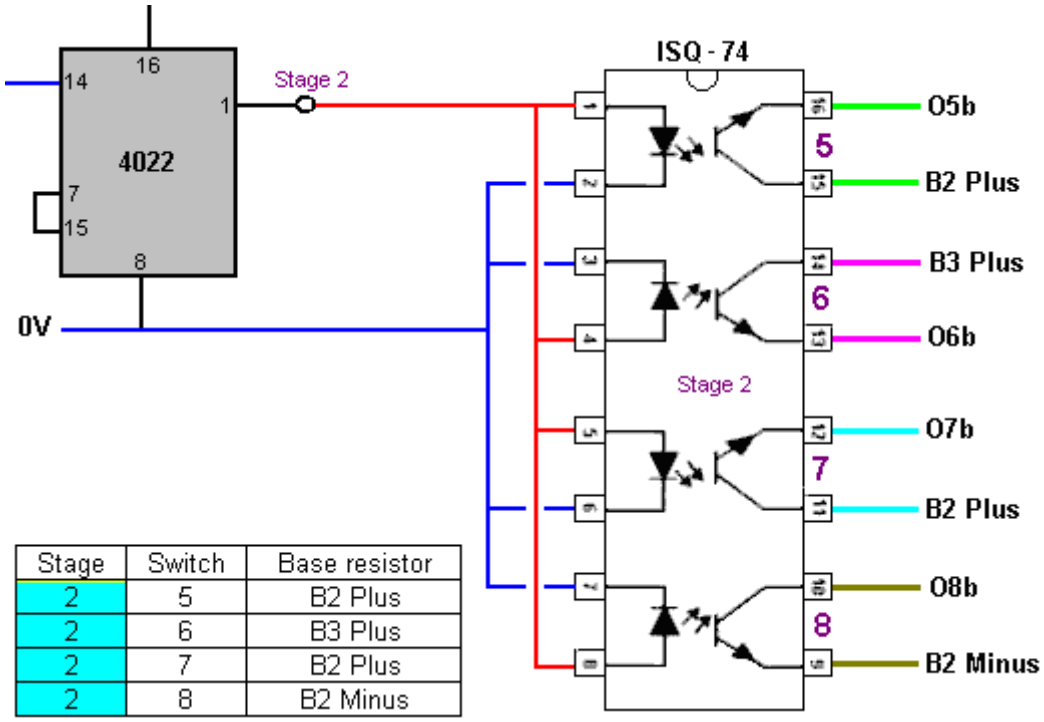
There are various opto-isolators available and while the rather expensive high-speed varieties are tempting, since we have to provide three sets of four, the ISQ-74 quad chip seems very suitable for this application although it is slower:



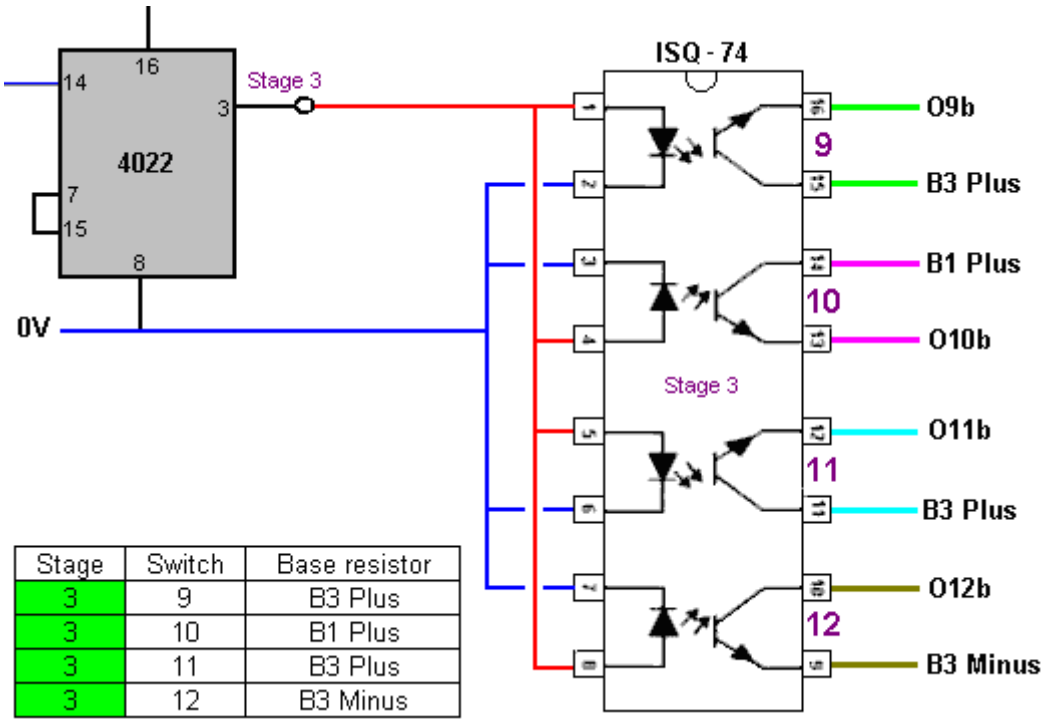
The overall circuit for the opto switching is then:



Stage	Switch	Base resistor
1	1	B1 Plus
1	2	B2 Plus
1	3	B1 Plus
1	4	B1 Minus



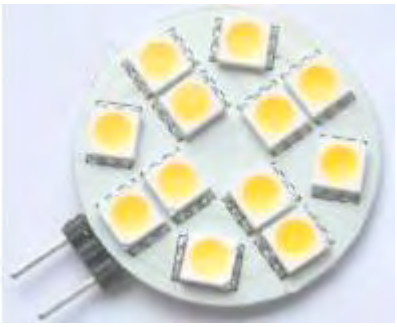
Stage	Switch	Base resistor
2	5	B2 Plus
2	6	B3 Plus
2	7	B2 Plus
2	8	B2 Minus



Stage	Switch	Base resistor
3	9	B3 Plus
3	10	B1 Plus
3	11	B3 Plus
3	12	B3 Minus

The output transistors are expected to switch 1 amp and so the TIP132 NPN and matching TIP137 transistors have been selected. These are cheap, Darlington transistors with current gains in excess of 1000 which means that the base current requirements are about 1 milliamp, which suggests that the base transistors could be 8.2K for a 12V system. These transistors can switch 12A at up to 100V and have a power dissipation of 70 watts, indicating that they will be running so far below their capability that they should run cool.

With this kind of circuit, it is desirable to have a fairly large current flow (relative to the battery capacity) in order to give a marked difference between the discharging and charging cycles for each battery. One possible load for this circuit could be the potentially low-cost G4 LED lighting unit shown here:



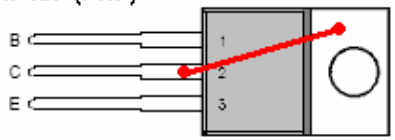
This quite remarkable device has a massive 160 lumens output with a lighting angle of 160 degrees and an input power of just 1.2 watts (100 milliamps at 12 volts). That light output is startling when seen in a completely dark place and some heat is generated, which is unusual for LEDs. A 100-watt filament light bulb has a light output of about 1600 lumens and a lighting angle of 360 degrees (some of which is normally not used usefully), and so, for the equivalent level of lighting, we would need ten of these G4 LED units, with a total power input of about 1 amp at 12 volts. Checking this in dark conditions, suggests that a much lesser lighting level would be more than acceptable. These LED lighting units are available in pure white and in “warm white” versions, and it is probable that the “warm white” version would suit most people better than the pure white versions.

Tests on one of the pure white units, shows a quite remarkable variation in the light output as the current is reduced, as the effect is highly non-linear:

At 12 volts the current is 1 amp. At 9.8 volts the current has dropped to a tiny 23 milliamps and while the light is nothing like as bright, there is still a considerable amount of light. At 9.42 volts, the light level is still significant and the current has dropped to a mere 10 milliamps. This suggests two main options: 10 LED units at 12 watts, providing a massive lighting level, or perhaps the same LED units run at 9.42 volts for just 1.2 watts of input power.

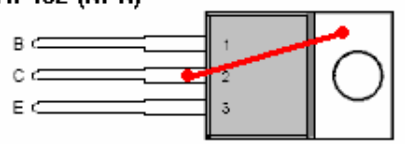
Using a 104 mm x 50 mm board size which will slot directly into a standard plastic slotted-side box, a stripboard layout (where the red circles indicate a break in the copper strip on the underside of the board) for the transistor switching section might be:

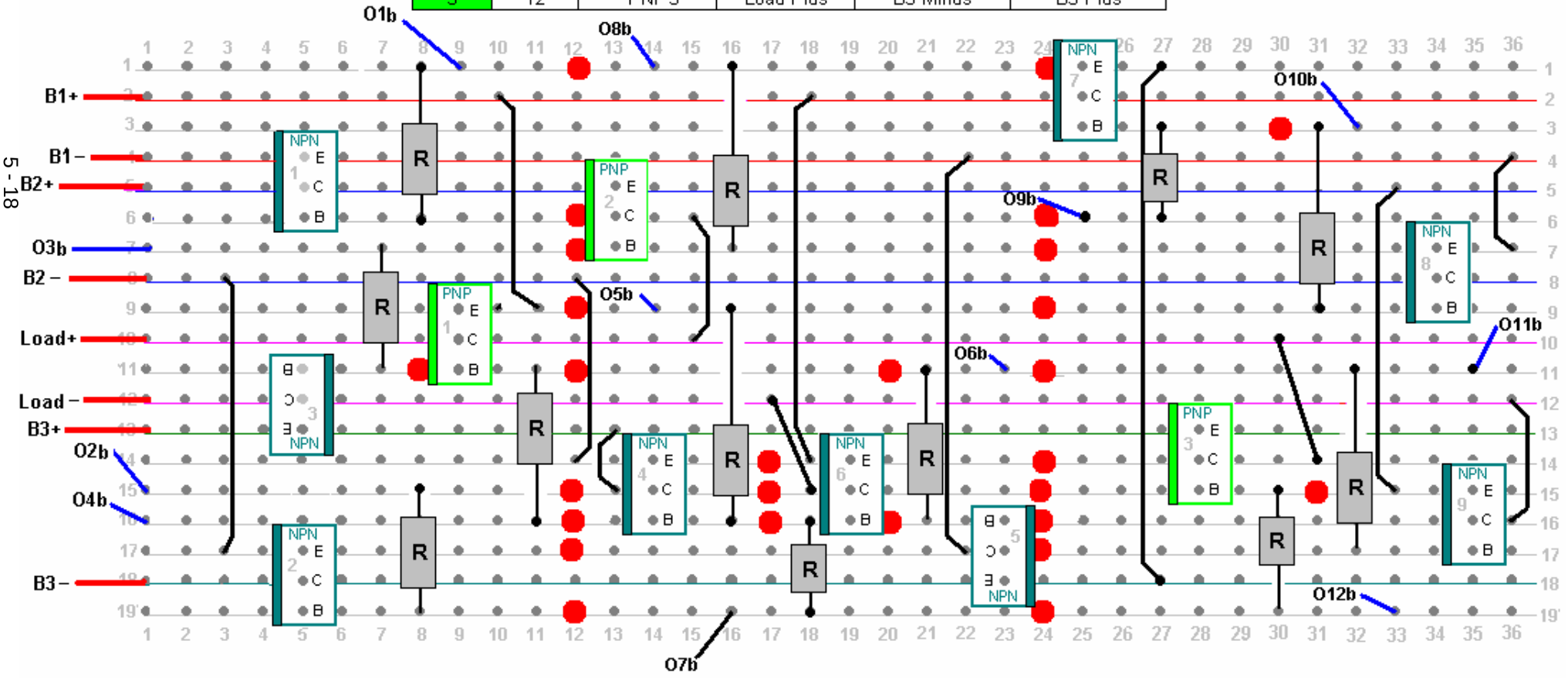
TIP137 (PNP)



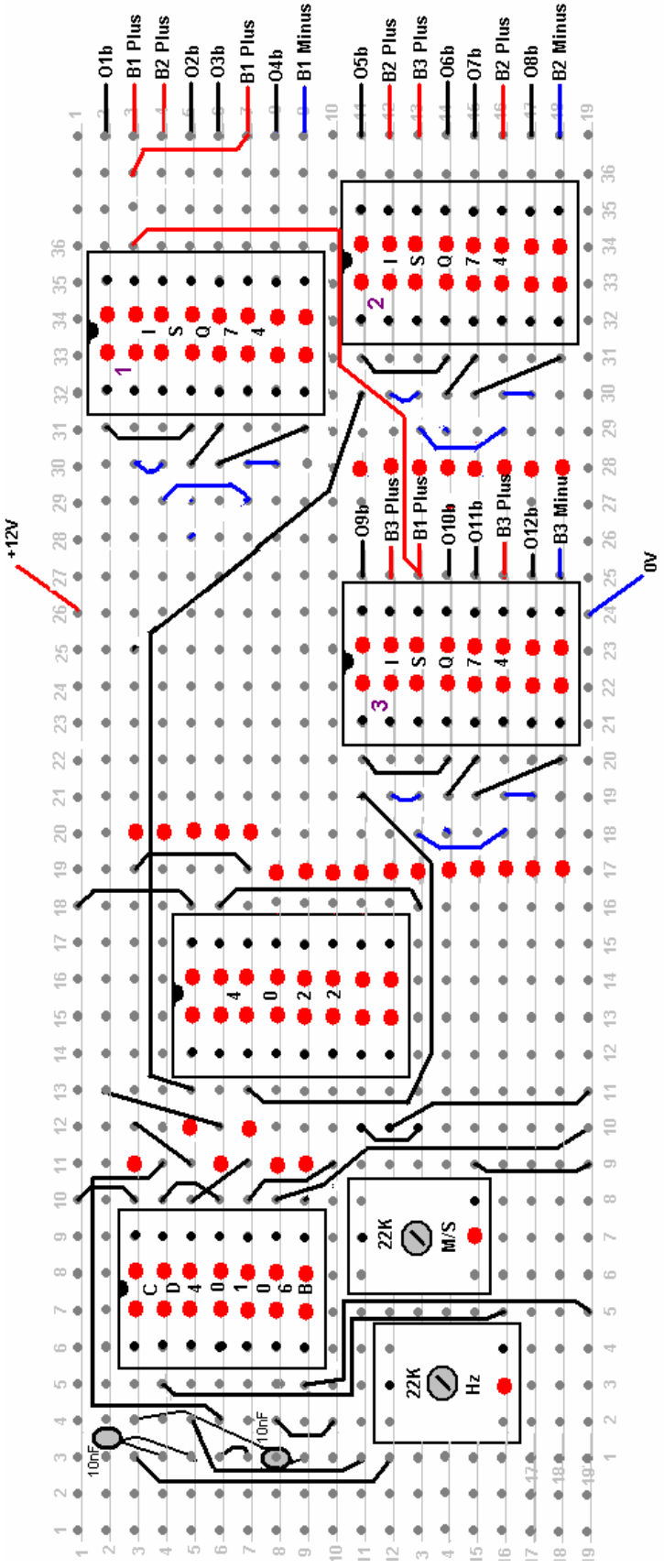
Stage	Switch	Type	Collector	Base resistor	Emitter
1	1	NPN 1	B2 Plus	B1 Plus	B1 Minus
1	2	NPN 2	B3 Minus	B2 Plus	B2 Minus
1	3	NPN 3	Load Minus	B1 Plus	B3 Plus
1	4	PNP1	Load Plus	B1 Minus	B1 Plus
2	5	NPN 4	B3 Plus	B2 Plus	B2 Minus
2	6	NPN 5	B1Minus	B3 Plus	B3 Minus
2	7	NPN 6	Load Minus	B2 Plus	B1 Plus
2	8	PNP2	Load Plus	B2 Minus	B2 Plus
3	9	NPN 7	B1 Plus	B3 Plus	B3 Minus
3	10	NPN 8	B2Minus	B1 Plus	B1 Minus
3	11	NPN 9	Load Minus	B3 Plus	B2 Plus
3	12	PNP3	Load Plus	B3 Minus	B3 Plus

TIP132 (NPN)





Each base resistor has an output link (O1b through O12b) which is connected through it's opto-isolator to the destination shown in the "Base" column in the table. Each set of three NPN transistors and one PNP transistor are switched together via a single ISQ-74 quad opto isolator chip. Each of the three ISQ-74 chips is powered in turn by one of the outputs from the CD4022 Divide-by-Three connected chip, which driven by the CD40106B hex Schmitt inverter chip wired as a clock as shown above. It is expected that a suitable clocking frequency would be about 700 Hz. A possible layout for the clock, Divide-by-Three and twelve opto-isolators on a 104 mm x 50 mm strip board, is shown here:



The timing and switching circuits form part of the load which is being switched. However, if we assume that there will be a power loss when running this system, then we should consider the very clever designs of Carlos Benitez in 1915.

The Self-powered Free-Energy Generators of Carlos Benitez

The Mexican Civil Engineer Carlos Benitez devised what is essentially the 3-battery switch discussed above. He was working at a time when solid-state electronics was not available and so his design is all the more impressive for that. Here is some of his patent information:

Carlos Benitez Patent GB 17,811 13th May 1915

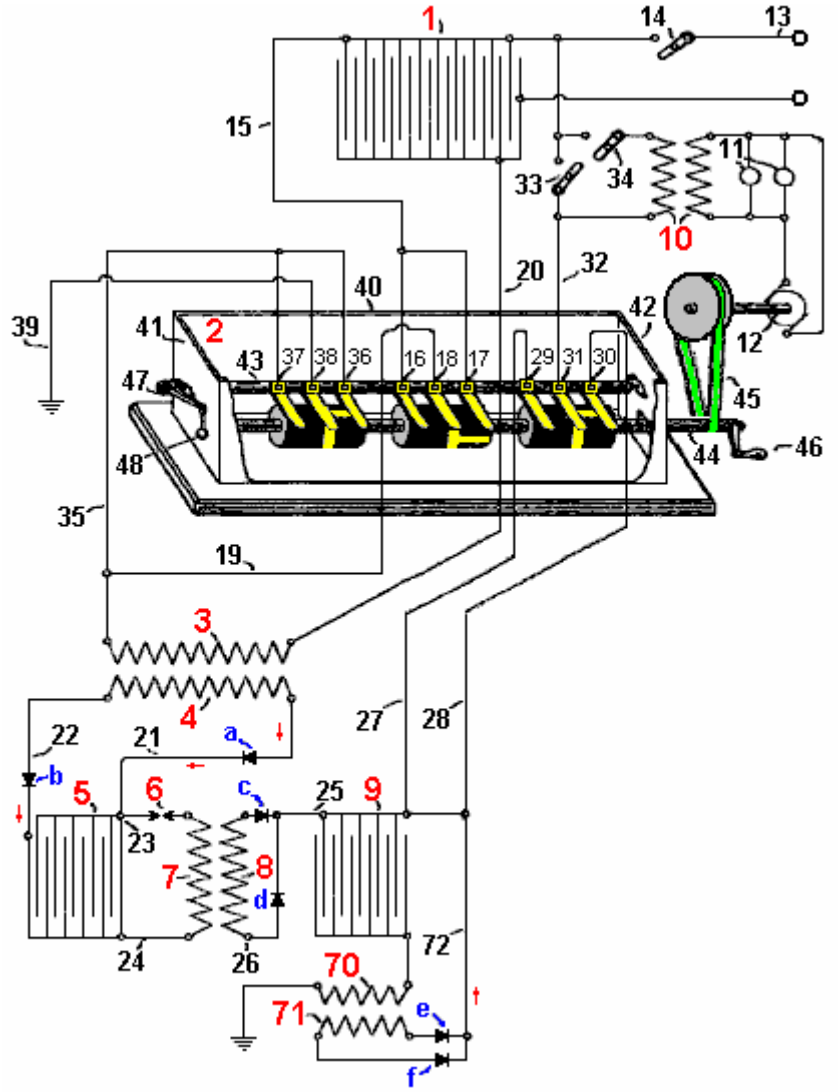
System for the Generation of Electric Currents

I, Carlos F. Benitez, Civil Engineer, 141 Ocampo Street, Guadalajara, Mexico, do hereby declare the nature of this invention:

The invention relates to a new process for obtaining electric currents under unusually simple, economic and practical conditions. I use synthetically in combination: apparatus for the production of electric currents for charging one or several capacitors whose collectors or inner coatings are connected to one of the terminals of the primary winding of one or several induction coils and the other coatings of those capacitors are alternately grounded through the primary winding of a transformer, or connected through the transformer primary to the above collectors, appropriate means for collecting the currents produced in the secondary windings of these transformers and for applying charging to the above collectors, and an appropriate means for discharging those collectors, and for the application of all or part of its energy to the successive replication of the process already described, in this way, increasing the generation of electrical energy, or maintaining a constant, pre-determined electrical output.

Furthermore, the invention consists of a new combination of parts from which are derived advantages which will be fully understood by the consideration of the two different cases illustrated in the accompanying drawing, in which:

1 is a bank of capacitors.
2 is a rotating commutator to make and break the circuit connections at the appropriate instants.
3 is the primary winding of a transformer or induction coil.
4 is the secondary winding of that transformer or induction coil.
5 is a second bank of capacitors.
6 is an oscillator spark gap.
7 is the primary winding of a second transformer.
8 is the secondary winding of that transformer.
9 is a third bank of capacitors.
70 is the primary winding of a third transformer
71 is the secondary winding of that transformer.
10 is a fourth transformer.
11 is a series of incandescent lamps.
12 is an electric motor.
(a), (b), (c), (d), (e) and **(f)** are mercury-vapour converters or cathodic valves, allowing electric current flow only in the direction shown by the arrows.



The capacitor bank **1**, is connected through wire **13** to a source of electric current, providing the initial charge to capacitor bank **1**. This initial charge is used to start the system running and can be disconnected at any time by means of switch **14**.

Wire **15** connects the inner leaves of capacitor bank **1** with poles **16** and **17** of the commutator **2**, and it's pole **18** is connected via wire **19** to one of the terminals of transformer primary winding **3**, whose other end is connected through wire **20** to the outer leaves of capacitor bank **1**. The secondary winding **4**, of this transformer, is connected by wire **21** to the inner plates of capacitor bank **5**, and by wire **22**, to the outer plates of capacitor bank **5**. In the same manner, wires **23** and **24** pass those connections on to the two sides of the primary winding **7** of the second transformer. Wire **23** also contains an oscillator spark gap **6**, and wires **21, 22, 25, 26** and **72** contain the one-way cathode valves **a, b, c, d, e** and **f**. The secondary winding **8**, of this second transformer, connects to the inner plates of capacitor bank **9**, whose outer plates are connected to ground through the primary winding **70** of the third transformer. The secondary winding **71**, of this third transformer is also connected through wire **72**, to the inner plates of capacitor bank **9**. Wires **27** and **28** also connect these inner plates to commutator poles **29** and **30**, which form a change-over switch through commutator contact **31** which is connected to the inner plates of capacitor bank **1** through wire **32**. Switches **33** and **34**, allow the connection or disconnection of the primary winding of transformer **10**, whose secondary winding supplies current to the incandescent lamps **11** and motor **12**. Finally, one of the ends of primary winding **3** is connected through wire **35** to both pole **36** and pole **37** of the commutator **2,** and their corresponding commutator contact **38** is connected to ground by wire **39**.

As the construction and use of all of these components (with the exception of the commutator) is perfectly understood, it would be pointless to describe them. The commutator **2**, is enclosed in a tank **40**, whose end walls **41** and **42** support the ends of the contact-mounting bar **43,** and the bearings of the rotating shaft **44**. The contact bar is made of a non-conducting material to which are immovably attached, the copper contact strip brushes **16**, **17** and **18**, **36**, **37** and **38**, and **29**, **30** and **31**. Secure contact between these brushes and the rotating cylinders mounted on shaft **44** is ensured by the rotating lever arm **47** and its associated weight **48**.

The three rotating cylinders mounted on shaft **44**, are made of a non-conducting material and have a conducting strip around their centre. This strip has two conducting spurs running outwards, one to the right and one to the left, positioned 180 degrees apart around the circumference of the cylinder. When shaft **44** is rotated, this causes the central contact (for example, **18**) to connect first to one of its associated contacts (say, **17**) and then disconnect and connect to the other contact (say, **16**) forming a change-over switching mechanism.

The switching strips on the central cylinder are positioned 90 degrees around the circumference when compared to the position of the switching strips on the two outer cylinders which are aligned with each other.

This can be seen in the diagram, where in the shaft position shown, **38** and **36** are connected and **31** and **30** are connected, while **18** is not connected to either **16** or **17**.

When shaft **44** is rotated through 90 degrees, **18** will be connected to **17**, while **31** and **38** will both be isolated.

When shaft **44** is rotated through an additional 90 degrees, brush **18** will be isolated while brush **38** will be connected to **37** and brush **31** will be connected to **29**.

When shaft **44** is rotated through an additional 90 degrees, brush **18** will be connected to brush **16**, while brushes **31** and **38** will be isolated.

[Note: if the diagram is correctly proportioned, there will be four positions in each rotation where the three central brushes are not connected to any of the outer brushes, producing the switching sequence Make, Break, Make, Break, Make, Break, Make, Break for each revolution. These breaks in the switching sequence have been shown to have a significant effect when batteries are being charged.]

Shaft **44** is elongated and projects through the end wall **42**, so that a belt drive **45**, or other suitable method, may be used to rotate the shaft, driven by motor **12** or possibly by crank handle **46**. The tank **40**, is filled with oil or any other insulating liquid, in order to prevent sparking between the brushes, which would lower the efficiency of the system.

This system is operated as follows:

With the commutator in the position shown in the drawing, that is, with brush **36** connected to **38** and brush **30** connected to **31**, and supposing capacitor bank **1** is connected through wire **13** to a source of electric power (say, a Wimshurst Machine), switch **14** being closed, electric current passes through wire **13** to charge capacitor bank **1**, causing a current to flow through wire **20**, primary winding **3**, wire **35**, brush **36**, brush **38** and wire **39** to earth. This current flow through primary winding **3** induces an inverse current in secondary winding **4**, which flows through wire **21**, charging capacitor bank **5** and then a direct induced current flowing through conductor **22**, charging capacitor bank **5**.

The current flow to both sets of plates in capacitor bank **5**, charges it and creates a spark across the spark gap **6**, causing a very sharp current pulse through primary winding **7**. This in turn, causes a considerable number of high-frequency oscillating current flows in the secondary winding **8** and these pass along wires **25** and **26**, and via diodes **c** and **d**, thus charging capacitor bank **9** and causing a corresponding set of high-frequency pulses to flow to ground through the primary winding **70**. This induces current flow in the secondary winding **71**, which flows through diodes **e** and **f**, and onwards via wire **72**, further boosting the charge on capacitor bank **9**.

Therefore, capacitor bank **1** being charged by an external source, capacitor bank **9** will be charged indirectly and successively re-charged several times, resulting in a quantity of electricity considerably greater than that of capacitor bank **1**. By this means, the system can be self-powered with no need for the power source used to start it, which means that switch **14** can be opened.

When shaft **44** rotates through 90 degrees, brush **17** connects with brush **18**, while brushes **31** and **38** are both disconnected. This causes capacitor bank **1** to be completely discharged through transformer primary **3**, causing the already described process by which capacitor bank **9** receives a substantially larger electrical charge. This results in capacitor bank **1** being totally discharged and capacitor bank **9** being highly charged with a great deal of electricity. Consequently, if we now close switch **33** and rotate shaft **44** through another 90 degrees, the following situation results:

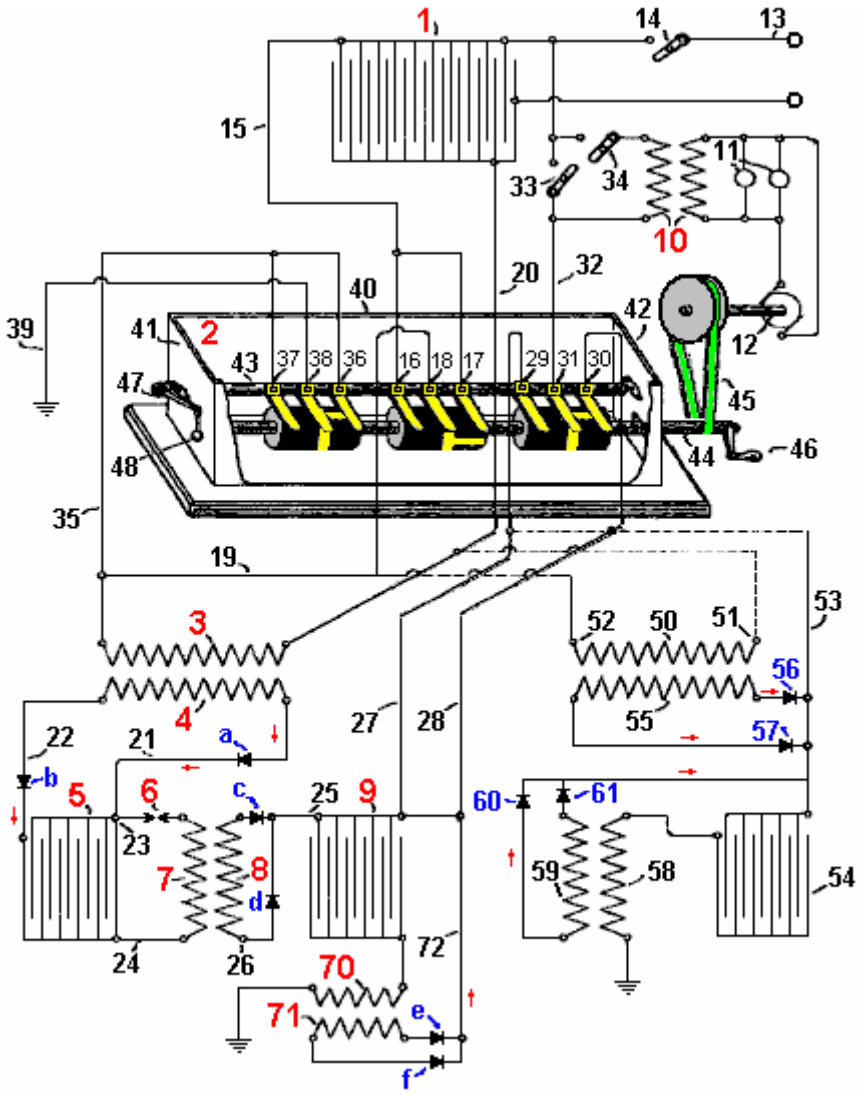
1. Brushes **17** and **18** will be disconnected.
2. Brushes **37** and **38** become connected which in turn connects the outer plates of capacitor bank **1** to ground. Brushes **29** and **31** are connected, which then connects the inner plates of capacitor bank **9** to the inner plates of capacitor bank **1**.

3. Part of the high charge on capacitor bank **9** will flow as an electric current, through wire **32** and into capacitor bank **1**.
4. This current flow from the inner plates of capacitor bank **9** causes an immediate matching negative charge to flow from ground through primary winding **70**, to offset the charge imbalance.
5. This induces a current flow in the secondary winding **71**, which passes additional electrical power to the outer plates of both capacitor bank **9** and capacitor bank **1**, and that further intensifies the current flow through primary winding **70** quite considerably.
6. Further, as capacitor bank **1** has just been newly charged, it drives additional current through primary winding **3**, causing new induced currents which will, as before, produce much increased charge on the inner plates of both capacitor bank **5** and capacitor bank **9**, as described earlier.

If shaft **44** is rotated through a further 90 degrees, then a connection between brushes **16** and **18** will be made and all other circuits will be opened, causing capacitor bank **1** to be discharged again, thus repeating the entire process described above provided that shaft **44** is rotated continuously.

This system produces a constantly increasing supply of electric current flowing through wire **32**, and so, switch **34** can be closed, allowing transformer **10** to provide the electrical power to run motor **12** which maintains shaft **44** in continuous rotation, making the system self-powered with no requirement for any form of outside power supply. Additional transformers inserted in wire **32** can be used to power additional equipment.

Without employing the high-frequency currents described above, similar results may be attained by means of the arrangement shown at the lower right hand side of the following drawing:



Here, primary winding **50** is connected as shown by the dashed lines, with wires **19** and **20** of the former arrangement, and wire **53** connects to both wire **27** and wire **28**. This arrangement has primary winding **50** connected through its end **51** to wire **20** and so is permanently connected to the outer plates of capacitor bank **1**, and its other end **52** being connected to wire **35** will be intermittently connected to ground. Wire **53** being

connected to wires **27** and **28** will intermittently connect together, the inner plates of capacitor bank **54** and the inner plates of capacitor bank **1**.

With this arrangement, both ends of the secondary winding **55** are connected through diodes **56** and **57**, to wire **53**. The outer plates of capacitor bank **54** are permanently connected through primary winding **58** to ground. Both ends of the secondary **59** are connected through diodes **60** and **61**, back to wire **53**. As a result, if capacitor bank **1** is charged, it drives a current through wire **20** and so, through primary winding **50**, and on to ground through wires **35** and **39**. This induces current in secondary winding **55** which gets stored in the inner plates of both capacitor bank **54** and capacitor bank **1**, as in this moment, the circuit is closed between brushes **29** and **31**, and so, wire **53** is connected to wire **32**. On receiving these new charges, both capacitor bank **1** and capacitor bank **54** will create new induced electric currents flowing through primary windings **50** and **58**. These multiple charging pulses will decrease with time until they are insignificant, at which time, due to the rotation of shaft **44**, the connection between brushes **29** and **31** and between brushes **36** and **38** will no longer be maintained, and instead, brushes **18** and **17** will become connected, discharging capacitor bank **1** through primary coil end **50** which is a very strong discharge, charging capacitor bank **54** as wire **53** is now disconnected from wire **32**. This, in turn, causes powerful current flow through primary winding **58**, further charging capacitor bank **54** which then feeds capacitor bank **1** when shaft **44** rotates further, making the system both self-powered and capable of supplying useful electrical power to other equipment.

It should be clearly understood that the use of mercury-vapour converters or cathodic valve diodes as described, are not in any manner indispensable in the system as those devices can be replaced by a suitable arrangement of capacitors which would receive separately, the direct and inverse currents of the secondaries.

When a bell is struck just once, it vibrates many times, passing those vibrations to the air and so, making the sound which we hear. The bigger the bell, the slower the vibrations and the lower the pitch of the note which we hear. The same thing happens when a sharp voltage pulse is applied to a coil of wire as just one pulse causes many vibrations in the coil. Like the bell, the frequency of the vibrations depends on the structure of the coil and not on how it is pulsed, although, like a bell, a sharp pulse for a coil or a sharp blow for a bell, produces a greater effect.

You will notice here that Carlos uses the resonant ‘ringing’ of an air-core coil to get an energy gain which is then used as positive feedback to further charge a capacitor bank. A single sharp pulse generated by one spark, causes a large number of coil oscillations, each of which contributes output power, producing an energy gain. The ringing frequency is liable to be around 3 MHz. It is also worth noting that with this design, electricity generation can be achieved without any battery and just the manual turning of a Wimshurst electrostatic generator and the initial operation of the commutator shaft **44**.

Carlos also produced another design, this time working with batteries (although he tended to think in terms of 60-volt battery banks rather than 12-volt batteries) and his patent includes what we tend to call the Tesla Switch nowadays. Perhaps we should call it the Benitez Switch. However, instead of switching it rapidly, Carlos uses a switching time interval of one hour. The lower voltage overcomes the need for the switching contacts to be submerged in oil.

Carlos Benitez Patent GB 14,311 17th August 1916

System for the Generation of Electric Currents

I, Carlos F. Benitez, Civil Engineer, 141 Ocampo Street, Guadalajara, Mexico, do hereby declare the nature of this invention:

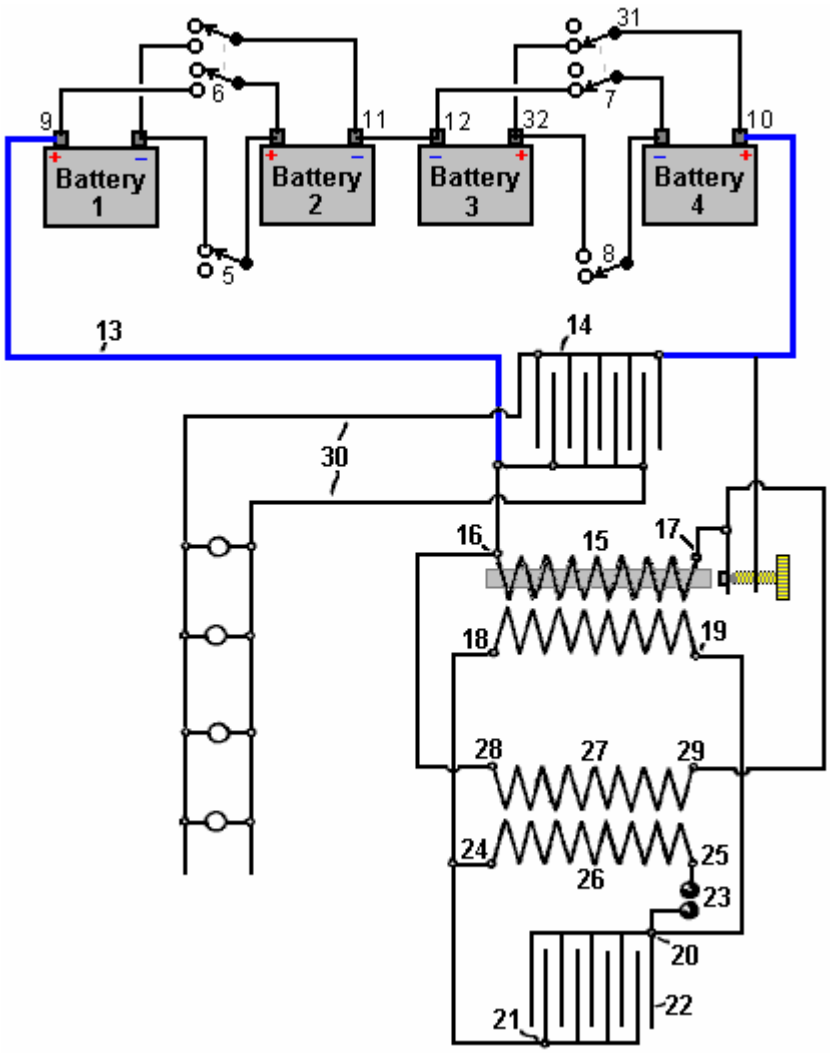
The invention which forms the object of this Patent of Addition, relates to new improvements in the system for the generation of electric currents, described in the main patent No. 17,811 and in the patent of addition No. 5591, filed 14th April 1915.

The system may be still further simplified and improved by the addition of batteries, which, suitably adjusted in conjunction with system previously described, can be charged and discharged alternately, producing an excess of electrical energy which can be used in any desired manner.

In other words, in this arrangement, I use in combination: two batteries connected in series and two batteries connected in parallel, these pairs being used so that the discharge of one pair is used to charge the other pair, and *vice versa*.

Another object of this new arrangement is to allow the use of low voltages, small capacity capacitors and additional facilities for starting the system.

The advantages of such an improvement will be better understood by considering the following drawing which illustrates one method of carrying out the invention:



In the diagram, **1, 2, 3** and **4** are batteries which, when charged from some external source, will maintain their charges indefinitely, in the following manner:

With the connections established as shown in the diagram, that is, with batteries **1 & 2** connected in series through switch **5** (switch **6** being open), batteries **3 & 4** are connected in parallel through switch **7** (switch **8** being open). Under these conditions, assuming that the four batteries are similar, having similar voltages, batteries **1** and **2** being in series will have a combined voltage greater than batteries **3** and **4** which are connected in parallel, and so, a load connected between them will have a current flowing from batteries **1 & 2** and into batteries **3 & 4**.

In other words, if wire **13** is connected to the positive terminal of the battery **1 & 2** combination, and to the positive poles **10 & 32** of the battery **3 & 4** combination, then an electric current will be established between the two sets of batteries, until their voltages match. Of course, the current provided by batteries **1 & 2** would produce a smaller charge in the batteries **3 & 4**, but that current can be increased by any of the methods described in my earlier patents (No. 17,811/14), and by these means it is always possible to alternately charge and discharge the battery pairs from each other, keeping a constant, pre-determined charge, and furthermore, producing an excess of electrical energy which can be used for any chosen purpose.

With these objectives in view, and using as an illustration, the arrangement shown in Figure 1 of the Patent of Addition No. 5591/15, wire **13** connects to capacitor **14**. The primary winding **15** of an ordinary induction coil provided with an interrupter, is connected by its ends **16** and **17**, to wire **13**. The ends **18** and **19** of the secondary winding of that induction coil **15** are connected to connections **20** and **21** of capacitor (or bank of capacitors) **22**. Connections **20** and **21** are also connected through spark gap **23**, to ends **24** and **25** of the primary winding **26** of a high-frequency transformer. The ends **28** and **29** of the secondary winding **27** of that transformer are connected

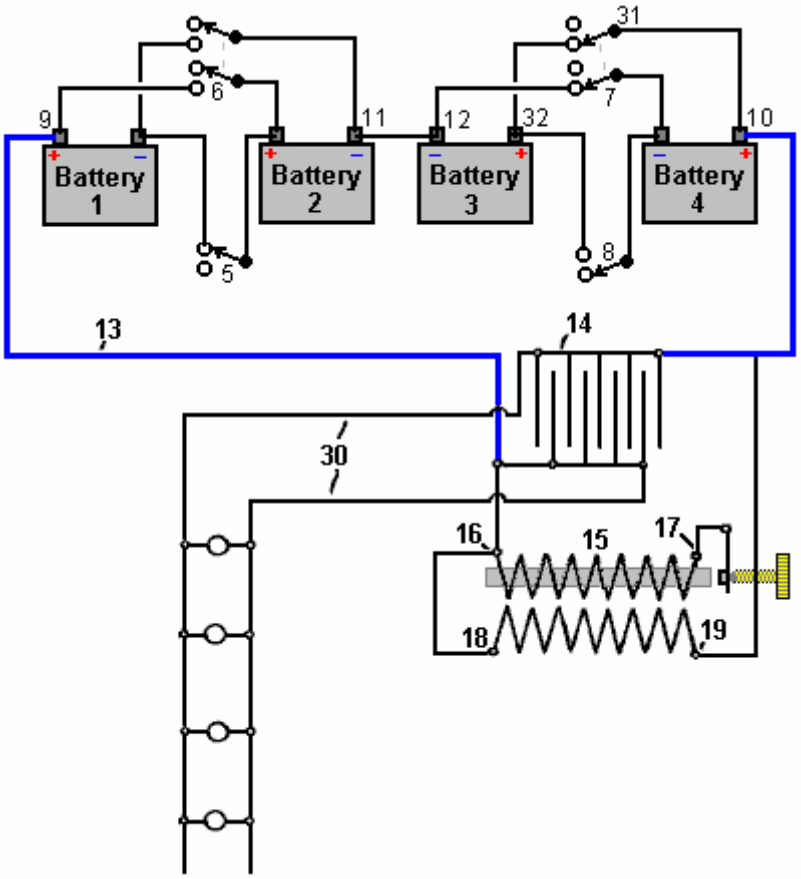
to ends **16** and **17** of the induction coil **15**. Finally, wires **30** are connected across capacitor **14** and they are used to power external loads such as the incandescent lamps shown in the diagram.

This arrangement being made, the electrical energy stored by batteries **1 & 2**, passing through terminal **9**, wire **13**, primary winding **15**, terminal **10** of battery **4**, pole **31** of switch **7** and terminal **32** of battery **3**, will go back through terminal **12** of battery **3** to the battery **1** and **2** combination.

As a consequence of the current passing through primary winding **15**, high-voltage current is produced in its secondary winding and collected in capacitor **22**, discharging through spark gap **23**, generates high-frequency currents in the high-frequency transformer primary and secondary coils **26** and **27**. As coil ends **28** and **29** are connected to coil ends **16** and **17**, this greatly augments the current flow provided by batteries **1 & 2** and so the storage batteries **3** and **4** now receive adequate charging current to keep them fully charged as well as driving additional loads via wires **30**.

Under these conditions, as the voltage in one of the batteries is decreasing while the other is increasing, after some hours, both voltages match and it is then impossible to produce any current flow unless switches **5, 6, 7** and **8** are operated, reversing the functions of the batteries and allowing the process to continue entirely as before with batteries **1 & 2** being connected in parallel and batteries **3 & 4** being connected in series.

When the resistance of the primary winding of transformer **15** is not high, it is possible to simplify the above circuit by obtaining the high-frequency currents directly from induction coil **15**, in which case, ends **18** and **19** of the secondary winding are connected directly to ends **16** and **17** of the same coil and the second bank of capacitors **22** and the high-frequency transformer **26/27** can be omitted. Under these conditions, the breaker or interrupter which is part of the induction coil construction, acts as a spark gap, and capacitor **14** discharges in the form of oscillations through the primary winding **15** of the same coil, thus directly increasing the amount of electrical energy furnished by the accumulators.



An essential part of this design which is not indicated clearly in the patent, is that what was common practice a hundred years ago, namely, that the input power connection to point **17** of the primary of the (step-up) transformer **15**, is fed through an “interrupter” contact. This connection opens when the coil **16-to-17** is energised, causing the insulated iron wire bundle core of the coil to become magnetised, and attract the pivoted arm of the interrupter, which breaks the current to the coil very sharply, causing high frequency resonant oscillations in both windings of the transformer **15**, which generates the excess power which runs the system and its additional loads. In Benitez’s day, door bells used this style of interrupter to produce a hammering action on a metal bell. These were very cheap, very simple and very reliable.

As I understand it then, the voltage difference between the two pairs of batteries, charge capacitor **14** and apply power to the primary winding **16-17** of step-up transformer **15**. This causes a current to flow in this winding, making the core attract the pivot arm of the interrupter in the same way that a relay operates. This breaks the current flow very sharply, causing a powerful back-EMF pulse in the primary winding. The primary winding has a resonant frequency, very considerably lowered by the presence of the iron core which itself is a low-frequency material, and the coil oscillates at its resonant frequency, not for just one cycle but for many cycles. Each of those cycles generates a high voltage in the secondary winding **18-19** and every one of those cycles contributes high-voltage power to the system. That power is fed to three outlets. Firstly, it flows back to provide charging power to one of the battery pairs. Secondly, it adds additional power to the capacitor driving its own primary winding. Thirdly, it provides power for the load which is shown as a series of lamps connected in parallel.

That is just for the first interrupter pulse. The broken current through the primary winding **16-17** causes its core to cease to be an electromagnet and so it ceases to attract the pivot arm of the interrupter and while that seems very quick in human terms, it is very slow compared to the multiple oscillations ringing in the winding. When the pivot arm returns to its starting position, it establishes the current flow through the primary winding once again. However, the charge on the capacitor powering the primary winding has been boosted by those resonant oscillations in the secondary winding and so is more highly charged than when the interrupter contact opened previously. This process continues repeatedly, providing battery charge and power to the load.

According to Carlos, there is a slight overall drain on the battery system and so, after about an hour, the switches are operated, changing the series-connected batteries to become parallel-connected and the parallel-connected batteries to become series-connected. This timing seems odd as switching the batteries over much more frequently only requires batteries with a much lower capacity.

As we are not familiar with induction coils and interrupters now that solid-state electronics is available, we can get constructional and operational information from that period in the book “Wireless Telegraph Construction for Amateurs” by Alfred Powell Morgan, published in 1913, which is available as a free download from here: http://www.free-energy-info.tuks.nl/Morgan.pdf.

For example, the interrupter details include:
“Some means of charging the capacitor is necessary. An induction coil is the most practical for the amateur. The induction coil consists of a primary coil of wire wound around a central iron core and surrounded by a secondary coil consisting of many thousands of turns of carefully insulated wire. The primary coil is connected to a source of direct current which also includes an interrupter to “make” and “break” the current in rapid succession. Every “make” of the circuit and consequent magnetisation of the core, induces a momentary inverse current in the secondary winding, and every “break” and corresponding demagnetisation induces a momentary direct current. Normally, the induced currents would be equal, but by means of a capacitor connected across the interrupter, the circuit when “made” requires considerable time for the current and magnetisation of the core to reach a maximum value, while when broken, the demagnetisation and current drop are nearly instantaneous. The value of the induced electromotive force in a circuit, varies with the speed at which the magnetic lines of force cut the circuit, and so, the induced e.m.f. at “break” becomes high enough to leap across a spark gap.

The formulas connected with induction coils depend on conditions which are never met in actual practice and cannot be relied on. To construct a coil of a given size, it is necessary to use dimensions obtained empirically. Therefore, the amateur should stick closely to the lines or hints given here, or which appear in some *up to date* book on induction coil building.

For a long time, the induction coil was an expensive, inefficient instrument, until wireless telegraphy demanded of it more rigid and efficient design and construction. It was the aim of manufacturers to produce the longest possible spark length with a minimum amount of secondary wire. As a result of this demand, wireless coils are now made with a core of larger diameter and give heavier and thicker sparks. The secondary in this case, is short and uses wire of large cross-sectional area in order to reduce the resistance and minimise the heating.

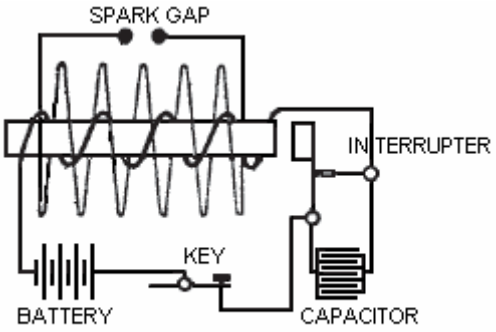


Diagram of an Induction Coil.

No one part of an induction coil may be developed to its maximum efficiency without seriously influencing and lowering the efficiency of the other parts. The following suggestions regarding the construction are given so that they may prove to be a useful guide for the amateur coil builder. The parts will be considered in their natural order of construction.

Core: Some experimenters who are not quite familiar with the principles of magnetism, think that if an induction coil were provided with a closed core like that of a transformer, then the efficiency of the coil would be materially increased. But that is not the case because then the magnetisation and demagnetisation of the iron cannot take place rapidly enough in a closed core when an interrupted direct current is used instead of an alternating current.

The core of an induction coil is therefore always straight. For the same reason, it is never solid but instead is always made up of a bundle of soft iron wires in order that rapid changes in magnetism may take place. The wires are always of as high a permeability (magnetic conductance) as possible so as to create a strong magnetic field. Swedish or Russian iron of good quality is the best as its hysteresis losses are small. The smaller the diameter of the wire, the less will be the eddy current losses and consequent heating, but the permeability is also reduced and the core will not be so effective, as the amount of iron is thereby decreased and the oxidised surface increased. No. 22 gauge wire is the best size for the average core.

Wires of a good quality may be purchased already cut to various lengths. To buy them in this form will save a great deal of the labour required when building a core. If the wires are not quite straight, they may be straightened by rolling them, one at a time, between two boards. It is best to re-anneal the wires. To do this, place the wires in an iron pipe and plug the ends of the pipe with clay. Then lay it in a coal fire until the whole mass reaches a red heat. The fire is then allowed to die out gradually, with the pipe and wires remaining in the ashes until cool. When cool, remove them from the pipe and rub each one with emery paper until bright. After this cleaning, the wires are dipped in hot water and then dried. They are then dipped in a good quality varnish and allowed to dry again.

The varnish provides a resistance to the flow of eddy currents in the core and reduces those losses very considerably. A strong paper tube having an internal diameter equal to the diameter of the finished core is made by rolling the paper on a form and cementing it with shellac. When it is perfectly dry, the tube is removed and the wires packed tightly inside it. The following table gives the core dimensions for practical coils of different sizes:

CORE DIMENSIONS

Spark Length	Core Diameter	Core Length
0.5 inch (12 mm)	0.5 inch (12 mm)	6 inches (150 mm)
1 inch (25 mm)	1 inch (25 mm)	8 inches (200 mm)
2 inches (50 mm)	1.25 inches (31 mm)	10.5 inches (262 mm)
4 inches (100 mm)	1.375 inches (34 mm)	12 inches (300 mm)
6 inches (150 mm)	1.5 inches (38 mm)	14 inches (350 mm)

Primary Winding: The ratio of the number of primary turns of an induction coil to the number of secondary turns, bears no relation to the ratio of the primary and secondary currents. It has been found in practice, that two layers of wire wound tightly on the core, form the best primary. The primary should always be thoroughly covered with shellac or other insulating varnish. Since there is almost no ventilation in the primary, the wire must be large enough to avoid all heating. A table containing the various sizes of primary wires is given here:

PRIMARY WINDINGS

Spark Length	Wire Gauge	Number of Layers
0.5 inch (12 mm)	18 (1.219 mm)	2
1 inch (25 mm)	16 (1.63 mm)	2
2 inches (50 mm)	14 (2.03 mm)	2
4 inches (100 mm)	12 (2.64 mm)	2
6 inches (150 mm)	12 (2.64 mm)	2

In large coils, the inductance of the primary causes a “kick-back” and sparks are liable to pass between the adjacent turns. For this reason, it is always a good idea to use double cotton covered wire and to further thoroughly insulate it by soaking the primary and core in a pan of melted paraffin wax and allowing the wax to harden with them inside it. Afterwards, the pan is slightly warmed to loosen the cake of paraffin and the excess wax removed by scraping with a blunt instrument so as to avoid damaging the wires. Paraffin wax contracts when it hardens and the proper method for impregnating a porous substance is to allow it to soak and become set in it upon cooling.

A good method of reducing the “kick back” as well as the size of the capacitor connected across the interrupter, is to make the primary with a number of turns of smaller diameter wire wound in parallel, the effect being to produce a conductivity equal to a wire of large diameter and at the same time, make a more compact winding of the primary on the core. This method of winding is very desirable in large coils, as it reduces the cross-section of the primary and allows the secondary to be placed nearer to the core where the magnetic field is the strongest.

The primary winding should cover nearly the whole length of the core, since there is no advantage in carrying the core very far beyond the end of the primary as most of the magnetic lines of force bend at the end of the primary winding and return without passing through the extreme ends of the core.

Insulating Tube: The successful operation of an induction coil without breaking down when heavily stressed, depends largely on the insulating tube which separates the primary and the secondary windings. Hard rubber tubes are perhaps the best. A tube may easily be built up of several layers of half inch sheet hard rubber by steaming it so as to soften it, and then wrapping it around a former. The tube should fit tightly on the primary and be about one inch (25 mm) shorter than the core. After the tube is in place, it is poured full of beeswax and rosin in order to fill all interstices and prevent sparks due to the capacitor effect of the windings from jumping from the inside of the tube to the primary.

Secondary: A coil used as a radio telegraph transmitter must have wire of a large cross-sectional area in its secondary, so as to produce a heavy disruptive discharge. Number 34 and number 32 gauges are generally used for small coils and number 30 and number 28 gauge for large coils. Silk covered wire is the usual practice, but enamelled wire is coming into use. Cotton covered wire takes up too much space and has poor insulating qualities.

Enamelled wire is insulated with a coating of cellulose acetate, which has a dielectric strength of about twice that of cotton and it takes up much less room than silk covered wire, giving a great saving in space and a greater number of turns may be wound in the secondary without increasing its mean distance from the core.

When winding enamelled wire, it must be taken into consideration that the insulation of enamelled wire is rigid and has no give. Consequently, to allow for expansion, enamelled wire must be wound more loosely than fibre or silk covered wire. The occasional insertion of a layer of paper in the winding gives room for expansion and does not add greatly to the diameter. The length of the secondary is generally not more than half the length of the core.

Coils producing sparks up to 2 inches (50 mm) in length, may be wound in two sections or in layer windings, but the layer winding is not recommended for coils giving sparks over one inch in length. It is best in a coil of this type, to insert an occasional layer of paper. The paper should be well shellacked or paraffined and be a good grade of linen. It should project about one quarter of an inch (6 mm) from the ends of the secondary as shown in this cross-section:

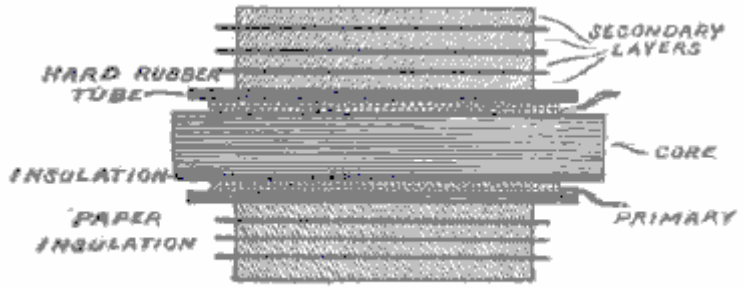
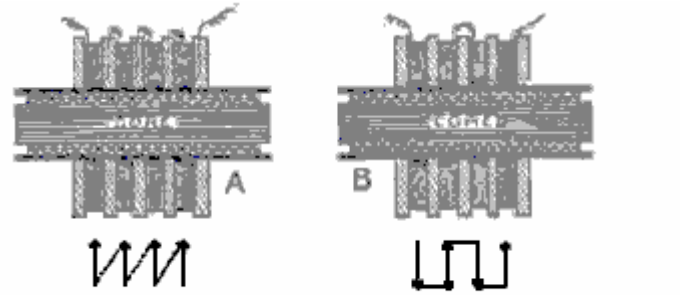


Fig. 28. Layer Winding for Small Coils.

This insertion of paper increases the insulation and reduces the possibility of sparks jumping from layer to layer when the layers are very long. The secondaries of large coils are made up of "pies" or "pancakes" from one eighth (3 mm) to three eighths of an inch (9 mm) in thickness. The "pies" are separated from each other by a triple thickness of blotting paper which has been thoroughly dried and then soaked in melted paraffin. When each "pie" is completed, it is tested for continuity and rejected if not perfect. They are connected in series:



Methods of Connecting the Secondary Sections

If connected as shown in example "A", where the inside of one section is connected to the outside of the next section, the maximum voltage which can exist between the adjacent sections in this case is equal to the e.m.f. generated by one "pie" and is equal throughout. Connecting as shown in case "B" where the outside of one coil is connected to the inside of the next one, the voltage ranges from zero at the points where they are connected, to twice the e.m.f. developed by any one section. This is the better method and every second coil is turned around horizontally to allow for the reversed direction of current flow.

After the secondary is assembled, the coil should be submerged in an airtight tank containing melted paraffin wax. The tank is then connected to a vacuum pump and the air pumped out. This causes any air bubbles in the windings to be pumped out. After standing for a while, the vacuum is released and the air pressure then causes the bubble gaps to be filled with paraffin wax.

SECONDARY DIMENSIONS

Spark Length	Wire Gauge	Amount required	No. of sections	Length (1 inch = 25.4 mm)
0.5 inch (12 mm)	36 (0.193 mm)	0.5 lb (0.227 Kg)	1	3 inches
1 inch (25 mm)	34 (0.234 mm)	1 lb (0.454 Kg)	2	5 inches
2 inches (50 mm)	34 (0.234 mm)	2.5 lb (1.134 Kg)	4	7 inches
4 inches (100 mm)	32 (0.274 mm)	8 lb (3.629 Kg)	12	8.5 inches
6 inches (150 mm)	32 (0.274 mm)	12 lb (5.443 Kg)	20	10 inches

It should be noted that the very experienced Alfred Morgan flatly contradicts the standard theory of symmetrical transformer operation when he states that "*the ratio of the number of primary turns of an induction coil to the number of secondary turns, bears no relation to the ratio of the primary and secondary currents*". That is a highly significant statement.

Carlos Benitez also produced another very clever design, still using four batteries and a very slow switching rate, although no longer using the series and parallel switching which we think of as the Tesla Switch. In this design, he shows a remarkable high-frequency power-gain system where 400 watts of input power produces 2400 watts of output power (COP=6):

Carlos Benitez Patent GB 121,561 24th December 1918

New Process for the Generation of Electrical Energy

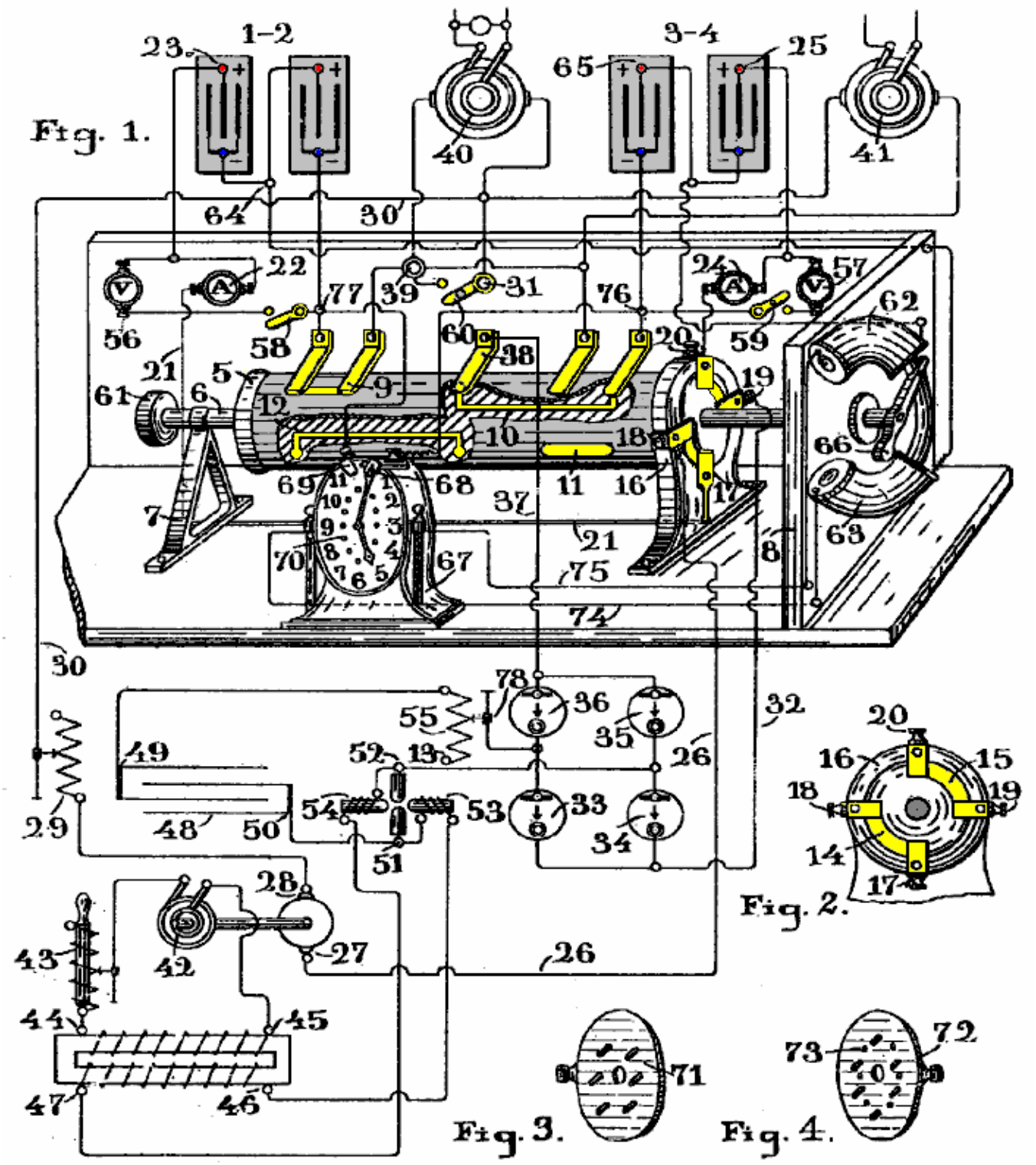
I, Carlos Benitez, Civil Engineer, of 141, Ocampo Street, in Guadalajara, Mexico, do hereby declare the nature of this invention and in what manner the same is to be performed, to be particularly described and ascertained in, and by, the following statement:

This invention relates to a new utilisation of the high frequency currents or electrical oscillations, by means of which, a constant production of electrical energy can be secured, under unusually simple, economic and practical conditions.

Such results are attained by means of the procedure disclosed in present applicant's English Patent specification No. 14,311, filed on October 9th, 1915, but in order to obtain a better utilisation the high-frequency currents and the automatic operation of the mechanisms employed in such a procedure, I have invented a novel arrangement of parts from which is derived several other advantages that will be here described and set forth.

Two batteries are used, one of which has already been charged. This charged battery discharges through a circuit which utilises the greater part of this power and the remainder drives an oscillating circuit connected to the second battery. This oscillating circuit contains a rectifier which directs the electrical oscillations in such a manner that they are forced to pass constantly through the second battery from the positive to the negative terminal. As the number of oscillations per second in that circuit can be varied at will, the current intensity developed by those oscillations can be regulated so that the second battery is fully charged in the same period of time during which the first battery is discharged. Therefore, it is only a matter of reversing the connections to the two batteries in order to obtain the continuous production of electrical energy.

The invention is illustrated here:



This circuit shows the circuit connections of the arrangement and a perspective view of a commutator. Figure 2 is a lateral view of a part of the same apparatus, and figures 3 and 4 are parts of the same machine, whose location and use will be explained later.

The figure shows two battery banks **1 & 2** and **3 & 4,** both of which are connected in series. Their positive poles **23** and **25** are connected to terminals **17** and **20** of the commutator, through the ammeters **22** and **24**. The commutator is a cylinder **5**, mounted on a rotating shaft **6**, which passes through two supports **7** and **8**. The cylinder is provided with conductive paths **9, 10, 11** and **12** insulated from the cylinder **5**. It also has conductive paths **14** and **15** which can be seen better in **Figure 2** and which make contact with brushes **17** & **18** and **19** & **20**, properly insulated and fixed to ring **16** which encircles cylinder **5** and is fastened to the base of the apparatus. These brushes connect to the batteries and terminal **17** is connected to the positive pole **23** of battery **1 & 2**. Terminal **20** is connected to the positive pole **25** of battery **3 & 4** through ammeter **24**. Terminal **18** connects through wire **26** to terminal **27** of a DC motor whose terminal **28** connects through the variable resistor **29** and wire **30** with terminal **31** of the commutator. Finally, terminal **19** is connected through wire **32** to terminals **33** and

34 of an electrolytic rectifier whose other two rectifiers **35, 36** connect via wire **37**, to brush **38** of the commutator.

This apparatus is also provided with terminal **39** which is alternately connected to the negative poles of the batteries through the brushes which can be seen in the figure, in contact with cylinder **5**. Under these conditions, the apparatus **40** and **41**, which are employed to provide a practical output independent of the running of the "plant", and which are connected to both terminals **31** and **39**, will be alternatively in circuit with the discharging battery, and thus, part of the power produced by the battery discharging, is used by this apparatus without impairing the normal running of the whole mechanism, as explained below.

On the other hand, an alternator **42**, is coupled to the motor **27-28**, which in this manner can be rotated at the speed needed to attain the desired frequency. Both terminals of the alternator are connected through the inductive resistance **43**, with the poles **44, 45** of the primary winding of a step-up transformer, whose secondary winding **46, 47**, connected to capacitor **48**, will finally produce the high-voltage alternating currents needed to charge the capacitor. However, the use of this motor and alternator is not essential in this process as the same results could be attained if the primary **44, 45** of the transformer, were connected through an ordinary interrupter to the discharging battery.

The terminals **49, 50** of capacitor **48** are connected through an arc lamp of the Poulsen type **51, 52**, or through an adequate spark-gap with the rectifiers **33, 34, 35** and **36**, and with the battery under charge (**3 & 4** in this case). Extending into the arc chamber are the poles of a strong electromagnet, the coils of which are in series with the arc, so that their excitation current is the arc current.

Connected in this manner, the strong magnetic field which exists between the poles of the magnets, acts on the arc, and owing to this action, combined with the influence of the voltage developed by the secondary of the transformer **46, 47**, it happens that this action and influence being about equal in value, are each of them alternately greater than the other, so that the voltage across the arc automatically rises and falls. Therefore, when the action of the magnetic field is greater than the voltage produced by the secondary winding of the transformer has not enough power to pass across the arc and so the capacitor **48** is charged to a higher voltage. But, an instant later the action decreases in value and the capacitor discharges again across the arc.

On the other hand, owing to the oscillatory nature of the circuit (which is arranged with proper values of capacity, inductance and resistance), the charging and discharging of the capacitor can occur several million times per second if desired, and in this manner, a current of a great number of amperes can be obtained through the oscillating circuit, even with a small amount of electricity stored in the capacitor.

In the same manner, one gallon of water could produce a flow of one thousand gallons per second through a pipe if such a pipe were connected with two different vessels and the gallon of water could be forced through the pipe by a piston which could transfer that gallon of water from one vessel to the other one thousand times per second. Obviously, what can be easily performed with electricity is not so feasible with water.

In other words, the small amount of electrical power taken by the DC motor **27, 28**, appears (with a small loss) at the alternator **42**, and that power is delivered to the transformer primary **44, 45**. Again, the inductive action of this transformer produces at the secondary **46, 47**, a similar amount of power (diminished slightly due to the efficiency of the transformer), and finally, the capacitor is charged with a small amount of electricity which is then converted into oscillitatory energy. Obviously, if such electric power, instead of being stored by the capacitor, had simply been rectified and used to charge one of the batteries, such power would produce only a very small effect on the battery and the entire discharge of one battery would never cause the complete charge of the second battery.

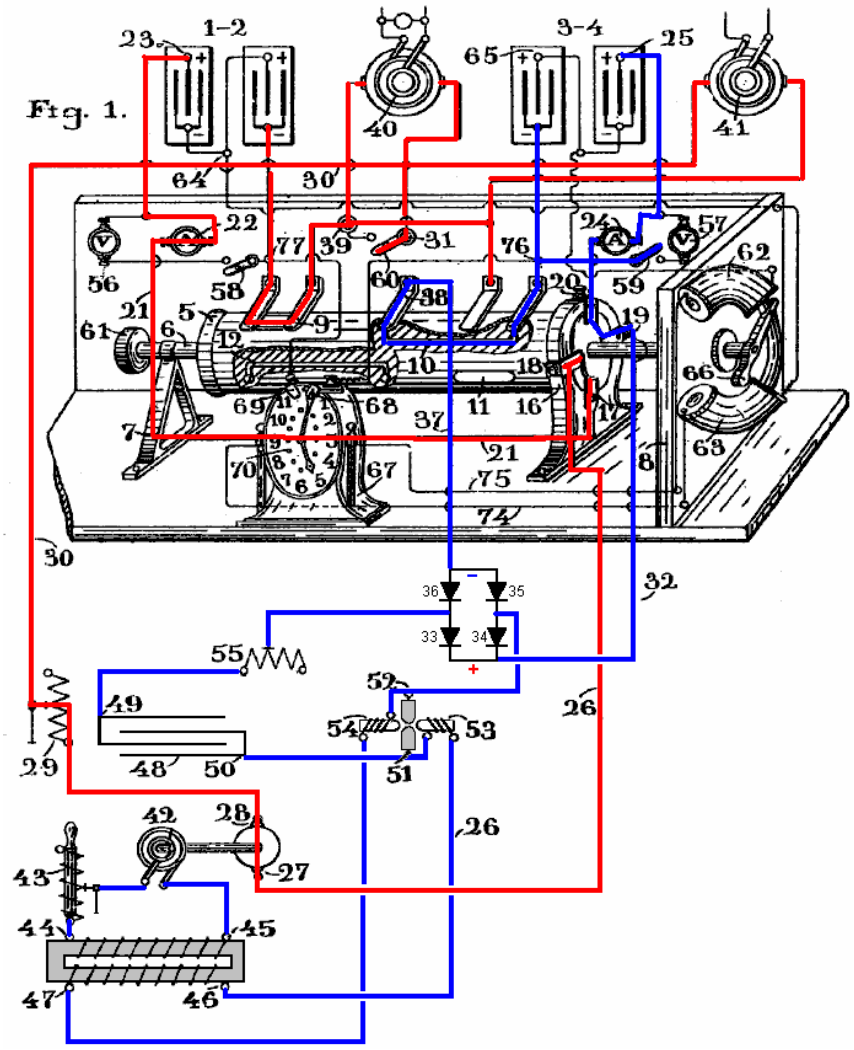
Contrary to that, if that same power is stored in capacitor **48**, and that capacitor is properly connected to an oscillatory circuit in which one of the batteries can be joined, and furthermore, if by means of a rectifier, the high-frequency currents produced in such an oscillatory circuit are forced to pass from the positive to the negative pole through the battery, it is obvious to state that it is always possible to secure by these means, the number of amperes required to charge the battery in the available time. That is to say, with a small number of coulombs stored in capacitor **48**, it is possible to produce in the oscillatory circuit, a great number of amperes, if that same small number of coulombs are forced to pass and repass through the circuit, thousands or millions of times per second, just as was explained in the water analogy.

On the other hand, the values of the above named quantities: capacity, resistance, inductance and voltage can be varied within very wide limits, and therefore it is always possible to attain the required conditions in each case, in order to produce a given number of oscillations per second: The capacity of the capacitor can be adjusted to a certain value by increasing or reducing the inter-meshed surface area of its plates. The resistance of the circuit can be adjusted to the required value by varying the length of the arc in the Poulsen lamp, or varying the number of lamps connected in series or parallel in the circuit. The inductance of the circuit can be varied by winding part of the circuit on an insulating frame, in such a manner as to obtain the number of turns required to produce the

desired inductance, and finally, the charging voltage can be regulated by increasing or decreasing the number of turns in the secondary winding of the transformer or by varying the diameter of the wire used in the coil. In order to obtain a better efficiency from alternator **42**, it is convenient to use the resonance coil or adjustable inductive resistor **43**. By so doing, it is possible to adjust the resistance in order to obtain a state of resonance in the circuit, and in that state, the current produced by the alternator will be in phase with the impressed E.M.F., therefore the effective watts will be a maximum in the primary and secondary circuits.

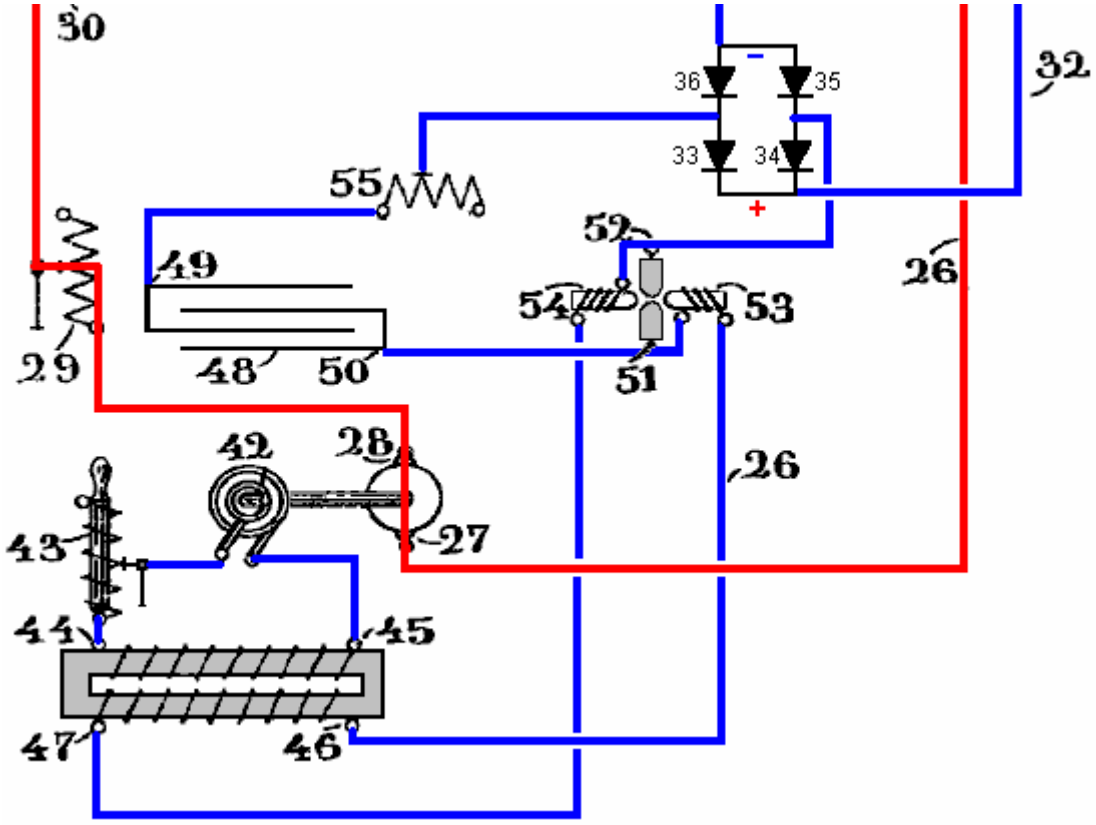
The variable resistor **29**, which is placed in the circuit which connects each battery with motor **27-28**, is used to adjust the current of the discharging battery to a fixed value, since it is important to obtain a constant number of revolutions per second of the alternator **42**.

The commutator is also provided with voltmeters **56** and **57**, and by means of the switches **58** and **59**, the circuits connecting both poles of each battery, can be closed and the voltage of the discharging current may be determined when desired. Finally, by means of switch **60** connected across the terminals of the apparatus **40** and **41**, the apparatus may be switched off when not required.



It is easy to understand the whole operation of this mechanism. Suppose initially that battery **1-2** has been charged and that cylinder **5** has been rotated to the position shown in the drawing. Battery **1-2** will immediately discharge via battery terminal **23**, ammeter **22**, wire **21**, contacts **17-18**, wire **26**, motor **27-28**, variable resistor **29**, wire **30**, rotary transformers **40** and **41**, (or through any other apparatus which may be used instead of those transformers in order to use part of the current flowing from the battery), pole **39**, and commutator path **9-12**, through which the whole discharging circuit is closed.

As a result of this discharge from the battery, the electrical power produced will be dissipated in three ways: A first part is wasted in overcoming the internal resistance of the various components in the circuit. A second part is used to power the DC motor **27-28**, and the third part powers the apparatus connected to poles **31-39** of the commutator, that is, the useful equipment powered in addition to the running of the system.



It is well known how the mechanical power developed by the rotary motion of motor **27-28** can be converted to electrical energy by alternator **42**, and in the same manner, it is well known how such electrical energy of low voltage can be transformed into one of high voltage using a transformer, and it is well known too, how such high voltage electrical energy can be converted into oscillatory energy, and how the alternating currents produced by such oscillations can be rectified in order to produce a direct current. However, all such combinations of apparatus arranged to charge a battery while another similar battery is discharging, was perfectly unknown before this invention, and a short consideration of the calculations involved in determining the different values of capacitance, resistance, inductance and voltage required by the various combined circuits of a small working plant of this kind, may be useful in giving an idea of the procedure which can be followed in general practice.

Suppose that both of the batteries are accumulators with a voltage of 60 volts each and a capacity of 40 AHr. Under such conditions, if battery 1-2 discharges at the rate of 40 amperes, it will be fully discharged in one hour. [Please note that this is not so and is only being stated here for discussion purposes. A lead-acid battery will be damaged if discharged at a rate greater than the "C20" rate which is the rated AHr value over a period of 20 hours, and so, a lead-acid battery of 40 AHr should not be discharged at more than 40 / 20 = 2 amps. Also, batteries are highly non-linear and discharging a 40 AHr battery at 40 amps will result in a fully discharged battery in a lot less than one hour.] The power developed over that time period will be 60 volts x 40 amps = 2,400 watts.

On the other hand, in order to charge battery **3-4** fully in one hour it is necessary to provide a current of at least 40 amps. Suppose that in order to develop such a current, it is desired to take from the power produced by the discharging battery only a portion, say 40 amps x 10 volts = 400 watts. With this objective in view, the DC motor should be arranged so as to create a current of 40 amps causing a drop of 10 volts on the line.

Suppose that the electrical efficiency of the DC motor is 95%, then the desired 400 watt output will not be attained but instead will be reduced to 400 x 0.95 = 380 watts.

Then, suppose that the electrical efficiency of alternator **42** is 95% then that will reduce the output to just 361 watts. Then, if this 361 watts is passed to the transformer and that transformer has an efficiency of say, 89%, then the resulting energy output will be further reduced to just 321 watts, and that is the amount of power passed to the oscillating circuit in order to obtain the required 40 amps of current.

Now, suppose that the frequency of alternator **42** is 500 Hz. As is well known, with an alternating generator, a high voltage is obtainable twice in each cycle and so there will be 1000 voltage peaks per second. As capacitor **48** discharges at the instant of maximum voltage, it will discharge 1000 times per second. Therefore, the quantity of electrical energy which must be stored in it can be calculated as follows: Suppose that the oscillating circuit has an impedance of 15 ohms. As the required current is 40 amps, the required voltage to develop such current will be 40 amps x 15 ohms = 600 volts. But in order to produce 600 volts from the discharge of the capacitor, it is

necessary to get it charged up to 1200 because the average voltage during discharge is V / 2.

If the energy produced by the transformer secondary is assumed equal to 321 watts and the voltage required for the charge of the capacitor is 1200 volts, then the current delivered by the secondary will be 321 watts / 1200 volts which is 0.267 amps.

On the other hand, as the frequency of the alternator is 1000 peaks per second, then in one thousandth of a second the secondary has to deliver to the capacitor 0.000267 coulombs which is 267 microcoulombs.

Hence, the capacity of capacitor **48** must be adjusted to store that amount of electricity, and its value can be determined by **Q = K x V** where K is the capacity of the capacitor in microfarads when Q is given in microcoulombs and V is the maximum voltage, and so, K = 267 / 1200 which is **0.222 microfarads**.

It is also well known that if you want to produce an oscillatory discharge through a circuit, the capacitance, resistance and inductance of the circuit must be arranged so that the square root of 1000 x L millihenrys / K microfarads is greater than the resistance of the circuit in ohms (K being the capacitance of the capacitor).

It has been supposed that the resistance of the oscillatory circuit is 15 ohms. Using 20 ohms in the above equation will allow a suitable value of inductance to be calculated, one which will satisfy the required conditions for the production of the oscillatory discharge in the circuit, and so the inductance in millihenrys is 400 x K / 4000 or 400 x 0.222 / 4000 which is **0.0222 millihenries**, or equal to 22,200 centimetres of wire.

It is possible to determine the number of oscillations per second which can be achieved in any such circuit, and that is given by Hz = 5033000 / the square root of L x K where L is in centimetres and K is in microfarads. And so, knowing the inductance and the capacitance we get Hz = 5033000 / sqrt(22200 x 0.222) which is **71900 Hz**. That represents 72 oscillations in each of the 1000 sparks per second.

The current strength developed by these oscillations can be determined from the formula below which uses the Voltage **V** in volts, the resistance of the circuit **R** in ohms, the inductance **L** in henrys, and the capacitance **K** in farads:

Current = V / sqrt(R x R + (6.28 x Hz x L – 1/(6.28 x Hz x K)^2) or in our case:

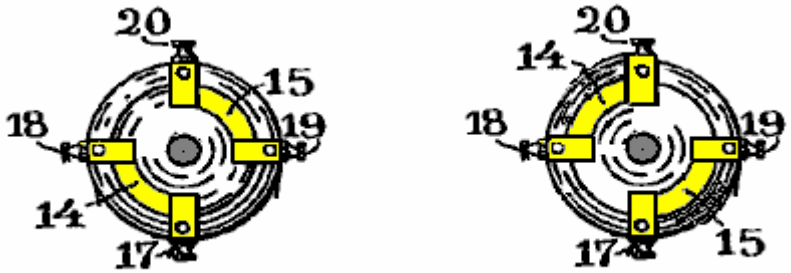
Current = 600 / $\sqrt{}$ 15 x 15 + (6.28 x 71900 x 0.0000222 – (1/(6.28 x 71900 x 0.00000222)))^2

Which works out as 40 amps, meaning that the impedance of the oscillating circuit is equal to the resistance in ohms of that circuit, since the inductive reactance and the capacitative reactance are so combined that the resultant value of the total reactance is equal to zero, and the only voltage required to produce the current of 40 amps is that needed to overcome the ohmic resistance of the circuit which is 15 ohms. This also means that the E.M.F. is in phase with the current, and therefore, the watts are a maximum.

Hence, there can be no doubt that battery **3-4** will be fully charged during the period when battery **1-2** is discharging, especially since the charging current can be further increased at will, even without taking more power from the discharging battery. In fact, it is easy to increase the number of turns in the transformer secondary **46-47** in order to increase the voltage. Obviously, as the amount of power delivered to the primary **44-45** of this apparatus is always 321 watts, if the voltage is increased, the quantity of electricity which the capacitor **49** receives will be reduced accordingly. Consequently, the value of that capacitor must also be reduced and so the number of oscillations per second will also be increased. Finally, by increasing the voltage, the current intensity is increased proportionately.

Therefore, it is always possible to combine, in the manner described, the values of resistance, inductance, capacitance and voltage in the oscillating circuit, so as to obtain the required current strength to fully charge one of the batteries during the period when the other battery is discharging.

Once this battery charging has been attained, if the plant is to continue in operation, then the batteries need to be swapped over by altering their connections to the circuit. To accomplish this, cylinder **5** is rotated until the conductive paths **11** and **12** come into contact with the brushes which are connected to the negative poles of the batteries, and then, battery **3-4** which is fully charged will be connected with the motor **27-28**, and its discharge will now be produced by this path: battery terminal **25**, ammeter **24**, commutator poles **20** and **18** (now connected together due to the rotation of the cylinder **5** through 90 degrees), wire **26**, motor **27-28**, variable resistor **29**, wire **30**, apparatus **40** and **41**, and commutator path **11** which closes the circuit to the negative pole of battery **3-4**.



In other words, the motor **27-28**, apparatus **40** and **41**, and alternator **42**, will still be operating in exactly the same way as described above, and in the same manner, the high-frequency currents are still developed, producing the same current intensity which is now passing via terminal **50** of capacitor **48**, arc-lamp **51-52**, diode **34**, wire **32**, commutator contacts **19-17**, wire **21**, ammeter **22**, and positive terminal **23** of battery **1-2** (which is now connected to contact **38** of the commutator through path **12**), wire **37**, diode **36**, inductor **55**, and capacitor terminal **49**; also through duplicate path from capacitor terminal **49**, diode **33**, wire **32**, commutator contacts **19-17**, wire **21**, ammeter **22**, positive terminal **23** of battery **1-2**, commutator path **12**, commutator pole **38**, wire **37**, diode **35**, arc lamp **52-51**, and capacitor terminal **50**.

It is obvious to state that this same procedure can be repeated indefinitely by the simple altering of the battery connections from time to time in accordance with the battery capacity and the rate of discharge. Only 400 watts is taken from the discharging battery for battery recharging, leaving 40 amps at 50 volts (2000 watts) available for doing continuous useful work.

The patent continues with a description of how a modified clock can be made to move the commutator once every hour. This is a brilliant 2 kilowatt, self-powered, free-energy design. However, operating the design as described would not be realistic. Batteries nowadays have limited operational lives offering, typically, between 400 and 1000 charge/discharge cycles within the C20 discharge current limits. Exceeding the C20 discharge rate will reduce the battery life by a major amount, that amount being determined by the degree of abuse which the battery suffers. If we ignore that factor and say that our batteries will manage 1000 cycles, at the proposed rate of say, one hour discharge and one hour charging time, then battery replacement is liable to be required in just 500 hours of operation. That is, within three weeks of continuous operation.

The essential, rapid switching of this circuit is performed by the spark-gap but in contrast to that, the commutator switching of the batteries does not require high speed operation. It is feasible then, to replace the commutator with simple solid-state switching and swap the batteries over every second or two. That way, the batteries are never discharged and long battery life can be expected.

Bozidar Lisac's Power-Boosting System.

Recently, a patent application has been lodged on what is effectively the **Ron Cole** one-battery switch and the Tesla Switch. I am including the re-worded patent here. Some experimenters have reported overall battery energy gains with switching speeds of 0.5 Hz or less, which means that in circuits of that type, mechanical switching should give a reasonable switch contact life. This patent has needed a fair degree of attention as the person writing it does not have a full grasp of English and confused the word "load" with the word "charge". Let me say again, that the following patent application is included here primarily for interest sake, rather than being the definitive way of making a circuit of this type.

Patent Application US20080030165 7th February 2008 Inventor: Bozidar Lisac

METHOD AND DEVICE FOR SUPPLYING A LOAD WITH ELECTRIC ENERGY RECOVERY

ABSTRACT

In the invention an electric current circulates from the battery **UB**, through the electric motor **M**, and the diode **D1** charges the capacitors **CA** and **CB**, connected in parallel, which, once charged, are connected in series, giving rise to a difference in voltage in relation to the battery, causing half the charge of the capacitors to be returned to the battery through the diode **D2**, whilst with a new parallel connection, the capacitors recharge, this charge being

equal to that which had been previously transferred from the capacitors to the battery, so that by means of the cyclic connection of the capacitors in parallel and series the energy is transferred from the battery to the capacitors and from the capacitors to the battery, thus considerably extending the range of the battery and operation of the motor.

OBJECT OF THE INVENTION

This invention relates to a method and device enabling the electrical energy with which a charge is supplied to be recovered using a self-rechargeable electricity source in which, which by means of a circuit, the current circulating from an accumulator or battery through a load, e.g. a motor, is fully returned to the same energy level, thereby considerably extending its range.

More specifically, two capacitors that are connected cyclically from parallel to serial and vice versa are charged through a motor during the connections in parallel, whilst in series connection, when its voltage doubles, they return the electricity, recharging the battery. This source represents a closed system which does not require an energy supply from the outside, except to compensate for the losses produced, the range of the battery being limited by the number of charges and discharges that the same technically permits.

BACKGROUND TO THE INVENTION

A load, such as an electric motor, is connected to a battery or accumulator with a certain charge, which will be progressively discharged by it, this discharge being directly proportional to the connection time and to the current circulating through the motor. It is therefore necessary to supply fresh energy from an external source to recharge it. Systems that enable the energy consumed by the load to be reused are not known in the state of the art.

DESCRIPTION OF THE INVENTION

A first aspect of the invention relates to a method for supplying a load with recovery of electrical energy, which comprises supplying a load with electrical energy deriving from the first electrical energy accumulator, and returning at least a proportion of that electrical energy after it passes through the load to the first accumulator for the purpose of recovering the energy supplied.

The electrical energy, after passing through the load, is recovered by second electrical energy accumulator, from where it is transferred to the first accumulator, giving rise to cyclic transfer of electrical energy between the first and second energy accumulators.

The recovery of energy from the second accumulator and transfer to the first accumulator may be achieved without passing the energy through the load. In another alternative implementation, the energy is recovered from the second accumulator and passed to the first accumulator through the load, in which case the polarity of the load is reversed during the recovery of energy through the load.

The transfer of energy is brought about by cyclically connecting two or more electrical energy accumulators between parallel and serial connections.

A second aspect of the invention relates to a device for supplying a load with recovery of electrical energy, which comprises a first electrical energy accumulator and a second electrical energy accumulator, where the load is connected between the first and second accumulators. The device may be provided in one embodiment with a unidirectional connection device, for example, a diode which is connected in parallel to the load, causing circulation of the electrical energy recovered after passing through the load, and via which the electrical energy is returned to the first accumulator.

The first electrical energy accumulator may be a battery. The second electrical energy accumulator might be two or more capacitors with switching to cyclically connect them between parallel and serial connection configurations.

The invention constitutes a self-rechargeable source of electrical energy which enables the range of a battery to be considerably extended so that the current circulating from the same through a motor charges two capacitors connected in parallel, up to the voltage level of the battery, by means of contacts. These capacitors, once charged, are connected in series, producing double their voltage, and they then return the energy to the battery, thereby extending its range. Once the losses have been compensated for, the duration of the extended range depends on the charging and discharging properties of the capacitors.

The existence of the difference in voltage between the battery and the capacitors connected both in parallel and in series, and which give rise to the displacement of energy from the battery to the capacitors and vice versa, is

used to supply the motor connected between the battery and the capacitors, comprising the self-rechargeable source of electrical energy.

When connected in parallel, the capacitors are charged through a motor and a diode, and when connected in series, they are charged through another diode, the voltage of the motor being half that of the battery. On the other hand, if the motor is connected between the battery and the serially-connected capacitors, the latter, which are charged in parallel through a diode and are discharged by means of the motor and the other diode, will supply the motor with a voltage equal to that of the battery, whilst a capacitor connected in series to the winding of the motor guarantees its operation without loss of power.

Instead of the two capacitors, two batteries connected in series and another two connected in parallel may be used, between which batteries a motor is connected, the current circulating in this case from the batteries connected in series through the motor to the batteries connected in parallel. The serially-connected batteries are then connected in parallel, by means of switching contacts, and the other two parallel-connected batteries are then connected in series, reversing the direction of the current, whilst the connections of the motor are inverted by means of the simultaneous switching of other contacts in order to maintain the polarity and direction of rotation of the motor.

In one possible embodiment of the invention, another two capacitors and a transformer with two primary windings, or a motor with two windings are added to the device previously described, each pair of capacitors cyclically switching from parallel to serial connection and vice versa so that during the parallel connection cycles, two of the capacitors are charged through one of the windings up to the voltage level of the battery at the same time that the other two capacitors are connected in series, double their voltage and are discharged by means of a second winding to the battery.

The reduced level of energy losses brought about mainly by the dissipation of heat and in the capacitors, as well as by the charge factor of the batteries, is compensated for from an external source, and because the sum of the current circulating through a winding of the motor or transformer charging two of the capacitors and the current simultaneously circulating from the other two capacitors through the second winding, recharging the battery, plus the current which is supplied from the external source, is equal to zero, because of the work carried out by the motor or the loads which are connected to the alternating voltage induced in the secondary of the transformer, no discharge of the battery takes place.

DESCRIPTION OF THE DRAWINGS

In order to supplement the description now being given, and with the aim of contributing to a better understanding of the characteristics of the invention, according to a preferred practical embodiment, a set of drawings is attached as an integral part of this description, in which, for informative and non-restrictive purposes, the following is shown:

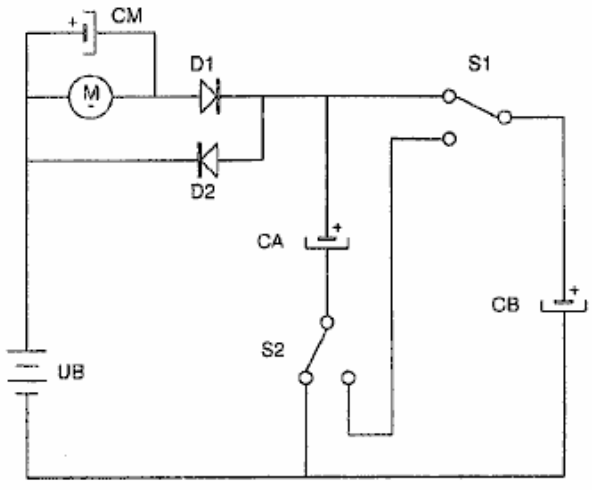


FIG.1

Fig.1 shows a practical circuit in which, by means of switching, two capacitors connected in parallel are charged from a battery through a motor and a diode, and after the contacts are switched, they are connected in series, thereby discharging the battery through another diode.

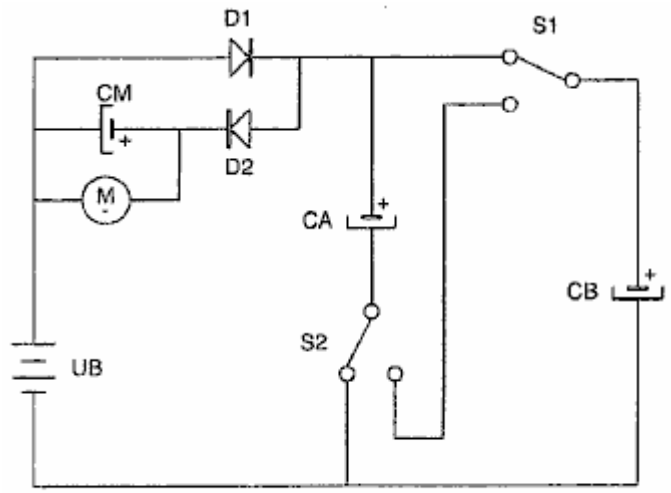


FIG.2

Fig.2 shows a practical circuit in which, through switching, the two capacitors are connected in parallel and are charged from a battery through a diode, and after the switching of the contacts they are connected in series, thereby charging the battery through the motor and the other diode.

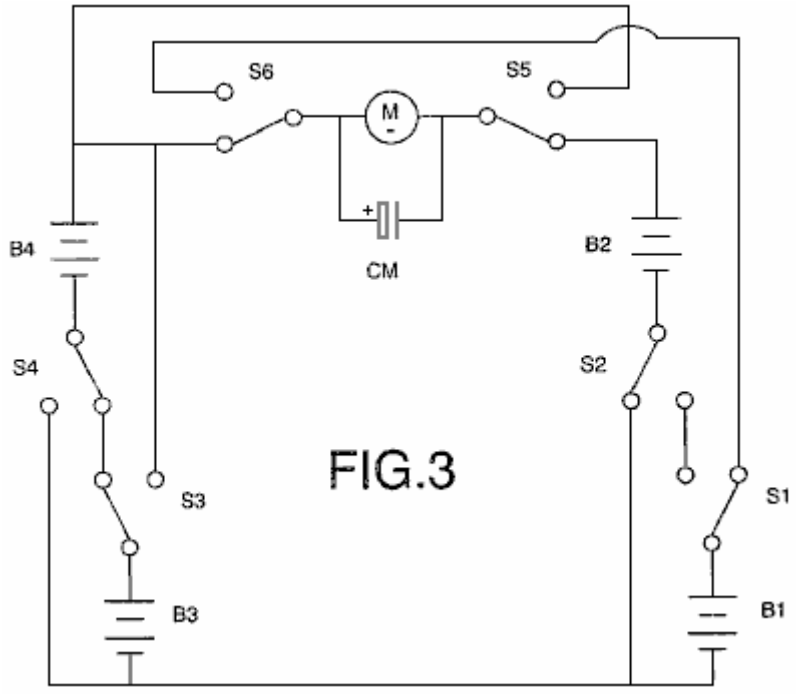


FIG.3

Fig.3 shows the connection of the two batteries in series, connected through a motor to another two batteries connected in parallel, and which, by means of contacts, switch alternatively, this giving rise to effects similar to those described in relation to the use of the capacitors.

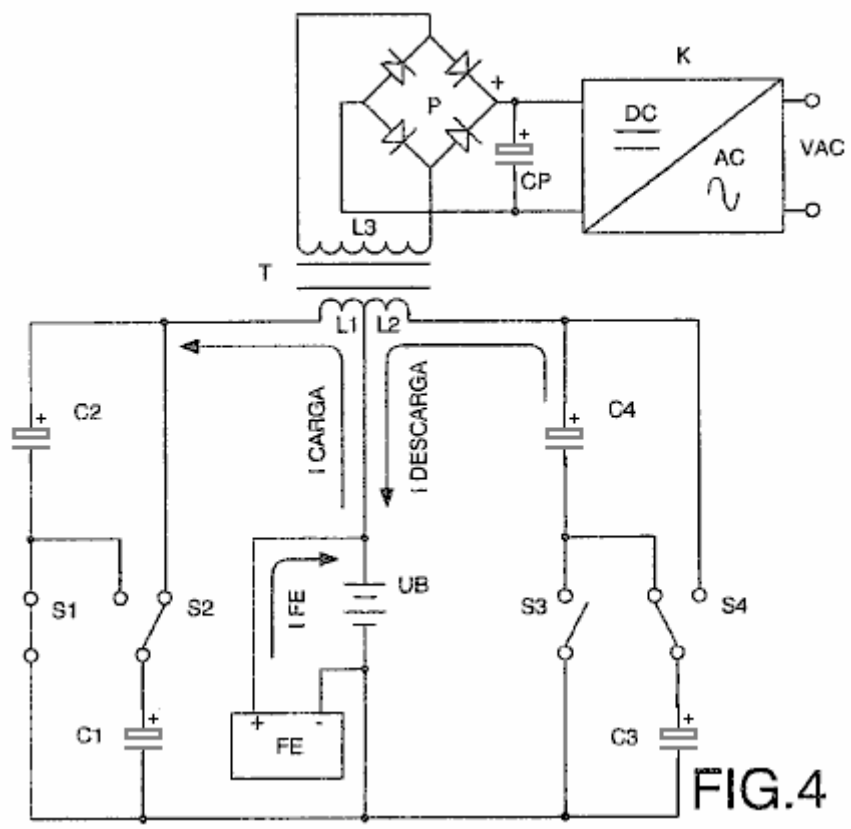


Fig.4 shows the electrical diagram corresponding to the connection between the battery and the two pairs of capacitors of a transformer with two primary and one secondary winding, in which an alternating voltage is induced which is rectified, filtered and converted to a sinusoidal voltage.

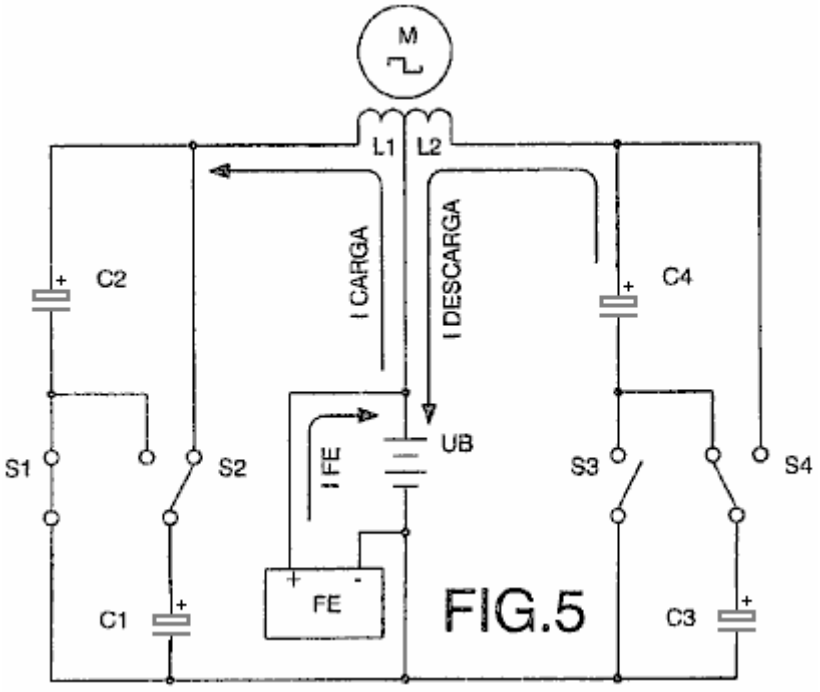


Fig.5 shows the electrical diagram of an alternating current motor with two windings connected between the battery and two pairs of capacitors.

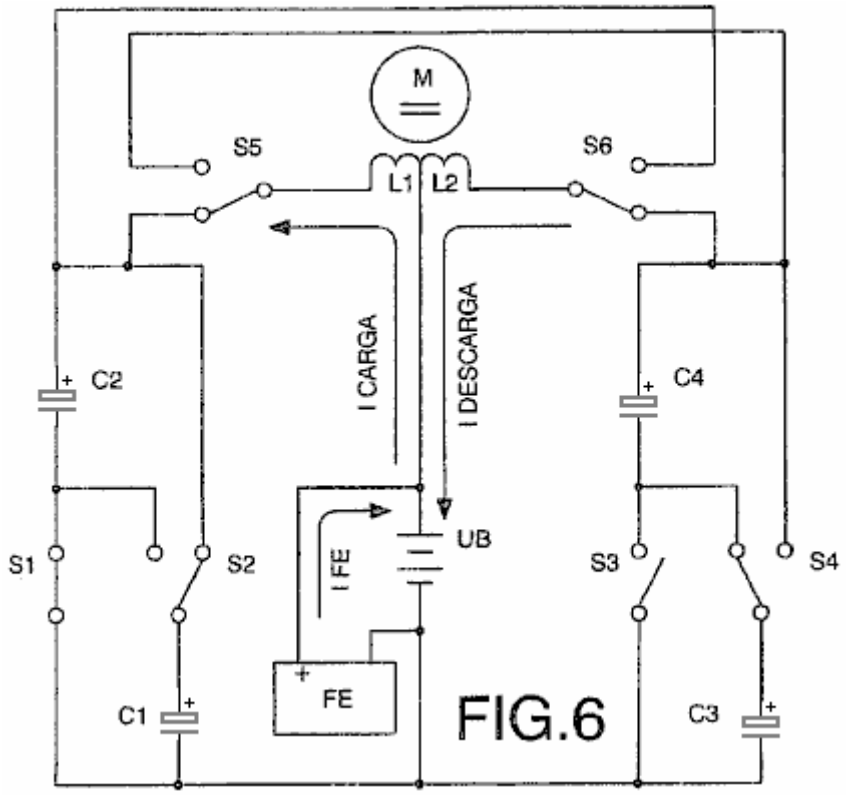
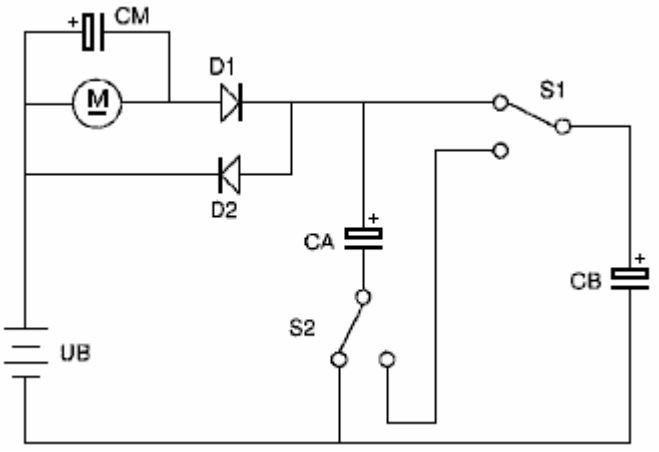


Fig.6 shows the electrical diagram of a direct current motor with two windings connected between the battery and two pairs of capacitors, in which two switch contacts ensure their correct polarisation and direction of rotation.

PREFERRED EMBODIMENT OF THE INVENTION

In a preferred embodiment shown in **Fig.1**, the load consists of a direct current motor **M**, the battery **UB**, and the second accumulator which consists of a pair of capacitors **CA** and **CB**. The capacitors **CA** and **CB** are connected to each other in parallel by means of two switches **S1** and **S2**. These capacitors are charged through the motor **M** and diode **D1** to a voltage level equal to that of the battery **UB**, the charge being Q = (CA+CB)UB, and while these capacitors are being charged, the motor **M** is rotating.

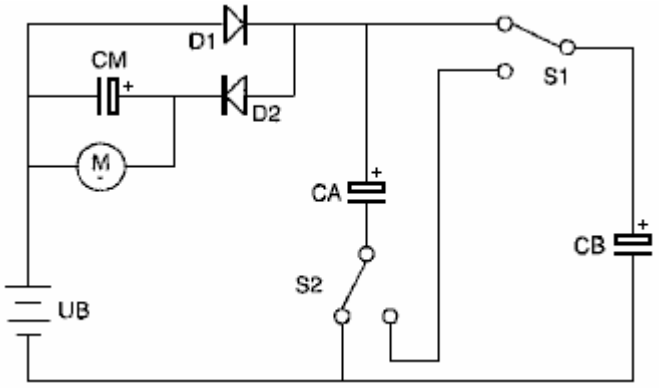


When both capacitors are fully charged, they are connected in series by the switch contacts **S1** and **S2**. This produces a voltage which is twice the value of the voltage of the battery **UB**, resulting in the charge which is given by Q = 2 x UB x (CA+CB) / 2 which is Q = (CA+CB)UB, which shows that once charged, the charge Q of both capacitors is identical both in parallel and in series.

Diodes **D1** and **D2** ensure that current flow through the motor **M** is only ever in one direction. Immediately after capacitors **CA** and **CB** are connected in series, they return half of their charge through diode **D2**. Switches **S1**

and **S2** then connect the capacitors **CA** and **CB** in parallel. In this arrangement, they start off with half of the battery voltage. They charge immediately, regaining the battery voltage through the motor **M** and the diode **D1**.

By means of repeated cyclic switching of the capacitors **CA** and **CB** from parallel to serial connection mode, the current circulating from the battery **UB** through the motor **M** to the capacitors, and from these to the battery, recharging it and extending its range, constitutes a self-rechargeable source of electrical energy.



In a second practical embodiment shown in **Fig.2**, the motor **M** is connected between the battery **UB** and the capacitors **CA** and **CB** by means of the diode **D2**. The capacitors are charged directly through the diode **D1** and are discharged through the motor **M** and the diode **D2**, the values of the charges on the capacitors **CA** and **CB** previously described in the example shown in **Fig.1** remain unchanged, the difference in this circuit is that the voltage applied to the motor **M** is the full battery voltage in this case.

The charging rate of the capacitors **CA** and **CB** is determined by the intensity of the current flowing through the motor **M**, to which is connected in parallel, the capacitor **CM** which guarantees that the operation of the motor is maintained at maximum power. It is possible to substitute a battery, preferably a rapid charge battery, for capacitor **CM**.

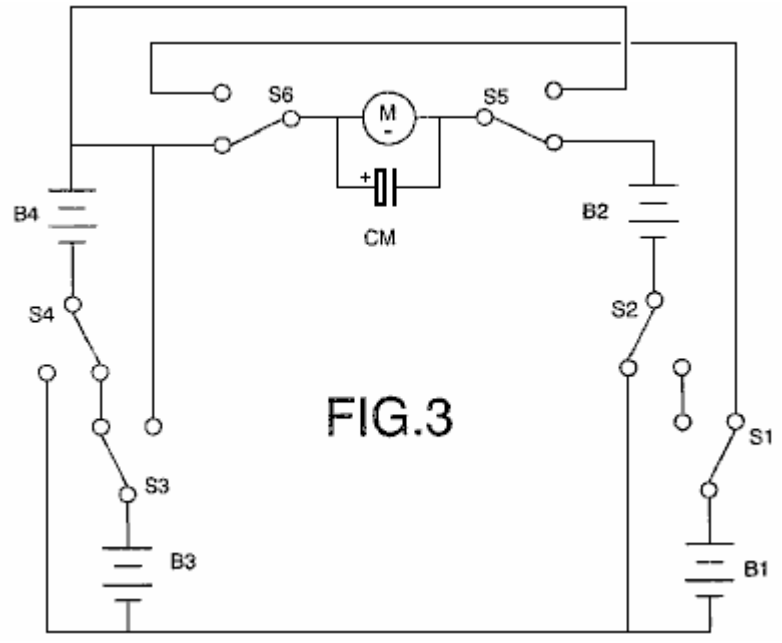


FIG.3

In another embodiment shown in **Fig.3**, the first and second accumulators consist of pairs of batteries **B1**, **B2** and **B3**, **B4**. Therefore, in this embodiment, two pairs of batteries are used instead of the capacitors **CA** and **CB**. Batteries **B1** and **B2** are connected to the switches **S1** and **S2**, and the pair of batteries **B3** and **B4** are connected to the switches **S3** and **S4**. The switches **S1** to **S4**, connect the pairs of batteries with which they are associated, into series or parallel configurations, depending on the position of the switches.

While the batteries **B1** and **B2** are connected in parallel, the other two batteries **B3** and **B4** are connected in series, and the motor **M** rotates as a result of the difference in voltage between the batteries, as it is connected

between both pairs of batteries. At the same time, the current circulating through the motor from the serial connected batteries recharges the two parallel-connected batteries. The switches **S1** to **S4**, which connect the batteries **B1** and **B2** in series and the batteries **B3** and **B4** in parallel then switch, thus reversing the direction of the current flow, and at the same time, the switches **S5** and **S6** change positions in order to maintain the correct polarity for the motor and its direction of rotation.

The two capacitors and the batteries may be switched by means of any mechanical, electromechanical, electrical, electronic or other element that meets the conditions described with the purpose of obtaining a self-rechargeable electrical energy source. These switching operations may be controlled by any known method, for example, a programmable electronic circuit.

In the preferred embodiments previously described, the load consists of a direct current motor, but as an expert in the field may understand, the load may also consist of any type of resistive (?) and/or inductive load.

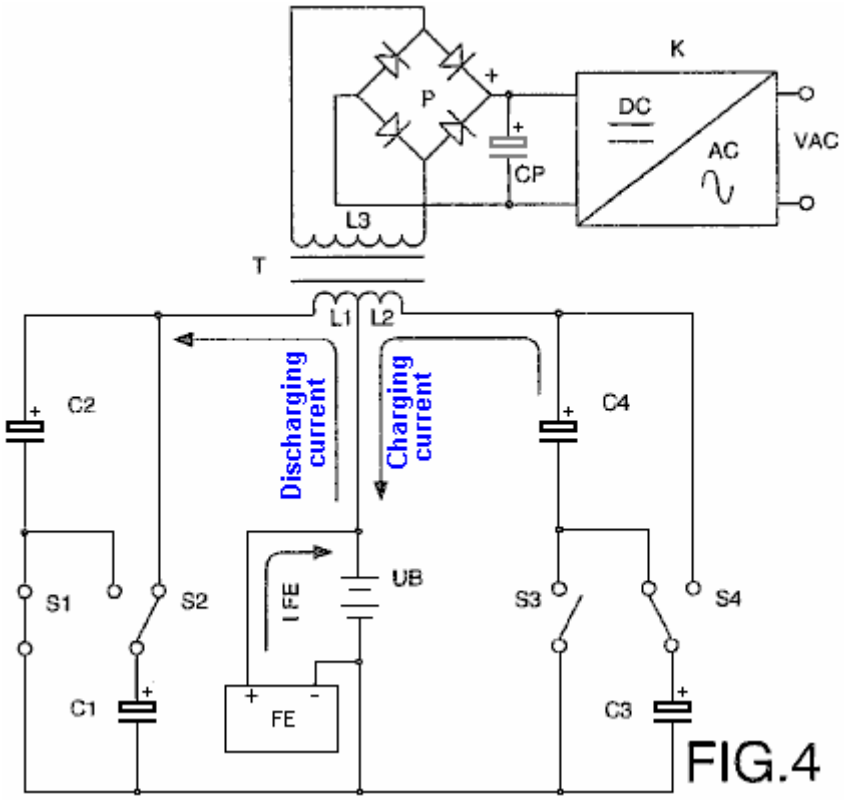


FIG.4

Another preferred embodiment is shown in **Fig.4**, where a transformer **T** with two primary windings **L1** and **L2** is connected between the battery **UB** and the two pairs of capacitors **C1** and **C2**, plus **C3** and **C4**, causing the two capacitors **C1** and **C2** to switch their connections from parallel to serial and back again by means of the contacts **S1** and **S2**, and causing the capacitors **C3** and **C4** to switch by means of contacts **S3** and **S4**, so that during the cycles of connection of the capacitors **C1** and **C2** in parallel, the latter are charged via the winding **L1** up to the voltage level of the battery, whilst at the same time the capacitors **C3** and **C4** are connected in series and provide double their voltage, the battery being discharged by means of the winding **L2**, in which case the charging and discharging currents to circulate in the same direction. On the other hand, during the cycles of connection in parallel of the capacitors **C3** and **C4**, which are charged through the winding **L2** up to the battery voltage level, the capacitors **C1** and **C2** are connected in series to provide double their voltage and are discharged into the battery through the winding **L1**. The direction of the charging and discharging current therefore changes, thus inducing in the secondary winding **L3** an alternating voltage whose frequency depends on the speed of switching of the contacts mentioned, and after being rectified by means of the bridge of diodes **P** and filtered by the capacitor **CP**, the resultant DC voltage is converted to a sinusoidal voltage by means of a circuit **K**.

The connection in parallel of one pair of capacitors and the connection in series of the other pair take place at the same time. Therefore the sum of the current circulating from the battery through one of the windings, charging two of the capacitors, and the current circulating from the other two capacitors through the other winding to the battery, is approximately zero.

From an external energy source **FE** the minimum energy losses caused essentially by dissipation of heat and in the capacitors, as well as by the charging factor of the battery, are compensated for, with the result that the sum

of the current circulating from this source external to the battery and the charging and discharging currents of the capacitors is equal to zero. Therefore the battery is not discharged and its range does not depend on the work developed by the motors or the loads connected to the secondary winding **L3** of the transformer **T**, since the greater the power of the loads, the higher the intensity of the charging and discharging currents of the capacitors.

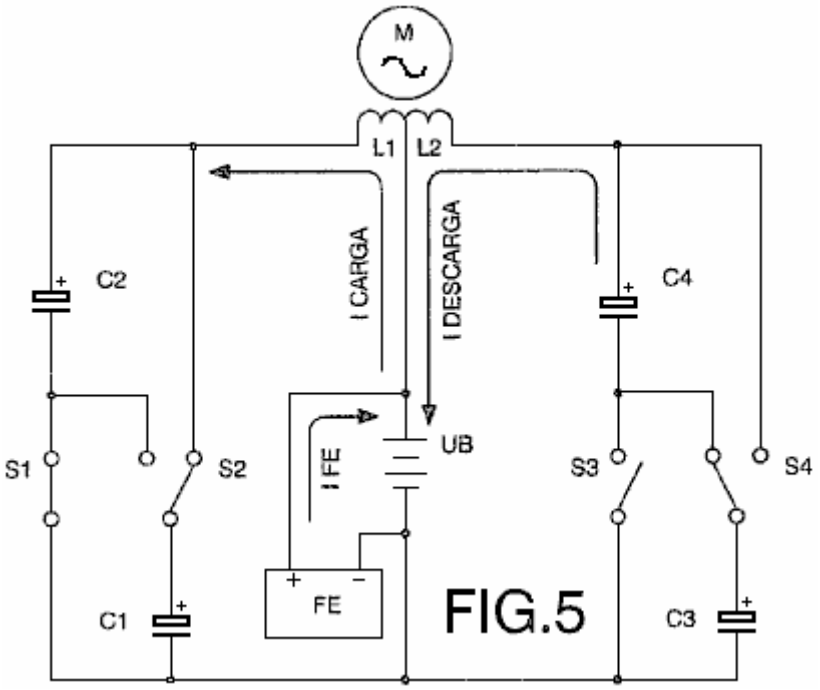


FIG.5

Fig.5 shows another embodiment in which an alternating current motor **M** is connected to two windings **L1** and **L2** so that during the connections in parallel of the capacitors **C1** and **C2**, the latter are charged by means of the winding **L1** at the same time that the capacitors **C3** and **C4**, connected in series, are discharged by means of the winding **L2** to the battery **UB**, the charging and discharging current circulating through the windings in the same direction. The capacitors **C1** and **C2** are then connected in series and the capacitors **C3** and **C4** are connected in parallel. The direction of the charging and discharging current of the capacitors is therefore reversed, thus producing at terminals of the motor an alternating voltage with a frequency that depends on the speed of switching of the contacts. The energy losses caused are compensated for from an external source **FE**, the sum of the current circulating from this source to the battery and the currents circulating through the two windings during charging and discharging of the capacitors being equal to zero. The battery is therefore not discharged as a result of the work developed by the motor.

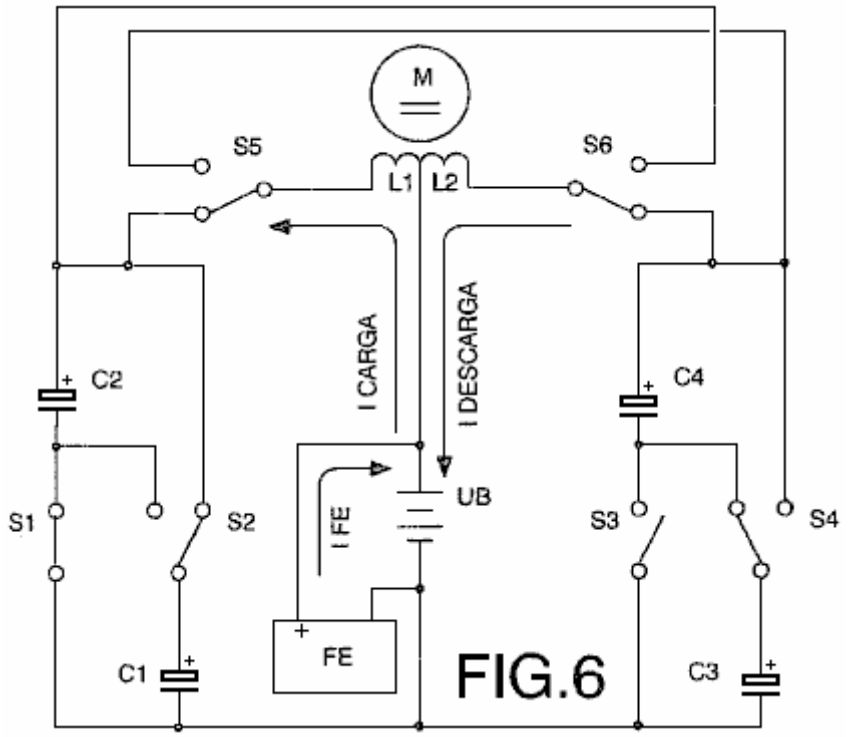


FIG.6

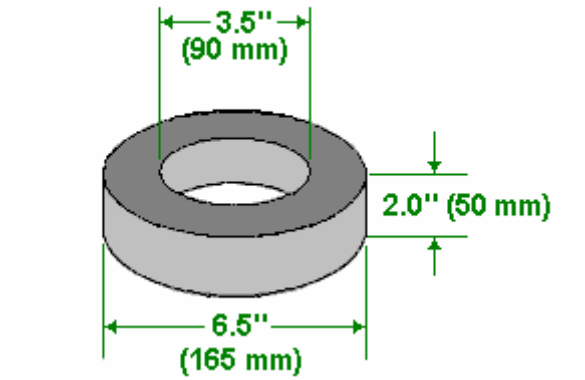
Fig.6 shows the connection of a direct current motor **M** to two windings **L1** and **L2** between the battery **UB** and the two pairs of capacitors **C1** and **C2** plus **C3** and **C4**, so that during the connections in parallel two of the capacitors are charged by means of the winding **L1**, and during the simultaneous connections in series, the other two capacitors are charged by means of the winding **L2** to the battery. Coinciding with the switching of the contacts **S1**, **S2**, **S3** and **S4**, which connect to each pair of capacitors from parallel to serial and vice versa, the contacts **S5** and **S6** switch, polarising the windings of the motor so that the charging and discharging currents of the capacitors circulate in the same direction, producing a direct voltage. The sum of the current supplied from the external source **FE** and the charging and discharging currents of the capacitors is equal to zero, and thus there is no battery discharge.

Bob Boyce's Toroid.

Consider also, Bob Boyce's very effective pulsed toroid system. As the waveform fed to the toroid has to have very sharp rising and falling voltages, the toroid needs to be able to handle very high frequency signals, far higher than the number of pulses per second fed to the toroid. If the rising edge is very sharp (and it needs to be so fast that it won't show on a 150 MHz oscilloscope), then as far as the toroid is concerned, there may be a similar falling edge one nano second later and so it needs to be able to respond to that sort of frequency. Consequently, the material and the windings need to be selected very carefully.

The toroid is a 6.5 inch iron-dust unit from MicroMetals, part number "T650-52" and it can be purchased through their web site: http://www.micrometals.com/pcparts/torcore7.html and it can be purchased in small quantities via their "samples requests", which can be submitted at http://www.micrometals.com/samples_index.html

The Micrometals T650-52 Toroidal Core



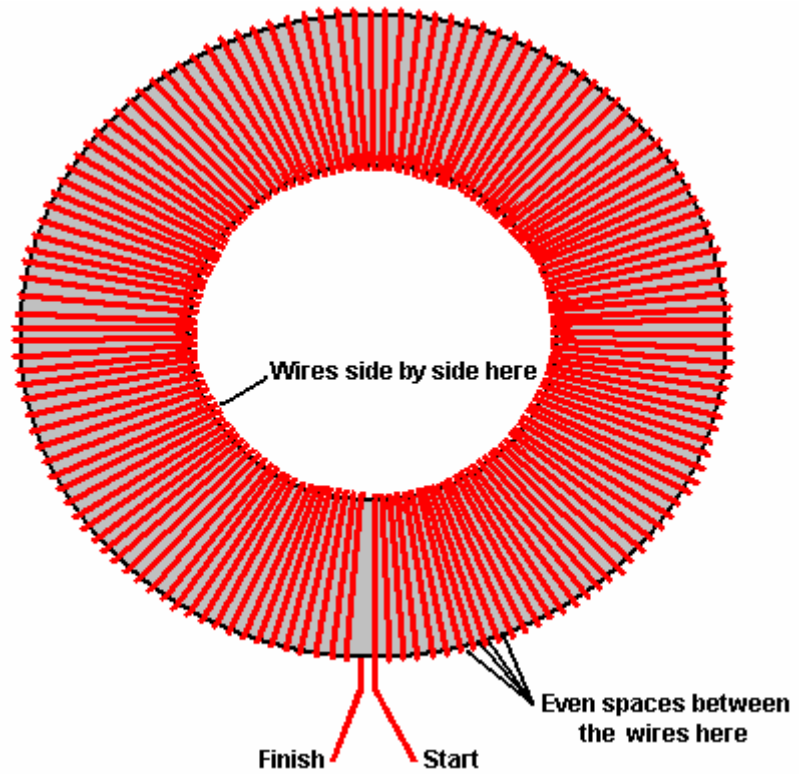
There are four windings made on this core. The wire chosen to wind the transformer is most important. Bob uses **solid** teflon-covered silver-plated copper wire. It is very important that this wire is solid core and **not stranded** as stranded wire does not work here (due to the generation of inter-strand, phase-differential induced eddy currents). At this time, a supplier of this wire is http://www.apexjr.com.

Before any winding is done, the toroid is given a layer of the yellow 1P802YE winding tape available in 3" rolls, both the 1" and 2" widths from: http://www.lodestonepacific.com/distrib/pdfs/tape/1p802.pdf. It is very important to avoid using fiberglass winding tape anywhere in the construction of this wound toroid. Bob comments on this as follows: "Big warning here !!!! **DO NOT USE FIBERGLASS WINDING TAPE !!!!** A big box of 3M winding tape was ordered by accident so I tried it to see if it would work. It not only suppressed the acousto-resonance response of the entire wound toroidal core, but for some strange reason it also caused the electrostatic pulse response of the secondary to **reverse polarity** as well as reducing the signal amplitude to a mere 10% of what it was !! It totally negated the benefit of the teflon insulation".

Having covered the toroid with a layer of the 1P802YE winding tape, the secondary winding is made. Again, it is very important that the teflon-covered, silver-plated solid copper wire is used. This is not a system which provides COP>1 performance if any old components are thrown together carelessly during the building process.

The winding turns must be evenly spaced where they fan out from the center of the core. They are tightly packed side by side in the centre opening and they must be wound tightly and the gaps between adjacent turns along the outer edge must be exactly the same. This is not to make the winding look "pretty" but if this is not done, then it will cause magnetic field errors that will lower the overall efficiency when the toroid is being used.

The secondary winding is made using 16 gauge wire which covers the entire length of the toroid as shown here:



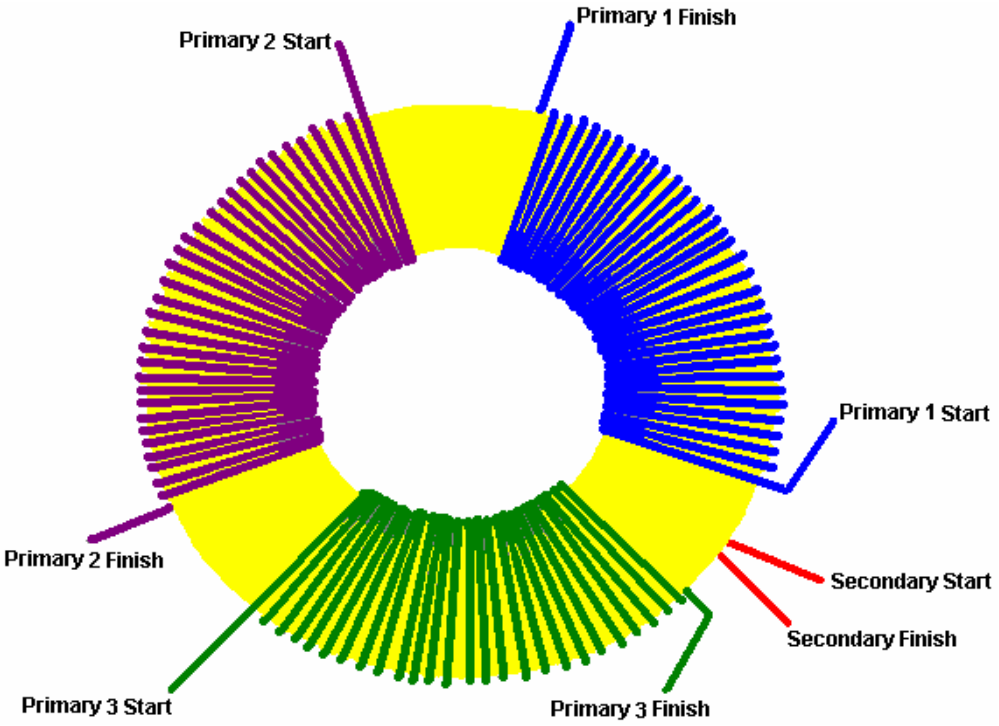
If the spaces between the wires are not quite even, then the turns can be pushed into exactly the right place. It is sometimes convenient to use two-inch lengths of plastic strimmer line placed between the turns in order to get the spaces between the turns exactly the same. These can be held in place with a strip of the yellow winding tape:

The picture above has been taken to show what a partially prepared secondary winding looks like when its windings are being moved into their exact positions. When a section of the windings has been spaced accurately, the triangular gaps between the evenly-spaced turns are filled in with beeswax, made pliable using a heat gun. A plastic bottle pushed into the central hole can be helpful when doing this filling. When the beeswax has hardened on both sides of the toroid, the process is then repeated for the next group of turns.

When the winding is complete, with even spacing of the turns and the gaps filled with beeswax, the whole of the toroid is then covered with a layer of the yellow winding tape, as shown here:

So, to recap: the toroid is wrapped in tape, the secondary winding completed, extending the entire way around the toroid, the windings carefully spaced out so that the gaps around the outer edge of the toroid are exactly equal, the winding gaps filled with beeswax, and then the toroid covered with a layer of the yellow tape. There will normally be anything from 127 to 147 turns in the secondary winding due to manufacturing tolerances in the insulation of the wire and so the overall wire length will be about 100 feet.

The primary windings are now wound on top of the tape layer covering the secondary winding. As with the secondary, the direction of the turns is very important.



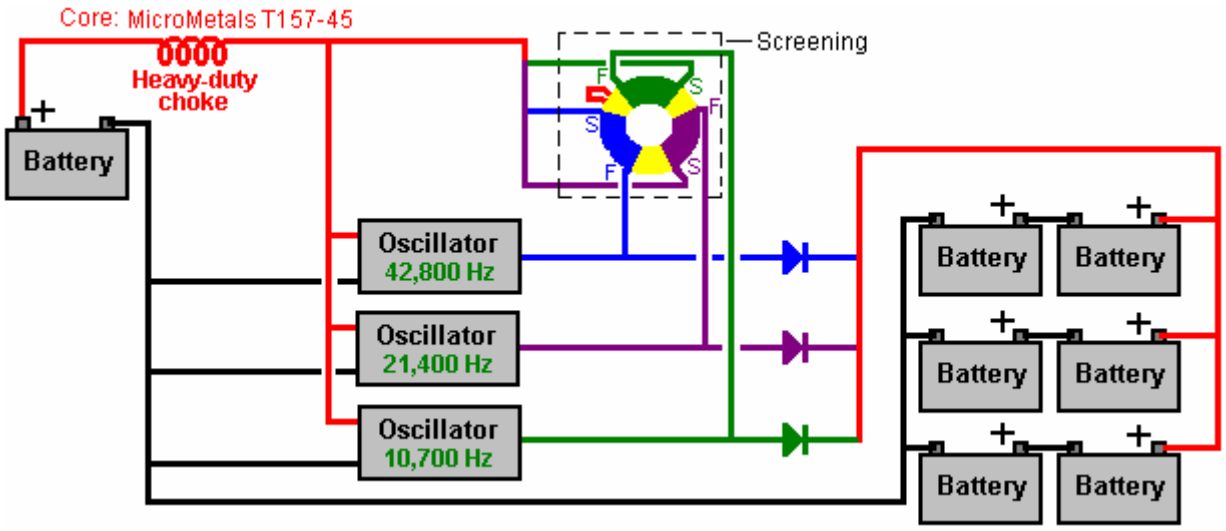
Please note that every winding starts by passing **over** the toroid and then being brought up on the outer side ready for the next turn. Each of the following turns proceed in a counter-clockwise direction, and finishes by passing **under** the toroid. Every winding is created in this way and the quality of workmanship is very important indeed when making these windings. Each winding needs to be tight and positioned exactly with turns touching each other in the centre of the toroid and positioned on the outer edge with exactly equal spaces between each turn. Your construction work **has** to be better than that of a commercial supplier and needs to reach the quality demanded by the military.

The three primaries are wound on top of the tape which covers the secondary winding. These three primary windings are spaced out equally around the toroid, that is, at 120 degree centres and the leads of the secondary winding are taken out through the gap between two of the primary windings and **not** taken out through the middle of a primary winding. As with the secondary winding, the primary winding turns are spaced out exactly, held in place with beeswax, and then tightly taped over. The primaries can have more than a single layer, and they are

wound with the same direction of winds as the secondary, and with the same care for exact turns spacing as needed for the secondary winding. Tape the entire core well with tightly-stretched PVC electrical tape after winding, to ensure that the primary windings do not move and then add an outer layer of the winding tape.

This toroid pulls in additional energy from the immediate environment when driven by very high quality voltage pulsing applied to each of the three primary windings. The full details of this system have not yet been disclosed, but Bob has said in open forum that he has demonstrated his toroid being pulsed with the secondary not connected to anything and the output is triple the current at twice the input voltage, with is COP=6. When the ends of the secondary are joined together the output current doubled, giving COP=12, that is, twelve times as much power output as the input which Bob had to supply to get that output. This is, of course, not a case of energy being created (which is not possible) but instead, it is a case of eleven times the input power being drawn in from the surrounding environment.

I have never seen the circuitry for this, but it may be as shown here:



As the output voltage is doubled, the battery bank being charged can be double the voltage of the battery supplying the input power. The choke in the lead from the input battery positive is to direct the power generated to the charging batteries. Each primary winding is driven by it's own separate oscillator which has adjustable frequency and adjustable Mark/Space ratio, or "Duty Cycle". The Duty Cycle is set initially to about 25% which means that the input power is turned off for three quarters of the time. The highest frequency oscillator is adjusted to give the largest charging output current. Then, the middle frequency oscillator is adjusted to give the highest charging output. Finally, the lowest frequency oscillator is adjusted to give the highest charging output. When that has been done, the Duty Cycle of each oscillator is lowered as far as possible without lowering the charging rate. An oscilloscope is not needed for any of these adjustments.

It should be noted that what was marked as the "Start" of each primary winding in the diagrams showing how each winding is done, is connected to positive power supply line and the end marked "Finish" is connected to the output of the oscillator. This connection arrangement is very important because connecting them the other way round is likely to give a much lower performance.

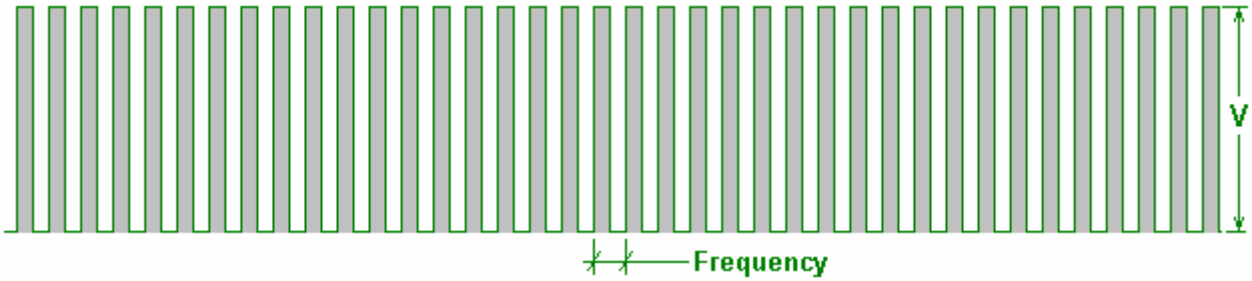
The three oscillators are quite independent of each other and are free-running. In other words, they are not synchronised in any way and will produce all sorts of different combinations of composite output waveforms as a result of this arrangement. It must be noted that this is **not** a system which produces a rotating magnetic field as the windings are **not** pulsed sequentially. As already mentioned, the voltage output waveform from each of the oscillators must have very sharp rise and fall times and the output must, of course, be capable of supplying sufficient current to power the primary windings.

I have never seen an explanation of how this system works, and please remember that the circuit shown above is based on my guesswork and does not come from Bob Boyce. However, I will now suggest a possible mechanism for the way that the circuit works, and for it, let us assume that there is only the highest frequency oscillator connected: When there is an output pulse from the oscillator, a powerful current passes through the primary winding to which it is connected. This generates a strong magnetic pulse. But, since the winding is around a high-quality toroidal core, nearly all of the magnetic flux races around the toroid instead of radiating outwards. That magnetic pulse induces an electrical pulse in both of the other two primary windings, and so all three windings provide a current surge to the battery bank being charged. One pulse in, produces three pulses out,

which matches the observed tripling of the input current.

(This suggests that if there were four primary windings that there would be an increased COP result for that arrangement. If that were done, then the fourth oscillator might run at around 85,600 Hz). The short-circuiting of the ends of the secondary winding produces a further increase in the output. I would suggest that this may be due to the fact that the turns ratio of the primary-to-secondary winding, produces a much higher voltage in the secondary winding. If the ends of the secondary winding are connected, then that induced voltage will generate a strong electrical current flowing through the secondary winding. That current will in turn, generate an even greater magnetic pulse, both in the toroid and in the primary windings which are wound around the secondary winding. This enhanced magnetic pulse may account for the enhanced electrical output to the battery bank being charged. Please remember that this not a matter of fact, but just a suggestion which I am putting forward as being a possible explanation of how the circuit is functioning.

Please remember that the toroid has to be able to handle frequencies far higher than the pulsing rate which is applied to it. A high-frequency waveform looks like this:



If you apply that frequency to Bob's toroid, then the toroid needs to be able to handle the waveform cleanly, without degrading it in any way. For this sort of application, an iron-dust toroid such as the MicroMetals product is essential. What many people have difficulty in seeing, is that even if the overall frequency of the signal is lower, as shown here:

in order for the rising edge of the waveform to be handled cleanly, the toroid has to be able to handle a very high frequency signal. The toroid doesn't "know" that the leading edge of the waveform is not going to be followed by a whole stream of very short, very fast, high-frequency pulses. So the toroid has to be able to handle high-frequency waveforms in order to deal with the very sharply rising leading edge which is essential for the successful operation of this and many other free-energy devices.

Bob Boyce has experimented with the much more dangerous rotating magnetic fields type of circuitry and he says:
While working as an engineer for a government subcontractor, I became aware of a problem with a switcher power supply, which under certain temperature and load conditions, would go into over unity operation. Sometimes these would fail in a big way and totally destroy the load. There was far more damage than the power source to the power supply was capable of delivering. The power supply used an on-board toroidal coil wound with teflon insulated, silver-plated solid copper wire on Honeywell powdered iron cores. I guess that this was what started my interest in this. When I went into business for myself, I had the time to devote to research and experimentation.

My early radiant energy research and experimentation were with attempting replications of Tesla and Moray devices, as well as many others. I found a copy of Dr. Hans A. Niepers' book "Revolution" ISBN 3-925188-07-X. In that book there is information about many devices related to the conversion of gravity field energy.

My first foray into a 3-phase device was during experimentation with a g-strain energy absorber as theorised by professor Shinichi Seike. Needless to say, the results were not what I expected when I connected a 3-phase air core toroid to the 3-phase outputs of the g-strain energy absorber board.

I had been doing these experiments at home and that was where my lightning strike injury occurred in 1995. That put an end to my 3-phase research for a number of years. When I did start back into it, I soon tired of the expense of burned out electronics and loads. I went back to the lower powered single-phase transformers and stuck with that until I worked out some of the control issues which would allow me to get back into the 3-phase design in a more controlled manner. I wish that I had heard of Steven Mark sooner as that would probably have saved me some time. Like Steven, I learned to have a very healthy respect for this stuff.

The interesting part is that the experiment was not running at the time of the lightning strike, but the strike did enter the room where I had been running an earlier experiment with the 3-phase toroid driven by the g-strain energy absorber board. I am under the impression that the running of the experiment had somehow imprinted an energy pattern upon me which may have attracted the lightning to me. The lightning entered the room, entered my right hand, exited my right ankle and left the room via the telephone line. I was quite understandably shaken up, but I was able to drive myself to St Mary's Hospital (in West Palm Beach, Florida), where I was treated in the emergency room, and released. The pain in my chest afterwards was something which I would not wish upon anyone! All I could really do was to take pain pills until the pain subsided. Considering that I had a well-grounded (by three 20-foot long half-inch diameter copper pipes, one connected to each leg of the base, and a 20-foot long half-inch diameter copper pipe at each guy point) 160-foot radio tower only twenty feet away from that room, the entire house should have been protected under the "cone of protection" provided by that tower. That particular lightning strike defied all common sense of good lightning protection. It was as if it targeted me directly, despite the well-protected location. I have had avalanches occur since then, but I make darned sure that I am very well protected against lightning before running any of these experiments now. I have had the metal building struck but it has not penetrated inside. There are 8-foot long, five-eighths inch diameter copper clad steel ground rods at each corner of the building. I have an 8-foot by 10-foot PVC utility building in which I can run an experiment while I watch on a closed-circuit TV monitor, from the safety of my metal building.

The core does provide stability. I use a low-permeability core in order to avoid saturation from the bias. The lower the frequency, the fewer impulses (kicks) there are per second which results in a lower power density but it also lowers the relativistic effects in the core. So sure, it is a trade-off of control versus power. If the controller were to be fitted totally within the "eye of the storm", that is, in the centre of the toroid space, then the power density could be greatly increased. My main concern would be losing control in a runaway situation with the controller ignoring command signals from the outside due to these relativistic timing changes. Energy induced in even DC control wiring can override those control signals. You would not want to be anywhere near that if it happened.

I do not like to use ferrite or laminated iron cores. With their high permeability, they can only be used at very low frequencies and very low power densities.

One of the features of this power source is that it seems to adapt to the load (within reason, of course). Load impedance is fairly important for getting maximum power output, due in part, to the High-Frequency portion of the energy riding on the DC output. Momentary short-circuiting of the output does result in a plasma-like discharge arc.

From what I can see, my toroid and that of Steven Mark do look to be very similar. The primary differences that I can see are the core material and the number of poles. It sounds and looks like Steven Mark used stranded copper wire and 4 poles, while I use powdered iron and 3 poles in my current device. Otherwise, both seem to build up a swirling electromagnetic vortex during operation.

Over the years, I have used other core materials with varying degrees of success. I started out with laminated iron cores on 2-phase devices bin the mid 1980s and ten progressed to air-core 3-phase devices in the early 1990s. One of the first units on which I tried 3-phase operation was a Seike "g-strain energy absorber" which I had connected to a hand-wound 3-phase air-core toroidal coil. Talk about uncontrolled operation! But I kept at it, unaware of the dangers. Since I was no longer doing the HHO gas research at that time, I was obsessed with trying to replicate Tesla and Moray research on a shoestring budget. After my lightning strike injury in 1995, I mothballed that line of research until I had the time, energy and funds to continue. I refused to apply this 3-phase design to practical application until the control issue cold be solved.

One technique which can be used to improve the power density at low frequencies is to go from 3 primaries to 6 primaries, that is, two phased sets of 3-phase drive coils.

In March 2012 a similar style of circuit design is to roll out the first 200 commercial units from a company in South Africa. The http://pesn.com/2012/02/22/9602042_South_African_Fuel-Free_Generator_Preparing_for_Market/ web site of Stirling Allan covers many of the details of these units. The expected price for a self-powered 5 kilowatt unit is US $6000 and there are other units of up to 40 kilowatt output. Stirling has visited the company in South Africa and witnessed the device in operation and was due to receive one of these units in March 2012, but

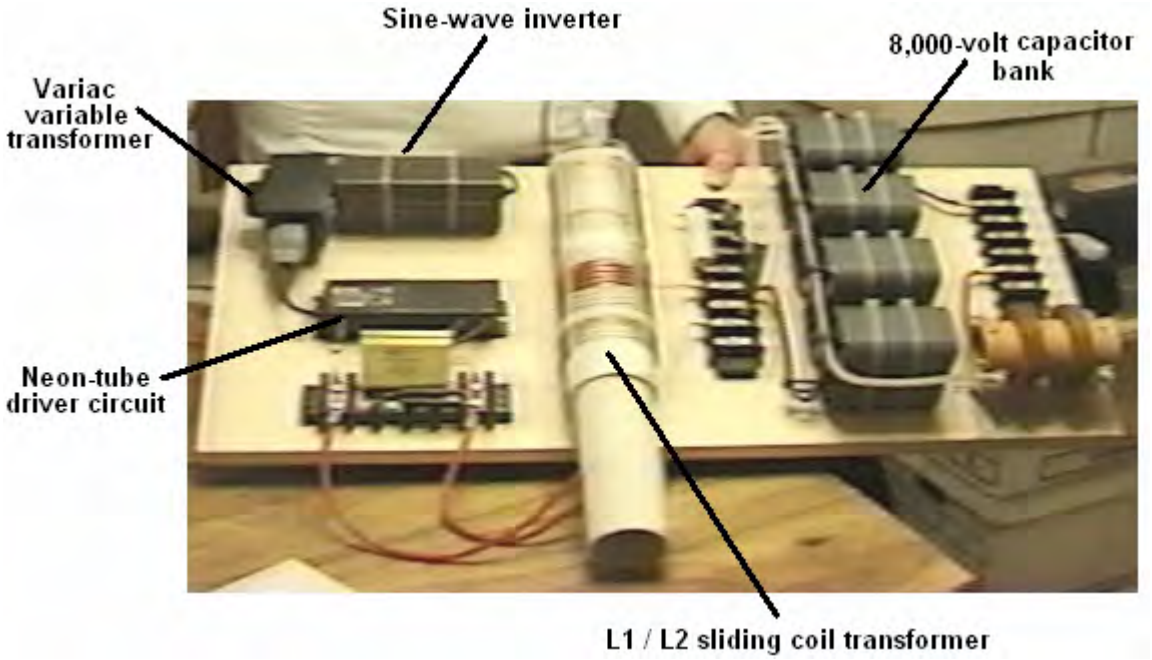
delivery has been delayed due to the batteries stopping operating after three months of use.

Don Smith's High-Power Devices.

Don Smith was a very talented American who has understood all of Tesla's work and has produced literally dozens of practical devices based on his understanding. You will find more specific details in chapter 3, but in broad outline, a twelve-volt battery can be used to generate the pulsing magnetic field needed to nudge the local environment into providing massive amounts of electrical energy. The device described in detail in chapter 3, has an output of around 160 kilowatts, which is far, far more than any individual would need. In other words, it is a device which could easily power your home, and considering that an electric car needs about 65 kilowatts, one could easily power a vehicle, making it into a fuel-less mode of transport. This is not magic, just standard electrical theory being applied correctly for a change.

The key component in many of Don's devices is the humble, commercial power supply used to drive neon sign displays. This module produces some 9,000 volts at a frequency of 35,100 Hz (cycles per second). As Don points out, when you double the pulsing frequency and double the pulsing voltage, the available power goes up by a factor of sixteen times, because the effect of both of these things is squared. You will recall that Bob Boyce is pulsing his toroid very sharply at 42,000 Hz and that high frequency has a major effect on the power produced in his system.

Don then boosts his working voltage further with a step-up transformer called a Tesla Coil. This brings us into an area of massive power. People have the very mistaken idea that a Tesla Coil can only produce voltage and not current. The reality is that if the primary coil is positioned in the centre of the secondary coil, then the voltage and current produced will be roughly the same, and that is a very, very high level of power. One device of Don's looks like this:



This prototype is actually more complicated than it needs to be. It uses three very high-voltage capacitors which are not necessary if you opt for a slightly different method of construction. However, in this version, the twelve-volt battery (which is not shown), powers a true sine-wave inverter in order to provide the mains voltage and frequency needed by the neon-tube driver circuit. The voltage limitations of the capacitors, in particular, the 8,000-volt output storage capacitors, make the 9,000 volt output of the neon-tube driver too much for safe usage. To deal with this, Don uses a Variac-style variable transformer to lower the voltage supplied to the neon-tube driver circuit, and this lets him limit the output voltage to the 8,000 volts of the output storage capacitors.

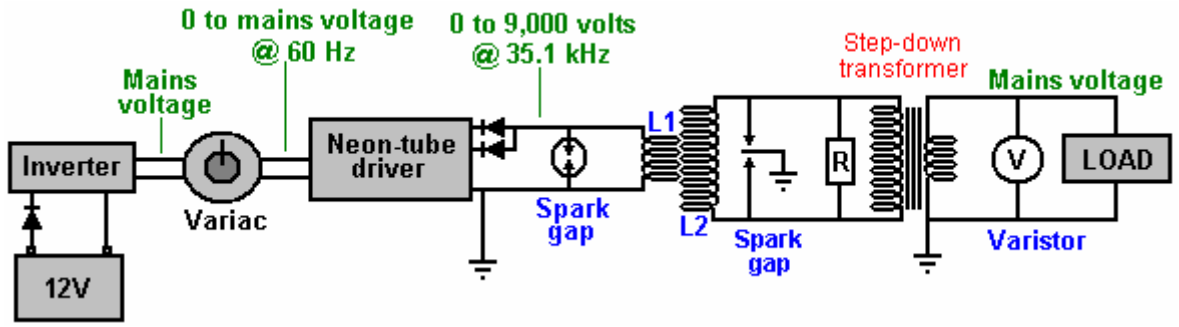
A key detail is that the wire length in the turns of the short primary winding of the Tesla Coil is exactly one quarter of the wire length of the turns in the long secondary winding. This makes the coils resonate which is a vital factor in the operation. The final, exact tuning, can be done by sliding the primary coil to a slightly different position. In this prototype, Don chose to do the final fine tuning by attaching a small capacitor across each of the windings. This is not necessary.

In the prototype shown above, Don then uses four diodes to rectify the output to DC to feed the storage capacitors. This results in an 8,000 volt supply which can provide 20 amps of current. That is an output power of

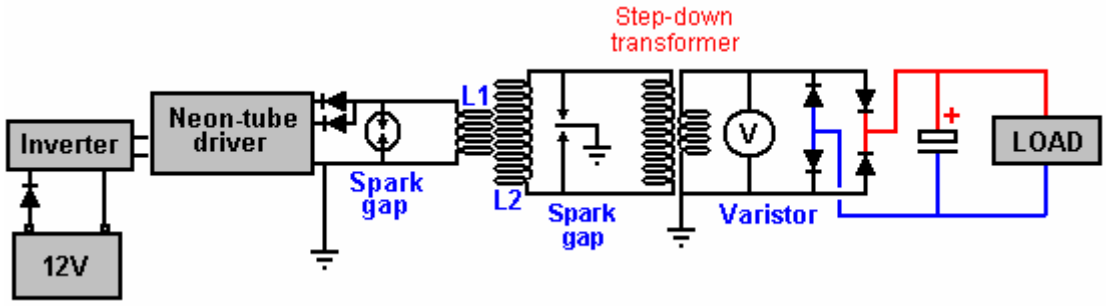
160 kilowatts, and is limited by the output capacitor voltage rating.

Don points out that it is not necessary to do it that way and instead, a step-down transformer can be used to lower the output voltage and boost the available current. If this is done, then the voltage limitations disappear (provided that you are using very high-voltage cables) and so no Variac is needed and no high-voltage capacitors are needed either.

There are two options. Either you can aim for a mains-voltage, mains-frequency, AC output, or you can produce a DC output and use an off-the-shelf inverter to run any mains equipment powered by the device. With the first option, Don connects a single resistor across the primary of the step-down transformer and that drags the frequency down to the wanted level, provided that the resistor has the right value:



The alternative method which aims for a DC output does not need to alter the frequency:



In both of these cases, the twelve volt driving battery can be charged continuously by part of the output power, and there are various ways of doing that. However, care needs to be taken that the battery is not overcharged as the input power is very low.

You will notice the similarity between Bob Boyce's toroid system and Don Smith's Tesla Coil system. In each case, a very carefully wound circular winding is pulsed at high frequency, and in each case, substantial amounts of excess electrical power becomes available, flowing in from the surrounding environment, courtesy of the pulsating magnetic field.

Tariel Kapanadze's Self-Powered Generators.

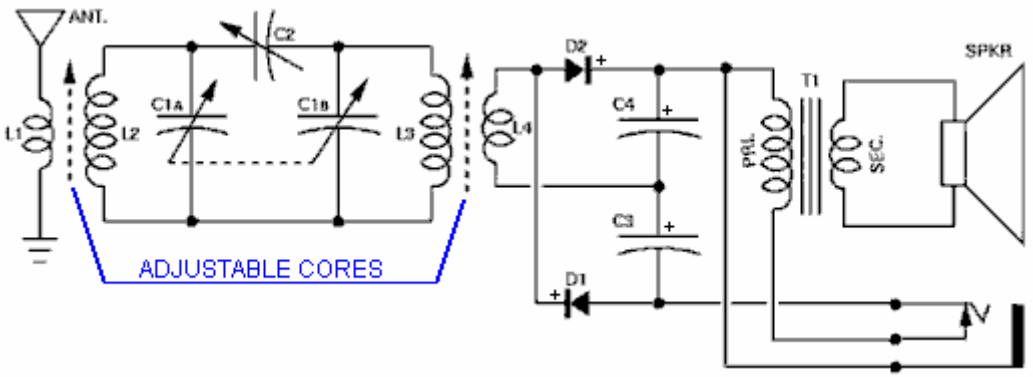
Tariel Kapanadze produced a similar style of device which is self-powered and produces a mains electricity output. He has demonstrated this for a TV documentary. Further details available are in chapter 3.

The Insights of Vladimir Utkin

Vladimir has issued a paper in which he describes some of the very important work done by himself and members of a Russian forum. He has major insights into the work of Tesla, Don Smith and others. With his kind permission, his paper is freely shared here http://www.free-energy-info.tuks.nl/VladimirUtkin.pdf.

Walter Ford's High-powered Crystal Set

In the 1961 edition of the Electronics Experimenter's Handbook, there is an interesting circuit from Walter B. Ford for a high-power crystal set capable of powering a 2.5-inch loudspeaker:



PARTS LIST

C1a/C1b-- 2-gang, 365-uuf. variable capacitor
(Lafayette MS-142 or equivalant)
C2-- 180-uuf. compression-type trimmer capacitor
C3, C4-- .005-uf. fixed capacitor
D1, D2-- 1N34A diode
J1-- Closed-circuit phone jack
L1, L4-- 22 turns of No. 24 cotton-covered wire
(see text)
L2, L3-- Ferrite antenna coil (Miller 6300 or
equivalent)
R1-- 47,000 ohm, 1/2 watt resistor
(*optional, see text)
T1-- Replacement-type output transformer; 3000-
to 10,000-ohm primary; 4-ohm secondary
Spkr.-- 2 1/2" speaker, 4-ohm voice coil (Lafayette
SK-65 or equivalent)
Misc.-- Hardware, wood, sheet aluminum, Fahnestock
clips, etc.

He says: Here's a pint-sized crystal radio with enough power to drive a 2·5" speaker. This little unit's selectivity is far better than you would expect to find in a crystal receiver and volume is equal to that obtained with sets using a transistor. No external power source is required.

The unusual selectivity of this radio is due to its special double-tuned circuit. A pair of diodes connected as a voltage-doubler provides the extra power to operate the small speaker. An output jack is provided for headphone listening and for connecting the set to an amplifier.

Construction: The model was built on a 2.5" x 4.5" wooden chassis with a 3.5" x 4.5" metal front panel. However, size is not critical, and other materials can be substituted if desired.

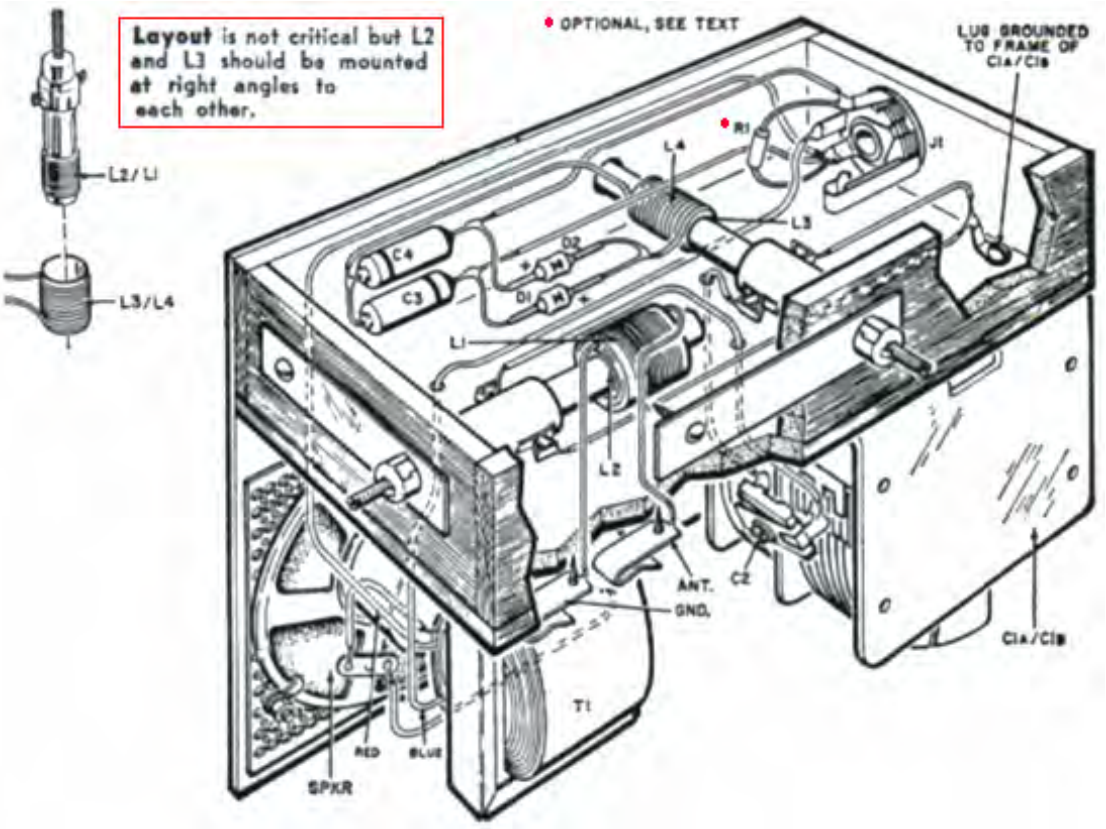
Two standard ferrite loopsticks, L2 and L3, are used. Both must be modified by the addition of a second winding. L1 and L4, respectively. Each of the added windings consists of 22 turns of No. 24 cotton-covered wire wound on a small cardboard tube as shown in the picture. (Actually, any wire size from No. 22 to No. 28 with cotton or enamel insulation will do the job).

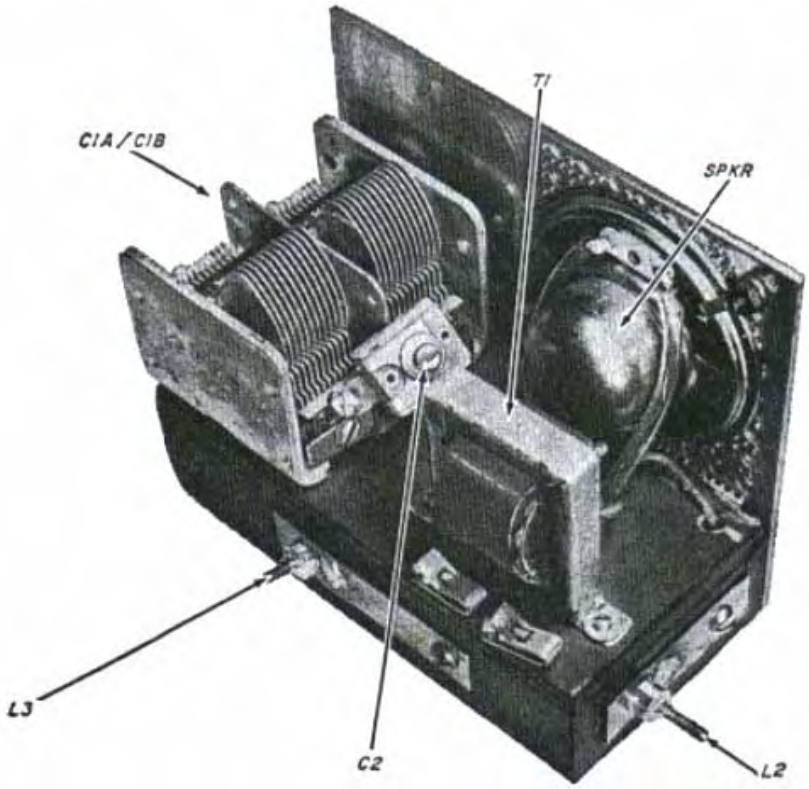
The diameter of the cardboard tube should be slightly larger than L2 and L3 so that L1 and L4 will slip overL2 and L3 easily.

Resistor R1 is used only for feeding the set into an amplifier; it should be omitted for both earphone and loudspeaker operation. Trimmer capacitor C2 should be soldered across the stator terminals of two-gang variable capacitor C1a/C1b as shown. The speaker and output transformer can be mounted wherever convenient.

If a metal chassis is used, then be sure to insulate the aerial and earth connection sockets from the chassis. When all of the parts have been mounted on the chassis, wire them together following the schematic and pictorial diagrams. Be sure that diodes D1 and D2 and capacitors C3 and C4 are connected correctly, paying attention to their polarity.

While this is interesting, what appears to be a key factor is contained in the drawings, where he states that the only important thing is that it is essential for the two sets of coils to be mounted perpendicular to each other:





Alignment and Operation. To align the receiver, connect it to an antenna and ground. The optimum length of the antenna varies with location, but 50 feet will usually be suitable in areas which receive several broadcast stations. Next, plug a high-impedance earphone into jack J1. Tune a station near the high frequency end of the broadcast band – say, 1500 kHz – and adjust the trimmer capacitors on variable capacitor C1a/C1b to get the loudest signal.

Trimmer capacitor C2 should then be adjusted for the best selectivity and volume over the entire broadcast band. Finally, Coils L1 and L4 can be moved to their optimum positions by sliding them backwards and forwards over coils L2 and L3. If a nearby station interferes with the reception of a weaker station, tune the slug of L2 to get minimum interference. For loudspeaker operation, simply unplug the earphone. Strong local stations should be received at fair volume.

How It Works: The receiver employs a double-tuned circuit feeding a crystal diode voltage-doubler/detector which drives a small loudspeaker. Radio frequency signals picked up by the antenna system are induced into coil L2 from coil L1. The desired signal is selected by the tuned circuit C1a/L2 and coupled through capacitor C2 to a second tuned circuit C1b/L3, which improves the selectivity by narrowing the radio frequency bandpass. The twice-tuned signal is then induced into coil L4 from coil L3.

The positive half of the radio frequency signal appearing across L4, passes through the 1N34A germanium diode D2 to charge capacitor C4. The negative half of the signal passes through diode D1 to charge capacitor C3. The polarity of the charges on C3 and C4 are such that the effective voltage is doubled. This voltage appears across the primary of output transformer T1 which converts the high-impedance signal to a low-impedance output suited to the loudspeaker.

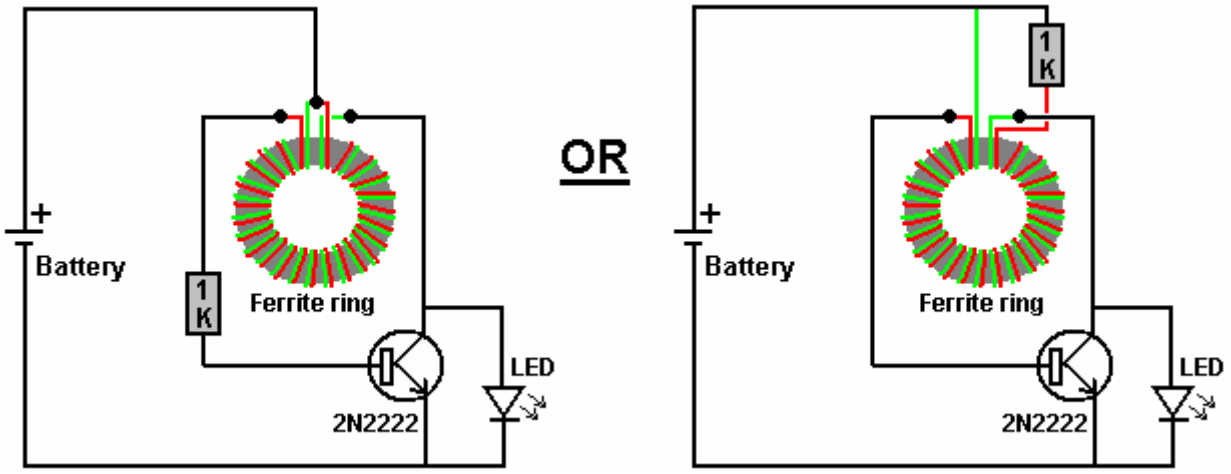
While this looks like a very good design for a crystal set, the fact that it is insisted that the coil pairs **must** be mounted at right angles to each other raises an interesting parallel with the above work of Vladimir Utkin where he states that if the high-frequency excitation field is at right angles to the output coil, then there will be a free-energy inflow into the circuit from the local environment. Perhaps this crystal set design gains extra power to drive it's loudspeaker from an inflow of environmental energy.

Lawrence Tseung's Self-Powered "FLEET" Generator.

The "FLEET" ("Forever Lead-out Existing Energy Transformer") device is a self-powered electrical generator which has no moving parts and which can be constructed cheaply. It has been developed by a Hong Kong based team of people: Mr Lawrence Tseung, Dr. Raymond Ting, Miss Forever Yuen, Mr Miller Tong and Mr Chung Yi Ching. It is the result of some years of thought, research and testing and it has now reached an advanced stage of testing and demonstration and is nearly ready for commercial production.

Mt Tseung has applied his "Lead-out" theory to the category of low-power circuits known as the "Joule Thief" circuits. These circuits originated with an article by Mr Z. Kaparnik, in the "Ingenuity Unlimited" section of the November 1999 edition of the "Everyday Practical Electronics" magazine.

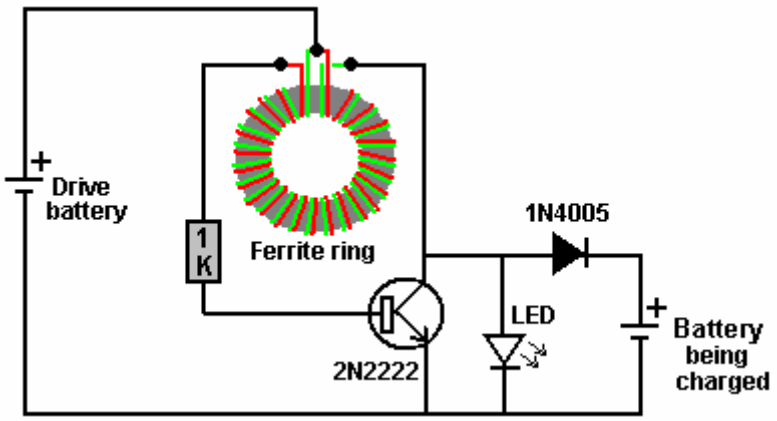
The initial circuit allowed the very last energy to be drawn from any ordinary dry-cell battery, and used to light a white Light-Emitting Diode ("LED") for use as a small torch. It allows a battery which is considered to be fully discharged, to drive the circuit until the battery voltage drops right down to 0.35 volts. The initial circuit uses a bi-filar coil wound on a ferrite ring or "toroid". Bi-filar means that the coil is wound with two separate strands of wire side by side, so that each adjacent turn is part of the other coil. A coil of that type has unusual magnetic properties. The Joule Thief circuit is like this:



It is important to notice how the coil is wound and how it is connected. It is called a "toroid" because it is wound on a ring. The ring is made of ferrite because that material can operate at high frequencies and the circuit switches On and Off about 50,000 times per second ("50 kHz"). Notice that while the wires are wound side by side, the start of the red wire is connected to the end of the green wire. It is that connection which makes it a "bi-filar" coil instead of just a two-strand coil.

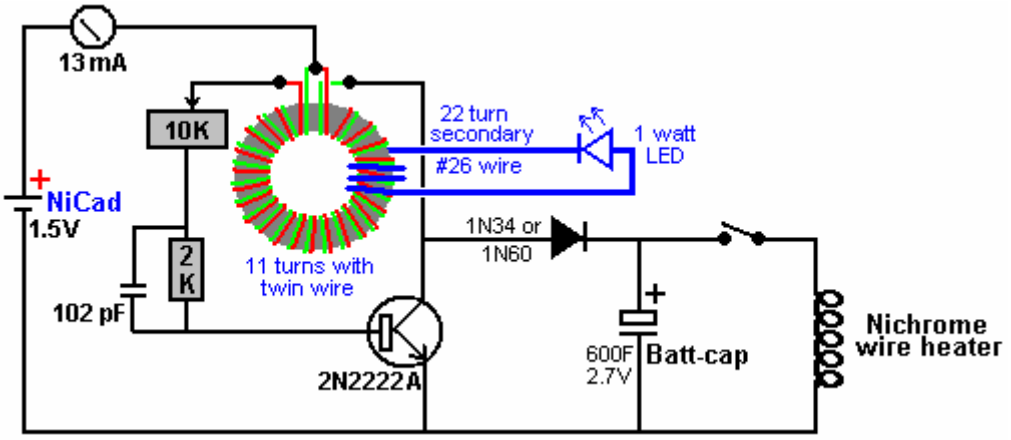
This "Joule Thief" circuit was then adapted by Bill Sherman and used to charge a second battery as well as lighting the Light-Emitting Diode. This was achieved by adding just one more component - a diode. The diode used was a 1N4005 type because that was to hand at the time, but Bill suggests that the circuit would work better with a very fast-acting Schottky-type diode, perhaps a 1N5819G type.

The circuit produced by Bill is:



When driven by a 1.5 single cell battery, this circuit produces about 50 volts with no load and can supply 9.3 milliamps of current when the output is short-circuited. This means that you could charge a 6-volt battery using a 1.5 volt battery.

“Gadgetmall” of the www.overunity.com Joule Thief forum has taken the circuit further and found a very interesting situation. He has modified the circuit and used a “batt-cap” which is a very high capacity, very low-loss capacitor. This is his circuit:



He has added an additional winding to his one-inch (25 mm) diameter ferrite toroid, and he uses that to power a 1 watt LED. Why he has done this is not immediately clear to me, except possibly, that it shows when the circuit is operating. He runs the circuit driven by a small rechargeable battery, which feeds 13 milliamps into the circuit, for a period of fourteen hours. At the end of that time, the batt-cap has gathered enough energy to fully recharge the driving battery in a minute or two, and then power a heater winding of nichrome wire (as used in mains-powered radiant heaters) for four and a half minutes. Alternatively, that amount of extra power could boil a kettle of water. The really interesting thing about this is that the driving battery gets recharged every time and so the circuit is self-sustaining although it is not a powerful circuit.

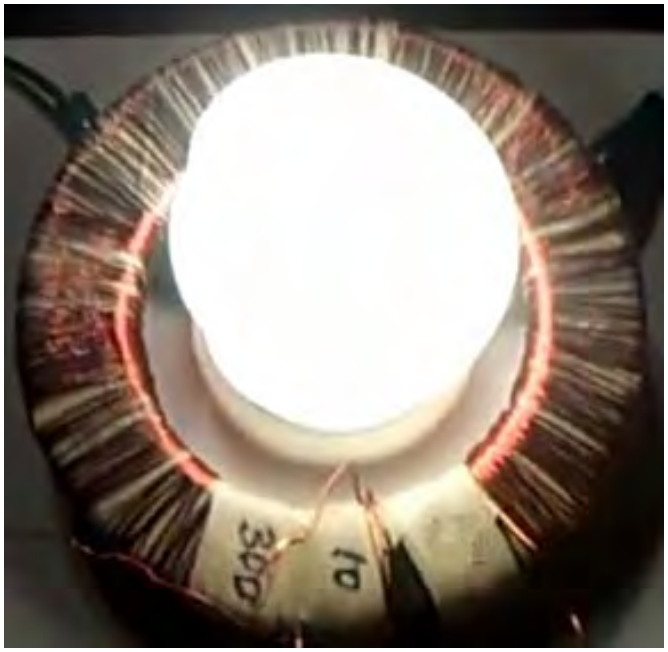
However, Jeanna has developed the circuit significantly as she shows in her series of videos:

http://www.youtube.com/watch?v=Y4IMgDRGpHE
http://www.youtube.com/watch?v=1tVlCJiuWH4
http://www.youtube.com/watch?v=y6pbzrhBR-8
http://www.youtube.com/watch?v=tNoZrV3w4f8
http://www.youtube.com/watch?v=XzhbsLBwc54
http://www.youtube.com/watch?v=4gj7IFCXw9Q

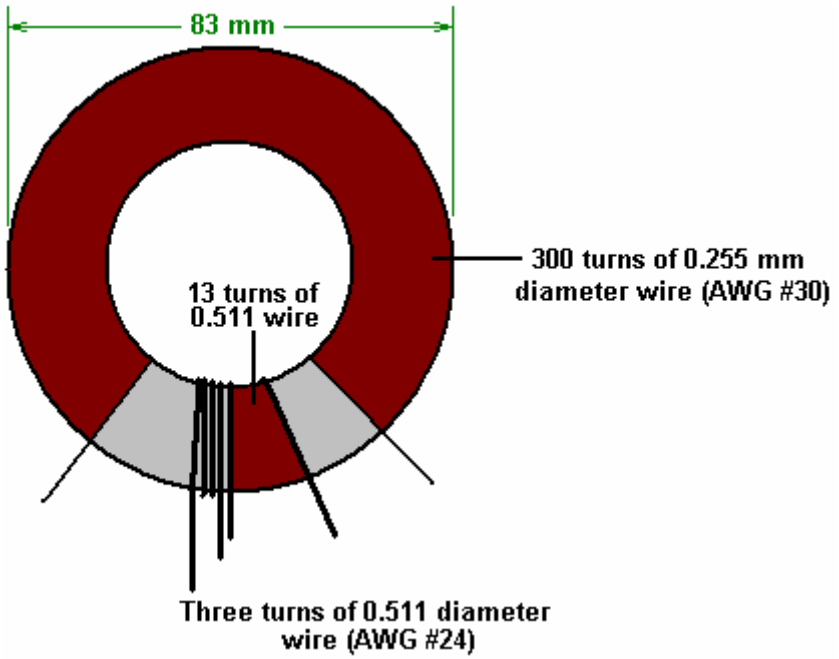
Her main point is that using the collector of the transistor as the power take-off point of the circuit, is inefficient as that draws a lot of input current without a corresponding increase in output current. She adds a 74-turn secondary winding on top of her two 11-turn Joule Thief bi-filar windings, and that appears to give a far better power output. She uses the very small AAA size 1.2V battery and further drops the output (because “the light is too blinding”) by putting a resistor in series with the battery and using many LEDs in series. She has recorded the following results:

With no resistor, the output voltage is 58V peaks at 62.5 kHz (open circuit output, with no load at all)
With a 10 ohm resistor, the output voltage is 49V peaks at 68 kHz.
With a 33 ohm resistor, the output voltage is 25V at 125 kHz.

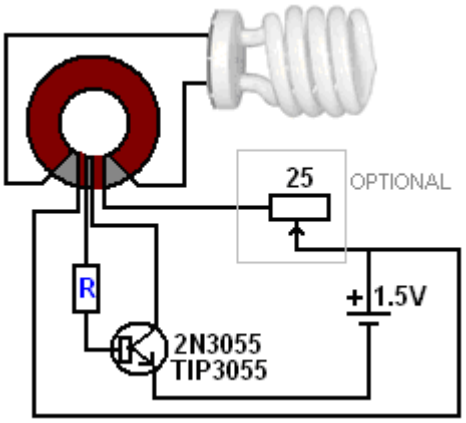
‘LidMotor’ states that Jeanna also produced a Joule Thief circuit which could light a 15-watt straight fluorescent tube for about five hours when being driven by a single AA battery. He states that he was not satisfied with that level of lighting and at https://www.youtube.com/watch?v=KAakZTR_4LE he shows a version (which he thinks is Jeanna’s design and which Jeanna thinks is his design) driving a 10-watt Compact Fluorescent Light which has had the ballast circuitry removed. The build uses an expensive 3.25 inch (83 mm) outer diameter ferrite toroid, and the lighting from a single AA battery looks like this:



The ferrite ring is wound like this:



The main winding is 300 turns of AWG #30 enamelled copper wire with a diameter of 0.255 mm. Please notice the gap between the ends of that winding. That gap is important as high voltage is developed between the two ends of the winding and if the winding were continued all the way round the toroid, then the insulating enamel coating the wire would be liable to burn out due to the very high voltage difference between the first and last turns, causing a short-circuit. The two other windings are with AWG #24 enamelled copper wire which has a diameter of 0.511 mm and those two windings are positioned closely side by side in the middle of the gap between the ends of the 300-turn winding. The circuit is like this:



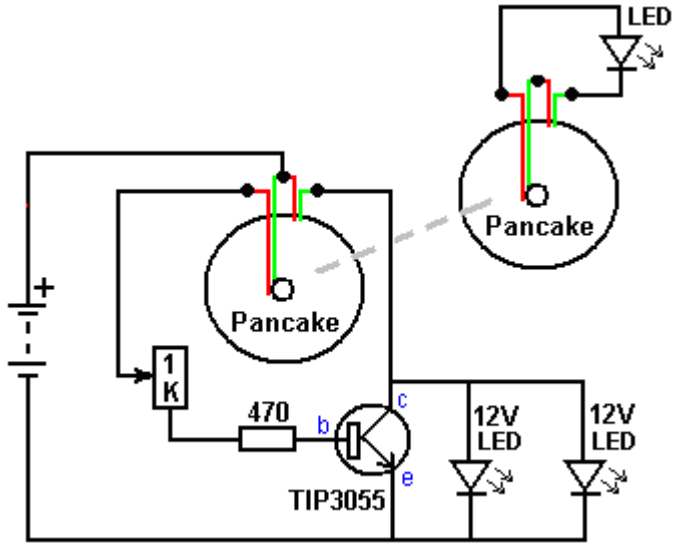
The optional 25 ohm wire-wound variable resistor wastes power but creates a voltage drop across it, reducing the voltage reaching the circuit and so, dimming the light progressively, all the way down to zero. The base resistor "**R**" has been set at 22 ohms by 'Lidmotor' who says that it really should be 100 ohms but he has lowered it to get brighter lighting. Please note which side of the 3-turn and 13-turn windings are connected in the circuit as the direction of winds is very important for those two windings.

It is not uncommon for people to comment on the faint whistling sound make by a Joule Thief circuit (especially a low-voltage version like this). It is my experience that the sound is caused by the transistor resonating with the frequency of oscillation of the circuit, the TIP3055 being particularly prone to this. I suggest therefore, that bolting on a heat sink (which is most definitely not needed to dissipate heat produced by this circuit) will alter the resonant frequency of the transistor/heat sink combination and so stop the whistling.

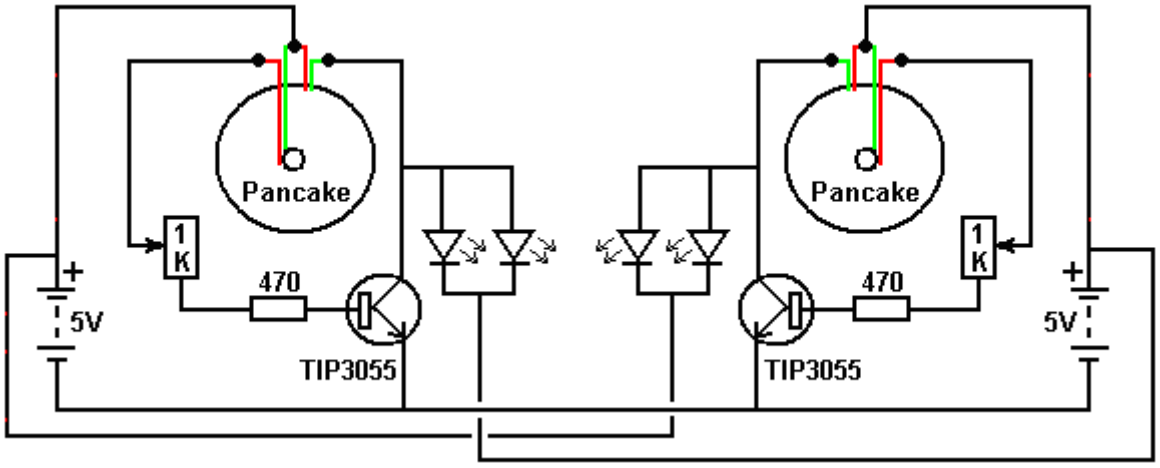
Personally, I have always had the greatest difficulty in getting any satisfactory light out of a Compact Fluorescent Lamp when driven by a Joule Thief circuit and so, for me, the best light source is one of the "G4" LED arrays with a "5050" driver chip. These look like this:

Quite good lighting can be had from a single Joule Thief circuit driving up to twelve of these simultaneously. The larger lighting area produces a more even and more gentle light which is quite effective in total darkness.

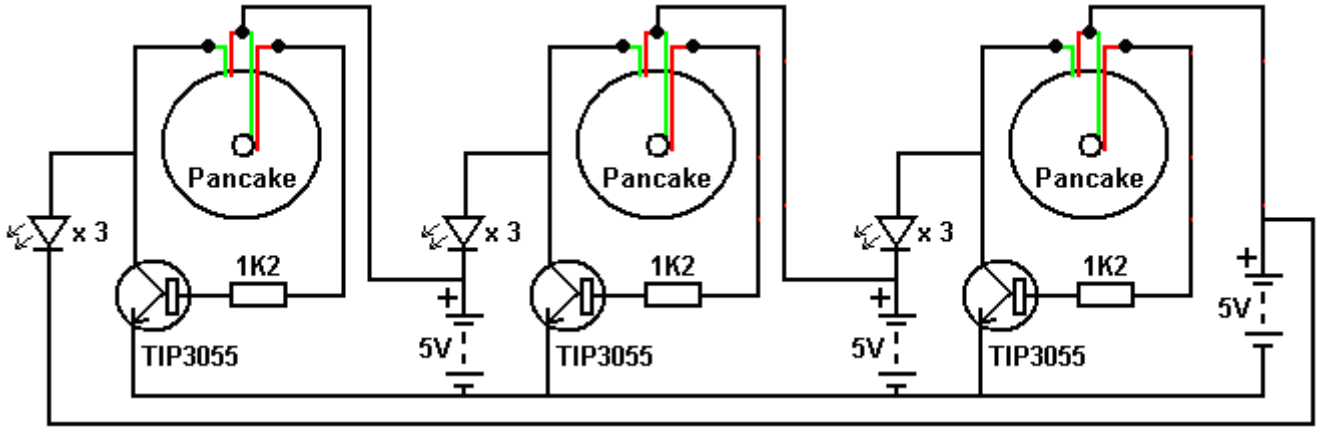
Most Joule Thief circuits specify a ferrite toroid, but a 75 mm diameter pancake coil also works well, and interestingly, pressing a second pancake coil tightly against the Joule Thief pancake coil, allows an additional LED array to be powered without increasing the current draw of the Joule Thief circuit:



It is also possible to cross-connect two or more Joule Thief circuits so that as well as producing 12V LED array lighting, each charges the battery being used by the other circuit:



And this works well with three circuits cascaded:



When working with these circuits, I bought a light-meter to take the guesswork out of assessing light levels as the human eye is very bad at doing that, although the visual impression of the lighting produced by any arrangement is actually more important than the measured lighting level. For example, there are 1-watt tiny LEDs, which have a theoretically high light output, but due to their tiny lit area are pretty useless for home lighting.

When I used the light meter (which measures in lux) I got a considerable shock. I used it with a light box to measure the light produced by two G4 LED arrays side by side, first with straight battery input and then with a Joule Thief input. The major shock was that the G4 LED arrays are actually more efficient at converting electric

current to light than when a Joule Thief is used to drive those same LED arrays. That was completely unexpected. The figures for voltage / current draw / light produced using 1.2V (nominal) NiMh batteries were:

9 batteries 11.7V 206 mA 1133 lux: 2.41 watts 470 lux per watt (the manufacturer's intended performance)

8 batteries 10.4V 124 mA 725 lux 1.29 watts 562 lux per watt

7 batteries 9.1V 66 mA 419 lux 0.60 watts **697** lux per watt (a very realistic performance level)

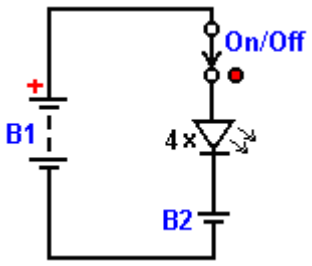
6 batteries 7.8V 6 mA 43 lux 0.0468 watts 918 lux per watt

A Joule Thief circuit driving two G4 LED arrays powered by 4 batteries giving 5.2V with the current draw controlled by selection of the transistor's base resistor:

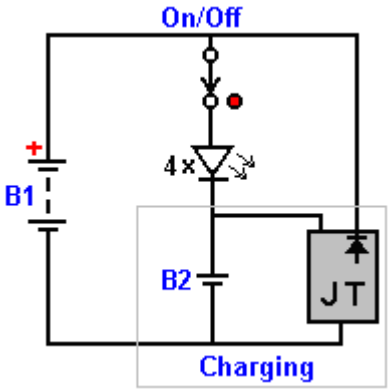
358 mA 259 lux 1.86 watts 139 lux per watt
200 mA 212 lux 1.04 watts, 204 lux per watt
180 mA 200 lux 0.936 watts 101 lux per watt
158 mA 182 lux 0.822 watts 221 lux per watt

This was quite startling, and the surprising conclusion is that using just four LED arrays driven by a 9V battery supply, produces a very respectable 800 lux for just 135 milliamps which is about 1.2 watts total – a very unexpected result. When the four LED arrays are fitted into a desk lamp and covered with frosted plastic to give a diffused light, the result is an excellent level of desk lighting which also lights the rest of the room quite well.

If you choose to do that, then it is possible to collect all of the current passing through the Joule Thief circuit, like this:

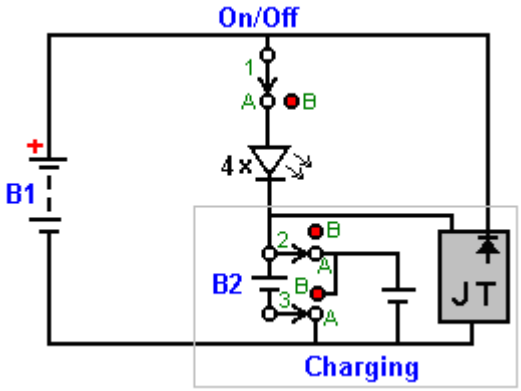


Here, the drive battery "B1" is made up with one battery more than the wanted voltage and all of the current passing through the four LED arrays is used as charging current for one additional battery which can be used to power a Joule Thief charging circuit:

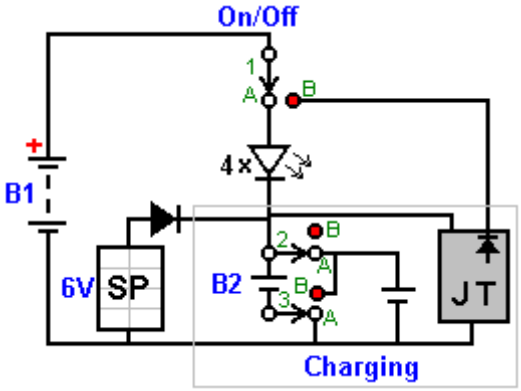


As the LED array lights are essentially diodes anyway, they have no problem with being supplied with current from a battery which is being pulse-charged by a Joule Thief, so there is the option of leaving the Joule Thief circuit connected all the time as shown above. That, of course, is optional.

One further thing which can be done is to allow for doubling the Joule Thief voltage when the light is off. While the Joule Thief charging circuit is perfectly capable of charging the "B1" battery when driven by a 1.2V battery, it can charge faster if its voltage is doubled, which can be done quite easily using a standard 3-pole, 4-way rotary switch:



or



Here, the “B2” battery is made using two 1.2V batteries connected in parallel when the light is on, and when the light is switched off, the two batteries are connected in series, driving the Joule Thief charging circuit with double the voltage. Another option is to connect in a 6V or higher solar panel to charge the “B2” battery during the day. It appears that although the charging circuit can run at all times, it is actually more effective if the battery charge is stored and the pulse-charging only starts when the light is switched off.

It needs to be remembered that NiMh batteries are only 66% efficient, which means that when they are driving a load, you will only ever get back again, two thirds of the current fed into them. Where the above circuits show just one Joule Thief charging circuit, there will normally be two, three or more charging circuits to raise the rate of charging during daylight hours. Also, if the battery powering the Joule Thief has a higher voltage than the battery being charged by it, extra diodes will be needed so that their combined voltage drop will prevent the charging battery feeding a large current directly into the battery being charged.

Mr Lawrence Tseung has taken the Joule Thief circuit and modified it to become a circuit with a very serious output, moving it into a completely different category.

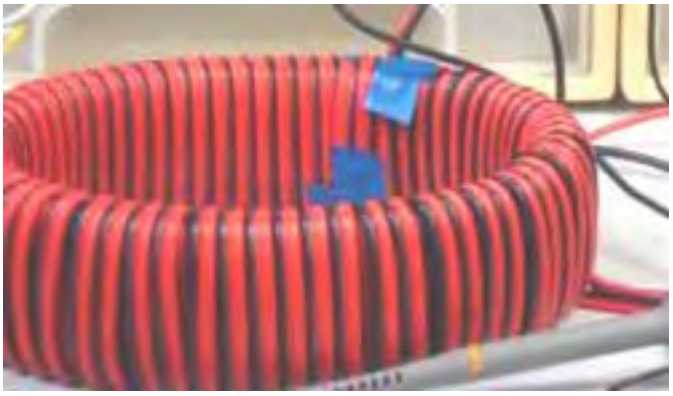
As a first step towards what the team calls their "Fleet" device, the toroid has been enlarged to a much greater diameter. The coil is now wound on a section of plastic pipe, 170 mm (6.5 inches) in diameter and 45 mm (1.75 inch) deep:

This section of pipe is "bi-filar" wound with two wires side by side as already described for the Joule Thief construction. As before, the start of one wire is connected to the end of the other wire. Then, the winding is given a layer of electrical tape to hold it in place and to provide an easy working surface for a second winding.

The wire used for the winding is the widely available red and black pair of wires, sometimes called "figure of eight" because the cut end of the wires looks like the numeral 8. The wire should be able to carry 2.5 amps. It must be side-by-side wire and **not** one of the twisted varieties. It looks like this:

The second winding is made in the same way but the connections are slightly different. As before, the end of the first wire is connected to the start of the second wire, but that connection is then insulated and not used in the following circuitry. This just connects the two windings one after the other, known technically as being connected

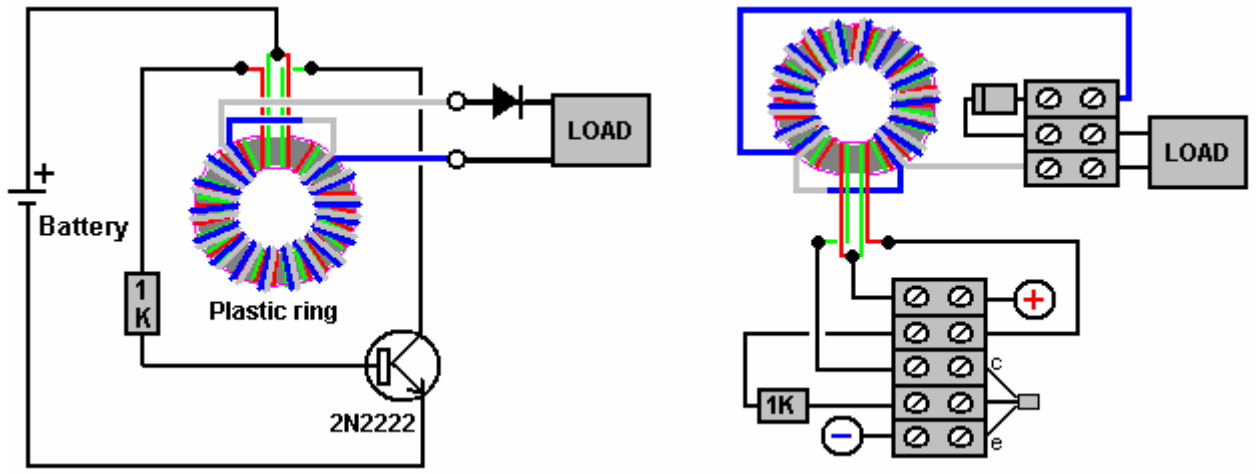
"in series" and is the equivalent of making the winding with just a single strand of wire. The completed coil may look like this:



This particular design is still in it's early stages and so many different coils sizes and constructions are being tested:

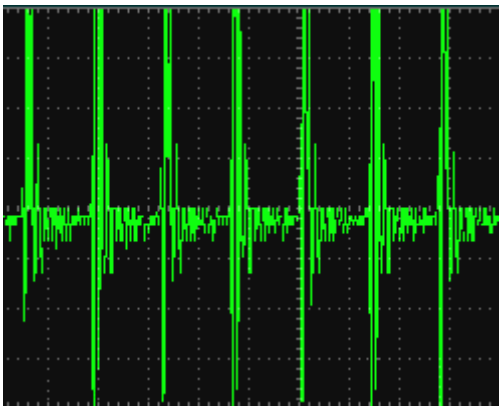
The arrangement is for the inner winding of the toroid to be oscillated by the Joule Thief circuit already described. This causes a pulsating magnetic field to envelope the outer winding of the toroid, producing an electrical output which is capable of doing useful work. The really important thing about this arrangement, is the fact that the amount of power coming out of the circuit is very much greater than the amount of power needed to make the circuit operate. The additional power is led out of the local environment and drawn into the circuit, becoming available to do useful work.

The overall circuit then looks like this:



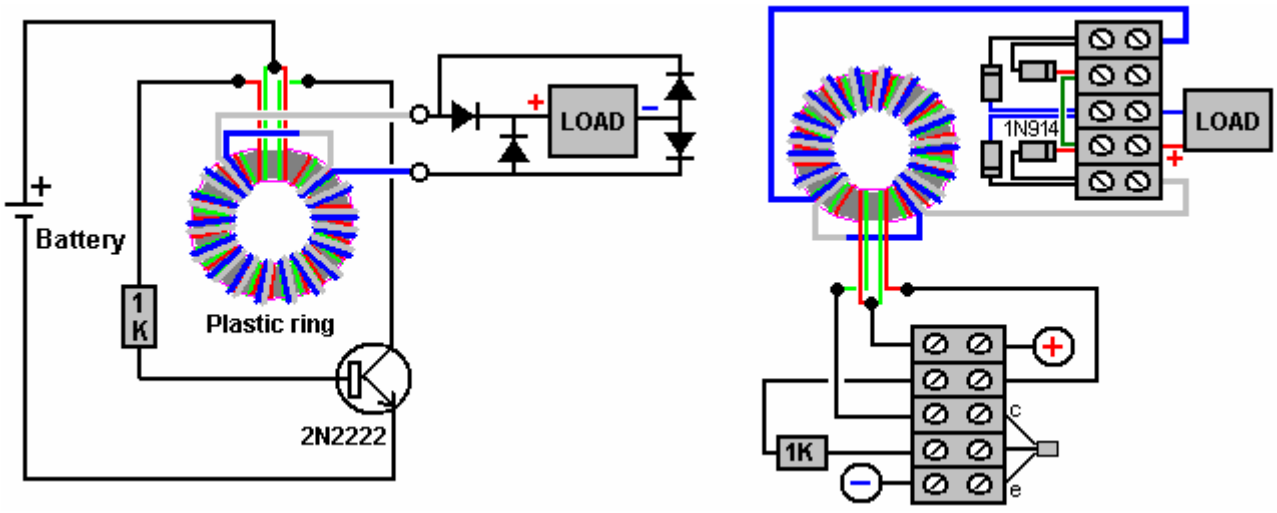
While the outer winding is shown here with thicker wire of a different colour, this is only to make the arrangement easier to understand. In reality, the outer winding is with exactly the same wire as the inner winding, and it will normally go all the way around the toroid. The total amount of wire needed to make the windings is about 70 metres and so it is normal to buy a full 100 metre reel of the twin-core wire, which allows both windings to be made and leaves spare wire for other things.

For those of you who are **very** technically minded, the output waveform looks like this:



and the voltage pulses in this output are occurring about 290,000 times per second.

What has worked better for me is using a bridge of four diodes rather than a single diode:



I have used this circuit, driven by a 1.5 volt battery, to charge 12-volt batteries, but the best results are in the five to six volt range. I have used this circuit to confirm COP>1 by charging one small 12V lead-acid battery with an identical battery, swapping the batteries over and repeating the process several times. The result was that both batteries gained genuine, usable power. I suspect that the effect would have been much greater if I had charged two or more batteries in parallel. The toroid was an 8-inch diameter, 10 mm by 12 mm offcut from a plastic pipe

which happened to be to hand and the wire used was plastic covered 6-amp equipment wire, again, because it was to hand at the time. Winding the toroid and setting up the circuit was done in a single evening.

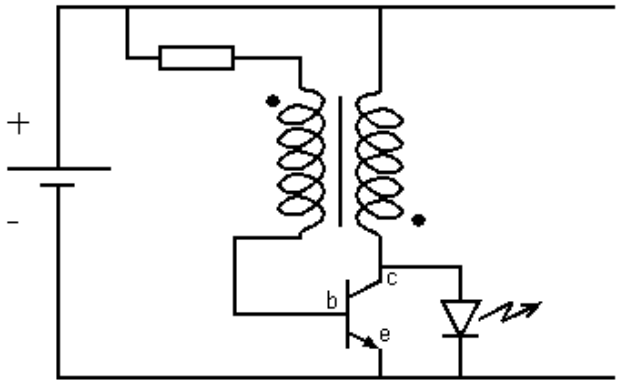
Overall, this is a very simple, cheap and easily constructed COP>10 device which has the potential of providing large amounts of free, useable, electrical power. With further development, it may well be possible to produce a version which could deliver the power needed by a whole household. It is also likely that these devices will become available for purchase a quite a low cost. All in all, this is a very important device and full credit must go to the development team who have carried the research to this point and who are continuing to refine the design to produce more and more power.

Advanced versions of the Joule Thief Circuit.

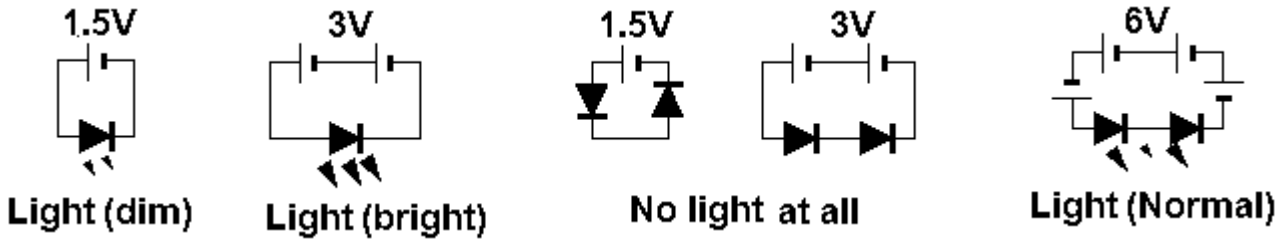
A contributor who prefers to go by his ID of “Ace_Propulsion” shows here a series of clever, innovative variations on the well-known Joule Thief circuit.

What is a Joule thief? A Joule Thief circuit is a minimalist self-oscillating voltage boosting circuit which is small, low-cost and easy-to-build. It is normally used for driving light loads. It can use nearly all of the energy in a single-cell battery, even those which are already far below the voltage level where other circuits consider the battery to be ‘fully discharged’ (or "dead").

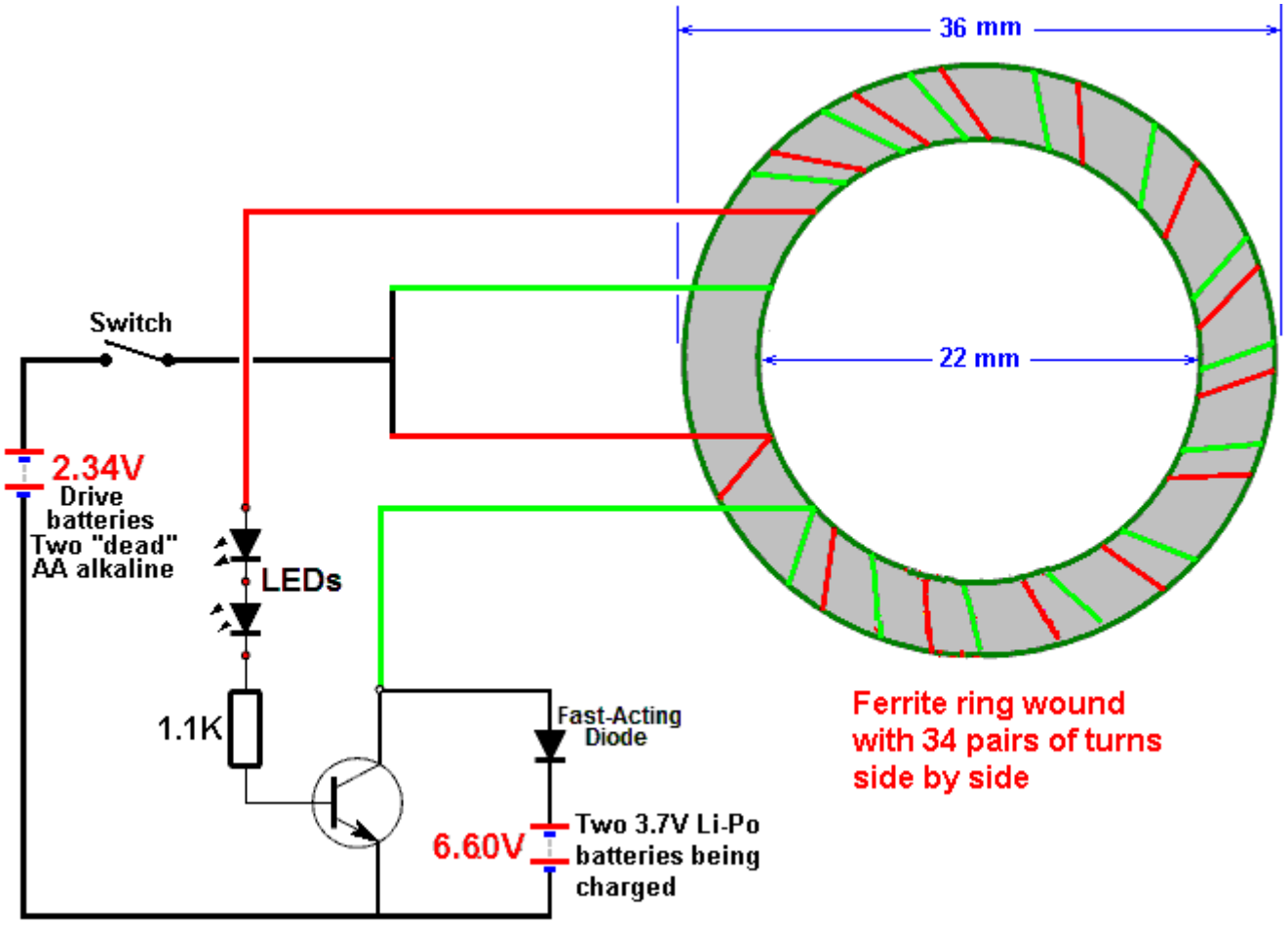
Notice the description "voltage booster". This means that the output voltage is increased at the expense of a higher input current draw. Conventional science says that a Joule Thief circuit can never ever achieve COP>1. A conventional Joule Thief circuit as shown below, will always have an energy loss between the Collector and Emitter of the transistor.



With proper modifications this circuit can acquire energy from the environment to power the output. This is quite simple to achieve. First, before we start discussing circuits, I'm going to tell you the weird stuff about LEDs:



Notice that you can illuminate an LED with just 1.5 Volts and get a brighter light by using 3 Volts. But if you connect two LEDs in series then that 3 Volts is too low to light them and so there is no light at all and zero current draw. Well, the weird stuff is that you can power ONE LED with 1.5V and can't power TWO LEDs in series with 3volts?!? And, while resistors control the brightness of the light they do not change the voltage required in any way. Now that's it! I used this on a Joule Thief and when I did, I got COP>1 using this circuit:



This circuit has an input current of 12.5 milliamps at 2.35 volts (which is 30 milliwatts) and an output current of 8 milliamps at 6.60 volts (which is 52.8 milliwatts) and that is COP=1.8 or 80% greater output power than the input power. The ferrite toroid is wound using 0.4 mm diameter enamelled copper wire (#26 AWG) and while the trrns are shown slanted in the diagram, the actual direction of the turns is radial, and anyway, the direction of the turns has no effect on circuit performance. It is expected that the diameter of the ferrite ring is not at all critical but only one diameter was available for testing. The fast-acting diode could be an FP607, UF5408 or similar, and it is also possible to use a transistor with the Base and Emitter connected together, instead of a high-speed diode. The LEDs used are 8 mm diameter types.

In this circuit it was found that the input voltage is important. The best input voltage is between 2.2 volts and 2.5 volts, so two Ni-Cad or two Ni-Mh batteries are about the optimum input as higher voltages just cause higher input currents without any improvement in the output power.

The key to free energy from this circuit is to use at least two LEDs connected in series. Place them in the current flow to the base of the transistor and the current fluctuations caused by their ‘weird stuff’ will boost the energy coming from the output.

one very important point is that there must be at least two LEDs and the circuit must **not** self-start, because if it does, then the input voltage is too high and the circuit will operate at COP<1. Because of this, you need to start the circuit operating by hand, and another very important point is that the output voltage should at least twice as much as the input voltage.

Features of this technique:

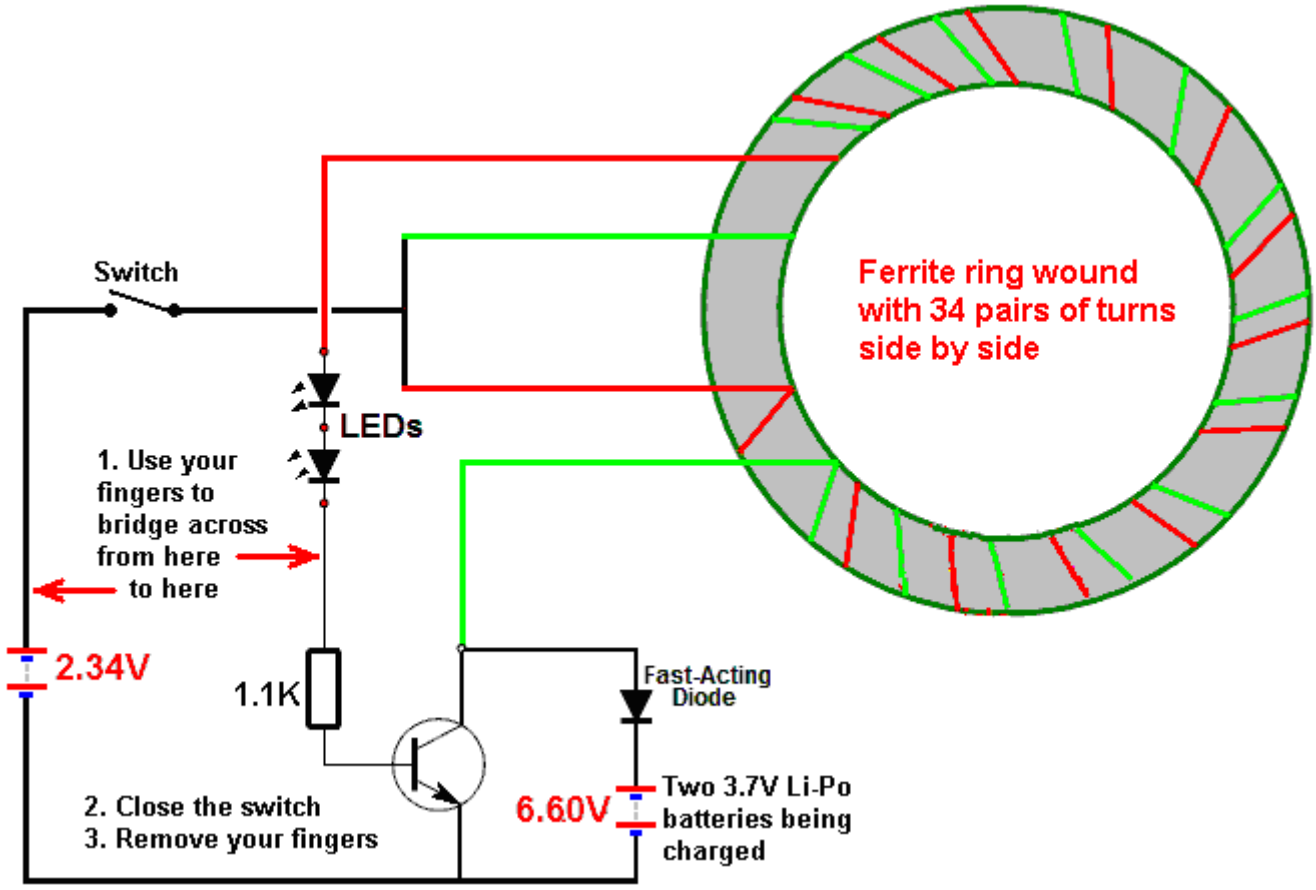
- COP>1 can be achieved and then with proper modification the circuit can become self-powered.
- You can extract energy from “dead” batteries and supplement it with additional energy drawn from the environment.
- Go to work and leave the circuit charging batteries at home while you are working.
- Interestingly, high-frequency sound is emitted by the circuit, and that sound tends to keep mosquitoes away!

Fine-tuning the circuit:
There are 5 parameters involved in the operation of this circuit:

1. Input voltage,
2. Output voltage, The coil winding,
3. The diameter of the toroid,
4. The number of LEDs, and
5. The resistor which feeds current to the base of the transistor.

The first step in assembling the circuit is to check the LEDs which you are going to use. These LEDs are going to be connected in series, so start with two and connect them across the battery. If the LEDs light up, then add an extra LED until the chain of LEDs do not light when connected across the battery. Doing this, will raise the COP of the circuit over 1 as the output power will then exceed the input power.

When setting up the toroid assembly, remember that winding more turns on the toroid will make a coil which has higher impedance which will increase the COP value, but too many turns may result in lower current which means a slower rate of charging from the output. The output voltage should **always** be more than twice the input voltage (for example: Input 2.35V Output 6.60V).



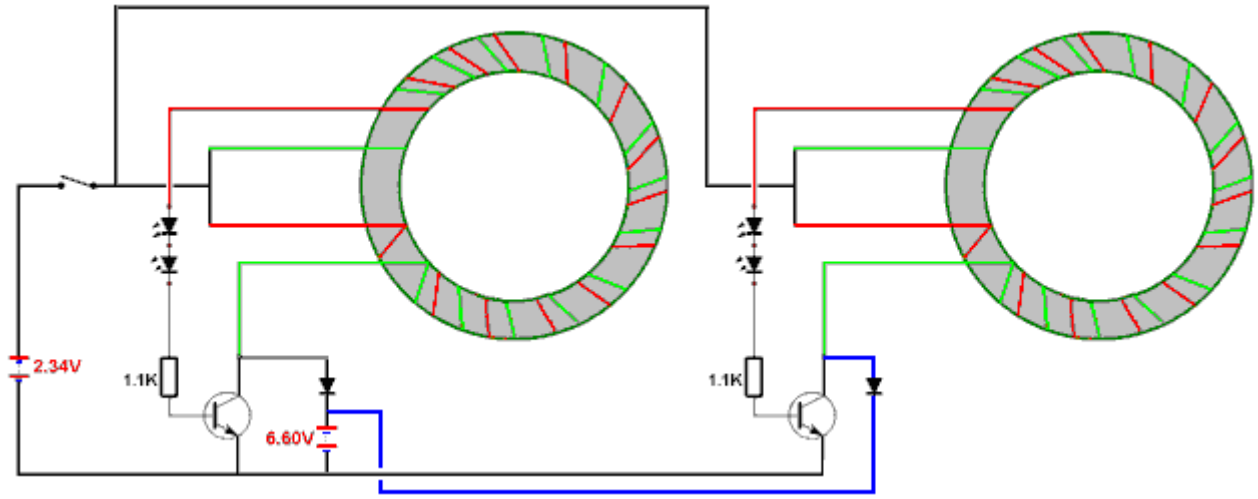
When the circuit has been built as shown above, if it starts running when the switch is closed, then the input voltage is too high, so keep adding an additional LED until the circuit will no longer self-start. Then, use your fingers to start it running by using your body resistance to bypass the LED chain very briefly to get the circuit oscillating. This is a low voltage section of the circuit and so there is no danger at all of getting a shock by doing this. An alternative would be to put a resistor in place instead of your fingers and use a press-to-make switch to trigger the circuit.

A further refinement is to add even more LEDs until the point is reached where the circuit will not start, even if you use your fingers to try to get it operating. When that point is reached, remove one of the LEDs and get the circuit running. Compare the input and output power levels, and then remove one more LED and repeat those power measurements. Keep doing that while you still have more than two LEDs until you determine what is the most effective number of LEDs in your circuit, that is, you find the best COP performance that your circuit can achieve.

In this circuit the LEDs are there to cause base current fluctuations as a mechanism to get COP>1 performance, and so, they are not there for lighting. You may increase the resistor value and so, lower the amount of input current used, but doing that results in lower output power. In my circuit, I used a 1100 ohm resistor.

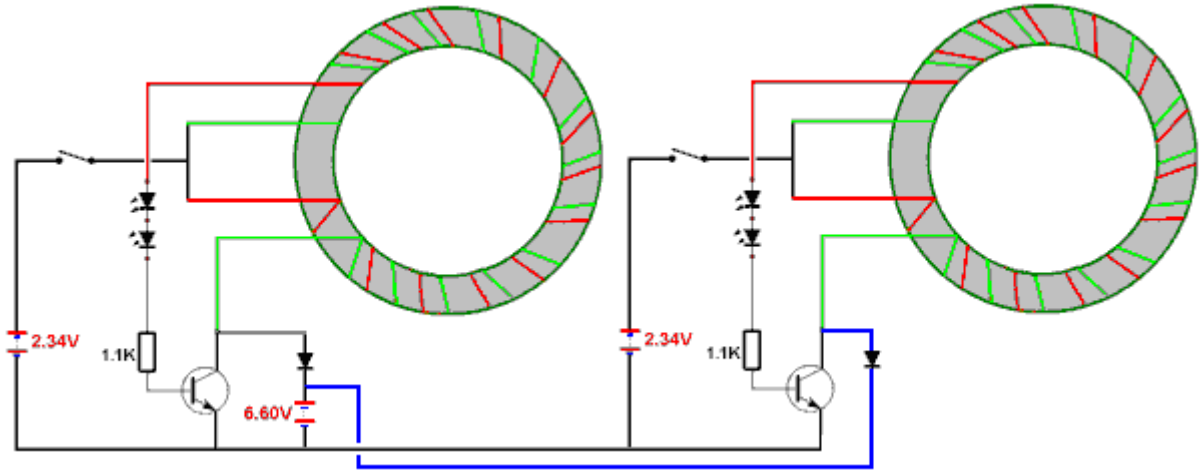
Further Experiments:
This section is about some experiments which I have done to see if I can improve the performance of the Joule Thief. Obviously, I have not tried every possible arrangement and so I invite you (the reader) to do further experiments as the Joule Thief is clearly a very good circuit with which to experiment.

Charging at just 8 milliamps is too low a rate for a 1000 Milliamp-Hour 7.4 Volt Li-Po battery pack, and so, it is necessary to increase the rate of charge. This can be done by using two or more of these circuits connected in parallel as shown here:

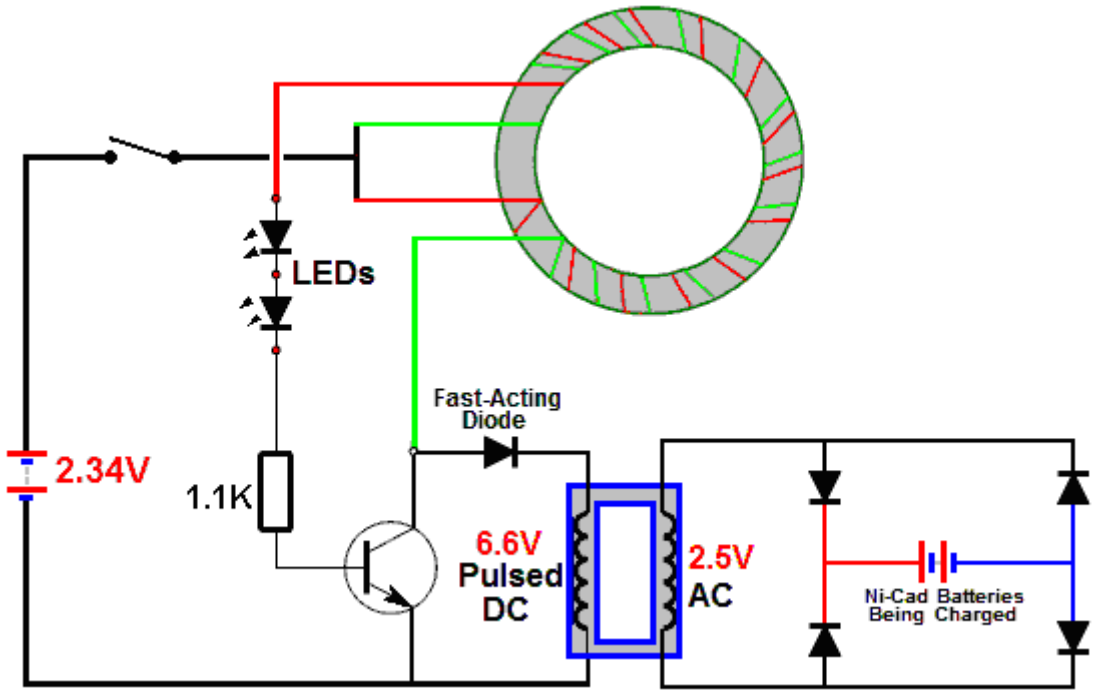


Batteries which are almost fully discharged, have a higher internal resistance than a new battery, and so the higher the current drawn by the circuit the lower the efficiency of the circuit will be. As a result of this, the effective input voltage used by this circuit is actually lower than the 2.34 volts shown in the circuit diagram.

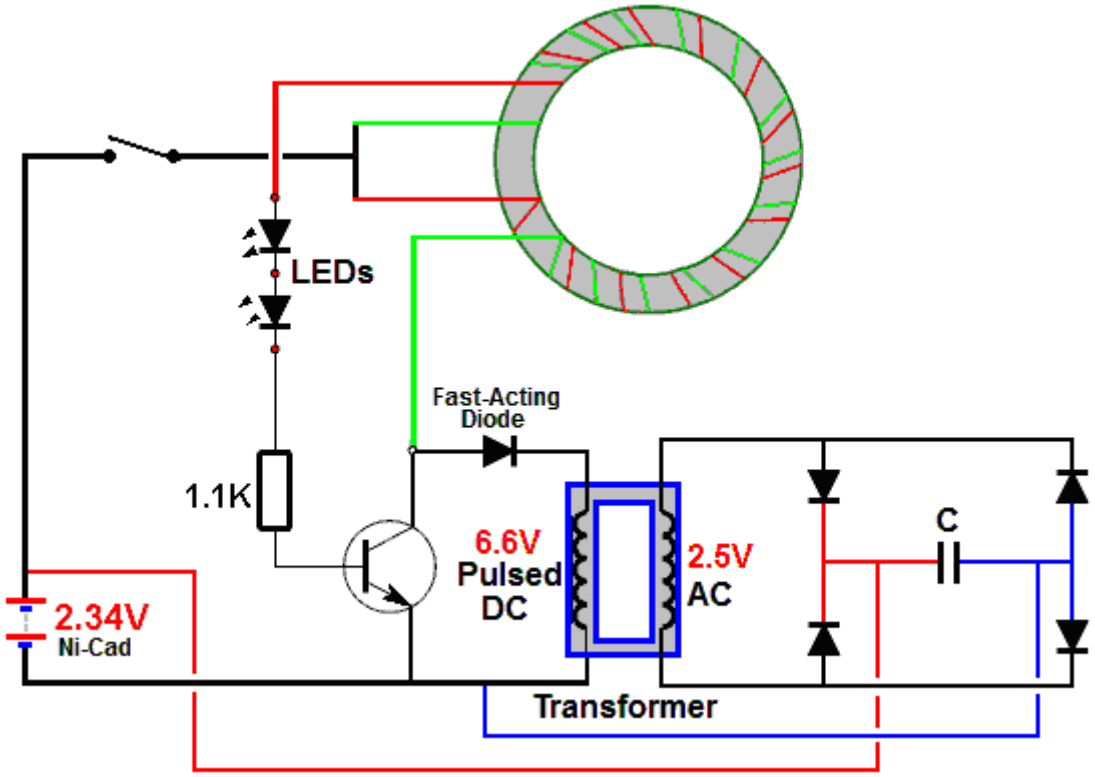
So perhaps it should be arranged like this:



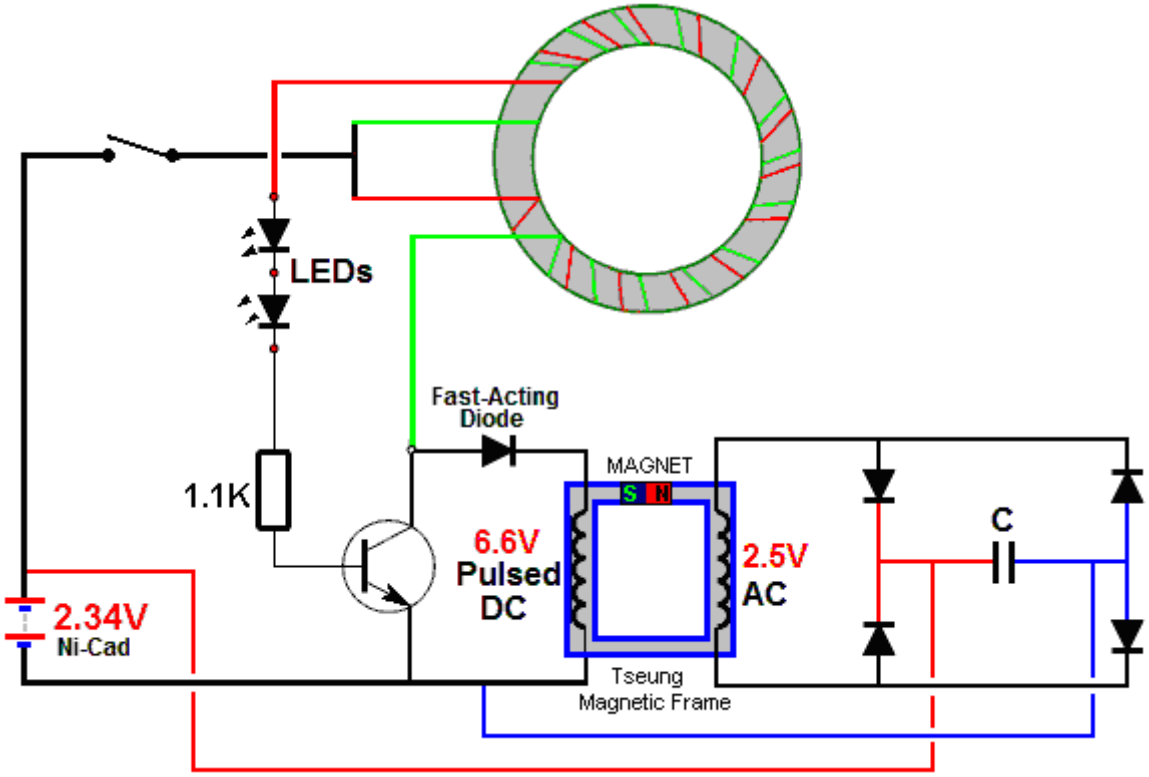
Remember that the output voltage should always be more than twice the input voltage. So if you want to charge Ni-Cad batteries which have a lower voltage, then you should use a step-down transformer as shown here:



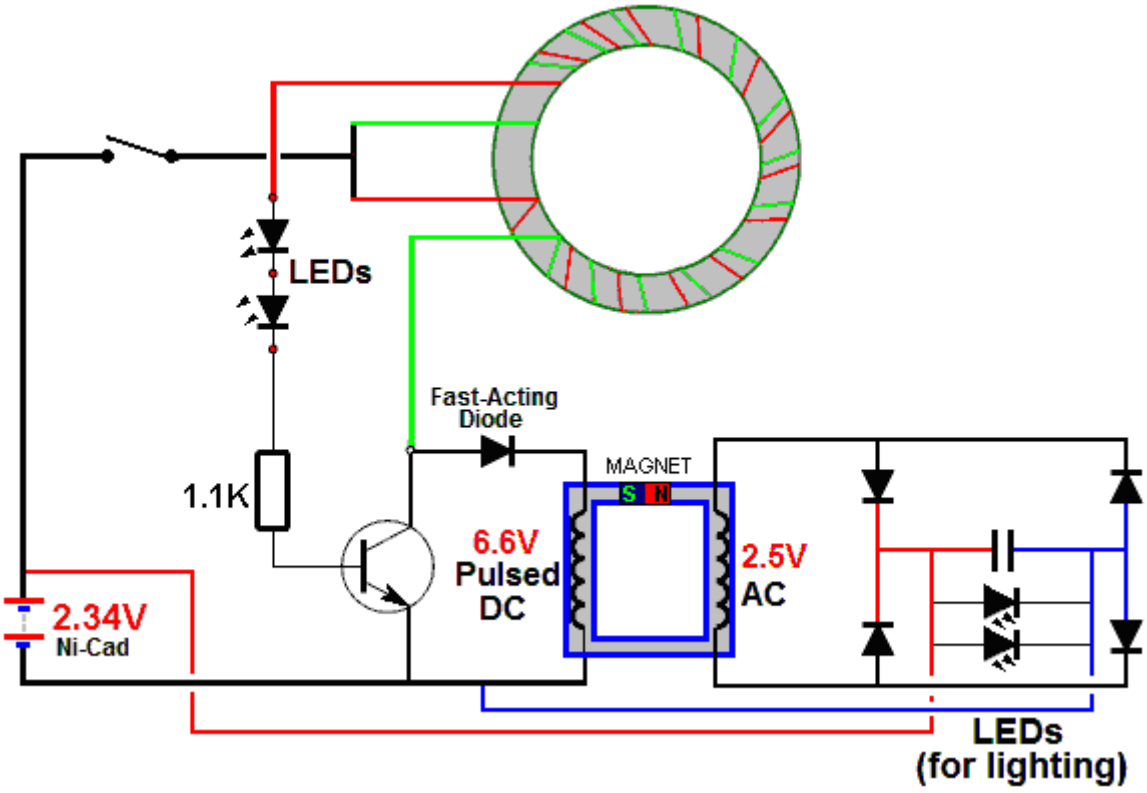
With the appropriate modifications it can become self-powered and self-charging as shown here:



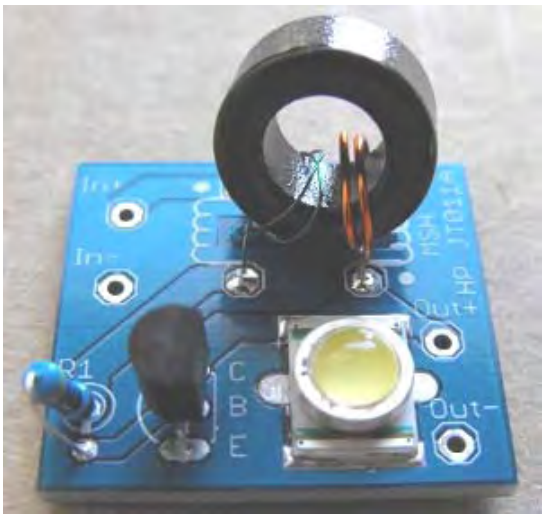
The performance of this circuit can be boosted further by using an asymmetrical transformer such as Lawrence Tseung's Magnetic Frame or Thane Heins' transformer as shown here:



Perpetual Lighting circuit:
Higher voltage results in lower LED resistance and so will decrease the efficiency of the circuit through negative feedback and so the circuit can become stabilised as shown here:



Also, check out the high-power Joule Thief circuits with 1-watt output and various other useful products at http://www.madscientisthut.com/Shopping/agora.cgi?product=Energy%20Harvesting&user4=Joule%20Thief%20Kits.

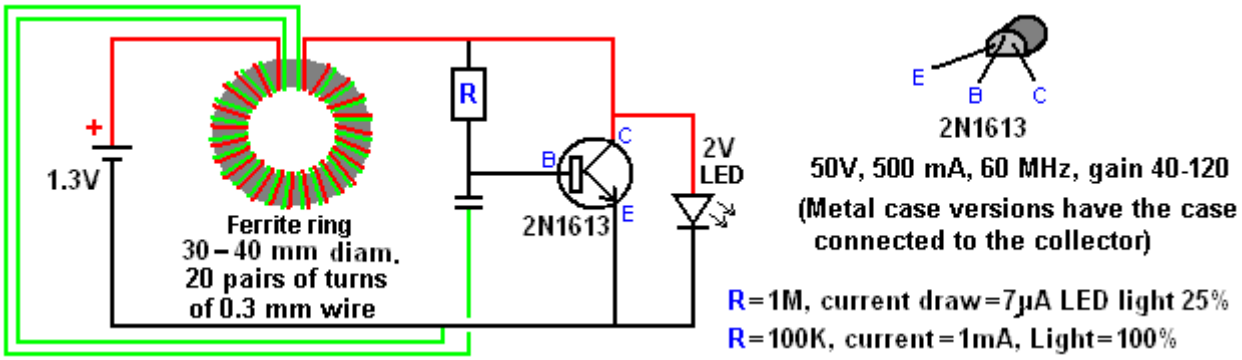


We come now to circuits designed by **Johnny Aum** of Romania who is an independent free-energy researcher since 1982 in spite of everyone around him saying that free-energy is not possible, shows some of his work on his johnnyaum3 YouTube channel. Johnny has shared three of his many high-efficiency designs, including his successful permanent magnet motor design, on the JL Naudin website since 1999.

The following two circuits are still under development. They can be considered to be Joule Thief circuits as they do roughly the same thing, but strictly speaking, these are not Joule Thief circuits but instead are completely new innovative designs, initially developed between 2009 and 2010.

The following circuit is aimed at minimum current draw and while it looks quite like a Joule Thief, you will notice that the end of one of the toroid windings does **not** connect to the start of the other winding, and while the wires

are bi-filar wound side by side as in a Joule Thief, the different connection of the windings makes it a quite different circuit. It runs from 44 to 49 kHz with increased COP at the higher frequency.



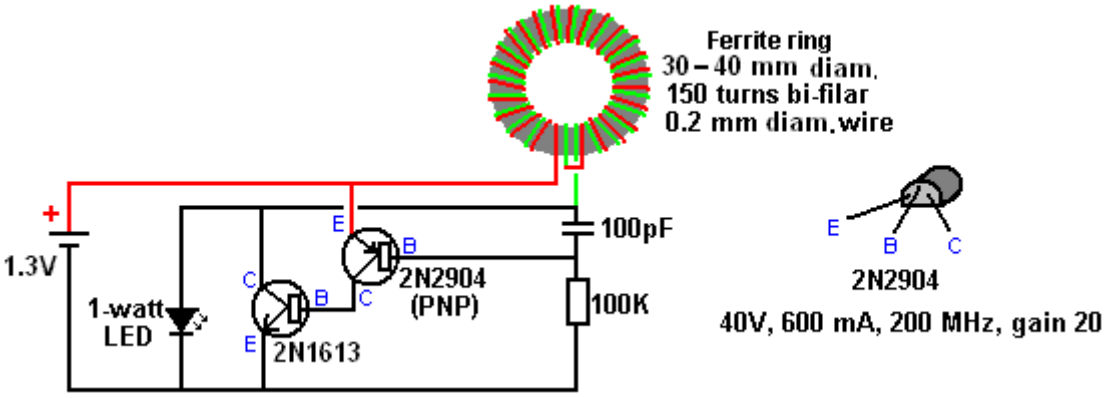
This circuit has a spectacular COP of around 650 at very low currents and you will notice that with a resistor value of 1Megohm, the current draw is only 7 microamps. Both of these circuits can recharge the battery to a certain degree and can recondition the driving battery. It is generally found that in any device, as the power level increases, the COP drops off. To get the full light output from the LED, the current increases to 1 milliamp, which, of course, is very impressive performance, and you can imagine how long a NiMh 3000 milliamp-hour AA battery could keep the LED lit at full power.

The second circuit from Johnny Aum is intended for more useful lighting levels for use where the mains is not available and can run for ten days on one battery due to the slight degree of battery recharging produced by the circuit. Here are some of Johnny's prototypes in operation:

Johnny stress that this circuit generates healing properties and the light is whiter than that produced by a classic circuit. This circuit runs at around 15 kHz. It should be pointed out that an obvious way to increase the lighting level is to have additional LEDs lit, whether by using more than one connected in parallel, and/or using two or more circuits, this is quite feasible as the circuits are very small, lightweight and cheap to make. One-watt LEDs in many different varieties are readily available from different suppliers.

The circuit for driving these powerful LEDs is somewhat different, with a PNP transistor being connected directly to a 2N1613 NPN transistor, boosting its gain by a factor of about 20 times. This connection method has no

appreciable voltage drop when switched on and is convenient for circuits which use very low voltages, such as this circuit. The ferrite toroid in this circuit is wound around its whole circumference with thin 0.2 mm diameter wire. On the prototypes, this resulted in 150 turns of the two wires side by side. This winding is a genuine bi-filar winding, but it is not connected like a Joule Thief. Instead, the end of one wire is connected to the start of the other wire, but not to anything else, leaving just two wires coming away from the toroid winding. It looks like this:



As you can see, this is a very simple-looking circuit with very few components, and yet it is very effective in driving a powerful LED.

A Suggestion

This is an arrangement of readily available, cheap parts for operating a useful light. Probably the most efficient circuit for this task is the 'Joule Thief' circuit and the light bulb usually chosen for this circuit is the compact fluorescent lightbulb ("CFL") which is widely available around the world and although there are many variations, looks something like this:

The arrow is pointing to a join in the construction where the bulb section joins the lower section which houses circuitry and the mains connector which can be one of many different types. However, that type of bulb has mains circuitry built into it which means that in order to get the higher efficiency which we would like, each bulb needs to be physically modified which is far from ideal. The circuitry inside a bulb of this type, converts the alternating current of the mains into Direct Current and then uses that direct current to supply an oscillator circuit which generates high-frequency pulses which power the bulb. Unfortunately, that circuitry gets in the way of low-power operation and so needs to be removed. People who have done this, say that the bottom can be popped off just by inserting a screwdriver into the crack and twisting the screwdriver to force the two pieces apart. If you try that, then let me wish you luck as it has never worked for me, even when applying enough force to permanently damage the material on both sides of the crack. Instead, I do it by cutting the plastic base 8 mm below the arrow line resulting in this:

The 8 mm clearance is because the glass tubes project down below the join and we need the remaining section of the plastic housing to support the bulb when we mount it on our circuit box. The cut should be made when holding the base section as the glass tubes are very fragile and easily damaged. If available, a Dremel or similar tool which has a small cutting disc very well suited to making this cut. The cut should be just deep enough to go through the plastic wall but not further than that. There is a tiny circuit board contained in the lower part, generally, with some very nice components which can be used for other circuits. The cut-down light bulb can be mounted on the circuit box using a strong adhesive, or a hole can be drilled very carefully in the centre, between the tubes (holding the base and not the tubes when doing this) and the base can then be bolted to the component box.

The glass tubes are U-shaped and there are two inter-tube horizontal connecting tubes in order to make all of the U-tubes function as one long zig-zag discharge path and get all of the tubes lit up at the same time. The two tubes which do not have this horizontal interconnecting tube, have two wires coming out of the bottom of them, and used to connect the circuit to the tubes. These four wires need to be cut, leaving them as long as possible, then each pair has the insulating enamel scraped off them and then soldered to a length of wire which will be used to connect to the new circuit, or alternatively, direct to the board if very small 'fiddly' work is no problem.

This adaption makes this low-cost method unsuitable for applications where an off-the-shelf bulb is needed so that replacements can be made without the need for any technically skilled person to be available. What we need therefore, is a bulb which does not have the mains circuitry (called "ballast") built into it, and the earlier 'PLD' bulbs are in that category. Unfortunately, they are more expensive and not so widely available. They look like this:

4-pin PLD bulb

2-pin PLD bulb

These have the enormous advantage of not requiring any work on the bulb in order to operate immediately with our circuit. We can build a suitable circuit from scratch, but it is very popular with home constructors making a one-off experiment to use the very cheap circuit found in Fuji's "Quick Snap" disposable camera which is very widely available. It looks like this:

There are various ways of adapting the circuit board found inside the camera, and thanks is due to 'Gadgetmall' of the http://www.overunity.com/10723/fuji-joule-thief-full-instructions-video-and-pics/new/#new forum for sharing his method of adaption and expertise with these circuits, which has allowed him to run a fluorescent light for 38 hours powered by just one AA size 1.5 volt battery

Disclaimer: This document must not be considered to be a recommendation for you to actually attempt to undertake any of the following modifications, and should you decide to do so, then any loss, damage or injury are wholly your responsibility and not that of anyone else.

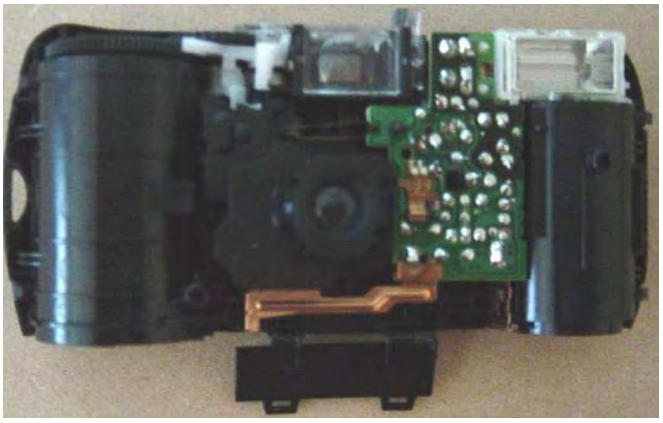
The camera needs to be taken apart in order to get at the circuitry inside it. A word of warning here, there is a high-voltage capacitor inside the camera and if it happens to be charged, then it is quite capable of giving you a really nasty shock, so as soon as the circuit board is exposed, I strongly recommend that you take great care to avoid a shock, even though it is not likely to be a fatal shock. As soon as the capacitor is exposed, then short across it's wires using a metal tool which has a plastic handle, such as a screwdriver or pair of pliers with a well-insulated grip. If the capacitor happens to be charged, then that may produce a bright spark which makes a loud crack.

The camera is taken apart like this:

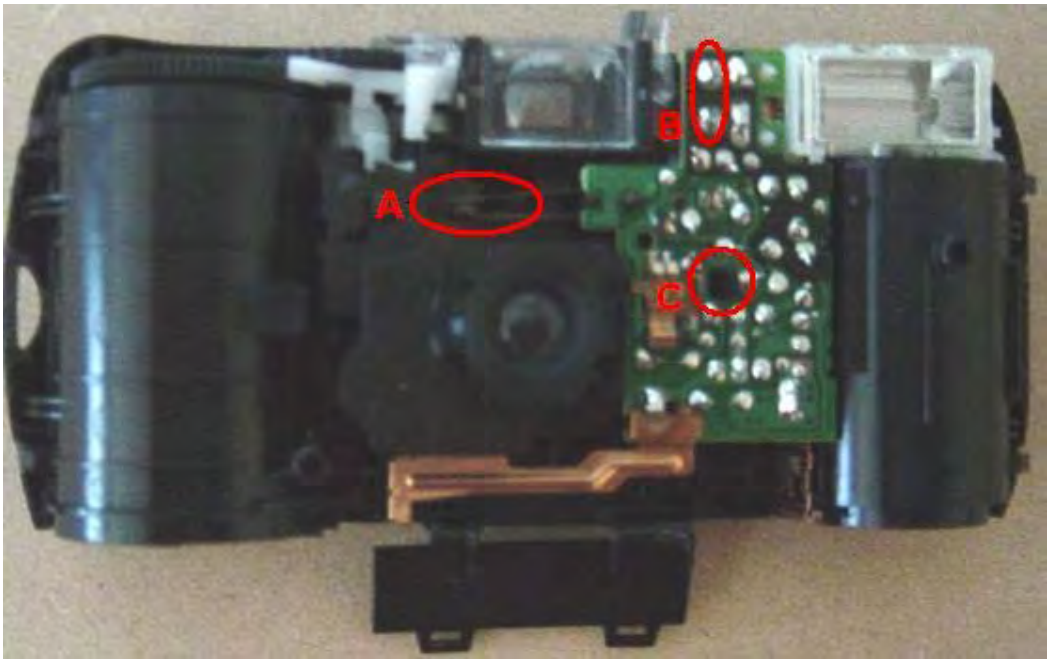
1. The green covering which is a piece of very sticky, strong plastic is peeled off. The join is underneath, where the black colouring ends.

2. In the middle of the bottom, there is a flap which you lever up, uncovering the battery. There are various varieties of Fuji "Quick Snap" camera with different circuitry, the one shown here was supplied by Asda (Wal-Mart in the UK) in 2012 and the circuit board is marked A07 or A60 while some earlier versions have a different layout for some components and even have the battery inserted the other way round. Before you remove the battery which in the UK is a 1.5V AAA alkaline battery, make a careful note of which way round the battery is inserted. In this case, the plus of the battery connects with the long copper arm. Remove the battery.

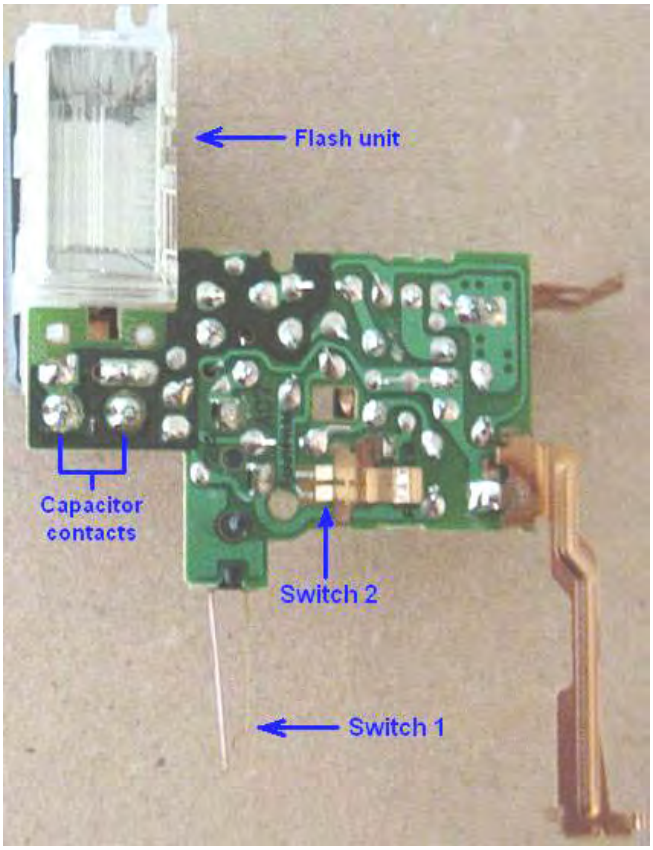
3. Pull off the black plastic covers on the underside of the camera, located at each end of the battery compartment, and then using a screwdriver, force the two halves of the black camera case apart, which leaves the front of the camera looking like this:



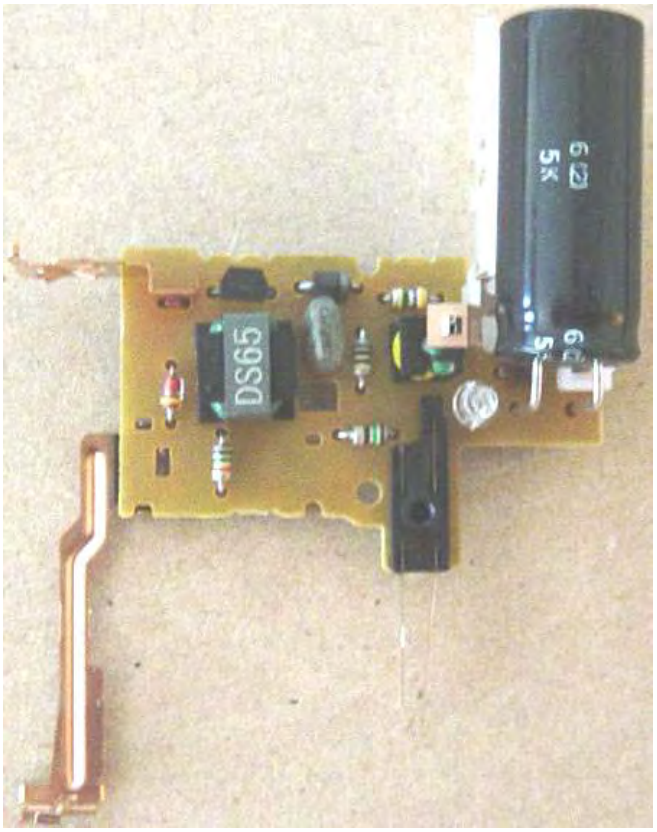
4. Make sure that the flash is not charged, first, by using a non-conducting item to press together the switch contacts marked "A" in the following picture, and then using some metal object with a plastic handle, bridge across the gap between the soldered points ringed around and marked "B" as they are the ends of the high-voltage capacitor. If the capacitor happens to be charged, then there will be a spark and a loud sound, but this is unlikely with a new camera unless you have been pressing the buttons since unwrapping it.



5. Press the black plastic latch marked “C” in the picture above, over towards the left and that releases the circuit board which can be lifted out and looks like this:



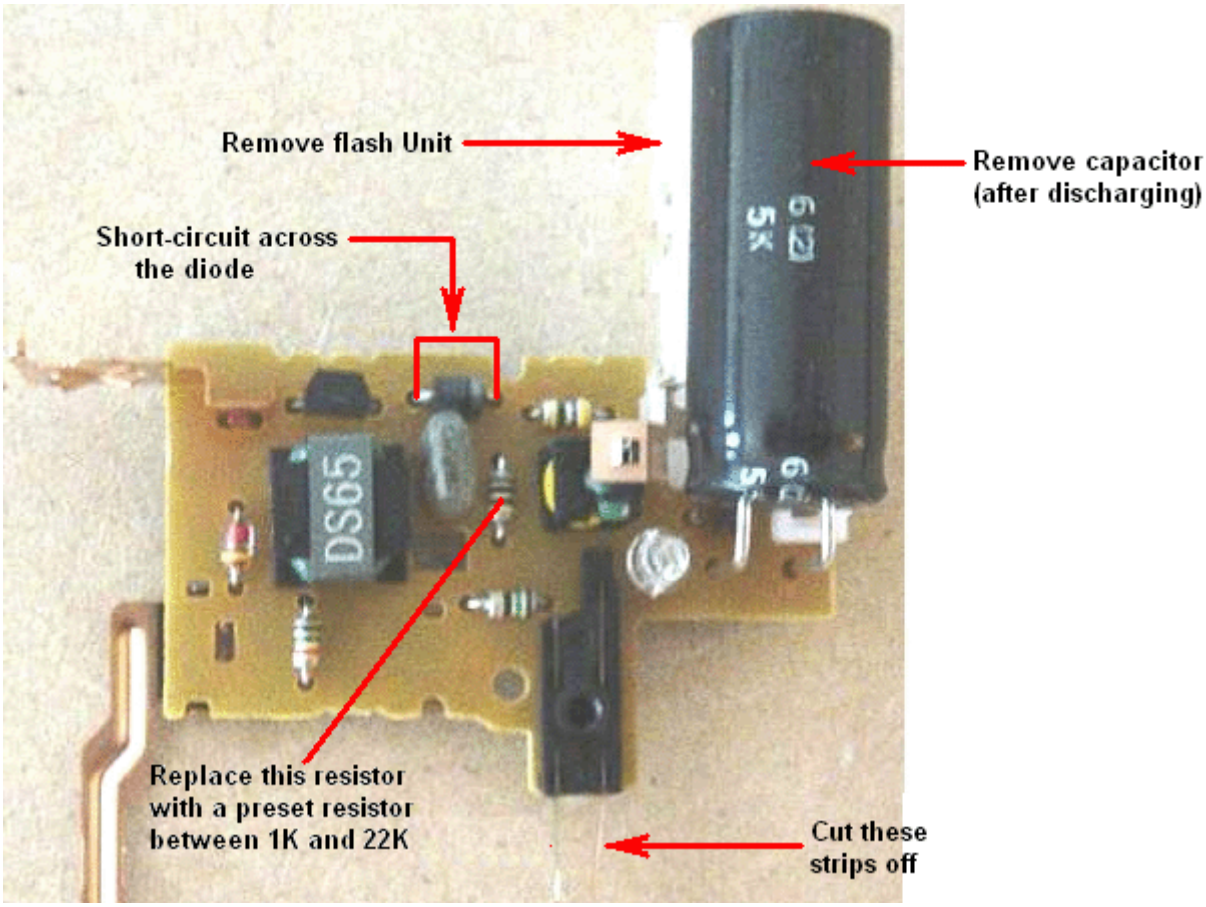
And seen from the top:



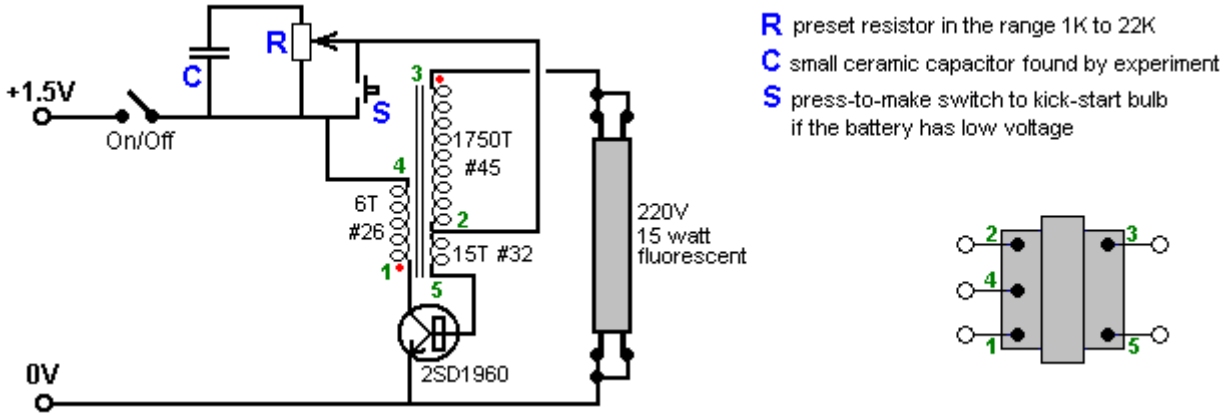
6. The board is quite small, being about 40 mm x 25 mm when the capacitor and flash unit are removed which is the next thing to do, probably by cutting the very tough capacitor leads and then cutting off the plastic pins holding the flash unit to the board, levering it upwards and cutting away it's metal contacts which link it to the board.

7. The two switches which form the camera's flash-charging switch and it's shutter release switch need to be wired permanently closed. These are marked "Switch 1" and "Switch 2" in step 5. above. I suggest that you cut the arms of Switch 1 to about half length, clamp them together with a pair of long-nosed pliers and bend them over to crimp them flat, and then solder them together. Switch 2 needs to be bridged across to make it permanently closed. A clip can be used to clamp the upper and lower contacts together so that they can be bridged with solder.

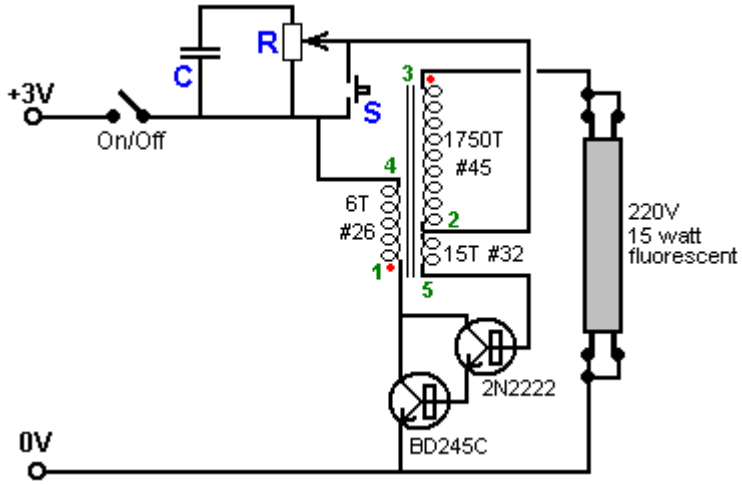
8. The rest of the modification is seen from the top of the board:



This is to achieve this 'Gadgetmall' circuit:

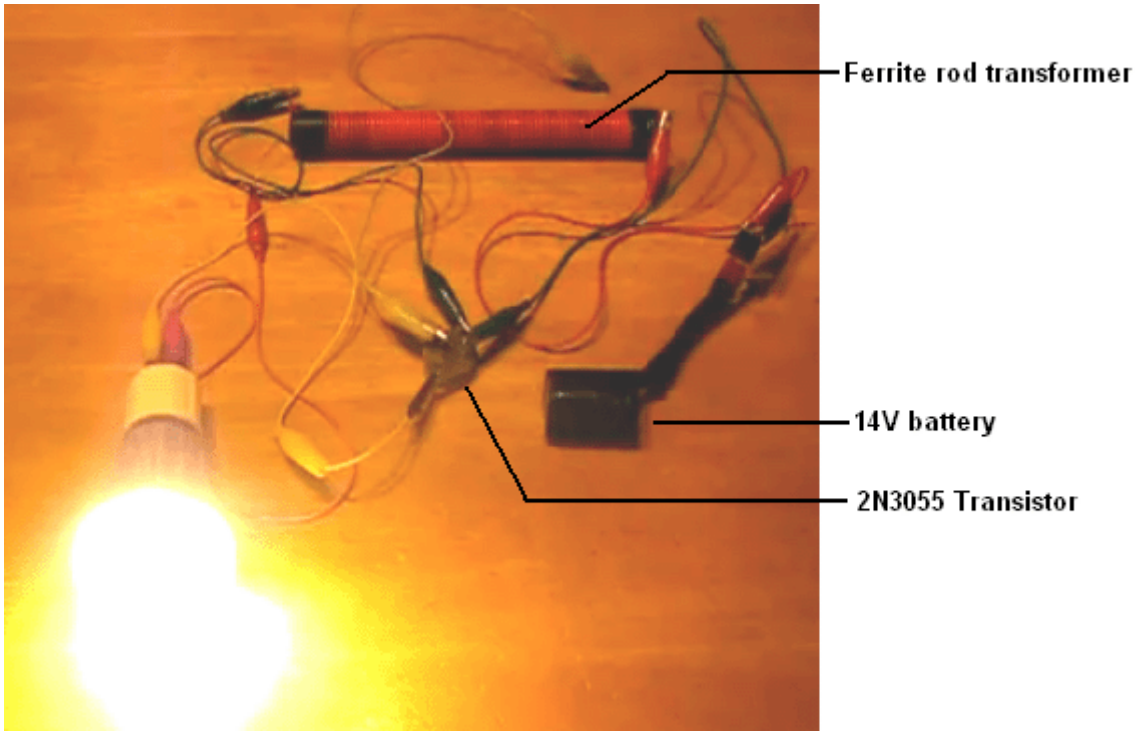


Gadgetmall comments that increasing the voltage supplied to this circuit, puts the 2SD1960 transistor at risk as it will overheat. That transistor is tiny, has no heat sink or space to fit one and was only intended to be on for a few seconds while the capacitor charges up in preparation for operating the flash bulb. This circuit modification runs the transistor continuously for very long periods and so we are already stepping outside the Fuji circuit designer's operating conditions. Also, we would like to run the circuit with somewhat higher voltage in order to get improved operation of the fluorescent tube or bulb. Consequently, we might consider using a more powerful transistor. The 2SD1960 transistor is rated at 30-volts, 5-amps, 170 MHz and 0.75 watts, so we might consider swapping it for, say, a BD245C transistor rated at 100-volts, 10-amps, 3 MHz and 80 watts as our circuit runs at under 0.1 MHz and the BD245C transistor can be mounted on a heat sink, although with it's much greater handling capacity, it should stay cool at these tiny powers. We can boost the gain of the BD245C by a factor of 200 or so, by using a BC109C or a 2N2222 transistor to form a Darlington pair, making the circuit:

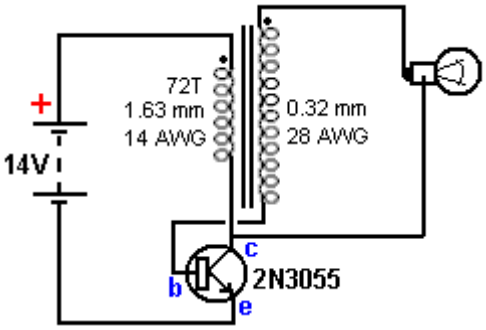


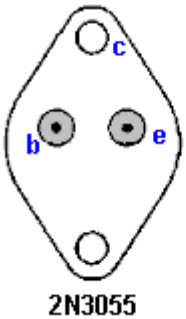
LaserSaber's Joule Thief Lighting Circuits.

Another very successful variation on the basic Joule Thief was put into the Public Domain on 4th October 2012. Details are on his web site www.laserhacker.com. His variation is extremely simple, both in concept and construction:



and in his video: http://solarpower.energygratis.com/2012/10/09/solar-electricity-super-joule-ringer-3-0-real-world-power-made-easy/ he demonstrates it lighting a mains-voltage LED bulb (seen above), a mains filament bulb, a halogen bulb and a compact fluorescent bulb with the ballast still in place, that is, as is sold in shops without any modification. The circuit is:



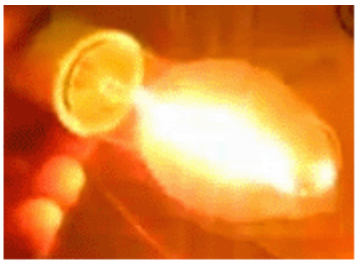


This “Super Joule Ringer 3.0” circuit is unusual in that the feedback to the base of the 2N3055 transistor is through the load (the bulb). The circuit is a high-frequency DC inverter which could hardly be any more simple, but please be aware that the high-frequency output voltage spikes are not sine wave, nor are they voltage controlled, and so this is not a circuit for driving things like TV sets. The transistor oscillates at high frequency governed by the characteristics of the 72-turn primary winding of the transformer. The output voltage level is a combination of the battery voltage and the turns ratio of the transformer.

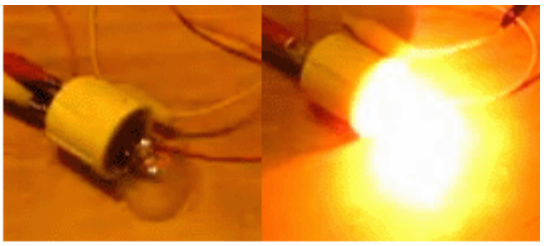
To raise the output voltage, the number of turns in the secondary can be increased. The number of turns is not specified, other than to say that the turns are laid close, side-by-side along the whole length of the 8-inch (200 mm) long ferrite rod, using enamelled copper wire of 0.32 mm diameter (28 AWG or 30 swg). Using arithmetic, that would suggest that there are some 600 turns wound directly on to the ferrite rod. The ferrite rod itself has a diameter of 0.625 inches (15.88 mm) which makes it an item which is likely to be difficult to find. However, I strongly suspect that the diameter of the ferrite rod is not at all critical.

Both of the coils are wound on the ferrite rod in the same direction, the secondary being wound first, placing it underneath the primary which is wound with 72 turns of insulated wire of 1.63 mm diameter (14 AWG or 16 swg). No current draw is indicated and it seems likely that it is quite high there being just two helical coils in the drive section. The (110-volt) bulbs demonstrated in his video are:

(BALLASTED) CFL



HALOGEN



SMALL INCANDESCENT

LED

MEDIUM AND LARGE INCANDESCENTS

A second, very practical lighting system from ‘LaserSaber’ is a Joule Thief driven LED lighting box, running from just one AA rechargeable battery. It looks like this:

Mounted on the end of the box is a single small solar panel:

which is able to charge the battery during the day. The unit is built into a small box which opens up to allow easy wiring, and different angles for positioning the LED lights. Not seen in these pictures is one other large LED cluster light which is also lit. It is a slight pity that the case does not allow the LED lights to shine on the solar panel when the lights are on, as that would provide some battery charging when the light is being used. The LED lights can provide a good deal of light:

A video of this unit being demonstrated is at http://www.youtube.com/watch?v=sFpzkyP6DCU.

Ed Gray's Power System.

The power tube presented to the public by Edwin Gray snr. operates by generating a series of very short, very sharp pulses using a spark gap. This device is reputed to have a power output which is one hundred times that of the power input. Ed Gray and his electric pulse motor are very famous, but as far as I am aware, nobody has successfully replicated this claimed performance. Further, an in-depth examination of the background details by Mr Mark McKay have turned up a number of facts which present a very different picture, and while it is perfectly correct to say that spark-gap pulses generate a good waveform for shocking the local zero-point energy field into the sort of imbalance which can provide a massive power inflow into a device or circuit, we need to be careful to get the full facts in this case.

First, let us put the whole thing in its proper perspective. In May 1973, Cal-Tech in the US performed an independent assessment of an engine provided to them by Edwin Gray. They measured the input and the output and certified that the output power was 275 times greater than the input power. This demonstrates clearly that excess power can be drawn into an engine and provide a performance which can power both the engine as well as doing additional useful work.

Having said that, it needs to be made clear that Edwin Gray did not build that small motor, did not understand how it worked, nor did he ever disclose the design in any of the patents which he obtained afterwards. We need to follow the sequence of events and notice when each thing happened. The history is as follows:

In 1957, a Russian immigrant to the USA, one Alexei Poppoff, showed Edwin Gray a circuit which he said that he had been shown by Nikola Tesla. Edwin Gray did not understand the circuit and had no idea how to create anything useful based on it. He then joined up with his next-door neighbour Marvin Cole, who held a Masters degree in Mechanical Engineering and who, unlike Gray, was able to understand the circuitry.

In 1958, Ed Gray left the Los Angles area in a hurry.

From 1958 to 1967 Marvin Cole, working alone, designed and built ever more powerful prototype engines, and it was a small one of these which was tested by Cal-Tech. In this period, Marvin also developed ever more powerful power supplies, which are the really important item in all of this.

In 1967, Ed Gray rejoins Marvin Cole and together from 1967 to 1972 they solicited venture capital and promoted the technology.

Early in 1972, Marvin Cole disappeared and never saw Gray again. It is not clear if he was intimidated, died, or just did not want to be involved in all the publicity and effort needed to turn the prototype engines into a commercial product. No matter what the reason, the result was that Edwin Gray was suddenly disconnected from the brains behind the project, and that left him in a very difficult position. He didn't want to let go of the dream of becoming rich through this spectacular development, and so he tried to continue the development on his own.

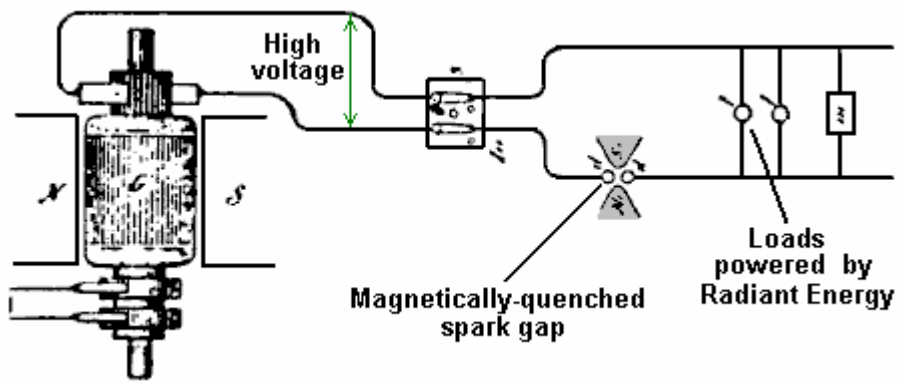
As already mentioned, in May of the following year (1973), Gray had a small Marvin Cole motor independently third-party tested at the famous Cal-Tech laboratory in Los Angles, where a measured input of just 27 watts produced a measured output of 10 horsepower (7460 watts). The objective was to provide solid evidence of a new technology which was capable of changing the world and so would attract investors. To further boost his image and convince potential investors, in that same year of 1973, Edwin staged demonstrations which jumped electromagnets up into the air, showing the strength of the power which drove the Marvin Cole engines.

It is very important to understand that all of Edwin Gray's patents were applied for after the departure of Marvin Cole. These do not disclose the technology tested by Cal-Tech and it must be understood that Edwin was very much afraid of revealing anything important in any of the patents in case some other person would understand the things which were a mystery to him and snatch away the prize of commercial success. Try watching Peter Lindemann's informative video at http://www.youtube.com/watch?v=dmf10hggUm4 for considerable additional information.

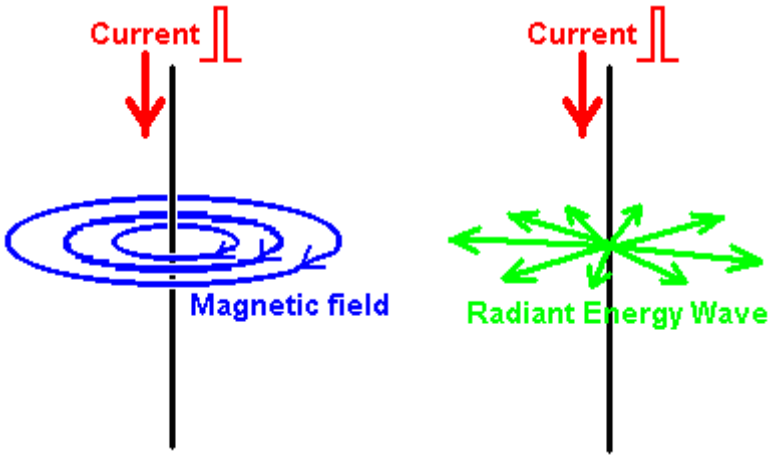
Marvin Cole's power system produced "cold electricity" which could power lights and other devices. It was frequently demonstrated that the output was not conventional electricity and powered light bulbs which were placed under water and at the same time, it was quite safe for a hand to be put into that same water along with the lit bulb. The glass of the conventional bulbs used in these demonstrations would have shattered when placed under water if they had been powered by conventional "hot electricity" as the sudden change in temperature

would have broken the glass. Powered as they were by "cold electricity", they ran cool and so there was no stress on the glass when submerged in water.

Peter Lindemann points out that Ed Gray's power conversion tube circuit is effectively a copy of Nikola Tesla's circuit for doing the same thing:



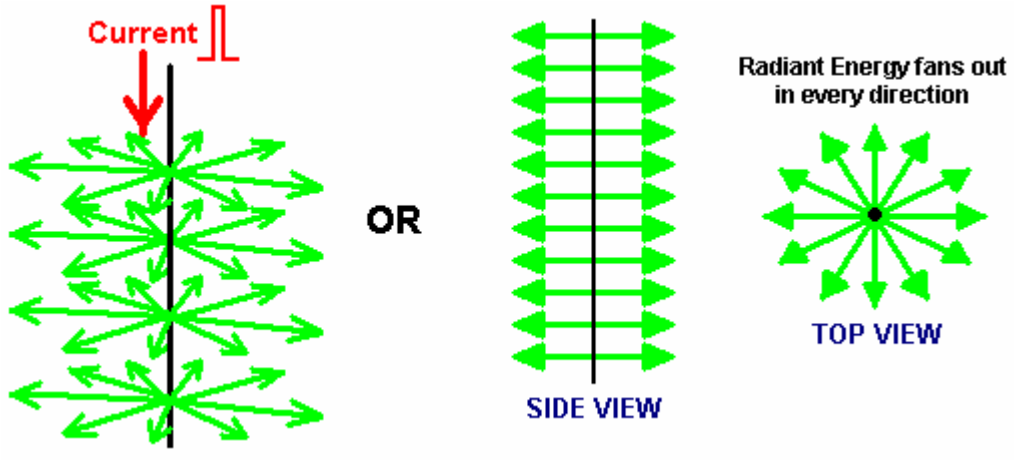
This was disclosed by Tesla in his 'Philadelphia and St Louis' lecture in 1893 and shows how loads can be powered when a high voltage source is pulsed by a magnetically-quenched sparks - this creates DC pulses of **very** short duration.



The diagram above, illustrates the difference between the Magnetic field generated around a conductor fed with a pulse of Direct Current and the Radiant Energy waves created by that pulse. If a sharp current pulse is driven down a vertical wire, it causes two different types of field. The first field is magnetic, where the lines of magnetic force rotate around the wire. These lines are horizontal, and rotate clockwise when viewed from above. The magnetic field remains as long as the current flows down the wire.

The second field is the Radiant Energy wave. This wave will only occur if the current pulse is in one direction, i.e. it will not occur if the wire is fed with alternating current. The wave radiates out horizontally from the vertical wire in every direction in the form of a shock wave. It is a one-off event and does not repeat if the current in the wire is maintained. The Radiant Energy briefly unbalances the zero-point energy field and that causes an energy flow as the field moves back into equilibrium again.

The Radiant energy wave is not restricted to a single plane as shown in the diagram above, which is intended to indicate the difference between the electromagnetic field circling around the wire, and the Radiant Energy field which radiates away from the wire. Both of these fields occur at all points along the full length of the wire as shown here:

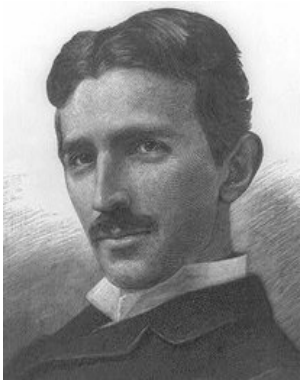


Radiant Energy, when converted to electrical power, produces a different kind of electrical power to that produced by batteries and by the mains supply. Power a motor with conventional electricity and it gets hot under load. Power the same motor by Radiant Energy electricity and under load the motor gets cold. Really overload it by stalling it and the motor housing is likely to be covered with frost. That is why this form of electricity is referred to as "cold" electricity.

In his book "Cold War Secrets - HAARP and Beyond", Gerry Vassilatos quotes research work done in this area by Tesla and others:

Tesla's Experiments.

In 1889 Tesla began experimenting with capacitors charged to high voltages and discharged in very short time intervals. These very short pulses produced very sharp shockwaves which he felt across the front of his whole body. He was aware that closing a switch on a high-voltage dynamo often produced a stinging shock. This was believed to be static electricity and it occurred only at switch-on and only for a few milliseconds. However, in those few milliseconds, bluish needles of energy stand out from the electrical cables and they leak to ground, often through the bodies of any people standing nearby, causing immediate death if the installation is large. While the generators of that time were rated at some thousands of volts, these discharges were millions of volts in intensity. The generator problem was eliminated by the used of highly insulated switches which were provided with a very large ground connection.



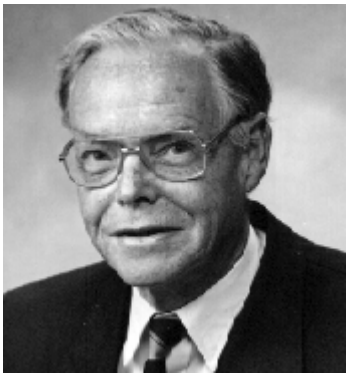
Tesla was intrigued by this phenomenon which appeared to match the effect of his capacitor discharges. He calculated that the voltages produced were hundreds of times greater than could be supplied by the capacitor or generator. It was clear that the power supplied was being amplified or augmented in some way, but the question was, from where was the extra energy coming?

Tesla continued to investigate through experiments, taking precautions against the high voltages being produced. He was soon able to produce these shockwaves whenever he wanted to. The shockwaves produced a stinging sensation no matter where he stood in his laboratory, and hands and face were particularly sensitive to the wave. These waves radiated out and penetrated metal, glass and every other kind of material. This was clearly not an electromagnetic wave, so he called the new wave 'Radiant Electricity'.

Tesla searched the literature to find references to this radiant energy but he could not find much. In 1842, Dr. Joseph Henry had observed that steel needles were magnetised by a Leyden Jar spark discharge located on a different floor of the building. The magnetising wave had passed through brick walls, oak doors, heavy stone and iron flooring and tin ceilings to reach the needles located in a vault in the cellar.

In 1872, Elihu Thomson took a large Ruhmkorrf Spark Coil, attached one pole of the coil to a cold-water pipe and the other pole to a metal table top. This resulted in a series of massive sparks which electrified the metal door knob of the room and produced the stinging shockwaves which Tesla was investigating. He found that any insulated metal object anywhere in the building would produce long continuous white sparks discharging to ground. This discovery was written up briefly in the Scientific American journal later that year.

Tesla concluded that all of the phenomena which he had observed, implied the presence of "a medium of gaseous structure, that is, one consisting of independent carriers capable of free motion - besides the air, another medium is present". This invisible medium is capable of carrying waves of energy through all substances, which suggests that, if physical, its basic structure is much smaller than the atoms which make up commonplace materials, allowing the stream of matter to pass freely through all solids. It appears that all of space is filled with this matter.



Thomas Henry Moray demonstrated this energy flow passing through glass and lighting standard electric light bulbs. Harold Aspden performed an experiment known as the "Aspden Effect" which also indicates the presence of this medium. Harold made this discovery when running tests not related to this subject. He started an electric motor which had a rotor mass of 800 grams and recorded the fact that it took an energy input of 300 joules to bring it up to its running speed of 3,250 revolutions per minute when it was driving no load.

The rotor having a mass of 800 grams and spinning at that speed, its kinetic energy together with that of the drive motor is no more than 15 joules, contrasting with the excessive energy of 300 joules needed to get it rotating at that speed. If the motor is left running for five minutes or more, and then switched off, it comes to rest after a few seconds. But, the motor can then be started again (in the same or opposite direction) and brought up to speed with only 30 joules **provided** that the time lapse between stopping and restarting is no more than a minute or so. If there is a delay of several minutes, then an energy input of 300 joules is needed to get the rotor spinning again.

This is not a transient heating phenomenon. At all times the bearing housings feel cool and any heating in the drive motor would imply an increase of resistance and a build-up of power to a higher steady state condition. The experimental evidence is that there is something unseen, which is put into motion by the machine rotor. That "something" has an effective mass density 20 times that of the rotor, but it is something that can move independently and take several minutes to decay, while the motor comes to rest in a few seconds.

Two machines of different rotor size and composition reveal the phenomenon and tests indicate variations with time of day and compass orientation of the spin axis. One machine, the one incorporating weaker magnets, showed evidence of gaining strength magnetically during the tests which were repeated over a period of several days.

This clearly shows that there is an unseen medium which interacts with everyday objects and actions, and confirms Tesla's discovery. Tesla continued to experiment and determined that a very short uni-directional pulse is necessary to generate the radiant energy wave. In other words, an alternating voltage does not create the effect, it has to be a DC pulse. The shorter the pulse time and the higher the voltage, the greater the energy wave. He found that using a capacitor and an arc discharge mechanism with a very powerful permanent magnet placed at right angles to the spark, improved the performance of his equipment by a major factor.

Additional experiments showed that the effects were altered by adjusting the duration of the electrical pulse. In each instance, the power of the radiated energy appeared to be constant irrespective of the distance from his apparatus. The energy was in the form of individual longitudinal waves. Objects placed near the equipment became powerfully electrified, retaining their charge for many minutes after the equipment was switched off.

Tesla was using a charging dynamo as a power source and he found that if he moved his magnetic discharger to one side of the dynamo, the radiant wave was positive. If he moved the magnetic discharger towards the other side of the dynamo, the radiant wave became negative in sign. This was clearly a new electrical force which travelled as light-like rays, showing them to be different in nature to the electromagnetic waves of Maxwell.

Investigating the effects of adjusting the duration of the pulses, Tesla found that a pulse train which had individual pulses with durations exceeding 100 microseconds, produced pain and mechanical pressures. At this duration, objects in the field visibly vibrated and were even pushed along by the field. Thin wires subjected to sudden bursts of the radiant field, exploded into vapour. When the pulse duration was reduced to 100 microseconds or below, the painful effect was no longer felt and the waves are harmless.

With a pulse duration of 1 microsecond, strong physiological heat was felt. With even shorter pulse durations, spontaneous illuminations capable of filling rooms with white light, were produced. Even shorter pulses produced cool room penetrating breezes with an accompanying uplift in mood and awareness. These effects have been verified by Eric Dollard who has written about them in some detail.

In 1890, Tesla discovered that if he placed a two-foot long single-turn deep copper helix coil near his magnetic disrupter, the thin-walled coil developed a sheath of white sparks with long silvery white streamers rising from the top of the coil. These discharges appeared to have much higher voltages than the generating circuit. This effect was greatly increased if the coil was placed inside the disrupter wire circle. The discharge seemed to hug the surface of the coil with a strange affinity, and rode up its surface to the open end. The shockwave flowed over the coil at right angles to the windings and produced very long discharges from the top of the coil. With the disrupter charge jumping one inch in its magnetic housing, the coil streamers were more than two feet in length. This effect was generated at the moment when the magnetic field quenched the spark and it was wholly unknown at that time.

This train of very short uni-directional pulses causes a very strange field to expand outwards. This field resembles a stuttering electrostatic field but has a far more powerful effect than would be expected from an electrostatic charge. Tesla was unable to account for the enormous voltage multiplication of his apparatus using any of the electrical formula of his day. He therefore presumed that the effect was entirely due to radiant transformation rules which would have to be determined through experimental measurements. This he proceeded to do.

Tesla had discovered a new induction law where radiant shockwaves actually auto-intensified when encountering segmented objects. The segmentation was the key to releasing the action. Radiant shockwaves encountered a helix and “flashed over” the outer skin, from end to end. This shockwave did not pass through the windings of the coil but treated the surface of the coil as a transmission path. Measurements showed that the voltage increase along the surface of the coil was exactly proportional to the length travelled along the coil, with the voltage increase reaching values of 10,000 volts per inch of coil. The 10,000 volts which he was feeding to his 24 inch coil were being magnified to 240,000 volts at the end of his coil. This was unheard of for simple equipment like that. Tesla also discovered that the voltage increase was mathematically linked to the resistance of the coil winding, with higher resistance windings producing higher voltages.

Tesla then began to refer to his disrupter loop as his special “primary” and to the long helical coil as his special “secondary” but he never intended anyone to equate these terms to those referring to electromagnetic transformers which operate in a completely different way.

There was an attribute which baffled Tesla for a time. His measurements showed that there was no current flowing in the long copper ‘secondary’ coil. Voltage was rising with every inch of the coil, but there was no current flow in the coil itself. Tesla started to refer to his measured results as his “electrostatic induction laws”. He found that each coil had its own optimum pulse duration and that the circuit driving it needed to be ‘tuned’ to the coil by adjusting the length of the pulses to give the best performance.

Tesla then noticed that the results given by his experiments paralleled the equations for dynamic gas movements, so he began wondering if the white flame discharges might not be a gaseous manifestation of electrostatic force. He found that when a metal point was connected to the upper terminal of the ‘secondary’ coil, the streamers were directed very much like water flowing through a pipe. When the stream was directed at distant metal plates, it produced electronic charges which could be measured as current at the receiving site but in transit, no current existed. The current only appeared when the stream was intercepted. Eric Dollard has stated that this intercepted current can reach several hundred or even thousands of amps.

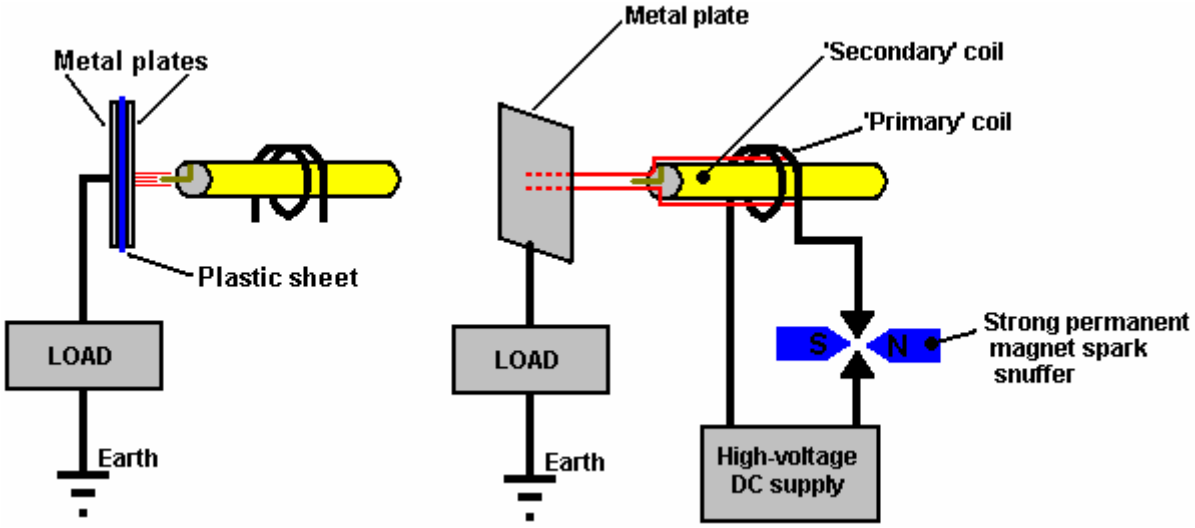
Tesla made another remarkable discovery. He connected a very heavy U-shaped copper bar directly across the primary of his disrupter, forming a dead short-circuit. He then connected several ordinary incandescent filament bulbs between the legs of the U-shaped bar. When the equipment was powered up, the lamps lit with a brilliant cold white light. This is quite impossible with conventional electricity, and it shows clearly that what Tesla was dealing with was something new. This new energy is sometimes called “cold electricity” and Edwin Gray snr. demonstrated how different it is by lighting incandescent-filament bulbs directly from his power tube, submerging them in water and putting his hand in the water. Cold electricity is generally considered to be harmless to humans. Ed Gray’s power tube operates by generating radiant electricity waves by using a spark gap, and collecting the energy using three encasing copper cylinders surrounding the spark gap. The cylinders are drilled with many holes as that enhances the pick-up and the load is driven directly from the current in the cylinders. When lighting bulbs, Ed used an air-cored transformer made of just a few turns of very heavy wire. I, personally, am aware of two people who have independently reproduced Ed’s power tube.

Tesla viewed the streamers coming off his coils as being wasted energy so he tried to suppress them. He tried a conical coil but found that this accentuated the problem. He then tried placing a copper sphere at the top of his

coil. This stopped the streamers but electrons were dislodged from the copper sphere, creating really dangerous conditions. This implied that metals generate electron flows when struck by the coil streamers (as had been seen when the streamers had been aimed at remote metal plates and current was generated as a result).

Tesla designed, built and used large globe lamps which required only a single external plate for receiving the radiant energy. No matter how far away these lamps were from the radiant source, they became brilliantly lit, almost to the level of an arc lamp and far, far brighter than any of the conventional Edison filament lamps. By adjusting the voltage and the pulse duration of his apparatus, Tesla could also heat or cool a room.

Tesla's experiments suggest that a method of extracting free-energy is to use a Tesla coil which has a metal spike instead of the more common metal sphere at the end of the 'secondary' coil. If the Tesla coil is fed with sufficiently short uni-directional pulses and the 'secondary' coil pointed at a metal plate, then it should be possible to draw off serious levels of power from the metal plate, just as Tesla discovered. This has been confirmed by Don Smith who uses two metal plates separated by a layer of plastic dielectric, forming a capacitor. He states that a well designed Tesla coil is capable of producing currents as high as the voltages and he demonstrates a hand-held 28 watt Tesla Coil played on the first plate producing a substantial continuous spark discharge between the second plate and ground. I estimate that the spark produced would have to be thousands of volts at a significant current, which puts it in the kilowatt range, like most of Don's other devices. Video: http://www.metacafe.com/watch/2820531/don_smith_free_energy/ Don's patent is in Chapter 3 and his .pdf document here: http://www.free-energy-info.tuks.nl/Smith.pdf in which he explains many of his high-power designs.



Don also points out that the positioning of the primary coil relative to the secondary coil of a Tesla Coil determines the amount of current which can be provided. Contrary to most opinion, it is possible to have Tesla Coil current as high as the voltage. Don always stresses that you have the option of picking the electrical component (as conventional science has done) which leads to "heat death" while the alternative option of selecting the magnetic component makes "the world your oyster". With a magnetic ripple imposed on the zero-point energy field, which Don prefers to call the 'ambient background energy', you can make as many electric conversions as you wish, without depleting the magnetic event in any way. In other words, you can draw off serious amounts of current from capacitor plates positioned at right angles to the magnetic flow, and every additional pair of plates gives you an additional source of major current without any need to increase the magnetic disturbance in any way. With his single metal plate, Tesla mentioned currents of a thousand amps being available. Please remember that a Tesla Coil produces seriously high voltages and is **not** a toy. Great care is needed around a Tesla Coil so, when it is running, keep well away from it.

Don also states that the collection and transfer of energy requires temporary storage which occurs as the capacitors and coils of a resonant circuit are cycled on and off. The frequency at which the capacitors and coils are pumped, determines the amount of electrical energy that moves onwards. The amount of Energy transferred relates directly to the density of lines of magnetic flux present. The Kinetic Energy formula is helpful in establishing the amount of energy present. This formula points to mass multiplied by the square of the velocity. In the case of electrical energy, intensity of voltage and amperes multiplied by cycles per second, replace velocity. Note that the "acceleration" of the Voltage and the Amperage, increases in a non-linear fashion as the Law of Squares applies, with each unit of increase causing a squaring of the flux lines present. In resonant air-core coil energy transfer, the increase in flux lines present disturbs more electrons than previously and this results in greater output energy than input energy being present and available.

Energy stored, multiplied by the cycles per second, is the energy being pumped by the system. Capacitors and inductors (coils) temporarily store electrons.

Capacitor formula: $W = 0.5 \times C \times V^2 \times Hz$ where:

W is the energy in Joules (Joules = Volts x Amps x seconds)
C is the capacitance in Farads
V is the voltage
Hz is the cycles per second

Inductor formula: $W = 0.5 \times L \times A^2 \times Hz$ where:

W is the energy in Joules
L is the inductance in Henrys
A is the current in amps
Hz is the frequency in cycles per second

Both one Henry and one Farad equal one volt. The higher the frequency, including the squaring of the flux lines, causes a large increase in the amount of energy being produced. This, combined with the use of a resonant energy induction system (all electrons moving in the same direction at the same time), make the move into COP>1 practical.

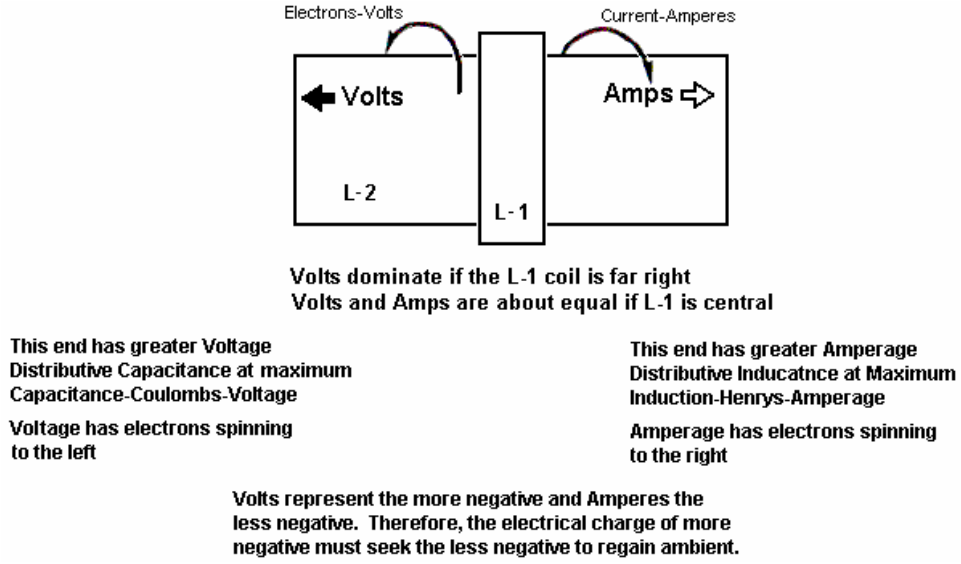
The damping process of conventional electrical power generation, has all of the available electrons bouncing randomly, mostly cancelling out each other, and so the useful energy available is only a very small percentage of the energy which is present. In a resonant induction system, a very high percentage of the energy present is useful. When resonating, (ohms-impedance-Z) becomes zero and all of the energy present becomes available, without being degraded in any way. Ohms is load or wasted energy and amperes is the rate of that wasting.

Now, apply this information to an air-core coil resonant transformer energy system. L-1 and L-2 coils are now present. L-1 has fewer turns and is several times the diameter of L-2. Input from a 12-volt 'gelcel' high-voltage laser module, produces 8,000 volts with low (wasted energy) amperage into 4 turns of coil L-1. Each turn of L-1 then acquires 2,000 volts of resonant potential. Each turn of L-2 is then exposed to an electric flux of 2,000 volts. Each turn at the bottom end of L-2 acquires 2,000 volts. The flux lines are squared and are additive as the voltage and amperage progress towards the top end of L-2's many turns.

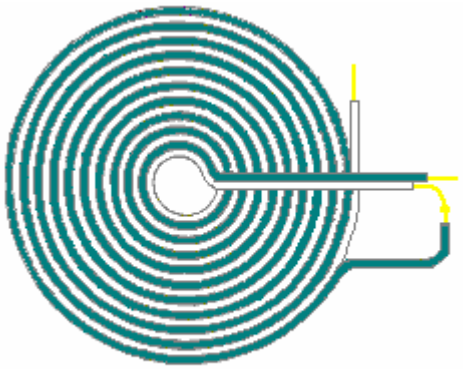
A huge number of flux lines which were not previously present, occur at the top end of L-2. These flux lines excite the electrons nearby in it's earth and air and groundings. This high level of excitement above the ambient, causes a large number of electrons to become available, electrons which previously, were not part of the energy present. At this point, large amounts of excess energy is present. This COP>1 device produces energy at radio frequencies in the megahertz range and this allows it to be small in size and yet produce large amounts of energy. A megawatt sized unit will sit comfortably on a breakfast table. The energy is changed to direct current, and then, to the desired working frequency.

The energy powering these devices is drawn from the surrounding energy field and is not conventional electricity and it does not flow through the wire of the ‘secondary’ coil, but instead, it runs along the outside of the coil and through space to strike the surface of the metal plate, where it generates conventional electric current. Thomas Henry Moray demonstrated that this energy flowing along the outside of the wire can pass through glass without being affected in any way.

In his 1995 paper Don Smith presents the following diagram:



While Tesla's experiment used a metal plate, he patented (US 512,340) a coil type which he said is very effective in picking up this radiant energy. This "pancake" coil type goes by the rather impressive name of "bi-filar serial-connected coil", which, despite it's impressive name is not difficult to wind using two separate strands of wire as shown here:



Flat 'pancake' coil winding method patented by Nikola Tesla

If a strong magnetic field is positioned across the spark gap as shown above, it sharpens the cut-off of the spark and enhances the uni-directional character of the pulse of current. It should be remembered that if a very short sharp pulse of uni-directional current such as is produced by a spark jumping across a spark gap as in the arrangement shown above, occurs in a conductor, then a strong wave of radiant energy radiates out in a plane at right angles to the pulse of current.

This radiant energy wave is quite different from the electromagnetic field generated around the wire carrying the pulse of current. In the Tesla coil arrangement shown above, it should be possible to gather additional free energy through one or more co-axial (like layers of an onion) cylindrical coils around the spark gap leads. These coils will be better if they are would as bi-filar serially-connected coils. The reason for this arrangement is that the magnetic field component of the coils is (nearly) zero as the current flowing through the wire is flowing in opposite directions in alternate turns, and so the magnetic fields produced should cancel out:

Tesla was granted US Patent 685, 957 "Apparatus for the Utilisation of Radiant Energy" in which he shows various ways of handling the energy collected by the metal plate. It is likely that the pick-up techniques shown in the patent of Hermann Plauston, which is in the Appendix, would also work very effectively with this collected energy. Old patents sometimes mention a "condenser" which is the original term for what is nowadays called a "capacitor".

After careful consideration and many experiments, Tesla concluded that the radiant rays which he was utilising, radiated out so rapidly that electrons were unable to keep up with them. The rays were being carried via a medium consisting of extremely mobile, almost mass-less particles, very much smaller than electrons and which,

because of their size and speed, could pass easily through most materials. In spite of their small size, their extreme speed caused them to have considerable momentum. A fact which is **very** difficult to come to terms with is that these rays seem to propagate outwards instantly, with no time delay at all, as if transmitted through matter which is wholly incompressible. It is sometimes called “Radiant Energy” or “RE” for short and appears to have no net charge in conventional terms. This is a unique feature of the universe, with unique characteristics, which if utilised, provides a whole host of new applications and capabilities.

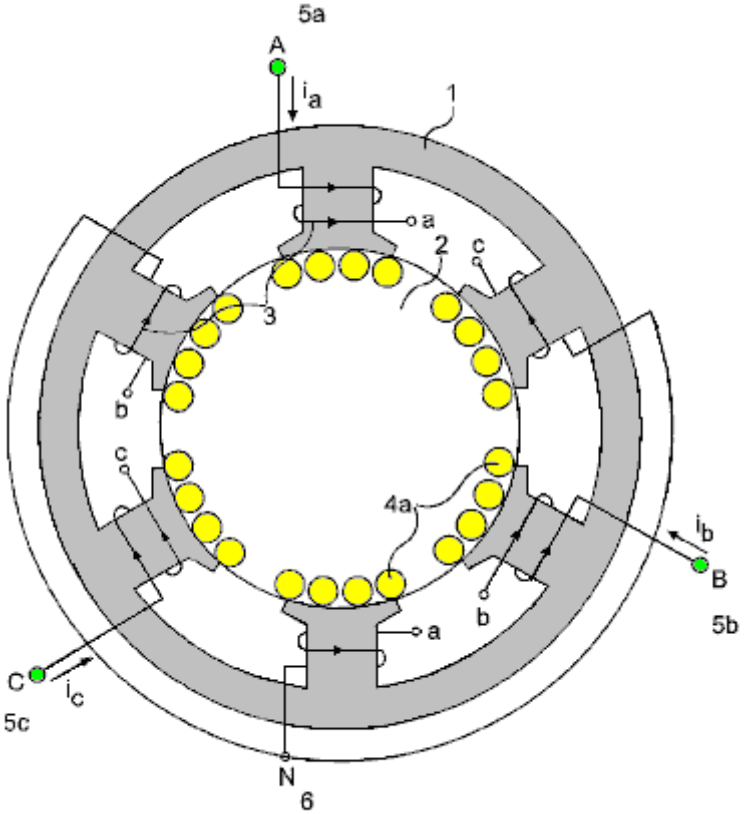
Tesla considered that this newly discovered field acted like a fluid. A hundred and fifteen years later, the cover story of the December 2005 edition of the ‘Scientific American’ journal states that experimental models hint that space-time could be a kind of fluid. It has taken a long time for modern science to start catching up with Tesla. In actual fact, it was Michael Faraday (1781 - 1867) who came up with the idea in the first place.

The Alberto Molina-Martinez Generator.

US patent application US 20020125774 of 6th March 2002, shows a self-powered electrical generator. Like that used by Bob Boyce, this is a toroidal (ring-shaped) frame with several windings on it, as shown in the diagram below. Once it has been powered up with AC mains frequency voltage, it produces so much power that it can supply it’s own input power requirement as well as powering other loads such as light bulbs. This patent application is shown in full in the Appendix.

It is said that the Toroid device built by Stephen Mark and shown in web videos, is a replication of this generator design. The forum at present at http://www.overunity.com/index.php/topic,2535.0.html is dedicated to replicating Stephen Mark’s device and considerable progress has been made. This group is operating on the basis that instead of a metallic toroid core as shown here, that a Mobius-loop toroidal wire core is used. At this point in time, their efforts have not yet produced a circuit which exhibits a COP>1 performance

You will notice that very many different devices, aimed at doing different things, all operate by generating very sharp DC pulses



So, a wide range of different devices have the same background technique for making them work. Meyer used the pulsing for water-splitting in a hydroxy gas cell. Bedini uses the pulsing to charge batteries with cold electricity. Tesla used the pulsing to charge batteries, provide heating, cooling and lighting. Boyce uses pulsing to obtain electrolysis at 1,200% of Faraday’s stated maximum rate of electrolysis. Gray used the pulsing to capture cold electricity to drive a powerful electric motor. Many different applications all based on using very short, very sharp, high-voltage pulses.

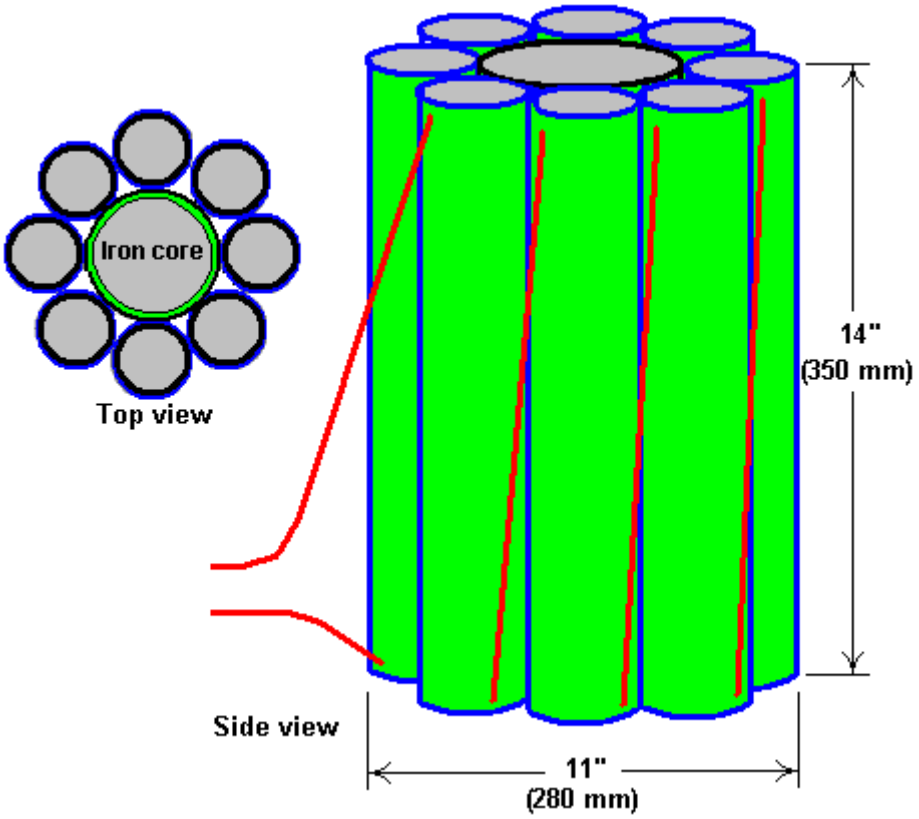
Alfred Hubbard's Self-Powered Generator.
At Portage Bay on Lake Union, Seattle, Washington in America, Alfred Hubbard, an acquaintance of Nikola Tesla, demonstrated in 1919 a self-powered electricity generator design. The generator was about 14 inches (350 mm) tall and 11 inches (280 mm) in diameter. It powered a 35 H.P. electric motor, which pushed an 18-foot boat which contained no batteries, continuously around the bay for several hours. This demonstration was witnessed by thousands and ended because the wiring was beginning to overheat. It was said that the cable used contained seven strands of 0.09-inch (2.286 mm) diameter. Each of those strands would be able to carry 12 amps and so if this is correct, the cable had a current-carrying capacity of about 84 amps. The diameter of the wire including the insulation was said to be 0.34 inches (8.5 mm). The inner core was said to be made of a pipe containing 16 iron rods with 43 turns of wire around it, which if correct, would suggest 43 turns in 14 inches or 3 turns per inch, implying a cylindrical coil with the turns side by side, touching each other.

However, a great deal of misleading information, not to mention a good deal of speculation has been spread around concerning the Hubbard design, which Alfred took three years to develop. Several years after the demonstration, when Hubbard was employed by the Radium Company, he said that radium was used in the device, which is something which I personally, find very hard to believe, and strongly suspect that Hubbard was persuaded to say that by his employers who were selling radium at that time.

Hubbard made a sketch of one of his smaller generators which was used for ordinary household electrical appliances and that showed a very simple design which had eight cylindrical primary coils each of which was wound on a solid iron bar and connected in series. These primary coils surrounded a slightly larger secondary coil of some 35 turns wound around a hollow tube filled with metal bars or wires (presumably of soft iron). This smaller device was about six inches (150 mm) tall and about five inches (125 mm) in diameter. Each core had only one layer of thick insulated wire and not many turns were used.

I understand that when a patent was applied for, the patent application was seized and a spurious "Of National Security Importance" order slapped on it, acting as an unlawful gag order on Hubbard, prohibiting him from ever developing, using, showing or selling it or anything akin to it. The US Patent Office is a privately owned commercial company, and while they will probably be using the design themselves, they certainly have no intention of ever allowing the public to have access to it as energy freedom is a major step towards complete freedom. Consequently, we know next to nothing about Hubbard's successful design.

The general arrangement was something vaguely like this:



In Joseph Cater's book "The Awesome Life Force" he attempts to explain the theory of its operation, but it must be clearly understood that what Cater says is just speculation on his part as Hubbard's actual design was never disclosed publicly.

What Cater says is certainly plausible, and even if it is not Hubbard's design, it is worth investigating and experimenting with. The mechanism put forward by Cater is based on the well-known and widely accepted graph of the magnetisation of soft iron versus applied levels of magnetic force. This graph is highly non-linear and the central section of the graph rises steeply, indicating that there is a considerable increase in the magnetisation of the iron for relatively little increase in energy input.

Cater stresses that the input waveform should be pulsating DC. The method of applying pulsing DC is then, almost the same as for the Clemente Figuera design shown in chapter 3, with an offset base level of DC current flow which needs to be maintained at all times.

Here is the magnetisation graph for soft iron:

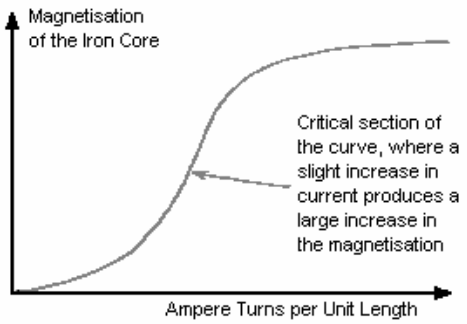


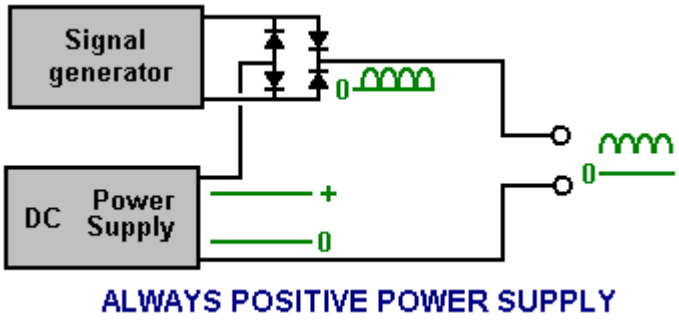
Fig. 29

Fig. 29 shows a graph of the magnetisation of an iron core plotted against ampere turns per unit length. The term "ampere turns" is the number of turns of the coil per unit length of the coil multiplied by the number of amps of current flowing through the coil.

The steep section of the curve appears to start at around 3.5 Tesla, and so, a constant DC current in the magnetising (Hubbard primary) coil needs to provide that level of magnetisation at all times, and the applied

pulsing DC half-sinewave waveform applied on top of that and since the induced EMF in a coil is directly proportional to the rate of change of magnetic flux, it follows that the higher the frequency of this sine wave supply, the better. Using a ramp waveform might well be more effective.

Normal working transformers have ampere-turns which are well below this critical point. The additional EMF induced in the coils by the magnetisation of the iron offsets the natural inductive impedance of the coils. This is why transformers have such a high degree of efficiency. If any material other than iron or special steel were used for the core, the efficiency would drop significantly. Hubbard used part of the output power to provide the input power, and so he only needed to provide input power for less than a second to get the device running. The power supply might well be of this nature:



ALWAYS POSITIVE POWER SUPPLY

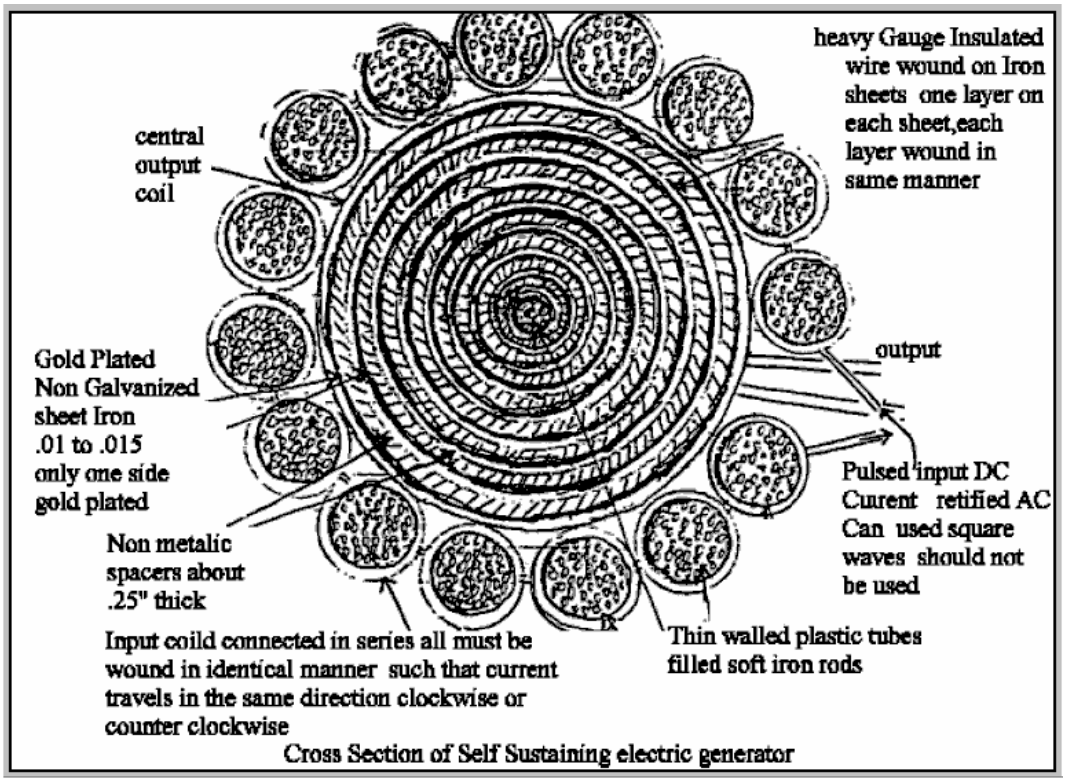
Here, instead of letting the high frequency rectified sine wave (or ramp generator signal) reach zero volts, and additional DC current supply is maintained, and while the signal generator pulses add to the overall voltage applied to the device, the voltage is never allowed to reach zero.

There is possibly another factor which could contribute to the success of the Hubbard device. At that time, the only insulated wire available had thick and heavy insulation. This means that adjacent turns of wire in the coil were separated by a distance equal to twice the thickness of the insulation. Consequently, the gap resulted in a cancellation of magnetic effects produced by electrons flowing in the wire. Since inertia is dependent on the ability to generate a magnetic field, the inertial properties of the electrons would be almost nullified.

There is an optimum distance between the wires which would produce the maximum effect. It seems likely that the thick insulation on Hubbard's wire produced this optimum distance. Most of the resultant magnetic field was that which encircled both wires and that would be the weaker part of the field. This means that a relatively low EMF could accelerate a larger number of electrons to a high velocity during a very short period of time. As the electrons leave the coil, inertia returns. This would result in a backup of a high concentration of electrons in the coil. Since electrostatic repulsion is not affected, electrons would be ejected from the coil at a high velocity despite their increased inertia. This would produce an output of both high voltage and high amperage.

Joseph Cater's Version of the Hubbard Generator.

Although containing conflicting information, there is what appears to be an implementation of the Hubbard coil system, or perhaps a very closely related device from Joseph H. Cater. As usual, information on it is limited and not particularly clear, so the following is just my attempt to piece together some information from different sources. Much of this information comes from a document which has Geoff Egel's name on it and although it seems likely that Geoff is quoting some other source, my thanks goes to him for sharing what we have here. The diagrams give the names of various minor websites none of which exist any longer and so these have been removed as they have no useful purpose any longer. Here is an original diagram from this information:



As it seems to me that there are many conflicting details in this information, I am presenting it here in pretty much the same form in which it reached me. You will notice that the composite central coil is now presented as the secondary rather than the primary. It should be stressed that Hubbard never disclosed his design publicly and so this, and similar information elsewhere, has to be considered to be guesswork.

Mr. Cater claims that a group in California built this device which, it is claimed, performed very well, but he does not claim that he has personally seen or tested such a device. This design is published for researchers and experimenters in order that a working prototype may be developed. Mr. Cater says "I would be willing to give big odds that if my instructions are carried out to the letter, then sensational results will be obtained. It should easily outperform any other generator that has ever been built including the Moray and the Hubbard devices. It could easily be mass-produced.

Some years ago I got word from someone in Germany who built a similar configuration (a very poor replica of this one, where the output coil consisted of only windings on a solid iron bar which in turn was surrounded by smaller coils on smaller bars which constituted the input. Even this was quite successful as the output was three times the input. I do not know what happened to the builder but such a crude device as this could give the world free energy. The output of a small unit could be used as the input for a larger one and so on".

Please bear in mind that these plans are not meant to be explicit in every fine detail, but are provided as the best guide that the author can make with the available data. Therefore you will need to use some of your own ingenuity and design skills in the construction of this rather unusual coil configuration.

The Primary Coil Input-driver: Suggestions for the Bench-test Prototype

I would suggest the construction of an input power supply which can vary Frequency, Voltage and Current. A frequency range of 50 Hz to 1,000 Hz would be a good starting point. The higher the frequency of the input current (the amperage and voltage being held constant) the greater the induced output E.M.F. as it is directly proportional to the frequency (the rate of change of the magnetic flux). A frequency of 50 or 60 Hz would be more convenient to experiment with as these frequencies are standard power mains frequencies, however a frequency of 360 Hz or higher is recommended.

Mr. Cater suggests that for experimental purposes in determining the input needed to get the desired output, that rectified 12 volt AC is used. Sinusoidal waves should be used and not square waves. Because of its tremendous potential, care should be taken to limit the amount of input current. One should start with a low frequency (50 or 60 Hz) and low amperage, then gradually increasing the current until the desired input / output is obtained.

Such caution was not followed with a previous model built by a group in California and it resulted in the

disintegration of the output coil. The iron sheets in this model were not plated and did not have the caps fitted. Nevertheless, it was still an effective orgone accumulator. The gold plating of the iron sheets and the addition of the caps enables it to operate with a much lower input current and lower frequency.

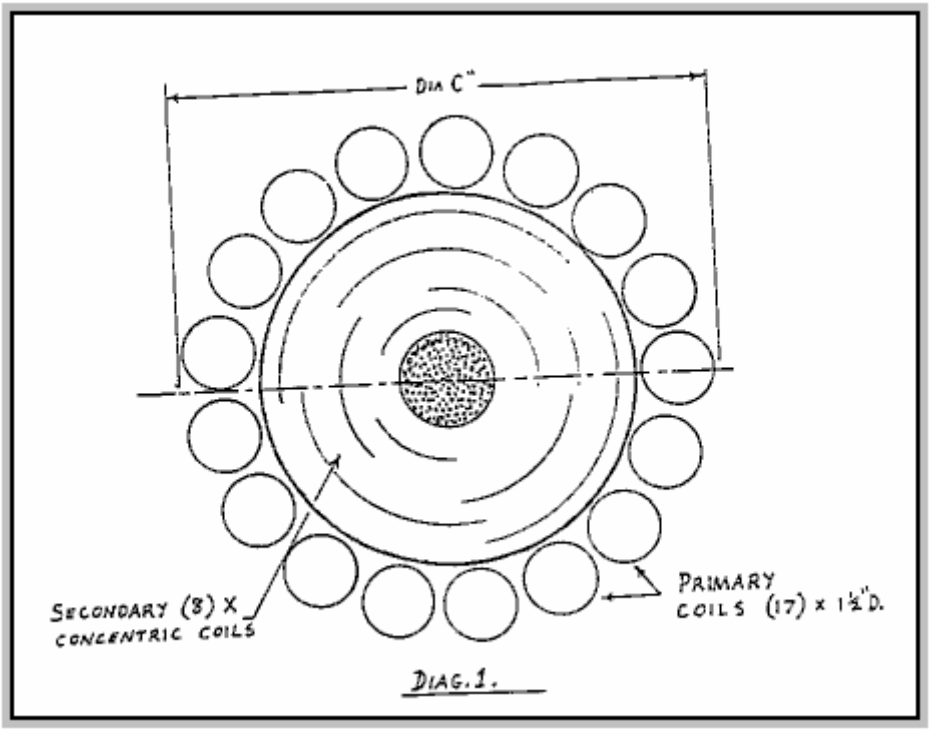
The Primary Coils

If the outer body of your secondary coil is eight inches in diameter, then you won't fit the recommended seventeen primary coils around its perimeter. If your primary coils are one and a half inches in diameter then these will fit nicely around the perimeter of an 8-inch diameter secondary coil. However, it is preferable to have larger primary coils as mentioned in Mr Cater's opening comments, so it may be advisable to stick to the recommended 2-inch diameter size for the primary coils, but settle for one less and use only 16 primary coils.

Experimentation will decide which is the best way to go. For the purposes of this article I will refer to 2-inch diameter coils.

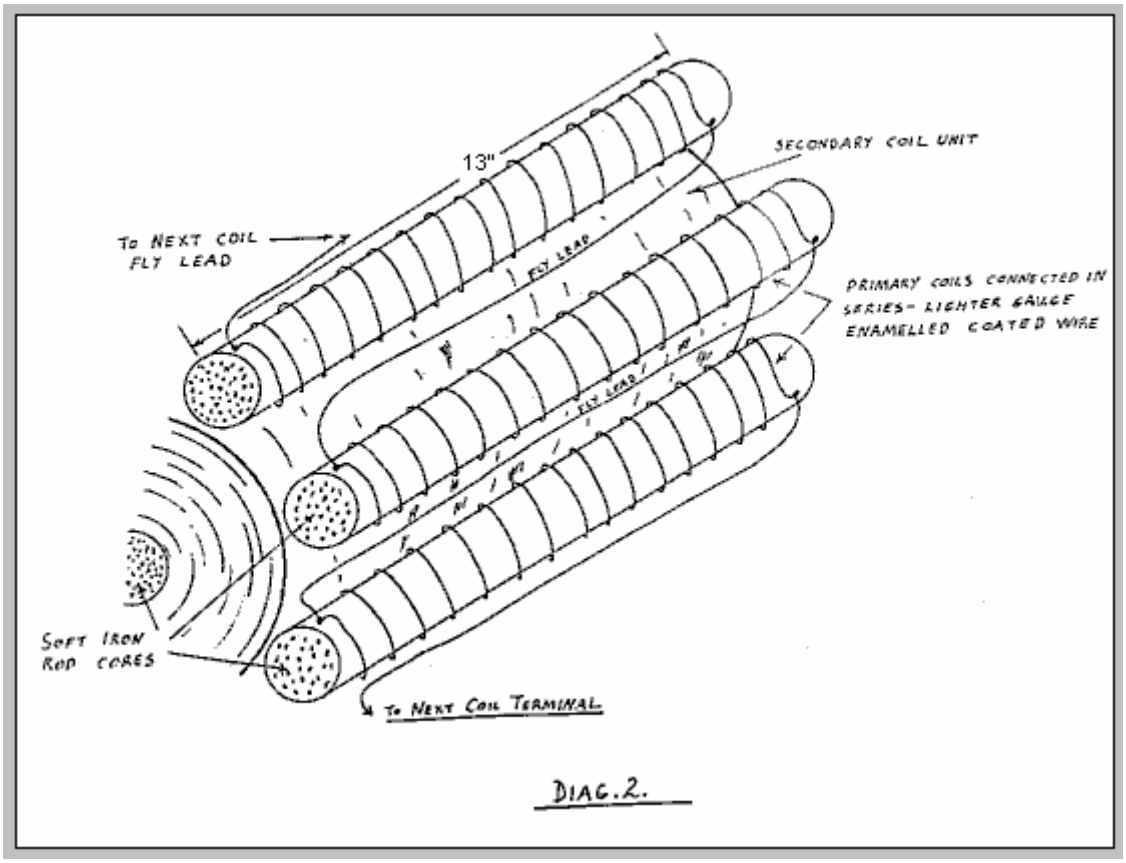
Cut medium gauge soft iron rods (oxy-welding rods will do) to 13-inch lengths. Be sure to de-burr the cut rods so that a compact fit is achieved.

Next, wind each coil separately with one terminal at each end (no gap 'G' is required for the primary coils). Then the primary coils are physically mounted around the large secondary coil - refer to Diagram 1.



DIAG. 1.

The primary coils are then interconnected with suitable leads of the same gauge as the coil wire to form a series coil configuration. Refer to Diagram 2.



DIAG. 2.

All coils must be wound in an identical manner so that the current in each one travels in a clockwise or counter clockwise direction. It is essential that the current flows in the same direction.

The Secondary Coil: Construction Notes

The secondary coil consists of a number of concentric cylinders and coils of three varying types repeated in a special sequence as detailed here.

1. You begin with the soft iron core in the same way as the primary iron cores were constructed. Use two inch diameter (2" OD) thin-walled PVC tubing cut to thirteen inches (13") in length, and packed with soft iron rods (oxyacetylene welding rods will do).

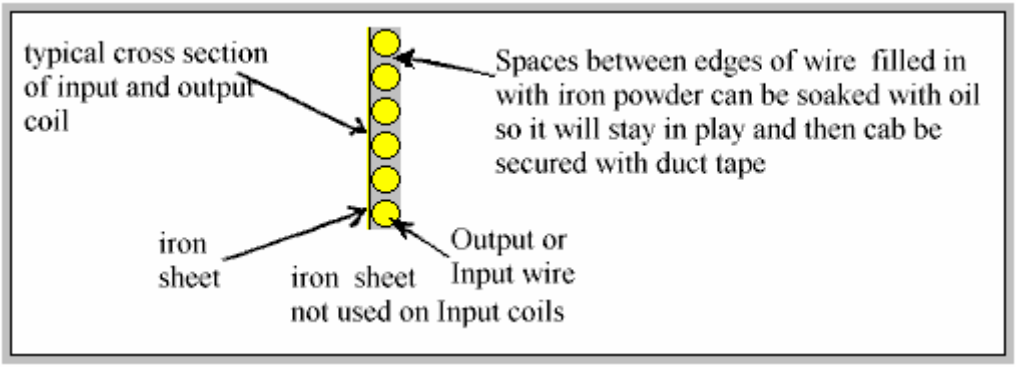
2. Around the central PVC tubing wrap the gold-plated iron sheeting so that the gold is facing outwards. The iron sheeting needs to be in the range 0.010" to 0.015" in thickness. The iron sheeting should be as thin as possible as you want to get the most powerful fluctuating magnetic field possible, induced as close to the wire as can be physically and electrically achieved.

This is the reason for the oil-soaked iron powder. The purpose of the oil is, of course, to make the iron powder physically manageable. The thinner the iron sheeting the more completely magnetised it will be. The gold plating is only the frosting on the cake so to speak. It certainly does not need to be very thick and no, you don't have to pay thousands of dollars for gold plating. A simple chemical process is used. Ask your local electroplater for a lead in the right direction. As to the suppliers of the iron sheeting, you certainly wont find it down at your local hardware store as it is a rather a specialised item. Try transformer manufacturers or electric motor and generator suppliers.

You will need eight (8) concentric iron cylinders. Each one will be thirteen inches (13") wide with varying lengths depending on the circumference of each concentric layer. Allow a quarter inch over the circumference length to give a small overlap. You will need to devise a method for keeping the iron sheeting in position ready for the next stage of construction. Several spots of super glue should do the job nicely.

3. Now that you have wrapped your first iron layer around the central PVC tube containing the soft iron core, you are now ready to wind your first secondary coil. Use a heavy gauge enamel coated wire somewhere near the gauge of house wiring. If this is not available, then insulated single core wire will do. As with all the coils that

are to be wound, whether primary or secondary, only one layer of wire is wound. When you are winding the secondary coil leave a small space between each turn. Refer to Diagram 3.



The gap 'G' reduces the inertia of flowing electrons as well as providing room for the oil-soaked iron powder which is to be packed between each winding. Perhaps 1 mm to 1.5 mm would be a sufficient gap between adjacent turns of the winding. However, before packing each coil with the iron powder, it would be advisable to lacquer the coil winding to seal it in position on the iron sheeting. This also provides extra insulative protection. The purpose of the non-metallic concentric spacers within the secondary coil serves two purposes:

a. To minimise the cancellation effects.

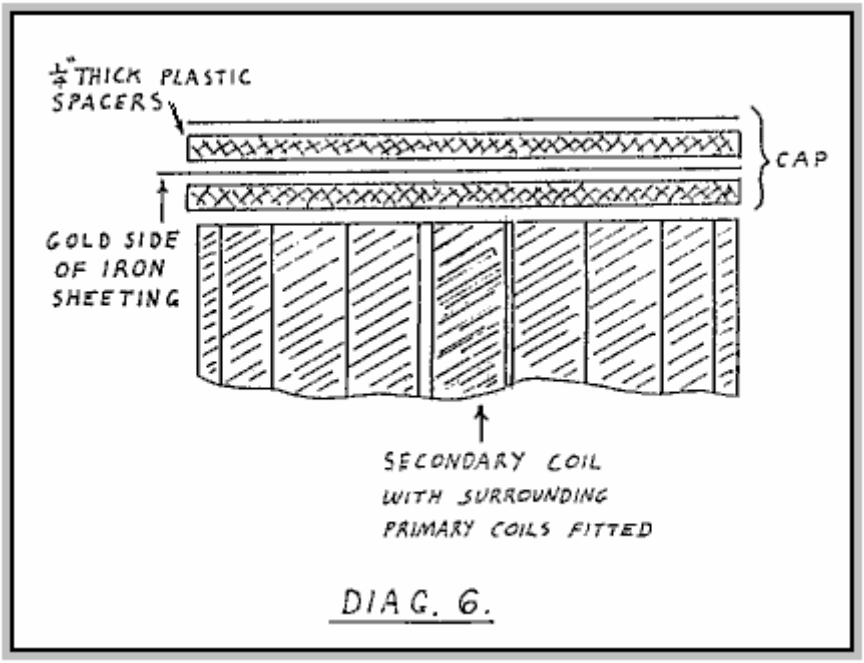
b. To produce an Orgone accumulator effect.

The material used could be heavy-duty PVC tubing with quarter-inch thick walls or quarter-inch thick sheeting, possibly heat treated, wrapped around the coils. You may be lucky for one or two of the concentric rings required, and have a piece of PVC tubing which is just the right diameter. For the remaining diameters you could reduce the circumference of a larger piece of tubing, thus converting it to the desired diameter. Be sure that the butt joint is perfect or that any gaps in the join are filled in with a suitable plastic filler. Some innovation and ingenuity may be required for this part of the construction. The general strategy for building this multi-layered secondary coil is to build it by winding each coil on separate concentric cylinders consisting of the gold-plated iron sheeting wrapped around the non-metallic spacer. The inner diameter of one cylinder will be the outer diameter of another. They are then joined together one inside the other. Fly wires are then used to interconnect the ends of each coil. For initial experiments this may be done in several ways, two of which are recommended by Mr. Cater:

1. Each concentric coil may be connected in series so that the current will flow in the same direction, either clockwise or counter clock wise as if it is one continuous coil.

or

2. Each adjacent pair of coils is wired so that the current flows in the opposite direction to the adjacent pair of coils. In other words, the first two adjacent coils are connected in the clockwise direction, and then the next pair of adjacent coils is connected counter clock wise. The third pair will be clockwise and the fourth pair counter clockwise. Changing the wiring configuration can be achieved quite simply by rearranging the external fly leads which are used to interconnect each of the secondary coils.



DIAG. 6.

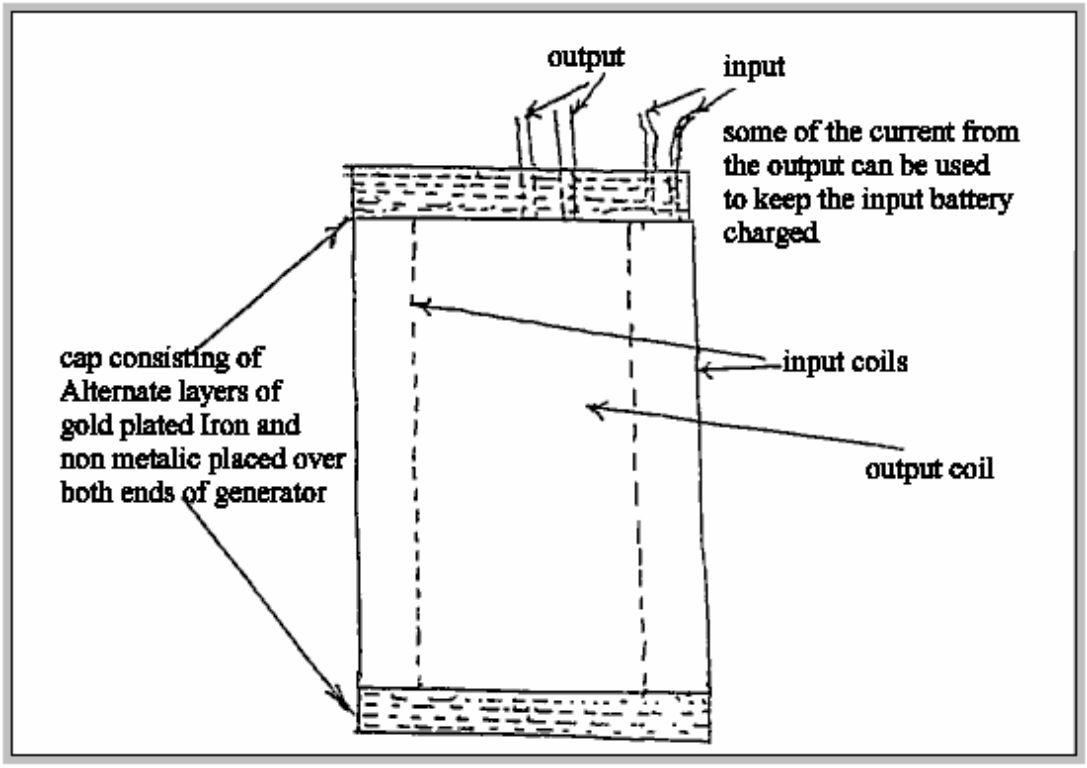
The leads should take the shortest path around the outer face of the secondary coil and of course they should be of the same gauge as the actual coil winding itself. Refer to Diagram 4

The Side Caps

Now that you have completed the secondary coil and wound the primary coils, the next step is to cut the caps to their correct size so that their diameter will be big enough to cover in the entire primary and secondary coil assembly. Refer to Diagram 1 above where the required dimension is marked as "Dia. C"

1. Cut eight pieces of quarter-inch thick plastic sheeting to the diameter "Dia. C" dimension, 4 per cap, so 8 in total.
2. Cut eight pieces of gold-plated iron sheeting in the same manner.
3. Glue together the plastic and iron sheeting as illustrated in the expanded drawing Diagram 6.

Devise a method to attach the caps to the sides of the unit and a means of positioning the outer primary coils so that they are all held in their correct positions. Bear in mind that powerful magnetic forces will be present and that the unit itself will be quite heavy, so a strong form construction is needed. One suggestion is to use dowels to hold the caps in position and use suitably shaped plastic spacers to position and hold the primary coils in place. Once the caps have been fitted, the generator becomes a highly potent orgone accumulator.



Gold-plated iron is many, many times more effective than any other metallic material. The accumulator effect greatly increases the effectiveness of the generator.

Testing

Now that you have actually completed all the construction work, you now need a suitable input driver unit which should have been thoroughly tested and ready for driving the unit. Let's be optimistic and hook up a good size load for the secondary, a couple of radiator bars (electric heaters) should do to begin with. Across the output terminals you can connect all the usual test gear.

Summary

The construction of the secondary coils may be carried out by completing the following steps:

1. Fill a thin-walled PVC tube of 2-inch diameter and 13-inches long, with soft iron rods.
2. Wrap the PVC tubing with the iron sheeting cut to 13" size with a 1/4" overlap along the tube, flush with the ends. Ensure that the **gold side** is facing outwards.
3. Wind the single-layer heavy-gauge coil with a suitable spacing between each turn of the winding and attach suitable terminals at each end of the wire.
4. Coat the coil winding with lacquer, sealing it in position.
5. Pack between each turn of the coil windings with oil-impregnated iron powder.
6. Wrap the coil and iron powder with ducting tape.
7. Fit the quarter-inch thick non-conductive spacer as described above.
8. Repeat step 2 to step 7, eight times and finish off by fitting an outer casing of the quarter-inch thick non-conducting material.

This Article first saw the light of day several years ago and it is believed, was first published in the Australian Free-Energy Newsletter called "Tuning In".

Another source comments on the Cater device as follows:

A self-sustaining electric generator was demonstrated at Seattle, Washington in 1919 by an inventor named Hubbard. His invention was supposedly 14 inches long and 11 inches in diameter. It powered a 35 horsepower electric motor which pushed a boat continuously around the bay for several hours. This demonstration was witnessed by thousands. During the time of his demonstrations, Hubbard made a sketch of one of his smaller generators used to power ordinary electrical appliances shown in Fig. 28:

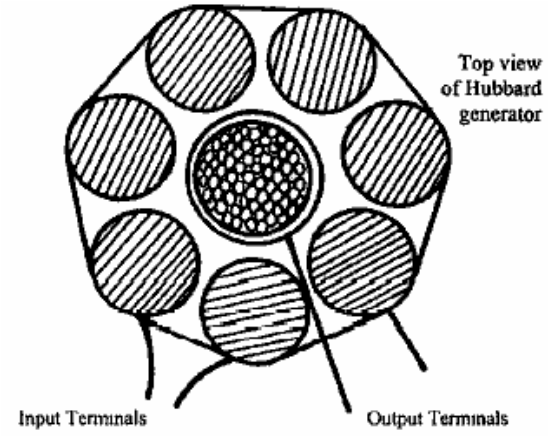


Fig. 28

It was approximately six inches long and about five inches in diameter. It consisted of eight coils in series, wound on iron cores which in turn surrounded a slightly larger central coil. The central coil was wound on a hollow tube which contained many small rods of soft iron. Four terminals extended from the unit, two connecting to the outer coils which received the input current, while the other two came from the central coil.

It is highly significant that both wires used in the generator appeared to be of heavy gauge like those used in power lines with the same kind of insulation. Each coil had only one layer of this wire which means that only a moderate number of turns were used in the entire generator. It is known that the generator produced a fluctuating current of an undisclosed frequency and had no moving parts.

The basic principle on which the generator operated is apparent. A small current passed through a coil with a moderate number of turns per unit length will magnetise an iron core to a surprising degree. This principle is utilised to great advantage in electromagnets. What apparently hasn't been realised is that during the brief interval in which the current builds up after it is turned on, an induced EMF (voltage) is produced in the coil by the changing magnetic flux, which is in the same direction as the current. This induced EMF is the result of the magnetic field produced by the magnetisation of the iron core. If this induced EMF were in the opposite direction to the current, then a sizeable current could never be produced in the coil as the EMF opposing the current would automatically cancel it before it could increase.

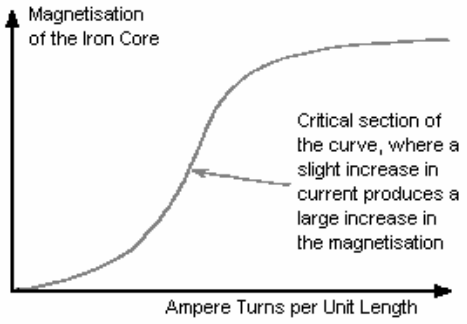


Fig. 29

Fig. 29 shows a graph of the magnetisation of an iron core plotted against ampere turns per unit length. The term "ampere turns" is the number of turns of the coil per unit length multiplied by the number of amps of current flowing through the coil. For example, a current of 1 amp flowing through a coil of 100 turns will produce the same effect as 2 amps flowing through a coil of the same length which has only 50 turns.

There is a section on the curve where a slight increase in ampere turns will produce a tremendous in magnetisation of the iron core. The cause of this phenomenon should be analysed. It seems strange that just a few ampere-turns can produce extensive and significant magnetisation of the iron core. Yet, the observable magnetic field produced by the current without the magnetic core is tiny by comparison. A similar field produced by a permanent magnet, would be unable to induce a noticeable magnetisation of the iron. This is something which conventional science has found convenient to ignore.

If an alternating current is passed through an electromagnet and the ampere-turns exceed a critical point, a chain reaction takes place in the coil, producing a tremendous increase of current in the coil. This is responsible for transformers which occasionally burn out during current surges. In some cases the sudden increase in current is sufficient to push the ampere-turns value into this critical range. The chain reaction results from an increase in the magnetisation of the iron which produces an increase in the current, which then produces an additional large increase in magnetisation, and so on until the iron reaches its maximum degree of magnetisation.

This process occurs during the first half of the AC cycle. The EMF is flowing in the direction opposite to that of the current after it reaches its maximum value and the second part of the cycle begins. This EMF, which is the same magnitude as that which brought the current to its maximum value during the first part of the cycle, now acts as a brake and stops the current. The applied alternating EMF then starts the current in the opposite direction and the identical process occurs again with the current flowing in the opposite direction.

Normal working transformers have ampere-turns which are well below this critical point. The additional EMF induced in the coils by the magnetisation of the iron offsets the natural inductive impedance of the coils. This is why transformers have such a high degree of efficiency. If any material other than iron or special steel were used for the core, the efficiency would drop significantly.

A normal square-wave pulsed current cannot be used in such a device due to the very short time of the rise and fall of the applied voltage, so a sine wave power supply is needed to produce this effect. Since the induced EMF in a coil is directly proportional to the rate of change of magnetic flux, it follows that the higher the frequency of this sine wave supply, the better.

There is possibly another factor which could contribute to the success of the Hubbard device. At that time, the only insulated wire available had thick and heavy insulation. This means that adjacent turns of wire in the coil were separated by a distance equal to twice the thickness of the insulation. Consequently, the gap resulted in a cancellation of magnetic effects produced by electrons flowing in the wire. Since inertia is dependent on the ability to generate a magnetic field, the inertial properties of the electrons would be almost nullified.

Floyd Sweet's "VTA" Self-Powered Generator.

Another device in this category of pulsed devices which tap external energy was produced by Floyd ("Sparky") Sweet. The device was called "Vacuum Triode Amplifier" or "VTA" by Tom Bearden. There is very little practical information available on this device, though there is a video of it in operation on the web, with an input power of just 0.31 milliwatt and a continuous power output of more than 500 watts (112 volts AC at 60 Hz) which is a COP of more than 1,612,000 which is spectacularly impressive.

The device was capable of producing more than 1 kW of output power at 120 Volts, 60 Hz and can be connected so as to be self-powered. The output is energy which resembles electricity in that it powers motors, lamps, etc. but as the power increases through any load there is a temperature drop instead of the expected temperature rise, which is why it is called "cold" electricity.

When it became known that he had produced the device he became the target of serious threats, some of which were delivered face-to-face in broad daylight. It is quite possible that the concern was due to the device tapping zero-point energy, which when done at high currents opens a whole new can of worms. One of the observed characteristics of the device was that when the current was increased, the measured weight of the apparatus reduced by about a pound. While this is hardly new, it suggests that space/time was being warped. The German scientists at the end of WWII had been experimenting with this (and killing off the unfortunate people who were used to test the system) - if you have considerable perseverance, you can read up on this in Nick Cook's inexpensive book "The Hunt for Zero-Point" ISBN 0099414988.

Floyd found that the weight of his device reduced in proportion to the amount of energy being produced. But he found that if the load was increased enough, a point was suddenly reached where a loud sound like a whirlwind was produced, although there was no movement of the air. The sound was heard by his wife Rose who was in another room of their apartment and by others outside the apartment. Floyd did not increase the load further

(which is just as well as he would probably have received a fatal dose of radiation if he had) and did not repeat the test. In my opinion, this is a potentially dangerous device. It should be noted that a highly lethal 20,000 Volts is used to 'condition' the magnets and the principles of operation are not understood at this time. Also, there is insufficient information to hand to provide realistic advice on practical construction details.

On one occasion, Floyd accidentally short-circuited the output wires. There was a bright flash and the wires became covered with frost. It was noted that when the output load was over 1 kW, the magnets and coils powering the device became colder, reaching a temperature of 20 degrees Fahrenheit below room temperature. On one occasion, Floyd received a shock from the apparatus with the current flowing between the thumb and the small finger of one hand. The result was an injury akin to frostbite, causing him considerable pain for at least two weeks.

Observed characteristics of the device include:

1. The output voltage does not change when the output power is increased from 100W to 1 kW.
2. The device needs a continuous load of at least 25W.
3. The output falls in the early hours of the morning but recovers later on without any intervention.
4. A local earthquake can stop the device operating.
5. The device can be started in self-powered mode by briefly applying 9 Volts to the drive coils.
6. The device can be stopped by momentary interruption of the power to the power coils.
7. Conventional instruments operate normally up to an output of 1 kW but stop working above that output level, with their readings showing zero or some other spurious reading.

It appears that Floyd's device was comprised of one or two large ferrite permanent magnets (grade 8, size 150 mm x 100 mm x 25 mm) with coils wound in three planes mutually at right angles to each other (i.e. in the x, y and z axes). The magnetisation of the ferrite magnets is modified by suddenly applying 20,000 Volts from a bank of capacitors (510 Joules) or more to plates on each side of it while simultaneously driving a 1 Amp 60 Hz (or 50 Hz) alternating current through the energising coil. The alternating current should be at the frequency required for the output. The voltage pulse to the plates should be applied at the instant when the 'A' coil voltage reaches a peak. This needs to be initiated electronically.

It is said that the powering of the plates causes the magnetic material to resonate for a period of about fifteen minutes, and that the applied voltage in the energising coil modifies the positioning of the newly formed poles of the magnet so that it will in future, resonate at that frequency and voltage. It is important that the voltage applied to the energising coil in this 'conditioning' process be a perfect sinewave. Shock, or outside influence can destroy the 'conditioning' but it can be reinstated by repeating the conditioning process. It should be noted that the conditioning process may not be successful at the first attempt but repeating the process on the same magnet is usually successful. Once conditioning is completed, the capacitors are no longer needed. The device then only needs a few milliwatts of 60 Hz applied to the input coil to give up to 1.5 kW at 60 Hz at the output coil. The output coil can then supply the input coil indefinitely.

The conditioning process modifies the magnetisation of the ferrite slab. Before the process the North pole is on one face of the magnet and the South pole on the opposite face. After conditioning, the South pole does not stop at the mid point but extends to the outer edges of the North pole face, extending inwards from the edge by about 6 mm. Also, there is a magnetic 'bubble' created in the middle of the North pole face and the position of this 'bubble' moves when another magnet is brought near it.

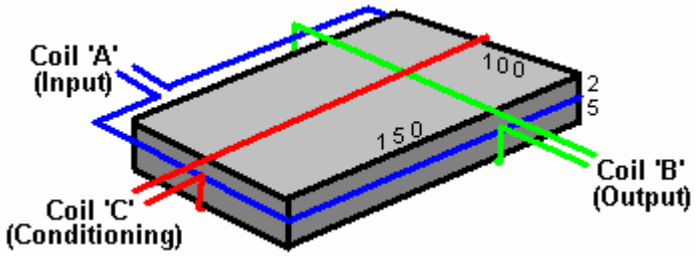
The conditioned slab has three coil windings:

1. The '**A**' coil is wound first around the outer perimeter, each turn being 150 + 100 + 150 + 100 = 500 mm long (plus a small amount caused by the thickness of the coil former material). It has about 600 turns of 28 AWG (0.3 mm) wire.

2. The '**B**' coil is wound across the 100 mm faces, so one turn is about 100 + 25 + 100 + 25 = 250 mm (plus a small amount for the former thickness and clearing coil 'A'). It has between 200 and 500 turns of 20 AWG (1 mm) wire.

3. The '**C**' coil is wound along the 150 mm face, so one turn is 150 + 25 + 150 + 25 = 350 mm (plus the former thickness, plus clearance for coil 'A' and coil 'B'). It has between 200 and 500 turns of 20 AWG (1 mm) wire and should match the resistance of coil 'B' as closely as possible.

Coil '**A**' is the input coil. Coil '**B**' is the output coil. Coil '**C**' is used for the conditioning and for the production of gravitational effects.



At time of writing, information and photographs of the original device can be found on the website: http://www.intalek.com/Index/Projects/Research/Construction%20of%20the%20Floyd%20Sweet's%20VTA%20by%20Michael%20Watson.htm where a paper by Michael Watson gives much practical information. For example, he states that an experimental set up which he made, had:
The '**A**' coil with a resistance of 70 ohms and an inductance of 63 mH,
The '**B**' coil, wound with 23 AWG wire with a resistance of 4.95 ohms and an inductance of 1.735 mH, and
The '**C**' coil, also wound with 23 AWG wire, with a resistance of 5.05 ohms and an inductance of 1.78 mH.
Chapter 3 has additional information on the VTA.

Rosemary Ainslie's COP=17 Heater.
Rosemary Ainslie has produced a pulsed heater system which has been measured at a performance of COP = 17. This is a recent design and as far as I am aware, has not yet been replicated by other people. Panacea-bocaf.org are working with Rosemary's original developers to produce an independent implementation of the heater. At this point in time, the heater has been built to a prototype testing scale for laboratory examination and measurement and not been produced in the kilowatt range, which, hopefully, will come at a later date.

Panacea have produced a 250-page document describing the research, the testing, the theory, etc. and that can be downloaded free using this link:
http://www.panaceatech.org/Rosemary%20Ainslie%20COP17%20Heater%20Technology.pdf

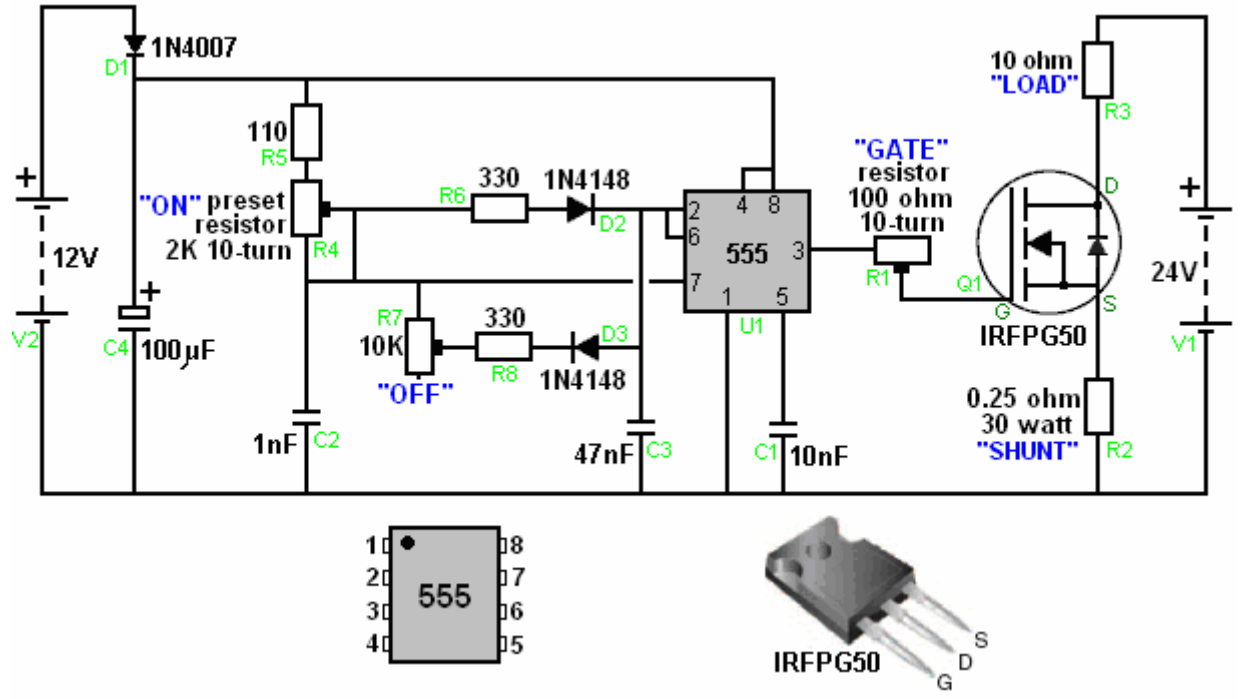
As that document contains the details which scientists need to see for serious testing and development, it may be a little technical for some people, so Panacea have produced a simplified version aimed at the average home-build investigator and that can be downloaded free using this link:
http://www.free-energy-info.tuks.nl/Ainslie_heater_circuit.pdf.

In very broad outline, the circuit produces the same very short, very sharp voltage pulses that are the basis for so many "free-energy" devices. The circuit used looks very simple but in spite of that, the way that it operates is not at all simple. The circuit is shown below and to a quick glance, it looks like a standard 555 timer chip circuit, used in many existing applications. However, if the circuit is operated as a 555 pulsing circuit, then the output is not COP>1.

Looking more closely, we notice that the link between the output of the 555 chip on pin 3 and the input gate pin of the Field-Effect Transistor, is unusual as it is not the usual voltage divider between pin 3 and the 0-volts ground line. Instead, the gate is directly coupled to the 555 chip output by a single, low-resistance preset resistor.

Normally, an NE555 chip struggles to reach 50,000 cycles per second and a large number of 555 chips on the market can't even operate at even that frequency. To get Rosemary's circuit into it's COP>1 operation, the resistor marked "**GATE**" is adjusted very slowly to find the point at which the circuit becomes unstable, over-rides the normal operation of the 555 chip and starts oscillating at the resonant frequency of the overall circuit, forcing the 555 chip to become a feedback component. The circuit then produces the sharp, short voltages spikes at more than ten times the operating speed of the 555 chip and pulsing the 10-ohm heating element marked "**LOAD**" at about 500,000 pulses per second.

That rate of operation is clearly well outside the possible performance of an NE555 chip, besides which, the timing elements of the chip should be producing a much lower frequency, as indeed it does before the "**GATE**" resistor adjustment causes the circuit to break out of its normal design-mode operation and start the high-speed spike generating, resonant performance. The circuit used is shown here:



As Panacea-bocaf are working to test and develop this circuit further, it would be a good idea to download their free documentation on the design and keep an eye on their progress in this field. The two documents give very considerable detail on the work which has already been done, and of course, you can yourself experiment with this circuit and see what results and adjustments you can discover yourself.

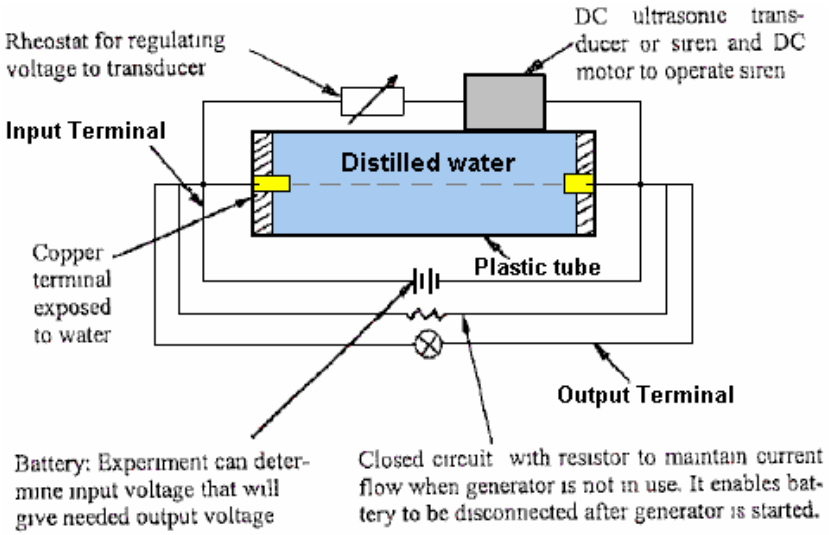
Joseph H. Cater

Joseph comments: The experiments of Schauberger and others have confirmed the enormous and almost unlimited quantities of electricity housed in water. The following is an absurdly simple and practical method of extracting this energy. It employs the “Coanda” or “cloud-buster” effect.

A plastic tube 14” to 16” (350 mm to 400 mm) long and about 2.5” (65 mm) in diameter is filled with distilled water. At each end, exposed to the water, is a copper terminal which is used for both the electrical input and output. Rechargeable dry cells of suitable voltage are connected in series with the input terminals. When the two output terminals are short-circuited or connected to a load, electricity starts flowing. This is current entrained by the input current. When high voltage is applied, the output voltage is almost as great as the input voltage. However, the amperage is inadequate. The answer to the problem is ultrasonics. It is an experimental fact that ultrasound of 600,000 Hz focussed on a container of water causes the water to boil. This means that sound of this frequency disintegrates large quantities of “soft” electrons in the water. The sudden release of “hard” electrons produces tremendous thermal agitation of the water molecules.

A DC ultrasonic transducer attached to the tube would produce sufficient free electrons to be entrained for the unit to have almost unlimited output potential. The tube functions like a sounding board. Mr Cater has been given powerful evidence that two different individuals who received this information got sensational results from the generator. They had access to such a transducer. They tried to set up in business but the vested interests saw to it that they were put out of business and persuaded to remain silent ever since.

An associate of Mr Cater built a fist-sized siren which generated a frequency of 600 kHz. When focussed on a small container of water, the water boiled. This demonstrated that it could be used instead of a solid-state DC ultrasonic transducer on the water generator. A small DC motor could operate the siren. It would be far more effective as it produces a much more intense sound. The construction is shown here:



Self-sustaining Electrical Generator Employing the Cloud-buster Effect

Dr Oleg Gritskevitch's 1.5 Megawatt Self-Powered Generator.

Dr Oleg V. Gritskevitch of Vladivostok in Russia, the holder of some seventy patents, designed and fully tested an electrical generator along the same lines as Joseph Cater's device mentioned above. It uses no fuel and has given a DC output of 220 volts at 6,800 amps (1.5 megawatts) for more than two years. As built by Dr Gritskevitch, this is not a home-builder's ideal project as massive electrical input is needed to get the device started, and his prototype weighs 900 kilograms (nearly 2,000 lbs). Details are given on the very good RexResearch web site: http://www.rexresearch.com/gritskevich/gritskevich.htm but in broad outline, the device is a toroidal pipe some two metres (6'-6") in diameter, coated on the inside with barium titinate and filled with ultra-pure distilled water mixed with 'heavy water'. Inside the toroid are electromagnetic coils and surrounding it, copper pipes carrying cooling water to keep the temperature down to 50 degrees Centigrade. Also inserted into the toroid at intervals around the circumference are electrical contacts.

The device is started by giving the water a massive high-voltage discharge of some 100,000 volts at 50 mA for three to five minutes. This power input gets the water ionised and circulating. The circulation is maintained by the electromagnetic coils and the power output is around COP=100.

Oleg died without ever getting funding for his design (a typical method of blocking free-energy devices from reaching the market). A more detailed description of the device and it's operation comes direct from Oleg:

This is a description of the construction and operation of Oleg V. Gritskevitch's hydro-magnetic dynamo, which is an example of a very powerful new energy system. The prototype in Armenia has averaged some 1,500 kilowatts of power over a period of several years.

Oleg was born on 14 August 1936 and grew up in Vladivostok, Russia. He married and has a son Boris. Gritskevitch was a physicist by education. He worked in the Far-East branch of the USSR Academy of Sciences. Since 1985 he worked independently as an inventor. He has more than 70 patents on inventions ranging from household engineering up to high technologies, which he has been trying to apply in our country although he

encountered major difficulties in this. After numerous attempts to obtain the patents, he became convinced that the information had become widely known. Therefore he received the state certificates of know-how (a French way of patenting), for all his inventions.

Introduction

During the 1999 Symposium of the Institute for New Energy, he lectured on his hydro-magnetic dynamo. This paper is his attempt to explain the construction and operation of his dynamo. To protect his secrets from investigators, he, on occasion, provided misleading information. For example, the drawing accompanying the Russian patent mentioned below, shows a cylinder across the toroid to mislead readers. The real dynamo has the toroid alone, without the cylinder. Even its name "hydro-magnetic dynamo" is somewhat deliberately misleading.

Oleg stated that he had some familiarity with the new energy field. Nearly all purported new energy devices are fairly small electrical generators. The dynamo may be the only new electrical generator which most nearly meets all the requirements of an ideal large-scale electrical generator. Oleg claimed that his dynamo really is the single most valuable invention the world has ever known.

Alexander V. Frolov of St. Petersburg recommended Oleg to contact Dr. Patrick Bailey of the Institute for New Energy since Patrick has lots of contacts who could possibly help with patenting his invention of a new source of energy in USA.

Oleg worked on the theory and creation of the electrostatic generator-converter the "Hydro-Magnetic Dynamo" for about 20 years. The first primitive equipment was created when Oleg worked in the Academy of Sciences. During that time, various changes were introduced in the generator and in the theory of how it works. As a result, it is now ready for manufacture, installation, and applications in industry.

Oleg made the first public report on his work in 1991 at a symposium in Volgodonsk city. His report received positive reactions and reviews from the experts in the nuclear industry in USSR. That same year, he was accepted into the International Nuclear Society. During those years he offered development of this technology to different state bodies and private enterprises. But the answer from everybody was "It is a very interesting and important project, but we cannot fund it.

Eventually, Oleg tried to transfer this technology to the USA through the embassy in Moscow. The former ambassador to the USSR, Dr. J. Matlock knows about it. He wanted to meet Oleg, but at that time there were forces opposed to the carrying out of his plans. So he started to look for other possible investors. He was ready to consider any offers of co-operation, joint patenting, sale of technological information, creation of a joint venture, etc. etc. Oleg was awarded some 70 Russian patents covering a wide range of important technical topics.

History

This project was the result of one article in the August 1972 issue of a popular Russian magazine *Tehnika Molodiozhi.* The article written by A. Kaldamasov was entitled *Ball Lightning in a Liquid*. The article came to the attention of Michail Razovsky and Oleg in 1974. Oleg's group of volunteers and enthusiasts was looking for a new source of energy and so this article served as a starting point for the understanding of chemical-physical processes occurring in water. During the period 1976 to 1978 one year was spent in the radiological lab of the Vladivostok city hospital, including Vladilen Bulgakov, radiology physician, and Michail Razovsky, theoretician in the plasma physics field and others, assembling a device, which was supposed to separate water into oxygen and hydrogen more efficiently. During the experiments, instead of the expected results, it produced electricity very efficiently! The input power during the experiment was one 800-watt water pump. The output was 1,400 watts (COP=1.75). This device was assembled using plastic pipes connected with hoses, where the water was circulating in a loop. This then led to the idea of creating the second device as a generator-toroid.

The second generator was assembled in the workshop of the Ocean Research Institute in Vladivostok (Director Academician Viktor Ilichov), and in the summer of 1990 it was transported to the testing station of the Ministry of Electronics Industry in Vladivostok. This lab was well equipped with all necessary sets of instruments. At the same time, patent papers were filed in the USSR State Committee of Inventions. In the spring of 1991 the State Commission, led by Yurii Lebedev, chairman of the Innovation Council and Chairman of the Russian Federation Council of Ministers, arrived in Vladivostok. This commission arrived to the town for two reasons: to recommend a financial request for manufacturing the dynamo; and to classify this energy source as a "Discovery". (document #14-451).

After the next change in Russian government the financing for the project was terminated. The first article on the dynamo was published in the Russian magazine (*Tehnika Molodyozhi* 1990, #3, March issue, Page 17, entitled

"*Innovator's Ideas*".

Several Armenian physicists, after reading that article, sent Oleg a letter asking to meet with him in Vladivostok for negotiations about the dynamo. They arrived in March 1991 and ran tests on the second generator, which was operational at that time. Oleg flew to Armenia, and work on the third generator started at the end of 1991. It was completed at the end of 1992. It was operating and producing energy until January 1997, when it was destroyed during the war. Some people were also killed and other people moved to the USA. This version of dynamo created an output which averaged 6,800 amperes at 220 volts DC (1.496 megawatts). It's input power was only approximately 1% of the output power.

Oleg was an invited speaker of the Meeting of the Alternative Energy Institute (Dr. Hal Fox) in Salt Lake City in August 1999. The official announcement about his speech had several mistakes (for example, the name of Armenia was changed to Romania).

Dynamo Theory

The **H**ydro-**M**agnetic **D**ynamo is a large-scale, emission-free electrical generator, which does not require external fuelling. The dynamo is capable of powering large transportation vehicles such as buses, trucks, ships, locomotives, and aeroplanes. Doubt remains about making dynamos compact enough to power cars.

While three experimental prototypes have been built with Russian and Armenian expertise and equipment, a fourth demonstration prototype needs to be built with more modern Western engineering expertise and equipment to verify the dynamo's performance claims and to further explore the dynamo's potential capabilities.

The claimed performance is as follows:

> Dynamos are scaleable from 100 kilowatts to 1,000 megawatts. One 1000-megawatt dynamo is about the size of a two-car garage. For comparison, Hoover Dam's 17 generators have a total capacity of 2,000 megawatts. A dynamo can reliably run continuously for 25 years or more with little or no maintenance, no external fuel source, and no pollution. If a dynamo's output is 1,000,000 watts, its total input power is approximately 10,000 watts and so the dynamo's energy efficiency is about 10,000%.

The source of the dynamo's huge electrical output is a nuclear reaction, which is not generally known to mainstream science. However, it is known that the dynamo produces alpha particles, which are helium nuclei, made from fused deuterium, an isotope of hydrogen with one proton and one neutron. The electrons missing from the helium nuclei are what seem to provide a copious 'sink' of electricity, and that is the secret of the dynamo's ability to generate an exceptionally large amount of electricity. It is also known that the dynamo uses high-density charge clusters. High-density charge clusters are thought by some theorists, to be the basis of plasma-injected transmutation of elements and the neutralisation of radioactive materials. Unlike hot fusion and fission reactors, the dynamo does not accumulate any radioactive components.

The result of the dynamo's processes is conversion of electrostatic fields to direct current. It should be noted that a clear understanding of terms like "Coulomb's conversion" and "liquid Van de Graff generator" is very important.

Schematically, the dynamo is an electrostatic transformer, or in other words an electrostatic voltage multiplier. One version of the dynamo uses lasers to start up. There were three dynamo prototypes built. The first two small experimental prototypes were built in Vladivostok, Russia. The third and last prototype generated electricity continuously, (except when turned off to incorporate improvements), from 1992 to January 1997 in Armenia.

As mentioned above, the Armenian prototype generated a direct current of 6,800 amperes at 220 volts which is about 1.5 megawatts. Minimum power output has been 500,000 watts, and maximum power output has been 2,500,000 watts during winter experiments due to better cooling. The Armenian prototype dynamo's toroid weighed 900 kilograms and had a diameter of approximately 2 meters. Cooling water is circulated through copper pipes wrapped around the toroid. The heat is expelled from the cooling water with a heat exchanger. The working temperature was typically 36 degrees Centigrade.

After a dynamo is assembled, the water is literally 'jump-started' (by discharging a large bank of capacitors) to get it circulating inside the toroid. The starting impulse pressure is as high as 400 atmospheres. The dynamo's controls are temporarily set to generating a modest amount of electricity sufficient to sustain itself, possibly even while being transported from the factory to the place of its future operation. The control circuits are simple as only sensors and a control computer are used. We do not need any technical-maintenance personnel.

For the Armenian prototype dynamo, two 10-Farad capacitor-batteries were used to provide the initial water motion (acceleration and excitation of water). The capacitors were 20 kilograms each, with diameters of 50 centimetres, and were borrowed from Russian military radar stations. Using a total of 20,000 Joules, 100,000

Volts at 0.05 Amperes of current were applied to the Armenian dynamo for 3 to 5 minutes to ionise and polarise the water, which then started the generation of electricity.

The reason for the very high voltage provided by the large Russian radar capacitors, when starting the generator, appears to be to polarise the crystals of barium titanate. One comparison is with the electronic ignition on a gas stove. Once the barium titanate crystals are polarised, the generator is running.

After these capacitors had been used to 'jump-start' the Armenian prototype dynamo, a bank of buffer batteries sustained continuous operation when water motion and ionising began. This battery bank contained 8 powerful 12-volt, 150-ampere lead batteries. The Armenian dynamo's sustaining input power was 14,400 watts. The nominal maximum output power was about 1,500,000 watts. On one occasion, the output current was accidentally increased to 40,000 amperes for almost a minute. Fortunately, the power was reduced to a safe level before the water started to boil. Internal coils (windings) provide the control of water velocity and therefore control dynamo power. The faster the water is moving, the more electricity the dynamo generates. Once the water stops circulating around the toroid, the dynamo must be 'jump-started' again to a minimum power level before it can sustain its electricity generation using it's own output power.

The following is a condensed summary, with some editing and additional commentary, of the "Description" of the dynamo's Russian patent IPC H 02 K 44/00 "*Method of Deriving Electrical Energy and Realisation of Gritskevich's MHD-Generator*":

The dynamo is a sealed polystyrene toroid filled with ultra-pure distilled water with heavy water (deuterium oxide) added. The movement of water inside the closed loop and the use of the unique properties of water as a polar liquid, cause a release of electrical energy as an outcome of a rupture of the hydrogen bonds. Additional electrical energy is drawn from nuclear reactions and micro-cavity processes. The liquid is ionised, polarised, and moving around the toroid at start-up time by a running magnetic field with the help of stimulating electromagnetic windings.

Electrostatic generator-transformer "Hydro-magnetic dynamo". ("GT HMD") works due to the process of amplification and maintenance of a stationary (oscillating in particular) electromagnetic field by hydrodynamic movements of the conductive medium. The stator (i.e. the toroid) is made of materials with a high dielectric permittivity. Liquid rotor is a recombined water ('pure' water with high-molecular compounds), which moves due to the high-voltage discharges and running electromagnetic field.

The main processes in GT HMD are:

A principle of Van-der-Graff's electrostatic generator, where the solid insulating tape was changed to the liquid one.

A perpetual washout of the surface electrons from the spacer layer takes place;

The Coulomb's transformations take place;

A single-turn low-frequency generator works as a coaxial turn with 4 resonance points and energy carrying substance inside it that has very high resonance properties;

The electrostatic breakdowns of cavitation-vacuum structures in water take place.

The polar liquid (pure water) consists of dipoles only, i.e. strictly oriented charged molecules. During the interaction of ionized pure water with the layer $BaTiO_3$ the electrostatic field of above 10 million volts/cm is formed. During this process the breakdown of physical vacuum takes place.

The electrostatic field, coupled with the action of the $BaTiO_3$ layer (if we apply electrical filed to $BaTiO_3$, then this layer creates the sound vibrations of about 25,000 Hz, this vibration helps to break down the water molecules) and facilitates the further break down of the molecular-atomic structures of water. Also, due to the perpetual electrostatic discharges, the breakdown of the cavitation-vacuum structures occurs and the cold fusion nuclear reaction continues. With this fusion the energy of 500 kJ/mole is liberated in a vacuum and an energy of 6 kJ/mole is liberated in water. Thus, new hydrogen bonds form in vacuum with the energy liberation of about 20 kJ/ mole. Due to this process the acceleration of ionisation of polar liquid takes place. In addition, the constant 'washing out' of the incomplete electron bindings from the layer of barium titanate occurs and free electrons form. Due to this process, the polar liquid transforms into an ordered flow of electrons and negative ions, which can be described very simply, as an ionic-electric current.

Work on construction of the experimental generator started in September 1991 in Armenia and came to an end in March 1992. The active working weight of the prototype of dynamo (torus + water) was about 900 Kg. The diameter of the torus was about 2 meters. The torus was made from impact-resistant optic polysterol. This torus consisted of two halves, which were turned on the merry-go-round machine.

The monocrystallic barium titanate $BaTiO_3$ was sprayed on the internal surface of torus, its dielectric permittivity was 6000. The thickness of the layer was about 1 micron. The water was purified to the specific resistance of 18,000,000 Ohm/cm.

As we mentioned above, to start the dynamo two condenser jars of 10 Farad each were used. The energy of a starting battery constituted 20,000 Joules, the voltage was 100,000 Volts and the current was 0.05 Ampere to provide the initial movement of water (acceleration and disturbance).

The electrodes were made of metal tubes with diameter of about 5 mm. The dynamo is started using these electrodes. A total of 32 of these electrodes were installed evenly spaced around the circumference of the toroid. The toroid's cooling system formed a closed circuit of copper piping with purified water circulated through it. The copper tubes used in this system were covered with glass insulation. They also were the turns of load winding. The temperature of toroid was maintained not higher than 50 Celsius degrees.

An average output power was 220 volts x 6,800 ampere = 1,490 kilowatts. The current was DC. Periodically the power could be increased to 2,500 Kw when sufficient cooling of the generator could be provided. The additional power was drawn from four resonant windings. This alternating current, after rectification, was used to charge the back-up battery. Thus, the total output power constituted more than 1,500,000 watts. The low-frequency voltage was obtained from the load windings and the direct current was obtained from the stabilisation chamber.

It should be noted that the high-voltage discharges of the 32 electrodes, ionise the partially pre-ionised water further. By means of the stimulation windings, a circulating magnetic field is created which moves the water in one direction inside the toroid. An electromotive force is created by the electromagnetic induction in a separate set of windings. As we already mentioned, during the movement of the water stream, free electrons are created, and an additional energy gets emitted because of the water's friction against the coating layer on the inside surface of the toroid, because of electrostatic breakdowns of cavity-vacuum structures, and because of the ongoing nuclear reaction.

If the dynamo's output is 1,000,000 watts, it's total input power is approximately 10,000 watts. So therefore the dynamo's energy efficiency is about 10,000%.

In addition to the barium titanate deposited on the teflon-coated inner surface of the polystyrene toroid, the water itself also contains tiny barium titanate crystals which are suspended in the water. Ultrasound at 25,000 cycles per second is propagated through the water to form micro-bubbles on the surfaces of the suspended barium titanate crystals. Again due to the barium titanate's piezoelectric action, very high electrostatic fields are also developed within the micro-bubbles at the surface of the crystals. The electrons from the nuclear reaction are added to the electrons generated at the toroid's interior surface. The total amount of mono-crystalline barium titanate in the Armenian dynamo was nearly 1000 grams. Satellites, locomotives, heavy trucks, airplanes, and ships are obvious transportation applications.

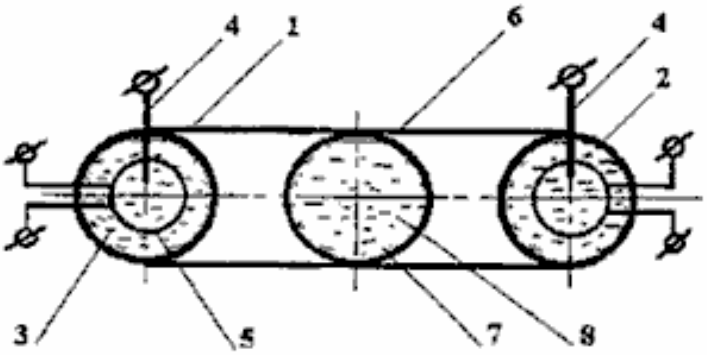
Dynamo Economics

The dynamo's production cost is estimated at $500 per kilowatt which is very competitive when compared to nuclear power's capital costs of $5,000 per kilowatt, windmill capital costs of $4,000 per kilowatt, etc. A well-run nuclear power plant can generate power for 1.5 cents per kilowatt-hour, coal 1.8 cents, natural gas 3.4 cents, and oil 4.1 cents, on average. The dynamo's operating cost would be approximately 0.1 cent per kilowatt-hour with no external fuel needed and without any pollution being created.

These dynamos could replace all nuclear power plants, solar installations, wood-burning furnaces, hydro-electric generation, etc. A recent IEEE *Spectrum* article stated that the world's demand for electricity increases by approximately 500 megawatts every day. To put this in perspective, that is the equivalent of building another Hoover Dam every four days to keep up with the world's increasing electricity demand. Or, a dynamo manufacturing company would have to build another 500-megawatt dynamo every single day to keep up with world electricity increased demand (in addition to replace all existing generators fuelled by hydro, nuclear, and fossil fuels.)

The text of the patent application mentioned above is not in English although the abstract of the patent number WO 01/15305 A1 has been translated into English:

WO 01/15305 A1



(57) **Abstract:** The invention is directed at increasing output, reliability and environmental safety of MHD generators as well as at simplifying the design of said generators. The inventive method for the production of energy comprises the following steps: a polar liquid (8) is circulated in a predetermined direction along a hermetically sealed toroidal channel (1) by means of a travelling magnetic field, and electric power is collected by means of electromagnetic winding. The liquid is ionized at least at the stage of launching, by means of electrodes (4), for example. The internal walls (2) of the channel have a dielectric constant which is higher than the dielectric constant of said liquid.

Patrick Kelly
www.free-energy-info.com
www.free-energy-info.co.uk
www.free-energy-info.tuks.nl
www.free-energy-devices.com

Chapter 6: Pulse-Charging Battery Systems

Note: If you are not at all familiar with basic electronics, you might find it easier to understand this chapter if you read chapter 12 first.

It is possible to draw substantial amounts of energy from the local environment and use that energy to charge batteries. Not only that, but when this method of charging is used, the batteries gradually get conditioned to this form of non-conventional energy and their capacity for doing work increases. In addition, about 50% of vehicle batteries abandoned as being incapable of holding their charge any longer, will respond to this type of charging and revive fully. This means that a battery bank can be created for almost no cost.

However, while this economic angle is very attractive, the practical aspect of using batteries for any significant home application is just not practical. Firstly, lead-acid batteries tend to get acid all over the place when repeatedly charged, and this is not suited to most home locations. Secondly, it is recommended that batteries are not discharged more rapidly than a twenty hour period. This means that a battery rated at a capacity of 80 Amp-hours (AHr) should not be required to supply a current of more than 4 amps. This is a devastating restriction which pushes battery operation into the non-practical category, except for very minor loads like lights, TVs, DVD recorders and similar equipment with minimal power requirements.

The main costs of running a home are those of heating/cooling the premises and operating equipment like a washing machine. These items have a minimum load capacity of just over 2 kW. It makes no difference to the power requirement if you use a 12-volt, 24-volt or 48-volt battery bank. No matter which arrangement is chosen, the number of batteries needed to provide any given power requirement is the same. The higher voltage banks can have smaller diameter wiring as the current is lower, but the power requirement remains the same.

So, to provide a 2 kW load with power, requires a total current from 12-volt batteries of 2000 / 12 = 167 amps. Using 80 AHr batteries this is 42 batteries. Unfortunately, the charging circuits described below, will not charge a battery which is powering a load. This means that for a requirement like heating, which is a day and night requirement, there needs to be two of these battery banks, which takes us to 84 batteries. This is only for a minimal 2 kW loading, which means that if this is being used for heating, it is not possible to operate the washing machine unless the heating is turned off. So, allowing for some extra loading like this, the battery count reaches, perhaps, 126. Ignoring the cost, and assuming that you can find some way to get over the acid problem, the sheer physical volume of this number of batteries is just not realistic for domestic installation and use. In passing, you would also need two inverters with a 2500 watt capability

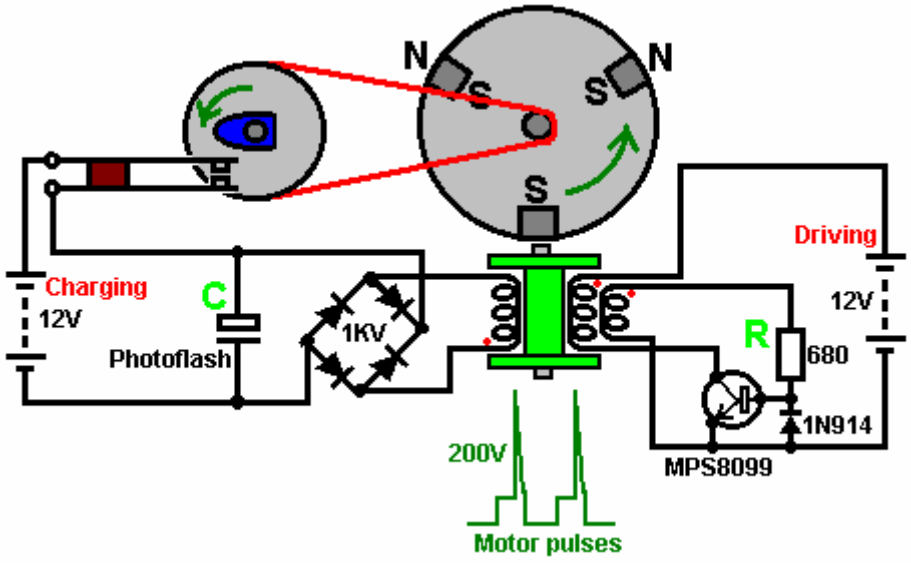
The recent charging system shown by ‘UFOpolitics’ in chapter 3, provides a very good and simple charging method which uses cold electricity. This can overcome the previous constraints imposed by using batteries, probably both with regards to current draw and with regards to recharging time. The Electrodyne Corp. staff who experimented extensively with the Tesla Switch circuitry, found that when a battery was fully conditioned to used cold electricity, that a battery could be disconnected, discharged independently to it’s full capacity, and then re-charged completely in under one minute. That style of operation completely overcomes the objections to using battery banks to power household equipment of any power.

Battery banks are used to power standard inverters which can look like this:

The battery connects at the back, using very thick wires, and one or more mains sockets on the front provide a power supply similar to the mains, matching it in both voltage and frequency. There is one variety of inverter called a "True Sine-Wave" inverter and costing much more than the ordinary non-sinewave inverters. Most equipment works well on the ordinary variety. It is usually the power available from the battery bank which is the limiting factor, combined with the long time taken to recharge the battery bank after use.

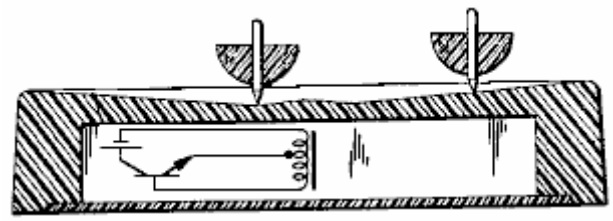
John Bedini's Battery-Charging System.

John Bedini has designed a whole series of pulse-generator circuits, all based on the 1:1 multi-strand choke coil component disclosed in his patent US 6,545,444



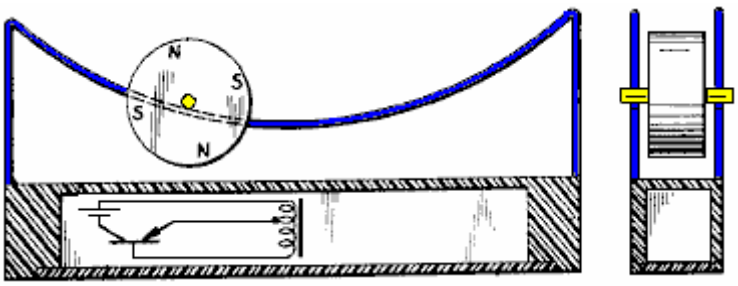
Roger Andrews' Switching System.

The very neat switching arrangement used by John is shown in detail in the earlier patent US 3,783,550 issued in 1974 where the same magnet-triggered boosting electromagnet pulse is used to power a whole series of movements. One of these is two magnetic spinning tops made to spin in a shallow dish:



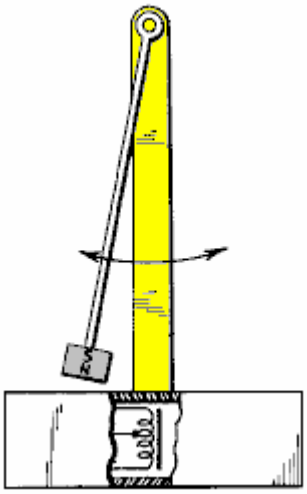
When the tops spin fast, they rise up the sloping base of the dish and spin near the outer edge. When they slow down they move back towards the centre of the dish and that triggers the battery/transistor/electromagnet built into the base of the dish. The pulse from the electromagnet boosts the spin of the top, sending it back up the slope. This is a very neat arrangement as the transistor is off most of the time and yet the two tops keep spinning.

Another of Roger's systems is shown here:

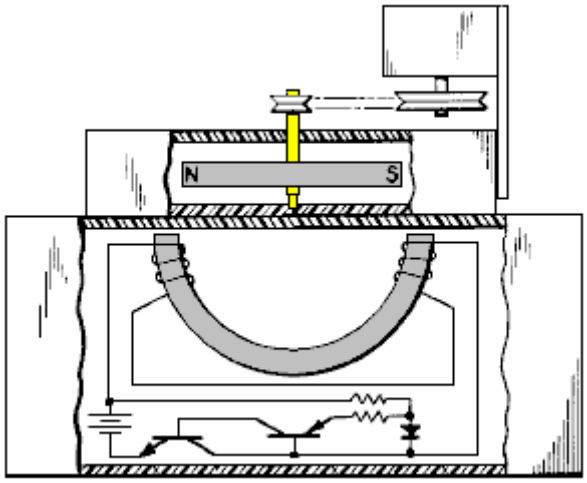


It operates in almost the same way, with a magnetic wheel rolling backwards and forwards along a curved track. At the lowest point, the electromagnet is triggered by the induced voltage in some of the turns of the coil, powering the transistor and boosting the magnetic roller on it's way.

Another Andrews device is the pendulum where the passing magnet of the pendulum triggers a boosting pulse from the solenoid, keeping the pendulum swinging. John Bedini has also used this mechanism for a pulsed battery charging system and Veljko Milkovic has demonstrated that substantial mechanical power can be extracted from a lever which is powered by a pendulum.



Andrews also shows a switching arrangement for a motor. This design is essentially the same as used by John Bedini in many of his pulsing systems:



Here, as the rotor magnet passes the curved electromagnet in the base, it switches on the two transistors which produce a pulse which keeps the rotor spinning and the tiny generator turning. Andrews produced this for amusement as the rotor appears to spin on it's own without any drive power.

As with the Andrews system, the Bedini rotor is started spinning by hand. As a magnet passes the triple-wound "tri-filar" coil, it induces a voltage in all three coil windings. The magnet on the rotor is effectively contributing energy to the circuit as it passes the coil. One winding feeds a current to the base of the transistor via the resistor 'R'. This switches the transistor hard on, driving a strong current pulse from the battery through the second coil winding, creating a 'North' pole at the top of the coil, boosting the rotor on its way. As only a changing magnetic field generate a voltage in a coil winding, the steady transistor current through coil two is unable to sustain the transistor base current through coil one and the transistor switches off again.

The cutting of the current through the coil causes the voltage across the coils to overshoot by a major amount, moving outside the battery rail by a serious voltage. The diode protects the transistor by preventing the base voltage being taken below -0.7 volts. The third coil, shown on the left, picks up all of these pulses and rectifies

them via a bridge of 1000V rated diodes. The resulting pulsing DC current is passed to the capacitor, which is one from a disposable camera, as these are built for high voltages and very rapid discharges. The voltage on the capacitor builds up rapidly and after several pulses, the stored energy in it is discharged into the "Charging" battery via the mechanical switch contacts. The drive band to the wheel with the cam on it, provides a mechanical gearing down so that there are several charging pulses between successive closings of the contacts. The three coil windings are placed on the spool at the same time and comprise 450 turns of the three wires (mark the starting ends before winding the coil).

The operation of this device is a little unusual. The rotor is started off by hand and it progressively gains speed until its maximum rate is reached. The amount of energy passed to the coil windings by each magnet on the rotor stays the same, but the faster the rotor moves, the shorter the interval of time in which the energy is transferred. The energy input per second, received from the permanent magnets, increases with the increased speed.

If the rotation is fast enough, the operation changes. Up to now, the current taken from the 'Driving' battery has been increasing with the increasing speed, but now the driving current starts to drop although the speed continues to increase. The reason for this is that the increased speed has caused the permanent magnet to move past the coil before the coil is pulsed. This means that the coil pulse no longer has just to push against the 'North' pole of the magnet, but in addition it attracts the 'South' pole of the next magnet on the rotor, which keeps the rotor going and increases the magnetic effect of the coil pulse. John states that the mechanical efficiency of these devices is always below 100% efficient, but having said that, it is possible to get results of COP = 11. Many people who build these devices never manage to get COP>1.

It is important that a standard mains powered battery charger is never used to charge these batteries. It is clear that the 'cold electricity' produced by a properly tuned Bedini device is substantially different to normal electricity although they can both perform the same tasks when powering electrical equipment. When starting to charge a lead-acid battery with radiant energy for the first time, it is recommended that the battery is first discharged to at least 1.7 volts per cell, which is about 10 volts for a 12 volts battery.

It is important to use the transistors specified in any of John's diagrams, rather than transistors which are listed as equivalents. Many of the designs utilise the badly named "negative resistance" characteristics of transistors. These semiconductors do not exhibit any form of negative resistance, but instead, show reduced positive resistance with increasing current, over part of their operating range.

It has been said that the use of "Litz" wire can increase the output of this device by anything up to 300%. Litz wire is the technique of taking three or more strands of wire and twisting them together. This is done with the wires stretched out side by side, by taking a length of say, three feet, and rotating the mid point of the bundle of wires for several turns in one direction. This produces clockwise twists for half the length and counter-clockwise twists for the remainder of the length. Done over a long length of wire, the wires are twisted repeatedly clockwise - counter clockwise - clockwise - counter clockwise - ... along their whole length. The ends of the wires are then cleared of their insulation and soldered together to make a three-strand cable, and the cable is then used to wind the coils. This style of winding modifies the magnetic and electrical properties of the windings. It has been said that taking three long strands of wire and just twisting them together in one direction to make a long twisted three-strand cable is nearly as effective as using Litz wire. The websites www.mwswire.com/litzmain.htm and www.litz-wire.com are suppliers of ready made Litz wire.

A website which shows pictures of John's devices is: www.rexresearch.com/bedini/images.htm

CAUTION: Care must be taken when working with batteries, especially lead-acid batteries. A charged battery contains a large amount of energy and short-circuiting the terminals will cause a very large current flow which may start a fire. When being charged, some batteries give off hydrogen gas which when mixed with air is highly dangerous and which could explode if ignited by a spark. Batteries can explode and/or catch fire if grossly overcharged or charged with an excessively large current, so there could be danger from flying pieces of the casing and possibly acid being thrown around. Even an apparently clean lead-acid battery can have caustic traces on the case, so you should be sure to wash your hands thoroughly after handling a battery. Batteries with lead terminals tend to shed small fragments of lead when clips are put on them. Lead is toxic, so please be sure to wash your hands after handling any part of a lead-acid battery. Remember too that some batteries can develop slight leaks so please protect against any leakage. If you decide to perform any experiments using batteries, that you do so entirely at your own risk and on your own responsibility. This set of documents is presented for information purposes only and you are not encouraged to do anything other than read the information.

Also, if you get one of John's pulse motors tuned correctly, it will accelerate to perhaps 10,000 rpm. This is great for picking up energy but if ceramic magnets are used, the speed can cause them to disintegrate and fly in all directions. People have had magnet fragments embedded in their ceiling. It would be wise to build a housing enclosing the rotor and magnets so that if the magnets disintegrate, all of the fragments are contained safely.

Ronald Knight's Professional Advice on Battery Safety.
Ronald Knight has many years of professional experience in handling batteries and in pulse-charging them. He comments on battery safety as follows:

I have not heard of anyone having a catastrophic failure of a battery case in all the energy groups to which I belong and most of them use batteries in the various systems which I study. However, that does not mean that it cannot happen. The most common reason for catastrophic failure in the case of a lead-acid battery, is arcing causing failure in the grids which are assembled together inside the battery to make up the cells of the battery. Any internal arcing will cause a rapid build up of pressure from expanding Hydrogen gas, resulting in a catastrophic failure of the battery case.

I am a former maintenance engineer for U.S. Batteries, so I can say with confidence, that when you receive a new battery from at least that manufacturer, you receive a battery which has undergone the best test available to insure the manufacturer that he is not selling junk which will be sent back to him. It is a relatively easy test, and as it takes place during the initial charge, there is no wasted time nor is there one battery that escapes the pass-or-fail test. The battery is charged with the absolute maximum current which it can take. If the battery does not blow up due to internal arcing during the initial charge it is highly likely that it will not blow up under the regular use for which it was designed. However, all bets are off with used batteries that have gone beyond their expected life.

I have witnessed several catastrophic failures of battery cases daily at work. I have been standing right next to batteries (within 12 inches) when they explode (it is like a .45 ACP pistol round going off) and have only been startled and had to change my under shorts and Tyvek jump-suit, and wash off my rubber boots. I have been in the charge room with several hundred batteries at a time positioned very closely together and have seen batteries explode almost every working day and I have never seen two side by side blow, nor have I ever seen one fire or any flash damage to the case or surrounding area as a result. I have never even seen a flash but what I have seen tells me it is wise to always wear eye protection when charging.

I have my new gel cells in a heavy plastic zip-lock bags partly unzipped when in the house and in a marine battery box outside in the garage, that is just in the remote chance of catastrophic failure or the more likely event of acid on the outside of the battery case.

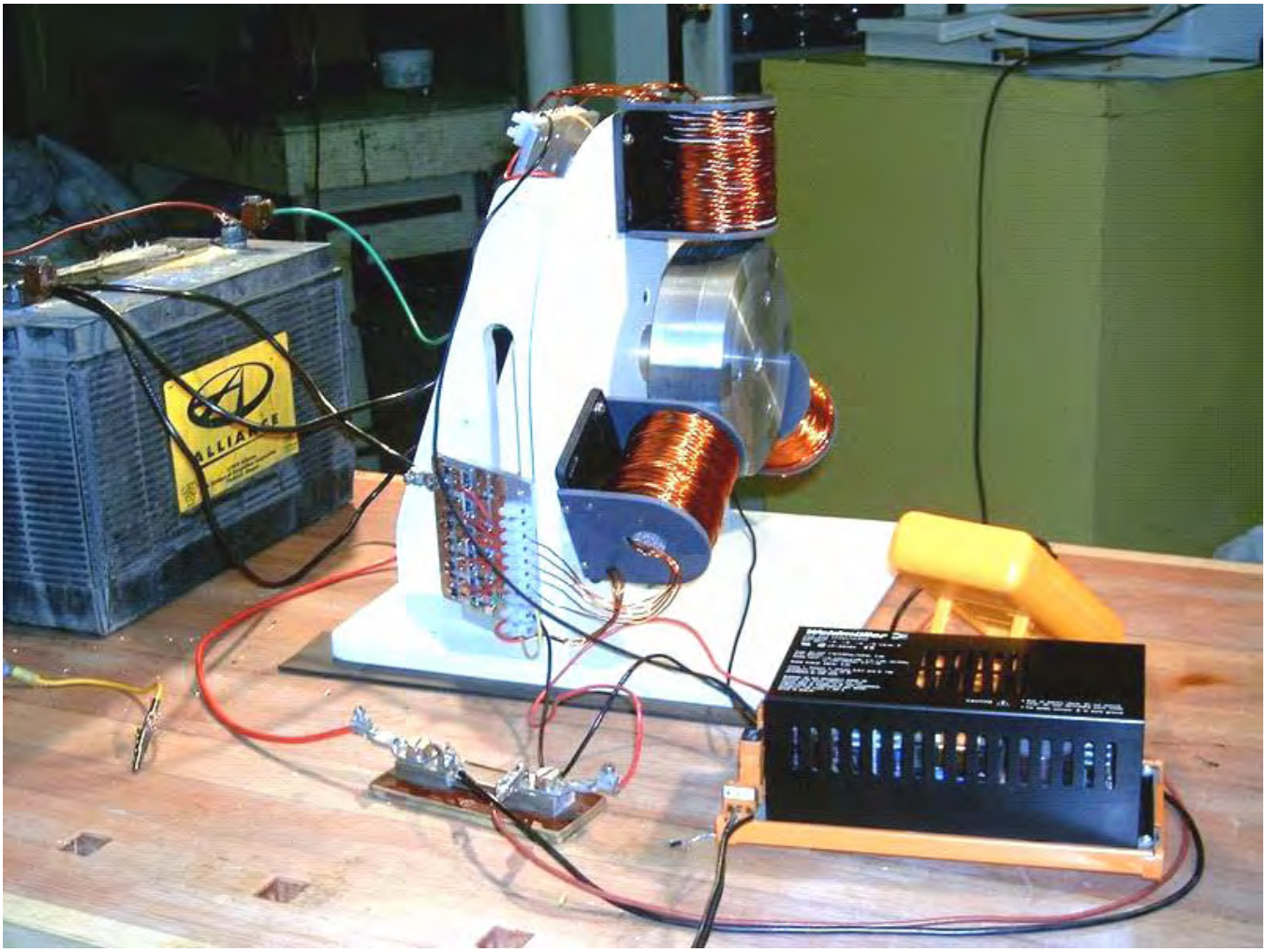
Vented batteries are always a risk of spillage which is their most common hazard, they should always be in a plastic lined cardboard or plastic box with sides taller than the battery and no holes in it. You would be surprised at how far away I have found acid around a vented lead acid battery under charge.

Have an emergency plan, keep a box of baking soda and a water source around to neutralise and flush the acid in case of spillage. It is best to have plastic under and around wherever your lead-acid batteries are located.

Ronald Knight gets about fifteen times more power from his Bedini-charged batteries than is drawn from the driving side of the circuit. He stresses that this does not happen immediately, as the batteries being charged have to be "conditioned" by repeated cycles of charging and discharging. When this is done, the capacity of the batteries being charged increases. Interestingly, the rate of current draw on the driving side of the circuit is not increased if the battery bank being charged is increased in capacity. This is because the power which charges the batteries flows from the environment and not from the driving battery. The driving battery just produces the high-voltage spikes which trigger the energy flow from the environment, and as a consequence of that the battery bank being charged can be a higher voltage than the 12-volt driving battery, and there can be any number of batteries in the charging bank.

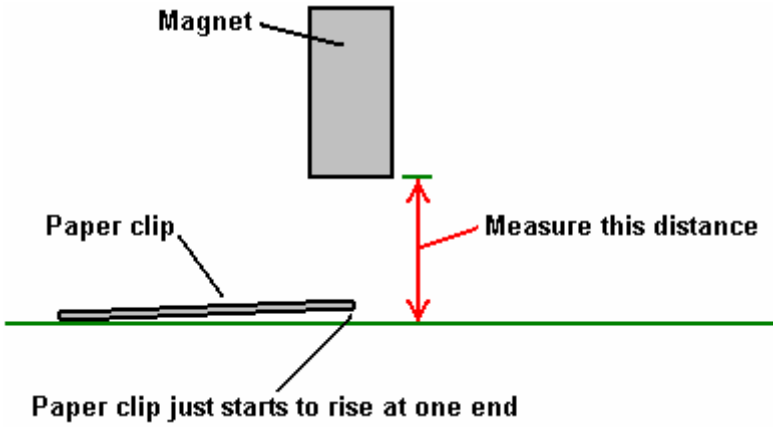
Ron Pugh's Battery Charger.
John Bedini's designs have been experimented with and developed by a number of enthusiasts. This in no way detracts from fact that the whole system and concepts come from John and I should like to express my sincere thanks to John for his most generous sharing of his systems. Thanks is also due to Ron Pugh who has kindly agreed for the details of one of his Bedini generators to be presented here. Let me stress again, that if you decide to build and use one of these devices, you do so entirely at your own risk and no responsibility for your actions rests with John Bedini, Ron Pugh or anyone else. Let me stress again that this document is provided for information purposes only and is not a recommendation or encouragement for you to build a similar device.

Ron's device is much more powerful than the average system, having fifteen coil windings and it performs most impressively. Here is a picture of it rotating at high speed:



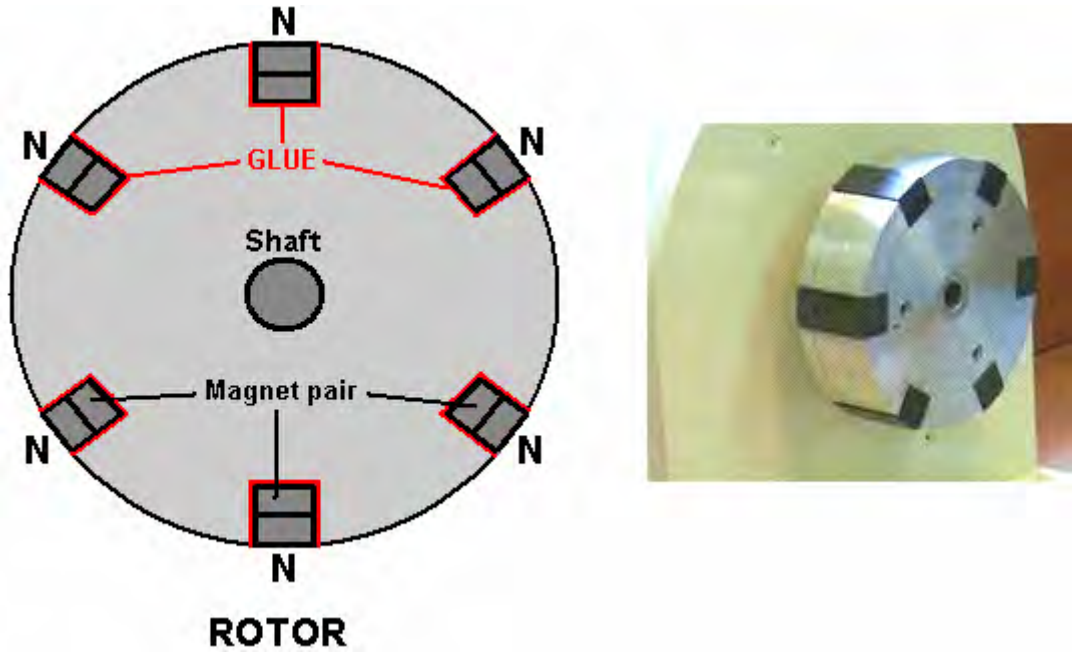
This is not a toy. It draws significant current and produces substantial charging rates. This is how Ron chose to build his device. The rotor is constructed from aluminium discs which were to hand but he would have chosen aluminium for the rotor if starting from scratch as his experience indicates that it is a very suitable material for the rotor. The rotor has six magnets inserted in it. These are evenly spaced 60 degrees apart with the North poles all facing outwards.

The magnets are normal ceramic types about 22 mm wide, 47 mm long and 10 mm high. Ron uses two of these in each of his six rotor slots. He bought several spare ones and then graded all of them in order of their magnetic strength, which varies a bit from magnet to magnet. Ron did this grading using a gauss meter. An alternative method would have been to use a paper clip about 30 mm in size and measure the distance at which one end of the clip just starts to rise up off the table as the magnet is moved towards it:

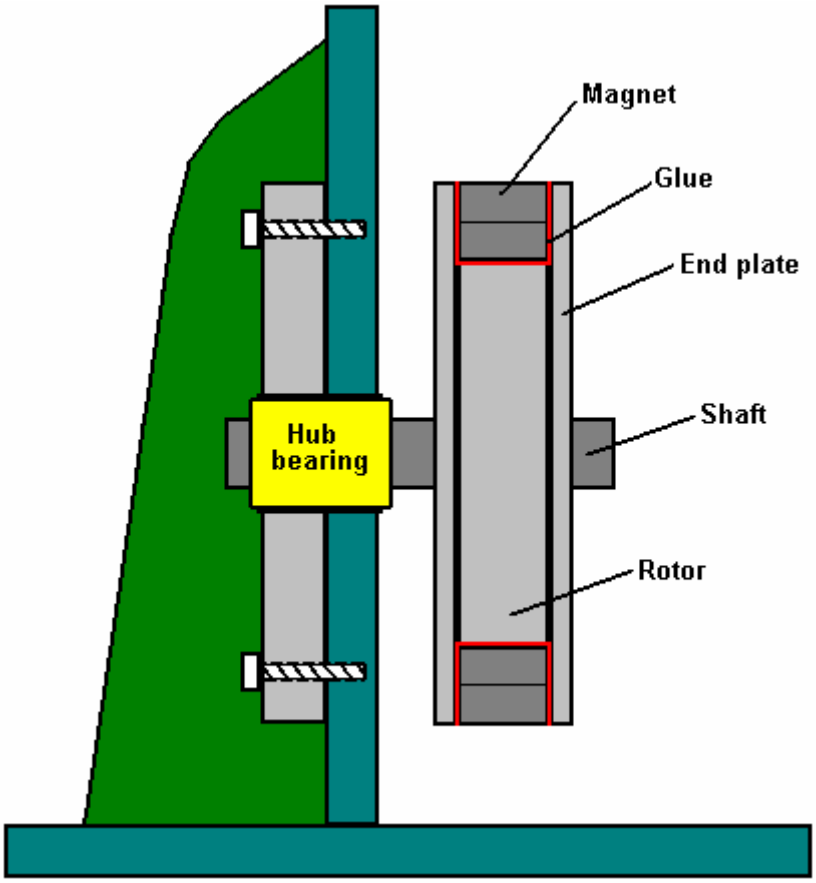


Having graded the magnets in order of strength, Ron then took the best twelve and paired them off, placing the

weakest and strongest together, the second weakest and the second strongest, and so on. This produced six pairs which have fairly closely matching magnetic strengths. The pairs of magnets were then glued in place in the rotor using super glue:

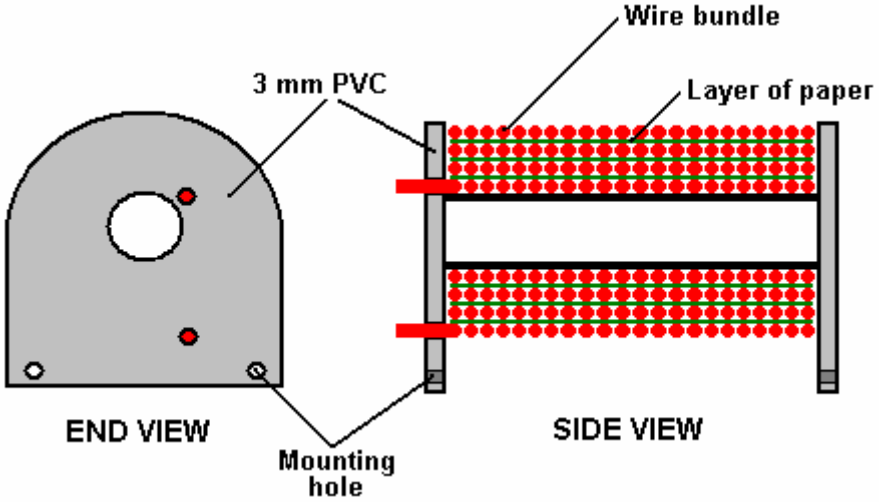


It is not desirable to recess the magnets though it is possible to place a restraining layer around the circumference of the rotor as the clearance between the magnet faces and the coils is about a quarter of an inch (6 mm) when adjusted for optimum performance. The North poles of the magnets face outwards as shown in the diagram above. If desired, the attachment of the magnets can be strengthened by the addition of blank side plates to the rotor which allows the magnet gluing to be implemented on five of the six faces of the magnet pairs:

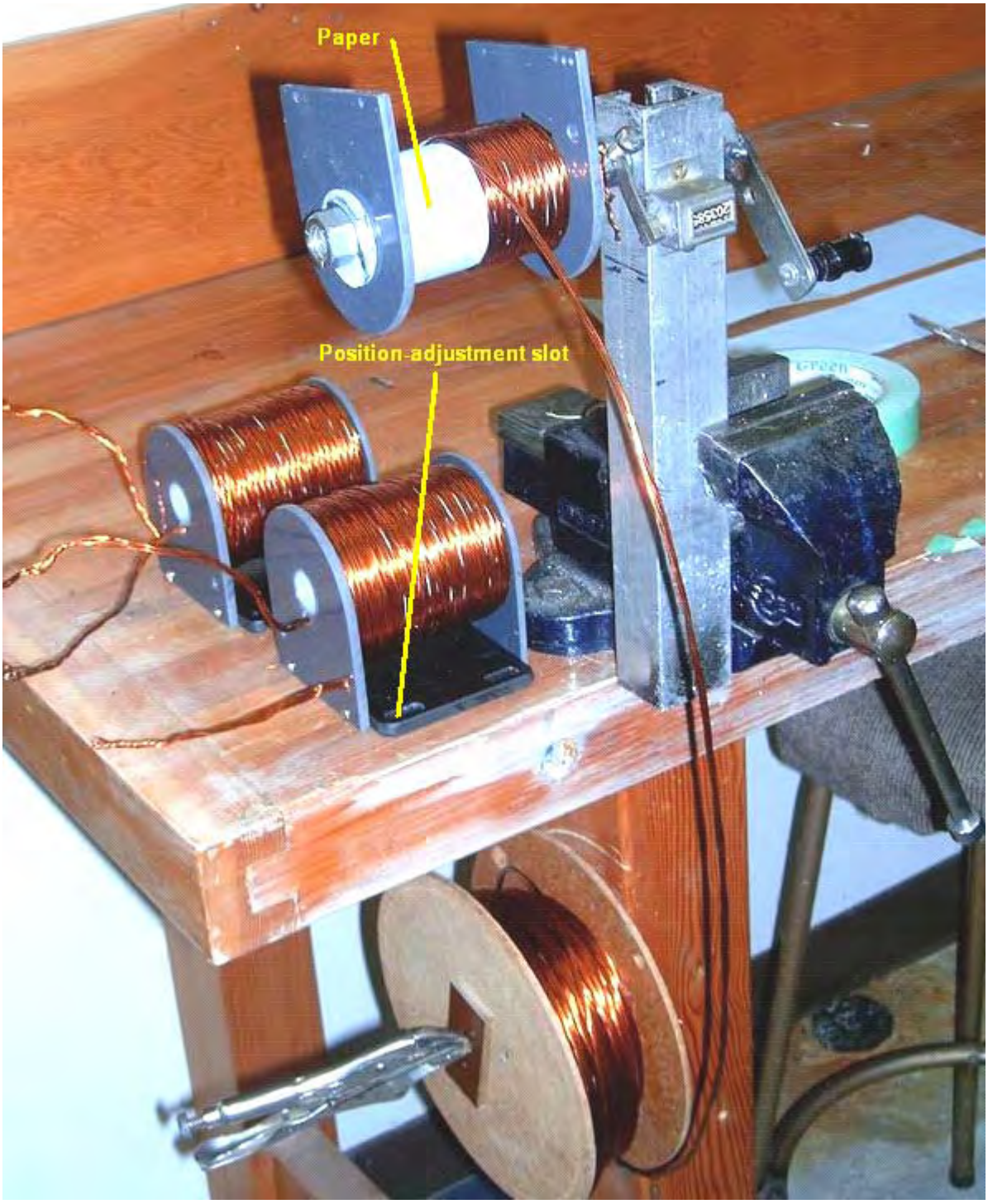


The magnets embedded in the outer edge of the rotor are acted on by wound “coils” which act as 1:1 transformers, electromagnets, and pickup coils. There are three of these “coils”, each being about 3 inches long and wound with five strands of #19 AWG (20 SWG) wire. The coil formers were made from plastic pipe of 7/8 inch (22 mm) outer diameter which Ron drilled out to an inner diameter of 3/4 inch (19 mm) which gives a wall thickness of 1/16 inch (1.5 mm). The end pieces for the coil formers were made from 1/8 inch (3 mm) PVC which was fixed to the plastic tube using plumbers PVC glue. The coil winding was with the five wires twisted around each other. This was done by clamping the ends of the five wires together at each end to form one 120 foot long bundle.

The bundle of wires was then stretched out and kept clear of the ground by passing it through openings in a set of patio chairs. A battery-powered drill was attached to one end and operated until the wires were loosely twisted together. This tends to twist the ends of the wires together to a greater extent near the end of the bundle rather than the middle. So the procedure was repeated, twisting the other end of the bundle. It is worth remarking in passing, that the drill turns in the same direction at each end in order to keep the twists all in the same direction. The twisted bundle of wires is collected on a large-diameter reel and then used to wind one of the “coils”.

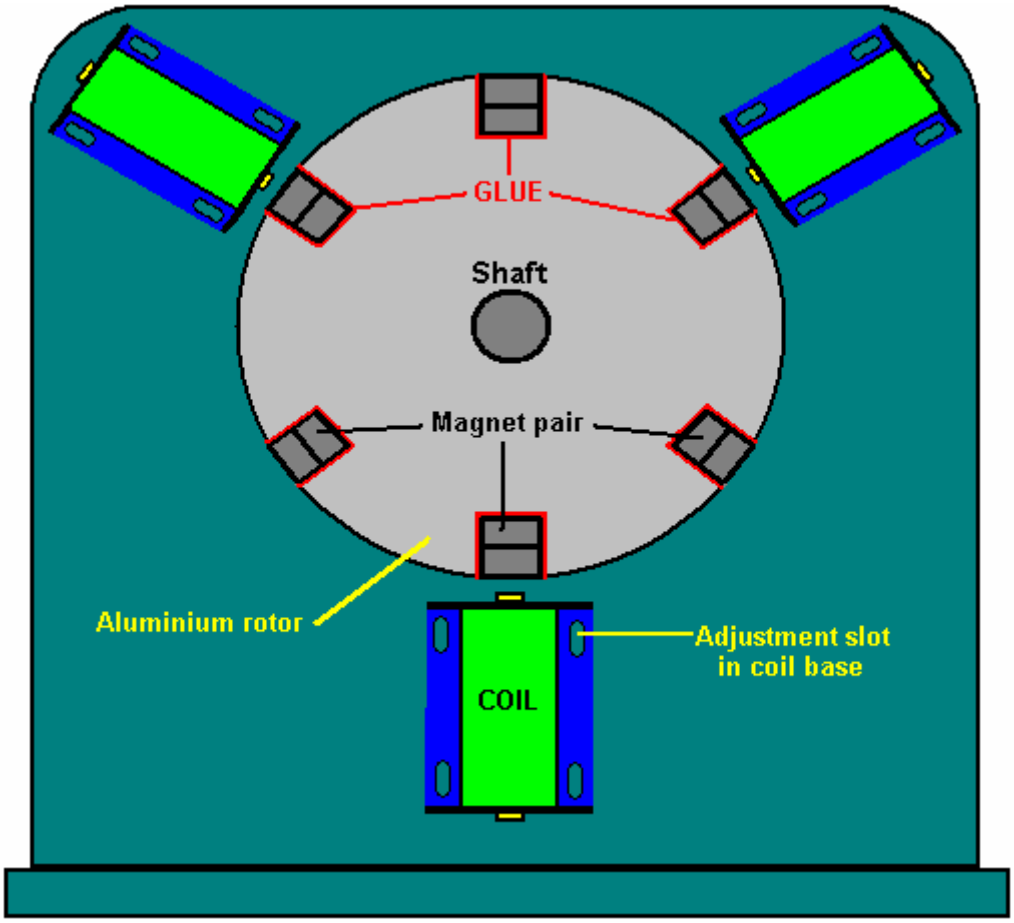


The coils are wound with the end plates attached and drilled ready to screw to their 1/4 inch (6 mm) PVC bases, which are the bolted to the 3/4 inch (18 mm) MDF supporting structure. To help the winding to remain completely even, a piece of paper is placed over each layer of the winding:



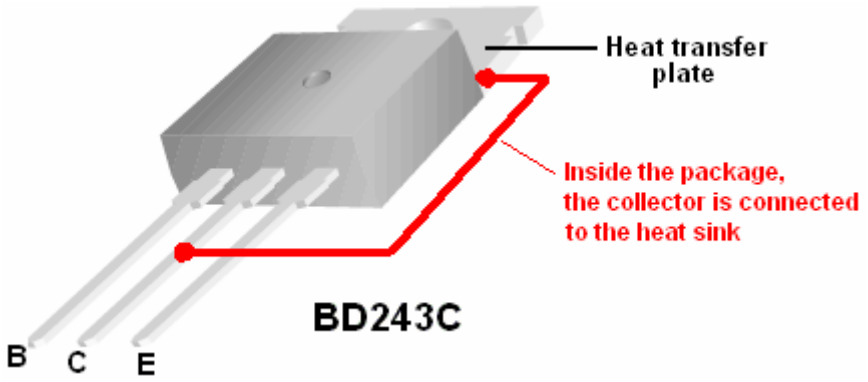
The three coils produced in this way were then attached to the main surface of the device. There could just as easily have been six coils. The positioning is made so as to create an adjustable gap of about 1/4 inch (6 mm) between the coils and the rotor magnets in order to find the optimum position for magnetic interaction. The magnetic effects are magnified by the core material of the coils. This is made from lengths of oxyacetylene welding wire which is copper coated. The wire is cut to size and coated with clear shellac to prevent energy loss through eddy currents circulating inside the core.

The coils are positioned at equal intervals around the rotor and so are 120 degrees apart. The end pieces of the coil formers are bolted to a 1/4 inch (6 mm) PVC base plate which has slotted mounting holes which allow the magnetic gap to be adjusted as shown here:

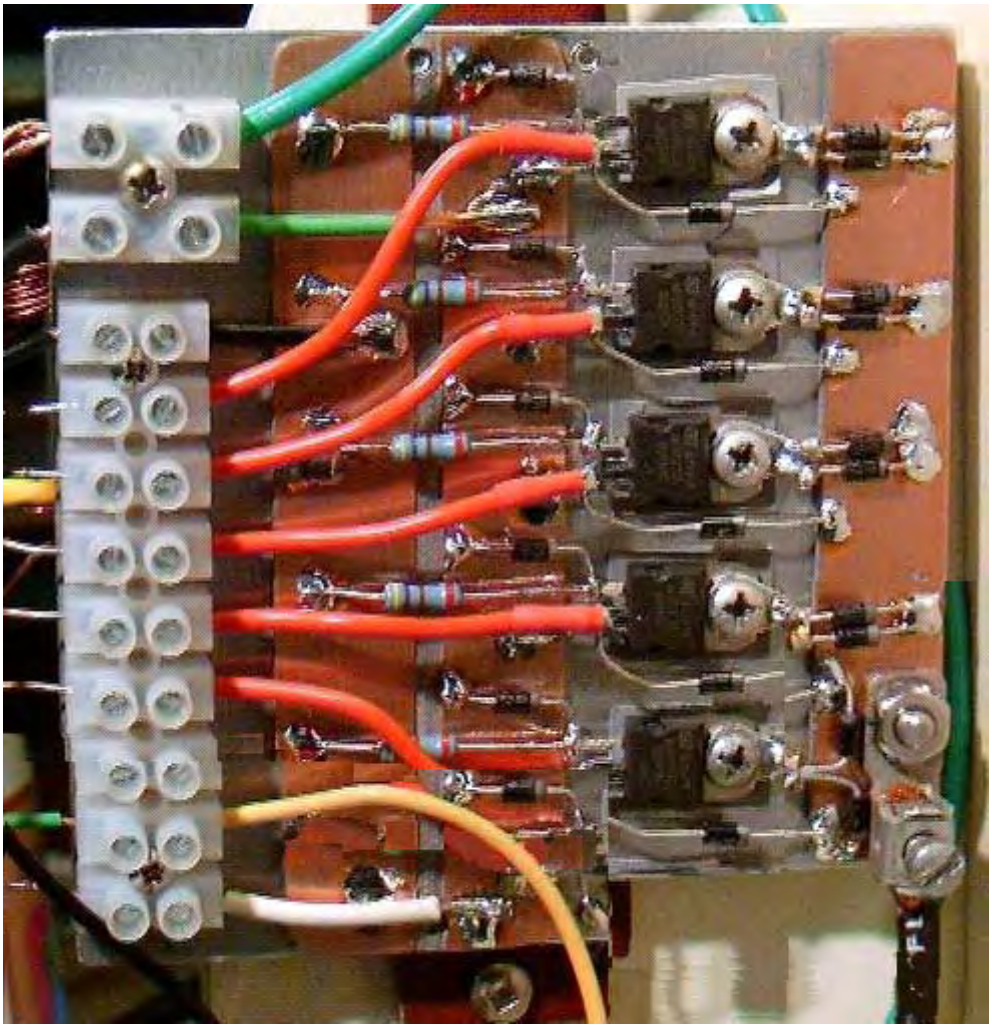


The three coils have a total of fifteen identical windings. One winding is used to sense when a rotor magnet reaches the coils during its rotation. This will, of course happen six times for each revolution of the rotor as there are six magnets in the rotor. When the trigger winding is activated by the magnet, the electronics powers up all of the remaining fourteen coils with a very sharp, pulse which has a very short rise time and a very short fall time. The sharpness and brevity of this pulse is a critical factor in drawing excess energy in from the environment and will be explained in greater detail later on. The electronic circuitry is mounted on three aluminium heat sinks, each about 100 mm square. Two of these have five BD243C NPN transistors bolted to them and the third one has four BD243C transistors mounted on it.

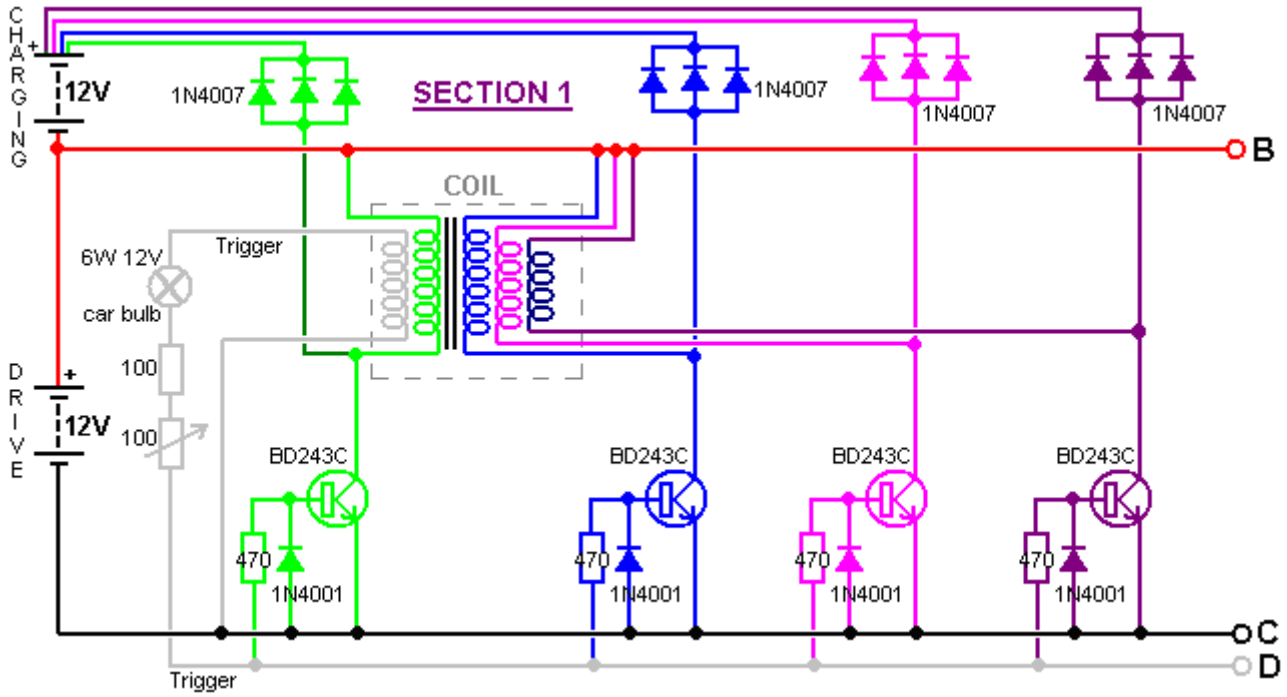
The metal mounting plate of the BD243 transistors acts as its heat sink, which is why they are all bolted to the large aluminium plate. BD243C transistors look like this:

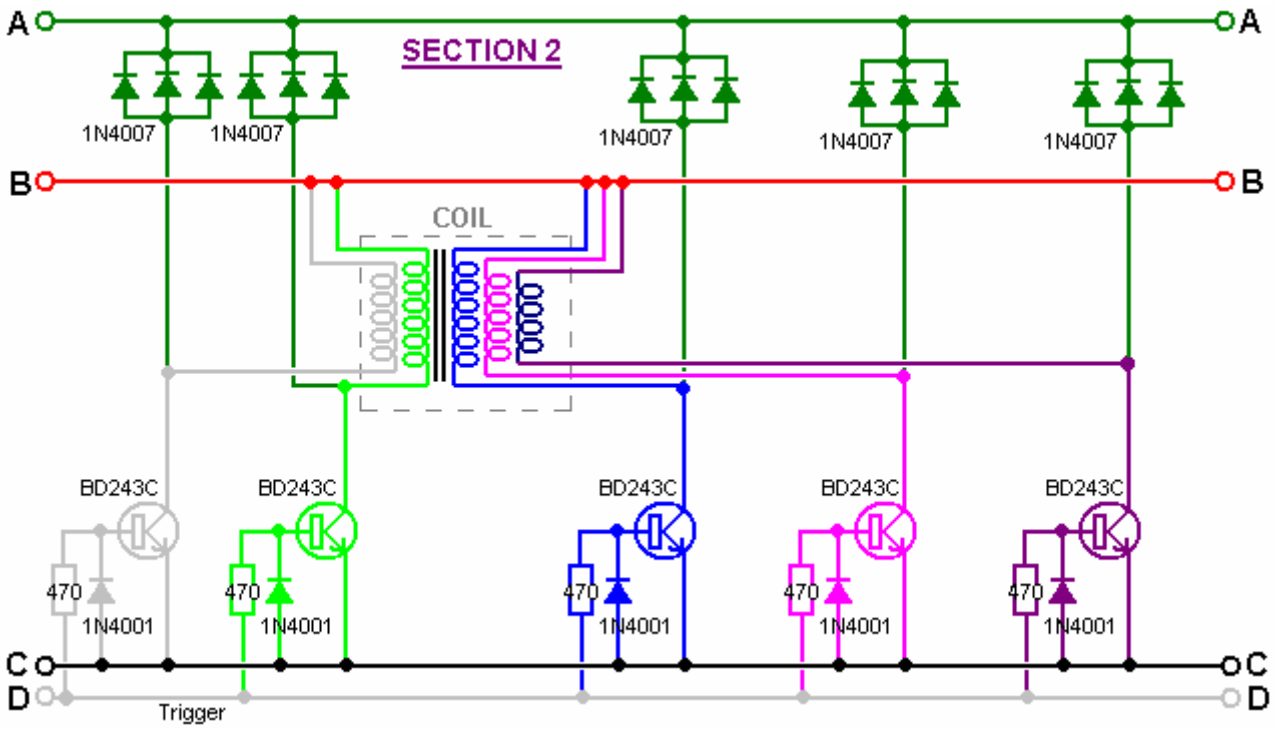


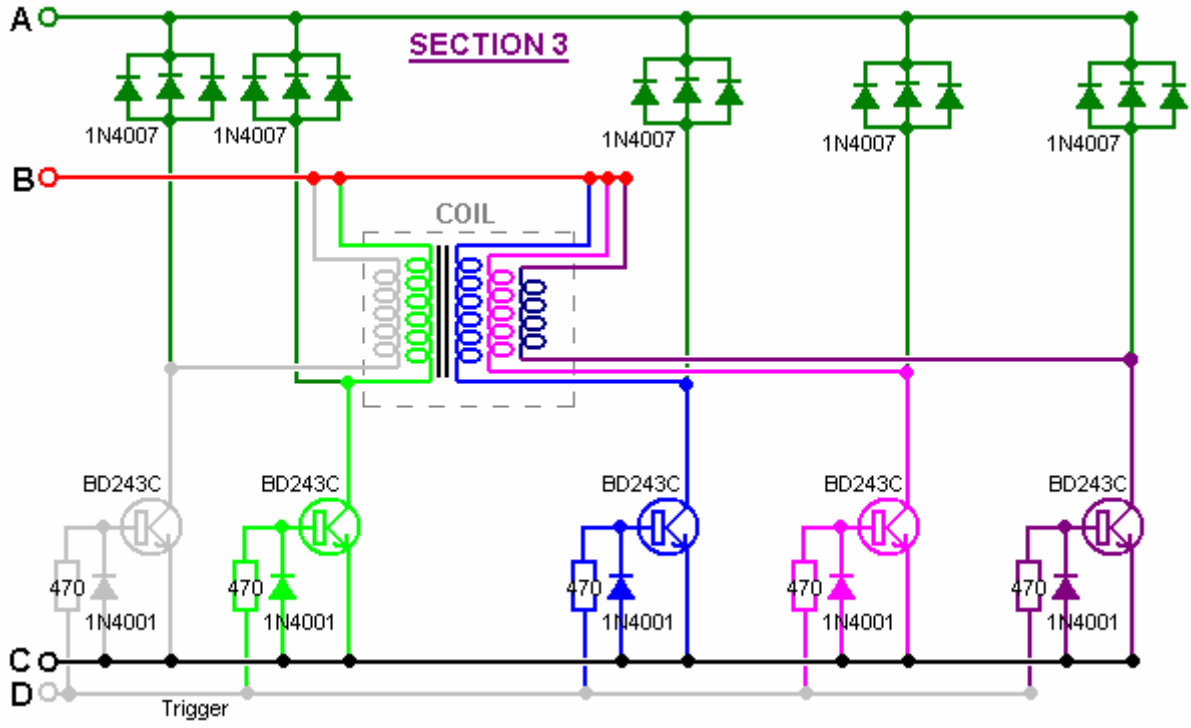
The circuit has been built on the aluminium panels so that the transistors can be bolted directly on to it, and provided with insulating strips mounted on top of it to avoid short circuits to the other components. Standard strip connector blocks have been used to inter-connect the boards which look like this:



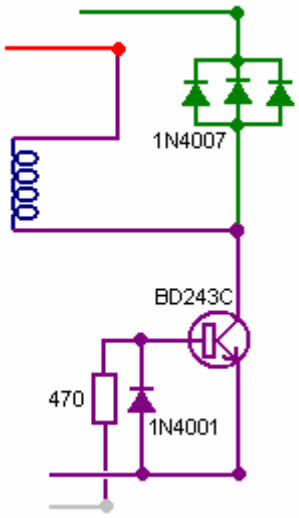
The circuit used with this device is simple but as there are so many components involved, the diagram is split into parts to fit on the page. These diagrams are usually drawn with a common charging wire going to the top of the battery which is being charged. However, it needs to be understood that drawing it that way is just for convenience and better performance is achieved if each charging circuit has its own separate wire going to the charging battery as shown in Section 1 here:


CHARGING
12V
1N4007
SECTION 1
1N4007
1N4007
1N4007
B
COIL
Trigger
6W 12V
car bulb
100
100
DRIVE
12V
BD243C
BD243C
BD243C
BD243C
470
1N4001
470
1N4001
470
1N4001
470
1N4001
C
D
Trigger


A
A
SECTION 2
1N4007
1N4007
1N4007
1N4007
1N4007
B
B
COIL
BD243C
BD243C
BD243C
BD243C
BD243C
470
1N4001
470
1N4001
470
1N4001
470
1N4001
470
1N4001
C
D
C
D
Trigger

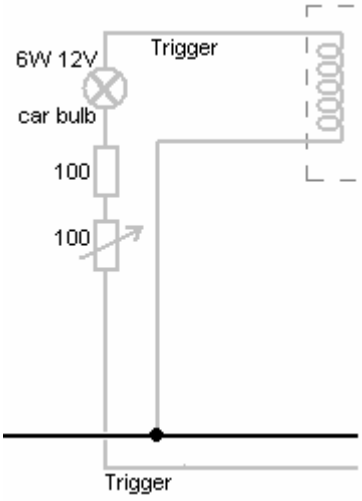


While this looks like a fairly large and complicated circuit, it actually is not. You will notice that there are fourteen identical circuit sections. Each of these is quite simple:



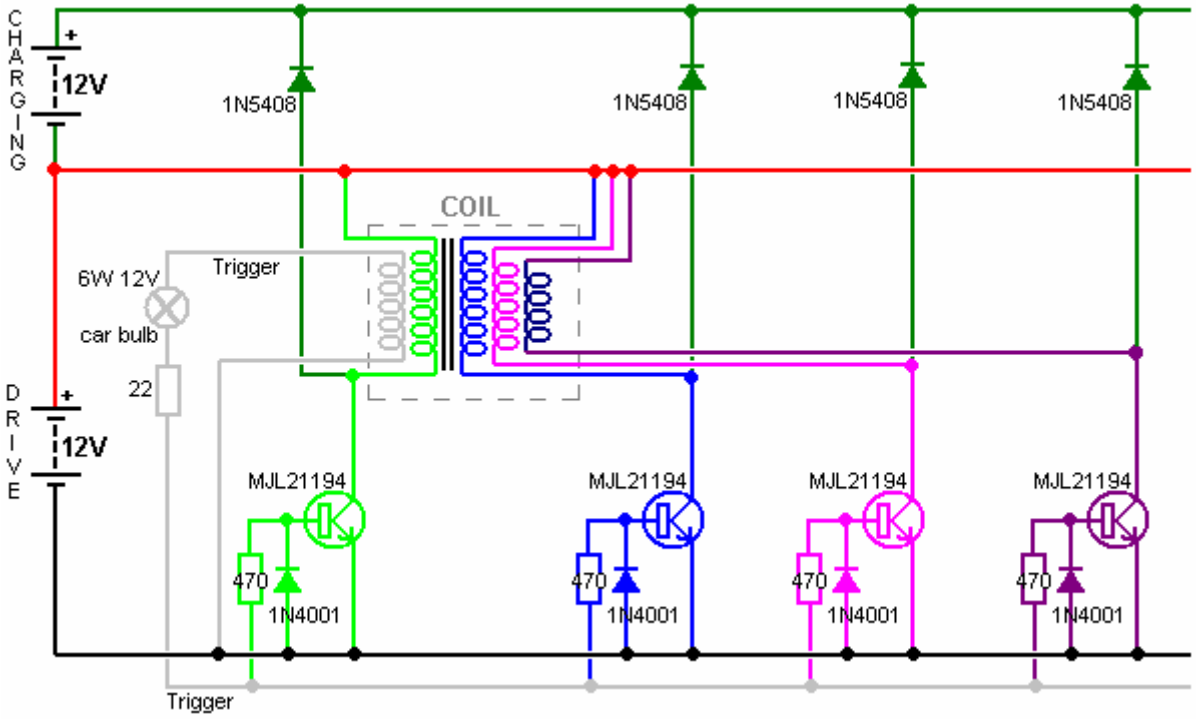
This is a very simple transistor circuit. When the trigger line goes positive (driven by the magnet passing the coil) the transistor is switched on hard, powering the coil which is then effectively connected across the driving battery. The trigger pulse is quite short, so the transistor switches off almost immediately. This is the point at which the circuit operation gets subtle. The coil characteristics are such that this sharp powering pulse and sudden cut-off cause the voltage across the coil to rise very rapidly, dragging the voltage on the collector of the transistor up to several hundred volts. Fortunately, this effect is energy drawn from the environment which is quite unlike conventional electricity, and thankfully, a good deal less damaging to the transistor. This rise in voltage, effectively “turns over” the set of three 1N4007 diodes which then conducts strongly, feeding this excess free-energy into the charging battery. Ron uses three diodes in parallel as they have a better current-carrying capacity and thermal characteristics than a single diode. This is a common practice and any number of diodes can be placed in parallel, with sometimes as many as ten being used.

The only other part of the circuit is the section which generates the trigger signal:

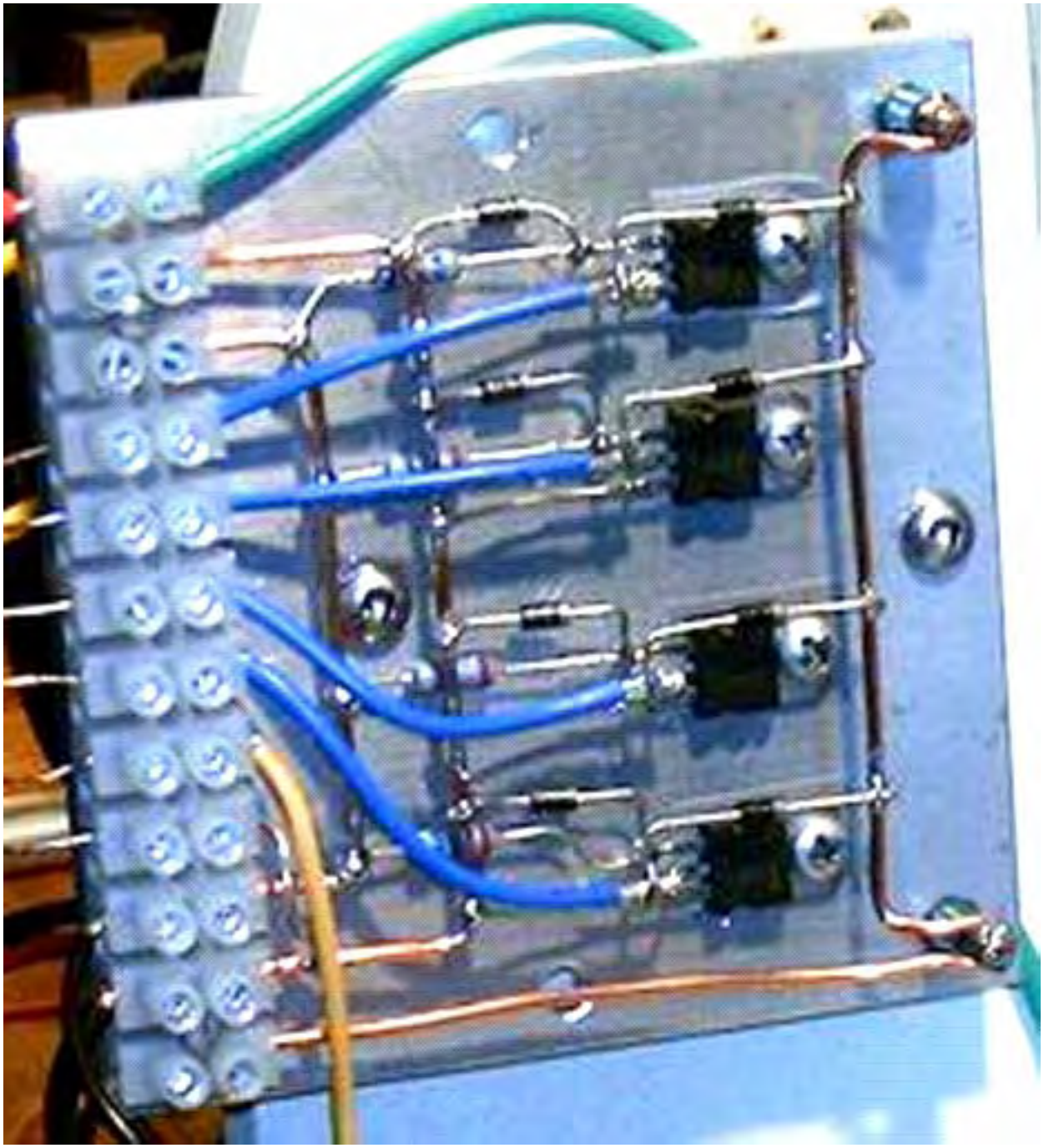


When a magnet passes the coil containing the trigger winding, it generates a voltage in the winding. The intensity of the trigger signal is controlled by passing it through an ordinary vehicle 6 watt, 12 volt bulb and then further limiting the current by making it pass through a resistor. To allow some manual control of the level of the trigger signal, the resistor is divided into a fixed resistor and a variable resistor (which many people like to call a "pot"). This variable resistor and the adjustment of the gap between the coils and the rotor are the only adjustments of the device. The bulb has more than one function. When the tuning is correct, the bulb will glow dimly which is a very useful indication of the operation. The trigger circuit then feeds each of the transistor bases via their 470 ohm resistors.

John Bedini aims for an even more powerful implementation, wiring his circuit with AWG #18 (19 SWG) heavy-duty copper wire and using MJL21194 transistors and 1N5408 diodes. He increases the trigger drive by dropping the variable resistor and reducing fixed resistor to just 22 ohms. The MJL21194 transistor has the same pin connections as the BD243C transistor. This is the starting section of John's circuit:



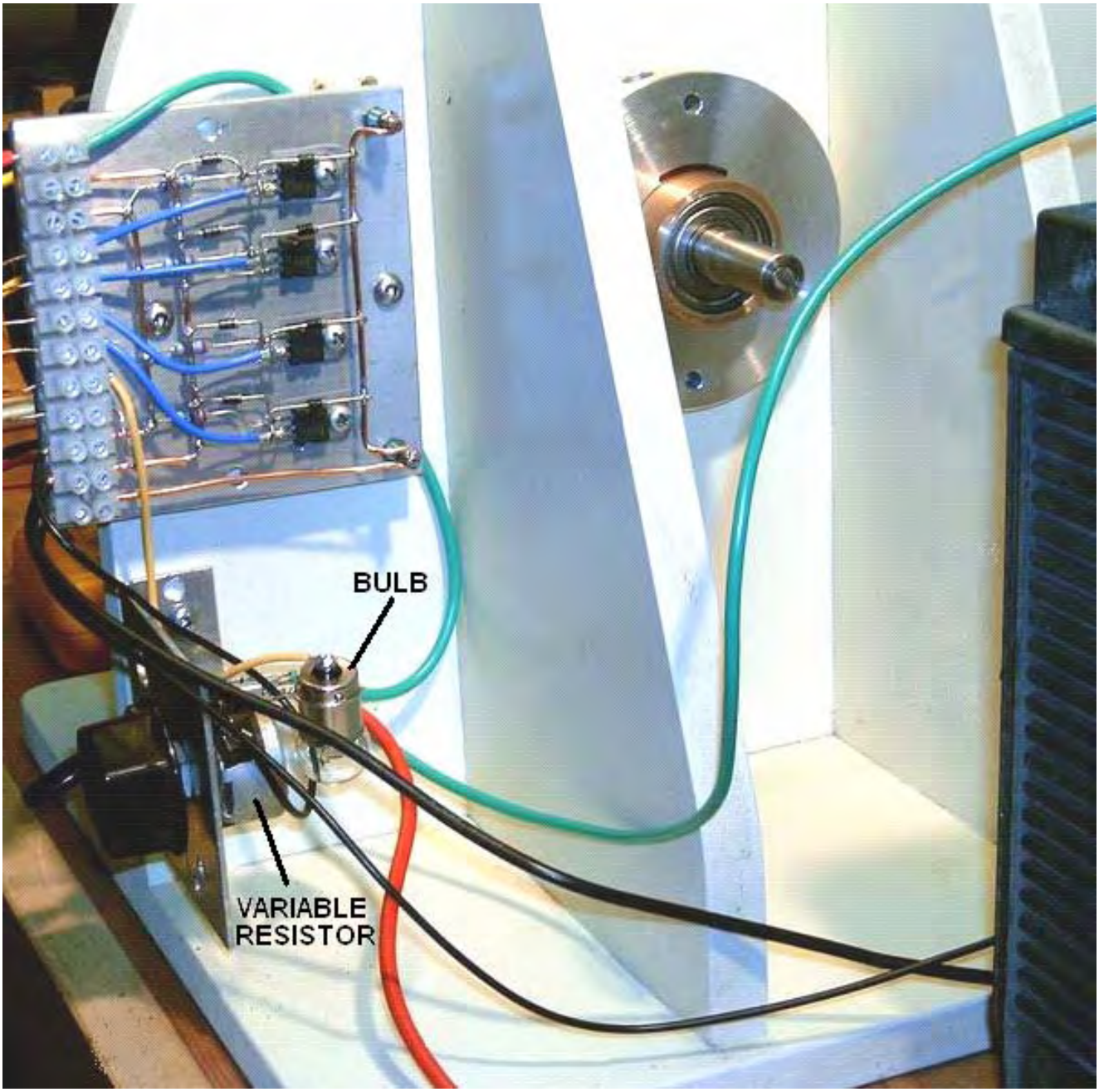
There are various ways of constructing this circuit. Ron shows two different methods. The first is shown above and uses paxolin strips (printed-circuit board material) above the aluminium heat sink to mount the components. Another method which is easy to see, uses thick copper wires held clear of the aluminium, to provide a clean and secure mounting for the components as shown here:



It is important to realise that the collector of a BD243C transistor is internally connected to the heat-sink plate used for the physical mounting of the transistor. As the circuit does not have the collectors of these transistors connected together electrically, they cannot just be bolted to a single heat-sink plate. The above picture might give the wrong impression as it does not show clearly that the metal bolts fastening the transistors in place do not go directly into the aluminium plate, but instead, they fasten into plastic tee-nuts.

An alternative, frequently used by the builders of high-powered electronic circuits, is to use mica washers between the transistor and the common heat sink plate, and use plastic fastening bolts or metal bolts with a plastic insulating collar between the fastening and the plate. Mica has the very useful property of conducting heat very well but not conducting electricity. Mica “washers” shaped to the transistor package are available from the suppliers of the transistors. In this instance, it seems clear that heat dissipation is not a problem in this circuit, which in a way is to be expected as the energy being drawn from the environment is frequently called “cold” electricity as it cools components down with increasing current as opposed to heating them up as conventional electricity does.

This particular circuit board is mounted at the rear of the unit:



Although the circuit diagram shows a twelve volt drive supply, which is a very common supply voltage, Ron sometimes powers his device with a mains operated Power Supply Unit which shows a power input of a pretty trivial 43 watts. It should be noted that this device operates by pulling in extra power from the environment. That drawing in of power gets disrupted if any attempt is made to loop that environmental power back on itself or driving the unit directly from another battery charged by the unit itself. It may be just possible to power the unit successfully from a previously charged battery if an inverted is used to convert the power to AC and then a step-down transformer and regulated power rectification circuit is used. As the power input is so very low, off-grid operation should be easily possible with a battery and a solar panel.

It is not possible to operate a load off the battery under charge during the charging process as this disrupts the energy flow. Some of these circuits recommend that a separate 4 foot long earthing rod be used to earth the negative side of the driving battery, but to date, Ron has not experimented with this. In passing, it is good practice to enclose any lead-acid battery in a battery box. Marine chandlers can supply these as they are used extensively in boating activities.

When cutting the wire lengths for coating and pushing into the coil formers, Ron uses a jig to ensure that all of the lengths are identical. This arrangement is shown here:

The distance between the shears and the metal angle clamped to the workbench makes each cut length of wire exactly the required size while the plastic container collects the cut pieces ready for coating with clear shellac or clear polyurethane varnish before use in the coil cores.

Experience is particularly important when operating a device of this kind. The 100 ohm variable resistor should be a wire-wound type as it has to carry significant current. Initially the variable resistor is set to its minimum value and the power applied. This causes the rotor to start moving. As the rate of spin increases, the variable resistor is gradually increased and a maximum speed will be found with the variable resistor around the middle of its range, i.e. about 50 ohm resistance. Increasing the resistance further causes the speed to reduce.

The next step is to turn the variable resistor to its minimum resistance position again. This causes the rotor to leave its previous maximum speed (about 1,700 rpm) and increase the speed again. As the speed starts increasing again, the variable resistor is once again gradually turned, increasing its resistance. This raises the rotor speed to about 3,800 rpm when the variable resistor reaches mid point again. This is probably fast enough for all practical purposes, and at this speed, even the slightest imbalance of the rotor shows up quite markedly. To go any faster than this requires an exceptionally high standard of constructional accuracy. Please remember that the rotor has a large amount of energy stored in it at this speed and so is potentially very dangerous. If the rotor breaks or a magnet comes off it, that stored energy will produce a highly dangerous projectile. That is why it is advisable, although not shown in the above photographs, to construct an enclosure for the rotor. That could be a U-shaped channel between the coils. The channel would then catch and restrain any fragments should anything break loose.

If you were to measure the current during this adjustment process, it would be seen to reduce as the rotor speeds up. This looks as if the efficiency of the device is rising. That may be so, but it is not necessarily a good thing in this case where the objective is to produce radiant energy charging of the battery bank. John Bedini has shown that serious charging takes place when the current draw of the device is 3 to 5+ amps at maximum rotor speed and not a miserly 50 mA draw, which can be achieved but which will not produce good charging. The power can

be increased by raising the input voltage to 24 volts or even higher - John Bedini operates at 48 volts rather than 12 volts

The device can be further tuned by stopping it and adjusting the gap between the coils and the rotor and then repeating the start-up procedure. The optimum adjustment is where the final rotor speed is the highest.

The above text is intended to give a practical introduction to one of John Bedini's inventions. It seems appropriate that some attempt at an explanation of what is happening, should be advanced at this point. In the most informative book "Energy From The Vacuum - Concepts and Principles" by Tom Bearden an explanation of this type of system is put forward. While the description appears to be aimed mainly at John's motor system which ran continuously for three years, powering a load and recharging it's own battery, the description would appear to apply to this system as well. I will attempt to summarise it here:

Conventional electrical theory does not go far enough when dealing with lead/acid batteries in electronic circuits. Lead/acid batteries are extremely non-linear devices and there is a wide range of manufacturing methods which make it difficult to present a comprehensive statement covering every type in detail. However, contrary to popular belief, there are actually at least three separate currents flowing in a battery-operated circuit:

1. Ion current flowing in the electrolyte between the plates inside the battery. This current does not leave the battery and enter the external electronic circuit.

2. Electron current flowing from the plates out into the external circuit.

3. Current flow from the environment which passes along the external circuitry and into the battery.

The exact chemical processes inside the battery are quite complex and involve additional currents which are not relevant here. The current flow from the environment follows the electron flow around the external circuit and on into the battery. This is "cold" electricity which is quite different to conventional electricity and it can be very much larger than the standard electrical current described in conventional textbooks. A battery has unlimited capacity for this kind of energy and when it has a substantial "cold" electricity charge, it can soak up the conventional energy from a standard battery charger for a week or more, without raising the battery voltage at all.

An important point to understand is that the ions in the lead plates of the battery have much greater inertia than electrons do (several hundred thousand times in fact). Consequently, if an electron and an ion are both suddenly given an identical push, the electron will achieve rapid movement much more quickly than the ion will. It is assumed that the external electron current is in phase with the ion current in the plates of the battery, but this need not be so. John Bedini deliberately exploits the difference of momentum by applying a very sharply rising potential to the plates of the battery.

In the first instant, this causes electrons to pile up on the plates while they are waiting for the much heavier ions to get moving. This pile up of electrons pushes the voltage on the terminal of the battery to rise to as much as 100 volts. This in turn, causes the energy to flow back out into the circuit as well as into the battery, giving simultaneously, both circuit power and serious levels of battery charging. This over potential also causes much increased power flow from the environment into the circuit, giving augmented power both for driving the external circuit and for increasing the rate of battery charge. The battery half of the circuit is now 180 degrees out of phase with the circuit-powering half of the circuit.

It is important to understand that the circuit-driving energy and the battery-charging energy do **not** come from the sharp pulses applied to the battery. Instead, the additional energy flows in from the environment, triggered by the pulses generated by the Bedini circuit. In other words, the Bedini pulses act as a tap on the external energy source and are not themselves the source of the extra power.

If the Bedini circuit is adjusted correctly, the pulse is cut off very sharply just before the tapped energy inflow is about to end. This has a further enhancing effect due to the Lenz law reaction which causes an induced voltage surge which can take the over-voltage potential to as much as 400 volts. This has a further effect on the local environment, drawing in an even higher level of additional power and extending the period of time during which that extra power flows into both the circuit and the battery. This is why the exact adjustment of a Bedini pulsing system is so important.

Ossie Callanan's Free-Energy System

In 2007, Ossie Callanan published a document showing how and why he was getting COP>1 battery charging. Ron Pugh's system kindly shared in detail above, with careful tuning and running on 24 volt input and 24 volt

output operates at COP>10 which is likely to be due to Ron's skill in building and adjusting, both of which are very good indeed, coupled with the use of many transistors working in parallel and tripled charging diodes to improve their performance. The John Bedini SSG pulse-charging system is very easy to construct and works very well, even with many sulphated batteries which have been discarded as useless. However, most people will not get COP>1 performance from their own SSG build. Ossie explains why this is in the following section which is his Copyright. He says:

I believe I may have this radiant energy system worked out to the point where anyone can build it and when you build all of it, it can provide you with free and continuous energy. There are two sides to the circuit and just having either side is no good, you must have both. One side is the pulse-charger side, and the other is the battery and accumulator-converter side. At present, I am charging batteries at a performance level between COP=2 and COP=10 and battery swapping is not a problem.

First we need to focus on the charger side of the system. Basically, you must build a charger which produces large amounts of radiant energy in the form of radiant pulses. Radiant energy pulses are Back-EMF pulses provided that they have very fast rising edges and falling edges and occur at high frequencies. They are **not** transistor transients or switching transients! A spark-gap produces classical radiant energy pulses. They are chaotic events but they are radiant energy events nevertheless. Switching a coil on very rapidly using a transistor will produce one radiant energy pulse from the coil's Back-EMF, but one pulse on its own is no good. You need thousands or better still, millions of those pulses for them to be of any practical use.

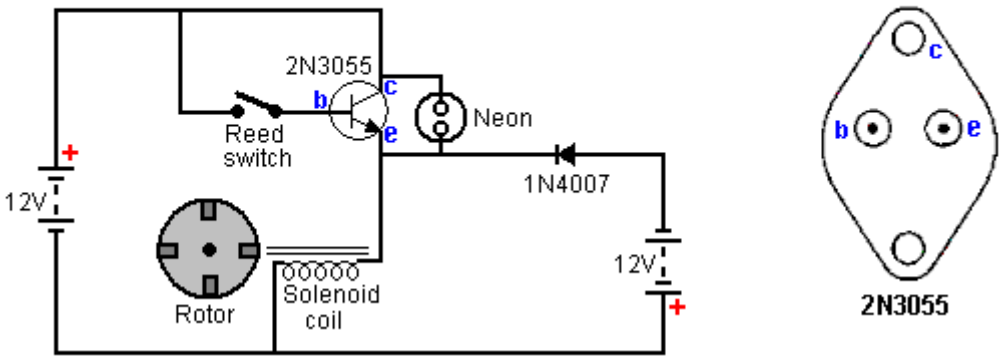
John Bedini's Simplified School Girl motor (the "SSG") only produces a very small amount of radiant energy when you have the base of the transistor tuned so that you get the longest self oscillating pulse train per magnet pass. It is not very efficient as the base-clamping diode wastes this energy, passing it back through the base resistor-bulb circuit but that circuit it is necessary so that you can synchronise the operation and provide a driving force to the passing magnet. Without the base-clamping diode, you get an oscillator and the motor won't turn - catch 22. But still, even when tuned with the base-clamping diode in place, the self-oscillating pulse train is usually at most 3 to 6 pulses which is not very much and so, not much radiant energy is produced. For all those people trying to tune the motor for only one pulse per magnet pass, they are wasting their time and only building a pulse motor and not an efficient radiant energy generator.

The SSG is not a very powerful or good radiant energy generator and apart from being educational, it is really a waste of time unless someone can explain to you how to tune it to get the greatest possible amount of radiant energy from it with a long pulse train, and then tell you what to do with that radiant energy. John Bedini's method of using a bulb in the base-feed circuit is to keep the motor tuned to a particular number of pulses in the pulse train per magnet pass or for the longest pulse train as the motor increases speed, also, the battery's impedance changes the speed as well. The two-strand coil where one coil is used just as the transistor trigger, along with the wasted energy in the base circuit, add to the overall inefficiency and make construction more difficult.

In saying that about the SSG I would like to show you now a very simple and basic pulse oscillator or motor driver circuit that you can build with off the shelf parts, one which will produce very large amounts of radiant energy when adjusted correctly. Here is that circuit:

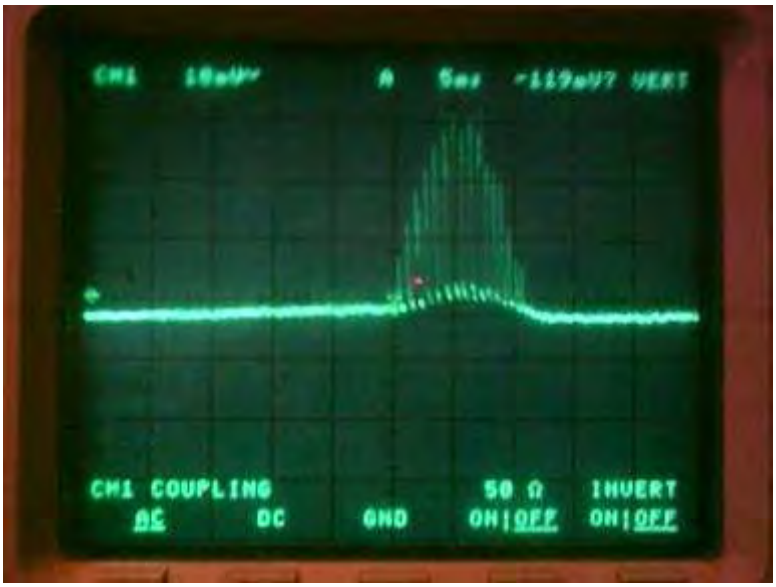
Basic Radiant Pulse Oscillator / Motor-Driver Circuit

By Ossie Callanan

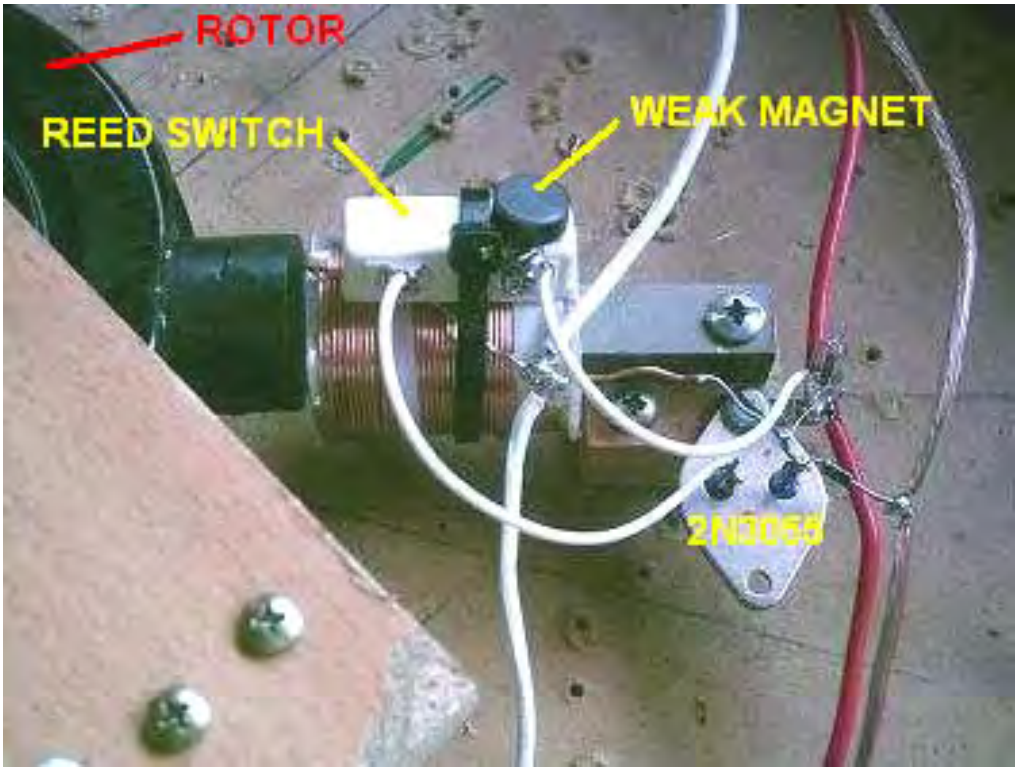


Do not be fooled by appearances - this is as close to a controlled spark gap circuit as you are going to get and it is extremely efficient in the production of radiant energy! But most importantly, you must place and adjust the reed switch appropriately and properly!!! Take a look at this picture:

As you can see, the trick is to place the reed switch so that it runs along the length of the solenoid coil and so is located in the coil's magnetic field. This couples the coil's magnetic field to the reed switch as well as the magnetic field of the passing rotor magnet. This provides a magnetic feedback and turns the reed switch into an oscillator. So when the magnet passes the reed oscillates with the coil's field and causes very many pulses, typically 20 to 50 pulses per magnet pass. Amazingly, unlike the clamping diodes in the SSG, this is not wasteful. This reed switch oscillation actually reduces the input current. Instead of the reed remaining closed for the whole of the pulse duration, it switches on and off and so, less input power is drawn from the driving battery. I run the motor so that when the coil is energised, the magnet is attracted to the coil. Here is a trace across the charging battery:



The trace above is from a motor that is only drawing 50 milliamps but is charging the battery many times faster than if it were drawing 300 milliamps with a single pulse per magnet pass! But there is more. Take a look at the following picture:



Using a very small and weak magnet, you can now control and adjust the reed switch. This allows you to adjust the switching so that the circuit oscillates continuously but still powers the passing rotor magnet. Below, is the trace across the charging battery and the battery is charging very rapidly although you are still paying for this as the input current will increase but nevertheless, you are really producing a vast amount of radiant energy for what is effectively, very little input current! Also, when you do this, the coils HISS very loudly! Yes, the coils HISS, not with a tone or frequency but with a hissing NOISE.

My prototype motor uses four of these circuit's, positioned 90 degrees apart around the rotor, and all connected in parallel. You can use just one reed switch to switch all four transistors and coils but it is more load and the reed switch does not hold up for very long. In fact, either way, if you use small reed switches, they will get worn and start to stick. I have bought larger reed switches but I am also working on an electronic switching version of this, although that is easier said than done. I have worked on this for a few months now so I have tried very many things and have not managed to match it with electronic switching yet. Limiting the current passing through the reed switch does not necessary increase the length of its operational life, besides which, doing that produces less radiant energy.

Now having shown you all this, we are only half way to a complete radiant energy system which will provide continuous free energy. The above circuit and motor, even though they do provide large amounts of radiant energy, will still only give you a COP equal to, or close to 1 when regularly swapping over between the source

battery and the charging battery. For battery swapping to work, you MUST have the second and equally important side to the system. The second side of the system is the radiant energy accumulator-converter.

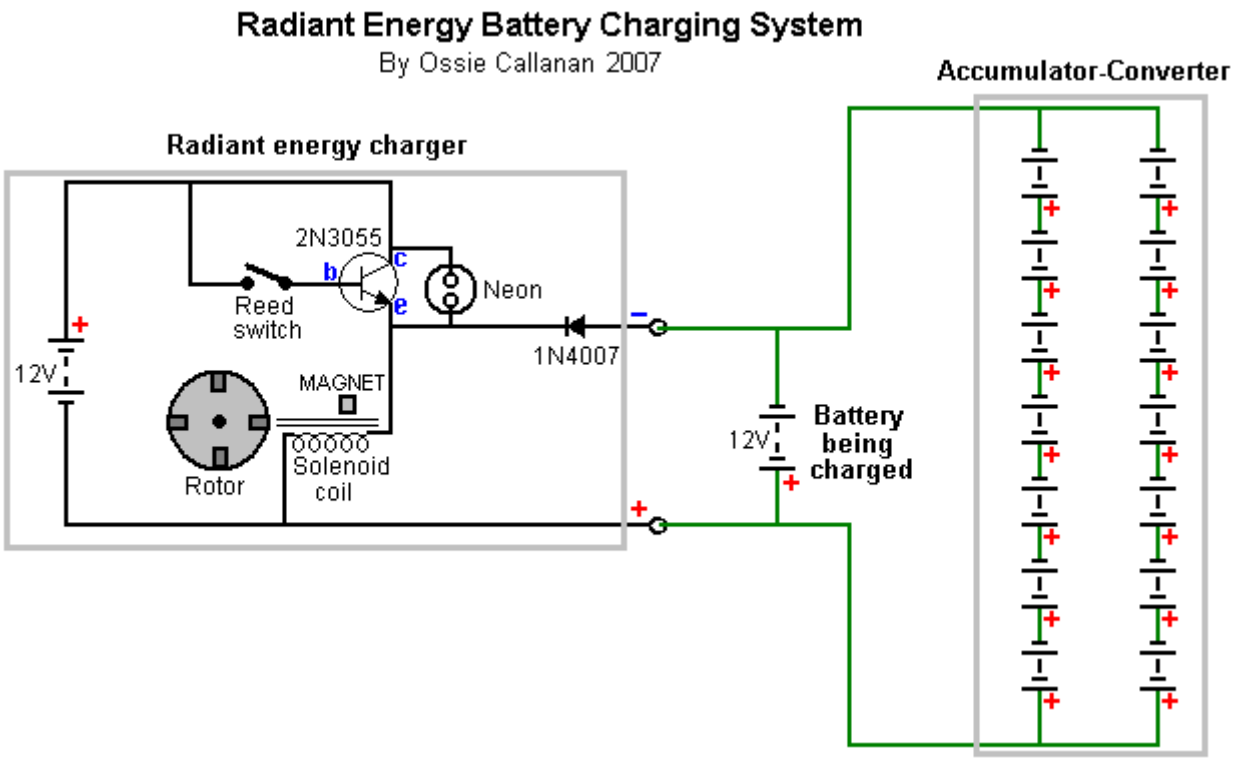
Before I describe the radiant-energy accumulator-converter, I want to highlight how important it is to build and experiment with the circuit described above. Only after adjusting and watching the input current and how the charging battery charges can you truly see how these radiant energy pulses are affecting the charging battery. In terms of why it works or how it provides so much radiant energy you have to understand that if the reed switch stays closed, which will happen when they get worn and start to stick, with the 9 millihenry crossover choke solenoid coil (bought off the shelf) which I use, the transistor is turned fully on and at virtually it's lowest resistance and so the current draw is about 6 to 8 AMPS for a single coil circuit!!! It seems this is a key requirement for generating these radiant energy pulses with a transistor. You have to switch the transistor fully on at maximum current for the coil and input voltage. This is another thing the SSG does not do well. But yet, in this circuit, when the reed switch is adjusted correctly, you can get the input down to just a few milliamps if you want!

OK, now to the radiant accumulator-converter. The requirement for this appears to be because the charging battery is not very efficient at absorbing all of the radiant energy pulses. John Bedini had related this as being due to impedance matching but I am not sure of this at this stage, but impedance may be a factor. Because the charging battery does not absorb much of the radiant energy by itself, you MUST have an accumulator-converter to absorb and convert the radiant energy for the charging battery to use. Ok, that said, what is a radiant energy accumulator-converter?

A Radiant Energy Accumulator-Converter (“REAC”) is nothing more than a dipole! But the larger the dipole the better! The dipole can be a battery but that is silly when we are already charging a battery. Well in that case, it is a special battery. It is a battery that is mostly composed of potential but little current. The higher the potential the better the accumulation/conversion but some current is still required to be able to pass the energy back and charge the charging battery.

There are a number of traditional dipole's that fit what is needed. There is a simple long and raised wire antenna and ground, but this will not provide enough back current to charge our battery. There is an earth-battery setup, but unless you want to put in the effort and amount of materials to raise the voltage and still have some required current this needs a lot of work and materials. Finally, I found the best compromise to be "old, dead, sulphated lead acid batteries". At this stage of my research the condition of the battery does not really matter as long as it is “dead” and sulphated. As long as they are old and dead so that they can barely light up a 12V 100ma bulb, then they will do just fine. Gee I am glad I never threw away my old dead batteries that kept piling up.

If you go down to a battery recycler or junk yard, you can buy pallet loads of old and dead Uninterruptible Power Supply (“UPS”) batteries for very little cost. When I say pallet load, I mean pallet load. The bigger the bank of these you get the batter. Connect them both in series and parallel so that if they where good, you would get anywhere from 48-120 volts. When connecting in parallel make sure that each 12 volt segment has an approximately even capacity in Amp-Hours. You can put this bank under your house or table or even bury them in the ground. It is no problem as you will never have to do anything to them again (as long as they are sealed). They will NOT keep running down. They are already run down. All you need them for is to use their potential as a dipole and their hidden capacity. The very small amount of current they will provide for the size of the bank due to the crystalline resistance of the sulphation is all that is needed to provide the free energy that the will convert the radiant energy pulses and feed it back into your good battery which is being charged. I believe that these sulphate crystals may indeed be the main component that is doing the radiant energy conversion for us. Now for how to connect up your ‘REAC’. See the following diagram:



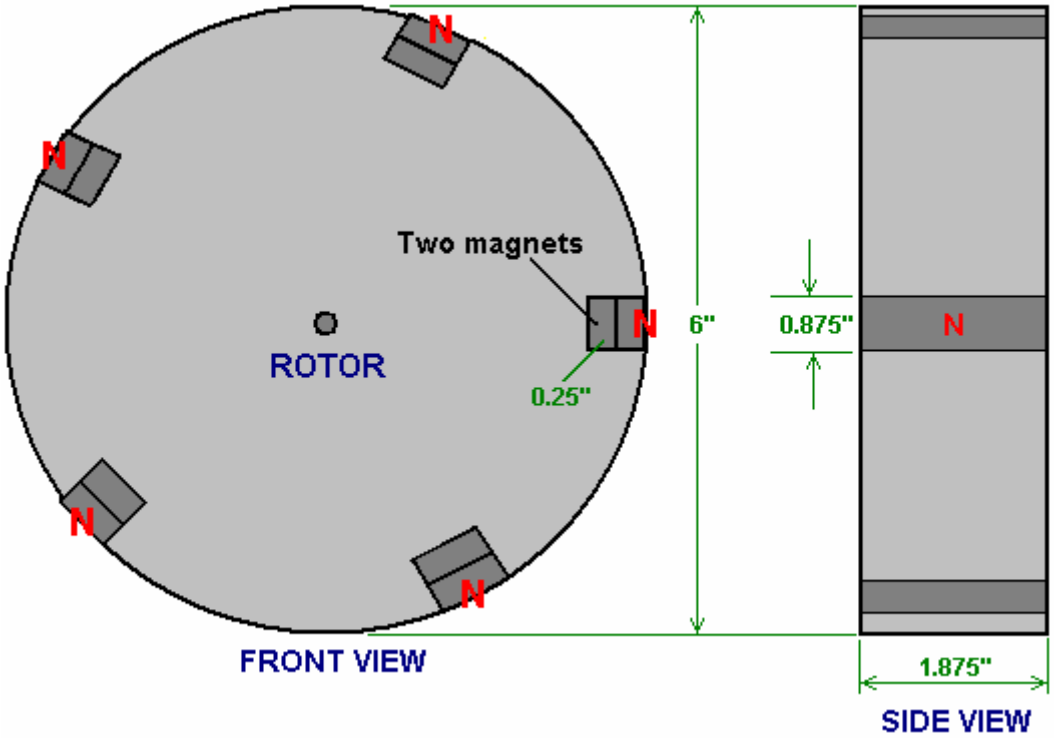
You must connect the REAC directly to the charging battery as I have shown above. Amazingly, there is a great voltage different when you measure the voltage directly across the charging battery compared to the voltage measured across the REAC while the reed motor is running. This voltage difference is seen even with thick cables connecting them, but distance does affect it as well. You must have two separate sets of cables. One set going directly from the radiant energy charger to the charging battery and the other set from the charging battery to the REAC. I have run the above setup for over one month now. Below is a picture of the "dead" batteries which I use as an REAC.

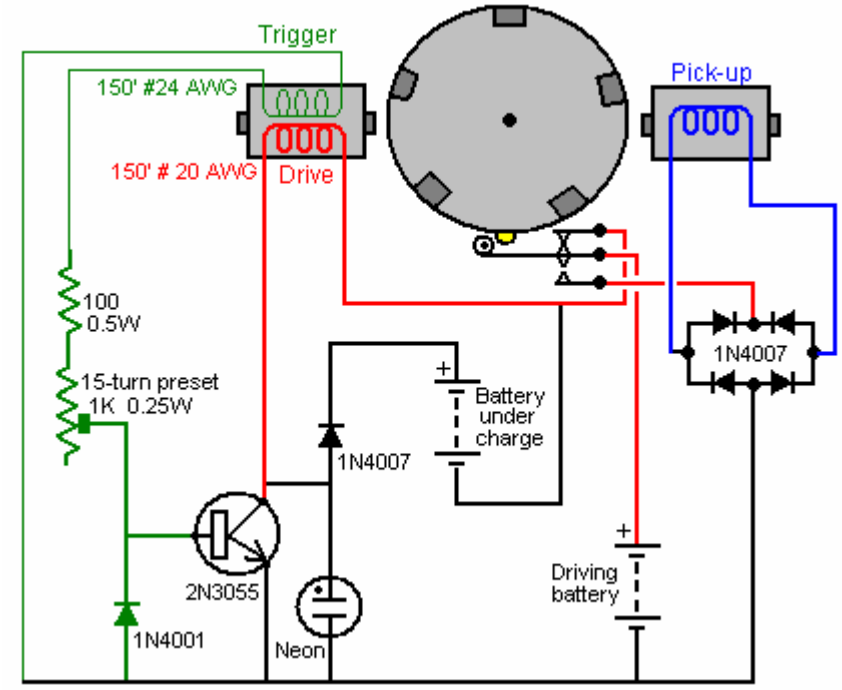
Using my good 33 Amp-Hour UPS batteries, I can charge them up from 10 volts to 14 volts in about 6 hours with the radiant energy reed motor running 4 coils drawing only 600ma. I can then swap the source battery with the charging battery and keep doing this until I have both batteries charged in about 24 hours. I have done this very many times and the charging rate appears to be improving over time.

But one thing I want to make clear. If you think that I am somehow just using the stored energy in the REAC bank, if I do not use my reed motor, then the charging battery will not charge. If I try replacing the reed motor with a regular battery charger, the battery will take as long to charge as a normal battery charger would to charge it. When using the reed motor, the REAC is converting most of the radiant energy and providing the energy back to the charging battery. There you have it, a fully working radiant free energy system. Enjoy! -- Ossie Callanan

The Self-Charging Battery Charger.

One major disadvantage of some of these battery pulse-chargers is the fact that it is thought that it is not possible to self-power the device nor to boost the running battery during the battery charging process. There is one variation of the pulse-charger which does actually boost the driving motor as it runs, and one particular implementation of this is shown here:





The rotor weighs about five pounds (2 Kg) and is very heavy for its size, because it is constructed from flooring laminate, and has a thickness of 1.875 inches (48 mm) to match the width of the magnets. There are ten magnets size 1.875" x 0.875" x 0.25" (48 mm x 22 mm x 6 mm) which are assembled in pairs, to produce the most evenly matched magnetic sets possible. That is, the strongest is put together with the weakest, the second most strong with the second weakest, and so on to produce the five sets, each half an inch (12 mm) thick. These pairs are embedded in the rotor at equal 72^{O} centres around the edge of the rotor.

The battery pulsing produced by this circuit is the same as shown in John Bedini's patent already mentioned. As the rotor turns, the trigger winding energises the 2N3055 transistor which then drives a strong pulse through the winding shown in red in the diagram above. The voltage spike which occurs when the drive current is suddenly cut off, is fed to the battery being charged. This happens five times during a single revolution of the rotor.

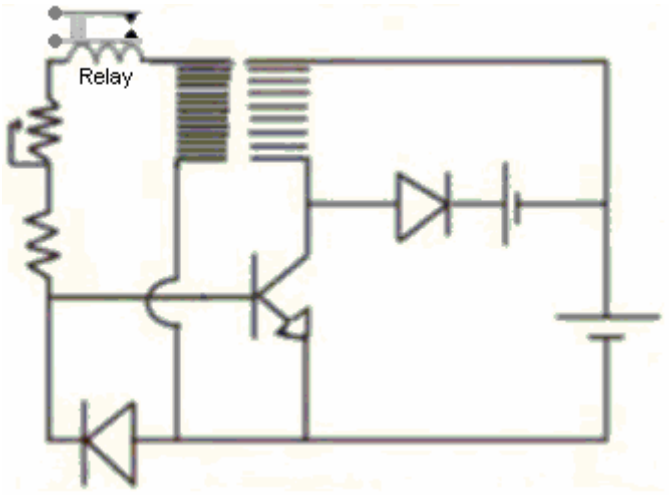
The clever variation introduced here, is to position a pick-up coil opposite the driving/charging coil. As there are five magnets, the drive/charging coil is not in use when a magnet is passing the pick-up coil. The driving circuit is not actually active at this instant, so the micro switch is used to disconnect the circuit completely from the driving battery and connect the pick-up coil to the driving battery. This feeds a charging pulse to the driving battery via the bridge of 1N4007 high-voltage diodes. This is only done once per revolution, and the physical position of the micro switch is adjusted to get the timing exactly right.

This arrangement produces a circuit which in addition to pulsing the battery bank under charge, but also returns current to the driving battery.

Another variation on this theme is shown on YouTube where an experimenter who calls himself "Daftman" has this video explaining the circuit he uses in his Bedini-style battery-charging motor: http://uk.youtube.com/watch?v=JJillOTsmrM&feature=channel and his video of his motor running can be seen at: http://www.youtube.com/watch?v=S96MjW-isXM and his motor has been running for months in a self-powered mode.

The Relay Coil Battery Charger.

One experimenter on the Energetic Forum has posted a video of his adaptation of the Bedini circuit at http://uk.youtube.com/watch?v=4P1zr58MVfI. He has found that adding a 6-volt relay coil into the feed to the base of the transistor has halved the power used and yet keeps the rotor at about the same rate of rotation. The circuit is shown here:

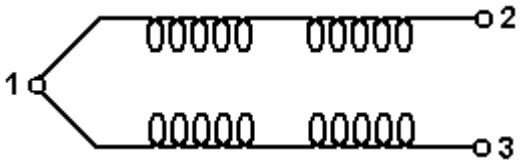


The build used has three electromagnet coils placed around a horizontal rotor:

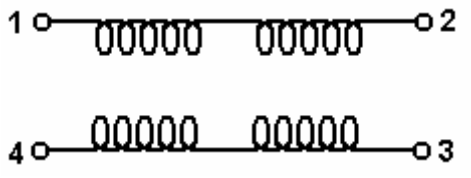
The Modified Fan Battery Charger.

Other more simple methods of getting this radiant energy charging of batteries are also available. One simple method is to skip most of the mechanical construction and use a slightly adapted synchronous fan. This method is shown by "Imhotep" in his instructional video which is located at http://uk.youtube.com/watch?v=eDS9qk-Nw4M&feature=related. The original idea comes from John Bedini and the fan idea from Dr Peter Lindemann.

The most common choice for the fan is a computer cooling fan - the larger the better. These fans usually have four windings connected like this:

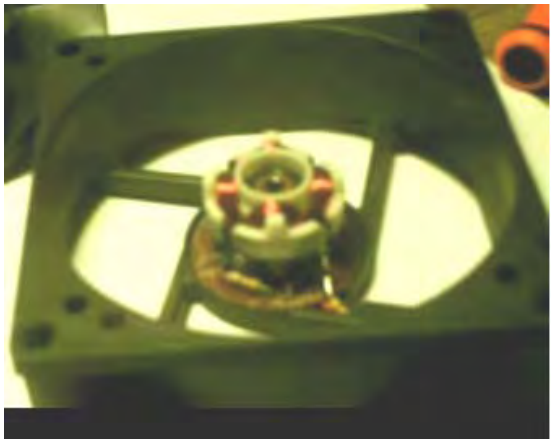


To use these windings as both drive and pick-up coils, the fan is opened up by lifting the label covering the hub of the fan, removing the plastic clip holding the fan blades on the spindle and opening the casing to expose the coils. The wire post with two wires going to it then has one wire removed and a fourth post improvised by drilling a small hole and inserting a short length of wire from a resistor. The fourth wire end is then soldered to it to give this arrangement:

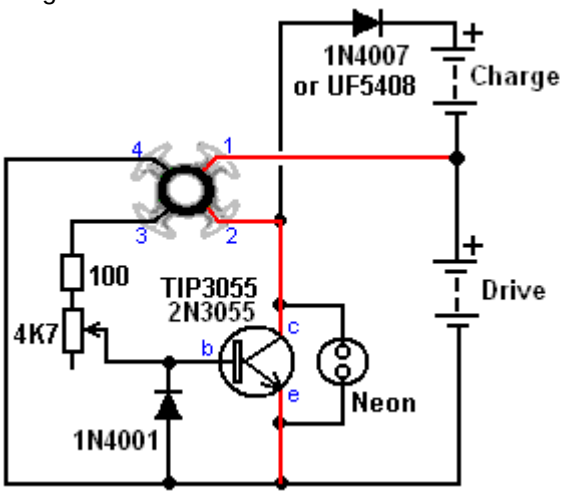


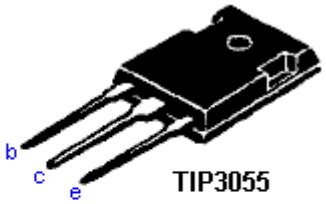
This produces two separate coil chains: 1 to 2 and 4 to 3. One can then be used as the drive coil and the other as the power pick-up coil which passes the very short high voltage pulses to the battery which is being charged.

When opened up, the fan looks like this:



And the circuit arrangement is:





The fan is started by hand and then continues to spin, working as a fan as well as charging a battery. The current draw from the driving battery is very low and yet the radiant energy charging of the other battery (or battery bank) is not low. Please remember that batteries which are to be used with this radiant energy, need to be charged and discharged many times before they become adapted to working with this new energy. When that has been accomplished, the battery capacity is much greater than specified on the label of the battery and the recharging time also becomes much shorter. The circuit is adjusted with the variable resistor, which changes the transistor drive current, which in turn, alters the speed of the fan. It should be stressed that this device and the relay charger shown below, are simple demonstration devices with small coils and to get serious charging, you need to use a large-coil battery pulsing systems with a bank of lead-acid batteries being charged.

This circuit is a clever implementation of John Bedini's Simple Schoolgirl ("SSG") design. As it can be a little confusing to know which of the four wires coming out of the modified fan to use, let me explain how they operate. You now have two pairs of series-connected coils inside the fan. An ohm-meter (or battery and bulb) will let you see which of the four wires are the two ends of each of those coils. The coils are symmetrical and so it does not matter which coil feeds the base of the transistor and which coil is driven by the collector of the transistor. It does

not matter either, which way round the coil feeding the base of the transistor is connected but it does matter very much, which way round the drive coil is connected. Connecting it the wrong way round will not cause any harm but the fan will not operate because instead of the coil repelling the rotor magnets and pushing them on their way, it will attract them and oppose the rotation. So, if the fan does not spin when you give it a push, swap the drive leads over and it should work perfectly.

The neon bulb protects the transistor but it also give a good indication of how well the battery being charged is being fed. Adjust the variable resistor to get the minimum current draw from the drive battery while still having the neon lit well and that should give a good performance.

A very neat build of an 80 mm computer fan conversion to a pulse charger built by Brian Heath is shown here:

This unit runs with a PP3 9V battery as the drive battery, and charges a PP3 9V rechargeable battery when it is running. Both batteries are enclosed in the box in this very neat construction.

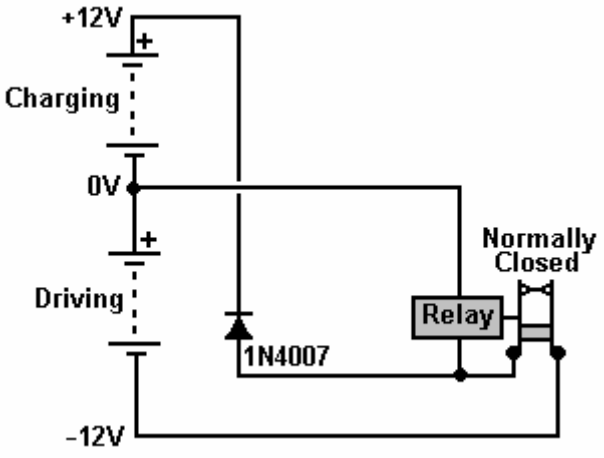
The Automotive Relay Battery Charger.

An even more simple charging method is also shown by “Imhotep” in another of his instructional videos at http://d1190995.domaincentral.com.au/page6.html. Here he adapts an ordinary 40 amp car relay, converting it from having a “normally open” contact, to operating with a “normally closed” contact. It is not necessary for you to do this as automotive relays with “normally closed” contacts are readily available and are not expensive.

The relay is then wired up so that it powers itself through its own contacts. This causes a current to flow through the relay coil winding, operating the contact and opening it. This cuts off the current through the relay’s own coil, causing the contacts to close again and the process starts all over again.

The repeated opening and closing of the relay contacts happens at the resonant frequency of the relay and this produces a buzzing noise. Actually, buzzers were originally made this way and they were used in much the same way as a doorbell would be used today.

The circuit used is shown here:



As you can see, this very simple circuit uses only two components: one relay and one diode. The key feature is the fact that when the relay contacts open and current stops flowing through the relay coil, a very high voltage spike is generated across the relay coil. In transistor circuits which drive a relay, you will see a diode wired across the relay coil in order to short-circuit this high voltage at switch-off and stop the transistor getting destroyed by the

excessively high voltage. In this circuit, no protection is needed for the relay. Any number of batteries can be charged at the same time.

An ordinary 40 amp automotive relay like this:

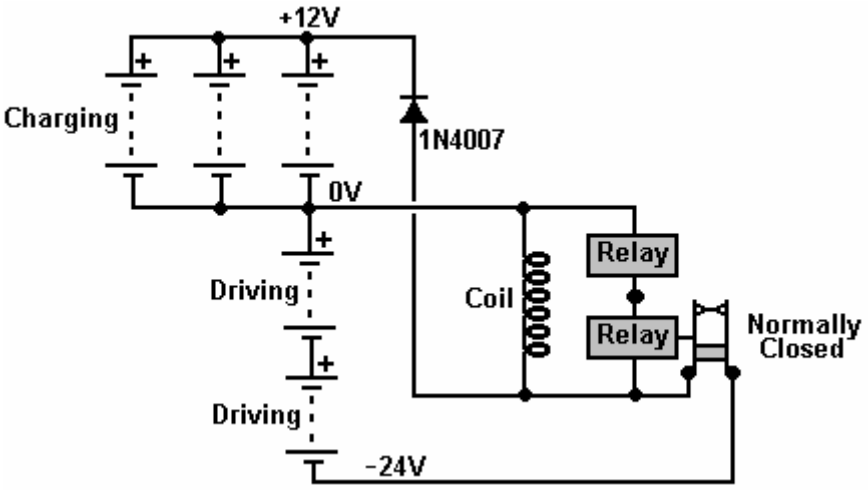
can have a “changeover” contact, which means that it has a “normally closed” contact and so can be used directly without any need to open or modify the relay itself.

In this circuit, however, that reverse voltage is being used in a very productive way. These voltage spikes are very sharp, very short and have a very fast voltage rise. This is exactly what is needed to trigger an inflow of radiant energy from the local environment, into the battery. This battery charging current is not coming from the driving battery but is coming from the environment. The small current from the driving battery is just operating the relay as a buzzer.

One user of this circuit commented that he was using a non-automotive relay with a greater number of turns in the coil, and he found that both of the batteries were getting charged at the same time, but of course, the drive battery was gaining charge at a slower rate. This will have been because the current draw with the more efficient relay was less that the rate of charging and so, even the drive battery gained power.

Please remember that at this time, we have no instrument which can directly measure the flow of radiant energy into the charging battery. The only reliable way of assessing the inflow is to see how long it takes to discharge the charged battery through a known load.

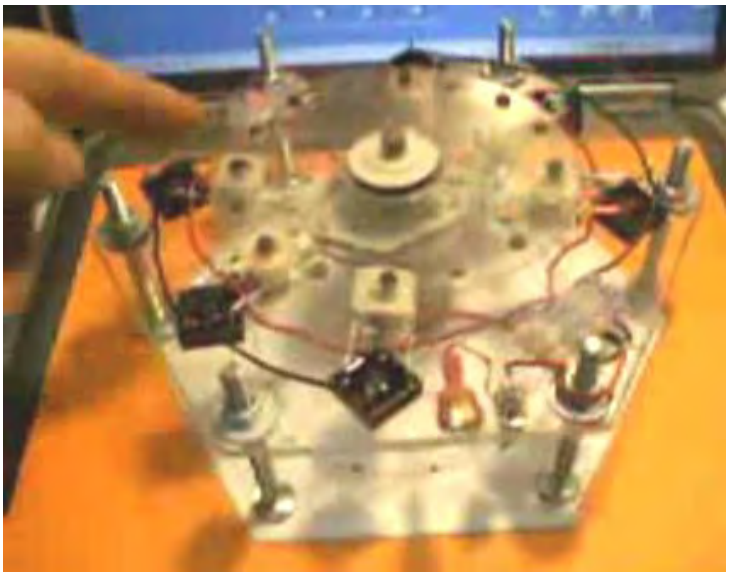
My experience with using relays for battery charging indicates that you get a better result if 24 volts is used to drive the circuit and as vehicle relays don’t have that much of a coil winding, there is a considerable improvement if a large coil is connected across the relay coil or coils as shown here:



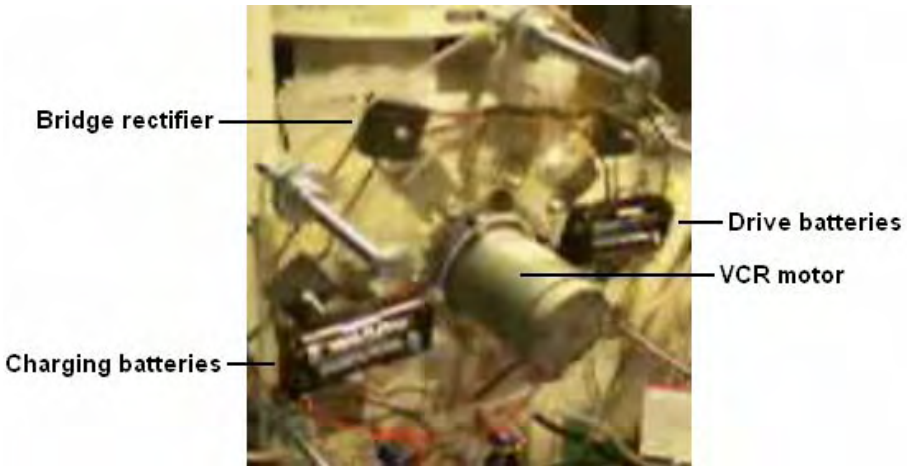
When using one of these relay charging systems you will find that quite a lot of noise is generated. This can be reduced quite easily with a little padding and it does have the advantage of indicating that the charging system is running correctly.

The Self-Charging Motor.

A video at http://www.youtube.com/watch?gl=GB&v=AWpB3peU3Uk&hl=en-GB&feature=related shows an interesting home-built device which uses the motor out of an old video recorder, the bearing out of an old computer CD drive and pick-up coils made by removing the case and contacts from standard relays:



The construction is very straightforward with a simple, uncluttered, open layout:

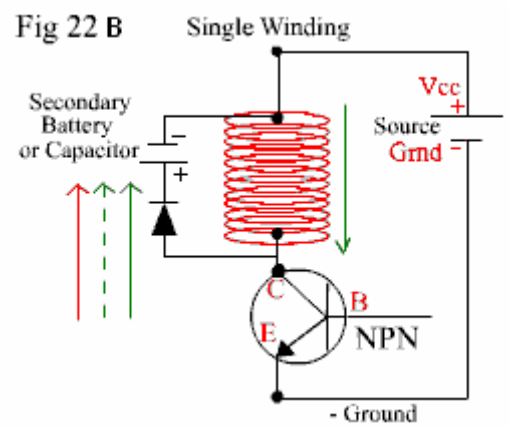


With this arrangement, one pair of AA-size NiCad batteries drives the motor, spinning the motor, moving its magnets rapidly past the ring of converted relays, producing charging DC current via the bridge rectifiers and that current is sufficient to keep the device running continuously.

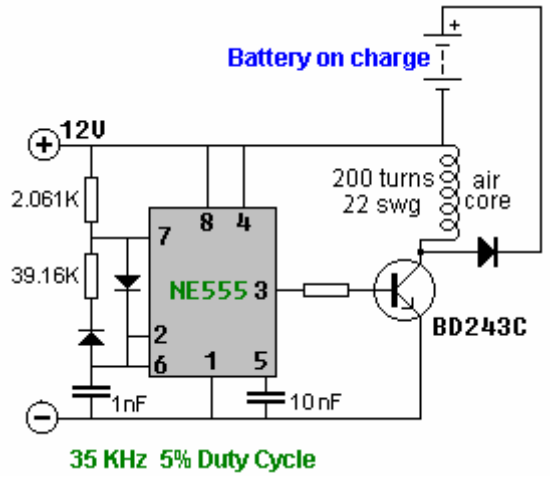
A comment made on the video is that if the ferrite magnets were replaced with neodymiums, then the charging voltage rises to around 70 volts. Unfortunately, the present rotor is too flexible and the neodymium magnets actually flex the rotor down towards the relay cores as they pass, so a more robust rotor is needed.

The 'Alexkor' Solid-State Battery Charger Circuits.

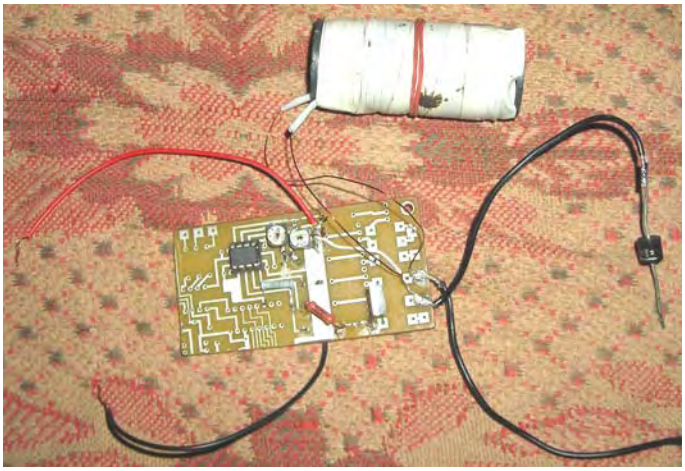
The "Alexkor" battery-charging system is very effective, cheap and easy to build. It is a version of the system described in Fig. 22B on page 7 of the http://www.totallyamped.net/adams/ web page:



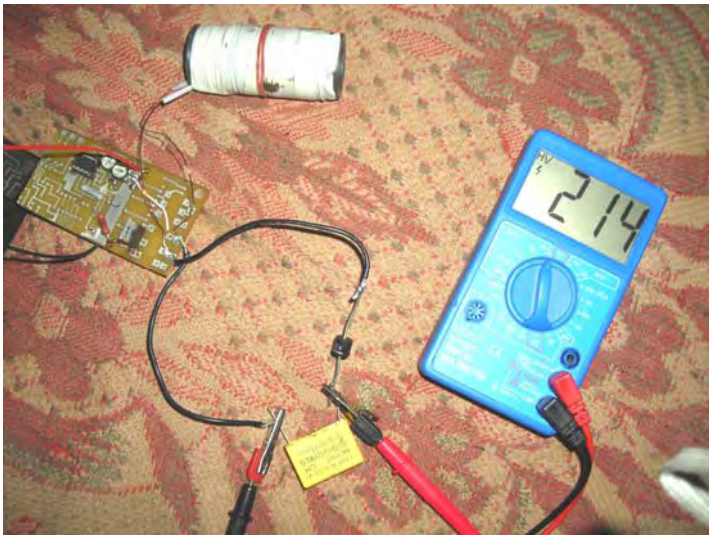
While this description has been around for years, it is part of a discussion on the principles of the operation of EMF magnetic fields and pulsing in coils. 'Alexkor' has developed a practical circuit which he says works very well. It can be constructed as a single unit as shown here:



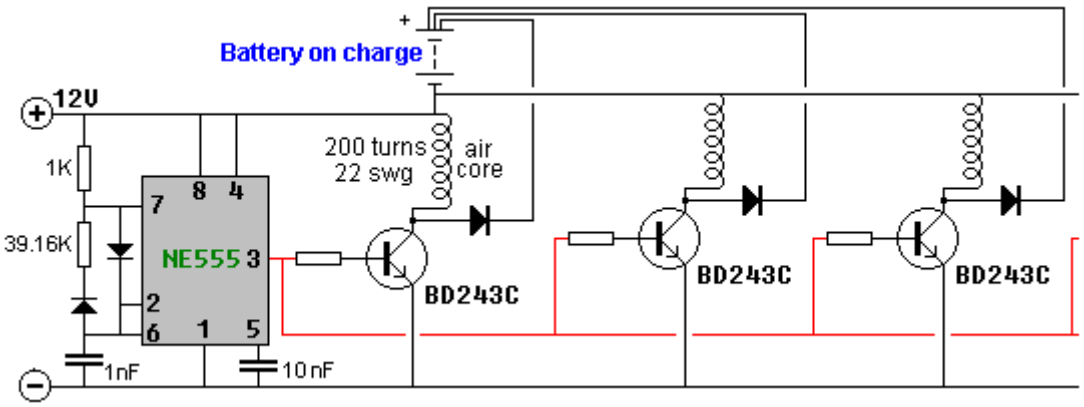
Here, the coil is wound with 200 turns of 0.7 mm enamelled copper wire and the actual construction is compact:



And to get an idea of the performance, Alex uses a capacitor to see the size of the voltage spikes produced by the circuit:

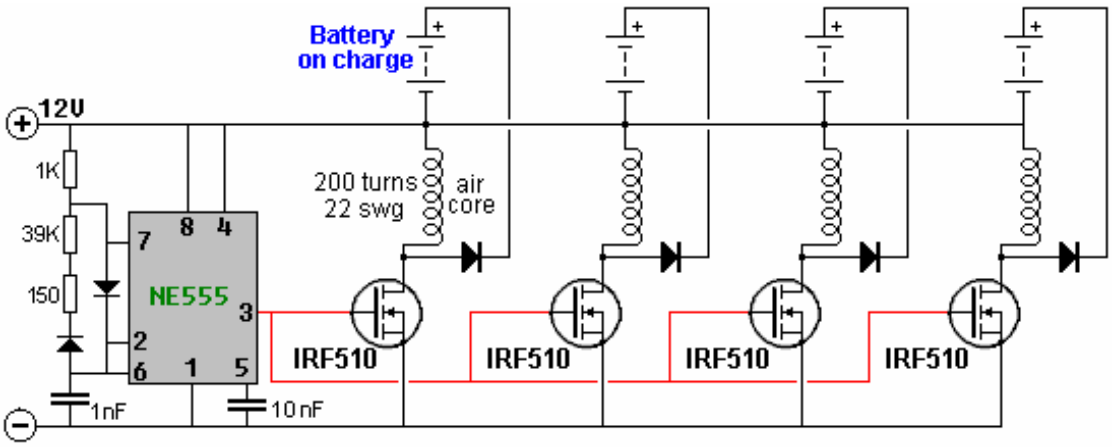


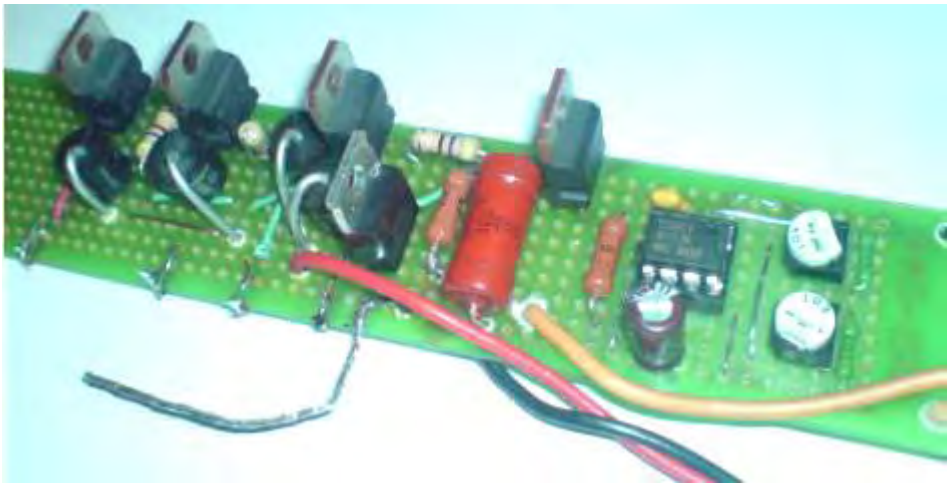
This is the first step in the process as the same circuit can be used to drive many coils of this type. The resistor feeding the base of the transistor is about 500 ohms for the prototype, but using a 390 ohm resistor in series with a variable resistor of say, 1K, would allow a good standard resistor value to be selected for each transistor/coil pair:

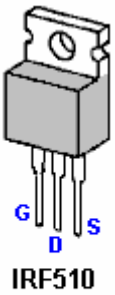


As can be seen from the photographs, Alex uses preset resistors to adjust the settings to their optimum values. The simplicity of this circuit makes it very attractive as a construction project and using more than one coil should make for impressive performance figures. Alex says that the best results are achieved with just the one (1000V 10A) diode and not a diode bridge, which is borne out by the teaching comments on the above web site. Multiple transistor chargers like the one above, work best when there is a separate wire from each coil to the battery being charged.

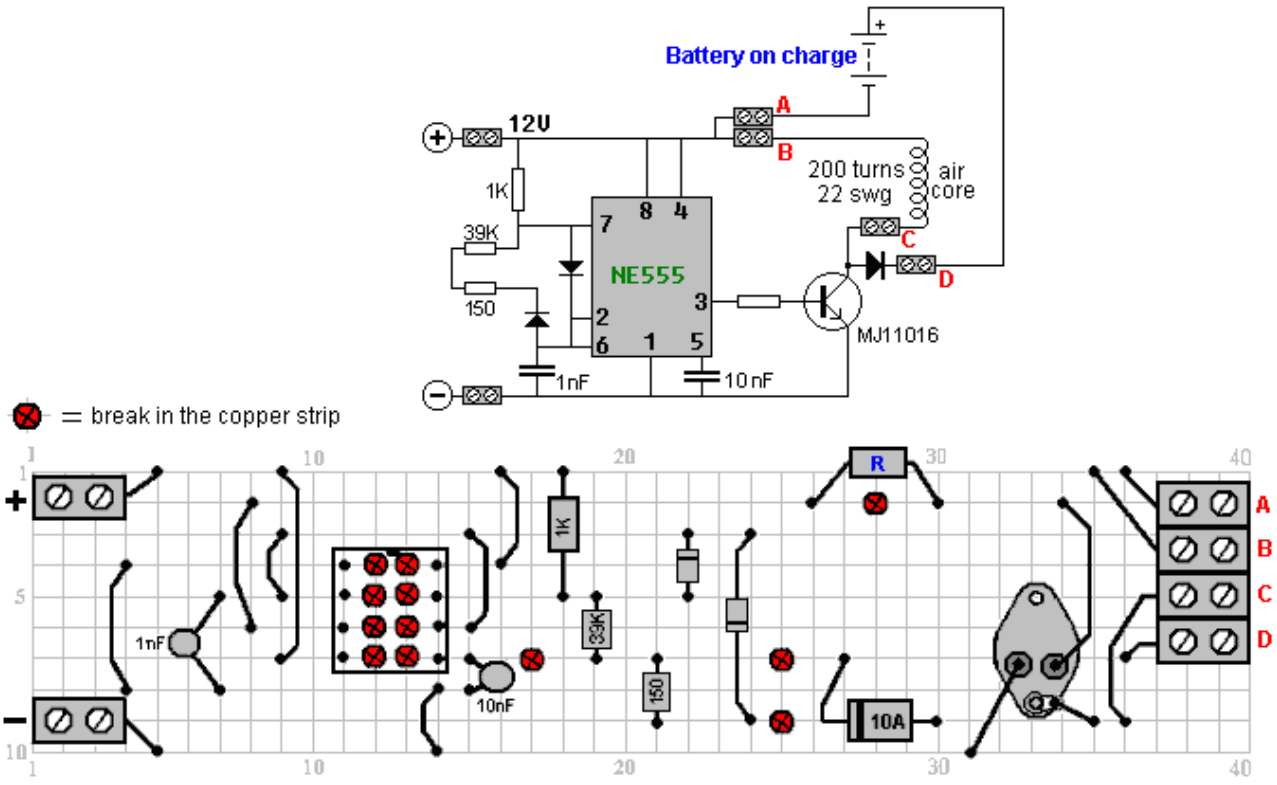
Further development by Alex shows better performance when using the IRF510 FET instead of the BD243C transistor. He also has found it very effective charging four separate batteries and he has revived an old NiCad drill battery using this circuit:





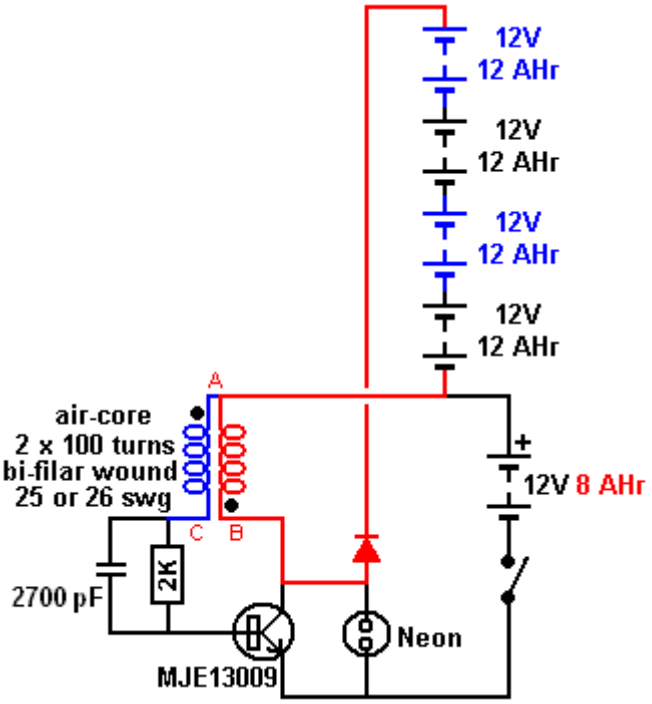


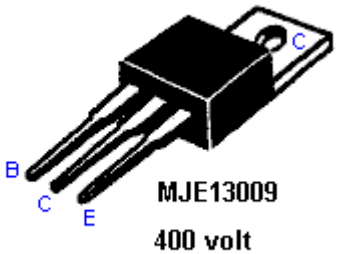
It is possible to use various different transistors with these circuits. As some people have difficulty in working out a suitable physical construction for a circuit, here is a suggestion for a possible layout using an MJ11016 high-power high-gain transistor on stripboard.



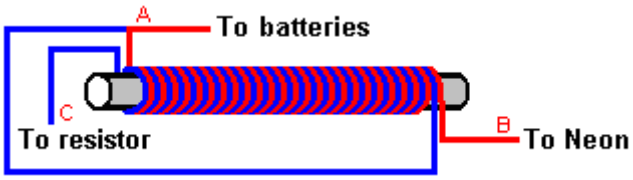
Alexkor's Self-Charging Circuit.

This is a particularly simple circuit which allows a 12V, 8 amp-hour battery charge a 48V, 12 amp-hour battery with radiant energy, in 20 hours using twelve times less current than a conventional charger would. The circuit can charge lithium, NiCad or lead-acid batteries The circuit used is:

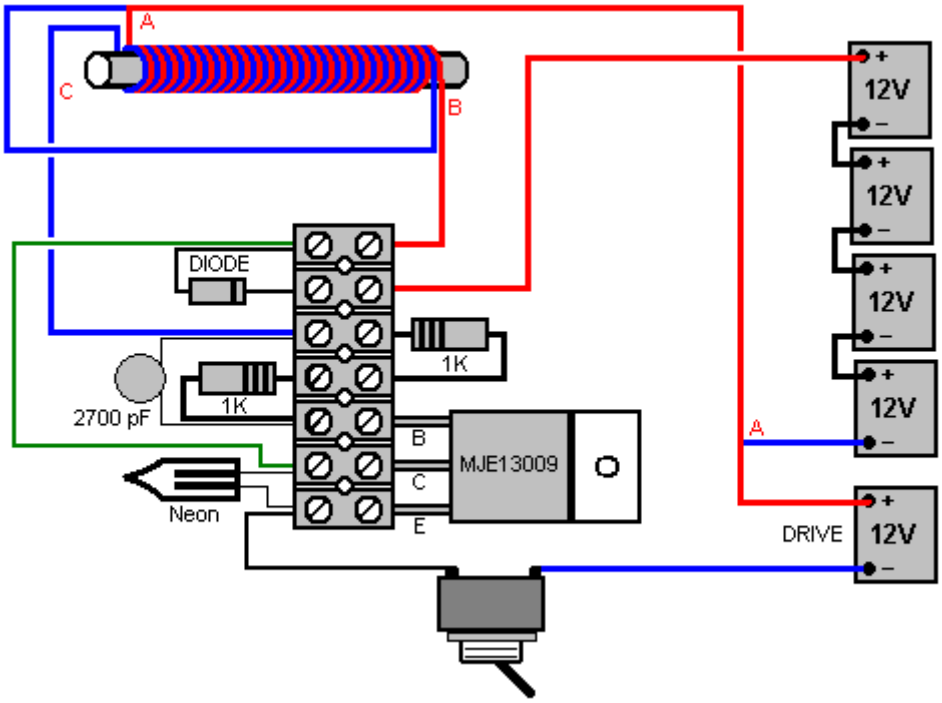




The coil is wound on a hollow former, using two separate strands of wire of 0.5 mm diameter, giving a resistance of just 2 ohms. The strands of wire are placed side by side in a single layer like this:

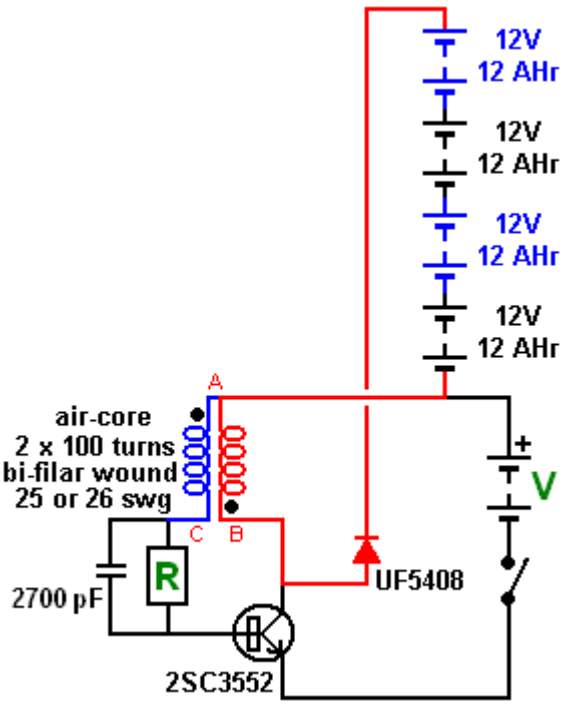


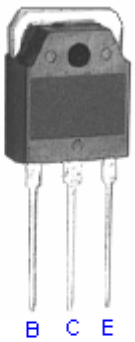
A possible physical layout using a small standard electrical connector strip might be:



If the coil is wound on say, a 1.25-inch or 32 mm diameter plastic pipe, then the outside pipe diameter is 36 mm due to the wall thickness of the plastic pipe, and each turn takes about 118 mm, so around 24 metres of wire will be needed for the 200 turns. If 13 metres (14 yards) of wire is measured off the spool and the wire folded back on itself in a sharp U-turn, then the coil can be wound tightly and neatly with close side-by-side turns. A small hole drilled at the end of the pipe allows the folded wire to be secured with two turns through the hole, and the 200 turns will take up a length of about 100 mm (4-inches) and the two loose ends secured using another small hole drilled in the pipe. The starting ends are cut apart and the ends of each coil determined using a continuity test.

An even more advanced circuit from Alex has even higher performance by using a high-speed transistor and a very fast-action diode, and a neon is not needed to protect the transistor:





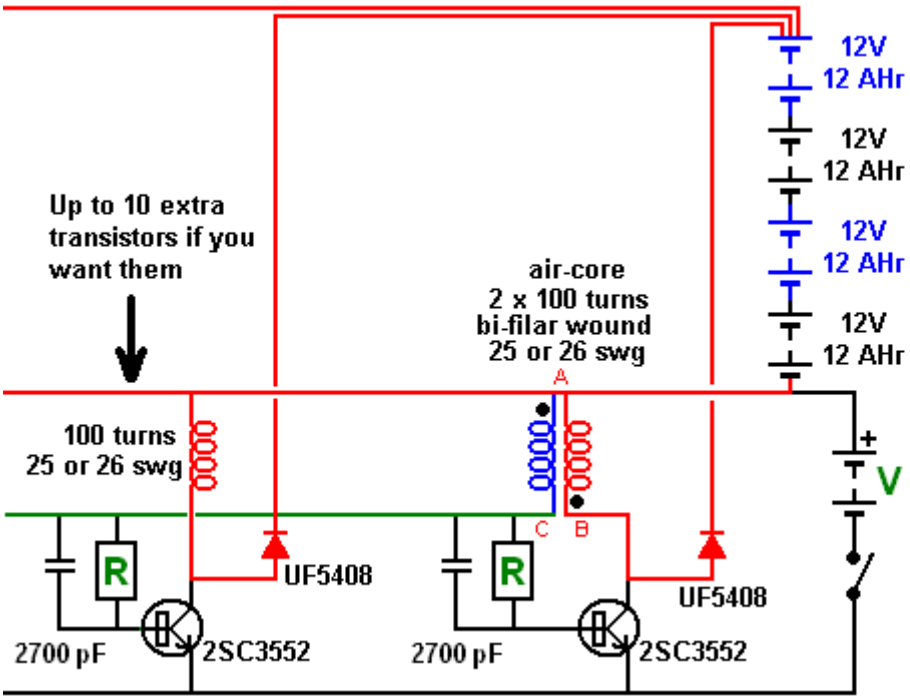
2SC3552

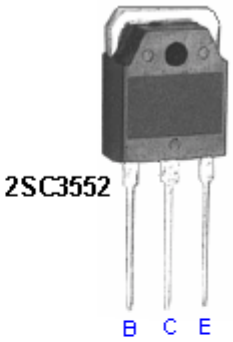
500 to 1100 volt
12 amp continuous
150 watt
Fast-acting

V	12	24	36	48
R	1K	2K	3K	4K

The fast UF5408 diode used in this circuit is available, at the present time, on www.ebay.co.uk in packs of 20 for £3.84 inclusive of postage.

The transistor drive to the battery bank can be replicated for additional drive and an additional ten transistors could be used like this:



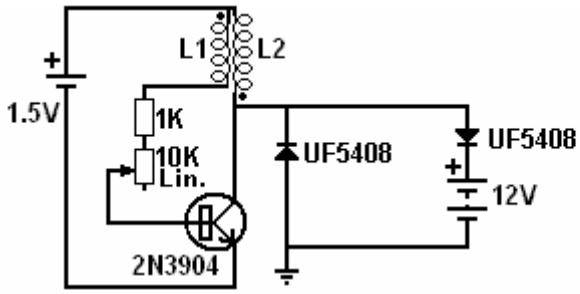


500 to 1100 volt
12A continuous
150 watt
Fast-acting

V	12	24	36	48
R	1K	2K	3K	4K

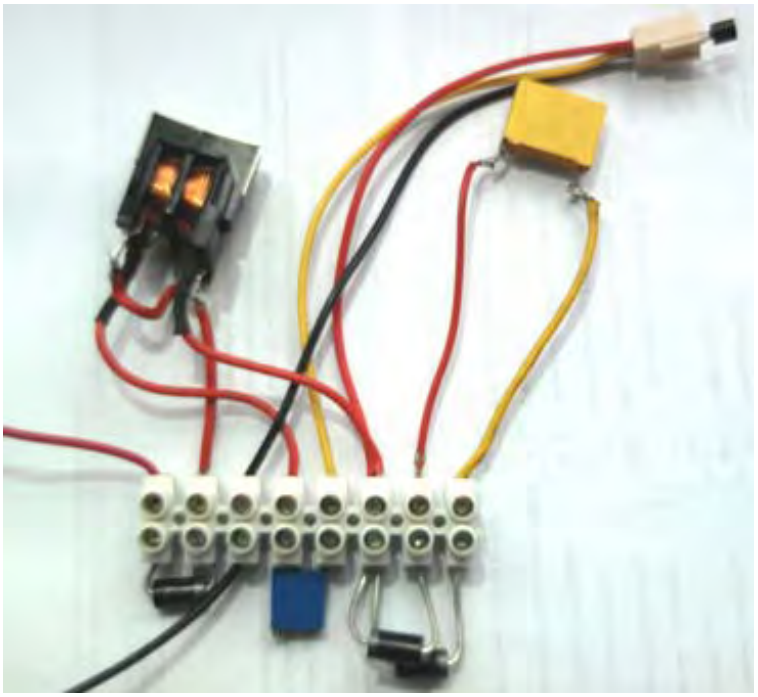
The 2700 pF capacitor is recommended for each additional transistor, but it is not an essential item and the circuit will operate ok with just the one on the bi-filar coil drive section.

A recent circuit design from Alexkor uses the tiniest of inputs; just 1.5 volts at a current which can be adjusted down from 4 milliamps to just 1 milliamp. This tiny circuit can charge a 12-volt battery, although admittedly, the charging rate is not very high as it takes ten hours per Amp-Hour to charge the battery. However, it is spectacular to get a input of just 1.5 milliwatts to charge a 12V battery. The circuit has very few components:



Coils: 0.5 to 1.0mm diameter solid copper wire length: 1 to 2 metres bi-filar wound

Variable resistor adjusted for minimum current 1 to 4 mA
Output is 40V spikes



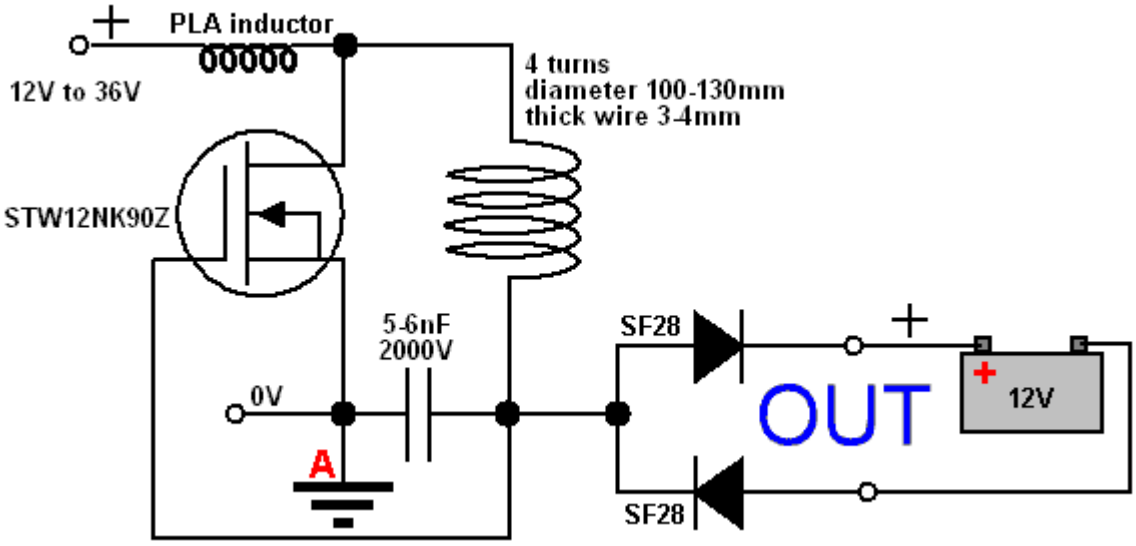
The coil is tiny, bi-filar wound on ferrite or with an air-core. In the circuit diagram, the dots on the coil windings indicate the start of the two side-by-side windings. This makes it clear that the start of one winding is connected to the end of the other winding as well as to the positive side of the 1.5V battery. The variable resistor could be omitted and various fixed resistors tried until the 1 milliamp current level is reached. It should be emphasised that there is just one earthing point and it is a real connect-to-the-ground type of connection. Simple arithmetic will show you that if there is a charging current flowing into the battery to charge it, then even with an imagined 100% efficiency of the battery, the battery charge is many times greater than the draw from the battery driving the circuit. The circuit runs at a frequency between 200 MHz and 300 MHz.

Alex uses a commercial "choke" from http://it.farnell.com/murata/pla10an1522r0r2b/choke-common-mode-2x1-5mh-2-0a/dp/9528423?whydiditmatch=rel_3&matchedProduct=3532290 as shown here:

Type	PLA
Execution	with a single rail
Rated current, a	2
Rated voltage, v	300
Winding inductance, mH	1.5
Active resistance, Ohm	1500
Hull length, mm	18

Jes Ascanius of Denmark has replicated this circuit and he makes these comments: The 10K variable resistor and the additional 1K resistor need to be 250 mW types as larger wattages cause a greater current draw. Also, the quality of the earth connection is important as his very efficient earth produces 60-volt pulses from the circuit (70-volts at night) and just by touching the earth connection can boost those pulses right up to 92-volts and so further experimentation may produce some other interesting effects.

Alexkor's most advanced circuit to date is the one shown here:



This circuit uses the PLA inductor shown above. The initial reaction of somebody familiar with electronic circuits might well be “this is impossible as the battery being charged is ‘floating’ as it is not connected to either side of the driving battery”. While that is true, the circuit works very well indeed and a battery bank of ten 1.2V Ni-Mh batteries rated at 1100 mAHr capacity which had been charged and discharged ten times before, is now charged by this circuit in just half an hour.

The input voltage can be anything from 12V to 36V without the need to change any of the circuit components. The choice of transistor is important and the STW12NK90Z is a very high-performance, high-voltage transistor (available at the present time from www.mouser.com), and while it is not cheap, I would strongly recommend its use if you decide to replicate this circuit. The SF28 diodes are also special components, rated at 600 volts and 2 amps, these are high-speed diodes, not to be replaced with any diode which happens to be available.

The coil is most unusual in that it is just four turns of very thick copper wire, 3 mm to 4 mm in diameter, although aluminium wire can also be used. This power cable is wound on to a spool of 100 mm to 130 mm (4-inch to 5-inch) diameter. The tiny 5 nF capacitor needs to be rated at a very high 6000 volts. The real Earth connection at point “A” gives a 20% to 30% improvement in performance but if the circuit has to be portable, then it will work with the lower level of performance if the earth connection is omitted and point ”A” is connected to the 0V line of the input battery.

While the coils shown above are air-core to allow high frequency operation, coils, most other coils are generally much more efficient with some form of magnetic core, such as iron-dust or ferrite. While it is not likely to be able to operate at frequencies as high as 35 KHz, a very good material for coil cores is the metal of masonry anchors or "sleeve anchors" which look like this:

This metal is immune to rusting, easy to work and loses all magnetism as soon as the magnetic field is removed. You can confirm this for yourself by placing a permanent magnet on one end of the bolt or the tube and using the other end to pick up a steel screw. As soon as the permanent magnet is removed, the screw falls off as the metal does not retain any of the magnetism from the permanent magnet. These anchors are cheap and readily available from builder's supplies outlets, including those on the internet. It is unlikely that this material could operate at more than 1,000 Hz and the circuit above gains a lot of it's performance from the high speed, fast switching and very short "On" time duty cycle.

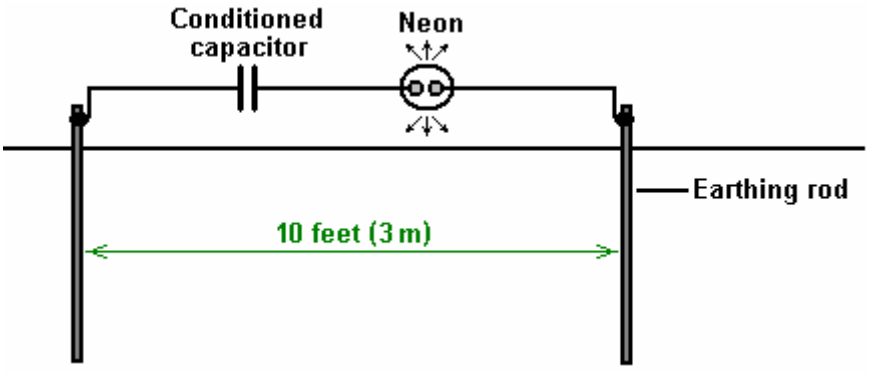
If you use the bolt section of one of these anchors, the conical bump at the end of the shaft will have a delaying effect on the build-up and release of the magnetic field and so it might be advisable to either file it down gently by hand, or to cut off the conical section. There will always be eddy current losses in any solid metal core, but that does not stop them being very effective in operation. As with everything else, testing an actual device is the key to good performance and sound knowledge.

Howerd Halay's Battery Charging Techniques

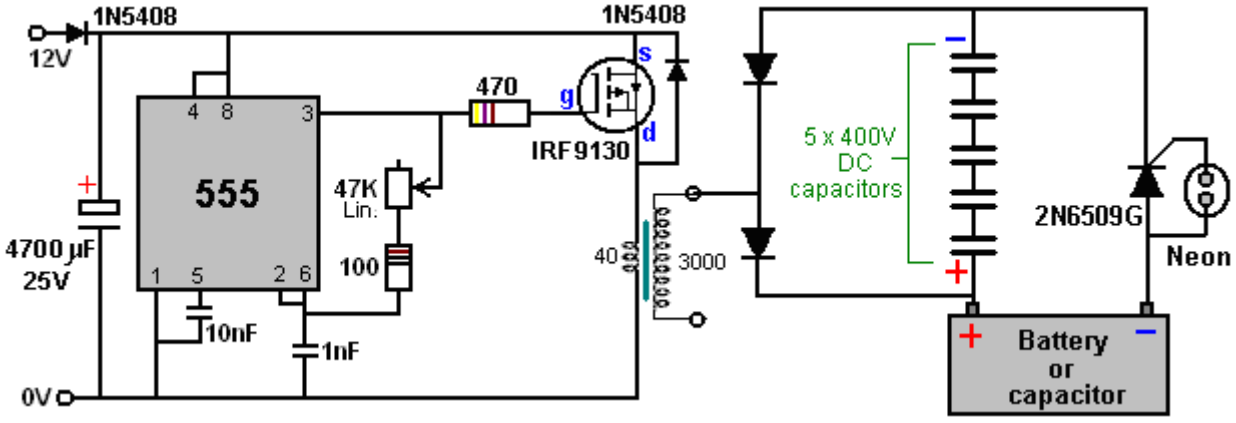
Howerd Halay of the UK stresses the major difference between "conditioned" batteries and all batteries which have not been conditioned. He says: to condition a battery or capacitor, it needs to be repeatedly charged with 'cold' electricity and discharged again. Cold electricity is either high frequency AC electricity or alternatively DC at high voltage. With cold electricity, the electricity flows outside the wires (Steinmetz) and so, Current does not equal Voltage divided by Resistance as Ohm's Law suggests. Instead, Current equals Voltage x Resistance x a Constant "C" which has to be determined by experimentation. It is also possible to get cold electricity from pulsed DC, provided that the DC voltage is over 80 volts. If using that technique, then the sharper and faster the pulses, the better.

When you first pulse an AC or DC capacitor, it behaves normally. After approximately 12 hours of continuous pulsing a change occurs in the behaviour of the capacitor. In the case of the water capacitor, it develops a nano coating on one side only. When measured with a resistance meter it shows no resistance at all. One can say that one side becomes quasi superconducting. In the case of an ordinary capacitor, there is no reason to believe that it behaves differently. The capacitor also charges much faster than before and when the power source is switched off it continues charging! Yes you read that correctly. In my case it fires pulses for up to 3 minutes after the power is switched off, which is why they are dangerous. The firing decays exponentially although I haven't yet tabulated it scientifically – I'll leave that to other people to do.

The result of this is that you can have two identical capacitors side by side. One behaves as if it is plugged into a charger, while the other capacitor behaves normally. All capacitors self-charge to a certain extent but "conditioned" capacitors are in a league of their own! I have tested a neon on a conditioned capacitor through two earth rods 10 feet apart. I gave up looking at the lit neon after half an hour!



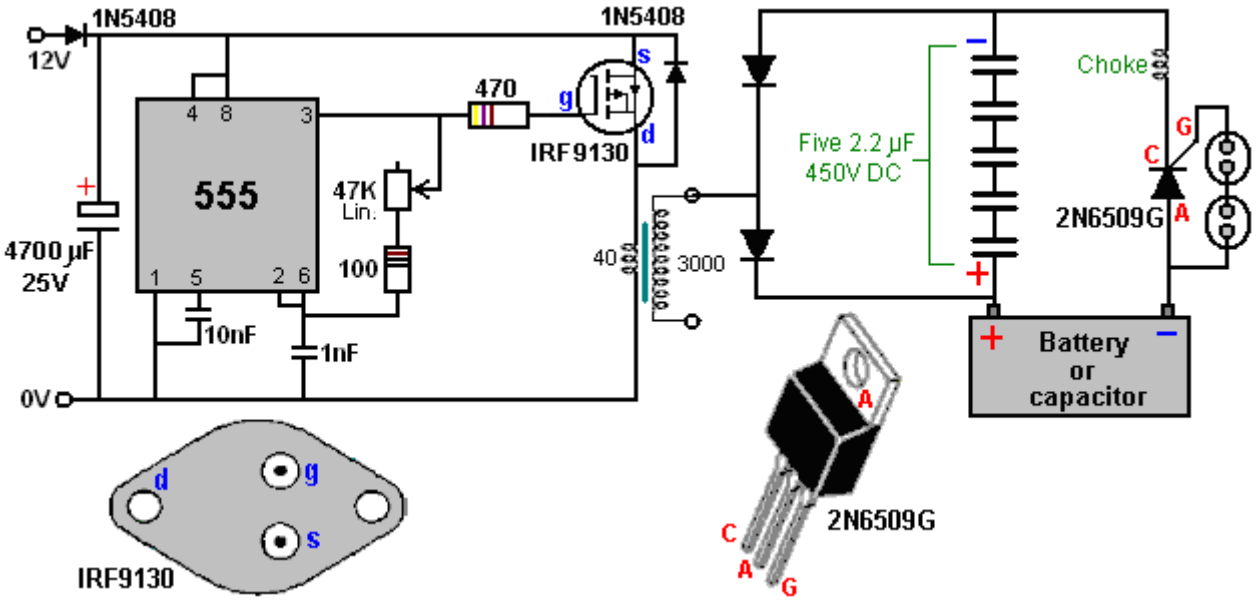
I use a very low-powered high-voltage source with a power output of only 1.2 watts as I like to play safe with these things. With a low power source, I have charged batteries using pulses of up to 800 volts without the batteries showing any ill effects. Also, using one-wire electricity is safer as that transmits mostly voltage and so minimum current is fed. So, to condition a battery or a capacitor using cold electricity, you can use a circuit like this:



Here, the size of the voltage pulses fed to the battery or capacitor to be conditioned, is controlled by the strike voltage of the neon. The ordinary NE2 type neon lamps strike around 90V and so the 2N6509G SCR will feed pulses of about that voltage to the battery or capacitor. If two neons are connected in series and used instead of the single neon shown above, then the voltage pulses will be around 180V. This type of circuit appears to work better if several capacitors are used in series as shown here, as they seem to charge up faster and discharge faster as well. You have to leave the device running for a day to get the full benefit. I regularly charge a 1.6 Kw car battery bank, and after switching off, the battery bank voltage goes up!!

I have also tried 5 seconds of ON time and two minutes of OFF time, and the capacitors continue firing pulses. However the rate of firing is much less when the power is off than when the power is on. If you fail to use the capacitors for a while - in my case it was three weeks or so - you have to start the conditioning process all over again. In my case conditioning them again was harder and seemed to take days rather than hours. The capacitors are COLD. The wires leading up to them and out of them are COLD, but if you get a shock from them, then that shock is HOT !!

Because this charging process uses cold electricity, non-rechargeable batteries can be charged this way. In my case two out of three batteries recover their charge OK, and curiously they charge to a much higher voltage than their rated value. The battery can be replaced with a capacitor. Obviously, any battery or capacitor which is to be conditioned, needs to have be able to be charged with a voltage of not more than 70 volts per neon, so for example, a 96V battery bank would need two neons in series across the SCR of the charging circuit. This circuit will keep on charging the battery for up to three minutes after the input power is switched off. An even more powerful version of the circuit boosts the cold electricity power by using a choke. The neons will light much more strongly. The neons should pulse or you've got a short-circuit. In other words, if the neon(s) is lit continuously, it is a bad sign.



You can use a variable resistor in series with the input power to vary the pulse rate. Negative radiant energy is delivered which produces cold electricity and conditions all capacitors in the output section of the circuit.

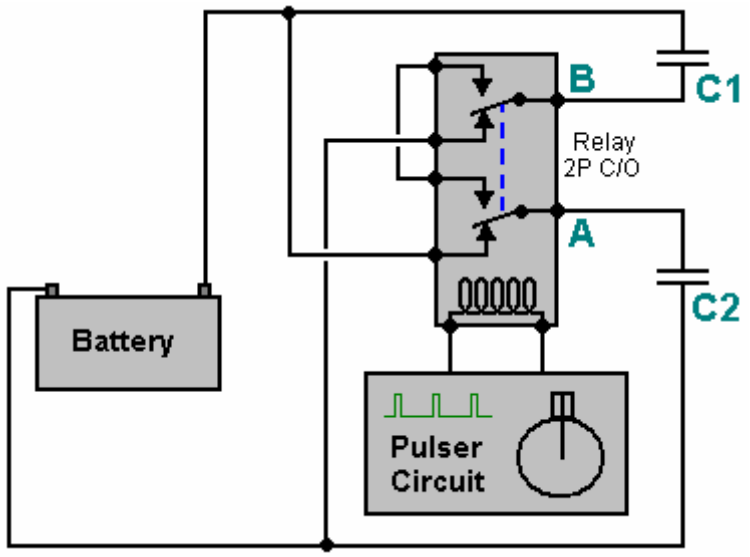
Be very careful with this circuit as it can kill you. This circuit is only for experienced experimenters. Capacitors will take about a day to get conditioned. This circuit is good for bringing dead car batteries back to life. When a battery is conditioned and the charging circuit input power is switched off, the battery will continue charging ! Once they are conditioned, you can charge 4 car batteries in parallel using just a 6 watt 12 volt power supply, or a solar panel. However, this description must not under any circumstances be considered to be a recommendation that you should actually build this circuit as this presentation is for information purposes only.

The Ron Cole One-Battery Switch.

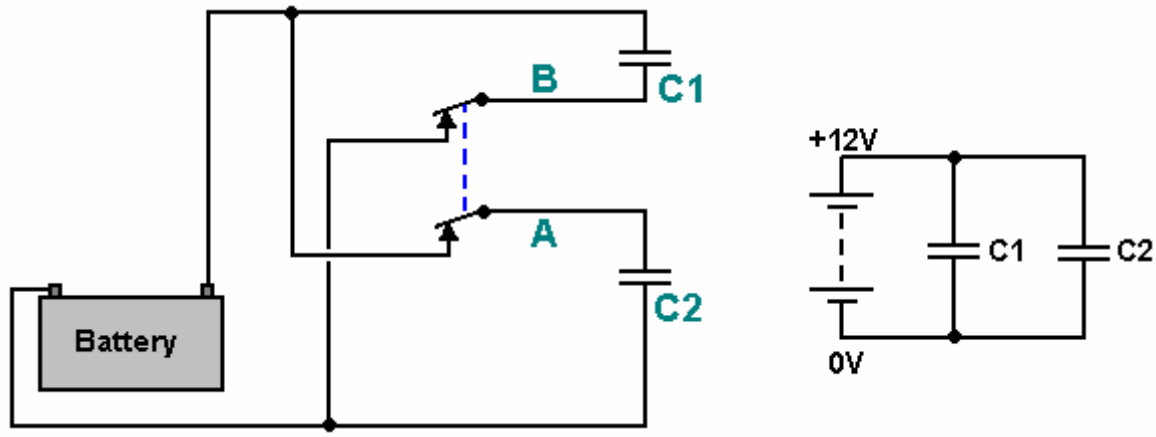
The following circuit is unproven as far as I am aware, but it is an interesting idea. Also, I am not sure if the idea came from John Bedini or from Ron Cole. It has the potential advantage of being a battery charger which operates on its own driving battery. It may also be possible to operate it while it is powering a load. At this time, this is not a fully tested circuit, so please treat it as an idea for experimentation if you are so inclined.

The idea is to use two capacitors which are charged up to the battery voltage and then suddenly connected together to apply twice the battery voltage to the battery. The idea is that the sudden pulse may be sharp enough

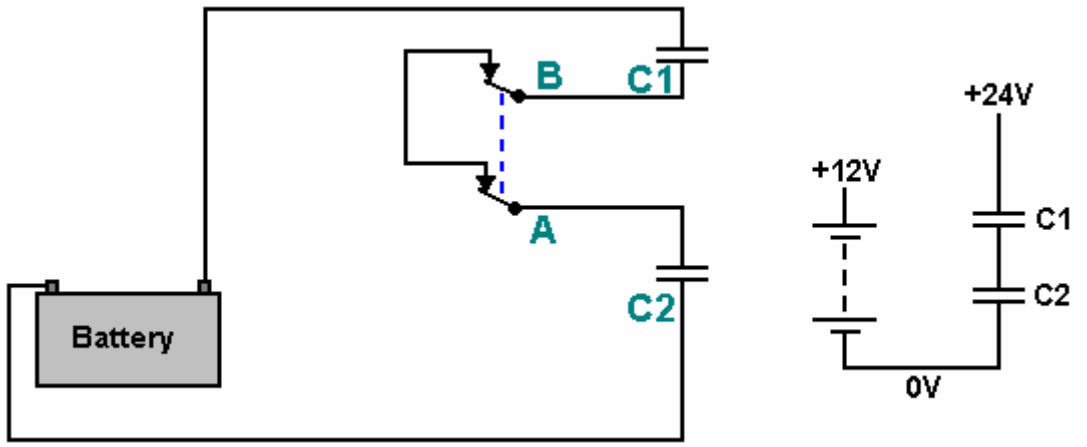
to cause an inflow of radiant energy from the local environment. To be successful, that energy inflow has to be greater than the current draw of the circuit and the capacitors. The circuit is essentially like this:



Here, the pulser circuit is set to give short, very sharp pulses to drive the relay cleanly. The relay has two changeover contacts "**A**" and "**B**". The operation is very simple. Initially, the capacitors "**C1**" and "**C2**" are charged up when the relay is in it's unpowered state and no current is flowing through the relay coil:



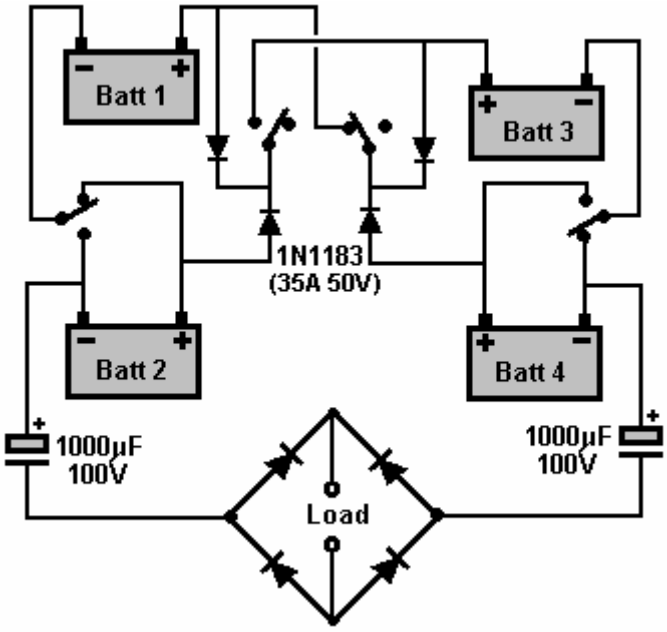
As you can see, the "normally closed" relay contacts have each of the capacitors connected directly across the battery supply rails. This gives the circuit shown above on the right. When the relay is powered up, the situation changes very suddenly to give this arrangement:



Here, the two charged capacitors are disconnected from the opposite supply rails and connected together to form a combined voltage of, in the case of a 12 volt battery, 24 volts connected across the 12 volt battery. This will cause a sudden inflow of current into the battery. However, before practically any capacitor current has flowed, the relay is operated again, repeating the sequence.

The Tesla Switch.

The Tesla Switch is covered in more detail in Chapter 5, but it is worth mentioning it again here as it does perform battery charging. The similarity ends there, because the Tesla switch does the battery charging while the circuit is providing serious current into a load. Also, the Tesla switch uses only four batteries, and still is capable of driving a thirty horsepower motor, which is the equivalent of 22 kilowatts of electrical power.



The simple circuit shown here was used by testers of the Electrodyne Corp. over a period of three years using ordinary vehicle lead-acid batteries. During that time, the batteries were not only kept charged by the circuit, but the battery voltage climbed to as much as 36 volts, without any damage to the batteries.

If the voltage on a battery under load actually increases, it is reasonable to assume that the battery is receiving more power than that delivered to the load (a load is a motor, a pump, a fan, lights, or any other electrical equipment). As this is so, and the circuit is not connected to any visible outside source of energy, it will be realised that there has to be an outside source of energy which is not visible. If the circuit is provided with powerful enough components, it is perfectly capable of powering an electric car at high speeds, as has been demonstrated by Ronald Brandt. This indicates that the invisible source of outside energy is capable of supplying serious amounts of additional power. It should also be remembered that a lead-acid battery does not normally return anything like 100% of the electrical energy fed into it during charging, so the outside source of energy is providing additional current to the batteries as well as to the load.

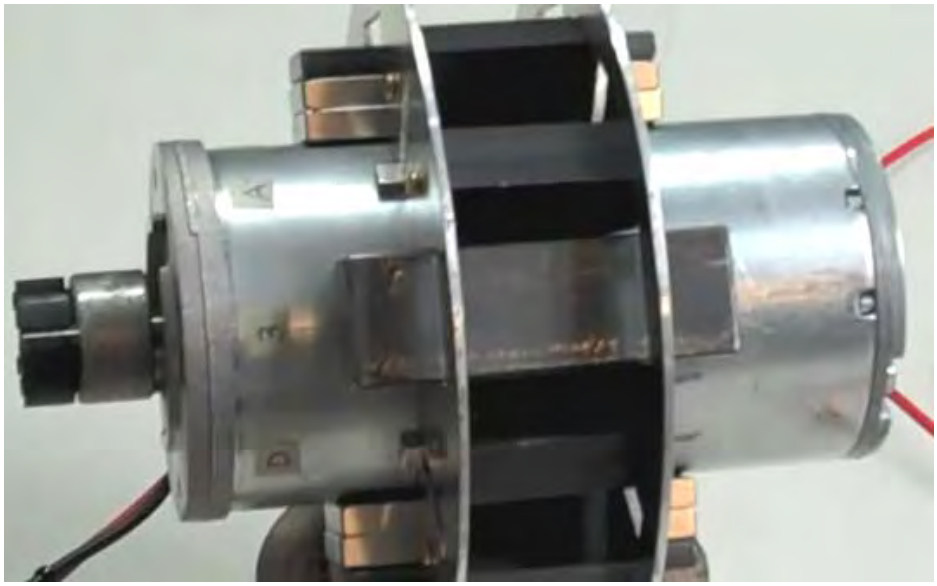
So, how does this circuit manage to do this? Well, it does it in exactly the same way as the battery pulse-charging circuits in that it generates a very sharply rising voltage waveform when it switches from its State 1 to its State 2 (as shown in detail earlier). This very rapid switching unbalances the local quantum energy field, causing major flows of energy, some of which enters this circuit and powers both the circuit and the load. Although it does use four batteries, and the batteries do get charged through the generation of sharp pulses, this is not a circuit which charges massive battery banks so that they can power a load at some later time.

The Self-charging Motor

It is possible to enhance the performance of a DC motor by attaching sets of neodymium magnets to the outside of the body of the motor. One example of this is shown in a video at http://www.youtube.com/watch?v=NoLbphJkxMM&list=LLIpt7ksyRVQi3ITZwSeQxaw&feature=mh_lolz shows such an arrangement and the video at http://www.youtube.com/watch?v=5Xv-req4U8U&feature=related shows how the motor can be set up to produce this enhancement of it's performance. The motor looks like this:

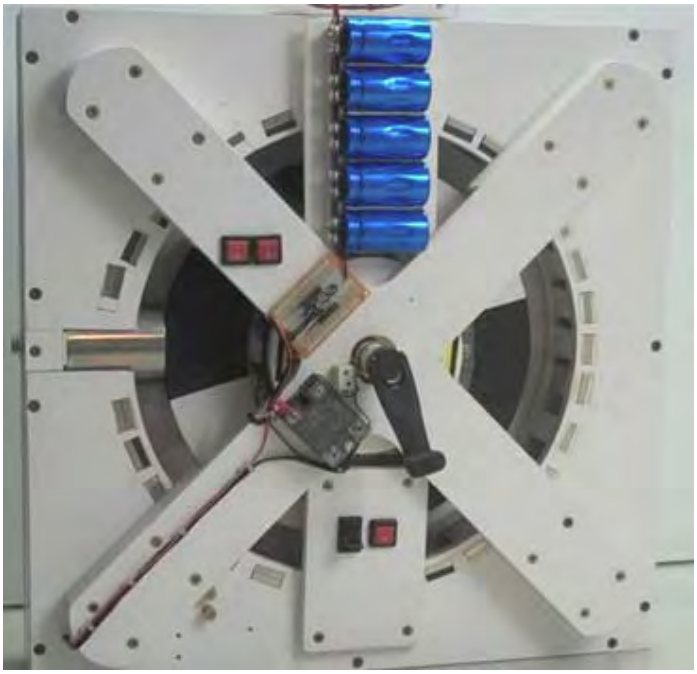
And a frame is used to hold four sets of bar magnets in position around it:

The frame is made from two pieces of 3 mm aluminium with plastic spacers holding the two aluminium discs apart. The plastic spacers are secured with brass screws. As the motor casing is about 3 mm thick steel, that tends to deflect the added magnetic field outwards, which is the reverse of what is wanted. So, a 6 mm thick strip of steel is placed outside the magnets in order to direct the field inwards. The magnets and steel strips are then inserted to complete the arrangement:



The motor is demonstrated when set up like that. Then four sets of very narrow magnets are added in intermediate positions and the number of broad magnets increased to three at each location, the steel strips discarded and four broad magnets used at each of the four locations around the motor:

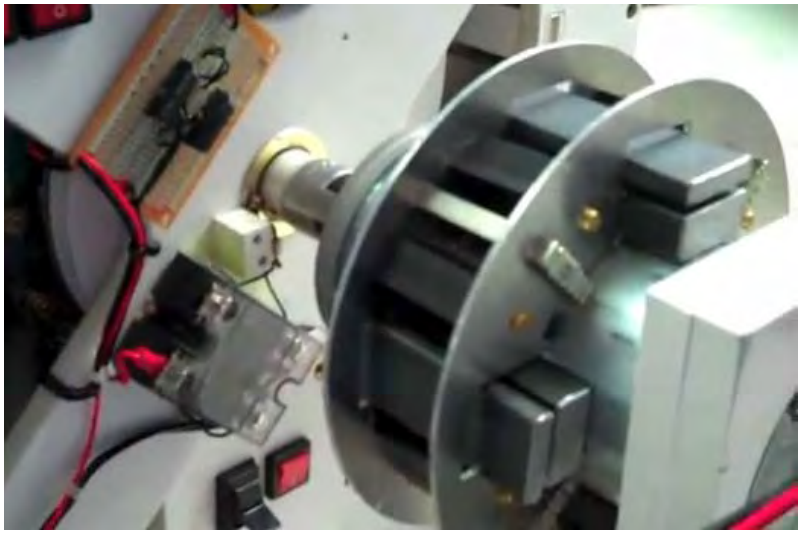
With this arrangement, the motor runs at ten times it's design speed (which would destroy it very quickly), so it is run with just one sixth of it's design voltage. It is used to drive an electric generator shown in two videos at: http://www.youtube.com/watch?v=0dkiHUasERA&feature=related and http://www.youtube.com/watch?v=ZTLvqCiKadI&feature=related which explain the sophisticated design features of this generator which has been built with great care:



Given a couple of minutes of gentle turning of the crank, charges up the five batt-caps sufficiently to run the motor for anything up to two hours. The arrangement is clever with the stator magnets in a Howard Johnston sequence which is also angled inwards like the magnetic Wankel motor. One electromagnet is powered up briefly once per revolution and the back-EMF at switch-off is rectified and passed back to the capacitors, forming a very efficient generator. Other design features are explained in the videos which are definitely worth watching.

The generator, when running can discharge a rapid sequence of high-voltage sparks which the builder just uses for an interesting display. However, those spark discharges are quite capable of charging a bank of batteries (not

to mention, running a Don Smith style power device). The builder then goes on to add an ordinary capacitor to the series-connected batt-cap bank, and couples the motor to it:



This allows the enhanced motor to become fully self-running as well as producing excess power which could easily charge a bank of large batteries. Incidentally, conventional science says that this is "impossible" and so they will only look at this machine with their eyes closed because they can't afford to believe it – after all, that would require them to modify their present theories and the people who pay their salaries would not allow any such change.

This motor is shown running on capacitors, but if it is powered by an ordinary battery and run at it's design speed of 3,300 rpm instead of the extremely low speed shown, then it should be able to charge up a serious bank of large batteries, one of which could then be used to power it for the next charging run.

The 'UFOpolitics' Motor

In chapter 2, there is a section on how an experimenter alters the wiring inside DC motors. This alteration has a major effect, boosting the output power by a major amount as well as providing an additional generated electrical output. One person who followed his instructions and rewound a tiny 3-volt 3-pole motor, then tried running that motor on a discharged 6-volt battery. The motor ran, slowly at first and then picked up speed. That seemed impossible as the motor draws 300 milliamps when running and the battery was just not capable of providing that current. However, impossible or not, the motor ran and not only did it run but it started recharging the battery driving it. This suggests that this is a self-sustaining system which provides output power and yet never needs to have the battery recharged.

The way that this happens is that what we think of as "electricity" is actually a more complex thing called "electromagnetism". We think of electricity and magnetism as being two different things, while in fact, they are two different aspects of the single entity electromagnetism. The electric component always has an efficiency of less than 100% which the magnetic component always has an efficiency which is greater than 100% - something which we usually don't notice as we tend to ignore the magnetic component.

In the case of this tiny 3-volt motor, it draws it's running 'hot electricity' current from the battery. That causes rotating magnetic fields inside the motor and these generate 'cold electricity' which flows back along the supply wires and charges the battery.

Some Battery Charging Suggestions

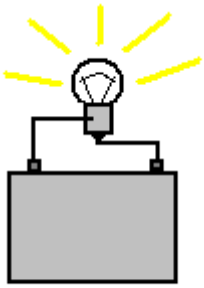
Using a vehicle to charge batteries.

It is very easy to overlook the obvious. It is possible for people who do a fair bit of driving, possibly to and from work, or maybe dealing with a delivery round, to charge some additional batteries from the car alternator. Once the engine is started, additional batteries can be connected in parallel with the car battery, that is, the plus terminal of the extra battery or batteries connects to the car battery plus.

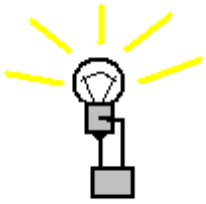
This does draw some additional energy from the car engine and in theory should cause some additional fuel to be used, but the extra fuel should be fairly minor as most of the power of the engine is used to push air aside as air resistance goes up with the cube of the speed of the vehicle. The charged batteries can be removed from the vehicle in the evening and used to power lights, TVs, DVD players, etc. in the evening by using a small commercial inverter.

Experimenting with a battery charger.

There is an interesting possibility when using one of the pulse battery chargers shown earlier in this chapter, say, an Alexkor solid-state pulser. Commonsense tells you that if you want to light a lamp using a battery, then the longer the lamp is to be lit, the larger the battery needs to be.



Long period lighting



Short period lighting

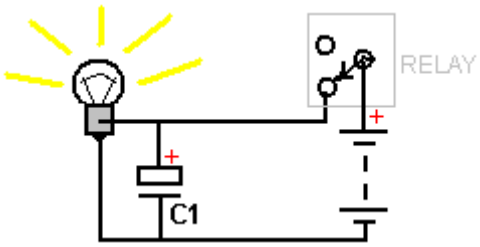
Looking at this from a slightly different viewpoint, we could say that the shorter the period of time when the battery has to power the bulb, the smaller the battery can be. So, for example, if the bulb only needed to be lit for ten seconds, then the battery could be quite small.

But suppose we were to reduce that time period to just one second, then the battery could be very small indeed and if we were to reduce the time period to one tenth of a second, then the battery power would be tiny compared to the power needed to run the light for say, one hour. This may seem trivial and obvious, but it this is leading up to a strategy which could be useful.

Fairly recently, super capacitors have become available at reasonable price, with a 2.5V 5F capacitor costing as little as £3. Just to put this in perspective, in electronics, a 1000 microfarad capacitor is considered large and a 10,000 microfarad capacitor is considered to be very large. Well, a 5F capacitor is 500 times larger than a 10,000 microfarad capacitor. These capacitors charge up fully in just a few seconds, are 100% efficient and should remain operational for many years if not many decades.

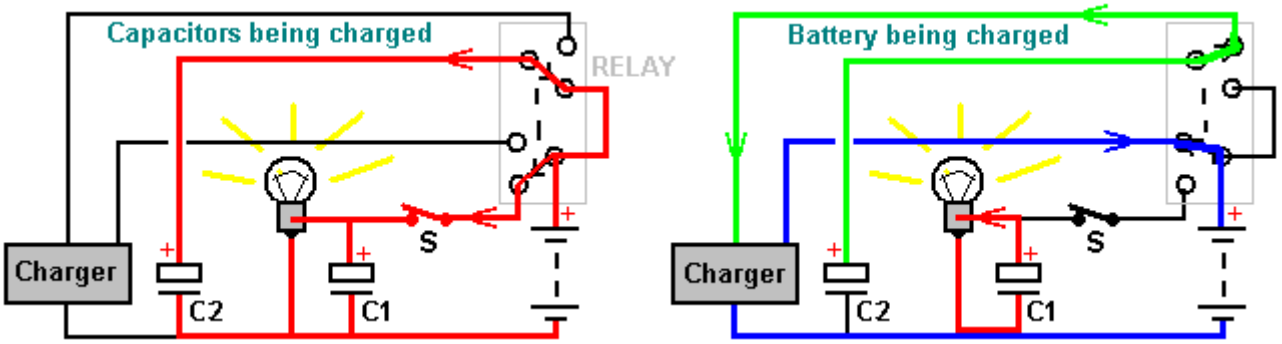
Suppose then, that we were to connect a 6-volt bulb across a 6-volt lead-acid battery. If the battery were charged to say, 6.6 volts (measured one hour after being disconnected from a battery charger) and the lamp were connected across the battery and left lit until the battery voltage dropped to 5.0 volts (on load, as it will gain a good deal of voltage when the load is disconnected). Then that determines how long the battery can power the lamp when the battery is discharged that much.

It would be interesting to see if that length of time could be extended without using a larger battery. Suppose we were to connect a set of three of those super capacitors in series to form a 7.5V capacitors of 1.67F capacity, and then implemented the following circuit:



using a simple NE555 circuit to drive the relay On and Off once per second. Then, the battery would only be connected to the lamp for half of the time and it would be 'resting' for half of the time. That sounds like a mad idea as when the battery is connected it will have to both light the lamp and top up the charge in the capacitor, and some current is needed to run the NE555 circuit and drive the relay. That results in a poorer performance than before.

However, as the battery is disconnected for half of the time, we can use that fact to our advantage by extending the circuit to this arrangement:



This makes the battery draw even worse as there are now two large capacitors which have to be topped up every second while the lamp is being powered and the NE555 circuit is also being powered. But, for half of the time, the relay is in its other position which connects capacitor "C2" to the battery pulse-charger, and at the same time, the 'resting' battery is connected to the output of the battery charger, feeding additional charge to it.

You might feel that a charging period of half a second is far too short to be of any use, but that is actually not the case. In chapter 5, there is mention of the 1989 patent US 4,829,225 granted to Yury Podrazhansky and Phillip Popp, their evidence is that batteries charge much better and have a longer life if they are pulsed in a specific way. Their formula is that the battery should be given a powerful charging pulse lasting for a period of time between a quarter of a second and two seconds, the pulse being the Amp-Hour rating of the battery. That is, for an 85 AHr battery, the charging pulse would be 85 amps. That pulse is then followed by a discharging pulse of the same, or even greater current but only maintained for only 0.2% to 5% of the duration of the charging pulse. Those two pulses are then followed by a resting period before the pulsing is repeated. They quote the following examples of their experiences when using this method:

Battery:	**9V alkaline**	**1.25V NiCad 0.5 AHr**	**15V NiCad 2 AHr**	**12V lead-acid 40 AHr**
Charging current:	0.5 Amps	1.2 Amps	3.0 Amps	48 Amps
Charging period:	550 mS	700 mS	500 mS	850 mS
Discharge current:	6 Amps	6 Amps	14 Amps	85 Amps
Discharge period:	2 to 3 mS	2 mS	2 mS	3 mS
Rest period:	15 to 20 mS	7 to 10 mS	10 mS	15 mS
Charging range:	50% to 100%	20% to 100%	20% to 100%	20% to 100%
Total charging time:	12 to 15 mins	20 mins	35 to 40 mins	40 mins

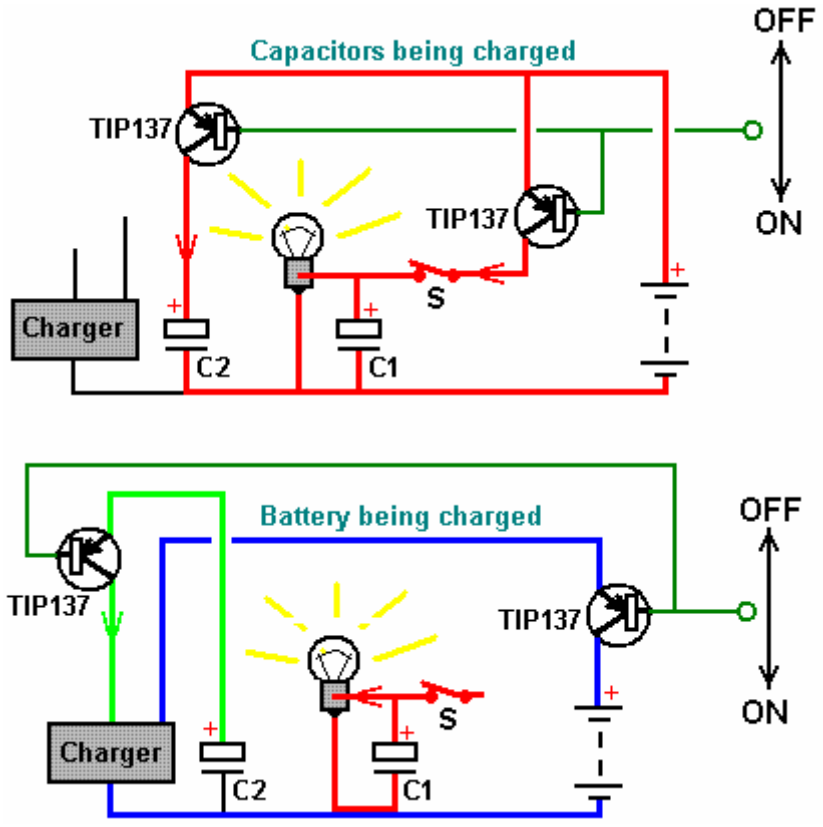
Of course, while this table of results does not apply directly to our proposed circuit, it does indicate that very satisfactory results can come from a very intermittent charging sequence and while I have, so far, spoken of a one-second cycle for our test circuit, that is purely due to the initial test use of a mechanical relay. If the test is

satisfactory, showing a longer length of time for the battery to reach an on-load voltage of 5 volts, and the relay circuit is replaced with a solid-state (lower current draw) circuit, then a switching time of maybe one tenth of a second might be used.

If the battery charger runs at 300,000 Hz (which is only 10% of some of the Alexkor circuits), then with a timing of one tenth of a second On time and one tenth of a second Off time, then the battery would receive about 30,000 charging pulses five times per second, which is 150,000 charging pulses every second, which is very close to some very successful battery chargers which operate all the time.

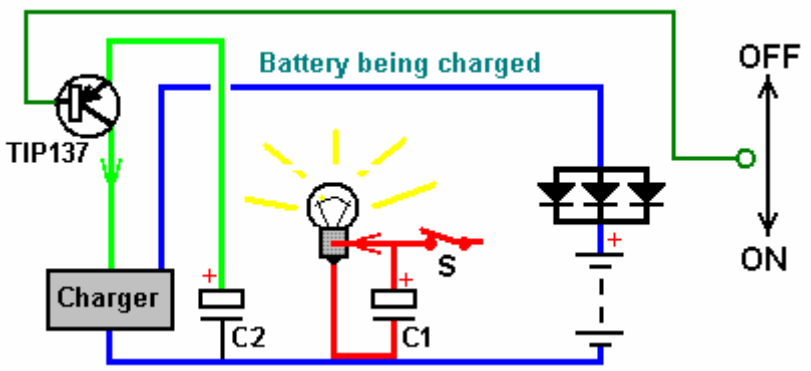
This is an untested idea, but it looks as if it might produce a usefully extended battery life and if the switch "S" is included in the circuit, then when it is opened, the battery may well be able to self-charge. As this arrangement is so simple, it might well be worth testing. The NE555 integrated circuit can provide up to 200 milliamps and so it could probably drive most 6V relays directly if the relay is connected between pin 3 and the plus of the battery. Initially, the NE555 circuit can be powered by a separate battery or power supply while the effect of the circuit is being checked out. Later on, if the circuit happens to work well, it could be powered by being connected to the battery via a choke and a smoothing capacitor (with a varistor across the capacitor if you want to play really safe) to block the charging pulses from the timing circuit, or it could be powered via an additional super capacitor or from one of the two existing capacitors. There is, of course, no reason why the NE555 circuit should have an On time equal to the Off time and so the timer's Mark/Space ratio could be made adjustable and tests run to see what the optimum setting would be. This could be an interesting experiment.

However, If you decide to test this idea, while you could use a relay to do the initial testing, it might be a good idea to test solid-state switching at the same time, as solid-state switching is likely to be much more reliable and have a much longer life. We might choose to use bi-polar transistors as they are the most reliable for lower voltage circuits, and so we might treat the circuit like this:

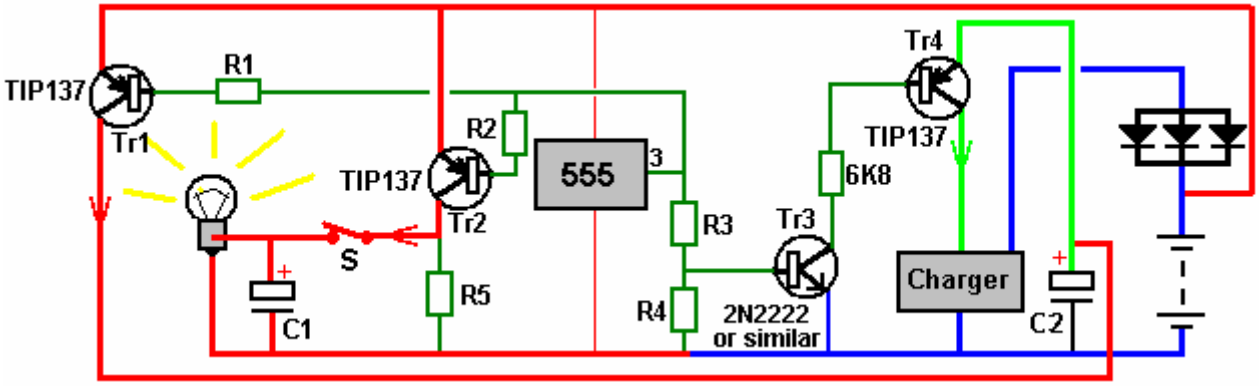


Maplins have 2.7V super capacitors at very reasonable prices. It would be necessary to connect four or five in series in order to meet the voltage requirements, although doing that reduces the overall capacitance substantially. It would be quite easy to add a dimmer control using a switched variable resistor and an emitter-follower transistor where the switch short-circuits the dimmer but when open, allows the voltage (and so, the current) applied to the LED bank, to be adjusted, as shown in the solar panel light suggestion at the end of chapter 14.

Initially, a slow-running 555 circuit could be used to drive one of the above circuits and invert the 555 output to drive the other circuit, using a transistor with a base resistor to protect it. A more practical charging circuit could be:



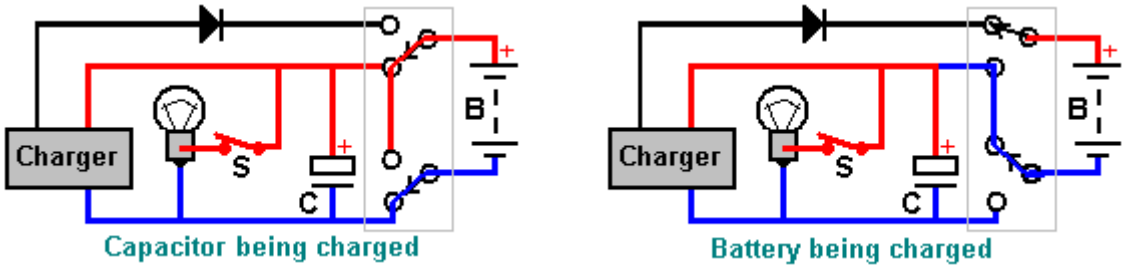
Here, the charger is connected to the battery at all times, through a diode (or perhaps more practically, through three or four diodes, such as UF5408, in parallel which feeds current better than one diode on its own), leading to a circuit something like this:



When pin 3 of the 555 goes low, resistors R1 and R2 control the current flow to transistors Tr1 and Tr2, switching them fully on in order to charge capacitors C1 and C2. Resistor R5 is a high value of say, 470K and is there just to provide a current flow path when the switch S is opened to turn the light off.

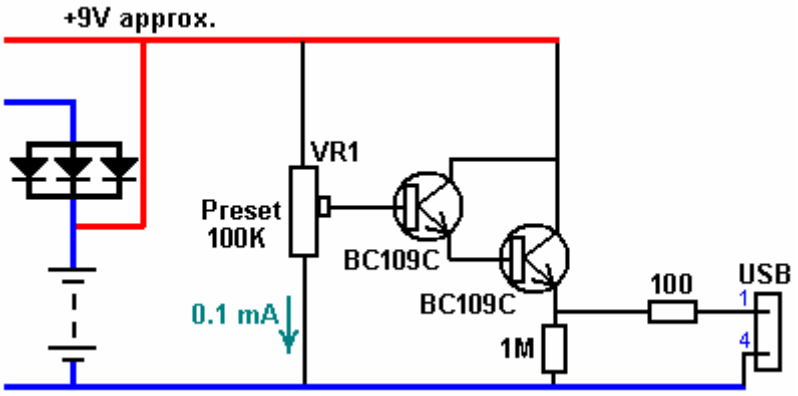
Resistor R3 is chosen to switch the 2N2222 transistor (or BC109, or other small-signal transistor) Tr3 fully on, when pin 3 goes high, and R4 is chosen to ensure that Tr3 switches fully off when pin 3 is low, switching the charger off by turning Tr4 off and starving it of current.

It is quite common for a charging circuit such as those from John Bedini or from Alexkor, to need the battery being charged to be connected in series to the battery which is driving the charging circuit (capacitor C2 in our case). This can be arranged like this:



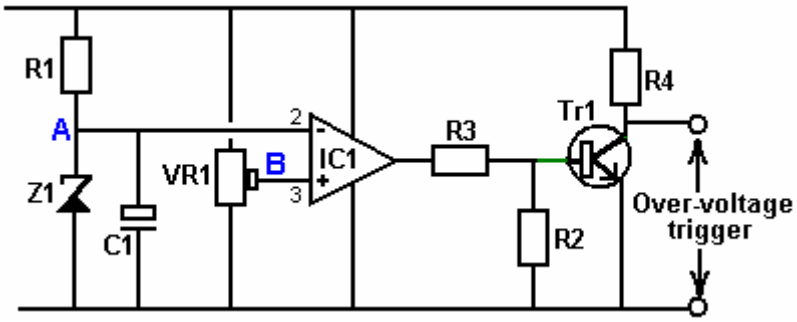
Here, the size of capacitor “C” is increased as it is now supplying both the light and the charging circuit, which runs all of the time. As the voltage spikes are high voltage and the resulting charging current very high, solid-state implementations of this arrangement require prototype testing before being shown here.

Ideally, for a USB output, we want an arrangement which draws no significant current when not in use and yet, which does not need a mechanical switch. For that we might do something like this:



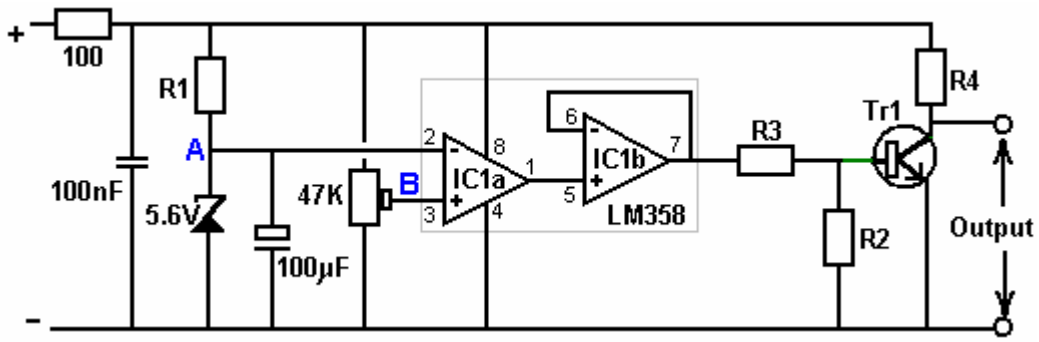
Here, the battery voltage of around 9V or 9.8V or whatever, is connected across a 100K variable resistor which will waste just under 0.1 milliamps. Two (low-cost) BC109C transistors are connected as a Darlington pair in emitter-follower mode with a 1M resistor as their load. These transistors have a gain which typically exceed 200, and so, two together like this draws only one 40,000 th of the proposed 10 mA current, and that is just 0.25 microamps which is about 0.25% of the current flowing through VR1 and so, even under load, the VR1 voltage will be steady. This wastes another 0.006 milliamps when VR1 is adjusted to give about 5.3V as an output when under load. A 100 ohm resistor is connected in series with this output and pin 1 of the 'type A' USB socket. As pins 1 and 4 are open-circuit, no current flows through this resistor until such time as a load is connected between pins 1 and 4.

If pins 1 and 4 were short-circuited – which is a most unlikely occurrence, then the current through the 120 ohm resistor would be 52.5 milliamps, which indicates that that resistor should be a half-watt type as that dissipation is 330 mW. The most effective way to make the initial adjustment would be to set the output voltage low, connect a mobile phone to the socket and adjust VR1 to give 5.3V or so on the socket. Preventing over-charging of the battery might be done with a circuit of this type:



Point "A" is a reference voltage provided by resistor R1 and zener diode Z1 and held at a steady voltage by capacitor C1. The Op-amp IC1 acts as a voltage comparator between points "A" and "B" (which is set by the preset variable resistor VR1 and reflects the overall voltage of the battery). The output voltage of IC1 will change sharply if the voltage difference between "A" and "B" swaps over, but by no means to the rail voltages, so transistor Tr1 is there to give a full logic swing. The transistor base current is set by resistor R3 and resistor R2 is there to ensure that Tr1 switches off properly. Resistor R4 has quite a high value as it is just there to form a voltage divider pair with Tr1.

The op-amp might be an LM358 as it will operate with voltages as low as 4.5V, it is very cheap, and it draws almost no current. As the package has two op-amps in it, the second one can be connected as a buffer:



As we want the circuit to switch as the battery voltage increases, point B must start at a lower voltage than point A. If connected as shown, then Tr1 will be on when the battery is charging and will switch off when the battery is fully charged at a voltage set by the position of the slider of VR1.

As the battery is receiving high-voltage spikes when being charged, the input to this section of the circuit is buffered by the 100 ohm resistor and the 100 nF capacitor is there to attempt to short-circuit spikes which get through the 100 ohm resistor. An inductor would probably be better than the 100 ohm resistor for suppressing voltage spikes and connecting a much larger capacitor in parallel with the 100 nF capacitor might be an additional help in keeping the voltage rails of this section of the circuit at a steady average voltage. The choice of resistor R1 and the voltage of the zener diode are not critical. The current through R1 can be very low as IC1 needs almost no current and as there is a reservoir capacitor across the zener diode and that will keep the voltage steady.

Resistors R2 and R3 are chosen to suit Tr1, R3 to ensure that it switches on properly when connected to the charger circuit and R2 to make sure that it switches off properly when then output pin 7 of IC1b drops to its lowest value of about 2 volts. It is likely that R4 will not be needed as TR1 will probably be connected direct into the charging circuit in order to power it down or stop it oscillating.

Patrick Kelly
www.free-energy-info.tuks.nl
www.free-energy-info.com
www.free-energy-info.co.uk
www.free-energy-devices.com

Chapter 7: Aerial Systems and Electrostatic Generators

Note: If you are not at all familiar with basic electronics, you might find it easier to understand this chapter if you read chapter 12 first.

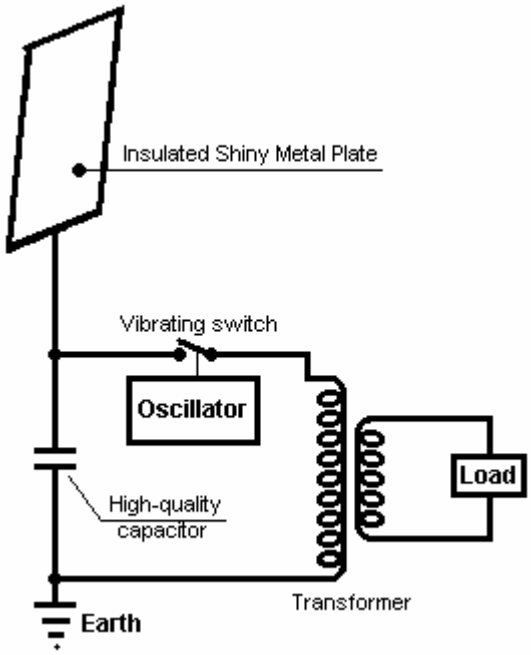
It is generally thought that aerials are not capable of gathering much power. The popular conception is that the only power available is low level radio waves from distant radio transmitters, and while it is certainly true that radio waves can be picked up with an aerial, the real sources of power are not radio transmitters.

For example, we will be looking at information from Hermann Plauson and he considered any aerial system of his which did not produce more than an excess power of 100 kilowatts, as a "small" system (admittedly, with many aerials). Thomas Henry Moray demonstrated his system to audiences repeatedly, pulling in power levels of up to 50 kilowatts from a single sort aerial. These power levels are not produced by radio station signals.

Nikola Tesla's Aerial System.

Nikola Tesla produced an aerial device which is worth mentioning. It was patented on May 21st 1901 as an "Apparatus for the Utilisation of Radiant Energy", US Patent number 685,957.

The device appears simple but Tesla states that the capacitor needs to be "of considerable electrostatic capacity" and he recommends using the best quality mica to construct it as described in his 1897 patent No. 577,671. The circuit draws power via an insulated, shiny metal plate. The insulation could be spray-on plastic. The larger the plate, the greater the energy pick-up. The higher the plate is elevated, the greater the pick-up.



This system of Tesla's picks up energy day and night. The capacitor gets charged up and a vibrating switch repeatedly discharges the capacitor into the step-down transformer. The transformer lowers the voltage and raises the current available and the output is then used to power the electrical load.

It seems probable that this device operates primarily from static electricity, which some people believe is a manifestation of the zero-point energy field. Tesla's equipment might well operate when fed by a motor-driven Wimshurst machine instead of a large aerial plate. Details of home-built Wimshurst equipment are available in the book 'Homemade Lightning' by R.A. Ford, ISBN 0-07-021528-6.

However, it should be understood that Tesla described two different forms of energy pick-up. The first is static electricity, picked up from very slight interaction of the pick-up plate with the zero-point energy field flowing through it, and the other being pick-up of dynamic radiant energy events, typically from lightning strikes. At a casual glance, the average person would not consider lightning as being a viable source of energy, but this is not the case as there are about two hundred lightning strikes per second - mainly in the tropics - and what is generally not understood is that they are radiant energy events and their effects are felt instantly everywhere on earth as transmissions through the zero-point energy field are instantaneous at any distance. To clarify the situation a little more, here are two of Tesla's patents, one on pick-up of the static field which Tesla remarks appears to be unlimited in voltage, and one patent on pick-up of dynamic energy.

This is a slightly re-worded copy of this patent, as some words have changed their meaning since this patent was issued. If you wish to see the original, then http://www.freepatentsonline.com will allow you to download a copy without any charge.

Patent US 685,957 **5th November 1901** **Inventor: Nikola Tesla**

APPARATUS FOR THE UTILISATION OF RADIANT ENERGY

To all whom it may concern:
Be it known that I, Nikola Tesla, a citizen of the Unites States, residing at the borough of Manhattan, in the city, county and State of New York, have invented certain new and useful improvements in Apparatus for the Utilisation of Radiant Energy, of which the following is a specification, reference being had to the drawings accompanying and forming a part of the same.

It is well known that certain radiations - such as those of ultra-violet light, cathodic, Roentgen rays, or the like - possess the property of charging and discharging conductors of electricity, the discharge being particularly noticeable when the conductor upon which the rays impinge is negatively electrified. These radiations are generally considered to be ether vibrations of extremely small wave lengths, and in explanation of the phenomena noted, it has been assumed by some authorities that they ionise, or render conducting, the atmosphere through which they are propagated. However, my own experiments and observations lead me to conclusions more in accord with the theory heretofore advanced by me that sources of such radiant energy throw off with great velocity, minute particles of matter which are strongly electrified, and therefore capable of charging an electrical conductor, or, even if not so, may at any rate discharge an electrified conductor, either by bodily carrying off its charge or otherwise.

My present application is based upon a discovery which I have made that when rays or radiations of the above kind are permitted to fall upon an insulated conducting-body connected to one of the terminals of a capacitor, while the other terminal of the capacitor is made to receive or carry away electricity, a current flows into the capacitor so long as the insulated body is exposed to the rays, and under the conditions specified below, an indefinite accumulation of electrical energy in the capacitor takes place. After a suitable time interval during which the rays are allowed to act, this energy may manifest itself in a powerful discharge, which may be used for the operation or control of mechanical or electrical devices, or rendered useful in many other ways.

In applying my discovery, I provide a capacitor, preferably of considerable electrostatic capacity, and connect one of its terminals to an insulated metal plate or other conducting-body exposed to the rays or streams of radiant matter. It is very important, particularly in view of the fact that electrical energy is generally supplied to the capacitor at a very slow rate, to construct the capacitor with the greatest care. I prefer to use the best quality of mica as the dielectric, taking every possible precaution in insulating the armatures, so that the instrument may withstand great electrical pressures without leaking and may leave no perceptible electrification when discharging instantaneously. In practice, I have found that the best results are obtained with capacitors treated in the manner described in Patent 577,671 granted to me on 23rd February 1897. Obviously, the above precautions should be the more rigorously observed the slower the rate of charging and the smaller the time interval during which the energy is allowed to accumulate in the capacitor. The insulated plate or conducting-body should present to the rays or streams of matter, as large a surface as is practical, I having ascertained that the amount of energy conveyed to it per unit of time is, under otherwise identical conditions, proportional to the area exposed, or nearly so. Furthermore, the surface should be clean and preferably highly polished or amalgamated. The second terminal or armature of the capacitor may be connected to one of the poles of a battery or other source of

electricity, or to any conducting body or object whatever of such properties or so conditioned that by its means, electricity of the required sign will be supplied to the terminal. A simple way of supplying positive or negative electricity to the terminal is to connect it to an insulated conductor supported at some height in the atmosphere, or to a grounded conductor, the former, as is well known, furnishing positive, and the latter negative electricity. As the rays or supposed streams of matter generally convey a positive charge to the first terminal of the capacitor mentioned above. I usually connect the second terminal of the capacitor to the ground, this being the most convenient way of obtaining negative electricity, dispensing with the necessity of providing an artificial source. In order to use the energy collected in the capacitor for any useful purpose, I also connect to the capacitor terminals, a circuit containing an instrument or apparatus which it is desired to operate, and another instrument or device for alternately closing and opening the circuit. This latter device can be any form of circuit-controller with fixed or moveable parts or electrodes, which may be actuated either by the stored energy or by independent means.

My discovery will be more fully understood from the following description and drawings, where **Fig.1** is a diagram showing the general arrangement of the apparatus as usually employed.

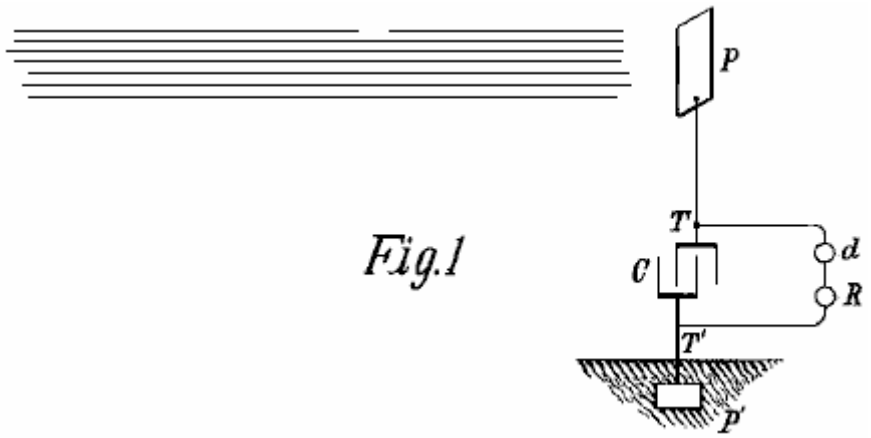


Fig.1

Fig.2 is a similar diagram, illustrating in more detail, typical forms of the devices or elements used in practice.

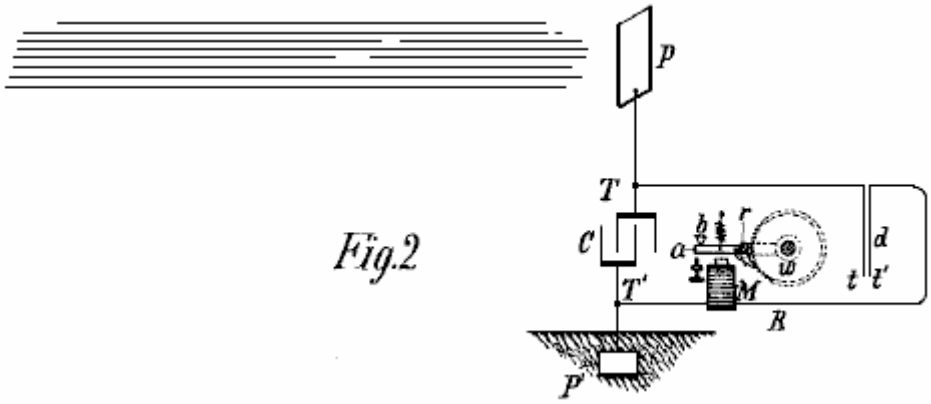


Fig.2

Fig.3 and **Fig.4** are diagrams of modified arrangements suitable for special purposes.

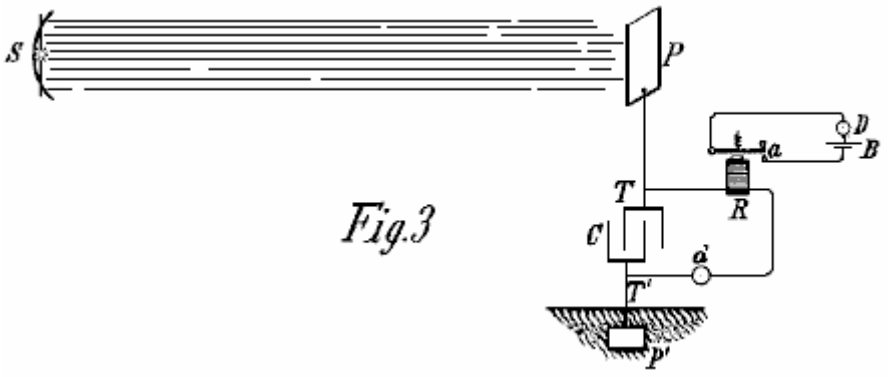


Fig.3

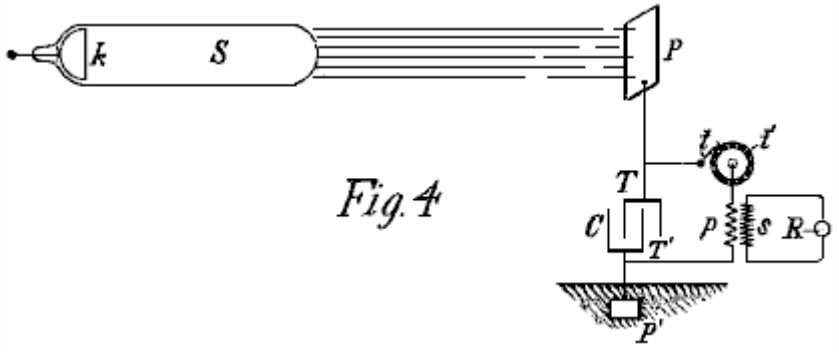


Fig.4

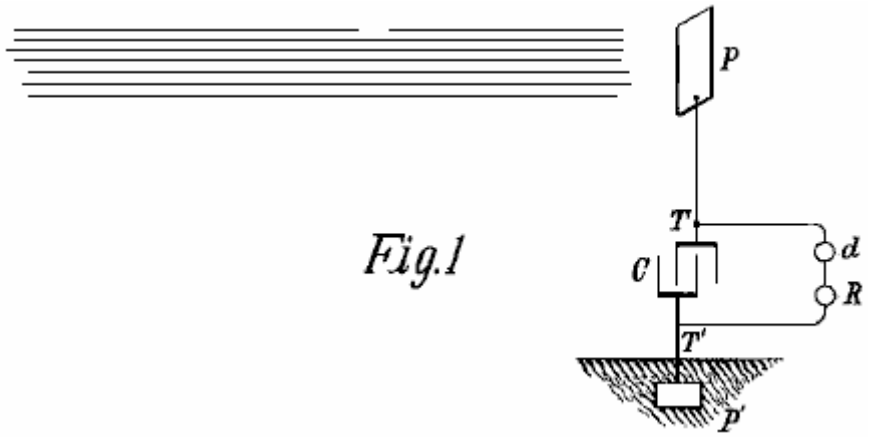


Fig.1

Fig.1 shows the simplest form, in which **C** is the capacitor, **P** the insulated plate or conducting-body which is exposed to the rays, and **P'** another plate or conductor which is grounded, all being connected in series as shown. The terminals **T** and **T'** of the capacitor **C** are also connected to a circuit which contains a device **R** which is to be operated, and a circuit-controlling device **d** as described above.

The apparatus being arranged as shown, it will be found that when the radiation of the sun, or any other source capable of producing the effects described above, fall on plate **P**, there will be an accumulation of energy in capacitor **C**. I believe that this phenomenon is best explained as follows: The sun, as well as other sources of radiant energy , throws off minute particles of positively electrified matter, which striking plate **P**, create an electrical charge on it. The opposite terminal of the capacitor being connected to the ground, which can be considered to be a vast reservoir of negative electricity, a feeble current flows continuously into the capacitor, and since these supposed particles are of an inconceivably small radius or curvature, and consequently, charged to a very high voltage, this charging of the capacitor may continue as I have actually observed, almost indefinitely, even to the point of rupturing the dielectric. If the device **d** be of such character that it will operate to close the circuit in which it is included when the capacitor voltage has reached a certain level, then the accumulated charge will pass through the circuit, operating the receiver **R**.

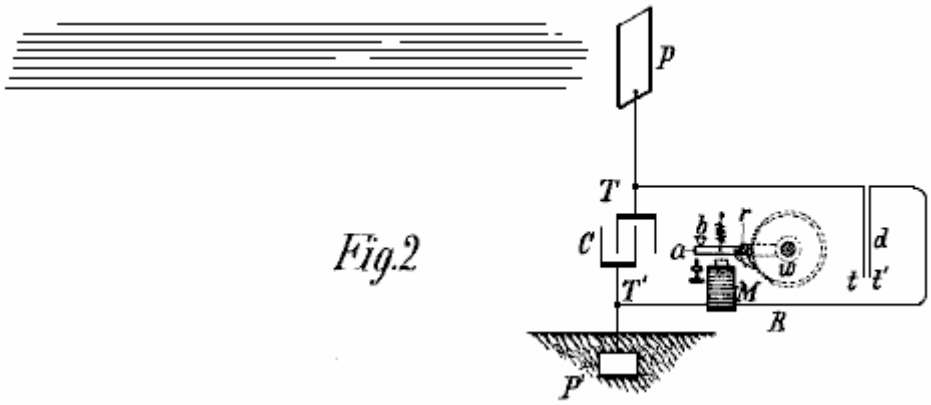


Fig.2

In illustration of this effect, **Fig.2** shows the same general arrangement as in **Fig.1**, and the device **d** is shown composed of two very thin conducting plates **t** and **t'** which are free to move and placed very close to each other. The freedom of movement can be either through the flexibility of the plates or through the character of their support. To improve their action they should be enclosed in a housing which can have the air removed from it.

The plates **t** and **t'** are connected in series in a working circuit which includes a suitable receiver, which in this example is shown as an electromagnet **M**, a moveable armature **a**, a spring **b**, and a ratchet wheel **w**, provided with a spring-pawl **r**, which is pivoted to armature **a** as illustrated. When the radiation falls on plate **P**, a current flows into the capacitor until its voltage causes the plates **t** and **t'** to be attracted together, closing the circuit and energising the magnet **M**, causing it to draw down the armature **a** and cause a partial rotation of the ratchet wheel **w**. When the current flow stops, the armature is retracted by the spring **b**, without, however, moving the wheel **w**. With the stoppage of the current, the plates **t** and **t'** cease to be attracted and separate, thus restoring the circuit to its original condition.

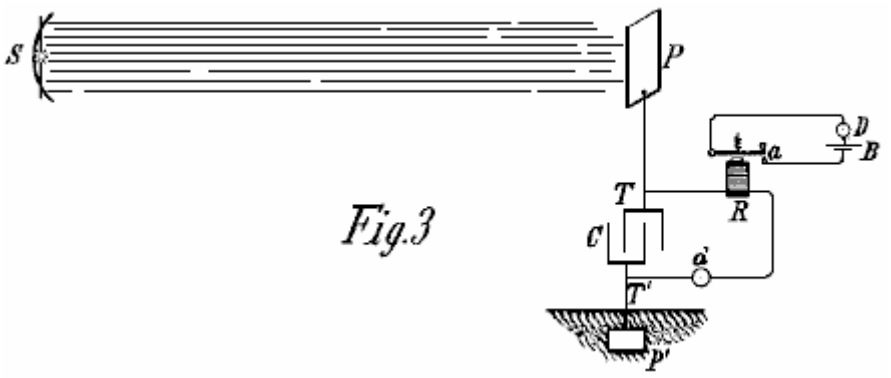


Fig.3

Fig.3 shows a modified form of apparatus used in connection with an artificial source of radiant energy, which in this case may be an arc emitting copious ultra-violet rays. A suitable reflector may be provided for concentrating and directing the radiation. A magnet **R** and circuit-controller **d** are arranged as in the previous figures, but in this case, instead of performing the whole of the work, the magnet performs the task of alternately opening and closing a local circuit, containing a source of current **B** and a receiving or translating device **D**. The controller **d** may, if desired, consist of two fixed electrodes separated by a minute air gap or weak dielectric film which breaks down more or less suddenly when a definite voltage difference is reached at the terminals of the capacitor, and returns to its original state when the discharge occurs.

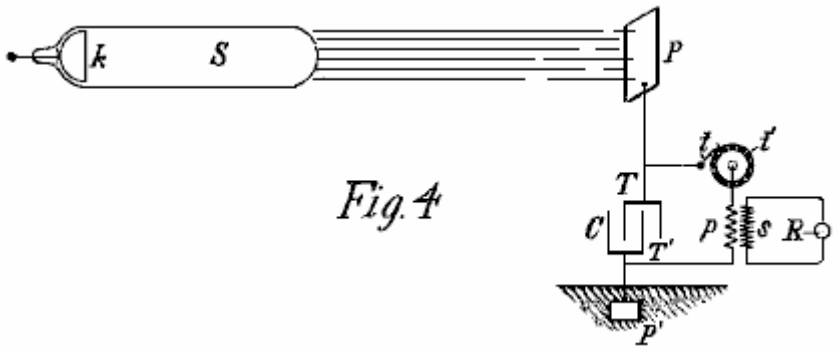


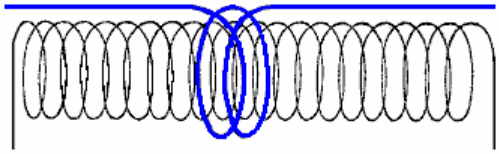
Fig.4

Still another modification is shown in **Fig.4**, in which **S**, the source of radiant energy is a special form of Roentgen tube devised by me, having only one terminal **k**, generally of aluminium, in the form of half a sphere, with a plain polished surface on the front side, from which the streams are thrown off. It may be excited by attaching it to one of the terminals of any generator with sufficiently high electromotive force; but whatever apparatus is used, it is important that the tube has the air inside it removed to a high degree, otherwise it might prove to be entirely ineffective. The working, or discharge circuit connected to the terminals T and T' of the capacitor, includes, in this case, the primary winding **p** of a transformer, and a circuit-controller comprised of a fixed terminal or brush **t** and a moveable terminal **t'** in the shape of a wheel, with conducting and insulating segments, which may be rotated at an arbitrary speed by any suitable means. In inductive relation to the primary winding **p**, is a secondary winding **s**, usually of a much greater number of turns, to the ends of which is connected a receiver **R**. The terminals of the capacitor being connected as shown, one to an insulated plate **P** and the other to a grounded plate **P'**. When the tube **S** is excited, rays or streams of matter are emitted from it and these convey a positive charge to the plate **P** and capacitor terminal **T**, while the capacitor terminal **T'** is continuously receiving negative electricity from plate **P'**. As already explained, this results in an accumulation of electrical energy in the capacitor, and this continues as long as the circuit including the primary winding **p** is interrupted. Whenever the circuit is closed by the rotation of the terminal **t'**, the stored energy is discharged through the primary winding **p**, giving rise to induced currents in the secondary winding **s**, which operates the receiver **R**.

It is clear from what has been stated above, that if the terminal **T'** is connected to a plate supplying positive instead of negative electricity, then the rays should convey negative electricity to plate **P**. The source **S** may be any form of Roentgen or Leonard tube, but it is obvious from the theory of action that in order to be very effective, the impulses exciting it should be wholly, or mainly of one sign. If ordinary symmetrical alternating currents are employed, then provision should be made for allowing the rays to fall on plate **P** only during those periods when they can produce the desired result. Obviously, if the source radiation is stopped or intercepted, or the intensity varied in any manner such as periodically interrupting or rhythmically varying the current exciting the source, there will be corresponding changes in the action upon the receiver **R** and thus signals may be transmitted and many other useful effects produced. Further, it will be understood that any form of circuit-closer which will respond, or be set in operation when a predetermined amount of energy is stored in the capacitor, may be used instead of the device already described in connection with **Fig.2**.

The second patent requires the equipment to be tuned to one quarter of the wavelength of the energy pulses being collected. This patent shows a transmission method as well as a receiving method, but our main concern here is the receiving section shown on the right of the diagram as that can receive naturally occurring energy pulses in the environment and so provides free usable energy.

As it may be a little difficult to visualise the coil arrangement in this patent as many people are familiar with the "Tesla Coil" arrangement where a few turns of thick wire or copper tubing are used as a winding placed around an ordinary cylindrical coil, much like, this illustration from Tesla's patent US 568,178:

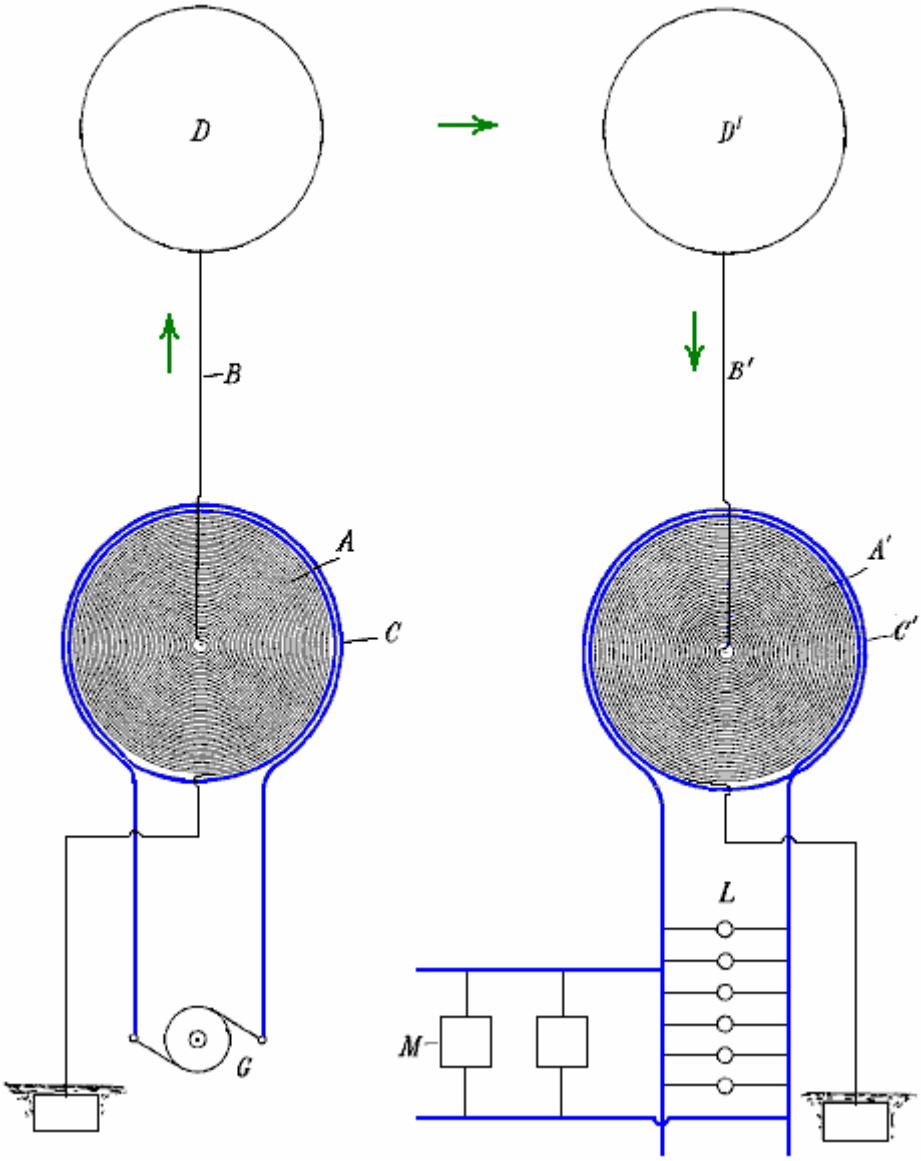


In this case it should be understood that Tesla is speaking about his flat "pancake" coil design and not the well-known Tesla Coil configuration.

Patent US 649,621 **15th May 1900** **Inventor: Nikola Tesla**

APPARATUS FOR THE TRANSMISSION OF ELECTRICAL ENERGY

To all whom it may concern:
Be it known that I, Nikola Tesla, a citizen of the Unites States, residing at the borough of Manhattan, in the city, county and State of New York, have invented certain new and useful improvements in Apparatus for the Transmission of Electrical Energy, of which the following is a specification, reference being had to the drawing accompanying and forming a part of the same.



This application is a division of an application filed by me on 2nd September 1897, US 650,343 entitled "Systems of Transmission of Electrical Energy" and is based on new and useful features and combinations of apparatus shown and described in that patent application.

This invention comprises a transmitting coil or conductor in which electrical currents or oscillations are produced and which is arranged to cause these currents or oscillations to be propagated by conduction through the natural medium from one location to a remote location, and a receiving coil or conductor adapted to be excited by the oscillations or currents propagated by the transmitter.

This apparatus is shown in the accompanying diagram where **A** is a coil, generally of many turns and of a very large diameter, wound in spiral form, either around a magnetic core or not as may be desired. **C** is a second coil formed by a conductor of much larger size and smaller length, wound around and in proximity to coil **A**.

The apparatus at one point is used as a transmitter, the coil **A** in this case forming a high-voltage secondary of a transformer, and the coil **C** the primary which operates at a much lower voltage. The source of current for the primary winding is marked **G**. One terminal of the secondary winding **A** is at the centre of the spiral coil, and from this terminal the current is led by a conductor **B** to a terminal **D**, preferably of large surface, formed or maintained by such means as a balloon at an elevation suitable for the purpose of transmission. The other terminal of the secondary winding **A** is connected to earth, and if desired, to the primary winding also in order that the primary winding may also be at substantially the same voltage as the adjacent portions of the secondary winding, thus ensuring safety.

At the receiving station, a transformer of similar construction is used, but in this case the coil **A'** constitutes the primary winding and the shorter coil **C'** is the secondary winding. In this receiving circuit, lamps **L**, motors **M**, or other devices for using this current, are connected. The elevated terminal **D'** connects with the centre of the coil **A'** and the other terminal is connected to earth and preferably, also, to the coil **C'** again for safety reasons as mentioned above.

The length of the thin wire coil in each transformer should be approximately one quarter of the wave length of the electric disturbance in the circuit, this estimate being based on the velocity of propagation of the disturbance through the coil itself and the circuit with which it is designed to be used. By way of illustration, if the rate at which the current flows through the circuit containing the coil is 185,000 miles per second, then a frequency of 925 Hz would maintain 925 stationary nodes in a circuit 185,000 miles long and each wave would be 200 miles in length.

For such a low frequency, which would only be resorted to when it is indispensable for the operation of ordinary motors, I would use a secondary winding wound from a wire 50 miles in length. By adjusting the length of wire in the secondary winding, the points of highest voltage are made to coincide with the elevated terminals **D** and **D'**, and it should be understood that whatever wire length is chosen, this length requirement should be complied with in order to get the best possible results.

It will be readily understood that when these relationships exist, the best conditions for resonance between the transmitting and receiving circuits are attained and owing to the fact that the points of highest voltage in the coils **A** and **A'** are coincident with the elevated terminals, the maximum current flow will take place in the two coils and this implies that the capacitance and inductance in each of the circuits have the values which produce the most perfect synchronism with the oscillations.

When the source of current G is in operation and produces rapidly pulsating or oscillating currents in the circuit of coil **C**, corresponding induced currents of very much higher voltage are generated in the secondary coil **A**, and since the voltage in that coil gradually increases with the number of turns towards the centre, and the voltage difference between adjacent turns is comparatively small, a very high voltage is generated, which would not be possible with ordinary coils.

As the main objective is to produce a current with excessively high voltage, this objective is facilitated by using a current in the primary winding which has a very considerable frequency, but that frequency is in a large measure, arbitrary, because if the voltage is sufficiently high and the terminals of the coils be kept at the proper height where the atmosphere is rarefied, the stratum of air will serve as a conducting medium with even less resistance then through an ordinary conductor.

As to the elevation of terminals **D** and **D'**, it is obvious that this is a matter which will be determined by a number of things, such as the amount and the quality of the work to be performed, the condition of the atmosphere and the character of the surrounding countryside. Thus, if there are high mountains in the vicinity, then the terminals should be at a greater height, and generally, they should be at an altitude much greater than that of the highest objects near them. Since, by the means described, practically any voltage which is desired may be produced, the currents through the air strata may be very small, thus reducing the loss in the air.

The apparatus at the receiving station responds to the currents propagated by the transmitter in a manner which will be well understood from the description above. The primary circuit of the receiver - that is, the thin wire coil **A'** - is excited by the currents propagated by conduction through the intervening natural medium between it and the transmitter, and these currents induce in the secondary coil **C'**, other currents which are used to operate the devices connected to that circuit.

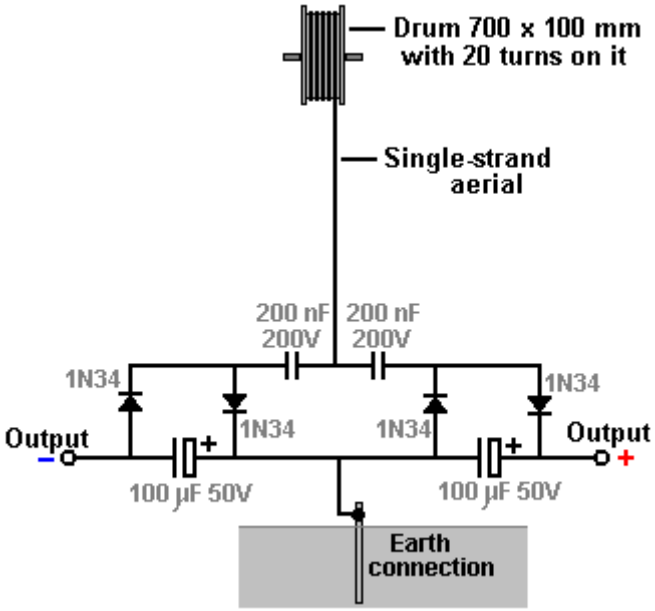
Obviously, the receiving coils, transformers, or other apparatus may be moveable - as for instance, when they are carried by a vessel floating in the air or by a ship at sea. In the former case, the connection of one terminal of the receiving apparatus to the ground might not be permanent, but might be intermittently or inductively established.

It should be noted that Tesla's suggestion of using the conductive envelope of a specially constructed balloon as a good method of increasing the active area of the elevated receiving plate, is one that was taken up by Hermann Plauson when he was building power stations operating on naturally occurring energy.

Jes Ascanius' Version of Nikola Tesla's Aerial System.

This sort of information may seem confusing and maybe a little too technical for you, so let me tell you about the practical and useful applications used by Jes Ascanius, a Danish developer, to whom thanks is due for sharing his design. Initially, he set up a system to charge his mobile phone battery overnight from an aerial. Then he went on to produce a full-size Tesla Aerial System as described at the start of this chapter. Let's start with the very simple system and progress from that to the more powerful arrangements.

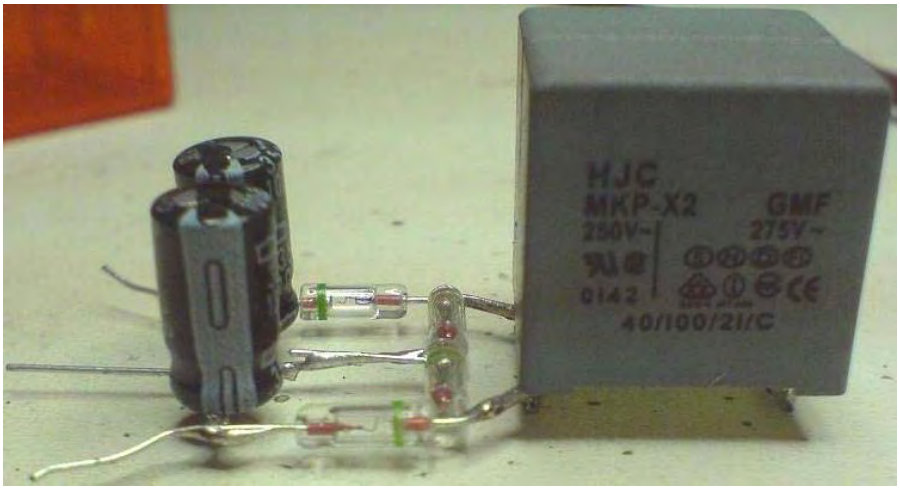
The initial circuit uses one strand of solid wire which rises vertically to a 700 mm diameter drum where there are some twenty turns. The arrangement is like this:

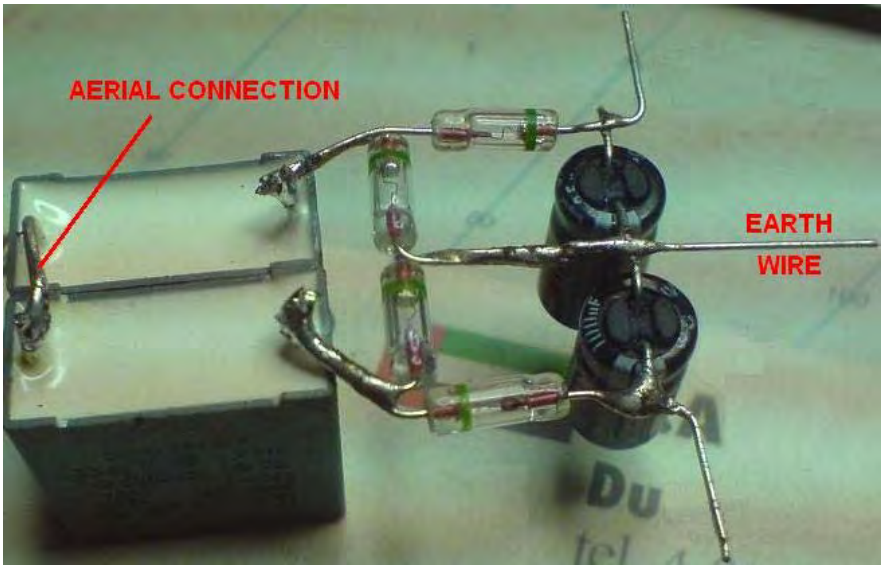


The aerial wire is several metres long, and in the prototype, was supported by (and insulated from) the eaves of a house. The aerial should be vertical or near vertical and a proper earth connection provided by driving a metal rod into the ground or connecting a wire to a metal plate and burying the plate in the ground as a good electrical connection is needed here. The earth connection used here is a 12 mm copper pipe 3 metres long, driven into the ground and the ground around it saturated with water:

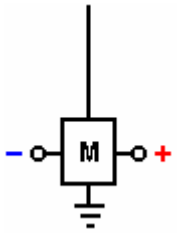
The wire used to connect with the earthing rod is very important and should not be less than 8 swg copper wire, that is, 4 mm diameter and 13 sq. mm. cross-sectional area. As with all free-energy devices, the exact constructional details are vital.

The diodes used are germanium 1N34 or 1N34a which are particularly suited to this application. Ceramic disc types are recommended for the 200 nF capacitors. The prototype build looked like this:

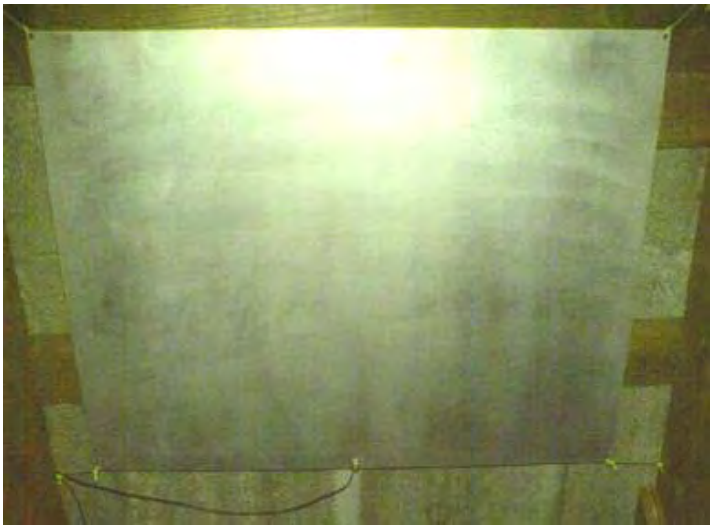




Now, consider this circuit as described, to be one modular building block which can lead to unlimited power from an aerial. I will represent the circuit shown above as a rectangle, showing the above circuit as:

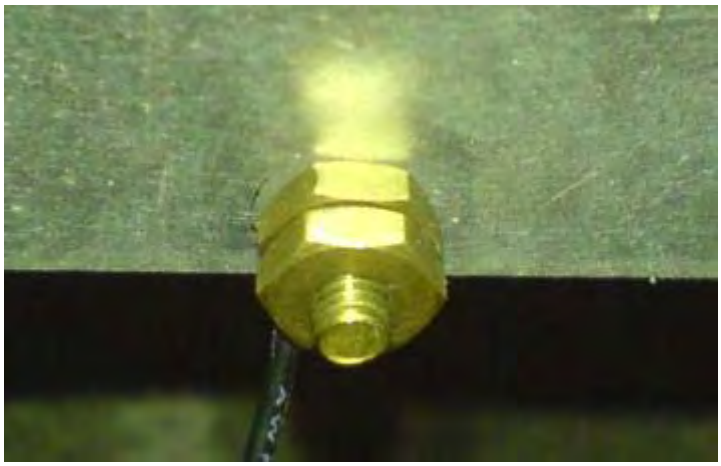


While it is possible to use more than one module with the aerial to get more power, the Danish developer then switched to the full-blown Tesla arrangement by attaching a 600 x 800 x 2 mm aluminium plate inside the sloping roof of his house:



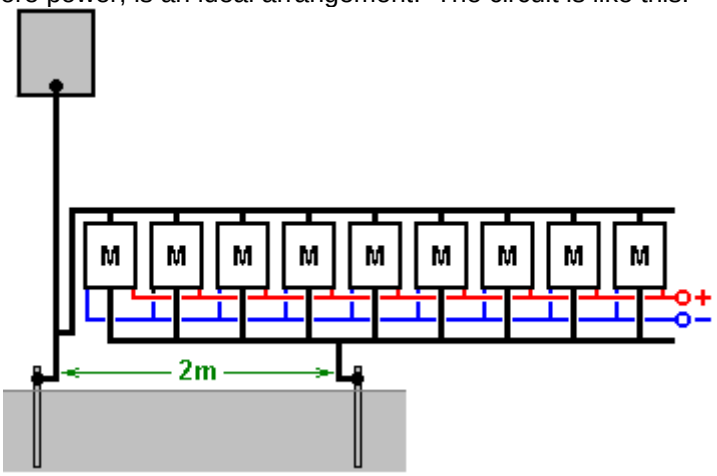
The plate being suspended using nylon cord to prevent it touching the roof or anything else:

The plate is positioned between 3 and 3.5 metres (10 to 12 feet) above the ground and the attachment to the plate is also heavy-duty 8 swg cable:



The cable is connected to the aluminium plate using a brass bolt and nuts which the builder thinks may be significant, quite apart from avoiding any galvanitic connection to the circuit. The cable is then run vertically downwards to the circuit. For this arrangement a second earthing point is also used. This is a galvanised iron pipe 3 metres long, driven vertically into the ground which is saturated with water. The second earth is 2 metres away from the first earth and there is no known significance in the use of an iron pipe as it was used because it was to hand at the time..

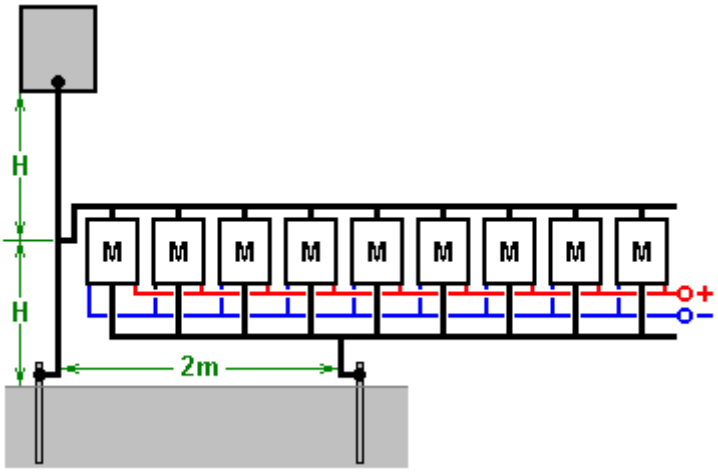
This arrangement provides serious power, enough to cause injury to, or kill a careless human. With two modules, it will light an LED very brightly, driving it to 2.6 volts. If the LED is removed, then the voltage climbs to about twenty volts and is easily sufficient to charge a 12V battery or battery bank although that takes time. With twenty modules a 12V battery can be charged over night. It is estimated that with two hundred modules, the power would be sufficient to power a household although that has not yet been done. It should be borne in mind that each module is easy and cheap to make, so arranging for a stack of them where additional modules can be added at a later date for more power, is an ideal arrangement. The circuit is like this:



This circuit looks completely mad as the aerial input to the circuit appears to be directly short-circuited by the second earth connection. In spite of this, the circuit works very well when connected this way. Additional modules can be added without any known limit. Increased power can be had by either raising the aluminium plate higher above the ground, to say, 10 metres (33 feet), or by adding one or more additional aerial plates. As you

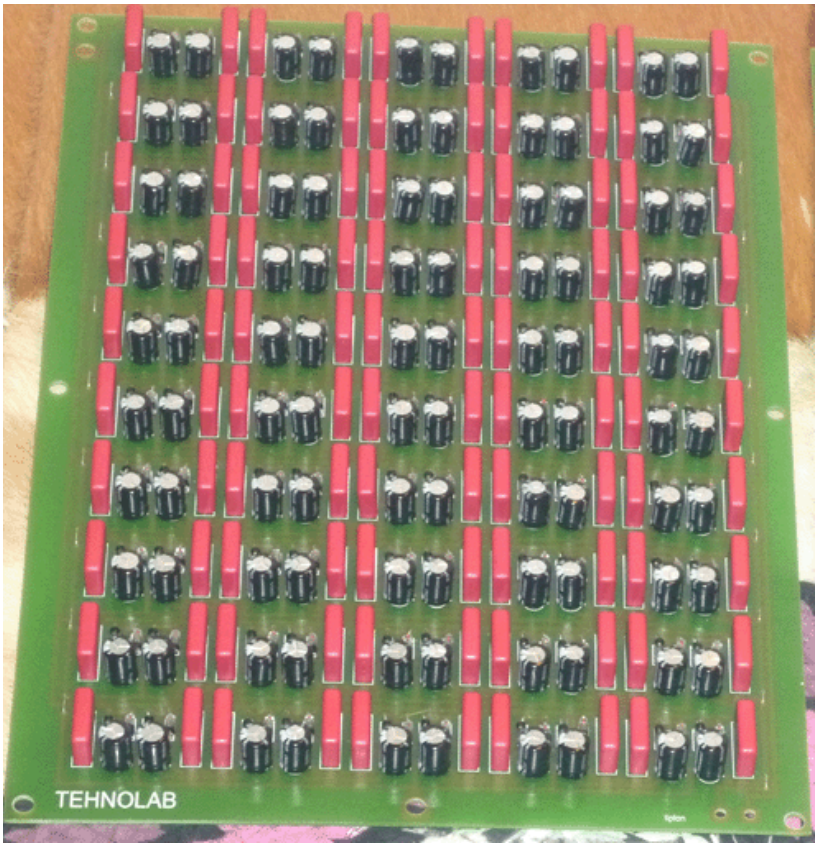
have a good aerial connected through to a very good earth, there has to be the possibility of the equipment being hit by lightning, and so it is recommended that a protective spark-gap is installed between the aerial and the earth, close to the circuit, so that if high-voltage is suddenly applied to the aerial, the spark gap will fire and shunt the excess power through to the earth. Alternatively, possibly a better solution is to install a standard lightning rod system a few metres away from the aerial and a metre or two higher up, so that it forms a more attractive point for a lightning strike.

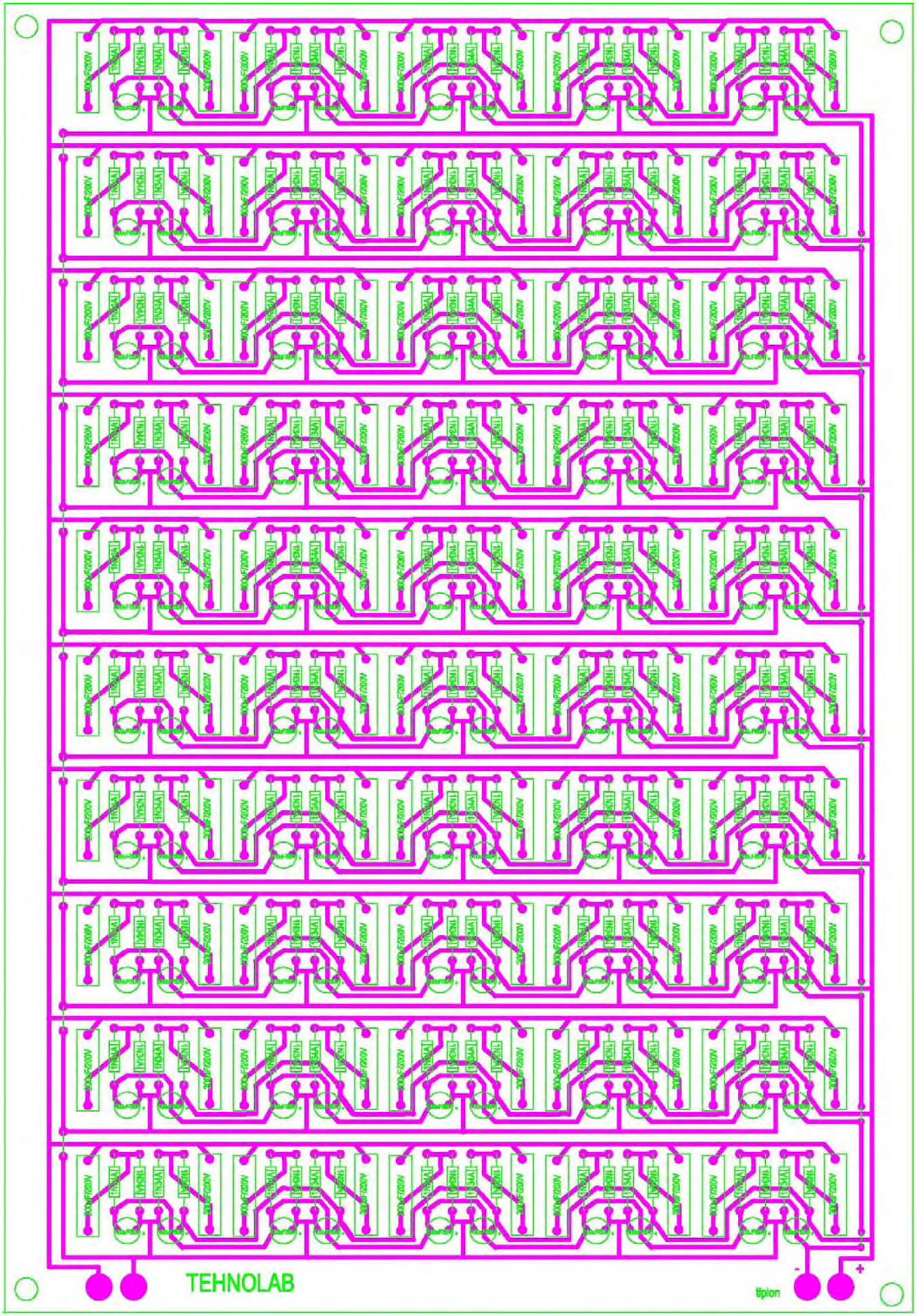
Further experimentation has shown that altering the connection point for the aerial has a significant effect on the results. If the connection is made at the mid point between the aerial plate and the earth connection, it produces a greater output:



With this arrangement a single module produces around 30 volts while the original method of connecting near the earth was giving about 26 volts with two modules. Jes Ascanius has carried out further experimentation and he states that diodes with response times under 30 milliseconds produce a greater output and he recommends the use of BYV27 diodes which have a 200-volt 25nS rating as he gets three times the output from them. He also recommends using them in Joule Thief circuits.

Dragan Kljajic has been experimenting with this circuit and has started by building many of these modules on a printed circuit board like this:




TEHNOLAB
tipion
-
+

Using two of these boards, Dragan is pulling 96 watts continuously from his aerial plate. He intends to extend this arrangement much further, but is being hindered at present by a local civil war.

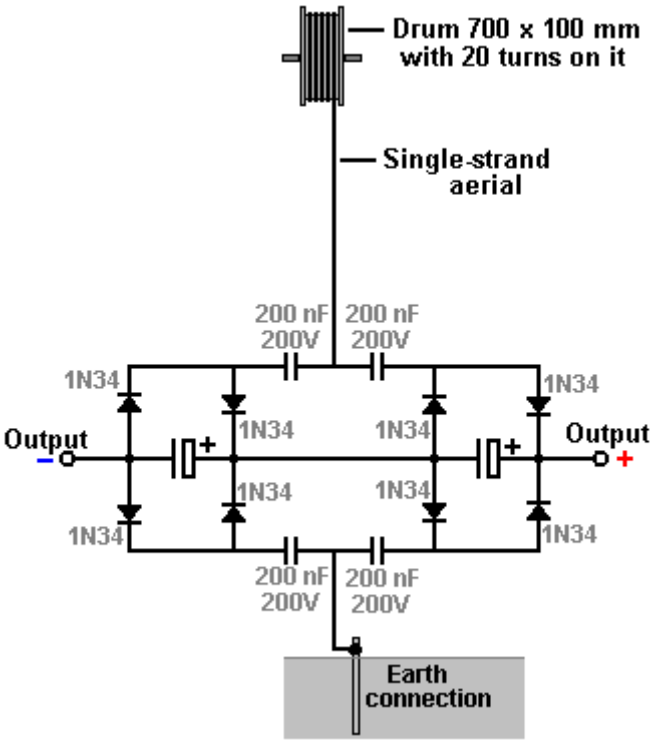
Here is a forum: http://www.energeticforum.com/renewable-energy/10947-jes-ascanius-radiant-collector.html where some builders of this system share comments. One comment is that there is an increased risk of a lightning strike where you have an earthed aerial, and so it is advisable not to place the aerial plate inside a house, but perhaps suspended between two trees. Also, using a car spark plug connected across the module set can protect against lightning strikes damaging the circuitry.

As a result of queries, Jes stresses the following points:

1. The plate **must** be high off the ground.
2. The plate **must** be polished and insulated.
3. The wire **must** be single-strand solid wire.
4. There **must not** be any part of the wire above the circuit, which is not insulated.

He further comments: you can use aluminium foil and cling film to make many collector plates 0.4 m x 5 m and connect them close together to feed the aerial wire. Remember, no uninsulated wire anywhere. Any queries should be asked on the forum shown above.

A modification of this circuit of Jes Ascanius' by a developer who prefers to remain anonymous, doubles the output of each module by adding a mirror image of the circuit like this:



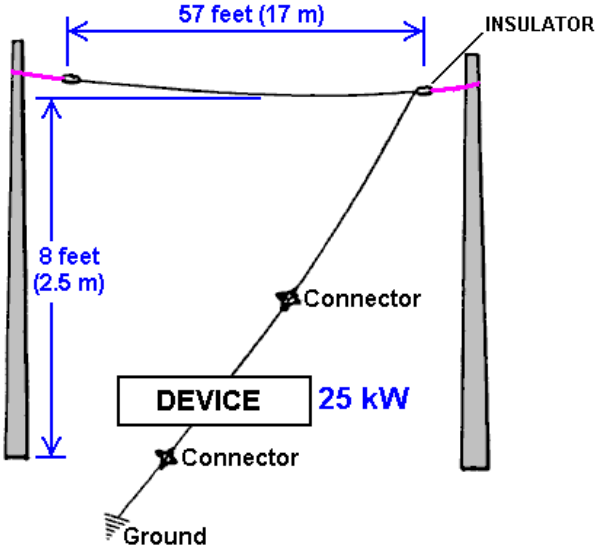
As can be seen, the addition is of four diodes and two capacitors. Presumably, using BYV27 diodes rather that 1N34 diodes would create a further enhanced output.

Thomas Henry Moray's Aerial System.

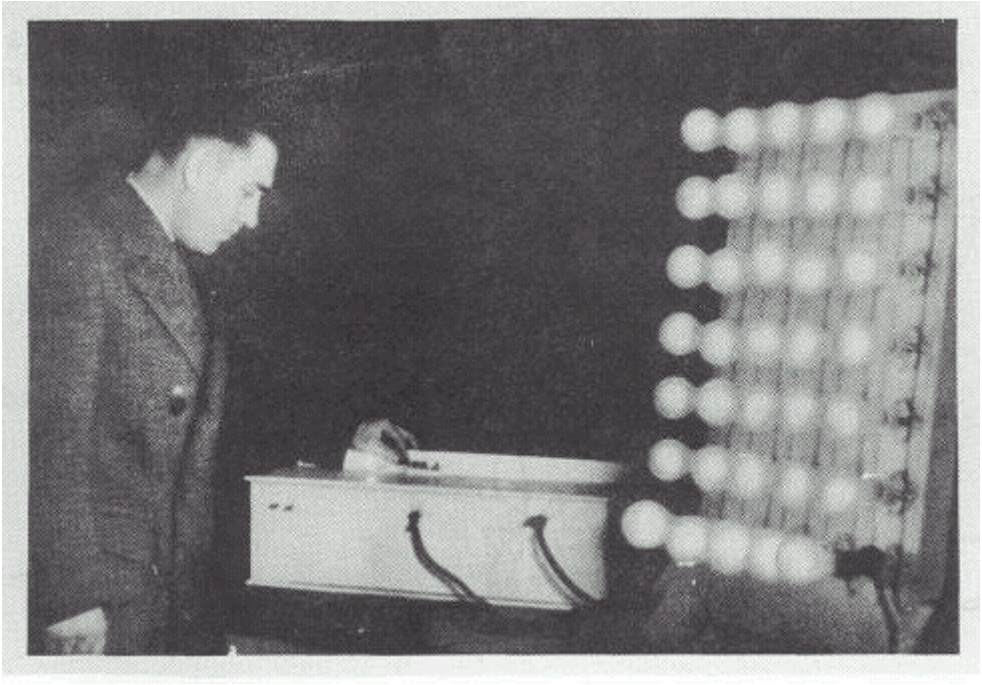
In this field, Thomas Henry Moray is outstanding. By 1936 he had developed a piece of apparatus which was capable of putting out high power with no human-generated input power at all.

Moray's equipment is said to have contained a germanium diode which he built himself in the days before solid-state devices became readily available. The equipment was examined and tested many times. On dozens of occasions, he demonstrated the equipment driving a bank of twenty 150W bulbs, plus a 600W heater, plus a 575W iron (a total of 4.175 kW). The power picked up by this device needed only small diameter wires and had characteristics different from conventional electricity. One demonstration which was repeated many times, was to show that the output power circuit could be broken and a sheet of ordinary glass placed between the severed ends of the wire, without disrupting the supply. This type of power is called "Cold electricity" because thin wires carrying major power loads, do not overheat. This form of energy is said to flow in waves which surround the wires of a circuit and not actually through the wires at all. Unlike conventional electricity, it does not use electrons for transmission and that is why it can continue through a sheet of glass which would stop conventional electricity dead in its tracks.

On one occasion, Moray took his equipment away from all urban areas to a place chosen at random by a critic. He then set up the equipment and demonstrated the power output, well away from any man-generated electrical induction. He disconnected the aerial and showed that the power output stopped immediately. He connected the aerial again to generate the output as before. He then disconnected the earth connection which stopped the output again. When the earth wire was connected again, the output power returned. He found that the power output level fell somewhat at night. The aerial used in his demonstrations was reported to be like this:



From which it can be seen that even a relatively short aerial wire suspended not far from the ground, is capable of collecting substantial amounts of usable electrical power. The earth used in the remote demonstrations was a length of gas pipe which was hammered into the ground. It was noted that as each hammer blow drove the pipe deeper into the earth, the light bulbs (which formed the electrical load) shone more brightly, indicating that the quality of the earth connection is very important when serious power is being drawn from an aerial.



Thomas developed various versions of the device, the latest of which did not need the aerial or earth connections, weighed 50 pounds and had an output of 50 kilowatts. This device was tested in both an aeroplane and a submarine, thus showing the device to be fully self-contained and portable. It was also tested in locations which were fully shielded from electromagnetic radiation.

Moray was shot and wounded in an assassination attempt in his laboratory. This caused him to change the glass in his car to bullet-proof glass. He was threatened many times. His demonstration equipment was smashed with a hammer. When threats were made against his family, he stopped rebuilding his equipment and appeared to have turned his attentions to other things, producing a device for 'therapeutic' medical treatment.

In his book "The Energy Machine of T. Henry Moray", Moray B. King provides more information on this system. He states that Moray was refused a patent on the grounds that the examiner couldn't see how the device could output so much power when the valve cathodes were not heated. Moray was granted US Patent 2,460,707 on 1st February 1949 for an Electrotherapeutic Apparatus, in which he included the specification for the three valves used in his power device, apparently because he wanted them to be covered by a patent. As far as can be seen, the valve shown here is an oscillator tube. Moray claimed that this tube had the very high capacitance of 1 Farad when running at its resonant frequency. Moray liked to use powdered quartz as a dielectric in the capacitors which he made, and he had a habit of mixing in radium salts and uranium ores with the quartz. These materials may well be important in producing ionisation in these tubes and that ionisation may well be important in tapping the energy field.

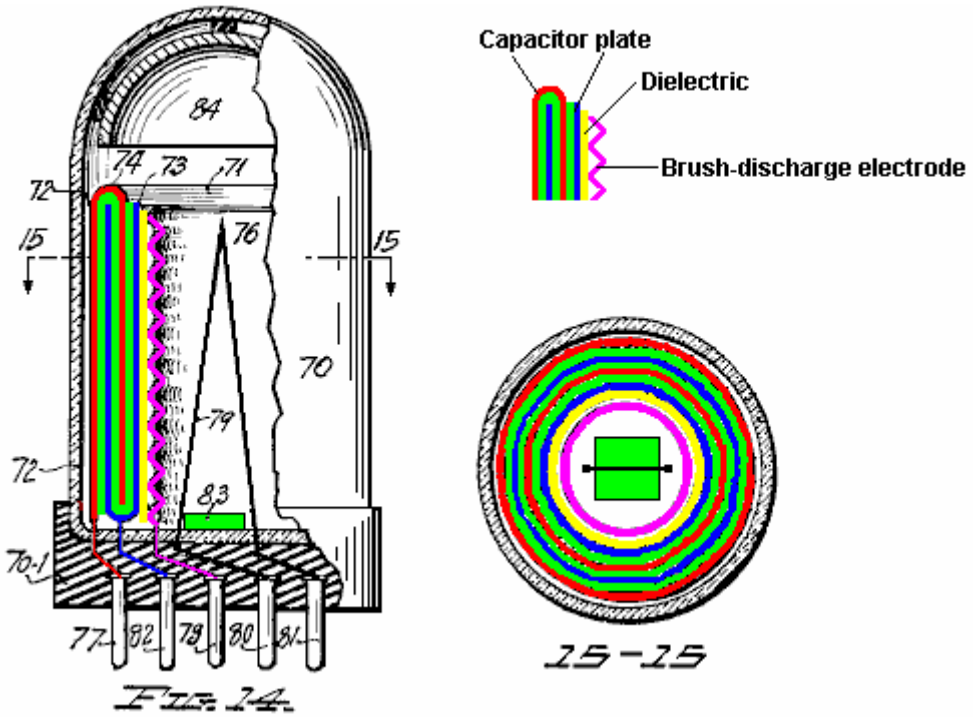


FIG. 14.

The tube shown above has a six-layer capacitor formed from two U-shaped circular metal rings with the space between them filled with a dielectric material. The plates are shown in red and blue, while the dielectric is shown in green. Inside the capacitor, there is a separate ring of dielectric material (possibly made from a different material) and an inside ring of corrugated metal to form an ion brush-discharge electrode. The capacitor and electrode connections are taken to pins in the base of the tube.

Quartz is suggested for the material of the outer covering of the tube and the wire element numbered **79** in the diagram is said to be a heating element intended to be powered by a low-voltage current source. However, as Moray had an earlier patent application refused on the grounds that there was no heating element in his tubes, it is distinctly possible that the heating element shown here is spurious, and drawn solely to avoid rejection by the examiners. In his patent, Moray refers to the capacitor in this tube as a “sparking” capacitor, so he may have been driving it with excessively high voltages which caused repeated breakdown of the capacitor material.

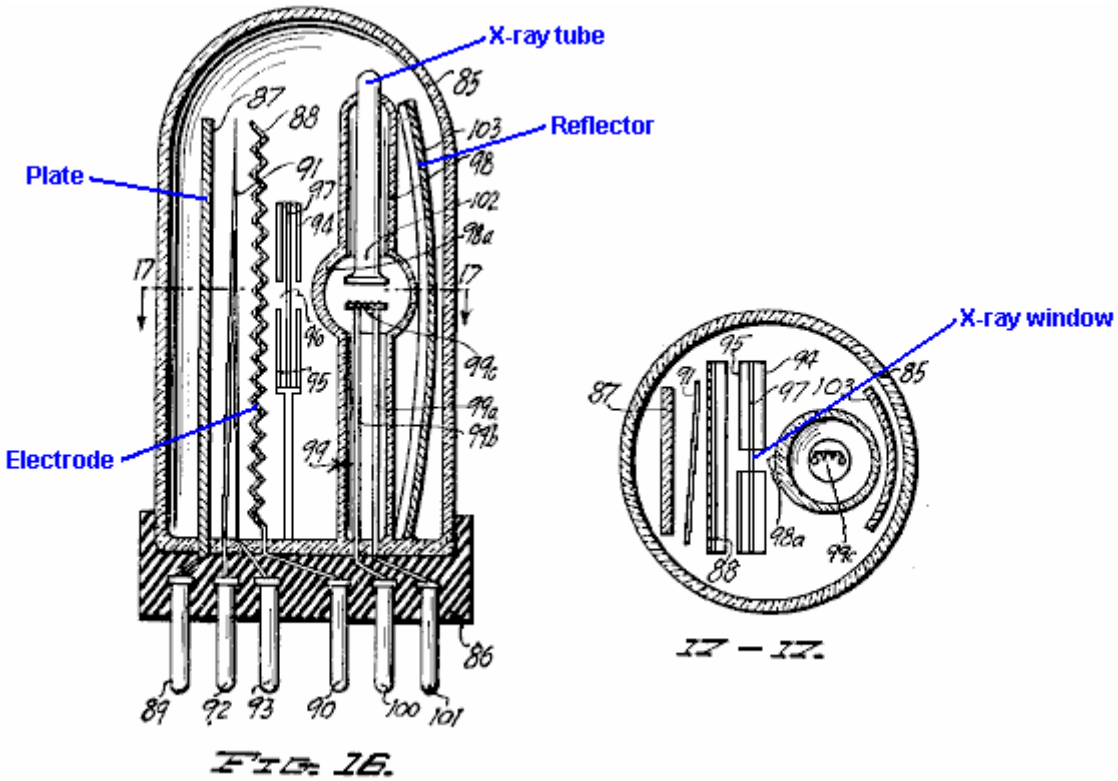


FIG. 16.

The tube of Fig.16 above, uses a different technique where an X-ray tube is used to bombard a corrugated electrode through a screen containing an X-ray window. It is thought that a brief burst of X-rays was used to trigger very short, sharp bursts of ions between the anode and cathode of the tube and these pick up extra energy with every burst.

An alternative version of this tube is shown in Fig.18 below. Here the construction is rather similar but instead of an X-ray window, a lens and reflector are used to cause the ionisation of the switching channel between the anode and cathode. In both tubes, the corrugated electrode supports a corona build-up just prior to the short X-ray switching pulse, and it is thought that the ions contribute to the intensity of the resulting pulses which emerge from the tube. Very short uni-directional pulses are capable of causing conditions under which additional energy can be picked up. From where does this extra energy come? In 1873, James Clerk Maxwell published his "Treatise on Electricity and Magnetism" and in it he pointed out that the vacuum contains a considerable amount of energy (Vol. 2, p. 472 and 473). John Archibald Wheeler of Princeton University, a leading physicist who worked on the US atomic bomb project, has calculated the flux density of the vacuum. Applying Einstein's $E=mC^2$ formula indicates that there is enough energy in every 1 cc of "empty" space, to create all of the matter in the visible universe which can be seen with our most powerful telescopes. That amount of energy is so great as to be beyond imagining. This energy field is referred to as "Universal Energy", "Cosmic Energy" or "Zero Point Energy". At this time, we do not have any instrument which responds directly to this energy and so it is almost impossible to measure.

The existence of this energy field is now widely accepted by mainstream science and it is borne out by the situation found at quantum levels. It is generally thought that this energy is chaotic in form and for useful energy to by drawn from it, it needs to be restructured into a coherent form. It appears that uni-directional electromagnetic pulses of one millisecond or less, can be used to cause the necessary restructuring as they generate an outward coherent wave of radiant energy, from which energy can be extracted for use in most electrical devices, if a suitable receptor system is used. Tom Bearden states that at the quantum level, the seething energy of this field appears continuously as positive and negative charges. As these are evenly distributed, the net charge at any point is always zero. If a "dipole" (two opposite charges near each other) is created anywhere, then it polarises the energy field disrupting the previously even distribution of charges and causing massive streams of energy to radiate outwards from the dipole.

A voltage pulse acts as a dipole, provided the voltage rise is fast enough, and that is what causes a wave of radiant energy fanning out from the location of the voltage pulse. Batteries and magnets create continuous dipoles and so cause the local quantum energy field to send out continuous streams of massive power which can be utilised if (and only if) you know how to do it. The search for mechanisms to capture and use even a tiny fraction of these energy streams is what the "free-energy" field of research is all about. Some people say that there is no such thing as "free-energy" because you have to pay for the device which captures it. That is like taking a bus trip to a car dealership where they are giving away new cars, and saying that your new car was not a "free" car because you had to pay a bus fare to reach the car dealership.

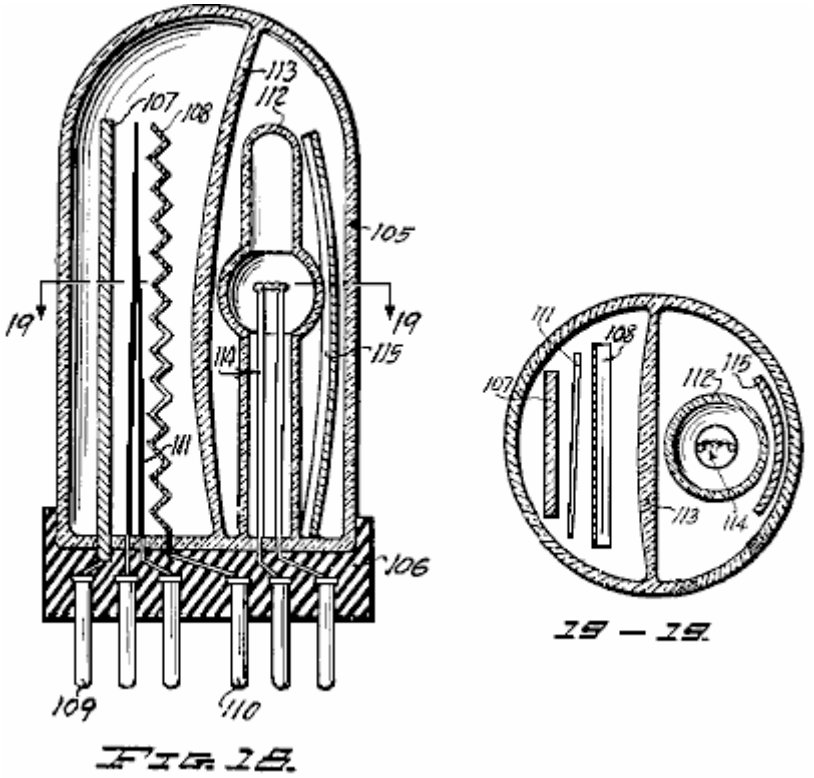
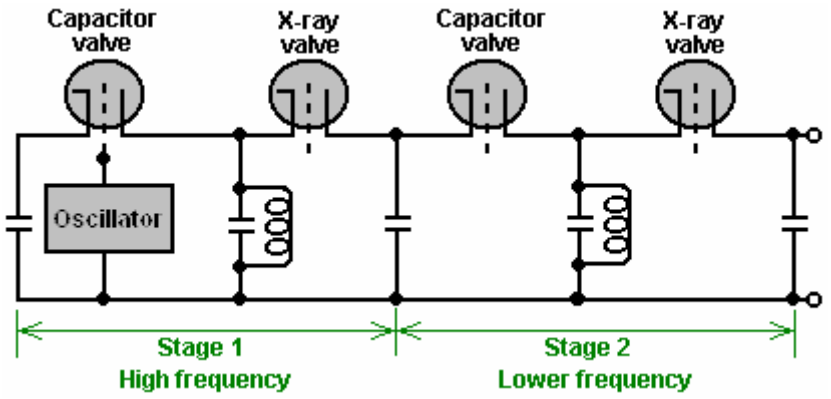


Fig. 2.

Moray King suggests that the circuit used by Thomas Henry Moray was as follows:



There can be little doubt that Thomas Henry Moray built several versions of his apparatus, each of which produced output power well in excess of any input power needed. It seems highly likely that most of them used no input power whatsoever, and if there were any others, they will have been powered by a tiny fraction of the output power. If mild radioactive material was used as described, then the output power could in no way be attributed to that source alone, since the output power was thousands of time greater than any power available from the radioactive materials.

It is perhaps time to explain a little more about, voltage, power and current. We have been raised with the notion that it is necessary to “burn” a fuel to get power, that batteries “run down” when used and that you have to keep turning the shaft of an electrical generator to be able to draw current from it. These things are not actually true. The relatively recent field of Quantum Mechanics shows that if a charge, such as an electron has, is positioned in what is supposed to be “empty” space, it is not alone. The “empty” space is actually seething with energy, to the extent that “virtual” particles are popping into existence for a fraction of a second and then disappearing again. They are called “virtual” because they exist for such a short time.

Because of the negative charge of the electron, the particles appearing and disappearing around it will all be positive in charge. The electron has “polarised” the space around itself because it has a charge. The instant that a positive “virtual” particle appears, there are two charges near each other - minus on the electron and plus on the particle. When you have two opposite charges near each other, they form a “dipole”. Dipoles form a gateway through which energy from the environment flows continuously. An instant later, the particle disappears, but it’s place is immediately taken by another virtual particle. The result is a continuous stream of energy flowing out from the dipole.

Batteries with their positive and negative terminals are electrical dipoles, so too are generators when the input shaft is spun. Permanent magnets with their North and South poles are magnetic dipoles. Both of these have continuous streams of energy flowing through them. So, why then do batteries run down and lose their charge? The reason is that we power circuits using a closed loop. The energy flowing out of one terminal flows into the opposite terminal and instantly destroys the dipole. A new dipole has to be created every split second if the circuit is to deliver power, and it is that self-destructive method of use which causes the battery to discharge or which needs the generator shaft to be rotated continuously.

If a different operating technique is used, where the dipole is not continuously destroyed, then devices which can provide a continuous stream of energy drawn from our natural environment can be constructed. This is not magic, just the next step in conventional science and engineering. Thomas Henry Moray managed it, initially with an aerial and earth like a crystal set to provide the dipole, his device was able to draw many kilowatts of power from the environment. No fuel was needed, the energy is already there surrounding us all, all of the time. As far as I am aware, nobody has managed to replicate Moray's device (which was the reason for it being violently suppressed) but knowing that it existed and was repeatedly demonstrated to work perfectly well, is useful in that it shows that it is possible to tap the massive zero-point energy field with a practical, home-constructed device.

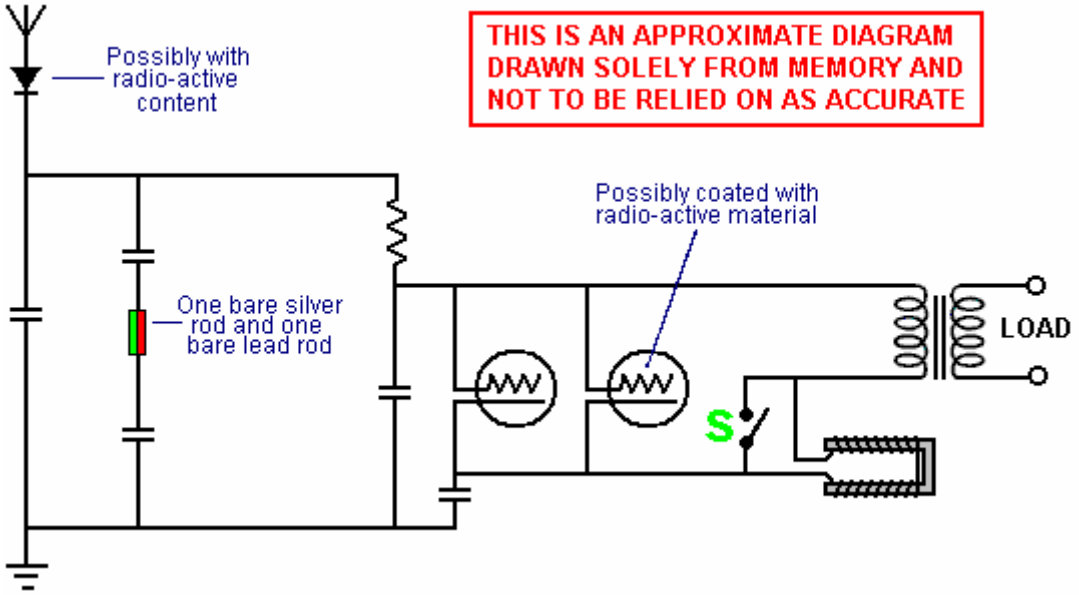
Here is a collection of additional items of information gathered from several different sources:

Moray started his experiments with 'the taking of electricity from the ground', as he described it, during the summer of 1909. By autumn 1910 he had sufficient power to operate a small electrical device, and demonstrated his idea to two friends. The early stages of this demonstration consisted of operating a miniature arc light. It soon became clear to him that the energy was not static and that the static of the universe would be of no assistance to him in obtaining the power for which he was searching.

During the Christmas Holidays of 1911, he began to realise fully, that the energy with which he was working, was of an oscillating nature. He also realised that the energy was not coming out of the earth, but instead, was coming to the earth from some outside source. These electrical oscillations in the form of waves were not simple oscillations, but were surging like the waves of the sea, coming continually to the earth but more in the daytime than at night, but always arriving as vibrations from the reservoir of colossal energy out there in space. By this time Moray was able to gather enough power to light a 16-candlepower carbon lamp to about a half of it's capacity, but he did not manage to gain any further improvement until the spring of 1925.

In 1912 Moray was called to go on a mission for the Church of Jesus Christ of Latter-Day Saints, and under a visitor's visa was allowed to enter Sweden during the Exhibition of 1912 in Stockholm. In his notebook, dated November 1, 1913, he included a note saying that he had obtained material from a railroad car at Abisco, Sweden the previous summer, also some more material from the side of a hill. He made electric tests of these materials, taking them home to try each as a detector for his energy machine. Tests indicated that this soft, white stone-like substance might make a good "valve-like detector". This "valve-like detector" is what led him to do research into semi-conductive materials, and from this soft white stone he developed his first valve and the valve which was used in some of his early Radiant Energy devices (silver wire touching a stone can act as a rectifier).

Moray demonstrated that energy was available by its actions on a resistive load, such as a flat-iron or a space heater, and by lighting bulbs. A resistive device acts as a load which is directly proportional to the amount of energy delivered to it. In heating a heater, or lighting a bulb, the number of watts produced can be calculated as equal to the number of watts provided to the device. This energy is fed into a load to give either heat, light, or power. A motor can be operated but it must be designed to run on a high frequency power supply. The Radiant Energy device used an antenna and a ground connected to his solid state Radiant Energy circuit:



The diagram shown above is reproduced from a rough sketch drawn from memory after seeing Moray's circuit diagram. The person who drew it does not understand how the circuit works, so please treat this diagram as being just an overall suggestion as to what Moray's circuit might have been like. It is actually much more likely that is was a cascade of pairs of tank circuits containing Moray's valve, each pair being one series tank circuit followed by one parallel tank circuit, the oscillating frequency dropping with each tank pair and the output power rising with each tank pair. Moray's circuit was started oscillating by stroking the U-shaped coil with a permanent magnet for a few seconds, and when the circuit started operating, then switch '**S**' was closed, effectively removing the U-shaped coil from the circuit.

Moray was able to demonstrate that none of the output energy came from within his device. Internally the device was electrically dead when it had not been connected and tuned to the antenna. When his device was set up, he could connect it to an antenna and ground, and by priming it first and then tuning it as he primed it, the device would draw in electrical energy. This high frequency electrical energy produced up to 250,000 volts and it powered a brighter light than witnesses had ever seen before. Heavy loads could be connected to the device without dimming the lights already connected to it. This device worked many miles from any known source of electrical energy such as power transmission lines or radio signals. The device produced up to 50,000 watts of power and worked for long periods of time.

Moray initially assumed that this energy was electromagnetic in nature however, he never claimed that it was. He assumed at first that this energy came from the earth but later he believed it was flowing in from the universe. Finally he began to believe that it was present throughout all space, intermolecular space as well as terrestrial and celestial space. He did not necessarily understand how his detectors operated, only that if he built the device very carefully according to his calculations it would work. He was able to demonstrate the existence of an energy that today, though it has not been identified or proven, has been theorised by many researchers.

The largest instrument was about 6 inches high, circular in shape and about 8 inches in diameter. We went out on the roof of the chicken coop carrying the device on a small drafting board, erected an antenna upon the roof of the coop, the antenna being about 100 feet from the house. We pulled the main line switches in the house before going out upon the roof. Mr Judd had Moray move the drafting board from place to place and he also examined the inside of the coop for hidden equipment. The machine was then assembled in his presence and the device was started. Mr Judd timed me to see how long it would take to bring get the light operating. I was able to light the CGE lamp to its full brightness and to heat an old-style Hotpoint electric flat iron to sizzling point, which required 655 watts. Mr Judd asked for the antenna to be disconnected. When this was done, the light went out. The aerial was connected again and the light reappeared. We drove a new grounding rod at a spot selected by Mr Judd, made a connection to the new grounding point and the light burned dim, but came brighter and brighter as the new grounding rod was driven deeper and deeper into the ground.

If the ground or antenna is left disconnected for too long a time, the device becomes electrically dead and must be retuned in order to regain the energy flow. Dr Eyring found no fault with the demonstration and the worst that he could say about it was that it might be induction, but that if Moray would take the device out in the mountains away from all power lines, a distance of three or four miles, and it then operated, he would then acknowledge that it could not be induction and that his theory was wrong.

At last they decided to go up Emigration Canyon, as there are no power lines in that canyon. All three gentlemen were very well satisfied and pleased with what they saw. The antenna wire was put up without any aid or instructions whatever from Mr Moray, except that he suggested that the wire be stretched tighter to prevent so much sag at the centre. This was done and the wire then appeared to clear the ground by about 7 or 8 feet at its lowest point. The ground pipe was of half-inch water pipe consisting of two sections. The lower section was pointed at the end to make its driving into the creek bed easy. It was about 6 feet long and after being driven down about 5 feet the second section, which was about 4 feet long, was screwed on with a wrench and the pipe driven further down until it struck a hard object, so about 7 feet of pipe was in the ground.

The antenna wire was insulated from the poles with two glass insulators about 6 inches long and having holes in both ends. A piece of wire about 2 feet long connected each insulator with the pole. The lead-in wire was fastened to the antenna wire at a point about 10 or 15 feet from the east pole. I helped Mr Moray solder the connection. I paced the distance between the two antenna poles and estimated it to be 87 feet. Mr Moray's equipment, apart from the antenna and ground wires, consisted of a brown box about the size of a butter box, another slightly smaller unpainted box, a fibre board box about 6" x 4" x 4", which Mr Moray referred to as containing the tubes, and a metal baseboard about 14" x 4" x 1" containing what appeared to be a magnet at one end, a switch near the middle and a receptacle for an electric light bulb at the other end. There were also several posts for connecting wires on the baseboard.

When all of the wires were connected and everything was ready, Mr Moray began tuning in. Before tuning, he placed the key on the post: he said it would be in contact while the light burnt, but no light appeared. The tuning consisted of stroking the end of a magnet across two metal projections protruding from what I referred to above as being 'a magnet'. After tuning for slightly more than 10 minutes the key was put on the operating post and the light appeared immediately. Mr Moray put the key on the operating post two or three times before during the tuning operation but no light appeared. We allowed the light to burn for 15 minutes. In my opinion, the brilliance of the light coming from the 100-watt bulb, was about 75% as bright as a 100-watt bulb connected to an ordinary house socket. It was an steady light, without fluctuations of any kind.

While the light was burning Mr Moray disconnected the antenna lead-in wire from the apparatus and the light went out. He connected it again and the light appeared. He also disconnected the ground wire and the light went out. He then connected it and the light appeared again.

In another demonstration, Mr Moray opened the device and let everyone see everything except one small part that he placed his hand over and hid in his fist. This part he cut off and put in his vest pocket. Everything else, people were allowed to examine to their hearts' content. "If that part is able to make such power itself, then it's some device and worth selling. Such a battery would be worthwhile", were some of the remarks made.

On several occasions Dr Moray would disconnect the antenna wire momentarily, but not long enough to lose the light. In disconnecting and connecting the antenna wire a flash of electricity could always be seen at the

At a demonstration in 1928, the aerial used was about 200 feet long and positioned about 80 feet above the ground: the wire is a copper cable approximately a quarter of an inch in diameter, and well insulated. The earth connection used was the water pipe in the basement of Dr Moray's home. The device was assembled in a trunk through the sides of which were holes for the connections to ground and to the antenna and for observation; the holes were about one-half inch in diameter. There were two boxes about 10 by 20 by 4 inches, one on top of the other; both were closed and the covers fastened with screws. On the upper box was lying an insulating panel about an inch thick by 15 inches long and 3 inches wide; it was made of slate or hard rubber or some material of similar appearance. On this were two binding posts which could be connected together by a small switch; also mounted on this panel was an object about 2.5 inches square, wrapped in friction tape, from which protrude two poles about 1/4-inch in diameter, apparently of sift iron. Two light bulb sockets were connected in the circuit. In one of these there was a 20-watt bulb, and in the other a 100 watt bulb.

Dr Moray then took a magnet, which was a very broad, short limbed U, and began to stroke one pole of it on the poles in the taped body. Mr Jensen placed his fingers on the binding posts several times, and at last received a rather vigorous shock. Mr Moray then threw the switch and the bulbs lit up. As a further proof that the conversion of the energy was due to the mechanism in the box, Dr Moray hit the table on which the trunk was standing, a moderate blow with a hammer whereupon the light flickered and went off, due to the detector being shaken out of adjustment. The boxes, in which the mechanism had been housed during the test, were opened and the contents examined. There were capacitors, the detector, a transformer, and two tubes in them but nothing else. Nothing that in the least resembled a battery.

It is to be noted that after a total run of 158 hours the device supplied 635 watts; inasmuch as a horsepower is but 746 watts this equals 0.878 of a horsepower or slightly more than 7/8 horsepower. This alone is sufficient to dispose of any suggestion of a battery.

A report from 1929 says: It is now more than 2 years since I first became acquainted with Dr T. H. Moray and the work he is carrying on, and in that time he has demonstrated inventive ability of an exceptional order. Perhaps the most wonderful of his inventions is a device whereby he is able to draw electric power from an antenna. This energy is not derived by induction from power lines, as has been suggested by some, nor is it derived from radio stations, as has been demonstrated by taking the apparatus more than 26 miles from the nearest power line and over a hundred miles from the nearest radio station and showing that it operates just as well as anywhere else.

This device was subjected to an endurance test in which it was operated continuously for a week, and at the end of that time a 100-watt lamp was lighted simultaneously with the heating of a 575 watt standard Hotpoint flat iron, making a total of 675 watts; it is very evident that no batteries could sustain such a drain as this.

He has also invented a very sensitive sound detector whereby it is possible to hear conversations carried on in an ordinary tome of voice at a distance of several blocks. He has also worked out numerous radio hook-ups which eliminate many of the parts now considered necessary for good reception, yet there is no diminution in quality or volume; in fact, there is a notable elimination of interference from static when some of these are used. He has devised a means by which he is able to measure with some degree of accuracy the energy evolved during mental activity; that is, he gets definite, variable deflections of the needle of a sensitive galvanometer which appeared to be related to the vigour of mental activity. There are a great many other equally remarkable things which he has done, such as reducing old rubber from truck tires to the state of a viscous fluid which is readily vulcanisable without the addition of smoke sheet as is necessary with other processes; also a high frequency therapeutic device, and numerous other devices which show great ingenuity.

The 6 lamps are set up in parallel and a small diameter wire is used as the current enters the tube prior to and connecting with the step-down transformer, this takes the very high voltage to the transformer. This voltage will jump across a spark gap of at least six inches. The operating frequency is so high that I have no instrument in my laboratory that is able to measure the amperage or the voltage at this frequency. (Signed, Murray O. Hayes, PhD.).

Dr Milton Marshall was attempting to identify the material that Moray called his "Swedish Stone". Moray described the radio detector which he had developed. He compared it to what was commonly known as the crystal of a crystal set. However, his detector was superior since it could drive a loudspeaker without the use of a battery. He used the most easily demonstrated device, the germanium diode, that worked on the same principle to illustrate how he thought the Radiant Energy Detector worked (Moray originally built the radio simply for the purpose of showing how he was able to pick up radio signals with a solid-state device, producing sufficiently strong signals which could drive a loudspeaker, which was something unheard of in that day. His circuit did not have batteries, and it was very similar to the old crystal-set circuitry.

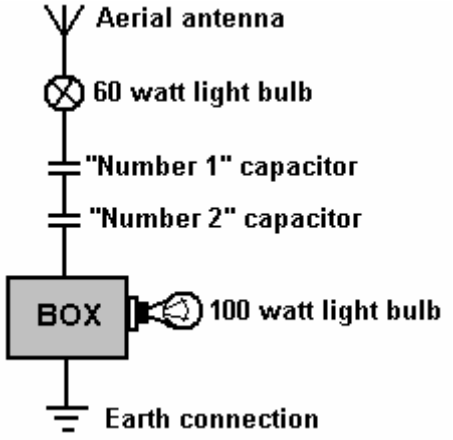
The device was housed in a wooden box something like 12" by 18", with an antenna and a ground going into it. Wires leading out of the box led to a bank of some forty 100-watt light bulbs and to an electric iron. Moray touched a switch at the top of the box with a hand electrostatic plate and the globes all lit up brilliantly. We all noted that the bulbs burned cold except each had a hot spot about the size of a dime on the top slightly off centre. I also recall that I could turn the lights on and off by approaching and retreating to and from the device, either with my whole body or my hand. If my memory is clear, the machine had to be tuned with a dial to be placed in this condition. (Chester M. Todd, 1971)

In 1938, after examining the transformer of the device, Mr E. G. Jensen stated that he considered that the amount of current which he had seen taken from the device was positive proof that the current developed by or in the machine was different to any in use at that time. This was because the transformer would have burned out if it had been carrying normal current, but the transformer showed no signs of even ever having been warm. He was informed by Dr Hayes that the transformer had been in use under the same loading conditions during many demonstrations in the past.

The "Number 1" capacitor consisted of two small sheets of aluminium of about 30 gauge, separated by and making contact with a piece of one-quarter inch thick plate glass. The plate glass was larger than the aluminium sheets and overlapped them.

The "Number 2" capacitor was a commercial unit manufactured by Igred Condenser & Mfg. Co. and had a capacity of 0.025 mfd.

They were used as shown here:



With the 60 watt lamp and the two capacitors attached to the antenna and the antenna and ground attached to the box containing the Radiant Energy equipment as shown in the sketch, the 100-watt lamp on the secondary or output side was lighted. Unscrewing the 60-watt lamp from its socket caused the 100-watt amp to go out, but it immediately lighted when the 60-watt lamp was screwed into its socket again. The 60-watt lamp did not light. Shorting the antenna and ground by placing a wire across them, caused the 100-watt lamp to go out. Similar shorting with the hands also caused the 100-watt lamp to go out. No electricity could be felt when shorting with the hands. If either the ground or the antenna wires were disconnected from the box, the 100-watt lamp would go out. Neither of the capacitors or the 60-watt lamp on the primary side of the box were necessary but were simply put there to show that the high frequency power will jump or pass through them.

Moray's patent application on this device was filed in 1931 and rejected on a number of grounds. Firstly, "Because no means was provided for causing the cathode to emit an appreciable number of electrons, the current produced in the cathode by the antenna will not heat the cathode to a temperature at which an appreciable number of electrons per second are emitted". In other words, according to Thomas E. Robinson, Commissioner of Patents, a solid state device, such as a transistor, cannot possibly work. Secondly, because "No natural source of electric wave energy is known to the Examiner and proof of the existence of such a source is required". In other words, it was not enough for Moray to demonstrate the effect of the energy source; he also had to identify it, which he could not do. None of the original patent applications that Henry made are any longer available at the US Patent Office. Although their file jackets are there, the contents and applications themselves are gone.

In 1942, Moray attempted to rebuild a Radiant Energy device, using the remaining bit of what was known as the "Swedish Stone". This material, which was the heart of his original RE detector, he had never managed to duplicate, and the shortage of this material limited the amount of power which he could draw. Consequently, in the large unit, he developed a second detector that forced him into extensive research involving nuclear materials and radioactive reactions. He became deeply involved in the study of synthetic radioactivity as described by Gustave LeBon in his book "The Evolution of Matter". The years slipped by and Moray spent most of his time working on what he called the "counter-balance" to eliminate the need for an aerial antenna.

Moray said:
Enough energy is coming to the earth to light over 1,693,600 100-watt lamps for every human being on the earth today. No fuel of any kind need be taken as this energy can be picked-up directly by ocean liners, railroads, airplanes, automobiles, or any form of transportation. Heat, light and power can be made available for use in all kinds of buildings and for all kinds of machinery. An example would be to pump water onto the desert lands, the power source being only a fraction of the weight of any steam plant or any kind of engine in use today and all this at a fraction of the current cost.

The total energy involved in "cosmic" radiations is very large. The mechanism of its generation involves a basic relationship with the total structure and action of the universe. Today it is believed that cosmic radiation consists primarily of protons and some heavier nuclei. At times this cosmic energy packs a wallop of around 100 quadrillion volts. Coming continuously with slight variations in time, the radiations have a uniformly directional isotropy. The earth is, therefore, surrounded in an atmosphere of radiation with cosmic rays coming continually to the earth from all directions, although there may be a slight deflection of the weaker rays by the earth's magnetic field. There is every indication that our sun is not the source of any appreciable amount of this radiation. The origin, therefore, is from the universe as a whole. The total energy of cosmic radiation is more than the entire luminous output of all the stars and nebulae of the universe combined. Unlimited power is being delivered to everyone's doorstep.

The Moray Radiant Energy discovery, using radiations from the cosmos as its power source, gives the greatest amount of energy per pound of equipment of any system known to man. Electrical power through an electric

motor or an electric jet far exceeds any form of energy in any engine in the delivery of power. There is no dead centre of lost motion in an electric motor nor loss of push in an electric jet. Also, the starting torque is much higher in the electrically powered engine than in the combustion engine.

Harnessing cosmic energy is the most practical method yet discovered by man. Furthermore, it is possible to utilise this vast source of energy from the universe without a prime mover at any point on the earth --- on the ground, in the air, on the water, under the water, or even underground. If one considers that an electrical generator is not in the true sense a generator - as electricity is not made by the generator - but is merely an electrical pump, the Moray Radiant Energy device may then be referred to as a cosmic ray pump: that is, a high speed electron oscillator serving as a detector of cosmic radiations which causes a pumping action or surging within its circuitry.

To account for the propagation of heat and light - two of the forms of Radiant Energy - man has postulated the existence of a medium filling all space. But, the transference of the energy of radiant heat and light is not the only evidence in favour of the existence of such a medium. Electric, magnetic, and electromagnetic phenomena and gravitation itself point in the same direction.

Attractions and repulsion take place between electrified bodies, magnets, and circuits conveying electric currents. Large masses may be set in motion in this manner, acquiring kinetic energy. If an electric current is started in any circuit, corresponding induced currents spring up in all very closely neighbouring conductors. To originate a current in any conductor requires the expenditure of energy. How, then, is the energy propagated from the circuit to the conductors? If we believe in the continuity of the propagation of energy - that is, is we believe that when it disappears at one place and reappears at another it must have passed through the intervening space and, therefore, have existed there somehow in the meantime - we are forced to postulate a vehicle for its conveyance form place to place.

When a particle is electrified, what one must first observe is that a certain amount of energy has been spent; work has been done. The result is an electrified state of the particle. The process of electrifying a conductor is, therefore, the storing of energy in some way in or around the conductor in some medium. The work is spent in altering the state of the medium, and when the particle is discharged, the medium returns to its original state, and the store of energy is disengaged. Similarly, a supply of energy is required to maintain an electric current, and the phenomenon arising from the current are manifestations of the presence of this energy in the medium around the circuit. It used to be that an electrified particle or body was supposed to have something called "electricity" residing upon it which caused electrical phenomena. An electric current was regarded as a flow of electricity travelling along a wire (for example), and the energy which appeared at any part of a circuit (if considered at all) was supposed to have been conveyed along the wire by the current. But, the existence of induction and electromagnetic interactions between bodies situated at a distance from each other leads one to look upon the medium around the conductors as playing a very important part in the development of these electrical phenomena. In fact, it is the storehouse of the energy.

It is upon this basis that Maxwell founded his theory of electricity and magnetism, and determined the distribution of the energy in the various parts of an electric field in terms of electric and magnetic forces. The medium around an electrified body is charged with energy and not of an imaginary electric fluid distributed over the electrified body or conductor. When we speak of the charge of an electrified conductor we are referring to the charge of energy in the medium around it, and when we talk of the electric flow or current in the circuit we are referring to the only flow we know of, namely, the flow of energy through the electric field within the wire.

The work in producing the electrification of a conductor is spent on the medium and stored there, probably as energy of motion. To denote this we shall say that the medium around the conductor is polarised, this word being employed to denote that its state or some of its properties have been altered in some manner and to a certain extent depending on the intensity of the charge. If the charge is negative the polarisation is in the opposite sense, the two being related, perhaps, like right-handed and left-handed twists or rotations.

Now consider the case of a body charged alternately, positively and negatively in rapid succession. The positive charge means a positive polarisation of the medium, which begins at the conductor and travels out through space. When the body is discharged the medium is once more set free and resumes its former condition. The negative charge now induces a modification of the medium or polarisation in the opposite sense. The result of alternate charges of opposite sign is that the medium at any point becomes polarised alternately in opposite directions, while waves of opposite polarisations are propagated through space, each carrying energy derived from the source or agent supplying the electrification. Here, then, we have a periodic disturbance of some king occurring at each point, accompanied by waves of energy travelling outwards from the conductor.

The phenomenon of interference leads to the conclusion that light is the result of periodic disturbances or vibrations of the medium, but as to the nature of these vibrations, as to the exact nature of the periodic changes or what it is that changes them, we possess no knowledge. We know that alternating electric charges are

accompanied by corresponding changes of state or vibrations of the medium, and if the charge is varied periodically and with sufficient rapidity, we have a vibration at each point analogous to, perhaps identical with, that which occurs in the propagation of light - a combination of wave and particle properties. This then is the electromagnetic theory of the luminous vibration.

In the older elastic-solid theory, the light vibrations were supposed to be actual oscillations of the elements or molecules of the medium about their positions of rest, such as takes place when waves of transverse disturbance are propagated through an elastic solid. Such limitation is unwarranted to some extent, but one cannot afford to entirely disregard the particle theory of light either. A combination of the theories has merit. We know that the change, disturbance, vibration, polarisation, or whatever we wish to term it, is periodic and transverse to the direction of propagation. The electromagnetic theory teaches us nothing further as to its nature, but rather asserts that whatever the charge may be, it is the same in kind as that which occurs in the medium when the charge of an electrified body is altered or reversed. It reduces light and heat waves to the same category as waves of electrical polarisation. The only quality of the later required to constitute the former is sufficient rapidity of alteration. These speculations were given the strongest confirmation by experiments of Prof. Hertz many years ago.

When a resilient substance is subjected to strain and then set free, one of two things may happen. The substance may slowly recover from the strain and gradually attain its natural state, or the elastic recoil may carry it past its position of equilibrium and cause it to execute a series of oscillations. Something of the same sort may also occur when an electrified capacitor is discharged. In ordinary language, there may be a continuous flow of electricity in one direction until the discharge is completed, or an oscillating discharge may occur. That is, the first flow may be succeeded by a backrush, as if the first discharge had overrun itself and something like recoil had set in. The capacitor thus becomes more or less charged again in the opposite sense, and a second discharge occurs, accompanied by a second backrush, the oscillation going on until all the energy is either completely radiated or used up in heating the conductors or performing other work.

When capacitors are filled with energy captured by the Moray Radiant Energy device and then discharged through a circuit of proper impedance, reactance and inductance, thereby synchronising the oscillation of the device with those of the universe, electrical inertia is set up. In the reversal of the current, the capacitors are charged, discharged and recharged slowly until the energy stored in them is radiated in kinetic energy through the device, and this energy can be kept alive indefinitely by establishing resonance with the oscillations of the universe.

Considering oscillations from a mechanical, electrical and mathematical point of view, we find that electrical resistance is the same as mechanical friction and current is comparable to mechanical velocity. Inertia and inductance may then be considered analogous terms. In mechanics the greater the inertia of a body, the longer it will stay in motion. In the Radiant Energy device's resistance-inductance-capacity (REC or RLC) circuit, the greater the electrical inductance, the longer the current continues to flow once it is established by synchronisation with cosmic surges.

Expressed mathematically, the equations are the same for electrical or mechanical phenomena. Which means, that $R < \sqrt{(4L / C)}$, where R is the resistance in ohms, L is the inductance in henries, and C is the capacitance in farads. When this is true, an oscillatory discharge will occur and a very powerful inductance inertia will assert itself. For low values of R, the frequency of the oscillations can be shown by $f = 1/2 \pi \sqrt{(CL)}$. The rapidity of the oscillations is governed by the capacitance and inductance.

In the vibrational forces of the universe, we find the key to the source of all energy. How we can utilise this energy for modern industry without being limited to mechanical prime movers is the question. And, the answer may be an energy generator, balanced so as to oscillate in synchronisation with the oscillations of the universe.

Dr Ross Gunn, a civilian scientist for the US Navy, stated years ago that the earth is a huge generator, generating over 200 million amperes of electric current continuously. For example, the aurora borealis is considered to be a very large definite electrical phenomenon produced by the passage of electric charges through the rarefied gases of the higher atmosphere. The earth has since been shown, by Dr Gunn and others, to have a negative charge amounting to 400,000 coulombs. Yet, six feet above the ground the air is charged with more than +200 volts with respect to the ground.

It is known that air conducts electricity away from charged objects. This being true, how does the earth maintain its charge since it is a charged object exposed to the surrounding atmosphere? If the air conducts electricity, the earth's charge must be constantly passing into the atmosphere. And it has been calculated that the earth has a continuous discharge into the atmosphere of 1,800 amperes. At this rate, the earth should lose 90% of its charge into the air in one hour, yet the earth's charge does not diminish. From where does the earth's energy come?

The conversion of matter to energy in the stars is accepted, and, reasoning from what occurs in radioactive disintegration during which energy waves are radiated, one may conclude that energy waves of very high

frequency are sent out from the stars (one of which is our sun). Now, of course, the conversion of energy into matter must equally be accepted.

It has been found that ionisation, which could be the medium for the flow of energy, increases with increasing altitude, instead of decreasing as would be expected. Since the source of energy is the universe, the generation of energy by rotary action and by all prime movers is an effect and not a cause. Oscillatory energy action, be it in a Leyden jar, another man-made capacitor, or in what we may call natural capacitors, always behaves the same. The oscillations will continue until they have reach their cycle of height and then there will be a backrush returning to where the oscillations originated. Every oscillation, whether large or small, is completed during the same interval of time. These oscillations all prove the same great fact, that they are governed by the same cycle of time, completed during the same interval of time. Waves of energy have a regular beat note, coming and going as the waves of the sea, but in a very definite mathematical order - coming to the earth from every direction with a definite rhythm.

Energy has a definite elastic or resilient rigidity and density, which is subject to displacement and strain. When strain is removed, the medium will spring back to its old position and beyond, surging back and forth, and will continue to oscillate until the original pressure is used up. If the internal impedance is too great, there will be no oscillations, but it will merely slide back in a dead beat to its unrestrained state.

By cutting down resistance to a minimum and synchronising the resilient ionic actions of the Moray device with the wave actions of the universe, periods of oscillation can be made to come quicker and quicker until inertia asserts itself, thus lengthening out the time of final recovery. This is done by carrying the recoil beyond the natural oscillations and prolonging the vibrations by capturing the in oscillatory action. When the recovery becomes distinctly oscillatory, a harmonic pattern is initiated and the oscillations continue, resonance thereby being established with the universe.

In the universe we see the same laws being obeyed as in our laboratories. As one traces down to the almost infinitesimal constituents of the atom, one finds that matter does not exist at all as the realistic substance which we have supposed it to be. There at the very foundation, it consists of nothing more than energy charges emitted at various wavelengths or frequencies. It is becoming more and more certain that the apparent complexity of nature is due to our lack of knowledge. And, as the picture unfolds, it promises a marvellous simplicity.

One of the most marvellous relationships that has ever been revealed in the entire science of physics is that between light and electricity and the existence of electronics in atoms of matter. Knowing what we do at the present time with regard to the structure of atoms, this relationship is not quite so surprising. However, considering the total absence of this knowledge about a half century ago, the discovery that light, and radiation in general, are vibratory phenomena was revolutionary.

Speaking of radiation, "Radiant" here means proceeding from a centre in straight lines in every direction. Energy is internal and inherent. "Energy" is defined as a condition of matter, by virtue of which, any definite portion may affect changes in any other definite portion. This was written in 1892, and discoveries since confirm it. Energy then is a state of matter, or rather, the result of a particular state or condition in which matter may be when any observed phase of energy appears.

In addition to possessing kinetic energy, the atom is capable of absorbing energy internally. This internal energy is associated with the configuration of the particles of which the atom is composed. Under ordinary conditions an atom is in what is known as a state of equilibrium, in which there is neither a giving off, nor an absorbing of energy. But, the internal energy of the atom can be altered. When the internal energy of the atom exceeds that of its normal state it is said to be excited. Excitations may be caused in several ways, e.g., the collision of an atom with rapidly moving positive or negative particles or the breaking of lines of force in an electromagnetic generator. Kinetic energy is released when excitation causes a particle to give up some or all of its kinetic energy to the atom during collisions. This is taking place in the universe all the time.

The electric motor and generator would never have been discovered if a dielectric (insulator) had not been discovered. If one discovers a dielectric valve for the energy of the universe, one has the answer to harnessing the energy of the universe! A limiting case of excitation is ionisation, wherein energy is absorbed by the atom sufficiently to allow a loosely bound electron to leave the atom, against the electrostatic forces which tend to hold it within the atom. An atom which has given up one or more electrons is said to be ionised. It is possible that ionisation, i.e., excitation, may take place in successive steps through absorption of quanta energy. The return of an ionised atom to a state of lower energy is associated with electromagnetic radiation. Also, from the process of ionisation, electrical energy may become associated with the vibrational forces of the universe coming into the earth as cosmic radiation. The higher the frequency, the greater the ionisation or excitation, a form of energy which is kinetic in nature. There are tremendous energies coming to the earth from outer space. These energies are only different manifestations of the energies we see in operation all around us. In most cases we are not even aware of their existence. They penetrate everything including our own bodies. Every one of us is alive by virtue

of these energies. Every part and particle of the universe is alive with them. The generators that now furnish our electric power do not create or originate any power or electricity; they merely direct, pump, the existing energy or electricity.

As in musical notes of high and low "C", the vibrational rates (frequencies) are different, but all "C" notes are essentially the same (harmonically related). This is the foundation upon which much of my investigation of vibratory phenomena is based.

It has been agreed that all forms of matter are vibrating at a particular rate or frequency. And, so it is with the various forms of energy - heat and light, magnetism and electricity. These are but forms of vibratory motion connected with and being generated from the same source, the universe. Matter vibrates at a particular rate, according to its character, and may be transmitted into other substance by lowering or raising its rate of frequency. If the frequency is raised high enough, the molecules will separate and the atoms become free. Raising the frequency still higher, the atoms resolve themselves into their original components. Matter then becomes a form of energy. Frequencies may be developed which will balance the force of gravity to a point of neutralisation. One can then go beyond the force of gravitation. Understanding the principles of vibration is truly understanding energy.

In gamma rays, we find potentials which are equivalent to as much as 1,000,000 volts, yet their wave lengths are not the shortest known. In octaves still higher there are rays which are known as "cosmic rays". Who can draw a definite line and say how much higher other octaves exist than those knows as the cosmic rays? Our starting point from the discovery of these different waves was electrical conductivity of the air, and it has been found that this conductivity is just as strong by night as by day. Radiations emitted by the sun can scarcely be the sole cause of this energy. All space is saturated with vibration, energies, which are no doubt electrical in character. The relation of matte to energy and energy to matter then becomes the potential of the universe - one continuous series of oscillations.

Atoms maintain an equilibrium by oscillations, rotations, attractions and repulsions, but this does not interfere with a transformation of equilibrium, which, when the transformations of equilibrium are rapid enough, become energy, i.e., matter is turning into energy and energy into matter.

There can be no generation of electrical current and no kinetic energy if there is no disturbance of equilibrium, i.e., change of potential or change of energy levels. When one thinks of the oxygen and nitrogen molecules of the air all about us moving with the speed of bullets and striking us and everything else at this speed, one can form some idea of the agitation taking place here and in the universe.

The oscillations from outer space are emitting electromagnetic waves of many wavelengths and frequencies. The Moray device is so constructed that the frequency is very much lower on the secondary side than on the primary side, and almost complete resonance is established. I am convinced that the energies from the universe are active radiations produced by the evolution of matter into energy and energy into matter.

Dr Anderson's cloud chamber at the California Institute of Technology, in which the positron was discovered, has furnished much information about cosmic ray energies. He found that some positrons are born of cosmic rays smashing into matter. The cosmic ray energies deduced from the tracks left in the Anderson cloud chamber range from 100 volts to 3,000,000,000 volts. The Lemaitre-Vallarts theory, together with Dr Johnson's asymmetry measurements, give definite values for the energy of half of the cosmic radiation, and shows it continuously distributed between 5 billion and 50 billion volts.

The figure of 100 billion volts is a result of Dr W. Kolhorster's measurement of penetrating radiation in the depths of the Strassfurt salt mines. He found that the minimum energy of these rays had a penetration which was greater than ever before demonstrated. Dr Axel Corlin of Sweden's Lund Observatory found radiation that still had energy after passing through somewhat greater depths and, therefore, the voltage figures can be made even higher. Energies of 100 billion volts or more are indicated by the great bursts set off by cosmic ray collisions, called the stosse, which have been observed particularly in Germany. The Moray RE devices have worked equally well in deep mines, under water or high in the mountains and in an airplane.

It is about 100 years since science began to consider light, heat, magnetism, galvanism, and electricity as natural forces. In the early part of the 19th century school books termed these things "imponderable substances". The corpuscle theory of light was taught, the sun was supposed to provide an endless supply of those corpuscles. After the corpuscle theory faded, scientists turned to the wave theory, but even that was based on a crude concept of movement of the ultimate principles or atoms, of matter. The electron theory has superseded the earlier ones now, and while the electron theory explains the observed and theoretical "facts" better than the previous concepts did, could it be that, as the greater light of knowledge leads us on, the electron theory in turn will fall short of providing "absolute" knowledge? The Einstein Theory may stand in need of revision or amendment; or, in time, it may join the theories of corpuscles and waves on the back shelf.

A specific case, in which the electric field performs the double function of molecular excitation and the creation of intermolecular and atomic ions, is being given by the system used by the inventor. It is a system utilising the principles of the wire corona with a concentric cylinder at different pressures. The system is modified in conformity to the concept that chemical reactions must take place when the oppositely charged molecular ions from an appropriate activated catalyst are accelerated against one another in the wire corona. It consists of a cylinder made of a suitable catalyst from which positive ions are emitted. The reactants (gases) streaming through the chamber parallel to the length of the wire attain the polarity of the negative molecular ions by the high electric field close to the wire. As these negative molecular ions are accelerated at the right angles to the wire in the direction of the electric field toward the positively charged catalyst cylinder, they are met by an avalanche of onrushing atomic ions from the catalyst. A certain amount of reaction takes place in that instant, 10^{-8} seconds. However, some of the negative molecular ions outside the mean free path of the positive atomic ions are free to rush headlong toward the positive cylindrical field where they are neutralised, and instantly given a positive charge by the avalanche of out-rushing positive ions. These positive molecular ions are accelerated back into the field and collide against the negative molecular ions coming from the direction of the negative electrode corona. This melee continues until the reaction has come to a point where the individual participants are either all gone or the mixture is outside of the electric field: backrush oscillations.

The Moray apparatus combined with other equipment, consists of a combination of specially constructed tubes which we will refer to as valves, "pressure transmitters", interceptors and oscillators. The valves are not rectifiers in the sense that they operate as radio valves in changing Alternating Current or High Frequency oscillations into Direct Current. They have an actual valve action in stopping the "flow" of energy which may be thought of as oscillatory action similar to the waves of the sea, without rectification, from returning to the outer circuit, much as a retaining wall could stop the waves of the sea from returning. The other modalities and "tubes" of the device are equally unique in their performance. Although no new laws of energy are being advanced or claimed as having been discovered, the application in the method of utilisation of the energy throughout space is unique in that "generation" is accomplished by oscillatory utilisation rather than by the conventional prime mover. These detector tubes have a synchronised pull with the specially developed oscillators of high faradic capacity and provide a means through which oscillating energy may pass to specially constructed valve oscillators whose relation to the first stage valve is such as to permit oscillations to come in from but not return to the outer circuit with an automatic variable relation to the oscillations from the universe, and capable of setting up within their circuits initial oscillations which coincide with the oscillations of the universe.

Special provision is provided to stop RE tubes from becoming blocked in their dissipation of the charges created by the oscillations that continually accumulate based on the oscillatory capacity backrush effect common to capacitors and are herein applied in vacuum tubes. This action of these devices has the effect of enlarging and prolonging the time of charge and discharge of the capacitors and the capacity energy in the circuit to an appreciable interval in perfect harmony with the natural energy wave through the interceptor's valves and oscillators in the circuit which set up in the circuit electrical pulsations corresponding to the energy waves captured by the interceptor and again kept from returning to the second outer circuit by "multi-walled" valves. The final tubes act as energy pressure transmitters with a means to prevent "shunting" condensation by a special form of "getter". This stops condensation accumulating at the base of the tubes which would block their ionic action.

One must "split" the energy discharge band into lines of variation (call this what you will), lines of energy or lines of light beyond the "light rays". The oscillations, therefore, do not become simple oscillations but through the action of the universe set up an energy flow which might be referred to as the assertion of inertia. When inertia sets in, the action will continue because of the oscillations of the cosmos, otherwise one would have a complete dissipation of energy and no oscillations. The oscillation will vibrate during the same period of time regardless of the potential, but the rate of vibration of the device depends on the "capacity" of its modalities, i.e., condensers, etc.

Put together in pure energy resonance, certain energy responding apparatus which synchronise with the resonance of certain vibrations in the universe, and what do you have? Useable energy from the universe. This energy may come to the planets as oscillations similar to the oscillations and tides of the sea. The Radiant Energy tubes receive this energy in surges which may last only a few microseconds by the pressure and current in those surges are so strong that sufficient energy is delivered to the equipment in resonance to be useable in multiples of flashes and in a magnitude which competes with the light of day. Remember resonance and pressure can do a lot to amplify energy. Also remember that the vibrations going out from the sources in the universe must also return to their sources. Nothing is lost. There is only a lowering of potential like water flowing over a water wheel.

The Radiant Energy tubes present no new laws of physics. They simply expand the application of known laws, thereby obtaining results not at first thought possible. This is the history of science. Radiant Energy tubes possess greater ability to obtain "saturation" and thus charge the accompanying capacitors at a more steady rate.

When a certain voltage is reached, ionisation occurs in the gases of the discharged tube and causes the capacitors of the valve circuit to discharge into other capacitors of the valve circuit, to discharge into other capacitors of the oscillators and the other modalities of the circuit.

When ionisation in the preceding tubes is no longer possible because of the reduced voltage, the process starts all over again. The first valve passes vibrations of energy into an oscillatory circuit; ionisation sets in, a discharge occurs, and energy passes through another valve into other oscillators. The process is repeated from the first stage on to the second stage, on to the third and so on, much like a bucket brigade. That is why I asked years ago, "Cannot a steady flow of water be obtained from the waves of the sea or energy from the vibrations of the cosmos?"

When a vibration of any kind strikes a boundary between two media of different vibratory impedances at an angle of less than 90 degrees, a transformation of the vibratory rate may be changed into another vibratory rate. The Radiant Energy device therefore will continue to capture energy by resonance, or call it what you will, as long as the "keep alive" vibration of the cosmos continues to oscillate the various stages of the valves and oscillators in the circuit. Simple, is it not? Just a case of the trapping of energy which is everywhere present in the primary circuit and causing it to oscillate through the secondary circuits through a blocked circuit of no return.

Our experiments have proved that there is an energy which exists in the universe which, by proper development of equipment, can be made available for commercial use.

Such an energy transformer or converter has been built. It has been operated, at full load continuously with no expenditure of fuels of any type, without a mechanical prime mover, kept alive by the oscillations of the energies from the cosmos; an energy converter, or transformer, which would be capable of converting the high frequency, high level energy of the cosmic radiation into current of usable frequency and voltage.

Basically the theory of operation is as follows:

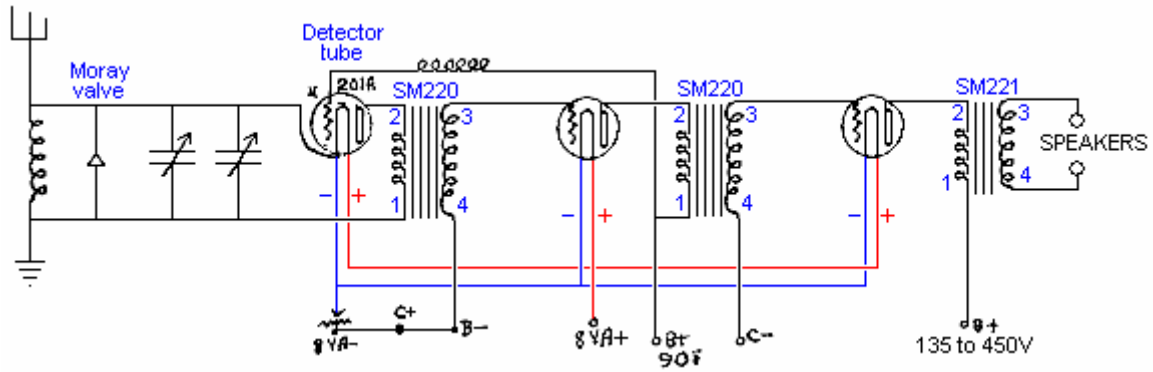
Oscillations are started in the first stage or circuit of the device by exciting it with an external energy source.

The circuit is "tuned" until the oscillations are sustained by harmonic coupling to the cosmic wave frequencies.

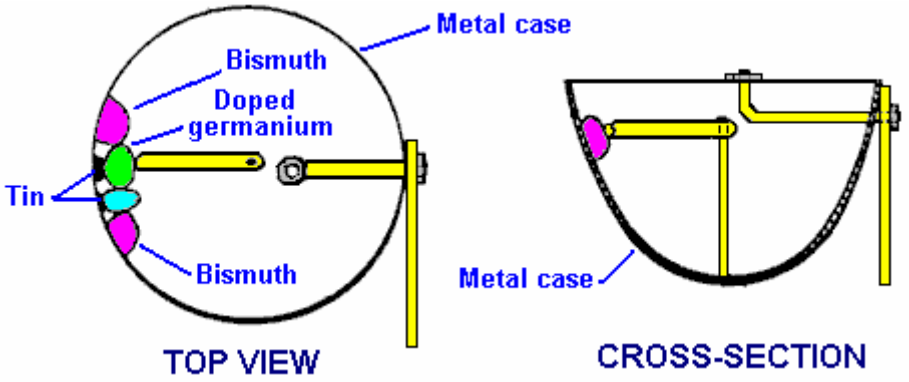
The reinforcing action of the harmonic coupling increases the amplitude of the oscillations until the peak pulses "spill" over into the next stage through a special detector or valve which prevents the return or feedback of energy from succeeding circuits.

These "pulses" drive this stage, which oscillates at a lower frequency and is again reinforced by harmonic coupling with the ever present cosmic waves.

The second stage drives a third stage, and additional stages are coupled until a suitable power level at a useable frequency and voltage is obtained by means of special transformers.



The specific information on Moray's system is very limited, especially since his patent application has been removed. The diagram above and the diagram below have been reproduced from what is alleged to be the notes from which the patent application was composed. These notes are not very clear both in wording and in the quality of reproduction, however, the diagrams shown here are an attempt to show clearly anything which is reasonably certain in those diagrams.



This is supposed to be the construction information on the Moray Valve which was capable of being set to either rectify a signal or to amplify a signal. The casing is a metal cup which also forms one of the contacts for the valve. inside the cup there are four pellets attached to the side. The outer two pellets are made of bismuth and are fused directly on to the metal case. The two inner pellets are attached to the case with tin instead of solder. Judging from the drawing, it looks as if the metal arm contacting the pellets can only connect with the inner two pellets. The arm presses sharply against the pellets in the same way that the "cat's whisker" diodes of the time were touched with a silver wire to make a point contact and produce rectification.

If it is correct that the rotatable arm only contacts one of the two inner pellets, then the reason for those outer bismuth pellets must be as an indirect part of the valve. So, this section of the case is an arrangement of the metal of the case, tin, bismuth and five junctions between different materials, not counting the contact arm. One of the two inner pellets is made of purified germanium with the addition of very small amounts of a doping material. Iron Sulphide (FeS), Molybdenum Sulphide (MoS), Bismuth, Uranium and Silver have been mentioned as possible doping agents. Another material mentioned is Lead which has had it's structure altered by the process described in Moray's patent US 2,460,707. The pellets are said to be produced under high pressure.

From this it can be seen that we do not have anything remotely like the full information on Moray's system. However, there are a number of important things which we can learn from this. Firstly, using just a good earthing connection and an aerial of just ninety feet (30 m) or so in length suspended only some eight feet off the ground, it is possible to draw significant current from the environment. The photograph shows 35 light bulbs being lit by Moray and that is a substantial amount of power. It is unlikely that we will be able to reproduce Moray's exact method of extracting power, but it is highly unlikely that his method is the only possible way of achieving efficient power extraction. So, if we experiment with the components and materials to hand today, it is distinctly possible that we could extract major amounts of power from a relatively small aerial wire positioned at quite a convenient height above the ground, and a good quality earth.

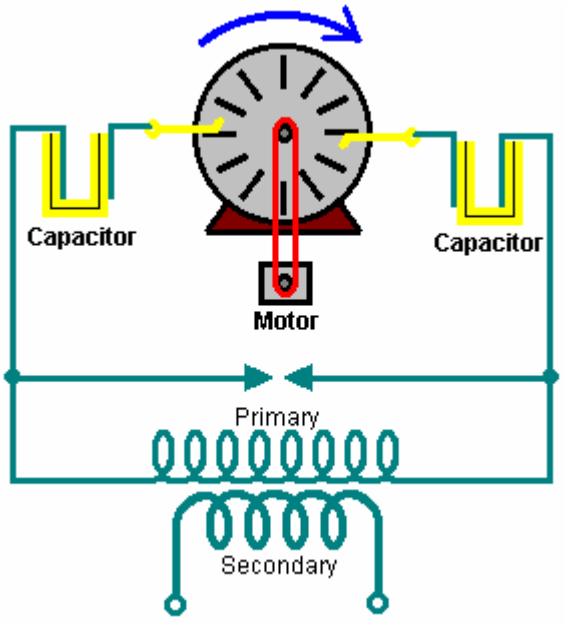
Moray's book "The Sea of Energy in which the Earth Floats" can be downloaded as a free 64-page eBook from http://www.free-energy-info.tuks.nl/P26.pdf

Hermann Plauson's Aerial Systems.

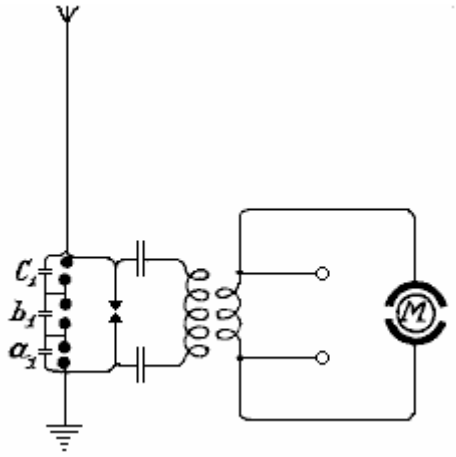
Hermann Plauson was granted US Patent 1,540,998 in June 1925. The patent is similar in style to Tesla's pick-up system and it illustrates the principle with a system which is very much like Paul Baumann's "Testatica" device hidden away in a Swiss religious commune. The patent is very detailed with 37 drawings showing different arrangements, and a more simply worded version is shown in full in the Appendix. In fact, the patent reads more like a tutorial rather than a patent.

A system of this type should most definitely be taken seriously: Hermann considers one of his systems with an output of 100 kilowatts as being a "small" system. However, it needs to be understood that each of Hermann's aerials, unlike those of Thomas Henry Moray, contributed less than one kilowatt each, and when Hermann speaks of a 100 kilowatt installation, he is referring to an array of more than one hundred separate aerials.

Hermann illustrates several different methods of energy capture and several methods of increasing the effectiveness of the captured energy. While an installation to capture a continuous supply of 100+ kilowatts is unrealistic for an individual, requiring amny tall aerials, there is the distinct possibility of making a scaled-down version which is capable of providing serious levels of free power. Reading his patent through carefully is definitely to be recommended.



Herman starts by illustrating how working electricity can be taken from a Wimshurst machine. The Wimshurst output voltage is very high and the current capacity is very low and most people would dismiss it out of hand as being totally inadequate for any kind of practical work. However, Hermann boosts the power level by feeding the output into a step-down transformer which lowers the output voltage to a convenient level and raises the available current in proportion to the reduction in voltage. This is the same technique patented by Nikola Tesla. The apparatus which Herman illustrates is shown here:



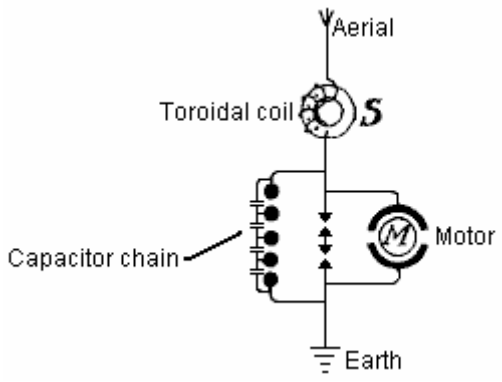
His patent says: “By suitably selecting the ratio between the number of turns in the primary and secondary windings, with regard to a correct application of the coefficients of resonance (capacitance, inductance and resistance) the high voltage of the primary circuit may be suitably converted into a low voltage high current output”. It should be remembered that a spark produces a very sharply rising voltage pulse and that unbalances the local quantum energy field, as described earlier, producing very large energy flows as the local environment returns to its balanced steady-state. The spark, which is produced by relatively low power, is used as a trigger for vastly larger energy flows, which feed the step-down transformer, producing serious current at reasonable voltage, capable of doing useful work, without the requirement for any input power from the user.

You will notice how simple this circuit is. Three capacitors “a1”, “b1” and “c1” in a chain, form a single high-voltage capacitor. The blobs shown connected across these capacitors are emergency discharge spark gaps put there to deal with unusual events like the aerial being hit by a lightning strike. This circuit is very much like the Wimshurst machine circuit which Hermann uses as an illustration of the principle of operation of these kinds of circuits. In this circuit, he shows a special motor marked “M” which is driven by the circuit and he also shows output terminals which can have other equipment connected across them.

When the oscillatory discharges in the primary circuit become weaker or cease entirely, the capacitors are charged again by the static electricity until the accumulated charge again breaks down across the spark gap. All this is repeated as long as electricity is produced by the static machine through the application of mechanical energy to it. Herman states that without the spark gap arrangement across the three capacitors connected

between the aerial and the earth, “it is impossible to collect and render available large quantities of electrical energy.”

In addition to the use of spark gaps in parallel, a second measure of security is also necessary for taking the current from this circuit. This is the introduction of protective electromagnets or choking coils in the aerial circuit as shown by **S** in the diagram below. A single “electromagnet” having a core of the thinnest possible separate laminations is connected with the aerial. In the case of high voltages in the aerial network or at places where there are frequent thunderstorms, several such toroidal-wound coils may be connected in series.



In the case of large units, several such magnets can be employed in parallel or in series parallel. The windings of these electromagnets may be simply connected in series with the aerials. In this case, the windings should be made up from several thin parallel wires, which together, make up the necessary cross-sectional area of wire. The winding may be made of primary and secondary windings in the form of a transformer. The primary winding will then be connected in series with the aerial network, and the secondary winding more or less short-circuited through a regulating resistor or an induction coil. In the latter case it is possible to regulate, to a certain extent, the effect of these choking coils.

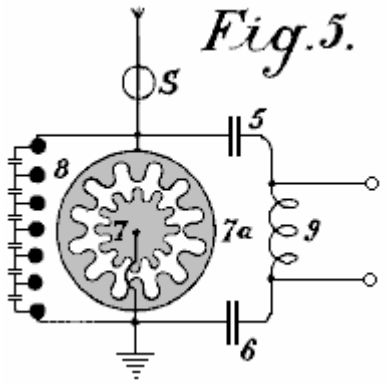


Fig.5 shows an arrangement for producing large currents which can be used directly, without motors, to provide heating and lighting. The main difference here is that the spark gap consists of a star-shaped disc **7** which can rotate on its own axis and is rotated by a motor opposite similarly fitted electrodes **7a**. When separate points of the stars face one another, discharges take place, thus forming an oscillation circuit with capacitors **5** and **6** and inductor **9**. A motor may also be connected directly to the ends of inductor **9**.

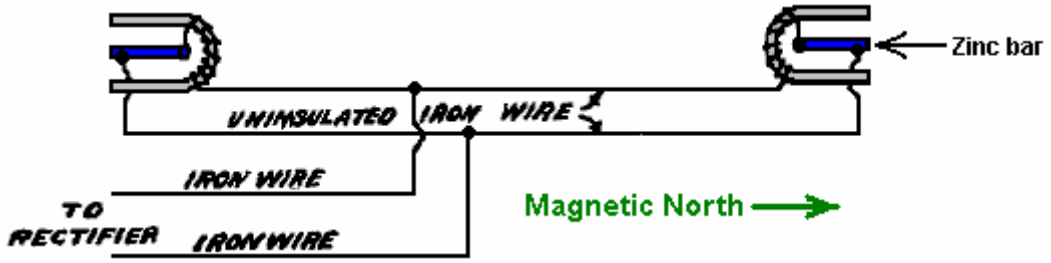
The patent continues by showing many ways to increase the power of the aerial system and many ways of applying the output to practical electrical devices. It contains 37 diagrams, a wealth of practical information, and a copy of it is in the Appendix.

Roy Meyer’s Aerial Device.

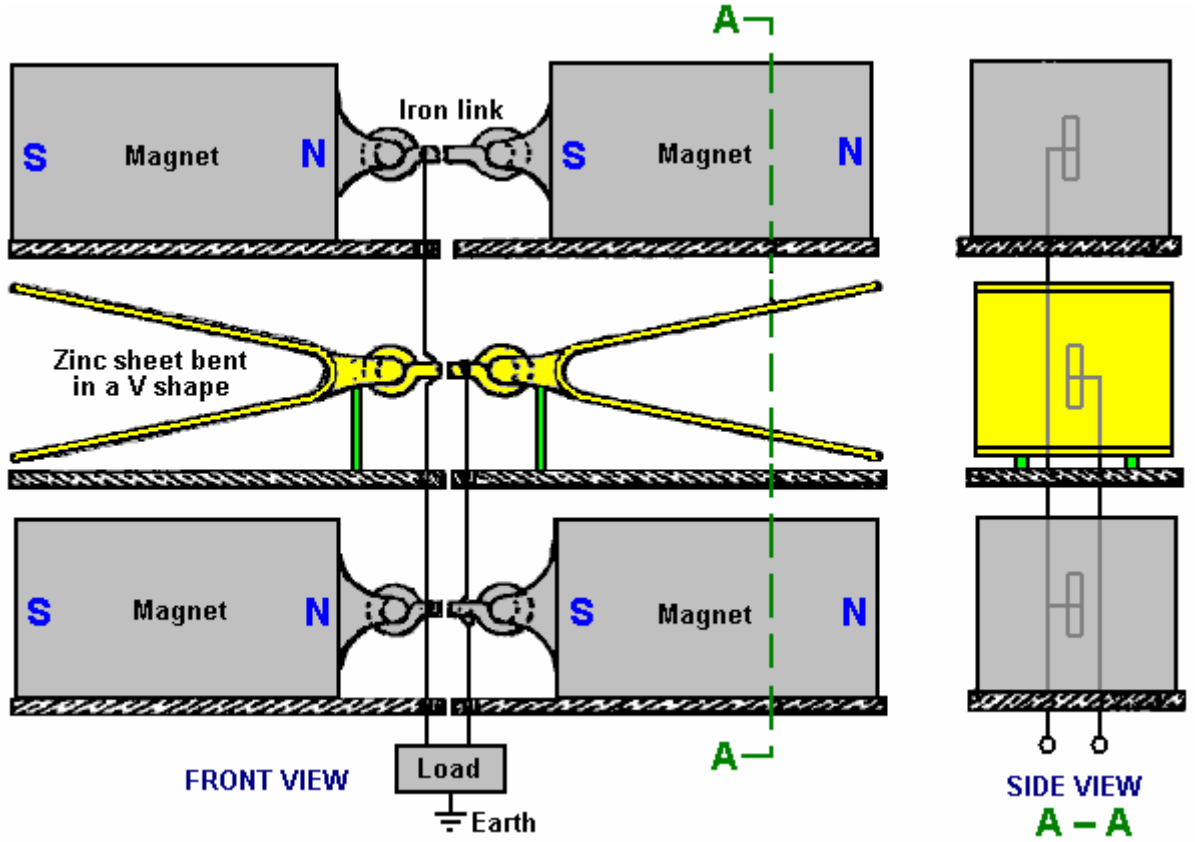
Roy Meyers was granted UK Patent 1913,01098 in January 1914. The patent, which is included in the Appendix, shows an extremely simple device which produces an electrical output without any form of visible input whatsoever. This intriguing device was discovered when testing a very simple form, where two horseshoe magnets were interconnected with soft iron wire and two bars of zinc placed between the legs of the magnets. Roy found that he got an output of 8 volts using just two 4-inch magnets with 1-inch square legs and zinc bars of similar size. The physical orientation of the device is very important. The patent says that current is collected if the open ends of the magnets are pointing in a North - South direction and not if they are positioned in the East -

West direction. However, replication attempts seem to indicate the reverse of this with energy pick-up occurring when the alignment is East-West. Indications are that this is not an easy device to get operating correctly.

The first arrangement is shown in the following diagram:



Roy developed his system further and found that while it works indoors, it does perform better if located outdoors and raised to a height of fifty or sixty feet. However, that is by no means essential, and the output power and voltage can be increased by increasing the number of collector units. Roy developed these to produce the style shown here:



The zinc acts more effectively if installed as sheets bent into a V shape. The magnets and zinc sheets can be stacked vertically and/or horizontally and the greater the number used, the greater the electrical output. A good earth connection is recommended and presumably, the average cold water pipe of any house provides a more than adequate earth connection which is convenient to use, provided the pipework is made of metal.

The Aerial System of Raymond Phillips Senior.

Presents an interesting patent US 4,685,047 of 4th August 1987, entitled “Apparatus for Converting Radio Frequency Energy to Direct Current”. While this patent speaks of radio-frequency energy, I can see no particular reason why that would be the only energy which could be picked up by this circuitry. The patent information is as follows:

Abstract:

This patent describes an apparatus and methods for converting radio frequency energy into direct current for generating electric power. It includes a dipolar antenna for receiving radio frequency energy and a circuit for converting the radio frequency energy to direct current. The circuit has a positive output line connected to one pole of the antenna and a negative output line connected to the other pole of the antenna. A positive transmitting diode is in the positive output line and a negative transmitting diode is in the negative output line. First and second bus lines and a pair of tuned circuits of opposite polarity couple the positive output line and negative line to the bus line with one of the bus lines being connected to ground. Each tuned circuit includes a first bridging line connecting the positive output line to the first and second ground lines and a second bridging line connecting the negative output line to the first and second ground lines. Each bridging line has in it, a diode connected with a polarity which is reversed with respect to the input diode. The bridging lines of each tuned circuit are connected to one another by an inductor and have capacitors placed between the diode and the bus lines. A Direct Current device is connected to the positive line of the circuit.

Background of the Invention:

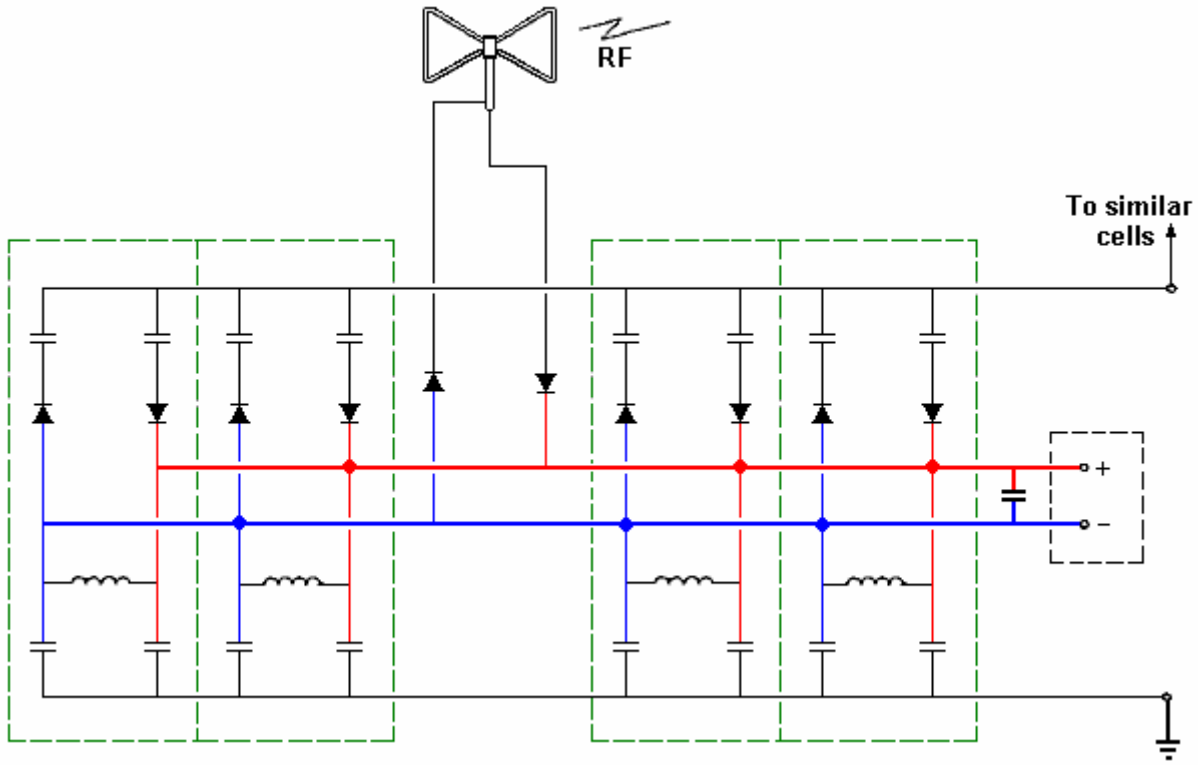
This invention shows an apparatus for converting radio frequency energy to Direct Current of sufficient magnitude to power devices such as battery chargers and electric motors without the use of amplification.

There has long been interest in technology directed to transmitting electrical energy over a distance without using wires. Development of such a technology has enormous potential. This was first recognized by Nikola Tesla who in 1899 constructed a 200 foot Tesla coil rated at 300 kilowatts at 150 kilocycles. Tesla hoped to set up standing waves of electrical energy around the whole surface of the earth, so that receiving antennas set at optimum points could tap the power when needed. Tesla was able to light hundreds of lamps at a distance of about 40 kilometers with his device without using wires. The scheme has generally remained a scientific curiosity but has provided the initial groundwork for current developments wherein attempts are being made to transmit power using microwaves. However, power transmitted by microwaves is envisioned in the form of a beam of very high intensity which is focused from a microwave generator to a receiving antenna. This technology is envisioned as being used for many types of purposes, however, the focused microwave beam is not suitable for many applications because the beam must be directed toward a receiving antenna and cannot be transmitted through most objects, including living objects, without destroying those objects.

This invention relies on converting energy from standing waves which are emitted from radio frequency antennas in the RF range rather than the microwave range. Of particular interest are very low frequencies which are not used in communications and are available for transmitting power. Also of interest are the low frequency waves emitted by the earth due to pulsing of its magnetic field. These low frequency standing "earth" waves can be picked up by receivers tuned to them.

Summary of the Invention:

This invention shows an RF antenna for receiving radio waves. The RF antenna connected to a circuit configured to convert the RF signals to Direct Current. The radio frequency signals received by the antenna are transmitted to two leads, one being rectified to produce positive voltage and the other rectified to produce negative voltage. The positive voltage lead is connected directly to a positive output line and the negative voltage lead is connected directly to a negative output line. The positive output line is connected to a pair of bus lines through a first pair of capacitors, while the negative output line is connected to the pair of bus lines by a second pair of capacitors. Placed between the first bus line and the positive output line is a reverse diode of negative polarity, while placed between the negative output line and first bus line is a reverse diode of positive polarity. The positive and negative output lines are connected to one another through an inductor which is in parallel with the capacitors of the first and second pair connected between the second bus line and the positive and negative output lines.



In one implementation of the invention this circuit is duplicated for each positive and negative output line. In another the circuit is coupled to additional identical circuits in order to increase the direct current output of the arrangement. In another implementation, the antenna used is a dipolar antenna of aluminium wire arranged in a "butterfly" configuration.

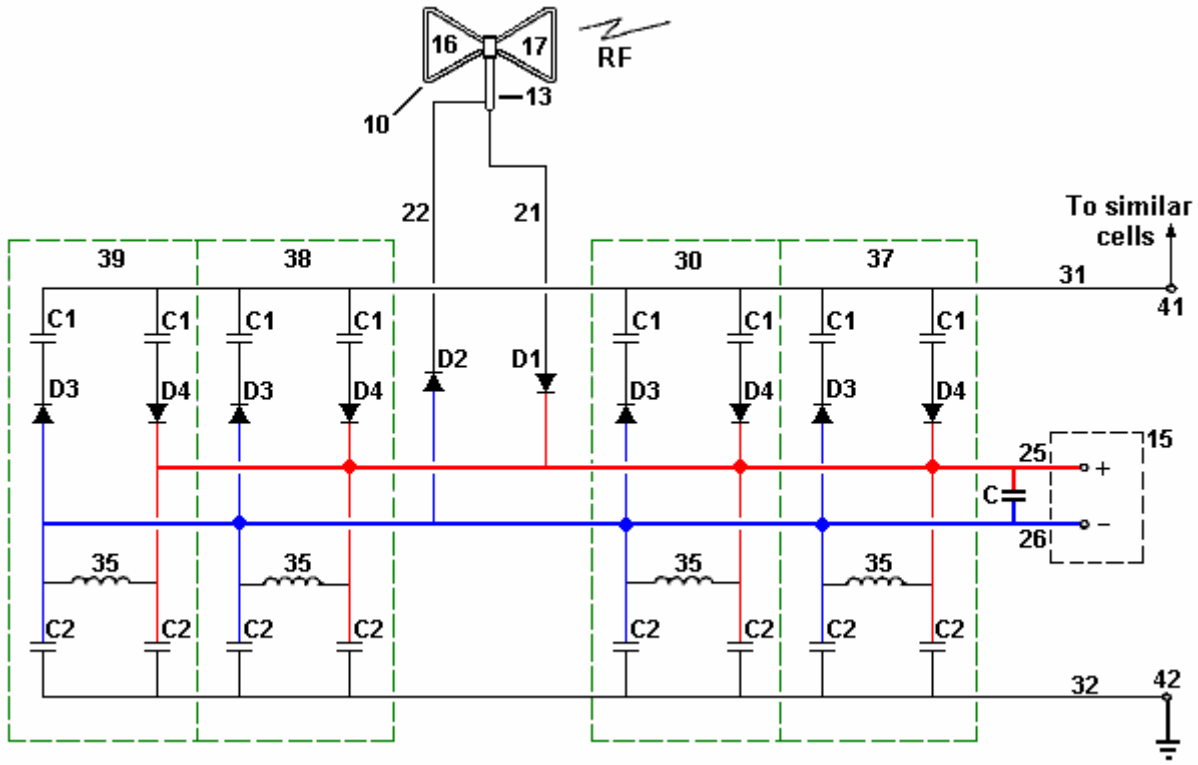
This invention uses these elements to generate direct current of sufficient power to perform tasks such as charging batteries, lighting lamps and powering direct current electric motors without the use of amplifiers.

Brief description of the drawing:

The reference characters used show the same or similar parts in each views, and what is shown is a diagram of a circuit, a driven device and a dipolar antenna which receives radio frequency waves which are then converted to DC current for powering the driven device.

Description of the preferred embodiment:

Referring now to the drawing below, there is shown a dipolar antenna, designated generally by the numeral **10**, which receives radio frequency waves from an RF transmitter.



These waves are passed to the conversion circuit through a coaxial cable **13** and the DC output current of the circuit is used to power an output device **15**, which may be, for example, a battery charger, DC motor, or lighting device. The circuit has no other power inputs and so has no amplifiers for boosting the RF energy.

The source of frequencies which can be converted to Direct Current by the circuit shown may include sources of high frequency (HF), low frequency (LF), very low frequency (VLF) and extremely low frequency (ELF) radio waves as well as seismic vibration of the earth's magnetic fields.

Preferably, the dipolar antenna **10** is formed of two triangular loops of aluminium wire **16** and **17**, one of which is connected to the annular conductor **21** of the coaxial cable **13** and the other of which is connected to the centre conductor **22** of the coaxial cable. The size of the bipolar antenna **10** is dependent on the particular application to which it is put. In one embodiment of the invention, the antenna **10** is approximately **12** inches in width and **18** inches in length. Such an antenna is used to receive five watt energy, such as that generated by a walkie-talkie or citizen-band radio.

The outer conductor of the aerial is connected to positive lead **21** and the centre conductor of the coaxial cable is connected to the negative lead **22** of the circuit. A positive transmitting diode **D1** is placed between the lead **21** and the positive output line **25**. A negative transmitting diode **D2** is placed between the lead **22** and the negative output line **26**.

In order to provide a DC output of sufficient power, a number of inductance-capacitance RF tuned circuits **30**, each forming a positive cell, or a negative cell, are used to connect the positive output line **25** and negative output line **26** to the first and second bus lines **31** and **32**, respectively. Bus line **32** is connected to ground while bus line **31** can be connected to similar circuits. Inductor **35** serves as a radio frequency choke. As is seen of the drawing, the RF tuned circuit cell 30 is repeated a plurality of times. In the specific example shown, the circuit has separate cells **30, 37, 38** and **39**. The cells **30** and **38** are of opposite polarity and balance one another, while the cells **37** and **39** are of opposite polarity and also balance one another. In order for the system to function, a pair of opposite polarised cells must be used. The particular number of cells **30** and the value of the components in it, are determined by the configuration of the dipole antenna **10** and the power and frequency of the RF transmitter.

This radio frequency to direct current conversion circuit may itself be connected to a duplicate circuit via pin **41** so as to provide additional direct current output on lines similar to positive output line **25** and negative output line **26** the output lines may be connected together in order to boost the total output of the system.

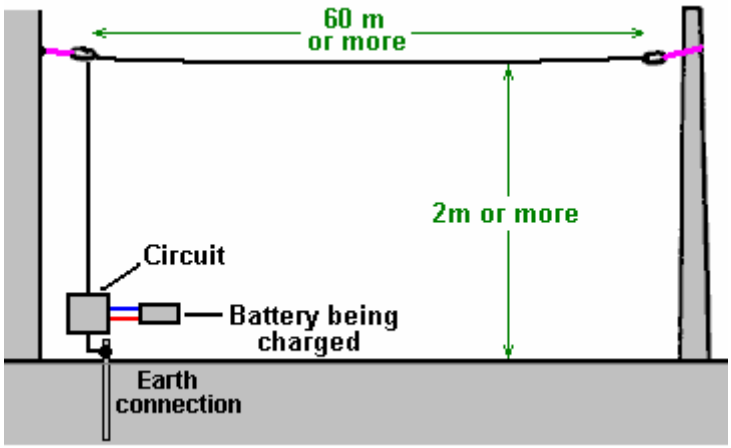
An operative embodiment of the invention uses the following elements:

Diodes: D1, D2, D3 and D4 - Germanium Diodes, type **1N34A**,
Inductors: **35-47 millihenry** R. F. Choke
Capacitors: C1 and C2 - **0.47 Pico Farad**s at 200 volts
Coaxial Cable: 13 - 50 ohms
Dipolar Antenna: aluminium wire triangular loops approximately 12 inches by 18 inches.

From the foregoing description, one skilled in the art can easily ascertain the essential characteristics of this invention, and without departing from the spirit and scope of it, can make various changes and modifications of the invention to adapt it to various uses and conditions.

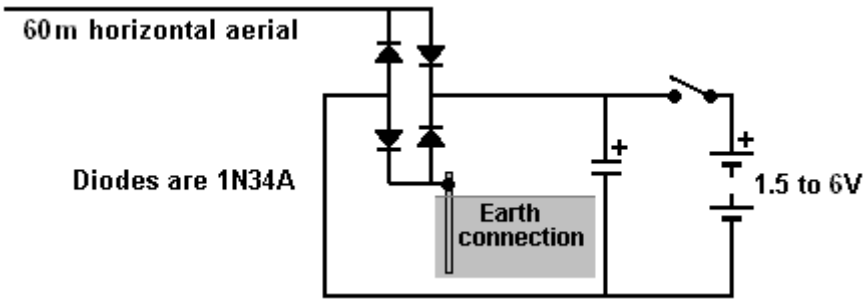
The "Alexkor" Aerial System.

'Alexkor' who provided some of the charging circuits shown in chapter 6, also uses an aerial system for charging batteries in the 1.5V to 6V range. It is a simple system which uses an aerial of the type used by Thomas Henry Moray, that is, the bulk of the aerial is horizontal:

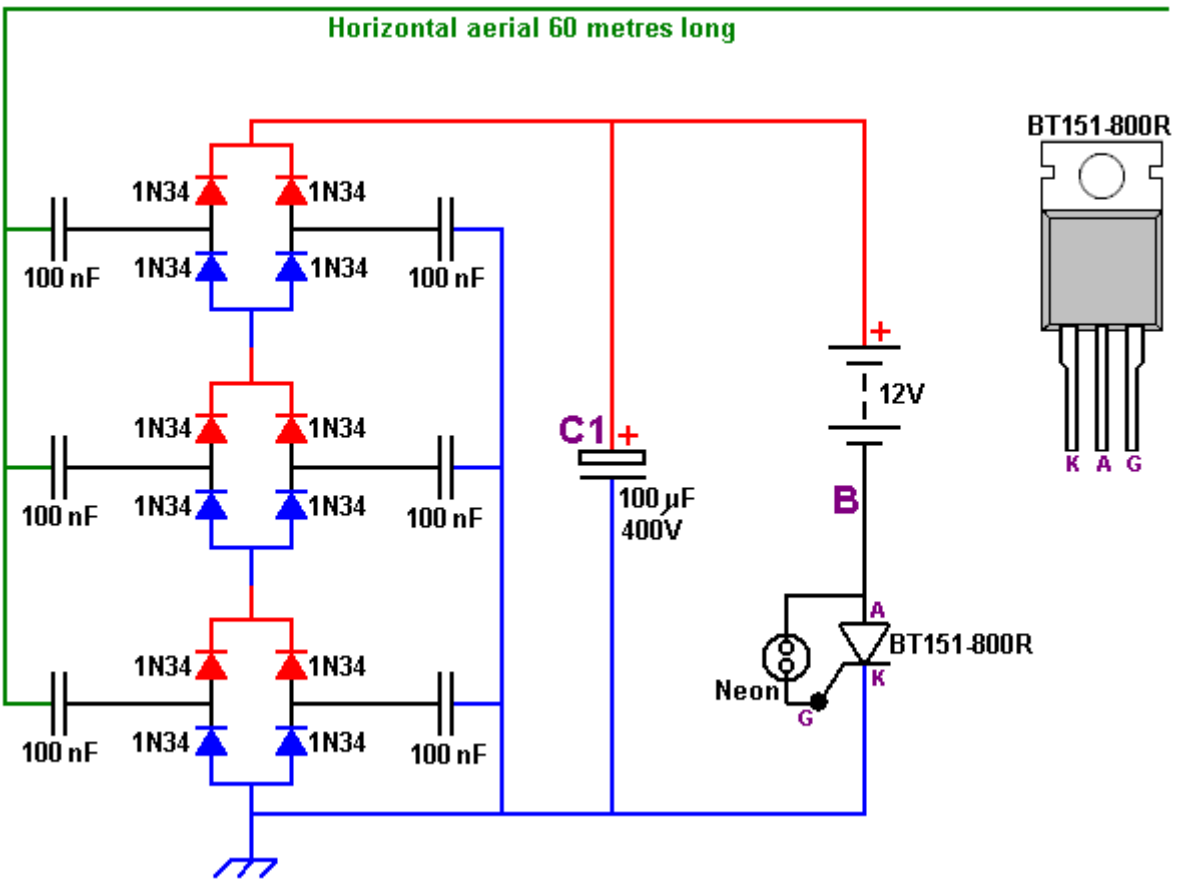


The suggestion is that the aerial is suspended between the eaves of a house and a nearby tree, but I don't know anybody who would be able to do that. The longer the aerial or the greater the number of aerials connected, the greater the charging power available. The aerial wire should be not less than 0.5 mm in diameter and it needs to be insulated from it's supports – plastic cord can be used for that.

The circuit used is:



A more powerful version of the circuit is:



Here, there are three sets of diodes placed between the aerial and the earth. Let me stress again that the earth connection is a thick copper wire connected to a long rod or pipe driven into the ground or a large metal plate buried in dampened ground.

As the circuit operates, the three sets of diodes with their isolation capacitors, build up an increasing voltage on capacitor “C1”. The voltage at point “B” will also increase and be about twelve volts less than the voltage on capacitor “C1”. Eventually, that voltage will rise high enough to cause a discharge through the tiny neon tube and that current pulse flows through the Gate connection “G” of the thyristor, switching it on. Once switched on, the thyristor stays on until capacitor “C1” has been discharged, after which, the thyristor switches off and the process starts all over again. The amount of power in these pulses is considerable and the thyristor gets quite warm when the circuit is running. The BT151-800R thyristor can handle as much as 800 volts and pass 7.5 amps of current continuously.

An important point to note is that the power available from this circuit increases with additional aerials. With two aerials connected, the power is doubled and with three aerials the power is tripled. That is, each additional aerial provides as much power as the first aerial did and there does not appear to be any limit to the number of aerials which can be connected.

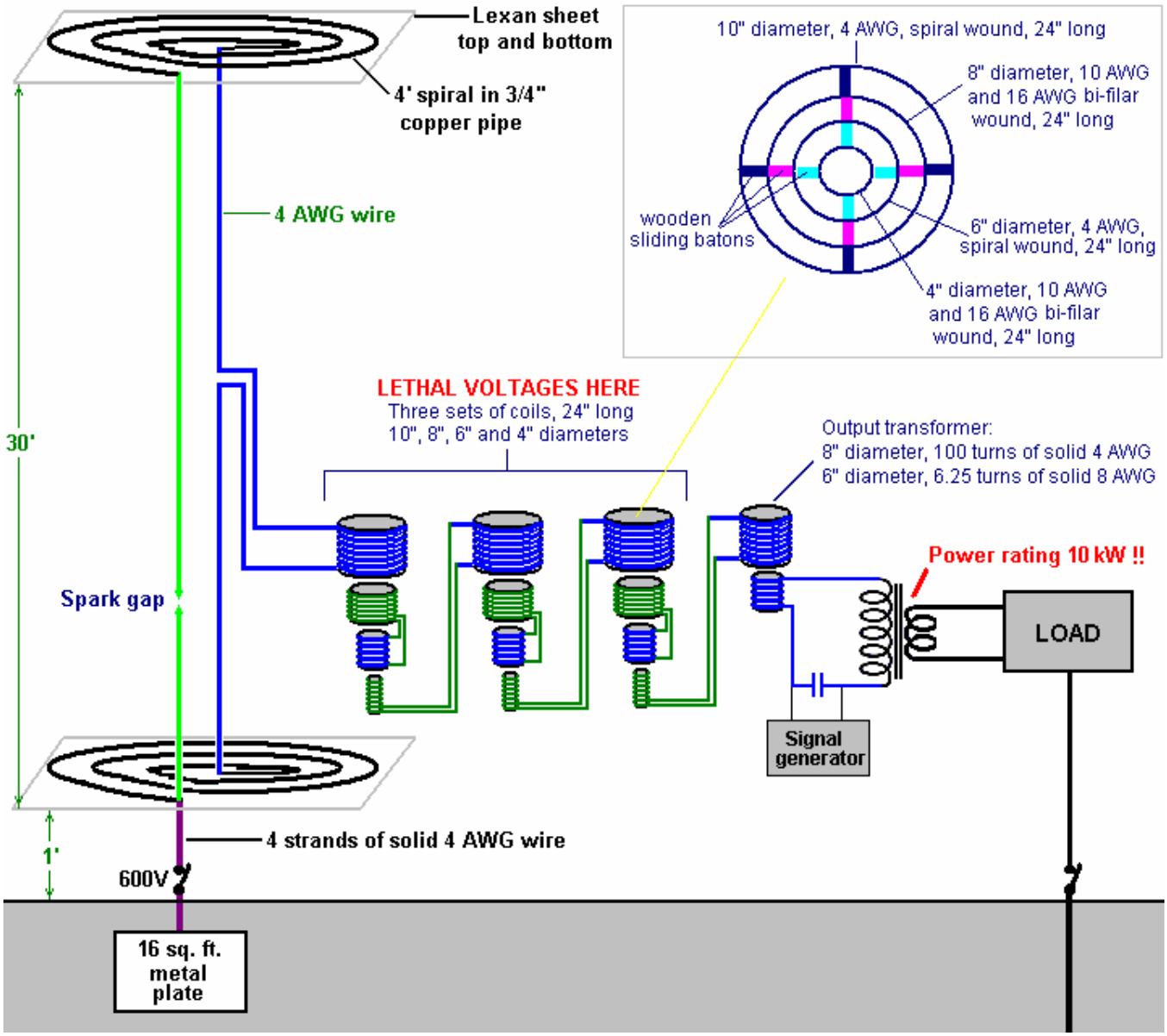
The TREC

There is a powerful aerial system which has been developed by Lawrence Rayburn of Canada. Initially, Lawrence intended to market kits to make it easy for people to replicate his aerial design which was providing him with ten kilowatts of power. However, Lawrence shelved that idea as he decided that the risk to users was just too high. He was also concerned that the aerial/earth combination might attract lightning strikes, causing considerable local damage. So, it is VERY important that you understand that this is a dangerous and potentially LETHAL aerial system which is quite capable of killing you if you are careless. If you are not already experienced in working with high-voltage/high-current devices, then this is NOT something for you to experiment with, and anyone who does fool around with this design does so wholly at his own risk as this material is presented “for

information purposes only" and there is **no** recommendation whatsoever that you should actually construct one of these aerial systems.

The technique used is quite different from the other aerial systems mentioned earlier in this chapter. Here, the objective is to create a tuned, resonant cavity reaching up to the ionosphere where there is a massive amount of excess energy supplied by the sun.

The tuning mechanism consists of two spirals of 3/4-inch (20 mm) copper pipe wound so that they cover a four foot (1220 mm) diameter area. These spirals are covered above and below with a sheet of Lexan plastic. The diameter of the copper pipe is important and smaller diameter should not be used even though it would be much easier to bend into shape. Actually, bending the large-diameter copper pipe is unlikely to be an easy task. Presumably, a pipe-bending machine would be used and a conical spiral produced and then the cone flattened to form the flat spiral. The spiral has even spaces between the turns, at all points along it's length.



The two spirals are mounted, one vertically above the other with thirty feet (9145 mm) between them and the lower spiral being one foot (305 mm) above the ground. This means that this aerial system is only suited to people who can erect a structure of this height without inconveniencing neighbours or contravening local planning regulations.

The tuneable cavity is created by connecting the innermost ends of the two spirals together using 4 AWG wire which has a 5.19 mm diameter copper core. The outer end of the lower spiral is then connected to a large ground plate buried two feet (600 mm) below the surface of the ground, with four separate strands of the 4 AWG wire.

An adjustable spark gap is used. It can be positioned in the vertical wire between the two spirals, or in the middle of an extra 4 AWG wire run vertically between the outer ends of the two spirals. The power take-off is from the middle of the central vertical wire.

The cavity is tuned by sliding coils inside each other. This arrangement not only tunes the cavity but it also steps down the very high aerial voltage and raises the available current at the same time. This is where the danger comes in. The intermediate parts of this step-down tuning arrangement can have voltages of 600 volts at high current, and those voltages can easily kill you, which is why this is NOT a system for beginners and why Lawrence did not go ahead with selling kits for this aerial system.

The aerial output is fed into a final step-down transformer and a 50 Hz (cycles per second) or 60Hz signal from a signal generator is used to modulate the incoming power and allow it to be used with normal mains equipment which may need that frequency of AC. Resistive heaters and lights don't need AC or care about the frequency of any AC which they are fed.

The tuning of the system is very sharp and whatever is being used as the load affects the tuning. The wire used for the tuning array is insulated, single solid-core copper wire. 4 AWG has a core diameter of 5.19 mm, 8 AWG has a core diameter of 3.26 mm, 10 AWG has a core diameter of 2.59 mm and 16 AWG has a core diameter of 1.29 mm. Page 1 of the Appendix shows other characteristics for these wires. These heavy wires are an essential part of this design.

You will notice that two separate earth connections are needed for this aerial system. The second earth connection should be a long copper rod driven deep into the ground and some distance away from the buried plate earth. For the separation of two earth connections, ten metres (eleven yards) is generally considered to be a reasonable distance between them. The main earth is a 4-foot x 4-foot plate buried exactly under the lower spiral and the connection runs to the centre of that plate.

It is suggested that the spark-gap be constructed using copper-clad carbon welding rods, mounted in a nylon housing which allows the gap to be adjusted with a **nylon** screw. Remember that this is not a connection which you can disconnect from the power supply when making connections or adjustments. Also, on one occasion when a 500-watt light bulb was connected across the spark gap, it burned out instantly with a bright flash of light. This system is not a toy so thick rubber gloves and rubber footwear should be used.

The 4 AWG copper wires can be connected to the copper rods forming the spark electrodes, using copper pipe clamps. A tapered nylon rod could be used to adjust the spark gap screw while keeping well away from the copper. The spark gap should start at a one-inch (25 mm) gap and adjusted to no less than a half-inch gap. The spark gap and it's adjusting gear should be enclosed in a weather-proofing container and it has been suggested that it could be and advantage to fill that container with helium gas.

Please remember that coils, such as those in the tuning section of this system, have capacitance as well as inductance. The coil insulation is a dielectric and you standing on the ground form a good earth connection, so please don't imagine that you can't get a serious shock from handling an insulated coil carrying high-voltage high-frequency AC current, and in one implementation after modulation in the mixer at 60 hertz the pre-output transformer was measured at 3496 volts!

If you are not experienced in working with high-voltage circuitry, and still decide to try building and using this circuit, then please put your affairs in order and pre-pay your funeral expenses before you start building. (That statement is not intended to be humorous).

It is said that if the upper spiral is positioned at a height of just four feet (1200 mm) then the output is much lower and suited to charging a battery bank. As far as I am aware, this system has not been replicated and so there has been no practical feedback on building or using it.

The TREC appears to be not unlike the Patent Application US 2008/0191580 from Harold Deyo:

Harmonic Energy Exchange Device

Patent application US 2008/0191580 **Inventor: Harold Stanley Deyo** **14th August 2008**

ABSTRACT

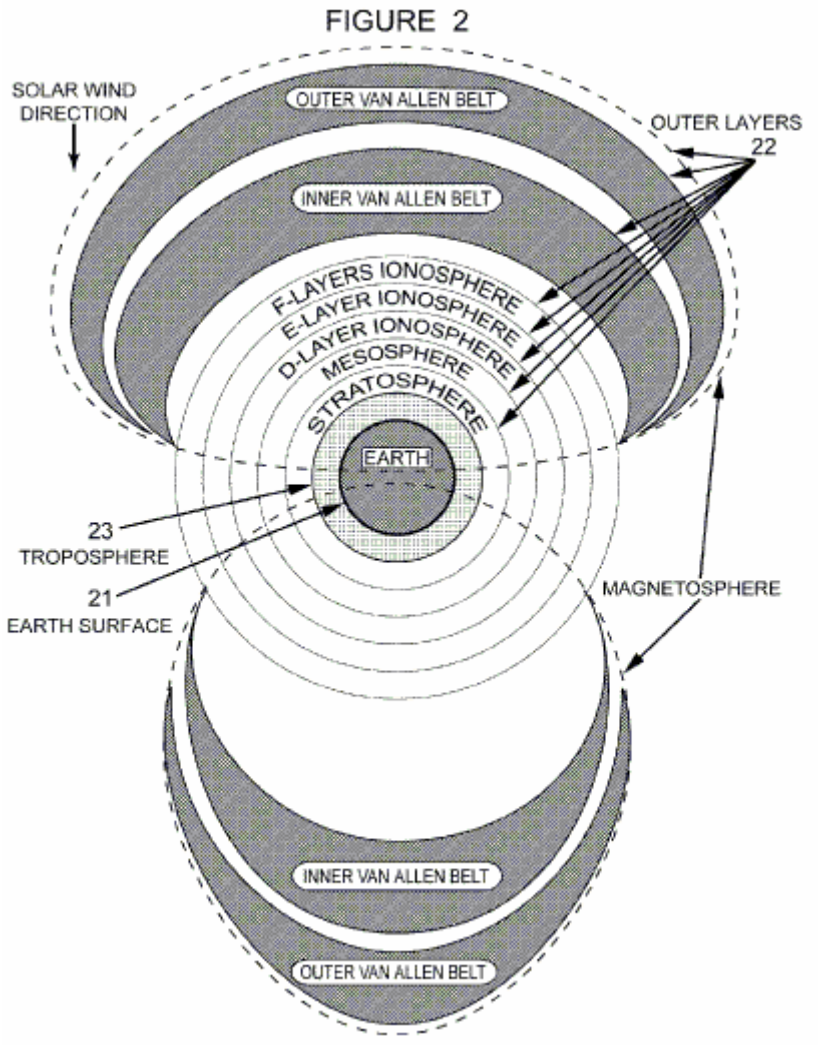
This invention converts inertial impulses into electric currents. Specifically, it converts impulses created by the impacts of high-energy particles from the Sun and other cosmic sources into the Earth's Magnetosphere and the varying D, E, F1 and F2 layers of its Ionosphere to controlled electric currents. This invention presents a new method of utilising energy from the Sun and other sources of high energy articles as a virtually, inexhaustible, alternative-energy source for the world.

BACKGROUND OF THE INVENTION

This invention relates to the conversion of impact energies created by the collision of high-speed cosmic particles and electromagnetic radiations with “Earth's Outer Layers” to produce inertial waves in the dielectric Troposphere which are subsequently converted into electricity by this invention.

The term “Earth's Outer Layers” refers to: Earth's Magnetosphere, Van Allen Belts, Ionosphere, Mesosphere, and Stratosphere as illustrated in Fig.2:

FIGURE 2



DESCRIPTION OF THE PRIOR ART

One day while reading an old scientific magazine I saw a small article on the research of Dr. John Trump of MIT (the basis for an electrostatic generator patented later by Onezime P. Breaux in U.S. Pat. No. 4,127,804). The article talked about a series of experiments which Dr. Trump had performed with a 2-plate, capacitor in a vacuum dielectric. Briefly, he discovered that by varying the distance between the two plates one could produce current flows in one direction or the other between either of the plates and ground.

The Solar Wind strikes “Earth's Outer Layers” constantly as do many other cosmic particles. At any given instant of time, the vector product of these impacts produces either a net pressure increase or decrease in the Troposphere. This creates random waves of pressure in the “Earth's Outer Layers” as one charged “plate” moves in relation to the oppositely-charged Earth's surface as the other “plate”. This variation of pressure in the “Earth's Outer Layers” is equivalent to moving an outer “plate” back and forth toward the Troposphere and the Earth's surface as the inner “plate” - thus giving rise to variations in voltage on both “plates”. In this case the Troposphere (see Fig.2) acts as the dielectric medium instead of the vacuum in the Trump experiments. Furthermore, the

Troposphere, itself, is also compressed and decompressed by these random waves of pressure on the “Earth's Outer Layers”. Thus, I reasoned that if one could create a charged envelope or field of sufficient voltage within the Troposphere, one could convert these random pressure waves in the dielectric Troposphere into current flow on the Earth's surface.

As “Earth's Outer Layers” surround the planet, any impact waves propagate throughout all of them when they are present - even to the night-time side of the planet. Thus, I also reasoned that one could extract power from these impacts anywhere on Earth's surface or in it’s atmosphere whether it be day or night. The pressure waves on the dark side of Earth would be less energetic than those occurring on the day-time side, because the night-time layers of the Ionosphere are compressed so much that the D-Layer of the Ionosphere disappears at night and the F1 and F2-Layers of the Ionosphere combine into one F-Layer. I calculated that the available energy from these impact waves would be significantly less by 30-45% depending upon one's location on the night side of the planet in respect to the terminator.

Many years ago, around 1900, Dr Nikola Tesla built and tested a device to extract energy from the Sun using an apparatus which forms part of my own invention (see U.S. Pat. Nos. 685,957 an 685,958 dated 5 Nov. 1901).

Subsequently, Dr. Thomas Henry Moray, another inventor and student of these Tesla patents made a device which apparently accessed the same source of energy but with a method different than my own. As Dr. Moray was never granted a patent for his device I cannot be certain that his work is considered prior art but I list it here as part of my due diligence. Apparently, he had trouble explaining to the patent reviewer, the source of energy which his device was converting, and thus, was not granted a patent for his impressive work. Dr. Moray's public disclosure of certain aspects of his invention are public knowledge through his published lectures and his book, “The Sea of Energy in Which the Earth Floats' - published in 1931.

Since the beginning of the 20th century, mankind has been looking for new sources of electric power to feed the ever-increasing energy demands of the human population. In the last half of the 19th century, coal, whale oil, human and animal labour, moving water, wind and wood were the main sources of energy. However, in the first quarter of the 20th century, mankind began to use more electricity produced by hydroelectric generators and generators producing electricity by combustion of hydrocarbon fuels. With increasing use of petrol engines and the increasing use of electrical appliances the demand for electricity accelerated the addition of heat and greenhouse gasses into the Earth's ecosystem.

In the last half of the 20th century, many forward-thinking individuals began looking for alternative energy sources to reduce the emissions of greenhouse gasses. All of the alternative energy systems investigated and/or developed in the last 56 years have had significant drawbacks to their usage as the main energy supply for the needs of humanity.

Some of those alternative energy systems were nowhere near as efficient as the existing systems. Some of the alternative energy systems still introduced extra heat into the environment like the existing fossil fuel and nuclear energy systems. And others produced even more greenhouse gasses than the existing energy systems. Some of the new systems used food resources to produce combustible fuels which, consequently, reduced the food resources of the planet. Among the leaders in the new energy sources was, and is, nuclear energy which produces extra heat, toxic wastes and materials which can be used in nuclear weapons. Hydroelectric energy sources do not add heat to the ecosystem but they are not an inexhaustible resource as they depend upon rainfall which is diminishing as global warming increases. Wind energy resources do not add heat to the ecosystem but they are not constant and require expensive conversion and storage equipment.

The ideal energy source for mankind will not add heat to the planet's ecosystem, will not produce by-product green house gasses; will not deplete our food and water resources; will not produce toxic wastes; however, it will be portable, self-replenishing and constantly available in real time anywhere on Earth and in useful quantities.

Our major source of natural energy on Earth is the Sun. Energy from the Sun enters the Earth's ecosphere by particle bombardment, radiated electromagnetic energy, gravitational variations and magnetic processes. Mankind has developed various methods of capturing energy from the Sun already. Silicon-based solar cells convert light in a narrow bandwidth into electron flows. These cells are, at most, only 50% efficient and only work when the Sun is shining on them at an optimum angle. Wind turbines only work when there is wind to power them and are, for the most part, not portable. Devices to convert wave action in the oceans only work when there are waves created by the wind and tidal effects caused by gravitational variations; and they are not portable enough to be used in common transportation devices.

I concluded therefore, that if mankind could find an inexpensive method of obtaining electricity from the Sun in some form which worked both day and night and which was already flowing into the Earth's ecosystem, we might be able to develop a clean and self-replenishing energy source which would add no extra heat to our ecosystem, would not alter our planetary albedo and might well be made in a portable form.

SUMMARY OF THE INVENTION

The Harmonic Energy Exchange Device (or “HEED”) offers an interim solution to Earth's fossil fuel resources. It utilises a previously untapped energy resource which exists throughout the known Universe. By the very nature of the invention, it only interrupts the flow of energy from the Sun and all the other cosmic sources of high-energy particles on their normal journey to the Earth by way of The Outer Layers.

The invention does not add energy to the normal thermodynamic equilibrium of the planet and its associated outer layers. The invention does not produce wastes - toxic or otherwise. It does not produce harmful gasses; and in some embodiments it could even assist in the rebuilding of our thinning ozone shield. Use of this invention will eventually remove mankind's dependence on fossil fuels and create new occupations. It will be cheap enough that homeowners will be able to install their own HEED and sell back energy to the grid until such time as every home has their own HEED and civilization has completely weaned itself from the use of fossil fuels and nuclear energy in its current form.

BRIEF DESCRIPTION OF DRAWINGS

FIGURE 1

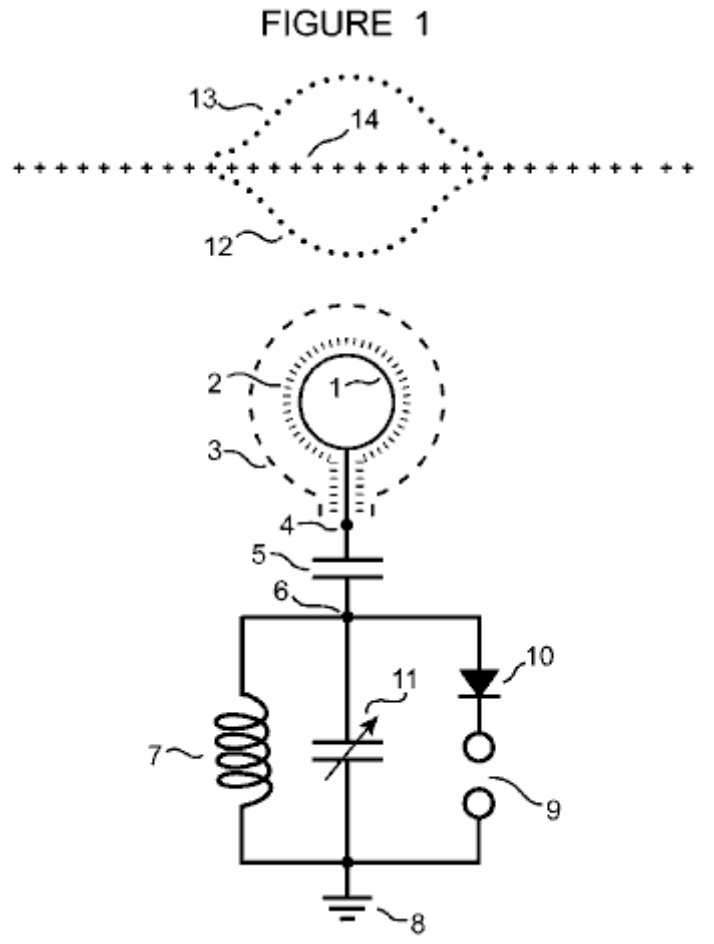


Fig.1: Is the preferred embodiment of the circuit diagram of my invention incorporating references to external stimuli which are used to activate the circuit. It represents a parallel tuned circuit to store the energy.

FIGURE 2

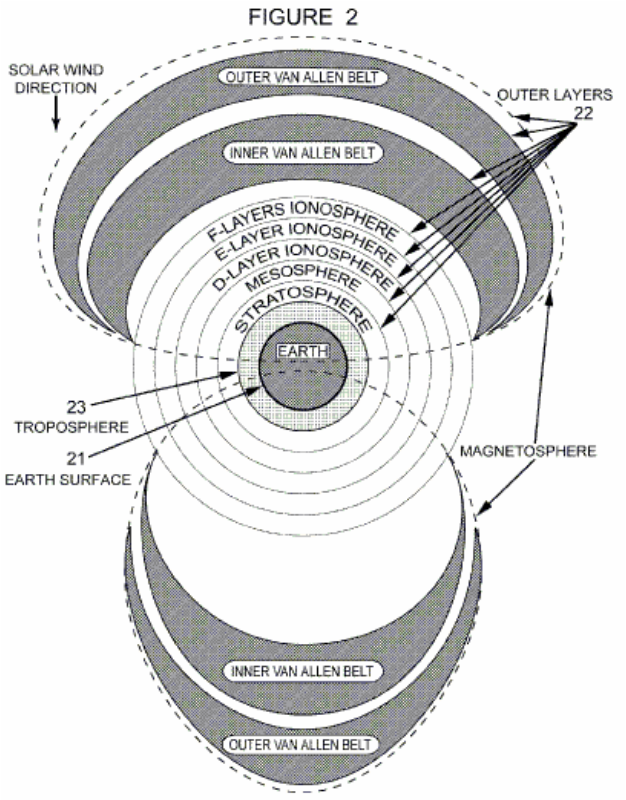


Fig.2: Is an illustration of the Earth and it immediate environment as a global capacitor where The Outer Layers form one plate and the Earth's surface form the other plate and both are separated by the Troposphere as a dielectric medium.

FIGURE 3

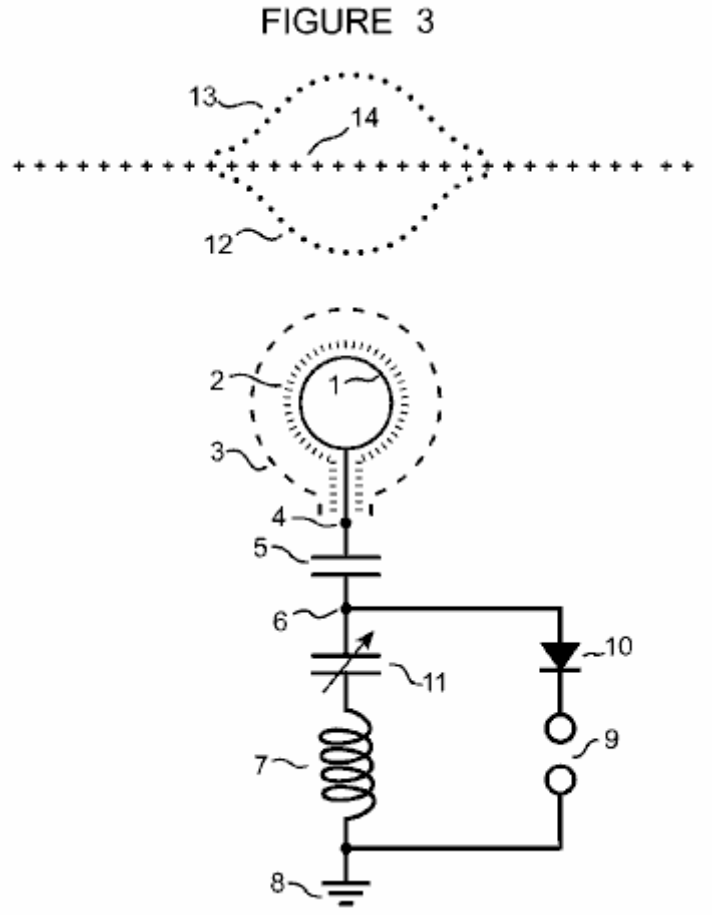
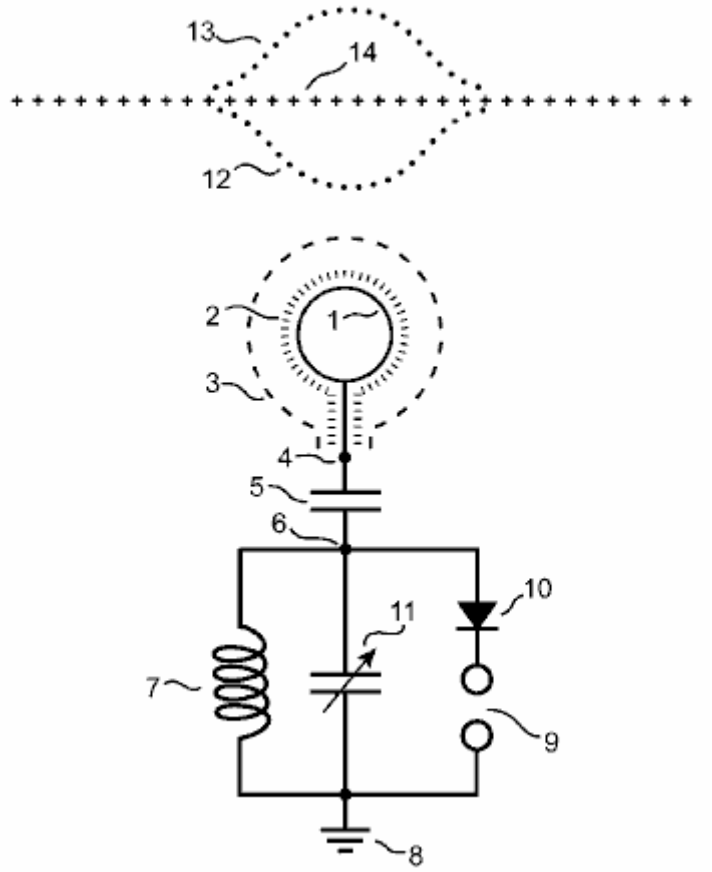


Fig.3: Is an alternative embodiment of the circuit diagram of my invention incorporating references to external stimuli which are used to activate the circuit. It represents the use of a series tuned circuit to store the energy.

DETAILED DESCRIPTION OF A PREFERRED EMBODIMENT OF THE INVENTION

Be it known that I, Harold Stanley Deyo, Jr, a citizen of both The United States of America and The Commonwealth of Australia, residing in the community of Pueblo West in Pueblo County, Colorado have invented a Harmonic Energy Exchange Device which converts dynamic pressures in the ambient media around Earth into controlled electric currents.

FIGURE 1



This invention as represented in **Fig.1** is unique in that it is designed to extract electricity from random pressure waves propagated in the Troposphere **14** by impacts of the solar wind and other cosmic particles with the Earth's "The Outer Layers" **22** of **Fig. 2**.

In **Fig.1**, the Troposphere **14** in its rest state is represented by the line of "+" marks **14**. The convergent or compression state **12** of the random waves in the Troposphere **14** is represented by the lower line of dots while the divergent or decompression state **13** of the random waves in the Troposphere **14** is represented by the upper line of dots. The preferred embodiment of my invention **Fig.1** is a method of coupling a parallel, resonant, electrical circuit to these random pressure waves to extract electricity from them.

As Dr John Trump's research and Onezime P. Breaux's U.S. Pat. No. 4,127,804 show, when one plate of a charged capacitor is moved closer to or farther from the other plate of that capacitor a voltage change appears on both plates of that capacitor. Furthermore, the electric field gradient between those plates changes as the plates are moved in this manner. Thus they have described a method of converting mechanical energy into electrical current by simply varying the plate spacing over time which can be expressed as ds/dt where "ds" is the change in spacing and "dt" is the change in time.

As illustrated in **Fig.2**, The Outer Layers can be viewed as one plate of a capacitor comprised of the Earth's surface itself as one plate **21** and Earth's outer layers as the other plate **22** where the Earth's Troposphere **23** serves as the dielectric medium separating the two plates. There is a charge between these plates which varies extremely rapidly but not with a constant period. At any given instant, the vector product of all the impacts by the cosmic particles with The Outer Layers **22** will create a pressure wave in The Outer Layers **22** which will manifest in the Troposphere **23**. Whatever that vector product is, it will alter the field gradient in the dielectric Troposphere **23**. That alteration will either increase or decrease the effective spacing between "plates" **22** and **23**.

In order to capture and convert the resulting voltage variations produced at or near the Earth's surface **21** of **Fig.2**, this invention creates its own, localised stress field **3** of **Fig.1** established in the Troposphere **23** of **Fig.2** with a charged, conductive surface **1** of **Fig.1** which is encased in high-voltage insulation **2** of **Fig.1** and extending into the Troposphere **23** of **Fig.2**.

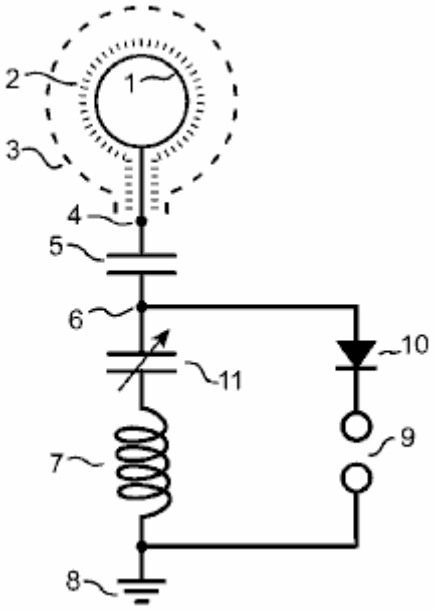
In **Fig.1**, a series of high-voltage, starting pulses is applied across points **4** and **6** on opposite sides of a capacitor **5** to create the localised stress field **3**. As these high-voltage starting pulses are applied, the parallel resonant circuit formed by inductor coil **7** and variable capacitor **11** both referenced to ground **8** is stimulated into resonance within the bandwidth determined by the values of these circuit components. Tuning of this circuit is effected through variable capacitor **11**. The high voltage charge on conductive surface **1** of **Fig.1** is maintained by the parallel tuned circuit formed by inductor coil **7** and variable capacitor **11**.

Then as the random, pressure waves propagate throughout the Troposphere **23** of **Fig.2**, the localised stress field **3** of **Fig.1** is oscillated by the compression wave front **12** and the decompression wave front **13** (both in **Fig.1**) which creates voltage changes on capacitor **5**.

In **Fig.1**, the resulting voltages changes on capacitor **5** will add energy to the parallel resonant circuit formed by inductor coil **7** and variable capacitor **11** which acts as a tank circuit to store the energy which has been passed to it. As energy builds in the parallel circuit the voltage of the circuit rises until a spark discharge occurs across the gap **9** also referenced to ground **8**. The current will only discharge in one direction as determined by diode **10**. The circuit will work without diode **10** but a diode is used here as one method to keep from draining all the energy from the tank circuit when a discharge occurs.

This resonant circuit can be tuned to various bandwidths to maximise the efficiency of the conversion process depending upon the location of the device, the time of day, temperature variations, relative humidity and other variables in the ambient Troposphere **23** of **Fig.2** around the conductive surface **1** of **Fig.1**. A resistive load to extract power from the resonant circuit can be used instead of the spark gap **9** of **Fig.1**.

This preferred embodiment uses a parallel, tuned circuit to access a wide range of frequencies usually to be found in the range of 4.5 to 7 MHz. This range encompasses the major, naturally-occurring, resonant frequencies found in the ionosphere.



A second embodiment of this invention as shown in **Fig.3** replaces the parallel tuned circuit formed by inductor coil **7** and variable capacitor **11** in **Fig.1** placed in parallel to each other and referenced to ground **8**. This second embodiment forms a series-tuned circuit formed by inductor coil **7** and variable capacitor **11** in **Fig.3** placed in series to each other and referenced to ground **8**. It differs from the preferred embodiment only in the placement of variable capacitor **11** so that it is in series with coil **7**.

This embodiment limits the frequency range and, hence, the energy which the system will store when compared to the preferred embodiment. It produces higher voltages across the spark gap **9** than those produced in the preferred embodiment across the spark gap **9** of **Fig.1**.

*****

What is said in this patent application adds some interesting factors to the TREC design of Lawrence Rayburn which pre-dates the Deyo patent by some years. His comment about replacing the spark gap with a resistive load suggests some alternative arrangements which might be tried out with the TREC system.

Also, somewhat along the same lines:

Translated from a Spanish-language original:

Many people think that it is not possible get power from the Earth magnetic field because that field intensity is too low and so the energy levels aren't high enough for any form of useful application. this is not the case. I have built many coils and I get many kilowatts for useful purposes from Earth's magnetic field and here I present the basic concept and some formulas fundamental to all of this.

The influence of the Earth's magnetic field should not be ignored. In times of strong sun activity, the Earth's magnetic field oscillates and in any long power transmission line there are voltage surges and over-voltages which can cause technical troubles, breakdowns and stoppages of the electrical sources. The Faraday law for that induction is given for the following:

V = **2** x **pi** x **f** x **B** x **A**

Where:

B is the Earth's magnetic field,

f the frequency of the fluctuations, and

A the surface area across which the field flows.

For calculation purposes we can approximate the Earth's magnetic field as being 1 Gauss (or 10^{-4} Teslas)

If we consider an ordinary, long transmission line of about 10 Km in length with the power lines spaced 1 metre apart, then we have:

$V = 6.28 \times 10^{-4} \times 10^{4} \times f$ which equals $6.28 \times f$

If the fluctuations are f = 10 Hz, that produces an over-voltage of 63 V. If the fluctuations are 100 Hz then the over-voltage is 630 V, etc.....

So if the Earth's magnetic field is lower in intensity, the effect is considerable in a great surface and volume range.

For energy and power considerations, we can see that the Earth's magnetic field is lower than common permanent magnets, but the volume of space which it covers is very large. The energy in a magnetic field is not just the field intensity alone, but it depends also on the volume across which that field acts.

The energy stored in a magnetic field **B** across a space volume **V** is:

$\mathbf{U} = 1 / (2\mathbf{muo}) \times \mathbf{B}^2 \times \mathbf{V}$ ..... (1)

Where **muo** is the magnetic permittivity of the vacuum.

Common permanent magnets channel energy. If we can use them to get unlimited power like the Bearden's MEG, then the Earth's magnetic field across an air core coil can achieve the same result.

Now we can do a comparison between a permanent magnet and a coil oriented to the Earth's magnetic field to get the same energy levels.

Let's consider a powerful permanent magnet, as used in a MEG, of 5,000 gauss and dimensions of 50 mm x 20 mm x 10 mm. According to the equation (1) above, the energy stored in the permanent magnet will be:

$\mathbf{U} = 1 / (8 \times \mathbf{pi} \times 10^{-7}) \times (0.5)^2 \times (5 \times 10^{-2}) \times (2 \times 10^{-2}) \times (10^{-2})$, so

U = 0.995 Joules - that is to say, roughly 1 Joule of energy.

Devices like the MEG with permanent magnets don't get too many kilowatts, the reason is because that magnetic energy is constant. If we close that magnetic field in a core or magnetic circuit and we pulse that field we get 1

joule of energy at any desired time rate because the permanent magnet stores that energy unlimited and so if we want a power output of 1 KW as the power **P** we calculate:

P = **dU**/dt

For P = 1 KW , we need pulse 1 joule of energy for only 1 millisecond.

In the same way, if we can get power of the same levels from Earth's magnetic field, we must calculate the volume of the air core coil. By using the same equation, we see that

$(0.5)^2 \times (5 \times 10^{-2}) \times (2 \times 10^{-2}) \times (10^{-2}) = (10^{-4})^2 \times \mathbf{V}$

V is the volume of the coil we need for get the same magnetic energy levels, and in this case, V = 250 m^3

That is to say, a coil of 6.3 m diameter and 6.3 m length, placed parallel to the Earth's magnetic field, can store the same energy as that little 5,000 gauss permanent magnet which we considered for a MEG device.

But it is not necessary build a huge coil, we can use a smaller coil. The enclosed magnetic energy will be lower, but as P = dU/dt we must raise the frequency of the pulses to obtain the same power level coming from a bigger coil. For example, an air core coil of 1 meter diameter and 1 meter length according to equation (1), stores an energy of:

$\mathbf{U} = 1 / (8 \times \mathbf{pi} \times 10^{-7}) \times (10^{-4})^2 \times \mathbf{pi} \times 1 / 4 \times 1 = 0.003$ Joules

If we pulse that energy level at 330 kHz, then we will get 1 kW, and at 660 kHz, 2 kW, etc., thus a higher frequency yields more power.

Then the question becomes, how can we pulse the constant magnetic field inside the coil? The answer is simple: by using an external source, we can cancel the Earth's magnetic field inside the coil. There must be power and energy amplification with respect to the external input source. To realise that power amplification, we must do the following:

Let the magnetic field variation inside the air coil be given by:

B(t) = **Bo** + **Bf** x Sin(**w** x t)

Where

Bo is the constant of Earth's magnetic field,
Bf is the magnetic field in the coil created by the external power source, and
w is the angular frequency of the external source.

Replacing B(t) from equation (1) we get the energy variation with time, U(t), and then we can calculate the power as P = dU/dt resulting in:

P(t) = **Bf** x **w** x **V** x (**Bo** + **Bf** x Sin(**w** x t) x Cos(**w** x t)) / **muo** ..... (2)

Remember that V is the volume inside the coil.

We see here that the output power depends on Bo, the Earth's magnetic field, just as in the case of Bearden's MEG it depends on the magnetic field intensity of the permanent magnet in the circuit.

So we can now calculate a COP value with Bo and without Bo, or Bo = 0

Calculating the RMS power for both cases (not reproduced here because it corresponds to a case of basic differential calculus) and using the ratio, the result for the COP is:

$COP = (1 + (2 \times \mathbf{Bo} / \mathbf{Bf})^2)^{0.5}$

We see then power amplification, and of course if Bo=0 and not a permanent magnetic field, the maximum COP is 1, input and output powers are equal. In the case of Bearden's MEG, the condition is Bo = Bf for not degaussing the permanent magnet and in that case we have a COP = square root of (5), which is a value between 2 and 3 which corresponds to the practical results for this classic calculation.

But for the Earth's magnetic field, we can get higher values because we can never cause a demagnetisation of the Earth's magnetic field.

How many turns on the coil, pulsing frequency, coil diameter, coil length, etc., do we need?

The input power to cancel the Earth's magnetic field inside the coil, or the condition Bf = Bo , we calculate using the RMS component of equation (2) above, setting Bo = 0, so we have:

P = 0.05 x **f** x **A** x **L** , (where P is in watts).

A is the section of the coil in m^2,
f the frequency in Hertz,
L the coil former length in metres

To cancel the Earth's magnetic field inside the coil, the condition based on Ampere's law is:

N x **i** / **L** = 100

Where **N** is the number of turns on the coil and
i is the current in amperes.

Finally, the induced voltage in the coil will be:

V = 2 x **pi** x **f** x **N** x **Bo** x **A**

So, with the following values, we have for example:

N = 1000 turns,
f = 100 kHz,
Coil diameter = 1 m,
Coil length **L** = 1 m,
i = 100 mA or 0.1 A,
Bo = Earth's magnetic field

This gives V = approximately 4,000 Volts.

The input power for this example is about 4 kW, the output power for the COP factor will be a maximum of 12 kW. In closed-loop operation we provide the input 4 kW from the generated output and we get a self-powered generator producing 8 kW of output power. My practical results match these calculations.

How is a practical device built?

Using a large plastic pipe, we wind the primary input coil for use with the external source. For extracting the output power, we use another pipe placed inside the first pipe, and of approximately the same diameter and length with an adequate number of turns and wire size to step down the voltage, to, for example, 110 VAC. For the high voltage oscillator of the input source I use a resonant LC circuit taken from a ferrite power transformer.

It is very interesting to see this device work, and on days of geomagnetic storms, power levels go too high and I must use protective devices to prevent damage to the system and the loads connected to it.

Any questions to the following email: enertec2200@yahoo.es (which may not still be operational as this information is quite old).

Electrostatic Generators

Electrostatic generators are not at all unlike aerial systems in that a conversion to ordinary current is needed. The Swiss device designed by Paul Baumann (chapter 13) which has kilowatts of conventional electricity output power and which is self-powered demonstrates that very useful systems can utilise electrostatic charge. Paul's design is built around a Wimshurst machine which is self-revolving due to the sloping charge-collection strips which form an electrostatic motor as well as collecting the charge which is passed on to the remainder of the circuit.

Quite a number of electrostatic generators have been designed and used. They generally use a small, low-power motor to repeatedly alter the capacitance between a series of metal plates and so, create alternating current flow. Some of these designs are really clever in the way in which they operate. For example, the following patent from Charles Goldie has an input voltage source which is generated by the device itself and the voltage used can be adjusted by the user, while the device provides two separate DC outputs in series, which can be used as a single higher voltage if desired.

The Le May patent shown below has useful techniques which raise the output power to more than five times what it would have been without the use of those techniques. The general impression of electrostatic generators is that they are cumbersome and generally ineffective. That may be the case for some of them, but it is certainly not the case for all of them. For example, the electrostatic generator design of William Hyde, shown in chapter 11, has a COP=10 performance where the electrical output is ten times greater than the mechanical power needed to make it operate, and I'm told that the generator output can be 10 kilowatts which is a more than respectable output.

A particularly clever design comes from Onezime Breaux (US 4,127,804) where a high-voltage charge is generated briefly, and then that same charge is shuttled backwards and forwards through the output load, without that charge ever being 'used up' (to put it in layman's terms). As no system is ever 100% perfect, every ten minutes or so, the high-voltage charge is very briefly reinforced. A system of that type has to be capable of being a self-sustained generator. In a minimal prototype which was not perfectly constructed by any means, the power needed to spin the rotor was 0.8 watts, while the electrical output was 20 watts. This shows clearly, that electrostatic generators can be seriously useful, especially if built in the larger sizes.

The Charles Goldie Electrostatic Generator

US Patent 3,013,201 12th December 1961 Inventor: Charles H. Goldie

SELF-EXCITED VARIABLE CAPACITANCE ELECTROSTATIC GENERATOR

This invention relates to electrostatic generators wherein the charge transfer mechanism makes use of capacitive effects between charged bodies and electrodes upon which charge is to be induced, and in particular, to the self-excitation of such generators.

The invention may best be understood from the following detailed description, having reference to the accompanying drawings in which:

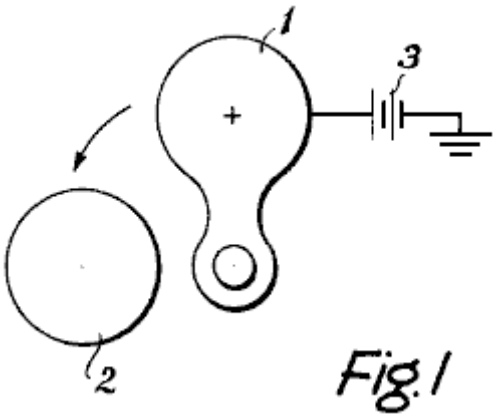


Fig. 1

Fig.1 is a diagram illustrating one form of a variable-capacitance electrostatic generator;

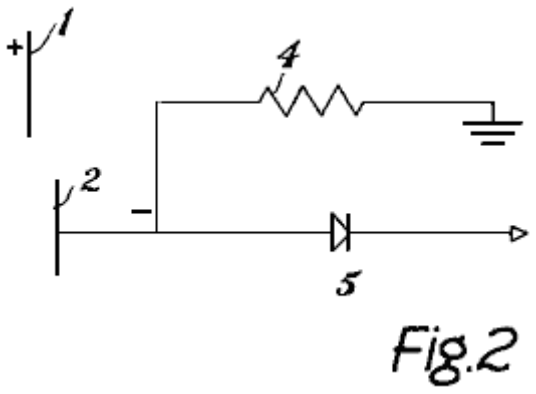


Fig.2

Fig.2 is another diagram further illustrating the arrangement of **Fig.1**;

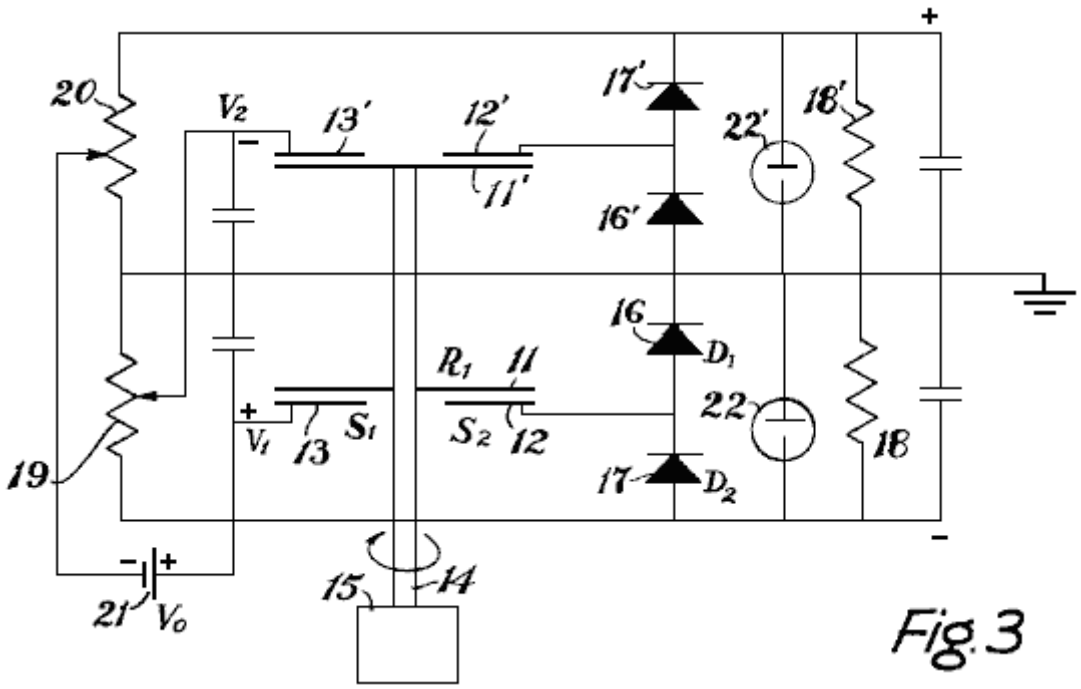


Fig.3

Fig.3 is a diagram illustrating a modification of the apparatus of **Fig.1** and embodying the invention; and

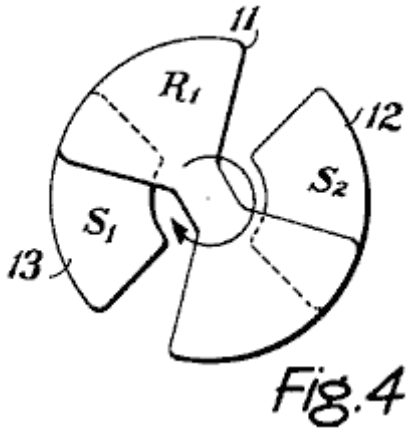


Fig.4

Fig.4 is a diagram illustrating a portion of the apparatus of **Fig.3**.

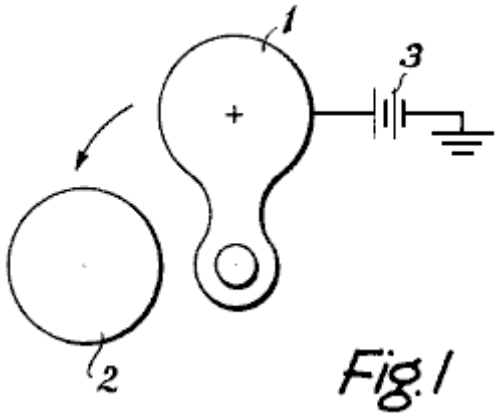


Fig.1

Looking at the drawings in more detail, **Fig.1** shows a rotor **1** and a stator **2**. The rotor **1** has a positive charge on it. A power supply **3** maintains the charge on rotor **1**.

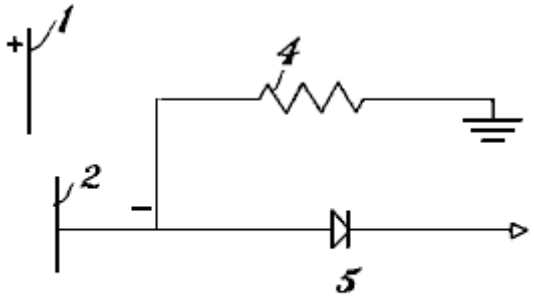


Fig.2

In **Fig.2** it can be seen that the stator **2** is connected to ground by a resistive path **4** which is connected in parallel with at least one rectifier **5** so that current can only flow in one direction between the stator **2** and ground.

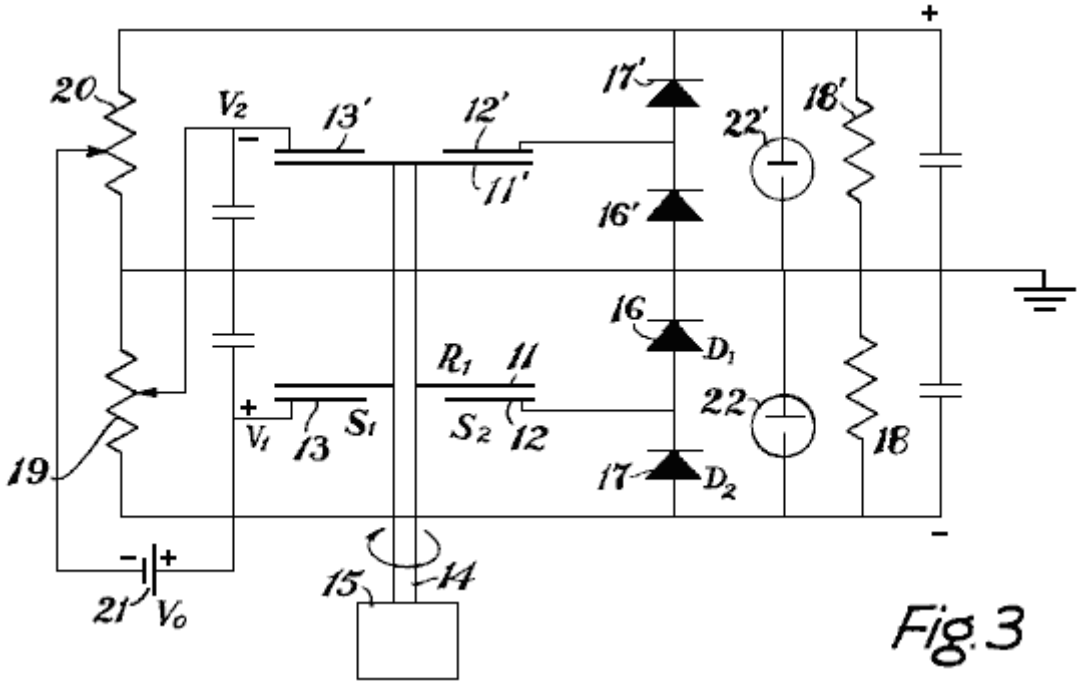


Fig.3

In the device shown in **Fig.3** and **Fig.4**, the rotor serves as a capacitive link between a stationary charging electrode and a stator. A rotor **11** periodically passes close to a stator **12**. However, instead of carrying it's own charge as in the device shown in **Fig.1** and **Fig.2**, the necessary inducing charge on rotor **11** is itself induced on

rotor 11 by the charging electrode **13**. So, when one end of rotor **11** is near the stator element **12**, and the opposite end of rotor **11** is close to the charging electrode **13**, a negative charge is induced on that portion of rotor **11** which is near the charging electrode **13** with the result that positive charge appears at the opposite end of rotor **11** since rotor **11** must itself be electrically neutral.

The invention will now be described with reference to the apparatus shown in **Fig.3** and **Fig.4**, but it will be clear from the comments above, that the invention could equally well be used with a device of the type shown in **Fig.1** and **Fig.2**. The principle advantage of the device shown in **Fig.3** and **Fig.4** is the fact that it eliminates the necessity of a brush on the rotor to supply it with charge, as the charge on the rotor is induced capacitively.

Referring back briefly to **Fig.1** and **Fig.2**, it will be recalled that the invention is concerned with the construction of a self-excited generator with no external energy other than shaft power supplied. That is to say, the objective is a completely self-charging generator. Referring to **Fig.1** and **Fig.2**, it will be seen that direct feedback of the output to the induction plate is ruled out by the polarity switch inherent in this type of machine. The difficulty is avoided by the two-section unit shown in **Fig.3** and **Fig.4**. Here, a fraction of the output voltage of each section is fed to the induction plate of the other.

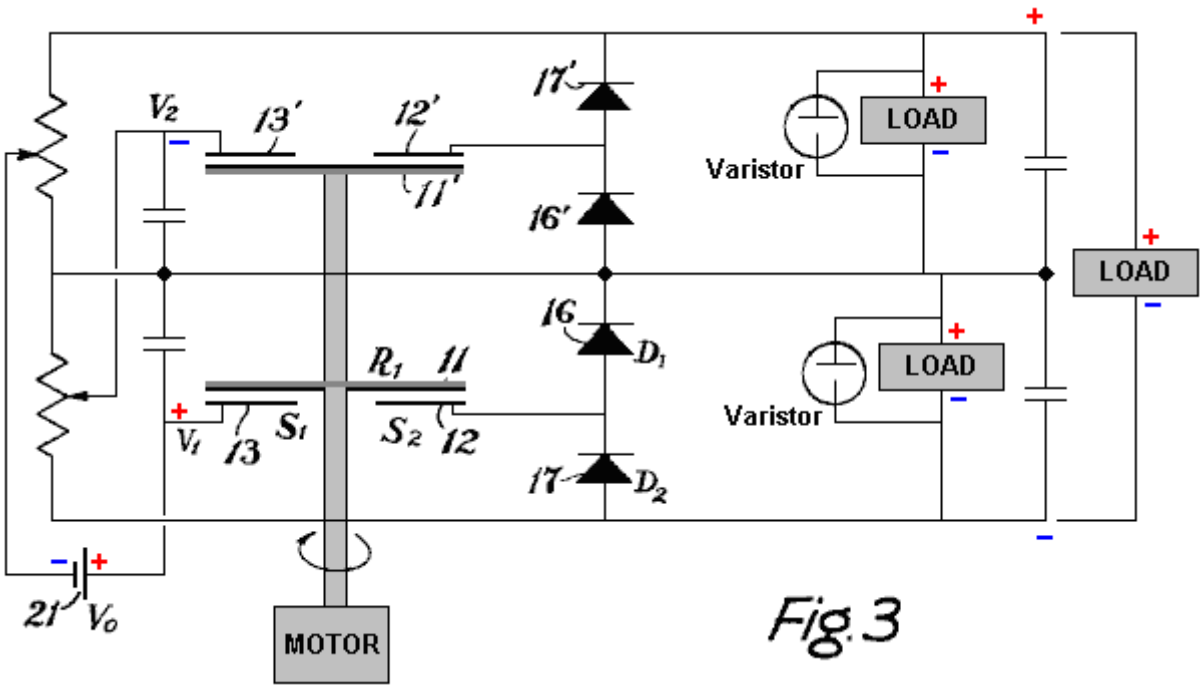


Fig. 3

The action of the circuit is best understood by looking first at the lower section only. Rotor **11** is an electrically isolated, flat fan-like rotor mounted on an insulating shaft **14** which is driven by a suitable motor **15**. As rotor **11** rotates, it periodically covers stator **2** and the charging electrode **13**, both of which are insulated and sector-shaped. Charging electrode **13** is maintained at a positive DC voltage with respect to ground. As rotor **11** rotates into a position where it covers both the charging electrode **13** and the stator **12**, the capacitance between the charging electrode **13** and stator **12** increases and negative charge is induced from the ground through diode **16** and on to stator **12**. As rotor **11** rotates away from stator **12** and charging electrode **13**, the capacitive coupling decreases and the voltage on stator **12** rises, negative with respect to ground. Diode **16** blocks any flow of negative charge from stator **12** and as the negative potential of stator **12** rises, current flows through a second diode **17** to the load **18**. A fraction "**V2**" of this negative load voltage is applied to the charging electrode or induction plate **13'** of the second or upper section by means of the variable resistor **19**. The charging action of this second section is identical to that in the first section except for reversed polarities, and the output of the second section is correspondingly fed back to induction plate **13** of the first section via a second variable resistor **20**. Variable resistors **19** and **20**, control the output voltage and power by adjusting the feedback ratios.

The circuit is regenerative and so it is only necessary to supply a small reference voltage to start in the right direction. The use of stator and rotor materials with suitable contact potentials is one solution. Another solution is using a small battery **21** in one induction line. Suitable voltage-limiting devices **22** would be added either across induction plate **13**, or across the load **18** to prevent excessive voltage build-up.

A number of features in this circuit are of interest. The usual slip rings and brushes are eliminated by the use of stator **12** and charging electrode **13** and an isolated rotor **11** rather than a single stator **2** and a rotor **1** maintained at constant induction potential as shown in **Fig.1** and **Fig.2**.

In spite of the fact that the second section was added for the purpose of self-excitation, both output voltages are added together across the load **18**, **18'**. The arrangement shown in **Fig.3** is merely for explanation of the self-excitation principle. The ground shown, for example, is not required and for higher voltages, the units may be cascaded with a common insulating drive shaft **14**. For a given total voltage output, the number of rotor/stator sets is the same for this self-charge generator as for a generator requiring auxiliary charging supplies.

Two poles are shown for simplicity. In an actual machine a greater number would be more practical. The system applies equally well to other geometrical arrangements, such as the concentric or segmented drum design mentioned in US Patent application 829,823.

The Le May Electrostatic Generator

US Patent 3,094,653 18 June 1963 Inventors: D. B. Le May et al.

ELECTROSTATIC GENERATOR

The present invention relates to an energy conversion machine of the type in which a variable capacitor is cyclically driven while being charged and discharged in synchronism with the cyclic variations in capacitance.

There are several different types of electrostatic machines which, for example, function to generate electrical energy from kinetic energy or kinetic energy from electrical energy. One broad classification for these machines is based on whether conductive or non-conductive apparatus is used to transport electrical charge within the system. The present invention relates to the class of machines in which electrical charge is transported through conductive apparatus. More specifically, the present invention relates to electrostatic systems in which a form of variable capacitor is mechanically driven to provide a cyclically-varying capacitance which is charged during intervals of high capacitance and discharged at intervals of low capacitance, to provide electrical energy.

In general, previously-proposed variable-capacitor electrostatic generators (sometimes called electrostatic induction generators) have been capable of very high efficiencies and could be embodied in relatively light apparatus. These considerations result primarily from the fact that electrostatic fields may be sustained in air, vacuum, or other very light medium, while electromagnetic fields necessitate the use of heavy ferrous material. Further, the electrostatic fields are caused by charges carried on surfaces, rather than by currents in conductors, so that the relatively heavy copper requirements of electromagnetic machines may be avoided. It has been found that electrostatic machines operating in a vacuum, do not have the iron loss, copper loss or windage loss of electromagnetic machines, thereby achieving their extremely high efficiency.

However, other considerations have prevented these machines from coming into widespread use. Specifically, prior electrostatic machines of this class have normally been of large size relative to their power capacity. Furthermore, practical use of prior machines has normally required extremely high voltage operation. Both of these drawbacks result from the nature of the electrostatic field, it's requirement of high potential gradients for field strengths comparable to electromagnetic fields, and the limitations on potential gradients imposed by the breakdown of the dielectric medium. Therefore, considerable insulation and safety equipment has been required in prior electrostatic systems, and operation at moderate voltages has not been practical. Still further, certain forms of electrostatic induction generators have necessitated rather complex switching systems to charge and discharge the capacitors; these systems have not been commercially feasible.

In general, the present invention comprises an efficient electrostatic induction system capable of economical operation which may be embodied in a unit that is small in size relative to power capability. The system includes two sets of rod-like members, one set of which is revolved relative to the other, so that certain of the relatively-moveable rods provide a variable capacitance. The other rods in the structure are then used to shape the electrostatic field in a manner which produces a greater difference between the minimum and maximum capacitance of the apparatus, resulting in increased power conversion capability for the size of the unit. The variable capacitance structure is connected to an electrical system which charges and discharges the capacitance in a cyclically-varying manner. Reactive elements are then coupled to the system to cause voltage phase shifts in the system relative to the variable capacitance, to increase the charge applied to, and taken from, the capacitance structure during the operating cycle, thereby further increasing the power-conversion capabilities of the system.

Another object of the present invention is to provide an improved electrostatic induction system.

Another object of the present invention is to provide an electrostatic induction generator or motor which has a higher power conversion capability relative to it's physical size.

Still another object of the present invention is to provide an electrostatic energy conversion system which may be embodied economically in a relatively small size apparatus.

A further object of the present invention is to provide an electrostatic induction motor or generator in which energy is more effectively converted to another form by improving the shape of the electrostatic fields in the system.

Still a further object of the present invention is to provide an electrostatic induction generator or motor in which energy is effectively transformed to another form by the provision of reactive elements to tune the capacitive elements that are inherent in the system.

One further object of the present invention is to provide an improved system for efficiently converting energy from one form to another, wherein an electrostatic field is used to inter-couple the different forms of energy.

These, and other objects of the present invention will become apparent from a consideration of the following, taken in conjunction with the drawings, wherein:

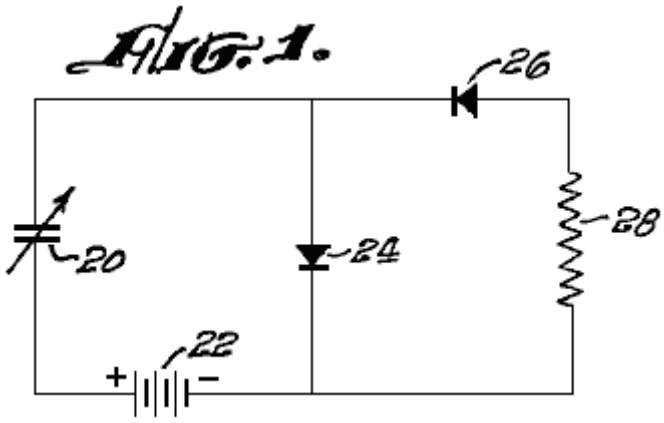


Fig.1 is a schematic diagram illustrating the basic operation of a preliminary system to which the present invention may be applied;

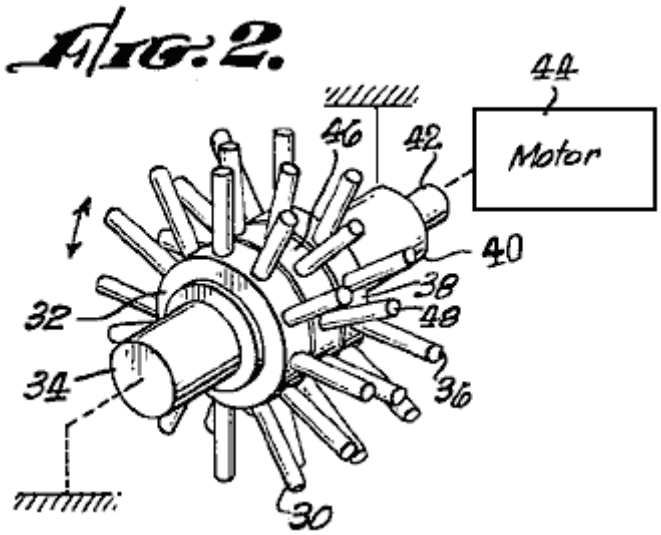


Fig.2 is a perspective and diagrammatic representation of one form of variable capacitor which may be incorporated in an apparatus of the present invention;

FIG. 3.

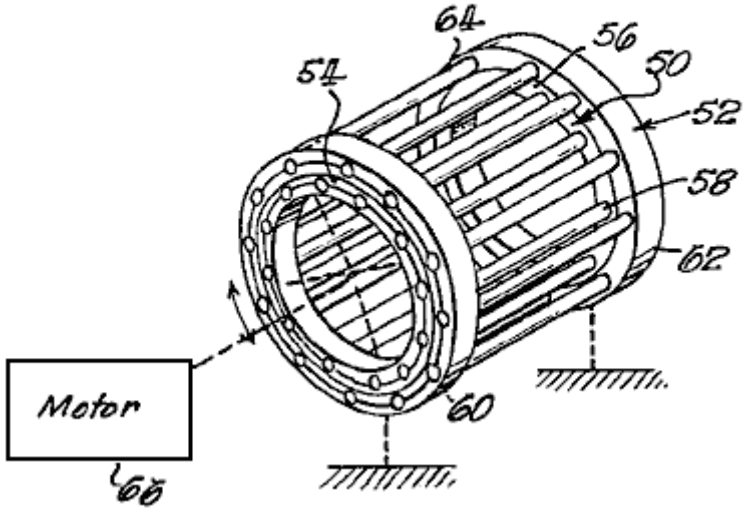


Fig.3 is a perspective view of another form of variable capacitor which may be incorporated in an apparatus of the present invention;

FIG. 4a.

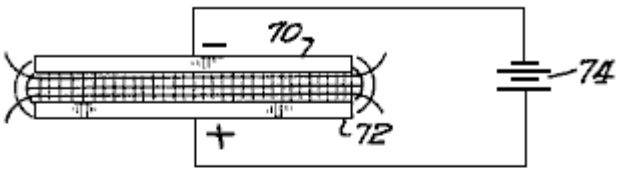


FIG. 4b.

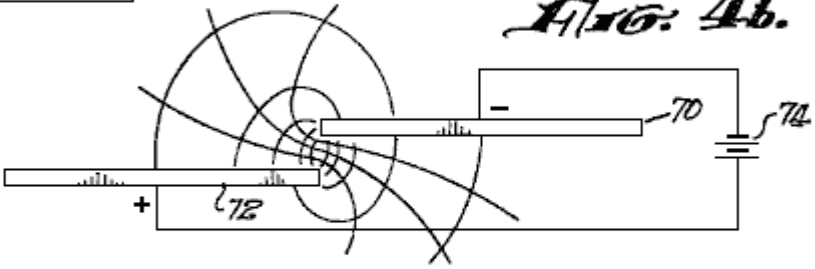


FIG. 4c.

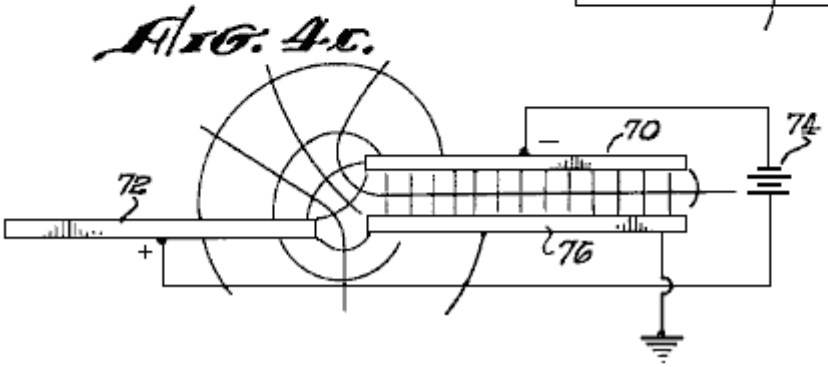


Fig.4a, 4b and 4c are diagrammatic representations illustrating an aspect of the operation of a system incorporating the present invention;

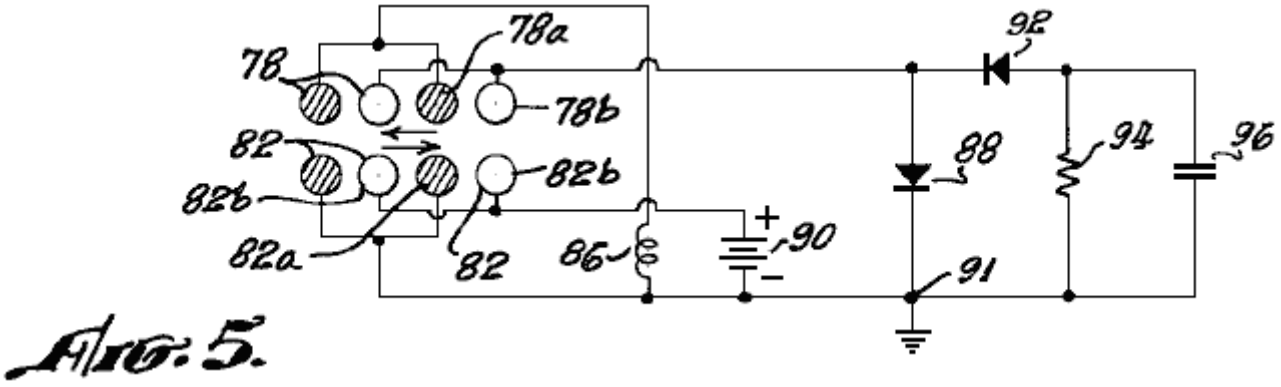


Fig. 5.

Fig.5 is a circuit diagram incorporating one form of the present invention;

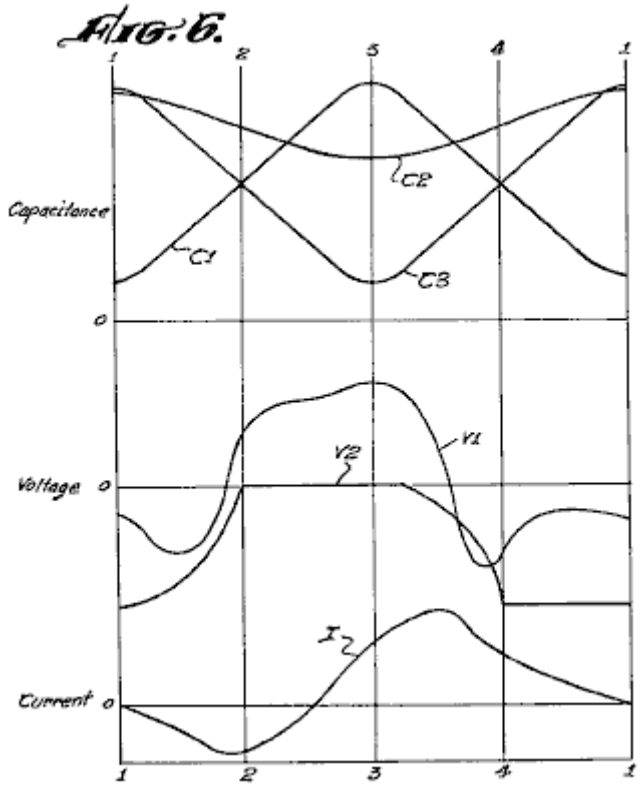


Fig.6 is a set of graphs illustrating the operation of the system shown in **Fig.5**;

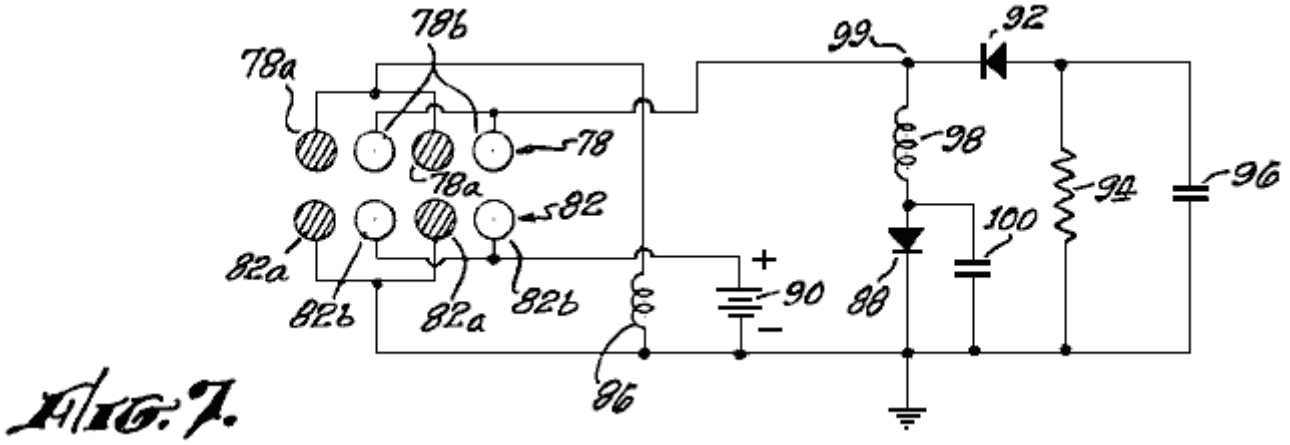


FIG. 7.

Fig.7 is a circuit diagram of another system incorporating the present invention;

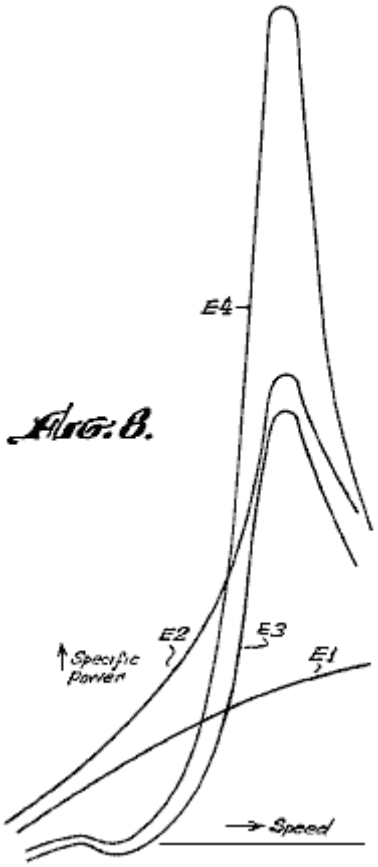


FIG. 8.

Fig.8 is a series of curves illustrating the effectiveness of systems constructed in accordance with the present invention;

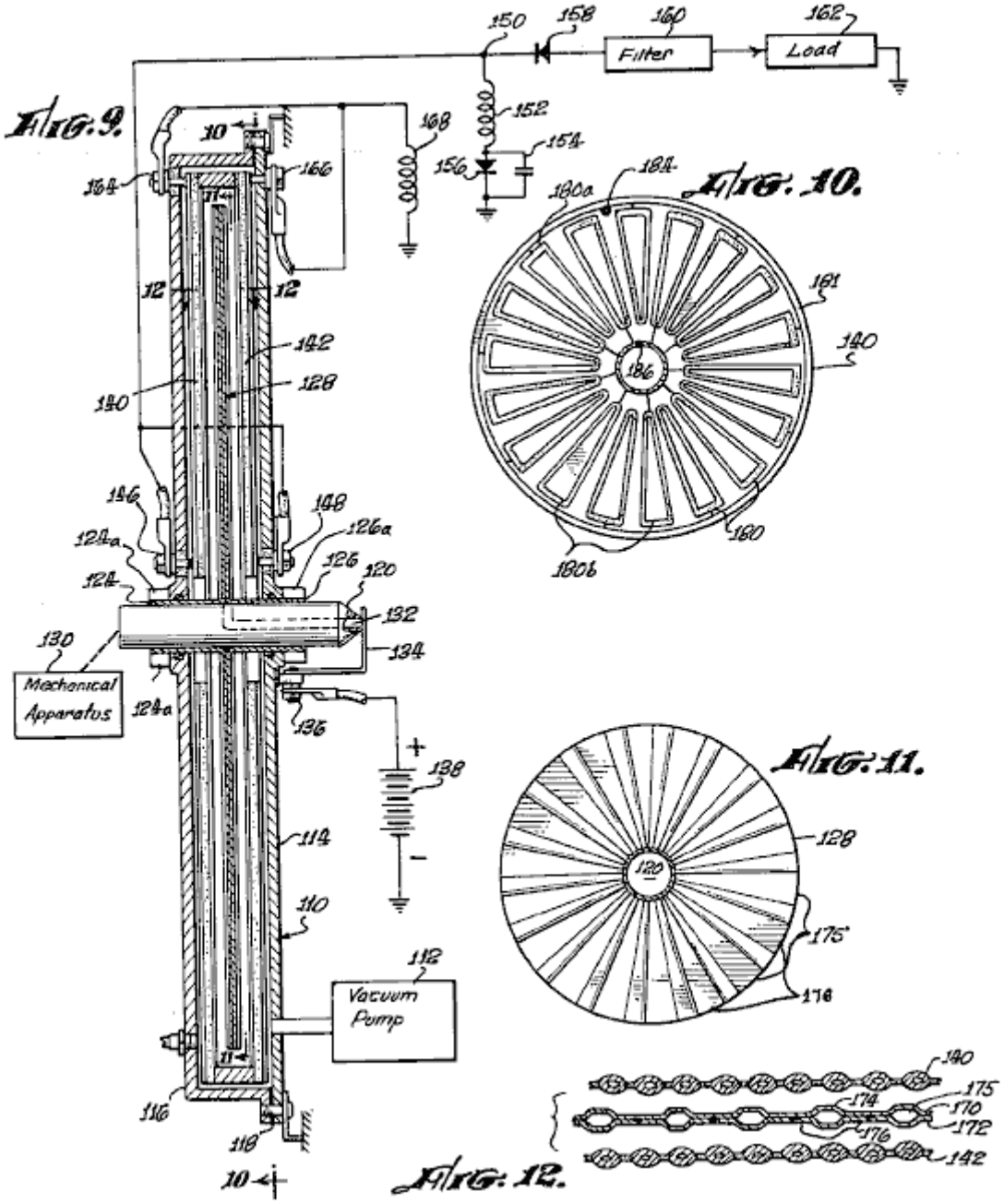
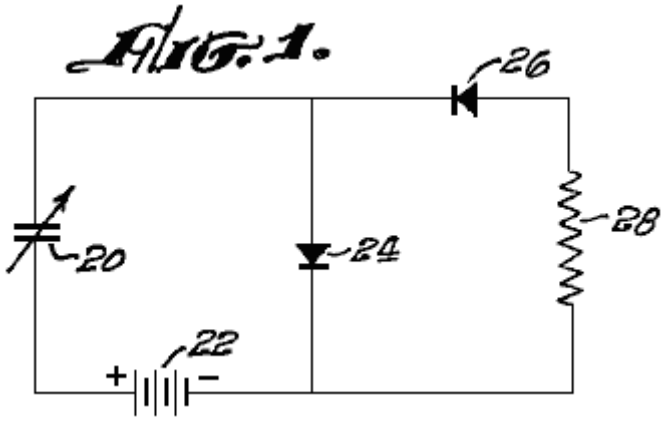


Fig.9 is a vertical sectional view and diagrammatic representation of an apparatus constructed in accordance with the present invention;

Fig.10 is a vertical sectional view along line **10 --- 10** of **Fig.9**;

Fig.11 is a vertical sectional view along line **11 --- 11** of **Fig.9**;

Fig.12 is a horizontal sectional view along line **12 --- 12** of **Fig.9**.



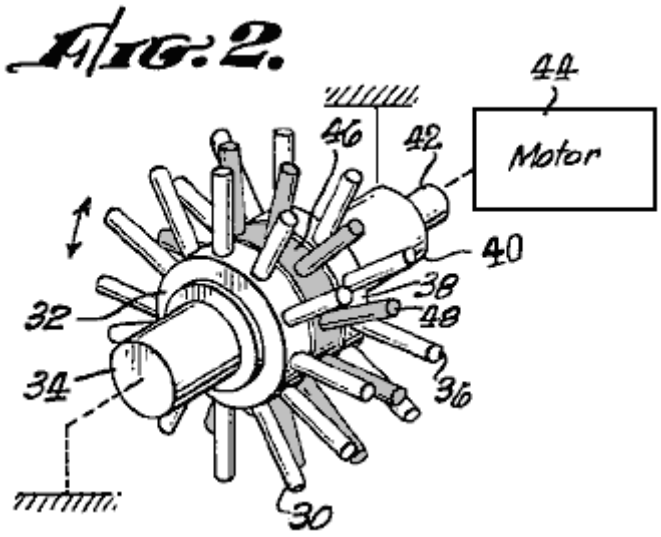
Referring initially to **Fig.1**, there is shown a preliminary form of an electrostatic induction generator. A variable capacitor **20** is connected in a serial loop with a battery (or other source of voltage) **22** and a diode **24**. A series circuit including a diode **26** and a resistor **28** is then connected across the diode **24**. The diodes **24** and **26** are connected in a way which allows current flow in opposite directions through the battery **22**. Variable capacitor **20** is driven by a motor to provide a cyclically-variable capacitance, with the result that the mechanical energy used to drive capacitor **20** is converted to electrical energy supplied to the load resistor **28**.

Considering the operation of the system of **Fig.1** in detail, capacitor **20** is driven to vary between maximum and minimum levels of capacitance. At a maximum level of capacitance, capacitor **20** is charged to a charge level "**Q**" by the battery **22** through diode **24**. The voltage across capacitor **20** then equals Q/C where C is the capacitance of capacitor **20**. Of course, this voltage coincides substantially with the voltage of battery **22** minus the voltage drop across diode **24**.

As the capacitance C of capacitor **20** decreases towards it's minimum value, the charge Q remains constant; therefore the voltage across the capacitor increases to accommodate the decrease in capacitance as V = Q / C. The increased voltage across capacitor **20** produces a current through battery **22**, load resistor **28** and diode **26**. This current tends to recharge battery **22** as well as providing energy to the load resistor **28**.

In this manner, the mechanical energy used to drive capacitor **20** cyclically, is manifest in the form of electrical energy passed to the load resistor **28**. Therefore, in general, this basic operation of the electrostatic generator is similar to an electromagnetic generator, except that the field which couples the mechanical input energy to the electrical output energy is electrostatic rather than electromagnetic.

A wide variety of structures may be used as the variable capacitor in an electrostatic generator. However, in the selected structure, it is normally desirable to reduce the ratio of the peak field voltage gradients to the average gradients by using rounded capacitor elements, e.g. tubes or bars, rather than capacitor plates with sharp corners. Two variable capacitor structures using these conductive bars are shown in **Fig.2** and **Fig.3** and will now be considered in detail.



The structure of **Fig.2** is a radial-bar variable capacitor configuration where rows of radially-extending bars are mounted to be relatively moveable and so provide a variable capacitance between the rows. Specifically, a first row of bars **30** is mounted in a hub **32** which is carried on a fixed shaft **34**. Similarly, another row of radially-

extending bars **36** is supported in a hub **38** which is mounted on fixed shaft **40**, which telescopically receives a concentric shaft **42** which is coupled to motor **44** and which carries hub **46** supporting the row of bars **48** which is positioned between rows **30** and **36**. The bars in each of the rows **30, 36** and **48** may be interconnected electrically by a conductor supported in the respective hubs. However, as the individual hubs **32, 38** and **46** are formed of insulating material, the separate rows are not interconnected electrically. Therefore, a variable capacitance exists between the individual rows **30, 36** and **48**. This capacitance is somewhat similar to a common rotating plate capacitive structure, however, in order to reduce the peak field voltage gradients between the rows, the bars or capacitive elements are rounded to avoid sharp corners.

At a time when the row of bars **48** is in angular alignment with the bars in rows **30** and **36**, the bars are closest and there is maximum capacitance. As the row **48** revolves it reaches a position of least alignment of the bars in rows **30** and **36** (as shown in the drawing) and there, the capacitance between the rows is at a minimum. Therefore as motor **44** rotates the bars in row **48**, a cyclically varying capacitance is provided between these bars and the bars in rows **30** and **36**. In the application of this structure to the present invention, certain of the bars may be used to shape the electrostatic field as described below, while the remaining bars provide the desired variable capacitance. Of course, the number of rows provided in a configuration of this type may be varied to increase the capacitance, as may the size of the bars. Of course, these considerations are determined by the particular application of the system.

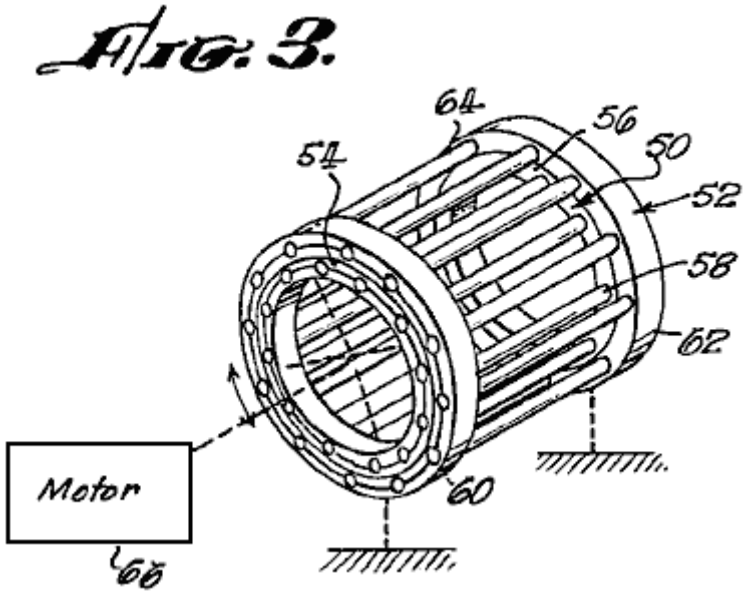


Fig.3 shows another form of variable-capacitance structure which may be used in a system of the present invention. The structure shown in **Fig.3** includes two concentric cylindrical structures **50** and **52**, which are mounted so that they rotate relative to each other. These two structures are similar except for their size. The smaller structure **50** is telescopically fitted concentrically into the larger structure **52**. The smaller structure **50** includes a pair of insulating rings **54** and **56** which are held spaced apart by conductive bars **58** and are angularly offset in a circular configuration. A pair of insulating rings **60** and **62** of the structure **52** are concentrically mounted outside the rings **54** and **56** respectively and are held spaced apart by angularly offset conducting bars **64**. The structure **52**, comprising the rings **60** and **62** and the bars **64** is held fixed, while the structure **50**, comprising the rings **54** and **56** and the bars **58**, is revolved making bars **58** and **64** pass each other in parallel relationship. As a result of this movement, a variable capacitance exists between the bars of the two structures. As in the structure shown in **Fig.2**, the individual bars in the separate structures may be interconnected electrically to provide a composite capacitance, or alternatively certain of the bars may be used for field shaping as described below. Thus the structure shown in **Fig.3**, as well as the structure shown in **Fig.2**, may be used as the cyclically-varying capacitance in the system shown in **Fig.1** to provide electrical energy from kinetic energy. In such a system, the energy would be provided by motors **44** or **66**, however, the source of the kinetic energy is not important for the present invention.

In the operation of an electrostatic generator as shown above, the ratio of power conversion per unit size (and weight) of machine may be increased by increasing the maximum capacitance, or by decreasing the minimum capacitance of the variable capacitor. Of course, it is apparent that as the energy conversion is accomplished in systems of the present invention by a varying capacitor, the effectiveness of the conversion depends on the range over which the capacitor varies. In general, the efficiency of these systems is quite high even when operating with a small capacitive range, however, in such systems the energy conversion capability or power

rating of the system is low relative to it's physical size. One major aspect of the present invention resides in a structure for increasing the range of capacitive change, and will now be considered in detail. For purposes of illustration, a moveable plate variable capacitor will be considered with reference to **Figures 4a, 4b** and **4c**. The flat plate capacitor provides a good example of the universal principle to be described because of the simple electrostatic field present in such a structure.

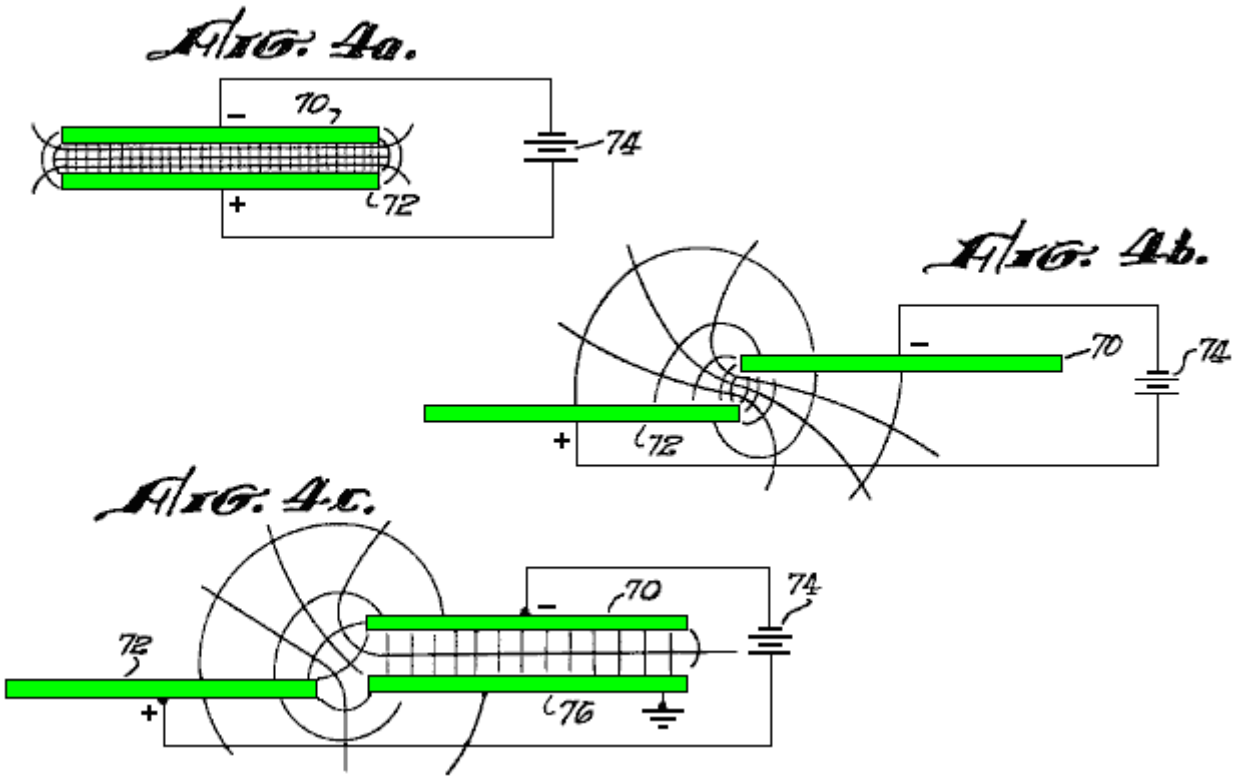


Fig.4a shows a pair of spaced-apart conductive plates **70** and **72** in face-opposing relationship and individually connected to the terminals of a battery **74** or other voltage source. The electrostatic field between plates **70** and **72** is indicated by lines of equal potential, and orthogonal flux lines which indicate the capacitance between the plates. **Fig.4b** shows plates **70** and **72** in an offset position, displaced out of face-opposing relationship to result in reduced capacitance between the plates. Again, the electrostatic field between plates **70** and **72** is indicated by lines of equal potential and the orthogonal flux lines which indicate the capacitance between the plates.

As previously mentioned, any arrangement which can result in greater capacitance for the plates in the configuration shown in **Fig.4a**, or reduced capacitance between the plates in the position shown in **Fig.4b**, results in more effective energy conversion in an electrostatic generator. The system of the present invention includes a structure for shaping the electrostatic field between plates **70** and **72**, to reduce the capacitance between them when these plates are in their position of minimum capacitance. This field-shaping is accomplished by adding extra conductors to influence the field between the plates.

Referring to **Fig.4c**, plates **70** and **72** are again shown in their position of minimum capacitance, however, another plate **76** is shown in face-opposing relationship to plate **70**. Plate **76** is connected to ground or other independent potential. As a result of this, and electrostatic field is provided between plates **70** and **76**, altering the field shape and the potential gradients, as shown in **Fig.4c**. The electric field and the capacitance between plates **70** and **72** is reduced substantially, while the maximum capacitance between these plates (when aligned in face-opposing relationship as shown in **Fig.4a**) remains substantially unchanged.

These same considerations apply in a similar fashion to bar conductors or other configurations of variable capacitors. Of course, the grounded or field-shaping plates as the plate **76** may be provided in both the stator and the rotor of a variable capacitor, or may be carried by only one of these structures.

Referring now to **Fig.5**, there is shown a system incorporating field-shaping electrodes and utilising capacitive structures comprising bars which may take the form generally illustrated in **Fig.2** and **Fig.3**.

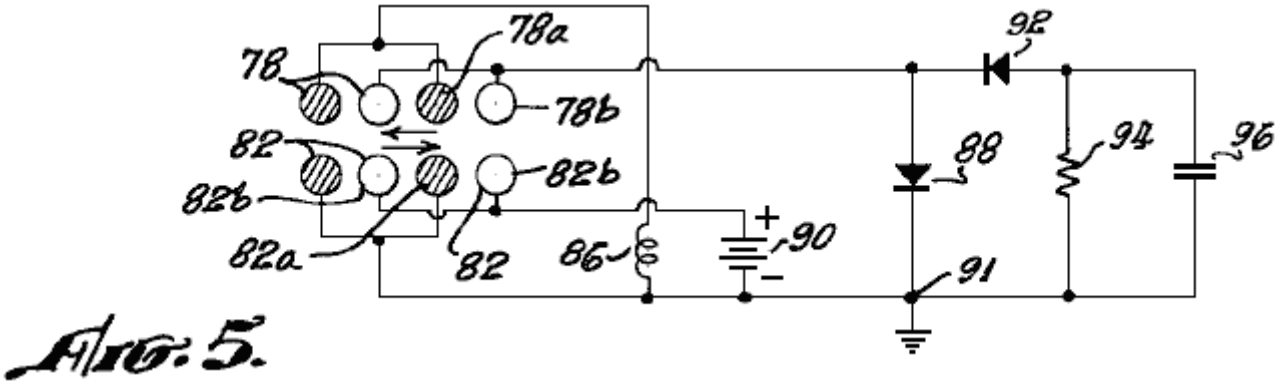


Fig. 5.

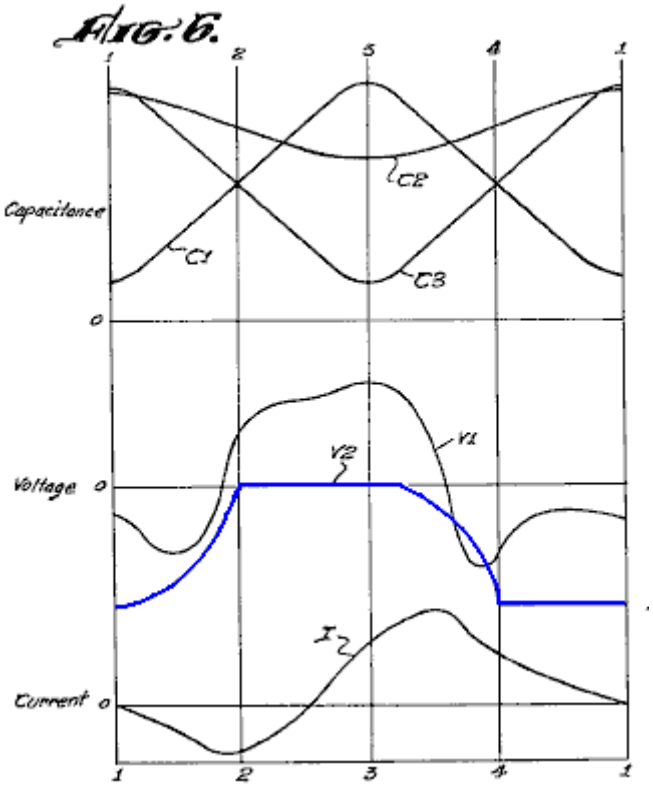
In **Fig.5**, the stator bars are diagrammatically represented aligned in one row **78** and the rotor bars are similarly shown to be aligned in an adjacent row **82**. Of course, either of the rows **78** and **82** may comprise a stator while the other comprises the rotor since relative motion is all that is necessary. However, in the prior art it has become somewhat common to term the rotor conductors as charge inductors while describing the rotor elements as charge conveyors.

In the system shown in **Fig.5**, grounded field-shaping electrodes or conductors are provided in both the row **78** and the row **82**, and are designated bars **78a** and **82a** shown shaded. The electrodes **82a** are connected directly to ground, however the electrodes **78a** are connected to ground through an inductor **86**. The function of this coil will be discussed later.

The conductive bars serving as capacitance elements in the two rows are designated **78b** and **82b** respectively. The bars **78b** (in the row **78**) and the bars **82b** (in the row **82**) are connected in a circuit configuration somewhat as shown in **Fig.1**. Specifically, the bars **78b** are connected in a serial loop with a diode **88** and a battery **90**. The junction point **91** between diode **88** and battery **90** is grounded, and a diode **92** serially connected with a load resistor **94** is connected across diode **88**. A capacitor **96** is connected across the load resistor **94**, and serves as a filter for the load current.

In the operation of the system, relative movement is provided between row **78** and row **82** producing a cyclically variable capacitance to cause the system to function as the system shown in **Fig.1**. That is, during the interval of high capacitance between the rows **78** and **82** (as shown in **Fig.5**), the capacitive structure is charged by battery **90** through diode **88**. Then, as the bars are separated to reduce the capacitance, the voltage across them increases, causing a current to flow in the opposite direction through battery **90**, the load resistor **94** and diode **92**. As alternate bars in each of the rows **78** and **82** are grounded, the electrostatic field is shaped to reduce the minimum capacitance between the bars **78b** and **82b** as disclosed in **Fig.4**. As a result, the capacitance variation is greater and more charge is transferred through the load during each electrical cycle.

In the operation of the system shown in **Fig.5** to convert mechanical energy into electrical energy, a generally capacitive electrical system is provided. Therefore the phase relationships within the system are generally those of a capacitive apparatus. In view of this, the present invention includes the provision of inductors to produce phase shifts and resonant voltage amplifications which increase the effective power-conversion capabilities of the system. In **Fig.5**, such an inductance **86** is connected between the bars **78a** and ground. The operation of the system of **Fig.5**, including the inductor **86** is graphically illustrated in **Fig.6** which includes plots of capacitance, voltage and electrical current versus the position of the capacitive structure.



The upper family of curves are capacitance plots, in which, curve **C1** is a plot of varying capacitance between the conductive bars **78b** and **82b**. The curve **C2** is a plot of the capacitance between bars **78a** and **82a**, while the curve C3 is a plot of the capacitance between bars **78a** and bars **82b**. These curves are plotted against a uniform variation in relative positions between row **78** and row **82**, the curves beginning at a point of minimum capacitance for the structure. **Fig.5** illustrates the relative position marked as **3** in **Fig.6**.

The voltage curves in **Fig.6** are plotted on the same base as the capacitance curves, and include a curve **V1** which represents the voltage across coil **86**, and a curve **V2** which represents the voltage on the bars **78b** relative to ground potential. The curve **I**, illustrates electrical current flow through inductor **86**, and it is also plotted with respect to capacitative displacement.

Considering the operation of the inductance **86**; in view of the capacitance between the grounded bars **78a** and the capacitively-charged bars **82b** (connected to battery **90**) a charge is deposited on the bars **78a**. Of course, in the absence of the inductance **86**, this charge would be immediately passed to ground. However, by providing the inductance **86**, the voltage of curve **V1** is developed across this element. This voltage serves to produce further field shaping between the bars and results in more efficient power conversion. That is, the voltage of curve **V1** appears on the bars **78a** in phase with the capacitance **C1** so that the field influence of bars **78a** serves to increase the charge induced at or near the maximum of capacitance **C1** and also serves, by reducing the field strength to reduce the minimum charge held on bars **78b** and **82b** at, or near the instant of minimum capacitance. Thus the variation in the electrostatic field causes the system to absorb more torque, resulting in increased power-conversion capability for the system.

In addition to the provision of inductor **86** as shown in **Fig.5**, other circuit configurations using inductors are to further affect resonant current surges which increase and decrease the field voltages in proper phase relationship with the capacitance variation, and so increase the amplitude of the capacitor alternating current. One such other variation is shown in **Fig.7** which uses bars similar to those of **Fig.5** and in which, like elements are similarly identified.

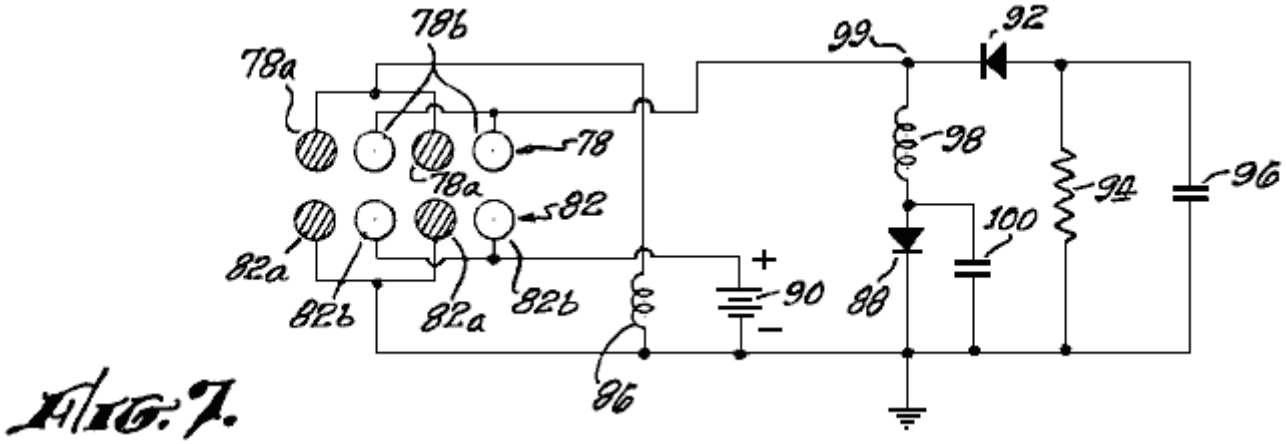


FIG. 7.

In the system shown in **Fig.7**, an inductor **98** is connected between diode **88** and the junction point **99** between diode **92** and the bars **78b**. Furthermore, a capacitor **100** is connected across diode **88**. In general, the added resonant circuit provides further current surges in the desired phase relationship with the variation of capacitance. Also, the resonance between inductor **98** and capacitor **100**, tends to produce higher voltages in the electric field of the variable capacitor relative to the voltages across the excitation source **90** and the load **94**. The capability of the machine to achieve high power from a small size of variable capacitor is thereby greatly improved.

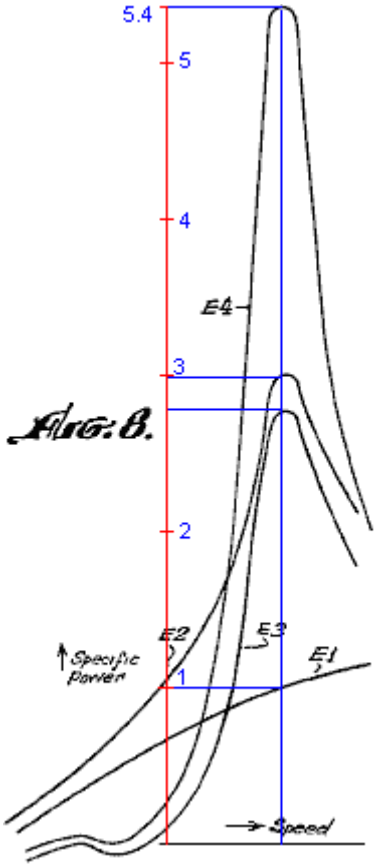
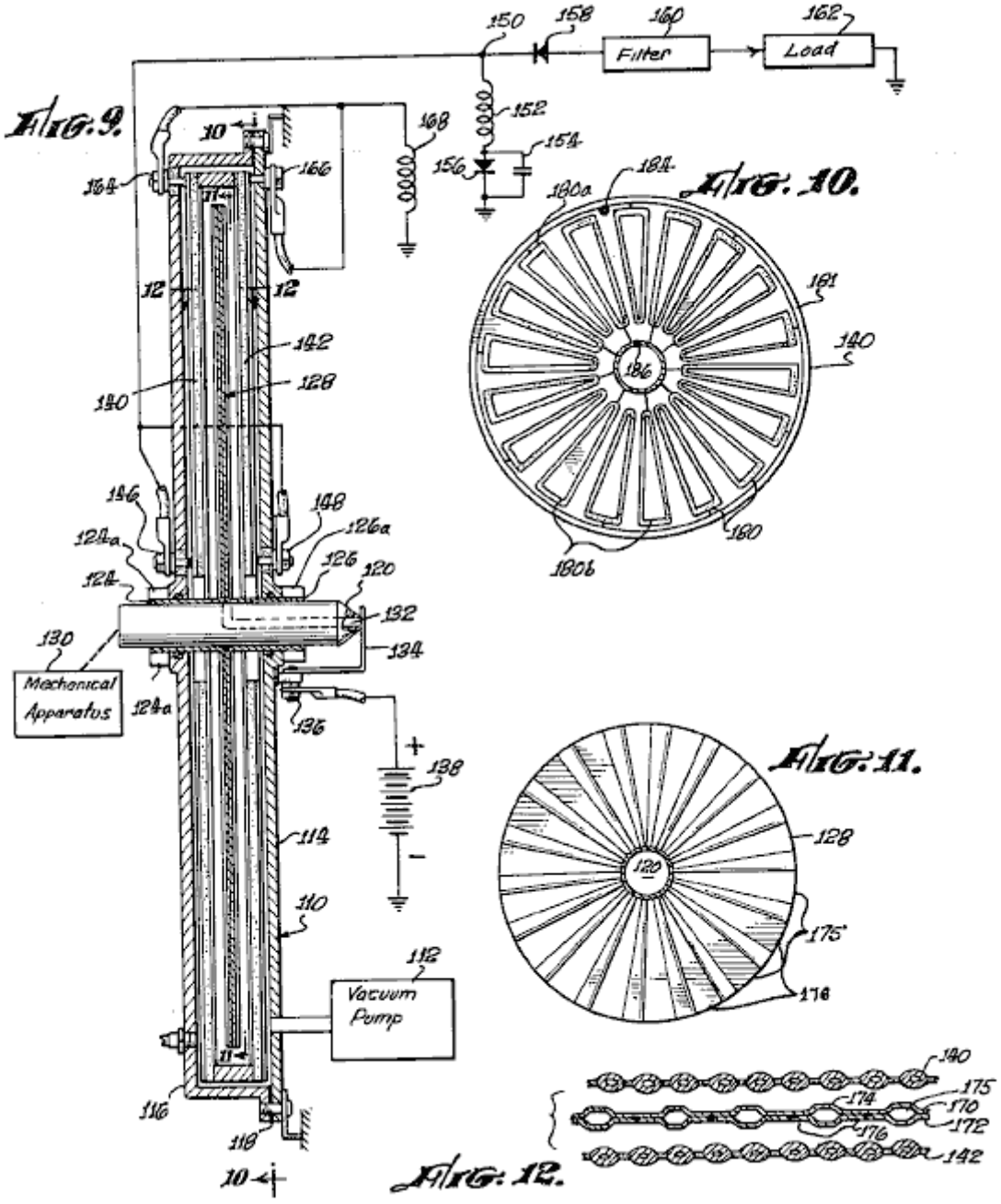


FIG. 8.

A graphic indication of the improvement to a system resulting from the application of certain principles of the present invention may be seen in **Fig.8** which includes several curves descriptive of various apparatus, and wherein speed is plotted along the horizontal axis versus specific power which is plotted along the vertical axis. Specific power is herein defined as the power output per unit of excitation voltage squared (P/E_o^2) for a given size and configuration of machine.

The curve **E1** is indicative of the operating characteristic of the simple system shown in **Fig.1**. The curve **E2** (indicating a substantial improvement in performance) is indicative of the system shown in **Fig.5**. The curve **E3** indicates the improvement of the system of **Fig.7** (without the aspect described with respect to **Fig.5**), while the curve **E4** indicates the operation of the system shown in **Fig.7** as illustrated there.

It is apparent that the power conversion capabilities of the system of **Fig.7** (indicated by the curve **E4**) are greatly improved over the simple system which does not include these aspects of the present invention, particularly if the generator is operated at a fixed speed, or if means are provided to adjust the inductance or capacitance values of components **86, 98** and **100** to give the tuned peak power at whatever speed the machine is being operated.



Although various aspects of the present invention may be differently applied to electrostatic machines, to accomplish the desired results, one specific illustration of an operating machine is shown in **Figures 9,10, 11 and 12**, and will now be considered in detail.

The system is mounted in a housing **110** which is sealed and maintained evacuated by a vacuum pump **112**. By operating the cyclically-variable capacitor apparatus in a relatively high vacuum, insulation problems are reduced and windage losses are minimised. The housing **110** has a flat cylindrical form and consists of a circular plate bolted to a cup-shaped member **116** with an annular seal **118** provided at the junction between these members.

A mechanical shaft **120** passes concentrically through housing **110** and is journal led into bearings **124** and **126**, which incorporate vacuum seals **124a** and **126a**. The shaft **120** has a rotor **128** coaxially fixed on it and is connected to a mechanical apparatus **130** by which the shaft is driven. Shaft **120** (made from insulating material) has a conductive section **132** formed in it, which electrically connects rotor **128** through a brush **134** and a lug **136** to one terminal of a battery **138**, the other terminal of which is grounded.

In general, the basic electrical system is similar to that of **Fig.7** with the stator elements being connected by lugs **146** and **148** to a junction point **150** which is connected through an inductor **152** and a capacitor **154** to ground. A diode **156** is connected in parallel with capacitor **154**. The junction point **150** is also connected through diode **158** to a filter **160** which is in turn connected to a load **162**. The field-shaping electrodes in the stator elements **140** and **142**, are connected from lugs **164** and **166** through an inductor **168** to ground.

Considering the cyclically-varying capacitance apparatus in greater detail, the rotor comprises a pair of plates **170** and **172** (**Fig.12**), which are formed with angularly-offset radiating grooves **174** which are interconnected by spot welds **176** so that the grooves **174** essentially form radiating bars **175**. In this structure, the rotor is entirely conductive material, e.g. metal, formed to comprise radiating tubes **175** (**Fig.11**) which are connected through the shaft conductor **132** and the brush **134** to the battery **138**.

The stator components **140** and **142** are similar and comprise a plurality of tubes **180** (**Fig.10**) radiating in a plane from an axial location in the stator element and supported in a circular stator mounting **181** by being imbedded therein. Tubes **180** are formed of metal and the mounting **181** supporting the tubes is formed of insulating material as a casting of epoxy material. Alternate of the tubes **180** are designated **180a** and are connected electrically to a terminal **184** which is contacted by one of the lugs **146** or **148** at a point **186**.

The lugs **164** and **166** are connected to ground through inductor **168**, with the bars **180a** used as field-forming bars while bars **180b** are the actual variable-capacitance elements. In this structure, only the stator is provided with field-shaping electrodes; however, it is to be understood, that in various other embodiments, field-shaping electrodes may be provided in both the stator and the rotor, or in only one structure, as taught with respect to **Fig.5** and **Fig.7**. Also, in the system shown in **Fig.9**, only a single rotor plate is provided; however it is to be clearly understood that a large number of intermeshed rotor and stator plates may be provided in systems of the present invention, depending on the rated power conversion capability of the system.

In the operation of the system shown in **Fig.9**, rotor **128** is revolved relative to the stator elements **140** and **142** by the mechanical apparatus **130**. Therefore, the tubular rotor bars are driven in and out of alignment with the radiating stator bars **180b**. Therefore the connections to these elements (lugs **146** and **148**, and brush **134**) experience a cyclically-varying capacitance. The stator bars **180a** extend the range of this capacitance in the manner described with respect to **Fig.4**.

During intervals of high capacitance, the cyclic capacitive structure is charged by battery **138** through inductor **152** and diode **156**. As the capacitance of the structure decreases, the voltage therein increases, providing a current through diode **158**, and filter **160** to the load **162**. Thus, mechanical energy from the mechanical apparatus **130** is efficiently converted into electrical energy and supplied in that form to the load **162**.

While the previous discussions have pertained primarily to electrostatic generators, it is to be understood that the important features of the present invention apply equally to electrostatic motors.

An important feature of the present invention resides in the structure disclosed in **Fig.9**, which may be economically manufactured and used in practical electrical systems.

Another important feature of the invention resides in the use of field-forming electrodes in the stator or rotor, or both the stator and the rotor to accomplish more effective energy transitions.

One other important feature of the invention resides in the provision of inductance elements to further increase the effectiveness of the system.

The Onezime Breaux Electrostatic Generator

US Patent 4,127,804 28 Nov. 1978 Inventors: Onezime Breaux

ELECTROSTATIC ENERGY CONVERSION SYSTEM

This is an important design, but as the physical details in the patent seem rather limited, instead of reproducing the patent here, I will attempt to provide a description which should be more understandable for the average person.

The Principle

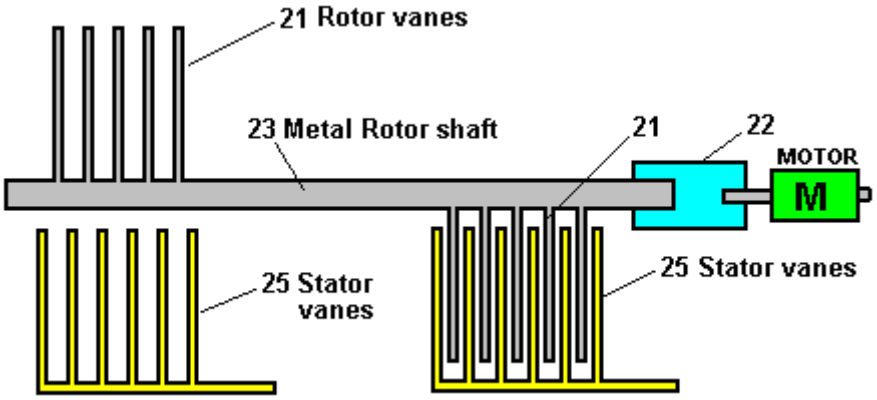
For there to be a greater energy output than the energy input required to make any generator operate, that extra energy must come from somewhere. There is no magic about any of these systems as we ourselves are immersed in a massive energy field. We do not notice this any more than we notice the radio waves which pass by, and through us, all the time. This energy field is very powerful as Thomas Henry Moray demonstrated many times by gaining fifty kilowatts of electrical power from just a simple aerial and a ground connection. We don't know how he did that, but we would very much like to know how. Nevertheless, Moray demonstrated that there is massive power all around us – power which is free for the taking **if** we know how to do it. Unfortunately, most of the large number of people who do know how to extract this energy, refuse to share that information, leaving us to work out from basic principles, how to gain this free-energy. It should be noted in passing, that we already use this free energy when we use "mains electricity" as none of the energy supplied by the local Power Company actually comes out of your mains wall socket since all of the current supplied by the Power Company flows directly back to that company. All the Power Company does is to set up the conditions which cause some of the free-energy in your location to flow through the electrical equipment which you plug into that wall socket. What free-energy devices do, is to set up those same conditions to draw in electricity from our massive surrounding energy field, without the need for any help from the local Power Company.

The Design

When you try to start a car engine, you connect the car battery to an electrical Starter Motor which turns the engine over and causes sparks to be generated inside each of the cylinders of the engine. The current draw from the battery passes through the Starter Motor and is lost for ever. If the car engine starts, then some of the energy stored in the fuel burnt by the engine is used to replace the electrical current taken from the battery. If the car engine does not start, then after a fairly short time, all of the electrical energy stored in the car battery will be fed to the Starter Motor and the battery will no longer be able to supply any more current to turn the engine over.

The generator described here is a very clever design as any output power which it provides, it keeps and does not lose any of it. It is somewhat like that car Starter Motor returning the current to the battery, allowing any number of attempts to start the engine. The way that it operates is to pass current backwards and forwards between two capacitors. That current is made to pass through the primary winding of an output transformer, exactly mimicking the signal which the Power Company charges you for. However, with our generator, that current is never lost and can be used over and over and over again. A battery is used to create that initial current, and when the generator starts operating, it recharges the battery and if desired, the battery can be removed completely and the generator is then self-sustaining. The battery is only used to start the generator running and is not used at any other time.

The outline design is very straightforward and can be visualised as shown here:

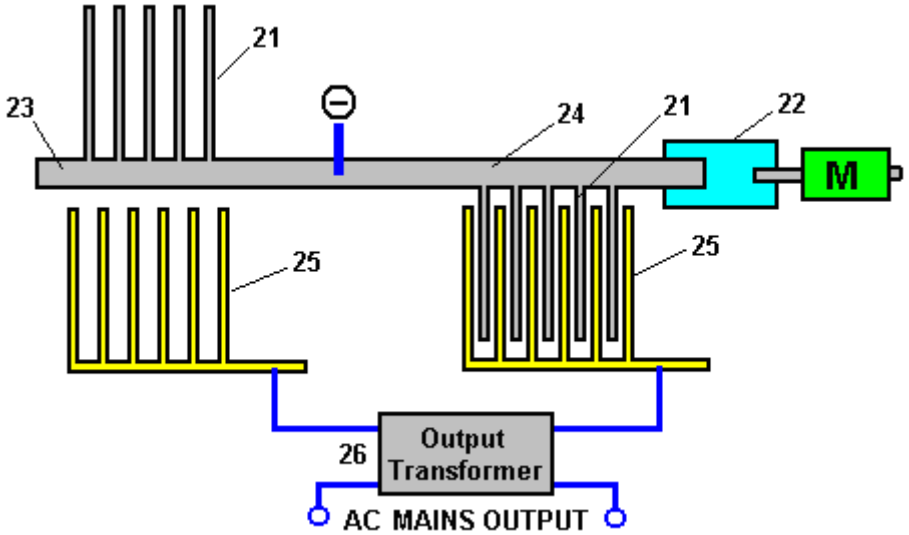


Here, the motor marked **M** in the diagram, spins a metal shaft **23**. The shaft is mounted on ceramic roller bearings (as used on skateboards) and nothing on the shaft touches any other part, which means that the shaft spins very freely and very little power is needed to spin it.

Shaft **23** has metal 'vanes' or plates **21** attached to it and these pass though somewhat similar metal vanes **25** which are fixed in place and do not move. These four sets of metal vanes form two separate variable capacitors.

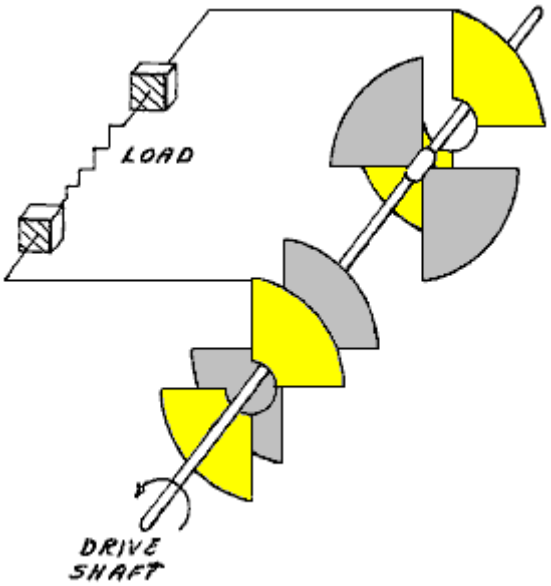
The vanes **21** are arranged so that when one set reaches it's maximum capacitance (being fully inserted into the non-moving vanes **25** as shown on the right of the diagram) then the other set of vanes is at it's minimum capacitance being as far away as possible from it's matching set of vanes **25**.

Initially, at start-up, both of the sets of vanes marked **25** are fully charged to a high positive voltage by an electrical system (not shown here) and the output transformer is connected between these two sets of vanes, producing this arrangement:



The sets of vanes **21** are connected to a strongly Negative charge, passed to them through the metal shaft **23**. When the negative plates **21** are inserted between the Positively charged plates **25**, all of the positive charge on the unmeshed plates **25** gets sucked towards the negative charge on the meshed plates **21**, but that charge is forced to flow through the output transformer on it's journey. This arrangement causes the positive charge on the fixed plates to shuttle backwards and forwards through the output transformer, never getting 'used up' in the process. This means that the generator has an electrical output of 220 volts at 50 cycles per second and the only input power needed is the tiny amount required to spin the rotor shaft. The force needed to spin the rotor shaft is not affected in any way by the intensity of current drawn from the output transformer, which means that the generator is wholly independent of the load. While this is a very clever design, a number of practical construction details can be altered to improve the performance.

The Breaux patent shows a vane shape which doubles the frequency at which the drive motor spins and which is balanced on the shaft, producing no shaft vibration when spun rapidly. The vane shape is shown here:



You will notice that the rotor vanes (shown in grey), are positioned ninety degrees apart, providing even better shaft balance. All of the rotor vanes have the same shape and all of the stator vanes have the same shape, which is helpful for mass production. While just two rotor vanes and two stator vanes are shown in this introductory diagram, there can be as many pairs of rotor/stator plates as desired. Each additional pair of plates increases the output power of the generator.

This design calls for the vanes to be in a vacuum, which strongly suggests that this device was never prototyped as operating in a vacuum, unless all conducting surfaces have exceptional insulation, would be a major disaster

as the coronal discharge would be massive – after all, that is how thermionic valves or 'vacuum tubes' are made and they carry large electrical currents through a vacuum. For this design to function properly, it needs to be operated in dry air.

The power output of a device of this nature increases with increasing vane area, and with decreased inter-vane gap and with increased voltage. Consequently, a working voltage considerably higher than 220 volts will be used when charging the vanes. This means that the output voltage will be much higher than is convenient. Generators of this type generally show a resistive load and only a resistive load. A practical implementation would require a step-down transformer on the output in order to reach a working AC voltage of 220V. It is generally thought that the inductive nature of the primary winding of any such transformer would create a major choking effect on the oscillating flow of charge between the vanes, due to the back-EMF of that winding. That is not the case if a Thane Heins style of transformer is used (as shown in chapter 3) as a transformer of that type does not have any significant back-EMF, making this a very interesting design, especially since the output current draw does not affect the input power which is solely used to spin the rotor shaft.

In theory, the static charge initially placed on the capacitor vanes is never depleted. However, in this world, perfect conditions never seem to be encountered. In this instance, there is a **very** slight loss of charge over time, and that charge is boosted very briefly after a substantial time has elapsed. A brief high-voltage pulse once every ten minutes is likely to be sufficient. There is no corona energy loss as corona streamers only form at high frequency and this generator does not exceed an absolute maximum of 60 Hz which is a very low frequency indeed, so there is never any corona discharge at all.

Patrick Kelly
http://www.free-energy-info.tuks.nl
http://www.free-energy-info.com
http://www.free-energy-devices.com

Chapter 8: Fuel-less Engines

We have been raised with the idea that it is necessary to burn a fuel to produce power which we can use. We are sold coal, coke, timber, paraffin/kerosene, petrol/gasoline, diesel, propane, etc. for us to burn in order to "get" energy. While it is perfectly true that burning these things will indeed result in energy in a form which we find convenient to use in heating, cooling, powering engines, etc. what is carefully avoided is the fact that it is not at all necessary to burn a fuel in order to run the things which we want to power. This 'inconvenient' fact has been concealed and denied for more than fifty years now (very surprisingly, by the people who want to sell us these fuels to burn – do you perhaps think that they may have some motive for this, other than our best interests about which they are no doubt, very concerned?).

This chapter is about 'fuel-less' motors. Strictly speaking, they are not self-powered but as they don't burn a fuel of any kind, in everyday language they can be described as 'self-powered'. In the same way that a solar panel in sunlight uses no fuel and yet puts out electrical power, these motors draw energy from the environment and provide us with mechanical power. In actual fact, power is never "used up" but just converted from one form into another. In the case of our trusty solar panel, some 17% of the radiation from the sun (mainly ultraviolet) is **converted** into electrical power and 83% goes in heating and other losses, but as we don't have to supply the sunlight, and the solar panel pours out the electricity which we want without us having to do anything to make it happen, we really don't care very much about its extremely low efficiency. As far as we are concerned, the electricity flowing from the panel is "free-energy".

It is really amazing that we have been persuaded that we must burn a fuel in order to get power. Take the case of a heavy-displacement sailing yacht. The skipper can voyage using his inboard diesel engine:

This matches perfectly with the thinking that you need to burn a fuel in order to get power as the yacht is moving along, pushed by the engine which is powered by burning diesel fuel. But, what if the skipper decides to switch the engine off and set the sails?:

Now, the same boat, weighing exactly the same with the same crew, is now continuing the voyage at the same speed, but no fuel is being burnt. The really interesting thing is that while we know this perfectly well, and we are aware that people have sailed right around the world in boats which do not have engines, it does not seem to occur to us that this shows conclusively that it is not necessary to burn a fuel to power some item of equipment or form of transport.

In the case of our yacht, the energy comes from the sun which heats the atmosphere unevenly, causing winds to blow and the yachtsman uses the sails to make those winds power his boat through the water. So, a sailing boat is actually powered by the sun although we don't usually think about it that way.

There are many hydro-electric "power stations" where electricity is 'generated' by machines driven by water pressure. In actual fact, no power is 'generated' at all, but instead, the potential energy of the body of water is **converted** into electricity by having the water fall and spin the shaft of a machine. So, how did the water get up there in the first place? Well, it came from rain. And how did the rain get up there? It rose up there due to evaporation caused by the heat of the sun. So, the bottom line again is that hydro-electric 'power' stations are powered by the sun.

Windmills are also powered by the sun. But, and here is the really interesting thing, if I state that it is perfectly possible for a compressed-air engine to produce mechanical power without burning any fuel, then there is an immediate and strong reaction where people will say "Impossible – that is perpetual motion !!" They imply that perpetual motion is impossible but never supply any rational evidence to support that implication. The Earth has been spinning on its axis for millions of years, so when exactly do they expect it to stop? All the planets in the Solar System have been orbiting for millions of years, how long do they have to orbit before they can be considered to be in perpetual motion? Why then are people so opposed to the idea of perpetual motion? Presumably, because perpetual motion shows clearly that a fuel does not have to be burned to 'produce' power and that would not be good for people who sell fuels, and so, we are all told from an early age that perpetual motion is "impossible".

Well, that does not matter here as we are going to look at compressed-air engines which run off the heat of the sun. That is, they are heat-pumps which are a well accepted engineering fact and they work on wholly accepted standard scientific principles. An ordinary refrigerator outputs three or four times as much heat power as the electrical power driving it, and it could be twice that efficient if it were used properly. This is a Coefficient Of Performance (COP) of 3 or 4, which is supposed to be "impossible" but unfortunately, all refrigerators work like this and you can't exactly say that refrigerators don't exist, just because their performance does not appear to fit in with some theories.

Actually, there is no magic involved here as the extra energy is being drawn from the heat content of the air in the immediate locality. The refrigerator is not operating in isolation and there is a heat exchange with the air surrounding it. This outside energy causes the COP>1 performance. In passing, all COP>1 devices operate by drawing energy in from an external source (usually the zero-point energy field) and none of them actually break the 'rules' of science. But, enough of that.

The people who don't want self-powered engines used in the world today, pin their hopes on a continued ignorance of Engineering facts relating to heat pumps. A self-sustaining compressed-air engine is actually running off power from the sun just as sailboats, windmills and hydro-electric power stations do. Sorry folks, no magic here, just bog-standard Engineering. Admittedly, very few people know or realise the implications of this standard Engineering:

1. All work done in compressing air into a storage tank is converted into heat and then lost to the atmosphere, so the energy in the compressed air inside the tank is the same as that produced by atmospheric heating of that air, but as more of it is now in the tank, there is additional potential for work to be done. This extra energy was fed into the air by atmospheric heating before the air was compressed.

The First Law of Thermodynamics states that where heat is converted into mechanical energy, or mechanical energy is converted into heat, the quantity of heat is exactly equivalent to the amount of mechanical energy. We then have the intriguing situation where all of the mechanical energy put into compressing air into a storage tank is lost as heat, and yet, the tank contents now has a higher potential for doing work. This information comes from Engineering textbooks.

2. If the expanded cold air leaving the engine is used to cool the intake air of the compressor, then there will be an added gain when it warms up inside the cylinder, pulling heat in from the local environment.

3. If the heat of compression is transferred to the air container feeding the engine and not given time to dissipate, then there is a further power gain for the engine.

4. If compressed air is allowed to expand rapidly, there is a marked drop in temperature. The Leroy Rogers engine design, shown later in this chapter, uses this fact to create air-conditioning for a car driven by a compressed-air engine.

OK then, in broad outline, the energy available from a tank of compressed air comes directly from the heat contained in the atmosphere, in spite of the fact that we always imagine that the energy in the tank was put there by our energetic pumping.

Let's check this out by taking a look at some of the engines which use these principle to provide fuel-less operation, starting with the design of Bob Neal. The full patent for Bob's design is included in the Appendix.

Bob Neal's Compressed Air Engine.

Bob Neal's design is a compressed-air operated engine and compressor where the operation of the engine keeps re-supplying the compressed air tank:

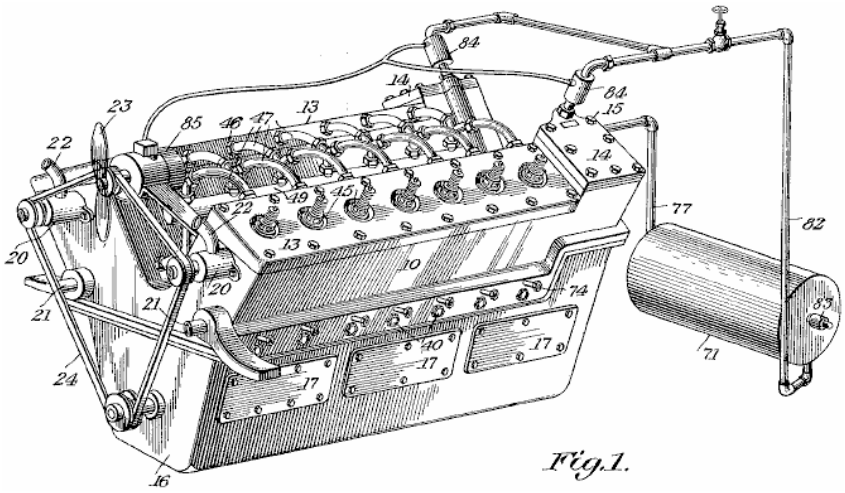
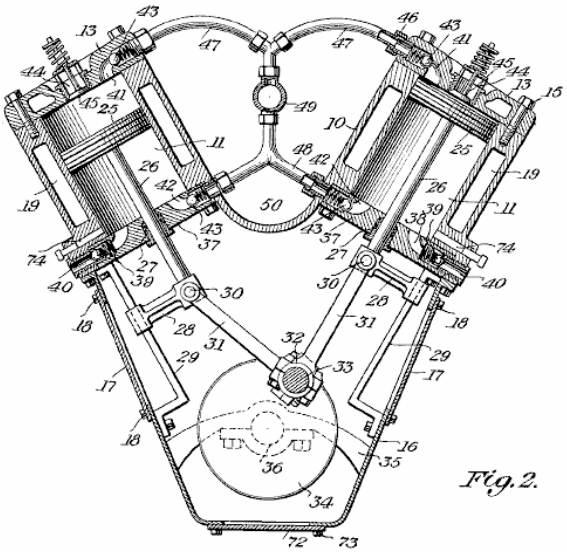


Fig.1.

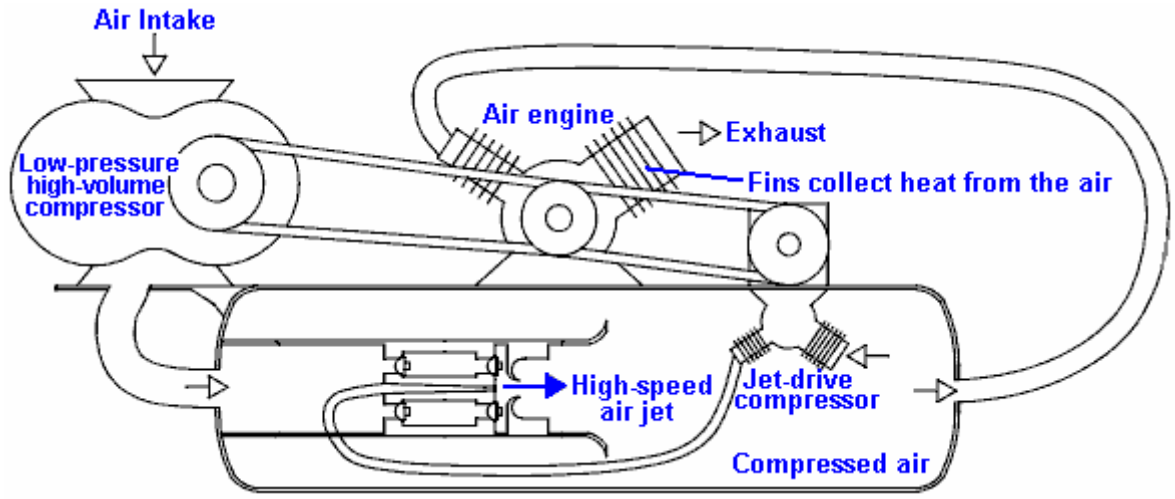
This is a perspective view of the engine and this:



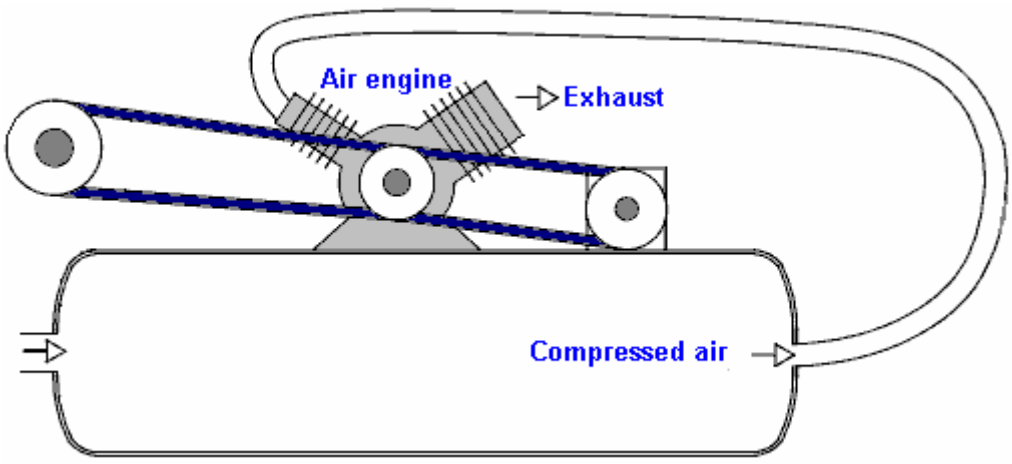
is a vertical transverse cross-section view through the compressor part of the engine. In his patent, Bob has avoided any direct mention of the fact that his engine design is fuel-less. That sort of statement is not popular with Patent Examiners even if it is perfectly true.

Scott Robertson's Compressor System.

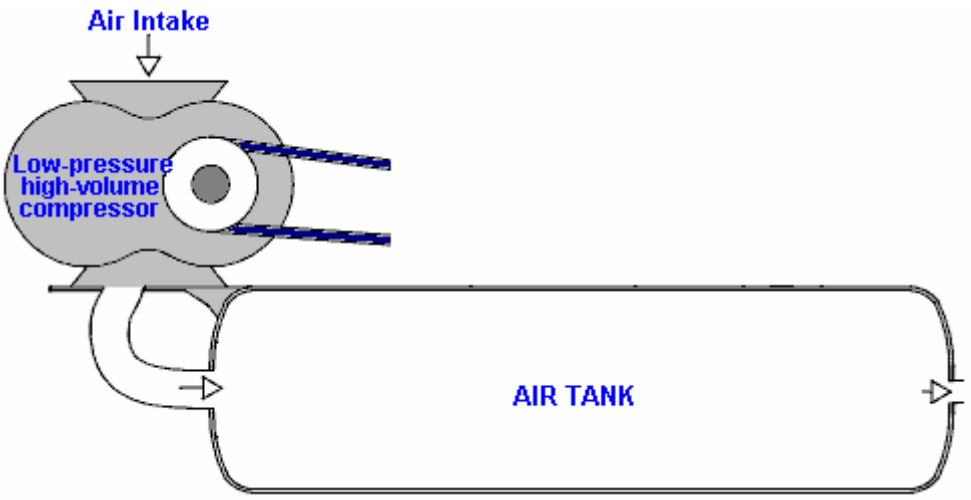
Bob Neal's system could do with some further explanation, so here is an idea from Scott Robertson whose web site is http://www.aircaraccess.com/index.htm, for a possible working compressor system using a leaf-blower:



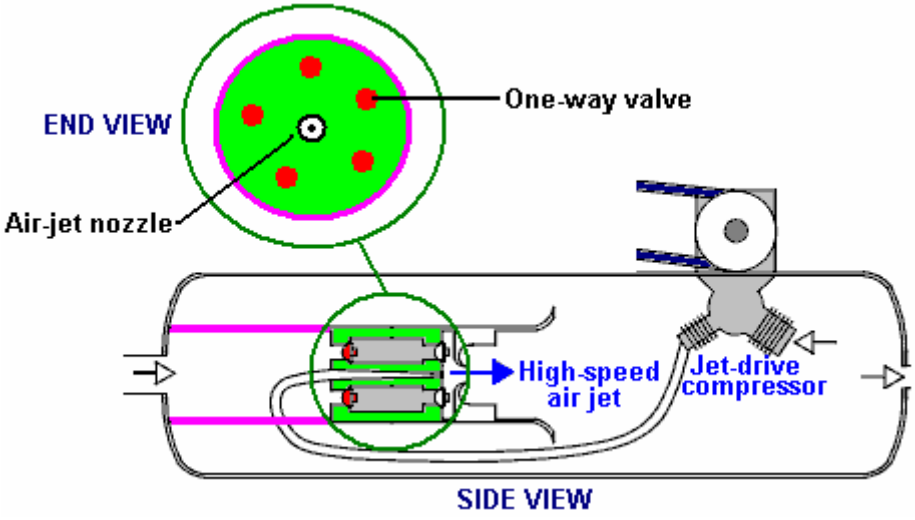
While this looks rather complicated, in reality it really isn't. Let's take the different sections in order:



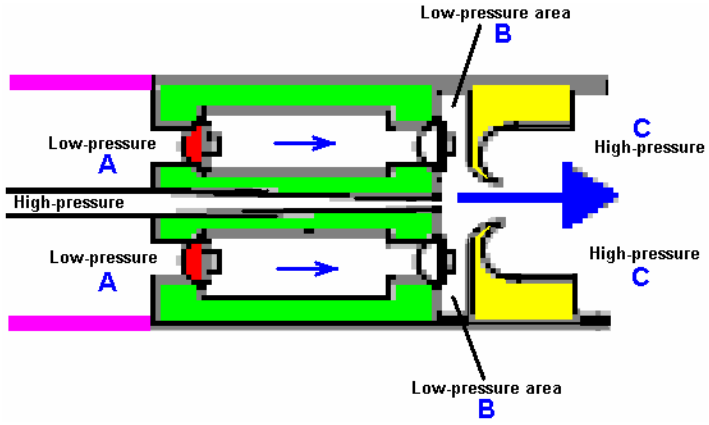
First, you have an ordinary air engine, supplied with compressed air from a pressure tank. This engine exhausts its (cold, expanded) air to the atmosphere. The engine powers two compressors which between them keep the tank full of compressed air.



The first compressor is a simple 'leaf-blower' type which produces a large volume of low-pressure air. The big question is "how do you get this large volume of low-pressure air into a tank which has high-pressure compressed air inside it?". Well this seemingly impossible task is performed by the second compressor aided by a cunning, ultra-simple design:



Here, low-pressure air is fed into the low-pressure area marked in pink. Separating it from the high-pressure area is a metal plug marked in green. Set into this plug is a ring of five one-way air valves marked in red. These one-way valves let the low-pressure air into the high-pressure area because of a high-speed jet of air produced by the 'jet-drive compressor'. At first glance, this seems impossible, but it is actually just an application of a standard Engineering technique. The high-speed air jet is directed through a specially shaped nozzle, creating a local low-pressure zone around the jet:



The low-pressure air at point “A” flows through the ring of five one-way valves into the disc-shaped low pressure area “B” and is blasted into the high-pressure area “C” by the high-power air jet ripping through the doughnut-shaped ring marked in yellow. The high-speed air jet causes the low pressure ring “B” by its rapid movement which creates a vortex due to the shape and positioning of the doughnut-shaped ring marked in yellow. This clever arrangement allows large volumes of low-pressure air to be drawn into a tank which contains high-pressure air.

You will also note that the two-stage compressor which generates this high-speed jet of air, has its working area actually inside the tank. This means that the heat of compression is used to heat the air inside the tank and raise its pressure, enhancing the operation further. It should be borne in mind that the new air entering the system has been heated by the sun and contains the energy which powers the system.

The Retro-fit Compressed Air Vehicle System of Leroy Rogers.

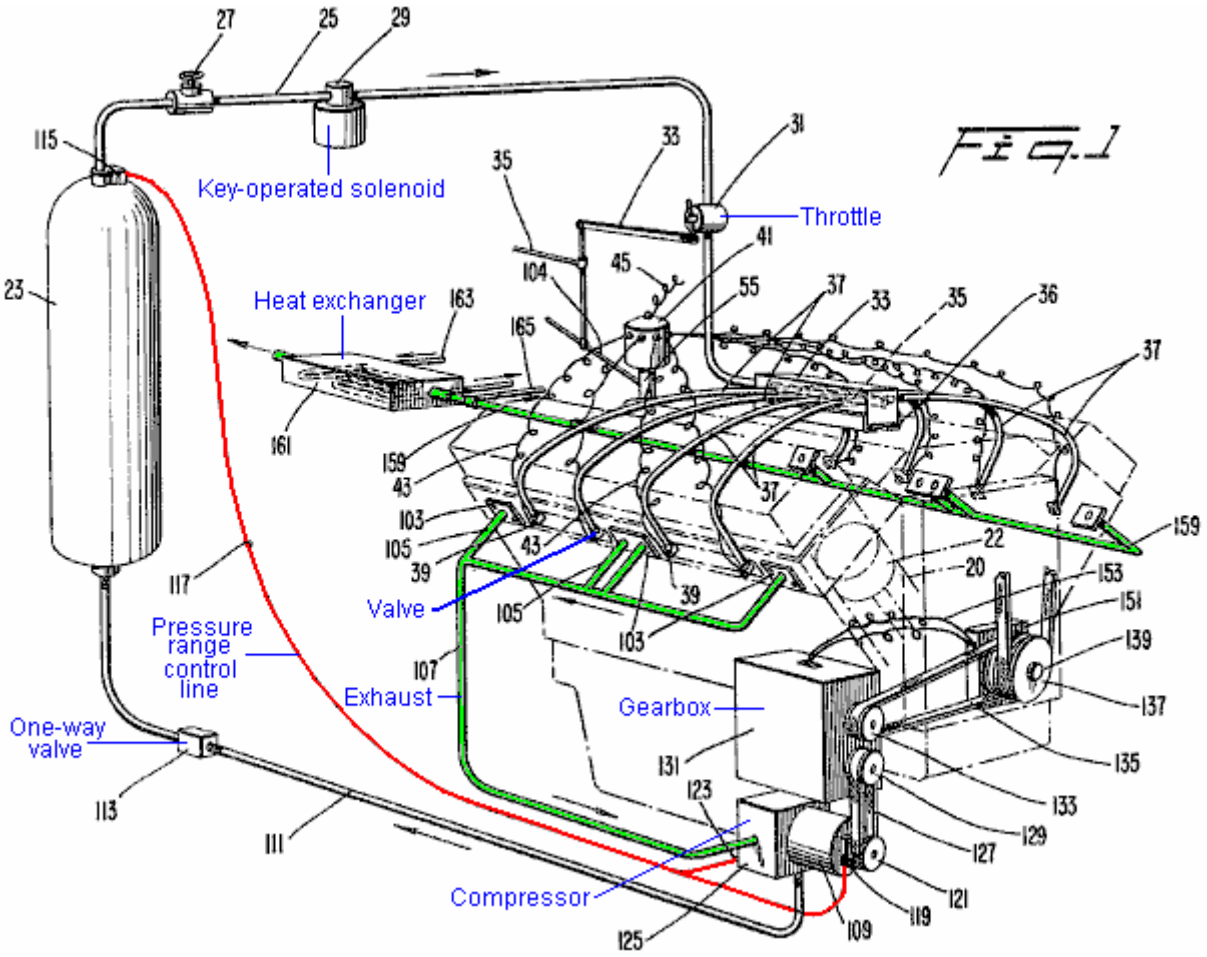
The Rogers motor shown here makes no claims to spectacular operation, but in spite of that, Leroy did admit in an interview that this motor does indeed have a greater output than the applied input, provided that the motor is not left just ticking over. This motor is like the US patent 3,744,252 “Closed Motive Power System Utilising Compressed Fluids” by Eber Van Valkinburg shown below. However, the Rogers patent shown here has the distinct advantage that it uses off-the-shelf motors and readily available hardware and there is nothing really exotic or difficult about the Rogers engine that a person couldn’t get from a valve supplier or get a metal fabrication company to construct.

However, while Leroy did state that his design was self-sustaining when going over 30 miles per hour, a key design feature is his very high performance compressor unit which he later patented as shown below. Present day vehicle engines are under-geared and run at fairly low revs. These same engines operate much more efficiently at higher revs, if they are given different gearing. With the Rogers motor, the air contained in the high-pressure tank is sufficient to drive the pistons up and down. Air can be pumped back into the high-pressure tank by a compressor which has a much higher gearing and much lower capacity per piston stroke. The expanded air exiting from the engine is at much lower temperature than the surrounding air and if it captured in a buffer tank and used as the input of the compressor, then recharging the air tank is more efficient, provided that the tank absorbs heat from the surrounding environment, raising it’s temperature inside the tank and so giving an extra boost to the tank pressure, over and above the compression provided by the compressor.

One really nice feature of Leroy’s design is that he envisages it as being an adaption of an ordinary vehicle engine and he provides a considerable amount of practical detail as to how the adaption can be carried out.

Using a RotoVerter (as described in Chapter 2) to drive a compressor would lower the power requirements of the compressor drive to the extent that a motor adaption of this kind should be self-sustaining. The RotoVerter provides a major energy gain in its own right and is particularly suited to driving mechanical loads such as the compressor and it particularly ‘likes’ constant-load applications such as a compressor.

The adapted engine shown in the patent is like this:



This patent shows how the practical details of running an engine on compressed air can be dealt with. What it does not show is background details of the actual energy flows and the effects of compressing air and then letting it expand. These things are not normally encountered in our daily lives and so we do not have an immediate intuitive feel for how systems like these will operate. Take the effects of expansion. While it is quite well known that letting a compressed gas expand causes cooling, the practical effect is seldom realised.

Leroy's compressor patent is shown here:

United States Patent 4,693,669 Supercharger for automobile engines

Inventor: Rogers Sr., Leroy K. (Rte. 13, P.O. Box 815-DD, Briarcliff Rd., Fort Myers, FL, 33908)
Publication Date: 15th September 1987

Abstract:
A supercharger for delivering supercharged air to an engine, comprising a shrouded axial compressor, a radial compressor which is located downstream of the axial compressor and a housing. The housing is comprised of four sections, including a section which is a highly converging, 'frustoconical' transition duct which favourably directs the discharge of the axial compressor to the inlet of the radial compressor and a hollow, highly convergent, exhaust cone section immediately downstream of the radial compressor which converges into the exhaust port of the supercharger. An annular flow deflector is provided for directing the discharge of the radial compressor into the exhaust cone.

Description:
Superchargers impart additional pressure to the air or the air/fuel mixture of an engine so that the cylinders receive a greater weight per unit volume of air or air/fuel mixture than would otherwise be supplied. As a result, the volumetric efficiency and power output of the engine are improved.

According to prior practices, superchargers generally comprise a single air-blower which forces air or an air/fuel mixture into the cylinders of an engine. Typically, the air-blower is driven by a gear train which is connected to the crankshaft of the engine with a gear ratio of about 6 to 1. These prior types of superchargers have been used extensively in racing engines and radial aircraft engines. However, by reason of their high operating speeds and their gear trains, these superchargers have been considered too complicated, too heavy and too costly for use with mass production engines such as are found in cars and trucks.

Recently, some car manufacturers have been offering turbocharged engines which expand to exhaust gases of the engine through a turbine to drive a centrifugal compressor. Although turbochargers are advantageous in that the turbine can deliver large amounts of power to the compressor, their extreme operating speeds require special bearings, lubrication and maintenance. In addition, turbochargers require special ducting, such as by-pass arrangements, which only add to their cost and maintenance requirements. Consequently, turbochargers are only offered as expensive options in cars.

Further, there is current interest in a new type of car engine which operates from tanks of compressed gas to effect reciprocation of its pistons. An example of such an engine can be found in the U.S. Pat. No. 4,292,804 issued to the same inventor of the present invention. In the referenced patent, at least a portion of the partially expanded exhaust gas from the cylinders is directed to a compressor where it is recompressed and then returned to the storage tanks from whence it originally came. It would be desirable that at least some, if not all of the aforementioned recompression of the exhaust gas could be achieved with a belt-driven, rotary supercharger that is easily manufactured and maintained, yet is capable of providing ample recompression.

Objects of the Invention:
Accordingly, an object of the present invention is to provide a supercharger suitable for improving the performance of engines of cars, helicopters or the like, which supercharger is inexpensive to produce and easy to maintain.

It is another object of the present invention to provide a supercharger which provides sufficient boost without resort to extreme operating speeds and accordingly avoids the costly complications associated with high speed operation.
It is yet another object of the present invention to provide a relatively compact and lightweight supercharger which is inexpensive to manufacture and maintain.

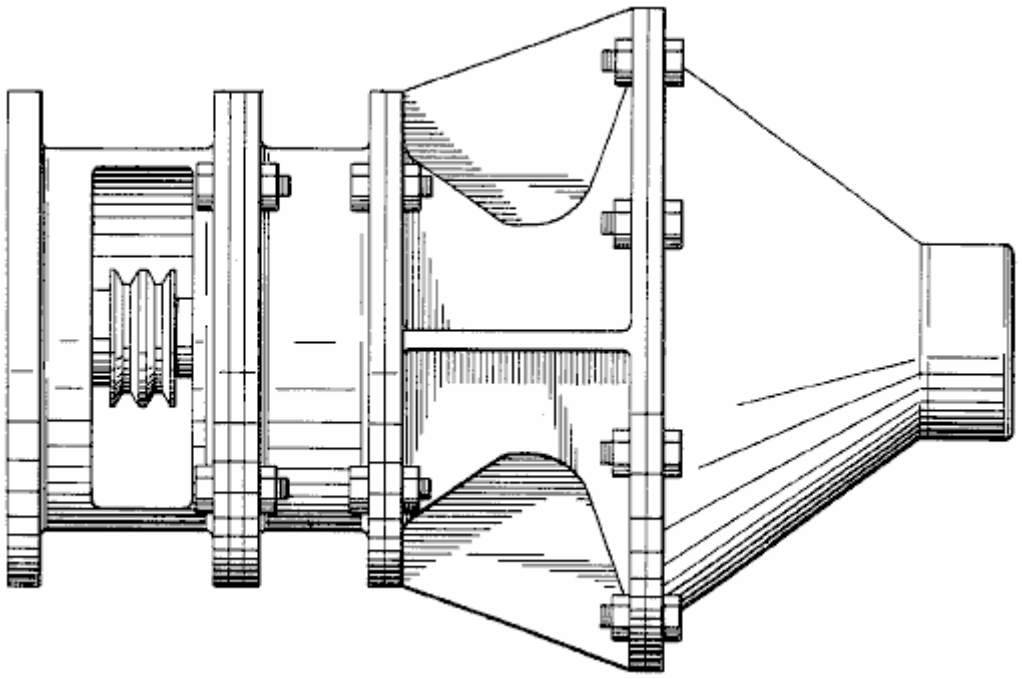
Another object of the present invention is to provide a belt-driven supercharger having a design which provides supercharging compression at relatively low operating speeds.

It is still another object of the present invention to provide a supercharger which can be quite readily disassembled and reassembled for purposes of low cost maintenance and repair.

Still another object of the present invention is to provide a supercharger which can be constructed from mass producible parts to thereby reduce the cost of its manufacture.

It is still another object of the present invention to provide a belt-driven supercharger which provides supercharging compression without resort to a larger number of compressor stages.

Yet another object of the present invention is to provide a rotary supercharger for a gas operated engine, which supercharger is easily manufactured and maintained, yet capable of providing ample recompression of the recirculating drive fluid.



Summary of the Invention:
These and other objects are achieved by the present invention which provides a supercharger comprising a housing having an inlet and an outlet, a shrouded axial compressor and a radial compressor rotatably mounted within the housing, a highly convergent shallow, frustoconical transition duct for favourably directing the discharge of the axial compressor to the inlet of the radial compressor.

In accordance with a further aspect of the invention, the above-described supercharger further comprises an exhaust cone at a location downstream of the radial compressor and a flow deflector for directing the discharge of the radial compressor to the exhaust cone.

In the preferred embodiment, the housing itself comprises four sections: a cylindrical front housing section which defines an axially directed inlet; a second, cylindrical ducting section enclosing the axial compressor; a rear housing section defining the transition duct as well as the inlet and casing for the radial compressor; and the exhaust cone section which defines at its terminus the outlet of the housing. For driving the compressor shaft, a double-tracked pulley wheel is secured to the forward end of the common shaft, which pulley wheel is adapted to receive one or more drive belts from the crank-shaft wheel of the engine. A lateral opening in the front housing section accommodates the connection with the drive belts.

With the disclosed arrangement, compression can be achieved for supercharging purposes without resort to a large number of compressor stages or high operating speeds. Additionally, the design of the disclosed supercharger avoids the need for guide vanes between the axial compressor and the radial compressor. The exhaust cone section also favourably avoids the build-up of back pressure against the radial compressor. The design is also very simple and therefore inexpensive to manufacture and maintain.

Other objects, advantages and novel features of the present invention will become apparent from the following detailed description of the invention when considered in conjunction with the accompanying drawing.

Brief description of the Drawings:
A preferred embodiment of the present invention is described in greater detail with reference to the accompanying drawing wherein like elements bear like reference numerals, and where:

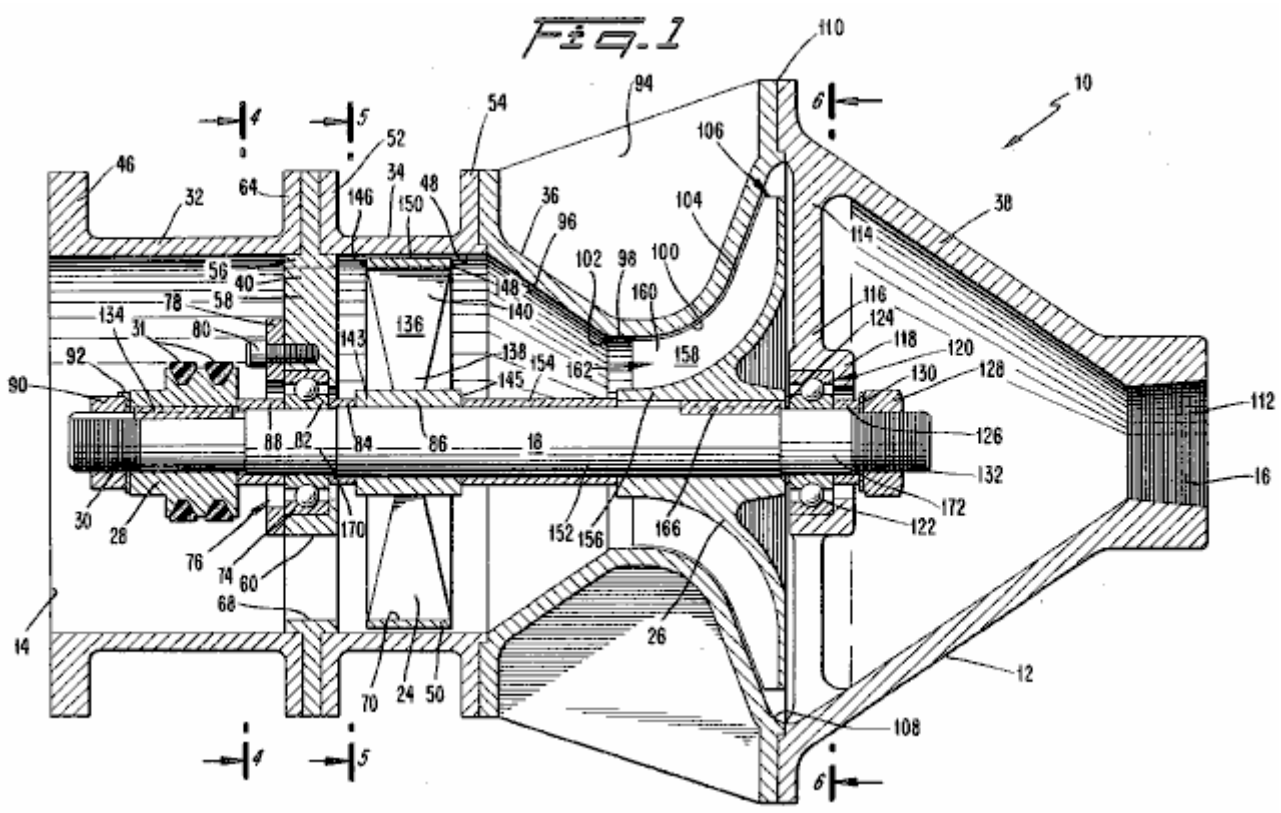


Fig.1 is a cross-sectional side view of a supercharger constructed in accordance with the preferred embodiment of the present invention;

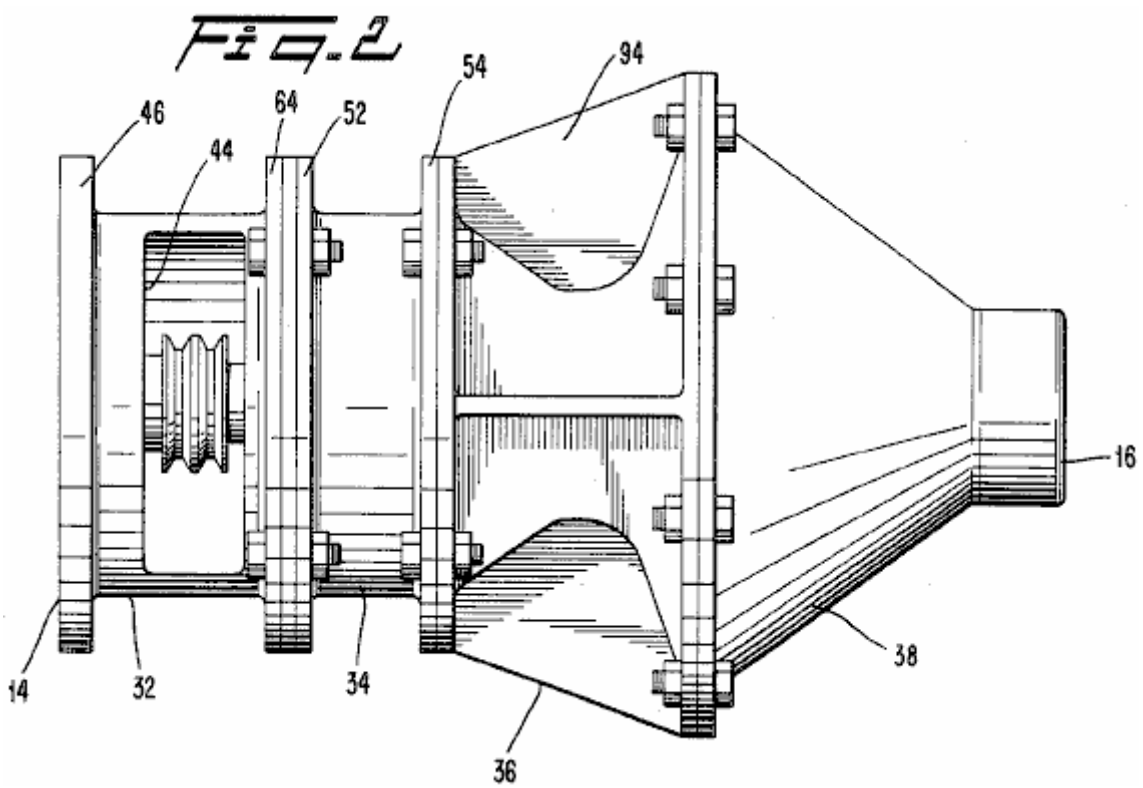


Fig.2 is a side view of the supercharger of Fig.1;

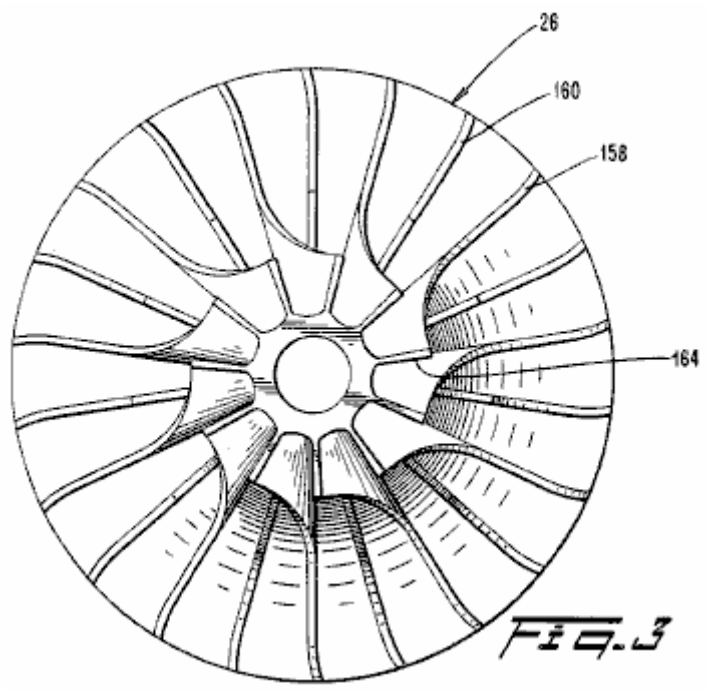


Fig.3 is a frontal view of the impeller of the supercharger of Fig.1;

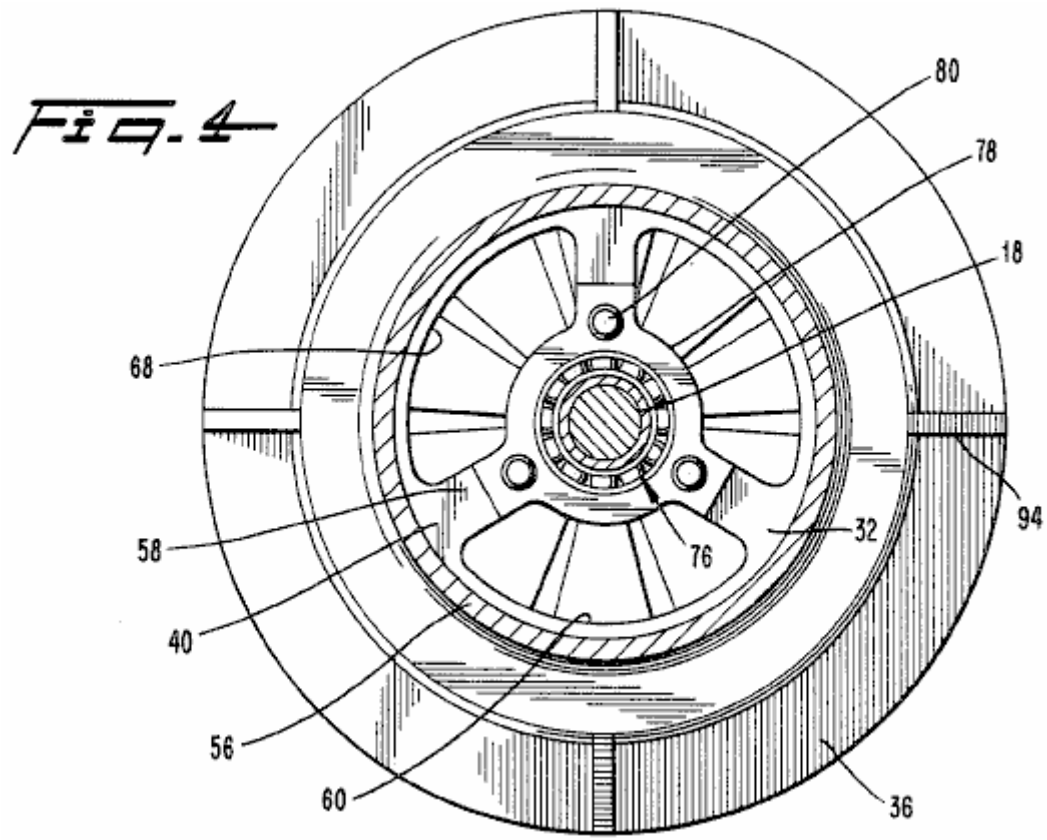


Fig.4 is a cross-sectional view taken along line **4--4** of Fig.1;

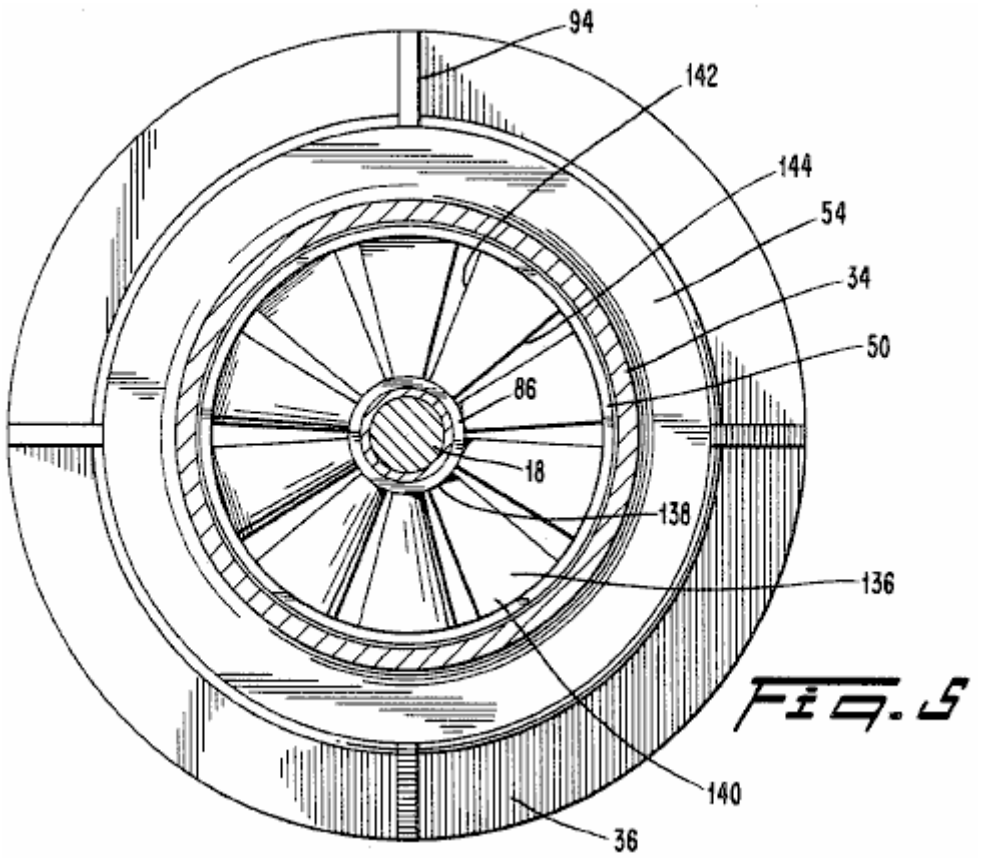


Fig. 5

Fig.5 is a cross-sectional view taken along line **5--5** in Fig.1;

Fig. 6

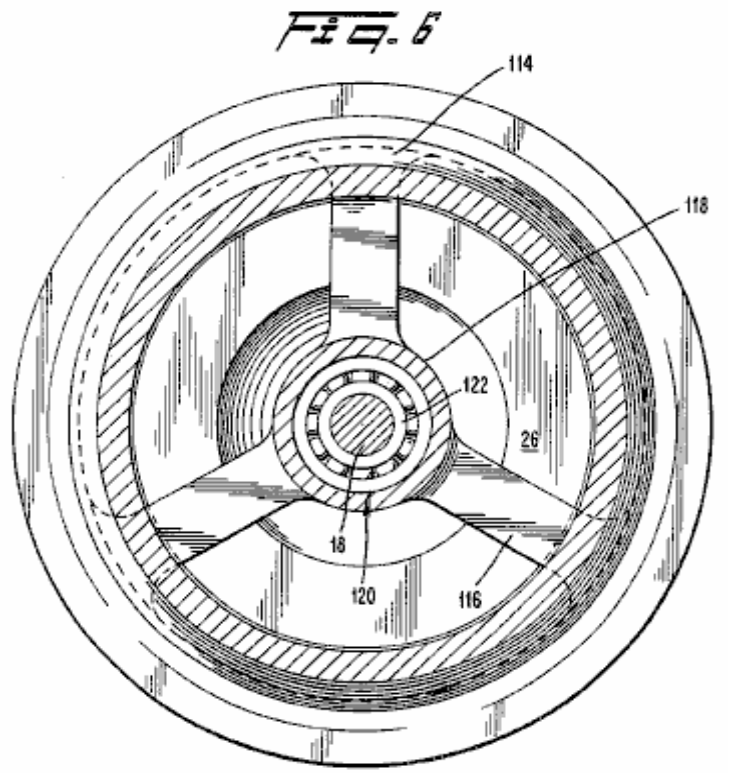


Fig.6 is a cross-sectional view taken along line **6--6** in Fig.1;

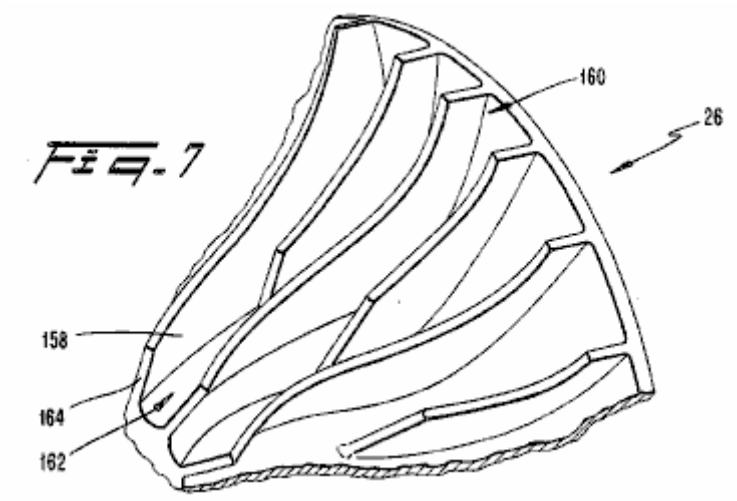


Fig.7 is a perspective view of a segment of the impeller of the supercharger of Fig.1; and

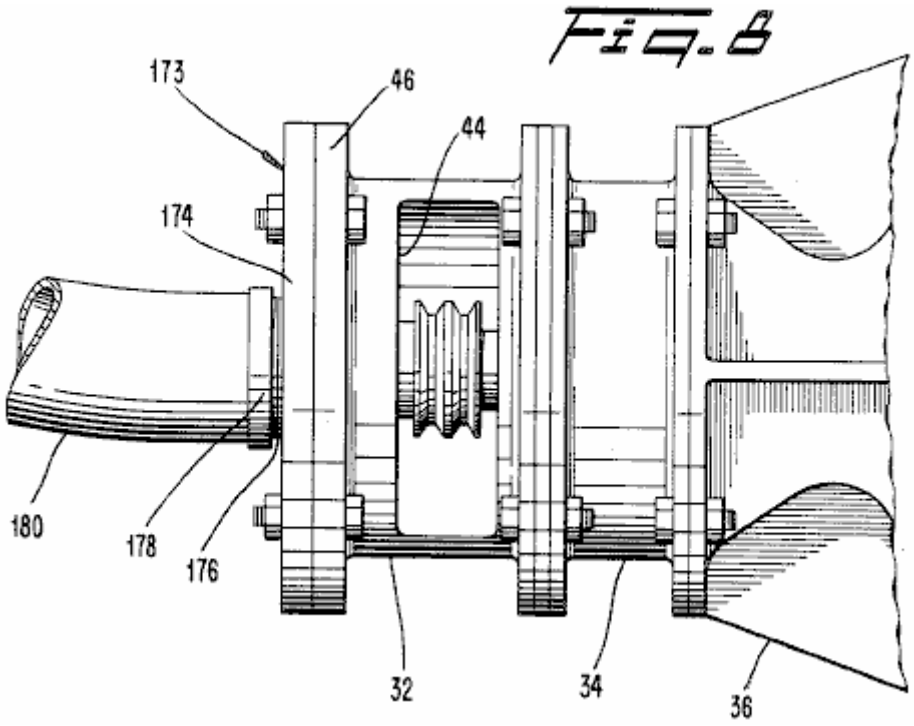
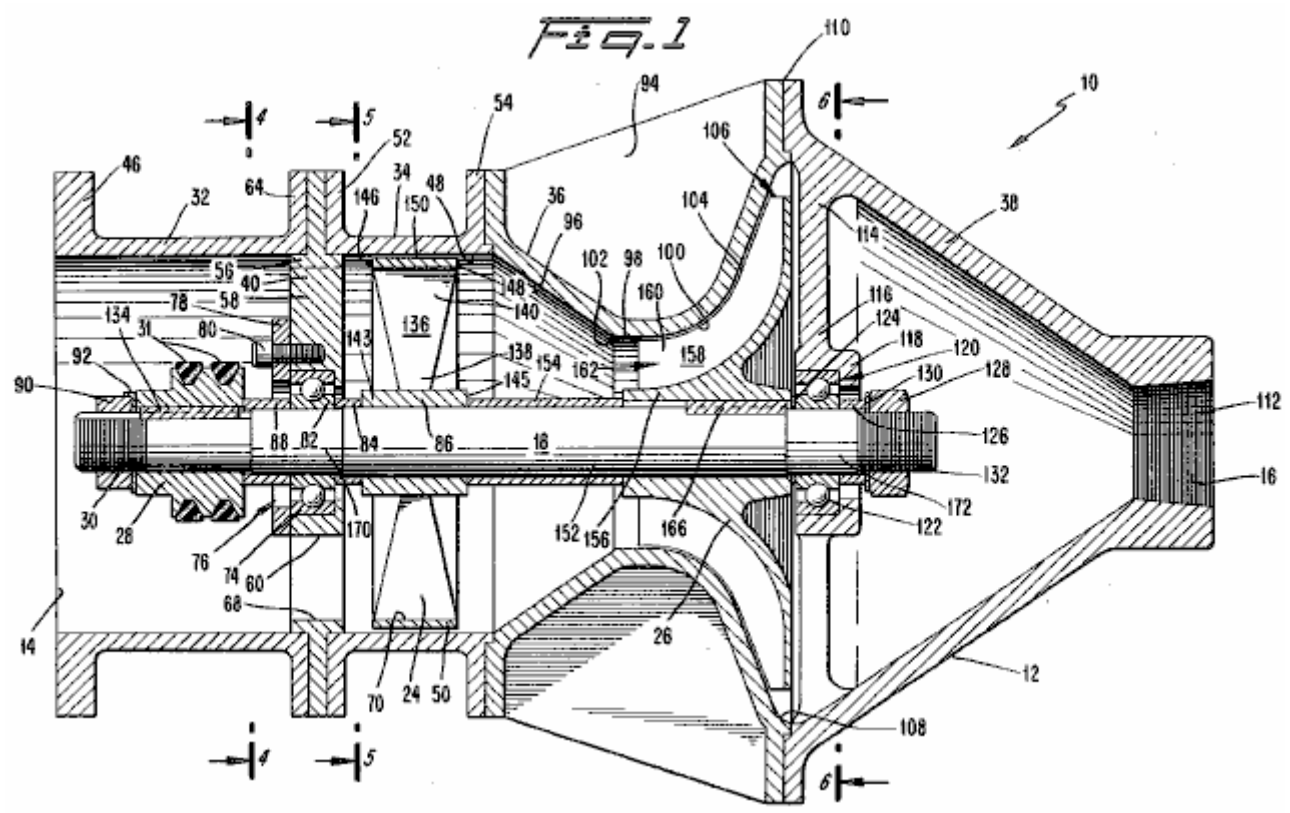


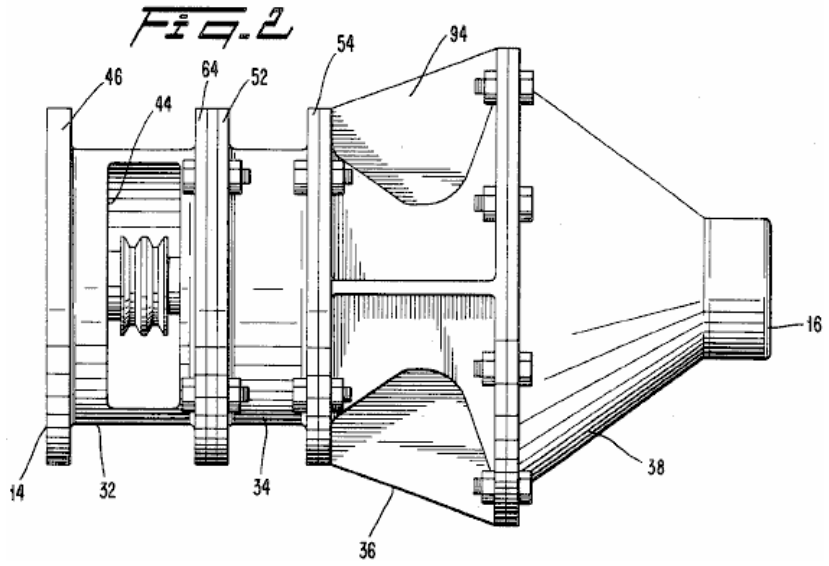
Fig.8 is a partial side view of the supercharger of Fig.1 with an adaptor .

Detailed Description of the Preferred Embodiment:



Referring to **Fig.1** and **Fig.2**, a supercharger **10** is provided for supplying supercharged air to a car engine or the like, so that the engine receives a greater weight per unit volume of air or a fuel/air mixture than would be otherwise supplied. In accordance with a preferred embodiment of the present invention, the supercharger **10** comprises a housing **12** having an axially directed inlet **14** for receiving ambient air and an axially directed outlet **16** for delivering supercharged air to the intake of the car engine. Rotatably mounted within the housing **12** is a shaft **18** on which are secured an axial compressor **24** and a radial compressor **26**, which is positioned downstream of the axial compressor. A pulley wheel **28** is secured to a forward end **30** of the shaft for receiving drive belts **31**, which drive belts connect the shaft **18** to a pulley wheel on the crankshaft of the engine (not shown). The drive belts **31** deliver torque to the shaft **18** as required for driving the compressors **24** and **26** of the supercharger **10**.

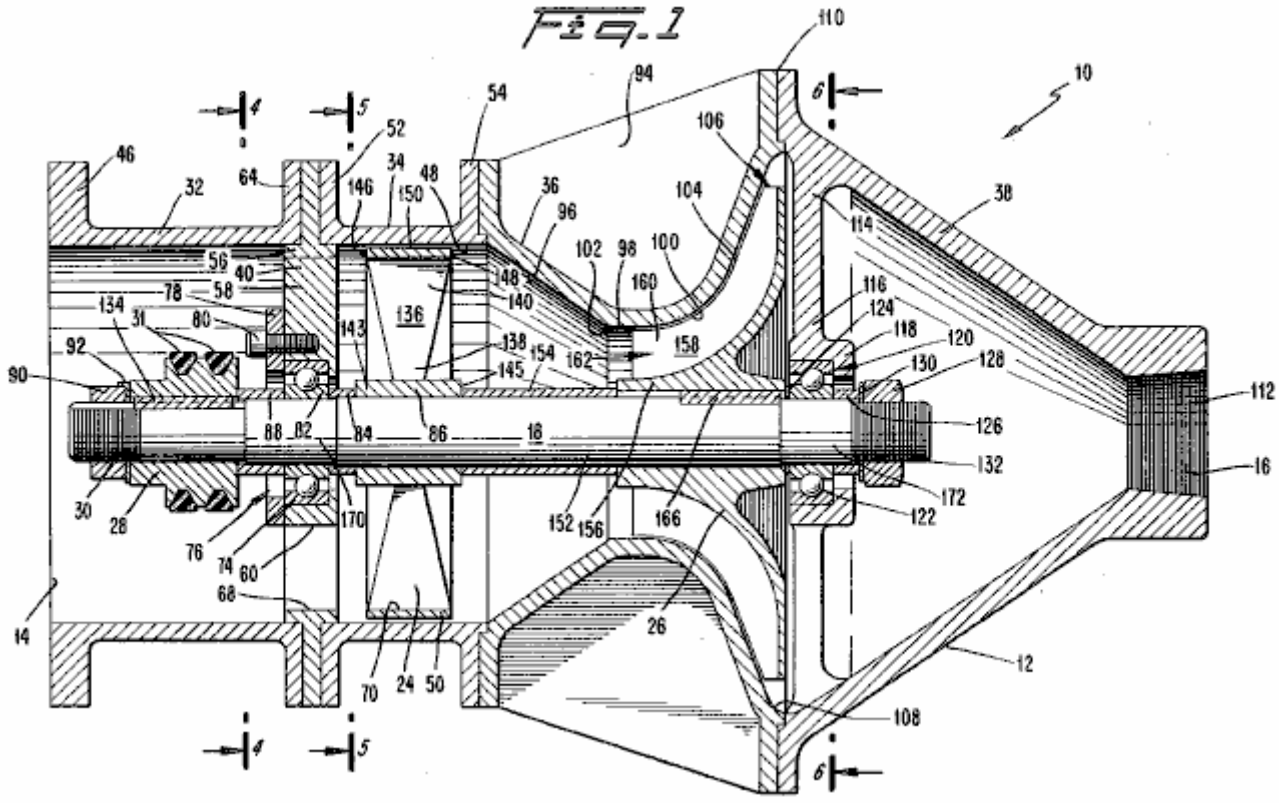
Housing **12** is constructed from four sections which are preferably bolted together at flanged connections in an end-to-end relationship. These sections include a front housing section **32**, an axial compressor duct section **34**, a rear housing section **36** and an exhaust cone section **38**. The shaft **18** extends along the longitudinal axis of the housing **12**.



The front housing section **32** is a hollow cylinder which extends forward of a front bearing support **40**. The front housing section **32** encloses the forward end **30** of the shaft **18** and the associated pulley wheel **28**. At its forward end, the front housing section **32** defines the inlet **14** for receiving air from an external source (not shown).

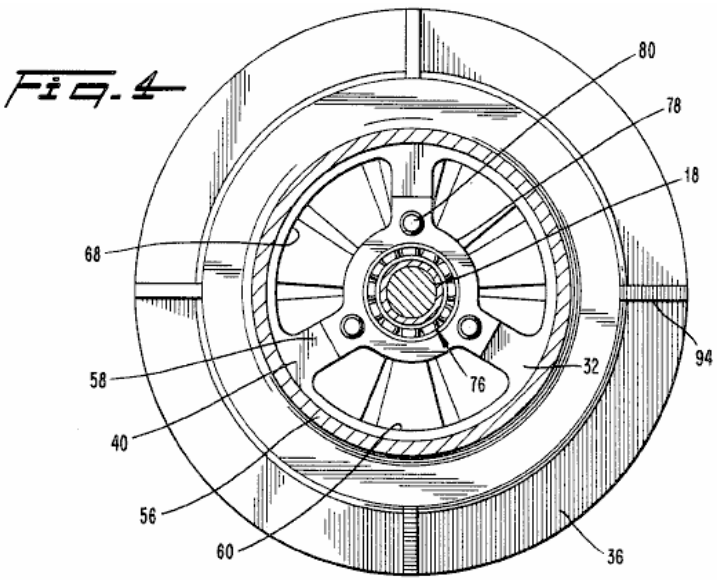
Referring particularly to **Fig.2**, the front housing section **32** includes a lateral opening **44** on one side in order to accommodate the connection of the drive belts **31** to the pulley wheel **28**. The front housing section **32** also includes a forward flange **46** for accommodating the connection of air filters, carburettors, air scoops or the like upstream of the supercharger **10** according to the particular engine layout.

It is to be understood that in the usual engine layout, the supercharger **10** receives air or a fuel/air mixture from an external source through its inlet **14**, compresses the air or fuel/air mixture and then delivers it to the intake of the engine.



Referring again to **Fig.1**, the pulley wheel **28** is interference-fitted upon the forward end **30** of shaft **18** and a key **134** is used to lock the pulley wheel **28** in place. The pulley wheel **28** is preferably a double-track design which is suitable for the attachment of twin drive belts, although a single-belt type pulley wheel would be adequate. The pulley wheel **28** is preferably sized so that the ratio of it's diameter with respect to the diameter of the drive wheel of the engine's crankshaft provides an effective gearing ratio in the range of approximately two and one-half to four and one-half. Thus at idle, when the car engine is running at approximately 700 rpm, the supercharger **10** is running at approximately 2,400 rpm, and at cruise, when the engine is running around 2,500 rpm, the supercharger **10** is preferably turning over in the range of 6,000 to 8,000 rpm. It is to be noted that although the diameter of the pulley wheel **28** may be substantially reduced in order to achieve a desired gearing ratio, the double-track wheel **28** presents a sufficient sum total of surface area to avoid slippage of the belts **31**.

The next adjacent section of housing **12** is the axial compressor duct **34** which is a short cylinder coaxially disposed about the axial compressor **24**. Preferably, the axial compressor duct **34** is constructed from cast aluminium, with the interior surfaces **48** machined to assure uniform clearance between the duct **34** and shroud **50** of the axial compressor **24**. As with other sections of the housing **12**, the axial compressor duct **34** is provided with flanges **52** and **54** for effecting connection to the adjacent housing sections. The axial compressor duct **34** guides air delivered from the front housing section **32** towards the axial compressor **24**.



Referring now to **Fig.1** and **Fig.4**, a front bearing support **40** is placed between the front housing section **32** and the axial compressor duct **34**. The front bearing support **40** includes an outer annulus **56** and three radial arms **58**. Between these arms are defined passages **60** for allowing air to pass through the bearing support **40**. The outer annulus **54** is secured by bolts connecting a rear flange **64** of the front housing section **32** and the flange **52** of the axial compressor duct **34**. By this arrangement, the front bearing support **40** is rigidly secured to the housing **12** so that loads and shocks to the shaft **18** can be transferred through the front bearing support **40** to the housing **12**.

In the preferred embodiment, the outer annulus **56** of the bearing support **40** extends into the region of the inlet **14** of the front housing section **32** in such a way that it's inner rim **68** coincides with the inner rim **70** of the shroud **50** of the axial compressor **24**. In this way, the outer annulus **56** contributes to the guiding of the flow of air toward the axial compressor **24**.

An outer raceway **74** of the front roller bearing assembly **76** is secured between the front bearing support **40** and a bearing retainer plate **78**, which is secured by the removable bolts **80**. In this preferred embodiment, the front bearing assembly **76** is of the sealed, high speed type. A suitable commercially available bearing assembly is marketed under the reference: model Fafnir 405KDD. Preferably, a lower raceway **82** of the front bearing assembly **76** is secured to the shaft **18** with an interference fit. A spacer **84** is provided on one side of the lower raceway **82**, which spacer **84** also abuts a hub **86** of the axial compressor **24** in order to position the axial compressor **24** at a predetermined distance downstream of the bearing support **40**. Similarly, a spacer **88** is provided on the other side of the lower raceway **84**, and it abuts the pulley wheel **28** so as to space apart the pulley wheel **28** from the front bearing support **40** to ensure that there is sufficient clearance between them.

It should be appreciated that the bearing retainer plate **78** allows ready access to the front bearing assembly **76** for purposes of maintenance or repair. To service the front bearing assembly **76**, a nut **90** and lock-washer **92** on the forward end **30** of the shaft **18** are loosened and removed together with the pulley wheel **28** and the spacer **88**. Then bolts **42** and the bearing retainer plate **76** are removed, leaving the whole bearing assembly **76** exposed for servicing and/or removal.

The rear housing section **36** is connected by bolts to the downstream end of the axial compressor duct **34**. Preferably, the rear housing section **36** is constructed from a single section of cast aluminium and includes external longitudinal ribs **94** for enhancing the structural rigidity of the rear housing section **34**. The walls of the rear housing section **36** define three elements of the supercharger **10**: a highly conical transition duct **96** which favourably directs the output of the axial compressor to an inlet **98** of the radial compressor **26**; the inlet **98** of the radial compressor **26**, itself; and a casing **100** for the radial compressor **26**.

The transition duct **96** is a hollow, frustoconical portion having a half-apex angle (from the generatrix to the axis of symmetry) of approximately 35°. The angle is selected such that the inlet to the radial compressor **26** is as close as possible to the outlet of the axial compressor, without causing undue back-pressure. In the preferred embodiment, the transition duct **96** begins a short distance downstream of the axial compressor **24** and ends at the beginning of the inlet **98** of the radial compressor **26**. The highly conical shape of the

transition duct **96** is believed to roll-in the higher volume of air being discharged from the more radially outward portions of the axial compressor **24**. This rolling-in action is believed to promote a favourable flow regime at the inlet **98** of the radial compressor **26** so that there is no need for inlet guide vanes for the radial compressor **26**. It is also believed that the highly conical shape of the transition duct **96** affects upstream flow conditions at the axial compressor **24** in such a way that it's performance is improved. It has also been found that there is no need for a stator (or exit guide vane) for the axial compressor **24**.

In essence, it is believed that the transition duct **96** performs the functions of the exit vanes of axial compressors and inlet guide vanes of radial compressors, but without the pressure losses commonly associated with them. Avoiding these pressure losses and the expected improvement in the performance of the axial compressor, allows the supercharger **10** to impart a higher overall pressure ratio than would otherwise be achieved without the transition duct **96**. As a result, adequate compression is achieved at moderate operating speeds without resort to a bank of several axial compressors. It should be understood however, that when connecting the supercharger **10** to a relatively slowly reciprocating diesel or a very large engine, it may be desirable to include two or more axial compressors in order to boost the supercharger's overall pressure ratio. In such cases, the present invention would then include the placement of a transition duct downstream of at least the last axial compressor.

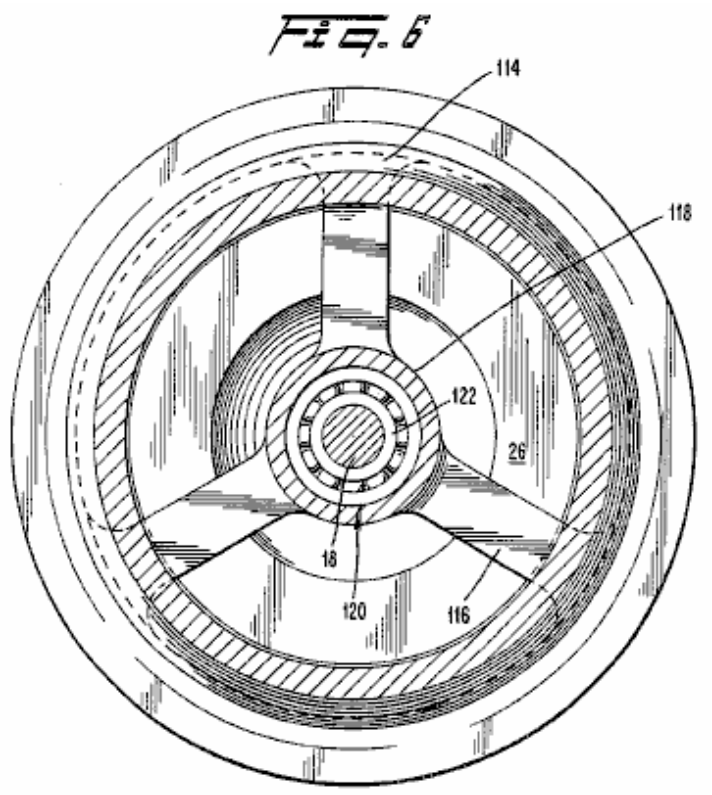
At the inlet **98** of the radial compressor **26**, the walls of the rear housing **36** are cylindrical and coaxially positioned around shaft **18**. It should be noted that in the preferred embodiment, the surface transition **102** from the transition duct **96** to the inlet **98** is rounded-off.

The casing portion **100** of the rear housing section **36** closely follows the contour defined by blade edges **104** of the radial compressor **26** in a close, substantially sealing manner as is well known in the art of radial compressors. The casing portion **100** of the rear housing section **78** channels air between the rotating blades of the radial compressor **26** so that the blades can impart work to the passing air. The casing portion **100** also defines a discharge outlet **106** for the radial compressor **26**.

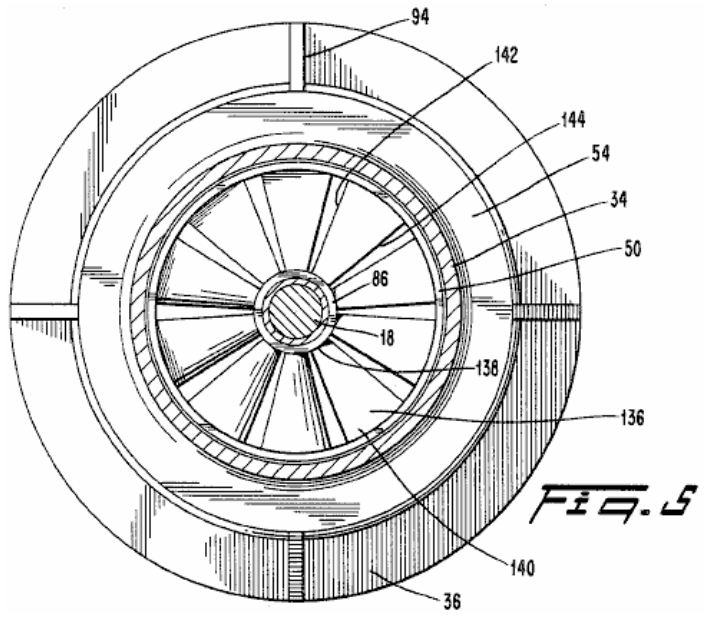
Just beyond the discharge outlet **106** of the radial compressor **26**, the interior surfaces of the rear housing section **36** begin to curve immediately inwardly to provide a transition into the next adjacent section of the housing **12**, the exhaust cone **38**. In this fashion, the interior surfaces at the rear-most portion of rear housing section **36** and those of the forward portion of the exhaust cone **92** define internally a flow deflector **108**. In the preferred embodiment, the flow deflector **108** is closely and concentrically positioned around outlet **106** of the radial compressor **26** so that the air being discharged from the radial compressor **26** does not have the opportunity to diffuse significantly prior to its arrival at the annular flow deflector **108**. The annular flow deflector **108** directs the output of the radial compressor **26** into the exhaust cone **38** by providing a smooth surface transition from the interior of rear housing section **36** to the interior of the exhaust cone **38**.

The exhaust cone **38** is a highly convergent, hollow, conical section placed immediately downstream of the radial compressor **26** for receiving the output of the radial compressor **26** from the annular flow deflector **108**. In the preferred embodiment, the exhaust cone **38** is a single section of cast aluminium which is joined to the downstream end of the rear housing section **36** at a flanged joint **110**. Preferably, the exhaust cone **92** converges according to a half-apex angle of approximately 35° and defines the exhaust port **16** at its terminus. A threaded section **112** at the exhaust port **16** allows the attachment of the appropriate external ducting (not shown) leading to the intake of the engine.

During operation of the supercharger **10**, the space enclosed by the exhaust cone **92** prevents the build up of an elevated back pressure which might otherwise arise and detract from the operation and efficiency of the radial compressor **26**. The enclosed space of the exhaust cone **92** is also of sufficient volume to absorb pulses and to average out unsteady flow conditions so to promote a smooth and continuous output from the supercharger **10**.



Referring now to **Fig.1** and **Fig.6**, the exhaust cone **38** includes a rear bearing support **114** which comprises members **116** which extend radially inwardly from the outer walls of the exhaust cone **38**. At a radial inward location close to the shaft **18**, the members **116** converge to form a cupped annulus which serves as a housing **118** for the rear bearing assembly **120**. The housing **118** is open towards the rear face of the radial compressor **24** to facilitate disassembly of the supercharger **10**. The rear bearing assembly **120** is the same type and size as the front bearing assembly **76**. The inner race **122** of the bearing assembly **120** is set in place on the shaft **18** by spacers **124** and **126** in conjunction with a nut **128** and washer **130** on the rearward end **132** of the shaft **18**. In this preferred embodiment, the members **116** are formed to be integral with the walls of the exhaust cone **38**.

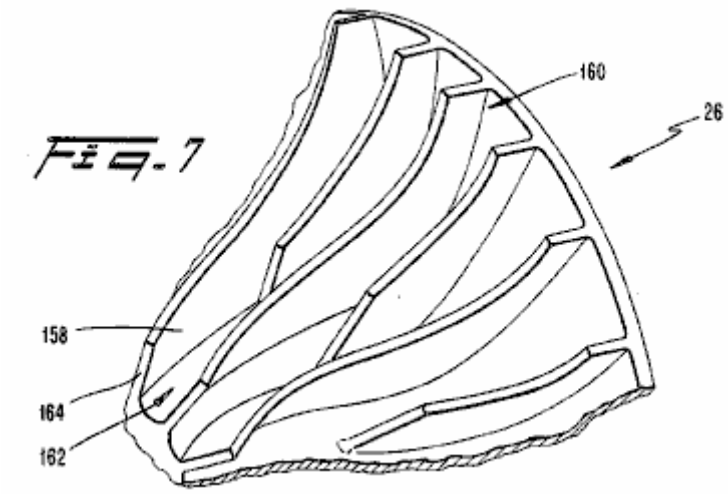


Referring to **Fig.1** and **Fig.5**, upon rotation, the axial compressor **24** draws air through the inlet **14** and imparts an initial amount compression to the air as it forces the air into the transition duct **96** of the rear housing section **36**. In the preferred embodiment, the axial compressor **24** comprises a hub **86**, the shroud **50** and a series of ten (10) equally spaced, radial blades **136**. Ideally, each blade **136** increases in cord from a root **138** to a tip **140** and includes a trailing edge **142** and a leading edge **144**, where these edges are both

slightly curved. The blades gradually increase in pitch from approximately 12° at the root **138** to approximately 36° at the tips **140**. However, the particular values of pitch and other geometrical aspects of the blades **136** might be varied in accordance with different operating speeds or other parameters as would be apparent to one skilled in the pertinent art and familiar with this disclosure.

The axial compressor **24** is preferably constructed from a single, cast aluminium section with the faces **143** and **145** of the hub **86** being machined for purposes of achieving accurate, axial positioning of the axial compressor **24** on the shaft **18** relative to the housing **12**. The faces **146** and **148** of the shroud **72** are also machined flat. Additionally, the outer periphery **150** of the shroud is machined to assure uniform clearance between the shroud and the adjacent interior surfaces **48** of the axial compressor duct **34**. Preferably, the axial compressor **24** is secured to the shaft **18** by an interference-fit on to a stepped portion **152** of the shaft **18**. The spacers **84** and **154** axially position the axial compressor **24** relative to the front bearing support **40** and the radial compressor **26**, respectively.

Dynamic balance test machines of the conventional type may be used to test the balance of the axial compressor **24** prior to its installation. If an imbalance is detected, material can be removed at the outer periphery **150** of the shroud **50** so as to achieve proper balance.



Referring now to **Fig.1**, **Fig.3**, and **Fig.7**, the radial compressor **26** is constructed from a single section of cast aluminium and includes a hub **156** and curved blades **158**. Interposed between each pair of blades **158** are a second set of blades **160** which terminate short of the intake **162** of the radial compressor **26** so that the intake **162** is not crowded by both sets of blades. Accordingly, the radial compressor **26** features both a large total number of blades and an intake of relatively small diameter, ad these features enhance the performance of the compressor **26**. In the region of the intake **162**, the blades **158** present leading edges **164** and undergo a twist into the direction of rotation so as to prevent a favourable angle of attack at the intake **162**.

Preferably, the radial compressor **26** is positioned upon the stepped section **128** of the shaft **18** with an interference-fit and locked against rotational slippage by a key **166**. The spacer **124** assures clearance between the rear face of the radial compressor **26** and the rear bearing assembly **120**.

The shaft **18** is constructed from a hardened steel and is threaded at both ends **30** and **132** to receive nuts **90** and **128**, respectively. In addition to the central stepped portion **152**, which receives the compressors **24** and **26**, the shaft **18** also features stepped portions **170** and **172** for receiving the front and rear bearing assemblies **76** and **120**, respectively. The stepped arrangement of the shaft **18** facilitates assembly and disassembly in that the stepped portion **152** of the greatest diameter is centrally located on the shaft **18** and all the stepped portions are greater than the diameter of the threading at ends **30** and **132**.

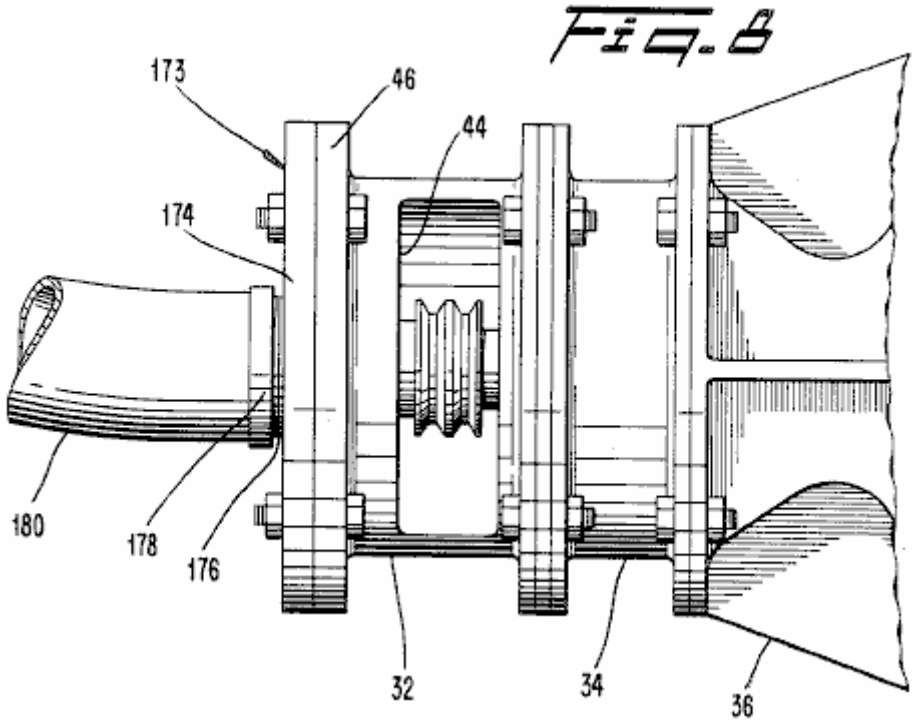
Please note that the bearing supports **40** and **114** are in a fixed position relative to the housing **12** and that the compressors **24** and **26** are held in position between the bearing supports **22** and **40** by spacers **84**, **124** and **154**, which have predetermined lengths. Consequently, the placement of the compressors **24** and **26** relative to the longitudinal axis of the housing **12** is fixed by the spacers and not by the axial location of the shaft **18** relative to the housing **12**. Please also note that the stepped portions **152**, **170** and **172** of the shaft **18** are each provided with extra lengths so that the respective components (the bearing assemblies and compressors) can each be situated over a relatively wide range of locations in the respective stepped

portions. Thus, the shaft **18** need not be positioned accurately along the longitudinal axis of the housing 12 in order to achieve proper assembly of the supercharger **10**. For instance, if nuts **90** and **128** had been tightened differently than they appear in **Fig.1**, then shaft **18** might have been displaced slightly in the axial direction from where it is shown in **Fig.1**. However, the relative positioning of the various components on the shaft **18**, i.e., the pulley wheel **28**, the compressors **24** and **26** and the bearing assemblies **76** and **120**, would have remained the same relative to themselves and the housing **12**. This feature eases the process of manufacture and accordingly, reduces costs. It also reduces the amount of labour required for reassembly after repair.

In operation, the supercharger **10** is suitably connected at it's outlet **16** to an intake of a car engine, with the drive belts **31** from the crankshaft of the engine being attached to the pulley wheel **28** of the supercharger **10**. Then, as the engine is operated, torque is transferred by the drive belts **31** to the pulley wheel **28** for driving the compressors **24** and **26**. Upon rotation, the axial compressor **24** draws air through the inlet **14**, imparts an initial amount of compression to the air and discharges it into the transition duct **96** with a swirl. By reason of it's design, the axial compressor **24** is believed to move a greater volume of air in the region of its blade tips **140** than at it's more radially inward locations. Accordingly, there is a greater of mass of air situated at the outer annular region behind the axial compressor **24** than at the inner annular region. As the discharge from the axial compressor 24 is caused to leave the axial compressor duct **34**, the highly convergent, transition duct **96** is believed to cause the outer annulus of air which is discharged from the axial compressor **24** to roll-in. This action is believed to have two favourable results. First, the roll-in action causes a flow regime to be established at the inlet **98** of the radial compressor **26** such that the need for a guide vane is wholly avoided. Secondly, and of equal importance, the rolling-in action, in conjunction with the large volume of space enclosed by the transition duct **96**, is believed to affect the performance of the axial compressor **24** favourably, so that a higher pressure ratio is obtained from it.

Since the overall pressure ratio of the supercharger **10** is the product of the pressure ratios of the two compressors, it can be seen that the increase in performance of the axial compressor **24** results in a corresponding improvement in overall performance of the supercharger. It should also be noted that the deletion of inlet guide vanes for the radial compressor **26** and of exit vanes for the axial compressor **24** greatly simplifies the design of the rear housing section **36** and therefore provides savings in costs of manufacture. It also avoids the pressure losses associated with such guide vanes, which are often quite significant.

Upon leaving the transition duct **96**, the pre-swirled flow of air enters the inlet **98** of the radial compressor **26** and then into the compressor **26** itself. In passing through the radial compressor **26**, the air is turned and whirled such that the airflow is centrifugally discharged with a substantial radial velocity component, whereupon the resultant flow is abruptly turned by the annular flow deflector **108** and caused to enter the exhaust cone **38**. As previously explained, the large volume of space enclosed by the exhaust cone **38** induces flow conditions behind the radial compressor **26** such that elevated back pressures are avoided, pressures which might otherwise impair the performance of the radial compressor **26**. Pulses in the output of the radial compressor **26** are also moderated. The air is then delivered in a compressed state to the exhaust port **16** of the exhaust cone **38**. The supercharged air then flows down the appropriate intake system of the engine until it reaches the cylinder or cylinders of the engine.

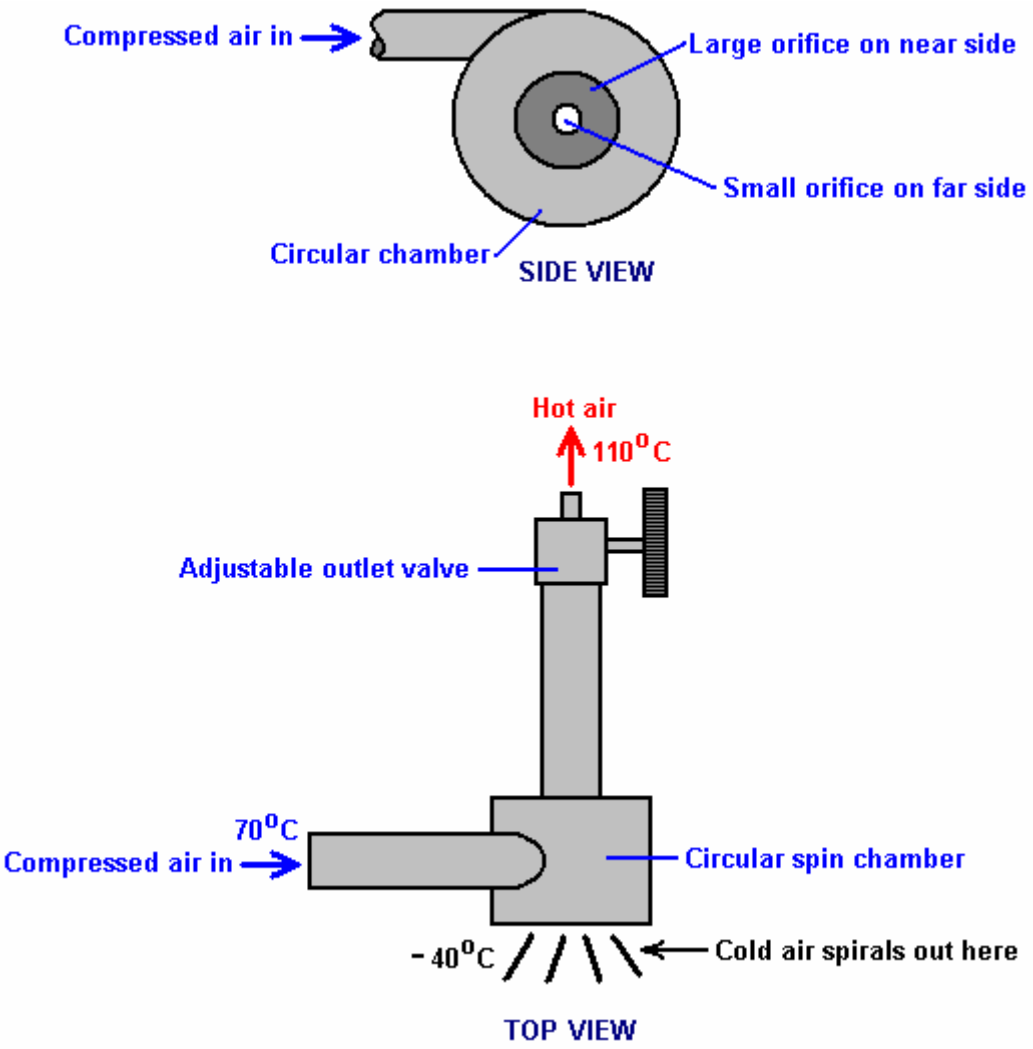


With respect to the application of the supercharger **10** to air-tank powered engines, such as disclosed in U.S. Pat. No. 4,292,804, the supercharger **10** functions in the same manner as described above, but is connected to the engine differently. In the air tank powered engine, at least one of the exhaust manifolds of the engines delivers partially expanded air to a line connected to the inlet **14** of the supercharger **10**. Referring to **Fig.8**, in most of such applications, this line will be of a smaller diameter than the housing **12** at the inlet **14** of the supercharger, such that an adaptor **173** is needed. The adaptor **173** comprises an annular plate **174** having a threaded aperture **176** sized to receive a mating, threaded end **178** of the line **180**. The plate **174** is secured to the flange **36** of the front housing section **32** by a plurality of bolts. Because the air coming from the line **180** is usually less than the full capacity of the supercharger, additional air is introduced through the lateral opening **44** along the side of the front housing section **32**. In this application, the opening **44** thus serves as an air intake port as well as a means for accommodating the drive belts **31** and must therefore be sized upon the additional criteria that it not be so large as to upset the flow of the incoming air in the line **180**. Upon the passage of the air through the supercharger, the air is directed through the exhaust port **16** and into a suitable line connected to it, which line may lead directly to the engine or to the storage tanks of the engine. If directed to the tanks, this recompressed air is used to supplement the required recharging of the storage tanks.

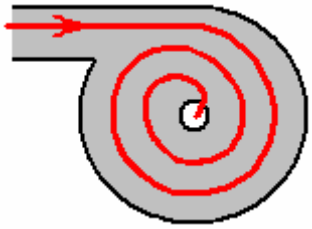
It is to be appreciated that savings in the cost of manufacturing the supercharger **10** are achieved by reason that the housing **12**, the bearing supports **40** and **114**, the axial compressor **24** and the radial compressor **26** are all constructed from cast aluminium parts and require only a minimum amount of machining. Moreover, the roller bearing assemblies 76 and **120** are commercially available components, and the supercharger **10** is easily assembled. These aspects further reduce the cost of manufacture and render the disclosed supercharger inexpensive to maintain and overhaul. More importantly, the supercharger **10**, despite its simple design, provides supercharging at relatively low operating speeds. With it's lower operating speeds, the service life of the supercharger **10** is extended and the risk of it suffering mechanical failure is reduced. The need for special bearing designs and lubrication is also avoided. Accordingly, the supercharger **10** is highly suitable for mass production and for use in cars, trucks, helicopters or the like.

The Vortex Tube.

The web site http://www.airtx.com/how-does-a-vortex-tube-work.htmv show "vortex tubes" which are completely passive devices with no moving parts:



This device does things which you would not expect. Compressed air at a temperature of, say, seventy degrees Centigrade is fed into the circular chamber where the shape of the chamber causes it to spiral rapidly as it exits the tube:



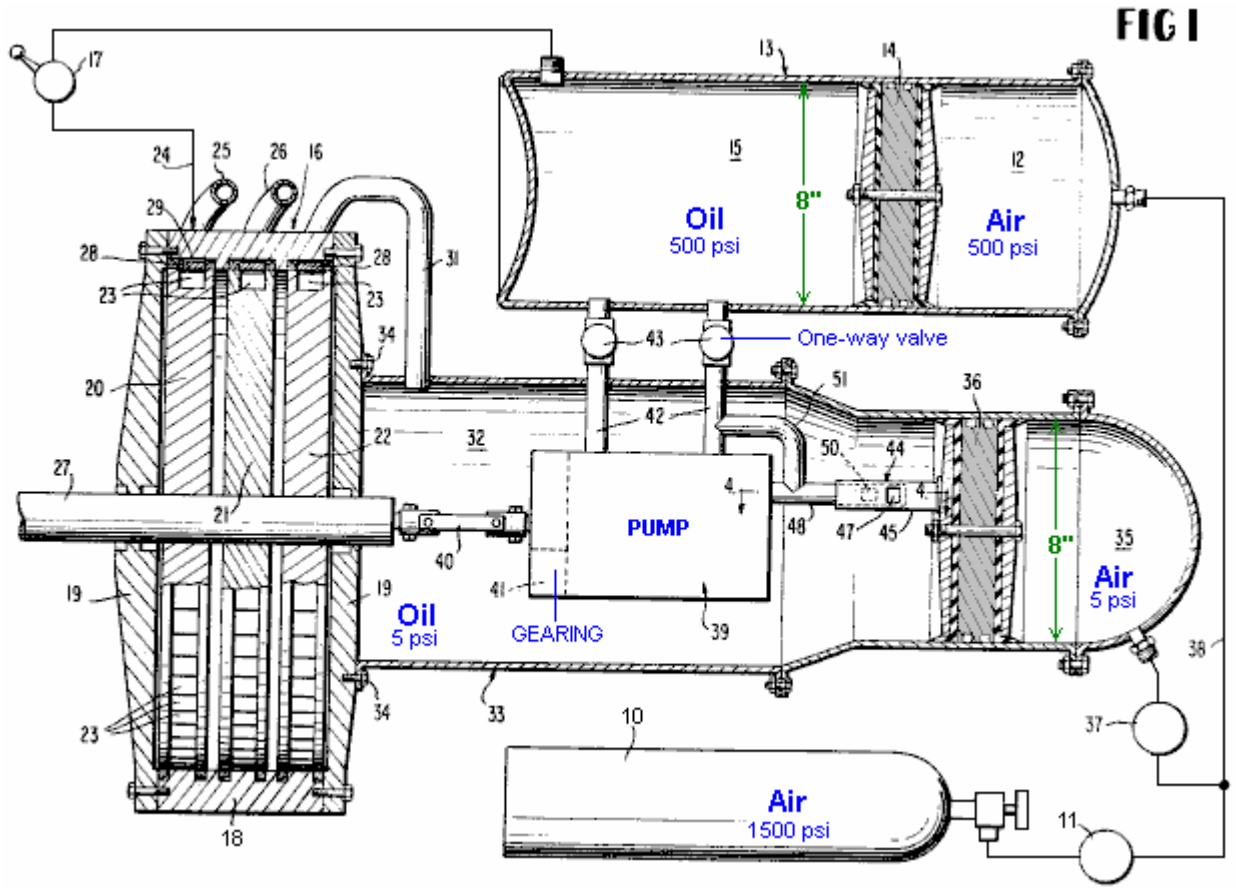
There is an energy gain in a vortex, as can be seen in a hurricane or tornado, but the really interesting thing here is the dramatic change in temperature caused by the change in pressure as the air expands. The ratio of heat gain to heat loss is controlled by the ratio of the sizes of the openings, which is why there is an adjustable nozzle on the small opening.

The air exiting through the large opening is much higher volume than the air exiting through the small opening and it expands very rapidly, producing a massive drop in temperature. The density of this cold air is now much higher than the air entering the vortex chamber. So there has been both a drop in temperature and an increase in density. These features of the expansion are made use of in the Leroy Rogers engine design, where some of the expanded air exhaust of the engine is compressed and passed back to the main air storage tank. While the compressor does raise the air temperature as it pumps the air back into the tank, it does not reach its original temperature instantly.

This results in the air temperature inside the tank dropping as the engine operates. But, the lowered tank temperature causes an inflow of heat from its immediate environment, raising the overall tank temperature again. This warming of the chilled air causes the tank pressure to increase further, giving an energy gain, courtesy of the local environment. It is important to understand that it takes less energy to compress air than the kinetic energy which can be generated by letting that compressed air expand again. This is a practical situation, courtesy of the local environment and is not a breach of the law of Conservation of Energy. It is also a feature which has not yet been exploited to any great degree and which is just waiting to be used by any adventurous inventor or experimenter.

Eber Van Valkinburg's Engine.

Eber presents a custom engine based on these principles. His engine uses both compressed air and compressed oil to manipulate pressures within the system and provide an engine which is self-powered. In the Appendix is a slightly re-worded copy of the Eber Van Valkinburg patent, which remarks that "stored energy in a compressed elastic fluid is utilised in a controlled manner to pressurise an inelastic fluid and to maintain such pressurisation. The pressurised inelastic fluid is throttled to the impeller of a prime mover. Only a portion of the output energy from the prime mover is utilised to circulate the inelastic fluid so as to maintain a nearly constant volumetric balance in the system".



Richard Clem's Motor.

The Clem Engine is based on an entirely different principle, and one which is not spoken about very often. Hurricanes or "twisters" as they are sometimes called, are large rotating air masses of incredible power which develop in hot areas which are more than eight degrees North or South of the equator. The distance from the equator is essential as the rotation of the Earth is needed to give them their initial spin. They usually develop over water which is at a temperature of twenty-eight degrees Centigrade or higher as that allows the air to absorb enough heat energy to get started. That is why there is a distinct "hurricane season"

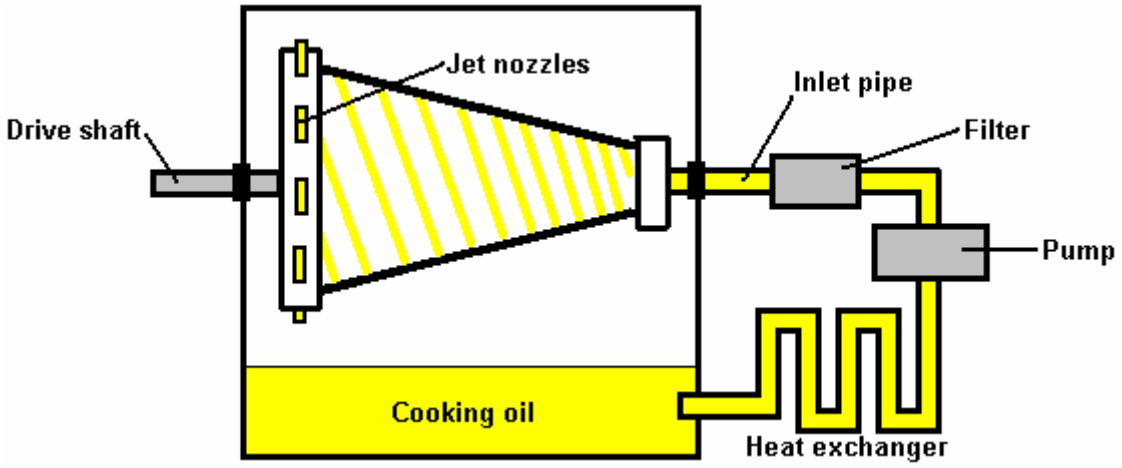
in these areas, since at certain times of the year the ocean temperature is just not high enough to trigger a hurricane.

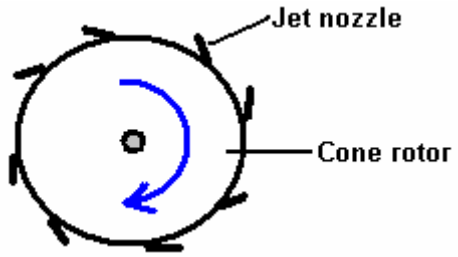
What is not generally realised is that a hurricane develops excess energy due to its swirling circular movement. The generation of this extra power was observed and documented by Viktor Schauberger of Austria, who also used his observations to great effect. I think that what Schauberger says makes some people uncomfortable as they seem to think that anything "unorthodox" has to be weird and too peculiar to be mentioned. This is rather strange as all that is involved here is a simple observation of how our environment actually works. A hurricane is wider at the top than at the bottom and this concentrates power at the base of the swirling mass of air. This tapered rotation is called a "vortex" which is just a simple name to describe the shape, but any mention of "vortex power" (the power at the base of this rotation) seems to make many people uncomfortable which is most peculiar.

Leaving that aside, the question is "can we use this energy gain from the environment for our own purposes?". The answer may well be "Yes". Perhaps this principle is utilised by Richard Clem. In 1992, Richard Clem of Texas, demonstrated a self-powered engine of an unusual type. This engine, which he had been developing for twenty years or more, weighs about 200 pounds (90 kilos) and generated a measured 350 horsepower continuously over the full period of a nine-day self-powered test. Although this engine which runs from 1,800 to 2,300 rpm is especially suited to powering an electrical generator, Richard did install one in a car, and estimated that it would run for 150,000 miles without any need for attention and without any kind of fuel. Richard said that his prototype car had reached a speed of 105 mph. Just after receiving funding to produce his engine, Richard died suddenly and unexpectedly at about 48 years of age, the death certificate having "heart attack" written on it as the cause of death. Remarkably convenient timing for the oil companies who would have lost major amounts of money through reduced fuel sales if Richard's motor had gone into production.

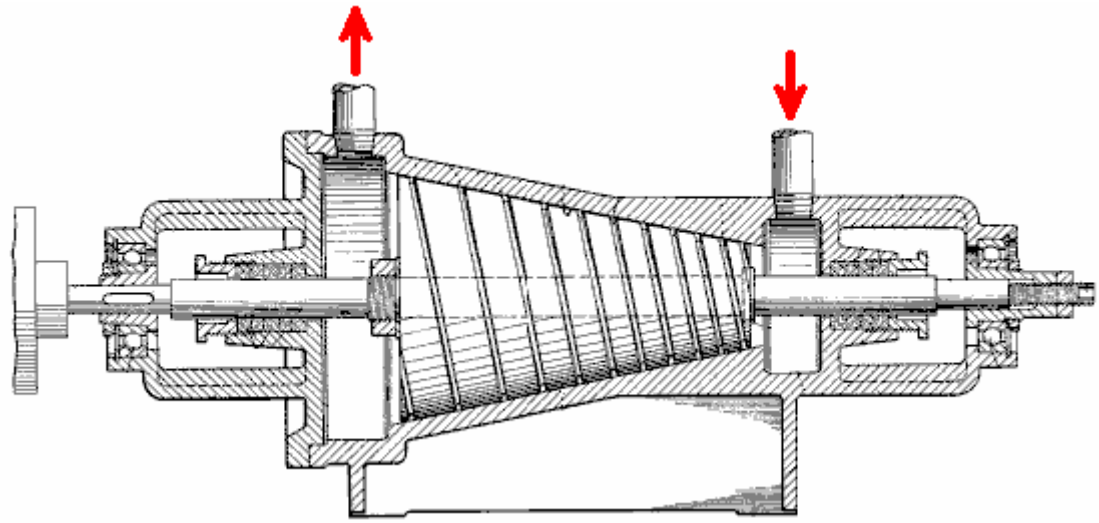
The motor is unusual in that it is a rotary turbine style design which runs at a temperature of 300^{o}F (140^{o}C) and because of that high temperature, uses cooking oil as its operational fluid, rather than water as the oil has a much higher boiling point. To a quick glance, this looks like an impossible device as it appears to be a purely mechanical engine, which will definitely have an operating efficiency which is less than 100%.

In broad outline, the oil is pumped through a pipe and into the narrow end of the cone-shaped rotor. The engine is started by being rotated by an external starter motor until it reaches the speed at which it generates enough power to be sustain its own operation. The rapid spinning of the cone, causes the oil to run along spiral grooves cut in the inner face of the cone and exit through angled nozzles placed at the large end of the cone:





The operating pressure produced by the pump is 300 to 500 psi. Richard did not attempt to patent his engine as US Patent 3,697,190 “Truncated Conical Drag Pump” granted in 1972 as a liquid-asphalt pump is so close in detail that Richard felt that there was insufficient difference for him to be granted a patent:



There appears to be considerable scope for anyone who wishes to build or manufacture this engine and it is capable of acting as a heater as well as device for producing mechanical power. This suggests that water purification could be an additional “extra” option for this engine.

Prof. Alfred Evert of Germany has produced an analysis of the operation of the Clem Engine and turbines in this general category. His website http://evert.de/indefte.htm has a good deal of information on the subject.

Josef Papp’s Inert Gas Engine Conversion.

The Hungarian, Josef Papp, invented an unusual engine system which genuinely appears to be very nearly “fuel-less”. His design modifies an existing vehicle engine to operate on a fixed amount of gas. That is to say, the engine has no air intake and no exhaust and consequently, no inlet or exhaust valves. The engine cylinders contain a mixture of gases which have an Atomic Number below 19, specifically, 36% helium, 26% neon, 17% argon, 13% krypton, and 8% xenon by volume. The control system causes the contained gas to expand to drive the pistons down the cylinders and then contract to suck the pistons back up the cylinders. This effectively converts the engine into a one-stroke version where there are two power strokes per revolution from every cylinder.

A small amount of radioactive material is used in the engine, and I have seen it suggested that the engine should be screened to protect the user from radiation. I’m not sure that this is correct, but if it is, then it suggests that a matter to energy conversion is indeed taking place. It seems most unlikely that the minor amount of radioactive material in the engine itself could cause any significant radiation. The patent describes the material as “low-level” which suggests to me, material no more dangerous that the luminous paint that used to be used on the hands of clocks and watches.

Suitable engines must have an even number of cylinders as they operate in pairs. Josef’s first prototype was a four-cylinder, 90 horsepower Volvo engine. He removed the intake and exhaust components and

replaced the engine head with his own design. During a thirty-five minute test in a closed room, the engine generated a constant 300 horsepower output at 4,000 rpm. The electrical power needed to run the engine was produced by the standard engine alternator, which was also able to charge the car battery at the same time. Interestingly, an engine of this type, quite apart from having zero pollution emissions (other than heat), is quite capable of operating under water.

Josef, a draftsman and ex-pilot, emigrated from Hungary to Canada in 1957 where he lived until his death in April 1989. There is solid evidence that Josef built an engine of over 100 horsepower (75 kilowatts) that was "fuelled" by a mixture of inert (or “noble”) gases. With no exhaust or cooling system, it had huge torque even at low rpm (776 foot-pounds at only 726 rpm in one certified test). Dozens of engineers, scientists, investors and a Federal judge with an engineering background saw the engine working in closed rooms for hours. This would not have been possible if the engine had been using fossil fuel. There was absolutely no exhaust and no visible provision for any exhaust. The engine ran cool at about 60°C (140°F) on its surface, as witnessed by several reliable observers. All these people became convinced of the engine's performance. They all failed to discover a hoax. Ongoing research in the United States (totally independent of Papp) has proved conclusively that inert gases, electrically triggered in various ways, can indeed explode with fantastic violence and energy release, melting metal parts and pushing pistons with large pressure pulses. Some of the people performing this work, or who have evaluated it, are experienced plasma physicists. Contemporary laboratory work has established that inert gases can be made to explode

In a demonstration on 27th October 1968 in the Californian desert, Cecil Baumgartner, representing the top management of the TRW aerospace corporation and others witnessed the detonation of one of the engine cylinders. In full public view, just a few cubic centimetres of the inert gas mixture was injected into the cylinder using a hypodermic needle. When the gas was electrically triggered, the thick steel walls of the cylinder were burst open in a dramatic way. William White, Edmund Karig, and James Green, observers from the Naval Underseas Warfare Laboratory had earlier sealed the chamber so that Papp or others could not insert explosives as part of a hoax. In 1983, an independent certification test was carried out on one of the Papp engines. Joseph Papp was issued three United States patents for his process and engines:

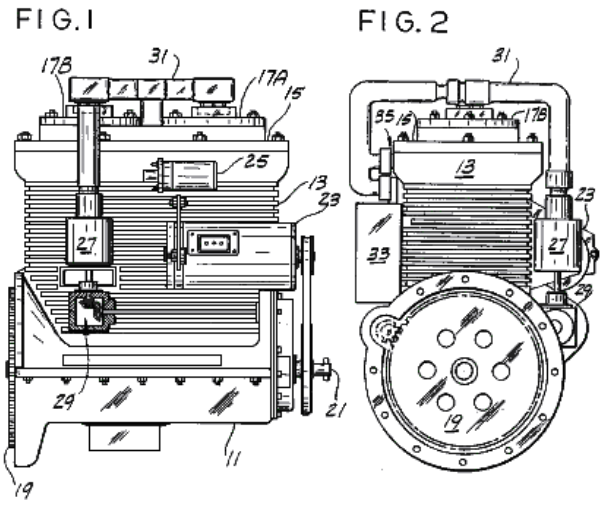
US 3,680,431 on 1st August 1972 "Method and Means for Generating Explosive Forces" in which he states the general nature of the inert gas mixture necessary to produce explosive release of energy. He also suggests several of the triggering sources that may be involved. It appears that Papp is not offering full disclosure here, but there is no doubt that others who have examined this patent and followed its outline have already been able to obtain explosive detonations in inert gases.

US 3,670,494 on 20th June 1972 "Method and Means of Converting Atomic Energy into Utilisable Kinetic Energy" and

US 4,428,193 on 31st January 1984 "Inert Gas Fuel, Fuel Preparation Apparatus and System for Extracting Useful Work from the Fuel". This patent shown here, is very detailed and provides information on building and operating engines of this type. It also gives considerable detail on apparatus for producing the optimum mixture of the necessary gasses.

At the time of writing, a web-based video of one of the Papp prototype engines running on a test bed, can be found at http://www.youtube.com/watch?v=N4li_z4Jpso although it must be said that a good deal of the footage is of very poor quality, having been taken many years ago. The video is particularly interesting in that some of the demonstrations include instances where a transparent cylinder is used to show the energy explosion. Frame-by-frame operation on the original video shows energy being developed outside the cylinder as well as inside the cylinder, which does seem to suggest that the zero-point energy field is involved. I have recently been contacted by one man who attended some of the engine demonstrations run by Papp and he vouches for the fact that the engine performed exactly as described.

Papp’s Patent US 4,428,193 is shown in full in the Appendix.



Josef never managed to get his engine design into commercial production before his death, primarily due to the opposition of vested interests. However, his design principles have been picked up and advanced by John Rohner and Haik Biglari.

The “Plasmic Transition Process” is the subject of various patents pending by PlasmERG Inc. of Iowa. John Rohner founded this company in 2008 to be the means to disseminate, develop and license this technology to other motor manufacturers for their own use. This process originally called the "Papp Engine" did run in 1982 and was then lost until John, and his partner Haik Biglari rediscovered it and applied modern science to the system to explain the process and filed their patents, presently pending. The original process was based on information originally patented by the late Joseph Papp, whose patents have now expired. John Rohner, a well known new-product design engineer, was originally contacted in 1979 by his brother Robert, with a schematic for the controller which Papp had designed. Unfortunately, John was busy with several other projects so he turned it over to his brother Tom.

PlasmERG has designed two motors for Own-Equipment-Manufacturers to use. One is an opposed, 2-cylinder, 120 cubic inch engine which produces some 300 horsepower. The second is a 6-cylinder 360 cubic inch engine which can produce around 1,500 horsepower. These motors are being co-developed with a sister company in Canada. John Rohner has personally provided the total investment for this development. As the company moves toward manufacturing, they are seeking investment partners by trading stock for investment. Their first commercial manufacturing plant will cost about 10 million dollars.

An alternative strategy is to create licenses for existing car and truck motor manufacturers until they can fund their own production. The current plan is to provide 500 to 1,000 test sites in underdeveloped nations for water pumping and power generation as "humanitarian" test sites. This should allow the time needed to get production understood and patents completed.

The expected run time of a motor from a single inert gas charge is over 3 months of continuous operation and gas re-charge should cost less than US $50. John stresses that the PlasmERG motor is **not**, (just as the original Papp engine was **not**), a "Pulsed Plasma motor". Plasma is not retained and "pulsed" as some people have supposed. What actually happens is that the plasma is recreated with each power stroke and then returns to a steady state gas on each return stroke, from which the name “Plasmic Transition” is derived.

The initial power and creation of plasma for expansion, is produced by a fusion event with a side-effect of a limited "chaotic" fission event causing a “plasmic transition” which is contained in a sealed 2-cycle rotating crankshaft motor.

There are two parallels to Plasmic Transition and power production of this motor. The first is natural lightning, which uses an almost identically similar Plasmic Transition process; and the second is steam which provides the same torque over rotational speed event characteristics. There is nothing in the ordinary internal combustion motor's operation that is comparable to either of these processes. The most crucial part of PlasmERG's motor operation is the Electronic Control System (ECS), comprising the following elements:

Programmable micro computers;

Radio frequency power generator;

HV spark coil initiation driver;

Various electromagnetic coil voltage switches providing base (resting) or variable (engine speed) voltages for all cylinder or reaction chamber electromagnetic coils;

On-controller DC to 12 volt DC converter;

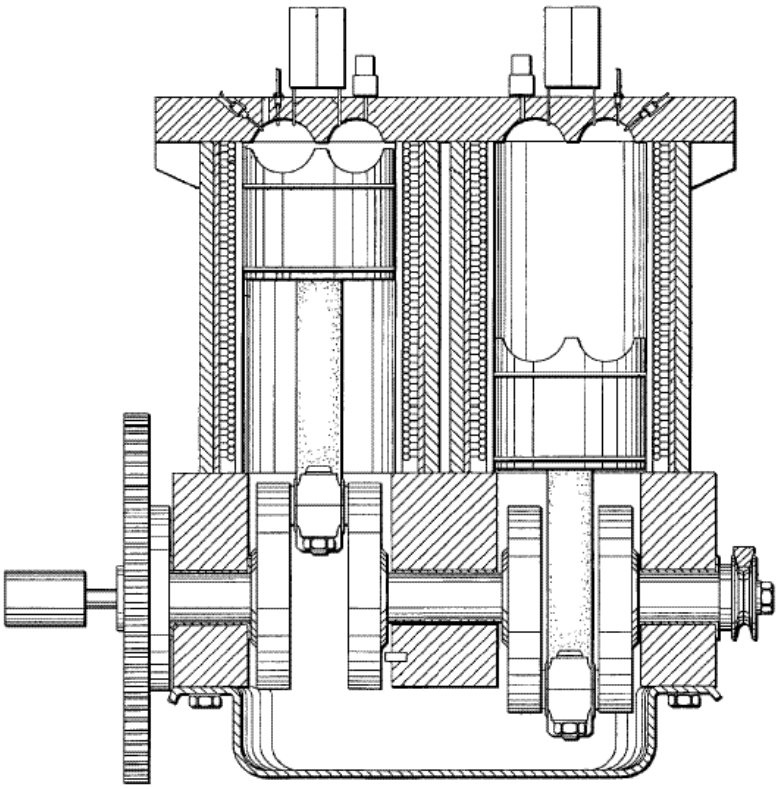
Engine speed DC voltage (accelerator) to programmed variable voltage DC converter,

Inter controller communications port;

Instrument support for user panel and action port which receives commands from the user comprising by not limited to things like Run, Start, Throttle position, Hold speed, Brake application, Brake hard, various motor inputs and fuel container information.

The PlasmaERG's website is at http://plasmerg.com/ but it needs Internet Explorer to display properly as it has major display problems when Firefox is used.

John Rohner's Patent Application US 2011/0113772 A1 entitled "Plasmic Transition Process Motor" can be downloaded from www.freepatentsonline.com or from www.free-energy-info.com. It shows a 2-cylinder motor as an example of the operation:



Robert Britt's Inert Gas Motor.

Robert Britt designed a very similar engine to that of Josef Papp, and he was also awarded a US patent for an engine operating on inert gasses. William Lyne remarks that this engine design may be replicated using a Chevy "Monza" 6-cylinder engine or a VolksWagen 4-cylinder engine. The heads are removed and the new heads cast using the "pot metal" used for "pseudo chrome" automotive trim. That alloy contains aluminium, tin, zinc and possibly antimony and is particularly suitable as the insides of the cavities can be polished to the high reflectivity specified in the patents.

A full copy of Robert Britt's patent US 3,977,191 is in the Appendix.

Michael Eskeli's Turbines.

In April 1989, Michael Eskeli was annoyed by a newspaper article published in the Dallas Times Herald which commented on the failure of science to come up with alternative power systems which do not rely on

petroleum products to operate. Michael responded in a letter to the Editor, stating that he holds patents for fuel-less power generators, work-free heat pumps, and other related items, 56 patents issued in the mid-70s.

Michael does hold many patents, one of which is shown in Chapter 14, as a work-free fuel-less heater. However, as I am not aware of any working prototype being shown, I must recommend that you consider the following information as "an idea" rather than a proven fact. As far as I am aware, in the 1970s, the US Patent Office did not demand to see a working prototype before granting a patent, especially if the patent related to a device based on accepted Engineering principles.

However, as Michael's claim is for self-powered devices, his claim seems too important to be ignored, prototype or no prototype, as competent people reading this may well understand the principles suggested and be in a position to build a self-powered device as a result. If that is the case, then I should really appreciate feedback information on any successful replications and the construction methods used.

As I understand it, Michael's self-powered devices are Heat Pumps where the additional energy is flowing from the heat contained in the air, courtesy of the heating effects of sunshine. Standard engineering, but with a design which utilises this available energy to provide practical mechanical output power for vehicles and electrical generators.

The Eskeli patents which I have been able to locate are:

3,650,636 Rotary Gas Compressor
3,719,434 Rotary Ejector Compressor
3,748,054 Reaction Turbine
3,748,057 Rotary Compressor with Cooling
3,758,223 Reaction Rotor Turbine
3,761,195 Compressing Centrifuge
3,795,461 Compressor with Cooling
3,809,017 Heat and Steam Generator
3,834,179 Turbine with Heating and Cooling
3,854,841 Turbine
3,861,147 Sealed Single-rotor Turbine
3,874,190 Sealed Single-rotor Turbine
3,879,152 Turbine
3,889,471 Dual-rotor Dual-fluid Turbine
3,895,491 Turbine with Dual Rotors
3,919,845 Dual-fluid Single-rotor Turbine
3,926,010 Rotary Heat Exchanger
3,931,713 Turbine with Regeneration
3,933,007 Compressing Centrifuge
3,933,008 Multi-stage Heat Exchanger
3,937,034 Gas Compressor-Expander
3,938,336 Turbine with Heating and Cooling
3,939,661 Power Generator
3,949,557 Turbine
3,961,485 Turbine with Heat Intensifier
3,962,888 Heat Exchanger
3,972,194 Thermodynamic Machine of the Vane Type
3,972,203 Rotary Heat Exchanger
3,981,702 Heat Exchanger
3,986,361 Turbine with Regeneration
4,003,673 Fluid Pressuriser
4,005,587 Rotary Heat Exchanger with Cooling and Regeneration *****
4,012,164 Rotor with Recirculation
4,012,912 Turbine
4,030,856 Rotor with Jet Nozzles
4,044,824 Heat Exchanger
4,047,392 Dual Rotor Heat Exchanger *****
4,050,253 Thermodynamic Machine
4,057,965 Thermodynamic Machine with Step-type Heat Addition
4,060,989 Thermodynamic Machine with Step-type Heat Exchangers
4,068,975 Fluid Pressuriser
4,077,230 Rotary Heat Exchanger with Cooling

4,106,304 Thermodynamic Compressor
4,107,944 Heat Pump with Two Rotors *
4,107,945 Thermodynamic Compressor
4,124,993 Refrigeration Machine
4,167,371 Method of Fluid Pressurisation
4,178,766 Thermodynamic Compressor Method
4,574,592 Heat Pump with Liquid-Gas working Fluid

And there are presumably 7 others not listed here, to raise the total to the 56 mentioned by Michael. I do not have the expertise to tell which of these may be self-powered just by reading the patent information, which generally does not mention anything along those lines, (the Patent Office staff not believing that COP>1 exists). Practically any of these patents might fit Michael's description, so I will pick the following patents to reproduce here:

4,107,944 Heat Pump with Two Rotors (continuing 4,005,587 and 4,047,392)
4,012,912 Turbine, and
3,931,713 Turbine with Regeneration

****

US Patent 4,107,944 **22nd August 1978** **Inventor: Michael Eskeli**

HEAT PUMP WITH TWO ROTORS

ABSTRACT

A method and apparatus for generating heating and cooling by circulating a working fluid within passageways carried by rotors, compressing the working fluid in them and removing heat from the working fluid in a heat-removal heat exchanger and adding heat into the working fluid in a heat-addition heat exchanger, all carried within the rotors. The working fluid is sealed in, and may be a suitable gas, such as nitrogen. A working fluid heat exchanger is also provided to exchange heat within the rotor between two streams of working fluid. In one arrangement, the unit uses two rotors, both rotating; in an alternate arrangement, one of the rotors may be held stationary. Applications include air conditioning and heating applications.

US Patent References:

2,490,064	Thermodynamic Machine	Dec 1949	Kollsman
2,490,065	Thermodynamic Machine	Dec 1949	Kollsman
2,520,729	Machine for producing Heat Energy	Aug 1950	Kollsman
2,597,249	Thermodynamic Engine	May 1952	Kollsman
3,470,704	Thermodynamic Apparatus and Method	Oct 1969	Kantor
3,834,179	Turbine with Heating and Cooling	Sep 1974	Eskeli
3,861,147	Sealed Single-rotor Turbine	Jan 1975	Eskeli
3,889,471	Dual-rotor Dual-fluid Turbine	Jun 1975	Eskeli
3,895,491	Turbine with Dual Rotors	Jul 1975	Eskeli
3,919,845	Dual-fluid Single-rotor Turbine	Nov 1975	Eskeli
3,931,713	Turbine with Regeneration	Jan 1976	Eskeli
4,005,587	Rotary Heat Exchanger with Cooling & Regeneration	Feb 1977	Eskeli
4,044,824	Heat Exchanger	Aug 1977	Eskeli

Cross References to Related Applications

This application is a continuation-in-part application of "Dual Rotor Heat Exchanger" filed Nov. 18, 1973, Ser. No. 407,665, now U.S. Pat. No. 4,047,392.

This application also is a continuation-in-part of "Heat Pump" filed June 30, 1975, Ser. No. 591,881, now abandoned.

And this application also is a continuation-in-part of "Rotary Heat Exchanger with Cooling and Regeneration" filed Oct. 1, 1975, Ser. No. 618,456, now U.S. Pat. No. 4,005,587.

BACKGROUND OF THE INVENTION

This invention relates generally to devices for heat transfer from a lower temperature to a higher temperature by using a working fluid enclosed within a centrifuge rotor as an intermediate fluid to transport the heat.

Heat pumps have been known in the past but are complex and costly, and usually use a working fluid that is evaporated and condensed, which results in poor efficiency, and thus high energy cost.

SUMMARY OF THE INVENTION

It is an object of this invention to provide apparatus that is low in initial cost and has high thermal efficiency thus reducing cost of the power required to run it. It is further the object of this invention to provide a device and process wherein the losses that normally occur in bearings and seals, due to friction, are applied to the working fluid for its circulation, thus in effect eliminating the power loss due to such friction losses. Also, it is an object of this invention to provide the rotor with a working fluid heat exchanger to reduce needed rotor speeds.

BRIEF DESCRIPTION OF THE DRAWINGS

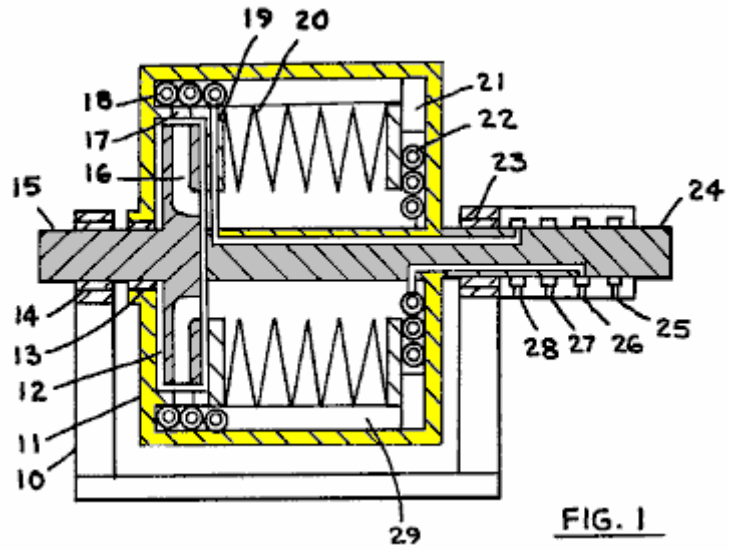


Fig.1 is a cross section of the device.

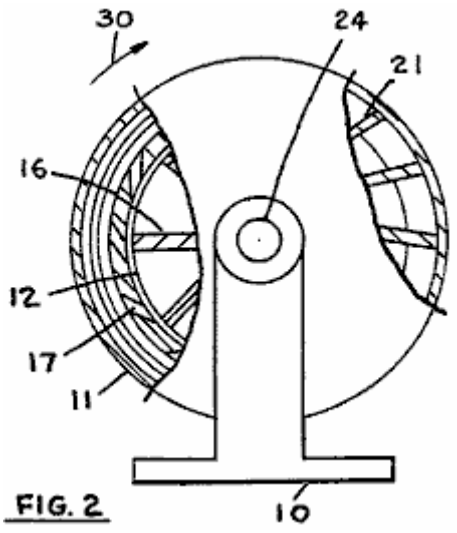


Fig.2 is an end view of the device.

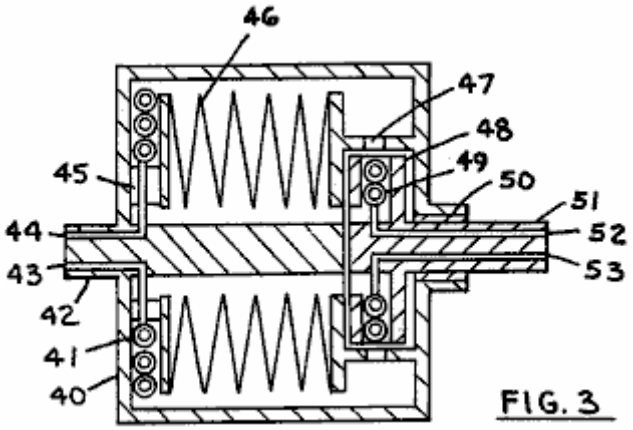


Fig.3 is an axial cross section of another form of the device.

DESCRIPTION OF PREFERRED EMBODIMENTS

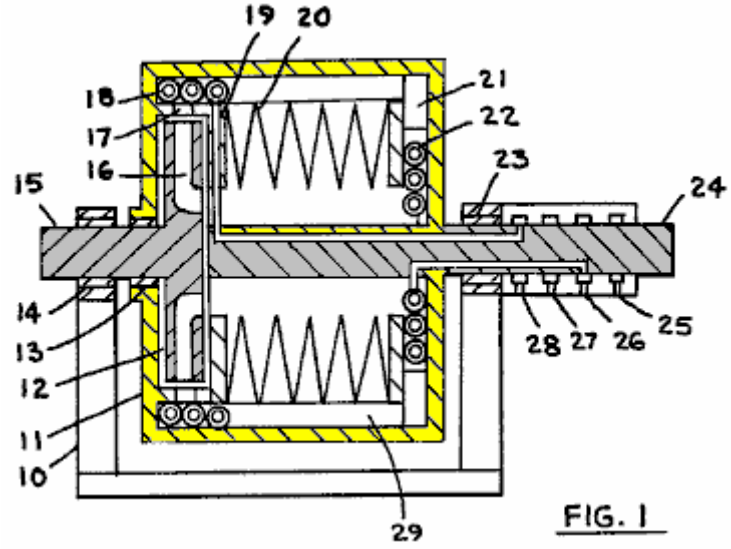


Fig.1 shows an axial cross section of the device, where **10** is the base, **11** is the first rotor, **12** is the second rotor, **13** is a seal and **14** is the bearing supporting shaft **15**, **16** is fluid passage in the second rotor, **17** is the

working fluid opening which may be a nozzle, **18** is the first heat exchanger for heat removal from working fluid, **19** is first heat transfer fluid conduit, **20** is working fluid heat exchanger, in this instance formed from sheet metal like bellows, **21** are vanes, **22** is second heat exchanger for heat addition to working fluid, **23** is bearing supporting shaft **24**, **25** and **26** are entry and exit for second heat transfer fluid, **27** and **28** are entry and exit for first heat transfer fluid, and **29** is a vane in peripheral passage.

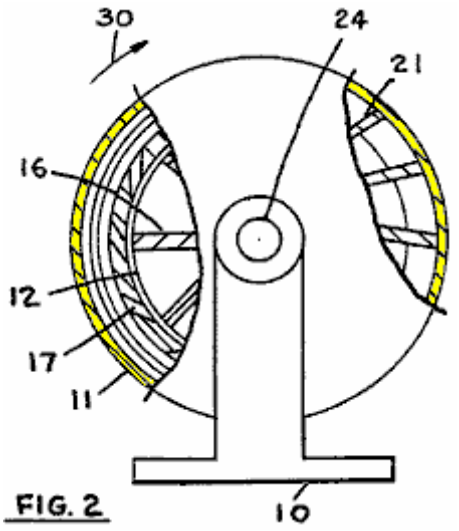


FIG. 2

In **Fig.2**, an end view of the unit shown in **Fig.1** is illustrated. Where **10** is base, **11** is first rotor, 17 are fluid openings, **12** is second rotor, **16** are second rotor fluid passages with vanes, **30** indicates direction of rotation, **24** is first rotor shaft, and **21** are vanes.

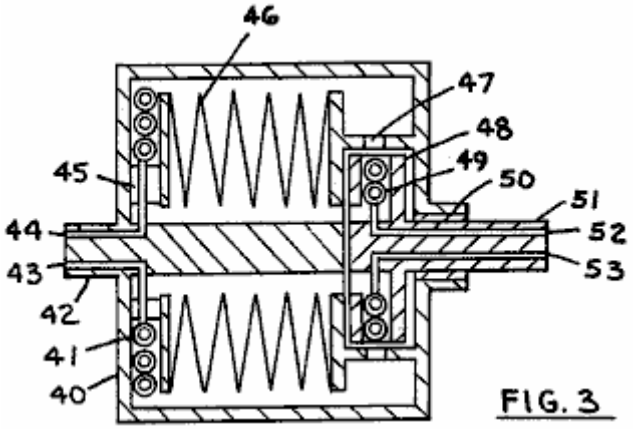


FIG. 3

In **Fig.3**, the rotors are arranged differently, but perform the same functions, approximately, as in the unit of **Fig.1**. Where **40** is first rotor, **41** is first heat exchanger for heat removal from first fluid, **42** is first rotor shaft, **43** and **44** are entry and exit for first heat transfer fluid, **45** is conduit, **46** is working fluid heat exchanger, **47** are fluid openings which may be nozzles, **48** is second rotor, **49** is second heat exchanger for adding heat to the working fluid, **50** is bearing and seal, **51** is second rotor shaft, **52** and **53** are entry and exit for second heat transfer fluid.

In operation, the rotors are caused to rotate and the rotor cavities are filled with a suitable working fluid, which is usually a gas, such as nitrogen, air or other gaseous or vapour substance. Referring to **Fig.1**, the second rotor rotates usually faster than the first rotor, and the working fluid is compressed by centrifugal force in passages **16**, and in the first rotor to some extent, after which heat is removed in heat exchanger **18**, with such heat then being transported by the first heat transfer fluid out of the device. The working fluid then passes along the peripheral passage **29** and releases heat in heat exchanger **20**, after which the fluid is expanded against centrifugal force in vanes **21** and in heat exchanger **22** where heat is added to the working fluid. After expansion, the working fluid passes along centre passage and receives heat from heat exchanger **20**, thus completing its work cycle.

The operation of the unit in **Fig.3** is similar, except that the second rotor usually rotates slower than the first rotor, and the second rotor may be kept stationary, if desired. Note that if the second rotor is held stationary, one may use dirty water as the second heat transfer fluid; normally, in rotating heat exchangers,

the heat transfer fluid must be free of solids, which will collect in the heat exchanger due to centrifugal force, and block the heat exchanger, and by having a stationary heat exchanger, ordinary water may be used, such as water from a cooling tower.

In the unit of **Fig.1**, the power input is normally to the second rotor, and the first rotor is allowed to rotate freely. In such usage, the rotor diameters are selected to provide, together with the friction loss in bearings, for the needed speed differential between the two rotors. With the second rotor rotating faster, necessary push for the working fluid is provided to keep the working fluid circulating. Alternately, the speed differential may be maintained by using a power transmission between the two rotors, such as a gearbox. In the unit of **Fig.3**, the second rotor speed is slower than the speed of the first rotor, and where the rotor diameters are suitable, the second rotor may be held stationary, providing needed push for the working fluid for its circulation.

The working fluid heat exchanger **20** and **46**, employ centrifugal force and varying gas density to obtain heat exchange between the two working fluid streams. Hot gas in the peripheral passage is lighter, and colder gas between the folds of the heat exchanger is colder, thus the cold gas is displaced by lighter gas by centrifugal force. Similarly, at the centre passage, cold gas at centre displaces hot gas between folds. Other types of heat exchangers may be used for the heat exchanger **20**, including heat pipes, sheet metal discs, and finned tubing filled with a liquid.

The rotor may be encased within a vacuum tank, if desired, to reduce friction on rotor outer surfaces. The use of the working fluid heat exchanger **20** will reduce required rotor speeds to obtain required temperature differentials between the two heat transfer fluids, which then reduces friction losses on the rotor, which may eliminate the need for a vacuum tank.

Various modifications of this device may be made, and different types of heat exchangers used. Also, working fluid radial passages may be curved in various directions, one being the slope for vanes shown as item **21** in **Fig.2**. By using vane slopes and sloped passages, one can adjust the amount of work exchange between the working fluid and the rotor. Nozzles **47** are usually positioned so as to discharge backwards, in order to generate some torque on the first rotor, and similar nozzles may also be used in passages **21** of the unit shown in **Fig.1**. Further, the heat exchanger **22**, of **Fig.1**, may be mounted on a stationary member, if desired, in manner shown in **Fig.3**, and heat exchanger **18** may be mounted within rotor **12**, if desired. The various components of the units may be interchanged, as desired.

CLAIMS

1. In a heat pump wherein a compressible working fluid is circulated radially outwardly in a first fluid passage, said first passage contained in a first member, and radially inwardly towards centre of rotation in a second fluid passage, said second passage contained in at least one of said first and second members, said first and second members coaxially arranged, at least one of said members being supported by a shaft for rotation;

said first and said second radial working fluid passages communicatingly connected at their respective outward ends by an outer passage and at their respective inward ends by an inner passage, said radial and outer and inner passages forming a closed loop extending at least partially through both of said members, a working fluid adapted to be circulating through said loop, means for compressing said working fluid by centrifugal force within said loop with accompanying temperature increase, first heat exchange means for cooling said working fluid after compression, said first heat exchange means being carried by one of said members, a second heat exchange means, carried by one of said members, for regeneratively exchanging heat between said working fluid within said inner and outer passages, and a third heat exchange means carried by one of said members for heating said working fluid after said heat exchange between said working fluid within said inner and outer passages.

2. The heat pump of claim 1 wherein a first heat transfer fluid is circulated within said first heat exchange means to remove heat with said first heat exchange fluid entering and leaving via conduits near the centre of rotation of said members.

3. The heat pump of claim 1 wherein a second heat transfer fluid is circulated within said third heat exchange means entering and leaving via conduits near the centre of rotation of said members.

4. The heat pump of claim 1 wherein both of said members are rotors.

5. The heat pump of claim 4 wherein the two rotors rotate at different angular speeds.

6. The heat pump of claim 1 wherein at least one of said members is a rotor.

7. The heat pump of claim 6 wherein said second heat exchange means includes a plurality of folds.

8. The heat pump of claim 7 wherein said second heat exchange means is of bellows configuration.

****

US Patent 4,012,912 **22nd March 1977** **Inventor: Michael Eskeli**

TURBINE

ABSTRACT

A method and apparatus for the generation of power wherein a working fluid is compressed within outward extending rotor passages, and then passed inward in other rotor passages with accompanying expansion and deceleration, with work being generated by the decelerating fluid. Heat may be added into the working fluid near the rotor periphery, and in closed rotors, heat is removed from the working fluid after expansion. A regenerator may also be used, mounted on the rotor, exchanging heat between two streams of the working fluid. During the deceleration, the working fluid passages are curved backwards, while the working fluid passages for acceleration are usually radial. The working fluid may be either a liquid or a gas, and the heating fluid and the cooling fluid may also be either a liquid or a gas.

US Patent References:

3,761,195 Compressing Centrifuge	Sept 1973	Eskeli
3,834,179 Turbine with Heating and Cooling	Sept 1974	Eskeli
3,926,010 Rotary Heat Exchanger	Dec 1975	Eskeli

Cross References to Related Applications:

This application is a continuation-in-part application of "Turbine," Ser. No. 566,373, filed 4-9-75 now U.S. Pat. No. 3,949,557.

BACKGROUND OF THE INVENTION

This invention relates to power generators where a working fluid is circulated from a higher energy level to lower energy level, generating power.

In my earlier U.S. Pat. Nos. 3,874,190 and 3,854,841, I described a closed and open type turbines, and using centrifuge design. These turbines used forward facing nozzles within the rotor; in the apparatus disclosed here, such nozzles have been replaced by other methods.

SUMMARY OF THE INVENTION

It is an object of this invention to provide a single rotor centrifuge type turbine stage, where vanes or fins, with suitable contours, are used to extract power from the working fluid, using either an open type or a closed type rotor.

BRIEF DESCRIPTION OF THE DRAWINGS

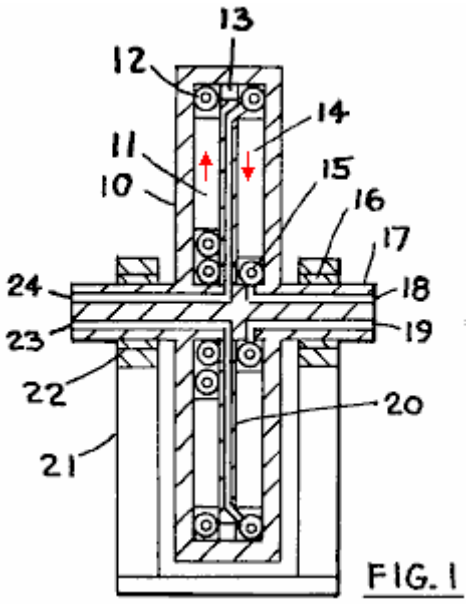


Fig.1 is a cross section and

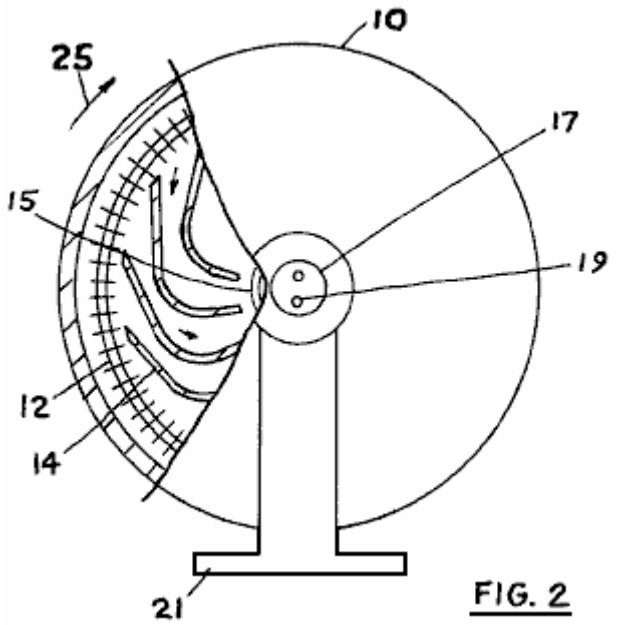


Fig.2 is an end view of a closed type rotor.

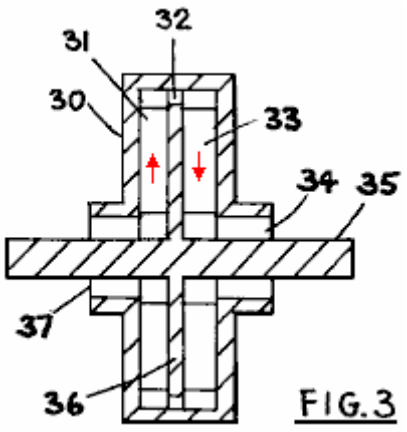


Fig.3 is a cross section and

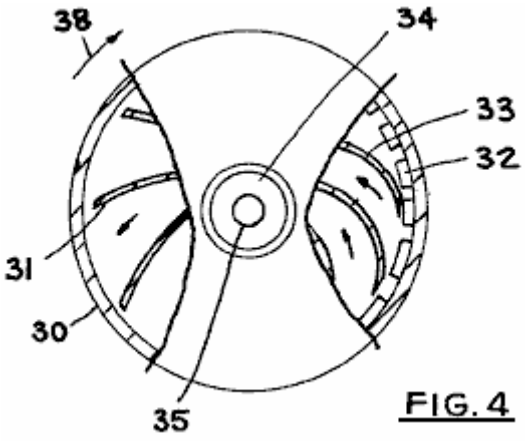


Fig.4 is an end view of an open type rotor.

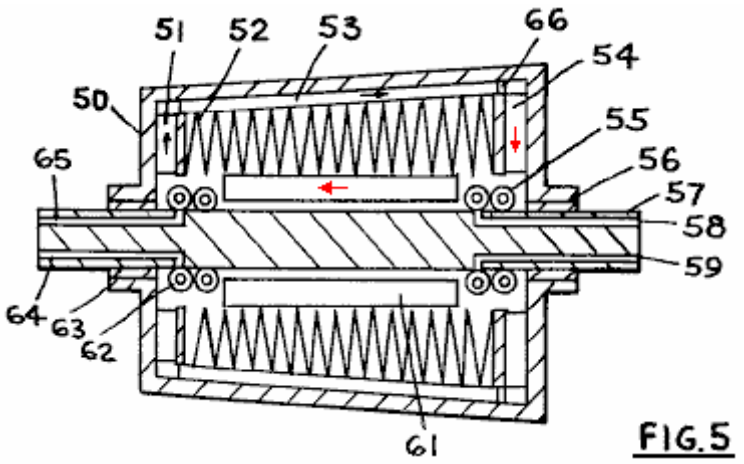


Fig.5 is a cross section of a unit using a closed type rotor and also using a regenerator.

DESCRIPTION OF PREFERRED EMBODIMENTS

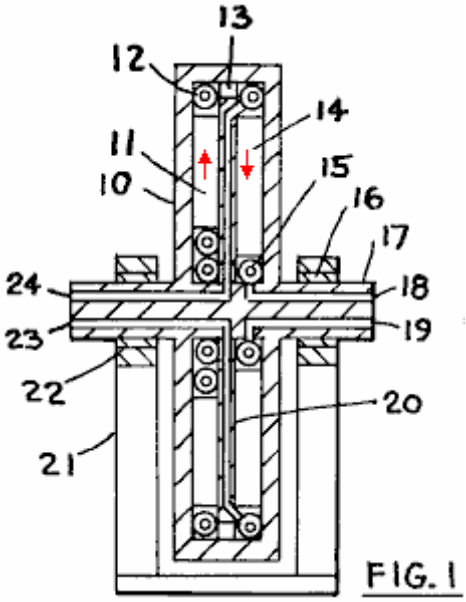


FIG. 1

Referring to **Fig.1**, there is shown a cross section of one form of the unit. Where **10** is the rotor which is supported by bearings **16** and **22**, shaft **17** and base **21**. **12** is a heat supply heat exchanger and **15** is cooling heat exchanger, **14** and **11** are vanes or fins, **18** and **19** are coolant entry and exit, **20** is a dividing wall, **23** and **24** are heating fluid entry and exit, and **13** is a working fluid passage which may be used to regulate the flow of working fluid within the rotor.

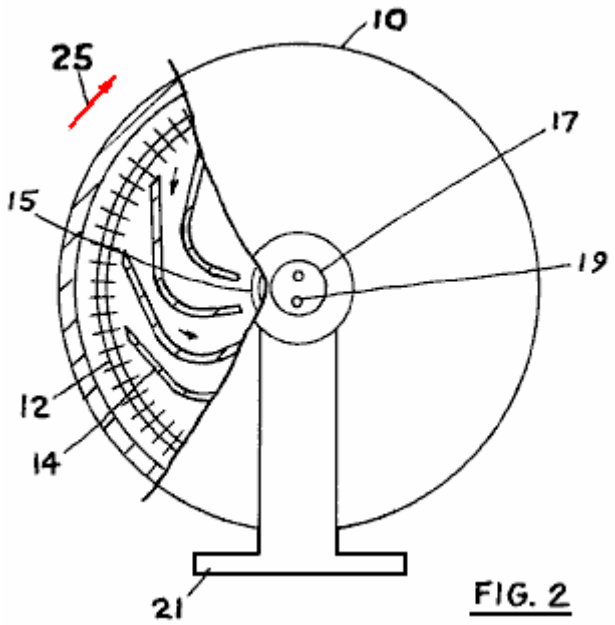


FIG. 2

Fig.2 is an end view of the unit shown in **Fig.1**. Where **10** is the rotor, **17** the shaft, **19** is a coolant passage, **21** is the base, **14** are vanes positioned so that they slope away from the direction of rotation as indicated by arrow **25**, while simultaneously passing the working fluid inwards, **12** is the heating heat exchanger, and **15** is the cooling heat exchanger.

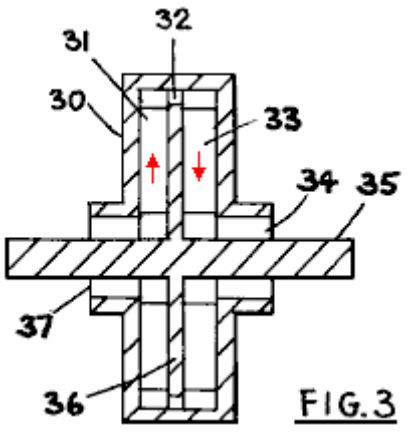


FIG.3

In **Fig.3**, a rotor for a unit using open cycle is used, where the working fluid enters and leaves the rotor. Here, **30** is the rotor, **31** is the vane situated in a passage which extends outwards, **32** is the fluid passage, **33** is a vane in the passage for inward bound working fluid, **34** is the working fluid exit, **35** is the rotor shaft, **36** is a rotor internal divider and **37** is the working fluid entry into the rotor.

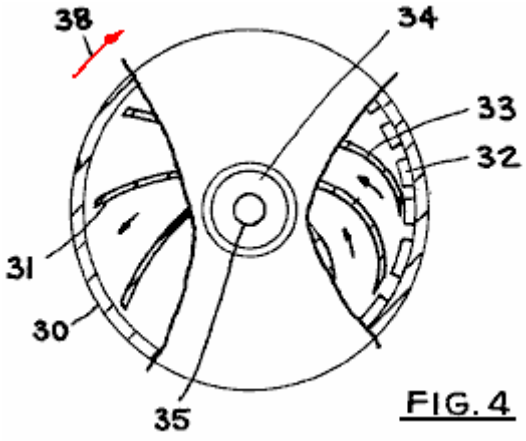


FIG. 4

Fig.4 shows an end view of the unit of **Fig.3** where **30** is the rotor, **35** is the shaft, **31** are vanes in the passages for outward bound fluid, and are shown here to be curved backwards, when the rotor rotates in the direction shown by arrow **38**. After passing openings **32**, the working fluid passes inwards guided by vanes **33**, and then leaving via exit **34**. Vanes **33** are curved as indicated, with the curvature being away from the direction of rotation, so the working fluid provides thrust against the rotor components as it decelerates when passing inwards toward the centre of the rotor.

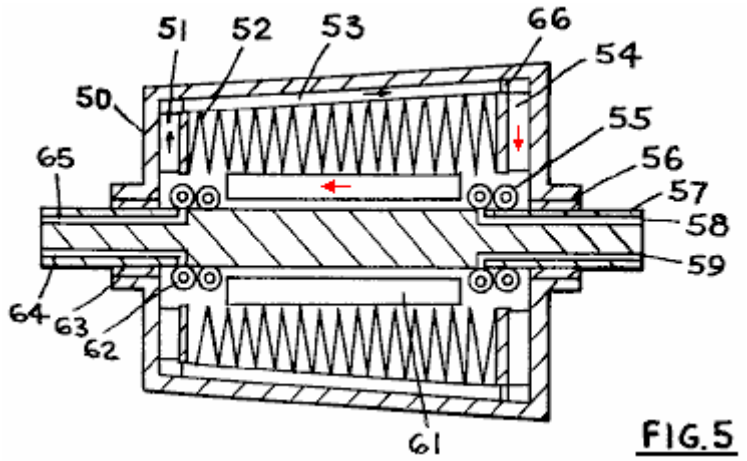


FIG.5

In **Fig.5**, a rotor with a regenerator is shown, and also the rotor shaft is arranged so that it can be kept stationary if desired. **50** is the rotor which is supported by bearings **56** and **63** and shaft **57**. Vanes **51** may be radial or curved as desired, and vanes **54** are curved in a manner similar to vanes **33** in **Fig.4**. **52** is a

regenerative heat exchanger, exchanging heat between the working fluid streams flowing in passages **53** and **61**. Heat supply heat exchanger **55** and cooling heat exchanger **62** are attached to the shaft, so that the shaft may be kept stationary or rotated at a different speed than the rotor **50**. **58** and **59** are the entry and exit points for the heating fluid while **64** and **65** are the entry and exit points for the cooling fluid, and **66** is an opening.

When operating, the rotor rotates, and a working fluid within the rotor passes outwards in passage **11**, and is compressed by centrifugal force, and accelerated to a tangential speed that may be the same as that for the rotor periphery. In a closed rotor such as is shown in **Fig.1**, heat is added into the working fluid near the rotor periphery, and then the working fluid decelerated in the fluid passages **14** extending inwards toward rotor centre, with the passages being curved backwards away from the direction of rotation as shown in **Fig.2**. As the working fluid is decelerated in the inward extending passages, the work associated by such deceleration is transferred into the rotor and this provides the thrust and torque to rotate the rotor. After deceleration and expansion, the working fluid is cooled in heat exchanger **15** and then passed to the outward extending passages thus completing its working cycle.

The operation of the unit of **Fig.3** is similar, except that the working fluid enters the rotor via opening **37** from external sources. For the unit shown in **Fig.3**, the heat addition heat exchanger is omitted; for this unit, there is a pressure drop between entry **37** and exit **34**. A heat exchanger similar to that shown in **Fig.1**, item **12**, may be used in the unit of **Fig.3**, and then the entry and exit pressure for the working fluid may be the same, if desired.

The operation of the unit shown in **Fig.5**, is similar to that described for the other units. The rotor rotates, and by centrifugal force, compresses the working fluid in passages **51**, and then the working fluid gains heat in the regenerative heat exchanger, with the heat being supplied by another working fluid stream returning from the high temperature end of the unit. The working fluid is expanded and decelerated in passages **54** and heat is added in the heat exchanger **55**. Then the working fluid passes through the regenerative heat exchanger and then is cooled in the cooling heat exchanger and then is passed into passages **51** thus completing its cycle.

The various components of the units shown can be exchanged to make additional forms of the apparatus. As noted, the unit of **Fig.3** may be provided with a heat exchanger similar to that shown in **Fig.1** for adding heat into the working fluid near the rotor periphery. Further, a regenerator may be provided with the units of **Fig.1** and **Fig.3**, if desired, between the outward extending and the inward extending working fluid passages. Also, the cooling coil of **Fig.5**, item **62** may be eliminated, and the working fluid taken into the unit from outside the unit, if desired.

The openings **32**, **13** and **66** may be made into nozzles, if desired, and the nozzle oriented in different directions as desired. In particular, these nozzles may be positioned so as to discharge the working fluid tangentially backwards, if desired.

The regenerator of **Fig.5** is shown to be tapered. This taper may be as shown, or the taper may be made such that the regenerator portion diameter is smaller at the end which has the heat exchanger **55**, than the end which has the heat exchanger **62**. Also, the regenerator may be made without a taper.

Passages **53** and **61** are usually provided with vanes, as indicated in **Fig.5**, to prevent tangential movement of the working fluid.

Applications for this power generator are those normally encountered in power generation.

The working fluid is usually a gas for units such as those shown in **Fig.1** and **Fig.5**, but the working fluid may also be a liquid for a unit such as shown in **Fig.3**. The heating and cooling fluids may be either gases or liquids, as desired.

The heat exchangers for heating and cooling are shown to be made of finned tubing. Other forms of heat exchangers for adding heat and for removing heat may be used. The regenerative heat exchanger is shown to be made of sheet metal; other forms of heat exchangers may be also used.

US Patent 3,931,713 **13th January 1976** **Inventor: Michael Eskeli**

TURBINE WITH REGENERATION

ABSTRACT

A method and apparatus for generating power by passing a motivating fluid from a higher energy level to a lower energy level by compressing the fluid in a centrifuge-type first rotor and discharging the fluid via nozzles near the periphery of the first rotor, forwards in the direction of rotation to a second rotor which is an inward flow type reaction turbine, then passing the fluid through a regeneration type heat exchanger to transfer heat from the inward bound fluid into the outward bound fluid, after which the fluid is cooled in a heat exchanger to its original temperature and is passed outward again thus completing its cycle. Heat is added to the fluid near the periphery of the second rotor, or the heat may be added near the periphery of the first rotor, or both. Additionally, the fluid may be supplied to the unit from outside source, and returned to such outside source, and the cooling may thus be eliminated from the unit. Further, the fluid entering from an outside source may be at an elevated pressure. The fluids used may be gaseous, which is normal for a closed type unit, or they may be liquids at entry for the open type unit.

US Patent References:

2,490,064 Thermodynamic Machine	Dec 1949	Kollsman
2,514,875 U-passage Gas Turbine	July 1950	Kollsman
2,597,249 Thermodynamic Engine	May 1952	Kollsman
3,236,052 Closed-cycle Gas Turbines	Feb 1966	Guin
3,530,671 Regenerative Air Turbines	Sep 1970	Kolodziej

This application is a continuation-in-part application of "Turbine with Dual Rotors," Ser. No. 405,628, filed 10/11/73, and uses material of a previous U.S. Pat. No. 3,834,179, "Turbine with Heating and Cooling".

BACKGROUND OF THE INVENTION

This invention relates generally to devices for generating power in response to a fluid flowing from a higher energy level to a lower energy level passing through a turbine for generating the power.

There have been various types of turbines previously, in some of which a fluid is accelerated in a single or multiple stationary nozzles and then passed to vanes mounted on a rotating rotor wheel, where the kinetic energy contained by the moving fluid is converted to power by deceleration of the fluid.

These conventional turbines normally have a high energy loss due to fluid friction, especially between rotor vanes and the fluid where the velocity differential is usually large. Also, these turbines often require complex shaped turbine vanes making the unit costly.

SUMMARY OF THE INVENTION

It is an object of this invention to provide a turbine for power generation in which heat is converted to power, in an efficient and economical manner, and with high thermal efficiency. It is also an object of this invention to provide a means for transferring heat from the motivating or working fluid, which is the first fluid, during its passage from rotor periphery to rotor centre into the first fluid which is passing from the rotor centre towards the rotor periphery. This heat transfer improves the efficiency of the turbine, and reduces the necessary rotational speed of the rotor, allowing less costly rotor construction.

BRIEF DESCRIPTION OF THE DRAWINGS

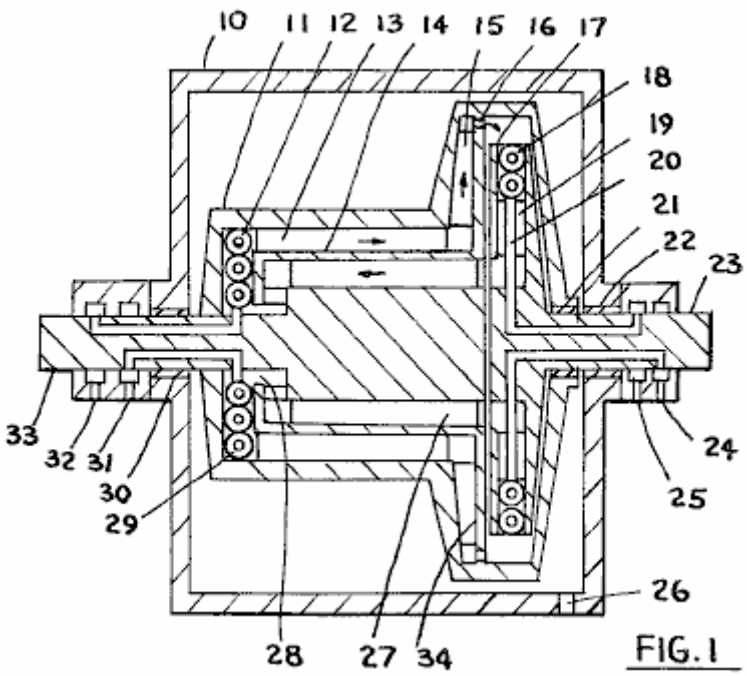


Fig.1 is a cross section of one form of the device, and

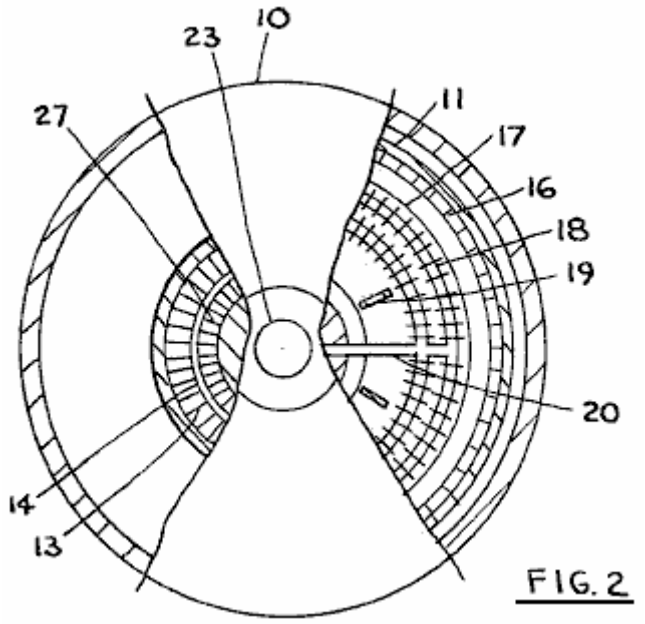


Fig.2 is an end view of the unit shown in **Fig.1**.

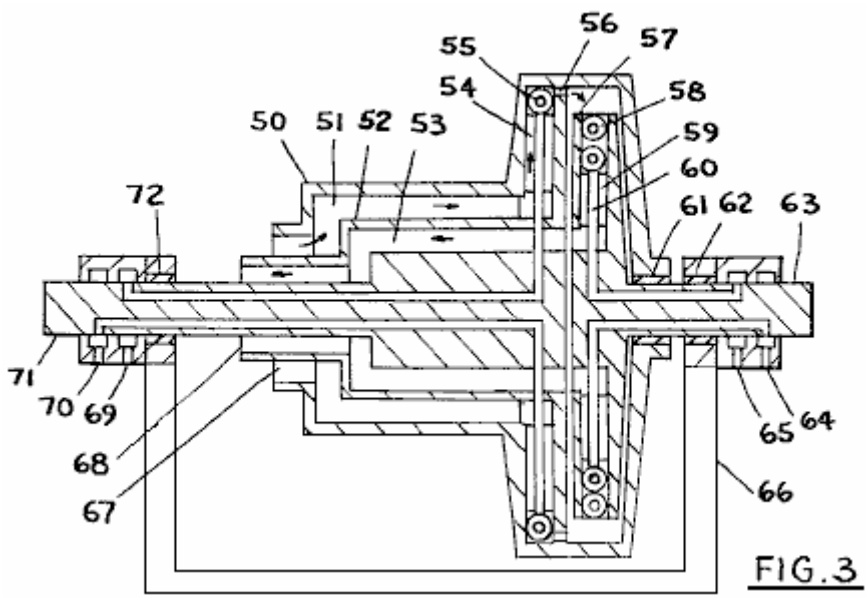


Fig.3 is a cross section of another form of the device.

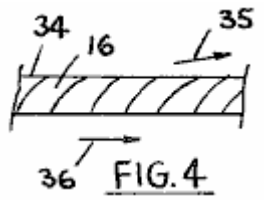


Fig.4 is a detail of rotor nozzles.

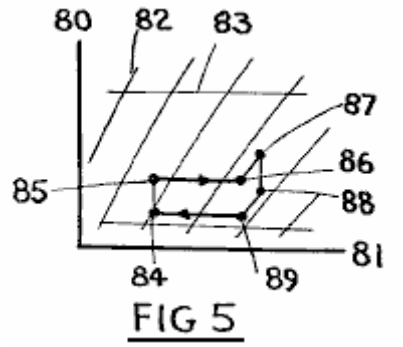


Fig.5 is a pressure-enthalpy diagram of the first fluid with working cycle illustrated for the first fluid.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

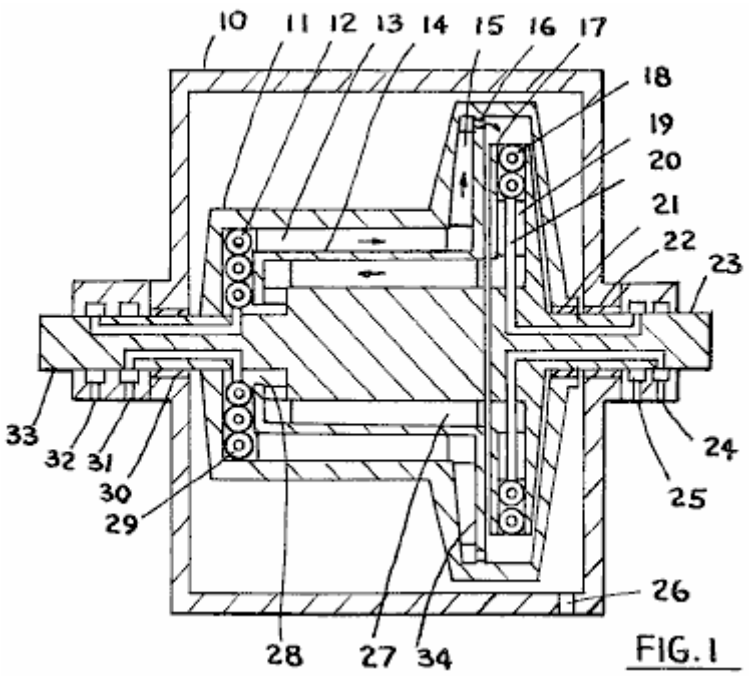


FIG. 1

Fig.1 shows a cross section of one form of the turbine. In this form, the first fluid is sealed within the rotor with a second fluid which supplies heat to the first fluid, and a third fluid which cools the first fluid, being circulated from external sources.

The first fluid is accelerated and compressed within the first rotor, and after discharge from the nozzles of the first rotor, into the second rotor, where it receives heat from the second fluid, and after deceleration and expansion the first fluid passes in heat exchange relationship with the first fluid flowing outward so that heat is transferred from the inward bound first fluid to the outward bound first fluid. Cooling is then provided for the first fluid to bring the first fluid temperature to an initial predetermined value.

In **Fig.1**, **10** is the casing, **11** is the first rotor, **12** is the third fluid heat exchanger, **13** is the vane which also serves as a heat exchange member, **14** is a heat-conductive wall, **15** is a vane, **16** is a nozzle, **17** is the second rotor, **18** is the second fluid heat-exchanger, **19** is a vane, **20** is the second-fluid conduit, **21** is a combined bearing and seal, **22** is a combined bearing and seal, **23** is a second rotor shaft for the delivery of power, and for support of the second rotor, **24** and **25** are supply and return for the third-fluid, **26** is a vent opening in the casing into which a vacuum source may be connected, **34** is a dividing wall, **27** are vanes serving also as heat-exchange members, **28** is a first-fluid passage, **30** is a combined bearing and seal, **31** and **32** are the second-fluid entry and exit points, and **33** is the first rotor shaft.

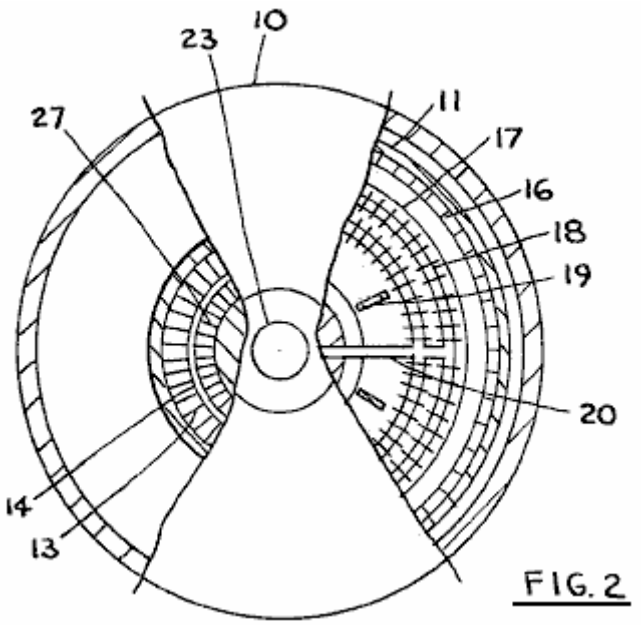


FIG. 2

Fig.2 shows an end view of the unit of **Fig.1** where **10** is the casing, **11** is the first rotor, **17** is the second rotor, **16** are the first-fluid nozzles, **18** is a heat exchanger, **19** are vanes, **20** is a conduit, **13**, **14** and **27** form a heat exchanger for the first-fluid and **23** is the second rotor shaft.

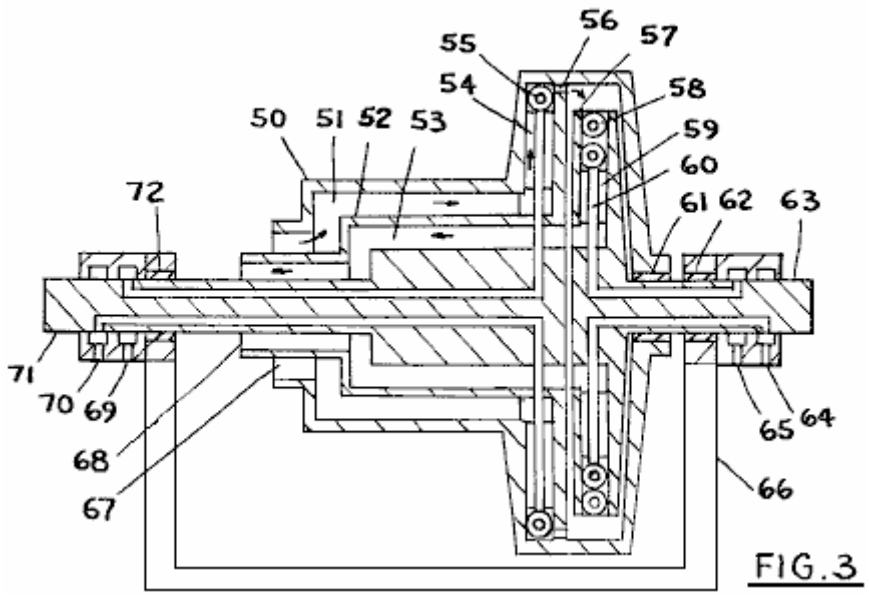


FIG. 3

Fig.3 shows another form of the turbine, where the first-fluid is supplied to the turbine from outside sources thus eliminating the third-fluid heat exchanger. **50** is the first rotor, **51**, **52** and **53** form a heat exchanger for the first-fluid, **55** and **58** are heating heat exchangers for adding heat to the first-fluid and may use a second-fluid at the same temperature or at a different temperature as the heating fluid, **54** are vanes within first rotor, **56** are first-fluid nozzles oriented to discharge forwards, **57** is the second rotor, **59** are vanes, **60** is a conduit for the second-fluid, **61**, **62** and **72** are bearings, **64**, **65**, **69** and **70** are entries and exits for the second-fluid, **63** is the second rotor shaft, **71** is first rotor shaft, **66** is the base, while **67** and **68** are the exit and entry points for the first-fluid.

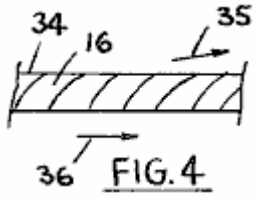


FIG. 4

Fig.4 shows a detail of the first-fluid nozzles where **34** is wall on which nozzles **16** are mounted, **35** is the approximate direction of leaving of the first-fluid, and **36** indicates direction of rotation of first rotor.

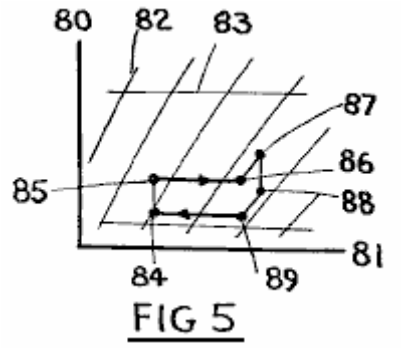


FIG 5

In **Fig.5**, a pressure-enthalpy diagram for the first fluid is shown, with the working cycle for the first-fluid where **80** is the pressure axis and **81** is enthalpy axis, **82** are constant entropy lines, **83** are constant pressure lines, and for the cycle, compression with heat removal, or without heat removal, occurs from **84** to **85**, heat is added from returning first-fluid from **85** to **86**, further compression is from **86** to **87**, then expansion from **87** to **88** and **89**, and heat removal to the first-fluid from **89** to **84**, thus completing the cycle. Heat is normally added between **87** and **88**, from the second-fluid. The heat addition between **85** and **86**, and heat removal between **89** and **84** may be at constant or varying pressure as desired; pressure may be varied conveniently by increasing or decreasing the diameter of the first-fluid to first-fluid heat exchanger, making the heat exchanger tapered.

In operation, the rotors are filled to a desired pressure with a suitable first-fluid, and the first rotor is caused to rotate. The first-fluid is first compressed with heat removal, and then is passed in heat exchange relationship with the inward bound first-fluid with addition of heat, and after this the first-fluid is further compressed and accelerated and after this compression, the first-fluid is passed via nozzles mounted on the first rotor forwards in the direction of rotation, after which the first fluid enters the second rotor's inward extending passages for deceleration, with heat being added to the first-fluid in the second rotor inward passages for reduction of density of the first-fluid. After passing inwards and decelerating, the first-fluid is passed in heat exchange relationship with the outward bound first-fluid, and after that, the first-fluid may be further decelerated, and then the first-fluid enters the outward extending passages of the first rotor thus completing the cycle.

The operation of the open turbine of **Fig.3** is similar to that described, except that the first-fluid is supplied from external sources, and is then returned to said external source, with cooling then being deleted.

The work input to the first rotor is the work required to accelerate the first-fluid, and the work output by the second rotor is the work of deceleration received by the second rotor. The work output by the turbine is the work differential of these two rotors.

The rotational speed of the second rotor may be higher than the rotational speed of the first rotor. To provide for inward flow of the first fluid within the second rotor, the fluid density is reduced by adding heat to the first fluid either within the second rotor, or also within the first rotor.

The addition of heat from the inward bound first fluid to the outward bound first fluid increases the temperature of the first fluid during latter part of compression and during expansion, and thus has the effect of improving the thermal efficiency of the turbine. Also, another effect is the reduction in the needed rotational speed for the turbine rotors, thus reducing the required strength for the rotors, and making the rotors more economical to make and operate.

Working fluids for this turbine are usually gases for the first-fluid, and liquids for the second and third fluids. Gaseous second and third fluids may be also used, and the first-fluid may be a liquid in some instances. Also, the first fluid may undergo a phase change within the turbine, if so desired, when using a suitable fluid. Applications for this turbine include normal power generation service using various heat sources.

The first rotor shaft and the second rotor shaft are normally connected via a power transmission device so that a part of the power produced by the second rotor is used to rotate the first rotor. Starting of the unit is by a starting device.

The vanes of the rotors may be made curved if desired. In many instances, the first rotor vanes may be curved backward to increase compression of the first-fluid, and the vanes of the second rotor may be also curved, to improve performance, and to suit the design and fluid selected. In this connection, the fins for the heat exchangers are considered to be vanes.

The pressure-enthalpy diagram shown in **Fig.5**, is approximate only. This diagram may be varied, depending of the amount of heat added in the second rotor, or in the first rotor, and depending on the specific location of the second fluid and third fluid heat exchangers. In particular, heat may be added to the first-fluid during expansion to make the first-fluid actually increase in temperature; this will normally improve the overall thermal efficiency of the turbine. Also, heat removal by the third fluid may be conducted in places other than that shown in **Fig.1**, as desired.

It should be also noted that the heat addition to the first-fluid may be from sources other than the second fluid, and similarly, some other means may be used to cool the first-fluid other than the third fluid. Such heating sources may include electricity, or other rotors mounted in proximity to this turbine; these will not change the spirit of this invention.

The heat exchanger mechanism for transferring heat from the inward bound first-fluid to the outward bound first-fluid can also be located within the second rotor, and also the entry and exit for the first-fluid into the turbine may be within the second rotor. Such arrangements are not shown specifically in the drawings since they are considered to be within the capabilities of a skilled designer, in view of the descriptions given herein.

Much of this information on Michael Eskeli is taken, with the kind permission of Scott Robertson, from his web site http://www.aircaraccess.com .

James Hardy's Self-Powered Water-jet Generator.

Repeated here from Chapter 2, a device which needs to be in this list of self-powered devices is the ultra simple water-jet generator. There is a video on Google which shows a self-powered electrical water-pump driven, electrical generator at the location: http://www.youtube.com/watch?v=zlinM1wAI5U

This is a very simple device where the jet of water from the pump is directed at a simple water-wheel which in turn, spins an electrical alternator, powering both the pump and an electric light bulb, demonstrating free-energy. What is of particular note is the utter simplicity of this device. It uses off-the-shelf parts almost exclusively and can be constructed by almost anyone.

It should be noted that the implementation shown in this video uses the most basic of turbine blades which must have a very low efficiency, and yet the output power generated is well above the level needed to sustain its own operation. Given well shaped conventional turbine blades of much higher efficiency would appear to raise the performance further, while one would think that using a Tesla Turbine with its simple discs should give a really spectacular performance. However, this may very well not be the case a the irregular, pulsed drive of the wheel will be leading-out additional energy as in the case of the Chas Campbell flywheel and the John Bedini flywheel. As it is, with its present form of construction, this device is already capable of producing additional power able to run other pieces of standard mains equipment.

This is clearly a development platform and it would benefit from having the areas which contain water, fully enclosed, and the electrical diversion from mains power to the output alternator operated by a switch.

Initially, the generator is got up to speed, driven by the mains electrical supply. Then, when it is running normally, the mains connection is removed and the motor/generator sustains itself and is also able to power at least one light bulb. The generator output is normal mains current from a standard off-the-shelf alternator. Power generation could hardly get any more simple than this.

The Heat Pump System of Arthur Cahill and John Scott.

Arthur Cahill and John Scott have patented a heat-pump system which draws heat energy from the surrounding environment and uses that energy to produce mechanical and/or electrical energy for powering a household. Why most people have a refrigerator they are generally not aware that it is a heat-pump and moves three times as much heat from inside the refrigerator compared to the necessary input power (COP=3 but could be up to COP=11 when used differently).

This heat-pump system appears to run without any form of energy input, but the energy comes indirectly from the sun heating the surrounding environment and there is no magic involved. Mind you, when the system runs and provides power, generally, without the need for any fuel, the user can be forgiven for

thinking of it as a fuel-less or self-powered system even though strictly speaking, that is not the case. The inventors have made allowances for unusual conditions where environmental conditions can't provide the temperature difference needed to make the system operate as intended. A liquid or gaseous fuel is provided along with a burner to provide the heat difference if those conditions are encountered.

Here is an extract from their patent:

Patent US 4,309,619 5th January 1982 Inventors: Arthur Cahill & John Scott

SOLAR ENERGY SYSTEM

ABSTRACT:

A dynamic, self-sustaining and self-perpetuating device for the production of motive force by combining cryogenic and thermodynamic principles into one system, keeping the systems separated, two open to atmosphere, the other closed, sealed, pressurised and using special compounded fluids, which when alternately exposed to the heat of atmospheric temperature, then, to the coldness of a liquid or air-cooled condenser, first evaporates, then condenses. Rapid expansion during evaporation produces a high pressure vapour which operates an engine and a generator, which are an integral part of the closed system. Rapid condensation drastically reduces back pressure on the aft side of the engine, and the engine operates on the difference between the two pressures, producing electricity, or, the engine can be used as a direct drive for vehicles or equipment. Built-in safeguards and alternatives are a part of the systems, assuring continued operation despite adverse conditions.

US Patent References:

2,969,637	Converting solar to mechanical energy	Jan 1961	Rowekamp
3,495,402	Power system	Feb 1970	Yates
3,995,429	Generating power using environmental temperature differentials	Dec 1976	Peters
4,110,986	Using solar energy carried by a fluid	Sep 1978	Tacchi
4,214,170	Power generation-refrigeration system	Jul 1980	Leonard

BACKGROUND OF THE INVENTION

1. Field of the Invention

This invention relates to a closed-cycle, sealed, pressurised, energy producing system, utilising the sciences of thermodynamics and cryogenics to convert liquid into gas, then back to liquid.

2. Description of the Prior Art

There is no exact prior art, as cryogenics have been used primarily for air conditioning and refrigeration purposes, and thermodynamic efforts have been directed in the area of low efficiency ocean thermal energy conversion systems. A few attempts have been made to combine some form of cryogenics and thermodynamics, without notable success, mainly using sea water for evaporation and condensing. While using no fuel and requiring little in the area of labour, these ocean thermal energy conversion systems are of necessity, low pressure systems and require large sea going platforms to support the huge turbines and heat exchangers which are necessary to produce reasonable electrical power, resulting in excessive capital costs for minimal electrical output, since such stations only have the ability to service a small portion of the populace along the seaboards. None of these contrivances serve or benefit the populace as a whole, while the whole bears the burden of financing through taxes, or government grants.

Proposals to heat gases and cool gases in an endeavour to improve the efficiency of home heating and cooling systems, have been previously advanced, some operating on the heat pump principle. All such previous proposals and inventions have had one thing in common, they all plug into the Utility Company's electric line to obtain the electricity necessary to run the system.

In cryogenics the knowledge that certain liquids, when heated, change into a high pressure vapour, which is the heart of all air conditioning and refrigeration systems, has been known for many years. Thermodynamics were pioneered by the 19th century French physicist Nicholas Carnot. Attempts have been advanced during the years to harness one or the other and sometimes both, for the purpose of heating and cooling, resulting

in the invention of the heat pump in a much earlier year, but none of the systems yet devised for use by the general public have been able to operate without the use of an outside source of electricity, or, fuel, such as oil, or gas fired boilers, resulting in a considerable consumption of fuel and a cataclysmic effect on the earth's environment.

SUMMARY OF THE INVENTION

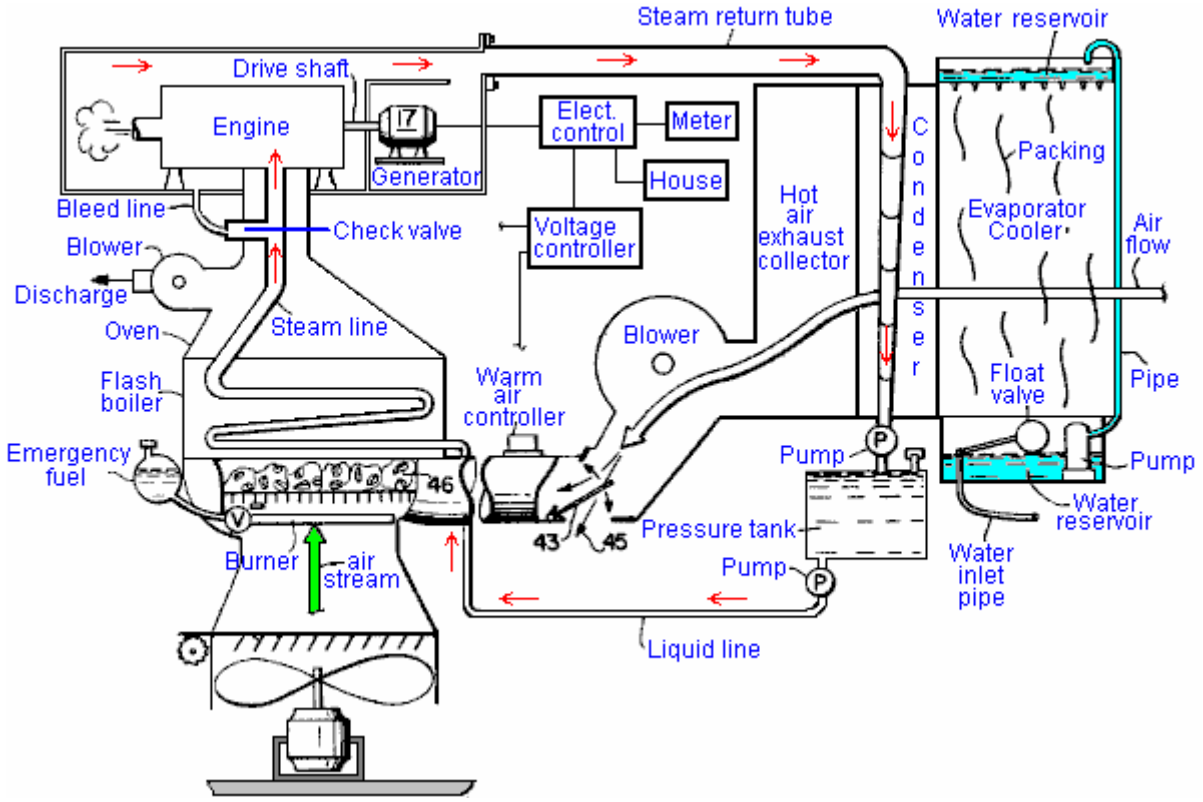
In accordance with the present invention, the device will operate in hot sunshine; on cloudy days without sunshine; during rainstorms; during snowstorms; during changes in temperature from day to night; during changes in seasons from winter, to spring, to summer, to fall; when it's cold, even below zero; for the power generated is that energy produced when a compounded fluid changes form, first to vapour, then back to liquid, by application of controlled temperatures within the sealed cycle. Thus, by combining cryogenics and thermodynamics into one system, keeping the two separated, one open to atmosphere and the other closed, sealed and pressurised, and by using fluids specifically compounded for the given area, or climate, these fluids, when exposed to atmospheric temperatures, in accordance with the kinetic theories of matter, gases and heat, provide the kinetic energy to operate an engine.
The condenser can be either liquid or air cooled, although for the embodiment depicted herein, the condenser is air cooled.
Generally speaking, there is up to an approximate 2.5 PSI increase associated with each degree of temperature rise in most cryogenic fluids and gases. However, using commercially available fluids, here are a few examples:

Temperature F.	Fluid	Pressure in psi.
125	R-22	280
125	R-500	203
125	R-502	299
125	R-717	293
80	R-13	521
80	R-22	145
80	R-500	102
80	R-502	160
80	R-700	128

You will notice that R-13 at 80° F. produces 521 psi., or 35.4 times atmospheric pressure and at 125° F. would produce thousands of psi. At 95° F. R-22 produces 185 psi., or a thrust on a five inch diameter piston of 3,633.4 pounds. Even at 30° F., with R-22 a thrust of 583.2 pounds is obtained. R-13 at 30° F. produces 263 psi. or 5,112.7 pounds of thrust on a five inch diameter piston. The pressures are there by using the Casco Perpetuating Energy System, utilising proprietary formulated liquids for the area and temperatures to be encountered. It is not intended that any of these mentioned fluids will be used in the present invention; the comparisons being made herein with popular and well known liquids, for comparison purposes only.

THE INVENTION

The present invention relates to a device to supply pollution free power to operate a generator for the producing of electrical power, or, to supply power as a direct drive to a shaft, transmission, clutch, differential or the such, the invention being independent of outside sources of power such as electricity supplied by a Public Utility Company. This is not to be considered perpetual motion, as will be explained later in the text.



Proprietary liquids, specifically compounded to produce the desired results in a given area, or climate, under pressure in the reservoir to keep them in a liquid state, will, when directed through tubes exposed to atmospheric temperature, change from a liquid state into a gaseous state (from here on referred to as steam), such conversion resulting in tremendous expansion, thus producing high pressure steam with which to drive the engine, or turbine.

It is a general object of this invention to provide a pollution free device for public utilisation, that will produce electrical power or, direct drive power. One object is to produce electrical power with which to heat, cool, cook, run electrical appliances and light a home. Another object of the invention is to provide industry with a pollution free means to not only heat, cool and light factories, but to supply electrical or direct drive power with which to operate factory equipment. A still further object of the invention is to provide a pollution free source of power to propel cars, trains, trucks, buses, equipment, steamships, aeroplanes, and other forms of transportation, without the use of fossil fuels as the primary power source. It is also an object of the present invention to provide the means whereby individuals can produce electrical power for their own use, and as a small power production, sell their surplus electrical power to the local electricity utility company. A further object of the invention is to provide a self-sustaining, small apparatus, that provides ample power from the engine to operate a car or other conveyance, or to supply sufficient electrical power to a home or factory, without having to plug the apparatus into a Public Utility electrical supply.

BRIEF DESCRIPTION OF THE DRAWING

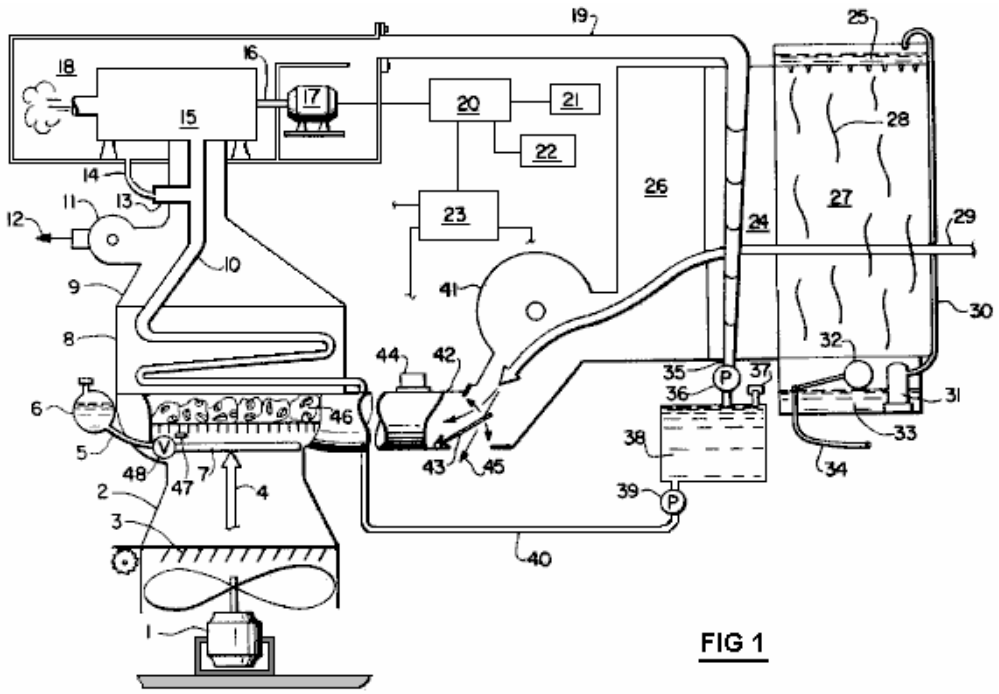


FIG 1

Fig.1 is a partially sectioned schematic view of the system:

DETAILED DESCRIPTION OF A PREFERRED EMBODIMENT

In the drawing, **Fig.1**, the invention is shown in a preferred embodiment for home use. The liquid pump **39**, pumps the cryogenic fluid from the pressurised liquid reservoir **38**, into liquid line **40**, where the fluid gravity feeds into flash boiler **8**. Pump **39** also prevents back pressure from flash boiler **8** from entering the pressurised liquid reservoir **38**, and since the pressure within steam line **10** and liquid line **40** are equal, the cryogenic fluid gravity feeds down liquid line **40** into steam line **10**. The fins on flash boiler **8** are heated to atmospheric temperature by air stream **4**, which converts the fluid within steam line **10**, inside flash boiler **8** into high pressure steam. To maintain the pressure during the passage of steam to the engine **15**, steam line **10** from flash boiler **8** is housed inside the oven **9**, which is exhausted when necessary by discharge **12** from centrifugal blower **11**. Constant temperature within oven 9 and flash boiler 8 is maintained by the admission of fresh atmosphere via air stream 4 passing through the finned flash boiler **8** and up through oven **9**. Centrifugal blower **11** is thermostatically controlled to exhaust air within oven **9** which has cooled below a predetermined temperature. Any excess pressure within steam line **10** is by-passed through check valve **13** and bleed line **14** into the exhaust collector box **18**, thus, a pressurised, closed system is maintained, which, once charged, unless an accident damages or ruptures a line, should not have to be replenished. Pressure within the exhaust collection box **18** will be less than the inlet pressure from steam line **10** to engine **15**, because the condenser **24** is at a lower pressure, as is steam return tube **19**, than PSI input to engine **15** from steam line **10**, thereby creating a suction on the back of the exhaust collector box **18**.

Airstream **29**, which has been cooled by evaporator cooler **27**, flows over the finned surfaces of condenser **24**, instantly lowering the temperature of the steam within condenser **24** below a predetermined condensation point, thus turning the steam back to a liquid, such conversion and instant reduction of volume within condenser **24** causing a pressure reduction at the back of engine **15**. This condensed liquid drains down into liquid coil return **35**, where it is immediately pumped into the pressurised liquid reservoir **38** by liquid pump **36**.

While under pressure in pressurised liquid reservoir **38**, the fluid is maintained in a liquid state regardless of exterior temperature, until it is re-circulated back into the system by liquid pump **39**, through liquid line **40** to flash boiler **8**, where it again converts into steam.

The capacity of evaporator cooler **27** and packing **28** is sufficient to cool intake air stream **29** to a predetermined temperature below atmospheric temperature at any given time, even with humidity rise at night, or during rainstorms, or just during high humidity weather. This temperature differential is maintained as the atmospheric temperature rises and falls, with an anti-freeze liquid being added to the water in the evaporator cooler when temperatures drop below 32° F., to keep it from freezing.

Additional warm air to augment air stream **4** is obtained by directing the warm air exhaust collected in hot exhaust air collector **26**, by centrifugal blower **41**, through T-assembly **42**. Air flow through T-assembly **42** is directed and controlled by the warm air control mechanism **44**, which regulates flow-damper **43** to either exhaust through exit **45** into the atmosphere, or alternatively, recirculate air flow **29** through restricted-T **42**, thereby compressing and further heating air flow **29** before injecting it through flash boiler **8**. Warm air control mechanism **44** also controls fan **1** and louvers **3**, selecting the optimum heat from either T-assembly **42** or variable-Venturi **2**, to perpetuate the system.

The evaporator cooler **27** has an inlet water pipe **34** which supplies cold water from the normal house water supply, or well, (neither shown). The bottom water reservoir **33** is kept at a constant level of water by float valve **32**. The water is pumped by pump **31** up tube **30** into the top water reservoir **25**, where it runs through the perforated bottom of reservoir **25** down on to the packing **28**, keeping packing **28** constantly wet, which cools air flow **29** as it is drawn through packing **28** and over the fins of condenser **24** by the partial vacuum in the hot air exhaust collector **26**, such partial vacuum being created by centrifugal blower **41** exhausting the air from the hot exhaust air collector **26** slightly faster than air flow **29** can replace it.

As long as there is warmer air on the flash boiler **8** side of the system than the cooler condenser **24** side of the system, this device will continue to operate and produce electricity and/or power. The heat recirculating system and the use of three separate, distinct sub-systems within the system, one sealed, permits the system to perpetuate itself. As mentioned above, this device is not to be considered perpetual motion, for somewhere in the ranges of temperature differentials and weather conditions, there may be a no-man's land where the system could shut down, then the variable venturi **2**, in conjunction with motor and fan **1** and louvers **3**, will come into use automatically upon a signal from warm air controller **44** and **12** voltage controller **23**, and be used for a period of time. Motor and fan **1** forces an air stream **4** upward through the variable Venturi **2**, with air stream **4** controlled by warm air controller 44 and 12 volt controller 23, adjusting louvers 3. As air stream 4 is forced through the restriction of variable Venturi **2**, air stream **4** is compressed as it funnels up the narrowing walls of the variable Venturi **2**, such compression causing the air to heat, thus overcoming possible deadlocked or identical temperatures between the condenser **24** and air flow **4**. This slight temperature rise in air stream **4** will enable the system to perpetuate itself until the atmospheric temperature itself changes enough to permit a continued operation. Since fan and motor **1** is run by battery power from 12 volt supply **23**, even though the batteries are constantly being charged during operation, the batteries may become exhausted because of an extended time the fan and motor **1** are run, then, or, if for any other reason the system starts to run down, a small burner **7**, operating on liquid or gaseous fuel **6**, through line **5** and valve **48**, is ignited by spark mechanism **47** and supplies the heat necessary to support and perpetuate the system until atmospheric temperature and condensing temperature permits the system to operate normally. Fired clay heat retainers **46** are arranged on the grate within burner **7**, to retain heat.

The cryogenic system is charged by filling with liquid under pressure through fill pipe **37**. Re-charging, if necessary, is done the same way. Engine **15** turns drive shaft **16**, which turns generator **17**, producing electrical power (110V or 220V) via electric control system **20** which passes the electricity into three channels:

One: To 12-volt controller **23**, to run the electrical parts of the system and keep the batteries charged.
Two: To the house **22**, to supply the electricity with which to cook, run appliances, light, heat and cool the home.
Three: All remaining electricity is channelled through utility meter **21** into a local utility company's electrical line for sale and use elsewhere.

Patrick Kelly

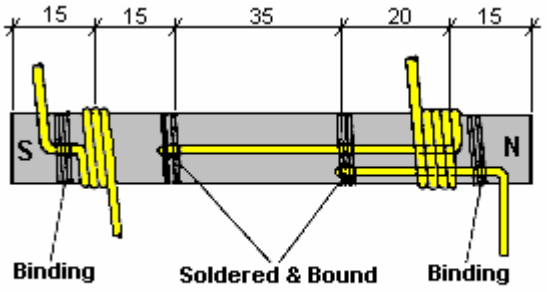
http://www.free-energy-info.tuks.nl/

Chapter 9: Passive Systems

The Devices of Hans Coler.

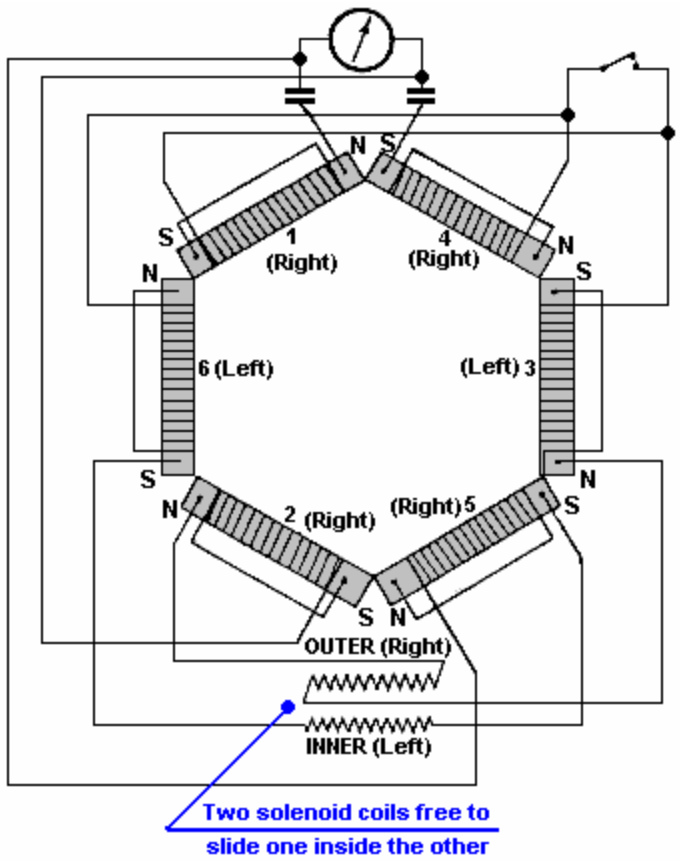
A German naval captain called Hans Coler invented a COP>1 generator in 1925. He called this device the 'Stromerzeuger' and for a few watts from a dry battery it provided 6 kW continuously. He was refused development support because it was "a perpetual motion machine".

Hans also invented a passive device which he called the 'Magnetstromapparat'. His unit required very careful and slow adjustment to get it operating but when it started it continued on test in a locked room for three months of continuous operation. Nobody, including Hans, seems any too sure how this device works but it is presented here in case you wish to research it further. It comprises six bar magnets wound as shown here. Some are wound in a clockwise direction when looking at the North pole and these are called "Right" those wound in an anticlockwise direction are called "Left":

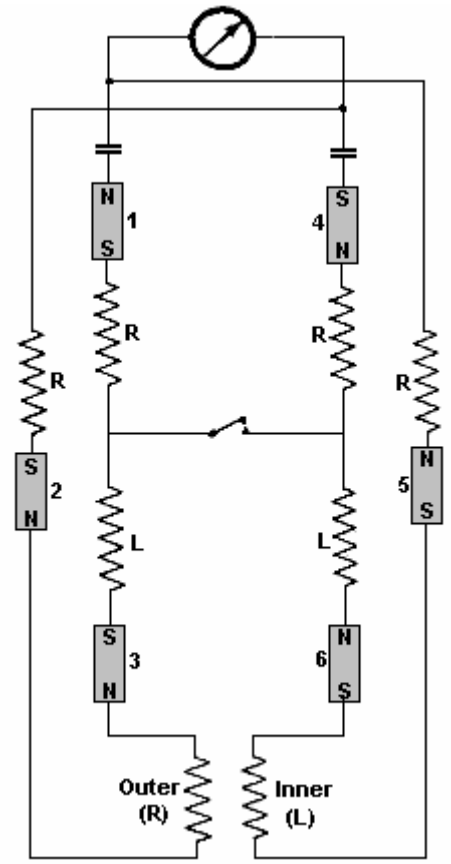


Note: There is paper insulation between magnet and coil

These six magnets are arranged in a hexagon and wired as shown here:



And the schematic diagram is:



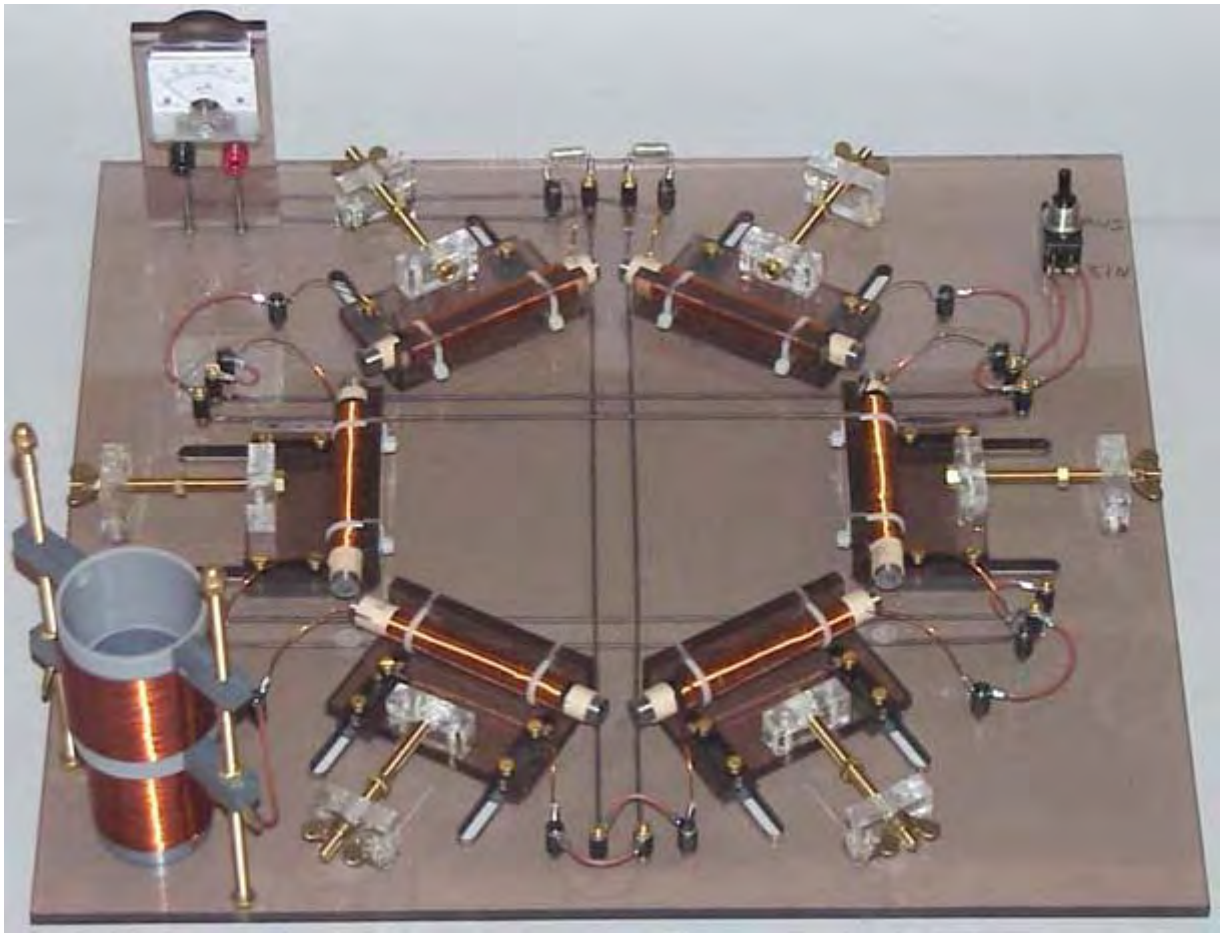
One extremely interesting feature of this passive device is that it has been witnessed producing 450 mV for several hours; it was capable of developing up to 12 Volts. The witnesses were quite sure that it was not picking up radio or mains input. So, what was it picking up? With magnets as the key component, it seems clear that it is the zero-point energy field which is being accessed, but clearly, the access represents a vanishingly small percentage of the actual power available

To operate the device, the switch is left in the open position, the magnets are moved slightly apart and the sliding coil set into various positions with a wait of several minutes between adjustments. The magnets are then separated still further and the coils moved again. This process is repeated until at a critical separation of the magnets, a voltage is developed. The switch is now closed and the process continued more slowly. The voltage then builds up to a maximum which is then maintained indefinitely. The position of the apparatus in the room and the orientation of the device had no effect on the output.

The magnets were selected to be as nearly equal in strength as possible and the resistance of the magnet and coil were checked after winding to make sure they were as nearly equal as possible (about 0.33 ohms).

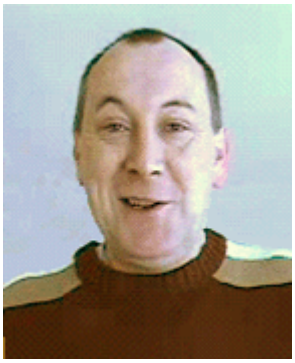
A very neat construction of the Coler 'Magnetstromapparat' by an unknown German experimenter is shown below - I'm afraid without permission as I have no idea who he is or how to contact him to ask his permission. The quality of workmanship is impressive and the result is a very professional looking device. Notice the sliding coil arrangement at the bottom left with one coil being positioned closely inside another and held in place where the experimenter chooses.

Recently, a construction video has come on offer at http://www.kohlermagnet.com/ and it shows step by step, one method of implementing this design. The website design screams 'con artist' and the final tuning is shown is such an apparently simple and quick way that it suggests that it could well be faked (by AAA batteries embedded in the underside of the very thick base board and wired in underneath), but the construction instructions make this a video which is worth watching. Scaling up the output from one tiny bulb to kilowatts of power is by no means the simple matter implied in the video, and I personally doubt that a passive device of this type could ever produce kilowatts of output power – at twelve volts you are looking at 165 amps of current requiring copper wire with a diameter of more than 7 mm to carry it.



Thomas Trawoeger's Electricity-Producing Pyramid.

One thing which is quite certain, and that is the fact that at this point in time, our technical know-how has not yet encompassed the zero-point energy field properly. It is by no means obvious how the Hans Coler device operates, and if we understood the technology properly, we would be able to say with certainty, exactly how and why it operates, and ways to improve it would be obvious. As it is, all we can do is look at it and wonder, possibly try a few experiments, but the bottom line is that we do not yet understand it. This is the normal situation in the early days of any new field of technology.



It is also quite usual for pioneers in any new field to encounter a good deal of opposition, mistrust, and generally disheartening treatment from other people. That is certainly the case for Thomas Trawoeger from Austria, who has progressed well in the passive energy field. He has suffered repeated web-based attacks with his display material being destroyed and web sites being made inoperable.

So, what makes some people so afraid of Thomas? The answer is that he is experimenting with shapes. That doesn't sound too terrible does it? Well, it certainly bothers some people, which suggests that he must be on the verge of uncovering a mechanism for drawing serious amounts of power from the zero-point energy field.

Thomas is by no means the first person to examine this area, but he is one of the first to consider drawing serious amounts of electrical energy from the local environment using shape and an appropriate detector. Obviously, this is the same area that Hans Coler was investigating, and it appears that Thomas has managed to tap a continuous 8 watts of electrical energy using a wholly passive device.

As we are not all that familiar with this type of technology, we tend to dismiss it as being a “crackpot” area, not worthy of investigation by serious scientists. It is actually, very far from being that in reality, and it just indicates our serious lack of technical understanding if we dismiss it out of hand. Two hundred years ago, the idea of a television set would definitely have been considered a “crackpot” pipe dream, far, far away from reality. Today, any schoolchild would be horrified at the thought of a TV set being considered “crackpot”. So, what has changed? Only our level of technology, nothing else. In another two hundred years time, when the zero-point energy field is fully understood, people will look back with a smile at the though of people like us who didn’t know how to draw any amount of energy, freely from the environment, and they will laugh at the thought of burning a fossil fuel to produce energy from a chemical reaction. That, of course, does not help us at all in this time of our ignorance, and we still have to deal with the sort of people who thought that the horse-drawn cart would never be superseded.

The scientific method has been established for a long time now. Essentially, observations are made, experiments are performed and a theory is produced which fits all of the known facts. If additional facts are discovered, then the theory needs to be modified or replaced by another which includes all of the new facts. Established scientists find it difficult to adhere to the scientific principle. They are afraid of losing their reputation, their job or their funding and so are reluctant to investigate any new facts which indicate that some of their best-loved theories need to be revised. Fortunately, not being in the business, we can take new facts on board without any problem. In the light of what certain shapes do, this is just as well.

Let us see if we can put this in perspective. Consider an intelligent, well-educated person living several hundred years ago. Looking skyward at night, he sees the stars. At that time, the theory was that the stars were fixed to a ‘celestial sphere’ which rotates around the Earth. That was a perfectly good theory which matched the known facts of the time. In fact, the concept matches the observed facts so well that some people who teach Astro Navigation to sailors still find it to be useful in teaching the subject today. If you told the average person of those days, that the stars were not very small but very large indeed, that the Earth is orbiting around the Sun and in fact, the Sun is one of those ‘tiny’ stars, then you would have been considered one of the ‘lunatic fringe’.

Next, if you were to tell that person that there were invisible forces passing through the walls of his house and even through him, he would most certainly rate you as a bona fide member of the ‘lunatic fringe’. However, if you then took several compasses into his house and demonstrated that they all pointed in the same direction, he might start to wonder.

Now, just to really establish your membership of the ‘lunatic fringe’ you tell him that one day there will be invisible rays passing through the walls of all buildings and that these rays will allow you to watch things happening on the other side of the world. Finally, to complete the job, you tell him that there is a substance called uranium, and if he were to carry a piece around in his pocket, it would kill him by destroying his body with invisible rays.

Today, school children are aware of, the Solar System, magnetic lines of force, television and X-rays. Further, as the scientific theory has caught up, these children are not considered part of the ‘lunatic fringe’ but this knowledge is expected of them as a matter of course. The only thing which has changed is our understanding of the observed universe.

At the present time, we are faced with a number of observations which do not fit in with the scientific theories of some of the current educational establishments. If we consider these things seriously, we run the risk of being considered part of the ‘lunatic fringe’ until such time as scientific theory catches up with us again. So be it, it is better to examine the facts than to pretend that they don’t exist.

Present theory has worked well enough up to now, but we need to take on board the fact that since it does not cover all of the facts, it needs to be extended or modified. So, what observed facts are causing a problem? Well:

1. In Quantum Mechanics it has been found that some pairs of particles are linked together no matter how far apart they are physically. If you observe the state of one of the pair, the state of the other changes

instantly. This happens far, far faster than the speed of light and that does not fit neatly into present theory.

2. If a substance is cooled down to Absolute Zero temperature, it should be completely motionless, but that is not the case as movement can be observed. This movement is caused by external energy flowing into the frozen material. That energy, observed at Absolute Zero temperature is called 'Zero-Point Energy'. So where does that fit into the theory?

3. There are several devices which are self-powered and which are capable of powering external loads. These things appear to act in defiance to the Law of Conservation of Energy.

4. The Aspden Effect (described below) indicates that current theory does not cover all of the facts.

5. It is now known and fully accepted by science that more than 80% of our universe is composed of matter and energy which we cannot see.

6. Even though our Sun is losing some five tons of mass per second, it radiates more energy than can be accounted for by the fusion of the amount of matter which would cause this loss of mass.

7. The inner core of the Earth is hotter than present theory would expect it to be.

These things indicate that there is something in our universe which is not properly covered by current theory. The present theory thinks of space as being a volume which contains no matter, other than perhaps, a tiny amount of inter-stellar dust. And while space can be traversed by radio waves and many other types of radiation, it is essentially empty.

This concept is definitely not correct. All of the odd observed facts suddenly fit in if we understand that there is an additional field which streams through all of space and passes unnoticed through all matter. This field is composed of particles so tiny that they make an electron appear enormous. These particles may in fact be the 'strings' of String Theory. What is sure, is that this stream of matter contains virtually unlimited energy.

It is the energy seen at Absolute Zero as it is continually streaming in from outside the cold area. It flows to us from every direction and the sun being a major source of it, augments the flow we receive during the daytime. This accounts for the variations seen by T. Henry Moray during the night when the energy he was picking up decreased somewhat.

This matter stream acts like a very dense gas except for the fact that effects in it have effectively zero propagation time. This accounts for the widely separated particles having what appears to be simultaneous reactions to a stimulus. Einstein's idea of the speed of light being an absolute maximum is definitely wrong, as has been demonstrated in the laboratory.

In the early stages of investigating a new field, it can be quite difficult to work out how to approach it, especially if the field is entirely invisible and can't be felt. The same situation was encountered in the early days of magnetism as lines of magnetic force are not visible and cannot be felt. However, when it was observed that iron was affected by magnetism, a mechanism was discovered for displaying where the invisible lines are located, by the use of iron filings. Interestingly, the presence of an iron filing alters the lines of magnetic force in the area as the lines "have a preference for" flowing through the iron. Also, the iron filings used in school demonstrations do not show the actual lines of magnetic force correctly as they themselves become tiny magnets which alter the lines of force which they are supposed to be showing.

We are still in the early stages of investigating the Zero-Point Energy field, so we have to consider anything which has an effect on this invisible field. One observed effect was found by Harold Aspden and has become known as the 'Aspden Effect'. Harold was running tests not related to this subject. He started an electric motor which had a rotor mass of 800 grams and recorded the fact that it took an energy input of 300 joules to bring it up to its running speed of 3,250 revolutions per minute when it was driving no load.

The rotor having a mass of 800 grams and spinning at that speed, its kinetic energy together with that of the drive motor is no more than 15 joules, contrasting with the excessive energy of 300 joules needed to get it rotating at that speed. If the motor is left running for five minutes or more, and then switched off, it comes to rest after a few seconds. But, the motor can then be started again (in the same or opposite direction) and brought up to speed with only 30 joules **provided** that the time lapse between stopping and restarting is no

more than a minute or so. If there is a delay of several minutes, then an energy input of 300 joules is needed to get the rotor spinning again.

This is not a transient heating phenomenon. At all times the bearing housings feel cool and any heating in the drive motor would imply an increase of resistance and a build-up of power to a higher steady state condition. The experimental evidence is that there is something unseen, which is put into motion by the machine rotor. That "something" has an effective mass density 20 times that of the rotor, but it is something that can move independently and its movement can take several minutes to decay, while in contrast, the motor comes to rest in a few seconds.

Two machines of different rotor size and composition reveal the phenomenon and tests indicate variations with time of day and compass orientation of the spin axis. One machine, the one incorporating weaker magnets, showed evidence of gaining magnetic strength during the tests which were repeated over a period of several days.

Nikola Tesla found that uni-directional electric pulses of very short duration (less than one millisecond) cause shockwaves in this medium. These Radiant Energy waves passed through all materials and if they strike any metal object, they generate electrical currents between the metal and ground. Tesla used these waves to light glass globes which had just one metal plate. These lights do not have to be near the source of the Radiant Energy waves. He discovered many other features of these 'longitudinal' waves but one which is of particular interest is that when using his famous Tesla Coil, the waves produced visible streamers which showed what they were doing. What they were doing was running up the outside of the long inner wire coil, not through the wire, mark you, but along the outside of the coil, and when they reached the end of the coil, they continued on out into the air. Interestingly, Tesla believed that this flow of energy "preferred to run along the corrugations of the outside of the coil". That is to say, somewhat like magnetic lines showing a preference for running through iron, this energy field shows a preference for flowing along certain physical shapes.

Thomas Henry Moray developed equipment which could tap up to fifty kilowatts of power from this field. There are two very interesting facts about Moray's demonstrations: Firstly, the valves which he used to interact with the field, had a corrugated cylindrical inner electrode - an interesting shape considering Tesla's opinion on the corrugated outer surface of his coil. Secondly, Moray frequently demonstrated publicly that the power obtained by his equipment could flow uninterrupted through sheet glass while powering light bulbs. Quite apart from demonstrating that the power was definitely not conventional electricity, it is very interesting to note that this power can flow freely through materials. I venture to suggest that Moray's power was not flowing through the wires of his apparatus but rather it was flowing along the outside of the wires, or perhaps more accurately, flowing along near the wires.

Edwin Gray snr. managed to draw large amounts of power from a special tube designed by Marvin Cole. The tube contained a spark gap (like that used by Tesla) and those sparks produced Radiant Energy waves in the Zero-Point Energy field. He managed to collect energy from these waves, very interestingly, by using perforated (or mesh) cylinders of copper surrounding the spark gap. His 80 horsepower electric motor (and/or other equipment such as light bulbs) was powered entirely from energy drawn from the copper cylinders while all of the electrical energy taken from the driving battery was used solely to generate the sparks.

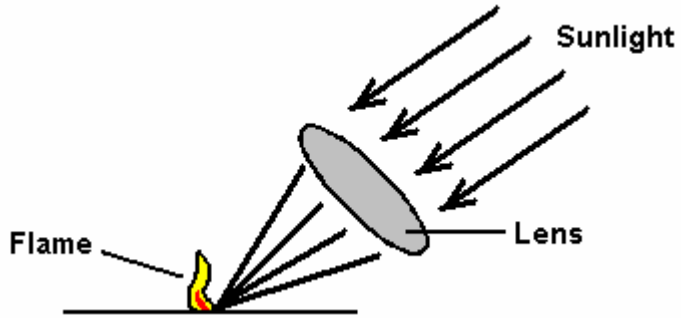
It is very interesting to note that Tesla, Moray and Gray all indicate that corrugated or rough-surface cylinders seem to direct the flow of this energy. Dr Harold Aspden also indicates that once the field is set in motion in any locality, it tends to continue flowing for some time after the influence which is directing it is removed.

Please remember that we are starting to examine a new field of science, and while we know a very limited amount about it at this point in time, at a later date, every schoolchild will be completely familiar with it and find it hard to believe that we knew so little about it, at the start of the twenty-first century. So, at this time, we are trying to understand how energy can be extracted from this newly discovered field. The indications are that the physical shape of some objects can channel this energy.

If you think about it, you suddenly realise that we are already familiar with shape being important in focusing energy. Take the case of a magnifying glass. When the sun is high in the sky, if a magnifying glass is placed in just the right position and turned in just the right direction, then it can start a fire. If the principles behind what is being done are not understood, then the procedure sounds like witchcraft:

1. Make a specially shaped object with curved faces, out of a transparent material

2. Discover the ‘focal-length’ of the object
3. Wait until Noon
4. Place some kindling on the ground
5. Position the object so that it looks directly at the sun
6. The kindling will catch light without you even having to touch it.



Sounds like something out of a book on magic, doesn’t it? Well, you need to know all about that if you want to pass any basic physics examination, and it comes in under the title of “Optics”. Please notice that the shape of the lens is vital: it must have a convex face on both sides. Also, the positioning is vital, the lens must be exactly its focal length away from the kindling material: a little too near or a little too far away and it just does not work. Magic? Well it may seem like it, but no, it is just scientific understanding of the nature of radiation from the sun.

Take the case of a satellite dish. This familiar object needs to be an exact shape to work well. It also needs to be made of a material which reflects high-frequency radio waves. Make one out of wood and it will look just the same but it will not work as the TV transmission will pass straight through the wood and not be reflected on to the pick-up sensor connected to the television set.

However, obvious and all as this is, it still did not cut any ice with the patent office in Czechoslovakia on the 4th November 1949. A radio engineer called Karel Drbal turned up with a patent application for a cardboard pyramid shape which kept razor blades sharp and was promptly told to get lost. The patent authorities demanded that he have a theory to show how the device worked. Karel was not particularly put out, and spent years investigating before he determined a theoretical basis for the device. He returned to the patent office, much to the disbelief of the Chief Patent Officer. He was granted his patent, not because his theory was compelling, but because the Chief Patent Officer took a pyramid home and tested it with his own razor blades. When his practical tests confirmed that the pyramid did exactly what Karel claimed, he was granted Patent No. 91304, “Method of Maintaining Razor Blades and the Shape of Straight Razors” and here is a translation:

Republic of Czechoslovakia
Office For Patents And Inventions
Published August, 1959
Patent File Number 91304

The right to use this invention is the property of the State according to Section 3, Paragraph G, Number 34/1957
Karel Drbal, Prague
Method of Maintaining Razor Blades and the Shape of Straight Razors.

Submitted 4 November, 1949(P2399-49)
Patent valid from 1 April, 1952

The invention relates to the method of maintaining of razor blades and straight razors sharp without an auxiliary source of energy. To sharpen the blades therefore, no mechanical, thermal, chemical or electrical (from an artificial source) means are being used. There are various mechanical sharpening devices being used up to now, to sharpen used razor blades. The blade is sharpened by crude application of sharpening material, which always results in certain new wear of the blade during the sharpening process. Furthermore, it is known that the influence of an artificial magnetic field improves the sharpening of razor blades and straight razors, if their blades are laid in the direction of the magnetic lines.

According to this invention, the blade is placed in the earth's magnetic field under a hollow pyramid made of dielectric material such as hard paper, paraffin paper, hard cardboard, or some plastic. The pyramid has an opening in its base through which the blade is inserted. This opening can be square, circular, or oval. The most suitable pyramid is a four sided one with a square base, where one side is conveniently equal to the height of the pyramid, multiplied by π / 2. (which is pi or 3.14 / 2). For example, for the height of 10 cm, the side of 15.7 cm is chosen. The razor blade of a straight razor is placed on the support made also of dielectric material, same as the pyramid, or other such as cork, wood, or ceramics, paraffin, paper, etc. Its height is chosen between 1/5 and 1/3 of the height of the pyramid, this support rests also on a plane made of dielectric material. The size of this support should be chosen as to leave the sharp edges free. Its height could vary from the limits stated above. Although it is not absolute necessary, it is recommended that the blade be placed on the support with its sharp edges facing West or East respectively, leaving its side edges as well as its longitudinal axis oriented in the North / South direction. In other words to increase the effectiveness of the device it is recommended lie in essence in the direction of the magnetic lines of the horizontal component of the earth's magnetism. This position improves the performance of the device, it is not however essential for the application of the principle of this invention. After the blade is properly positioned, it is covered by the pyramid placed in such a way that it's side walls face North, South, East, and West, while its edges point towards North-West, South-West, South-East, and North-East.

It is beneficial to leave a new blade in the pyramid one to two weeks before using it. It is essential to place it there immediately after the first shave, and not the old, dull one. But it is possible to use an old one, if it is properly resharpened. The blade placed using the method above is left unobstructed until the next shave. The west edge should always face West. It improves the sharpening effect.

Example: When this device was used, 1778 shaves were obtained using 16 razor blades, which is 111 shaves per blade on the average. The brand used was "Dukat Zlato" made in Czechoslovakia. The lowest count was 51, the highest was 200. It is considered very easy to achieve up to 50 shaves on the average. (for a medium hard hair).

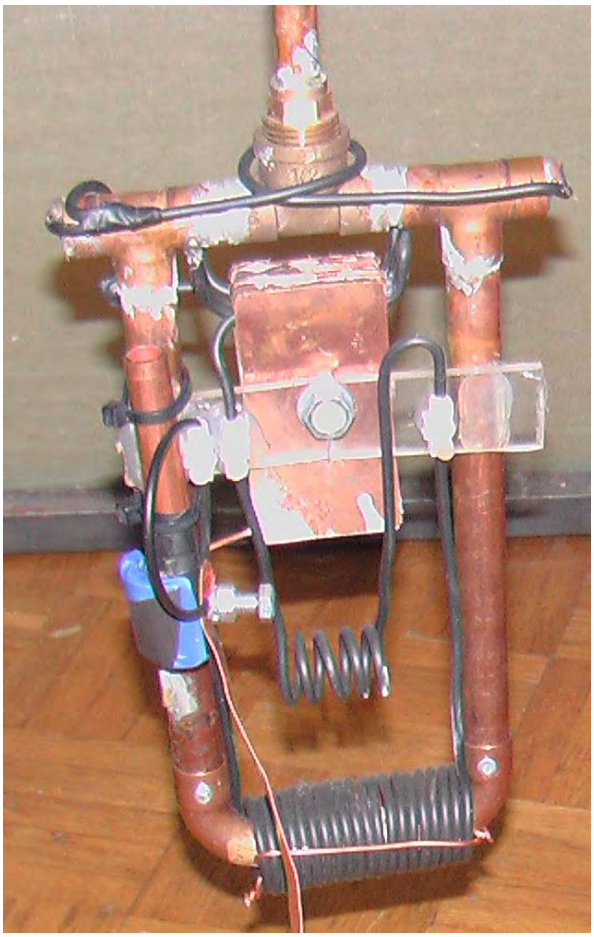
The following shows how the invention could save both valuable material and money. One of the razor blades mentioned above, weighs 0.51 grams. We will consider 50 shaves on average when placed in the pyramid against 5 shaves when it is not. It is obvious that the number of shaves, degree of wear, and the ability to regenerate the dull edge depends on the quality of the material, quality of sharpening process, and hardness. ....given that the numbers are averages and could be in fact much better. In the course of the year one therefore uses 73 razor blades without the aid of the pyramid while only eight razor blades while using the pyramid. The resulting annual saving would be 65 razor blades or 33.15 grams of steel per person.

Only the pyramid shape has been used for this invention, but this invention is not limited to this shape, as it can cover other geometric shapes made of dielectric material that was used in accordance with the invention. And that this shape also causes regeneration of sharp edges of shaving blades by lowering of stresses and reducing the number of defects in the grids of crystal units, in other words recovering and renewing the mechanical and physical properties of the blade.

This is interesting, as it confirms by independent test that a pyramid shape produces an effect, even if it is not possible to say with absolute certainty what exactly the effect is and how exactly the pyramid shape manipulates that energy.

Thomas Trawoeger has produced a video of a pyramid which he constructed. The video commentary is in German and it shows a computer fan being operated when connected to his pyramid which looks like this:

Sceptics will immediately say that as there are wires connected to the device, that the power for the fan is being fed through those wires, even though they appear to be connected to monitoring equipment. This is possible, but in my opinion, it is not actually the case. The pick-up used is shown here:

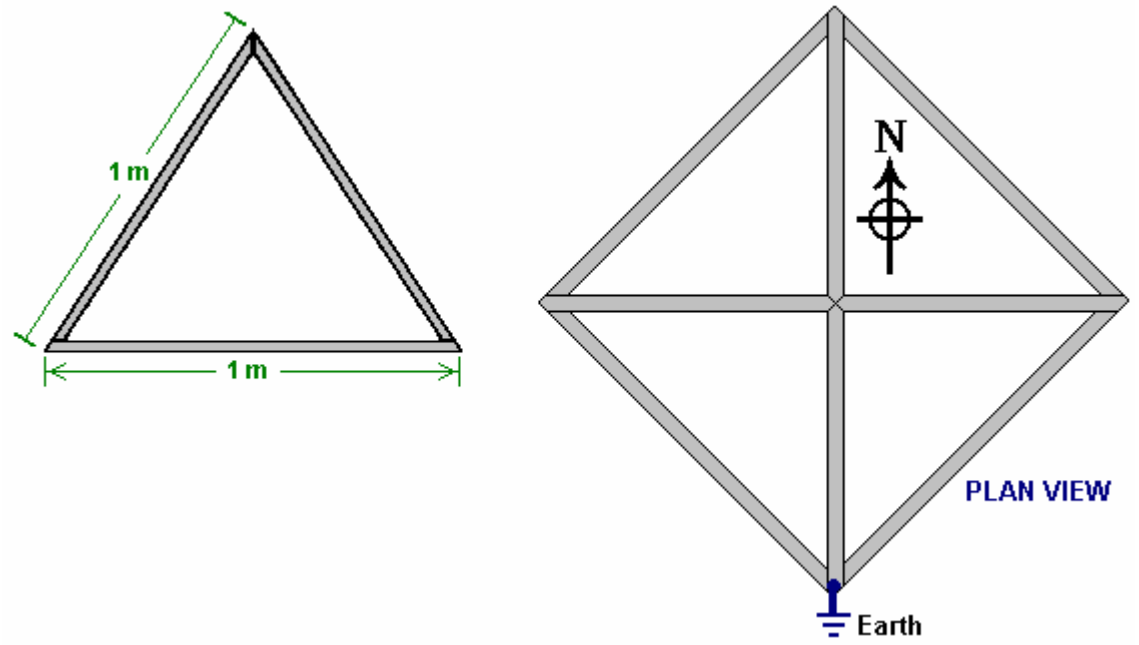


It should be remembered that these pictures are quite old and all inventors keep working on their inventions in an effort to improve their operation and to investigate the effects caused by alterations. At the close of 2007 the design has progressed considerably and now features a number of most unusual things ranging from construction to orientation. The http://www.overunity.com/index.php/topic,695.300.html forum is working on replicating this design thanks to the generosity of Thomas Trawoeger who speaks German and the exceptional work of Stefan Hartmann who has produced an English translation and who hosts the web site.

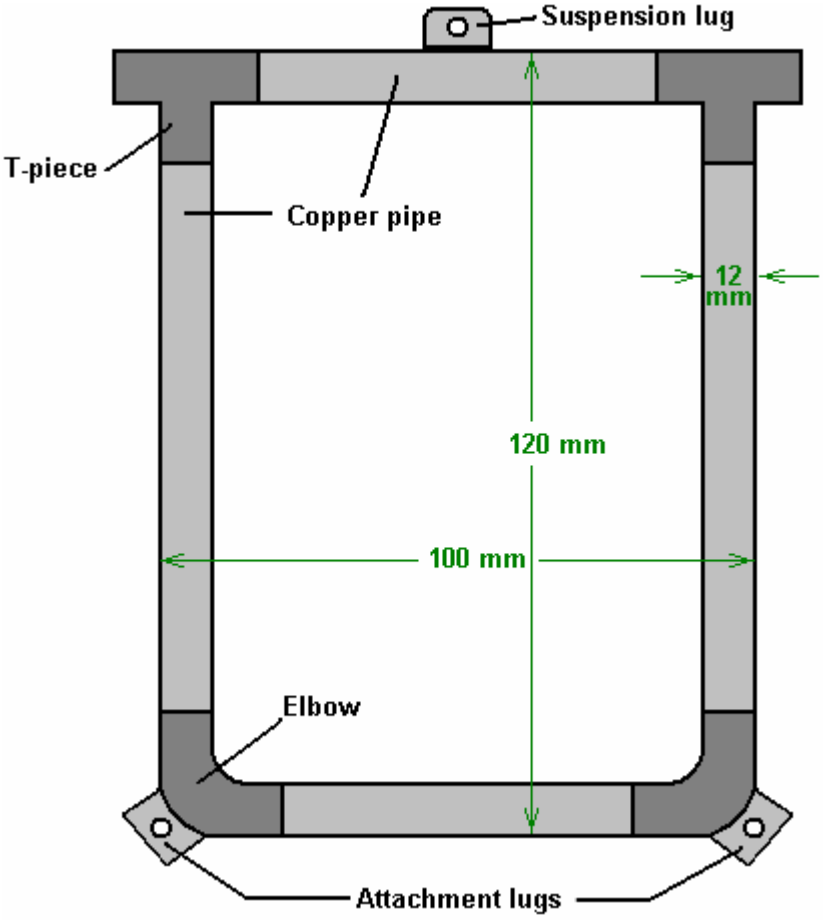
The following is an attempt to present the basic information from that forum in a clear and concise manner, but I recommend that you visit and contribute to the forum if you decide to experiment with this design.

The frame of the pyramid is not the same shape as the well-known Egyptian pyramids and has a sloping face some 5% longer than those in Egypt. The materials used in constructing the pyramid are very important. The frame is made of 20 mm x 20 mm x 2 mm square-section steel tube. While the exact size of the pyramid is not critical, the exact proportions are critical. The base must be exactly square, with each side of the base being exactly the same length, 1 metre in this case. The sloping sides are exactly the same length as the base pieces being 1 metre long also. Eight one-metre lengths of steel section will therefore be needed for building the frame.

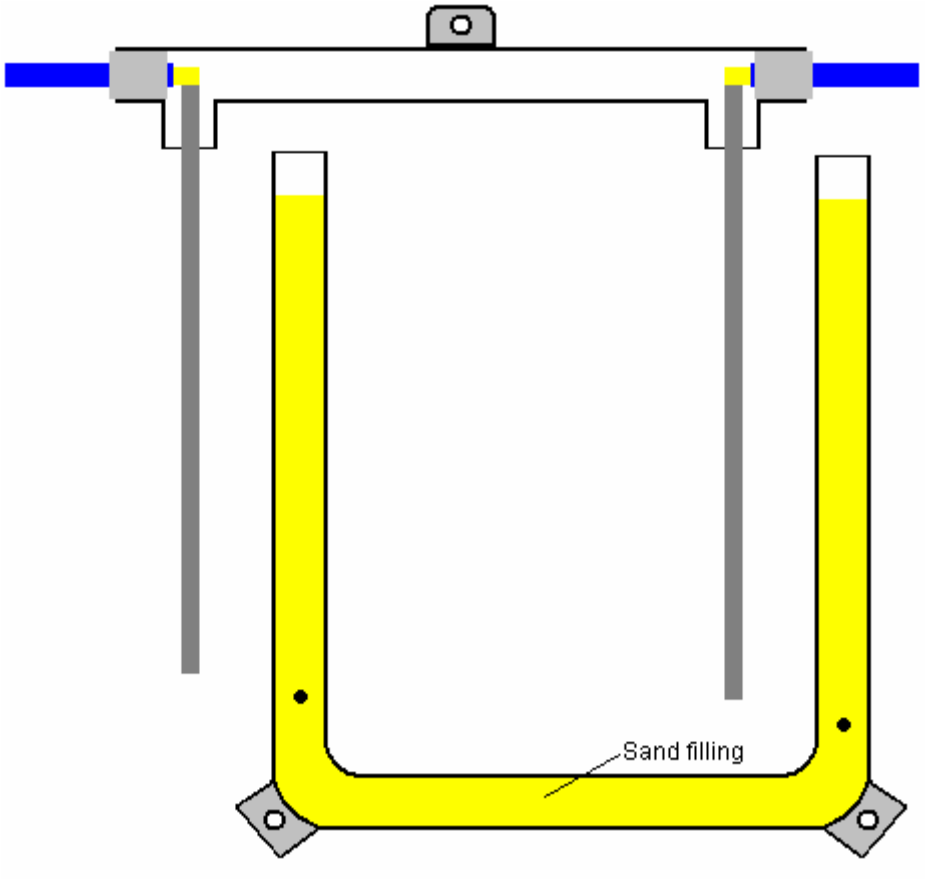
The sides of the pyramid need to be covered with a rigid sheet and here again, the material used is critical, with only gypsum/paper boards (plasterboard with no foil) being satisfactory - other materials just don't work. If no sides are added, then the pyramid is very difficult to adjust to get proper operation. When the frame has been constructed, its is positioned in a most unusual way being forty-five degrees away from the conventional positioning of a pyramid. This sets this pyramid so that one pair of corners face North - South, and the frame should be connected to a good electrical ground as shown here:



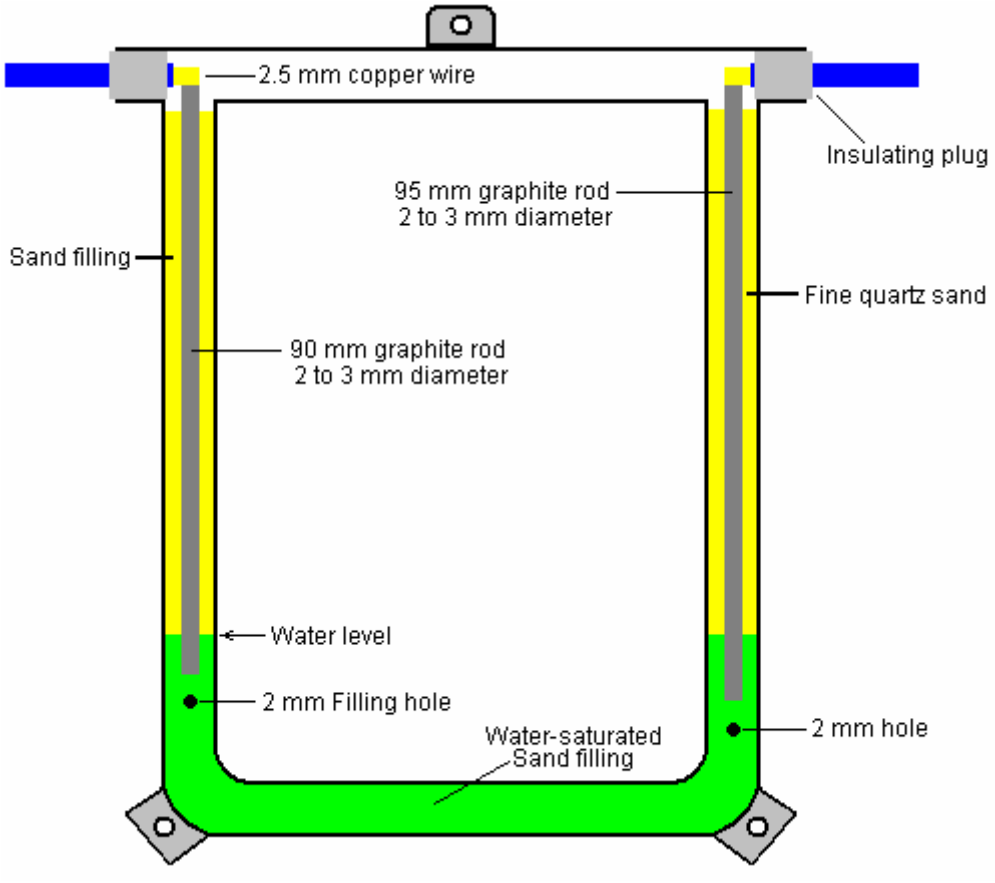
The pick-up is constructed from 12 mm outside diameter copper pipe and fittings and is hard soldered together. It has an overall size of 120 mm x 100 mm hard soldered together as shown here:



This frame of copper piping is not assembled as shown straight off as there is a requirement for a long graphite rod, 2 to 3 mm in diameter, to be positioned vertically inside each vertical leg of this frame and that can't be done after assembly. So the bottom section is assembled as one piece, and the top section is assembled separately with the graphite rods sticking down out of the T-sections, held in place by their wires and insulating plugs. The graphite rods can be bought from art materials supply shops.

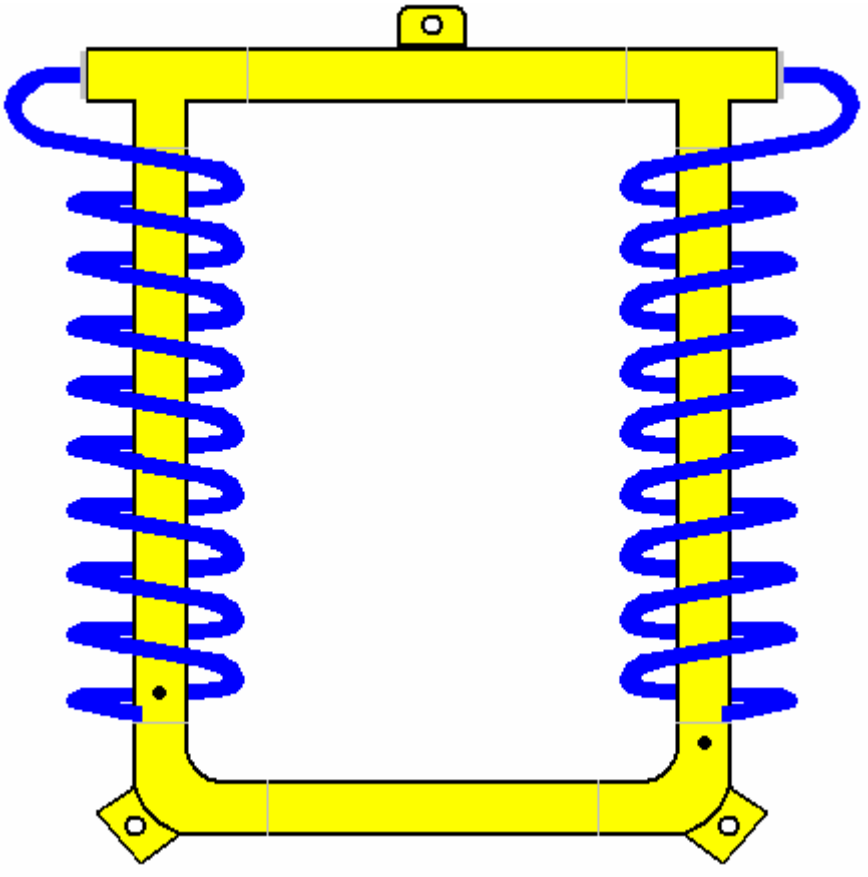


The very fine filter-grade quartz sand filling for the tubes is inserted and the graphite rods carefully positioned so that they do not touch the side walls of the vertical copper tubes, and the two parts joined by hard soldering:

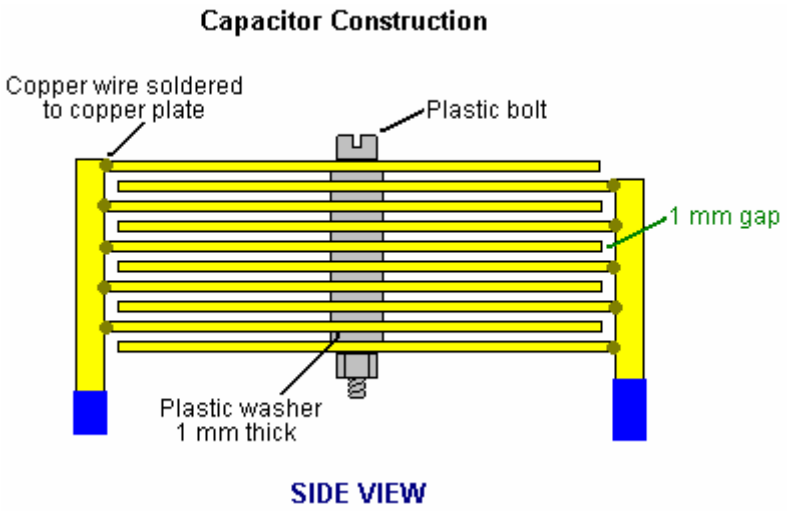


The left hand side hole in the copper pipe is used to inject a 5% salt / water solution, using a hypodermic syringe, until the water starts to come out of the hole at the right hand side. The right hand side hole is 5 mm lower down than the one on the left.

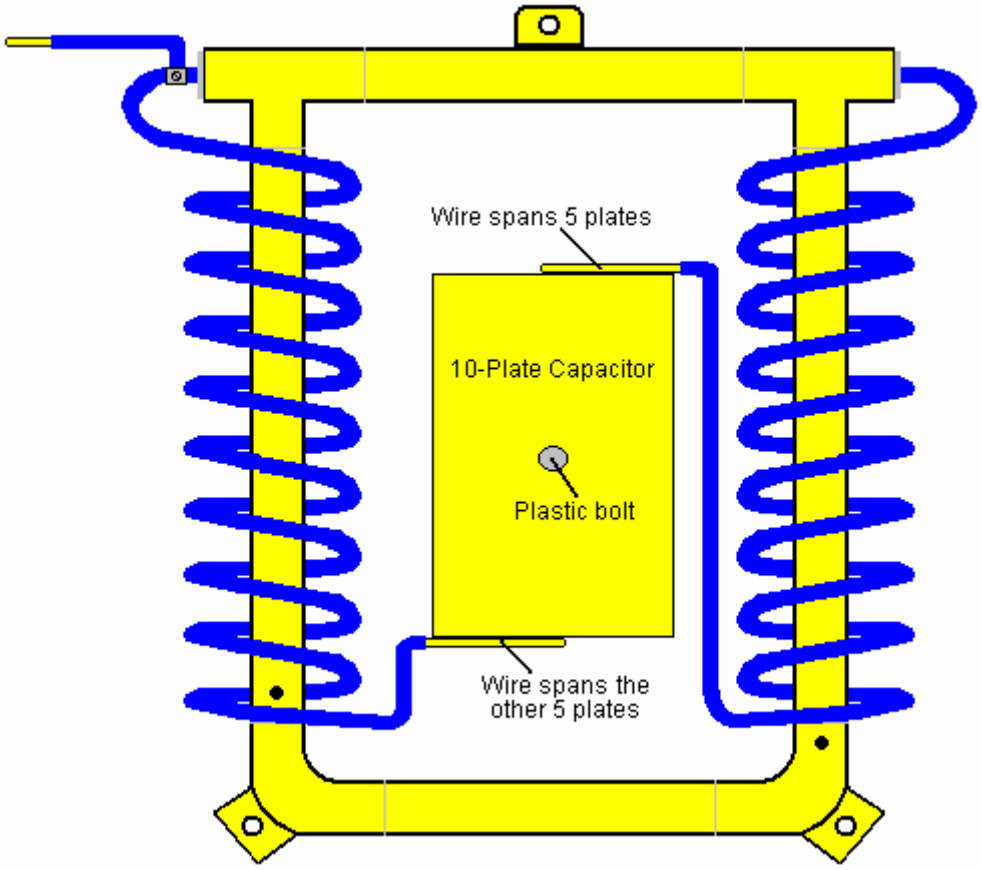
Next, the wires are bent around to produce a 9-turn coil with a 25 mm diameter, around the vertical copper pipes. The windings are in opposite directions on the opposite sides of the frame:



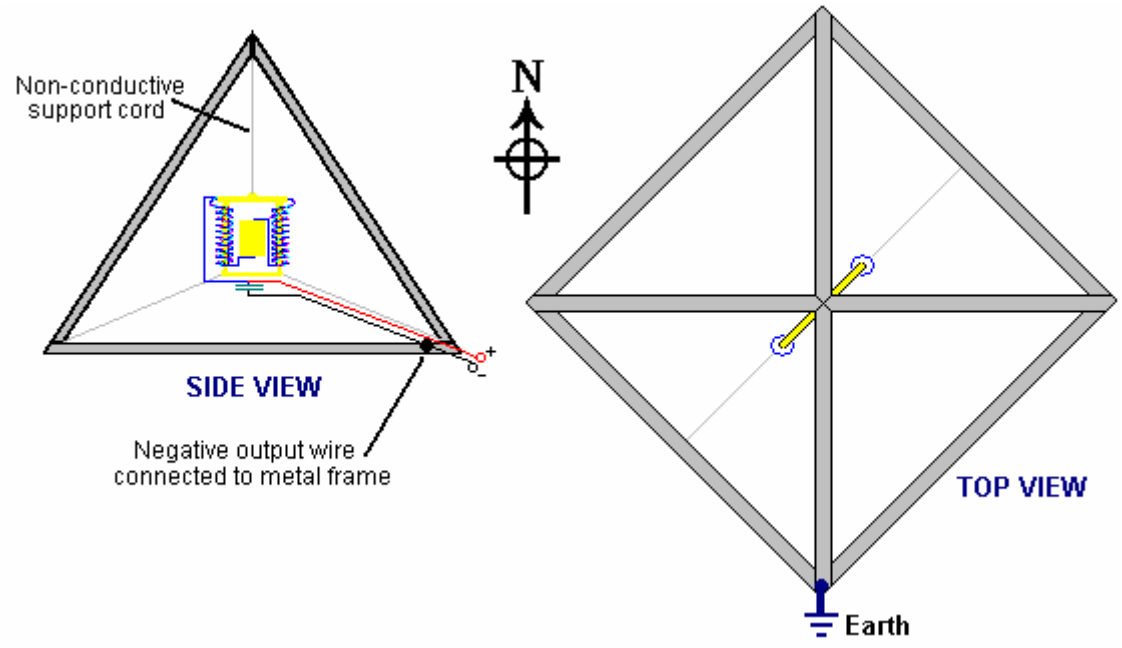
Next, a ten-plate capacitor is made from copper sheets 1 mm thick. As copper is very expensive, the copper plates can be produced from spare lengths of copper pipe, cut along the axis and flattened careful to produce a smooth, unmarked surface 70 mm x 35 mm in size. The plates are stacked and accurately aligned, and a hole is drilled 1 mm off-centre. Then each alternate plate is turned around to produce two sets of plates bolted together with a 6 mm diameter plastic bolt, 1 mm thick plastic washers and a plastic nut. A plastic threaded rod and a plastic nut can be used instead of a plastic bolt. Because the hole is not quite central, the plates stick out at each end, giving clearance for attaching the plates together with the copper wire coming out of the copper pipe framework:



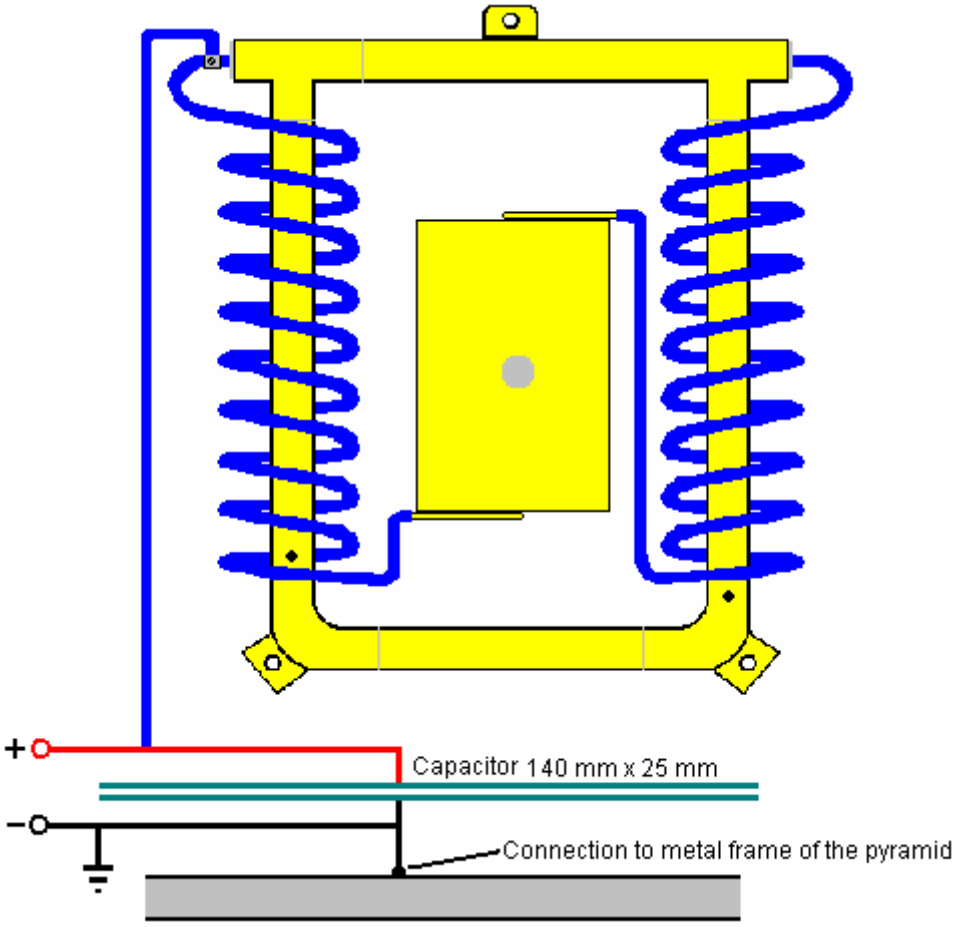
The capacitor is positioned inside the copper pipe frame and held in place by the strength of the 2.5 mm thick copper wire coil around the vertical pipes in the frame:



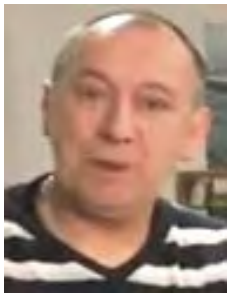
The pick-up sensor is now attached to the pyramid frame. Using a non-conductive cord, it is suspended by the top lug and it's orientation controlled using the lower two lugs. The positioning in the pyramid is unusual, being North-East to South-West, as is shown here:



Next, a second capacitor is constructed from 1 mm thick copper sheet. Again, sections of copper pipe can be used after being cut along their long axis and carefully opened out and flattened. This capacitor is just two plates 140 mm x 25 mm spaced 1 mm apart (one inch = 25.4 mm).



A voltmeter can be used to check the exact alignment of the pyramid. There is a video (with a commentary in German, at http://www.secret.tv/player_popup.php?id=1307723&movieid=1308850 showing an earlier version of this pyramid set-up driving an electrical fan taken from a computer). If this device interests you, then you should join the enthusiast research and development forum mentioned earlier.



In June 2011, Thomas issued instruction videos which show how to construct, use and troubleshoot his newer design of pyramid. These videos are in English and they are very detailed and instructive. One person has split those two very long videos and placed them on YouTube as a series of thirteen smaller videos. In them, he states that the salt water in the design above is actually counter-productive and should not be used.

I will not attempt to show the construction details described in those videos as the information is very extensive, but a few comments may be appropriate here. Thomas refers to a 'wheel' but unless I am mistaken, he means a solid plastic rod of circular cross-section. The coil wire which he uses is 1.5 mm diameter copper wire with plastic insulation. With an inner coil of 104 turns, that can produce 1.5 amps at 14 volts, which is 21 watts, and with no input power being provided by the user, that is a COP of infinity. However, if I understand what he is saying, he positions his pyramids North-South (unlike the design above) and more importantly, at a particularly good point on what he calls a 'water-line' which I take to be a ley line located by a dowser. That may be due to the fact that he lives in a village in Austria which is a long way from the Equator, which, according to Joseph Cater's analysis of pyramids, reduces their effectiveness.

In this new design, Thomas uses 20 watts of power from a Citizen's Band radio transmitter, boosted by an amplifier, and applied to his detector tube while he fills it very slowly with fine-grain, high-quality quartz sand. This signal orientates most of the quarts grains and probably replicates most of Thomas Henry Moray's detector which allowed such high power extraction from a simple aerial. In this design of Thomas' no earth wire is needed for operation. An earth wire is provided, but this is for the protection of the user and is not part of the energy gathering system. Thomas agrees with Joseph Cater, that the energy being tapped flows out of the top of the pyramid.

Thomas' website which is in German is http://www.comshop.tv/

Thomas picks a fairly small size of pyramid for his demonstration tutorial. It is made using eight pieces of steel channel, each piece being exactly one meter long. He welds these together but remarks that bolting them together is perfectly all right provided that each piece makes good electrical contact with the pieces which it touches as the whole frame acts as a single component in his design.

Thomas also talks about 'welding' the inner copper pipe to one end cap, but what he means is 'soldering' the joint as he spreads resin on the join, heats it with a gas torch and then runs plumber's solder around the join.

Thomas' video series was removed from the web but has been replaced here:

Part 1: http://www.youtube.com/watch?v=QmngHEhu3wI
Part 2: http://www.youtube.com/watch?v=gNrUHIwtqBY
Part 3: http://www.youtube.com/watch?v=EBrcFiO20GY
Part 4: http://www.youtube.com/watch?v=nTeehORmh0s
Part 5: http://www.youtube.com/watch?v=qs1O3YKkMl4
Part 6: http://www.youtube.com/watch?v=hiY4TJlIRRs
Part 7: http://www.youtube.com/watch?v=5MEp08P_vJs
Part 8: http://www.youtube.com/watch?v=PwJK81eW_0k
Part 9: http://www.youtube.com/watch?v=ohzxjmhv3W0
Part 10: http://www.youtube.com/watch?v=u0LQRJbi240
Part 11: http://www.youtube.com/watch?v=2MFRRHuv5S8
Part 12: http://www.youtube.com/watch?v=WkVd3viiQCU
Part 13: http://www.youtube.com/watch?v=QLqUwM-PTok

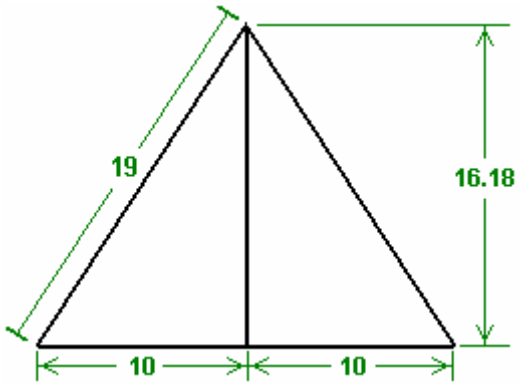
* * * * * * *

Antoine Bovis' Discoveries.

Confirmation of the dehydrating effect of a pyramid was provided by the Frenchman Antoine Bovis who went on holiday to Egypt in the 1930s and visited the Great Pyramid which was constructed exactly in the North - South direction (almost certainly not by accident) and built to an accuracy of 0.01% or better. He discovered that a number of small animals had wandered into the pyramid, got lost and starved to death. The really interesting point was that all of these animals had been mummified through dehydration and none of the bodies had rotted away. When he returned home, he built a model pyramid with base edges three feet long.

He found that his pyramid duplicated the dehydration effect. He, and others who followed him, investigated the effect of pyramids. They found:

1. The best shape is that which matches the dimensions of the Great Pyramid, whose faces slope at an angle of 51 degrees, 51 minutes and 10 seconds. Pyramids with other slopes will work, but not quite as well. If you would like to make one yourself and test the effects, then each of the four sides can be cut from stiff cardboard to these proportions:



So if the base length is to be 20 units, then the height at the mid point of the base will be 16.18 units.
If the base length is to be 25 units (mm, cm, inches, or whatever), then the height should be 20.22 units.
If the total base width is to be 30 units, then the height at it's mid point should be 24.27 units.
If the total base width is to be 35 units, then the height should be 28.32 units, and so on.
The right-angled triangles formed have a height of 1.618 (the "golden ratio") times the base width.

Just cut out four of the triangles and tape the edges together. It would be a good idea to add a square base piece (or triangular gussets) to ensure that the base is exactly square and not skewed.

Thomas Trawoeger states that the exact angle is not particularly important and his latest designs have angles of about sixty degrees.

2. There is no need for the pyramid faces to be solid, provided that there are four base sides and four sloping edges. Having solid sloping faces gives a slight improvement and Thomas considers that at least three of the sloping faces should have panels, the material type being important..

If building a framework pyramid, then the dimensions for the four base pieces and the four sloping edges would be:
Base: 20, length of the sloping edges: 19
Base: 25, length of the sloping edges: 23.76
Base: 30, length of the sloping edges: 28.52
Base: 35, length of the sloping edges: 33.27 and so on.

3. The best material from which to construct the pyramid is copper, but as it tends to be rather expensive, almost any other material can be used: plastic piping, timber laths, steel alloy pipes, wire, etc. Giving the pyramid a sheet-copper cap which runs down about 5% of the face length, giving a short solid face on the open framework also gives a slight improvement. Les Brown states that the best material is sheet iron plated with gold, though that sounds very expensive.

So, what can a pyramid do? Well, nothing, actually, except for directing and possibly concentrating and focusing the Zero-Point Energy field. Perhaps the question should be ‘what effects are caused by using a pyramid?’.

Well, as seen above, Flavio Thomas Trawoeger has managed to get a continuous electrical output via a pyramid for a period of at least thirty days. I understand that he uses a magnet just as an on-off switch, but having a magnet as part of the pick-up makes a lot of sense as the magnetic dipole of any magnet has a distinct effect on the zero-point energy field. The low-tech investigators have noted that an effect caused by a pyramid may be repeated for maybe nine times in a row, and then inexplicably, one day it will not work. They surmise that the effect may be caused by magnetic variations due to solar flares or the like. They may well be right in this as they are not using a magnet but just simple cardboard, or more frequently, simple

frame pyramids. This area is wide open to investigation with very low-tech apparatus and passive electronic components.

What has been found repeatedly:

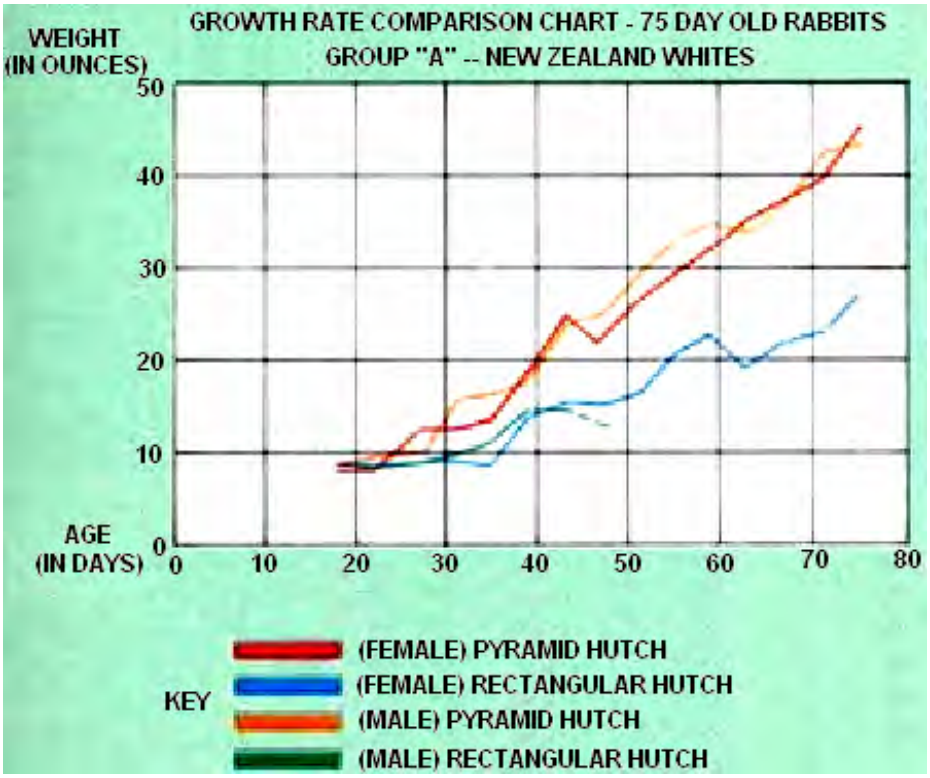
James Brock's Experiments.

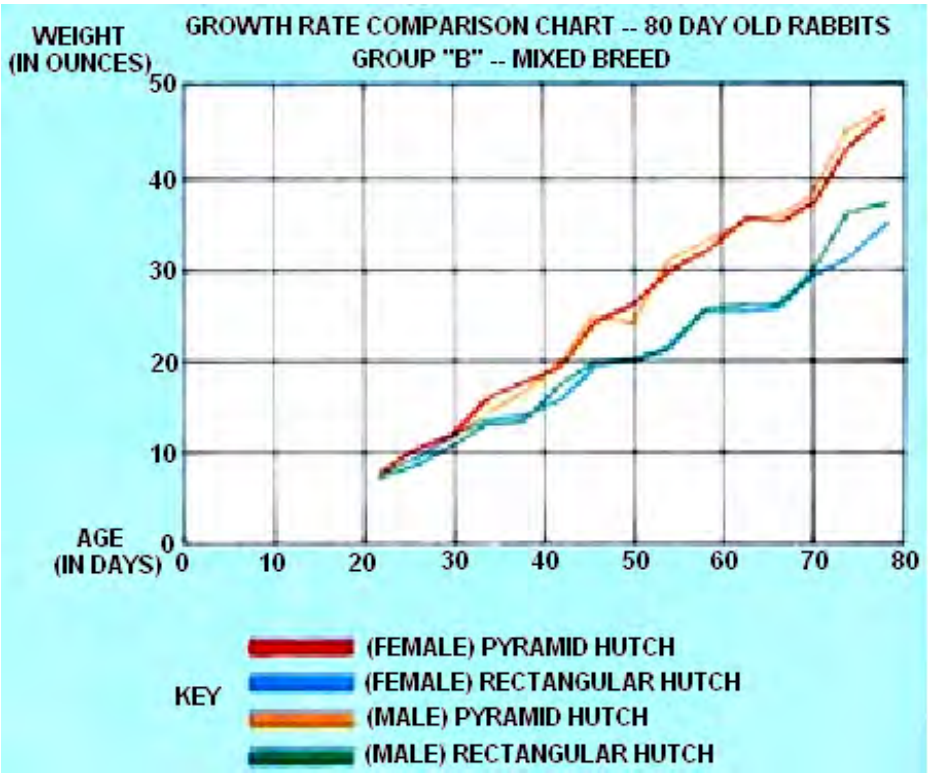
1. Living things placed under a pyramid shape are boosted in health and growth. You can test this easily for yourself by taking two identical plants or animals and keeping one under a pyramid and one outside the pyramid. An example of this is given on the website: http://www.motherearthnews.com/Sustainable-Farming/1977-11-01/Raising-Rabbits.aspx where James Brock of Texas reports on tests he has run on a group of rabbits. It would be incredibly easy to fake this kind of information, so you need to make up your own mind on the validity, and ideally, run some simple tests of your own. James states that he built a pyramid-shaped hutch with 4-foot long sloping edges out of timber, and a rectangular hutch:

Each of the hutches had a transparent door. He then borrowed eight rabbits aged about 20 days old, taken from two different litters and placed them in matched groups of four in each hutch, and fed them equally, weighing them every four days.

By the end of the experiment, 57 days later, the rabbits which had been housed in the pyramid hutch weighed an average of 46.5 ounces, compared to an average of 34.5 ounces for those in the rectangular hutch. That is, the rabbits in the pyramid hutch were nearly 35% heavier and side by side they looked like this:

James presents the results like this:





James invites you to run this test for yourself to verify that this does indeed occur. It should be noted that as the test ran over a period of 57 days, any days lost through magnetic variation would not have been detected by him.

2. Pyramid users also state that they find the following effects on a consistent basis (provided that the pyramid is kept away from strong electromagnetic fields, so do not put a pyramid on top of a TV set or a refrigerator):

(a) Fruit is preserved. When a purchase of fresh fruit or vegetables is made, if they are placed under a pyramid for about an hour and then stored as they normally would, it is said that they stay fresh for at least twice as long as normal and the flavour is enhanced. It is believed that unhelpful micro-organisms are killed by the pyramid. If the fruit and vegetables are kept indefinitely under the pyramid they eventually dry up instead of rotting.
(b) Food quality is enhanced. If frozen meat, fish or fowl is thawed out under a pyramid, the quality of the meat is said to be noticeably improved.
(c) Coffee quality is improved. If a cup of coffee is left under a pyramid for about twenty minutes, it is said to gain a much more mellow flavour. Leaving ground coffee or a jar of instant coffee under a pyramid over night is also said to change it so the coffee made from it is of a much higher quality.
(d) A glass of wine placed under a pyramid for twenty minutes is said to undergo a distinct change with great improvement seen in both the taste and the aroma. Other alcoholic drinks are also said to be improved by this process.
(e) A twenty to thirty minute treatment of fruit juices is said to reduce the acidic “bite” of the drink, and in many cases, alter the colour of the juice.
(f) Any item pickled in vinegar, such as olives and pickles, gain a greatly enhanced natural flavour and are greatly mellowed by the process.
(g) The rapid growth of mould on Cheddar cheese can be overcome by the cheese being kept under a pyramid at normal room temperature. It is recommended that the cheese be wrapped in plastic to reduce the rate at which it dries out.
(h) Rice and wheat can be kept in open jars under a (twelve-inch open frame wire) pyramid for at least four months without any form of deterioration or infestation by insects or flies - which are repelled by the energy inside the pyramid. A test was run outdoors with a six-foot base pyramid with food placed in the centre to attract ants. It was found that ants heading for the food followed a curved path out of the pyramid without ever reaching the food.
(I) Water left under a pyramid is altered. Cut flowers placed in it tend to last 30% longer than normal while growing plants watered with it grow more strongly and are hardier. The water appears to hold the

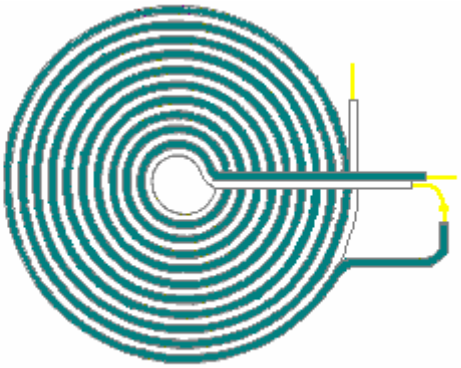
energy indefinitely, a glassful takes twenty minutes, a quart (two pints) takes one hour and larger amounts should be left over night. Animals given the choice of pyramid water or untreated water almost always choose the treated water.

3. In the 1940s, Verne Cameron of America discovered that the beneficial pyramid energy could be transmitted. He placed a pyramid at each end of a row of plants, connected a wire to the apex of each pyramid and ran the wire underneath the plants. He placed a clump of steel wool on the wire under each plant. The pyramids were, aligned North--South and he found that even better results were obtained if the row of plants was also aligned in a North--South direction.

4. There are reports of instances where dogs suffering from old age, lameness and hair loss have been cured and rejuvenated in about six weeks by the use of a pyramid.

I suggest that the Great Pyramid in Egypt was most definitely not built as a burial place but that the chamber inside it was used to treat people with large amounts of the energy picked up by the shape of the pyramid. It is also likely that the pyramid was used as a communications device, but that is outside the scope of this document.

The really important thing is that there is clearly an energy field (perhaps the ZPE field) which flows continuously, is very beneficial to life and which can be tapped to produce unlimited motive power without the need for any kind of input from us. Just like the early discovery days of radio waves, TV signals, X-rays, Gamma rays, etc. we are in the discovery days of the Zero-Point Energy field. You, personally, have as much chance of being successful in harnessing this energy as any large research laboratory with unlimited financial resources. Remember that Flavio Thomas can drive an electrical fan using equipment which costs next to nothing. A cone shape with the same face slope as a pyramid is also an effective shape, and no matter how you position it, it always has a face pointing North--South. May I also remark that it might be worth experimenting with the “pancake” coil (called a bi-filar series-connected coil) patented by Tesla because he found that it was particularly effective in picking up Zero-Point Energy:



Flat 'pancake' coil winding method patented by Nikola Tesla

The Patent of Peter Grandics.

Other people have also investigated pyramid and cone shapes and they confirm that there is indeed a considerable effect from these shapes. Peter Grandics has been awarded US patent 6,974,110 for the collection of electrical energy from a pyramid shape. He tested the system both with an applied high voltage and without any applied voltage, and discovered electrical pick-up in both cases. Here is a digest of part of his patent:

SUMMARY OF THE INVENTION

This invention describes a simple technique to convert the energy of a DC electrostatic field into an alternating current by wrapping a coil around a pyramid. The resulting AC current can be rectified and used for practical purposes. A pyramid-shaped capacitor can also be used in an inverse mode of operation for the generation of propulsive force.

Accordingly, one embodiment of the present invention is a method for converting DC electrostatic energy into usable electrical energy, the method comprising the steps of:

(1) Providing a capacitor of pyramidal shape;

(2) Placing an insulated coil on the surface of the capacitor, the coil having leads;

(3) Attaching a rectifier to the leads of the coil, the rectifier having leads; and

(4) Attaching a capacitor or a battery to the leads of the rectifier so that DC electrostatic energy is converted into usable electrical energy.

DESCRIPTION OF THE PREFERRED EMBODIMENT

This invention describes a novel method is of converting DC electrostatic energy into an AC current which can be rectified and used for practical purposes. The shape of the capacitor and the body of such device is designed to convert the DC electrostatic energy into the AC current for maximum effect.

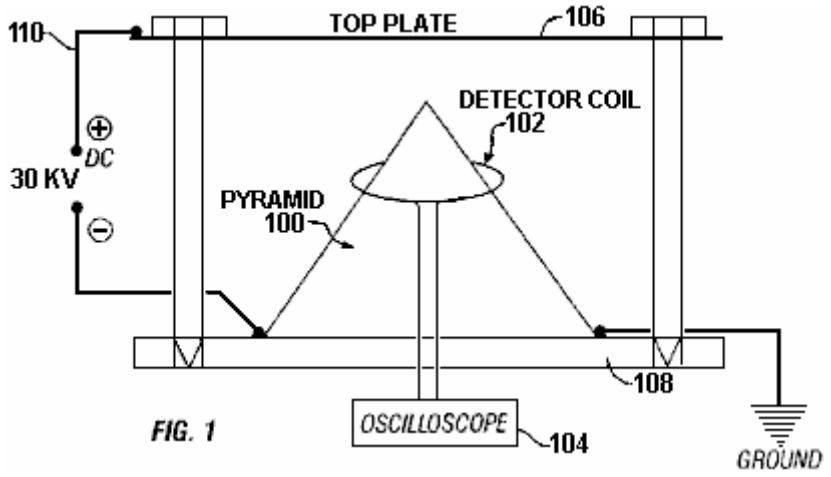


FIG. 1

A pyramidal or conical shape is preferred for one of the capacitor electrodes. In **Fig.1**, a detector coil **102** is provided that which connected to an oscilloscope **104**. The coil surrounds the metallic pyramid **100**. In the experimental set-up shown, the field is established between a top plate **106** and the pyramid **100** by using a ground **108** connected to a source of DC electrical energy **110**. When a high voltage DC field (30 kV) is established on such capacitor, a regularly repeating, clock-like signal is detected in the coil placed on the pyramid's surface (**Fig.2**). This is an unexpected observation as corona discharges are irregular by nature.

The alternating current from the coil can be rectified and used for practical purposes. If a suitable DC electrostatic field could be found in nature, this principle would be useful in tapping the energy of such field. To test for this possibility, I have measured the rectified signal from the coil without an external power source. The rectified coil output was collected in a capacitor and voltage measured at intervals of one hour. The voltage measured is significantly higher if the capacitor electrode is pyramid-shaped as opposed to a box-shaped electrode of the same height and volume. When the pyramid is placed inside a Faraday cage, the signal is excluded (see details in the Example). The data have demonstrated in principle that with this experimental set-up, electrical energy can be extracted from the Earth's electrostatic field. The Earth's surface and the ionosphere substitutes for the two charged electrodes, which exhibit negative and positive polarities, respectively.

EXAMPLE 1

Demonstration of the pyramid generator: For the experiments, I have selected a one-foot base length foam pyramid from a pyramid vendor (The Pyramid Project, Ft. Wayne, Ind.). The outside of the pyramid was covered with aluminium foil. The pyramid was placed on a 2 foot × 2 foot insulating polyethylene platform equipped with an adjustable height 2 foot × 2 foot size aluminium top plate, 1/16" thick. The height of the aluminium plate was adjusted as needed and a gap of 1.25" between the plate and the tip of the pyramid was used in the experiments. In some experiments, an aluminium pyramid was used with a wall thickness of 1/16".

A high voltage (HV) CRT power source producing 30 kV DC was taken from a colour monitor. I have assumed that an actual energy-producing pyramid should be relatively high in order to obtain a large voltage drop from its tip to the ground. Therefore, assuming a height of 100-150 m for a life-size pyramid and a voltage drop of 200-300 V/m near the surface of the Earth, the 30 kV is in the range of the voltage-drop expected for the height of a life-size pyramid.

The positive pole was attached to the top aluminium plate. This simulated the positive charge of the atmosphere. One corner of the pyramid was attached to the negative pole of the high voltage power source, while the opposite corner of the pyramid was grounded. This set-up served as a model for the electrostatic field distribution around a potential life-sized pyramid. As controls, either a 1 foot × 1 foot sheet of aluminium foil or an aluminium foil-covered box, having the main dimensions of the test pyramid (1' × 1' × 7.625"), was used as a negative pole. The detector coils were made by winding 20 turns of 24 gauge enamel-coated magnet wire, approximately 8 cm in diameter. A Tektronix high-frequency oscilloscope, Model no. 2236 was used for signal analysis.

The first sets of experiments were control measurements with a box of the same height and base length as the test pyramid. The detector coil was placed on the top of the box. Measurements were taken with or without the high voltage applied. One corner of the box was attached to the HV power source (negative pole) and the opposite corner to the ground. The same arrangement was used for the flat square (1'×1') foil. The peak-to-peak signal amplitude for the box was 8 mV and the signal frequency was 2 MHz. For the flat foil sheet, the signal amplitude was 12 mV with a frequency of 1.43 MHz. The signal form was of a decaying sine wave.

When high voltage was applied to these shapes, signal amplitude of 14 mV was obtained for the flat sheet and of 16 mV for the box. The signal frequency was 1.54 MHz for the flat sheet and 2 MHz for the box. The waveforms were of decaying sine waves in all these experiments.

When the pyramid was tested without HV, the peak-to-peak signal amplitude was measured at 60 mV with a frequency of 2 MHz. When the high voltage was applied, the signal amplitude increased up to 180-200 mV, while the frequency remained at 2 MHz. The pyramid produced signal intensity significantly higher than the controls. The signal is regularly repeating, clock-like in nature (**Fig.2**). When a metal (aluminium) pyramid of the same size (wall thickness 1/16" inch) was tested in the same high voltage field using the same detection coil, a voltage of 1 to 1.5 V was detected at the frequency of 2 MHz.

To collect energy from the coil, a bridge rectifier (1000 V peak voltage at 6 A) was attached to the leads of the coil. The rectified current was fed into a capacitor (1500 microfarad, 250 V DC max.), and a direct current of 45 V was obtained. This has demonstrated a simple method to convert electrostatic energy into a continuous direct current. An 8-turn coil having an output of 200-300 V AC (peak-to-peak) was also used for energy conversion. The rectified current from the 8-turn coil powered a 0.186 W light bulb (**Fig.3**). Ideally, the bridge rectifier is made of fast-recovery diodes.

Even in the absence of an externally applied voltage, current is always present in the circuit. Charge builds up in the capacitor and 1 V was obtained overnight using the 20-turn coil set-up. Over 48 hours, a voltage of 5 V was measured. Faraday shielding practically prevented the phenomenon.

The preferred shape of the pyramid as a charge collector was again demonstrated in further experiments using the 8-turn coil-bridge rectifier-capacitor (1,500 microfarad) assembly placed on the 1 ft base length aluminium pyramid. A same-volume and height aluminium box was used as control. Charging times of 1-2 hours were used under fair-weather conditions. For the pyramid, 550 mV was measured on the capacitor while on the box 100 mV was obtained. This demonstrated the superiority of the pyramidal shape in capturing atmospheric electrostatic energy. It also demonstrated that we could tap into the electrostatic field of the atmosphere and draw electric energy. For the collection of energy, a battery could substitute for the capacitor.

You should also check out the pyramid aspects of the cutting-edge work of Paulo and Alexandra Correa as detailed in Chapter 11.

The Pyramids of Les Brown.

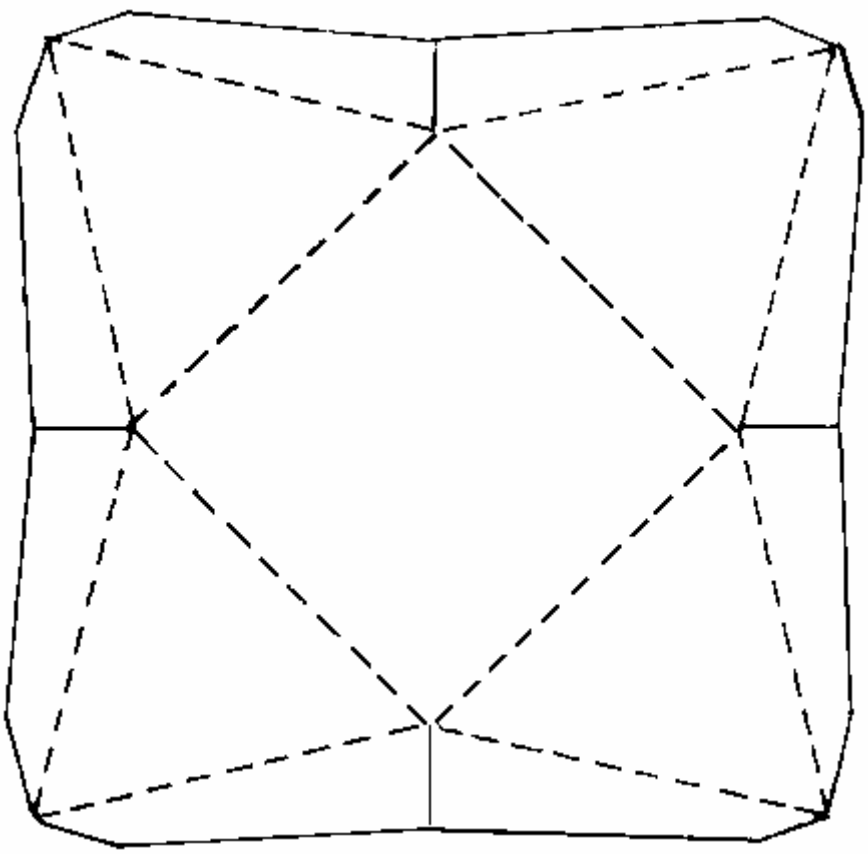
Les Brown experimented extensively with pyramids and related devices. It should be stressed at this point that while various facts have been observed, the action caused by a pyramid is not fully understood at this time and no "laws" have yet been deduced. We have to work here on the basis of "this is what was done, and these are the results". Because of this, the following extract from the work of the late Les Brown is reproduced here and you must decide for yourself if what he says is true and whether or not it might be worth your while trying out some of what he says:

In conducting pyramid experiments you should look daily for signs of change and note them meticulously, and above all be patient. Don't plant a seed one day and expect to have a plant six feet high the next. A plant takes just as long to develop inside a pyramid as it does outside,

but in time you will see the tremendous difference in size. Also, don't keep moving plants around inside your pyramid during an experiment; leave pots stationary so you can see what the results are.

Here, you will see how to make your own pyramids for experimentation; they can be constructed entirely from inexpensive materials equally as well as from costly ones. Cardboard, wire, plywood or anything rigid enough to retain the pyramid shape will suffice. The pyramid does not necessarily have to be solid; in many experiments just the outline shape is sufficient, provided that it is jointed at all corners and at the apex.

An "energiser" array of small pyramids can be made from cardboard and positioned at the corners of a large pyramid to enhance its operation. An array might have 5 rows of four little pyramids or perhaps twelve rows of six pyramids. Individual small pyramids can be made from cardboard and then assembled on a base to form the array. The following template could be used for this, just ensure that the base is square:



A pyramid energizer has multiple uses. The one shown above consists of 72 small cardboard pyramids and a top plate of cardboard covered with foil. By placing the top plate on top of the pyramids and orienting the entire energizer to magnetic north, you can use the energy generated for such purposes as mummification, energizing water, or as a beneficial surface on which to place growing plants for outstanding growth.

Remember that with all types of pyramids, positioning is all-important. One of the sloping sides must at all times be facing towards magnetic North - not a corner but the flat face of the pyramid must face **magnetic** North. Use a compass to determine the direction of magnetic north which will be nearly in line with the Pole star, (but not quite). In this respect, a pyramid is like a radio. For maximum performance you must tune it in properly, pointing it right at the transmitter. Similarly, several pyramids with one face towards magnetic North and stacked on top of each other will produce increased energy, cell activity and growth.

Using pyramids, I sincerely believe that I can grow 36 times more and better plants in a given area than any farmer or market gardener can in the same area, using conventional methods.

All kinds of shapes contain energy, the very nature of the shape itself determines the degree of energy which it contains, that is, the shape determines how receptive it will be to energy flows. By shapes I refer mainly to cubes, spheres, triangles, pyramids and the like. Each shape has potential, but they all have different limits and we should seek the one which offers the most potential. Of all shapes, the pyramid gives us the best performance because it receives the greatest amount of energy. It must be four-sided, of specific measurements and correct angles, and it must have the correct compass orientation. There are dangers in using a pyramid blindly without knowledge of its functions and its great potential. When all four sides are put together you must have a pyramid which is leaning in at 51 degrees, 51 minutes, 14 seconds.

The energy inside the pyramid is said to come down through the peak and continues coming in until it reaches a certain intensity, or the limits of safety, at which point the pyramid releases all the energy and begins collecting it again. The pyramid is said to release 80% of its energy through the peak and the other 20% via the four base corners. Only a pyramid performs this way, and even then the maximum performance is obtained only by a perfect pyramid. So, when you start to build one, aim at perfection. The nearer you can get to a perfect pyramid, the more you will benefit.

You can grow bumper crops with the use of a pyramid. Anything growing beside an iron fence will be bigger and better than one nowhere near metal. The reason is that the iron fence picks up static or magnetic energy and feeds it to the plant. I remember my mother and grandmother placing large nails in the soil of their potted house plants because plants thus treated always grew bigger and better. They had no idea why. In fact, if you asked them, the standard answer was that, as the nail rusted, the plant fed on it. This reasoning is fallible, however, because for one thing, plants can only absorb minerals in liquid form and for another, the rust would kill some plants. Rather, the nails picked up the magnetic energy and boosted the house plants' growth. When a plant receives an extra dose of energy to that already floating free in the atmosphere, that dose acts as a stimulant and causes better growth. What really happens is that the living cells are increased in size, and naturally when each cell is larger, since there are still the same number of cells, the final plant is a lot larger than normal.

If you grow a plant in a pyramid, it absorbs energy at a much higher intensity than that produced by the nails in the plant pots, and so the end result is enormous growth. When this is applied to vegetables and fruits, the plants, as well as their products, are immensely oversized. My own experiments have convinced me that this energy creates a special reaction in living cells of plants, resulting in larger blooms, leaves and fruits on whatever plants are propagated within the pyramid.

The normal life cycle of lettuce, for instance, from seed to maturity, is six to eight weeks. Grown under a pyramid the life cycle is still the same, but the plant is considerably larger. If one allows the vine type of tomato to mature to six or seven trusses under a pyramid while simultaneously allowing an identical plant to do the same outside the pyramid, giving both plants precisely the same feeding and watering, a startling difference in yield occurs. I should mention that if you put your outside plant too near the pyramid, it will reach for, and receive, some of the pyramid's energy, so keep it well away to get a fair test. The outside tomatoes would weigh out at approximately 10 to 14 pounds per plant, whereas the plant grown in the pyramid would produce between 50 and 60 pounds of tomatoes. Not every type of plant grown under a pyramid will produce this increase; this is the average that I have come to expect from tomatoes.

A few more averages I have obtained repeatedly were: lettuce two to three times larger than average; beans 25 inches long by 1.25 inches wide; cabbage - when controls were three pounds each, the pyramid-grown plants were 12 to 13 pounds per head; radishes that normally would be the size of a quarter were four inches in diameter; controlled cucumbers that averaged 14 inches in length and weighed up to one pound normally, were 21 inches long and weighed up to four pounds when grown in the pyramid and the pyramid also warms your whole house and cleanses the air which you breathe.

Energised air in the pyramid also appears to repel small insects; though, there is no need for pesticides to be

used within its glass walls. Pest-free plants grow to maturity inside, with none of the setbacks plants subject to normal attack from pests suffer in the garden outside. This also means that pyramid-grown vegetables need no washing upon harvesting. The mere appearance of such plants is more appetising than that of those grown normally. Greens are more vivid, and many leaves have a sheen which is noticeably absent from plants in kitchen gardens. Artificial fertilisers will never be used in my pyramid. Since many fertilisers apparently are becoming short in supply themselves, the ability to grow plants without their use is a double blessing for all mankind. I will use natural farmyard manures, the best way to regain the succulent taste and nutrition which are missing from plants grown with chemical fertilisers.

An egg broken out of its shell and left within the confines of the pyramid will gradually congeal and become like plastic, as the interior energy works on its cells - harmlessly. The cells do not die nor induce putrefaction. After a period of even weeks or months these congealed eggs can be reconstituted in water to the point where they can be eaten with complete safety, and they taste even more delicious than eggs produced in the usual way.

One peculiar phenomenon which I have observed under my large pyramid is the formation of dew on the plants inside it. This happens early in the morning. During all my years of experience with greenhouses, I never noticed dew forming on any plants in conventional greenhouses. This dew gently dissipates as the sun grows stronger, exactly as it would outside. Also, after a recent thunderstorm, my pyramid cucumbers grew two to two and a half inches in a matter of a few hours.

I built a pyramid purely for research purposes. For large-scale production, one which is much greater in size would be needed. When building the prototype, I encountered and overcame, virtually all the problems one can expect to meet in a construction of this type. Building a pyramid is nothing like building a house, and while a slight difference in measurements can be overcome when building a house, it is not possible to make a mistake in a pyramid and just carry on building. The particular piece containing the error must be pulled out and replaced correctly, as any mistake is transferred all the way around the pyramid.

My test pyramid is 30 feet high along a perpendicular line from ground to peak. The sides from base corner to peak measure 44 feet 4.5 inches, with a baseline of 46 feet 10.5 inches. It contains two additional floors above the ground level, and the sum of the areas of these two floors equals or is greater than that of the ground floor. Thus, the two additional floors virtually double the growing area. My first floor is 12 feet above the ground, and there is a reason for this. I calculated that when the sun was at its highest point the first floor would have to be positioned at 12-feet high so as to allow the sun to shine on to the back north edge of the ground floor. The 12-foot height was perfect, its achievement resulting from a mixture of good judgement and a bit of luck. However, I would not need to build a growing level at this height again because there are as many plants that grow well in shade as there are those that prefer growing in the sun. In future, my floors will all be eight feet apart, and I will put my sun-loving plants in the southern half and my shade-loving ones in the back northern half.

By placing the floors at eight-foot intervals there is much more growing area available than there is in placing the first floor at a 12-foot height. In addition, choosing this lower height allows the upper floors to decrease more slowly in size than they would if placed farther apart. Obviously, as the peak of the pyramid is approached, the floors decrease in area. One of the benefits obtained with these extra floors is that since heat rises, there are higher temperatures on each floor nearer the top of the pyramid. This allows a wide range of crops to be grown.

To illustrate the temperature difference, if the ground floor were 75^{O} F, then the second floor would reach a temperature of 90^{O} F, and the third floor would be around 105^{O} F to 115^{O} F, each floor having higher humidity. This means that anything from cool to tropical crops can be grown within one pyramid. The ground floor is perfect for such crops as radishes, lettuce, carrots, beets, tomatoes, etc., while the second floor is ideal for cucumbers, squash, peppers, and plants which like it hotter and more humid than the first floor. The top floor can be used for lemons, oranges, figs and especially, orchids.

The pyramid draws in its own water on the ground floor; I have never had to water that level being built directly on the ground. It never draws too much or too little water, always just the right amount for growth. Naturally, I have to pump water to the upper floors, but because the first floor provides its own water supply, at least half of my pyramid is watered for nothing automatically. I grow right in the ground the pyramid stands on, but upstairs I have placed wooden planting troughs all around the floors, leaving room to walk, and I grow plants in these. It is a major job getting the soil to the upper floors initially, but it is only a one-time task. The troughs are 16 inches deep and 14 inches wide, and contain a bottom. There is a run all around the

edge of each of the upper floors and another inside, leaving enough room to work, with a gap on each side allowing passage from one run to the other.

Space in the pyramid is used to the utmost. At the perimeter of the low areas, I plant the kind of plants that need little headroom, and then plant bigger crops toward the middle. This is a matter of common sense, but using vine type tomatoes and stringing them up, one can work between the rows better, and if the lower leaves are removed, there is sufficient space to grow lettuce, cabbage or any low-lying crop in between the tomato plants. The trusses may be left on the tomatoes; they will not shade the low-lying plants. To ensure a steady supply of food, it is wise to plant only a few plants of each variety at intervals, which means that in the beginning it will take several weeks to reap a full harvest, but subsequently there will be a continuous yield.

By planting in such a manner, the grower will reap about six full crops each year. This method applies only to an enclosed pyramid, which would also require heating in the winter. The means of heating is up to the individual. Personally, I use a wood stove because I have my own wood supply. However, a wood and oil combination is best because it allows one to be away for a couple of days when necessary. If the wood fire gets low, then the oil burner takes over.

As for growing potentials of the floors of the pyramid, I have found that the second floor is the best place for germinating new seeds. I find that I can get germination in three days on that floor, whereas it usually takes five days otherwise. Thus far I have described the pyramid from the growing aspect only to the point of showing you how and where to grow crops, but let's take a look at a few statistics on production. Your garden, for instance, will only give one crop a year, but the pyramid through the means I suggest will give you six crops per year. Now let's compare two pieces of ground, both the same size, one with a pyramid on it. Say the garden is 50 feet square and the pyramid is 50 feet square. The upper floors inside the pyramid give you approximately 2,500 square feet of growing area upstairs to go along with your ground floor area of 2,500 square feet, and you are getting six crops per floor as opposed to one, or twice times six, the equivalent of 12 times your garden yield. Nor is that all of the advantage. Recall my discussion of the increase in size of the plant cells; this itself gives an average of three times the size of a normal crop. Multiplying the 12 times by the size of the crop, you are 36 times better off than with a single garden.

In addition to food growth, the pyramid also has application in food preservation. I have read statistics stating that 40 percent of all food grown in my home country of Canada is lost to putrefaction, whether at the place of storage, in transport, in wholesale and retail, or finally in the home. Regardless of how this spoilage occurs, this state of affairs can be remedied. The energy of the pyramid which grows plants so amazingly well, can also be used for the purpose of mummification of food, which can be dehydrated and kept in storage for an indefinite period without losing any of its taste or nutritional properties. There are absolutely no ill effects on any food stored in a pyramid. In fact, in many instances it is far better when reconstituted than it was in the first place. It has the water taken out of it; but it also repels bacteria and as a result, nothing will rot in a pyramid. For instance, I cannot make a compost heap inside my pyramid; I have to do it outside; otherwise the ingredients in the compost all remain in good shape and will not break down. For further proof, the grain grown in Manitoba today is a direct descendant of the grain found in the Great Pyramid, grain that had been there for centuries and had kept perfectly.

Earlier I mentioned mummifying eggs. I conducted an experiment in mummifying an egg, using a pyramid energizer instead of a single wire pyramid. The energizer consisted of a small batch of one-inch-high pyramids, 20 in all, positioned in a group. Using such a grid of pyramids provides quicker results than using just one pyramid. I broke an egg into a dish, placed the dish on top of the energizer and observed what happened in the ensuing days. In about six hours I noticed the lower perimeter of the yolk turning a pale yellow, and this continued through each day, the pale colour gradually moving up to the top of the yolk. In the meantime, the white was becoming less fluid, thickening, so to speak. In two weeks the whole egg was just like glass; the yolk was hard and the white was now in crystal form. At no time would flies or any insect approach it, even though it was open to them. Flies were as prevalent as usually, but they would not go near the egg. There was never any smell present from start to finish.

I left the egg in this state for about three months and showed it to many of my visitors, but then it was time to reconstitute it and try it for taste. I added some water (an egg loses about 30 grams of water in such a period) and left it 24 hours. I then boiled some water with the intention of poaching the egg. When the water was ready, I tipped the egg in, and immediately the albumen turned snowy white and the yolk a perfectly natural yellow colour. Had I dropped the egg before reconstituting it, it would have shattered, but now it was simmering away, looking just like a fresh egg. After cooking the egg I put it on a plate, salted and peppered it, then cut the yolk through with my knife, and it flooded across my plate. I admit I was not in too big a hurry

to eat it, but if I was going to prove something, I had to taste it. I smelled the egg, and it was no different from normal, so I ate it. I can honestly say it was one of the nicest eggs I have ever tasted; it seemed to have more flavour than usual.

I do not know of any food that cannot be treated in such a manner. I have tried preserving every food I can think of, and it all keeps indefinitely, with no refrigeration necessary.

Egg processed in a glass ashtray

Since publicity was first given to my large, wooden pyramid some three years ago, I have been inundated with letters, phone calls and personal visits. People from all walks of life and many different places in North America and Europe and indeed a few from further afield - Australia, Africa and the East Indies - all have expressed an interest and have given me kind encouragement. Many of these letters, calls and visits have had as their objective the gathering of practical information on the building of pyramids and the problems likely to be encountered.

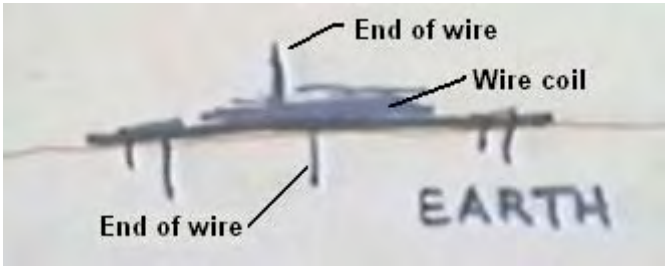
My pyramid is made from rough sawn timber, cut on and near my own property and milled by a neighbour. But it is not necessary for pyramids to be made of wood. They can be made of any rigid material which will support permanent glazing: cardboard, strong wire, sheet steel or metal, angle irons, logs - anything which will not curve and that can be precisely measured and fitted.

Nor do pyramids have to be solid for many uses; open-sided shapes will do, so long as all corners are joined and the angles are correct. My present pyramid is made of timber and covered with heavy-gauge plastic sheet. Future ones will be sheathed in fibreglass, acrylic or glass. They will be closed pyramids solely because I propose to grow food during the depths of Canada's frigid winters.

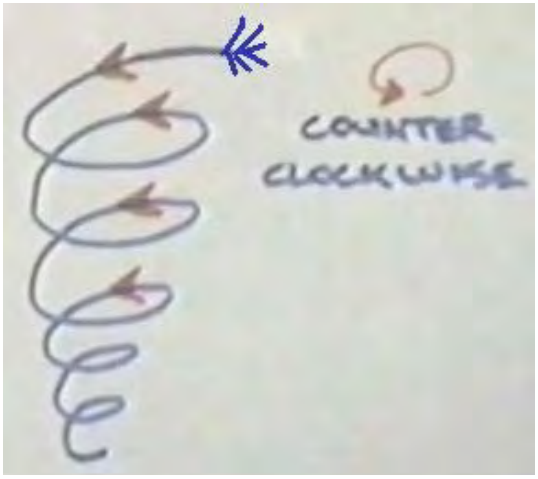
My pyramid frame is built mainly of wood measuring two inches by four inches and two inches by eight inches rough sawn (not planed all over). Pyramids can be built to any scale as long as the proportions are correct.

There is a video of Les Brown on the web at this time. It shows him describing some of his experiences with pyramids. It is at http://www.youtube.com/watch?v=P7VN6B2GjVI and in it, Les describes an experiment which he carried out on his very large agricultural pyramid. First, he constructed a helical coil using co-axial cable which he sketches this way:

The coil was then placed on a 1/4" (6 mm) thick sheet of plywood and the end of the wire in the centre of the coil was inserted through the plywood and the central conductor pushed into the ground. The other end of the wire was bent up to act as an aerial, and the board itself was stapled through with two inch long wire staples driven into the ground and acting as further grounding for the coil, as he sketched here:



Les then climbed up a ladder and fastened another coil under the apex of the pyramid. This coil he wound out of multi-stranded copper wire. The wound it in the shape of an inverted cone with the wire going in a counter-clockwise direction and he spread the top strands out like an aerial array. He drew it this way:



He then decided to connect the two coils together using a strand of natural wool, so he tied it to the bottom of the coil at the apex. He climbed down the ladder and standing on the ground, he took hold of the piece of wool and was thrown several feet away by the energy flowing through the wool. This frightened him so much that he took the pieces apart and never investigated that type of arrangement ever again. He believes that it was because he was grounded when he grasped the wool, that there was such an energy flow through him. If you decide to try this with a smaller pyramid, then I suggest that you connect the wool to the lower coil first, making the earth connection that way, before tying it to the upper coil. However, let me stress again that we just don't know what is happening in and around the pyramid, so any experiments which you choose to carry out are entirely your responsibility and at your own risk. Although a good deal of practical information is given here, it must not be construed as being a recommendation that you make or use any of the devices described in this document as the presentation is for information purposes only.

Les Brown mentions other uses for the devices which he has developed. He says: My wife, who suffered from migraine headaches, had a severe attack for a week; and during that time she took an abundant supply of pills, which proved to be of no benefit. She placed a pyramid on her head in the early hours of the night, and in about 20 minutes had been relieved of all pain. She said nothing of this to me, but two weeks later suffered another headache. Instead of suffering for a lengthy period, she used the pyramid again, and without the use of pills the headache vanished again in about the same length of time. She told me about this second episode. About a week later she suffered an ordinary headache and repeated the performance, with the same result. She now has gone three years with no headaches at all. Around this time, I had been working all day in the hot sun, and the back of my neck became quite painful. I tried placing a small pyramid on the back of my neck while I sat still for a while. It produced the same relieving results for me. I believe it increases blood circulation to the affected part.

After about ten minutes I could feel the pain being drawn out upwards. At the same time my head felt cold on top, not cold to the touch, but simply felt cold to me. I mentioned this to my wife, and she said she had

had the same sensations of coldness but was waiting for me to remark about them first. I am not suggesting that this treatment will do the same for everyone, but merely mentioning that it worked for us.

The elderly mother of a friend of mine, her hands badly crippled with arthritis, was persuaded to place her hand on an energiser. A wire pyramid was put over both hand and energiser for one hour. She continued this treatment for a few days, experiencing a diminution of pain and a loosening of her fingers each time. Before she returned to her home in England she actually sat and played the organ. She played haltingly and rustily, it is true, but she had been unable to play at all for several years prior to using the pyramid. Significantly, her hands returned to their crippled state shortly after she returned to England as her pyramid was stolen at London Airport.

We also find that when we are feeling low and depressed, my wife and I sit for an hour or so, each with a small pyramid on our head, facing North, our depression lifts and we feel quite rejuvenated. My wife and I regularly drink pyramid-treated water. If we are feeling low we have a snifter of pyramid water and in a short time we feel a marked lift. On the other hand, if we find ourselves in what used to be described as a choleric state and sit with the pyramid on our head - but facing south -our irritability rapidly leaves and is replaced by a state of calm.

A prominent acquaintance of mine who does not wish to be named, always drives with a pyramid under his car seat. He claims he feels less "bushed" after a long drive than he did before using a pyramid. Many investigators have found that by putting a pyramid or an energiser over, or under, their beds, they experience better and more refreshing sleep, some of them claiming they need far less sleep than they required before using a pyramid for this purpose.

The real keynote when trying these pyramid experiments is persistence. As with many other experiences, pyramids do not always "work" the first time you apply one to a situation, and a person must "stay with it" and make subtle corrections until the desired results are obtained. Putting energized water into a vase containing tulips caused them to last for about three and a half weeks. Not only did they last this phenomenal length of time, but they grew about nine inches during that period ! I have never before seen flowers grow after having been cut.

Many people claim that insect stings stop itching after the application of energised water; this has been our own experience. On one occasion I had a nasty cut on my hand. Immediately after washing it with pyramid water, the pain stopped and when it was fully healed, I had no scar. The cut also healed much more quickly than is "normal" for me.

Please be aware that the energy channelled by pyramids and the Joe Cell described below is very powerful and more than capable of running a vehicle without the use of any kind of visible fuel. Dealing with this power is not an established science and we are just blundering around the edges of a science which has not been properly investigated, so at this time it is a case of discovering what tests have been carried out and analysing the results to see if there is a solid pattern leading to repeatable tests which can be performed by anyone.

We are at about the same place as the Wright brothers were in 1903 just after their successful powered flight at Kitty Hawk. They took their "Flyer" back to Dayton, Ohio and discovered that it wouldn't fly. The problem was not immediately obvious to them. Being down at sea level and in an uncomfortably high wind, was essential for their device to be able to get off the ground, but understanding why and how to overcome the problem was not immediate as the theory of practical heavier-than-air flight was not established at that time. The same situation applies today to pyramid power.

Joseph Cater's Pyramid Analysis.

There has been a good deal of disagreement between people in different locations, on how effective or otherwise, a pyramid is in day to day use. A very astute scientist by the name of Joseph H. Cater explains that this is due to the pyramid being powered primarily by energy flowing from the Sun, and so, people who live near the Equator will find a pyramid much more effective than people living in high latitudes.

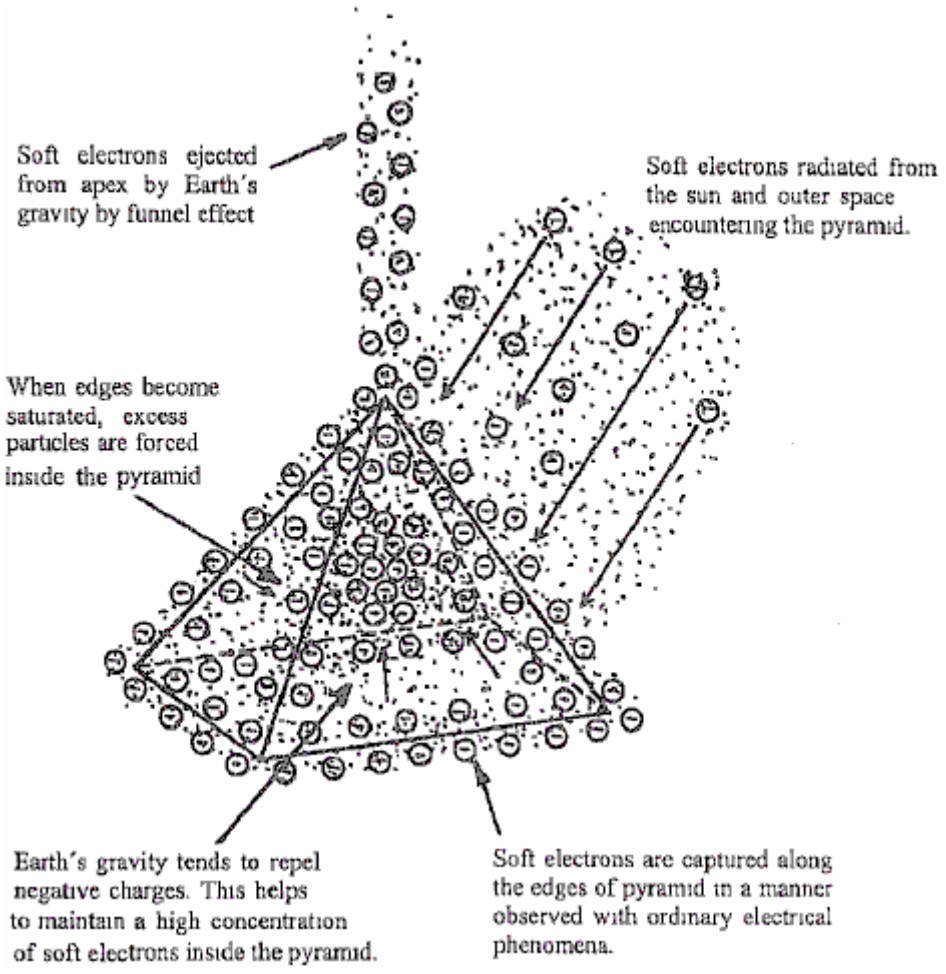
Mr Cater explains exactly how a pyramid operates, but to understand his explanation fully, you will probably need to read the extract from one of his books, which is included in Chapter 15., to discover what "soft particles" are, where they come from and what effect they have on everything around us.

Mr Cater says that the energies concentrated inside a pyramid have been shown to be extremely beneficial to humans. Soft particle bombardments from outer space and especially from the Sun, concentrate inside

the pyramid. Some, passing through the surface of the pyramid are slowed down to such an extent that the Earth's gravitational field, repelling the negative charges, tends to keep them inside until collisions with other particles drives them out.

Most of the particles collected by the pyramid, concentrate along the edges as would be expected, since electricity on any charged body tends to do much the same thing, with concentrations at points and along edges. In fact, pyramid frames have been found to be nearly as effective as the closed pyramid, if, and only if, there is a continuity in the framework and no breaks in any of the joining parts.

The soft electrons collected on a pyramid frame or closed pyramid, soon reach saturation point and continued bombardment causes the excess to drop down inside the pyramid. This, coupled with the gravity-repelling forces, causes a high concentration inside the pyramid. The proportions of the pyramid are apparently a factor in it's performance. If the sides are too steep, many of the soft electrons will move along the edges into the ground outside instead of being forced inside the pyramid. If the sides are not steep enough, not many particles will be collected as they strike the material at nearly a right angle which causes only a small reduction in velocity. If they strike at a sharper angle, there is a greater tendency for them to be retained by the material.



If two side of the base are aligned with magnetic North, it is allegedly more effective. Pyramids can be rendered more potent by lining the interiors of a non-metallic enclosed pyramid, with metal foil such as aluminium or copper. The foil allows a greater quantity of soft electrons to accumulate around the non-metallic outer portion because the soft particles do not pass through the metallic substance as easily, causing a back-up of soft particles. During the process, the foil absorbs large quantities of soft particles before many of them can enter the pyramid. pyramids also radiate soft electrons upwards from the peak.

Many of the soft particles which are stopped briefly on the outside of the pyramid, are repelled upwards by the Earth's gravitational field, and as well, by soft electrons attached to the pyramid. This produces a funnelling effect which ejects soft electrons from the apex of the pyramid. The Earth's gravity accelerates soft particles at a far greater rate than it does ordinary matter as soft particles are associated with ethers

which are much closer to those of the gravity-inducing particles than is the case for ordinary matter. After the pyramid becomes saturated, a greater quantity of soft particles than ever, will concentrate inside. The foil will continue to radiate a high concentration of soft particles during the night when the number of particles bombarding the pyramid is considerably reduced.

It is found that pyramids work better during the summer than at any other time of the year. They are also more effective in the lower latitudes because most of the energy concentrated by the pyramid comes from the Sun. There are conflicting opinions as to the effectiveness of pyramids because of this as there is little understanding of the principles involved. For example, those who experiment with pyramids in Canada may claim that they don't work while those in Southern California will contradict them. A pyramid does not increase the flow of soft particles through the area covered by the pyramid as the same concentration flows outside the area. What a pyramid does, is impede the general flow of soft particles and produce a back-up of particles inside and below the pyramid, and consequently, a higher concentration of soft electrons in these regions. The material used in a pyramid is of great importance. This was demonstrated when a wealthy man in the Midwest built a pyramid-shaped house five stories high, which was then covered with gold-plated iron. The phenomena produced were completely unprecedented. For example, ground water was forced to the surface and flooded the first floor. This was because the soft particle concentration inside and below the pyramid was so great that ground water was impregnated with such an abnormal concentration of negative charges that it was repelled upwards by the Earth's gravity.

Gold atoms have extremely high positive electrostatic field effects, more so than any other atom. This is why gold is the most malleable of all substances. This means that soft electrons will have a greater affinity for gold than for any other metal. As a result, abnormally high concentrations of soft electrons will concentrate around gold. This effect is greatly enhanced when gold is in contact with iron. These dissimilar metals produce an EMF which is turn, causes a flow of electricity or eddy currents resulting in the iron being magnetised. The magnetic field produced, captures additional soft electrons. A higher concentration of soft electrons is created by this combination then could be produced by a similar thickness of gold foil alone. It follows that by far the most effective material that could be used for pyramids is gold-plated sheet iron (galvanised iron should not be used).

With everything else being the same, the greater the size of a pyramid, the better the performance. The reason for this is that the thicker the layer of concentrated soft electrons through which the incoming soft particles must pass, the more they are slowed down when passing. This results in a greater back-up of soft electrons and an increase in the concentration inside the pyramid. Another reason is that a large pyramid has a greater ratio of volume to surface area. Soft electrons are continuously leaking away from the surface of the pyramid, the larger the pyramid, the lower the percentage of soft electrons which is lost. Consequently, very small pyramids are ineffective.

The Devices of Pier Luigi Ighina.

Pier Luigi Ighina was a remarkable Italian scientist who died in 2004 at the age of 95. He was a colleague of Guglielmo Marconi who was well-known for his work in the early days of radio. Pier, in common with many famous scientists, had an enquiring mind, major intelligence and a very high level of patience and persistence. He too, discovered the energy flows which affect life on this planet and he had his theories and descriptions for what he observed. From his very wide range of inventions and fields of interest, just two will be mentioned here in this chapter as they relate directly to the energy flows channelled by pyramids and Joe Cells, and whether we call this energy flow 'orgone', 'soft electrons', 'OD', 'ethers', 'ambient background' or whatever, the effects are the same. Pier describes (very briefly) two important devices which have a direct effect on our health and well-being, and on our local environment generally. The first of these passive devices, he has named "**ERIM**".

He describes this device as a small apparatus which concentrates and develops what he calls "the Magnetic Rhythm Solar Earth Energy, which regenerates cells and returns them to their normal functioning". While I have not seen any specific constructional details for this device we do have a general description and several photographs which show three or four different constructions. On the surface, the device appears to be very simple and straightforward and the photographs give the impression that the construction does not have to be exact in order for it to function correctly. However, it needs to be remembered that these energy flows are directly influenced by the user and just because Pier has had spectacular results with his constructions, some of the performance may well be caused by Pier himself. Having said that, there appears to be very little to be lost by attempting a replication and quite possibly, a major amount to be

gained if the replication is successful. The following four photographs show three different constructions of the device which looks like this:

Or like this:

Photographs 1 and 2 show two different views of one build. The nine spiral coils have seven turns and the narrow end is a flat turn. The gap between the yellow and blue spirals is quite large, appearing to be just under the depth of two turns of the spirals. The base is the same shape and the mounting bolt is not painted, the colour suggesting that it might be a brass bolt. The size of the mounting ring can be estimated from the man in photograph 2 and I would put it at 250 mm (10") as Pier was Italian and so probably used metric units of measurement. I doubt that the dimensions are critical and the spirals appear to be about 100 mm (4") tall with the widest turn being about 45 mm (1.75"). We are not told what metal was used for any of the construction, but as the second of Pier's devices to be described here specifies aluminium, it can't be ruled out here, although the metal does look like painted steel.

Both aluminium and copper are generally considered to be non-magnetic because a magnet does not normally attach itself to either of those metals. But, this is highly misleading as both have a major effect when placed in a magnetic field, and the energy flow being channelled by this device of Pier's is itself affected by magnetic fields, so the use of aluminium is likely to be highly significant. However, that being said, we have no direct information on what is used in this "ERIM" device.

You will notice in photographs 1 and 2, that the lugs on the arms of the bars across the circular mounting are quite short, causing their bolts to be widely spaced, which in turn, causes the outer blue and yellow spirals to be widely separated. Those lugs are longer in the build shown in photograph 3, moving the bolts closer together and bringing the blue and yellow spirals closer together. In photograph 3 the long mounting bolt at the bottom of the ring appears to be thicker and it is painted blue. The spirals have the smallest turn flat across the length of the spiral.

Photograph 4 shows a third build of the device, with an even thicker supporting bolt painted blue, wider spacing between the blue and yellow spirals and this time, the smallest turn of the spiral turned up axially and filed to a point. It appears that the device is assembled and then painted afterwards although the spirals may be dipped beforehand. Photograph 4 shows how the spirals are secured under the head of a bolt which is screwed into a threaded hole in the circular frame. Photograph 3 seems to suggest that constructional accuracy is not critical with the side spirals not aligned correctly (unless the lower spirals are supposed to point slightly inwards which seems unlikely, especially since the top spiral of photograph 3 is not truly vertical and does not align with the other two yellow spirals.

Photograph 4 seems to suggest that the green spirals can be slightly more compact than the others although they have exactly the same number of turns. Every spiral is wound the same way so when viewed from above, the upper spirals move in an counter-clockwise direction as they rise and the lower spirals move in a clockwise direction as they get lower. It might be that the colouring of the components might just be for referencing them, but I do not think that this is so. In the following device, Pier stresses the different colours and considering that different colours radiate different frequencies of light, there may well be significance in the colours. Piers describes the operation in this way:

The properties of the device are determined by it's shape. The three yellow spirals pointing upwards concentrate Solar Energy while the three blue spirals pointing downwards, concentrate Earth's bio energy which is the negative reflection from the Earth of the Solar Energy. When these two meet, they produce a wave on the green spirals called the Sun-Earth Magnetic Rhythm (which is the rhythm of everything which exists on Earth - both animate and inanimate). This wave enhances that essential basic rhythm for everything within it's operating radius, normalising and balancing cells and cleansing the energy of a room.

Ideally, the device should be placed in the bedroom so that it can act during the night, but it can also be effective almost anywhere. It should be positioned about 10-20 cm away from walls. At the beginning, to accelerate the process of cellular balancing, you can place your hands about 10-15 cm from each side of the device, with your palms turned towards the green spirals and at the same level as they are. The time required will depend upon individual sensitivity

If you have hot water central heating, you can place the device about 10 cm from the boiler with the green spirals parallel to it; the circulating water will carry information throughout the system and cleanse rooms within a few days. To energise water and return it to it's original, natural state, place a bottle of water at each side of the device, about 10-15 cm from the green spirals. It can also energise foodstuffs such as fruits, vegetables and meats by placing them level with, and about 10-15 cm in front of the green spirals, using a

non-metallic support, if necessary. Flowers placed close by, will keep fresher longer, or will blossom sooner than normal.

The second passive device designed by Pier is what he calls his "**Elios**" device. It is constructed from a long small-diameter coil of aluminium wire which is snaked backwards and forwards through a circular wooden frame supported on three short wooden legs. The wooden frame is painted as shown here:

The aluminium wire coil is fed through the frame using plastic tubular beads. It is not clear if the beads pass the whole way through the frame with the upper bead meeting the lower bead in the centre of the frame, or if the beads are completely external to the wood, or perhaps partially embedded in the wood.

Pier says that the Elios device concentrates the maximum energy output in the cavity. The device is able to destroy every type of toxic pollution, including nuclear radiation. Food exposed to it's action changes completely because every chemical added to the food gets dissolved by balancing it's own Rhythm (vibration) by a multi-wave signal comprising a full octave of colours. Every type of matter receives a certain type of enhancement from it. For example, our body and all types of biological entities are healed, and growth is boosted by it.

This signal is a stationary type phenomenon of rotating particles which come from the Sun and other stars. These particles strike the Earth, store energy, and return to their sources in a reflected state which is the opposite phenomenon of discharge. Everything which was born or which is growing, changes when inside this column of light-like particles. This phenomenon is called "Solar-Terrestrial Rhythm" and it is the power supply of the Planet, continuously causing balancing and growth. Everything on the planet is continuously bathed in this stream of rotating particles.

Every type of matter has its own basic Rhythm plus the Solar-Terrestrial Rhythm. This rhythm is unique and it represents the state of tension of the specific matter. The matter grows and evolves because of this incoming energy. If you modify the energy with various types of pollution, radiation, or other types of

destructive phenomena, you destroy the normal work of Nature. This device concentrates this phenomenon in it's immediate local area and as a result, it improves the Health and Life Energy of that local environment.

The Elios device is only a passive resonator, tuned by it's shape to resonate with a harmonic of the cosmic event signal. It consists of a ring of painted wood with seven holes positioned in a heptagon shape in it's perimeter. This ring is supported on three legs evenly spaced around the circle. On every hole there is a piece of a coloured plastic tube, each one having a different colour and positioned in a particular order.

The sinusoidal conductor is made of aluminium wire, wound clockwise, and inserted through every hole in a serpentine movement, leaving one end facing upwards and the other end facing downwards, parallel to each other. The signal coming from the sky, moves in a clockwise direction. The two ends of the aluminium wire coil finish in adjacent holes. One end points towards the sky and the other end points towards the ground.

The device should be located on open ground and not on concrete which generates an out-of-phase component of the earth signal which destroys the good rhythm. The phenomenon occurs in the cavity especially, and around the perimeter of the ring. To increase the amplitude of the signal you can put **glass** bottles filled with water, one inside the cavity and six around the outside in a hexagon shape. Wait some days and observe the environment. Everything changes! There is an explosion of life! Everything which I have described happens, depending only on the length of elapsed time and the amplitude of the signal. The device needs sunlight for it's best operation. Take it away from every artificial electrical and magnetic field because this is only a passive resonator and a strong EMF can distort its' signal.

Do not try nuclear cleansing with Elios until it is highly charged. The device is a passive resonator and it needs to build up a charge before bringing any isotopes near it. When dealing with isotopes, always use water containers but do **not** contact the substance with the apparatus as you can destroy the ambient rhythm by doing that, and cause very big trouble in the area. This device has a powerful resonant effect on the surrounding land, influencing the environment up to a kilometre in every direction.

The Joe Cell.

The device called the "Joe Cell" used to be one of the most difficult devices for any experimenter to get operating properly but new design data has changed all that. It is a passive device for concentrating energy

drawn from the local environment and it takes great perseverance and patience to use one to power a vehicle. Here is some practical information on the Joe Cell.

In 1992 in Australia, Graham Coe, Peter Stevens and Joe Nobel developed previously patented units which are now known by the generic name of the "Joe Cell". Peter introduced Joe to Graham and they rehashed the patented cells which Graham knew about, using materials from the Local Dairy Production Facility NORCO. A two hour long video showing the Joe Cell was produced by Peter and Joe and the unit shown operating in the video was attached to Peter's Mitsubishi Van. Joe had his equipment stolen and his dog killed, so he decided to keep a low profile, moving out into the wilds and not generating much publicity, in spite of fronting the two hour video recording. A search on the Joe Cell will locate many videos on the subject. This document is an attempt to provide detailed information on a recent Cell built by Bill Williams in the USA and the subsequent constructional advice which has arisen from his experiences.

First, you need to understand that, at this point in time, building and using a Joe Cell of any variety, is as much an art as a science. It might best be explained by saying that creating building plans for it is rather like producing plans for painting a copy of the famous Mona Lisa painting. The instructions for the painting might be:

1. Buy a canvas, if one is not available, then here is how to make one.
2. Buy some oil-based paints, if none are available, then here is how you make them
3. Buy an artists brush, palette and charcoal, if none are available then this is how you make them.
4. Here is how you paint the picture.

Even given the most complete and detailed instructions, many people, including myself, are unlikely to produce a top-quality copy of the Mona Lisa. It is not that the instructions are lacking in any way, it is the skill and ability of the person attempting the task which are not up to the job. It used to be that not everybody who built a Joe Cell had instant success. However, recent advances have changed all that, but let us run through the earlier constructional information as a lead-in to the current technology.

A Joe Cell is capable of powering a vehicle engine without needing to use conventional fossil fuel. So, what does the engine run on? I suggest that it runs on a newly discovered energy field not yet understood by mainstream science. In another couple of hundred years time, it will be a routine subject which every child in school will be expected to understand, but today it looks like the 'witchcraft' of the magnifying glass starting a fire.

It is not unusual for newcomers to the subject to get confused by the Cell itself. The Cell consists of a metal container with tubes inside it. The container has what looks like ordinary water in it and it sometimes has a DC voltage applied across it. This causes many people to immediately jump to the false conclusion that it is an electrolyser. It isn't. The Joe Cell does **not** convert water to hydrogen and oxygen gasses to be burnt in the engine. The water in a Joe Cell does not get used up no matter how far the vehicle travels. It **is** possible to run a car on the gasses produced by electrolysis of water, but the Joe Cell has absolutely nothing whatsoever to do with electrolysis. The Joe Cell acts as a concentrator for a new energy field, in the same way that a magnifying glass acts as a concentrator for sunlight, and both have to be done just right for them to work.

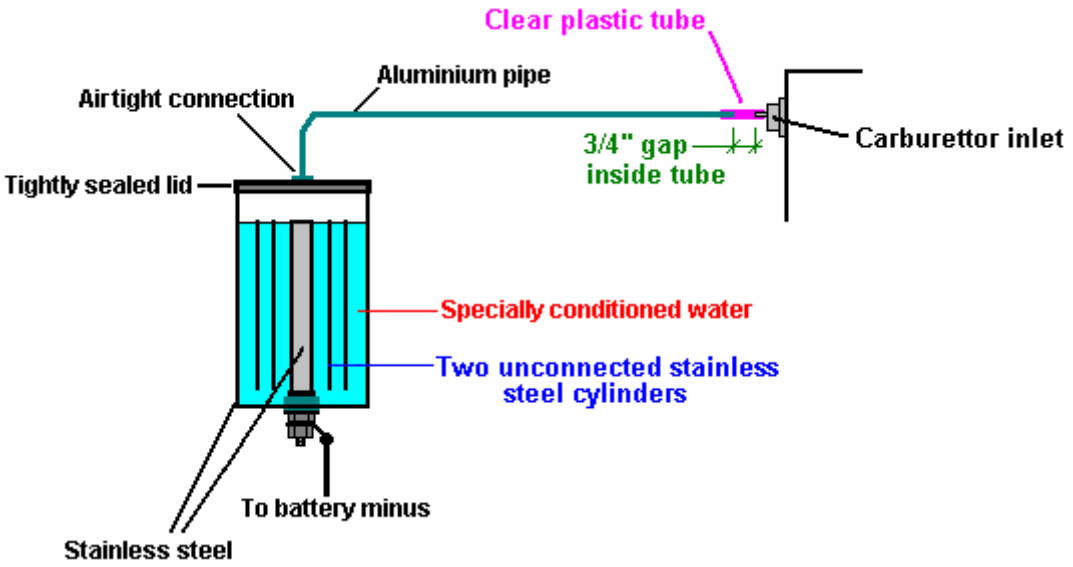
At the present time, there are at least fifteen people who have built Joe Cells and managed to power vehicles using them. Several of these people use their Joe Cell-powered vehicles on a daily basis. Most of these are in Australia. The first Cell-powered vehicle was driven some 2,000 kilometres across Australia.

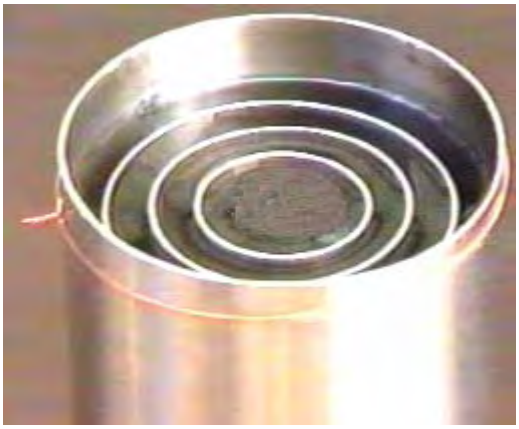
Disclaimer: The remainder of this document contains considerable specific detail on the design and construction of a Joe Cell. This presentation is for information purposes only and must not be construed as a recommendation that you actual physically construct a device of this nature. The author stresses that he is in no way liable for any damage, loss or injury caused by your future actions. It should also be borne in mind that any alteration to an automotive vehicle, such as changing the fuel on which it runs to hydroxy gas, natural gas, Joe Cell energy, or anything else, might void the vehicle insurance unless the insurer is informed beforehand and agrees to continue insurance cover on the modified vehicle.

In broad outline, a Joe Cell is a 316L-grade stainless steel container, with a central cylindrical electrode, surrounded by a series of progressively larger stainless steel cylinders, and filled with specially treated water. This arrangement of steel shells and treated water acts as a focusing mechanism for the energy field used to power the vehicle.

The Cell itself is made up with the battery negative taken to the central electrode. The connection to this stainless steel electrode is made at the bottom with the electrical connection passing through the base of the cell container. This obviously needs careful construction to prevent any leakage of the conditioned water or the energy focused by the Cell.

Surrounding the central electrode are two or three cylinders made of either solid or mesh stainless steel. These cylinders are not connected electrically and are held in position by insulating material which needs to be selected carefully as the insulation is not just electrical insulation but is also energy-field insulation. The outside stainless steel cylinder forms the container for the cell:





The picture above shows the general construction of a cell of this type although, unlike the description below, this one does not have the lip which is used for attaching the lid. It is included here just as a general illustration of how the cylinders are positioned relative to each other.

The following information on constructing a Joe Cell, is broken down into the following sections:

1. The Materials needed for construction.
2. Constructing the Cell
3. Getting the Cell working
4. Installing the Cell in the vehicle
5. Getting the vehicle running
6. Suppliers
7. Workarounds

The Materials needed for Construction.

Various vehicles can be powered by a Joe Cell. If you have not built and used a Joe Cell before, then it is worth using the easiest type to convert. The most suitable is an older type vehicle with no computer control of the combustion, a carburettor and a water-cooled engine. If the engine block is aluminium rather than steel then that is also a slight additional advantage.

The Cell is built from stainless steel pipes. The lower the magnetism of the finished unit the better, so 316L grade stainless steel is preferred. However, there is no need to become obsessed with this as most varieties of stainless steel can be persuaded to operate. The length of the tubing is not critical, but about 8 inches (200 mm) is a reasonable choice for the overall length of the inner tubes. The outer pipe which forms the casing, needs to be about 10 inches in length so that there is clearance above and below the inner pipes.

The innermost pipe diameter is 2 inches (50 mm) and the others can be 3 inch, 4 inch, and 5 inches in diameter as that creates a gap of just under half an inch between the pipes, which is a suitable spacing. The wall thickness of the pipes is not critical but it needs to be a practical size with 1 mm being the minimum thickness with the most common thickness being 1/16 inch (1.6 mm or 0.0625 inch). It is important that the walls of the outermost cylinder are completely rigid, so using a greater thickness for that cylinder is an advantage.

Some stainless steel plate is needed for the ends of the outer cylinder. Ideally, the top and base should not overhang the sides but that is difficult to achieve if the cell is to be airtight, so the end pieces will need to be slightly larger than the outside tube and 1/8 inch (3 mm) thick sheet is suggested. The base size is 5 inch square, or possibly slightly larger to facilitate cutting a circular shape out of it. The lid and lip blanks will need to be 6 inch squares, or again, slightly larger to facilitate cutting circles out of them.

The plinth component at the base of the 2-inch inside tube needs to be cut from a piece of stainless steel. If the option of machining the whole plinth as a single piece is chosen, then the piece of 316L stainless steel needed to do this will be substantial, perhaps a section of solid bar 2.25 inches (57 mm) in diameter and some 3 inches (75 mm) long. If the easier and cheaper option of using a standard half-inch (12 mm) 316L stainless steel bolt (if one is available) is selected, then a piece of 316L stainless steel some 2.25 inches (57 mm), or slightly larger, 2 inch (50 mm) thick will be needed. The exact details of this will need to be discussed with the person who will undertake the machining as practical issues come into play, and the optimum size will depend to a certain extent on the lathe being used. If a screw thread is being machined on the spigot of the plinth, then the thread should match the locally available nuts, unless nuts are also being made up.

Some additional steel will be needed for constructing a mounting bracket inside the engine compartment, also, some double-laminated hessian sacking (“burlap”) and about 36 inches (1 m) of half-inch (12 mm) wooden dowel to use in the mounting bracket.

Some Ultra-High Molecular Weight Polyethylene material as found in kitchen chopping boards will be needed to insulate between the engine mounting and the cell and between the inside tube’s plinth and the base plate.

A length of aluminium tubing typically three quarters of an inch (20 mm) in diameter will be needed for connecting the Cell to the engine, and a short length of strong, clear plastic pipe for the actual final connection to the engine, needed to prevent an electrical short-circuit between the Cell and the engine. This plastic pipe needs to be a tight push-fit as clamping clips are not used. A stainless steel compression fitting to fit the pipe is needed to make the seal between it and the lid of the Cell. It is very important that this fitting is stainless steel as other materials such as brass will prevent the cell from operating. The wrong material for this fitting has been the reason for many Cells not operating. Neither brass nor any other material (other than stainless steel) should not be used anywhere in the construction, whether it be for nuts, bolts, fittings, metal connections, or anything else.

Ideally, natural rubber with no additives or colouring, failing that “Buna-n” (nitrile rubber) o-ring, or teflon, is needed for inter-cylinder bracing and some sheet to make the circular lid gasket. Also some white marine-grade Sikaflex 291 bedding compound. Natural rubber with no colouring or additives is the best insulator and should be used if at all possible. After extended use, Bill has found that teflon spacers work better than the rubber and so has switched to teflon.

Seven or eight stainless steel cones will be needed for the water-conditioning process. These are usually manufactured for machines which separate cream from milk and it is possible to buy them via eBay from time to time. If none are available, then it is perfectly possible to construct them yourself.

There will also be minor items like a few bolts, lengths of electrical wire and the like. To summarise this then:

Stainless steel pipes in 316L grade steel:

5-inch (125 mm) diameter 10 inches (250 mm) long, one off
4-inch (100 mm) diameter 8 inches (200 mm) long, one off
3-inch (75 mm) diameter 8 inches (200 mm) long, one off
2-inch (50 mm) diameter 8 inches (200 mm) long, one off

Stainless steel plate in 316L grade steel:

5.25 inch (133 mm) square 1/8 inch (3 mm) thick, one off
6.25 inch (157 mm) square 1/8 inch (3 mm) thick, two off
3 inch (75 mm) strip, 16 gauge thick, two feet (600 mm) long
One plinth blank as described above, size depending on the lathe and style of construction.

Stainless steel bolts:

1/4 inch (6 mm) diameter, 3/4 inch (18 mm) long, twelve off with matching nuts
One 1/2 inch (12 mm) diameter, 2.25 inch (57 mm) long with two nuts and three washers

Aluminium tubing 3/4 inch (20 mm) in diameter, 3 feet (1 m) long
Plastic tubing to form a tight fit on the aluminium tubing and some 4 inches (100 mm) long
One stainless steel compression fitting to seal the pipe-to-lid connection

Natural rubber with no additives, (or “Buna-n” insulation if natural rubber just cannot be got):

O-ring tubing, 3 feet (1 m) long
Sheet, 6 inch (150 mm) square, one off

Miscellaneous:

White Sikaflex 291 bedding compound (available from ships chandlers), one off
Double-laminated hessian sacking (“burlap”) 1 foot (300 mm) wide, 6 feet (2 m) long
Wood (ramin) dowel three quarter inch (18 mm) diameter, 36 inches (1 m) long
UHMWP plastic food-chopping board, one off
Sundry connecting wire and ordinary engine compartment mounting bolts, and the like
Stainless steel cones and canister as discussed below

Don’t polish the tubes and never, ever use sandpaper or wet-and-dry paper on any of these components as the result is scored surfaces and each score reduces the effectiveness of the Cell.

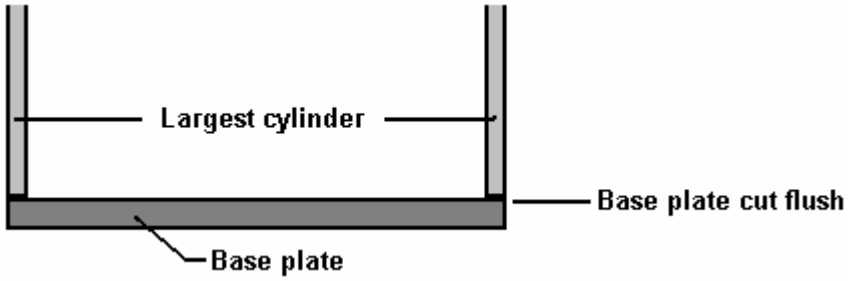
Constructing the Cell

The Joe Cell looks like a very simple steel construction which could easily be made by any amateur. While it can be constructed by an amateur, it is not a simple construction as it is important to keep any acquired magnetic properties to a minimum. Consequently, it is suggested that an angle grinder is not used for any of the metalwork, and hand tools used for cutting and shaping. Also, if the cutting tool has previously been used to cut anything other than stainless steel it should not be used, or at the very least, thoroughly cleaned before use as contamination of your Cell components through particles of another material is critical and can prevent the Cell from working. It should be stressed again that the materials used in the construction of a Cell are absolutely critical if success is to be assured. If you have an experienced friend who has made many Cells work, then you can experiment with different materials, but if this is your first Cell and you are working on your own, then use the exact materials shown here and don’t end up with a Cell which doesn’t work.

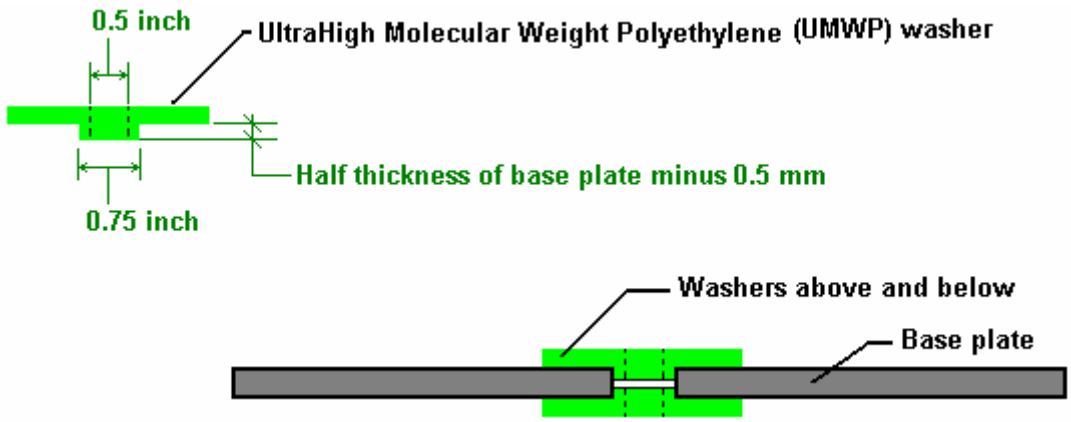
Bill Williams started building a 5 cylinder cell comprising 1", 2", 3", 4" and outer tube 5" but Peter Stevens later advised him to remove the 1" centre tube and go with only two neutrals being the 3” and 4" tubes as the 1-inch diameter is too small for optimum energy pick-up.

Please accept my apologies if the following suggestions for construction seem too basic and simple. The reason for this is that this document will be read by people whose first language is not English and who will find it much easier if plenty of detail is provided.

The first step is to construct the base plate, used to form the bottom of the container. Cut the largest diameter pipe to a 10-inch (250 mm) length. (If you have difficulty in marking the cutting line, try wrapping a piece of paper around it, keeping the paper flat against the tube and making sure that the straight edge of the paper aligns exactly along the overlap, then mark along the edge of the paper). Place the pipe on one of the end blanks and mark the blank around the bottom of the pipe. Cut the blank to form a circular plate which sits flush with the bottom of the tube:



The next step is to mount the innermost 2-inch (50 mm) diameter pipe rigidly to the base plate. Cut the pipe to an 8-inch (200 mm) length. The pipe mounting needs to be exactly in the centre of the plate and exactly at right angles to it. This is probably where the most accurate work needs to be done. To complicate matters, the mounting needs to be connected electrically outside the base, be fully insulated from the base plate, and make a completely watertight fit with the base plate. For that reason, the arrangement looks a little complicated. Start by drilling a three quarter inch (18 mm) hole in the centre of the base plate. Construct and fit two insulating washers so that a half-inch stainless steel bolt will fit through the base plate while being securely insulated from it. The washers are made from Ultra-High Molecular Weight Polyethylene (plastic food-chopping boards are usually made from this material):

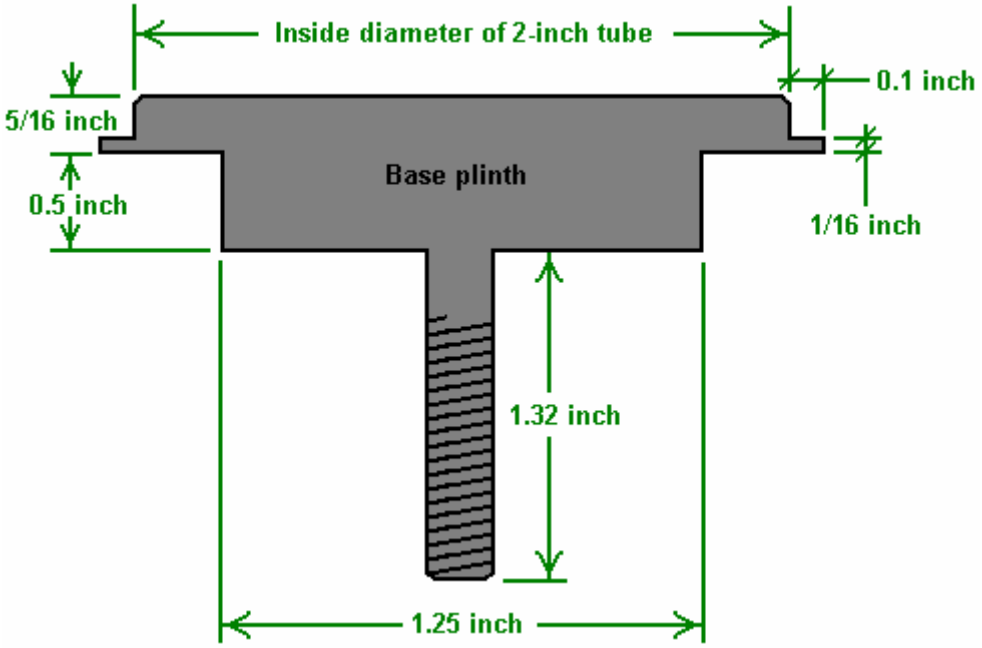


The washers which fit into the hole in the base plate need to be slightly less than half the thickness of the plate so that they do not actually touch when clamped tightly against the base plate, as shown in the lower

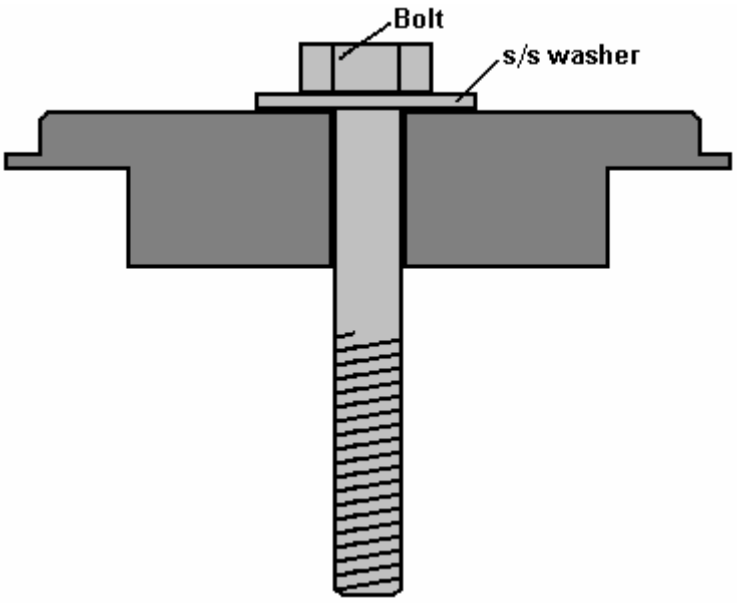
part of the diagram. Cut another washer, using the full thickness of the plastic sheet. This will act as a spacer.

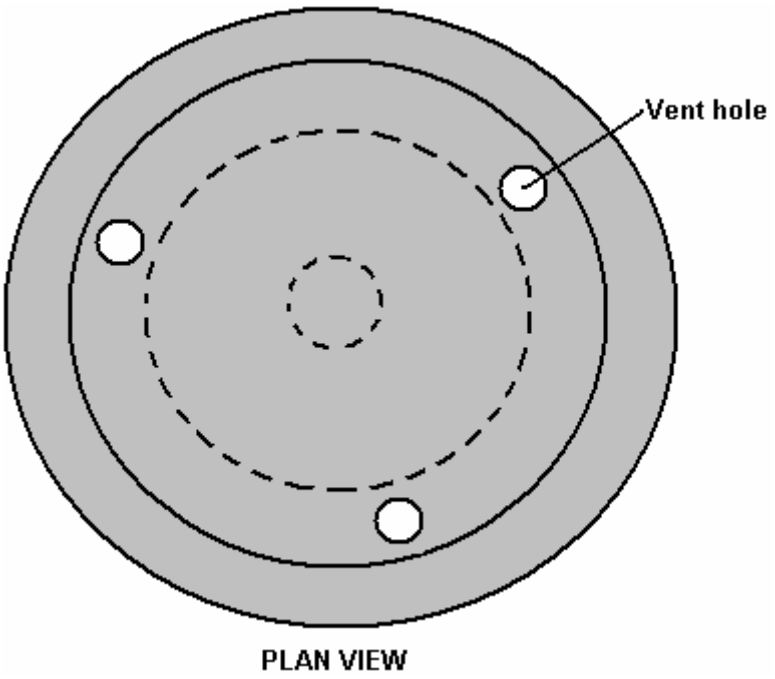
Next, the plinth for the central 2-inch diameter cylinder needs to be made. This is the only complicated component in the construction. It is possible to make this component yourself. The local university or technical college will often be willing to allow you to use their lathe and their staff will usually do the job for you or help you to do it yourself. Failing that, your local metal fabrication shop will certainly be able to do it for you. If all else fails and this equipment is just not available, then the 'workarounds' section below shows how to fabricate an alternative version which does not need a lathe.

A large piece of 316L stainless steel needs to be machined to produce the plinth shown below. The actual 2-inch diameter central cylinder needs to be a tight push-fit on the top of this component. To facilitate assembly, the central boss is given a slight chamfer which helps alignment when the tube is forced down on top of it. Peter Stevens recommends that tack welds (in stainless steel using a TIG welder) are used to connect the plinth to the outside of the cylinder. Three evenly-spaced vent holes are drilled in the plinth to allow the liquid inside the Cell circulate freely inside the central cylinder.

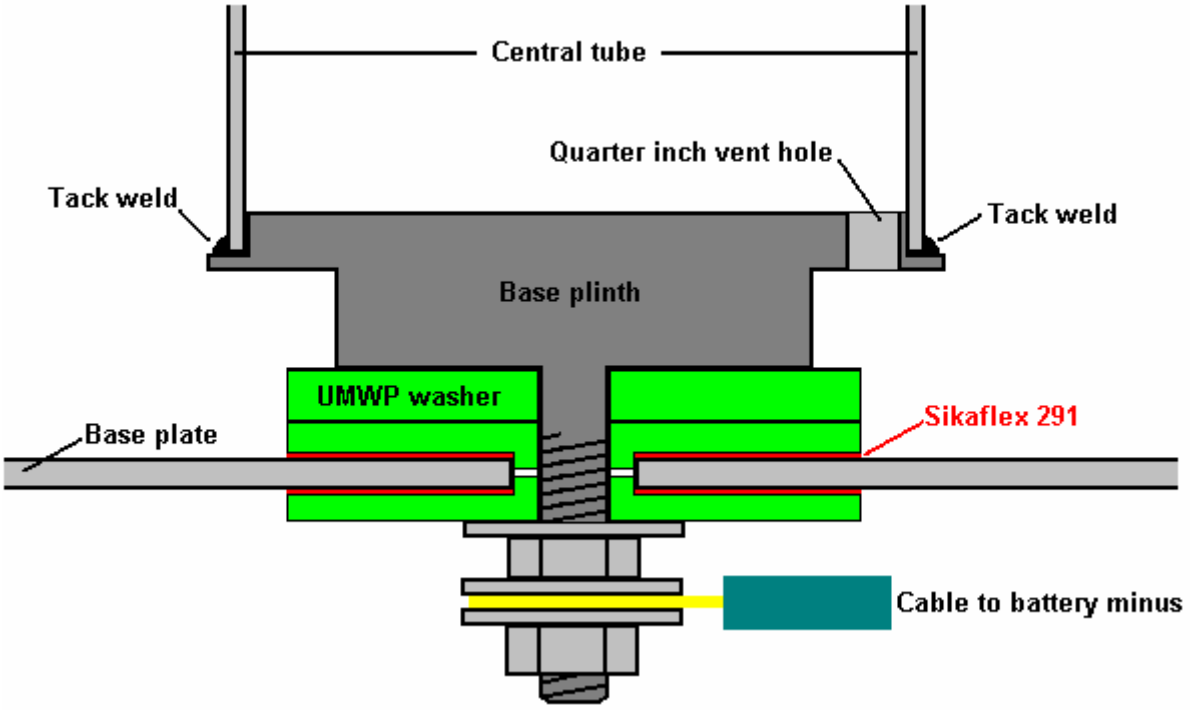


An alternative method of construction which does not call for such a large amount of machining is to machine the plinth to take a standard stainless steel bolt as shown here:



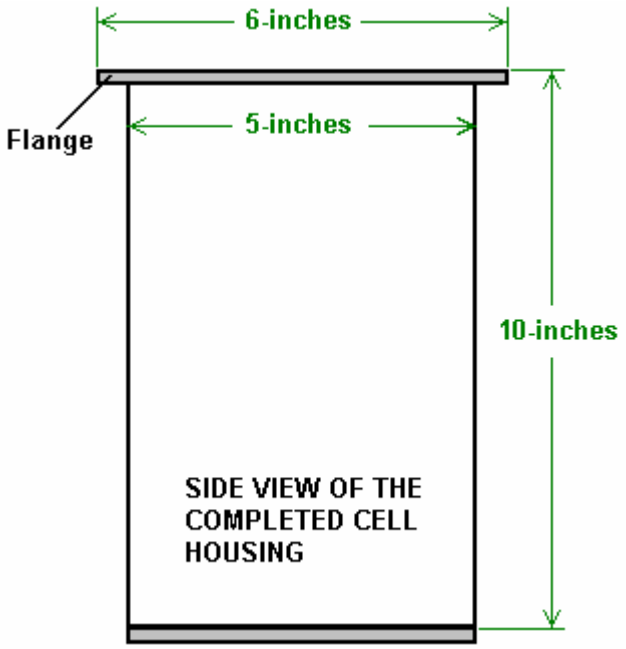


When assembled, the arrangement should look like this:



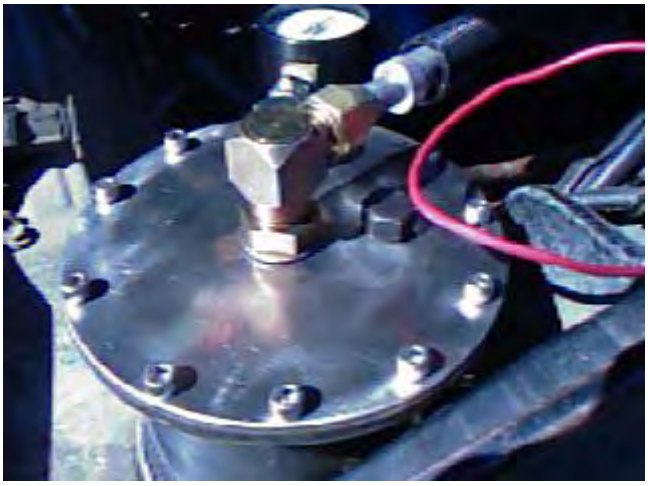
This arrangement looks more complicated than it really is. It is necessary to have a construction like this as we want to mount the innermost tube securely in a central vertical position, with the battery negative connected to the cylinder, by a connection which is fully insulated from the base plate and which forms a fully watertight seal with the base plate, and to raise the central cylinder about one inch (25 mm) above the base plate.

However, as the plastic washers would be affected by the heat when the base plate is joined to the outermost pipe, when all of the components shown have been prepared, they are taken apart so that the base plate can be fuse-welded to the outside tube. Unless you have the equipment for this, get your local steel fabrication workshop to do it for you. Be sure that you explain that it is not to be TIG welded, but fuse-welded and that the joint has to be fully watertight. At the same time, get them to fuse-weld a half-inch wide lip flush with the top edge of the tube. You cut this piece as a 6-inch (150 mm) circle with a 5-inch (125 mm) circular cut-out in the centre of it. When it is welded, it should look like this:



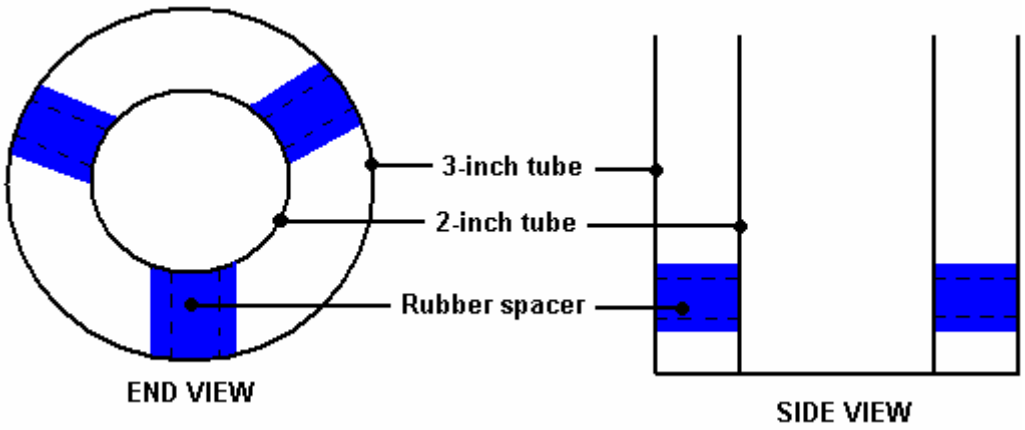
Cut a six-inch (150 mm) diameter lid out of 1/8 inch (3 mm) stainless steel. Cut a matching ring gasket of natural rubber (Buna-n material if natural rubber can't be obtained), place it on top of the flange with the lid on top of it and clamp the lid firmly down on the flange. Drill a hole to take a 1/4 inch (6 mm) stainless steel bolt, through the lid and the middle of the flange. Insert a bolt and tighten its nut to further clamp the lid in place. An alternative to this for the more experienced metalworker, is to drill a hole slightly smaller than the bolt, and when all holes have been drilled, remove the lid, enlarge the lid holes to allow free passage of the bolts, and cut a thread inside the flange holes which matches the thread on the bolts to be used. This gives a very neat, nut-free result, but it calls for a greater skill level and more tools.

If using nuts and bolts, drill a similar hole 180 degrees away and fasten a bolt through it. Repeat the process for the 90 degree and 270 degree points. This gives a lid which is held in place at its quarter points. You can now complete the job with either four more evenly-spaced bolts or eight more evenly-spaced bolts. The complete bolting for the twelve-bolt choice will look something like this when the cell is installed:



The lid can be finished off by drilling its centre to take the fitting for the aluminium pipe which will feed the output from the cell to the engine. This fitting, in common with every other fitting **must** be made of stainless steel. Video at http://youtu.be/-7075bVmDQo.

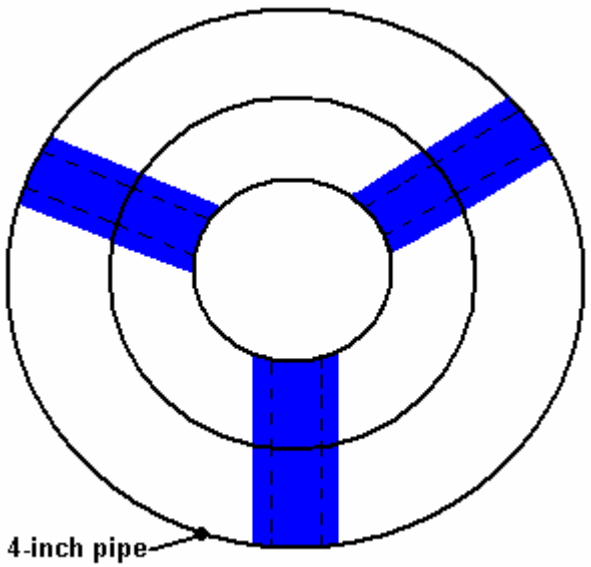
The next step is to assemble the neutral pipes. Cut them to 8-inch (200 mm) lengths. These pipes are held in place by the natural rubber insulators. This material comes in an o-ring strip which is like a hosepipe with a large wall-thickness. The gap between the pipes will be approximately half an inch (12 mm), so cut each piece of pipe to a length which makes it a very tight fit in that gap. Cut six spacers, locate the 3-inch diameter pipe exactly over the inner pipe and push three of them between the pipes, about a quarter of an inch from each end and evenly spaced 120 degrees apart around the circumference of the pipes. The hole through the centre of the insulating strip points towards the centre of the cell and the ends of the insulator pieces press against the cylinder walls. These pieces are **not** placed lengthwise:



Place similar insulators at the other end of the two-inch pipe, directly above the ones already in place. If you look down the length of the tubes, then only three of the six insulators should be seen if they are correctly

aligned. The spacers will be more effective if the ends are given a thin layer of the Sikaflex 291 bedding compound before the ends get compressed against the cylinder walls.

Do the same for the four-inch pipe, pushing tightly squeezed natural rubber insulators strips between the three-inch and four-inch pipes. Place them directly outside the insulators between the two-inch and three-inch pipes so that when viewed from the end, it looks as if the rubber forms a single strip running through the middle pipe:



END VIEW

Spark off each of the cylinders in the inner assembly. This is done by connecting a 12V battery negative to the inside surface (only) at the bottom of the tube and with a wire from the battery positive, sparking the outside surface of the cylinder at the top of the tube. Give each four sparks in rapid succession.

If you are using a bolt rather than a machined spigot, insert the stainless steel bolt and washer through the bottom of the base to the central pipe. Wedge the bolt in place by inserting a piece of the dowel, or some similar material into the centre of the 2-inch pipe and tape it temporarily in place. Alternatively, force the innermost cylinder tightly over the machined plinth. Turn the inner pipe assembly upside down and place the full-depth UMWP plastic washer on the threaded shaft. Apply a thin layer of white Sikaflex 291 bonding compound to the face of one of the shaped UMWP washers and place it on the threaded shaft with the bonding compound facing upwards.

Carefully clean the surface of the base plate of the outer casing around the central hole, both inside and outside. Under no circumstances use sandpaper or wet-and-dry paper, here or anywhere else, as these abrade and score the surface of the steel and have a major negative effect on the operation of the Cell. Carefully lower the 5-inch outer casing on to the assembly so that the threaded shaft goes through the central hole and the shaped washer fits tightly into the hole in the base of the outer housing. Apply a thin layer of the bonding compound to the face of the second shaped washer, place it over the shaft of the bolt and press it firmly into place to completely seal the hole in the base plate. Add a stainless steel washer and bolt and tighten the bolt to lock the assembly together. If using a bolt, a long-reach box spanner may be needed inside the central pipe for tightening the locking bolt. If one is not available, use a longer bolt through the washers, screw a second nut up on to the shank of the bolt, file two flats on the end of the bolt, clamp them in a vice to hold the bolt securely and tighten the locking nut. When the spare nut is unscrewed, it pushes any damaged fragments of the bolt thread back into place.

Finish the assembly by adding three further rubber insulators between the top of the 4-inch tube and the outer 5-inch casing. Use a thin layer of Sikaflex 291 bonding compound on the cut faces of the insulators as this improves the insulation. Line the new insulators up with the insulators already in place and make them a tight fit. These extra insulators support the end of the tube assembly and reduce the stress on the plinth fitting at the base of the central tube when the unit is subjected to knocks and vibration when the vehicle is in motion.

The construction of the basic unit is now complete, with the exception of the lid fitting for the aluminium pipe which feeds the engine. The construction so far has been straightforward engineering with little complication, but the remaining steps in getting the Cell powering a vehicle are not conventional engineering. If you do not feel confident about this construction, then advice and help can be got from the experienced members at the Yahoo Group http://groups.yahoo.com/group/joecellfreeenergydevice/ or alternatively, the companion Group http://groups.yahoo.com/group/JoesCell2 both of which are very active.

Getting the Cell working

The Cell is not just the container and the inner tubes. A major active ingredient of the "Cell" is the liquid placed inside the container. To a casual glance, the liquid appears to be water and loosely speaking it is water. However, water is one of the least understood substances on the planet. It can have many different molecular configurations which give it widely different characteristics. For example, in one configuration, it will actually burn, but this "burning" is nothing like the burning experienced in an ordinary log fire. The water flame is not hot and it is quite possible to hold your hand just over the flame without feeling any heat from it.

We do not want to "burn" the liquid in the Cell. The "conditioned water", for want of a better description, is not consumed when a Cell powers an engine. Instead, the engine is powered by external energy flowing into it. Here, the Cell acts like a lens, concentrating the external energy and focusing it to flow along the aluminium pipe to the engine. This action is not unlike the way in which a magnifying glass gathers and concentrates the sun's energy into a small area to raise the temperature there. The "conditioned water" in the cell, along with the materials and shapes in the Cell, cause the gathering and concentration of this external energy and channel it into the engine.

At this point in time, nobody knows for sure, what the energy is. Earlier, I called it the Zero-Point Energy field, but I have no direct evidence for that, some people call this energy "orgone". Nobody knows exactly how this energy makes the engine run. Engines powered by this energy sound pretty much the same as when they are running on fossil fuels but they run a lot colder and it is usually necessary to advance the timing of the spark. These engines can tick over at a much lower rate than normal and they have much greater power than when running on fossil fuels. Bill Williams in America found that when he fitted a Joe Cell to his Ford pickup, the performance suddenly became like a Formula One racing car and very gentle use of the throttle was needed. He says:

"Over the summer, I used the truck to haul firewood for this winter's wood supply. I added 5 gallons of fuel to bring the fuel level to the half tank mark. I ran the truck with the cell which I installed a month earlier. Basically, I tried to forget about the cell being installed in the truck. The Ignition timing was set at about 25 degrees before TDC with no vacuum connect to the distributor. The fuel line was still connected so "shandy"

mode was being used. The surprising thing is that the truck did not use any fuel during the two and a half months of driving in the woods. In fact, when I parked the truck at the end of the wooding season, I physically sounded the fuel tank (it is a 'behind the seat' tank). It was still showing the half full mark. I pulled the cell for the winter and have it sitting on the bench waiting for spring to arrive for it to be installed again. I don't even pretend to understand this technology, but I keep hoping that someone will come up with a viable explanation of how the cell works".

Anyway, how do we get “conditioned water”? It can be generated inside the Cell, but as the conditioning process usually generates an unwanted residue on top of the water and on the bottom of the Cell, there is an advantage to do the conditioning in a separate container. If water conditioning is done in the Cell, then when the residue is removed, the Cell does not have the correct amount of water and needs to be topped up. That has to be done with non-conditioned water which promptly puts the Cell back to square one. So, use a separate conditioning vat which contains considerably more water than the Cell needs. In the documentary video produced by Peter and Joe, the conditioning procedure is described in some detail.

Joe explains that he conditions the water by suspending an electrode array in the water and applying 12 volts DC to it. Using the water found local to Joe, the current is initially about 10 amps and if left overnight the current drops to anywhere between 2 amps and 4 amps. This indicates that his local water contains a large amount of dissolved material since completely pure water will carry almost no current when 12 volts DC is placed across it. It is almost impossible to get pure water as so many things dissolve in it. Raindrops falling through the atmosphere pass through various gasses and some of these dissolve in the droplets. If the pollution in the atmosphere is particularly bad, then the rain can become acidic and this “acid rain” can rot the trees and vegetation on which it falls. Water on and in the ground, picks up chemical elements from nearly everything with which it comes in contact, so water, any water, needs treatment to reach its “conditioned” state.

Joe’s conditioning electrode array is made up from truncated stainless steel cones, positioned vertically above one another. Joe describes it as being made up from seven cones (not strictly true) with the central cone connected to the battery positive and the top and bottom cones connected to the battery negative. That leaves two unconnected cones positioned between the positive and each of the two outer negative cones. His array looks like this:

What Joe does not mention, but what can be seen in the video, is that there is an eighth cone cut-down and tack-welded in an inverted position underneath the bottom cone:

The inverted cone section appears to project underneath the rim of the bottom cone by an amount of about one inch (25 mm), or perhaps slightly less:

The electrical straps connecting to the cones are insulated to prevent contact with either the other cones or the inside of the metal drum which Joe uses to hold the water being 'conditioned'. He says that if this array is suspended in a tank of water (his happens to be a vertical metal cylinder - a significant shape) and provided with 12 volt DC electrical power for a few minutes, then the water becomes 'charged' as he expresses it. Although the water is supposedly clean, Joe gets gas bubbles coming off the surface of the water. These will explode if lit, so it is very important that this process is carried out in the open air and there is no possibility of the gas ponding on a ceiling.

Joe states that the cleaner the water the better the result. Also, the longer the array is immersed and powered up, the better the result. It is likely that the shape of his powered array is causing the energy field to flow through his water in a concentrated fashion. The water absorbs this energy, and the effect increases with the length of time it is being conditioned, until a maximum level is reached. The objective is to achieve unusually pure water in one of its least usual molecular configurations. The overall procedure is as follows:

1. A vertical stainless steel cylinder, with an open top, is obtained and filled with water. Joe uses a steel beer keg but he selects the keg very carefully indeed from a very large choice of kegs, and then cuts the top off it. There is no need to have such a large container, or cones as large as the ones which Joe uses.

2. The array of cones is suspended vertically in the middle of the water and 12 volts applied to it. The Cell is most definitely **not** any form of electrolyser and should never be confused with one. An electrolyser

operates by breaking water down into hydrogen and oxygen gasses which are then used for combustion inside an engine, and it requires rapid and continuous replacement of the water which gets used up as the engine runs. The Joe Cell never operates in that way, instead it channels outside energy through to the engine and the water inside a Joe Cell is never used up by the engine running. However, in this conditioning process, some hydrogen and oxygen are produced as a side effect of the purification process. Consequently, the conditioning should be carried on out of doors to prevent any hydrogen ponding on the ceiling and forming an explosive mixture there. The more impure the water, the higher the current which flows and the greater the unwanted electrolysis of some of the water.

3. The procedure for applying the 12V supply to the conditioner electrodes is unusual. First, connect the negative supply, and **only** the negative supply. After 2 to 20 minutes, make the positive connection for just 2 or 3 minutes. A residue of impurities will form from this process. Some, being lighter than water, rise to the surface and form a layer there. Some being heavier than water, sink to the bottom. The surface residue is removed and the process repeated until a surface layer no longer forms. This may take 24 hours. The clean water from the middle section of the container is used to fill the Cell.

Many people are of the opinion that a current of about one amp should flow through the conditioning vat in the early stages of the process. If the current is much less than this, then it may take a considerable length of time to get the processing completed - possibly one or two weeks if the water needs a good deal of work done on it. The process can be speeded up by using higher voltage, 24 volts or 36 volts by adding extra batteries or using an electronics bench power supply. The water can also be pre-processed by placing it in a glass jar in an orgone accumulator for a day or two, but that process is outside the scope of this description.

As the impurities get ejected from the water by this process, the electrolysis element gets stifled progressively and as a consequence, the current drops. As completely pure, molecularly-reconfigured water is the goal, no additives of any kind are normally added to the water used to fill the Cell. However, if citric acid is used to clean the cylinders before assembly, there is no harm in allowing them to be assembled in the Cell with traces of the acid on them.

The Cell is filled to just under the level of the top of the inside tube array. This is very important as we need to have separate cylinders of water divided by the steel cylinders. If the water level is over the top of the cylinders, then the whole charging arrangement is destroyed. Further water conditioning inside the Cell may be needed as the cylinders also need to be conditioned. This is done with an easily removable cover replacing the lid of the Cell. The Cell should be kept covered while it undergoes its further conditioning and the lid only lifted briefly to examine the bubbles (unless a glass lid is used). The positive connection to the cell is made to the outside of the 5-inch cylinder and at the top of the cylinder. A length of copper wire tightened around the top of the cylinder is a convenient way to make the connection to the outside (and only the outside) of the cell. Place the cell on a wooden workbench or failing that, on a sheet of high-density plastic such as a chopping board. Connect the negative wire and wait two minutes before connecting the positive wire.

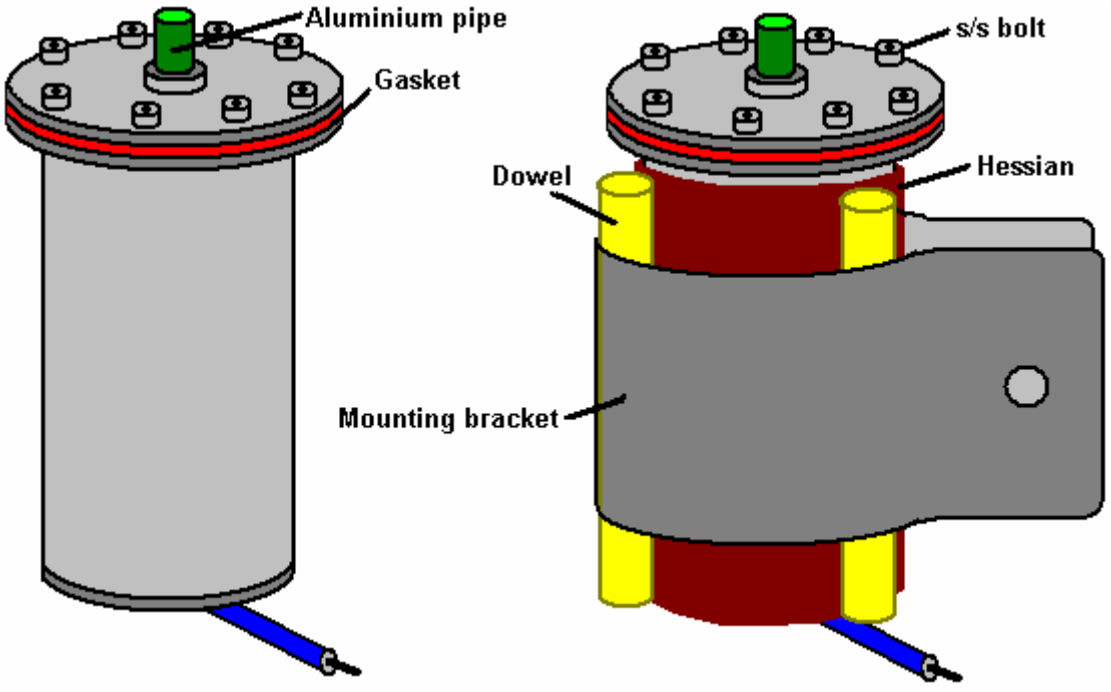
The Cell is ready for use, when it continues to produce surface bubbles for hours after the 12 volt DC power supply is removed from the Cell. The bubbles produced are not part of the energy-focusing process and are themselves unimportant, but they act as an indicator of the outside energy flowing through the Cell. When the Cell is running correctly, the flow of outside energy is sufficient to keep the water in its conditioned state without the need for any external electrical supply. It also maintains its own energy flow through the Cell. There is no point in proceeding any further until the Cell has reached its self-sustaining condition. If it is not happening for you, check out the information in the “workarounds” section below and if that does not get your Cell operational, ask for advice and assistance through the Yahoo groups mentioned above.

Some people concern themselves with the pH of the water. The pH really is not important as the cell will take up the correct pH as conditioning proceeds. A cell of the type described in this document, will have water which is very slightly acid with a pH of about 6.5, but it is not important to know this or to measure it. Do not put litmus paper in the cell water as that will contaminate the cell. Just rely on the action of the bubbles to determine how the cell conditioning is progressing.

Installing the Cell in the Vehicle

When the Cell has reached its self-sustaining condition, it can be mounted in the vehicle. The first step is to insulate the Cell from the engine components. This insulation is not just electrical insulation which is easily accomplished, but it is a case of introducing sufficient separation between the Cell and the engine to stop the

concentrated (invisible) energy leaking away instead of being fed to the engine through the aluminium tube. So, wrap the Cell walls in three layers of double-laminated hessian sacking ("burlap"), pulling it tightly around the 5-inch diameter outer tube. Tie (a minimum of) three wooden dowels along the length of the Cell and bend the mounting bracket around the dowels. The purpose of this is solely to ensure that there is at least a three quarter inch air gap between the walls of the Cell and everything else, including the mounting bracket:



The mounting details depend on the layout of the engine compartment. The really essential requirement is that the aluminium pipe running to the engine must be kept at least 4 inches (100 mm) away from the engine electrics, radiator, water hoses and air-conditioning components.

The last four inches or so, of the tube going to the engine cannot be aluminium as that would cause an electrical short-circuit between the (occasional) positive outer connection to the outside of the Cell and the engine itself which is connected to the battery negative. To avoid this, the final section of the pipe is made using a short length of clear plastic piping, forming a tight push-fit on the outside of the aluminium tube and on the connection to the intake of the engine's carburettor. There should be a 3/4 inch (18 mm) gap between the end of the aluminium pipe and the nearest metal part of the carburettor. If it is just not possible to get an airtight fit on the intake to the carburettor and a hosepipe clamp has to be used, be sure that the fitting is non-magnetic stainless steel. If such a fitting cannot be found, then improvise one yourself, using only 316L grade stainless steel.

In the installation shown above, you will notice that the aluminium tube has been run well clear of the engine components. A vacuum gauge has been added but this is not necessary. For the early stages of installation, the aluminium pipe runs to the vacuum port of the carburettor but stops about 3/4 inch (20 mm) short of it, inside the plastic tubing. This method of connection is advisable for the initial setting up of the vehicle modification. At a later date, when the engine has been running with the Cell and is attuned to it, the Cell operates better if the pipe is connected to one of the bolt heads on the engine block, again using the plastic tube and a gap between the aluminium tube and the bolt head. Some people feel that a safety pressure -release valve with a safe venting arrangement should be used if the pipe feeding the engine, terminates on a bolt head. If it is still available, the http://www.youtube.com/watch?v=DexBoYfDoNw video shows Bill Williams operating his Joe Cell.

Getting the Vehicle Running and Driving Techniques

The Joe Cell is not a 'turnkey' system. In other words, just building a Cell and installing it in the vehicle is not nearly enough to get the vehicle running without the use of a fossil fuel. Some adjustments need to be made to the timing and the engine has to become 'acclimatised' to the energy.

Mount the Cell in the engine compartment and connect the Cell to the battery negative. After two or three minutes, take a lead from the battery plus and touch it briefly to the lid of the Cell. This should produce a spark. Repeat this until four sparks have been produced. This 'flashing' process aligns the Cell electrically and directs the energy to flow in the direction of the metal which has been 'flashed'.

The next procedure is dangerous and should only be carried out with the greatest of care. The engine crankshaft also needs to be 'flashed' four times. This is carried out with the engine running and so can be hazardous - take extreme care not to get caught up in the moving parts. Connect the lead from the battery positive to the shaft of a long-handled screwdriver and keep your hands well clear. The procedure is to get a helper to start the engine, then arc the current to the exposed pulley on the crankshaft (where timing adjustments are made). There should be a total of four sparks to the crankshaft in a period of about one second.

Next, for three or four seconds, flash along the length of the aluminium pipe. This encourages the energy to flow along the pipe, reinforcing the natural attraction between aluminium and this energy. Remove the wire coming from the battery positive as the Cell operates with only the negative side of the battery connected (remember that this is NOT electrolysis and the cell just directs the unseen energy into the engine).

Mark the present position of the distributor cap. Loosen the bolt holding it in place and rotate it to advance the timing by 10 degrees. Disconnect the fuel to the carburettor (do not use an electrically operated valve for this). The engine will continue to run on the fuel left in the carburettor and the engine will start to cough. Turn the distributor cap a further 20 degrees (that is now a total of 30 degrees from its original position) and have your helper use the starter motor to assist the engine to keep turning.

Rotate the distributor cap to further advance the spark until the engine starts to run smoothly. There will be a gasping sound and the engine will slow nearly to a stop, then it will pick up again and then slow down. The action is wave-like, something like breathing. Fine-tune the timing to get the smoothest running and then fasten the distributor cap in place. Do not touch the Cell, but leave it undisturbed. You are now ready to drive away in a vehicle which is not using any fossil fuel.

The procedure described here may not end successfully as just described. Some cars are more difficult to get operating on a Cell than others. Experience helps enormously when getting the vehicle started for the first time. Joe mentions in the video that it has taken him a couple of days of sustained effort to get a particular car going for the first time, which is quite something considering that he has years of experience and has got many vehicles and Cells operational.

When the vehicle has been run and is operating correctly on the Cell, it is time to make the final adjustment to the set-up. For this, the pipe connection to the vacuum inlet of the carburettor is moved from there to terminate on a bolt head on the engine block. The Cell works best when completely sealed off from the air in the engine compartment and as no gas is actually being moved from the Cell to the engine, there is no need for any kind of connection to the carburettor. If the engine is a V-type, then the bolt head chosen should be one in the valley of the V, otherwise, any convenient bolt head on the head of the engine block will be satisfactory. Don't forget that the connecting pipe must still be kept well clear of the engine's electrical leads and other fittings as described earlier. Also, the 3/4 inch (18 mm) gap between the end of the aluminium pipe and the top of the bolt head must be maintained inside the clear plastic tube, and the pipe fitting should

remain airtight. A slight timing adjustment may be necessary with the new connection in order to get the very best running.

The energy which powers the engine has a tendency to run along magnetic fields. Driving under high voltage overhead power lines can position the vehicle in an area where the energy level is not sufficient to maintain the energy flow through the Cell. If the energy flow through the Cell is disrupted, then it is likely to stop functioning. If this were to happen, then the Cell would have to be set up again in the same way as for a newly built Cell which has never been used before. This can be avoided by attaching an AA ("penlight") dry cell battery across the Cell with the battery plus going to the lid of the Cell. A battery of this type has such a high internal resistance and so little current capacity that no significant electrolysis will take place on the very pure conditioned water in the Cell. But the battery will have the effect of maintaining the integrity of the Cell if it is temporarily moved away from its source of power.

Suppliers

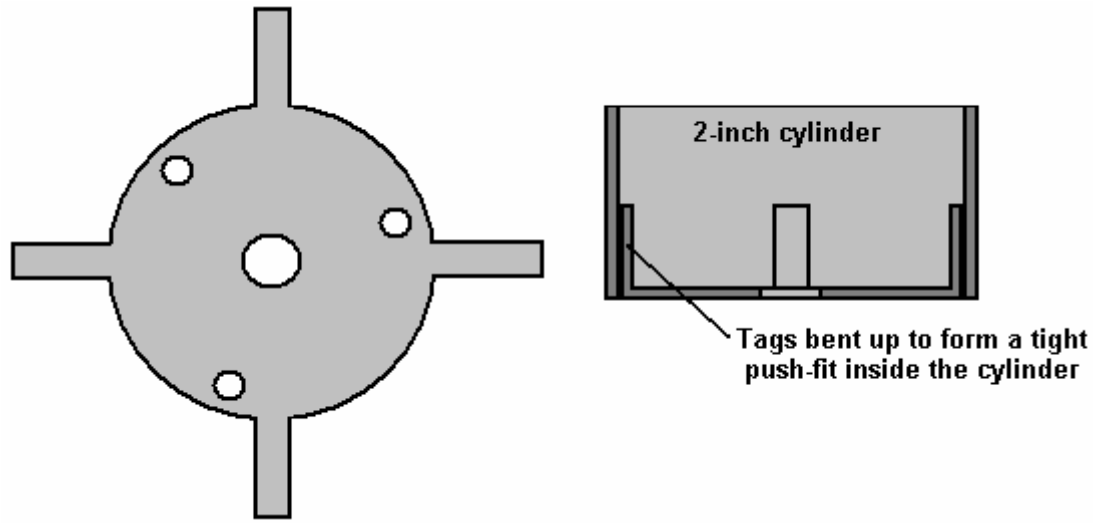
Sheets of nitrile rubber NB70 ("Buna-n") : http://www.holbourne.co.uk
Nylon rod: http://www.holbourne.co.uk
Stainless steel tubing: http://www.stabarn.co.uk
A4 Bolts (316 S31 stainless): http://www.a2a4.co.uk

Workarounds

If it is not possible to get pipes of the desired diameters, then they can be made up by rolling stainless steel sheet and using a TIG welder with completely inert gas, to tack weld at each end and in the middle of each cylinder. Don't weld along the full length of the join unless it is the 5-inch outer casing.

If it is found to be particularly difficult to make the four circular cuts in 1/8 inch (3 mm) steel using hand tools, then I would suggest using a plasma cutter. Make a template to guide the cutting head and clamp it securely in place. You can hire the cutter and compressor quite cheaply as you will only need them for a very short time. If they are not given to you as a pair and you have to select each from a range, take the smallest cutter and a twin-cylinder compressor rated at nearly double the input quoted for the cutter. This is because the cutter is rated by the volume of compressed air, and the compressors are rated by the volume of their uncompressed air intake as that sounds more impressive.

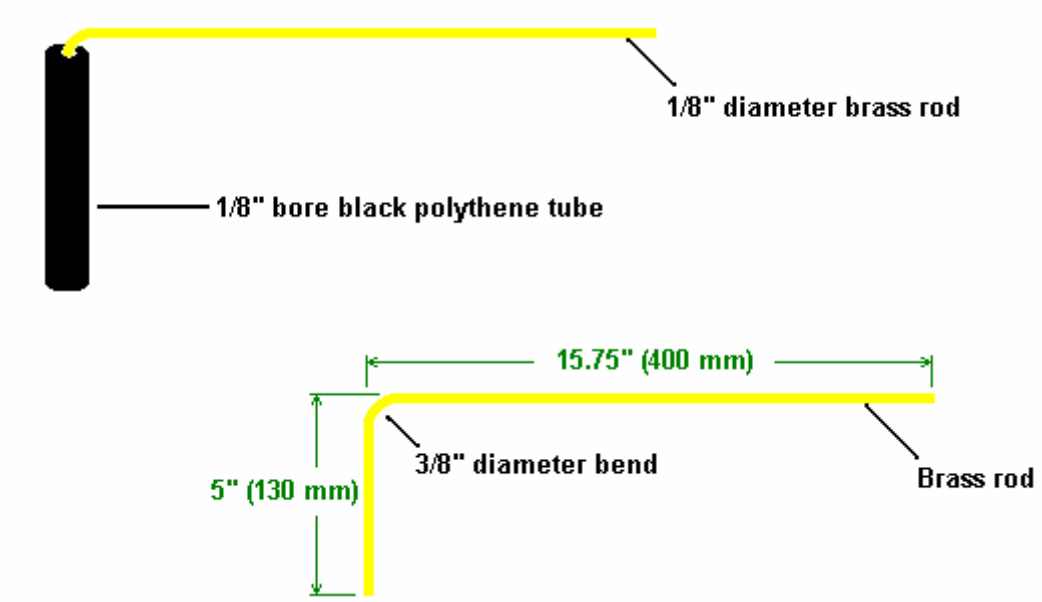
If no lathe is available for machining the base plinth for the central cylinder, then take a piece of 16-gauge stainless steel sheet and cut the plinth out of it as shown below. Bend the projecting tags upwards by holding each tag in the end of the jaws of a vise and tapping the body section square, with a flat-faced hammer and if you consider it necessary, tack-weld the top of the tags to the outside of the central cylinder to give rigidity to the mounting. Extreme heat such as is generated by welding or cutting tends to create permanent magnetism in any ferrous metal being heated, so avoid high temperature operations such as welding wherever possible. If a tight push-fit can be obtained with the base of the 2-inch cylinder, then I suggest that the optional spot welds are omitted.



If tack-welded cylinders have to be used, then it is usually best to line all of the seams up as the seam area does not work as well as the remainder of the tube, so if the seams are all aligned, then there is only one small line in the Cell which is not operating at its optimum value.

Cylinders are best aligned in the same direction. This sounds odd as they are physically symmetrical. However, these cylinders will be used to channel an energy field and each cylinder has a direction along which the energy flows best. To find this, stand all of the tubes upright in a tight group on a table. Leave them for a minute and then place your hand on top of the whole set. If any tube feels hotter than the others, then it is out of energy alignment with the rest and should be inverted. Repeat this test until no tube feels hotter than the rest.

An alternative way to do this test is to use a pair of L-rods. These can be made from two short lengths of rigid black polythene tubing often found in garden centres for use in garden irrigation. This tubing has 1/8 inch internal diameter and so takes 1/8" brass welding rod very nicely. The welding rods should be bent with a radius as shown here:



The curved bend in the brass welding rod helps to prevent the rod fouling the top of the plastic tube handle and it allows free rotation of the brass rod. It is essential that the rod can move completely freely in the handle. If two of these are made up, they can be used to check the cylinders before they are assembled for insertion into the Cell. Place a tube standing vertically on a table well away from all other objects (especially magnetic and electrical items). Hold an L-rod handle in each hand so that the rods are parallel in front of you. The rods must be exactly horizontal so as to avoid any tendency for them to turn under the influence of gravity. Approach the cylinder. The rods should either move towards each other or away from each other as the cylinder is approached.

Repeat this procedure at least three times for each cylinder so as to be sure that a reliable result is being obtained. Invert any cylinder if necessary, so that every cylinder causes the rods to move in the same direction. Then assemble the Cell, maintaining that alignment of the cylinders during the assembly.

If you are having difficulty in getting the Cell operational, then try striking and sparking the cylinders again. This is done as follows:

1. Take a 12V lead-acid battery and position it so that it's negative terminal is pointing towards East and it's positive terminal is pointing towards West (i.e. at right angles to the Earth's magnetic field).
2. Attach a lead from the battery negative to the outside of the base of the tube.
3. Lay the tube on a table and strike it with a hammer along its length. If the tube has a seam, then strike the tube along the length of the seam.
4. Connect a lead to the positive terminal of the battery and spark the inside of the top of the tube. It is essential to spark each tube if they have been polished. It is better not to polish any of the tubes.
5. Repeat this procedure for each tube.

If you consider it necessary to clean the cylinders, then, considering the lengths you went to remove all of the things dissolved in the water, be sure to avoid using any kind of chemical or solvent. You can electro-clean them by using the following procedure:

Starting with the largest cylinder;

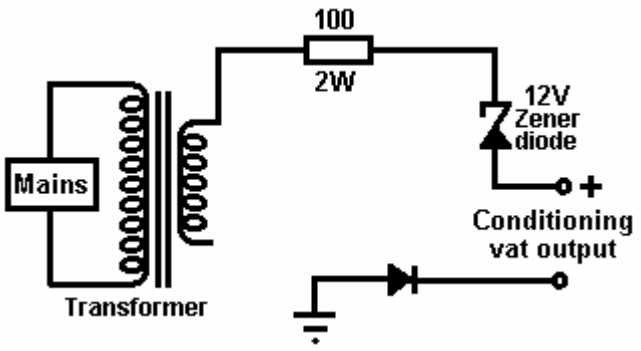
1. Put the battery positive on the inside of the top of the cylinder, and the negative on the outside at the bottom, and leave them in place for one minute.

2. Put the negative on the inside of the top of the cylinder, and the positive on the outside at the bottom, and leave them in place for one minute.

3. Repeat step 1: Put the battery positive on the inside of the top of the cylinder, and the negative on the outside at the bottom, and leave them in place for one minute.

Do this for all cylinders, working inwards.

It has been suggested that an improved method of conditioning water to fill the Cell can be achieved if pulsed DC is used instead of straight DC from a battery. This has not been proven but there is a reasonable amount of information to suggest that this is likely. The following, most unusual circuit, has been suggested, but it must be stressed that it is untried and anybody who is unfamiliar with working with electronics should not attempt to construct or use this circuit without the assistance of a person who is experienced in building and using mains equipment.



This is a most unusual circuit. A 12V step-down mains transformer provides 12V AC which is taken through a limiting resistor and a zener diode which would not normally be connected as shown. The really odd thing is that the circuit which contains the secondary of the transformer appears not to be connected. The expected output from this very odd circuit is pulsing DC of odd waveform, all of which is positive relative to the ground connection, which is a literal, physical connection to an earthing rod driven into the ground.

Notes:

Engines running while powered by a Joe Cell act in a somewhat different manner. They can idle at a very low number of revs per minute, the power available on acceleration is much greater than normal and they appear to be able to rev very much higher than ever before without any difficulty or harm.

The type of Cell described in this document was built by Bill Williams in the USA with the help and assistance of Peter Stevens of Australia. Bill describes his first driving experience with his 1975 F 250, 360 cu. in. (5.9 litre) Ford pickup:

Well, all I can say is "who needs an Indy car when you can drive an old FORD" – WOW!!!! The first five miles after leaving home were wild. I had to be extremely careful on how I pressed the accelerator. I gingerly crept up to 45 mph and that was with moving the pedal maybe half and inch. The throttle response was very crisp or touchy. With about a 1/8" of movement the next thing I new I was close to 80 mph. If I lifted off ever so slightly on the throttle, it felt like I was putting the brakes on and the speed would drop down to 30 mph or so. "Very erratic". If I barely even touched or bumped the pedal it felt like I had pushed a nitrous oxide booster button. WOW !!!

As stated earlier, the first 5 miles were wild and things started to change. The engine started to buck or surge with very large rpm changes and literally threw me against my seat belt. It got so bad I just took my foot completely off the pedal and rode the brakes to stop the truck. The truck left skid marks on the pavement every time the engine surged in rpm. Well anyway, I manage to get it stopped and shut it off with the ignition key - thank GOD !

I retarded the timing, turned the gasoline back on, crossed my fingers and hit the ignition key, and the engine took right off, revving to maybe 4,000 rpm and then gradually decreased to 700 rpm. I took a deep breath and put it into drive and the truck responded close to normal again. I made it into work a little late, but late is better than never the way I see it. After working during the day at the job and thinking what I could do to stop this erratic rpm oscillation, I decided to disable the cell and drive home on gas. WOW !!!

Peter Stevens states that the main reason for the erratic behaviour of the Cell was due to outside air leaking into the Cell, and he stresses that Cells need to be completely airtight. It is also clear that the timing was not set in the correct position. All properly built Cells give enhanced engine power.

Water Conditioning:

Note: With the Cell design shown later, no water conditioning is needed. Please be aware that water quality and purity varies enormously from place to place. One experienced cell builder says: I use water taken from the start of rivers. Further down the river, the water will have encountered influences which are not helpful. My favourite water catchment area well is outside Melbourne, Australia, where there are no roads, power lines, dams, pipes or any man made intrusions, the water flows how and where it wants to in natural, twisty downhill paths it has created, the whole area is green all year round and you can feel the vitality and Nature at work.

This water has a pH of 6.5. That means it is slightly acidic, and perfect for Joe Cells. I bring this water home making sure that I protect it from excessive sloshing and the heat of the sunlight whilst in the car. At home, I store it in 20 litre Pyrex bottles. Do not store it in plastic containers even if the container is marked "suitable for water". Earthenware or wood containers would also be very suitable.

I make an electrolyte solution by dissolving 500 grams of food-grade phosphoric acid and 100 grams of sodium perborate, in three litres of de-ionised water or distilled water. Just a few drops of this solution will provide a current of 1 amp at 12 volts in the conditioning vat. An alternative is to use a 90% acetic acid solution which has no stabiliser in it.

When conditioning the water in the cell, you will need a lid, or some way of sealing of the cell from air. A lid loosely sitting on top of your test jar is sufficient. The seeding and breeding process is hampered by having too great an area of the top of the cell being exposed to air. All lids are not the same as regards to being a obstruction to orgone. If the lid does not seem to be working, place a layer of aluminium foil underneath the lid and use the foil and lid as one unit.

The aim is to modify the conductivity of the water by the addition of acid, so as to get a suitable current flow. If we used de-ionised water with a pH of 7.0, we would have a very low current flow for our electrolysis, and would have to add something to increase the conductivity of the water if we wanted observable results in a short period of time. As we lower the pH, the current flow and electrolysis process will increase together with a heat increase.

We are trying to achieve electrolysis action with the minimum heat generation. As the propagation of orgone is reasonably slow, there is not much to be achieved with excessive current. Slow and steady does it. For the patient experimenter or one that is using neat water, i.e. water without electrolyte, excellent results are achieved with currents as low as 50 milliamps.

The procedure is:

1. Place your cell on a wooden work bench or on a sheet of plastic type material or, as a last resort, on a newspaper. We are trying to insulate the cell from metal paths that may impede the seeding process. Keep the cell well away from electrical sources such as a television set, refrigerator, electric cooker, etc.

2. With a multimeter, measure the resistance between the innermost and the outermost cylinders of your cell. It should be in the high Megohm range. If not, the insulators are conductive or there is a short-circuit. Check for a short-circuit and if there is none, remove the insulators and reassemble the set,

checking the resistance between the innermost and outermost cylinders as each cylinder is added. The resistance between every pair of cylinders should be very high.

3. When all is okay in the above step, fill the cell using a funnel containing a paper coffee filter. Fill it only to a level just under the top of the cylinders and no more. The effect that we want to create is a set of water cells separated by metal cylinders. These are your alternate organic and inorganic chambers. Of course, the submerged section of you chambers are flooded, but with this simple cell, the top will be doing all the work. This is why the cylinders should be completely horizontal and true at the top, otherwise the meniscus formed by the water would not work and the water would flow from compartment to compartment. This level is only critical during the seeding process, as we require maximum orgone capture to seed the cell. Naturally, with a charged cell, the water is sloshing all over the place whilst you are driving the car.

4. Turn on the power supply, and if it is adjustable, set it to 12 volts. Connect the negative end of your power source to one end of your meter that is set up to read a minimum of 2 amps and connect the other end of the meter to the bottom of the central cylinder. Wait for two minutes and then connect the positive end of your power source to the top of the outer cylinder. What you have done is set up the meter to read any current flow into your cell from the power source.

   At this stage, if your water is close to a pH of 7, as previously discussed, the current flow will be zero, or in the low milliamp region. If the current flow is amps, then you are doing something wrong! It is impossible to pass a huge current through ordinary pure water when using 12 volts. Think about it. To draw even 1 amp at 12 volts, the resistance of the water would have to be 12 ohms! No way! You are doing something wrong. Correct the problem and then move on.

5. Presuming that the current is only milliamps, you now want to introduce electrolyte to increase the current flow through the water. The aim is to get a current flow of about one amp. To do this, drip a small amount of your chosen electrolyte into the cell water whilst stirring and watching the current measurement. Use a glass, Perspex or wooden dowel rod as the stirrer - do not use your handy paint-stirring screw driver! Throw away the stirrer when finished as it will have absorbed some of the cell contents. Do plenty of gentle stirring of the water as you add the electrolyte, otherwise you will add too much electrolyte. Stop adding electrolyte when the meter indicates 1 amp. Your water level may rise as a consequence of the addition of electrolyte. Remove some water from your cell. I use a pipette, so as not to disturb the cell. Remove enough water to again just expose the top of the cylinders. At this stage, disconnect your meter and power source and have a bit of a clean up as the next stages are guided by observation.

The charging process is separated into three distinct stages which are called Stages 1, 2 and 3. These stages have both some obvious differences and some subtle ones. For the rest of the charging process, you will be only connecting your power source to the cell for a maximum of 5 minutes at a time. As orgone lags electricity by about 30 seconds, you will know the state of the cell in less than a minute. Do not be tempted to leave the power connected to the cell for long periods! Yes, I know that you are in a hurry and more is better, but in this case you only generate heat, steam, waste power and overheat the cell. You can pick the failures by seeing their cells running non-stop for days with 20 or more amps turning the water to steam, etching the cylinders and ending up with a barrel full of scum. What else would you expect? After all, electrolysis is time and current related. If you have had the misfortune of having your cell left on for a long period with high current, you have probably destroyed your cylinders. You cannot retrieve the situation so throw the cell away and start again. I bet you don't do it next time!

Danger: Do **not** charge any cell that is totally sealed! The cell will explode, with all the resulting consequences. An airtight seal **is not required !** At no stage do I prescribe any form of airtight container.

Stage 1: This stage is plain old electrolysis. Due to passing direct current through a liquid which contains ions, chemical changes will occur. In our case, you will see small bubbles and a cloud of activity that is greater nearest the outside of the innermost negative cylinder. The important observation points are that the activity is greatest nearest the central cylinder and gets progressively less as we move outward via the different chambers formed by the rest of the cylinders. Also, within a short period of turning the power off, all activity stops, the water becomes clear and the bubbles disappear.

Every fool and his dog can reach Stage 1. The secret for progressing further is to restrain your impatience and not increasing the electrolyte concentration to raise the current (and/or leaving the cell on for days on end). Be patient, leave the cell on for no longer than 5 minutes, turn the power source off, remove the leads to the cell, and put the top on the test cell, or partially block off the exit of the car cell. It does not have to be

airtight! Go and do something else. It is like waiting for a tree to grow from the seed. Do this on a daily basis for days, or a week, or longer, until you get to Stage 2. You will find that the more "alive" the water is , the quicker is the seeding of the cell. I have found that the storage, age, and source of the water all affect the seeding speed. I have also found that by changing the structure of the water by various means e.g. vortexing, shaking, filtering, etc., you can greatly enhance the water quality to make it more "alive".

Stage 2: You will now notice on your initial powering up of the cell, that the bubbles are getting larger and the white cloud of tiny bubbles in the water are much smaller or more transparent. Also in Stage 1, you had the action occurring mainly near the central cylinder. Now the bubbles form in a regular fashion irrespective of their location in the cell. More importantly, on turning the power off from the cell, the bubbles do not go away immediately but stay there for minutes rather than seconds as in Stage 1. Also, the top of the water assumes a glazed look and the meniscus is higher due to a change in the surface tension of the water. At this stage you may have some brownish material amongst your bubbles. Don't panic - it is only the impurities being removed from the cell. I find that if I wipe the top surface of the water with a paper towel, the bubbles and the deposit will adhere to the paper and can be removed easily. Top up the cell with water from your charging vat, if required, after the cleaning, so that again, the top edges of the cylinders are just showing. No more electrolyte is added! In cleaning the top of the cell as described, it has been observed that some people react unfavourably with the cell. If so, keep that person away, or if it is you, try changing your hand i.e. use your right hand instead of your left or vice versa. If the presence of your hand seems to collapse the surface bubbles, I would suggest you get a friend to do the work for you.

Summary of Stage 2: The result is very similar to Stage 1, but now we have a more even bubble distribution and an increase of surface tension and a longer presence of the bubbles when the power is turned off. There will be no scum in the bottom of the cell and the water will be crystal clear. At this stage the orgone has seeded the cell, but as yet, is not “breeding”, that is, the orgone concentration is not yet great enough to attract additional orgone flow to itself. With the right cell, water and operator, it is possible to go straight to Stage 2 on the first turn on of a new cell.

Stage 3: Not many people get to this stage, or what is worse, get here incorrectly. If you get here following the above steps, your water is still crystal clear with no deposits in the sump. If you get here by brute force, you will have stripped appreciable amounts of material from the cylinders and this material will now be deposited on the insulators and suspended in the water as tiny particles which never settle out, and finally, the material will form a deposit at the bottom of the cell. The low resistance insulators and the metallic particles in the water will create a cell which leaks orgone and consequently it will cause endless mysterious car stoppages or refusals of the car to start.

Right, the miracle of Nature is now breeding in your cell. Upon turning your power on to the cell, within 30 seconds copious beautiful white bubbles will rise from all the surface area of the cell. Before these bubbles cover the water surface, you will notice a slowly rotating and pulsing front in all cylinders, that is synchronised and has a regular rhythm of about 2 pulses per second and a clockwise rotation speed of about 1 revolution every 2 seconds. These effects are very hard to observe for a first time viewer who does not know what to look for. I find it easier to watch these effects with the aid of a fluorescent light, as the 100 cycles per second pulsations of the light "strobe" the water surface and help the observation.

The bubbles may overflow the container and show great surface tension. One of the definite proofs that the cell is breeding is that, on turning the power source off and coming back the next day, most of the bubbles will still be on top of the water as opposed to Stage 1 or Stage 2 where they disappear in minutes. There is no way that you can mistake this stage. The bubbles are larger and pure white, the surface tension is greater, the bubbles are pulsating and most importantly the surface tension remains days after the power has been removed.

I do not recommend any additional tests or measurements. But for those who are incapable of leaving things be, they may measure the voltage across the cell after it has been left standing with the power off for at least 24 hours. A Stage 3 cell will have a residual voltage, or more correctly, a self-generated voltage of around 1 volt. A Stage 1 cell measured under similar conditions will read 0.1 to 0.2 volts. Remember, that unless you know what you are doing, these voltage measurements can be very misleading due to probe materials and battery effects that can easily mask your true measurement. As the cell reaches the maximum density of orgone that it can hold, the result of the breeding process is the conversion of this excess orgone into the formation of electricity. As such, electrical measurement with the correct instruments is a very valuable method in the verification of the efficiency of the cell. If you are conversant with the work of William Reich, you may care to make an orgone meter and thus remove all guesswork. This meter is fully described on some web sites.

I do not recommend any form of bubble exploding. As noted earlier, noise and vibration are orgone-negative. Therefore, these explosions applied during the delicate seeding period will kill your cell. Apart from a dead cell, the chance of fire igniting other gasses in the workshop and injuries to the ears etc. makes this exercise highly unnecessary. I must admit that I too fell for the "go on, ignite it!" feeling. I had a cell that had been at Stage 3 for seven months. It was my favourite test cell. My hands and matches fought my brain and they won. There was a huge "ear-pulling, implosion/explosion", and yes, I killed the cell. It went back to Stage 2 for four days. I will not do it again.

As all water we are using so far has been electrolysed, this water is not suitable for use in non-stainless steel or glass containers due to reaction with the container and the resultant corrosion, but if you have to, or want to, you can use juvenile water with no electrolyte added and still charge it to Stage 3. As the ion count is much lower, the water is not as conductive, i.e. you cannot get as much current flow with 12 Volts as you would if you electrolysed the water. However, if you obtain a power supply of approximately 60 to 100 Volts at about 1 Amp, you will be able to charge "plain old ordinary water". The down side is the additional waiting, in some cases, over 3 weeks, and the cost of the fairly expensive power supply. The advantage is that you will be able to pour it into the radiator of a car with no increase in corrosion as compared to water containing acids.

Do not at any stage short circuit, i.e. join any of the cell cylinders to each other electrically with your charging leads, wedding ring, etc. If you do, the cell will "die"! Your only option, if this occurs, is to connect the cell to your power source and see if you are still running at Stage 3. If the cell does not revert to running in Stage 3 mode within 1 minute, your only option is to completely dismantle the cell and re-clean and re-charge. Huh???, you are kidding us, right??? No, I am serious, that is your only option! So do not do it, do not short out your cell! You will have similar, but not as severe problems if you reverse your leads to the cell.

When the cell is running at Stage 3, you can tip the charged water out of the cell into a glass container and clean, adjust or maintain your now empty cell. Try to keep all cylinders in the same relation that they were in before you dismantled the cell, i.e. keep all cylinders the same way round and in the same radial alignment. This is mainly relevant when dismantling cells over 6 months old as the metal parts develop a working relationship that can be weakened or destroyed by careless re-assembly.

When finished, pour the charged water back and you are back in business. Of course you can pour this charged water into other cells, or use it as you see fit, but, remember, do not leave it out of the cell for periods longer than 1 hour at a time as the breeding has now stopped and you are slowly losing charge.

Troubleshooting.

It is usually quite difficult to get an engine running from a Joe Cell. Many people find it difficult to get their Cell breeding (“at Stage 3”). The following suggestions from various experienced people who have succeeded are as follows:

1. The metal construction of the Cell needs to be of stainless steel and nothing else. Using copper or brass, even for something as simple as the connector between the Cell and the aluminium tube running to the engine is sufficient to cause serious problems as the energy is not directed to the engine and just leaks away sideways.

2. The water is best charged in a separate vat which has a larger capacity than the Cell itself. That way, when the Cell is being conditioned and scum removed from the surface of the water, the cell can be topped up with charged water from the vat. If, instead, ordinary, uncharged water is used, then the whole process is liable to be put right back to square one.

3. Be very sure that the mounting in the engine compartment is electrically insulated from the engine and chassis and be sure that there is serious clearance between the Cell and everything else. Also, the aluminium pipe running to the engine **must** be kept at least four inches (100 mm) clear of the main engine components. Otherwise, the energy which should be running the engine, will leak away sideways and not reach the engine.

4. It can take up to a month to get a steel engine acclimatised to a Cell. Run the engine as a “shandy” where fossil fuel is still used but the Joe Cell is also attached. This usually gives greatly improved mpg, but more importantly, it is getting the engine metal and cooling water ‘charged’ up ready for use with the Joe Cell alone. Once per week, try advancing the timing and see how far it can be advanced before the

engine starts to ping. When the timing gets to a 20 or 30 degree advance, then it is time to try running on the Joe Cell alone.

5. Finally, having conditioned the Cell, the water, the engine and the coolant, if there is still difficulty, then it is probably worth conditioning yourself. Both the idea and the procedure sound like they have come from Harry Potter's classes in Hogwarts School of Witchcraft and Wizardry. However, there is a serious scientific basis behind the method. Use of the Bedini battery-pulsing devices shows that lead/acid batteries act as a dipole for Radiant Energy. Also, the energy flow which powers the Cell appears to move from West to East. Bearing those two facts in mind, makes the following rather bizarre procedure seem slightly less peculiar:

(a) Get a car battery and position it so that it's terminals line up East/West with the negative terminal towards the East and the positive terminal towards the West (along the main energy flow line)

(b) Stand on the North side of the battery, facing South.

(c) Wet the fingers of your right hand and place them on the battery's negative terminal (which is on your left hand side).

(d) Keep your fingers on the terminal for two minutes.

(e) Wet the fingers of your left hand. Place your left arm under your right arm and place the fingers of your left hand on the positive terminal of the battery. Do not allow your arms to touch each other.

(f) Keep the fingers of your left hand on the positive terminal for three minutes.

(g) Remove your left fingers from the positive terminal, but keep the fingers of your right hand on the negative terminal for another 30 seconds.

This procedure is said to align your body with the energy flow and make it much easier for you to get a Cell to "Stage 3" or to get a vehicle engine running. In passing, some people who suffer continuing painful medical conditions state that they have got considerable pain relief from this procedure.

Recent Joe Cell Developments.

One of the greatest problems with using a Joe Cell has been to get it operational. The reason for this has probably been due to the lack of understanding of the background theory of operation. This lack is being addressed at this time and a more advanced understanding of the device is being developed. These design dimensions cause ordinary tap water to go immediately to the fully functional "Stage 3" and remain in that state indefinitely, the only way of stopping the Cell is to physically take it apart.

While it is still rather early to draw hard and fast conclusions, a number of results indicate that there are three separate, unrelated dimensions which are of major importance in constructing a properly "tuned" Joe Cell. It needs to be stressed that these measurements are very precise and construction needs to be very accurate indeed, with one sixteenth of an inch making a major difference.

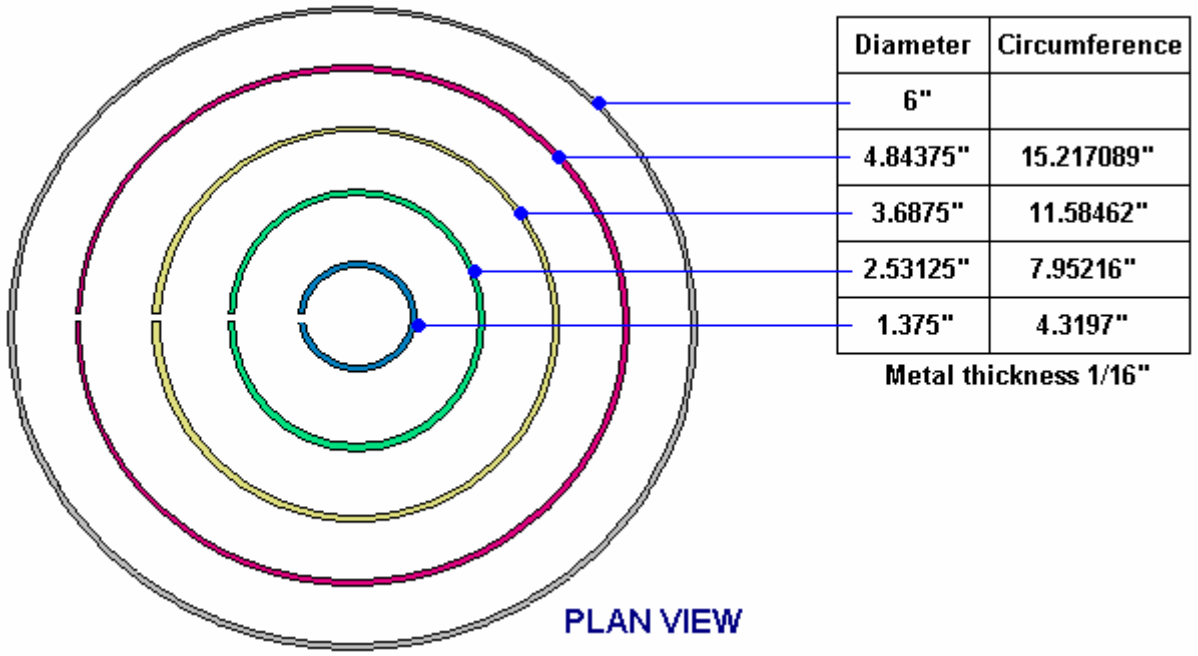
The dimensions are specified to this degree of accuracy as they represent the tuning of the Cell to the frequency of the energy which is being focussed by the Cell. The fact that there are three separate dimensions, suggests to me that there are probably three components of the energy field, or possibly, three separate energy fields.

These three dimensions have been assigned names and are as follows:

Golden dimension: 1.89745" (48.195 mm)
Blue dimension: 3.458" (87.833 mm)
Diamagnetic dimension: 0.515625" (13.097 mm)

It is suggested that a Joe Cell should be constructed with cylinder heights which are a multiple of either the 'Golden' or 'Blue' length. Also, the water height inside the container should be below the tops of the inner cylinders and be a multiple of the basic length chosen for construction. The inner cylinders should be positioned the 'Diamagnetic' dimension above the base of the Cell. They should also be constructed from stainless steel of thickness 0.06445" (1.637 mm, which is very close to 1/16") and there should be a horizontal "Diamagnetic" gap between all of the vertical surfaces.

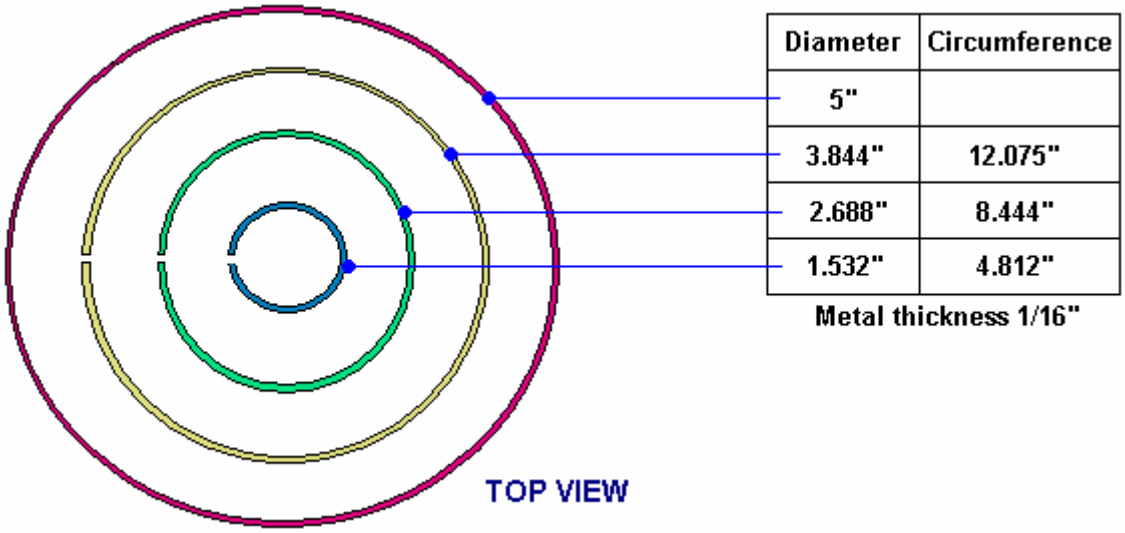
The inner cylinders should be constructed from stainless steel sheet which is tack welded at the top and bottom of the seam, and all of the seams should be exactly aligned. The lid should be conical and sloped at an angle of 57^{O}, with it's inner surface matching the inner surface of the housing and the inner surface of the outlet pipe. The outer casing should not have any dome-headed fasteners used in its construction. The length of the outlet pipe should be made of aluminium and should be 15.1796" (385 mm) for 'Golden' height cylinders or 20.748" (527 mm) for 'Blue' height cylinders. That is 8H for Golden and 6H for Blue and should there be a need for a longer pipe, then those lengths should be doubled or tripled as the single dimensions no longer apply (this being a fractal effect). At this point in time, these are only suggestions as the science has not yet been firmly established. One possible arrangement is shown here



Diameter	Circumference
6"	
4.84375"	15.217089"
3.6875"	11.58462"
2.53125"	7.95216"
1.375"	4.3197"

Metal thickness 1/16"

It is not necessary for there to be four inner cylinders so an alternative might be:

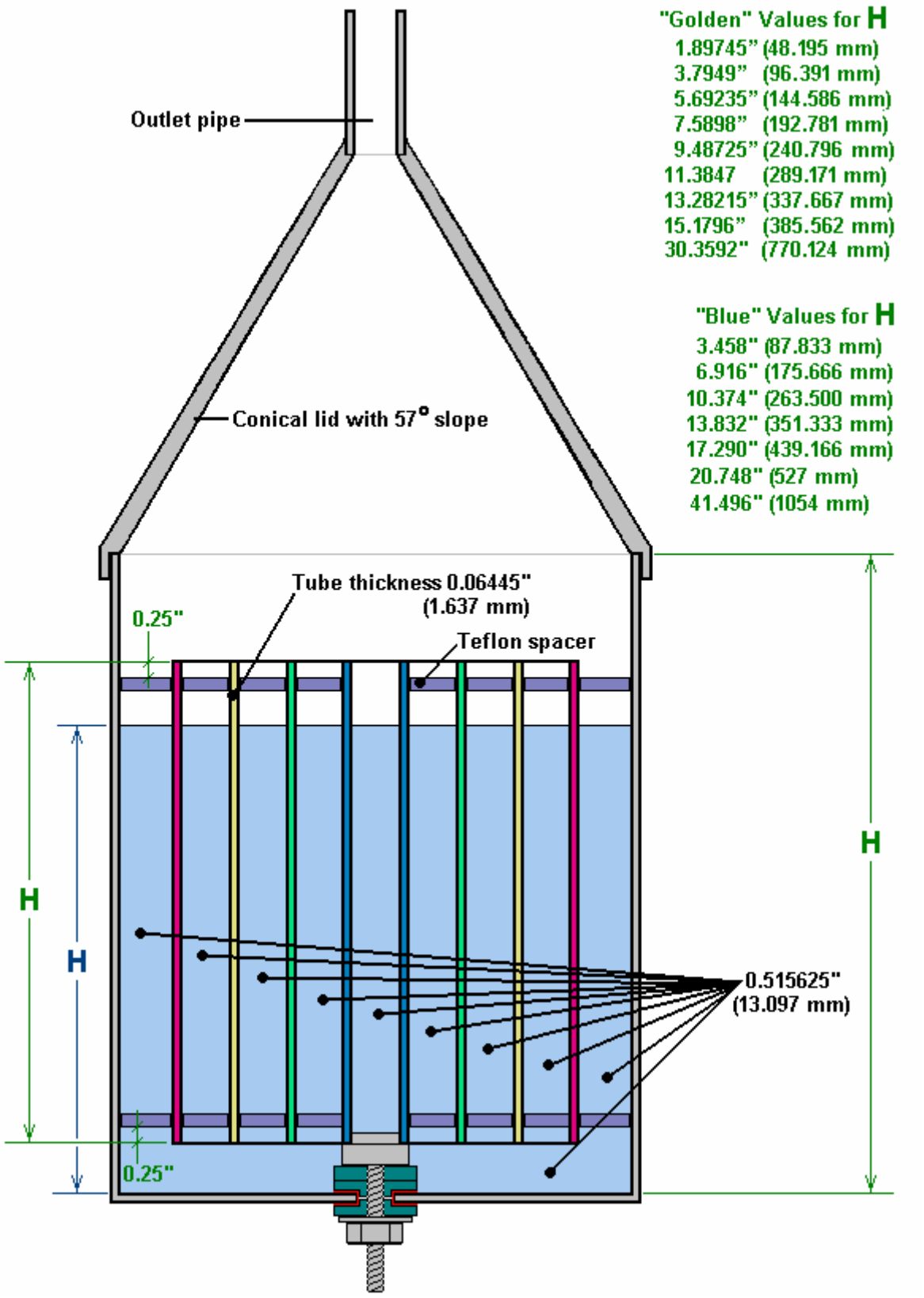


Diameter	Circumference
5"	
3.844"	12.075"
2.688"	8.444"
1.532"	4.812"

Metal thickness 1/16"

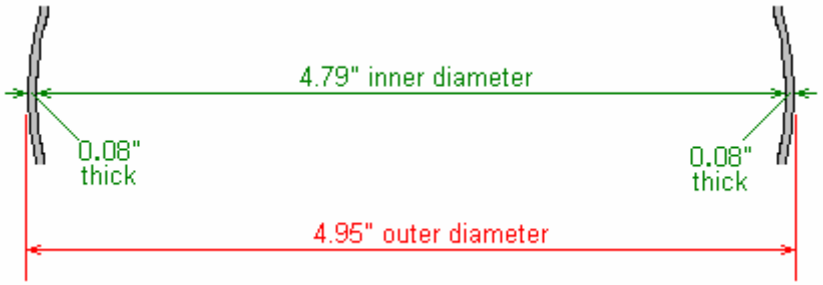
A suggested Joe Cell design is shown below. This diagram shows a cross-section through a Joe Cell with four inner concentric stainless steel tubes. These tubes are positioned 0.515625 inches (13.097 mm) above the bottom of the Cell and the gap between each of the tubes (including the outer casing) is exactly that same 'Diamagnetic' resonant distance.

It should be clearly understood that a Joe Cell has the effect of concentrating one or more energy fields of the local environment. At this point in time we know very little about the exact structure of the local environment, the fields involved and the effects of concentrating these fields. Please be aware that a Joe Cell which is properly constructed, has a definite mental / emotional effect on people near it. If the dimensions are not correct, then that effect can be negative and cause headaches, but if the dimensions are correct and the construction accurate, then the effect on nearby humans is beneficial



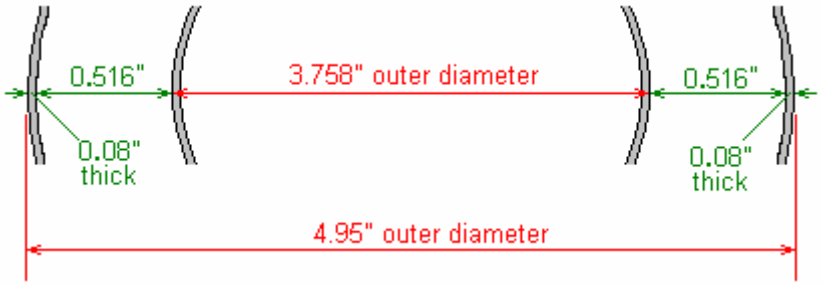
It should be pointed out that Joe Cells will be constructed with the materials which are readily to hand and not necessarily those with the optimum dimensions. If picking stainless steel sheet which is not the suggested optimum thickness, then a thinner, rather than a thicker sheet should be chosen. In case the method of calculating the diameters and circumferences of the inner cylinders is not already clear, this is how it is done:

For the purposes of this example, and not because these figures have any particular significance, let's say that the steel sheet is 0.06" thick and the outer cylinder happens to be 4.95" in diameter and it is 0.085" thick. People wanting to work in metric units can adjust the numbers accordingly where 1" = 25.4 mm.

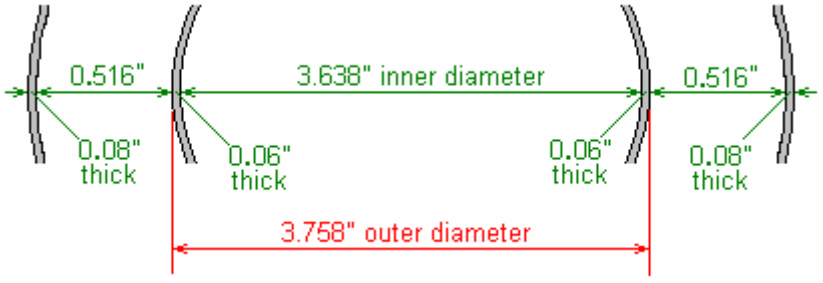


Then, the inner diameter of the outside cylinder will be its outer diameter of 4.95", less the wall thickness of that cylinder (0.08") on each side which works out to be 4.79".

As we want there to be a gap of 0.516" (in practical terms as we will not be able to work to an accuracy greater than that), then the outside diameter of the largest of the inner cylinders will be twice that amount smaller, which is 3.758" :



And, since the material of the inner cylinder is 0.06" thick, then the inner diameter of that cylinder will be 0.12" less as that thickness occurs at both sides of the cylinder, which works out to be 3.838" :



The length of stainless steel needed to form that cylinder will be the circumference of the outer diameter of 3.758" which will be 3.758" x 3.1415926535 = 11.806 inches.

The dimensions of the other inner cylinders are worked out in exactly the same way, bearing in mind that every steel thickness is 0.06". The results for three inner cylinders would then be:

Diameter	Circumference
4.95"	
3.758"	11.806"
2.606"	8.187"
1.454"	4.568"

OR

Diameter	Circumference
125.7 mm	
95.5 mm	299.9 mm
66.2 mm	208.0 mm
36.9 mm	116.0 mm

Comments from an expert in July 2012:

That information is really ancient history and it was Dave's best guess at optimising a cell at that time. Later testing left us disappointed with the 'Blue' lengths, however the 'Golden' lengths make a very nice cell, but not one which is particularly good with an internal combustion engine but one which is better as a learning tool, or for use in healing.

We are into an entirely different approach now, one which entails introducing specific vibrations into the cell. An optimum implementation involves cutting each tube to a specific length so as to make it self-exciting, but that's not necessary because the frequencies can be introduced just using a caliper, or a precise length of metal touched against the tubes in a sequence. Since this approach was totally different from traditional Joe Cell work, we set up a discussion group specifically for it: http://tech.groups.yahoo.com/group/vibrational_combustion_technology/

The nice thing about this approach is that it's ultra stable. Once the vibration is set up the only way to stop it is to take the cell apart. This construction method totally eliminates the human influence factor problem! In fact, a cell can affect the engine even without there being water in the cell. Another nice thing about it is the mathematical design process is implemented in a couple of spreadsheets. My thinking at this time, is that we now need to incorporate specific engine parameters into the design to tune the cell to a particular engine.

We have been a bit sidetracked lately and have been working a lot on the healing aspects of Torsion fields: http://groups.yahoo.com/group/awaken_to_vibration/ but I hope to get back into engine testing soon.

Advances in 2011. In an effort to develop a device to emulate the function of a Joe cell without it's inherent stability issues, Dave Lowrance came up with the idea of a set of 3 concentrically-wound torsion field coils. In early testing it has become apparent that a field **is** being generated, as demonstrated by their effect on two test engines, even with no power being applied to the coils.

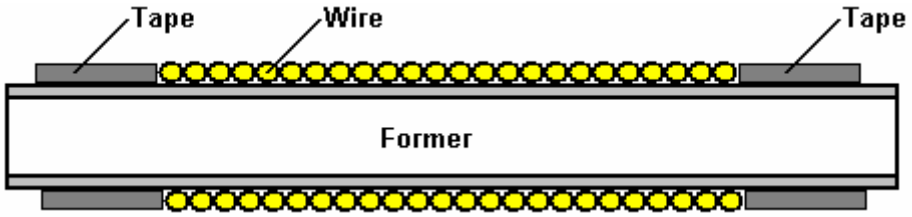
This is the very early stage of the investigation so this initial design is being released with the hope that others will wind and test similar coils and report their results to the appropriate groups, so that we can learn more about them through further experimenting on a variety of different engines.

The initial set of coils were wound on 7/8" (22 mm) diameter stainless steel tubing which happened to be to hand. The use of stainless steel is not significant and two successful replications have used half-inch (12 mm) PVC plastic pipe, as using a non-ferrous material is the main requirement.

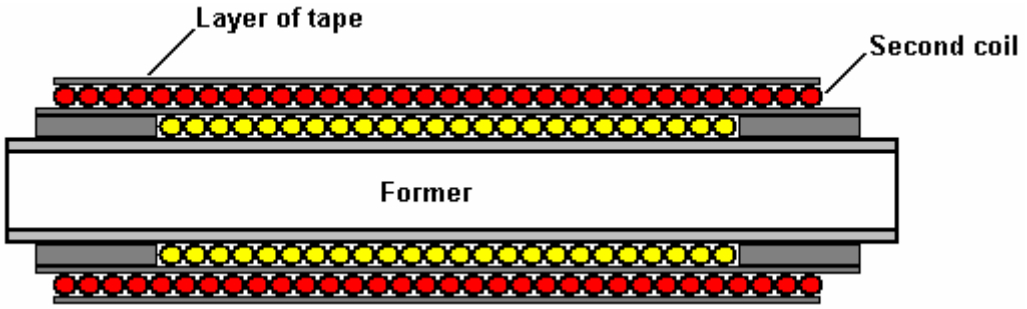
The wire diameter has an effect and while 20 gauge (0.812 mm diameter) enamelled copper wire was used for the coils shown here, coils wound with 12 gauge (2.05 mm diameter) copper wire work much better and it is now thought that the weight of copper in the winding is important.

For the first layer, a length of 311 cm is used and wound on the former in a clockwise direction. The ends of the wire are secured with tape, leaving three or four centimetres of wire exposed at each end of the coil, for connection purposes. This is the first layer wound and secured:

The wire for the second layer is cut to a length of 396 centimetres. This second coil layer will be longer than the first layer, so before winding it, it's necessary to build up the area at both ends of the first layer with tape:



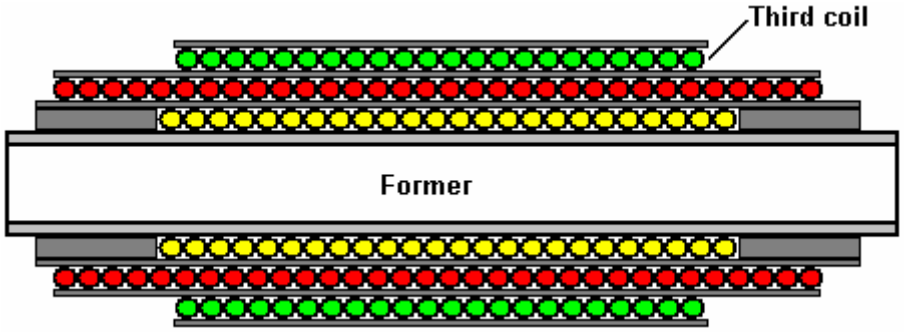
This is so that the second layer of wire will have the same diameter along it's entire length. It is probably a good idea to completely cover the first layer of wire with tape to ensure good electrical insulation.



The second wire layer is also wound in a clockwise direction:

The wire for the third layer is cut to a length of 313 centimetres. Since it will be covering less length along the former, there is no need to build up the ends of the earlier layers. So, simply cover the second winding with tape, and then wind on the third layer, but this time, the coil is wound in a counter-clockwise direction and then the entire coil is covered in tape to protect it.

To be sure that the second and third layers are centred over the earlier layers, it is a good idea to locate the centre of the wire and start winding from the middle outwards in both directions:



It has been found that one end of the centre winding is similar to the centre tube of the Joe cell, and the opposite end of the outer winding functions like the canister of a Joe cell. In theory, this can be tested by connecting a small capacitor between these two points, and checking for a low DC voltage using a digital voltmeter. Like a Joe cell, polarity is really the important issue to test for, since we do want the positive polarity end to transfer the energy, and the negative polarity end to be connected to engine ground. If the polarity is wrong, simply use the opposite ends of both coils.

In the testing the negative end was connected to chassis ground, and the positive end to a Hull-effect type oil probe already installed in each test vehicle. The oil probe is Robert Hull's contribution to this technology. He found that if you apply a torsion field to the oil, it will charge up an engine in a way similar to a Joe cell, but more consistently than a Joe cell would. There are two basic types of Hull-effect probe - the simplest is just a wire inserted down the dipstick tube. However, the preferred method is to remove the oil-pressure sensor and insert a T-fitting, then slide an insulated stainless steel rod into the high-pressure oil at that point. By using an oil probe, one can eliminate the aluminium transfer tube in favour of a length of wire.

The experimenter who wound the 20-gauge coils then wound a larger diameter set using 12-gauge wire on a 1.5-inch (38 mm) diameter former. He fitted these over the original set and connected just two wires, one end of the innermost of the six coils and the opposite end of the outermost coil. This gave about a 25% reduction in the fuel used by an old Honda Accord car with an Electronic Fuel Injection system.

Fuel-less operation has not yet been achieved, but that could just be a matter of getting the engine set up properly. Some of the issues we need to deal with are things like antifreeze, which destroys the dielectric properties of water, and inhibits it from charging up. This has never been discussed, but it is one of the key things which limited the ability of people to succeed with their cells. Oil is a similar issue. Some oils, particularly the ones with all the additives and detergents, simply won't charge up.

There still needs to be a lot of testing done. For instance, with this setup it might be better to connect one end of each coil to ground. Or possibly the coils would do better if the windings were all connected in series. This is all uncharted territory! Dave’s original concept was to use a set of these coils to replace each tube of a Joe cell.

The engine from an old Pinto car is also being used as a test bed. Attempts were made to run it completely fuel-less. It would kick repeatedly, but just wasn't quite there. It would only kick at a very specific timing setting - somewhere between 50-60 degrees before Top Dead Centre. The Pinto has antifreeze and with just water it's more likely to run fuel-less. But that should be a last-resort option, since most people do need antifreeze.

Devices such as the Joe cell tend to work really well on engines which have a carburettor because the spark timing can be adjusted quite easily. They work well on older EFI engines (probably those prior to OBD2) but they can be a real problem on the newer EFI models as they are liable to cause a fuel injection error state to

be reached almost immediately. The newer ECUs control everything so tightly that they are almost impossible to work with (which was probably a design objective of the ECU design).

The Pinto engine had not been started for over six months. No T-field devices were connected to the engine during this period, so we can assume that there was little or no residual charge on the engine. The cooling system had only water in it. The crankcase was filled with NAPA brand 30-weight oil. We fiddled with the engine to get it started. At that time the car had a little motorcycle carburettor on it, rather than the stock carburettor and the timing was set quite a bit advanced.

After just a few minutes of idling we realised that the engine was getting extremely hot with the exhaust manifold glowing red. So we shut it down. Being the optimist that I am, we went ahead and connected the coils at this time.

The next morning I took a little compass and found that it didn't point to North anywhere within about 2 feet of the car body - a very good sign! So we went ahead and started it up, and carefully monitored the head temperature with an infra-red thermometer. The temperature rose slowly to about 170 degrees F which is a little below normal. After verifying that the temperature held steady at that value, I tested with the compass again, and now it was messed up out to about 10 feet from the body. So the field strength had jumped up about 500% after starting the engine.

We then played with the carburettor and timing to get the smoothest operation at the lowest RPM at which it would idle smoothly. The RPM appeared to be well below a normal idle RPM., and when I went back and checked the timing, it was very close to 60 degrees before Top Dead Centre. At this point everything was looking so good that we tried a few attempts at fuel-less operation, but the engine died each time.

Due to pressure of other work, the car was ignored for a couple of months. When I finally got back to doing a little further testing, I found it surprisingly easy to get it started again. I didn't have to reset the timing to get it running. It actually started up with little effort, which was amazing, since the timing was still way advanced. It should be nearly impossible to start an engine with the timing set like that. The spark is just occurring at the wrong time in the cycle so it should try to push the pistons in the wrong direction.

Anyhow, it was starting to get cold here, so I decided to install some antifreeze, and that just set everything way back. It reduced the field strength by over 80%.

Since then Dave has come up with a coil-set designed to charge up antifreeze, but I was disappointed when I tried it. It did better with the antifreeze than the original set did, but we came to the conclusion that the antifreeze destroys water's diamagnetic properties to the point that the mixture is just hard to charge up. Working on this problem is the reason why I didn't release the coil info sooner. I kept hoping that we might solve this problem as well, but we didn't. However, this just might not be as big a problem as I thought, because I've heard that well-charged water just might have a significantly lower freezing point. This has not been tested yet to verify it.

An interesting side issue is the fact that the water which I drained out when adding antifreeze, showed no sign of rust. It was perfectly clear. Under normal circumstances, with no additives in the cooling system, this water should have been a horrible orange mess. It wasn't, and that has to be because of the field on the engine.

The Pinto is not roadworthy, so I have no way of knowing what kind of fuel consumption is possible with this setup or what power it might be capable of producing. At this time, I just use it to test different devices, and to try for fuel-less operation. However, if I was to achieve a consistent, repeatable fuel-less operation, it could become roadworthy very quickly, so I could do some actual road testing.

Electrets Constructed from Co-axial Cable.

There is a device which is not widely known. It is called an "electret" and I have to confess that my knowledge of them is almost zero. Essentially, an electret is a passive device which pours out electrical energy. I do not know where that electrical energy comes from. The Wikipedia encyclopaedia has some highly technical information on the subject remarking that "tunnel ionisation" is a process in which the electrons in an atom can pass through the atom's potential (voltage) barrier and escape from the atom. In an intense electric field, the potential barrier of an atom is distorted drastically and so the length of the barrier

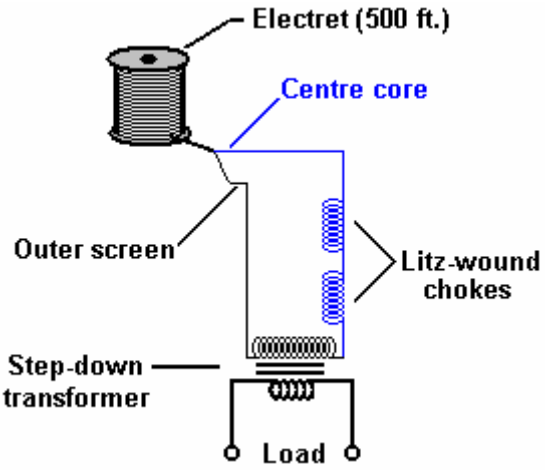
through which electrons have to pass, decreases and electrons can escape quite easily. The atoms spoken of here, might be those of a dielectric which could form an electret.

One method which has been used in the past to make an electret, has been to alter the structure of certain types of wax. A more convenient method is to use a reel of standard co-axial cable which is the sort of cable used to connect television aerials to television receivers:

An electret of that type can produce 10,000 volts at 10 milliamps. The current flow of 10 milliamps sounds trivial and of no consequence, but that is not actually the case as the power of 10 milliamps at 10,000 volts is 100 watts, so imagine a 100 watt light bulb brightly lit and not needing any power input at all to make it shine. That is actually, quite impressive.

PLEASE NOTE THAT 10,000 VOLTS WILL KILL YOU AND INVESTIGATING A DEVICE OF THIS TYPE IS NOT FOR PEOPLE WHO ARE NOT ALREADY FAMILIAR WITH WORKING SAFELY WITH VERY HIGH VOLTAGES. MEASUREMENTS MUST ONLY BE MADE WITH HIGH-VOLTAGE EQUIPMENT. LET ME STRESS AGAIN THAT I AM NOT ENCOURAGING YOU TO MAKE OR EXPERIMENT WITH ANY FORM OF HIGH VOLTAGE DEVICE AND THAT THIS INFORMATION IS FOR YOUR INTEREST ONLY.

The arrangement with a single reel of cable is:



Unfortunately, life being what it is, it has been found that when you try stepping that voltage output down to a more convenient level, there are liable to be losses which can lower the output power to just 50 watts. That sounds disappointing until you put it in perspective. This is a device which has the same output as a 50 watt solar panel in full sunlight, mounted at the optimum angle and positioned near the equator, but a home installation of such a panel gives far lower output, especially so when your home is a long way from the equator. But, note that the electret costs far less, produces that full output at any latitude and at night, while the solar panel is restricted by cloud cover, distance from the equator, needs an expensive mounting system, ideally should rotate to track the position of the sun, and only works when there is a high light level. So, the electret's fifty watts of continuous power is not an insignificant thing when you compare it to the other options available. These electrets can be stacked in parallel and an output in the kilowatts range is possible.

Let me stress that I personally have not yet made or used a co-axial cable electret, and so the information here comes from an experimenter who has done this. Also, while the information here is intended to help anyone who wishes to experiment along these lines, the fact that it is here must not be interpreted as my encouraging you personally to try to make or use an electret of this or any other type. If you choose to do that, then you do so entirely at your own risk and nobody other than yourself is liable should any mishap occur.

The following procedure has been used to convert a full reel of 1/4", type RG6/U 75-ohms, 18-AWG co-axial cable into an electret:

1. Make sure that neither end of the cable has the screen touching the central core.

2. Make an electrical connections to both the screen and the core at both ends of the cable.

3. Place the whole of the spool of cable inside an oven.

4. Heat the oven (a genuine oven and not a microwave) slowly to 350^{O}F (180^{O}C).

5. Maintain the heat until the inner plastic insulation is so soft that it can be permanently indented. This plastic must not get too soft and reach the flowing stage, nor must it get burnt or develop holes which allow arcing - if that happens, then the reel of cable is a throwaway. The objective here is to get the plastic to lose it's polarisation memory.

6. When the inner plastic sleeve has reached this level of softness, apply a steady DC voltage of about 10,000 volts to the connections already made to one end of the cable (to the screen and to the core). Although any voltage from 12V to 20,000V can be used, a 10 mA current draw can be expected when using 10,000V. Maintain this applied voltage at the high temperature for about ten minutes.

7. Turn off the heat and let the oven cool down gradually at its own rate to the 25^{O}C to 30^{O}C region, keeping the high voltage attached to one end of the cable.

8. Disconnect the DC voltage.

9. Connect the cable screen to the central core at both ends of the cable.

10. Leave the cable at room temperature for five to seven days. During this time, the polarisation of the plastic is reorganising. After this time, the electret is ready for use as a power source.

Patrick Kelly
http://www.free-energy-info.co.uk
http://www.free-energy-devices.com
http://www.free-energy-info.com

Chapter 10: Automotive Systems

Note: If you are not at all familiar with basic electronics, you might find it easier to understand this chapter if you read chapter 12 first.

There are two or three main objectives for people who create automotive devices – increasing the mpg performance and reducing the harmful emissions are the top two priorities, while running the vehicle on water alone is the aim of a few people.

The first two objectives are readily achievable, but running a vehicle on water alone is not going to happen for almost everybody. This idea is peddled by con artists who sell worthless “plans”, claiming that these will run a vehicle on water for anybody who wants to construct these simple devices. This is just not true. You are welcome to download the "HydroStar" and "HydroGen" plans free from http://www.free-energy-info.com/P61.pdf and http://www.free-energy-info.com/P62.pdf. However, most experienced people looking at these plans are convinced that they could not possibly produce enough hydrogen/oxygen gas mix to run an engine. While I have never heard of anyone, anywhere, ever getting an engine to run on these plans, the present day science of water is so inadequate, that I am not in a position to be certain that they could not work, and so I am just highly doubtful as to them being viable devices.

Just before getting on to explain the construction details of practical systems, let me put the running of an engine on water alone in its proper context. The internal combustion engine which you own has an efficiency less than 50%. This means that at least half of the energy available from the fuel which you use is wasted and does not produce any useful mechanical output power. In many cases, that percentage can be as high as 90% wasted, but let’s be generous and assume that your particular engine is especially good and manages 50% efficiency.

The main way of running an engine with water as the only fuel, involves splitting water into hydrogen and oxygen and then burning those gases to power the engine. To be self-sustaining, the splitting of the water has to be done by the electrics of the vehicle and that means that the efficiency of the water splitting has to be more than 200% efficient. That just doesn’t happen with simple systems, so please forget the notion of building some device in your garage with a couple of hours work and waving goodbye to filling stations forever – it ain’t going to happen.

Just to set the record straight, it **is** possible to appear to run a car on water alone, but the difficulty level is about the same as building a rocket capable of going into orbit, something well beyond the capabilities of most people, including me. This document does tell you how it can be done, but please understand that it calls for exceptional skills, very considerable expenditure and a great deal of patience, so for the time being, please forget about it.

What can be done quite readily and at low cost, is to construct a device which will raise the efficiency of your engine. This is done by feeding a hydrogen/oxygen gas mix (called “hydroxy” gas) into your engine along with the air which is drawn in to make the engine run. A device of this type is called a “booster” as it boosts the fuel burn, extracting a greater percentage of the fuel’s available energy. An important side effect of this improvement in the burn quality of the fuel is the fact that unburnt fuel no longer gets pushed out of the exhaust as harmful emissions.

Another effect is that the engine has greater pulling power and runs smoother. Inside your engine, carbon deposits will have built up from previous un-boosted running and these deposits get burnt away when you use a booster and that internal cleaning extends the engine life.

Some people worry about the fact that burning hydroxy gas produces water and they imagine this water causing rusting inside the engine. What they don’t realise is that the ordinary fuel used in the engine is a “hydrocarbon” which is a compound of hydrogen and carbon and that fuel actually splits up to form hydrogen which the engine burns. Actually, it is the carbon part of the hydrocarbon fuel which is the problem, producing Carbon Dioxide, Carbon Monoxide, and physical carbon deposits inside the engine. A normal fuel burn produces water anyway, but you don’t get rusting inside the engine as the temperature there is so high that any water is in the form of steam or vapour which dry out completely when the engine is switched off. Adding a small amount of hydroxy gas has no adverse effects at all.

This document describes different types of booster. Let me stress that each engine is different and it depends on how inefficient the engine is to begin with, what sort of mpg improvement is likely to be produced by a booster. Just to make sure that you understand what is involved, a booster is a simple container which holds a set of plates submerged in water which probably has an additive to make the water conduct electrical current better. A pipe from the top of the container feeds the gas into the air filter of the vehicle, via one or two simple safety devices. Adding this gas causes a major improvement in the quality of the fuel burn inside the engine and cuts harmful emission to near zero.

As a consequence of this, it is possible to reduce the amount of fossil fuel being sent to the engine, which is **not** something which should be done **if** hydroxy gas is **not** being added, as the engine is liable to overheat and some valve damage could occur. It is a completely different matter if hydroxy gas is being added. However, all recent engine designs have an Electronic Control Unit ("ECU") which controls the amount of fuel being sent to the engine. The ECU accepts input signals from an "oxygen sensor" placed in the exhaust stream, and often a second sensor after the catalytic converter to make sure that the catalytic converter has not failed.

Unfortunately, the much improved exhaust caused by the better fuel burn caused by the hydroxy gas, causes the ECU to think that the engine fuel-air mix must be too low, and so it pumps in more fuel in an effort to compensate. Ideally, this can be dealt with by adding a circuit board which adjusts the signal coming from the oxygen sensor so that it is correct for the improved fuel burn. Details of how to do this are in a companion document.

So, to recap, the only practical device which you can build yourself and use to improve automotive performance is a 'booster'. Using a booster improves the efficiency of the fuel burn inside your engine and that results in more power, better torque, smoother running and vastly improved exhaust emissions. If the ECU is not adjusted or its input signal not controlled, the mpg figures may actually get slightly lower due to unwanted excess fuel being pumped into the engine. If a control circuit is used to correct this ECU error, then mpg gains will be produced.

So, what mpg gains can be expected? The worst I have ever heard of was 8% which is very rare. The lowest likely gain is 20%. Typical gains are in the 25% to 35% bracket. Not particularly unusual is 35% to 60%, while gains up to 100% and over have been achieved but they are rare. A realistic expectation would be a 33% gain.

This chapter is divided up into the following sections:

1. Simple DC boosters, using a 12-volt electrical input.
2. Advanced DC boosters using much higher DC voltages.
3. Water-splitters which use pulsed electrical signals to change water into "hydroxy" gas.
4. Running engines without fossil fuels.
5. Other useful devices.

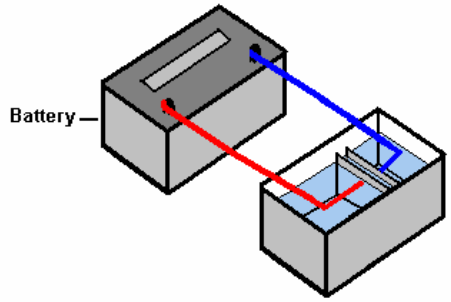
One thing which needs to be understood:

Caution: A booster is not a toy. If you make and use one of these, you do so entirely at your own risk. Neither the designer of the booster, the author of this document or the provider of the internet display are in any way liable should you suffer any loss or damage through your own actions. While it is believed to be entirely safe to make and use a properly built booster, provided that the safety instructions shown in this document are followed, it is stressed that the responsibility for doing this is yours and yours alone.

Simple DC Boosters.

It is important that you understand the basic principles of electrolysis if you are to be successful in building and operating a booster, or alternatively, buying and operating a booster. A "DC booster" operates on "Direct Current" which is the sort of electrical power delivered by a car battery.

The method is very simple in basic outline. Two metal plates are placed in water and an electric current is passed between the plates. This causes the water to break down into a mixture of hydrogen gas and oxygen gas (The two components used in the Space Shuttle). The greater the flow of current, the larger the volume of gas which will be produced. The arrangement is like this:



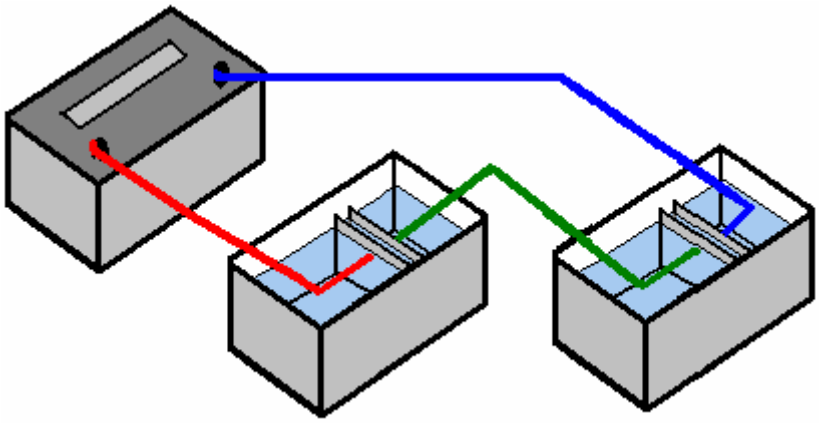
Remembering that the result of doing this is to produce fuel for the Space Shuttle, you should avoid doing this indoors and letting the gas produced by the process collect on the ceiling. There are many videos on the web where people act in a dangerous manner and perform electrolysis indoors using a container which is open at the top as shown above. Please, please don't do that as it is highly dangerous - it is not a party popper which pushes

the Space Shuttle into space! If you were to collect a cupful of hydroxy gas and light it, the resulting explosion would probably damage your hearing permanently, so don't do it under any circumstances. Just like the fact that a very useful chain saw is a dangerous device which needs to be treated with respect, so too, please understand that the very useful hydroxy gas mix contains a lot of energy and so needs to be treated with respect.

This style of electrolysis of water was investigated by the very talented and meticulous experimenter Michael Faraday. He presented his results in a very technical and scientific format which are not understood by most ordinary people. But in simple terms, he tells us that the amount of hydroxy gas produced is proportional to the current flowing through the water, so to increase the rate of gas production, you need to increase the current flow. Also, he found that the optimum voltage between the two "electrode" plates is 1.24 volts.

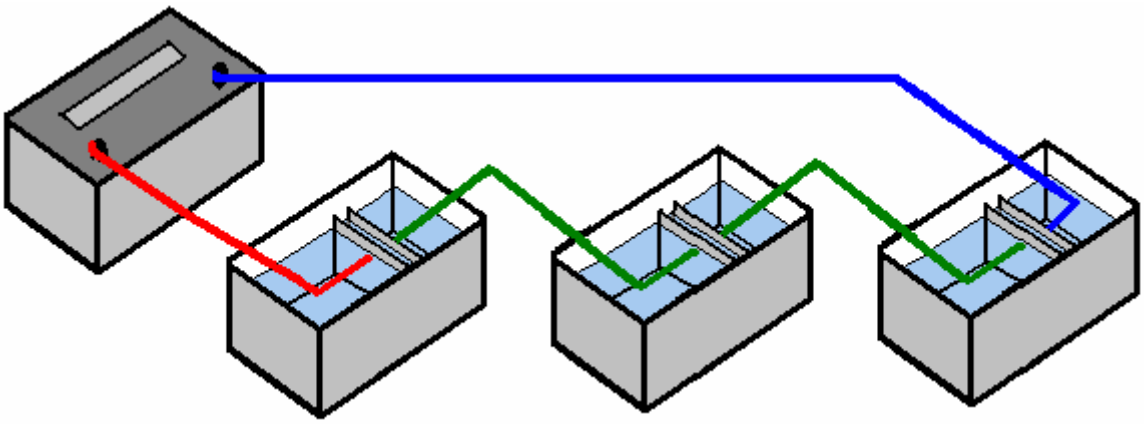
This sounds a bit technical, but it is a highly useful piece of information. In the arrangement shown above, twelve volts is being connected across two plates in water. Faraday tells us that only 1.24 volts of that twelve volts will go to make hydroxy gas and the remaining 10.76 volts will act as an electric kettle and just heat the water, eventually producing steam. As we want to make hydroxy gas and not steam, this is bad news for us. What it does tell us is that if you choose to do it that way, then only 10% of the power taken by the booster actually makes hydroxy gas and a massive 90% is wasted as heat.

We really don't want a low electrical efficiency like that. One way around the problem is to use two cells like this:



This arrangement uses our 1.24 volts twice while the twelve volts stays unchanged and so the electrical efficiency goes up to 20% and the heat loss drops to 80%. That is quite an improvement but even more important is the fact that twice as much hydroxy gas is now produced, so we have doubled the electrical efficiency **and** doubled the gas output, giving a result which is four times better than before.

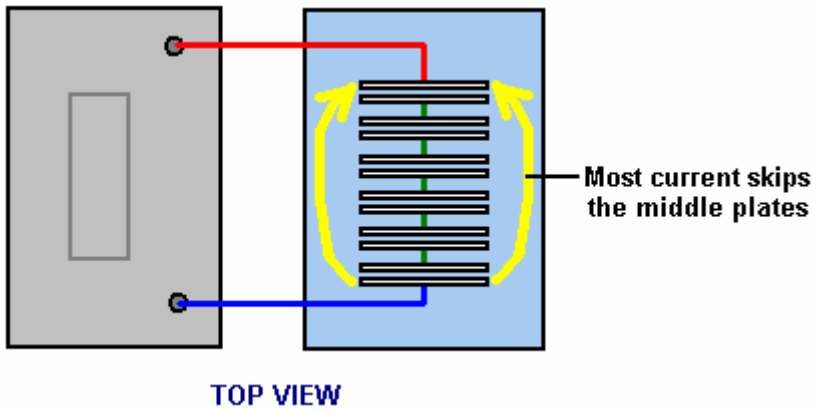
We could go one step further and use three cells like this:



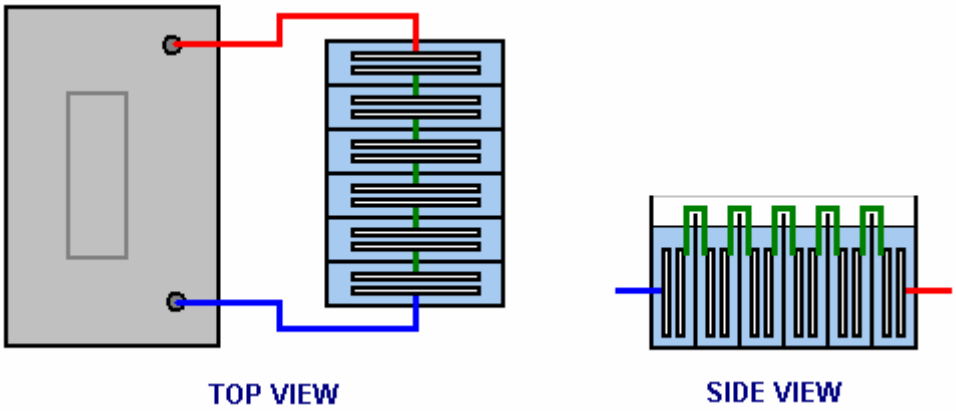
This time we are using three of our 1.24 volt sections and this gives us an electrical efficiency of 30% and three times the amount of gas, making the system nine times more effective.

This is definitely going in the right direction, so how far can we take it when using a twelve volt battery? When we use the construction materials which years of testing has shown to be particularly effective, there is a small voltage drop across the metal plates, which means that the very best voltage for each cell is about 2 volts and so with a twelve volt battery, six cells is about the best combination, and that gives us an electrical efficiency of 62% and six times as much gas, which is 37 times better than using a single cell, and the wasted electrical power drops down from 90% to 38%, which is about as good as we can get.

Of course, it would not be practical to have six boxes each as large as a car battery as we would never manage to fit them into most vehicles. Perhaps we could just put all the plates inside a single box. Unfortunately, if we do that, a good deal of the electric current would flow around the plates and not make much gas at all. A top view of this arrangement is shown here:



This is a disaster for us as now we will not get your six times the gas production or our massively reduced heating. Thankfully, there is a very simple fix for this problem, and that is to divide the box up into six watertight compartments using thin partitions like this:



This gives us back our high efficiency by blocking the current flow past the plates and forcing the current to flow through the plates, producing gas between every pair of plates.

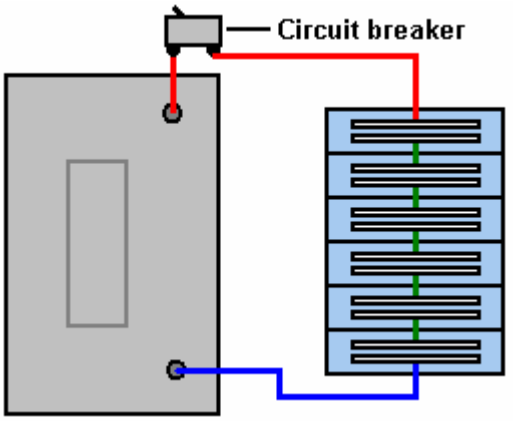
In passing, if this booster were to be powered by the electrics of a vehicle, then the voltage although called "twelve volts" will actually be almost fourteen volts when the engine is running so that the "twelve volt" battery will get charged. This would allow us to use seven cells inside our electrolyser, rather than the six cells shown above and that would give us seven times the gas volume that a single pair of plates would give. Some people prefer six cells, and others, seven cells - the choice is up to the person constructing the unit.

We have been discussing the methods of increasing the gas production and reducing the wasted energy, but please don't assume that the objective is to make large volumes of hydroxy gas. It has been found that with many vehicle engines, very good performance gains can be had with a hydroxy gas production rate of less than 1 litre per minute ("lpm"). Flow rates of as little as 0.5 to 0.7 lpm are frequently very effective. Remember, the hydroxy gas from a booster is being used as an igniter for the regular fuel used by the engine and not as an additional fuel.

The big advantage of an efficient booster design is that you can produce the wanted volume of gas using a much lower current, and so, a lesser extra load on the engine. Admittedly, there is not much additional engine load needed by a booster, but we should reduce the extra amount by intelligent design.

In the discussion above, the battery has been shown connected directly across the booster or "electrolyser". This should **never** be done as there is no protection against a short-circuit caused by a loose wire or whatever. There should be a fuse or a circuit-breaker as the first thing connected to the battery. Circuit breakers are available from any electrician's supply outlet as they are used in the "fuse box" in homes, to provide protection for each lighting

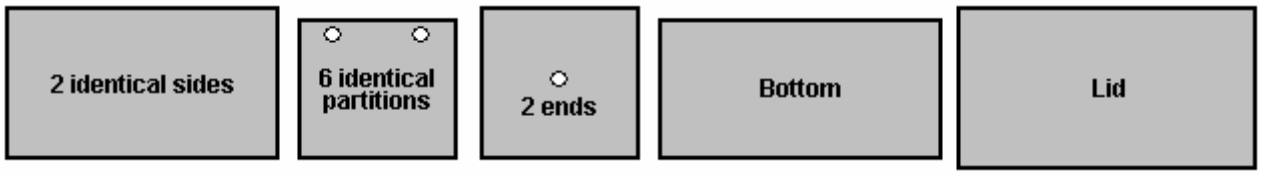
circuit and each power socket circuit. They are not expensive as they are manufactured in very large volumes. They are also available on eBay. The circuit breaker is wired like this:



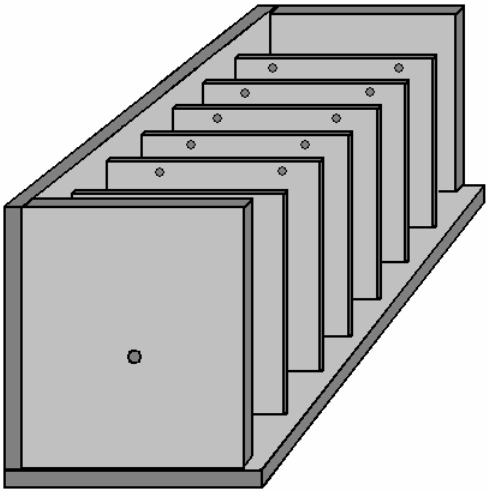
a common design (rated at 32 amps) looks like this:

Some would-be constructors feel that some aspects of the construction are too difficult for them. Here are some suggestions which might make construction more straightforward.

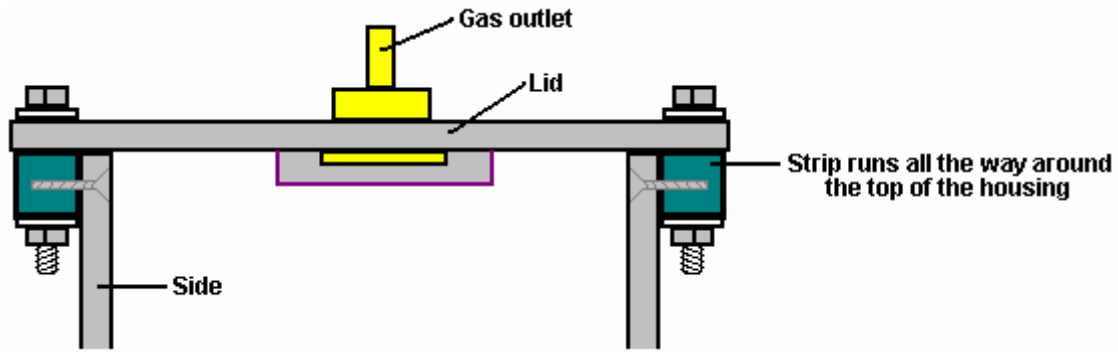
Constructing a seven-cell housing is not difficult. Pieces are cut out for two sides, one base, one lid and six absolutely identical partitions. These partitions must be exactly the same so that there is no tendency for leaks to develop. If you decide to use the bent-plate system of electrodes shown on the next few pages, then drill the bolt holes in the partitions before assembling them:



The bottom piece is the same length as the sides, and it is the width of the partitions plus twice the thickness of the material being used to build the housing. If acrylic plastic is being used for the construction, then the supplier can also provide an “adhesive” which effectively “welds” the pieces together making the different pieces appear to have been made from a single piece. The case would be assembled like this:

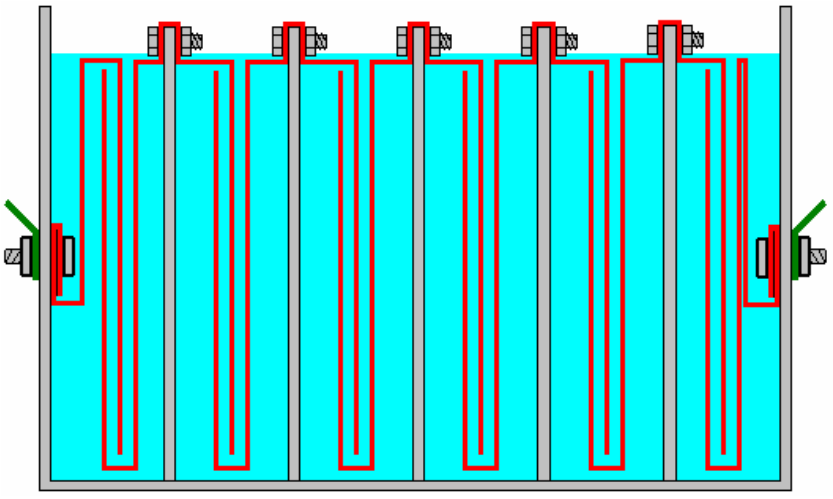


Here, the partitions are fixed in place one at a time, and finally, the second side is attached and will mate exactly as the partitions and ends are all exactly the same width. A simple construction for the lid is to glue and screw a strip all the way around the top of the unit and have the lid overlap the sides as shown here:

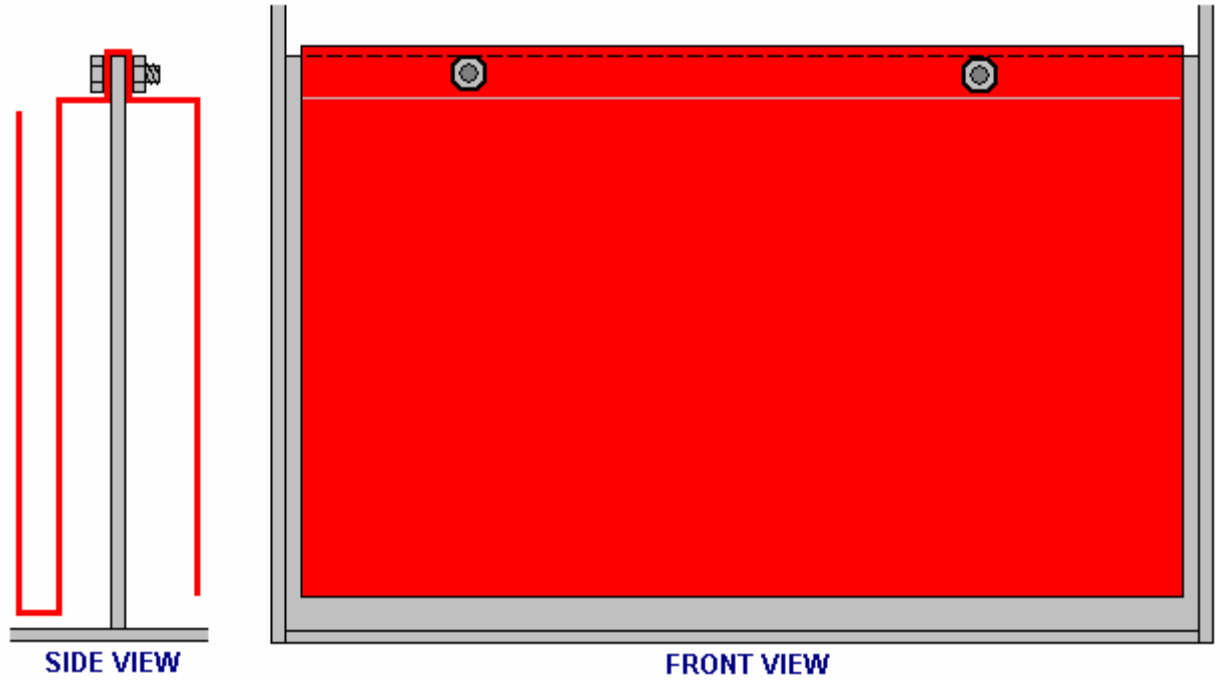


A gasket, perhaps of flexible PVC, placed between the sides and the lid would assist in making a good seal when the lid is bolted down. The gas outlet pipe is located in the centre of the lid which is a position which is not affected if the unit is tilted when the vehicle is on a steep hill.

Years of testing have shown that a really good choice of material for the electrode plates is 316-L grade stainless steel. However, it is very difficult to connect those plates electrically inside the cells as you need to use stainless steel wire to make the connections and bolted connections are really not suitable. That leaves welding the wires to the plates and welding stainless steel is not something which a beginner can do properly as it is much more difficult than welding mild steel. There is a good alternative, and that is to arrange the plate material so that no wire connections are needed:



While this six-cell design may look a little complicated to a quick glance, it is really a very simple construction. Each of the plates used in the central cells is just this shape:

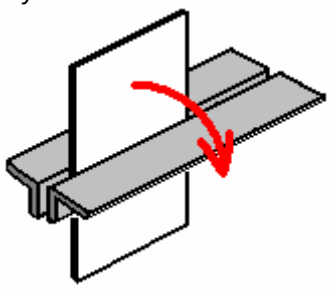


The plate shapes shown above are arranged so that there is access to the bolts from above and they can be reached by a spanner and held steady while the other nut is being tightened.

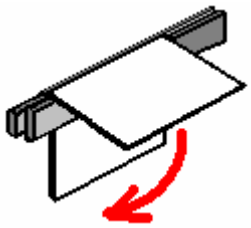
Unless you are skilled in bending plates, I suggest that you use stainless steel mesh for the plates. It works very well, can be readily cut using tin snips or any similar tool and it can be bent into shape by the home constructor using simple tools - a vice, a piece of angle iron, a small piece of mild steel sheet, a hammer, etc.

You will find a skip outside any metal fabrication shop where scrap pieces are tossed for recycling. There will be off-cuts of various sizes of angle iron and all sorts of other small sections of sheet and strip. They are in the skip mainly to get rid of them as the fabrication business gets paid almost nothing for them. You can use some of these pieces to shape your booster plates, and if you feel bad about costing the business about a penny, then by all means put them back in the skip afterwards.

If you clamp your plate between two angle irons in a vice, then careful, repeated gently tapping with a hammer close to the bend location, will produce a very clean and neat bend in the plate:

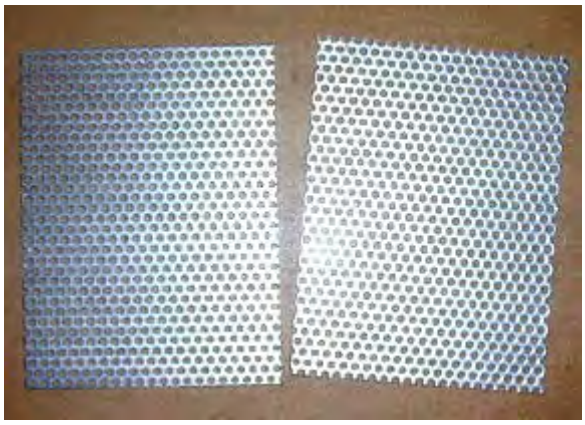


The bent sheet can then be clamped between two steel strips and a sharp U-shaped bend produced by tapping with a hammer, again, along the line of the required bend:



The thickness of the steel bar on the inside of the bend has to be the exact width of the required gap between the finished plate faces. This is not particularly difficult to arrange as 3 mm, 3.5 mm, 4 mm, 5 mm and 6 mm are common thicknesses used in steel fabrication, and they can be combined to give almost any required gap.

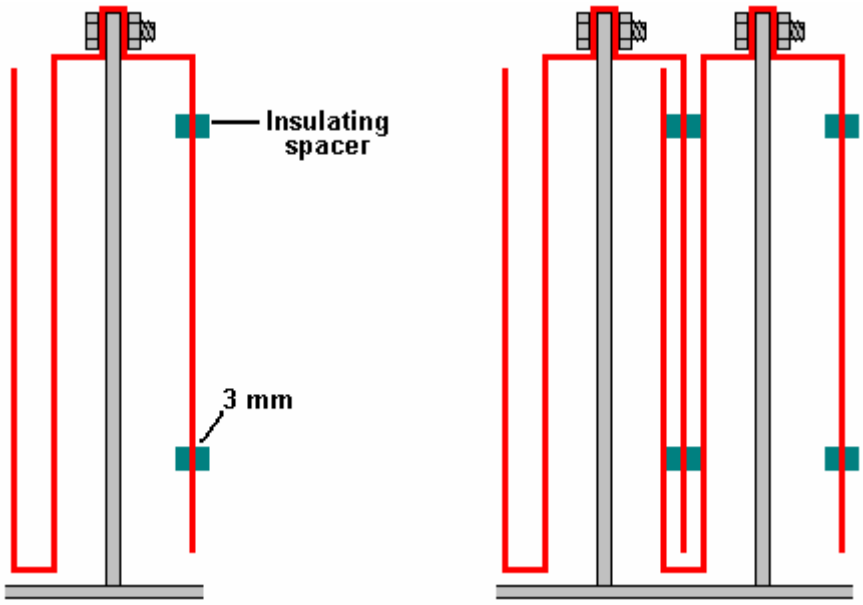
There are many varieties of stainless steel mesh. The style and thickness are not at all critical but you need to choose a type which is reasonably stiff and which will hold its shape well after it is bent. This style might be a good choice:



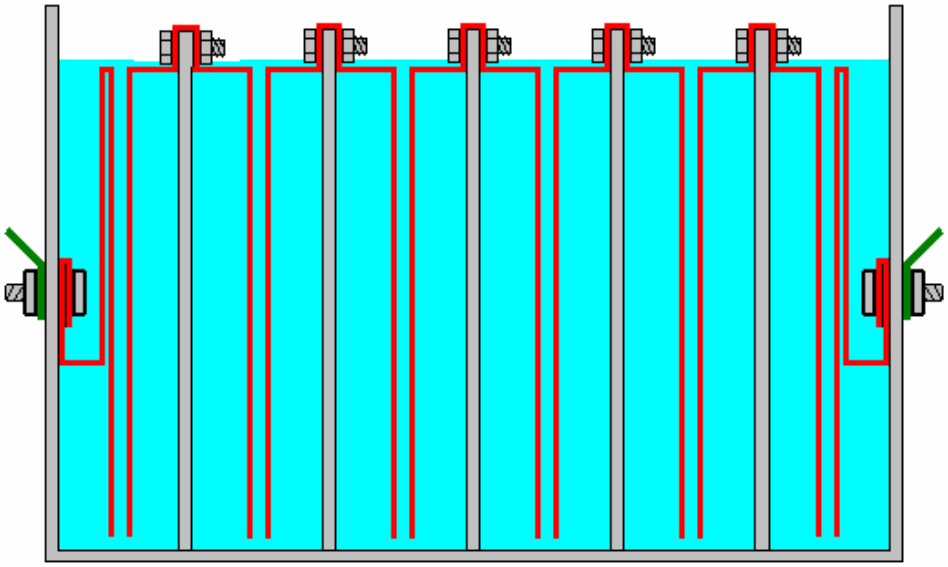
Your local steel supplier probably has some types on hand and can let you see how flexible a particular variety is. The shape shown above is for a "three plate per cell" design where there are two active plate faces. Ideally, you want two to four square inches of plate area per amp of current flowing through the cell, because that gives very long electrode life and minimum heating due to the plates.

This style of construction is reasonably easy to assemble as the two bolts which pass through the partitions and which hold the plates rigidly in place, can be accessed from above, two spanners being used to lock them tight. Lock nuts are optional. If you feel that your particular mesh might be a little too flexible or if you think that the bolts might eventually loosen, then you can attach two, or more, separator insulating pieces - plastic washers, plastic bolts, cable ties or whatever to one of the plate faces.

These will hold the plates apart even if they were to become loose. They also help to maintain the gap between the plates. This gap has to be a compromise because the closer the plates are together, the better the gas production but the more difficult it is for the bubbles to break away from the plates and float to the surface and if they don't do that, then they block off some of the plate area and prevent further gas production from that part of the plate as the electrolyte no longer touches the plate there. A popular choice of gap is 1/8 inch which is 3 mm as that is a good compromise spacing. Circular spacers would look like this:

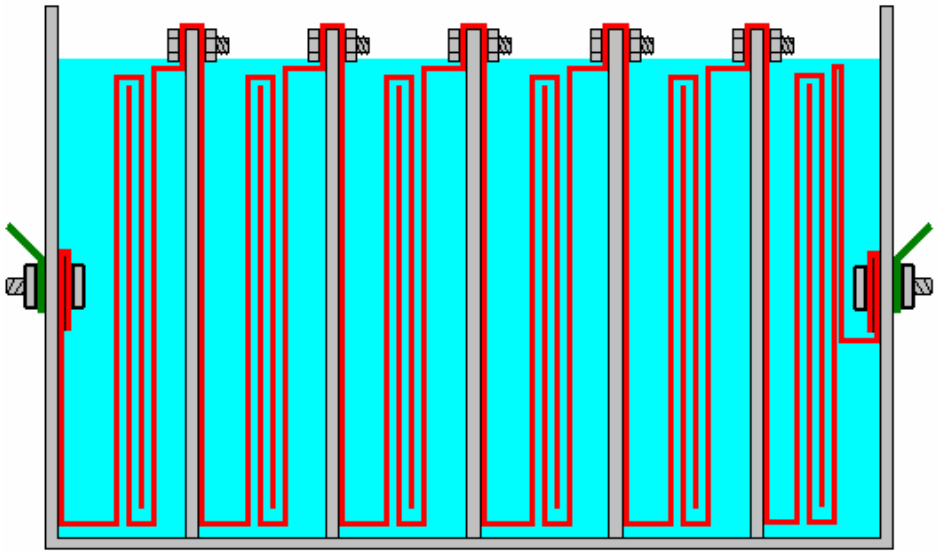


If the current is low enough, an even more simple shape which has just a single pair of active plate surfaces per cell, can be used as shown here:



Any of these designs can be 6-cell or 7-cell and the plates can be constructed without outside help. You will notice that the electrical connections at each end of the booster are submerged to make sure that a loose connection can't cause a spark and ignite the hydroxy gas in the top of the housing. There should be a gasket washer on the inside to prevent any leakage of the electrolyte past the clamping bolt.

If you want to use three active plate pairs in each cell, then the plate shape could be like this:



The electrolyte is a mix of water and an additive to allows more current to flow through the liquid. Most of the substances which people think of to use to make an electrolyte are most unsuitable, producing dangerous gasses, damaging the surfaces of the plates and giving uneven electrolysis and currents which are difficult to control. These include salt, battery acid and baking soda and I strongly recommend that you do not use any of these.

What is needed is a substance which does not get used up during electrolysis and which does not damage the plates even after years of use. There are two very suitable substances for this: sodium hydroxide, also called "lye" or "caustic soda". In the USA, this is available in Lowes stores, being sold as "Roebic 'Heavy Duty' Crystal Drain Opener". The chemical formula for it is NaOH.

One other substance which is even better is potassium hydroxide or "caustic potash" (chemical formula KOH) which can be got from soap-making supply shops found on the web. Both NaOH and KOH are very caustic materials and they need to be handled with considerable care.

Bob Boyce of the USA is one of the most experienced people in the construction and use of boosters of different designs. He has kindly shared the following information on how to stay safe when mixing and using these chemicals. He says:

These materials are highly caustic and so they need to be handled carefully and kept away from contact with skin, and even more importantly, eyes. If any splashes come in contact with you, it is very important indeed that the affected area be rinsed off immediately with large amounts of running water and if necessary, the use of vinegar which is acidic and so will neutralise the caustic liquid.

When making up a solution, you add small amounts of the hydroxide to distilled water held in a container. The container must **not** be glass as most glass is not high enough quality to be a suitable material in which to mix the electrolyte. The hydroxide itself should always be stored in a sturdy, air-tight container which is clearly labelled "DANGER! - Potassium (or Sodium) Hydroxide". Keep the container in a safe place, where it can't be reached by children, pets or people who won't take any notice of the label. If your supply of hydroxide is delivered in a strong plastic bag, then once you open the bag, you should transfer all of its contents to sturdy, air-tight, plastic storage containers, which you can open and close without any risk of spilling the contents. Hardware stores sell large plastic buckets with air tight lids that can be used for this purpose.

When working with dry hydroxide flakes or granules, wear safety goggles, rubber gloves, a long sleeved shirt, socks and long trousers. Also, don't wear your favourite clothes when handling hydroxide solution as it is not the best thing to get on clothes. It is also no harm to wear a face mask which covers your mouth and nose. If you are mixing solid hydroxide with water, **always** add the hydroxide to the water, and not the other way round, and use a plastic container for the mixing, preferably one which has twice the capacity of the finished mixture. The mixing should be done in a well-ventilated area which is not draughty as air currents can blow the dry hydroxide around.

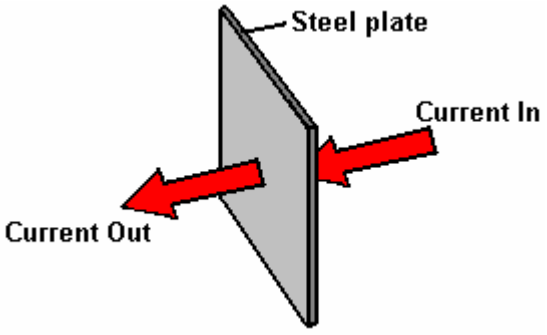
When mixing the electrolyte, **never** use warm water. The water should be cool because the chemical reaction between the water and the hydroxide generates a good deal of heat. If possible, place the mixing container in a larger container filled with cold water, as that will help to keep the temperature down, and if your mixture should "boil over" it will contain the spillage. Add only a small amount of hydroxide at a time, stirring continuously, and if you stop stirring for any reason, put the lids back on all containers.

If, in spite of all precautions, you get some hydroxide solution on your skin, wash it off with plenty of cold running water and apply some vinegar to the skin. Vinegar is acidic, and will help balance out the alkalinity of the hydroxide. You can use lemon juice if you don't have vinegar to hand - but it is always a good idea to have a bottle of vinegar handy.

The concentration of the electrolyte is a very important factor. Generally speaking, the more concentrated the electrolyte, the greater the current and the larger the volume of hydroxy gas produced. However, there are three major factors to consider:

1. The resistance to current flow through the metal electrode plates.
2. The resistance to current flow between the metal plates and the electrolyte.
3. The resistance to current flow through the electrolyte itself.

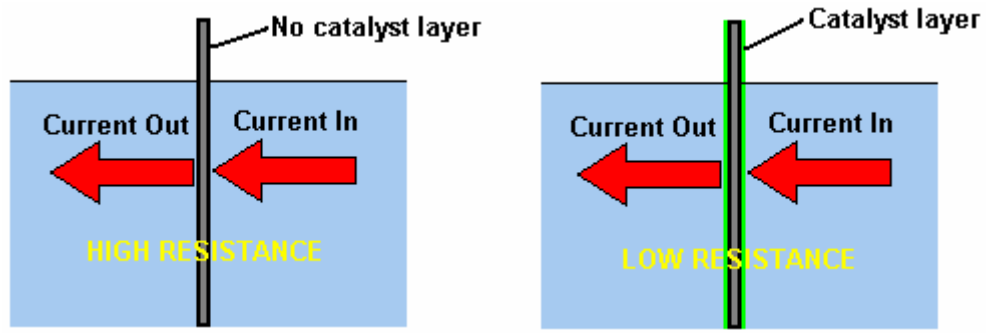
1. In a good electrolyser design like those shown above, the design itself is about as good as a DC booster can get, but understanding each of these areas of power loss is important for the best possible performance. We were taught in school that metals conduct electricity, but what was probably not mentioned was the fact that some metals such as stainless steel are quite poor conductors of electricity and that is why electrical cables are made with copper wires and not steel wires. This is how the current flow occurs with our electrolyser plates:



The fact that we have folds and bends in our plates has no significant effect on the current flow. Resistance to current flow through the metal electrode plates is something which can't be overcome easily and economically,

and so has to be accepted as an overhead. Generally speaking, the heating from this source is low and not a matter of major concern, but we provide a large amount of plate area to reduce this component of power loss as much as is practical.

2. Resistance to flow between the electrode and the electrolyte is an entirely different matter, and major improvements can be made in this area. After extensive testing, Bob Boyce discovered that a very considerable improvement can be made if a catalytic layer is developed on the active plate surface. Details of how this can be done are given later in the companion "D9.pdf" document as part of the description of Bob's electrolyser.



3. Resistance to flow through the electrolyte itself can be minimised by using the best catalyst at its optimum concentration. When using sodium hydroxide, the optimum concentration is 20% by weight. As 1 cc of water weighs one gram, one litre of water weighs one kilogram. But, if 20% (200 grams) of this kilogram is to be made up of sodium hydroxide, then the remaining water can only weigh 800 grams and so will be only 800 cc in volume. So, to make up a 20% "by weight" mix of sodium hydroxide and distilled water, the 200 grams of sodium hydroxide are added (very slowly and carefully, as explained above by Bob) to just 800 cc of cool distilled water and the volume of electrolyte produced will be about 800 cc.

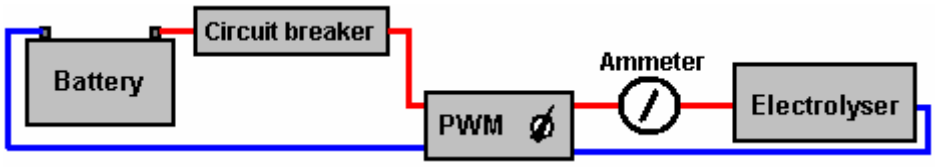
When potassium hydroxide is being used, the optimum concentration is 28% by weight and so, 280 grams of potassium hydroxide are added (very slowly and carefully, as explained above by Bob) to just 720 cc of cold distilled water. Both of these electrolytes have a freezing point well below that of water and this can be a very useful feature for people who live in places which have very cold winters.

Another factor which affects current flow through the electrolyte is the distance which the current has to flow through the electrolyte - the greater the distance, the greater the resistance. Reducing the gap between the plates to a minimum improves the efficiency. However, practical factors come into play here as bubbles need sufficient space to escape between the plates, and a good working compromise is a spacing of 3 mm. which is one eighth of an inch.

However, there is a problem with using the optimum concentration of electrolyte and that is the current flow caused by the greatly improved electrolyte is likely to be far more than we want. To deal with this we can use an electronic circuit called a "Pulse-Width Modulator" (or "PWM") circuit. These are often sold as "DC Motor Speed Controllers" and if you buy one, then pick one which can handle 30 amps of current.

A PWM circuit operates in a very simple way. It switches the current to the electrolyser On and Off many times every second. The current is controlled by how long (in any one second) the current is On, compared to how long it is Off. For example, if the On time is twice as long as the Off time (66%), then the average current flow will be much greater than if the On time were only half as long as the Off time(33%).

When using a PWM controller, it is normal to place its control knob on or near the dashboard and to mount a simple low-cost ammeter beside it so that the driver can raise or lower the current flow as is considered necessary. The arrangement is like this:



There is a more sophisticated circuit controller called a "Constant-current Circuit" and that allows you to select the current you want and the circuit then holds the current at your set value at all times. However, this type of circuit is not readily available for sale although some outlets are preparing to offer them.

Some of the most simple boosters don't use a PWM circuit because they control the current flow through the booster by making the concentration of the electrolyte very low so that the resistance to current flow through the electrolyte chokes off the current and holds it down to the desired level. This, of course, is far less efficient and the resistance in the electrolyte causes heating, which in turn, is an operational problem which needs careful handling by the user. The advantage is that the system appears to be more simple.

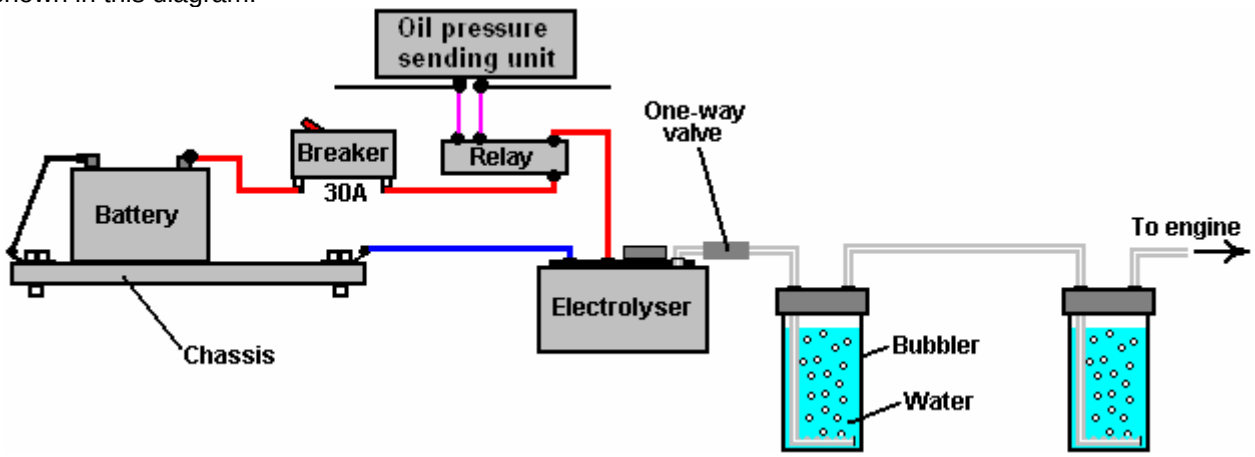
Feeding HHO Gas to an Engine.

When using a booster of any design you need to realise that hydroxy gas is highly explosive. If it wasn't, it would not be able to do it's job of improving the explosions inside your engine. Hydroxy gas needs to be treated with respect and caution. It is important to make sure that it goes into the engine and nowhere else. It is also important that it gets ignited inside the engine and nowhere else.

To make these things happen, a number of common-sense steps need to be taken. Firstly, the booster must not make hydroxy gas when the engine is not running. The best way to arrange this is to switch off the current going to the booster when the engine is not running. It is **not** sufficient to just have a manually-operated On/Off switch as it is almost certain that switching off will be forgotten one day. Instead, the electrical supply to the booster is routed through the ignition switch of the vehicle. That way, when the engine is turned off and the ignition key removed, it is certain that the booster is turned off as well.

So as not to put too much current load on the ignition switch, and to allow for the possibility of the ignition switch being on when the engine is not running, instead of wiring the booster directly to the switch, it is better to wire a standard automotive relay across the oil pressure unit and let the relay carry the booster current. The oil pressure drops when the engine stops running, and so this will also power down the booster.

An extra safety feature is to allow for the (very unlikely) possibility of an electrical short-circuit occurring in the booster or its wiring. This is done by putting a fuse or contact-breaker between the battery and the new circuitry as shown in this diagram:



If you choose to use a contact-breaker, then a light-emitting diode ("LED") with a current limiting resistor of say, 680 ohms in series with it, can be wired directly across the contacts of the circuit breaker. The LED can be mounted on the dashboard. As the contacts are normally closed, they short-circuit the LED and so no light shows. If the circuit-breaker is tripped, then the LED will light up to show that the circuit-breaker has operated. The current through the LED is so low that the electrolyser is effectively switched off when the contact breaker opens. This is not a necessary feature, merely an optional extra:

A good source for general components needed in building boosters is The Hydrogen Garage in the USA, website: http://stores.homestead.com/hydrogengarage/StoreFront.bok A very important safety item for any booster is the "bubbler" which is just a simple container with some water in it. The bubbler has the gas coming in at the bottom and bubbling up through the water. The gas collects above the water surface and is then drawn into the engine through an outlet pipe above the water surface. To prevent water being drawn into the booster when the booster is off for any length of time and the pressure inside it reduces, a one-way valve is placed in the pipe between the booster and the bubbler.

If the engine happens to backfire, then the bubbler blocks the flame from passing back through the pipe and igniting the gas being produced in the booster. A bubbler is a very simple, very cheap and very sensible thing to install. It also removes any traces of electrolyte fumes from the gas before it is drawn into the engine. In practice, it is a very good idea to have two bubblers, one close to the booster and one close to the engine. The second bubbler makes sure that every last trace of electrolyte fumes are washed out of the hydroxy gas before it enters the engine.

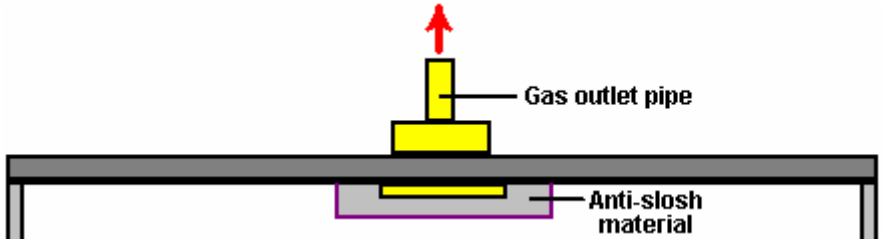
There are various ways to make a good bubbler. In general, you are aimed at having a five-inch (125 mm) depth of water through which the hydroxy gas must pass before it leaves the bubbler. It is recommended that a bubbler is built inside a strong container such as this one:

These strong containers are generally sold as water filters. They can be adapted to become bubblers without any major work being done on them. At this point, we need to consider the mechanism for moving the hydroxy gas out of the booster and into the engine.

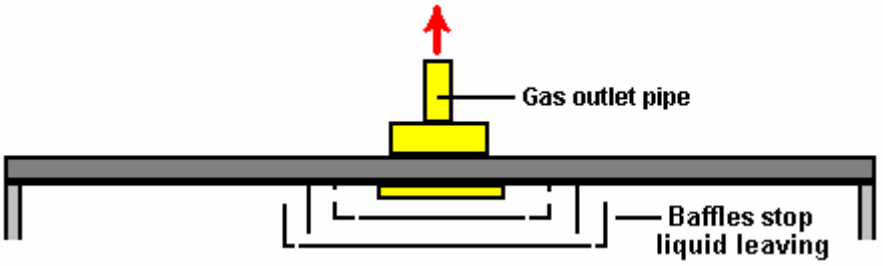
It is generally a good idea to position the gas take-off pipe in the centre of the lid so that if the booster gets tilted due to the vehicle operating on a sloped surface, then the surface level of the liquid remains unchanged underneath the gas pipe. A common mistake is to use a gas pipe which has a small diameter. If you take a length of plastic pipe of a quarter inch diameter (6 mm) and try blowing through it, you will be surprised at how difficult it is to blow through. There is no need to give your booster that problem, so I suggest that you select a gas pipe of half an inch (12 mm) or so. If in doubt as to how suitable a pipe is, then try blowing through a sample length of it. If you can blow through it without the slightest difficulty, then it is good enough for your booster.

One other thing is how to deal with splashes and the spray from bubbles bursting at the surface of the electrolyte. You want some device which will prevent any spray or splashes caused by the vehicle going over a very rough road, from entering the gas pipe and being drawn out of the booster along with the hydroxy gas.

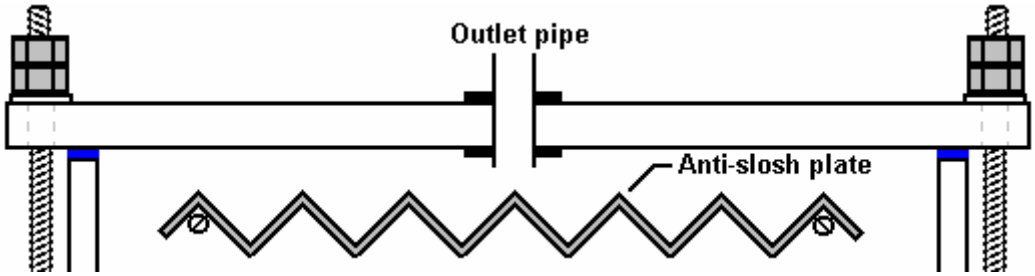
Various methods have been used and it is very much a matter of personal choice as to how you decide to deal with the issue. One method is to use a piece of suitable material across the end of the pipe. This is generally called anti-slosh material because of the job which it does. The material needs to let the gas pass freely through it but prevent any liquid getting through it. Plastic pot-scrubbers as a possible material as they have an interlocking mesh of small flat strands. The gas can flow around and through the many strands, but splashes which go in a straight line will hit the strands and drip back into the booster again. Another possible device is one or more baffles which will catch the liquid but let the gas pass freely by:



OR



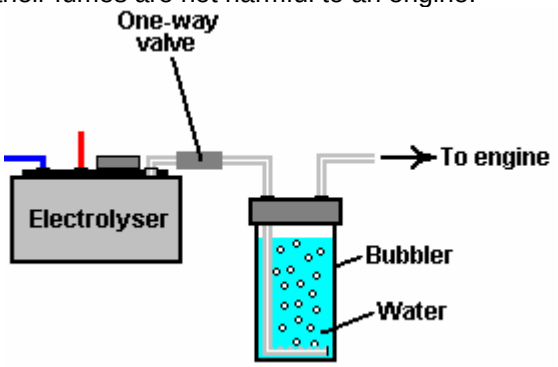
OR



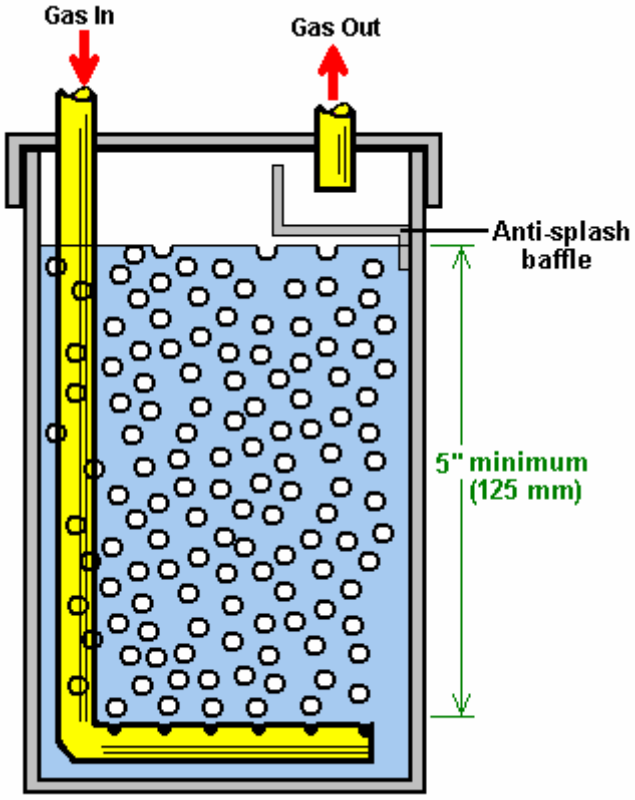
The hydroxy gas produced by a DC booster of this type contains about 30% monatomic hydrogen, which means that 30% of the hydrogen is in the form of single atoms of hydrogen and not combined hydrogen pairs of atoms. The monatomic form is about four times more energetic than the combined form and so it takes up a greater volume inside the booster housing.

If the booster is left turned off for a long period of time, then these single hydrogen atoms will eventually bump into each other and combine to form the less energetic diatomic form of the gas. As this takes up less space inside the booster, the pressure inside the booster drops and this has been known to suck water out of the bubbler back into the booster. We don't want this to happen as it dilutes our carefully measured electrolyte concentration and it can make the bubbler ineffective due to lack of water.

To deal with this, a one-way valve is put between the booster and the bubbler, positioned so that it does not allow flow back into the booster. In very cold climates, a 28% by weight potassium hydroxide electrolyte will not freeze until -40^{O} C, it is more difficult to stop the bubblers freezing. While it is possible to have equipment which unplugs and can be taken indoors overnight, an alternative is to use alcohol or paraffin (kerosene) instead of water and they generally do not freeze and their fumes are not harmful to an engine.



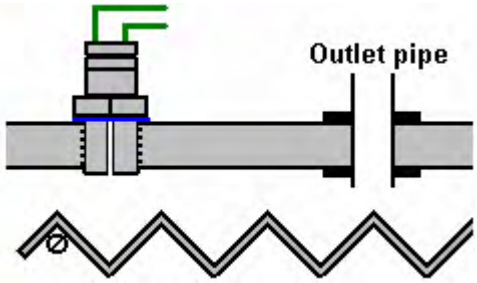
The bubbler design is not difficult. Ideally, you want a very large number of small bubbles to be formed and float upwards through the water. This is because it gives the best connection between the gas and the water and so can do a really good job of washing any traces of hydroxide vapour out of the hydroxy gas before it gets fed to the engine. Small bubbles are also better separated from each other and so there is no real chance of a flame passing through the water where large bubbles might merge together and form a column of gas as they rise to the surface.



In this good bubbler design, the pipe which feeds the hydroxy gas into the bubbler is bent into an L-shape. The end of the pipe is blocked off, and many small holes are drilled in the horizontal section of the pipe. Only a few holes are seen in this diagram, but there will be a large number in the actual construction. Like the booster itself, the gas outlet pipe needs to be protected from splashes of water caused by the vehicle going over a bump. It is very important to make sure that water is not drawn into the engine along with the gas, so anti-slosh material or one or more baffles are used to prevent this happening. So the overall protection for the gas flow is:

Where the first bubbler is close to the booster and the second one is placed close to the engine. Once in a while, the water from the first bubbler can be used to top up the water inside the booster so that any traces of hydroxide which may have reached the bubbler are returned to the booster, keeping its electrolyte concentration exactly right and making sure that the water in the bubbler is always fresh.

There is one final item which is an optional extra. Some people like to add a gas-pressure switch. If, for any reason, the pressure starts to rise - say that the outlet pipe became blocked - then the pressure switch would disconnect the electrical supply and stop the pressure rising any further:

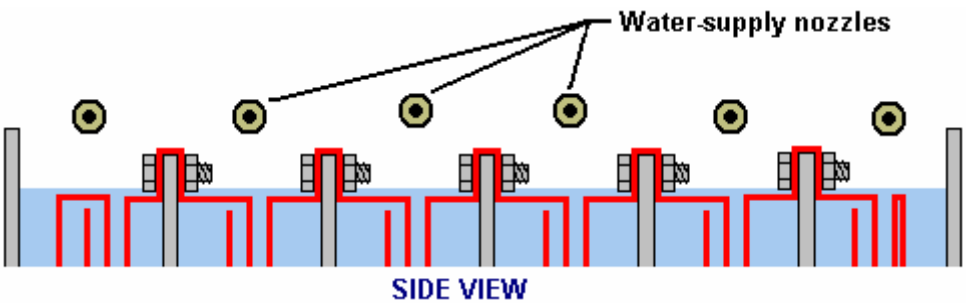


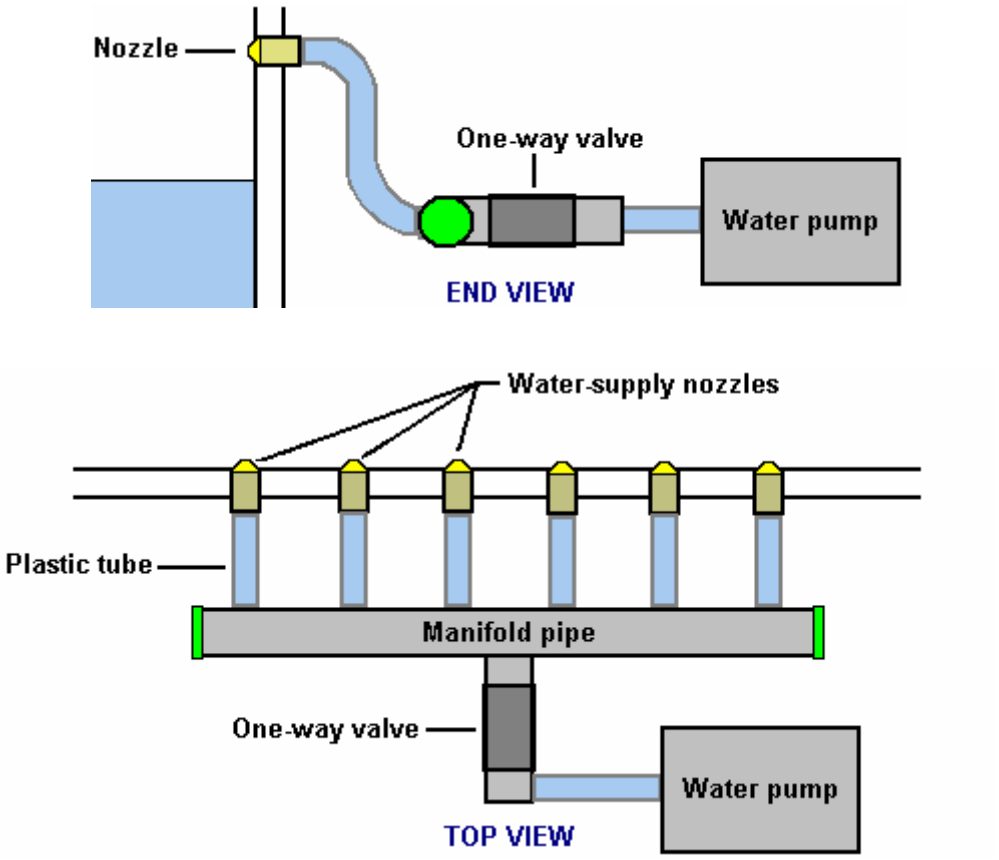
One decision which has to be made is the rate of hydroxy gas production which is the best for you. Most people seem to think that the larger the volume of hydroxy gas the better. That is not necessarily true because a very effective use of the gas is to make it act as an igniter for the engine's normal fuel and very satisfactory results have been achieved with hydroxy gas flow rates in the range of 0.4 to 0.7 litres per minute. You control the rate of gas production by controlling the current, either by the concentration of the electrolyte or by adjusting the current flow using an electronic circuit.

Each litre of water produces about 1,750 litres of hydroxy gas, so you can estimate the length of time the booster can operate on one litre of water. If, for example, your booster is producing 0.7 litres of gas per minute. Then, it will produce 1,750 litres in 1,750 / 0.7 minutes and that is 2,500 minutes or 41 hours 40 minutes. As the booster only operates when you are driving, you are looking at 41 hours of driving time and if you drive about two hours per day, it would take three weeks to use one litre of water. The internal dimensions of your booster allow you to calculate how far the electrolyte level will drop if one litre of water is taken out of it.

Generally speaking, it is normally considered that topping up the booster with water by hand every so often, is a perfectly good method of operation. The booster design described above has a good electrolyte capacity in each cell and so topping up with water should not be a major task. As tap water and well water have a good deal of dissolved solids in them, when the water is taken away by electrolysis, these solids drop out of solution and fall to the bottom of the housing, and/or coat the plates with an layer of unwanted material. For this reason, life is so much easier if distilled water is used for making electrolyte and for topping up the booster after use.

It is possible to have an automatic water supply for your booster even though that is probably over-kill for such a simple device. If you decide to do that, then you need a water supply nozzle for each of your six or seven cells. It is not necessary for the electrolyte level to be exactly the same in each cell, but you would normally have them at roughly the same height. Your automated water supply could be like this:





A point which might not be immediately obvious is that because the gas pressure inside the booster is probably about 5 pounds per square inch ("psi"), once the water pump stops pumping, it is possible for the gas pressure to push out the remaining water in the inlet pipes and escape through the body of the pump. To prevent this, an ordinary one-way valve is put in the water supply pipe to prevent flow back towards the pump.

Up to now, the hydroxy gas feed to the engine has just been indicated in a vague way in spite of the connection point being important. With most engines, the hydroxy gas should be fed into the air filter where it mixes well and is fully dispersed inside the air being drawn into the engine. You sometimes see diagrams which show the connection point being close to the engine intake manifold. This is not a good idea because the lowered pressure there causes reduced pressure inside the booster which in turn produces more unwanted hot water vapour, so stick with feeding the gas into the air filter. If there is a supercharger on the engine, then feed the hydroxy gas into the low-pressure side of the supercharger.

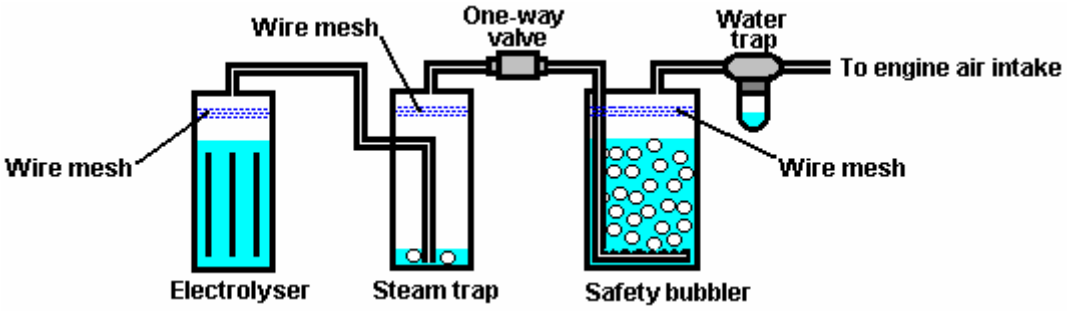
The “Smack’s” Booster.

The style of booster described above has the advantages of high electrical efficiency, easy construction, very few specialist parts and a large electrolyte volume per cell. There are many other very successful booster designs which have very different forms of construction. One of these is the "Smack's Booster" where electrical cover plates are clamped together and placed inside a length of plastic pipe:

The advantages of this design are the very simple construction, compact size, reasonable performance and the fact that you can buy one ready-made if you want to. You can download a copy of the construction details free from http://www.free-energy-info.com/Smack.pdf The electrical efficiency of this design is lowered a bit because only a single body of electrolyte is used and so current can bypass the plates. The overall performance is a respectable 1.3 lpm for 20 amps, though you may wish to lower the current and settle for about half that rate of hydroxy gas production. The construction of a 5 lpm version is at http://www.youtube.com/watch?v=cqjn3mup1So

The "Hotsabi" Booster.

Another design which is very easy to build is the "HotSabi" booster, which is a single threaded rod inside a length of plastic pipe with a stainless steel inner lining. It has the lowest possible electrical efficiency, being just a single cell with the full vehicle voltage connected directly across it, but in spite of that, it's performance in actual on the road use has been remarkable, with a reported 50% improvement on a 5 litre capacity engine. This excellent performance is probably due to the design having a steam trap which removes the hot water vapour produced by the excessive heating caused by having only a single cell with so much voltage across it (remember, 90% of the power supplied to this booster design goes in heating the electrolyte).



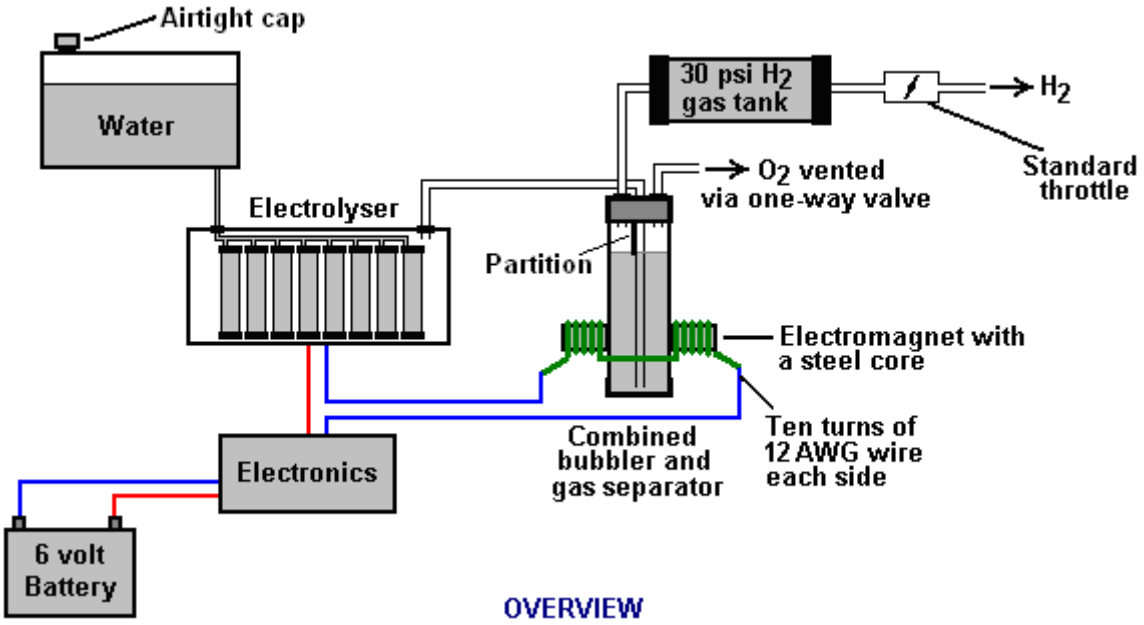
As the designer of this booster has freely shared his design, the free construction plans can be downloaded from http://www.free-energy-devices.com/Hotsabi.pdf

The "Zach West" Electrolyser.

Zach West of the USA has produced a motorcycle electrolyser. Zach's 250 cc motorcycle can run on the output of his electrolyser and Zach estimates the output as being 17 litres per minute of HHO gas, which seems to me to be far too high for the current flow. This is not a COP>1 system as the output from the electrical system of the motorcycle is very limited, and so the battery will slowly run down as time goes by. However, Zach's design of electrolyser is interesting, both for it's simplicity and it's high rate of gas output. The increased gas volume which would be produced if this design were adapted for, and driven by, a 12-volt input could be very useful, especially if combined with David Quirey's system which allows the resulting modified gas to operate in unmodified engines, as shown later on in this chapter.

The method which Zach uses is somewhat unusual as he manages to bleed off and discard most of the oxygen produced. This means that the remaining gas is mainly hydrogen which is far less explosive than HHO which is already in the perfect proportions for combination back into water and so is highly reactive. Instead, the resulting gas can be compressed reasonably well, and Zach compresses it to 30 psi (pounds per square inch) in a storage container. This helps with acceleration from stationary at traffic lights.

Zach uses a simple, modular style of construction where a series of coiled electrode pairs are each placed inside an individual length of plastic pipe. This is a design which is neither difficult nor particularly expensive to build. In overall broad outline, Zach's electrolyser is fed water from a water tank to keep it topped up. The electrolyser box contains several pairs of electrodes which split the water into hydrogen and oxygen when fed with pulsed electrical current generated by the electronics, which is powered by the electrical system of the motorcycle. The gas produced by the electrolyser is fed to a dual-purpose bubbler, which prevents any accidental igniting of the gases from travelling back to the electrolyser and in addition, removes most of the oxygen from the gas by acting as a gas "separator". The arrangement is like this:



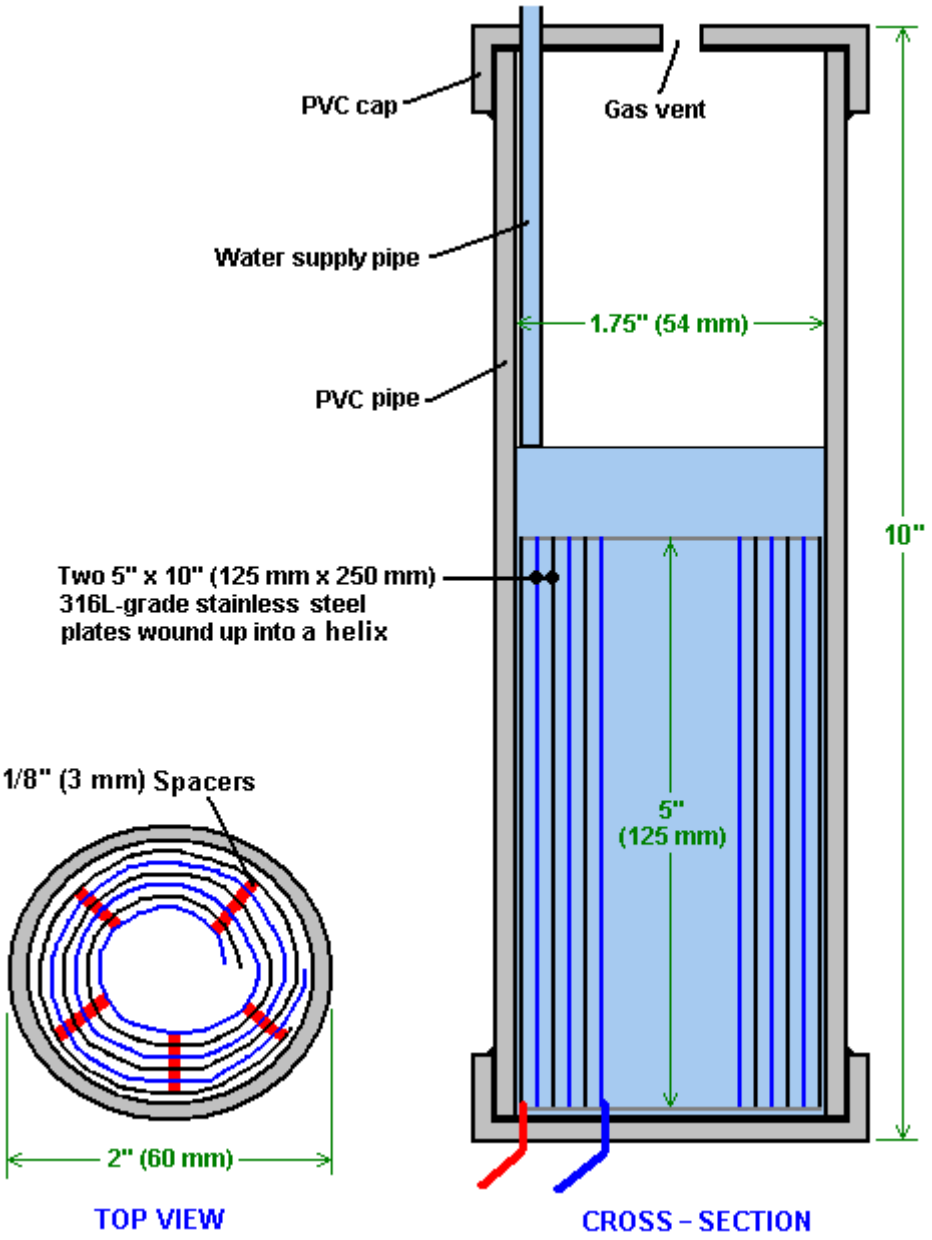
OVERVIEW

The hydrogen gas output from the electrolyser is not fed directly to the engine but instead it goes to a pressure tank which is allowed to build up to thirty pounds per square inch before the engine is started. The majority of the oxygen produced by the electrolysis is vented away through a 30 psi one-way valve which is included to keep the pressure inside the bubbler (and the electrolyser) at the 30 psi level. That pressure is excessive for a high-performance electrolyser which produces HHO which is highly charged electrically and so will explode spontaneously when compressed, due to it's own electrical charge. In this simple DC electrolyser, the HHO gas is mixed with quite an amount of water vapour which dilutes it and allows some compression.

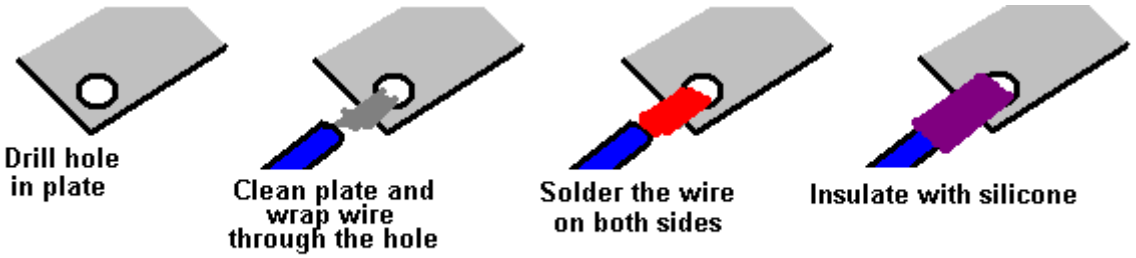
The water supply system operates by having an air-tight supply tank positioned at a higher level than the electrolyser. A small diameter (1/4" or 6 mm) plastic tube coming from the supply tank feeds through the top of the electrolyser and straight down, terminating at exactly the electrolyte surface level wanted in each of the electrolyser tubes. When the electrolysis lowers the electrolyte level below the bottom of the pipe, bubbles of gas pass up the tube allowing some water to flow from the tank to raise the electrolyte surface level back to it's design position. This is a very neat passive system needing no moving parts, electrical supply or electronics but yet one which accurately controls the electrolyte level. One essential point to understand is that the water tank needs to be rigid so that it will not flex and the filler cap needs to be air-tight to prevent the entire water supply discharging into the electrolyser. Another point to remember when topping up the water tank is that the tank contains a mix of air and HHO gas above the water surface and not just plain air, and that gas mix is at 30 psi pressure.

Now, to cover the design in more detail. This 6-volt electrolyser contains eight pairs of electrodes. These electrode pairs are coiled around in "Swiss-roll" style and inserted into a length of 2 inch (50 mm) diameter plastic pipe, ten inches (250 mm) tall. The electrodes are each made from a 10 inch (250 mm) by 5 inch (125 mm) of 316L-grade stainless steel shimstock which is easy to cut and work. Shimstock is available from a local steel supplier or metal fabrication company and is just a sheet of very thin metal.

Each electrode is cleaned carefully, and wearing rubber gloves, cross-scored using coarse sandpaper in order to produce a very large number of microscopic mountain peaks on the surface of the metal. This increases the surface area and provides a surface which makes it easier for gas bubbles to break away and rise to the surface. The electrodes are rinsed off with clean water and then coiled round, using spacers to maintain the necessary inter-plate gap, to form the required shape which is then inserted into a length of plastic pipe as shown here:



As the springy metal pushes outwards in an attempt to straighten up again, spacers are used to keep the electrodes evenly separated along their whole length by inserting 1/8" thick vertical spacer strips. The connections to the plates are made by drilling a hole in the corner of the plate and inserting the wire several times through the hole, twisting it back around itself and making a wire-to-wire solder joint on both sides of the steel. The joint is then insulated with silicone or any other suitable material. It is, of course, essential that the joint does not short-circuit to the other electrode even though that electrode is very close by.

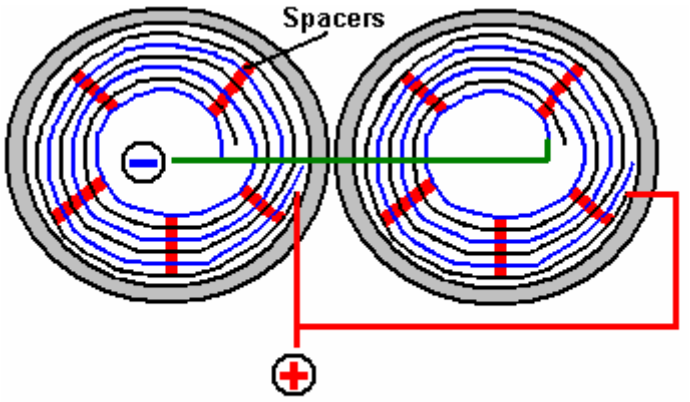


CONNECTING TO THE PLATES

It is always difficult to make a good electrical connection to stainless steel plates if space is restricted as it is here. In this instance, the electrical wire is wrapped tightly through a drilled hole and then soldered and insulated. The soldering is only on the wire as solder will not attach to stainless steel.

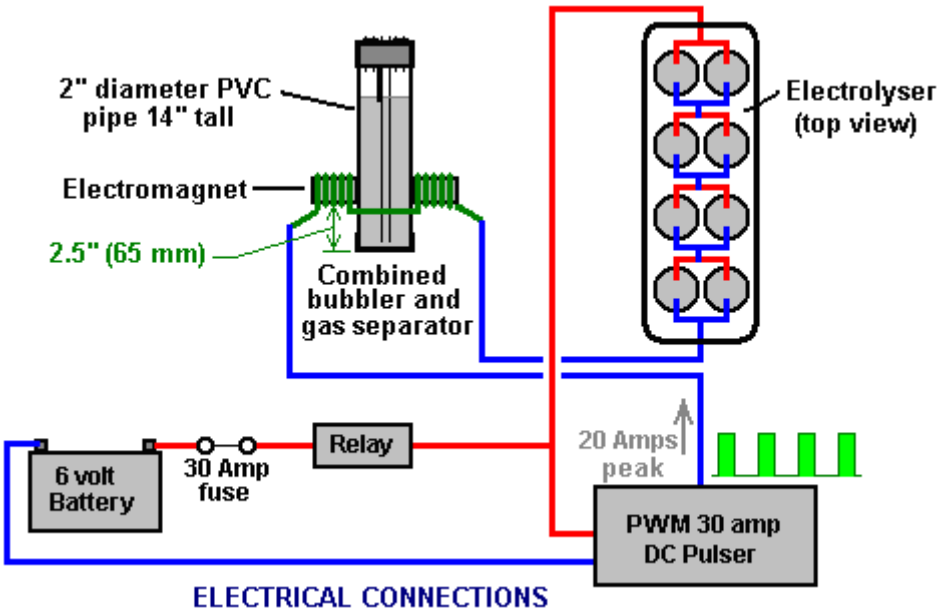
An unusual feature of this design is that each of the electrode pairs is effectively a separate electrolyser in its own right as it is capped top and bottom, and effectively physically isolated from the other electrodes. The water feed comes through the top cap which has a hole drilled in it to allow the gas to escape. The electrical wires (#12 AWG or swg 14) are fed through the base and sealed against leakage of electrolyte. Each of these units has some electrolyte stored above it, so there is no chance of any part of the electrode surface not being able to generate gas. There is also a large amount of freeboard to contain splashes and sloshing without any being able to escape from the container. The end caps are standard PVC caps available from the supplier of the PVC piping, as is the PVC glue used to seal them to the pipe.

Eight of these electrodes are placed in a simple electrolyser case and connected together in pairs as shown here:



CELLS ARE WIRED IN PAIRS (Top View)

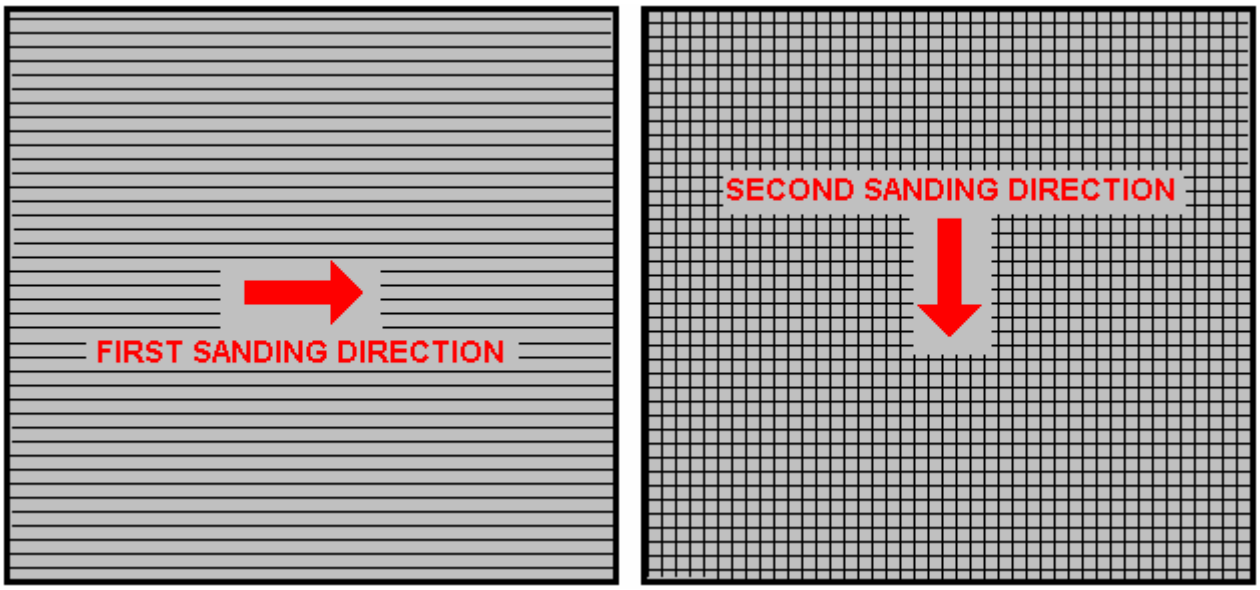
Pairs of pipe-enclosed electrode spirals are then connected in a chain inside the electrolyser as shown here:



Many years of experimentation and testing have shown that 316L-grade stainless steel is the most suitable material for electrodes, but surprisingly, stainless steel is not highly electrically conductive as you would expect. Each electrode causes a voltage drop of nearly half a volt, and so careful surface preparation, cleansing and conditioning are needed to get top performance from the electrodes. This process is described in detail by the very experienced Bob Boyce who says:

The preparation of the plates is one of the most important steps in producing an electrolyser which works well. This is a long task, but it is vital that it is not skimped or hurried in any way. Surprisingly, brand new shiny stainless steel is not particularly suitable for use in an electrolyser and it needs to receive careful treatment and preparation before it will produce the expected level of gas output.

The first step is to treat both surfaces of every plate to encourage gas bubbles to break away from the surface of the plate. This could be done by grit blasting, but if that method is chosen, great care must be taken that the grit used does not contaminate the plates. Stainless steel is not cheap and if you get grit blasting wrong, then the plates will be useless as far as electrolysis is concerned. A safe method is to score the plate surface with coarse sandpaper. This is done in two different directions to produce a cross-hatch pattern. This produces microscopic sharp peaks and valleys on the surface of the plate and those sharp points and ridges are ideal for helping bubbles to form and break free of the plate.



When doing hand sanding the sandpaper is drawn across the plates in **one** direction only and not backwards and forwards, as the backwards stroke always destroys the perfectly good ridges created on the forward stroke. Also,

you only need two strokes in one direction before turning the plate through ninety degrees and completing the sanding of that face of the plate with just two more strokes (again, with **no backstroke**).

Always wear rubber gloves when handling the plates to avoid getting finger marks on the plates. Wearing these gloves is very important as the plates must be kept as clean and as grease-free as possible, ready for the next stages of their preparation. Any particles created by the sanding process should now be washed off the plates. This can be done with clean tap water (not city water though, due to all the chlorine and other chemicals added), but only use distilled water for the final rinse.

While Potassium hydroxide (KOH) and Sodium hydroxide (NaOH) are the very best electrolytes, they need to be treated with care. The handling for each is the same:

Always store it in a sturdy air-tight container which is clearly labelled "DANGER! - Potassium Hydroxide". Keep the container in a safe place, where it can't be reached by children, pets or people who won't take any notice of the label. If your supply of KOH is delivered in a strong plastic bag, then once you open the bag, you should transfer all its contents to sturdy, air-tight, plastic storage containers, which you can open and close without risking spilling the contents. Hardware stores sell large plastic buckets with air tight lids that can be used for this purpose.

When working with dry KOH flakes or granules, wear safety goggles, rubber gloves, a long sleeved shirt, socks and long trousers. Also, don't wear your favourite clothes when handling KOH solution as it is not the best thing to get on clothes. It is also no harm to wear a face mask which covers your mouth and nose. If you are mixing solid KOH with water, always add the KOH to the water, and not the other way round, and use a plastic container for the mixing, preferably one which has double the capacity of the finished mixture. The mixing should be done in a well-ventilated area which is not draughty as air currents can blow the dry KOH around.

When mixing the electrolyte, **never** use warm water. The water should be cool because the chemical reaction between the water and the KOH generates a good deal of heat. If possible, place the mixing container in a larger container filled with cold water, as that will help to keep the temperature down, and if your mixture should "boil over" it will contain the spillage. Add only a small amount of KOH at a time, stirring continuously, and if you stop stirring for any reason, put the lids back on all containers.

If, in spite of all precautions, you get some KOH solution on your skin, wash it off with plenty of running cold water and apply some vinegar to the skin. Vinegar is acidic, and will help balance out the alkalinity of the KOH. You can use lemon juice if you don't have vinegar to hand - but it is always recommended to keep a bottle of vinegar handy.

Plate cleansing is **always** done with NaOH. Prepare a 5% to 10% (by weight) NaOH solution and let it cool down. A 5% solution 'by weight' is 50 grams of NaOH in 950 cc of water. A 10% solution 'by weight' is 100 grams of NaOH in 900 cc of water. As mentioned before, never handle the plates with your bare hands, but always use clean rubber gloves.

A voltage is now applied across the whole set of plates by attaching the leads to the outermost two plates. This voltage should be at least 2 volts per cell, but it should not exceed 2.5 volts per cell. Maintain this voltage across the set of plates for several hours at a time. The current is likely to be 4 amps or more. As this process continues, the boiling action will loosen particles from the pores and surfaces of the metal. This process produces HHO gas, so it is very important that the gas is not allowed to collect anywhere indoors (such as on ceilings).

After several hours, disconnect the electrical supply and pour the electrolyte solution into a container. Rinse out the cells thoroughly with distilled water. Filter the dilute NaOH solution through paper towels or coffee filters to remove the particles. Pour the dilute solution back into the cells and repeat this cleaning process. You may have to repeat the electrolysis and rinsing process many times before the plates stop putting out particles into the solution. If you wish, you can use a new NaOH solution each time you cleanse, but please understand that you can go through a lot of solution just in this cleaning stage if you choose to do it that way. When cleansing is finished (typically 3 days of cleansing), do a final rinse with clean distilled water. It is very important that during cleansing, during conditioning and during use, that the polarity of the electrical power is always the same. In other words, don't swap the battery connections over as that destroys all the preparation work and requires the cleansing and conditioning processes to be carried out all over again.

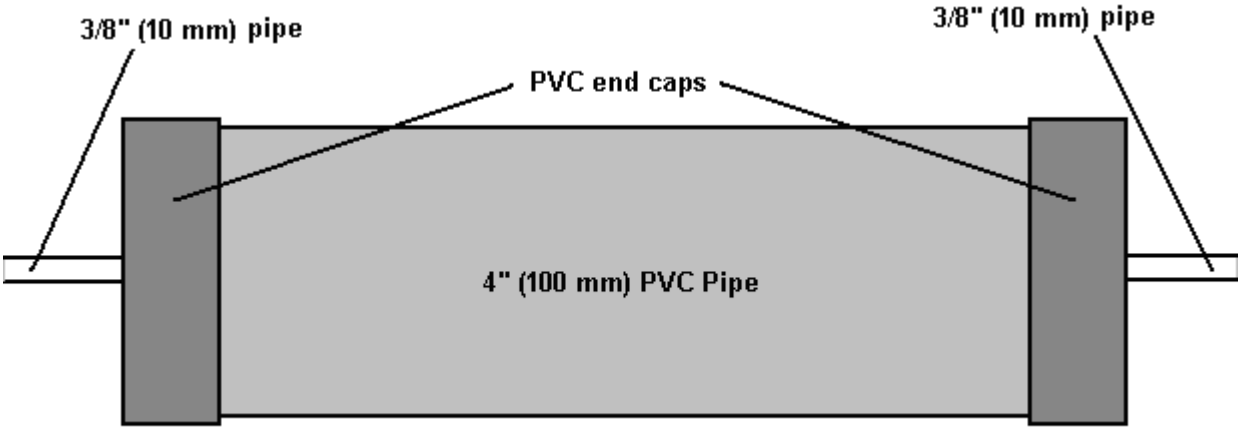
Using the same concentration of solution as in cleansing, fill the cells with dilute solution. Apply about 2 volts per cell and allow the unit to run. Remember that very good ventilation is essential during this process. As water is consumed, the levels will drop. Once the cells stabilise, monitor the current draw. If the current draw is fairly stable, continue with this conditioning phase continuously for two to three days, adding just enough distilled water to replace what is consumed. If the solution changes colour or develops a layer of crud on the surface of the

electrolyte, then the electrodes need more cleansing stages. After two to three days of run time, pour out the dilute KOH solution and rinse out the cells thoroughly with distilled water.

The construction which Zach has used is very sensible, utilising readily available, low-cost PVC piping. The spiral electrodes are inside 2" diameter pipe and Zach says that the bubbler is also 2" diameter PVC pipe. I seriously doubt that a two-inch diameter bubbler could handle a flow as high as 17 lpm which is a substantial amount. Also. You want the bubbles in the bubbler to be small in order that the gas comes into good contact with the water. Consequently, using more than one bubbler where the diagram shows just one, would be sensible.

At this time, Zach only uses one bubbler, but a second one is highly desirable, located between the storage tank and the engine and positioned as close to the engine as possible. This extra bubbler does two things, most importantly, it prevents the gas in the storage tank being ignited by a backfire caused by a valve sticking slightly open and secondly, it removes every last trace of potassium hydroxide fumes from the gas, protecting the life of the engine. This is a big gain for such a simple addition.

The gas storage tank is also made from PVC pipe, this time, 4 inch (100 mm) diameter, 14 inches (350 mm) long with standard end caps fixed in place with PVC glue as shown below. This is a compact and effective arrangement well suited for use on a motorcycle. The majority of this extra equipment can be mounted in bike panniers, which is a neat arrangement.



GAS PRESSURE RESERVOIR

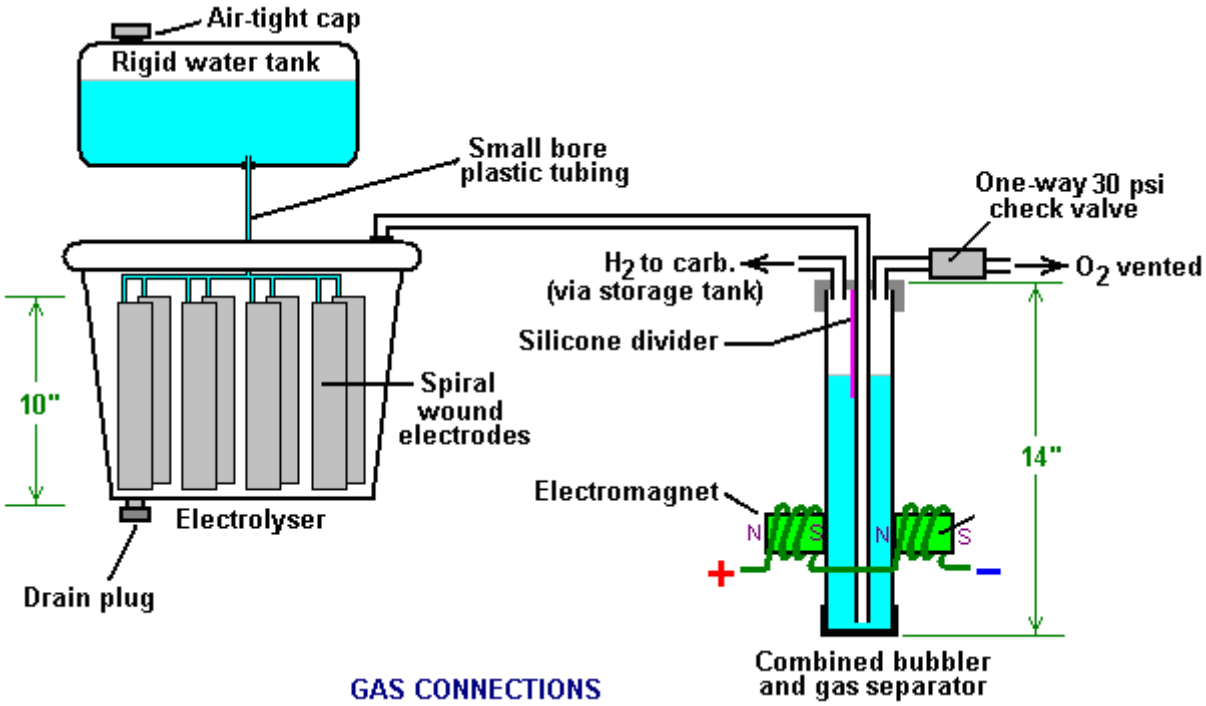
The electric drive to the electrolyser is from a Pulse Width Modulator ("DC Motor speed controller") which was bought from the Hydrogen Garage as Zach is in America. That particular PWM board is no longer available, so especially for those people in Europe the choice might be rmcybernetics.com, although there are many suppliers and the module should not be expensive.

As this unit was rated at just 15 Amps maximum, Zach added another 15 Amp rated FET transistor in parallel to the output stage to raise the current capacity to 30 Amps. A fuse protects against accidental short circuits and a relay is used to control when the electrolyser is to be producing gas. The connecting wire is #12 AWG (swg 14) which has a maximum continuous current capacity of just under ten amps, so although the current peaks may be twenty amps, the average current is much lower than that.

Two electromagnets outside the bubbler, positioned 2.5 inches (65 mm) above the base, are connected as part of the electrical supply to the electrolyser, and these cause most of the oxygen and hydrogen bubbles to separate and exit the bubbler through different pipes. There is a divider across the bubbler to assist in keeping the gases from mixing again above the water surface. The bubbler also washes most of the potassium hydroxide fumes out of the gas as the bubbles rise to the surface, protecting the engine as these fumes have a very destructive effect on engines.

The objective with any hydroxy system is to have the minimum amount of gas between the bubbler and the engine in order to block the ignition of the gas in the unlikely event of a backfire. In this system, the gas storage tank contains a very large amount of gas, though admittedly it is not full HHO gas thanks to the electromagnet separation system, but nevertheless, it would be most advisable to have a second bubbler between the gas storage tank and the engine, positioned as close to the engine as possible. HHO gas produces a very high-speed shock-wave when it is ignited so the bubbler needs to be of strong construction to withstand this. No pop-off bubbler cap or blow-out device acts fast enough to contain a HHO shock-wave, so make the bubbler housing strong enough to withstand the pressure wave.

Zach's electrolyser arrangement is like this:



GAS CONNECTIONS

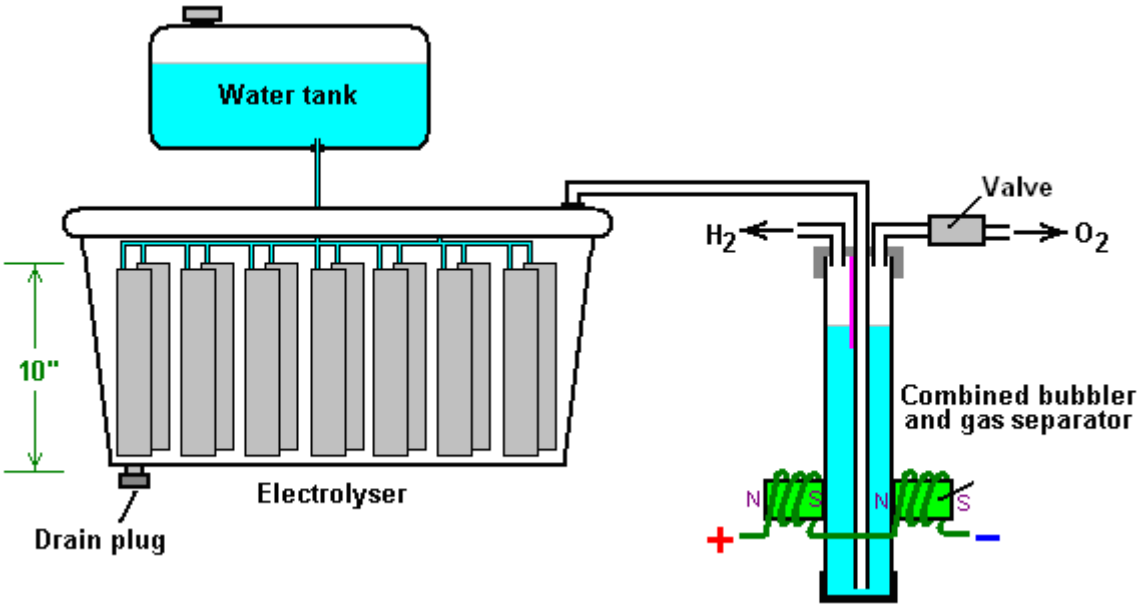
It must be realised that the water tank, electrolyser, bubbler/separator and hydrogen holding tank, all operate at thirty pounds per square inch. This means that each of these containers must be robust enough to withstand that pressure quite easily. It also means that the 30 psi one-way check valve on the oxygen venting pipe is an essential part of the design as well as being a safety feature. As a bubble of gas from the electrolyser escapes into the water tank every time a drop of water feeds to the electrolyser, the contents of the water tank above the water surface becomes a stronger and stronger mix of air and HHO. Consequently, it soon becomes an explosive mixture. It is common for static electricity to build up on a tank of this nature, so it will be very important to earth both the tank and it's cap before removing the cap to top up the tank with more water.

The electrolyser has a potassium hydroxide (KOH) solution in it. The electrolysis process produces a mixture of hydrogen, oxygen, dissolved gases (air) and potassium hydroxide fumes. When the system is being used, the water in the bubbler washes out most of the potassium hydroxide fumes, and in doing so, it gradually becomes a dilute electrolyte itself. Potassium hydroxide is a true catalyst and while it promotes the electrolysis process, it

does not get used up during electrolysis. The only loss is to the bubbler. Standard practice is to pour the contents of the bubbler into the electrolyser from time to time, filling the bubbler again with fresh water. Potassium hydroxide has been found to be the most effective catalyst for electrolysis but it has a very bad effect on the engine if it is allowed to enter it. The first bubbler is very effective in removing the potassium hydroxide fumes, but many people prefer to take the scrubbing process a step further by placing a second bubbler in the line, in this instance, between the hydrogen pressure tank and the engine. With two bubblers, absolutely no potassium hydroxide fumes reach the engine.

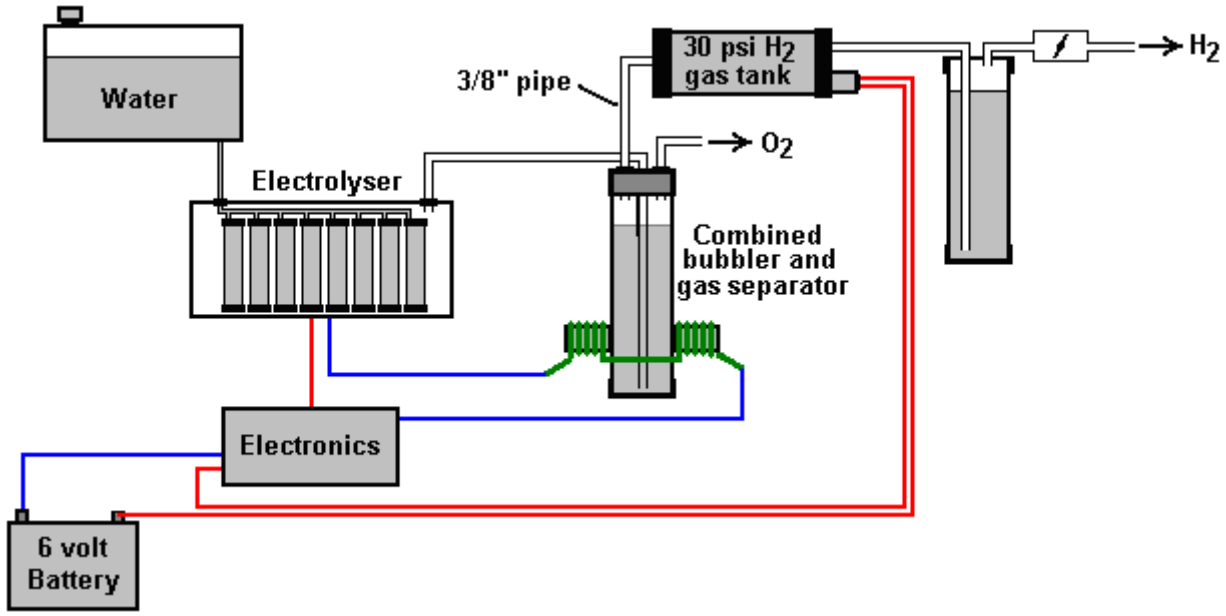
When running with HHO gas as the only fuel, it is essential to adjust the timing of the spark so that it occurs after Top Dead Centre. The timing on this bike is now set at 8 degrees after TDC. However, if David Quirey's style of bubbling the HHO through a liquid such as acetone, then no timing alterations would be needed.

This electrolyser is designed to run off the nominal six volts of a motorcycle electrics (about 7.3 volts with the engine running), but increasing the number of tubes, each containing electrode coils, would convert the design to a 12V system and then the electrolyser housing would probably be like this:



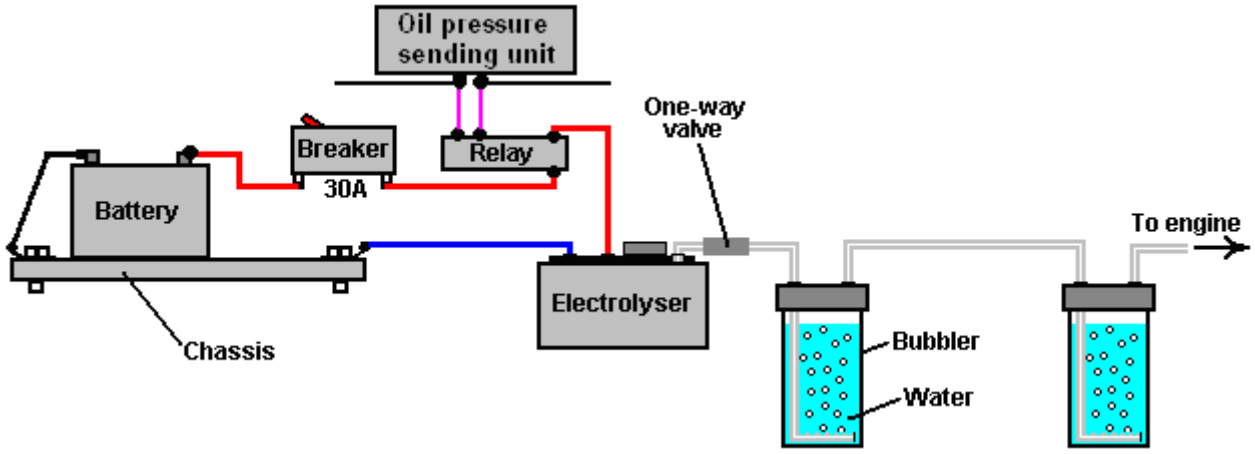
It is possible that seven sets of three or four spirals wired in parallel would be used for larger engines with their 13.8 volt electric systems. Zach uses the very simple method of allowing excess gas to be vented via the oxygen valve if gas production exceeds the requirements of the engine. When operating on a twelve volt system it might be more convenient to use a standard pressure switch which opens an electrical connection when the gas pressure rises above the value for that switch:

The pressure switch just mounts on one of the end caps of the pressure tank and the switch electrical connection is placed between the relay and the electrolyser. If the gas pressure reaches it's maximum value of 30 psi. then the switch opens, stopping electrolysis until the pressure drops again:



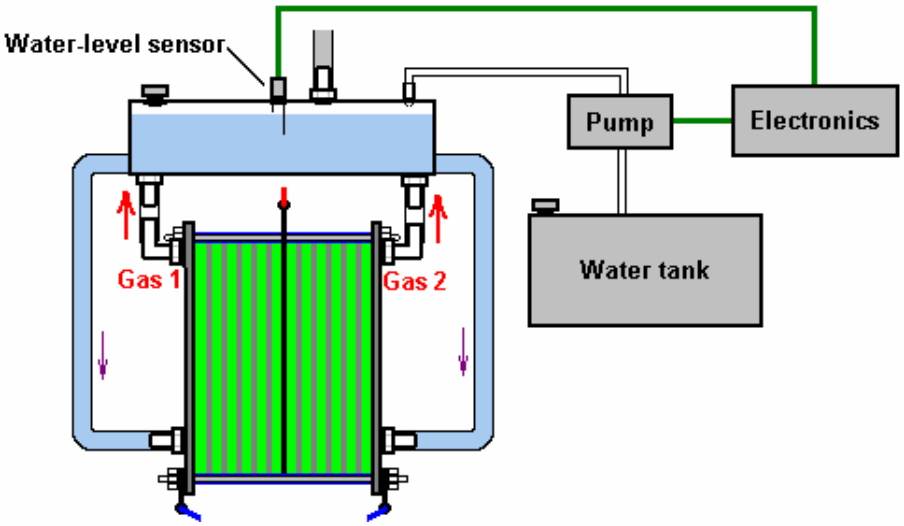
Caution: This electrolyser is not a toy. If you make and use one of these, you do so entirely at your own risk. Neither the designer of the electrolyser, the author of this document or the provider of the internet display are in any way liable should you suffer any loss or damage through your own actions. While it is believed to be entirely safe to make and use an electrolyser of this design, provided that the safety instructions shown below are followed, it is stressed that the responsibility is yours and yours alone.

An electrolyser should not be considered as an isolated device. You need to remember that both electrical and gas safety devices are an essential part of any such installation. The electrical safety devices are a circuit-breaker (as used by any electrician when wiring a house) to protect against accidental short-circuits, and a relay to make sure that the booster does not operate when the engine is not running:



The "DuPlex" Booster designed by Bill Williams.

A fully-submerged design from Bill Williams in the USA is another different style of booster:



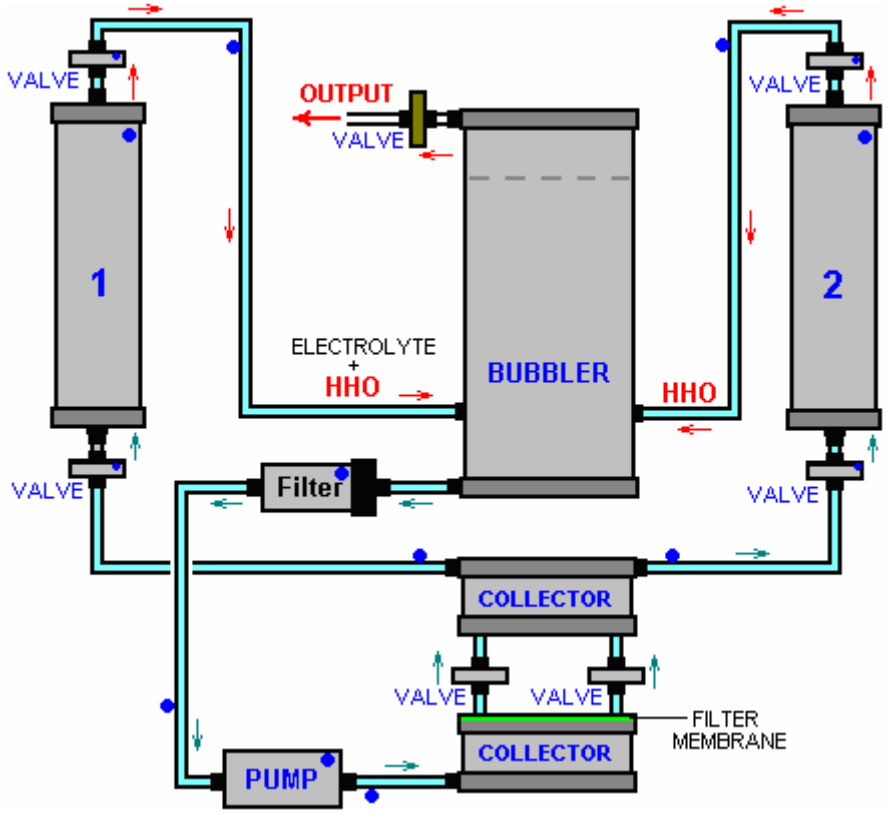
The construction details for this booster design, can be downloaded free from the web using the link: http://www.free-energy-devices.com/DuPlex.pdf

There are many other designs, including those with concentric pipes, each having its own advantages and disadvantages, some being commercially available as ready-made devices, and there are links to these boosters on the web sites mentioned above and a general booster forum at http://tech.groups.yahoo.com/group/watercar/ and another at http://tech.groups.yahoo.com/group/HHO/ where people will answer any queries.

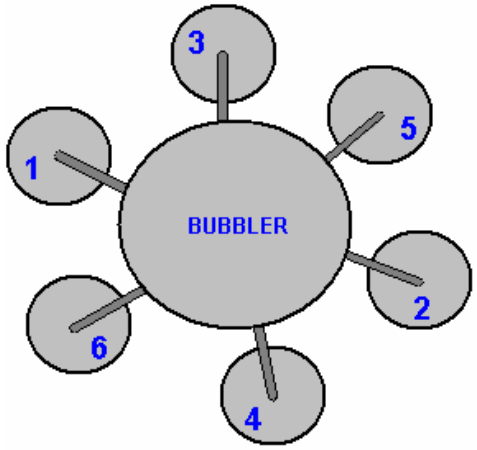
One problem with the use of boosters is that if the hydroxy gas volume is higher than it needs to be, the vehicle's Electronic Control Unit ("ECU") is liable to detect the improved fuel burn and start pumping in excess fuel to offset the improved conditions. How to deal with this situation is covered in the free document which can be downloaded from http://www.free-energy-devices.com/D17.pdf

The "Hogg" Electrolyser from Selwyn Harris.

An interesting design is the Hogg electrolyser as described by Selwyn Harris of Australia. The Hogg cell has two stainless steel mesh electrodes coiled around each other. This gives a large electrode surface area in a very compact container. In this version of the design there are six identical cells which feed into a large bubbler. For clarity, only two of the six electrolyser cells are shown here:



The items marked with a blue dot form just one of three identical sets. That is, three filters feed out from the bottom of the bubbler, the flow through them being caused by three separate pumps and the two electrolyser cells marked “1” and “2” along with their associated pipes and four one-way valves, are replicated to give electrolysis cells “3” to “6” which are not shown in this diagram. Those three identical sets are connected to the central bubbler, spaced around it evenly at 120-degree positions horizontally as shown in rough outline here:



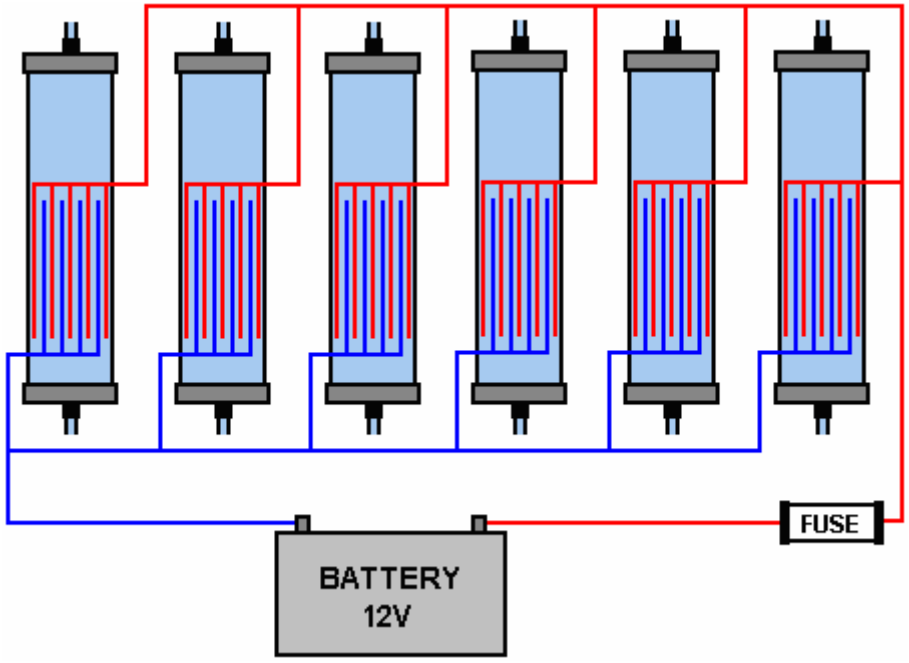
The water is circulated through the set of cells using three small pumps and there are two water collectors built on to the bottom of the bubbler. Also, as the ‘electrolyte’ used is rainwater, and each electrolysis cell is completely full of electrolyte as this is an ‘electrolyte circulation’ style of electrolyser.

Each of the three pumps has it’s own filter to trap any particles coming from the cells as experience has shown that the water can contain a considerable amount of material. The filters are standard irrigation in-line filters made from transparent plastic filled for three quarters of their length with fine plastic sponge material.

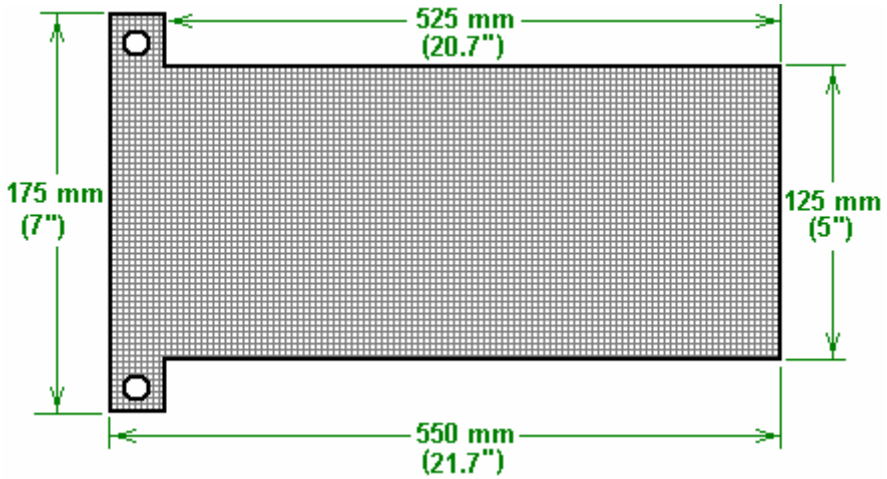
A key feature of the cell design is the use of two powerful neodymium magnets per cell. These act directly on the water and that causes a major increase in the gas production rate. The magnets have their North poles facing towards each other.

The two mesh electrodes are made from stainless steel wire of 0.32 mm diameter and woven to give 2 mm holes between the wires and a 0.65 mm overall sheet thickness. These dimensions are important as other mesh sizes and styles do not give so good a performance. The electrodes are wider at one end to form a connection tab

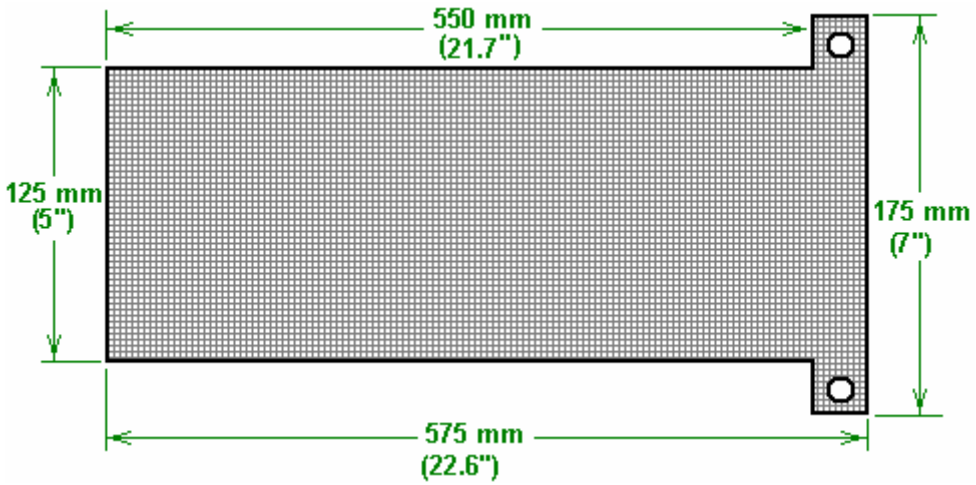
which allows easy electrical connection to each electrode and they are then wired in parallel so that each cell gets 12 volts across it as shown here:



The two mesh electrodes are cut like this:

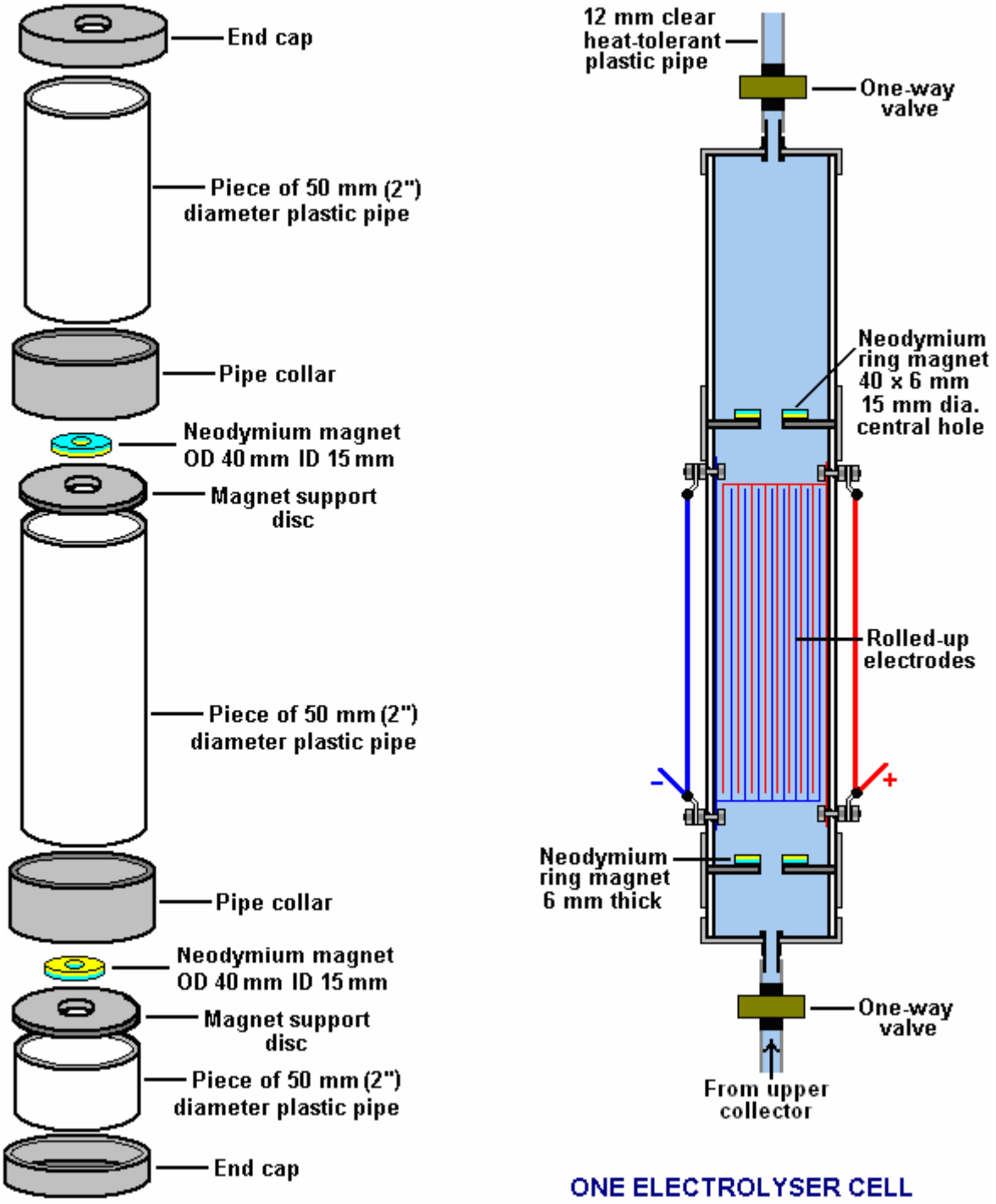


and:

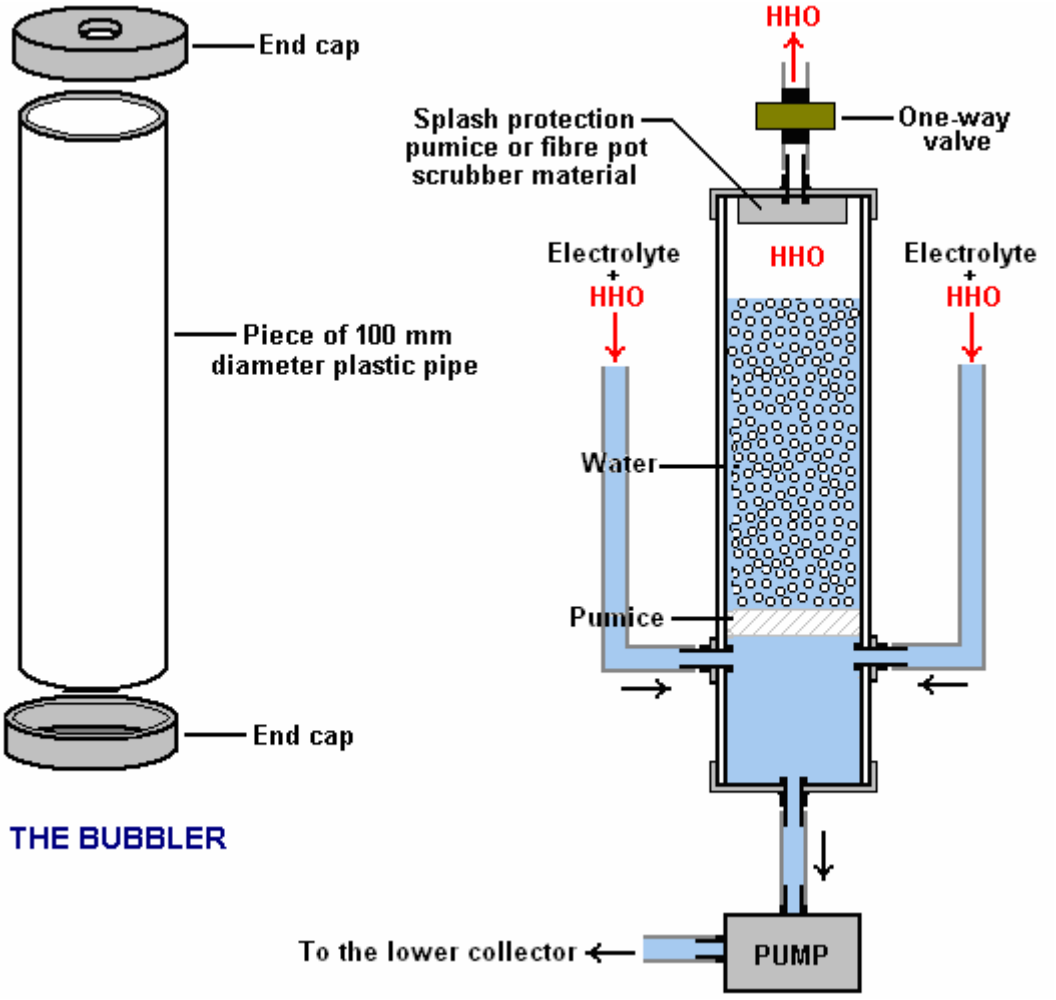


The mesh material looks like this:

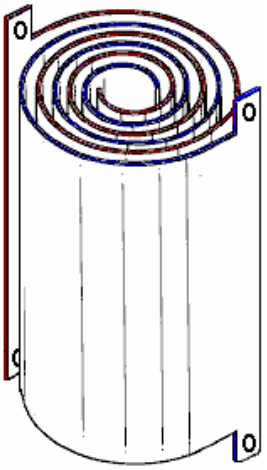
The six electrolyser cells and the single large bubbler are constructed using standard plastic plumbing materials:



The connecting pipes are a clear plastic, heat-tolerant type of 12 mm (0.5") diameter. The bubbler is also constructed from plastic pipe fittings:



As some people have difficulty in visualising the way in which the electrodes are combined, this simplified concept sketch may be helpful:



The two electrodes are kept separated by the use of small diameter fibre washers held in place between them at strategic places using super glue. The mesh itself is then treated by being immersed in citric acid in order to make it work well with rainwater.

There are three of these pairs of electrolyser cells, each pair being connected to the upper collector reservoir. The water being pumped out of each cell is passed through one of the three filters before entering the collector reservoir which feeds the small pump which keeps the water circulating which in turn keeps removing any particles which have got into the rainwater.

When used with rainwater from a barrel, this electrolyser is said to draw just 1.4 amps per cell, giving a total input of about 115 watts when run on a 12-volt electrical supply. While rainwater is supposedly pure, the reality is that it seldom is and it's ability to carry a current varies dramatically from place to place and even more widely from country to country. If you decide to build this electrolyser and find that you do not get anything like 1.4 amps

flowing through any one cell, then you may well have to add a small amount of electrolyte to the water in order to get the current flowing. That is, assuming that you want a current flow of 1.4 amps per cell on 12 volts. The output of this electrolyser is said to be capable of running a small electrical generator but that has not been confirmed at this time.

Advanced DC Boosters.

All of the practical construction details on electrical safety, gas safety, engine connections, type of water, safe mixing of electrolyte, etc. already discussed, apply to all kinds of electrolysers and boosters of every design. So, please understand that these are universal features which need to be understood when using any design of booster.

It is possible to produce large volumes of hydroxy gas from a DC booster, enough gas to run a small motor directly on it. For this, we need to pay attention to the efficiency factors already covered in this document. The person who is outstanding in this field is Bob Boyce of the USA who has kindly shared his experience and expertise freely with people who want to use serious electrolysers.

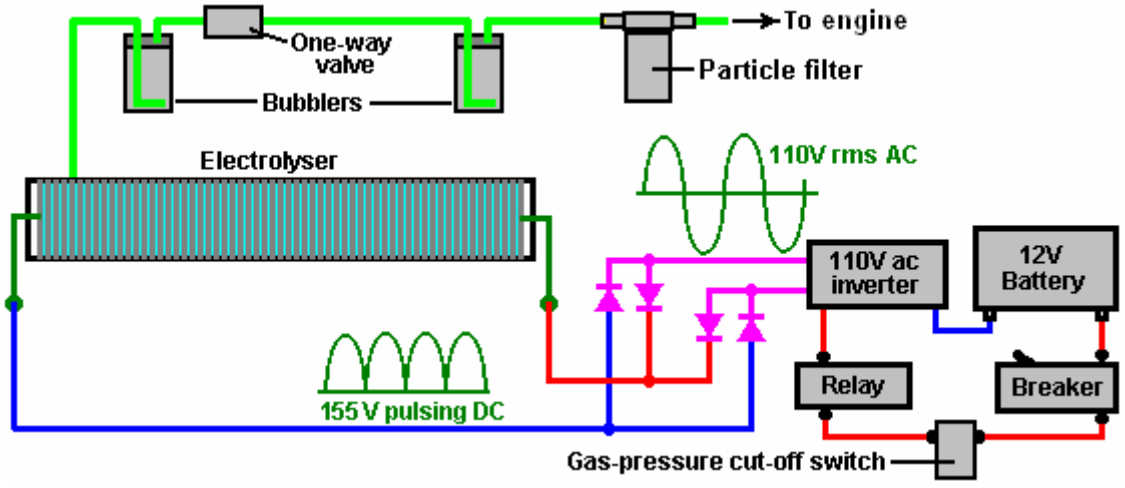
Bob's attention to detail when constructing high-performance electrolysers has resulted in efficiencies which are more than double those of the very famous Michael Faraday whom most scientists consider to be the final word on electrolysis.

Bob Boyce's High-Efficiency Electrolyser.

We are now moving from the "casual" style of booster to the "serious" style of electrolyser. In this category, you will find that the units built are not cheap, weight a considerable amount, require considerable skill to make and usually are quite large physically. I will mention two designs here. First, the very well-known design from Bob Boyce. For this electrolyser, Bob makes solid stainless steel electrode plates act as cell partitions as well as being electrodes. This is a clever technique but it takes a very high level of construction accuracy to make a box with slots in the side and base, so that the stainless steel plates can be slid into the box and when there, form a watertight seal between the cells, preventing electrical current bypassing the places by flowing around them.

The number of cells in the electrolyser depends on the electrical DC voltage supply which is produced from the electrics of the vehicle. This higher voltage is created by using a standard off-the-shelf "inverter" which produces high-voltage alternating current ("AC") meant to be the equivalent of the local electricity mains supply. In the USA, the voltage produced is in the 110 to 120 volt region, elsewhere, it is in the 220 to 230 volt region.

If you are not familiar with electrical jargon, then check out chapter 12 which explains it step by step. The AC output from whatever inverter you buy, is changed back into DC by using a component called a "diode bridge" and a reservoir device called a capacitor. When this is done, the resulting DC voltage is 41% greater than the quoted AC voltage, so a 110-volt inverter will produce about 155 volts and a 220-volt inverter about 310 volts. As you want about 2 volts per cell, the number of cells would be about 80 or 150 depending on which inverter is used. This large number of stainless steel plates each sized at six-inches (150 mm) square, creates a substantial weight which then is increased by the weight of the case, and the electrolyte. The overall arrangement (without the capacitor) is like this:



A very high-precision box for this style of electrolyser can be had from **Ed Holdgate** of Florida who has also shared the construction methods if you fancy yourself as a skilled fabricator:

Ed's website is at http://www.holdgateenterprises.com/Electrolyzer/index.html and each case is hand-made.

The gas production rate is so high that the gas outlet pipe has to have holes drilled along the top in order to try to exclude spray and moisture from the massive rate of bubbles bursting at the surface of the electrolyte. The high efficiency of Bob's electrolysers is due to his meticulous preparation and construction methods. You will notice that Bob recommends the use of a particle filter with a 1-micron mesh, between the engine and the hydroxy system. Apart from ensuring that everything entering the engine is very clean, the particle filter with a mesh of that small size, also acts as a flashback-preventer as flame can't pass through it.

Firstly, the stainless steel plates are cross-scored with sandpaper to create a specially shaped plate surface which helps high-speed bubble release. Secondly, the plates are put through a rigorous "cleansing" process where they are subjected to repeated periods of electrolysis followed by rinsing particles off the plates and filtering the electrolyte solution. When no further particles break free from the plates, they are then put through a "conditioning" process which develops a catalytic layer on the plate surfaces.

This processing and the various construction details are provided in the following free download document, thanks to Bob's generosity in sharing his experience with us: http://www.free-energy-info.com/D9.pdf and there is a forum for Bob's design: http://tech.groups.yahoo.com/group/WorkingWatercar/ where questions are answered.

Pulsed Water-Splitters.

There is a much more efficient way of converting water into a hydroxy gas mix. Unlike the electrolysis devices already described, this method does not need an electrolyte. Pioneered by Stanley Meyer, pulse trains are used to stress water molecules until they break apart, forming the required gas mix. Henry Puharich also developed a very successful system with a somewhat different design. Neither of these gentlemen shared sufficient practical information for us to replicate their designs as a routine process, so we are in a position today where we are searching for the exact details of the methods which they used.

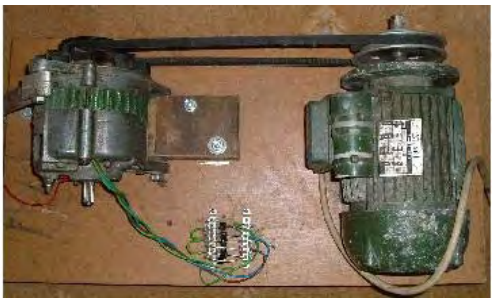
Dave Lawton's Replication of Stan Meyer's "Water Fuel Cell".

The first significant replication of which I am aware, came from Dave Lawton of Wales. By using very considerable tenacity, he discovered the practical details of how to replicate one of Stan Meyer's early designs which is called by the rather confusing name of the "Water Fuel Cell". Dave's work was copied and experimented with by Ravi Raju of India who had considerable success and who posted videos of his results on the web. More recently, Dr Scott Cramton of the USA has adapted the design construction slightly and achieved very satisfactory rates of electrical efficiency, producing some 6 lpm of hydroxy gas for just 3 amps of current at 12 volts.

Dave Lawton

The video of Dave Lawton's replication of Stanley Meyer's demonstration electrolyser (not Stan's production system) seen at http://www.free-energy-info.com/WFCrep.wmv has caused several people to ask for more details.

The electrolysis shown in that video was driven by an alternator, solely because Dave wanted to try each thing that Stan Meyer had done. Dave's alternator and the motor used to drive it are shown here:



The technique of DC pulsing requires the use of electronics, so the following descriptions contain a considerable amount of circuitry. If you are not already familiar with such circuits, then you would be well advised to read through Chapter 12 which explains this type of circuitry from scratch.

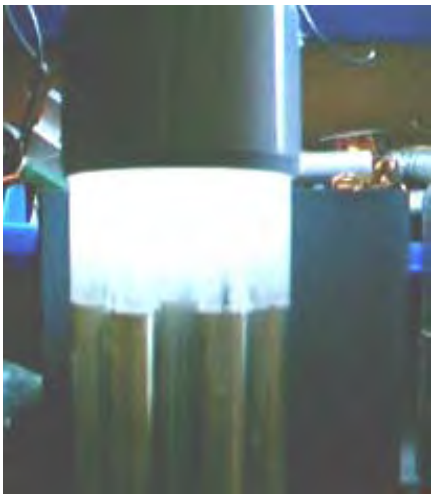
The field coil of Dave's alternator is switched on and off by a Field-Effect Transistor (a "FET") which is pulsed by a dual 555 timer circuit. This produces a composite waveform which produces an impressive rate of electrolysis. The tubes in this replication are made of 316L grade stainless steel, five inches long although Stan's tubes were about sixteen inches long. The outer tubes are 1 inch in diameter and the inner tubes 3/4 inch in diameter. As the wall thickness is 1/16 inch, the gap between them is between 1 mm and 2 mm. The inner pipes are held in place at each end by four rubber strips about one quarter of an inch long.

The container is made from two standard 4 inch diameter plastic drain down-pipe coupler fittings connected to each end of a piece of acrylic tube with PVC solvent cement. The acrylic tube was supplied already cut to size by Wake Plastics, 59 Twickenham Road, Isleworth, Middlesex TW7 6AR Telephone 0208-560-0928. The seamless stainless steel tubing was supplied by: http://www.metalsontheweb.co.uk/asp/home.asp

It is not necessary to use an alternator - Dave just did this as he was copying each thing that Stan Meyer did. The circuit without the alternator produces gas at about the same rate and obviously draws less current as there is no drive motor to be powered. A video of the non-alternator operation can be downloaded using this link: http://www.free-energy-info.co.uk/WFCrep2.wmv.

Dave's electrolyser has an acrylic tube section to allow the electrolysis to be watched, as shown here:

The electrolysis takes place between each of the inner and outer tubes. The picture above shows the bubbles just starting to leave the tubes after the power is switched on. The picture below shows the situation a few seconds later when the whole of the area above the tubes is so full of bubbles that it becomes completely opaque:

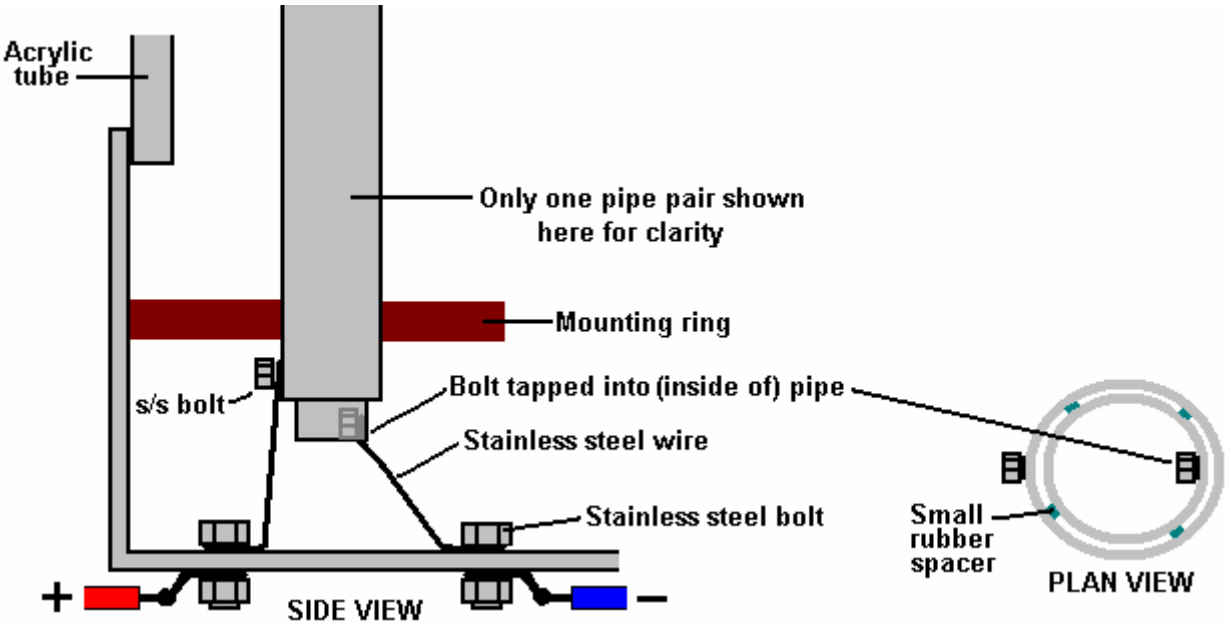


The mounting rings for the tubes can be made from any suitable plastic, such as that used for ordinary food-chopping boards, and are shaped like this:

And the 316L grade stainless steel, seamless tubes are held like this:

Here is the assembly ready to receive the inner tubes (wedged into place by small pieces of rubber):

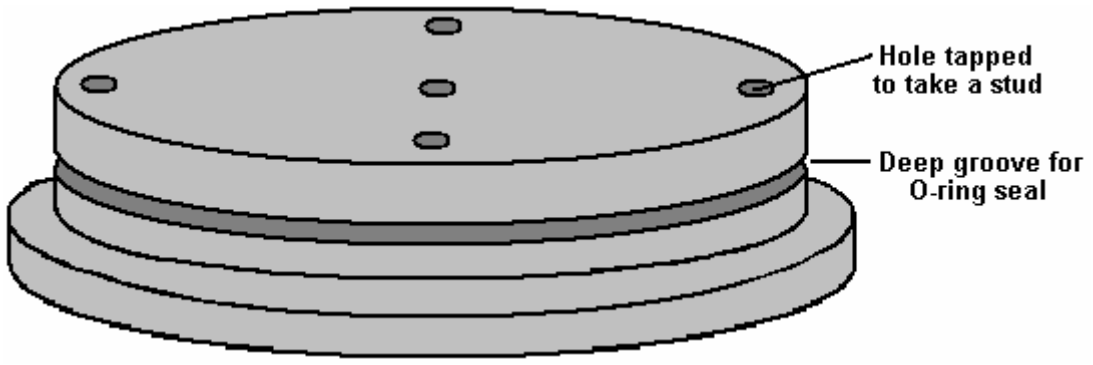
The electrical connections to the pipes are via stainless steel wire running between stainless steel bolts tapped into the pipes and stainless steel bolts running through the base of the unit:



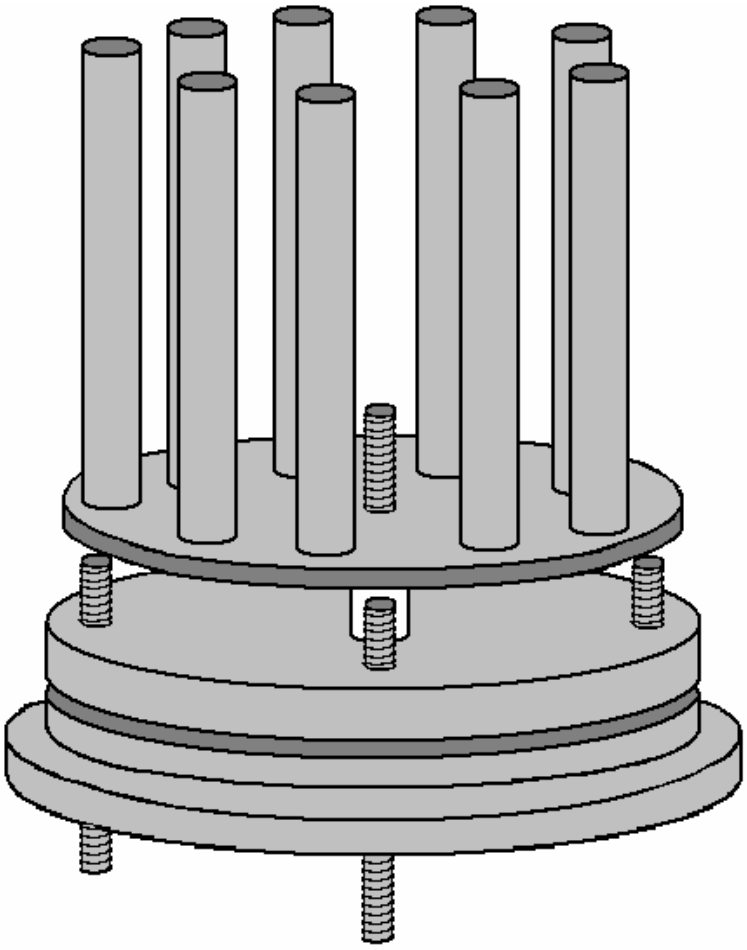
The bolts tapped into the inner tubes should be on the inside. The bolts going through the base of the unit should be tapped in to give a tight fit and they should be sealed with Sikaflex 291 or marine GOOP bedding agent which should be allowed to cure completely before the unit is filled for use. An improvement in performance is produced if the non-active surfaces of the pipes are insulated with any suitable material. That is, the outsides of the outer tubes and the insides of the inner tubes, and if possible, the cut ends of the pipes.

Stan Meyer's Style of Construction.

While Dave's style of construction is simple and straightforward, recently, a copy of one of Stan Meyer's actual construction drawings has surfaced. The image quality of this copy is so low that much of the text can't be read, so the replication presented here may not be exact or might be missing some useful item of information. Stan's construction is unusual. First, a piece of plastic is shaped as shown here:

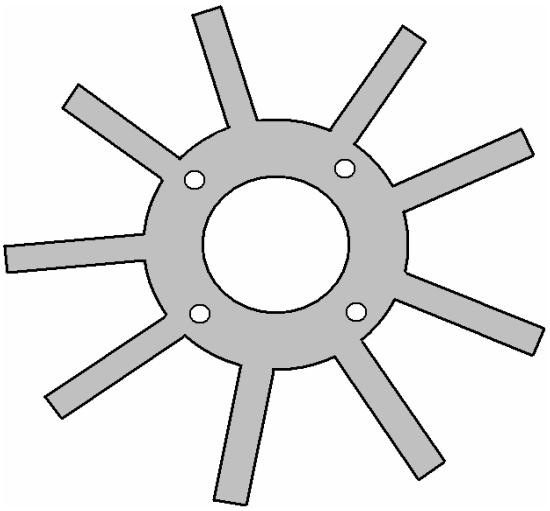


The size of this disc is matched exactly to the piece of clear acrylic used for the body of the housing. The drawing does not make it clear how this disc is attached to the acrylic tube, whether it is a tight push fit, glued in place or held in position with bolts which are not shown. The implication is that a ring of six bolts are driven through the top and tapped into the acrylic tube, as these are shown on one of the plan views, though not on the cross-section. It would also be reasonable to assume that a similar ring of six bolts is also used to hold the base securely in position. There is a groove cut in the plastic base to take an O-ring seal which will be compressed tightly when the disc is in place. There are either two or three threaded stud recesses plus two through holes to carry the electric current connections. The pipe support arrangement is unusual:



A ring of nine evenly-spaced inner pipes are positioned around the edge of a steel disc which is slightly smaller than the inside dimension of the acrylic tube. The pipes appear to be a tight push-fit in holes drilled very accurately through the disc. These holes need to be exactly at right-angles to the face of the disc in order for the pipes to be exactly aligned with the acrylic tube – definitely a drill-press job. The disc is mounted on a central threaded rod which projects through the plastic base disc, and a plastic spacer is used to hold the disc clear of the studs positioned at ninety degrees apart around the outer edge of the base disc.

The mounting for the outer tubes is also most unusual. A piece of steel plate is cut with nine projecting arms at evenly-spaced positions around a circular washer shape as shown here:

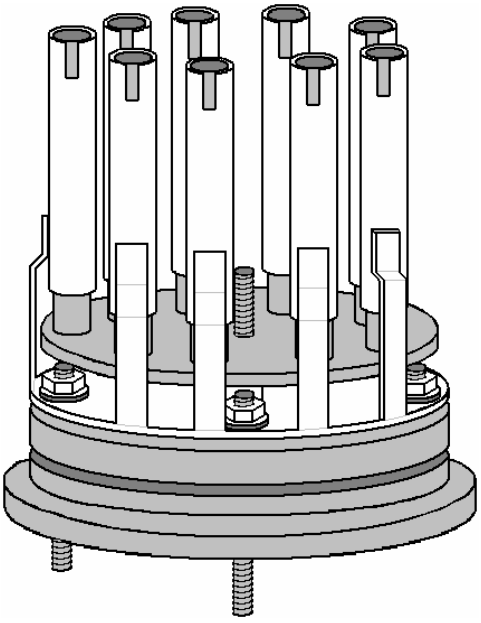


This piece has four holes drilled in it to match the stud positions of the plastic base piece. The number of studs is not specified and while I have shown four, the plate resonance might be helped if there were just three. The size is arranged so that when the arms are bent upwards at right-angles, they fit exactly against the inner face of the acrylic tube.

These arms get two bends in them in order to kink them inwards to form mounts for the outer tubes. The degree of accuracy needed her is considerable as it appears that there are no spacers used between the inner and outer tubes. This means that the very small gap of 1.5 mm or so has to be maintained by the accuracy of these mounts for the outer tubes.

It should be noted that the inner tubes are much longer than the outer tubes and that the outer tubes have a tuning slot cut in them. All of the inner tubes are mechanically connected together through their steel mounting disc and all of the outer tubes are connected together through the ring-shaped steel disc and its kinked arm mounts. It is intended that both of these assemblies should resonate at the same frequency, and they are tuned to do just that. Because the inner tubes have a smaller diameter, they will resonate at a higher frequency than a larger diameter pipe of the same length. For that reason, they are made longer to lower their natural resonant frequency. In addition to that, the slots cut in the outer tubes are a tuning method which raises their resonant pitch. These slots will be adjusted until every pipe resonates at the same frequency.

Looking initially at the mechanical design, suggests that the assembly is impossible to assemble, and while that is almost true, as it will have to be constructed as it is assembled and it appears that the inner and outer pipe assembly can't be taken apart after assembly. This is the way they are put together:



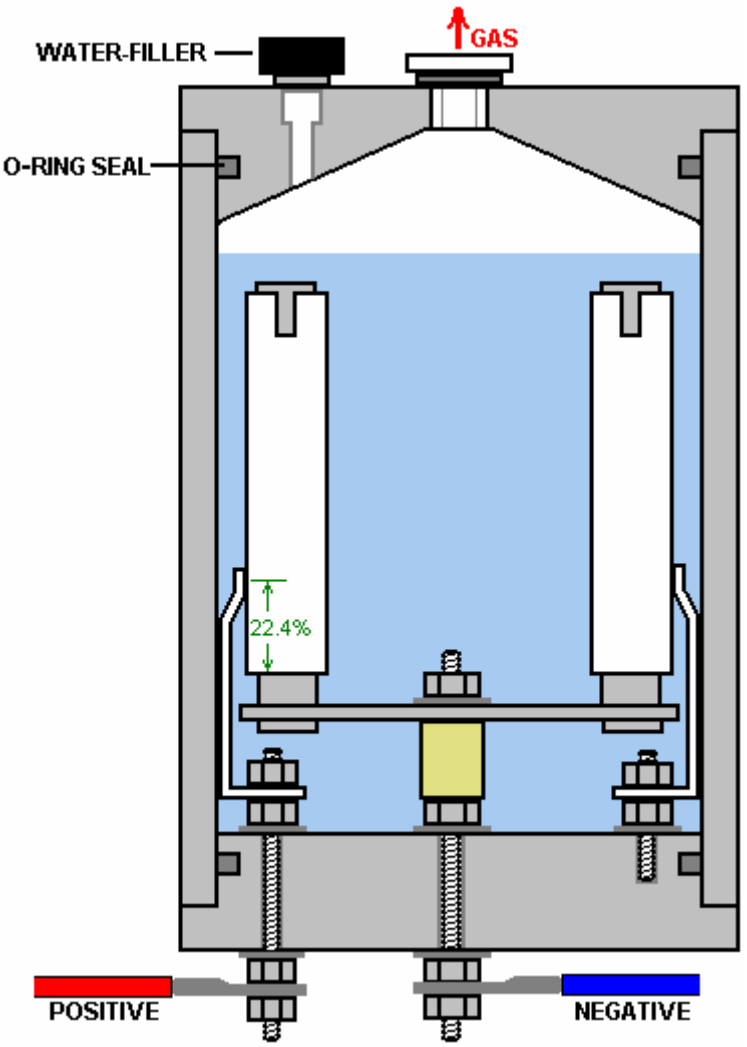
The ring support for the outer pipes is not bolted securely to the plastic base but instead it is spaced slightly above it and mounted on just the stud points. This ring is underneath the slightly smaller diameter disc which holds the inner pipes. This makes it impossible for the two components to be slid together or apart, due to the length of the

pipes. This suggests that either the inner pipes are pushed into place after assembly (which is highly unlikely as they will have been assembled before for tuning) or that the outer pipes are welded to their supports during the assembly process (which is much more likely).

One of the “studs” is carried right through the plastic base in order that it can become the positive connection of the electrical supply, fed to the outer pipes. The central threaded rod is also carried all the way through the plastic base and is used to support the steel plate holding the inner pipes as well as providing the negative electrical connection, often referred to as the electrical “ground”.

Another plastic disc is machined to form a conical lid for the acrylic tube, having a groove to hold an O-ring seal and the water inlet for refilling and the gas output tube. The drawing mentions the fact that if tap water is used, then the impurities in it will collect in the bottom of the electrolyser when the water is removed by being converted to hydroxy gas. This means that the cell would have to be rinsed out from time to time. It also draws attention to the fact that the gasses dissolved in the tap water will also come out during use and will be mixed with the hydroxy gas output.

When these various components are put together, the overall cell construction is shown like this:



This cross-sectional view may be slightly misleading as it suggests that each of the nine outer pipes has its own separate bracket and this is probably not the case as they are connected together electrically through the steel ring-shaped disc and should vibrate as a single unit. It is tempting to use separate brackets as that would allow the assembly to be taken apart quite easily, but the electrical contacts of such a system would be much inferior and so it is not to be recommended.

Because of the way that all of the inner pipes are connected together and all of the outer pipes are connected together electrically, this form of construction is not suited to the three-phase alternator drive shown below, where the nine pipes would have to be connected in separate sets of three. Instead, the solid-state circuit is used, which

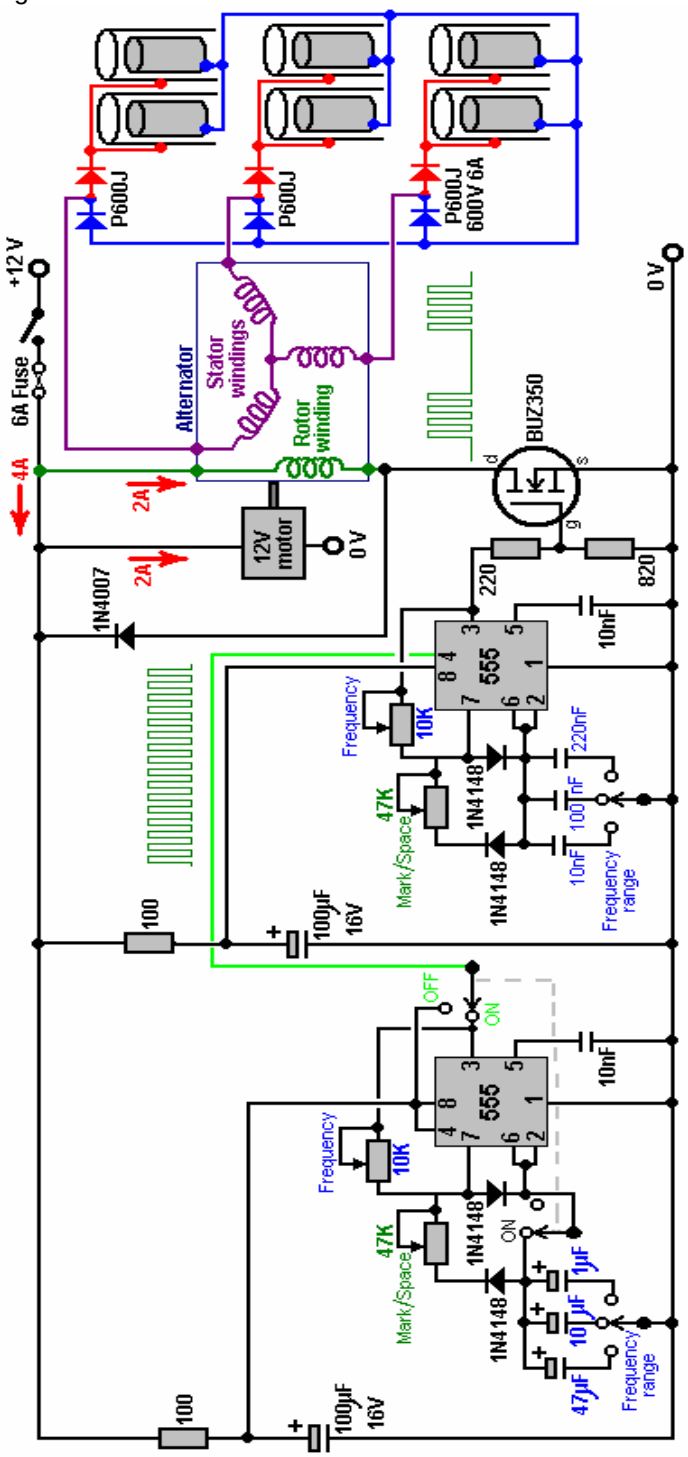
is very effective and which does not have the size, weight, noise and increased current of the alternator arrangement.

If accuracy of construction is a problem, then it might be possible to give the outer pipes a deliberate slope so that they press against the inner pipes at the top, and then use one short spacer to force them apart and give the desired spacing. It seems clear that Stan worked to such a degree of constructional accuracy that his pipes were perfectly aligned all along their lengths.

Dave Lawton points out that the connection point of the brackets for the outer pipes is highly critical as they need to be at a resonating node of the pipes. The connection point is therefore at 22.4% of the length of the pipe from the bottom of the pipe. Presumably, if a slot is cut in the top of the pipe, then the resonant pipe length will be measured to the bottom of the slot and the connection point set at 22.4% of that length.

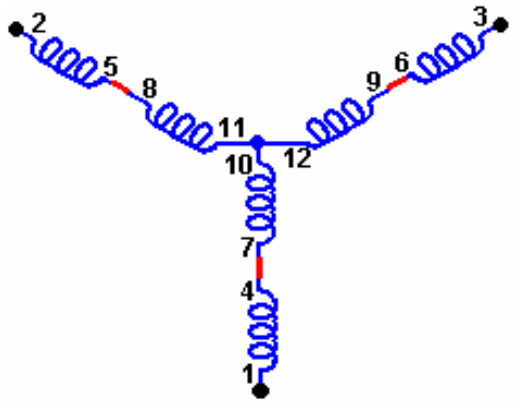
Dave Lawton's 3-Phase Circuit.

Dave Lawton's pipe arrangement can be driven either via an alternator or by an electronic circuit. A suitable circuit for the alternator arrangement is:



In this rather unusual circuit, the rotor winding of an alternator is pulsed via an oscillator circuit which has variable frequency and variable Mark/Space ratio and which can be gated on and off to produce the output waveform shown below the alternator in the circuit diagram. The oscillator circuit has a degree of supply de-coupling by the 100 ohm resistor feeding the 100 microfarad capacitor. This is to reduce voltage ripple coming along the +12 volt supply line, caused by the current pulses through the rotor winding. The output arrangement feeding the pipe electrodes of the electrolyser is copied directly from Stan Meyer's circuit diagram.

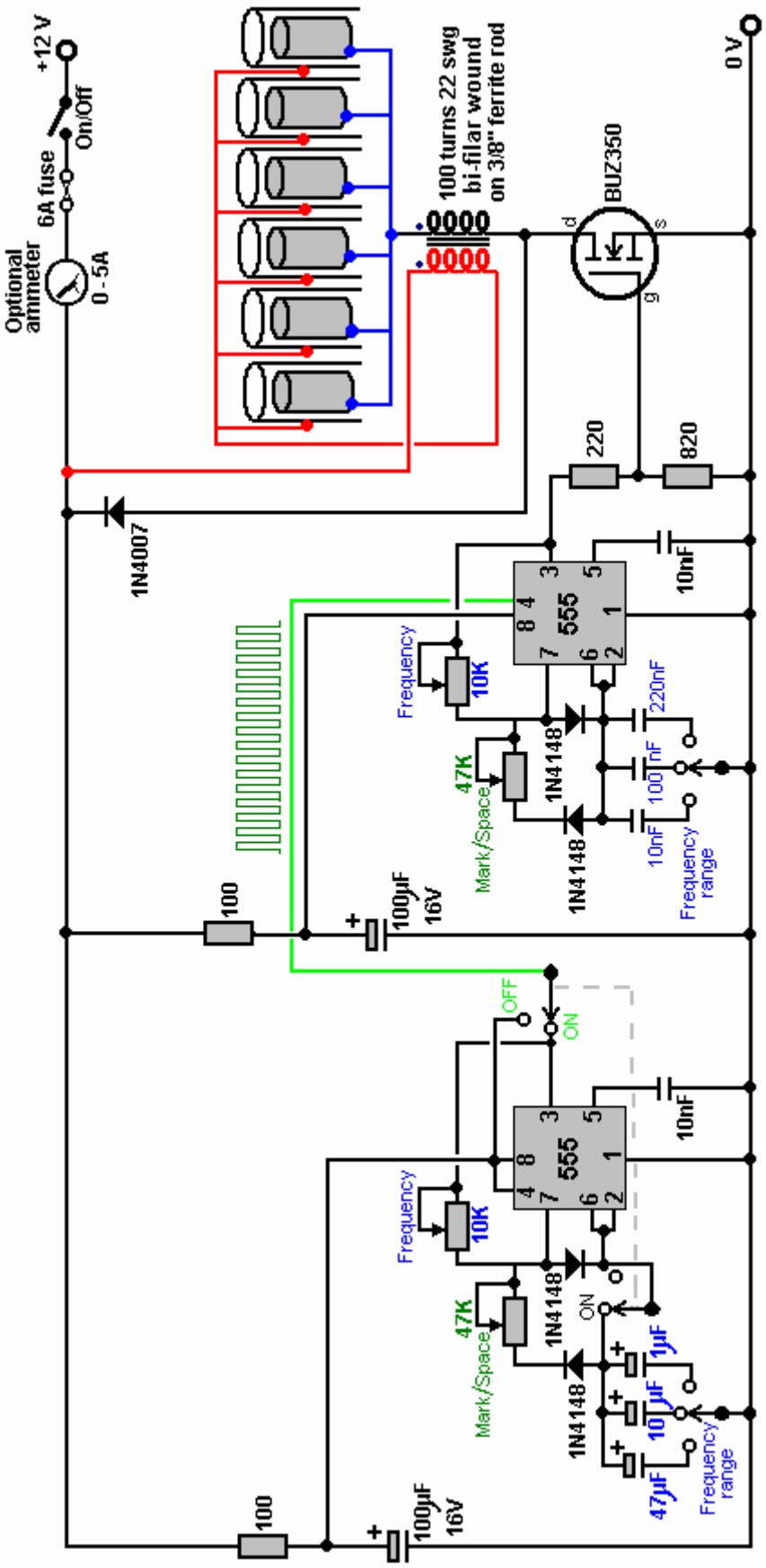
It is not recommended that you use an alternator should you decide to build a copy of your own. But if you decide to use one and the alternator does not have the windings taken to the outside of the casing, it is necessary to open the alternator, remove the internal regulator and diodes and pull out three leads from the ends of the stator windings. If you have an alternator which has the windings already accessible from the outside, then the stator winding connections are likely to be as shown here:



The motor driving Dave's alternator draws about two amps of current which roughly doubles the power input to the circuit. There is no need for the size, weight, noise, mechanical wear and current draw of using a motor and alternator as pretty much the same performance can be produced by the solid-state circuit with no moving parts.

Both circuits have been assessed as operating at anything from 300% to 900% of Faraday's "maximum electrical efficiency", it should be stressed that the inductors used in this circuit, form a very important role in altering and amplifying the voltage waveform applied to the cell. Dave uses two "bi-filar wound" inductors, each wound with 100 turns of 22 SWG (21 AWG) enamelled copper wire on a 9 mm (3/8") diameter ferrite rod. The length of the ferrite rod is not at all critical, and a ferrite toroid could be used as an alternative, though that is more difficult to wind. These bi-filar coils are wound at the same time using two lengths of wire side by side. The solid-state circuit is shown here:

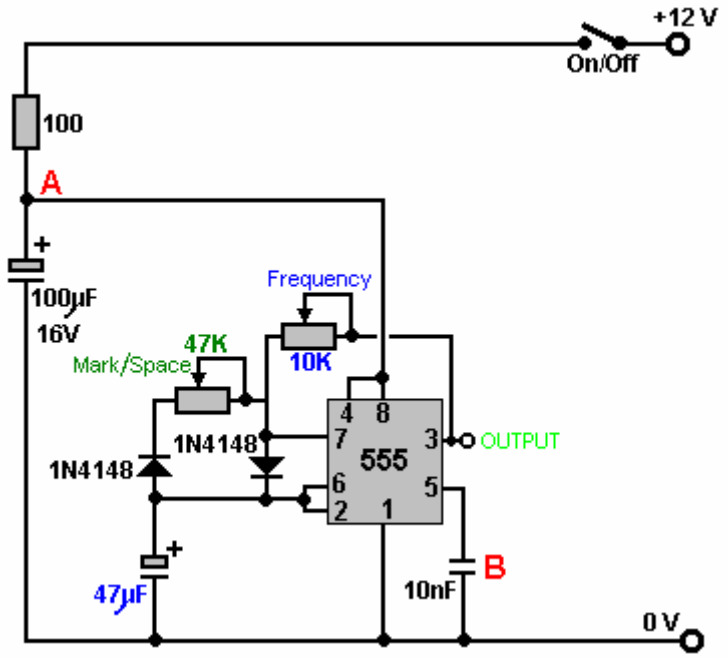
Dave Lawton's Single-Phase Circuit:



Circuit operation:

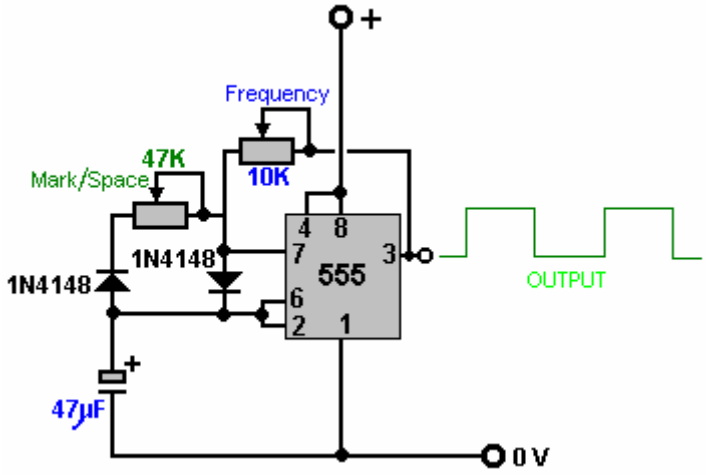
The main part of the circuit is made up of two standard 555 chip timers. These are wired to give an output waveform which switches very rapidly between a high voltage and a low voltage. The ideal waveform shape coming from this circuit is described as a “square wave” output. In this particular version of the circuit, the rate at which the circuit flips between high and low voltage (called the “frequency”) can be adjusted by the user turning a knob. Also, the length of the ON time to the OFF time (called the “Mark/Space Ratio”) is also adjustable.

This is the section of the circuit which does this:



The 100 ohm resistor and the 100 microfarad capacitor are there to iron out any ripples in the voltage supply to the circuit, caused by fierce pulses in the power drive to the electrolysis cell. The capacitor acts as a reservoir of electricity and the resistor prevents that reservoir being suddenly drained if the power supply line is suddenly, and very briefly, pulled down to a low voltage. Between them, they keep the voltage at point "**A**" at a steady level, allowing the 555 chip to operate smoothly.

The very small capacitor "**B**" is wired up physically very close to the chip. It is there to short-circuit any stray, very short, very sharp voltage pulses picked up by the wiring to the chip. It is there to help the chip to operate exactly as it is designed to do, and is not really a functional part of the circuit. So, for understanding how the circuit works, we can ignore them and see the circuit like this:

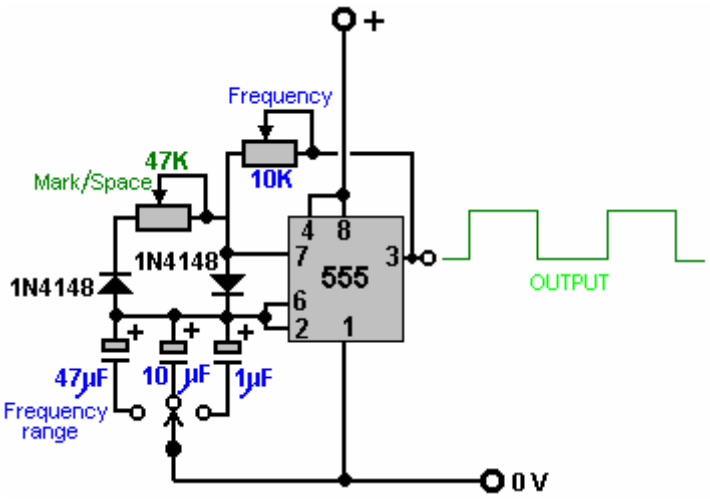


This circuit generates output pulses of the type shown in green with the voltage going high, (the "Mark") and low (the "Space"). The 47K variable resistor (which some people insist on calling a "pot") allows the length of the Mark and the Space to be adjusted from the 50 - 50 shown, to say, 90 - 10 or any ratio through to 10 - 90. It should be mentioned that the "47K" is not at all critical and these are quite likely to be sold as "50K" devices. Most low cost components have a plus or minus 10% rating which means that a 50K resistor will be anything from 45K to 55K in actual value.

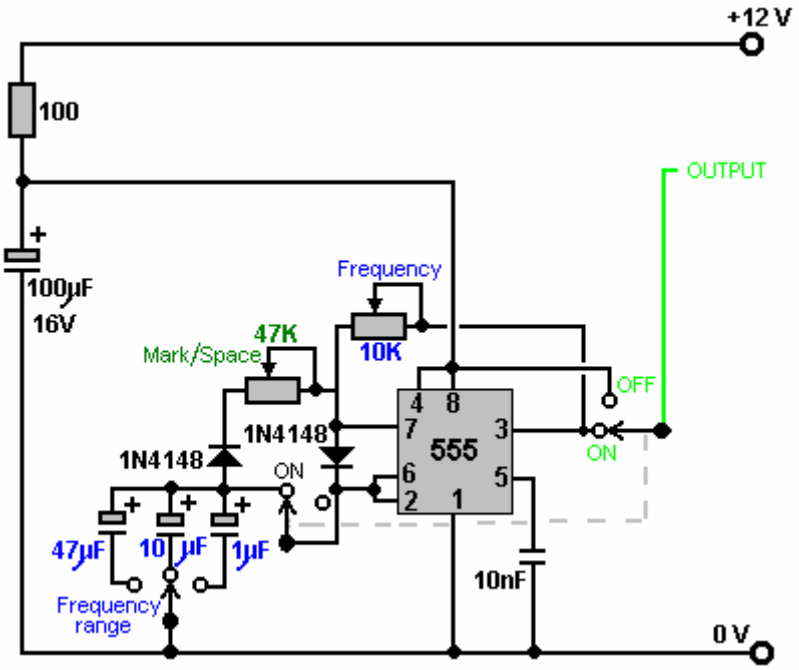
The two "1N4148" diodes are there to make sure that when the Mark/Space 47K variable resistor is adjusted, that it does not alter the frequency of the output waveform in any way. The remaining two components: the 10K variable resistor and the 47 microfarad capacitor, both marked in blue, control the number of pulses produced per

second. The larger the capacitor, the fewer the pulses per second. The lower the value of the variable resistor, the larger the number of pulses per second.

The circuit can have additional frequency tuning ranges, if the capacitor value is altered by switching in a different capacitor. So the circuit can be made more versatile by the addition of one switch and, say, two alternative capacitors, as shown here:

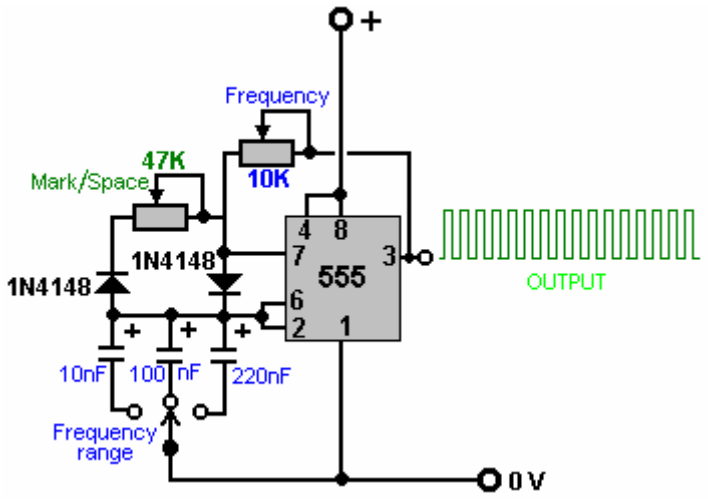


The capacitors shown here are unusually large because this particular circuit is intended to run relatively slowly. In the almost identical section of the circuit which follows this one, the capacitors are very much smaller which causes the switching rate to be very much higher. Experience has shown that a few people have had overheating in this circuit when it is switched out of action, so the On/Off switch has been expanded to be a two-pole changeover switch and the second pole used to switch out the timing elements of the 555 chip. The complete version of this section of the circuit is then:

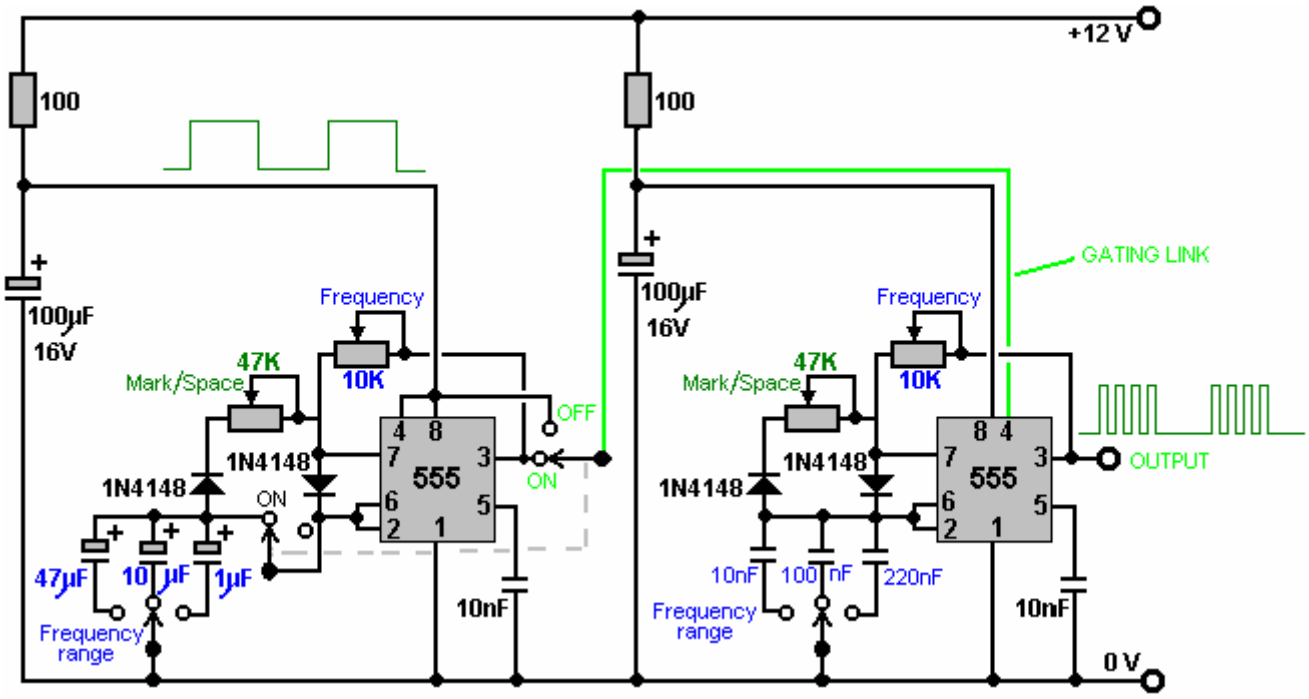


which just has one additional switch to allow the output to be stopped and the 12-volt supply line to be fed instead. The reason for this is that this part of the circuit is used to switch On and Off an identical circuit. This is called “gating” and is explained in Chapter 12 which is an electronics tutorial.

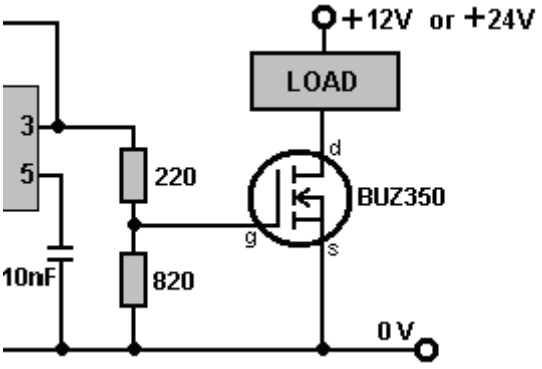
The second part of the circuit is intended to run at much higher speeds, so it uses much smaller capacitors:



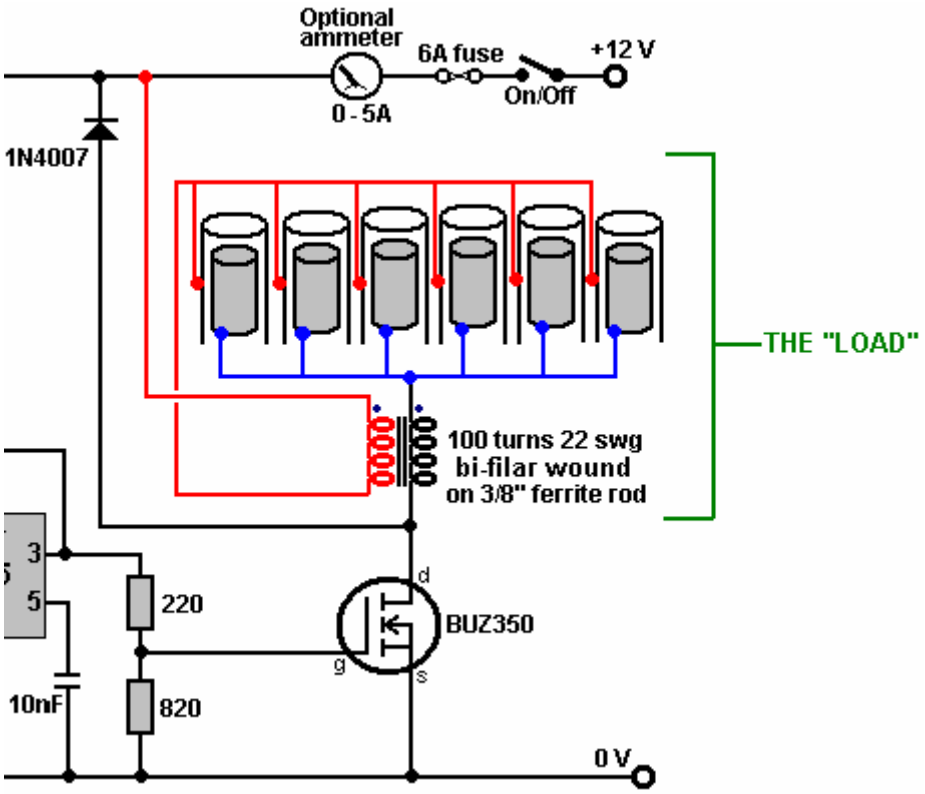
So, putting them together, and allowing the first circuit to switch the second one On and Off, we get:



The final section of the circuit is the power drive for the electrolyser cell. This is a very simple circuit. Firstly, the output of the second 555 chip is lowered by a basic voltage-divider pair of resistors, and fed to the Gate of the output transistor which while it can run on the 12 volts which the pulse-generation circuit needs, Dave does prefer to run on 24 volts as that generates a greater gas flow:



Here, the 555 chip output voltage is lowered by 220 / 820 or about 27%. When the voltage rises, it causes the BUZ350 transistor to switch on, short-circuiting between its Drain and Source connections and applying the whole of the 12-volt supply voltage across the load, which in our application, is the electrolyser cell:



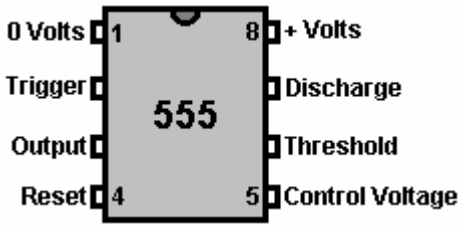
The transistor drives the electrolysis electrodes as shown above, applying very sharp, very short pulses to them. What is very important are the wire coils which are placed on each side of the electrode set. These coils are linked magnetically because they are wound together on a high-frequency ferrite rod core and although a coil is such a simple thing, these coils have a profound effect on how the circuit operates. Firstly, they convert the 555 chip pulse into a very sharp, very short, high-voltage pulse which can be as high as 1,200 volts. This pulse affects the local environment, causing extra energy to flow into the circuit. The coils now perform a second role by blocking that additional energy from short-circuiting through the battery, and causing it to flow through the electrolysis cell, splitting the water into a mix of hydrogen and oxygen, both gases being high-energy, highly charged atomic versions of those gases. This gives the mix some 400% the power of hydrogen being burned in air.

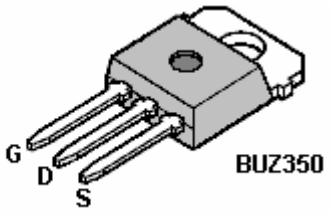
When the transistor switches off, the coils try to pull the transistor Drain connection up to a voltage well above the +12-volt battery line. To prevent this, a 1N4007 diode is connected across the cell and its coils. The diode is connected so that no current flows through it until the transistor Drain gets dragged above the +12-volt line, but when that happens, the diode effectively gets turned over and as soon as 0.7 volts is placed across it, it starts to conduct heavily and collapses the positive-going voltage swing, protecting the transistor. You can easily tell that it is the environmental “cold” electricity which is doing the electrolysis as the cell stays cold even though it is putting out large volumes of gas. If the electrolysis were being done by conventional electricity, the cell temperature would rise during the electrolysis. A John Bedini pulser circuit can be used very effectively with a cell of this type and it adjusts automatically to the resonant frequency as the cell is part of the frequency-determining circuit.

The BUZ350 MOSFET has a current rating of 22 amps so it will run cool in this application. However, it is worth mounting it on an aluminium plate which will act as both the mounting and a heat sink but it should be realised that this circuit is a bench-testing circuit with a maximum current output of about 2 amps and it is **not** a **P**ulse-**W**idth **M**odulation circuit for a high-current DC electrolyser. The current draw in this arrangement is particularly interesting. With just one tube in place, the current draw is about one amp. When a second tube is added, the current increases by less than half an amp. When the third is added, the total current is under two amps. The fourth and fifth tubes add about 100 milliamps each and the sixth tube causes almost no increase in current at all. This suggests that the efficiency could be raised further by adding a large number of additional tubes, but this is actually not the case as the cell arrangement is important. Stan Meyer ran his VolksWagen car for four years on

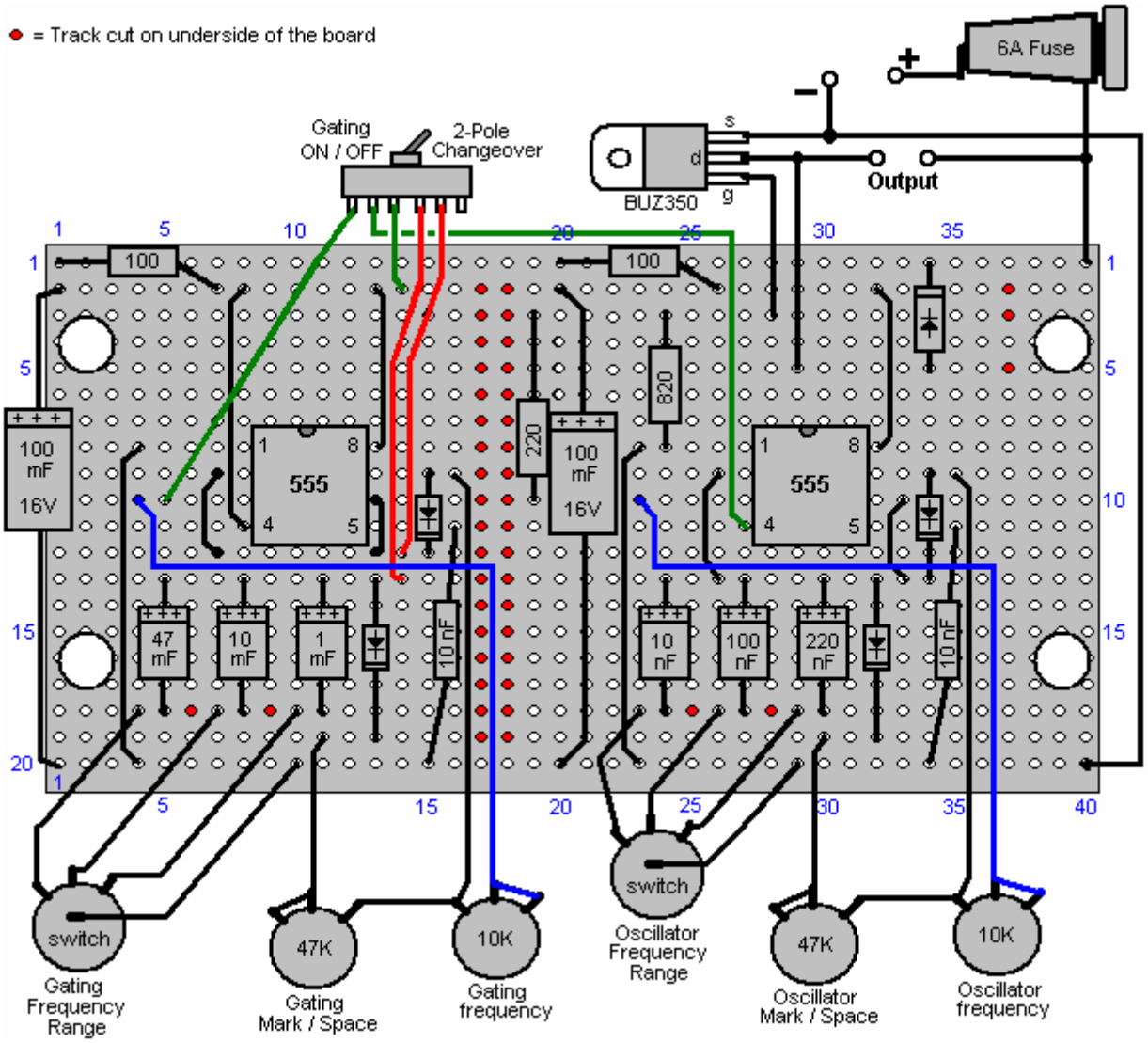
the output from four of these cells with 16-inch (400 mm) electrodes, and Stan would have made a single larger cell had that been feasible.

Although the current is not particularly high, a five or six amp circuit-breaker, or fuse, should be placed between the power supply and the circuit, to protect against accidental short-circuits. If a unit like this is to be mounted in a vehicle, then it is **essential** that the power supply is arranged so that the electrolyser is disconnected if the engine is switched off. Passing the electrical power through a relay which is powered via the ignition switch is a good solution for this. It is also **vital** that at least one bubbler is placed between the electrolyser and the engine, to give some protection if the gas should get ignited by an engine malfunction.

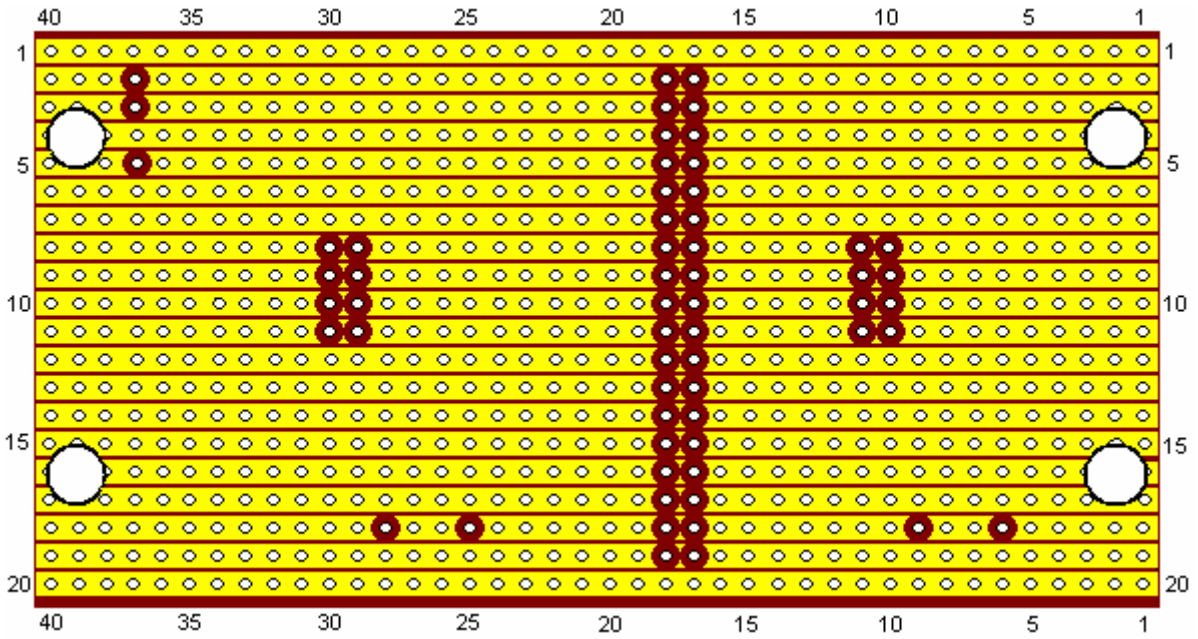


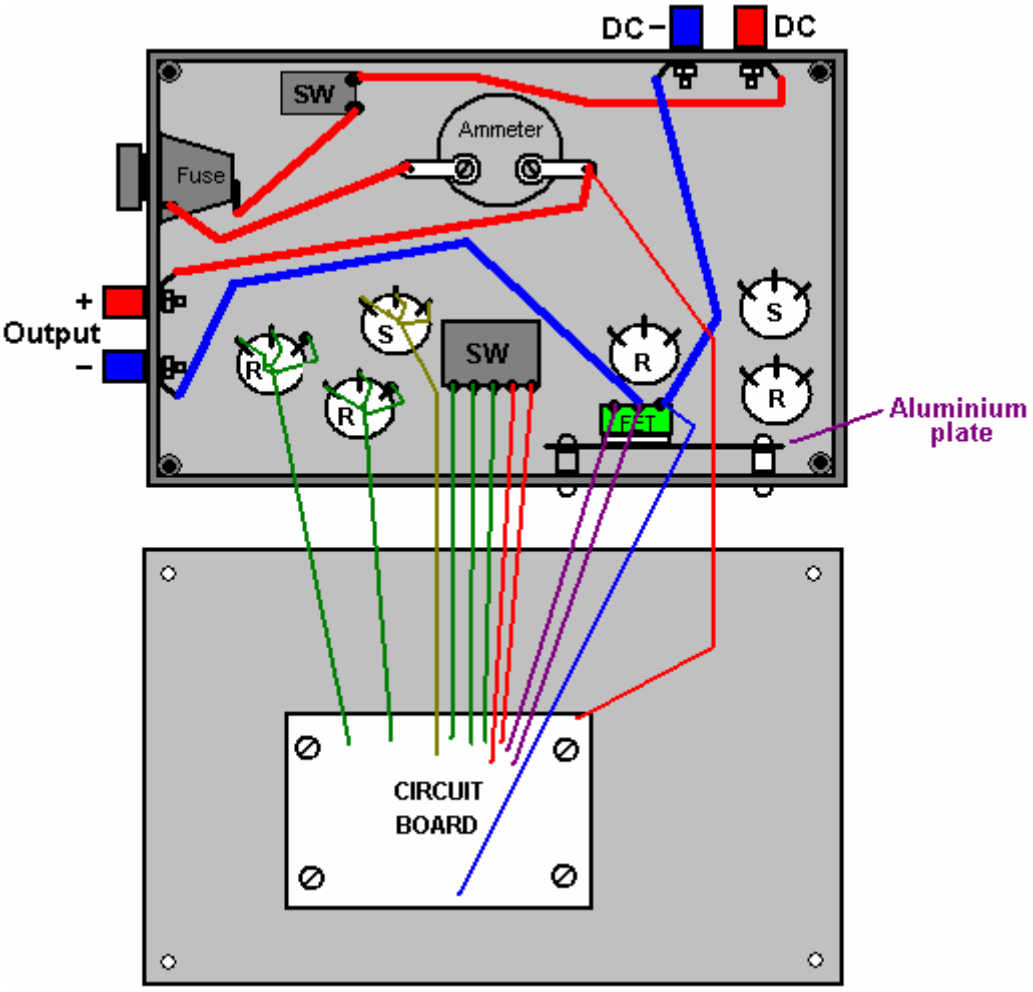


Although printed circuit boards have now been produced for this circuit and ready-made units are available commercially, you can build your own using stripboard if you want to. A possible one-off prototype style component layout for is shown here:

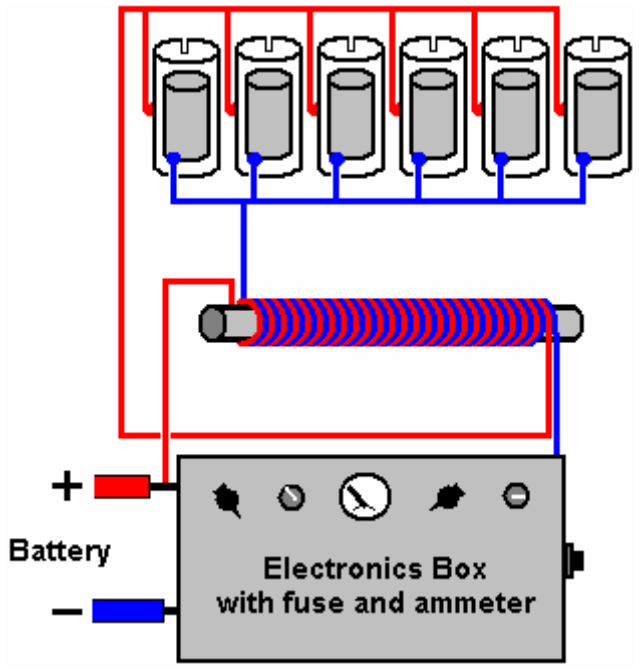


The underside of the strip-board (when turned over horizontally) is shown here:



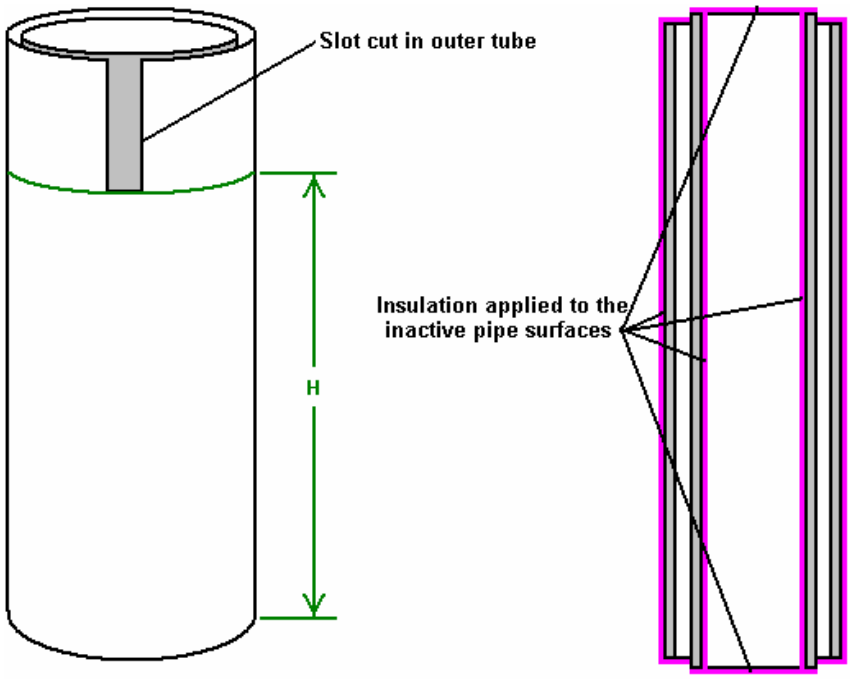


Although using a ferrite ring is probably the best possible option, the bi-filar coil can be wound on any straight ferrite rod of any diameter and length. You just tape the ends of two strands of wire to one end of the rod and then rotate the rod in your hands, guiding the strands into a neat side-by-side cylindrical winding as shown here:



Component	Quantity	Description	Comment
100 ohm resistors 0.25 watt	2	Bands: Brown, Black, Brown	
220 ohm resistor 0.25 watt	1	Bands: Red, Red, Brown	
820 ohm resistor 0.25 watt	1	Bands: Gray, Red, Brown	
100 µF 16V capacitor	2	Electrolytic	
47µF 16V capacitor	1	Electrolytic	
10 µF 16V capacitor	1	Electrolytic	
1 µF 16 V capacitor	1	Electrolytic	
220 nF capacitor (0.22 mF)	1	Ceramic or polyester	
100 nF capacitor (0.1 mF)	1	Ceramic or polyester	
10 nF capacitor (0.01 mF)	3	Ceramic or polyester	
1N4148 diodes	4		
1N4007 diode	1		FET protection
NE555 timer chip	2		
BUZ350 MOSFET	1	Or any 200V 20A n-channel MOSFET	
47K variable resistors	2	Standard carbon track	Could be screw track
10K variable resistors	2	Standard carbon track	Could be screw track
4-pole, 3-way switches	2	Wafer type	Frequency range
1-pole changeover switch	1	Toggle type, possibly sub-miniature	Any style will do
1-pole 1-throw switch	1	Toggle type rated at 10 amps	Overall ON / OFF switch
Fuse holder	1	Enclosed type or a 6A circuit breaker	Short-circuit protection
Veroboard	1	20 strips, 40 holes, 0.1 inch matrix	Parallel copper strips
8-pin DIL IC sockets	2	Black plastic, high or low profile	Protects the 555 ICs
Wire terminals	4	Ideally two red and two black	Power lead connectors
Plastic box	1	Injection moulded with screw-down lid	
Mounting nuts, bolts and pillars	8	Hardware for 8 insulated pillar mounts	For board and heatsink
Aluminium sheet	1	About 4 inch x 2 inch	MOSFET heatsink
Rubber or plastic feet	4	Any small adhesive feet	Underside of case
Knobs for variable resistors etc.	6	1/4 inch shaft, large diameter	Marked skirt variety
Ammeter	1	Optional item, 0 to 5A or similar	
Ferrite rod 1-inch long or longer	1	For construction of the inductors	bi-filar wound
22 SWG (21 AWG) wire	1 reel	Enamelled copper wire, 2 oz. reel	
Sundry connecting wire	4 m	Various sizes	

Dave, who built this replication, suggests various improvements. Firstly, Stan Meyer used a larger number of tubes of greater length. Both of those two factors should increase the gas production considerably. Secondly, careful examination of video of Stan's demonstrations shows that the outer tubes which he used had a rectangular slot cut in the top of each tube:



Some organ pipes are fine-tuned by cutting slots like this in the top of the pipe, to raise it's pitch, which is it's frequency of vibration. As they have a smaller diameter, the inner pipes in the Meyer cell will resonate at a higher frequency than the outer pipes. It therefore seems probable that the slots cut by Stan are to raise the resonant frequency of the larger pipes, to match the resonant frequency of the inner pipes. If you want to do that, hanging the inner tube up on a piece of thread and tapping it, will produce a sound at the resonant pitch of the pipe. Cutting a slot in one outer pipe, suspending it on a piece of thread and tapping it, will allow the pitch of the two pipes to be compared. When one outer pipe has been matched to your satisfaction, then a slot of exactly the same dimensions will bring the other outer pipes to the same resonant pitch. It has not been proved, but it has been suggested that only the part of the outer pipe which is below the slot, actually contributes to the resonant frequency of the pipe. That is the part marked as "H" in the diagram above. It is also suggested that the pipes will resonate at the same frequency if the area of the inside face of the outer pipe ("H" x the inner circumference) exactly matches the area of the outer surface of the inner pipe. It should be remembered that as all of the pipe pairs will be resonated with a single signal, that each pipe pair needs to resonate at the same frequency as all the other pipe pairs.

It is said that Stan ran his VolksWagen car for four years, using just the gas from four of these units which had pipe pairs 16-inches long. A very important part of the cell build is the conditioning of the electrode tubes, using tap water. Ravi in India suggests that this is done as follows:

1. Do not use any resistance on the negative side of the power supply when conditioning the pipes.
2. Start at 0.5 Amps on the signal generator and after 25 minutes, switch off for 30 minutes
3. Then apply 1.0 Amps for 20 minutes and then stop for 30 minutes.
4. Then apply 1.5 Amps for 15 minutes and then stop for 20 minutes.
5. Then apply 2.0 Amps for 10 minutes and afterwards stop for 20 minutes.
6. Go to 2.5 Amps for 5 minutes and stop for 15 minutes.
7. Go to 3.0 Amps for 120 to 150 seconds. You need to check if the cell is getting hot...if it is you need to reduce the time.

After the seven steps above, let the cell stand for at least an hour before you start all over again.

You will see hardly any gas generation in the early stages of this conditioning process, but a lot of brown muck will be generated. Initially, change the water after every cycle, but **do not** touch the tubes with bare hands. If the ends of the tubes need to have muck cleaned off them, then use a brush but do not touch the electrodes!! If the brown muck is left in the water during the next cycle, it causes the water to heat up and you need to avoid this.

Over a period of time, there is a reduction in the amount of the brown stuff produced and at some point, the pipes won't make any brown stuff at all. You will be getting very good gas generation by now. A whitish powdery coat of chromium oxide dielectric will have developed on the surfaces of the electrodes. Never touch the pipes with bare hands once this helpful coating has developed.

Important: Do the conditioning in a well-ventilated area, or alternatively, close the top of the cell and vent the gas out into the open. During this process, the cell is left on for quite some time, so even a very low rate of gas production can accumulate a serious amount of gas which would be a hazard if left to collect indoors.

Further Developments

When producing hydroxy gas from water, it is not possible to exceed the Faraday maximum unless additional energy is being drawn in from the surrounding environment. As this cell runs cold and has substantial gas output, there is every indication that when it is running, it is drawing in this extra energy.

This idea is supported by the fact that one of the key methods of tapping this extra energy is by producing a train of very sharply rising and sharply falling electrical pulses. This is exactly the objective of Dave's circuit, so it would not be too surprising if that effect were happening.

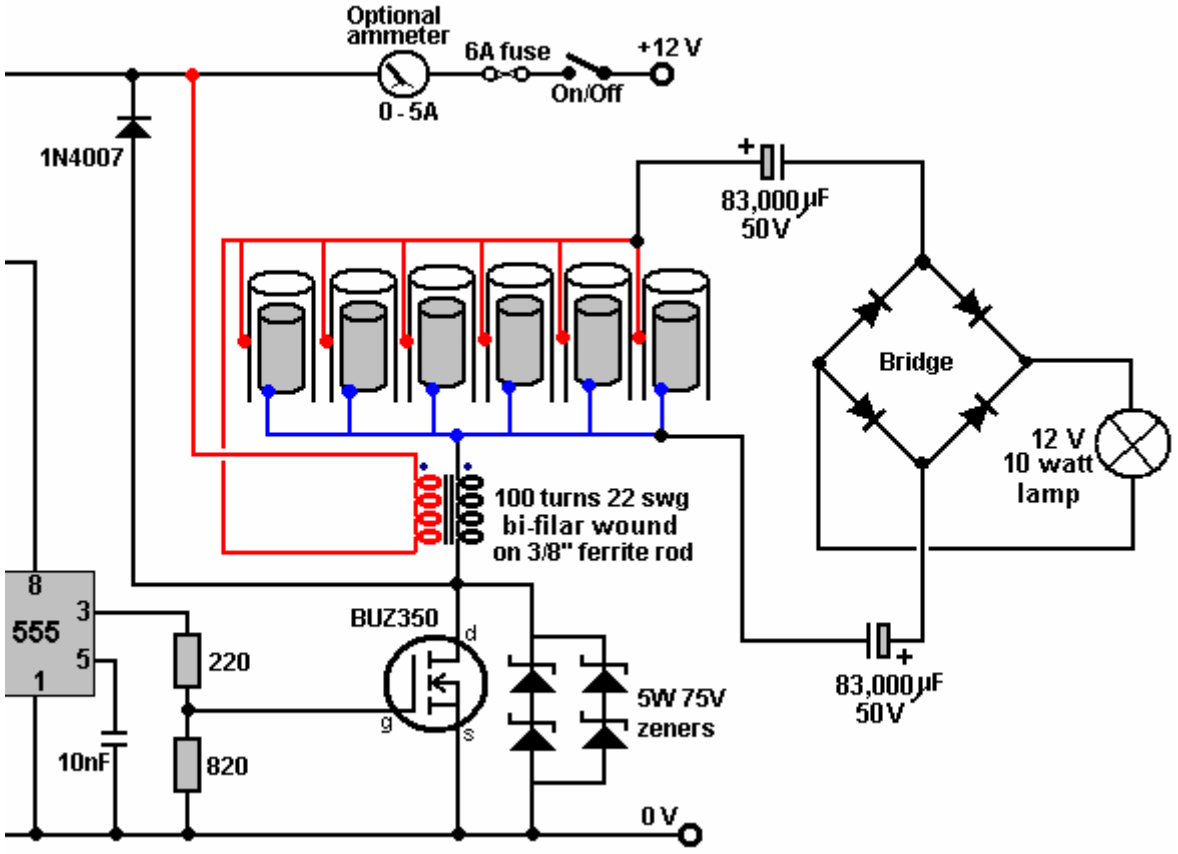
The additional energy being accessed is sometimes referred to as "cold" electricity, which has very different characteristics to normal conventional electricity. Where normal electrical losses cause local heating as a by-product, "cold" electricity has exactly the opposite effect, and where a normal electrical loss would take place, an extra inflow of useful "cold" energy enters the circuit from outside. This flow causes the temperature of the circuitry to drop, instead of increase, which is why it is called "cold" electricity.

This remarkable occurrence has the most unusual effect of actually reducing the amount of conventional power needed to drive the circuit, if the output load is increased. So, increasing the load powered by the circuit causes additional energy to flow in from the environment, powering the extra load and as well, helping to drive the original circuit. This seems very strange, but then, "cold" electricity operates in an entirely different way to our familiar

conventional electricity, and it has its own set of unfamiliar rules, which are generally the reverse of what we are used to.

To test his cell system further, Dave connected an extra load across the electrodes of his cell. As the inductors connected each side of the cell generate very high-value, sharp voltage spikes, Dave connected two large value capacitors (83,000 microfarad, 50-volt) across the cell as well. The load was a 10-watt light bulb which shines brightly, and interestingly, the current draw of the circuit goes down rather than up, in spite of the extra output power. The gas production rate appears undiminished.

This is the alteration to that part of the circuit which was used:



It has also been suggested that if a BUZ350 can't be obtained, then it would be advisable to protect the output FET against damage caused by accidental short-circuiting of wires, etc., by connecting what is effectively a 150-volt, 10 watt zener diode across it as shown in the above diagram. While this is not necessary for the correct operation of the circuit, it is helpful in cases where accidents occur during repeated testing and modification of the cell components.

Dr Scott Cramton's Cell Construction.

Dr. Cramton and his team of Laesa Research and Development scientists have been investigating and advancing this technology and they have reached an output of six litres per minute for an electrical input of 12 watts (1 amp at 12 volts). In addition, Dr. Cramton's cell has stable frequency operation and is being run on local

well water. The objective is to reduce the amount of diesel fuel needed to run a large capacity standard electrical generator.

The style of design is similar to Stan Meyer's original physical construction although the dimensions are slightly different. The cell body is transparent acrylic tube with end caps top and bottom. Inside the tube are nine pairs of pipes, electrically connected as three sets of three interspersed pipe pairs. These are driven by a three-phase pulsed supply based on a replication of Stan Meyer's original cell. It consists of a Delco Remy alternator driven by a 1.5 horsepower 220 volt AC motor. This arrangement is, as was Stan Meyer's, for demonstration purposes. In a working application, the alternator is driven by the engine being supplied with the hydroxy gas. The 120 degree phase separation is the critical component for maintaining the resonant frequency. It should be noted that the alternator must maintain a rate of 3,600 rpm while under load.

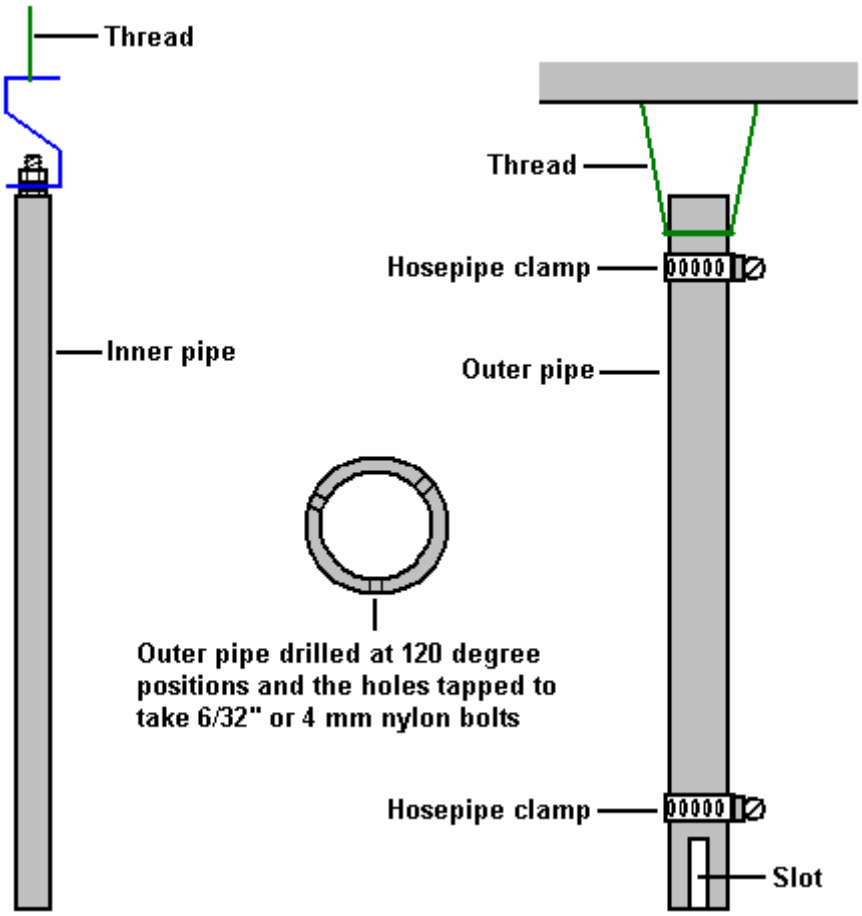
It needs to be stressed that Dr. Cramton's cell is very close in construction principles to Dave Lawton's cell and the quality of construction is very important indeed. The first and foremost point which can be easily missed is the absolutely essential tuning of all of the pipes to a single, common frequency. This is the equivalent of tuning a musical instrument and without that tuning, the essential resonant operation of the cell will not be achieved and the cell performance will not be anything like the results which Dr. Cramton and his team are getting.

Dr. Cramton is using 316L-grade stainless steel pipes 18 inches (450 mm) long. The outer pipes are 0.75 inches in diameter and the inner pipes 0.5 inches in diameter. This gives an inter-pipe gap of 1.2 mm. The first step is to get the pipes resonating together. First, the frequency of an inner pipe is measured. For this, a free internet frequency-analyzer program was downloaded and used with the audio card of a PC to give a measured display of the resonant frequency of each pipe. The download location is
http://www.softpedia.com/get/Multimedia/Audio/Other-AUDIO-Tools/Spectrum-Analyzer-pro-Live.shtml

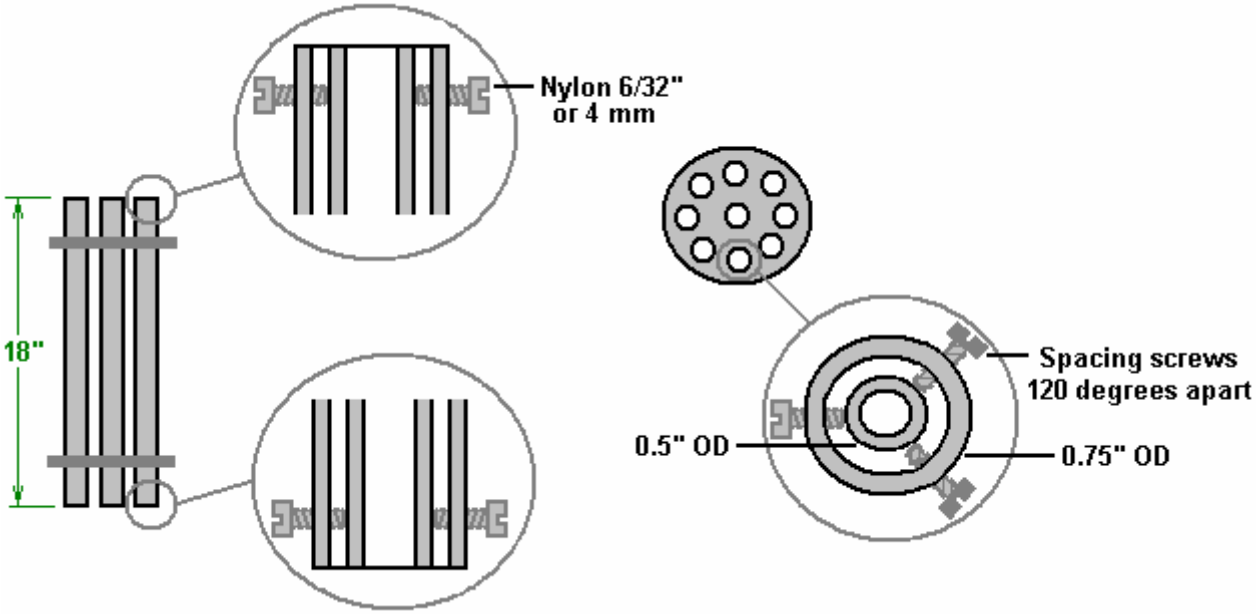
The method for doing this is very important and considerable care is needed for this. The quarter-inch stainless steel bolt is pressed into the inner pipe where it forms a tight push-fit. It is very important that the head of each nut is pressed in for **exactly** the same distance as this alters the resonant frequency of the inner pipe. The steel connecting strip is then bent into its Z shape and securely clamped to the bolt with a stainless steel nut. The assembly of pipe, steel strip, nut and bolt is then hung up on a thread and tapped gently with a piece of wood and its resonant frequency measured with the frequency analyzer program. The frequency is fed into the program using a microphone. All of the inner pipes are tuned to exactly the same frequency by a very slight alteration of the insertion length of the bolt head for any pipe with a resonant frequency which is slightly off the frequency of the other pipes in the set of nine inner pipes.

Next, the outer tubes are slotted to raise their resonant frequency to match that of the inner pipes. Their frequency is also measured by hanging them up and tapping them gently with a piece of wood. If the frequency needs additional raising, then the tube length is reduced by a quarter of an inch (6 mm) and the testing continued as before. Adjusting the width and length of the slot is the best method for adjusting the resonant frequency of the tube. A small file can be used to increase the slot dimensions. This procedure is time consuming and tedious but it is well worth the effort. The average finished length of the outer pipes is 17.5 inches (445 mm) and the slot dimensions 0.75 inch long and 0.5 inch wide (19 mm x 13 mm).

When setting the resonant frequency of the outer pipes, it is important to have the clips in place. These "hosepipe", "jubilee" or "Terry" clips are used to make electrical connections to the outer pipes as shown in the diagrams and they do have an effect on the resonance of the pipes, so fit them before any tuning is done. The pipe arrangement is shown here:

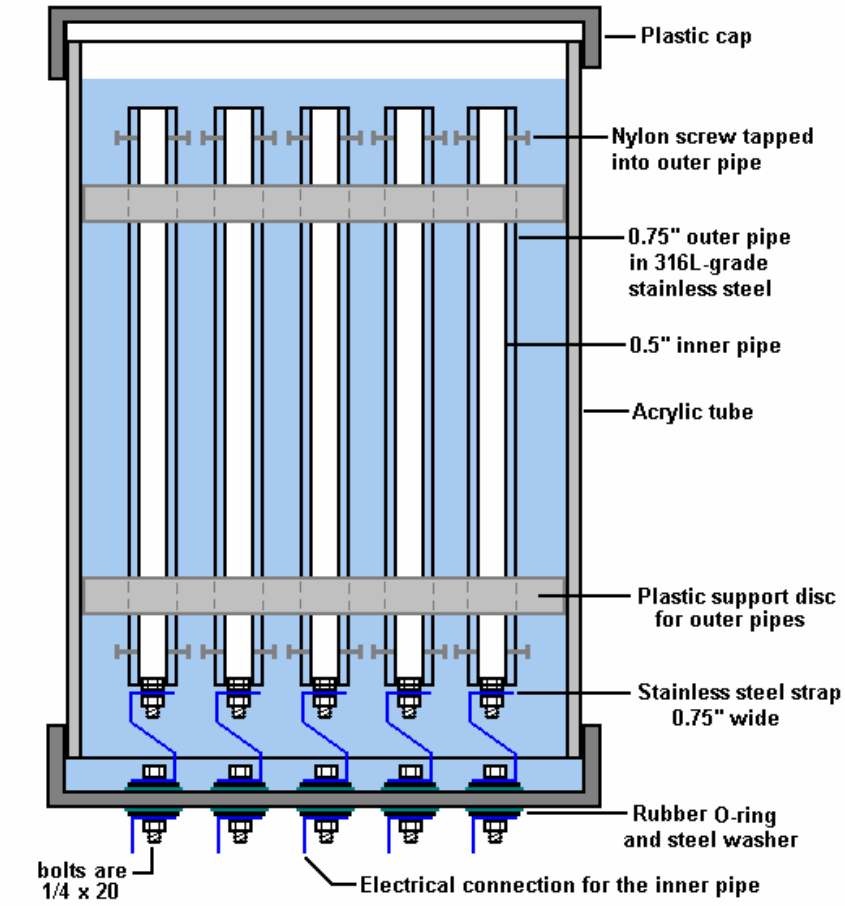


The outer pipes are drilled and tapped to take either a 6/32" nylon bolt available from Ace hardware stores in the USA, or alternatively, drilled and tapped to take a 4 mm nylon bolt. Three of these bolt holes are evenly spaced around the circumference of each end of all of the outer pipes.



These nylon bolts are used to adjust and hold the inner pipe **gently** in the exact centre of the outer pipe. It is very important that these bolts are not over tightened as that would hinder the vibrations of the inner pipe. The bolts are adjusted so that a feeler gauge shows that there is exactly the same 1.2 mm gap all round, both top and bottom. The weight of the inner pipe is carried by a 3/4 inch (18 mm) wide strip of stainless steel bent into a Z-shape, and **none** of the weight is carried by the nylon bolts. Dr Cramton describes this Z-shaped steel strip as a

"spring" and stresses its importance in constructing a set of resonating pipe pairs. The arrangement is shown here:



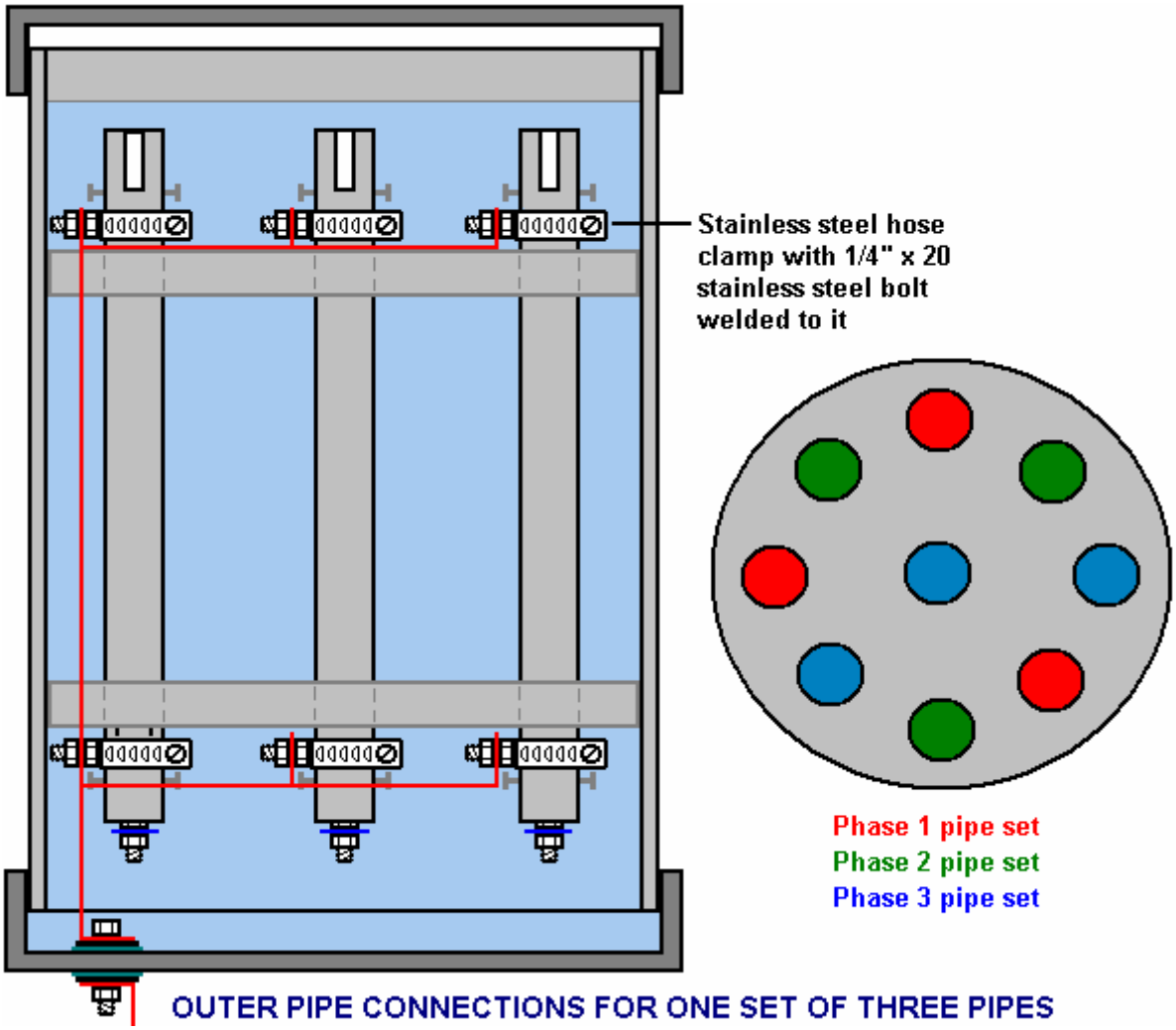
The supporting springy strip of steel is shown in blue in the above diagram as it also forms the electrical connection for the inner tubes. The outer tubes are held securely in position by two plastic discs which form a tight push-fit inside the 6" (150 mm) diameter acrylic tube which forms the body of the cell. The cell is sealed off with plastic caps (ideally, the upper one being screw threaded for easy maintenance) and the electrical connections are carried through the lower cap using 1/4" (6 mm) x 20 stainless steel bolts. The bolts are sealed using washers and rubber O-rings on both sides of the cap.

For clarity, the diagram above shows only the electrical connections for the inner pipes. The electrical connections for the outer pipes are shown in the following diagram. The connections are made at both the top and the bottom of each outer pipe by attaching a stainless steel hose clamp with a stainless steel bolt attached to each clamp. The wiring is then carried across inside the cell so that all six connection points (three at the top plus three at the bottom) for each set of three pipes are carried out through the base of the cell with just one bolt, again, sealed with washers and rubber O-rings. The nine pipe pairs are electrically connected in three sets of three, and each set is fed with a separate phase of a 3-phase waveform. This sets up an interaction through the water and produces a complex pulsing waveform with each set of pipes interacting with the other two sets. The sets are arranged so that the individual pipes of each set are interspersed with the pipes of the other two sets, making the sets overlap each other as shown in the next diagram. For clarity, the diagram does not show the electrical connections for the inner pipes and it omits the pipes of the other two groups of three, the water-level sensor, the gas take off pipe and the gas-pressure sensor.

At this time, Dr. Cramton is driving the pipe arrays with the circuit shown below. It uses an AC sinewave generated by a pulsed alternator. The current fed to the motor driving the alternator accounts for about 24 watts of power while the current drive to the alternator winding is just 12 watts. It should be realised that the alternator can easily drive many cells, probably without any increase in power required. Dr. Cramton is investigating methods of producing the same waveform without the need for an alternator and while that would be useful, it should be realised that a gas output of six litres per minute for a power input of only 36 watts is a very significant result. Others have shown that it is possible to power a 5.5 kilowatt electrical generator on hydroxy gas alone with a flow rate of this sort of magnitude, and obviously, the 36 watts can very easily be provided from that 5.5 kilowatt output.

It is absolutely essential that the pipe pairs are "conditioned" as there will be very little gas production until the white conditioning layer is built up on the active surfaces of the pipes. As has already been described, one

method is by powering the cell up for a few minutes, and then letting it rest unused for a time before repeating the process. Dr. Cramton emphasises that at least a hundred hours of conditioning will be needed before the gas output volume starts to rise, and it will be three months before the white conditioning layer reaches its full thickness and the gas production rate increases dramatically.

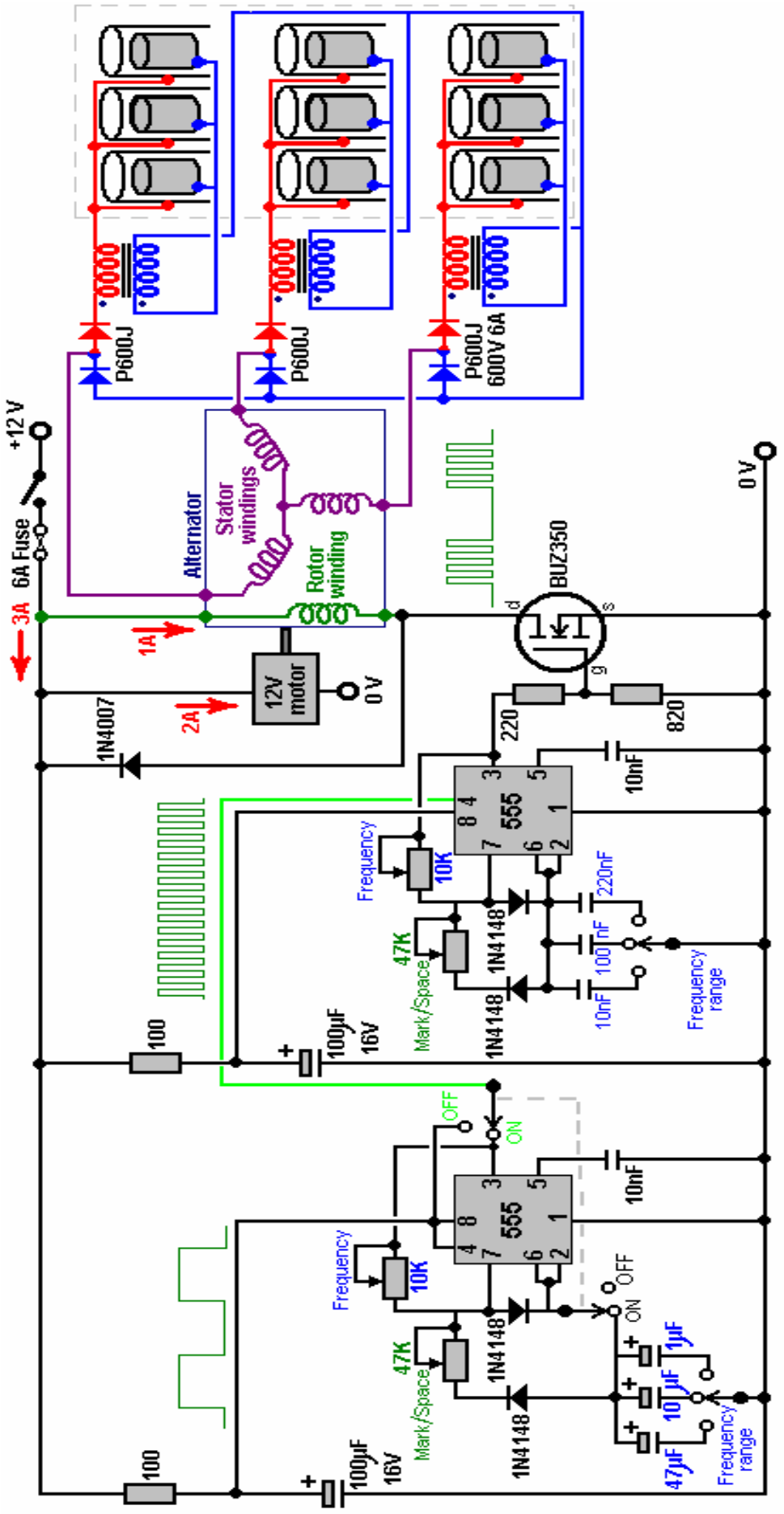


Dr Cramton stresses that it is the mechanical construction which will make the difference in the gas production rate. The inner and outer pipes must be tuned to a common frequency. It is vital that the pipe pairs must be conditioned, which can be done through repeated use over a period of time. A very important alternative to this long conditioning process is coating the whole of the pipe surfaces with the insulating material "Super Corona Dope" (http://www.mgchemicals.com/products/4226.html) as this gives immediate conditioning of the pipes. When a complete set of tuned tubes has been achieved, then the electronics must be built and tuned to the resonant frequency of the tube sets. Voltage builds up on the pipes from the repeated pulsing of the low voltage circuit and the action of the bi-filar wound coils each side of each pipe set and allowed by the insulation of the pipes. With Super Corona Dope this voltage has been measured at 1,480 volts but with the insulating layer from a local water supply, that voltage is around 1,340 volts.

It should be understood that the bi-filar wound coil (that is, wound with two strands of wire side by side) generates very sharply rising, very short voltage spikes, typically in excess of 1,000 volts in spite of the electrical supply being less than fourteen volts. The coils used by Dr Cramton are wound on ferrite rods, 300 mm (11.8”) long and 10 mm (3/8”) in diameter. As only 100 mm long rods were available, these were constructed by placing three inside a plastic tube. The coil winding is of enamelled copper wire and to allow sufficient current carrying capacity, that wire needs to be 22 swg (21 AWG) or a larger diameter, that is, with a lower gauge number such as 20 swg. These coils are wound to give an inductance of 6.3 mH on each of the two windings.

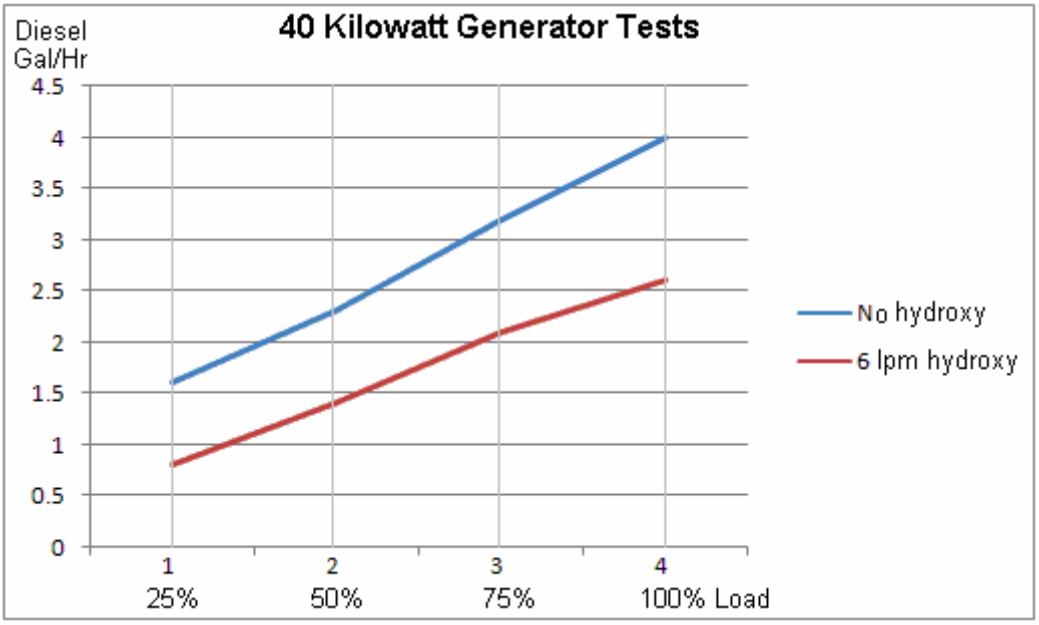
The circuit below is the one being used at this time. You will notice that an additional pole has been added to the Gating On/Off switch so that the timing components are switched out when the gating signal is turned off. This gives added protection for the Gating 555 chip in the circuit, preventing overheating when it is running but not being used. The frequency used with Dr. Cramton’s cell is 4.73 kHz although this is not the optimum frequency for the cell. The alternator imposes a certain limitation on the highest possible frequency, but the frequency used

has been shown to be the most effective and it is a harmonic of the optimum frequency. This is a bit like pushing a child on a swing and only pushing every third or fourth swing, which works quite well.



Dr. Cramton says: “I would like people to know that the scientific community is working on these projects and this technology is now a fact of science and not conjecture”.

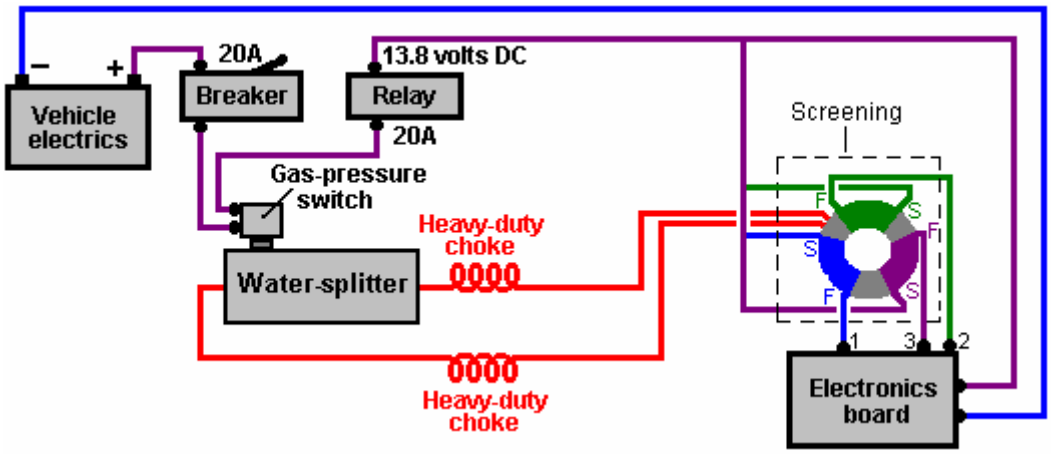
Dr Cramton has performed repeated performance tests on a 40 kilowatt diesel generator and the results were highly consistent, coming in within 1% each time on ten successive tests. Here is his graph of the results of this preliminary work:



The gains at full 40 kW load are about 35%, representing a reduction of 1.4 gallons of diesel per hour. As the generator is part of the equipment of a major power supplier, it is likely that the number of generators will be manipulated in relation to the demand and so the continuous overall gain is likely to be about 33% even with such a low hydroxy input as 6 lpm. The investigation and development is continuing.

Bob Boyce’s High-Gain Toroidal System.

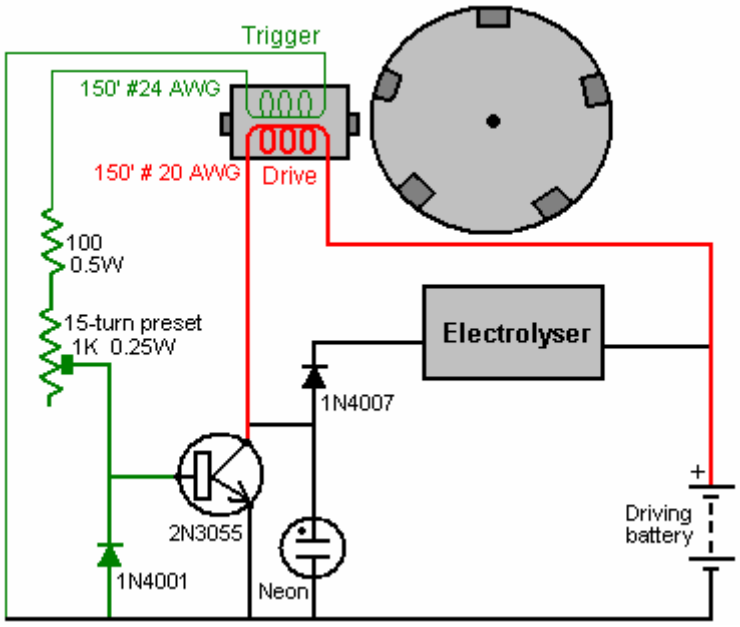
Bob Boyce has recently released a different method for water-splitting using his flat-plate electrolyser-style construction and pulsed with just twelve volts as in the above water-splitter designs. Bob's circuit is:



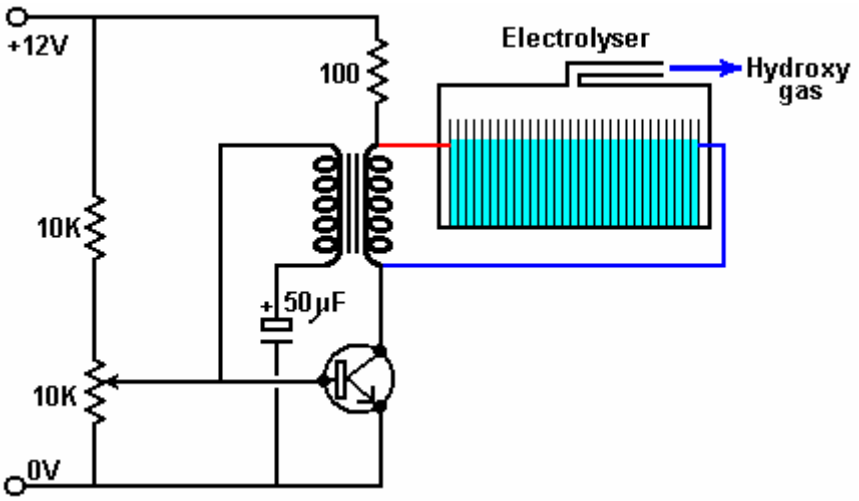
Here, the electronics board produces three separate, tuneable, very sharp square waveforms as described in the D9.pdf document mentioned earlier. These three waveforms are integrated into a single complex waveform when each is fed into a separate high-precision, high-specification winding on an iron-dust toroidal transformer core. This signal is stepped up to a higher voltage in the secondary coil of the transformer and then applied to the electrode plates via a choke coil on each side of the unit in exactly the same way as in the previous designs.

Resonance.
Water-splitters only operate properly if are held on their resonant frequency. Stan Meyer has a patent on his electronics system which would locate, lock on to and maintain the electronic pulsing at the resonant frequency of his cell. Unfortunately, Stan's patent just gives broad outlines for the methods used.

The John Bedini battery-charging pulse circuits have been very successfully applied to water-splitter cells. Here, the cell itself is part of the frequency control of the oscillator circuit and the arrangement might look like this:



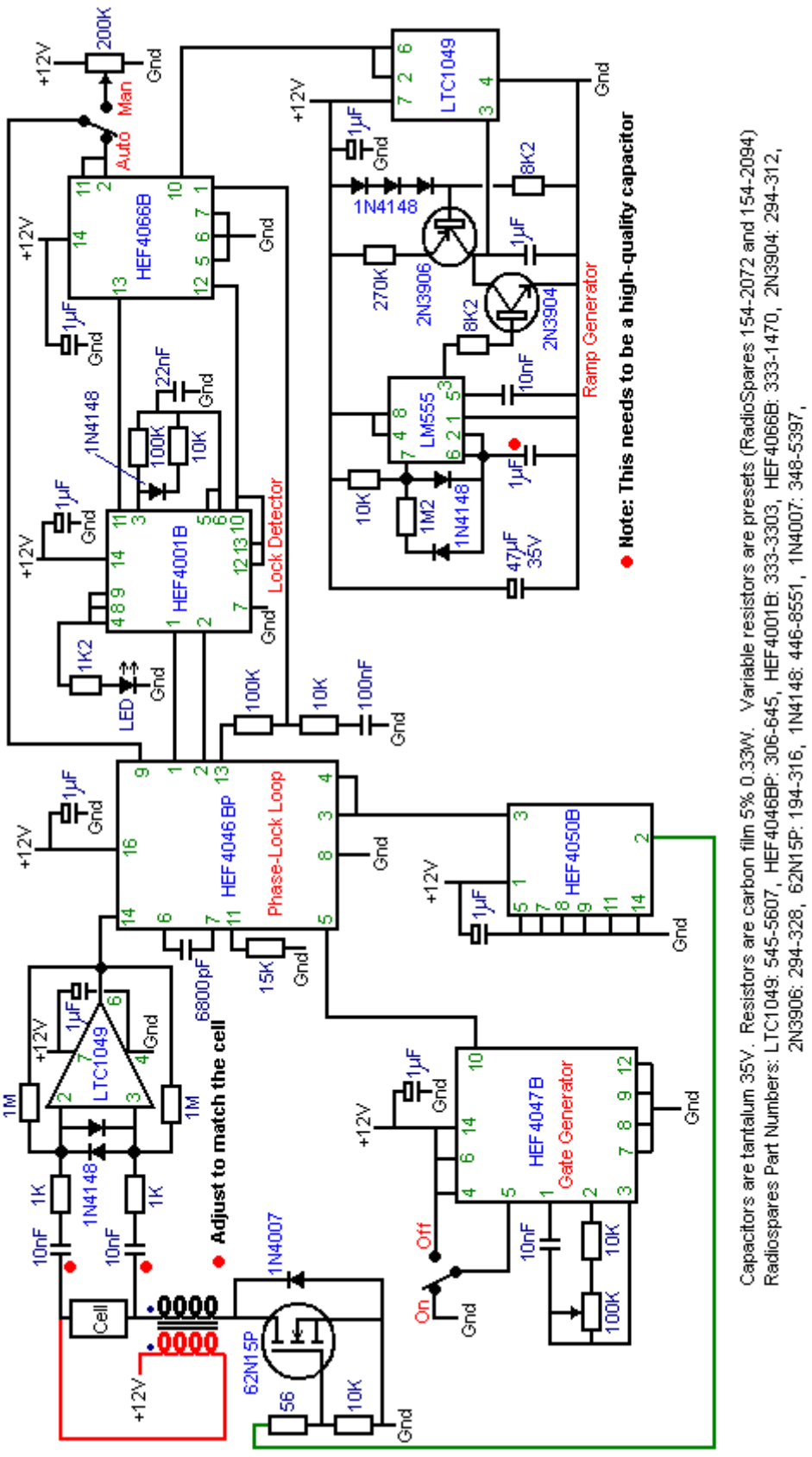
This idea is advocated on a YouTube video put up by a user whose ID is "TheGuru2You" where this arrangement is suggested:



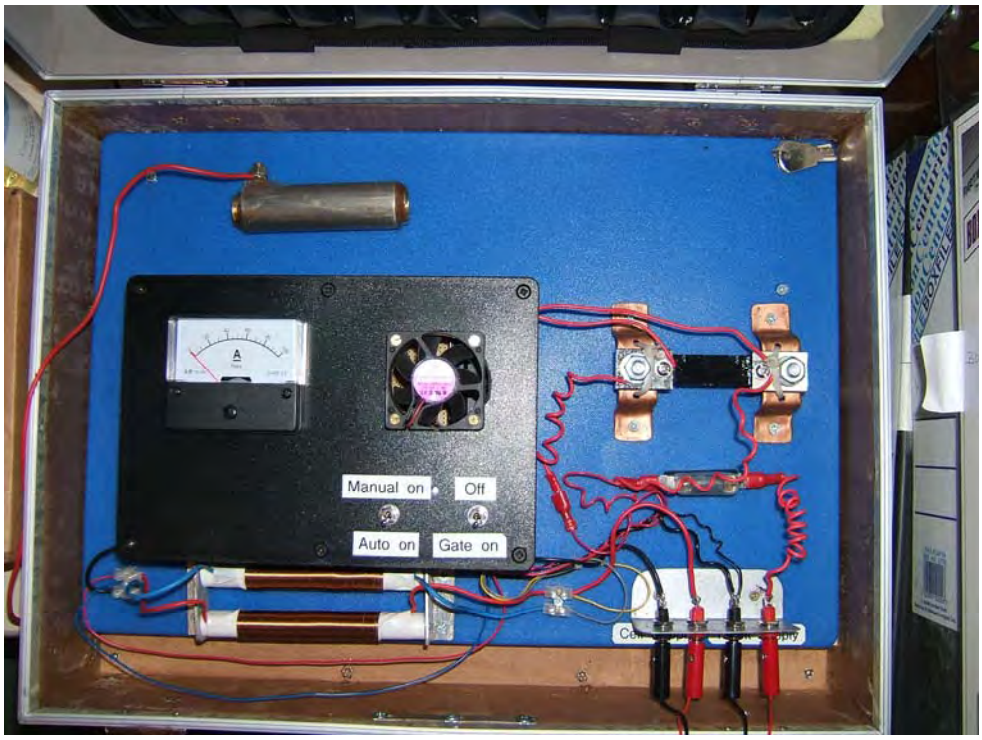
TheGuru2You states that he has built this circuit using a capacitor instead of the water-splitter and he says that he can confirm that it is self-powering, something which conventional science says is impossible (unless perhaps, if the circuit is picking up radiated power through the wiring of the circuit). Once a twelve volt supply is connected briefly to input terminals, the transistor switches on powering the transformer which feeds repeating pulses to the base of the transistor, sustaining the oscillations even when the twelve volt supply is removed. The rate of oscillation is governed by the resonant frequency of the water-splitter unit. Consequently, as the resonant frequency of the cell alters because bubbles form, the pressure changes, the temperature changes, or whatever, the circuit automatically tracks and maintains that optimum frequency.

Dave Lawton's Auto-Tune Circuit.

Dave Lawton uses a different method as he has designed and built a Phase-Lock Loop ("PLL") circuit which does the same thing that Stan Meyer's automatic circuit did. This is Dave's circuit:



This circuit has been used very successfully by a number of people. One experimenter had the circuit built by a friend as he is not very confident with building electronic circuits. The construction looks like this:



The two air-core coils are wound separately rather than bi-filar wound, and some experimentation with different types will be undertaken to see the effect on overall gas production. This circuit is shown in the following video, driving a 2.6 inch long pair of electrodes with a 2 mm gap between them, sitting in a test cell. The electrodes have seams and are made of an unknown quality of stainless steel and can be seen at the top of the photograph above. The video: http://youtu.be/XMizRAYdGwA shows considerable gas production with almost no current draw and the cell staying completely cool.

Running Electrical Generators on Water Alone.

After many decades of being lied to, most people believe that it is necessary to burn a fuel (typically, a 'fossil fuel' such as petrol or diesel) in order to make an engine run. 'Scientific experts' demonstrate their ignorance by proclaiming that their calculations show that there is just not enough energy in hydrogen released through electrolysis, to provide enough power to run an engine which can provide sufficient electrical energy to perform the electrolysis in the first place.

Their calculations are completely wrong as they are based on a major level of ignorance of the real facts:

1. Ignoring HHO altogether, engines can run extremely well on environmental energy channelled through a Joe Cell as shown in chapter 9, and when doing that, no fuel at all is consumed..

2. They are not aware that properly made HHO has typically four times the energy content of hydrogen gas.

3. They are not aware that a properly built electrolyser running on DC has more than double the efficiency that Faraday considered to be the maximum possible production rate of HHO for any given current flow.

4. They are not aware that resonant pulsed electrolysis has several times the water-to-HHO conversion efficiency that straight DC can produce, resulting in more than ten times the Faraday 'maximum' conversion rate.

5. They are not aware that the majority of the energy produced by HHO being converted back into water does not come from the hydrogen but instead comes from charged water clusters. It is likely that they have never even heard of charged water clusters.

6. They are probably not aware that introducing cold water mist to the air entering an internal combustion engine, makes that engine operate as an internal combustion steam engine as the mist gets converted into flash-steam, raising the pressure inside the cylinder and boosting the engine efficiency very considerably.

Because of these things, the calculations of the 'scientific experts' produce completely wrong results due to lack of knowledge and the flawed assumptions on which the calculations are based.

Interestingly, quite apart from the fact that generators running on water as the only visible fuel, have been powering off-grid locations 24 x 7 for many years now, and ignoring that little detail, check out this newspaper comment:

US Navy will turn seawater into fuel

THE US Navy claims it has found a way to turn sea-water into fuel, having spent decades conducting scientific experiments.

If true, the development could see military ships developing their own fuel and staying operational 100% of the time. Navy scientists say they have already used the new fuel to fly a model aircraft.

Calling it "a huge milestone", Vice Admiral Philip Cullom pointed out that "in the Navy, we have some pretty unusual kinds of challenges."

One of these is the need at present to refuel from an oil-tanker. But "developing a game-changing technology like sea-water-to-fuel reinvents a lot of the way we can do business when you think about logistics, readiness," said Cullom.

The US Navy has 289 vessels and most rely on oil-based fuel, apart from some aircraft carriers and 72 submarines that use nuclear power.

The breakthrough came after scientists found a way to extract carbon dioxide and hydrogen gas from seawater. The gasses are then turned into a fuel with the help of catalytic converters.

"We are in challenging times where we have to think in new ways to look at how we create energy, how we value energy and how we consume it," said Cullom.

"We need to challenge the assumptions that of the last six decades of constant access to cheap, unlimited amounts of fuel."

This is a very interesting comment from Vice Admiral Cullom. Firstly, he confirms that the Navy's massive engines can be powered by gas produced by electrolysis. Second, he implies very strongly that refuelling with oil-based consumables is no longer needed, and that means that the electrolysis is well in excess of 100% efficient, making those massive engines self-powered. Third, it seems reasonable to presume that if those massive engines can be self-powered through electrolysis, that the very much smaller engines in cars could also be run on electrolysis, even if the same method of electrolysis were not used. Anyway, we have to deal with the tiny, very inefficient motors which are used in generators which normally use petrol as the fuel:

In the UK, three men managed to run a generator on water alone, using just simple things which are within the scope of the average handyman in his workshop. They bought a standard petrol-driven electrical generator off eBay and managed to run it without using any petrol. They used a hydroxy gas flow which they measured at just 3 lpm and they test loaded the 5.5 kilowatt generator with 4 kilowatts of equipment. Afterwards they abandoned the generator and moved on to a much bigger engine as their plans are to sell electricity to the local power company. Their representative said: the equipment was put together by my associate, who supplied the water/electrolyte (not a standard electrolyte, 0.4% by volume). It was a 5 kW petrol generator (about 300cc). We attached a Chevrolet alternator which constantly charges a 12-volt 55 Amp-Hour battery, which in turn powers six HHO tubes, each of which draws 6 watts, for a total of 30 watts. The outputs of these electrolysers are connected in series and they feed the gas into a low-pressure butane gas "camping" tank which has an 18 psi pressure release valve. This tank feeds the air intake of the generator which is adjustable with a choke. We loaded the

30-amp socket of the generator with various drills, heaters etc. working, for more than four hours. The maximum load which we tried was 4 kilowatts, being a bar heater, a kettle and two drills. The electrolyser tubes are heavy-duty plastic (rated for 80 psi pressure). Inside are 4 tubes of stainless steel (3 positive and 1 negative). Each tube created 1 litre of gas every two minutes which is a total of 3 litres per minute. They eventually become warm to the touch, but they do not get hot.

Running an Electrical Generator without Fossil Fuel

In Broad Outline

In order to achieve this objective, very much like Stan Meyer, we need to feed the engine three things:

1. Air - this is fed in as normal through the existing air filter.
2. Hydroxy gas - how to make this has already been explained in considerable detail.
3. A mist of very small water droplets, sometimes called "cold water fog".

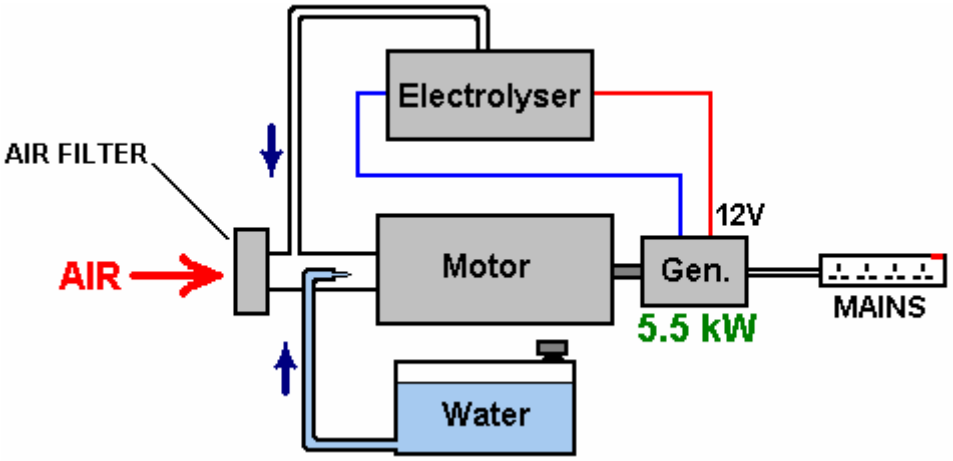
Also, we need to make two adjustments to the engine:

1. The spark timing needs to be retarded by about eleven degrees.
2. If there is a "waste" spark, then that needs to be eliminated.

To summarise then, a good deal of work needs to be done to achieve this effect:

1. An electrolyser needs to be built or bought, although the required gas production rate is not particularly high.
2. A generator of cold water fog needs to be made or bought.
3. Pipes need to be installed to carry these two items into the engine.
4. The engine timing needs to be retarded.
5. Any waste spark needs to be suppressed.
6. Water tanks are needed for the cold water fog and to keep the electrolyser topped up.
7. Ideally, some form of automatic water refill for these water tanks should be provided so that the generator can run for long periods unattended.

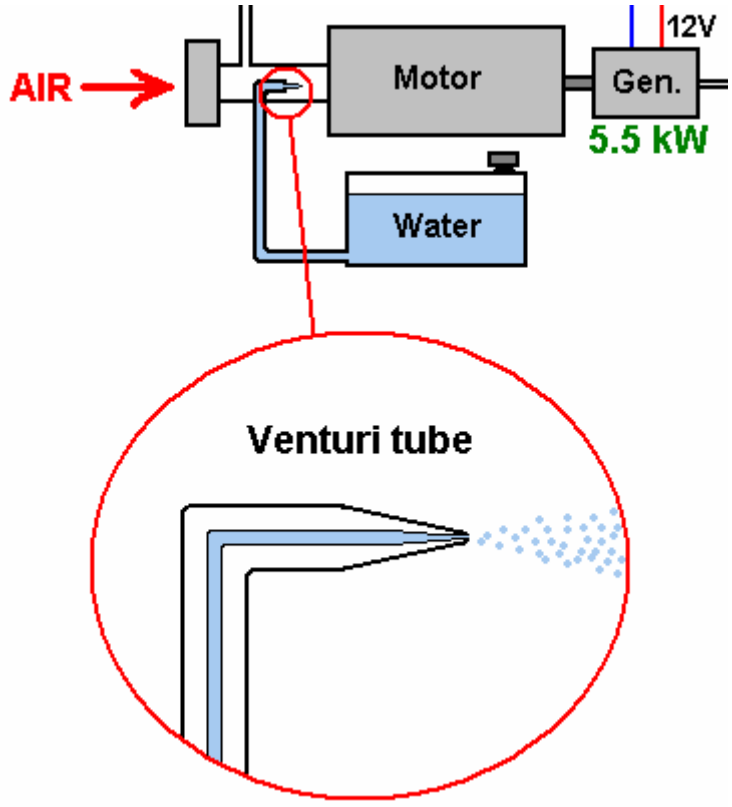
If we omit the electrical safety equipment which has already been explained in detail, and omit the hydroxy gas safety equipment which has already been explained in detail, and skip the automated water supply details and the starting battery, then, a generalised sketch of the overall arrangement looks like this:



Here, they have opted to feed the hydroxy gas into the air system after the air filter (a thing which we normally avoid as it is not helpful for the hydroxy gas production efficiency, but the first step is to reproduce their successful method exactly before seeing if it can be improved further). Also fed into this same area is the cold water fog which is comprised of a very large number of very tiny droplets. The air enters this area as normal, through the existing air filter. This gives us the three necessary components for running the generator engine without using any fossil fuel.

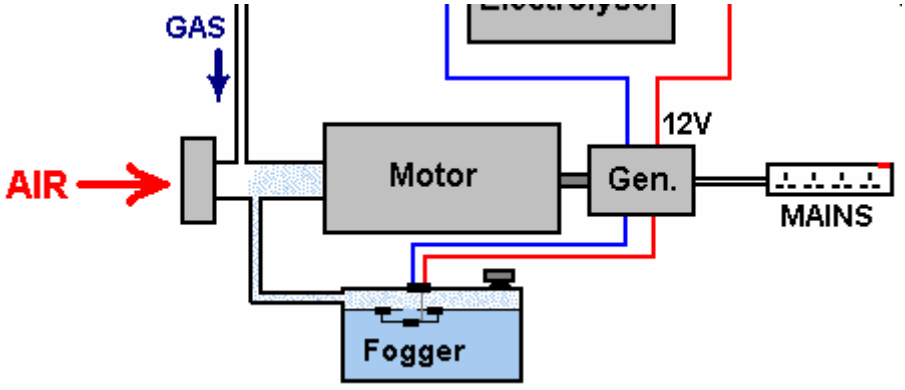
Creating the cold water fog

There are three different ways to generate the spray of very fine water droplets which are a key feature of the success of this way of running the engine. One way is to use a Venturi tube, which, while it sounds like an impressive device, is actually very simple in construction:



It is just a pipe which tapers to a point and which has a very small nozzle. As the engine draws in the air/hydroxy mix on it's intake stroke, the mixture rushes past the nozzle of the Venturi tube. This creates an area of lower pressure outside the nozzle and causes water to exit through the nozzle in a spray of very fine droplets. Some perfume spray bottles use this method as it is both cheap and effective.

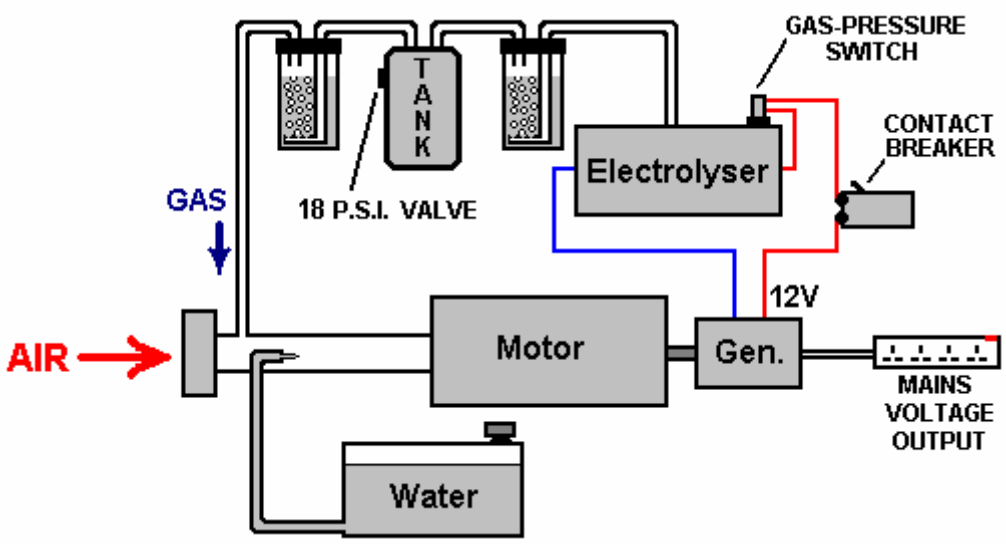
An alternative method of making the cold water fog is to use one or more "pond foggers". These are small ultrasonic devices which are maintained at the optimum operating depth in the water by a float. They produce large amounts of cold water fog which can be fed into the engine like this:



A third method is to use a small carburettor of the type used with model aircraft. This does the same job as a regular engine carburettor, feeding a spray of tiny water droplets into the engine air intake. The physical arrangement of this option depends on the construction of the air filter of the generator being modified. You will notice that the people in the UK who did this, used a small gas tank with an eighteen pounds per square inch pressure release valve. This is not possible with the highest quality of HHO gas as it cannot be compressed that much. However, with a lower grade of HHO which has some water vapour mixed in with it, it is possible to have a gas reservoir with that sort of pressure in it. In this case, except possibly for starting, their gas production rate is probably not high enough to allow much raised pressure inside the tank. Obviously, the gas-pressure switch on the electrolyser and the one on the gas storage tank will have similar operational pressures.

Some Safety Features

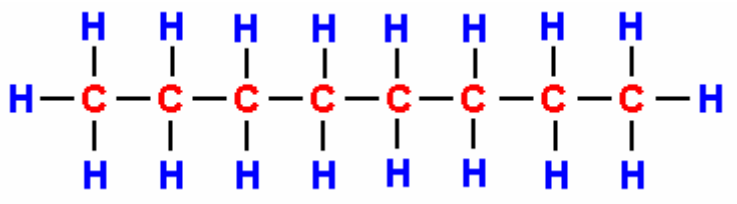
Up to this point, the electrolyser has been shown in bare outline. In practice, it is essential that some safety features are incorporated as shown here:



These safety devices should be familiar to you by now as they have already been explained earlier in this document.

The Reason for Changing the Timing

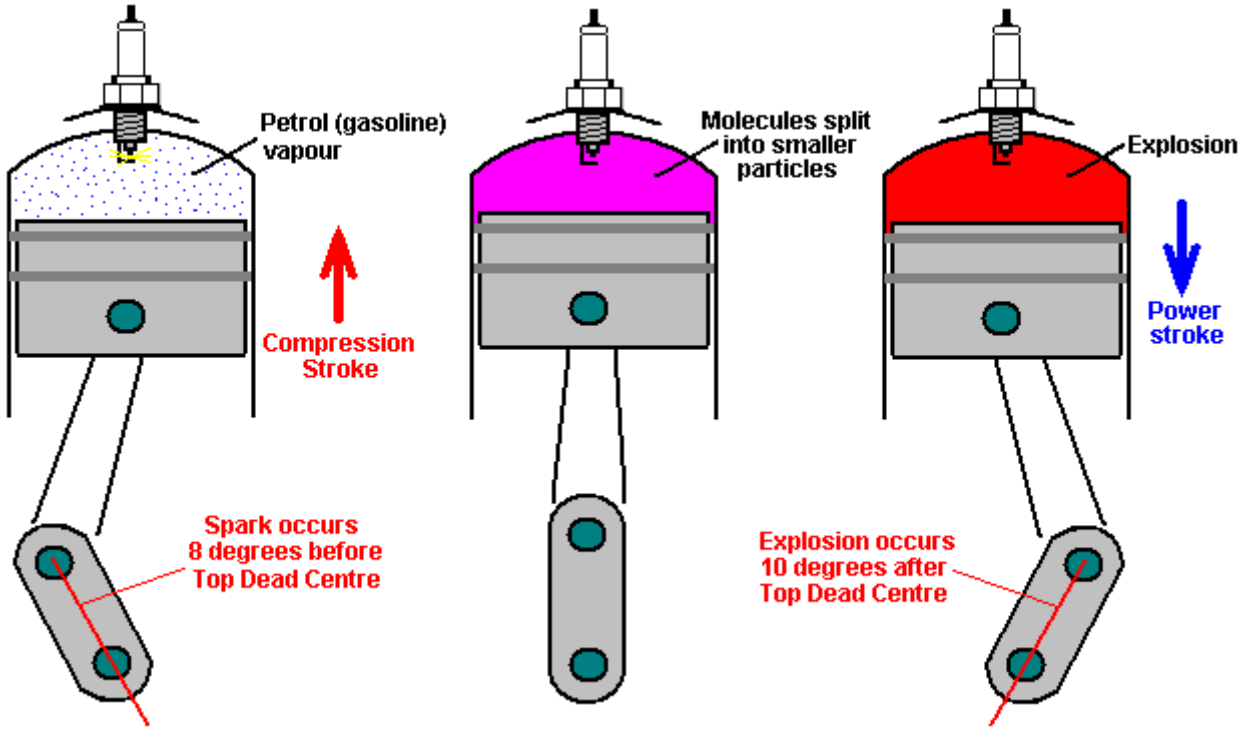
The fuels used with most internal combustion engines are either petrol (gasoline) or diesel. If you are not interested in chemistry, then you are probably not aware of the structure of these fuels. These fuels are called "hydrocarbons" because they are composed of hydrogen and carbon. Carbon has four bonds and so a carbon atom can link to four other atoms to form a molecule. Petrol is a long chain molecule with anything from seven to nine carbon atoms in a chain:



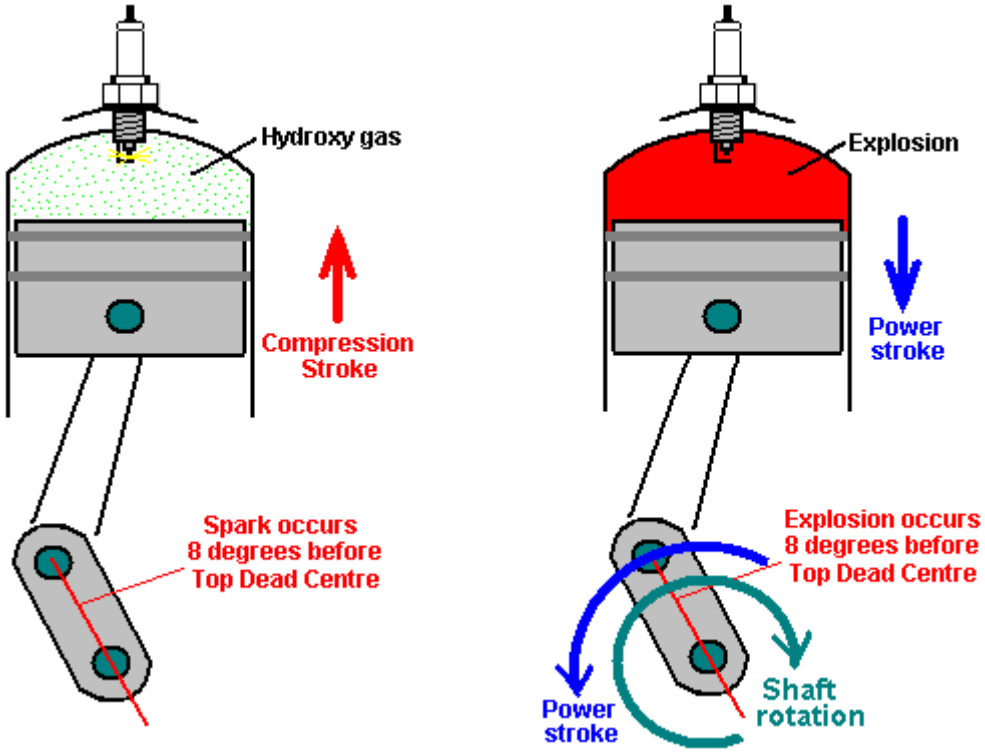
Diesel has the same structure but with eleven to eighteen carbon atoms in a chain. In a petrol engine, a fine spray of petrol is fed into each cylinder during the intake stroke. Ideally, the fuel should be in vapour form but this is not popular with the oil companies because doing that can give vehicle performances in the 100 to 300 mpg range and that would cut the profits from oil sales.

The petrol in the cylinder is compressed during the compression stroke and that reduces its volume and raises its temperature substantially. The air/fuel mix is then hit with a powerful spark and that provides enough energy to start a chemical reaction between the fuel and the air. Because the hydrocarbon chain is such a large molecule, it takes a moment for that chain to break up before the individual atoms combine with the oxygen in the air. The main engine power is produced by the hydrogen atoms combining with oxygen, as that reaction produces a large amount of heat. The carbon atoms are not particularly helpful, forming carbon deposits inside the engine, not to mention some carbon monoxide (CO) and some carbon dioxide (CO_2) as well.

The key factor here is the slight delay between the spark and the combustion of the fuel. The combustion needs to happen a few degrees after Top Dead Centre when the piston is about to start its downward movement in the power stroke. Because of the delay caused by the hydrocarbon chain breaking down, the spark occurs a few degrees **before** Top Dead Centre:



If you were to replace the petrol vapour with hydroxy gas, then there would be a major problem. This is because hydroxy gas has very small molecule sizes which do not need any kind of breaking down and which burn instantly with explosive force. The result would be an explosion which occurs far too soon and which opposes the movement of the rising piston as shown here:



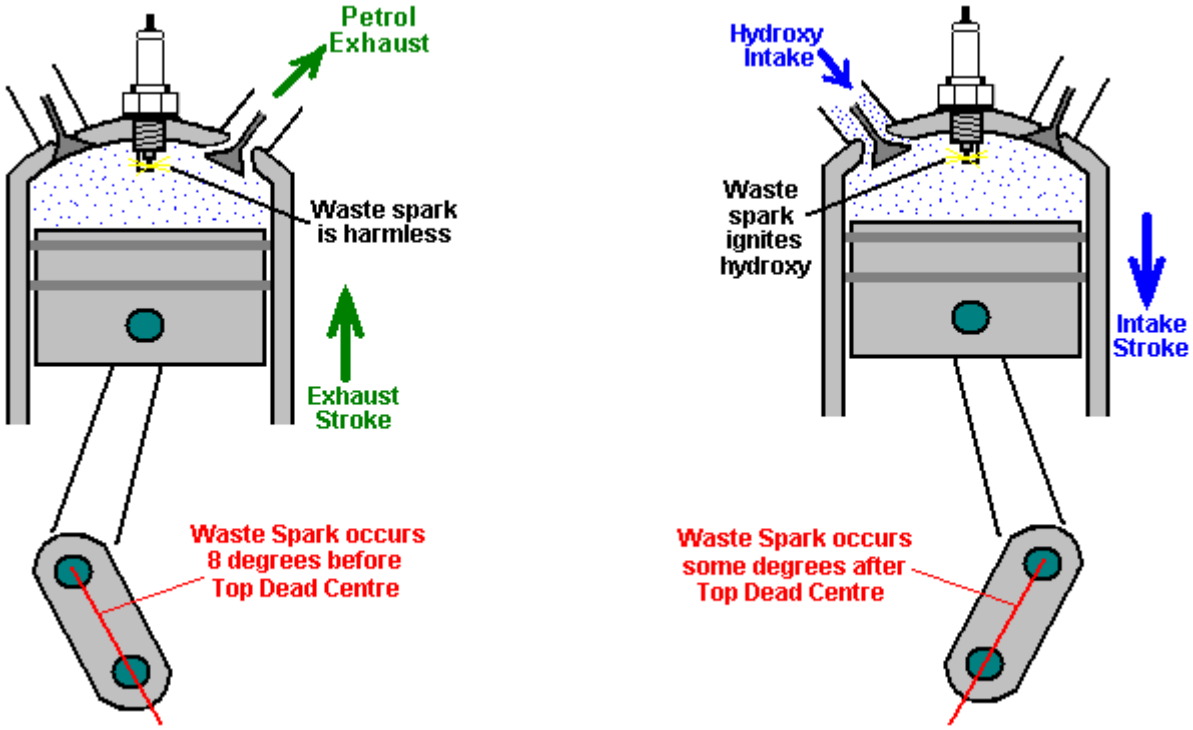
The forces imposed on the piston's connecting rod would be so high that it would be quite liable to break and cause additional engine damage.

In the case of our electrical generator, we will not be feeding it a mix of air and hydroxy gas, but instead, a mix of air, hydroxy gas and cold water fog. This delays the combustion of the hydroxy gas by a small amount, but it is still important to have the spark occur after Top Dead Centre, so the ignition of the generator needs to be retarded by eleven degrees.

Engine design varies considerably in ways which are not obvious to a quick glance at the engine. The timing of the valves is a big factor here. In the smallest and cheapest engines, the engine design is simplified by not having the spark timing taken off the cam-shaft. Instead, production costs are cut by taking the spark timing off the output shaft. This produces a spark on every revolution of the engine. But, if it is a four-stroke engine, the spark should only occur on the power stroke which is every second revolution of the output shaft. If the fuel is petrol, then this does not matter as the extra spark will occur near the end of the exhaust stroke when only burnt gasses are present in the cylinder.

Some people are concerned when they think of hydroxy gas burning and producing water inside the engine. They think of hydrogen embrittlement and rusting. However, because of the nature of the hydrocarbon fuel already being used, the engine runs primarily on hydrogen anyway and it always has produced water. The water is in the form of very hot vapour or steam and the engine heat dries it out when the engine is stopped. Hydrogen embrittlement does not occur as a result of using a hydroxy gas booster.

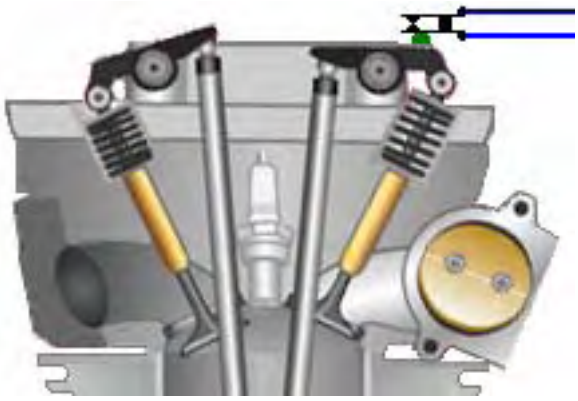
Anyway, if we were to delay the spark until after Top Dead Centre as we must, then the situation is quite different as the waste spark will also be delayed by the same amount. With most engines, at this point in time the exhaust valve will have closed and the intake valve opened. Our very flammable gas mix will be being fed into the engine on it's intake stroke. This means that our gas supply system is openly connected to the cylinder through the open intake valve, and so, the waste spark would ignite our gas supply system (as far as the bubbler which would smother the flashback). The situation is shown here:



we definitely do not want that to happen, so it is very important that we suppress that additional "waste" spark. So, this leaves us with two engine adjustments: timing delay and waste spark elimination. There are various ways in which these can be done and as each engine design is different, it is difficult to cover every possibility. However, there is a technique which can be used with many engines and which deals with both issues at the same time.

Most engines of this type are four-stroke engines with intake and exhaust valves, perhaps something like this:

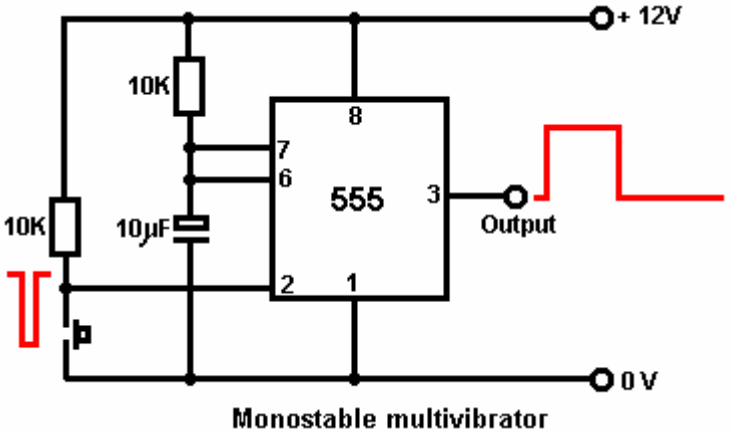
The intake valve (shown on the right in this illustration) is pushed down by a cam shaft, compressing the spring and opening the inlet port. The exact arrangement will be different from one engine design to the next. What is fixed is the movement of the valve itself and that movement only takes place every second revolution. There are various ways of using those movement to eliminate the waste spark and retard the timing. If a switch were mounted so that it opens when the intake valve opens and closes when the intake valve closes, then the switch closure shows when the piston starts upwards on its compression stroke and a simple electronic circuit can then give an adjustable delay before firing the coil which produces the spark. This, of course, involves disconnecting the original electrical circuit so that no waste spark is generated. The current flowing through the switch contacts can be arranged to be so low that there will be no sparking at the contacts when the circuit is broken again. The switch positioning might be like this:



An alternative is to attach a strong permanent magnet to the rocker arm, using epoxy resin, and then position a solid state "Hall-effect" sensor so that it triggers the delay before the spark is generated.

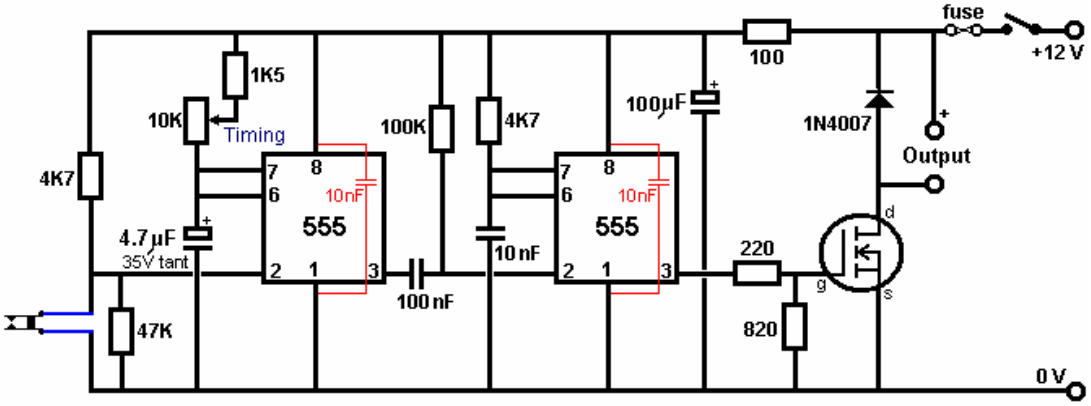
If the engine did not have a waste spark, then in theory, the timing mechanism of the engine could be used to retard the spark. However, in practice, the timing mechanism is almost never capable of retarding the spark to the position that is needed for running without fossil fuel, and so, some kind of delay circuit will be needed anyway.

The sort of delay circuit needed is called a "monostable" as it has only one stable state. A basic circuit of that type is:



Monostable multivibrator

If you are not at all familiar with electronic circuits, then take a look at the beginner's electronics tutorial found in the Chapter12.pdf document on the http://www.free-energy-info.co.uk website as that explains how circuitry works and how to build any simple circuit from scratch. We can use two of these circuits, the first to give the adjustable delay and the second to give a brief pulse to the ignition circuit to generate the spark:

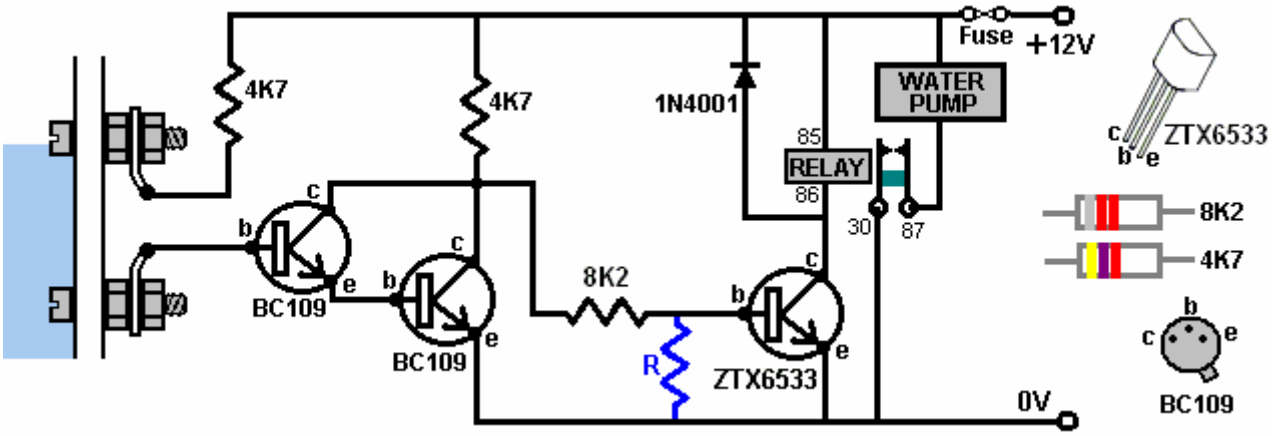


Making the hydroxy gas

When the generator is running, we have a ready supply of electrical energy, coming from a piece of equipment which has been specifically designed to supply large quantities of electricity for any required application. We are not dealing with the spare capacity of some low-grade alternator in a car, but we have substantial electrical power available.

Having said that, the electrolysers described at the start of this document are efficient and it is unlikely that an excessive amount of power would be needed when using one of those designs. Another convenient factor is that this is a stationary application, so the size and weight of the electrolyser is not at all important, and this gives us further flexibility in our choices of dimensions.

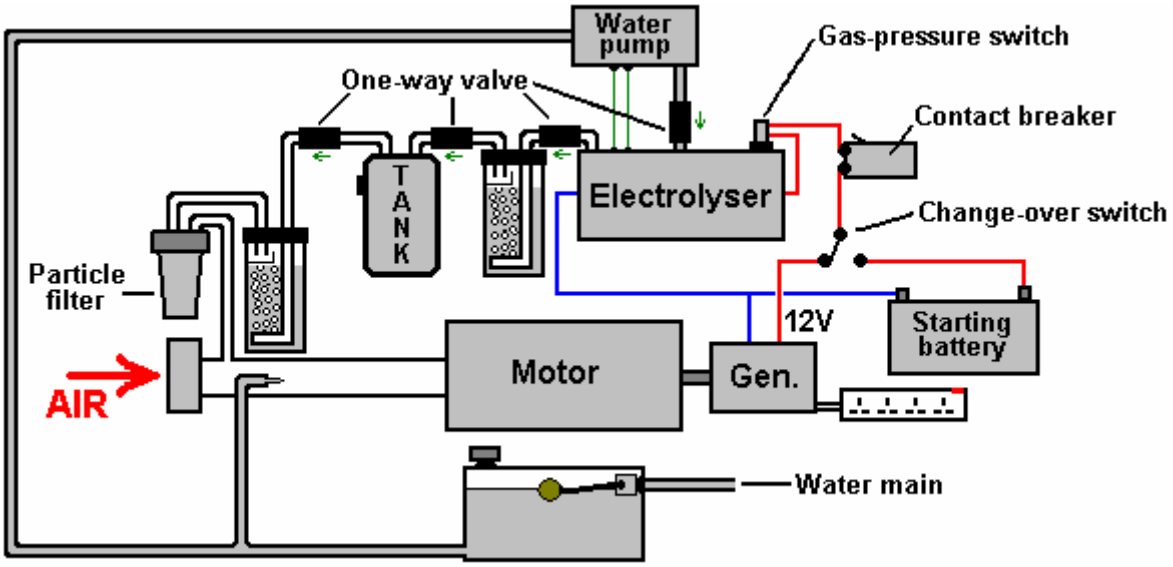
As this is an application where it is highly likely that the electrolyser will be operated for long periods unattended, an automated water supply system should be provided. The main details of such a system have already been covered, but what has not yet been dealt with is the switching for the water pump. The water pump itself can be an ordinary windscreen-washer pump, and we need some form of switch which operates on the electrolyte level inside the electrolyser. It is sufficient to sense the level in just one of the cells inside the electrolyser as the water usage will be pretty much the same in every cell. If you make the electrolyser in a suitable size or shape, then a simple off-the-shelf miniature float switch can be used. If you prefer, an electronic level sensor can be operated, using two bolts through the side of the electrolyser as the level sensor. A suitable circuit for this simple switching task could be:



When the electrolyte level inside the electrolyser is in contact with the upper bolt head, the circuit is switched off and the water pump is powered down. The electrolyte has a low resistance to current flow, and so it connects the 4.7K resistor through to the base of the BC109 Darlington pair (as described in Chapter 12). This keeps the two transistors switched fully on which keeps the 8.2K resistor connection well below the 0.7 volts needed to switch the ZTX6533 transistor on. If you are concerned about the ZTX6533 transistor being partially on, then resistor "R" could be added, although the prototype did not need one. The value would be about 2K. When the electrolyte level falls below the upper bolt head, the first two transistors switch off, and the ZTX6533 transistor is then powered fully on by the 4.7K resistor and the 8.2K resistor in series, providing the 150 mA needed for the relay to be switched fully on. The circuit draws about 5 mA in it's standby state. The numbers on the relay symbol correspond to the numbers on a typical automotive 12 volt relay. Using two BC109 transistors as the front end allows this circuit to be used with tap water if you wish. However, the water-level control for the water supply to the pond fogger or Venturi tube misting device does not need any form of fancy mechanism. The standard ball-cock valve mechanism which is used with toilets is quite adequate, especially if a floating pond fogger is being used as it maintains its own optimum depth below the surface and so the overall depth is not in any way critical provided, of course, there is sufficient depth for the fogger to float correctly.

Starting:

When left for any length of time, the gas pressure inside the electrolyser will drop because the nature of the hydroxy gas alters. This means that there will not be sufficient hydroxy gas available to start the engine and no more gas will be generated until the engine drives the generator. So, to deal with this situation, a lead-acid car battery is included so that it can be switched in to replace the generator for a brief period before the engine is started. That inclusion gives this overall arrangement:



This arrangement is perfectly capable of running a standard generator without the use of any fossil fuel. It should be noted that while no fossil fuel needs to be bought to run this generator system, the electrical output is far from free and is actually quite expensive as there is the purchase cost of the generator, the electrolyser and the minor additional equipment. Also, generators have a definite working life and so will need to be refurbished or replaced.

It might also be remarked that if a generator of this type is going to be used in an urban environment, then the addition of sound-reducing baffles and housing would be very desirable. At this point in time I am aware of nine different electric generators which have been adapted to run on water. At least four of these are from different manufacturers. The method of altering the timing and dealing with the waste spark is different from one adaption to the next. One user has altered the spark timing of his generator to after Top Dead Centre by rotating the timing disc to a position not envisaged by the manufacturer. The timing disc is held in place by a locking ("key") bar which fits into a channel cut in the shaft of the engine, matching it to a similar channel cut in the disc. The alteration was achieved by cutting a new channel in the shaft, allowing the timing disc to be positioned further around the shaft, producing the required timing delay. This arrangement also makes the waste spark ineffective and so it can be ignored. While this method requires the cutting of a slot, it does away with the need for any electronics and it is a very simple solution.

If you feel that the construction of a suitable electrolyser would be a problem or that the amount of electric current needed to operate it would be excessive, let me show you the actual figures involved:

Michael Faraday was an exceptional and highly respected researcher who investigated the electric current needed to convert water into hydrogen gas and oxygen gas by electrolysis. His results are accepted by pretty much every scientist everywhere. While he expressed the results of his work in terms which would be meaningless to the average person, his result is that an electrical input of 2.34 watts produces one litre of hydroxy gas in one hour.

In practical terms, that means that a current of 0.195 amps at 12 volts will produce 1 litre of hydroxy gas in one hour. In passing, only a nearly discharged lead-acid battery would have a voltage of 12 volts as the fully charged state is 12.85 volts and a vehicle alternator produces about 14 volts in order to charge the battery.

It is easier then, to compare the gas output of electrolysers directly to the figures produced by Faraday as shown here, based on a gas output of 15 litres per minute which is 900 litres per hour:

Faraday: 900 litres in one hour, takes 2,106 watts or 100% Faraday
Boyce: 900 litres in one hour, takes 998 watts or 211% Faraday without pulsing
Boyce: 900 litres in one hour, takes 180 watts or 1,170% Faraday with pulsing
Cramton: 900 litres in one hour, takes 90 watts or 2,340% Faraday

Much of this is not very important as it has been demonstrated that a gas production rate of around 3 lpm (180 lph) is sufficient to run a generator which produces 5,500 watts. Let us assume that the measured figure is 100% wrong and that it takes 360 lph of hydroxy gas, plus cold water fog, plus air, to run the generator, then:

Faraday would need 843 watts

Boyce would need 400 watts without pulsing
Boyce would need 72 watts with pulsing
Cramton would need 36 watts

None of these figures are important for running a generator because with an electrolyser efficiency of only 50% Faraday still leaves a massive generator excess of nearly 4 kilowatts on a 5.5 kilowatt generator. The gain is in running a generator as an internal combustion steam engine and not in the great efficiency of the electrolyser. It is distinctly possible that the pessimistic figures shown above are twice what is actually needed, but who cares? - the facts speak for themselves, with several people scattered around the world, already running generators on water. Many different generator designs have been adapted, typically, by modifying the flywheel, filling in the keyway and cutting another one to give a spark 2 degrees after TDC. Experience has shown that the 6.6 kVA Honda V-twin petrol motor generator and the Vanguard V-twin work very well long-term when adapted to run on water only.

Wear and Tear Issues

A man who lives in Alaska is very experienced in the use of renewable energy sources and unconventional fuel systems. His experiences are likely to be helpful for anyone who intends to use an electrical generator, whether running on water or on a fossil fuel. He recalls the experiences of a friend:

He decided to live off-the-grid because it was going to cost him $20,000 to get connected to the grid and as his house was not that large, he decided to go the alternative route. We designed a system which would use a 4 kW inverter and have an 8 kW Briggs & Stratton generator with a 13 kW surge capacity, for back up. The system has 6 solar panels and a 24 volt battery bank with 400 amp capacity. Having long summer days here in Alaska, the solar panels have more than enough capacity for charging the battery bank on sunny days. However, but when the day is overcast or when it is winter when there are only six hours of sunlight, the battery bank does not get charged fully. At these times, the generator is used to top up the battery bank.

American generators normally have either two or four 120-volt outputs each rated at 15 amps, plus one 240-volt output rated at 33 amps. If one of the two 120-volt outputs is used to charge the battery bank, then you get left with just the other 120-volt output for any other power needs during the time when the battery bank is being charged. This is not a satisfactory arrangement as operating with one field at maximum power and the other one lightly loaded or unused, causes a field imbalance in the generator, engine crank imbalance and ring or regulator failure within six months. It also causes noisy running and excessive fuel consumption.

Run in this way, providing a 60-amp charge rate, the generator ran hard and loud for two to two and a half hours per day, and running it was costing $350 per month for gasoline. The generator failed after four months.

In order to balance the loading on the replacement generator, a 15 kVA step-down transformer costing less than $1000 was purchased so that the 240-volt output could be used to drive 120-volt equipment. A transformer to be used for this needs to have a power-handling capacity which is greater than the surge capacity of the generator. A major advantage is that the generator current is halved for any given level of equipment current drawn because the equipment is running at only half of the generator voltage.

Using this transformer made a massive difference, giving a balanced output and providing a 90-amp charging rate for the battery bank as well as having ample power to run other household equipment when the battery bank was being charged. The result was a charge time of just one hour twenty minutes per day, with the generator running quietly and smoothly. The fuel consumption also dropped to just $70 per month which is just one fifth of what it was, covering the cost of the transformer in under four months. This generator has been running now for two years without any problems at all.

The Step-by-Step Conversion of a Generator

Selwyn Harris of Australia has kindly agreed to share detailed information on how he performs the conversion of a standard electrical generator to enable it to run on water alone. The generator which he uses as an example for this tutorial is a GX4000i generator:

The supplier is AGR Machinery which is an Australian company on eBay which buys up stock from collapsed companies and resells the equipment. The supplier says: GX4000i portable type generators have smoother output power, comparable to public utility sources. Ideal for powering medium loads such as:

- Power tools - Both Single & 3 Phase
- Game consoles, Digital Cameras
- Laptops, Camcorders
- Lighting and Microwave Ovens
- Drills, Grinders
- Resistive Load Kitchen Appliances (i.e. Coffee Maker, Toaster)
- Emergency Home Back up power where 240v power is required

Also, these units are significantly quieter than others due to refined engine technology

Features:

- Commercial Grade Engine: 196 cc 4-stroke, 7 horsepower, overhead camshaft, T.D.I. ignition
- Maximum output 4.0 kVA at 240 or 415V AC (Rated output: 2.7 kilowatts)
- Quality Heavy-Duty Construction
- AVR (Auto Voltage Regulator)
- Three 240V and one 415V Protected Outlets
- 100% Pure Copper Core
- Gearless direct drive
- Robust Square frame Design
- Easy – Recoil Start
- Oil capacity: 0.7 litres

- Powder Coated Finish
- Light and Compact for easy manoeuvrability (38.5 Kg)
- Noise level: 69 dB

The first step of the conversion is to remove the fuel tank which is held in place with four bolts:

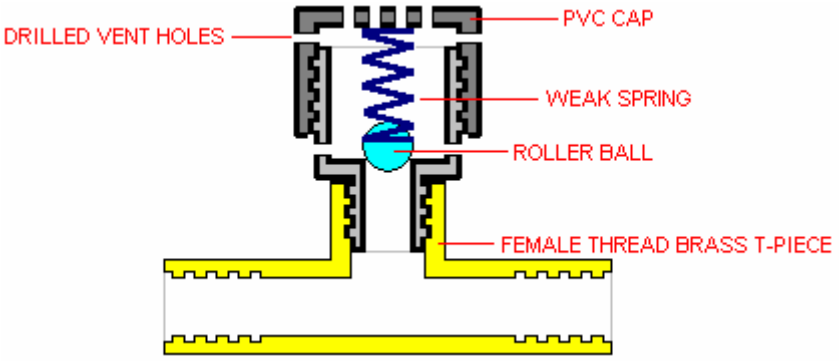
This allows access to the carburettor which is then removed as it will not be used:

The next step is to construct a pressure-release valve mechanism which will protect the equipment from damage in the unlikely event of a major, sudden rise in pressure caused by the unwanted ignition of the HHO gas mix used to power the generator. For this, parts are purchased from the local hardware store. The brass fittings are a 12mm barrel, a 12mm female T-fitting and a 12mm to 9mm hose reducer as shown here:

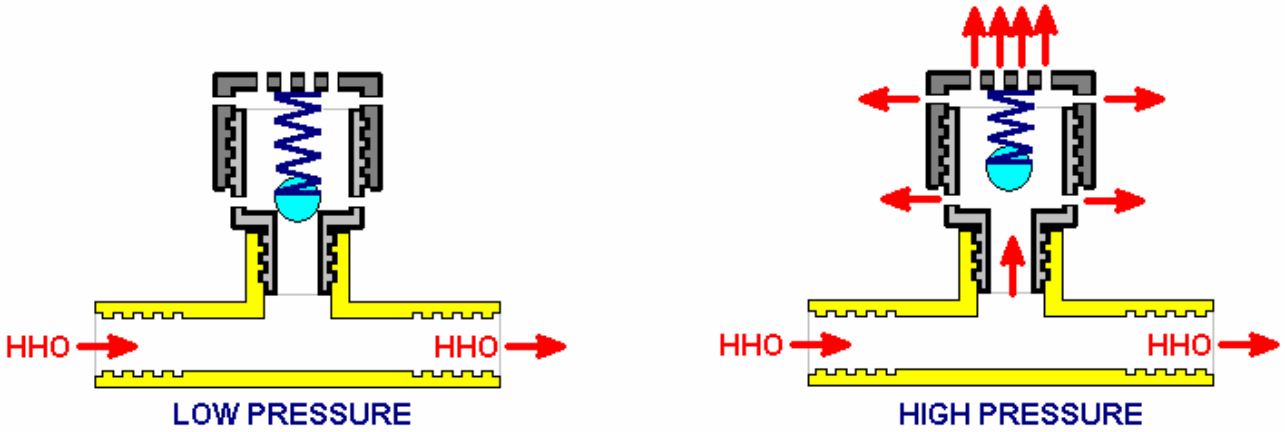
The PVC plastic fittings are a ½" to 1-1/4" reducer and a 1-1/4" End Cap, along with the roller ball from an old-fashioned mouse and a relatively weak compression spring to hold the ball in place during normal operation where the gas pressure is low:

These components are then assembled to produce the pressure-release valve:

The inside of the flash-arrestor looks like this:



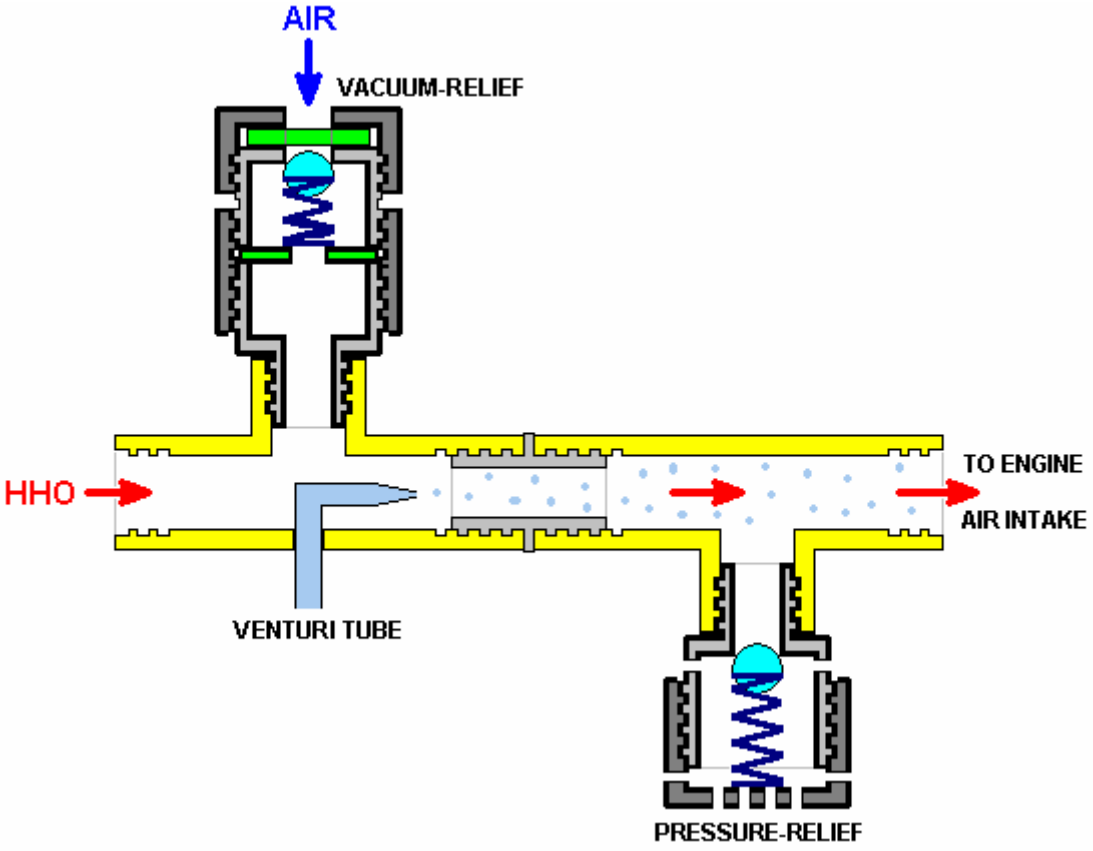
The ball is held in place by the spring allowing the HHO to flow past it, but if a sudden increase in pressure should occur, then the ball is forced upwards, opening a path to the many holes drilled in the plastic fittings:



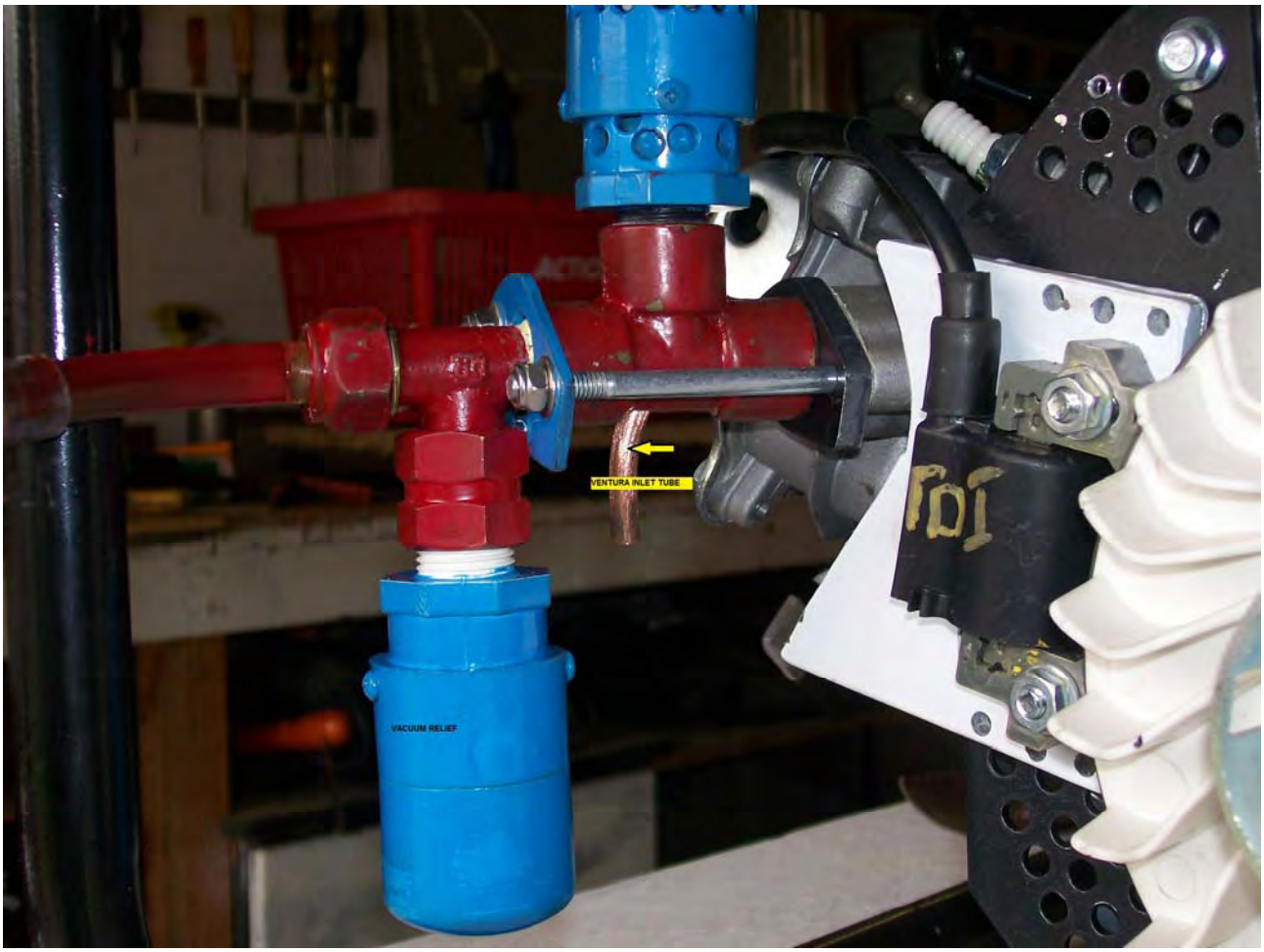
When the gas pressure drops again, the spring pushes the roller ball down to seal off the pressure-release holes.

However, Selwyn adds an additional spring-loaded valve to the arrangement. This one is there in case the electrolyser fails to produce a sufficient volume of gas in the event of a sudden increase in demand. This valve is marked as a “vacuum-relief” valve although, strictly speaking, it deals with reduced pressure rather than an actual vacuum. The arrangement is shown below. Please take note of the fact that Selwyn uses the Hogg style of electrolyser and that design has a bubbler built into it, so if you are using some other design of electrolyser, please be very sure to use at least one bubbler between the electrolyser and the engine, in spite of the fact that there is very little chance of the engine mis-firing and igniting the HHO gas in the electrolyser. For an engine of this size, an electrolyser which produces 4.5 or 5 lpm of HHO should be adequate.

The addition of cold water mist through a Venturi tube as shown, both lowers the engine temperature and increases the engine power as the mist converts instantly to flash-steam when the HHO gas ignites, raising the pressure inside the cylinder and boosting the power output.

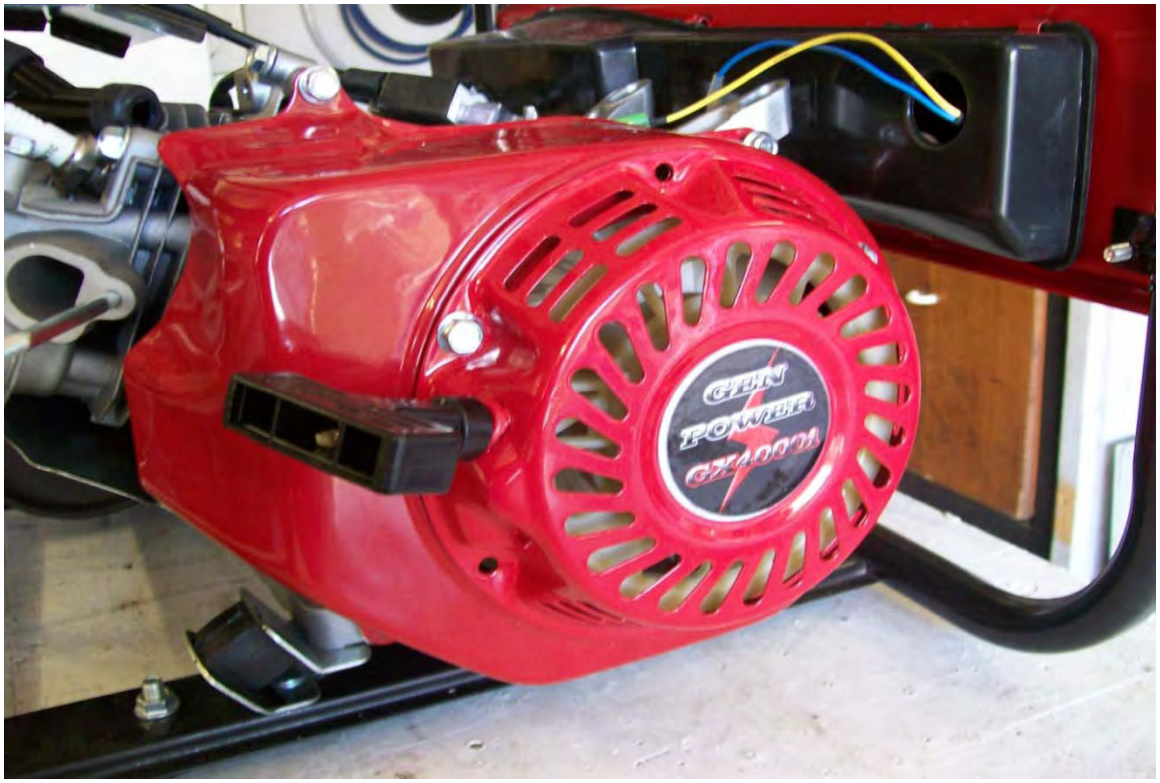


Next, a piece of ¼" (6 mm) thick aluminium plate is cut and shaped to the size of the carburettor gasket which is not a symmetrical item. This is done by tracing the gasket and transferring it on to the aluminium plate, drilling the holes and then cutting out the outline shape. The edges are then filed to create a nice fit on the engine port.

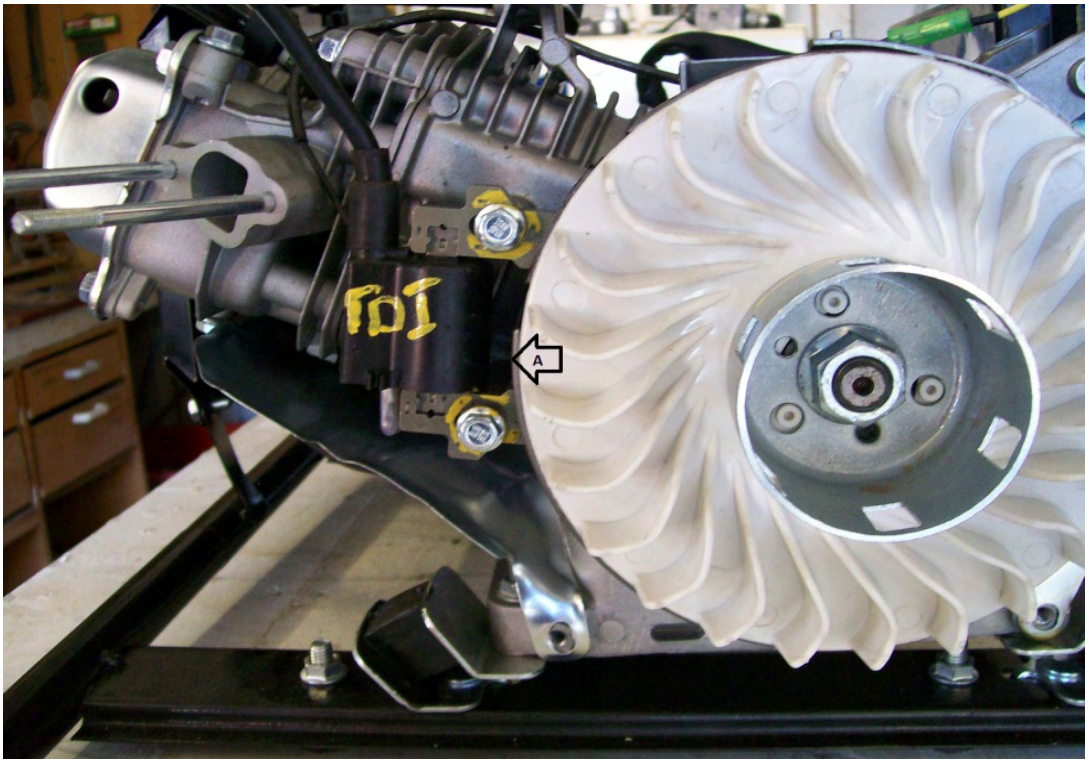


The pipes, backing plate, pressure-relief, vacuum-relief, gaskets, nuts and bolts are then assembled as shown above. Most of the pressure-relief valve components shown in the photograph have been painted, which happens to conceal the different materials being used.

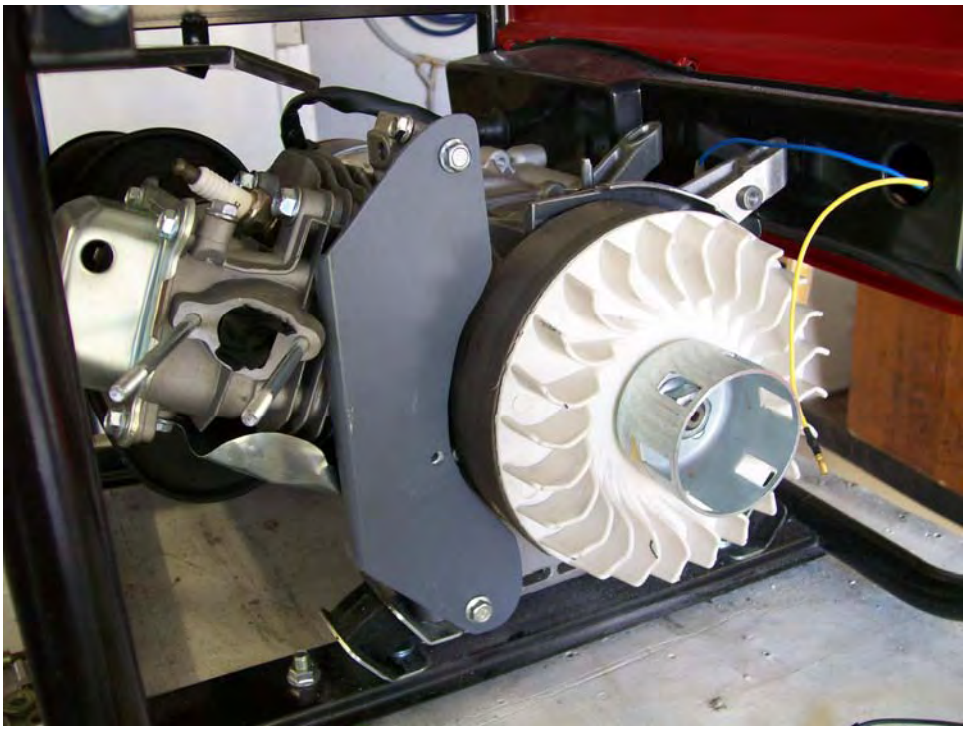
At this point an electrolyser of any design which can produce at least 4.5 litres of HHO gas mix per minute is connected to the intake. The electrolyser most often used by Selwyn is the Hogg design disclosed by him earlier in this chapter.



The manual Pull-start and the generator cover are now removed. It is only necessary to remove four of the bolts to take the cover off:



This is the engine with the starter pull and the blower cover removed. At “A” you can see the magnetic pulse type **T**ransistor **D**ischarge **I**gnition pick-up in it’s original position, bolted in place at 8 degrees before Top Dead Centre. This needs to be removed and an aluminium plate inserted to allow the TDI to be mounted in it’s new position. Because of the new fuel, it is necessary to retard the ignition system. This can be done in one of two ways, neither of which is particularly easy, so you may need the help of an engineering shop. The easiest way is to modify the installed ignition to Top Dead Centre. This is Selwyn’s aluminium TDI adaptor plate which he made from 2mm thick aluminium sheet:

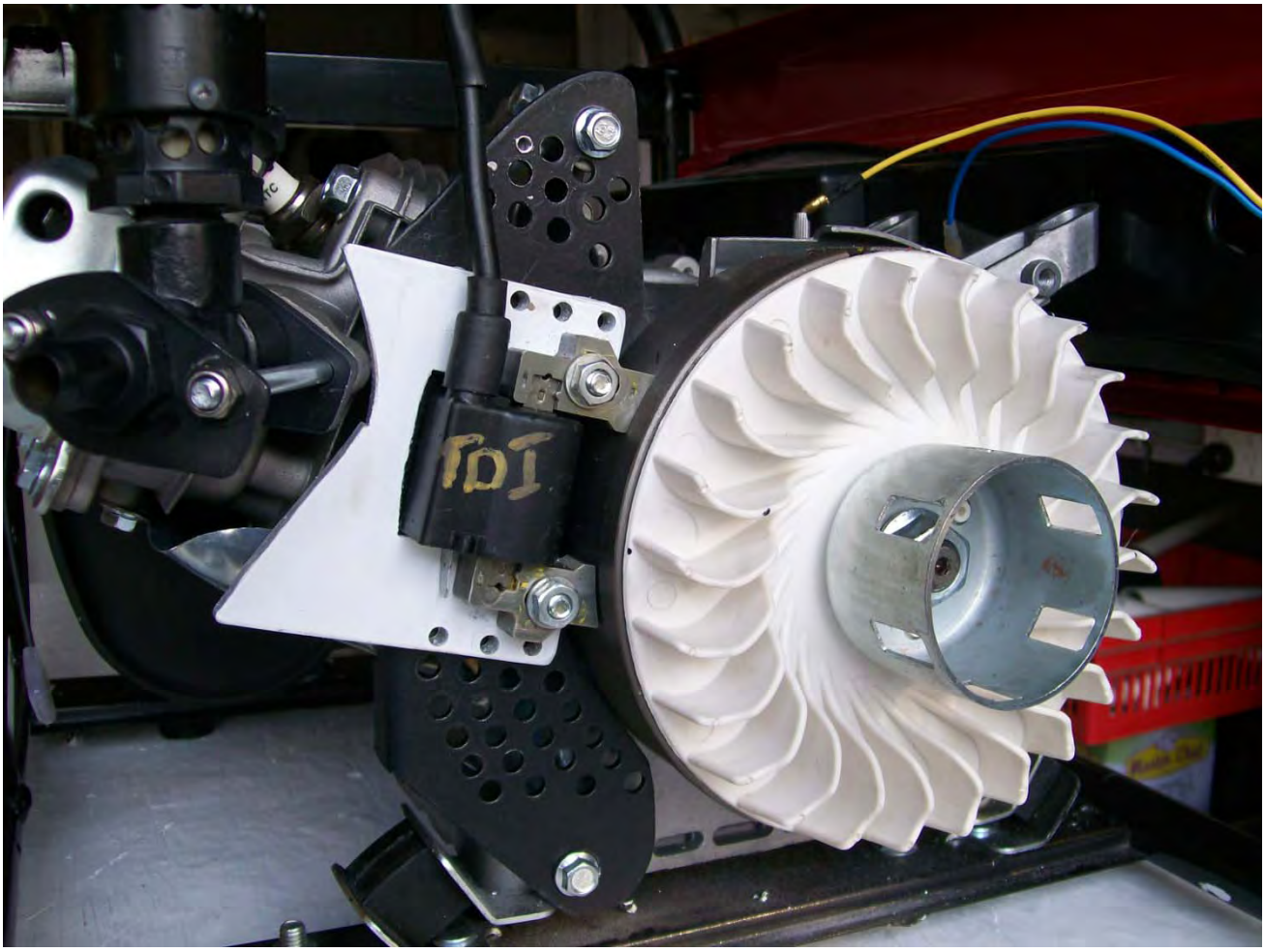


In this picture, the outline of the fuel intake port is obscured due to it having been temporarily blocked off during the construction. The tools required for constructing these components are a drill press and a jig saw fitted with a metal blade. Selwyn used this timing alteration method on his own smaller generator which has run trouble-free

for a year. The objective is to delay the ignition spark from 8 degrees before Top Dead Centre to either Top Dead Centre or to 1 degree after TDC. This allows for a good spark on the compression stroke and when the waste spark occurs, the inlet valve has not yet opened and so there is no HHO is in the ignition area. That is to say, the exhaust valve has just closed and the inlet valve has not yet opened. This results in a good compression stroke for the HHO and does not try to send the piston backwards due to premature ignition of the gas mix. The above picture shows the aluminium plate mounted and ready to accept the pick-up. This plate needs to have air holes drilled in it in order to allow cooling air to flow over the engine fins behind it.

The TDI adaptor plate looks like this:

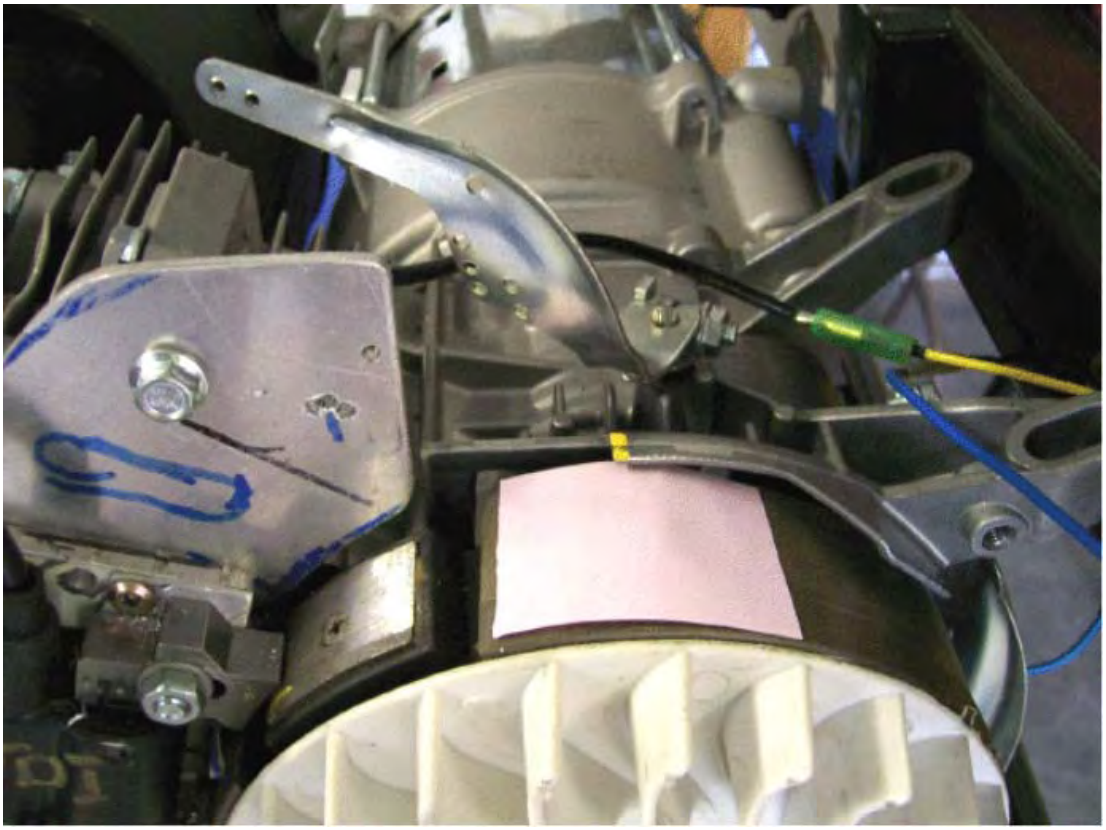
And as shown below, the support plate is drilled with the ventilation holes. In this photograph the adaptor plate is just resting on the support plate. Later, when the TDC timing position is established, the adaptor plate will be bolted to it using the three holes top and bottom on the white plate. This locks the timing to that setting and the timing is never changed. In 2010, when adapting a previous generator, an experienced mechanic was asked to establish the TDI plate position and he charged sixty Australian dollars for doing that.



Finally, the covers and the Starter handle need to be bolted back in place.

Instead of paying somebody else to set the new spark timing, it is perfectly possible to do that yourself. One effective method is as follows:

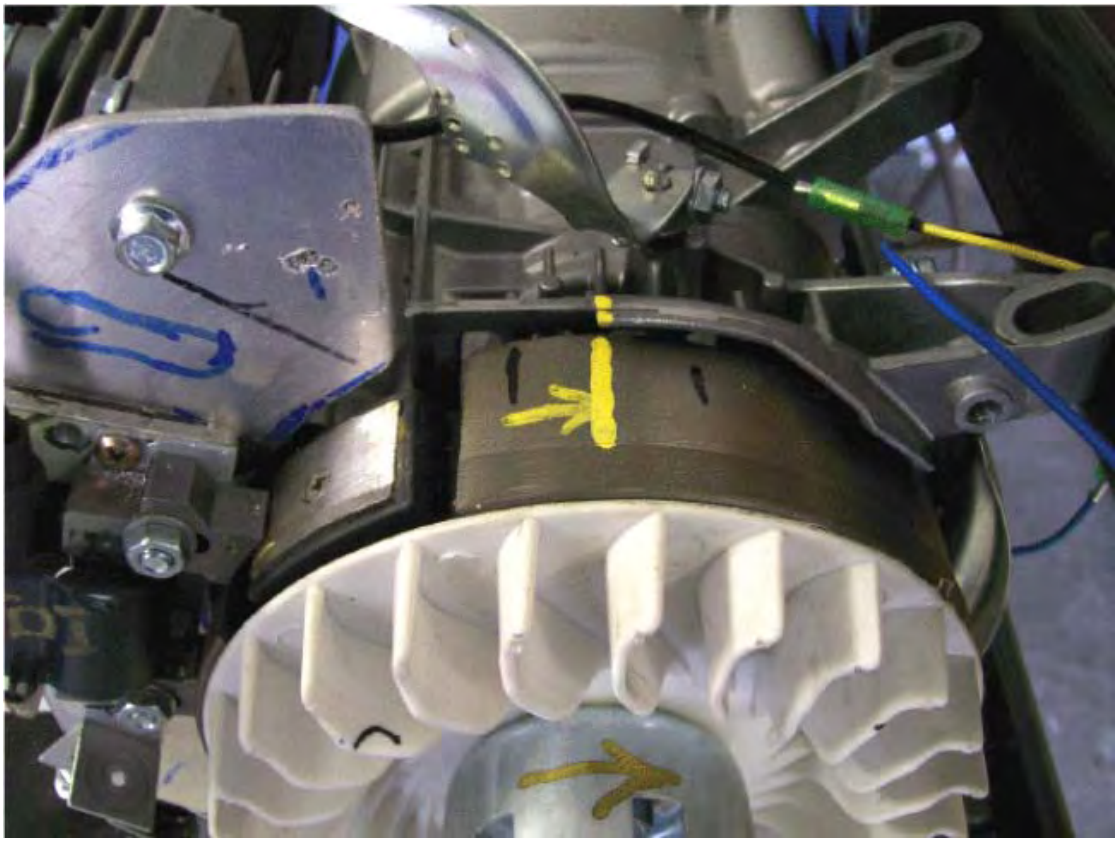
1. Mark the casing of the engine in a convenient location as shown in yellow in this photograph:



2. Remove the spark plug and insert a long screwdriver until the top of the piston is felt. Manually rotate the engine (clockwise for this generator as can be seen from the curved fan pieces on the flywheel) until the screwdriver is no longer pushed upwards. It may take more than one rotation to find this point accurately. When that point is found, mark the flywheel directly in line with the casing mark which you just made. This marking needs to be very accurate.

3. Continue rotating the flywheel very slowly until the screwdriver starts to go down again and mark that point on the flywheel. Again, this marking needs to be very accurate.

4. Measure the distance along the flywheel between the two flywheel marks which you have just made and then make a larger mark on the flywheel exactly half way between your two marks. If accurately done, this new point is where the flywheel is when the piston is exactly at Top Dead Centre, which is where we want the spark to occur. This marking on Selwyn's flywheel is like this:



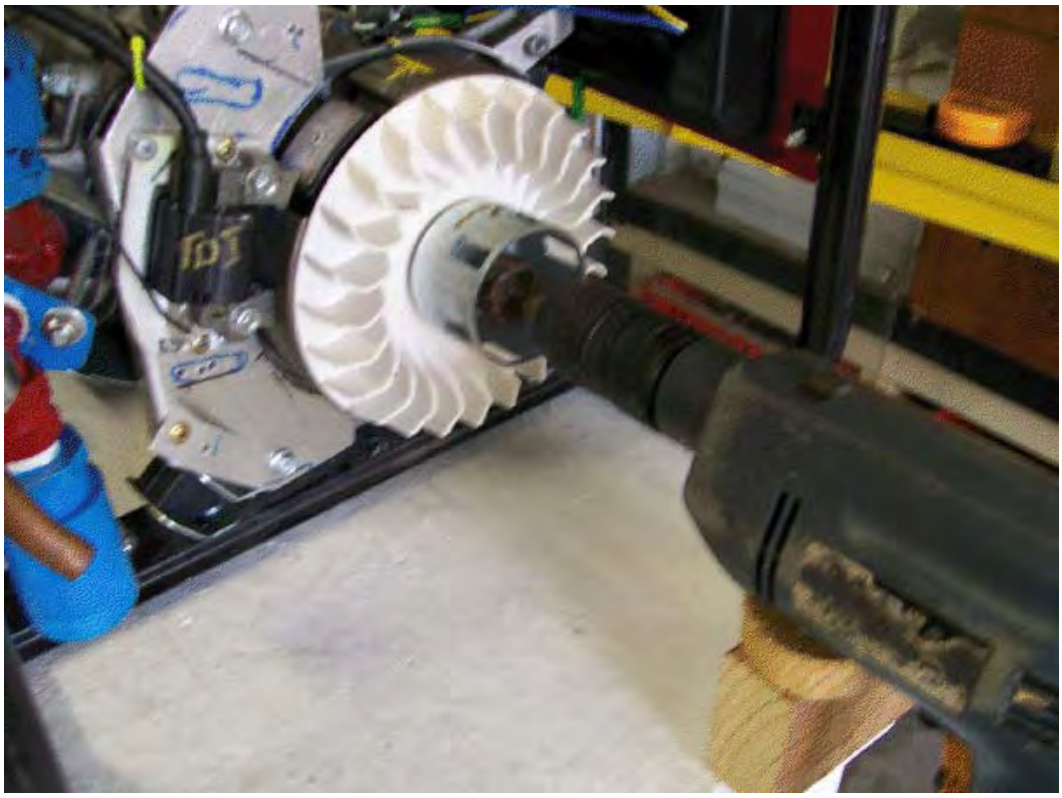
5. Next comes a bit of arithmetic. The diameter of the flywheel is 180 mm which means that it's circumference is 3.14159 x 180 = 565.5 mm and as there are 360 degrees in each rotation of the flywheel, then the outside edge of the flywheel will move 1.57 mm for each of those degrees.

   The engine specification states that the spark timing is 8 degrees before Top Dead Centre and we want the spark to occur exactly at TDC, which means that we want 8 x 1.57 = 12.5 mm of the flywheel circumference to have passed by before the spark occurs.

6. To achieve this delay in the spark timing, the TDI needs to be moved 12.5 mm in the direction which the flywheel rotates. You will notice that for this major timing change, the TDI adjustment is very small, only half an inch.

7. When the TDI adjustment has been made, the timing can be checked using an automotive timing light connected to the spark plug lead. The engine can be spun using an electric drill. As the flywheel is spinning fast and the flash of light from the timing light is very short, it makes the flywheel mark appear to be stationary in spite of the fact that it is passing by very rapidly. If the TDI adjustment is correct, then the central mark made on the flywheel will appear to be stationary and exactly aligned with the mark made on the casing.

   This is exactly what happened when Selwyn's motor had it's timing adjusted, but the important factor is to have the spark close to the Top Dead Centre point to make sure that the inlet valve is fully closed before the spark occurs. Two degrees after Top Dead Centre is a popular point for the spark with many of the existing generator conversions which I have been told about, possibly to reduced the loading on the piston's connecting rod. Here is a photograph of Selwyn's latest generator conversion having it's new spark timing checked out:



8. Most small petrol engines have the spark timing set between 8 degrees and 10 degrees before Top Dead Centre. If it so happens that you do not know what the timing of your particular generator is, then complete the flywheel marking procedure of step 4 above, but make three additional marks on each side of the TDC mark. Space those marks 1.5 mm apart as they will then make a scale which shows each degree from 3 degrees before TDC to 3 degrees after TDC. When the timing light is used, it then shows exactly where the spark occurs and if the engine had an original spark timing which was not 8 degrees before TDC, then the scale shows immediately how much further the TDI needs to be moved to set the spark exactly where you want it to occur.

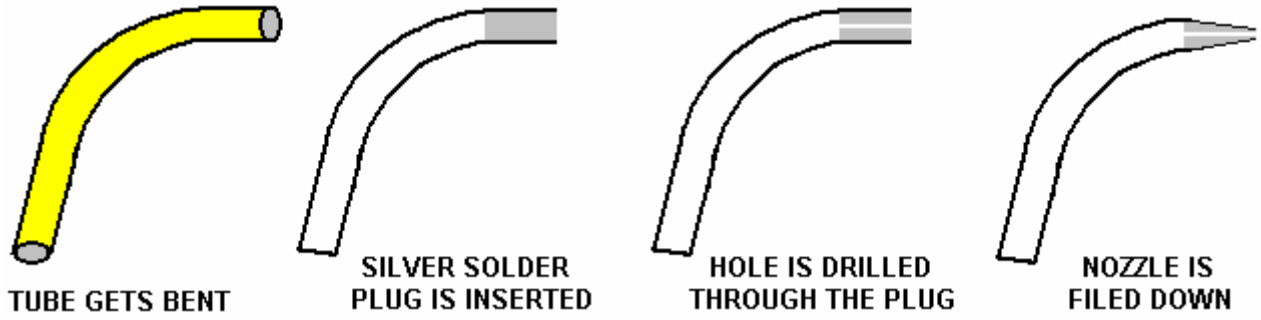
The Cold Water Mist.

Getting the fine droplets of water into the engine can be done two different ways. The first way is to use a Venturi tube which generates a fine spray of droplets when air moves rapidly past a small water-filled hole. You may not have noticed, but this method has been used extensively in perfume sprays and it is very effective. Selwyn describes how he constructs a Venturi tube:

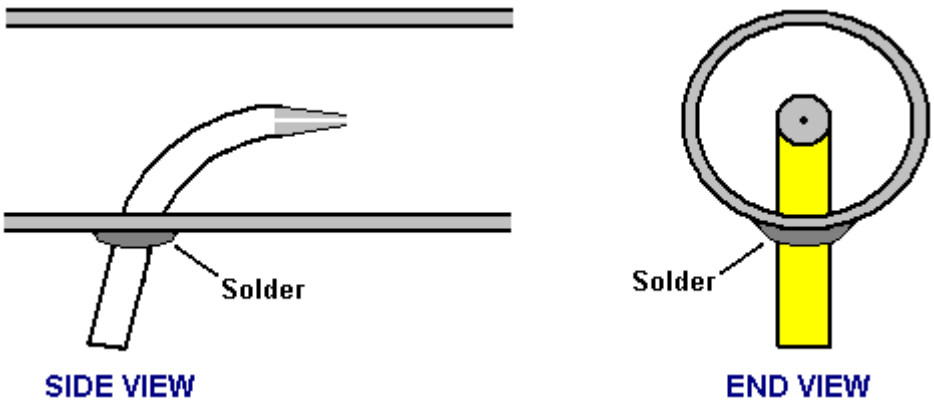
A short length of 1/4” (5 or 6 mm) diameter copper tubing is used. This is generally available as central heating supplies and if there is any difficulty in finding some, then your local garage can probably direct you to a supplier (if they don’t just give you a short length from their own supply).

The copper pipe is then heated with a plumber's gas torch and bent very slowly and carefully to the shape shown above. Some people find it helpful to insert a length of suitable flexible material into the pipe before starting the bending – something like the coiled steel spring material used to support net curtains – as that helps to keep the copper pipe from kinking when being bent.

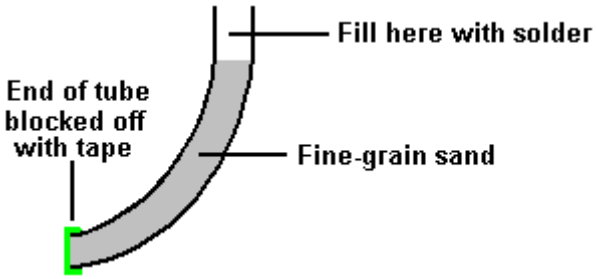
Next, the end of the copper pipe which will form the nozzle, is filled with silver solder and the end filed flat. Then, a small hole is drilled through that silver solder plug. The smallest possible drill bit should be used for this, although the hole may need to be drilled out to a slightly larger diameter, depending on what the engine requires (which is found by successive trials):



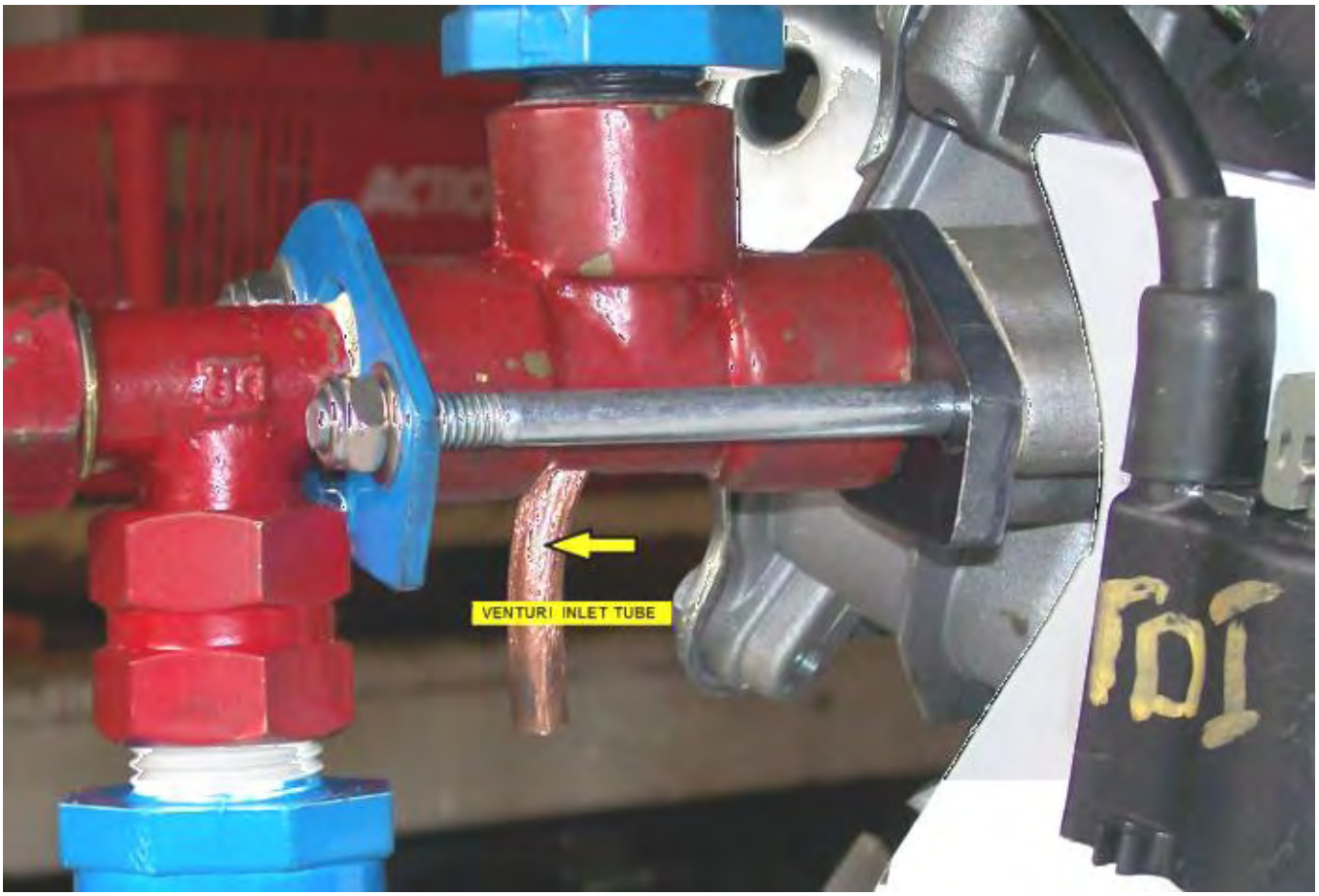
This Venturi tube is to be inserted into the last brass fitting before the engine, so a 1/4" hole is drilled through the brass and then the drill is removed very slowly at a slight angle, the angle of drag being down the axis length of the brass fitting. The copper Venturi tube is then inserted through the hole and positioned so that the Venturi hole is aligned up exactly with the centreline of the brass fitting and positioned exactly in the middle of the cross-section of the brass fitting and then soldered in place:



The method which Selwyn uses to block off the end of the copper tube with silver solder is to seal the far end of the tube with tape and fill the tube with fine-grained sand like this:



And then the tube is heated with the gas torch flame and the solder run into the top part of the tube. When the solder has cooled, the tape is removed and the sand removed by tapping the tube. When the hole has been drilled through the solder, air is blown through it to dislodge any remaining sand, and then water is forced through the hole. As the tube is short, any remaining sand can be removed with a pipe-cleaner or any similar slender cleaning device. The installed Venturi tube can be seen here:



The second way to introduce cold water mist into the airstream entering the engine is to use a commercial "pond fogger" which can be bought at pet supplies outlets. These have to be powered electrically and housed in their own water container. Some of the more advanced versions float on the surface of the water so that the fog-generating section is always submerged to the ideal operational depth below the surface of the water.

The generator should run well with 5 lpm of HHO gas plus cold water mist. Any design of electrolyser can be used. However, when used with rainwater, the Hogg electrolyser will draw about 1.4 amps per cell, giving a total input of about 115 watts when run on a 12-volt electrical supply. While rainwater is supposedly pure, the reality is that it seldom is and it's ability to carry a current varies dramatically from place to place and even more widely from country to country. However, regarding the water, Selwyn says:

The water I use is treated in a special way to make sure that the electrolyser runs at the lowest temperature and amperage possible. For this, using rain water is a must and the rain water coming off a steel roof is the best.

The water is then treated by inserting a double coil of stainless steel wire into a volume of about 5 litres of water. A supply of 12-volts DC is applied to the coils, and the resulting current allowed to run through the coils for about 5 hours. This results in hot and very dirty water. The water is then filtered using a 0.5 micron filter making the water ready for use in the electrolyser. If more water is needed, say 30 litres, then leave the coils running for at least 24 hours.

I use an old 35 litre beer keg and prepare 30 litres at a time. A major reason for doing this is to remove all the solids suspended in the water so that they will not clog up the stainless steel mesh inside the electrolyser.

After the construction of the Hogg electrolyser is completed, then the stainless steel mesh electrodes need to be treated and cleaned. For this I use distilled water and fill the electrolyser enough to cover all the plates, and then add 1 packet of citric acid for each 3 litres of water used to fill the electrolyser. I got the citric acid from www.hho-research.com.au which is an Australia-only supplier and each packet has about 22 grams of citric acid in it:

The pumps are then run for about an hour after which the Hogg tubes are washed out completely with distilled water and then allowed to dry completely. This removes any residue from the stainless steel mesh electrodes, making the gas production rate much greater.

I use an ordinary car battery to generate the HHO gas needed to start the generator running, after which, a standard battery charger powered by the generator output is used to keep the starting battery topped up.

Please Note: This document has been prepared for information purposes only and must not be construed as an encouragement to build any new device nor to adapt any existing device. If you undertake any kind of construction work, then you do so entirely at your own risk. You, and only you, are responsible for your own actions. This document must not be seen as an endorsement of this kind of generator adaption nor as providing any kind of guarantee that an adaption of this kind would work for you personally. This document merely describes what has been achieved by other people and you must not consider it as being a foolproof blueprint for replication by anyone else.

There are YouTube videos which show generators being operated on what appears to be just HHO gas alone and while the operation does not appear to be anywhere close to full power, the addition of cold water mist would probably make a major difference to the performance, but it does demonstrate that a generator can certainly be run without using any fossil fuel. The spark circuit in the first video appears to be powered by a small mains unit, but as the generator is lighting a powerful lamp, that electrical input could almost certainly be met by the output from the generator when it is running.

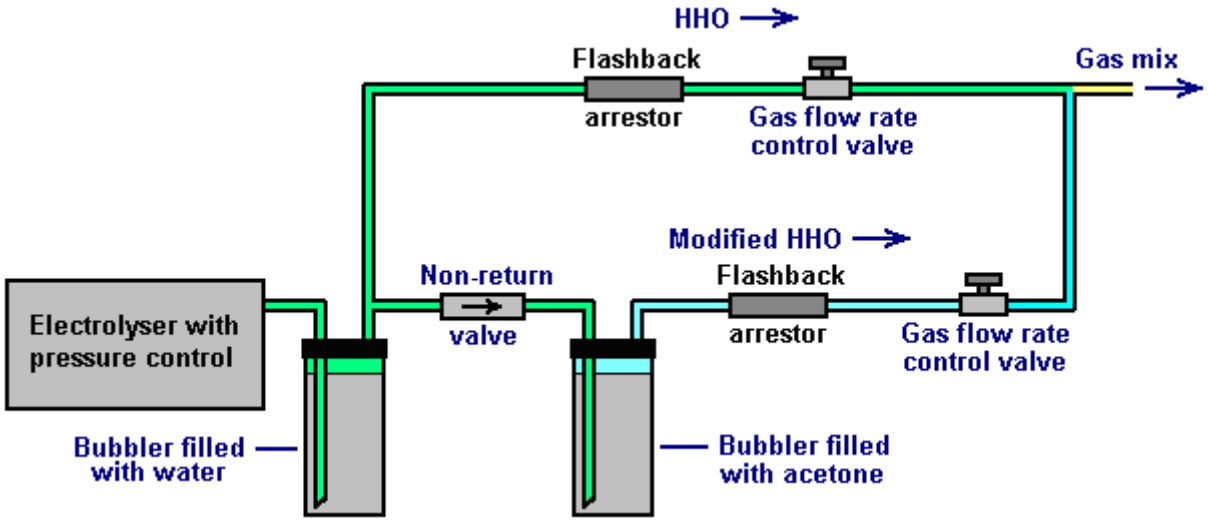
Running an Unmodified Generator on HHO

The reason for the modification of standard generators as shown above is due to the fact that the HHO gas mix produced by an electrolyser, ignites about a thousand times faster than a hydrocarbon fuel, and because of that, the spark which ignites the fuel needs to be delayed. That mechanical adaption of the generator can be avoided if the HHO gas mix is modified so that it ignites more slowly. This can, and has been done.

David Quirey of New Zealand has been operating an unmodified generator and a welding torch on the HHO output from his 6 lpm own-design of electrolyser, for many years now. Henry Paine's US Letters Patent No. 308,276 dated 18th November 1884, states that HHO gas can be converted into a more convenient gas which is much easier to handle, by the simple process of bubbling it through a suitable liquid such as turpentine or linseed oil. Although unaware of Henry Paine's patent, David discovered the technique independently and he has extended the technology further so that the gas ignition speed can be set manually.

One important point which David stresses is that it is essential that the HHO coming from the electrolyser is passed through an ordinary bubbler containing water, before it passes through the second bubbler containing the modifying liquid. David finds that the lighter liquid, acetone, works better than the liquids suggested by Henry Paine although white spirit, carbon tetrafluoride, aviation fuel, hexane or even petrol can be used and any of them will slow the flame speed right down to that of butane. If the flame is being used for a specialist task such as jewellery making or glass blowing, then there may be an advantage in using one particular modifying liquid. Please note that the bubbler holding the acetone needs to be made of stainless steel as acetone can dissolve some plastics.

David has further modified the characteristics of the output gas by adding in a percentage of the unmodified HHO gas. Although it is actually, subtle and sophisticated, David's overall system is easy to understand. The ratio of the two gasses is adjusted by the settings of the two control valves as shown here:



Adjusting the ratio of modified HHO to unmodified HHO allows a high degree of control over the characteristics of the resulting gas mix. Added to that, David has developed an electronic control system which oversees and manages the gas flow rate according to the user's needs at any given moment. The result is a system which allows water and electricity to be the means of supplying a gas which can be used as a safe, general purpose fuel. If it is used to run a generator, then the system appears to become self-powered if part of the generator output is used to drive the electrolyser. It should be possible to substitute the modified gas mix for propane or butane and so operate a wide range of existing equipment for heating, cooking and/or lighting.

David runs a 4 horsepower Honda generator using this system:

The generator runs very well for David, however, I suspect that if cold water mist were introduced into the incoming air, then the power output would be increased due to the mist turning into flash-steam and providing

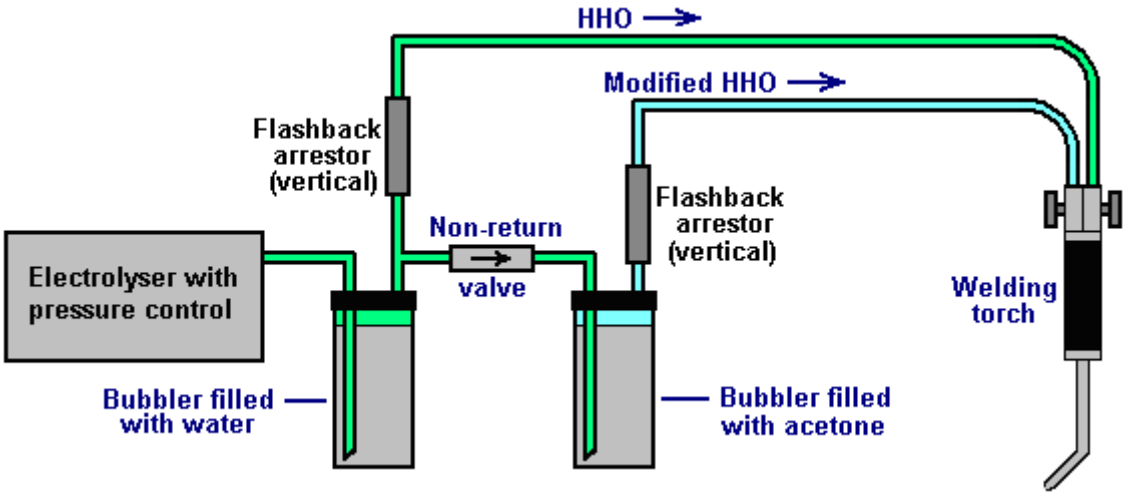
greater pressure on the piston during it's power stroke. Alternatively, it might be possible to match the present performance with a lesser gas flow rate, possibly powering a much larger generator if that were a requirement.

It needs to be understood that David uses electronics which manages and controls the gas flow volume, suiting it to whatever the needs are at any given moment. Consequently, it is probable that the six litres per minute which David's electrolyser can produce, is not actually used for most of the time. David also does welding, brazing and cutting with the same modified electrolyser gas mix which can provide adjustable flame heat and a flame length of anything up to two feet in length:

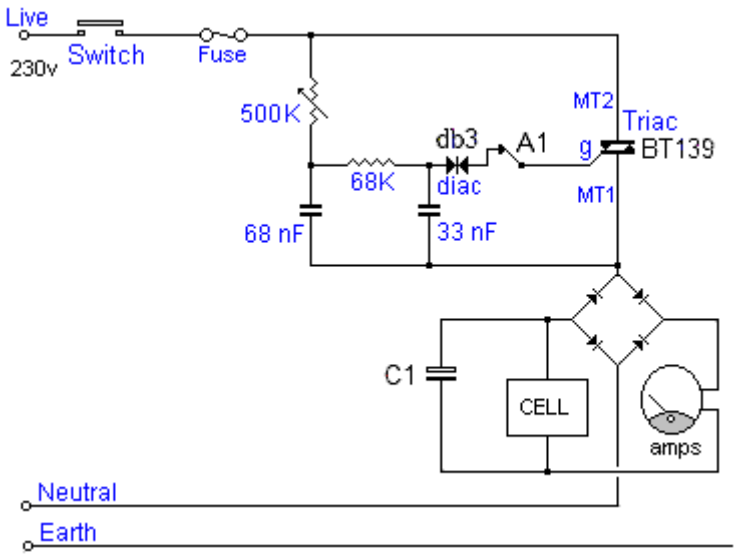
It is a good idea to use a proven design with full control electronics. David can help here with detailed step by step construction plans and instructional videos.

You can contact David at dahq@clear.net.nz for information on what is available to help you at the present time.

When using the system for welding, David uses the mains to power the electrolyser, the arrangement being like this:



The flashback arrestors are a sand-filled design and so are mounted vertically. The gas production rate is knob controlled using this circuit:

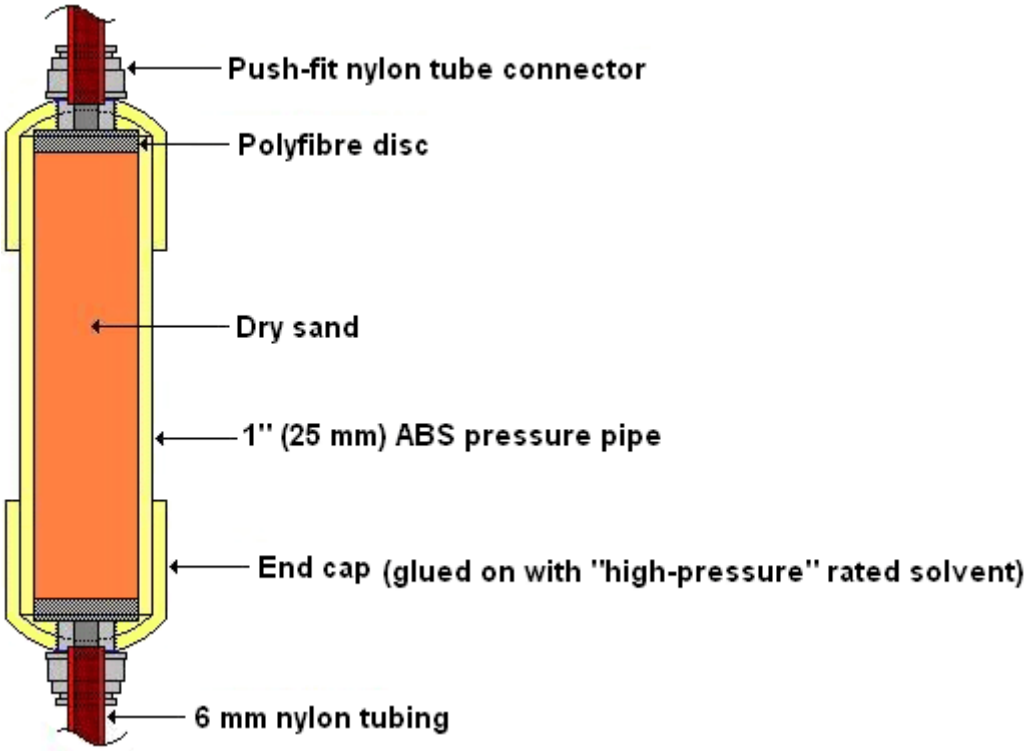


The first part of David Quirey's circuit acts very much like a dimmer light switch. The 230-volt AC mains is fed through an On/Off switch and then an ordinary mains fuse. The current flow on through the circuit is blocked by the BT139 triac until it receives a pulse from the db3 diac (which is a component specifically designed to feed pulses to a triac).

As the voltage builds up on the 68 nanofarad capacitor it eventually reaches the point at which it triggers the triac, which then switches on and remains on until the mains voltage drops down to zero again. The 500K variable resistor sets the rate at which the capacitor charges up, and so it controls the length of time that the triac is on in any given second (and so, the level of power fed onwards to the rest of the circuit). This happens on both the positive-going half of the AC waveform and the negative-going half of the mains sinewave voltage supply. Both the diac and the triac operate with AC and trigger either 100 or 120 times per second depending on the frequency at which the local mains runs.

The current flow is then passed to a bridge rectifier in order to convert the AC into pulsing DC and the capacitor C1 which is 400-volt rated, smoothes the resulting DC. David's cell has a large number of plates and so, operates off the 300 volts produced by this system. The ammeter between the diode bridge and the cell indicates the current flow and so, the amount of gas being produced at any given moment.

The flashback arrestors are constructed as shown here:



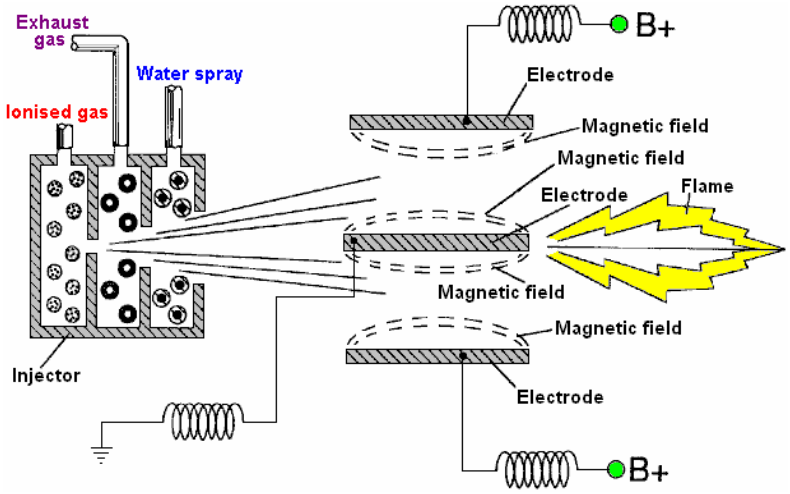
Sincere thanks are due to David Quirey for freely sharing his design and experiences, and for his willingness to provide direct additional support and further details should they be needed.

Stan Meyer's Water Injection System.

While the above simple adaption will work for an electrical generator which is stationary and which is designed to provide large amounts of electrical power, some of which can be used for producing hydroxy gas from water, it is much more difficult to run a large capacity petrol/gasoline engine continuously with just water as the "fuel".

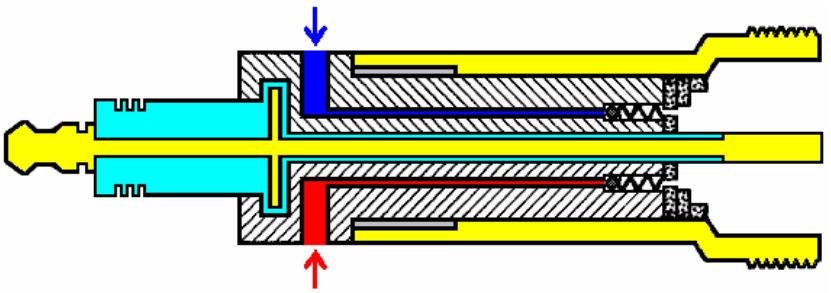
For larger engines aimed at producing mechanical output, we need a more powerful system, such as that developed by Stan Meyer of America. Although many years have passed since the sudden, highly suspicious death of Stan Meyer, as far as I am aware, his design has not been replicated, mainly because people do not understand the information which he left behind. However, recently, a man whose forum ID is "H2Opower" has helped explain what Stan meant, and much of the following description is down to his sharing his understanding publicly and freely.

Stan's Canadian patent 2,067,735 has an arrangement where the injector schematic shows the injection into the engine of three separate components:



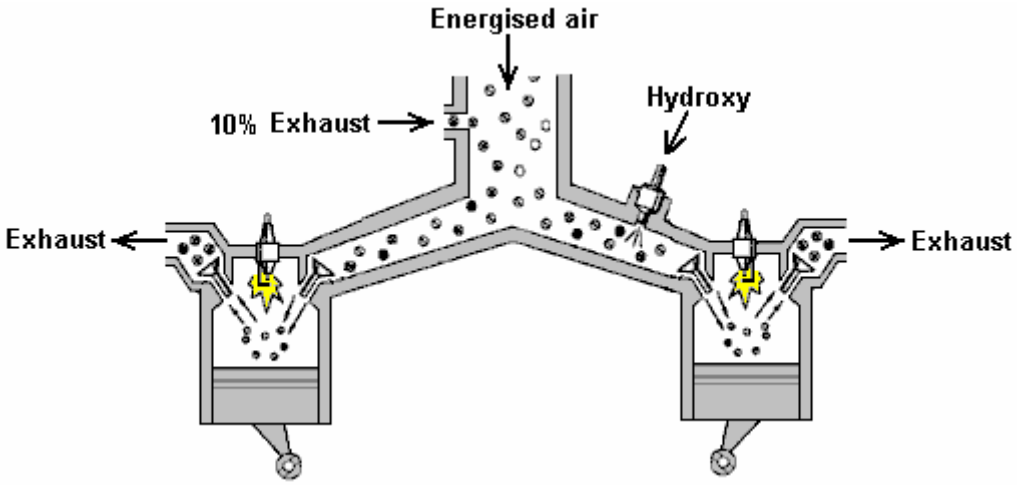
One component is described as ionised gas and ambient air is mentioned. The second component is part of the exhaust gas which is hot water vapour fed through a limiting valve, referred to as "inert gas". The third injection component is a very fine spray of water droplets or cold water “mist”. This three-component mixture is passed between high voltage electrodes and if the mix does not ignite spontaneously, then it is ignited with a spark.

One version of Stan's recommended injectors for an existing engine are shown like this:



This is a cross-section which only shows two of the three gas inputs flowing through this injector/spark plug. Each gas input has it's own feed passage which leads out between the central electrode and the circular outer electrode, and each feeder tube has it's own one-way valve to prevent the power stroke from forcing the incoming fuel back up it's feed tube.

This is only one of the ways that Stan shows that it can be done. Here is his diagram for a two cylinder engine fuel input, although it applies equally as well to any number of cylinders:

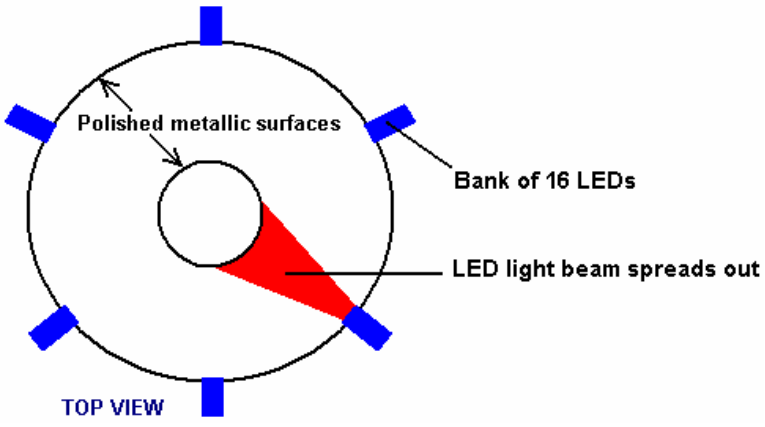


Let me stress that this is just an explanatory diagram and you do **not** have the intake and exhaust valves open when the spark plugs fire. Also, the pistons do not go up and down together but their movement is staggered to give a less uneven drive to the crankshaft. The point to note here is that the fuel intake is through standard valves and ordinary spark plugs are shown. However, this diagram for US Patent 5,293,857 is based on the use of an electrolyser and Stan found a way to avoid the need for an electrolyser.

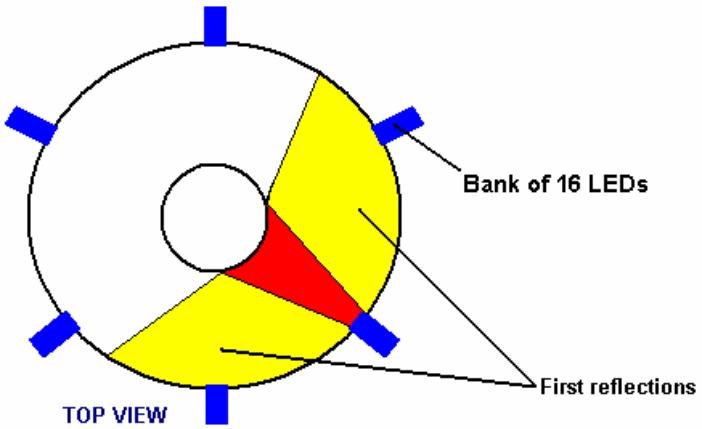
For this, Stan's "Gas Processor" is a really key device in spite of it being what appears to be a fairly simple construction. It operates by pumping light energy into the fuel components. This may not sound like a big deal but it most definitely is, in fact, it is the heart of Stan's system. In addition to the extra energy, the fuel components are also stressed by high-voltage pulsed DC applied in such a way that it tends to pull the molecules apart as they flow past into the engine.

The Gas Processor is constructed from two highly reflective metal cylinders, 96 Light-Emitting Diodes ("LEDs") and a carefully adjusted high-voltage pulsed DC source. The reflecting tubes are used to make the light bounce backwards and forwards indefinitely until it is absorbed by the passing molecules. The choice of LEDs is very important as the wavelength of the light needs to be matched to the resonant frequencies of the fuel components passing by. Laser LEDs can be used but they need to be angled slightly in order for them to bounce and completely fill a section of the Gas Processor gas passage.

The operation is simple in concept. Six banks of sixteen LEDs are installed in the outer cylinder which has a polished inner surface. To see the effect of each LED, consider just one LED on it's own:

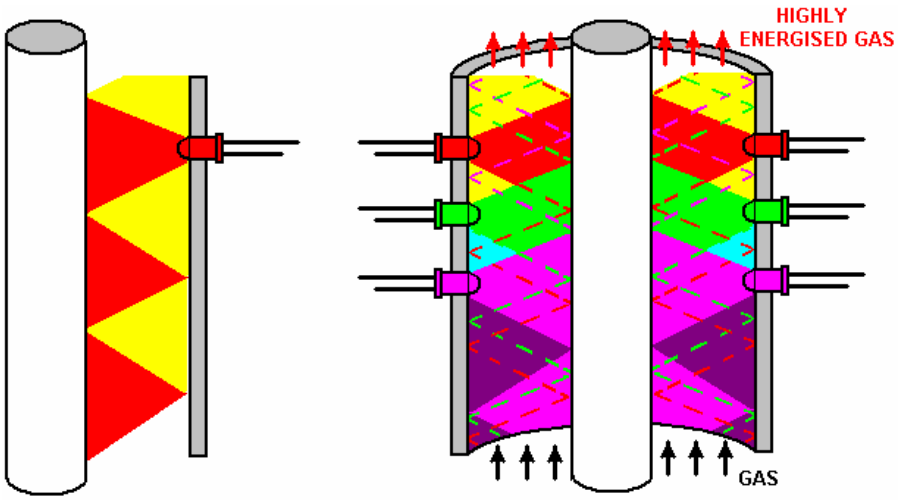


The LED light shines on to the inner tube which has a highly polished outer surface. In the diagram above, the light is shown in red, and the uppermost LEDs of the six columns of LEDs can be seen. The light is then reflected back to the outer cylinder again:



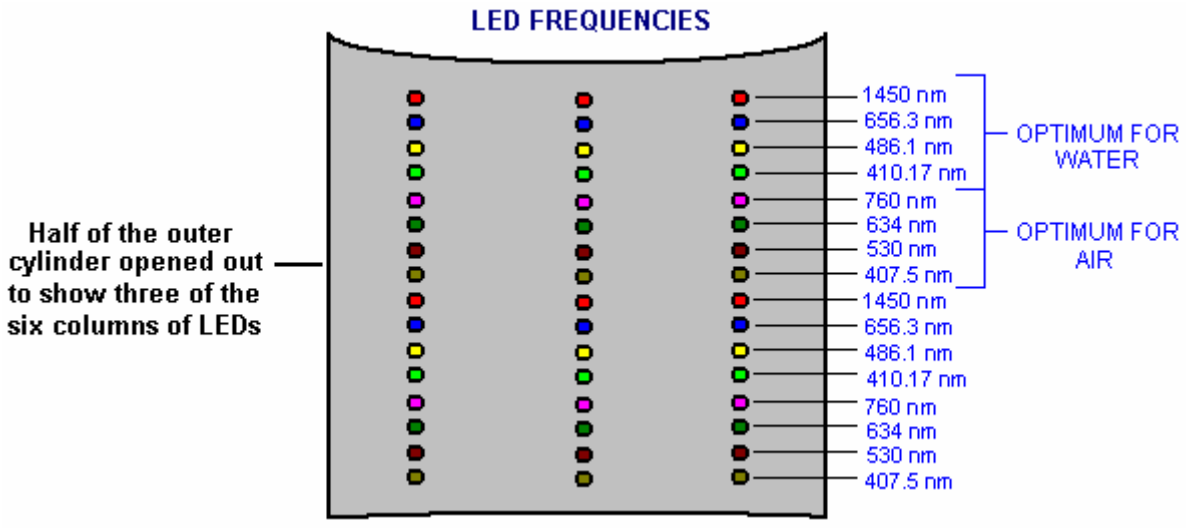
The reflected light is shown here in yellow although it is identical in wavelength to the section shown in red. This reflected light is again reflected by the outer tube and that repeats over and over again until the whole of the area between the two tubes is completely filled with the light. This will happen with just one LED, and that light will continue to be reflected backwards and forwards indefinitely if it does not collide with a gas molecule. That is the effect of just one LED, but there are six LEDs at that level, each producing light of the same frequency and reinforcing the power of each of the other five LEDs, producing a very powerful overall band of light.

The same thing happens in the vertical plane with the light bouncing all the way down the tubes, and as the LEDs at different levels generate different wavelengths, there is a powerful interaction between the different frequency waves, producing intermediate frequencies by a process with the technical name of "heterodyning":



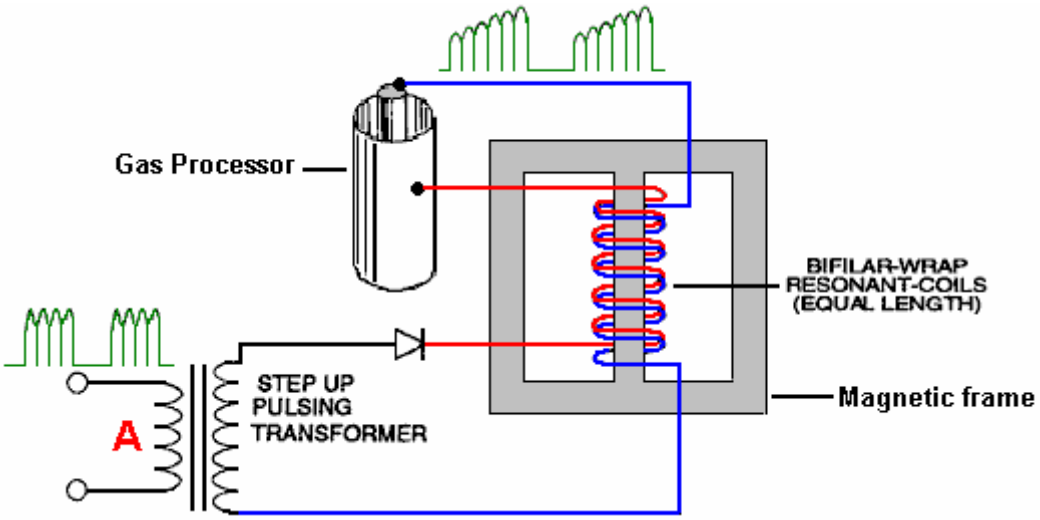
In this diagram it has not been possible to show the way in which the reflected light from each LED interacts with the light from all of the other LEDs, but there is a complete intermixing of the beams. The LEDs are shown with greatly exaggerated size and spacing in order to give a reasonably understandable diagram.

The diagram above, barely indicates the level of light intensity inside the Gas Processor, and that is from only six of the ninety six LEDs actually installed. It is a little difficult to visualise this device, but the gas (of whatever type) flows up through a circular doughnut-shaped space between two pipes and is hammered by a very high level of light energy of the appropriate frequency. "H2Opower" believes that the optimum LED frequencies are as shown here:



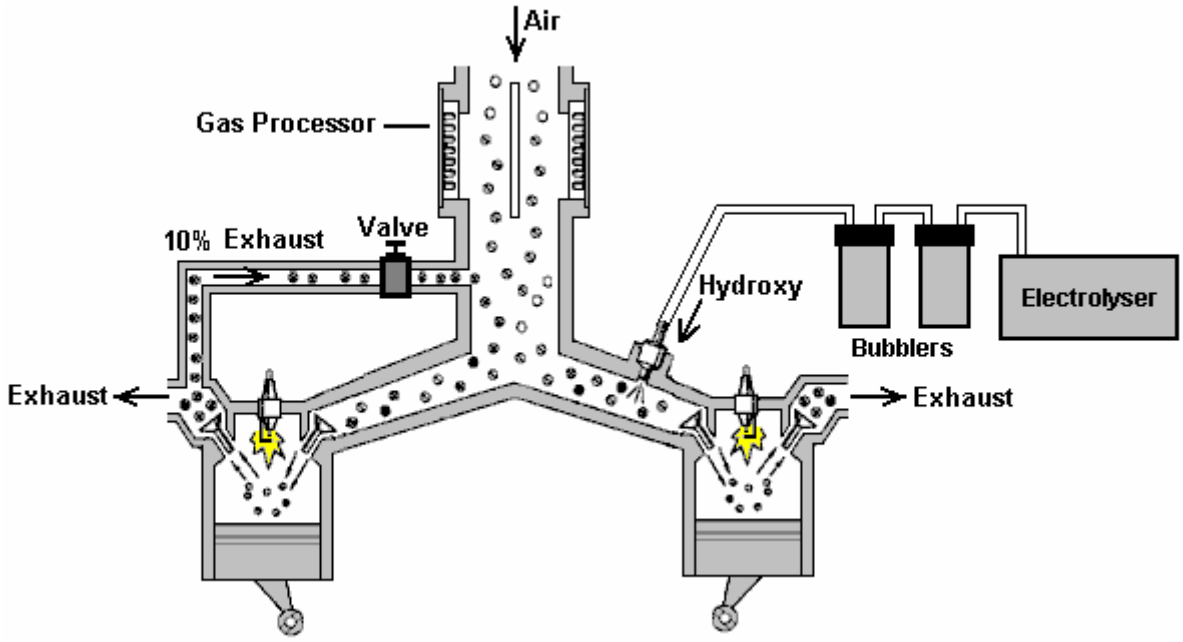
and while, not surprisingly, I have not been able to find an LED supplier for these exact frequencies, most LEDs emit a band of frequencies rather than just a single frequency, so the working LEDs need to include the frequencies shown here.

The operation of the Gas Processor is further enhanced by applying a pulsed high-voltage between the inner and outer metal cylinders. This voltage stresses the molecules of the gas flowing between the two cylinders and because it is pulsed, it applies a tugging action on the molecules, tending to pull them apart:



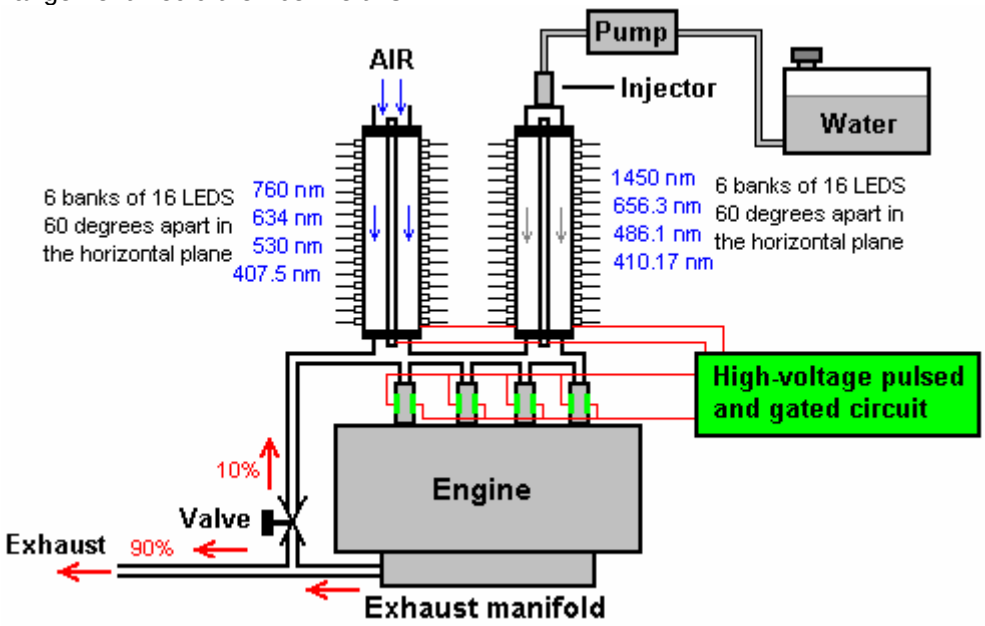
As with all high-power free-energy systems, having the device operate in resonance is very important. If it is still available, the excellent video: http://www.youtube.com/watch?v=kQdcwDCBoNY shows what resonant circuit operation is all about. Briefly, the frequency of the driving waveform passing through winding "**A**" in the diagram above, is adjusted until almost no current flows through that winding. This is the same as tuning a radio to a particular radio station - in both cases, the signal at that frequency finds it almost impossible to drive any current through the winding. This would make you suspect that almost no current would flow in the output winding, but that is not the case because that current flows into the secondary winding from the local environment (after all, here is no electrical connection between the two windings, so common sense tells you that the electrons flowing in winding "**A**" do not flow out of the other winding). This resonant operation gives you a major energy gain courtesy of the local environment.

The objective of the Gas Processor is to modify whatever gas is flowing through it, air, water vapour, hydroxy, or whatever, to one of it's highly charged, highly energetic states. The high voltage applied to the metal cylinders of the Gas Processor does not flow through the gas between the cylinders. Instead, it provides a very high electrostatic Plus charge on the outer cylinder and a very high Minus charge on the inner cylinder, and these two opposing charges pull the charged parts of each molecule towards themselves. The positively charged gas ions get pulled towards the Negative inner cylinder and the negatively charged gas ions get pulled towards the Positive outer cylinder. Stan uses this technique on the incoming air when he uses an electrolyser to power an engine:



However, Stan's preference is to use air, water mist and some of the exhaust gas from the previous power stroke to power his engines. This is an effective method as "H2Opower" states that he has blown part of a car engine clean through the bonnet of the car, and that indicates that there is plenty of power available from those three components (although, ideally, you want to get the timing right and not over-stress the engine!!).

One thing that "H2Opower" points out is that the injectors supplied for converting any car to run on alternative fuels such as natural gas, can be adapted to be the equivalent of Stan's specialised injectors, if a pair of high-voltage electrodes are added to each injector and fed from a pulsed circuit such as Stan's "Voltage Intensifier Circuit". The arrangement would then be like this:



A 220-page "Technical Brief" from Stan Meyer at http://www.free-energy-info.co.uk/MeyerData.pdf can be downloaded free, although it may well provide you with more information than you ever wanted to know about the subject.

Peter Lindemann's Lawnmower Running on Water Alone.

Peter Lindemann has shown that a much simplified version of Stan's injection system can operate small engines directly on water alone. Let me suggest that it is not the water which causes the explosion, but instead, that the high-power spark causes disassociation of some of the water vapour, producing hydrogen and oxygen which then ignite, turning the remainder of the water vapour into flash-steam, making the engine operate as an internal combustion steam engine. The http://www.youtube.com/watch?v=p3NE8P0sPS8 video shows a lawnmower engine operating on a spark which occurs ten or eleven degrees before Top Dead Centre:

Developed in conjunction with the EnergeticForum, the technique is based on the Aaron/Gotoluc style of operation with the circuitry shown in this video: http://www.youtube.com/watch?v=vOhNtRhJ5Rw although the "DirectHits" spark-booster shown at http://www.pulstar.com/directhits.cfm might well make the circuitry easier. Obviously, using this technology to operate a standard electric generator would be a major objective, especially since there appears to be little need to adjust the existing spark timing much.

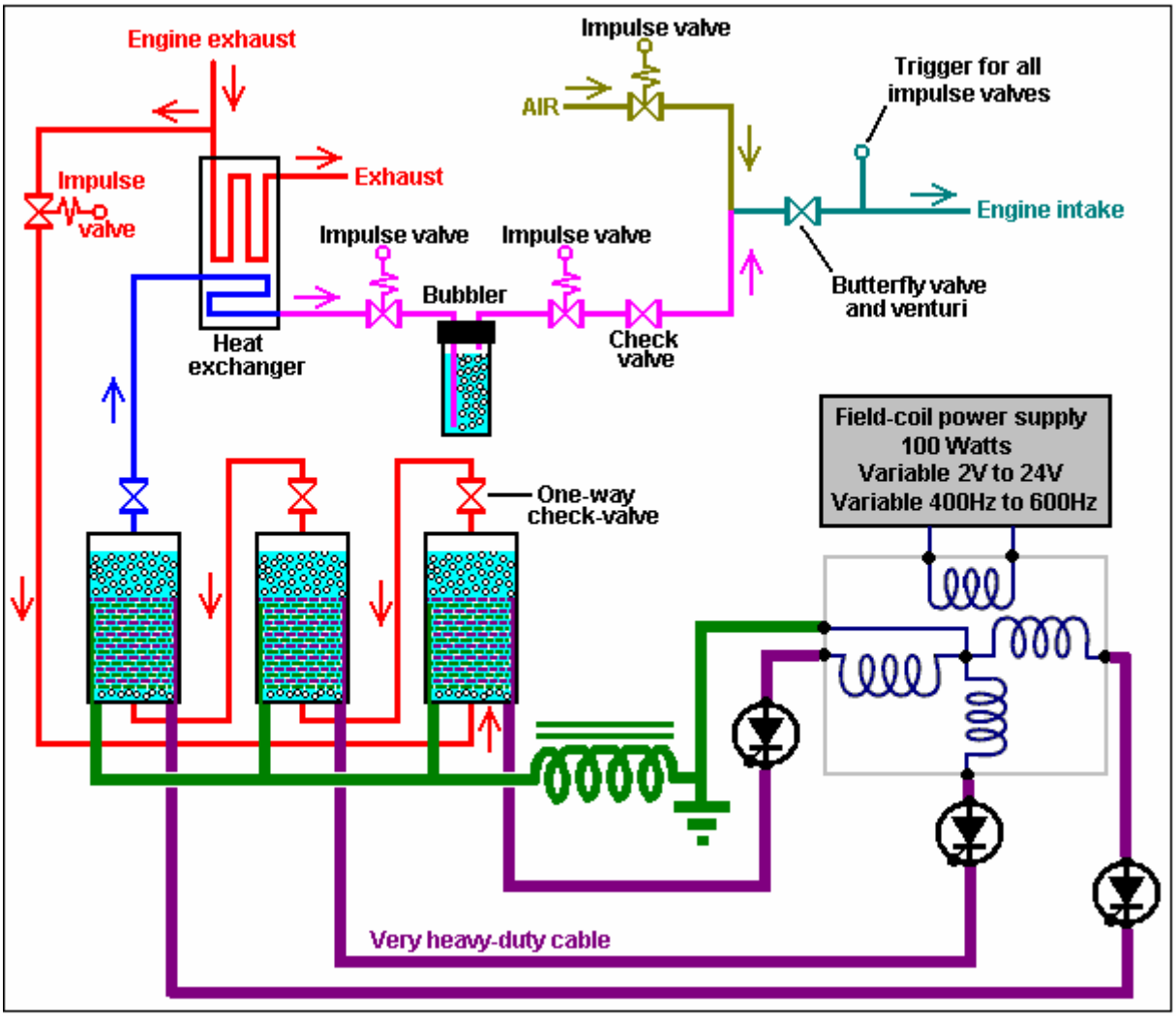
Peter Lowrie's High-Current Electrolyser System.

Peter Lowrie of New Zealand developed an electrolysis system for use with internal combustion engines. Like the previous systems, Peter feeds a spray of fine water droplets into the engine, using a carburettor, supplied by a water tank. He also feeds in some exhaust gas and heated hydroxy gas, which is a technique which is almost identical to Stan Meyer's method.

Peter also produces a very large volume of hydroxy gas with a most unusual method. He uses a delta-wound, GEC marine alternator (though he says that the alternator from a truck would do). He modifies the alternator by removing the diodes inside it and leading each of the three phase-windings out to his electronics. He uses each of the three phase-windings to power one electrolysis cell. He applies only 2 volts or so, to the DC winding of the alternator, which is about the minimum which allows the alternator to work.

The DC current supplied is less than one amp while the pulsed current to the electrolysis cells is much higher. When a snap-on ammeter surrounds the wires to the cells, a current of at least 800 amps is displayed. A point of particular interest is the inductor (coil) placed between the electrolysis cells and the windings of the alternator. Peter describes this as a choke out of a 3-phase industrial power supply. It is comprised of a laminated steel core with a sheet of copper wound around it. This is remarkably like the arrangement used by Stan Meyer and already described earlier in this document.

Peter has run a 1,600 cc car engine at 5,500 rpm with the gas output from his cells. He believes that his method of cascading the gas output from the cells through each other, produces a more active form of hydroxy gas. He also uses a heat-exchanger which allows the exhaust to pre-heat the hydroxy gas before it is passed to the engine (a method also used by Stan Meyer for running a car on water alone). Peter also uses impulse-operated valves to control the flow of gas to the engine as shown here:



The Use of Water in Engines.

I can fully understand someone having difficulty with accepting the idea that water can be used as part of the fuel inside an internal combustion engine. However, there is an old saying that "fact is stranger than fiction" and that certainly appears to apply in this case. First, it is essential to accept the fact that it can be an important component inside the engine, as this has been demonstrated by several different people who do not know each other and who live in different countries around the world. Second, having accepted that fact, the next step is to ask what exactly happens with the water.

I have come across an interesting document dated May 2009. I have been unable to trace its origin in order to ask permission to reproduce it here, so if you know how to contact the author, then please let me know. This is a report on an actual experimental tests on an internal combustion engine:

Ambient heat into work conversion.

Water-steam car, simple concept: simple Otto engine running from water mist using standard off-the-shelf ultrasonic humidifier (4 stroke diesel engine will work as well...could even perform better).

http://home.howstuffworks.com/humidifier4.htm

Replication log: Ultrasonic humidifier 0.25 litter water vaporisation per hour was set to 1/3 of its power. Use the water humidifier at air intake. Remove carburettor (or injection stuff if any) and filter for first. The way from duster to the engines air intake should be as short as possible to avoid the micro water droplets created by the humidifier collecting on the inside of the tube which feeds the water mist to the engine air intake. The engine used was a 1-cylinder 200 cc

electrical generator. The engine ran very slowly at first. In the case of a 4-stroke Otto engine, it is possible to speed it up using a propane cartridge. The engine then tends to "lock" into a higher rpm band, where the propane can be cut off permanently. The air going into the engine has a temperature of 22° C, while on average the air temperature leaving the engine was -16° C (a 38° C drop).

This is amazing – but this concept really does work!

You can make the car engine idle using this method. The water may need to be a bit warm. Ignition is not needed but it does facilitate the engine running. You may need to "turn the key" several times (based on this replication test). In the case of a diesel engine, you may need to replace the injectors with some kind of spark plug like on an A.E.R.O.P.S engine or use Firestorm plugs or any other stuff which is able to give the water an additional blast via plasma discharge, to increase performance further. You can make a gasoline generator work with water steam, also a lawn mower, etc.

Now – when it idles you have gone 'over the edge' so you need to add some more power, possible options are: adding hydrogen or HHO, preheating the water (possibly with solar energy), adding some fuel or spirit, powerful ignition or maybe some magnetic polarisation. This concept is now public and you can experiment with it on your own. Follow safety precautions and use common sense.

Do not expect high rpm results on the very first test run. Its a very basic concept which shows just **one** kind of transformation of environment heat into pressure and usable work. Also, do not pre-heat the engine before using it. The engine must not be hot!

A bit of physical background:

During the first stroke of a 4-stroke Otto engine, the piston moves down and sucks ambient air into the cylinder. During the second stroke, the piston moves up and compresses this air to a pressure of 25 bar (atmospheres). According to the "ideal gas law":

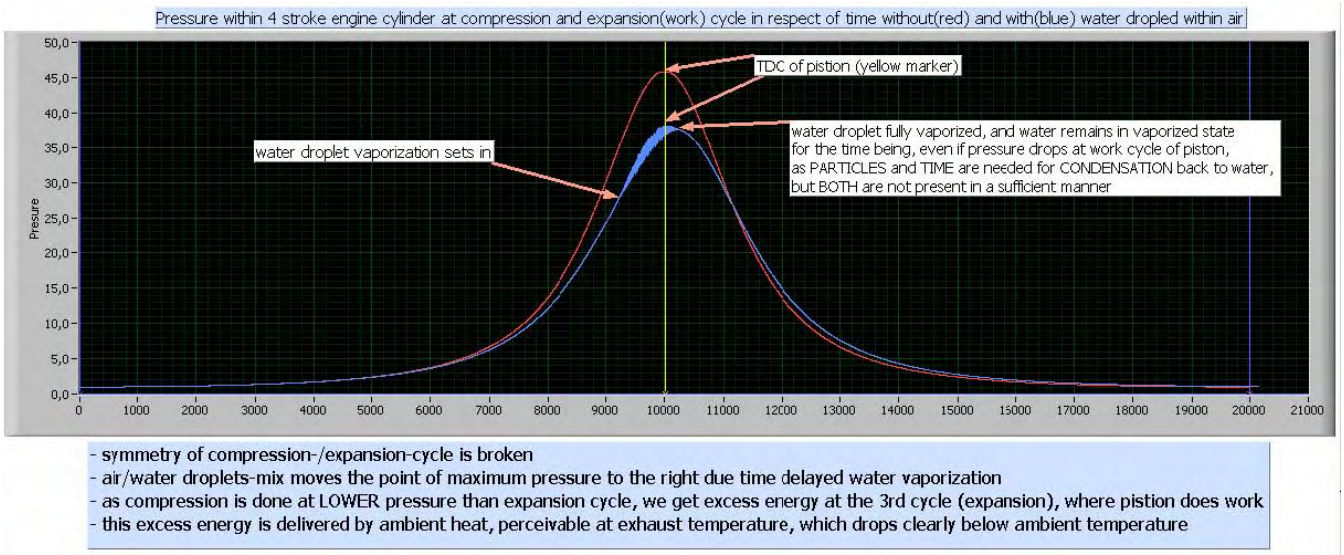
$$p \times V = n \times R(m) \times T$$

the temperature rises due to compression above its initial ambient temperature. The compression ratio of the engine will boost the temperature, typically up to 450°C to 500°C. This second stroke of an Otto engine **NEEDS ENERGY FOR COMPRESSION!** Now consider what will happen to microscopic water droplets contained in the air inside the cylinder. If the water converts instantly to flash-steam, then its volume increases dramatically, boosting the pressure inside the cylinder and powering the piston during its power stroke and storing energy in the flywheel.

Please note, that the water mist isn't water vapour...it's not a gas! It is still a liquid! The **important DIFFERENCE** becomes evident during the compression stroke! As the piston starts to move upwards to compress air which contains the water droplets, pressure AND HEAT, as described before, start to rise. BUT the compression itself is NOT able to raise the water temperature directly, as the water is still a LIQUID and therefore the temperature of the water isn't affected from higher pressure! So initially, only the air temperature rises due to compression. But tiny water droplets are present inside the cylinder, and as the air temperature gets greater and greater than the temperature of the water droplets, heat starts to flow from the air into the water droplets, heating them up! But as long the water temperature does not reach its boiling point, the droplets will not expand due to this rise in temperature and they will simply keep the same volume.

So the water droplets acting as a heat-absorber during the compression stroke! Lower heat implies lower pressure, and lower pressure during the compression stroke implies ENERGY SAVING DURING THE COMPRESSION STROKE!

Please have a look at the following graph (no losses are included in the calculation):



The graph shows the pressure within a 4-stroke engine cylinder in respect of time, starting from the beginning of the compression stroke (second one), ending with the end of the power stroke (third one). The work that is needed for the compression stroke is represented by the integral of the graph, from the very left to the yellow marker. The work which is done by the piston during the power stroke is represented by the integral of the graph from the yellow marker to the very right of the graph (20,000 at the time scale).

Let's assume that engine is an ideal one (loss-less). The red graph line represents the pressure within engine cylinder, if no water is present, and no ignition occurs. The graph is symmetrical to the Top Dead Centre (the yellow marker in the middle), so we put in energy at compression and we get the same amount back during the 3 stroke (power stroke). No loss, no gain. We get plus and minus zero.

The blue graph line shows what happens, if microscopic water droplets are present. The pressure does not rise as much as in case without water droplets, because the water works as a heat absorber during the compression stroke, and this lowers the pressure as well. The shape of the graph changes, so that symmetry of the compression and power strokes isn't there any longer, resulting in a gain of energy.

Factors which affect this process include the water droplet diameters, droplet distribution, the engine's rpm which controls the length of time, the ambient air temperature, the engine's compression ratio and even the water hardness and it's physical properties. It is certainly **not** a simple case where you can say "more water, means higher revs giving more power". For example, if the water droplets are too large, then there will be too little conversion to steam and insufficient power will be produced. Alternatively, if there are too few water droplets, then there may be no effect at all or the conversion to steam may take place too early to provide useful power.

The theory given here is a very simplified one, but should suffice for the first steps towards a better technology and a better understanding of the concepts of "energy from the environment".

It is tempting to conclude that the power gained from water droplets inside an engine cylinder are caused by the water turning into flash-steam and nothing else. However, that is probably not the case, nor is it the case that energy placed in the water by the sun heating it is extracted (which does happen) and that is the only additional source of energy.

In the scientific paper entitled "**Possibility of Liberating Solar Energy via Water Arc Explosions**" by George Hathaway and Peter Graneau, they discovered that when they produced an electric arc discharge in cold water fog: "The principal discovery made in the past two years was that it is a collection of fog droplets in the water which explodes and not the liquid water itself. The term 'fog' is meant to include not only the tiny droplets which float in air but also larger droplets which fall in the atmosphere and would be more correctly described as 'mist'.

The sole explanation of the explosions so far put forward contends that the intermolecular bonding energy in fog is less than 540 callg, the latent heat of bulk water. The bonding energy difference is then liberated in a quantum jump when the fog is formed in micro-seconds". Summing up their experiments, they conclude that "virtually all the kinetic energy developed by the explosion must be internal water energy"

In the earlier scientific paper "**The Anomalous Strength of Cold Fog Explosions Caused by High-Current Water Arcs**" by N. Graneau, he comments: "The unusual strength of explosions caused by a pulsed current flowing through water plasma was first noticed in 1907 by Trowbridge in his early high-voltage laboratory at Harvard University. When he passed as arc through a spray of water, the resulting explosion was louder than in ordinary laboratory air. During the Second World War, Früngel measured the strength of water arc explosions and published his results in 1948. He concluded that they were not caused by heat and steam and freely admitted that he was unable to explain the phenomenon. In 1969, the US Bureau of Mines issued a long report on their investigation into using water arc explosions for rock fragmentation. In one experiment, the investigators at the Twin City Mining Research Centre noticed that the energy output was apparently 156% of the input. This result was reported but treated as an experimental error".

The bottom line appears to be that using cold water mist droplets in an internal combustion engine has a minimum of the following energy additions:

1. Reduction of the energy required during the compression stroke due to the water droplets absorbing some of the heat generated by the compression and so reducing the increase in air volume during the compression.

2. On ignition, the very rapid conversion of the droplets to steam, caused by their massive surface area, produces a very rapid rise in pressure inside the cylinder.

3. The internal energy of the water caused by absorbing energy from the sun before entry into the engine may well be contributed to the power generation process.

4. Surprisingly, it has been shown that under these conditions, at the moment of explosion, the water itself contributes energy, and this process is one which most people are at a loss to explain, in spite of observing and measuring it happening.

The conclusion has to be that it appears possible that an internal combustion engine could be made to operate using cold water mist as the fuel, if a sufficiently powerful ignition plasma spark is provided using something like one of Robert Krupa's "FireStorm" spark plugs described below. Alternatively, with a lesser spark from an ordinary spark plug and the addition of quite a modest amount of a gas mix of hydrogen and oxygen from the electrolysis of water, the same effect can be produced. Consequently, although it appears so unlikely to a quick glance, that an internal combustion engine could be run on a mix of hydroxy gas, air and cold water mist, the reality is that the process is actually based on sound scientific principles and readily understood processes.

There is at the present time, a video of a car that runs effectively on just water, using aluminium plates: http://www.youtube.com/watch?v=g_2tlnf6y_k.

5. Other Useful Devices

The Panacea series on improving engine running is highly recommended. You can view that information here: http://www.panacea-bocaf.org/fuelsaving-pollutionreduction.htm

Ted Ewert's "Vortex" Air-Feed System.

Ted Ewert has developed and tested a very effective and simple device which can improve the running of some vehicles. This device works best with four-cylinder vehicles because the pulsed air intake of vehicles with fewer cylinders, enhances the beneficial effect.

This is a silent, simple and cheap device which enhances the airflow into the engine. This can have a dramatic effect on the performance of the engine. For example, Ted has an old Datsun 310 which has been sitting unused for years. Petrol ("gasoline") loses it's lighter fractions in six months or so and that makes it far less volatile and more difficult to burn. Ted's Datsun has gasoline in the tank which is five years old and the car will not run on that fuel with it's normal air intake. However, when Ted put one of his turbines on it, it starts immediately and runs fine with that old fuel. That particular vortex turbine has been dubbed "The Respirator". The Datsun has a carburettor which shows that this turbine works well with carburettors.

The "Respirator"

This simple device is a vortex tube made from a short piece of PVC pipe which has been slotted and shaped. It fits between the air filter and throttle body, or carburettor, and causes the incoming air to spin at a relatively high rate, creating a vortex. Angular velocity is crucial in the formation of a strong vortex and the air entering the Respirator from the air filter, arrives at right angles to the turbine slots, giving an immediate and powerful spin inside the tubing.

Ted says: "Most people think of a spinning air mass as having no particularly unusual properties. This is not correct. A spinning air mass has some very unique and useful properties. Standard aerodynamics, and linear Newtonian physics are unable to explain the properties of a flow of air spinning at high speed. In fact, when compared to a static flow of air within a pipe, a vortex behaves in almost completely the opposite way.

All spinning objects, whether they are solid, liquid or gaseous, contain two opposing forces: centrifugal and centripetal. Centrifugal is the expanding force travelling away from the centre axis, and centripetal is the contracting force pulling in towards the centre. This concept of dual forces is key to understanding a vortex. "Modern" physics has decided that the centrifugal force doesn't exist and now refers to it as a 'false' or 'phantom' force. This illustrates how detached from the real world academia has become and why it has stagnated.

The combination of these two forces, acting together in a vortex, create some unique conditions. One of these conditions is a laminar configuration. Co-axial laminations form throughout the vortex, creating numerous layers of air spinning virtually independently of each other. These layers are separated by zones of extremely low, virtually zero, friction and this allows them to spin at different rates.

As the vortex spins faster, the two opposing forces become stronger. This further laminates the flow as well as compressing the layers. The low-friction zones allow the compressed central air mass an unimpeded pathway for it's axial flow through the pipe. This is the reverse of the flow conditions for a straight, non-coherent air mass which has a tendency to develop friction and resistance, due to turbulence, in direct proportion to its velocity through a pipe.

Spin rate determines the degree of air compression and the linear flow rate of the mass. The faster that the vortex spins, the more it does just what we want, which is to create a dense, compressed and fast-moving flow. This is why we take the flow of air from the air box and use its speed and direction (90 degrees) to initiate the spin in our tube. This is by far the simplest and most efficient way to get the air spinning fast. The properties of a vortex are increased in step with the angular velocity. Just as a top wobbles and falls when spun slowly, so a vortex will not exhibit any strong properties until spun really fast.

As you may know, an important part of supplying air to an engine is the ability to supply a lot of air in a short burst.

This turbine creates a spinning air mass which is uniquely able to supply this air due to its virtually frictionless laminar composition and pressure built up through compression. The vortex provides compressed, dense air to the cylinder, which takes significantly less energy to draw in due to it's stored inertial energy, and it's ability to move freely in the direction of it's axis of rotation.

Between engine cycles, when air is not needed, the vortex continues to spin and build up additional pressure. This spinning air mass acts like a flywheel and stores energy which is put to use on the next intake stroke. A static air flow has no such stored energy and has to be accelerated by the engine intake stroke every time air is needed, thereby wasting energy. This flywheel property is key to understanding why the vortex works as well as it does. Unless the vortex is pulsed, or modulated, no extra energy can be developed.

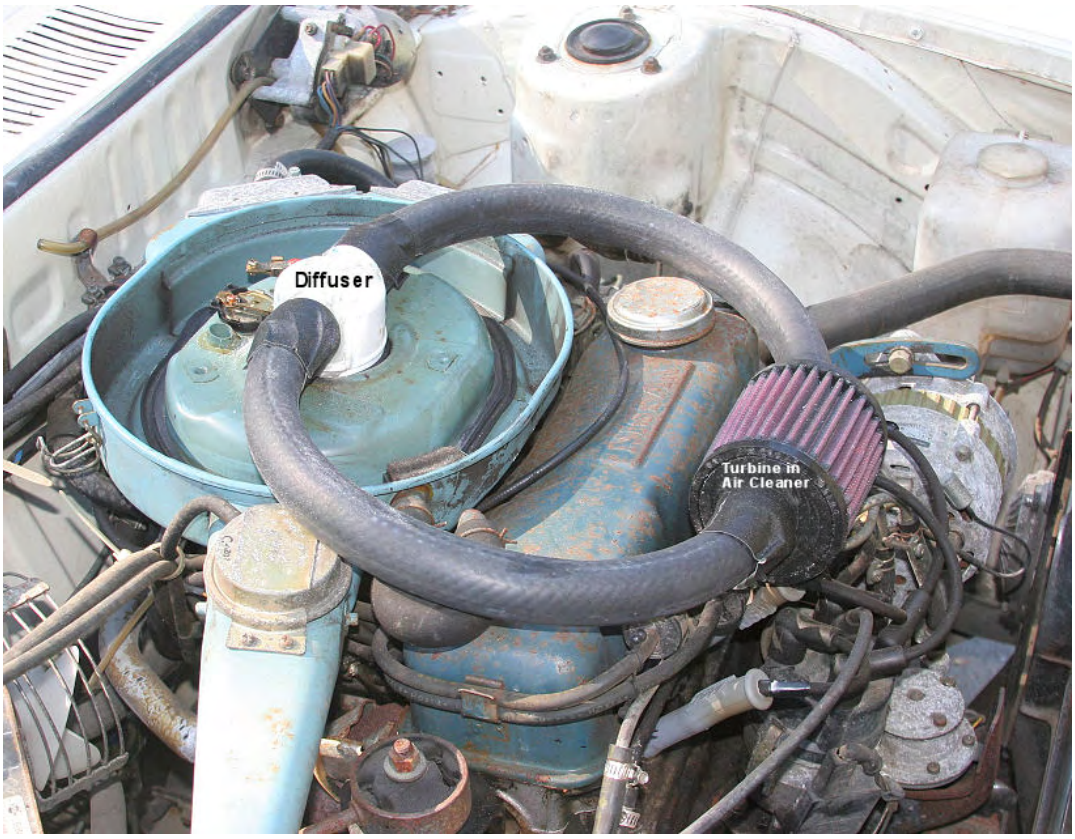
In a multi-cylinder car the flow of air becomes so steady that no effect is produced with just the turbine because there is no pulsing in the airflow into the engine. The rapidly rotating air within the turbine acts as a flywheel. When it is pulsed by the cylinder on the intake stroke, force is applied to the vortex as air is sucked down the pipe and into the cylinder. As soon as the intake valve closes, the pulse ends, the air stops its linear movement, but increases it's angular spin velocity. This is where the extra power is generated. While the intake valve is closed, the vortex continues to draw more air into the pipe, where it is accelerated and compressed, until the intake valve opens again.

Power cannot be accessed until the pulse ceases. In a steady flow this never happens. Force has to be alternately applied and relaxed. To help visualise this imagine a coil spring attached to a shaft. When a sharp pulse is applied to the shaft, the spring expands. Only when the pulse ceases, and the spring starts to contract does the power get translated into movement. This also applies to a flywheel. You can also see with the coiled spring that the pulse must be timed to coincide with the resonant frequency of the spring for the highest efficiency. Random pulses, or pulses that are badly timed, will not have nearly the effect that correctly timed pulses have.

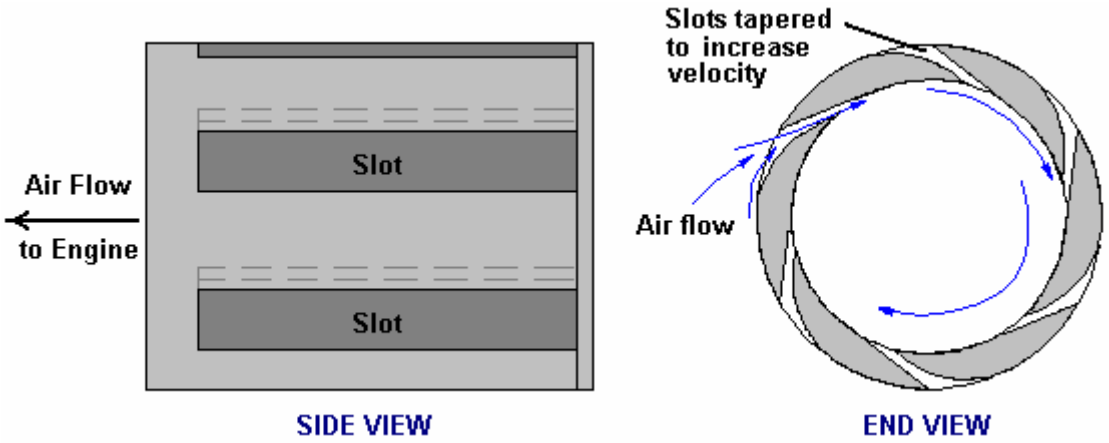
The air turbine doesn't rely as much on resonance as it does on large, well spaced pulses. This is because the power of the pulse is huge in comparison to the inertia of the air. Resonance is essential for anything that has a fair amount of mass - solids or liquids. In the case of a multi-cylinder engine, the pulses become less distinct the greater the number of cylinders. A six-cylinder vehicle barely sees any gain from the turbine, and an eight-cylinder little to none. With this type of engine the vortex needs to be modulated to gain energy.

This enhancement can be done through manipulating the shape of the intake tube. A round tube gives no gain but if the tube is "egg shaped" it produces an alternate centripetal / centrifugal pulse which imparts extra energy to the vortex. Just as the Earth draws energy from it's elliptical orbit, so in the same way, the vortex gains energy with each rotation it makes through an elliptical, or egg shaped tube.

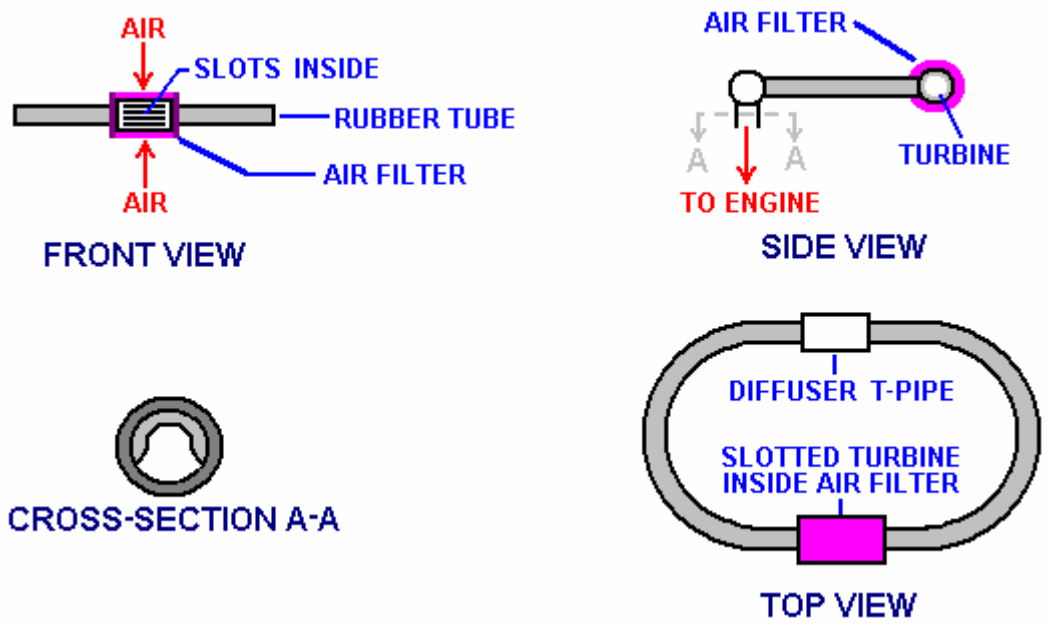
I put a slice of a smaller diameter pipe along the inner top of my tube. This small addition accomplished a noticeable increase in performance for the unit in my car. A curve in the pipe will also act like an ellipse since the rotation is compressed on the inside of the curve and expanded around the outside. Another interesting thing with the turbine is that it works much better when the engine gets hot. I notice a large increase in power in my bike as soon as the engine gets hot. This is because the heat adds energy to the vortex, just like a hurricane travelling across warm water. The heat added by the intake tube adds velocity and compression to the vortex as it spins waiting for the intake valve to open".



The vortex is created by cutting angled slots into a piece of PVC piping as shown here:



The air enters through each of six tapered slots cut parallel to the axis of the pipe. These give the air an initial spin inside the pipe and the pulsating intake of the engine, combined with the oval shape of the PVC exit T-piece, accelerates the air into a serious vortex which improves the intake to the engine, raising it's efficiency and giving more engine power.



Ted created the egg-shaped part of the final PVC T-piece by adding an extra cut section of PVC pipe to a standard T-piece as shown here:

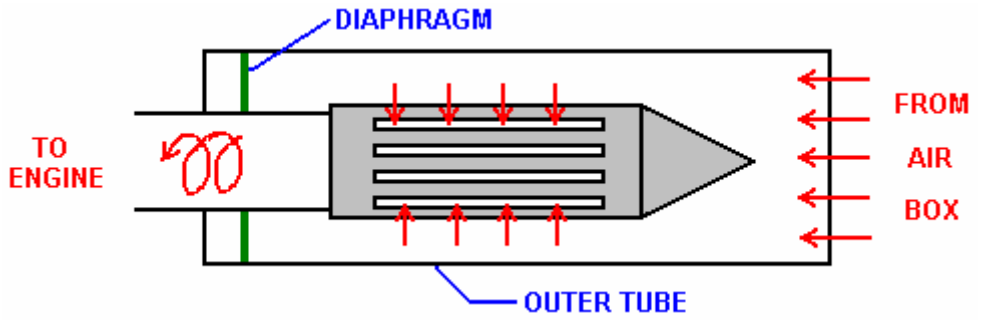
The turbine which Ted put on his bike works outstandingly well. The torque curve is extended well below it's former efficiency range. It is possible to substantially enrich the fuel / air ratio and still maintain the same mpg results as before. When the turbine is removed, both the mpg and the engine performance go way down. The turbine adds more air to the engine. To take full advantage of the possible increase in performance, the mixture should to be enriched.

Ted also put one in his 1995 Toyota Corolla car which has an 1800 cc 4 cylinder engine and a 5 speed gearbox and he is getting over 40 mpg on the open road and the low 30s around town. Originally, those figures were 34 on the open road and 27 around town. The performance has also increased very noticeably. Another nice feature is the lack of knocking and pinging under load. Performance in the mountains at high altitude is also significantly improved.

Ted has spent only a couple of months testing and evaluating this device on his cars and bike. A problem with this device is that it cannot be run directly through a carburettor, as it can with a fuel injection system. A carburettor works with a venturi which develops a low pressure zone in the throat with respect to the float bowl pressure. A vortex has no respect for a venturi and creates it's own pressure gradient which screws up the fuel metering. Ted has somewhat solved this issue by diffusing the vortex just before it enters the carburettor. Pressure and velocity are built up before the carb then sent through a diffuser.

There is still plenty of research to do with this device. And there will be many improvements and beneficial modifications still to be made to it. Ted remarks that he does not have access to any engine test facilities and that makes it difficult for him to assess accurately the results of any design variations which he may make. Ted is hoping that someone will take his design and improve it further. There is great potential in this little piece of plastic pipe.

Ted has fitted a different style of his turbine to his Toyota as shown here. The turbine section is marked “PMT” which stands for “Poor Man’s Turbo”, though obviously, you don’t need to be poor to benefit from a turbine system like this which has no moving parts:



Cam-Timing Issues in America.

A deceptively simple way of improving mpg performance has been discussed recently in the watercar forums, and that is the adjustment of the cam settings on American cars made since 1971. This sounds most unlikely, but it is a proven fact. For example, a 2004 Jeep Wrangler 2.4 litre received a 10 degree advancement on both cams, and that gave a 70% improvement on the mpg, much more engine power and an exhaust which runs much cooler.

Over the years, one man experienced a 50% to 100% improvement in mpg over a range of personally owned cars and trucks, and the emissions were improved by nearly 90%. It is not suggested that everybody should make a cam adjustment, just to be aware that an adjustment of that nature can have a dramatic effect.

Another example: “Advancing the cam timing will make the engine run cooler. I have been messing with cam timing for about 25 years. I had a 1985 Ford Ranger with a 2.8 litre engine - it was a dog. The same engine used in the 1970 Mercury Capri had lots of power. The Ranger was a dog because the cam timing was set almost 10 degrees retarded. I gave it an 8 degree advance and the Ford Ranger came to life and hauled ass. Also, after-market ratio-rocker arms help a lot on late model cars. I changed the cam timing on my 1998 Chevy truck by 10

degrees. With it's 350 cubic inch engine and ratio rocker arms installed, it gained almost 90 horsepower and brought the power band lower giving more torque because the rocker makes the cam have higher lift and longer duration on the cam which makes it breath better."

Comment from a man with 25 years experience in this field: "Cam timing is when the valves open and close in relation to the crank shaft and piston movement. The number 1 piston is set at true Top Dead Centre. At this point the degree wheel is set to the front of the engine against the front pulley at the zero degrees mark and you install a pointer mounted to the engine block pointing at the zero mark on the wheel. When the crank is turned to about the 108 to 112 degree mark, the intake valve is fully opened. That is where most engines are set nowadays. This what I call retarded cam timing. The engine seems to run well but doesn't really to seem to have much low and mid-range pulling power. When racing, you would retard a cam for high RPMs, they also could breath and had no restriction in the exhaust. The power may come in at, lets say, 3000 - 6500 RPM and advancing a cam for more torque and power, that same cam may produce power at 1000 - 4000 RPM and after all, who drives over 4000 rpm on the road?"

Another comment: "Our jeep has twin overhead cams. Advancing them does not make them stay open longer, they just open and close sooner. My reason for advancing both cams was, if I only advance the intake cam, the intake would open earlier causing more overlap if the exhaust wasn't advanced. Normally the intake valve closes after Bottom Dead Centre. Just by looking at the piston, sometimes it's almost one quarter of the way up on the compressing stroke before the intake closes. By advancing the cams, the intake closes closer to BDC. This produces higher compression. Years ago, when I did this to some of the V8s, I would switch to adjustable rocker arms and a solid lifter cam. I was able to adjust the overlap by backing off on the rockers. On an engine with one cam, advancing the cam will adjust both the intake and the exhaust. Rule of thumb is: lets say most engines are retarded by 4 degrees or more, you really don't want to advance the cams more than 4 degrees advanced. I sometimes push this as far as 6 degrees advanced for improved mpg. That is a total difference of 10 degrees from 4 degrees retarded to 6 degrees advanced. This works well with low compression engines. I also don't see a need to go to a higher compression ratio. Think about it: if you had a compression ratio of 12 to 1 and the intake closes a quarter of the way up the compression stroke, how much is compression will there be, compared to a 8 to 1 compression ratio where the full stroke compresses the mixture? If you had a engine that made it easy to get to the cam or cams by just removing a dust cover, like on our Jeep 4-cylinder, I would say to install adjustable timing gears. Then you could just remove the cover and play with the cam timing until you came up with the best power and mileage

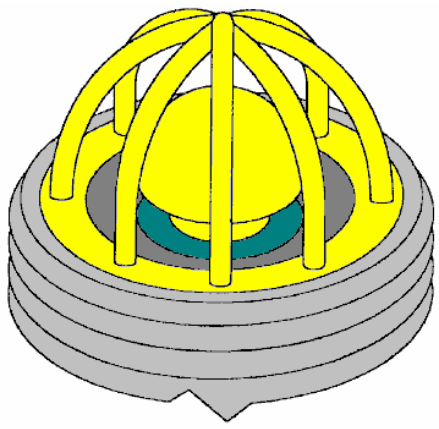
Robert Krupa's "FireStorm" Spark Plug.

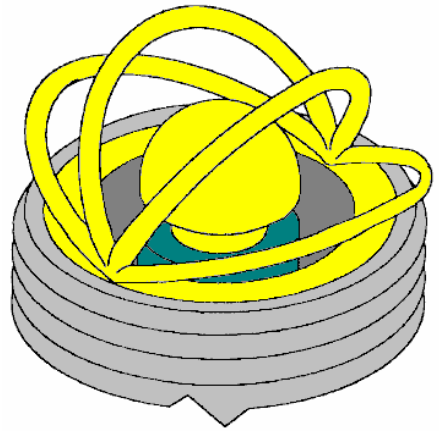
The "FireStorm" plug was developed by Robert Krupa and it is an innocuous looking spark plug which can be used to replace a standard spark plug in an ordinary production engine:

However, this plug is far from ordinary. The central electrode has been changed from a cylindrical post to a hemispherical dome, surrounded by four arched electrodes, each of which being positioned at a constant distance from the hemisphere. This allows a much greater spark area and results in very much improved performance.

The fuel/air mixture can be made leaner without any harmful side effects. If this is done using standard plugs, then the engine will run at a much higher temperature which can damage the engine. But when using FireStorm plugs, a leaner fuel/air mix actually results in the engine running at a lower temperature. Robert has measured this effect and found that under identical running conditions, the engine exhaust was 100^{o}F cooler when using FireStorm plugs. A mixture ratio of 24:1 is used rather than the current 14.7:1 mix and polluting emissions are very much reduced by the use of this plug design. Mixtures of up to 40:1 can be used with this plug.

Robert has been awarded two patents for this plug design: US 5,936,332 on 10th August 1999 and US 6,060,822 on 9th May 2000. These show variations of the basic dual arch electrodes, two of which are shown here:



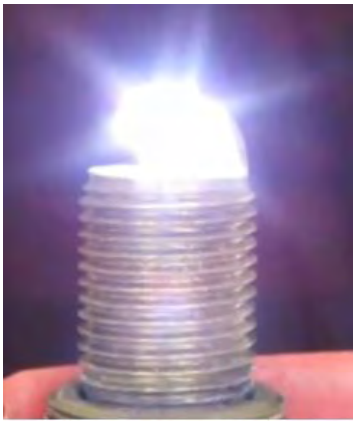


It was hoped that these plugs would go into production early in 2008 but there is no word of manufacturing starting. Robert gave Bosch of Germany a set of FireStorm plugs to test. After ten weeks of testing, their response was “This is unbelievable - we have never seen anything like this in all the time we have been building sparkplugs”. When standard spark plugs fire for a long time, the spark gap increases and the spark is weakened. Bosch ran an eight-week endurance test on the FireStorm plugs and found that there was zero gap growth. They concluded that FireStorm plugs would never wear out (which may well be why they are not yet in production - after all, who wants to manufacture something which never wears out?).

Robert’s first FireStorm plug was made in 1996 and he has encountered strong opposition to their introduction and manufacture ever since. This plug will not be popular with the oil companies as less fuel is burnt. This is probably a fallacy because, human nature being what it is, people are likely to keep spending the same amount on fuel and just drive more. For the same reason, the plug will not be popular with governments who tax fuel. The companies who make spark plugs will not like it as it does not wear out like standard plugs do. It uses less fuel and cuts harmful emissions dramatically, so it will be popular with motorists and environmentalists, if Robert can get it into production.

Plasma Ignition.

If any form of construction work with a vehicle is beyond your abilities, then the plasma ignition system offered through the http://www.bluephoenixignition.com/products.htm web site may be of interest to you. The company offers a system which can be used with any cheap non-resistor spark plugs and they claim a 40% improvement in mpg performance.



The video at http://www.youtube.com/watch?v=gYCr4p5QDEA explains the details of the system and claims that the resulting plasma from the plugs is so powerful that it can ignite cold water mist. Also, because of the very much improved fuel burn, the engine timing can be adjusted closer to the Top Dead Centre position, increasing the engine efficiency further. The very cheap spark plugs are liable to need replacing after five thousand miles or so, but that is hardly a serious overhead. The system can be used with 6-cylinder vehicles, 4-cylinder vehicles and electrical generators with one or two cylinders, and with outboard engines.

Roger Maynard's Water Vapour Injection System.

Fifty years ago car engines were not nearly as powerful as they are now. In those days it was quite common for a driver to remark that his car ran smoother and more powerfully on wet days. This was not imagination as water vapour drawn into the engine along with the air, turned to steam at the moment of ignition, and expanding provided additional thrust to the pistons while lowering the running temperature slightly.

This fact was utilised in World War II when units which were effective standard bubblers used with hydroxy boosters were added to the vehicles. Roger Maynard has built and used these units extensively since 1978, and my thanks goes to him for providing this information and illustrations.

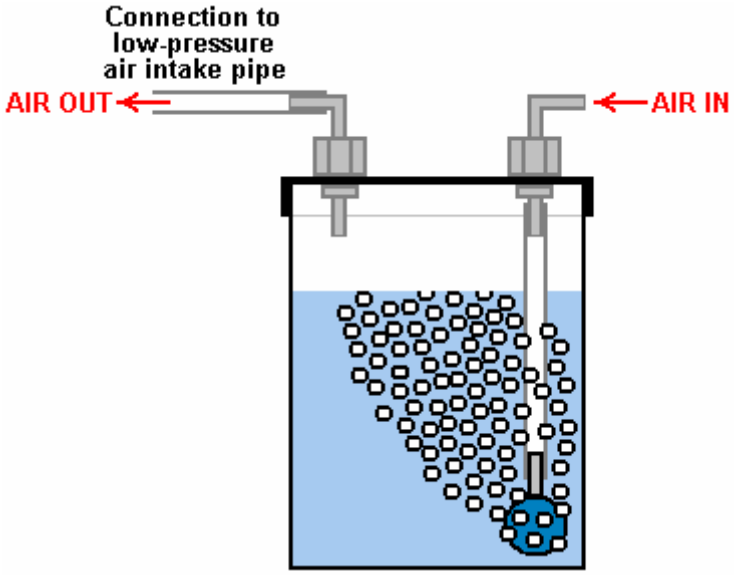
The unit is attached to the air intake of the vehicle, between the air filter and the engine. A small diameter plastic pipe is lead from there to a glass or plastic container holding water. In the above picture Roger is using a glass Mason jar with a screw-on metal lid which has a seal. Sometimes called a preserving jar, these jars are very convenient.

The air feed into the jar is by a length of the same plastic piping and terminated with a standard air-stone or "soap-stone" as used in a home aquarium, as this causes a large number of separate bubbles. It is good practice to glue the plastic fittings to the lid of the jar, but this can make the jar too airtight and if that happens it may be necessary to remove the rubber seal which is around the neck of the jar.

A glass jar has the advantage of not being affected by the heat produced by the engine. This is a very simple unit and it uses ordinary water which is not exactly a hazardous substance. The effect of using it is far greater than would be imagined. On Roger's 4-cylinder KIA car, the mpg rose from 320 miles per tank full of fuel to 380 miles around town (18%) and 420 miles on the open road (31%) which is a very marked improvement. On his 6-cylinder Tacoma shows an 8% increase around town and a 12% increase on the open road. The water is topped up every 1200 miles or so.

However, some engines are suited to the air-stone and some are not. Smaller engines may work much better if a stainless steel screw is used instead of the air-stone:

Just to clarify the operation of the device:

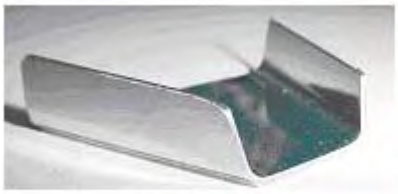


The container has a plastic elbow connection in the lid through which the outside air is drawn into the container. The air flows down through a plastic tube to either an air-stone from a pet shop, or a loose bolt in the end of the plastic tube. The air-stone has many small holes in it and these break the incoming air up into many streams of small bubbles.

There is a second elbow in the lid and the air, which is now very damp, is drawn out through it by the reduced pressure in the normal air intake of the engine. The lower pressure there is caused by the intake strokes of the engine and the air going to the engine now comes from two sources – the normal path through the air filter, and the new path through the bubbler. Most of the air flows through the air filter as normal, but there is now a small percentage which flows through the water, adding cold moisture to the airflow.

Some people feel that this couldn't possibly make any difference, but experience has shown that the addition of this extra stream of damp air can and usually does have a beneficial effect, improving the mpg, making the engine run a little cooler and generally improving the operation of the engine. It is a very simple low-tech device which does not cost much, so if you feel inclined, then try it out and see what effect it has on your vehicle, after all, if it does not provide a useful improvement, then you can easily remove it.

“Fuelsavers” Streamlining Fins.
A similar system used to be on offer from the website **http://www.fuelsaver.com.au/** where they offered small aluminium fins which mount on top of the trailing edge of the bodywork of a vehicle. The devices are reckoned to save some 10% to 12% on fuel consumption, they can be home-made, nine per vehicle is the recommended number. The device and mounting look like this:


A "Fuelsaver"

The mounting method

Robert Patterson’s “Ram Implosion Wing” Streamlining System.
The next device may not be a “free-energy” device as such, but if not, it is very close to being such. It is a structure, which when mounted on top of a motor vehicle, improves the airflow to such an extent that the fuel consumption is said to be reduced by a major factor. The device was invented by Robert Patterson and is said to create a vortex which not only decreases wind resistance but may also create a forward propulsion force.

It is claimed that the effect created by one of these wings reduces the amount of dust stirred up when driving along a dirt road and if there is a paper bag sitting in the middle of the road, it is left unmoved when the vehicle passes over it at high speed. About a dozen people are testing this device at the present time. The biggest effect is at speeds of 60 mph or more. One researcher states that he installed the wing on the roof of his Lincoln Town car using a roof rack which allowed the wing to hang over the rear window by some six inches. He states that his fuel consumption has improved from 17 mpg to 56 mpg.

Positioning of the wing, texturing of the wing surface, and the speed of the vehicle appear to be important factors in gaining an improvement. There is a research group and the website is:
http://www.pureenergysystems.com/news/2005/03/08/6900067_RamWingUpdate/

High-mpg Carburettors.
The very poor mpg figures produced by most US vehicles is a quite deliberate arrangement forced on drivers by the oil companies. In 1997, an engineer working at a US Ford company plant witnessed a 351 CID V8 started at about 4:30 pm. with a 1 litre bottle of fuel as an exactly measured amount. The next morning when he went to the factory floor, that engine was still running and had only consumed about one third of the one litre bottle. On asking about the fuel consumption, he was shown a display that read, "248.92 mpg". He was shocked and said, "This must be a mistake" but the engineer said that it was true. He then asked when they would have it ready to be put in a new Ford, he was told that he would **not** see it in his lifetime. This is company policy and has nothing to do with engineering which is easily capable of this level of performance. That 249 miles per US gallon is 298 miles per European gallon since the European gallon is 20% bigger than the US gallon.

There have been more than 200 patents granted for high-mpg carburettors. These designs all give between 100 and 250 mpg on a US gallon of fuel. Not a single one of these designs has made it to the marketplace due to the fanatical opposition of the oil companies. Last year, the Shell oil company posted typical earnings for the year, which showed that that they had made US $3,000,000 profit **per hour** for every hour of every day of the entire year. Did you enjoy contributing to that profit every time you bought fuel to burn?

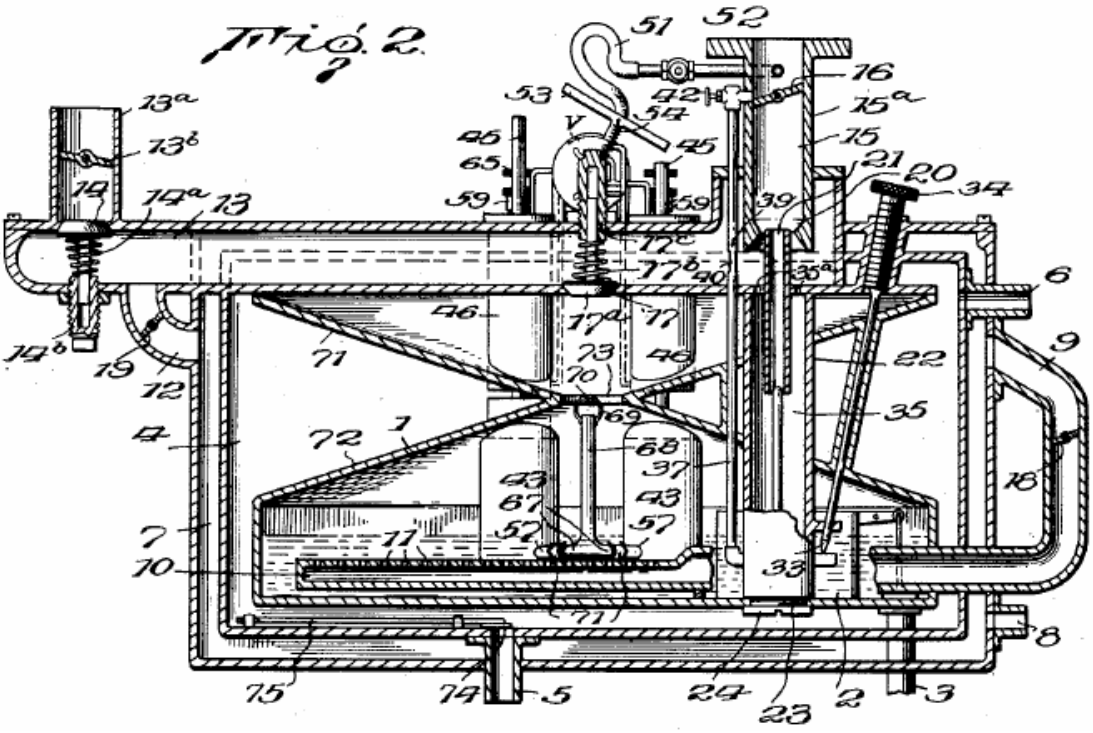
Nearly all of these high-mpg carburettor designs convert the fuel to vapour form before it enters the engine. There is no magic about this performance, just good engineering practice. It will probably come as a great surprise to you that the oil companies now put additives into the gasoline sold in the USA. They have 103 varieties of additives and I expect that they will explain that these are used to reduce evaporation in summer (as if they care about that !) and combat freezing in the winter. An “unfortunate” side effect of these additives is that they clog up any carburettor which converts the fuel to vapour form. Instead of 200 mpg, it is now quite common for US vehicles to have a 15 mpg performance and that effectively increases the cost per mile by more than ten times.

I am confident that it would be possible to design a high-mpg carburettor which deals with the additive sludge left over when the fuel is converted to vapour. In passing, the present situation gives added encouragement to stop burning oil-based products and switch to electric, compressed air, or water-powered vehicles. That is a perfectly viable option technically, but it would create frantic opposition from the oil companies and most governments which raise massive revenues from taxing oil products. The energy problem is not technical, it is financial and political.

Bob Boyce in America was handed a 3.5 year jail term for “running a vehicle on an unapproved fuel”. Bob was testing a car engine on HHO in his workshop behind his house and not on the road. Bob beat the charge in court. Since then, Bob has been told by the US Military that it is ok for him to build and use HHO boosters but he **must not** run a car on water alone. Guess who gives the US Military orders. Guess who does not care about the American people. A book on building your own high-mpg carburettor written by Larry D. Wagner can be downloaded from http://www.free-energy-info.com/Wagner.pdf but please be aware that thirty years has elapsed since it was written and gasoline in the USA is liable to be contaminated by additives which were not used thirty years ago and so, some variations may be needed. Wagner remarks that Pogue said that the new additives prevented him getting the same high mpg figures that he used to get.

In the Appendix, you will find eight out of the many high-mpg carburettor patents. The most famous of these are the three patents by Charles Pogue:

www.free-energy-info.com/PatCarb1.pdf
www.free-energy-info.com/PatCarb2.pdf
www.free-energy-info.com/PatCarb3.pdf



It would never have occurred to me that a carburettor could be used for anything other than feeding an engine, but a very experienced and practical man who lives in Alaska has used them for other things. He says:

I started out in 1976 when the first oil crunch came along and the Pogue system seemed to make the most sense at the time. We were looking to get a system started for making alcohol fuel and we were looking to develop a continuous-batch alcohol plant. This can produce feed for animals, fertilizer for the farms and fuel for vehicles or houses. It was based around growing barley and as barley is not good animal feed when grown, it needs to be cracked in some form and it was learned that this was a suitable method which had been used throughout Europe in ancient times. Two years later, the state passed a law banning alcohol fuel production. However, some of us wound up making heaters to use the alcohol produced in this way and those heaters were very efficient and they helped a lot of people, so something good did come of it anyway.

It was about this time that I met a man who had done research on the Pogue carburettor for the Sun oil company in 1928. He was very knowledgeable and helpful and in his 80's at the time. He agreed that going to alcohol, would be a better idea as it would solve many of the problems which had been engineered into gasoline fuels.

If you look at present day engines, they are as far as possible from what is needed for using this method. High compression engines will fire from heat generated under compression before top dead centre and will knock the top out of the pistons as it is an explosion rather than a burn. So, low compression engines are preferred, and firing at least 10 degrees after top dead centre. A long-stroke engine is best and some added flywheel weight increases the engine power. The incoming air needs to be pre-heated to reduce the amount of heat drawn from the combustion by the nitrogen in the air. It is the nitrogen heat content that suppresses the burn of any open fuel.

The GEET system is very similar to what we did with a few exceptions. We heated a small amount of fuel using an electric element, converting the fuel to vapour, pressurising it using the exhaust which made controlling the amount of vapour much easier. All the rest is basic GEET or Pogue.

There are many devices which I have made over the years, including the Pogue carburettor which can get very good energy efficiency from burning fuels. It is used by many industries from smelting furnaces to trains for transporting goods but it is not used by car manufacturers.

I like to study industry, because they use what works in the real world, even if they refuse to share the knowledge with others. Our group built a few metal smelters for cleaning aluminium and casting ingots. They use waste oil and roar like mad, smelting 50 lb engines in 15 minutes or less, burning less than a pint (half a litre) of oil to do it.

Allen Caggiano

If you think that this is exaggeration, take the case of Allen Caggiano whose high-mileage carburettor patent and constructional details are at www.free-energy-info.tuks.nl/Caggiano.pdf.

Allen is an inventive and stubborn man not readily blackmailed or browbeaten. In 1978 he produced the design for his first high-mileage carburettor, assembled the prototype and installed it in a 1973 Dodge Coronet station wagon. This gave spectacular results of 111 miles per gallon. Unfortunately, it failed after a short time. In 1979, he installed the second generation device on a 318 cubic inch (5.2 litre) Dodge V-8 engine and called it the “FIVS Gen II” standing for “Fuel Implosion Vaporisation System Generation 2”. This second prototype proved very reliable and produced results as remarkable as the first, getting as much as 113 miles per US gallon (European gallons are 20% larger than US gallons so 113 miles per US gallon is 135 miles per European gallon).

The installation of the ‘FIVS Gen II’ required modifications to the carburettor and the removal of the catalytic converter. This was prohibited by EPA regulations. It was therefore a violation of Federal Law. Al ignored the

regulations because he knew tail pipe emissions from his FIVS vehicle were much lower than the law required. He was upholding the spirit of the law, and he was willing to argue his case in court if it came to that. He wanted a confrontation, he wanted a chance to tell the world that his FIVS made pollution control devices obsolete. He painted the station wagon bright yellow and in bold black letters along the sides he wrote: “This Car Gets Over 100 Miles Per Gallon and Doesn’t Pollute the Air”.

On the third day of his new campaign, as Al got into the station wagon he noticed a car pulling up behind him. He got out to greet two men in suits flashing FBI credentials. While he spoke to one, the other slipped away, climbed in to his station wagon, and drove it off. Astonished, he turned to watch his vehicle going down the street. Then he heard the FBI car pulling out behind him. Al just stood there watching the two vehicles disappear around the corner. An old friend, his attorney, later called the FBI office. The FBI denied any knowledge of the incident. Angry and frustrated, but undaunted, Al said good-bye to the Dodge Coronet, and found another similar Dodge station wagon and set to work installing another FIVS system. He painted this one yellow, too, with bold black lettering.

Not long after the first car was stolen, he received an interesting offer from a California based corporation. This corporation wanted to purchase exclusive rights to his FIVS Gen II. Al asked his attorney to check it out. The corporation turned out to be a subsidiary of several other corporations, all of which were owned by an oil company. This arrangement is typical of the way contemporary monopolies are structured. Al had been doing some reading about other inventors in the past and other fuel saving devices that had never seen the light of day and he was determined that he would never allow the oil companies get control of his device.

After he refused the offer, two different FBI agents came calling. He was careful not to leave keys in his unattended vehicle this time. They informed him that he was violating Federal laws and should cease and desist. Defiant and excited that he might soon make his case in court, he told his wife, Deb, not to worry. A couple of weeks after the second FBI visit, unmarked brown paper envelopes began arriving, containing 8” x 10” photographs of the children and Deb. A child on the playground at school. A child getting off the school bus. Deb in the supermarket, and so on. She was terrified and the marriage broke up.

The FBI was behaving like the Mafia. He wrote on the side of his Dodge: “The Big Boys Are Trying To Make Me And This Car Disappear! HELP ME!”. One his oldest friends, his attorney, who was like a brother, refused to have anything else to do with him: “Wake up!” his attorney said, and then abruptly hung up the phone.

The Feds weren’t going to give him his day in court to defend his FIVS. They had stolen his first prototype vehicle and they knew it worked as claimed. He had refused to relinquish his control, so they were going to send him to prison, but not for violating federal emissions regulations. On the face of it, the Environmental Protection Agency appears to be imposing regulations on the auto makers and the oil companies in the public interest of protecting the quality of the air we breath, and the quality of the air we breath is improved over what it used to be. But in fact, these special interests often write the legislation themselves. The regulations then create a profitable new area of business which allows the special interests to increase their control over the market. The public interest is best served by creative innovation in a free market. Al was learning Politics 101 the hard way. In the business of vehicles and oil, there IS no free market. In a monopoly-controlled market, there are anti-competitive regulations, dirty tricks, and active suppression. The Big Boys protected their turf and took control of or destroyed any potential competition in the hallowed tradition of John D. Rockefeller Snr. and the Robber Barons of the past. That wasn’t hard to understand. But he was unprepared for the sophisticated tactics of today’s faceless Robber Barons.

The Chief of Police for Brockton, planted stolen cocaine in Al’s home during a drug raid that ultimately put Al in prison in 1986, for 15 years on a cocaine trafficking conviction, in spite of the fact that Al didn’t use drugs, nor associate with those who did. He fought back. In prison, he fashioned a key in the prison shop and simply let himself out. He contacted a friend on the police force and then turned himself in on the same day. This police officer friend was able to uncover evidence of the chief’s corruption. Two days later, the Chief was arrested for stealing cocaine from the evidence locker, most of which he had taken home to feed his addicted wife. He went to prison himself, which led to the reversal of over 300 drug convictions that had been decided during his tenure. The Massachusetts Supreme Court of Appeals overturned Al’s conviction. For a moment, he thought he’d beaten the Big Boys and was a free man again.

But then the Federal Prosecutor stepped in and indicted him on new charges relating to the seizure of two shot guns during the phoney drug bust. A spurious interpretation of the US Code was applied. Al was sentenced to a total of 30 years in Allenwood Federal Prison, without parole.

Al’s time in Allenwood was not wasted in self-pity or bitterness towards the Big Boys or the corrupted system that put him there. He was popular in Allenwood right away because of his role in exposing the dirty cop that overturned so many drug convictions. He developed a good relationship with the warden of facilities. As a licensed HVAC contractor, Al was able to fix the prison heating and cooling system that had never worked

properly, saving the government a large amount of money. Honeywell Corporation trained him in the use of computers so that he could operate and maintain the system. The prison had an excellent machine shop which allowed him to continue working with his FIVS devices. He designed small FIVS for the prison lawn mowers and produced numerous FIVS GEN II's, with the collaboration of the warden of facilities, that were secretly distributed outside.

Al made many useful contacts, one of whom helped him secure US Patent 5,782,225, awarded July 21, 1998, for the FIVS Gen II. He designed a new FIVS, the "Gen III", which did not violate any Federal regulations, and he put together a plan to manufacture and distribute the Gen III. And then one day he was out, free for real in 1997. Sentenced to 30 years without parole, he was suddenly released after ten years with five years parole. The Federal appeals court had finally ruled that his possession of two shotguns was legal and that it had no jurisdiction over the matter in the first place. Several years later, because he was curious, he asked a police officer friend to do a background check on him. No record of his conviction and incarceration in Allenwood was found. The stain of systemic corruption had been discretely removed.

He didn't look back and went to work to develop a prototype Gen III device, applied for his new patent, and implement the strategy he had dreamed up in prison. He was no longer politically naive, no longer the blindly patriotic American he'd once been. He did not believe it would be possible to build the Gen III in the home of the brave and the land of the free, so he made arrangements to manufacture parts in the Ukraine, a former satellite of the defunct Soviet Union. He would then assemble the devices in Mexico. He had developed a global perspective in Allenwood. His network of supporters and investors was now called: "FIVS Gen III International" and he set up a website which was generating 70,000 hits a month from all over the world. He also offered the complete blue prints for manufacturing the earlier FIVS Gen II as a free download from his site so that anyone who wanted to could build their own. He thought this might distract the Feds and tie up their manpower as he implemented the Gen III strategy.

By 2002, the delivery date for the first beta testing group was set. The parts were shipped from the Ukraine to Mexico where they were assembled. It was necessary for Al to travel South of the border to oversee the operation. He made the punishing drive from Massachusetts to Mexico several times in his FIVS equipped Pontiac Catalina and it functioned flawlessly, delivering more than 70 mpg with its 400 cubic inch engine. His friends warned him not to drive alone, but he made the last trip by himself and on the return leg of the journey, he noticed an 18-wheel truck following him. The intentions of this truck were soon obvious when it overtook him and forced him off the road. Al anticipated the manoeuvre, however, and was able keep control of the Pontiac. He breathed a sigh of relief and continued on, believing he'd outwitted them once again. He made it all the way to Massachusetts and was nearly home again before the truck found him a second time and caught him unawares. The Pontiac rolled over several times, but landed upright. The driver's side door was crushed and the roof caved in, but the car still ran and Al was able to drive it home without further incident in spite of his injuries. He had to be cut out of the car with a torch. He had several broken ribs and a punctured lung and was immediately rushed to hospital.

The Gen III's for the first group of licensees were shipped from Mexico on time, however, by means of several different shippers. Some devices for US licensees were shipped via United Parcel Service. A total of 137 units were shipped around the world. Only those that went UPS in the continental US and Canada, a total of 44 units, did not arrive at their destinations. Every shipped item has a tracking number, of course, and when Al inquired about the missing 44 units and provided the tracking numbers he'd been given, he was informed the numbers he had did not exist.

The attempt to turn him into road kill was not completely unexpected, but Al was shaken just the same. He maintained his bravado, however, while friends and sympathisers reacted more predictably. When the intent of the suppression escalated from malicious to deadly, most began to slip quietly away and it was 'High Noon' again. Another complicating and aggravating factor was the appearance of a discussion group at the Yahoo website called "Get 113to138mpgNOT". This Yahoo Group was established by an individual calling himself "David Rodale". He was not a Gen III licensee. He (or she) was a freelance public servant dedicated to helping those who had been ripped off by the promoter of impossibilities, the unscrupulous scoundrel, Allen Caggiano. He provided advice and counsel to those disappointed licensees who had not received their Gen III devices. He assured them that they could find justice in the courts. Al spent much time and energy fighting back against this defamation.

Al was fully recovered from his "accident" by this time and had repaired the Pontiac. He was feeling every day of his 59 years, but he soldiered on with a grim determination towards whatever final confrontation awaited him. When a careful, bloodless voice on the phone proposed a compromise one day, he felt ready to bargain. His website was experiencing growing traffic. T he voice told him that if he would just remove the Gen III from his site he would be left alone. It felt like a small victory, but he didn't relish the idea of backing down. If Gary Cooper had received such an offer in High Noon, he would have taken it.

He knew a bargain with the devil could never work in his favour, but he had to catch his breath, so he played along and removed the Gen III from his web site. It was a strategic retreat. If they would leave him alone, the beta testing of units already out there could go forward. The program was smaller than he had originally intended, but it was a start and if he could relax and gather his data, then he might ultimately win the game. However, a careful examination of the FIVS in his Pontiac one afternoon made his heart jump into his throat. He found a tiny hairline crack in the aluminium/titanium alloy canister. This Gen III unit had many thousands of miles on it. It presaged a potential disaster and he immediately notified all the licensees of the problem and recalled the units. He worked feverishly and discovered that he didn't have to redesign the canister. A simple alteration appeared to be the solution.

Al was then told that he required immediate surgical treatment. After the fact, this prognosis was shown to be false. While the surgery was underway, Al suffered a stroke. His heart stopped and he was technically dead on the operating table. In addition, the surgeon had damaged nerves in his spinal column. It was not his time to go, however, and he revived, but then lay in a coma for 30 days. When he finally came around in his hospital room, feeling more dead than alive, he was astonished to discover that he could not move his legs. It is possible, that Al's medical experience was not accidental.

On the local TV noon news on a sunny day in the spring of 2003 he watched a dramatic live report of a SWAT team in action. They were closing in around a familiar looking building. He thought to himself: "Hey! That's looks like my condo! Hey! That IS my condo!". He watched the police seizing his yellow Pontiac in the parking lot as the Channel 7 reporter explained that Chelmsford, Massachusetts, resident, Allen Caggiano, had defrauded investors in a fuel saver scam and then fled the country. He didn't see how that could be true since he was in the Intensive Care ward of the local hospital, not 20 miles away.

Meanwhile "David Rodale" at Yahoo Group "Get 113to138mpgNOT" had found 20 disappointed Gen III licensees and was patiently building consensus for legal action at the state level in Massachusetts. It wasn't easy to turn disappointment into outrage and a desire for revenge. In spite of the resources available to the Big Boys, they hadn't been able to otherwise identify most of the testing program licensees. Al returned home to his condo to find his Pontiac with the repaired Gen III's in the trunk gone from its parking space. His premises had been ransacked, his computer hard drives removed. With his mind foggy from pain killers, Al tried to concentrate on getting used to a wheelchair. Nurses from the Visiting Nurses Association were with him around the clock. Gradually he stopped using the pain killers. He began to notice sensation returning to his legs.

Even as he felt himself improving, his diabetic condition inexplicably worsened. Twice he was rushed to the hospital in a comatose state. The third time this happened, a nurse checked his pill box and discovered insulin pills that should not have been there. He was now taking insulin through injection, but the old insulin pills were still in the medicine cabinet and had been put in his pill box with his other medications. The result was insulin shock. The nurse, Michele, who had done this, not once, but three times, did not again appear for her shift. Al tried to reach her to ask for an apology for her mistakes, but she had vanished. The Visiting Nurses Association denied having any record of her employment.

"David Rodale" was having success convincing the disappointed licensees to file suit, and with the newly acquired information about the FIVS Gen III International operations taken during the SWAT assault, a postal inspector launched a preliminary inquiry into the feasibility of action at the federal level for mail fraud. Rodale was confidant that the threat to society posed by Allen Caggiano was now neutralised. He announced to the Yahoo Group members that he'd done his best and there was nothing more to do. He would leave the Yahoo Group in place for a while, but he planned to take it down in a couple of months. He was sorry that so many people had been taken in, and he hoped they'd be less gullible in the future. He was glad he could help.

That summer, the judge dismissed the charges against Al in the Massachusetts court. His lawyer petitioned for the return of his property, the Pontiac the local police had seized a year earlier. He was told it had been taken to Washington, DC, and was being examined to determine if it violated any Federal regulations. A grand jury in DC was convened to investigate the Federal charges of mail fraud, but it failed to return an indictment. The licensees who had paid their money and signed their licensing agreements had agreed to assume the risks of a testing program and most of them did understand that part of the risk involved the historical efforts of the oil/auto cartel to suppress new technology that could affect their profitability or control over the markets which they ruthlessly dominate.

Through his attorney, Al received an offer for exclusive rights to the Gen III. The amount of money involved beggars belief, and suffice it to say that Al again refused, as he had done in the early 80's when an offer was made for the Gen II device. The Big Boys have never attempted to prosecute him for the violation of Federal Emission Control regulations. He is clearly guilty on this score so far as the Gen II is concerned. To do so would result in the exposure of the fraud they are perpetrating on the public. Their technology is obsolete. As Al has pointed out on his website, they do not want a reduction in the demand for oil. This would mean a reduction in oil company profits. If the consumer used half or less of the oil now being used, government tax revenues would be

reduced accordingly. If the Gen III were to become available the public interest in fuel economy and clean air would be served and Allen Caggiano would become rich beyond wild imaginings, but the oil/energy cartel and its partner in government would suffer. Therefore, the Big Boys will continue to do all that they can to stop Al and his FIVS and to keep the public ignorant of any technology they do not themselves control. If they don't control it, and if you don't buy it from them, then it doesn't work, or it's a fraud.

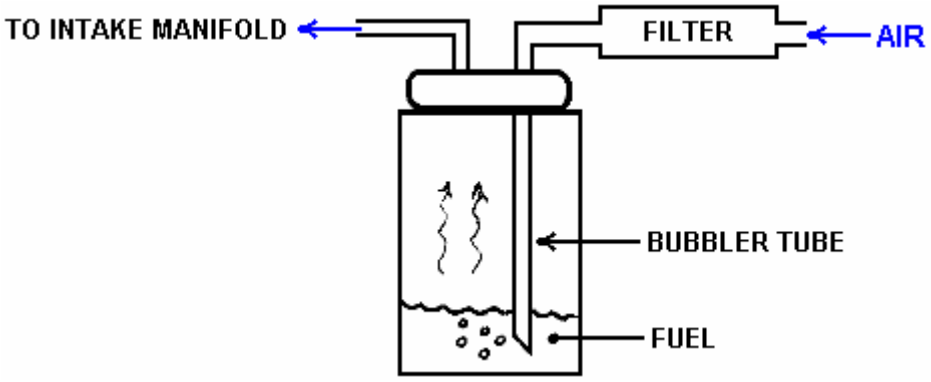
The "FIVS Gen III International" enterprise has been successfully suppressed. Al's fight to manufacture and distribute his invention and enjoy the great wealth it would have given him is over. The Big Boys have broken his bank, and his health. The struggle has nearly destroyed him. The money would have been nice, but it never was the money that kept him going. He is now donating his work of a life time to the American people. The Big Boys can harass, intimidate, and attempt to kill one man and his American Dream, but can they do the same to many thousands of Americans and others around the world? Al now gives away his "FIVS Gen III" plans free.

Al will allow his patent application for the Gen III to expire. He can no longer afford the large investment required for a patent here and in other countries. His main concern now is to prevent the FIVS from being patented by anyone and to keep the device "open source", so to speak, so that it cannot come under the control of the Big Boys and will remain freely available to the public. Though Al will not profit from his invention through licensing fees or royalties, there is considerable satisfaction for him in knowing that the Big Boys have not and will not ultimately win this game and destroy the work of a lifetime.

Reference site: http://peswiki.com/index.php/OS:Caggiano%27s_Fuel_Vaporizor_System_%28FIVS%29
Forum: http://groups.yahoo.com/group/fivsgenv Patent: whttp://www.free-energy-info.tuks.nl/US5782225

Fuel Vapouriser Bubbler.

There is one technique originally fromt www.alternativefuelsnow.com and reproduced here by kind permission, which appears to be effective in spite of the additives. The method is very much like the technique used by Roger Maynard for adding moisture to the incoming air as described above. The difference is that instead of using water in the container, petrol is used. Improved performance of up to 60% has been achieved and experiments are continuing. The general arrangement is like this:



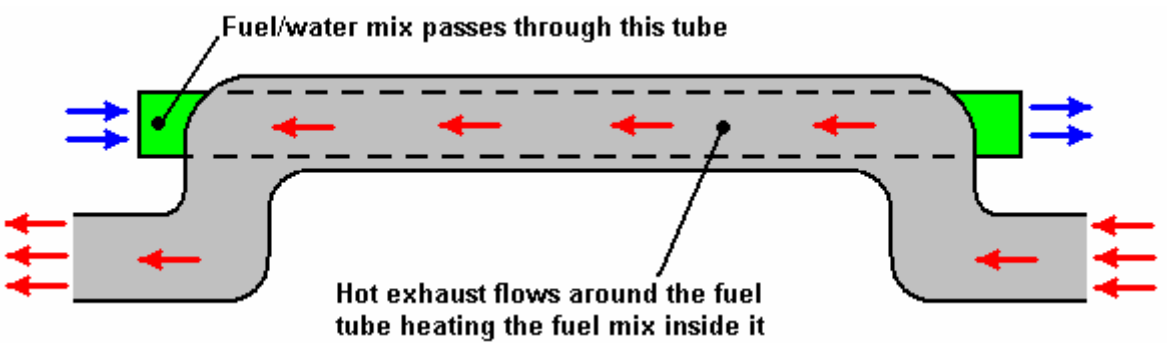
You will notice that the fuel level inside the container is kept fairly low in order to allow considerable space to contain the bubbles, allowing them to pop and not be drawn into the air tube which feed the engine.

Vortex Fuel Reforming.

This is a very important technology which has been around for more than a hundred years. The objective is to increase mpg not just by the vaporisation of the fuel but also by "cracking" the water / fuel mixture into smaller molecules before being fed into the engine. This is more advanced than the fuel 'conversion-to-vapour' technique of the high-mpg carburettors. To get a better understanding of this, you can try a Google search for "fuel reformer" or "steam reforming" which will provide additional information which may help you to understand the basic principles.

The fuel-reforming method can be highly effective and its effectiveness has been proven beyond all doubt with designs from Cal-Tech, Philips Petroleum, Nissan Motors, NASA, universities and other very serious contributors. Some years ago Cal-Tech spent millions proving that on board fuel reformers would give us all better fuel economy and cleaner air. They did long-term testing on buses and cars to provide proof. They teamed up with the very large auto-parts supplier Arvin Meritor to put them in production vehicles. Then "One Equity Partners" bought out Arvin Meritor's division that did all the final work to get fuel reformers in all our vehicles. They created a new company, EMCON Technologies, and that company dropped the fuel reformer from their product line, not because it did not work but because it did work.

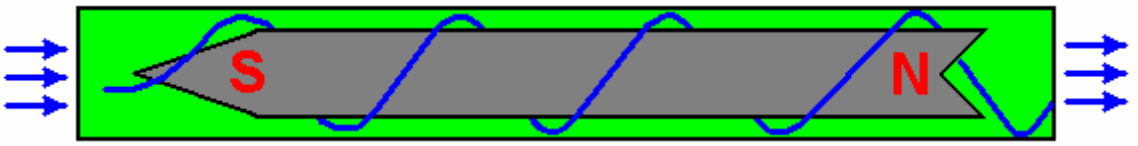
There are various techniques for achieving this process. One which is easy to understand is shown here:



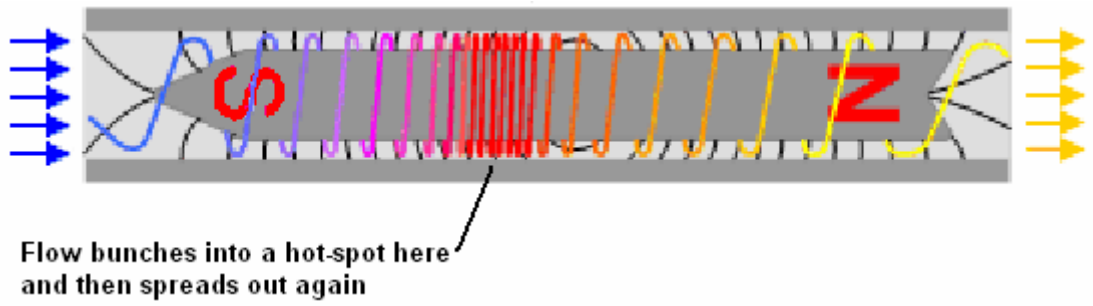
Here, the standard exhaust pipe is given a kink to move it clear of its normal run in order to allow an extra straight pipe of smaller diameter to be positioned inside it so that the hot exhaust gases are used to heat the incoming fuel

flow. This is a useful energy gain as it uses some of the waste heat, raising the overall efficiency of the engine very considerably.

This extra fuel-flow pipe has a solid magnetised ferromagnetic metal rod mounted inside it, blocking off most of the pipe area. This change in available flow area causes the fuel flow inside the pipe to speed up, and as well as that, it causes the flow to spiral around the rod in a vortex flow:

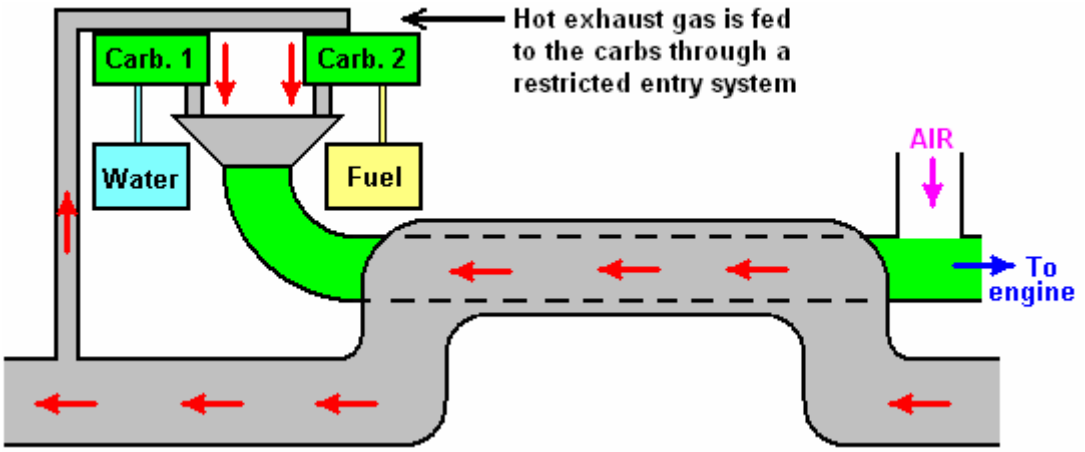


However, the magnetism of the solid rod, causes a most unusual effect and instead of the helical gas flow being as shown above, a highly uneven flow pattern is created. This causes the fuel flow to bunch up in the centre of the tube, generating a hot spot which creates quite unexpected results:



The really incredible result of this peculiar effect is that the fuel mix exiting from the tube, contains chemical components which did not enter the tube - impossible according to present day physics. This goes to demonstrate once again, that we really don't yet understand the world in which we live.

The fuel mix for use in this system is best provided by two tiny carburettors, one feeding a fine mist of water droplets and the other a fine mist of fuel droplets. These are fed directly into the intake of the fuel reformatter tube. These carburettors are of the miniature type used for radio control model aircraft and their venturi intakes are blanked off with a plate with a small hole in it. Air is **not** fed into the reforming tube – after all, this is a fuel reforming system. The air is mixed in with the reformatted fuel after it exits from the reformatter, as shown below. Some of the hot exhaust gas is fed into both of the carburettors in order to help prepare the mixture for the reformatting process. The blanking plates on the carburettors are there in order to reduce the amount of the exhaust gas being drawn in with the fuel:



The use of carburettors is important since using a bubbler as suggested in the free plans on the internet, creates problems as the lighter fractions of the fuel get taken first which is exactly what we don't want to happen. The

carburettors have the massive advantage that they feed all fractions of the fuel together and so the remaining fuel is always in the correct proportions.

The ratio of water to fuel (typically gasoline or diesel) can be adjusted over a very wide range, with some people using 90% water. Actually, there have been claims of running on 100% water, using multiple reactors in parallel with energy being drawn either from transmutation of elements or perhaps spin interactions with the local environment. Jean Chambrin's patent give details of running only on water.

There are several forums where members are researching and using various designs of fuel reformers, with the GEET designs being popular. The http://tech.groups.yahoo.com/group/VortexHeatExchanger/ forum is one such research forum and one of the files there lists 214 different patents for these devices. There are several different types of reformer. Another forum is http://tech.groups.yahoo.com/group/geet-pantone/.

Almost any hydrocarbon fuel can be used - vegetable oil, old motor oil, etc. the normal fuels are the most popular. A forum member named 'bryishere' says in a YouTube video comment: "Everyone should really try this. IT WORKS. I have spent a lot of time on this device. It's very simple. Just follow the plans and experiment as much as you can. Currently I am using 90% water and 10% crude oil/waste oil on a 1-ton, 1969 Chevy truck ..... Get out in your shop !!!!!"

These devices are often used on stationary generators. These devices have been popular in France for some years now. Jean Chambrin found that the gases needed to swirl inside his reactor in the same direction that the crank was turning.

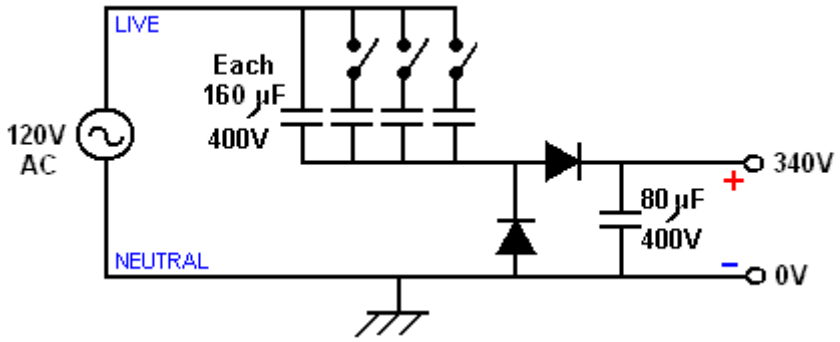
There is a massive 175 Mb file called 'FuelReformerTechnology.zip' which you can search for and download from the internet if you are very enthusiastic. That file contains the contents of more than 220 patents and applications. These patents are also listed in the 'Files' section of the Yahoo VortexHeatExchanger forum mentioned above.

Electric Power.

We have been talking about internal combustion engines as there are so many of them around the world, but electric power is an alternative. One hi-tech option is a Chinese SUV-style 5-seat saloon car which has a top speed of more than 120 mph, a range of 250 miles per charge and a re-charging time of just one hour, thanks to newly developed iron batteries which can be recharged up to 2000 times before they start to degrade. Also, these new vehicles look perfectly normal. I suspect that the cost of such a vehicle may prevent most people from buying one.

However, low-tech solutions can work well for some people. For example, Nader Hoville in Hawaii has adapted a truck for electric operation. He uses a battery bank of eighteen 8-volt Exide golf-cart batteries of 150 AHr each, giving a 144-volt 150 AHr bank. This provides a forty mile journey in fairly flat terrain and twenty-five mile range in hilly country and while that would not suit many people, it suits places like Hawaii where travel distances are not normally very great.

Running costs are lowered still further because Nader has a solar panel array of ten 250-watt panels which provide more than enough electricity for his house. As his solar system has larger capacity than is needed for household equipment, the additional capacity can be used to charge his truck's battery bank for free. Nader sometimes charges his truck's battery bank from the 120-volt AC mains when he is away from home, and to do this he uses an ultra-low-tech charging system which does not use a mains transformer. If, like Nadir, you are not familiar with electrical circuitry, then reading chapter 12 may help. The circuit is like this:



This is an unusual voltage-doubler circuit where the mains input has an average of 120 volts and a voltage peak of 170 volts and as that occurs on both sides of the output, the output is around 340 volts of pulsing DC current which both charges and de-sulphates the batteries.

The current fed to the battery bank is controlled by the capacitors inserted in the mains 'live' line. These capacitors have to be very high quality and typically cost US $10 each. They look like this:

each of them being 80 microfarads in capacity and rated for 400-volt use. Nader uses these in pairs, giving 160 microfarads for each pair and each pair passes 133 watts of charging power to the battery bank. These are non-polarised capacitors (meaning that they don't have a Plus and a Minus side) and electrolytic capacitors will NOT do and are liable to explode if used in a circuit of this type.

The diodes are 40-amp 400-volt rated and they cost $3.50 each. Please understand that this circuit has DANGEROUS voltages EVERYWHERE and those voltages can cause injury or death if you are careless. So please be warned, this is not a circuit where carelessness can be tolerated at any time. The charger is connected to the batteries BEFORE the mains supply is connected, and the mains disconnected before the charger is disconnected from the battery bank.

This is not a circuit which I see as particularly safe to use. You will notice that one of the terminals which you handle and connect to the battery bank is actually one side of the mains. I am not inclined to recommend that anyone handles mains wires. It is vitally important that it is the "Neutral" mains wire, which in theory is connected to ground and so is perfectly safe. However, in practice, that is by no means always the case and it has been known for the 'Neutral' mains wire to float 180 volts away from the earth voltage (which you are standing on!). So, I suggest that the "Neutral" wire is connected to a real, physical, earth connection and that mains power is not switched on while that connection is being handled.

Having said that, the circuit works very well for Nader, and is silent in operation. The charging power depends on the number of capacitors used and so switching in the extra pairs of capacitors gives the choice of charging at 133 watts, 266 watts, 399 watts or 532 watts. Nader's build looks like this:

The Weird Nature of Water and Advanced Electrolysis.

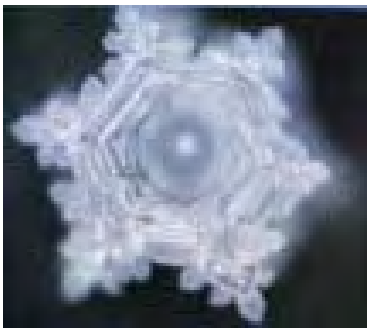
This chapter has been dealing with systems for enhancing vehicle operation with the use of water, so it seems appropriate to finish it with a brief note on water itself. To a casual glance, it appears that we know all about water. It's composition is H_2O and when it breaks down, we get two hydrogen atoms and one oxygen atom - right? Well maybe, and maybe not.

The longer you spend looking at systems which use water, the more you get to realise that water is by no means as simple as you might initially think. There is a much maligned branch of alternative medicine called "Homeopathy" which is based on giving patients very dilute water-based solutions various chemicals. Sceptical investigators have run professional-quality tests intended to show that homeopathy is fraudulent and has no medical benefits whatsoever. Unfortunately, the tests did not work out the way that the investigators wanted. The tests showed that there actually was some benefit from the treatments being examined, and unfortunately, because a placebo control group was being used, the placebo effect was definitely not the cause of the effects recorded during the trials.

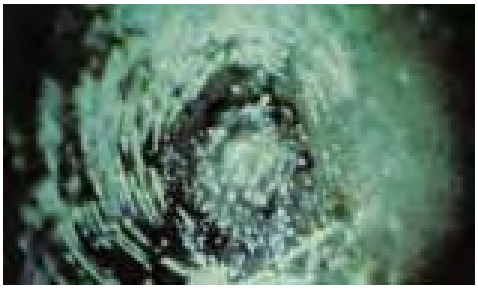
Determined not to just accept the results which went against their expectations, the testers started testing ever more dilute samples on the patients. They eventually got down to the level where there no longer remained a single atom of the chemical in the liquid being fed to the patients, but to their consternation, the medical effect remained. They tried water which had never had the chemical in it, and there was no medical effect. They returned to the apparently "pure" and definitely chemical-free water and the medical effect was seen again, in spite of the fact that there was not even one atom of the chemical remaining in the water.

This showed clearly that the water was different after having had the chemical in it, even when no chemical remained. They were forced into the opinion that water has "memory". That, of course, is a conclusion based on the facts which are hard to explain. You may wish to deduce something else from those facts, and that is entirely up to you - just be aware of the facts.

Studies carried out by Mr Masaru Emoto here: http://www.vidaplena.net/Videos_VP/Videos_A-B/Agua.htm have shown that the thoughts of ordinary members of the public can alter the structure of water without there being any actual physical contact with the water, and many other very important effects. If the water receives positive thoughts and is then frozen, the resulting crystal structure will be like this:



While on the other hand, if negative thoughts are aimed at the water, whether just by looking at it and thinking, or by writing those thoughts down on paper, the resulting crystal shape is quite different when the water is frozen, as shown here:



It is not all that startling if you consider that the quantum mechanics researchers have been saying for a long time that experiments can be affected by the observer. People who build Joe Cells which operate through environmental energy focused by specially treated and structured pure water, record the fact that certain people can affect a Joe Cell in a negative way from a distance of fifty yards (or metres) away.

Personally, I am quite sure that we do not understand the fundamental nature of our environment and that we have very little idea of how we as individuals impact on our surroundings.

There is an extremely honest and reputable researcher called George Wiseman who operates through his company Eagle-Research (http://www.eagle-research.com/). George is very experienced in producing "Brown's Gas" and he publishes excellent instruction books on the subject. The really interesting thing is that Brown's Gas is produced from water and that gas has the most remarkable properties which are not readily explained by our present day "conventional" science. When Brown's Gas is used as the gas to power a cutting torch (like an oxy-acetylene torch) the resulting flame is nearly colourless and can be waved across a bare hand without any ill effects - the hand is not burnt. But when applied to a fire brick which is intended to resist high temperatures, it burns a neat hole through it. It will vaporise a tungsten rod which normally takes 6,000^{O}C to do that, which indicates that the flame temperature depends on what it touches (!).

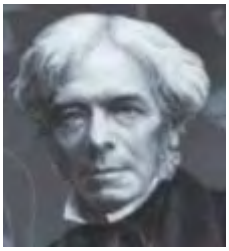
It can also weld aluminium to aluminium without the need for an inert gas. It will weld aluminium to brass and it can weld a steel rod to an ordinary building brick. It can fuse glass to a building brick. This is not "normal" for a chemical combustion reaction, showing that Brown's Gas is not a "normal" chemical substance. As Brown's Gas comes from water, does that perhaps suggest that water is not a "normal" chemical substance? I will leave you to make up your own mind about that, perhaps helped by the following, most enlightening presentation by Moray B. King.

A Presentation by Moray B. King

Moray King has produced a substantial document covering many aspects of free-energy with special emphasis on the more unusual systems and some of the difficult-to-explain things which people have discovered. As Moray's presentation has 166 pages containing many pointers to video clips and specialist web sites, this is just a brief summary of his pdf document which can be seen in full here: http://www.free-energy-info.com/MorayKing.pdf and which has the uncommon feature of an icon at the top left hand corner of each page, and if you click on that icon, then additional comments can be seen. He starts with:

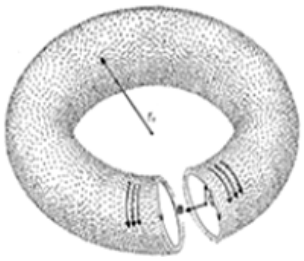
Steve Ryan, a man living in Auckland, New Zealand, who was shown running a motorcycle on modified water, shown on a cynical TV video presentation at http://www.youtube.com/watch?v=POJQKg9CRJc where the complete ignorance of the presenters is demonstrated yet again. Steve has disappeared from public view for a very long time now. However, Moray raises the most relevant question: "How can water store enough energy to make it appear to be a fuel?" The first step in answering this question is to realise that causing cavitation (tiny stress bubbles) in the water inside an electrolyser, produces excess energy, and so, circulating the water inside an electrolyser while causing cavitation in the water is a major step forward, and one which seems to indicate that the majority of the energy in HHO gas does not actually come from the hydrogen. Instead, zero-point energy gets trapped in the water when the turbulence caused by the circulation, charges the water electrostatically, raising it's energy content as it circulates repeatedly through the electrolyser.

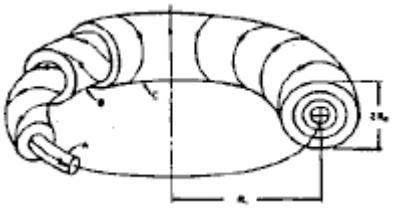
Mark LeClair, the founder of the NanoSpire Corporation (https://nanospireinc.com/), discovered a microscopic crystalline form of water which has an extraordinary energy density. This crystalline form of water is similar to the microscopic plasmoids discovered by Ken Shoulders and the larger plasmoids discovered by Adamenko's team at the proton-21 laboratory in the Ukraine. When a plasmoid strikes any element, the result is transmutation of that element. The excessive energy and most unusual properties of Brown's Gas come from charged water gas clusters which are stored in a stable ring-form of the microscopic water crystals. This feature completely by-passes the standard electrolysis of water as examined by the great Michael Faraday where more energy is needed to split water than can be regained when the resulting gas mixture is burned. This is an entirely different process as explained at http://peswiki.com/index.php/Video:Water_as_Fuel_%28via_ZPE%29 in considerable detail on Sterling Allan's PESWiki web site, and it shows why water can indeed act as a fuel. If you search for "water fuel" on YouTube, more than 41,000 videos can be found, most of which are demonstrating electrolysers, which shows that there is a growing awareness of the potential of water as a fuel. There are many commercial electrolysers available.



Brown's Gas (HHO) has most unusual properties as demonstrated by Denny Klein's use of it in a welding torch as seen here: http://www.youtube.com/watch?v=6Rb_rDkwGnU and Denny has also run his car with water appearing to be the only fuel – the power comes from the zero-point energy field but that energy is transported by the water. Surprisingly, nearly everyone who is working with, or experimenting on the gas produced by electrolysis, believes that the resulting energy comes from the hydrogen in the gas mix, while the reality is that this is not actually the case. Brown's Gas has a cool flame of just 130 degrees Fahrenheit (water boils at 212 degrees Fahrenheit), and yet that same flame can vaporise tungsten which requires more than 10,000 degrees Fahrenheit and burning Hydrogen will never, ever, reach that temperature. Brown's Gas can also dramatically reduce radioactivity in radioactive materials and burning hydrogen can't do that. Also, when Brown's Gas is analysed in a high-tech laboratory, very little hydrogen is found and instead there are gaseous water clusters with excess electrons.

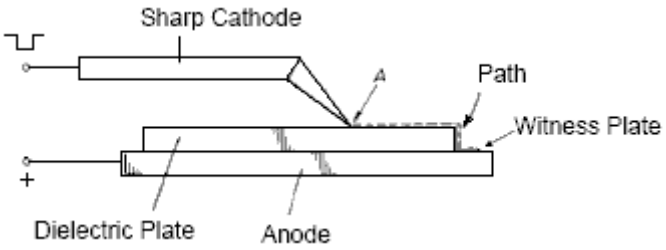
These charged water gas clusters have the same strange energetic effects that the plasma-charged have and that appears to be a microscopic form of ball lightning, studied extensively by Ken Shoulders who named them "Exotic Vacuum Objects" or "EVO" when he became convinced that their excess energy was being drawn in from the zero-point energy field of 'the vacuum'. These water gas clusters have a self-organised grouping of matter, plasma and zero-point energy. The typical grouping caused by this turbulent plasma is a vortex ring called a plasmoid (which has often been proposed as a model for ball lightning):





In a plasmoid, the electrons and ions spiral around the vortex ring and the force-free vortex creates a natural stability which sustains the plasmoid shape. These charge clusters can be produced quite easily as shown by Ken Shoulders in his US Patent 5,018,180 of 1991 where he shows that an abrupt electric discharge from a capacitor through a sharp pointed electrode on to a dielectric surface creates a charge cluster which travels on the surface of the dielectric to the anode. It appears to be a micron-sized form of ball lightning and it can punch a hole through the 'witness' plate, leaving a crater made by a high-energy event. The patent is well written and

describes many possible applications of his discovery.



The Exotic Vacuum Object formed by this process is believed to contain 100,000,000,000 electrons plus some 100,000 ions, giving it a charge-to-mass ratio similar to an electron and a really interesting fact is that it contains more energy than the energy which was stored in the capacitor which created it. These charge clusters adhere to dielectrics and they can remain for a long time. Many of them can clump together into a formation like a necklace. They can create holes through high melting-point ceramics such as aluminium oxide. Ken believes that the creation of these holes in ceramics is caused by the disruption of the electrons in the ceramic and so the hole 'melted' through the ceramic is not actually produced by heat. He has performed experiments which demonstrate the transmutation of one element into another and others which show radioactive materials being converted into benign elements.

Water gas clusters have the same characteristics as Ken's Exotic Vacuum Objects and they cause the most unexpected effects of using Brown's Gas where the cool flame (266 degrees F) does not boil water (which needs 212 degrees F to boil) http://www.watertorch.com/, and yet, that same flame can vaporise Tungsten which requires 10,031 degrees F, mind you, the Brown's Gas flame does not vaporise Tungsten by heating it above 10,000 degrees Fahrenheit but instead, does it by disrupting the bonding of the molecules in the metal. Here are some comparisons:

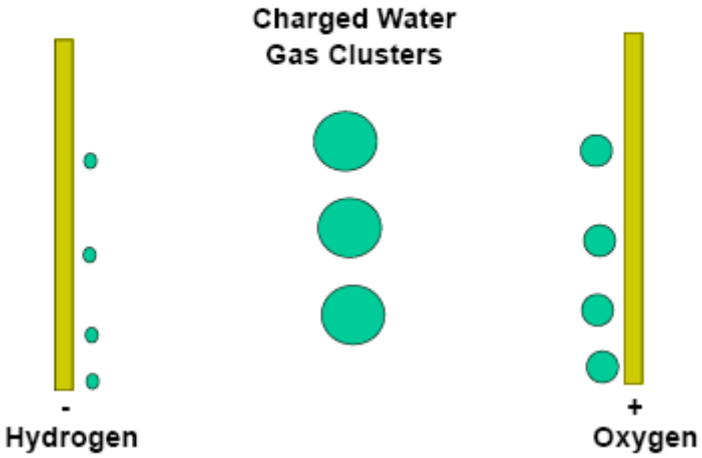
Tungsten		
Melt	6192 F	3422 C
Vaporise	10031 F	5555 C
Browns Gas	266 F	130 C
Torches		
Acetylene	5972 F	3300 C
Hydrogen arc	7232 F	4000 C
Cyanogen	8477 F	4525 C
Dicyanacetylene	9009 F	4987 C

This heating can't possibly come from burning hydrogen. Some tests produce interesting results. For example, fill a balloon with the gas produced by electrolysis and leave the balloon sealed for some time. The tiny hydrogen atoms and molecules can, and do, escape through the material of the balloon, causing it to fall to the ground. But, the contents remaining in the balloon still produces a burning flame when pushed out through a small tube and lit. A similar experiment is to fill a paper bag with the gas. Seal the bag and leave it for twelve hours to allow the hydrogen to escape. What remains in the bag is a gas which is heavier than air and which can be ignited.

George Wiseman (http://www.eagle-research.com/) is a leading Brown's Gas researcher who has found that browns Gas burns downwards in an imploding ring. In 2008, Chris Eckman measured the characteristics of Brown's Gas at Idaho State University. The measurements showed that there was very little hydrogen (monatomic or diatomic) present. Instead, the gas was found to be a form of water with excess electrons, effectively, a gas which was neither water vapour nor steam. When ignited, the flame temperature was found to be 266 F. or 130 C. (*Extraordinary Technology*, vol 2(6), pp 15-25, 2008).

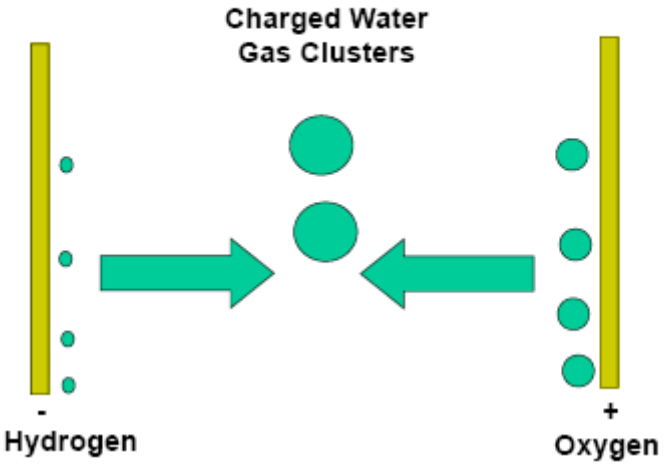
When using his acrylic electrolysers, George Wiseman has made an observation which he says is never mentioned in any textbook describing electrolysis:

Wiseman: Middle Bubbles

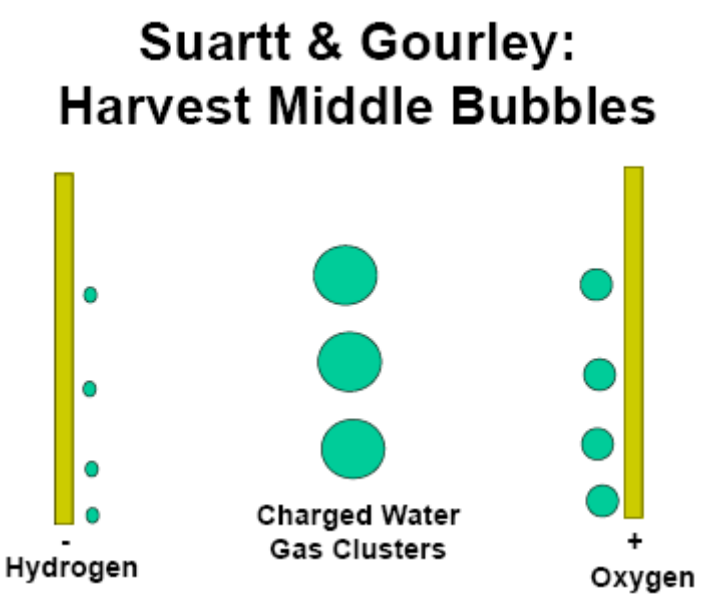


Between the electrolysis electrode plates which have a wide spacing of more than 10 mm, three sets of bubbles are produced. Hydrogen is produced on the negative electrode plate. Oxygen is produced on the Positive electrode plate. But, in the middle of the gap between those electrode plates a third set of bubbles are generated. Many researchers believe that these additional bubbles form the most energetic component of the gas – the charged water gas clusters. Bob Boyce has made a similar observation, noting that when electrolysis first starts, there are two jets which start from the plates and collide in between the plates where the middle bubbles are formed.

Bob Boyce: Two Colliding Jets



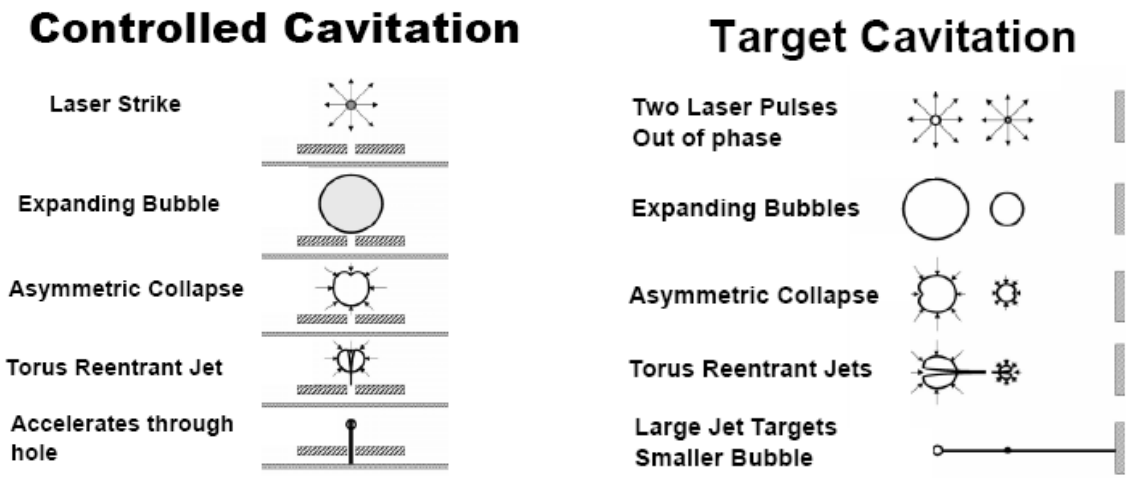
Ted Suartt and Rob Gourley (http://www.wateriontechnologies.com/) have not only made the same observation, but have developed a process and applied for a patent where they intentionally harvest just the middle set of bubbles:



Suartt and Gourley realised that the dominant electrolyser gas was not Hydrogen, and claiming that they are the first to discover this, named the gas after themselves as "SG Gas". Their extraction process involves widely separated electrode plates and a method of extracting the gas bubbles produced in the middle region between those two plates and excluding the hydrogen and the oxygen produced. They have investigated the properties of water infused with the gas and claim that it has health benefits. They state that Rhodes Gas and Brown's Gas are both "dirty cocktails" which include H_2 or O_2.

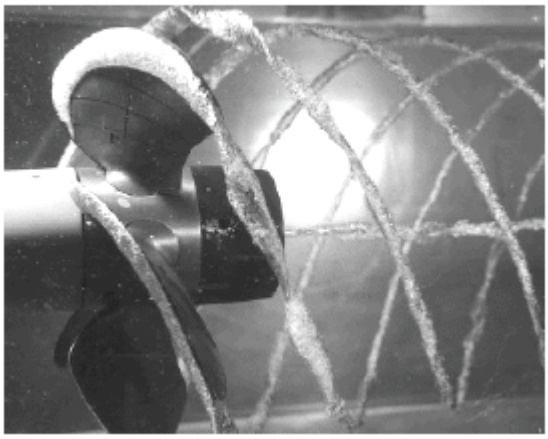
The anomalies of Brown's Gas are similar to those of plasma charge clusters (Ken Shoulders' EVOs). It adheres to matter and is electrically polarised. It gives an electric shock if it implodes to form water again. The isolated gas tends to implode instead of explode in piston experiments. However, if air is added to the mixture, the air is heated and that can cause overall expansion. In a welding torch it has a cool flame but it can vaporise tungsten. The flame cuts cleanly through solid, high melting-point materials including wood and ceramics, can weld dissimilar metals together and can even weld steel to clay brick. The claims of neutralising radioactive materials as well as the transmutation of elements are extraordinary. Todd Knudston comments on these properties at http://www.amasci.com/freenrg/hydroxy.html. At the 2011 Tesla Conference, Vernon Roth announced that he has observed element transmutation in his electrolysis cell. Details of this are given on Sterling Allan's web page at http://peswiki.com/index.php/OS:Vernon_Roth%27s_Alchemical_Hydrogen.

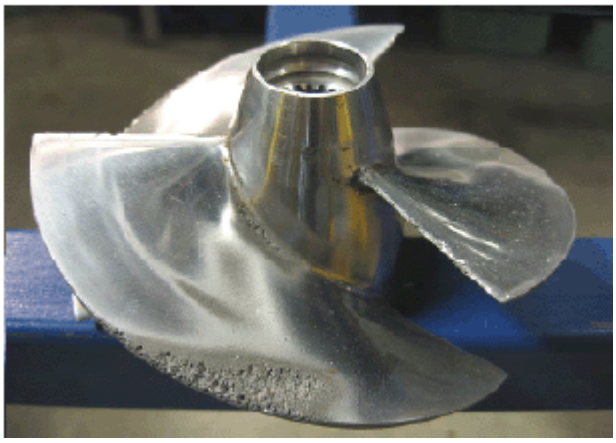
Mark LeClair can explain how water cavitation creates microscopic craters in metallic surfaces, carves trenches in high melting-point ceramics, transmutes elements, and produces excess energy. Cavitation bubbles have been studied and they have been shown to produce unexpected excess energy. Sonoluminescence occurs when water mixed with an inert gas such as argon or xenon, is excited by ultrasonic waves. A blue light is emitted as each bubble collapses suddenly and symmetrically. If that blue light were produced by a heating effect, then the bluish spectrum would indicate temperatures of over 10,000 degrees Kelvin, which caused many scientists to suggest that it might be used for hot fusion. Nobel laureate, Julian Schwinger, suggested instead that the light is caused by the zero-point energy. Here, the abrupt scalar compression of the bubble walls activates a Zero-Point Energy coherence, emitting the blue light at a far lower temperature (http://en.wikipedia.org/wiki/Sonoluminescence). Mark LeClair has four patents on controlled cavitation (typically for the precision cutting of materials): US 6,932,914, US 6,960,307, US 7,297,288 and US 7,517,430.



Cavitation bubbles form in the low-pressure region behind any rapidly moving surface in water. Ship's screws are notorious for making cavitation bubbles and being damaged by them:

Cavitation Erosion





However, the most useful application of cavitation is inside an electrolyser. Archie Blue raised the effectiveness of his electrolyser by blowing air upwards through the electrolyte. The technique can be applied to many different electrolyser designs. The electrode plates should have a rough clean surface with a tiny inter-plate gap of less than one millimetre. A gap that small allows the electrolysis gas to cause cavitation. The electrical stimulation can use pulsed DC waveforms but with minimal current and electrolyte (remember that we are **not** trying to make hydrogen). Circulating the water rapidly can charge it electrostatically and possibly even enough to allow the removal of the external DC pulsing. There are many ways to produce cavitation in water: make electrolysis gas in narrow gaps, blow air through the electrolyser, create a Venturi vacuum, vibrate the water by mechanical, acoustical or ultrasonic means, oscillate an electric field via a toroidal coil or via pulsed waveforms. Here, charged or polarised clusters or bubbles will oscillate with the field causing turbulence and cavitation.

When a cavitation bubble collapses near a hole or irregularity, it forms a torus and all of the energy of the collapsing bubble gets concentrated into a re-entrant jet. The extreme pressure in the jet creates a new solid-state of water, a water crystal with a plasma bow shock wave which draws in Zero-Point Energy. If they collide, water crystals can form small rings, trapping the energy in a meta-stable torus form. This is the seed of the charge water gas cluster. When ignited, the ring breaks to re-launch the LeClair effect water crystal or alternatively, becomes a plasmoid Exotic Vacuum Object, either of which have excess energy.

Rapidly circulating water through an electrolyser causes numeric energetic effects. It charges the water by electrostatic rubbing, it causes turbulence and cavitation as it flows through the tight rough gaps, it can vibrate the plates producing reed cavitation and best of all, cycling the water repeatedly through the electrolyser integrates it's energy content, producing an ever increasing energy level. With water which is sufficiently charged, spraying it as a mist into the carburettor of an engine can give the illusion that water is a fuel.

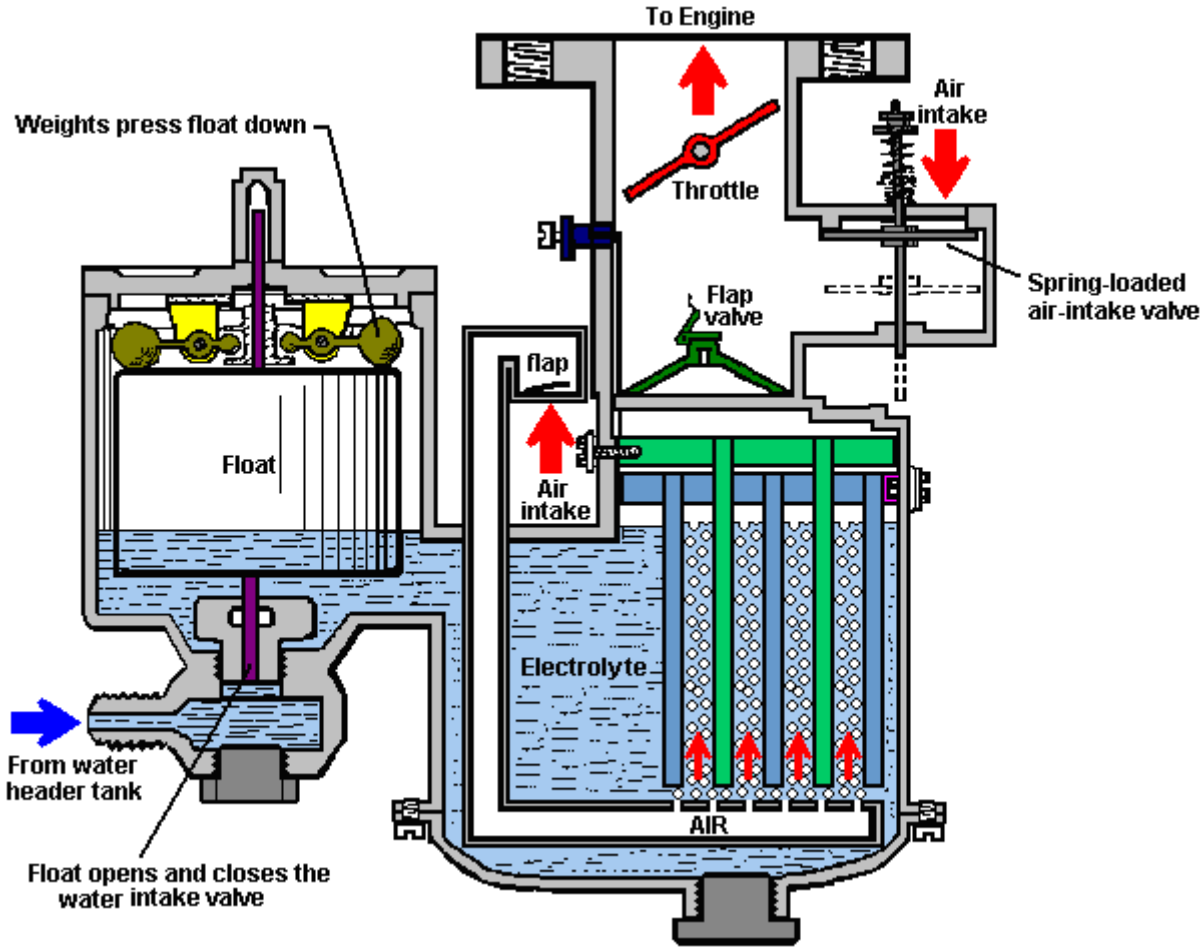
This is just a brief summary of part of the contents of Moray King's presentation pdf document which you can read in full here: http://www.free-energy-info.com/MorayKing.pdf. An interesting, seemingly directly related video is: http://www.youtube.com/watch?v=i-T7tCMUDXU&feature=youtu.be. Considering what Moray King has described, we need to consider more carefully the patented electrolyser designs of Charles Garrett and Archie Blue:

The Charles Garrett Electrolyser.

Charles Garrett was granted US Patent 2,006,676 on 2nd July 1935 in which he shows some impressive details. Firstly, he generated an extra electrical input by fitting a second (6 volt) alternator to his car. While the drawing shows the applied voltage swapping over in polarity, this was not done rapidly, just occasionally to even up any deterioration of the electrodes.

He maintained the water level in the electrolysis chamber with a neat carburettor-style float and pin valve arrangement. He improved the electrolysis by introducing a perforated tube below the electrode plates which allows the engine to suck air up past the plates. This cools the electrolyte (water with a few drops of hydrochloric acid) introduces water vapour to the gas mix and dislodges any bubbles on the plates, without the need for any extra mechanical device. Considering that he did this seventy-five years ago, it is an impressive piece of work. Please note that while only five electrode plates are shown in the diagram, in reality it is probable that many such

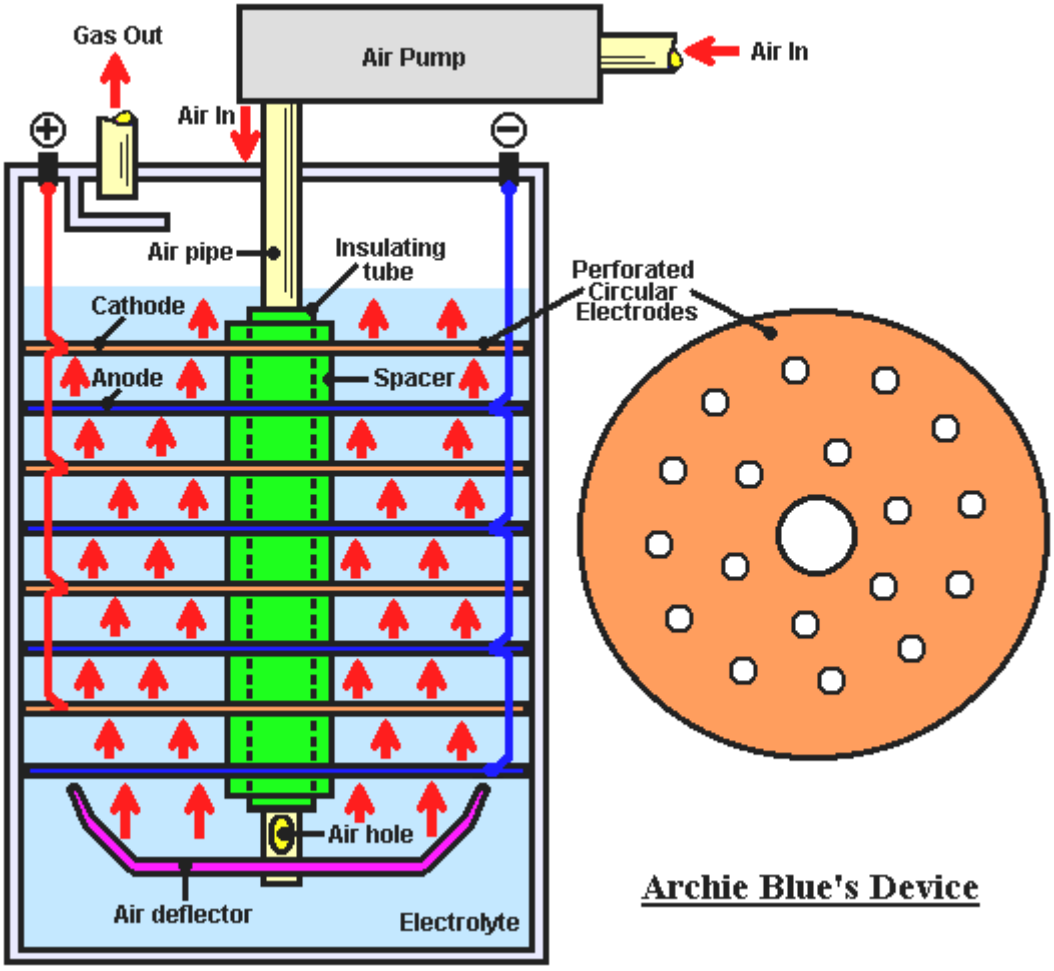
plates were used since the gas volume is directly proportional to the plate area.



One point which should be noted is that the cars of that time had very much smaller capacity engines and so they will have needed far less HHO gas mixture in order to run adequately.

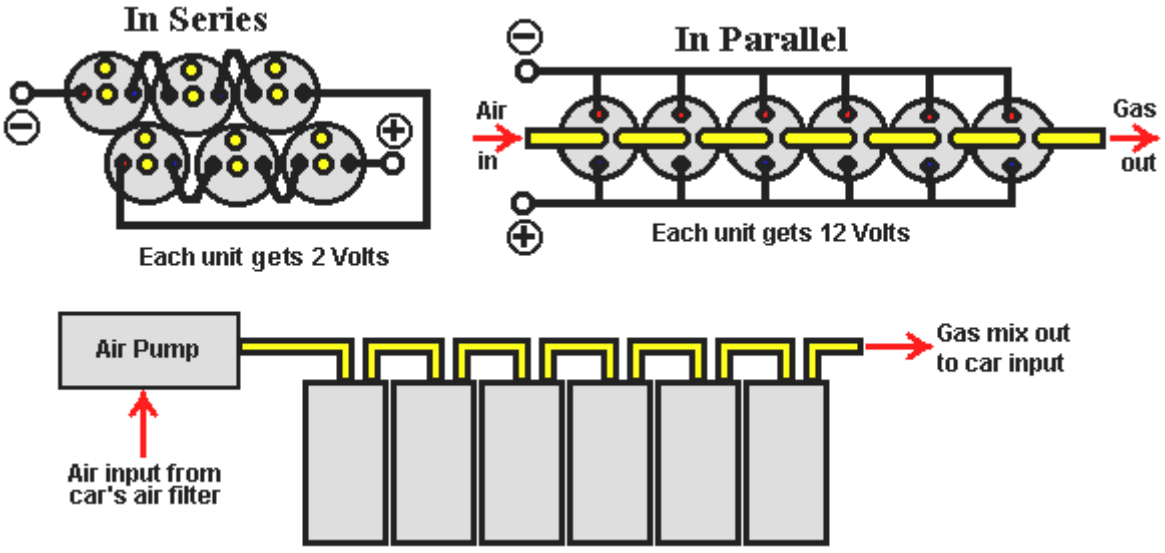
The Archie Blue Electrolyser.

More than fifty years after Charles Garrett was granted his patent, another one was granted to Archie Blue. The equipment described in the two patents operates in more or less the same way. Archie's equipment is very simple to construct and uses straight electrolysis with no attempt at pulsing the electrical supply. Like Charles Garrett, Archie Blue claimed to have run a car on water alone, using his electrolyser design, which is shown here:



With this unit, air is sucked out of the exit pipe by the vehicle engine, while being pumped into the electrolyser by an air pump. The air flows down through the central pipe and is forced up through the non-aligned holes in the electrode plates, causing turbulence and probably, the formation of water-gas crystals. The air bubbles also stir the electrolyte into vigorous motion, dislodging the hydrogen and oxygen bubbles which form on the plates as a result of the electrolysis current flow through the electrolyte.

It is said that six of these electrolysis units are sufficient to run a car using just water as the fuel. It has been stated that electrolysis of water is optimum at 1.5 Volts, so it might be more efficient to connect the units in series where each units receives 2 Volts rather than in parallel where each unit receives 12 Volts (unless, of course, the heating caused by connecting them in parallel is a factor in the very high efficiency of Archie Blue's system):



The air connection is the same for either method of wiring the cells. If wired in series, the voltage drop across each cell may not be the same although they were constructed in an identical fashion.

Please bear in mind that should you modify a vehicle to run on hydrogen, either as an additive or as a replacement for petrol, you need to clear it with your insurance company before using it on a public road, otherwise, you will be driving without insurance since any alteration to the vehicle automatically invalidates the insurance if the insurer is not notified and agrees the change. You may, of course, modify any stationary engine or any vehicle which you only run on private property. In the USA, the oil companies have influenced the local courts to such a degree that in some States, it is an offence to "run a vehicle on a non-approved fuel".

In passing, you may be interested to hear that I have been told that the Prohibition era in America had nothing at all to do with people drinking alcohol. The reality was that in the early days, Henry Ford was going to have his Model-T car running fuel-less by using a Nikola Tesla designed magneto system and an electric engine, but he was pressured into using an internal combustion engine to burn the gasoline which was an unwanted component of the local oil industry. This caused a problem for people on long journeys as there were very few gasoline filling stations at that time. To overcome the problem, the early cars were set up so that they could run on either gasoline or on alcohol produced by some 50,000 farmers scattered around the country. When the oil industry discovered how profitable it was selling gasoline, they opened many gasoline filling stations. They then wanted to exclude the farmers and have all of the profits for themselves and so Prohibition was introduced, not to stop people drinking alcohol (although that was the pretext), but in reality, to shut down the 50,000 alcohol stills which were their competition. When the stills were gone, then Prohibition was dropped as it had achieved it's goal of a vehicle fuel monopoly.

The Electrolyser System of Paul Zigouras

In his document, Moray King draws attention to the HHO cell design of the American, Paul Zigouras who became very well known in the year 2011 due to his cell design which is perfectly capable of running a 320 horsepower marine engine. Paul, at age thirty, had a history of rebuilding marine engines and then selling them. He then was involved in helping a friend who had entered a competition for the car with the highest mpg performance. Paul found that by adding HHO to the air entering the car engine, that they could get 70 mpg on the flat when driving carefully. Their target was 100 mpg, and so they were disappointed by being 30 mpg short of their target.

Paul decided to work on the HHO angle to see what could be achieved and enlisted the help of a friend and sometime freelance employee who was expert in electronics as Paul was not familiar with electronics. Probably without understanding the underlying causes, they managed to utilise several operating principles – cavitation, charged water gas clusters, resonance, mechanical vibration and a 'splitting the positive' style power supply. The overall result was spectacular, culminating in a small cell, into which water could literally be hosed and only gas came out of the other end.

Paul never revealed the exact circuit design and it is reported that he sold the rights to the design for US $6,000,000. The purchasers then contacted the eBay buyer of the latest version of Paul's circuit board and paid $20,000 to buy it back. The eBay buyer was happy with the deal as he had paid just $1,100 for it and so made a profit of $18,900 on the deal. At this time of writing (2013) some eighteen months have gone by and it seems quite clear that the present owners of Paul's circuit design have no intention whatsoever, of sharing or manufacturing the design and so it has been effectively shelved, never to be seen again. They were not interested in the cell, but just the electronics board.

However, a number of things are known about the design, probably sufficient to allow a similar design to be produced. These items are as follows:

Paul made units in two sizes. The smaller version had between 20 and 30 plates each being 2-inches (50 mm) wide and 8-inches (200 mm) long, material 316L-grade or 318L-grade stainless steel one sixteenth of an inch thick (1.6 mm), stacked with a gap of just 0.635 mm between them. This small version could blast 2.5 (US) gallons of water per minute into gas, which is about 17,500 litres of gas per minute. The water was hosed in at one end, and no water reached the far end, which is quite spectacular performance.

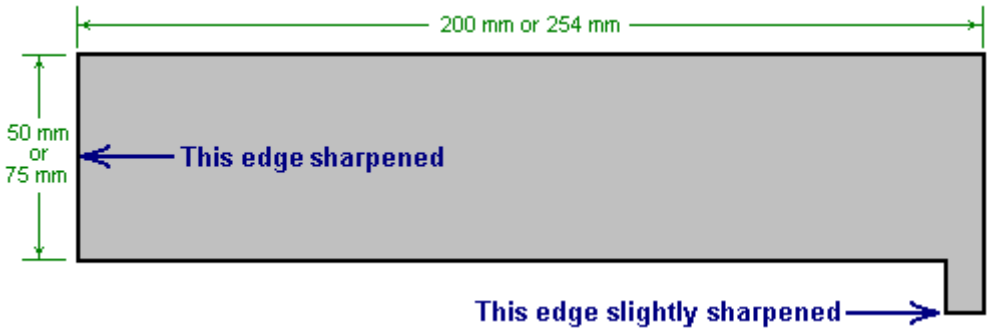
The larger cell had 36 plates 3-inches (75 mm) wide and 10-inches (254 mm) long, also 316L or 318L-grade stainless steel 1.6 mm thick and with a gap of 0.635 mm between the plates. That size of cell could convert water to gas at the rate of 5 (US) gallons per minute (35,000 litres of gas per minute)

The techniques used with these cells is nothing at all like any of the various other electrolyser designs discussed in this chapter. This is because the cell operation is nothing like conventional electrolysis or even like DC pulse driven water splitting as used in the Stan Meyer "Water Fuel Cell"

Firstly the plates are grit-blasted with 60-grade silicon carbide at an angle of 45 degrees to the face of the plate, making sharp-edged craters in the surface of the plate. When water is forced through the very narrow gap

between these plates, these craters on both sides of that very narrow water flow causes turbulence and cavitation. Cavitation produces tiny bubbles in the water and in the book "Ultrasonics Methods and Applications" by Jack Blitz, it is stated that each cavitation bubble has a positive charge on one side and a corresponding negative charge on the opposite side. As the bubbles are very small, those charges are not far apart and it would not seem to be unrealistic to suggest that those charges cause electrolysis of the water on a very tiny scale. But, as there is a very large number of these bubbles, the overall effect might not be insignificant. The violent cavitation almost certainly produces charged water gas clusters, so what comes out of the end of the cell will be HHO gas, charged water gas clusters, plus whatever gases which were dissolved in the water and possibly water vapour.

The plates were shaped like this:



The edge facing the incoming water is sharpened to a knife edge, and the projection to take a push-on electrical spade connector has it's outer edge slightly sharpened to make it easier to push the connector on to the plate and make a good electrical connection to the plate. Due to the very small gap between the plates, every second plate is turned over to give some clearance between the connectors. This places all of the positive connections on one side and all of the negative connections on the opposite side.

When being used to run a car, the HHO cell is placed inside the standard, rectangular, plastic air box which connects the air filter to the manifold air intake. This causes the incoming air to mix well with the HHO and other gases produced by the cell, before the mixture enters the engine.

The electronic circuit (valued at $6,000,000) has an ordinary automotive power supply of around 14 volts. This is supplied by a standard MSD 200-amp high performance alternator driven by the engine. The initial current draw for the larger (35,000 lpm) cell is 190 amps, but when the cell gets going, that current draw drops to a constant 10 amps, and the gas production rate is **not** related to the current draw. This process is **not** any kind of conventional electrolysis and has nothing at all to do with Faraday's excellent lab work. The Coefficient of Performance is said to be between 5 and 10 although how that figure could be derived is by no means at all clear.

The circuit is said to produce a very clean square wave with very sharply rising and falling edges to the waveform. The frequency of the wave is in the 40 kHz to 44 kHz region and 30 separate transistors are used to drive the plates – presumably, one transistor per plate for the 30-plate version of the smaller cell. The waveform does not drop to zero volts, but instead, has a voltage offset of +1 volt. That is, the voltage oscillates between +1 volt and +14 volts and so there is always a voltage applied to the plates. As the circuit has never been disclosed, it is quite possible that the voltage is boosted well above the +14 volt level, however, that seems unlikely if 190 amps is the starting current. It is stated that at 44 kHz, the current needed is only one eighth of what would be expected for the HHO flow rate.

In a vehicle, the gas flow rate is controlled by improvising a linkage between the throttle and the valve which controls the rate of water flow into the cell. Not surprisingly, if the water inflow rate is cut down, the gas production rate has to fall also as there just isn't any water remaining to be converted into gas. Each litre of water produces around 1,860 litres of HHO gas, and so if the cell output is 17,500 lpm, then the water inflow rate would be about 9.4 litres per minute or 157 ccs per second. However, it seems unlikely that when mixed with air, as much as 17,500 litres of HHO would be needed per minute. It is remarked that using hot water just under boiling point, is an advantage, although why that should be is not specified.

The cell buzzes quite loudly when operating. This is definitely not caused by a 40 kHz frequency signal as human hearing only goes up to 20 kHz at most. It might be a lower harmonic (20 kHz, 10 kHz, 5 kHz,…) or it might be caused by mechanical forces generated by the water flow. If tap water is being used, then dissolved solids will be left behind when the water becomes gas. This residue can be washed out of the cell by turning the water on and keeping the electrical power off as that washes the plates.

With this amount of information on the cell and electronics design, it could well be possible to replicate the cell and run vehicle engines from it. However, please be aware that Bob Boyce in America was handed a 3.5 year jail term for “running a vehicle on an unapproved fuel”. He beat that charge, but it should be clearly understood that an action like that is wholly unlawful and is part of the scam which attempts to force all vehicle users to burn oil.

Also in America, Bill Williams was running his Ford pick-up truck with a Joe Cell connected as a booster in “shandy” mode. He found that his truck used no fuel at all even though it was perfectly capable of drawing fuel in from the fuel tank. Bill destroyed his cell and does not talk about it due to intimidation from armed thugs. Details of his design along with more advanced Joe Cell designs are in chapter 9.

The Petro Dollar

However, things did not stop there by any means. To discover more, perhaps you should consider the following video at http://www.safeshare.tv/w/gQnBDHTCDs which states:

Why did the United States attack Libya, Iraq, Afghanistan and Yemen? Why are US operatives helping to de-stabilise Syria? And why is the United States government so intent in taking down Iran, in spite of the fact that Iran has not attacked any country since 1798?

And, what’s next? What are we headed for? When you look at the current trajectory that we are on, it doesn’t make any sense at all if you evaluate it on what we are taught in school. And it doesn’t make any sense if you base your world view on the propaganda that the mainstream media tries to pass off as news. But it makes perfect sense once you know the real motives of the powers that be. In order to understand those motives, we first have to take a look at history:

In 1945, Britain, with agreement, established the dollar as the world’s Reserve Currency, which meant that international commodities were priced in dollars. The agreement, which gave the United States a distinct financial advantage was made under the condition that those dollars would remain redeemable for gold at a consistent rate of $35 per ounce.

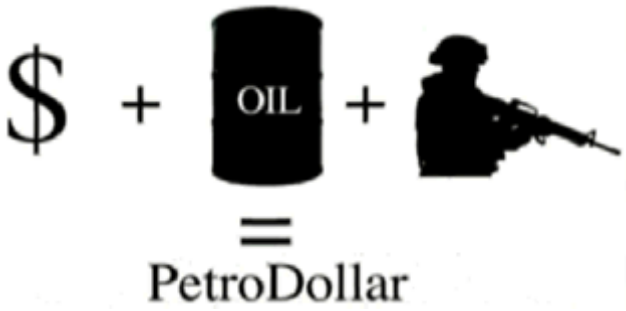
The United States *promised* not to print very much money but this was on the honour system because the Federal reserve refused to allow any audit or supervision of it’s printing presses.

In the years leading up to 1970, expenditures in the Vietnam War made it clear to many countries that the US was printing far more money than it had in gold, and in response, they began to ask for their gold back. This, of course, set off a rapid decline in the value of the dollar. The situation climaxed in 1971, when France attempted to withdraw it's gold and Nixon refused. On 15th August, he made the following announcement:

"I have directed the Secretary of the Treasury to take the action necessary to defend the dollar against the speculators. I directed Secretary Connolly to suspend temporarily, the convertibility of the dollar into gold or other reserve assets except in amounts and in conditions determined to be in the interests of monetary stability and in the best interests of the Unites States".

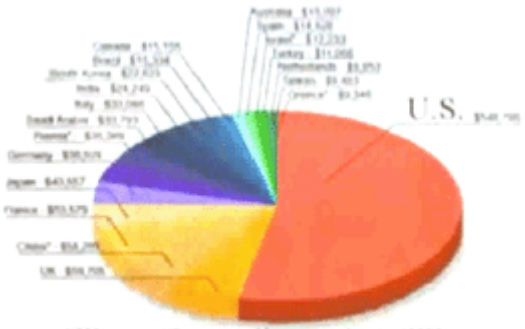
This was obviously not a temporary suspension as he claimed, but rather a permanent default, and for the rest of the world who had entrusted the United States with their gold, it was outright theft. In 1973, President Nixon asked King Faisal of Saudi Arabia to accept only US dollars in payment for oil, and to invest any excess profits in

US Treasury Bonds, Notes and Bills. In return, Nixon offered military protection for Saudi oil fields. The same offer was extended to each of the key oil-producing countries, and by 1975, every member of OPEC had agreed to only sell their oil in US dollars.



The act of moving the dollar off gold and tying it to foreign oil, instantly forced every oil-importing country in the world to maintain a constant supply of Federal Reserve paper, and in order to get that paper, they would have to send real physical goods to America. This, was the birth of the Petro Dollar. Paper went out, everything America needed came in, and the United States got very, very rich as a result. It was the largest financial con in recorded history.

The Arms Race of the Cold War was a game of poker. Military Expenditures were the chips, and the US had an endless supply of chips. With the Petro Dollar under it's belt, it was able to raise the stakes higher and higher, outspending every other country on the planet, until eventually, US military expenditure surpassed that of all of the other nations in the world combined – the Soviet Union never had a chance.



Military Expenditures in Millions

The collapse of the communist bloc in 1991, removed the last counterbalance to American military might. The United States was now an undisputed Super-power with no rival. Many hoped that this would mark the start of a new era of peace and stability. Unfortunately, there were those in high places who had other ideas. Within that same year, the US invaded Iraq in the first Gulf War, and after crushing the Iraqi military, and destroying their infrastructure, including water-purification plants and hospitals, crippling sanctions were imposed which prevented that infrastructure from being rebuilt.

These sanctions which were initiated by Bush Senior, and sustained throughout the entire Clinton administration, lasted for over a decade and were estimated to have killed over five hundred thousand children. The Clinton administration was fully aware of these figures.

A TV interviewer speaking to Madeleine Albright, Secretary of State for Clinton, asked: “We have heard that half a million children have died, I mean, that’s more children that died than in Hiroshima, and, and is the price worth it?”. To which Madeleine Albright replied: “I think that this is a very hard choice. We think that the price is worth it”.

Miss Albright, what exactly was worth killing 500,000 kids for? In November of 2000, Iraq began selling it’s oil exclusively in Euros. This was a direct attack on the dollar and on US financial dominance, and it wasn’t going to be tolerated. In response, the US government with the assistance of the mainstream media, began to build up a mass propaganda campaign claiming that Iraq had weapons of mass destruction and was planning to use them. In 2003, the US invaded and once they had control of the country, oil sales were immediately switched back to dollars. This is particularly noticeable as switching back to the dollar meant a 15% to 20% loss in revenue due to the Euro’s higher value. It doesn’t make any sense at all unless you take the Petro Dollar into account.

On 2nd March 2007, US General Wesley Clark said: “So I came back to see him a few weeks later and by that time we were bombing in Afghanistan. I said: ‘Are we still going to war with Iraq?’ and he said ‘Oh it’s worse than that’. He said as he reached over on his desk and picked up a piece of paper and he said ‘I just got this down from upstairs (meaning from the Secretary of Defence’s Office) today, this is a memo which describes how we are going to take out seven countries in five years, starting off with Iraq and Syria, Lebanon, Libya, Somalia, Sudan and finishing off Iran”.

Let’s take a look at the events of the past decade and see if you see a pattern. In Libya, Gadaffi was in a process of organising a block of African countries to create a gold-based currency called the “Dinar” which they intended to use to replace the dollar in that region. US and NATO forces helped to destabilise and topple the government in 2011 and after taking control of the region, US armed rebels executed Gadaffi in cold blood and immediately set up the Libyan Central Bank. Iran has been actively campaigning to pull oil sales off the dollar for some time now, and it has recently secured agreements to trade it’s oil in exchange for gold. In response, the US government with mainstream media assistance has been attempting to build international support for military strikes on the pretext of preventing Iran from building a nuclear weapon. In the mean time they established sanctions which they openly admit are aimed at causing a collapse of the Iranian economy.

Syria is Iran’s closest ally and they are bound by mutual defence agreements. The country is currently in the process of being destabilised with covert assistance from NATO and although Russia and China have warned the United States not to get involved, the White House has made statements in the past month indicating that they are considering military intervention. It should be clear that military intervention in Syria and Iran isn’t being considered – it is a foregone conclusion. Just as it was in Iraq and Libya, the US is actively working to create the context which gives them the diplomatic cover to do what they have already planned. The motive for these invasions and covert actions becomes clear when we look at them in their full context and ‘connect the dots’. Those who control the United States understand that if even a few countries begin to sell their oil in another currency, it will set off a chain reaction and the dollar will collapse. They understand that there is absolutely nothing else holding up the value of the dollar at this point and so does the rest of the world. But instead of accepting the fact that the dollar is nearing the end of it’s lifespan, the powers that be have made a calculated gambit. They have decided to use the brute force of the US military to crush each and every resistant State in the Middle East and Africa.

That in itself would be bad enough, but what you need to understand is that this is not going to end with Iran. China and Russia stated publicly and in no uncertain terms that they will not tolerate an attack on Iran or Syria. Iran is one of their key allies, one of the last independent oil-producers in the region, and they understand that if Iran falls, then they will have no way to escape the dollar without going to war. And yet, the United States is pushing forward in spite of the warnings. What we are witnessing here is a trajectory which leads straight to the unthinkable. It is a trajectory which was mapped out years ago in full awareness of the human consequences. But who was it that put us on this course? What kind of psychopath is willing to intentionally set off a global conflict which will lead to millions of deaths, just to protect the value of a paper currency? It obviously isn’t the President. The decision to invade Syria, Libya and Iran was made long before Obama had risen to the national

spotlight, and yet, he is carrying out his duties just like the puppets who preceded him. So who is it that pulls the strings?

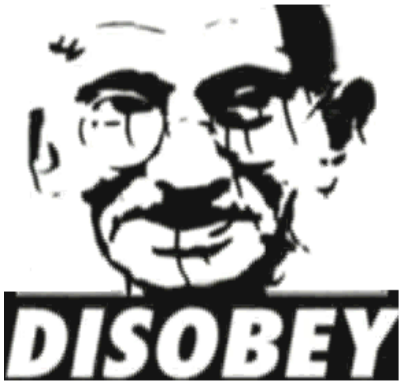
Often, the best answers to questions like this are found by asking another question “Cui Bono?” - “Who Benefits?” Obviously, those who have the power to print the dollar out of thin air have the most to lose if the dollar were to fall, and since 1913, that power has been held by the Federal Reserve. The Federal reserve is a private entity owned by a conglomerate of the most powerful banks in the world and the men who control those banks are the ones who pull those strings. To them, this is just a game. Your life, and the lives of those you love are just pawns on their chessboard. And like a spoiled four-year-old who tips the board on to the floor when he starts to lose, the powers that be are willing to start World War Three to keep control of the global financial system.

Remember that when these wars extend and accelerate. Remember that when your son, or your neighbour’s son comes back in a flag-draped coffin. Remember that when they point the finger at the new ‘bogeymen’ because the madmen who are running this show, will take this as far as you allow them to.

So, how much time do we have left? It’s a question which I hear constantly. But it is the wrong question. Asking how much time we have left is a passive posture. It is the attitude of a prisoner who is waiting to be taken out to a ditch and shot in the back of the head.

What are our chances? Can we change course? Also, the wrong question. The odds don’t matter any more. If you understand what we are facing, then you have a moral responsibility to do everything in your power to alter the course we are on, regardless of the odds. It is only when you stop basing your involvement on the chances of your success, that success actually becomes possible. To strip the ill-begotten power from the financial elites and bring these criminal cartels to justice, will require nothing less than a revolution. The government is not going to save us. The government is completely infiltrated and corrupt to the core. Looking to them for a solution at this point is utterly naive.

There are three stages of revolution and they are sequential. Stage One is already underway Stage One is the ideological resistance. In this stage we have to actively work to wake up as many people as possible about what is happening and the direction we are headed. All revolutions originate from a shift in the mind-set of the population and no other meaningful resistance is possible without it. Success in this stage of the game can be measured by the contagion of ideas. When idea reaches critical mass, it begins to spread on its own and seeps into all levels of society. In order to achieve that contagion, we need more people in this fight. We need more people speaking up, making videos, writing articles, getting this information on to the national and international stage, and we especially need to reach the police and the military.



Stage Two is civil disobedience, also known as Non-violent Resistance. In this Stage, you put your money where your mouth is, or more accurately, you withhold your money and your obedience from the government and you do

everything in your power to bring the gears of the State to a halt. Practiced in mass, this method alone is often enough to bring a regime to it's knees. However, if it fails at this stage, Stage Three is inevitable.

Stage Three is direct physical resistance. Direct physical resistance is the last resort and it should be avoided and delayed as long as possible and only invoked when all other options have been thoroughly exhausted. There are those who 'talk tough' and claim that they will resist when the time comes, but what they fail to realise is that if you are inactive during the first two Stages and save your efforts for the last resistance, then you will fail.

When the Nazis were moving from door to door, dragging people out of their homes in Germany, that was the time to fight back physically, but due to the lack of ideological resistance and civil disobedience leading up to that moment, even an armed uprising would have likely failed at that point. An armed uprising can only succeed if the people have established an attitude of active resistance. And active resistance is only possible after their minds have broken free from the mainstream propaganda. If you want to fight back, then it's now or never – you're not going to get another chance, and the stakes are far higher than they were in Nazi Germany.

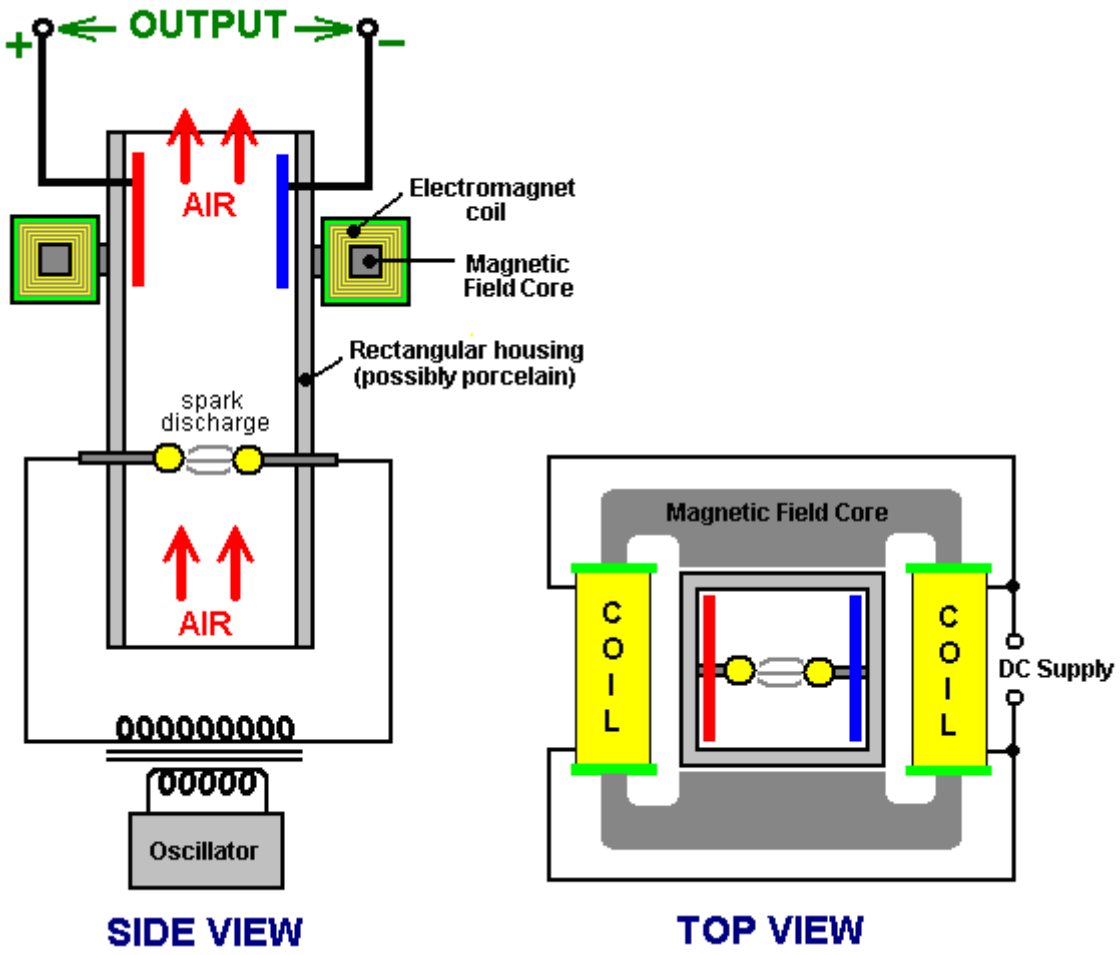
If you want to know more about the present situation, then watch the very informative web video at http://www.youtube.com/user/ThriveMovement, which also shows what you can do to improve things.

Patrick Kelly
http://www.free-energy-info.com
http://www.free-energy-info.tuks.nl
http://www.free-energy-info.co.uk

Chapter 11: Other Devices and Theories

Nikola Tesla's Ionised-Air Electrical Generator.

Tesla also designed a device for picking up energy from the air. As far as I am aware, it was never patented and I have never seen a specification of its output. Perhaps it was one of Tesla's failures but personally, I doubt that. It might make a very interesting experiment so see what level of output can be achieved using it. The construction is shown here:



It is essentially, a rectangular cylinder which contains two spherical electrodes like a Wimshurst machine. The cylinder is positioned vertically, so that when the electrodes are powered up with high voltage to create spark discharges, the air inside the cylinder is heated which causes it to rise up the cylinder. The heated air is ionised, so a magnetic field generated by a surrounding electromagnet, causes the charged ions to move to opposite sides of the cylinder. Electrode plates positioned inside the cylinder, provide an electrical path for the excess positive and negative charges to flow together through the load - lighting, heating or motor circuits typically.

On the surface, this system would appear to be less than 100% efficient, in that the amount of power applied to the device to make it operate should be less that the amount of power drawn from it to drive useful loads. I am not sure that this is necessarily so. Firstly, the air already contains charged ions before this device starts to generate more. These naturally occurring ions gain in number when a thunderstorm is likely, even to the extent of giving many people a headache by their presence. These naturally occurring ions will be picked up by this device and without any input power needed to create them, they are capable of providing output power.

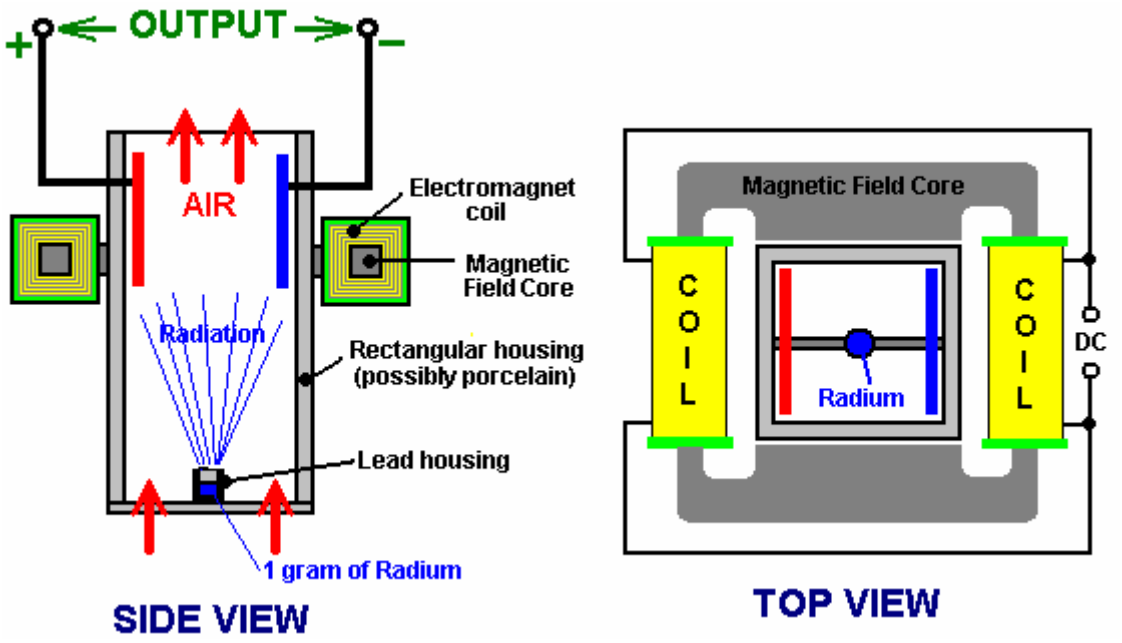
Also, the whole earth is immersed in the zero-point energy field. This is seething energy at the quantum level whose effects can be seen even at 'absolute zero'. This field is made of small random effects which

makes it hard to obtain useful energy directly from it. The field needs to be structured before energy can be drawn from it. One way to do this is to align the field with an event which causes coherent waves of energy to radiate outwards as a 'radiant energy' wave - something like the ripples caused on the surface of a pond of still water when a large stone is dropped vertically into the water. The ripple 'waves' move outwards from the 'event' until they reach the bank of the pond. If there was a generator attached to a float in the pond, it would be possible to pick up some energy from the ripples. The same can be done with 'radiant energy' waves if you can create them and know how to pick up energy from them.

Radiant energy waves can be formed by very short sharp uni-directional electrical pulses. Pulses less than one hundredth of a second are suitable for this. One way of creating pulses of that type is using a spark gap. In Tesla's device shown above, sparks are generated continuously. These sparks will generate radiant energy waves radiating out at right angles to the spark. Without a doubt, the vertical cylinder will have a mass of radiant energy shooting up it when it is being operated. This is in addition to the air ions which are being picked up. The only question is whether or not the electrode plate arrangement shown is capable of picking up any of this excess energy. Considering the metallic pickup device used by Edwin Gray to capture radiant energy as described below, it seems highly likely that some of that additional energy is, in fact, picked up and used to power the loads.

It should be noted that Tesla's device shown above, will generate UV radiation in the same way as any MIG or stick welder does, so care should be exercised to avoid looking at the arc or allowing the UV to shine on your skin, even if the skin is covered by clothing. You can get serious sunburn through thin clothing if it is subjected to strong UV radiation. Also, radio interference is likely to be generated by the arc, so screening should be provided during any tests. **WARNING**: Tesla accidentally discovered that electric spark discharges in air, ignite and burn atmospheric oxygen and nitrogen, producing 12,000,000 volt waves. The oxygen and nitrogen, both below atomic number 19 are thereby transmuted into alpha and beta charges (stripped helium nuclei with +2 charge each, and electrons with -1 charges each) by the powerful radiation produced, having a voltage potential of 12 Mev. This is almost three times the Mev level of gamma radiation emitted by radium, it may well be the reason why Tesla did not publicise the device shown above, and should you decide to experiment with it, please be aware of the potential hazard of this radiation.

A variation on the above device of Tesla's is given in the book "Physical Chemistry" by E. A. Moelwyn-Hughes, Pergamon Press, Oxford 1965, page 224. Rutherford and Geiger determined the fact that radium puts out alpha particles at the rate of 34,000,000,000 per second, each having two units of positive charge at 4.5 million electron-volts. This is a staggering amount of energy which ionises the air inside the housing and produces enough power to be capable of replacing the entire Four Corners power complex indefinitely.



The variation of Tesla's device shown above, supports the lead container with its gram of radium on a strap across the bottom of the housing. The radiation ionises the air and the magnetic field separates the charges and directs them to opposite sides of the housing, to be collected and used via the electrode plates. There

does not appear to be any reason why strong permanent magnets should not be used instead of the DC electromagnet shown.

Dr Harold Aspden's Electrical Power Generator Patent.

Scientists freely acknowledge that more than 80% of the matter and energy in the universe is "dark matter" and "dark energy" where "dark" only means that we cannot readily see that form of matter and energy. The highly respected British scientist Dr Harold Aspden, has been awarded a patent for a system to collect this energy directly. The patent, which is one of several similar patents included in this eBook, is reproduced here:

Patent GB2390941 ***21st January 2004*** ***Inventor: Dr. Harold Aspden***

ELECTRICAL POWER GENERATING APPARATUS

Abstract

An electric generating device includes two capacitors **1** and **2**, each having a pair of concentric electrodes and in-series connection to inductors **3** and **4**. Each capacitor has an electrode connected to a high voltage DC source **5** and another connected to a low-voltage or earth terminal **6**. An AC Power output may be produced from terminals between each capacitor and inductor or from a transformer where the inductor is the primary winding. Electricity production may be sustained by drawing energy from the vacuum medium surrounding the electrodes.

Field of the Invention

This invention relates to a new and non-conventional means for the generation of electrical power. The energy source is the quantum underworld of space, the aether medium of the vacuum state, long recognised for its ability to allow the storage of electric field energy by reacting as its intrinsic charge is displaced, a process understood by physicists by reference to the research findings of Clerk Maxwell.

Background of the Invention

The current state of the art of electrical power generation does not recognise the possibility of ultimately tapping energy from the aether. Physics is taught on the basis that energy cannot be created or destroyed, inasmuch as it is conserved in all physical processes, though it can be degraded in its usefulness, as by burning of hydrocarbons and conversion into heat which dissipates as by radiation into outer space. The aether as a source or as an absorber of energy is not deemed to serve any specific role in the physics of energy deployment, it having been dismissed from consideration by invoking the notion of 'field energy' without admitting the specific physical reality of something in space that accounts for the properties involved.

Theoretical physicists have, however come to suspect that space devoid of matter is nevertheless a seething sea of activity subject to sporadic energy fluctuations which can create electron-positron pairs that exist momentarily before decaying back into their quantum underworld. Yet those same physicists deny all possibility that this energy resource of space itself can be exploited to provide useful power on a scale large enough to rival the role played by atomic power plants and fossil fuel generating installations.

Curiously, they do subscribe to the belief that one day they may be able to generate power on a viable commercial scale from fusion reactors by processes replicating what they believe sustains the Sun's heat output as hydrogen is transmuted into different atomic forms. In contrast with this rather elusive objective, it having proved beyond reach even after half a century of effort, this invention is based on success in generating power by replicating, not the Sun's onward energy decay, but rather a process akin to that by which the Sun itself was created from energy drawn from the enveloping aether medium.

The invention to be described below has emerged from an in depth theoretical investigation into the properties of the aether and quite independently of any of the well known claims of published record which feature at the fringe of mainstream scientific literature. A recent and very well-presented account of what amounts to a century of relevant energy history is the book 'The Search for Free Energy' by Keith Tutt, published in 2001 by Simon Schuster (ISBN 0-684-86660-9). Here in this book is a comprehensive background of information concerning the energy devices of several researchers but the references to Nikola Tesla and T. Henry Moray are particularly pertinent to the subject of this invention and, though imposing a

limitation on what can be legitimately claimed by this patent application, they serve also as a basis for a very important lesson to those engaging in this field of invention.

The lesson is that it is not sufficient to build and demonstrate something that works, if you do not fully understand why what you have devised actually does work. This is especially the case here where one is claiming a source of energy hitherto unknown. The invention to be described below will, in its broadest sense, appear to be quite similar to what T. Henry Moray is said to have demonstrated in showing that substantial electrical power could seemingly be drawn from the aether using a simple wire antenna strung between two poles.

However, as will be seen, the antenna is not needed and the reason is that the energy source is not the radiant emission by some process involving radio wave propagation through the anther, but rather what can best be described as a phase-lock that couples the apparatus with the quantised motion of electric aether charge. There is a technique, to be described below, by which it is possible to exploit this phase-lock condition by setting up an energy oscillation involving an apparatus component and its enveloping aether, the result being that energy in an immediately useful electrical form is imported into the apparatus from that aether.

Brief Description of the Invention

According; to one aspect of the invention, an electric power delivery circuit comprises two capacitors, each having a pair of electrodes formed by a pair of metal cylinders having concentric axes, each capacitor having an associated inductor series-connected to it to form a capacitor-inductor unit, DC voltage excitation means connected to a parallel combination of the two capacitor-inductor units, whereby to apply between corresponding electrodes of the capacitors a DC bias voltage which primes them with electric charge, and power output terminals, one at each point of connection between a capacitor and its associated inductor, whereby to provide for an AC power output owing to oscillations of electric charge between the two capacitors at the resonant frequency of the capacitor-inductor units.

According to another aspect of the invention, an electric power-delivery circuit comprises two capacitors, each having a pair of electrodes formed by a pair of metal cylinders having concentric axes, each capacitor having an associated inductor series-connected to it to form a capacitor-inductor unit, DC voltage excitation means connected to a parallel combination of the two capacitor-inductor units, whereby to apply between corresponding electrodes of the capacitors, a DC bias voltage which primes them with electric charge, each inductor being the primary winding of an electrical transformer, the secondary winding of which serves to provide an AC power output owing to oscillations of electric charge between the two capacitors at the resonant frequency of the capacitor-inductor units.

According to a feature of the invention the capacitors have no intervening solid or liquid dielectric medium separating their concentric electrodes.

According to another feature of the invention, two inductors are coupled electromagnetically by having a common ferrite core and their primary windings are connected to their associated capacitors in the polarity configuration which assures that, in their mutually resonant state, electric charge is exchanged between the two capacitors.

According to yet another feature of the invention, the central axes of both cylindrical electrode capacitors are mutually parallel.

According to a further feature of the invention, an electrical power delivery system comprises a plurality of these electric power delivery circuits, where the central axes have different angular orientations as between the different circuits.

According to a still further feature of the invention, in such a power delivery system, the difference in angular orientation of the central axes is at least 60^0.

Brief Description of the Drawings

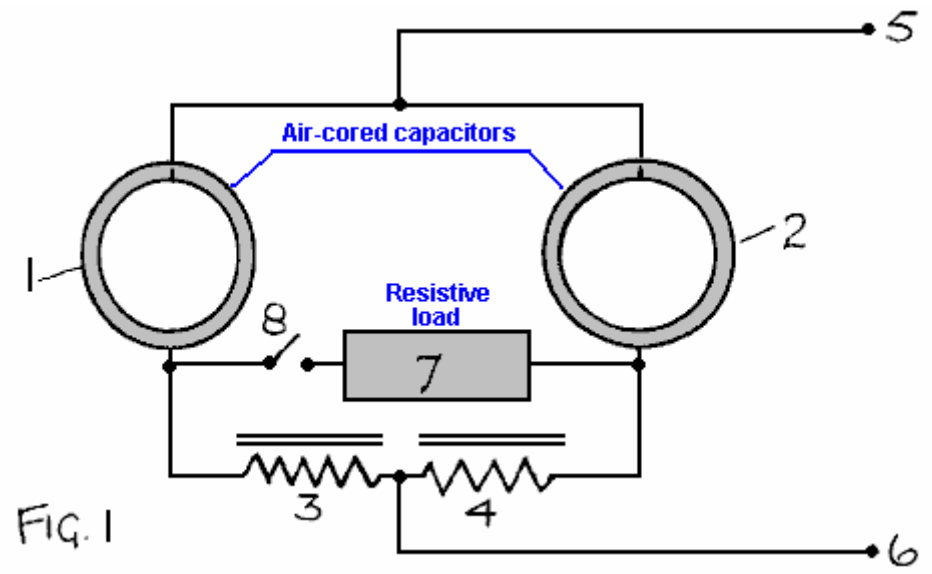


Fig.1 shows an electrical power generating circuit incorporating two concentric cylindrical capacitors having central axes which are parallel.

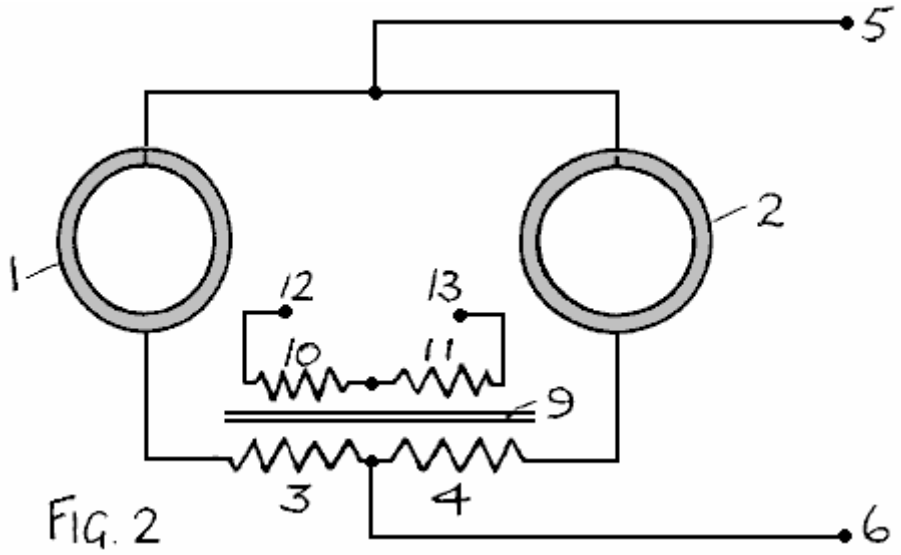


Fig.2 shows a modified version of the circuit of **Fig.1** with a transformer system providing the inductors and an output winding.

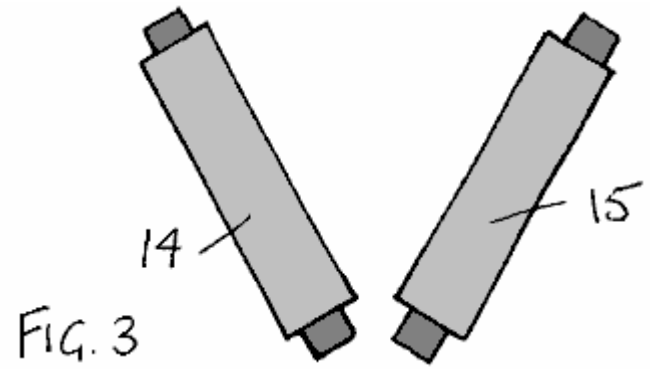


Fig.3 illustrates a mutually inclined capacitor system comprising two pairs of concentric cylindrical capacitors.

Detailed Description of the Invention

The invention draws energy from the aether. To understand why the invention works, one needs to understand the process by which the aether stores energy when an electric field is set up across the dielectric separating two capacitor plates. Moreover, one needs to understand the means by which the aether determines the quantum of action, specifically in the form of the Bohr magnetron and the unit of angular momentum linked to Planck's constant.

It is not sufficient to imagine that electric charge in the aether is displaced from a rest position in a background continuum of opposite charge polarity to which it is attracted by a restoring force. Indeed, one must consider such action to be superimposed on a system of charge which has an underlying jitter motion, a quantum theory theme associated with the German physicist Heisenberg (Zitter-bewegung, which has the dictionary meaning 'Circular fluctuation movement, of spin'). When these two factors are combined, and the constraint added of there being a phase-lock which keeps that jitter motion in synchronism as between the charges, one finds that the physical theory involved has some very interesting consequences.

One of these consequences is that a spherical or cylindrical volume of aether, if spinning bodily about a central axis, will acquire a magnetic moment and set up an electric field inside that sphere or cylinder that is directed radially with respect to the spin axis. A summary analysis is presented in the Appendix to this specification, being, in part a quotation from pages 31-33 of a booklet entitled 'The Theory of Gravitation' which the Applicant of this invention, Dr. Harold Aspden, authored in 1959 and duly published early in 1960.

The induction of electric charge by 'aether spin' was there shown to give a physical basis, both qualitative and quantitative, for the geomagnetic moment, the property of body Earth of setting up a magnetic field which created magnetic North and South poles at latitudes offset from the geographic poles, with the geomagnetic polar axis precessing slowly around the Earth's spin axis at a rate of several hundred years per revolution. By identifying its source as a rotation of a sphere of aether coextensive with body Earth, a volume of aether relative to which the Earth could have a component of motion even though the aether spin frequency is equal to that of the Earth, this axial tilt of some 17 degrees has a physical explanation. However, that aspect of the aether's role was not seen at the time as offering anything of promise technologically. The physics involved is nevertheless very relevant and directly pertinent to the experiments on which this invention is based, the findings of which would otherwise be quite baffling scientifically.

The applicant has, over the 40 or so years since the theory was first published, given a great deal of consideration to the theoretical implication that, just as aether spin can set up electric charge displacement inside coextensive matter, so the setting up of an electric field directed radially with respect to an s axis can induce aether spin about that axis and with it develop angular momentum. Indeed, in the author's onward publications on this subject, as, for example, 'Physics Unified' published in 1980 by Sabbeton Publications, P.O. Box 35, Southampton, England (ISBN 0 85056 0098), it is shown how the onset of the force of gravitation when a disordered aether consolidated into an orderly structured form caused protons to accrete more rapidly than electrons, owing to their higher mutual rate of gravitational acceleration. This created stars with all initial positive charge and the associated aether spin resulted in the stars acquiring their spin states and shedding matter which consolidated into planets which share the angular momentum so generated. The aether with its property of spin as related by its electric charge density according to the formula presented in the Appendix is therefore the key factor if we attempt to account for the creation of the stars which populate our universe.

That same formula, however, is equally valid if applied to the circumstance where a radial electric field is set up between the concentric cylindrical electrodes of a capacitor formed around a hollow dielectric cylinder. It tells us how fast the aether within that dielectric will spin. The related theoretical analysis shows that the quantum phase-lock feature of the aether imports from the external aether world an amount of energy equal to that supplied in setting up aether charge displacement, this imported energy being the dynamic energy corresponding; to the acquired aether angular momentum. Guided by the argument concerning stellar creation one can see that this aether angular momentum can be transferred to matter and this process also has its energy transfer implications.

However, one can wonder what happens if, after setting up a radial electric field in that capacitor having concentric electrodes, the applied voltage is reduced, thereby withdrawing electric field energy from the capacitor. The imported energy present in kinetic energy form as a cylindrical shell of aether spins about the central axis of the capacitor will tend to sustain electric charge displacement. To conserve energy, since the

aether phase-lock cannot force the expulsion of energy by obliging the enveloping aether universe to keep in step, this energy can only be shed by augmenting that released electrostatically. In other words, the net result is that an up and down fluctuation of the electric charge condition of the capacitor must give rise to an electric energy output that is, for the lowest dielectric constant (the permittivity of the vacuum), double the input in each cycle of change. One can then envisage an oscillation escalating in energy content powered almost wholly by aether input before one taps into that source of power to draw off energy at a rate consistent with stable operation.

This is, of course, a bewildering prediction that no physicist could imagine as being at all possible and yet, given the relevance of the theoretical argument involved, as applied to the phenomenon of geomagnetism and stellar creation, which are supported by strong evidence in that book 'Physics Unified', once such a notion is conceived it surely has to be put to the test by experiment. This then, after decades of effort before this realisation has dawned, is the basis on which the Applicant has only now come to appreciate the amazing technological possibilities that lie before us and is asserting by this patent specification that energy can in fact be tapped from the aether on a commercially viable scale.

Given that aether theory indicates that the special form of capacitor described above will, if subject to an oscillatory charge condition, generate an excess of energy, a question to consider is why such a phenomenon has not manifested itself in bench-type experiments performed in numerous electrical laboratories over the past one hundred years. Ostensibly the implication is that the capacitor will exhibit a negative resistance if used with an inductor as a component in what would become a self-resonating circuit. The answer to this may be that if such a phenomenon has occurred it has passed unnoticed or been regarded as spurious or noise-related, being something connected with radio interference etc. Alternatively, and as a function of the size and scale of the apparatus, the effect may have lacked an exciting trigger needed to overcome an energy threshold set by such factors as circuit contact resistance or contact potentials as well as the basic resistance of the inductors which, with the capacitors, form the resonant circuit.

Note that, even for a capacitor of quite large physical dimensions, having regard to its accommodation on top of a laboratory bench, the actual capacitance is necessarily quite small. being of the order of a billionth of a farad. This means that a capacitor charge fluctuation of the order of a volt would only imply energy fluctuations that are of the order of a billionth of a joule per cycle. The situation is quite different if perchance a DC bias voltage of, say, 5,000 volts is applied to the capacitor. Then a small superimposed voltage fluctuation makes the related energy fluctuations very much larger with much greater prospect of an escalating self-resonance being triggered.

With this in mind the applicant perceived a possible prior art link with the experimental claims reported by Dr. Moray who, in 1929 is said (see pages 46-50 of the above-referenced recently-published book by Keith Tutt) to have powered six 100 watt light bulbs plus a standard 575 watt electric flat iron, merely by providing an earth connection and coupling an input lead to an overhead wire antenna. The apparatus involved had no other source of input power but included a special arrangement of capacitors and presumably some kind of high frequency inductor/transformer unit.

In spite of the attention given to the Moray demonstrations, it seems that the secrets involved in the design and construction of the apparatus remain unknown and so cannot feature in the prior art of published record. Nor, indeed, can the anecdotal evidence of Moray's efforts serve to show that the subject invention has been put to prior use. The technology as to how to replicate the Moray device, always assuming it did perform as claimed, has therefore to be rediscovered and, indeed, given that there is reference to his detectors incorporating some special substance which was referred to as ‘Swedish stone', possibly the dielectric he used in his capacitor construction, there is a considerable mystery to unravel. More to the point, however, one is led to believe that Moray was implying that the energy he was tapping was radiant energy drawn from the aether, with that antenna featuring prominently because, without it being connected, the energy output fell to zero. However, as he surely may well himself have known, one just cannot draw power on such a scale from a simple overhead wire strung between two poles and so, without know how, he would have suspected that the energy inflow was coming into his capacitors via the action of that mystery substance he called 'Swedish Stone'.

The applicant here suggests that, based on an insight into the quantum workings of the aether medium as outlined above, the curious discovery demonstrated decades ago by Dr. Moray may have been attributable to setting up an oscillation in a resonant circuit including, a concentric cylindrical electrode capacitor which had a voltage bias of the order of a thousand and more volts fed from a connection to that overhead antenna but drawing no significant current from that antenna other than enough to prime his capacitor with charge

and stimulate a high frequency fluctuation which could initiate an escalating circuit oscillation tapping aether energy from the aether spin induced in the capacitor dielectric.

This is speculation, but it is sufficient to justify the Applicant's interest in constructing a capacitor and seeking to verify the assumptions just made. Notwithstanding, the reference alcove to Dr. Moray and the note below concerning Nikola Tesla, what it leads to is new invention by virtue of full disclosure of details of operation and manufacture of something hitherto unknown, the actual means by which to harness a source of energy latent in the aether medium and deemed by those familiar with state of the art knowledge to be beyond man's reach. Furthermore, there are supplementary inventive features of a special nature because of the way the subject invention exchanges energy between two capacitors and also because the optimisation of aether power output from the capacitors is found to be a function of the orientation of the capacitor axes relative to the cosmic background owing to the Earth's rotation.

It seems here appropriate to mention something described by Nikola Tesla in his U.S. Patent No. 685,958. This was filed on 21 March 1901 and granted on 5 November 1901. It was entitled: 'Apparatus for the Utilisation of Radiant Energy'. By installing two metal plates, one high above the ground and the other at ground level, with wires connecting the plates to separate electrodes of a capacitor, it was stated that the capacitor became charged to a very high potential, the energy input being that radiated to Earth from outer space. This may well have motivated the efforts of T. Henry Moray but, so far as this Applicant's invention is concerned, no such input from overhead components is necessary as a quite different energy source is at work, namely the zero-point vacuum energy activity of our quantum underworld.

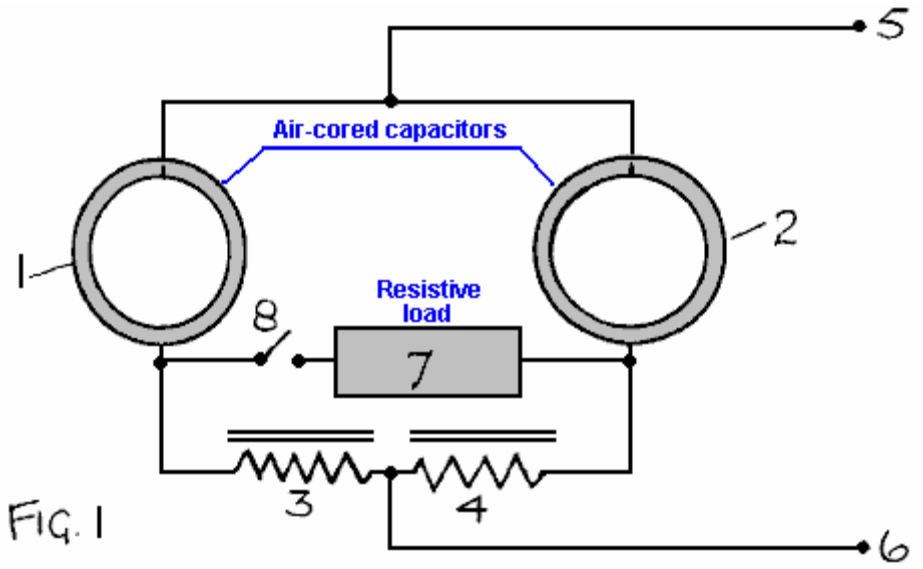


FIG. 1

Referring now to **Fig.1**, two capacitors **1, 2** formed by concentric cylindrical metal electrodes and having their central axes parallel, form part of a resonant circuit combination by each being series-connected to an inductor **3, 4** having a ferrite core. Their inner electrodes are connected to a high-voltage DC source **5** and their outer electrodes are separately connected through their corresponding inductors to a low-voltage or earth terminal **6**. A resistive load device **7** is connected via switch **8** between the junction points of the capacitors and inductors.

In operation, owing to spurious electrical signals induced in the inductors, or to an imposed electrical stimulus provided by means not shown, the priming electric charge of the two capacitors will develop oscillations as charge is exchanged between the two capacitors. There is energy inflow owing to the quantum coupling of electric charge displaced between the concentric electrodes of each capacitor and the quantum activity of the underworld of the enveloping aether. This affords an electrical energy output which is supplied upon closure of switch **8**.

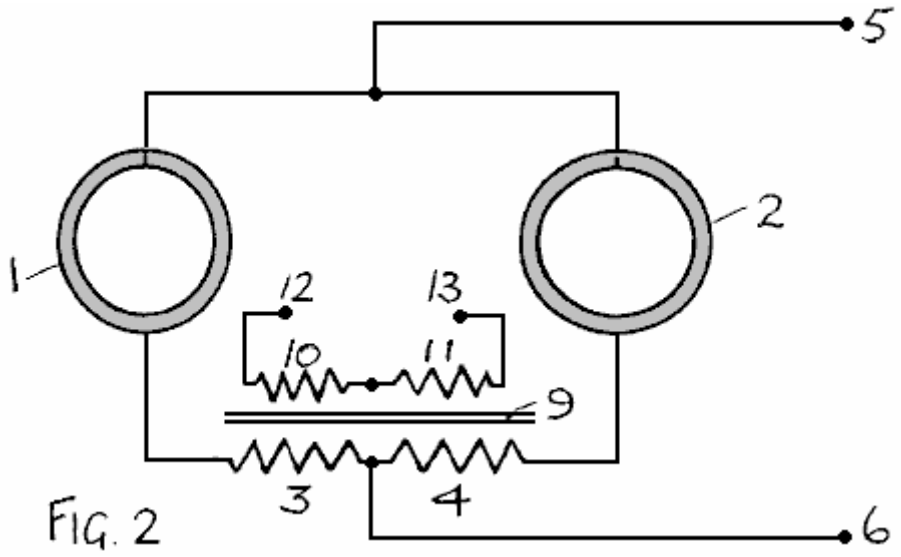


FIG. 2

Referring to **Fig.2**, the inductors **3, 4** are shown to have a common ferrite core **9** and to have secondary windings **10,11**, which, by transformer action, can supply electrical power output between terminals **12** and **13**.

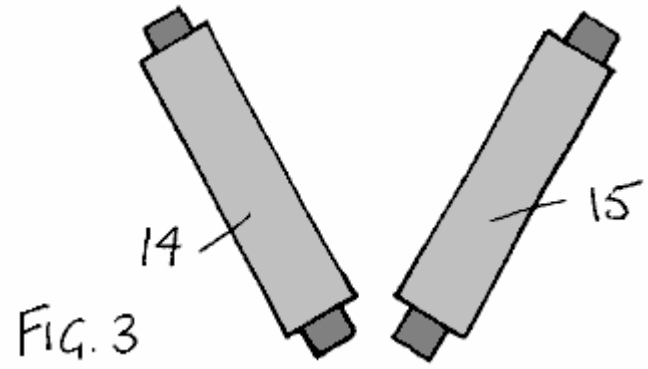


FIG. 3

The apparatus of **Fig.1** and **Fig.2** will, when viewed in side elevation, appear as having a capacitor form with an outer cylindrical electrode within which there is a slightly elongated inner cylindrical electrode, to facilitate the high-voltage connection to that inner electrode. **Fig.3** shows, in very simple diagrammatic form, two such arrangements **14, 15**, with the central axes of the two pairs of capacitors mutually inclined. There may, however, be three or more such pairs of capacitors, each pair constituting a circuit such as is depicted in **Fig.1** or **Fig.2**.

The reason for configuring multiple capacitor systems, each with its own power output, in a combined manner with the outputs merged to supply an overall energy producing system is that the aether energy output of each capacitor unit is a function of axis orientation. This is because the quantum activity of the aether has its own preferred axis and, as the Earth rotates there is variation of the relative axial orientation in a daily cycle. Also, one needs to cater for systems applying, this invention in a mobile application, which also implies change of orientation and by having; the mutually inclined capacitor axis configurations one can be assured that the potential power output avoids the null situation that can occur if the capacitor axes of a stand-alone unit of **Fig.1** or **Fig.2** were to be at right angles to the aether quantum spin axis.

The capacitor electrodes can be of thin metal sheet foam and so of light weight and preferably are not spaced apart by any dielectric medium, whether liquid or solid. They need to be held apart by a simple insulating frame structure. The reason is, that the only dielectric medium that is operative in the functioning of the invention is the vacuum medium and to have a normal dielectric present implies more capacitance and so extra current oscillation without extra energy gain per cycle of oscillation. The key factor assuring operation is the need for circuit resistance to be low compared with capacitance that is solely attributable to the vacuum medium combined with the high voltage priming which greatly enhances the power output to weight factor.

The two capacitors of a pair are preferably of identical capacitance and structure, as are the inductors, so that the oscillation period of the two resonant sectors of the circuit is the same. The common ferrite core feature of the **Fig.2** configuration assists in this role.

The apparatus will normally be designed to operate at a capacitor frequency of the order of 100 KHz or more, and a voltage of 10,000 V or higher, and so the transformer output of **Fig.2** will be preferable with voltage duly adjusted to suit the application. The high frequency AC so produced can then be converted as needed by using the appropriate technology of known form.

Appendix

Extract from pp. 30-31 of 'The Theory of Gravitation', 1960 printed publication by the Applicant. Note that the earlier pages explained that the aether comprises a system of electric particles in a cubic crystal-like distribution set in a uniform background continuum of opposite charge polarity, the particle system and the continuum both sharing a common circular orbital motion of radius **r** and the relative velocity between the particles and continuum being the speed of light.

The Effect of Aether Rotation

Consider what happens when a large volume of the aether is rotating bodily. The continuum and particle system rotate together. There will be no resultant magnetic moment unless the particle distribution is disturbed. An evident disturbance is the centrifugal effect arising from aether rotation, but for the angular velocities of magnitude found in the solar system this effect is of negligible consequence. A much more important effect arises from the synchronising interaction between particles in the rotating volume. This requires that the particles shall move about their neutral points at the same angular velocity. Thus if a particle is to have a velocity component **V** directed in the plane of its orbit, whilst retaining a mean velocity **C**/2, its speed along its orbit must be of the form **C** / 2 + **V** cos(**P)**, where **P** is the angle subtended by a line joining the particle and the centre of its orbit relative to a fixed reference datum in the inertial frame. To satisfy the above requirement the centre of the orbit cannot be the neutral point. Evidently the particle is distant from this neutral point by **r** + (2 **V r** / **C**) cos(**P)**. As **V** is much less than **C** the effect of this is that the particle is moving around a circular orbit whose centre has been displaced a distance 2 **V r** / **C** perpendicular to **V** in the plane of the orbit. If **V** is much less than **w x** cos(**A)**, where **w** is the angular velocity at which the aether rotates, **x** is the distance of the aether particle from the axis of rotation, and **A** is the angle of tilt of the axis to the common axial direction of the aether particle system, this displacement distance is 2 (**w x r** / **C**) cos(**A)**. Consider a disc-like section of the rotating aether of radius **x** and unit thickness. Then, the effective charge displacement arising from the effective physical displacement of the particles is 2 **pi x s** (2 **w x r** / **C**) cos(**A)**. The disc has acquired a uniform charge density of 4(**w r s** / **C**) cos(**A)** esu/cc. The polarity of this charge depends upon the direction of rotation of the aether.

When evaluated from the aether data already presented, the charge density is found to be: 4.781 **w** cos(**A)** esu/cc. This charge density represents a charge component which rotates with the aether.

Calculation of the Geomagnetic Moment

For Earth, **w** is 7.26 x 10^{-5} rad/sec and **A** is 23.5^{O}. Thus the Earth's charge density is, from the above expression, 0.000319 esu/cc. The rotation of this charge gives rise to a magnetic moment of:

(0.000319)(4 **pi** / 15)**w** $\mathbf{R}^5$ / **C** where **R** is here the radius of the Earth's aether.

If **R** is greater than the Earth's radius (6.378×10^8 cm) by a small factor **k**, the Earth's theoretical magnetic moment becomes (1 + 5**k**)6.8 x 10^{25} emu. This may be compared with the measured value of the Earth's magnetic moment of 8.06 x 10^{25} emu.

An upper limit of 0.035 is imposed on **k** suggesting the Earth's aether terminates at a mean height of about 140 miles above the Earth's surface. This suggests that the ionosphere may be a phenomenon arising at the aether boundary.

Claims

1 An electric power delivery circuit comprising two capacitors each having a pair of electrodes formed by a pair of metal cylinders having concentric axes, each capacitor having an associated inductor series-connected to it to form a capacitor-inductor unit, DC voltage excitation means connected to a parallel

combination of the two capacitor-inductor units, whereby to apply between corresponding electrodes of the capacitors, a DC bias voltage which primes them with electric charge, and power output terminals, one at each point of connection between a capacitor and its associated inductor, whereby to provide for an AC power output owing to oscillations of electric charge between the two capacitors at the resonant frequency of the capacitor-inductor units.

2 An electric power delivery circuit comprising two capacitors, each having a pair of electrodes formed by a pair of metal cylinders having concentric axes, each capacitor having an associated inductor series-connected to it to form a capacitor-inductor unit, DC voltage excitation means connected to a parallel combination of the two capacitor-inductor units, whereby to apply between corresponding electrodes of the capacitors a DC bias voltage which primes them with electric charge, each inductor being the primary winding of an electrical transformer, the secondary winding of which, serves to provide an AC power output owing to oscillations of electric charge between the two capacitors at the resonant frequency of the capacitor-inductor units.

3 An electric power delivery circuit according to Claim 1 or 2, wherein the capacitors have no intervening solid dielectric medium separating their concentric electrodes.

4 An electric power delivery circuit according; to Claim 1 or 2, wherein the capacitors have no intervening liquid dielectric medium separating their concentric electrodes.

5 An electric power delivery circuit according to Claim I or 2, wherein the two inductors are coupled electromagnetically by having a common ferrite core and their primary windings are connected to their associated capacitors in the polarity configuration which assures that, in their mutually resonant state, electric charge is exchanged between the two capacitors.

6 An electric power delivery circuit according to Claim 1 or 2, wherein the central axes of both cylindrical electrode capacitors are mutually parallel.

7 An electric power delivery system comprising a plurality of electric power delivery circuits according to Claim 6, wherein the central axes have different angular orientations as between the different circuits.

8 An electric power delivery system according to Claim 7, wherein the difference in angular orientation of the central axes is at least 60^0.

Comments by Dr Aspden on 19th March 2006:

OUR ENERGY FUTURE

A Message of Vital Importance

The world needs a new source of energy, one that is not an exhaustible commodity subject to power-play as between nations. Yes, one can dream and then awake to say this is impossible, but I urge those with the necessary skills to heed what I have to say in my three messages below.

First, however, let me introduce myself. My name is Dr. Harold Aspden. I am retired and elderly but have had a lifelong scientific interest in fundamental physics relevant to the energy theme. My 6-year university education in U.K. was at Manchester University and Cambridge University (Trinity College). My 33-year working career in U.K. comprised 9 years with English Electric and 24 years with IBM. Though having high technical qualifications (see below), being interested in the specialised field of protecting inventions pertaining to electrical engineering, I became a Chartered Patent Agent and later a European Patent Attorney. My last 19 years with IBM were spent as Director of IBM's European Patent Operations. This was followed, in my early retirement, by 9 years as a Visiting Senior Research Fellow at Southampton University and thereafter my scientific interest has been a private pursuit evidenced by my writings as on this and my related websites. My formal qualifications are: B.Sc., Ph.D., C.Eng., F.I.E.E., F.I.Mech.E., C.Phys., M. Inst.P., C. Sci., Wh.Sc.

Message No. 1: Physicists have come to recognise that there exists a quantum underworld alive with energy and permeating all space. However, their related research aims merely at probing experimentally the spectrum of elementary particles that have a transient existence as a product of that energy activity. The reward they seek is recognition should new particles be discovered and, by their properties, reveal

connections with other particles that help in formulating a new theory or verifying an existing theory. Sadly, they do not see that quantum underworld as a potential source of energy that we can harness. Nor have they understood how most of the energy shed in creating matter formed the elementary particle which bears the name proton and which, together with the electron, constitutes the hydrogen atom.

There is also a secret they have yet to fathom. It is the effect of creating a radial electric field centred on electrical charge around which that quantum underworld can develop a state of spin that causes it to shed energy. In the presence of a radial electric field set up by an electrically charge body, whatever constitutes that quantum underworld that permeates all space shares a motion like that of sequence dancers who keep in step with one another as they move around the dance floor, a synchronous motion, which, in the presence of that radial electric field can only be held if a secondary motion develops around an axis centred in that radial field.

How else could the Sun spinning about its own axis have come into existence? Here we have gravity attracting hydrogen atoms and pulling them so closely together that ionisation occurs, meaning freeing some electrons from their proton bonding, and so, because the mass of a proton is very much greater than that of the electron, creating a Sun having a body that is positively charged sitting within an outer shell of negative electron charge. Two free protons experience a mutual rate of gravitational acceleration that is 1836 times that experienced by the interaction of two electrons. The body of the Sun, therefore, has a uniform mass density and a uniform positive charge density enclosed within a compensating negative charge at its surface. This is because gravitational compaction forces balance the expansion forces attributable to electrostatic repulsion. It further means the presence of a radial electric field within the body of the Sun and, in turn, owing to the effect of this field on the space medium of the quantum underworld, this induces a state of spin accompanied by release of energy from that medium to feed the kinetic energy of that spin.

In depth analysis of the physics involved, meaning the effect of the resulting radial electric field on that quantum underworld, then allows one to calculate the resulting rate of spin and thereby understand how the solar system was created.

So, if the reader is a physicist, here is the way forward and full guidance on this is to be found on my parallel website www.aspden.org or in a new book of mine entitled Creation - The Physical Truth, that will be published in the near future. However, if the reader is not a physicist but has the technological aptitudes of the university-trained electrical engineer then it is Message No. 2 below that warrants attention.

Message No. 2: If it were possible to generate electrical energy by tapping an omnipresent medium it is surely to be expected that the occasional natural phenomenon might already have hinted at this possibility. Consider, therefore, the thunderball, a glowing spherical object sometimes seen, especially following a lightning storm. It appears aethereal in the sense that it can move unimpeded through matter, yet remains an enigma, an unsolved mystery of record in the annals of science. Lightning strokes are high current discharges which, as electrical engineers well know, can develop a 'pinch effect' squeezing the electron-carried current into a filamentary flow within a cylindrical channel of positively charged air. That implies a radial electric field, a pulsating radial electrical field if the discharge surges, a sure recipe for something to happen that could form a miniature Sun, the thunderball. So when we look at a thunderball we are looking at a natural phenomenon that has drawn energy from that quantum underworld of space, energy which is then dissipated, but energy shed by a process we can surely harness, once we understand the physics involved.

Scientists lacking the necessary imagination do not seek to understand how the thunderball is created and so they seldom write about it. So here we have something to think about. It is Nature's message telling us: "Produce a radial electric field, one that pulsates, and you can develop a spin that taps energy from the quantum underworld of space." As engineers, however, we need to be practical and, if possible, we should avoid trying to replicate a phenomenon that involves powerful electric discharges, if there are better ways in which to proceed.

So now I come to my primary theme in this Message No. 2. It is a brief survey of a few of the claims of record that have declared a mysterious energy gain and have features which I see as relevant to what has been said above. In particular I draw attention to the research findings of four different pioneers in what has come to be termed 'The Search for Free Energy', this being the title of a really excellent book by Keith Tutt, published by Simon & Schuster in 2001. Three of these are described in considerable detail in that work. I now ask you to keep in mind my reference to a radial electric field as I mention each of them below and do realise that electrical structures of cylindrical form are a key feature.

Nikola Tesla is famous for his research concerning electromagnetic induction and high voltage solenoidal transformer apparatus (Tesla coils) and he is said to have demonstrated an automobile which derived its

power by tapping energy from space. He did not disclose its design details and died leaving us with a mystery. Tesla coils comprise large solenoidal windings concentrically mounted and operate with high voltage pulsations between their cylindrical forms which must produce a pulsating radial electric field between those windings. So, although electromagnetic induction effects are the primary focus of attention, there is here scope for the electrical action described in Message No. 1 above. Tesla may well have stumbled experimentally upon a way of tapping energy from space, but without understanding the true underlying physical process.

Dr. Henry Moray, a pioneer of the 1920-1930 era, demonstrated something which merely needed a kind of antenna, a wire connected from tree tops to earth via electrical apparatus in the boot (trunk) of his automobile. It is said that the latter included several capacitors and that a kilowatt level of power was generated. In this case the automobile merely carried the test apparatus for demonstration at a location remote from a built-up area and any electrical power line interference. No doubt Moray was seeking to follow in Tesla's footsteps by drawing energy from the Earth's electric field, known to be measured in hundreds of volts per metre. It is likely that those capacitors were of Leyden jar type configuration, that is cylindrical in structural form, and that the wire linked to tree tops tapped charge at a kilovolt voltage level. However, the output power claimed could surely not have come from that source. Therefore one must assume that Moray used that treetop voltage input merely to prime the voltage across his capacitor electrodes, whilst incorporating some special feature in the operation of his electrical circuit that gave access to the energy of the quantum underworld. Capacitors having concentric electrodes of cylindrical form will, when charged electrically, have a radial electric field in the space between the electrodes. Several capacitors coupled together could give rise to oscillations of charge as between the capacitors and so lead to a pulsating radial electric field. Yet though demonstrating as possible something that should not be possible, a mysterious inflow of energy able to illuminate several light bulbs, Moray could surely not have understood the true physical process that was feeding energy into his apparatus. Again I see this as relevant to what is stated in Message No. 1.

Stan Meyer demonstrated apparatus that included sets of concentric tubular electrodes enclosed in a cylindrical container filled with water, the electrodes being fed by high voltage (5 KV) pulses. Combustible gas was generated, a mixture of hydrogen and oxygen, the burning of which generated far more heat than could be accounted for by the electrical energy input. Energy was being tapped as if from nowhere unless the source was the ambient medium of space itself. Here there was a pulsating radial electric field and electric charge oscillating between different components in Meyer's apparatus. Meyer did not offer any useful explanation as to the physical process underlying what he could demonstrate but persisted in conveying the message that the invention was wonderful and talking about a multiplicity of applications such as powering automobiles, ships etc. This is the project not mentioned in Keith Tutt's book. As for the Tesla and Moray projects Meyer's research was a U.S. based activity. It did, however, attract the interest of a British Admiral, Admiral Tony Griffin who was concerned with the impact of new technology upon the marine industries. Griffin witnessed Meyer's demonstrations and was interested in its development. Indeed an article on the subject mentioning Admiral Griffin and entitled 'Free Energy for Ever' was published in the January 1991 issue of the U.K. magazine Wireless World. The importance of the article was evident from the fact that the Editor of that magazine was the author.

Paul Baumann, a member of a Christian community in a isolated valley high in the Swiss Alps has constructed working free energy devices which have been demonstrated to visitors. The first working prototype was relatively small and included a pair of glass Leyden jars, concentric capacitors. Keith Tutt in his book devotes 30 pages to this subject. The high voltage needed to prime the capacitor operation was generated by a Wimshurst machine driven by the electric power generated. The community has, however, kept design details secret. In spite of such information as is available the underlying physical process governing its operation remains a mystery. Yet I can but feel confident that what I say in my Message No. 1 provides the answer.

Message No. 3: My Message No. 1 has drawn attention to the physical process by which the vast amount of energy needed to create the Sun was extracted from the quantum underworld that permeates all space. My Message No. 2 has drawn attention to the reported efforts of just some of the several energy research pioneers who actually demonstrated apparatus that, contrary to accepted scientific principles, drew energy from a mystery source. My Message No. 3, based on recognising the common physical feature can but be the suggestion that technology for generating our power needs from the hidden underworld of space has to be possible. Accordingly, I will now outline what I see as the basis on which to build the ultimate power generating device that harnesses the physical principles presented in Message No. 1.

Being 78 years of age and no longer having access to university research laboratory facilities, I can but leave it to others to take note and, hopefully, prove me right. If proved right then the world will benefit and the

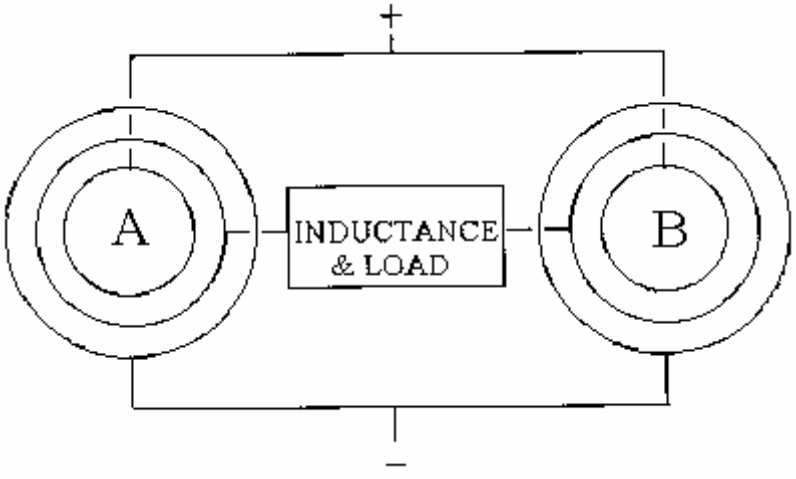
impending energy crisis will be avoided. Hopefully also, the scientific community might then be willing to accept my claim as to how the quantum underworld deploys its energy into proton creation and is active in producing the phenomenon of gravitation. I know of no other theory that has been able to derive theoretically the value 1836.152 of the proton/electron mass ratio. I would like to see that recognised as my contribution to man's knowledge.

Consider a capacitor formed by a pair of concentric cylindrical electrodes, something many of us remember from the school physics laboratory, the Leyden jar. However, the capacitor structure I have in mind is very much larger and has to be operated at a quite high voltage. When that voltage is applied between the electrodes electric charge is displaced in the underlying vacuum medium located between those electrodes. A commensurate amount of electric charge is thereby held in place on those electrodes, a negative polarity charge on one and a positive polarity charge on the other. Given my claim that this is accompanied by 'vacuum spin', aether rotation, which has imported an equal amount of energy owing to a quantum phase-lock as between the charge of the vacuum medium, we have the energy gain we seek to exploit.

The problem, however, is that, with this simple capacitor configuration, the only control parameter available is the reduction of the voltage between the electrodes. This will shed energy within the circuit of the apparatus used, the outflow of electric charge at the voltage difference merely delivering energy equal to that originally supplied by our voltage source. The added energy imported from space is merely dispersed by the 'vacuum spin' slowing down but expanding beyond the bounds of the capacitor electrodes as it conserves its angular momentum. The energy imported from the quantum underworld of space has no way of enhancing the energy output of the capacitor circuit and so is left to dissipate itself and eventually be reabsorbed by that quantum underworld that pervades all space.

However, now consider a concentric electrode capacitor having a third cylindrical electrode intermediate the inner and outer electrodes. Here we have a control parameter other than the voltage between the outermost and innermost electrodes, because we can wonder about the voltage of the central electrode whilst retaining the other voltage difference at a constant high level. In fact, by keeping the latter voltage difference constant but varying the voltage of the intermediate electrode we can decrease the capacitor energy of one half of the overall capacitor as that of the other half decreases. The imported energy shed by one half of the overall capacitor can then contribute to the action that energises the other half and thereby induce oscillations from which energy can be extracted and deployed as a power source.

One needs two such capacitors having their central electrodes coupled through a load circuit in order to capture the 'free energy' inflow and get it to do useful work rather than being dissipated. An inductance in the coupling circuit can determine the oscillation frequency and, since the energy inflow increases with frequency, this should no doubt be well into the kilocycle region. The figure below is a simple schematic diagram of the electrical apparatus that I have in mind.



So my Message No. 3 is what I may describe as a 'thought experiment', one that I cannot verify myself, owing to my age and lack of facilities. I therefore can but record my thoughts and hope that others will prove me right and not wrong.

The capacitors depicted in the figure should have their electrodes spaced so that the capacitance C as between their central and outermost electrodes is the same as the capacitance C between their central and innermost electrodes. Suppose that the outermost electrodes are maintained at a voltage of 20,000V relative

to the innermost electrodes. This means that the two central electrodes will be at an intermediate voltage which we expect to be 10,000V in the absence of oscillations. However, as with any ever-active electrical system, there will be minor voltage fluctuations affecting the central electrodes. So we may ask what happens if the voltage of the central electrode of capacitor A decreases owing to electric charge being shed by the inner capacitance C but gained by the outer capacitance C. Think about that for a moment. You will see that it implies reciprocal action in the opposite sense by capacitor B, as current flows from A to B via the central inductor coupling. Yet no net current flows from the 20,000V power source.

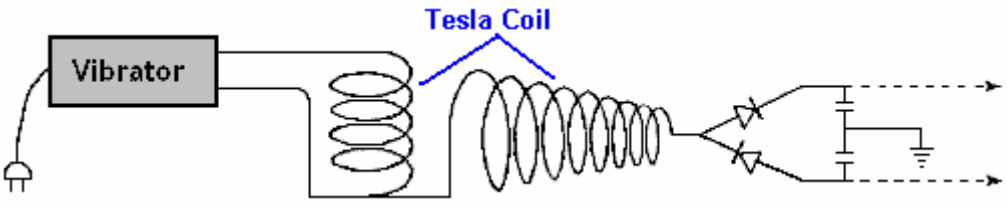
Now, of course, common sense backed by our scientific training assures us that this system can but keep its equilibrium without those minor voltage fluctuations building up in some way. Yet, if we heed Message No. 1 and keep in mind Message No. 2, there is a question we must ask. If current does flow through that central link between A and B, one half of A and one half of B both shed energy and so release the imported 'vacuum spin' energy, if such is present. This occurs as other halves of A and B have to gain energy and as angular momentum of the imported 'spin energy' spreads into the other sections of the capacitors. The question then is: "Does that imported energy escape, as it does for the two-electrode capacitor configuration, or might it be retained and so augment the action?"

I submit the answer can only be provided by actual experiment. If the energy does escape then there is nothing further to discuss. However, if some of that energy is captured then we can expect an escalation of oscillations in that inductive link and so can then say that a new source of energy has been discovered. Those oscillations will be a function of the capacitance C and the inductance of the load circuit. Given a high frequency and a high voltage a significant level of power per unit volume of capacitor structure will be produced. If power output at a level commensurate with the claims of Tesla, Moray, Meyer and Baumann results the world's energy future is then assured. A pollution-free energy resource powered by the quantum underworld of space will be at hand wherever we are on body Earth.

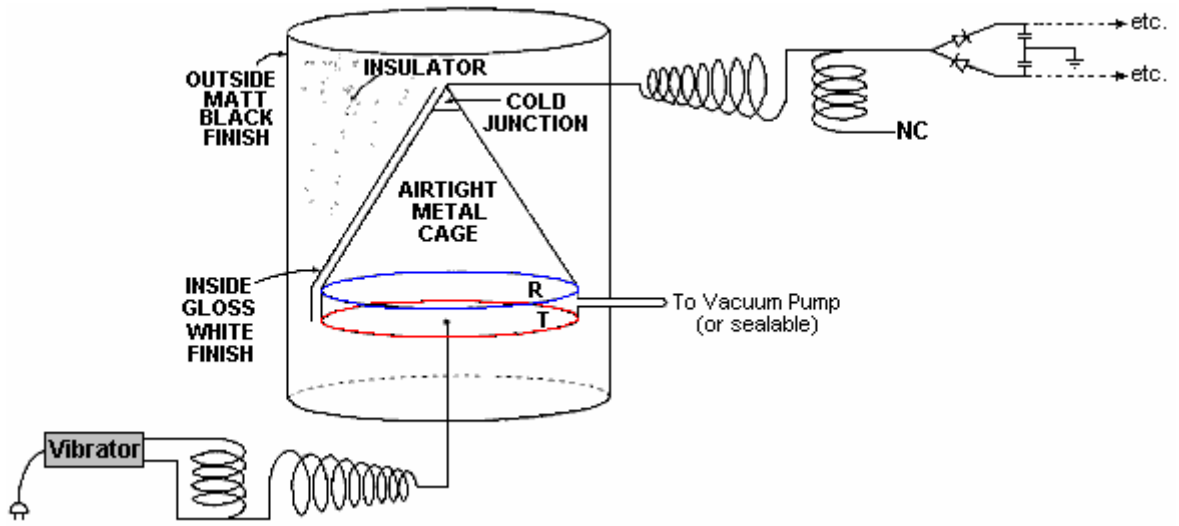
The Design of Paulo and Alexandra Correa.

Paulo & Alexandra Correa have discovered a way of converting Tesla's longitudinal waves into ordinary electrical power. They have made US Patent Application 2006/0,082,334 entitled "Energy Conversion Systems" in which they show various ways of achieving this energy-type conversion.

Their techniques range from applying the longitudinal wave energy coming from a Tesla Coil directly to two capacitors via diode rectification and the voltages generated are related directly to actual ground earth potential:



The patent application forms part of this set of documents so the full details can be examined. A theory of operation is presented based on their many experiments and observations, and the practical form of one of their conversion devices is:



Where the active pick-up plates **R** and **T** are encased in a cylinder and are provided with a cone shape to assist the procedure. The patent application contains a good deal of information and is worth reading.

Professor Konstantin Meyl.

Another key person in the advancement of current theory and analysis is Professor Konstantin Meyl who has described how field vortices form scalar waves. He has described how electromagnetic waves (transverse waves) and scalar waves (longitudinal waves) both should be represented in wave equations. For comparison, transverse EM waves are best used for broadcast transmissions like television, while longitudinal scalar waves are better for one-to-one communication systems like cell phones.

Konstantin Meyl

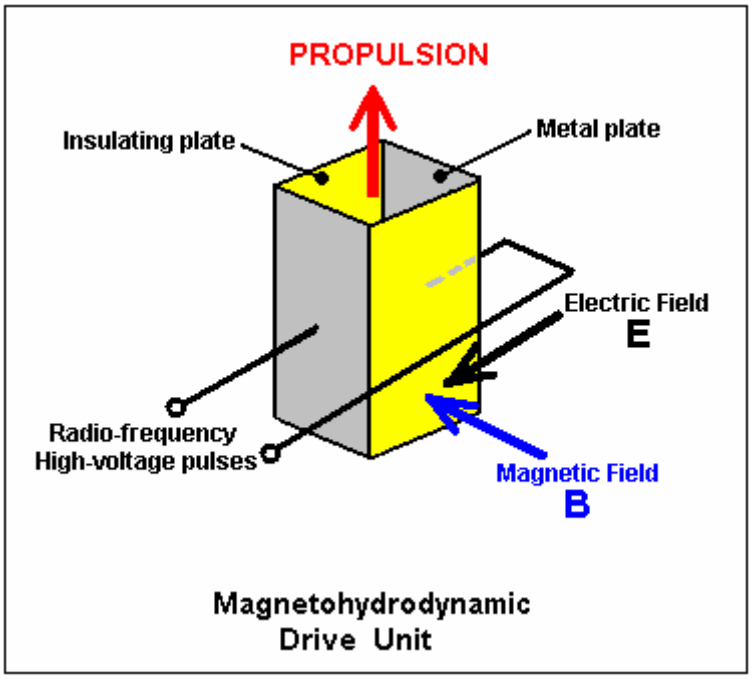
He also presented the theory that neutrinos are scalar waves moving faster than the speed of light. When moving at the speed of light, they are photons. When a neutrino is slowed to below the speed of light, it becomes an electron. Neutrinos can oscillate between e- and e+. Fusion involves e-, and a lightning flash involves e+. Energy in a vortex acts as a frequency converter. The measurable mixture of frequencies is called noise.

Dr. Meyl has pointed out that Tesla measured the resonance of the Earth at 12 Hz. The Schumann resonance of the Earth is 7.8 Hz. Meyl shows how one can calculate the scalar wave of the Earth to be 1.54 times the speed of light. He has developed a model which ties the expansion of the earth to be the result of the earth's absorption of neutrino energy. The ramifications of this model are that neutrino energy can be tapped. He took this to the next step and postulated that Zero Point Energy is neutrino power – energy from the field; available at any time, and present everywhere. To show the place of neutrinos in conventional science, Meyl noted that the 2002 Nobel Physics prize was in regards to work on neutrinos. The English-language version of Dr. Meyl's web site is at http://www.meyl.eu.

Nikola Tesla's Magnetohydrodynamic Propulsion Drive.

Tesla performed an experiment in which he applied high-voltage high-frequency alternating current to a pair of parallel metal plates. He found that the 'space' between the plates became what he described as "solid-

state" exhibiting the attributes of mass, inertia and momentum. That is, the area transformed into a state against which a mechanical push could be exerted. This implied that, using this technique, it should be possible to produce a spaceship drive anywhere in space, if the mechanism for thrusting against the 'solid-state' space could be determined. Further experiments convinced Tesla that powerful electromagnetic waves could be used to push against (and pull against) what appears to be 'empty space'. The drive principle is based on the Hall-effect used in semiconductor magnetic sensors, and is called the magnetohydrodynamic ("MHD") effect. This might be illustrated like this:



Here, a box is constructed with two metal plates forming opposite sides and two insulating plates holding them in position and surrounding an area of 'space'. High-frequency, high-voltage alternating current is applied to the metal plates and this creates an electric field "**E**" acting between the plates as shown in black. A magnetic field "**B**" is generated by the electrical field. The magnetic field acts at right-angles to the electric field, as shown in blue. These two fields produce a propulsion thrust "**F**" shown in red in the diagram. This propulsion force is not produced by ejecting any matter out of the box, instead, it is produced by a reaction against the 'solid-state' condition of space-time caused by the high-frequency electromagnetic pulsing of that area of space. This is enormously more effective than a jet engine. The thrust increases with the fourth power of the frequency, so if you double the frequency, the effect is sixteen times greater.

To put this into perspective, consider the force being applied against gravity to lift an object into the air. The force pulling the object downwards is gravity and its strength is given by:

Gravitational force:

$\mathbf{F} = \mathbf{g} \times \mathbf{M} \times \mathbf{m} / \mathbf{r}^2$

where
g is the gravitational constant (6.672×10^{-8} cm^3 g^{-1} s^{-2})
M is the mass of the first body
m is the mass of the second body and
r is distance between the two centres of mass

The lifting force is given by:

Lorentz Force: Force on an object = Electric force + Magnetic force

$\mathbf{F} = \mathbf{q} \times \mathbf{E} + \mathbf{q} \times \mathbf{v} \times \mathbf{B}$

where
q is the charge on the object,
B is the magnetic field,
v is the velocity of the object and
E is the electric field

How do these forces compare? Well, the electromagnetic force is stronger than the gravitational force by a factor of about 2,200,000,000,000,000,000,000,000,000,000,000,000,000 times. That number (2.2×10^{39}) is too big for anybody to really visualise, so let me put it another way.

If the amount of energy used to mechanically lift an object a distance of one hundredth of an inch (one quarter of a millimetre) off the ground, were used as an electromagnetic lifting force, then that amount of energy would lift the object more than 3,472,222,000,000,000,000,000,000 miles off the ground, or in metric units, more than 5,588,001,700,000,000,000,000,000 kilometres off the ground. This kind of drive is an entirely different kind of animal. This Hall-effect type of drive if used in a spaceship would require only a very small amount of input power to drive the ship at great speeds and over great distances.

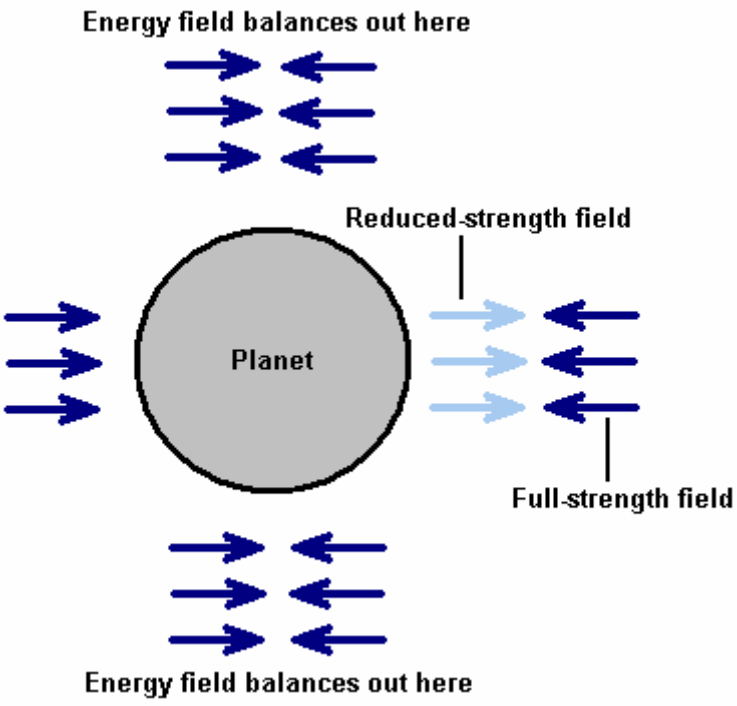
As the device shown above operates directly on the space-time field which penetrates all matter, there would appear to be no reason why it should not be used to drive a conventional vehicle by positioning it in a horizontal position rather than the vertical position shown in the diagram. Throttle operation could be by very slight adjustment to the frequency of the AC pulses applied to the metal plates. However, Bill Lyne indicates that horizontal movement is better achieved by producing Tesla's very short, high-voltage high-frequency DC pulses at the front of the vehicle while at the same time generating very high-voltage high-frequency AC waves at the back of the vehicle. This style of drive is said to pull the vehicle along rather than push it along.

The **Unified Field Theory** is being searched for by scientists who want to come up with a theory which encompasses the force of gravity with the electromagnetic force. In my opinion, they would have more chance of success in trying to find a needle in a haystack which does not contain a needle since when the entire haystack has been disassembled, it becomes clear that there never was a needle in it. In my opinion, there is no such thing as a "force of gravity", in fact, there is no such thing as gravity. Find that hard to believe? Well, let me explain.

If when standing, you hold an object a waist level and let it go, it "falls" and lands near your feet. Yes agreed, and yet I suggest that there is no such thing as gravity. If you suspend a pendulum close to a mountain, the pendulum does not hang down vertically but moves slightly towards the mountain. This is said to be because the mountain attracts the pendulum. Sorry Chief, but I suggest that it just ain't true and the mountain does **not** attract the pendulum. The Moon orbits around the Earth which requires a continuous acceleration inwards towards the Earth and this is said to be caused by the attraction of gravity pulling the two bodies of matter together. Well, yes the Moon does orbit the Earth but not because of "the force of gravity".

The reason why "the force of gravity" is so tiny compared to electromagnetism is because there is no such force at all. Yes, indeed, all of the observed phenomena which are supposed to be gravitational, do exist exactly as seen, but I suggest that there is no such thing as "the force of gravity" and the Unified Field Theory is not needed. Let me explain:

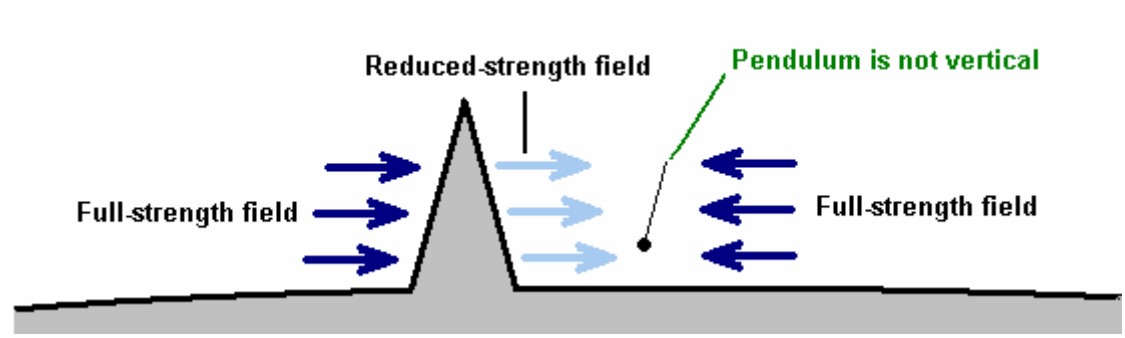
The Zero-Point Energy field exists everywhere in the universe and it flows in every direction equally. It acts like a flow of particles thousands of times more tiny than electrons, and so, it flows through matter. No matter can shield completely from the flow of this energy field. But, a tiny percentage of the flow does happen to collide with the electrons, atoms and molecules of matter as the energy flow moves through matter. The bigger the chunk of matter, the more of the energy flow collides with it. The collisions convert the energy into additional mass, which is why our Sun is not losing mass as rapidly as theory would predict. The situation is like this:



The force of the Zero-Point Energy field is slightly reduced having passed through (and interacted with) the large mass of the Earth. This reduced strength in indicated in the diagram by the light-blue arrows. The incoming Zero-Point Energy field is not reduced in strength in any significant way as the molecules in the atmosphere are not nearly as tightly packed as those in the matter which makes up the Earth itself. The imbalance of these two thrusts causes a net push towards the surface of the Earth.

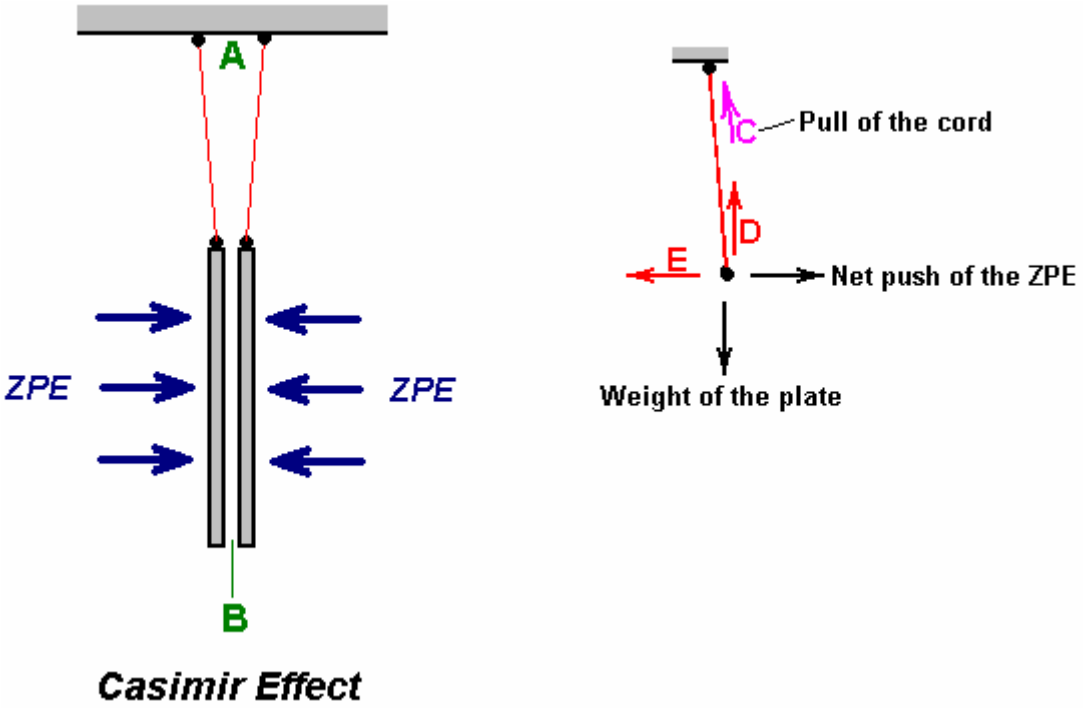
For clarity, the diagram only shows the field acting in one direction, while in reality, the same situation applies in every possible direction around the planet. When you let an object go and it moves towards the surface of the planet, it is not being pulled down by “the force of gravity”, but instead, the downward push of the Zero-Point Energy field is greater than the upward push of the Zero-Point Energy field which has just passed through the planet. The object moves “downwards” because the push from above is greater than the push from below.

Exactly the same thing applies to cause the effect that a mountain appears to have on a pendulum. In reality, the mountain has no effect on the pendulum, apart perhaps from a minor electrostatic influence. The main effect is caused by the flow of the Zero-Point Energy field:



Here, the (very roughly drawn) mountain, reduces the push of the Zero-Point Energy field which passes through it, due to its interaction with the matter with which it collides on its trip through the mountain. The push of the Zero-Point Energy field on the side of the pendulum is not diminished, so there is a net push towards the mountain and that makes the pendulum move in the direction of the mountain. The effect is not very large, so the pendulum does not move much out of the vertical as the downward push towards the surface of the planet is quite marked, so the pendulum needs to be very near the mountain for this effect to be observed.

This can also be seen in the Casimir Effect where two non-magnetic metal plates, which are not carrying an electrostatic charge, are suspended very close to each other. The plates do not hang straight down but move towards each other. This is the same effect as is caused by a mountain near a stationary pendulum, or plumb-line. Each plate screens out a little of the Zero-Point Energy field which passes directly through both plates, so the second plate gets slightly less of a push:



The result is that between the plates, the horizontal force pushing them is unbalanced. Hang just one plate up and the horizontal Zero-Point Energy ("ZPE") forces coming from the right exactly balance the ZPE forces coming from the left, and the plate hangs vertically below its point of suspension with the supporting cord (shown in red in the diagram above) hangs vertically. But with two plates as shown, the push from the left is reduced very slightly as it passes through the left hand metal plate. This means that there is a lesser push from left to right on the right hand plate. This causes the plate to move very slightly to the left, until the horizontal pull caused by the red cord not being vertical, just balances the difference in the ZPE thrusts on that plate. So, the right hand plate moves slightly to the left.

The same thing happens with the left hand plate. The ZPE thrust coming from the right is slightly reduced as it passes through the right hand plate, and the left hand plate moves slightly to the right until the angled pull of its supporting cord balances the net thrust on that plate. The overall effect is that the gap at point "**A**" in the diagram is very slightly larger than the gap at point "**B**", though the amount is not nearly as great as suggested by the diagram, which has been deliberately exaggerated to show the effect clearly. There is nothing complicated about this, it is just simple common sense. Remember that the pull of the supporting cord "**C**" is the exact equivalent of a vertical force "**D**" along with a horizontal force "**E**". Here, the vertical force **D** exactly matches the weight of the plate, and the horizontal force **E** exactly matches the unbalanced ZPE force (if they did not match exactly, then the plate would move until they did). The further away from the vertical that the plate moves, the greater the resulting horizontal force caused by the pull of the supporting cord.

Tesla expressed this in a very slightly different way in his **Dynamic Theory of Gravity** (1897) which states that all bodies emit microwaves whose voltage and frequency are determined by their electrical contents and relative motion. He measured the microwave radiation of the earth as being only a few centimetres in wavelength. He said that the frequency and voltage were influenced by the velocity and mass of the earth, and that its "gravitational" interaction with other bodies, such as the sun, was determined by the interaction of the microwaves between the two bodies.

If you find the concept of producing a driving force through pushing against the space-time continuum to be difficult to accept, then perhaps you should consider the US Patent granted to Boris Volfson on 1st November 2005. The important thing about this patent (which is crammed full of long words) is not whether or not it presents a realistic mechanism for a practical space drive, but the fact that the US Patent Office in the year 2005, granted the patent after what presumably was careful consideration. With that in view, it is hardly possible to consider Tesla to have been totally confused when he designed (and built) his "electric flying machine" which operated by pulling on the space-time field.

Tesla used high voltage at gigahertz frequencies for his electropulsion system. The propulsion of a vehicle powered by a Tesla drive is by the use of an additional AC generator at the back (which stiffens the space-time continuum behind the vehicle) and a DC 'brush' generator at the front (which weakens the space-time continuum in front, causing the vehicle to be pulled forwards).

Tesla was very astute. He deduced that 'empty space' actually contained:

1. Independent carriers which permeate all space and all matter and from which all matter is made. These carry momentum, magnetism, electricity or electromagnetic force, and can be manipulated artificially or by nature.

2. 'Primary Solar Rays' (starlight) which travel at the speed of light, having frequencies far above X-rays, gamma and UV radiation.

3. 'Cosmic Rays', particles in space propelled by the Primary Solar Rays.

4. X-rays, Gamma rays and UV electromagnetic waves, all of which travel at the speed of light.

5. Ordinary visible and Infra-Red electromagnetic waves which travel at the speed of light.

6. Rapidly varying electrostatic force of enormous potential, emanating from the earth and other gravitational bodies in space.

When we grasp the actual nature of the universe, it becomes clear that we have a much larger range of opportunities for producing usable energy in large quantities and at minimal cost.

Additional information can be found in Boris Volfson's US Patent 6,960,975 of November 2005 "Space Vehicle Propelled by the Pressure of Inflationary Vacuum State" which is reproduced in the Appendix.

If you find the thought of generating a gravitational field, difficult to come to terms with, then consider the work of Henry Wallace who was an engineer at General Electric about 25 years ago, and who developed some incredible inventions relating to the underlying physics of the gravitational field. Few people have heard of him or his work. Wallace discovered that a force field, similar or related to the gravitational field, results from the interaction of relatively moving masses. He built machines which demonstrated that this field could be generated by spinning masses of elemental material having an odd number of nucleons -- i.e. a nucleus having a multiple half-integral value of h-bar, the quantum of angular momentum. Wallace used bismuth or copper material for his rotating bodies and "kinnemassic" field concentrators.

Aside from the immense benefits to humanity which could result from a better understanding of the physical nature of gravity, and other fundamental forces, Wallace's inventions could have enormous practical value in countering gravity or converting gravitational force fields into energy for doing useful work. So, why has no one heard of him? One might think that the discoverer of important knowledge such as this would be heralded as a great scientist and nominated for dynamite prizes. Could it be that his invention does not work? Anyone can get the patents. Study them -- Wallace -- General Electric -- detailed descriptions of operations -- measurements of effects -- drawings and models -- it is authentic. If you are handy with tools, then you can even build it yourself. It does work.

Henry was granted two patents in this field: US Patent 3,626,605 -- "Method and Apparatus for Generating a Secondary Gravitational Force Field", Dec 14, 1971 and US Patent 3,626,606 -- "Method and Apparatus for Generating a Dynamic Force Field", Dec 14, 1971. He was also granted US Patent 3,823,570 -- "Heat Pump" (based on technology similar to the above two inventions), July 16, 1973.

Dr Peter Lindemann gave a lecture at the TeslaTech conference which is very informative and which I would highly recommend. It is available on DVD from http://www.free-energy.ws/products.html under the title of "Tesla's Radiant Energy". He makes a number of important points, some of which are repeated here.

We tend to think of the battle for industrial electricity to have been between Thomas Edison's DC system and Tesla's AC system, with Tesla winning. Unfortunately, while true, that is not the full story as Tesla moved on from AC to more powerful systems, and Tesla lost out on those systems, leaving us today, with a workable, but much inferior system. We need to see the overall picture clearly. James Clerk Maxwell produced his famous equations, relating electricity and magnetism (which are actually two faces of a single entity called "electromagnetism"). Subsequently, H.A. Lorentz damaged those equations, throwing away the parts which showed that free-energy was available for use if we knew how to access it.

Out present day position has grown up where, while we realise that "gravity" is 10^{39} times less powerful than electromagnetism, we see "static electricity" as a weak and useless thing which needs to be avoided. The reality is quite different as Tesla points out and demonstrates. Tesla describes "static" and "radiant" energy as being a force which appears to have no ultimate limits at all. That is, it is capable of providing unlimited power. Our knowledge of this power is so inadequate that we believe that power can only flow in a circuit which is a closed-loop and the power flow has to be a stream of electrons. This is most definitely not the case.

The more powerful radiant energy flows like an electrically charged sound wave passing through incompressible air and which can be fed down a single wire without the slightest difficulty. In fact, you can even skip that one wire and use the earth instead, transmitting power with what appears to be no wires at all. The actual, final and most important contest was between closed-loop electricity and single-wire power transmission, and that contest was one which Tesla lost.

The nature of this radiant energy is so different to what we think of as conventional electricity that it is perfectly possible to light a filament light bulb held in one hand, while grasping a single wire in the other hand. When doing this, there is no sensation whatsoever and nothing at all is felt. Videos on this suibject are at https://www.youtube.com/watch?v=3uXL4_Yas2k and https://www.youtube.com/watch?v=6BnCUBKgnnc at this time. An excellent set of teaching videos are at http://journal.borderlandsciences.org/videos/.

In his lecture, Dr Lindemann remarks that his understanding of the subject has been helped considerably by the book (http://www.free-energy-info.tuks.nl/TeslaBook.pdf) "The Inventions, Researches and Writings of Nikola Tesla" and the book (http://www.datafilehost.com/download-c74378fc.html) "The Secrets of Cold War Technology - HAARP and Beyond" which describes some of the early work done by Tesla. He also praises the book "Tesla's Vocabulary for Dummies" which is a joke on his part as there is no such book in spite of his appearing to show some quotations from it in his DVD lecture.

One feature of radiant energy which becomes clear from Tesla's description of it, is that the most useful effects which can be gained from it, start at a DC pulse frequency of 1 MHz which is far higher than experimenters use today. He stresses that we do not actually know the exact nature of electricity and that all of our present day measuring instruments are based on electron theory and so just do not measure radiant energy. In a way, it is a bit like the difference between AM radio and FM radio. Both are perfectly valid and work well, but an AM radio will not receive an FM radio signal and an FM radio will not receive an AM radio signal. Unfortunately, radiant energy is much more powerful than conventional electricity and it is not dangerous like electricity is. It should be noted that Hermann Plauston's very detailed patent - US 1,540,998 (http://www.free-energy-info.tuks.nl/PatD8.pdf) is on methods of capturing and using this radiant energy, and he describes any multi-aerial systems which produce a net output of 100 kilowatts as being a "small" system. I don't know about you, but I would settle for a system which produced less than 10% of that fuel-less output.

The best information on radiant energy comes from the writing of Tesla and Dr Lindemann draws attention to one of Tesla's patents, US 685,957 (http://www.free-energy-info.tuks.nl/PatD37.pdf) which explain how this radiant energy can be captured and used. Tesla also used a motor design which is effective with this type of energy. The motor has two windings, the first being fed directly and the second one receiving a 90 degree delayed pulse through a capacitor.

One thing that Tesla points out is that there is an incompressible gaseous medium filling the universe and which is composed of particles which are much smaller than hydrogen atoms. Mendeleev who constructed the table of elements indicates quite clearly that there should be two gaseous elements which are lighter than hydrogen, but he did not put them in his table because he did not know what they are.

The Nature of Gravity by Maurice Cotterell.

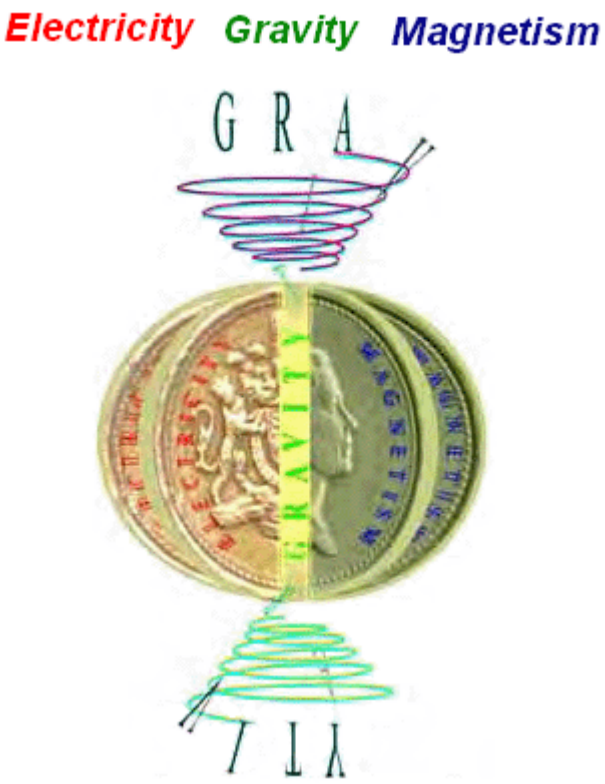
The comments by Joseph H. Cater later in this chapter include the statement that gravity is electromagnetic radiation of a frequency just below that of infra red. The extract from Karl Schappeller's work, also later in this chapter, states exactly the same thing and the same frequency. The extracts from Joseph Newman's

book, also later in this chapter, point out the effect of electron spin on attraction and repulsion. The document http://www.free-energy-info.tuks.nl/Cotterell.pdf is a paper on gravity by the very able Maurice Cotterell, part of which is quoted here:

How Gravity Works

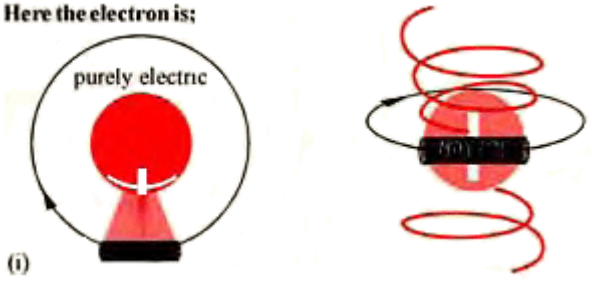
"How Gravity Works" explains how Electricity and Magnetism work together to produce the force of Gravity and in doing so, they reveal the underlying science behind Newton's equation which he tried in vain to uncover, namely, the reason why Gravity is proportional to the masses of two attracting objects, why it grows weaker in proportion to the square of the distance between them, and the nature of the "Gravitational Constant" which continues to perplex researchers to this day. It explains why all objects accelerate to Earth at 32 feet per second every second. It explains why the atom is comprised of eight orbital shells and why those shells fill with electrons in the way that they do. It explains how permanent magnetism works at atomic level, the nature of so-called 'Dark Matter', the reason why spiral galaxies are spiral, why the centre of the Earth is boiling hot and how the Earth's magnetic field is generated.

Scientists cannot understand why the positive protons on the centre of atoms simply do not spring apart – their positive charges repelling each other. In 1935, a researcher named Yukawa suggested that another substance or particle must exist to 'glue' the protons in the nucleus together - a kind of atomic 'velcro'. But an objector to Yukawa pointed out that if indeed that were the case, then there would have to be two kinds of velcro - a velvet side and a hook side (so to speak). But another researcher pointed out that protons covered in the velvet-side would not stick to other protons covered in the velvet-side, and that protons covered in the hook-side would not stick to other protons covered with the hook-side. So, another researcher pointed out that for such a scheme to work, there would have to be three types of each of the three types. Thus, the crazy field of Particle Physics was born. In the decades that followed, Particle Physicists were granted billions of dollars to find out what makes the atom stable. Since then, they have been searching for more than 300 imaginary sub-atomic particles. The Concise (Oxford) Science Dictionary comments:
"… the whole elaborate theory is circumstantial… none of the smaller particles have ever been identified in experiments… the theory does not claim to have been verified".

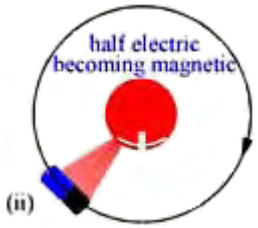


Cotterell shows that none of the particles actually exist, and more importantly, that none are necessary. To explain why the atom does not spring apart, he returns to firm ground and to the 1930s pioneering work of Ernest Rutherford and James Chadwick and shows that by simply changing the perceived shape of the electron and neutron, the gravitational mechanism is exposed, explaining away the mysteries of the atom and many of the mysteries of the cosmos.

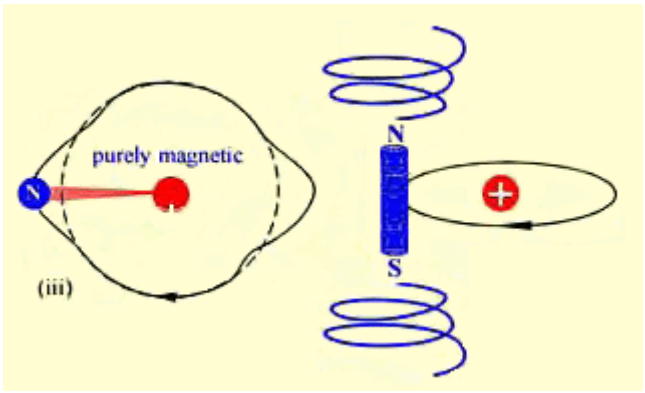
Summary Part 1:



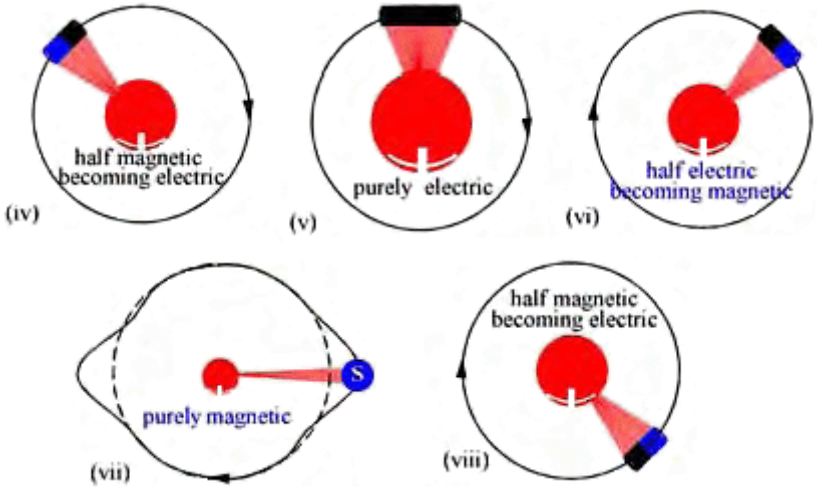
(**A**) A coil-shaped electron (shown in black) makes the hydrogen atom dynamic [come alive]. The amount of helically polarised electric field forced from the poles of the proton is maximum when the electron is horizontal, that is, after 180 or 360 degrees of orbital travel. Shown at the top left is the plan view of a hydrogen atom. The horizontal electron slices through the electric field (shaded pink). A magnetic field is induced into the coil which pushes against the electric field, causing it to topple on it's own axis (like the opposing blades of a windmill) as it orbits the positive proton.



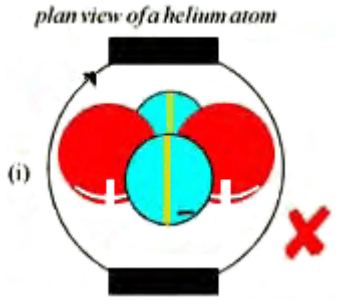
The orbiting coil-like electron sucks energy from the proton causing the proton to compress, release heat and cool rapidly.



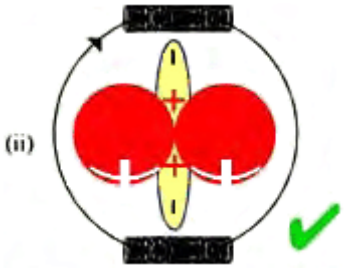
Once the spinning electron-magnet passes the vertical position, the super cold proton sucks in ambient heat and expands rapidly; compression and expansion of the proton results in cyclical changes of capacitance [electric field] proportional to the surface area between the particles.



(iii and vii) The **ends** of the electron-magnet radiate maximum amounts of helically polarised magnetic energy after 90 degrees and 270 degrees of orbital travel [when vertical]. Hence the magnetic and magnetic radiation is displaced by 90 degrees and the electromagnetic (EM) radiation from the northern sector is displaced in phase by 180 degrees from that of the southern sector.

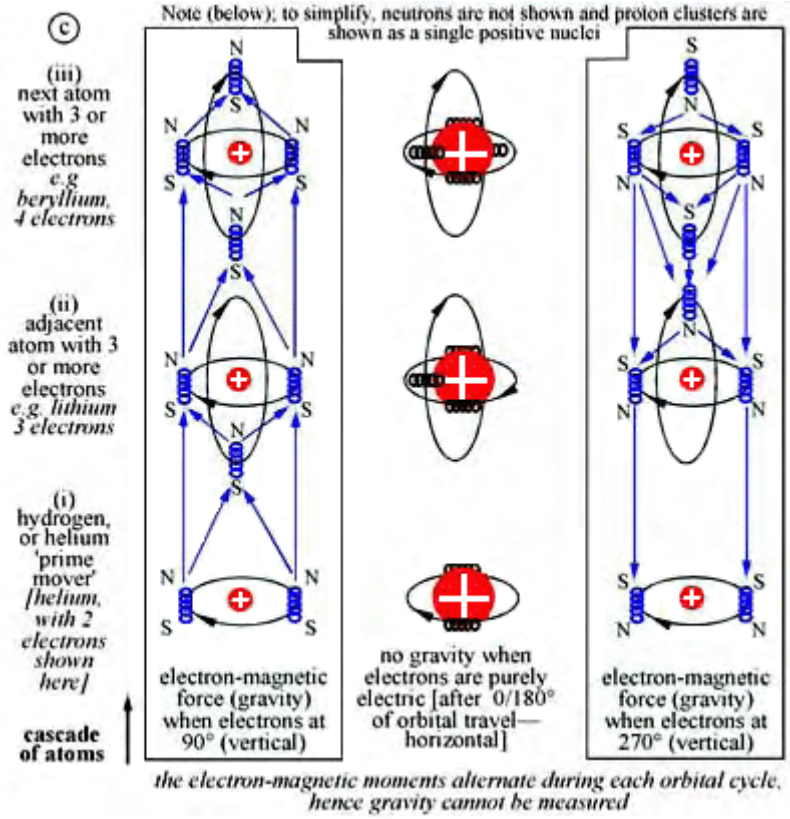


(**B**) (i) The orthodox view of a helium atom (containing two neutrons) precludes an understanding of gravity - each electron is passive [dead] and the spherically shaped neutrons render the atom unstable. Moreover, the electron is shown as a purely 'electric' particle when in fact it is electric for only half the time, and magnetic for half of the time. **The electron is an 'electromagnetic' particle**.



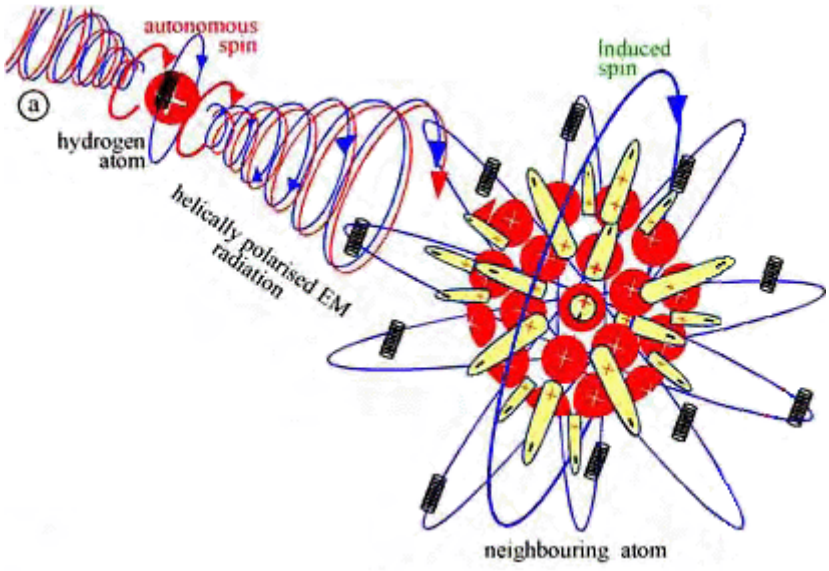
(ii) This new view of the helium atom accommodates a gravitational mechanism - the coil-shaped electrons bring the atom to life and the spike-shaped neutrons make the atom (the contents of the old neutron and the spiked neutron being identical). Supporting evidence: 31/3/08 researchers of the Hohn-Meitner Institute in Berlin announced that "neutrons behave like compass needles" – Science Daily.

(**C**) hydrogen spins axially and autonomously and so it radiates helically polarised ElectroMagnetic energy into space. [Supporting evidence: 'hydrogen radiates helically polarised EM energy' www.sciencenews.org].

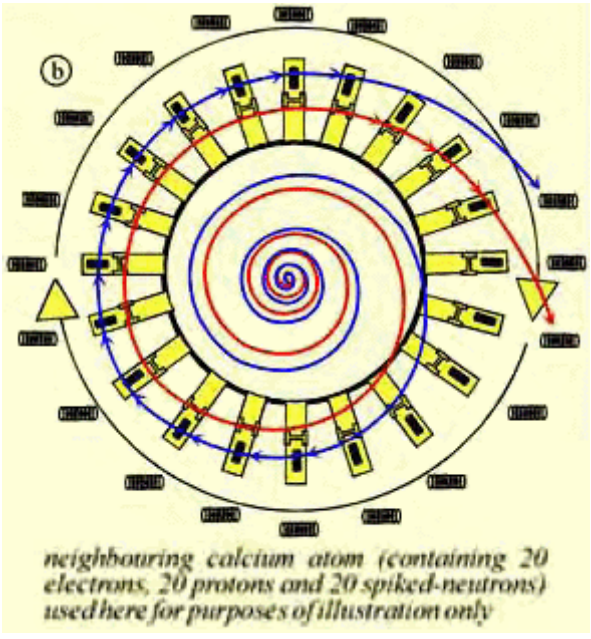


Helium [the only electrically geometrically symmetrical atom to contain spiked neutrons] also spins axially and autonomously and hence, likewise radiates helically polarised electromagnetic energy. In free space, hydrogen and helium are prime movers in the gravitational mechanism. Other atoms cannot **autonomously** radiate helically polarised electromagnetic energy because they do not spin axially or autonomously. However, when helically polarised electromagnetic radiation from hydrogen or helium bombards neighbouring atoms, it causes the nucleus of those atoms to spin in the same direction and, at the same time, causes the orbiting electrons in a cascade of atoms to synchronize their spin. As a result, the magnetic moments of orbiting electron-magnets in a cascade of atoms attract each other. This is the force of Gravity.

Summary Part 2:

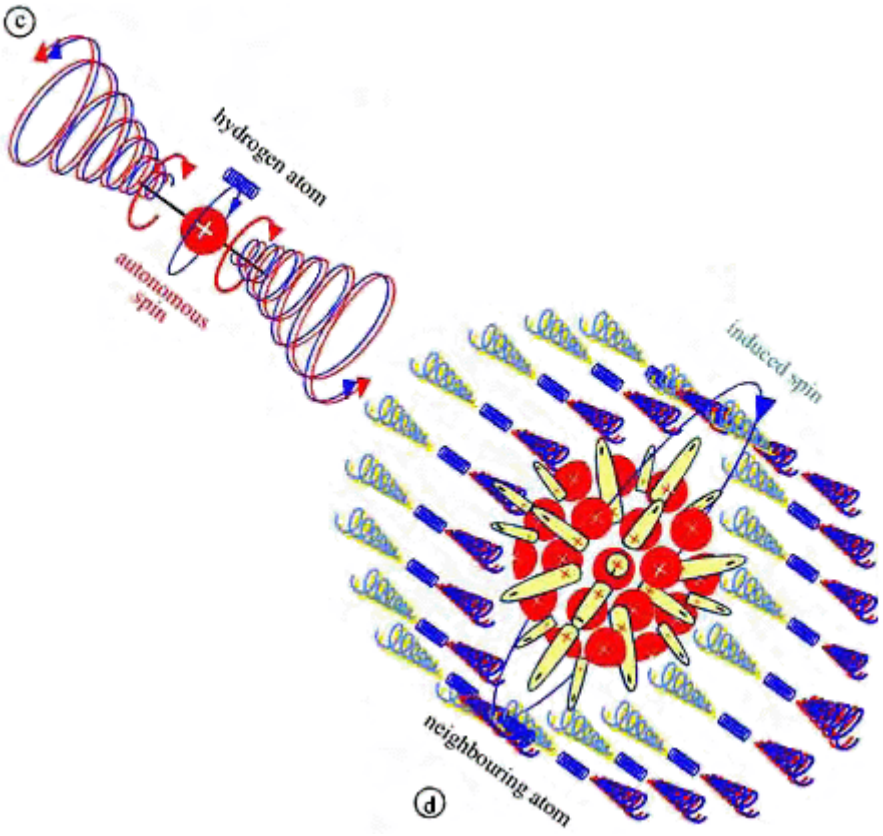


(**2A** and **2B**) The 'Motor Effect'; helically polarised electromagnetic radiation from the hydrogen atom acts upon the negative charges of the neighbouring atom, causing the nucleus and the electron cage to spin axially and at the same time, synchronizes the spin of the electrons in both atoms [Supporting evidence: Nature 2009;458 (7238):610 DOI:10.1038.nature07871 (persistent spin helix)]. The mass of the electron is twice that of the neutron-negative charge, hence the torque on the electron-cage is twice that on the neutron-negative charges in the nucleus. Thus, the electron-cage spins faster than the nucleus.



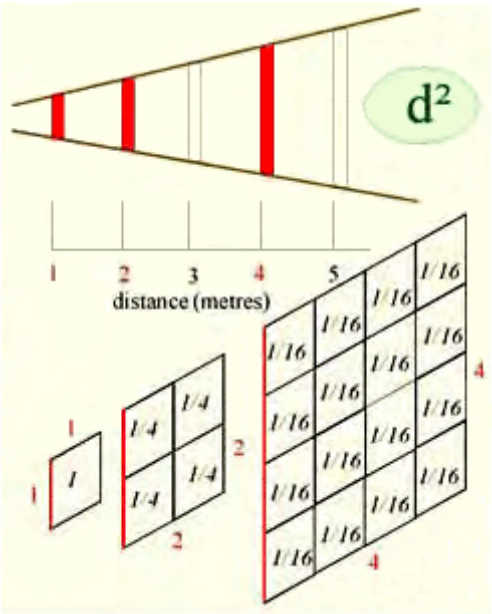
(**2B**) The 'Generator Effect; orbiting electrons in the neighbouring atom now 'generate' helically polarised electromagnetic radiation. The power of the atomic 'generator' is proportional to the differential speed between the nucleus and the electron cage and to the number of orbiting electrons (atomic mass).

Electromagnetic output is a function of the number of electrons [atomic mass] ('g') AND of the difference in rotational speed between the nucleus and the electron cage ('f').

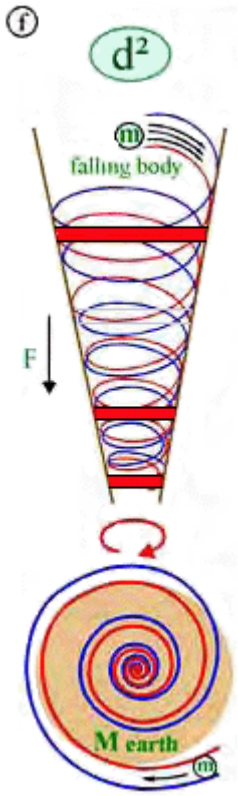


(**2C**) Gravity radiation from the hydrogen atom bombards neighbouring atoms sucking them towards the source of the radiation. *[Here, to facilitate explanation, the magnetic wave (shown in blue) is shown radiating from the nucleus whereas in actuality, it radiates from the orbiting electron]*.

(**2D**) The corkscrew-style electromagnetic radiation from the neighbouring atom now, likewise, sucks nearby atoms towards itself. Both atoms suck in the same direction, thus gravity from the Sun and the Moon pull in the same direction. *[To facilitate explanation, the electric waves (shown in red) are shown radiating from orbiting electrons]*.



(**2E**) As gravity radiation propagates and moves away from the atom, it decreases in field strength [per metre squared]. For every unit of distance travelled, the radiated energy diverges geometrically. This causes the radiated energy to reduce by the square of the distance travelled ("d^2"). The numbers inside the square in the diagram above show the field strength of the radiating electromagnetic energy in volts per square metre. *[Squares are used here just for illustration purposes as the diverging wave is actually conically-helical as seen in Figure 2F]*.



(**2F**) The frequency of the helically polarised electromagnetic radiation from the Earth remains constant. However, as **m** approaches **M**, the spiralling electromagnetic radiation accelerates the differential rotation between each atomic nucleus and electron-cage of which **m** is made. As a result, the 'relative' atomic frequency increases and hence the output of the 'atomic generator' increases, uniformly increasing the attracting electromagnetic **F**orce between **m** and **M** in accordance with a square-law scale as **m** proceeds along the electromagnetic spiral. **M** thus attracts **m** with square-law [$\mathbf{d^2}$] uniform acceleration. In the diagram on the right, **m** is shown spiralling towards **M** with the electromagnetic wave stationary, but in reality, m approaches **M** in a straight line as the wave spirals across **m**.

(**2G**) Hence Isaac Newton's equation for the force of gravity; where the force is proportional to the masses (**m** and **M**) of the two attracting bodies and the strength of the force decreases inversely with the square of the distance between them ($\mathbf{d^2}$).

$$\mathbf{F = GmM/d^2}$$

G is Newton's gravitational constant **6.67428×10^{-11}** $N\ m^2\ kg^{-2}$ - the instantaneous alternating magnetic force between any two electron-magnets in neighbouring atoms.

Maurice Cotterell gives a more advanced and in-depth explanation of this process in his latest book "FutureScience" in which the introduction says:

The 12 great mysteries of Physics

You can be forgiven for believing that Isaac Newton discovered the force which makes apples fall to the ground, but he did not. He simply gave the force (F) a name ("Gravity") and noted that the strength of the force, whatever it might be, depends on the size of the Earth (Mass M2). He also figured out that the strength of force on the apple depends on how far away the apple is from the Earth to begin with and he noted, like Galileo before him, that as an apple falls it speeds up and travels faster and faster until it hits the ground.

At first, Newton set down the relationship between the force and the masses in a shorthand way (a formula) saying that F (the strength of the force) could be found by multiplying m1 by M2 and then by dividing the result by the square of the distance between them (d^2). There was just one problem: every time he did the calculation he finished up with the wrong answer. He eventually discovered that the only way to get the right answer was to multiply the wrong answer by '667 billionths'. He did not know where this tiny amount of force came from, but just like he had done with Gravity, he gave the mysterious number a name – "the Gravitational Constant" **G**. This made his formula F = G x m1 x M2 / d^2. It is hard to believe that since then, that is, for more than 350 years, still no-one understands what gravity is or what causes it or why we have to multiply the wrong answer by G to get the right one. To this day, no-one understands what G is or even where it comes from. That's how sophisticated we are. No-one understands why objects fall to the ground.

The second most embarrassing problem for modern Science is that no-one understands why positive charges in the middle of atoms simply do not just spring apart which is what you would expect, given that positive repels positive. To overcome the contradiction, physicists decided that the positive charges must be glued-together in some way and so, for the past 75 years they have been searching for the 'glue' or more precisely, more than 300 types of glue – but as yet have not found any.

Also, no-one understands (given that positive attracts negative), why the orbiting negative electrons of atoms do not simply get sucked into the positive bits in the middle of the atom. They just 'cannot understand it'.

When you place a magnet on the door of a refrigerator, you probably imagine that scientists understand how a magnet actually works, but they do not. As far as they are concerned, it seems to work by *magic* so they gave it the name *magnet*.

Switch on the light and you no doubt take it for granted that the people who supply the electricity understand how it works, but they do not. Physicists are pretty sure that an electrical current is simply the movement of atomic negative charges along a wire. The problem is that no-one understands why, whenever a current flows, a magnetic field appears around the wire.

When they look skywards at night, astro-physicists cannot understand what makes the stars cluster together into galaxies, because their calculations contradict what their eyes can see. To overcome this, they say that there must be other sources of gravity out there, like planets and dust which cannot be seen because it is dark. So they named their fantasy planets 'dark matter', at least for the time being, until they can figure out why their calculations do not make sense. Neither can they understand why stars spiral into galaxies, but they cling to the hope that when (or if) they understand how gravity actually works, that the reason for the double-spiral of galaxy formation will become clear.

"FutureScience" explains in a simple way, how Electricity and Magnetism work at atomic level, and how they work together to produce the force of Gravity and in so doing, reveal the underlying science behind Newton's equation – the underlying science which he tried in vain to uncover – the reason why Gravity is proportional to the masses of the two attracting objects, why it grows weaker in proportion to the square of the distance between them, and the nature of the mysterious 'Gravitational Constant' which continues to perplex researchers to this day. It explains why all objects accelerate to Earth at 32 feet per second every second. It explains why the atom is comprised of 8 orbital shells and why the shells fill with electrons the way that they do. It explains the nature of so-called 'dark matter', the reason why spiral galaxies are spiral, why the centre of the Earth is boiling hot, how the Earth's magnetic field is generated, how to generate gravity waves, how gravity waves can be used to produce unlimited quantities of free energy, how to generate antigravity waves, how antigravity waves can be used to de-couple hydrogen bonds in water, separating hydrogen from oxygen and providing unlimited supplies of free hydrogen to fuel hydrogen-powered cars, and how antigravity radiation can be used to cause the disintegration of matter in water and living tissue (phasor weapons technology).

The book goes on to explain the cause of the Sunspot cycle and how the Sun affects life on Earth: the cause of global warming and global cooling, how the 289-day spinning Sun regulates fertility in females, how the Sun's radiation causes 12 types of genetic mutations which result in 12 types of personality (sun-sign astrology), how sunspots cause schizophrenia, how the Sun controls bio-rhythms and our day-to-day behaviour, how the Sun causes catastrophe cycles and the rise and fall of civilizations, and how mobile phones and power lines cause cancer.

How the discoveries were made

In 1935, in an attempt to explain why protons in the centre of atoms do not simply spring apart, a researcher named Yukawa suggested that another substance or particle must exist to 'glue' the protons in the nucleus together - a kind of atomic 'velcro'. But an objector to Yukawa pointed out that if indeed that were the case, then there would have to be two kinds of velcro - a velvet side and a hook side (so to speak). But another researcher pointed out that protons covered in the velvet-side would not stick to other protons covered in the velvet-side, and that protons covered in the hook-side would not stick to other protons covered with the hook-side. So, another researcher pointed out that for such a scheme to work, there would have to be three types of each of the three types. Thus, the crazy field of Particle Physics was born. In the decades that followed, Particle Physicists were granted billions of dollars to find out what makes the atom stable. Since then, they have been searching for more than 300 imaginary sub-atomic particles. The Concise (Oxford) Science Dictionary comments:
"… the whole elaborate theory is circumstantial… none of the smaller particles have ever been identified in experiments… the theory does not claim to have been verified".

FutureScience shows that none of the particles actually exist and, more importantly, that none are necessary. To explain why the atom does not spring apart, the Author returns to firm ground and to the 1930s pioneering work of Ernest Rutherford and James Chadwick and shows that by simply changing the perceived shape of the electron and the neutron, the gravitational mechanism is exposed, explaining away the mysteries of the atom and many of the mysteries of the cosmos. The book is aimed at the general reader with no prior knowledge of Science and it uses many illustrations to convey the concepts simply and concisely.

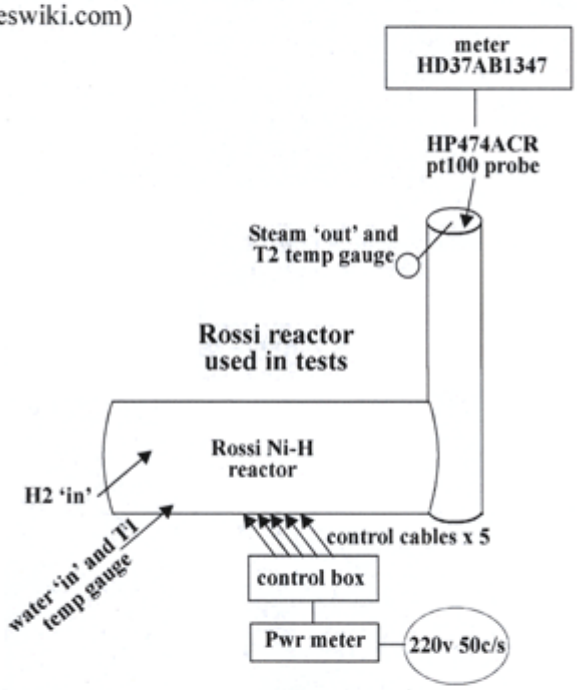
Maurice Cotterell's web site is http://www.mauricecotterell.com/

Maurice also comments on the recent Cold fusion generator development of Rossi as follows:

E-Cat generator not 'cold fusion'

The Energy Catalyzer (also called E-Cat) is a purported cold fusion or Low-Energy Nuclear Reaction (LENR) heat source built by inventor Andrea Rossi with support from physicist Sergio Focardi. An Italian Patent, which received a formal but not a technical examination, describes the apparatus as a'process and equipment to obtain exothermal reactions, in particular from nickel and hydrogen'. Rossi and Focardi say the device works by infusing heated hydrogen into nickel, transmuting it into copper and producing heat. An international Patent Application has received an unfavorable International preliminary report on patentability because it seemed to 'offend against the generally accepted laws of physics and *established* theories' and to overcome this problem the Application should have contained either experimental evidence or a firm theoretical basis in *current* scientific theory. (http://www.peswiki.com)

FREE ENERGY?
YES
but it's not
'Cold Fusion'

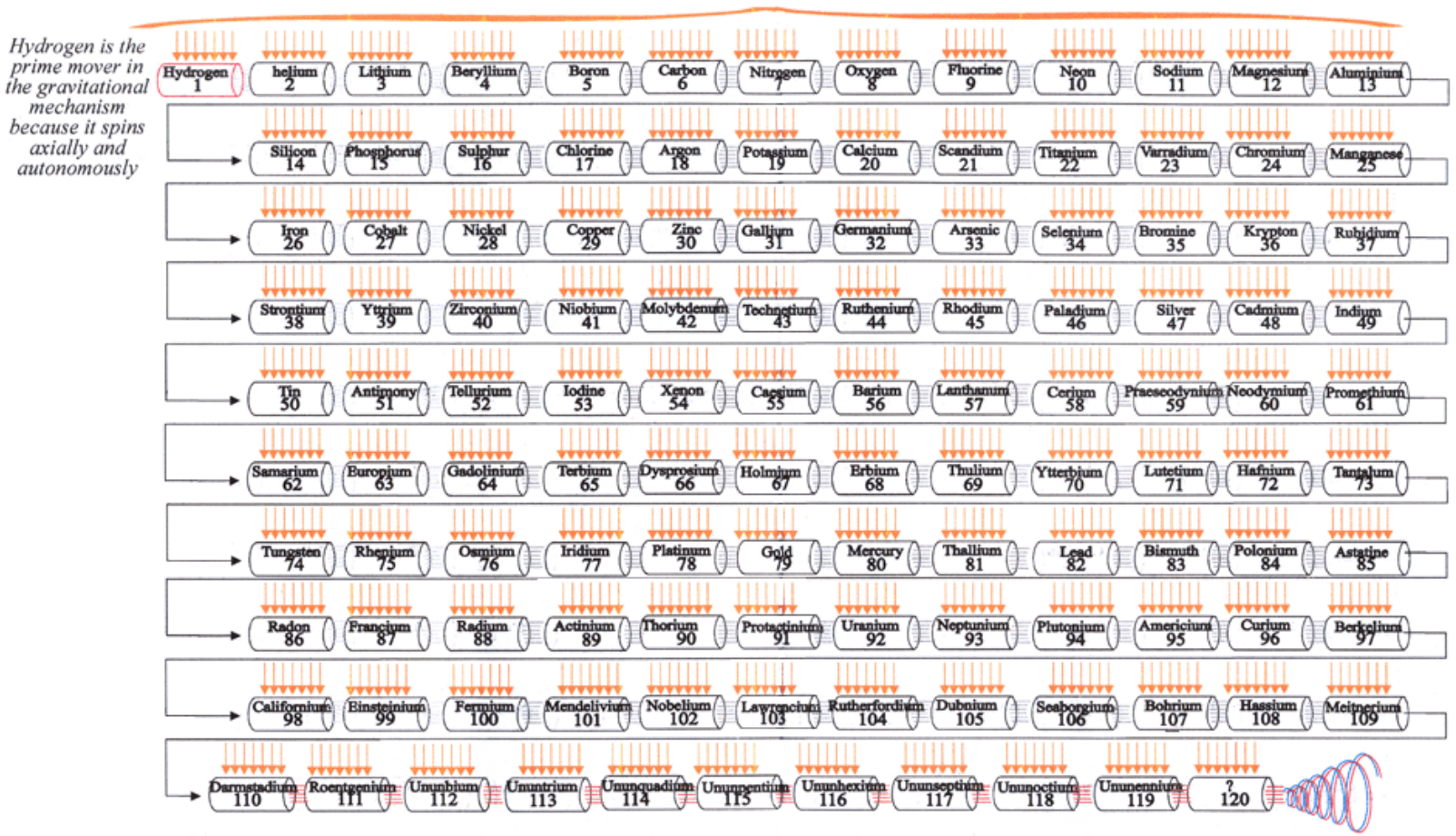


How they think it Works: The Energy Catalyzer works—they believe—by placing nano to micro sized particles of nickel powder in a reactor along with pressurized hydrogen gas and special undisclosed (for proprietary reasons) catalysts. When the contents of the reactor is heated (using a series of resistors in the reactor) to approximately 400 to 500° C., *nuclear reactions* start taking place. The strength of the reaction can be changed by varying the pressure of the hydrogen in the reactor. The output energy can be up to 400 times the input energy. No precious metals or radioactive substances are placed in the reactor. After the reactor is turned-off the reactor can be opened and no radiation can be detected. The process appears to transmute nickel into copper and trace amounts of other elements such as zinc. It will not work with deuterium (heavy hydrogen)—but they don't know why.

if it's not 'Cold Fusion', how does it really work?'

...it works by 'Periodic Scaling' ⟶

How to generate gravity waves, and free energy, using 'Periodic Scaling'



Fill each of 120 containers with one each of the elements and align the containers in a straight line in rising atomic number order. How does it work? Gravity waves from the hydrogen atoms in the hydrogen container will bombard the container next in line containing helium. The helically polarized electromagnetic gravity waves from the hydrogen container will act upon the spiked-neutrons in the helium atoms causing the helium atoms to spin axially in the same direction. The helium atoms will now generate and radiate helically polarized gravity waves, sucking-in ambient heat to fuel the process. In this way ambient heat is converted to gravity waves and more heat. Helically polarized gravity waves from the helium atoms will act upon the spiked-neutrons in the lithium atoms causing those atoms to spin axially in the same direction. The lithium atoms will now generate and radiate helically polarized gravity waves, again, sucking-in ambient heat to fuel the process. This sets up a chain reaction along the line of containers; hydrogen, the prime-mover, activates every atom in the chain. Containers along the chain will become increasingly hotter and massive amounts of gravitational energy and heat will radiate from the end of the chain following atom 119, ununennium. [from *FUTURESCIENCE—forbidden science of the 21st-century* by Maurice Cotterell]

Note; the line of containers will be influenced by gravity waves from the Earth and so should be stacked vertically, silo-like, with hydrogen uppermost. Water may be used in place of hydrogen, and dangerous elements may be omitted, at a cost to efficiency.

The E-Cat generator uses 'Periodic Scaling' to generate heat in the same way; the hydrogen fuel, used in the process, sucks-in ambient heat and converts the heat to gravity waves—that's why liquid hydrogen is the most effective super-coolant. The corkscrew-style electromagnetic gravity waves from the hydrogen atoms then excite the nickel atoms causing the nickel to get hotter. The hydrogen fuel then sucks-in more heat *from the nickel* and radiates more gravity waves that again bombard the nickel. Some of the fuel is consumed in the process. This cumulative excitation continues until the availability of heat in the ambient air, and nickel, is constrained by energy transfer considerations.

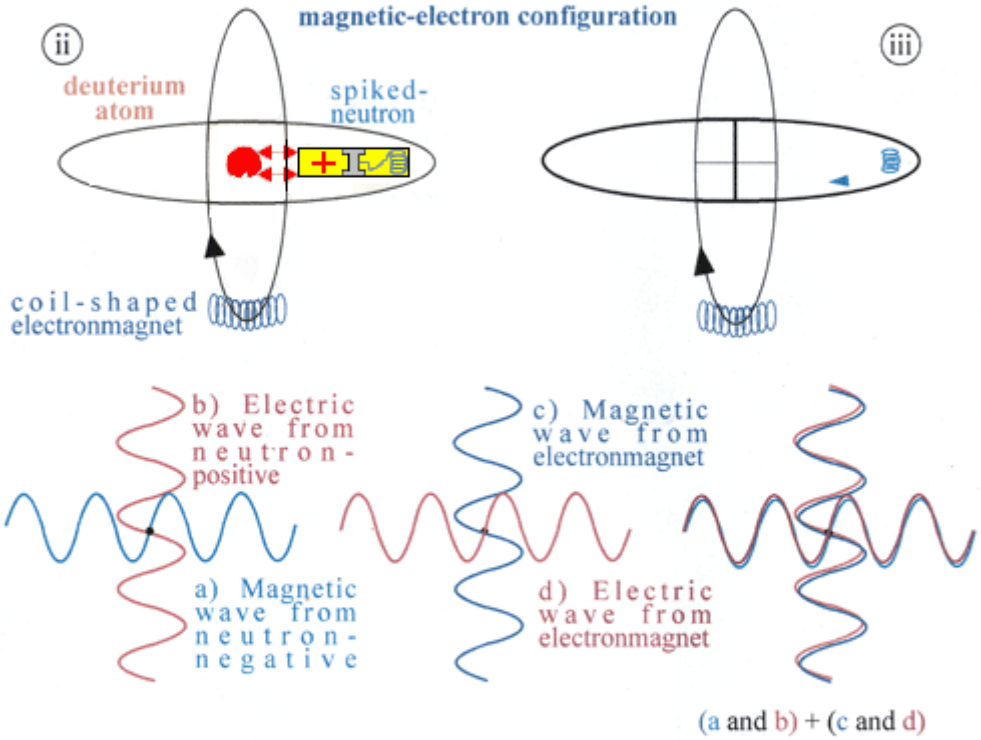


Figure 7 (in *How Gravity Works* [*see* www.MauriceCotterell.com]) showed how the first two atomic orbital shells are offset by 90°. In this case, the neutron-negative occupies the first *virtual* orbital 'radius'. The electronmagnet orbit must therefore be offset by 90°, as shown. Because of this, the electric wave from the neutron-positive clashes with the magnetic wave from the orbiting electronmagnet, and vice-versa, and the two electromagnetic waves cancel. Hence, deuterium cannot radiate gravity waves.

here's the 'smoking-gun'; the E-Cat generator will not work with deuterium [heavy hydrogen]—because deuterium cannot radiate gravity waves

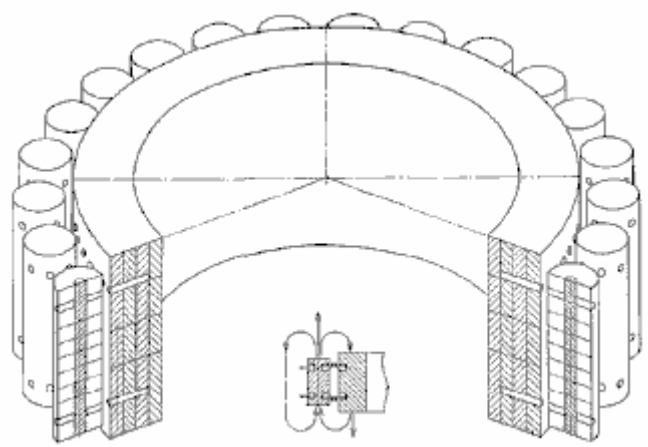
proof, if proof were needed, that the process must work by 'Periodic Scaling'

www.MauriceCotterell.com

© M.Cotterell 2013

John Searle's Permanent Magnet Drive/Generator.

Professor John R.R. Searle of Britain developed an electrical generation system based on two rings of magnets being spun relative to one another. The magnet orientations oppose each other to produce a magnetic splatter field.



The outer magnets in the diagram above are referred to as “rollers”. When three rings of rollers are placed one inside the other, then the outer ring rotates of its own accord, without any external power being applied. If pick-up coils are placed around the outside, then electrical current is generated with a COP of infinity. The method of imprinting the necessary magnetic pattern on both the rollers and the stators is a difficult and expensive process.

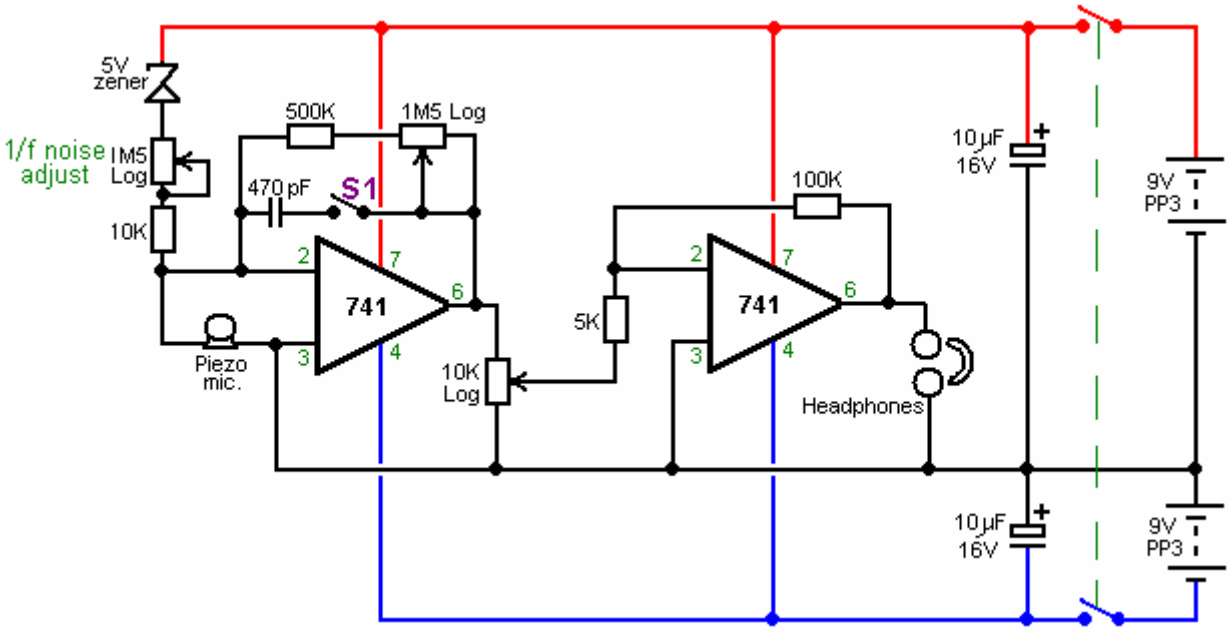
Dr. Terry Moore has recently built a replication model of this Searle technology and his model video is available at http://www.youtube.com/watch?v=bb3N1epMG7A. The Searle device also demonstrates a gravitic effect and John has built what would loosely be described as a “flying disc” using this technology. If high voltage is applied to the device when it is rotating, then a surrounding corona develops and strong upward electrogravitic forces are generated.

Dave Lawton’s Gravity-Wave Detector.

It has been reported that Nikola Tesla made a device which allowed him to hear sounds at great distances. I have never seen any details of the circuitry used by Tesla. However, Dave Lawton has produced such a device, and he reports that he could hear conversations taking place four and a half miles away from him. Interestingly, the sounds from that distance were also travelling through a solid stone wall some three feet thick. The circuit for this device is described in this document.

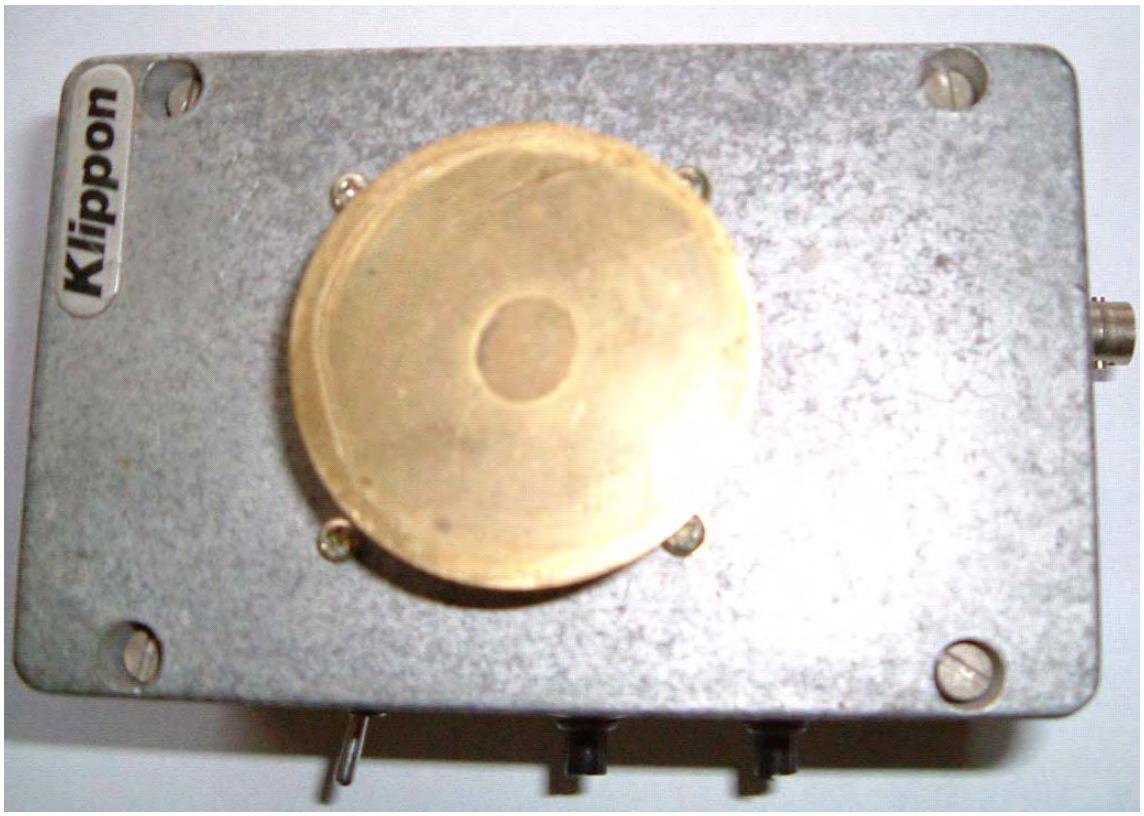
In my opinion, the device is not picking up audio signals in the manner of a conventional microphone where air pressure waves vibrate a transducer, creating an electrical signal which is then amplified. The interesting thing is that it is distinctly possible that some other mechanism is coming into play here. This opinion is supported by the fact that Dave’s circuit is an upgraded version of a monopole gravity-wave detector. Dave used this device to record the “sound” of the Shumaker-Levy comet colliding with Jupiter.

The circuit shown here is quite conventional electronically speaking, comprising of two 741 operational amplifiers connected as a two-stage amplifier. The unusual feature is where a small amount of white noise is being fed into the microphone input:



The white noise is generated by the 5-volt zener diode. The level of this white noise component is controlled by the 1.5 megohm variable resistor plus the 10K fixed limiting resistor. While the range of these two components is 10K to 1.501 Meg. the working setting is normally very high and so only a very small amount of white noise is fed into the input of the first 741 op. amp. to modify the microphone input.

The adjustment of this injection of white noise is the main control of this most unusual circuit, and it has been found that when the setting is just right, the circuit has the feel of a public address system just about to go unstable from positive feedback. The unit build looks like this:

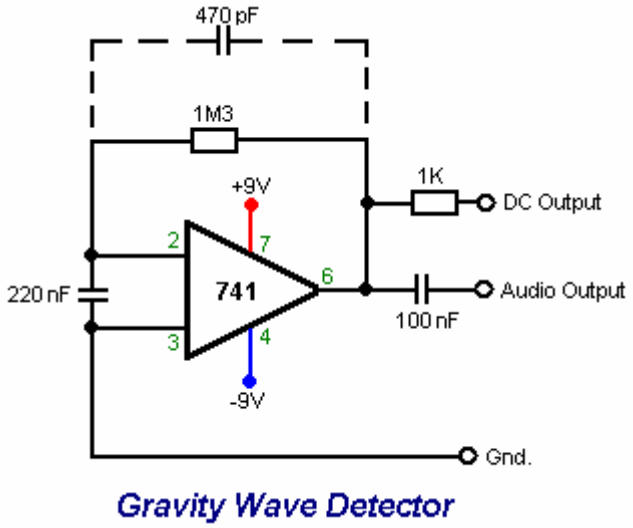


The theory of operation was put forward by Gregory Hodowanec in the April 1986 issue of the Radio-Electronics Magazine, where he puts forward the theory that the source of noise in electronic devices is caused by gravitational waves and he suggests that there are monopole gravity waves. This does not oppose the gravity waves predicted by Einstein. Gregory views these monopole gravity waves as being much stronger than those suggested by Einstein, and consequently, much easier to detect.

He also suggests that monopole gravity waves have been seen for many years and have been described as “low frequency noise” signals or “flicker noise”. These signals have also been called Microwave Background Radiation, supposedly caused by the “Big-Bang” though this cause is disputed by some.

Gregory views our universe as a finite, spherical, closed system, i.e. a black body. Monopole gravity waves propagate in Planck time so their effects appear everywhere almost simultaneously. Gravity wave energy can be imparted to ordinary objects. So it is suggested that the fact that a fully discharged electrolytic capacitor can develop a charge when disconnected from all circuitry, is down to the interaction of the capacitor with monopole gravity waves.

Gregory suggests the following circuit for examining monopole gravity waves:



Gravity Wave Detector

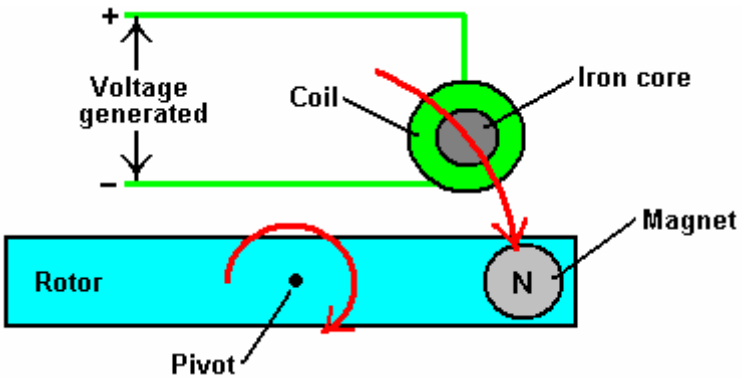
Details of this and the theory can be found at www.rexresearch.com/hodorhys/remag86/remag86.htm Dave has taken that circuit and extended it substantially to give added gain plus a controlled feed of white noise, without relying on the characteristics of a capacitor, capacitors being notoriously variable in precise characteristics.

The unit is operated by turning the gain up until the circuit just reaches self-oscillation, and then backing the gain off very slightly. The white noise source is then adjusted until the unit is producing a somewhat echoing quality to the sound. The result is a device which has unusual characteristics. The circuitry is so simple and cheap, that you can easily try it out for yourself.

Butch Lafonte’s Motor/Generator.

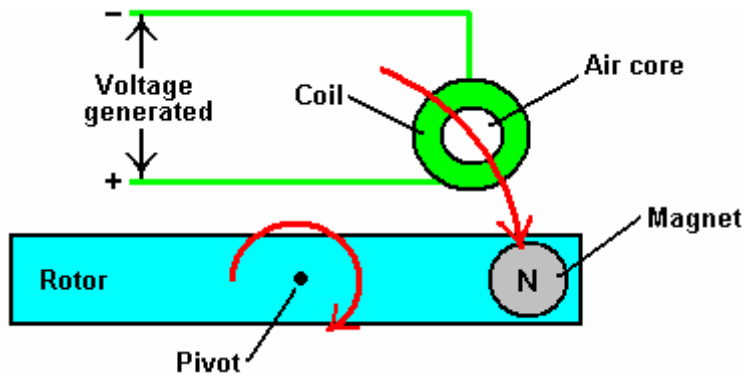
Butch has designed an intriguing Motor / Generator system based on the balancing of magnetic and electrical forces. This clever design operates according to the following statements made by Butch:

1. If a magnet is moved away from an iron-cored coil, it generates a voltage:



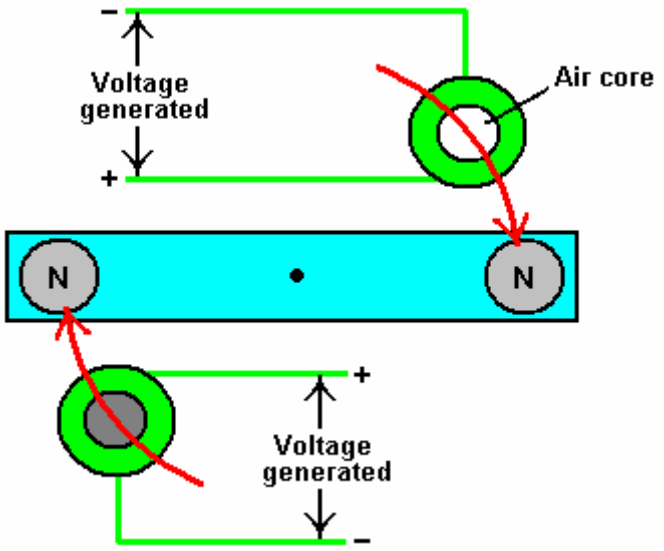
The voltage generated for any given magnet and speed of movement, is directly proportional to the number of turns of wire which make up the coil.

2. If a magnet is moved away from an air-cored coil, it also generates a voltage. However, the big difference is that the voltage is of the opposite polarity. In other words, the plus and minus connections are swapped over:

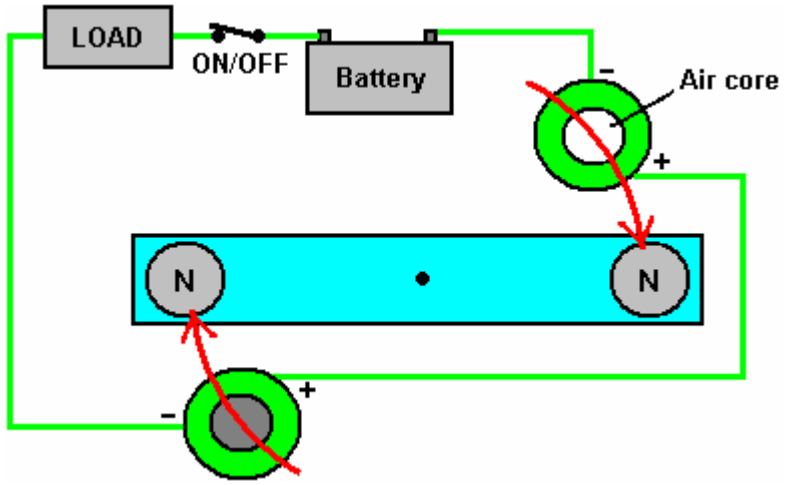


Again, the voltage generated for any given magnet and speed of movement, is directly proportional to the number of turns of wire which make up the coil.

So, if these two arrangements are joined together, they produce a system where the voltages cancel each other exactly, provided that the number of turns in each coil are adjusted to produce exactly the same voltages. The mechanical attraction and repulsion forces also balance, so the circuit can be arranged to have no net effect when the rotor is rotated:



It follows then, that this motor arrangement could be introduced into an existing circuit without affecting the operation of that circuit. The arrangement would look like this:

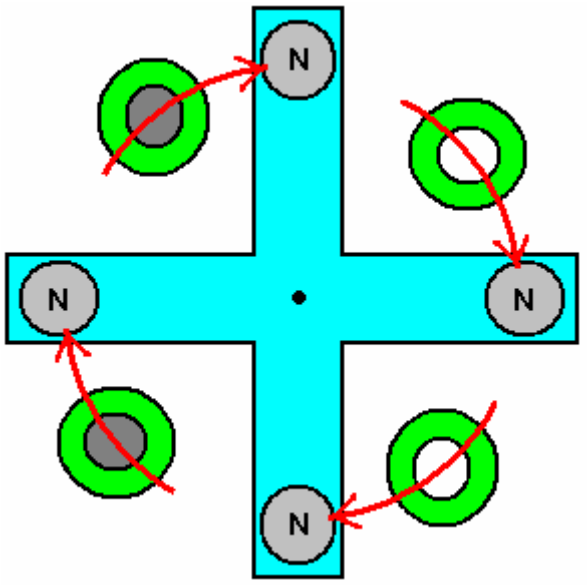


Here, there is no net electrical or magnetic drag on the rotor as the magnets move away from the coils. The battery supplies current to the load in the normal way and rotor arrangement has no effect on the operation of the circuit.

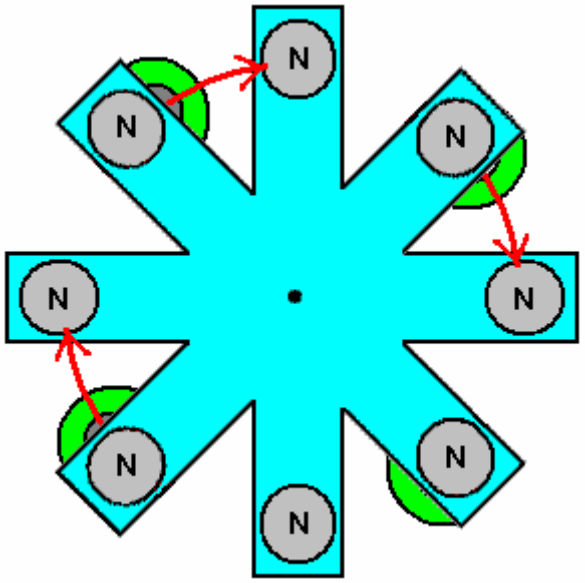
However, when the rotor reaches 100^{O} or so, past the coils, the On/Off switch can be opened. This leaves the rotor in an unbalanced condition, with there being an attraction between one magnet and the iron core of one coil. There is no matching repulsion between the other magnet and the air core of the other coil. This produces a rotational force on the rotor shaft, keeping it spinning and providing useful mechanical power which can be used to generate additional power. This extra mechanical power is effectively free, as the original circuit is not affected by the inclusion of the rotor system.

From a practical point of view, to give high rotational speed and long reliable life, the On/Off switch would need to be an FET transistor with electronic timing related to the rotor position.

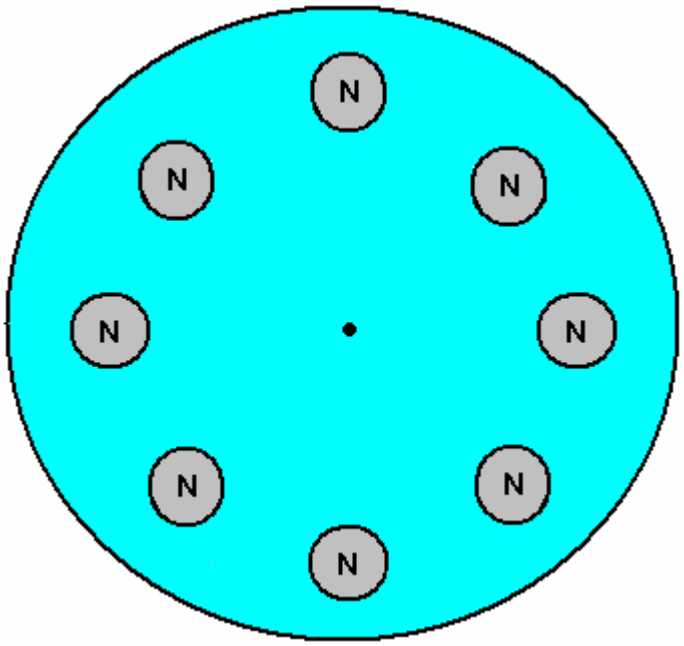
There is no need for the rotor to have only two magnets. It would be more efficient if it had four:



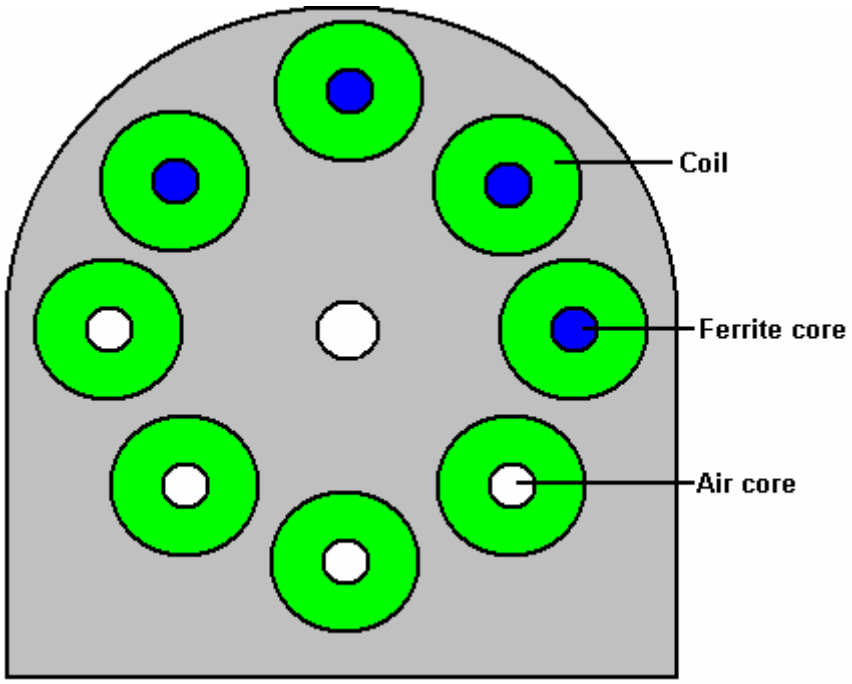
Or better still, eight:



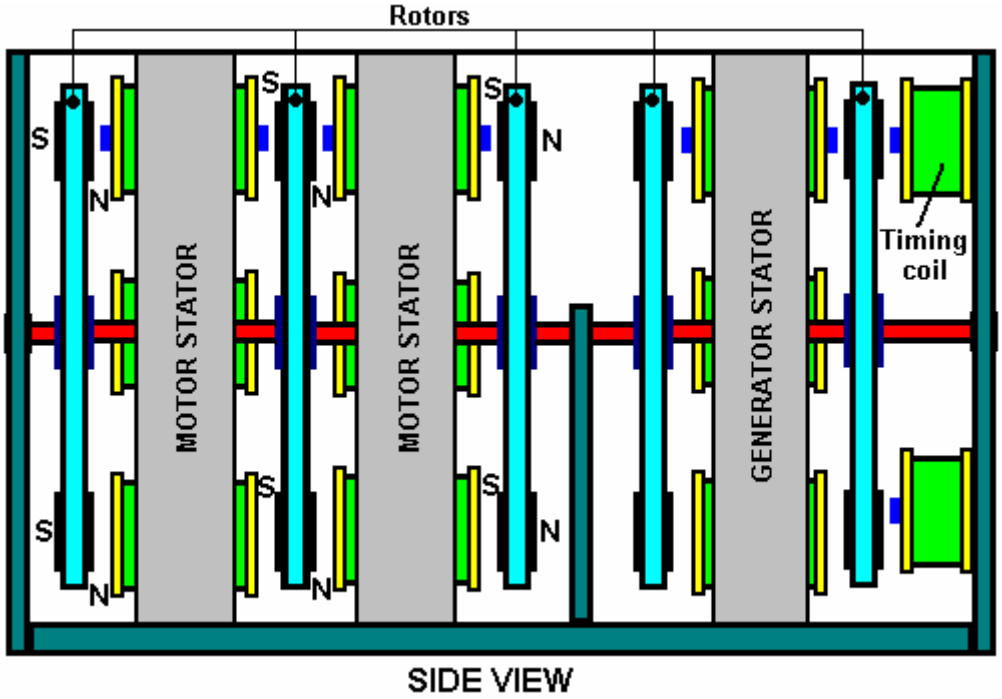
And if you are going to have eight, there is no need to have the V-shaped cut-outs which just create turbulence when spinning, so make the rotor circular:



And the stator supporting the coils matches the rotor:

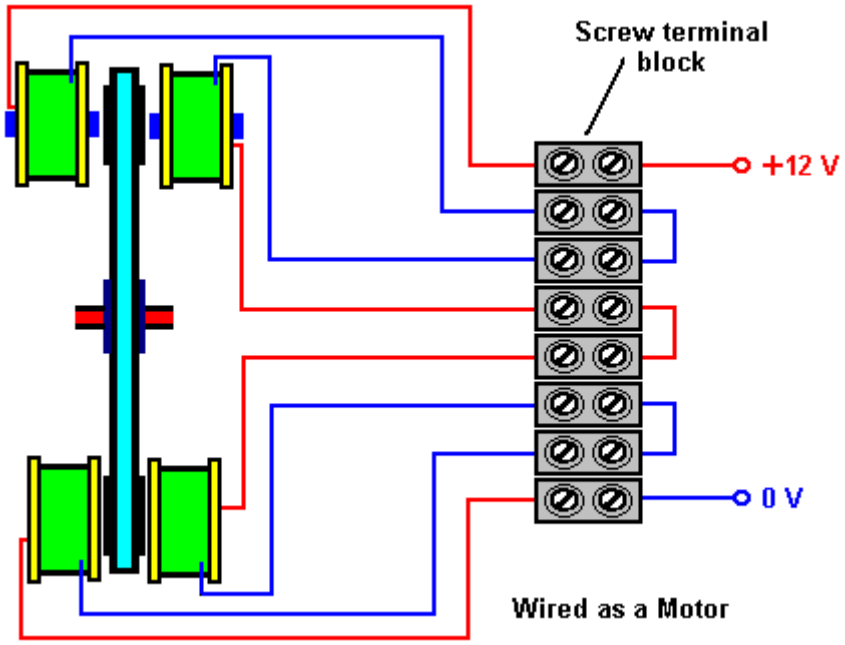


Ferrite is a better material for the cores of the coils. The stators go each side of the rotors and the hole in the middle of the stators is to give clearance for the shaft on which the rotors are mounted:

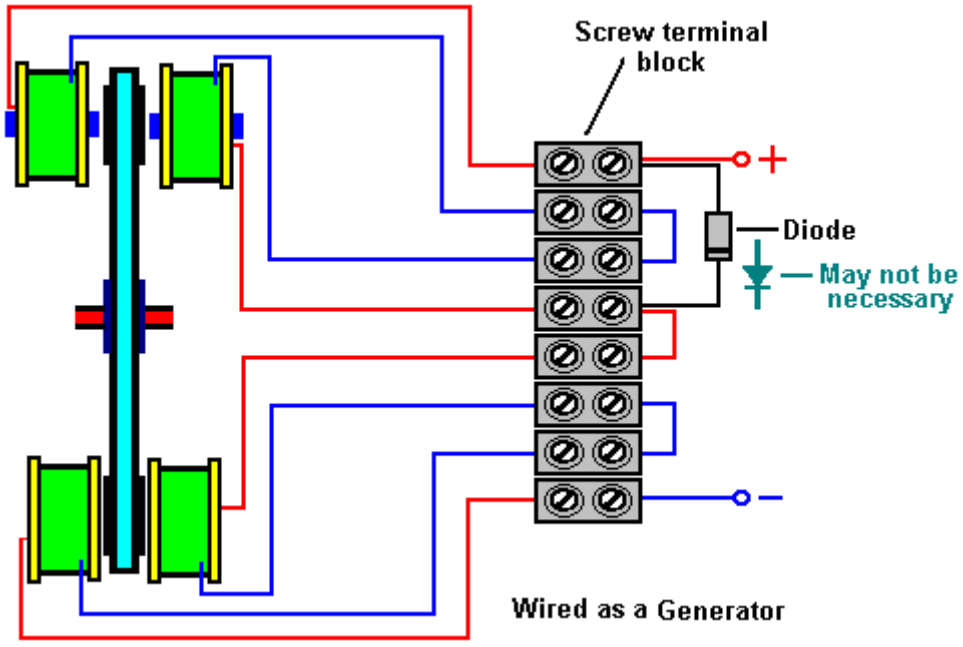


A system of this type needs accurate timing which is solely related to the rate of rotation. This is best arranged by the use of a bistable multivibrator as described in the Electronics Tutorial of Chapter 12. You will notice the two Timing Coils shown at the right hand side of the diagram above. These are used to toggle the bistable On and Off and they are adjustable in position so that both the On and the Off can be set very precisely. The output of the bistable is set to switch an FET transistor On and Off to give circuit switching which is not affected by either the switching rate or the number of times the switch is operated.

The Rotor / Stator combination can be wired to act as either a driving Motor or an electrical Generator. The difference is the addition of one diode:



With this arrangement, for each rotor, all four pairs of Cored coils are wired in parallel across each other, and all four Air-cored coils are wired in parallel across each other. To improve the clarity, the above diagram shows only one of the four pairs, but in reality, there will be four wires coming into the left hand side of each of the screw terminals.



In the case of the Generator arrangement, you have the option to connect each of the four pairs in parallel as in the Motor arrangement or to connect them in series. Connected in parallel, the coils can sustain a greater current draw, while if connected in series, they provide a higher voltage. The voltage could be further increased by increasing the number of turns on each coil.

Joseph Newman's Motor.

Joseph Newman is a man who impresses me. He performs experiments, reports the results and then bases theoretical conclusions on the results of his own experiments. This is the true scientific method.

Joseph has been granted a patent and he has written a book. I would recommend that you buy a copy of his book and help support his work by doing that, but unfortunately, as I understand it, the printing plates for the book were destroyed in a fire and printed copies of his book are effectively unobtainable. You can download a .pdf version from the http://www.free-energy-info.tuks.nl/ web site but please be aware that the overall file size is 100 Mb and so the download will take quite some time. A background download can be had from http://www.mediafire.com/view/?oe5obw6offoqfs3 or the alternative http://www.filefactory.com/file/b4g6b3d/Newman.pdf link while the link to Joseph's own web site is http://www.josephnewman.com/.

In very brief outline, Joseph has built a motor which can access free energy. He has a theory about where the excess energy is coming from and how it is acquired by his designs. He has also built a large stationary motor to demonstrate his theory and he has built a motor into a car. The car engine runs on very minor battery power. Joseph's patent is included in the Appendix.

With the kind permission of the Joseph Newman organisation, I am going to attempt to introduce you to the important scientific conclusions reached by Joseph and the Energy Machine which he designed and which is based on those conclusions. Joseph has a keen enquiring mind and thinks things through for himself rather than blindly accepting everything he is told. This description contains illustrations and wording taken from parts of Joseph's book published in 1984, and I should like to express my thanks for being given permission to use this material.

Joseph Newman's motors all consist of a very powerful permanent magnet which rotates or oscillates in or near a coil with a very large number of turns of copper wire. The coil is energised by a battery pack, and the magnetic field produced by the coil provides the force needed to move the permanent magnet. A mechanical switching device or "commutator" reverses the direction of current flow through the coil every half cycle, and in some models, it also cuts off the current input between the current reversals.

The main difference between Joseph's designs and previous motors is one of scale as Joseph uses very large coils and very large ceramic magnets weighing up to 700 pounds. His smaller motors use powerful rare earth magnets and the coils are wound with 100,000 turns of copper wire. This creates a very high coil resistance and the battery pack voltages are correspondingly high, being in the hundreds to thousands of volts range.

The torque or turning power applied to the magnet in these motors is proportional to the magnet strength, the number of turns in the coil and the current flowing in the coil. In Joseph's motors, very large torques can be developed by very small currents. In one demonstration, a motor running on 3,000 volts at 0.8 milliamps has such power that it is not possible to stop the motor by holding its two-inch (50 mm) diameter shaft, though the current can be raised by trying to stop it, to 3 milliamps, or nine watts of power.

Joseph's motors are different in other ways. If fluorescent tubes are connected across the motor coil, they light up due to the coil's collapsing magnetic field each time the current direction is switched. These fluorescent tubes are used to protect the mechanical switch from arcing damage. The additional power produced in these tubes is at a very high frequency of 10 to 20 MHz. This radio-frequency current has been accurately measured and it exceeds the battery input current by a factor of five to ten times in the different motors. The measured current and voltage were in phase, indicating a real power output.

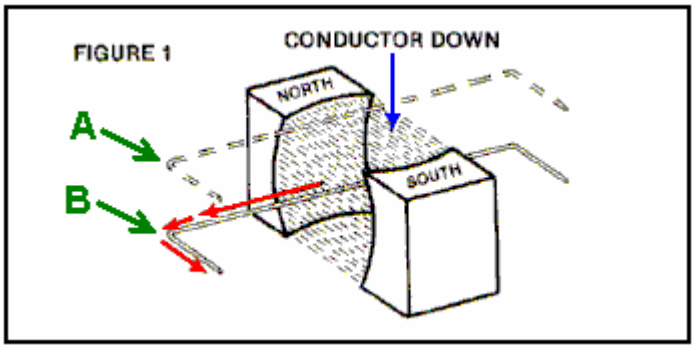
To understand the thinking behind these motors, we need to follow Joseph's experiments and the deductions which he made from those experimental results. Joseph considered, and thought carefully about statements made by the two scientific giants James Clerk Maxwell and Michael Faraday, and this led him to valuable

insights:

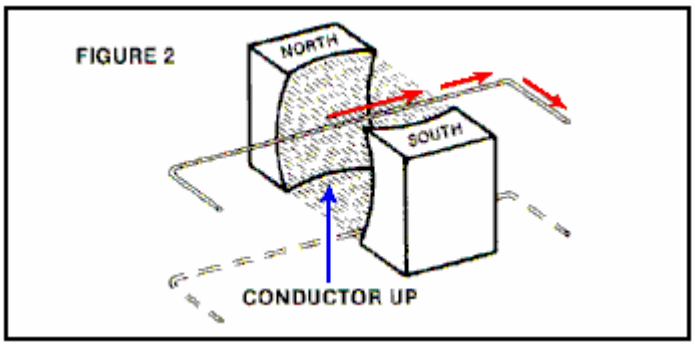
It appears that Maxwell and Faraday were the only people who considered that “lines” of magnetic force are actual physical entities and not just a method of representing notional forces and those “lines of force” are actually streams of **matter in motion**.

Maxwell says: “**In speaking of the Energy of the field, however, I wish to be understood literally**. All energy is the same as mechanical energy, whether it exists in the form of motion or in that of elasticity, or in any other form. **The energy in electromagnetic phenomena is mechanical energy**”.

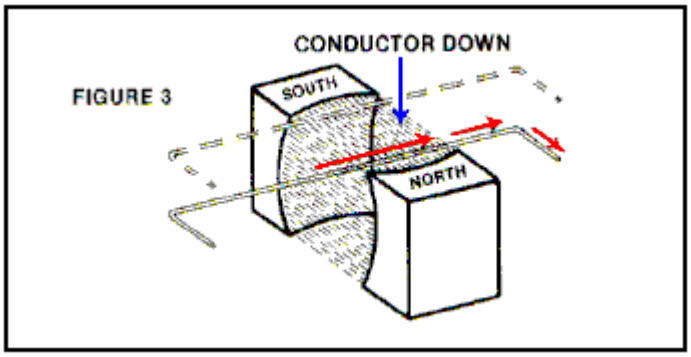
Joseph then considered Michael Faraday’s Electrical Generator and the implications of the way in which it operated:



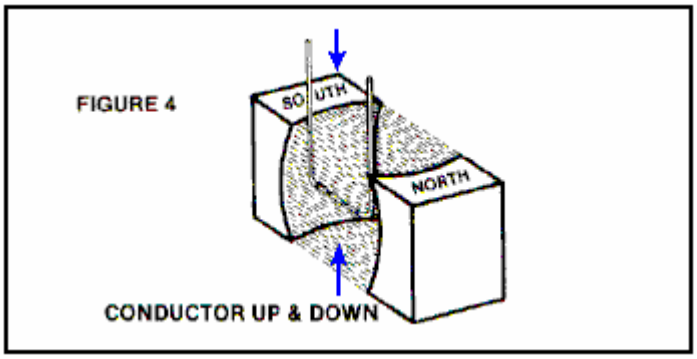
Here, a loop of wire is moved downwards from level “**A**” to level “**B**”. This movement causes an electrical current to flow leftwards along the wire as shown by the red arrows. Joseph’s question was “why does it go in that direction every time the wire is moved in that way?”



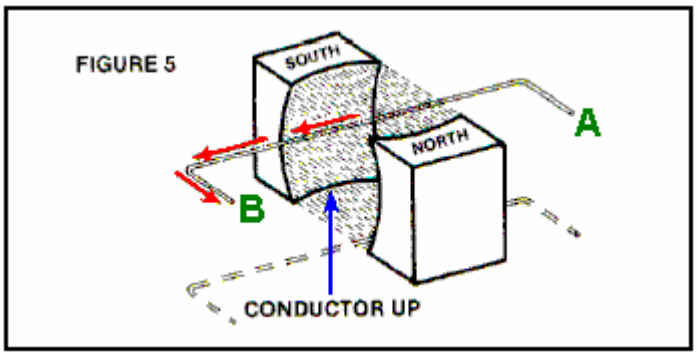
If the wire is moved upwards through the same magnetic field, then the current flowing in the wire moves in the opposite direction. Why? How does the current “know” which way to go?



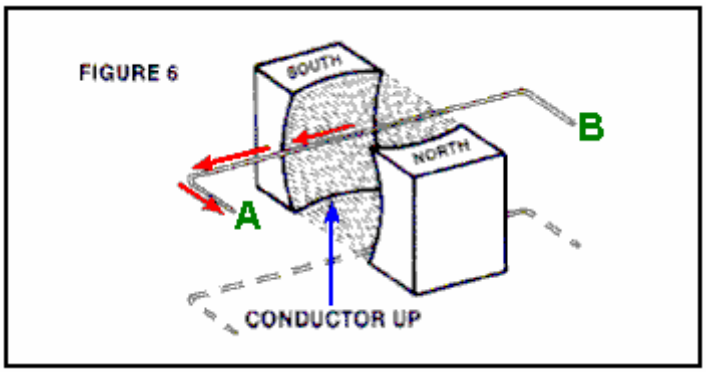
If you turn the magnetic field round by reversing the position of the magnetic fields and then move the wire loop in the same way as before, the current flows in the opposite direction. How does the current “know” which way to flow, or which way round the magnets are turned as it does not touch them?



The next interesting point is that if the wire loop is moved up and down between the magnets, but turned to be parallel to the flux flowing between the poles, then no current flows in the wire, no matter how quickly the wire is moved up and down.



Another point is that if the wire loop is moved **slowly** up through the magnetic flux, the electric current which flows as a result of that slow movement, **moves at the speed of light**, flowing from "**A**" towards "**B**".

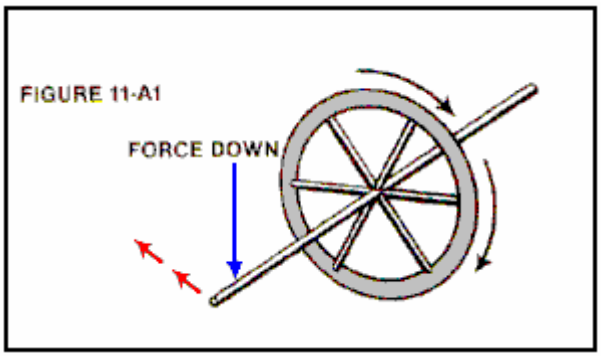


Now, if the wire loop is disconnected and turned over, the part which was at "**A**" now moved to "**B**", and the same movement of the wire carried out - the current flow is in exactly the same direction although its path along the wire is reversed (because the wire has been reversed). This shows that the direction of current flow is not affected by the wire itself.

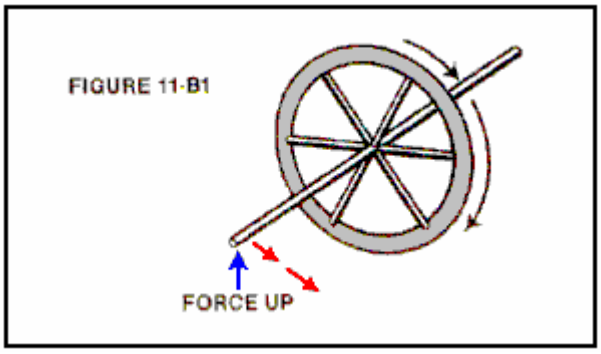
According to conventional teachings, this electric current flow was not a result of the magnetic field as the magnetic lines of force were supposed to be imaginary, consisting of Potential Energy and **no** Kinetic Energy. It became clear to Joseph that this conventional teaching was wrong. Instead, it seemed clear that the magnetic field consists of particles which have mechanical characteristics, and those particles must be moving at the speed of light within the magnetic field.

A key question seemed to be: "how does the current 'know' which direction to flow?" as the direction was always consistent. After careful consideration, it occurred to Joseph that the answer was provided by the

actions of a gyroscope:

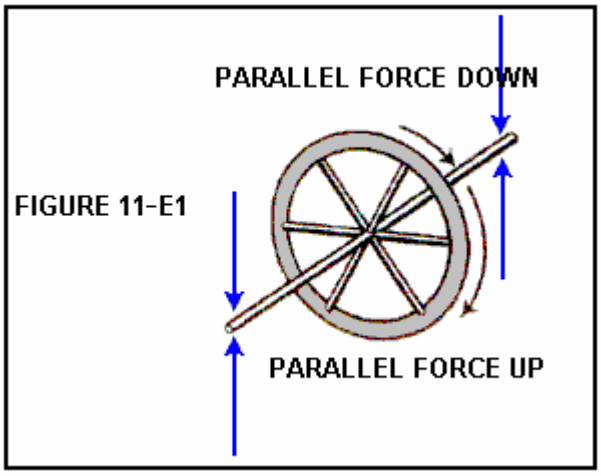


Here, if the axle of the spinning flywheel, or gyroscope, is pressed downwards it moves off in the direction shown by the red arrows. However, if the axle is pressed upwards:



then the axle moves in the opposite direction as shown by the red arrows. This effect is, of course, reversed if the direction of rotation of the gyroscope is reversed (as it will be if viewed from the other side, in the same way as the current flow direction in the wire is reversed if the magnetic poles are swapped over).

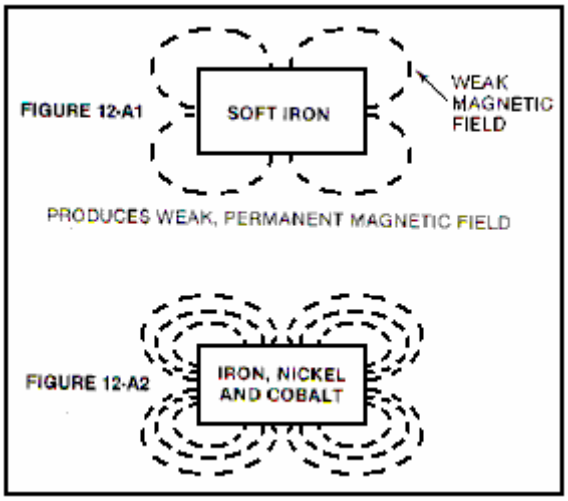
Now, if the gyroscope axle is moved up and down equally on both sides, there is no resulting sideways force:



The action of the gyroscope axle matched the current flow in the wire in every respect, so it became clear to Joseph that the particles flowing between the poles of the magnet were spinning as well as moving at the speed of light. This gyroscopic mechanical motion of the particles accounts for all of the characteristics of the current flow in a wire which is being moved through a magnetic field. This is a major insight on the part of Joseph.

May I remark that these particles are not coming from the magnet itself, but are flowing in from the zero-point energy field, that flow being caused by the broken symmetry of the zero-point energy field generated by the dipole effect of the poles of the magnet. That is why energy can (appear to) be drawn from magnets for years on end.

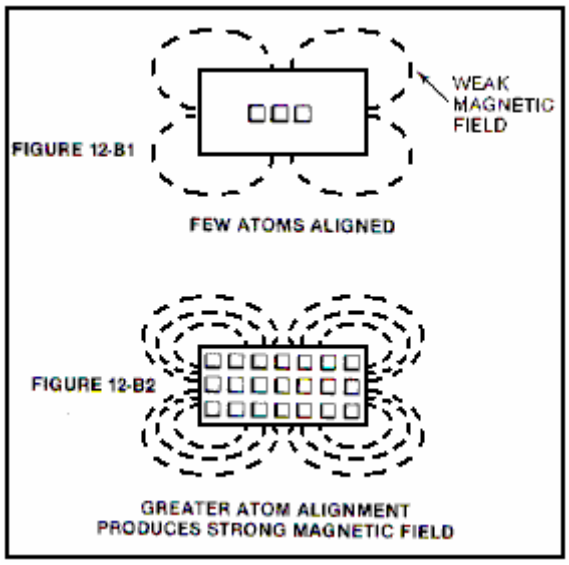
Joseph then went on to consider the physical aspects of permanent magnets. There were two very significant facts which had to be considered. The first of these is that different materials have markedly different magnetic characteristics:



A bar of soft iron does become a permanent magnet when pulsed briefly with a strong magnetic field, but if exactly the same level of magnetic pulsing is applied to a similar bar of an alloy of iron, nickel and cobalt, a permanent magnet is also produced, but the magnetic field of the alloy is very much stronger than that of the soft iron bar. This shows that the molecular structure of the bar has a major effect on the resulting magnet.

In passing, please be aware that the more powerful magnets available nowadays are so strong that they can easily injure you. If you pick up a magnet and inadvertently get close to a second one, the loose magnet will jump some inches and try to connect to the one in your hand, crushing your fingers in the process and proving very hard indeed to shift in order to deal with the injury. I have also seen it alleged that US 'AlNiCo' (Aluminium / Nickel / Cobalt alloy) magnets are deliberately doped with K40 isotope which renders them useless fairly quickly. The source of this information is highly dubious, but the extra sales advantages to the magnet manufacturers would be significant. Also, the advantages for the people wanting to suppress the creation of free-energy magnet motors would be major as many talented US inventors are likely to think that their successful magnet motors were failures because the magnets appeared to be "drained of power" by being used in their design, when in fact, the design is perfectly good. So I will leave you to make up your own mind about the matter and remark that Bill Muller found that his powerful Chinese-manufactured magnets were in perfect condition after eleven years of use.

Another point which Joseph considered was the fact that when successive magnetic pulses are applied to a ferromagnetic metal bar, the resulting magnetic field strength reaches a definite maximum value, and further pulsing has no further beneficial effect:



This is considered to be the magnetic pulse aligning atoms in the metal. Eventually, all of the atoms are aligned and so not further effect can be produced by further pulsing. This alignment can be destroyed if the metal bar is heated to a sufficiently high temperature, forcing the atoms into such an energetic state that the alignment is lost.

It should perhaps, be stressed here, that the magnet itself does not have any power, in spite of seeming to have. Tom Bearden explains this clearly by pointing out that what happens is that the opposite poles of the magnet created a “dipole” which unbalances the random ‘quantum foam’ nature of the local environment (the zero-point energy field) and that causes continuous energy flows from the environment. The “magnet” power is coming from the environment and not from the magnet itself.

If you find this hard to believe and think that you are just getting back the electromagnetic energy which you pumped into the metal when creating the magnet in the first place, then apply simple arithmetic. Assume that you get back exactly 100% of the original power and calculate how long that amount of power would allow the magnet support its own weight against gravity, when attached to a vertical metal surface. Then ask yourself how come the magnet can do it for years and years on end. Point proved conclusively?

Joseph concluded that the attraction of “unlike” magnetic poles and the repulsion of “like” poles is caused by the gyroscopic spin direction of the actual physical streams of the “lines of force”, which he has shown that both of the scientific giants, Maxwell and Faraday were convinced were actual physical entities. The intuitive genius Nikola Tesla described the zero-point energy field as having the physical characteristics of a gas, capable of having motion, exerting pressure, and yet having particle size so small that it can flow through any physical material. Joseph has concluded that this field flow has a specific spin direction as it flows, certainly for flows caused by the magnetic dipole of a magnet. It should be remembered that the scientific teaching of present day educational institutions is at least fifty years out of date. We have the most unusual situation where the scientific literature of a hundred years ago is actually of better quality than that of today which does not describe the actual world at all well. Currently, misconception is alive and very well.

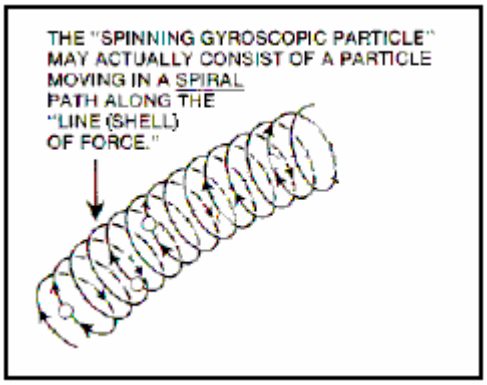
For example, Maxwell produced equations describing how the world works. Admittedly, these equations are very difficult for people to understand. H. A. Lorentz simplified these equations and his results are mistakenly described as Maxwell’s which they most certainly are not. Tom Beardon illustrates it this way; consider a sailing boat being driven along by the force of the wind against the sails:

Maxwell says that there is a vast swathe of wind blowing across the ocean, capable of powering a long row of a thousand sailing boats side by side. This is the actual physical case. Lorentz has 'simplified' things by saying "we will consider one boat and only one boat. As the rest of the wind does not touch any part of the boat we can ignore it". While that is true for that one boat, what science teaching now says is that the wind can **only** power a single boat. This is not the actual case, as the environmental wind is **not** limited to powering just one boat (sailing regattas would not be much fun if that were the case!). This, of course, is just an illustration. Maxwell's equations cover energy and power for the whole universe, and deal with all cases. Lorentz has taken a sub-set of the conditions described by Maxwell's equations, just the group which apply to "closed systems" – just one boat on the ocean. Science has latched on to this and now confidently states that everything is a "closed" system, when in fact, as the zero-point energy field flows through everything, everywhere at all times, and is capable of supplying unlimited additional energy anywhere at any time, there is probably not a single instance of a "closed" system anywhere in the universe.

Joseph Newman, and all other serious inventors, have to fight against this "conventional" science teaching, which is now so entrenched that it has become the equivalent of religious dogma, and 'scientists' are unwilling to consider valid observations which do not fit in with the very limited Lorentz concept of the environment. They say "perpetual motion is impossible" which means that Newton was wrong when he said that a moving body will keep on moving indefinitely unless some force acts on it to stop it. Presumably, then, the Earth can't keep on orbiting the Sun (gosh, I hope it doesn't stop today as that would be very awkward). Come on – get real !!

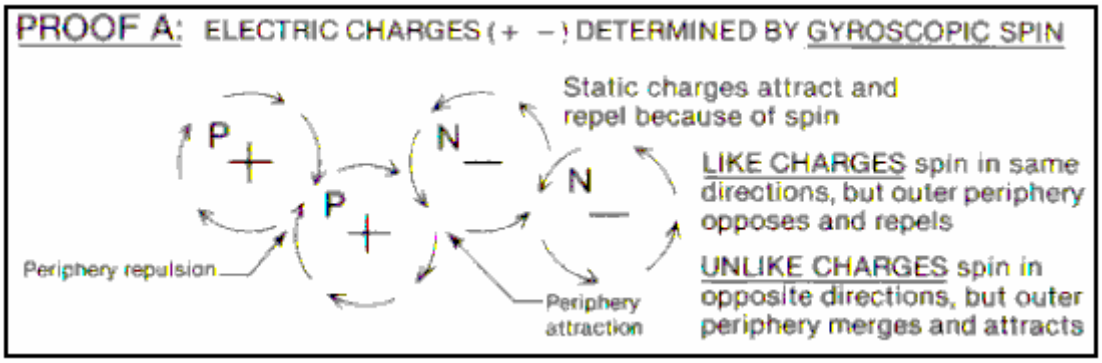
You can see then, that when Joseph performs tests and then bases his conclusions on the results of those tests, that he is applying the true scientific method, and people who say that his verified results are impossible because Lorentz says so, are not being scientifically honest. No honest person can ignore genuine scientific observations.

Joseph's deduction that magnetic lines of force are formed of actual physical particles spinning in gyroscopic motion as they move along their magnetic path at the speed of light, was not something which was obvious to scientists, in spite of the fact that both Maxwell and Faraday had both explicitly described these lines of force as being 'kinetic magnetic energy':



As a wire passes in front of and across the end of a bar magnet, the current flows in one direction, pauses, and then flows in the opposite direction. This occurs due to the gyroscopic flow direction of the particles. For instance, on one side of the South end of the magnet, the lines of force spin "up" while on the other side

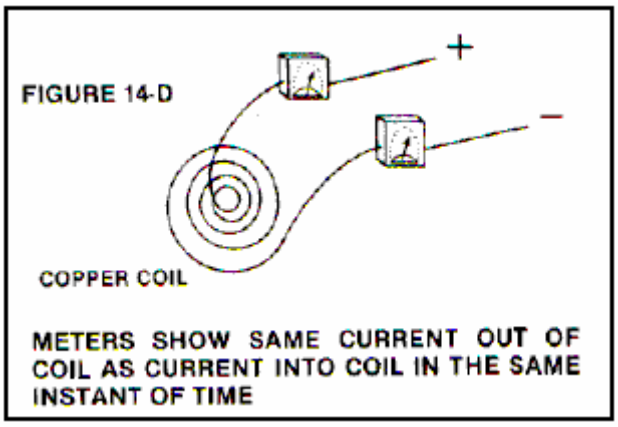
of that same South end, they spin "down". A spinning gyroscope will move at right angles to the force acting on it, so as the gyroscopically spinning particles encounter the particles of the wire, they move "up" or "down" the wire at right angles to the direction in which they first encounter the wire. Please note that it is the gyroscopic spin direction of the particles which determines magnetic 'attraction' or 'repulsion' and not the direction of flow of those particles along their line of force:



It should also be realised that although we draw 'lines' of force around a magnetic device, the reality is that these are really shells of force and the magnetic flow is really like water flowing in a river. While we may draw arrows to indicate direction and strength of currents in a diagram of the river, the reality is, of course, that there is water flowing at all points in the river and not just along the lines which we decide to draw. The same applies to the magnetic flow around a magnet, it exists like a solid mass flowing through and around the magnet. You don't see it or feel it because the particles are so small.

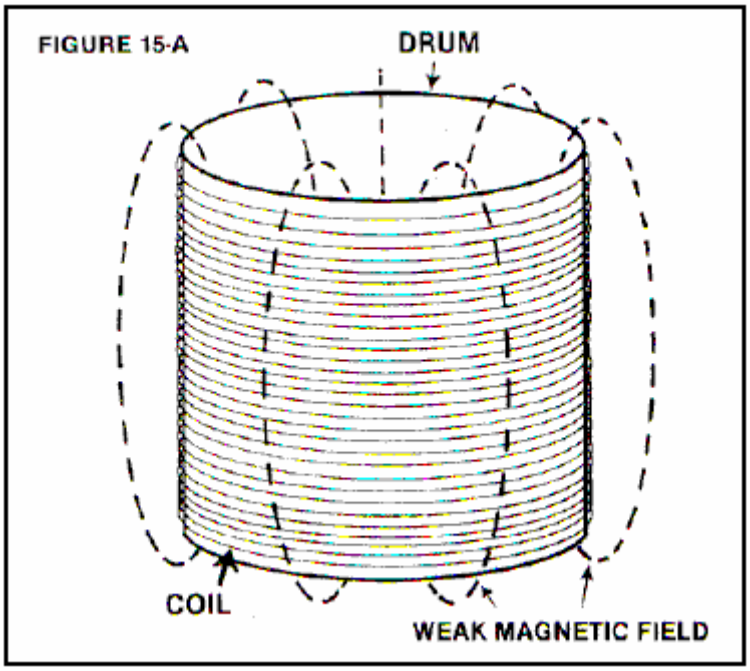
Now to the details of how to construct a device to take advantage of this magnetic movement and output more power than is required to make it operate. Let me remind you again that we are talking here of a Coefficient of Performance (COP) which is greater than 1 in a system which has an overall power efficiency of less than 100%. This is, of course, due to the additional energy flowing in from the zero-point energy field. Joseph visualises the apparent energy gain as being conversion of a small amount of matter into its energy form ($E = mC^2$), and while this is probably correct, it will be particles of the zero-point energy field which are being converted into their energy form and not particles from the metal of the magnet. It must be remembered that the particles of the zero-point energy field keep swapping over from energy to physical form all the time anyway. Energy is never "used up" but merely converted from one form to another and the zero-point energy field contains such a staggering amount of energy that all of the visible matter in the whole of the universe could be created from the energy in a single cc of the zero-point energy field. So, if a few sub-sub-sub-microscopic particles of the zero-point energy field switch into their energy form to produce what looks like excess power to us, that is an item so trivial to the field that it is not even worth mentioning – less than the effect of taking one grain of sand off a beach one hundred miles long. The conventional conception of the way things are is so far away from reality that it is ridiculous, (and that is even without saying anything about the effect that the time axis dimension has on the energy balance and flow of energy).

But back to Joseph's design. Firstly, he points out that it is generally agreed (courtesy of Gustav Kirchhoff) that in the situation shown here:



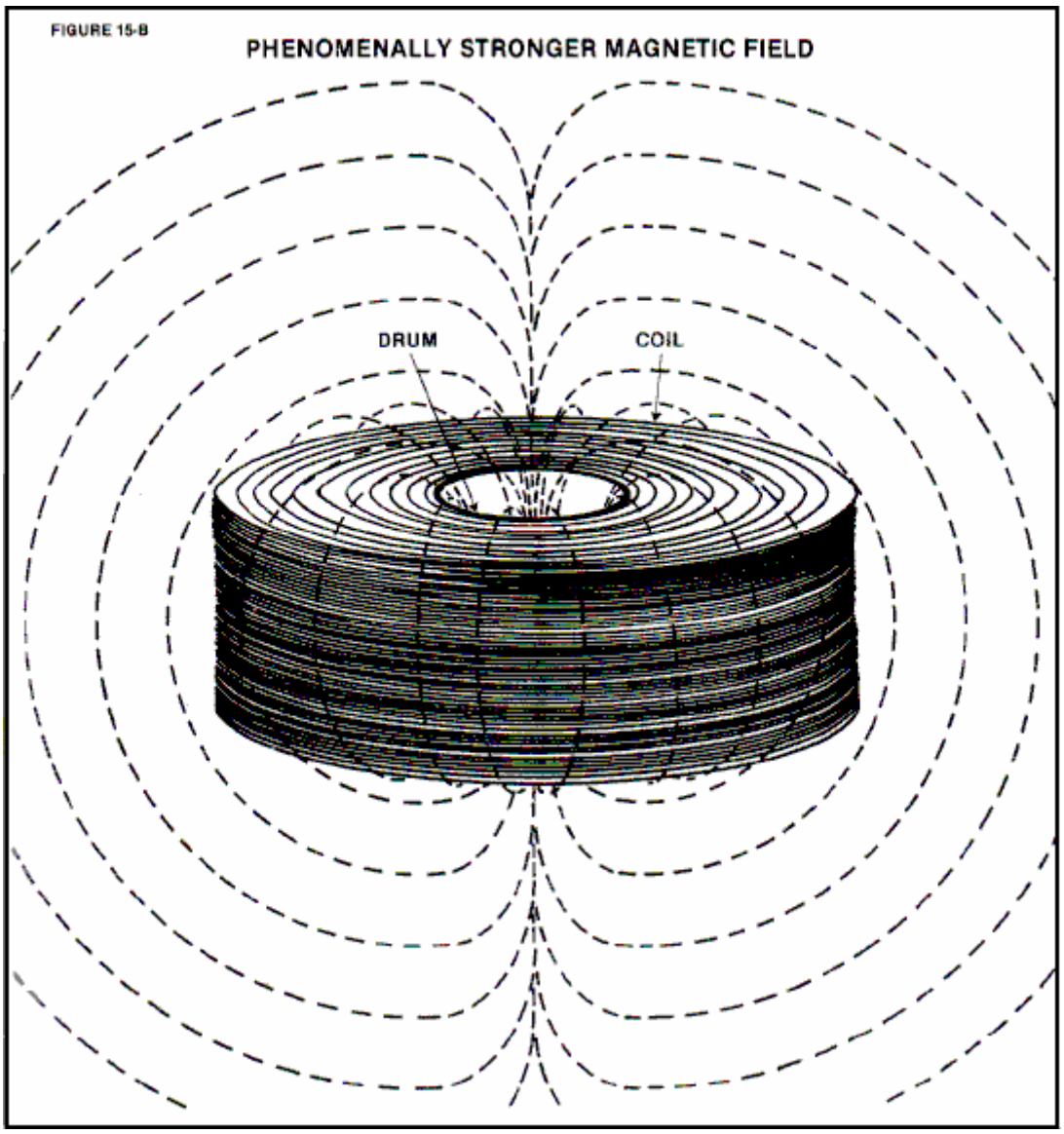
In any given instant of time, the amount of current flowing into the system ("X") is exactly matched by the amount of current flowing out of the system ("X"). **But**, if measuring equipment is attached across the coil at the moment of switch-off, an additional amount of current ("X") flows out of the coil. This is generally agreed, and it suggests that a quantity of current "X" flows into the coil and yet a quantity of "2X" flows out of it (COP=2).

Joseph examines this situation in practical detail as follows:



Consider an air-cored coil with an interior diameter of 10 feet, a height of 8.32 feet and would with 1,000 feet of 40-gauge copper wire. That length of wire has a resistance of 1,049 ohms and weighs 0.02993 pounds. If 100 volts DC is connected across it, then a current of about 95 milliamps will flow, which is a power input of 9.5 watts. With just 31.8 turns, it will produce a weak magnetic field of 0,012 Gauss, with a mere 0.000014 Joules of energy stored in it. With a tiny inductance of just 0.003 Henries, if the current is stopped and the ends of the coil shorted together, only an insignificant current would flow.

Now, repeat the experiment, but this time, use 5-gauge copper wire. As it has a resistance of 0.3133 ohms per 1,000 foot length. To equal the same resistance and match the previous current flow, a massive length of 3,348,000 feet needs to be used. This length of wire will weigh 335,469.6 pounds which is 16.77 tons. The 10-foot interior diameter coil, 8.32 feet tall, wound with this wire will have about 90,000 turns. If 100 volts DC is now connected across the coil, the same 95 milliamp current will flow with an input power of 9.5 watts, the same as before. But due to the massively larger coil, it has a magnetic field of 23.7 Gauss, which is 1,905 times larger than the previous coil, and with 116 Joules of energy stored in the magnetic field. This is a phenomenal 8,000,000 times more energy than in the 40-gauge coil of the previous example. A phenomenally larger current flow would now occur if the current input was stopped and the coil shorted out, as that would generate an inductance of 25,700 Henries which is more than eight million times the inductance of the previous coil:



Joseph then built a smaller version of his design, as shown here:

this prototype used 5-gauge insulated copper wire weighing 4,200 pounds and 300 pounds of 30-gauge copper wire wound over the 5-gauge winding, and a massive 4-foot long, 20-inch diameter permanent

magnet of some 600 pounds in weight. The coil was wound with an inner diameter of 4-feet and a height of about 3-feet, wound on a fibreglass tube. The overall weight was about 5,000 pounds.

Everybody who was it was asked: “Based on your expertise, how much power would be necessary to simply operate this device mechanically?”. Answers ranged from 200 watts to 1,000 watts. On learning that it had an air-cored coil, other skilled individuals stated that in their expert opinion, the unit would be highly inefficient since it contained no iron core. However, from the design information already presented here, it can be shown that the actual necessary power input is less than 1.5 watts providing a power output far in excess of 100%.

Dr. Roger Hastings, Principle Physicist at Sperry Univac Corporation and former Associate Professor of Physics at North Dakota State University, tested this prototype and showed that it had an 800% efficiency – that is a Coefficient of Performance of 8.0 which is impressive. In addition, Dr Hastings estimated that with a 1.5 watt power input, the back emf exceeded 80,000 watts. In operation, the 600-pound, hand-made magnet rotates at just 200 rpm.

Joseph’s patent which is in the Appendix, indicates four different ways of implementing his design principles. It is very clear that Joseph has proved his point by producing and constructing a device which Oliver Lorentz considered to be impossible, thanks to his throwing out the free-energy sections of Maxwell’s equations. Joseph Newman has clearly earned our respect.

You can see J L Naudin’s builds and tests of small models at http://jnaudin.free.fr/html/qm11bp.htm.

Daniel Cook’s Motionless Generator.

In 1871, Daniel Cook obtained US Patent 119,825 for “An Improvement In Induction Coils”. It is by no means obvious how the device described could possibly operate, and it has been suggested that the patent information is incomplete, having been edited at a later date. But as I have no direct evidence that it does not, or cannot, operate, it is shown here. Interestingly, the highly-respected Dr Harold Aspden considers this a very serious piece of equipment, operating as paired cross-linked capacitors, and his opinion carries very considerable weight.

It is a very simple device which could be interesting to test, especially as it does not involve any electronics or complicated construction. The patent can be summarised as follows:

My invention relates to the combination of two or more, simple or compound, helical coils with iron cores or magnets, in such a manner as to produce a constant electric current without the aid of a battery.

Fig.1 represents the different parts of a compound helical coil and iron core.

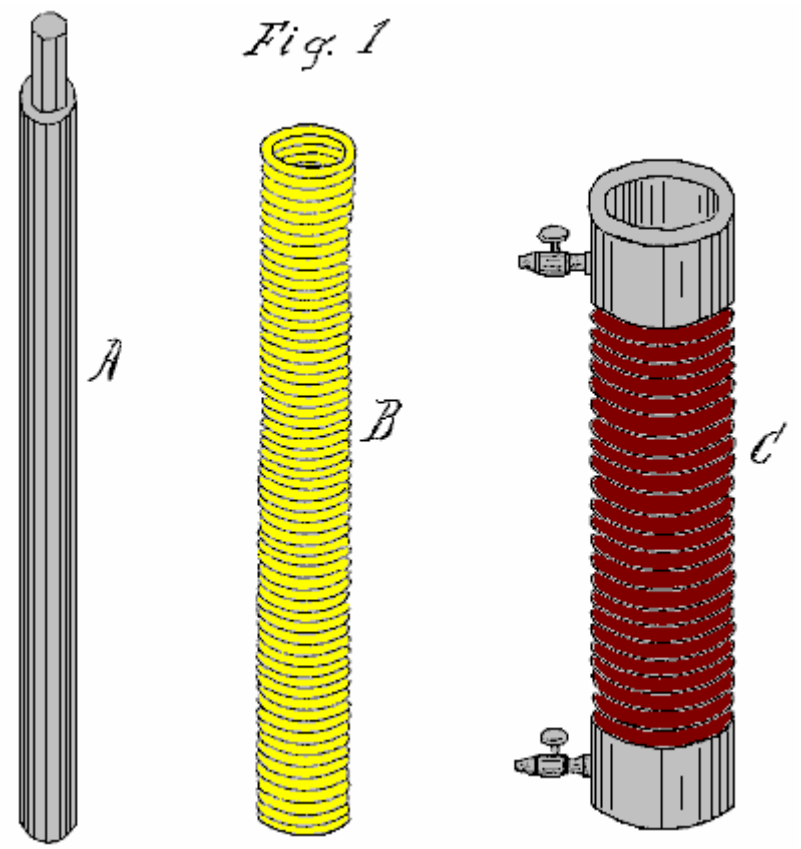
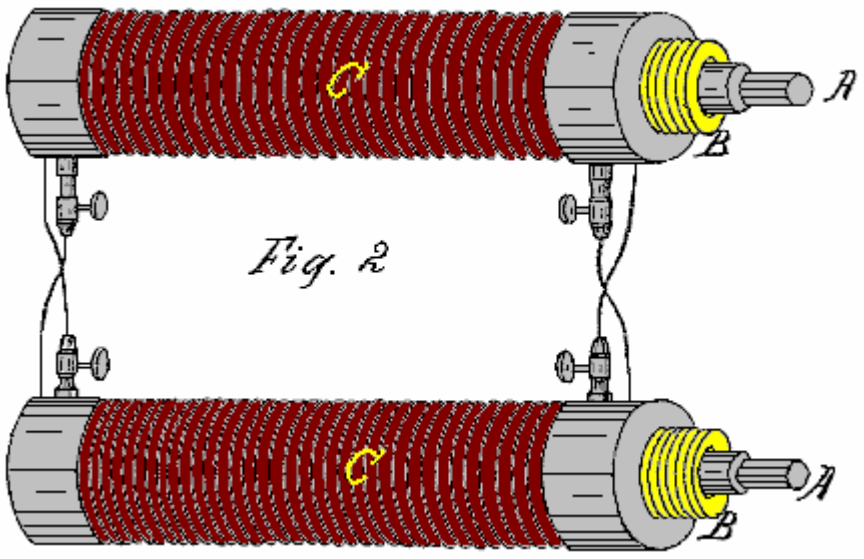


Fig.2 is a perspective view of my invention.



In carrying out my invention, I do not confine myself to any particular mode of coil construction or to any particular size of wire, observing only that the quantity of wire in the various coils must be sufficient to produce the required result; also, the material used to insulate the wires must be suitable for producing the required result. However, I generally prefer to use the same size of wire in the construction of both simple and compound coils.

When constructing simple coils, to produce the required voltage and current, it is desirable to use a long iron core as shown as **A** in **Fig.1**. This iron core may be two, three or even six feet in length, and two, three or more inches in diameter. The coil should be wound from good quality copper wire, insulated with silk or

shellac. The iron core **A** may be a solid bar or a bundle of separate iron wires, the latter giving better results and providing more current for any given wire diameter. While the wire may be fine or coarse, I prefer to use No. 16 or even heavier wire, as the power output is in proportion to the length and diameter of the wire.

When using compound coils, it is preferable in some cases to use a small wire, say, No. 30 or even less, for the primary coil, and No. 16 or even larger for the secondary coil. With this combination, the initial secondary current of the primary coil being very small in comparison to the terminal secondary current of the secondary coil, offers little resistance to the terminal secondary, hence a quicker action is obtained. Alternatively, the primary coil may be of uninsulated wire coiled into a solid helix, being insulated only between the coils, in which case there is little or no opposing initial secondary current.

Helically wound coils alone with large quantities of wire will produce similar results. A ribbon spiral may be substituted for the secondary coil **C**, say, of three, six, twelve or twenty-four inches in width and of any convenient length, but always of sufficient length to raise its output current to the level necessary to sustain itself through its action on the primary coil **B**. In the use of compound coils, it is important that the secondary coil should be wound in the same direction as the primary coil, and the primary and secondary coils be cross-connected as shown in **Fig.2**. The action will then be as follows:

The secondary current of the secondary coil **C**, will circulate through the opposite primary coil **B**, while at the same instant, a secondary current from the primary coil **B** will be generated and circulate through the opposite secondary coil **C**, both currents flowing in the same direction in the opposite coils **B** and **C**, producing a combined magnetic action on the iron core **A** in the centre. The opposing initial secondary currents of the two coils **B** and **C** being overpowered, do not show in the main circuit **D** of the device, there being eight distinct currents developed in the action of one entire circuit of the two pairs of coils, two terminal and two initial secondary currents to each pair of coils, the four initial secondaries constantly opposing the circulation of the four terminal secondary currents, but the initial secondaries being of much lower voltage and current than those of the terminal secondary, are overcome, leaving a sufficient surplus terminal power to overcome the resistance of the primary wire and charge the bar **A** to the degree needed to reproduce itself in the opposite secondary coil. By this means, a constant current is kept flowing in all of the coils.

These coils may be constructed using 500 feet to 1,000 feet or more for each of the primary and secondary coils. The longer, and better insulated the wire, the greater is the power obtained from the device. The larger the wire diameter, the greater the current obtained.

If only single coils are to be used, it is preferable to have a wire length of 1,000 feet or more in each coil. The action is the same as with the compound coils, but only four currents are produced: two initial and two terminal currents, the latter flowing constantly in the same direction - in effect, there being only one current in the same direction.

The action in the coils may be started by using a permanent magnet, an electromagnet or by pulsing an extra coil wound around the outside of one of the coils of the device. If the load circuit is broken for any reason, the current stops immediately. It is then necessary to perform the start-up procedure again to get the device restarted. This can be overcome by permanently connecting a resistor across the terminal of the load so that if the load circuit is broken, the device can continue under very much reduced current until the load is restored. By this means, the device becomes the direct equivalent of a battery.

A rheostat **D** may be introduced into the main circuit to limit the current and prevent the overheating of the coils through the drawing of excessive amounts of current. The iron cores may also be used for producing electromagnetic motion when the device is operating.

Note: Interesting replication attempts are shown at
http://www.overunity.com/2630/the-brnbrade-coiloverunity/#.VClsAncXJRw.

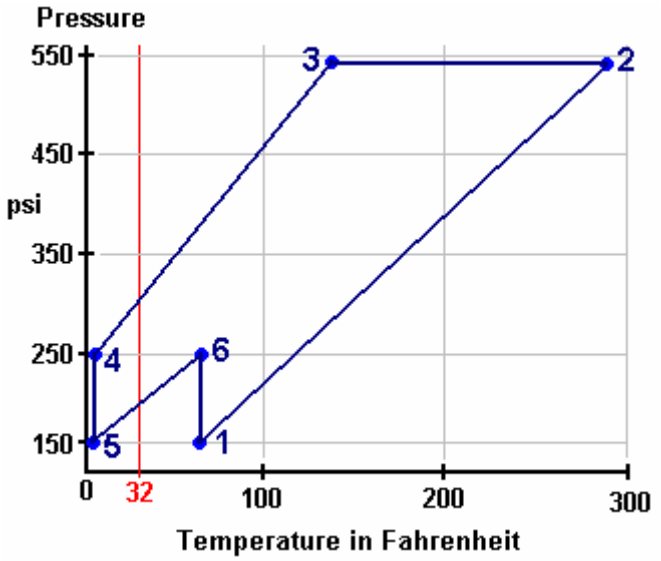
Michael Eskeli's "No Work" Generator.

One of the greatest expenses for many families is the cost of heating or cooling a home. Any device which can help with this task is definitely welcome. Michael Eskeli has produced several most interesting designs which may have been overlooked due to lack of emphasis of what they do.

Normally, a central heating system uses an expensive method of heating a liquid, typically oil, which is then pumped through radiators around the building by a low energy pump. The vast majority of the cost is in heating, typically, a furnace and very little is spent on moving the heated liquid through the radiators. In this

design from Michael, the cost of the heating is zero, and all that is left is a low-power (quarter to half horsepower) input, needed for spinning a rotor against the friction of its bearings and stuffing box.

As this seems impossible, a little very technical explanation is given here. This information has come from the web site of Scott Robertson at http://www.aircaraccess.com/ with his kind permission. In broad outline, the device comprises of a disc-shaped housing with a closely-fitting rotor spinning inside it. A gas under pressure and a liquid under pressure are both fed into the device and they intermingle in a pulsating sequence which alternately compresses and releases the pressure on both fluids. This heats both fluids very effectively, and most interestingly, without the use of any user-supplied heating power and without the use of any heating fuel. This next paragraph is for Engineers, so if you don't understand it, then just ignore it, as the important thing is to understand what the device does, rather than exactly how it does it.



The Heat-Pump Work Cycle: The example diagram above shows the sequence of events caused by the rotation of the disc inside the device housing. This "**Pressure / Enthalpy**" or "**Pressure / Internal-Energy**" diagram shows the pressures and temperatures during a single pressure cycle of the device. Using nitrogen as the gas, the cycle starts at point "**1**" which has a pressure of 150 psi and a temperature of sixty degrees F. A pressure wave now hits the mix of nitrogen and the liquid. This pressure wave moves us to point "**2**" where the pressure has been boosted to 540 psi which raises the temperature to 280 degrees F.

Moving to point "**3**" is where the wanted heat is passed throughout the gas a the liquid (performing the heating task which is the whole object of the exercise), even though the pressure is maintained, so at point **"3**" there is a pressure of 540 psi and a temperature of 138 degrees F. Next, comes a major drop in pressure, taking us to point "**4**" pulling the temperature down to below freezing: 250 psi at just 4 degrees F. At point "**5**" the pressure is dropped further to 150 psi, still at 4 degrees F. Point "**6**" takes us to 250 psi at 60 degrees F from where the cycle takes us back to point "**1**", and the sequence starts all over again.

The compression takes place on leg **1** to **2** and leg **5** to **6**. The actual amounts are 53.2 and 13.5 respectively, giving a **Compression Total** of **66.7** B/lb.

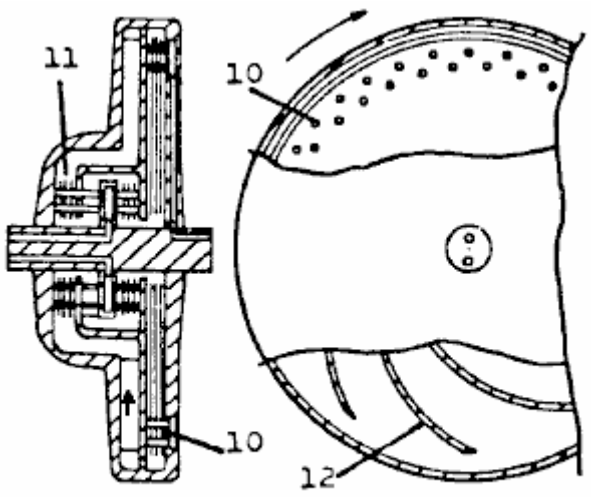
The expansion takes place on leg **3** to **4**, leg **4** to **5**, and leg **6** to **1**. The actual amounts are 31.6, 16.6 and 18.7 respectively, giving an **Expansion Total** of **67.0** B/lb.

As these two are virtually identical, the overall result of a complete cycle is effectively **work-free**.

This work cycle can be readily performed by the **Centrifuge-Type Heat Pump**. This is a unit which has only one moving part, the rotor, the working fluid, such as nitrogen, is sealed in with the rotor and circulates in passages in the rotor. The circulation of the working fluid inside the rotor is accomplished by density control alone, in accordance with the work-cycle shown above, and **there is no work input to the working fluid from the rotor shaft**. Thus the work input for the heat transfer is nil, and a work-free heat pump results.

In the diagram shown below, an axial cross-section and an end view with sections removed, shows a typical heat pump rotor suitable for use with the work-cycle discussed above.

In the diagram, **10** is the heated-fluid heat exchanger, **11** is the heat-supply heat exchanger in two parts, and **12** are the vanes in passages which return the working fluid from the periphery to the centre of the disc. The work cycle process is of the non-flow type for the working fluid in this rotor, and this provides higher performance that that in the example shown above.



Centrifuge-type Heat Pump

The fluid to be heated is usually a liquid, such as water, which enters and leaves the rotor via the rotor shaft. Similarly, the heat-supply fluid circulating through heat exchanger **11**, is a liquid which enters and leaves via rotor shaft passages.

The work-free heat pump obviously has many uses. One such use is in heating all types of buildings and homes, resulting in cost-free heating, since no fuel is needed, and the power usage is nearly nil. In the heat pump shown above, power is needed to drive the rotor against friction which may require a quarter to half horsepower motor.

Another use is in power generation, resulting in cost-free power since the unit uses no fuel, the energy source being either ambient air, or water from some natural source. (Attached turbine generates the power; part of this is used to overcome the heat pump friction loss and the remainder is available for generation of electricity). Further uses are in portable power and transportation vehicles, etc.

The apparatus and methods and work cycles are patented. For basic heat pump, see US Patent 3,926,010 and Canadian Patent 984,827.

Michael Eskeli.

Here is one of Michael's many patents:

US Patent 3,650,636 21st March 1972 Inventor: Michael Eskeli

ROTARY GAS COMPRESSOR

ABSTRACT

Method and apparatus for a compressor for compressing air, gases and vapours isothermally using a liquid stream to compress the gas; the liquid issuing from an impeller intermittently, with the gas being entrained between these liquid pulses and compressed by the liquid; the liquid having high kinetic energy when leaving the impeller and in slowing the kinetic energy is converted to pressure for both the liquid and entrained gas.

Also, this compressor may be used advantageously to compress vapours, wherein the liquid is the same fluid as the gas, in which case condensation of the gas to the liquid occurs, and work of compression is reduced.

US Patent References:

1009908		November 1911	Lafore
1115942		November 1914	Kieser
1192855		August 1916	Buss
1488388	Centrifugal pump	March 1924	Hariveau
1521270	Vacuum pump	December 1924	Bogdanoff
2007138	Boiler feed pump apparatus	July 1935	Becker
3001691	Jet operated device for circulating or compressing a fluid	September 1961	Salmon et al.
3081932	Gas or fluid compressor	March 1963	DeLancey

BACKGROUND OF THE INVENTION

This invention relates generally to devices for compressing gases, air and vapours, in which a liquid is in intimate contact with the gas or vapour to be compressed.

DESCRIPTION OF PRIOR ART

There are numerous devices and machines available for compressing a gas or a vapour. In some of these machines a liquid is rotated inside an eccentric casing, so that the machine rotor will cause the liquid to pulsate and the space between the rotor blades is increased or decreased, and this variation compresses the gas. These machines are called liquid piston type machines. Another device is the jet ejector compressor, where a stream of liquid or gas is used to entrain the gas or vapour to be compressed, and the kinetic energy of the stream is converted in a diverging nozzle to a pressure.

The main disadvantage of the liquid piston type machine is its poor efficiency, since the liquid is rotated in the machine and requires relatively large power input for compressing the gas. In the ejector compressor, the velocity of the liquid stream is limited and it entrains poorly of any gas; therefore the efficiency of the device is very poor. The available kinetic energy in the liquid stream is high, but due to poor entrainment of the gas by the liquid, results for the device are poor.

BRIEF DESCRIPTION OF THE DRAWINGS

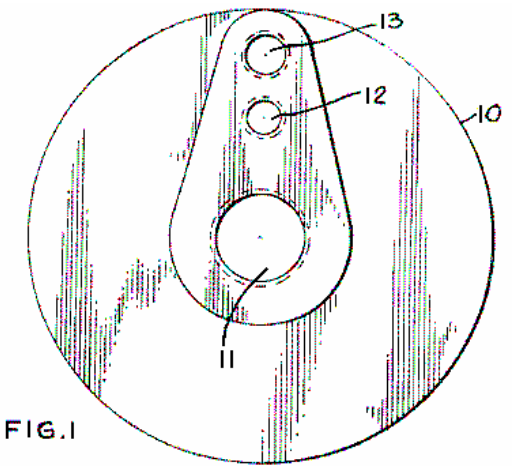


FIG.1

Fig.1 is an end view of the compressor casing, showing the exterior.

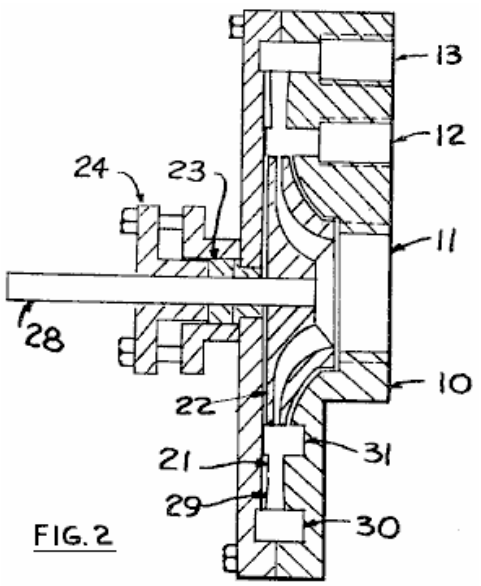


Fig.2 is a side view and a section of the casing and the impeller of the compressor.

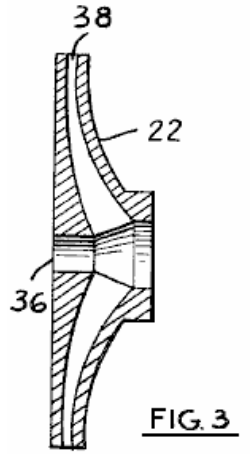


Fig.3 is a side view and a section of the impeller, and

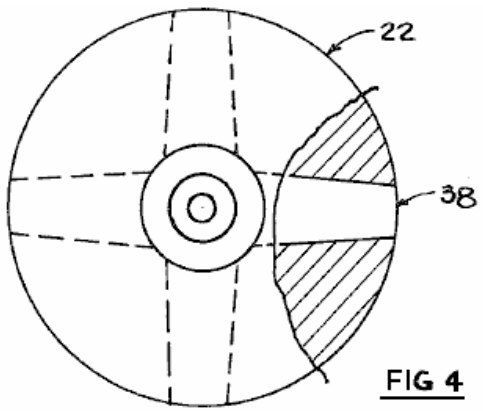


Fig.4 is an end view of the impeller, showing the fluid passages.

DESCRIPTION OF PREFERRED EMBODIMENTS

It is an object of this invention to provide a method and a device for compressing gases or vapours essentially isothermally in which the kinetic energy contained by a liquid stream is used to compress said gas to a higher pressure where the liquid in slowing in speed will increase its pressure and increase the

pressure of the gas being entrained in it. Also, it is an object of this invention to provide a method and a device in which the gas may be partially or fully be condensed in the liquid stream thereby lowering the work of compression; this occurring when the gas or vapour being compressed is the same fluid as the liquid; that is, the gas being compressed is the vapour phase of the fluid, and the liquid being used for as the motive fluid is the liquid phase of the fluid.

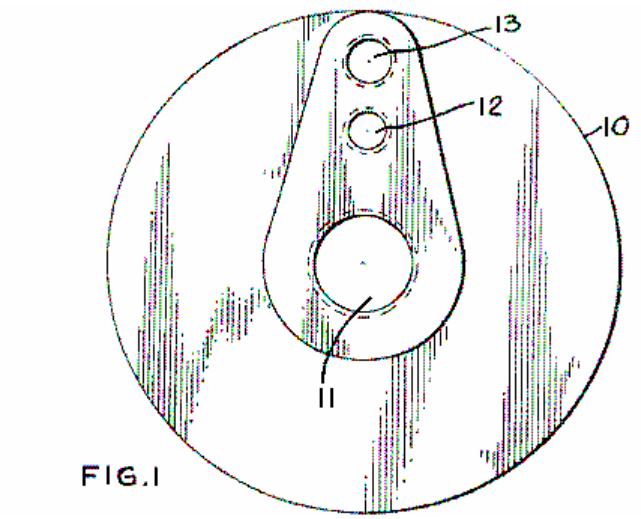


FIG.1

Referring to **Fig.1**, there is shown an end view of the compressor, where **10** is the compressor casing, **11** is the liquid inlet, **12** is the gas or vapour inlet, and **13** is the outlet.

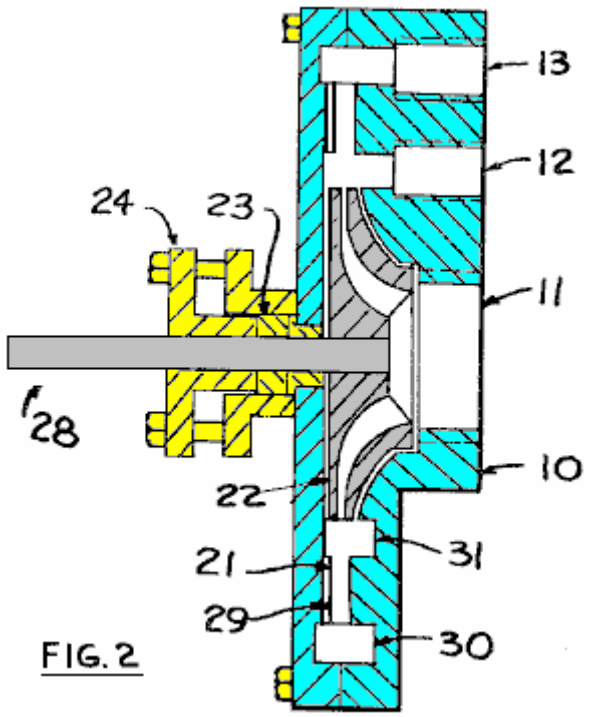


FIG. 2

In **Fig.2**, a side view of the compressor is shown. The impeller **22** is rotated by shaft **28**, supported by bearings and sealed by packing **23** and stuffing box **24**. Alternately a mechanical seal could be used. The liquid that is used as the motive fluid enters through opening **11**, passes through the impeller **22** and leaves the impeller at a high velocity and entering the throat section **21** and from there the diffuser section **29** in the casing **10**. After leaving the diffuser at a higher pressure, and at a lower velocity, the gas and liquid mixture is collected in annular space **30**, and from there passes out through opening **13**. The liquid entrains gas from annular space **31**, and the gas enters the annular space from outside through opening **12**.

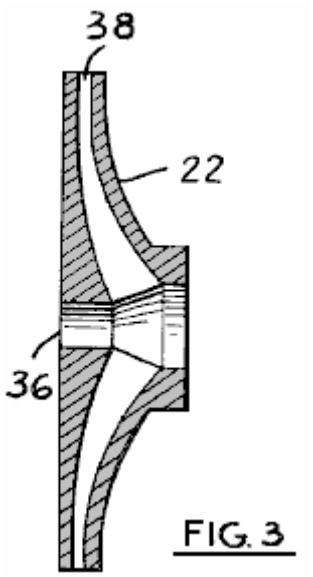


FIG. 3

In **Fig.3**, the impeller **22** is shown in more detail, where **38** is the fluid passage, and **36** is the opening for the drive shaft.

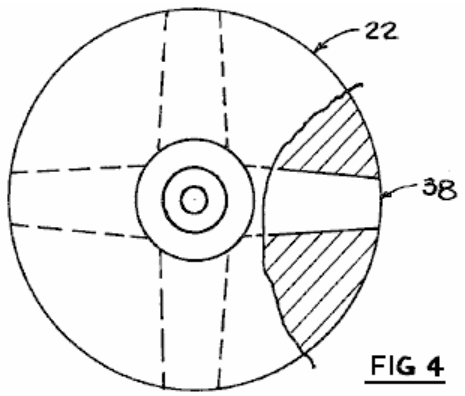


FIG 4

In **Fig.4**, the impeller is shown, with **22** being the impeller and **38** being the fluid passage.

In operation, the compressor functions in a manner similar to a jet ejector compressor. A motive fluid is accelerated in a passage in the impeller to a high velocity; this corresponds to the motive fluid nozzle in a jet ejector. However, the fluid stream issuing from the impeller, when it rotates, is not continuous as seen by the compressor casing, since in this particular instance, the impeller has four fluid passages, with solid material between them. Therefore, the flow from impeller, as seen by the compressor casing, is pulsating, with empty spaces between the high speed liquid; these empty spaces being filled by the gas from the annular spaces, item **31**, **Fig.2**, and the gas being rapidly moved with the liquid to the outer annular space **30**, and from there to discharge. This pulsating action improves the entrainment of the gas by the liquid, and more fully utilises the kinetic energy available in the liquid stream.

The sizing of the fluid passages and the calculations related to them, are fully described in thermodynamics literature for jet ejectors and for steam injectors. The space of the passage **38** in **Fig.3**, would be either converging for liquids that do not vaporise when leaving the passage; or the passage could be diverging at its outlet for fluids which will vaporise either partially or fully when leaving the passage. Of the non-vaporising liquids, water would be an example, and of the partially vaporising types, butane would be an example, both at atmospheric temperatures, and at low pressures. As illustrated in **Figs. 2-4**, passageways **38** comprise a converging section nearest the centre of the impeller but are at least non-converging at the discharge section. Preferably, the at least non-converging section is a diverging section for better taking advantage of the energy available in the motive fluid to achieve higher exit velocities.

The fluid passages shown in **Fig.4**, item **38**, can be radial as illustrated, or be forward or backward curved, depending on the fluid used. Also, the throat section **21**, of **Fig.2**, may have vanes of proper shape to prevent circular motion of the fluid after it leaves the impeller. Vanes of this type are commonly used in

turbines and pumps and so are not described here. There are four fluid passages shown in **Fig.4**, but this number will be determined when calculations are made for the size of the passages, and the frequency of pulses of liquid required to maintain suitable pressure and volume relationships inside the compressor; also, the rotational speed of the impeller would enter into these calculations.

Normally, the amount of liquid is large when compared to the amount of gas or vapour. Therefore, when compressing a gas, the heat of compression from the gas is transferred to the liquid, resulting in a temperature increase of the liquid, as well as of the gas. This temperature increase is much less than it would be for the gas alone, resulting in nearly isothermal compression, and therefore reduced work of compression, as compared to isentropic compression which is often used in rotary compressors. Also, if a liquid that will expand in the impeller is used, with an expanding fluid passage, the temperature of the motive fluid is lowered, and the fluid velocity greatly increased, resulting in much better efficiency for the compressor; this is similar to the function of converging-diverging diverging nozzles in jet ejectors.

The operation of the compressor may be inferred from the above descriptive matter. A liquid source is connected to the impeller inlet **Fig.1, 11** and a gas or vapour source is connected to the gas inlet **Fig.1, 12**. Discharge from the compressor is from **Fig.1, 13**. A suitable power source, such as an electric motor, is connected to shaft **Fig.2, 28**, causing the shaft to rotate. The liquid is accelerated by the action of the impeller, and as it passes through the annular space **Fig.2, 31** in a pulsating flow, it entrains the gas and carries it to annular space **30**, from where it discharges.

Materials of construction for the compressor would be similar to those used to make pumps for pumping liquids. Cast iron, steel, bronze, brass, stainless steel and various plastics could be used.

CLAIMS

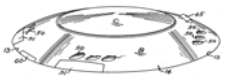
What is claimed new is as follows
1. A machine for compressing gaseous fluid and having the major components of:
2. The machine of claim 1 wherein said at least non-converging section is diverging.

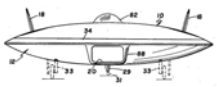
The Work of Karl Schappeller.

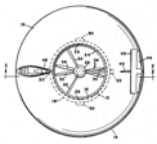
There have been a number of quite outstanding men who have had great insight as to how the universe is and how it operates. One of these is Karl Schappeller who is virtually unknown. One of the reasons for this is the fact that publicising his work has been strongly opposed by people who do not want his understanding to become widely known. One device produced by Karl in order to prove that his understanding of things was correct, produced substantial amounts of excess energy, and while I do not know of anybody who has replicated his device, I am including here, a short presentation on the subject, written by Henry Stevens and the book by Cyril Davson can be downloaded from http://www.free-energy-info.tuks.nl/Davson.pdf and read in full.

In the presentation by Henry Stevens, he mentions UFOs or flying discs. There has been a propaganda campaign waged against the general public for more than fifty years now, with the objective of persuading people that “UFOs” are not real, and if they were, then they would be the vehicles of “little green men”. This campaign has been remarkably effective and members of the public in general will immediately dismiss flying discs as being “impossible” and not something which any sane person would consider for a moment. This attitude is based on an almost total lack of knowledge of the facts. There are at this time, large numbers of flying discs, built by humans and capable of spectacular flight abilities. There are two varieties: those which need the atmosphere to operate and those which don’t.

If you feel that this is a “load of rubbish” then take a look at the following US patents:

US 2,718,364, Ernest Crabtree 

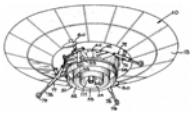
US 2,772,057, John Fischer 

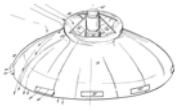
US 2,876,965, Homer Streib 

US 2,912,244, Otis Carr

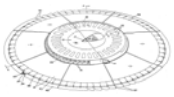
US 2,927,746, Walter Mellen

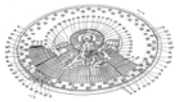
US 2,935,275, Leonard Grayson

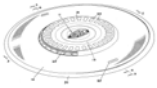
US 2,953,320, Robert Parry 

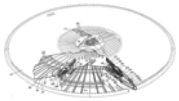
US 2,997,254, Thomas Mulgrave 

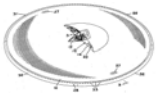
US 3,018,068, Frost & Earl

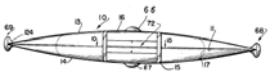
US 3,020,002, John Frost 

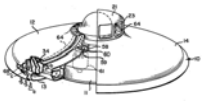
US 3,020,003, Frost & Williams 

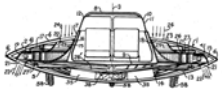
US 3,022,963, Frost & Earl 

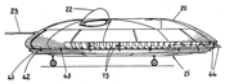
US 3,024,966, John Frost 

US 3,065,935, Duberry/Frost/Earl 

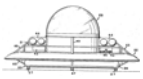
US 3,066,890, Nathan Price 

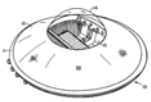
US 3,067,967, Irwin Barr 

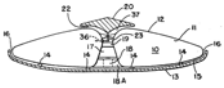
US 3,123,320, Eldon Slaughter 

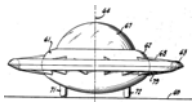
US 3,124,323, John Frost 

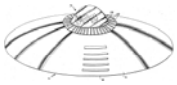
US 3,243,146, Paul Clover

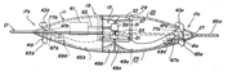
US 3,312,425, Lennon & Varner 

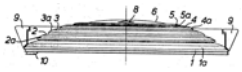
US 3,395,876, Jacob Green 

US 3,397,853, William Richardson 

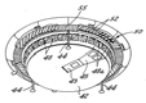
US 3,410,507, Paul Moller 

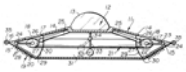
US 3,432,120, Efrain Guerrero 

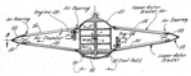
US 3,442,469, Troy Davis 

US 3,469,802, Roberts & Alexander 

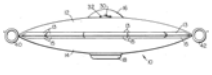
US 3,514,053, Gilbert McGuiness

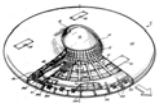
US 3,519,224, Boyd/Mallory/Skinner 

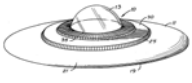
US 3,750,980, Samuel Edwards 

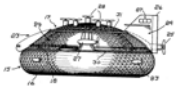
US 3,774,865, Olympio Pinto 

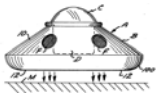
US 3,946,970, Ben Blankenship

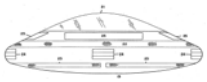
US 4,014,483, Roderick MacNeil 

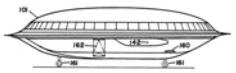
US 4,193,568, Norman Heuvel 

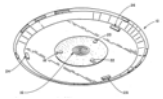
US 4,214,720, Edwin Desautel 

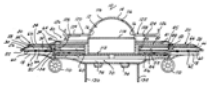
US 4,269,375, John Hickey 

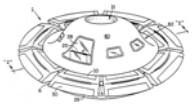
US 4,457,476, Frank Andresevitz 

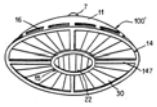
US 4,804,156, Rodney Harmon 

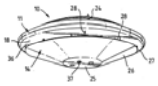
US 4,824,048, Kyusik Kim 

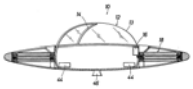
US 4,955,962, Christian Mell 

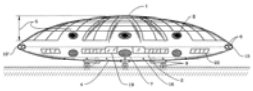
US 5,072,892, Alfred Carrington 

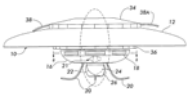
US 5,170,963, August Beck 

US 5,178,344, Vaclav Dlouhy 

US 5,203,521, Terence Day 

US 5,344,100, Allan Jaikaran 

US 5,351,911, George Neumayr 

US 6,270,036 Charles Lowe 

This small selection of forty-six patents is restricted to just those which have the well-known circular “flying saucer” shape. Do you seriously think that not a single one of these patents had a test prototype which flew or that they were all piloted by “little green men”?

I have no interest in flying machines which need an atmosphere as they are just advanced versions of conventional aircraft. In the early 1900s, Nikola Tesla designed and built what he described as his “flying machine”. This was a small device without wings and which ‘flew’ without the use of a fuel. This design of Tesla’s was taken by the Germans and during World War II, developed and experimented with. After the war, it was taken to the USA and developed further at Groome Lake, and at this time, the US, UK, Canadian and Russian governments have got large working copies which they keep as secret as they possibly can. One major cover story is that these craft belong to “extraterrestrials” who have such an advanced level of technology that we will never be able to understand it. It is a good story, as it is not possible to disprove it. If

you want a good deal of specific information on this, then read “The Hunt for Zero Point” by Janes researcher and writer Nick Cook.

While I have no particular interest in electrogravitic propulsion (or "anti-gravity" as it is popularly known and incorrectly termed) if the subject interests you then take a look at the YouTube video sequence of Stan Deyo who was one of the people who developed electrogravitic propulsion for an international group, several decades ago. Stan explains "gravity" as being the product of spin affecting the space-time continuum.

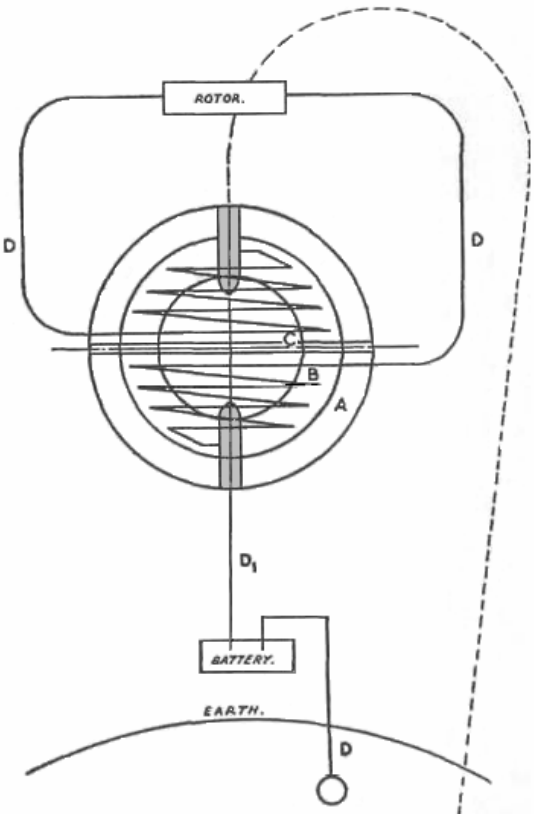
The Karl Schappeller Device

Author: Henry Stevens

Was the Schappeller device an engine used in German field propulsion saucers? This is a possibility. Because so little has been reported about this device in the English language, the following is a report describing Karl Schappeller and his device in some detail.

Karl Schappeller (1875-1947) literally went from being born in poor-house to owning a castle during his lifetime. His economic success was mirrored in his experiments in energy as a lay-scientist, culminating in the invention of a free-energy device which attracted considerable attention around 1930. Schappeller made no secret of his invention and actively sought private financing to manufacture and distribute the results of his research. He was in touch with financial concerns and he even spoke with a representative of the British Admiralty concerning the utilisation of his device to power Royal Navy's ships (1).

At this time, 1930, the device was somehow appropriated and further worked upon by a governmental organisation of the German Weimar Republic, the Reichsarbeitsgemeinschaft or Reich Works Association (RAG). At least one aim of the RAG was to make Germany self-sufficient in energy production. Specifically, they published their intentions to utilise many Schappeller devices in a system of broadcast energy distribution throughout Germany which would result in the entire elimination of the electrical grid (2). As we know, Adolf Hitler assumed power three years later and for strategic reasons, he was also very interested in making Germany independent of foreign sources of energy. It is known that political and scientific structures were set up to work on the energy problem as evidenced later by the synthesising of gasoline and oil products from coal by the 3rd Reich. One of these political and scientific structures was contained within the SS and it is known that Karl Schappeller actually met with SS Reichsfueher Heinrich Himmler in Vienna in 1933 (3).



Left: Inventor Karl Schappeller Right: Karl Schappeller’s Device. **A**. Steel outer casing. **B**. Special ceramic

lining in which tubes are embedded. **C**. Hollow centre, filled by glowing magnetism when in operation. **D**. Tubes, circuit and earthling.

Fortunately, there are good descriptions of the Schappeller device upon which to draw in both German and English. Per Vril-Mythos is a complete discussion of Schappeller, his device, the history and the controversy surrounding it. "Vril, Die Kosmische Urkraft Wiedergeburt von Atlantis" and "Weltdynamismus Streifzuege durch technisches Neuland an Hand von biologischen Symbolen" represent an attempt by the RAG to popularise their ideas in booklet form. Finally, British electrical and mechanical engineer, Cyril W. Davson, visited Karl Schappeller in Austria and spent three years learning of his device and his theory before the Second World War. Davson's descriptive book, "The Physics of the Primary State of Matter", was written in 1955, after the war and the death of Schappeller.

Before describing the device itself it should be understood that Schappeller and all writing about his device believe that the energy-source being tapped is aether energy, sometimes called "Raumkraft" or "Raumenergie", that is space-energy (4)(5)(6). This device was also said to be capable of, perhaps with some tuning, emitting aether as a radiant energy (7). The physics of aether energy is described by Davson as "primary physics" as opposed to "conventional physics" which he believed could only be considered to be a secondary, derivative understanding.

Aether Theory

For readers who have never heard of "aether", perhaps the simplest explanation for aether physics is that of the late Dr. Hans A. Nieper (7) entitled "Revolution in Technology, Medicine and Society". Aether could be thought of as an energy source emanating from everywhere equally at once. The universe could be considered, as is often said, to be "a sea of energy". It forms a background of energy everywhere, and since it is everywhere all the time, it is difficult make independent measurement of it. This aether energy is in constant motion. All energy is radiant energy, according to this theory. This can easily be appreciated as to electromagnetic radiation but it is also true of that very elusive thing called gravity. Newton described the effects of gravity but he never told us exactly what it was. Dr. Nieper tells us that gravity is really a push, and not a pull. Gravity is acceleration and is caused by the aether field. Again, all energy is radiant-energy whose fundamental basis is aether radiation.

From the aforementioned book by Dr. Nieper:

In addition, Nieper established the axiom that, "all natural accelerations can be attributed to a single unified basic principle, namely, the interception (or braking) of a field energy penetrating from the outside (gravity acceleration, magnetic, electromagnetic, electrostatic and radiesthesic acceleration)".

In trying to explain aether, it might be thought of as an all-pervasive liquid, occupying all of space. This liquid concept is useful because a liquid can not be compressed but can only transfer the energy attempting to compress it from one location to another. This is how the brakes of a vehicle work. The driver presses the brake pedal when he wants to slow down or stop. The plunger of the brake pedal attempts to compress the liquid in the master cylinder. The master cylinder is connected to each wheel by metal tubes full of liquid. When pressure is put on the master cylinder by the driver it is transmitted to each of the four wheel cylinders full of the same fluid which transmits the force, moving the brake mechanism, slowing the wheels of the vehicle.

In a similar way, the aether serves to transmit energy through this "non-compressible" quality. In a primary electric coil and secondary electric coil, for instance, induction in the secondary does not take place directly from the primary as is now said by current physics, but instead, the induction between the two windings is due to the aether field. This concept of the energy transfer function of the aether field is also expressed by Davson.

Using this perspective, that all energy is radiation, the braking of aether radiation, that is the slowing down or stopping of this radiation, can cause a transfer to other forms of energy. The word "energy" means the entire electromagnetic spectrum. That includes, electric, magnetic and electrostatic fields. This means heat. This also means gravity. Again, gravity is the primary radiation of the aether field. It radiates from every point in the universe equally.

This concept seems ridiculous until it is given some thought. One might ask: "How can gravity be a push when we know better?" After all, things fall to earth, don't they? The answer is that the effects which we feel and call "gravity" are due to aether shielding. Aether radiation can be braked, that is slowed down and absorbed by mass. It is then re-radiated or converted into mass. It is re-emitted as slower aether radiation

or even as heat. Some of it can, and is, converted into mass inside a planet. If there is a loss of aether radiation, then there is shielding. Thus, a planet would shield from this radiation in one direction. That direction is always toward its centre which is the direction of greatest mass and that is what we describe as "down". This is simply the area which contains the maximum amount of shielding. In all other directions the aether radiation continues to exert its push on us. The area of minimum shielding is directly opposite the area of maximum shielding, so things fall (or more correctly are accelerated or "pushed") towards the earth.

Think about this for a minute. Being in deep space is rather like being underwater. Underwater, the pressure at all points is so similar that we feel weightless. We are weightless in deep space because the aether field exerts a push on us from all directions equally. In space, the nearer one gets to a large body the stronger the push is from the opposite direction since the body shields or converts the aether radiation. The result of this thinking is a mechanism totally different from "gravity" as we know it but appearing as exactly the same observed phenomena.

The beauty of this aether theory of gravity is that gravity functions like every other form of radiation. Its underlying cause, aether radiation, can be converted to mass or, in certain circumstances, re-radiated or converted to other forms of energy. No Unified Field Theory is necessary. The aether field **is** the unified field. Further, there is no need to look for something separate called "anti-gravity". If gravity is a push then it is all anti-gravity. All we have to do to make a UFO, is to find this particular gravity frequency and find out how to generate it.

Aether physics was a lost physics. Physics was hijacked early in the 20th Century by the alleged results of the Michelson-Morley experiment. This experiment assumed that "aether" was matter. There is some confusion here. We know now that particles moving near the speed of light are measured as waves, that is energy, rather than as matter. Nevertheless, aether theory has been discredited among physicists who, in turn, discredit others who raise the subject. It is only through the efforts of "free-energy devices" and free-energy researchers, that this knowledge is being returned to us. Without this aether theory, the reason these devices work cannot be explained at all. Rejection of aether theory allows these devices to be dismissed as "theoretically impossible" and so “fraudulent” by implication. They are marginalised and dismissed as "perpetual-motion devices". According to established physics, perpetual-motion devices violate the physical laws of conservation of energy. Without an aether theory as an explanation, they do violate the laws of conservation of energy and so their detractors are able to simply dismiss them out of hand. The simple fact that some of these free-energy devices actually work, does not seem to bother these scientists in the least. Rather than change the theory to accommodate the observed facts, the facts are ignored and substituted by dogma. Whether we like it or not, we are living in an energy Dark Age.

Instead of aether theory, we have all been led to focus upon Einstein and his Theories of Relativity. Two or three generations of scientists have wasted themselves on "trying to prove Einstein right". This misguided thinking has resulted in stagnation. One need go no further than the many "free-energy" devices which have arisen to the level of notice in spite of accepted scientific theory to see that this statement is true.

Needless to say, German scientists of the Nazi period laboured under no such illusions. They never abandoned aether physics. This was the fundamental reason why field propulsion UFOs were first developed in Germany. After the Second World War two different sciences developed, both called "Physics". One was the relativity-based concept taught in schools, while the second, more esoteric type, was used secretly, by the secret government, for deep “black projects”.

Structure of the Schappeller Device

According to Davson's description, upon which we will rely, the Schappeller device is really composed of two separate units, the rotor and the stator. The stator is constructed as follows: Its surface is round or ball-shaped, being composed of two half-shells of steel. These half-shells contain the internal structure and are airtight. Attached at the "pole" of each half-shell is an iron bar-magnet, most of which is positioned inside the sphere. This means that the bulk of each magnet is inside the steel ball, one opposite the other. There is a space between the two bar magnets at the very centre of the sphere.

An Insulating ceramic material, is placed on the inside of the steel ball, leaving a hollow central area. Within this hollow area, and around the space between the magnets, are two coils are mounted. These start at the pole of the bar magnet and finish at the centre of the sphere, with a connection leading out of the sphere to the rotor. These coils are wound using a hollow copper tube filled with a special, secret substance called the "electret". Upon leaving the sphere, the electret-filled copper tubes are replaced by conventional copper wire. An electrical connection is made from the outside surface of one pole to one pole of a special type of

battery which is grounded at the other pole or, alternatively, to a special device called an "Ur-machine" which will be discussed later.

This electret is a permanent magnet within the sphere. This type of magnetism is not identical with ferro-magnetism or electromagnetism, it is much stronger (8). The actual composition of Schappeller's electret remains a secret, but another electret has been made by Professor Mototaro Eguchi. It consists of carnauba wax and resin, perhaps also containing some beeswax. It was kept in a strong electric field while baking slowly until it solidified. For purposes of the production of Schappeller spheres, a complete electret manufacturing plant would have to be set up, which has no parallel in present science (9).

Before being set into operation, all the air is pumped out of the hollow core of the sphere. This whole ball is mounted on a swivel mechanism so that the poles can be moved from the vertical to the horizontal. The stator is not attached to the rotor. The stator can function without the rotor and the stator is capable of producing electrical energy without the rotor. The rotor could also be used to generate additional electrical energy.

The rotor consists of a steel wheel of special design fixed on the shaft to be driven, and surrounded on its outer surface by magnets which are attracted and repelled by the force of the stator. The copper wire attached to the internal electret-filled copper tubing, runs through this wheel and supplies electric power to the magnets. The magnets are hollow and filled with the same electret. There is always an odd number of magnets.

A variant of this rotor comes to us from Taeufer, who refers to this further development as the "Ur-Machine". This machine is composed of six sphere units as described above, five revolving around a sixth one which is positioned either above or below the plane of the other revolving spheres. A seventh unit would be employed to rotate the five rotating spheres and so would be offset, and not attached to the others. The five rotating spheres would charge the sixth stationary sphere. The sixth and seventh spheres would function as an anode and cathode and so ground the unit. The Ur-machine could be used to activate other spheres instead of a battery-earthing procedure (10).

As a prime mover, an engine, the rotor would be used to turn a drive shaft. The stator would be offset, that is, moved off centre in relation to the rotor. Schappeller worked out various angles of efficiency (11). The drive shaft could be used to power any number of machine applications such as, for instance, the propellers of a ship.

Means of Operation

The device is started through a connection to a totally unique battery and a connection to the earth (12). A specific excitation impulse must be given to the device (13). This electric impulse is conducted through the iron magnet and jumps the gap in the centre of the sphere to the other iron magnet.

What occurred then sets this device apart from all others. In the vacuum of the sphere, in the centre space between the two bar magnets a field of "glowing magnetism" is set up. This glowing magnetism is something entirely unique. It is recognised as a magnetic field, but much more powerful and unlike any magnetic field produced by an iron bar or an electric coil. Once the initial input had been made to start the device, the battery and ground can be disconnected. The device then continues to operate on its own (14).
For an understanding of what is really happening here we have to consider the bar magnet. We think of a bar of iron with two poles, one positive and one negative or perhaps one north pole and one south pole. But there are really three components to the bar magnet. There are the two poles and the neutral zone between the poles. If we cut the magnet in half we get two new poles. For the Schappeller device, this neutral zone is very important. Imagine a bar magnet running through the vertical axis of the ball. Then imagine the centre section cut out. We now have a north pole at the top of the ball, a south pole at the bottom of the ball just as we do with the planet Earth. In the centre we have a missing section with a south pole, opposing the north pole at the top of the ball and, likewise, a north pole opposite the south pole at the bottom of the ball. We have now four poles and a split bar magnet with a gap in its centre section.

It is this gap in the centre where Schappeller's "glowing magnetism" is generated by grounding, that is, charging the device via a special battery and an earth connection. This glowing magnetism is the mystery. Davson cites Schappeller's calculations and gives this form of magnetism as being a thousand times more powerful than that produced by present magnetism (15). He also states that in this form of magnetism the electricity is stationary while the magnetism is radiated (16).

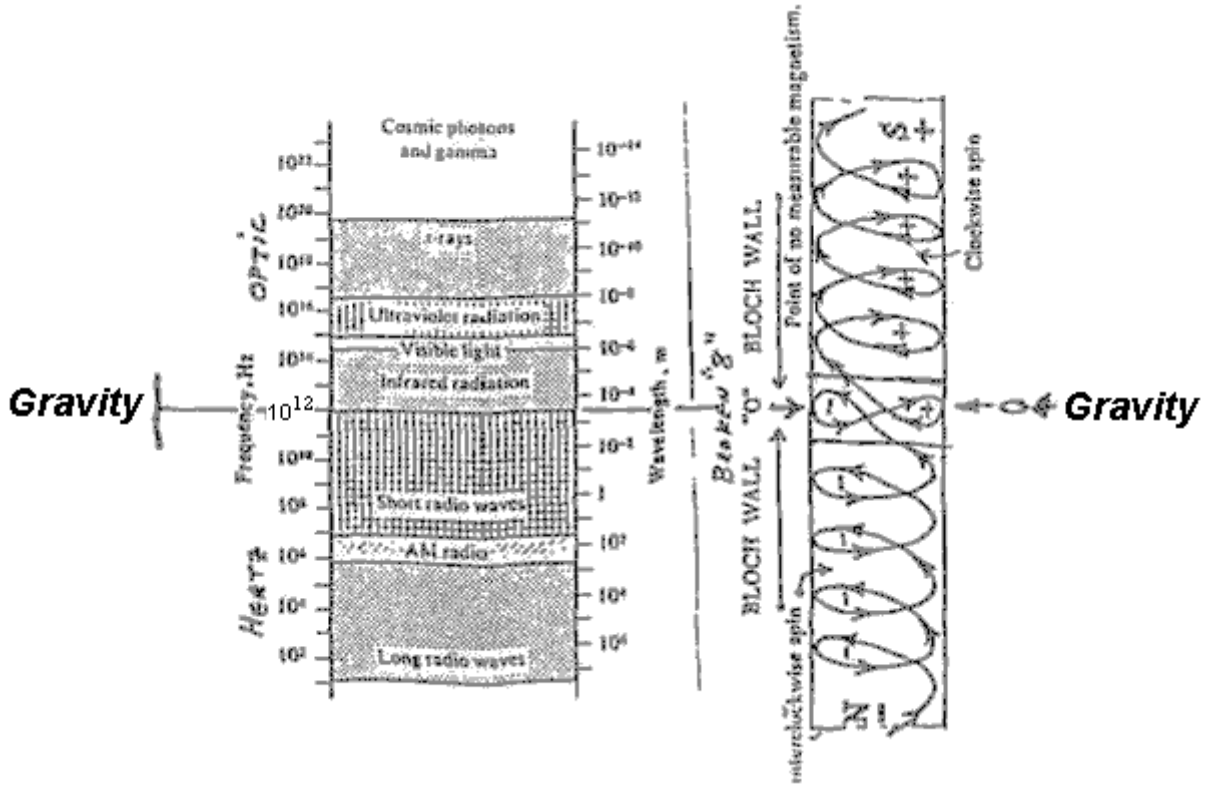
To state this again, Davson contends throughout his book that this glowing magnetism is not found in secondary physics, that is, in modern physics, and that this glowing magnetism is a manifestation of primary physics. As a phenomenon of primary physics, it is responsible for, and can generate, heat, electricity and magnetism.

After initial stimulation and in a state of glowing magnetism, no further input of energy is needed from the battery. The device is able to draw in energy to it directly from the surrounding aether, bind this energy though its magnetic electret material, that is the filling in the hollow copper coils of the internal coil, and then re-radiate energy producing heat, electricity, magnetism or mechanical work depending upon the application.

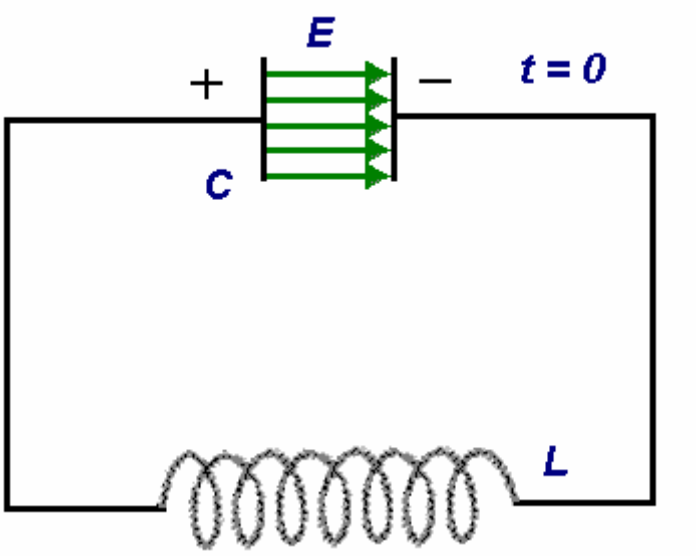
Stated another way, this is an implosion device and it is described as such (17) (18). Unlike the Schauberger device which is associated with the word implosion, the Schappeller device operates purely at the energetic level. Energy is drawn towards the centre, through the magnets, into the field of glowing magnetism, and then radiated outward.

My first explanation for this output of radiant energy involves the concept of the Bloch Wall. A Bloch Wall is defined by Van Norstrand's Scientific Encyclopedia, 1958 edition, pages 201 and 202, as: "This is a transition layer between adjacent ferromagnetic domains magnetised in different directions. The wall has a finite thickness of a few hundred lattice constants, as it is energetically preferable for the spin directions to change slowly from one orientation to another, going through the wall rather than to have an abrupt discontinuity" (18).

In electromagnetics the Bloch Wall is external to the hardware itself. It is the point of division of the circling vortex, or spin, of the electronic magnetic energies of the north and south poles. The negative north pole magnetism spins to the left while the positive south pole spins to the right. Energy is being conducted into the Schappeller device through the un-insulated poles and being conducted and spun on its way to the centre of the unit. The point of zero magnetism, no spin and magnetic reversal, where the two spin fields join, is the Bloch Wall (19).



Bloch Wall, a gravity wave source as a function of the electromagnetic spectrum?
(Dr. Richard Le Fors Clark)



Oscillating circuit. Charged capacitor (electric field) discharges, current carried through insulated wire to charge coil (magnetic field) which discharges, charging capacitor. Oscillating electric and magnetic fields yield electromagnetic waves.

The Block Wall radiates energy. Remember, if energy is coming in then it must be going out. The Bloch Wall may generate radio, radar and other electromagnet frequencies but what is most interesting is that it is actually able to radiate gravity as according to Dr. Richard Le Fores Clark. According to this interpretation, the conjunction of two dipolar generated force-field vectors, a quadropole force field or gravity is generated according to Dr. Clark. Gravity being a quadropole source, it radiates in a circular, 360^{O}, pattern of two cycles. Dr. Clark has fixed the point of emission as below that of radar and above infrared at 10^{12} Hz (20). Dr. Richard Le Fors Clark believes that gravity is a radiation (20) and so it is a "push".

Another Opinion

In late 2001, I wrote a magazine article on the Schappeller device (20) which contained most of the material described above. In that article, I requested alternative explanations for the Schappeller device. I received a letter from Mr. Michael Watson, BSc, Charted Physicist and Member of the Institute of Physics in the United Kingdom. But there was something in Mr. Watson's background even more impressive than his professional credentials. Cyril W. Davson was a family friend whom Mr. Watson knew well in his youth and with whom he had discussed Schappeller and his ideas at some length on many occasions. In Mr. Watson's letter was a brief summary of Schappeller's theory in which he cut through most of the confusing terminology.

This summary is important for a couple of reasons. Mr. Watson's summary of Schappeller's aether theory as described by Davson dovetails nicely into the ideas of Schauberger yet seems to allow for Tesla's experimental results on aether as explained by Bill Lyne. The following is what I learned from Mr. Watson's letter:

Most of us have heard of the two Laws of Thermodynamics. These are laws of heat. The First Law of Thermodynamics states that energy is conserved, meaning that the total amount of energy in the universe always remains the same. This is no surprise for most of us and it is not the real concern here.

What is of concern is the Second Law of Thermodynamics which discusses heat and entropy. The word “entropy” might be thought of as a state of randomness or chaos. Negative entropy would then mean movement toward a less random or more ordered state of any particular thing. If we apply this to a system, then entropy tends to increase until the system breaks down in utter chaos. This will occur unless the system is re-charged with additional outside energy. A concrete example of this might be helpful:

Imagine a new car just coming off an assembly line. It has taken a great deal of energy to find, refine, forge, weld, and paint the metal parts of this car. This same concept also applies to all the other components of the car. This energy and organisation constitute a highly organised state, or, in other words, a state of negative entropy.

What happens next illustrates entropy. The car is purchased. Whether it is driven hard or just sits in the

garage does not matter in the long run because what happens to the car is that it starts to fall apart. This change may be small at first and may only occur at the molecular level, but it occurs nevertheless. The engine, transmission, paint, rubber, electronics, etc. all will fail with time. Even it the car just sits in the garage, in a thousand years the metal will eventually oxidise. Finally, the car rusts away forming a reddish brown heap. This is exactly the opposite of the organisation and energy used to put the car together. This disorganisation is entropy. The only thing which will reverse this, as we all know, are additional inputs of energy by the owner in the form of maintenance and repairs.

All things in a relative state of order move toward a state of disorder. In terms of heat, heat will always flow into a colder place from a warmer place. When something is heated there is a rise in its entropy. With increasing heat its molecules move faster and faster in random chaos, just as a bomb does when it explodes. Increasing heat means increasing randomness and chaos which is entropy. Cold, then, can be seen in terms of negative entropy. Any cold object is simply more organised and less random than the same object once it is heated.

Schappeller had something to say about the Second Law of Thermodynamics. He said there was another and unknown thermodynamic cycle which runs opposite the Second Law. To name this idea we will call it "Reverse Thermodynamics". It is the reverse of the Second Law of Thermodynamics in that it leads to an increase in entropy. Not only is there an increase in order but there is an increase in cold! Schappeller, according to Mr. Watson's letter, built his spherical device primarily to demonstrate the principles behind this Reverse Thermodynamics. It was not designed as a practical machine.

To demonstrate the difference between the Second Law of Thermodynamics and Reverse Thermodynamics two theoretical machines shall be examined. Actually, a machine running according to the Second Law of Thermodynamics is not theoretical at all. Combustion machines are of this type. For simplicity sake we will use a wood burning stove such as the type invented by Benjamin Franklin for the heating of a house.

Wood is put in a hollow iron vessel with an adjustable hole at one end. The adjustable hole admits atmospheric oxygen. An initial small input of heat is added to the wood and oxygen until burning occurs. A great deal of heat is produced once the wood begins to burn. We know heat expands. Carbon, carbon dioxide and water vapour are also produced as by-products of the combustion. Entropy is increased. Since entropy is increased, so is pollution so perhaps we all can agree that this is a good example of the destructive technology so characteristic of the world in which we live.

In our example of a theoretical Reverse Thermodynamic machine the by-products of the previous example can be used as fuel. But Schappeller's machine has the additional property of being creative, that is, negatively entropic. Schappeller believed this creative process to be individualistic, so we need a specific template to use as a pattern for this creation. Heat, water, and carbon dioxide are fed into this machine. Quite amazingly, oxygen is yielded as a by-product of this reaction! The heat is also absorbed in Schappeller's Reverse Thermodynamic machine! This absorption of heat is another way of saying that the machine is implosive in nature rather than expansive or explosive as was the heat producing machine. What is most amazing, however, is that entropy is actually reduced yielding, something which has been created - wood!

Actually, this machine is not theoretical either. It exists and works as we speak. These machines are all around us. We call these machine "life". In this case our machine is a tree. In the tree, energy, sunlight, is absorbed and combined in a cold process with water and carbon dioxide to form wood. The template used as a pattern for this seemingly intelligent, creative, process is simply a seed. In this type of reaction the "cold" force is something other than the absence of heat. This cold is an active cold. It is a "densifying", implosive cold. It is a life-giving cold. This is a cold, life giving force. To quote Watson: "This process is life force and the reverse of the second law of thermodynamics; it is the vital force: Vril."

This is one huge difference between the physics of Schappeller and Schauberger and the physics of the Nineteenth Century. The physics of the Nineteenth Century explain everything in terms of the inanimate. Laws of physics are written using inanimate examples. Chemical reactions are described which stem from inanimate models. Animate models are simply made to conform with the inanimate assuming that life is just a special case which eventually will be shown to be nothing but chemistry and so subject to the same Second Thermodynamic Law as the inanimate. Schappeller and Schauberger both say in their own ways that this is not so. They say, each in their own ways, that a new and different law of thermodynamics applies to living forces. They say that this more akin to a life process than previous theories allow. They say this force is creative. Those who subscribed to these new ideas claimed that it was not only a new physical law but a new science and that Germany would lead the way in this new science. Let us take a closer look at what is claimed to be the physics behind this new science.

The first concept to be considered is cold. Cold in this sense does not mean the mere absence of heat. This is interstellar cold, the cold found in the vacuum of space. In this relative vacuum, matter is not found in sufficient quantity to be used to measure this cold. Think about how we measure cold. We measure matter which is cold. We measure the heat in air or water for instance. In the absence of matter how would cold be measured? There is no doubt that if we could, for instance, place a thermometer in a glass of water in deep space, the temperature recorded would be at or very near absolute zero, 0^{O} Kelvin or -273^{O} Centigrade or -460^{O} Fahrenheit.

The presence or absence of matter in deep space may be the subject of conjecture. The presence or absence of energy in deep space is something universally accepted. For instance, we all know that light passes through interstellar space. We see the proof when we look up at the stars, planets or the moon. Besides visible light, other electromagnetic radiations freely pass through space. These include x-ray, gamma and cosmic rays. Yet besides electromagnetic radiation many people now believe that in the depths of space there resides another form of energy with is found there as well as everywhere else all around us. This energy sometime goes by the name of "zero-point energy" but for our purposes we can simply call it "aether energy". It is sometimes argued that this energy is really the result of aether rather than the aether itself and that aether really is matter. For a moment, let us postpone this discussion and focus on the vast, stretches of interstellar space which are filled with aether energy, near or at absolute zero.

Mr. Watson points out Dawson's words on page 83 of “The Physics Of The Primary State Of Matter” where he says: "Cold is not therefore the absence of heat, primary heat and cold having nothing to do with molecular action (in the cosmos) there are no molecules available".

The reader may recall that something strange happens to electrical energy at absolute zero. For instance, if a disc of conducting material is held at absolute zero and the disc is given an electric charge, the electric current will circulate around and around the disc forever, never losing its energy as it would if the disc were sitting on an office desk at room temperature. This property of cold is instrumental in the storage of at least one form of energy. The vast stretches of cold interstellar vacuum must be seen as a vast energy storage sea in a state of heightened negative entropy. Schappeller called this undirected matter-energy reserve potential "latent magnetism”. Out of this latent magnetism, both energy and matter could be produced with the corresponding stimulation. The non-excited electromagnetic field was viewed by Schappeller as simply latent magnetism. Matter is a condensation out of bipolar aether. Therefore, electromagnetism is a product of matter and is nothing more than bipolar aether in a different condition. Latent magnetism could be, then, excited into matter. Latent magnetism could be influenced by either of the thermodynamic principles discussed, the Second Law of Thermodynamics or by Reverse Thermodynamics. This vast aether field, whose most notable characteristic is the property of cold, latent and awaiting stimulus, is the predecessor of both energy and matter as we know them.

Since primary cold, this vast reserve of negative entropy potential, is responsible for both matter and energy and since all energy eventually degenerates into heat, it follows that, as Davson puts it, again on page 83: "Primary heat, as may now be understood, is composed of cold energy". This is seems like a surprising play on words, especially from a man of science, nevertheless, this statement follows perfectly from Schappeller's reasoning.

We turn now to Schappeller's concept of "stress". Both heat stress and cold stress can be applied to an electromagnetic field. Heat stress is the usual type of stress applied to electromagnetic fields in secondary physics. Secondary physics is the physics of our everyday world according to Schappeller. Primary physics is the physics dealing with the cold force and aether yielding matter and energy, which constitute the secondary reactions and so Schappeller uses the term "secondary physics" to describe our world as we know it.

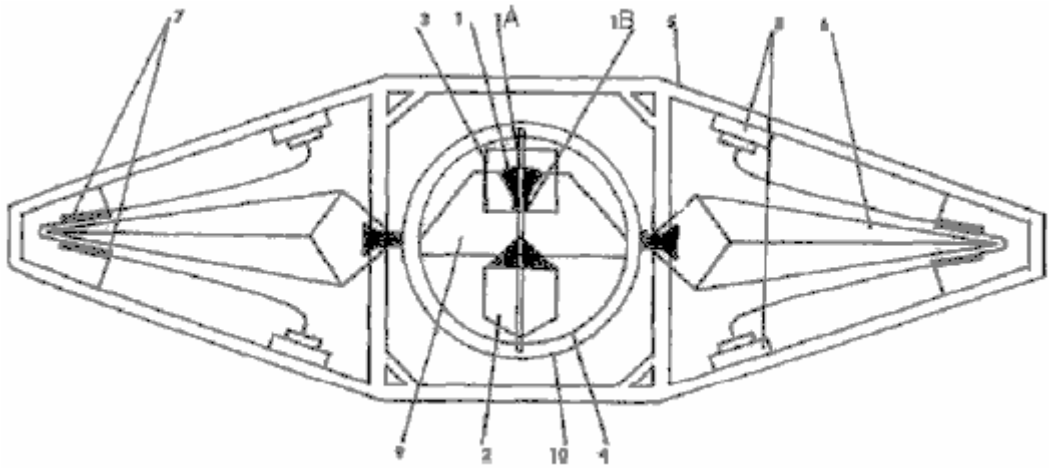
An example of heat stressing of the electromagnetic field is the capacitor and the coil. A charged capacitor produces an electric field and a charged coil produces a magnetic field. A charged capacitor and coil, connected by a wire circuit alternately charge and discharge each other, producing electromagnetic radiation unit the heat caused by the resistance of the wire degrades the whole process into heat. Heat stress on the electromagnetism is +/-.

Cold stress on the electromagnetic field is something totally new to our science and technology. It is also seen in terms of +/- but the machines used to produce it are not known in our world. Mr. Watson did not say this but if we return to our examples of heat stressed machines, the capacitor and the coil, the corresponding cold stressed machines might be the Schappeller sphere and the Schappeller coil electret. The sphere collects the charge through the magnets, holds and condenses it in its glowing centre corresponding to the

electric field of the capacitor. The internal coils filled with electret produce a magnetic field in the presence of the intense and pulsing electric field. According to my interpretation, the whole Schappeller sphere is a combined capacitor/coil combined into one machine made possible through an initial input of cold stress.

As in our example of the capacitor/coil interaction producing an electromagnetic wave, so an attraction exists between a machine obeying the Second Law of Thermodynamics and one obeying the law of Reverse Thermodynamics. This attraction can lead to interaction. For example, an imploding or centripetal vortex can couple with an exploding or centrifugal vortex. The centripetal vortex is an example of a system following the law of Reverse Thermodynamics while the centrifugal vortex represents system following the Second Law of Thermodynamics. We have all seen these two systems working together in everyday life. The common toilet is such a machine although the centrifugal side forms inside the drain pipe which is out of sight.

Perhaps there is another example which is more germane to our discussion. It is the diagram of the Vril power plant. (This engine diagram is used here as an example for discussion and is not a blind endorsement of the diagram's existence or accuracy.)



Internal plans for a "Vril-1" saucer, according to Polish historian Igor Witkowski.

In this interpretation of this diagram, we are really dealing with two separate devices. First, is the central spherical device which may be a refined version of the Schappeller sphere. An initial charge would be imputed into the sphere to start it after which the unit would continue to gather up the surrounding energy. This is a Reverse Thermodynamic machine. The sphere generates a magnetic field which could be offset by rotating the Schappeller device. The offset field would feed and so rotate the arms of the electric generator surrounding the sphere. The electric generator would gather electric energy, feeding the four large broadcasting fixtures on the walls of the saucer. These fixtures might be, for instance, Tesla pancake coils. The electric generator is an example of a machine complying with the Second Thermodynamic Law.

Both components of the power plant are bonded together in a single system since the output energy of the broadcasting fixtures on the walls of the saucer constitute additional input energy for the sphere. The two components attract one another and use and depend upon one another as they circulate and recirculate energy. As the energy level of one component increases so does the energy level of the other. Indeed, the biggest problem facing the use of such an engine may be getting some means of stopping it.

The actual levitation might be the particular electromagnetic radiation coming out of the sphere. In this interpretation, the broadcast fixtures are used to steer the saucer. Davson gives output frequencies for the sphere as 10^6 (20).

Mr. Watson points out in his letter that one reason machines utilising the Reverse Thermodynamic principle have not been recognised is that a cold stressed magnetic field is a cold machine. Even a centripetal vortex cools rather than heats. All our devices of measurement ultimately measure heat in some form. Measurement of cold is more difficult. The example already given, the problem of measuring temperature in interstellar space in the absence of matter is an example of this problem.

Finally, the reader will recall that Mr. Watson points out that electromagnetism itself manifests bipolarity, yielding four components in all. These are +/- hot electromagnetism and +/- cold electromagnetism. The reader will recall that two hot electromagnetic components can be joined (the capacitor and the coil) and set

into a cycle producing an electromagnetic wave. Is it possible that two complementary hot electromagnetic and cold electromagnetic machines could be set into cycle producing not a bi-polar but a quadropolar, 360^{O} radiation to produce gravity, such as the one described by Dr. Richard Le Fores Clark?

Planned Uses for the Schappeller Device

If the above discussion has any meaning at all in the quest for an answer to the UFO question, one use for which the Schappeller device must have been destined was that of a power plant for a flying machine. Was this so? The Schappeller device had many planned uses. In 1930 this device was planned as a source of broadcast energy, reminiscent of Tesla, for both German homes and industry. The device could also be used as a generator, battery, transformer, or antenna (21). It is reported that toward the end of the war the SS researched the possibility of using this device in the form of a death ray (22). But additionally, and in answer to our question, the Schappeller device was envisioned as a levitation device for a flying machine. Here is some of that discussion from our sources:

"The new dynamic technology will, in the future, be able to drive electric locomotives and cars without the manufacture of costly armatures and everywhere through connection to the atmospheric voltage network. Hypothetically, is certainly the installation of a sufficient number of central amplification facilities which transports from the Ur-Machine the specific magnetic impulse from the dynamic spherical element. New types of aircraft with magneto-static power devices and steering, which are completely crash and collision proof, could be built for a fraction of the cost of today's aircraft and without the lengthy training of everyone who will be servicing these aircraft". (From "Vril Die Kosmische Urkraft Wiedergeburt von Atlantis" by Johannes Taeufer, page 48).

"Our goal must be to drive forward the space ship problem to new understanding and realisation! Here a definite postulate can be established: "A spherical space ship with its own atmosphere" also technical creation of small planets with world dynamic propulsion and buoyancy!. Will this be possible? -- Major powers in the world prepare themselves in any case presently, especially in Germany."

The above from "Weltdynamismus Streifzuege durch technisches Neuland an Hand von biologischen Symbolen" pages 11 and 12. Please note the use of the words "spherical space ship"(Kugelraumschiff).

From Davson's "The Physics Of The Primary State Of Matter", page 240: "The Rotor is laminated to prevent eddying and the magnets do not project; the Rotor periphery is thus entirely equi-radial. The Rotor is fixed to the shaft to be driven and the Stator is fixed about a metre above the earth's surface. The latter is, of course, flexible because the earth can include the sea or even the floor of an aether-ship."

From Davson, page 199: "As has already been explained, the new Technique will not concern itself with the air as a supporting medium, but directly with the aether. Therefore, the body may be a vertical sealed cylinder with conic ends or any other suitable form. Such a body is obviously rigid and inelastic, and it must contain an aether stress of sufficient intensity to support its mass against thither stress of the earth's stress field, which means that the glowing magnetism core in the Stator, provided in the body to be lifted, must be able to vary its intensity according to the height at which the aether-ship is to be raised and supported whilst in transit, as the aether stress or field, itself, varies inversely as the square of the distance from the earth's surface. The actual design and solution of all the various problems in the production of such ships, the choice of methods of propulsion, whatever independent or directional, belong to the new Technique, whereas here we are only interested in the principle as applied to the problem of Gravitation."

Finally, from Davson, page 177: "Now the reason that an unsupported body falls to the ground is primarily because it has "no hold" on the medium. It was previously explained that any inert mass or body has only a latent stress field which functions merely as the force of cohesion and has no mobility and thus only a latent internal stress field and no external stress field. This means that it has no "hold" on any elastic medium such as the aether or the air, therefore it must fall, and it falls towards the greater inductive energy.

If the inductive energy, through some exterior cause, could be made suddenly to increase enormously, there would come a point when the body would be supported, or rather suspended, before it reached the earth's surface. The new Technique could accomplish this by placing a Schappeller Stator in the body in question, where the body is suitably constructed, thus setting up a glowing magnetic stress field which would hold or keep the weight or mass of the unit body suspended, not in the air "the stress field would have no reaction on the air" but only on the earth's magnetic stress field. This is the basis of the new principle for 'aether ships'".

Employing the Schappeller mechanism is only half the total explanation. In a field propulsion saucer there are possible two types of "drive" needed. The first is the "Auftrieb" or levitation. Employment of levitation makes the craft buoyant. It weighs nothing. If it weighs nothing it can be moved very easily. "Antrieb", impulse or motive power is the second drive involved. It moves the craft directionally. Levitation only would be supplied by the Schappeller system. Directional movement is so far best explained, in my mind, using the Tesla pancake coils as explained by Bill Lyne.

Concluding Thoughts on the Schappeller Device

In the end, what can be said of the Schappeller device? Certainly, it did exist. It drew attention and funding from people within the German government of the time. It was studied by a qualified outsider, a British engineer, for a period of three years and was judged to be genuine.

However, there are some obvious problems. Exotic energies have been evoked which have not been explained satisfactorily. Therefore, the facts are not yet proven. Certainly more proof is required before the claims made for this device or the energies involved can be wholly accepted. For the time being we must put this discussion aside, awaiting further correlations.

There are some solutions connected with this device also. If we accept the idea that both the Schauberger and the Schappeller devices worked on the theory of implosion, then one explanation will serve to explain them both. It also allows for an aether-as-matter explanation. This may fit into the evidence gathered by Nikola Tesla. The commonality of these devices could then be sought and perhaps a more efficient device built as a result. We will pick up this theme again in the discussion section of this book.

It should be pointed out that the quest for this "new science" is not specific to Schappeller or Schauberger. Mr. Watson passed on these words from Ehrenfried Pfeiffer, a scientist who collaborated with Dr. Rudolf Steiner around 1920. Although he is not happy with the translation, he sent it as he found it which is as it is presented here:

"...the method of science, in a materialistic sense, is based on analysis splitting apart, disintegration, separation, dissecting and all the procedures which have to destroy and take apart, to work on the corpse rather than to grow, to develop, to synthesise. That the human mind was captured by these methods of braking apart: in that I saw the source of our present situation. My question (to Rudolf Steiner) was therefore: is it possible to find another force or energy in nature, which does not have in itself the objective of atomising and analysis but instead builds up, and synthesises? Would we discover that constructive force, which makes things alive and grow, develop adequate building up of methods investigation, eventually use this force for another type of technique, applied to drive machines, than because of the inner nature of this force or energy we might be able to create another technology, social structure, constructive thinking of man rather than destructive thinking? This force must have the impulse of life, of organisation within itself as the so-called physical energies have the splitting, separating trend within themselves."

My question to Rudolf Steiner in October 1920 and spring 1921 therefore was: "Does such a force or source of energy exist? Can it be demonstrated? Could an altruistic technology be build upon it?"

My questions were answered as follows: "Yes, such a force exists, but is not yet discovered. It is what is generally known the aether (not the physical aether) but the force which makes things grow, lives for instance in the seed as Samenkraft. Before you can work with this force you must demonstrate its presence. As we have reagents in chemistry, so you must find a reagent for the aetheric force. It is also called formative aetheric force because it is the force which relates the form, shape, pattern of a living thing - growth. You might try crystallisation processes to which organic substrata are added. It is possible then to develop machines, which react upon, and are driven by, this force. Rudolf Steiner then outlined the principles of the application of this force as source of a new energy..."

Since this quest for a new science with the accompanying new machines had a relatively long history in Germany, certainly pre- dating the 3rd Reich, it is almost certain that the Schappeller device or others built along a similar understanding were further developed during the Nazi period. What became of it after the war is not known. It can be assumed that this device did not escape the scrutiny of the numerous Allied intelligence units tasked with combing Germany for examples of German science. Perhaps someday a government report will be de-classified explaining all this as it was in the case of another free-energy machine, that being the Hans Coler device, which was declassified by the British in 1978 (23) and which worked, according to Mr. Watson, using the same principles of cold magnetism. Until that final reckoning comes, aspects of the Schappeller device will still remain a mystery. And until a more final reckoning comes,

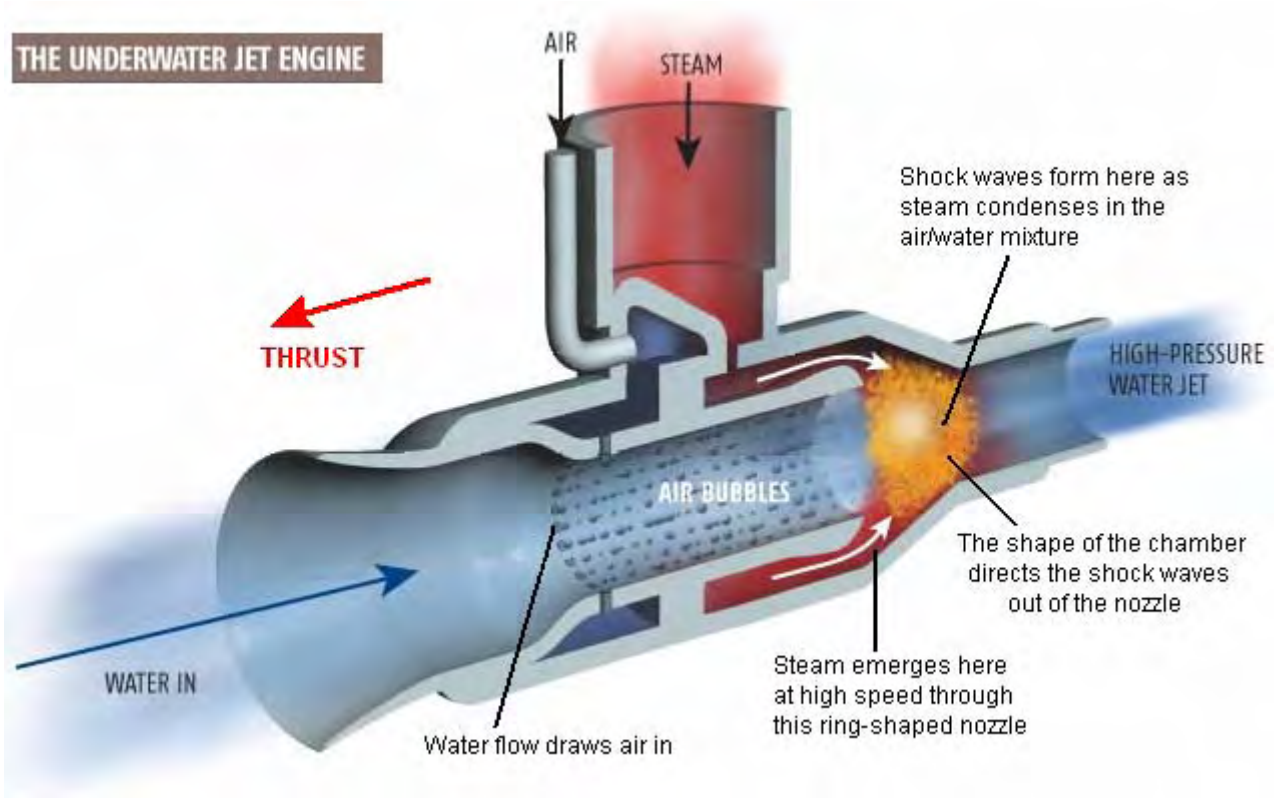
the question of whether or not the Schappeller device was used as a source of field propulsion in German flying saucers, must be deferred.

Sources and References

1. Bahn, Peter, Ph.D. and Heiner Gehring, 1997, pages 120-131, Per Vril-Mythos Eine geheimnisvolle Energieform in Esoterik,Technik und Therapie, Omega Verlag, Duesseldorf
2. Taeufer, Johannes, 1930, page 31."Vril" Die KosmischeUrkraft Wiedergeburt von Atlantis, commissioned and distributed by the Reichsarbeitsgemeinschaft "Das kommenden de Deutschland", Astrologischer Verlag Wilhelm Becker, Berlin-steglitz
3. Bahn/Gehring, 1997, page 131
4. Bahn/Gehring, 1997, pages 120-124, 130
5. Weltdynamismus Streifzuege durch technisches Neuland an Hand von bioloaischen Symbolen, 1930. pages 14-15, commissioned and distributed by the Reichsarbeitsgemeinschaft "Das Kommenden de Deutschland", Otto Wilhelm Barth Verlag, Berlin
6. Davson, Cyril W., 1955, pages 50-59, The Physics Of The Primary State Of Matter And Application Through the Primary Technique, Elverton Books, London
7. Nieper, Hans A., Ph.D., 1985, Conversion of Gravity Field Energy/Revolution in Technology. Medicine and Society. M.I.T. Management Interessengemeinschaft fuer Tachyonen-Geld-EnergyGmbH, Friedrlch-Rueder-Strasse 1, 2900 Oldenbuurg, Germany (available in German and English language versions)
8. Davson, Cyril W., 1955, pages 212-213
9. Davson, Cyril W., page 231
10. Davson, Cyril W., pages 217, 223
11. Taeufer, 1930, pages 30-32
12. Davson, 1955, page 230
13. Davson, 1955, page 226
14. Taeufer, 130, page 30
15. Taeufer, page 32
16. Davson, 1955, page 231
17. Davson, 1955, page 231
18. Davson, 1955, page 57
19. Taeufer, 1930, pages 38-40
20. Clark, Richard Le Fors, Ph.D., 1987, page 64, "The Earth Grid, Human Levitation And Gravity Anomalies", contained in Anti-Gravity And The World Grid edited by David Hatcher Childress, Adventures Unlimited Press, Stelle, Illinois
21. Stevens, Henry, 2001, "Infinite Energy", pages 9-13, Volume 7, Issue 40
22. Davson, 1955, page 244
23. Bahn/Gehring, 1997, page 115
24. British Intelligence Objectives Sub-Committee Final Report Number 1043, item number 31, "The Invention Of Hans Coler, Relating To An Alleged New Source Of Power, Bryanston Square, London

Condensation-Induced Water-Hammer.

There is another little-known effect which has a high potential for being a useful technique, and this is the water hammer effect produced by the sudden condensation of steam. Under suitable conditions, the effect can be harnessed to provide motive power.



One of the techniques which has been used is shown above. Here, steam under 55 psi pressure is forced into a tubular structure through a ring orifice. That steam then passes through a doughnut-shaped nozzle where it hits a stream of heavily aerated water. The bubbles in the stream of water are drawn in by the venturi effect of the water flow past small apertures in the tube. The result is a rapid series of water hammer shock waves which, because of the shape of the shock wave chamber, boost the water as it exits from the tube. This produces a thrust in the opposite direction, effectively forming a jet engine which is particularly suited to water-borne vessels. In the diagram shown above, the device shown is designated as an Underwater Jet Engine. The diagram is from the web display at:
http://www.newscientist.com/data/images/ns/cms/dn3321/dn3321-1_843.jpg and is copyright of the New Scientist.

Further information on this form of energy can be found at:
http://www.kirsner.org/pages/condInduceWatHamText.html and
http://www.energeticforum.com/renewable-energy/3093-condensation-induced-water-hammer

William Hyde's 10 Kilowatt, COP=10, Electrostatic Power Generator.
This is best described by his patent, a slightly re-worded version being shown here:

This patent describes a device which can be a little difficult to visualise and so some colour shading of parts has been used to help matters. Essentially, it is two circular rotors spinning inside a section of plastic pipe. These rotors generate electrostatic energy which people have mistakenly been led to believe is not a source of significant power. This design by William Hyde has an electrical output which is some ten times greater than the mechanical input power required. A **C**oefficient **O**f **P**erformance = 10 result like this, has to be significant, especially since the device is of fairly simple construction.

Electrostatic energy field power generating system

Patent US 4,897,592 **30th January 1990** **Inventor: William W. Hyde**

Abstract:

Externally charged electrodes of an electrostatic generator induce charges of opposite polarity on segments of a pair of confronting stators by means of electric fields within which a pair of rotors are confined during rotation to vary the charge binding field linkages between confronting rotors and stators by a shielding action of the rotors in a plane perpendicular to the field flux. A high electric potential difference induced between the stators resulting from such rotation of the rotors, is transformed by an output circuit into a reduced DC voltage applied to a load with a correspondingly increase current conducted through it.

US Patent References:

2522106	Electrostatic machine	Sep 1950	Felici	310/309
3013201	Self-excited variable capacitance electrostatic generator	Dec 1961	Goldie	322/2A
4127804	Electrostatic energy conversion system	Nov 1973	Breaux	322/2A
4151409	Direct current variable capacitance electric generator	Apr 1979	O'Hare	250/212
4595852	Electrostatic generator	Jun 1986	Gundlach	310/309
4622510	Parametric electric machine	Nov 1986	Cap	322/2A

Description:

This invention relates to the generation of electrical power by conversion of energy from an electrostatic field. The conversion of energy from a static electric field into useful electrical energy by means of an electrostatic generator is already well known in the art as exemplified by the disclosures in U.S. Pat. Nos. 2,522,106, 3,013,201, 4,127,804, 4,151,409 and 4,595,852. Generally, the energy conversion process associated with such prior art electrostatic generators involves the input of mechanical energy to separate charges so that a considerable portion of the output is derived from the conversion of mechanical energy.

It is therefore an important object of the present invention to provide an electrostatic generator in which electrical power is derived from the energy of static electric fields with a minimised input of mechanical power.

Summary:

In accordance with the present invention, static electric fields are established between electrodes externally maintained at charge levels of opposite polarity and a pair of internal stator discs having segmental surfaces that are dielectrically spaced to confine thereon charges induced by the electric fields. A pair of rotor discs are rotated within continuous electric fields in planes perpendicular to the field flux to locationally vary the charge linkage established by the electric fields between the electrodes and stator discs. Such changes in charge linkage are effected by rotation of electrically conductive segments of the rotor angularly spaced from each other to partially shield the stator discs from the electric fields. The segments of each rotor disc have charged faces confronting the electrodes in its field to shield the stator disc over a total face area that is one-half the total area of the confronting segment surfaces on the stator disc to which the induced charges are confined. Charges on the rotors and stators are equalised by electrical interconnections established through the rotor shafts. The stator discs are electrically interconnected with an electrical load through an output circuit transforming a high potential between the stator discs into a reduced dc voltage to conduct a correspondingly multiplied current through the load.

Brief Description of the Drawings:

These and other objects and features of the present invention will become apparent from the following description taken in conjunction with the preferred embodiments thereof with reference to the accompanying drawings in which like parts or elements are denoted by the same reference numbers throughout all of the different views shown in the drawings and where:

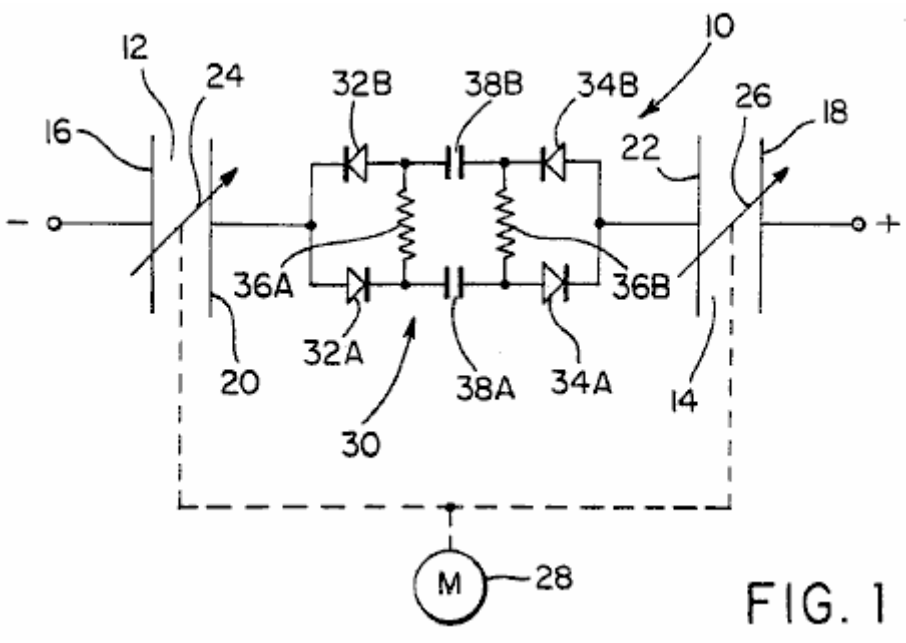


FIG. 1

Fig.1 is a simplified electrical circuit diagram corresponding to the energy conversion system of the present invention.

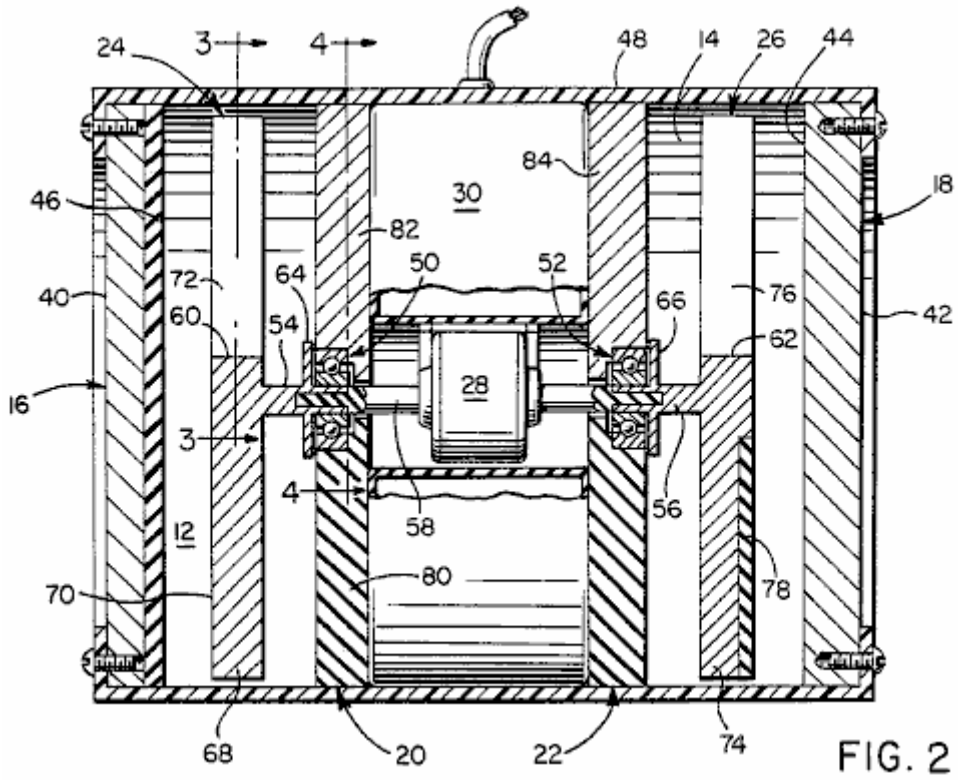


FIG. 2

Fig.2 is a side section view of an electrostatic generator embodying the system of Fig.1 in accordance with one embodiment of the invention.

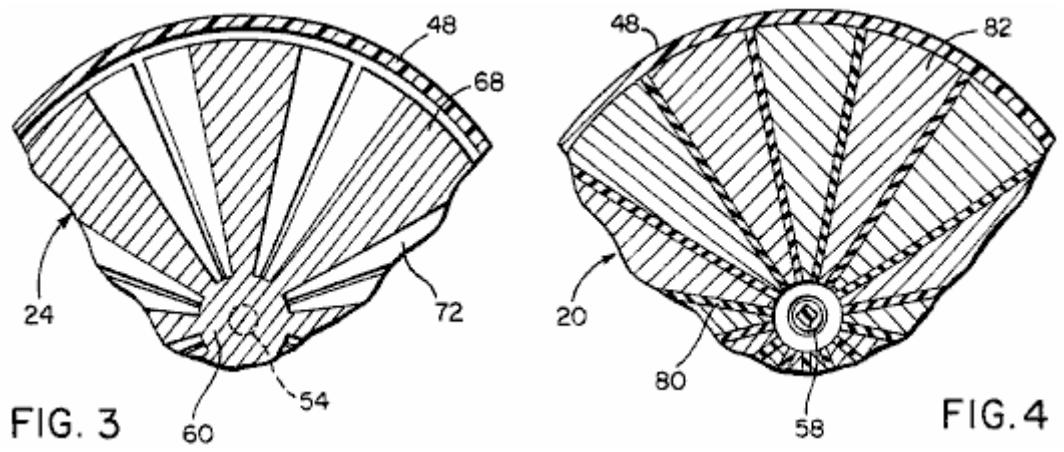


Fig.3 and **Fig.4** are partial section views taken substantially through planes indicated by section lines **3--3** and **4--4** in Fig.2.

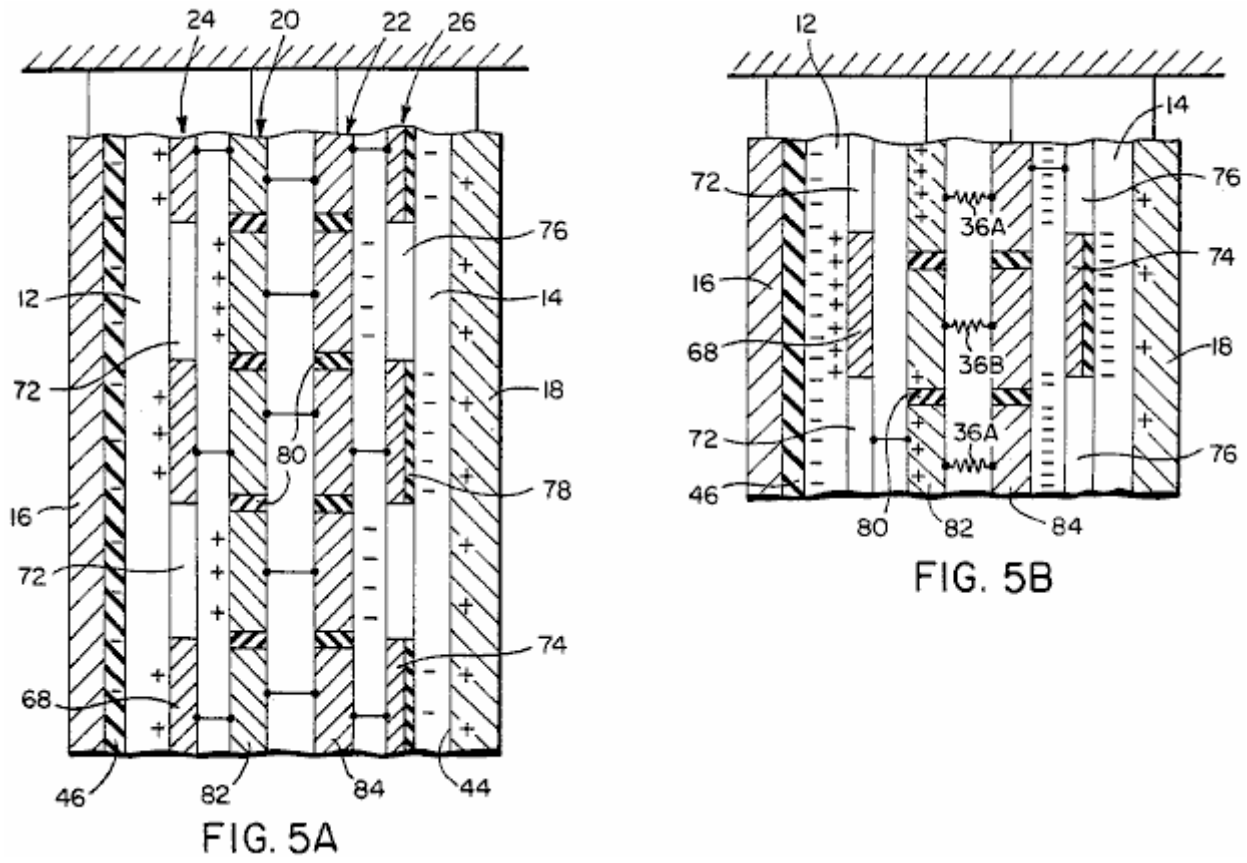


Fig.5A and **Fig.5B** are schematic partial laid out top views of the electrostatic generator of Figs.2-4, under static and dynamic charge distribution conditions, respectively.

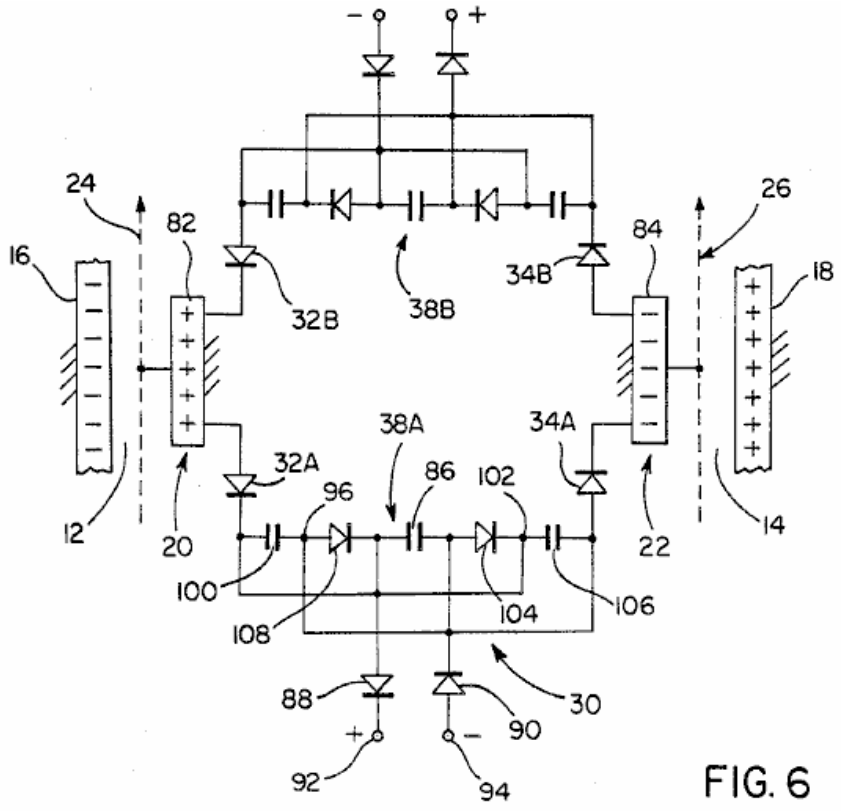


FIG. 6

Fig.6 is an electrical circuit diagram of the output circuit of the generator shown in Fig.2, in accordance with one embodiment.

Detailed Description of the Preferred Embodiment:

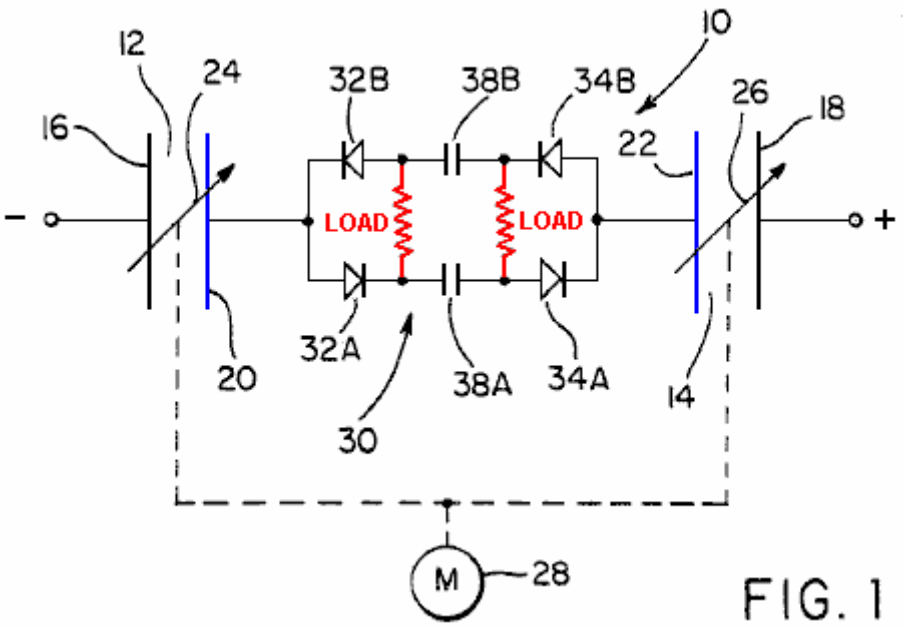


FIG. 1

Referring now to the drawings in detail, **Fig.1** diagrammatically depicts the energy conversion system of the present invention generally referred to by reference numeral **10**. As shown in **Fig.1**, the system includes a pair of electrostatic fields **12** and **14** established by electrostatic charges of opposite polarity applied to electrode plates **16** and **18** from some external energy source. Thus, the electrostatic field **12** is established between electrode **16** and a stator disc **20** while the electrostatic field **14** is established between electrode **18** and a stator disc **22**. In accordance with the present invention, electrostatic charge linkages established by the flux of the fields between the electrodes and stators are periodically varied by displacement within the continuous energy fields **12** and **14** in response to rotation of rotors **24** and **26** aligned with planes perpendicular to their common rotational axis and the field flux, as will be described.

The rotors are mechanically interconnected with an electric motor **28**, as diagrammatically illustrated in **Fig.1**, which rotates them around their common axis. Electrical energy may be extracted from the electric fields **12** and **14** during rotation of the rotors **24** and **26** (by motor **28**) through an output circuit generally referred to by the reference number **30**. The output circuit **30** as shown in a simplified fashion in **Fig.1**, includes two pairs of current-conducting diodes **32A, 32B** and **34A, 34B**. The diodes of each pair are connected with opposite polarity and each pair is connected in parallel to one of the stators **20** and **22**. The diodes of each pair are also electrically connected across an electrical load represented by resistors **36A** and **36B** with capacitor networks **38A** and **38B** interconnected between each pair of diodes by means of which the voltage potential between the stators **20** and **22** is reduced in favour of an increased current through the electrical load.

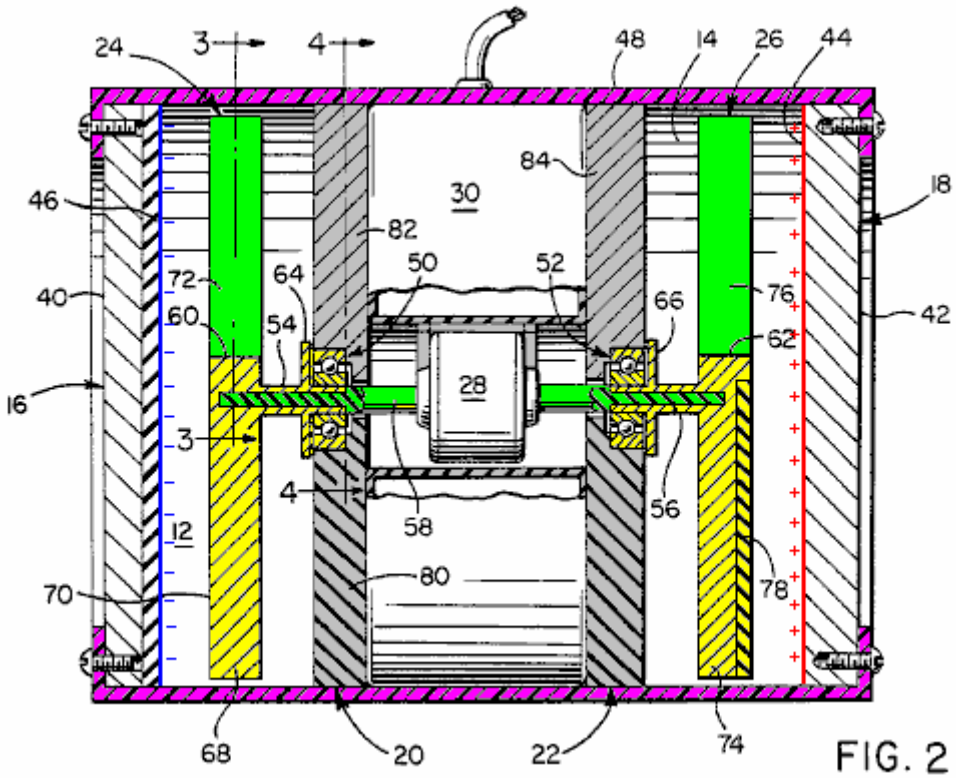


FIG. 2

Referring now to **Figs 2, 3** and **4** in particular, a physical embodiment of the energy conversion system shown in **Fig.1** is shown. The electrodes **16** and **18** are in the form of circular plates or discs made of an electrically conductive metal having external surfaces **40** and **42** adapted to be charged from the external source as already mentioned. The internal surface **44** of electrode **18** is thereby adapted to maintain a positive charge opposite in polarity to the negative charge of the electrode **16** which is maintained in a stable ion form within a dielectric surface portion **46** of the electrode **16**. The energy conversion system may be enclosed within an outer housing **48** to which the electrodes **16** and **18** are secured.

With continued reference to **Fig.2**, the stators **20** and **22** mounted by housing **48** in axially fixed spaced relation to the electrodes **16** and **18** are provided with bearings **50** and **52** supporting the powered rotor shaft driving the shaft assembly which has electrically conductive shaft sections **54** and **56** to which the rotors **24** and **26** are respectively connected. In the embodiment illustrated in **Fig.2**, the drive motor **28** is mechanically interconnected with the shaft sections **54** and **56** through an electrically nonconductive shaft section **58** of the power shaft assembly for the simultaneous rotation of both rotors **24** and **26** at the same speed and in the same direction about their common rotational axis perpendicular to the parallel spaced planes with which the electrode and stator discs are aligned. The electrically conductive shaft sections **54** and **56** are respectively keyed or secured in any suitable fashion to hub portions **60** and **62** of the rotors and are provided with flange portions **64** and **66** forming electrical wipers in contact with confronting surfaces of the stators **20** and **22**, which are inductively charged by the static electric fields **12** and **14** to equal levels of opposite polarity.

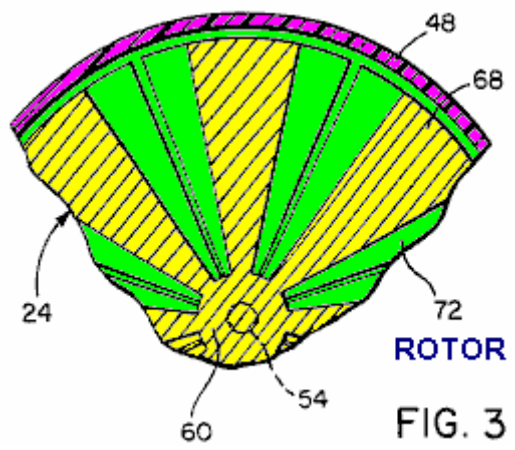


FIG. 3

As more clearly seen in **Fig.2** and **Fig.3**, the rotor **24** has several angularly spaced, field linkage controlling segments **68** projecting radially outwards from the hub portion **60**. Each rotor segment **68** is made of an electrically conductive metal having a face **70** on one axial side confronting the adjacent electrode **16**. The faces **70** confronting the electrode **16** are charged positively by the electric field **12** extending between the dielectric surface portion **46** of electrode **16** and the stator disc **20**. While the electric field **12** projects through the spaces **72** between the rotor segments **68**, the rotor segments **68** themselves shield portions of the stator disc **20** from the electric field.

The rotor **26** is similarly formed with rotor segments **74** angularly spaced from each other by spaces **76** through which the electric field **14** extends between the positively charged surface **44** of electrode **18** and the stator **22**. The rotor segments **74** of rotor **26** as shown in **Fig.2**, are provided with dielectric surface portions **78** confronting the internally charged surface **44** of electrode **18**. While the rotor segments **74** are negatively charged by the electric field **14** within the surface portions **78**, they also shield portions of the stator disc **22** from the electric field as in the case of the rotor segments **68** already described. The internal dielectric surface portion **46** of electrode **16** and dielectric surface portions **78** of rotor **26** act as a stabiliser to prevent eddy currents and leakage of negative charge. Further, in view of the electrical connections established between the rotors and the stator discs, the charge on each stator is equalised with that of the charge on its associated rotor.

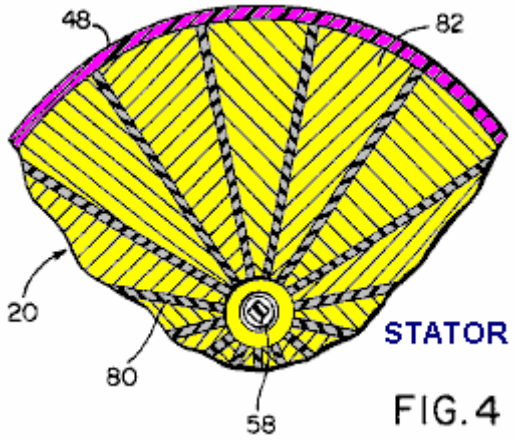


FIG. 4

As shown in **Fig.2** and **Fig.4**, the stator disc **20** includes several segments **82** to which charges are confined, closely spaced from each other by dielectric spacers **80**. The segments **82** are electrically interconnected with the rotor segments **68** through rotor shaft section **54**. Similarly, the segments **84** of the stator **22** are electrically interconnected with the rotor segments **74** through rotor shaft section **56**. The stator segments **82** and **84** are therefore also made of electrically conductive metal. Each of the segments **82** of stator **20** is electrically interconnected through the output circuit **30** with each of the segments **84** of the stator. The stator discs being fixedly mounted within the housing **48**, centrally mount the bearings **50** and **52** through which the electrically nonconductive motor shaft section **58** is journaled as shown in the embodiment of the invention illustrated in **Fig.2**. Further, the total area of the charged segment surfaces on each of the stator discs is greater than the total area of the faces **70** or **78** on the segments of each associated rotor disc **24** or **26**. According to one embodiment, the total charged stator surface area is twice that of the rotor face area.

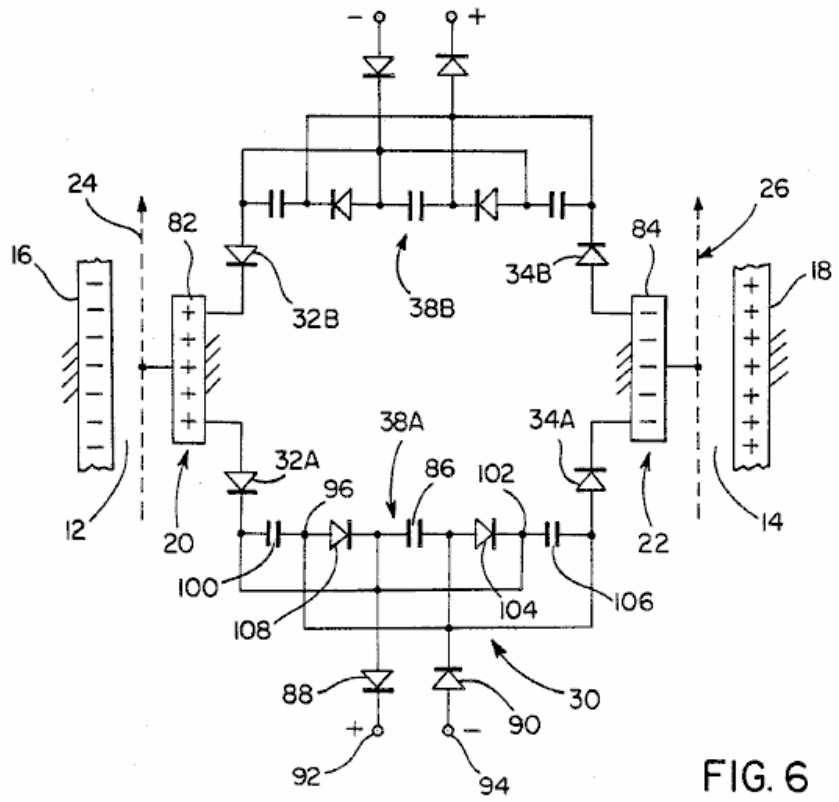


FIG. 6

According to the embodiment of the invention illustrated in **Fig.6**, the output circuit **30** includes the two oppositely poled capacitive circuit networks **38A** and **38B** connected across each aligned pair of stator segments **82** and **84** on the stators **20** and **22** by means of the oppositely poled diodes **32A** and **34A**. Each of these capacitive circuit networks includes a capacitor **86**, the opposite sides of which are connected by oppositely poled diodes **88** and **90** to positive and negative load terminals **92** and **94** across which a suitable electrical voltage is established for operating an electrical load. The diode **88** is connected to the junction **102** between diode **104** and one side of capacitor **106**. The diode **88** is also connected to the junction between one side of capacitor **100** and the diode **32A**. The diode **90**, on the other hand, is interconnected with the junction **96** between diode **108** and capacitor **100**. Also, diode **90** is connected to the junction between the other side of capacitor **106** and the diode **34A**. The foregoing circuit arrangement of capacitive network **38A** is the same as that of network **38B** by means of which aligned pairs of the stator segments **82** and **84** have the electrical potentials between them transformed into a lower voltage across the load terminals **92** and **94** to conduct a higher load current.

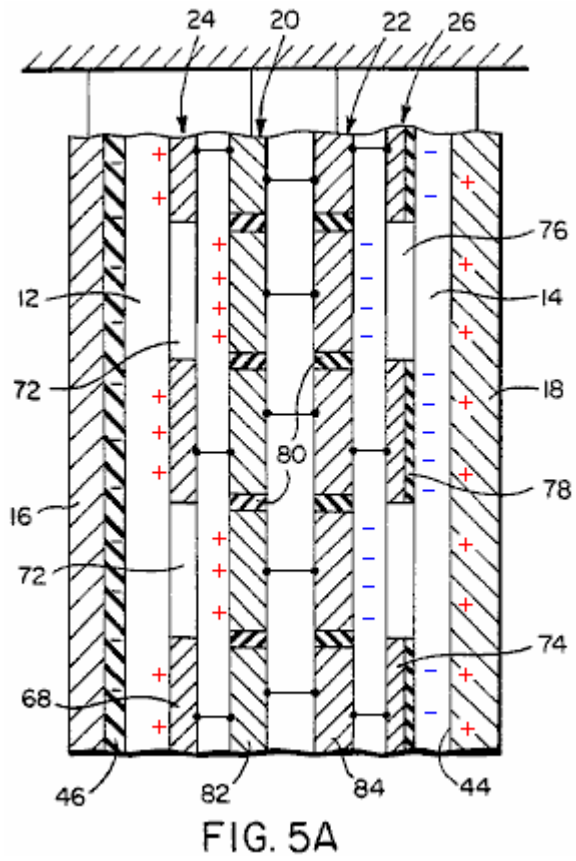


FIG. 5A

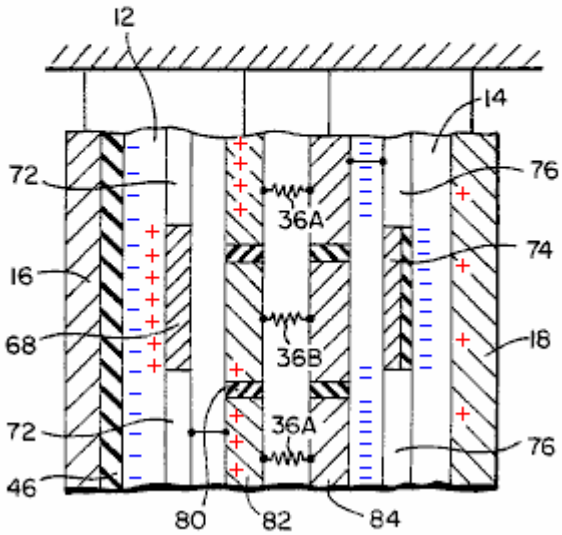


FIG. 5B

Fig.5A illustrates the distribution of charges established in the electric fields **12** and **14** between the electrodes and stators under static conditions in which each of the rotor segments **68** and 74 are positioned in alignment with one of the stator segments **82** and **84** to thereby shield alternate stator segments from the electric fields. The charges established by the electric fields are therefore confined to the faces of alternate stator segments confronting the electrodes and are equalised with the charges established on and confined to the shielding faces of the rotor segments confronting the electrodes by virtue of the electrical interconnection between the rotors and stators as already mentioned. As depicted in **Fig.5B**, when the rotors are rotated, the charge linkages established by the electric fields between the electrodes and alternate stator segments **82** or **84** are interrupted by the moving rotor segments **68** or **74** so that previously shielded stator segments become exposed to the fields to re-establish field energy linkages with the associated electrodes. Such action causes electrical potentials to be established between the stator segments **82** and **84**.

It will be apparent from the foregoing description that the electrostatic energy fields **12** and **14** of opposite polarity are established maintained between the externally charged electrodes **16** and **18** and the internally charged stators **20** and **22** under static conditions as depicted in **Fig.5A**. During rotation, the rotors **24** and **26** continuously positioned within the energy fields **12** and **14**, exert forces in directions perpendicular to the field flux representing the energy linkages between electrodes and stators to cause interruptions and reestablishment of energy linkages with portions of different stator segments as depicted in **Fig.5B**. Such energy linkage locational changes and the charge binding and unbinding actions between electrodes and stators creates an electrical potential and current to flow between stators through the output circuit **30**. Thus, the output circuit when loaded extracts energy from the electric fields **12** and **14** as a result of the field linkage charge binding and unbinding actions induced by rotation of the rotors. The stator segments **82** and **84** shielded from the electric fields by the moving rotor segments **68** and **74** as depicted in **Fig.5B**, have electric potentials of polarity opposite to those of the external electrodes **16** and **18** because of the field linkage charge unbinding action. Previously shielded stator segments being exposed to the electric fields by the moving rotor segments, have the same electric potential polarity as those of the external electrodes because of field linkage binding action. Since the forces exerted on the respective rotors by the electric fields **12** and **14** of opposite polarity act on the common rotor shaft assembly perpendicular to these fields, such forces cancel each other. The energy input to the system may therefore be substantially limited to mechanical bearing losses and windage during conversion of electrostatic field energy to electrical energy as well as electrical resistance losses and other electrical losses encountered in the output circuit **30**.

Based upon the foregoing operational characteristics, rotation of the rotors in accordance with the present

invention does not perform any substantial work against the external electric fields **12** and **14** since there is no net change in capacitance thereby enabling the system to convert energy with a reduced input of mechanical energy and high efficiency, as evidenced by minimal loss of charge on the electrodes. It was therefore found that working embodiments of the present invention require less than ten percent of the electrical output energy for the mechanical input. Further, according to one prototype model of the invention, a relatively high output voltage of 300,000 volts was obtained across the stators. By reason of such high voltage, an output circuit **30** having a voltage reducing and current multiplying attribute as already described, was selected so as to render the system suitable for many practical applications.

The Suppression of Knowledge.

Early in October 2011, Sir Christopher Woodhead, former Chief Inspector of Schools in England, stated that children should leave school at age 14 and become apprentices. While I personally have a high opinion of apprenticeships for certain professions, this appears to me to be a direct attempt to lower the education levels of the UK population even further than it already is now, where many sales staff are unable to perform even the most basic arithmetic as part of their jobs.

The basis for how we live our lives and understand the things around us, society, personal relationships, family values and the like, is rooted in what we are told by "the authorities". Unfortunately, much of what you have been told is just not true. You have been lied to about science, astronomy, the environment, global warming, government, taxation, war, energy, inventions, education, terrorism, health, finance and the media, to name just a few topics which spring to mind.

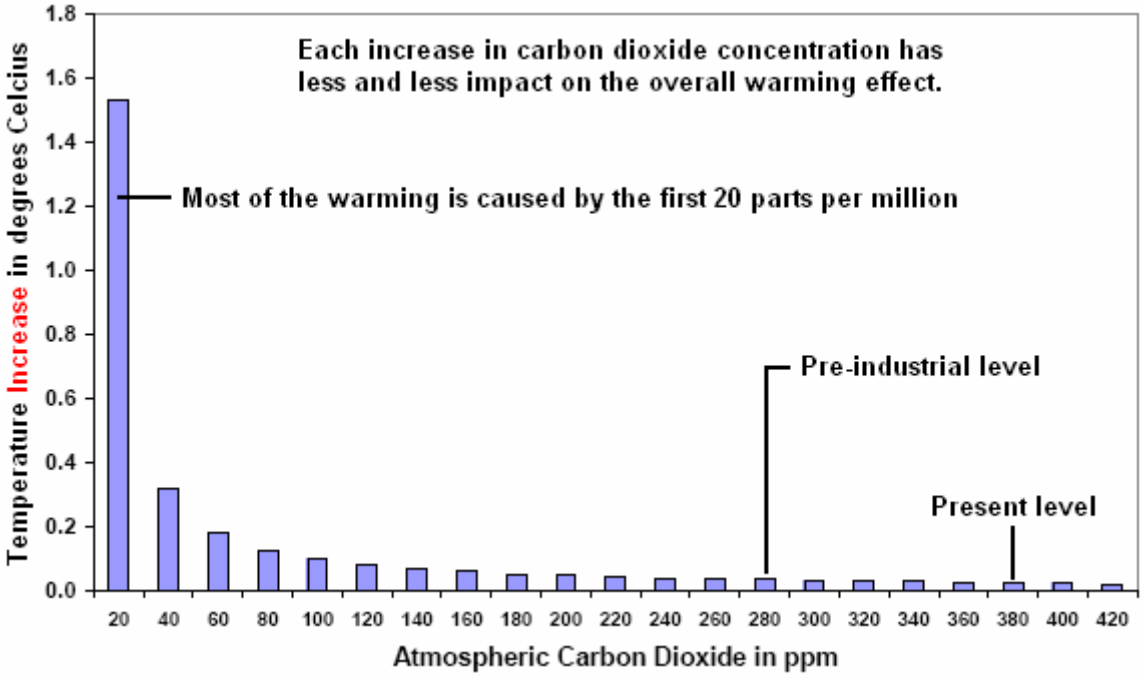
Science: There has been, and still is, systematic, deliberate, and frequently brutal suppression of scientific fact. We are being told that you have to "burn a fuel" in order to produce power which is then "consumed" when we use it to power our lives. This is a deliberate lie. We are told "scientific" things which are sheer fantasy and designed to keep the truth from us. The objective is to keep the people of this planet ignorant, weak, and subject to the people who know the real facts, and to remove all possible wealth from the ordinary person.

Astronomy: We are told that the only way we can get into space is to use rockets. The people who say this not only know it is a lie, but they themselves have had electrogravitic drive vehicles for at least fifty years now. This sort of rubbish is roughly the equivalent of telling people that they can only get around by hopping on one foot. We are told that the speed of light is a constant and that is a lie. We are told that nothing can go faster than the speed of light, and that is a lie. We are told that gravity depends on mass, and that is a lie. We are told that the Earth is the only planet in our solar system which has an atmosphere, and that is a lie. In fact, almost all of what we are told is incorrect.

The Environment: We are told that we need to run our lives by burning "fossil fuels" which are limited in quantity and so there is "an energy crisis". While some oil is produced by the decay of organisms and vegetation laid down in earlier eras, oil wells which were pumped supposedly dry are actually filling up again with oil created in the crust of the Earth and which does not come from decaying vegetation and organisms. Nor is there the slightest need to burn oil for power, or have central power stations and fragile wires running all over the countryside. Those things are wanted by the powerful elite in order to control ordinary people and to part them from their money.

Global Warming: We are told that man's burning of fossil fuels is causing global warming through the production of Carbon Dioxide. This is the most ridiculous rubbish possible. Carbon dioxide is a natural gas, essential for plant growth, which in turn is essential for our survival. Carbon dioxide levels have long since reached the level where any increase has almost no further effect on warming. The percentage produced by man is trivial, as the vast majority comes out of volcanoes, the next largest amount is produced by animals and insects, and man's contribution is minor. Global warming and global cooling are natural events which were taking place long before man was around in any significant numbers and certainly, centuries before the burning of fossil fuels started. The contribution from all of the cars in all of the would is only about 1.7% of man's minor contribution and the fussing about "carbon footprint" is only a con which is being used to extract more money from the ordinary person.

David Archibald in this document http://www.davidarchibald.info/papers/Climate_Outlook_to_2030.pdf, published the graph shown here in 2006:



This graph clearly shows that even a big increase in the present carbon dioxide level has almost no impact on atmospheric temperature. It has also been calculated that an increase to 620 ppm would give a temperature increase of only $0.16^{O}C$. The graph shown above, together with its calculations should have ended any further discussion about reducing carbon dioxide emissions. However, this has most definitely not been the case.

For financial reasons, a decision has apparently been taken to keep up the pressure on carbon dioxide reductions, resulting in senseless investments on new technology for capture and storing carbon dioxide, “beautiful” windmills absolutely everywhere and the mandatory use of bio fuel (which has a negative impact on food production). Don’t forget about all the revenue which ‘governments’ are getting from "climate taxes".

All these taxes and the massive spending of taxpayers money, really threatens the economic foundations of the whole of the western world, but worse still are the consequences of the commitment to bio fuel. This mad strategy has caused a near doubling of the cost of cooking oil and essential foods such as rice, because farmland is now being used for the production of bio fuel. These increases are hitting poor countries the worst as the people can no longer afford to pay for basic necessities.

If the intention of the IPPC and the world’s politicians is to starve people to death, then they are definitely on the right track through their continued insistence on the supposedly harmful effects of carbon dioxide. In reality, the greenhouse gas which has the most effect is actually water vapour which accounts for about 98% of all warming. I wonder how the politicians are going to prevent water vapour from getting into our atmosphere !

In his video “An Inconvenient Truth” Al Gore points out the close correlation between the fluctuations of global temperatures and the levels of carbon dioxide in the atmosphere. He point out strong rises and falls in the average global temperature (nearly all of which are before mankind started burning oil) and the related strong rises and falls of carbon dioxide concentration in the air. What he is very careful to avoid mentioning is that the changes in carbon dixide concentration, lag behind the global temperature changes by some 600 years. The carbon dioxide concentrations are a **result of** the global warming and **not** the cause of it.

Government and Taxation: It appears that the sole purpose of government is to enrich the people in power at the expense of the ordinary person. People in a democracy are supposed to be “free” but this is a gross illusion. If you were working all week in a company and at the end of the week you were paid £1,000 cash. When you start going home, an armed gang of thugs threatens you and steals £800 of your £1,000 and that happens every week without fail. So, what do you think of that arrangement? Good or bad? That is what government does for you at this time. But, you say, “we live in a democracy, so we can fire the politicians”. That is the equivalent of having three exits from your workplace and each exit has a different gang of thugs ready to rob you. Yes, indeed, you have choice. You can choose which exit to use and therefore, which particular gang of thugs robs you.

You think that is overstating the case? I have seen it calculated that 80% of everything you earn is taken away from you in "taxes" and there is justification for that figure when you consider what contributes to the cost of everything you do. You pay taxes. Can you remember being asked if you would be willing to pay tax?

Oh, but you say, we need taxes in order to pay for public services, health, education, transport, etc. Yes we do, but how much of your money actually goes on that. In the UK there is a large annual "car tax" which was introduced with the excuse that it was needed to finance the building and repair of roads. I understand that at least 85% of that money is not spent on the roads. There is a massive tax on vehicle fuel and I have seen the figure of 85% mentioned. If that is the case, fuel should only be one sixth of the price charged and that extra cost finds it's way into almost everything else, food, clothes, electrical goods, all manufactured items, heating, electricity supply, gas supply, etc. etc. In addition, there are major taxes on alcohol, tobacco, service industries, travel, and everything else that 'politicians' can think up. So you're a free person living in a free democracy? Who are you kidding? Do you actually have "sucker" stamped on your forehead? For further information, visit http://www.yourstrawman.com where the facts are laid out clearly.

Take the case of America. The objective of the Civil War was to make money for the unscrupulous. A really major fraud was perpetrated on the American people at that time. They believed (and most still do) that the United States is the government of the American nation. That is not the case. "The United States of America" is actually a commercial conglomerate of private corporations, corporations with names like "The State of Arizona", or, "The State of Texas". Federal taxes are imposed on the American people, but that goes directly into the private finances of these companies and not a single dollar of it goes to benefit any ordinary American citizen. There is much more detail on this in Chapter 15.

War: The general public is held subject to what passes for 'government' through fear and ignorance. The government announces that some country needs to be attacked and invaded in order to "free" the people and establish democracy. So, generally without any realistic reason, that country is attacked and invaded, much of the civilian population killed, massive property damage inflicted, and great distress and hardship caused for many people. This is organised by the politicians. Do they themselves take part in the fighting? No. Who benefits from the war? Answer: the arms manufacturers, and companies such as Halliburton which are paid taxation money to rebuild what the local army has just spent time destroying. The country attacked is likely to have it's assets stolen and then be saddled with a massive fake debt so that it can be forced to pay incredible amounts of "interest" for decades afterwards. The lucky country has been attacked, damaged, citizens killed, infrastructure destroyed, assets stolen and then robbed for years afterwards by the people who caused that devastation. You did say that you live in "a free democracy", didn't you?

Energy: The energy companies are very keen to sell you fuel for you to burn to "get" energy to run your vehicles, home appliances, heating, etc. They are also very keen that you should not discover that you are surrounded by energy which can be used for all those things without paying the local energy company anything. Most people imagine that the electrical power which flows out of the wall socket in your home (and for which you have to pay a good deal), comes along the wires from the generator station which burns fossil fuel to drive the generators. That is not the case, as not a single electron of the current you use actually comes from the power station but instead is collected locally from free environmental energy. If you knew how to do it, then you would have no need for the power company, power lines across the country, power stations or oil drilling and shipment around the world. Cars don't need fuel to run but that information is heavily suppressed and you are lied to when you are told that you need to burn a fuel to power your life.

Inventions: When they discover that there is such a thing as free-energy, most people think that what is needed is a scientific breakthrough in order to develop the technology needed to let everyone have free energy. The reality is quite the reverse as the technology has been discovered and implemented literally thousands of times. You are being lied to when you are told that free-energy is "impossible". I am reliably informed that in America alone, more than 40,000 free-energy patents have been confiscated by the (private company called the) Patent Office. Inventors are then silenced by a gag order under the absurd excuse that the invention is "of National Security importance". Inventors who don't try to patent their inventions are intimidated, some are murdered, some are attacked and their reputation smeared, some are illegally put in mental institutions, some are framed and all suffer massive financial opposition if they try to start manufacturing a product for sale – not even Cal-Tech managed to get past that financial block. I personally know several researchers who have been unlawfully intimidated.

Education: There is a massive program of "dumbing-down" education combined with the deliberate teaching of things which are known to be completely wrong. After only a few decades of this, many adults can't perform basic arithmetic or construct simple sentences in their native tongue. Science and Engineering are neglected and most people don't understand how things work. Most schools no longer allow

schoolchildren to perform experiments in chemistry or physics, usually with the excuse that it is 'too dangerous'. University students are told that there is no such thing as perpetual motion, which is as laughable as saying that water isn't wet. Newton actually said that gravity is a push and not a pulling force as indicated by the mis-translation of his Latin original text. This incorrect notion of gravity is still taught in spite of the fact that many governments have had electrogravitic drive vehicles since the early 1960s, driven by electromagnetic waves just below the frequency of infrared, disproving what is supposed to be Newton's theory, in a most resounding fashion. Any academic who steps out of line and admits that students are being taught rubbish, is immediately attacked with every possible means to discredit, ridicule and isolate him. The objective in all of this is to manoeuvre the general public into a weak position where they are ignorant, uneducated, fed false information and distracted with things of no importance. The ultimate aim is the slavery of the masses without them even realising that they are being enslaved.

Terrorism: The main way to make people give up their rights and do whatever you want is through fear. The book "1984" (dated that way because it was written in 1948) explains how this is done. Fictitious enemies are invented in order to frighten people and stop them complaining when vast amounts of their money is given to vested interests under the guise of being steps to protect them from this imagined danger. The "cold war" was completely fake with both sides using it as an excuse. The invasion of smaller countries is never, ever, for the reason given publicly, but is always for financial gain by the people in power and their commercial colleagues. The reasons given for these invasions are so ridiculous that it never ceases to amaze me that the public actually accept and believe these wild stories. One of the strands of this continuous deception waged against the public is fake "terrorism", frequently bolstered up by "false flag" attacks to discredit the perfectly innocent people who are supposed to have made those attacks. Any real terrorist attacks are normally a response to atrocities committed against the supposed "terrorists". The commercial companies which pass for governments, are the biggest terrorists around.

Health: The pharmaceuticals industry is one of the biggest money-making operations on the planet. It may come as a shock to you, but they don't want to cure people. Instead, they want people to go on buying expensive medication at heavily inflated prices. Long term health treatments are their ideal situation as that provides them with a continuous revenue stream. Several people have come up with very effective cures for cancer and other serious illnesses, and those people have received the same unjustifiable, outrageous treatment as do the inventors of free-energy devices. The pharmaceuticals industry is well aware of the highly beneficial effects of colloidal silver, but they will never produce any as they can't patent it and charge highly inflated prices for it. Consequently, they are not interested as their only objective is to make money, and it is definitely not to cure people.

Bob Beck's Electronic Pulser.

The late Bob Beck emphasised this in his video lecture which may still be available at:
http://video.google.com/videoplay?docid=-3389948315844437935&ei=XdqBSrucK5_sqAPr28irBg&q=Suppressed+Medical+Discovery%3A+Dr.+Robert+C.+Beck+%28+Cancer%2CAIDS%2C+anything+viral%29#

Here is an excerpt from that lecture:

For the last five years I have spent my own money (no government or university funding) in investigating a thing which has proven to be the most remarkable thing for all diseases that I have ever heard of personally, and I have been in this field for about forty years.

I have a stack of IRB Studies – PCR laboratory reports from major hospitals around the United States, and until these are peer-reviewed published, I am not supposed to show them to anyone who is not a medical doctor because of patient confidentiality.

Now, how many of you have read that there is no cure for AIDS? That is an absolute lie. There has been a perfectly workable, 95% accurate cure for HIV and cancer and herpes and hepatitis and Epstein-Barr and about a dozen other incurable diseases, which was invented on March 11th 1990 at the Albert Einstein College of Medicine in New York City by Drs Kelly and Wymans.

And what happened to this breakthrough which is far more important than penicillin, antibiotics, anything you can name? - it has been suppressed. And why has this information been withheld from you – information which you can easily prove for yourself (don't take my word here for anything: check it out)? I found recently that the Mafia owns about 51% of major pharmaceutical houses as well as working the other side of the track with the ***illegal*** drugs from South and Central America and China.

The medical cartels in this country (USA) will charge you from $50,000 to $200,000 if you have terminal

cancer and this is for surgery and chemotherapy and radiation and hospital care.

Consider a packet of Wrigley's Spearmint chewing gum. That packet of chewing gum cost me a nickel (5 cents or $0.05) when I was a young man and I paid $1.32 for that packet at the Chicago Airport recently - $1.00 for the gum plus $0.32 for taxes and handling. The price of that package of chewing gum is every penny (cent) that it takes to cure [and I'm not supposed to use this word as only a doctor may use the word 'cure'] most of the known diseases in America today. And why haven't you heard about this? – because General Electric, Westinghouse, and the other companies have billions of dollars tied up in X-rays, CAT scanners and MRIs which are leased to health organisations, and for $1.32 you can cure yourself without doctors, without pharmaceuticals, without medicines, certainly without surgery, certainly without chemotherapy, without any drugs, without any herbs, any homeopathy remedies – you can do it for yourself and it has been done.

People keep asking "why haven't I heard about this, why are there no hospital studies?". There are hundreds of them. Here is one test result: At this date, the eleventh of August 1998 this patient had only about 3,063 particles of HIV in his blood. But now, as this test develops, the PCR on the third of November 1998, (several months later), the count was less than 1 in a hundred, which means that the test result was zero as far as this test was concerned. We have a stack of these reports a foot and a half (45 cm) deep, where **every one** of these full-blown AIDS patients, every one even though some of them were on life-support systems or needed attendants to lift them from the bed, these people are symptom-free now. Every single one of them is back at work. Not one of them has the symptoms of AIDS. So, if you read in your newspaper 'the proceeds of the baseball game this Saturday will be donated to a search for a cure for AIDS' that is an absolute lie.

There is, and has been, a cure for AIDS – **all** AIDS – a cure for about 95% of cancer, a 100% cure for Epstein-Barr, hepatitis, lupus, about a 50% cure for herpes, and I've been funding these out of my own pocket. I have absolutely nothing whatsoever for sale. God has been exceedingly good to me – I think that I am going about my Father's business here, and I am paying for it myself and I am giving to the world.

There is a free paper available entitled "Take Back Your Power" (a web download link for this is http://www.free-energy-info.tuks.nl/Beck.pdf) you do not need doctors, nurses, chiropractors, herbs, pharmaceuticals, surgery. Only if you are going along with the politically-correct solutions do you need these things that your neighbours have been depending on for all these many generations. There was paper entitled "Blood Electrification and Immune System Restoration with Microcurrents – A Proven, Startling, Rapid, Inexpensive and Safe Discovery for Positive, Controllable Remissions". Now we found that this information has been suppressed. The doctors who discovered this, and let me tell you about the discovery: In 1990, they put a couple of small platinum wires into a Petrie dish that contained highly infected human blood, and the infection there was a very, very strong dose of HIV which was supposedly the cause of AIDS. They found that when they had electrified this blood, the HIV could no longer attach to any receptor sites in the blood. That means that the healthy cells could not be infected by the HIV. The only mention of this incredible discovery was in *Science News* March 30th 1991, page 207, '*Shocking Treatment Proposed for AIDS*' and it tells about how these two doctors had given a paper on this process saying, 'the experiments described on March 14th at Washington DC at the first International Symposium on Combination Therapies, showed that the shocked viruses lost the ability to make an enzyme crucial for their reproduction and could no longer cause the white cells to clump together – the true key signs of viral infection.

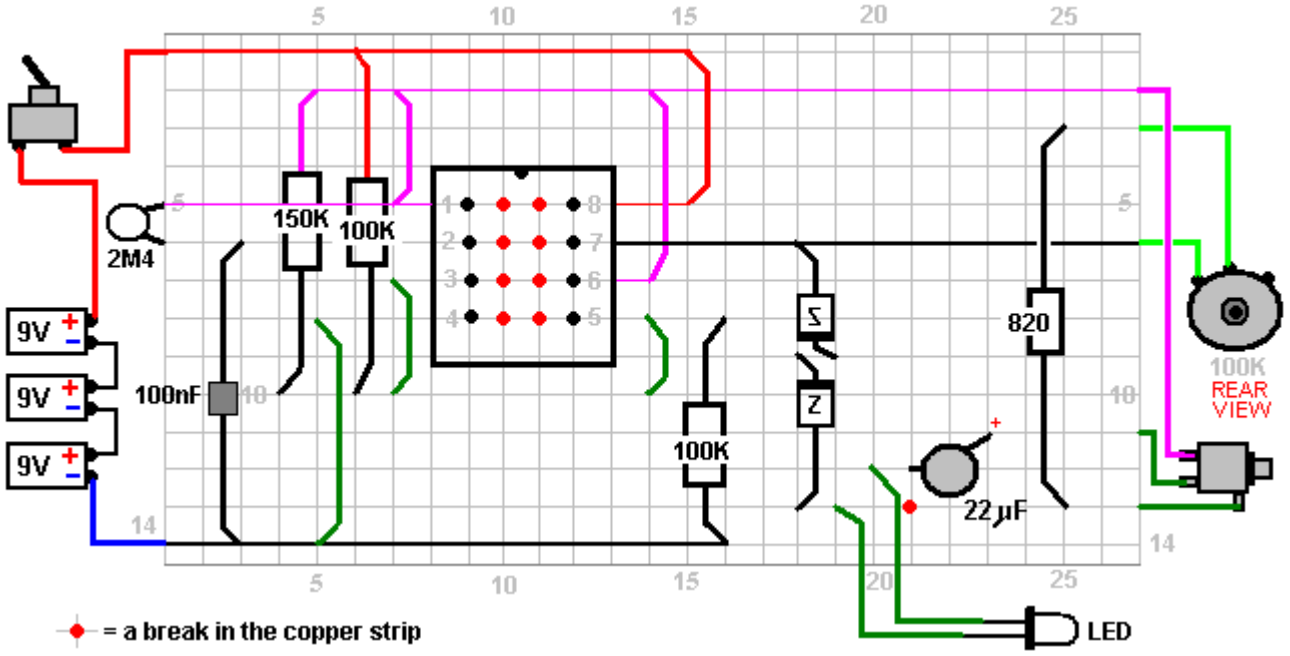
The only papers which I was able to find that got into the American press were *Science News* and article in *Longevity News* '*Electrocuting the AIDS Virus*'. But what happened in the *Longevity* article (which was December 1992, page 14), was that Dr Colley said 'it will be 15 years before this process is ready for human experimentation'.

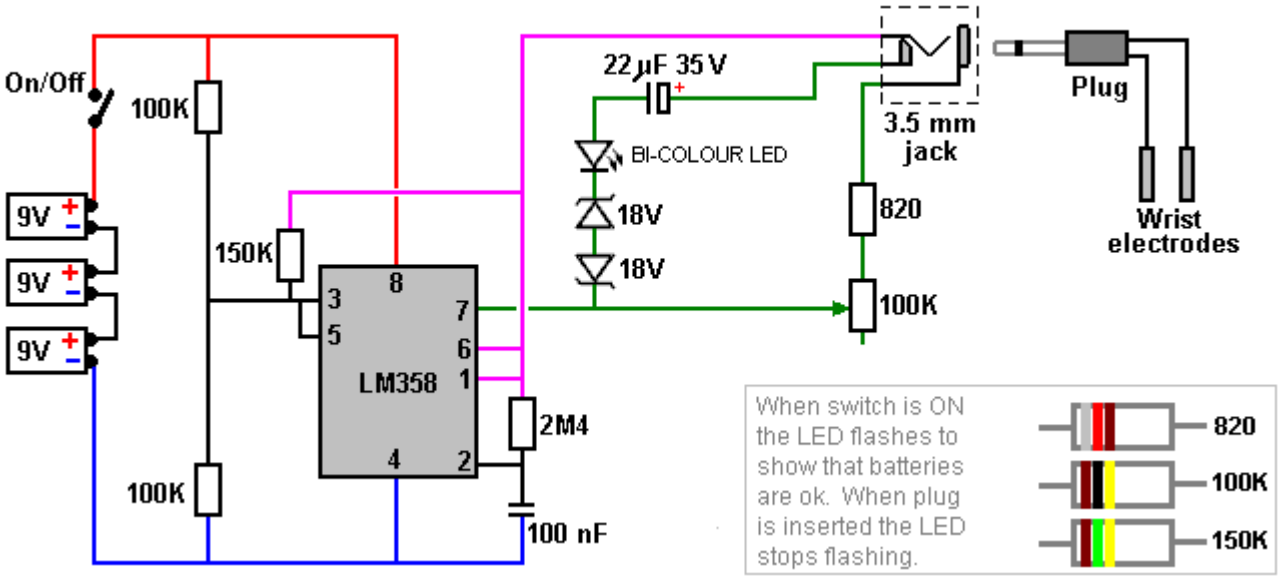
We have been doing it, sub-rosa, with patient consent, for over five years here, and we have the test results to prove it. When I was lecturing to a room of about 800 people, an 80-year old man came up and said 'Bob, I was sent home today, my cancer had metastasised all through my system. I was told to go home and make my will, as I had at the most, about two or three weeks to live. I went home and I made my will and I saw your article in some magazine and I did it. I am now free of all cancer. My biopsies are clean, my CAT scans are clean. I went back to work, and being an attorney, I had to work sixteen hours per day to catch up on all the work I had missed while I was in intensive care. Now I'm jogging three miles per day' (and there were tears running down his face).

We have had women who have had lupus for about ten years, barely able to make it from the couch in front of the television to the bedroom. They are back at the malls now, spending their money, walking around and able to hold jobs full-time. Do any of you know Linda Wright ? She brought in a clean blood test one day

and said 'my doctor had to send this to two different labs – he's never seen a person with clean blood after that person has had lupus'.

We have dozens of these people, and not one of these people went to the doctor for the cure, except for the people who are on this Institutional Review Board study at a hospital in Huntington Beach and Newport Beach, California. Now, the point is, in my humble opinion, and we can put anything against this because we know that it is true at this time, when I first started this I thought 'this is kind of interesting'. I have looked into five hundred things that never worked, the Rife Machine never worked, Holderclarke's Machine does not work, many, many, many of these devices work only marginally, and I felt 'I'm about to be taken in again', I have built Mucalsky multiple-wave oscillators as you know. I have built almost every electro-medical device since I've been in California (since World War Two began), and these things worked at best, marginally. But the thing which you can build for yourself, there is nothing that you have to buy. The circuit diagram and all of the instructions for building this, are in the paper. This is the one which I designed in 1991 to make an experimental study with, and you can get a kit from an electronics store. I get not one nickel from the tens of thousands of dollars worth of these kits which they have sold.





There is a parts list on page 4 that tells you what is needed. Even if you don't know what a resistor is, or a capacitor is, put the list on the counter at Radio Shack and tell them to give you what is on the list. The retail price ready-made is around $150. If you build it yourself it will cost $15 to $20 and this machine will not only make colloids, which is the most powerful medicine known to man, (which is why the FDA has stamped it out recently).

SOTA Instruments Inc. Canada (1-800-224-0242) builds these devices because the owner read the article in PACE magazine and his wife had chronic fatigue syndrome and Epstein-Barr so badly that she was non-functional. She is now up and around and is totally well today because my husband built this little box for about $15. The box straps on the arm with the electrodes positioned on the radial and ulna arteries and puts a very small electrification into the body which removes **all** parasites, viruses, fungus, microbes, pathogens – everything in the blood which does not belong there and was not there the day that you were born, eventually (after two or three weeks) being neutralised and discarded by your body, dead or alive by the spleen, the kidneys and the liver. This is darned exciting but when the hospitals tried to patent this, US Patent 5,188,738 presented to Dr Colley on February 23rd 1993, describes the process, as do fourteen other patents from Harvard MIT, page 1 column 1 says 'because of this problem (the fact that there is no known antibiotic or vaccine for any of these incurable diseases) the present invention has been devised to attenuate any bacteria, virus (including the AIDS HIV virus), parasites and fungus contained in blood contributed by a donor, these being rendered ineffective for infecting a normally healthy human cell'. Before these claims were accepted by the patent examiners, they had to provide the examiners with massive proof.

Now there has been a cure, a known cure and it costs $1.32 per patient so there is no money in this if they are trying to sell you combination therapies for $20,000 to $60,000 per year.

I strongly recommend that you get the paper as it contains the exact electrode locations which have been found to work the best and get the maximum current into the blood. It is almost as if God the Creator had locked the back door approach to the problems and with an adversary like the Holy Spirit which is within – the Bible says 'look within, not without'. And this time, the adversary is within. The 'aliens' in your blood are the things which are holding down your lifespan to seventy or eighty years. Many, many generations ago, according to the sacred texts, they lived to be several hundred years old, Solomon, David, Methuselah, etc.

So, we have raided about a dozen different medical textbooks including Grey's *Anatomy*, and we've published the exact spots where you can access the lymph tissue where these germs are going to be hibernating. After you clear yourself and get a clean bill of health, after three to five years you will re-infect yourself when many of these viruses which are latent, start germinating. And with this information, and these circuit diagrams and these instructions, we have literally given you back to yourself.

We are going to show you all of this before I leave here. This is not my opinion any more, we have enough proof that it is a statement of fact and we can bank on it. If you are too lazy, too stupid, too electronically impaired or too frightened to build your own, there are forty known companies (and probably a hundred that I know nothing about) that are building these things full-time and they are getting the most remarkable reports back. I do meet these people, and NO, I do not get a nickel from any single one of these companies, and all of this information is in your paper.

Now what else do we have to watch for here. Nobody has ever to this day died of AIDS. Did you know that? They have died of the opportunistic infections, pneumonia, Carkinson's sarcoma, etc. etc. that can attack and ruin your body when your immune system cannot handle the load, in fact it is called '**HIV**' (**H**uman **I**mmune-deficiency **V**irus) '**AIDS**' is **A**cquired **I**mmune **D**eficiency **S**yndrome. Now this is the only way that **I** know of, I know the people who have written the books on ozone therapy, I won't mention their names, I know the doctors who are using all of the herbs and drugs. I have talked to the people who worked with Dr David Ho who was on last week's *Time* magazine cover as the Scientist of the Year, who has discovered that if you use two or three chemotherapy agents at the same time, you have a cocktail effect and it reduces the HIV count, the PCR count in the blood tediously. But what he didn't tell you, and I think I'll read this which was in *The Los Angeles Times* December 10th 1996: "Drug combo knocks HIV down, not out". And a basic thing happened, they took some of these patients who were on combination therapy, which incidentally cost $20,000 to $60,000 per year, as soon as they stopped taking these cocktails, they got immensely elevated counts up in the millions or billions of AIDS counts, as though these chemotherapy agents, (AZT, DDI, DD5, etc.) had held this viral load down in the blood as long as they were taking it. The minute they stopped, it rebounded and many of the people were critically ill and they had to put them back on it immediately.

So, once you start on this cocktail therapy, you are stuck with it for life, as long as you live, and as you know from the International AIDS Conference in Japan last year, they found that none of these AZTs etc. prolonged the life of one single AIDS patient by even one day. Their's is simply a promise, they are engineered to get your money while they give you hope. I was beaten up and bloodied; I was hit over the head and had blood running down my face, by 'AIDS Act Up' David Buller it was that was running this crusade in New York city in February. He called the New Yorker Hotel and said that they would burn down the hotel, so they gave me a bodyguard (who wasn't around there all the time) and said 'Bob, do not use the

front door but go around the back and use the service entrance – never go out the front door of this hotel. The last day when I was there and I had to run out to get a shuttle to take me to the New Jersey airport to bring me home, there was a guy standing across the street with a video camera and a guy who hit me over the head and bloodied me, and I didn't chase him down the street because my knees aren't very good as I don't have any cartilage in them.

But when's the last time that you saw me Ivan – about three years, four years ago? Do you notice anything that's different? I lost 130 pounds – I weighed 290 and I have pictures of me to prove it. My hair on top was thinner than some of my old good buddies here. As soon as I got rid of the adversaries – the aliens in my blood, my immune system took over and I got rid of the half a pound or so of parasites which every man, women and child in America (if he's more than four days old) is carrying. A new-born baby, after four hours has breathed enough air to become infected. If you live in America, there are 140 **known** indigenous parasites. If you live in Africa or the tropics, there are 500 identified and known parasites and probably half as many again that are unknown. When we got these parasites out of my body by wearing this device on my wrist, an amazing thing happened. I began losing weight like crazy. I thought 'my God, have I got stomach cancer or am I wasting away with something?'. No, I was in better shape than I was before I started.

This is a theory: about 30% of the parasites which live in human beings, and these can range from fourteen feet (4.27 m) long if they are a tape-worm down to microscopic size (which is why they have to use microscopes to do the stool analysis if you go in for a check), these parasites were setting my appetite. If you will think about this for a minute, it is not what you eat or how much you eat that has anything whatsoever to do with your weight. It is how much of what you eat you store as fat because the parasites have changed the P51, the leptin and about five or six known neuro-transmitters in the hypothalamus to feed them tomorrow. Many of these parasites have co-existed with mammals – apes, elephants …. for twenty million years and the anthropologists will tell you that they find them whenever they find a carcass frozen in the ice floes, these things have had parasites for as long as ever man has been on the planet. These parasites have learned to use you as a meal and not kill you because then they are going to die themselves, like ebola which runs it's course in a couple of weeks. Many of these viruses do that, but these parasites almost kill you and use you for a meal in the same way that we send cattle to a feed loft before we take them to market. And when these parasites vanished from my body, and this is provable by dark-field microscopy, phase-contrast microscopy, any doctor who does this analysis can prove it, you don't have to take my word for it, I lost all of that weight and before this time I had offered $10,000 cash currency 'under the table' to anyone who could get this weight off of me.

I had tried dozens of diet plans. I had tried these canned-meal Jenny Craig type adventures. I had been injected with pregnant mare's urine. I know that some of you have been through the same things. I had taken legal amphetamines to control my appetite … I had tried everything until I had nearly killed myself. And until I found the true cause of what I believe is about 30% of the people who are overweight, and this was certainly me, I was terribly embarrassed to find the airline stewardess running down the aisle with a seat-belt extender because the seat-belt wouldn't fit around me. I couldn't sit in the booth in a restaurant because I'd have to sit on a chair, etc. etc.

The lady at the back of the room, Jane had inoperable cancer when I met her. She said "I'm going along with blood-cleansing here, and not the chemotherapy and radiation which the doctor had said that if I do not have I will surely die". And her daughter Cathy was outraged because her mother didn't do all of this chemotherapy and have her hair drop out … Now she is totally well. Her last several physicals – biopsies and CAT scans – have shown not one speck of cancer anywhere in her body. Neither have dozens and dozens of 'terminal' cancer patients.

Now, why does this work for cancer? We started out working with viral and microbe and parasitic diseases. When oncologists in the East began reporting many, many, many cancer cures, we had to have a telephone conference between about for or five of us. Now what's really happening here? I don't believe it until I see the medical proof. So I guess you know that about three months ago *Explorer* magazine published an article entitled '*Total Cancer Remissions through Blood Electrification combined with Silver Colloid*'. And this article of two pages is reprinted in it's entirety in the paper which you should be taking home with you.

Social Security does not want you to live if you get to be 65. I'm now way past that. I don't act like it, but I am way past that, I'm in my seventies. The Director of Social Security – I will not mention his name because he picks me up at the airport when I'm in Washington and drives me where I'm going – I've stayed in his house – says 'Bob, we want anyone who reaches 65 to be stored like cord-wood in an old-age home and given mind-control drugs like valium and let him die because Social Security is bankrupt **now**, it won't be *bankrupt in five years time* as it is bankrupt now. And if you restore perfect health to these people with a device that will fit in your shirt pocket, the job market will not absorb the people getting out of college today.

Even if McDonalds has a smart cash register, they can't add or subtract, they can hardly write a paragraph and you want to get these experienced people back in the jobs market?'. I said 'hell yes!' I was 65 nearly ten years ago.

Another problem: this steps on the toes of all the vested interests. They are expecting to make hundreds of thousands of dollars off of you: Health Management organisations, insurance companies, hospitals, … If you own a boat that thing is written into the budget of the boat builder down there in San Pedro. These medical people expect to see you come in there some day and spend a lot of money, but if you do this for yourself, you will get well. There are seventeen of us who I know of, and I know sixteen of them personally, who this day have immortal blood. Now this is a rather dramatic claim, so let me tell you what it means. Toward the end of last year, there was a dark-field microscopist 'Mike', you know his name, he is world famous, he's the fellow who trains doctors all over the world how to use this type of technology for diagnostic work. He invited a chap in who had been on one of these devices for about six months, and he said 'Let's look at your blood', so he pricked his finger, put that drop of blood in a microscope slide and looked at it. And he said 'you don't have any background clutter in your blood'. Of course, they couldn't see viruses as this was a light microscope with a magnification of about three thousand times, back-illuminated. He said 'just a minute – I must have done something wrong', so he stuck another finger, then he went to the earlobe and made a slide and said 'what have you been doing? anyone who has ever come to me, man, woman or child, in the last thirteen years, you can see the parasites, many of the germs, microbes and bacteria swimming around in the blood but you don't have any at all. The patient said that he had been on the blood cleaner so Dr Mike said 'give me some of those, I've got to test them on some of my patients – I don't believe you'. He did, and they worked.

Now, in Colorado, blood is considered 'toxic waste' or hazardous waste. It must be picked up and disposed of by incineration by registered toxic waste disposal people. Dr Mike had thrown these three or four slides in a box underneath his laboratory table. He looked down at this box about twenty-seven days afterwards and checked the slides again under the microscope. That blood was still alive. The half-life of human blood on an unsealed microscope slide is between two and a half to four days maximum. Twenty-seven days had gone by and the red cells were swimming in the plasma exactly the same as the minute he had drawn that blood. It had not crenated (serrated around the edge) – a fibrin reaction had not set in although it had evaporated around the edges. Now let's get to the point. One month after that, that is some fifty-one days later, those slides had not yet been picked up for disposal, so Mike checked them again. The blood had 'circled the wagons': there was about a half a millimetre spot in the middle where the cells had not evaporated and were still alive.

The people in his practice who had been tested upon with the blood cleansers have now got immortal blood. Now, what does this mean? If 'the adversary is within' he might not be a man with a red tail and horns out there somewhere with the number 666. Just as the Holy Spirit dwells within this temple, so might the adversary, at least the aliens which have been involved with us over many, many moons are in there and you can see them if you go to a dark-field microscopist, if you get rid of those, who knows what is going to happen? What happened to me. I was so fat that I was in a wheelchair and I had been given up for dead. The doctors who came to see me, friends of mine who were Medical Doctors who hadn't made a house call in Beverley Hills in thirty years, were coming down to my apartment saying 'Bob, you'd better dispose of whatever you've got – you'll never walk again. I was given so many tests. I brought Dr. Alexander Everett from Washington, paid his way, put him up in the *Red Lion* hotel. He worked on me until he said 'Bob, I'll have to go back now, there is nothing that anyone can do for you'. And I was in a wheelchair, but I'm sure as hell not dead! Everything is working well again.

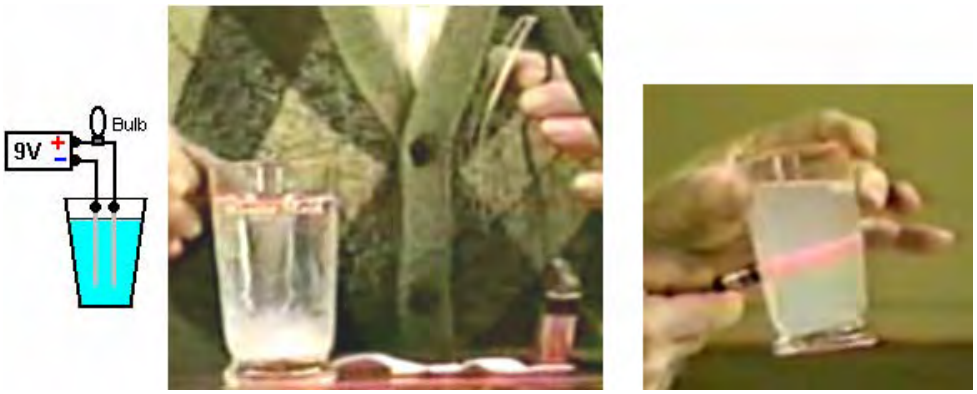
My blood pressure when I started this averaged 219 over 190 and was borderline diabetic with blood sugar 425 to 475. Today my blood pressure is 130 to 140 over 70 to 75 and after just three weeks my blood sugar had dropped to 150, which at that time was considered normal.

So I'm here to share this information with you. I have nothing for sale. There are people who are selling things, I'm not one of them. Part of that is because I believe that God's work should be given freely. I expect to get my reward somewhere else if I don't come back to this planet of painful endeavour. I have a lot of theories about this and also the Food and Drugs Administration has come to my door with drawn guns at three o'clock in the morning. I have been beaten up by the people supporting the pharmaceutical houses who don't want a $1.32 cure for AIDS.

Bob also speaks of his use of colloidal silver, (a document on which can be downloaded using the download link http://www.free-energy-info.tuks.nl/Silver.pdf) although Bob's method of production looks more simple. He says:

This is a glass of water which I got out of the kitchen. I've taken a single 9-volt battery and put a little "grain

of wheat" bulb (6 to 12 volt 55 milliamp) in series with the battery and these two silver wires. There is no known germ, bacteria or microbe which can survive what I am about to do right now. I'm going to put these two silver wires into this glass of water, about half an inch (12 mm) apart, and leave them there for about three minutes. That is all it takes. This is tap water and you can make this silver colloid for about one tenth of a cent per gallon and that is the most powerful medicine known to man. After three minutes Bob stirs the colloid with a plastic knife and demonstrates how a laser beam reflects off the microscopic particles of silver suspended in the water and then drinks the water. Bob drinks two or three glasses of colloid each day. The colloid prevents you from contract secondary infections during the first three weeks while you are getting well with the blood electrification process.



Finance: Money only has value when backed by something of worth. The largest purchase that most people make is buying a property. This is often done by taking out a mortgage with a financial institution such as a bank or a Building Society. For this, most people sign a mortgage agreement with the bank, agreeing to pay back the mortgage amount plus a large amount of interest, generally over a period of many years. If you default on the repayments, typically after many years of payments, the bank takes your property and sells it in order to gain as much money as they can. What you are not told is that the mortgage agreement is a contract and no contract is valid unless something of worth is offered by both parties and full disclosure of all circumstances is made. The bank does not offer anything of value as it just invents the money, creating it out of thin air and without any backing of something of worth. The bank regulations actually forbids the bank from lending out money deposited by it's members, so a bank cannot legally enter into any such mortgage agreement backed by the funds deposited with it. Further, no contract is valid unless there is a "wet ink" signature by both parties and a wet ink signature can only be made by a human. A bank cannot make a wet ink signature, nor can some other individual sign a contract on behalf of somebody else, so, no mortgage agreement is actually valid in law. Also, the bank being aware of this and not making it clear to you the mortgagee, is in breach of the "full disclosure" requirement of a contract, which is enough to invalidate the contract anyway. Therefore, if a bank tries to take your property for lack of payments, they are acting unlawfully and are not entitled to do so. They usually get away with it, because most people are not aware of the real facts involved.

The Media: You are expected to believe that the media is an honest and unbiased source of information. The reality is that all of the main media outlets are owned by just a few, very rich individuals, and their objective is **not** to provide unbiased reporting. It appears that they have three main objectives.

First, they want to make money and objective and truthful reporting does not necessarily help towards that goal. You will never, ever, find honest reporting of anything to do with free-energy or any other positive thing which individuals can do for themselves. For example, if a "report" is made on say, an HHO booster to improve mpg and reduce harmful emissions dramatically, it will be arranged that the supposed results show that these thing are not worthwhile. Contrast that with the reality that Dr Scott Cramton can use HHO gas from the electrolysis of water, to reduce the fuel requirement of any diesel engine by at least 60%. No matter how you try, no major news outlet will every report the honest facts as that could damage the excessive revenue stream of the oil companies.

Second, they want to generate as much fear and worry as possible because doing that over an extended period, causes the average person to be more amenable to having their freedoms taken away from them. It facilitates the introduction of surveillance on everybody, identity papers (birth certificate, driving licence, passport, social security number, voting register, National Census record, etc. – very much in the Nazi occupation style), going to war with countries which have never harmed you, considering people to be dishonourable just on the basis of their nationality, and the like. Try counting the items reported in the daily

news which are negative – murders, attacks, dishonesty, deaths through accidents, deaths through natural disasters, kidnappings, piracy, violence, environmental damage, etc. and compare that number to the count of positive things reported on – happy events, acts of kindness, successful ventures, etc. and see what ration you get, day after day, after day. This is not just sales being boosted because people love to read about disaster. Instead, it is a deliberate, long-term policy of psychological warfare against ordinary people.

Third, they deliberately divert the attention of people away from anything of real importance by promoting and encouraging excessive interest in things of no importance. There is a deliberate dumbing-down of almost all TV programs in the UK, the promotion of excessively violent video games, the promotion of trivia such as celebrities and their activities, dieting, gossip, fashion, and other irrelevant topics. They want you to give all of your attention to these things and not notice what is actually going on all around you.

Joseph Cater's Comments.

Joseph H. Cater has produced a book entitled "Ultimate Reality" which I find impossible to buy at any reasonable price level. In it, he points out many things which he supports with strong evidence. These things seem startling because, and only because, the present educational system deliberately encourages us to believe things which are clearly not true. Google books have a partial copy of the "The Ultimate Reality" at:

http://books.google.com/books?id=ySsNiCPUPecC&printsec=frontcover&dq=the+ultimate+reality+cater&cd=1#v=onepage&q&f=false

The things which Joseph Cater states are so unusual that it would be easy to write him off as a crank. However, he backs up what he says, with a considerable body of realistic evidence which makes it very difficult to ignore what he says, in spite of the fact that most of his findings flatly contradict what we have been taught from an early age, and so accept as being reality. Whether or not you accept what he says is entirely up to you, but it is difficult for an honest person to reject his presentation out of hand.

He puts forward a strong case for there being a deliberate programme of scientific misinformation and suppression aimed at keeping the general public completely ignorant as to the actual physical realities of the solar system and the universe in general, and as a result, reality is very far from the popular conception. Mr Cater's description of matter provides reasoned explanations for a whole range of anomalies which conventional science can't adequately explain, and he quotes numerous experiments which provide firm evidence that what he is saying has a substantial basis in fact.

Here is a very brief summary of some of what he says in his book "The Ultimate Reality":

The biggest single factor is in the erroneous theory of sub-atomic particles. Mr Cater states that reality is actually much more simple than conventional theory suggests. The universe is filled with a range of energetic particles which he describes as "higher order ethers" and "lower order ethers". These ether particles are in continuous random movement at different frequencies and they produce a number of different kinds of composite particles including "Hard electrons" (with which we are already familiar) and "Soft electrons" which have very different properties. Soft electrons can draw hard electrons into themselves, masking the properties of those hard electrons. Combinations of ether particles form photons and all matter is composed of photons and these ether particles.

The operational forces which govern all matter in the universe are only electrostatic force and magnetic force, and the actions of both of these are modified by many different combinations of the two types of electron. As light is composed of photons and as they interact with the two types of electrons, many observed scientific facts have led to wholly incorrect deductions. Mr Cater indicates that the New World Order "elite" carefully foster and support these false conclusions, suppress information and physically alter observations before they reach the public. Mr Cater points out a number of instances where the alteration of data has not been sufficient to suppress the facts.

It is not possible to mention all of the points which Mr Cater makes, so please understand that the following is only a minor selection from a cohesive whole and much of the supporting evidence which he provides in his book is omitted here.

Mr Cater says:

1. The current theory of gravity is completely wrong, and gravity is caused by a component of the electromagnetic spectrum of about one trillion cycles per second (0.3 to 4.3 mm wavelength; located above

radar and below the infra-red region). The theory of gravity put forward by Sir Isaac Newton does not account for the **amount** of deflection of a plumb bob near a mountain. An asteroid as small as 150 miles in diameter, can have a surface gravity about the same as Earth's and some asteroids have moons of their own orbiting around them. This would be impossible if Newton was right.

As a result of the real cause of gravity, our Moon has a much higher surface gravity than was expected, (and consequently, a serious atmosphere). NASA had a major problem with the lunar landing of 1969 because if the real facts became known, then it would demonstrate that a major section of physics, as it is currently presented, is incorrect and they want to keep things exactly as they are at present. The Moon, in common with most planets, is not solid. When part of the lunar lander was dropped back on to the surface of the Moon, a seismic detector left on the surface showed that the Moon vibrated like a bell for about an hour afterwards. That would not have happened if the Moon were solid.

There is very clear evidence that NASA has censored the Moon landing very heavily, but their attempts at suppression have been flawed and some obvious pointers exist. For example, NASA claimed that the space suits of the astronauts weighed 185 lbs when on Earth. Photographs exist, showing an astronaut leaning over backwards and then regaining his balance. That would be a physical impossibility even under reduced gravity, and this implies that the "life support" systems were actually empty and not needed because there is an atmosphere. The high surface gravity is also seen when the astronauts ran. Even the slowed-down version released by NASA can't conceal the length of the steps and the height off the ground which are the same as they would be on Earth.

If the Moon had one sixth of the Earth's gravity as is claimed by conventional science, then the point at which the gravitational pull of the Earth balances that of the Moon would be about 22,000 miles from the Moon. The Encyclopaedia Britannica states this distance as being about 40,000 miles, which agrees with various other sources. That could only be so if the Moon's surface gravity were much higher than the supposed one sixth of Earth's gravity.

On more than one occasion, an Apollo astronaut tripped and fell on his face. Under one sixth gravity, that just would not happen, especially with a fit and active astronaut. Also, the rover vehicle used in later landings was 10' long, 7.5' wide and 4' high with 32-inch diameter wheels. The Earth weight was 460 pounds and under one sixth gravity, that would only be 75 pounds, but the astronauts had great difficulty unloading it from the lunar module. Engineers on Earth had already determined that to operate under one sixth gravity, the rover would have to be 20' long and have a 20' tread. With a loaded earth weight of 1,600 pounds, it would need a turning radius of well over 80 feet to avoid tipping over at 10 mph or more than 20 feet at 5 mph and descending steep hills would not be possible without major problems. But, the astronauts did descend steep hills and they made very sharp turns at maximum speed.

One of the photographs brought back by the Apollo 12 trip, showed an astronaut carrying instruments hanging from a bar. The instruments had an Earth-weight of 190 pounds, supposedly 31 pounds on the Moon, but the pronounced bowing of the bar would not have been caused by just 30 pounds.

During the early Moon trips, the astronauts stated that when they left the atmosphere, the stars were not visible. This is understandable as the atmosphere scatters starlight, making stars appear larger and so become visible to the naked eye. Outside the atmosphere, there is no scattering and the stars are too small to be seen without a telescope. On the Apollo 11 trip, shortly before reaching the Moon, Collins stated "Now we're able to see starts again and recognise constellations for the first time on the trip. The sky's full of stars ... it looks like it's night side on Earth". This demonstrates that the Moon has a significant atmosphere caused by much higher gravity than one sixth that of Earth, although the refraction of light through that atmosphere is less than the refraction caused by Earth's atmosphere.

Mr Cater points out that NASA is well aware of the real nature of gravity and had electrogravitic drives long before the 1969 Moon shot. Not only that, but due to the fact that the Moon is bigger than currently believed, further away and possessing a higher surface gravity, that rocket power used during the flight was supplemented by an electrogravitic drive. Any honest person who has studied the evidence is well aware that there are many craft with electrogravitic drives and most of these seen in the last sixty years, are man-made. (All governments are very keen to suppress this information as national boundaries could not be maintained if electrogravitic drive vehicles were available to the public).

2. Relativity, proposed by Albert Einstein, is not correct and Mr Cater spends quite some time demonstrating that relativity is wrong. Prior to Einstein, the Transverse Wave Theory of light was universally accepted. Waves cannot exist without a medium which vibrates in some manner to transmit them. Therefore, the "ether" which permeates all of the universe was accepted. The Michelson-Morley experiment was set up to

test this. A ray of light was split into two parts which were made to traverse different paths of equal length. The motion of the Earth through the ether should then cause the recombined rays to show diffraction patterns. They didn't. It did not occur to anyone that if light were retarded by passing through the ether, then bodies such as planets would be seriously retarded by their passage through it and would slow down and stop. This experiment also gave rise to the ridiculous idea that the speed of light is a constant, in spite of the well-known fact that the speed of light through water is only 75% the speed of light through space. It was also proposed that the speed and direction of movement of an observer didn't matter, That time slows down on a moving system, that a body will shorten in the direction of motion and that the mass of a body will increase the faster that body moves. These are ridiculous suggestions. The famous $\mathbf{E = mC^2}$ equation was actually derived from the Lorentz equations in 1903, two years before Einstein got into the act.

Physicists argue that particle accelerators demonstrate the increase of mass with speed. This is not the case and the experiments actually demonstrate a very important principle which provides a better understanding of many physical phenomena. It is an established fact that a magnetic field develops around a charged body when it is given a velocity. Where did the magnetic field come from? In the particle accelerator, as the particles accelerated, magnetic fields developed around them. As the total energy of the system remains constant, the magnetic field must have developed at the expense of the electrostatic field - the transformation of one kind of energy into another kind. This conversion from repelling electrostatic charges to magnetic fields causes the particles to clump together, giving the false impression of an increase in mass. Further, as the electrostatic component drops to almost zero, the accelerating force diminishes to near zero also, giving the false impression that a material body can't travel faster than the speed of light. The reality is that bodies can travel many times the speed of light.

According to General Relativity, a gravitational field will tend to slow the passage of time and the stronger the gravitational field, the more marked the effect. It was found that Caesium clocks run faster at high elevations than they do at ground level. This has been taken as a proof of the validity of Einstein's ideas. The concentration of soft particles is higher near the ground than at high elevations and that makes clocks run faster at high elevations. As to the speed of light not depending on the velocity of it's source, the Sagnac experiment of 1913 provides direct proof that the observed velocity of light **is** dependent on the velocity of it's source, disproving Relativity. Mr Cater provides extensive demonstrations (as do other people) that Einstein's deductions are not correct.

3. It is clear that gravity is responsible for the tides, but the standard explanation is wholly inadequate, being based on the assumption that gravitational effects have unlimited penetration. In other words, the only attenuation when passing through matter is due to the inverse square law, which actually would be a violation of the law of conservation of energy.

It is well known that a body cannot be given an acceleration relative to another body if both bodies experience the same acceleration. It follows then that since large bodies of water are accelerated relative to the Earth to cause tides, the water is being experiencing a different acceleration than the Earth as a whole, otherwise, there would be no tidal movement of water across the surface of the Earth. Assuming that gravity has unlimited penetration causes problems when accounting for tidal movements. Since the distances between the Sun and the Earth and the Moon and the Earth are large in comparison to the diameter of the Earth, all parts of the Earth will experience nearly the same gravitational attraction from these external bodies if gravity has unlimited penetration. High tides tend to occur when the Moon is at it's zenith, both directly underneath the Moon **and** simultaneously, on the opposite side of the Earth.

The Earth's orbit is inclined to the Equator by 28 degrees and so the Moon is never further North or South than 28 degrees. According to Newton's theory, the highest tides should occur near the Equator but the reality is that the highest tides are experienced much further away from the Equator, both North and South of it. Mr Cater provides an in-depth discussion of these effects, demonstrating that Newton's concept of gravity is wrong.

4. It is generally accepted that energy, in any form, flows from a higher potential to a lower one. The law of redistribution of energy states that when radiant electromagnetic energy interacts with matter, the resulting radiation as a whole, is of a lower frequency than the original light. This is why temperatures at lower elevations are generally higher than those at higher elevations, as sunlight passing through air converts to lower frequencies including infrared which activates the thermal motion of atoms and molecules, thus producing heat. Any dynamic unit is less active as a whole, than the individual parts comprising it. The higher ethers consist of the smaller, more active particles while the lower ethers are composed of the large, more complex and consequently less active particles. Both ethers occupy the same 3-dimensional space (which is the only space that there is).

When light of a given frequency range is produced, only the ethers associated with this light are directly activated. Light photons are composed of combinations of ether particles. Photons combine to form the penetrating particles which accompany this light. Particles composed of light in the lower frequency ranges are referred to as “soft” particles while those associated with gamma rays and above are referred to as “hard” particles.

Soft particles are more penetrating than the photons from which they are made, because, unlike the larger soft particles, photons have a relatively great surface area in proportion to their mass. Soft particles, and particularly soft electrons, play a vital role in all life processes and in other chemical reactions. The energy or field intensity of and around the higher ether particles is greater than that of the lower ethers. The diameter of a particle is inversely proportional to the average frequency of it’s constituent light.

5. The energies radiated from the Sun are continuously transformed into ever lower frequencies as they penetrate into the Earth. In this manner, nearly all of the original ultraviolet is transformed into lower frequency radiation by the time it penetrates the shell of the Earth. It is the transformation of some of the radiation from the Sun into gravity-inducing radiation which holds the Earth and the other planets in orbit around the Sun and give the illusion that the Sun has about thirty times Earth gravity. It should be mentioned that soft particles penetrate solid matter more readily than hard particles which are, of course, an integral part of matter.

All matter continuously radiates soft particles of many different kinds due to the interactions of the fundamental particles. These radiated particles undergo a transformation effect, according to the transformation law, when passing through large concentrations of matter. When this occurs, some of the radiation is transformed into gravity-inducing radiation. This is the source of some of the surface gravity of both the Earth and the Moon. The greatest contributing factor to Earth and Moon gravity is the transformation of radiation resulting from the thermal agitation of atoms and molecules. The particles resulting from this activity are comprised of lower-frequency photons. Such radiation is more readily transformed into gravity-inducing radiation because it is closer to this frequency band to begin with. A significant portion of this radiation, originating miles below the surface, is converted into gravity-producing energies by the time it reaches the surface. Most of the gravity radiation of the Earth and the Moon is created in the topmost fifty miles of their crusts. Below that level, much of the energy from the Sun has been transformed into softer particles, and that material of the Earth and Moon is permeated with them.

These soft particles screen out gravity radiation more effectively than solid matter does. This is because the ethers with which they are associated, are closer in frequency to the gravity radiation band. This explains why Moon gravity is nearly equal to Earth gravity. At the same time, it is clear why the Cavendish Experiment for determining the so-called “gravitational constant” was misleading – there wasn’t enough material in the bodies used in the experiment to produce any radiation transformation. The gravitational effects produced by the bodies were due entirely to the thermal agitation of the molecules, without any transformation of radiation. The thermal agitation of molecules produces infrared and only an infinitesimal portion of this radiation is in the gravity-producing frequency range. This “gravity constant” plus the idea of unlimited gravity penetration, requires scientists to assume that the Earth has a tremendous mass and an iron core four thousand miles in diameter.

It is significant that some of the Cavendish Experiments indicated that gravity effects varied with the temperature. When the large sphere used ion the experiments was heated, the smaller sphere had a greater tendency to move towards the larger sphere. When the larger sphere was cooled, the smaller sphere receded. This was explained away as being caused by convection currents although they failed to explain how convection currents could produce such an effect. A detailed account of this can be found in the 11th edition of the *Encyclopaedia Britannica* in the section entitled “Gravity”. (If they felt that air currents were skewing the results, then the experiment should have been repeated inside a box which had the air removed).

As mentioned before, matter produces infrared radiations which are partially transformed into gravity radiations. In the case of mountain ranges, there is not enough matter to transform significant portions of such radiations into gravity radiations. Much of the radiation will escape from the tops and slopes of the mountains before they can be transformed, since their average heights are generally small compared to their horizontal dimensions. The gravity radiations produced deep in the interior of the mountains are partially dispersed by the overlying mass. This is the cause of the plum bob enigma which is a source of annoyance to conventional physicists because the plumb bob is not pulled towards the mountains to the extent demanded by Newtonian laws.

Another problem is that, in comparison to the Sun, the Earth radiates only an infinitesimal amount of

radiation per unit of surface area but it is able to keep the Moon in it's orbit around the Earth. Even allowing for infrared radiation passing from the Earth to the Moon and converting to additional gravitational radiation there, it still wouldn't be enough to keep the Moon in orbit unless the Moon were hollow and had a shell not more than a hundred miles thick.

In 1978, scientists were shocked to discover that some of the asteroids have moons which orbit around them at respectable velocities. According to Newton, this is impossible as the gravity of an asteroid would be far too feeble to allow this. When a body is a few miles across, it is large enough for gravitational radiation to be produced. This effect increases rapidly as the size of the body increases as far more infrared is transformed than is screened out by the outer layers of the mass. The effect continues until the body is about 150 miles in diameter and beyond that point, the screening effect of the outer layers keeps pace with the rate of increase of the transformation of infrared into gravity radiation. This means that all planets have practically the same surface gravity.

6. Mr Cater explains how soft and hard particles and the limited penetration of gravity account for Erath upheavals, continental drift, earthquakes and volcanoes. He also remarks that if the Earth were a completely solid ball and the Newtonian version of gravity were correct, then the Earth would be completely rigid and no Earth changes would occur other than some minor erosion, and there would certainly be no mountains left by now.

7. One of the most fundamental physical laws involves the relationship between the electrostatic and magnetic fields. One transforms into the other and vice versa. Inertia is a third factor involved in the relationship between the electrostatic and magnetic fields. The kinetic energy of a moving charge is manifested in it's magnetic field. The magnetic field increases at the expense of it's electrostatic field (as dictated by the law of conservation of energy). The role of inertia and the conditions governing it's magnitude are now apparent. The inertia of a body is dependent on it's ability to generate a magnetic field when it is given a velocity. The greater the inertia, the greater this ability.

The magnitude of the inertia of a body is directly proportional to the energy of the magnetic field which the body develops for a given increase in velocity. It follows then that inertia is dependent on the total electrostatic charge of a body. This is also true for so-called "uncharged" matter. In the supposedly uncharged state, all atoms and molecules have a net positive charge. Therefore, even atoms and molecules develop a magnetic field when they are given a velocity.

In 1901, Max Planck found that he could only derive the correct distribution in frequency of the radiant energy in the cavity of a black body as a function of the temperature of that body, if he assumed that energy exists in discrete units. He came up with NHV where N is an integer, V is the frequency of the light involved and H is a universal constant (expressed in terms of energy multiplied by time, that is, erg-seconds). This is now known as Planck's Constant and is 6.6×10^{-27} erg-seconds.

The kinetic energy of a light photon is inversely proportional to the frequency. The lower frequency light, consists of larger and more massive photons travelling at the same velocity as the higher frequency photons. On average, the number of photons in any given ray, and the number of accompanying soft electrons will be a constant, regardless of the frequency. This is in accordance with the conclusion that the average distance, or mean free path between ether particles of the same kind, is a constant, regardless of the ethers involved. The average number of photons comprising a soft electron will also be independent of the frequency. This means that the diameter of the surface area of a soft electron, will also be inversely proportional to the frequency. Soft electrons accompanying light, travel at a velocity which is less than that of light. The soft electrons pick up speed, by bombardments of faster moving photons.

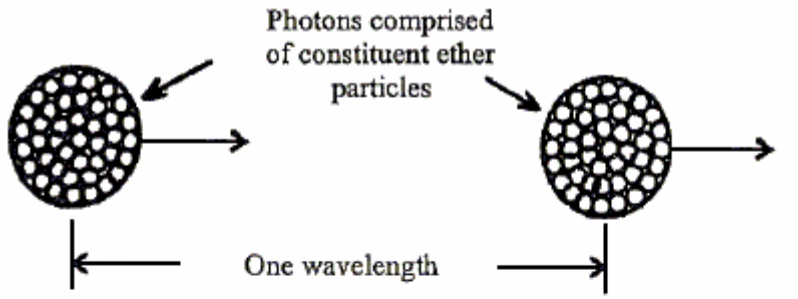
From a superficial glance, it seems that the average velocity of soft electrons should be independent of the frequency of the light associated with them. This is not so. The soft electrons associated with the higher frequency, travel at a higher velocity, and herein lies the key to the photo-electric effect. Although the lower mass of the higher frequency soft electrons is offset by the lower kinetic energy of the bombarding higher frequency photons, the surface area is greater in proportion to mass. This means that in proportion to mass, the electrons associated with the higher frequency light will receive a greater bombardment of photons and so, a greater accelerating force. The ratio between surface area and volume, or mass, is inversely proportional to the ratio between the diameter of two given spheres. Since the other factors balance out, it follows that the resultant average kinetic energy of soft electrons in proportion to mass, is directly proportional to the frequency of the light with which they are associated. As soft electrons collide with a surface, the hard electrons which they contain, are released and they bombard the surface, producing the photo-electric effect. They will be travelling at the same velocity as the soft electrons which housed them, so their average kinetic energy will be proportional the frequency of light.

Quantum mechanics is considered the most monumental achievement of twentieth century physics. In view of the principles presented above, it is not surprising that mathematical juggling with Planck's constant would account for many experimental results (in a quantitative sense). Quantum mechanics experts have enjoyed considerable success in this respect, especially in the realm of atomic spectra, without knowing why. In reality, quantum mechanics does not even qualify as a theory or a concept. It is merely an attempt to give mathematical descriptions of certain phenomena with Planck's constant and his valid assumption as a starting point. Modern "theoretical" physicists have absolutely no conception of why their mathematics agrees with certain experimental results. Yet, they have led themselves to believe that by giving mathematical descriptions of such phenomena, they have actually explained them.

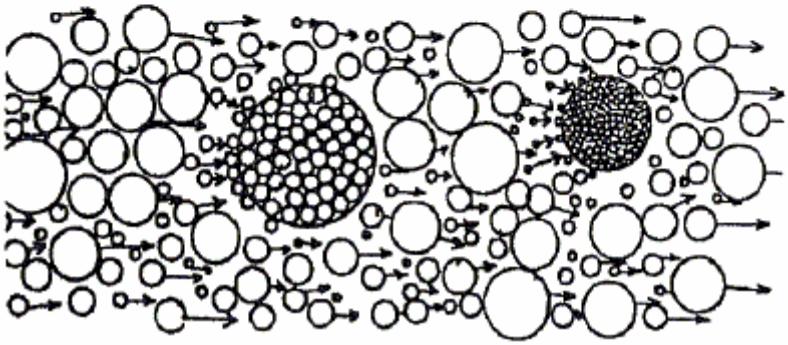
It now becomes evident, why a mass can travel through space at a constant velocity, and encounter no decelerating force. The ether particles are so active that the closing forces at the back of the moving body, tend to equal the resistive forces encountered at the front. The rear portion creates a temporary void which is rapidly filled in by the surrounding ether particles, producing an effect very much like the Coander Effect. During the filling in process, the fundamental particles comprising the rear of the body are bombarded with ether particles travelling at a higher velocity than is normal. Also, the ether particles of which the mass is comprised are so relatively sparsely distributed throughout space, the situation is equivalent to a great mass travelling through a highly rarefied atmosphere.

8. During the creation of a photon, the ethers in the vicinity are suddenly compressed. Some of the ether particles are forced close enough together to adhere to each other. This aggregate is then propelled outwards with great force in a manner similar to a compressed spring being released. The photon reaches the speed of light after this accelerating force has been expended, which happens in a distance equal to the so-called wavelength. This process is repeated in the same region and another photon is produced which follows the first one, just one wavelength behind. A wide range of ethers are periodically affected during the production of ordinary light. This results in a countless variety of such particles being propagated in all directions with many different wavelengths. Since many photons are projected in all directions, many collisions will result, causing a significant portion to adhere to each other in aggregates.

The great majority of soft electrons are created during fluctuations in light velocity when passing through media of varying density, and even in it's passage through outer space. Any slowing down, produces a backing up of photons and a consequent combining into relatively huge aggregates. In the beginning, these aggregates move much more slowly than the free photons. Consequently, some of the photons which were created at a later time, catch up and adhere to the aggregate. Their collisions with the aggregate particles causes the particles to speed up. This is the origin of the particles which always accompany light. Particles formed in this manner will vary greatly in size, stability and penetrating ability. It has been shown that soft particles will penetrate ordinary matter more readily than the hard particles. So, ether particles combine to form photons which in turn, combine to form light particles. This, light particles will disintegrate into photons.



A typical beam of light
made up of photons and soft particles



Atoms are comprised of hard particles which are uniform in size and structure and it follows that they were produced by an entirely different process. When light enters a medium, it encounters a conglomerate of soft particles created by the activities of the fundamental particles of the atoms which comprise the medium. This causes the light to slow down and the particles of light to crowd together inside the medium. If a beam of light enters a medium at an angle, the portion entering first will travel a shorter distance than the rest of the beam during the same interval of time. The portion of the beam entering the medium later is pulled by magnetic attraction of the particles, towards the side which reached the surface first. This causes the beam of light to be bent or change direction, accounting for the refraction of light which has never before been adequately explained.

Mr Cater then goes on to resolve the famous wave-particle paradox, and the also points out that the famous Michelson-Morley Experiments actually disprove the Transverse Wave Theory of light.

9. It must be realised that nature's laws are basically simple. To gain a deeper understanding of the nature of electrons, protons and electrostatic forces, it is necessary to look for an uncomplicated picture of the fundamental particles and the cause of their behaviour patterns. The collision laws involving the molecules of a gas can be applied to the ethers. Also, it can be deducted that electrostatic forces are the result of an imbalance of ether particles bombarding fundamental particles of matter.

It seems logical to assume that electrons and protons have a spherical shape as a sphere is the most stable and efficient geometrical form. It also has the smallest surface area for any given volume. However, such an assumption leads to insurmountable difficulties. Electrons and protons have a preferred direction of spin in relation to their direction of motion. The electron follows the left hand rule, while the proton spins according to the right hand rule. With a perfect spherical shape they could not have any preferred direction of spin. However, the preferred directions of spin can be readily accounted for if the particles are pear-shaped or egg-shaped and they are hollow.

When ether particles have a preferred direction of motion away from the electrons due to reflections, a pulsating electric field results. The excessive flow away from the electron tends to reduce the bombardment of incoming ether particles. A temporary low ether pressure around the particle is a result of this and in turn, this reduced pressure reduces the reflections and that causes the ethers to move in again and a sudden increase in ether bombardment results. This is something akin to the Coander Effect. The cycle is then repeated. It is to be expected that an electrostatic field is no exception and in this respect, "electrostatic" is a misnomer. The fluctuations are at such a high frequency that experimental results will see the (average) force as being a constant.

The behaviour of beams of electrons and protons in strong magnetic and electric fields indicates that protons have about 1836 times the inertial mass of electrons. Inertia is directly proportional to charge, indicating that the total charge of a proton is 1836 times as great as that of an electron. The idea that the hydrogen atom consists of one electron and one proton has never been questioned. To quote from a science magazine: "When protons crash into each other, they release showers of electrons, which suggests that protons are made up of particles more basic than themselves".

On the basis of relative charge effects alone, it follows that a hydrogen atom, instead of having only one electron orbiting a proton, that there are at least 1836 orbiting electrons. However, since the proton has relatively little movement in comparison to the electron, a far greater percentage of the electrostatic field of the electron has been transformed. This means that in order for the hydrogen atom to have close to a neutral charge, there must be thousands of electrons in one hydrogen atom. This seems to create a paradox as the amount of electricity required to liberate a given amount of hydrogen in electrolysis indicates that only one electron is necessary for every atom of hydrogen.

Scientists have never comprehended the source of the electricity that powers electrical equipment. There are unlimited quantities all around us, permeating all of known space. This hard electricity is camouflaged by softer particles which are also distributed throughout space. The flow of this limitless source of electrons can easily be set into motion. The electricity employed in electrolysis merely triggers the flow of far greater quantities. Also, when a hydrogen atom is ionised, it only needs to lose a very small percentage of it's electrons instead of being reduced to only a proton.

Matter is rendered visible by the steady formation of soft particles generated by the activities of the fundamental particles. It is then apparent that frozen hydrogen would be completely invisible if electrostatic fields were not cyclic and the hydrogen atom had only one electron. Cyclic electrostatic fields are largely responsible for the complex spectral pattern of all of the elements. The cyclic pattern of hard-particle interactions is complex. This complexity increases rapidly as the number of fundamental particles in the atom increases.

Since electrons move at much higher velocities in the atom than protons do, they cover much more territory and so a higher percentage of their electrostatic charge is transformed into magnetic energy. This means that the positive charge in the atom will overbalance the negative charge and so, give the atom an overall positive charge. This explains why electricity tends to move towards ground and the Earth must posses a positive charge.

The electrostatic field effects near the atom in close proximity to the electrons, will be negative. Moving outwards, this negative effect diminishes quickly and a zone of positive field effect exists. The position and intensity of these zones, determines in part, the chemical and physical properties of the atom. There are regions where the atoms will attract each other and regions where they will repel each other. Ether particles have a similar structure and follow the same pattern.

The velocity of orbiting electrons in atoms is not uniform. There are periodic fluctuations resulting from mutual interferences within the atom itself and from adjacent atoms, in addition to the pulsating electrostatic fields. It must be noted that the properties of the atom are not observed individually, but as a collective group. The region of activity for the protons is relatively small and a significant number of electrons are trapped here. This region is the origin of neutrons, which are actually collapsed hydrogen atoms. It is interesting to note that when hydrogen is subjected to ultra high pressures, it behaves like a high concentration of neutrons and passes through the container which is being pressurised as though it didn't exist.

A more detailed discussion of the structure of the neutron is in order. The new concept of thousands of electrons comprising the hydrogen atom (to say nothing of the other atoms), provides, for the first time, a means of accounting for the properties of the neutron.

When a cloud of electrons orbiting the proton is forced into close proximity with the zone of repulsion, as described earlier, their motions become restricted. As a result, there is a lowering of the average velocity with a consequent increase in their negative electrostatic charge. This provides a stronger bond between the proton and the electrons. The orbital speed cannot be increased because of the zone of repulsion surrounding the proton, and the crowding of the electrons. The higher overall negative charge of the electrons almost completely cancels out the positive charge of the proton. The result is a particle which is electrically neutral, as far as most experiments can determine.

The electron cloud comprising the hydrogen atom is further removed from the proton and the individual electrons are not restricted in their orbital motions. The average velocity is much higher and consequently, the hydrogen atom has a high positive charge. The atoms of the gaseous elements, such as hydrogen and oxygen, are highly magnetic. Therefore, two atoms combine in much the same way as two bar magnets, to form a molecule consisting of two atoms. This is the reason why the molecules of nearly all the gaseous elements consist of two atoms. The combination has a still higher overall positive charge than a single atom has. As a result of this, the molecules have a strong mutual repulsion which keeps them widely separated at normal temperatures and pressures. Thus, they remain a gas even at extremely low temperatures.

The presence of electrons in the "nucleus", nullifying repulsive forces, along with the magnetic fields resulting from the motions of neutrons, is the major source of the so-called "mysterious force holding the nucleus together". In reality, the pinch effect of magnetic fields is the prime force which holds the atom together. Orthodox physicists have complicated the picture by claiming that many different forces exist: magnetic, electrostatic, gravitational, nuclear, and others to which they have ascribed odd names. In reality, only electrostatic and magnetic forces exist and there are two, and only two, basic particles – electrons and protons. Since the electrostatic field effects around the electron and proton are cyclic, the magnetic fields which they generate will also have a cyclic intensity.

10. Although neither spin when at rest, both the electron and the proton start to spin in a definite direction when they are given a velocity. This is contrary to the assertions of modern theorists who talk about particle spin with reckless abandon. The electron always follows the left-hand rule, while the proton follows the right-hand rule.

When placed in an electrostatic field, they move in such a manner that the large end is facing in the direction of their motion, regardless of their original orientation. The reason for this is not difficult to discern. If they are hollow and the shell is of a certain thickness in proportion to it's diameter, then the larger end will have more surface area in proportion to it's mass than the smaller end will have. The thickness of the shell at the smaller end will be much greater in proportion to it's diameter. This means that ether bombardment at the larger end will tend to give it a greater acceleration than that imparted to the smaller end and as a result, the larger end will be forced ahead in the direction of motion.

The picture is still incomplete. In order for the particle to have a preferred direction of spin, the frontal surface must be grooved in the manner of a right-hand or left-hand screw. Such a shape is consistent with recent experiments at the *Argonne National Laboratory*, which studied the shattering of proton beams aimed at target protons. The results indicated that protons are not spherical. A detailed account of such experiments can be found in the article "The Argonne Experiments and The End of Quarkery" by Eric Lerner which appeared in the Oct-Nov 1997 issue of *Fusion Magazine*. In the article he showed that some of the basic assumptions of quantum mechanics are contradictory, and he dispensed with the popular theory in particle physics which assumed an ever-growing family of hypothetical particles called "quarks".

It has been noted that a magnetic field surrounds a moving charge. The magnetic lines are in the form of circles. An electron or proton tends to carry ether particles around with it in a circular motion as it moves through the ethers. This is due to the mutual repulsion between the ether particles and the ether particles comprising the particle. The reactive forces cause the particle to spin and they produce a vortex motion in the ether itself. The greater the velocity of the particle, the faster it spins and the more ether particles are caused to flow around it in the direction of the spin. It is this flow of ether particles around a moving charge which produces the magnetic field effects observed. A three-dimensional view of this magnetic field shows that it resembles a corkscrew spiral or vortex.

The ether particles which would normally cause repulsion between two adjacent particles at rest, spin when they both move and the electrostatic repulsion drops off and is replaced by a magnetic field which draws the two particles together. This effect is also seen in two adjacent wires carrying a heavy current flowing in the same direction. The wires are drawn towards each other.

If two unlike charges move along together, they spin in opposite directions, generating magnetic fields of opposing polarity which tends to push the particles apart.

An electron or proton moving in a magnetic field has two forces acting on it. One force tends to push it down the magnetic lines of force because of excessive ether particle bombardments in one direction of flow. The other force is perpendicular to the lines of force. If the velocity of the particle is high, then the latter force is by far the more significant. This force is a result of Bernoulli's principle. Magnetic fields tend to capture large quantities of soft electrons.

11. The conventional theory of geomagnetism lacks merit. According to it, the major portion of geomagnetism is the result of electric currents flowing in a molten iron core 2,000 miles beneath the surface of the Earth. Even if such a core did exist, the conclusion would still be false. Scientists are somewhat vague as to how a magnetic field could extend 2,000 miles beyond an electric current considering that it takes a very powerful current to produce even weak magnetic effects a short distance from the current flow. The electrical resistance of iron at the alleged temperatures of the core would be staggering, which raises the question of how the necessary massive potential difference is produced to drive a large current in the core in the first place.

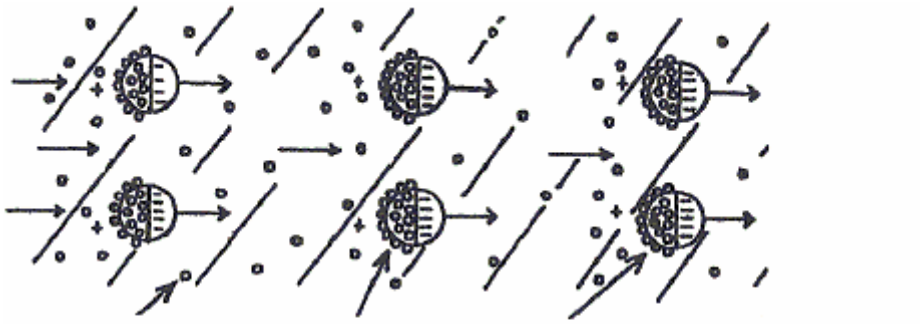
A great wealth of evidence supports the conclusion that geomagnetism is produced by the Earth's rotation. The intensity of the field is dependent on the concentration of negative charges in the atmosphere and the crust and on the rotational velocity. Since the concentration of charges in the atmosphere fluctuates in a 24-hour cycle, the magnetic field can be expected to fluctuate accordingly. This is an established fact.

Supposedly uncharged atoms and molecules are not electrically neutral, but possess a positive charge. It has always been assumed since the days of Newton, that inertia is directly proportional to mass. This has been shown to be incorrect as inertia is dependent on total charge and is therefore independent of mass. It follows that an atom has less inertia than any of the fundamental particles of which it is composed. The small overall charge of an atom is the result of the equalising of positive and negative charges. It's slight ability to generate a magnetic field for a unit increase of velocity is due to electrons following the left-hand rule while protons follow the right-hand rule. The inertia of an atom is limited because the magnetic fields of the electrons and protons from which it is constructed, cancel each other out to a major degree. Stripping electrons from an atom will give it a strong positive charge and much greater inertia even though it now has less mass. Adding electrons to it can also raise it's inertia if the extra electrons cause it to end up with a greater overall charge than it had before. The Nobel Prize winner Gabriel Lippman confirmed this when he found that bodies in a charged state have greater resistance to acceleration than they have in their uncharged state.

Since matter behaves like a positive charge, it follows that gravity radiation will accelerate positive charges in a direction opposite to that of the direction of its propagation. A gravitational field repels negative charges. When the hair on a person's head is given a strong negative charge, it stands straight up due to the negative charges on the hairs carrying the hairs with them as the charges are pushed upwards by gravity.

The bulk of the radiations and soft particles of matter cover only a relatively narrow part of the electromagnetic spectrum. They are produced by the interactions of the fundamental particles of the atom in addition to the interactions of the atoms themselves. Incidentally, it is the soft particles comprised of photons close to, and in, the visible range which permeate matter that make matter visible. If only the hard particles were present, solid matter would be invisible, although completely tangible.

The leading part of the gravity radiation front produces negative charge effects, while the trailing portion which has passed a given body must have substantially reduced negative charge effects. The spin of the particles in gravitational radiation have a gyroscopic effect which keeps the particles orientated in the same position and the particles have little tendency to scatter.



Photons of gravity-inducing radiations not appreciably scattered while passing through positive matter

Soft electrons comprised of gravity-inducing radiations

Photons of gravity-inducing radiations collide and attach themselves to the back portion of soft particles comprised of gravity-inducing radiations. This in part, tends to nullify the negative charge effect on the back side of the soft electrons.

The faster moving photons overtake the soft particles and tend to adhere to them in irregular patterns, creating a perforated and extremely rough surface on this part of the particle, not unlike that of a positive charge. This has a tendency to cancel out much of the negative field effects on this part of the particle. The bombardments accelerate the particles to such an extent that no more photons can adhere to them. Therefore, the rear part maintains a positive charge, or at least, a much reduced negative charge.

Another important factor which contributes to a reduced negative charge at the rear is that there is a far greater photon concentration in this region than exists at the frontal portion. This is a result of a backup of photons caused by the lower velocity of the soft particles accompanying the radiation. This photon concentration tends to divert the normal ether bombardments (which produce the electrostatic effects) from their usual paths. Since gravity radiations produce forces, it follows that there are interactions which eventually disperse the radiation, accounting for the limited penetration of gravity radiation. Gravity is an electrostatic effect, not some space warp or mysterious force with imponderable properties. If matter is impregnated with sufficient quantities of negative charges, especially soft electrons, it will become weightless and even levitate.

Some individuals have the ability to do the reverse of levitation, possibly by expelling large quantities of negative charge from their bodies. A dwarf who had a normal body weight of 123 pounds demonstrated under strict anti-fraud conditions that he could increase his weight to 900 pounds.

The physics of levitation was demonstrated in part when missiles were found to have lost most of their original weight after travelling through the Van Allen Radiation Belt and returning to Earth. The weight loss continued for some time and containers in which pieces of the missile were placed, also lost weight. The radiation belt contains high concentrations of negative charges of all kinds, from hard electrons to the very soft ones. The missile became impregnated with negative charges as it passed through this region, absorbing an abnormal quantity. The more penetrating softer particles opened the door for the harder particles to enter. The loss of weight of the container would have been caused by the missile gradually losing some of it's excess negative charges and those charges being absorbed into the container.

12. Faster than light travel is possible because the accelerating gravity beam travels with the mass being accelerated. At ultra-high velocities, or where most of the electrostatic potential of matter has been transformed, cohesive forces will tend to break down and the material will cease to be a cohesive solid. However, spaceships can travel many times the speed of light provided that the ship and the occupants are impregnated with the right combination of negative charges which would prevent any extensive transformation of the electrostatic mass into magnetic energy. At ultra-high velocities, the closing forces on the rear of the craft no longer compensate for the forward resistance, so it requires a steady application of accelerating forces to maintain velocities many times that of light.

The evidence concerning spaceship propulsion demonstrates that the famous Einsteinian equation $E = mC^2$ falls far short of representing the energy potential of matter. From the kinetic energy equation $E = 0.5mV^2$ it follows that a body travelling at only 1.5 times the speed of light (which isn't even a cruising speed for most spaceships) has a kinetic energy which exceeds the value of Einstein's celebrated equation. At this velocity, only a miniscule part of the energy potential of the mass has been released. The meaninglessness of the famous equation is also evident, because inertia is dependent only on net charge and not necessarily on mass or quantity of material.

13. Another item which demonstrates the validity of the information presented here is the fact that determinations of the gravity "constant" "g" is always significantly higher when measured in mines. This is to be expected as the soft electron concentration is much higher below the surface than it is above the surface. Another fact which disturbs physicists (and consequently given little publicity) is that objects at the bottom of mine shafts weight significantly less than they should according to Newton's concept of gravity.

Another enigma which is damaging to the academic viewpoint is that experiments indicate that gravity doesn't impart the same acceleration to all substances. To try to deal with this fact, they have to introduce a mysterious fifth force which is supposed to be an extremely feeble repulsive force with a limited range. It is supposed to be more prevalent in some substances than in others. The concepts already explained here show that this is to be expected. Different atoms and molecules have different total positive charge effects in proportion to the number of fundamental particles from which they are made. Consequently, they will not be given the same gravitational acceleration even if the mass is identical.

14. We now come to questions such as; How does the law of redistribution of energy work? What are the real principles behind colour perception? Why is the velocity of light independent of it's frequency? Why is this velocity equal to the ratio between an electromagnetic and electrostatic unit of charge? The answers to

these questions have never been given before.

When materials are subjected to high temperatures, great fluctuations occur in the velocity of electrons in their orbits. This in turn, creates interactions and interference effects between electrons moving in opposite directions inside atoms and between electrons in adjacent atoms. These interactions generate changes in the electrostatic field effects of the electrons, which will cause regular and distinct changes in their so-called orbits. This is because the charge on the electron varies with it's velocity. Abrupt changes in the velocity of a particle disrupt the ethers which results in the formation of photons.

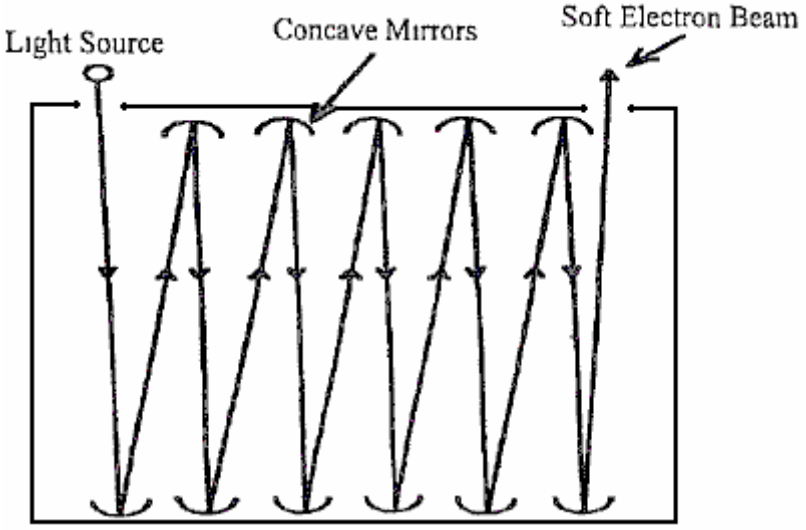
The larger the atoms, the more complex the interactions, and consequently, the more intricate the spectral pattern. The photons comprising the electrons, determine the range of the ethers which will be disrupted. These ethers are intimately connected with the electrostatic field intensity of the electrons. Thus it follows from this new definition of Planck's constant, just why this definition must be taken into consideration in all of the calculations concerning the frequency of light produced in interactions.

The electrostatic field effects around an electron depend on the range and level of the ethers deflecting off the sides of the particles. This range is not necessarily the same as the range of ethers affected by sudden changes in the velocity of the electrons, but there is a direct correlation between the two. Planck's constant has a role throughout the procedure as all ether particles have the same kinetic energy.

The law of redistribution of energy states that when light interacts with matter, new light with a lower average frequency than the original is produced. One of the most simple demonstrations of this is shining a blue or violet light through a large number of filters of any type. The emerging light is always red. All of the colours which we normally see are combinations of different frequencies of photons. A wide frequency range of soft particles and photons will tend to be grouped together. This means that bands of light in the blue, indigo and violet ranges will contain other colours down to the reds, although the reds will make up only a very small percentage. The human eye sees only the dominant colour, and prisms cannot separate them. The colour experiments of Edwin Land proved this to be the case.

In the May 1959 issue of *Scientific American*, Land showed that two black and white transparencies produced from black and white film, were exposed to filtered light from two different (reasonably spaced apart) parts of the spectrum, the resulting superimposed images were in full colour. This shows that it requires subtle combinations of frequencies for the eye to perceive colour if the light isn't in a narrow frequency band. Otherwise, the eye will see things in various shades of black and white, which contain all the colours but in the wrong combinations to be seen as colours. This is what occurs for people who are subject to 'colour blindness'.

Under certain conditions, light reflected from a mirror can have greater intensity than the incoming light. The light has to be of high intensity. When the particles of the incoming light collide with the mirror, those that are reflected are instantly brought to a halt. This produces a tremendous disturbance of the ethers which results in the creation of new photons which are added to the photons which are reflected. In addition, many of the photons reflected and created, combine to form soft electrons, and so the reflected light has a higher percentage of soft electrons than the incoming light beam.



It follows that repeated reflections of a light source such as the Sun, would result in a highly lethal laser-like

beam. This has been demonstrated on numerous occasions. Perhaps the most notable demonstration occurred near White Sands, New Mexico in the early 1950s. It was witnessed by an acquaintance of a personal friend of Mr Cater's, and it was something which he wasn't supposed to have seen. About thirty-five four foot diameter mirrors were arranged so that the resulting beam was directed at the rock wall of a mountain. It immediately created a neat hole through 200 feet of solid rock.

An associate of Mr Cater's found that by putting a strong negative charge on the mirrors that their reflective power is considerably increased. He charged a series of metallic mirrors to 20,000 volts and found that after 10 repeated reflections from the Sun, the resulting beam was very lethal. This shows that it is the negative charges deposited on a mirror surface which enables it to reflect most of the light that falls on it. Incoming light immediately deposits negative charges on the surface and those charges repel the rest of the light. The more intense the incoming light, the higher the concentration of negative charges placed on the surface. This accounts for the fact that highly lethal beams reflecting from the surface do not destroy the mirror. The mirrors must be metallic and preferably concave. Glass mirrors do not work as much of the incoming light is lost before it reaches the reflecting surface and much of the shock effect of the light reflection is lost due to the glass slowing down the incoming beam. The incoming light must strike the mirror in as nearly a perpendicular direction as is possible. If soft electrons associated with colours known to be highly beneficial could be concentrated using this method to produce rapid healing

15. The question arises: Why is the velocity of light independent of it's frequency? This is implicit in Maxwell's equations but it still isn't explained. When the ethers are disturbed to produce a photon, a fraction of them are compressed and a great number are forced close enough together to adhere to one another. The higher the ethers affected, the more rapidly and suddenly this displacement has to occur in order for a photon to be produced, otherwise, the ether particles will escape this compression since they are very active. This momentary compression quickly returns to normal, rather like a compressed spring being released. This rebound hurls the aggregate photon forward at the speed of light. The distance of this rebound is equal to the so-called wavelength, or distance over which the photon is accelerated to the speed of light.

This is exactly what happens when lower ethers are disturbed to form lower frequency photons, except that the rebound takes place over a greater distance with a lower average acceleration of the photon. Since the warped pattern is identical in both cases, both photons reach the same velocity, which is independent of the actual wavelength produced. As both photons receive the same thrust, it can be seen that lower frequency photons must have a greater mass, that is, the frequency of light is inversely proportional to the mass of the protons which form that light.

The behaviour of electrons and protons in a particle accelerator shows that at the speed of light, all of their electrostatic potential has been transformed into magnetic energy. This shows that the velocity of light relative to it's source is the ratio of it's Electromagnetic Unit of charge ("EMU") and it's Electrostatic Unit of charge ("ESU"). The ratio EMU / ESU is equal to the speed of light "C". Calculating from these details, shows that the total pressure exerted on a single electron by the surrounding ethers is 14.4 dynes which represents a pressure beyond normal comprehension when the minute size of an electron is considered.

16. We now need to consider the role of soft electrons in promoting chemical changes and maintaining life. It has been repeatedly confirmed that magnetic fields have considerable healing properties and will stimulate plant growth. What has not been realised is that it is not the magnetic fields themselves which are responsible for this effect, but it is the soft electrons which they capture and concentrate. One pole of a magnet has beneficial effects for certain ailments, while the opposite pole is not as effective.

One of the most significant properties of soft electrons is their ability to promote chemical changes. A change in a molecule is not likely to take place without lessening the chemical bond or attraction among it's constituent atoms. Soft particles interpenetrating the molecule will bring about this condition by carrying harder electrons in with them, which in turn weakens this bonding by offsetting the positive charge effects of the nucleus. Soft particles tend to camouflage a variety of harder particles. This is a vitally important property because in this manner, other atoms which are going to take part in the chemical change, also have their zonal effects temporarily altered so that they can come into more intimate contact during the reaction. The soft particles tend to act as catalysts for the reacting particles and the soft particles tend to get disintegrated in the process, releasing additional energy which expedites the reaction and allows the normally powerful electrostatic field effects within the atom to return to their original state. The release of the hard electrons contained within the soft particles which disintegrate is the source of much of the heat produced during chemical reactions.

17. The properties of water: water is a universal catalyst because of it's unique ability to collect and concentrate an enormous quantity of soft electrons of all kinds. This is the reason why water has the highest

specific heat of any know substance. The large amount of energy contained in water in the form of soft particles, has been demonstrated by experimenters on many occasions. For example, a number of reports show internal combustion engines running with water as the fuel. No reasonable explanation for this has been given as it appears to be contrary to all the rules of chemistry. However, the disintegration of the more unstable soft particles contained in the water when subjected to compression and ignition inside the engine, accounts for this seeming impossibility.

Water is a unique substance being comprised of two of the most chemically active elements, both of which are gaseous elements. The fact that three oxygen atoms can combine to form ozone, indicates that the oxygen atom is extremely magnetic, indicating that a higher percentage of it's orbital electrons are moving in approximately the same plane. This leaves fewer orbital electrons tending to offset the positive charge of the nucleus and other portions of the atom. Consequently, two side of the oxygen atom possess an inordinately strong overall positive charge. When hydrogen atoms combine with an oxygen atom, the electro9ns on the side of the hydrogen atoms adjacent to the oxygen atom are brushed aside. This is on the segment of the oxygen atom where most of the electrons of the oxygen atom are orbiting. The normal flow of electrons around the proton of the hydrogen atom is diverted to become a flow which encircles the oxygen atom and the outer periphery of the hydrogen atoms. This results in a powerful magnetic and electrostatic bond between the hydrogen atoms and the oxygen atom. The electron flow around the hydrogen atoms is extremely rapid, resulting in a very high overall positive charge on the hydrogen atoms. As there is a very strong mutual repulsion between the hydrogen atoms, they will line up on opposite side of the oxygen atom, giving water the structure H-O-H. This molecule has strong and extensive positive zones, so the attraction zone is a considerable distance from the molecules. This is why the specific gravity of water is low, despite the strong positive charge of the molecules.

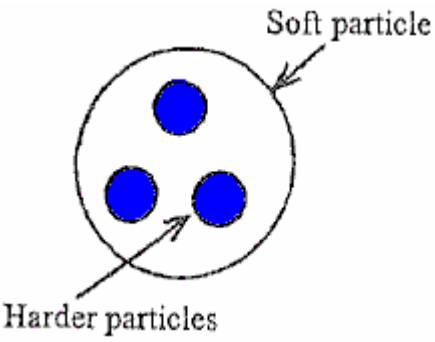
The great affinity of water for soft electrons is now apparent. The large positive zones between molecules are havens for soft electrons, drawn there by the attenuated, but significant, attraction of the hard electrons captured by the soft electrons. Although soft electrons are large compared to hard electrons, they are still very small compared to an atom. Therefore, the spaces between water molecules can harbour large quantities of soft electrons, without them being bound to the water molecules.

Perhaps the most baffling feature of water is that it expands when it freezes. The high concentration of soft electrons greatly weakens the forces of attraction and repulsion between the molecules. As a result, the average kinetic energy of the molecules at the freezing point are still sufficiently large to allow the molecules to move in and out of the zones of attraction and repulsion, without being confined in the attraction zone. The cooling must continue until the soft electron concentration reaches the stage where the attractive forces become strong enough to confine the molecules to the attractive zone. When this occurs, the water becomes a solid. Since the attractive zone is an inordinate distance from the molecules, the average distance between molecules becomes greater than it was when the water was in a liquid state. At the freezing point, the molecular activity is low enough to permit soft electrons to enter or leave the substance without disintegrating. In order for the water to be transformed from a solid back into a liquid, the same quantity of soft electrons must be injected into it as were removed when it changed from a liquid to a solid.

The melting and freezing temperatures of water vary considerably due to the differing amounts of soft electrons contained in it. Another unusual feature is that in cold weather, hot water pipes have a greater tendency to freeze than cold water pipes do. This is because the heating of the water drove off many of the soft electrons normally contained in the water and due to the low temperature of the surroundings, these soft electrons were not replaced, and as a consequence, freezing to become a solid happens more easily.

One tends to think of colloids as ultra-small particles of solid matter. However, molecules of water can adhere to each other to form aggregates of water molecules, which are effectively, colloids as well. Colloids have strong electrical properties as indicated by the fact that they are not affected by gravity. The field zones around any such colloidal group will be much stronger than that around a single water molecule. Water with a high percentage of such colloidal groups can capture a very large number of soft electrons which are beneficial to health. Abnormal conditions in certain places can favour the formation of water colloids and that can account for the healing properties of water found in some places, such as Lourdes in France.

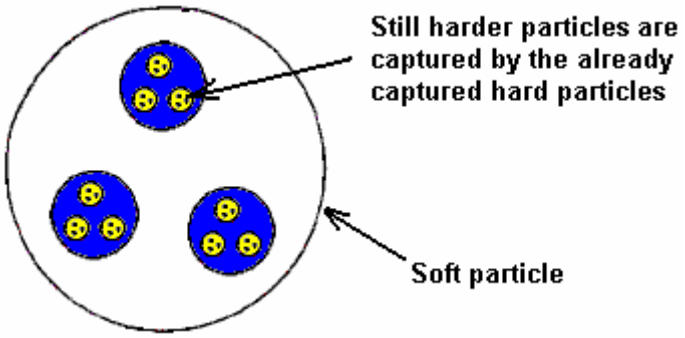
18. Hard particles can be captured by softer particles and this is deeply involved in a wide range of phenomena, from the transmission of heat and electricity, to the formation of clouds.



Two particles differing greatly in the frequency range of photons of which they are comprised, have little tendency to repel or attract each other. Therefore, they tend to interpenetrate.

Ether particles have zones of attraction and repulsion. Since photons are composed of ether particles, they will in turn, possess zones of attraction and repulsion. In the case of ether particles, these zones will be correspondingly smaller in proportion to the diameter of photons. When protons combine to form electrons or protons, the same zones are present between these particles. However, the zones of attraction are minute when compared to the diameter of the electron or proton, and like particles, seldom if ever get close enough together at sufficiently low velocities for the attractive forces to become effective.

The situation is entirely different when two similar particles composed of photons but with widely differing frequencies, approach each other. Electrostatic attraction or repulsion is considerably lessened because each is associated with ethers which differ considerably from each other. When they are ion direct contact with each other, electrostatic repulsion tends to vanish, since there can be little or no bombardments on the sides facing each other. Since each particle associated with ethers is somewhat different, they will tend to interpenetrate. This means that they will be completely within the ether attraction zones of one another. As a result, the harder particle is captured by the softer one. In a similar manner, the captured harder particles will, in turn, capture still harder particles and this process continues until electrons normally associated with electricity are confined. This combination of particles tends to nullify the electrostatic forces which are normally produced by the confined particles, camouflaging the captured harder particles so that their presence is not readily apparent.



The ether particles normally bombarding the hard electrons and protons which produce electrostatic field effects, tend to be diverted from their normal paths by the presence of softer particles or media between the repelling like charges and/or the attracting unlike charges. These interpenetrating softer particles produce an ultra-high concentration of ether particles around the hard particles. The motion of these ether particles is greatly restricted. This offers a barrier to the higher ether particles which normally bombard the hard particles. This has a tendency to slow them down, and any which do collide with the hard particles, do so with considerably less impact than normal, therefore they tend to become electrically neutral and their motion slows to nearly a halt.

Soft particles permeate matter as well as the spaces between matter, yet they do not to any great extent, neutralise the electrostatic field effects of the fundamental particles, because they are more concentrated and their rapid motion tends to prevent capture. However, additional concentrations of soft particles of the right kind, injected into matter, can render the elementary particles within the atom, electrically neutral and the matter becomes what is known as “dematerialised”. This conglomeration of soft and hard particles renders the soft particles electrically neutral.

It should be noted that only hard particles or the fundamental particles of the atom, are hollow. All other particles, including photons, do not have this tendency because of the nature of their formation. If the softer particles were hollow, they would be unable to capture harder particles. Hard particles entering a hollow, soft particle, would maintain their charges and force a mutual repulsion. Therefore, they would escape

immediately. Photons, if hollow, would tend to be less stable, and the probability of forming other particles would be lessened.

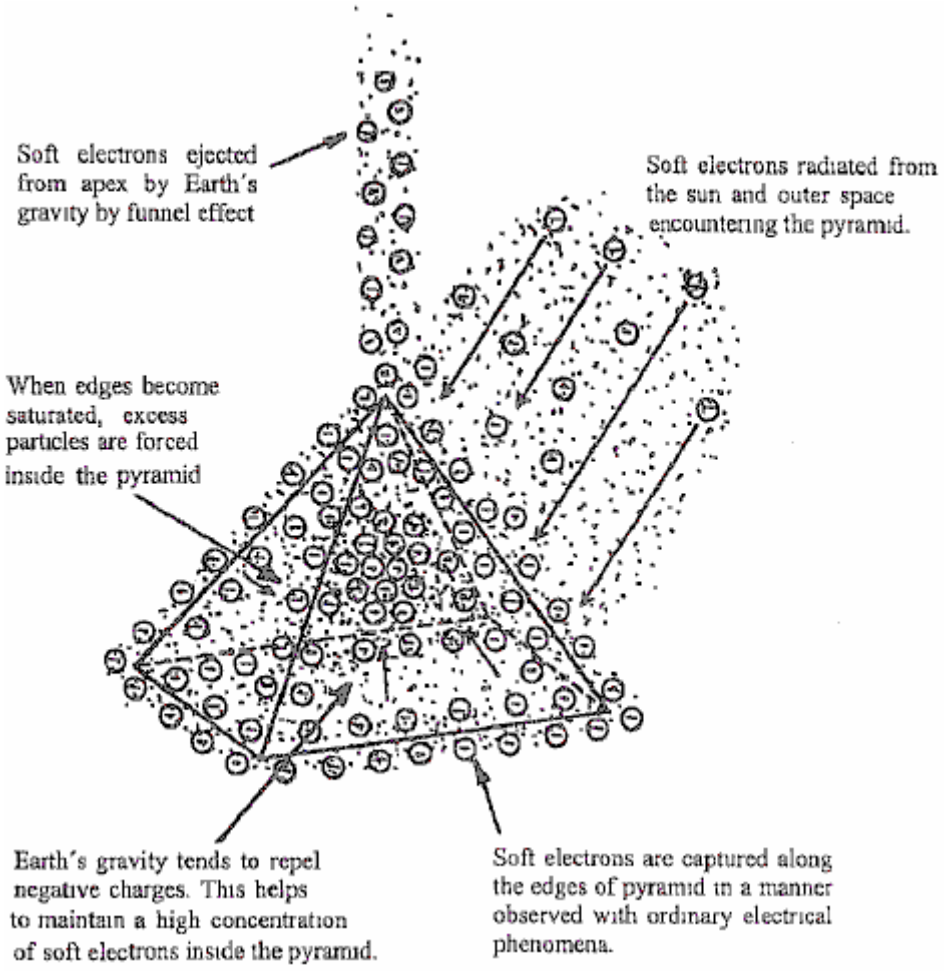
When a soft particle disintegrates, a chain reaction occurs. The disintegration releases the confined, harder particles. The energy released during the disintegration is generally sufficient to disintegrate the weaker hard particles which it originally captured. This, in turn, results in the disintegration of still harder particles, until the very hard and stable electrons of electricity are released. Highly interesting experiments performed in Poland by two scientists; Howsky and Groot, demonstrated the ability of soft electrons to house and camouflage harder electrons, and to release them under certain conditions. These experiments were also a great confirmation of other principles already mentioned here, especially those involved with levitation.

A small quartz crystal was attached to an oscillator which generated radio frequencies of several kilowatts. This caused the crystal to lose it's transparency and increase it's volume 800%. The crystal then levitated and carried the oscillator, as well as a 55 pound weight, to a height of two metres above the floor. An account of this was given in an issue of *Science and Invention* magazine and it included a photograph of the levitation.

19. The energies concentrated inside a pyramid have been shown to be extremely beneficial to humans. Soft particle bombardments from outer space and especially from the Sun, concentrate inside the pyramid. Some, passing through the surface of the pyramid are slowed down to such an extent that the Earth's gravitational field, repelling the negative charges, tends to keep them inside until collisions with other particles drives them out.

Most of the particles collected by the pyramid, concentrate along the edges as would be expected, since electricity on any charged body tends to do much the same thing, with concentrations at points and along edges. In fact, pyramid frames have been found to be nearly as effective as the closed pyramid, if, and only if, there is a continuity in the framework and no breaks in any of the joining parts.

The soft electrons collected on a pyramid frame or closed pyramid, soon reach saturation point and continued bombardment causes the excess to drop down inside the pyramid. This, coupled with the gravity-repelling forces, causes a high concentration inside the pyramid. The proportions of the pyramid are apparently a factor in it's performance. If the sides are too steep, many of the soft electrons will move along the edges into the ground outside instead of being forced inside the pyramid. If the sides are not steep enough, not many particles will be collected as they strike the material at nearly a right angle which causes only a small reduction in velocity. If they strike at a sharper angle, there is a greater tendency for them to be retained by the material.



If two sides of the base are aligned with magnetic North, it is allegedly more effective. Pyramids can be rendered more potent by lining the interiors of a non-metallic enclosed pyramid, with metal foil such as aluminium or copper. The foil allows a greater quantity of soft electrons to accumulate around the non-metallic outer portion because the soft particles do not pass through the metallic substance as easily, causing a back-up of soft particles. During the process, the foil absorbs large quantities of soft particles before many of them can enter the pyramid. pyramids also radiate soft electrons upwards from the peak.

Many of the soft particles which are stopped briefly on the outside of the pyramid, are repelled upwards by the Earth's gravitational field, and as well, by soft electrons attached to the pyramid. This produces a funnelling effect which ejects soft electrons from the apex of the pyramid. The Earth's gravity accelerates soft particles at a far greater rate than it does ordinary matter as soft particles are associated with ethers which are much closer to those of the gravity-inducing particles than is the case for ordinary matter. After the pyramid becomes saturated, a greater quantity of soft particles than ever, will concentrate inside. The foil will continue to radiate a high concentration of soft particles during the night when the number of particles bombarding the pyramid is considerably reduced.

It is found that pyramids work better during the summer than at any other time of the year. They are also more effective in the lower latitudes because most of the energy concentrated by the pyramid comes from the Sun. There are conflicting opinions as to the effectiveness of pyramids because of this as there is little understanding of the principles involved. For example, those who experiment with pyramids in Canada may claim that they don't work while those in Southern California will contradict them. A pyramid does not increase the flow of soft particles through the area covered by the pyramid as the same concentration flows outside the area. What a pyramid does, is impede the general flow of soft particles and produce a back-up of particles inside and below the pyramid, and consequently, a higher concentration of soft electrons in these regions. The material used in a pyramid is of great importance. This was demonstrated when a wealthy man in the Midwest built a pyramid-shaped house five stories high, which was then covered with gold-plated iron. The phenomena produced were completely unprecedented. For example, ground water was forced to the surface and flooded the first floor. This was because the soft particle concentration inside and below the pyramid was so great that ground water was impregnated with such an abnormal concentration of negative

charges that it was repelled upwards by the Earth's gravity.

Gold atoms have extremely high positive electrostatic field effects, more so than any other atom. This is why gold is the most malleable of all substances. This means that soft electrons will have a greater affinity for gold than for any other metal. As a result, abnormally high concentrations of soft electrons will concentrate around gold. This effect is greatly enhanced when gold is in contact with iron. These dissimilar metals produce an EMF which is turn, causes a flow of electricity or eddy currents resulting in the iron being magnetised. The magnetic field produced, captures additional soft electrons. A higher concentration of soft electrons is created by this combination then could be produced by a similar thickness of gold foil alone. It follows that by far the most effective material that could be used for pyramids is gold-plated sheet iron (galvanised iron should not be used).

With everything else being the same, the greater the size of a pyramid, the better the performance. The reason for this is that the thicker the layer of concentrated soft electrons through which the incoming soft particles must pass, the more they are slowed down when passing. This results in a greater back-up of soft electrons and an increase in the concentration inside the pyramid. Another reason is that a large pyramid has a greater ratio of volume to surface area. Soft electrons are continuously leaking away from the surface of the pyramid, the larger the pyramid, the lower the percentage of soft electrons which is lost. Consequently, very small pyramids are ineffective.

20. Viktor Schauberger of Austria was puzzled by the fact that large mountain trout could remain motionless for as long as they liked in the fastest flowing water in streams. When disturbed, they escape upstream with fantastic speed. He also noticed that water gets charged up through swirling vortex action as it flows around obstructions. As the water is highly agitated, it gives up large quantities of hard and soft electrons to the fish, causing the entire outer surface of the fish to get a high negative charge. This charge repels the outer electrons of the water molecules, totally eliminating drag and as a result, the water exerts almost zero force on the fish. This effect is even more enhanced as the fish moves upstream, much more so than if the fish went downstream. The negative charge also helps the fish jump as the Earth's gravity boosts it upwards.

21. Brown's gas, produced by one form of the electrolysis of water has properties which seem bewildering to most scientists. Using it, allows steel to be welded to a clay brick and the flame is not harmful to human flesh. The flame temperature depends entirely on what it is applied to. It can also reduce nuclear radiation by 96%. The properties of Brown's gas confirm the information above. Water has a very large capacity to store soft electrons in addition to those already present in the structure of water. Brown did not separate water into hydrogen and oxygen. Instead, he added additional soft electrons to the water molecules. These additional charges greatly weakened the cohesive forces between the molecules, converting the water to an unstable gas. All of the properties of Brown's Gas follow naturally from this. Under welding conditions, the vast concentrations of soft electrons supply the release of sufficient quantities of hard electrons to produce the needed heat. In addition, the soft electron concentrations enable iron atoms to partially interpenetrate brick molecules to produce a bond between brick and iron. Also, with the ultra-high concentration of soft electrons, the gas can readily neutralise the positive charges of nuclear radiation.

22. We need to examine the source of the Sun's radiant energy. One thing that all suns seem to have in common is their great size. The astrophysicists speak of white dwarf suns of planetary size or less. It is clear that any claims made by astronomers or astrophysicists concerning celestial determinations, have about the same degree of merit as the other scientific claims which have already been mentioned. There is nothing to justify the existence of a white dwarf. For one thing, due to it's allegedly small size and limited gravitational influence, it could only hold very small bodies of asteroid size in orbit around it and those would have to be only a short distance away from it. According to the fallacious theories of orthodox science, a white dwarf consists of atoms with practically all of their electrons stripped away, giving it enormous gravity. It will be shown that astrophysicists have no way of accurately determining the distance or the size of any celestial body.

The larger the body, the greater it's mass or volume in proportion to it's surface area. This means that as the size increases, it is less probable that the energies produced by the normal activity of the atoms in the body's interior will escape from the surface without a resulting increase of temperature at the surface. The energy radiated from the surface will be in the form of photons and other particles of all types. Below a critical size, the surface area is sufficient to allow all of the radiant energy created in it's interior, to escape without an increase in temperature. In fact, such a body will lose heat unless it receives sufficient energy from it's surroundings.

As a body increases in size, it's surface area becomes increasingly inadequate to allow the radiated energy in it's interior to escape without a build up of heat at, and below, the surface. The surface will not radiate the

heat or energy outwards as quickly as it is created in the interior. The rate at which energy is radiated from a surface increases rapidly with a resulting increase in surface temperature. This varies as the fourth power of it's absolute temperature. For example, within a certain temperature range, if the temperature is doubled, the rate at which energy is radiated in the form of photons and soft particles, increases by a factor of 16.

The critical size of such a body will depend on it's composition. For example, if it contains a high concentration of mildly radioactive substances, this critical size will be less. If the body is hollow, then the dimensions would have to be greater. The red giants, if they are even close to the dimensions claimed, would have to be hollow and have relatively thin shells, otherwise, they would not be red as their surface temperatures would be astronomical.

The actual source of the energy which is finally radiated out into space is the soft particles and photons which are normally radiated by the atoms of the material inside a sun. This is due to the activities of the fundamental particles. Because of the great mass of a sun, an abnormal concentration of these soft particles is always present in the interior. This concentration is greatest near the surface. There is a steady increase in intensity, from the centre toward the outside. This results in a continuous disintegration of a high percentage of those particles near the surface, accompanied by a great increase in temperature, which in turn, results in a greater rate of disintegration, with the release of harder particles which produce the higher temperatures. At the same time, there is an increase in the rate at which the soft particles are created. The temperature will decrease steadily as the centre is approached and any sun will have a relatively cool interior.

The principle that size is the major factor in a celestial body's ability to radiate is confirmed by the behaviour of very large planets such as Jupiter and Saturn. An application of this principle indicates that bodies of such size should start radiating more energy than they receive from outside sources. Recent determinations indicate that Jupiter and Saturn do, in fact, radiate more energy than they seem to receive from the Sun. A probe showed a surprisingly higher temperature in Jupiter's upper atmosphere than was formerly believed to exist.

It now becomes apparent that the conventional theory which states that the radiant energy of the Sun is produced by thermonuclear reactions is complete nonsense. One thing to consider is that if this were the case, the Sun's radiation would be so lethal that no life could exist on any of the planets in the solar system.

Occasionally, throughout the universe, the gradual build up of heat in the interior of suns becomes very much greater, possibly due to the quantity of radioactive elements in the interior caused by transmutation. In such cases, relief valves in the form of sunspots, no longer take care of the excess energy increases and large portions blow apart, releasing astronomical quantities of radiation. After the explosion, the supernova becomes a burnt out body in comparison to it's former state. Considering the countless billions of stars within our field of vision, and since only a few supernovas have been observed down through history, it is logical to conclude that it is not the fate of the great majority of stars.

One of the phenomena concerning the Sun, which completely baffles all of the scientists, is that it seems to rotate faster at the equator than it does in the higher latitudes. Sunspots in the vicinity of the equator make a revolution about the Sun in less time than those in the higher latitudes. This is an annoying paradox which can't be pushed aside by these scientists as it is out there for all to observe.

The part of the Sun which we see is a highly fluidic blanket. The region around the Sun's equator could rotate faster if, and only if, a steady external pull is exerted on that region. Otherwise, internal friction would eventually produce a uniform motion. This means that bodies in orbit near the equator and close to the surface, are generating a high concentration of gravity-inducing radiations. It becomes evident that such bodies could not consist of normal matter and are probably composed of atoms and molecules made up of softer particles which are little affected by the Sun's radiation. Such bodies could generate a concentration of gravity radiations considerably out of proportion to their masses. Being constructed of this kind of material, they would be practically invisible.

23. Errors have been made in determining the size and distance of planetary bodies. Charles Fort cited many instances of fiascos which belied astronomers' claims of extreme accuracy in determining stellar and astronomical distances. His revelations did little to enhance their reputations as paragons of integrity.

The principles employed by astronomers in their measurements are essentially the same as those used by surveyors in measuring distances and elevations. However, some surveyors admit that they are unable to determine the height of mountains with any degree of precision and their measurements may be off by as much as 10%. Mr Cater has tested this using an altimeter which was set to zero at sea level and then driven

to the top of a mountain at 42^{O} 30' North latitude, which is supposed to have an elevation of 9,269 feet. The altimeter reading agreed closely with the established elevations of towns along the route ranging from 1,000 to over 4,000 feet. However, at the top of the mountain, the reading was only 8,800 feet. Mr Cater then reset the altimeter to the 9,269 feet attributed to the mountain and retraced his route. At every spot on the return trip, the altimeter consistently indicated elevations more than 400 feet higher than before. Even after several months, the altimeter reading was still more than 400 feet higher than it should be. A similar test was carried out on a mountain with a recorded elevation of 4,078 feet and at the top, the altimeter showed 3,750 feet although it agreed with other established elevations much lower down.

The fact that the altimeter was accurate at all places except the top of the mountain (whose official height was found by triangulation) shows that the methods employed by surveyors and astronomers are far from being accurate. The heights of mountains determined by triangulation will always be considerably more than the true value. There are two factors involved. First, the atmosphere becomes steadily denser as one descends from the top of the mountain. Second, the orgone concentration becomes greater closer to the ground. This means that light rays from a mountain top will be refracted and so appear to be originating from a point well above the top of the mountain. This was also confirmed by a barometric test at the top of Mount Everest which indicates that it is actually 27,500 feet in elevation and not the 29,000 feet previously supposed.

A friend of Mr Cater had his property surveyed to determine the acreage. Afterwards, he checked some of the distances determined by triangulation, using a tape measure, and found significant errors. Refraction of light is clearly throwing triangulation results off. The bulk of refraction effects are caused by orgone concentration. The measurement of mountain elevations taken at different times give different values and this is due to fluctuations in orgone concentrations, which are higher on hot sunny days than on cool cloudy days. Also, they are generally higher during summer months rather than at other times of the year.

The examples above show the unreliability of results obtained by triangulation. Astronomers are faced with additional factors when they try to apply triangulation, such as the Van Allen Radiation Belt, varying concentrations of orgone throughout space, etc. It is not realistic to assume that astronomers can determine planetary and astronomical distances with great precision.

There are several factors which astrophysicists and astronomers have not taken into consideration in their calculations. Perhaps the most important of these is the fact that all electromagnetic radiations including gravity in free space, suffer an attenuation effect which is well above that of the inverse square law. Everywhere in the universe is permeated with soft and hard particles of all kinds. These particles have been radiated by planetary systems for countless ages. This principle is demonstrated by fluctuations in the velocity of light and gravity attenuation.

There is a steady decline in the velocity of light as it travels through space. The reasons for this can be seen from the following considerations. Normal light, or light which has travelled a relatively short distance from it's source, immediately resumes it's original velocity after passing through a dense medium such as glass or water. As shown earlier, this is due to the close bunching of photons and soft electrons in any given ray. The concentrations of particles in a ray of light tends to decrease after travelling great distances. The father it travels, the more attenuated the ray becomes. This means that its ability to increase it's velocity after passing from a medium of a given density to one of a lesser density, will be reduced. This is, of course, due to the scattering and dissipation of particles within the ray as it encounters the conglomeration of particles moving in random directions throughout space.

Since conglomerations of soft particles permeate all known space, and the distribution is not uniform, it follows that light will experience refraction effects, even when passing through free space. Therefore, even under the best conditions, with observations being made beyond the atmosphere, astronomical observations cannot be made with any degree of accuracy. The difficulty is, of course, compounded when the observations are made inside the atmosphere. It is small wonder that Charles Fort found a wealth of evidence that completely debunked the astronomer's claims of great precision.

The fluctuation in soft particle distribution, along with the refraction effects of the atmosphere, rules out the possibility of averaging out errors by making many observations and applying the mathematical method of least squares. Conventional statistical theory obliterates actual small variations and distorts data by such averaging out processes. The gross errors which crop up despite these methods speak for themselves.

In order to measure the orbital distance of various planets, it was necessary to find the distance of the Earth from the Sun. Originally, this was allegedly found by measuring the angles that two widely separated observation points on the Earth made with the Sun. This is known as the parallax method. The distance to

the Sun was calculated from these angles and the distance between the observation points. The size of the Sun could then be determined, and knowing the orbital period of the Earth around the Sun, the Sun's mass and surface gravity were calculated by applying the false Newtonian concept of gravitation.

More recently, the distance to the Sun, known as the "astronomical unit" was supposedly determined to a high degree of "precision" by measuring the distance of the body Eros by the parallax method when it was closest to the Earth. Knowing the period of Eros' orbit, the distance to the Sun was calculated by the use of Kepler's law which states that "the square of the periods of any two planets are proportional to the cube of their mean distances from the Sun". Since the orbital periods of the planets are known to a reasonable degree of accuracy, most of the other unknowns within the solar system could be calculated by knowledge of the Sun's alleged mass and surface gravity. By now, it should be apparent that it would be a miracle, or at least, one of the strangest coincidences ever, if the actual distances coincided even approximately with the calculated values.

If the Newtonian concept were valid and the planets were held in orbit by only the effects of the Sun's surface gravity, then the orbital periods of the planets would be a reliable means of determining planetary distances. Since it has been proven that the concepts on which these calculations were made are false, it can be safely concluded that the size of the orbits is considerably different from what the astronomers claim. As a result of the dissipation effects of radiation, well beyond that which can be expected from the inverse square law, it follows that planetary distances are very much different from the accepted values.

This excessive attenuation of the gravity effects of the Sun is reflected in the alleged rapid increase of orbital distances of the outer planets. The supposed orbital distances are as follows:

Earth: 1.0 astronomical units.
Mars: 1.52 (difference 0.52)
The asteroids: 2.76 (difference 1.24)
Jupiter: 5.2 (difference 2.44)
Saturn: 9.58 (difference 4.38)
Uranus 19.16 (difference 9.68) and
Neptune 30.24 (difference 11.08)

It does not follow that the longer the orbital period, the greater the planetary distance. For example, within certain limits, the larger and more massive the planet is beyond a certain critical amount, the slower it must move in order to stay in a given orbit. This is because the total gravity effects of the Sun are unable to permeate and affect the entire mass to the extent that they would with a smaller planet. For example, a planet like Saturn could be placed in a stable orbit inside Earth's orbit. Yet it would have to move so slowly in it's orbit that it's orbital period would be much greater than that of Earth. This means that orbital periods are not a reliable gauge for relative orbital distances.

Although planetary and stellar distances are completely unknown as far as astronomers are concerned, and at this time there are no reliable means available of determining them, the diameters of some of the inner planets, including Jupiter and Saturn, can be calculated far more accurately than any of the other values in the realm of astronomy. The orbital distances of the planetary satellites in proportion to planetary diameters as well as their periods can be accurately determined. The determination of these constants is not affected to any significant degree by the dissipating factors of light already mentioned since a planet and it's satellites are about the same distance from the Earth. The main factor which makes it possible to approximate the diameter of any of these planets is the knowledge that they have practically the same surface gravity as Earth does.

If a satellite is very small as is the case with the satellites of Mars, a planetary diameter can be calculated with a high degree of accuracy. In fact, Mars is the only planet in the solar system whose diameter can be reliably determined. Astonishingly, Mars turns out to have a diameter of about 11,300 miles. Using unusual methods, Mr Cater has estimated the diameter of the Sun as over 2,500,000 miles and at a distance of about 277,000,000 miles from Earth. The Moon diameter at 5,200 miles at an average distance of 578,000 miles, shell thickness 115 miles and surface gravity 98% that of Earth. With a lesser degree of accuracy, the diameter of Venus is assessed at 23,000 miles and Mercury at over 8,000 miles. Jupiter diameter about 230,000 miles and Saturn about 200,000 miles. It is most unlikely that the accepted distances to the stars are even approximately correct.

24. Hard electrons travel through metals more readily than through non-metals. This indicates that they encounter more extensive positive electrostatic fields between atoms and molecules than in non-metals. At the same time, the atoms in metals are usually more mobile or free to move around than is the case with

solid non-metals. This is why the best conductors of electricity are also the best conductors of heat. It is significant that all of the heavier atoms are metals, with the exception of radon which is a gas. This means that such atoms have a higher net positive charge, which causes a stronger mutual repulsion for greater distances on atoms which are not directly connected to each other. This greater extension of the positive zone around such atoms gives them more freedom without breaking the bond which holds them together. The repulsive forces of nearby atoms, increases the mobility of any given atom.

The heavier atoms contain more protons and neutrons bunched together. The outside pressure needed to hold a group of mutually repulsive particles together is independent of the number of particles present.

One might conclude that the heaviest atoms make the best conductors, but this is not the case. Silver, copper and aluminium are the best conductors although their positive field zones are not as extensive, they have less inertia and so are more easily pushed out of the path of a flow of hard electrons. Electrons which flow along conductors are continually colliding with atoms in motion. Therefore, it require a steady application of electromotive force at the ends of the conductor in order to keep them flowing. The atoms of non-metals are more firmly locked into position and therefore do not have that much of a tendency to move out of the way and this is why they make good insulators. Electrons follow the lines of least resistance and so they tend to move on the surface of the conductor where there is less tendency to collide with atoms.

The rules governing the conductivity of soft electrons are somewhat different from those of hard electrons. Soft electrons are enormous when compared to hard electrons. This can be seen when considering that the average diameter of a particle is directly proportional to it's so-called wavelength of the light comprising it (or inversely proportional to the frequency). The ethers associated with atoms and their fundamental particles are much higher in frequency than those associated with soft particles. This means that atoms will offer little resistance to the passage of soft electrons. However, the magnetic fields resulting from thermal agitation of certain atoms and molecules are involved with ethers which are closer in frequency to the ethers directly associated with soft electrons. Consequently, soft electrons will interact with these fields. This explains why metals in general offer greater resistance to the passage of soft electrons than do non-metals.

The ordinary electrical transformer presents an enigma. The secondary of the transformer continues to pour out or eject electrons from a seemingly unlimited source. There is a limited quantity of free electrons in conductors which should be exhausted quite quickly. The standard argument used to account for the source of current is that free electrons in the circuit supply the electrons and are used over and over again. A simple calculation demonstrates that free electrons in conductors are not the source of electricity.

Consider a wire two millimetres in diameter which carries about 10 amps of current. The electron flow is concentrated near the surface of the wire. Since the electricity in a conductor travels at about the speed of light, such a wire 186,000 miles long would have 10 coulombs of electricity distributed over it's surface at any instant. The surface area of this wire is 1,840,000 square metres. A parallel plate capacitor having this plate area and a separation of one millimetre, would have a capacity of 0.016 farads. Even with a potential across it's plates of 100 volts, it would still only be able to concentrate an equivalent of 1.6 coulombs, and a good part of this electrostatic charge would be due to the displacement of the electrons and protons of the atoms. This voltage is more than enough to concentrate all of the free electrons on the surface of the plates. Similarly, all of the free electrons in the wire example would be involved if the current were maintained with 100 volts. Of course, a wire this long would have too much resistance to carry any appreciable current with 100 volts, but this has nothing to do with the argument just given. As a matter of fact, even 6 volts is far more than enough to produce a current of 10 amps in a wire of 2 mm diameter. Therefore, there aren't enough free electrons in any conductor to supply any appreciable current. This means that the source of electrons in current flow is not coming from free electrons in the conductor. The conclusion is therefore that the hard electrons somehow manage to get through the insulation of the conductor and flow into the wire from outside.

By the law of action and reaction, since a current has inertia, any change in the primary current of a transformer produces a force in the opposite direction in the secondary. This reactive force produces a disturbance of the ethers which produce the voltage or Electromotive Force as a result of increased ether bombardment. The EMF induced in the secondary winding of the transformer, creates a temporary electric void in the wire which draws all kinds of negative charges to the wire. The softer electrons quickly penetrate the insulation and stop at the surface of the wire as they do not travel as readily through a hard electron conductor. These softer electrons absorb most of the electrostatic forces in the insulation which impede the flow of hard electrons, allowing the hard electrons to pass through the insulation and enter the wire.

Electrical charges, composed of photons in nearly all the frequency ranges, permeate all space, since they are continually radiated by stars throughout the universe. They are no easily detected as they are in the

form of conglomerates with the harder particles residing inside the softer ones. The resulting combinations are highly penetrating and it takes something like a voltage induced in a conductor to separate the harder particles from the softer ones. The performance of a transformer can be greatly impaired by completely shielding the secondary winding with a good conductor of electricity such as copper or pure aluminium. This is because the shield tends to impede the flow of soft particles to the secondary. This effect has been verified by experiment.

The terms “EMF” and “voltage” need clarification. The true nature of the phenomena associated with these terms has never been fully understood. All that has been known is that if a conductor is exposed to an EMF, a flow of electricity is produced. Also, voltage is associated with the amount of energy or work which a current is capable of producing. An EMF of a given value can induce a current with a definite voltage. The voltage produced is directly proportional to the EMF impressed on the conductor. Also, the energy of the current is directly proportional to the voltage. The amperage of a current is a measure of the number of electrons passing through each segment of a conductor per second. Since wattage, or the total kinetic energy of this current flow is equal to the amperage multiplied by the voltage, it follows that the amperage is also directly proportional to the energy of the current flow. Therefore, Voltage is a measure of the average kinetic energy of the electrons flowing along the conductor. This in turn, is directly proportional to the square of the average velocity of the electrons. This simple definition of voltage is sadly lacking in all standard textbooks.

An EMF induces an accelerating force on an electron. What is the nature of this force? Basically, there are two methods of producing an EMF. One is by subjecting the conductor to a fluctuating magnetic field, and the other is by exposing the conductor to a difference of potential, such as connecting it between the opposite poles of a battery. In that instance, one battery pole has a negative charge while the opposite pole is positive. The flow of electrons is the result of an electron concentration at one point tending to flow to an area where there is a shortage.

The EMF is produced by direct electrostatic force, which in turn, has a dual nature. There is the tendency for negative charges to be attracted to positive charges, and then there is also the mutual repulsion between negative charges. The voltage attained is directly proportional to the difference of potential existing between the poles of that battery. The difference of potential is equal to the kinetic energy gained by the electrons in moving from one potential to the other.

The EMF produced by a fluctuating magnetic field gives the same results but the process is different. When a conductor is subjected to a fluctuating magnetic field, as with the secondary winding of a transformer, the “free” electrons of the conductor and the outer electrons of the atoms which are not as intimately associated with the atoms, are exposed to differential ether bombardments. It is equivalent to an electrostatic force. When a magnetic field changes, the change does not take place simultaneously throughout that volume of space occupied by the field but it progresses from one portion to another. This creates differential electrostatic ether particle bombardments on electrons within the field. When a conductor cuts magnetic lines as with an AC generator, the electrons are subjected to the same conditions experienced by electrons moving between the poles of a magnet. The accelerating force will be in a direction perpendicular to the direction in which the electrons in the conductor are found to move.

If there were even a small fraction of the free electrons existing in the matter as is believed by our physicists, the negative charge effects of matter would be so great that these bodies would be unable to get close to each other. Much of the charge on capacitors comes from outside the capacitor, as is the case with the flow of electricity in conductors. Actually, free electrons in a conductor are practically non-existent. Hard electrons which are not a part of the atoms are captured by soft particles which permeate matter. The soft particles release hard electrons when subjected to the EMF in a current, or the voltage across the plates of a capacitor.

The current in a straight wire is evenly distributed along the surface where the electron flow encounters the least resistance. The released hard electrons which are directly affected by the EMF, tend to move as a unit partially held together by mutual magnetic attraction. This unit leaves a temporary void behind it which is quickly filled by surrounding hard electrons. Many such groups are started almost simultaneously in a conductor at about the speed of light, although the electrons themselves travel at a much lower velocity. When an EMF is applied to a conductor, something akin to the domino effect is set up in the ethers. This effect travels at the speed of light since it is produced in a similar manner.

That the source of electricity flowing in power lines as well as that produced by generators, comes from soft particles which permeate and surround the area, has been proven during auroral displays. When aurora activity is unusually high, transformers in Canada have been known to burn out and even explode. At the

same time, increase of current flow in power lines has been great enough to trip circuit breakers as far south as Texas. As explained earlier, the concentration of soft electrons in the atmosphere is greatly increased during auroral phenomena. Some areas receive higher concentrations than others at the same latitude.

A loop of wire or a coil offers impedance to alternating current. This property is known as “inductance”. Since a single loop of wire has inductance, it follows that the effect can be explained in terms of one loop. Electrons tend to travel along the surface of a conductor as that is the path of least resistance. The major source of this electricity is the high concentration of soft electrons which gather around a conductor and permeate the material. This is due to the relatively high positive charge of the conductor. The greatest concentration is found at the surface and a short distance below the surface. When an EMF is applied to the conductor, free electrons are set in motion. During this process, soft electrons concentrated at and just below the surface tend to disintegrate and release more hard electrons. This is enhanced by the concentration of the soft electrons, which is turn causes an agitation of the soft particles, causing them to become highly unstable.

In a straight wire, most of this disintegration and nearly all of the electron flow takes place below the surface. This condition greatly shortens the mean free path of the electrons and the flow stops immediately after the applied EMF is shut off. Consequently, an alternating current will encounter the same ohmic resistance in a straight wire as will a direct current. However, the situation is different when the conductor is looped.

When an EMF is applied to a loop, the free or released hard electrons below the surface are forced to the outside by centrifugal force, whence a still greater disintegration of soft electrons occurs because the greatest concentration is at the surface. The mean free path of the electrons is greatly increased and the flow continues for a brief period after the EMF travelling in the direction of the current flow ceases. When the EMF continues in the opposite direction as in the case of an alternating current, the force must oppose the momentum of the electron flow still continuing in the opposite direction to that of the new EMF direction. It follows that this impedance will be directly proportional to the number of turns and to the frequency of the AC. It is logical to assume that the deceleration rate of the electron flow is a constant when the EMF is zero. This means that the more quickly that the EMF is applied in the opposite direction, the higher the velocity of flow that will be encountered. It will be a linear function.

It would now seem evident that when the AC is rectified of has been changed to a pulsed DC, the coil will produce an increase in amperage where a straight wire will not. Experiments have confirmed this. It was found that the input amperage of a current was greatly increased after it passed through a coil. The increase was greatest during the initial stage of the applied EMF and soon dropped to a lower value as the concentration of soft electrons around the wire was reduced. It follows that a coil will offer impedance only to an AC current. It follows that pulsed DC has numerous advantages over AC. It can be used to operate transformers as well as AC without suffering impedance.

A steady direct current experiences the same resistance in a coil as it does in a straight wire of the same length. The fluctuating EMF produces extreme agitation of the soft electrons around and inside the wire, resulting in the disintegration of a large percentage of them, and the release of a high concentration of hard electrons. This does not occur during the steady flow of direct current. During the initial application of DC there is a surge of additional current during the build-up of the EMF. When the current is shut off, there will be a momentary surge of current in the opposite direction. The excess of electrons on the surface of the conductor and in the coil will naturally flow towards the void outside the coil and in the opposite direction to which the current was flowing. The concepts just outlined can be applied when building a self-sustaining electric generator.

When an alternating current is applied to a coil, the EMF must overcome the impedance each time the EMF changes direction. The greatest amount of resistance occurs at the beginning of each change and then steadily decreases as the current builds up. The resistance will be at a minimum when the current reaches it’s maximum. With AC, the EMF changes direction very frequently and so the maximum resistance is encountered for a high percentage of the time.

The flow of electrons in a wire results in a circular magnetic flow around that wire. As mentioned previously, the magnetic effects between electrons moving together tend to cancel each other out. They are drawn together and the resulting ethers encompass the entire group. This also occurs between adjacent wire segments in a coil. The magnetic effects are cancelled out between the segments and a continuous ether flow, encompassing the entire coil, perpendicular to the direction of the current flow, will occur. The solenoid will then behave like a bar magnet with continuous lines of force.

The Earth’s atmosphere produces geomagnetism in much the same way that a solenoid produces a

magnetic field. Changes in the atmosphere move along with the Earth in a circular motion. Although there is little motion of the charges relative to the surface, a magnetic field is still created. Magnetic lines, or ethers, flow from the South magnetic region to the North magnetic region as a result of these rotating charges.

25. Despite the fact that our illustrious physicists have managed to develop as highly a destructive device as a nuclear bomb, they still have no concept of the nature and source of the energy released after a detonation. As with all other well-known phenomena, they try to create the illusion that they comprehend and have explained it. As a matter of fact, academic science has not yet supplied satisfactory explanations for any of the simplest and most common everyday phenomena. The energy released by nuclear devices is explained away by stating that it is a conversion of matter into energy in accordance with the false Einstein relation $E = mc^2$. Many readers, especially those steeped in orthodoxy, may be shocked to learn there is no conversion of mass into energy during such a process, nor by any process in which energy is released! The tremendous heat produced in a nuclear blast means that an abnormal quantity of hard electrons were suddenly released by the complete disintegration of all the soft electrons within the area of the explosion. The intense light that accompanies the blast is the result of the photons set free by the disintegration of those soft electrons.

The key to the triggering of the reaction is the neutron. As indicated earlier, a neutron is equivalent to a collapsed hydrogen atom, and yet it is more than this. A hydrogen atom has a strong net positive charge, while the neutron has no net charge. This means that a neutron has collected far more hard electrons than a hydrogen atom. Since a neutron has no charge, it cannot add to the weight of an atom, as is commonly believed.

The concepts introduced in this treatise render all of the old beliefs concerning atomic structure invalid. The weight of an atom is dependent almost entirely on the number of orbital electrons and the number of protons in it's nucleus. This will be discussed in more detail later. There is an exception or two to the above rule in the case of certain radioactive elements where the presence of neutrons can actually reduce the weight of an atom. An interchange of excess electrons between protons and neutrons within the nucleus, and thus transformations of protons into neutrons and vice versa, can occur. The neutrons greatly outnumber the protons in the heavier atoms, especially those that are radioactive. During the interchanges between neutrons and protons, excess neutrons disintegrate into protons and hard electrons are ejected from some of the atoms. This results in a transformation of such atoms. Simultaneously, the tremendous interactions between electrons released in this manner as well as from the disintegration of soft electrons in the vicinity cause the higher ethers to be disturbed, ultimately resulting in the production of gamma rays.

The isotope of the more common uranium 238 atom known as U235 is lighter yet it is fissionable and more radioactive than the uranium 238. It is lighter because it supposedly has fewer neutrons than the ordinary uranium atom. The opposite is actually the case. The U235 having more neutrons is more radioactive. The greater interactions within the nucleus result in more hard electrons being released, which reduces the overall positive charge of it's nucleus.

There is a continuous interchange of ejected protons transforming back into neutrons and vice versa among the U235 atoms. A similar but less violent interchange takes place among the atoms of U238. A low percentage of the U238 atoms receive more than their share of these interchanges and thus transform into U235 atoms. Most of the hard electrons released which contribute to such interchanges and transformations is the result of the disintegration of soft electrons which permeate the atoms. It follows that the main contributing factor of radioactivity is the presence of soft electrons which house the hard electrons! Therefore, if the soft electron concentration throughout the vicinity of a radioactive substance is reduced, it will lose much of it's radioactivity. By now, it has no doubt occurred to the reader, that a Reich cloud-buster pointed at a radioactive material would cause it to lose it's radioactivity! This has been proven to be the case. For example, a glowing piece of radium stops radiating when it is placed in front of a cloud-buster.

The source of the energy released during a nuclear blast is now becoming clear. When a fissionable material like U235 or plutonium is bombarded with additional neutrons, the increased activity in the nuclei causes even the most stable soft electrons in the vicinity to disintegrate. A chain reaction of soft electron disintegration in areas well beyond the confines of the fissionable material results. All of the hard electrons and protons originally camouflaged by the soft particles are suddenly released. A tremendous gamma ray production also occurs. Adequate quantities of fissionable materials suddenly brought together can result in a sufficient increase of neutron bombardment of the interior atoms to produce such a result. It is known as the 'critical mass'. The proper fusion of hydrogen atoms can also cause enough soft electron disintegration to produce a similar result. It is now apparent there is no conversion of mass into energy during the process. All of the fundamental particles of the atoms involved remain intact. In fact, there is even more mass following a blast than there was previously, as a result of the additional hard electrons and protons released.

Once again, it is obvious that the Theory of Relativity is in no way concerned.

The monstrous hoax fostered on the public by the Defence Department of the Government, now becomes more than obvious. A Reich cloud-buster can completely deactivate nuclear devices for great distances by drawing away the soft electron concentration from the vicinity of such device. In fact, a cloud-buster can be used for downing fleets of planes carrying nuclear weapons. Combustion is also dependent on soft electron concentrations which of course includes jet engines. Therefore jet engines or missiles cannot function in an area affected by a cloud-buster. The fact that a simple cloud-buster can deactivate a nuclear reactor from a great distance has been proven on numerous occasions. For example, during the time Reich was carrying out intensive experiments with a cloud-buster in Arizona in the early 1950s, a large reactor several hundred miles to the southeast quit functioning. This means that hundreds of billions of tax dollars are being funnelled every year to support a multibillion dollar nuclear industry and other related industries which are rendered obsolete by the device used by Reich.

It is evident that the proper use of the cloud-buster could throw modern warfare back to the stone age. Obviously the drawing of soft particles away from any group would completely enervate each individual and even turn him into a block of frozen flesh. Although a cloud-buster could not completely deactivate a particle beam weapon it could bring down any craft carrying such a device before it could get into position. The potential of the cloud-buster is perhaps greater than even Reich himself realised. Since heat is transferred from one body to another by soft electrons which release harder electrons, the cloud-buster can be used as a highly efficient refrigeration system by drawing soft electrons away from a body. It has been made apparent that this simple device can render present fire fighting techniques obsolete. By use of the cloud-buster in the proper manner, the loss of life and property from fire and storms could become a thing of the past. It also provides dramatic proof of the validity of many of the new concepts introduced in this treatise.

Radioactivity was the subject of a ridiculous, if not amusing, fiasco more than two decades ago when two physicists, Lee and Yang, received the Nobel Prize in 1957. The incident, which was given wide publicity, concerned an error in the parity principle. The parity principle has been defined as "a mathematical concept impossible to define in physical terms". How such a concept could have any bearing on physical reality is not made clear. Generally, anything relating to reality can be defined in terms of reality, which is in conformity with the Law of Cause and Effect.

Incredibly, an experiment was devised to test the validity of this great revelation. It was based on the idea that a radioactive substance should eject more particles in one preferred direction, than in any other. Radioactive cobalt was chosen. It was cooled down to near absolute zero and exposed to a powerful magnetic field produced by a solenoid, in order to align the nuclei. Another physicist, a Dr. Wu, had devoted six months of hard work setting up the experiment. Indeed, it was found that more particles were ejected out of one pole of the solenoid than the other. Which pole was it? Of course, it was the pole out of which the magnetic lines flowed. Naturally, the experiment merely demonstrated that particles tend to travel down magnetic lines of force. The excess of particles that came out of the pole were those barely ejected from the atom. They had such a low initial velocity that, regardless of what direction they happened to be travelling initially, the magnetic field would dictate their final direction of travel.

Lee and Yang were accorded every scientific honour, including the Nobel Prize, as a result of this experiment. Instead of giving them the Nobel Prize, the awarding of an Oscar would have been more appropriate. Accompanying the extensive publicity given this comedy act was a photo appearing in a prominent magazine showing one of the recipients pointing to a grotesque mathematical equation containing over 100 terms! He was allegedly explaining the reasoning behind their great revelation.

The great nuclear scare scam should rate as the biggest lie of the century. This fabrication is so colossal even Mr Cater is somewhat embarrassed to admit he was also taken in by it, as everyone else was, except those directly involved in the conspiracy but Mr Cater never questioned it as he had other universally accepted beliefs. The long-hidden truth is this: radiation from radioactive substances is relatively harmless! One can actually swim in water from so-called high level nuclear waste, drink the water and actually thrive on it. Also one can handle pure U235 and even plutonium (the fuel for A-bombs) with bare hands all day long, and suffer no ill effects. Of course excessive exposure to this radiation can be inimical, as can occur from any other type of radiation such as sunlight. The effects, however, are immediate and not long-lasting if the source of the trouble is removed. The popular belief that mutations can result from this radiation that effects future generations, is complete nonsense. A study of the effects on the victims of Hiroshima and Nagasaki bear this out. The only mutations are immediate damage to cells and subsequent malfunctions from ultra-high concentrations.

For the above revelations we are indebted to a courageous and dedicated individual who worked for 35 years in the nuclear industry. He was intimately involved in every aspect of the production of nuclear fuels and the building of reactors. He is Galen Windsor, of Richland, Washington. The so-called lethal nuclear radiations are relatively harmless, with very limited penetration. In 1987, Galen had already lectured in 77 different cities over a two-year period. His credentials are impressive. He has worked in every major reactor decommissioning project in America. He was involved in the analytical process inventory control, which was responsible for measuring and controlling the nuclear fuel inventory for these projects. He has few peers in this field and all of them agree with him but are afraid to speak out. He is one of the world's greatest authorities of nuclear radiation measurement.

He and others did plutonium processing using their bare hands until radiation monitors were installed at every reactor. Rules were laid down as to the amount of exposure to radiation that must not be exceeded. If the rules were not followed, the worker disappeared and was never seen again by anyone. The reason is obvious. If word leaked out that radioactive materials below critical mass were harmless, then there would be widespread pilfering of such products. The lie that radioactive materials can be handled only with extreme safety precautions must be maintained at all costs. The high-level disposal act of 1982 calls for the permanent disposal of the so-called high level waste 3,000 feet underground. This high level waste consists of reusable uranium fuel and contains useful metal isotopes. Every ton of it is worth about 10 million US dollars.

Enough has been collected to pay off the National Debt. Portland General Electric owns the Trojan reactor and had a storage basin problem. Windsor offered to take all of their spent fuel off their hands. He would ship it, store it, and do everything that needed to be done at no expense to them, if they would give it to him. They told him "Go to Hell, Galen Windsor - we value it more than plutonium or gold. We are going to play the plutonium future ourselves".

Windsor was taken to California in 1965 to help design and build a nuclear fuel reprocessing plant. After it was built in 1973, a presidential order was issued which said the plant was not to be run. At that time, they had 170 metric tons of spent fuel stored in the basin. Also, the maximum allowed exposure was further reduced by a factor of 10. This was the beginning of Windsor's rebellion against "The Establishment". He began swimming in a 660,000 gallon pool containing the spent fuel. The radioactive materials it contained were enough to maintain a water temperature of 100° F and the water would glow with a bluish light in the dark. He discovered that the tank provided good drinking water.

Windsor was also asked about what was in the burial sites of so-called low-level nuclear waste. The answer was that there are no nuclear wastes, only materials produced in a reactor to be recovered and used beneficially. Low-level waste is an excuse for a Federally-mandated non-inspectable disposal system, so that organised crime can get rid of any evidence they want, and that it can never be dug up again. Also, so that no one finds out whose bodies are in those drums. This is what is contained in barrels that are dropped in the ocean.

Windsor also stated, (and he was in a position to know), that in 1947 the United States sent to the USSR all the necessary materials and technology to build A-bombs! This was under the approval of President Truman. In 1949, they exploded their first A-bomb. Later, the Rosenbergs were executed for allegedly turning over nuclear "secrets" to the Russians.

According to Windsor, by 1975 large reactors no longer had a future. They are being phased out. The notorious Three-Mile Island incident was no accident. It did no damage and no one was harmed. But, it did impress upon the public the alleged danger of reactors. Windsor advocated the use of many small reactors systematically distributed. They could be not only used for producing adequate quantities of electricity, but for obviating refrigeration of packaged foods. When briefly exposed to the radiation from such reactors they will keep indefinitely. A Federal energy cartel is the reason that the effective use of many smaller reactors is not implemented. They control the amount of electricity, the availability and the price. Windsor stated that no reactors have been built correctly. In any event, reactors can in no way, match the use of self-sustaining electric generators for electricity, but their use in preserving food is intriguing.

From the above, it becomes obvious that recent publicity concerning the danger of radon gas in homes is another monstrous hoax! It could be mixed with oxygen in sizeable quantities and breathed for long periods without damage.

The question now arising is why are the radiations from radioactive sources relatively harmless? Three types of radiation emanate from them: alpha, beta and gamma rays. Alpha rays consist of helium nuclei or a small group of protons, neutrons and electrons. The number of neutrons and protons are still unknown

despite the vaunted claims of nuclear physicists. Since they are totally ignorant as to their true nature, the nature of gravity and soft particle physics, they have no means of determining such values. Beta particles are the hard electrons of ordinary electricity, and gamma rays are negative charges comprised of ether particles in the same range as that of hard electrons. As such, they have no more penetrating ability than a stream of hard electrons and are no more damaging than the hard electrons when they do penetrate. Since they are comprised of the same ethers as atoms and molecules, it follows that they will not penetrate matter as readily as softer particles. This leaves the alpha rays. It has already been shown that positive charges are inimical to living tissues since they tend to absorb the negative charges living organisms require. However, alpha particles also have little penetrating ability for the same reasons as given above, and quickly absorb electrons to become harmless helium gas. Windsor gave demonstrations of such facts at his lectures by use of a Geiger counter.

The capture of large quantities of alpha particles by high concentrations of orgone energy is another matter. Reich termed it deadly orgone energy. The orgone enters the body and releases alpha particles throughout vital organs and produces what is known as radiation sickness. Reich and his assistants had this experience. When the source of the trouble was removed they quickly recovered. Occasionally, deadly orgone can remain in the body and at later periods, release the positive charges at various times, thus causing severe after-effects. This problem can be alleviated by treatments inside orgone accumulators or pyramids after being exposed to the deadly orgone. This is what Reich and his helpers did.

Galen Windsor also stated that the bomb dropped on Hiroshima contained 20 pounds of U235. The one exploded over Nagasaki had 2.5 kilograms or about 5 pounds of plutonium. This values were jealously guarded secrets for a long time. Windsor is likely the first one who has dared reveal it to the public. The author was acquainted with the late Stanislaw Ulam, a mathematician who had calculated the critical mass of these elements for bombs while working on the Manhattan project. It was a deep, dark secret and he would have parted with his life before he would reveal anything so sacred. The degree of the hang-up officialdom has on secrecy is of a magnitude difficult to comprehend, and the steps they have taken to insure it is even more mind-boggling. It is so infantile and childish that there are no words in the English language to describe it adequately.

Radioactive substances such as plutonium, produce their own heat and the larger the mass, the higher the temperature of the mass. This follows identically, the same principle by which the Sun generates it's energy or the ratio between mass and surface area. The essential difference is that plutonium can do in a few cubic inches what ordinary matter does in a ball 2,500,000 miles in diameter!

Interestingly enough, Windsor looks about 30 years younger than his chronological age would indicate. He also stated he always sustained a healthy tan from his work with radioactive substances. It is significant that Reich and his assistant had a similar experience in their work with orgone energy. It is clear Windsor and Reich were exposed to the same energies, high concentrations of soft electrons. The constant exposure to high orgone concentrations during his working hours is what has kept Galen Windsor young. Of course the high orgone concentrations was produced by soft electrons clustering around the radioactive substances which produced a constant flow of positive charges.

You may wonder why Windsor and certain others could handle radioactive substances with impunity while others have either died or suffered from serious physical problems after being exposed to the radiation. As mentioned earlier, radioactive materials radiate highly positively charged particles which are extremely inimical. Some are absorbed by soft electrons and become DOR (deadly orgone energy). They can enter the body and remain for long periods before disintegrating and releasing the deadly particles. These particles are very sluggish when compared to negative particles including soft electrons.

The victims of radioactive fallout are not exposed for any significant period to high concentrations of positive charges. Also, the area of exposure is very large; consequently, the Coanda effect does not come into play and there is no rush of orgone energy to alleviate the situation as was the case with Windsor when he was directly exposed to ultra high concentrations of radioactive materials. Any victim of radioactive contamination can be cured by being exposed to high concentrations of orgone for extended periods as was Reich and his assistants after experiencing radioactive sickness. They had received a heavy dose of DOR as mentioned earlier. Being exposed to high concentrations of radioactivity is equivalent to being placed in an orgone accumulator.

This is only a summary of part of Mr Cater's book which has 586 pages. A number of topics are not even mentioned here. Mr Cater also speaks of:

Superconductivity.
The properties of Helium at low temperatures.
Mystery spots on Earth, including the famous Oregon Vortex.
Gravitational anomalies.
The origin of the Van Allen Radiation Belt.
The research of William Reich.
Orgone energy.
The Oranur Experiment.
The Reich Cloud Buster (which is also a weapon of major power).
Atmospheric phenomena.
Three practical Free-Energy devices.
The great potential of crystals in the obtaining of Free-Energy.
The work of Nikola Tesla with Free-Energy.
The Searle Effect and many other topics.

If you wish to buy a copy of his book, it is supposedly available at these two outlets: http://www.healthresearchbooks.com/ or in the UK from http://www.amazon.co.uk. I have ordered and paid for a copy of this book on three occasions and every time, the supplier claimed to be unable to supply what was advertised for sale

So, having absorbed some of what Mr Cater has to say, do you feel that you have been treated fairly, and that the true scientific details were presented to you as part of your general education?

Patrick Kelly

http://www.free-energy-info.tuks.nl/

Chapter 12: Basic Electronics

Introduction

This document is not an in-depth presentation of the subject of electronics. Instead, it is intended to give you sufficient (empirical) knowledge of the subject to be able to understand, design and build simple circuits such as the control circuits used with the 'Free Energy' devices described in the later parts of this eBook.

Disclaimer

This material is provided for information purposes only. Should you decide to attempt construction of some device based on information presented here and injure yourself or any other person, I am not liable in any way. To clarify this; should you construct something in a heavy box and drop it on your toe, I am not liable for any injury you may sustain (you should learn to be more careful). If you attempt to construct some electronic circuit and burn yourself with the soldering iron, I am not liable. Also, I strongly recommend that unless you are expert in electronics, you do not construct any device using, or producing more than 30 Volts - high voltage circuits are extremely dangerous and should be avoided until you gain experience or can obtain the help and supervision of a person experienced in constructing high voltage circuits.

Voltage.

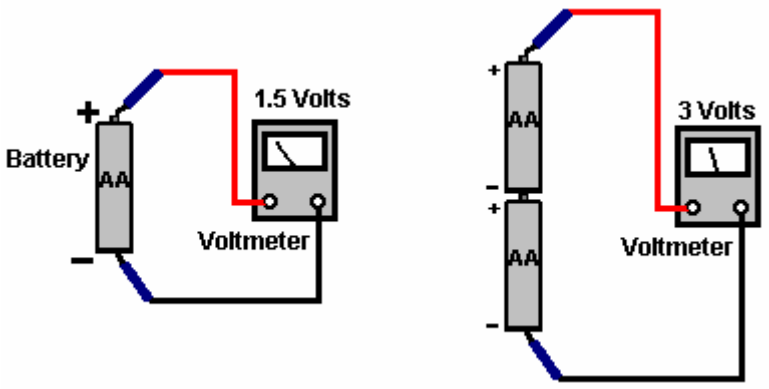
Voltage is the key to understanding electronics. Without voltage, nothing happens in electronics. What is it? Nobody knows. We know how to generate it. We know what it does. We know how to measure it, but nobody knows what it actually is.

It is also called "Electro Motive Force" or "EMF" which is no help whatsoever in knowing ***what*** it is. That, is roughly equivalent to saying "the thing that pushes is the thing that pushes" - very true but absolutely no help whatsoever. OK, having admitted that we really don't know *what* it is, we can start to say the things we do know about it:

A new battery has a voltage between its terminals. This voltage is said to cause a current to flow through any complete electrical circuit placed across it. The current flowing through the circuit can cause various things to happen such as creating light, creating sound, creating heat, creating magnetism, creating movement, creating sparks, etc., etc.

By using the current caused by a voltage, a device called a 'Voltmeter' can indicate how big the voltage is. The bigger the voltage, the bigger the current and the bigger the display on the voltmeter. The voltmeter can have a numerical display where you read the voltage directly from the display, or it can be an 'analogue' voltmeter where the voltage is shown by the position of a needle on a scale. The size of the voltage is stated in 'Volts' which is a unit of measurement named after the man Volta who introduced voltage to the world (it was always there, we just did not know about it).

Voltages add up if they are connected the same way round, i.e. with the + terminals all facing the same way:



The physical size of the battery usually determines the length of time it can supply any given current - the bigger the battery, the longer it can provide any given current. A battery is constructed from a number of 'cells'. The number of cells in the battery controls the voltage of the battery. For example, an 'AA' size battery (what used to be called a 'penlight' battery) has a single 'cell' and so produces 1.5 Volts when new. The very much larger and heavier 'D' battery also has just one cell and so it also produces 1.5 Volts when new. The difference (apart from the higher cost of the 'D' cell) is that the larger cell can provide a much higher current if both batteries are discharged over the same period of time.

There are several different types of battery construction. A rechargeable NiCad battery has a single cell but its construction method means that it produces about 1.35 Volts when fully charged. In passing, NiCad batteries have a 'memory' characteristic which means that if they are recharged before they are fully discharged, then the next time they are discharged they run out of power at the voltage level it had when the last charging was started. Consequently, it is a good idea to fully discharge a NiCad battery before charging it again.

Car and motorcycle batteries are described as Lead/Acid batteries. This type of construction is not very convenient being large, heavy and potentially corrosive. The big advantages are the ability to provide very high currents and giving 2.0 Volts per cell. These batteries are normally produced as 6 Volt or 12 Volt units. The Amp-Hours for lead/acid car batteries is usually quoted for a 20 hour discharge period, so a fully charged, new, 20 AHr battery can provide 1 Amp for 20 hours of continuous use. That battery loaded to give 5 Amps, will **not** provide that current for 4 hours but might only last 2 hours, or perhaps a little better. The manufacturers literature should give an indication of the performance, but if it is important, run your own test to see how the battery actually works in practice.

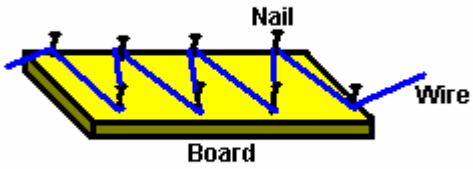
"Mains units" are known in the electronics world as "Power Supply Units" or "PSUs" for short. These convert the mains voltage (220 Volts in UK, 110 Volts in USA) to some convenient low voltage; 12 Volts, 9 Volts, 6 Volts, or whatever is needed. A mains unit can provide several different voltages simultaneously.

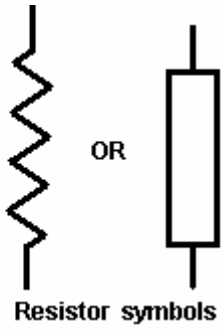
Resistance.

Being familiar with Voltage and Resistance is the key to understanding electronic circuitry. Resistance is a measure of how difficult it is for current to flow through something. Some materials such as glass, ceramics, wood and most plastics do not easily carry a current and so are considered to be 'insulators'. That is why you will see power lines hung from their pylons by a series of ceramic discs. Current flows easily through metals, especially along the surface of the metal, so cables are made from metal wires surrounded by a layer of plastic insulation. The higher grade cables have wire cores made up of many small-diameter strands as this increases the surface area of the metal for any given cross-sectional area of the metal core (it also makes the cable more flexible, and generally, more expensive).

There is a very important, third group of materials, silicon and germanium in particular, which fall between conductors and insulators. Not surprisingly, these are called 'semi-conductors' and the amount of current they can carry depends on the electrical conditions in which they are placed. Much, much more about this later on.

While a metal wire carries current very well, it is not perfect at the job and so has some 'resistance' to current flowing through it. The thicker the wire, the lower the resistance. The shorter the wire, the lower the resistance. The first researchers used this characteristic to control the way circuits operated. Sometimes, as higher resistances were needed, the researcher used to need long lengths of wire which would get tangled up. To control the wire, a board with nails along each side was used and the wire wound backwards and forwards across the board like this:

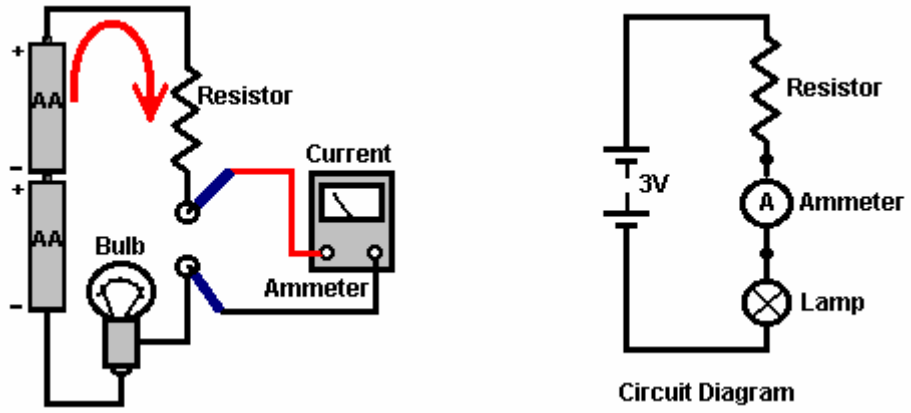




Resistor symbols

When drawing a circuit diagram, the researcher would sketch the wire on the board giving a zig-zag line which is still used today to represent a 'resistor' although different methods of construction are now used. An alternative symbol for a resistor is a plain rectangle as shown above.

If a resistor is connected across a battery, a circuit is formed and a current flows around the circuit. The current cannot be seen but that does not mean that it is not there. Current is measured in 'Amps' and the instrument used to display it is an 'ammeter'. If we place an ammeter in the circuit, it will show the current flowing around the circuit. In passing, the ammeter itself, has a small resistance and so putting it in the circuit does reduce the current flow around the circuit very slightly. Also shown is a bulb. If the current flowing around the circuit is sufficiently high and the bulb chosen correctly, then the bulb will light up, showing that current is flowing, while the ammeter will indicate exactly how much current is flowing:



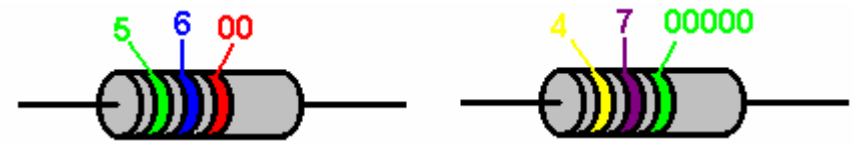
Shown on the right, is the way that this circuit would be shown by an electronics expert (the 'Resistor', 'Ammeter' and 'Lamp' labels would almost certainly not be shown). There are several different styles of drawing circuit diagrams, but they are the same in the basic essentials. One important common feature is that unless there is some very unusual and powerful reason not to do so, every standard style circuit diagram will have the positive voltage line horizontally at the top of the diagram and the negative as a horizontal line at the bottom. These are often referred to as the positive and negative 'rails'. Where possible, the circuit is drawn so that its operation takes place from left to right, i.e. the first action taken by the circuit is on the left and the last action is placed on the right.

Resistors are manufactured in several sizes and varieties. They come in 'fixed' and 'variable' versions. The most commonly used are the 'fixed' carbon 'E12' range. This is a range of values which has 12 resistor values which repeat: 10, 12, 15, 18, 22, 27, 33, 39, 47, 56, 68, 82 and then: 100, 120, 150, 180, 220, 270, 330, 390, 470, 560, 680, 820 and then: 1000, 1200, 1500, 1800, 2200, 2700, 3300, 3900, 4700, 5600, 6800, 8200, etc. etc. Nowadays, circuits often carry very little power and so the resistors can, and are, made in very small physical sizes. The higher the resistance value of a resistor, the less current will flow through it when a voltage is placed across it. As it can be difficult to see printing on small resistors clustered together on a circuit board and surrounded by other larger components, the resistor values are not written on the resistors, instead, the resistors are colour-coded. The unit of measurement for resistors is the 'ohm' which has a very small size. Most resistors which you encounter will be in the range 100 ohms to 1,000,000 ohms. The higher the resistance of any resistor, the smaller the current which will flow through it.

The colour code used on resistors is:

0 Black
1 Brown
2 Red
3 Orange
4 Yellow
5 Green
6 Blue
7 Purple (Violet if your colour vision is very good)
8 Grey
9 White

Each resistor has typically, three colour bands to indicate its value. The first two bands are the numbers and the third band is the number of noughts:



Green: 5
Blue: 6
Red: 2 noughts
Value: 5,600 ohms or 5.6K or 5K6

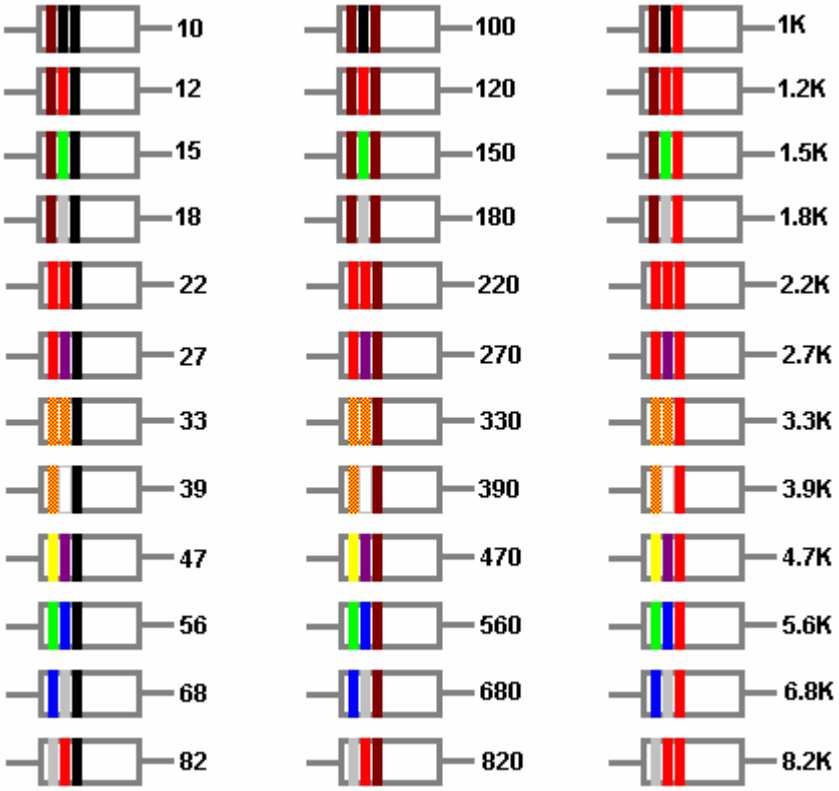
Yellow: 4
Purple: 7
Green: 5 noughts
Value: 4,700,000 ohms or 4.7M or 4M7

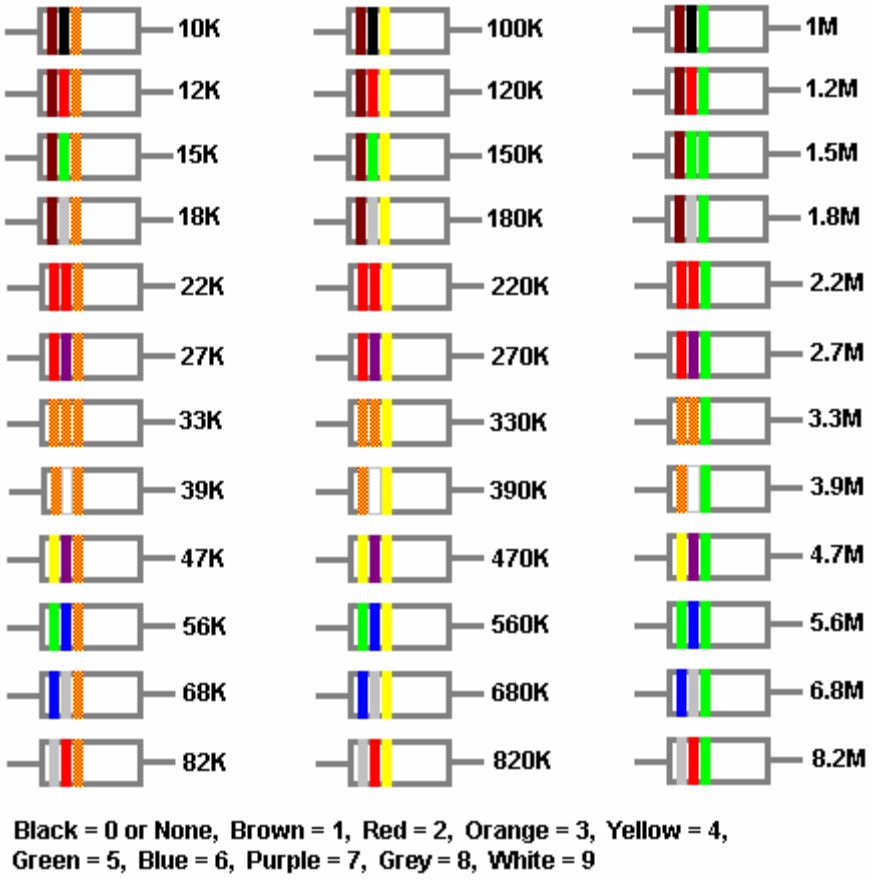
The colour bands are read from left to right and the first band is close to one end of the body of the resistor. There is often a fourth band which indicates the manufacturing tolerance: you can ignore that band.

Examples:

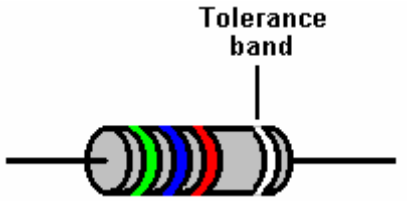
Red, Red, Red: 2 2 00 ohms or 2K2
Yellow, Purple, Orange: 4 7 000 ohms or 47K
Brown, Black, Brown: 1 0 0 ohms or 100R
Orange, Orange, Orange: 3 3 000 ohms or 33K
Brown, Green, Red: 1 5 00 ohms or 1K5
Brown, Green, Black: 1 5 no noughts, or 15 ohms
Blue, Grey, Orange: 6 8 000 ohms or 68K
Brown, Green, Green: 1 5 00000 ohms or 1,500,000 ohms or 1M5
Yellow, Purple, Brown: 4 7 0 ohms

As there are only 12 standard resistor values per decade, there are only 12 sets of the first two colour bands:
10: Brown/Black,
12: Brown/Red,
15: Brown/Green,
18: Brown/Grey
22: Red/Red,
27: Red/Purple
33: Orange/Orange,
39: Orange/White
47: Yellow/Purple
56: Green/Blue
68: Blue/Grey
82: Grey/Red





The details above give you all the basic information on resistor colour codes but there are a few additional refinements. There is an extra colour band further down the body of the resistor as shown here:

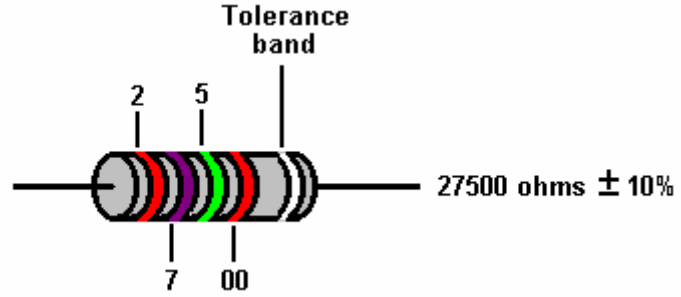


This extra band is used to indicate the manufacturing tolerance of the construction of the resistor. Resistor values are never exact and this rarely has any significant effect on their use in circuits. If some circuit needs very accurate resistor values in it, then buy several resistors of the same nominal value and use an ohm-meter to measure that actual value of each particular resistor and if none are perfect, then use two or more resistors to give the exact value wanted.

The tolerance band has the following codes:

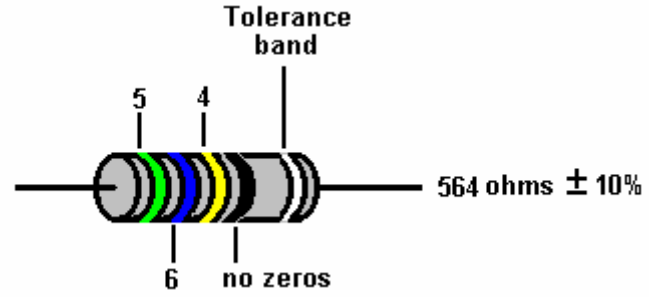
Silver is ± 10% (i.e. a 10K resistor of this type should be between 9K and 11K)
Gold ± 5% (i.e. a 10K resistor of this type should be between 9.5K and 10.5K)
Red ± 2% (i.e. a 10K resistor of this type should be between 9.8K and 10.2K)
Brown ± 1% (i.e. a 10K resistor of this type should be between 9.9K and 10.1K)
Green ± 0.5% (i.e. a 10K resistor of this type should be between 9.95K and 10.05K)
Blue ± 0.25% (i.e. a 10K resistor of this type should be between 9.975K and 10.025K)
Purple ± 0.1% (i.e. a 10K resistor of this type should be between 9.99K and 10.01K)

This type of resistor in the 10% and 5% ranges are the most common as they are the cheapest to buy and so tend to be the most popular. Recently, however, two additions to the coding have been introduced in order to allow for very high specification resistors which the average constructor may never come across. Each of these additions involves one additional colour band. The first additional colour band allows an extra digit in the resistor value, and looks like this:



As before, the colour coding is exactly the same, with the fourth colour band specifying the number of zeros after the digits indicated by the colour bands in front of it. So, in the example shown above, the first band being Red indicates a "2". The second colour band being Purple indicates a "7". The third colour band being Green indicates a "5" and the fourth colour band being Red indicates "2 zeros", so putting those together it produces the value of 27,500 ohms, which can also be written as 27.5 K or more briefly as 27K5.

Another example of this is:

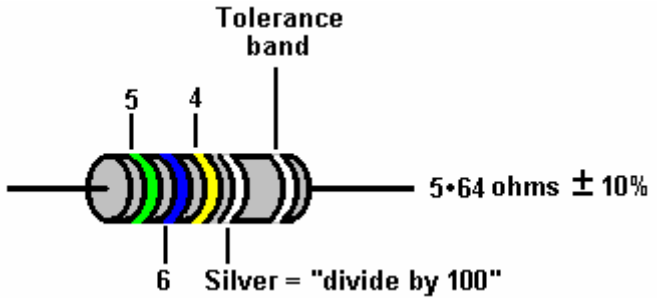


The fourth colour band coding has also been extended to include two other colours:

Gold: meaning "no zeros and divided by 10" so if the band in the example above had been gold, then the value would be 56.4 ohms.

Silver: meaning "no zeros and divided by 100" and if the example band had been silver then the value would have been 5.64 ohms.

So, for example, if the resistor had a fourth colour band which was silver, then the value would be:

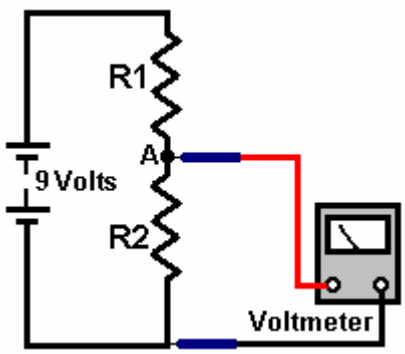


Finally, for very high-quality applications (typically military applications), there can be a sixth colour band positioned outside the tolerance band, and that final colour band states how much the resistance value can be expected to alter with changes in temperature. This is not something which is likely to be of any interest to you, but the codes for that final colour band are:

Brown: 0.01% of the resistor value for each degree Centigrade change in temperature.
Red: 0.005% of the resistor value for each degree Centigrade change in temperature.
Yellow: 0.0025% of the resistor value for each degree Centigrade change in temperature.
Orange: 0.0015% of the resistor value for each degree Centigrade change in temperature.

To put this in context, the worst of these represents a change of 1% in the resistor value when moving from the temperature of ice to the temperature of boiling water. Is this something which you really care about? I don't.

Leaving the details of identifying individual resistors, we now come to the interesting part: what happens when there are several resistors in a circuit. The important thing is to keep track of the voltages generated within the circuit. These define the currents flowing, the power used and the way in which the circuit will respond to external events. Take this circuit:



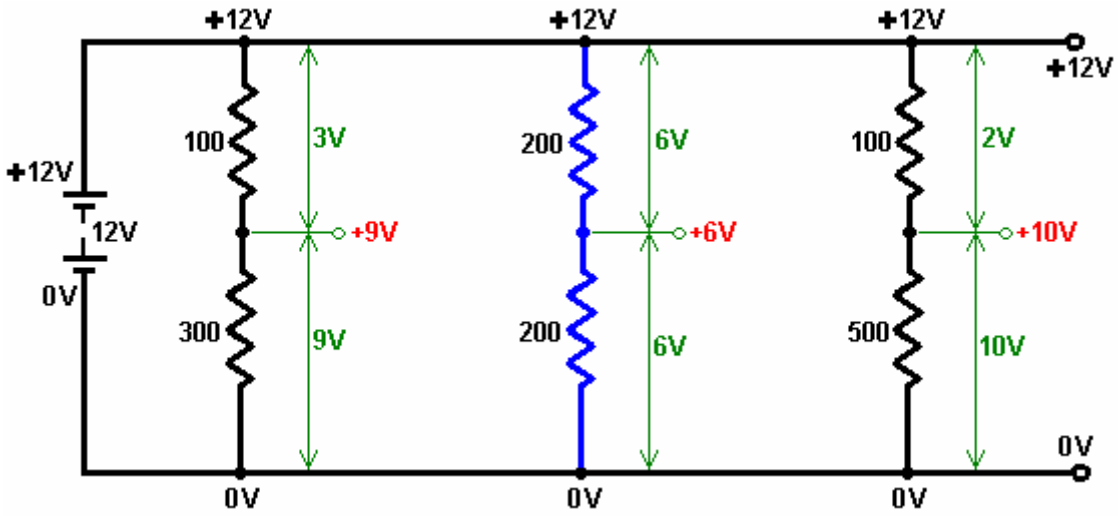
What is the voltage at point 'A'? If you feel like saying "Who cares?" then the answer is "you" if you want to understand how circuits work, because the voltage at point 'A' is vital. For the moment, ignore the effect of the voltmeter used to measure the voltage.

If R1 has the same resistance as R2, then the voltage at 'A' is half the battery voltage, i.e. 4.5 Volts. Half the battery voltage is dropped across R1 and half across R2. It does not matter what the actual resistance of R1 or R2 is, as long as they have exactly the same resistance. The higher the resistance, the less current flows, the longer the battery lasts and the more difficult it is to measure the voltage accurately.

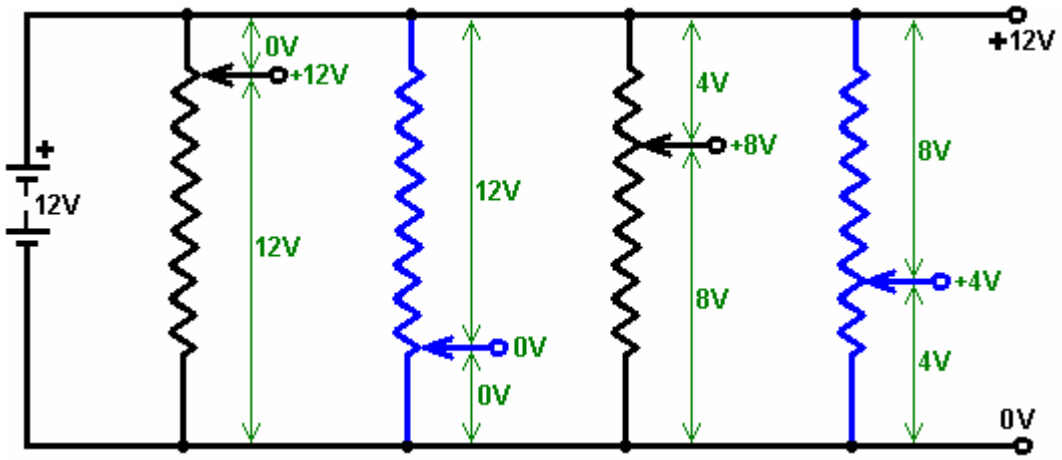
There is no need to do any calculations to determine the voltage at point "A" as it is the ratio of the resistor values which determines the voltage. If you really want to, you can calculate the voltage although it is not necessary. The method for doing this will be shown you shortly. For example, if R1 and R2 each have a value of 50 ohms, then the current flowing through them will be 9 volts / 100 ohms = 0.09 Amps (or 90 milliamps). The voltage drop across R1 will be 50 ohms = Volts / 0.09 amps or Volts = 4.5 volts. Exactly the same calculation shows that the voltage across R2 is exactly 4.5 volts as well. However, the point to be stressed here is that it is the **ratio** of R1 to R2 which controls the voltage at point "A".

If R1 has half as much resistance as R2, then half as much voltage is dropped across it as is dropped across R2, i.e. 3 Volts is dropped across R1, giving point 'A' a voltage of 6 Volts and that is what the voltmeter will show. Again, it does not matter what the actual value of R1 is in ohms, so long as R2 has exactly twice the resistance (shown by a higher number on the resistor).

If R1 has twice as much resistance as R2, then twice as much voltage is dropped across it as is dropped across R2, i.e. 6 Volts is dropped across R1, giving point 'A' a voltage of 3 Volts. Here are some examples with different resistors:

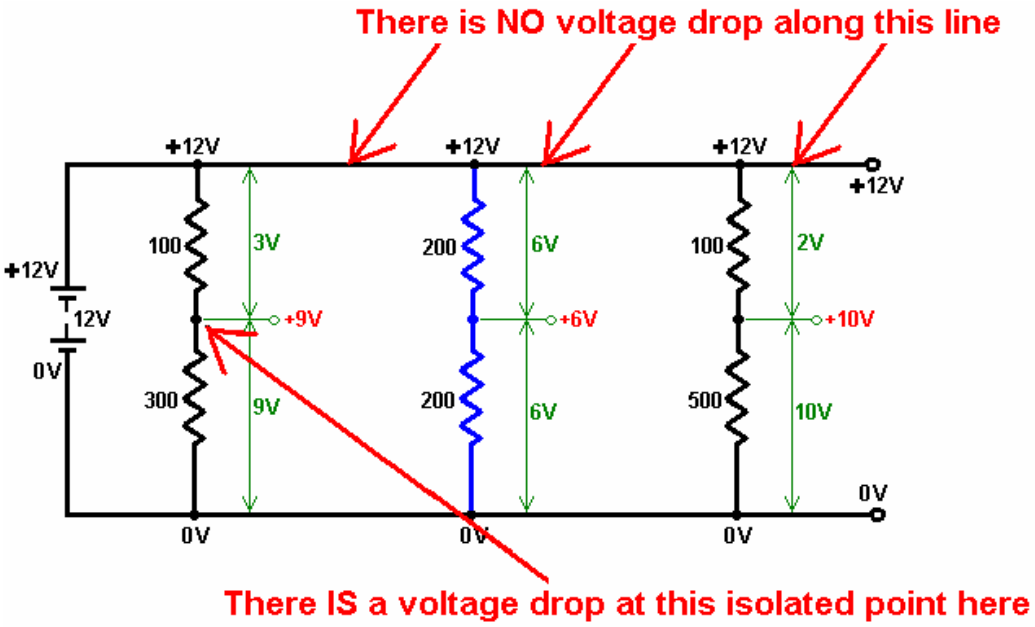


The same division of the supply voltage can be produced by positioning the slider of a variable resistor at different points by rotating the shaft of the device:



This determination of the voltage levels is the key factor to understanding electronic circuitry. The voltage levels control what currents flow and how every circuit will perform, so it is essential to understand what is happening. Stick with this section until you understand it, and if necessary, ask questions about what you find difficult.

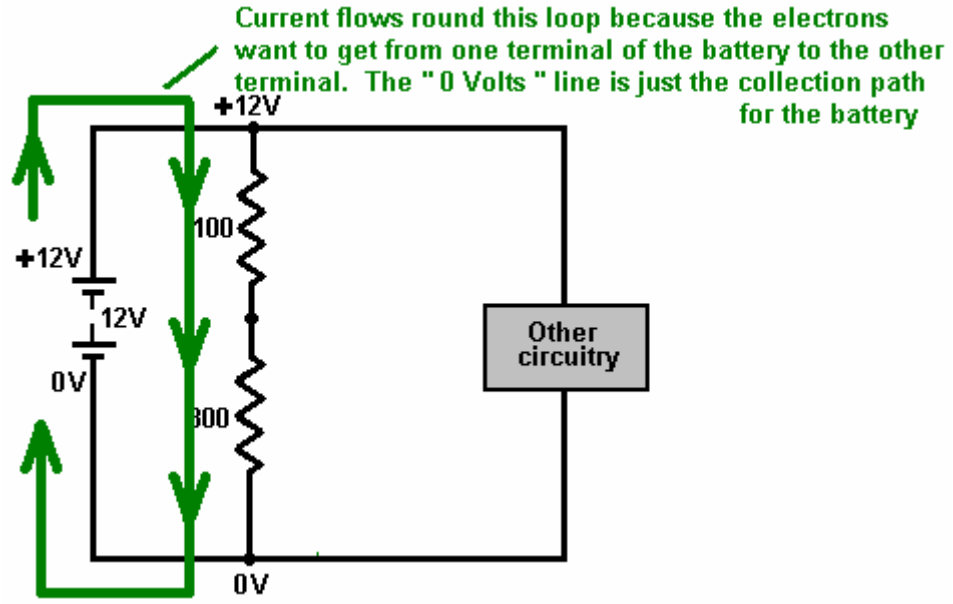
First, please understand that a good battery is an unlimited source of voltage and that voltage does not get "used up" when a resistor or whatever is connected across it:



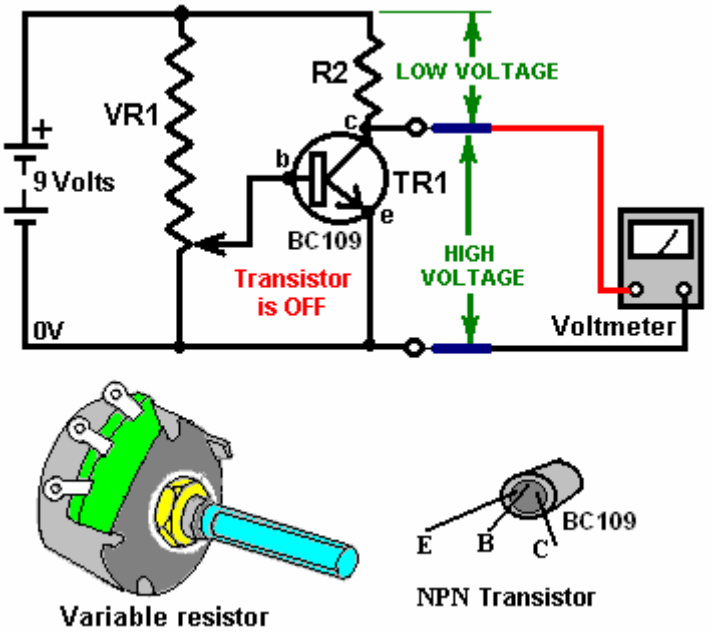
There can be some difficulty in understanding the "0-volt" connection in a circuit. All this means is that it is the return line for current flowing from the battery. Most conventional circuits are connected to both sides of the battery and that allows a current to flow around a closed "circuit" from one terminal of the battery to the other terminal.

It is normal practice to draw a circuit diagram so that the Plus terminal of the battery is at the top and the minus terminal is at the bottom. Many circuit diagrams show the negative line at the bottom connected to the ground or an "earth" connection, which is literally a metal rod driven into the ground to make a good electrical connection to the ground. This is done because the Earth is literally a vast reservoir of negative electricity. However, in reality, most circuits are not connected directly to the Earth in any way. The standard circuit diagram can be visualised as being like a graph of voltage, the higher up the diagram, the higher the voltage.

Anyway, when there is a circuit connected across the battery, the negative or "0V" line just indicates the return path to the battery for the current flow:



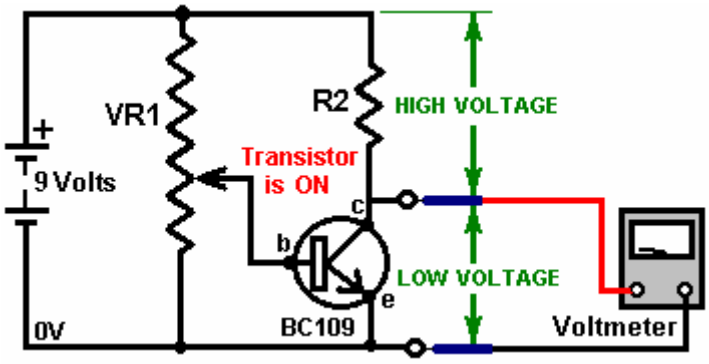
This principle applies immediately to the following circuit:



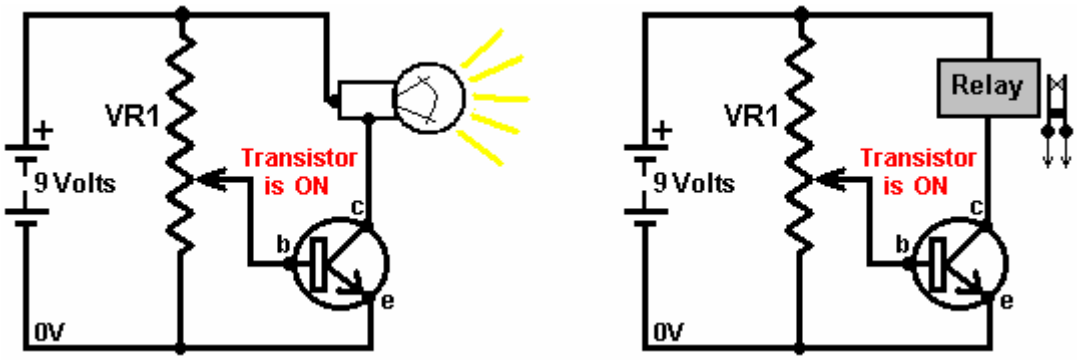
Here we encounter two new components. The first is 'VR1' which is a variable resistor. This device is a resistor which has a slider which can be moved from one end of the resistor to the other. In the circuit above, the variable resistor is connected across the 9 Volt battery so the top of the resistor is at +9 Volts (relative to the battery Minus terminal) and the bottom is at 0 Volts. The voltage on the slider can be adjusted from 0 Volts to 9 Volts by moving it along the resistor by turning the shaft of the component (which normally has a knob attached to it).

The second new device is 'TR1' a transistor. This semiconductor has three connections: a **C**ollector, a **B**ase and an **E**mitter. If the voltage on the base is below 0.7 volts, then the transistor is said to be "OFF" and in that state it has a very high resistance between the collector and the emitter, much higher than the resistance of resistor "R2". The voltage dividing mechanism just discussed means that the voltage at the collector will therefore, be very near to 9 Volts - caused by the **ratio** of the transistor's **C**ollector/**E**mitter resistance compared to the resistor "R2".

If the voltage on the base of the transistor is raised to 0.7 volts by moving the slider of the variable resistor slowly upwards, then this will feed a small current to the base which then flows out through the emitter, switching the transistor ON causing the resistance between the collector and the emitter to drop instantly to a very low value, much, much lower than the resistance of resistor 'R2'. This means that the voltage at the collector will be very close to 0 Volts. The transistor can therefore be switched on and off just by rotating the shaft of the variable resistor:

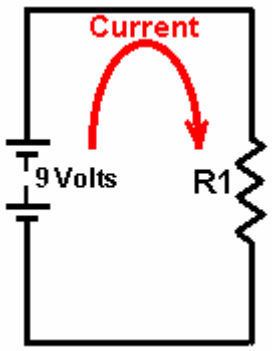


If a bulb is used instead of R2, then it will light when the transistor switches on. If a relay or opto-isolator is used, then a second circuit can be operated:



If a buzzer is substituted for R2, then an audible warning will be sounded when the transistor switches on. If a light-dependent resistor is substituted for VR1, then the transistor will switch on when the light level increases or decreases, depending on how the sensor is connected. If a thermistor is used instead of VR1, then the transistor can be switched on by a rise or fall in temperature. The same goes for sound, windspeed, water speed, vibration level, etc. etc. - more of this later.

We need to examine the resistor circuit in more detail:



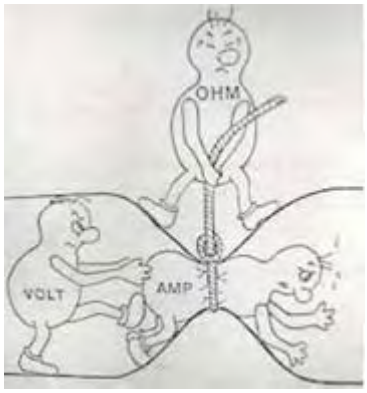
We need to be able to calculate what current is flowing around the circuit. If the circuit contains only resistors, then this can be done using "Ohms Law" which states that "Resistance equals Voltage divided by Current" or, if you prefer:

Ohm's Law (resistive Circuits only).

"**Ohms = Volts / Amps**" which indicates the units of measurement.

In the circuit above, if the voltage is 9 Volts and the resistor is 100 ohms, then by using Ohm's Law we can calculate the current flowing around the circuit as 100 Ohms = 9 Volts / Amps, or Amps = 9 / 100 which equals 0.09 Amps. To avoid decimal places, the unit of 1 milliamp is used. There are 1000 milliamps in 1 Amp. The

current just calculated would commonly be expressed as 90 milliamps which is written as 90 mA.



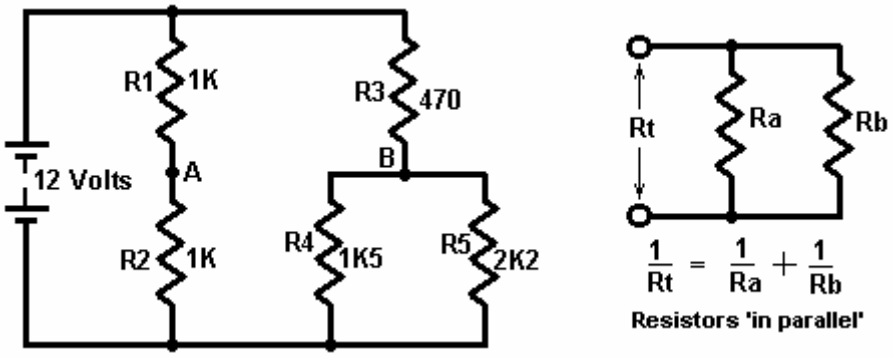
In the circuit above, if the voltage is 9 Volts and the resistor is 330 ohms, then by using Ohm's Law we can calculate the current flowing around the circuit as 330 = 9 / Amps. Multiplying both sides of the equation by "Amps" gives:
Amps x 330 ohms = 9 volts. Dividing both sides of the equation by 330 gives:
Amps = 9 volts / 330 ohms which works out as 0.027 Amps, written as 27 mA.

Using Ohm's Law we can calculate what resistor to use to give any required current flow. If the voltage is 12 Volts and the required current is 250 mA then as Ohms = Volts / Amps, the resistor needed is given by: Ohms = 12 / 0.25 Amps which equals 48 ohms. The closest standard resistor is 47 ohms (Yellow / Purple / Black).

The final thing to do is to check the wattage of the resistor to make sure that the resistor will not burn out when connected in the proposed circuit. The power calculation is given by:
Watts = Volts x Amps. In the last example, this gives Watts = 12 x 0.25, which is 3 Watts. This is much larger than most resistors used in circuitry nowadays.

Taking the earlier example, Watts = Volts x Amps, so Watts = 9 x 0.027 which gives 0.234 Watts. Again, to avoid decimals, a unit of 1 milliwatt is used, where 1000 milliwatts = 1 Watt. So instead of writing 0.234 Watts, it is common to write it as 234 mW.

This method of working out voltages, resistances and wattages applies to any circuit, no matter how awkward they may appear. For example, take the following circuit containing five resistors:



As the current flowing through resistor 'R1' has then to pass through resistor 'R2', they are said to be 'in series' and their resistances are added together when calculating current flows. In the example above, both R1 and R2 are 1K resistors, so together they have a resistance to current flow of 2K (that is, 2,000 ohms).

If two, or more, resistors are connected across each other as shown on the right hand side of the diagram above, they are said to be 'in parallel' and their resistances combine differently. If you want to work out the equation above, for yourself, then choose a voltage across Rt, use Ohm's Law to work out the current through Ra and the current through Rb. Add the currents together (as they are both being drawn from the voltage source) and use Ohm's Law again to work out the value of Rt to confirm that the 1/Rt = 1/Ra + 1/Rb + ... equation is correct. A spreadsheet is included which can do this calculation for you.

In the example above, R4 is 1K5 (1,500 ohms) and R5 is 2K2 (2,200 ohms) so their combined resistance is given by 1/Rt = 1/1500 + 1/2200 or Rt = 892 ohms (using a simple calculator). Apply a common-sense check to this result: If they had been two 1500 ohm resistors then the combined value would have been 750 ohms. If they had been two 2200 ohm resistors then the combined value would have been 1100 ohms. Our answer must therefore lie between 750 and 1100 ohms. If you came up with an answer of, say, 1620 ohms, then you know straight off that it is wrong and the arithmetic needs to be done again.

So, how about the voltages at points 'A' and 'B' in the circuit? As R1 and R2 are equal in value, they will have equal voltage drops across them for any given current. So the voltage at point 'A' will be half the battery voltage, i.e. 6 Volts.

Now, point 'B'. Resistors R4 and R5 act the same as a single resistor of 892 ohms, so we can just imagine two resistors in series: R3 at 470 ohms and R4+R5 at 892 ohms. Common-sense rough check: as R3 is only about half the resistance of R4+R5, it will have about half as much voltage drop across it as the voltage drop across R4+R5, i.e. about 4 Volts across R3 and about 8 Volts across R4+R5, so the voltage at point 'B' should work out

at about 8 Volts.

We can use **Ohm's Law** to calculate the current flowing through point 'B':

Ohms = Volts / Amps, (or **Amps = Volts / Ohms** or **Volts = Ohms x Amps**)

(470 + 892) = 12 / Amps, so

Amps = 12 / (470 + 892)

Amps = 12 / 1362 or

Amps = 0.00881 Amps (8.81 milliamps).

Now that we know the current passing through (R4+R5) we can calculate the exact voltage across them:

Resistance = Volts / Amps so

892 = Volts / 0.00881 or

Volts = 892 x 0.00881

Volts = 7.859 Volts.

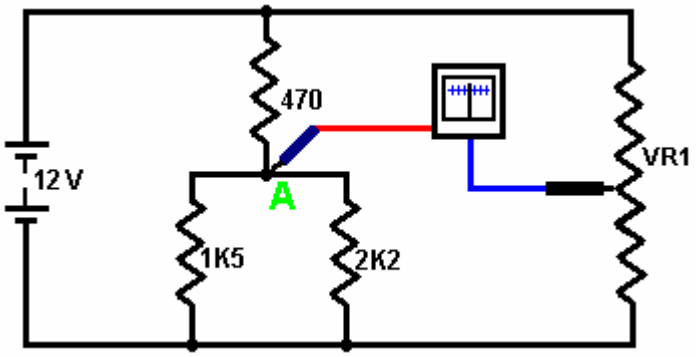
As our common-sense estimate was 8 Volts, we can accept 7.86 Volts as being the accurate voltage at point 'B'.

The Potentiometer.

Just before we leave the subject of resistors and move on to more interesting subjects, we come across the term 'potentiometer'. This term is often shortened to 'pot' and many people use it to describe a variable resistor. I only mention this so that you can understand what they are talking about. A variable resistor is not a potentiometer and really should not be called one. You can skip the rest of this part as it is not at all important, but here is what a potentiometer is:

A fancy name for voltage is 'potential', so a circuit powered by a 12 Volt battery can be described as having a 'potential' of zero volts at the negative side of the battery and a 'potential' of plus twelve volts at the positive side of the battery. Ordinary folks like me would just say 'voltage' instead of 'potential'.

When a voltmeter is used to measure the voltage at any point in a circuit, it alters the circuit by drawing a small amount of current from the circuit. The voltmeter usually has a high internal resistance and so the current is very small, **but** even though it is a small current, it **does** alter the circuit. Consequently, the measurement made is not quite correct. Scientists, in years gone by, overcame the problem with a very neat solution - they measured the voltage without taking **any** current from the circuit - neat huh? They also did it with a very simple arrangement:



They used a sensitive meter to measure the current. This meter is built so that the needle is in a central position if no current is flowing. With a positive current flowing, the needle deflects to the right. With a negative current flowing, the needle moves to the left. They then connected a variable resistor 'VR1' across the same battery which was powering the circuit. The top end of VR1 is at +12 Volts (they called that 'a potential of +12 Volts') and the bottom end of VR1 is at zero volts or 'a potential of zero volts'.

By moving the slider of VR1, any voltage or 'potential' from zero volts to +12 Volts could be selected. To measure the voltage at point 'A' without drawing any current from the circuit, they would connect the meter as shown and adjust the variable resistor until the meter reading was exactly zero.

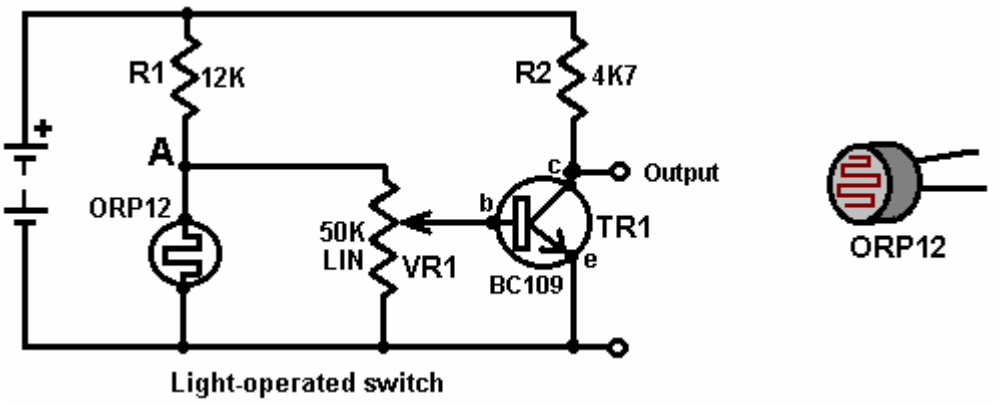
Since the meter reading is zero, the current flowing through it is also zero and the current taken from the circuit is zero. As no current is being taken from the circuit, the measurement is not affecting the circuit in any way - very clever. The voltage on the slider of VR1 exactly matches the voltage at point 'A', so with a calibrated scale on the variable resistor, the voltage can be read off.

The slick piece of equipment made up from the battery, the variable resistor and the meter was used to measure the 'potential' (voltage) at any point and so was called a 'potentiometer'. So, please humour me by calling a variable resistor a 'variable resistor' and not a 'potentiometer'. As I said before, this is not at all important, and if you want to, you can call a variable resistor a 'heffalump' so long as you know how it works.

Semiconductors.

This section deals with discrete semiconductors. A later section deals with 'Integrated Circuits' which are large-scale semiconductor devices.

ORP12 Light-dependent resistor. This device has a high resistance in the dark and a low resistance in bright light. It can be placed in a circuit to create a switch which operates with an increase in light level or a decrease in light level:



Light-operated switch

In this version, the voltage at point 'A' controls the circuit. In darkness, the ORP12 has a resistance ten times greater than that of R1 which is 12,000 ohms. Consequently, the voltage at point 'A' will be high. As the light level increases, the resistance of the ORP12 falls, dragging the voltage at point 'A' downwards. As the variable resistor 'VR1' is connected from point 'A' to the ground rail (the -ve of the battery), its slider can be moved to select any voltage between 0 Volts and the voltage of 'A'. A slider point can be chosen to make the transistor switch off in daylight and on at night. To make the circuit trigger when the light level increases, just swap the positions of R1 and the ORP12.

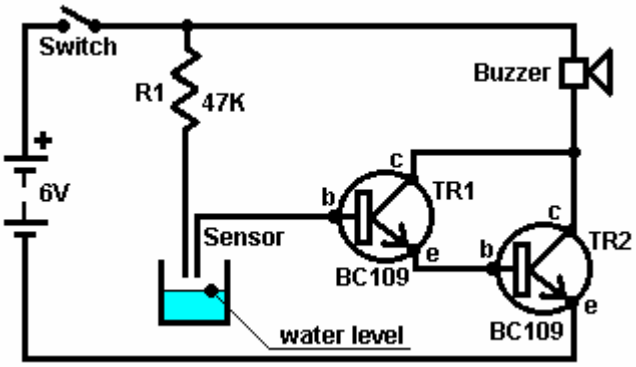
The transistor shown is a BC109 although most transistors will work in this circuit. The BC109 is a cheap, silicon, NPN transistor. It can handle 100mA and 30V and can switch on and off more than a million times per second. It has three connections: the Collector, marked '**c**' in the diagram, the Base, marked '**b**' in the diagram and the Emitter, marked '**e**' in the diagram.

As mentioned before, it has a very high resistance between the collector and the emitter when no current flows into the base. If a small current is fed into the base, the collector/emitter resistance drops to a very low value. The collector current divided by the base current is called the 'gain' of the transistor and is often called 'hfe'. A transistor such as a BC109 or a BC108 has a gain of about 200, though this varies from actual transistor to actual transistor. A gain of 200 means that a current of 200mA passing through the collector requires a current of 1mA through the base to sustain it. Specific information on the characteristics and connections of semiconductors of all kinds can be obtained free from the excellent website www.alldatasheet.co.kr which provides .pdf information files.

The BC109 transistor shown above is an NPN type. This is indicated by the arrow of the symbol pointing outwards. You can also tell by the collector pointing to the positive rail. There are similar silicon transistors constructed as PNP devices. These have the arrow in the transistor symbol pointing inwards and their collectors get connected, directly or indirectly, to the negative rail. This family of transistors are the earliest transistor designs and are called 'bi-polar' transistors.

These silicon transistors are so efficiently constructed that they can be connected directly together to give greatly increased gain. This arrangement is called a 'Darlington pair'. If each transistor has a gain of 200, then the pair give a gain of 200 x 200 = 40,000. This has the effect that a very, very small current can be used to power a load.

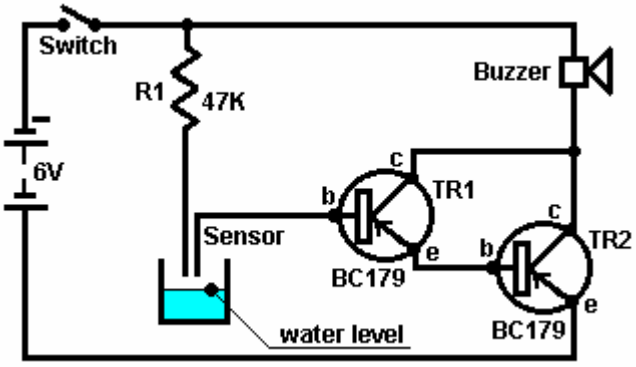
The following diagram shows a Darlington pair used in a water-level detector. This type of alarm could be very useful if you are asleep on a boat which starts taking on water.



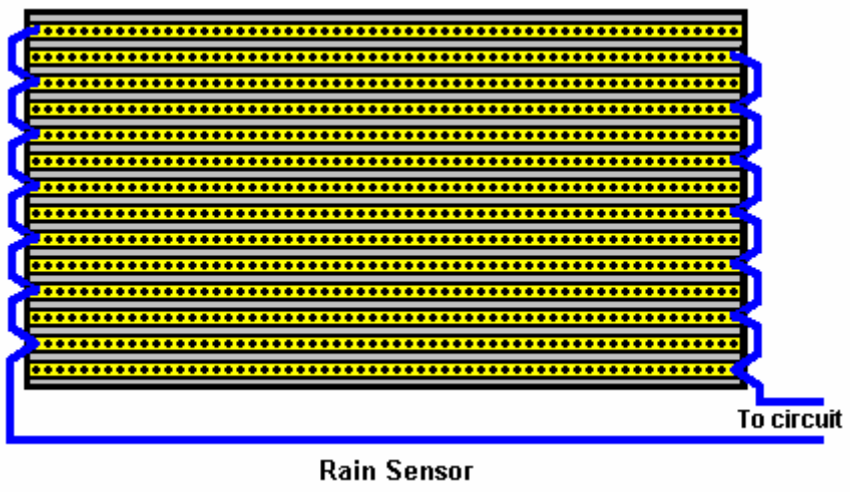
Here, (when the circuit is switched on), transistor TR1 has so little leakage current that TR2 is starved of base current and is hard off, giving it a high resistance across its collector/emitter junction. This starves the buzzer of voltage and keeps it powered off. The sensor is just two probes fixed in place above the acceptable water level. If the water level rises, the probes get connected via the water. Pure water has a high electrical resistance but this circuit will still work with pure water.

The odds are that in a practical situation, the water will not be particularly clean. The resistor R1 is included to limit the base current of TR1 should the sensor probes be short-circuited. Silicon bi-polar transistors have a base/emitter voltage of about 0.7V when fully switched on. The Darlington pair will have about 1.4V between the base of TR1 and the emitter of TR2, so if the sensor probes are short-circuited together, resistor R1 will have 6 - 1.4 = 4.6V across it. Ohms Law gives us the current through it as R = V / A or 47,000 = 4.6 / A or A = 4.6 / 47,000 amps. This works out at 0.098mA which with a transistor gain of 40,000 would allow up to 3.9A through the buzzer. As the buzzer takes only 30mA or so, it limits the current passing through it, and TR2 can be considered to be switched hard on with the whole battery voltage across it.

NPN transistors are more common than PNP types but there is almost no practical difference between them. Here is the previous circuit using PNP transistors:



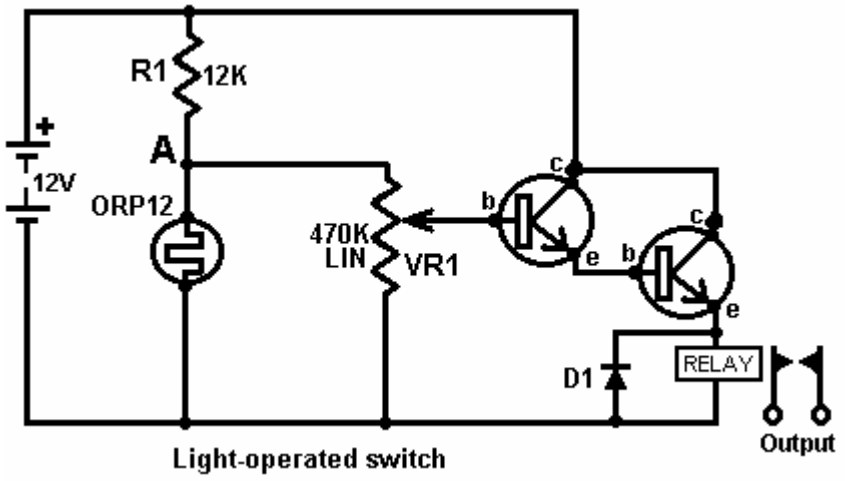
Not a lot of difference. Most of the circuit diagrams shown here use NPN types but not only are these not critical, but there are several ways to design any particular circuit. In general, the semiconductors shown in any circuit are seldom critical. If you can determine the characteristics of any semiconductor shown, any reasonably similar device can generally be substituted, especially if you have a general understanding of how the circuit works. Either of the two previous circuits can operate as a rain detector. A suitable sensor can easily be made from a piece of strip board with alternate strips connected together to form an interlacing grid:



Rain Sensor

Here, if a raindrop bridges between any two adjacent strips, the circuit will trigger and sound a warning.

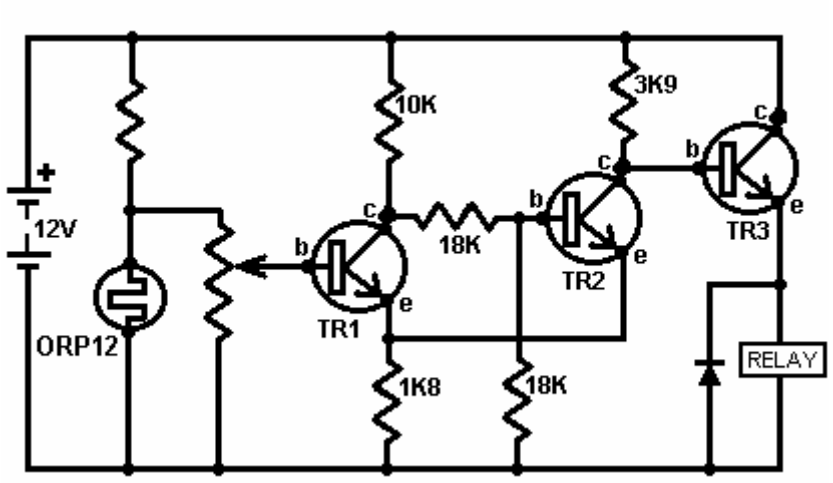
The transistors in the circuit above are connected with their emitter(s) connected to the ground rail (the lower battery line shown in any circuit is considered to be “ground” unless it is specifically shown elsewhere). This connection method is called ‘common emitter’. The following circuit uses the transistor connected in ‘emitter follower’ mode. This is where the emitter is left to *follow* the base voltage - it is always 0.7V below it unless the base itself is driven below 0.7V:



Light-operated switch

This is almost the same as the light-operated circuit shown earlier. In this variation, the transistors are wired so that they work as an ‘emitter-follower’ which follows the voltage at point ‘A’ which rises as the light level drops and the resistance of the ORP12 increases. This causes the voltage across the relay to increase until the relay operates and closes its contacts. A relay is a voltage-operated mechanical switch which will be described in more detail later on.

The disadvantage of the above circuit is that as the light level decreases, the current through the relay increases and it may be a significant amount of current for some considerable time. If it was intended to power the unit with a battery then the battery life would be far shorter than it need be. What we would like, is a circuit which switched rapidly from the Off state to the On state even though the triggering input varied only slowly. There are several ways to achieve this, one of them being to modify the circuit to become a ‘Schmitt Trigger’:



Here, an additional transistor ('TR2') has changed the circuit operation significantly, with transistor TR3 switching fully on and fully off, rapidly. This results in the current through the relay being very low until the circuit triggers.

The circuit operates as follows. When the voltage at the base of TR1 is high enough, TR1 switches on, which causes the resistance between its collector and emitter to be so low that we can treat it as a short circuit (which is a nearly-zero resistance connection). This effectively connects the 10K and 1K8 resistors in series across the battery. The voltage at their connecting point (both the collector and emitter of TR1) will then be about 1.8 Volts. The two 18K resistors are in series across that voltage so the voltage at their junction will be half that; 0.9 Volts.

This puts the Base of TR2 at about 0.9 Volts and its emitter at 1.8 Volts. The base of TR2 is therefore not 0.7 Volts above its emitter, so no base/emitter current will flow in TR2, which means that TR2 is switched hard off. This means that the TR2 collector/emitter resistance will be very high. The voltage at the base of TR3 is controlled by the 1K8 resistor, the TR2 collector/emitter resistance (very high) and the 3K9 resistor. This pushes the base voltage of TR3 up to near the full battery voltage and as it is wired as an emitter-follower, its emitter voltage will be about 0.7 Volts below that. This means that the relay will have most of the battery voltage across it and so will switch hard on.

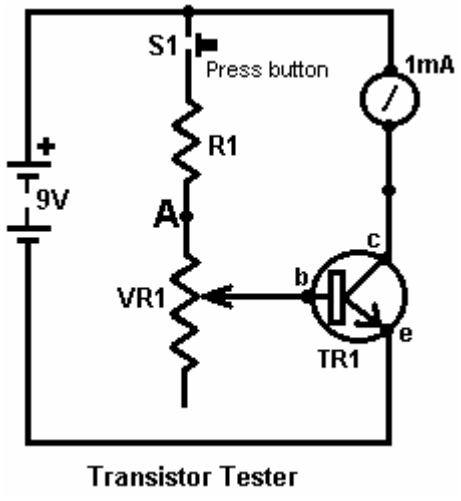
Some practical points: The current flowing into the base of TR3 comes via the 3K9 resistor. A 3K9 resistor needs 3.9 Volts across it for every 1 mA which flows through it. If the relay needs 150 mA to operate and TR3 has a gain of 300, then TR3 will need a base current of 0.5 mA to provide 150 mA of current through its collector/emitter junction. If 0.5 mA flows through the 3K9 resistor, there will be a voltage drop across it of some 2 Volts. The TR3 base/emitter voltage will be a further 0.7 Volts, so the voltage across the relay will be about 12.0 - 2.0 - 0.7 = 9.3 Volts, so you need to be sure that the relay will work reliably at 9 Volts.

If you used a Darlington pair of transistors, each with a gain of 300, instead of TR3, then their combined base/emitter voltage drop would be 1.4 Volts, but they would only need a base current of 150 mA / (300 x 300) = 1/600 mA. That current would only drop 0.007 Volts across the 3K9 resistor, so the relay would receive 10.6 Volts.

So, how do you work out the gain of any particular transistor? The main working tool for electronics is a multimeter. This is a digital or analogue meter which can measure a wide range of things: voltage, current, resistance, ... The more expensive the meter, generally, the greater the number of ranges provided. The more expensive meters offer transistor testing. Personally, I prefer the older, passive multimeters. These are looked down on because they draw current from the circuit to which they are attached, but, because they do, they give reliable readings all the time. The more modern battery-operated digital multimeters will happily give incorrect readings as their battery runs down. I wasted two whole days, testing rechargeable batteries which appeared to be giving impossible performances. Eventually, I discovered that it was a failing multimeter battery which was causing false multimeter readings.

Transistor Testers.

For the moment, let us assume that no commercial transistor tester is to hand and we will build our own (or at least, discover how to build our own). The gain of a transistor is defined as the collector/emitter current divided by the base/emitter current. For example, if 1mA is flowing through the collector and 0.01mA is flowing into the base to sustain that collector flow, then the transistor has a gain of 100 times at 1mA. The transistor gain may vary when it is carrying different current loads. For the circuits we have been looking at so far, 1mA is a reasonable current at which to measure the transistor gain. So let's build a circuit to measure the gain:



Transistor Tester

With the circuit shown here, the variable resistor is adjusted until a collector current of 1mA is shown on the milliammeter and the gain of the transistor is then read off the scale on the variable resistor knob. The circuit is built into a small box containing the battery and with a socket into which the transistor can be plugged. The question then is, what values should be chosen for the resistor R1 and the variable resistor VR1?

Well, we might choose that the minimum gain to be displayed is 10. This would correspond to where the variable resistor slider is taken all the way up to point 'A' in the circuit diagram, effectively taking the variable resistor out of the circuit. If the transistor gain is 10 and the collector current is 1mA, then the base current will be 0.1mA. This current has to flow through the resistor R1 and it has a voltage of (9.0 - 0.7) Volts across it as the base/emitter voltage is 0.7 Volts when the transistor is on. Ohms Law gives us Ohms = Volts / Amps, which for the resistor R1 means Ohms = 8.3 / 0.0001 or 83,000 ohms, or 83K.

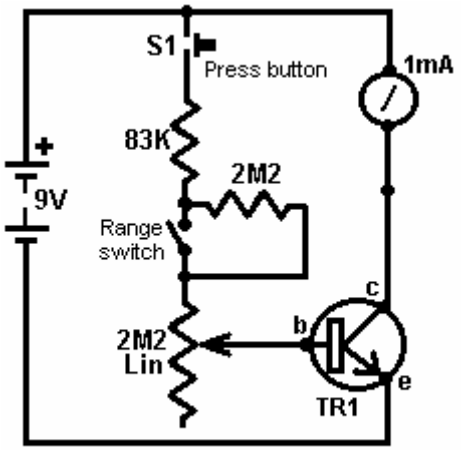
Rule of thumb: 1K provides 1mA if it has 1V across it, so 10K will give 0.1mA if it has 1 Volt across it. With 8.3 Volts across it, it needs to be 8.3 times larger to hold the current down to the required 0.1mA so the resistor should be 83K in size.

As 83K is not a standard size, we need to use two or more standard resistors to give that resistance. Nearest standard size below 83K is 82K, so we can used one 82K resistor and one 1K resistor in series to give the required 83K.

Suppose that we say that we would like to have 500 as the highest gain shown on our tester, then when VR1 is at its maximum value, it and R1 should provide 1/500 of the collector current of 1mA, i.e. 0.002mA or 0.000002 Amps. From Ohms Law again we get VR1 + R1 = 4,150,000 ohms or 4M15. Unfortunately, the largest value variable resistor available is 2M2 so the circuit as it stands, will not be able to cope.

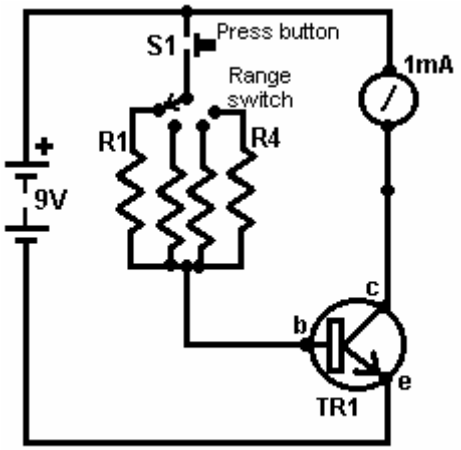
Suppose we were to just use a 2M2 variable resistor for VR1, what transistor gain range could we display? Well Ohms Law ... lets us calculate the base current with 8.3 Volts across (83,000 + 2,200,000) ohms and from that the maximum transistor gain which would be 277.77 (at 1mA). You would buy a 'linear' standard carbon track variable resistor so that the change in resistance is steady as the shaft is rotated. The scale which you would make up would be in even steps and it would run from 10 at the minimum setting, to 278 at the highest setting.

But that is not what we wanted. We wanted to measure up to 500. But they don't make variable resistors big enough, so what can we do? Well, if we wanted, we could lower the battery voltage, which in turn would lower the resistor values. As a 9V battery is very convenient for this kind of circuit, lets not go down that route. We could add extra circuitry to drop the 9V battery voltage down to a lower value. The most simple solution is to add an extra resistor and switch to give two ranges. If we switched in an extra 2M2 resistor above VR1 then the circuit would measure transistor gains from 278 to just over 500 and all we would need to do would be to add a second scale for the VR1 pointer knob to move over. We could, provide extra ranges which overlap and which have more convenient scales to mark. The design is up to you.



Transistor Tester 2

The design covered above is not the only way to measure the transistor gain. A second way, which accepts that it is not so accurate, picks a set base current and measures the collector current as a guide to the gain. In this simple method, one or more resistor values are chosen to give gain ranges, and the milliammeter used to read the corresponding gain:



Transistor Tester 3

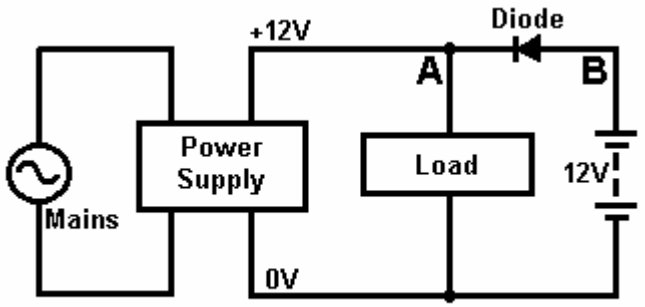
Here, resistor R1 might be chosen to give a collector current of 1mA (which is a full-scale deflection on the meter) when the transistor gain is 100. Resistor R2 might be picked to give a full-scale deflection for a gain of 200, R3 for a gain of 400, R4 for a gain of 600, and so on. Generally speaking, it is not essential to know the exact gain but any reasonable approximation to it is sufficient. You are normally selecting a transistor where you need a gain of 180, so it is not important if the transistor you pick has a gain of 210 or 215 - you are only avoiding transistors with gains below 180.

How do you work out the values of the resistors R1 to R4? Well, you probably won't expect this, but you use Ohms Law. Voltage drop is 8.3 Volts and the base current is given by the full-scale deflection's 1mA divided by the transistor gain for each range, i.e. 1/100 mA for R1, 1/200 mA for R2,... 1/600 mA for R4,...

The Diode.

One component which has been shown but not described is the diode or 'rectifier'. This is a device which has a very high resistance to current flowing in one direction and a very low resistance to current flowing in the opposite direction. The base/emitter junction of a transistor is effectively a diode and, at a push, can be used as such. A proper diode is cheap to buy and has far greater voltage and current handling capacities than the base/emitter junction of a transistor.

Diodes are mainly made from one of two materials: germanium and silicon. Germanium diodes are used with very small alternating currents such as radio signals coming from an aerial. This is because a germanium diode needs only 0.2 Volts or so to carry a current while silicon needs 0.6 to 0.7 Volts (same as a silicon transistor base/emitter junction). Germanium diodes (and transistors) are very sensitive to temperature change and so are normally restricted to low power circuits. One very neat application for a silicon diode is as an 'un-interruptible power supply' where mains failure is caught instantly:



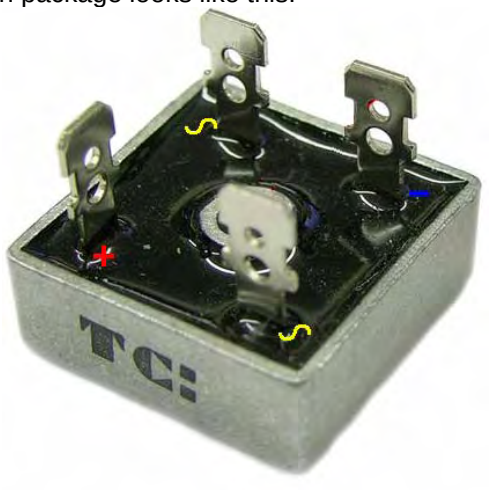
Mains Unit Back-up Circuit

In this circuit, the mains voltage drives the Power Supply Unit which generates 12 Volts at point 'A'. This provides current to the Load. The diode has +12 Volts at 'A' and +12 Volts at point 'B' so there is no voltage drop across it and it will not carry current in either direction. This means that the battery is effectively isolated when the mains is functioning. If the Power Supply Unit output were to rise above its design level of +12 Volts, then the diode would block it from feeding current into the battery.

If the mains fails, the Power Supply Unit ('PSU') output will fall to zero. If the battery and diode were not there, the voltage at point 'A' would fall to zero, which would power-down the Load and possibly cause serious problems. For example, if the load were your computer, a mains failure could cause you to lose important data. With a battery back-up of this type, you would have time to save your data and shut your computer down before the battery ran out.

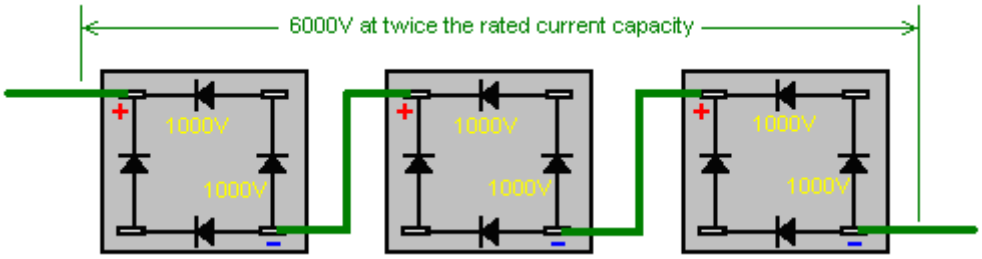
The circuit operates in a very simple fashion. As soon as the voltage at point 'A' drops to 0.7 Volts below the +12 Volts at point 'B", the diode starts feeding current from the battery to the Load. This happens in less than a millionth of a second, so the Load does not lose current. It would be worth adding a warning light and/or a buzzer to show that the mains has failed.

Diodes are also supplied packaged as a diode bridge, with four diodes enclosed inside. Usually intended for power supply rectification, they are not particularly fast-acting diodes, but are cheap and can carry a good deal of current. A common size is with the diodes rated at 1000 volts and able to carry 35 amps. Although there are many package types, a very common package looks like this:



The alternating signal is connected between two opposite corners and the pulsating DC is taken off from the other two terminals. The symbols shown above are normally marked on the flat face which is not seen in this picture. The package has a hole in the centre so that the metal case can be bolted to a heat-sink in order to keep the device reasonably cool when carrying large currents.

It is possible to connect the bridge in a different way and use it as a higher voltage double diode arrangement as shown here:



By skipping the alternating current ability and connecting to just the Plus and the Minus terminals, the package provides two pairs if diodes in connected in series. This gives twice the voltage handling in both current paths and the rated current handling capacity in both of those two paths which are now connected across each other, which doubles the current handling capacity. The diagram shows how three ordinary, cheap 1000V 35 amp bridges can be connected to give one 70 amp 6000V composite diode.

Diodes are specified by their voltage handling capacity and their current-carrying capacity and the speed at which they can switch on and off. For power supplies where the frequency is very low, any diode will do, but there are circuits where the switching is needed hundreds of thousand times per second and so the diode specification sheets need to be checked to see what frequency can be handled by any particular diode. Those data sheets can be downloaded free from http://www.alldatasheet.co.kr/.

One other thing which needs to be checked for some circuits is the voltage needed to get the diode to switch on. Two common materials used when making diodes are silicon and germanium. Germanium types have a low forward voltage of around 0.2 volts typically which silicon has about a 0.6 volt threshold generally. These voltage figures vary enormously as the current through the diode increases. Circuits which use very low voltages need germanium diodes such as the 1N34.

Light-Emitting Diodes.

There is a widely used variation of the diode which is extremely useful, and that is the Light Emitting Diode or 'LED'. This is a diode which emits light when carrying current. They are available in red, green, blue, yellow or white light versions. Some versions can display more than one colour of light if current is fed through their different electrical connections.

LEDs give a low light level at a current of about 8 or 10 mA and a bright light for currents of 20 to 30 mA. If they are being used with a 12 Volt system, then a series resistor of 1K to 330 ohms is necessary. LEDs are robust devices, immune to shock and vibration. They come in various diameters and the larger sizes are very much more visible than the tiny ones.

Thyristors ("SCR"s) and Triacs.

Another version of the diode is the Silicon Controlled Rectifier or 'Thyristor'. This device carries no current until its gate receives an input current. This is just like the operation of a transistor but the SCR once switched on, stays on even though the gate signal is removed. It stays on until the current through the SCR is forced to zero, usually by the voltage across it being removed. SCRs are often used with alternating voltages (described below) and this causes the SCR to switch off if the gate input is removed. SCRs only operate on positive voltages so they miss half of the power available from alternating power supplies. A more advanced version of the SCR is the 'Triac' which operates in the same way as an SCR but handles both positive and negative voltages.

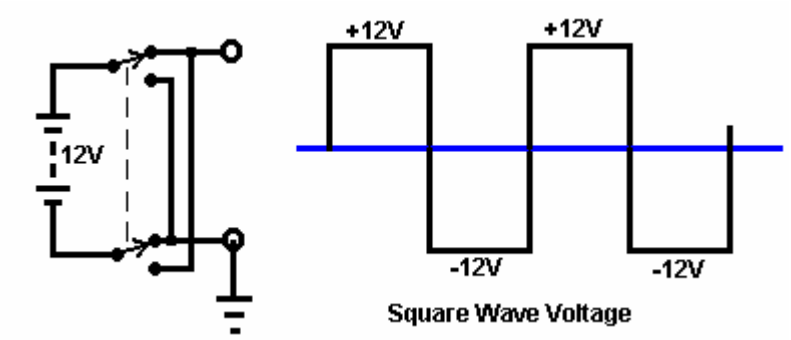
Opto-Isolators.

Another very useful variation on the LED is the Opto-Isolator. This device is a fully enclosed LED and light-sensitive transistor. When the LED is powered up, it switches the transistor on. The big advantage of this device is that the LED can be in a low voltage, low power sensing circuit, while the transistor can be in a completely separate, high voltage, high power circuit. The opto-isolator isolates the two circuits completely from each other. It is a very useful, and very popular, low-cost device.

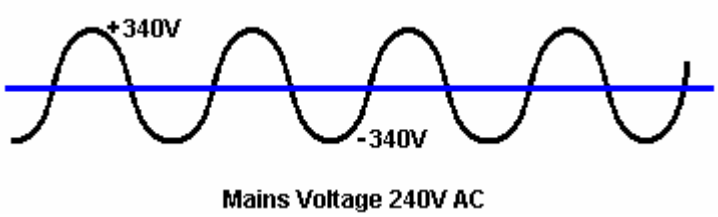
Alternating Current.

A battery provides a constant voltage. This is called a Direct Current or 'DC' source of power. When a circuit is connected to a battery, the positive rail is always positive and the negative rail is always negative.

If you connect a battery to a circuit through a double-pole changeover switch as shown here:



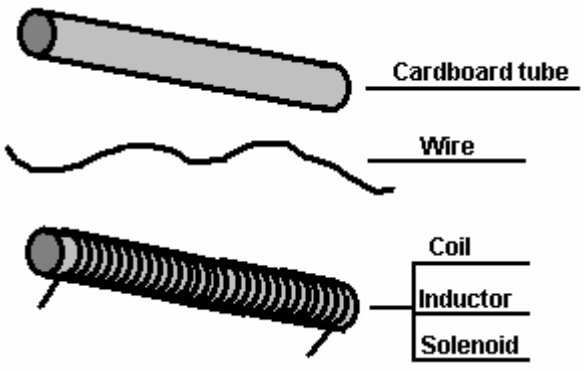
When the changeover switch is operated, the battery is effectively turned over or inverted. This circuit is called an 'inverter' because it repeatedly inverts the supply voltage. If the switch is operated on a regular, rapid basis, the graph of the output voltage is as shown on the right. This is a 'square wave' voltage and is used extensively in electronic equipment. It is called alternating current or 'AC' for short. SCRs and Triacs can be used conveniently with supply voltages of this type. Mains voltage is also AC but is rather different:



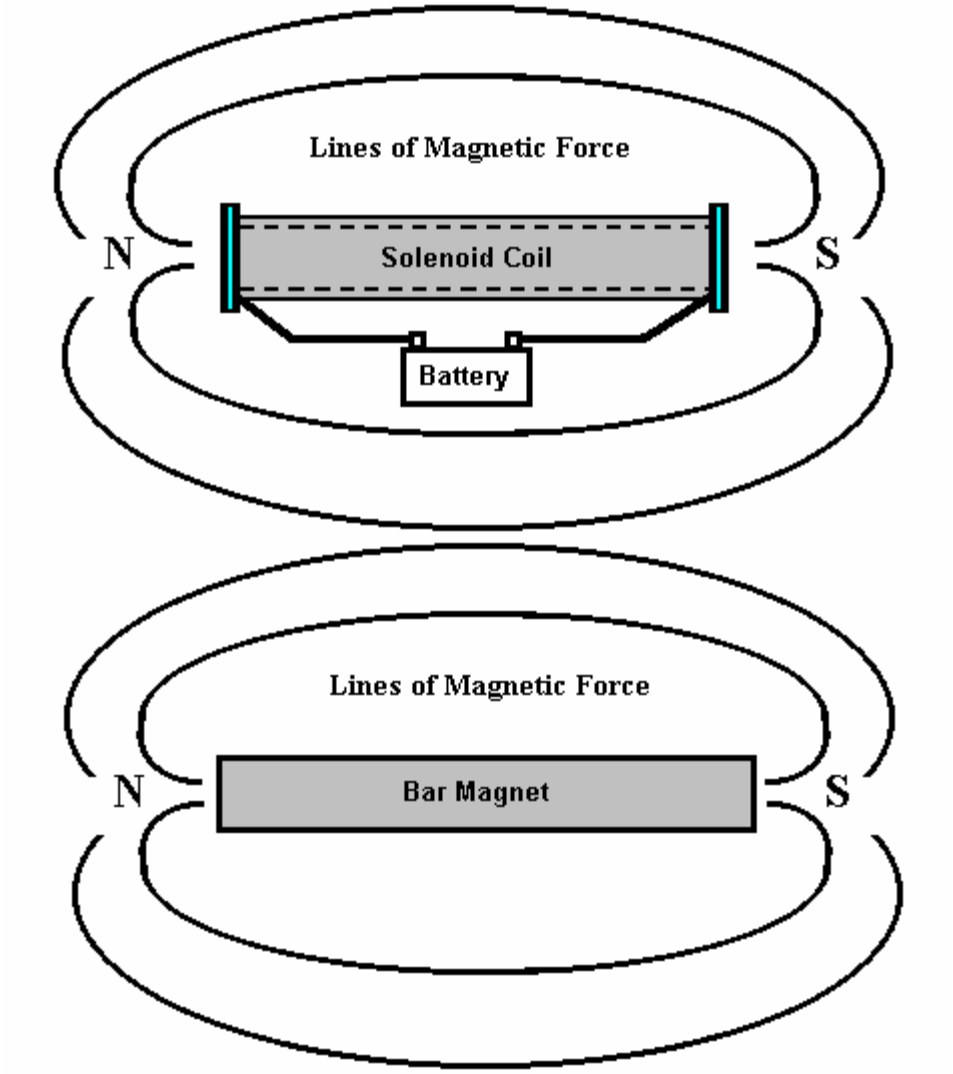
Mains voltage varies continuously in the form of a sine wave. In Britain, the mains voltage is described as '240 Volts AC' and it cycles up and down 50 times per second, i.e. 50 positive peaks and 50 negative peaks in one second. It would be reasonable to assume that each voltage peak would be 240 Volts but this is not the case. Even though the supply is described as 240 Volts, it peaks at the square root of 2 times greater than that, i.e. 339.4 Volts. The actual supply voltage is not particularly accurate, so any device intended for mains use should be rated to 360 Volts. In America, the supply voltage is 110 Volts AC and it cycles 60 times per second, peaking at plus and minus 155 Volts. Later on, you will see how one or more diodes can be used to convert AC to DC in a unit which is sold as a 'mains adapter' intended to allow battery operated equipment be operated from the local mains supply.

Coils ("Inductors") and Solenoids.

If you take a cardboard tube, any size, any length, and wind a length of wire around it, you create a very interesting device. It goes by the name of a 'coil' or an 'inductor' or a 'solenoid'.

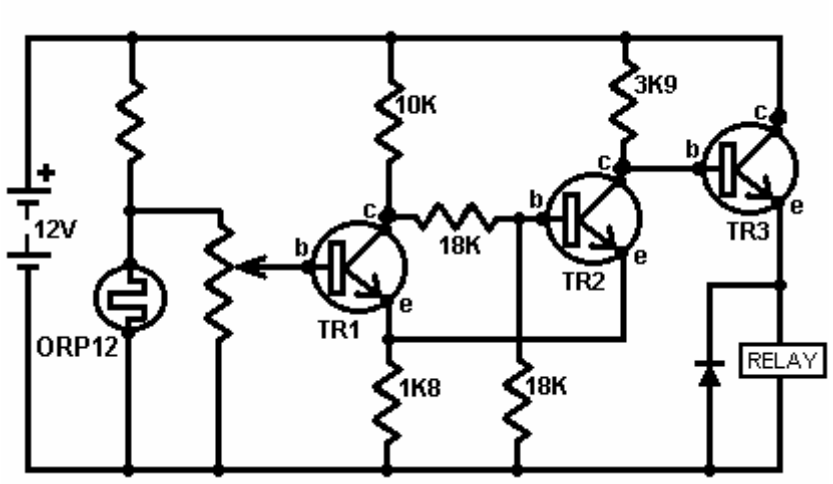


This is a very interesting device with many uses. It forms the heart of a radio receiver, it used to be the main component of telephone exchanges, and most electric motors use several of them. The reason for this is if a current is passed through the wire, the coil acts in exactly the same way as a bar magnet:



The main difference being that when the current is interrupted, the coil stops acting like a magnet, and that can be very useful indeed. If an iron rod is placed inside the coil and the current switched on, the rod gets pushed to one side. Many doorbells use this mechanism to produce a two-note chime. A 'relay' uses this method to close an electrical switch and many circuits use this to switch heavy loads (a thyristor can also be used for this and it has no moving parts).

A coil of wire has one of the most peculiar features of almost any electronic component. When the current through it is altered in any way, the coil opposes the change. Remember the circuit for a light-operated switch using a relay?:



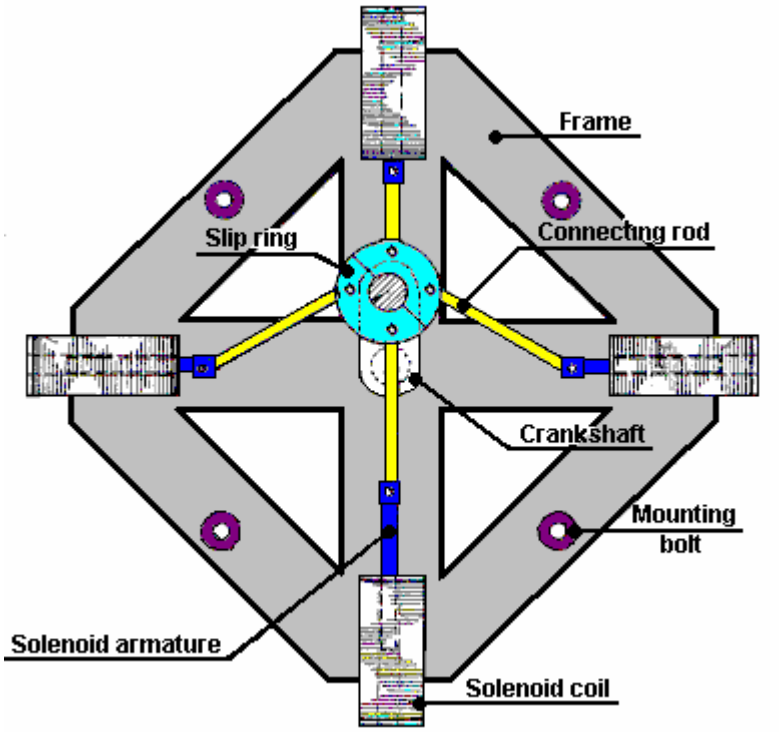
You will notice that the relay (which is mainly a coil of wire), has a diode across it. Neither the relay nor the diode were mentioned in any great detail at that time as they were not that relevant to the circuit being described. The

diode is connected so that no current flows through it from the battery positive to the 'ground' line (the battery negative). On the surface, it looks as if it has no use in this circuit. In fact, it is a very important component which protects transistor TR3 from damage.

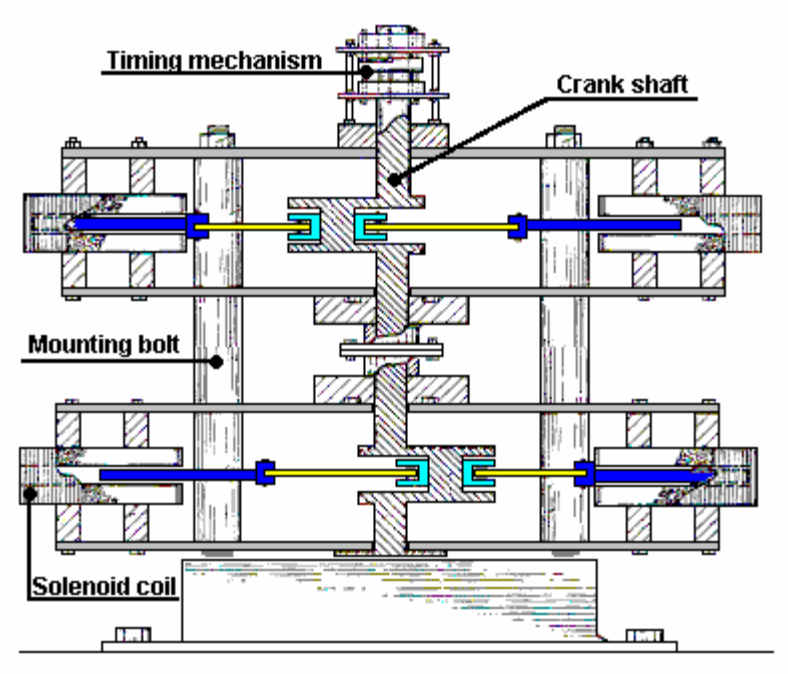
The relay coil carries current when transistor TR3 is on. The emitter of transistor TR3 is up at about +10 Volts. When TR3 switches off, it does so rapidly, pushing the relay connection from +10 Volts to 0 Volts. The relay coil reacts in a most peculiar way when this happens, and instead of the current through the relay coil just stopping, the voltage on the end of the coil connected to the emitter of TR3 keeps moving downwards. If there is no diode across the relay, the emitter voltage is forced to briefly overshoot the negative line of the circuit and gets dragged down many volts below the battery negative line. The collector of TR3 is wired to +12 Volts, so if the emitter gets dragged down to, say, -30 Volts, TR3 gets 42 Volts placed across it. If the transistor can only handle, say, 30 Volts, then it will be damaged by the 42 Volt peak.

The way in which coils operate is weird. But, knowing what is going to happen at the moment of switch-off, we deal with it by putting a diode across the coil of the relay. At switch-on, and when the relay is powered, the diode has no effect, displaying a very high resistance to current flow. At switch-off, when the relay voltage starts to plummet below the battery line, the diode effectively gets turned over into its conducting mode. When the voltage reaches 0.7 Volts below the battery negative line, the diode starts conducting and pins the voltage to that level until the voltage spike generated by the relay coil has dissipated. The more the coil tries to drag the voltage down, the harder the diode conducts, stifling the downward plunge. This restricts the voltage across transistor TR3 to 0.7 Volts more than the battery voltage and so protects it.

Solenoid coils can be very useful. Here is a design for a powerful electric motor patented by the American, Ben Teal, in June 1978 (US patent number 4,093,880). This is a very simple design which you can build for yourself if you want. Ben's original motor was built from wood and almost any convenient material can be used. This is the top view:



And this is the side view:



Ben has used eight solenoids to imitate the way that a car engine works. There is a crankshaft and connecting rods, as in any car engine. The connecting rods are connected to a slip-ring on the crankshaft and the solenoids are given a pulse of current at the appropriate moment to pull the crankshaft round. The crankshaft receives four pulls on every revolution. In the arrangement shown here, two solenoids pull at the same moment.

In the side view above, each layer has four solenoids and you can extend the crankshaft to have as many layers of four solenoids as you wish. The engine power increases with every layer added. Two layers should be quite adequate as it is a powerful motor with just two layers.

An interesting point is that as a solenoid pulse is terminated, its pull is briefly changed to a push due to the weird nature of coils. If the timing of the pulses is just right on this motor, that brief push can be used to increase the power of the motor instead of opposing the motor rotation. This feature is also used in the Adams motor described in the 'Free-Energy' section of this document.

The strength of the magnetic field produced by the solenoid is affected by the number of turns in the coil, the current flowing through the coil and the nature of what is inside the coil 'former' (the tube on which the coil is wound). In passing, there are several fancy ways of winding coils which can also have an effect, but here we will only talk about coils where the turns are wound side by side at right angles to the former.

1. Every turn wound on the coil, increases the magnetic field. The thicker the wire used, the greater the current which will flow in the coil for any voltage placed across the coil. Unfortunately, the thicker the wire, the more space each turn takes up, so the choice of wire is somewhat of a compromise.

2. The power supplied to the coil depends on the voltage placed across it. Watts = Volts x Amps so the greater the Volts, the greater the power supplied. But we also know from Ohm's Law that Ohms = Volts / Amps which can also be written as Ohms x Amps = Volts. The Ohms in this instance is fixed by the wire chosen and the number of turns, so if we double the Voltage then we double the current.

For example: Suppose the coil resistance is 1 ohm, the Voltage 1 Volt and the Current 1 Amp. Then the power in Watts is Volts x Amps or 1 x 1 which is 1 Watt.

Now, double the voltage to 2 Volts. The coil resistance is still 1 ohm so the Current is now 2 Amps. The power in Watts is Volts x Amps or 2 x 2 which is **4** Watts. Doubling the voltage has quadrupled the power.

If the voltage is increased to 3 Volts. The coil resistance is still 1 ohm so the Current is now 3 Amps. The power in Watts is Volts x Amps or 3 x 3 which is **9** Watts. The power is Ohms x Amps squared, or Watts = Ohms x Amps x Amps. From this we see that the voltage applied to any coil or solenoid is critical to the power developed by the coil.

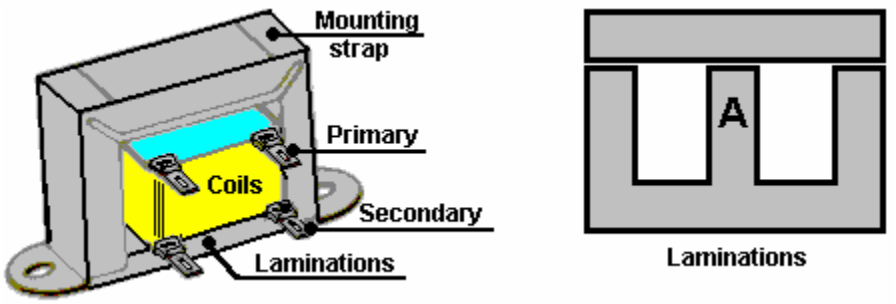
3. What the coil is wound on is also of considerable importance. If the coil is wound on a rod of soft iron covered with a layer of paper, then the magnetic effect is increased dramatically. If the rod ends are tapered like a flat

screwdriver or filed down to a sharp point, then the magnetic lines of force cluster together when they leave the iron and the magnetic effect is increased further.

If the soft iron core is solid, some energy is lost by currents flowing round in the iron. These currents can be minimised by using thin slivers of metal (called 'laminations') which are insulated from each other. You see this most often in the construction of transformers, where you have two coils wound on a single core. As it is convenient for mass production, transformers are usually wound as two separate coils which are then placed on a figure-of-eight laminated core.

Transformers.

Transformers are used to alter the voltage of any alternating current power source. If the alteration increases the output voltage, then the transformer is called a 'step-up' transformer. If the output voltage is lower than the input voltage then it is called a 'step-down' transformer. If the voltages are the same, it is called an 'isolation' transformer. A common construction looks like this:



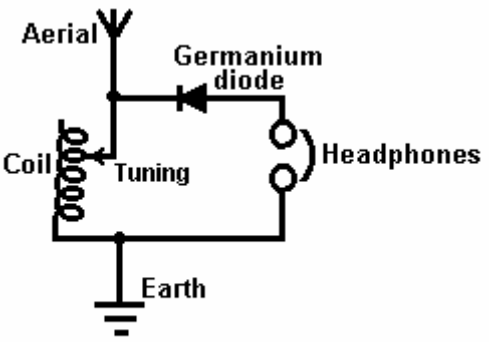
The Coil bobbin sits on the section of the laminations marked 'A' above. The coil is wound on its bobbin former, first one winding and then the second winding. The bobbin is then placed on the central part of the 'E' shaped laminations and then completely surrounded by the laminations when the crossbar is placed on the top. The mounting strap is used to hold the two sets of laminations together and provide mounting lugs for attaching the transformer to a chassis. There are typically, twenty laminations in each set and every lamination is insulated from the adjoining laminations.

If you want to change the voltage of a battery supply, it is possible to build an electronic circuit to generate an alternating voltage and then use a transformer to change that alternating voltage to whatever voltage you want. The most common form of this, is for generating mains voltage from a 12 Volt car battery, so that mains equipment can be run in remote locations, such as boats, caravans, etc. These circuits are called 'inverters' and they are very popular pieces of equipment. The voltage in the secondary coil of any transformer is determined by the ratio of the turns in the primary and secondary windings.

For example; if there is a 10 Volt alternating voltage available and you have a transformer which has 100 turns in the primary coil and 1000 turns in the secondary coil. If you connect the 10 Volts across the primary, there will be 100 Volts generated across the secondary coil.

Instead, if you connect the 10 Volts across the secondary coil, a voltage of 1 Volts will be generated across the primary winding. This is because there is a 10:1 ratio between the two windings. The Law of Conservation of Energy applies to transformers as it does to everything else. The power input to the primary winding will be the same as the power in the secondary winding minus the losses. The losses, in this case, will be a temperature rise of the whole transformer. If the current passed through the transformer is well below its rated capacity, then the losses will be small. The important point is that 10 Volts at 1 Amp into the primary winding will generate 100 Volts in the secondary, but at somewhat less than 0.1 Amps: Power Input is 10 Watts and Power Output is almost 10 Watts. The voltage has been raised to 100 Volts but the potential current draw has been reduced from 1 Amp to 0.1 Amps (100 mA).

In practice, the thickness of the wire used in the windings is very important. If the voltage to be placed across the winding is high, then the wire diameter will be small. Coil windings have fairly low resistances but this is not critical in circuits as coils operate in a peculiar way. Coils have AC 'impedance' in addition to their DC 'resistance'. While Direct Current (from a battery, say) can flow quite easily through a coil with low resistance, Alternating Current may have a hard job getting through the coil due to its high 'impedance'. Sometimes, coils are used to choke off any AC ripple (interference) coming along a DC power cable. When a coil is used for this purpose it is called a 'choke'. Each coil has its own resonant frequency and at that frequency it is very difficult for AC to get through the coil. Crystal set radios work on that principle:



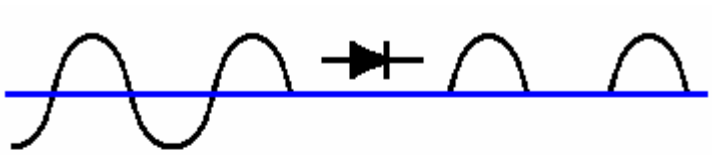
Crystal Set Radio

Here, the aerial picks up every radio station broadcasting in the area. These are all at different frequencies and they all head down the aerial wire, looking for the easiest path to the earth connection. Most of them run through the coil with no problem whatsoever. If the resonant frequency of the coil matches the frequency of one of the radio stations, then that radio signal (and **only** that signal) finds it very hard to get through the coil and looks for an easier path to earth. The next easiest path is through the diode and the headphones, so the signal goes that way. The diode blocks part of the signal which generates the sound of the radio broadcast in the headphones.

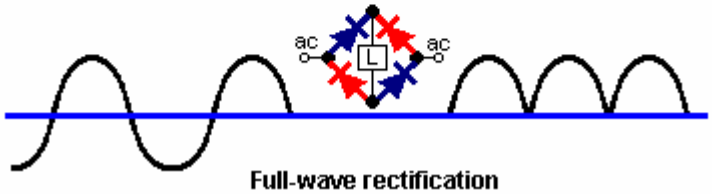
This system works very well indeed if there is a good radio signal. A germanium diode is used as the radio signal voltage is very small and a germanium diode operates on 0.2 Volts while a silicon diode needs 0.7 Volts to operate. That difference is significant at these very low voltages. The resonant frequency of the coil depends on the number of turns in the coil. In this design, the coil has a slider which allows the number of turns to be altered and so, different radio stations to be tuned in.

Rectification and Power Supplies.

We now have the question of how do we turn an alternating voltage into a constant 'direct' voltage. The crystal radio set operates by chopping off half of the alternating radio signal. If we were to do this to the output from a mains transformer with an output of say, 12 Volts AC, the result is not very satisfactory:



Half-wave rectification

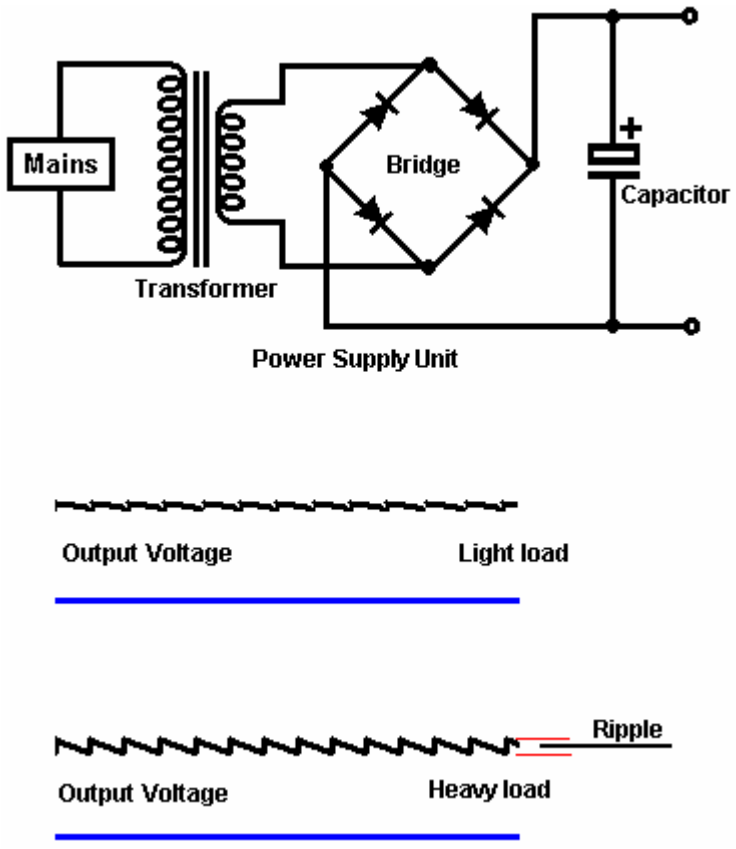


Full-wave rectification

Here, we have the situation shown in the upper diagram. The output consists of isolated pulses at 50 per second. You will notice that there is **no** output power for half of the time. The negative part of the waveform is blocked by the high resistance of the diode while the positive part of the waveform is allowed through by the low resistance of the 'forward-biased' diode. It should be remembered that the diode drops 0.7 Volts when conducting so the output of the half-wave rectified transformer will be 0.7 Volts lower than the transformer's actual output voltage.

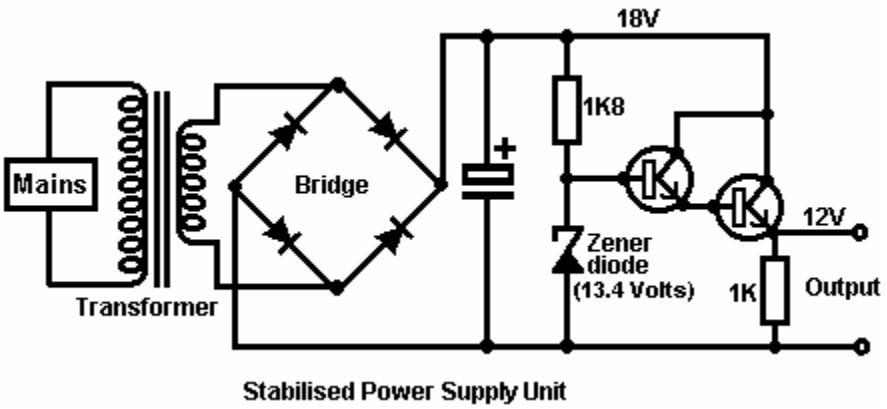
If four diodes are used instead of one, they can be arranged as shown in the lower diagram. This arrangement of diodes is called a 'bridge'. Here the positive part of the waveform flows through the upper blue diode, the load 'L' and on through the lower blue diode. The negative part flows through the left hand red diode, the load and then the right hand red diode. This gives a much better output waveform with twice the power available. The output voltage will be 1.4 Volts less than the transformer output voltage as there are two silicon diodes in the supply chain.

The output from even the full-wave rectifier is still unsatisfactory as there is a voltage drop to zero volts 100 times per second. Only a few devices operate well with a power supply like that, an incandescent bulb as used in a car can use this output, but then, it could use the original AC supply without any rectification. We need to improve the output by using a reservoir device to supply current during those moments when the voltage drops to zero. The device we need is a **Capacitor** which used to be called a 'condenser'. The circuit of a mains unit using a capacitor is shown here:



This produces a much better result as the capacitor stores some of the peak energy and gives it out when the voltage drops. If the load on the unit is light with not very much current taken from it, the output voltage is quite good. However, if the current drain is increased, the output voltage gets dragged down 100 times per second. This voltage variation is called 'ripple' and if the unit is supplying an audio system or a radio, the ripple may well be heard as an annoying hum. The larger the capacitor for any given current draw, the smaller the ripple.

To improve the situation, it is normal to insert an electronic control circuit to oppose the ripple:

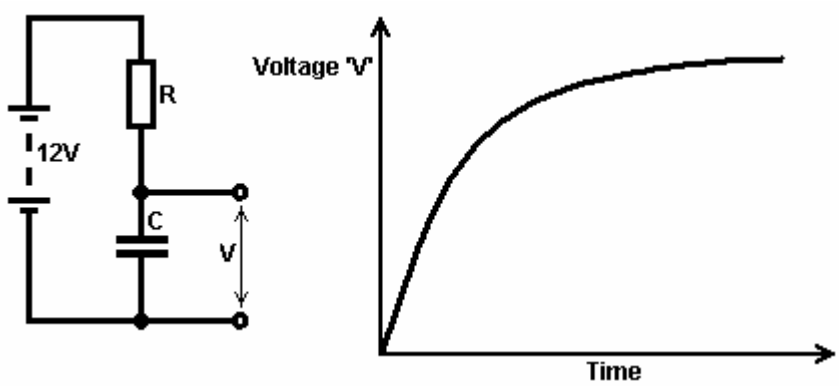


This circuit uses one new component, a new variety of diode called a 'Zener' diode. This device has an almost constant voltage drop across it when its current-blocking direction breaks down. The diode is designed to operate in this state to provide a reference voltage. The circuit merely uses a tiny current from the top of the zener diode to drive the Darlington pair emitter-follower transistors used to provide the output current.

With this circuit, when the output current is increased, the resistance of the transistor pair automatically reduces to provide more current without varying the output voltage. The 1K resistor is included to give the transistors a completed circuit if no external equipment is connected across the output terminals. The zener diode is chosen to give 1.4 Volts more than the required output voltage as the two transistors drop 1.4 Volts when conducting.

You should note that the output transistor is dropping 6 Volts at the full supply current. Watts = Volts x Amps so the power dissipated by the transistor may be quite high. It may well be necessary to mount the transistor on an aluminium plate called a 'heat sink' to keep it from overheating. Some power transistors, such as the 2N3055, do not have the case isolated from the active parts of the transistor. It is good practice to use a mica gasket between the transistor and the heat-sink as it conducts then heat without making an electrical connection to the metal heat-sink.

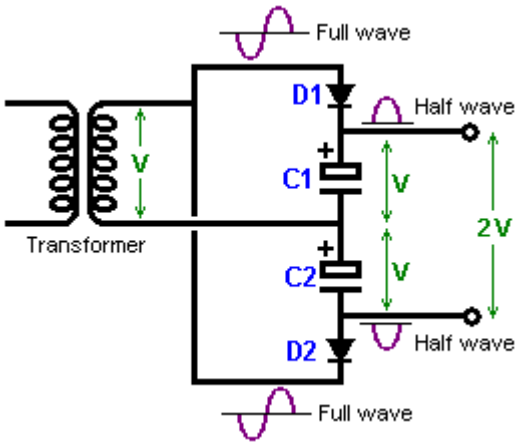
A capacitor, being an electrical reservoir, can be used as part of a timer circuit. If the current flow into it is restricted by passing it through a resistor. The length of time between starting the flow on an empty capacitor, and the voltage across the capacitor reaching some chosen level, will be constant for a high-quality capacitor.



As the voltage increase tails off, it becomes more difficult to measure the difference accurately, so if the capacitor is to be used for generating a time interval, it is normal to use the early part of the graph area where the line is fairly straight and rising fast.

The Voltage Doubler.

It is possible to increase the output voltage of a transformer although this does reduce its ability to supply current at that voltage. The way that this is done is to feed the positive cycles into one storage capacitor and the negative cycles into a second reservoir capacitor. This may sound a little complicated, but in reality, it isn't. A circuit for doing this is shown here:



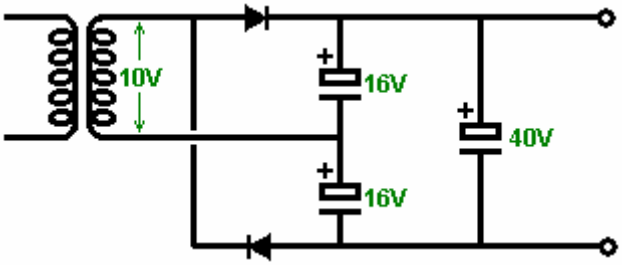
With this circuit, the transformer output is some voltage, say "V" volts of AC current. This output waveform is fed to capacitor "**C1**" through diode "**D1**" which lops off the negative part of the cycle. This produces a series of positive half-cycles which charge up capacitor "**C1**" with a positive voltage of "V".

The other half of the output is fed to capacitor "**C2**" through diode "**D2**" which cuts off the positive part of the cycle, causing capacitor "**C2**" to develop a voltage of -V across it. As the two capacitors are 'in series' and not placed across each other, their voltages add up and produce twice the transformer output voltage.

A word of warning here. The transformer is producing an AC waveform and these are marked with the **average** voltage of the waveform, which is usually a sine wave. The peak voltage of a sinewave is 41% greater than this, so if your transformer has an AC output of 10 volts, then the peaks fed to the capacitors will be about 14.1 volts. If

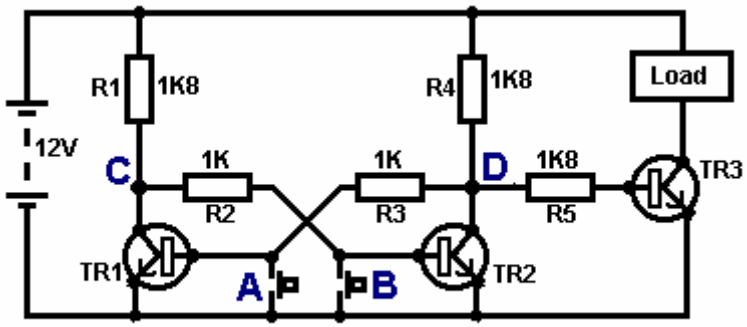
there is no current draw from the capacitors (that is, with the load switched off), then each capacitor will charge to this 14.1 volts and the overall output voltage will be 28.2 volts and **not** the 20 volts which you might expect. You need to understand that as this is only a half-wave supply, there will be considerable ripple on the output voltage if the current draw is high.

Using one additional smoothing capacitor and paying attention to the voltage ratings of the capacitors, the 28 volts supply circuit might be like this:



Multivibrators: The Bistable.

The number of electronic circuits which can be built with basic components such as resistors, capacitors, transistors, coils, etc. is limited only by your imagination and needs. Here is a circuit where two transistors operate as a pair:



Bistable multivibrator

This circuit has two stable states and so it is called a “bi” “stable” or “bistable” circuit. It is important to understand the operation of this simple and useful circuit.

If press-button switch ‘**A**’ is pressed, it short-circuits the base/emitter junction of transistor TR1. This prevents any current flowing in the base/emitter junction and so switches TR1 hard off. This makes the voltage at point ‘**C**’ rise as high as it can. This leaves transistor TR2 powered by R1 and R2 which have 11.3 Volts across them and switches TR2 hard on.

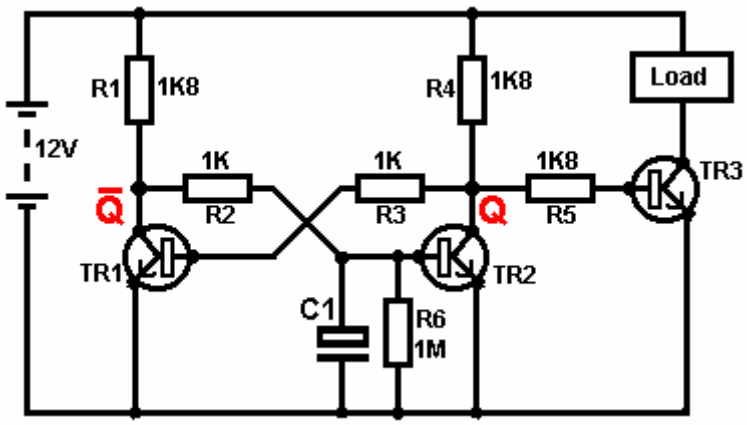
This pulls point ‘**D**’ down to about 0.1 Volts. This happens in less than a millionth of a second. When the press-button switch ‘**A**’ is released, transistor TR1 does not switch on again because its base current flows through resistor R3 which is connected to point ‘**D**’ which is far, far below the 0.7 Volts needed to make TR1 start conducting.

The result is that when press-button ‘**A**’ is pressed, transistor TR2 switches on and stays on even when press-button ‘**A**’ is released. This switches transistor TR3 off and starves the Load of current. This is the first ‘stable state’.

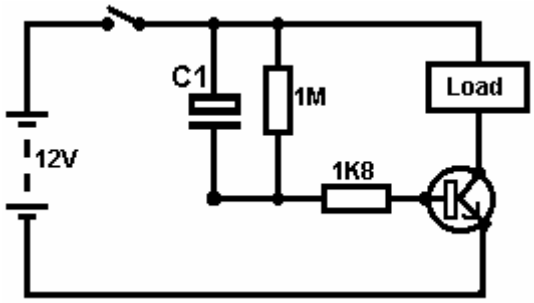
The same thing happens when press-button ‘**B**’ is pressed. This forces transistor TR2 into its ‘off’ state, raising point ‘**D**’ to a high voltage, switching transistor TR3 hard on, powering the Load and holding transistor TR1 hard off. This is the second of the two ‘stable states’.

In effect, this circuit ‘remembers’ which press-button was pressed last, so millions of these circuits are used in computers as Random Access Memory (‘RAM’). The voltage at point ‘**C**’ is the inverse of the voltage at point ‘**D**’, so if ‘**D**’ goes high then ‘**C**’ goes low and if ‘**D**’ goes low, then ‘**C**’ goes high. In passing, the output at ‘**D**’ is often called ‘**Q**’ and the output at ‘**C**’ is called ‘**Q-bar**’ which is shown as the letter Q with a horizontal line drawn above it. This is shown on the next circuit diagram.

A minor variation of this circuit allows a load to be energised when the circuit is powered up:



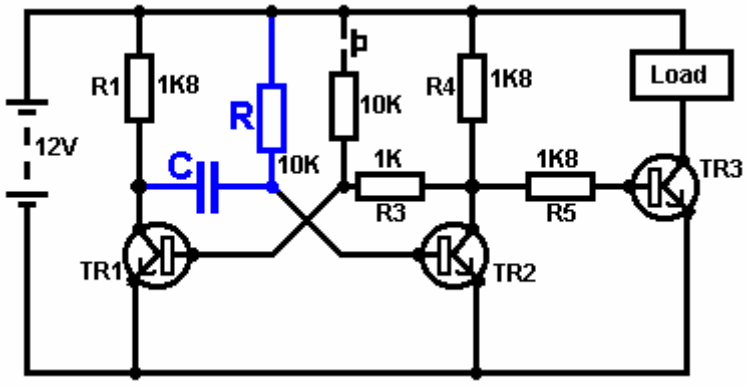
When powered down, the capacitor 'C1' in this circuit is fully discharged through resistor 'R6'. When the 12 Volts supply is connected to the circuit, capacitor C1 does not charge instantly and so holds the base of TR2 down below 0.7 Volts for much longer than it takes for transistor TR1 to switch on (which, in turn, holds TR2 hard off). Mind you, if it is not necessary to have the Load held powered on indefinitely, then an even more simple circuit can do this:



Here, when the switch is closed, both sides of the capacitor C1 are at +12 Volts and this causes the 1K8 resistor to conduct heavily, driving the transistor and powering the load. The capacitor charges rapidly through the transistor and reaches the point at which it can no longer keep the transistor switched on. When the battery is switched off, the 1M resistor discharges the capacitor, ready for the next time the battery is connected.

the Monostable Multivibrator.

The monostable has one stable state and one unstable state. It can be flipped out of its stable state but it will 'flop' back into its stable state. For that reason, it is also known as a 'flip-flop' circuit. It is similar to a bistable circuit, but one of the cross-link resistors has been replaced by a capacitor which can pass current like a resistor, but only for a limited amount of time, after which, the capacitor becomes fully charged and the current flow stops, causing the 'flop' back to the stable state once more.



In this circuit, the '**R**' resistor and the '**C**' capacitor values determine how long the monostable will be in its unstable state. The circuit operates like this:

1. In the stable state, transistor TR1 is off. Its collector voltage is high, pushing the left hand side of capacitor '**C**' to near +12 Volts. As the right hand side of capacitor '**C**' is connected to the base of TR2 which is at 0.7 Volts, the capacitor gets charged to about 11.3 Volts.

2. The press-button switch is operated briefly. This feeds current through its 10K resistor to the base of transistor TR1, switching it hard on. This drops the collector voltage of TR1 to near 0 Volts, taking the left hand side of the capacitor with it.

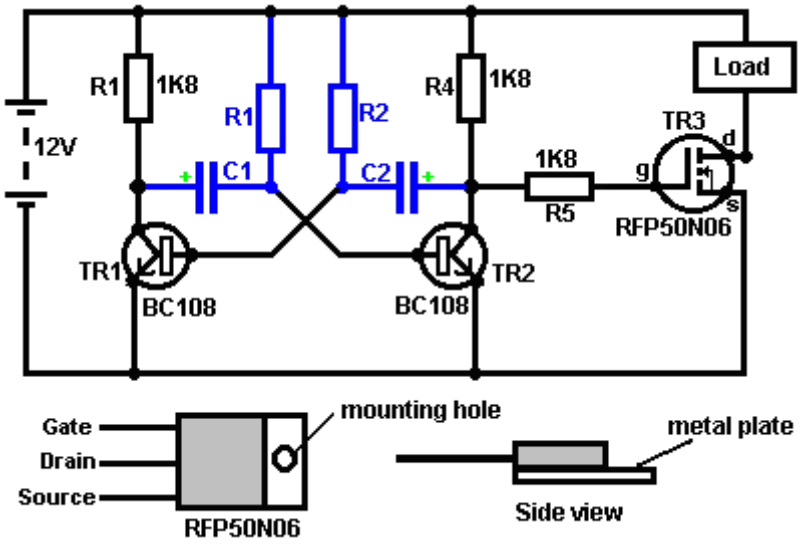
3. As the voltage across a capacitor can't change instantly, the right hand side of the capacitor drives the base of transistor TR2 down below 0.7 Volts, causing TR2 to switch off.

4. The circuit can't hold TR2 in its 'off' state for ever. The resistor '**R**' feeds current into the capacitor, forcing the voltage at the base of TR2 steadily upwards until the voltage reaches 0.7 Volts and transistor TR2 switches on again, forcing TR1 off again (provided that the press-button switch has been released). This is the stable state again. If the press-button switch is held on, then both transistors will be on and the output voltage will still be low. Another output pulse will not be generated until the press-button is let up and pressed again.

This circuit could be used to switch a microwave oven on for any chosen number of seconds, create a delay on your home-built burglar alarm, to give you time to switch it off after walking through your front door, operate a solenoid valve to feed a pre-determined quantity of beverage into a bottle on a production line, or whatever...

The Astable Multivibrator.

The astable circuit is the monostable with a second capacitor added so that neither state is stable. This results in the circuit flopping backwards and forwards continuously:

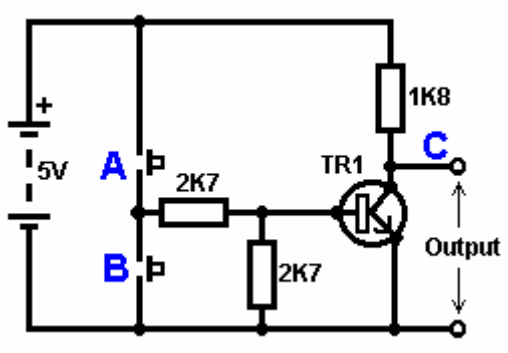


The rate of switching is controlled by the R1/C1 and R2/C2 combinations. The load's ON time to its OFF time is called the 'mark-space' ratio, where the ON period is the 'mark' and the OFF period is the 'space'. If you choose to use electrolytic capacitors which have their own polarity, then the +ve end of each capacitor is connected to the transistor collector.

While it is good to understand how these multivibrator circuits operate and can be built, nowadays there are pre-built circuits encased in a single package which you are much more likely to choose to use. These are called Integrated Circuits or 'ICs' for short. We will be discussing these shortly. Before we do, notice that in the circuit above, transistor TR3 has been changed to a new variety called a **F**ield **E**ffect **T**ransistor ('FET'). This type of transistor is newer than the 'bipolar' transistors shown in the earlier circuits. FETs come in two varieties: 'n-channel' which are like NPN transistors and 'p-channel' which are like PNP transistors.

FETs are more difficult to make but have now reached a level of cost and reliability which makes them very useful indeed. They require almost no base current (called 'gate' current with this type of transistor) which means that they have almost no effect on any circuit to which they are attached. Also, many of them can handle large currents and boast major power handling capabilities. Because of this, it is usual to see them packaged with a metal plate mounting, ready to be bolted to an aluminium heat-sink plate to help dissipate the heat generated by the large amount of power flowing through them. The 'RFP50N06' shown above can handle up to 50 Volts and carry up to 60 Amps, which is serious power handling.

Inverters and Truth Tables.
Consider the following circuit:

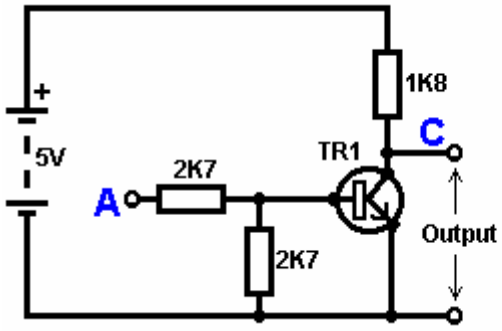


If neither of the press-button switches are operated, the transistor has no base/emitter current flow and so it is off. This places the collector voltage at '**C**' near the positive rail (+5 Volts).

If press-button switch '**A**' is operated, the base voltage tries to rise to half of the battery voltage but doesn't make it because the transistor base pins it down to 0.7 Volts. This feeds base current to the transistor, switching it hard on and causing the output at '**C**' to drop to nearly 0 Volts.

If press-button switch '**B**' is operated (don't do this when switch '**A**' is closed or you will get a very high 'short-circuit' current flowing directly through the two switches) it has no effect on the output voltage which will stay high.

If we re-draw the circuit like this:



TRUTH TABLE	
Input A	Output C
0	1
1	0

Key:
0 = Less than 0.5 Volts
1 = More than 3.5 Volts

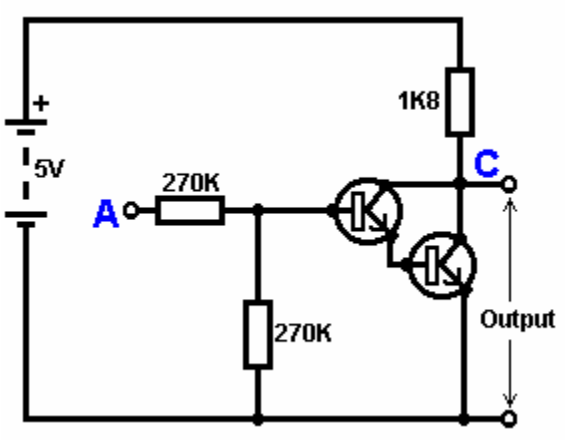
We can see that if the voltage at the input 'A' is taken high, then the output voltage at 'C' will be low. If the voltage at the input 'A' is taken low, then the output voltage at 'C' will be high. A circuit which does this is called an 'Inverter' because it 'inverts' (or 'turns upside down') the input voltage.

We can summarise this operation in a table. Personally, I would call the table an 'Input/Output' table, but for no obvious reason, the standard name is a 'Truth' table. The purpose of this table is to list all of the possible inputs and show the corresponding output for each input.

Another standard, is to substitute '1' for 'High Voltage' and '0' for 'Low Voltage'. You will notice that many items of electrical and electronic equipment have these symbols on the ON / OFF switch. In computer circuitry (hah! you didn't notice that we had moved to computer circuits, did you?), the '0' represents any voltage below 0.5 Volts and the '1' represents any voltage above 3.5 Volts. Many, if not most, computers operate their logic circuits on 5 Volts. This Inverter circuit is a 'logic' circuit.

A criticism of the above circuit is that its input resistance or 'impedance' is not particularly high, and its output impedance is not particularly low. We would like our logic circuits to be able to operate the inputs of eight other logic circuits. The jargon for this is that our circuit should have a 'fan-out' of eight.

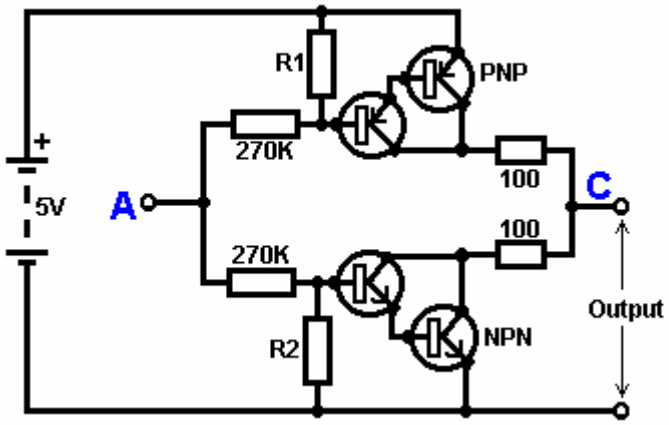
Let's go for a simple modification which will improve the situation:

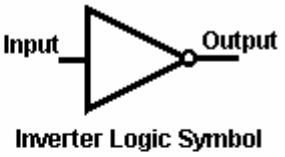


Here, The input impedance has been increased by a factor of 100 by using a Darlington pair of transistors which need far less base current, and so can have a much higher input resistor.

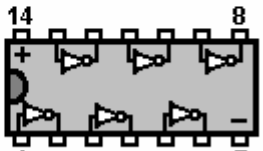
Unfortunately, the output impedance is still rather high when the transistors are in their OFF state as any current taken from the positive line has to flow through the 1K8 (1800 ohm) resistor. But we need this resistor for when the transistors are in their ON state. We really need to change the 1K8 resistor for some device which has a high resistance at some times and a low resistance at other times. You probably have not heard of these devices, but they are called 'transistors'.

There are several ways to do this. We might choose to use PNP transistors (we normally use NPN types) and connect these in place of the 1K8 resistor. Perhaps we might use a circuit like this:





Inverter Logic Symbol



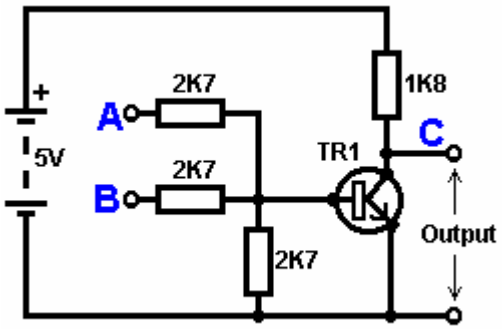
7414
Pin connections

This circuit is starting to look complicated and I don't like complicated circuits. It is not as bad as it looks. The NPN transistors at the bottom are almost the same as the previous circuit. The only difference is that the collector load is now two 100 ohm resistors plus the resistance of the two transistors. If the PNP transistors are OFF when the NPN transistors are ON, then the circuit loading on the NPN transistors will be negligible and the whole of the NPN transistors output will be available for driving external circuits through the lower 100 ohm resistor (a large 'fan-out' for the '0' logic state). To make sure that the PNP transistors are hard off before the NPN transistors start to switch on, the resistor 'R1' needs to be selected carefully.

The PNP transistors are an exact mirror image of the NPN side, so resistor R2 needs to be selected carefully to ensure that the NPN transistors are switched hard OFF before the PNP transistors start to switch ON.

You need not concern yourself unduly with that circuit, because you will almost certainly use an **I**ntegrated **C**ircuit rather than building your own circuit from 'discrete' components. An Integrated Circuit containing six complete inverters is the 7414 which is shown above. This comes in a small black case with two rows of 7 pins which make it look a bit like a caterpillar. Because there are two row of pins, the packaging is called "Dual In-Line" or "DIL" for short.

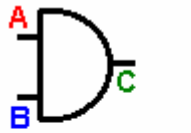
Now, consider the following circuit:



TRUTH TABLE		
Input A	Input B	Output C
0	0	1
0	1	0
1	0	0
1	1	0

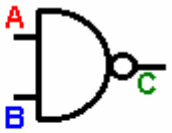
0 = Low voltage 1 = High Voltage

This circuit operates the same way as the Inverter circuit, except that it has two inputs ('**A**' and '**B**'). The output voltage at '**C**' will be low if either, '**A**' **OR** '**B**' or both, of the inputs is high. The only time that the output is high, is when both Input '**A**' and Input '**B**' are low. Consequently, the circuit is called an "**OR**" gate. Strictly speaking, because the output voltage goes Down when the input voltage goes Up, it is called a "**Not OR**" gate, which gets shortened to a "**NOR**" gate. In this context, the word "not" means "inverted". If you fed the output '**C**' into an inverter circuit, the resulting circuit would be a genuine "**OR**" gate. The digital circuit symbols for an **AND** gate, a **NAND** gate, an **OR** gate and a **NOR** gate are:



AND Gate

A	0	1	0	1
B	0	0	1	1
C	0	0	0	1

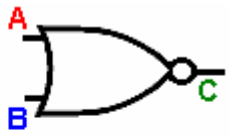


NAND Gate

A	0	1	0	1
B	0	0	1	1
C	1	1	1	0

OR Gate

A	0	1	0	1
B	0	0	1	1
C	0	1	1	1

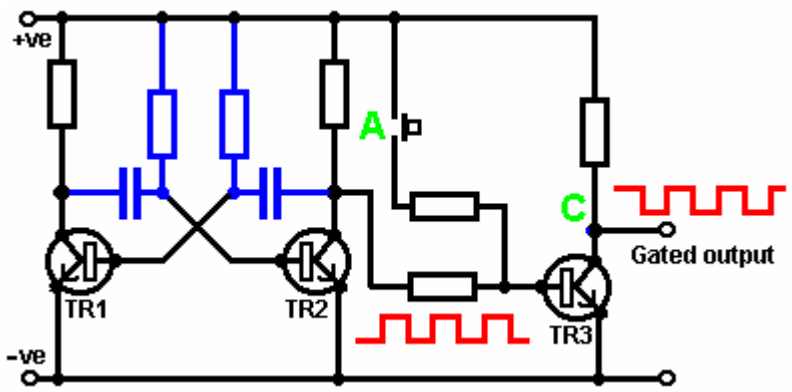


NOR Gate

A	0	1	0	1
B	0	0	1	1
C	1	0	0	0

Gating.

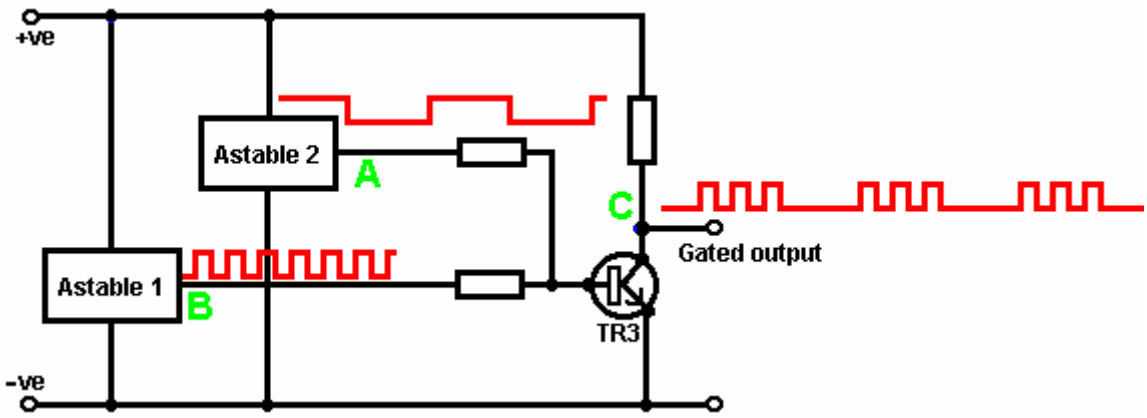
These common chips are usually supplied with 2, 4 or 8 inputs. So, why is it called a "Gate" - isn't it just a double inverter? Well, yes, it **is** a double inverter, but a double inverter acts as a gate which can pass or block an electronic signal. Consider this circuit:



Here, transistors 'TR1' and 'TR2' are connected to form an astable (multivibrator). The astable runs freely, producing the square wave voltage pattern shown in red. Transistor 'TR3' passes this voltage signal on. TR3 inverts the square wave, but this has no practical effect, the output being the same frequency square wave as the signal taken from the collector of TR2.

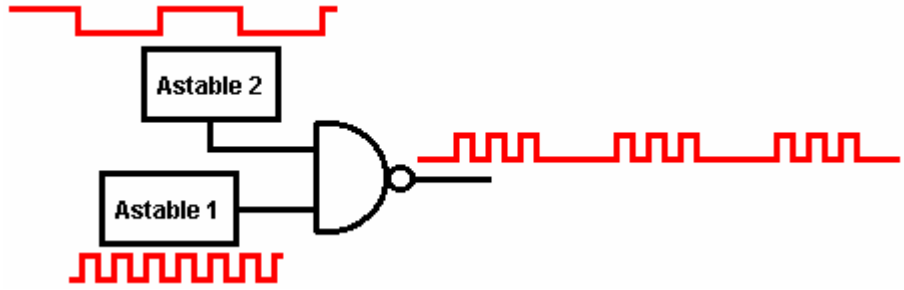
If the press-button switch at point '**A**' is operated, a current is fed to the base of TR3 which holds it hard on. The voltage at point '**C**' drops to zero and stays there. The square wave signal coming from the collector of TR2 is blocked and does not reach the output point '**C**'. It is as if a physical 'gate' has been closed, blocking the signal from reaching point '**C**'. As long as the voltage at point '**A**' is low, the gate is open. If the voltage at point '**A**' goes high, the gate is closed and the output is blocked.

There is no need for a manual switch at point '**A**'. Any electronic switching circuit will do:



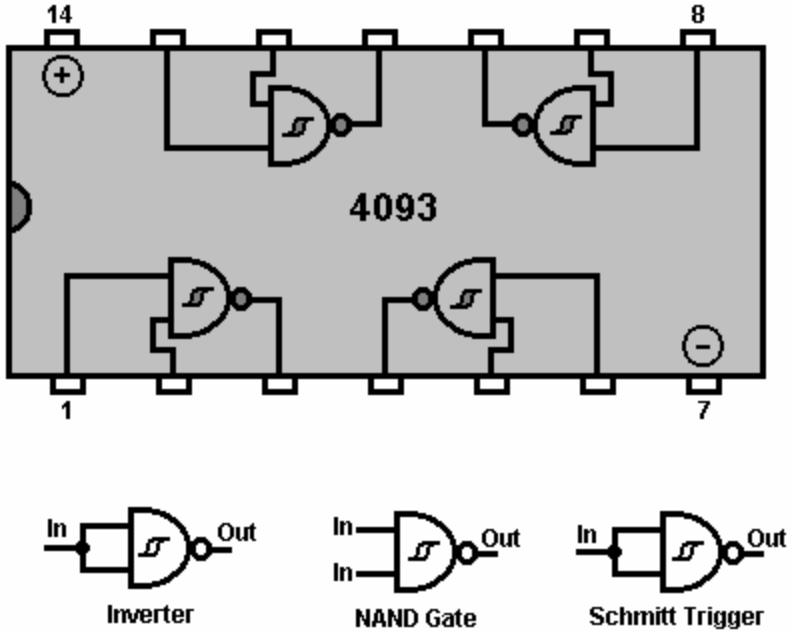
Here, a slow-running astable is substituted for the manual switch. When the output voltage of 'Astable 2' goes high, it switches the gate transistor 'TR3', holding it hard on and blocking the square-wave signal from 'Astable 1'. When the output voltage of 'Astable 2' goes low, it frees transistor 'TR3' and it then passes the 'Astable 1' signal through again. The resulting gated waveform is shown in red at point '**C**' and it is bursts of signal, controlled by the running rate of 'Astable 2'. This is the sort of waveform which Stan Meyer found very effective in splitting water into Hydrogen and Oxygen (see Chapter 10).

This circuit could also be drawn as:



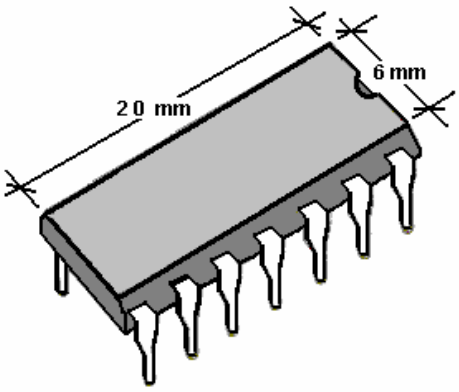
The small circle on the output side of logic devices is to show that they are inverting circuits, in other words, when the input goes up, the output goes down. The two logic devices we have encountered so far have had this circle: the Inverter and the NAND gate.

If you wish, you can use a NAND gate chip which has the circuitry also built as a Schmitt trigger, which as you will recall, has a fast-switching output even with a slowly moving input. With a chip like that, you can get three different functions from the one device:



If the two inputs of a NAND gate are connected together, then the output will always be the opposite of the input, i.e. the gate acts as an inverter. This arrangement also works as a Schmitt Trigger due to the way the NAND gate circuitry is built. There are several packages built with this type of circuitry, the one shown here is the "74132" chip which contains four "dual-input" NAND gates. Gates can have almost any number of inputs but it is rare to need more than two in any given circuit. Another chip with identical pin connections is the **4011** chip (which is not a Schmitt circuit). This 'quad dual-input' NAND gate package uses a construction method called "CMOS" which is very easily damaged by static electricity until actually connected into a circuit. CMOS chips can use a wide range of voltages and take very little current. They are cheap and very popular

The number of devices built into an Integrated Circuit is usually limited by the number of pins in the package and one pin is needed for one connection to 'the outside world'. Packages are made with 6 pins (typically for opto-isolators), 8 pins (many general circuits), 14 pins (many general circuits, mostly computer logic circuits), 16 pins (ditto, but not as common) and then a jump to large numbers of pins for Large Scale devices such as microprocessors, memory chips, etc. The standard IC package is small:



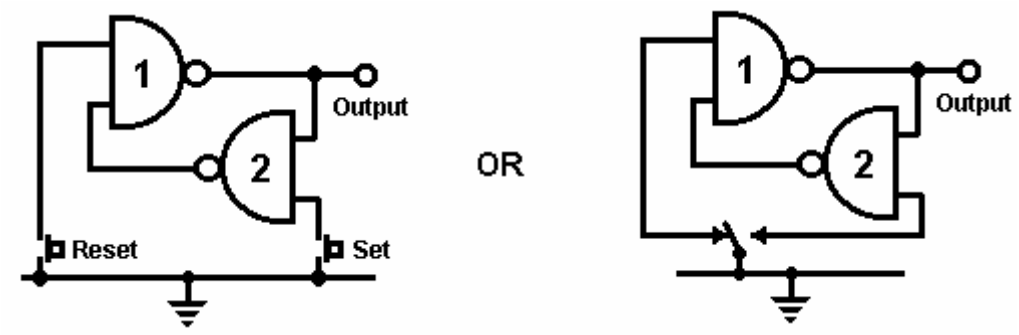
Prototype circuits are often built on 'strip board' which is a stiff board with strips of copper running along one face, and punched with a matrix of holes. The strips are used to make the electrical connections and are broken where necessary. This strip board is usually called "Veroboard":

Nowadays, the strip board holes are spaced 2.5 mm (1/10") apart which means that the gaps between the copper strips is very small indeed. I personally, find it quite difficult to make good solder joints on the strips without the solder bridging between two adjacent strips. Probably, a smaller soldering iron is needed. I need to use an 8x magnifying glass to be sure that no solder bridging remains in place before a new circuit is powered up for the first

time. Small fingers and good eyesight are a decided advantage for circuit board construction. The narrow spacing of the holes is so that the standard IC DIL package will fit directly on the board.

Circuits built using computer circuitry, can experience problems with mechanical switches. An ordinary light switch turns the light on and off. You switch it on and the light comes on. You switch it off and the light goes off. The reason it works so well is that the light bulb takes maybe, a tenth of a second to come on. Computer circuits can switch on and off 100,000 times in that tenth of a second, so some circuits will not work reliably with a mechanical switch. This is because the switch contact bounces when it closes. It may bounce once, twice or several times depending on how the switch is operated. If the switch is being used as an input to a counting circuit, the circuit may count 1, 2 or several switch inputs for one operation of the switch. It is normal to "de-bounce" any mechanical switch. This could be done using a couple of NAND gates connected like this:

The NAND Latch.



Here, the mechanical switch is buffered by a 'latch'. When the 'Set' switch is operated, the output goes low. The unconnected input of gate '1' acts as if it has a High voltage on it (due to the way the NAND gate circuit was built). The other input is held low by the output of gate '2'. This pushes the output of gate '1' high, which in turn, holds the output of gate '2' low. This is the first stable state.

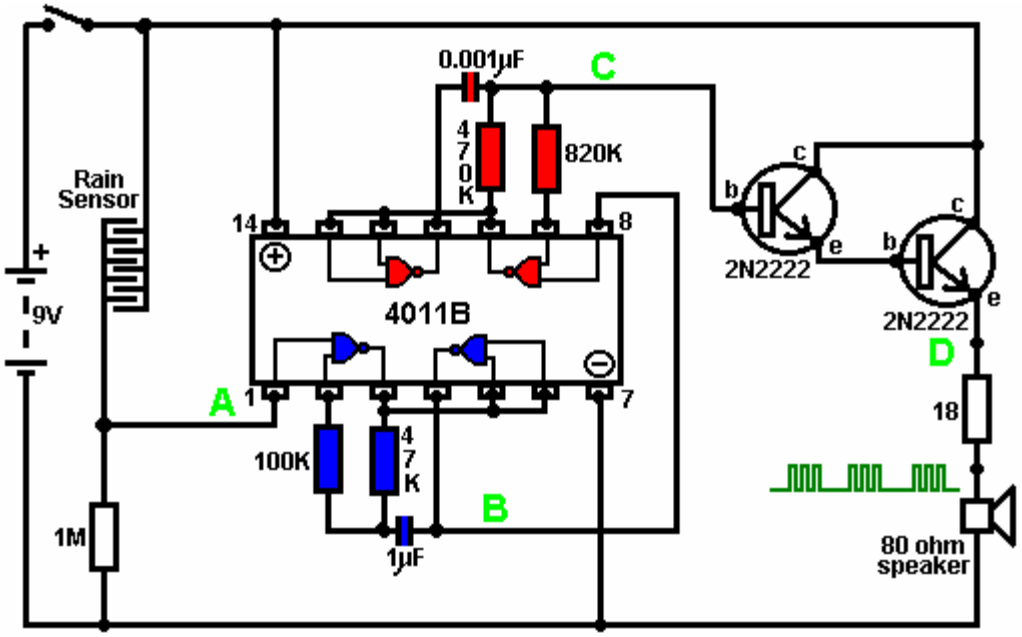
When the 'Set' switch is operated, the output of gate '2' is driven high. Now, both inputs of gate '1' are high which causes its output to go low. This in turn, drives one input of gate '2' low, which holds the output of gate '2' high. This is the second stable state.

To summarise: pressing the 'Set' switch any number of times, causes the output to go low, once and only once. The output will stay low until the 'Reset' switch is operated once, twice or any number of times, at which point the output will go high and stay there.

This circuit uses just half of one cheap NAND gate chip to create a bistable multivibrator which is physically very small and light.

Gate Circuits.

NAND Gates can be used as the heart of many electronic circuits apart from the logic circuits for which the package was designed. Here is a NAND gate version of the rain alarm described earlier. The '4011B' chip is a CMOS device which has a very high input impedance and can operate at convenient battery voltages (3 to 15 Volts):



This circuit is comprised of a rain sensor, two astable multivibrators and a power-driver feeding a loudspeaker:

1. The rain sensor is a wired-up strip board or similar grid of interlaced conductors, forming a voltage-divider across the battery rails.

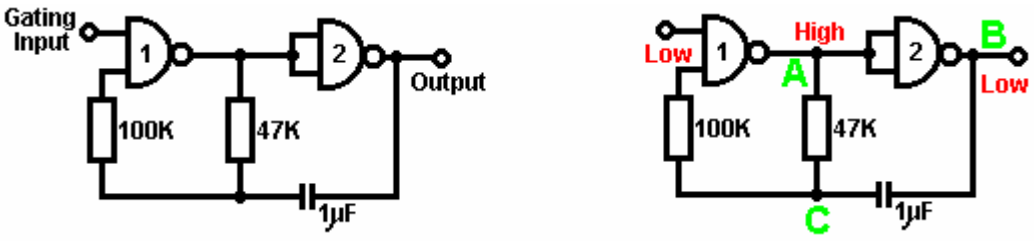
2. The output voltage from this, at point '**A**' in the circuit diagram, is normally low as the strip board is open-circuit when dry. This holds the first NAND gate locked in the OFF state, preventing the first astable from oscillating. This first astable is colour-coded blue in the diagram. Its frequency (the pitch of the note it produces) is governed by the values of the 47K resistor and the 1 microfarad capacitor. Reducing the value of either of these will raise the frequency (note pitch). If rain falls on the sensor, the voltage at point '**A**' goes high letting the astable run freely. If the voltage at '**A**' does not rise sufficiently when it rains, increase the value of the 1M resistor.

3. The output of the first astable is a low voltage when the sensor is dry. It is taken from point '**B**' and passed to the gating input of the second astable, holding it in its OFF state. The speed of the second astable is controlled by the value of the 470K resistor and the 0.001 microfarad capacitor. Reducing the value of either of these will raise the pitch of the note produced by the astable. The rate at which this astable operates is very much higher than the first astable.

When it rains, the voltage at point '**A**' rises, letting the first astable oscillate. As it does so, it turns the second astable on and off in a steady rhythmic pattern. This feeds repeated bursts of high speed oscillations from the second astable to point '**C**' in the diagram.

4. The Darlington-pair emitter-follower transistors cause the voltage at point '**D**' to follow the voltage pattern at point '**C**' (but 1.4 Volts lower voltage due to the 0.7 Volts base/emitter voltage drop for each transistor). The high gain of the two transistors ensures that the output of the second oscillator is not loaded unduly. These power-driver transistors place the output voltage across an eighty ohm loudspeaker, padded with a resistor to raise the overall resistance of the combination. The voltage pattern produced is shown at point '**D**' and is an attention-grabbing sound.

So, why does this circuit oscillate?:

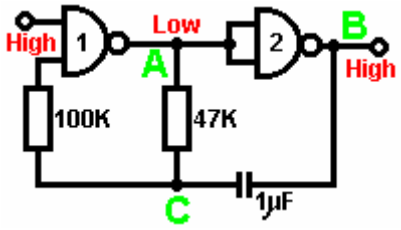


The circuit will not oscillate if the gating input is low, so assume it to be high. Take the moment when the output of gate 2 is low. For this to happen, the inputs of gate 2 have to be high. As the output of gate 1 is wired directly to the inputs of gate 2, it must be high, and for that to be true, at least one of its inputs must be low. This situation is

shown on the right.

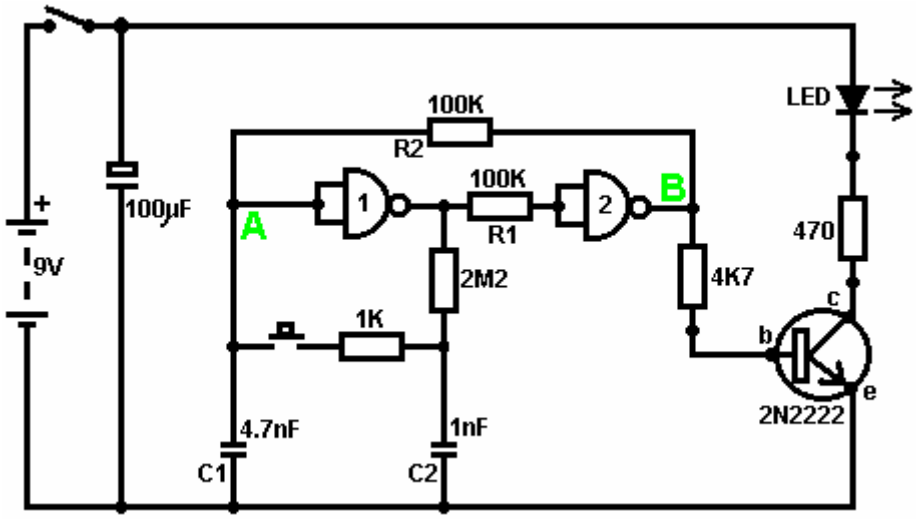
There is now a full voltage drop between point '**A**' and point '**B**'. The 47K resistor and the capacitor are in series across this voltage drop, so the capacitor starts to charge up, progressively raising the voltage at point '**C**'. The lower the value of the resistor, the faster the voltage rises. The larger the value of the capacitor, the slower the voltage rises.

When the voltage at point '**C**' rises sufficiently, the 100K resistor raises the input voltage of gate 1 far enough to cause it to change state. This creates the following situation:



Now, the voltage across '**A**' to '**B**' is reversed and the voltage at point '**C**' starts to fall, its rate governed by the size of the 47K resistor and the 1 microfarad capacitor. When the voltage at point '**C**' falls low enough, it takes the input of gate 1 low enough (via the 100K resistor) to cause gate 1 to switch state again. This takes the circuit to the initial state discussed. This is why the circuit oscillates continuously until the gating input of gate 1 is taken low to block the oscillation.

Now, here is a NAND gate circuit for a sequential on/off switch:



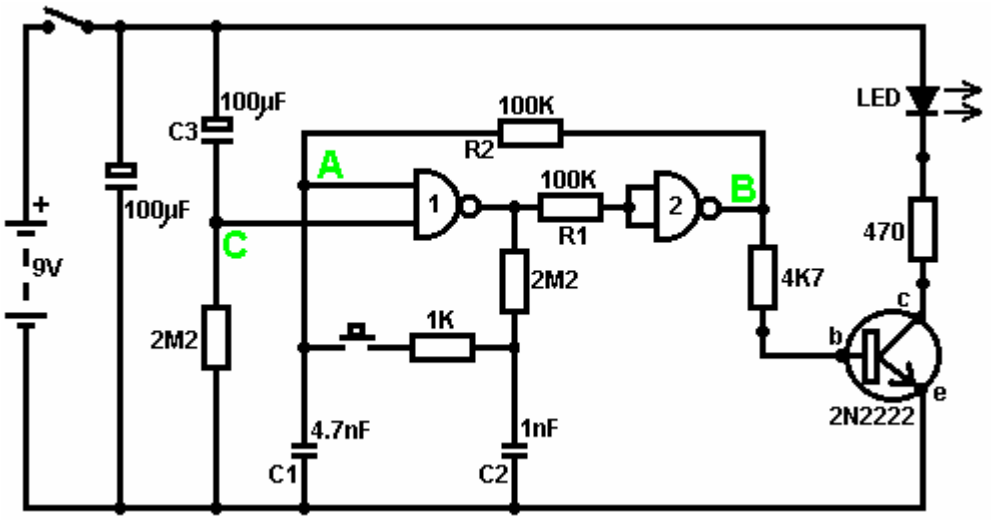
This circuit turns the Light Emitting Diode on and off repeatedly with each operation of the press-button switch. When the on/off switch is closed, capacitor 'C1' holds the voltage at point '**A**' low. This drives the output of gate 1 high, which moves the inputs of gate 2 high via the 100K resistor 'R1'. This drives the voltage at point '**B**' low, turning the transistor off, which makes the LED stay in its off state. The low voltage at point '**B**' is fed back via the 100K resistor 'R2' to point '**A**', keeping it low. This is the first stable state.

As the output of gate 1 is high, capacitor 'C2' charges up to that voltage via the 2M2 resistor. If the press-button switch is operated briefly, the high voltage of 'C2' raises the voltage of point '**A**', causing gate 1 to change state, and consequently, gate 2 to change state also. Again, the high voltage at point '**B**' is fed back to point '**A**' via the 100K resistor 'R2', keeping it high, maintaining the situation. This is the second stable state. In this state, point '**B**' has a high voltage and this feeds the base of the transistor via the 4.7K resistor, turning it on and lighting the LED.

In this second state, the output of gate 1 is low, so capacitor 'C2' discharges rapidly to a low voltage. If the press-button switch is operated again, the low voltage of 'C2' drives point '**A**' low again, causing the circuit to revert to its original stable state.

We could, if we wished, modify the circuit so that it would operate for three or four minutes after switch-on but then stop operating until the circuit was turned off and on again. This is accomplished by gating one of the gates instead of just using both as inverters. If we gated the second gate, then the LED would be left permanently on,

so we will modify the first gate circuit:



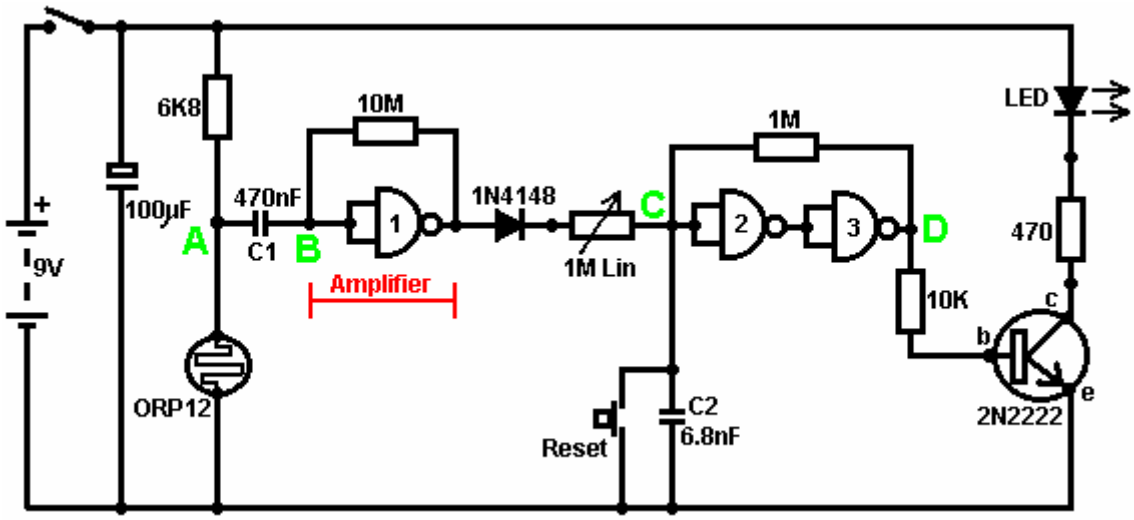
This circuit operates exactly the same way as the previous circuit if, and only if, the voltage at point '**C**' is high. With the voltage at point '**C**' high, gate 1 is free to react to the voltage at point '**A**' as before. If the voltage at point '**C**' is low, it locks the output of gate 1 at the high level, forcing the output of gate 2 to the low level and holding the LED off.

When the circuit is first powered up, the new 100 microfarad capacitor 'C3' is fully discharged, which pulls the voltage at point '**C**' to nearly + 9 Volts. This allows gate 1 to operate freely, and the LED can be toggled on and off as before. As time passes, the charge on capacitor 'C3' builds up, fed by the 2M2 resistor. This causes the voltage at point '**C**' to fall steadily. The rate of fall is governed by the size of the capacitor and the size of the resistor. The larger the resistor, the slower the fall. The larger the capacitor, the slower the fall. The values shown are about as large as are practical, due to the current 'leakage' of 'C3'.

After three or four minutes, the voltage at point '**C**' gets driven low enough to operate gate 1 and prevent further operation of the circuit. This type of circuit could be part of a competitive game where the contestants have a limited time to complete some task.

The NAND Gate as an Amplifier.

Gates can also be used as amplifiers although they are not intended to be used that way and there are far better integrated circuits from which to build amplifiers. The following circuit shows how this can be done:



This circuit operates when there is a sudden change in light level. The previous light-level switching circuit was designed to trigger at some particular level of increasing or decreasing level of lighting. This is a shadow-detecting circuit which could be used to detect somebody walking past a light in a corridor or some similar situation.

The voltage level at point '**A**' takes up some value depending on the light level. We are not particularly interested

in this voltage level since it is blocked from the following circuitry by capacitor 'C1'. Point '**B**' does not get a voltage pulse unless there is a sudden change of voltage at point '**A**', i.e. there is a sudden change in light level reaching the light-dependent resistor ORP12.

The first gate amplifies this pulse by some fifty times. The gate is effectively abused, and forced to operate as an amplifier by the 10M resistor connecting its output to its input. At switch-on, the output of gate 1 tries to go low. As its voltage drops, it starts to take its own inputs down via the resistor. Pushing the voltage on the inputs down, starts to raise the output voltage, which starts to raise the input voltage, which starts to lower the output voltage, which ..... The result is that both the inputs and the output take up some intermediate voltage (which the chip designers did not intend). This intermediate voltage level is easily upset by an external pulse such as that produced by the ORP12 through capacitor 'C1'. When this pulse arrives, an amplified version of the pulse causes a voltage fluctuation at the output of gate 1.

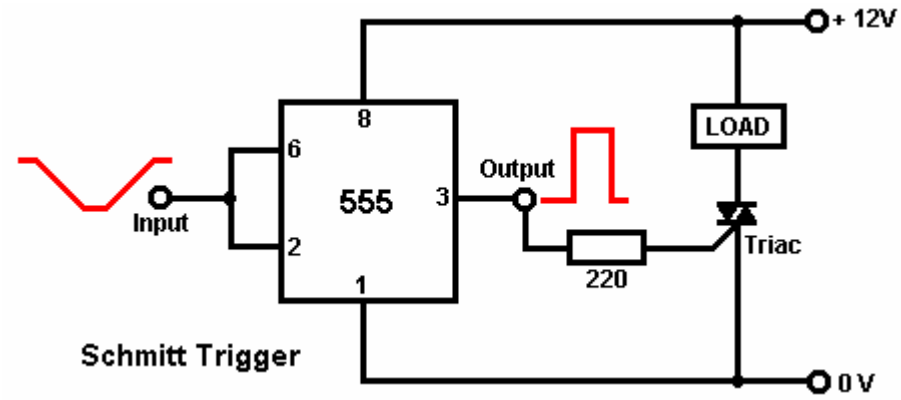
This voltage change is passed through the diode and variable resistor to the input of gate 2. Gates 2 and 3 are wired together as a makeshift Schmitt trigger in that the output voltage at point '**D**' is fed back to point '**C**' via a high value resistor. This helps to make their change of state more rapid and decisive. These two gates are used to pass a full change of state to the output stage transistor. The variable resistor is adjusted so that gate 2 is just about to change state and is easily triggered by the pulse from amplifier gate 1. The output is shown as an LED but it can be anything you choose. It could be a relay used to switch on some electrical device, a solenoid used to open a door, a counter to keep track of the number of people using a passageway, etc. etc. Please note that an operational amplifier chip (which will be described later) is a far better choice of IC for a circuit of this type. A gate amplifier is shown here only to show another way that a gate can be utilised.

The NE555 Timer Chip.

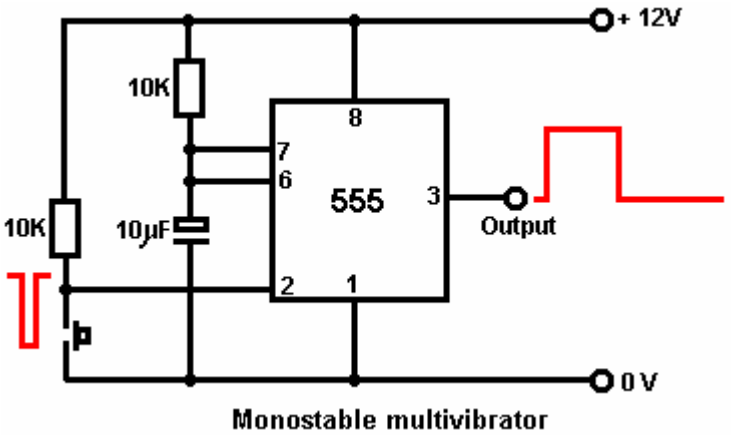
There is an exceptionally useful chip designated by the number 555. This chip is designed to be used in oscillator and timer circuits. Its use is so widespread that the chip price is very low for its capability. It can operate with voltages from 5 Volts to 18 Volts and its output can handle 200 mA. It takes 1 mA when its output is low and 10 mA when its output is high. It comes in an 8-pin **D**ual-**I**n-**L**ine package and there is a 14-pin package version which contains two separate 555 circuits. The pin connections are:

This device can operate as a monostable or astable multivibrator, a Schmitt trigger or an inverting buffer (low current input, high current output).

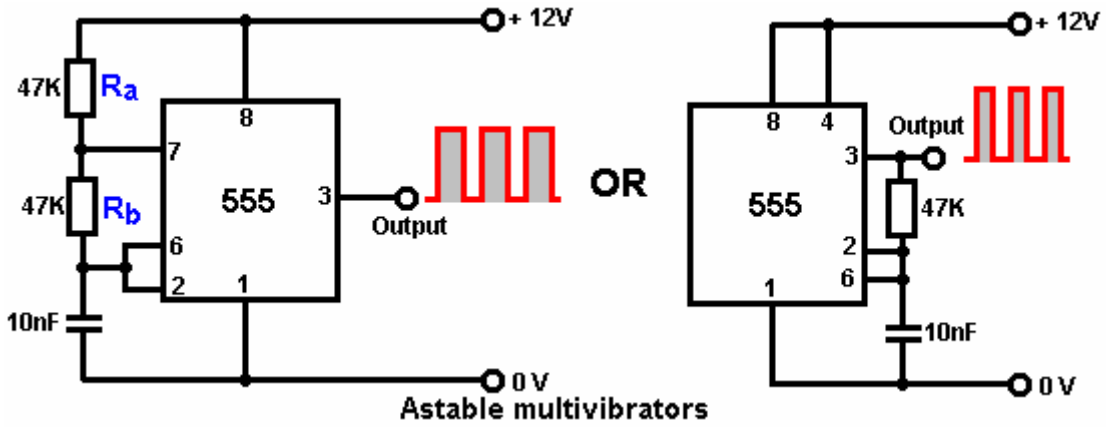
Here it is wired as a Schmitt trigger, and for variation, it is shown triggering a triac which will then stay on until the circuit is powered down (an SCR could be used just as well with this DC circuit):

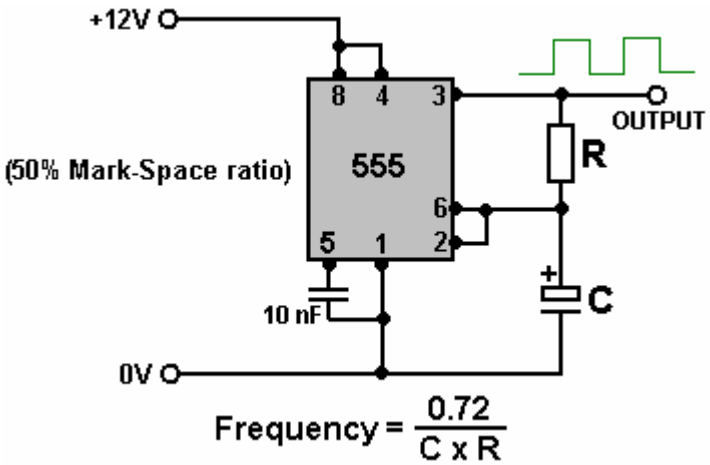


And here, a monostable:



And here are two astables, the second of which has fixed, equal mark/space ratio and the first a high output voltage time determined by Ra + Rb and a low voltage output time determined by Rb (2:1 in this case):



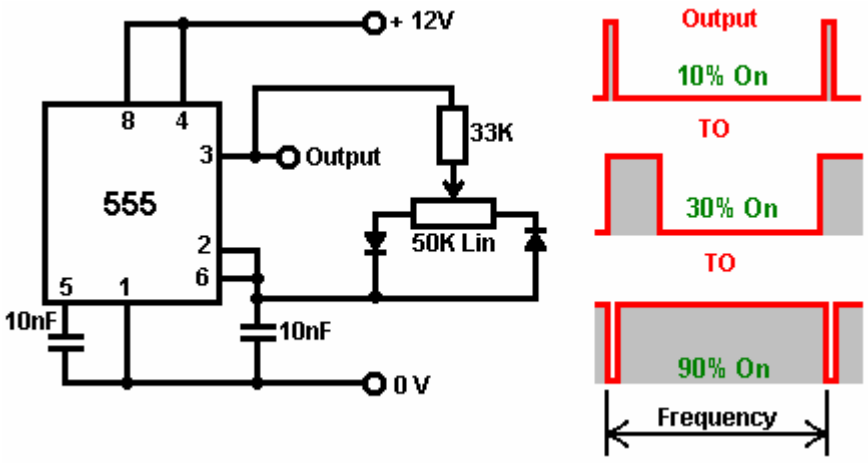


Astable Frequencies

	100	470	1K	4.7K	10K	47K	100K	470K	1M
0.1 µF	72,000 Hz	15,319 Hz	7,200 Hz	1,532 Hz	720 Hz	153 Hz	72 Hz	15 Hz	7.2 Hz
0.47 µF	15,319 Hz	3,259 Hz	1,532 Hz	326 Hz	153 Hz	33 Hz	15 Hz	3.3 Hz	1.5 Hz
1.0 µF	7,200 Hz	1,532 Hz	720 Hz	153 Hz	72 Hz	15 Hz	7.2 Hz	1.5 Hz	1.4 secs
2.2 µF	3,272 Hz	696 Hz	327 Hz	70 Hz	33 Hz	7 Hz	3.3 Hz	1.4 secs	3 secs
4.7 µF	1,532 Hz	326 Hz	153 Hz	33 Hz	15 Hz	3.3 Hz	1.5 Hz	3 secs	6.7 secs
10 µF	720 Hz	153 Hz	72 Hz	15 Hz	7.2 Hz	1.5 Hz	1.4 secs	6.7 secs	14 secs
22 µF	327 Hz	70 Hz	33 Hz	7 Hz	3.3 Hz	1.4 secs	3 secs	14 secs	30 secs
47 µF	153 Hz	33 Hz	15 Hz	3.3 Hz	1.5 Hz	3 secs	6.7 secs	30 secs	65 secs
100 µF	72 Hz	15 Hz	7.2 Hz	1.5 Hz	1.4 secs	6.7 secs	14 secs	65 secs	139 secs
220 µF	33 Hz	7 Hz	3.3 Hz	1.4 secs	3 secs	14 secs	30 secs	139 secs	307 secs
470 µF	15 Hz	3.3 Hz	1.5 Hz	3 secs	6.7 secs	30 secs	65 secs	307 secs	614 secs
1,000 µF	7.2 Hz	1.5 Hz	1.4 secs	6.7 secs	14 secs	65 secs	139 secs	614 secs	
2,200 µF	3.3 Hz	1.4 secs	3 secs	14 secs	30 secs	139 secs	307 secs		
4,700 µF	1.5 Hz	3.3 secs	6.7 secs	30 secs	65 secs	307 secs	614 secs		
10,000 µF	1.4 secs	6.7 secs	14 secs	65 secs	139 secs	614 secs			

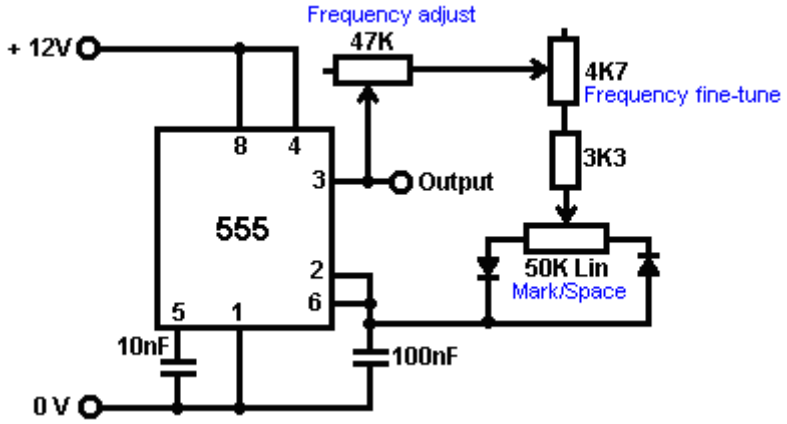
Note: The high leakage of large value electrolytic capacitors prevents them being used with high value resistors in timing circuits. Instead, use a smaller capacitor and follow the timing circuit with a “divide-by-N” chip to give accurately timed long periods. Not all 555 chips have a manufacturing quality sufficient for them to operate reliably above 20,000 Hz, so for the higher frequencies the chip needs to be selected after testing its actual performance.

We can also wire the 555 to give a variable mark/space ratio while holding the frequency of the oscillation fixed:



The output waveform changes drastically as the variable resistor is adjusted, but the frequency (or pitch of the note) of the output stays unaltered.

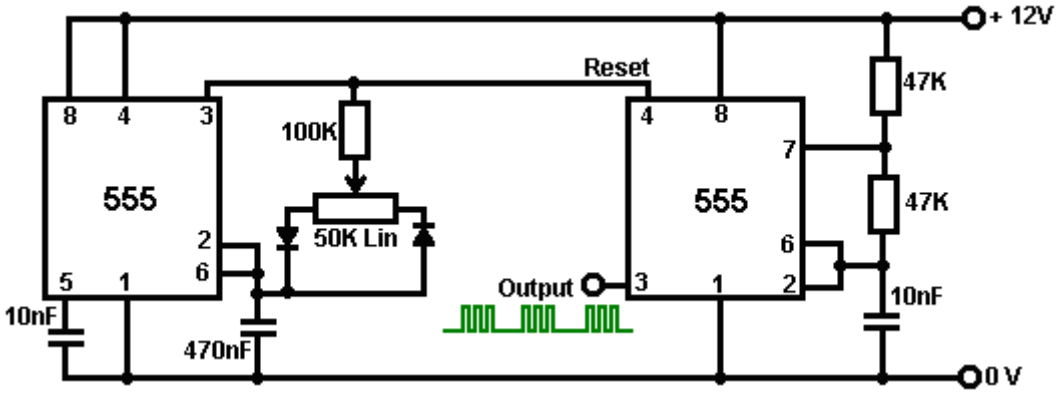
A variable-frequency version of this circuit can be produced by changing the 33K resistor to a variable resistor as shown here:



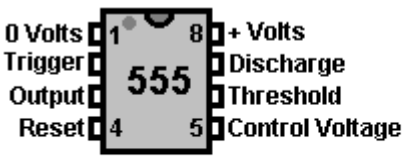
Here, the 33K resistor has been replaced by two variable resistors and one fixed resistor. The main variable resistor is 47K in size (an almost arbitrary choice) and it feeds to a second variable resistor of 4.7K in size. The advantage of this second variable resistor is that it can be set to it's mid point and the frequency tuning done with the 47K variable. When the frequency is approximately correct, the 4.7K variable can be used to fine tune the frequency. This is convenient as the small variable will have ten times more knob movement compared to the main variable (being just 10% of its value).

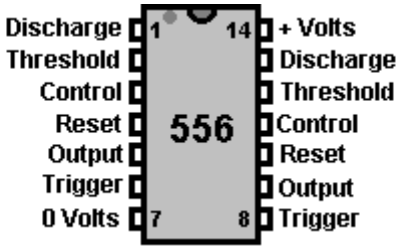
Obviously, it is not necessary to have the fine-tuning variable resistor, and it can be omitted without changing the operation of the circuit. As the 47K variable resistor can be set to zero resistance and the 4.7K variable resistor can also be set to zero resistance, to avoid a complete short-circuit between output pin 3 and the 50K Mark/Space variable resistor, a 3.3K fixed resistor is included. In this circuit, the frequency is set by your choice of the resistor chain 47K + 4.7K + 3.3K (adjustable from 55K to 3.3K) and the 100nF (0.1 microfarad) capacitor between pin 6 and the zero volt rail. Making the capacitor larger, lowers the frequency range. Making the resistors larger, also lowers the frequency range. Naturally, reducing the size of the capacitor and/or reducing the size of the resistor chain, raises the frequency.

One 555 chip can be used to gate a second 555 chip via its pin 4 'Reset' option. You will recall that we have already developed a circuit to do this using two astables and a transistor. We also generated the same effect using four NAND gates. Here, we will create the same output waveform using the more conventional circuitry of two 555 chips:



Both of the NE555 circuits can be bought in a single 14-pin DIL package which is designated 'NE556':



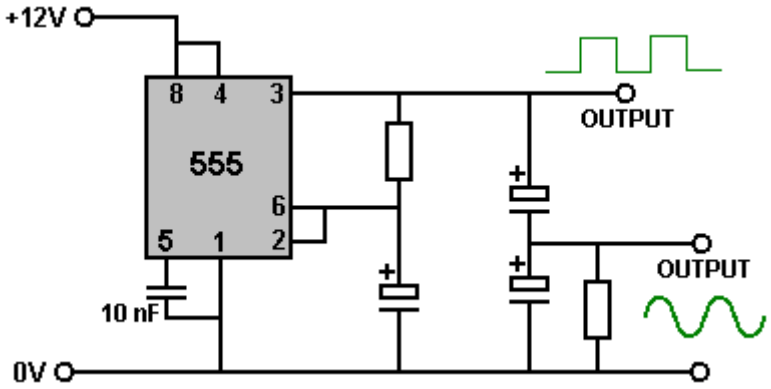


555 8-pin	555 No.1	555 No.2
2	6	8
3	5	9
4	4	10
5	3	11
6	2	12
7	1	13

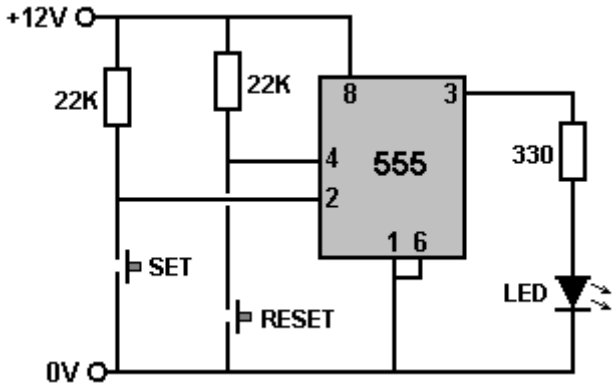
There are many additional circuit types which can be created with the 555 chip. If you wish to explore the possibilities, I suggest that you download the free pdf "50 555 Projects" from the web site:

http://www.talkingelectronics.com/projects/50%20-%20555%20Circuits/50%20-%20555%20Circuits.html.

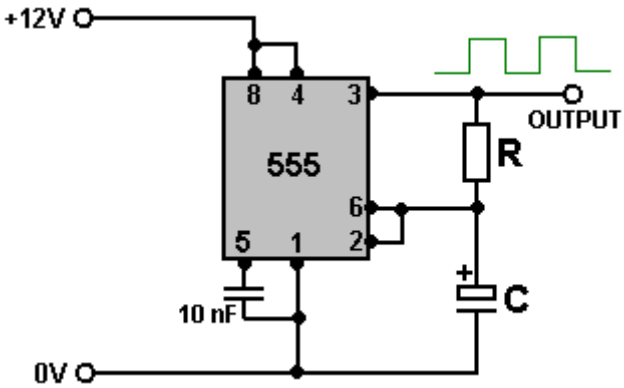
The 555 chip can also produce a sine wave output:



Or, if you wish, a bi-stable multivibrator:



All right, suppose that we want to design and build a circuit to do the same as Bob Beck's pulser circuit mentioned in chapter 11. The requirements are to produce a square wave output pulsing four times per second using a 27 volt power supply, the circuit being powered by three small PP3 size batteries. An obvious choice for the circuit seems to be a 555 timer chip which is small, robust and cheap and a suitable circuit would appear to be:



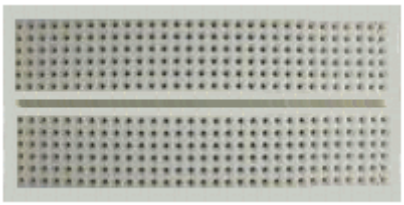
This leaves us with choosing a value for the capacitor and the resistor. We need to pay attention to the fact that the circuit will be running on 27 volts and while the capacitor will not charge up to anything like that voltage, we still will pick one which will survive 27V. Looking on the local eBay shows that a pack of ten capacitors of 1 microfarad rated at 50V can be bought for just £1 including postage, so take that as the value for "C". Looking at the 555 table of frequencies above shows:

Astable Frequencies

	100	470	1K	4.7K	10K	47K	100K	470K	1M
0.1 µF	72,000 Hz	15,319 Hz	7,200 Hz	1,532 Hz	720 Hz	153 Hz	72 Hz	15 Hz	7.2 Hz
0.47 µF	15,319 Hz	3,259 Hz	1,532 Hz	326 Hz	153 Hz	33 Hz	15 Hz	3.3 Hz	1.5 Hz
1.0 µF	7,200 Hz	1,532 Hz	720 Hz	153 Hz	72 Hz	15 Hz	7.2 Hz	1.5 Hz	1.4 secs

Which indicates that to get the circuit switching four times per second (4 Hz) the resistor "R" will need to be somewhere between 100K and 470K. With my capacitor, 120K is about right.

While the switching frequency does not have to be exact, let's aim at getting it correct. Most reasonably priced components have a tolerance of around 10% so we need to select our resistor/capacitor combination for the exact values of the actual components which we will use. For this, it is worth building the circuit on a solder-less 'breadboard', so looking on eBay again we find that a suitable small plug-in board can be bought and delivered for £3. It looks like this :



These type of boards allow ICs to be plugged in spanning the central divide, leaving up to five extra connections on every pin. Short lengths of solid-core wire can be used to connect between any two socket holes. This will allow us to plug in one of our capacitors and find what resistor (or what two resistors) make the circuit switch forty times in ten seconds.

However, if we go to http://www.alldatasheet.co.kr/ and download the data pdf for the NE555 chip, we find that the maximum 555 chip voltage is quite limited:

DC AND AC ELECTRICAL CHARACTERISTICS

T_A = 25°C, V_{CC} = +5V to +15 unless otherwise specified.

SYMBOL	PARAMETER	TEST CONDITIONS	SE555			NE555/SE555C			UNIT
			Min	Typ	Max	Min	Typ	Max	
V_{CC}	Supply voltage		4.5		18	4.5		16	V
I_{CC}	Supply current (low state)[1]	V_{CC}=5V, R_L=∞		3	5		3	6	mA
		V_{CC}=15V, R_L=∞		10	12		10	15	mA

This means that the chip is liable to burn out instantly if it is fed more than 16 volts. As we need to run our circuit on 27V this is a problem. As the 27V is being provided by three separate batteries, we could supply the 555 chip from just one of the batteries and run it on 9V which would be ok from the point of view of the chip as the table above shows that it can operate correctly with a supply voltage as low as 4.5 volts. The disadvantage of that arrangement is that one of the batteries will run down more quickly than the others and it would be nice to avoid that.

The table also shows that the current draw just to keep the 555 running can be anything from 6 to 15 milliamps. That is not a large current but the PP3 batteries have been chosen for their small size, allowing the whole circuit to be strapped to a person's wrist. A quick search on the internet shows that cheap PP3 batteries have a capacity of 400 milliamp-hours and the very expensive alkaline types 565 milliamp-hours. These ratings are the "C20" values, based on the battery being discharged at a constant current over a period of twenty hours, which would be ten days of use if Bob Beck's two hours per day protocol is followed.

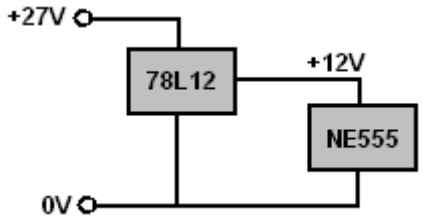
This means that the 'cheap' batteries should not be discharged at more than one twentieth of their 400 mAHr rating, which is 20 mA. The expensive alkaline batteries should be able to be discharged at 28 mA for twenty hours.

Our current draw is made up of two parts. The first part is supplying the circuit with the current which it needs to run. The second part is the current flowing through the body of the user. This second part is limited by the 820

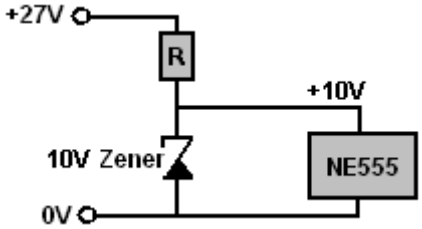
ohm resistor in the output line which limits that part of the current to a maximum of 33 milliamps (Ohm's Law: Amps = Volts /Resistance). This neglects the body resistance and assumes that the output control variable resistor is set to minimum resistance, which is unlikely.

Checking these values shows that the 555 chip is liable to draw as much current as the circuit supplies through the output electrodes. However, let's go ahead with the circuit, after all, we might decide to use rechargeable PP3 batteries which would overcome the need to buy new batteries every few days.

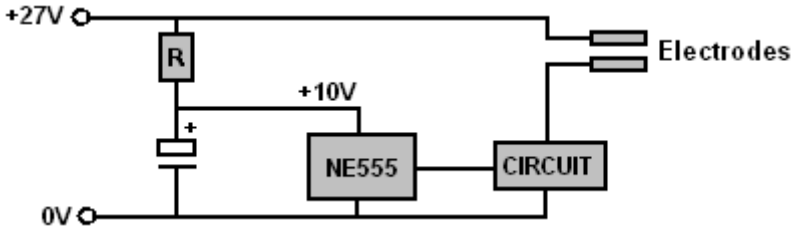
The first essential requirement is to provide the 555 chip with a voltage of, say, 10 volts when it is running in the completed circuit. That could be done with one of the voltage-stabiliser integrated circuits:



That is not a particularly expensive option, but those chips draw a current in order to provide the voltage stabilisation and an absolutely steady voltage is not needed by the 555 chip. Alternatively, we could use a resistor and a 10V zener diode:



But that method does waste some current flowing through the zener in order to provide the wanted voltage. The most simple method is to use a resistor and a capacitor:



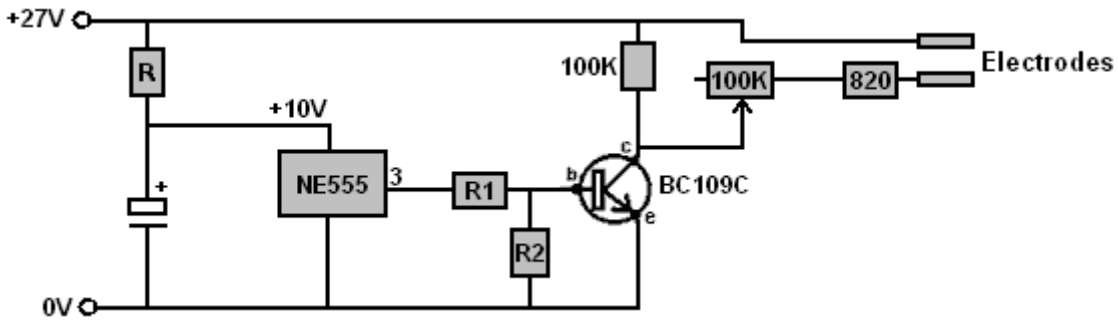
Considerable care is needed when selecting the resistor value "**R**". If the value is too low, then the voltage passed to the 555 chip will be too high and the chip will burn out. When selecting the resistor "**R**", start with a higher value than expected and then substitute slightly lower value resistors while monitoring the voltage across the capacitor to make sure that it stays low enough. The resistor value can be assessed using Ohm's Law. Assuming a current of about 6 mA, the voltage drop across the resistor being (27 – 10) = 17 volts, then a resistor of about 2.83K (as Ohms = Volts / Amps) which suggests that starting with a 4.7K resistor is likely to be ok, and then picking each lower standard resistor in turn until a satisfactory voltage across the capacitor is reached.

The capacitor could be 12V or 15V rated, but if one rated at a higher voltage is used, then if it is accidentally connected across the full 27V it will not be harmed in any way. The larger the capacitance, the better, say 220 microfarads which can be got for a few pence on eBay. If you want to play safe, you could connect a 12V zener diode across the capacitor. It will not draw any current under normal working conditions, but if anything should cause the voltage on the capacitor to rise, then it will fire up and hold the voltage down to a safe 12V level. I would be inclined to see the zener as being unnecessary, but the choice is always yours.

So what resistor power rating is needed? Well, if the resistor turns out to be a 2.7K and the capacitor voltage ends up as 9.5 volts, then the average voltage across the resistor is 17.5V which makes the current through it 6.48 mA and as Watts = Volts x Amps, the power rating needs to be 113 milliwatts, so the typical quarter-watt (250 mW) resistor should be perfectly ok. If two (nearly equal value) resistors in parallel are used to get some intermediate value of "**R**" then that increases the overall resistor wattage.

The output of the 555 chip is then used to drive the remainder of the circuit which operates at 27V. A BC109C transistor costs only a few pence, can handle the voltage and has a minimum gain of 200 although the gain can be anything up to 800 and a BC109 can handle the current quite easily. If you need to find out any of these things, then download a datasheet for the transistor from the internet.

The output of the 555 timer is on pin 3 and it can easily supply 200 mA which is far, far more current than we would ever need for this circuit. We can feed the 555 square-wave output to the 27V electrodes using a transistor:

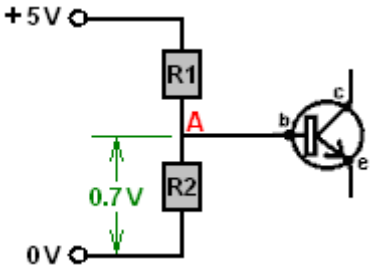


As the transistor is made of silicon, the switch-on voltage is when the base voltage is about 0.7 volts above the emitter voltage. That means that when the transistor is switched on, the top of resistor "**R1**" will be at around 10 volts and the bottom of "**R1**" will be at about 0.7 volts, which means that the voltage across "**R1**" will be (10 – 0.7) = 9.3 Volts. When that voltage is present across "**R1**" we want it to feed sufficient current to the transistor to switch it on fully. The transistor supplies a 100K resistor (which will carry 0.27 mA when 27 volts is across it) and the electrodes which will have a minimum resistance of 820 ohms across them (causing a current of 33 mA through them). So, the transistor might have to supply about 33 mA maximum. The BC109C transistor has a minimum gain of 200 so the current flowing into the base needs to be 33 / 200 = 0.165 mA and the resistor which will carry that current when it has 9.3 volts across it is 56.3K. A somewhat smaller resistor will suit.

A commonsense check that the resistor calculation is correct is:
A 1K resistor carries 1 mA per volt and so will carry 9.3 mA with 9.3 volts across it.
A 10K resistor will carry one tenth of that amount, or 0.93 mA with 9.3 volts across it.
A 100K resistor will carry one tenth of that again, or 0.093 mA with 9.3 volts across it.
This indicates that for a current of 0.165 mA which is about twice the 100K current, a resistor of about half of 100K should be about the right value, so 56.3K looks correct.

Considering that the gain of 200 is the minimum and three or four times that is typical, we could perhaps choose to use a 47K resistor for "**R1**"

As the electrode current is likely to be considerably less than 33 mA and as the BC109C gain is likely to be very high, it could be quite difficult to get the transistor to switch off as it can operate on very tiny amounts of input current. To get it to switch on and off cleanly when the 555 output voltage is say, about 5 volts, (at which point the NE555 voltage will be changing very rapidly), "**R2**" is included. With it in place, the output voltage of the NE555 is divided between "**R1**" and "**R2**" in the ratio of their resistances. The situation we want is:

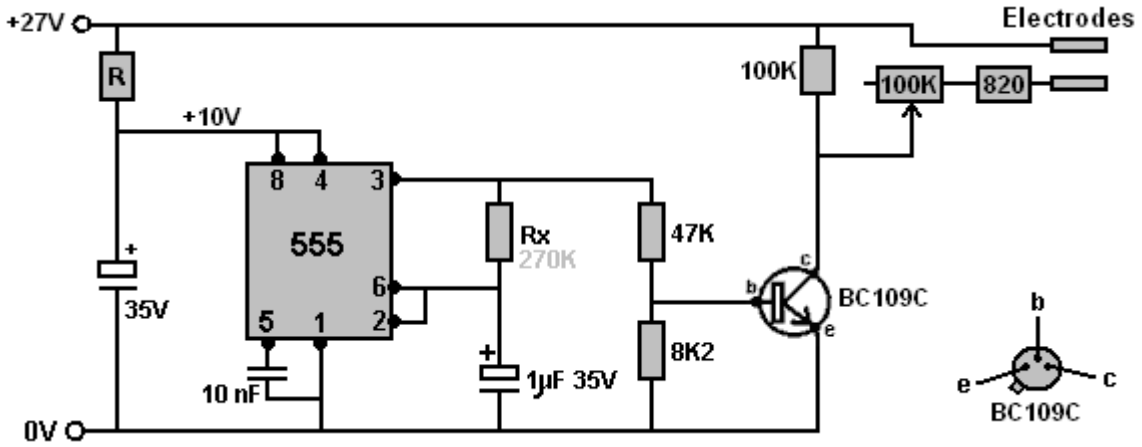


When The transistor is not switched on, it draws almost no current and so looks like a very high value resistor to the circuit. This allows the "**R1**" and "**R2**" resistors to act as a voltage-divider pair. This causes the voltage at point "**A**" to be determined by the ratio of "**R1**" to "**R2**" and the transistor can be ignored provided that the voltage at point "**A**" is below 0.7 volts. If the voltage at that point rises to 0.7 volts then the situation changes dramatically and Ohm's Law no longer holds as the transistor is not a passive resistor but instead, is an active semi-conductor device. If the voltage at point "**A**" tries to rise further it can't because the transistor base clamps it solidly there by appearing to be an ever lower resistor between the base and the emitter of the transistor. So for higher input voltages, resistor "**R2**" might as well not be there for all the difference it makes.

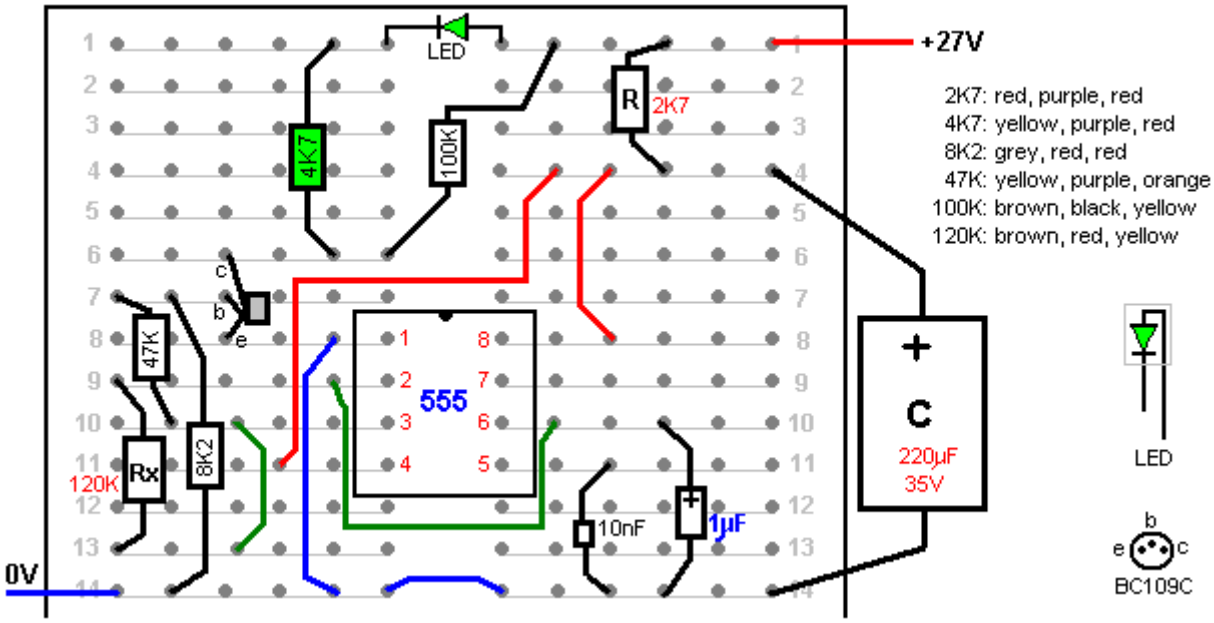
So, what value do we need for "**R2**" in order for the voltage at point "**A**" to be 0.7V when pin 3 of the NE555 reaches 5V? Well, that part of the circuit is acting in a resistive fashion and so Ohm's Law can be used. The resistor "**R1**" is 47K and has 4.3 volts across it, which means that the current through it must be 0.915 mA. That means that "**R2**" has 0.7V across it and 0.915 mA flowing through it which means that it has a value of 7.65K. A standard 8.2K or 6.8K resistor could be used as there is nothing dramatically important about the 5V switching point. If you were fussed about getting exactly 7.65K (and you shouldn't be), then you can get that value by combining two standard resistors, either in series or in parallel.

A common sense method of working out the value of "**R2**" is to use the fact that as the same current flows through them (no matter what that current happens to be), then the ratio of the voltage will be the same as the ratio of the resistors. That is: 0.7V / 4.3V = "**R2**" / 47K or "**R2**" = 47K x 0.7 / 4.3 which is 7.65K.

We have now reached the point where we can determine the resistor value needed to provide a reasonable voltage for the NE555 timer chip, the circuit being:



The "**Rx**" value is going to be fairly close to 270K so you can use that value when testing to find a suitable value for "**R**" (2.2K in my case). The capacitor across the NE555 chip should be as large a capacitance as is convenient, bearing in mind that the entire circuit, batteries, etc. is to fit into a small case to be strapped to a wrist. One way that the components could be positioned on the plug-board is:

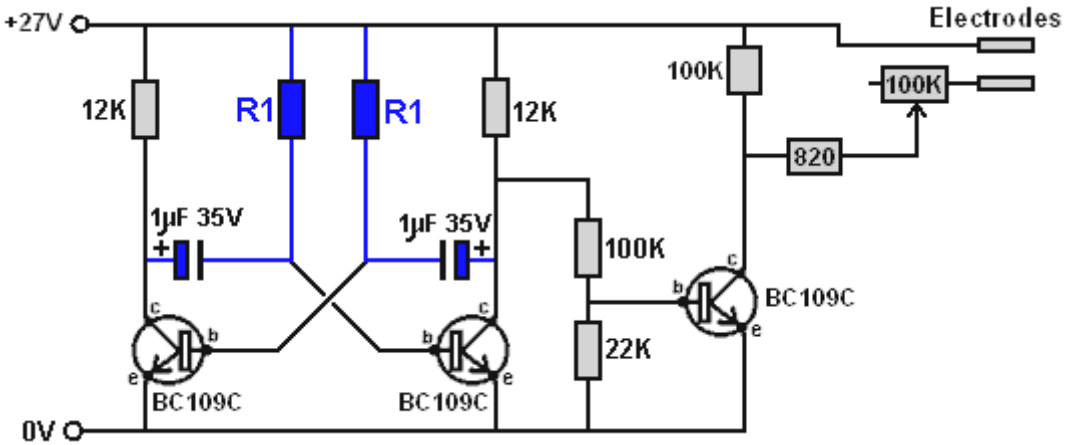


Remember that when trying various resistors for "**R**" you need to start high at about 4.7K and the resulting voltage on the capacitor shows the voltage drop across your first resistor choice and so, the actual current being drawn by your particular NE555 chip. That calculated current will allow you to calculate the resistor value needed to give 10 volts or so, allowing your next resistor to be tested to be almost exact in value.

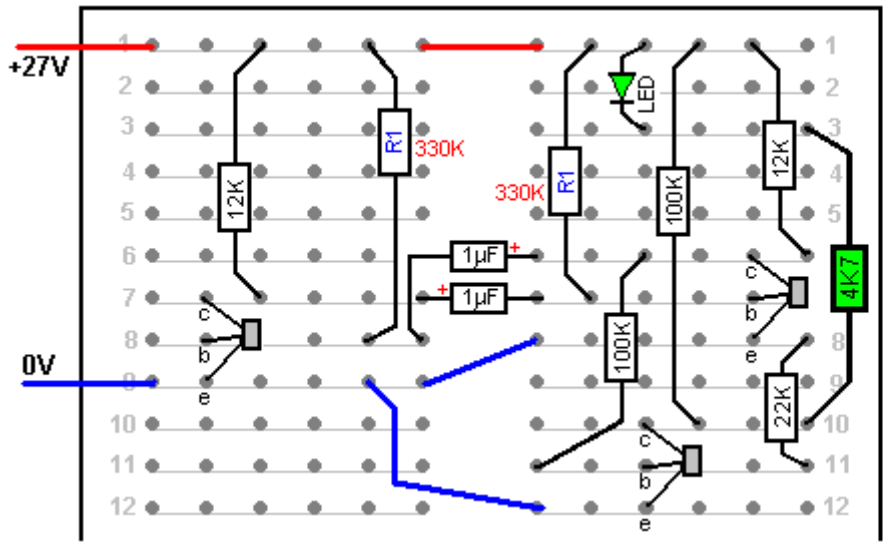
For checking the frequency produced by the circuit, any ordinary LED can be used as a temporary measure. It can be connected across the 100K 'load' resistor between the transistor collector and the +27V positive supply

line. A current-limiting resistor is essential to stop the LED burning out instantly. If we allow a current of 5 mA to flow through the LED then since the current-limiting resistor has some 26.3 volts across it, then it's value will be about 5.4K (1K would give 26 mA, 2K would give 13 mA, 3K would give 9 mA, 4K would give 6.5 mA) and so a 4.7K resistor works well. This LED and resistor are shown in the layout above. Please remember that if your BC109C transistor has a metal case, then that case is normally connected internally to the collector and so, care must be taken that the case does not short-circuit to anything else.

If it is considered important to maximise battery life by reducing the current draw to a minimum, then perhaps using an astable circuit might be a good choice. In common with most electronic circuits, there are many different ways to design a suitable circuit to do the required job. The BC109C transistor can handle the 27V and so we might aim at a current draw for the circuit of just 3 mA. If 2 mA were to flow through the astable transistors when they are switched on, then with 27V across them, the resistors would be 13.5K which is not a standard value. We might select 12K to give a 2.25 mA current, or 15K to give 1.8 mA. Either should be satisfactory. The circuit might then be:



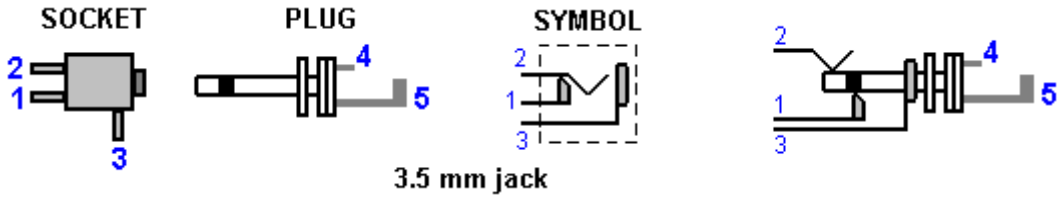
As the voltage swing feeding the output transistor has now risen from 10V to 27V the voltage-divider resistors can now increase in value by 2.7 times, giving around 127K and 22.1K for these resistors. However, the situation is not the same as for the NE555 chip which can supply at least 200 mA at the voltage-high output level. Instead, the transistor becomes such a high resistance that it can be ignored, but the 12K remains in the path which supplies the base current for the output transistor and it will in fact, add to the upper resistor of the voltage-divider pair. So while a 100K resistor is shown, it is effectively 112K due to that extra 12K resistor between it and the +27V supply line. The astable transistors will be switching fast at the point where the output transistor changes state, so the output square wave should be good quality. The BC109C transistor can switch on and off a hundred million times per second, so it's performance in this circuit should be very good. A test breadboard layout might be:



4K7: yellow, purple, red
12K: brown, red, orange
22K: red, red, orange
100K: brown, black, yellow
330K: orange, orange, yellow

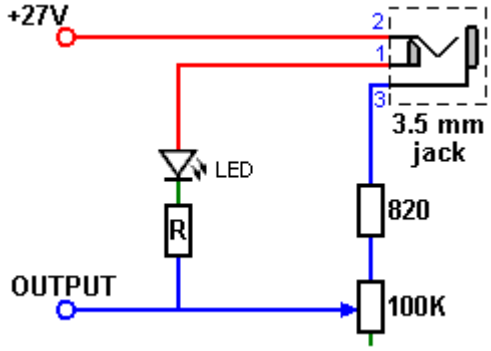
We now need to choose the timing components. For an even 50% duty cycle where each transistor is ON for half the time and OFF for half the time, the two timing capacitors can be the same size and then the two timing resistors will have the same value, in my case, 330K but it depends on the actual capacitors used.

Bob Beck's design calls for the LED display to be running when the unit is switched on and then be disconnected when the electrodes are plugged into a 3.5 mm socket mounted on the case containing the circuit. The switched socket looks like this:

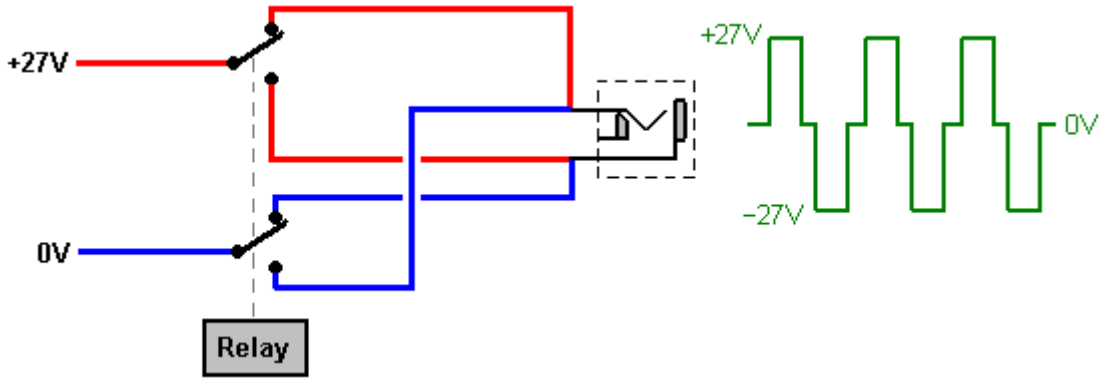


When the plug is not inserted into the socket, pin 1 connects to pin 2 and pin 3 is not connected to anything. When the plug is inserted, then pin 1 is isolated, pin 2 is connected to plug pin 4 and pin 3 is connected to plug pin 5.

The Beck circuit is connected to the output socket like this:

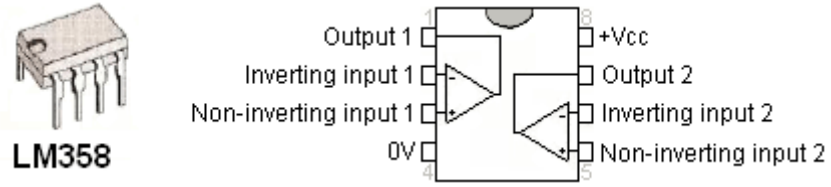


This arrangement will give a 27V 4Hz square wave output through the jack socket. But, Bob Beck's original circuit did not do that. Instead, it was like this:



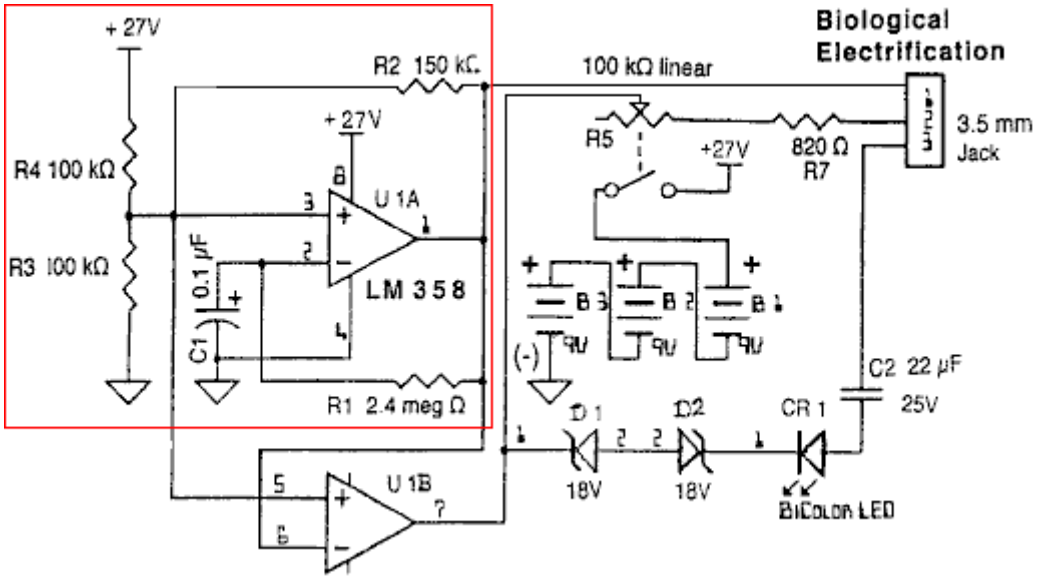
Here, a relay operates two change-over switch contacts which are used to reverse the battery bank contacts four times per second. That is different from just producing a positive-going square wave voltage between the two output terminals. If you were to consider a resistor connected across the output socket, then with the relay switching, the direction of the current reverses four times per second, but with the square wave, while it starts and stops four times per second, the direction of the current is always the same and there is no reversal of direction.

As Bob wanted to avoid using a relay which clicks four times per second all the way through the two-hour treatment described in chapter 11 and in the "Take Back Your Power" pdf on the http://www.free-energy-info.tuks.nl/ web site, he redesigned the circuit using the very impressive LM358/A integrated circuit:



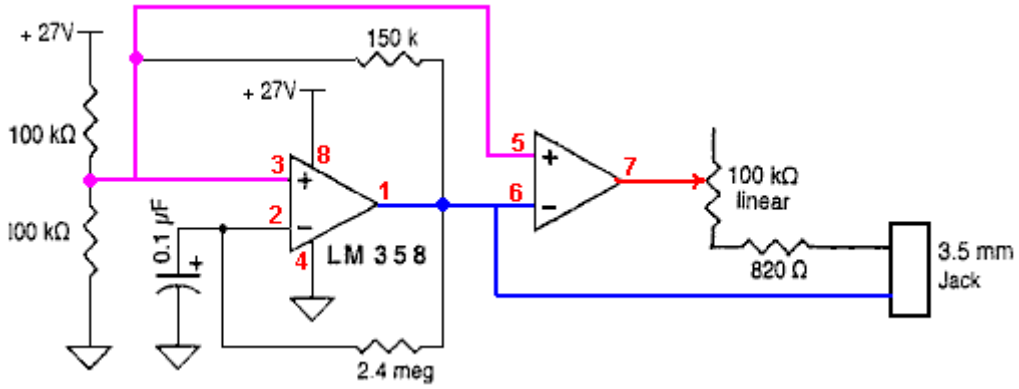
This chip draws only half of one milliamp, has two very high-gain operational amplifiers and can operate with a wide range of supply voltages. It is also inexpensive.

Bob displays the circuit as:



Bob states that the first section acts as a 4Hz square-wave signal generator, the frequency being controlled by the 2.4M resistor “R1” and the 100nF capacitor “C1”. The data sheet for the LM358 states that the output voltage swing is between zero volts and 1.5V less than the supply voltage “Vcc” (which is +27V in this case). That implies that, as would be expected, the pin 1 output voltage from the first stage will switch sharply from 0V to +25.5V and sharply back again, four times per second.

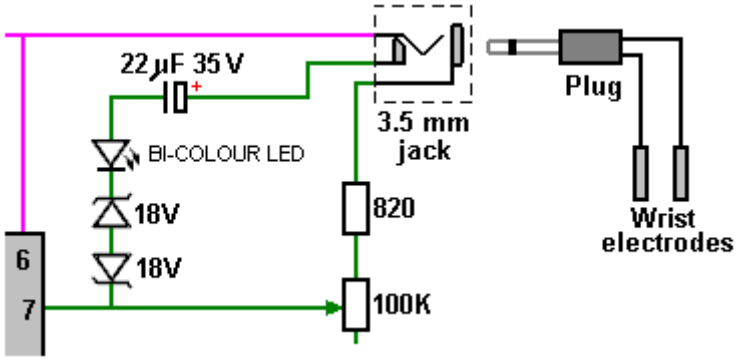
It is difficult to follow the circuit as it is drawn, so it might be a little easier to follow when drawn like this:



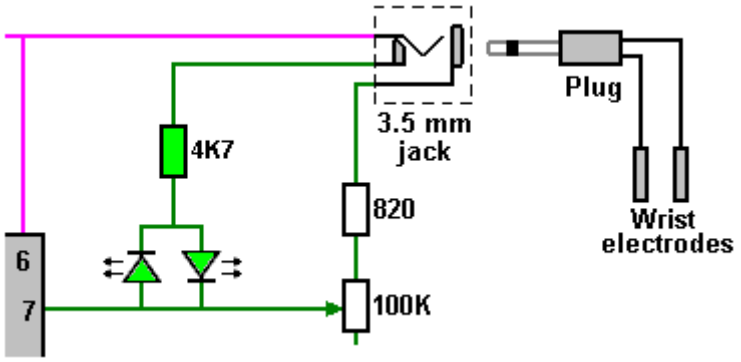
The output from the first amplifier inside the LM358 package is on pin 1 and it can supply a large amount of current (if a large current is ever needed). That output goes straight to one of the jack socket connections. It also goes the pin 6 input of the second amplifier inside the chip and that causes the high-power output of that amplifier on pin 7 to be the opposite of the pin 1 voltage. When pin 1 goes high to +25.5 volts, then pin 7 goes low, to about zero volts. That output is also fed to the other jack socket connection, placing 25.5 volts across the electrodes when they are plugged in to the jack socket.

When the oscillator circuitry connected to the first amplifier causes the voltage on pin 1 to go low, then the output on pin 7 inverts it and so it goes to +25.5 volts. You will notice that while the overall voltage of 25.5 volts is applied again to the jack socket, the polarity is now reversed, achieving what the relay circuit does (although 1.5 volts is lost in the process). This is a neat solution.

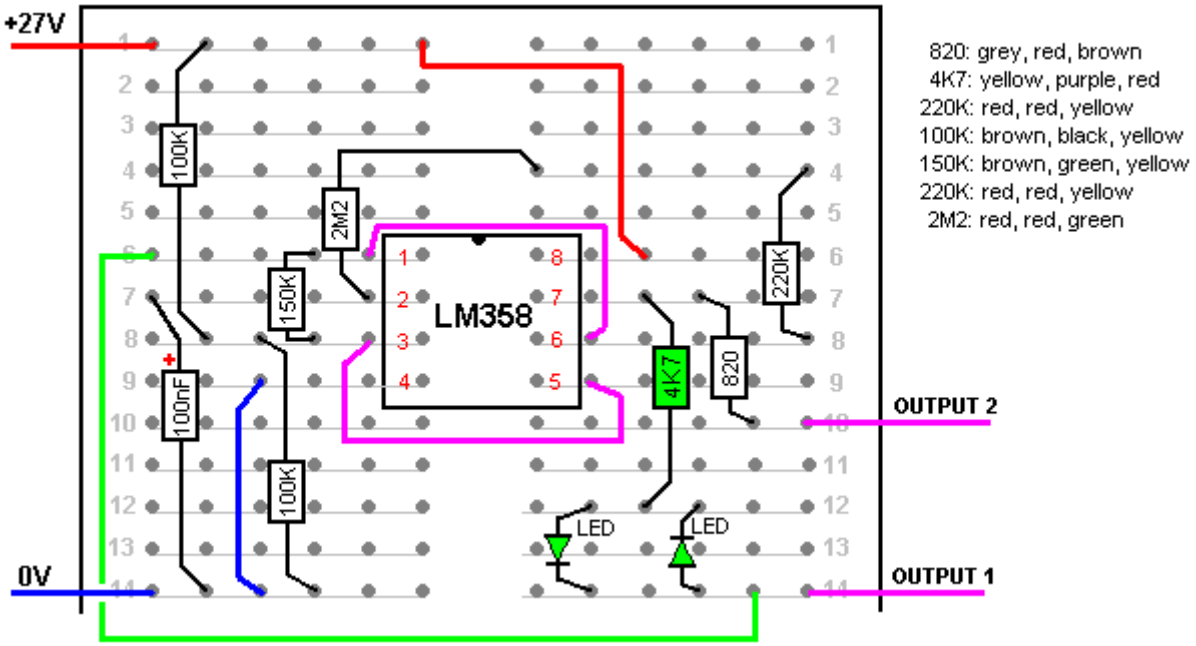
Bob uses a two-colour LED to confirm that the circuit is working correctly before the electrodes are plugged in. He chooses to do it this way:



The two 18V zener diodes drop off 18.7 of the 25.5 volts as one will be forward biased dropping 0.7 volts and the other reversed biased, dropping off 18 volts. That leaves a 7V drop for the LED, which is a bit excessive, so Bob says that he uses a capacitor to limit the current. As there is already an 820 ohm resistor in the LED current path through the socket, the capacitor is not needed. The variable resistor need to be set to it's minimum resistance by rotating it's shaft fully clockwise so that it does not affect the LED brightness as the zeners also show when the battery voltage has dropped as there will no longer be sufficient voltage to light the LED brightly, indicating that the batteries need to be replaced (or recharged if they are rechargeable batteries). When testing the circuit, an alternative to the two zeners is to use a 4.7K resistor and if a bi-colour LED is not to hand, then two ordinary LEDs can be used back to back like this:

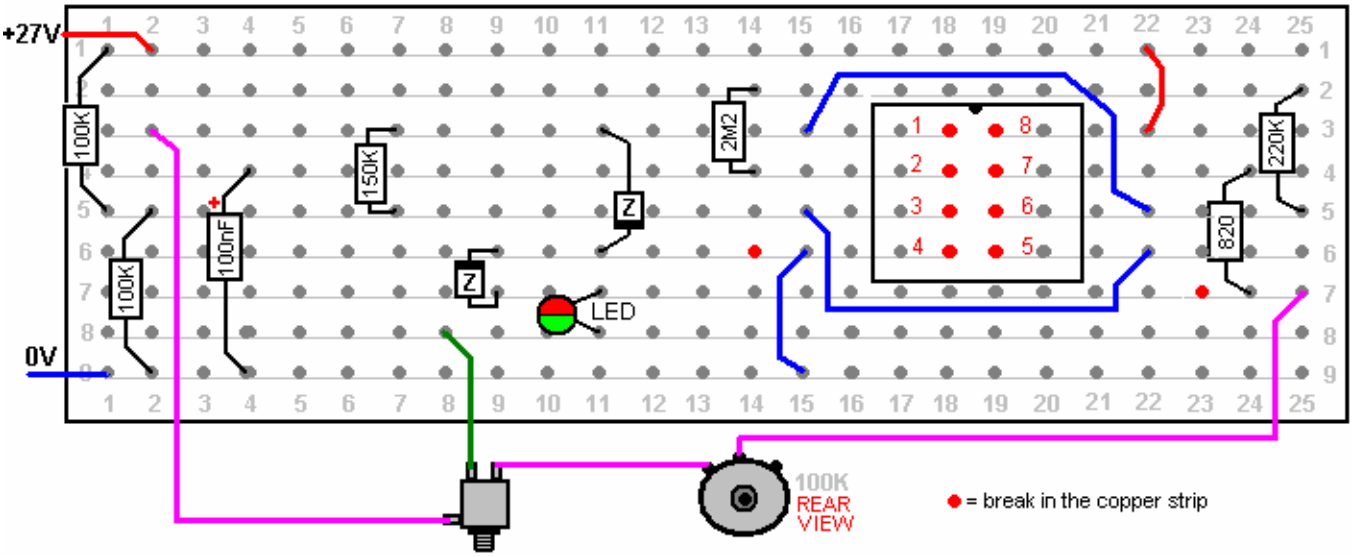


With this arrangement, the two LEDs flash alternately. In any circuit, a capacitor with a higher voltage rating can always be used if the capacitance values are the same. The Beck external circuit is completed through the body of the user, so there is just one electrode connected to each side of the output jack socket. A possible plug-board layout is:



The 4.7K resistor and LEDs are only on the board for testing purposes and when the circuit is built in permanent form, then the LED chain connects to pin 1 of the jack socket so that the LEDs are disconnected during the two hours of daily treatment recommended when using the device.

One stripboard layout using the standard 9-strip 25-hole board and incorporating the two 18V zener diodes for voltage sensing is:



When using a Beck device, it is very important to pay attention to the precautions which Bob sets out. These are in his “Take Back Your Power” pdf document: http://www.free-energy-info.tuks.nl/Beck.pdf which includes the following, which, while it refers to treatment to deal with HIV, presumably applies to all treatments with his device:

EXPANDED INSTRUCTIONS FOR EXPERIMENTAL / THEORETICAL HIV BLOOD NEUTRALIZATION

HYPOTHETICAL PROTOCOLS FOR EXPERIMENTAL SESSIONS

Revision March 20, 1997. Copyright 8 1991/1997 Robert C. Beck

PRECAUTIONS: Do NOT use wrist to wrist current flow with subjects who have cardiac pacemakers. Any applied electrical signals may Interfere with 'demand' type heart pacers and cause malfunction. Single wrist locations should be acceptable. Do NOT use on pregnant women, while driving or using hazardous machinery.

Users MUST avoid Ingesting anything containing medicinal herbs, foreign or domestic, or potentially toxic medication. nicotine, alcohol, recreational drugs. laxatives, tonics. and certain vitamins etc., for one week before starting because blood electrification can cause electroporation which makes cell membranes pervious to small quantities of normally harmless-chemicals in plasma. The effect Is the same as extreme overdosing which might be lethal. See *Electroporation: a General Phenomenon for Manipulating Cells and Tissues;* J.C. Weaver, Journal of Cellular Biochemistry 51:426-435 (1993). Effects can mimic increasing dosages many fold. Both the magnetic pulsar and blood purifier cause electroporation.

Do NOT place electrode pads over skin lesions, abrasions, new scars, cuts, eruptions, or sunburn. Do NOT advance output amplitude to uncomfortable levels. All subjects will vary. Do NOT fall asleep while using. The magnetic pulser should be safe to use anywhere on body or head.

Avoid ingesting alcohol 24 hours before using. Drink an 8 oz. glass of distilled water 15 minutes before and immediately following each session end drink at least four additional glasses daily for flushing during 'neutralization' and for one week thereafter. This Is imperative. Ignoring this can cause systemic damage from unflushed toxic wastes. When absolutely essential drugs *must* be ingested, do so a few minutes *after* electrification then wait 24 hours before next session.

If subject feels sluggish, faint, dizzy, headachy, light-headed or giddy, nauseous. bloated or has flu-like symptoms or rashes after exposures, reduce pulsing per session and/or shorten applications of electrification. Drink more water-preferably ozonized -to speed waste oxidation and disposal. Use extreme caution when treating patients with impaired kidney or liver function. Start slowly at first like about 20 minutes per day to reduce detoxification problems.

To avoid shock liability, use batteries only. Do NOT use any line-connected power supply, transformer, charger, battery eliminator, etc. with blood clearing device. However line supplies are OK with well-insulated magnetic pulse generators (strobe lights).

Health professionals: Avoid nicotine addicts, vegans, and other unconsciously motivated death-wishers and their covert agendas of 'defeat the healer'. Tobacco, the most addictive (42times more addictive than heroin) and deadly substance of abuse known, disrupts normal cardiovascular function. True vegetarian diets are missing essential amino acids absolutely necessary for the successful rebuilding of AIDS-ravaged tissues. Secondary gains (sympathy / martyrdom, work avoidance, free benefits, financial assistance, etc.) play large roles with many AIDS patients. “Recovery guilt” as friends are dying has even precipitated suicide attempts masked as 'accidents'. Avoid such entanglements, since many have unconscious death wishes.

SUPERIOR ELECTRODES: Excellent, convenient and vastly superior electrodes, reusable indefinitely can be made by butt-soldering lead wires to ends of 1” long by 3/32” dia. blanks cut from type 316 stainless steel rods available from welding supply stores (Cameron Welding Supply. 11061 Dale Ave., Stanton, CA 90680). Use 'Stay Clean' flux before soldering (zinc chloride/hydrochloric acid). Shrink-insulate TWO tight layers of tubing over soldered joints to prevent flexing/breaking and lead/copper ions from migrating. Wrap three or four turns of 100% cotton flannel around rods. Spiral-wrap with strong thread starting from wire side to end, tightly pinch cloth over the rod’s end so as to leave no metal exposed by wrapping 6 or 7 turns of thread TIGHTLY just off end of rod, then spiral wrap back to start and tie tightly with four knots then cut off excess cloth at end close to pinch -wraps. Treat end windings and knots with clear fingernail polish or Fray Check®(fabric & sewing supply stores) to prevent ravelling. Soak in a strong solution of sea salt (not table salt) containing a little wetting agent like Kodak Photo Flow, ethylene glycol, or 409 kitchen cleaner. Add a few drops of household bleach, sliver colloid, etc., for disinfectant. Store solution for reuse. Tape soaking-wet electrodes tightly over pulse sites with paper masking or Transpore™ tape or with 1”wide stretch elastic bands with tabs of Velcro ® at ends to fasten. Electrodes should closely conform *precisely* along blood vessels, not skewing ever so slightly over adjacent flesh. This insures better electrical conductivity paths to circulating blood and insures very low internal impedance. (~2000W). Rinse and blot-dry electrodes and skin after each use. NEVER allow bare metal to touch skin as this will cause burns manifested as small red craters that heal slowly. The objective is to get maximum current into blood vessels, not leak it over to adjacent tissue. Therefore never use any electrode wider that about 1/8 inch (3 mm).

ELECTRODE PLACEMENTS: Locate *maximum* pulse position *(NOT* to be confused with acupuncture, reflexology, Chapman, etc. points) on feet or wrists by feeling for maximum pulse on inside of ankle ~1”below and to rear of ankle bone, then test along top centre of instep. Place electrode on whichever pulse site on that foot that feels strongest. Scrub skin over chosen sites with mild soap and water or alcohol swab. Wipe dry. Position the electrodes lengthwise along each left and right wrists blood vessel. Note: with subjects having perfectly healthy hearts and not wearing pacers, it is convenient to use left wrist to right wrist exactly over ulnar arterial pulse paths instead of on feet. Recent (Dec. 1995) research suggests that placing both electrodes over different arteries on the same wrist works very well (see pg. 7), avoids any current through heart, and is much more convenient and just as effective. An 8” long, 1” wide elastic stretch-band with two 1.5” lengths of 3/4” wide Velcro ® sewn to ends of opposite sides makes an excellent wrist band for holding electrodes snugly in place. With electrode cable unplugged, turn switch ON and advance amplitude control to *maximum*. Push momentary SW. 2 'Test' switch and see that the red and green light emitting diodes flash alternately. This verifies that polarity is reversing about 4 times per second (frequency is *NOT* critical) and that batteries are still good. When LED's don't light replace all three 9V batteries. Zener diodes will extinguish the LEDs when the three 9V battery’s initial 27V drops below 18V after extended use. Never use any electrode larger than 1.125” (28 mm) long by 1/8” wide to avoid wasting current through surrounding tissue. Confine exactly over blood vessels only. Apply drops of salt water to each electrode’s cotton cover ~every 20 minutes to combat evaporation and insure optimum current flow. Later devices are solid-state, use only three batteries and no relays, and are much smaller.

Now rotate amplitude control to *minimum* (counter-clockwise) and plug In electrode cable. Subject now advances dial slowly until he feels a “thumping” and tingling. Turn as high as tolerable but don’t advance amplitude to where It is ever uncomfortable. Adjust voltage periodically as he adapts or acclimates to current level after several minutes. If subject perspires, skin resistance may decrease because of moisture, so setting to a lower voltage for comfort is indicated. Otherwise it is normal to feel progressively less sensation with time. You may notice little or no sensation at full amplitude immediately, but feeling will begin building up to maximum after several minutes at which time amplitude must be decreased. Typical adapted electrode-to-electrode impedance is on the order of 2000W. Typical comfortable input (to skin) is ~3mA, and maximum tolerable input (full amplitude) is ~7mA but this 'reserve' margin although harmless is unnecessary and can be uncomfortable. Current flowing through blood Is very much lower than this *external* input because of series resistance through skin, tissue and blood vessel walls, but 50 to 100 µA through blood is essential.

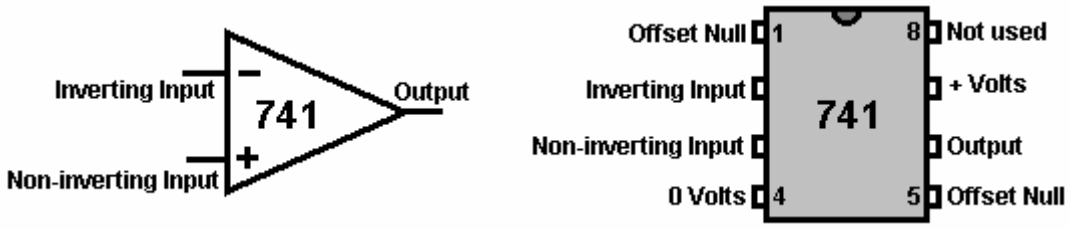
Apply blood neutralizer for about 2 hours daily for ~2 months. Use judgment here. The limiting factor is detoxification. Carefully monitor subject’s reactions (discomfort, catarrh, skin eruptions, weeping exudites, rashes,

boils, carbuncles, coated tongue, etc.). With very heavy infections, go slower so as not to overload body's toxic disposal capability. With circulation-impaired diabetics, etc., you may wish to *extend* session times. Again, ***have subject drink lots of water***. Recent changes in theoretical protocol being currently tested suggest following up the three weeks of treatments with a 24 hours per day (around the clock) continuous electrification of blood for two days to deal a knockout blow to the remaining HIV's 1.2 day life cycle. (A. Perelson; Los Alamos Biophysics Group, Mar. 16, 1996 "Science" Journal.) Remember to remoisten electrodes regularly. If you absolutely ***must*** ingest prescription drugs, do so immediately ***after*** turning off instrument and allow 24 hours before next treatment to let concentrations in blood plasma decay to lower levels.

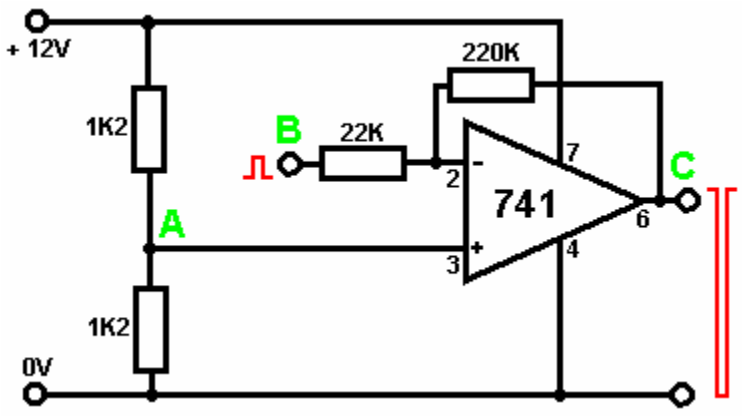
Remember, if subjects ever feel sleepy, sluggish, listless. nauseous, faint, bloated, or headachy, or have flu-like reactions they may be neglecting sufficient water intake for flushing toxins. We interpret this as detoxification plus endorphin release due to electrification. Let them rest and stabilize for about 45 minutes before driving if indicated. If this detoxing becomes oppressive, treat every *second* day. Treating at least 21 times should 'fractionate' both juvenile and maturing HIV to overlap maximum neutralization sensitivity windows and interrupt 'budding' occurring during HIV cells' development cycles. Treatments are claimed to safely neutralize many other viruses, fungi, bacteria, parasites, and microbes in blood. See patents US 5,091,152 US 5,139,684 US 5,188,738 US 5,328,451 and others as well as numerous valid medical studies which are presently little known or suppressed. Also. ingesting a few oz. of about 5 parts per million of silver colloid solution daily can give subjects a 'second intact immune system' and minimise or eliminate opportunistic infections during recovery phase. This miracle substance Is pre-1938 technology, and unlike ozone is considered immune from FDA harassment. Silver colloid can *easily* be made at home electrolytically in minutes and in any desired quantities and parts per million strength for under 14 cents per gallon plus cost of water. It is ridiculous to purchase it for high prices. Colloid has no side effects, and is known to rapidly eliminate or prevent hundreds of diseases. Sliver colloids won't produce drug resistant strains as will all other known antibiotics. No reasonable amount can overdose or injure users either topically, by ingestion, or medical professional injection.

The 741 Operational Amplifier.

An important and very useful group of Integrated Circuits is the "Operational Amplifier" or "op-amp" group. These devices have a very high gain, an 'inverting' input and a 'non-inverting' input. There are many op-amps but we will look at just one popular type called the "741" which has an 'open-loop' gain of 100,000 times. All operational amplifiers work in the same way in theory. The way they operate in a circuit is controlled by the external components attached to them. They can operate as inverting amplifier, a non-inverting amplifier (i.e. a 'buffer'), a comparator, an astable multivibrator, and various other things. The symbol and connections for a 741 op-amp are:



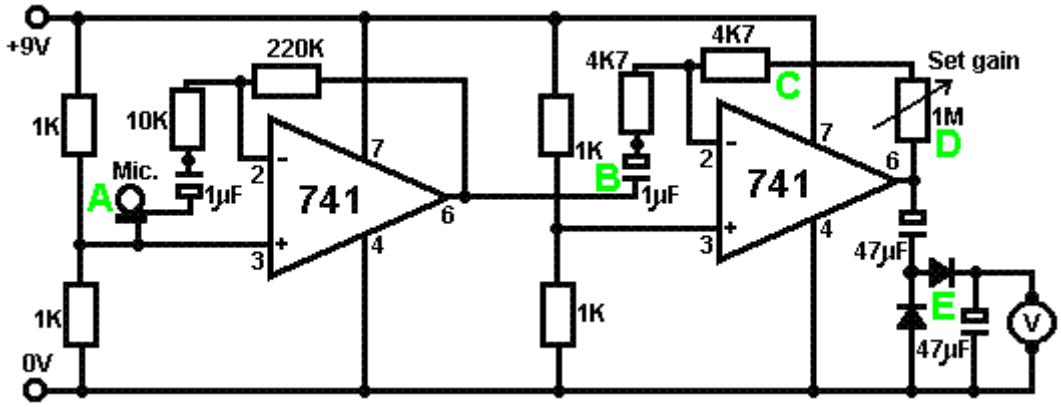
We can connect the 741 chip to act as an amplifier with any set gain level that we choose:



Here, the gain is set by the ratio of the 220K resistor to the 22K resistor. This circuit has a gain of 10 times, so the input signal at point 'B' will generate an output signal at point 'C' which is ten times larger, **provided** that the output signal does not approach the battery voltage. If it does, then clipping will occur with the top and the bottom of the output waveform chopped off at about a volt away from the battery voltage levels, approximately 1 Volt and +11 Volts in this example.

Operational amplifiers are generally designed to operate from a dual power supply. In the above example, the supply would be created by using two 6 Volts batteries instead of one 12 Volt battery. To avoid the inconvenience of this, a mid-point voltage is generated at point 'A' by using two equal resistors in series across the battery. This gives a central voltage of +6 Volts which is fed to the IC.

This circuit can be used in many applications. Here is a circuit for a meter to measure sound intensity:



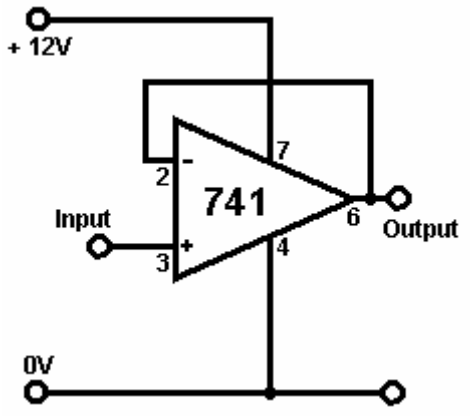
This circuit is two copies of the previous circuit. Each 741 chip has a reference voltage of half the supply voltage created by a voltage-divider pair of 1K resistors. This voltage is fed to pin 3 of the chip, which is the non-inverting input.

At point 'A', a microphone or small loudspeaker is used to generate a signal voltage when sound reaches it. This voltage is fed to the 741 op-amp via a 1 microfarad blocking capacitor. This passes the audio signal through while blocking the +4.5 Volts DC on pin 3. The first 741 has a gain of 22, set by the 10K and 220K resistors (220/10 = 22).

Point 'B' then receives an audio signal 22 times larger than the signal produced by the microphone. This signal is still quite small, so the second 741 boosts it further. The gain of the second 741 is variable and depends on the resistance set on the 1M variable resistor. If the variable resistor is set to zero ohms, then the gain of the second 741 will be controlled by the 4K7 resistor at point 'C' alone and so will be 1 (4.7/4.7 = 1). If the variable resistor is set to its maximum value, then the gain of the second 741 will be some 214 (1,004,700/4,700 = 213.8).

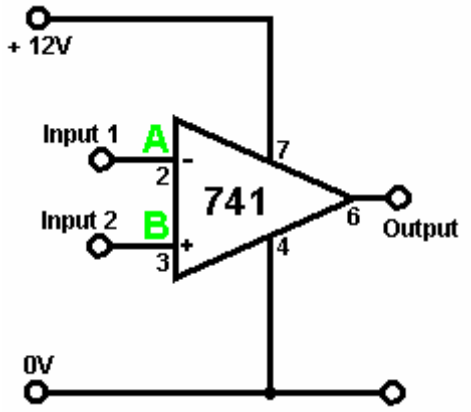
The two op-amps together have a combined gain which ranges from 22 to 4702. The amplified audio signal arrives at point 'D' and it can be adjusted to a respectable value. This alternating voltage is now rectified via the diodes at point 'E' and it builds up a DC voltage across the 47 microfarad capacitor there. This voltage is displayed on a voltmeter. The result is that the voltmeter shows a reading directly proportional to the sound level reaching the microphone.

The 741 can be wired as a **buffer**. This is the equivalent of an emitter-follower circuit when using transistors. The set up for the 741 is:



Difficult circuit - huh! Are you sure you can afford all the extra components? This circuit utilises the full gain of the 741 chip. The output follows the input waveform exactly. The input requires almost no current, so the circuit is described as having a 'high input impedance'. The output can drive a serious load such as a relay, so the circuit is described as having a 'low output impedance'.

The 741 chip can be wired to act as a comparator. This is the circuit:



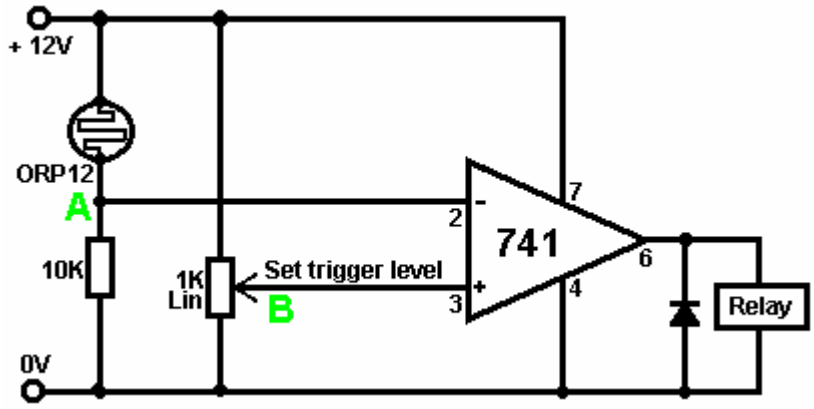
Are you sure you are up to such a difficult circuit? Bit complicated - huh! This is the basic operational form for an operational amplifier.

If the voltage at point 'A' is higher than the voltage at point 'B' then the output goes as low as it can go, say 1 or 2 volts.

If the voltage at point 'A' is lower than the voltage at point 'B' then the output goes as high as it can go, say 10 volts or so.

Having seen how transistor circuits work, you should be able to understand why the 741 chip circuitry (which is a transistor circuit inside the 741 package) needs some voltage inside the supply rails to provide an efficient high-current output drive.

Here is a 741 version of the light-operated switch:

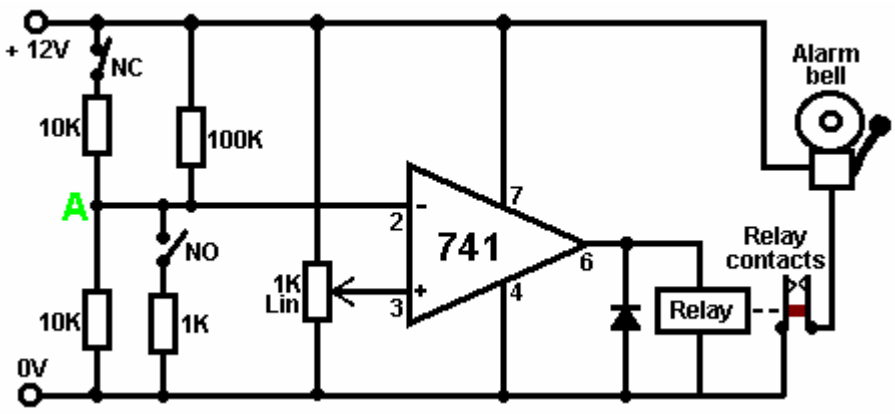


This circuit is set up as evening falls. We want the relay to have minimum voltage across it in daylight, so the voltage at point 'A' needs to be higher than the voltage at point 'B'. As the 1K variable resistor is across the supply voltage, its slider can be set to any voltage between 0 Volts and +12 Volts. To make this easy to do, we choose a 'linear' variable resistor as the logarithmic variety would be hard to adjust in this application. With the 'linear' version, each 1 degree of rotation of the resistor shaft causes the same change in resistance, anywhere along the range. This is not the case for the logarithmic variety.

Anyhow, we adjust the variable resistor downwards until the relay voltage drops to a minimum. When the light level has fallen to the level at which we wish the circuit to trigger, we adjust the variable resistor to make the relay click on. The 741 chip has a very rapid output voltage swing when the input voltages swap over, so the relay switching will be decisive. The switching can be made even more positive by adding a resistor between the output and point 'B'. This acts like a Schmitt trigger when switching occurs by providing some additional positive feedback, lifting the voltage at point 'B'.

If you wish the circuit to trigger on a rising light level, just swap the positions of the 10K resistor and the ORP12 light-dependent resistor. The same circuit will operate as a temperature sensing circuit by substituting a 'thermistor' (which is a temperature-dependent resistor) for the ORP12.

If we would like the circuit to act as a burglar alarm, we could use the same circuit like this:



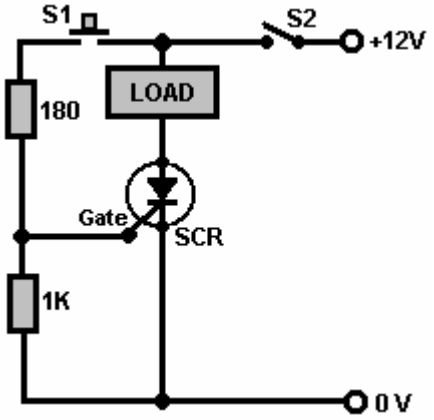
The circuit is still controlled by the voltage at point 'A'. Under normal circumstances, this voltage will be near +6 Volts (produced by the two 10K resistors and the 100K resistor). The upper switch marked 'NC' for 'Normally Closed', represents a chain of, say, magnetic switches attached to doors and windows. If any of these are opened, then the voltage at point 'A' will be dictated by the lower 10K resistor in series with the 100K resistor. This will cause the voltage at 'A' to fall instantly to a low value, triggering the circuit.

The 'NO' switch ('Normally Open') represents one or more pressure-operated switches under carpets or rugs and/or switches which get brushed when doors are swung open, etc. These switches are wired in parallel across each other and if any of them is closed for even a millionth of a second, the voltage at point 'A' will be pulled down by the 1K resistor and the circuit will be triggered.

The circuit can be latched on in any one of a variety of ways. One relay contact can be used to hold the relay on or hold the voltage at 'A' low. A transistor can be wired across the relay to hold the circuit on, etc. etc. If this is done, the circuit will remain in its triggered state until the supply voltage is interrupted. You might prefer to use a 555 chip to limit the length of time the alarm sounds to three minutes or so.

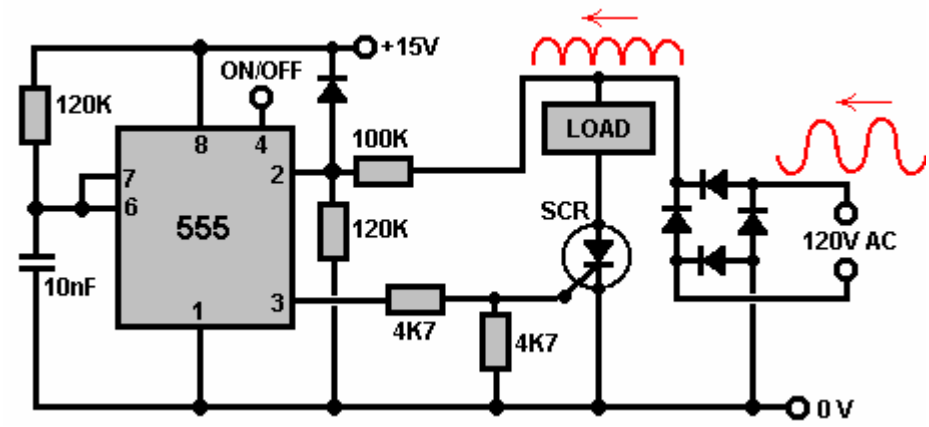
The SCR and Triac.

An alternative to using a relay or semiconductor latch is to use a **S**ilicon **C**ontrolled **R**ectifier usually referred to as an 'SCR' or 'Thyristor'. This device is normally "off" with a very high resistance to current flow. If it is switched on by applying a voltage to its Gate connection, it stays continuously on until some external device stops current flowing through it. The following circuit shows how it operates:



When the voltage is first applied to the circuit by closing switch S2, the SCR is in its OFF state so no current is supplied to the load. If the press-button switch S1 is pressed, a current is fed into the Gate of the SCR, turning it ON. When switch S1 is allowed to open, the SCR remains in its ON state and it will stay that way until the current through it is cut off. Opening switch S2 cuts off the current to the load and the SCR returns to its OFF state. A very valid question would be: "Why have an SCR at all and just turn the load on and off with switch S2?". The answer is that switch S1 might be the under-carpet pressure pad of a burglar-alarm and it might be operated some hours after switch S2 was closed to activate the alarm system. Stepping off the pressure pad does not stop the alarm sounding.

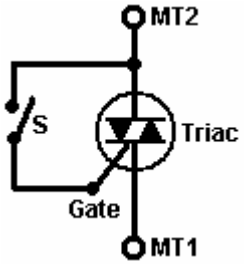
While this sort of DC latching action is useful, it is more common for an SCR to be used in an AC circuit. For example, take the circuit shown here:



The 120 volt AC supply coming in from the right hand side, is converted to positive-going sine-wave pulses by the diode bridge. This pulsing voltage is applied to the Load/SCR path. If the voltage at pin 3 of the 555 chip is low, then the SCR will remain OFF and no current will be fed to the load device. If the voltage on pin 3 goes high **and** the voltage applied to the Load/SCR chain is high, then the SCR will be switched ON, powering the load until the pulsing voltage drops to its zero level again some 1/120 of a second later.

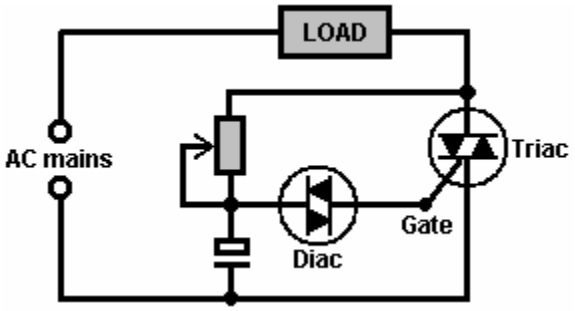
The 555 chip is connected to form a monostable multivibrator and the timing components (the 120K resistor and the 10nF capacitor) cause it to output a 1 millisecond pulse which is long enough to trigger the SCR into its ON state, but short enough to have finished before the mains pulse reaches its zero-voltage level again. The 555 chip is triggered by the rising mains voltage being passed to its pin 2 through the voltage-divider 100K and 120K pair of resistors, and that synchronises it with the AC waveform. Pin 4 of the 555 chip can be used to switch the load power on and off.

In the circuit shown above, the diode bridge is needed to convert the incoming AC waveform to pulsing DC as shown in red in the diagram, as the SCR can only handle current flowing in one direction. The AC load equipment works just as well with the pulsing DC as with a full blown AC waveform. A better semiconductor construction is the 'Triac' which acts like two SCR devices back-to-back in a single package. It is shown like this in circuit diagrams:



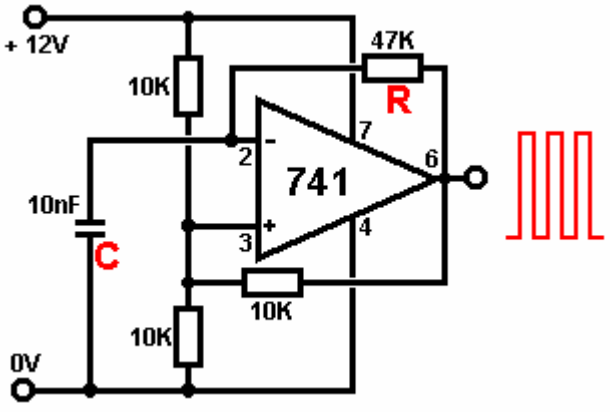
There are three connections to the device: Main Terminal 1, Main Terminal 2 and the Gate. When switch 'S' shown in the diagram is closed, the triac conducts on both positive and negative voltages applied to its MT1 and MT2 terminals. When the switch is open, the device does not conduct at all.

If the external circuit containing switch 'S' is placed inside the device as a permanently closed circuit, then the device becomes a 'Diac' which can be used to trigger a Triac and give a very neat circuit for controlling the power to an item of AC mains equipment as shown here:



Here, the variable resistor/capacitor pair controls the point on the AC waveform that the Triac is triggered and so controls how much of each sinewave cycle is passed to the mains equipment, and so it controls the average power passed to the equipment. A very common use for a circuit of this type is the 'dimmer-switch' used with household lighting.

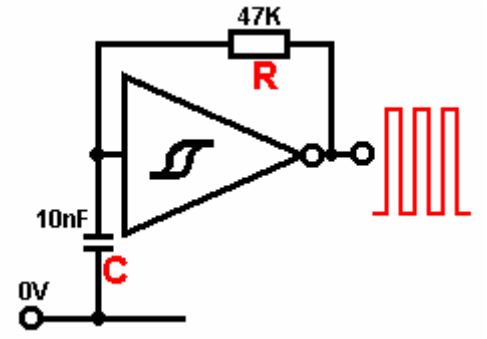
To return now to the 741 chip. The 741 can also be used as an **astable multivibrator**. The circuit is:



The rate of oscillation of this circuit is governed by the Resistor marked '**R**' in the diagram and the capacitor marked '**C**'. The larger the resistor, the lower the rate of oscillation, the larger the capacitor, the lower the rate of oscillation.

When the output goes high, capacitor '**C**' charges up until the voltage on it exceeds the mid-rail voltage on pin 3, at which time the 741 output goes low. The capacitor now discharges through resistor '**R**' until the voltage on it drops below the voltage on pin 3, at which time the output goes high again. The 10K resistor connecting the output to pin 3 provides some positive feedback which makes the 741 act quite like a Schmitt trigger, sharpening up the switching.

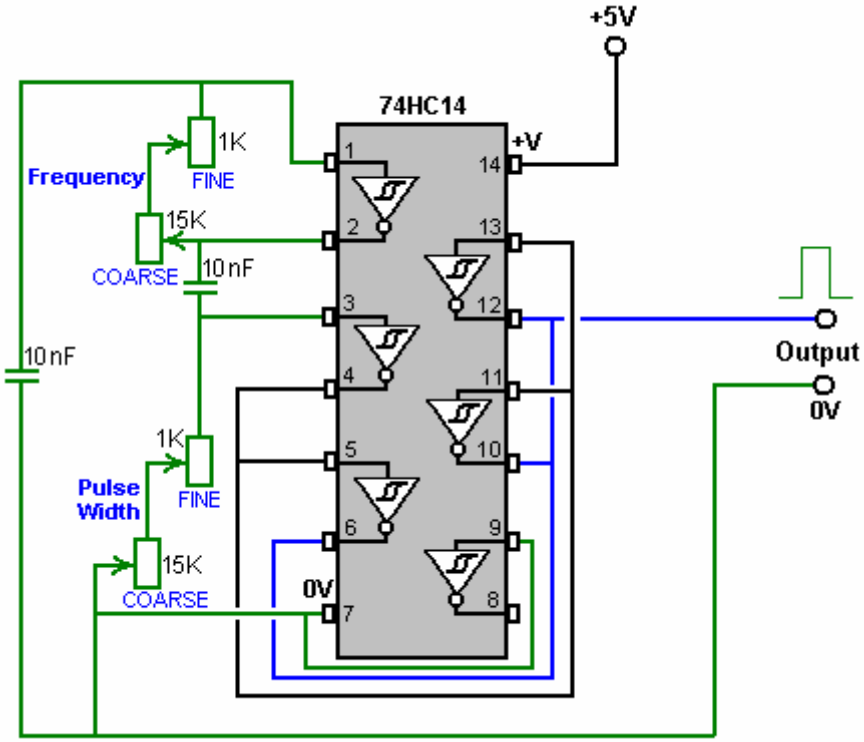
The same arrangement of resistor and capacitor applied to a Schmitt inverter or Schmitt NAND gate causes exactly the same oscillation:



If you would like to see additional ways of using 741 and 555 chips, I can recommend the excellent book "Elementary Electronics" by Mel Sladdin and Alan Johnson ISBN 0 340 51373 X.

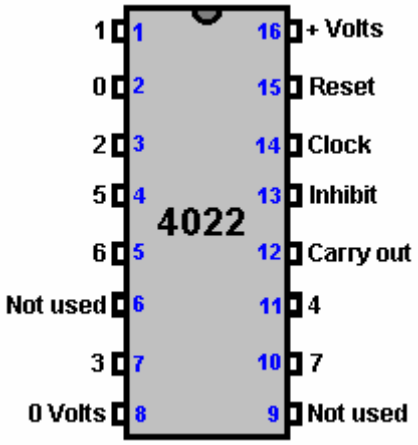
A Hex Inverter Signal Generator.

Here is a very well tested and highly thought of, low-cost oscillator circuit, using a 74HC14 Schmitt inverter chip (or the 40106B higher voltage CMOS chip). It allows fine tuning control of the frequency and the pulse width produced. Three of the inverters are connected together to give a more powerful output current drive:



The 4022 Divide-by-Eight Chip.

One very useful CMOS integrated circuit is the '4022' chip which is a 16-pin 'divide by 8' chip with built-in decoding. The connections are:



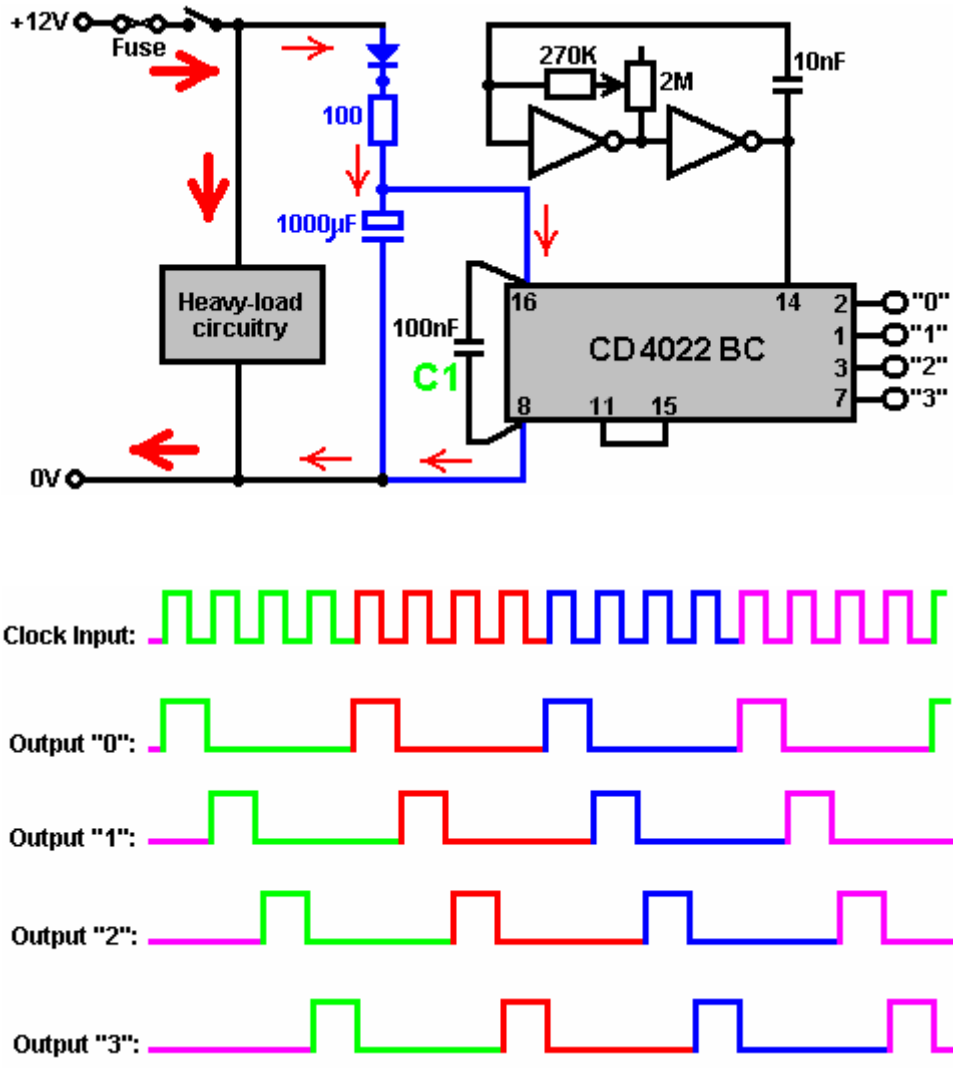
If pin 14 is provided with the output from some variety of astable multivibrator, on the first pulse, this chip sets the "0" output on pin 2 to High while the other outputs are Low. On the next pulse, the "0" output goes Low and the "1" output on pin 1 goes High. On the next pulse, output "1" goes Low and the "2" output on pin 3, goes High. And so on until on the eighth pulse, output "7" on pin 10 goes Low and output "0" goes high again.

The chip can also divide by lower numbers:

For 'Divide by 7' operation, connect pin 10 to pin 15 (this resets the output to '0')
For 'Divide by 6' operation, connect pin 5 to pin 15
For 'Divide by 5' operation, connect pin 4 to pin 15
For 'Divide by 4' operation, connect pin 11 to pin 15
For 'Divide by 3' operation, connect pin 7 to pin 15
For 'Divide by 2' operation, connect pin 3 to pin 15

If you want a 'Divide by 1' circuit, I suggest you cut down on the amount of alcohol you drink.

Here is an illustration of a 'Divide by 4' setup:



There are a number of things to notice in the above diagram. Firstly, the practical arrangements for circuitry have not been stressed before. If the circuitry has a pulsing circuit drawing heavy current, as shown by the thick red arrows, then it should be physically connected to the battery and any low-current circuitry should be further away from the battery. The supply from the battery should have a fuse or circuit breaker and a switch in the line before anything else is connected, so that if any component develops a fault and goes short-circuit, the fuse will blow and prevent any significant problems.

Secondly, it is a good idea to provide the other circuitry with a smoothed power supply as shown by the blue components in the diagram. This minimises the effect if the battery voltage gets pulled down by the pulsing of the

high-current circuitry. The diode (silicon, 1 Amp, 50 V) stops the heavy current circuit drawing current from the large smoothing capacitor. The 100 ohm resistor limits the current into the large capacitor on switch-on and provides a little more smoothing. This circuitry is called “de-coupling” as it de-couples the low current circuitry from the high current circuitry.

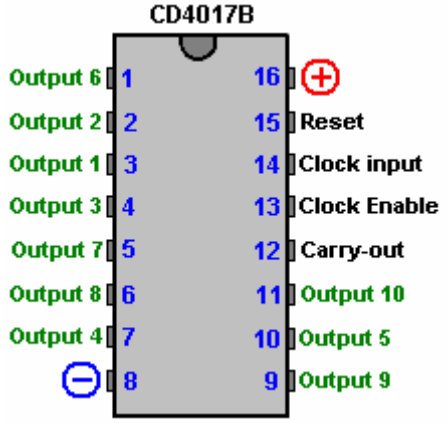
Thirdly, notice capacitor “C1” which is wired physically as close to the power supply pins of the integrated circuit as is possible. If a spike is superimposed on the battery supply, then this capacitor soaks it up and prevents it damaging or triggering the integrated circuit. A spike could be caused by a very strong magnetic pulse nearby as that can induce an extra voltage in the battery wires.

The lower part of the diagram shows the output voltages produced as the clock pulses reach pin 14 of the chip. The positive-going part of the clock signal triggers the change in state of the outputs. If necessary, a positive-going pulse on the reset pin, pin 15, causes output “0” to go high and the other outputs to go low.

The 4017 Divide-by-Ten Chip.

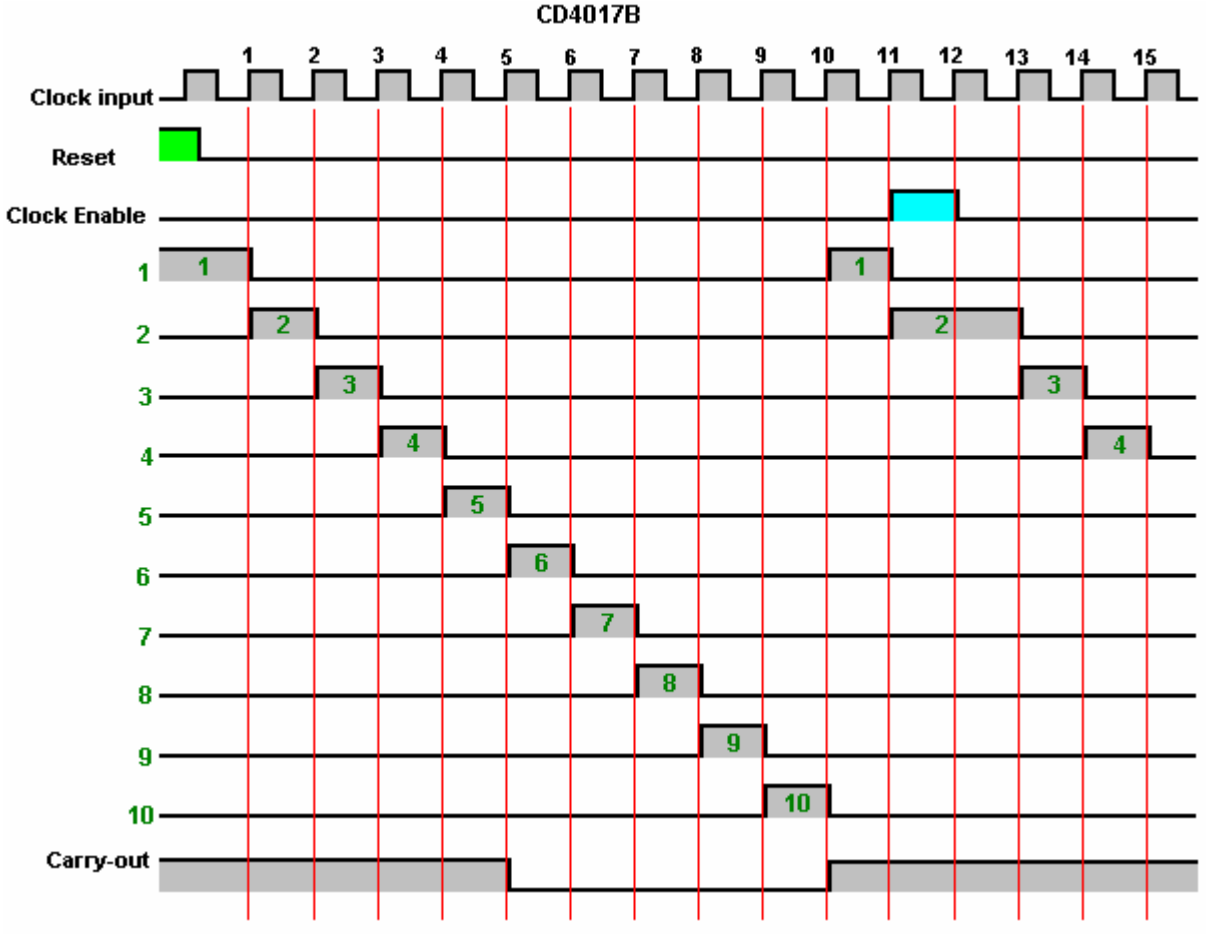
Now, to take this output sequencing a little further. For example, the Charles Flynn magnet motor shown in Chapter 1 needs coils to be powered up, one after the other and only one should be on at any one time. This calls for a circuit which has a large number of outputs. The CD4022BC chip gives up to eight outputs one after the other. The CD4017B chip gives up to ten outputs one after the other but there is no need to be limited by these numbers as more than one chip can be used. If you find this section difficult to understand, then just skip past to the next section as it is not important for you to understand these larger circuits.

The pin connections for the divide-by-ten CD4017B chip is shown here:



While this shows outputs 1 to 10, the manufacturers and some people who draw circuits, prefer to label the outputs as “0 to 9” which correspond to digital displays. In our style of operation, it is easier to think of the ten outputs as being from 1 to 10.

You will notice that there are two pin labels which we have not come across before, namely, the “Carry-out” pin and the “Clock Enable” pin. These allow us to use several of these chips in a row to give a much larger “divide-by” number. The “Clock Enable” pin can be used to block the clock input. The operation is like this:



In this example, the sequence is started by the Reset pin being given a high voltage as shown by the green shading. This pushes the output pin 1 to a high voltage and all of the other outputs to a low voltage and holds those voltages as long as the reset voltage is high.

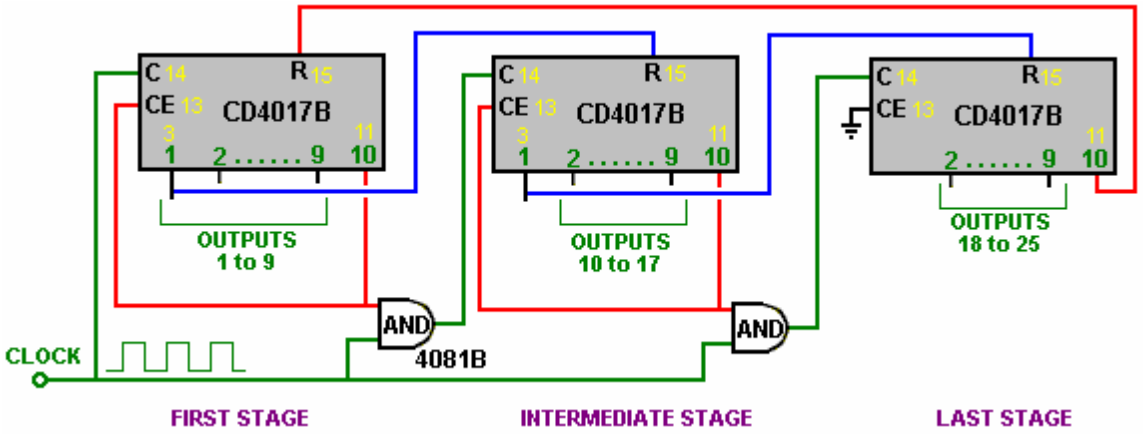
When the Reset voltage drops, the next rising edge of the clock pulse (marked “1” in the diagram) causes the output 1 to go low and output 2 to go high. Each of the successive clock pulses “2” to “9” moves the high voltage steadily along the outputs until output pin 10 is high.

The next clock pulse rising edge (marked “10” in the diagram) starts the sequence again with output 10 going low and output 1 going high again. If nothing changes, then that sequence of output voltage changes will continue indefinitely.

However, in the diagram above, the Clock Enable pin voltage is driven high on clock pulse “11”. Output 2 has just gone high and would have gone low when the rising edge of clock pulse “12” occured, but in this case, the Clock Enable feature blocks the clock pulse and prevents it reaching the rest of the circuitry. This causes the output 2 voltage to stay high as long as the Clock Enable remains high. In this example, the Clock Enable voltage stays high for just one clock pulse, causing the output 2 voltage to be high for twice it’s usual length, and then the sequence continues as before.

A Divide-by-Twenty-five Circuit.

Here is one way to get a large “divide-by” number. This example is divide-by-25 because there is only one ‘intermediate stage’ but there can be any number and each additional one adds another eight outputs to the total:



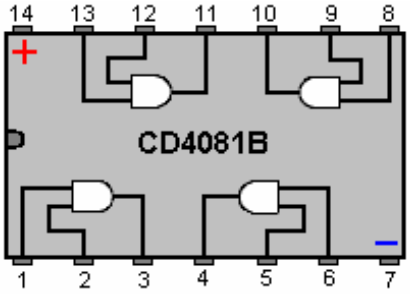
At startup, output 10 of the first stage (which is physical pin 11 of the chip) is at a low voltage. This holds the Clock Enable (pin 13) low, allowing the clock pulses to enter the first stage. Because the output 10 voltage is low, one input to the first AND gate is held low, preventing it from letting the clock pulse flow through it, i.e. the “gate” is closed to through traffic.

The first stage chip then operates as normal, producing outputs 1 to 9 in order as you would expect. The next clock pulse stes the first stage output 10 high, allowing the clock pulses through the first AND gate and holding the Clock Enable (pin 13) high, which in turn locks the output 10 high, dropping the first stage chip out of the operation.

As the output 1 of the first stage is connected to the Reset (pin 15) of the second chip, it will have been cleared and it’s output 1 set high, which in turn Resets the third chip and closes the second AND gate. So, when the first pulse gets through to the second chip, it pushes it from state 1 to state 2 where the output 2 goes high. For that reason, output 1 of the second chip is not one of the outputs which can be used by whatever following circuitry you choose to connect to this system. Consequently, only eight of the ten outputs of the second chip are available as counter outputs. That is, outputs 1 and 10 are taken up in passing the switching sequence between the various chips in the chain.

The same applies to all following chips in the chain, each extra chip adding up to eight extra sequential outputs. On the final stage chip, if you connect the red Reset wire (which goes back to fire up the first chip again) to output 9 instead of output 10 of the final chip, then you get a divide-by-24 result.

If the Reset is taken from output 8 of the final chip, then you get a divide-by-23 result, and so on. Using this method, you can have a divide-by circuit for any number you want. These chips are very popular and so their cost is low, making the entire circuit cheap to make. The pin connections for the AND gates is shown here:



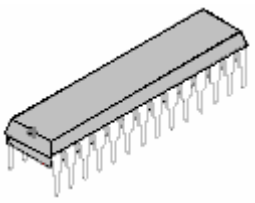
The PIC Revolution.

Over the years, there have been advances in the way that circuitry can be put together, prototypes built and tested. Initially, "valves" or "vacuum tubes" were used and circuits required a good deal of electrical power in order to operate. Mechanical vibrators or "reeds" were used to generate the switching needed to convert DC into AC. Then the transistor became widely available and the transistor replaced the mechanical vibrator reed, the circuit being called an "astable multivibrator" and comprising of two transistors wired back to back (as described in chapter 12). Then came the digital integrated circuit with it's "NOR gates" which could also be wired back to back to make a multivibrator. This was done so often that a special integrated circuit called the "555 chip" was designed to do the job all on its own. That chip has been a tremendous success and is now found in all sorts of different circuits, being very easy to use, very robust and very cheap. Surprisingly, the dominant position of the "555" chip is being challenged by a completely different type of chip, one which is essentially, a computer on a

single chip, and which is called a "PIC controller".

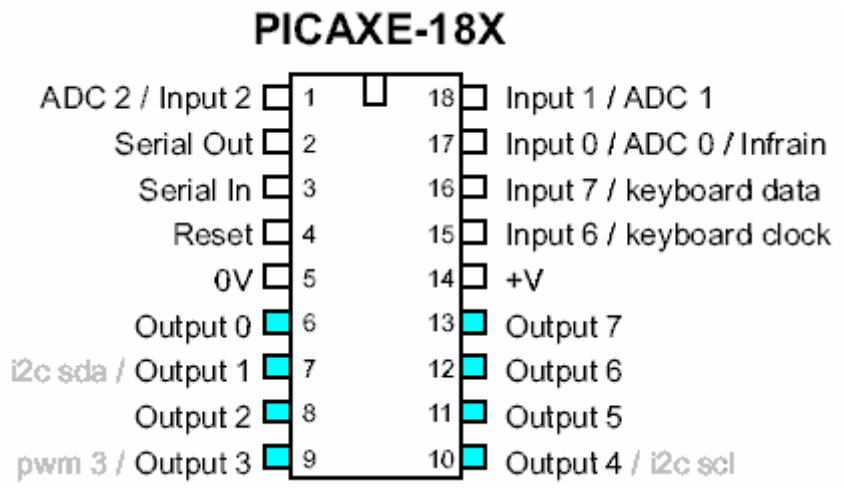
This new type of chip is not expensive, is easy to use, and can be changed to perform a different task in just a few seconds. It can perform timing tasks. It can act as a multivibrator. It can act as a "Divide-by-N" chip. It is a very impressive chip which is very useful. The reason that I mention it here is because it is at the heart of the fastest working Tesla Switch research forum around (the "energetic forum" group). The chip is something you need to know about as it will certainly take over more and more circuit applications in the coming years.

There is a whole family of these processor chips, but I will select just one for this description, and that will be the one being used by the "energetic forum" members, and I have to thank Jeff Wilson for his help in describing this circuitry, the programming and the methods which he uses.



First, however, some information on this new design of chip and the methods used with it. The one used by Jeff is called the "PICAXE-18X" and it looks like the chip shown here. From which you can see, it looks just like any other chip, although with eighteen pins. The powerful performance comes from the way that it operates. You are probably familiar with the "555" chip and understand that it operates by changing the voltage on just one of it's pins (pin 3) the output pin, from a low voltage to a high voltage. The PIC chip can do that as well, but even better still, it has more than one output pin and it can alter the voltage on any of those pins to either a high or a low voltage and it can do that in any order and with any timing that you choose. This makes it a very versatile chip indeed and one which is very well suited to be the central controller for a Tesla Switch test environment.

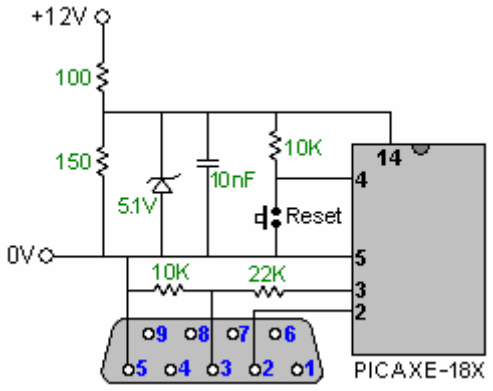
The chip is used by wiring it into a circuit in the same sort of way that a 555 chip would be used, except that the PIC has it's own internal timing clock and can operate in intervals of one thousandth of a second, that is, one millisecond.



The top eight pins are for making the chip work. The next two are for providing the chip with electrical power. The bottom eight pins are separate outputs, any one of which can operate switches, timers, etc., just as the output from a 555 chip can. Having been named by computer people, instead of the eight output pins being numbered from 1 to 8 as any rational person would do, they have numbered them from 0 to 7.

The voltage on those output pins will be either High or Low. PIC switching can be used with a wide range of different free-energy designs. The PIC chip is generally supplied with a socket, a connecting cable and a program for feeding instructions into the chip. The feed is generally from an ordinary PC. The programming instructions are very simple and anyone can learn how to use them in just a few minutes.

So let's look at a circuit which has been used by Jeff when he tests prototype circuitry. The first part of the circuit is for connecting the standard PC socket to the PIC chip and it looks like this:



A standard 9-pin computer socket has it's pin 2 connected to the PIC's pin 2, pin 3 connected to the PIC's pin 3 via a 10K / 22K voltage divider resistor pair (which lowers the incoming signal voltage), and pin 5 is connected to the PIC's pin 5. That is all that's needed to feed information into the PIC chip.

The chip is supplied from a 12-volt battery but as it needs a 5-volt supply, the 100 / 150 ohm (2 watt) resistor pair is used to drop the 12 volts down to about 7 volts and then the 5.1-volt zener diode clamps the voltage at 5.1 volts, which is just what the chip needs. The tiny 10 nF (0.01 microfarad) capacitor is there to trap any voltage spikes should any be picked up from some outside influence. Finally, the press-button switch used to short between pins 4 and 5 is used to wipe out the program inside the PIC, ready for a new program to be loaded.

The actual programming is not difficult and the feed into the chip is handled by the program supplied with the chip and which is run on your home computer. Let's take an example. Suppose we want the output on pin 10 to act as a clock signal. The people who made the chip expect that pin to be called "Output 4" in the program. Please don't ask me why it isn't called "10" in the program as I have no answer for you other than "it takes all sorts of people to make a world".

All right, suppose we want to produce an output signal like a 555 chip running at 50 Hz. We choose one of our output pins, say, the physical pin 10, that being the bottom right hand pin on the chip. As you can see from the pin diagram of the chip shown above, pin 10 is called "Output 4" in a set of commands, or just "4" to save typing. The program might be:

```
Main:
  high 4
  pause 10
  low 4
  pause 10
goto Main
```

Wow - really difficult stuff !! Only a genius could manage to program! Well, we'll see if we can struggle along with this "difficult" stuff.

The "Main:" at the start is a "label" which can be jumped to and that is done by the "goto Main" command which sends the chip back to repeat the commands in the loop indefinitely (or until the chip is powered down).

The second line "high 4" tells the chip to put the maximum possible voltage on the "Output 4" which is the physical pin 10 of the chip. The chip does this immediately, with no time delay.

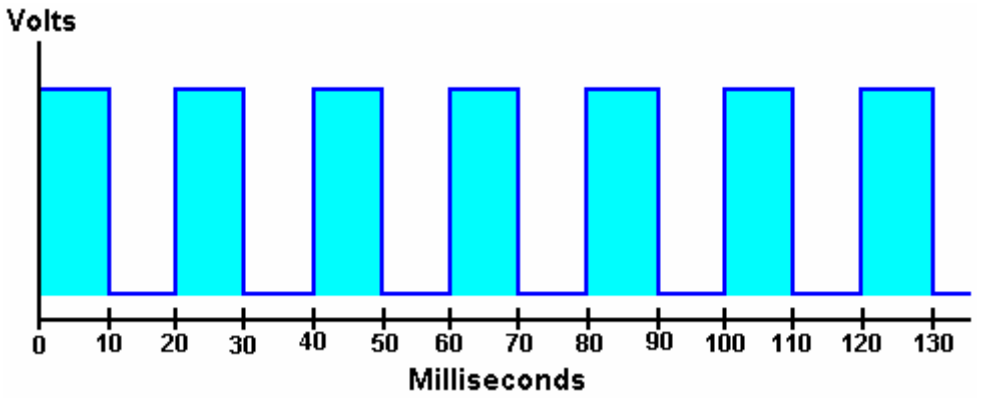
If we want the output to give a 50 Hz output signal, then the voltage on our chosen output pin will have to go high, pause, go low, pause and go high again, 50 times each second. As there are 1,000 milliseconds in one second, and the chip's clock runs with 1 millisecond ticks, then we need our complete cycle of "up, pause, down, pause" to happen 50 times in those 1,000 clock ticks. That is, once every 20 ticks, so each delay will be 10 clock ticks long.

The third line "pause 10" tells the chip to sit on it's hands and do nothing for the next 10 ticks of it's internal clock (which ticks 1,000 times per second).

The fourth line "low 4" tells the chip to lower the output voltage on it's "Output 4" (pin 10 in real life) to it's minimum value.

The fifth line "pause 10" tells the chip to wait for 10 milliseconds before doing anything else.

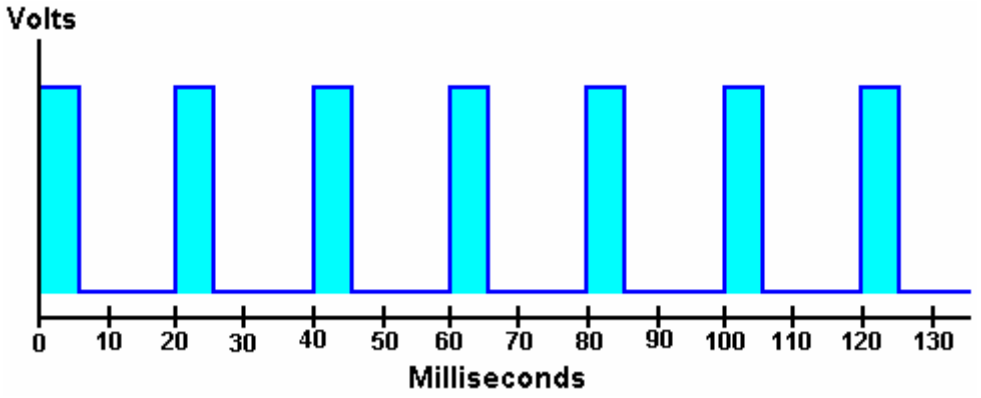
The last line "goto Main" tells the computer to go back to the label "Main:" and continue with whatever instructions follow that label. This puts the chip into an 'infinite loop' which will make it generate that output waveform continuously. The output will look like this:



This gives an even waveform, that is, one with a Mark/Space ratio of 50:50 or a Duty Cycle of 50%. If we want the same rate of pulsing but a Duty Cycle of just 25% then the program would be:

```
Main:
  high 4
  pause 5
  low 4
  pause 15
goto Main
```

which produces this waveform:

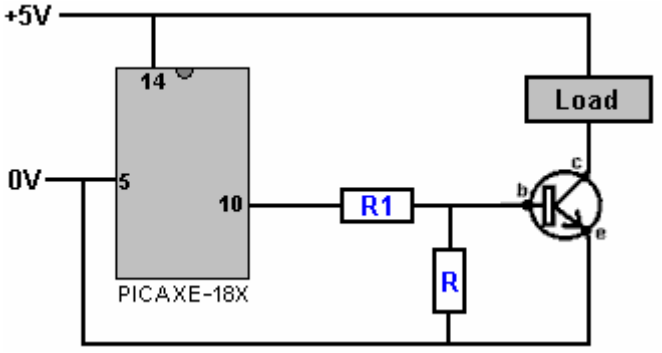


If you wanted "Output 7" (physical pin 13) to do the reverse of this at the same time - that is, when Output 4 goes high we want Output 7 to go low, and vice versa, then, for a 20% Duty Cycle the program would be:

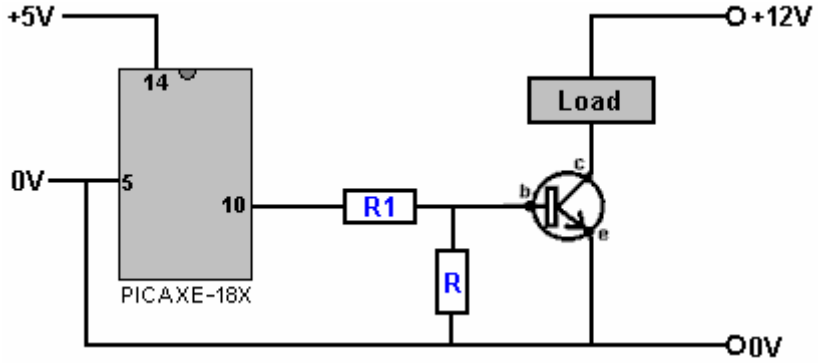
```
Main:
  high 4
  low 7
  pause 4
  low 4
  high 7
  pause 16
goto Main
```

These output voltages are then used in exactly the same way as the output voltages on pin 3 of a 555 chip, or any of the outputs of NAND gates, Hall-effect sensor chips, Schmitt triggers, or whatever. If the device to be powered requires very little current, then the easiest method is to connect the load directly to the output pin.

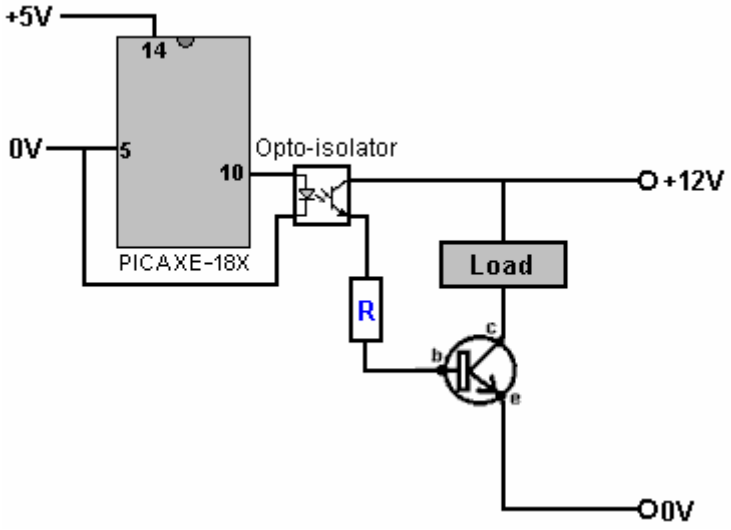
If, as is most often the case, the device to be powered needs a large current to make it work, then the output voltage is used to power a transistor, perhaps like this:



Here, the resistor "R1" limits the current fed into the base of the transistor when pin 10 goes high, but allowing enough current for the transistor to switch on fully, powering the load. The resistor "R" makes sure that the transistor switches off fully when the output on pin 10 goes low. The circuit as shown restricts the load to some piece of equipment which can operate on just five volts, so an alternative circuit could be:



This allows whatever voltage the load needs to be applied to the load, while the PIC chip remains running on it's normal 5-volt supply. However, the equipment to be powered may not be able to have a common zero voltage connection with the PIC. To deal with this, an optical isolation chip can be used like this:



Here a high output voltage on pin 10 of the PIC chip lights up the LED inside the opto-isolator chip, causing a major drop in the resistance between the other two pins. This causes a current controlled by the resistor "R" to be fed into the base of the transistor, switching it on and powering the load.

Recently, a very popular programmable chip has been introduced. It is called the “Arduino” and it is fast and versatile and very popular with experimenters. There is an extensive set of English-language Video tutorials on the Arduino chip, the first in the series by Jeremy Blum is http://www.youtube.com/watch?v=fCxzA9_kg6s. The board looks like this:

Capacitors.
We have avoided mentioning capacitors in any detail as it has not been necessary for understanding the circuitry covered so far. Capacitors come in many sizes, types and makes. Their size is stated in 'Farads' but as the Farad is a very large unit, you are unlikely to encounter a capacitor marked in anything larger than a microfarad, which is a millionth of a Farad. The symbol for a microfarad is mu-F where 'mu' is the letter of the Greek alphabet. This is a pain for normal text production as Greek letters do not occur in your average font. Some circuit diagrams give up on 'mu' and just write it as uF which looks like mu-F slightly mis-printed where the descender of the mu has not printed.

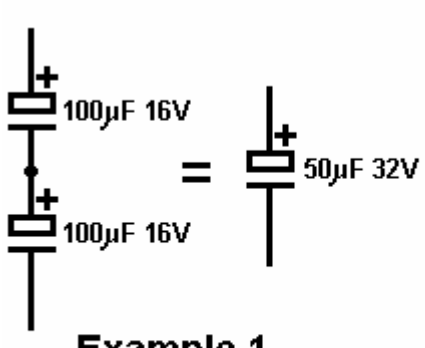
Anyway, very large capacitors which you may encounter range from 5,000 microfarads to maybe as much as 20,000 microfarads. Large capacitors range from 10 microfarads to 5000 microfarads. Medium sized capacitors run from 0.1 microfarad to about 5 microfarads and small capacitors are those below 0.1 microfarad.

1000 nanofarads ('nF') = 1 microfarad.
1000 picofarads ('pF') = 1 nanofarad

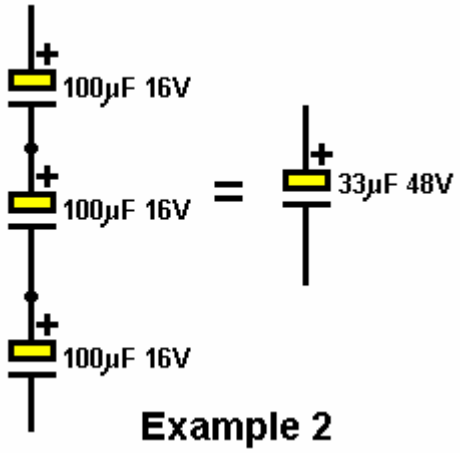
So:

0.01 microfarad can be written as 10nF
0.1 microfarad can be written as 100nF
0.1nF can be written as 100pF

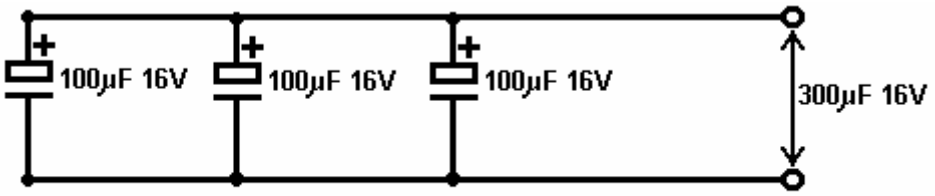
Capacitors larger than 1 microfarad tend to be 'polarised'. In other words, the capacitor has a '+' connector and a '-' connector, and it does matter which way round you connect it. The larger capacitors have a voltage rating and this should **not** be exceeded as the capacitor can be damaged and possibly even totally destroyed. Capacitors can be added together, but surprisingly, they add in the reverse way to resistors:



Example 1



Example 2



Example 3

If two capacitors are wired in series, as shown in Example 1 above, the overall capacity is reduced while the voltage rating increases. The reduction in capacitance is given by:

1/Ct = 1/C1 + 1/C2 + 1/C3 + ...

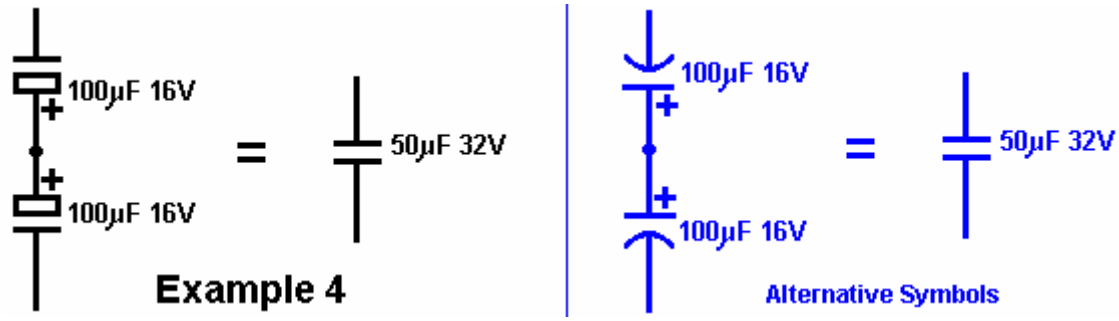
In Example 1, then, 1/total capacitance = 1/100 + 1/100 or 1/Ct = 2/100 or 1/Ct = 1/50
so the overall capacitance reduces from 100 microfarads to 50 microfarads. The advantage in wiring the capacitors like this is that the voltage rating has now increased to 32V (16V across each of the capacitors).

In Example 2, the overall capacitance has reduced to a third of 100 microfarads but the voltage rating has tripled.

In Example 3, the capacitors are wired in parallel. The voltage rating is unchanged but the overall capacitance is now the sum of the three capacitors, namely 300 microfarads.

There is no need for the capacitors to have similar values, there are merely shown that way in the examples to make the arithmetic easier and not distract you from the ways in which the capacitors interact together.

Occasionally, a circuit needs a large capacitor which is not polarised. This can be provided by placing two polarised capacitors back-to back like this:



When the capacitors are connected this way, it does not matter which end of the pair is connected to the positive side of the circuit and which to the negative side.

Large capacitors usually have their capacitance and voltage printed on the outside of the capacitor, but small capacitors are usually far too tiny for that to be an option. So, a code very much like that used for resistors id used for small capacitors. The code is a 2-digit code for capacitors up to 100 picofarads and for higher values it is a 3-digit code where the first two digits are the numerical value of the capacitor in picofarads and the thired digit is the number of zeros following the two digits. One thousand picofarads (pF) is one nanofarad (nF) and one thousand nanofarads is one microfarad. These are some common values:

Value	Code	Value	Code
10 pF	10	2.2 nF	222
22 pF	22	4.7 nF	472
47 pF	47	10 nF	103
100 pF	101	22 nF	223
220 pF	221	47 nF	473
470 pF	471	100 nF	104
1 nF	102	220 nF	224

The time has come for a serious warning: High voltages are very, very dangerous. Do not become so familiar with them that you treat them casually. **High voltages can kill you.** Capacitors are capable of building up high voltages and some good makes can hold the charge for several days.

In particular, do **not** try to make adjustments to, or take parts from, the inside of a TV set. A black and white TV set uses 18,000 Volts on the magnetic coils used to create the moving picture on the tube. A capacitor inside the set may well have that voltage on it three days after the set was last used. Don't fool around inside a TV set, it could kill you quick, or if you are really unlucky, it could injure you for life. A colour TV set uses 27,000 Volts to operate the coils inside it and that will fry you in jig time if you touch it.

Also, please don't think that you are safe if you don't quite touch it; 27,000 volts can jump across a gap to your hand. If you try to discharge a TV capacitor using a metal screwdriver with a wooden handle, please ensure that you medical insurance is up to date before you do it. You can receive a hefty shock through the screwdriver handle.

Voltages up to 24 Volts should be quite safe. **However**, some circuits will generate very high voltages even though the battery driving the circuit is low voltage. A standard off-the-shelf inverter circuit produces 240 Volts AC from a 12 Volt battery. Just because the battery is only 12 Volts does **not** mean that the circuit is not dangerous. Circuits which have inductors in them can produce high voltages, especially if they contain large capacitors. The voltage which produces the spark in your car engine is very high and it comes from the 12-volt car battery. You know enough about this by now, so **pay attention**!

The more advanced stuff:

You do not need to bother with this section if you are just starting out with some basic switching circuits of the type already described in this tutorial, so please feel free to skip this section and move on to the "Prototype Construction" section which you will find immediately useful.

This section is a lightweight introduction to Alternating Current circuits and pulsed DC circuits. Let me stress again that I am mainly self-taught and so this is just a general introduction based on my present understanding.

AC Power Factors.

Alternating Current, generally called "AC" is called that because the voltage of this type of power supply is not a constant value. A car battery, for instance, is DC and has a fairly constant voltage usually about 12.8 volts when in it's fully charged state. If you connect a voltmeter across a car battery and watch it, the voltage reading will not change. Minute after minute it says exactly the same because it is a DC source.

If you connect an AC voltmeter across an AC power supply, it too will give a steady reading, but it is telling a lie. The voltage is changing all the time in spite of that steady meter reading. What the meter is doing is **assuming** that the AC waveform is a sine wave like this:

and based on that assumption, it displays a voltage reading which is called the "Root Mean Square" or "RMS" value. The main difficulty with a sine wave is that the voltage is below zero volts for exactly the same length of time as it is above zero volts, so if you average it, the result is zero volts, which is not a satisfactory result because you can get a shock from it and so it can't be zero volts, no matter what the arithmetical average is.

To get over this problem, the voltage is measured thousands of times per second and the results squared (that is, the value is multiplied by itself) and then those values are averaged. This has the advantage that when the voltage is say, minus 10 volts and you square it, the answer is plus 100 volts. In fact, all of the answers will be positive, which means that you can add them together, average them and get a sensible result. However, you end up with a value which is far too high because you squared every measurement, and so you need to take the square root of that average (or "mean") value, and that is where the fancy sounding "Root Mean Square" name comes from – you are taking the (square) root of the (average or) mean value of the squared measurements.

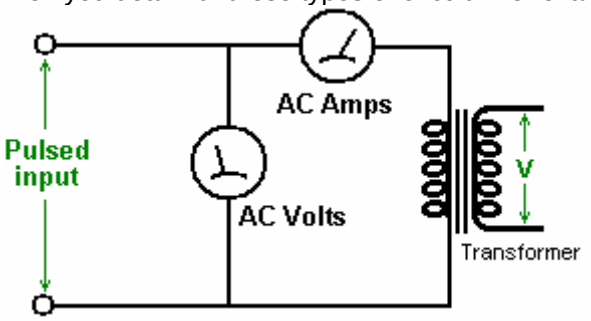
With a sine wave like this, the voltage peaks are 41.4% higher than the RMS value which everyone talks about. This means that if you feed 100 volts AC through a rectifier bridge of four diodes and feed it into a capacitor the

capacitor voltage will **not** be 100 volts DC but instead it will be 141.4 volts DC and you need to remember that when choosing the voltage rating of the capacitor. In that instance I would suggest a capacitor which is made to operate with voltages up to 200 volts.

You probably already knew all of that, but it may not have occurred to you that if you use a standard AC voltmeter on a waveform which is **not** a sine wave, that the reading on the meter is most unlikely to be correct or anywhere near correct. So, please don't merrily connect an AC voltmeter across a circuit which is producing sharp voltage spikes like, for instance, one of John Bedini's battery pulsing circuits, and think that the meter reading means anything (other than meaning that you don't understand what you are doing).

You will, hopefully, have learned that power in watts is determined by multiplying the current in amps by the voltage in volts. For example, 10 amps of current flowing out of a 12 volt power supply, represents 120 watts of power. Unfortunately, that only holds true for circuits which are operating on DC, or AC circuits which have only resistors in them. The situation changes for AC circuits which have non-resistive components in them.

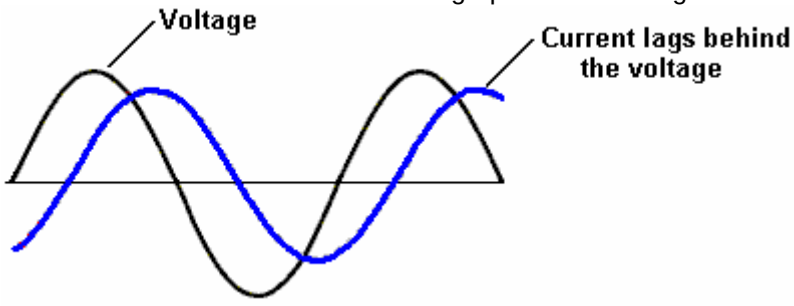
The circuits of this type which you are likely to come across are circuits which have coils in them, and you need to think about what you are doing when you deal with these types of circuit. For example, consider this circuit:



This is the output section of a prototype which you have just built. The input to the prototype is DC and measures at 12 volts, 2 amps (which is 24 watts). Your AC voltmeter on the output reads 15 volts and your AC ammeter reads 2.5 amps and you are delighted because 15 x 2.5 = 37.5 which looks much bigger than the 24 watts of input power. **But**, just before you go rushing off to announce on YouTube that you have made a prototype with COP = 1.56 or 156% efficient, you need to consider the real facts.

This is an AC circuit and unless your prototype is producing a perfect sine wave, then the AC voltmeter reading will be meaningless. It is just possible that your AC ammeter is one of the few types that can accurately measure the current no matter what sort of waveform is fed to it, but it is distinctly possible that it will be a digital meter which assesses current by measuring the AC voltage across a resistor in series with the output, and if that is the case, it will probably be assuming a sine wave. The odds are that both readings are wrong, but let's take the case where we have great meters which are reading the values perfectly correctly. Then the output will be 37.5 watts, won't it? Well, actually, no it won't. The reason for this is that the circuit is feeding the transformer winding which is a coil and coils don't work like that.

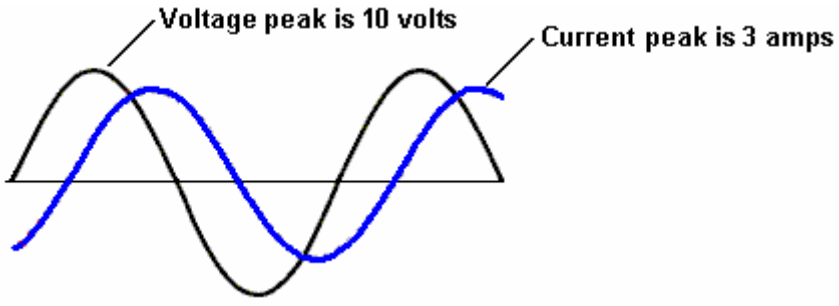
The problem is that, unlike a resistor, when you apply a voltage across a coil the coil starts absorbing energy and feeding it into the magnetic field around the coil, so there is a delay before the current reaches it's maximum value. With DC, this generally doesn't matter very much, but with AC where the voltage is changing continuously, it matters a great deal. The situation can be as shown in this graph of both voltage and current:



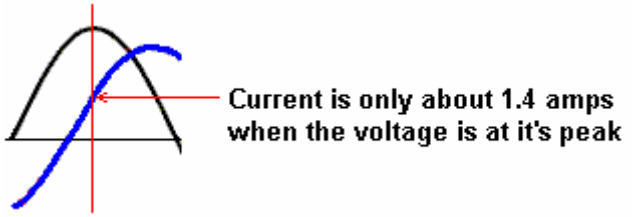
At first, this does not look like any great problem, but it has a very significant effect on the actual power in watts. To get the 37.5 watts output which we were talking about earlier, we multiplied the average voltage level by the average current level. But these two values do not occur at the same time and that has a major effect.

As this can be a little difficult to see, let's take the peak values rather than the averages as they are easier to see. Let's say that in our example graph that the voltage peak is 10 volts and the current peak is 3 amps. If this were

DC we would multiply them together and say that the power was 30 watts. But with AC, this does not work due to the timing difference:

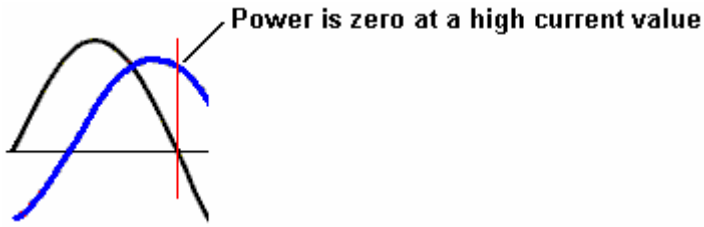


When the voltage is peaking, the current is nowhere near it's peak value of 3 amps:



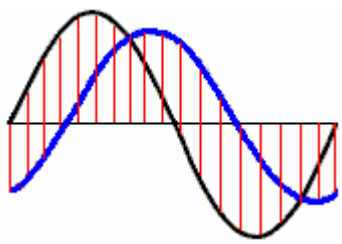
As a result of this, instead of getting our expected peak power at the top of the voltage peak, the actual power in watts is very much lower – less than half of what we were expecting. Not so good, but it gets worse when you look at the situation more closely. Take a look at what the voltage is when the current crosses the zero line, that is, when the current is zero. The output power is zero when the current is zero but this occurs when the voltage is at a very high value:

The same goes for when the voltage is zero. When the voltage is zero, then the power is also zero, and you will notice that this occurs when the current is at a high value:



The power is **not** the average current multiplied by the average voltage if there is a coil involved in the circuit – it will be less than that by an amount known as the "power factor" and I'll leave you to work out why it is called that.

So, how do you determine what the power is? It is done by sampling the voltage and current many times per second and averaging those combined results:



Both the voltage and the current are sampled at the times indicated by the vertical red lines and those figures are used to calculate the actual power level. In this example, only a few samplings are shown, but in practice, a very large number of samples will be taken. The piece of equipment which does this is known as a wattmeter as it measures watts of power. The sampling can be done by windings inside the instrument, resulting in an instrument which can be damaged by overloading without the needle being anywhere near full deflection, or it can be done by digital sampling and mathematical integration. Most digital sampling versions of these meters only operate at high frequencies, typically over 400,000 cycles per second. Both varieties of wattmeter can handle any waveform and not just sine waves.

The power company supplying your home measures the current and assumes that the full voltage is present all of the time that the current is being drawn. If you are powering a powerful electric motor from the mains, then this current lag will cost you money as the power company does not take it into account. It is possible to correct the situation by connecting one or more suitable capacitors across the motor to minimise the power loss.

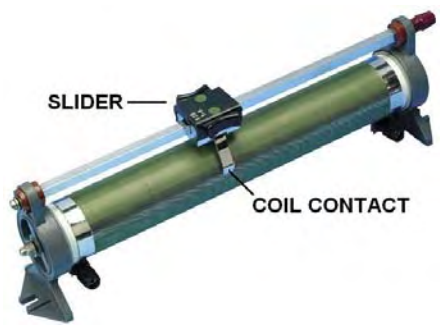
With a coil (fancy name “inductor” symbol “L”), AC operation is very different to DC operation. The coil has a DC resistance which can be measured with the ohms range of a multimeter, but that resistance does not apply when AC is being used as the AC current flow is **not** determined by the DC resistance of the coil alone. Because of this, a second term has to be used for the current-controlling factor of the coil, and the term chosen is “impedance” . The wire in any coil has a resistance and that opposes current flow through the coil irrespective of whether the voltage applied to the coil is DC or AC. The capacitance between the neighbouring turns of wire in a coil, introduces a feature of the coil which “impedes” AC current flow through the coil and the amount of that impedance depends on the frequency of the AC voltage being applied to the coil.

The impedance of a coil depends on it’s size, shape, method of winding, number of turns and core material. If the core is made up of iron or steel, (usually thin layers of iron which are insulated from each other), then it can only handle low frequencies. You can forget about trying to pass 10,000 cycles per second (“Hz”) through the coil as the core just can’t change it’s magnetisation fast enough to cope with that frequency. A core of that type is ok for the very low 50 Hz or 60 Hz frequencies used for mains power, which are kept that low so that electric motors can use it directly.

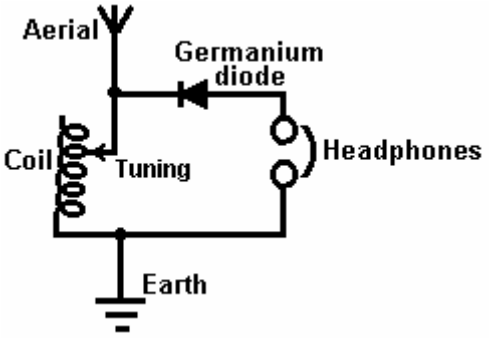
For higher frequencies, ferrite can be used for a core and that is why some portable radios use ferrite-rod aerials, which are a bar of ferrite with a coil wound on it. For higher frequencies (or higher efficiencies) iron dust encapsulated in epoxy resin is used. An alternative is to not use any core material and that is referred to as an air-core coil. These are not limited in frequency by the core but they have a very much lower inductance for any given number of turns. The efficiency of the coil is called it’s “Q” (for “Quality”) and the higher the Q factor, the better. The resistance of the wire lowers the Q factor.

A coil has inductance, and resistance caused by the wire, and capacitance caused by the turns being near each other. However, having said that, the inductance is normally so much bigger than the other two components that we tend to ignore the other two. Something which may not be immediately obvious is that the impedance to AC current flow through the coil depends on how fast the voltage is changing. If the AC voltage applied to a coil completes one cycle every ten seconds, then the impedance will be much lower than if the voltage cycles a million times per second.

If you had to guess, you would think that the impedance would increase steadily as the AC frequency increased. In other words, a straight-line graph type of change. That is not the case. Due to a feature called resonance, there is one particular frequency at which the impedance of the coil increases massively. This is used in the tuning method for AM radio receivers. In the very early days when electronic components were hard to come by, variable coils were sometimes used for tuning. We still have variable coils today, generally for handling large currents rather than radio signals, and we call them “rheostats” and some look like this:

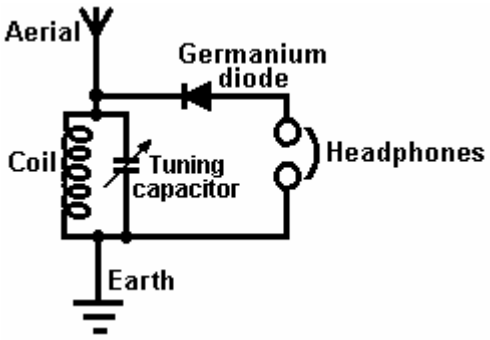


These have a coil of wire wound around a hollow former and a slider can be pushed along a bar, connecting the slider to different winds in the coil depending on it's position along the supporting bar. The coil connections are then to the slider and to one end of the coil. The position of the slider effectively changes the number of turns of wire in the part of the coil which is in the circuit. Changing the number of turns in the coil, changes the resonant frequency of that coil. AC current finds it very, very hard to get through a coil which has the same resonant frequency as the AC current frequency. Because of this, it can be used as a radio signal tuner:



If the coil's resonant frequency is changed to match that of a local radio station by sliding the contact along the coil, then that particular AC signal frequency from the radio transmitter finds it almost impossible to get through the coil and so it (and only it) diverts through the diode and headphones as it flows from the aerial wire to the earth wire and the radio station is heard in the headphones. If there are other radio signals coming down the aerial wire, then, because they are not at the resonant frequency of the coil, they flow freely through the coil and don't go through the headphones.

This system was soon changed when variable capacitors became available as they are cheaper and more compact. So, instead of using a variable coil for tuning the radio signal, a variable capacitor connected across the tuning coil did the same job:

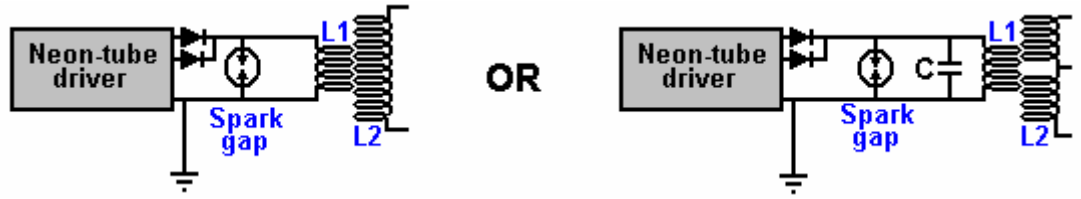


Resonance.

While the circuit diagram above is marked "Tuning capacitor" that is actually quite misleading. Yes, you tune the radio receiver by adjusting the setting of the variable capacitor, **but**, what the capacitor is doing is altering the resonant frequency of the coil/capacitor combination and it is the resonant frequency of that combination which is doing exactly the same job as the variable coil did on it's own.

This draws attention to two very important facts concerning coil/capacitor combinations. When a capacitor is placed across a coil "in parallel" as shown in this radio receiver circuit, then the combination has a very high impedance (resistance to AC current flow) at the resonant frequency. But if the capacitor is placed "in series" with the coil, then there is nearly zero impedance at the resonant frequency of the combination:

This may seem like something which practical people would not bother with, after all, who really cares? However, it is a very practical point indeed. In Chapter 3, some of the very high-power devices produced by Don Smith are described. Typically, he uses an off-the-shelf neon-tube driver module as an easy way to provide a high-voltage, high-frequency AC current source, typically, 6,000 volts at 30,000 Hz. He then feeds that power into a Tesla Coil which is itself, a power amplifier. The arrangement is like this:



People who try to replicate Don's designs tend to say "I get great sparks at the spark gap until I connect the **L1** coil and then the sparks stop. This circuit can never work because the resistance of the coil is too low".

If the resonant frequency of the **L1** coil does not match the frequency being produced by the neon-tube driver circuit, then the low impedance of the **L1** coil will definitely pull the voltage of the neon-tube driver down to a very low value. But if the **L1** coil has the same resonant frequency as the driver circuit, then the **L1** coil (or the **L1** coil/capacitor combination shown on the right, will have a very high resistance to current flow through it and it will work well with the driver circuit. So, no sparks, means that the coil tuning is off. It is the same as tuning a radio receiver, get the tuning wrong and you don't hear the radio station.

Choosing components which are not specified.

Some people find it difficult to select a suitable component where the exact component is not specified or where an alternative has to be selected, so perhaps a few general pointers might be helpful. The reason why component values are omitted may well be because a very wide range of alternative values can be used and if one particular is specified, the newcomers to electronics feel that they **have** to use that one value or the circuit will not work, (which is almost never the case). For example, I have been asked if a capacitor rated at 25V could be used instead of the same value capacitor rated at 16V shown in the circuit, to which the answer is 'yes, most definitely'. The lower voltage rating is adequate and the component cheaper to buy, but if one of a higher voltage rating is available, then it can be used.

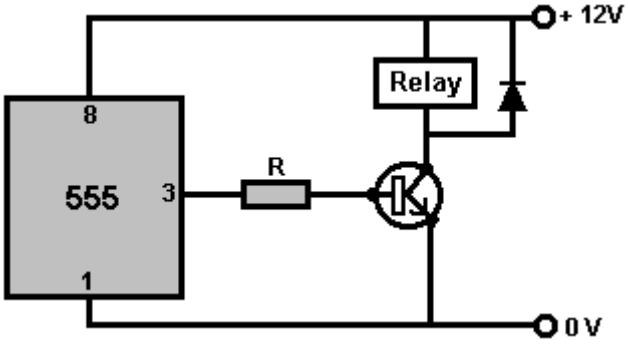
With **capacitors**, you need to consider the physical size and wire connections, the capacitance, the voltage rating, and the leakage. The cost and size of a capacitor is directly related to it's voltage rating, and once the voltage rating exceeds that normally used, the price shoots up rapidly as the sales volume reduces rapidly, which in turn, discourages further sales. This sometimes causes circuit builders to connect chains of cheaper capacitors together to make a smaller-capacity high-voltage capacitor. In the case of Tesla Coil builders, they then may connect several of these chains in parallel to boost the capacitance.

If the voltage rating is exceeded (usually by a very large amount), the capacitor will be damaged and become either a short-circuit, or more likely, an open circuit. Either way, it will never work as a capacitor again. In a household circuit, where the capacitor is being used as part of the power supply to the circuit, the voltage rating does not need to be much higher than the supply voltage, with say, 16V being used for a 12V circuit. You could use a capacitor rated at 25V, 40V, 63V, 100V or 400V and it would work perfectly well, but it will be much larger and have cost much more. But, if you have one sitting around and not being used, there is no reason why you should not use it rather than paying to buy another one.

If the capacitor is being used in a timing circuit where a high-value resistor is feeding current to it, then the leakage current of the capacitor becomes very important. Electrolytic capacitors are seldom suitable for such an application as they have a small, unpredictable leakage current which will vary with the age of the capacitor. For accurate timing with a capacitor, ceramic, polypropylene, mylar or tantalum should be used.

The voltage rating for an electrolytic capacitor is for DC, so if you use it for limiting current in an AC power supply, that is, where the current flows through the capacitor rather than the capacitor being placed across the supply and is acting to combat ripple, then great care is needed. The capacitor will heat up due to the power flowing through it, and it is possible for an electrolytic capacitor used in that way to rupture or 'explode' due to the electrolyte boiling. Instead, you need to use the very much more expensive oil-filled can capacitors (as shown near the end of chapter 10). That style of usage is unusual for home constructors.

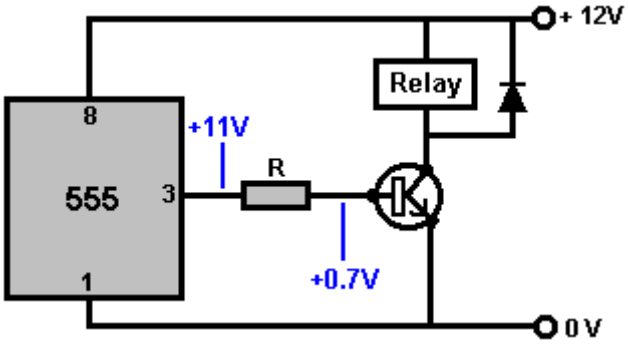
With bi-polar **transistors**, you need to use commonsense. Suppose a 555 timer chip is required to power a transistor which controls a relay:



For the moment, we will ignore the fact that the 555 could drive the relay directly without the need for a transistor. let's say that the relay draws a current of 30 mA when connected to a 12V supply. Therefore, the transistor needs to be able to handle a current of 30 mA. Any small switching transistor such as the BC109 or 2N2222 can easily handle that current. The transistor also needs to be able to handle 12 volts. If in doubt, look up the characteristics of your choice of transistor at http://www.alldatasheet.co.kr/ by entering the transistor name 'BC109' or whatever in the entry box at the top of the screen and clicking on the button to the right of it. Eventually, it will let you download a pdf document specifying the transistor, and that will show you the voltages which the transistor can handle. Both of the above transistors can handle far more than 12V.

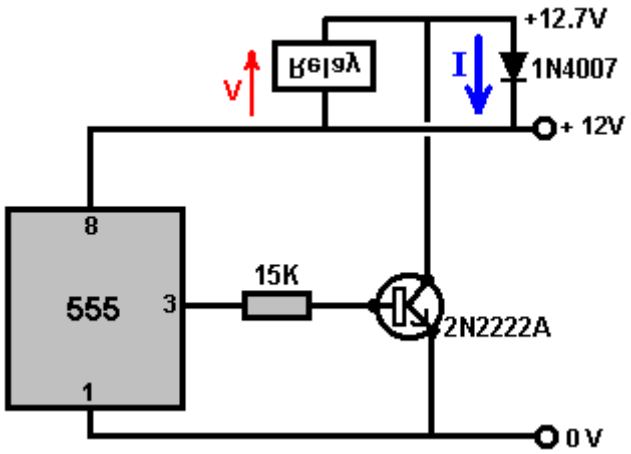
The next question is, 'can the transistor switch fast enough to work in this circuit?' and the data sheet will show that they can switch on and off a million times per second. As the relay can only switch on and off a few times per second, the transistor can easily operate fast enough to handle the switching.

Next, we need to know what size of resistor would be suitable. The data sheet will also show the DC current gain of the transistor. This is usually marked as "hfe" and for these transistors is likely to be a minimum of, say, 200. This means that the current flowing into the base of the transistor needs to be one two-hundredth of the relay's 30 mA which is 0.15 mA. The resistor will have about +11 volts at pin 3 of the 555 timer and around +0.7 volts at the base of the transistor when it is switched fully on. That means that the resistor will have about 10.3 volts across it when the relay is switched on:



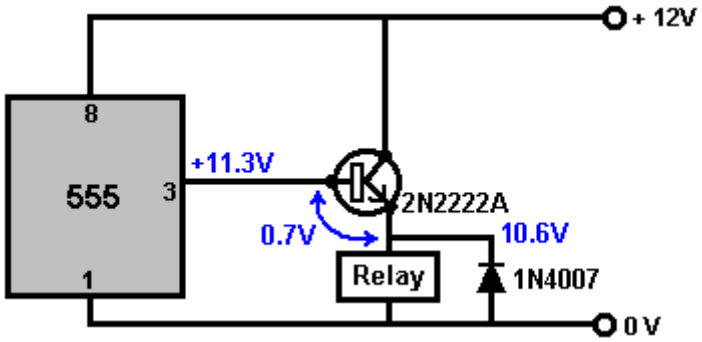
So, what size of resistor will have 0.15 mA flowing through it when there is a 10.3-volt drop across it? We know that a 1K resistor passes 1mA per volt and so would pass 10.3 mA with 10.3 volts across it. That is far more than we need. A 10K resistor would pass 1.03 mA which is still far too much but certainly could be used. As it is a resistor, we can use Ohm's Law: R = V / A (Ohms equals Volts over Amps), or R = 10.3 / 0.00015 which is 68K. So, any resistor between 68K and perhaps 15K should work well.

The diode is there to protect the transistor from excessive voltage caused by the coil of the relay. When a coil is switched off suddenly, it generates a reverse voltage which can be hundreds of volts, pulling the collector of the transistor far above the +12V power supply line. When that starts to happen, it effectively reverses the diode direction, allowing it to conduct and short-circuit that big voltage spike:



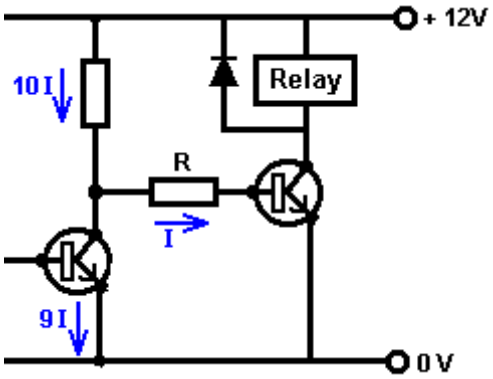
Due to the short-circuiting, the voltage can't get any higher and the current through the diode is not large, so most diodes such as the popular and cheap 1N4001 or 1N4007 types can be used.

When a transistor is connected like that and switched on, it is effectively a short-circuit between it's collector and emitter, and that places the full 12 volts across the relay, powering it very solidly. This connection method is called a "common-emitter" circuit because all of the transistors used have their emitters all wired in common to the 0V line. An alternative arrangement is the "emitter-follower" circuit:



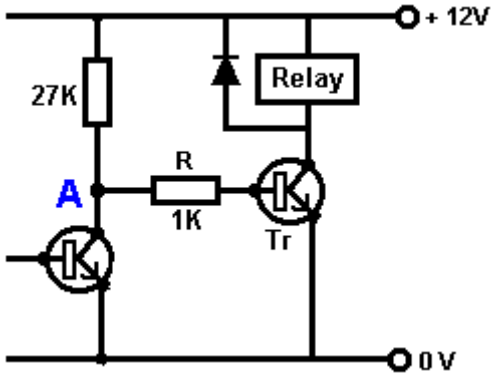
With this circuit arrangement, the emitter of the transistor "follows" the voltage on pin 3 of the 555 timer. It is always a constant voltage below it, typically about 0.7 volts. The output of the 555 timer has a maximum of about 0.7V below the supply voltage, and so it's maximum value is about 11.3V in this circuit. The transistor drops that by a further 0.7V, which means that the relay only gets about 10.6V across it instead of the full 12V of the supply, which means that it should be a 10-volt relay rather than a 12-volts relay.

Those are the easy cases because the 555 timer can supply at least 200 mA through it's output pin, while keeping the output voltage steady. That is not the case with simple transistor circuits. Take a situation like this:



For audio work - microphone pre-amplifiers and the like – the rule of thumb is that the current flowing through the first transistor should be at least ten times the current required by the base of the second transistor in order not to drag down and distort the audio waveform.

Relay switching is not so critical but the same general principle applies and attention needs to paid to the collector resistor of the preceding transistor. For example, if the current flowing through the preceeding transistor is small, say, 0.5 mA and the output transistor needs 1.5 mA flowing into it's base, then there can be a problem. In this circuit, for example:



Here, the voltage at point "A" goes high because the first transistor switches off and so becomes the same as a resistor of 1Meg or more. Normally, that resistance is so much greater than the 27K of it's resistor, that the voltage at point "A" would be nearly +12V, but if you were to connect the resistor "R" of just 1K in value, then the situation is changed completely. the base of "Tr" can't rise above 0.7V. The first transistor can be ignore due to its very high resistance. That leaves a voltage-divider pair of resistors, the 27K and the 1K, with 11.3 volts across them, stopping the voltage at point "A" from rising above 1.13V instead of the original 12V and transistor "Tr" will only get 0.43 mA instead of the 1.5 mA which was wanted. The transistor "Tr" has effectively a 28K resistor feeding it current from the +12V rail.

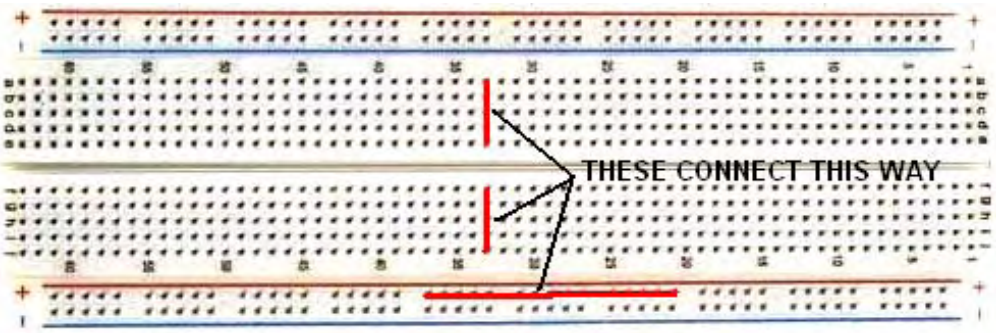
One solution would be to raise the current through the first transistor by using a resistor a good deal smaller than the present 27K. Another option is to lower the input current requirement of transistor "Tr" by making it a Darlington pair or by using a transistor with a much higher gain.

Constructing Prototypes.

The main options for building a prototype circuit are:

1. A (plug-in) breadboard
2. Electrical screw connector strips.
3. Stripboard
4. A printed circuit board.

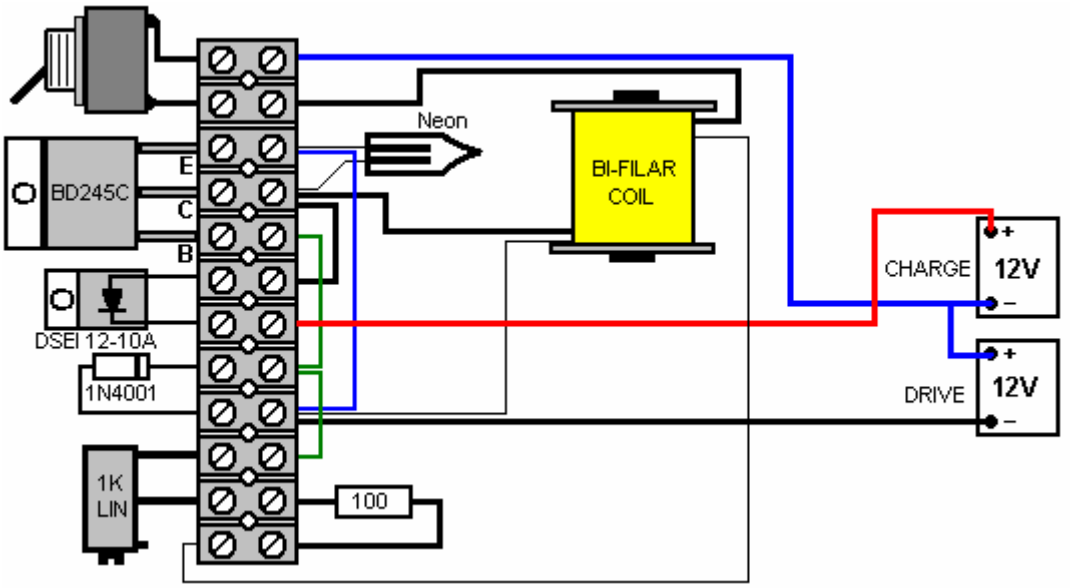
1. The typical breadboard unit consists of a matrix of clip holes wired in strips, into which component leads can be pushed to make a circuit. In my opinion, they are best avoided as it takes quite some effort to implement any significant circuit using them, some components do not fit well in the sockets which are small enough to take DIL IC packages, and when you do get a circuit working well on the breadboard, there is no guarantee that it will work well when you attempt to move it to a permanent soldered board:



While a plastic board of this type looks as if it should be quick and easy to use, I have never found it to be so ever since the boards were scaled down in size to take the closely-spaced pins of integrated circuits ("chips"). It is generally difficult to lay the components out in the same pattern as the circuit diagram, and if they are not, then it becomes slow to follow the circuit through on the breadboard layout.

2. The local hardware shop has cheap screw connectors which can be very effective. These come in several sizes and the smaller ones are very convenient for constructing transistor circuitry. They look like this:

Circuits can be assembled very easily, using these connectors and an example might be one of the John Bedini battery pulsing circuits which might have a layout like this:



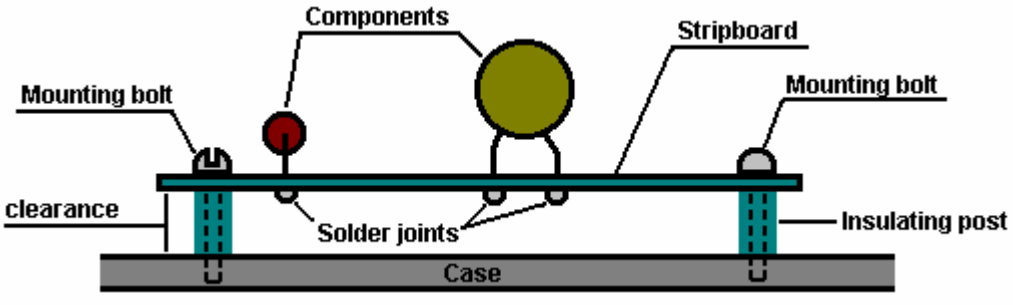
I have built this circuit using this style of construction and it was very successful indeed, being very quick and easy to construct and it proved to be very tough and effective over a long period of use. The plastic strip has a hole between each connector strip and that allows you to bolt the strip to a base board on which you mount other components, in this case, the pulsing coil and the rotor with the magnets attached. Each connection block can take two or three wires. The wires need to have the insulation removed and the wires scraped clean and shiny if they are not already in that state. If more than one multi-strand wire is being put into one side of a connector, then it is usually best to twist the wires together before tightening the clamping screw. If you want, you can give the twisted wires a thin coat of solder, but this has to be done neatly to avoid producing a joint which is too large to fit into the connector. One connector can be cut out of the strip quite easily, using a pair of scissors or a craft knife. Single connectors can join two wires very effectively without the need to solder them.

3. Stripboard, usually called ‘Veroboard’ even if it is not made by Vero, is a quick and satisfactory method, although you have to make very tiny solder joints. Please be aware that the fumes from the burning resin when soldering are most definitely not good for your health and should be avoided by making sure that the ventilation is adequate.

4. A printed circuit board is feasible for a one-off prototype and making one will increase your production skills, so it is also a reasonable option if you have the etching and drilling equipment to hand. Buying all of the necessary equipment if you do not have any, can cost a fair amount, but the skills gained are significant and the finished boards looks very professional.

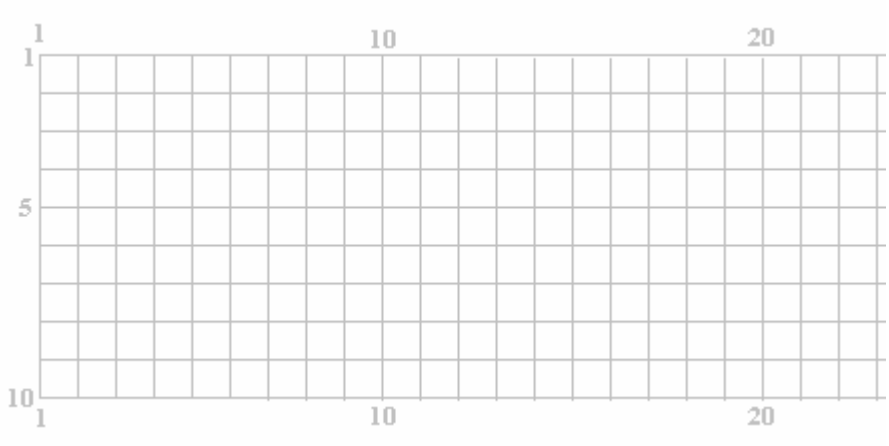
There are several other methods of construction, and many varieties of construction board and stripboard. Simple stripboard will be used in the following descriptions, although the method does apply to many different styles of construction.

The first step is to produce a layout for the components on the board. When designing the layout provision should be made for drilling holes to allow the completed board to be bolted to its case using bolts and insulating pillars to keep the soldered joints clear of all other surfaces.

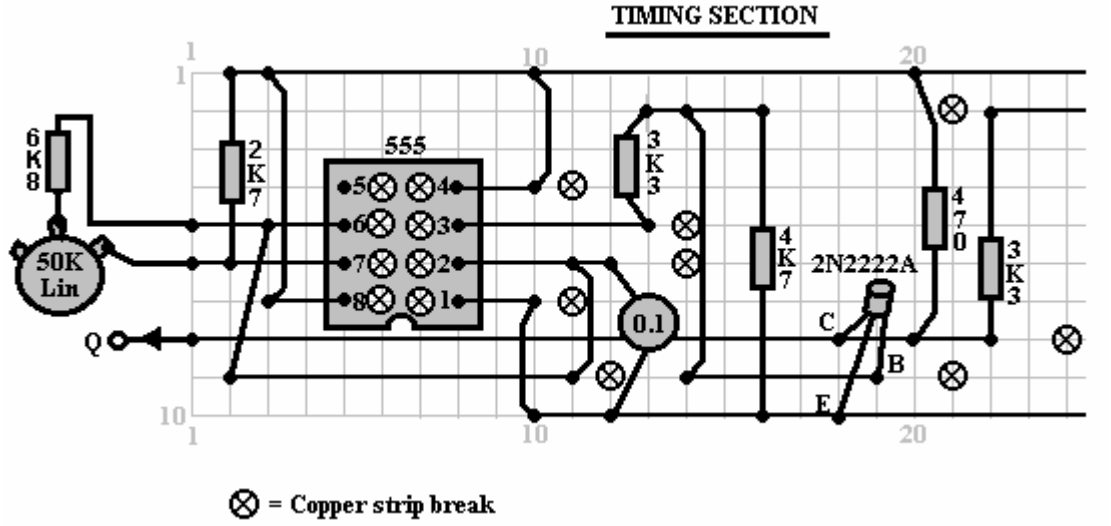


The circuit diagram of the circuit to be built is the starting point. You might wish to draw a light grid of lines to represent the matrix of holes in the strip board. This helps to visualise the run of the copper strips and the sketch can be made to show the exact number of holes available on the piece of strip board to be used. The strip board looks like this:

So you might wish to produce a layout sketch re-usable drawing like this:

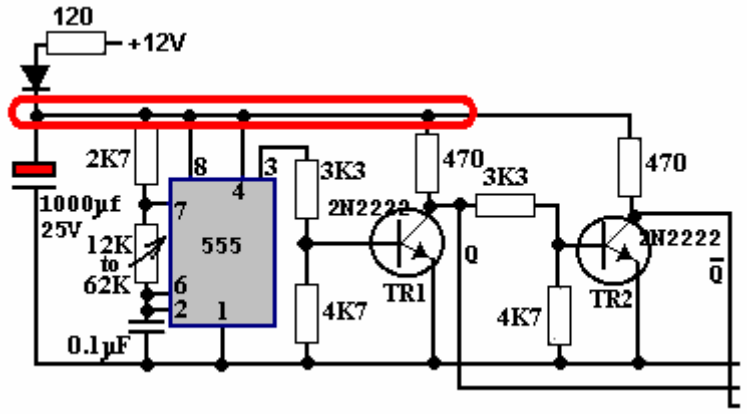


where the horizontal strips are numbered and the vertical lines of holes are also numbered. In this sketch, where the lines cross, represents a hole in the board. The sketch of a possible physical layout can then be prepared and it might look like this when seen from the top although the copper strips on the underside of the board are shown in the sketch:



It is very important when producing a sketch like this, that the copper strips making up the circuit are not accidentally used to connect components further along the board, without breaking the copper strip between the

two sections of the board. It helps to mark a copy of the circuit diagram when you are sketching a possible physical layout on the strip board. It might be done like this:

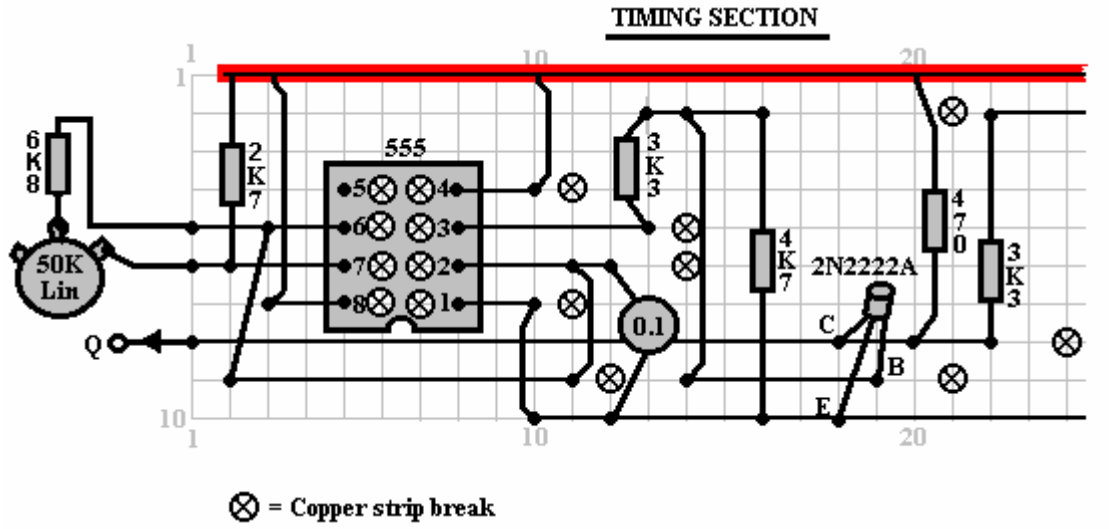


Here, the components just below the diode are ringed to show that they have been marked on the layout sketch and, if necessary, the copper strip broken to isolate the components. A component worth mentioning in passing, is the capacitor marked with red in the circuit diagram. This is a decoupling capacitor, fed from the 12V battery via a resistor and a diode (a diode is not normally used in this part of the circuit).

The decoupling is to provide the 555 chip and drivers with a supply which is reasonably isolated from the heavy current-draw circuit not shown in this small section of the circuit diagram. The pulsating heavy current draw of the rest of the circuit is capable of pulling the battery voltage down slightly many times per second. This creates a voltage ripple on the positive supply line from the battery and to smother the ripple, the resistor and diode are used to feed a large reservoir capacitor which smoothes out the ripple.

The circuit itself is not beyond criticism. Transistor 'TR2' and its associated components are redundant since pin 3 of the 555 chip already supplies the required signal (and with higher drive capacity) so the second output line should be taken directly from pin 3 of the 555 chip. This snippet of circuit is only shown here as an example of marking up a circuit diagram when making a components layout sketch.

As the layout sketch is produced, the circuit diagram should be marked off with a highlighting pen to make sure that every part of the circuit diagram has been successfully copied to the sketch. In the example below, not all of the highlighted strip is shown, since it runs off the small section of the board being shown here:



Many electronic components can be damaged by the high temperatures they are subjected to when being soldered in place. I personally prefer to use a pair of long-nosed pliers to grip the component leads on the upper side of the board while making the solder joint on the underside of the board. The heat running up the component lead then gets diverted into the large volume of metal in the pair of pliers and the component is protected from excessive heat. On the same principle, I always use a DIL socket when soldering a circuit board, that way, the

heat has dissipated fully before the IC is plugged into the socket. It also has the advantage that the IC can be replaced without any difficulty should it become damaged.

If you are using CMOS integrated circuits in any construction, you need to avoid static electricity. Very high levels of voltage build up on your clothes through brushing against objects. This voltage is in the thousands of volts range. It can supply so little current that it does not bother you and you probably do not notice it. CMOS devices operate on such low amounts of current that they can very easily be damaged by your static electricity. Computer hardware professionals wear an earthing lead strapped to their wrists when handling CMOS circuitry. There is no need for you to go that far. CMOS devices are supplied with their leads embedded in a conducting material. Leave them in the material until you are ready to plug them into the circuit and then only hold the plastic body of the case and do not touch any of the pins. Once in place in the circuit, the circuit components will prevent the build up of static charges on the chip.

Soldering is an easily-acquired skill. Multi-cored solder is used for electronic circuit soldering. This solder wire has flux resin contained within it and when melted on a metal surface, the flux removes the oxide layer on the metal, allowing a proper electrical joint to be made. Consequently, it is important that the solder is placed on the joint area and the soldering iron placed on it when it is already in position. If this is done, the flux can clean the joint area and the joint will be good. If the solder is placed on the soldering iron and then the iron moved to the joint, the flux will have burnt away before the joint area is reached and the resulting joint will not be good.

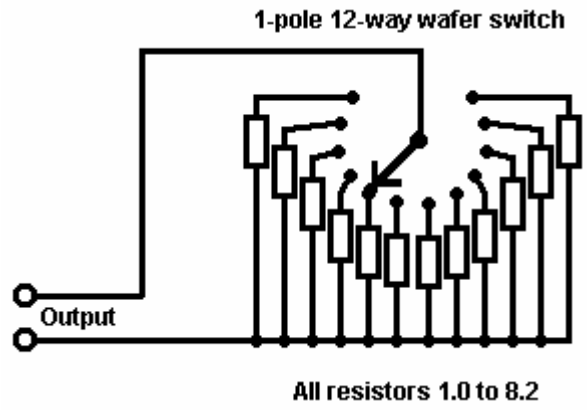
A good solder joint will have a smooth shiny surface and pulling any wire going into the joint will have no effect as the wire is now solidly incorporated into the joint. Making a good solder joint takes about half a second and certainly not more than one second. You want to remove the soldering iron from the joint before an excessive amount of heat is run into the joint. It is recommended that a good mechanical joint be made before soldering when connecting a wire to some form of terminal (this is often not possible).

The technique which I use is to stand the solder up on the workbench and bend the end so that it is sloping downwards towards me. The lead of the component to be soldered is placed in the hole in the strip board and gripped just above the board with long-nosed pliers. The board is turned upside down and the left thumb used to clamp the board against the pliers. The board and pliers are then moved underneath the solder and positioned so that the solder lies on the copper strip, touching the component lead. The right hand is now used to place the soldering iron briefly on the solder. This melts the solder on the joint, allowing the flux to clean the area and producing a good joint. After the joint is made, the board is still held with the pliers until the joint has cooled down.

Test Equipment.

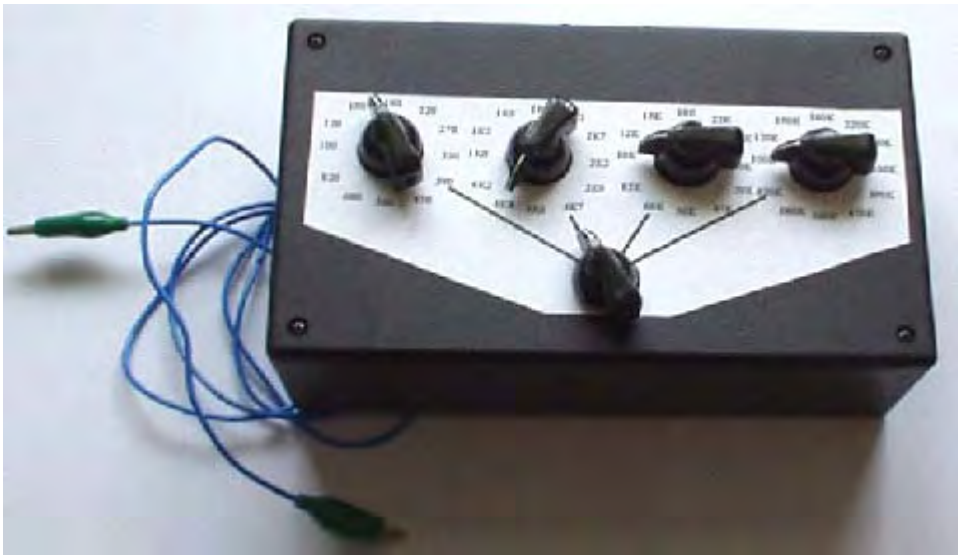
When developing new circuitry, it may be convenient to try different values of resistor in some position in the circuit (the resistor value may be dependent on the gain of a transistor or the actual resistance of an ORP12, or some such other situation). For this, it is very convenient to have a resistor-substitution box which allows you to select any standard resistor at the turn of a switch.

These are not readily available on the market. In years gone by, it was possible to buy custom wafer switches, where the number of wafers could be built up to whatever switch size was required, but these do not seem to be available any more. A slightly less convenient method of construction is to use four of these, selected by a second wafer switch:



In the above diagram, all of the resistors in one range (100 ohms to 820 ohms, 1K to 8K2, 10K to 82K or 100K to 820K) are wired to a single 12-way switch. The output wires then have any of these standard resistors across

them, depending on the setting of the switch. A second switch can then be used to select several of these groups, while still using the same output wires. When boxed, it might look like this:



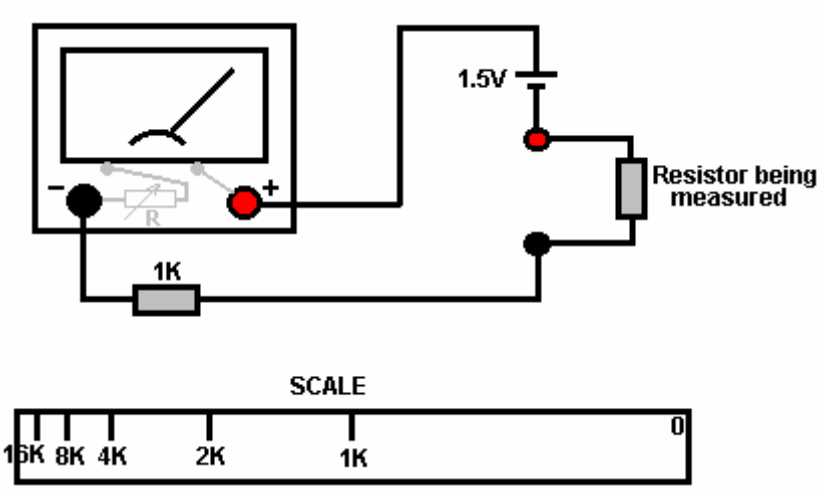
It can also be useful to have a versatile signal generator. You can easily construct your own with variable frequency, variable mark/space ratio and optional variable gating. If you do, you might as well make it with a low output impedance so that it can drive devices under test directly rather than having to provide additional buffering. It might look like this:

The really essential item of equipment is a multimeter. These come in many shapes, sizes and varieties and the cost varies enormously. The reliability also varies a great deal. The most reliable and the cheapest is the analogue type which does not use a battery (other than for the occasional measurement of resistance). Although these types are looked down upon nowadays, they are 100% reliable:

The meter shown above is rated at 2,000 ohms per volt, so connecting it to a circuit to make a measurement on the 10V range is the same as connecting a 20K resistor to the circuit. The big brother of this style of equipment is about five times larger and has 30,000 ohms per volt performance, so connecting it on a 10V range is the same as connecting a 300K resistor to the circuit being measured. This one is battery driven, so if you get one of these, may I suggest that you check its accuracy on a regular basis:

The really excellent non-battery (ex-professional) Avo meter multimeters are still available through eBay at affordable prices. These have 30,000 ohms per volt performance and are robust and accurate, having been built to very high standards.

A multimeter uses a 1.5V battery to measure resistance. Ohm's Law is used as the working principle and the operation is:



The meter shown in the diagram has a small resistance of its own. This has a small variable resistor added to it. This variable resistor will have a small knob mounted on the face of the multimeter, or it will be a thumbwheel knob projecting slightly from the right hand side of the multimeter case. The 1.5V battery will be positioned inside the multimeter case as is the 1K resistor. To use the resistance ranges, the multimeter probes are touched firmly together to form a short-circuit and the variable resistor adjusted so that the meter points to zero.

For the purpose of this discussion, let us assume that the internal resistance of the meter, when correctly adjusted, is exactly 1K. If the resistor under test is exactly 1K in value, then the current through the meter will be halved and the meter will show a needle deflection half way across the scale. If the resistor under test is 2K, then the current will be one third and the scale marking will be at the 1/3 position from the left. If the resistor is 4K, then there will be one fifth (1K+4K=5K) of the full-scale current and the 4K mark will be 20% from the left hand side of the scale.

Two things to notice: firstly, the scale has to read from right to left which can take some getting used to, and secondly, the scale is not linear, with the markings getting closer and closer together and consequently, more difficult to mark and read, the higher the value of the resistor being measured. The bunching up of the scale markings is why the more expensive multimeters tend to have more than one range.

A mains-operated oscilloscope is an excellent piece of equipment to own but they are expensive when new. It is possible to pick one up at a reasonable price second-hand via eBay. An oscilloscope is by no means an essential item of equipment. One of its most useful features is the ability to measure the frequency, and display the shape of a waveform. Most waveforms are of known shape so the frequency is the major unknown. The following meter is not expensive and it displays the frequency of a signal on a digital readout:

So, when you are deciding what multimeter to buy, consider the following points:

1. How reliable is it? If you are opting for a battery driven unit, what happens to the accuracy if the battery starts to run down. Does it display a warning that the battery needs to be replaced? Mains-operated digital multimeters are brilliant but are a problem if you want to make measurements away from the mains.

2. What DC voltage ranges does it have? If you are intending to work mainly with 12V circuits, it is inconvenient for the ranges to be 9V and 30V as successive ranges. Digital meters do not have this problem but the question then is, how accurate are they going to be in day to day use?

3. Transistor testing options you can ignore - you are better off making your own dedicated unit to check transistors if you think you will ever need to do this - you probably won't.

4. Measuring current can be very useful so see what ranges are offered.

5. Measuring capacitance is very useful, especially since many capacitors are not well marked to indicate their value.

6. Measuring the frequency of a waveform could be a significant bonus but the question is; are you every likely to need it?

7. Measuring resistance is very useful. Every meter has it. There is no need to be over fancy on measurement ranges as you usually only need to know the approximate answer - is it a 1K resistor or a 10K resistor?

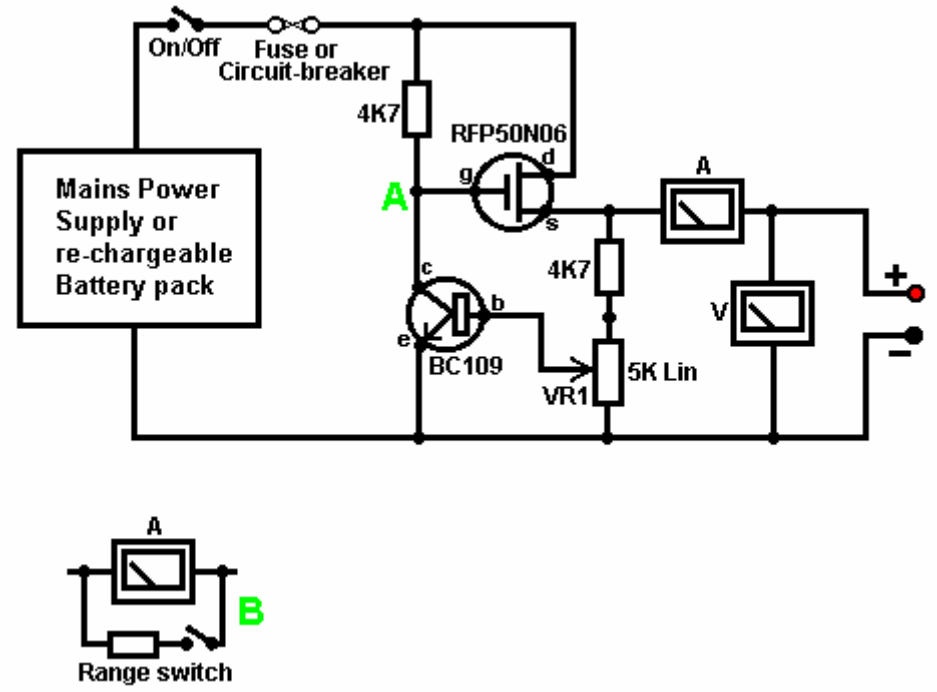
Look around and see what is available, how much it costs and what appeals to you. It might not be a bad idea to buy a really cheap multimeter and use it for a while to see if it has any shortcomings which are a nuisance, and if so, what improvements you personally want from a more expensive meter.

The 'Bench' Power Supply.

It might be worth getting a fancy bench power supply which allows you to set any voltage you want and which displays the current being drawn by your development circuit:

However, there is no need to spend money on a fancy unit when you can build an excellent unit of your own with voltage stabilisation, adjustable output, metered current, etc. etc. Personally, if developing a circuit to be used with a battery, I believe you are better off powering the development from a battery, that way the characteristics of the battery are included in any tests which you carry out.

If you wish, you can construct a very convenient development test bed power supply system. This has the advantage that you can make it in the most convenient style for your own use. You can also make the protection ultra-sensitive and build in additional circuitry such as transistor tester and resistor substitution box to produce an integrated test bed. You could perhaps use a circuit like this:



Here, the power is supplied by a pack of re-chargeable Ni-Cad batteries or possibly, a mains unit with voltage stabilisation. As in all actual circuits, the next thing in the circuit is **always** an on/off switch so that the power source can be disconnected immediately should any problem arise. Next, as always, comes a fuse or circuit breaker, so that should the problem be serious, it can disconnect the circuit faster than you can react. If you wish, you can build your own super-accurate adjustable circuit breaker to use in this position.

The two transistors and three resistors form an adjustable, stabilised output. The FET transistor has a high output power handling capacity and a very low input power requirement and so is good for controlling the output voltage.

Resistor 'VR1' is padded with the 4K7 resistor solely to reduce the voltage across the variable resistor. VR1 is adjusted to control the output voltage. If the current draw is increased and the output voltage is pulled down slightly, then the voltage on the base of the BC109 transistor is reduced. This starts to turn the transistor off, raising the voltage at point 'A', which in turn, raises the output voltage, opposing the variation caused by the load.

The output is monitored, firstly by a large milliammeter to show the current draw and secondly, on the output side of the milliammeter, a voltmeter. This allows very close monitoring of the power supplied to the prototype, especially if the milliammeter is placed alongside the prototype. You can build this circuit into a wide flat box which provides a working surface beside the milliammeter.

At point 'B' in the above diagram, a method for altering the current range of the milliammeter by placing a 'shunt' resistor across it. When the switch is closed, some current flows through the resistor and some through the milliammeter. This resistor has a very low value, so you are better off making it yourself. Let's say we wish to double the range of the meter. Solder the switch across the meter and for the resistor use a length of enamelled copper wire wound around a small former. Put a load on the output so that the meter shows a full-scale deflection. Close the switch. If the current displayed is exactly half of what it was, if not, switch off, remove some wire to lower the reading or add some wire to raise the reading and repeat the test until exactly half the current is displayed. The lower the value of the shunt resistor, the more current flows through it and the less through the meter, which then gives a lower reading.

Please note: it is very important to have a fuse or circuit breaker in the power being delivered to your test circuit. Any error in building the prototype can cause a major current to be drawn from the supply and **this can be dangerous**. Remember, you can't see the current. Even if you have a meter on the current being delivered, you may not notice the high reading. The first sign of trouble may be smoke! You can easily fry the circuit you are building if you do not have a safety cut-off, so use a fuse or other device which limits the current to twice what you are expecting the circuit to draw.

So, after all that, what equipment do you really need? You need a small soldering iron and multicore solder, a pair of long-nosed pliers and a multimeter. One other thing is some tool to cut wires and remove the insulation prior to soldering. Personal preferences vary. Some people prefer one of the many custom tools, some people use a knife, I personally use a pair of straight nail scissors. You pick whatever you are comfortable with.

Not exactly a vast array of essential equipment. The other items mentioned are not by any means essential so I suggest that you start by keeping things simple and use a minimum of gear.

If you are not familiar with electronics, I suggest that you get a copy of the Maplin catalogue, either from one of their shops or via the http://www.maplin.co.uk web site. Go through it carefully as it will show you what components are available, how much they cost and often, how they are used. The specifications of almost any semiconductor can be found free at http://www.alldatasheet.co.kr in the form of an Adobe Acrobat document.

Finally, because it is not important, all of the circuitry shown so far has indicated current flowing from the + of a battery to the - terminal. The discovery of voltage was made by Volta but he had no way of knowing which way the current was flowing, so he guessed. He had a 50 - 50 chance of getting it right but he was not lucky and got it wrong. Electrical current is actually a flow of electrons, and these flow from the battery minus to the battery plus. So, who cares? Almost nobody, as it has no practical effect on any of the circuitry. Some useful websites:

http//:www.esr.co.uk for components
http//:www.maplin.co.uk for components
http//:www.alldatasheet.co.kr for semiconductor specifications
http//:www.cricklewoodelectronics.com for components
http//:www.greenweld.co.uk for components

The Oscilloscope.

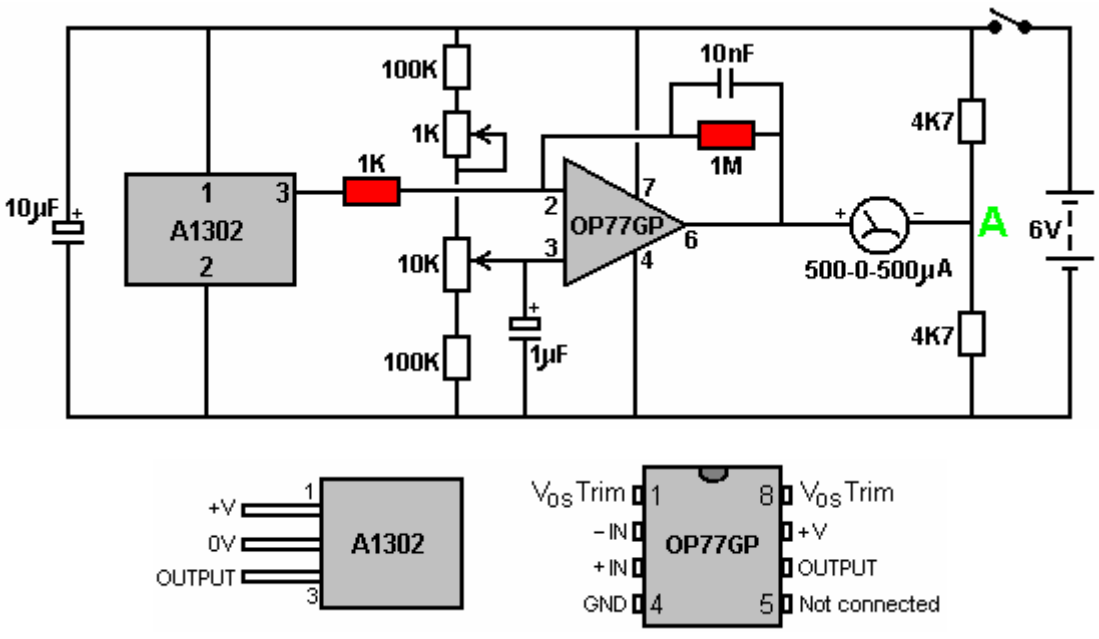
If you do decide that you are going to research new equipment, design and possibly invent new devices, then an oscilloscope is useful. Let me stress again that this is not an essential item of equipment and most certainly is not needed until you are quite familiar with constructing prototypes. It is quite easy to misread the settings of an oscilloscope and the methods of operation take some getting used to. The low-cost book "How to Use Oscilloscopes and Other Test Equipment" by R.A. Penfold, ISBN 0 85934 212 3 might well be helpful when starting to use a 'scope.

It is possible to get an oscilloscope at reasonable cost by buying second-hand through eBay. The best scopes are 'dual trace' which means that they can display the input waveform and the output waveform on screen at the same time. This is a very useful feature, but because it is, the scope which have that facility sell at higher prices.

The higher the frequency which the scope can handle, the more useful it is, but again, the higher the selling price. Not all scopes are supplied with (the essential) 'test probes', so it might be necessary to buy them separately if the seller wants to keep his. Getting the manual for the scope is also a decided plus. A low cost scope might look like this:

Measuring Magnetic Field Strength.

People who experiment with permanent magnets, can make use of an instrument which displays the strength of a magnetic field. Professionally made devices to do this tend to be well outside the purchasing power of the average experimenter who will already have spent money on materials for his prototypes. Here is a design for a simple and cheap circuit, powered by four AA dry cell batteries, and utilising a Hall-effect semiconductor as the sensor:



This design uses an OP77GP operational amplifier chip to boost the output signal from the A1302 chip which is a Hall-effect device. The gain of the DC-connected operational amplifier is set by the ratio of the 1K and 1M fixed resistors shown shaded in the circuit diagram, giving a gain of 1,000.

The circuit operation is simple. The six-volt battery charges the 10 microfarad capacitor which helps iron out any supply line fluctuations caused by varying current draw by the circuit. The 10K variable resistor is used to set the

output meter display to zero when the Hall-effect device is not near any magnet. The 1K variable resistor is there to make fine tuning adjustments easier.

When the A1302 chip encounters a magnetic field, the voltage on it's output pin 3 changes. This change is magnified a thousand times by the OP77GP amplifier. It's output on pin 6 is connected to one side of the display meter and the other side of the meter is connected to point "**A**". The voltage on point "**A**" is about half the battery voltage. It would be exactly half the voltage if the two 4.7K resistors were exactly the same value. This is rather unlikely as there is a manufacturing tolerance, typically around 10% of the nominal value of the resistor. The exact value of the voltage on point "**A**" is matched by the OP77GP tuning and so the meter reads zero until a magnetic field is encountered. When that happens, the meter deflection is directly proportional to the strength of the magnetic field.

The Weird Stuff.

You don't need to know the following information, so please feel free to skip it and move on to something else.

The presentation shown above is based on the conventional view of electronics and electrical power as taught in schools and colleges. This information and concepts works well for designing and building circuits, but that does not mean that it is wholly correct. Unfortunately, the world is not as simple as is generally made out.

For example, it is said that current is a flow of electrons passing through the wires of a circuit at the speed of light. While it is true that some electrons do actually flow through the metal of the wires, the small percentage of electrons which actually do that, do it quite slowly as they have to negotiate their way through the lattice of the molecules of metal making up the body of the wires.

In spite of this, when the On/Off switch of a circuit is flipped on, the circuit powers up immediately, no matter how long the wires are. The reason for this is that electrical current flows along the wires at very high speed indeed, but it flows rapidly **along** the outside of the wires, not rapidly **through** the wires. One thousandth of a second after switching on a circuit, the electrons flowing through the wires have hardly got started, while the current flowing along the outside of the wires has gone all around the circuit and back:

The above sketch does not show the proportions correctly, as the current flow spiralling along the outside of the wire should be hundreds of thousands of times longer than shown, which is not practical in a diagram.

The actual path taken by current flow makes the surface of the wire of particular importance, and the insulation material is also of great importance. In years gone by, wire manufacturers used to anneal (cool down) copper wires in air. This created a layer of cupric oxide on the outer surface of copper wires, and that layer gave the wire different characteristics than copper wire has today. William Barbat in his patent application claims that the cupric oxide layer can be utilised in making devices with greater power output than the power input from the user.

Unfortunately, the world is not quite as simple as that, as power flowing in a circuit has at least two components. The electrical current which we measure with ammeters is as described above and is sometimes referred to as "hot" electricity as when it flows through components, it tends to heat them up. But there is another component referred to as "cold" electricity, so named because it tends to cool components down when it flows through them. For example, if the output wires of Floyd Sweet's VTA device were short circuited together, frost would form on the device due to the heavy flow of "cold" electricity, and getting a "shock" from it could give you frostbite instead of a burn.

"Cold" electricity is not something new, it has always been there as it is just one aspect of "electricity". It has not been investigated much by conventional science because none of the instruments used to measure "hot" electricity, react to "cold" electricity at all. (Actually, "hot" electricity, "cold" electricity and magnetism are all features of a single entity which should really be called "electromagnetism").

Now the spooky bit: "cold" electricity does not flow along or through the wire at all. Instead, it flows in the space around the wire, possibly riding on the magnetic field caused by the "hot" current. Thomas Henry Moray is famous for building a device which captured "cold" electricity and produced a massive power output capable of powering a whole host of ordinary electrical pieces of equipment. In his many public demonstrations before he was intimidated into silence and his equipment smashed, he invited members of the audience to bring a piece of ordinary glass with them. Then, when his circuit was powering a row of lights, he would cut one of the wires and

insert the piece of glass between the cut ends of the wires. This had no noticeable effect on his circuit, with the power flowing happily through the glass and on through his circuit, powering the lights just as before. That does not happen with "hot" electricity, but as the "cold" electricity is not flowing through or along the surface of the wire, a break in the wire is not a major obstacle to it.

We still do not know very much about "cold" electricity. Edwin Gray snr. demonstrated light bulbs powered by "cold" electricity being submerged in water. Not only did the bulbs continue to operate unaffected by the water, but Edwin often put his hand in the water along with the lit bulb, suffering no ill effects from doing so. Neither of those two effects are possible with conventional electricity, so please **don't** try them to check it out.

Another interesting item is the water-powered car system produced by an American man Nathren Armour. His system, (among other things) involves feeding extra electrical power to the spark plugs. One thing which has always puzzled him is that the engine will not run with just one wire going to the spark plug cap. He has to have a second wire running from his extra power supply to the body of the plug where it screws into the engine block. Take that wire away and the engine stops. Put it back again and the engine runs. But according to conventional electrics, that wire cannot possibly be needed, because the engine block is grounded and the power supply output is grounded, so in theory, there is no voltage difference between the ends of the wire, therefore no current can flow along the wire, hence the wire is not needed and has no function. Well, that is true for "hot" electricity, but it seems possible that the Nathren Armour system is using "cold" electricity as well as "hot" electricity and the "cold" electricity needs the extra wire as a flow guide to the spark plug.

Enough about that for now. Let's go one step further into the "weirdness" of the actual world. If, three hundred years ago, you had described X-rays, gamma rays, nuclear energy and TV signals to the average well-educated person, you would have run a considerable risk of being locked up as being mad. If you do it today, your listener would probably just be bored as he already knows all this and accepts it as a matter of fact (which it is). Please bear that in mind when you read the following information. If it seems strange and far-fetched, that is only because conventional science today is lagging badly behind and still teaching things which have been conclusively proven to be wrong decades ago.

If you lived in a desert and every day a company drove in with a lorry-load of sand and sold it to you for a large amount of money, what would you think about that? Not a very good deal for you, is it? What's that you say, you would never do that? But you already do, because you don't realise that the sand is all around you ready for the taking at next to no cost at all. Several people have tried to publicise the fact, but the sand company has immediately silenced them by one means or another. The company does not want to lose the business of selling you the sand and definitely doesn't want you to start picking it up for yourself for free.

Well... to be perfectly fair, it is not actually sand, it is energy, and it is all around us, free for the taking. Sound a bit like X-rays did three hundred years ago? Doesn't mean that it is not true. It is perfectly true. The design of all computers made today is based on the equations of Quantum Mechanics, and while those equations are not yet perfect, they are easily good enough for practical purposes. The snag is that the world seen at the level of the quantum is not much like the world we think that we see around us and which we think that we understand fully. Examining the world at the quantum level shows that we live in a seething mass of incredible energy. Einstein is famous for stating that Mass equals a very large amount of Energy, a fact that is shown clearly when an atomic bomb is detonated. Put in different words, a small amount of matter is the equivalent of a very large amount of energy. Actually, Energy and Matter are two different aspects of a single thing (which could reasonably be called "Mass-Energy").

At the quantum level, it can be seen that particles of matter pop into existence and drop out again into energy on a continuous basis, everywhere in the whole of the universe. The whole universe is seething with energy. That energy doesn't bother us any more than water bothers a fish, as we evolved in this sea of energy and we just don't notice it. It doesn't harm us, but if we wanted, and knew how, we could use as much of that energy as we wanted for ever and ever. The amount of that energy is unbelievable. It has been calculated that one cubic centimetre anywhere in the universe contains enough energy to create all of the matter we can see in the whole of the universe. Think how many cubic centimetres there are in the Earth ... the Solar System ... our Galaxy ... If every person on Earth were to run their vehicles, power their homes, fly their planes, etc. etc. for the next million years, it would not make the slightest dent in the energy contained in one cubic millimetre of the universe. This is not a theory, it is a fact. (Would you like to buy a big pile of sand? - I've got a load just over here...). This big energy field has gone under different names over the years. A popular name at the present time is the "Zero-Point Energy Field" and it is responsible for everything that happens in the universe. It powers life itself. It balances out in equilibrium everywhere, which is one reason which makes it hard to realise that it is all around us.

Tom Beardon is an American man with very considerable abilities and considerable in-depth knowledge of how the world actually operates. His statements are generally based on laboratory-proven criteria backed up by his high level of mathematical skills which give him an additional grasp of things. He explains how electricity actually works in circuits, and it is nothing like the system taught in schools and colleges. We think that when we attach a

battery to an electrical circuit, the battery forces a current through the wires of the circuit. Sorry Chief - it is actually nothing like that at all. The power in the circuit comes directly from the Zero-Point Energy Field and has very little to do with the battery at all. We tend to think of "using up" power, but that is just not possible. Energy cannot be destroyed or "used up" the most you can do to it is to change it from one form to another. It will perform "work" (power equipment, generate heat, generate cold...) when it changes from one form to another, but if you reverse the process and convert it back to it's original form, it will perform another lot of "work" during the conversion and end up back in exactly the same state as it started out from, in spite of having performed two lots of "work" during the operation.

A battery does **not** provide energy to power a circuit. Instead, what happens is that the chemical action inside the battery causes negative charges to gather at the "minus" terminal of the battery and positive charges to gather together at the "plus" terminal of the battery. These two close-together "poles" of the battery are called a "dipole" (two opposite poles near each other) and they have an effect on the Zero-Point Energy Field which is everywhere. The "Plus" pole of the battery causes a massive cluster of Zero-Point Energy Field negative charges to cluster around it. In the same way, the "Minus" pole of the battery causes a massive gathering of ZPE ("Zero-Point Energy") positive charges to gather around it. Not only do these charges gather around the poles of the battery, but an imbalance in the energy field is created and the ZPE charges continue to arrive at the poles and they radiate out in every direction in a continuous stream of incredible energy.

So, there is your shiny new battery sitting there, not connected to anything and yet it causes massive energy streams to radiate out from its terminals in every direction. We don't notice it, because the energy flows freely through us and we can't feel it and none of our conventional instruments, such as voltmeters, ammeters, oscilloscopes, etc. react to it at all.

The situation changes immediately if we connect a circuit to the battery. The circuit provides a flow path for the ZPE energy to flow along, and a significant amount of energy flows near the wires of the circuit, actually powering the circuit for a split second until it reaches the battery "pole" at the far end of the circuit. When it gets there it promptly wipes out the pole, destroying it completely. The ZPE field calms down and the energy flow ceases. But our trusty battery immediately does it all again, using it's chemical energy to create the "dipole" once more, and the imbalance of the ZPE field starts again. It is because the battery has to use it's chemical energy all the time, creating and re-creating, and re-creating it's "dipole" that it runs down and eventually ceases to be able to create the dipole any more - result: no more power in the circuit.

Sorry to spoil the illusion, but the battery never did power the circuit itself, it merely acted as channelling device for the Zero-Point Energy Field. In passing, Direct Current ("DC") is actually not a continuous current at all, but instead it is a stream of DC pulses at an incredibly high frequency - way higher than we can measure at present. The speed of the pulses is so great that it looks continuous to us, a bit like the individual still pictures which are the frames of a movie, appear to be a moving image to us if they are played one after the other at a rate of 25 per second - it looks like continuous movement to us, but in reality, it is a rapid series of still pictures.

The way that a battery "dipole" works on the Zero-Point Energy Field is rather like the way that a magnifying glass acts on sunlight. The rays of the sun get concentrated into a point, focused by the lens. You can start a fire with the lens, and it would be easy to think that the lens started the fire, when in actual fact, it is the rays of the sun that started the fire and the lens just influenced a local area of the large "field" of sunlight, raising the temperature at just one point.

While we tend to think of a "dipole" being generated by a battery, the same effect is also created by a magnet, whether an electromagnet or a permanent magnet - remember that electricity and magnetism are two faces of the same entity. It is possible, but not easy, to capture the energy streaming out from the interference with the ZPE field caused by the poles of a magnet. For example, Hans Coler managed to do this with a completely passive device which, when set up correctly, could produce electrical power, hour after hour from apparently "nothing" (well, actually, the ZPE field). Roy Meyers also did it with his patented array of magnets and zinc plates - completely passive, with no moving parts at all, no battery and no circuitry.

Patrick Kelly

http://www.free-energy-info.tuks.nl/

Chapter 13: Doubtful Devices

This chapter covers a number of devices which either are unlikely to work, or which have too little practical information available to assist replication attempts. This selection, is of course, a matter of opinion.

Paul Baumann's "Thestatika" Machine.

This device is a perfect example of a free-energy device as it powers itself and provides kilowatts of excess mains electrical power. It is in this section, not because its operation is "doubtful" in any way, but because the design has never been fully disclosed. It was developed by the late Paul Baumann who was part of a Swiss commune which is not willing to explain its operation. This "Thestatika" or "Testatika" machine works beautifully and has a very high quality of workmanship. It has two electrostatic discs which are initially rotated by hand and which then continue to rotate driven by the power produced by the device.

There are various ideas as to how the device operates. The Swiss commune no longer shows this device to people as they have the theory that "mankind" is not ready to have, or use free-energy. They have always refused to show what is inside the large cylinders mounted on each side of the device. D. A. Kelly's 1991 document provides some very perceptive comments on this device. He says:

The "Swiss M-L Converter" is a fully symmetrical, influence-type energy converter which is essentially based on the Wimshurst electrostatic generator with its twin counter-rotating discs where metallic foil sectors generate and carry small charges of electricity to be stored in matched capacitors. In Wimshurst units, diagonal neutralising brushes on each opposite disc distribute the correct charges to the sectors as they revolve, but in the M-L converter this is carried out by a crystal diode which has a higher efficiency.

Two brushes collect the accumulating charges and conduct them to the storage capacitor located at the top of this device. The device has two horseshoe magnets with matched coils and a hollow cylindrical magnet as part of the diode function, and two Leyden jars which apparently serve as the final capacitor function for the converter. The use of top grade components such as gold-plated contacts, control electrodes and dual capacitor stages, insure much higher conversion efficiencies than those available with a Wimshurst machine. The details of the operating prototype are:

1. Efficiency: The unit is started by hand and no other input power is required.
2. Constant power output: 300 volts at 10 amps = 3 kilowatts.
3. Dimensions: 43.31" (1100 mm) wide, 23.62" (600 mm) tall, 17.72" (450 mm) deep.
4. Weight: 44 lbs (20 Kg).
5. Operating speed: 60 rpm. (low speed - one revolution per second).

The twin discs are made of acrylic (plastic) and the metallic segments are steel, which causes the Searle Effect with electromagnetic conversion made at the rim of the discs through passive electromagnets. This is an ideal converter since both high voltage AC and moderate AC amperage can be generated simultaneously via two separate electrical circuits from the discs. The conventional conductive brushes pick off the high voltage AC while the rim electromagnet coils produce useful amperage. When permanent horseshoe magnets with coils are used, then the output power is enhanced to a considerable extent as shown by the above output specifications.

The self-propulsion after hand-starting the discs is achieved through the adoption of the Poggendorff principle (a German scientist of the 1870s) in which slanted conductive brushes produce self-rotation in electrostatic motors (not generators).

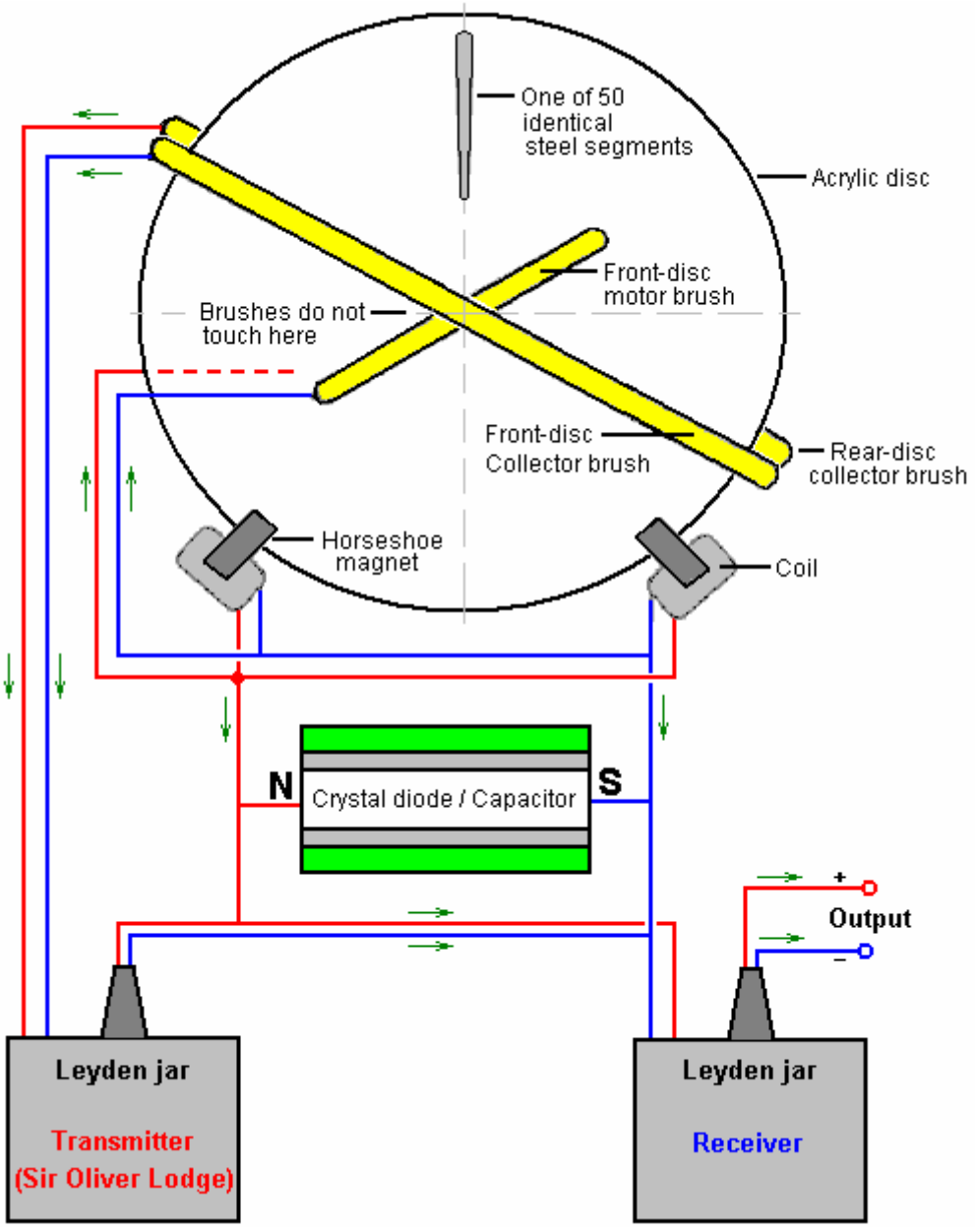
The special crystal diode module probably provides the dual functions of frequency regulation and capacitance amplifier - to the two Leyden jars - as part of the electrical resonance circuit, since it is connected with the horseshoe magnet coils.

This device is comprised of three separate electrical circuits:

1. The high voltage AC output from the twin electrostatic discs.
2. A moderate AC amperage circuit provided by the dual horseshoe magnet coils (Searle Effect) as the plus and minus discs pass by them. (Pulsed DC output at 50 Hz).
3. A resonant circuit in which the horseshoe magnet coils are connected to the diode capacitor so that frequency regulation is assured. The diode capacitor is then connected to the Leyden jar, transmitter unit.

The major physical principles involved in this outstanding composite unit are:

1. Electrostatic conversion using twin discs for positive output from one and negative output from the other.
2. The evidence of the Searle Effect from the use of multiple, identical steel segments inducing and EMF in electromagnets at the rim of the discs.
3. The Ecklin principle is also in evidence, since the steel segments pass by permanent horseshoe magnets, as in Ecklin's S.A.G. units.
4. The Poggendorff self-rotating electrostatic motor principle as described above.
5. The crystal capacitance function of the crystal diode module. The full operation of this unique component with its hollow cylindrical permanent magnet, is a composite component with the dual functions of distributing the correct charges to the sectors, and maintaining the output frequency at the desired value.



The M-L Converter is completely symmetrical with two acrylic discs, a light metal lattice, insulated copper wires, a secret crystal-diode rectifier, and gold-plated electrical connections. These machines have been developed over a period of twenty years.

In electrostatic generators, the air molecules between the two acrylic discs which counter-rotate closely side by side, become electrically activated by friction. This causes the discs to be continually charged until a flashover equalises the charge on them. To limit the voltage to the desired amount, the positively charged

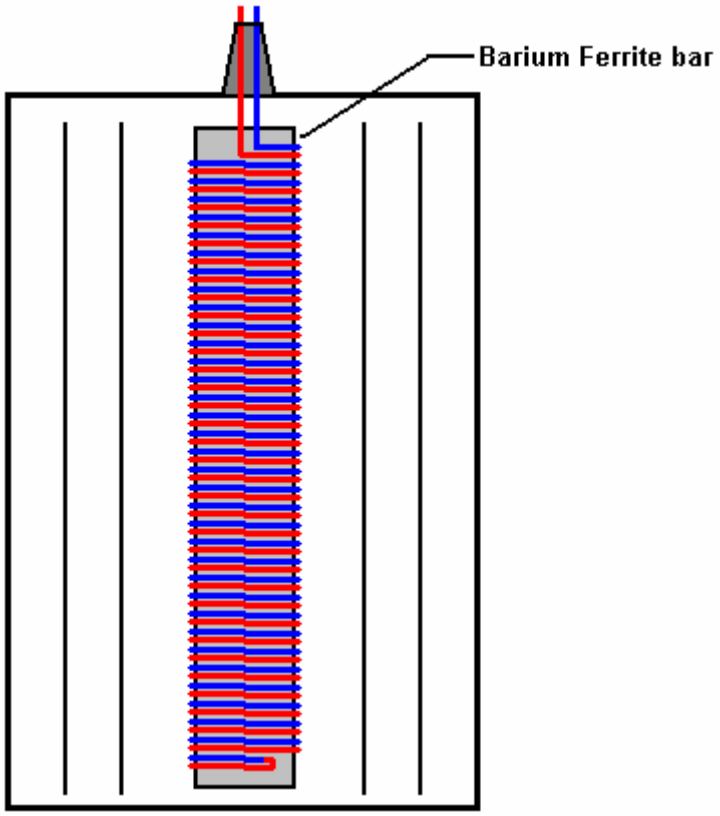
particles on one of the discs and the negatively charged particles on the other disc are each extracted by means of separately adjustable lattice-electrodes, and are fed into a Leyden jar which collects the energy. The speed of the discs which have 50 lattice electrodes, is 60 rpm which produces a 50 Hz pulsed DC output. This speed is synchronised by magnetic impulses.

The unit is hand started by revolving the two discs in opposite directions until the Converter is charged up enough to synchronise itself and continues to rotate smoothly and noiselessly without any external source of input power. A centrally mounted disc of about 4" (100 mm) in diameter glimmers with all the colours of the rainbow. After a few seconds the Leyden jars are ready for operation and 300 volts DC with a current of 10 amps can be drawn from the device for any desired length of time. On many occasions, demonstrations have been made of the power available from the device. Heating elements, lights and hand power tools can be run from the device.

This suggested explanation of the M-L Converter contains a number of very interesting points. It has seemed mysterious that the electrostatic discs continued to rotate on their own without any visible motor driving them. Mr Kelly, who has seen the device and its operation, suggests that there are sloping brushes pressing against the front and rear faces of the twin electrostatic discs and that these are supplied with electrical current from the horseshoe magnet coils and that acts as a motor which drives the discs onwards once they have been started. He also suggests that the fifty steel segments per second which pass between the poles of the horseshoe magnets cause a rapidly fluctuating magnetic field through the magnet coils, which makes them operate as an Ecklin electrical generator, as described elsewhere in this eBook.

Mr D. A. Kelly also suggests that the two cylinders seen on the M-L Converter, are Leyden jar capacitors and that they work together as described by Sir Oliver Lodge (whose book is on this website). This is a very interesting suggestion, but it does not explain why the people in the Swiss commune refuse point-blank to let anyone see what is inside those cylinders.

There is a video produced by Don Kelly (presumably, a different person) which puts forward another theory of operation. He suggests that each of the cylinders contains a bi-filar coil on a barium ferrite magnet:

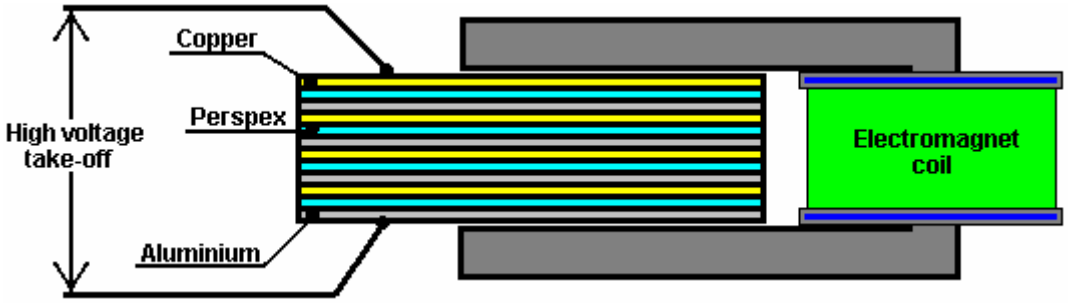


However, he describes the barium ferrite magnet as being the same type as used in radio receivers, and they are standard "ferrite rods" which are not permanent magnets as far as I am aware. Don suggests that the output from the high-voltage electrostatic discs gets fed directly to these coils and then on via a series connection to the coils around the horseshoe magnets. He envisages the bi-filar coil amplifying the current and the electrostatic discs being rotated by a standard low-voltage DC motor.

Another possibility is that the jars also contain a spark gap and surrounding copper pick-up shells and as the machine operates silently, the jars have a vacuum inside them. That would provide silent operation and explain why the people in the commune could not open them for inspection. It seems very clear that we just don't know exactly how this device operates.

One very interesting fact which has been reported by the Swiss group is that if a series of copper, aluminium and Perspex sheets are placed in a magnetic field, they generate a high voltage. This is worth investigating. It is not clear if the magnetic field should be constant or oscillating. The sequence of plates is said to be: cpacpacpacpa ("c" being copper, "p" being 'Perspex' (acrylic or 'Plexiglas') and "a" being aluminium).

The following set-up might be worth investigating:

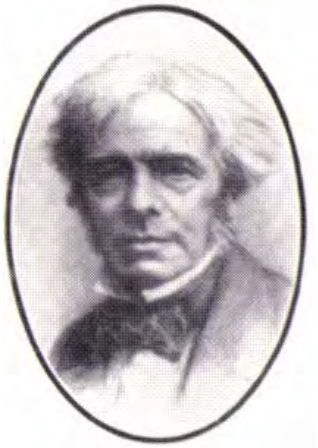


There is good information on the Testatika at http://peswiki.com/index.php/PowerPedia:Testatika but unfortunately, the bottom line is that nobody knows how to replicate Paul Baumann's excellent machine.

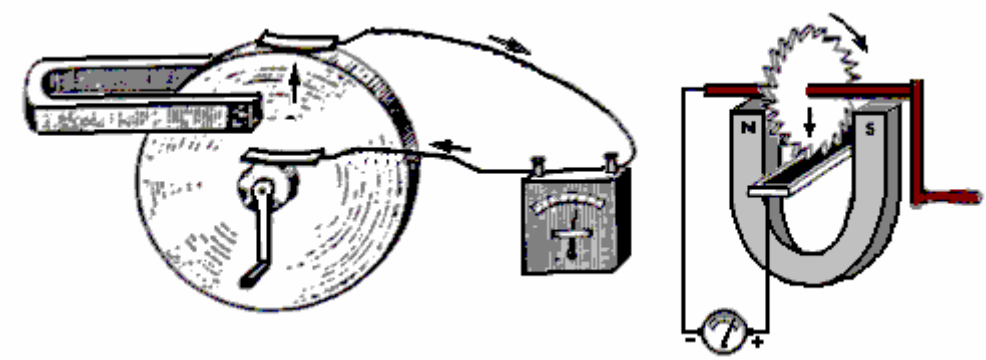
If you wish to understand the operation of electrostatic discs, then the McGraw-Hill book "Homemade Lightning" by R.A. Ford (ISBN 0-07-021528-6) gives full details of Wimshurst machines and plans for constructing your own, improved version. Ready-built Wimshurst machines are available from the web site: http://scientificsonline.com/product.asp?pn=3070070&bhcd2=1154180654

The Homopolar or "N-Machine".

This device was the brainchild of Michael Faraday in 1831 and has an intriguing method of operation and a remarkably large output.



The principle of operation is incredibly simple:

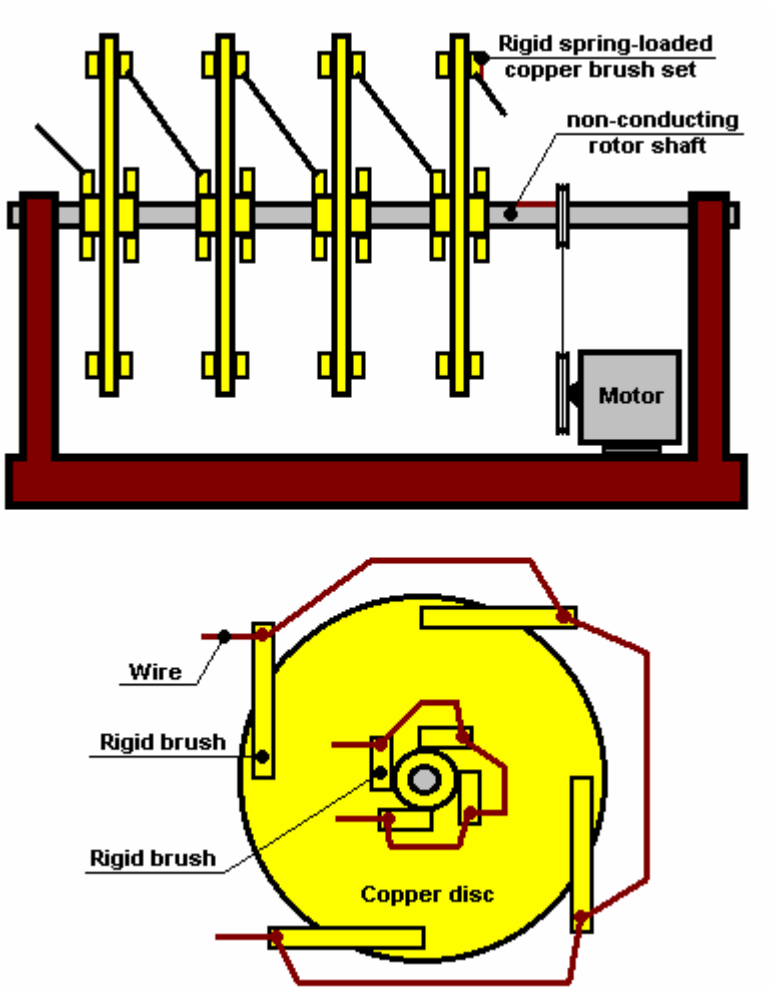


If a copper disc is rotated in a magnetic field, then power is developed between the shaft and the outer edge (or any intermediate position). It was then found that the device will still operate even if the magnet is attached to the copper disc and rotates with it - not something which is intuitively obvious. The power output is tremendous with the capability of extracting 1000 Amps but at a low voltage of less than 1 Volt. The power take-off can be from one face of the disc near the shaft rather than having to have a copper shaft integral with the copper disc. This device also works with a magnet just attached to the copper disc and rotating with it.

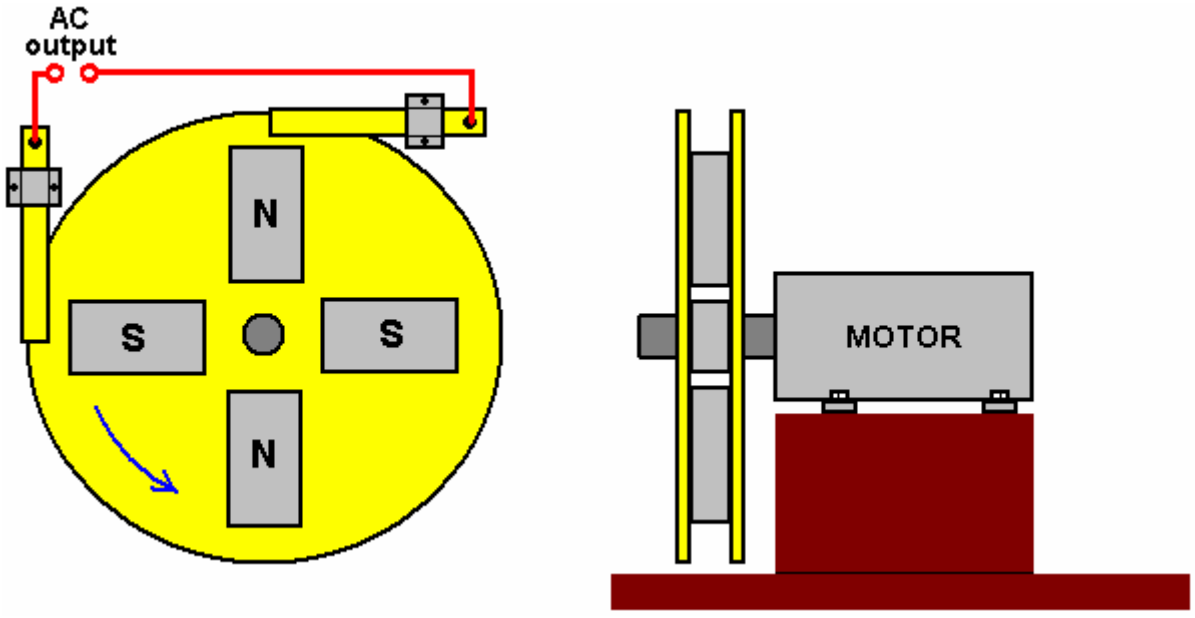
This looks like a very viable starting point to develop a device which can run itself and provide useful additional output, since a motor to rotate the disc will not require anything remotely like 1000A to drive it. The snag is, it is very difficult to provide reliable sliding contacts capable of handling large currents for extended periods of time. The second picture above shows the disc with its outer edge immersed in a bath of mercury. This is sufficient for a brief demonstration at low power but not realistic for a serious working device.

It might just be possible to get a reasonable working device by accepting that the current output is not going to be anything like 1000A. Long-life brushes could be made from solid copper bar and spring-loaded against the copper disc in matching pairs so that the brush thrusts oppose each other and so do not generate a sideways load. These could be made in multiple sets for each disc, say four or eight per disc, so that the effective electrical resistance between the brushes and the disc is reduced and the possible current draw increased.

Similar multiple brushes could be applied to the central shaft cylinder. Multiple discs could then be mounted on a non-conducting, non-magnetic shaft and their brushes wired in series as shown, to raise the output voltage:



However, when this is done, the brush drag is a serious problem. In 1987, three of the Borderlands Science team, Michael Know, Peter Lindemann, and Chris Carson experimented with this design and found that a much more satisfactory version could be produced. Their version produces sawtooth AC instead of DC and so the output could be fed directly into a step-up transformer. Their design has four ferrite magnets glued between two metal discs, and for additional mechanical strength, copper wire wound around the outer edges of the magnets in order to prevent magnets flying outwards if the glue bond should fail. Their arrangement is like this:

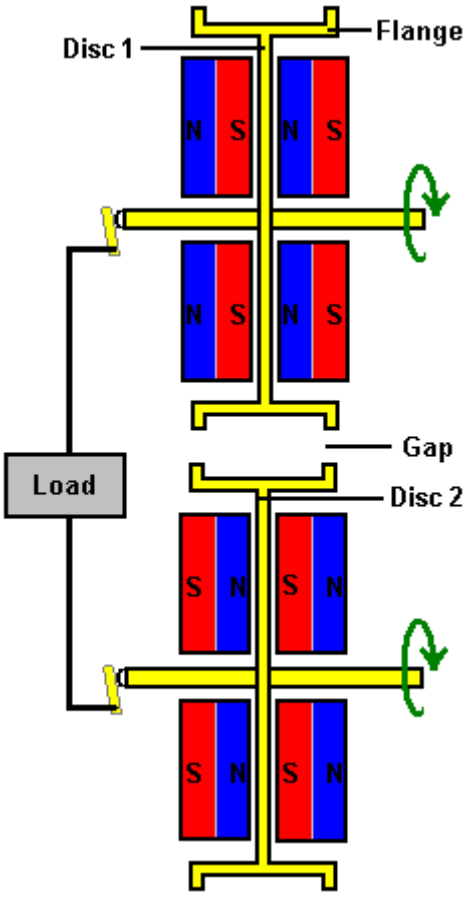


This appears to contradict the “laws” of conventional electricity as there is a very low resistance short-circuit directly across the brushes which pick up the AC voltage output. The output current from a small prototype was estimated at 100 amps. The frequency of the AC is directly proportional to the shaft speed of the motor, **but** the output voltage was almost independent of the shaft speed of the motor, increasing only very slightly with much greater speed. It was also found that putting the brushes at 90 degrees apart on the metal shaft of the motor gave the same output in spite of the contacts nearly touching each other. This design appears to have considerable potential for construction in a larger size and further investigation.

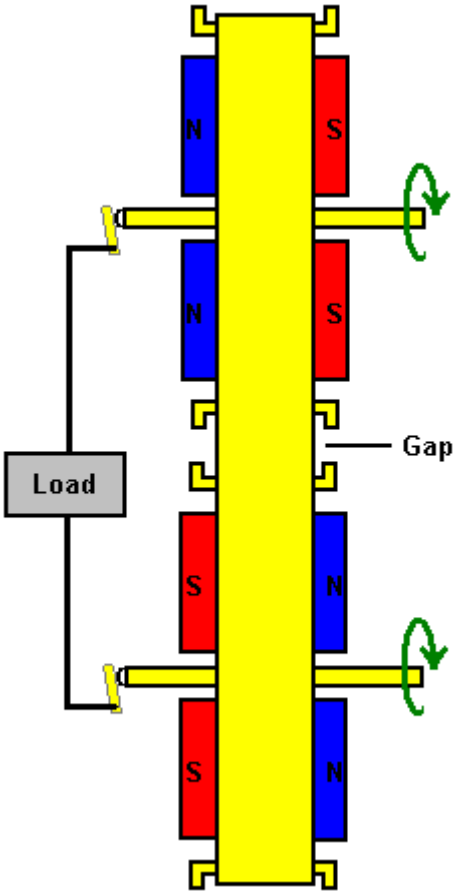
While the operation of these devices looks impossible at first glance, it needs to be understood that copper has some very unusual characteristics when interacting with magnetic fields. This is explained in the web site http://magnetism.vfedtec.com/SpinningCylinder.htm where it can be seen that a spinning copper cylinder exerts a large sideways force on a permanent magnet placed near it. This does not happen with spinning cylinders made from other metals.
Paramahamsa Tewari received an Indian patent (397/Bom/94) in 1994 for a COP=2.5 version, details of which can be seen at http://www.rexresearch.com/tewari/tewari.htm.

Nikola Tesla took Faraday’s 1831 design further as can be seen from his 1889 US patent No. 406,968. He remarked that to get any kind of useful power from the device would require a copper disc of very large diameter, or a disc which is spun very fast. A large copper disc would be an inconvenient size, and a high rate of rotation makes it very difficult to get a good, long-lasting, sliding contact at the outside edge of the disc. He also pointed out that current flowed from the shaft out to the outside edge if the magnetic field passing through the disc was in one direction, but if the direction of the magnetic field were reversed, then the current flow would be from the outer edge inwards to the shaft. The same change of direction of the current flow also happens if the direction of rotation of the disc is reversed.

Using those facts and considerable ingenuity, Tesla proposed an arrangement where the power take-off is from the axle alone, by using two separate copper discs and magnetic fields which moved in opposite directions. This arrangement has the advantage that it has an output voltage which is the sum of the two separate voltages. This basic concept uses four ring-shaped magnets and two discs of copper, brass or iron. Both of the discs are given a wide flange, as shown here:



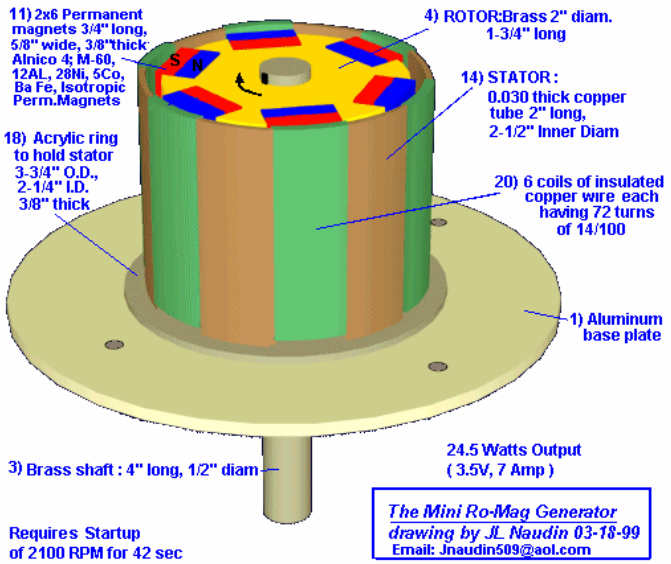
The circuit shown here is broken by the gap between the discs and Tesla dealt with that by using a flexible metal belt linking the two discs together:



While it is possible to use the belt to drive one of the discs, Tesla did not use that method. The belt overcomes the need for a sliding contact at the outside edge of the discs and so both sliding contacts are at the axles which is an easy place to have a sliding contact. Tesla shows the contact against the end of the axles as that is just a rotary movement with respect to the stationary contact, but even if the contact pressed against the outer face of the axle, the sliding movement would still be relatively slow. In spite of this clever design from Tesla, I have never heard of anybody building this style of generator in spite of the large currents which it can generate.

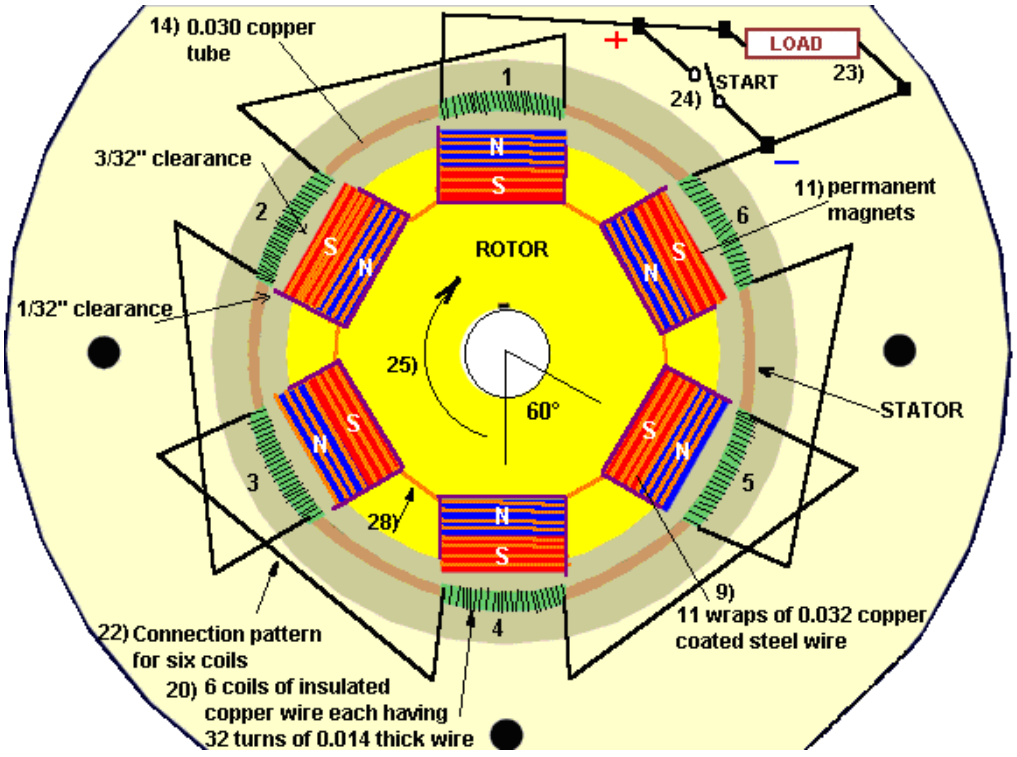
The 'Romag' and 'Mini-Romag' Generators.

These generators have been displayed on the internet for some considerable time now. They can be found on the Jean-Louis Naudin website:
http://jnaudin.free.fr/html/mromag.htm

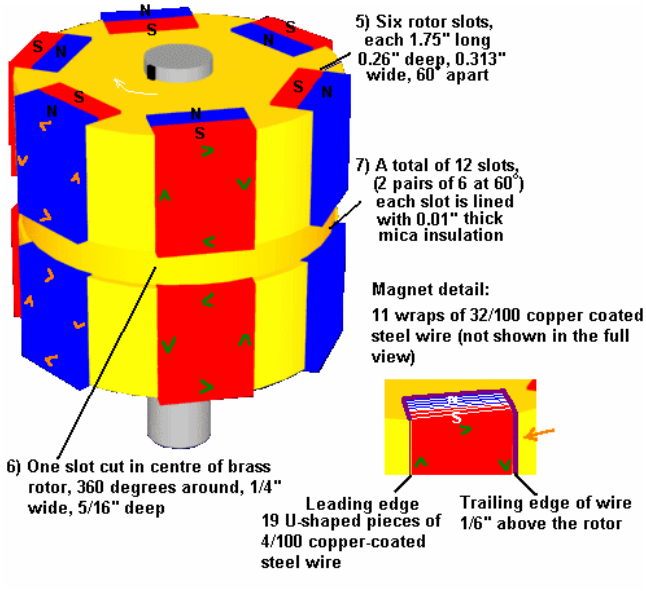


The Mini Romag generator from Magnetic Energy uses the principle of moving magnetic flow named "the magnetic current" for generating electrical power. According to Magnetic Energy this generator is able to produce 3.5 volts, 7A DC (24 Watts) of free electricity plus sufficient power to sustain itself.

This generator needs to be started by using an external motor to rotate it at 2,100 rpm for some 42 seconds. After this, the energy flow is established in the Romag generator and the external motor can be removed and the free electrical energy output can be used.



The starting procedure generates magnetic energy within the six coils of copper wire, the copper tube supporting these coils and the copper coated steel wires wrapped around the magnets. This charging is accomplished while the six coil connection wires, (shown as 22 in the above drawing), are making contact and setting up their alternating magnetic poles. After the 42 second start-up time one of these coil connection wires is opened by switch (24 above) leaving the working load in its place. The load (23 above) can draw 7 amps. As current is drawn from the six coils, it sets up magnetic poles which react with the rotor magnets maintaining the rotation. The main shaft is rotated by the 12 permanent magnets as they attract and build a release field. Then the driver unit (hand crank or motor) is disconnected allowing the unit to continue rotating with the load being the **activating** driving force.



Construction:

If you decide to attempt to build one of these units we suggest using the stated materials:

1. Aluminium Base Plate
2. Sleeve Bearing of oil impregnated brass, 1" long, 0.5" inside diameter.
3. Brass Shaft, 4" long, 0.5" outside diameter
4. Rotor, brass 1.75" long, 2" diameter,
5. Six rotor slots, each 1.75" long, 0.26" deep, 0.72" wide. These slots are spaced exactly 60 degrees apart.
6. One slot cut in centre of Brass Rotor, 360 degrees around, 0.25" wide by 0.313" deep.
7. 12 slots (produced from the six slots when the 360 degree cut is made). Each slot is lined with mica insulation, 0.01" thick.
8. A total of 228 pieces of U-shaped copper coated steel wires, 0.04" thick. Each slot (7 above) has 19 pieces of these wires fitted into the Mica, thus these wires do not contact the Brass rotor. The leading edge of these wires is flush with the Rotor's outer surface and the trailing edge protrudes 1/8" above the Rotor's outer diameter.
9. Each of the 12 magnets receives eleven complete turns of 0.032" thick copper coated steel wire. These 11 turns or 'wraps' accumulate to 3/8" wide and the same pattern is placed around all 12 magnets. When placed into the bent wires (8 above), they form a snug fit making firm contact.
10. Twelve pieces of mylar insulation, 0.005" thick, are inserted into the cores of the wires (9 above).
11. The twelve permanent magnets, insulated with the mylar, must not contact wires of 9). These magnets measure 3/4" long, 5/8" wide, 3/8" thick and are made of a special composition and strength. Alnico 4, M-60; 12 AL, 28 Ni, 5 Cobalt Fe, Isotropic permanent magnet material cooled in magnetic field, Cast 9100 TS. 450 Brin, 2.2 Peak energy product. When inserted in the rotor the outer faces of these 12 magnets are not to be machined to a radius. The centre of these magnets pass the centre of the coils with 3/32" clearance. The edges, where the wires are wrapped, pass 1/32" away from the coils. This 'changing magnet spacing' aids in not only the release cycle but also contributes to rotational movement. (Sharp magnet edges which are facing the coils are to be sanded to a small smooth radius.)
12. Make sure that the magnets are placed in the Rotor with the polarity shown in the diagram.
13. The 12 magnet wire wraps are divided into two sections; 6 upper and 6 lower. There are no connections between these sections. The magnetic flow direction between the upper 6 wraps and the lower 6 wraps is attained by the 'flow direction'. The wires are wrapped around the magnet starting at the top 'north' half and then after 11 complete turns the wire exits at the lower 'south' half. As this wire then goes to the next magnet it arrives at an attract wire which is its 'north' side. Thus all wires get interconnected from south to north magnet half or north to south magnet half. The actual connections should be crimped copper clips (**not** solder) with insulation tubing to prevent contact to the Rotor body.
14. A 0.03" thick copper tube (stiff material) 2" long by 2½" inside diameter.
15. Six slots are cut at the top of tube #14. These slots are 5/8" wide by 1/32" deep spaced at 60 degrees apart.
16. Six slots are cut at the bottom of tube #14. These slots are 5/8" wide by 5/16" deep and in line with the upper slots #15.
17. There are six copper tube mounting points.
18. An acrylic ring is used to hold Part #14, measuring 3.75" outer diameter and 2.25" inner diameter, 3/8" thick, bolted directly to Part #1. This ring has a 0.03" wide groove cut 0.25" deep to allow the six copper tube mounting points to be inserted (part 17).

19. Plastic insulation paper, 0.002" thick, is to be placed around the inside and outside of Part #14.
20. There are six coils of insulated copper wire, each coil having 72 turns of .014 thick wire. Each coil is wound with two layers, the bottom layer completely fills the 5/8" wide slot with 45 turns and the top layer spans 5/16" wide with 27 turns. To be sure each coil has the exact wire length of 72 turns, a sample length wire is wrapped then unwound to serve as a template for six lengths. A suggested coil winding method is to fill a small spool with one length then by holding the copper tube at the lower extension, then start at the plus wire in Figure 2 and temporarily secure this wire to the outer surface of the tube.
21. Next, place the pre-measured spool of wire inside the tube, wrapping down and around the outside advancing clockwise until the 5/8" slot is filled with 45 turns. Then, return this wire back across the top of the coil for 15/32" and winding in the same direction again advance clockwise placing the second layer spanned for 5/16" with 27 turns. This method should have the second layer perfectly centered above the first layer. After winding this coil, repeat the process, filling the small spool with another length of pre-measured wire. A very important magnetic response happens as all six coils have their second layers spaced in this way.
22. Item 22 above shows the connection pattern for six coils. When the unit is driven at start-up (hand crank) for 42 seconds at 2100 RPM, all six jumper wires must be together which means the plus wire goes to the minus wire connected by the start switch. After 42 seconds the load is added to the circuit and the start switch is opened. To double check your connections between the coils, note that the finish wire of coil #1 goes to the finish wire of coil #2, which is top layer to top layer. This pattern then has start of coil 2 (bottom layer) going to start of coil 3 (also bottom layer). When the copper tube with the coils is placed around the rotor, the distance from any magnet to any coil must be identical. If it measures different, acrylic holding shapes can be bolted to the aluminium base, protruding upward, and thus push the copper tube in the direction needed to maintain the spacing as stated.
23. Wires to load.
24. Wires to start switch.
25. Rotational direction which is clock—wise when viewing from top down.
26. Acrylic dome for protection against elements.
27. Coating of clear acrylic to solidify rotor. **Do not use standard motor varnish**. Pre-heat the rotor and then dip it into heated liquid acrylic. After removal from dip tank, hand rotate until the acrylic hardens, then balance rotor. For balancing procedure, either add brass weights or remove brass as needed by drilling small holes into rotor on its heavy side.
28. Insulation tubing on all connections.
29. Shaft for start purposes and speed testing (if desired).

The reason that this generator is included in this chapter is because the construction is quite complex. Also, the plans have been around for several years without my being aware of anyone constructing or operating one of these units other than J.L. Naudin.

Cold Fusion.

Cold fusion was initially accepted with great excitement. It then appeared to be discredited, primarily as a matter of fear of losing the funding for the decades old "hot" fusion research. At the present time, there are some two hundred labs. which have confirmed the findings and so there is no doubt as to the reality of the system. In essence, it is said that nuclear fusion can take place at room temperature, under certain conditions. However, a commercial working device is now reaching the market with details at Sterling Allan's website: http://peswiki.com/index.php/Directory:Andrea_A._Rossi_Cold_Fusion_Generator.

If you want the background details of cold fusion history, then there are several web sites which follow the progress in this field, including "Cold Fusion Times" at http://world.std.com/~mica/cft.html where considerable detail is available. However, although a cold fusion generator is unlikely to be something which you can knock together in your back yard, the process is edging steadily towards general use. In 2012, the Rossi "Ecat" cold fusion generator is expected to launch. The web site http://ecat.com/ is already taking orders and the home unit producing 10 kilowatts, is expected to sell for US $500 and have an annual running cost of just $20.

The Frolov / Moller Atomic Hydrogen Generator.

One already successful experiment has been shown at the J.L. Naudin website where many successful tests were performed. Alexander Frolov remarks that due to the water pump being externally powered, the results shown by J.L. Naudin are not quite accurate although the error is not significant for power levels below one kilowatt, and so the performance is actually very slightly less than that reported.

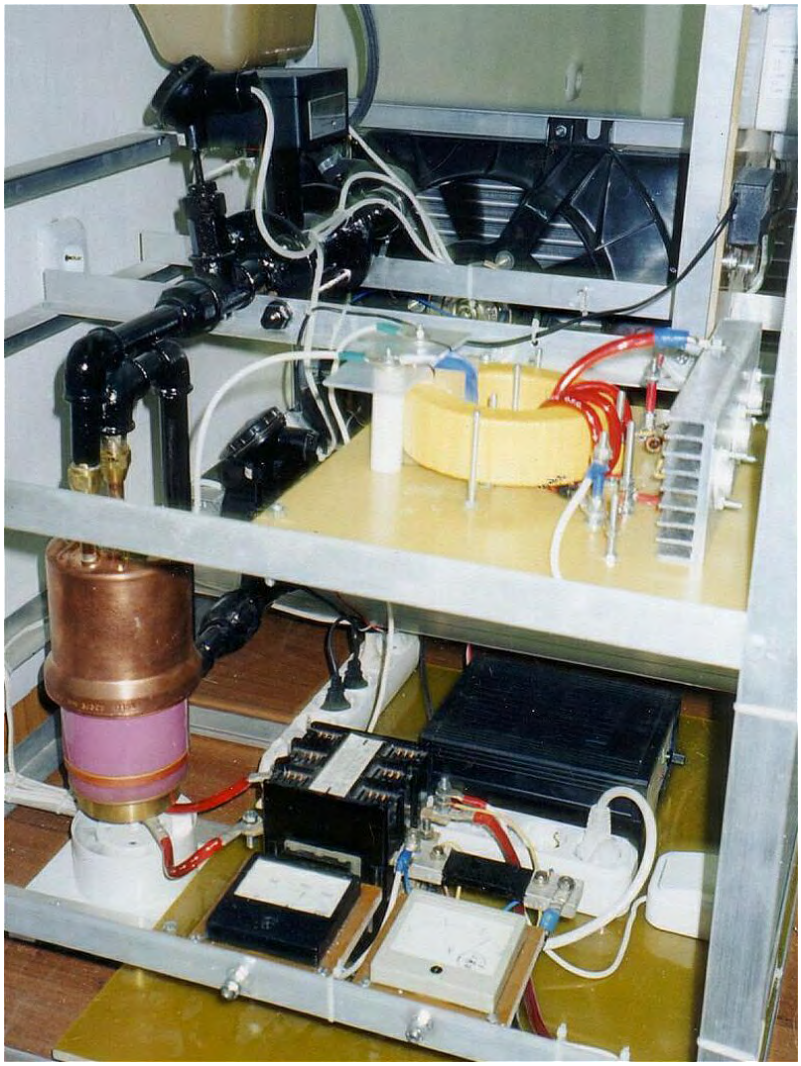
This system started with William Lyne's concept which he published in his book "Occult Ether Systems" in 1997. In 1999, Nikolas Moller bought a copy of Lyne's book and drew the attention of Alexander Frolov to the idea. Alexander then worked on the idea and produced both the present theory and design. A joint project was then started between Alexander's Faraday Company Ltd. of St. Petersburg in Russia (www.faraday.ru) and Nikolas' company Spectrum Ltd.

The prototype shown in the photograph below, was then built by Alexander and passed to Nikolas where it was tested extensively. The exact process involved in the energy gain has only recently been fully disclosed in Alexander's book entitled "New Energy Sources".

The technique used involves repeatedly converting a quantity of hydrogen gas from its diatomic state (H_2 where two hydrogen atoms are bonded together to form a stable molecule), to its monatomic state H-H (where two hydrogen atoms remain as separate atoms, not closely bonded together) and back again to it's original form.

No hydrogen is consumed. No additional gas is required. The gas is just converted from one state to the other, repeatedly. The problem for conventional science is that the output power measured in tests is considerably greater than the input power in carefully measured tests which were run for periods of more than half an hour each. The additional power is flowing in from the Zero-Point Energy field due to an energy extraction mechanism caused by the collision of molecules which have very different atomic weights. This is explained by Dr Frolov in his paper: http://alexfrolov.narod.ru/mac.html where the energy imbalance is described in detail. In it, he observes that the theory shows clearly that an input of 1,400 watts is capable of producing an output of 7,250 watts due to the impact energy imbalance between high-mass and low-mass molecules when they collide.

Here is a photograph of the development assembly built by Alexander Frolov and then used in the testing phase of the system:



Muammer Yildiz's "Ocean Star" Electrical Generator.

This is a purely mechanical device which is self-powered and which can provide electric current to drive other equipment. This device was designed and built in Turkey. It was demonstrated in Dortmund on 17th October 2005 where the demonstration was conducted by J. L. Duarte who ran an independent test and produced a report dated 17th July 2005 on behalf of the Department of Electrical Engineering, Electromechanics and Power Electronics of the Eindhoven Technische Universiteit. Muammer has obtained Patent Application WO2004091083 for his design. The demonstration was of his portable unit which outputs some 12 volts DC:

During the demonstration was used to light a car light bulb very brightly:

Muammer has also produced a larger version capable of powering a house:

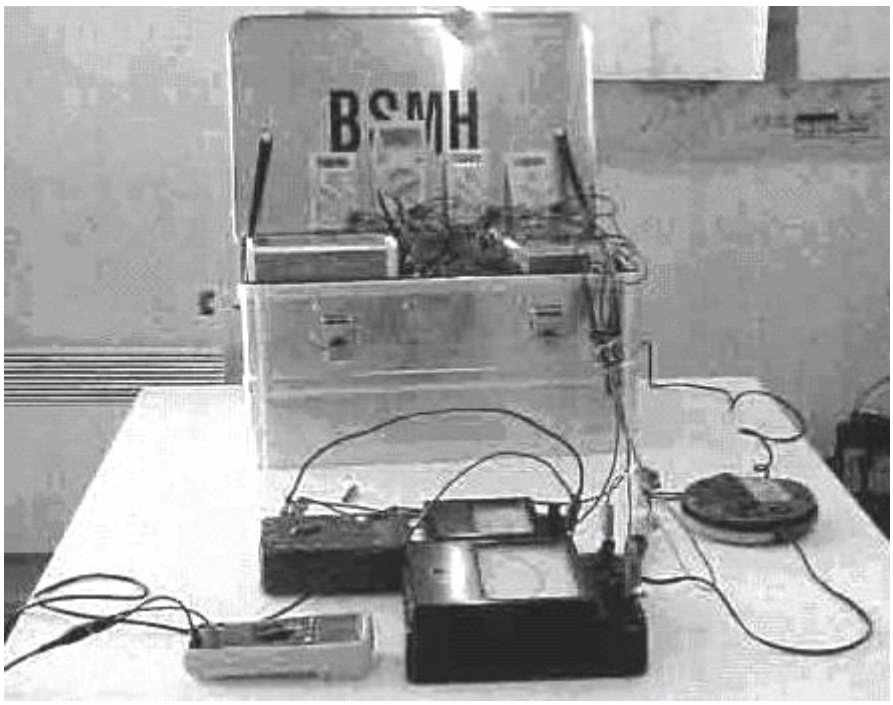
The demonstration unit was started using a 16 AHr battery for a few seconds. Once the unit reaches its running speed, it becomes self-powered and capable of delivering substantial electrical power and the starting battery is then disconnected. In theory, no mechanical system can produce 100% efficiency, let alone more than 100%. This system where a motor spins several generators, is very much like the Raoul Hatem system described in Chapter 2.

The report by Dr. J. L. Duarte on the smaller unit provides the following information:

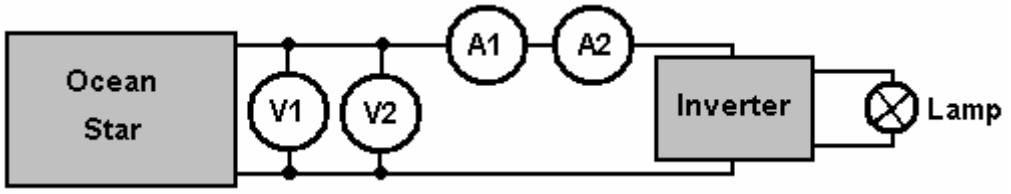
This technical note aims at describing a test which I personally conducted in Izmir, Turkey on 17th July 2005. The purpose of the experiment was to check the energy balance with respect to input and output of an apparatus which was the embodiment of the invention described in the international patent WO 2004/091083 A1 (shown below).

The apparatus was confined inside a metallic box sized 550 x 380 x 270 mm, weighing some 20 Kg, and I was allowed to inspect everything outside this box. However, in order to protect the core ideas of the invention, I was not supposed to check all the details of the internal parts. According to the inventor, the apparatus is predominantly a mechanical system, without any kind of energy storage inside the box (such as batteries, accumulators, flywheels, combustion motors, chemical or radioactive reactions). I believe the intentions of the inventor to be in good faith.

The experimental set-up was quite simple, as shown schematically in **Fig.1**. It consisted of placing the box with unknown contents, from which DC voltages and currents were expected to be generated, on a table in the middle of the room. A cable with two terminal contacts was run from the box and instruments were placed between the box and the load, which was a standard DC/AC inverter driving an incandescent lamp. The output power from the box was measured before the load connection as shown here:



The circuit connection method used is shown here:



After a short start procedure, the metallic box and the load were both fully isolated from the environment, ensuring that there was no physical contact or connection to external power sources such as the public electric mains supply, at any time during the whole duration of the measurements. As the start-up energy input to the apparatus was quite modest, the main issue was then to measure the delivered energy output.

I had prepared the power measurements with care, by using reliable instruments which I personally brought with me from my own University laboratory. In order to measure the DC voltage directly out of the positive and negative terminals, I used two different voltmeters connected in parallel. One voltmeter was an analogue type, constructed with permanent magnets and wires, while the other was a digital voltmeter. To measure the DC current I used two ammeters in series, one analogue and one digital. If electromagnetic waves should interfere with the measurements, then they would disturb one or other instrument, but not all four pieces at the same time and in the same way.

Before starting the test, no audible sound was being produced by the apparatus. The measured voltage and current at the terminals were zero. So, as far as I could observe, the apparatus was completely at rest.

The start-up procedure consisted of connecting a small 12V DC lead-acid battery to two contact points inside the box for a few seconds. I checked the time using my own watch and it was more than 5 seconds but less than 10 seconds. I consider it reasonable to consider the time to have been 8 seconds. After that time, no energy input was connected to the box by means of cables.

Immediately after the start-up procedure, I could hear noise such as would be produced by parts rotating inside the box. The inventor said that some ten minutes should be allowed to elapse before the load was connected. During that time, both of the voltmeters showed the output voltage dropping slowly from 12.9 volts to 12.5 volts. The two voltmeters matched accurately. In the following hours, I observed and recorded by hand, the voltage and current values displayed by the instruments. The displayed values were quite stable, so I initially decided to note them at 15 minute intervals, but later on at 30 minute intervals.

From time to time, using my hands, I attempted to find a temperature gradient inside the box, but I could not detect any variation or increase in the temperature compared to the room temperature. After five hours, I took the decision to stop the measurements. The results are shown in the following table:

Time	V1 (Digital)	V2 (Analogue)	A1 (Digital)	A2 (Analogue)
0:00	12:54	12.5	2.23	2.35
0:15	12.57	12.5	2.29	2.35
0:30	12.57	12.5	2.29	2.35
0:45	12.53	12.5	2.27	2.35
1:00	12.51	12.5	2.27	2.35
1:15	12.48	12.5	2.27	2.35
1:30	12.47	12.5	2.27	2.35
2:00	12.41	12.4	2.26	2.35
2:30	12.35	12.4	2.26	2.35
3:00	12.30	12.3	2.25	2.35
3:30	12.22	12.3	2.25	2.3
4:00	12.15	12.2	2.25	2.3
4:30	12.01	12.1	2.24	2.3
5:00	12.00	12.0	2.23	2.3

****

As far as I am concerned, the above table of results raises serious doubts. The voltage readings are absolutely typical of an inverter powered by a lead-acid battery. I have tested many batteries in exactly the same way and the table looks 100% familiar. If the box contained a genuine self-powered generator, then I would expect the output voltage to remain constant under the constant current drain. In my opinion, it was wholly irresponsible to have stopped the test after just five hours with the output voltage falling steadily. If the output voltage had been rock steady at 12.5 volts for the whole five hours, then that would not have been quite so bad but with it going down 12.3, 12.2, 12.1, 12.0 in the last four 30-minute intervals, and with a lead-acid battery voltage of 11.5 for a fully discharged battery, it was wholly unrealistic to stop the test. A further ten hours of testing should have been undertaken.

However, early in 2010, Muammer demonstrated a permanent magnet motor/generator of his own design. That type of device is notoriously difficult to get functional, let alone producing the 250 watts of power which Muammer demonstrated at Delft University. As part of the demonstration, Muammer took the motor apart completely to show that there was no hidden power source. This shows clearly that Muammer is a man of very considerable ability, and that lends considerable credibility to his earlier device shown here. Unfortunately, the patent application is not in English and some of the terms produced by a rough translation are not at all clear. Consequently, although the design appears quite simple, it can't be rated as a design which could be confidently replicated from the information here and for that reason, the OceanStar information remains here, among the "Unlikely to Result in a Workable Device" section.

Here is the information from the Patent Application WO2004091083 although the quality of reproduction and the clarity of the wording is not particularly good:

A SYSTEM WHICH GENERATES ELECTRICAL POWER VIA AN ACCUMULATOR THAT PROVIDES THE INITIAL MOTION FOR THE SYSTEM

ABSTRACT

This is a portable system that generates electrical power via an accumulator that provides the initial motion for the system. Two batteries are used in this system and the system is kept working via the initial motion provided by these batteries. There is no need for another transformer. This device works using its own

mechanism and there is no need for additional devices. In this way, a continuous electrical power generation is possible. This device can work without connecting it to a network so it is possible to use it at places where electricity does not exist. Moreover, when connected to the entry of a building, the need for a network is avoided. This system generates electrical power independent of a network.

DESCRIPTION

A system which generates electrical power via an accumulator that provides the initial motion for the system This is a portable system that generates electrical power via an accumulator that provides the initial motion for the system. Already existing systems can generate electric power of whose duration depends on the lifetime of the battery. In these systems, the battery has to be reloaded in order to restart the system. 12V electrical power provided by the batteries used in cars is increased to 220 V via transformers.

Two accumulators are used in our invention. The system works on a continuous basis after the initial start up via these accumulators. There is no need for another transformer. Our system, which generates electrical power, does not need any other devices and it keeps on generating energy via its own mechanism. Also, the system works without connecting it to a network.

Thus, it can be used at any place where no electricity exists. Nevertheless, when this system is connected to the entry of the buildings, there is no need for an additional network. The system can produce electrical power independent of a network.

DESCRIPTION OF THE DRAWINGS

Below are the explanations of the figures that provide a better understanding about this invention.

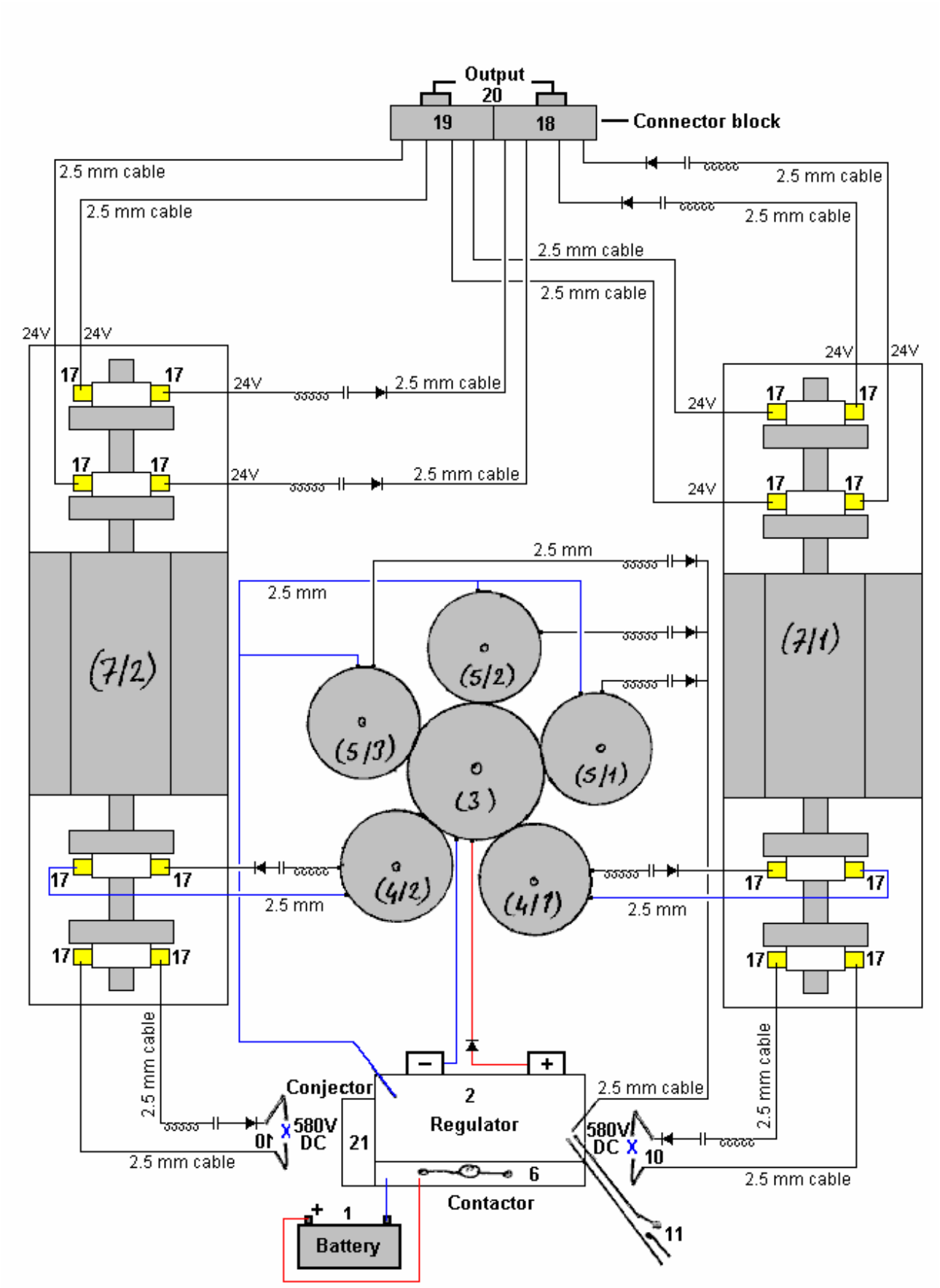


Fig.1 is a schematic view of the system.

Numbers used on the schematic:
1- Accumulator
2- Regulator
3- Big Gear 3/1-Starter dynamo
4- Small gear 4/1-2-Feedback dynamo
5- Small gear 5/1-2-3-Feedback dynamo
6- Contactor
7/l and 7/2- Commitatris
8- 29 DC input
9- 24 DC output
10- 580 DC output
11-Switch
12- Shunt
13- Rectifier
14- Capacitor
15- 2.5 mm cable
16- Collector
17- Charcoal
18- Fixing clamps
19- Fixing clamps
20- Lamp
21- Conjector
22- Starter dynamo
23- Feedback dynamo
24- Alternating current dynamo
25- Magnetic switch
26- Pulley
27- Pulley
28- V pulley
29- 380V current output
30- 220 V current input

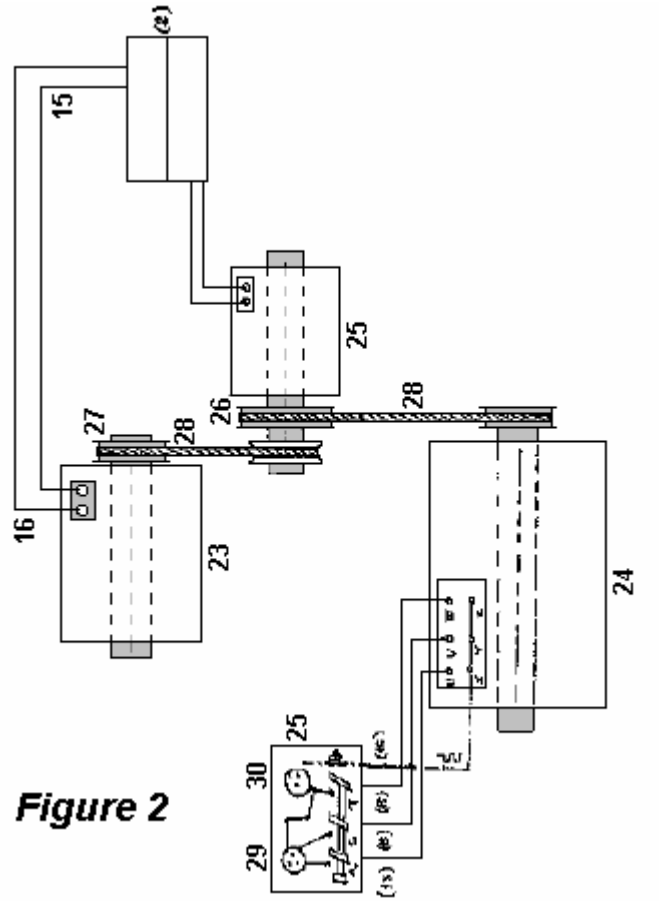


Figure 2

DESCRIPTION

This invention is a system that starts working via the motion of alternator. There exist two accumulators(**1**), and the first motion provided by the accumulator is carried to the regulator. Contactor (**6**) keeps the starter dynamo working by disconnecting the accumulator (**1**) once the regulator (**2**) is put in. The voltage coming from the accumulator (**1**) passes through the regulator and the start dynamo (**3/1**) starts working and thus the feedback alternators via the gears (**4/1-2-5/1-23-3**). Feedback dynamo start sending pure DC current to regulator via shunt (**12**), capacitor (**14**) and diode (**13**). It connects all the currents that reaches to the regulator in 4 seconds and sends to the contactor (**6**). Accumulator (**1**) is put out by this current that reaches to the regulator. This current is transformed to the started dynamo (**3/1**). There becomes a. transformation within the system. In case of electricity shortage, it keeps on working by using the current generated by the commitatris (**7/1**).

Via the starter dynamo(**3/1**), DC is generated in the alternators which are connected to the gears and this current is transformed to the commitatris (**7/1** and **7/2**) and DC voltage is generated at commitatris (**7/1** and **7/2**).

Second System: 3x24 DC voltage is transformed to the second starter dynamo (**22**). Once the start dynamo works (**22**), a feedback dynamo (**23**) having a pulley system and a feedback dynamo (**24**) generating alternating current starts working. The feedback dynamo (**23**) starts feeding back; the feedback dynamo (**24**) which generates alternating current is independently generating 6 KV, 18 Amp, 50Hz current. Moreover, first system produces 24 DC and 580 DC current on its own. The bigger the gears are, the more the generated current is.

This system, which is the subject of our invention, can be used at any place. You can use it at places where there exist no electricity, or at places such as villages, cities, buildings, greenhouses where there is no network. Moreover, network is no longer a must. Instead of a network, you can use our system. There is no need for gasoline when this system is used in vehicles.

Jesse McQueen's System.

There is a US patent which was granted to Jesse McQueen in 2006. This system looks too good to be true and, on the surface, appears impossible, even taking into account that it has been said that ordinary vehicle alternators have a Coefficient Of Performance over one (i.e. output energy is greater than the energy that the user has to put into the device to make it operate). I am not aware of anybody who has tried this system, so I have no evidence that it doesn't work - just a lack of belief in a system of this type being able to operate as described. As against that, it is not far from the proven system of Chas Campbell (chapter 4) and the US Patent office has granted this patent and they have a reputation of being highly opposed to admitting that there is any such thing as a "perpetual motion machine", which this system appears to be. So, I leave it up to you to make up your own mind, and test the system if you wish, which should be easy to do as it involves no real construction, but instead, uses off-the-shelf manufactured products which are readily available and not particularly expensive. Here is the patent:

US Patent 7,095,126 22nd August 2006 Inventor: Jesse McQueen

INTERNAL ENERGY-GENERATING POWER SOURCE

ABSTRACT

An external power source such as a battery is used to initially supply power to start an alternator and generator. Once the system has started it is not necessary for the battery to supply power to the system. The battery can then be disconnected. The alternator and electric motor work in combination to generator electrical power. The alternator supplies this electrical power to the two inverters. One inverter outputs part of it's power to the lamp, and part back to the electric motor/generator. This power is used to power the electric motor. The second inverter supplies power to the specific load devices which are connected to the system.

US Patent References:

5033565	July 1991	Abukawa et al.
5036267	July 1991	Markunas
5785136	July 1998	Falkenmayer et al.

BACKGROUND OF THE INVENTION

Electrical energy occurs naturally, but seldom in forms that can be used. For example, although the energy dissipated as lightning exceeds the world's demand for electricity by a large factor, lightning has not been put to practical use because of its unpredictability and other problems. Generally, practical electric-power-generating systems convert the mechanical energy of moving parts into electrical energy. While systems that operate without a mechanical step do exist, they are at present either excessively inefficient or expensive because of a dependence on elaborate technology. While some electric plants derive mechanical energy from moving water (hydroelectric power), the vast majority derives it from heat engines in which the working substance is steam. Roughly 89% of power in the United States is generated this way. The steam is generated with heat from combustion of fossil fuels or from nuclear fission.

In electricity, a machine is used to change mechanical energy into electrical energy. It operates on the principle of electromagnetic induction. When a conductor passes through a magnetic field, a voltage is induced across the ends of the conductor. The generator is simply a mechanical arrangement for moving the conductor and leading the current produced by the voltage to an external circuit, where it actuates devices which require electricity. In the simplest form of generator, the conductor is an open coil of wire rotating between the poles of a permanent magnet. During a single rotation, one side of the coil passes through the magnetic field first in one direction and then in the other, so that the induced current is alternating current (AC), moving first in one direction, then in the other. Each end of the coil is attached to a separate metal slip ring that rotates with the coil. Brushes that rest on the slip rings are attached to the external circuit. Thus the current flows from the coil to the slip rings, then through the brushes to the external circuit. In order to obtain direct current (DC), i.e., current that flows in only one direction, a commutator is used in place of slip rings.

A commutator is a single slip ring split into left and right halves that are insulated from each other and are attached to opposite ends of the coil. It allows current to leave the generator through the brushes in only one direction. This current pulsates, going from no flow to maximum flow and back again to no flow. A practical DC generator, with many coils and with many segments in the commutator, gives a steadier current. There are also several magnets in a practical generator. In any generator, the whole assembly carrying the coils is called the armature, or rotor, while the stationary parts constitute the stator. Except in the case of the magneto, which uses permanent magnets, AC and DC generators use electromagnets. Field current for the electromagnets is most often DC from an external source. The term dynamo is often used for the DC generator; the generator in automotive applications is usually a dynamo. An AC generator is called an alternator. To ease various construction problems, alternators have a stationary armature and rotating electromagnets. Most alternators produce a polyphase AC, a complex type of current that provides a smoother power flow than does simple AC. By far the greatest amount of electricity for industrial and civilian use comes from large AC generators driven by steam turbines.

SUMMARY OF THE INVENTION

It is an objective of the present invention to provide an energy source that generates more energy than the energy source requires in order to operate.

It is a second objective of the present invention to provide a system that uses the excess energy produced by the energy source to power other various devices.

It is a third objective of the present invention to provide an energy source for supplying power to various devices without the reliance on an external energy source for supplying power to the energy source of the present invention.

The present invention provides an energy source that is capable of producing more energy than it requires to operate. The excess energy is used to power devices. A feedback loop approach is used to channel a portion of the energy produce by the generator back to the generators power input port. This feedback loop approach enables the generator to use its own generated energy to operate. The additional energy generated by the generator is used to power other devices that can be connected to the generator.

In the method of the invention an external power source such as a battery is used to initially supply power to start an alternator and generator. Once the system has started it is not necessary for the battery to supply power to the system. The battery can then be disconnected. The alternator and electric motor work in combination to generate electrical power. The alternator supplies this electrical power to the two inverters. One inverter outputs part of its power to the lamp load device and part back to the electric motor/generator.

This power is used to power the electric motor. The second inverter supplies power to the specific load devices that are connected to the system.

DESCRIPTION OF THE DRAWINGS

Fig.1 is a configuration of an implementation of the internal power generating system of the present invention.

Fig.2 is a configuration of an alternate embodiment of the internal power generating system of the present invention.

DESCRIPTION OF THE INVENTION

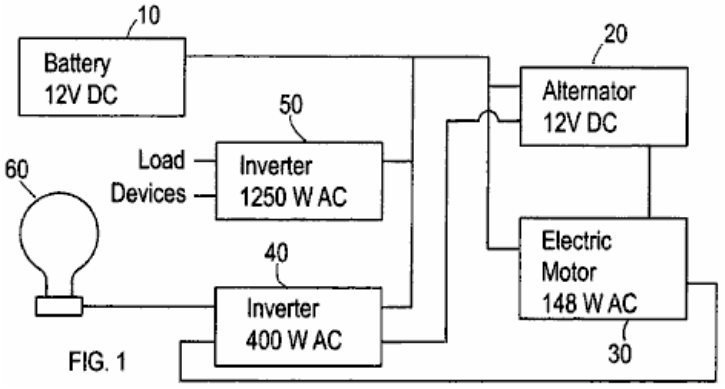
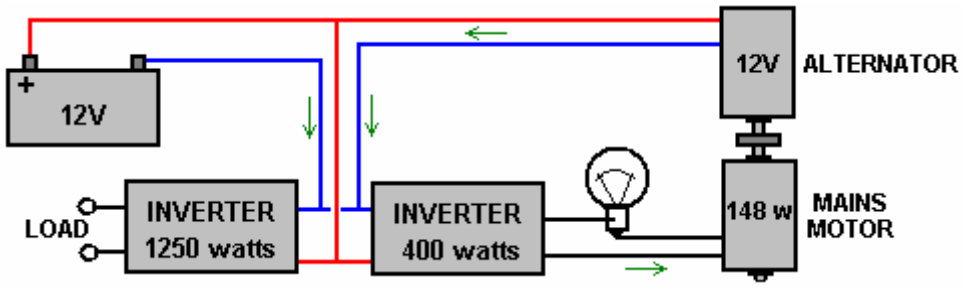


FIG. 1



This invention is an electric power-generating device that produces several times more power than it takes to operate this system. This invention comprises a first power source that is connected to a second power source. Referring to **Fig.1**, the system of the present invention comprises a battery source **10** (12 volt DC) that connects to an electrical alternator **20**. The battery supplies the initial power to the system to initiate/start the operation of the alternator. The present invention can implement other power sources in addition to the illustrated battery to supply the initial power to the system. In the initial model of the present invention incorporated an alternator from a 1997 Isuzu Trooper. The invention incorporates an electric motor **30** (148 watt AC). The electric motor connects to an inverter **40** (400 watt AC). The system also comprises a second inverter **50**. The battery **10** also connects to both inverters **40** and **50**. Each inverter has two outputs. For the first inverter **40**, one output feeds into the electric motor **30** to provide to the motor and alternator combination. The other output feeds into a lamp device **60**. The lamp device is a 60-watt AC lamp. This lamp device alters the current travelling from the inverter **40** such that the current feeding into the electric motor **30** is not purely inductive.

Although, **Fig.1** shows a lamp device, other loads can be used to accomplish this same a task. The inverter **40** has an input from which the inverter receives power from the alternator **20**. The second inverter **50** also has an input that also receives power from the alternator.

In operation, initially, the battery **10** is used to supply power to start the alternator **20** and generator **30**. Once the system has started, it is not necessary for the battery to supply power to the system. The battery can then be disconnected. Once started, the alternator **20** and electric motor **30** work in combination to generate electrical power. The alternator supplies this electrical power to the two inverters **40** and **50**. Inverter **40** outputs part of this power to the lamp **60** and part to the electric motor **30**. This power is used to

power the electric motor. The second inverter **50** supplies power to the specific load devices which are connected to the system. These load devices can be any devices which operate by using electrical power.

The key aspect of the present invention is the loop between the alternator **20**, electric motor **30** and the first inverter **40**. A portion of the power generated by the electric motor is recycled and is used to power the electric motor. In this way the system produces the power internally that is used to power the system. This concept makes this system a self-power generating system.

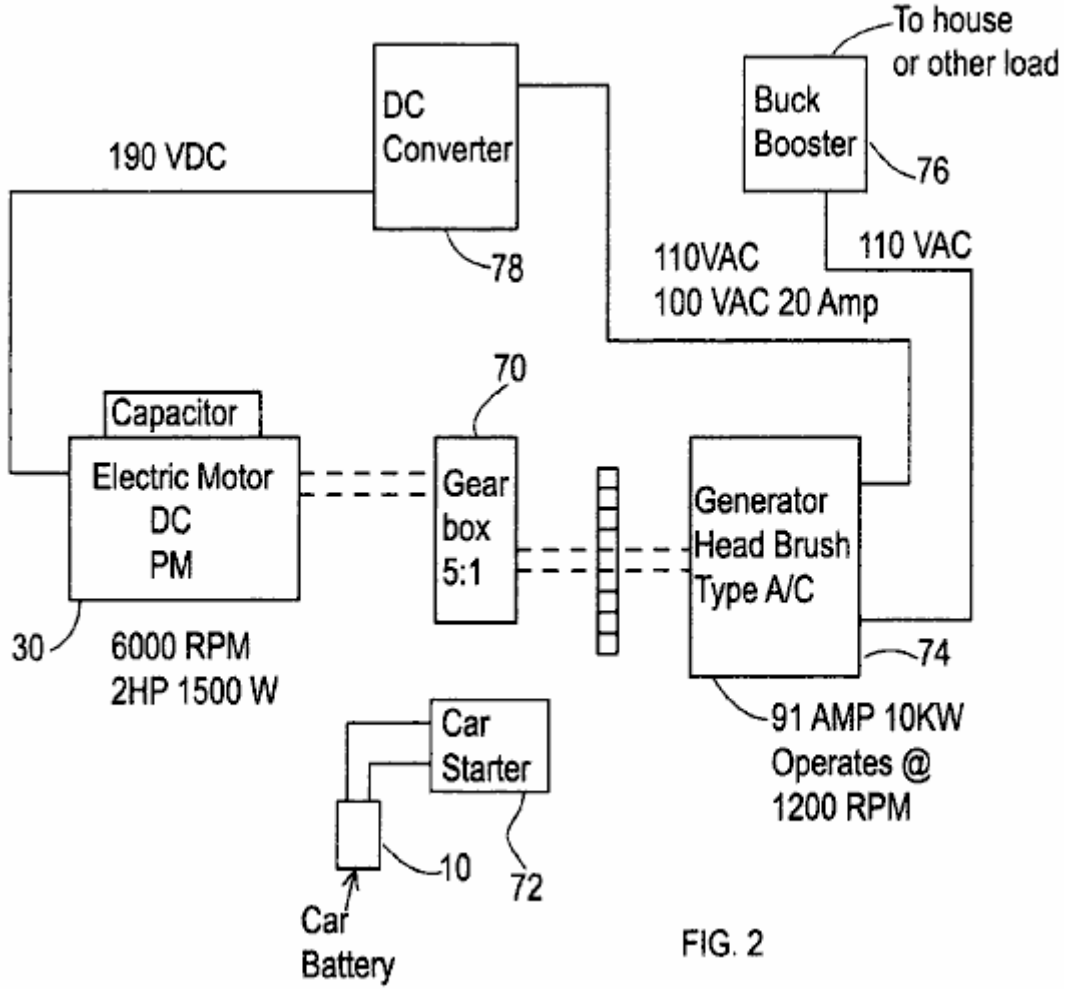


FIG. 2

Fig.2 shows an alternative embodiment of the power generating system of the present invention. This embodiment incorporates a gear box **70**, a car starter **72**, and a head brush generator **74**, and buck booster **76**. Initially, the car starter **72** works with the battery to supply power to the generator. This process is similar to the process of starting a car. The gearshift **70** increases the rpm of the generator. The Buck Booster **76** serves as the output to supply power to the various loads. This configuration also incorporates a DC converter **78**.

The "Nitro" Cell.

This document was originally produced at the request of an Australian man who said that the cell worked well for him but that he was afraid to publish the details himself. This document was prepared, approved by him and published. It proved very popular and an enthusiast group was set up to build and test this "Nitro Cell".

The results of this building and testing have been most unsatisfactory. As far as I am aware, not a single cell proved successful in powering an engine. I therefore, withdrew the document, since even though I believe it to be capable of working, the fact that many people failed to get it working indicates to me that this document should not be in a "practical" guide. I have been assured by two separate independent sources, both of which I rate as being reliable sources, that there are "hundreds" of these cells working in Australia and the USA. I have repeatedly been asked for copies of this document, so I am publishing it again, but requesting

you, the reader, to be aware that should you make one of these devices, that it is unlikely that you will get it operational. Having said that, I understand that it may work very well as a booster.

Simple arithmetic applied to the claimed performance of this device, shows that much of the claimed mileage has to have been covered without using any fuel at all. While this sounds impossible, in actual fact it is not, **but** that sort of operation comes from the Joe Cell which is notoriously difficult to get operational, requiring at least a week of fiddling around to get the metalwork of the vehicle aligned with the energy field used to provide the motive power. Also, each person acts as a "dipole" which produces an energy field around that person. Most people have a polarity which opposes the Joe Cell energy, and they will **never** get a Joe Cell to operate as they can disrupt such a cell from several paces away from it. Chapter 9, which describes the Joe Cell, includes information on how to reverse your own personal polarity, to stop blocking the cell performance.

This definitely sounds unbelievable, but as it happens to be the way that things actually are, there is little point in pretending otherwise. Personally, I never recommend anybody to build a Joe Cell for powering a vehicle, as the likelihood of success is so low. However, having said that, a friend of mine in the USA has his Joe Cell connected to his truck in "shandy" mode where the carburettor is left connected to it's normal fossil fuel supply. The vehicle is perfectly capable of drawing in fossil fuel to run the engine, but it just doesn't. His fuel consumption is literally zero and he is driving around powered solely by the energy channelled into the engine by the Joe Cell. This is most unusual, and I do not recommend you spending time and money on building such a cell. I mention these cells so that you can know all about them, but I would leave it at that.

Here is the original "D18" document, which is followed by important update information:

A Different Fuel

In the early days of heavier than air flight, observations were made and based on those observations, practical operating rules were deduced. After a time, those rules became called the "laws" of aerodynamics. These "laws" were applied to the design, building and use of aircraft and they were, and are, very useful.

One day it was observed that if you apply those laws of aerodynamics to bumble bees, then according to those laws, it was not possible for a bee to fly since there was just not enough lift generated to get the bee off the ground. But simple observation shows that bees do in fact fly and they can rise off the ground when they choose to do so.

Does that mean that the "laws" of aerodynamics are no good? Of course not, as they have been shown to be of great practical use when dealing with aircraft. What it did show was that the existing laws did not cover every instance, so research was done and the laws of aerodynamics were extended to include the equations for lift generated by turbulent flow. These show how a bee can develop enough lift to get off the ground. Do bees care about this? No, not at all, they just go on flying as before. What has changed is that the understanding of scientists and engineers has been extended to better fit the world around us.

Today, people who are trained in science and engineering are fed the idea that internal combustion engines need to consume a fossil fuel in order to operate. That is not strictly true and at the present time, engines using hydrogen gas as a fuel are becoming commonplace. Unfortunately, most of the hydrogen produced for this use, comes from fossil fuels, so these vehicles are still running on a fossil fuel, though only indirectly.

The "laws" of engineering say that it is not possible for an internal combustion engine to run without consuming some sort of fuel. Unfortunately, Josef Papp has demonstrated an internal combustion engine which has had it's intake and exhaust systems blanked off. Filled with a mixture of inert gasses, during one demonstration, that Volvo engine ran for half an hour, producing a measured 300 horsepower, and apparently consuming no fuel at all. Josef received US patent 3,680,432 for his engine and you can see a video of one of his engines running at http://video.google.com/videoplay?docid=-2850891179207690407. Robert Britt designed a similar sealed motor filled with a mixture of inert gasses, and he received US patent 3,977,191 for it.

Does this mean that the current laws of engineering are of no use? Certainly not, they are vital for everyday life today. What it does mean, however, is that the present laws need to be extended to include the effects shown by these engines.

Another thing widely accepted today is that an internal combustion engine can't use water as a fuel. Well.... let's leave that to one side for the moment and look at it from a slightly different angle. Engines can definitely

run using air and hydrogen as the fuel, there is no argument about that as there are many vehicle around which do just that. If you pass a current through water, the water breaks up into hydrogen gas and oxygen gas, this mixture is called “hydroxy” gas and that can most definitely be used, along with air, as the fuel for an internal combustion engine. But... this gas came from water, so is it really correct to say that water cannot be used as the fuel for an internal combustion engine?

Ah, says somebody with relief, that is not the case, because you are using water **and** electricity to get the fuel for the engine. But... the average vehicle powered by an internal combustion engine, has an alternator which produces electricity when the engine is running, so **there** is a source of electricity to do the electrolysis of the water and produce the gas to run the engine.

But the laws of engineering say that you can’t get enough electricity from the alternator to produce enough gas to run the engine. Engineers will point to the work of Faraday who examined the process of electrolysis in great detail and produced the “laws” of electrolysis. These laws show that you can’t get enough electrical power from an engine to make enough gas to run the engine.

Unfortunately, there have been several people who have done just that, so we have reached the point in time when these “laws” need to be extended to cover cases not covered by the work of Faraday. People have got from 300% to 1,200% of the gas output which Faraday considered to be the maximum possible. Several people have run vehicles on hydroxy gas produced by electrolysis of water using electricity generated by the vehicle’s alternator. This shows clearly that it can be done, and as a consequence, the “laws” need to be extended to include the newer techniques.

Leaving that aside for the moment, there have been at least two people who have managed to power an engine with water as the only fuel, and without using electrolysis. In this instance, a fine spray of water droplets inside the cylinder is acted on by the spark, and a secondary electrical supply from an inverter boosts the spark, forming a plasma discharge. The result is a power stroke nearly as powerful as using a fossil fuel. For the moment, let us also ignore that style of operation.

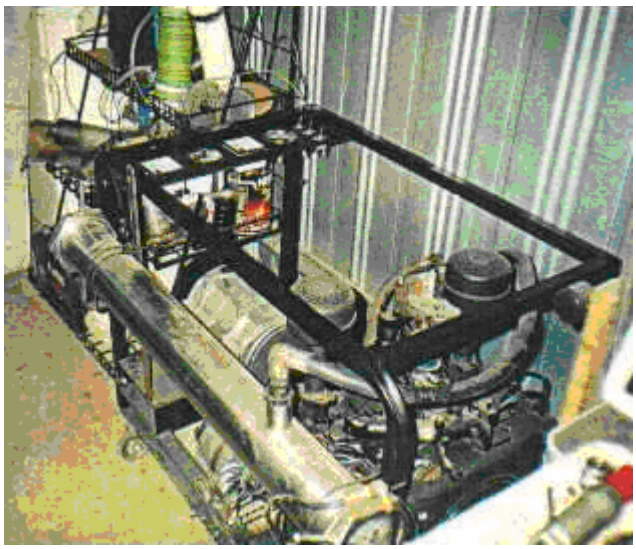
This document describes another system which uses water and air as the primary fuels, but again, does not use electrolysis to generate hydroxy gas for use in the engine. Instead, the objective is to create a continuous supply of Nitrogen Hydroxide (NHO_2) for use as the fuel. This system has worked well for a number of people but there has been considerable intimidation and most of these people are very reluctant to pass the information on. This document is an attempt to present those details clearly enough to allow the system to be replicated by anyone who wishes to do so.

So, how exactly is this fuel generated? The production method is described as the fuel gas being synthesised by a mixture of stream water and rock salt (the mineral "halite") in the presence of air, being acted on by engine “vacuum”, electrolysis and a strong magnetic field. This fuel is said to be more powerful than hydrogen and is a much more viable fuel source as less of it is needed to run an internal combustion engine.

This system may be used with any internal combustion engine, whether used in a vehicle or stationary when powering an electrical generator or other equipment. The additional equipment consists of one, or more, horizontal cylinders mounted near the engine. A single, horizontally mounted, cylinder can generate sufficient gas to power an internal combustion engine up to two litres in capacity. Larger engines will need two cylinders to generate enough gas for them to operate.

It must be stressed that this is **not** a hydroxy gas electrolysis cell. One test vehicle has been run on this system for a distance of 3,000 miles (4,800 kilometers) and the liquid fuel used was only 2 litres of water and 2 gallons of petrol. Two litres of water converted to hydroxy gas will definitely not power a vehicle engine for anything like 3,000 miles, so let me stress again that the fuel being generated in this cell is Nitrogen Hydroxide (NHO_2). It should be noted that if the cell described here is used as a booster for the original fossil fuel, then it will not be necessary to upgrade the engine by fitting stainless steel valves, piston rings, exhaust system, etc.

The person using this system which is shown in the following photograph, has opted for an exceptionally long generation tube attached to his stationary generator:



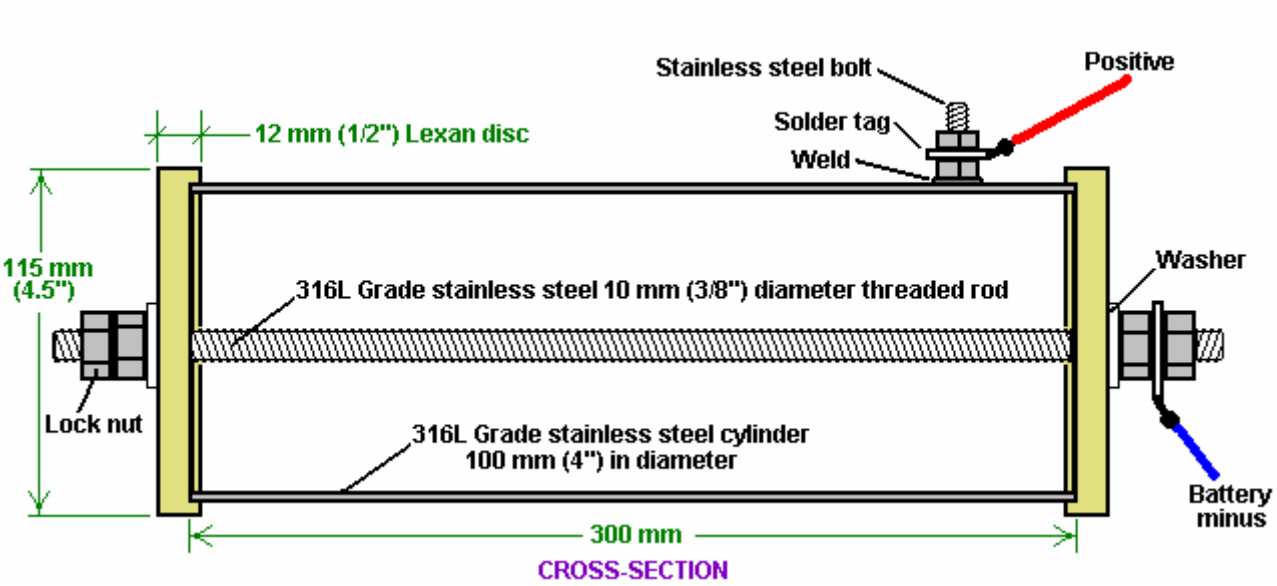
The versions of this cell design shown in the previous photograph and the following photograph, are early models which were in use before it was discovered that there was a considerable enhancement in gas production if a coil is wrapped around the cylinder.

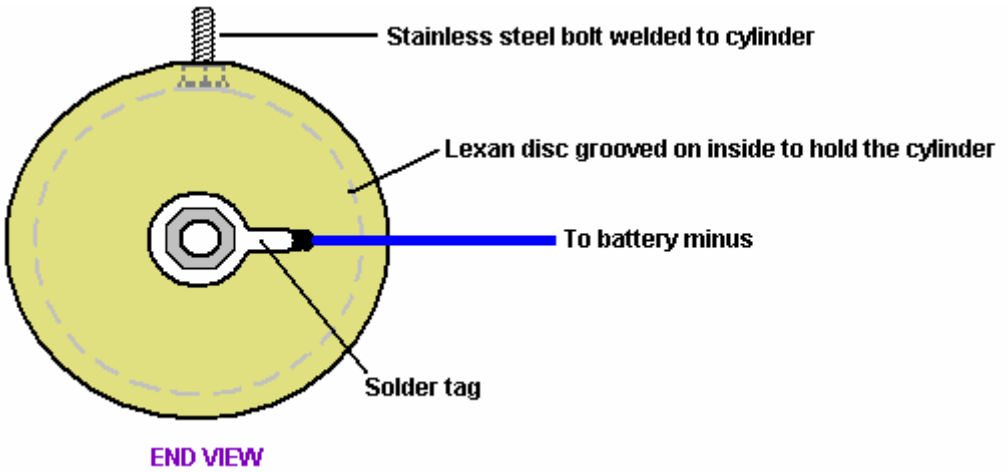
For vehicle operation, it is more normal to have a shorter cylinder, (or pair of cylinders if the engine capacity is large) as can be seen in the following photograph of a 4-litre, 8-cylinder vehicle engine which uses this system. Engines of up to 2 litre capacity can be powered by a single horizontal cell, while two cells are used for larger engines.

The construction details are not difficult to follow and the materials needed are not particularly difficult to find nor expensive to buy. The main body of the device is constructed as shown in the following diagram. A chamber is constructed from a piece of 316L Grade (food quality) stainless steel pipe, 300 mm (12 inches) long and 100 mm (4 inches) in diameter. The length of 300 mm is chosen for convenience of fitting in the engine compartment of a vehicle. If there is plenty of room there, the length can be extended for better gas performance and water capacity. If that is done, keep the 100 mm cylinder diameter and all of the clearance dimensions mentioned below.

The chamber is sealed at each end with 12 mm (half inch) thick discs made from “Lexan” (a very strong polycarbonate resin thermoplastic). These discs have a 3 mm (1/8”) deep groove cut into their inner faces. The groove is there for the cylinder to fit into when the discs are clamped in place and held by stainless steel

nuts tightened on a 10 mm (3/8") stainless steel threaded rod. To combat engine vibration, a lock nut is used to clamp the retaining nuts in place. The threaded rod also provides the contact point for the negative side of the electrical supply and a stainless steel bolt is TIG welded to the outside of the cylinder to form the connection point for the positive side of the electrical supply.





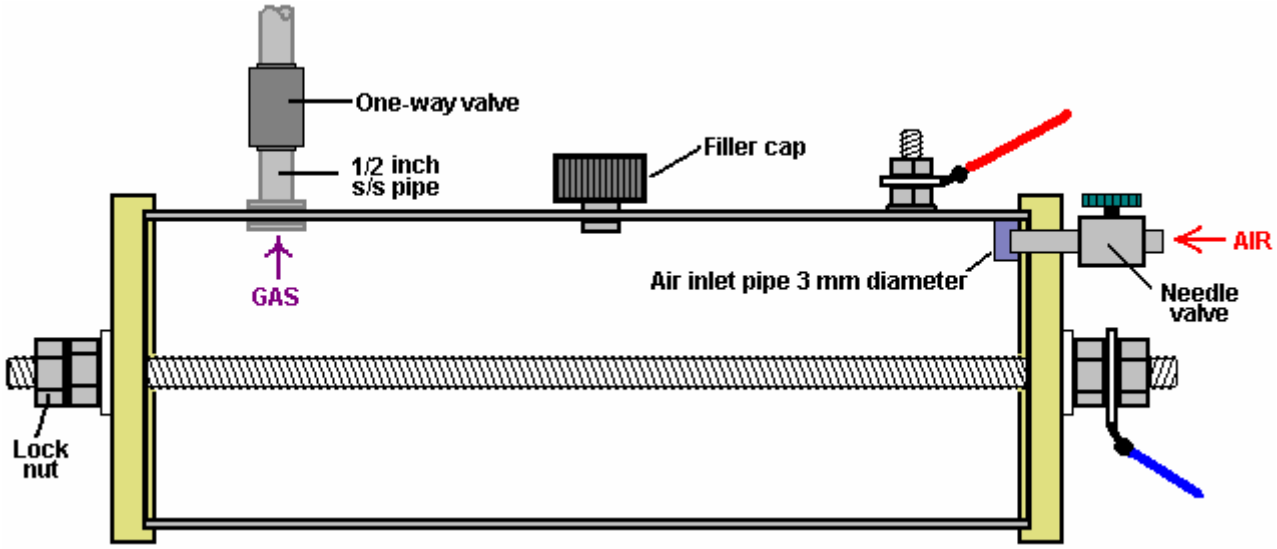
This basic container is modified in a number of ways. Firstly, a small 3 mm (1/8 inch) diameter air intake pipe is provided in one of the Lexan discs. This air intake is provided with a needle-valve which is screwed tightly shut for the early stages of testing and only eased slightly open when the engine is actually running.

Also fitted is an 12 mm (1/2") stainless steel pipe, attached to the stainless steel cylinder to form a gas supply feed to the engine. A one-way valve is placed in this pipe as the design calls for the cylinder to be maintained at a pressure which is less that the outside atmosphere. The lower the pressure inside the cell, the greater the rate of gas production. The one-way valve allows flow into the engine but blocks any flow from the engine into the cylinder. This valve is the same type as is used in the vehicle's vacuum brake booster system.

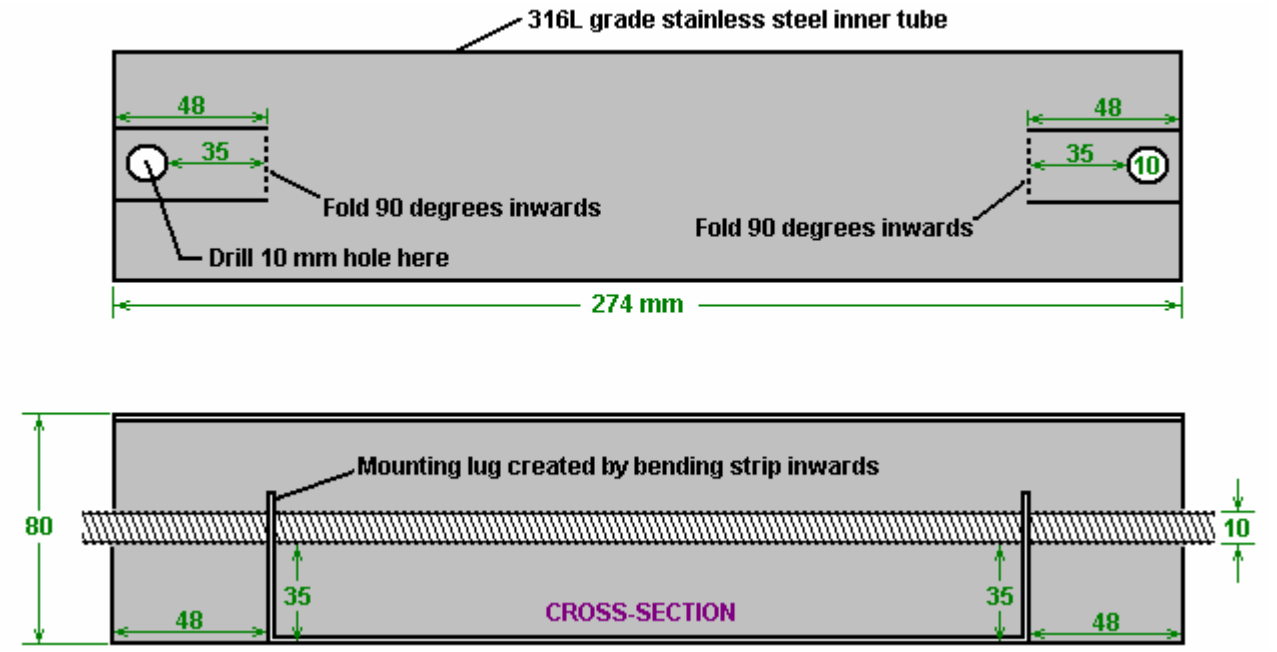
The gas outlet pipe is continued from the one-way valve using plastic tubing for a few inches. This is to prevent an electrical connection between the stainless steel cylinder which is connected to the positive side of the electrical supply, and the engine manifold which is connected to the negative side of the electrical supply. If this pipe were metal all the way, then that would create a direct electrical short-circuit. The pipe running to the engine intake manifold needs to be made of metal in the area near the engine, due to the high engine temperature, so stainless steel pipe should be used for the last part of the gas supply pipe running to the engine. The gas supply pipe fitting is made to the most central of the bungs fitted to the manifold.

For the initial testing period, a filling port with a screw cap is mounted on the top of the cylinder, in order to allow the water inside to be topped up as necessary. Later on, if long journeys are made on a regular basis,

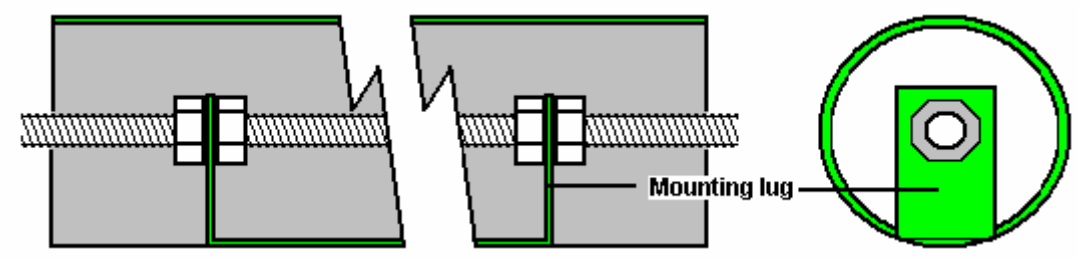
then it is worth fitting a separate water tank, water-level sensor and water injection system using a standard vehicle windscreen washer water pump. The topping up is done with water alone as the rock salt additive does not get used in the process and so does not need to be replaced. With these additional features, the gas generation cell looks like this:



There is one further step, and that is to add an inner cylinder of 316L grade stainless steel. This cylinder is 274 mm (10.75 inches) long and 80 mm (3.15") in diameter. Both cylinders have a wall thickness of 1 mm. The inner cylinder is supported on the central threaded bar and it is clamped in place with retaining nuts. A supporting lug is created by making two cuts at each end of the cylinder, drilling a hole and then bending the lug up inside the cylinder at right angles to its axis. This needs to be done accurately, otherwise the inner cylinder will not lie parallel to the threaded rod, or alternatively, not be centred on the threaded rod. The centre of the 10 mm (3/8") hole is positioned 8 mm (5/16") in from the end of the cylinder. Two 48 mm (1.9") long cuts are made each side of the hole, positioned to be about 5 mm (3/16") clear of the hole - this measurement is not critical. This is done at each end of the cylinder and the holes are positioned exactly opposite one another, along the axis of the cylinder, as shown here:

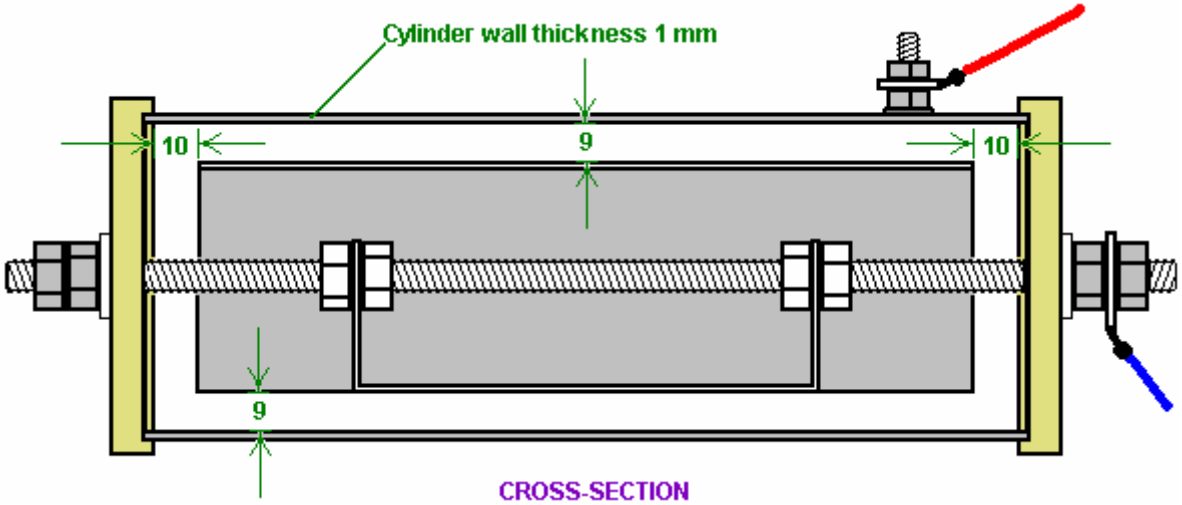


The inner cylinder is secured in position by two bolts as shown here:



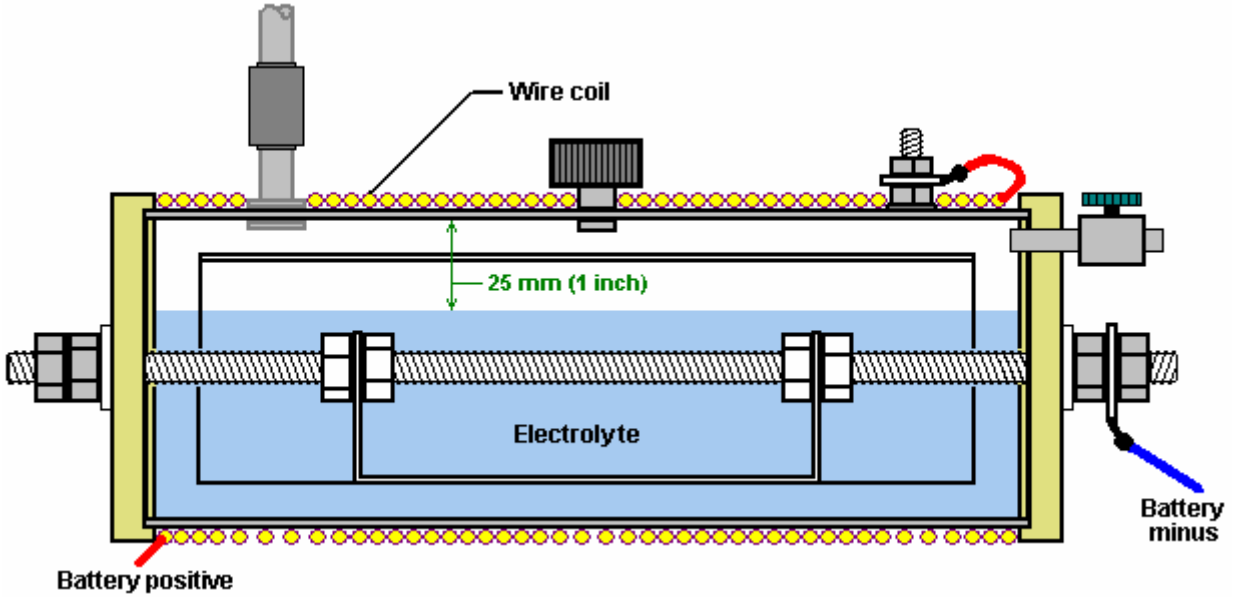
The inner nuts are manoeuvred on inside on of the lugs by hand and then the threaded rod is rotated to move one nut to the inside of the other lug, while the nearer nut is held to prevent it rotating. When the rod is positioned correctly and the inner nuts are pressed up hard against the lugs, then a box spanner is used to lock the outer nuts tightly against the lugs, forming a strong mounting lock.

The inner cylinder is inserted inside the outer cylinder, the Lexan end discs are then added and the outer lock nuts added to produce this arrangement:



This gives a 9 mm clearance between the two cylinders and this gap stretches 360 degrees around the cylinders. The inner cylinder is located 10 mm clear of the Lexan end discs.

The units is completed by winding a coil of 2 mm diameter insulated copper wire tightly around the full length of the outer cylinder and filling the unit with electrolyte to a level of 3 mm (1/8 inch) above the top of the inner cylinder as shown here:



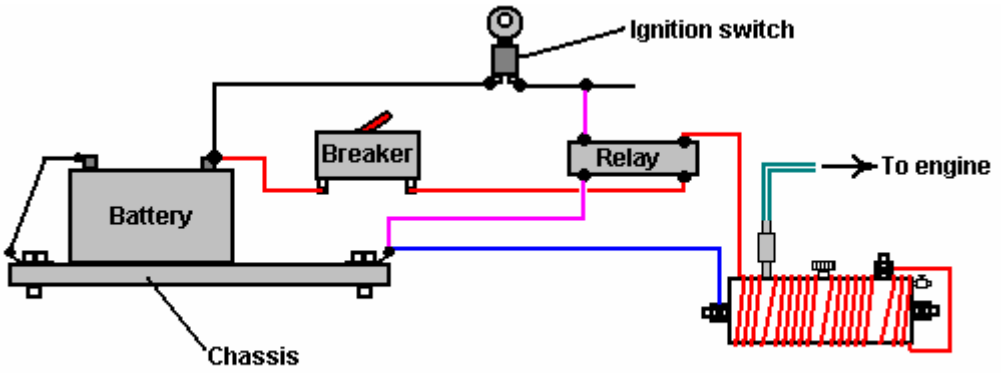
The wire used for the coil is heavy duty copper wire with an inner diameter of 2 mm, i.e. British 14 SWG wire or American 12 AWG wire. The coil is held in position at the ends of the cylinder, with plastic cable ties, as these are non-magnetic. This coil is of major importance in this design as the strong magnetic field produced by it has a very marked effect on the performance of the cell. The magnetic field produced by this coil, increases the gas production by anything from 30% to 50% and increases the production of Nitrogen Hydroxide by a factor of ten times. The electrical connection of the coil is in series with the cell, so the battery positive is not taken directly to the bolt welded to the outer cylinder, but instead it passes through the coil winding before being connected to the outer cylinder.

Installation and Use

The gas outlet pipe is connected directly to a vacuum port directly below the carburettor on the manifold of the engine. This connection is important as the cell relies on the "vacuum" (actually reduced air pressure) produced by the engine intake stroke, as part of it's gas-forming process.

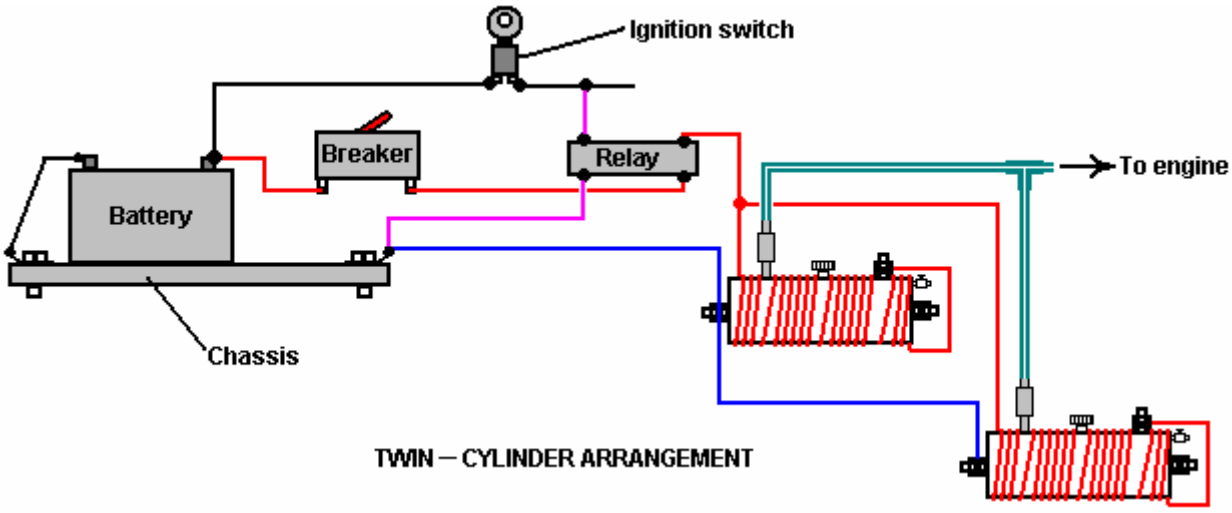
The exact method of mounting the cell in a vehicle depends on the vehicle, so this is something which you will need to think out for yourself. Be sure that you insulate the cell from the metal bodywork of the vehicle and I would suggest that you keep it away from the high-voltage electrical wiring (coil, distributor, spark plug leads, etc.).

The electrical connection arrangement is as shown here:



SINGLE – CYLINDER ARRANGEMENT

Or for larger engines:



TWIN – CYLINDER ARRANGEMENT

The method of electrical connection is important. It is vital that the electrical supply is disconnected when the engine is not running. For that reason, the power to the cell(s) is taken via the vehicle's ignition switch. In order not to load that switch unduly, a standard automotive relay is used to carry the main current, leaving

just the relay current to be handled by the ignition switch. Also, a 30 amp circuit-breaker or fuse is placed in the circuit, immediately after the battery connection. In the unlikely event of some physical problem with the cell occurring, this device will disconnect the power instantly and avoid any possibility of a short-circuit causing a fire, or of excess gas being produced when it is not needed

The water to be used in this cell needs to be selected carefully. Tap water is not acceptable as it will be contaminated with several additives - fluorine, chlorine, etc. put in it when going through the purification process of the supply company and many other chemicals picked up along the way. It is considered very important that the water be taken from a stream, preferably from where it rises, as that is the point of greatest purity. May I also suggest that the water be transported in either glass containers or stainless steel containers as these help to maintain the purity. Avoid plastic containers, because while these appear to be completely inert, they frequently are most definitely not and chemicals from their manufacture can, and do, enter any liquid contained in them.

The cell is filled to a depth of 25 mm (1 inch) below the top of the outer cylinder and then (on the first occasion only) one or two grains of rock salt are added to the cell. This addition needs to be minimal as it controls the current draw from the electrical system and the strength of the magnetic field created by that current. After using the cell for at least a week, if the gas rate is not adequate, then add one more grain of rock salt.

Getting the cell attuned to the vehicle is likely to take at least a week of use. The cell is put in place and the vehicle run using it's normal fuel. The needle valve on the cell's air intake is kept completely closed during this period. The inventor opted to continue running his engine on very small amounts of petrol plus this new gas fuel - the result being 3,000 miles covered on just two gallons of petrol. If you consider this as still being a petrol powered vehicle, then getting 1,500 mpg is quite an achievement - I certainly would settle for that.

When the cell is first connected, you will notice that the engine ticks over faster and tends to rev more than it did before. It will take several days for the system to settle down. Part of this is believed to be the effect of the new magnetic coil in the engine compartment. It may be that the metal parts of the vehicle have to take up a magnetic alignment which matches the magnetic field produced by the cell. Whether that is so or not, it will take a few days before the system settles down into its final state.

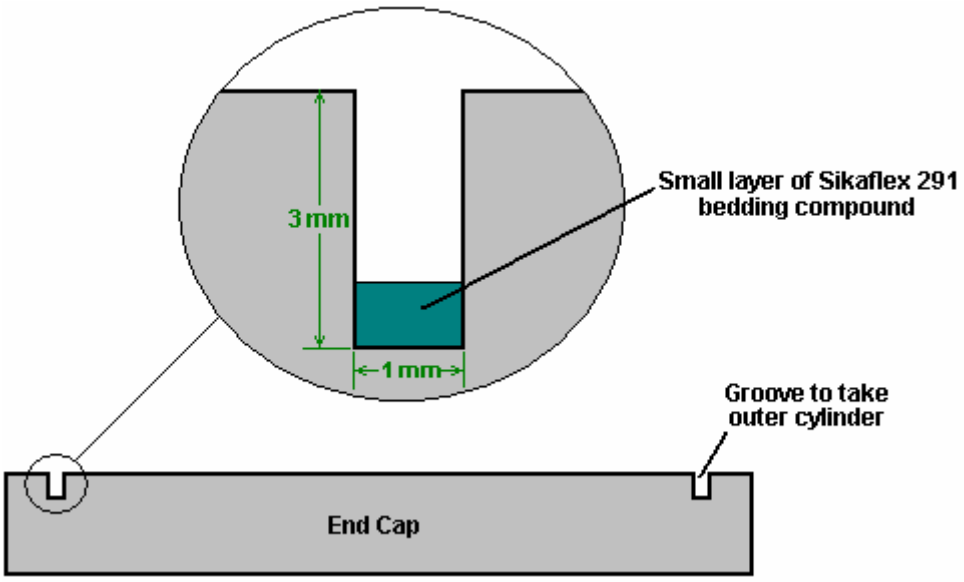
It should be realised that if the vehicle has a fuel-control computer with an oxygen sensor mounted in the exhaust stream, then the oxygen sensor signal will need to be adjusted. The D17.pdf document of this series, shows in detail how to do this, should it be necessary. If the vehicle has a carburettor, then there is an advantage in fitting a one inch bore carburettor of the type found on lawnmowers, as this promotes lower pressure inside the manifold and promotes good cell operation as the lower the pressure (or the greater the "vacuum"), the higher becomes the rate of gas production.

Practical Details

The original end pieces were cut and grooved using a lathe. Most people do not own or have access to a lathe so an alternative method of cutting the discs needs to be used. The essential part of this operation is to cut an accurate groove to take the 100 mm stainless steel outer cylinder. The groove needs to be cut accurately as it needs to form an airtight seal on the end of the cylinder. Consequently, the end of the cylinder and the bottom of the groove, both need to be straight and true if they are to mate securely.

An alternative method is to use an adjustable hole-cutter drill attachment. If this is used with a drill press or a vertical stand adaptor for an electric hand drill, then if care is taken, an accurate groove of the correct dimensions can be cut. As an extra precaution, a thin layer of marine grade white "SikaFlex 291" bedding compound can be used in the bottom of the groove. Two things here. Firstly, only use the genuine Sikaflex 291 compound even though it is far more expensive than other products which claim to be equivalents - they aren't, so pay for the genuine product. Secondly, we do not want the slightest trace of the Sikaflex contacting the electrolyte if we can avoid it, so be very sparing in the amount put into the groove, no matter what you paid for it. Make sure that the bedding compound is placed only in the very bottom of the groove and not on the sides. When the cylinder is forced into the groove, a very small amount of the compound will be driven into any gap between the cylinder and the sides of the groove.

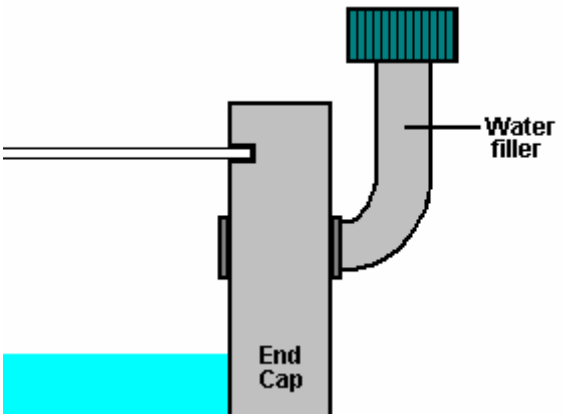
What is needed is a result which looks like this:



The other important part of this joint is the end of the outer cylinder. It is recommended that the cylinder be cut by hand with a hacksaw to avoid generating excessive heat which can affect the structure of the metal. To get the end exactly square, use a piece of printer paper. This has straight edges and square corners, so wrap it flat around the cylinder and manoeuvre it into place so that the overlapping edges match exactly on both sides. If the paper is flat and tight against the cylinder and the edges match exactly, then the edge of the paper will be an exact true and square line around the cylinder. Mark along the edge of the paper with a felt pen and then use that line as a guide to a perfectly square cut. To avoid excessive heat, do not use any power tool like an angle grinder on the cylinder. Just clean the edges of the cut gently with a hand file.

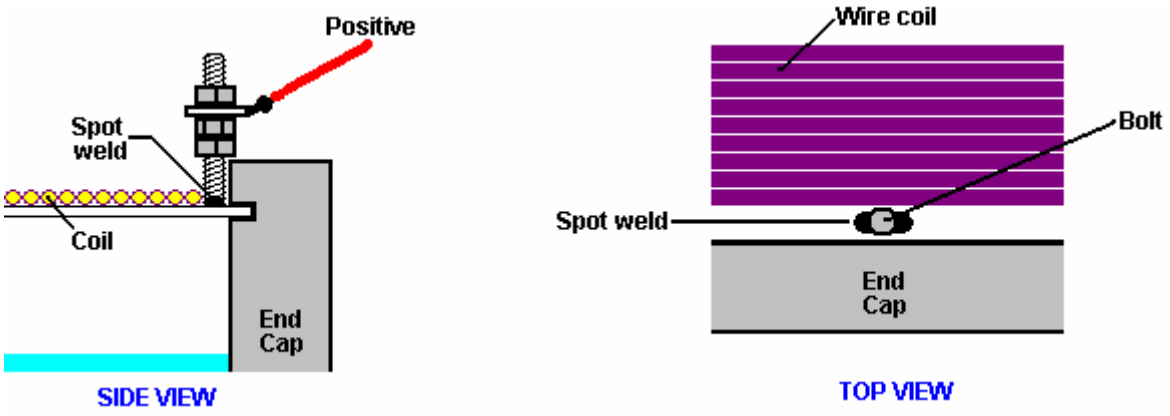
In the diagrams shown earlier, the gas pipe, water-filler cap and the battery positive connection bolt have all been shown on the top of the cylinder. This is only to show them clearly, and there is no need to have them positioned like that. You will notice that they all get in the way of the wire coil, which is not an advantage.

It is necessary for the gas pipe to be positioned at the top as that gives the best clearance above the water surface. The clearance should be maintained at 25 mm (1 inch). The water-filler cap which was shown on top of the cylinder, would be better positioned on one of the end caps as that would keep it out of the way of the coil of wire:



This arrangement has the advantage that it does not require a filler hole to be drilled through the steel cylinder.

It is necessary for the electrical connection to be welded to the cylinder, but it is not necessary to have a head on the bolt as that just gets in the way of the electrical coil. The best strategy is to use a longer bolt of small diameter, remove the head and weld the shaft in place with spot welds which will not get in the way of the coil, as shown below. Spot welds are very quick to make, but even they generate a good deal of heat in the pipe. Some people prefer to silver-solder the bolt shaft to the cylinder as the heating is less.



The bolt is kept just clear of the end cap to avoid fouling it when it is clamped on to the cylinder. A lock nut is used to keep the solder tag assembly clear of the outer edge of the end cap. This allows the wire coil to be wound right up to the bolt. It does not matter which end of the coil is connected to the outer cylinder, but common sense suggests that the end nearest the bolt is connected to the bolt. It is, however, important that once connected, the electrical connections to the coil are maintained ever afterwards, to ensure that the magnetic field stays in the same direction. Remember that the surrounding metal parts of the vehicle will take up a magnetic orientation matching that of the coil's magnetic field, so you do not want to keep changing the direction of the coil's magnetic field.

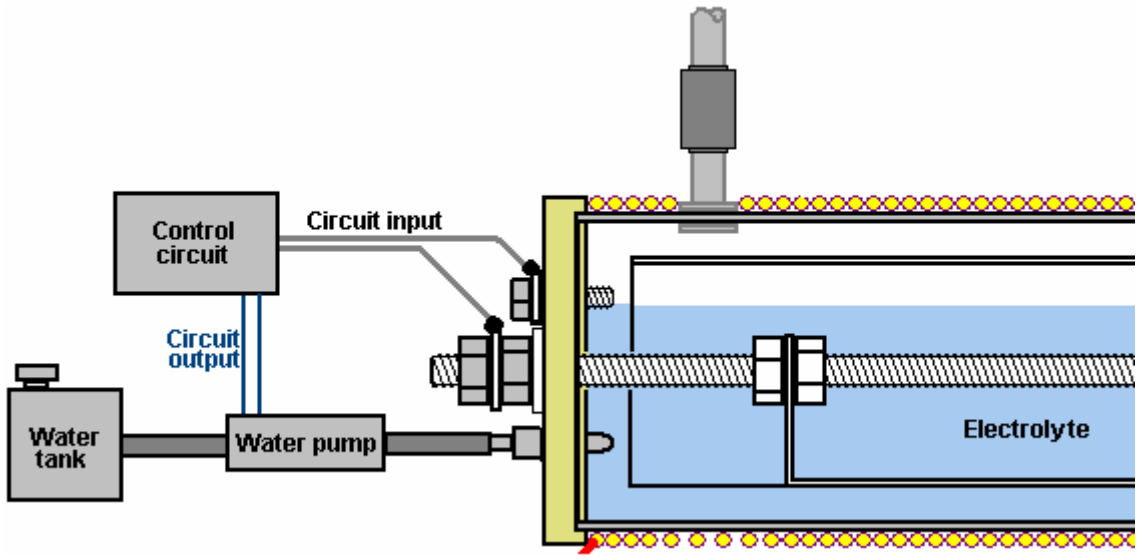
When welding the bolt to the outer cylinder, be sure you use stainless steel wire. The joint needs to be made with a MIG or TIG welder. If you don't have one and can't hire one, then your local metal fabrication shop will make the spot welds for you in less than a minute and probably not charge you for doing them.

The grade of stainless steel in the cylinders is important. Grade 316L is nearly non-magnetic, so if you hold the cylinder with it's sides vertical and place a magnet against the cylinder, the magnet should fall off under its own weight. Try this test no matter what grade the stainless steel is supposed to be, as some steels are not labelled correctly. There is a good chance that you will be able to find suitable tubing at your local scrap yard, but be careful on sizing. The 9 mm gap between the outer 100 mm diameter cylinder and the inner cylinder's 80 mm diameter, is very important indeed. This gap needs to be 9 mm (11/32 inch) so if really necessary to vary the diameters slightly up or down, be sure to pick material which gives the correct gap between the cylinders. Seamless piping is usually preferred to pipes which have seams as the seam welding tends to generate a magnetic effect in the steel. However, if a seamed pipe passes the magnet test with the magnet falling off it, it is definitely good material for the cell.

If you can get it, a good material for the 12 mm (1/2 inch) pipe running to the carburettor manifold, is aluminium. Please remember that the one-way valve on the cell's output pipe needs to be connected to this pipe with a material which insulates the two metal components. The suggested piping is therefore: the cell output is via a stainless steel pipe connector, connected directly to the one-way valve, which then has a plastic pipe connection to the aluminium tube which runs all the way to the manifold. Please remember to insulate the cell from the vehicle chassis and components to avoid a short-circuit.

An alternative to using the rather expensive "Lexan" for the end caps, is to use "UHMWP" - Ultra-High Molecular Weight Polyethylene which is cheap and easy to obtain as plastic food-chopping boards are usually made from it. The advantage of Lexan is that it is transparent and so the level of the electrolyte can be seen without the need for removing the water-filler cap.

It has been suggested that the topping up of water in the cell can be automatic if you wish it to be so. For this, a water-level sensor circuit is used to drive a standard windscreen-washer water pump when the level of the electrolyte falls below the design level. The sensor itself, can be a bolt running through one of the end caps as shown here:



When the electrolyte level drops below the upper bolt, the circuit contact to the control circuit is broken and the circuit responds by powering up the water pump, which injects a little water to bring the electrolyte level back up to where it should be. When the vehicle is moving, the surface of the electrolyte will not be steady as shown in the diagram, so the control circuit needs to have an averaging section which prevents the water pump being switched on until the circuit input has been missing for several seconds.

Circuitry suitable for this is shown in Chapter 12, and there is no reason why you should not design and build your own circuit for this.

In the initial stages of testing and installation, when adding rock salt, be very sparing indeed. Add just one grain at a time because the salt ions are very effective in carrying current through the electrolyte solution. Also, if too much is added, it is difficult to reduce the concentration as more water needs to be added, which involves draining off some of the water already in the cell. It is much easier to take your time and add very, very little salt. Give the salt grain plenty of time to dissolve and spread out throughout the electrolyte before checking the cell performance again.

Let me remind you that during the initial cell testing, the air intake needle valve is closed completely and it is not eased open until the engine is running satisfactorily. In the engine acclimatisation period, the engine should be run on it's normal fuel and the cell just used as a booster. Remember that it will take at least a week for the vehicle to settle down to it's new method of operation. There is no particular hurry, so take your time and don't rush things.

If the vehicle is fitted with computer control of the fuel supply, it may be necessary to apply some control to the unit by adjusting the signal coming from the oxygen sensor placed in the vehicle's exhaust system. The information on how to do this is shown in considerable detail in Chapter 10.

Some questions have been asked about this cell:

1. Does petrol have to be used or can the engine be run on the cell alone?
   Answer: No, you can eventually eliminate petrol altogether but the engine runs so cleanly that old carbon deposits around the piston rings and elsewhere will get cleaned away and the components may rust. These parts can eventually be replaced with stainless steel versions or instead of that, it is probably possible to avoid replacements by the use of the oil additive called "Vacclaisocryptene QX and Molybdenum Disulfide" - see http://www.clickspokane.com/vacclaisocryptene/ for details. This additive reduces wear to such a degree that engine life may be doubled, no matter what fuel is being used.

2. Why is the unit 300 mm long?
   Answer: Just for convenience in fitting it into the engine compartment. It can easily be longer if space allows it. The longer the unit, the greater the gas production and that is why two 300 mm cells are needed for engines over 2 litres in capacity.

3. Does the cell body need to be made from seamless pipe?
   Answer: Seamless 316L-grade stainless steel is preferred.

4. How do you determine the amount of rock salt to add to the water in the cell?
   Answer: The amount varies with the type and size of engine being dealt with. You want the minimum current through the coil so start with one grain and increase it only very gradually with tiny amounts. If the cell is being mounted in the engine compartment of a vehicle, then the make, model and size of the vehicle will affect the amount due to the magnetic effect of metal components near the cell.

5. Does it matter which end of the coil is attached to the outer cylinder?
   Answer: No, it can be either end.

6. Is the pipe diameter shown from the cell to the engine the best size?
   Answer: The 1/2 inch diameter is very good as it increases the "vacuum" inside the cell as the engine runs. When first testing the engine, remember that the needle valve is completely shut off, and when it is opened during tuning, it is only opened to a minimal setting.

7. Are the exhaust emissions damaging to the environment?
   Answer: Some years ago, a Mercedes car dealer ran his own emissions test on a new Mercedes diesel, using his own equipment. He found that the emissions were reduced by 50% and the engine power increased by 12%. The engine ran better, cleaner and quieter. He was fired for doing this.

   Other independent gas-analyser tests showed that there is an increase in water emissions and a drop in carbon emissions as less fossil fuel is used. It was also noted that the volume of gas produced by the cell was affected by where it was mounted in the engine compartment. This is thought to be due to the magnetic effect on the cell.

Update Information:

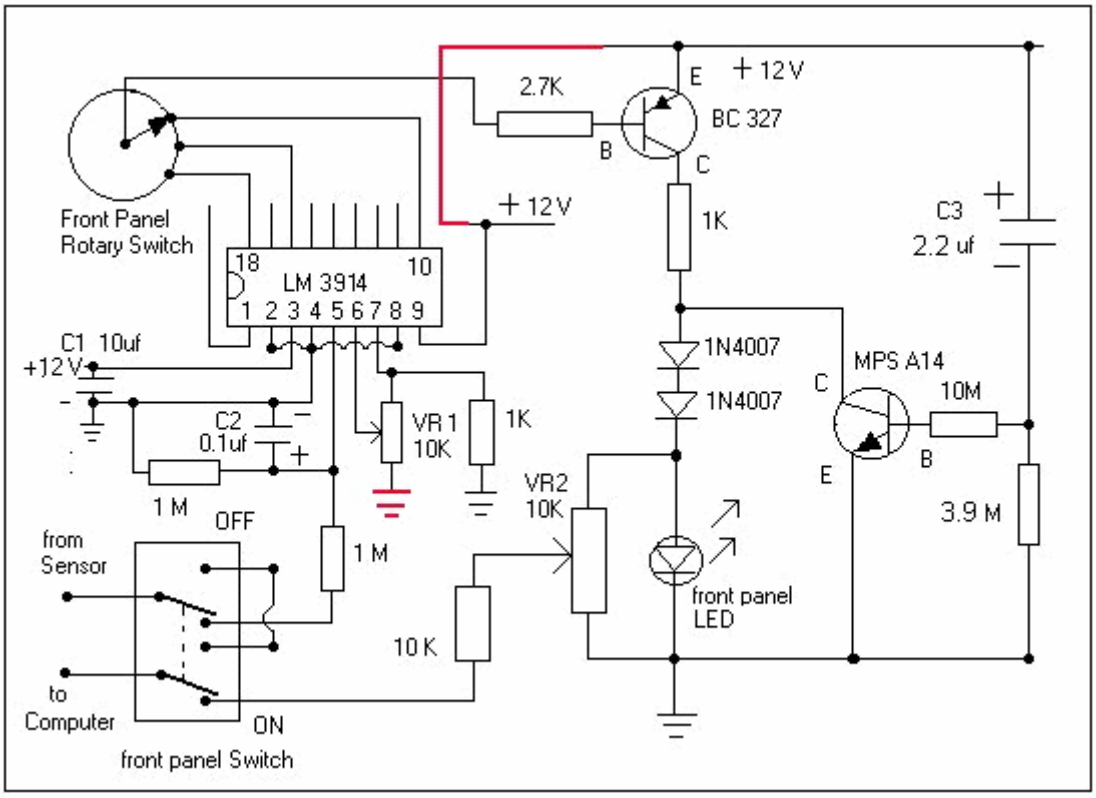
Question 1: Where do we connect the outlet hose from the D18 fuel system to the engine on a late model car with fuel injection system?

Response: There is a throttle body on the engine and it is connected to a rubber hose which goes to the air filter. Typically, the rubber hose attaches to the throttle body and is clamped in place. A hole needs to be punched through the rubber housing approximately two inches (50 mm) from the throttle body. A brass fitting needs to be put into this opening. It will have a flange on one end and the other end with be threaded to accept a nut to hold it in place. This brass fitting will be the attachment point for the incoming fuel line from the D18 system and/or any other booster. For the D18 horizontal system, the size of the fitting should be half-inch (12 mm) so as to be able to maintain the proper vacuum pressure to the D18 fuel system.

Important Note: Since the practice of using alternate fuels by the public is not widely accepted it would be expedient to locate the fuel inlet opening on the under side of the hose out of plain view. This will help the user pass vehicle inspections and keep inquisitive persons from asking too many questions.

Question 2: What do I do I have to do to make the on-board computer function properly with my new booster?

Response: You need to install an electronic mixer control system. Plans for such a system used to be downloaded from http://www.better-mileage.com. This control system will fool the on-board ECU into thinking that all is okay and it will continue to work as normal with no problems. There are two corrections that need to be made to the system to make it work properly. They are outlined in red on this diagram:



Note: In this application the D18 cell is only being used as a booster. Therefore the engine still is using a hydrocarbon fuel. There are numerous systems available such as “megasquirt”, which allow for tweaking the amount of fuel being injected into the engine, and for making numerous other on-board computer adjustments to your Electronic Control Unit, for those of you who want to use nitrogen hydroxide as your only fuel and/or want to reduce the amount of petrol being injected into the engine.

Air Inlet Port: None required!

Ageing of Cell / Cell Break In: Use only the proper water as described below. The cell needs to be drained every day during the ageing process. Filter the water five to seven times through a cotton T-shirt. Collect the water only in glass jars, and do not touch it with your bare hands. Re-use the water and top the cell up with the proper water. Use absolutely no electrolytes (such as salt or potassium hydroxide). You can use natural water which has not seen light and that has not been charged such as, well, cave, or spring water at it’s source. Age the cell until it becomes a slight bronze in colour and does not generate any more gunk inside the cell. The purpose of the break-in period is to purge impurities from the cell.

Cell Current: The peak electrical current with the proper water is approximately 10 amps.

The Positive Electrode: the inner cylinder should be connected to the battery positive. This should be done via an automotive relay to assure proper shutdown of the cell when the engine has been switched off.

The Negative Electrode: This is the outer cylinder, which is connected via a metal strap to the chassis.

Construction: The inner cylinder is separated from the outer cylinder by spacers made out of ebonite or any other material which will not deteriorate within the cell. The objective is to keep the plates at an equal 9 mm spacing throughout the cell. The inner cylinder is connected to the threaded rod via a stainless steel wire strap, which is silver soldered in place at both ends of the cylinder. The threaded rod forms the battery positive connection point on the outside of the cell.

Drain: There should be a drain at the bottom of one of the end plates, so that you can drain the cell without having to remove it from the vehicle. The water will need to be drained and filtered at least once every three weeks. Drain the cell contents into a glass container. Do not touch the water with your bare hands. Filter the water at least five times (seven is better). Use a cotton T-shirt for filtering. Never throw the water away but just filter it. Put the water back in the cell and top the cell off using only pre-charged water.

Electrical Generation: The cell will continue to produce electricity after engine shutdown which will also lead to gas production, so take the precaution of discharging the cell.

Electrolyte: Use absolutely no electrolyte (including salt) at any time. This has been found to decrease the fuel output of the cell and also to have caused unnecessary damage to the plates of the cell.

Engine Timing: Yes, you have to adjust it to your engine. This is a very important aspect of getting high mileage with this system. Each engine is different and therefore each engine has a different adjustment.

Fuel Outlet Piping: Copper piping is recommended as unlike plastic or rubber, it will reduce condensation of water into the fuel lines and thereby reduce the level of water getting into the engine.

Fuel Outlet Ports: There are two of them positioned on opposing ends of the cell. If you are using a 12" cell, then they are positioned 3" in from each end of the cell. If you are using an 8" cell then they are positioned 2" in from each end of the cell. Make one outlet port half-inch (12 mm) in diameter and the other three-quarter of an inch (18 mm) in diameter. Make sure that the piping from the ports extends into the cell by at least an eight of an inch (3 mm). This is to prevent water vapour accumulating at the top of the cell from entering the fuel outlet ports. This measure has been found to reduce water reaching the engine.

Leak Prevention: Use rubber gaskets - these can be the type used for domestic plumbing.

One-way Valves: One-way valves are not used on the fuel outlet pipes.

Outlet Pipe Connections: The half-inch (12 mm) pipe is connected to the engine after the butterfly valve, while the three-quarters of an inch (18 mm) pipe is connected to the engine before the butterfly valve.

Cylinder Preparation: The inside of the 4" (100 mm) outer cylinder and the outside of the inner cylinder, which are the opposing cell plates should be sanded very well with medium grit sand paper to rough up the surface. Two sanding directions at right angles to each other should be used. This will insure better cell productivity later. It is important that there should be no direct contact between the cell plates and your bare hands, so wear rubber gloves when sanding and then assembling the cell.

Voltage: Only 12 volts is required to run the cell, a typical car battery is all you need to power the cell.

Water Selection: Use only natural water that has come out of the ground and not seen light such as well, cave, or spring water at it's source. **Important**: Only add charged water to the cell. Water being used must have a pH of somewhere between 6.4 & 6.5 (slightly acidic). Do not use water with a pH of 7 or higher. The water is charged using a regular Joe Cell with electrodes separated by 3/16" (5 mm) for best results. The details of a Joe Cell can be found in D10.pdf which is a document in this series.

Water Level: Maintain the water level at approximately half full, that is, just covering the threaded rod.

E-mail from a contact:

Hi,

Thanks so much for shedding light into my cell cleaning concerns. I haven't been posting lately since right now the cell is already hooked up in my test car and I've been doing some tests with it day and night.

For once, I can personally tell you that the cell in fact works! However, with my results, it is hard to believe that the Nitrogen Hydroxide produced is enough to make the car get 1500 mpg. When the Nitrogen Hydroxide is allowed to enter the engine, the car starts to rev erratically for 2-3 minutes and then steadies itself thereafter. I noticed an increase of about 800-1400 rpm in my ECU data-logger once the Nitrogen Hydroxide cell is put into the equation. I then adjusted my Engine Management System and removed 15% of the petrol going into the system and drove around the block for a good 15 minutes or so. My exhaust gas temperature rose from 90 Celsius to 97 Celsius which is still fairly acceptable.

I went back to the garage and further adjusted the petrol to less 20% in total and at this point, the car began to vibrate erratically as if it was gasping for air. Upon noticing this, I concluded that not enough NOH must be getting into the ICE or something. The next thing I did was I mounted my old 304L cell alongside with the 316L currently installed. With two cells in the equation, petrol at 20% less did not cause vibrations at all, but mind you that even at 50% less petrol and without Nitrogen Hydroxide cells installed, the car will still run by petrol alone. It was getting dark and so I maxed out my engine management and removed 50% petrol from

the equation with the two NOH cells running side by side. Again, there were vibrations and it was very evident but my brother and I drove the car around the block anyway. Within just five minutes of leaving the garage, the engine temperature rose from 97 Celsius to 111 Celsius and was still rising. I also noticed that the car was underpowered to say the least. We drove up and down a parking complex to test out the non-sloshing design and from my guess it performed pretty well.

To cut a long story short, the cell produced some kind of fuel (NOH or HHO), but it was not enough to power the car when 50% petrol was removed even with 2 cells running. I am currently getting 22 mpg with this test car so I assume that 50% less petrol should give me something like 44 mpg on city driving and probably 60 mpg for long trips. These number are very small as compared to the 1500 mpg that the inventor reported. Maybe the cell needs more time to acclimatise to the test car... but I've been getting same results for 3 days now.

I am currently building two new 316L cells which will incorporate my non-slosh design and which have a vacuum-powered water top up system. I also believe that the gap inside the 3" cell should be sealed off since there is no reaction happening in this part of the cell and it only increases the resistance of the water to electricity. I also included this in my new cell design. I will probably publish it if I find that it produces more gas that the D18 design.

By the way, I've contacted someone in my city who sells 914L stainless steel. However, he told me that 914L requires special handling and special tools and it is much much harder to work on with hand tools alone. He gave me a 1" diameter tube as a sample to see if I can work with it. It's really, really expensive. One cut of a 4" diameter 914L will cost just as much as 2 years worth of gasoline (around 70 FULL TANKS).

And: Bore water is water pump out of the ground. It is similar to well water, the only difference being in the way the water is gathered. Well water is dug from the ground while bore water is SUCKED out of the ground by means of an electric or manual pump.

What I noticed so far is that there is really a lot of steaming going on inside the cell. Converting steam into Hydroxy Gas requires less power than water, so I suspect that this steaming is good. The suggestion to only fill the cylinder half-way actually makes sense as this would allow a greater storage space for the steam and pretty much eliminates water sloshing problems. I have also tried switching the polarities on my 304L cell a couple of times but it did not make any noticeable difference. I will try to make a test run with a half-filled cell and tell you my results.

And: The outlet tubes are positioned at 3" on both ends as this might be the optimum position where less water will accidentally splash into the ports. I was also told that these tubes extend inwards into the cell for about 3 mm so that the water building up in the top surface will not be allowed to slide accidentally into the outlet tubes. It makes sense because with the tubes positioned at 3" away from the edge, it actually gives you about 30 degrees of angle before one of the ports are completely submerged in the water. Also, if the cell is just half-filled, this could actually give us 45 degrees which is relatively a very steep slope to say the least.

The 1/2" and the 3/4" remains a mystery for me. The only reason I can think of is that one of these tubes might be directed before the butterfly valve and the other one placed after the butterfly valve. I would suspect
that the smaller tube (1/2") was placed after the butterfly valve and the larger tube was placed before the butterfly valve. This will make sense because the negative pressure during idle is naturally constant would only require less NOH, while stepping on the accelerator will result in variable pressures which will require larger amounts of NOH. This is only my theory and I am in no position to declare that this is in fact the reason behind the different tube sizes.

I cannot measure the amount of air entering my cell because my end caps are not see-through plexiglass. I only open my Air Inlet Valve halfway through. At this position, I could see a difference in the rpm and at the same time no water could be seen creeping into the tubes. If I open the valve all the way, the rpm will continue to increase but at the same time so does the water in the tubes.

This is my third installation and probably the most successful one. It has been on my test car for five days now but it was not switched on all the time. I found too much water creeping in into the tubes and so I had to
shut it off and run the car normally just to make sure no rusting will occur in my engine. I estimate that I have the cell switched on and running for a total of maybe 12 to 14 hours as of today.

From another user:

Hi,

thanks for the info about the EFI thing! it worked on my wife's Passat. After a few weeks of searching we were able to find smaller injectors for the car as the variable resistor trick only gave us CEL errors. I do notice that the engine's rpm is changing somewhat with every turn of the variable resistor but the relationship is far from linear.

The cell will in fact continue to produce fuel for some time after the power source has been cut. This will tell you that you have the correct water in your cell and you should be happy! What you should do, is to mount a small 12V computer fan beside your cell so that this fan will feed on the power created by cell and reduce the fuel build-up. If you want to be totally safe, then you should put another outlet port on top of the cell and open this every time you park your vehicle. If you want to automate things to avoid constant accessing of the cell, then you can get an electronic valve which will also feed on the excess power produced by the cell. I will not explain further on how this can be achieved, but basically, the valve and fan should only be activated when the engine is off. A few switches here and there will do the trick.

Regarding outlet ports, you are correct to assume that you should have separate lines. One line before the butterfly valve and another line after it, is quite correct and this is what I am doing right now. You should however, have the means to regulate these lines as you will soon realise that too much fuel is actually bad for the engine's health. Also make sure to top up the water regularly as too much empty space inside the cell will make the cell into a bomb!

My concern right now is that if our cells were made half-filled, then it would mean that more than a litre of empty space would be left inside the cell. One litre of Hydroxy or Nitrogen Hydroxide will definitely turn our D18s into a bomb. We should therefore provide a means of venting the NOH build-up when the car is parked. My cell does not produce 13v when shut off as of this moment, which obviously explains the inefficiency I am getting.

Another person:

Hi,

I've seen your set-up pictures at photobucket and I am surprised at the level of professionalism that you are dedicating into this project. I am even more surprised that you claim that your current set-up doesn't work at all! What gives???

Now for my take on your set-up: it seems that the pipes you are using are too small... is it 1/4" ? If so, try to use 1/2" as the minimum. Your fuel output on the end caps should be placed on top of the cell, as I previously stated in my messages. Your water inlet valve should be placed lower. I think the main problem of your set-up is that the introduction of air is placed very near your fuel output. Try to keep these two as far away from each other as possible.

Do not rely on the inventor's set-up as shown on the famous picture on his V8. This picture circulated years ago and to my knowledge, this is not the current set-up that gave him extreme mileage. Last I heard about this guy was that he also used a petrol vaporiser and this was one of the key components in achieving unimaginable mileage on his truck. I for one am not getting even half the mileage that this guy claims. With years of tweaking an old carby truck, I was able to get 225 mpg and this was good enough for me because sometimes I get 300+ on long drives to the country. You should also bear in mind that the longer that petrol stays in your tank, the more evaporation will take place. Upon installing a high mileage device, I realised that most of the petrol is wasted by just sitting in the tank and evaporating.

The air inlet port should be kept as far away from the output ports as possible. It is the water that does the work and not the stainless steel. It should be possible to drain the water completely without removing the cell from the car. The air inlet is a dual purpose port which is placed on the dead bottom of the end caps. Tuning the car to work efficiently with the cell can take a very long time. Make sure that the cell is producing gas aggressively before mounting it on the car.

If you always have the fuel tank full at all times, it reduces fuel loss through evaporation, since on a hot summer day, you are probably losing 12-18% of the fuel through evaporation and what will be left inside the

tank will be less volatile, with bigger molecule sizes which won't combust completely in the engine, which in turn, shortens the life of the catalytic converter and causes more pollution.

System Summary by Contact: I use two 8" cells on my truck with two gas ports on each cell for a total of four ports. Two Ports to Manifold and two Ports to air intake, and there are no one-way valves, instead I use small fuel filters to make sure that water entering the engine is minimised, and at the same time oil is prevented from going into the cell.

I drilled a small hole on the bottom of both fuel filters and sealed them with a small screw plus a rubber ring. From time to time, I remove the screw to drain the water from the filters. The water inside the filters is dirty and should not be recycled for use in the cell. No salt or KOH is used because once the cell is aged and ready, catalysts are no longer required as they will only produce more dirt inside the cells.

Now here comes the most controversial part... NO AIR INLET PORT OPENING!! I don't have an opening for air in my cells. I'm sorry that I've been keeping this from you since day one. I know I told you about properly proportioning your air to the amount of gas that your cell produces. This was the same information that I got from another guy years ago. Although this might be true, you can never be precise on how much gas your cell is producing as the temperatures and pressures you get in the engine vary from time to time... Right now you might be thinking that I may be using a different system all along... this is what I was wanting to avoid that is why I withheld this information from you. But don't worry, there's an explanation for everything...

Simple analysis of the cell design will tell you that it is plainly impossible to remove all the air inside the cell. Air will always enter the weakest point in your cell no matter how air-tight you think your cell is. Take your tires for example: air is continuously escaping your tires no matter how air-tight you may think they are. Needless to say, your cell is not air-tight to begin with so why the hell would you need another air opening? As controversial as I may sound, I found this design to the most efficient.

I use a Joe Cell to charge/clean my water. I don't have a working Joe Cell, I just use it for electrolysis to remove the junk out of the water before putting it in the cells in my car. I have a drain valve on the bottom of one cap and I usually drain and filter my cell water whenever I feel like it. If you have good water and an aged cell, you will produce hydroxy in no time. 304, 316, 317 stainless steel - it does not matter, just as long as you are able to produce gas and that it does not rust quickly. More expensive s/s will tend to outperform cheaper s/s but cheap s/s will still work!

I don't have a magnetic coil and I never heard of this until you pointed it out to me. It did not take me weeks to age the cell, the hard part is really the water. You can use plain old tap and maybe get some gas... hooray! You have just made a hydrogen booster! Or you can follow my lead, and use good water and make fossil fuel nearly obsolete. The water level inside the cell may not matter, but I find that the cell will produce more gas when there is less water is inside it. However, for safety reasons, I almost always make sure that the cell is 3/4 filled with water and 1/4 empty space. Another important thing to bear in mind is the steaming inside the cell. If you use plastic or rubber tubes, the steam might condense back into water before getting into the engine. Use copper tubes to make sure that steam will not condense. To my knowledge, the salt is just used to remove the protective layer on the s/s which actually prevents the bubbles from dislodging quickly. You can also age your cells in many other ways and this won't be a problem.

That main idea is, get your cells to produce hydroxy without using catalysts. The nitrogen part will come as an accident and I cannot explain how this happens. I am still a little sceptical about the nitrogen actually bonding with the hydroxide. Sometimes I think that it is only the hydroxy and steam which are doing all the work... You are entitled to your own opinion.

The wife gets around twice her previous mileage on the Passat. The injector change can only do so much. I only installed one 10" cell to keep all the stock parts intact. She is happy with it and so my EFI project stops here.

Please let me stress again that many people have built this device and tried to get it to work without any success whatsoever, and that is why it is in this chapter.

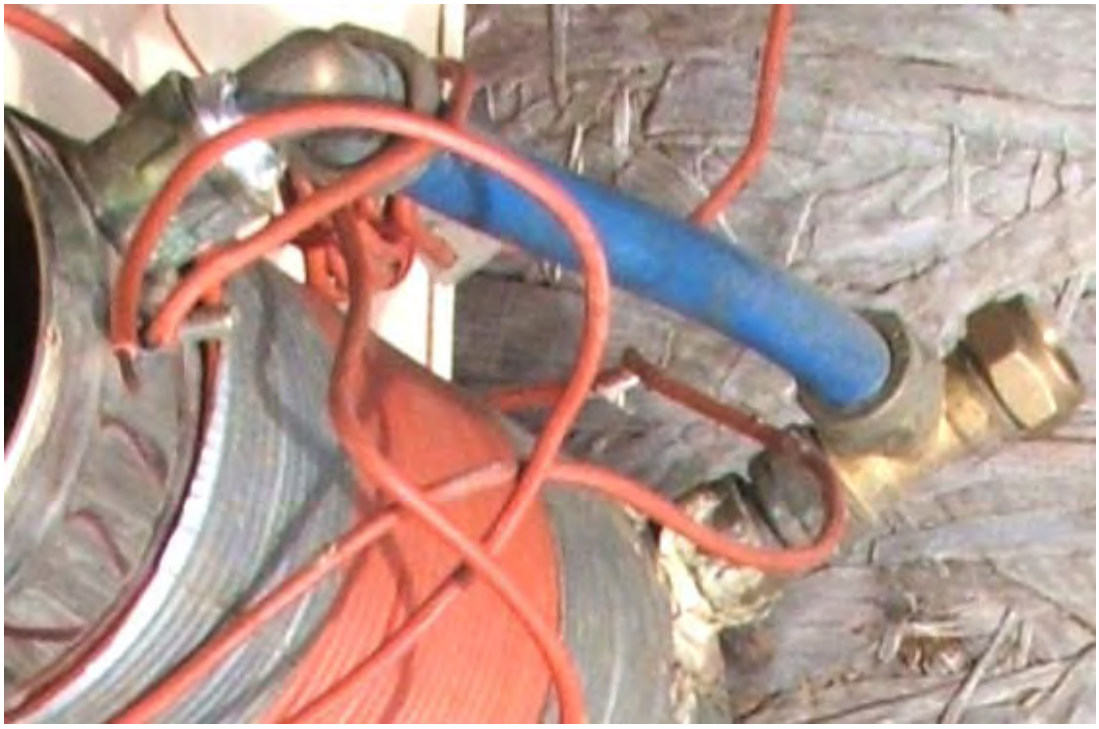
However, having emphasised those concerns and provisos, I was informed early in 2010, that Jim Bundock in the UK has had considerable success using this design as a booster. He estimates that using this device has saved him about £500 over a two year period. Here is some detail from him:

These two cells have been in use in two different vehicles for more than a year now. The vehicles are Ford Transit vans with 2.5 litre diesel engines. One is sixteen years old and the other is seventeen years old and the Nitro Cells have given a 10 mpg improvement from the day that they were installed in spite of the vans being heavily loaded when being used. The original mpg was 25, so the resulting 35 mpg represents a 40% improvement over a long period.

Jim says: these cells were built from the plans shown above, the only variation being that instead of bending the tabs for the inner tube, tabs were silver-soldered at the end of the inner tubes as shown here:

These cells are run without salt and they work equally well without the wire wrap suggested in the design. The vertical cell is 18 inches tall and 4 inches in diameter and it has a rubber diaphragm on top which acts both as a seal and as an anti-slosh device. The inner cylinder gives a 9 mm gap between the inner and outer cylinders. The cell ends and spacers are made from plastic food-chopping board. The top spacers are about 30 mm long as it was found that vibration tended to dislodge shorter ones. This vertical “slow” cell is very easy to build.

On one of the “fast” horizontal units, there is a length of blue pipe:



This prevents water from entering the engine when a certain boy-racer tries to negotiate corners at the speed of light. The horizontal cell was plumbed in using a plastic 12mm pipe connected to the inlet manifold. Personally, I don’t think these cells like speed and they seem to respond best when driving around 50 mph – this is for a single cell on a diesel engine. The cell is wired in as standard, only on a diesel, the relay is connected to the fuel cut-off valve, making it very easy to wire, and when the engine is switched off, then the cell is turned off at the same time. Using a miniature version of the upright cell, the bore-hole water which I use, is pre-charged on the bench for as long as I feel it needs to be, just to remove some of the material suspended in it. Topping up with water is done every 750 miles or so and takes about an egg cup full of water. All in all, the Nitro Cell is a good honest working device, it is a simple drive-and-go piece of kit that has saved me around £500 in two years and will carry on doing that for ever.

The “HydroStar” and “HydroGen” Systems.

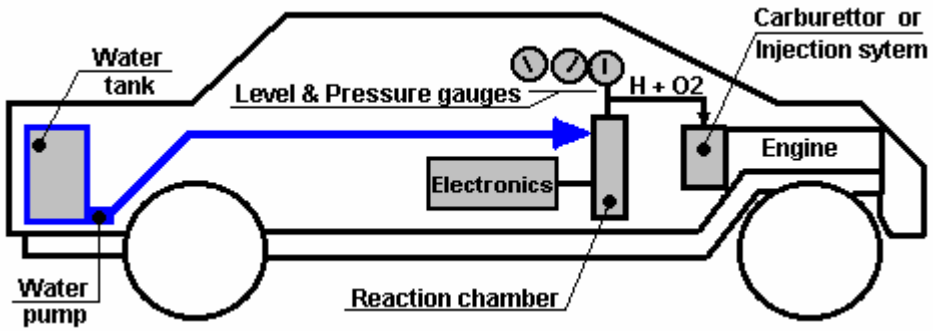
There are various sets of plans for car conversions and many of them are worthless and intended to waste the time and money of people who are interested in moving away from fossil fuel products. It is not possible for anyone to say with assurance that these plans do not work since even if you construct in exact accordance with the plans and your replication fails to come anywhere close to working, all that can be truthfully said is that your own replication was useless. We need to avoid this sort of comment, since for example, the Joe Cell does indeed work and can power a vehicle in a completely fuel-less mode, **but**, most people fail to get it operational. Consequently, it is completely wrong to write off the Joe Cell, but warnings on the difficulty of getting it working should always be given.

In the case of the HydroStar and HydroGen plans, I have never heard of anyone who has ever got either of them working. Also, experienced people are quite convinced that the design is seriously flawed and never worked in the first place. Still, it is up to you to make up your own mind on this, and so these plans are mentioned in this chapter.

The plans shown here can be downloaded free from http://www.free-energy-info.tuks.nl/P62.pdf and they are intended for free use by anyone who wants to use them. Please remember that should you decide to undertake any work of this nature, nobody other than yourself is in any way responsible for any loss or damage which might result. The full manual for an essentially updated version of the design is included under the name “HydroGen” and can be downloaded free from http://www.free-energy-info.tuks.nl /P61.pdf.

It is recommended that should experimental work be undertaken on a car, then the car chosen should be of little value and that all existing parts be kept so that the vehicle can be restored to its present fossil-oil

burning status should you choose to do so. It is also suggested that you use a car which is not important to your present transport needs. It is claimed that the modified car will travel 50 to 300 miles per gallon of water depending on how well it is tuned. The system is set up like this:

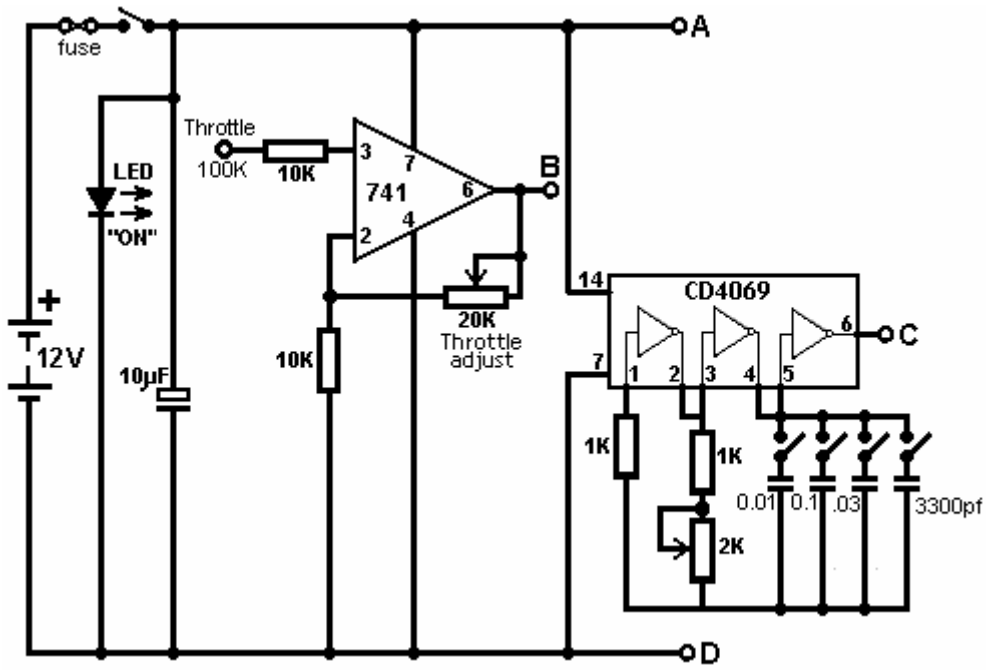


Here, the car has an extra tank installed to contain a reserve of water. This is used to maintain the water level in the reaction chamber which contains the electrode plates. The electrodes are driven by the electronics which applies a pulsed waveform to them in the 0.5 to 5.0 Amp range. The electronics box is powered directly from the existing car electrics. The Hydrogen/Oxygen mix which is the output from the reaction chamber is fed directly into the existing carburettor or fuel-injection system.

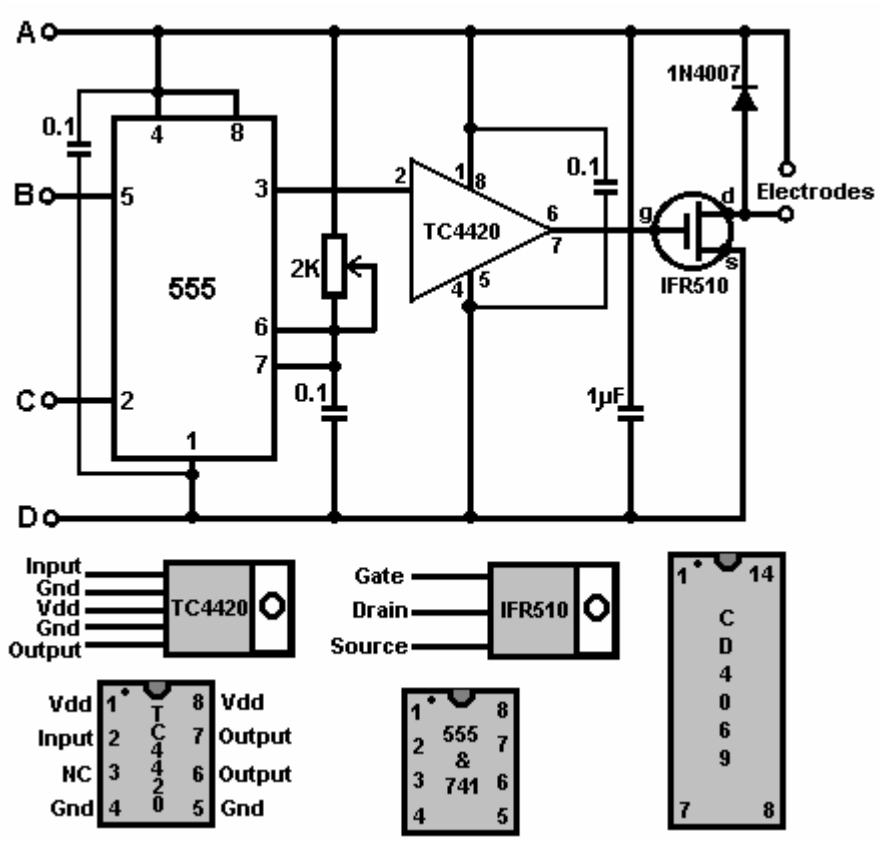
The start-up procedure is to power up the electronics and wait for the gas pressure to reach the 30 - 60 psi range. Then the car ignition is operated as normal to start the engine. The accelerator pedal is wired into the electronics to give more power to the electrode plates the further the pedal is pressed. This increases the gas production rate as the throttle is operated.

Electronic Control Circuit

The diagrams show a simple circuit to control and drive this mini-system. You are going to make a 'square-pulse' signal that you can watch on an oscilloscope. The premise given by the literature is: the faster you want do go down the road, the 'fatter' you make the pulses going into the reaction chamber. Duty cycle will vary with the throttle from a 10% Mark/Space ratio (10% on and 90% off) with the pedal up, to a 90% Mark/Space ratio with the pedal fully down.



There are many ways to generate pulses. This circuit uses an “NE555” integrated circuit. The output switching transistor must be rated at 5 Amps, 12V for pulsed operation.



The output of the 741 integrated circuit is adjusted via its 2K variable resistor, to give an output voltage (at point 'B' in the circuit diagram) of 1 Volt when the car throttle is fully up, and 4 Volts when the throttle is fully down.

The CD4069 is just an IC containing six inverters. It can handle a supply voltage of up to 18V and is wired here as an oscillator. Its four capacitors are likely to be used in just four combinations: C1, C+C2, C+C2+C3, and C1+C2+C3+C4 as these are the most widely spaced tuning ranges. There are, of course, eleven other capacitor combinations which can be switched with this arrangement of four switches.

Important Note

Gary of G. L. Chemelec commenting on "The HydroStar" circuit which sounds to be based on the same style of circuitry, states that the circuit and design are riddled with serious errors, some of which are:

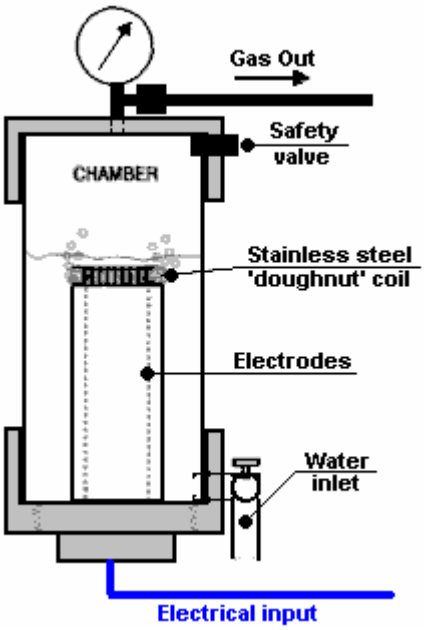
1. The use of the 741 WILL NOT WORK! Pin 5 is a Voltage Control pin that already has its own voltage of 2/3 of the Supply voltage so it requires a pull down resistor, not an IC to control it.
2. The 2K Pulse width adjust will blow the 555 timer if adjusted all the way down. It needs an additional resistor to limit current to those pins on the IC.
3. The output of the 555, Pin 3 is fed to the CD4059 as well as a TC4420CPA (Mosfet Driver). This driver is a waste of money as it is not needed.
4. The Output of the TC4420CPA is then fed to the IRF510 Mosfet which is now obsolete, however you can use an RFP50N06 (50V, 60A).
5. There is no schematic of the CD4059. They should have shown pin 1 as in, pin 23 as out, pins 3, 10, 13, 14, and 24 connected to 12 volts and pins 2, 4, 5, 6, 7, 8, 9, 11, 12, 15, 16, 17, 18, 19, 20, 21, and 22 connected to ground.
6. The "Strength Adjust" Only Needs the variable resistor connected to Pin 5 and the Ground. The Connection of this control to the Supply Voltage Make Absolutely No Difference in the Output Waveforms, as the IC only needs a 2/3's voltage on this Pin and this is supplied internally, Within the IC.
7. The "Frequency Adjust", Connects to Pins 6 & 7 of This 555. Supply to the battery Will Destroy the 555. so another resistor is needed to prevent this from happening.

This is just a small list of what is wrong. There is MUCH MORE and even after the thing is built it does NOT WORK! If you want to experiment then please do, but I would suggest you just make your own Pulse Width Modulator.

There are also many problems with the design of the reaction chamber and simply put, even if you did get it to work you would need more of these units than you could ever fit in your car to even think about running the engine. Simply put, the unit will NOT create enough gas to run much of anything. Don't get me wrong, I do think that the idea is GREAT and that it can be done.

Reaction chamber:

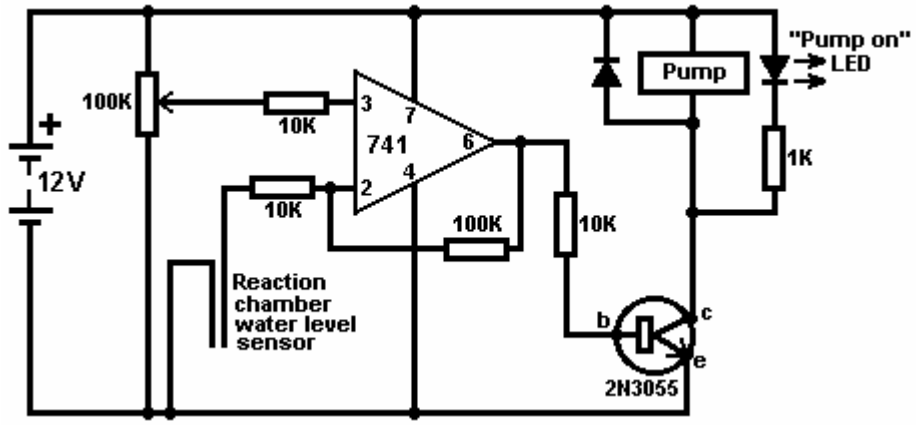
The suggested reaction chamber arrangement is:



It is suggested that you use a section of 4" PVC waste pipe with a threaded screw-cap fitting on one end and a standard end-cap at the other. Make sure to drill-and-epoxy or tap threads through the PVC components for all fittings. Set and control the water level in the chamber so that the pipe electrodes are well covered and there is still ample headroom left to build up the hydrogen/oxygen gas pressure. Use stainless steel wires inside the chamber or otherwise use a protective coating; use insulated wires outside. Ensure that the epoxy seals are perfect or alternatively, lay down a bead of water-proof silicone sufficient to hold the pressure.

The screw fitting may require soft silicone sealant, or a gasket. Its purpose is to maintain the pressure in the cylinder and yet allow periodic inspection of the electrodes. Make sure that there are no leaks and you will have no problems. Make sure you get a symmetric 1.5 mm gap between the 2 stainless steel pipes. The referenced literature suggests that the closer to 1 mm you get, the better. Check that the chamber water-level sensor is working correctly before you epoxy its cap in place. Make your solder connections at the wire/electrode junctions nice, smooth, and solid; then apply a waterproof coating, e.g. the epoxy you use for joining the pipes to the screw cap. This epoxy must be waterproof and be capable of holding metal to plastic under pressure.

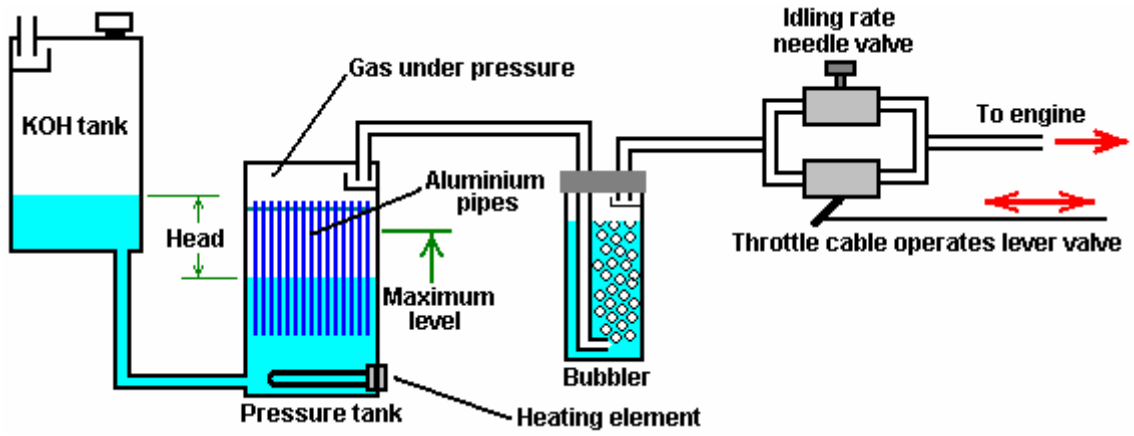
The suggested circuit for the reaction chamber water-level pump control is:



Hydrogen from Aluminium.

Since 2003 Rothman Technologies of Canada have been running a 12 HP petrol motor on hydrogen produced by a chemical process. This is a cheap process in which metal is consumed and so, although of great interest, this is not a 'free-energy' engine. A recent patent application by William Brinkley proposes a system where aluminium pipes are consumed by a 25% solution of Potassium Hydroxide heated to 180 degrees Fahrenheit. William remarks on the non-polluting nature of the system, but this is not really so in that a very large amount of energy has to be put into producing the aluminium metal in the smelting and refining process, and the pollution is just moved from the end user to the industrial plant, and much more importantly, the aluminium oxide produced is said to be highly toxic and cause a wide range of serious illnesses including Alzheimer's, although I have seen this disputed. Francis Cornish of the UK has a system where electrolysis of water is combined with a chemical process consuming aluminium wire. The system works well, but I have reservations about using consumables which tie you to industrial manufacturing, also concerns about the reliability of mechanical feed systems when they are being used by non-technical people (most car drivers). There is also the issue of removing and recycling the chemical residue generated by the process.

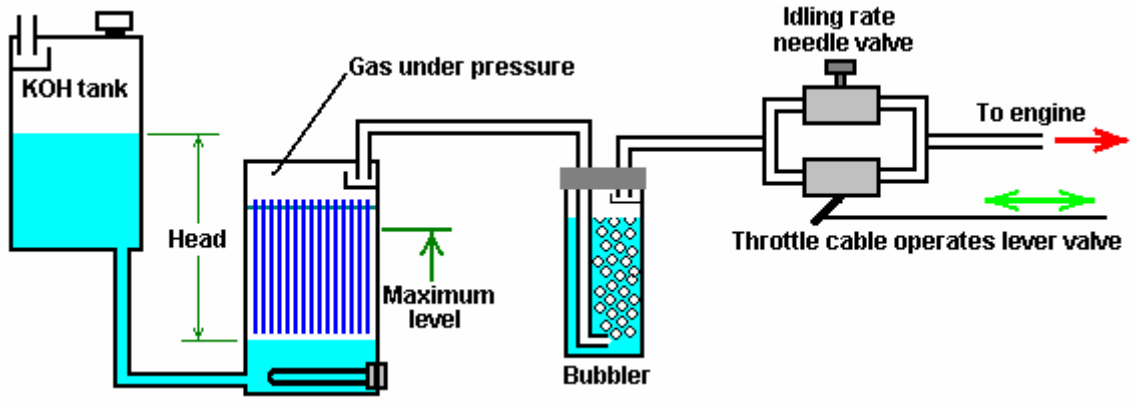
I personally am not keen on chemical processes and I do NOT recommend that you construct anything based on the following description. However, it might be possible to adapt the Brinkley system so that it operates with no moving parts:



Here, there is a header tank containing a 25% mixture of Potassium Hydroxide (KOH) in water. This tank is positioned higher than the pressure tank where the hydrogen gas is generated and the venting pipe is protected by a baffle. The venting pipe should provide an outlet to the air outside the vehicle or building which contains the system.

Initially, the KOH solution in the pressure tank is heated by the heating element, but when the process gets started, it generates heat to maintain the chemical reaction. The gas generation then builds up pressure in the strongly-built pressure tank. The raised pressure pushes some of the KOH solution back into the header tank, against gravity. This reduces the area of aluminium exposed to the KOH solution and reduces the rate of gas production. This effectively creates an automated gas production rate control which has no moving parts.

If the rate of gas taken by the engine increases, that lowers the pressure in the pressure tank, allowing more KOH solution to run into the pressure tank, increasing the rate of gas production. When the engine is stopped completely, then the KOH solution gets pushed into the header tank until all gas production stops, as shown here:



This looks as if the pressure tank is under considerable pressure, but that is not so, as the header tank is open to atmospheric pressure. I have concerns about controlling purely chemical processes rapidly enough for practical use. The above system would be more suited to a fixed engine, such as an electrical generator, where the gas requirement does not fluctuate greatly. The KOH tank shown above should be large enough to contain all of the KOH solution in case the gas production just does not stop when it should. The vent from the header tank should be capable of venting excess hydrogen with no possibility of it ponding on a ceiling and forming an explosive mixture with air. As far as I am aware, the above system has never been constructed and it is just shown here for discussion purposes.

Only 5 pounds per square inch of pressure is needed for electrolyser systems to feed a car engine satisfactorily, so a relatively low pressure is quite satisfactory, provided that the piping is of reasonable internal diameter. It should be remembered that the car engine will be applying a slight vacuum through the bubbler. As with all of these systems, it is vital that at least one bubbler is used between the gas production and the engine, to guard against flashback from the engine ignition if faulty ignition should occur. All bubblers should have a tightly fitting pop-off cap which can ease the effect of an explosion, and they should contain only a small amount of gas. The method of connection to the engine and the necessary timing adjustments are shown and explained in Chapter 10.

Francois Cornish Hydrogen System.

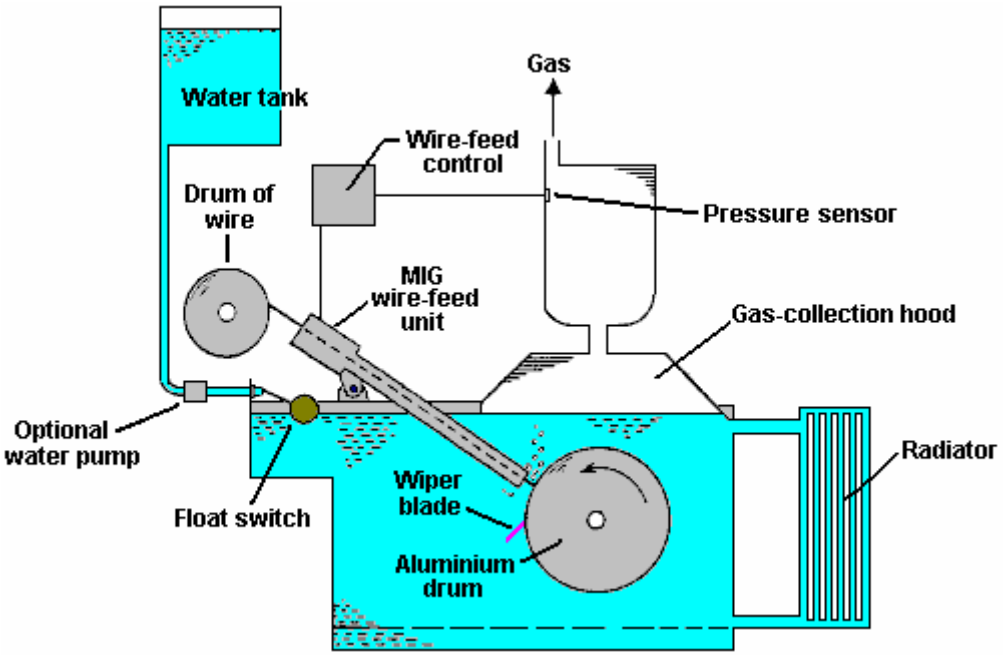
The method of using aluminium for a fuel in an on-demand hydrogen system for vehicle propulsion has been presented in detail by several people. One of the best known is the 1987 US Patent 4,702,894 by Francois Cornish, where he uses a feed mechanism for aluminium wire to maintain an underwater electrical arc which raises the water temperature high enough to make the aluminium react with the water. The rotating drum is made of aluminium but as it has a much larger thermal capacity than the aluminium wire being fed towards it, the drum temperature is much lower than that of the wire. As a result of this, the wire reaches the temperature required to make the aluminium react with the water. The chemical reaction releases hydrogen and converts the aluminium wire to aluminium oxide powder, which settles on the bottom of the tank, passing through a grid just above the bottom of the tank.

The bubbles of hydrogen gas released by the reaction tend to stick to the rotating aluminium drum, so a wiper blade is provided to sweep the bubbles off the drum. The bubbles then rise to the surface of the water and are directed into the gas collection chamber by a funnel located above the arc. If the engine demand drops and the pressure in the gas collection tank rises, a sensor located in the tank causes the wire-feed control electronics to stop the wire feed which cuts off the gas production.

At first glance, a system like this appears to have limited appeal. It uses aluminium wire which requires manufacturing by a process which uses substantial amounts of energy and while a vehicle using hydrogen produced by this method will generate very little pollution, the pollution occurs at the point of manufacture. Also, the device uses a mechanical wire feed and any device of that nature will need regular maintenance and may not be 100% reliable. In addition, the aluminium oxide powder will have to be cleaned out of the generating tank on a routine basis.

But, having said all that, the system has some very significant advantages. It does not use any fossil fuel (directly). It can be readily installed in a vehicle and the consumption of aluminium wire is surprisingly low. Figures quoted indicate that typical consumption is of the order of 20 litres of water, plus one kilogram of aluminium used to cover 600 kilometers distance (1 pound per 170 miles). This is probably a good deal cheaper than using fossil fuel to drive the vehicle. However, the aluminium oxide produced by this system is

a serious pollutant as it is **highly toxic**, producing a wide range of serious illnesses, including Alzheimer's. The system is set up like this:

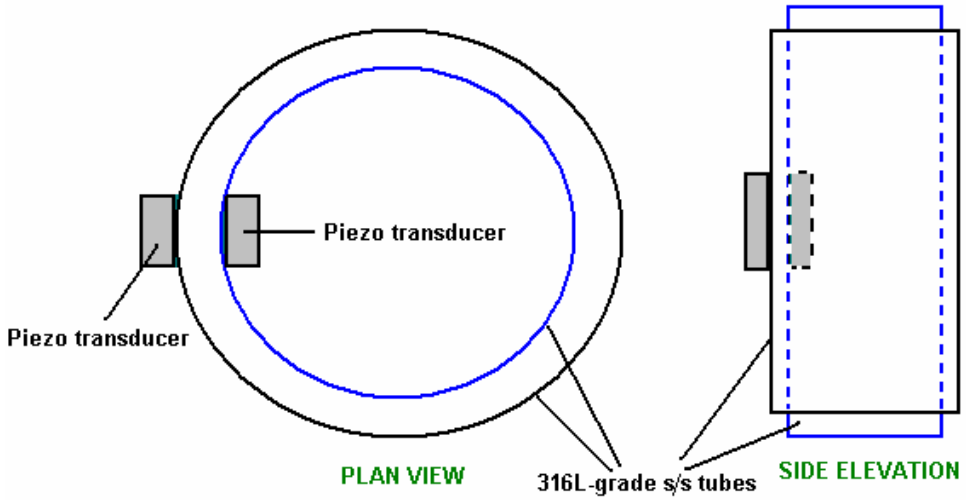


Another system of interest is the self-powered electrolysis system of the 1992 US Patent 5,089,107 granted to Francisco Pacheco where sacrificial anode plates of magnesium and aluminium are placed in seawater opposite a stainless steel cathode. Electrical power is generated and hydrogen produced on demand. There is also surplus electrical power available to run a standard electrolyser if so desired.

An Ultrasonic System for Producing Hydrogen.

I have been told (by a rather doubtful source) of a very high-performance water-splitting system which produces enough hydroxy gas to power a vehicle engine while only drawing 3 milliwatts at 3 volts which is a mere 9 milliwatts of power. I have never seen one of these units, and I have no evidence that the system works, other than word of mouth, so please treat the following entry as just a suggestion rather than a matter of hard fact.

The system is so interesting and simple that it is very attractive. Basically, you have two stainless steel pipes placed in a bath of tap water:



The objective is to get two stainless steel tubes resonating together at the same frequency. That is, they should both produce the same "musical" note when suspended on a thread and tapped. As the inner tube is smaller diameter, it will have a higher note than the larger diameter tube if they are the same length, so for them to match, it would be necessary for the inner tube to be longer, or the outer tube have a slot cut in it as Stan Meyer did and which is discussed in Chapter 10.

The piezo transducers are presumably glued to the cylinders, perhaps as shown above, and they are fed with a 2.24 MHz signal. The tubes need to resonate with the electronics signal, so they are ground down very slowly and carefully until they do resonate. This will presumably be at a much lower harmonic of the electronics signal, one in the standard ultrasonics range. Presumably, there will be three spacers top and bottom, maintaining the gap between the tubes. If the frequency were down in the mains region of about 50 Hz or 60Hz, then the device would just act as a water heater of the type designed by Peter Davey. At ultrasonic frequencies, the result is quite different as cavitation bubbles form in the water. A highly respected textbook on ultrasonics points out that these cavitation bubbles have a positive charge on one side and a negative charge on the other side and these charges cause electrolysis of the water surrounding the bubbles. Lots of bubbles - lots of hydroxy gas produced. So, background theory supports the possibility of this device working, however, I am not aware of anyone who has attempted to replicate it.

What we have not been told is:

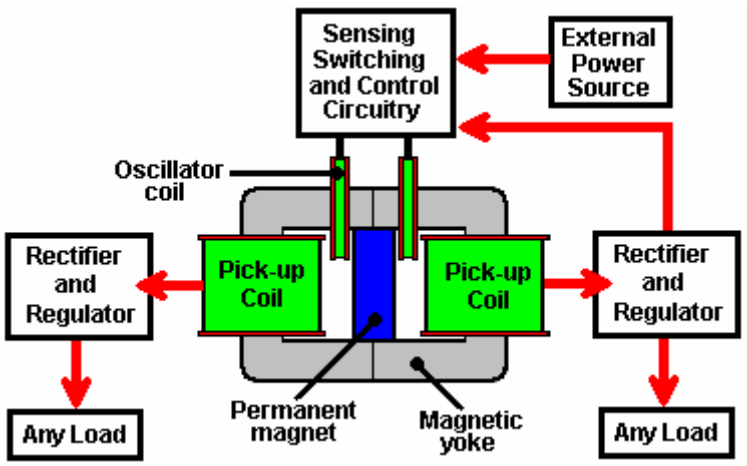
1. The size, length and thickness of tubes which work well.
2. The gap between the tubes.
3. The specific transducers used in the prototype.
4. What type of spacers were used.
5. Where and how the transducers were fixed to the cylinders.

However, even without this information, this could be an interesting investigation project using absolutely minimal power at trivial voltage levels.

The "MEG" Motionless Electricity Generator.

Tom Bearden, Stephen Patrick, James Hayes, Kenneth Moore and James Kenny were granted US Patent 6,362,718 on 26th March 2002. This patent is for an electromagnetic generator with no moving parts. This device is said be self-powered and is described and illustrated on JL Naudin's web site at **http://jnaudin.free.fr/meg/megv21.htm** where test results are shown. While this device has been claimed to have a greater output than its input and an output five times higher than the input has been mentioned, I am not aware of anyone who has attempted to replicate this device and achieved a COP>1 performance, and so, for that reason, it is described in this section describing devices which are unlikely to be worthwhile for the home-constructor to attempt to replicate.

The "**M**otionless **E**lectromagnetic **G**enerator" or "MEG" consists of a magnetic ring with output coils wound on it. Inside the ring is a permanent magnet to provide a steady magnetic flux around the ring. Superimposed on the ring are two electromagnets which are activated one after the other to make the magnetic flux oscillate. This is very much like Floyd Sweet's "VTA" device.



The external power source shown above is intended to be disconnected when the circuit starts operating, at which time, part of the output from one of the pick-up coils is fed back to power the circuit driving the oscillator coils. The circuit then becomes self-sustaining, with no external input but with a continuous electrical output.

If you should construct one of these, please be warned that it should **not** be started up unless there is an external load across the pick-up coils, otherwise dangerous, potentially lethal voltages can be produced. Don't get yourself killed or injured - please be very careful.

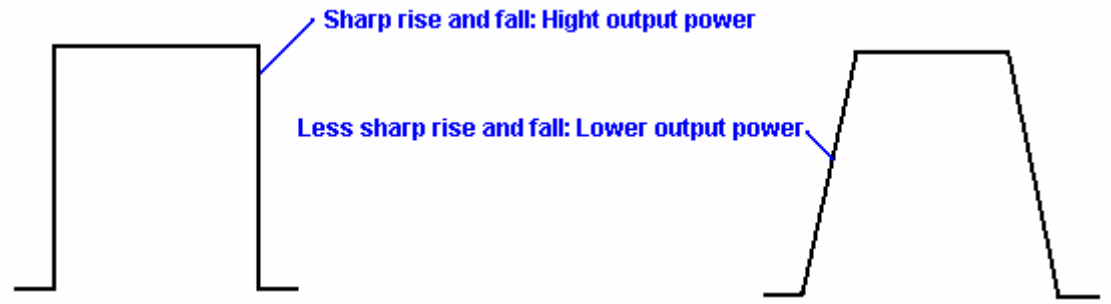
A re-worded excerpt from the patent for this system, is in the Appendix and it gives the construction details of the prototype: dimensions, number of turns, materials used, drive frequency, monostable pulse durations, etc. The prototype produced two outputs of 48 watts for one input of 12 watts. This allowed the input power to be taken from one of the outputs, while that same output was powering other loads.

This device is essentially, a custom-built transformer with two primary windings (the oscillator coils) and two secondary windings (the pick-up coils), with a permanent magnet inserted to create a standing magnetic field through the yoke (frame) of the transformer. However, a permanent magnet has two separate energy streams coming from it. The main field is the magnetic field which is very well known. It normally flows out in every direction, but in the MEG, a very good conducting path is provided by the frame of the device. This traps the magnetic energy flow and channels it around inside the frame. This prevents it masking the second energy field which is the Electrical energy field. With the magnetic field moved out of the way, it is now possible to tap this energy field for additional power output.

The MEG looks like a very simple device, but in actual fact, it is not. To act as a successful device with a Coefficient of Performance (COP) over 1, where the input power which is provided is less than the useful power output of the device, then Tom says that the frame needs to be made from a nanocrystalline material. This material has special properties which give the MEG it's exceptional output.

Care has to be taken with this device as the output power can be so high that it can burn the insulation off the wires and destroy the device if the output power is not controlled carefully. The output power is normally limited to a COP of 5.4 for practical reasons. If the necessary input power is taken from the output power via a rigorous control circuit which prevents runaway, then the device can provide output power while no outside input power is needed.

The output power is controlled by the waveform being sent to the oscillator coils. The power is controlled by the exact shape of the "square wave" drive:



This waveform is adjusted carefully to keep the COP down to 5.4 for safety sake. The waveform is also adjustable for frequency and Mark/Space ratio.

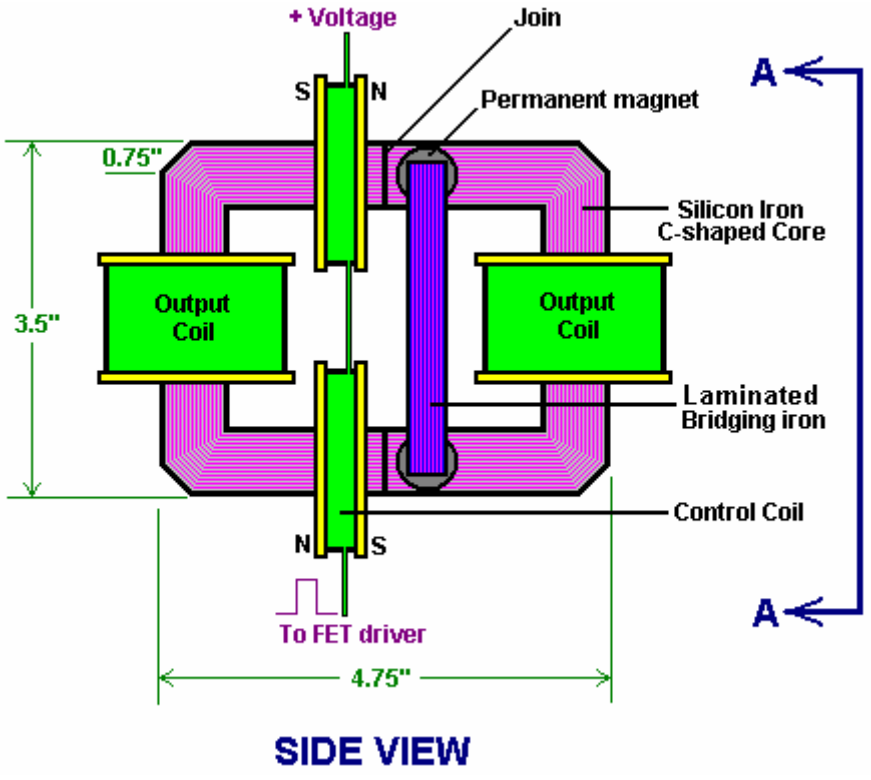
As it is some years since this device was patented, the question can be asked as to why it is not in production and offered for sale everywhere. The reason is that the MEG is a laboratory prototype which needs careful adjustment and tweaking. It has been replicated by others and it's performance verified as being COP>1, but it is not yet ready for production where it is necessary to have the design enhanced to the stage that it can be assembled in a factory and work immediately without the need for manual adjustments. That development is in hand and may be completed in the next year or two.

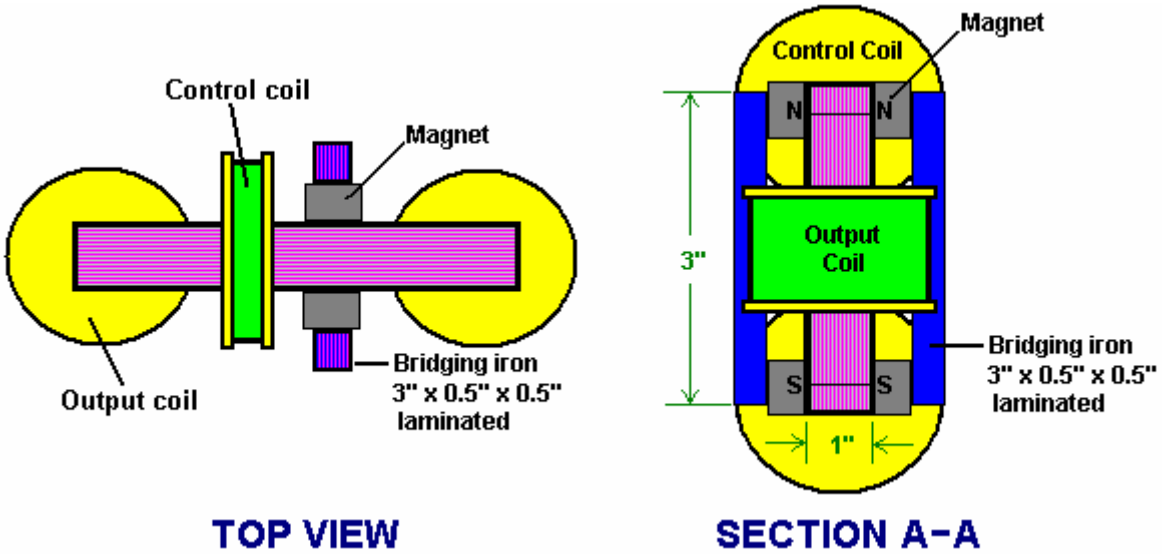
Some further explanation is in order. The MEG has an overall efficiency, well below 100% in spite of having a **C**oefficient **O**f **P**erformance well in excess of 1. The COP of 5.4 mentioned earlier is an arbitrary figure selected by the designers to prevent the insulation being burnt off the output wires. The actual maximum output is almost unlimited, certainly a COP of 100 is perfectly possible, but quite unnecessary in practical terms.

If a standard laminated iron yoke is used for the MEG, it will never have a COP>1 as input power will be needed to make it operate. The magnetic flux from a permanent magnet consists of two components. One component is rotary and it spreads out in every direction. The second component is linear and it gets swamped and hidden by the rotary field. If a torroidal yoke wound with an input winding over its whole length is used, then that traps all of the rotating magnetic field inside the torroid. The snag is that this requires considerable input power to energise the torroidal winding. The big advance with the MEG is that the inventors have discovered some standard off-the-shelf nanocrystalline materials which have the property of trapping the rotational magnetic field inside a torroid formed from them, **without** the need for any energising coil. This is a major boost to the functioning of the device.

Now, with the rotational magnetic field trapped inside the torroid, the liner field becomes accessible, and it is a very useful field indeed. It is electrical in nature. In actual fact, magnetism and electricity are not two separate things, but instead, they are different aspects of the same thing, so both should really be referred to as “electromagnetism”. Anyway, the linear field is easy to access once the rotational field has been removed. All that is necessary is to pulse it sharply. When that is done, real electricity is introduced into the MEG from the surrounding environment. The sharper the waveform, the greater the additional electrical input becomes. This is what makes the MEG have a COP of say, 5.4 which is a practical working output. If the output is then manipulated to provide the input power needed for the pulsing, the COP effectively becomes infinite as you do not have to provide any power to make it work and you have a substantial power output. The power output divided by the power input you have to provide to make the device operate, gives the COP rating, so any output divided by zero input, always gives infinity.

Dave Lawton has experimented with the MEG arrangement, using a professionally constructed custom laminated iron yoke. He found that using the standard arrangement, he found no difference when he removed the permanent magnet. Testing various configurations, he found that the most effective set-up for his components is:





Here, the drive coils are both put asymmetrically on one side of the frame and wired so that their pulses complement each other. Then two pairs of button magnets are placed on the other side of the centreline, each side of the yoke, and bridged together with two straight vertical sections of laminated iron bar. This arrangement is sensitive to the exact position of these magnets and tuning is achieved by moving the group of four magnets and two bars (effectively two “horseshoe” magnets) slightly left or right to find the optimum position. Introducing or removing these magnets then made a considerable difference to the operation of the device.

While most people find it very difficult to get COP>1 operation from the MEG, I am reliably informed that WITTS have succeeded in producing self-powered replications. There is, on the web, an exceptionally well-written presentation by Jacco van der Worp entitled “The Hitchhiker’s Guide to the ‘Free Energy’ MEG” at http://yowusa.com/scitech/2002/scitech-2002-06a/1.shtml and it explains how and why it can work. By kind permission, it is reproduced here:

A Hitchhiker's Guide to the "Free Energy" MEG

YOWUSA.COM, 25-June-02
Jacco van der Worp

Foreword by Marshall Masters

This article by Jacco van der Worp a Dutch physicist, explains the "free energy" Magnetic Energy Generator MEG simply, in layman's terms. I worked closely with Jacco on this project and sent a polite request to Lee Kenny, one of the MEG principals for assistance. His quick response was controlling and reeked of paranoia. "You are NOT AUTHORIZED by MEL to publish any information regarding the MEG." *(Readers can view and comment on the full text via the YOWUSA message board post, MEG -- Masters and Kenny.)* At first, his paranoid reply puzzled me, but as we further investigated the MEG we came to understand the paranoia and it chilled our blood. Perhaps this particular genie should be left in the bottle.

My first reaction to Kenny's suppressive controlling behavior was to shoot back a reply reminding him that there is a funny little piece of paper called The Constitution. After that, the urgent paranoia in Kenny's response continued to task me.

An obvious and simplistic explanation is that Bearden et al. are busy subscribing to Yachting Magazine and setting up bank accounts in offshore tax havens for their new to be wealth. The problem is that is too simplistic an explanation, because after working with engineers for twenty-five years I've learned to recognize a knee-jerk paranoid reaction when I see one. Kenny's reply was quintessential a knee-jerk

paranoid reaction. Perhaps, I would have reached this conclusion early on, were I not so enamored with the MEG technology. But as dug deeper, I found more tangible reasons for Kenny's paranoia.

The process of invention is a passionate one. At the end of the process are the hopeful financial rewards, but for something as complex as the MEG there is an equally large reward -- peer recognition. "Damn your eyes, but you did it." In this heat of invention the post reality dimensions are discussed but at a lower level and with less emphasis. I believe this is what may have happened to the MEG inventors. They simply got so carried away with the passion of invention that it clouded their view beyond the moment they would receive their patent, which in turn would embrace their paradigm shifting theories with plausibility.

However, once the patent was issued, the types of articles one would expect to see in magazines like Popular Science and Popular Mechanics failed to materialize. Further, what was made available by Bearden et al. was such geeky technobabble that it usually sails over the head of the average man. I now believe there was a reason for this.

The MEG opens a doorway into a new future, filled with new applications that will build upon the fundamental concepts of the MEG. Like the nuclear genie America unleashed upon Japan, the MEG genie offers us both a new and better world as well as a dead world all in the same breath. I believe this explains Kenny's paranoia. Now that he and his fellow MEG/MEL partners are looking at the practical applications of their invention they are starting to see it from "outside the box" as they say in the computer business. So what could they be seeing that makes them want to control the flow of knowledge about this new invention?

While the focus now is on the "free energy" application of MEG technology, its ability to interact and deform the space time continuum offers an alarming new range of applications to include: Weapon of mass destruction, interstellar space drive and time machine engine. And that is only the beginning of the hypothetical possibilities that could emerge from this paradigm shifting invention.

However, before we this genie is let out of the bottle, our elected representatives and all of us in the mainstream need to understand the MEG a little better. That is the goal of this superb article by Jacco van der Worp.

YOWUSA.COM
Marshall Masters, Publisher

A Hitchhiker's Guide to the "Free Energy" MEG

This fourth article in a YOWUSA series on the Motionless Electromagnetic Generator by Tom Bearden et al., now under patent, discusses this groundbreaking technology in non-technical layman's terms as opposed to delving deeply into the physical principles behind the MEG. Once you have finished reading the article, you will have a general idea of how the MEG works and why its inventors call it a "free energy" machine.

For those with bolder interests in the precise formulas and equations describing the MEG, those working on the project have already published several technical articles on the Internet. The best of these is a 69-page PDF file called The Motionless Electromagnetic Generator: Extracting Energy from a Permanent Magnet with Energy Replenishing from the Active Vacuum. This document lays out the most basic principles used for the MEG with some graphics added in.

Getting to know the MEG

The MEG represents a revolutionary approach to generating energy in our ever more demanding economy. However, it may also become a huge financial threat to big energy companies, as Marshall Masters outlined in his article Bearden's Free Energy MEG Destined For Chapter Eleven.

However, if the MEG grows into its destined role, it will conquer the market. It will do so by providing mankind with a durable and above all clean source of energy, which so far no other source of energy has achieved at an affordable price.

For the common person, the MEG is an obtuse technical wonder and understanding why it works requires one to dive headfirst into what seems to be a bottomless pool of technobabble. However, if we pare the issue down to five basic technological terms, we can create a basic understanding of the MEG.

In a MEG device there are five basic terms playing an important role in making it work. They are (energy) flux, efficiency, capacitance, magnetic field and shielding. This article will attempt to explain each MEG term

simply, using everyday examples. After the basic explanation, we will explore a full analogy in the form of a rain barrel.

Flux

Flux, or energy flux to be more precise, is essentially a by-product of energy creation. For example, an automobile engine uses a mixture of gasoline and air to power the car. Older engines in particular were not extremely efficient in burning all of this mixture and the by-product is the smog that comes out the tailpipe of the car.

Modern car engines are more economical; less unburned gasoline goes out their tailpipes. However, even the most efficient cars on the road today still push unburned gasoline out the tailpipe. On top of that, part of the heat generated by the burning of the gasoline inside the engine also leaves that way. In a manner of speaking, the smog that comes out of your car's exhaust is like flux. It is a by-product of the process of creating the energy needed to make your car travel down the road.

Flux and Efficiency

Before the oil embargo of 1973, American car designers largely ignored the issue of flux by building cars with larger engines to go faster. Consequently, this fuel economy suffered from the embargo. However, in 1973 the price of fuel did not make this an important issue. Once the fuel price rose after 1973, American auto manufacturers began losing sales to foreign manufacturers who were building more fuel-efficient engines.

Efficiency

We define efficiency as the useful application of energy flow. This means that if we build a more fuel-efficient engine, it will send less unburned fuel out the tailpipe and thereby increase the number of miles we can go on the same amount of gasoline.

The part we call useful here is the energy of the burned gasoline whose heat is used to move our car. That is what we aimed for, that part divided by the total energy that could be produced from the amount of gas we used, is called efficiency.

Efficiency and Unity

Since 1973, the goal of auto manufacturers is to use the available energy in gasoline as much as possible. The most they can hope for is what is called unity. Unity means that a system can use 100% of the energy input into it. In this case, that would be the energy stored in the gasoline to generate the desired output, the output in this case is the horsepower our engine delivers to move our car down the road.

No matter how efficient the designs are in pushing toward unity, the most a system can hope to achieve with present technologies is around 30% efficiency, which is some 70% short of unity.

Unity and Closed Systems

When we commonly think of unity (100% efficiency), we also tend to think in terms of closed systems. A closed system is a system that is completely isolated from the rest of the world. In a manner of speaking, an automobile engine is a closed system. The energy released from the burning of the gasoline to extract its energy is either captured to power the car or blown out the tailpipe as wasted flux. Likewise, you continually reduce the amount of gasoline in your tank as you drive down the road. The total amount of energy that is either still stored in unburned gasoline or is moving your car along or is leaving the tailpipe as waste flux is constant.

The essential point to keep in mind about a closed system is that it does not draw energy from the environment around it. On the other hand, an open system does draw energy from its surroundings.

Closed Systems and Open Systems

If we were to create automobiles with open systems, what would they look like? They would have the same engine, but an automobile with an open system would for example have a mast and a sail.

In this case, the driver would use the automobile engine to raise and lower the sail. The raised sail completes an open system by capturing motion energy from the wind and using it to propel the car down the road (provided the wind is with you.)

Once you have raised the sail, you can make the car go forward by angling the sail to the wind. To slow you can position the sail in line with the direction of the wind thereby neutralizing it as an active, open power system. Once you raise the sail, you do not need the car's engine for propulsion.

The essential point to keep in mind about an open system is that it draws energy from the environment around it like the sail on our modified car. In addition, it is controllable in that you can turn it on and off, just like a closed system.

Open Systems and the Coefficient of Performance (COP)

Because the auto sail is not limited to the energy of the fuel in the gas tank (which can never exceed unity or 100% efficiency), an open system (portrayed here by the auto sail) can in fact exceed unity. In other words, it can obtain more than 100% efficiency from the energy system it is interacting with, which in this case is the wind. So how do we measure systems than can achieve more than unity?

The term Coefficient of Performance (COP) compares the amount of energy input into a system versus the usable amount of energy output by the same system. For example, if using the engine in our auto to raise the sail requires one unit of energy and that when the sail is raised it generates or gathers 8 units of wind motion energy, the Coefficient of Performance is 800%. Another way of expressing this is 8:1, where 8 is the useful energy available and 1 is the amount of energy inserted into the system to trigger the generation of usable energy.

Coefficient of Performance and Capacitance

In our car example we used a sail to propel our car to achieve an 8:1 COP. Could this work with larger systems? For example, could it work with trucks as well? We already know that sails were used to propel huge wooden sailing ships centuries ago.

Let's go back in time to sailing ships for a moment. The builders of these great ships knew that the more and bigger sails they added to their ships, the faster they would go. However, great expanses of water surround that ship on the ocean. With this in mind, let's see how this applies to using sail power to propel cars and trucks.

Because cars and trucks travel in close proximity to each other and in the same direction on a highway, the amount of wind energy available in the open system must be divided amongst the various vehicles. Consequently, those in closest proximity to the wind direction will bleed off wind power from those ahead of them. In essence, they literally steal each other's wind.

In terms of capacitance, the essential concept here is that open systems may not be unlimited systems. In the case of our auto and truck sails, the proximity and number of sails accessing the same open wind system for energy will drain the open system. In other words, even if an open system is free, capacitance tells us that there is only so much of it to go around.

At this point, we've discussed some general terms used with the MEG. Now we're ready to discuss terms that are more specific to the MEG, beginning with fields.

Fields

In very simple terms, a field is any mechanism that serves as a means to an end. In the car example, we used a sail as our field, in that it gathered energy from the open wind system.

Unlike the sails we use as fields to capture the energy we need to propel our cars, in terms of the MEG, we must use something called a magnetic field.

Magnetic Fields

In the case of the auto sail, we used our sail as a field to pull energy from the open wind system around us so that we can propel out autos. The energy we needed was stored inside the wind.

With the MEG, the energy source we need to tap is not the wind, but electromagnetic forces of the universe that are just as omnipresent as the wind is on Earth, if not more so.

What the sail and the MEG have in common concerning the fields is that they must control their fields in order to prevent undesirable side effects.

Looking at our auto sail example, we attach lines to the sail and boom to control the trim of the sail so that it captures as much energy as possible without overstressing the sail and causing damage. Hence the popular sailing term, trimming the sails. Likewise, untrimmed sails can be dangerous, presenting a hazard to systems such as the boom and mast, or to the sailor if the wind moves the sail and boom violently across the ship, upsetting the balance of the ship.

In a manner of speaking, the MEG uses something called shielding to achieve the same thing as controlling the trimming of the sail if you will. Without it, violent effects may damage its surroundings.

Magnetic Fields and Magnetic Shielding

With the MEG, the magnetic fields are very powerful and must be controlled tightly at all times in order to prevent them from creating havoc in the space around them. This is why the MEG needs to use magnetic shielding.

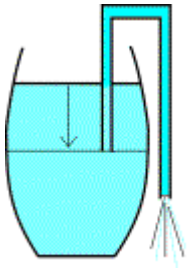
The most important concept of magnetic shielding is that it serves as a safety control for magnetic fields by containing and minimizing their negative effects.

The Rain Barrel Example

At this point, we've covered all the bases with the exception of the magnetic vector potential, which forms the crux of the MEG theory. To help you to understand the complexity of this concept, let's first review what we have covered up this point within the context of a simple rain barrel system. The reason for this is that one may understand the MEG magnetic vector potential more easily from a systemic viewpoint.

Flux

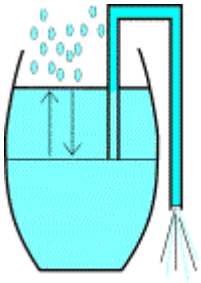
Some of us may have tried the following as kids or even later in life. If we take a barrel filled with water (or a gas tank filled with petrol) and we want to take some of that out, we do not have to suck it all out ourselves.



We take a piece of hose; simple garden hose will do, and stick it into the reservoir from which we want to take the liquid. On the outside, we lower one end of the hose a little lower than the opposite end sitting inside the tank. Then we gently suck on the hose (let's keep to water from here if only for the sake of taste) and the fluid will start to flow. Once it does, it will continue to flow until the other end of the hose inside the tank is no longer submerged. Therefore, with only a little effort we move a lot of fluid out.

The mechanism that makes this work is called the capillary effect. In other words, the weight of the column of fluid in the hose with a height equal to the difference in height of the two ends of the hose is providing the force that is needed to keep the fluid moving. However, what we do know is that the water barrel will run empty if we just pour it all out.

On the other hand, the MEG draws energy from a ‘barrel' that fills itself right back up! So it never runs empty! If you repeat the stimulated energy flow out of the MEG, energy flux will come out of it continually; it will not run dry like our rain barrel.

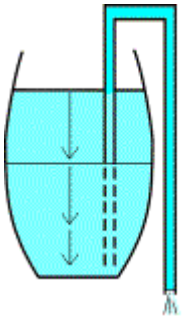


Therefore, a proper way to describe the MEG therefore in terms of this example would be a rain barrel into which more rain would fall the instant that that one draws water from it. Once you start the water flowing through the hose, the rain starts falling into the barrel and replaces the water you're pulling out at a similar pace.

For this reason, a MEG-style water barrel will never run empty and the water will flow forever out through the hose once you have brought it in motion because the MEG is

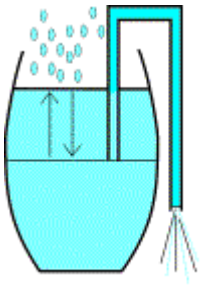
an open system, which brings us to the next point of consideration, what efficiency vs. COP means for our rain barrel.

Efficiency and the Coefficient of Performance



For the purpose of our rain barrel example, the term "efficiency" can be defined by the amount of water we can pull from the barrel by drawing it into motion with the siphon hose.

With the closed system water barrel, we learned that a closed system is isolated from the rest of the world, so no rain comes falling in from above to replace what we're taking out. In this case the best we can do is position the hose at the very bottom of the barrel on the inside and while letting it hand even lower on the outside. By doing this, we can siphon all of the water out of the barrel, which gives us maximum of 100% of the water. This 100% of efficiency is called "unity."



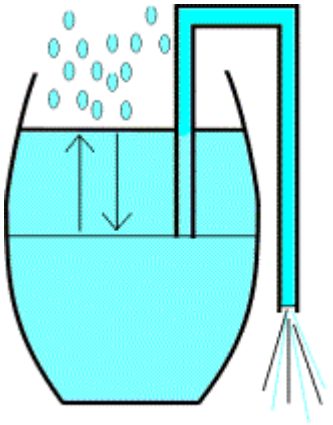
Keeping the unity of our closed rain barrel system in mind, let's shift back to our open system variant.

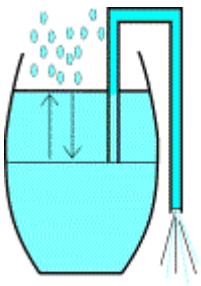
The moment we begin to siphon water from the open system rain barrel, fresh rainwater falls in through the open top of the barrel. No matter how much water we siphon from our open system water barrel, enough new rain falls through the top to replace what we are taking about.

Therefore, put in motion a never-ending stream of water with our open system rain barrel we can obtain results greater than that of unity. This is what the Coefficient of Performance (COP) is about. We use it to express the output result, which is greater that what we put in. Ergo, the COP for the open system rain barrel can exceed unity (100% efficiency) whereas the close system rain barrel can only hope to achieve unity. So then, what happens if we increase the size of our rain barrel?

Capacitance

Aside from the direct effect of having more water by increasing the size of our water barrel and siphon hose, there are indirect effects as well and they need to be carefully considered.





For starters, if we make our barrel bigger as well as our siphon hose what will that mean for us? Given that we'll be using a bigger hose to siphon out more fluid, we'll need a stronger suction force to begin with. We can do that simply (provided we have the lung power) without requiring a scaling-up of the entire system.

However, if we drain more water per second from a bigger barrel (e.g. the size of a lake) and we want it to keep running, it will have to rain harder to replenish the water that we take out and rainfall is bound to a natural limit.

At a certain barrel and hose size, not even a tropical storm will provide enough rain to keep the water level up and the system will start to collapse. On top of that, a normal barrel stands on a support structure. The bigger the barrel, the harder it will be to find a place for it to stand and remain standing. Otherwise it might fall over or break. So, how do we keep our bigger barrel from coming apart?

Fields and Shielding

If the barrel gets really big (let's assume for a moment it is the size of Lake Superior) and we start siphoning water out of it at the pace of four times the total water flux of the Sault Ste Marie Canals, then the water level will take time to readjust for the water poured out. The most important field in action here is the siphoning process, powered by gravity, which results in the water flow out of our bigger barrel lake. (Yes folks, now we're talking at a planetary scale.)

Normally, water level is horizontal (allowing of course for the curvature of Earth on a larger scale). However, if the pace of siphoning gets high enough, the normal water flow will become incapable of correcting the level quickly enough.

A permanent difference in height of the water level from one side of the lake to the other will arise. In that case stopping the siphoning action will not result in an immediate stop in water flowing towards the siphoning point. A sudden stop in pouring from a lake-sized barrel will cause at least a small tidal wave. The bigger the level difference across the lake, the worse the tidal wave will turn out. Although that looks like another field in action in the big barrel system, it isn't. It is a self-correcting mechanism for the lake surface after it has been disturbed.

In simplistic terms, what this all boils down to is that the flow energy like the flow water through our water barrel system represents a field. As a field increases in size it can likewise destabilize in greater amounts as well. Therefore, if we wish to increase the size of our fields we must find ways to shield them from those things, which could destabilize them.

Fields and Shielding

If our normal-sized rain barrel overfilled with water it could start to leak. In such a case, we would need measures to prevent a gushing flow of rainwater from damaging the immediate surroundings in a flood.

On the other hand, with our lake-sized version of water barrel we would need dikes to surround our lake to keep it from overflowing onto the land around it.

This flooding finds its cause in a pace difference between the raining in and pouring out of the water. These dikes must of course be able to withstand small tidal waves that emerge due to the starting and stopping of the siphoning action. In very simple terms, this is called shielding.

Up to this point, we've covered the most essential concepts we'll need before we tackle the big one -- vector potential. This is an important yet complex concept but it goes to the very heard of what a MEG is why it can do what it does.

MEG Components and Layout

The picture below, taken from the abstract by Magnetic Energy Limited as it published on the Internet. (We've added the colored placeholders to make it easier to view.) This illustration shows the basic layout of a lab prototype of the MEG used to successfully demonstrate the theory.

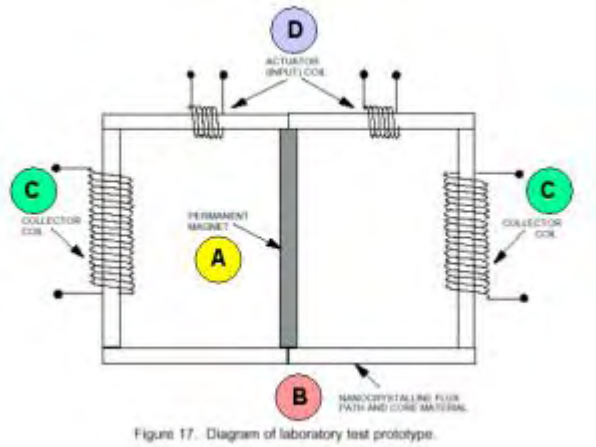


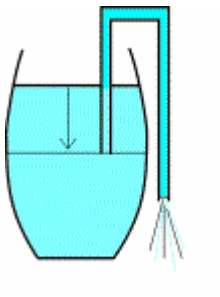
Figure 17. Diagram of laboratory test prototype.

- PERMANENT MAGNET (A): The most important element is the permanent magnet sitting in the middle of the schematic picture. The magnetic field lines come out of that bar magnet at the top and bottom side (in this picture). This magnet is what helps drive the entire machine.

- NANO-CRYSTALLINE FLUX PATH AND CORE MATERIAL (B): Instead of freely 'circling' from the North Pole of the magnet to the South Pole they enter a 'nano-crystalline flux path and core material.' That material

captures all of the magnetic field of the permanent magnet, so that no magnetic field is present free in the air any more.

- COLLECTOR COIL (C): The collector coils are the points where energy can be tapped from the MEG.
- ACTUATOR (INPUT) COIL (D): The actuators are the points of input of energy to put the much larger amount into motion.



To compare it to the rain barrel, the actuators are your mouth drawing water through the hose. The collectors are the hose ends hanging out of the barrel that starts to pour once you've generated a capillary with your siphon hose. And finally, the magnet and the coil containing the magnetic field are the barrel reservoir containing the water.

How the MEG Works

Now we come to the crux of the whole system, the reason why it works in the first place. The driving force in our large, lake-sized rain barrel was the force of gravity. That made the water that falls at some distance away from our hose move towards the initial siphon point and in turn caused the water to run through the siphon. In this case, gravity is the restoring force of the water level of our lake-sized rain barrel as it makes the water level go horizontal again.

This force and the correction mechanism attached to it have an equivalent in the magnetic arena. That force is the magnetic vector potential. If we look at the MEG, we see that it converts an energy flux that was stored in such a vector potential outside a closed magnetic field path. (Whoa, wasn't that a mouthful. Let's break it down into more simpler terms.)

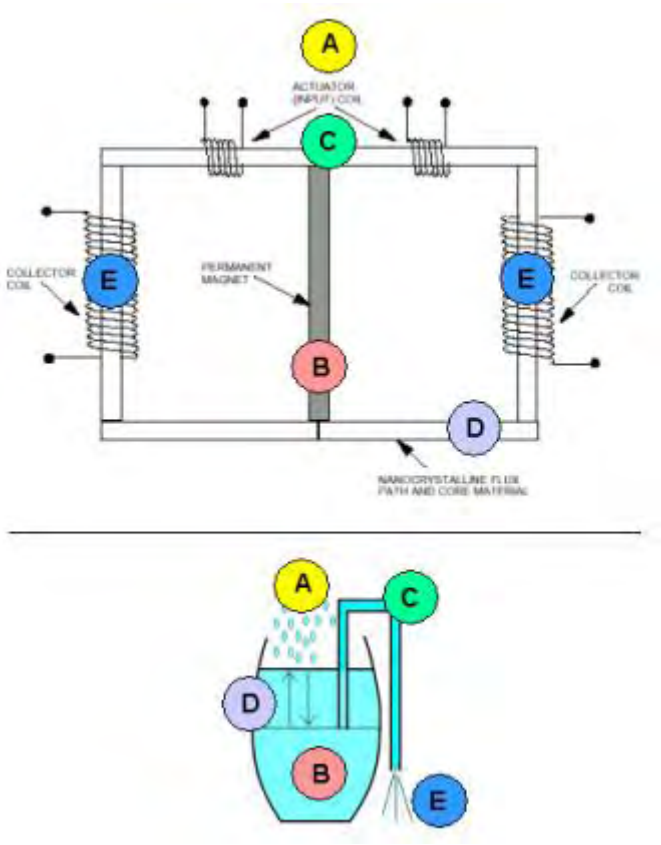
OK, So What is a Vector Potential Anyway?

To explain vector potential, we need to use something other than a rain barrel, but it must be familiar so why not the energy we use in our homes to run our computers, hairdryers, etc.

We all know about the electrical potential across the two wires of a wall outlet. This electrical potential is what makes a light bulb burn. If we're not careful with the outlet we could also find ourselves flat on our back as our family members frantically call for an ambulance. Thankfully though, the US, the electrical potential is 110 volts, which was chosen because it is not as lethal as the 220-240 volts standard found in most of the other countries in the world.

However, if we combine the numerical value of this electrical potential with a direction, we have a vector potential. In the case of our 110-volt outlet, if we change the direction from say the horizontal to the vertical, we can double our potential to 220 volts. Therefore, direction is important for the creation of any magnetic field; they all emerge from a magnetic vector potential where direction plays a critical role.

The following illustration compares the MEG with our rain barrel example for a very general layman's understanding of how the MEG works and why the magnetic vector potential is so important. (Please keep in mind a precise explanation would require an article several times the size of this one, so we'll just paint our picture with broad liberal brush strokes for now.)



- (A) An outside reservoir waits until the system is brought into motion and then starts to work to restore the balance that is broken by sucking the siphon hose. For the rain barrel it is the rain filling the barrel back up, for the MEG it is the vector potential converting its energy into magnetic field inside the closed path.

- (B) This adds an extra field to the magnet field inside the closed loop.

- (C) In essence the actuators work like the siphon hose in our rain barrel example and by changing the direction of the water it creates an outside vector potential.

- (D) Consequently, the closed path starts to interact with the magnetic field inside, to compensate for the change in the situation. It gives energy to the magnetic field inside the closed path.

- (E) We can then tap that energy from the collectors and we find that more electrical energy comes out the collectors than the amount we put in through the actuators. So energy from the vector potential field outside the closed path is 'flowing towards the siphon' to correct the 'field level' again. If we lead part of this energy back into the actuators again, the rest of it is free flux! Free flux?

With the MEG energy flux actually becomes the result product or output instead of a waste by-product as with fossil fuel powered systems like car engines What used to be waste is now useful output, just like today's electricity drawn from the net to light our houses.

In that respect, the MEG forms a new way of looking at energy flux and if you happen to install a MEG next to your home, it will require far less energy to provide you with far more electricity. Consequently, your energy costs will come down considerably.

Why the MEG is Commonly Misunderstood

The MEG uses an input energy flux to convert a far greater amount of thus far unusable energy flux into a controllable and more convenient form. This can leave some folks scratching their heads because this is a whole new twist on flux. For those who are familiar with the principles of conservation of energy this represents a paradigm shift in thinking that can defy years of heavily instructed thought about closed systems. And here is the rub. The rigid principles of conservation of energy apply only to closed systems such as automobile engines, whereas the MEG is an open system.

Because the MEG is an open system, it can turn flux into output because it is a system out of balance with the world around it and therefore constantly interacting with the environment around it! This way, it may result in a COP that is far greater than unity.

Another factor that makes it difficult for conventional thinkers to understand the MEG is that it does not use the Lorentz Gauge.

When Tom Bearden and his team of researchers discovered the principles behind the MEG when they chose to omit a commonly known calibration of an electromagnetic system, the so-called Lorentz Gauge.

The Lorentz Gauge is essentially a free choice of values for given parameters of an electrical system; this free choice makes mathematics simpler. At the same time however it discards a range of interesting (as it turns out now) solutions to a set of equations describing the same system. This range is the range of non-equilibrium states.

By keeping the MEG just off-equilibrium (out of balance) all the time, we can use it to pull a tremendous amount of energy out of a so far unusable reservoir into a convenient form. In essence, this is what the MEG is about.

Potential Problems with the MEG

We know that the MEG works, but it is also of interest to see just how much it can do. Most of you will likely be interested to know if a MEG can power a home. Can we scale it up without a limit, or could one such device even power a city?

The scale of many devices is only limited by practical design questions. The MEG needs a permanent magnet as well as a nano-crystalline material completely confining the magnetic field loops that leave this magnet. It also needs input and output coils. Electrical currents running through wires will produce heat, which will have to be dealt with at a high enough pace, but other than that, the potential size and productivity of the system is virtually unlimited. This type of generator should a priori be scalable to city-block level.

There are, however, possible side effects to its operation, which we want to take a closer look at before starting to operate a MEG for a city.

A few problems might exist for the MEG. Right now, it is uncharted territory, but we need to consider the possibility that above a certain level, the vector potential field cannot rearrange its energy fast enough for the working conditions to remain intact, thereby causing the MEG to fail. To fully illustrate this possibility, let's revisit some of the basic terms we discussed earlier in this article.

Magnetic Fields

The energy stored in the magnetic field and the vector potential field may interact with conducting materials outside the MEG as well, generating secondary magnetic fields and electric currents.

There's more. As energy leaves its surroundings, new energy comes flowing back in. We do not know if the pace of that is bound to a limit. It may have side effects that are currently unknown to us. Some pessimistic reactions have even spoken of an alteration of the space-time continuum surrounding the MEG. That would be a serious consequence indeed, but we have seen no proof of it so far.

Besides the argument presented above there is the coil material that is supposed to fully contain the permanent magnet's field and the additional field generated during interaction. A bigger MEG will also need a stronger magnetic field. Just how much can the coil material take before the field starts to break up the material itself? This bigger MEG may need superconducting materials to gain that stronger magnetic field. There will be additional conditions that are imposed by that material. We need to maintain that superconductivity to prevent damage to the material.

However, stronger magnetic fields do pose a health risk. For this reason, most people do not want to live under or very near high voltage power lines. They carry strong electromagnetic fields around them as well. It is however possible to shield magnetic fields.

Shielding

If someone were to switch off the input signal to a large MEG, the field may not die away instantly, which would result in a field spike as it follows suit to the signals. Then EM pulses may arise, which are (very) destructive to all electronic equipment. These EM pulses are so destructive in fact, that some nations have conducted extensive research into their possible application as a weapon.

A Faraday cage would form the "dike around the lake" for the MEG. This is a metal case enclosing the magnetic field plus the MEG completely. With the exception of strong EMP effects it will keep a semi-steady field contained so that no outside negative effects will occur. So in the course of normal operation, the MEG can be shielded quite easily. The worrisome moments occur when switching the MEG on or off. How will the shielding affect the ability of the vector potential field energy to replenish itself as energy leaves the unit?

We Need to Proceed With Caution

Concluding, we can say that the MEG is a means to pour energy from a tremendous reservoir with remarkable little effort. The result is almost limitless energy at practically no cost. The only drawback known in the current state of research is that on a large scale nasty, hard to control side effects may rear their ugly heads. Therefore while the MEG represents a ground breaking and innovative new technology it should not be rushed to market without exhaustive testing.

The Devices of Hans Coler

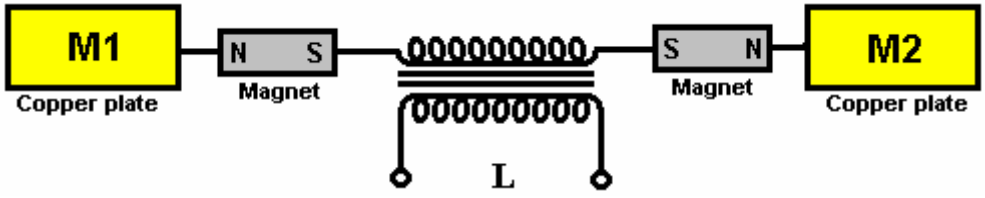
Hans Coler developed a device which he named the “Stromerzeuger” which consisted of an arrangement of magnets, flat coils and copper plates with a primary circuit powered by a small battery. The output from the secondary circuit was used to light a bank of lamps and it was claimed that the output power was many times the input power and could continue indefinitely.

The apparatus principally consists of two parallel connected spools which being bi-filar wound in a special way, are magnetically linked together. One of these spools is composed of copper sheets (the spool is called the ‘plate spool’). The other one is made of a number of thin parallel connected isolated wires (called ‘spool winding’), running parallel to the plates, at small intervals. Both spools can be fed by separate batteries (6 Volt, 6.5 AHr were used). At least two batteries are needed to get the apparatus operating, but subsequently, one battery can be removed.

The spools are arranged in two halves each by the bi-filar windings. The plate spool also contains iron rods with silver wire connections. These rods are magnetised by a special battery through exciter windings. Electrically, the exciter winding is completely isolated from the other windings. Hans said that the production of energy takes place principally in these iron rods and the winding of the spools plays an essential part in the process.

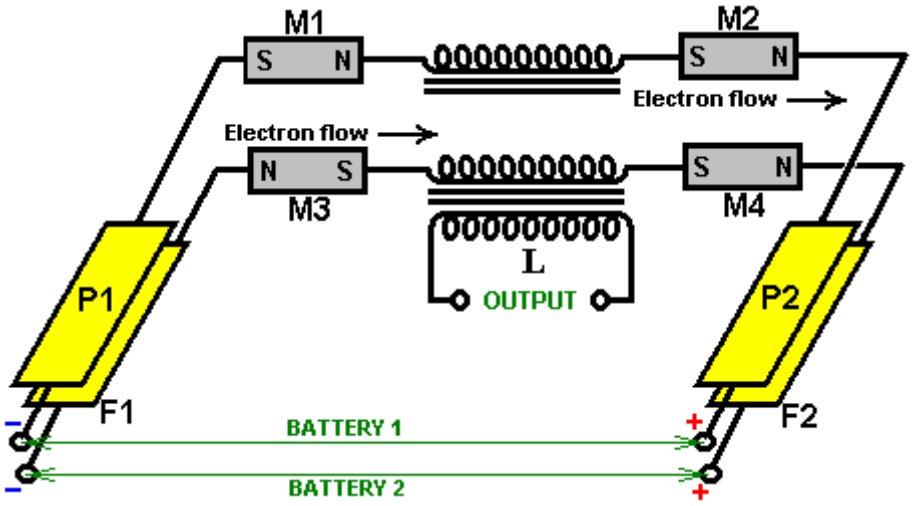
It should be mentioned that the spool circuit is powered up first. Initially, it took a current of 104 mA. The plates and exciter circuits are then switched on simultaneously. When this is done, the current in the spool circuit dropped from 104 mA to about 27 mA.

It is suggested that an electron be not only regarded as a negatively charged particle but also as a South magnetic pole. The basic Stromerzeuger element is that of an open secondary circuit, capacity loaded, inductively coupled to a primary circuit. The novel feature is that the capacities are connected to the secondary core through permanent magnets as shown here:



It is claimed that on switching on the primary circuit, “separation of charges” takes place with M1 becoming positively charged and M2 becoming negatively charged and that these charges are “magnetically polarised” when they formed, owing to the presence of the magnets. When the primary circuit is switched off, a “reversing current” flows in the secondary but the magnets “do not exert a polarising effect on this reversal”.

Two of the basic elements shown above are placed together making a double stage arrangement with the copper plates close together.



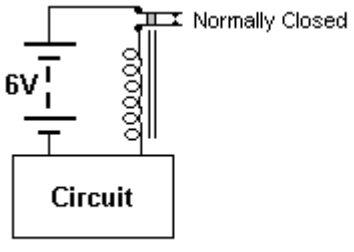
Note that the M1 coil is reversed. The magnets carry current and so it would be reasonable to presume that they are iron magnets rather than ferrite, iron magnets being common sixty years ago, combined with the fact that ferrite magnets do not conduct current and neodymium types were not available in 1925. The secondary windings are both exactly equal and "wound in a direction such that, on switching the primary coil on", the electrons flow from P1 to P2 and from F1 to F2. Again, this is a most peculiar statement as the direction of electron flow is not affected by the direction of winding of the coil, which suggests that the person making the report knew almost nothing about circuitry. It is highly likely therefore, that the brief report which has survived to today, is very vague and lacking in the essential details, the diagrams not indicating the bi-filar windings and straight wire lengths mentioned earlier.

One of the very unusual statements made is that the F1/F2 circuit only induces the effect and that the excess power is drawn from the P1/P2 circuit. This is the basic working arrangement. More of these double stages can be added to provide higher outputs, with their outputs connected in parallel for greater current.

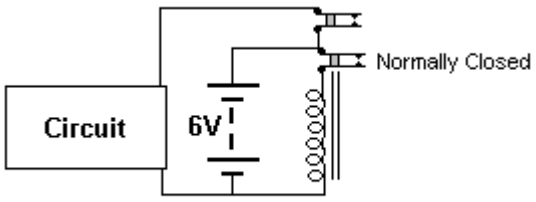
The original document, with masterly ambiguity, states that to make this arrangement, two of the original circuits are placed one on top of the other and then promptly draws the arrangement, showing the M1 magnet poles reversed, which of course, contradicts what has just been stated. So, it is not clear if the diagram or the statement is correct, but the likelihood is that the diagram is right.

For there to be any kind of continuous output from a transformer, the input cannot be DC in spite of the fact that dry-cell batteries were used to power the circuits (even though the output could be as high as six kilowatts). Coler stated that his research indicates that permanent magnets actually oscillate at about 180 kHz. If that is correct, then that could account for the magnetic induction produced in the circuit. If that is not what is relied on in this circuit, then the batteries have to be connected and disconnected continuously, and the technology being simple at that time, is seems likely that a relay would have been wired to self-oscillate as in the early electric bells. While that arrangement could have been used just to connect and disconnect the batteries with a relay contact, the relay coil will have generated high-voltage sharp pulses, generating short magnetic pulses of the type which causes an inflow of cold electricity into the circuit from the local ambient background.

Alternatively, although it is not mentioned or suggested, the connection could have been made using a relay which had only one contact, in which case, the back-EMF of the relay coil may have fed high-voltage pulses directly into the circuit. This is, of course, just ill-informed speculation. The arrangement **might** have been:



High voltage pulses



Low voltage pulses

In each case, the battery current flows the normally closed relay contact, powering the relay winding, which opens the switch contact, disconnecting the relay and so closing the switch contact again. This causes rapid pulsing of the battery voltage and the relay coil generates high voltages, typically of some hundreds of volts. If the circuit connection is as shown in the first arrangement, then when the relay switch opens, high back-EMF voltage pulses will be fed to the circuit. On the other hand, if a relay with two separate contacts is used, the second relay switch contact can be used to connect the battery to the circuit and disconnect it again in a very rapid sequence. There is a group of people at http://tech.groups.yahoo.com/group/testatika/ who are working towards making an operational version of this device. Information on the Coler devices is presented clearly at http://rimstar.org/sdenergy/coler/index.htm#STROMERZEUGER.

Patrick Kelly

http://www.free-energy-info.tuks.nl/
http://www.free-energy-info.com
http://www.free-energy-devices.com

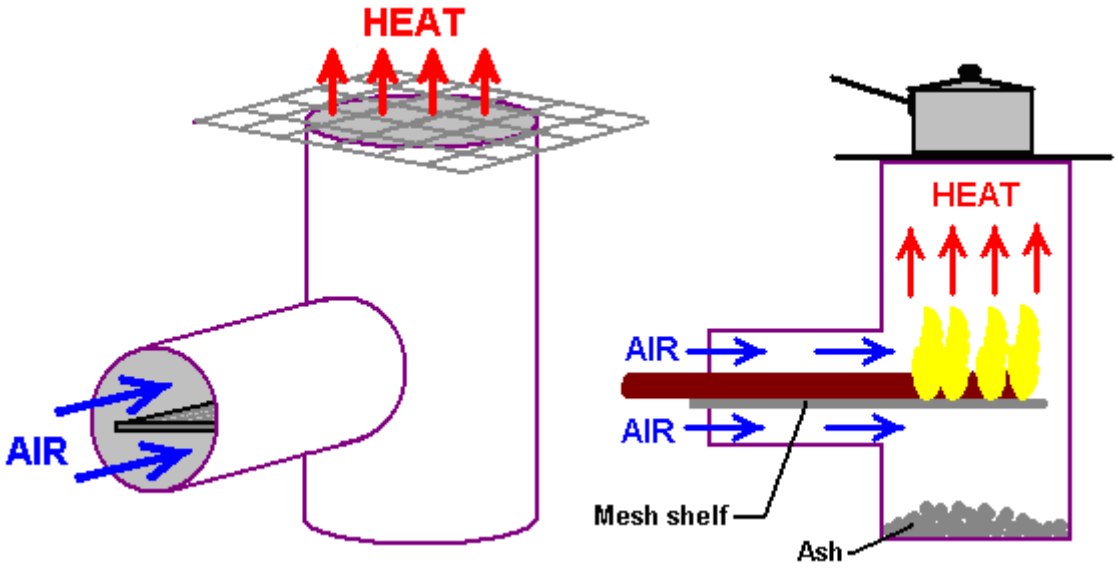
Chapter 14: Renewable Energy Devices

Note: If you are not at all familiar with basic electronics, you might find it easier to follow parts of this chapter if you read chapter 12 first.

Heaters.

The devices described here are not “free-energy” devices as such, but in spite of that, it is an area of considerable interest to many people, and the subject is included here because of that.

If you do not live in an urban area, then a solid fuel stove can be an economic solution, especially if the fuel can be collected free from wooded areas. Stove design has advanced considerably and it is now possible to make a simple stove with very high efficiency and very low emissions as shown here:

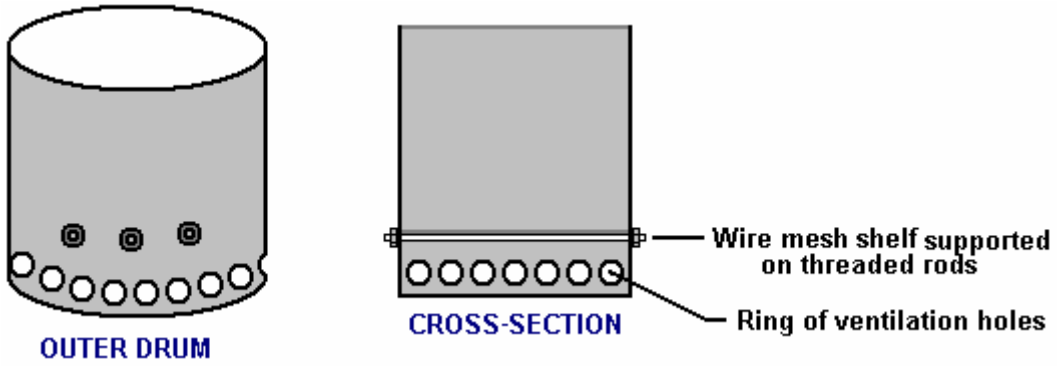


Although this stove is a very simple construction, it’s efficiency is very high indeed. The best fuel is made of smaller pieces which rest on a simple shelf. Branches work better than large pieces of wood as the consumption is more complete. As the fuel is consumed, it is pushed further into the stove, which gives the user an appreciation of the rate of consumption. Having the fuel resting on a shelf has the major advantage of allowing air to flow both above it and below it, which gives improved combustion. The operation is said to be so good that there is virtually no residue and no emissions.

Again, if land space is available, a solar oven (or Stirling motor) can be used, either to store energy for later use or generate heat for cooking or home heating, as can hot-water solar panels. However, it is only realistic to consider the application to be during the night in a built-up area with little or no spare space for equipment.

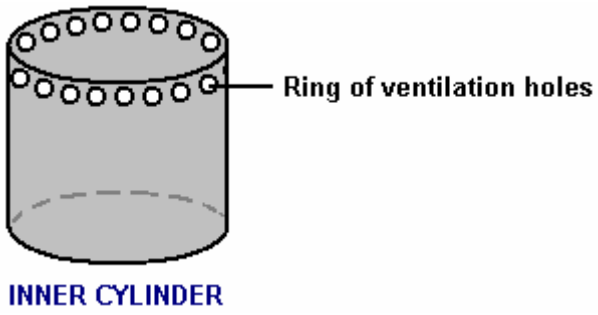
The Wood-Gas Stove.

There is another very effective type of wood-burning stove. There are several commercial versions of this stove on sale, but most of them use an electric fan to get the effect which is produced automatically by this design. This design has been replicated by Alberto Feliciano and found to be very effective in operation. It causes a load of wood which would normally burn up in fifteen minutes to burn for a whole hour, putting out a much greater amount of heat. The design is very straightforward. There is an outer drum which has a solid base, and the opposite end removed entirely as shown here:



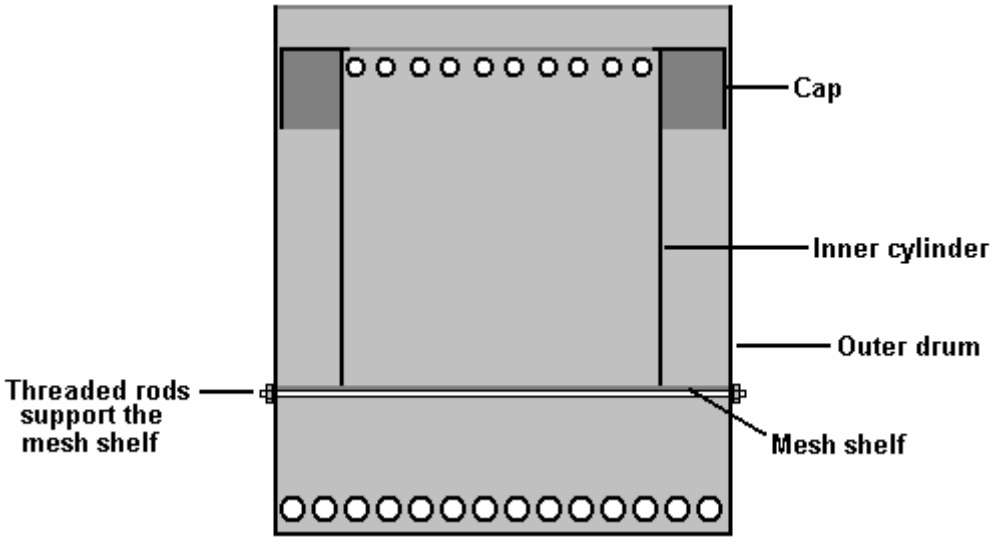
A ring of ventilation holes is drilled around the whole of the bottom edge of the drum and threaded rods are run through the drum. These support a circle of wire mesh which forms a supporting shelf.

A smaller drum then has both ends removed, to form a cylinder. This cylinder is to fit inside the outer drum. It rests on the mesh shelf of the outer drum. This cylinder has a ring of ventilation holes is drilled around the whole of it's upper edge as shown here:



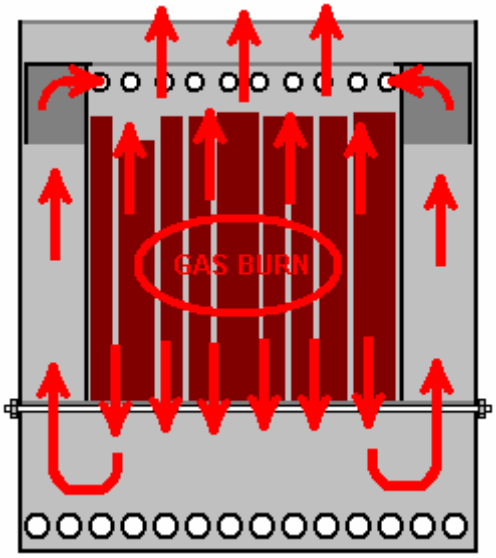
A third drum which is only slightly smaller than the outer drum is cut down to make a cap for the inner cylinder. This cap is not tight-fitting, but it effectively closes off the top of the gap between the sides of inner cylinder and the sides of the outer drum:

This cap has a circular hole cut in it, and this hole is only slightly smaller than the diameter of the inner cylinder. It is supported by the upper lip of the inner drum but the hole is large enough that it does not obstruct the flow of heat up through the top of the inner drum. These three pieces fit together like this:



Wood of any type, branches, sticks, firewood, etc. are placed inside the inner cylinder and set alight. Initially, the flames come out of the top of the stove as you would expect, but after a few minutes, the combustion alters completely. The burn now becomes that of wood-gas rather than of the wood itself. The wood is slowly converted to charcoal and the gas released by this process burns with greater heat than the wood would give as well as burning for a much greater length of time. This stove design can be made in any size. Small versions get a burn length of about three times that of the wood burning in the open air, while large versions can reach four times the burn length.

Flames no longer come out of the top of the stove, but instead, they come out of the ring of holes around the base of the outer drum. The process is like this:



A high-temperature gas burn takes place in the centre of the inner cylinder. This pushes heat out through the top and the bottom of the inner cylinder. The heat flowing out of the top is used for heating or cooking as before. The heat flowing out of the bottom gets diverted around the outside of the inner drum, flows upwards, is caught by the cap and fed back into the inner cylinder through the ring of holes at the top of the inner cylinder. This raises the temperature of the gas burn even further and it augments the heat coming out of the top of the stove. The result is a very hot burn which goes on far longer than would happen if the same wood was burnt on an open fire or in a single open drum. When the wood reaches the end of it's burn cycle, it can be replaced by feeding new wood through the hole in the cap which rests on the inner cylinder. The stove will need to be emptied of ash from time to time.

Mr Teslonian's Heating / Fuel-producing / Electricity and Refrigeration System

The Wood-gas stove shown above has been taken several steps further by "Mr Teslonian" as shown in his videos: http://www.youtubeskip.com/watch?v=arbXj9R6ZXw&list=PLw28_n7AgcmBtqfnxDV50M-V2JPj3vtFt and http://www.youtubeskip.com/watch?v=647R2R_VzG8 where he burns twigs in a home-made wood-gas furnace, heats his house, heats his water, produces engine fuel, runs an electricity generator off the gas and powers a refrigerator. All that, from a few twigs! This is very impressive and he is definitely to be congratulated for his development. His basic wood-gas stove can produce flames up to eight feet (2.5m) tall when operating very effectively on just a few handfuls of short twigs and branches. It looks like this:

This very effective wood burner produces enough heat to heat a house and produce hot water. The wood-gas can also be used to produce both crude oil and fairly well-refined oil which can be used in an engine:

And as you can see from the picture above, the wood-gas can run an electricity generator. There is a world of difference between burning wood like this and a typical camp fire or bonfire.

Wood-gas can also be used to run vehicles (as was done during World War 2). Links to a great deal more information include:

http://www.woodgas.net/index.html - an enthusiasts site with basic explanations and many links.

FEMA_emergency_gassifer.pdf - construction plans for any full size engine.

1938_Producer_Gas_Motor_Fuel.pdf - using wood-gas in a vehicle.

FOA 72.pdf - wood-gas as engine fuel.

Home_made_generator(12v).pdf - wood-gas generator for charging battery banks.

World bank tech paper 296.pdf - small scale biomass gasifiers for heat and power

http://www.free-energy-info.tuks.nl/Oil test.pdf – the best oils to use with wood-gas engines

Solid Fuel Producing Electricity

We then to think of small-scale solid-fuel burners as providing heat and perhaps a little light as well. The people in Siberia don't think like that and they produce a range of (just) portable solid fuel stoves which provide heating, cooking and 12-volt electricity up to 50 watts. While 50 watts doesn't sound like much, it is a major amount when there isn't any at all available. A small inverter provides mains AC power for smaller appliances:

If it is still there, the video of this is at http://www.youtube.com/watch?v=2Za98e7-Mxg&feature=related and the manufacturers (http://www.termofor.ru/eng/products/for_heating/normal_pg/index.html) show a video of this stove design being used in a snow-covered wood at -32 Centigrade temperatures, giving about +35 degrees Centigrade inside the tent. Mind you, weighing in at 54 kilograms which is a whopping 120 pounds or so, this is not a back-packing solution for camping trips. Sales outlets are in Siberia, New York, Australia, Kyrgyzstan, Ukraine, Belarus, Kazakhstan and Latvia at the present time.

The HHO Gas Option.

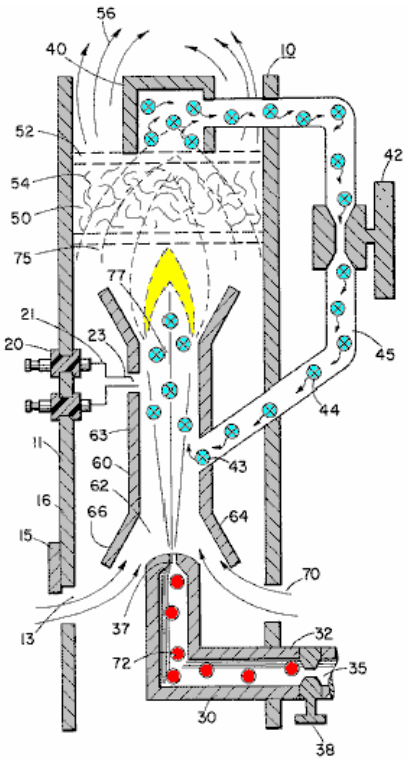
One significant advantage of this design is that less wood is needed for any give heating or cooking application and there is very little smoke or soot produced which is a distinct advantage for a person using the stove for cooking and standing nearby during the cooking process. This can cause a major medical

improvement as smoke inhalation for women who need to undertake long periods of daily cooking using wood as a fuel, are liable to suffer serious breathing problems as a result.

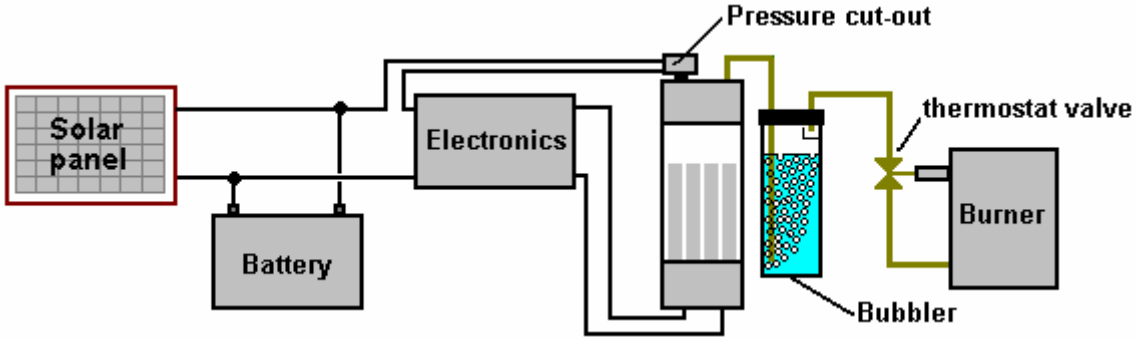
Electrical heating, while very convenient, is usually expensive, and it often seems that the effectiveness of an electric heater is not directly related to its power consumption. In theory it definitely is, but in practice it just does not seem that way. There are other alternatives.

One of the other documents in this set, shows how to construct a Stanley Meyer style electrolyser which uses ordinary tap water and splits it into burnable fuel using just a low power electrical input:

The difficulty in creating a heating system which uses the gas produced by this unit, is in the very high temperature produced when the gas is burnt. Stan overcame this problem with by designing a special burner which mixes air and burnt gasses in with the gas before it is burnt. That lowers the flame temperature to a level which is suitable for heating and cooking:



While this looks a bit complicated, it's construction is really quite simple. The combination of the Meyer electrolyser and Meyer burner form a system which has the potential of being operated from a solar panel and battery as shown here:



A system like this needs extreme care as the hydrogen / oxygen ("HHO") gas produced is explosive. So:

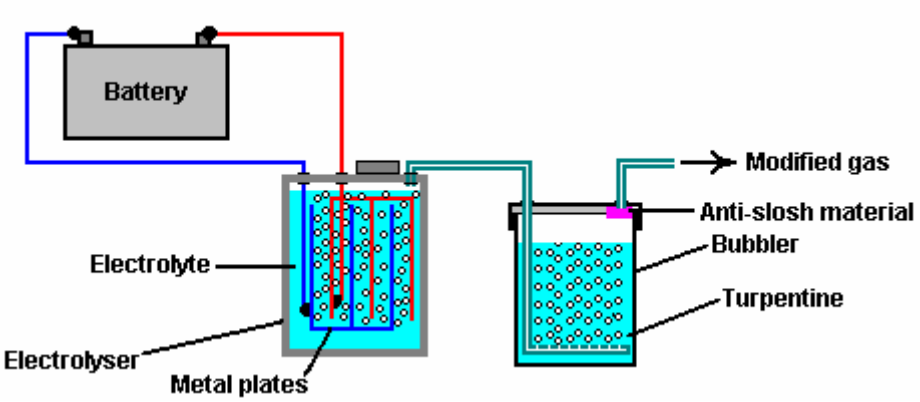
1. It is very important that the electrolyser has the ability to provide sufficient gas to keep the flame(s) sustained.

2. The electrolyser must be fitted with a pressure switch, typically operating at 5 pounds per square inch or so. This is included so that should the gas usage drop, then the drive from the electronics is cut off to stop further gas production, and incidentally, stopping the current draw from the battery.

3. It is absolutely essential that there be a flame-operated valve on the gas supply line to the burner, so that should the flame go out for any reason whatsoever, then the gas supply will be cut off. This type of valve is common on town-gas operated fires for use in homes.

Henry Paine's Patent.

There is an alternative method which it is claimed can convert the explosive HHO gas into a much more docile fuel, more suited to conventional burners and stoves. It must be stressed that this system is over 120 years old and it should not be used until you have carried out careful tests on it. David Quirey has been using this technique successfully for more than twenty years now. The method was patented by **Henry M. Paine** in US Letters Patent No. 308,276 dated 18th November 1884 and it is very simple:



The idea is to bubble the HHO gas produced by electrolysis of water, through a liquid hydrocarbon such as turpentine or acetone. The bubbler should have a large number of small holes in the incoming tube, so that a very large number of small bubbles of HHO gas pass through the hydrocarbon. This brings the majority of the HHO gas into intimate contact with the hydrocarbon and the process is claimed to convert the HHO gas into a new variety of gas which is not explosive, can be stored for later use, and which burns with the same characteristics as coal-gas ("town gas").

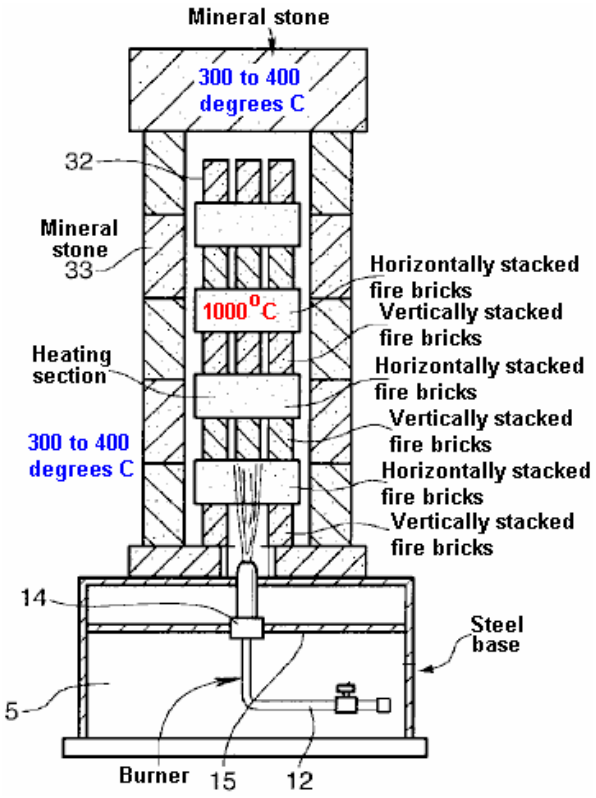
The Heaters of Sang Nam Kim.

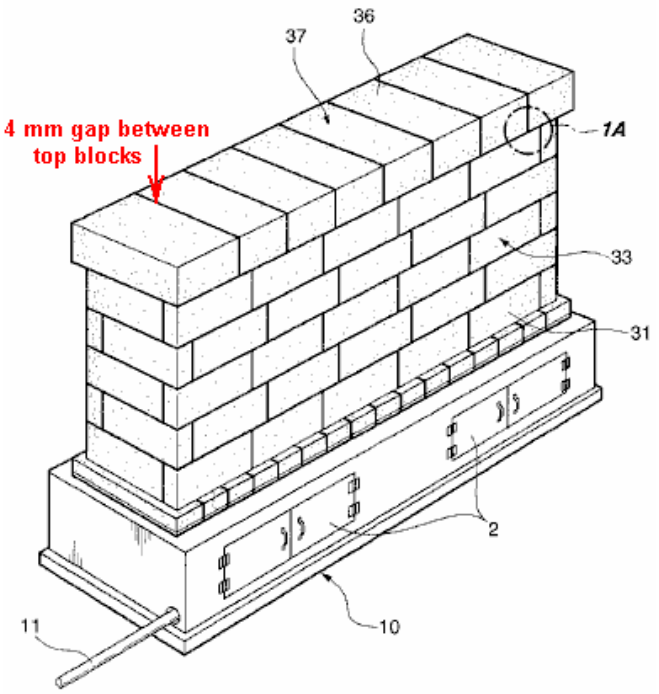
Mr Kim of Korea also proposes methods of using hydroxy gas for heating and using Henry Paine's hydrocarbon bubbling method. He has four patents on the subject of heating:

US 6,397,834 in June 2002 - Heating Furnace
US 6,443,725 in September 2002 - Energy Generation
US 6,761,558 in July 2004 - Heating Apparatus

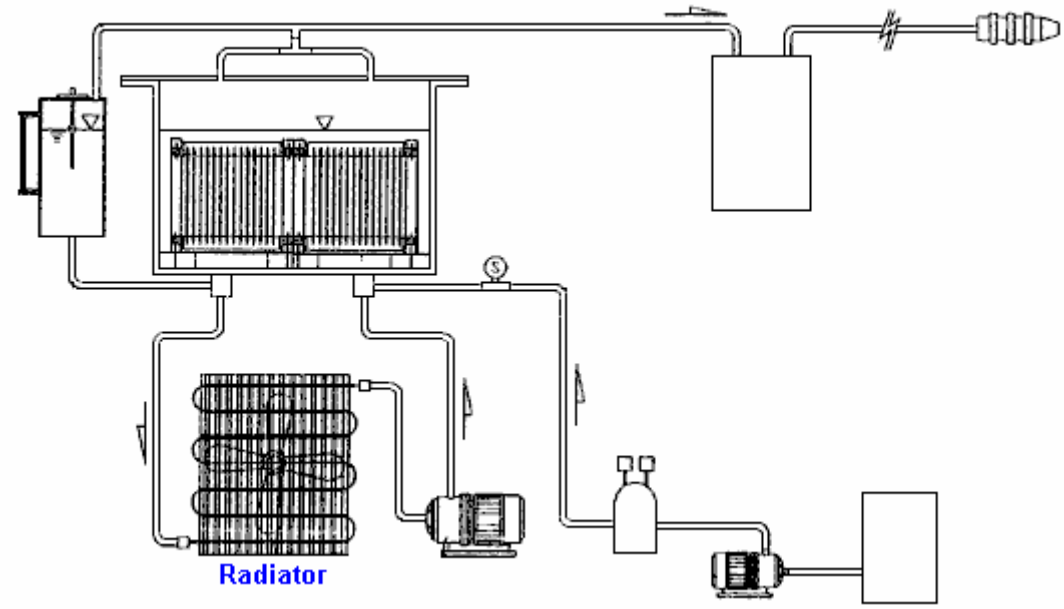
US 7,014,740 in March 2006 - Electrolyser

The first of these shows his method of getting both beneficial radiant heat and convention heating from a stone construction like this:



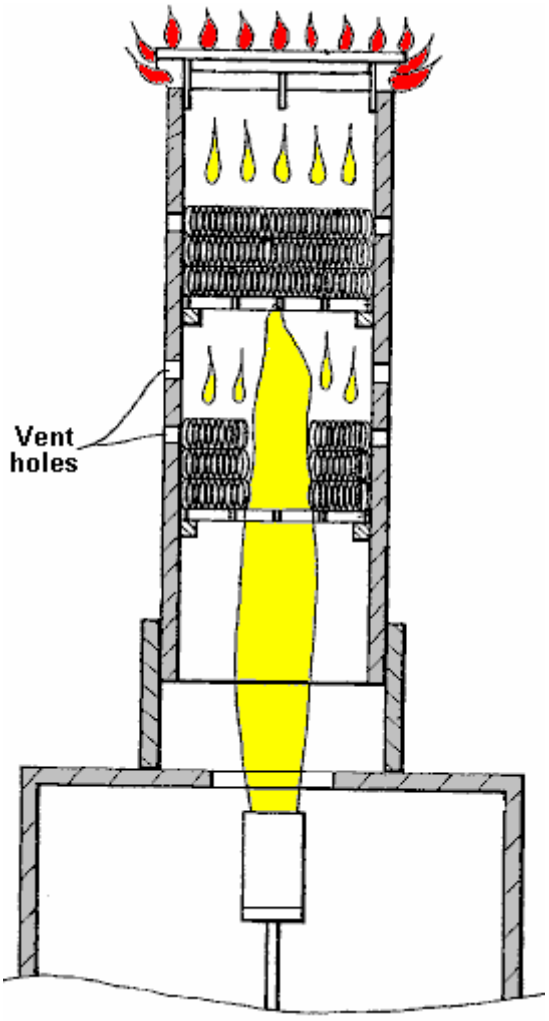


This unit is intended as a seriously powerful heating source for a minimum of one room. Mr Kim quotes a hydroxy gas requirement of 30 litres per minute which is a very considerable amount, and if the gas is produced by electrolysis of water at Faraday efficiency, it would need a current draw of 4.2 kilowatts. There is every indication that Mr Kim's method of electrolysis is low efficiency as his latest patent shows a radiator and fan:



It is virtually certain that there will be a high proportion of hot water vapour mixed in with the gas which is being burnt and that will reduce the explosive nature of the gas to near zero. It also means that the actual volume of hydroxy gas is likely to be well below the 30 lpm quoted by Mr Kim.

Mr Kim also shows a burner intended for use with an existing furnace. He remarks that the outer casing gets to be red hot, running at 1,000^{O}C or more, and so any replication of his design should be treated with care when mounting fixtures are being constructed. In this design, Mr Kim uses the Paine technique and recommends bubbling his hydroxy gas through hexane liquid (C_6H_{14}) where 0.3 litres of hexane per hour gets burnt as well as the hydroxy gas. He rates the hydroxy gas requirement of this burner as being 20 lpm, which at Faraday efficiencies, represents 2.8 kilowatts of electricity, although as mentioned before, it is likely that the actual amount of hydroxy gas in his 20 lpm volume is much lower than he thinks, and so will have a lower electrical requirement when using a more electrically efficient electrolyser. His burner is like this:



Mr Kim believes that the hexane prevents flashback ignition. He does not appear to specify the heating material inside the burner but it is probably stainless steel wool. He speaks of gas pressures of 1 Kg per sq. cm., which, if my calculations are correct is 14.22 psi. which is not possible for 100% hydroxy gas as it will explode spontaneously at 12 psi. due to its high energy state and electrical charge.

He states that the secondary flames at the top of the unit "burn with a blue flame colour" and that is different to the flame colour lower down. Mr Kim believes that the hydroxy burnt at the bottom of the burner forms water vapour which is then split into hydroxy gas again by the very high temperature and that is the reason for the blue flames at the top. Personally, I don't believe that this will take place and that the effect may have a good deal to do with the hexane liquid being burnt. However, this burner design appears to be a good one for lower grades of hydroxy gas.

If a higher grade of hydroxy gas is being used, please be aware that a hypodermic-size tiny burner orifice will be needed to avoid flashbacks and no commercial flashback arrester will work reliably with good quality hydroxy gas on every occasion and so a bubbler is absolutely essential.

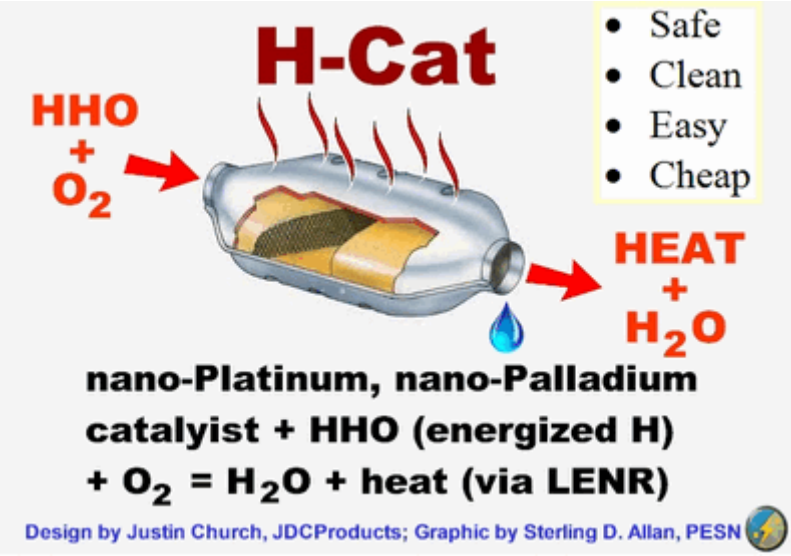
The Hydrogen Garage

In 2013, Andrew of the Hydrogen Garage in America stated:

We are heating up the shop with only 216 watts, 12 Volts at 18 amps using one “dual 7 plate” electrolyser cell with a water tank and bubbler and a 5-inch dryer, hose and torch which can be operated continuously 24/7 if desired. The flames produced are pointed into a central stainless steel tube 12-inches (300 mm) long and 1-inch (25 mm) in diameter. Wrapped all around the central pipe are 10 more stainless steel tubes held together as a group by 2 hose clamps. The set of pipes is placed on the concrete floor. In our work, we have to prepare HHO electrolyser cells and the HHO gas produced during that process is used to heat the shop. No more need for propane and no more headaches from 6 hours of heating the room. Now the HHO gas adds ozone to the air in the shop and there are no fumes and no smell. The burner used is shown at http://hydrogengarage.com/h2eat.html and it can be powered by one 14/7 HHO cell available at: http://stores.homestead.com/hydrogengarage/-strse-179/ICE-BLOCK-14-dsh-7-Cell/Detail.bok There is no need for a Pulse-Width Modulator unit as you can just use direct DC power from a DC power supply or battery charger. Solar panels can run HHO cells well.

Catalytic HHO Heating from Justin Church

On Sterling Allan’s website at http://peswiki.com/index.php/OS:Justin_Church%27s_H-Cat there is a description of an HHO heating process which has been under development since 2010. It is a particularly interesting system in that no flame is involved, but instead, a stream of HHO gas is fed into a standard vehicle catalytic converter along with air which can freely enter the converter. Justin has found that with a quite low flow of HHO, such as that produced by 13.8V at 5A (70 watts), the converter heats up to a level which can barely be touched and that is quite remarkable considering that the converter is has substantial weight of metal in its construction. Measurements of the temperature inside the converter show that it is running at more than five hundred degrees F. Justin calls his unit the “H-Cat”:



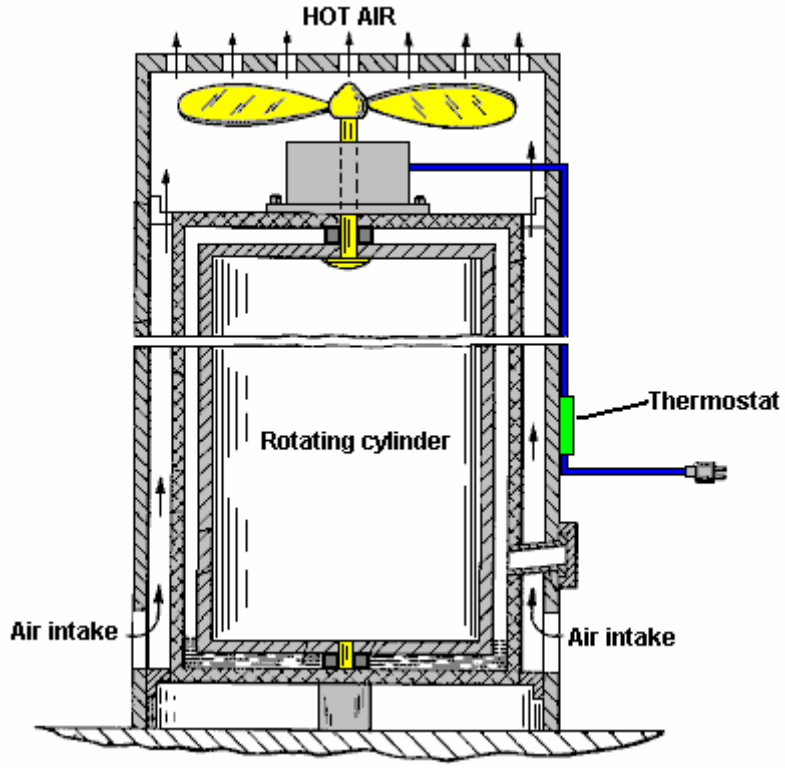
Sterling's web site is definitely worth a visit.

Heating With Electric Power.

Electric power is very popular for heaters. However, with most appliances, it is a very expensive form of heating. There is a technique which is reputed to improve the efficiency and lower the cost of electric heating. This method involves rotating a cylinder inside an outer cylinder and filling part of the narrow space between the cylinders with some variety of light oil.

Eugene Frenette's Patent.

This method has been patented more than once. In 1979, Eugene Frenette was granted patent 4,143,639 where a single motor is used to rotate the drum and power a fan to boost the motion of the hot air:



It is not immediately obvious why this arrangement should work well, but it appears that it does. As the inner drum spins around, the oil rises up between the two inner cylinders. It lubricates the bearing under the rotating drum and the rotation causes the oil to heat up. This heats the middle cylinder and air being drawn

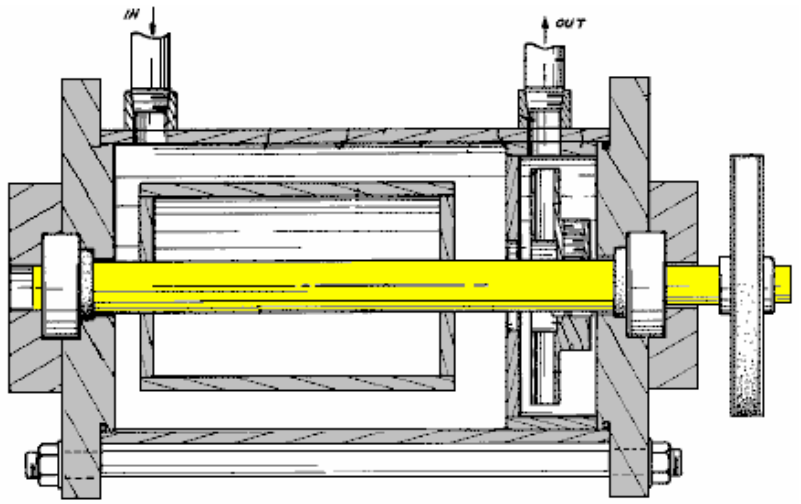
up around it by the action of the fan blade, is also heated before being pushed out of the top of the heater. After a few minutes, the outer housing becomes so hot that the thermostat attached to it, cuts off the electrical supply.

The heater does not stop heating at this time as air continues to circulate through the heater by ordinary convection. In my opinion, it would be more effective if the fan motor were operated independently and did not cut off when the heater reaches its operating temperature.

Eugene Perkins' Patent.

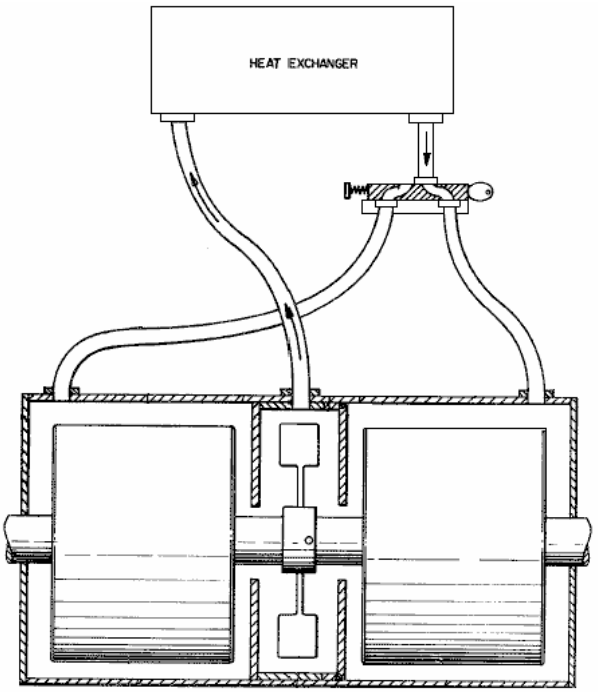
Very similar systems were patented by Eugene Perkins: January 1984 patent 4,424,797, November 1984 patent 4,483,277, March 1987 patent 4,651,681, October 1988 patent 4,779,575, and in January 1989 patent 4,798,176.

His first patent shows a horizontal drum which is completely immersed in the liquid:



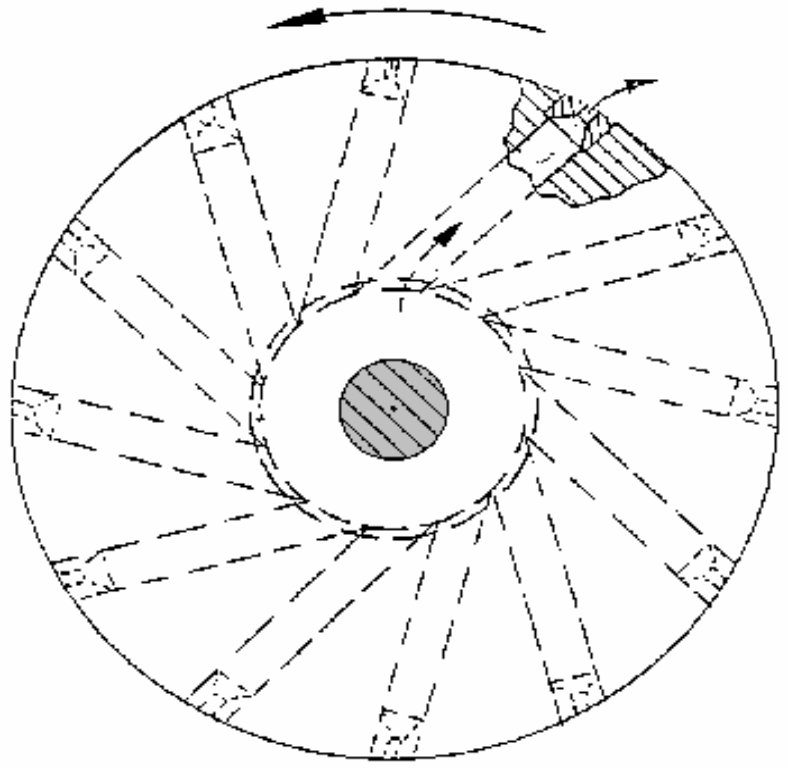
This calls for a much greater accuracy of construction in that the liquid has to be contained even though it has a rotating shaft running through the housing. This device pumps the heated liquid through central-heating piping and radiators.

In his later patent of the same year, he shows a modified version with two drums and an impeller:



The “heat exchanger” is a radiator or set of radiators.

He then progressed to a system where the shaft rotation forces the liquid to be expelled through the tips of arms radiating out from the centre of the impeller hub:



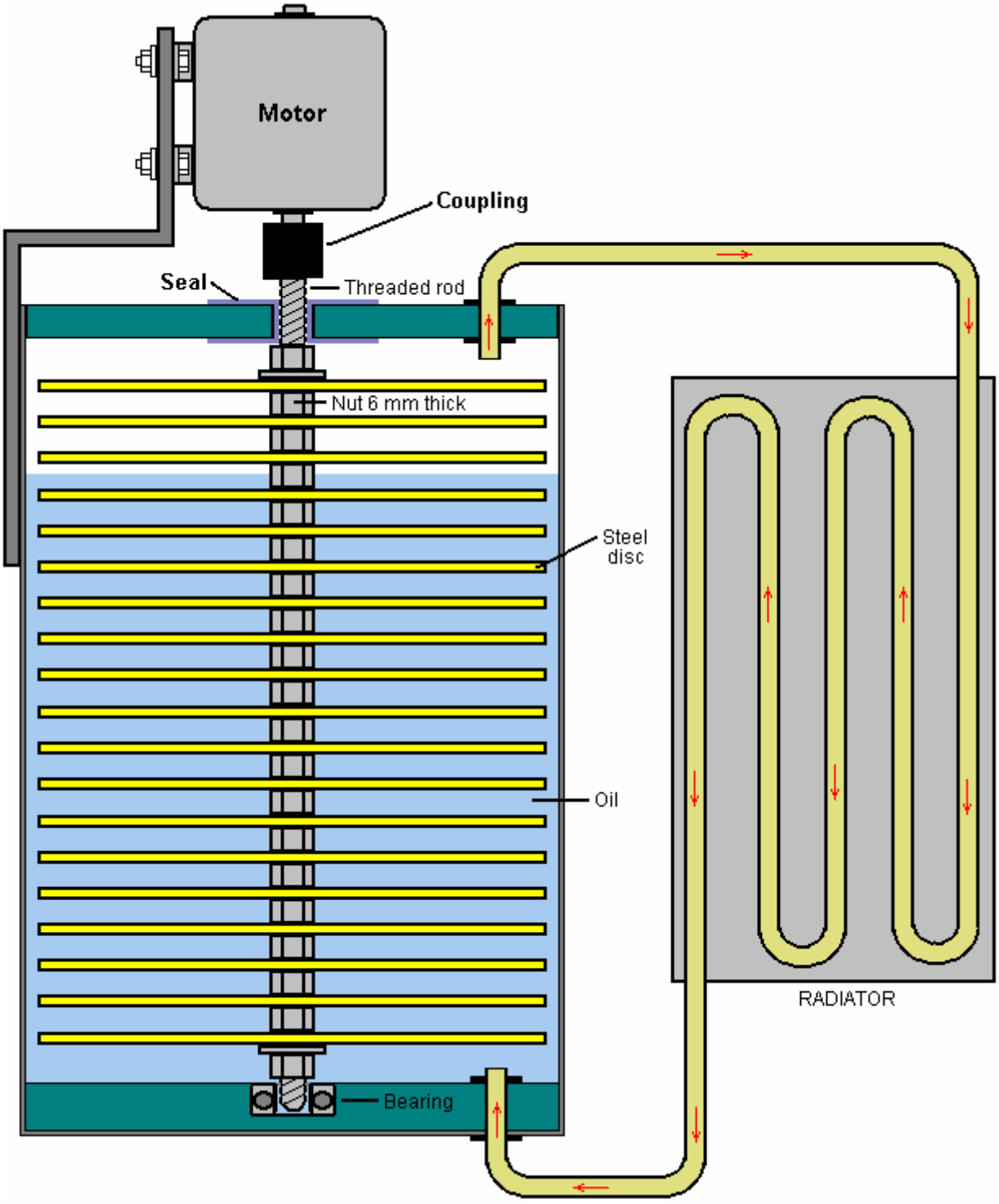
Here, the liquid is forced into a small space between the rotor and its drum housing. This system has been used very successfully for water heating and some measurements indicate that it is at least 100% efficient and some people believe that it is well over the 100% efficiency, though they don’t want get drawn into long discussions on methods of measurement. It is sufficient to say here, that this method is very effective indeed.

Frenette Variation: The Frenette heater design shown above with it’s two vertical cylinders, is not the easiest for the home constructor unless one of the cylinders (presumably the inner one) is constructed from steel sheet, as it is difficult to find two commercially available steel cylinders of just the right relative size to produce the wanted gap between them. A much easier variation replaces the inner cylinder with a stack of circular steel discs. As these can be cut from 20 gauge steel sheet fairly readily by the home constructor, or alternatively, cut by any local metalworking or fabrication company, any available size of outer cylinder can be used and the disc diameter chosen accordingly.

The discs are mounted about 6 mm (1/4”) apart on a central steel rod which is rotated in order to drive the discs through the oil contained inside the body of the heater. While this looks like a Tesla Turbine, it is not because the spacing of the discs creates a different effect. The wider disc spacing creates shear as they spin through the surrounding oil, and this shearing creates a high degree of heating. It must be remembered that this is a heater, and the outer canister gets very hot during operation (which is the whole point of the exercise in the first place). For that reason, oil is used as a filling and not water, which boils at a much lower temperature. The larger the diameter of the canister and the greater the number of discs inside it, the greater the heat developed.

To ensure that the discs do not come loose during prolonged operation, a hole can be drilled through them just outside the area covered by the locking/spacing nuts, and a stiff wire run through the holes and the ends either welded to the central rod or pushed through a hole drilled in it and bent over to hold it in place. The heat of the cylinder can be circulated by attaching a simple fan blade to the spinning shaft. This blows air down the hot sides of the canister, moving it towards the floor which is the most effective place for it circulate and heat the entire room.

As the discs spin, the oil is pushed outwards and moves upwards, filling the top of the canister and building up some pressure there. This pressure can be relieved by running an external pipe from the top of the cylinder back to the bottom, allowing the oil to circulate freely. This has the decided advantage the circulating oil can be passed through a radiator as shown in the following diagram:



The central rod can be rotated by any convenient motor, conventional, Adams type, pulse-motor, permanent magnet motor, or whatever. An alternative to this style of operation, is to use the rotating motor to spin a ring of permanent magnets positioned close beside a thick aluminium plate. The eddy currents cause very strong heating of the aluminium plate which then can have air blown across it to provide space heating.

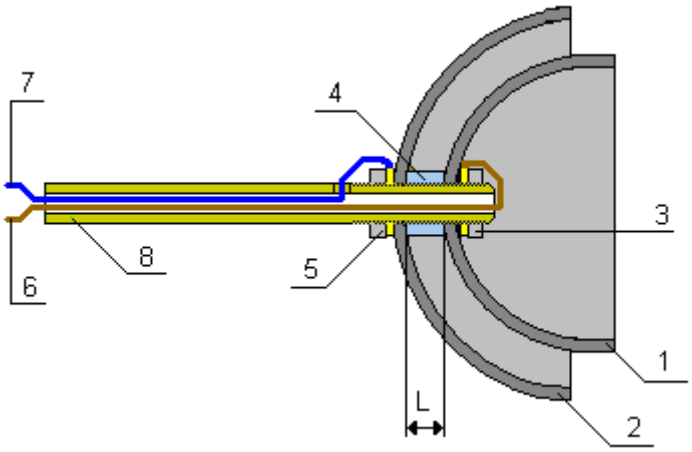
Peter Davey's AC Heater.

During World War II, Peter Daysh Davey, of Christchurch, New Zealand, a fighter pilot and musician, designed and built an unusual water heater. This design is not particularly well known and information is fairly thin on the ground although the basic principle and design details are known.

The device is intended to operate on New Zealand's mains supply of 220 volts 50 Hz and a requirement of the apparatus is that it resonates at that 50 Hz frequency. Resonance is a frequent requirement of free-energy systems, and the need for it is often overlooked by people who attempt to replicate free-energy devices. Properly built and tuned, this heater is said to have a COP of 20, which means that twenty times as much heat is produced as would normally be expected from that level of input power. This power gain is caused by additional energy being drawn in from the surrounding environment, and this is very important as a very large percentage of the energy used in cool climates is used for heating, so any reduction would be very welcome.

Peter was granted a New Zealand patent for his heater on 12th December 1944 but he found that opposition from the utility companies was so great that it prevented him from going into commercial production with it and although Peter kept trying for fifty years, he never succeeded.

The device is comprised of a hemispherical resonant cavity, formed by two metallic dome shapes, both of which resonate at 50 Hz. Initially, Peter used two bicycle bells and he found that when submerged in water, the device brought the water to the boil in a very short time indeed. The construction is like this:



If construction were to use two identical hemispheres, then the cavity between them would be anything but even width throughout, but the resonance would be the same. On the other hand, if you want the resonant cavity between the two hemispheres to be of constant width, then the outer hemisphere needs to be markedly larger than the inner hemisphere. The outside of both hemispheres needs to be insulated unless mounted in such a way that it is not possible to touch the hemispheres as both are attached to the mains.

In the diagram shown above, the mains live wire **6**, is fed through the connecting pipe **8**, and clamped to the inner hemisphere **1**, by nut **3** which screws on to the threaded section of tube **8**. It is important that it is the live wire which is connected to hemisphere **1**. The mains neutral wire **7**, is also fed through the connecting tube **8**, then exits through a small hole and is then clamped on to the outside of the outer hemisphere **2**, by nut **5**, also on the threaded section of tube **8**. The two hemispheres are held apart by a spacing washer **4**, which is made from a high-temperature non-conducting plastic. As the tube **8** connects electrically and mechanically to both mains wires through the two locking nuts **3** and **5**, it is essential that this tube is constructed from an electrically non-conducting material such as plastic. as the tube will be in boiling water

on a regular basis, it is also necessary that the tube material is able to handle temperatures over 100^{o}C and so possible materials include nylon and Teflon.

The washer is a key component of the heater and it's thickness is key to the efficiency of the whole device. This thickness **L**, is the tuning control for the cavity. The outer hemisphere is about 8 mm greater in diameter than the diameter of the inner hemisphere. Allowing for the thickness of the metal of the bowl, the resonant cavity will therefore be about 3 mm or one eighth of an inch.

The hemisphere **1** is also tuned to 50 Hz by grinding it carefully until it resonates freely at that frequency. Connecting a loudspeaker in series with a resistor of say, 100K ohms, will give a sound of the exact frequency with which this hemisphere needs to resonate. This tuning needs to be done with the unit fully assembled as the connections to the tube will alter the resonant frequency of the hemisphere. When this is being done, the resonance will be felt rather than heard, so hold the tube lightly so that it can resonate freely. The tuning is done by removing a small amount of metal from the face of hemisphere **1** and then testing for resonance again.

When hemisphere **1** resonates well at the mains frequency, (roughly G two octaves below middle C on a keyboard), the search for high-efficiency heating is carried out by very small adjustments of the gap **L**. The adjustment of the gap **L** is carried out by very careful grinding down of the separating washer **4** and the result is best determined by measuring the length of time needed to boil a known volume of water and the current taken to do that. Repeated tests and recorded results, shows when the best gap has been reached and the highest efficiency achieved. The heater can, of course, be used to heat any liquid, not just water.

This heater is unlike a standard kettle heating element. In the standard method, the water is not a part of the main current-carrying circuit. Instead, the mains power is applied to the heater element and the current flowing through the heater element causes it to heat up, and the heat is then conveyed to the water by conduction. In Davey's heater, on the other hand, the current flow appears to be through the water between the two hemispheres. It seems likely that the actual heating is not produced by current flow at all, but from cavitation of the water caused by the resonating of the cavity between the two hemispheres. This technique is used in small jewellery cleaners where and audio frequency is applied to a cleaning fluid in a small container.

A small amount of electrolysis will take place with the Davey heater as it in effect also forms a single parallel-connected electrolyser. The amounts should be very small as only 1.24 volts out of the 220 volts applied will be used in the electrolysis process.

An early construction of the original heater is shown in the photograph below. The coin shown in the picture is 32 mm (1.25 inch) in diameter. The heater is submerged in water when it is being used, and it brings that water to the boil exceptionally quickly. The unit was tested by New Zealand scientists who were able to vouch for its performance, but who were unable to state exactly how its operation allowed it to output such a high level of heat for such a low level of electrical input. You will notice from the photograph, how carefully the electrical connections and outer bowl are insulated.

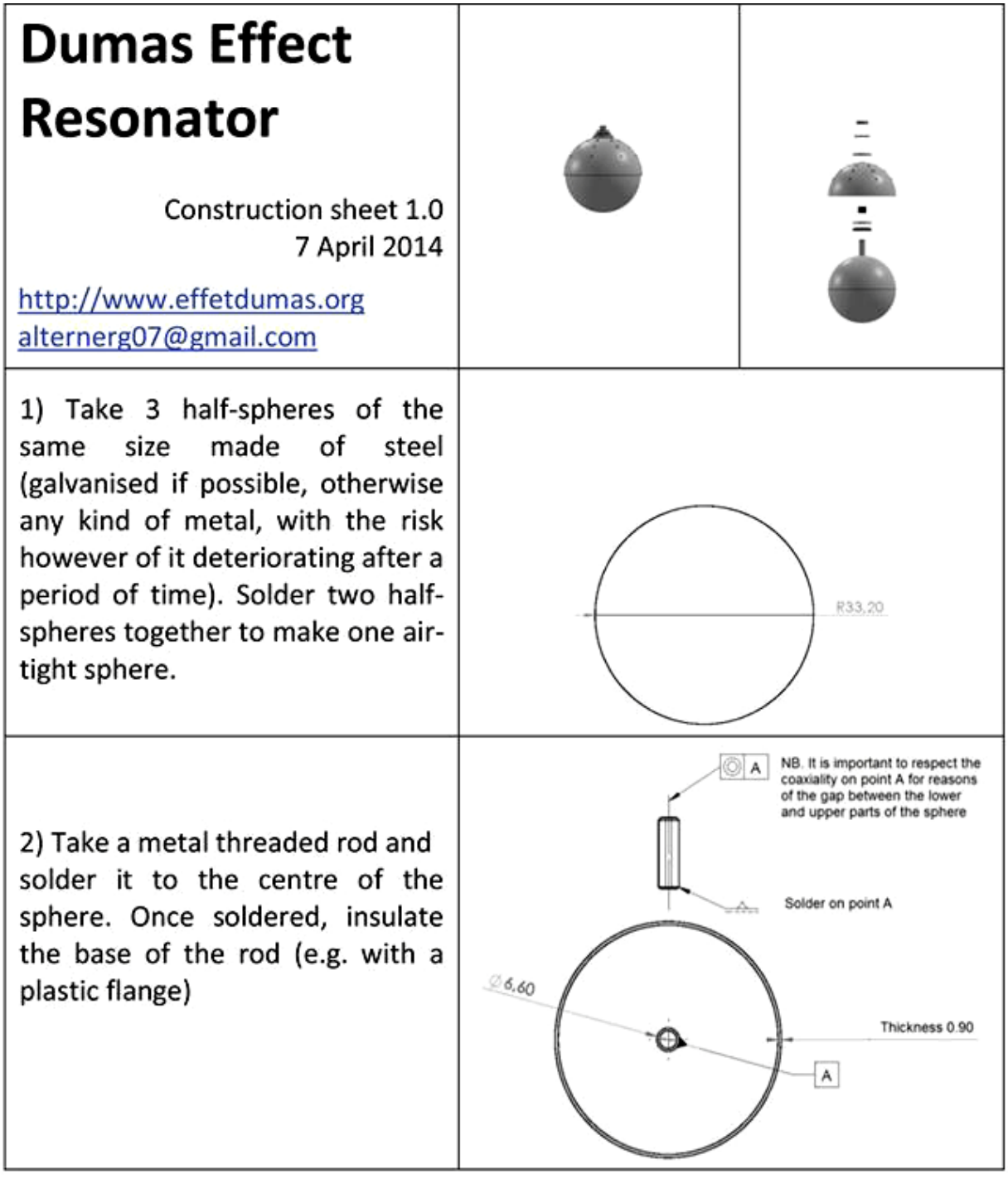
The original prototype which Peter made was constructed from the tops of two bicycle bells, only one of which was tuned to 50 Hz. This shows that the device will definitely work if the inner hemisphere is tuned correctly. You can find forum investigation at

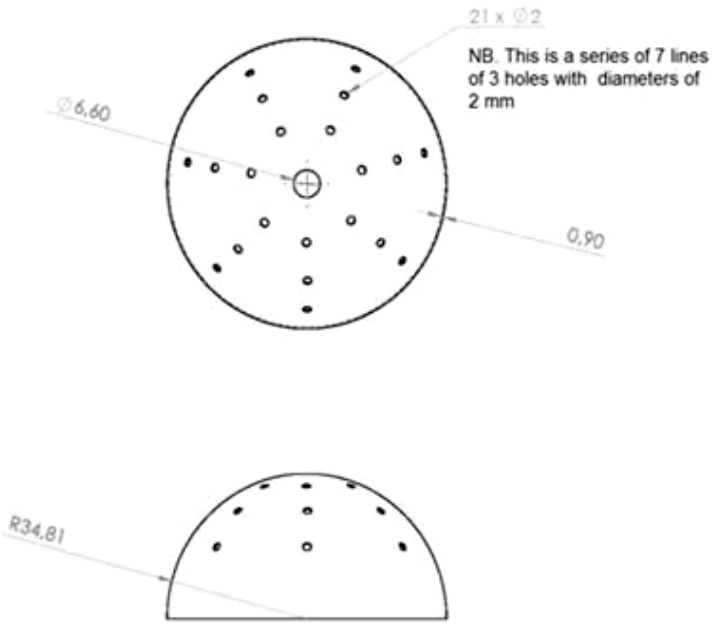
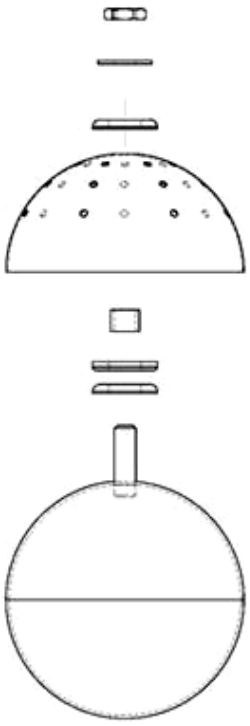
http://www.overunity.com/index.php?topic=4083.msg86151;topicseen and more recent information at http://merlib.org/node/5504.

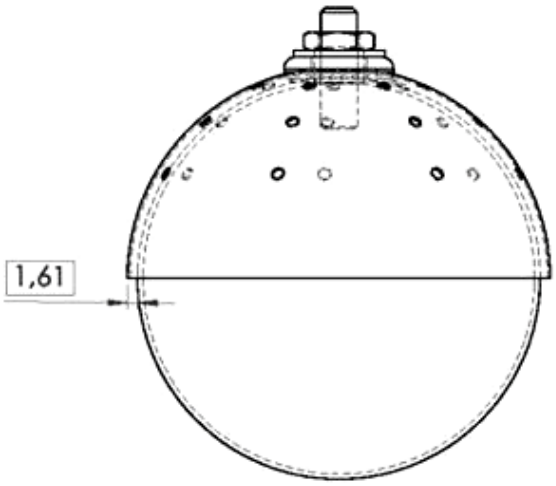
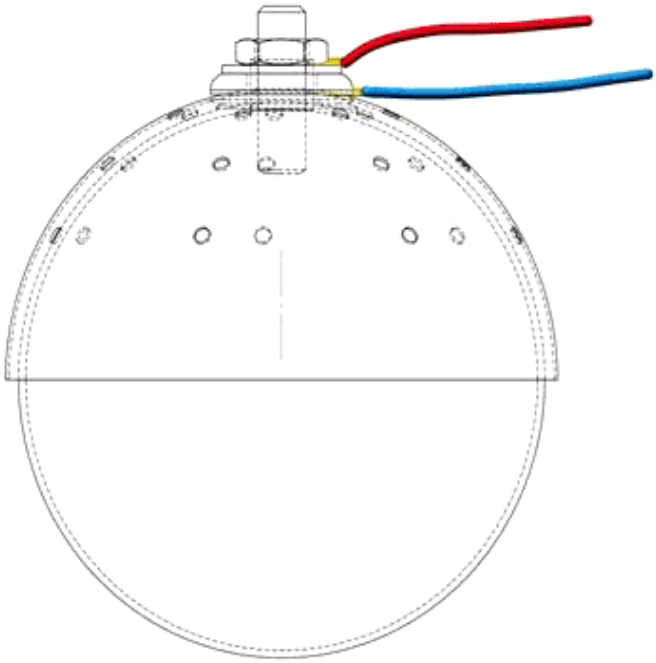
Jean-Christophe Dumas

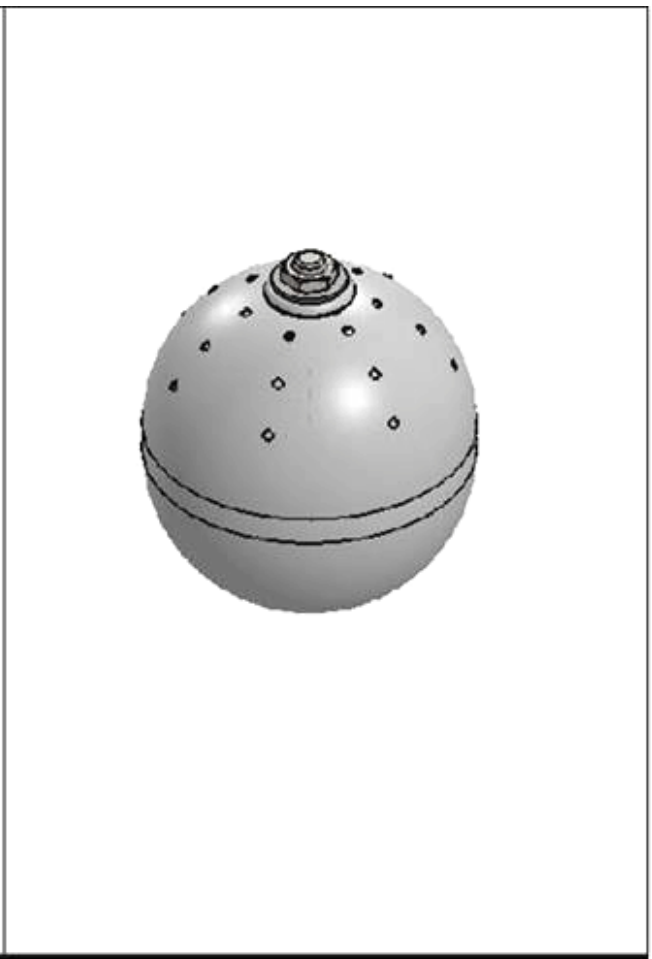
A Frenchman, Jean-Christophe Dumas, publicises his COP=1.17 heater design which appears to operate in a very similar way to that of Peter Davey. His website at http://leblogdejc.com/effetdumas/?page_id=165 gives an English translation of much of what he has to say.

In this design, steam is produced almost immediately after some 500 watts of input power is applied. He uses a metal sphere plus one metal hemisphere which has an adjustable spacing over half of the sphere's surface. Jean-Christophe wishes to share his design freely with everybody and he sees it as being used for efficient space heating in houses, given a radiator system. His drawings are:



3) Drill a hole in the centre of the remaining half-sphere the size of the rod (possibly increased to fit the thickness of the insulation). Drill a maximum of 20 more holes with a diameter of 2 mm, so that the sphere looks like a sieve.	
4) Place 2 x 2 mm plastic washers at the base of the soldered rod on the sphere, drill the half-sphere in the centre and adjust the gap to about 1.6-1.7 mm (For perfect precision, it would be ideal to create the spheres using a 3D printer)	

5) Screw a nut on the rod and adjust the gap between the upper half-sphere and lower part of the sphere to obtain an ideal gap of 1.61 mm (golden section). If the gap is much less (0.6 to 0.9 mm) there will be a short-circuit. If it is more than 1.8 to 1.9 mm, it will not produce any effet.	
6) Connect an electrical live wire (16A) to the top of the rod soldered on the sphere. 7) As shown in the drawing, connect the 2nd neutral wire (16A) to obtain direct contact on the upper half-sphere by means of a ring or spade connector. 8) Attach a switch to the wires and join the electrical wires to an ordinary electric switch made safe with a residual circuit breaker.	

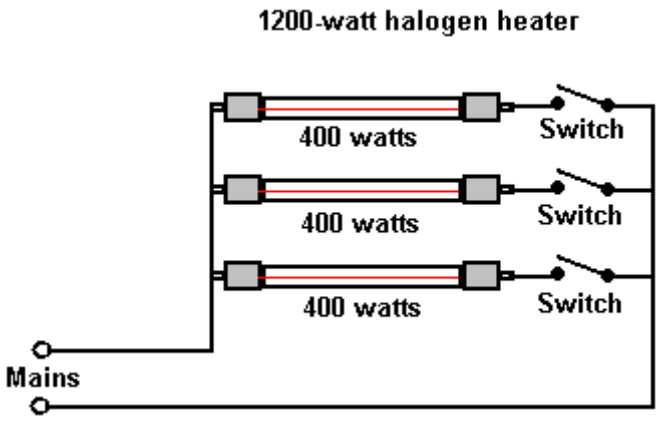


WARNING: The assembly of this process is for scientific experiments and offers the possibility of numerous appliations: water heater, domestic heating, production of drinking water, production of electricity by energy surplus, etc.) It must be done by a professional or a handyman experienced in handling electrical appliances. This process shows several physical effects that are still largely unknown. The inventor declines all responsibility for misuse of this application. NEVER TOUCH THE APPLIANCE WHILE IT IS WORKING, OR THE WATER, OR THE JAR (risk of burns and/or electrocution). Never carry out the experiment alone. Make sure you have a secure environment in case of short circuits, as well as protective gear (masks, gloves). There is a risk being splashed by boiling water or steam, or possible release of flammable gases (oxygen, hydrogen). JC Dumas

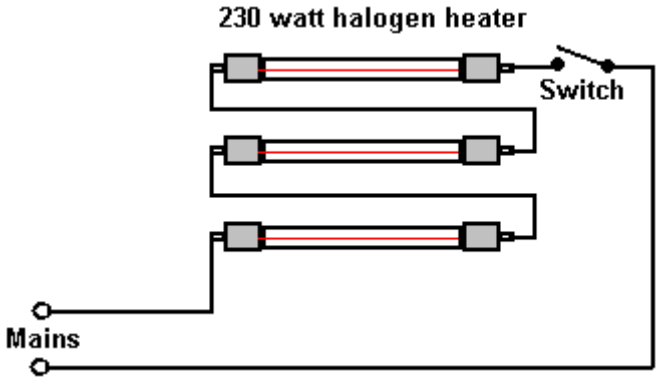
4

The Series-connected Heater.

While not a free-energy device, one simple arrangement which I use myself, is an adapted halogen heater. A standard, low-cost halogen heater consists of three separate 400-watt sections with a switching arrangement which allows one, two or three sections to be powered up:



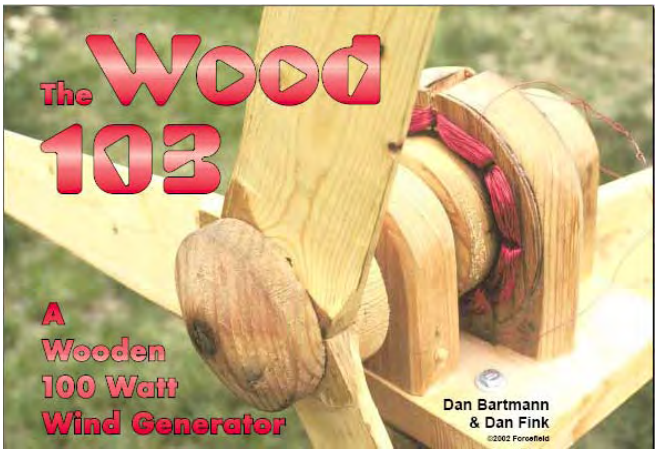
I changed the connections inside my particular heater, so that all three halogen lamps are connected in a chain. This did not involve cutting any wires or making any new connections as the wires connecting to the lamps have push-on 'spade' connectors to allow for both simple manufacturing and easy replacement of a halogen lamp. The new arrangement is like this:



This arrangement 'under-runs' the lamps as each lamp only gets one third of the voltage which it was designed for. This has the effect of increasing the working life of the lamp enormously. You would expect the heat output to be very poor, and perhaps it is. But it gives the impression of being quite effective and with all of the three sections working, it provides a gentle heat and light which seems very effective in keeping a room warm.

The Home Power, Home-Build Wind-Powered Electrical Generator.

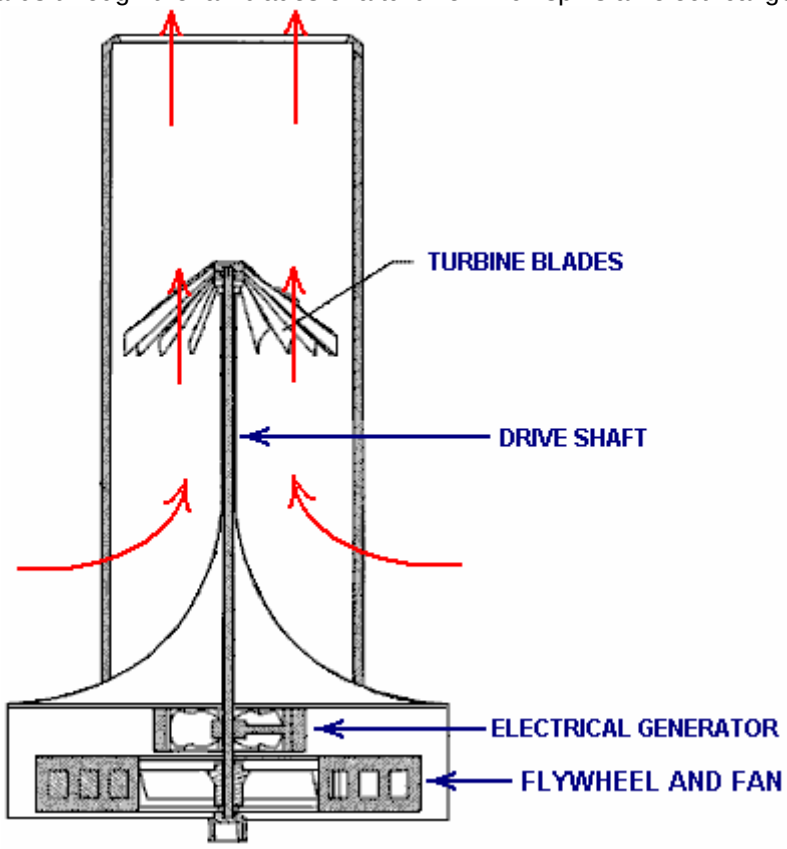
Here is an interesting article from the Home Power web site. If you are interested in renewable power, then I strongly recommend that you visit their web site http://www.homepower.com and consider subscribing to their magazine as they cover many practical topics using simple wording. Here is an example of the high quality material from Home Power – a 100 watt wind generator built from scratch in just one day:



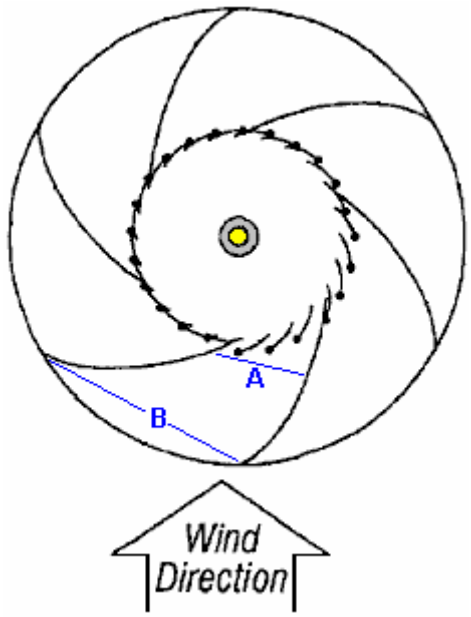
The link for this is: http://homepower.com/article/?file=HP88_pg22_Bartmann.

William McDavid's Wind-Powered Generator.

William McDavid junior remarks that a horizontal axis windmill of that type creates an area of slow-moving air behind the blades and that restricts the flow of air past the blades. A way to overcome that is to project the outgoing air in a direction which does not impede the incoming air. He shows how this can be done in his patent US Patent 6,800,955 of 5th October 2004. In this design, the wind blows into the generator housing and is deflected upwards through the fan blades of a turbine which spins an electrical generator:



One clever feature common to both of these designs is the use of a stationary circular housing with deflectors which uses the wind flow no matter what the wind direction happens to be at any given moment. Looking down from above, the housing looks like this:



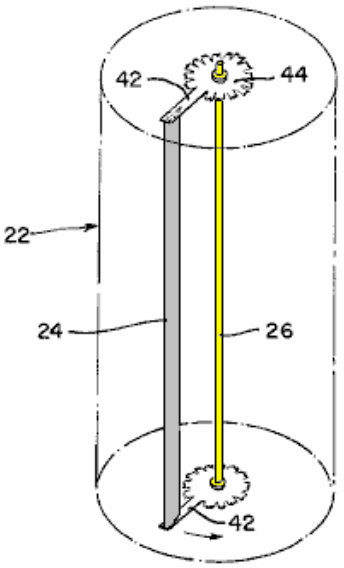
This view shows two important features which enhance the performance of the device. The first is that hinged flaps allow the (horizontal) inflow of air but block the air from flowing straight out of the other side of the central section. This forces the wind to turn and flow upwards, and not only that, but this arrangement causes the air to spin, creating a miniature tornado vortex which amplifies the power of the wind as can be seen from the devastation caused by full-size tornadoes in the environment. As can be seen from the upper diagram, an upward-curving conical piece on the base of the housing assists the airflow to turn upwards as it spins. The spinning air helps to spin the generator blades faster, giving additional power.

A major additional feature is the fact that the dimension "**A**" is considerably less than the dimension "**B**" due to the reduced diameter of the housing nearer the centre. This means that the air flowing past the vanes of housing gets squeezed into a smaller space as it flows. This forces the air to speed up, causing the flow inside the central housing to be higher than the wind-speed outside and that boosts the performance of the device. This wind-powered generator looks like a straightforward project for home construction and with the air being deflected vertically, there does not appear to be any reason why several should not be located near each other. William's full patent can be seen in the appendix of this eBook.

Frank Herbert's Wind-Powered Generator.

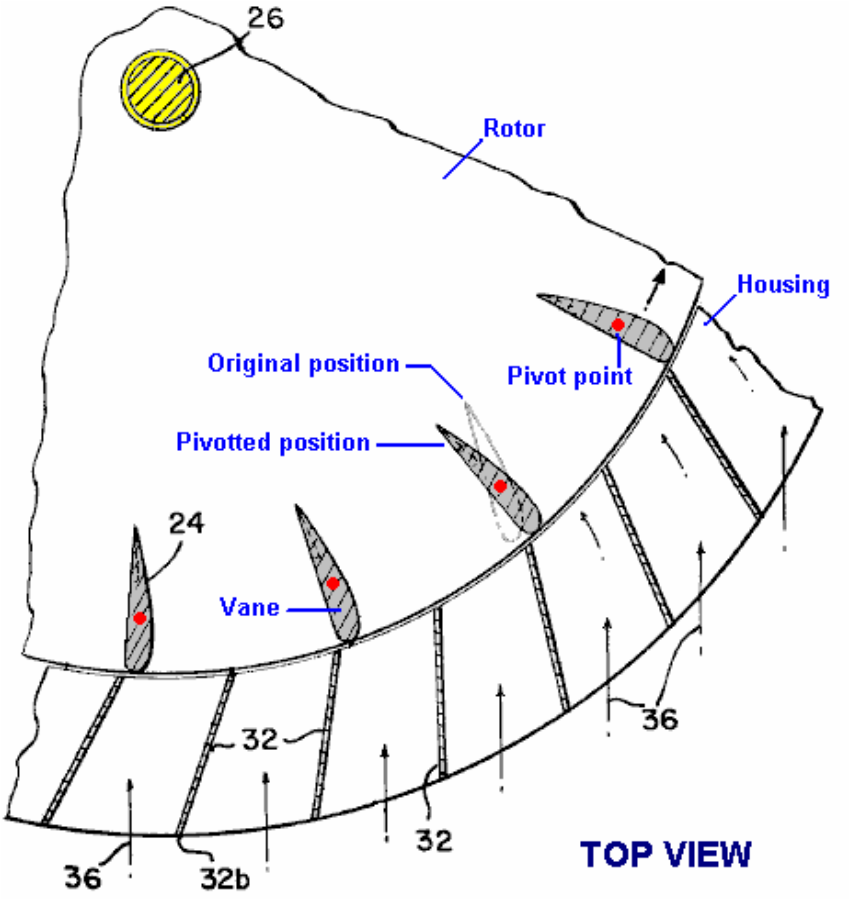
As has been carefully explained by the above article, if a windmill of the blade variety is mounted low down then it is dangerous, and people on sailing boats have been killed by them. Also, if the blade arrangement is designed to operate well in low wind conditions, then it is not unusual for there to be a problem if the wind rises to gale force or higher, with some generator designs giving up and switching off entirely, even though the available free energy is at its highest level.

This design by Frank Herbert is perfectly capable of being home-built and yet it overcomes these problems as well as being a high-efficiency wind turbine. It has an outside cage which prevents human access to the moving parts inside and the 'cage' is not just for protection but is there to enhance the performance of the device. In passing, windmills can be used to compress air and compressed air cylinders can be used to power vehicles and/or power electrical generators during periods of heavy power requirements. The following information is from Frank Herbert's US Patent 4,142,822 of 1979:



The vertical housing **22** shown dotted here, surrounds the vertical power take-off shaft **26**. The wind is allowed to flow through this housing at any angle, so there is no need for the housing to move. In the diagram above small discs **44** are shown at each end of the vertical shaft. These discs have arms **42** extending outwards to support a series of vertical vanes or pressure surfaces **24**. For clarity, just one vane is shown through there will actually be many of these (rather like the cutting blades on a cylinder lawnmower). In reality, there will be no arms on the discs **42** as it is much easier just to have a full-width solid disc supporting the vanes.

The outer housing has a series of vertical slats which are angled to direct the incoming wind on to the vanes at the best possible angle:

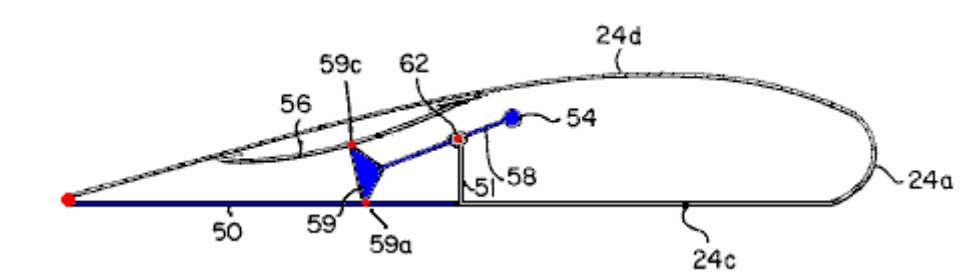


This top view of part of the device, shows the main mounting shaft **26** on which the top and bottom rotor discs are mounted. The red dots show the pivot points where the vanes **24** can turn to take the greatest advantage of the wind pressure. The incoming wind **36**, is deflected by the slats of the housing **32**, to give it

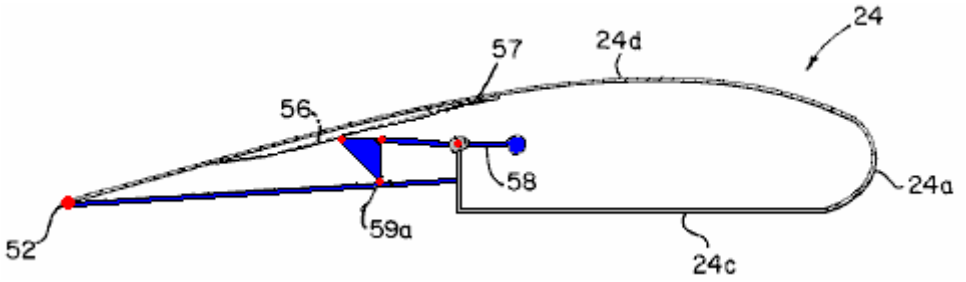
a good angle when flowing through the device as well as keeping humans away from the spinning mechanism. As the vanes and slats are located all the way around shaft **26**, sudden changes in wind direction and/or wind strength have no particular effect on this design as it operates with wind coming from any direction and no physical movement of any part of the device is needed for a change in wind direction.

The vanes can have various different profiles and still work well. The shape shown above is the shape of an aircraft wing, where a force acting towards the curved surface is generated when air flows around the shape. This is not a particularly difficult shape to construct and it is very effective in an airflow (which is why it is used to lift aircraft off the ground). There can be any convenient number of vanes and a device built as shown above should be very effective..

As the overall efficiency is improved if there is no turbulence inside the device, Frank has found a method of minimising this. For this, he uses a mechanism which can alter the shape of the vanes when the windspeed gets high. The higher windspeed whirls the vanes around faster, causing higher 'centrifugal' forces on the vanes which Frank uses as follows. Weight **54** gets pushed across by the spin rate of the rotor.



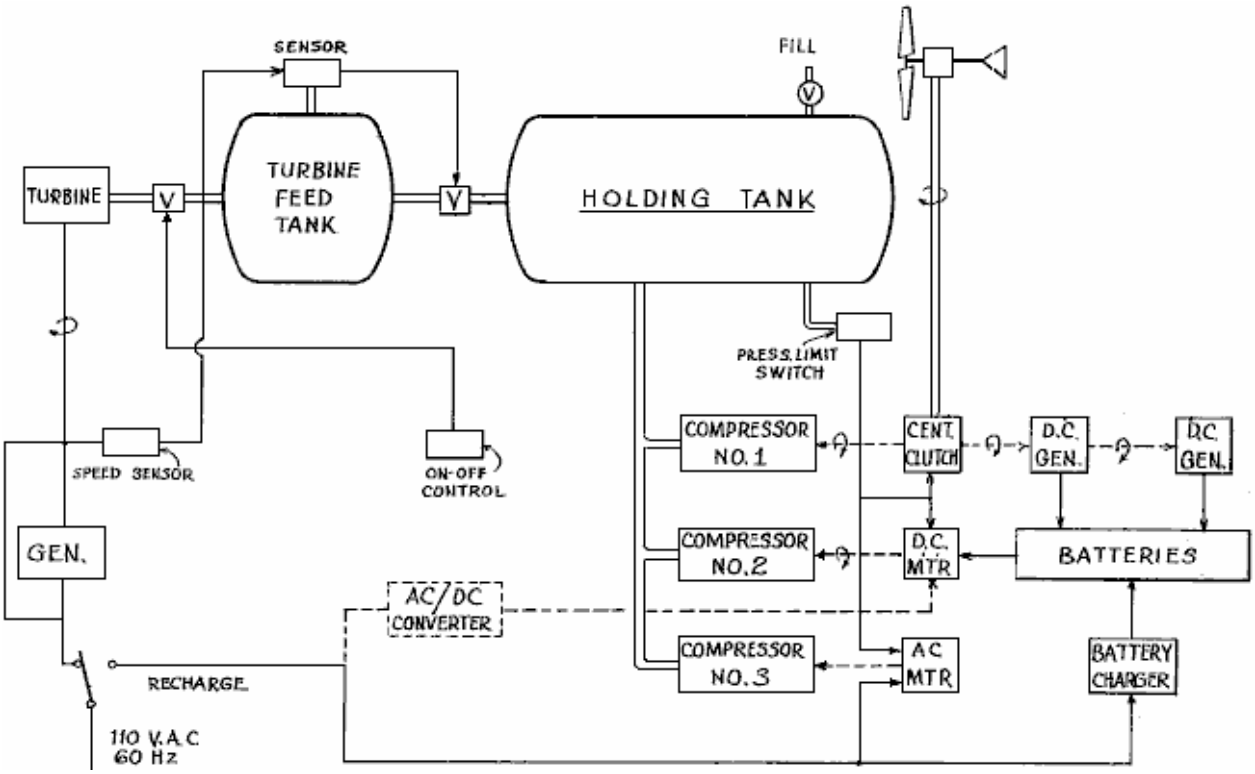
This pushes against the spring **56**, compressing it. The triangle link **59** moves upwards, pivoting at points **59a** and **59c**, and raising section **50** of the vane. This changes the shape of the vane as shown here:



The result of this changed shape is to reduce turbulence inside the device and raise the overall efficiency.

The Mead and Holmes Power System.

The US patent 4,229,661 dated 1980 from Claude Mead and William Holmes is entitled "Power Plant for Camping Trailer" proposes the use of a wind power generator to store compressed air for later use in providing household electrical current, and simultaneously charge batteries which can be used to drive the compressor in periods of very high electrical demand. There is also an option for a rapid system charge if AC mains power becomes available:



****

Solar Ovens.
This information comes from http://solarcooking.org/plans/funnel.htm and ownership remains with the original authors and the material is reproduced here with their kind permission.

The Solar Funnel Cooker

How to Make and Use The Brigham Young University Solar Cooker/Cooler

by Professor of Physics at Brigham Young University (BYU), with Colter Paulson, Jason Chesley, Jacob Fugal, Derek Hullinger, Jamie Winterton, Jeannette Lawler, and Seth, David, Nathan, and Danelle Jones.

Introduction

A few years ago, I woke up to the fact that half of the people in the world must burn wood or dried dung in order to cook their food. It came as quite a shock to me, especially as I learned of the illnesses caused by breathing smoke day in and day out, and the environmental impacts of deforestation - not to mention the time spent by people (mostly women) gathering sticks and dung to cook their food. And yet, many of these billions of people live near the equator, where sunshine is abundant and free. So.....

As a University Professor of Physics with a background in energy usage, I set out to develop a means of cooking food and sterilising water using the energy freely available from the sun. First, I looked at

existing methods.

The parabolic cooker involves a reflective dish which concentrates sunlight to a point where the food is cooked. This approach is very dangerous since the sun's energy is focused to a point which is very hot, but which cannot be seen. (Brigham Young University students and I built one which will set paper on fire in about 3 seconds!). I learned that an altruistic group had offered reflecting parabolas to the people living at the Altiplano in Bolivia. But more than once these parabolas had been stored next to a shed -- and the passing sun set the sheds on fire! The people did not want these dangerous, expensive devices, even though the Altiplano region has been stripped of fuel wood.

The box cooker: Is basically an insulated box with a glass or plastic lid, often with a reflecting lid to direct sunlight into the box. Light enters through the top glass (or plastic), to slowly heat up the box. The problems with this design are that energy enters only through the top, while heat is escaping through all of the other sides, which have a tendency to draw heat away from the food. When the box is opened to put food in or take it out, some of the heat escapes and is lost. Also, effective box cookers tend to be more complicated to build than the funnel cooker.

While studying this problem, I thought again and again of the great need for a safe, inexpensive yet effective solar cooker. It finally came to me at Christmastime a few years ago, a sort of hybrid between the parabola and the box cooker. It looks like a large, deep funnel, and incorporates what I believe are the best features of both the parabolic cooker and the box cooker.

The first reflector was made at my home out of aluminium foil glued on to cardboard, then this was curved to form a reflective funnel. My children and I figured out a way to make a large cardboard funnel easily. (I'll tell you exactly how to do this later on.)

The Solar Funnel Cooker is safe and low cost, easy to make, yet very effective in capturing the sun's energy for cooking and pasteurising water -> Eureka!

Later, I did extensive tests with students (including reflectivity tests) and found that aluminised Mylar was good too, but relatively expensive and rather hard to come by in large sheets. Besides, cardboard is found throughout the world and is inexpensive, and aluminium foil is also easy to come by. Also, individuals can make their own solar cookers easily, or start a cottage-industry to manufacture them for others.

Prototypes of the Solar Funnel Cooker were tested in Bolivia, and outperformed an expensive solar box cooker and a “Solar Coolkit” while costing much less then either. Brigham Young University submitted a patent application, mainly to insure that no company would prevent wide distribution of the Solar Funnel Cooker. Brigham Young University makes no profit from the invention. (I later learned that a few people had had a similar idea, but with methods differing from those developed and shown here). So now I'm trying to get the word out so that the invention can be used to capture the free energy coming from the sun - for camping and for emergencies, yes, but also for every day cooking where electricity is not available and where even fuel wood is getting scarce.

How it Works

The reflector is shaped like a giant funnel, and lined with aluminium foil. (Easy to follow instructions will be given soon). This funnel is rather like the parabolic cooker, except that the sunlight is concentrated along a line (not a point) at the bottom of the funnel. You can put your hand up the bottom of the funnel and feel the sun's heat, but it will not burn you.

Next, we paint a jar black on the outside, to collect heat, and place this at the bottom of the funnel. Or a black pot with a lid can be used. The black vessel gets hot, quickly, but not quite hot enough to cook with. We need some way to build up the heat without letting the outside air cool it. So, I put a cheap plastic bag around the jar -- and, the solar funnel cooker was born! The plastic bag, available in grocery stores as a "poultry bag", replaces the cumbersome and expensive box and glass lid of solar box ovens. You can use the plastic bags used in American stores to put groceries in, as long as they let a lot of sunlight pass. (Dark- coloured bags will not do).

I recently tested a bag used for fruits and vegetables, nearly transparent and available free at American grocery stores, that works great. This is stamped "HDPE" for high-density polyethylene on the bag (ordinary polyethylene melts too easily). A block of wood is placed under the jar to help hold the heat in. (Any insulator, such as a hot pad or rope or even sticks, will also work).

A friend of mine who is also a Physics Professor did not believe I could actually boil water with the thing. So I showed him that with this new "solar funnel cooker" I was able to boil water in Utah in the middle of winter! I laid the funnel on its side since it was winter and pointed a large funnel towards the sun to the south. I also had to suspend the black cooking vessel -- rather than placing it on a wooden block. This allows the weaker sun rays to strike the entire surface of the vessel.

Of course, the Solar Funnel works much better outside of winter days, that is, when the UV index is 7 or greater. Most other solar cookers will not cook in the winter in northern areas (or south of about 35 degrees, either).

I thought that a pressure cooker would be great. But the prices in stores were way too high for me. Wait, how about a canning jar? These little beauties are designed to relieve pressure through the lid -- a nice pressure cooker. And cooking time is cut in half for each 10ºC we raise the temperature (Professor Lee Hansen, private communication). I used one of my wife's wide-mouth canning jars, spray-painted (flat) black on the outside, and it worked great. Food cooks faster when you use a simple canning jar as a pressure cooker. However, you can also put a black pot in the plastic bag instead if you want. But don't use a sealed container with no pressure release like a mayonnaise jar -- it can break as the steam builds up (I've done it)!

How to Build Your Own Solar Funnel Cooker

What You will Need for the Funnel Cooker:

A piece of flat cardboard, about 2 feet wide by 4 feet long. (The length should be just twice the width. The bigger, the better).

Ordinary aluminium foil.

A glue such as white glue (like Elmer's glue), and water to mix with it 50-50. Also, a brush to apply the glue to the cardboard (or a cloth or paper towel will do). Or, some may wish to use a cheap "spray adhesive" available in spray cans. You can also use flour paste.

Three wire brads - or small nuts and bolts, or string to hold the funnel together.

For a cooking vessel, I recommend a canning jar ("Ball" wide-mouth quart jars work fine for me; the rubber ring on the lid is less likely to melt than for other jars I've found. A two-quart canning jar is available and works fine for larger quantities of food, although the cooking is somewhat slower).

The cooking jar (or vessel) should be spray-painted black on the outside. I find that a cheap flat-black spray paint works just fine. Scrape off a vertical stripe so that you have a clear glass "window" to look into the vessel, to check the food or water for boiling.

A block of wood is used as an insulator under the jar. I use a piece of 2" x 4" board which is cut into a square nominally 4" x 4" by about 2" thick. (100 mm square x 50 mm thick). One square piece of wood makes a great insulator.

A plastic bag is used to go around the cooking-jar and block of wood, to provide a green-house effect. Suggestions:

- Reynolds™ Oven Bag, Regular Size works great: transparent and won't melt. (Cost about 25 cents each in U.S. grocery stores.)

- Any nearly-transparent HDPE bag (High-density Polyethylene). Look for "HDPE" stamped on the bag. I've tested HDPE bags which I picked up for free at my grocery store, used for holding vegetables and fruits. These are thin, but very inexpensive. Tested side-by-side with an oven bag in two solar funnels, the HDPE bag worked just as well! **Caution:** we have found that some HDPE bags will melt should they contact the hot cooking vessel. For this reason, we recommend using the oven-safe plastic bag wherever possible.

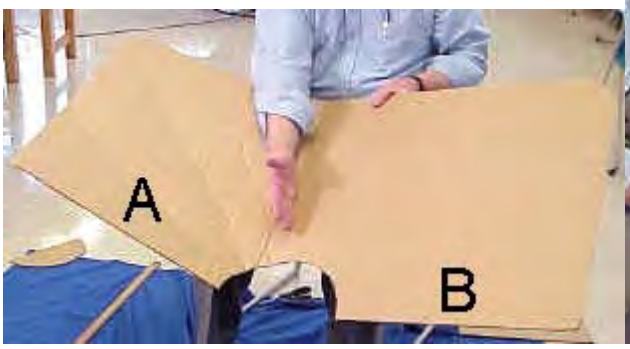
- An idea attributed to Roger Bernard and applied now to the BYU Funnel Cooker: place a pot (having a blackened bottom and sides) in a glass bowl, and cover with a lid. Try for a tight fit around the bottom to keep hot air trapped inside. The metal pot or bowl should be supported around the rim only, with an air space all around the bottom (where the sunlight strikes it). Put a blackened lid on top of the pot. Then simply place this pot-in-bowl down in the bottom of the funnel - no plastic bag is needed! This clever method also allows the cook to simply remove the lid to check the food and to stir. I like this idea - it makes the solar cooker a lot like cooking over a fire. See Photographs for further details.

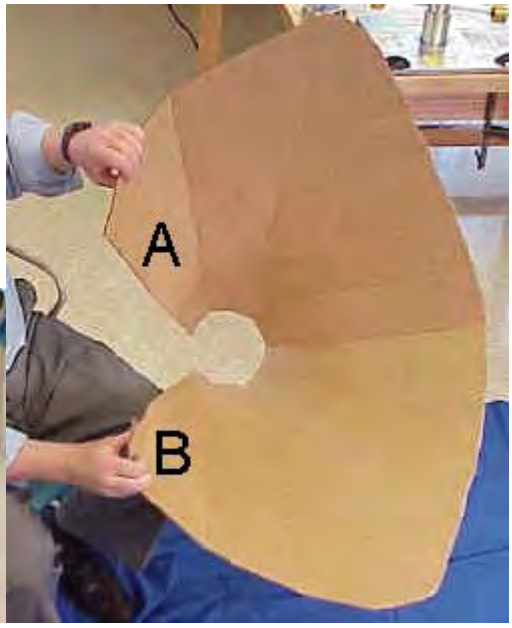
Construction Steps

Cut a Half-circle out of the Cardboard

Cut a half circle out of the cardboard, along the bottom as shown below. When the funnel is formed, this becomes a full-circle and should be wide enough to go around your cooking pot. So for a 7" diameter cooking pot, the radius of the half-circle is 7". For a quart canning jar such as I use, I cut a 5" radius half-circle out of the cardboard.

Form the Funnel





To form the funnel, you will bring side A towards side B, as shown in the figure. The aluminium foil must go on the INSIDE of the funnel. Do this slowly, helping the cardboard to the shape of a funnel by using one hand to form creases that radiate out from the half-circle. Work your way around the funnel, bending it in stages to form the funnel shape, until the two sides overlap and the half-circle forms a complete circle. The aluminium foil will go on the INSIDE of funnel. Open the funnel and lay it flat, "inside up", in preparation for the next step.

Glue Foil to Cardboard

Apply glue or adhesive to the top (inner) surface of the cardboard, then quickly apply the aluminium foil on top of the glue, to affix the foil to the cardboard. Make sure the shiniest side of the foil is on top, since this becomes your reflective surface in the Funnel. I like to put just enough glue for one width of foil, so that the glue stays moist while the foil is applied. I also overlap strips of foil by about 1" (or 2 cm). Try to smooth out the aluminium foil as much as you reasonably can, but small wrinkles won't make much difference. If cardboard is not available, one can simply dig a funnel-shaped hole in the ground and line it with a reflector, to make a fixed solar cooker for use at mid-day.

Join side A to side B to keep the funnel together.

The easiest way to do this is to punch three holes in the cardboard that line up on side A and side B (see figure). Then put a metal brad through each hole and fasten by pulling apart the metal tines. Or you can use a nut-and-bolt to secure the two sides (A & B) together.

Be creative here with what you have available. For example, by putting two holes about a thumb-width apart, you can put a string, twine, small rope, wire or twist-tie in one hole and out the other, and tie together.

When A and B are connected together, you will have a "funnel with two wings". The wings could be cut off, but these help to gather more sunlight, so I leave them on.

Tape or glue a piece of aluminium foil across the hole at the bottom of the funnel, with shiny side in.

This completes assembly of your solar funnel cooker.

For stability, place the funnel inside a cardboard or other box to provide support. For long-term applications, one may wish to dig a hole in the ground to hold the Funnel against strong winds.

Final Steps

At this stage, you are ready to put food items or water into the cooking vessel or jar, and put the lid on securely. (See instructions on food cooking times, to follow).

Place a wooden block in the INSIDE bottom of the cooking bag. I use a piece of 2” x 4” board which is cut into a square nominally 4" x 4" by 2" thick. Then place the cooking vessel containing the food or water on top of the wooden block, inside the bag.

Next, gather the top of the bag in your fingers and ***blow air into the bag, to inflate it***. This will form a small "greenhouse" around the cooking vessel, to trap much of the heat inside. Close off the bag with a tight twist tie or wire. Important: the bag should not touch the sides or lid of the cooking vessel. The bag may be called a "convection shield," slowing convection-cooling due to air currents.

Place the entire bag and its contents inside the funnel near the bottom as shown in the Photographs.

Place the Solar Funnel Cooker so that it Faces the Sun

Remember: Sunlight can hurt the eyes: so please wear sunglasses when using a Solar Cooker! The Funnel Cooker is designed so that the hot region is deep down inside the funnel, out of harm's way.

Put the Solar Funnel Cooker in the sun pointing towards the sun, so that it captures as much sunlight as possible. The design of the funnel allows it to collect solar energy for about an hour without needing to be re-positioned. For longer cooking times, readjust the position of the funnel to follow the sun's path.

In the Northern Hemisphere, it helps to put the Solar Funnel Cooker in front of a south-facing wall or window as this reflects additional sunlight into the funnel. A reflective wall is most important in locations farther from the equator and in winter. In the Southern Hemisphere, put the Solar Funnel Cooker in front of a North-facing wall or window to reflect additional sunlight into your cooker.

After Cooking

Remember that the cooking vessel will be very hot: so use cooking pads or gloves when handling it! If you are heating water in a canning jar, you may notice that the water is boiling when the lid is first removed - it gets very hot!

Open the plastic cooking bag by removing the twist-tie. Using gloves or a thick cloth, lift the vessel out of the bag and place it on the ground or table. Carefully open the vessel and check the food, to make sure it has finished cooking. Let the hot food cool before eating.

Helpful Hints

- Avoid leaving fingerprints and smudges on the inside surface of the cooker. Keep the inner surface clean and shiny by wiping occasionally with a wet towel. This will keep the Solar Funnel Cooker working at its best.
- If your funnel gets out-of-round, it can be put back into a circular shape by attaching a rope or string between opposite sides which need to be brought closer together.
- For long-term applications, a hole in the ground will hold the Funnel Cooker securely against winds. Bring the funnel inside or cover it during rain storms.
- The lids can be used over and over. We have had some trouble with the rubber on some new canning-jar lids becoming soft and "sticky." "Ball canning lids" do not usually have this problem. Running new lids through very hot water before the first use seems to help. The lids can be used over and over if they are not bent too badly when opened (pry off lid carefully).
- The jar can be suspended near the bottom of the funnel using fishing line or string (etc.), instead of placing the jar on a block of wood. A plastic bag is placed around the jar with air puffed inside, as

usual, to trap the heat. The suspension method allows sunlight to strike all surfaces of the jar, all around, so that heats faster and more evenly. This suspension method is crucial for use in winter months.

Adjust the funnel to put as much sunlight onto the cooking jar as possible. Look at the jar to check where the sunlight is hitting, and to be sure the bottom is not in the shadows. For long cooking times (over about an hour), readjust the position of the funnel to follow the sun's path. During winter months, when the sun is low on the horizon (e.g., in North America), it is helpful to lay the funnel on its side, facing the sun.

Tests in Utah

I have personally used the Solar Funnel Cooker to cook lunches over many weeks. My favourite foods to cook are potatoes (cut into logs or slices) and carrot slices. Vegetables cook slowly in their own juices and taste delicious. I also make rice, melted cheese sandwiches, and even bread in the Solar Funnel Cooker. I usually put the food out around 11:30 and let it cook until 12:45 or 1 pm, just to be sure that it has time to cook. I've never had any food burn in this cooker.

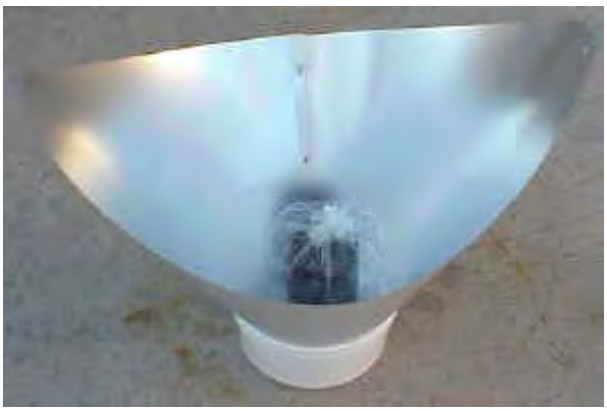
I have also cooked food in the mountains, at an altitude of around 8,300 feet. If anything, the food cooked faster there - the sunlight passes through less atmosphere at high altitudes.

I find that people are surprised that the sun alone can actually cook food. And they are further pleasantly surprised at the rich flavours in the foods which cook slowly in the sun. This inexpensive device does it!

Students at Brigham Young University have performed numerous tests on the Solar Funnel Cooker along with other cookers. We have consistently found much faster cooking using the Solar Funnel Cooker. The efficiency/cost ratio is higher than any other solar cooking device we have found to date. Mr. Hullinger also performed studies of transmissivity, reflectivity and absorptivity of alternate materials which could be used in the Solar Funnel Cooker. While there are better materials, such as solar-selective absorbers, our goal has been to keep the cost of the Solar Cooker as low as possible, while maintaining safety as a first priority.

Tests in Bolivia

The BYU Benson Institute organised tests between the Solar Funnel Cooker and the "old-fashioned" solar box oven. The solar box oven cost about $70 and was made mostly of cardboard. It took nearly two hours just to reach water pasteurisation temperature. The Bolivian report notes that "food gets cold every time the pots are taken from and into the oven." The solar box oven failed even to cook boiled eggs. (More expensive box cookers would hopefully work better.)



An aluminised-mylar Solar Funnel Cooker was also tested in Bolivia, during the Bolivian winter. Water pasteurisation temperature was reached in 50 minutes, boiled eggs cooked in 70 minutes, and rice cooked in 75 minutes. The Bolivian people were pleased by the performance. So were we! (La Paz, Bolivia, August, 1996).

I also donated two dozen solar funnel cookers for people in Guatemala. These were taken there by a group of doctors going there for humanitarian service. The people there also liked the idea of cooking with the sun's free energy. For an aluminised-Mylar Solar Funnel Cooker kit, please contact CRM (licensed manufacturer) at +1 (801) 292-9210.

Water and Milk Pasteurisation

Contaminated drinking water or milk kills thousands of people each day, especially children. The World Health Organisation reports that 80% of illnesses in the world are spread through contaminated water. Studies show that heating water to about 65º - 70º C (150º F) is sufficient to kill coliform bacteria, rotaviruses, enteroviruses and even Giardia. This is called pasteurisation.

Pasteurisation depends on how hot and how long water is heated. But how do you know if the water got hot enough? You could use a thermometer, but this would add to the cost, of course. When steam leaves the canning jar (with lid on tight) and forms "dew" on the inside of the cooking bag, then the water is probably pasteurised to drink. (The goal is to heat to 160º Fahrenheit for at least six minutes.) With a stripe of black paint scraped off the jar, one can look through the bag and into the jar and see when the water is boiling - then it is safe for sure.

Think of all the lives that can be saved simply by pasteurising water using a simple Solar Cooker!

Safety

Safety was my first concern in designing the Solar Funnel Cooker, then came low cost and effectiveness. But any time you have heat you need to take some precautions.

- The cooking vessel (jar) is going to get hot, otherwise the food inside it won't cook. Let the jar cool a bit before opening. Handle only with gloves or tongs.
- Always wear dark glasses to protect from the sun's rays. We naturally squint, but sunglasses are important.
- Keep the plastic bag away from children and away from nose and mouth to avoid any possibility of suffocation.

Cooking with the Solar Funnel Cooker

What do you cook in a crock pot or moderate-temperature oven? The same foods will cook about the same in the Solar Funnel Cooker - without burning. The charts below give approximate summer cooking times.

The solar cooker works best when the UV index is 7 or higher (Sun high overhead, few clouds).

Cooking times are approximate. Increase cooking times for partly-cloudy days, sun not overhead (e.g., wintertime) or for more than about 3 cups of food in the cooking jar.

Stirring is not necessary for most foods. Food generally will not burn in the solar cooker.

Vegetables (Potatoes, carrots, squash, beets, asparagus, etc.)
Preparation: No need to add water if fresh. Cut into slices or "logs" to ensure uniform cooking. Corn will cook fine with or without the cob.
Cooking Time: About 1.5 hours

Cereals and Grains (Rice, wheat, barley, oats, millet, etc.)
Preparation: Mix 2 parts water to every 1 part grain. Amount may vary according to individual taste. Let soak for a few hours for faster cooking. To ensure uniform cooking, shake jar after 50 minutes.
CAUTION: Jar will be hot. Use gloves or cooking pads.
Cooking Time: 1.5-2 hours

Pasta and Dehydrated Soups
Preparation: First heat water to near boiling (50-70 minutes). Then add the pasta or soup mix. Stir or shake, and cook 15 additional minutes.
Cooking Time: 65-85 minutes

Beans
Preparation: Let tough or dry beans soak overnight. Place in cooking jar with water.
Cooking Time: 2-3 hours

Eggs
Preparation: No need to add water. **Note**: If cooked too long, egg whites may darken, but taste remains the same.
Cooking Time: 1-1.5 hours, depending on desired yolk firmness.

Meats (Chicken, beef, and fish)
Preparation: No need to add water. Longer cooking makes the meat more tender.
Cooking Time: Chicken: 1.5 hours cut up or 2.5 hours whole; Beef: 1.5 hours cut up or 2.5-3 hours for larger cuts; Fish: 1-1.5 hours

Baking
Preparation: Times vary based on amount of dough.
Cooking Times: Breads: 1-1.5 hours; Biscuits: 1-1.5 hours; Cookies: 1 hour

Roasted Nuts (Peanuts, almonds, pumpkin seed, etc.)
Preparation: Place in jar. A little vegetable oil may be added if desired.
Cooking Time: About 1.5 hours

MRE's and pre-packaged foods
Preparation: For foods in dark containers, simply place the container in the cooking bag in place of the black cooking jar.
Cooking Times: Cooking time varies with the amount of food and darkness of package.

How to Use the Solar Funnel as a Refrigerator/Cooler

A university student (Jamie Winterton) and I were the first to demonstrate that the Brigham Young University Solar Funnel Cooker can be used - at night - as a refrigerator. Here is how this is done:

The Solar Funnel Cooker is set-up just as you would during sun-light hours, with two exceptions:

1. The funnel is directed at the dark night sky. It should not "see" any buildings or even trees. (The thermal radiation from walls, trees, or even clouds will diminish the cooling effect.).
2. It helps to place 2 (two) bags around the jar instead of just one, with air spaces between the bags and between the inner bag and the jar. HDPE and ordinary polyethylene bags work well, since polyethylene is nearly transparent to infrared radiation, allowing it to escape into the "heat sink" of the dark sky.

During the day, the sun's rays are reflected on to the cooking vessel which becomes hot quickly. At night, heat from the vessel is radiated *outward*, towards empty space, which is very cold indeed (a "heat sink"). As a result, the cooking vessel now becomes a small refrigerator. We routinely achieve cooling of about 20º F (10º C) below ambient air temperature using this remarkably simple scheme.

In September 1999, we placed two funnels out in the evening, with double-bagged jars inside. One jar was on a block of wood and the other was suspended in the funnel using fishing line. The temperature that evening (in Provo, Utah) was 78º F (25.5º C). Using a Radio Shack indoor/outdoor thermometer, a BYU student (Colter Paulson) measured the temperature inside the funnel and outside in the open air. He found that the temperature of the air inside the funnel dropped quickly by about 15º F (8º C), as its heat was radiated upwards in the clear sky. That night, the minimum outdoor air temperature measured was 47.5º F (8.6º C) - but the water in both jars had ICE. I invite others to try this, and please let me know if you get ice at 55 or even 60 degrees outside air temperature (minimum at night). A black PVC container may work even better than a black-painted jar, since PVC is a good infrared radiator - these matters are still being studied.

I would like to see the "Funnel Refrigerator" tried in desert climates, especially where freezing temperatures are rarely reached. It should be possible in this way to cheaply make ice for Hutus in Rwanda and for aborigines in Australia, without using any electricity or other modern "tricks." We are in effect bringing some of the cold of space to a little corner on earth. Please let me know how this works for you.

Conclusion: Why We Need Solar Cookers

The BYU Funnel Cooker/Cooler can:

- Cook food without the need for electricity or wood or petroleum or other fuels.
- Pasteurise water for safe drinking, preventing many diseases.
- Save trees and other resources.
- Avoid air pollution and breathing smoke while cooking.
- Use the sun's free energy. A renewable energy source.
- Cook food with little or no stirring, without burning.
- Kill insects in grains.
- Dehydrate fruits, etc.
- Serve as a refrigerator at night, to cool even freeze water.

(Try that without electricity or fuels!)

The burden for gathering the fuel wood and cooking falls mainly on women and children. Joseph Kiai reports :

From Dadaab, Kenya: "Women who can't afford to buy wood start at 4 am to go collecting and return about noon... They do this twice a week to get fuel for cooking... The rapes are averaging one per week."

From Belize: "Many times the women have to go into the forest dragging their small children when they go to look for wood. It is a special hardship for pregnant and nursing mothers to chop and drag trees back to the village... they are exposed to venomous snakes and clouds of mosquitoes."

And the forests are dwindling in many areas. Edwin Dobbs noted in *Audubon Magazine*, Nov. 1992, "The world can choose sunlight or further deforestation, solar cooking or widespread starvation..."

Americans should be prepared for emergencies, incident to power failures. A Mormon pioneer noted in her journal: "We were now following in their trail travelling up the Platte River. Timber was sometimes very scarce and hard to get. We managed to do our cooking with what little we could gather up..." (Eliza R. Snow) Now there's someone who needed a light-weight Solar Cooker!

Here's another reason to use a solar cooker. Many people in developing countries look to see what's being done in America. I'm told that if Americans are using something, then they will want to try it, too. The more people there are cooking with the sun, the more others will want to join in. A good way to

spread this technology is to encourage small local industries or families to make these simple yet reliable solar cookers for others at low cost. I've used this cooker for three summers and I enjoy it. Cooking and making ice with the funnel cooker/cooler will permit a significant change in lifestyle. If you think about it, this could help a lot of people. The BYU Solar Funnel Cooker uses the glorious sunshine -- and the energy of the sun is a free gift from God for all to use!

Answers to commonly-asked questions

Will the cooker work in winter (in the United States)?

As the sun moves closer to the southern horizon in the winter, the solar cooker is naturally less effective. A good measure of the solar intensity is the "UV index" which is often reported with the weather. When the ultraviolet or UV index is 7 or above – common in summer months – the solar cooker works very well. In Salt Lake City in October, the UV index was reported to be 3.5 on a sunny day. We were able to boil water in the Solar Funnel Cooker during this time, but we had to suspend the black jar in the funnel so that sunlight struck all sides. (We ran a fishing line under the screw-on lid, and looped the fishing line over a rod above the funnel. As usual, a plastic bag was placed around the jar, and this was closed at the top to let the fishing line out for suspending the jar.)

The solar "minimum" for the northern hemisphere occurs on winter solstice, about December 21st each year. The solar "maximum" occurs six months later, June 21st. Solar cooking works best from about 20th March to 1st October in the north. If people try to cook with the sun for the first time outside of this time window, they should not be discouraged. Try again when the sun is more directly overhead. One may also suspend the jar in the funnel, which will make cooking faster any time of the year.

It is interesting to note that most developing countries are located near the equator where the sun is nearly directly overhead all the time. Solar Cookers will then serve year-round, as long as the sun is shining, for these fortunate people. They may be the first to apply fusion energy (of the sun) on a large scale. They may also accomplish this without the expensive infrastructure of electrical power grids that we take for granted in America.

How do you cook bread in a jar?

I have cooked bread by simply putting dough in the bottom of the jar and placing it in the funnel in the usual way. Rising and baking took place inside the jar in about an hour (during summer). One should put vegetable oil inside the jar before cooking to make removal of the bread easier. I would also suggest that using a 2-quart wide-mouth canning jar instead of a 1-quart jar would make baking a loaf of bread easier.

What is the optimum "opening angle" for the funnel cooker?

A graduate student at Brigham Young University did a calculus calculation to assess the best shape or opening angle for the Solar Funnel. Jeannette Lawler assumed that the best operation would occur when the sun's rays bounced no more than once before hitting the cooking jar, while keeping the opening angle as large as possible to admit more sunlight. (Some sunlight is lost each time the light reflects from the shiny surface. If the sunlight misses on the first bounce, it can bounce again and again until being absorbed by the black bottle). She set up an approximate equation for this situation, took the calculus derivative with respect to the opening angle and set the derivative equal to zero. Optimising in this way, she found that the optimum opening angle is about 45 degrees, when the funnel is pointed directly towards the sun.

But we don't want to have to "track the sun" by turning the funnel every few minutes. The sun moves (apparently) 360 degrees in 24 hours, or about 15 degrees per hour. So we finally chose a 60-degree opening angle so that the cooker is effective for about 1.2 hours. This turned out to be long enough to cook most vegetables, breads, boil water, etc. with the Solar Funnel Cooker. We also used a laser pointer to simulate sun rays entering the funnel at different angles, and found that the 60-degree cone was quite effective in concentrating the rays at the bottom of the funnel where the cooking jar sits.

For questions regarding the complete Solar Funnel Cooker kit using aluminised Mylar and a jar for the cooking vessel, please contact CRM at +1 (801) 292-9210.

Tests of the Solar Funnel and Bowl Cookers in 2001

Christopher McMillan and Steven E. Jones
Brigham Young University

Introduction

With an increase in population and a decrease in available fuels such as wood and coal in developing countries, the need for alternative cooking methods has increased. Solar cookers are an alternative to conventional methods such as wood-fires and coal-fires. They provide usable heat for cooking and pasteurising water, without the harmful side effects such as smoke inhalation that non-renewable sources create. In many countries such as Haiti, Bolivia and Kenya, the need for cheap, effective, and safe cooking methods has increased due to poverty and deforestation. Solar cookers are ideal because they rely on the sun's free energy which is abundant in many of the world's poorest countries. Though there are good designs, more testing and improvement is desirable.

There are three areas of comparison that were focused on during the course of the study. The first area of comparison is in the reflective material used. The original material is a mirror-finished aluminium Mylar. Due to the mirror finish, the reflection light is very bright and can be difficult to work over when cooking. An alternative material is a matt-finish Mylar. This material diffuses the sunlight and is not as harsh on the eyes as is the mirrored finish.

The second area of concentration is on the method of containing the air that surrounds the cooker so that the cooker is kept from being cooled by convection currents. A common method is to use a clear plastic oven-safe bag around the cooking vessel. However, this method is rather tedious and awkward to use, and such bags are rarely available in developing countries. Another technique is to use a disk or window make out of a clear plastic or glass. This makes the cooker easier to use.

The third main area of focus is in the cooking containers used. The present cooking vessel for the Solar Funnel Cooker is a black-painted canning jar. This method is also tedious and awkward. The canning jars can be hard to clean, and they can break. Design changes are tested that would allow people to use their own cookware. This too would make the cooker more convenient to use.

The fourth area of testing pitted the wooden block support which we have been using for years against a rabbit-wire support. A rabbit-wire cylinder holds the cooking vessel up off the bottom of the cooker, and allows sunlight to strike essentially all surfaces of the cooking vessel, including the bottom.

The effectiveness of these methods is tested and compared both qualitatively and quantitatively. In addition to acquiring temperature-rise versus time data, we also cooked numerous meals in the solar cookers so as to get hands-on experience with cooking. Several students participated in these cooking tests.

Cooker Designs:

Several solar cooker designs were used during these tests. The Solar Funnel Cooker was the main cooker tested. A Solar CooKit and a bowl-shaped variation of the Solar Funnel Cooker were also tested. Most experiments were comparative tests between the various designs, and the cooker set-up was varied from test to test. The basic design of the Solar Funnel Cooker is a funnel-shaped aluminium Mylar collector. A highly reflective material is necessary to collect and concentrate the sun's rays. The funnel walls are at a 60 degree angle (with respect to the horizontal) since this collects sunlight for a two hour time period without requiring re-orientation to follow the sun. Due to the way the Mylar sheets are cut and folded, a pair of wings

on opposite ends of the funnel is formed. The wings increase the collector size and create an elliptical shape at top. At the tips of the wings, the cooker stands about 20 inches high and has a diameter of about 28 inches. At the top, along the minor axis of the elliptical funnel, the cooker stands about 15 inches high, and has a diameter of about 20 inches. Since the Aluminium Mylar does not support itself well, a nine inch diameter by five inch high bucket is used to support the funnel.

The cooking container primarily tested is a glass canning jar that has been painted flat black. The black paint allows the jar to absorb the sun's rays. The canning jar works well due to the added pressure-cooker effect caused by the rubber ring on the inside of the lid. A black-enamel pot and a black-painted stainless steel canister were also used. We found immediately that raising the vessel off the bottom of the cooker using a rabbit-wire stand provided more rapid and even heating than the wooden block used previously. Placing the jar or pot on a wire stand allows as much reflected light onto the cooking vessel as possible. This allows even the bottom of the cooking container to absorb thermal energy that is reflected off the lower portion of the funnel.

Two methods of closing the cookers off from convection currents were used. It is important to keep the air that surrounds the container from circulating, thus keeping the cooking container from being cooled by convection currents or breezes. This first method used was to enclose the cooking vessel and wire stand in a clear plastic bag, such as a heat resistant Reynolds Oven Bag. It is important to make sure that the bag is not touching the cooking vessel, so once the vessel is placed into the clear bag, air is blown into the bag and the bag is tied off. This is the most common method used for solar panel cookers, such as the Solar CooKit, because of the bags' ability to withstand the temperatures attained in these types of cookers. But these bags tear rather easily and they are not readily available in developing countries and must be imported.

The second method of closing off the cooking vessel from convection currents, designed by Dr. Jones, is to place a clear plastic disk down into the funnel above the cooking vessel. The funnel used in the test was a conventional-shaped funnel that was constructed out of thin sheet metal and aluminium-foil lined for better reflectivity. The diameter of this funnel is about 30 inches at the top, and it stands about 16 inches high. The walls also form about a 60 degree angle with respect to the horizontal. This funnel was designed to hold a larger cooking container such as a pot. The diameter of the plastic disk is large enough that the disk does not touch the top of the container. For the experiments that tested this method, a one-sixteenth inch (1.6 mm) thick Lexan disk was used.

Data Collection

To collect the temperatures as a function of time, a Texas Instruments Calculator Based Laboratory (CBL) was used. This portable interface is capable of recording real-time data from multiple channels. The data were downloaded into a graphing calculator, where they can be analysed and graphed immediately. From the calculator, the data can be transferred to a computer spreadsheet such as Microsoft Excel for further analysis. Due to the nature of these experiments and the low cost to purchase the CBL, this is an ideal data collector to use. A graphing calculator was used to program the CBL and to tell it what data to collect, how many points to collect, and the time period between data points collected. Since the CBL does not have any internal programs for data collection, a program must be written into the graphing calculator. There are

ready-made programs that can be uploaded into the calculator, or a custom program can be made to fit the needs of the test. The program that the CBL used allowed multiple thermocouples to collect data simultaneously. To ensure that the thermocouples were calibrated against each other, both were run on the same constant temperature sample in very close proximity. Both temperature probes agreed to within 0.21^{O}C of each other. For these experiments, this temperature difference was considered to be acceptable.

Procedure

Each experiment was conducted on the campus of Brigham Young University during mid-day, usually between 11:00 am and 2:00 pm to ensure that the sun was close to being directly over-head. This allowed as much sun light as possible to enter the solar collector. Each experiment included several steps, as listed below.

Before each experiment was set up, the volume of the water and the mass of the container were measured and recorded. The heat capacity of the water and the container were also found. The area of the cooker perpendicular to the sun's rays was also measured. To collect temperature data using thermocouple probes, small holes were drilled into the top of the canning jar and stainless steel canister lids. The jar and canister were both painted ultra-flat black to absorb as much of the sun's energy as possible.

On the morning of each test, the designated volume of water was measured out and poured into the cooking vessel. This volume ranged from 0.6 litre for one-quart jars, to 1.2 litters for half-gallon canning jars. For simultaneous testing, the same amount of water was poured into each container. The temperature probes were wired through the holes in the lids of the containers and secured about 13 mm into the water. For comparative tests, the probes were placed the same depth into the water to ensure that the probes did not read different measurements due to depth-related temperature differences within the containers. To enable later analysis; the time, ambient temperature, and solar irradiance were also noted and recorded. These numbers gave a reference point for each test. Each cooker that was to be tested was then completely set up. The temperature probes were secured through the lids, and the jar was placed into the clear oven bag – supported by a wire cage. Each bag was inflated so that no part of the bag touched the sides or top of the cooking container. The cord from the thermocouple to the CBL was passed through the top of the bag, and the bag was tied off with a twist-tie.

The test began once both cookers were completely ready and the CBL had been programmed. Care was taken to block the sun from radiating directly onto the cookers until both were ready to begin. This ensured that the water in both cookers started at very nearly the same temperature. Most tests were set up to collect one data point every four to five minutes, for up to two hours. This allowed the cooker temperatures to reach maxima and then remain at a nearly constant temperature. Once a test was complete, the cooker was disassembled and the data downloaded into the graphing calculator. Though the graphing calculator does allow analysis, a spread sheet such as Microsoft Excel is easier to use. Thus, the data from each test were downloaded from the calculator into Microsoft Excel. The elapsed time (in seconds) and the corresponding temperatures were listed next to each other. A graph of temperature versus time was made, with the Time being the horizontal axis for each test. For comparative tests, the Temperature versus Time data for both cookers was plotted on the same graph. As a reference, a trend-line was fitted to the linear portion of the graph, along with the linear regression and the coefficient of correlation (R2). It is important to have a coefficient of correlation close to one, as this is how close the linear regression fits the data. In a separate column, the temperatures were again listed, however only from 30^{O}C to 70^{O}C. The change in temperature for every ten or twelve minutes was found and logged next to the temperature column. The power output (in Watts) of each cooker could then be calculated.

To calculate the power output of the cookers for each specific test, the mass of the water and of the container were both measured. Though the thermal energy content of the container was relatively small compared to that of water (due to the large heat capacity of water), it was important to add it into the calculation. Also, since several different containers were compared, the energy content of the container was important. The power is found by:

$$Q_{(out)} = Q_{(water)} + Q_{(container)}$$

$$Q_{(out)} = (m_w c_w + m_c c_c)\Delta T$$

$$Power_{(out)} = \frac{Q_{(out)}}{\Delta t}$$

The power is found in Watts. A power output for each change in temperature for the time interval is calculated and logged next to the T column. Since there are uncertainties in all of the measurements, it is important to include the error in each power output. To do this, the error in the water's and container's measurements is taken into consideration. The error is found by:

$$\pm \Delta P = \sqrt{(\frac{\partial P}{\partial m_w}\Delta m_w)^2 + (\frac{\partial P}{\partial m_c}\Delta m_c)^2 + (\frac{\partial P}{\partial t}\Delta t)_w^2 + (\frac{\partial P}{\partial t}\Delta t)_c^2 + 2(\frac{\partial P}{\partial T}\Delta T)_w^2 + 2(\frac{\partial P}{\partial T}\Delta T)_c^2}$$

Where ±dP is the total error in the calculated error, dmw and dmc are the error in the mass of the water and container respectively, delta-Tp is the error in the temperature difference, and delta-t is the error in the time interval.

This simplifies to:

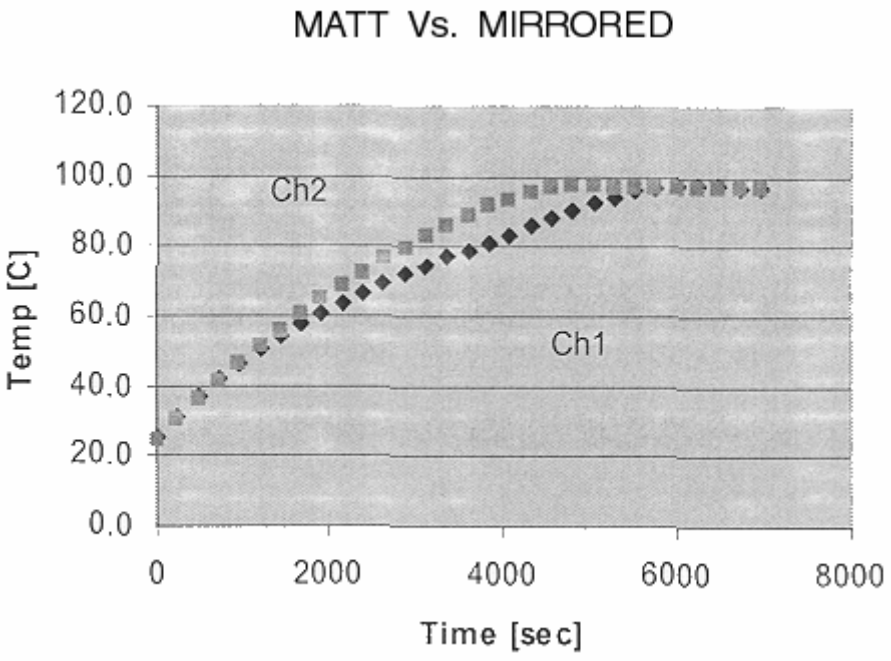
$$\pm \Delta P = \sqrt{[(\frac{c_w \Delta T}{t}\Delta m_w)^2 + (\frac{c_c \Delta T}{t}\Delta m_c)^2 + (\frac{m_w c_w \Delta T}{t^2}\Delta t)^2 + (\frac{m_c c_c \Delta T}{t^2}\Delta t)^2 + 2(\frac{m_w c_w}{t}\Delta T_p)^2 + 2(\frac{m_c c_c}{t}\Delta T_p)^2]}$$

The error was found only for the average change in temperature, rather than for each individual temperature measurement. Since the power output is dependant on the amount of energy coming in from the sun, the cooker efficiency is a good factor to calculate. To find the efficiency, the total amount of local solar radiation must be known. This should be given in watts per square metre, so that the input wattage can be found. To find the power coming in, the area of the cooker perpendicular to the sun's rays was multiplied by the solar radiation to give the amount of power that was being collected by the cooker. Since the Solar Funnel is able to be kept on track with the sun, and since the tests were done during mid-day, it was not necessary to calculate any angles. The efficiency is simply the power output divided by the power input. The solar radiation for each test was supplied by the Department of Physics and Astronomy weather station at Brigham Young University in Provo, UT, where the tests took place.

Results:

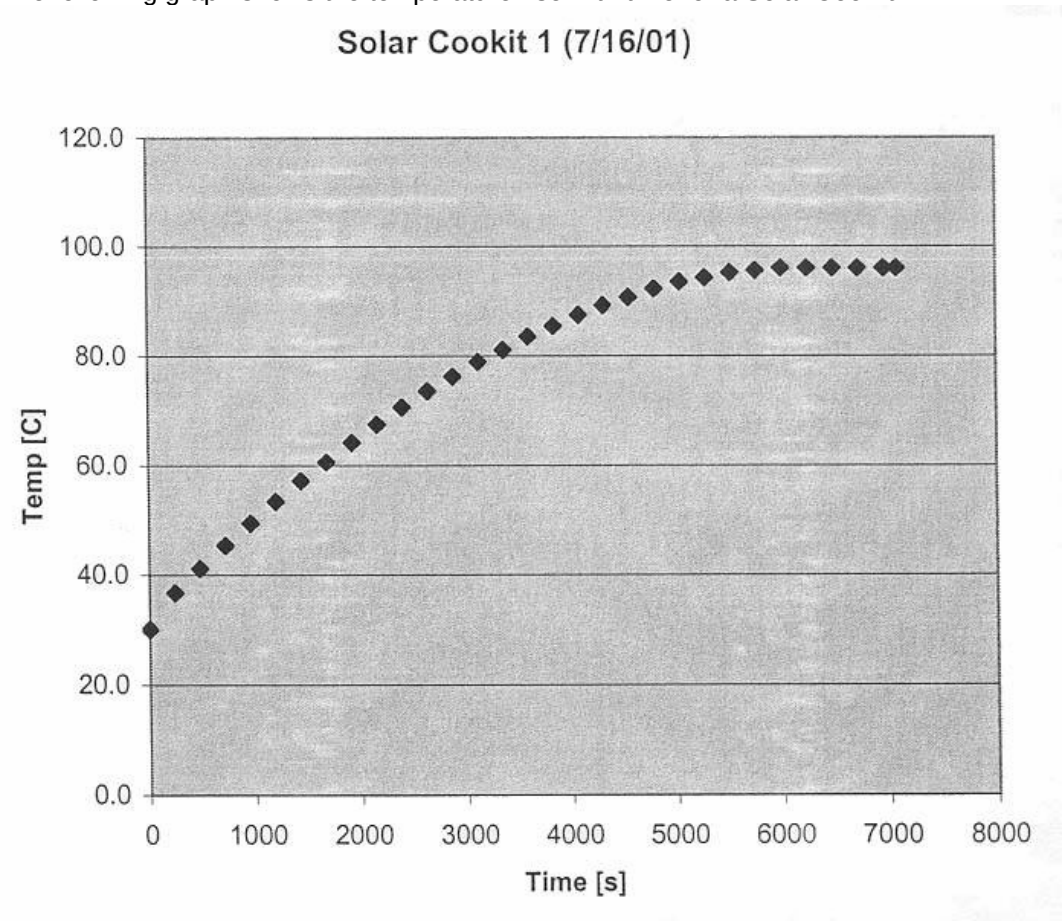
Matt vs. Mirror: Several tests were conducted on the matt versus mirror finishes. In each test, the matt finish outperformed the mirror finish. On 27 July, 2001, a matt funnel and a mirror funnel were simultaneously tested with 650 cc of water. The average power output for the mirror finish was 46.4 W ± 1.7 W, while the matt funnel put out an average of 59.4 W ± 2.1 W. The efficiency of the mirror funnel was 15.8%, while the matt was 20.2% efficient.

The following graph shows the temperatures reached by the matt and mirror funnels.



Channel 1 (Ch1) was the mirror finish, and channel 2 (Ch2) was the matt finish. This shows that both funnels peaked at about the same temperature: 97^{O}C (207^{O}F). The matt funnel peaked in about 76 minutes, whereas the mirror funnel peaked in 96 minutes, twenty minutes later. Though this perhaps a tolerable time difference for actual cooking, it is substantial. Every matt vs. mirror test performed in a similar way. These results are due to the way the matt funnel reflects the sun's rays. The mirror finish seems to focus a strip of light onto the cooking vessel more than the matt finish does. As a result, the matt finish diffuses the light more and the cooker is heated more uniformly. This is good, since the matt finish is easier to work with, delivering much less glare to the eyes.

The following graph shows the temperature rise with time for a Solar CooKit:



Comparing the two graphs above, we find that the Solar CoolKit performed very well, comparable to the Funnel Cooker. We should note that in both cases, we used a canning jar (pressurised) supported by a wire stand. We found that the wire stand improves the performance of the Solar CooKit significantly and hope that this support stand will be used in countries where the Solar CooKit is in use.

In tests where the use of the clear plastic disk was tested against the oven-bag, an aluminium pot was used in the disk-set-up. In these tests, the cooker with an oven bag outperformed the cooker using a plastic disk. On 10 August, 2001, a test was run which compared the disk/pot set-up against the oven-bag/jar set-up. Both cookers follow similar heating paths with time, but the oven-bag/jar did slightly better. Due to the higher mass of the jar compared to the mass of the aluminium pot, and the much higher heat capacity of the water, the average power output for the oven-bag/jar was 39.8 ± 1.4 W, while the disk/pot put out 30.3 W ± 1.2 W. The efficiency of the oven-bag/jar was 14.7% and the efficiency of the disk/pot set-up was 10.4% for this test. This is also partly due to the pressure-cooker effect that the canning jar produces. Though this is a considerable efficiency difference, the disk/pot set-up did very well in subjective tests where food was actually cooked and tasted. In all cases where the disk/pot set-up was used to cook food, the food cooked in about the same amount of time. The ease of the disk/pot set-up is also an important consideration. Overall, in tests where food was cooked, the disk/pot set-up was preferred over the oven-bag/jar set-up.

Conclusions:

As many countries are depleting their natural resources due to increased population and the resulting deforestation, methods other than burning wood are needed to cook food and pasteurise water. Solar

cookers provide a sustainable technology that relies on the sun's free energy. We report several advances to make them better. The need for cheap and effective solar cookers is very great and growing.

The Solar Funnel Cooker has been designed to meet the growing need by being inexpensive and effective. We determined that the Solar CooKit was nearly as effective when a rabbit-wire stand was used to support the cooking vessel. By collecting time vs. temperature data, quantitative analysis has been done. This analysis approach is useful for further development of the cookers.

Several areas of research were explored in 2001. Two finishes were tested for the reflector, a matt finish and a mirror finish. The benefits of the matt over the mirror finish are:

1) The matt finish is easier to work over because the sun's glaring reflection is diffused, and
2) the matt finish out-performs the mirror finish in temperature vs. time tests.

The method of closing off the cooker from convection current was tested and compared with an alternative method – a clear plastic disk. The use of a pot rather that a canning jar was also tested. Though the present oven-bag/jar method does outperform the disk/pot method, the disk/pot method is easier to use and seems to be nearly as efficient. Finally, we showed that a wire-mesh stand is a considerable improvement over the use of a wooden block or other opaque stand for the cooking vessel. We join with our fellow researchers around the world in pursuing further development of solar cookers, particularly to benefit people in developing countries.

References:

[1]. Jones, Steven E. et al., BYU. [2]. Wattenberg, Frank. Montana State University. 1996.
[2]. Wattenberg, Frank. Montana State University. 1996.

Recent Advances in Solar Water Pasteurisation

Boiling isn't necessary to kill disease microbes

The main purpose of solar cookers is to change sunlight into heat which is then used to cook foods. We are all familiar with how successful solar cookers are at cooking and baking a wide variety of foods. In this article I want to consider using the heat in solar cookers for purposes other than cooking. My main focus will be solar water pasteurisation, which can complement solar cooking and address critical health problems in many developing countries.

The majority of diseases in developing countries today are infectious diseases caused by bacteria, viruses, and other microbes which are shed in human faeces and polluted water which people use for drinking or washing. When people drink the live microbes, they can multiply, cause disease, and be shed in faeces into water, continuing the cycle of disease transmission.

World-wide, unsafe water is a major problem. An estimated one billion people do not have access to safe water. It is estimated that diarrhoeal diseases that result from contaminated water kill about 2 million children and cause about 900 million episodes of illness each year.

Boiling contaminated water

How can infectious microbes in water be killed to make the water safe to drink? In the cities of developed countries this is often guaranteed by chlorination of water after it has been filtered. In developing countries, however, city water systems are less reliable, and water from streams, rivers and some wells may be contaminated with human faeces and pose a health threat. For the billion people who do not have safe water to drink, what recommendation do public health officials offer? The only major recommendation is to boil the water, sometimes for up to 10 minutes. It has been known since the time of Louis Pasteur 130 years ago that heat of boiling is very effective at killing all microbes which cause disease in milk and water.

If contaminated water could be made safe for drinking by boiling, why is boiling not uniformly practised? There seem to be five major reasons:

1) people do not believe in the germ theory of disease,
2) it takes too long,
3) boiled water tastes bad,
4) fuel is often limited or costly,
5) the heat and smoke are unpleasant.

Some examples of the cost of boiling water are worth mentioning. During the cholera outbreak in Peru, the Ministry of Health urged all residents to boil drinking water for 10 minutes. The cost of doing this would amount to 29% of the average poor household income. In Bangladesh, boiling drinking water would take 11% of the income of a family in the lowest quartile. In Jakarta, Indonesia, more than $50 million is spent each year by households for boiling water. It is estimated that in the city of Cebu in the Philippines, population about 900,000, about half the families boil their drinking water, and the proportion is actually higher for families that obtain their water from an unreliable chlorinated piped supply. Because the quantities of fuel consumed for boiling water are so large, approximately 1 kilogram of wood to boil 1 litre of water, and because firewood, coal, and coke are often used for this purpose, an inadequate water supply system significantly contributes to deforestation, urban air pollution, and other energy-related environmental effects.

If wood, charcoal, or dung is used as fuel for boiling water, the smoke creates a health hazard, as it does all the time with cooking. It is estimated that 400 to 700 million people, mainly women, suffer health problems from this indoor air pollution. As a microbiologist, I have always been perplexed as to why boiling is recommended, when this is heat far in excess of that which is necessary to kill infectious microbes in water. I presume the reason boiling is recommended is to make sure that lethal temperatures have been reached, since unless one has a thermometer it is difficult to tell what temperature heated water has reached until a roaring boil is reached. Everyone is familiar with the process of milk pasteurisation. This is a heating process which is sufficient to kill the most heat resistant disease causing microbes in milk, such as the bacteria which cause tuberculosis, undulant fever, streptococcal infections and Salmonellosis. What temperatures are used to pasteurise milk? Most milk is pasteurised at 71.7^{O} C (161^{O} F) for only 15 seconds. Alternatively, 30 minutes at 62.8^{O} C (145^{O} F) can also pasteurise milk. Some bacteria are heat resistant and can survive pasteurisation, but these bacteria do not cause disease in people. They can, however, spoil the milk, so pasteurised milk is kept refrigerated.

There are some different disease microbes found in water, but they are not unusually heat resistant. The most common causes of water diseases, and their heat sensitivity, are presented in Table 1. The most common causes of acute diarrhoea among children in developing countries are the bacteria Escherichia coli and Shigelia SD. and the Rotavirus group of viruses. These are rapidly killed at temperatures of 60^{O} C or greater.

Solar water pasteurisation

As water heats in a solar cooker, temperatures of 56^{O} C and above start killing disease-causing microbes. A graduate student of mine, David Ciochetti, investigated this for his master's thesis in 1983, and concluded that heating water to 66^{O} C in a solar cooker will provide enough heat to pasteurise the water and kill all disease causing microbes. The fact that water can be made safe to drink by heating it to this lower temperature - only 66^{O} C - instead of 100^{O} C (boiling) presents a real opportunity for addressing contaminated water in developing countries.

Testing water for faecal contamination

How can one readily determine if the water from a well, pump, stream, etc. is safe to drink? The common procedure is to test the water for bacterial indicators of faecal pollution. There are two groups of indicators which are used. The first is the coliform bacteria which are used as indicators in developed countries where water is chlorinated. Coliform bacteria may come from faeces or from plants. Among the coliform bacteria is the second indicator, Escherichia coli. This bacterium is present in large numbers in human faeces (approximately 100,000,000 per gram of faeces) and that of other mammals. This is the main indicator used if water is not chlorinated. A water source containing 100 E. coli per 100 ccs poses a substantial risk of disease.

The standard method of testing water for the presence of coliforms and E. coli requires trained personnel and a good laboratory facility or field unit which are usually not present in developing countries. Thus, water supplies are almost never tested.

A new approach to testing in developing countries

In 1987, the Colilert MPM Test (CLT) was introduced as the first method which used a defined substrate technology to simultaneously detect coliforms and E. coli. The CLT comes as dry chemicals in test tubes containing two indicator nutrients: one for coliforms and one for E. coli. The CLT involves adding 10 ml of water to a tube, shaking to dissolve the chemicals, and incubating at body temperature for 24 hours. I prefer incubating tubes under my belt against my body. At night I sleep on my back and use night clothes to hold the tubes against my body.

If no coliform bacteria are present, the water will remain clear. However, if one or more coliforms are present in the water, after 24 hours their growth will metabolise ONPG and the water will change in colour from clear to yellow (resembling urine). If E. coli is among the coliform bacteria present, it will metabolise MUG and the tube will fluoresce blue when a battery-operated, long-wave ultraviolet light shines on it, indicating a serious health hazard. I have invited participants at solar box cooker workshops in Sierra Leone, Mali, Mauritania, and Nepal to test their home water supplies with CLT. One hundred and twenty participants brought in samples. In all four countries, whether the water was from urban or rural areas, the majority of samples contained coliforms, and at least half of these had E. coli present. Bacteriological testing of the ONPG and MUG positive tubes brought back from Mali and Mauritania verified the presence of coliforms/E. coli in approximately 95% of the samples. It is likely that soon the Colilert MPN test will be modified so that the test for E. coli will not require an ultraviolet light, and the tube will turn a different colour than yellow if E coli is present. This will make the test less expensive and easier to widely use in developing countries to assess water sources.

Effect of safe water on diarrhoea in children

What would be the effect if contaminated water could be made safe for drinking by pasteurisation or boiling? One estimate predicts that if in the Philippines, families at present using moderately contaminated wells (100 E. coli per 100 ml) were able to use a high-quality water source, diarrhoea among their children would be reduced by over 30%. Thus, if water which caused a MUG (+) test were solar pasteurised so it would be clear, this would help reduce the chance of diarrhoea, especially in children.

Water pasteurisation indicator

How can one determine if heated water has reached 65^{O} C? In 1988, Dr. Fred Barrett (USDA, retired) developed the prototype for the Water Pasteurisation Indicator (WAPI). In 1992, Dale Andreatta, a graduate engineering student at the University of California, Berkeley, developed the current WAPI. The WAPI is a polycarbonate tube, sealed at both ends, partially filled with a soybean fat which melts at 69^{O} C

("MYVEROL" 18-06K, Eastman Kodak Co., Kingsport, TN 37662). The WAPI is placed inside a water container with the fat at the top of the tube. A washer will keep the WAPI on the bottom of the container, which heats the slowest in a solar box cooker. If heat from the water melts the fat, the fat will move to the bottom of the WAPI, indicating water has been pasteurised. If the fat is still at the top of the tube, the water has not been pasteurised.

The WAPI is reusable. After the fat cools and becomes solid on the bottom, the fish line string is pulled to the other end and the washer slides to the bottom, which places the fat at the top of the tube. Another pasteurisation indicator has been developed by Roland Saye which is based on expansion of a bi-metal disc which is housed in a plastic container. This also shows promise and is in the early testing stages. The WAPI could be useful immediately for people who currently boil water to make it safe to drink. The WAPI will indicate clearly when a safe temperature has been reached, and will save much fuel which is currently is being wasted by excessive heating.

[Editor's note: Using Beeswax & Carnauba Wax to Indicate Temperature: In SBJ #15 we discussed using beeswax, which melts at a relatively low 62^{O} C, as an indicator of pasteurisation. We have now found that mixing a small amount of carnauba was with the beeswax (~1:5 ratio) raises the melting temperature of the beeswax to 70^{O} - 75^{O} C. Carnauba wax is a product of Brazil and can be bought in the US at woodworking supply stores. Further testing needs to be done to confirm that the melting point remains the same after repeated re-melting.

Different strategies for solar water pasteurisation

The solar box cooker was first used to pasteurise water. David Ciochetti built a deep-dish solar box cooker to hold several gallons of water. At this time of the year in Sacramento, three gallons could be pasteurised on our typical sunny days.

Dale Andreatta and Derek Yegian of the University of California. Berkeley, have developed creative ways to greatly increase the quantity of water which can be pasteurised, as we will hear about at this conference.

I am also excited about the possibility of pasteurising water using the simple solar panel cookers. By enclosing a dark water container in a polyester bag to create an insulating air space, and by using lots of reflectors to bounce light onto the jar, it is possible to pasteurise useful amounts of water with a simple system. It takes about four hours for me to pasteurise a gallon of water in the summer with the system I am using. Solar panel cookers open up enormous possibilities for heating water not only for pasteurisation, but also for making coffee and tea, which are quite popular in some developing countries. The heated water can also be kept hot for a long time by placing it in its bag inside an insulated box. In the insulated container I use, a gallon of 80^{O} C water will be approximately 55^{O} C after 14 hours. Water at a temperature of 55^{O} C will be about 40^{O} C after 14 hours, ideal for washing/shaving in the morning.

I will close with some advice from the most famous microbiologist, who pioneered the use of vaccinations in the 1890s: Louis Pasteur. When he was asked the secret of his success, he responded that above all else, it was persistence. I will add that you need good data to be persistent about, and we certainly have that with solar cookers; the work in Sacramento, Bolivia, Nepal, Mali, Guatemala, and wherever else the sun shines. Continued overuse of fuel-wood is non-sustainable. We need to persist until the knowledge we have spreads and becomes common knowledge world-wide.

For questions or comments contact Dr. Robert Metcalf at.

Dr. Robert Metcalf
1324 43rd St.
Sacramento, California 95819 USA.

IDEXX Laboratories, Inc. makes the Colilert kit and is located at this address:

IDEXX Laboratories, Inc.
One IDEXX Drive
Westbrook, ME 04092
USA
Voice: (800) 321-0207 or (207) 856-0496
Fax: (207) 856-0630

Editor's Note: Testing Water in Developing Countries

The Colilert system makes it possible to test water without the need for a laboratory. IDEXX Laboratories, the manufacturer, recommends that you use five test tubes for each sample. Bob Metcalf explains that five tubes would comprise 50 ml, which is the minimum sample size permitted by US law. This is an unrealistically high standard by which to judge the water in developing countries where you are examining water that is already being drunk, in spite of the fact that it may be making people sick. By using a single test tube (10 ml) there is a very small chance that your sample missed the small number of bacteria that might have been present.

IDEXX Laboratories will also tell you that you need an incubator to achieve valid results. Again, Bob Metcalf tells us that all that is needed is to keep the tubes close to your body for 36 hours, since body temperature is the correct incubation temperature.

What you are actually measuring in the test is the presence of 1) coliform bacteria, and 2) E. coli, a type of coliform bacteria that is largely found in faecal matter. A positive test for coliform bacteria might be due to coliform bacteria that has washed off of plant leaves , and thus be fairly innocuous. A positive test for E. coli, however, would indicate that any bacteriological contamination was from a faecal source, which might also contain Giardia, cholera, or other serious infectious microbes.

This document is published on The Solar Cooking Archive at
http://solarcooking.org/pasteurisation/metcalf.htm.

The Solar Puddle

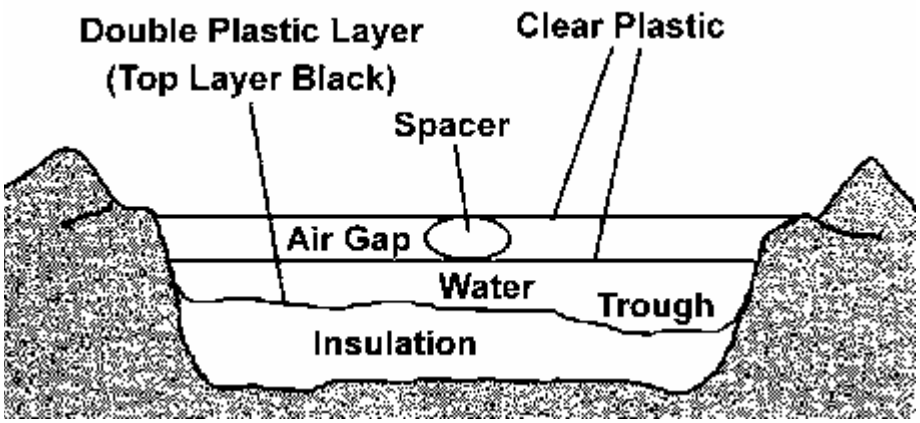
A new water pasteurisation technique for large amounts of water

The lack of clean drinking water is a major health problem in the developing world. To reduce this health risk ways of producing clean water at an affordable cost are needed, and people need to be educated about germs and sanitation, lest they accidentally re-contaminate their clean drinking water. Recently, several of us at the University of California at Berkeley have attacked the first of these requirements. Previous issues of this newsletter have included stories about our water pasteurisation indicator and our flow-through water pasteurises based on a design by PAX World Service. In this article we describe a new low-cost device that pasteurises water.

For those not familiar with the pasteurisation process, if water is heated to 149^{O} F (65^{O} C) for about 6 minutes all the germs, viruses, and parasites that cause disease in humans are killed, including cholera and hepatitis A and B. [Ed. We have reports from the field that at 145^{O} F (63^{O} C) in a solar puddle, bacterial growth might actually be increased. Since this temperature is very close to the minimum pasteurisation temperature mentioned in this article, we suggest that you heat the water to a higher

temperature and perform tests before adopting a solar puddle as your method of pasteurisation]. This is similar to what is done with milk and other beverages. It is not necessary to boil the water as many people believe. Pasteurisation is not the only way to decontaminate drinking water, but pasteurisation is particularly easy to scale down so the initial cost is low.

The new device is called a solar puddle, and it is essentially a puddle in a greenhouse. One form of the solar puddle is sketched in the figure below, though many variations are possible.



One begins by digging a shallow pit about 4 inches deep. The test device was a "family-size" unit, about 3.5 feet by 3.5 feet, but the puddle could be made larger or smaller. If the puddle is made larger there is more water to pasteurise, but there is also proportionately more sunshine collected. The pit is filled with 2 to 4 inches of solid insulation. We used wadded paper, but straw, grass, leaves, or twigs could be used. This layer of insulation should be made flat, except for a low spot in one corner of the puddle.

Put a layer of clear plastic and then a layer of black plastic over the insulation with the edges of the plastic extending up and out of the pit. Two layers are used in case one develops a small leak. We used inexpensive polyethylene from a hardware store, though special UV stabilised plastic would last longer. Put in some water and flatten out the insulation so that the water depth is even to within about 0.5 inch throughout the puddle, except in the trough which should be about 1 inch deeper than the rest. Put in more water so that the average depth is 1 to 3 inches depending on how much sunshine is expected.

A pasteurisation indicator (available from Solar Cookers International at 916/455-4499) should go in this trough since this is where the coolest water will collect. Put a layer of clear plastic over the water, again with the edges extending beyond the edges extending beyond the edges of the pit. Form an insulating air gap by putting one or more spacers on top of the third layer of plastic (large wads of paper will do) and putting down a fourth layer of plastic, which must also be clear. The thickness of the air gap should be 2 inches or more. Pile dirt or rocks on the edges of the plastic sheets to hold them down. The puddle is drained by siphoning the water out, placing the siphon in the trough and holding it down by a rock or weight. If the bottom of the puddle is flat, well over 90% of the water can be siphoned out.

Once the puddle is built it would be used by adding water each day, either by folding back the top two layers of plastic in one corner and adding water by bucket, or by using a fill siphon. The fill siphon should NOT be the same siphon that is used to drain the puddle, as the fill siphon is re-contaminated each day, while the drain siphon MUST REMAIN CLEAN. Once in place the drain siphon should be left in place for the life of the puddle.

The only expensive materials used to make the puddle are a pasteurisation indicator (about $2 for the size tested). All of these items are easily transportable, so the solar puddle might be an excellent option for a refugee camp if the expertise were available for setting them up.

Many tests were done in the spring and summer of this year in Berkeley, California. On days with good sunshine the required temperature was achieved even with 17 gallons of water (2 1/2 inch depth). About 1 gallon is the minimum daily requirement per person, for drinking, brushing one's teeth, and dish washing. With thinner water layers higher temperatures can be reached. With 6 gallons (1 inch depth) 176^{O} F was achieved on one day.

The device seems to work even under conditions that are not ideal. Condensation in the top layer of plastic doesn't seem to be a problem, though if one gets a lot of condensation the top layer should be

pulled back to let the condensation evaporate. Small holes in the top layers don't make much difference. The device works in wind, or if the bottom insulation is damp. Water temperature is uniform throughout the puddle to within 2^{O} F.

After some months the top plastic layers weaken under the combined effects of sun and heat and have to be replaced, but this can be minimised by avoiding hot spots. Another option would be to use a grade of plastic that is more resistant to sunlight. The two bottom layers of plastic tend to form tiny tears unless one is very careful in handling them, (that is why there are two layers on the bottom). A tiny hole may let a little water through and dampen the solid insulation, but this is not a big problem.

There are many variations of the solar puddle. We've been able to put the top layer of plastic into a tent-like arrangement that sheds rain. This would be good in a place that gets frequent brief showers. Adding a second insulating layer of air makes the device work even better, though this adds the cost of an extra layer of plastic. As mentioned the device can cover a larger or smaller area if more or less water is desired. One could make a water heater by roughly tripling the amount of water so that the maximum temperature was only 120° F or so, and this water would stay warm well into the evening hours. This water wouldn't be pasteurised though. One could help solve the problem of dirty water vessels by putting drinking cups into the solar puddle and pasteurising them along with the water. The solar puddle could possibly cook foods like rice on an emergency basis, perhaps in a refugee camp.

You can contact
Dr. Dale Andreatta
S. E. A. Inc.
7349 Worthington-Galena Rd.
Columbus, OH 43085
(614) 888-4160 FAX (614) 885-8014

This document is published on The Solar Cooking Archive at
http://solarcooking.org/pasteurisation/puddle.htm.

Important web link: http://solarcooking.org/plans/default.htm

The "Easy Lid" Solar Cooker.

Designed by Chao Tan and Tom Sponheim

Although designs for cardboard cookers have become more simple, fitting a lid can still be difficult and time consuming. In this version, a lid is formed automatically from the outer box.

Making the Base

Take a large box and cut it in half as shown in Figure 1. Set one half aside to be used for the lid. The other half becomes the base.

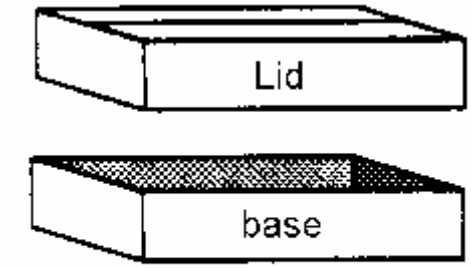


Figure 1

Fold an extra cardboard piece so that it forms a liner around the inside of the base (see Figure 2).

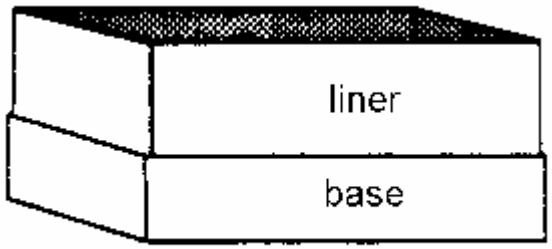


Figure 2

Use the lid piece as shown in Figure 3 to mark a line around the liner.

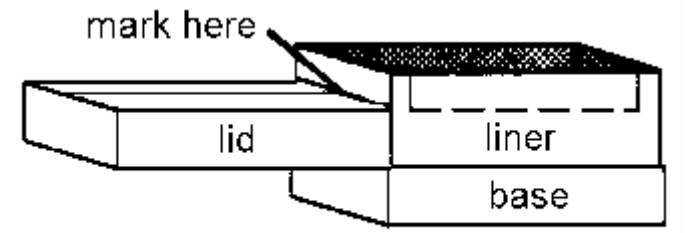


Figure 3

Cut along this line, leaving the four tabs as shown in Figure 4.

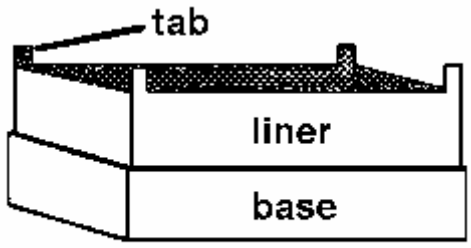


Figure 4

Glue aluminium foil to the inside of the liner and to the bottom of the outer box inside.

Set a smaller (inner) box into the opening formed by the liner until the flaps of the smaller box are horizontal and flush with the top of the liner (see Figure 5). Place some wads of newspaper between the two boxes for support.

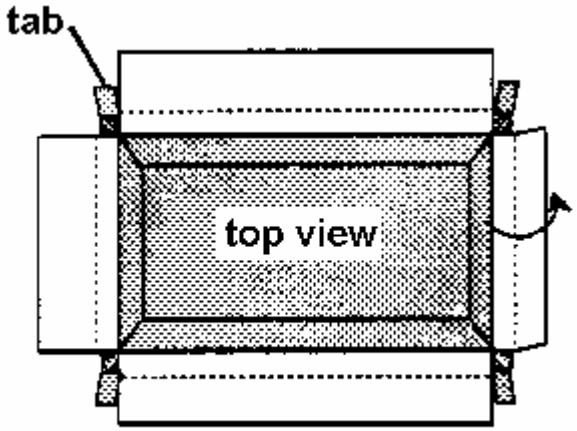


Figure 5

Mark the underside of the flaps of the smaller box using the liner as a guide.

Fold these flaps down to fit down around the top of the liner and tuck them into the space between the base and the liner (see Figure 6).

Fold the tabs over and tuck them under the flaps of the inner box so that they obstruct the holes in the four corners (see Figure 6).

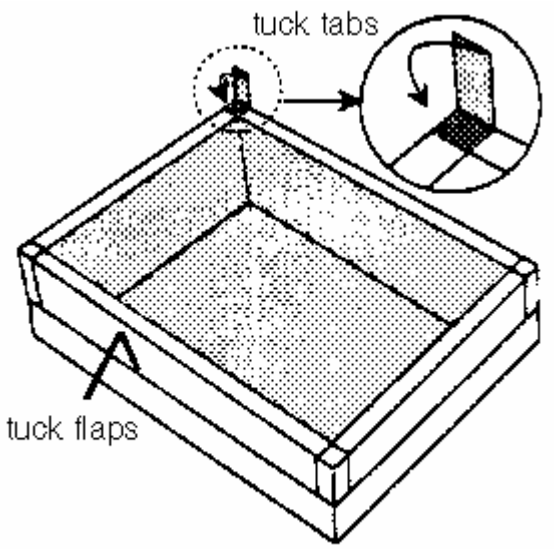


Figure 6

Now glue these pieces together in their present configuration.

As the glue is drying, line the inside of the inner box with aluminium foil.

Finishing the Lid

Measure the width of the walls of the base and use these measurements to calculate where to make the cuts that form the reflector in Figure 7. Only cut on three sides. The reflector is folded up using the fourth side as a hinge.

Glue plastic or glass in place on the underside of the lid. If you are using glass, sandwich the glass using extra strips of cardboard. Allow to dry.

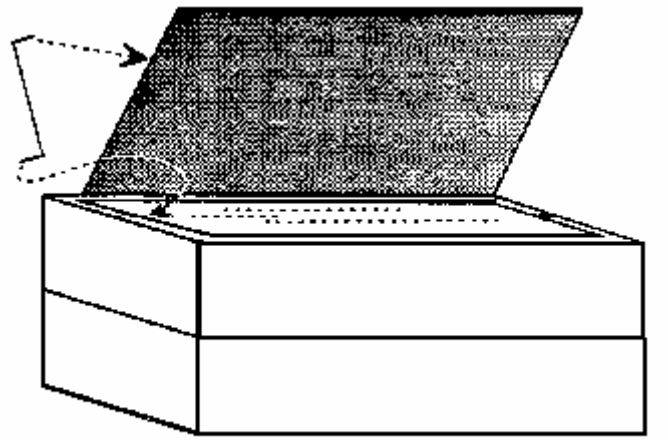


Figure 7

Bend the ends of the wire as shown in Figure 7 and insert these into the corrugations on the lid and on the reflector to prop open the latter.

Paint the sheet metal (or cardboard) piece black and place it into the inside of the oven.

Improving Efficiency

Glue thin strips of cardboard underneath the sheet metal (or cardboard) piece to elevate it off of the bottom of the oven slightly.

Cut off the reflector and replace it with one that is as large as (or larger than) the entire lid. This reflects light into the oven more reliably.

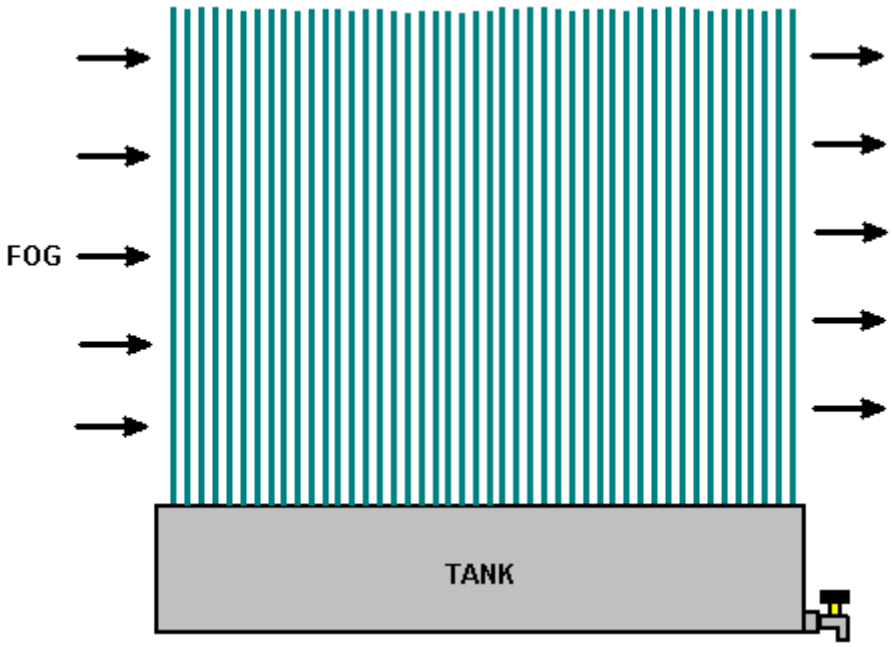
Turn the oven over and open the bottom flaps. Place one foiled cardboard panel into each airspace to divide each into two spaces. The foiled side should face the centre of the oven.

For more information contact:
Solar Cookers International
1919 21st St., Suite 101
Sacramento, CA 95811 USA

****

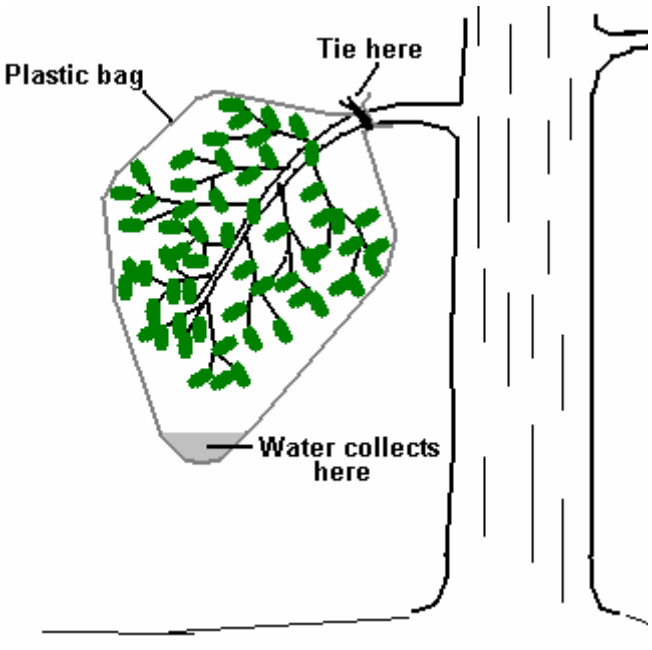
Drinking Water Collection Systems.
Getting adequate drinking water can often be a problem as there are many places where the local water supply is heavily contaminated with poisons, deadly organisms or both. While it is definitely much better to drink purified water than seriously contaminated water, it should be understood that most of the systems mentioned here, produce either distilled water or something very close to distilled water. Drinking distilled water for more than a few weeks is not ideal as distilled water is able to dissolve almost anything and it can do that perfectly well inside your body, removing essential minerals and other important items. So, if possible, avoid drinking distilled water for lengthy periods of time unless there is no other option other than seriously contaminated water carrying diseases and poisons.

One solution introduced for an area where there is almost never any rainfall is particularly interesting. This region gets fogs in the early morning, so plastic devices were constructed to take advantage of this fact. The devices were like plastic clothes brushes with long, slender vertical projections. The fog encountering these, condenses into freshwater droplets on the surface of these vertical fronds and run down the fronds into a plastic tank which forms the base of the device. No moving parts. No input power needed, but the result is large quantities of drinking water every morning. There is very little evaporation from the tanks, due to the small surface area of the stored water:

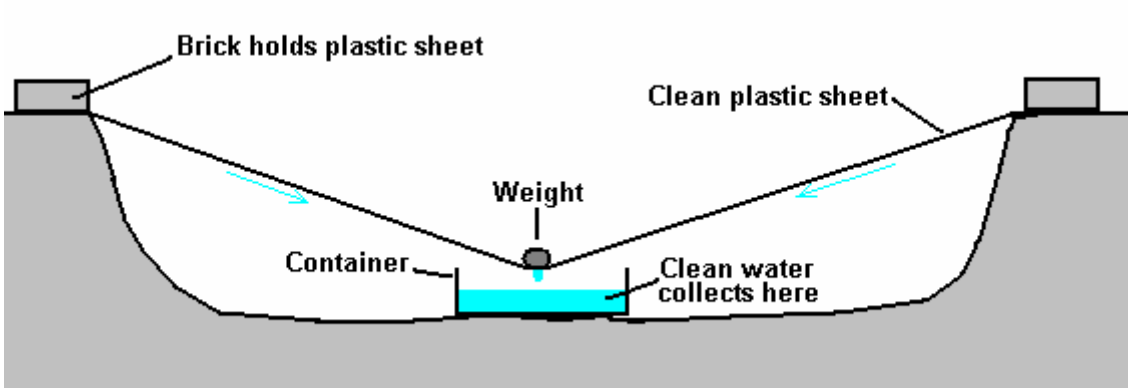


This effect is very noticeable on foggy days where trees drip water extensively due the fog depositing moisture on the leaves and branches.

An emergency measure where water is needed, is to fasten a clean plastic bag around a branch of a tree. Trees lift a large amount of water through their root systems and a good deal of that water exist from the leaves of the tree. The plastic bag intercepts that moisture loss and collects it as clean water:



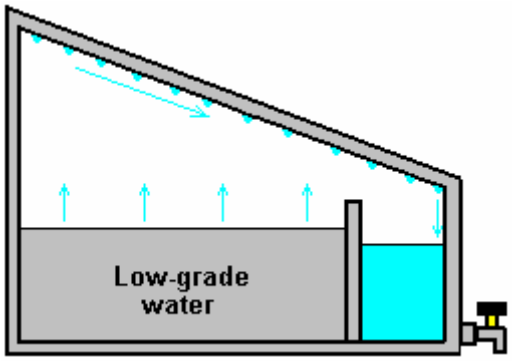
Another emergency measure is to use a clean plastic sheet and a hole dug in the ground. A clean container is placed in the centre of the hole and the plastic sheet used to cover the hold. The sheet is held around the edges of the hole with stones or any other suitable heavy material – bricks, timber, etc. A weight is then placed in the centre of the plastic sheet, pulling it down into a slope in every direction and forming an inverted peak over the container:



The area under the plastic sheet is heated by the greenhouse effect. Moisture also comes from the earth inside the enclosed hole. The moisture in the air in the cavity condenses on the underside of the plastic sheet. But as the plastic sheet is shaped into an inverted pyramid due to the weight just above the container, the water runs down and drips into the container. Again, no input power required and no moving parts.

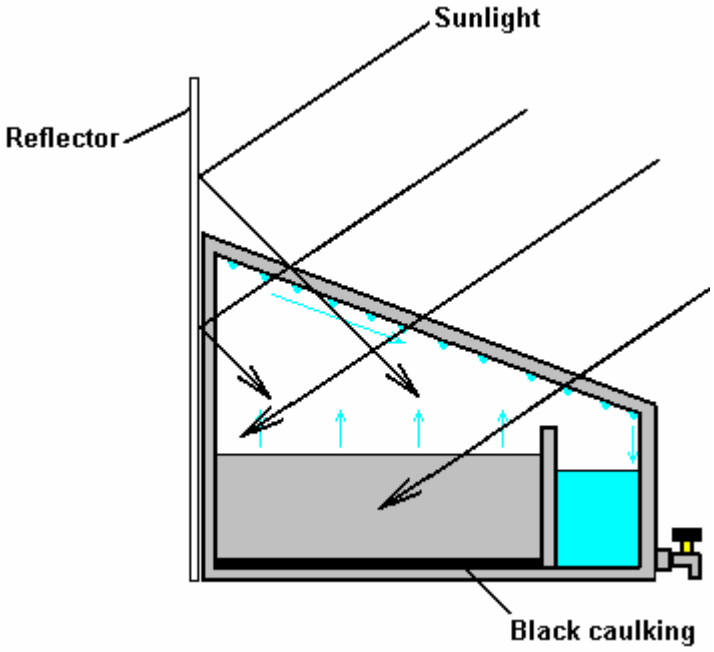
While these methods produce good quality water which is effectively distilled water, it should not be considered to be sterile and immediately ready for human consumption, even though any risk from drinking it 'as-is' is likely to be very low. There will always be air-borne pathogens, and the 'clean' components used to collect the water in the first place may not be as clean as was thought. The same applies to the excellent quality water produced by dehumidifiers, where the inner working surfaces cannot be considered sterile after the equipment has been used for any length of time. To raise the water quality, boiling briefly, microwaving the water or it to UV radiation should kill any remaining harmful organisms in the water and make it fit for consumption.

Applying these same methods on a more permanent basis, leads to the construction of devices of the following type:

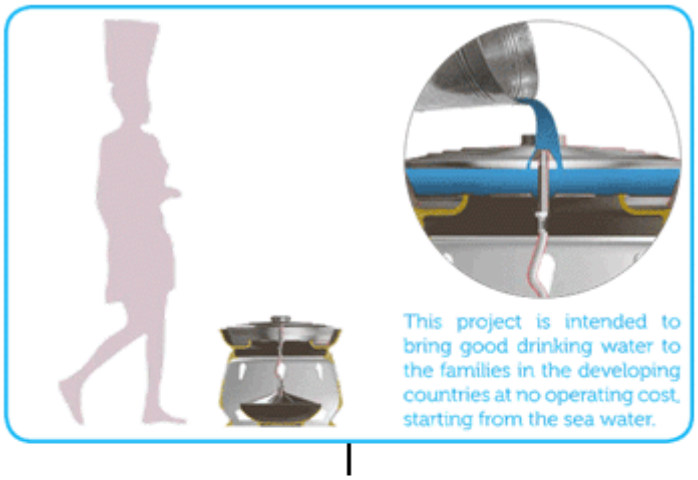


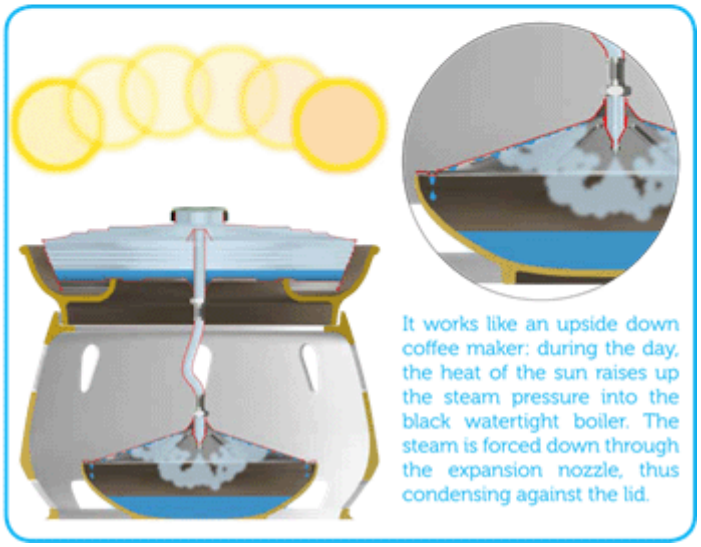
There can be many variations on this shape. These devices are generally built either with glass lids or the whole construction in acrylic sheet. Here, the greenhouse effect heats the inside of the box, causing evaporation of the water inside. This condenses on the walls and lid of the box, where it runs down and into the clean-water section. Please remember that after a long period of use, the device needs to be cleaned very carefully and to deal with air-borne bacteria, the water could be treated briefly with UV light.

This particular design can be further enhanced as shown on the http://www.permapak.net/solarstill.htm web site, where the heating inside the box is upgraded by using black high-temperature silicone to coat the inside of the bottom of the case. The black material absorbs sunlight particularly well and so helps to heat the water. Another enhancement is to place a reflector, possibly made from aluminium foil, behind the unit in order to increase the amount of sunlight or UV radiation reaching the water inside the box:



The web site http://www.gabrielediamanti.com/projects/eliodomestico---how-does-it-work/ shows details of what is considered to be a very effective, low-cost unit aimed at undeveloped countries. Please visit that site where there is a 'Donations' button which allows you to support this excellent work. The designer Gabriele Diamanti says "Eliodomestico is an open project, free to the people who need it. I would really appreciate if you would like to help me in the development of this project!". It works like this:

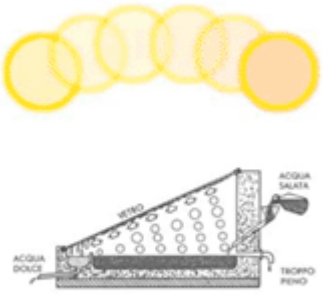




At the end of the day the **Eliodomestico** delivers 5 liters of fresh drinking water. The lower basin is specifically designed for the transport over the head, supporting this common habit.

Eliodomestico is entirely made from poor, widely available materials. The tecnologies involved in the production are very simple and popular. This also make the maintenance much easier.

- No electricity
- No filters
- Very easy maintenance
- Good impact on the local economy
- No impact on the environment

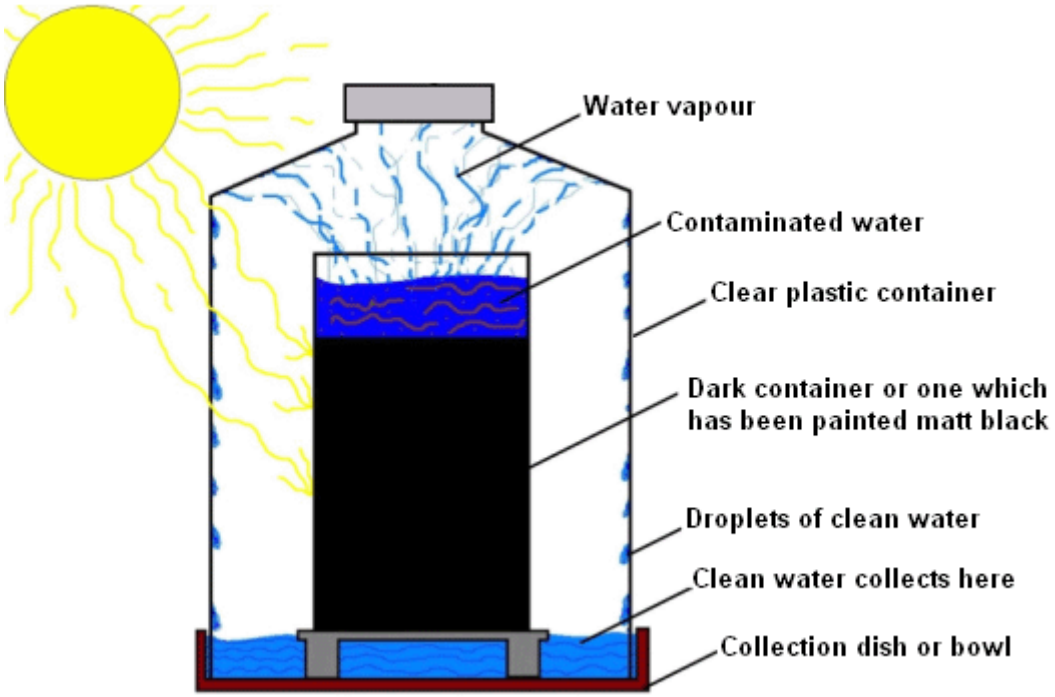
Eliodomestico	Normal solar still
	
5 liters/day	3 liters/day
estimated cost: 50$	average cost: 100$

Eliodomestico is intended to be an open project

attribution - non commercial - share alike

Very simple systems:

It is usually presumed that quite a bit of expenditure and construction is needed to make a solar still which will purify water. That is not necessarily the case. Here is a design which can cost nothing and which can work well in a sunny location:



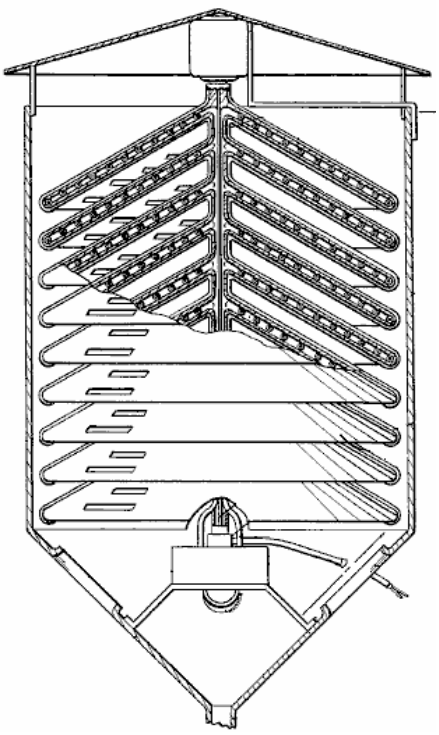
This arrangement could hardly be more simple than it is. An inner container is used and it is made either from a dark material or is painted black, preferably, matt black as a shiny black reflects more of the light falling on it than a matt black surface does. In the diagram above, the black stops well short of the top of the inner container, but that is only to show the liquid inside the container and in fact, the inner container will be dark from top to bottom. Because of this dark colour, the liquid inside the inner container gets hotter than the air outside it (which is hotter than the air outside the outer container, due to the greenhouse effect). There is considerable evaporation from the inner container, but as it can't escape, it forms droplets on the inside of the outer container and these droplets slide down and collect at the bottom of the container, forming a reservoir of distilled water which is safe to drink. The construction of this still can be from things which are already to hand. For example, the bottom can be cut off a transparent plastic drinks bottle and used to cover a dark glass drinks bottle standing in an ordinary bowl, as shown here:

It is, of course, essential for the outside of the glass bottle and the inside of the plastic bottle to be completely clean so that they do not contaminate the purified water.

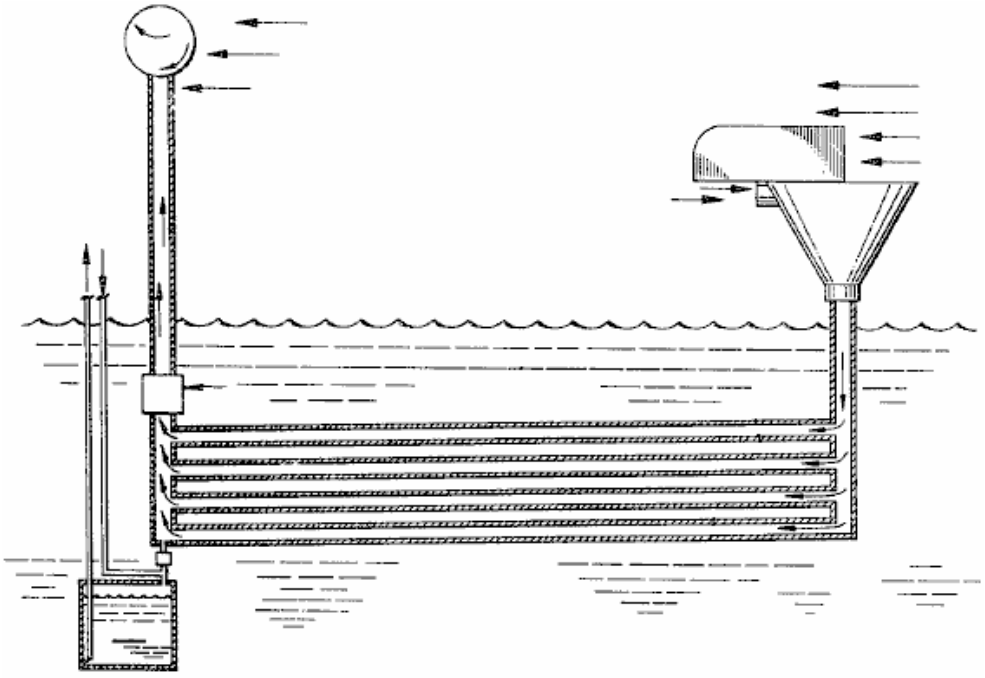
Another variation on this is to use a dark glass jar inside a plastic storage jar, again, placed inside an ordinary bowl as shown here:

Alternatively, a plastic jar with a screw lid can be used upside down and the lid used to replace the bowl. The inner container in this case, happens to be plastic. The capacity of the lid limits the amount of clean water which can be produced at any one time, unless the lid is carefully removed and a larger bowl used to catch the clean water (which means that it would have been easier to use the bowl instead of the lid):

On a larger scale: US patents 2,996,897 (1960) from Elmer Grimes:



and Patent 4,418,549 (1982) from Calice Courneya:



show methods of extracting large amounts of drinking water from moisture in the air:

Another system is using a large Fresnel lens to distil water which is not suitable for drinking. This is possible using the most simple equipment of two glass bottles and a piece of copper tubing. If it is still there, the video at http://www.youtube.com/watch?v=aXjMAItCMl0 shows the method, though I must admit that I would prefer to take the liquid which he drinks and pass it through the system again to improve it's quality further.

Toribio Bellocq's Water Pump.

A serious problem for farmers and individuals is the cost of pumping water up from a borehole or well. While the combined Lever / Pendulum system of Veljko Milkovic described in Chapter 4 can reduce the amount of effort required by a substantial margin, there are other methods which could be useful.

It was originally thought that water could not readily be pumped to a greater height than 32 feet or so unless the pump was located at the bottom of the pipe. Toribio Bellocq demonstrated in 1924 that this is actually not the case and that water can be pumped to any height using a pump mounted at the top of a vertical pipe. He showed a working system to the Patent Office where an 80-foot vertical pipe was used to demonstrate the principle and having proved the point, he was granted US Patents 1,730,336, and 1,730,337, and later, US Patent 1,941,593 in which he describes chamber devices which can enhance the sonic wave operation.

Toribio's system is very straightforward. He places a one-way ball valve at the bottom of the vertical pipe (item **V** in the diagram below). A crank rod is then used to vibrate piston **C** in its pumping cylinder. The pumping cylinder has no valves and the piston stroke is very short. Both the pipe and the pump cylinder are filled with water before the operation is begun.

The rapid movement of the piston creates a pressure wave in the water in the pipe. The pressure wave causes the water pressure inside the pipe to rise and fall rapidly. This altering water pressure at the one-way valve at the bottom of the pipe, causes water to be drawn into the pipe when the pressure is low and the valve prevents the water flowing out again when the pressure rises.

This repeating action causes water to be pumped up the vertical pipe and out through an adjustable valve **R**. When the pumping action is timed correctly, there is an almost continuous flow of water from the pipe.

Toribio quotes an example in his patent, where the vertical pipe has an internal diameter of one inch, placed in a well where the water is twenty metres below the ground level. The valve opening is 30 mm and the sealing ball of the valve has a diameter of about 38 mm and contained in an ordinary cage which allows some 20 mm of vertical movement of the valve.

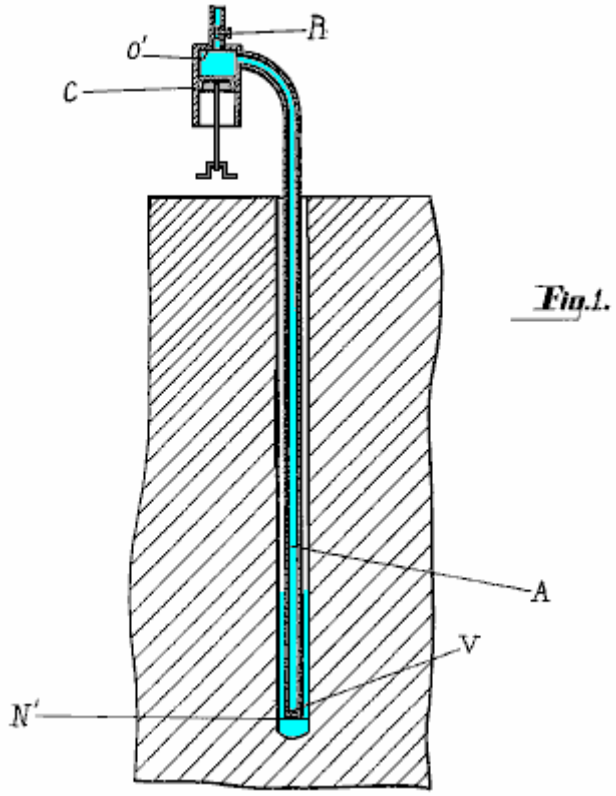


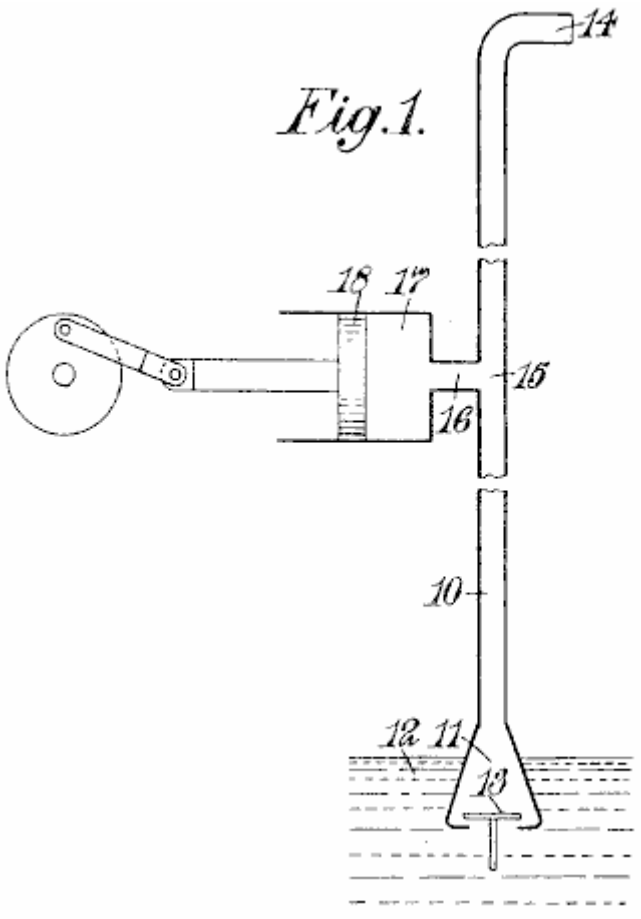
Fig.1.

With this arrangement, the piston at the surface has a diameter of 50 mm and a stroke of 38 mm and is driven by an electric motor at about 360 rpm. The outflow pipe has an internal diameter of half an inch and the valve **R** is used to control the rate of flow out of the system. When the valve is adjusted correctly, a continuous flow is achieved and the flow rate is about 1,000 litres per hour (265 US gallons per hour, or 220

Imperial gallons per hour). It is important that the initial filling of the pipe and fully-open piston avoids getting any air trapped along with the water. The compressor cylinder can be horizontal or vertical. The well can be of any depth and there is no need for the pipe to be straight or vertical. When the system is adjusted correctly, there is little or no wear on the valve at the bottom of the pipe. The liquid pumped does not have to be water.

Richard Dickinson's Water Pump.

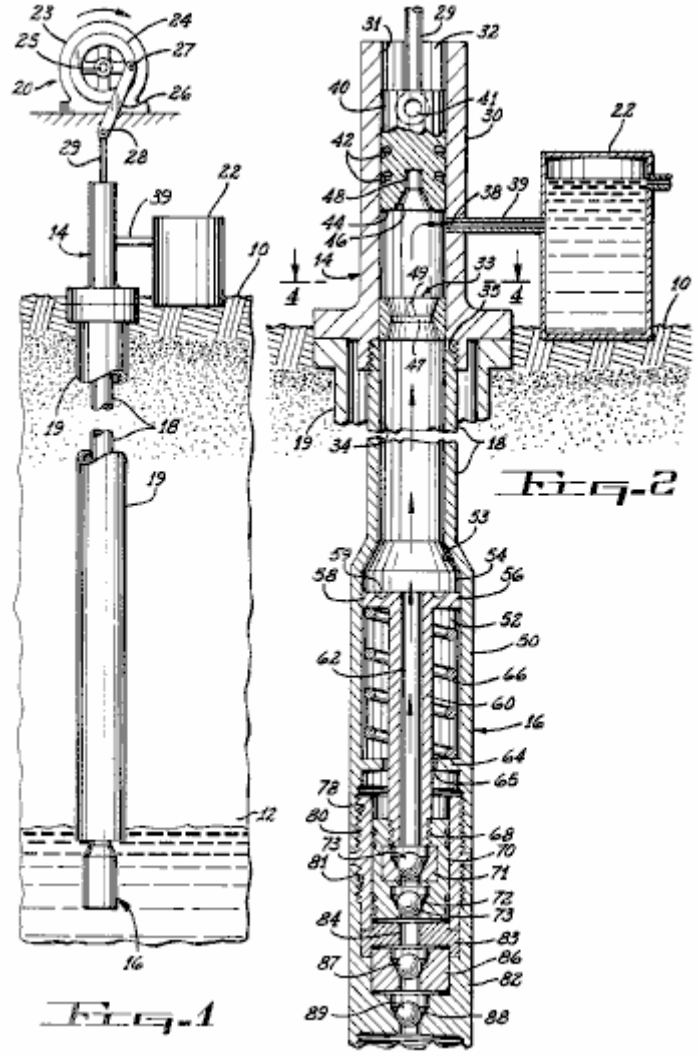
US Patent 2,232,678 of 1937, show a very similar system with a piston being driven in a cylinder without valves, creating a pressure wave in the vertical pipe which has a similar one-way valve at the bottom of the pipe. Interestingly, no mention of Bellocq's patent is made. Dickinson's patent drawing of the system outline is shown here:



Arthur Bentley's Water Pump.

The grandson of the car designer, Arthur Bentley has some 34 patents to his name, one of which (US 4,295,799) is very much like Toribio Bellocq's pump system. Richard Bruner writing in the Calgary Herald newspaper in 1989, tells how a prototype of the Bentley design was tested on a Navajo reservation in Arizona. Driven by four solar panels, a flow rate of 120 US gallons per hour was achieved, (about half that of Bellocq's rate at 20 metres depth), though the depth of the Navajo well was not mentioned. The manufacturers claim that the pump can operate at depths of up to 4,000 feet.

Neither this Bentley patent nor his earlier patent 3,804,557 makes any mention of Bellocq which seems somewhat strange, especially with the marked apparent similarity between the designs. Again, we see here, a piston being used to generate an acoustic wave in the vertical pipe and a series of one-way valves at the bottom of the tube being used to trap the rising column of water and prevent it from flowing out of the bottom of the pipe again. A variation in this patent is the addition of a spring loaded bottom section to the pipe which is alternately compressed and expanded by the sound waves as part of the pumping process as shown in the following diagrams:



The Ram Self-Powered Pump.

In hilly areas, it is frequently necessary to pump water up to locations where it is needed. These locations are usually considerably higher than the source of water. There is a simple device called a “Ram Pump” which is powered by water flow alone and needs no other form of power. In a way, it operates very much like the pumps just described, in that water flowing into a pressure chamber causes fluctuating pressure which with just two valves, and no other moving parts, pumps water to a considerable height.

A Ram Pump can be used if there is a fast-flowing stream of clean water, and more than 50% of the water flow into the pump can be lifted to a higher level. The remainder of the water flows back into the stream at a point lower down. These pumps are readily available commercially and interestingly, they have a COP of infinity as the user does not have to supply any input power and yet substantial pumping power is produced for an unlimited period. As this is a standard Engineering technique, nobody gets upset at the though of ‘perpetual motion’ or ‘free-energy’ even though the pump can go on pumping for years with absolutely no fuel being burnt. This is energy being drawn from the environment in the same way as a self-powered compressed air engine draws energy from the environment, and yet, the compressed air engine is considered to be “unbelievable” while the Ram Pump is accepted without question. Could there be a certain degree of bias being seen here? The power operating the pump comes from the water flowing down hill. The water arrives at this height by falling as rain. The rain gets up there by evaporation caused by water being heated by the sun. So, bottom line, the pumping power comes from the sun.

If a fast-flowing stream is not available but the terrain allows it, then a Ram Pump feeding system can be built. Ideally, there should be a drop of at least two metres (six feet) on the inlet pipe. This creates a fast flow into the pump by feeding it through a steeply sloping intake pipe, like this:

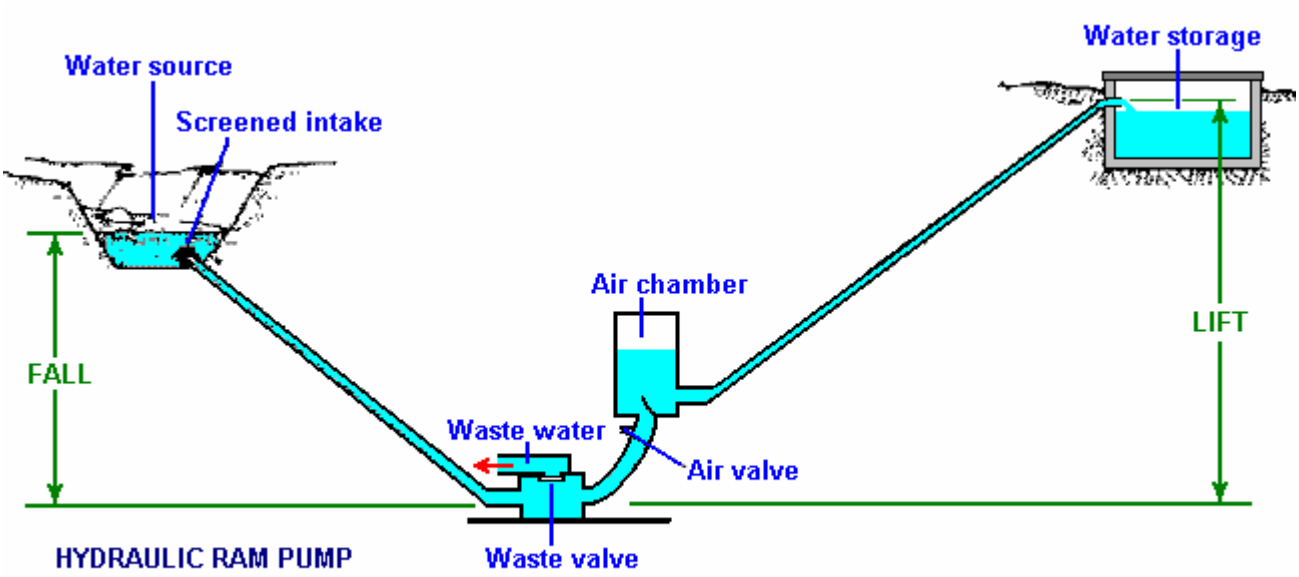


Diagram from: www.thefarm.org/charities/i4at/lib2/hydrpump.htm

The performance of a Ram Pump is impressive even though it has only two moving parts. With an input fall of just four metres and a small flow rate of just three litres per minute, a Ram Pump can deliver 69 litres per day to a massive height of 100 metres vertically above the pump. Or, 159 lpd to 60 metres above the pump, or 258 lpd to a height of 40 metres above the pump. This is impressive for such a simple device.

It operates by the water rushing into the air chamber. This raises the pressure until the valve at the base of the chamber slams closed. The increased pressure in the chamber pushes water out of the delivery outlet, lowering the pressure again. While this is happening, the closed valve causes a 'water-hammer' wave of reverse pressure which pushes excess water out of the 'waste' pipe and pushes water back up the intake pipe. When the pressure wave in the intake pipe dissipates, the water rushes back down the pipe, pushing the valve at the base of the air chamber, open again, to repeat the cycle. This oscillating pressure wave causes the pumping action, very much in the same way as the previous pumps which use a mechanical oscillator pump as no free-flowing water is available to create the oscillation.

Commercial ram pumps have an efficiency of about 66%. The calculation of performance is:

$$D = (S \times F \times E) / L$$

Where:

D = The quantity of water in litres delivered in 24 hours.
S = The quantity of water, in litters per minute, fed to the pump.
F = The height in metres of the water source above the pump intake.
E = The efficiency of the pump (assume 33% for home built units).
L = The height in metres, of the supply outlet above the pump.

Reproduced here by kind permission of US AID 1982 from the web site shown in the diagram above, is a table of values, calculated from the formula above, and assuming the 66% efficiency of a commercial unit. The input flow for these numbers is a tiny 1 litre per minute trickle. This is less than the hydroxy gas rate produced by the Smack's Booster shown in Chapter 10, so in practice, you will be multiplying the numbers in this table by a realistic number of inflow litres per minute.

Litres Delivered over 24 Hours

Working Fall (m)	Lift - Vertical Height to which Water is Raised Above the Ram (m)											
	5	7.5	10	15	20	30	40	50	60	80	100	125
1.0	144	77	65	33	29	19.5	12.5					
1.5		135	96.5	70	54	36	19	15				
2.0		220	156	105	79	53	33	25	19.5	12.5		
2.5		280	200	125	100	66	40.5	32.5	24	15.5	12	
3.0			260	180	130	87	65	51	40	27	17.5	12
3.5				215	150	100	75	60	46	31.5	20	14
4.0				255	173	115	86	69	53	36	23	16
5.0				310	236	155	118	94	71.5	50	36	23
6.0					282	185	140	112	93.5	64.5	47.5	34.5
7.0						216	163	130	109	82	60	48
8.0							187	149	125	94	69	55
9.0							212	168	140	105	84	62
10.0							245	187	156	117	93	69
12.0							295	225	187	140	113	83
14.0								265	218	167	132	97
16.0									250	187	150	110
18.0									280	210	169	124
20.0										237	188	140

Discouraging Mosquitoes and Small Flies

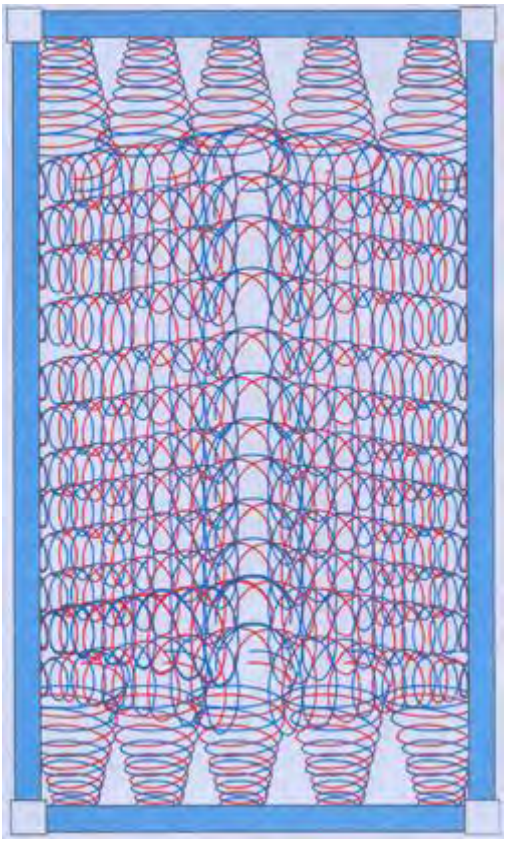
Maurice Cottrell (whose work is in chapter 11) presents what is a quite remarkable effect. He says:

What it is, is a church door in Ireland, surrounded with a white plastic pipe filled with water. When in Palenque in Mexico Maurice found that the local people discouraged flies, particularly mosquitoes and very

small flies, by hanging polythene bags filled with water, from the lintels of their open doors. They do not know why, but doing that discourages flies from entering the doorway.

Maurice considered the effect and his opinion is that small flies can be injured by raindrops and so they have learned to avoid them. A falling raindrop has an electrogravitic effect due to the hydrogen atoms in the water, and flies which can see in the ultraviolet spectrum, can detect that effect around water and their instinct makes them avoid water – whether falling as droplets or not. Large bluebottle flies do not seem to be affected, but there is a definite effect with smaller flies.

If a doorway is surrounded with a 1-inch (25 mm) diameter plastic pipe filled with water, as seen in the picture above, then Maurice believes that the resulting effect is like this:

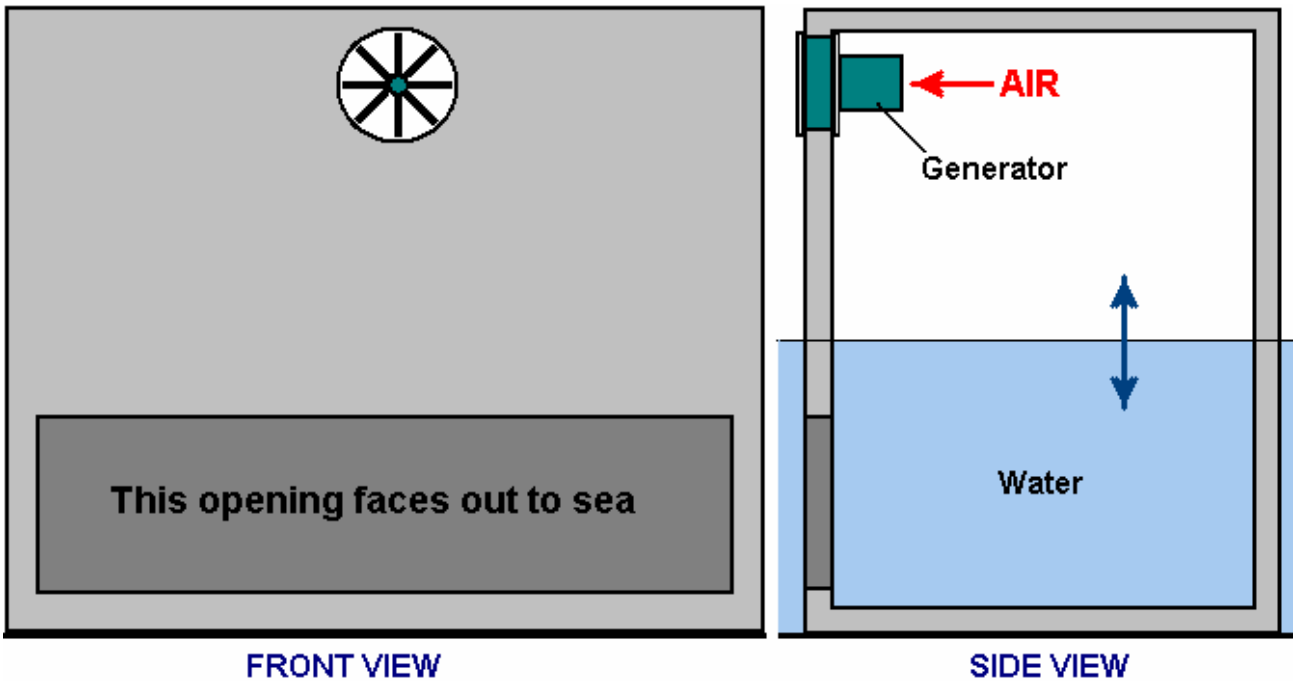


Here, gravity waves spread out from the water and create an effective barrier all the way across the doorway, discouraging the smallest flies, including midges. Obviously, any opening can be surrounded with a water-filled pipe, and not just doorways. While this is not a renewable energy device as such, it is a method which could well be useful where serious diseases are carried by mosquitoes.

Wave Power.

Although not generally thought of as an option for personal use, wave power does have a high potential, although, like wind power and unlike tidal power, not always available. We tend to think of wave power systems as being large scale and very expensive, but that is not always the case. At it's most basic level, most wave power system uses the varying distance between the surface of an ocean or sea and some fixed point on land or the sea bed.

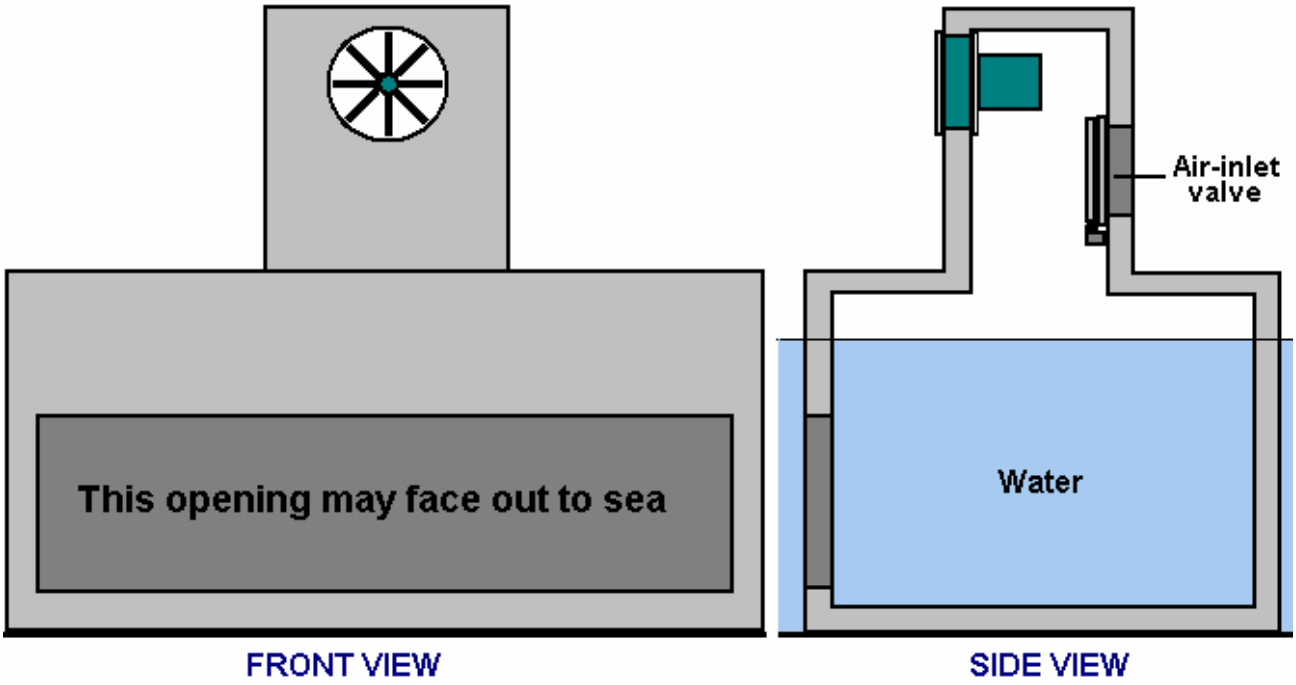
Ideally, there should be a minimum of moving parts. One neat design uses a simple rectangular concrete housing with an electrical generator mounted above sea level. The generator being above sea level is easy to reach for maintenance or replacement and there are no moving parts underwater. It is a very simple design which can be built quite easily. In it's most simple form, it is just a rectangular box with an underwater opening:



Here, a large opening allows the sea to flow into the structure which makes the water level inside the box move up and down with the wave motion outside. As a wave passes by, it compresses the air inside the box and the air is driven out through the generator opening, spinning the generator blades in the same way as a wind-powered generator has it's blades spun by the wind.

Some generators operate well with the air flowing backwards and forwards through the blades, generating electricity no matter which way the blades are spun. With a generator which works better with just one direction of spin, then a large flap valve is installed and it allows air to flow into the structure when the water level is falling but closes immediately the water level starts rising again.

Even though this style of wave power generator is so simple, it works very well in practice, provided that the vertical dimensions are arranged so that the top of the underwater opening is below the lowest neap tide and the bottom of the generator opening is above the highest spring tide. A baffle arrangement can be used to protect the generator from spray and storm debris. There is no need to have the structure full width above the water level:

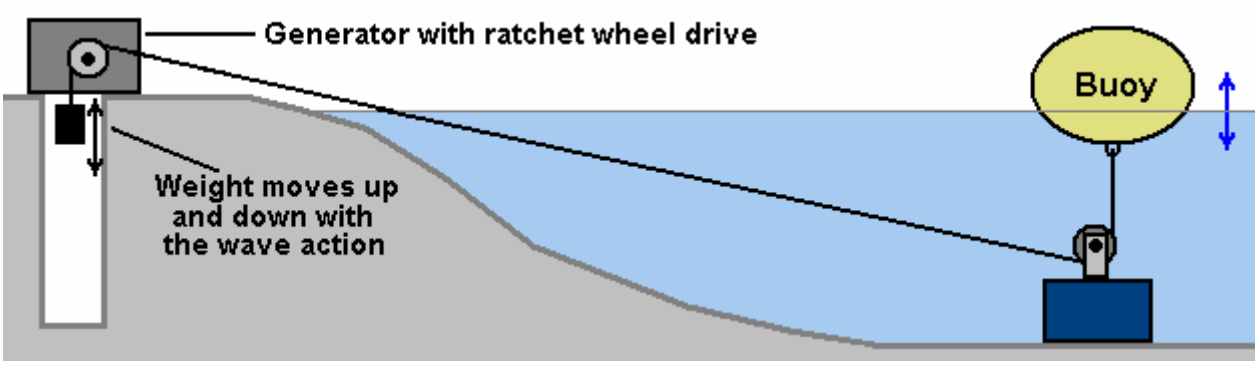


This has the very considerable advantage that the area of the water surface inside the structure is very much larger than the cross-sectional area of the generator housing column and so the air rushes out through the generator much faster than the wave rises. This amplification factor can be increased by increasing the length of the base of the unit, further enlarging the water surface area inside the structure. If wave action is frequently very strong, then it may be preferred to have the undersea opening facing inshore or sideways in order to reduce the amount of material driven into it by very strong surges.

Another fairly simple wave power generator system design which is based on simple principles, was suggested on Stefan Nystrom's website which has been discontinued. It is called the "WaveReaper" system. It operates using a large number of separate buoys. Plastic barrels are suggested as suitable buoys but almost any non-dangerous containers which will not corrode in the sea and which have a considerable internal volume, can be used in this system.

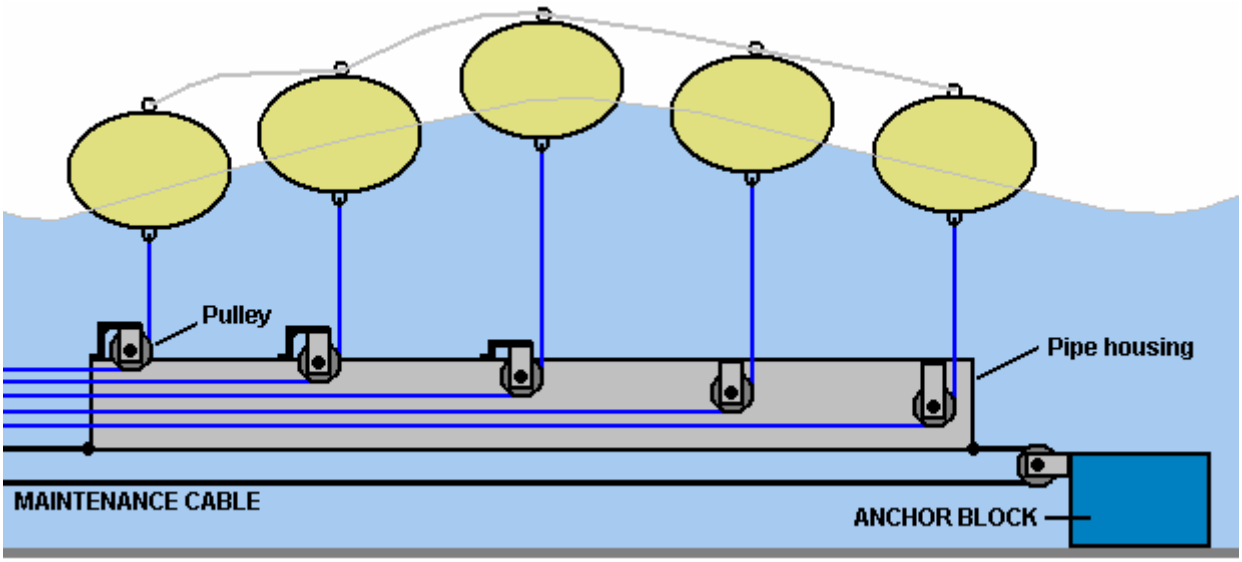
Each basic unit consists of a float, a pulley, a cable and a ratchet drive connection to the shaft of an electrical generator. The power provided by the movement of the buoy can be very substantial as sea water weighs a considerable amount. The connecting cable is kept taught by a heavy weight, and the cable runs over a pulley which is mounted on a shaft which connects to the shaft of the generator. Bicycle parts are suggested for this section of the drive as they are cheap and readily available in most places and they come with a toothed sprocket wheel which already has a suitable ratchet built into it.

The reason for the shaft is that a whole series of buoys are used. These buoys are positioned progressively further and further from the shore so that an incoming wave raises the buoys one after the other in a regular sequence. This means that while one buoy is letting its cable run back inshore (pulled by it's inshore weight), one or more of the other buoys will be rising and applying drive to the shaft linked to the generator. This arrangement allows the generator shaft to receive a continuous drive. Having a heavy flywheel on this shaft is an added advantage as it will smooth out the repeated drive strokes provided by the buoys:



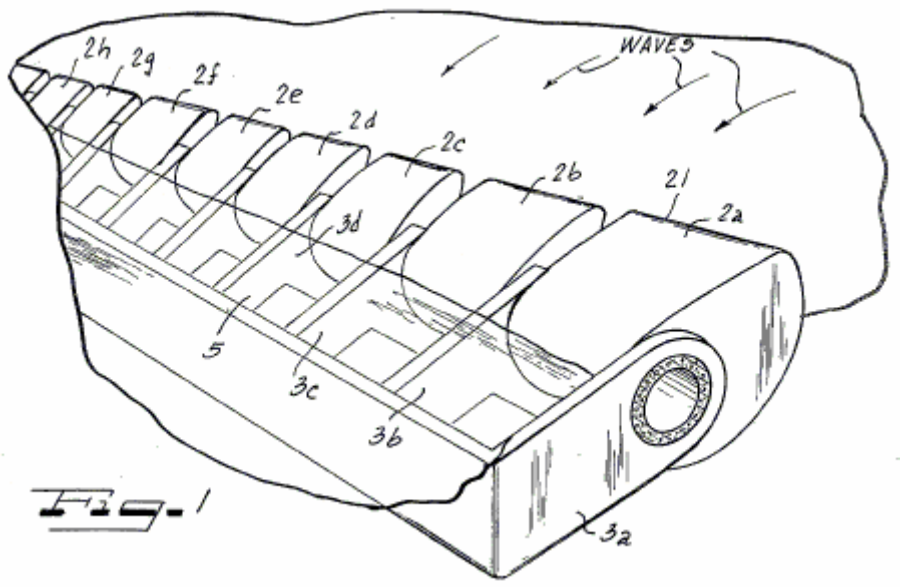
Having a sheaf of moving cables threaded through the sea near the shore is asking for a major tangle with seaweed and all kinds of other drifting material. Very sensibly then, Stefan suggests that the cables be housed in a protecting pipe. Considerable care needs to be taken to make sure that the cables do not rub against anything as the movement is constant and the forces involved are high. Each cable needs to have it's own space keeping it clear of all the other cables and having a pulley mounted at any points where there is a change of direction.

To make maintenance easier, it is also suggested that these protecting pipes are not fixed in position but are themselves on a pulley system so that they can be hauled ashore:



The buoys are also linked together loosely on top with a securing cord so that they always stay in a compact group, though there is little chance of any great sideways movement as the tension in the buoy cables is high. Stefan requests that anyone who constructs his design makes a donation via http://www.o2gruppen.se/ though how that is done is by no means clear to me as there does not appear to be a "Donate" button on that website.

There are many other wave-power devices, some with excellent efficiencies, but most are not generally capable of construction by the average amateur. One example is the "Nodding Duck" design by Stephen Salter of Edinburgh in Scotland, and shown in the US patent 3,928,967 where the wave power is extracted by a raft-like construction with cam-shaped floats. These floats have a rippling movement on the surface of the water and the movement of each section relative to the other sections is used to generate power. This is not exactly a back-yard construction.



Other Systems.

Not included in this eBook, but on the website http://www.free-energy-info.tuks.nl/ there are articles from the highly recommended Home Power website http://www.homepower.com/home/ which are on this general topic. There is a system for producing blocks of ice using sun power alone and no other energy input at all:

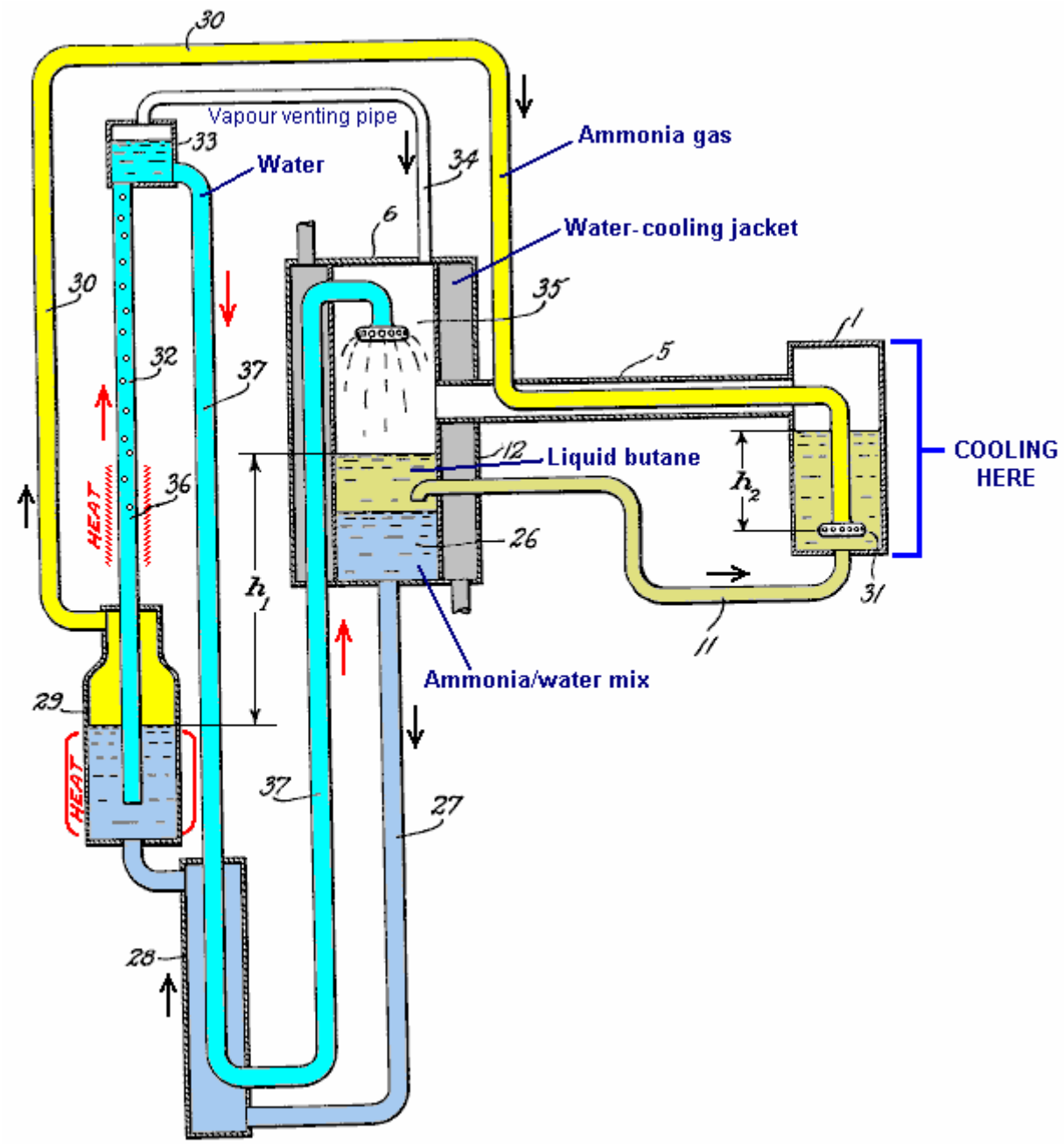
Above: Steven Vanek with his machine which uses solar thermal energy to make ice.

Also, a two-part article on Solar cooling, which concentrates on heat absorption with different colours, the strategic positioning of buildings and vegetation, practical roof overhangs and the like, to lower the temperature inside buildings in very hot locations.

There are articles is on cooking with hydrogen, heating your house with hydrogen and using hydrogen with a barbecue. On the Home Power site there is information on how to use solar power to heat household water and you may find the Google video on how to make your own hot-water solar panel interesting and useful. It makes sense to reduce your essential costs by doing a few simple things which help.

Cooling Using Heat.

Most of our current refrigerators use electricity to drive a compressor to achieve cooling. Here is a patent from Albert Einstein (whom you may have heard of) and Leo Szilard which uses heat to power refrigeration instead of electricity. It is US Patent 1,781,541 titled "Refrigeration" and dated 11th November 1930.



Our invention relates to the art of refrigeration and particularly to an apparatus and method for producing refrigeration where the refrigerant evaporates in the presence an inert gas and more particularly, to the type disclosed in the Von Platen and Munters Patent No. 1,685,764 of 25th September 1928 and our British Patent No. 282,428.

The objects and advantages of our invention will be apparent from the following description considered in conjunction with the accompanying drawing which shows more or less diagrammatically, a preferred embodiment of our invention.

In the drawing, **1** is an evaporator which is normally placed inside the chamber which is to be cooled. A pipe **5** connects the upper part of evaporator **1** to the more intermediate portion of the condenser **6**. Pipe **11** connects with the bottom of the evaporator **1** and extends into the condenser 6, at a level which is below the level of pipe **5**. A cooling water jacket **12**, surrounds the condenser and allows cooling water to flow through it.

Pipe **27** connects the bottom of the condenser **6**, to the lower part of a heat-exchanger jacket **28**. The upper part of jacket **28**, is connected to the lower part of generator **29** which is heated by any suitable method. Pipe **30** connects the upper part of generator **29** to a point near the bottom of evaporator **1** where it

terminates in a distributor head **31**. Pipe **30** runs inside pipe **5** so that there is a heat exchange between the fluids in those two pipes.

Pipe **32** runs upwards from the lower part of generator **29** to connect with a container **33** which is positioned at a level which is above that of condenser **6**. A source of heat **36**, is applied to pipe **32** at a point above generator **29**. Pipe **37** runs down from container **33**, passing through the heat-exchanger jacket **28** and then on up to the top of condenser **6** where it terminates in a distributor head **35**. Pipe **37** runs inside the cooling water jacket **12** so that the fluid passing through it will be cooled as it flows. A venting pipe **34** connects the upper part of container **33** with the upper part of condenser **6**.

The operation of the apparatus is as follows:

A suitable refrigerant, for example, butane in liquid form, is held inside the evaporator **1**. An inert gas, such as ammonia, is introduced into evaporator **1** through pipe **30** and it's distributor head **31**. The refrigerant evaporates in the evaporator in the presence of the inert gas due to the fact that the partial pressure of the refrigerant is reduced thereby and the resulting gaseous mixture passes through pipe **5** and into condenser **6**. Here, the mixture comes into intimate contact with an absorption liquid, for example, water, which is fed into the condenser through pipe **37** and it's distributor head **35**. The ammonia gas is very soluble in water but the butane is quite insoluble, so the ammonia is absorbed into the water freeing the butane from the gaseous mixture. Thus, the butane assumes substantially the entire pressure inside the condenser, and that pressure is sufficiently high to cause its liquefaction at the temperature maintained by the cooling water.

The specific gravity of liquid butane is less than that of the solution of ammonia in water and so stratification of the two liquids occurs with the liquid butane floating on top of the ammonia solution **26**. The liquid butane passes from condenser **6**, through pipe **11**, and returns to evaporator **1**, where it is again evaporated and the cycle repeated.

Gravity causes the ammonia solution to flow from condenser **6** through pipe **27** and heat-exchanger jacket **28**, into generator **29**. Here, the application of heat causes the ammonia to be expelled from the solution in the form of a gas, which then passes through pipe **30** and distributor head **31**, into evaporator **1**, where it reduces the partial pressure of the butane, causing it to evaporate as already described.

Water, containing very little ammonia in solution, passes from generator **29** through pipe **32** where it is further heated by the source of heat **36**. This heating causes the formation of vapour in pipe **32** which lifts the liquid through this pipe and into container **33** and on from there under gravity through pipe **37** to condenser **6** and during its flow, this hot, low-concentration liquid is cooled by the heat-exchanger jacket **28**. It is further cooled by the cooling water in jacket **12**, and so reaches a condition where it can rapidly absorb ammonia in the condenser **6**. Vapour entering container **33** through pipe **32**, continues on it's journey to the condenser **6** via the venting pipe **34**.

During the operation of this piece of equipment, the pressure existing in the various components is uniform with the exception of slight differences caused by columns of liquid needed to cause the fluids to flow. The pressure existing in generator **29** must be sufficiently greater than the pressure in the upper part of evaporator **1**, in order to make vapour flow through distributor head **31**. In other words, the pressure difference must be sufficient to overcome the liquid head marked $\mathbf{h_2}$. This excess pressure in the generator is balanced by the pressure created by the column of liquid marked $\mathbf{h_1}$ in the drawing. This means that $\mathbf{h_2}$ must be less than $\mathbf{h_1}$, otherwise there would be no flow.

This patent of Einstein and Szilard seems to indicate that any source of heat such as a fire or a solar oven, should be able to produce cooling using a device which has no moving parts. It would probably be necessary to provide a trickle of water through the water cooling jacket, but apart from that, it looks like a device which could be used effectively by people who live "off the grid" and have little or no access to electricity. All in all, it is an interesting design.

Solar Panels.

A very well-know method of producing electricity from what appears to be a renewable resource are solar panels. It hardly seems worth mentioning these as sales are so well promoted but it is possible that the problems with them are not understood. This is the most expensive way of producing electricity, and when there is a government scheme to cover one third or one half of the purchase price, it is quite likely that the chosen supplier has prices which are so high that it is cheaper to buy the panels direct from an outlet with competitive pricing rather than paying a much higher price with a subsidy.

What is often not realised is that solar panels have a working life span and will need replacement at some later date. The life is quite good, typically ten to fifteen years, but the price is so high that replacement costs must be allowed for. Solar panels are not the only part of your electrical system which will need replacing. Generally speaking, solar panels are used to charge batteries which then power equipment, usually by using an "inverter" which changes a DC battery voltage to a mains voltage AC supply. The big snag is that the frequently used lead-acid batteries generally have a lifespan of about four years. They are expensive, heavy and as they contain lead, they are not easy to dispose of (legally). This is a major recurring cost and disposal problem.

Another point which may not be realised is that the current which can be drawn from a battery without reducing its working life is very limited. The safe discharge rate is called the "C20" rate which just means that the allowable current is one which discharges the battery in a period of 20 hours. For example, if a new battery is rated as being a 100 Amp-hour ("100 AHr") battery, then the safe discharge rate is 100 amps divided by 20, which is 5 amps. At twelve volts, a current of five amps is a power output of sixty watts. A washing machine needs about 2,200 watts for at least some of the time, so to power it (and nothing else at the same time) would take 37 of those 100 AHr batteries. Could you conveniently house 37 large batteries? How much do you think it would cost to replace them every four years?

Another point which is often missed is that batteries are by no means 100% efficient and that means that you have to put more current into them than you can draw out of them afterwards. After some time in use, you may only get between 50% and 80% of the power back from the battery bank.

Please understand that I have nothing against solar panels and actually own three of them myself, but you need to be aware of the practical problems with using them and not imagine that buying them will give you free electricity for ever afterwards.

The next factor to be considered is the fact that the power "rating" of a panel such as "120 watts" is not the amount of power which will be provided by the panel. It will provide that amount of power IF it is positioned exactly square-on to very strong sunlight in a region near the Equator. If the panel is not aimed exactly at the sun, then the panel output will be considerably lower. The sun moves around at a rate of fifteen degrees per hour, so your panel will only be aimed directly at the sun for about four minutes unless you have it mounted on a rotating platform and you push the platform round slowly to face the sun at all times. While that sounds complicated, it is actually not a difficult or expensive thing to arrange although most people don't bother to do that.

Even though most solar panels operate on UV rather than visible light, any minor cloud cover drops the panel output quite noticeably. There is also the little detail that most solar panels only work in sunlight or with reduced output in very bright daylight. This means half the day does not contribute to electrical input and in winter, when the electrical need is greatest, the days are shortest, sunlight scarce and daylight quality very poor unless you live near the equator. The level of sunlight which you will receive drops off steadily the further you get from the equator as the light angle gets lower and lower. This also causes problems in winter with quite low obstacles casting a shadow on the panels and cutting the working time for the panel even further.

The effect of all this is that you will need a far greater claimed panel wattage than your needs seem to require and I would suggest that you probably need between two and three times as many panels as the stated panel wattage would suggest that you need. Considering that chapter 10 shows how to run a generator on water, it would be cheaper to buy and replace a generator on a regular basis than to buy batteries and panels and having to replace them. A generator also runs at night and through the winter.

However, if you are installing solar panels, please be aware that there are considerable differences between panels which have the same apparent rating. A key factor is the voltage produced by the cell under normal working conditions. This makes the difference between a good charging rate under poor conditions and a near-zero charging rate at those times. The good panels have a greater number of cells and the higher voltage produced can make a major difference, so check out the technical specification document on each panel you want to consider and pay attention to the voltage figures and not just the (maximum possible) "wattage" quoted by the manufacturer or sales person. The panels which I picked are made by Kyocera but do your own checking as designs change from year to year. At the present time, solar panels are only about 17% efficient which means that 83% of the power reaching them does not go to make electricity.

Solar panels change size during the day and so need to be mounted in such a way that allows for this expansion and contraction during the day and night. As you don't want to over-charge a battery, it is

generally considered essential to use a control circuit to prevent this happening. In real life, when you use the electricity generated on a daily basis, there is little or no chance of over-charging, but you can decide this for yourself.

One recent development is to use the solar panel(s) to drive a battery-pulsing charging system. This has the major advantage that battery charging does not stop when the light falling on the panel produces a voltage which is too low to charge the battery bank directly. With a pulsing unit, the charging pulse voltage is largely independent of the input voltage and so a good charging rate can be produced in poor lighting conditions, including twilight. This can extend the battery charging period very substantially, especially in winter when there is less light. At this time, one of John Bedini's companies is offering these units for sale.

Assessing solar panels generally from the standpoint of knowing what the snags are, unless you live in a very good location on the planet and have a good deal of spare money to spend, then they are not a great solution, but this is something which you also need to decide for yourself.

However, having said that, a scheme has recently been introduced in the UK and it appears to be a realistic option. Under this scheme, the householder does not buy the solar panels but merely pays to cover the cost of installing them. For example, this house:

has twenty-one panels mounted on the roof and the total cost to the home owner was just £500. The home owner receives a very much reduced electricity cost and expects to recover the installation cost within two years. Even in November at latitude 52 degrees North, these panels are performing well.

Not having to pay for the purchase of the panels themselves makes an enormous difference to the viability of such an installation. I understand that the strategy behind this scheme is to increase the presently tiny percentage of electricity which is produced from renewable sources, by offering people a scheme which actually makes it worthwhile to have an installation. No batteries are used in the installation.

A Solar Panel Desk Lamp

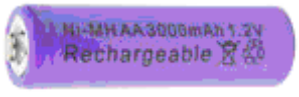
Solar panels can be very useful items in spite of their very low efficiency and high cost. When thinking about solar panels people generally imagine a set of many large solar panels mounted on the roof of a house. The cost of doing that is far too great for most people to consider it. However, at this time, there are one billion people in the world who do not have any electricity at all. It appears that the most useful electricity feature for them would be electric lighting at night. With the components which have become available recently, providing good lighting at realistic cost is now possible.

Small solar panels offered for sale as "10 watt, 12 volt" capacity can now be bought reasonably cheaply. Made in China, these panels can provide a current of just over half an amp These panels which have an aluminium frame are typically 337 x 205 x 18 mm in size and look like this:

Tests which I have run show that a 1000 lux very realistic level of lighting can be provided with a total of just 1.5 watts of electrical power. The best lighting source that I have found is the “G4” style, LED arrays made in China using the recent “5050” chip technology. These are cheap and have a very heavily non-linear light output for current draw, which is a fact which we can use to our advantage. These LED arrays come in “white” or “warm white” versions (my preference is the warm white variety) and they look like this:

With a diameter of 30 mm and pins which are easy to connect to, these are very convenient devices which have an excellent lighting angle of 160 degrees and a light output of 165 lumens for a 1.2 watt electrical input.

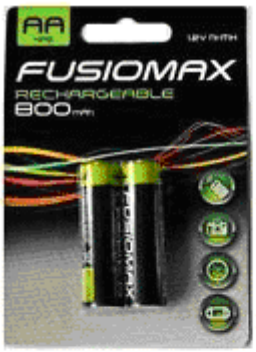
One of the problems with such a unit is the selection of a suitable battery. Lithium batteries are excellent but the cost of a suitable lithium battery is ten times greater than the cost envisaged for the whole unit, effectively excluding lithium batteries. Lead-acid batteries are far too large, too heavy and too expensive for this application. Surprisingly, what appears to be the best choice is the very popular AA size Nickel-Manganese rechargeable battery which is 50 mm long and 14 mm in diameter:



Rated at up to 3 Amp-Hour capacity, they are very low cost, are lightweight and can be placed in a battery box like this:

The battery box can be adapted to hold seven batteries rather than the intended eight batteries, producing a nine volt battery pack with 1.2V batteries. If three of these battery packs are used with the solar panel, then there is no need for over-charging protection as NiMh batteries can deal with overcharging current if it does not exceed 10% of the battery's milliamp-hour rating, and that simplifies the design very considerably.

However, some of these small NiMh batteries do not live up to the maker's claims and so you need to run a load test on any particular make of battery which you may consider using. For example, here are six different types of these batteries tested in groups of four, with a load of about 50 milliamps at five volts. The same load was used to test each of these batteries:



Fusiomax 800

Digimax 2850

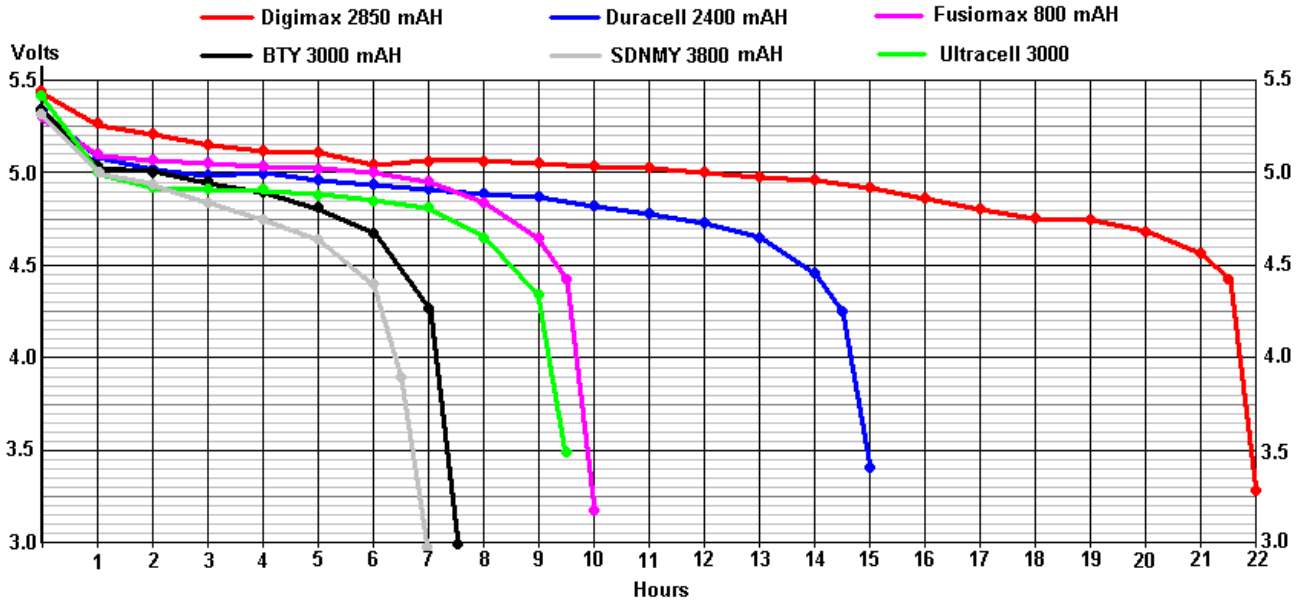
Duracell 2400

SDNMY 3800

BTY 3000

Ultracell 3000

The results were most revealing:



The BTY 3000 batteries do not actually claim on the battery to be 3000 mAHr (although the sellers do) and so, the "3000" could just be a trading name. The tests results for the BTY 3000 were so staggeringly poor that the test was repeated three times with longer recharging time for each test, and the one shown above is the 'best' result. You will notice how far short it falls when compared to the low-cost Fusiomax 800 mAHr batteries. The terrible performance of the BTY 3000 batteries is only exceeded by the incredible "SDNMY 3800 mAHr" batteries which show almost negligible capacity in spite of their amazing claims of 3800 mAHr.

NiMh batteries are 66% efficient. You should only ever charge a 3000 milliamp-hour NiMh battery at 300 milliamps or less and so with a 10-watt solar panel, overcharging is not a problem.

Light meter tests provide some very interesting results for the LED arrays. When using two LED arrays side by side in a light box, the figures for voltage / current draw / light produced using 1.2-volt NiMh batteries were:

9 batteries 11.7V 206 mA 1133 lux: 2.41 watts 470 lux per watt (the manufacturer's intended performance)

8 batteries 10.4V 124 mA 725 lux 1.29 watts 562 lux per watt

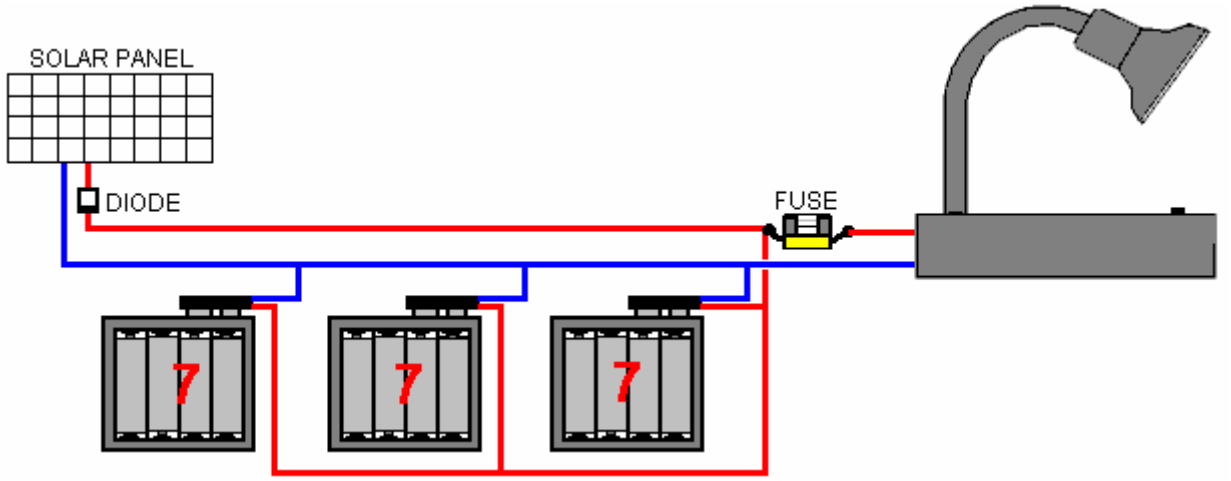
7 batteries 9.1V 66 mA 419 lux 0.60 watts **697** lux per watt (a very realistic performance level)

6 batteries 7.8V 6 mA 43 lux 0.0468 watts 918 lux per watt

This is very revealing information, showing that one of these LED arrays fed with just 33 milliamps can produce very impressive 210 lux lighting at a wide angle of illumination. To put that another way, feeding five LED arrays with 9 volts, generates a very acceptable 1000-lux lighting level for just 165 milliamps which is only 1.5 watts. That is spectacular performance.

Equally impressive is what happens when the battery voltage drops when the battery is nearly fully discharged. The LED performance rises to combat the loss of voltage and even at a ridiculously small 3 milliamps fed into each LED, there is a 21 lux light output from each LED array. The effect is that while the lighting does dim slightly, it does so very gradually in a barely noticeable way. With three sets of genuine high-capacity AA NiMh batteries, we can expect a minimum of eight hours of continuous 1000-lux lighting from our desk lamp. That is a total of twelve watt-hours, and the solar panel feeding 66% efficient batteries at nine volts, is capable of replacing one of those usable watt-hours in twenty minutes. In other words, just two hours forty minutes of good daytime lighting can provide eight hours of 1000-lux lighting every night.

The only moving component in this system is the On/Off switch and the circuit could not be any more simple than this:



All solar panels have a diode to prevent the panel drawing current from the batteries during the hours of darkness and it is not unusual for the panel to be supplied with a diode already connected in place. Personally, I would consider a fuse to be unnecessary but it is standard practice to fit one. The batteries are installed in a base box which supports the solar panel and gives sufficient weight to produce a very stable lamp. The five LED arrays are connected in parallel and fitted into a suitable lamp housing such as this one:

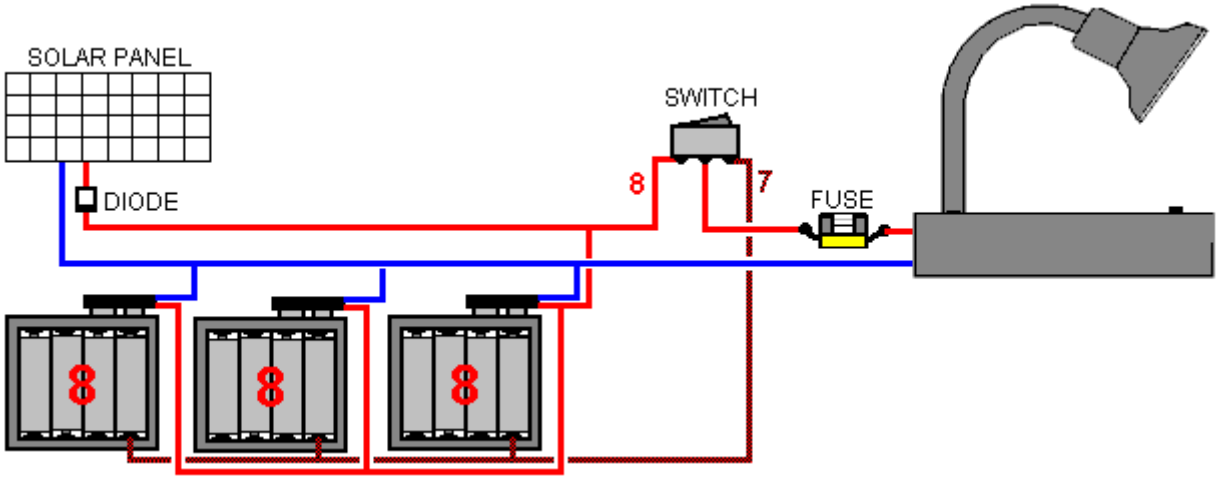
Only the flexible stem, 120 mm diameter lampshade and On/Off switch are used.

While this is an exceptionally simple and robust design, it is actually an affordable and very desirable unit which can provide years of cost-free lighting at a very satisfactory level. The prototype looks like this:

This is, of course, a perfectly ordinary and quite standard type of a solar-powered light. The difference here is that it is a very effective light suited to lighting a desk to a high level all night long. It is mobile and has a wide angle of lighting.

It is also possible to extend the design very slightly, to provide an even longer period of lighting or if preferred, a period of even brighter lighting. This can be done by using eight batteries in each battery holder – which has the advantage that standard battery holders can be used without any need to adapt them to hold just seven batteries.

This has the slight disadvantage that we do not want to supply the extra voltage to the LED arrays because doing that would cause a greater current draw than we want. We can overcome this by using an extra change-over switch and having two connections to each battery holder. The circuit could then become:



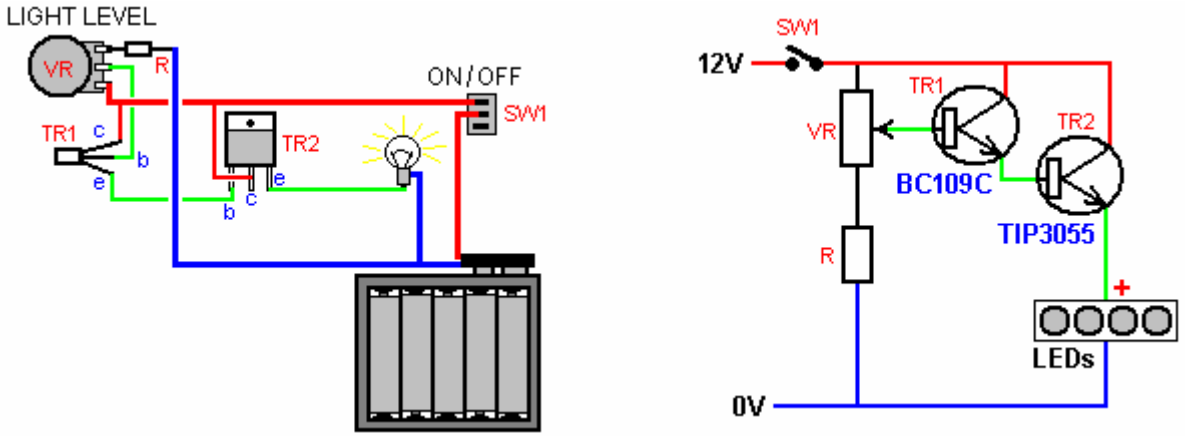
With this arrangement, the lighting unit is fed by either eight batteries or by seven batteries, depending on the position of the change-over switch. When the solar panel is charging the batteries, all eight batteries per holder get charged no matter what position the extra switch is in.

This has the advantage that when the battery voltage starts to drop after a few hours of powering the light, then the switch can be operated, raising the voltage reaching the lamp by the voltage of the extra battery, possibly producing a brightness exceeding the maximum when using just seven batteries in each battery holder. This arrangement has the slight disadvantage that the user could switch in all eight batteries from the beginning, producing a much higher current drain and while that would give a higher lighting level, the overall time is likely to be reduced. Mind you, it is possible that this might suit the user

If this style of operation is chosen, then I suggest that the extra switch is located well away from the On/Off switch so that the user does not get confused as to which switch does which job. Perhaps the second switch might be located near the stem of the lamp support, like this:

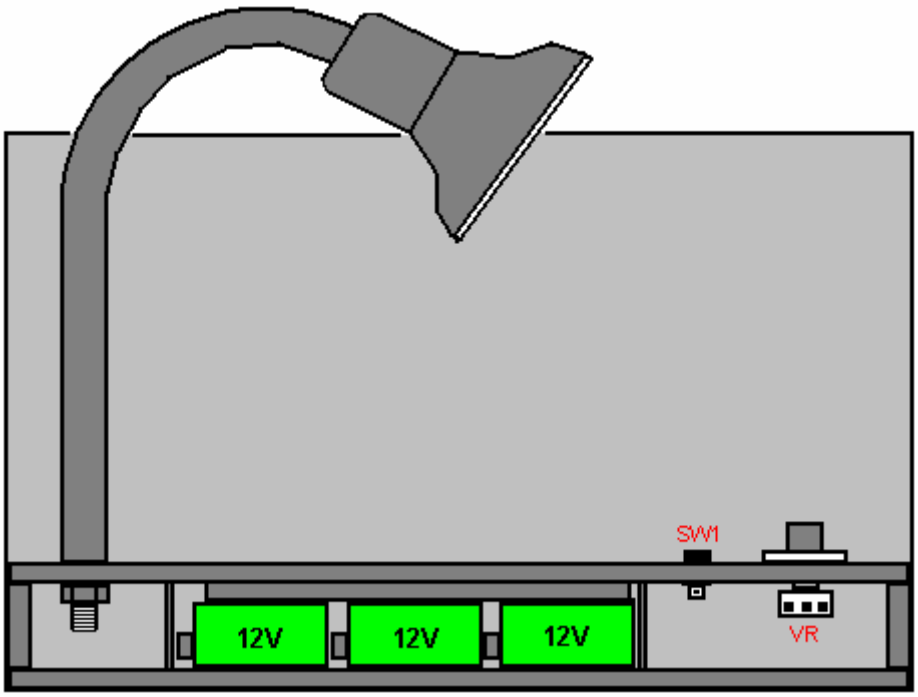
Without knowing much electronics, it is possible to make a more versatile version of this desk lamp. This can be done by providing a fully adjustable lighting level. For this, we use three 10-battery holders, each containing a full set of ten batteries, producing a nominal 12-volts but in reality, about 13.8 volts when fully charged.

We would like the user to be able to adjust the voltage applied to the LED arrays in order to get the very high current-to-light output efficiency or a higher level for a shorter time, if that style of operation is preferred. This can be done quite easily, in a very simple and cheap way, using this arrangement where just one battery is shown to represent all three battery packs wired in parallel to produce a higher battery capacity:



Here, instead of feeding the battery voltage directly to the LED arrays, two cheap and widely available transistors are placed between the battery and the LEDs. These transistors control the voltage applied to the LEDs and that controls the current through the LEDs and the level of light produced. The chosen light level is set using the variable resistor “VR” and the light level can be turned down gradually, all the way to zero, so the unit could also be used as a night light if that was wanted.

The physical layout of the components could be like this:

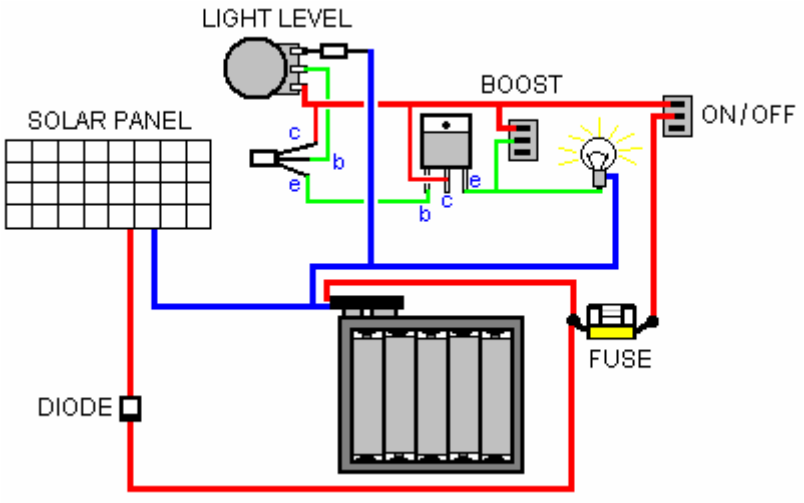


Here, the weight of the solar panel and the three battery packs give the unit stability if the lamp is bent in any direction. With four LED arrays, an excellent level of lighting results, however, I would suggest using five LED arrays as that gives an even wider range of lighting. A big advantage of this arrangement is that if the variable resistor is fitted with a knob like this:

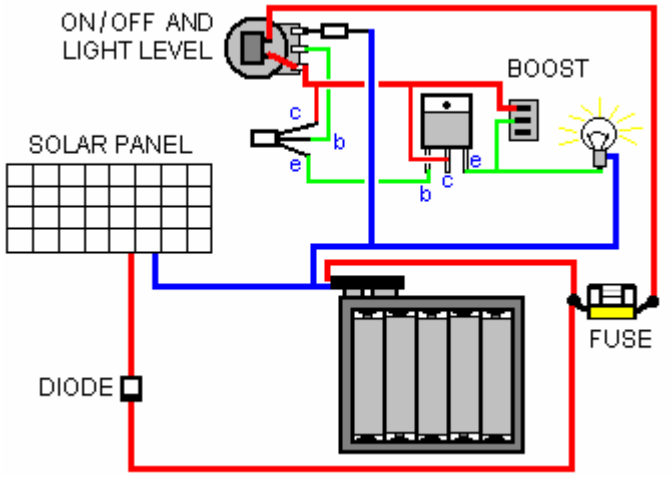
then each user will become familiar with the particular light-level setting which suits the recharging rate. It also allows for short periods of very high level lighting if that ever becomes necessary.

From a practical point of view, when the light is first switched on the battery voltage will almost certainly be over 13 volts. As we want about 9 volts applied to the LED arrays, some four volts need to be dropped off to stop the LEDs drawing excessive current and draining the batteries too quickly. In passing, it might be mentioned that the human eye is very bad at assessing light levels, and so, doubling the LED current (which would more than halve the length of lighting time) does not increase the visual effect by much, and so, controlling the current flow makes a major improvement in the length of time during which the light can be used each night. The lighting period each night is expected to be about eight hours.

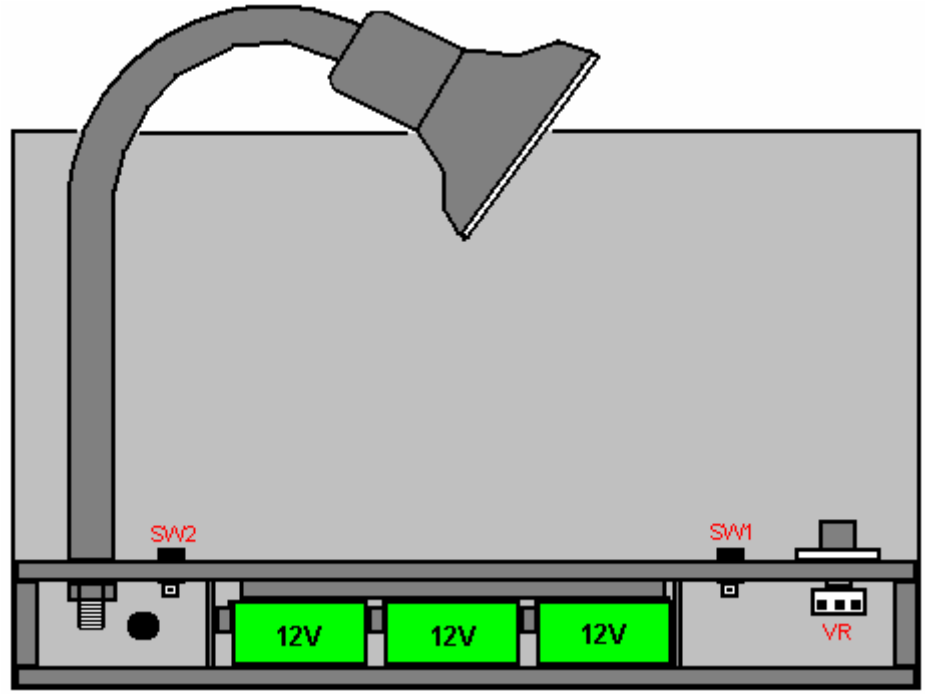
The transistors controlling the light level drop off some 1.4 volts, no matter what the current is. This is not a problem during the normal lighting period. However, if the lighting period is unusually long and the battery voltage is starting to drop, then it is possible to gain that extra 1.4 volts by using another switch to bypass the transistor control. It is wired so that the battery is connected directly to the LEDs in order to extract every last milliamp of current from the batteries. That extra 1.4 volts makes a major difference to a falling light level but it should not be used for any length of time on a fully charged battery (quite apart from the excessive current draw) as it is feeding the LEDs a voltage higher than the LED designer was expecting and while it does produce exceptional lighting, it is being very unkind to the LEDs. The bypass switch would be arranged like this:



Here, the extra switch marked “BOOST” short-circuits across the transistors and so connects the LEDs directly to the battery. The fuse shown is optional but if your wiring is not the best, then it is worth having in case your wiring creates a direct short-circuit across the batteries. A ‘slow-blow’ fuse rated at 1-amp would be a good choice as the normal current flow, even on BOOST with a fully charged battery, will be less than one amp if five LED arrays are used. It is also possible to use a variable resistor which has the On/Off switch built into it:



Also shown in the diagram is the solar panel which is connected to the batteries at all times, even bypassing the fuse. It is connected through a diode such as a 1N4007 so that the panel will not draw current from the batteries during the hours of darkness. To make sure that the two switches are not mixed up by the user, I suggest that the extra switch “SW2” is positioned well away from the On/Off switch:

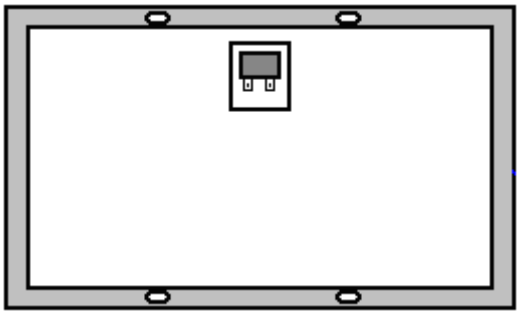


Of course, it is not essential to use a commercial, adjustable lamp such as this. Instead, the LED arrays could be mounted on a home-made strip attached to the back of the solar panel plate and possibly, angled slightly downwards if a desk or table is the main place to be lit. In either case, it is advisable to put a piece of frosted plastic across the LEDs as they can be lit so brightly that they can be difficult to look at. Also, just because this is intended primarily as a desk lamp, there is nothing to stop the user tipping the light upwards to light the whole room. Actually, the room will probably be fairly well lit even when the main lighting area is a desk or table even if that is located at one end of the room.

The construction sequence might be:

1. Lay the solar panel face up on the corner of a sheet of some suitable material such as 6 mm Medium Density Fibreboard ("MDF"). Run a pencil around the edge of the solar panel and cut the sheet along the pencil line to produce a sheet which can be attached to the back of the solar panel. Drill a 6 mm diameter hole about 30 mm in from one corner, that is 30 mm in and 30 mm up so that the hole will be clear of the sides and corner strips when they are attached later on. If it is your intention to paint the desk lamp, then painting this backing sheet before attaching it to the solar panel makes it much easier to do neatly.

2. Lay the solar panel face down on a soft surface and slide the cover off the electrical connection points. It should look like this:

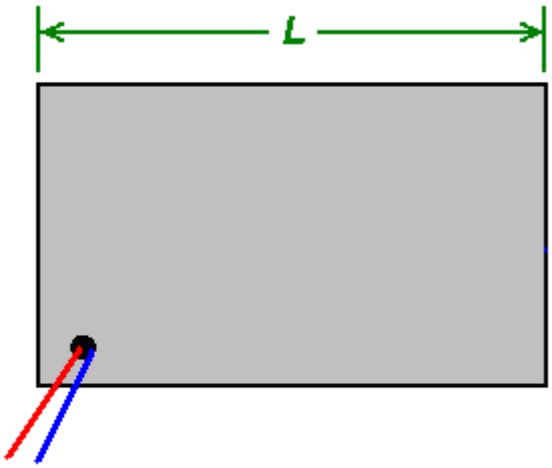


Under the cover, there will be two connection terminals, one marked with a Plus symbol. Some of these panels come with wires already connected and a diode already soldered in place. If there is a diode already there and you are going to remove the cover permanently, it is advisable to embed the diode region in epoxy resin as the diode leads are unduly flexible and could break easily.

If there are not wires already attached, then you need to solder a wire to each of these two terminals. If you do not know how to solder, and don't have a friend who can, then you can learn quite easily. On the web there are several videos showing how to do it. The connecting wire used should not be solid core but should have several tiny strands of wire inside the insulation. The wire should be able to carry at least 1-amp and ideally two or three amps. The thicker the wire, the better electrically but the more difficult it is to solder. The convention is to have a red wire connected to the Plus and a black wire connected to the Minus of the panel, but so long as you know which wire is which, you can use any colour of wire that is convenient. Even if you forget which is which, a voltmeter will tell you which is the Plus when light shines on the panel. These wires need to be about 600 mm. (2 feet) long.

3. Pass the two wires through the hole drilled in the solar panel backing sheet and with the hole positioned at the bottom left. Attach the sheet securely to the back of the solar panel using a good quality adhesive. I prefer to use Impact Evostick or epoxy resin for this, but each constructor will have his own preferred adhesive.

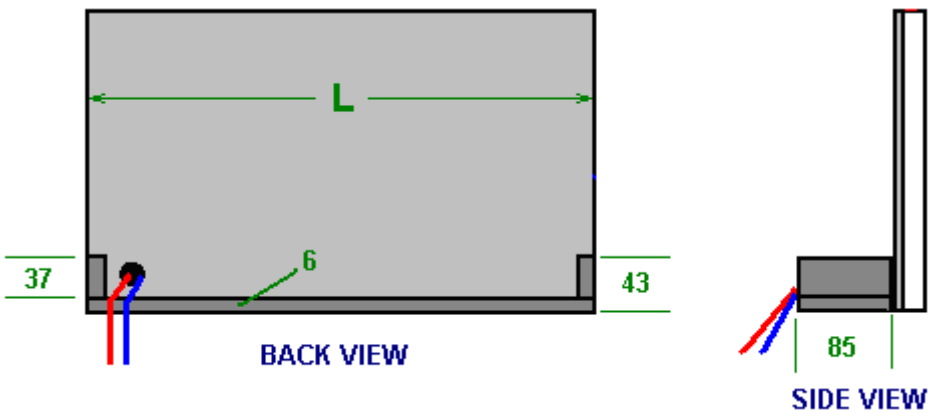
   You don't need the plastic cover which can slide over the terminals and it can be discarded. It is far more important to make good electrical connections to the terminals rather than bothering about keeping under the plastic cover which usually has very, very little clearance and which can get in the way of the backing sheet. The result should look like this:



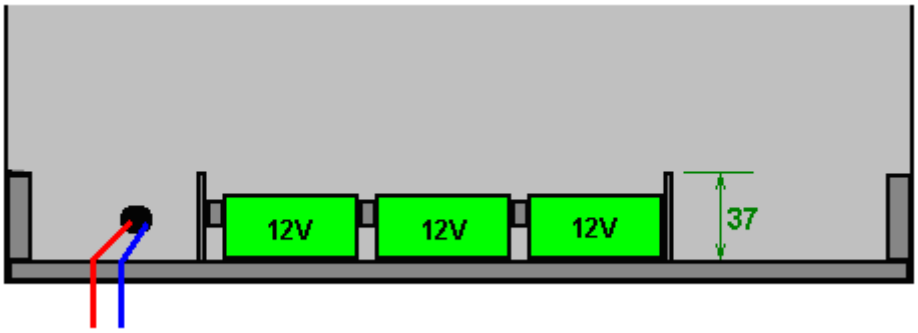
4. We now need to construct the housing for the batteries and other components. For convenience of description, I am going to assume that your sheet material is 6 mm thick. In America, that would be a thickness of one quarter of an inch. Cut a strip of the sheet material "**L**" long ('L' being the length of your particular solar panel). The width of the strip should be width of the battery box (typically 79 mm) plus one thickness of the material, which, with 6 mm thick material is likely to be 85 mm.

5. To make the ends of the box, cut two small strips with a length of that 85 mm (or whatever) and a width of 7 mm greater than the depth of your battery boxes, which are normally 30 mm deep, making the strip typically, 85 mm x 37 mm.

6. Epoxy these three strips to the bottom of the panel backing sheet to form an 85 mm wide, shallow channel like this:

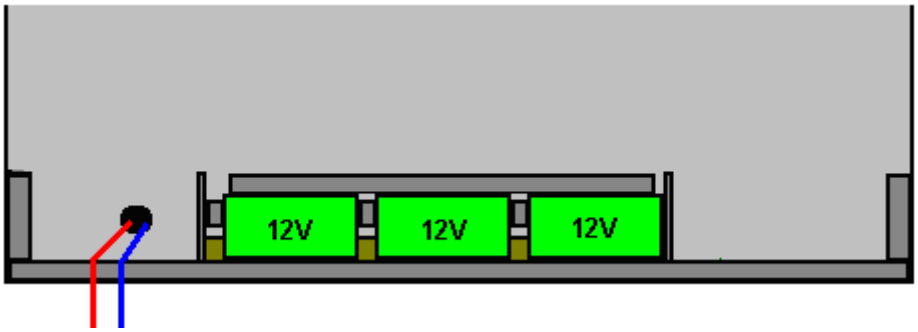


7. Using the same sheet material (or thinner material if convenient), cut two 70 mm x 37 mm strips, and epoxy them one on each side of the three battery boxes. The battery boxes need to have the batteries inserted and the connectors attached to them to ensure that the spacing is right. These dividers connect to the solar panel backing sheet, leaving a gap on the near side as wires have to run from left to right past those dividers. The result should be like this:

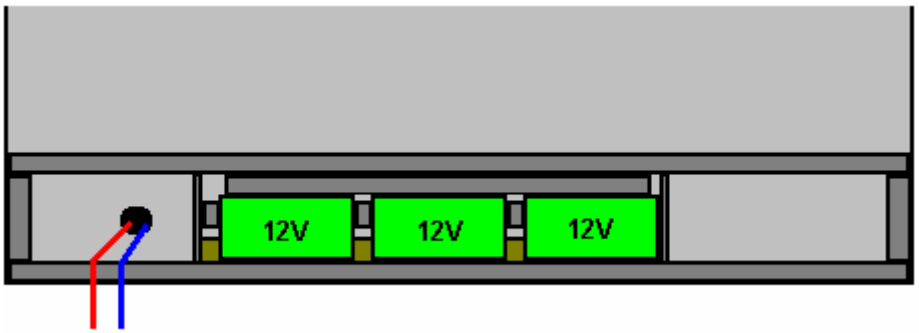


It is important that the batteries are secured so they cannot move at all. As shown above, they can still swivel through an angle due to the gap created by the battery connecting clips. To prevent this, three strips of timber or other suitable material should be attached very firmly to the base to hold the batteries rigidly in place, as shown below.

8. Next, cut a strip of your 6 mm. thick sheet, 65 mm wide and slightly less than the gap between your two dividers. This strip is to hold the batteries securely in place and still allow room for wires to be run past the batteries. The strip will **not** be attached to the solar panel sheet although it will be positioned very close to it. Check that it fits in place. The gap will be on the near side to make the final wiring easier:



9. The next step is to cut the top strip which will hold the lamp, switches and variable resistor. While the lamp fitting is quite light, the length of the arm is so very much greater than the diameter of the screw thread which attaches it to the top plate that any load applied to the lamp arm when adjusting its position, generates many times that stress in the area of locking nut holding it in place. Because of that stress in such a small area, it is suggested that the top plate be particularly strong. A strip of laminate flooring material is very tough and being only 7 mm thick, it is not difficult to work with, so cut a strip to run across between the two end pieces like this:

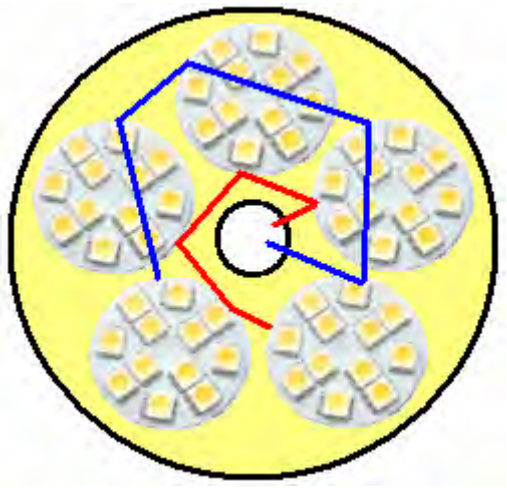


The objective here is to enclose the batteries so that they cannot move, nor can any individual AA cell move out of position. The laminate flooring strip will be used face down as the underside is normally a conservative and attractive matt finish which is particularly suited to this project.

10. It is much easier to work on the laminate strip before it is attached in place, so we will drill the holes for the lamp column and the variable resistor, and cut out the rectangular holes needed for mounting the two rocker switches. In the UK, the rocker switches and variable resistor look like this:

The switches need a rectangular hole of 13 x 19 mm, while the variable resistor needs a 10 mm diameter circular hole and the lamp shown also needs a 10 mm diameter hole. Hold the laminate strip in place and mark the locations carefully, making sure that each item will not foul anything underneath before cutting out the various apertures. Slide the knob on to the shaft of the variable resistor and note the length of shaft inside the knob. Cut the shaft, leaving that same length of shaft remaining. Attach the variable resistor and switches to the laminate strip and slide the knob on to the remaining variable resistor shaft. The skirt of the knob should now be close to the surface of the laminate strip and a V-shaped piece of material can be glued in place to give a reading point for the scale on the knob.

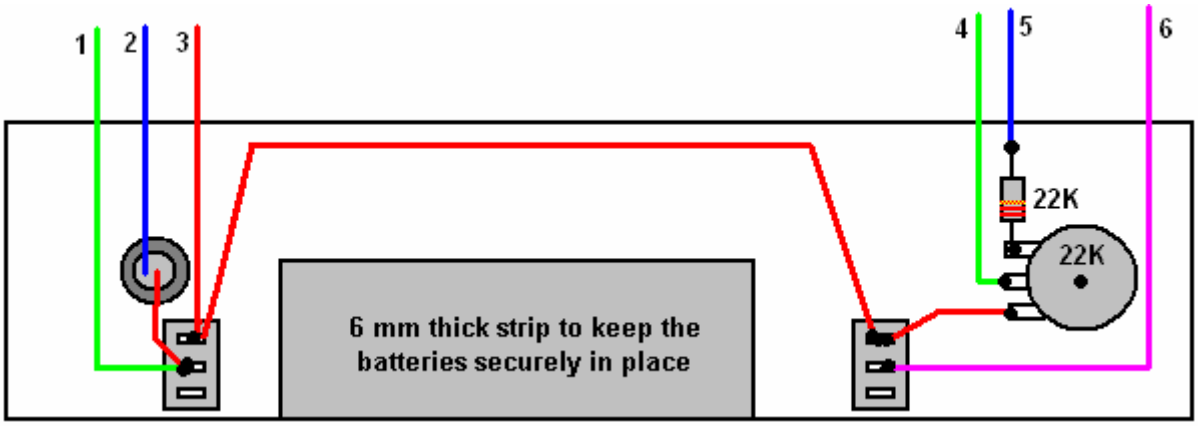
11. If a commercial lamp is being used, then it needs to be taken apart now and prepared for this project. The base is removed, the bulb holder is removed and two wires are fed through the remaining shaft so that the LED arrays can be fitted. A circular disc of any kind of rigid material is cut, the diameter being slightly less than the diameter of the mouth of the lamp. Four or five LED arrays (depending on your choice of numbers) are glued to the disc and wired up in parallel with all of the plus wires connected together and to one of the wires feeding through the shaft of the lamp, and all of the minus wires connected together and attached to the other wire passing through the column of the lamp:



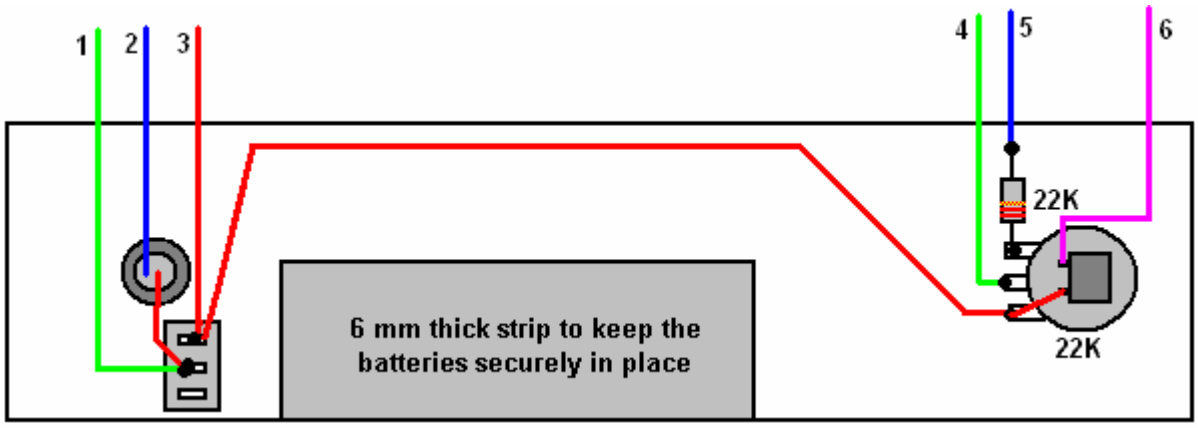
This disc is then eased through the mouth of the lamp shade where it sits about 10 mm below the rim of the shade due to the taper of the shade. Position the disc so that it is square on to the rim of the shade and glue it in position. If frosted plastic is to be used, then mark the sheet around the rim of the shade and cut out the resulting circle, drill some ventilation holes in it although the LED arrays always run cold, and glue the frosted plastic disc to the rim of the shade.

12. Connect the wires to the switches and variable resistor mounted on the laminate strip, and then position the lamp in its hole and fix it in place by tightening one nut. There are two nuts but I have found the locking nut to be less than adequate, so epoxy the tightened nut to the remainder of the screw thread as that keeps the nut securely in place and yet allows the lamp to be rotated quite easily.

    To make construction easier, before placing this laminate strip in place, if a separate On/Off switch is being used, then make the connections shown here:

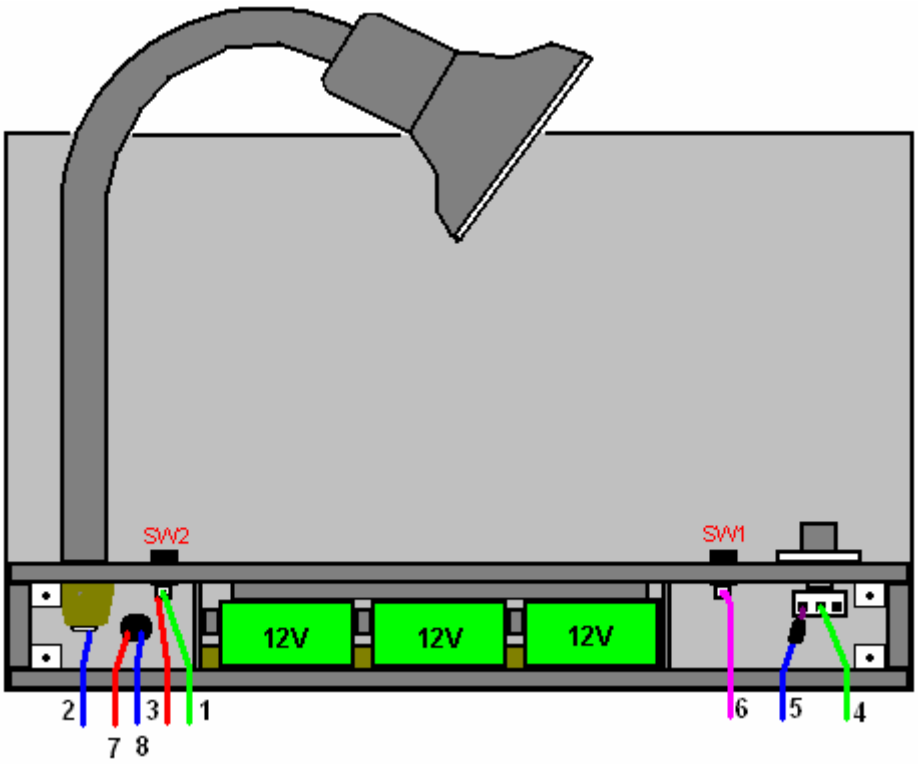


If a variable resistor with an integral On/Off switch is being used, the these connections become:



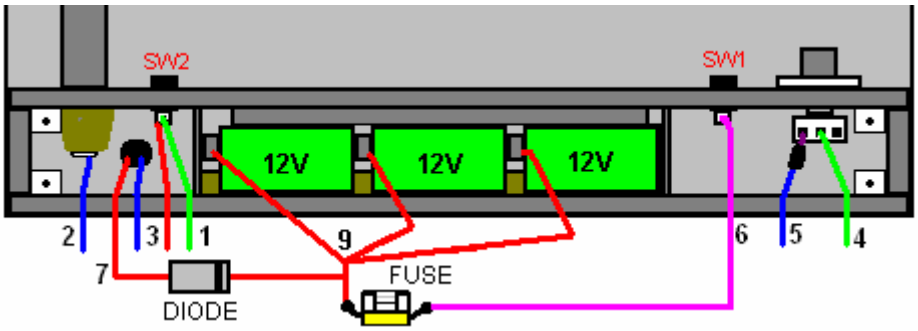
The 22K fixed resistor (colour bands red/red/orange) is connected to one of the outer tags of the 22K variable resistor and wire “5” is connected to the other end. I suggest that all wires coming off this panel are at least 200 mm (8-inches) long. A separate wire “4” is connected to the centre tag of the variable resistor, and the third tag is connected to the On/Off switch as shown, and when making the connection to the switch, also connect another wires to that same tag. The extra wire is run across and connected to the top tag of the “boost” switch along with another wire “3”. The wire connecting the switches is run close to the outer edge of the laminate strip so that it will clear the battery packs, as there is very little clearance past the batteries in order to hold them tightly and prevent them moving. The negative wire of the LED arrays is connected to the centre tag of the “boost” switch and another wire “1” is taken out from that central tag. When connected up as shown, there will be six wires coming out from the laminate strip and I will refer to these wires by the numbers shown above, so it might be worthwhile actually tagging those wires temporarily with those numbers.

13. If inclined, apply some glue to the variable resistor and switches (on the underside of the laminate strip) but do **not** epoxy the strip in place until the final connections have been made and the circuit confirmed to be working, including the solar panel charging the batteries.

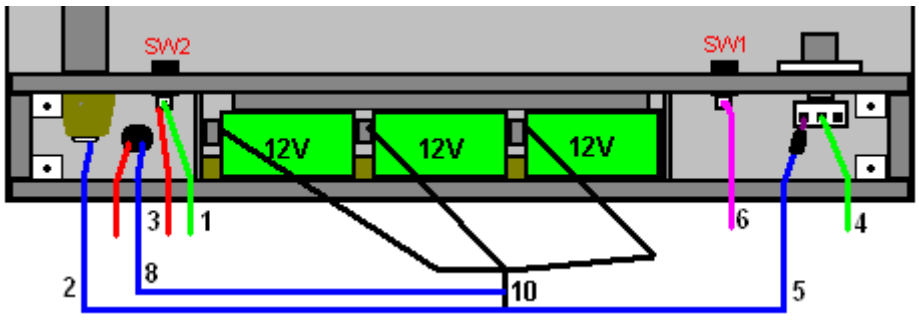


14.The final connections which need to be made are:

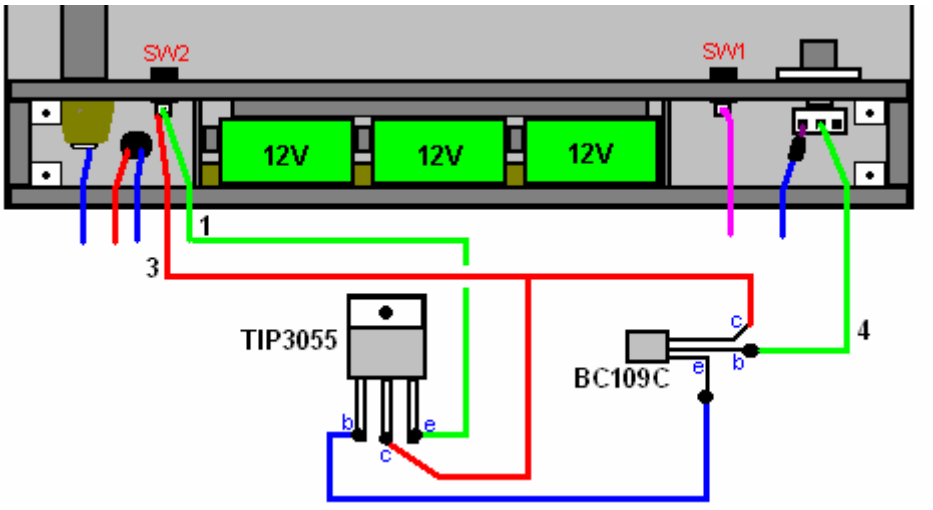
a. The non-striped end of the diode is connected to the Plus wire "**7**" coming from the solar panel and the Plus wire of the three sets of batteries are connected together "**9**" along with the striped end of the diode and one end of the fuse (if no fuse is being used, then wire "**6**" is connected directly instead of to the other end of the fuse). A fuse rated at 1-amp would be suitable as the working current should only be one sixth of that:

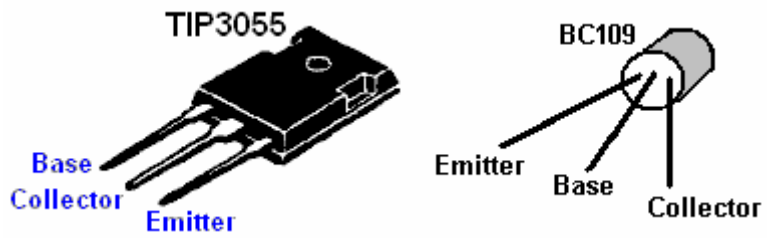


b. Remove the insulation from the Minus wires coming from all three battery packs and twist them together ("**10**"). Connect wires "**2**", "**5**" and "**8**" to this collection of Minus wires and solder them together make a solid joint, and insulate the joint with duct tape, epoxy, or any other robust form of insulation:



c. To complete the wiring, the three remaining wires are connected to the two transistors like this:





Please be aware that transistors are heat sensitive and so soldering transistor leads needs to be completed quickly and where possible, the transistor lead held in a pair of pliers while the solder joint is being made and when it remains hot, as the pliers diverts the heat away from the transistor.

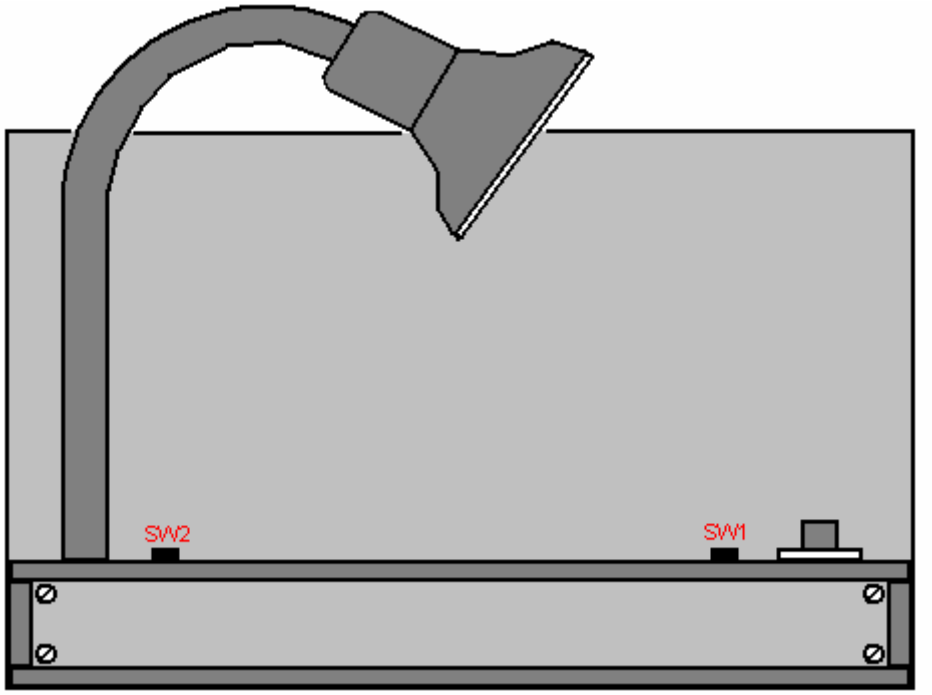
The desk lamp is now ready to be tested.

15. If the light goes on and off when the On/Off switch is operated and the light intensity alters when the variable resistor knob is turned, and if the light level increases when the “boost” switch is closed, then all is well and the laminate strip can be epoxied into its final position. If the desk lamp does not operate as it should, then follow the test procedures below, until the desk lamp is fully functional and then continue with step 16.

16. Epoxy the laminate strip in position and glue four short lengths of timber in the corners. These lengths of timber should reinforce the joint between the four pieces which form the housing and the solar panel backing sheet, and they should stop off 6 mm short of the outside edge so that they both support the final piece as well as holding it in place.

17. Connect a voltmeter across the LEDs and rotate the knob of the variable resistor to get 9V across the LEDs and note or mark the reading on the scale attached to the knob. This is the initial setting for the knob when starting with a fully charged battery.

18. Finally, insulate the transistor circuitry with tape, plastic or whatever method suits you and secure it in place with a little adhesive to make sure that the wiring does not break if the unit is moved a lot. Then, a strip is cut to close the box and screwed at the corners using woodscrews or bolts:



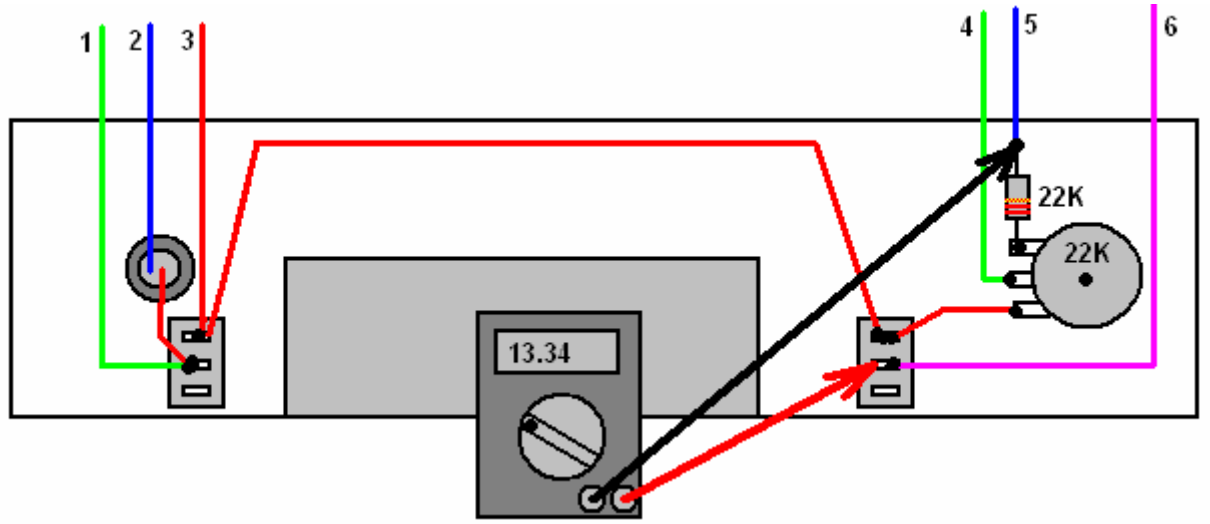
The gap between the two rocker switches provides an easy gripping point for carrying the unit from place to place.

While this solar desk lamp can be built this way at quite reasonable cost even with one-off prices on individual prices, a really major improvement is achieved if they are being manufactured, especially if made in China. Firstly, all the effort of making the housing disappears as a cheap custom plastic case does away with all that time and effort as the solar panel has its wires connected, fed through the existing hole in the case and then the panel is just pushed into place. The battery packs are then assembled and pushed into place, the circuit connections made and the case clicked closed on its own or held in place by self-tapping screws. The component prices drop enormously with bulk purchasing, reducing the not unreasonable one-off price by a major factor, especially since wages in China are very low by our standards. The result is an attractive, useful product with a very large market and a unit price capable of taking full advantage of that large sales base.

Test procedures:

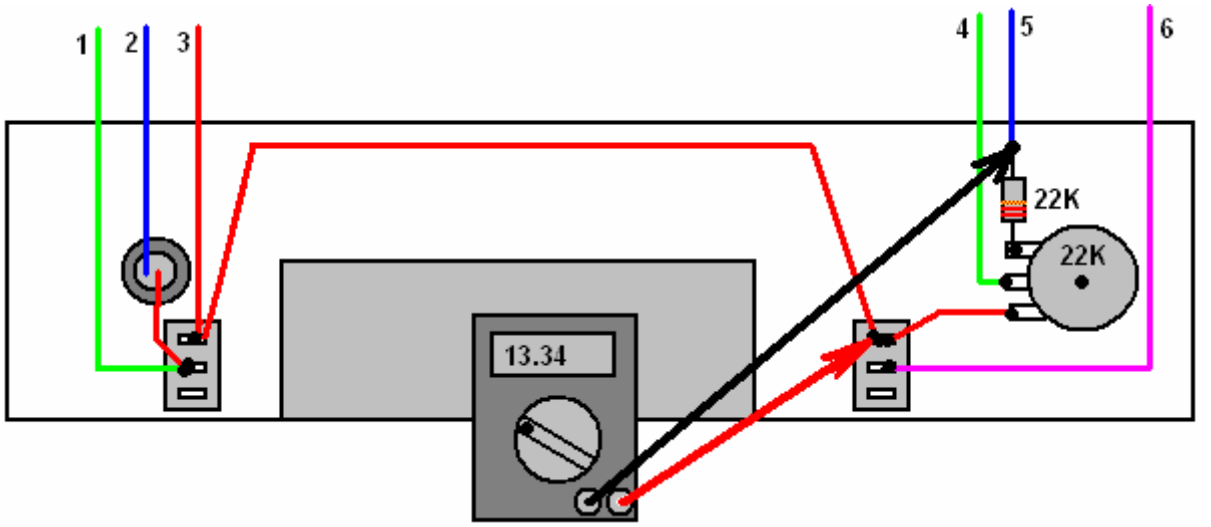
If the unit does not work straight off, then step-by-step tests are needed to locate and fix the problem. These are simple, common sense tests. If you do not have a voltmeter, then it is worth purchasing one as they are now very cheap.

1. We need to confirm that the battery has the electrical power to run the lamp, so open the “boost” switch “SW2” and connect a voltmeter as shown here:



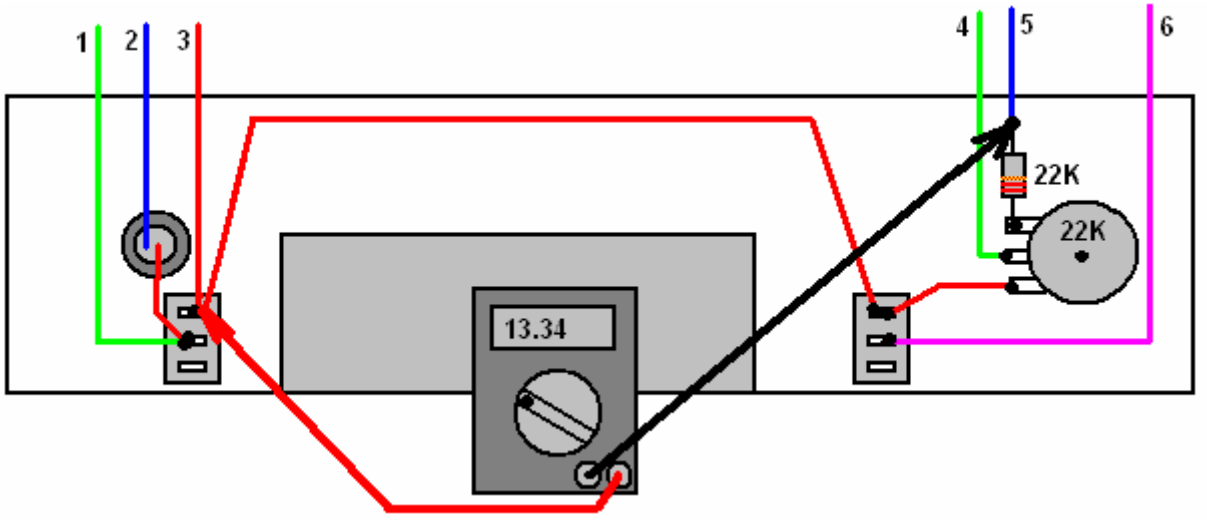
This connects the meter directly across the battery and so the meter should show some realistic voltage around the 12V mark. If it doesn't, then lift out one of the battery holders and check that the batteries are charged and confirm that the fuse has not blown. If the fuse has blown, then it will because of excessive current, probably caused by a direct short-circuit where the battery Plus is being connected directly to the battery Minus. Check the wiring from the battery to these two points until you do get a sensible voltage reading.

2. Next, connect the meter like this:



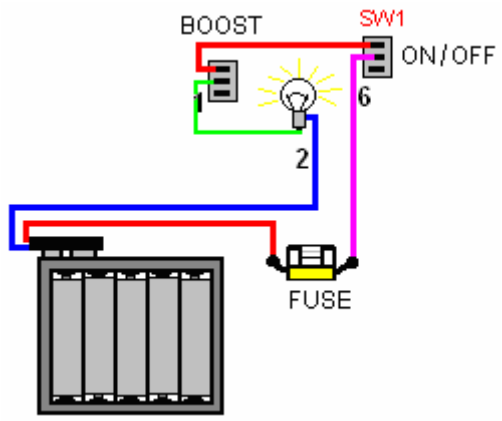
and confirm that the voltage appears and disappears when you operate the On/Off switch. If it does not, then it is highly likely that your connections to the two switch tags are touching and stopping the circuit being switched off. If that is the case, then adjust one of the connections to cure the problem and leave the On/Off switch switched On.

3. The next step is to connect the meter like this:



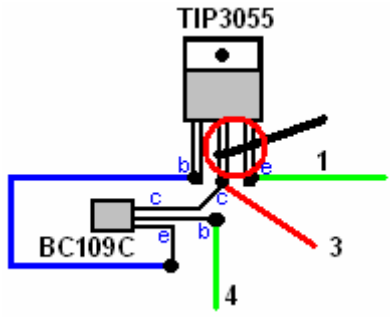
The same voltage reading should be seen. If it isn't, then the wiring between the two switches is almost certainly defective and needs to be corrected.

4. Now close the boost switch, and the light should come on at full intensity because the transistors are being bypassed and the battery connected directly across the LEDs. The circuit is then:



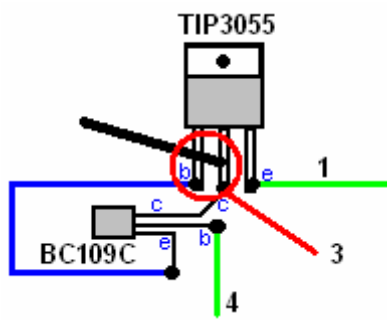
If the light does not come on brightly, then check the wiring around this path – battery Plus – fuse – wire "**6**" - On/Off switch – boost switch – wire "**1**" – LEDs Positive – LEDs Negative – battery Minus (wire "**2**").

5. When the boost switch circuit is working correctly, turn the boost switch off and check the operation of the transistors. First, use a screwdriver to short-circuit between the Collector and Emitter of the TIP3055 transistor:



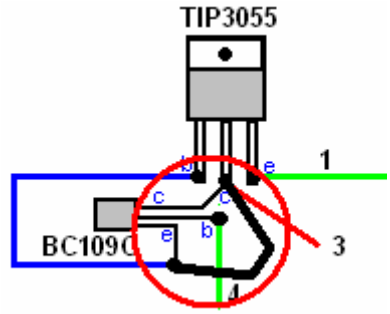
This should have exactly the same effect as closing the boost switch but it is testing your wiring connections to the power transistor and simulating the TIP3055 being switched on fully. If this does not produce the full light output, then you need to check (visually) the wire "1" connection between the centre tag of the boost switch and the Emitter of the TIP3055 transistor. Also check the connection between the Collector of the TIP3055 transistor and the top tab of the boost switch as one or both of those connections has to be faulty if the boost switch works when you switch it On.

6. When you have corrected the wiring and short-circuiting between the TIP3055 Collector and Emitter turns the light on at full brightness and the light-level control variable resistor is still not controlling the light level, then short-circuit between the TIP3055 Collector and Base:



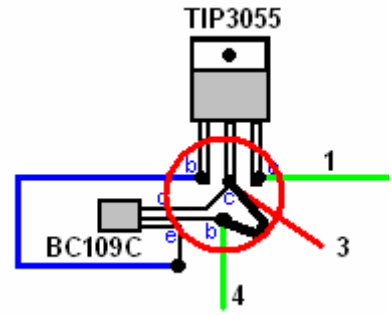
This bypasses the BC109C transistor and should produce the maximum light level. If it does not, then the TIP3055 transistor is defective. This is highly unlikely as the TIP3055 is enormously robust and can survive all sorts of mishandling. However, if this test definitely fails, replace the transistor without connecting its Base and confirm that the new one switches the light on fully when you short-circuit the Base to the Collector. Then connect the BC109C and repeat the TIP3055 Base to Collector short-circuit test to confirm that it is still working with the BC109C in place.

7. If the light-level control still does not work, then short-circuit the BC109C transistor's Emitter to its Collector:



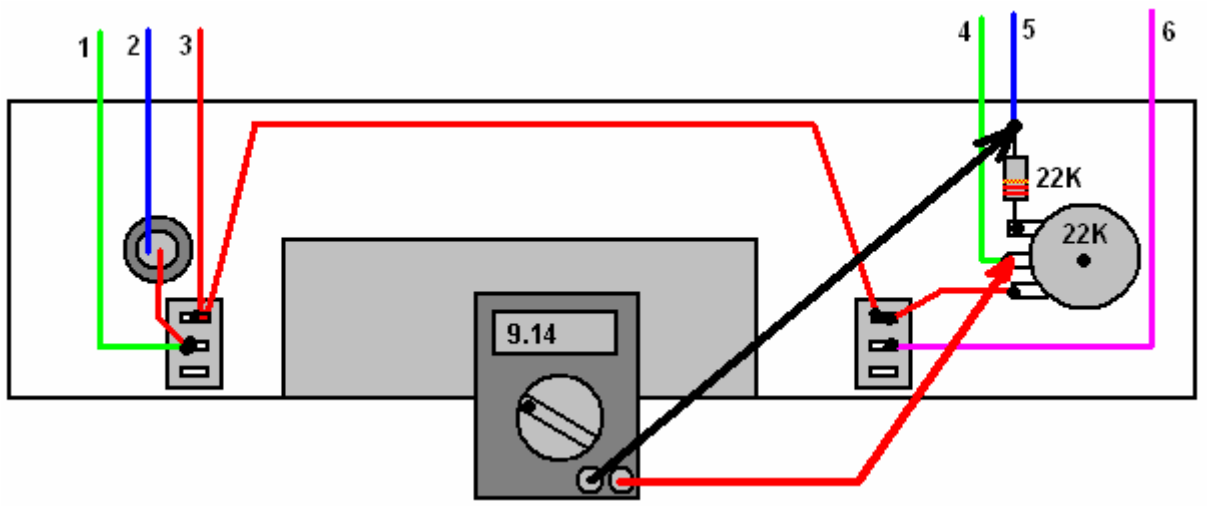
This checks the wiring between the Emitter of the BC109C and the Base of the TIP3055 (shown in blue in the diagram above). If the light comes on at full brightness, then the wiring is okay, otherwise, correct that connection.

8. Now we test the operation of the BC109C transistor. Turn the light-level control knob down to its minimum and short-circuit the BC109C Base to its Collector:

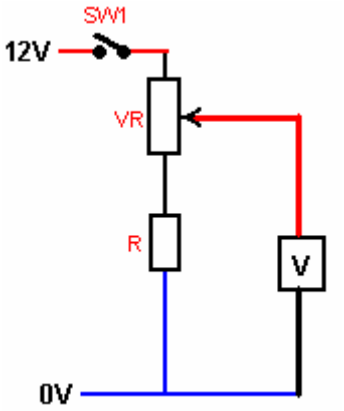


With the short-circuit in place, the light should be fully on and with the short-circuit removed (or if its base is short-circuited to its Emitter), the light should be fully off. If this does not happen, replace the BC109C transistor and repeat the test.

9. If the light-dimming control still does not control the light level, then the wiring of the variable resistor has to be suspect, so connect your voltmeter like this:



This is making a connection between the negative line and the slider of the variable resistor, and that should allow the slider to adjust the voltage fed to the BC109C transistor:



This is the test and unless there is a major problem with the resistor “R”, the variable resistor “VR” or the transistors short-circuiting the slider of the variable resistor, your voltmeter should show your full battery voltage when turned up as high as possible, and an evenly falling voltage to around half that voltage at the other end of the variable resistor shaft movement.

If this does not happen, then you need to check over the wiring connections for the variable resistor, fixed resistor and connection to the base of the BC109C transistor.

Components for home-build:

1. One 10-watt, 12-volt solar panel.
2. Thirty AA-size, 1.2V NiMh batteries with a capacity near 3000 mAHr if possible.
3. Three AA battery boxes each holding 10 batteries.
4. Three battery clip connectors to fit the battery boxes.
5. One 1-amp diode type 1N4007 or similar.
6. One 22K Linear variable resistor and graduated knob to fit the shaft diameter.
7. One 10% quarter watt (or higher rated) 22K resistor (colour bands red/red/orange)
8. One BC109C transistor.
9. One TIP3055 transistor.
10. Five G4 LED arrays.
11. One desk lamp (optional)
12. Two rocker switches.
13. One fuse holder (of any type) and a 1-amp fuse.

14. Ten metres of insulated, stranded wire rated for 1-amp or higher current.
15. Rigid sheet material 6 mm thick (possibly MDF).
16. One small off-cut of laminate flooring 85 x ***L*** (typically 337) mm
17. Adhesive.
18. Rubber feet or soft material for the underside of the desk lamp.

Notes:

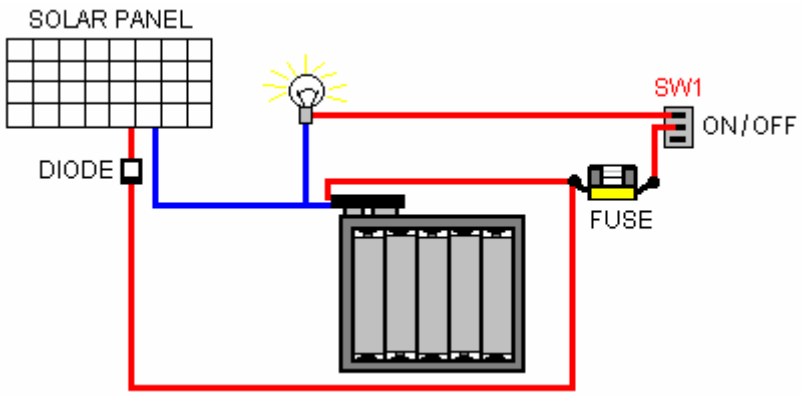
Rocker switches are generally cheaper as the Change-over version with three switch tags, and so these have been shown. For this design, only On/Off action is needed. Press the "1" end of the rocker down, and the two contacts used are then the two farthest away from the "1".

The TIP3055 transistor is the more recent version of the 2N3055 and is a particularly good transistor for this application being much easier to connect to. While the TIP3055 is strongly recommended, other power transistors may be used. No heat sink is needed if the transistor can handle 8 amps or more (the TIP3055 can handle 15 amps and 90 watts).

The BC109C transistor has a very high gain between 200 and 800, which is useful in this application, but a major advantage of it is that it switches off fully with pretty much zero leakage current, while I have some of the very popular 2N2222 transistors which don't switch off properly at all and so are effectively rejects. Most alternative transistors are highly inferior to the BC109C and so, if it is not readily available, then it is worth ordering.

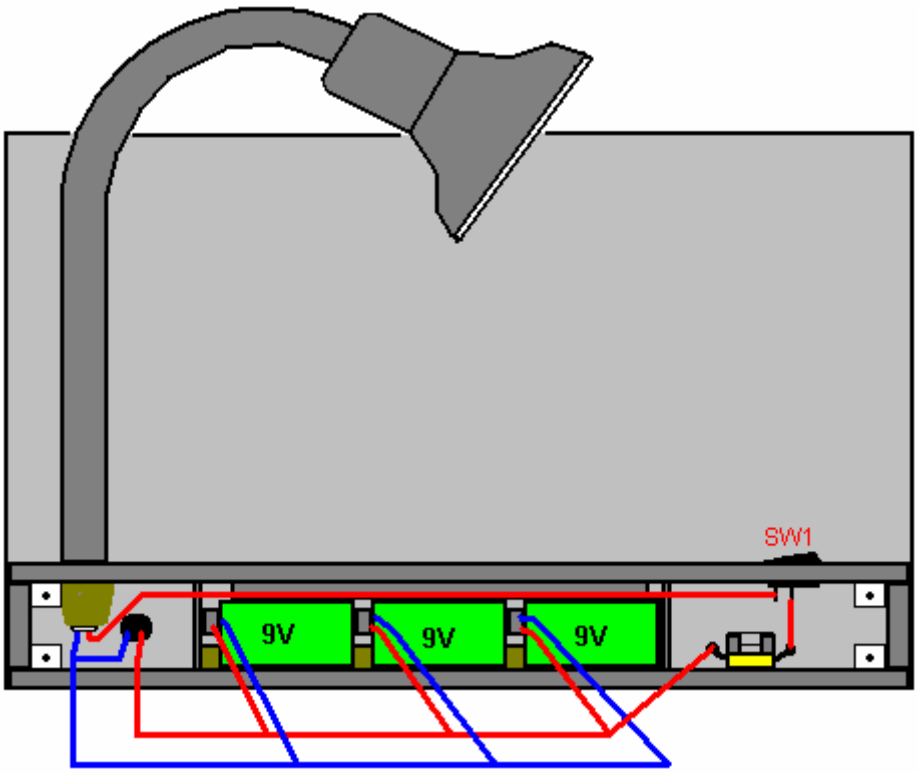
The 1N4007 diode is the cheapest readily available diode. It is rated for 1000-volt operation which is of course, of no interest here and so the lower voltage 1N4001 through 1N4006 diodes are perfectly suitable although slightly more expensive as they are less popular. While one diode is shown in this document, people who are expert in electronics are inclined to use three connected in parallel as that passes current more freely. That would cost an extra 6 pence at 1-off prices in the UK.

Some people are really put off by electronics of any kind, so here is a version of the solar desk lamp which effectively, has no electronics at all:

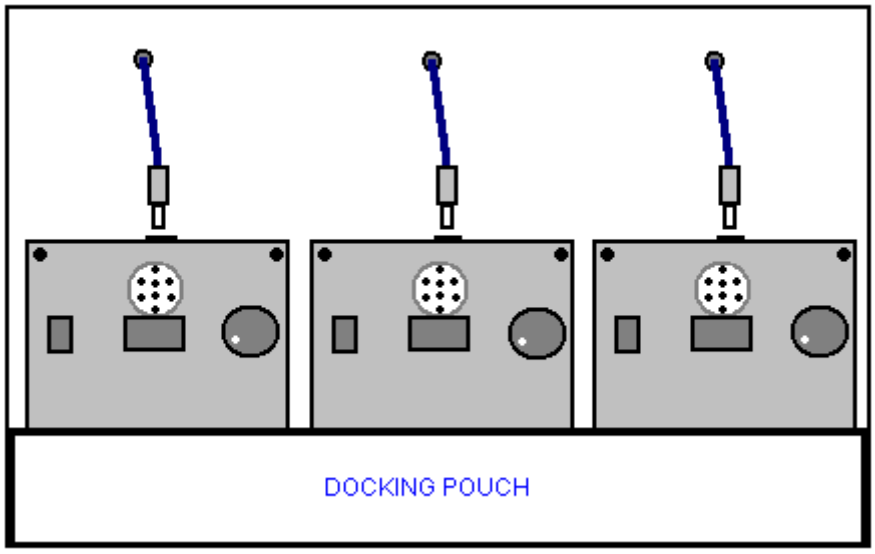


The only control on this version is the On/Off switch. As we no longer have any control over the voltage applied to the LED arrays installed in the lamp shade, and since that voltage controls dramatic changes in current draw, we need to reduce the number of 1.2V NiMh batteries from ten to seven. It is a matter of choice whether you use three or four sets of seven batteries. If 10-battery battery boxes are used, then a wire can be soldered across the gap left by the three omitted batteries.

The general construction is the same as before, with a rocker On/Off switch being installed instead of the variable resistor. As with all of these solar units, the important thing is to switch the light off when it is not in use, to avoid the next charging session being hindered by unwanted current draw through the light. The general layout of the unit can be the same:

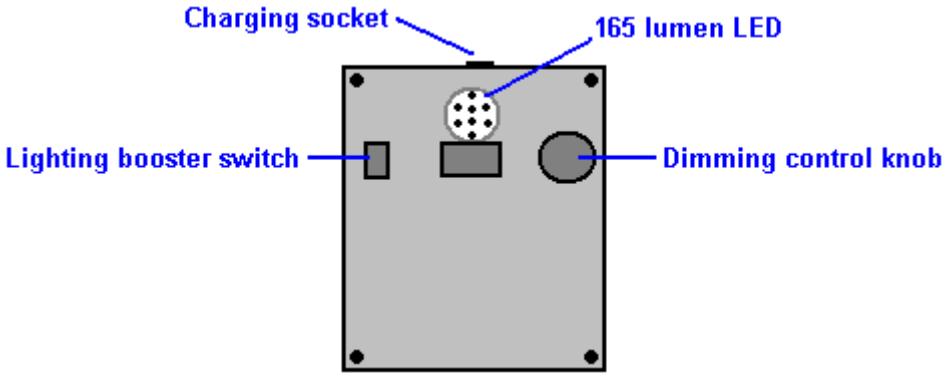


There are other variations which can be low-cost and very versatile. For example, instead of connecting the three battery packs together, they can be kept separate from each other, individually housed with the same circuit, each driving one of the LED arrays:

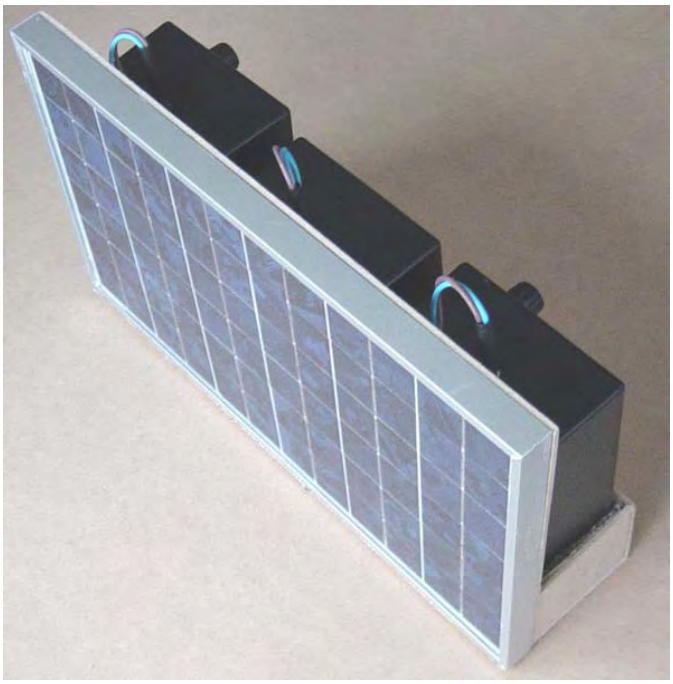


Here, three identical units are charged from the solar panel, which can be any size or wattage, although a ten-watt panel would be considered to be the minimum for daily use. Each unit is in effect, just a wide-beam, high-intensity, high-capacity solar torch which can be dimmed and which will stand securely when unsupported. It has the advantage that the units can be constructed one at a time, and if necessary due to time constraints, or perhaps, very poor daylight, just two or even one unit can be charged from the entire solar panel output. If manufactured and sold, then the units could be sold separately, allowing a lighting

system to be extended and improved when additional funds become available (or when the user has tried one unit and discovered how effective it is). Each unit is easy to understand and use:



A built unit of this type looks like this from the front:

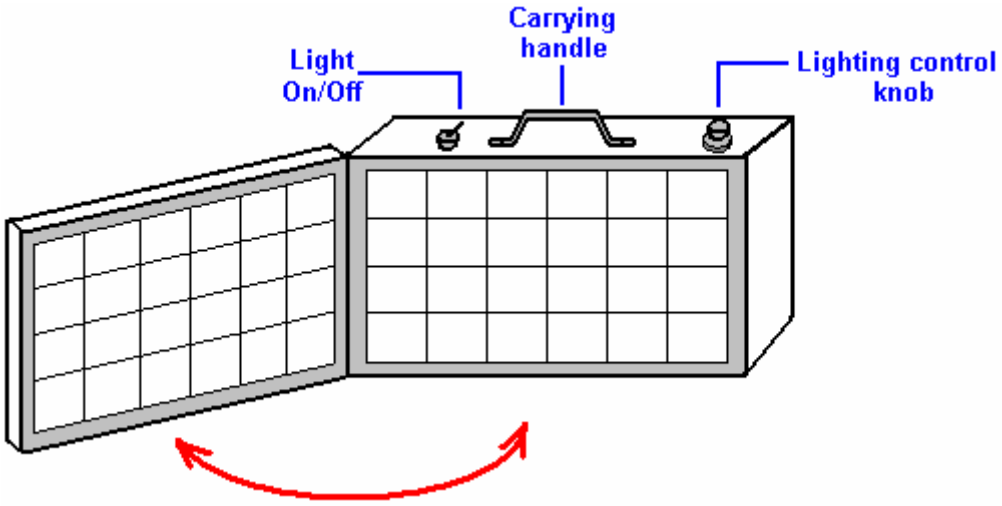


and like this from the back:

While these particular units have just one of the LED arrays each, tests run since that construction show that with manual dimming control, two LED arrays per unit would be more satisfactory. Also, if the second LED array is mounted on the side of the box, then the lighting angle increases from 160 degrees to 250 degrees which should be helpful for indoor lighting of a room. The units could, of course, be used in different rooms simultaneously and there is the option of carrying one around for lighting anywhere. Again, the construction is very simple and the same circuit as for the desk lamp shown above, can be used in each box.

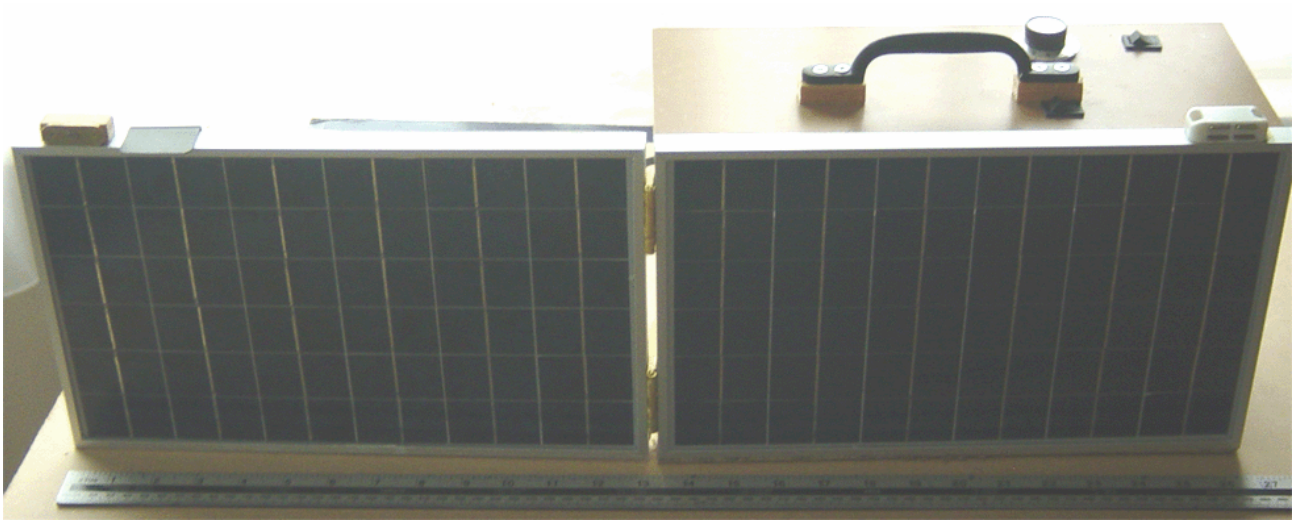
If good-quality NiMh AA-size batteries are used, each of the mobile units is capable of providing good light for eight hours, which means that if it gets dark at 4 pm (16:00), lighting is available until midnight as a minimum, at which point in time, the user should be going to sleep.

If a very powerful single source of lighting is needed, it is possible to use a larger solar panel, or for a more compact unit, two of the 12-volt 10-watt panels shown above. The arrangement can use the same simple manual control of lighting level and the same booster switch for even greater lighting for a few minutes. The arrangement can be like this:



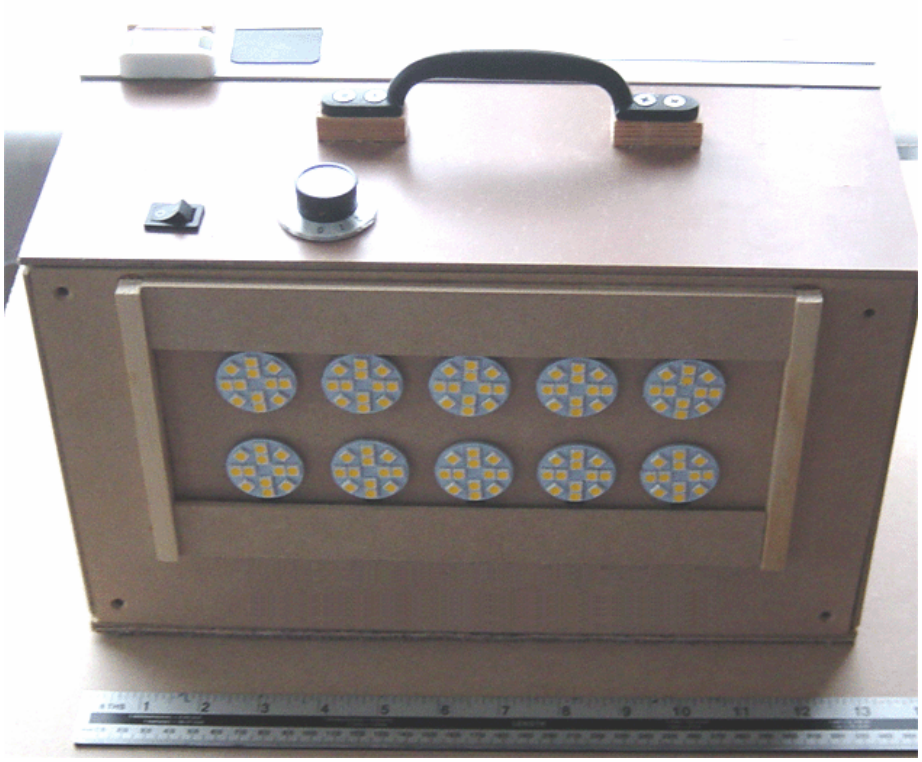
When closed over, the face of solar panel P1 faces that of solar panel P2, protecting both when the unit is being carried.

An early prototype of this style of construction with the panels open, looks like this:



A magnetic catch is used to hold the hinged panel securely closed when the unit is being carried and a small flap is attached beside the magnetic catch to overcome the slightly excessive looseness of the hinges. The unit does not need to be as deep as this experimental model was made.

The front view of the unit, ready to receive the frosted plastic cover for the LED arrays, looks like this:



The underside of the unit is covered with a soft protective layer to ensure that it does not scratch any surface on which it is placed. In order to avoid the need for overcharging circuitry, this unit has six battery packs and so, having ten LED arrays, the duration of lighting is about the same as that of the desk lamp, although obviously, the light output can be much greater. With the greater lit area of ten LED arrays, a lower actual current can be used while still providing a good level of lighting.

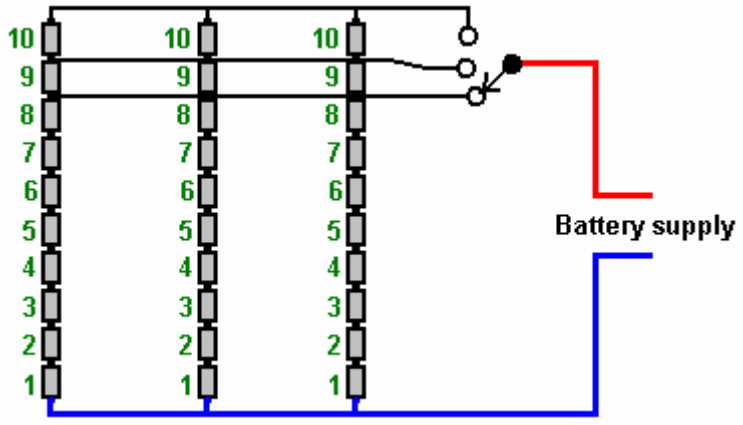
With fully charged batteries switched to the ‘boost’ setting, this unit puts out more light than a 100-watt incandescent bulb powered by the mains. Tested in daylight, that looks like this:

Let me stress again, that these units are not difficult to build and they don’t call for any great knowledge of electronics.

Alternatives:

While the systems shown above are effective and reasonable cost, it is possible to make them more efficient. With battery packs of ten NiMh 1.2V AA cells, when fully charged, we get more than 13-volts, when we actually want 9-volts for effective lighting. We get this by using two transistors to drop the voltage down to 9-volts. The current flow is about 135 milliamps and so, the voltage drop of 4-volts at that current flow is a power loss of 540 milliwatts, which is a significant percentage of the power used to generate the lighting.

By arranging things in a different way, we can reduce that power loss very considerably, IF we don’t mind the user being involved in controlling the lighting level. We could do this by tapping into our 10-battery pack and using an extra 3-way switch which allows us to start off with 8 batteries and when the voltage starts to drop, switch to 9 batteries as the extra fully-charged battery would make a major difference to the lighting. Then when the voltage of the 9 batteries starts to get low, the tenth battery can be switched in:



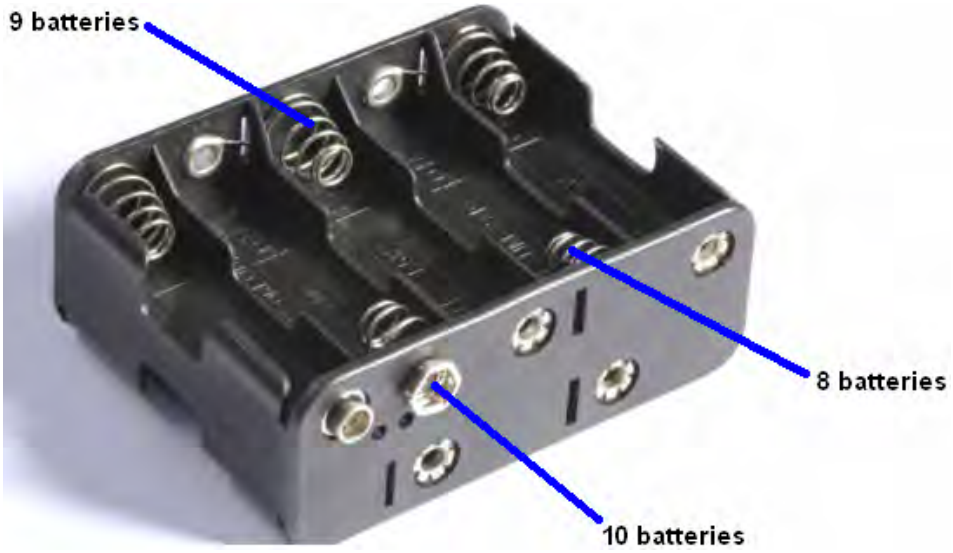
Whether or not this is worth doing is a matter of opinion in spite of the longer and/or brighter lighting. With 8 batteries in use, the power saving is 350 milliwatts and with 9 batteries the power saving is 175 milliwatts. The lighting power is 1215 milliwatts, so these represent 29% and 14.5% of that power. If it is decided to use this arrangement, then the battery boxes can be tapped with additional output wires connected between batteries 8 and 9 and between 9 and 10.

The operational sequence would be:

1. Eight batteries switched in: variable light control.
2. Nine batteries switched in: variable light control.
3. Ten batteries switched in: variable light control.
4. Boost switch closed: fixed lighting level as batteries discharge to their minimum level.

Each of these switched levels gives a new lease of life to the lighting system, but it does require the intervention of the user, although that intervention is very minor.

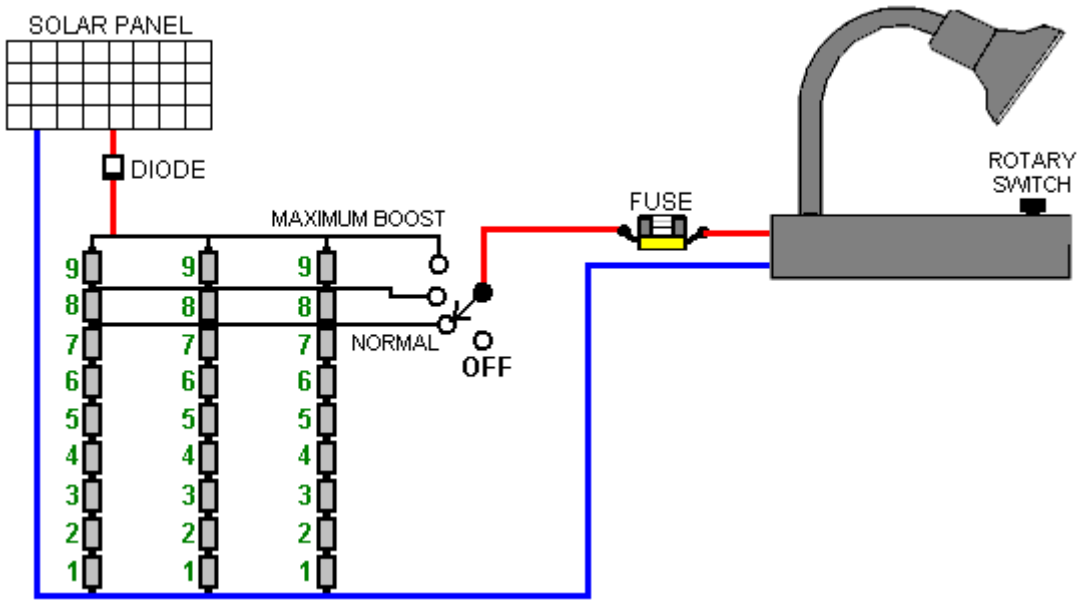
It does not matter whether it is the Plus end of the battery chain, or the Minus end which is switched. The way that the battery boxes are constructed, actually make it more convenient to make the two extra connections at the Minus end of the stack:



Medical usage

The applications of the solar desk lamp shown above are intended for home or office use. There si one other serious lighting requirement and that is for doctors working in isolated rural areas, far from any electrical grid but also, very far from fuel supplies for running a conventional generator which is the only known option for most of them. For example, it can take a doctor two days of travelling to get some more generator fuel. If we apply a slight alteration to the desk lamp and sacrifice the long period of lighting, we could boost the light output from 1000 lux right up to 2,700 lux for an hour or two of very brightly lighting. I can see no reason why two or three of those would not give more than adequate light for night time emergency surgery when placed on stands or suspended from the ceiling. Being self-contained and mobile, an assistant could hold and aim one to illuminate exactly where the light is needed.

The basic solar desk lamp operates off seven 1.2V batteries in series. Eight batteries in series would boost the light level to 1730 lux at the cost of much higher current draw and reduced period of lighting. Increasing the batteries to nine would raise the light level to a massive 2,700 lux with a current draw so high that there might only be lighting for perhaps one hour, but that doesn't matter in an emergency where a life could be saved. To achieve this emergency level of lighting all that is needed is an ordinary 4-way rotary switch which replaces the On/Off switch as it can do that job as well, otherwise the construction remains identical. The arrangement could be like this:



Low-cost Cooling Systems

When most people think of air-conditioning systems they think of large pieces of equipment which are expensive to buy and expensive to run and which are mounted in one fixed location. While those systems are certainly effective (although, often noisy in operation), there are other ways of improving living, travelling and working conditions. Ways which are not very expensive to buy and which are not expensive to run. Many people have made effective air coolers and shared their designs using web videos, and I should like to thank them for showing what they have achieved.

Simple cooling systems can use two different methods of operation. One method is to use ice which has been produced using an ordinary refrigerator or freezer (those being very common in places where air-cooling would be considered an asset). Another method uses the well-known principle that when water evaporates, it draws heat from its surroundings as part of the process.

The general idea is to cool some air and then use a fan to blow that cooler air to where it is most effective for the user. Typically, cooling is most needed inside buildings and inside vehicles. Inside buildings, mains electricity is often available. Inside vehicles, twelve-volt or similar levels of DC power are generally available. In both locations, a solar panel (possibly aided by a car battery for night time operation) can operate a useful system.

What we are talking about here, is a small system which will not be as effective as a large commercial unit, but yet is something which can make a great improvement in a small area. Something which does not seem to occur to many people is that you can have more than one of these units operating in a room or vehicle.

While I am delighted that people have shared their simple designs on the web, they usually assume that the components which they have used are available to everybody, and that is very much not the case as the items which are readily available locally vary enormously at different places around the world. For that reason, I should like to discuss the basic principles used so that when any chosen design uses something which is just not available locally, you can choose a suitable alternative which is readily available locally at low cost.

Let us start with indoor cooling. The unwanted heating comes mainly from the sun shining in through windows, and from heat leaking in through the walls because the outside of the walls is being heated by sunlight falling directly on it or by hot air flowing along the outer surface of the wall. Of these two main sources of indoor heating, the sun coming through the windows is generally the greater source of heat.

If a room has one or more windows facing the sun and one or more windows which do not face the sun, then a first step might be to use shiny silver-faced insulation cut to exactly fit the windows through which the sun

shines. Curtains can generally hold the insulation in place on their own, but if that is not convenient, then a few small pieces of adhesive tape work adequately and can be used again day after day. The outer shiny layer reflects most of the sunlight back out of the room. The outer surface of the insulation sheet heats up as not all of the light is reflected, but the insulation layer stops most of that heat actually entering the room, resulting in a major improvement in the room temperature, although the light level is reduced considerably as well. This is suitable for people living in apartment blocks where there is no convenient access to the outside of windows.

In the case of people living in bungalows or other low buildings, the use of shade should not be overlooked as it is a very effective way of reducing temperature. This effect is even very noticeable in cooler countries on sunny days as there is a very marked effect driving down a road which has large trees on the same side as the sun. With the car windows open, there is a very considerable drop in temperature when the car enters the area shaded by the trees. I am not suggesting growing trees to give shade as that is not convenient for most people, but a simple overhang or a temporary or retractable awning may be possible to block direct sunlight from coming in through the windows. An awning or overhang has no running costs.

Let's take a look at one of the systems shown on the web. This one is being used to lower the temperature in a house with a floor space of 1,700 square feet (158 square metres) and the video presentation is at https://www.youtube.com/watch?v=6ScZiMiva9M and deals with a house where the outside daytime temperatures exceed 100 degrees Fahrenheit (38 degrees Centigrade), making the indoor temperatures uncomfortable. With his cooler running and the outside air 90F (32C) the air entering the house is 69F (20.5C) which means that the warmest place in the house is around 75F (24C). As hot air rises, the hottest place in the house is the highest point and very sensibly, he has mounted an extraction fan which pulls the hottest air out of the house and pushes it into the attic (which is open to the outside and so vents to the outside air).

For the system to be effective, all of the windows and doors are kept closed except for the one which has the cooling system mounted on it. This blows the hottest air out of the house and at the same time, draws fresh air in through the cooling system. You can see the attic fan mounting here:

The entering air comes in through a large window which has evaporative pads clamped to the window frame by a simple timber mounting system:

The pads are kept wet by droplets of water pumped to the top of the window by a small submersible pump submerged in a large plastic storage box filled with water:

Excess water is returned to the reservoir by two short lengths of plastic guttering as shown here:

The particular fan and pump used draw a total of 120 watts, and so as one “unit” of electricity is 1000 watts for one hour, running the system all day and night is a power usage of 120 /1000 x 24 = 2.88 Kilowatt-Hours per day. The construction costs in America in 2013 were about US $60.

If solar panel power is to be used with a lead-acid battery, please remember that the battery is only 50% efficient, which means that it will only return half of the current that you put into it. Also, please remember that solar panels are rated for bright sunlight shining directly on them at exactly 90-degrees to the panel and in normal operation, the panel output is far less than the retailer’s claimed wattage, not to mention the fact that the panel will only be operational during daylight hours. As against that, many such installations do not need to be run day and night and so solar operation may be quite economical. Also, 120 watts is a large amount of input power and could almost certainly be reduced for smaller installations.

It is not unusual for the cooling of a single room to be satisfactory in many situations. For this, using a very small self-contained unit is often considered to be adequate. These units generally use a fan from an old computer, typically, running on 12-volts and drawing about 4-watts. For the cars which still do not have air conditioning, a fan of that type is very convenient as it can run directly off the ‘cigarette lighter’ socket or from a direct wire connection through a fuse, to the car battery. For indoor use, any tiny 12V mains power unit can power the unit for minimal cost.

These units have been made in all sorts of shapes and sizes and by all accounts, do very well when easing excessive heat. This very small unit uses a 12V 3.6-watt computer fan blowing air into a small insulated box containing ice and can be seen at https://www.youtube.com/watch?v=6pwHvXZr1A4. In spite of being rated as a 12-volt unit, the fan can actually run on a low-power USB socket. Cold air exits from the holes drilled beside the fan:

This one is at https://www.youtube.com/watch?v=Wf2Zhu4Fmzo uses a 21-inch (535 mm) square box fan with a simple evaporation system attached to it:

This one at https://www.youtube.com/watch?v=gawOWyhtthU also uses a computer fan:

This one at https://www.youtube.com/watch?v=LiS0W5peFBk uses a Wal-Mart battery-powered fan which means that with a solar panel to recharge extra rechargeable batteries when the unit is in use, the unit could be very, very cheap to run. An ordinary ice-box is used, so perhaps cold drinks are a possible additional option:

Also using an ordinary ice-box the https://www.youtube.com/watch?v=N1NwXBXpShQ unit has a wooden push-fit lid and uses a car fan:

The unit at https://www.youtube.com/watch?v=NUhs_1puHk0 is intended for use inside a car, although obviously, it could easily be operated inside a house using a small 12-volt mains unit. The preferred coolant is 2-litre plastic bottles filled with water which has been frozen solid in a freezer. The builder states that the bottles will cool a car down faster than ice does and the bottles also provide a longer period of cooling than ice can:

The fan is set to blow air into the container and the cold air exits through the four holes drilled around the fan. This design suggests that as a temporary measure, standing some frozen 2-litre bottles in front of an ordinary fan could well provide effective cooling in a room. Using ice or bottles of frozen water is energy efficient as refrigerators and freezers are generally about 300% efficient. That is, the cooling energy is three times greater than the electrical energy used to achieve that cooling.

These are only a few of the many ingenious and simple designs which could make life much more comfortable in conditions which are generally considered to be excessively warm. A little time spent examining these and similar units shown on in web videos should give you confidence to build your own, or if you prefer, to design and build your own simple cooler system.

Patrick Kelly
www.free-energy-info.tuks.nl
www.free-energy-info.com
www.free-energy-info.co.uk
www.free-energy-devices.com

Chapter 15: You Are Being Deceived

Without your knowledge, you were stolen from your parents at birth and automatically became the judicial property of a commercial company pretending to be your government. What you think of as "your" birth certificate is **proof** of this action. The biggest secret is that your government is a collection of commercial companies whose only real purpose is to take as much money from ordinary people as they can, using unlawful actions. Their sole intention is to keep you poor, in fear and doing whatever they tell you to do.

The Situation in the UK

This is a large subject, so let's take it step by step and start with people who were born in the UK. I will discuss the position of a typical native of what is called "The United Kingdom of Great Britain and Northern Ireland". The average person living in this area believes that he is a free person living in a democracy which is governed by elected parliamentary representatives of the people, protected by legal statutes introduced by Parliament and enforced by police officers who are there to ensure your protection where physically possible. Those who are better informed, believe that the Magna Carta signed in 1215 affords them substantial rights and freedoms.

None of that is true. Not a single item.

The inhabitants of this region fall into two categories which are "landowners" and "everyone else". The term "landowner" is, of course, ridiculous. No human can own land. The only way that you can own something is if you create it from materials which you own, if it is given to you as a gift by the rightful owner, or if you buy it from the rightful owner.

If I state "I own the whole of the planet Mars" it does not give me any genuine ownership of the planet. What has happened on planet Earth is that an individual stakes out an area and says "This is my land, I own it, and I will kill anyone who says that I don't". That can certainly frighten most people away from the area, but it does **not** give the squatter genuine ownership of the land. Consequently, no human owns any land whatsoever, and it is not possible for any land to be bought as there is no previous human owner of the land in question.

Killing a squatter on the land and claiming ownership of that land by right of conquest is farcical. It is true that violent bully-boys take control of areas of land through intimidation and peaceful people dare not oppose them. No human has any more rights than any other human, but violence and the threat of violence have been used to create a notional group of people called "landowners". Is it right? Is it just? No, of course not, but it is what has happened.

This violent situation resulted in there being two groups of humans in the country. The violent "landowners" called "Knights", "Lords", "Barons" or some other irrelevant title, who join together to maintain what they have taken, and then there are the other people who have been intimidated by them, sometimes called "serfs", whom the landowners feel free to rob on a regular basis through sheer intimidation.

Already, there was a certain degree of deception being practiced by the landowners as they said to the serfs "we will protect you, and in return, you can work for us and pay us to protect you. This, of course, was a complete con, because if it wasn't for the landowners, there would be no need at all for "protection" from other landowners. And, oh yes, the protection required the serfs to put their lives on the line as part of the "protection" racket.

There was one group of landowners who operated more by trickery than by sheer force of arms. Going under the general title of the "Church" their influence came from telling people that when they die that they would not go to "heaven" if they were not servants of the Church. That technique worked very well and even influenced the head honcho of the landowners - the Monarch who was the ultimate "landowner". Don't get me wrong. I'm not saying that it is wrong to worship God, follow a moral code, live an upright and blameless life, **but**, doing that has nothing whatsoever to do with being subject to any human or human organisation.

A point which you have probably missed is that the 'landowners' considered the serfs to be sub-human "things" and not people at all. Unfit to, and incapable of, owning land under any circumstances. Considered to be 'goods and chattels' owned by the landowners. At that time, "serf" was an alternative word for "slave".

It is not possible for one human to genuinely "own" another human. Parents do not own the children born to them. They have a duty and a responsibility to nurture them and care for them and ideally, love them, but they do not

own them. The weird notion which landowners have that they "own" humans is as false as anything can get. Slavery is another act of violence. Only a very ill-informed person believes that slavery has been 'abolished' as it is more widespread today than at any other time in all of recorded history, and it has nothing whatsoever to do with skin colour.

In the well-known year of 1066, another violent 'landowner' arrived in England and by force of arms, imposed Roman law on all of England. You will recall that Rome was very heavily into slaves and the slaves had no rights or privileges, no possessions and no protection, in spite of the fact that in reality, every one of them actually had equal entitlement to the rights claimed by their so-called "owners", both slaves and masters being human beings. So, please understand clearly, that Roman law encompassed the notion of people being called slaves, and owned as "things".

In 1213, King John was excommunicated and that frightened him so much that he signed a contract with the Pope. The contract is known as the "Treaty of Verona" and it undertook to introduce Rome's canon law to England, which made Archbishop Stephan Langton, the effective ruler of all of the country. The contract was supposed to give the Pope ownership of all of the lands, people, animals, etc. of all of King John's kingdom. According to this contract, the Vatican owns all of England and all of the British Commonwealth countries.

This is, of course, a rubbish document. Firstly, King John did not genuinely own any land of any description and never could have. Secondly, King John did not and never could have owned any of the people living on that land. Thirdly, a contract cannot be signed on behalf of a third party, so his signature could never, ever have any genuine effect on the people of the land. If that were not the case, then I could sign a contract on your behalf, guaranteeing that you would pay me a million pounds. It couldn't happen and if it did, it would be meaningless. The Treaty of Verona contract only applies to the signing parties, to have any effect on the people of the land would require them to be fully informed of the consequences and give their consent to the contract, and that never happened.

In an attempt to be cunning, the treaty was worded so as to make it appear to apply to the direct descendants of both the English monarch and the direct descendents of the Pope, namely, his successors. Personally, that seems highly doubtful to me as that would make it, in effect, a contract between people who have not agreed to it or signed it, and it is by no means obvious how a contract could be binding on people who are not party to it and who have not signed it.

However, just two years later in 1215 the landowners forced King John to sign a contract called the "Magna Carta". It is highly doubtful that a contract where one of the parties is intimidated into signing, has any actual validity, but that aside, the contract was made under the conditions prevailing at that time. The landowners, called "Barons" or "Free men" in the document, were insisting on a continuing agreement between them and the monarch. The contract has nothing to do with the serfs who were considered "things" with no rights at all. Today, members of the general public, who are actually the descendants of those serfs, read the Magna Carta and not understanding the situation when it was written, mistakenly think that it speaks of them. It doesn't. The objective was actually an attempt to strengthen the Treaty of Verona.

This power-play continued in 1600 when a private company was set up. It is still in place today and it is called "The Ministry of Justice" D-U-N-S Number 22-549-8526, Directors: Lord Falconer of Thoroton. Today, all Magistrate Courts are private for-profit companies subsidiaries of this main commercial company and not, as most believe, part of the government of the UK.

The next step came in 1689 with the "Bill of Rights" whose purpose was to make the monarch subject to Parliament. Again, this was introduced with the intention of ensuring that no monarch could set aside the Treaty of Verona and deny the Pope (supposed) ownership of the British empire and all the people in it. The Bill of Rights settled the succession of the crown, creating a Dynasty which has been maintained right up to now. The only powers of a Constitutional Sovereign created by this "Bill of Rights" is to warn, to guide and to be informed, just that and nothing else. In 1701 there was the Act of Settlement to confirm the line of royal succession.

In 1694 a privately owned company called "**The Bank of England**" was formed. The name gives the impression that it is a government organisation, but that is not the case. Actually, the UK has no government but instead there is a set of interlocking, privately owned, for-profit commercial companies designed to give the false impression of a government. For example, **the House of Commons** is a commercial for-profit company (number UC2279443). **The Labour Party** is a commercial company which trades under the name of "Allister Darling MP". **The House of Lords** which is the highest court in the land is a private company.

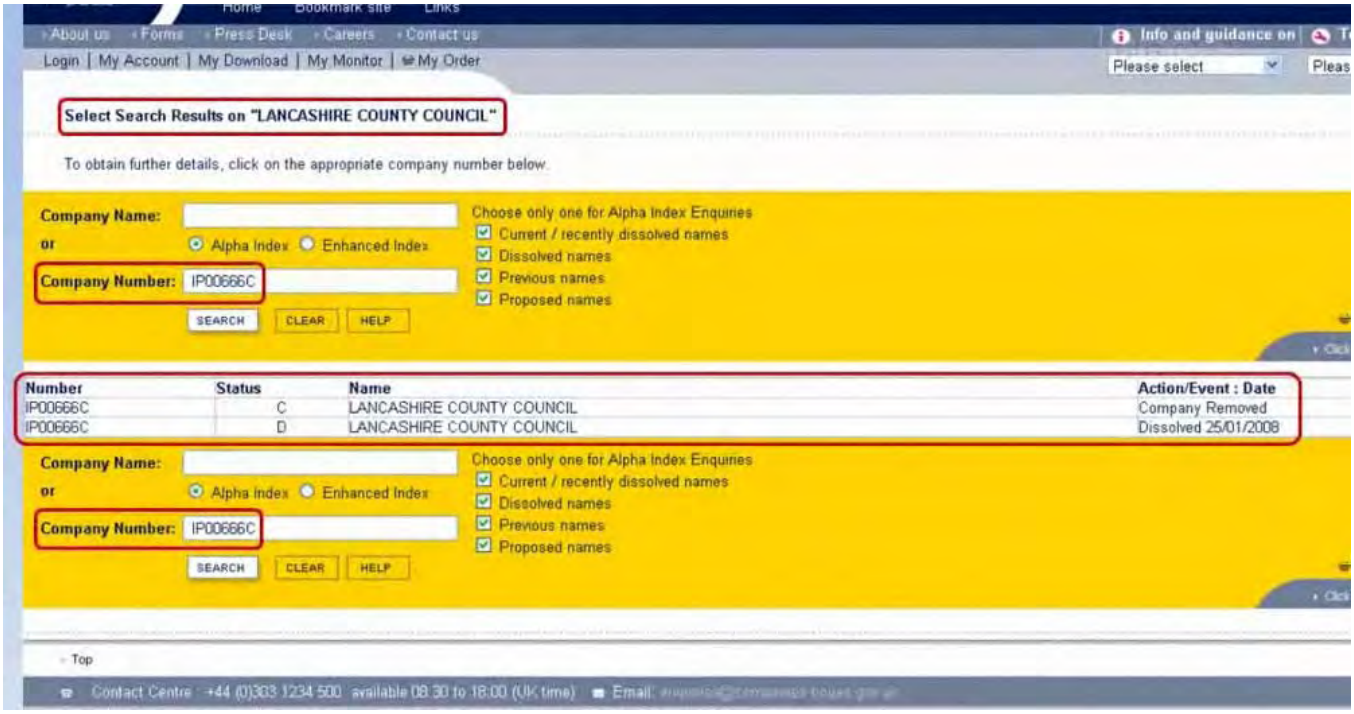
Then there is the **United Kingdom Corporation Ltd.**, formerly known as the "United Kingdom plc". In passing, anyone who has a National Insurance Number, is automatically deemed to be an employee of the United Kingdom Corporation Ltd. The Company policy of that corporation requires each employee, to pay tax and follow all the

legislative rules ("statutes") of that corporation. Anyone who breaks the rules of the corporation will be disciplined under the legislation of that corporation.

The police forces of the United Kingdom Corporation are all companies which can be seen from their company registrations, and their job is to enforce the rules of the Corporation. **Every court** in the land is a company. There is a commercial company called the "**Secretary of State for Trade and Industry**", which ceased trading on 4th November 2002 and yet it still operates TV licensing, taking £3 billion per year through it's subsidiary commercial private limited company called "Consignia (Customer Management) Ltd.", which is part of it's group of 35 subsidiary companies.

This continues to a ridiculous level where the company called "**The Devon and Cornwall Police**" has been taken over by a company called SouthWestOne which is one of a group of companies whose Ultimate Parent Company is IBM, but this Police Constabulary is a public service paid for by the taxpayer and given an annual budget of £256,800,000 taken from members of the public. Isn't IBM a lucky company to be given £256,800,000 per year to help it run it's business at a profit !

Equally ridiculous is the **Lancashire County Council** which was incorporated as a company (IP00666C) in 2002. It's registered office was 3rd Floor, Christ Church Precinct, County Hall, Preston. It was completely dissolved on 25th January 2008 and all of it's Assets and Liabilities were transferred on 12th November 2007 to another company - "The Blues and Twos Credit Union Ltd." whose registered address is Lancashire Police Headquarters, PO Box 77, Hutton, Preston.



If you feel that these things are all straightforward and aboveboard, then I would like you to change a bundle of £3 notes for me, and as well, perhaps you would be interested in buying the Eiffel Tower.

The ordinary people resident in the UK are descended from the slaves called "serfs" and the only protection which they have is the "Common Law", upheld by a policeman or a policewoman, each of whom receive authority through, and only through, their personal oath of allegiance to the monarch and the monarch's oath to uphold Common Law. The Police Authority for whom the policeman or policewoman work, actually has no authority itself as, being a company, it cannot swear an oath. Policemen and policewomen have a duty to uphold Common Law and to protect and serve members of the public.

UK Common Law deals with a human causing death or injury to another human, damage to property or acting fraudulently in dealings with others. It is based mainly on the earlier findings of courts and is modified from time to time if it is found to be unjust. Surprisingly, the main corrections are supposed to come, not from judges, but from juries. A jury is the most powerful element in any court of law for they have the power to dismiss a case and even to go against the law. A jury has "discretion" and they can order the judge to find against the law if they believe that in the particular case which they are trying, the law would be unfair, so they adjudicate on both the individual case **and** on the underlying law. Judges do not always bring this to the attention of jurors.

A system has been set up by the Law Society with the express intention of enslaving members of the general public. They have taken words from the English language and redefined them to create another language called "Legalese".

The translation for this language back into English is given in the publication "Black's Law Dictionary". The average member of the UK public is generally unaware that Legalese is being spoken and that what sounds like a simple, innocuous question in English, is actually a highly-loaded and underhand attempt to cause him both distress and unwarranted financial loss.

The objective of a lawyer is to make money for his commercial company and particularly for himself. To help them, lawyers have set up a system of "statutes" which are supposed to have the force of law, but which, being crafted by their sister parliamentary commercial companies, are **not** the law. These things are "statutes" and they don't actually apply to humans. Being wholly invented "legal" devices, they only apply to other "legal" things, which are essentially, just pieces of paper. Since a piece of paper does not own any money or goods, to get any real financial gain from their actions, they need to trick a human into taking on responsibility for a piece of paper. If they succeed in this, then they can demand money from that piece of paper and get the unfortunate human to actually provide the money even though the fictitious legal charges do not apply to that human.

Meet Your Strawman

In order to do this, they have to persuade the human to enter into a contract with them, agreeing to pay them money, essentially, for nothing in return. Their main trick is creating a fictitious legal identity using the name of the human. This is done by creating a "Birth Certificate" in the same name. Most parents mistakenly believe that they **have** to register the birth of their child, which is actually **not** a requirement under the law (which **does** apply to humans, as opposed to "statutes" which don't apply to humans).

In spite of their mistake, the fictitious legal "person" produced by creating the Birth Certificate document, the human who has the same first and last names as those recorded on the birth certificate, has no actual link to that document unless he can be tricked into agreeing to be linked to it.

All of the company statutes such as Income tax, Council tax, television licence, parking tickets, and the like are not being charged against a human but instead are being charged against the legal fiction with nearly the same name, sometimes called a "strawman". If the human refuses to agree that there is a link, then the invented charges collapse and the human is left alone until the next attempt to fool him. At this time there is a very nicely produced video http://www.youtube.com/watch?v=ME7K6P7hlko which explains the fiction of the strawman legal document.

The people who work for one of the commercial companies called Police Forces have two entirely different roles. As a police**man** or police**woman** they have a duty under common law to protect members of the public, and serve them by upholding common law. They have an entirely different role as a police **officer** dedicated to making money for their employer by imposing all kinds of fines, charges and penalties under a series of statutes invented by their sister companies. These financial penalties do not actually apply to humans, and so trickery must be employed in order to con the human into undertaking to pay when there is no requirement for the human to do so.

A Police **Officer** is a corporate employee who acts as a Revenue Collector for the Corporation and enforces Acts of Parliament.

A Police **Community Support Officer** is a Corporate employee who has no powers whatsoever, and holds no Oath of Office. They are given a Warrant Card and told to keep it because while they don't need it now, they will need it at some point in the future. Without an oath to the monarch, they have no authority at all.

The British people live in a "common law jurisdiction". This means that no Civil Law Statute applies to them at any point **unless the individual consents to it.** If asked to give your name, do not do so as that is a trick which places you into the realm of civil law which does not apply to you at all unless you consent to be placed under it. You are not obliged under common law to answer the question which asks your name and/or address. The police officer will be following training and may well not be aware that trickery is involved, never having been trained in common law and their duties under it, to each member of the public.

The Comments of John Harris

John Harris of the http://www.tpuc.org/ web site makes the following points:

The deception of "Joinder" or more aptly named "pretensive joinder": The human is asked for his full name, address and date of birth in order to establish the LEGAL PERSONALITY, the address of the LEGAL PERSONALITY and the date on which the LEGAL PERSONALITY was created. Once Joinder has been established, they can then offer that LEGAL PERSONALITY a contract. But, they need the consent of the human to act on behalf of the fictitious LEGAL PERSONALITY and so the human is asked "do you UNDERSTAND?" The human is then required to sign at least one piece of paper saying that he is the LEGAL PERSONALITY and that as the LEGAL PERSONALITY, he UNDERSTANDS.

This is deception as "UNDERSTAND" is taken to mean "STAND UNDER" and has nothing to do with the human comprehending anything. During any process where a human has any dealing with Corporation Officers, he is asked "do you understand?" and that is simply about **consent**. He is being asked to represent the LEGAL PERSONALITY and give the Corporate Officer authority over him.

The deception of "Notice": In Legalese, a "notice" is not a bill or a demand, instead, it is an offer of contract. When the offer has been made, the LEGAL PERSONALITY must be established through Joinder because the LEGAL PERSONALITY is only a piece of paper. Consequently, the human has to agree to represent that piece of paper by saying "I understand". Fixed penalty notices are the same as all, and any, statements made: they are simply, self-perpetuating Contracts. Making a statement to police is exactly the same - a self-perpetuating contract.

The deception of "Includes": The Inclusion of one is the Exclusion of another – this is a maxim (an established, unchallengeable fact). A summons (invitation) for watching TV without a licence has "We will present any evidence found of your watching TV without a licence, which may include a statement taken under caution by one of our Enforcement Officers". But the legalese "may" translates into "must", but if you make a statement with your name, address and confirmation of watching TV, then the court issues summary judgement as you have already incriminated yourself and the court has nothing to decide.

The deception of "Summons": Summons is synonymous with "invitation" which is a choice. They are inviting a LEGAL PERSONALITY to a Corporate place of business. The LEGAL PERSONALITY is just a piece of paper, a BIRTH CERTIFICATE created by the United Kingdom Corporation and it is **not** the human. You can't be forced into a contract, so they have to deceive you into entering into one without realising what you are doing. They are using deception as every Magistrates Court is a trading name of the commercial company called "The Ministry of Justice" D-U-N-S Number 22-549-8526 which does not have a Parent Company listed meaning that it is a Parent Company. Directors: Lord Falconer of Thoroton, date of incorporation 1600. Legal people on being shown this registration, responded by saying that if this information is genuine (which it is), then the UK has been lawless for more than 400 years because the whole Justice System is being dealt with by a company.

Deception of the Name: If you go to a Court (which is a Corporate place of Business) as the Accused, you will be asked to confirm your Name, quoting the full name shown on your birth certificate, which is the LEGAL PERSONALITY. Titles such as Mr, Dr, Lord, PC, QC, or whatever are not asked for as they are not required. The "Accused" is actually the LEGAL PERSONALITY which is the name on the birth certificate, so when they ask for the person's NAME, they are talking to the LEGAL PERSONALITY and **not** the human. This is because a human **cannot** exist in the legal world - only pieces of paper can, and that is something which they are very careful not to tell you.

This is a really key issue. Natural Law and Common Law are the only laws which apply to humans and they deal only with harming other people or causing them loss, and outside of those restrictions, a human has free and unlimited entitlement to do anything he chooses which complies with these principles. As opposed to this, Acts of Parliament, "Statutes" and Statutory Instruments "Contracts" do not apply to the human but only to the piece of paper which is the LEGAL PERSONALITY and which has no reality. As the legal fiction of the LEGAL PERSONALITY was created by the company called "the United Kingdom Corporation", it is that company which gets to say what the rights and duties are for that piece of paper.

The deception of the Birth Certificate: When a person is born in Britain, the mother and father submit a Birth Certificate Registration Form, which is a piece of paper. There is no requirement under common law to do this. When any limited company or corporation is set up, there is always a Certificate of Registration in order to create it's LEGAL PERSONALITY and that is a piece of paper. Please note that a British Birth Certificate states quite clearly that it is **not** evidence of identity, that is, it has nothing to do with a human. Marked on it is "Crown Copyright" showing clearly that it does not belong to an individual and was created by "the crown" which is not the monarch but is the independent 'City of London' represented by the Law Society and acting on behalf of the private owners. This act of Registering a child, makes that child a "ward of the court" which means that the child can be taken away from the parents at any time.

The deception of Registration: If I lawfully own a vehicle, then it cannot be taken away from me and crushed as that would constitute a criminal act and considered theft, no matter who takes it. However, if I Register my vehicle, then, without realising it, I hand ownership of the vehicle over to the registering company as a free gift. They acknowledge the gift by issuing me with a piece of paper stating that I am "the Registered Keeper" of their vehicle, to which they assign a number in their registration records of the vehicles which they own (although they will not pay for the upkeep of their vehicles).

John also asks:

Do you believe Parliament is elected to Represent the people and their wishes? It isn't, they are Directors of a commercial 'for-profit' corporation and so, nothing will ever change.

Do you believe in Politics? There is no such thing.

Do you believe that the only way to sort the UK out is to elect a different Political Party? It makes no difference as all you are doing is choosing employees of a commercial company, and the owners dictate what the company does, no matter who the corporate serving officers may be.

Do you believe that Acts of Parliament ("Statutes") are Law? They aren't.

Do you believe that Statutory Instruments are Law? They aren't.

What The UK Police Can, and Can't Do

Posted September 26th, 2008 by Guy_Euden

Do you feel intimidated when pulled over? Do you feel intimidated when stopped in the street? Do you feel that you are obliged to give them information or answer their questions? If your answer is "yes" to any of these questions, then you **don't** know your rights. A police officer is not what he was many moons ago. In bygone days they cared for their employers (yes you), they actually wanted to help you, they wanted to solve crimes, they had compassion, a sense of fair play and in response to this they earned our respect and admiration.

ALL of that has changed! Police officers are now Revenue Agents plain and simple and if any police officer is silly enough to think that's not true, or just wish to profess it isn't, then let him try not giving out fixed penalties for a month and see if a talking to is not the result! The truth is that they are now your adversary and that is not meant in a violent way, more like a game of who can get who to do as they're told, and they are trained to get from you what they need in order to actually fine or arrest you by actually getting you to consent to it. Let's quickly look at all laws that don't include Death or Injury to others, Damage to another's property or using fraud in your contracts.

OK, these three very heavy crimes are all you can truly be guilty of - YES it's all you can truly be guilty of!!! Anything else is just STATUTE LAW and the simple to remember definition of this law is "laws that only apply if you consent to them"!! Here, look at the definition of a Statute: "Legislative rule of society given the force of law **by the consent of the governed**". Yes look at the last four words, "CONSENT OF THE GOVERNED" that is you!! So, if you don't consent, then they can't do anything. Yes, I know what you're thinking, but it really **is** that simple. Ok, an officer walks up to you and says "can I ask you some questions?" Don't consent. Ask him "what is the nature of the intended detention?". Record the conversation and let him see that you are doing this. Use your mobile, as 99% of mobile phones available today have recorders so go for it! If he says that you cannot record this conversation, then what he is in fact saying is, "yep you guessed it - I don't consent to be recorded!!". So, you simply say "Ok, am I free to go?"

The simple fact is that they **need** your consent, and they have some very cool tools to use to get it, one is intimidation, and another is acting. They act like they are your friend. They act like they care. They just ACT! This is to put you off guard and most people sing like canaries!! Their other weapon of choice is Inconvenience. You don't have time to waste playing their game, so you talk in order to get away quicker, and wham! - they have their result! If you are pulled over for an offence under an act, for example, speeding, seatbelt or some other infraction, it is all part of the "Road Traffic Act of 1988 keyword is "ACT" and believe you me what a merry old act it is, it is pure stage play.

When issuing a ticket they **need** your name so SHUT UP! Don't argue with them as that is "adversarial" so just SHUT UP! They need your name and date of birth to confirm which corporation they are going to give the bill of exchange to (bills of exchange defences are simple but we are not doing those right now). So don't provide a name, or date of birth, or address. Don't give them anything to work with, don't say a single word no matter what coercion they try, hell they might even go for some "good cop bad cop" stuff this is great to watch when you are

fully aware of their lack of power, it's like watching "EastEnders". If they read you your rights you have to say that you understand in order for handcuffs to be put on you. **Again, your consent is necessary** (get it?) don't say a word!!

Now these Officers of Revenue don't like the grim realisation that they have no power. You could get arrested, yes I am afraid that they can still arrest you, but the worst that they can do is keep you for a few hours, so get some sleep, oh and if you have a “notarised claim of right, with a schedule of charges for arrest and detention”, then sleep even sweeter because these fools are paying you to learn that they have no power! I know that getting arrested isn't nice, they know this, that's why its one of their tools to get you **to consent to their unnecessary and unlawful penalties**.

But if you don't stay quiet, then they will always have power over you. Now if they are going to arrest you, one nice way of them playing into your hands is if you say these few of words: "Officer, I am a peaceful man, there will be no violence, but you are obliged to note that if you are going to arrest me it will be under protest and duress at all times and I reserve all of my rights at all times, and I waive none of those rights at any time for any cause or reason”. If they take you to the station remember DON'T SAY A SINGLE WORD. Hell you might get a new name! "JOHN DOE". If they try to fingerprint you, tell them that you do **not** consent to this. If they try to get your DNA then tell them that you do **not** consent to this. If they touch you, tell them, "Officer I do not consent to be touched at all".

If you make it known that you know your rights then they will leave you alone. Trust me, they will only act illegally if you continue to let them do so. Remember that any “legal” acts that they choose to quote are powerless, a clawless, toothless tiger.... they have no power if you don't consent.

Summary:

The present UK (and world) situation is designed to confuse you. The UK is a “Common Law jurisdiction” which has had a spurious and fake “legal” system added on to it in order to rob you blind. The average inhabitant of the UK is bound to keep the common law or suffer the consequences. The common law requires a human not to kill or injure another human, not to damage or steal his property and not to swindle him.

The spurious “legal” system introduced by the Law Society (who consider all other humans as inferiors), is designed to steal from those humans by deception and fraud. It operates by tricking humans into inadvertently and unknowingly entering into a contract to pay money to someone who will not provide anything of worth in exchange.

This is done by using the legalese language which has taken words from the English language and changed their meanings without making the changes of meaning made clear to the public. This covert and surreptitious method is used to get a human to represent a legal fiction which is a piece of paper and undertake to pay invented debts attached to that piece of paper.

The average UK resident believes that there is a UK government, when in fact, there is no actual government, just a commercial company pretending to be a government, namely, the company called “The United Kingdom Corporation Limited”. People are fooled into “voting” for the election of corporate officers in this company in order to make them believe that there actually is a government where no government exists.

This hateful “legal” system is there solely to enslave and rob ordinary people living in the UK and there appears to be no way to circumvent it. I have seen a calculation which estimates that on average, ordinary people in the UK have 80% of their income taken from them in taxes. So, for every £100 they earn, they only receive £20.

A self-employed person may be able to avoid paying Income Tax, Council Tax, TV licence and some other minor taxes. An employed person is tied to a “legal entity” employer which cannot escape the legal snare and so is forced to take Income Tax, National Insurance and Child Support amounts off a person’s earned payment before the remainder is given to the employee.

Unfortunately, even self-employed people can’t escape the thieves because they add their massive charge on to essential items which are not sold to people by name. For example, I have been told that 85% of the cost of a gallon of fuel for a vehicle is tax. That means that if I buy £100 of fuel, it really should have only cost me £15 as that is the cost of finding it, extracting it, refining it, transporting it and making a profit on selling it. My £85 has been taken against my will, by people who are not entitled to it.

If I buy almost anything in a shop, then 17.5% of the selling price is taken from me and given to someone who doesn’t deserve it. In reality, it is much more than that because the item purchased was transported to the shop using fuel which was heavily taxed, and I have to pay that tax indirectly in the raised cost of the item. Also, the

shop has to have heating and lighting, both of which are powered by heavily taxed fuel, which raises the price which I have to pay on the item. Then, there is the tax on the company and the tax on the company for having employees, both of which have to be covered in the price which I am charged. You can see why the person who costed out how much is really paid in tax, came up with an overall figure of 80% of a person's earned income. Even a self-employed person can't escape paying a major amount of indirect tax. The often-quoted phrase "we live in a free country" does not seem to have any basis in fact. So, how about "the land of the brave and the land of the free" across the Atlantic ocean, do they fare any better?

The Situation in America

In 1606, King James I the King of England set up a commercial enterprise called "The Virginia Company" with himself as the main stockholder. The Virginia Company was formed in anticipation of the imminent influx of white Europeans, mostly British at first, into the North American continent. The Virginia Company of London (also called the "London Company") was chartered by King James with the object of colonising the eastern coast of North America between latitudes 34°N and 41°N. It's shareholders were Londoners. The Virginia Company claimed ownership of most of the land that we now call America. The Virginia Company (the British Crown and the bloodline families) claimed rights to 50% of all gold and silver mined, plus percentages of other minerals and raw materials, and 5% of all profits from other ventures.

As has been already remarked, no human can own land as there is no way that genuine title to the land can be bestowed on any human, short of God giving him a written deed of ownership as a gift. No human can buy land from another human or organisation as they didn't genuinely own the land in the first place. King James was essentially, no different from any other man and he had no justification for blithely claiming ownership of another country. Even being the first squatter on the land (which he wasn't) does not give any man ownership of that land. Consequently, his was, and is, a spurious claim with no basis in fact or any form of justification.

At this time, America does not have a government because "The UNITED STATES" is a commercial company. The term "United States" means a Federal corporation. The US President is the President of a commercial corporation and American citizens are, in effect, the employees of that Corporation. This present situation was sneaked in by the covert actions of the Americans who organised and ran the civil war with the intention of taking lands and assets from people who lived in the South. Most people imagine that, say, Virginia, is a part of the American nation under the American government. However, this is not true as every one of the "States" is actually set up as a commercial, for-profit, company, under the parent company "The United States of America". So the reality is that instead of being a State of the Nation, the "State of Virginia" is the name of a commercial for-profit company.

The Vatican claims that it owns Britain and all of the commonwealth countries. Britain claims that it owns America as a British colony. The King of England provided financial backing to both sides of the Revolutionary war and believed that the lands of the Virginia Company were granted (on lease) to the colonies under a Deed of Trust and therefore they could not claim ownership of the land. They could pass on the perpetual use of the land to their heirs or sell the perpetual use, but they could never own the land itself as ownership was supposedly retained by the British Crown.

You may be interested to know that the Vatican's system of law is based on the ancient Roman system and in Rome, the seat of power for Caesar was called "Capitol Hill" where Caesar would meet with 'the Senate'. The Pope believes that he can abolish any law in the United States (Elements of Ecclesiastical Law Vol. 1, 53-54). The Pope's laws are supposed to be obligatory on everyone.

The conglomeration of privately owned commercial corporations know as "the United Sates of America" go along with the same old ridiculous scam of considering American human beings to be "slaves" under Roman law. Consequently, no American can own property because slaves can't own property. If an American reads the Deed to a property which he thinks that he owns, he will discover that he is listed as a "Tenant" (Senate Document 43, 73rd Congress, 1st Session).

After the first 21 years from the formation of the Virginia Company, all 'duties, imposts and excises' paid on trading activities in the colonies were paid directly to the British Crown through the Crown treasurer. Queen Elizabeth II of Britain, controls and has amended U.S. Social Security: Statutory Instrument 1997 No. 1778 The Social Security (United States of America) Order 1997.

The American Inland Revenue Service Publication 6209 is form 1040 which is to be used for tribute paid to Britain. Americans are considered to be slaves to the Queen and own absolutely nothing (Senate Document 43). America's Social Security is not insurance or a contract, nor is there a Trust Fund. The criminal courts on the lands of the Virginia Company were to be operated under Admiralty Law, the law of the sea, and the civil courts under common law, the law of the land. This is a crucial point. The United States of America is not a country, it is

a corporation owned by the same Brotherhood bloodlines who owned the Virginia Company, because the USA is considered to be the property of the Virginia Company.

There is a “USA” which is the 68 square miles of land west of the Potomac River on which is built the federal capital, Washington DC and the District of Columbia, also the US protectorates of Guam and Puerto Rico. And then there is a “usA” the “united states of America” which are the lands of the various states. These lands are still supposedly owned by the British Crown as the head of the Virginia Company.

The Act of 1871 created a new, altered Constitution. The title was changed to capital letters and the word “for” was changed to the word “of”. That is, “The Constitution for the united states of America” was altered to “THE CONSTITUTION OF THE UNITED STATES OF AMERICA”.

An American cannot use the Constitution to defend himself because he is not a party to it. “The People” does not include ordinary U.S. Citizens. The Act of 1871 also created a separate form of government for the District of Columbia, which is a ten mile square parcel of land which is governed with British Admiralty law, U.C.C. (Acts of the Forty-First Congress, Section 34, Session III, chapters 61 and 62).

When Americans agree to have a Social Security Number, they surrender their sovereignty and agree to become franchisees of the United States (which is supposedly the Virginia Company of the British Crown). Americans are led to believe that there is only one United States and that the Federal government is the rightful government. The use of lower/upper-case characters is making a legal statement. If you are American, then you should notice that when you receive any correspondence relating to government, law or anything to do with finance or taxation that your name is always spelt entirely in block capitals.

Your name in block capitals is not you. It is a corporation set up by the ‘government’ Corporation through the treasury department at the time of your birth. Any time that a child is born, a corporation/trust is created using his or her name in block capitals. One definition of “a human being” is “a corporation treated as having the rights and obligations of a person” while the definition of a “natural person” is ‘a real human being’, as distinguished from a corporation, which is often treated at law as being a fictitious person.

Everything in the “United States” is “For Sale”: roads, bridges, schools, hospitals, water, prisons, airports, etc. (Executive Order 12803). Americans are Human capital (Executive Order 13037).

The Emergency Banking Relief Act of 9th March 1933, defined the American people as legally, the **enemy** of the United States Government and the private, international Federal Reserve System became the Government of the United States.

The national flag of the United States always has a gold fringe when displayed in court or federal buildings, federally-funded schools and the uniforms of US troops. Under the International Law of the Flags, a gold fringe indicates the jurisdiction of commercial law, also known as British Maritime Law, and in the US as the uniform Commercial Code or “UCC”.

George Bush launched a ‘war on terrorism’ on behalf of a private Corporation, to further the goals of that Corporation. It had nothing to do with ‘America’ or ‘Americans’, because those are very different legal entities. The United States Corporation owns the United States military and everything else that comes under the term ‘federal’. It is not the duty of the US police to protect Americans. Their job is to protect the Corporation and arrest code breakers (Sapp v. Tallahassee, 348 So. 2nd. 363, etc.). There are no Judicial courts in America and there have not been since 1789. Judges do not enforce Statutes and Codes but instead, Executive Administrators enforce Statutes and Codes. The most powerful court in America is not the United States Supreme Court but instead, it is the Supreme Court of Pennsylvania (42 Pa.C.S.A. 502). Pennsylvania is the keystone state.

The FCC, CIA, FBI, NASA and all similar organisations were never part of the United States government, even though the “US Government” held shares of stock in the various Agencies. (U.S. v Strang, 254 US 451, Lewis v US, 680 F.2d, 1239).

Americans may think that their government and legal system is pegged in some way to the Constitution, but it is not. The United States, like Britain and elsewhere, is ruled by commercial law to overcome the checks and balances of common law. It is a monumental fraud.

The web site http://www.youtube.com/watch?v=klwWcp9eiPw&feature=player_embedded lists what it considers to be a series of little-known facts about America. It says:

“The IRS is not a U.S. government agency. It is an agency of the IMF which in turn is an agency of the United Nations. The United States has **not** had a Treasury since 1921 as the U.S. Treasury is now the IMF. The United States does not have any employees because there no longer is a United States. The FCC, CIA, FBI, NASA, etc.

were never part of the U.S. government but are privately owned companies. US Social Security numbers are issued by the UN through the IMF, (the SS5 form does not state who issues it). There have been no Judicial Courts in America since 1789. There have been no Judges in America since 1789. According to GATT, Americans must have a Social Security Number. New York City is defined in the Federal regulations as the United Nations. Social Security is not insurance or a contract and Social Security cheques come directly from the IMF which is an agency of the UN.

All Americans are slaves and so can own absolutely nothing, not even what they think of as being their own children. The most powerful court in America is the Supreme Court of Pennsylvania. The King of England financially backed both sides of the Revolutionary War. Americans can't use the US Constitution to defend themselves because they are not a party to it. Britain is owned by the Vatican and the Pope claims to own the entire planet through the laws of conquest and discovery (Papal Bulls of 1495 and 1493). The Pope has ordered the genocide and enslavement of millions of people (Papal Bulls of 1455 and 1493). The Pope's laws are obligatory on all Americans. "The People" does not include ordinary Americans. (Barron v City Council of Baltimore 32 US 243)".

When you become aware of these things, it becomes easier to understand why voting in a Presidential election can never make any difference - after all, he is only the President of a commercial company whose corporate policy is set by the owners of the company and not by the President. It also explains what these commercial companies have done and are still doing through their agencies such as the CIA.

The fractional reserve policy perpetrated by the Federal Reserve, which has spread in practice to the great majority of banks in the world, is, in fact, a system of modern slavery. Think about it: money is created out of debt. What do people do when they earn debt? They submit to employment to pay it off. But if money can only be created out of loans, how can society ever be debt-free? It can't, and that's the point. And it is the fear of losing assets coupled with the struggle to keep up with the perpetual debt and inflation in the system, compounded by the inescapable scarcity in the money supply itself, created by the interest which can never be repaid, that keeps the wage-slave in line, powering the pyramid which benefits only the elite at the top of the pyramid. At the end of the day, for whom are you really working? The banks. Money is created in a bank and inevitably ends up in a bank. They are the true masters along with the corporations and governments which they support. Physical slavery requires people to be housed and fed, economic slavery requires people to house and feed themselves.

Michael Rivero has put together a concise statement on the situation in America, and his statement is at http://whatreallyhappened.com/WRHARTICLES/allwarsarebankerwars.php where he says:

I know that many people have a great deal of difficulty with understanding just how many wars are started for no other purpose than to force private central banks on to nations, so let me share a few examples, in order that that you can understand why the US Government is engaged in so many wars against so many foreign nations. There is ample precedent for this.

The United States fought the war of the American Revolution primarily over King George III's Currency Act, which forced the colonists to conduct their business only using printed bank notes borrowed from the Bank of England at interest. After the revolution, the new United States adopted a radically different economic system in which the government issued its own value-based money, so that private banks like the Bank of England were not siphoning off the wealth of the people through interest-bearing bank notes.

"The refusal of King George 3rd to allow the colonies to operate an honest money system, which freed the ordinary man from the clutches of the money manipulators, was probably the prime cause of the revolution". -- Benjamin Franklin, Founding Father

But bankers are nothing if not dedicated to their schemes to acquire your wealth, and they know how easy it is to corrupt the leaders of a nation. Just one year after Mayer Amschel Rothschild had uttered his infamous statement: *"Let me issue and control a nation's money and I care not who makes the laws"*, the bankers succeeded in setting up a new Private Central Bank called the First Bank of the United States, largely through the efforts of the Rothschild's chief US supporter, Alexander Hamilton.

Founded in 1791, by the end of its twenty year charter the First Bank of the United States had almost ruined the nation's economy, while enriching the bankers. Congress refused to renew the charter and signalled their intention to go back to a state-issued value-based currency on which the people paid no interest at all to any banker. This resulted in a threat from Nathan Mayer Rothschild against the US Government: "Either the application for renewal of the charter is granted, or the United States will find itself involved in a most disastrous war". Congress still refused to renew the charter for the First Bank of the United States, whereupon Nathan Mayer Rothschild railed, "Teach those impudent Americans a lesson! Bring them back to colonial status!"

Financed by the Rothschild-controlled Bank of England, Britain then launched the war of 1812 to re-colonise the United States and force them back into the slavery of the Bank of England, or to plunge the United States into so much debt they would be forced to accept a new private central bank. The plan worked. Even though the War of 1812 was won by the United States, Congress was forced to grant a new charter for yet another private bank issuing the public currency as loans at interest, and that was the Second Bank of the United States. Once again, private bankers were in control of the nation's money supply and cared not who made the laws or how many British and American soldiers had to die for it.

Once again the nation was plunged into debt, unemployment, and poverty through the plundering of the private central bank, and in 1832 Andrew Jackson successfully campaigned for his second term as President under the slogan, "Jackson And No Bank!" True to his word, Jackson succeeds in blocking the renewal of the charter for the Second Bank of the United States.

"Gentlemen! I too have been a close observer of the doings of the Bank of the United States. I have had men watching you for a long time, and am convinced that you have used the funds of the bank to speculate in the breadstuffs of the country. When you won, you divided the profits amongst you, and when you lost, you charged it to the bank. You tell me that if I take the deposits from the bank and annul its charter I shall ruin ten thousand families. That may be true, gentlemen, but that is your sin! If I let you go on, you will ruin fifty thousand families, and that would be my sin! You are a den of vipers and thieves. I have determined to rout you out, and by the Eternal, (bringing his fist down on the table) I will rout you out!" -- Andrew Jackson, shortly before ending the charter of the Second Bank of the United States. From the original minutes of the Philadelphia committee of citizens sent to meet with President Jackson (February 1834), according to Andrew Jackson and the Bank of the United States (1928) by Stan V. Henkels

Shortly after President Jackson (the only American President to actually pay off the National Debt) ended the Second Bank of the United States, there was an attempted assassination which failed when both pistols used by the assassin, Richard Lawrence, failed to fire. Lawrence later said that with Jackson dead, "Money would be more plenty".

Of course, the present day American public school system is as subservient to the bankers' wishes to keep certain items of history from you, just as the corporate media is subservient to Monsanto's wishes to keep the dangers of Genetically Modified foodstuffs from you, and the Global Warming cult's wishes to conceal from you the fact that the Earth has actually been cooling for the last 16 years. Thus is should come as little surprise, that much of the real reasons for the events of the Civil War are not well known to the average American.

When the Confederacy seceded from the United States, the bankers once again saw the opportunity for a rich harvest of debt, and offered to fund Lincoln's efforts to bring the south back into the union, but at 30% interest. Lincoln remarked that he would not free the black man by enslaving the white man to the bankers and using his authority as President, he issued a new government currency, the “greenback”. This was a direct threat to the wealth and power of the central bankers, who responded quickly.

"If this mischievous financial policy, which has its origin in North America, becomes entrenched, then that Government will issue it’s own money without cost. It will pay off debts and be without debt. It will have all the money necessary to carry on it’s commerce. It will become prosperous to a degree which is without precedent in the history of the world. The brains, and wealth of all countries will migrate to North America. That country must be destroyed or it will destroy every monarchy on the globe". -- The London Times responding to Lincoln's decision to issue government Greenbacks to finance the Civil War, rather than agree to private banker's loans at 30% interest.

In 1872 New York bankers sent a letter to every bank in the United States, urging them to fund any newspapers which opposed government-issued money (Lincoln's greenbacks): *"Dear Sir: It is advisable to do all in your power to sustain such prominent daily and weekly newspapers... as will oppose the issuing of greenback paper money, and that you also withhold patronage or favours from all applicants who are not willing to oppose the Government issue of money. Let the Government issue the coin and the banks issue the paper money of the country...To restore to circulation the Government issue of money, will be to provide the people with money, and will therefore seriously affect your individual profit as bankers and lenders".* -- Triumphant plutocracy; the story of American public life from 1870 to 1920, by Lynn Wheeler

"It will not do to allow the greenback, as it is called, to circulate as money any length of time, as we cannot control that". -- Triumphant plutocracy; the story of American public life from 1870 to 1920, by Lynn Wheeler

"Slavery is likely to be abolished by the war, power, and chattel slavery destroyed. This, I and my European friends are in favour of, for slavery is but the owning of labour and carries with it the care for the labourer, while the European plan, led on by England, is for capital to control labour by controlling the wages. THIS CAN BE

DONE BY CONTROLLING THE MONEY". -- Triumphant plutocracy; the story of American public life from 1870 to 1920, by Lynn Wheeler

Goaded by the private bankers, much of Europe supported the Confederacy against the Union, with the expectation that victory over Lincoln would mean the end of the Greenback. France and Britain considered an outright attack on the United States to aid the confederacy, but were held at bay by Russia, which had just ended the serfdom system and had a state central bank similar to the system the United States had been founded on. Left free of European intervention, the Union won the war, and Lincoln announced his intention to go on issuing greenbacks. Following Lincoln's assassination, the Greenbacks were pulled from circulation and the American people forced to go back to an economy based on bank notes borrowed at interest from the private bankers.

Finally, in 1913, the Private Central Bankers of Europe, in particular the Rothschilds of Great Britain and the Warburgs of Germany, met with their American financial collaborators on Jekyll Island, Georgia to form a new banking cartel with the express purpose of forming the Third Bank of the United States, with the aim of placing complete control of the United States money supply once again under the control of private bankers. Owing to hostility over the previous banks, the name was changed to "The Federal Reserve" system in order to grant the new bank a quasi-governmental image, but in fact it is a privately owned bank, no more "Federal" than Federal Express. Indeed, in 2012, the Federal Reserve successfully rebuffed a Freedom of Information Lawsuit by Bloomberg News on the grounds that as a private banking corporation and not actually a part of the government, the Freedom of Information Act did not apply to the operations of the Federal Reserve. The year 1913 proved to be a transformative year for the nation's economy, first with the passage of the 16th "income tax" Amendment and the false claim that it had been ratified.

"I think that if you were to go back and try to find and review the ratification of the 16th Amendment, which was the internal revenue, the income tax, I think that if you went back and examined that carefully, you would find that a sufficient number of States never ratified that Amendment". – U.S. District Court Judge James C. Fox (Sullivan vs United States, 2003).

Later in 1913, apparently unwilling to risk another questionable amendment, Congress passed the Federal Reserve Act over the Christmas holiday, while members of Congress who were opposed to the measure were at home. This was a very underhanded deal, as the Constitution which explicitly grants Congress the authority to issue the public currency, does not authorise it to delegate that authority, and so it should have required a new Amendment to allow Congress to transfer that authority to a private bank. But Congress did pass it, and President Woodrow Wilson signed it (as he had promised the bankers that he would, in exchange for generous campaign contributions). Wilson later regretted that decision, stating: *"I am a most unhappy man. I have unwittingly ruined my country. A great industrial nation is now controlled by its system of credit. We are no longer a government by free opinion, no longer a government by conviction and vote of the majority, but a government by the opinion and duress of a small group of dominant men".* – Woodrow Wilson 1919.

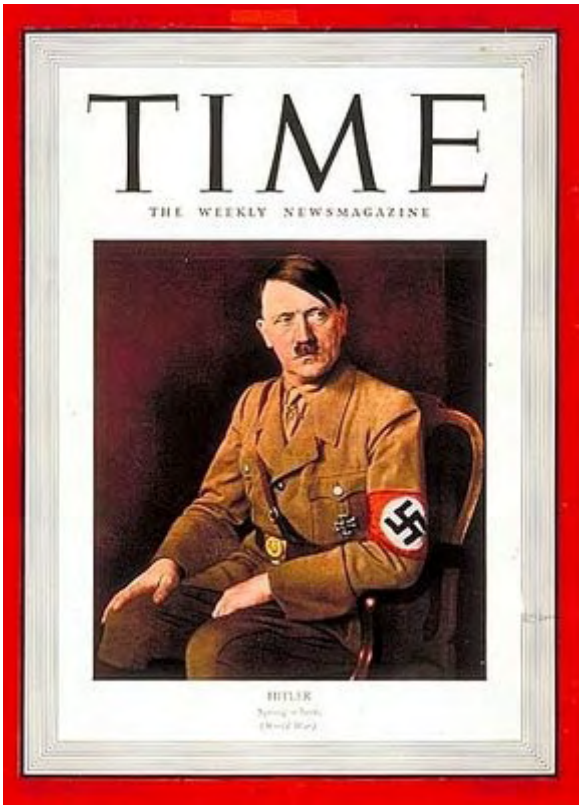
The next year, World War One started, and it is important to remember that prior to the creation of the Federal Reserve, there was no such thing as a world war.

The First World War started between Austria-Hungary and Serbia, but quickly shifted to focus to Germany, whose industrial capacity was seen as an economic threat to Great Britain, who saw the decline of the British Pound as a result of too much emphasis on financial activity to the neglect of agriculture, industrial development, and infrastructure (not unlike the present day United States).

Although pre-war Germany had a private central bank, that bank was heavily restricted, and inflation kept to reasonable levels. Under government control, investment was guaranteed to internal economic development, and Germany was seen as a major power. So, in the media of the day, Germany was portrayed as the prime instigator of World War One, and subsequently, not just defeated, but had its industrial base flattened.

Following the Treaty of Versailles, Germany was ordered to pay the war costs of all of the participating nations, even though Germany had not actually started the war. This amounted to three times the value of all of Germany itself. Germany's private central bank, to whom Germany had gone deeply into debt to pay the costs of the war, broke free of government control, and massive inflation followed (mostly triggered by currency speculators), permanently trapping the German people in endless debt.

When the Weimar Republic collapsed economically, it opened the door for the National Socialists to take power. Their first financial move was to issue their own state currency which was not borrowed from private central bankers. Freed from having to pay interest on the money in circulation, Germany blossomed and quickly began to rebuild its industry. The media called it "The German Miracle". TIME magazine lionized Hitler for the amazing improvement in life for the German people and the explosion of German industry, and even named him TIME Magazine's "Man Of The Year" in 1938.



Once again, Germany's industrial output became a threat to Great Britain, resulting in these comments:

"Should Germany merchandise (do business) again in the next 50 years, we have led this war (WW1) in vain". **- Winston Churchill in The Times (1919)**

"We will force this war upon Hitler, if he wants it or not". **- Winston Churchill (1936 broadcast**)

"Germany becomes too powerful. We have to crush it". **- Winston Churchill** (November 1936 speaking to US - General Robert E. Wood).

"This war is an English war and its goal is the destruction of Germany". **- Winston Churchill (- Autumn 1939 broadcast) .**

Germany's state-issued value-based currency was also a direct threat to the wealth and power of the private central banks, and as early as 1933 they started to organise a global boycott against Germany to strangle this upstart ruler who thought he could break free of private central bankers!

BUY BRITISH ballito STOCKINGS

Daily Express

To-day's Weather: Fair Mild.

NO. 10,258. FRIDAY, MARCH 24, 1933. ONE PENNY.

St. Ivel CHEESE Aids digestion 2d., 6d. & 8½d. each.

JUDEA DECLARES WAR ON GERMANY

Jews Of All The World Unite In Action

BOYCOTT OF GERMAN GOODS

MASS DEMONSTRATIONS IN MANY DISTRICTS

DRAMATIC ACTION

"Daily Express" Special Political Correspondent.

ALL Israel is uniting in wrath against the Nazi onslaught on the Jews in Germany.

Adolf Hitler, swept into power by an appeal to elemental patriotism, is making history of a kind he least expected. Thinking to unite only the German nation to race consciousness he has roused the whole Jewish people to a national renaissance.

The appearance of the swastika symbol of a new Germany has called forth the Lion of Judah, the old battle symbol of Jewish defiance.

Fourteen million Jews dispersed throughout the world have banded together as one man to declare war on the German persecutors of their co-religionists. Sectional differences and antagonisms have been submerged in one common aim—to stand by the 600,000 Jews of Germany who are terrorised by Hitlerist anti-semitism, and to compel Fascist Germany to end its campaign of violence and suppression directed against its Jewish minority.

World Jewry has made up its mind not to rest quiescent in face of this revival of medieval Jew-baiting.

appealing for an end of the Hitler "terror."

Every Rabbi in the City of New York has been placed under a sacred obligation by Rabbinical decree to devote Saturday's sermon to the plight of the Jews in Germany.

The "New York Times" this morning says a list of a thousand German immigrants who have come to the United States during recent years has been compiled by an overseas Nazi organisation, the object being to use these people for Nazi propaganda in the United States.

SPECIAL SESSION

The organisation of Jewish youth in Britain is organising demonstrations in London and the provinces during the week-end.

The Board of Deputies of British Jews, representing the entire Jewish community in Great Britain, is meeting in special session on Sunday to discuss the German situation and decide on what action is to be taken to counteract the attacks on their German fellow Jews.

£811,333 a minute to £4,000,000 a minute.

This is the record of the public's thirst for sound investments with a taxing statute.

When the West Riding County Council issued £1,000,000 of 3½ per cent. stock at £99 per cent. last week they obtained £14,000,000 in fifteen minutes.

Yesterday Middlesbrough asked the public to subscribe to £1,000,000 of its 3½ per cent. stock, and raised the price to £99 10s. per cent.

The lists were closed in five minutes, during which time applications had been received for £20,000,000.

THE BIRTH OF AN IDEA

DENNIS GLENNY, the young British dress designer, finds inspiration for his fashions by draping living models. Every man to his own trade!

OFFICER'S DAYS OF LOVE

HIGHER WAGES FOR STEEL WORKERS

AN INCREASE OF THREE SHILLINGS A WEEK

BRIGHT SPOT IN A BLACK TOWN

THE "Daily Express" has sent a special representative on an unusual mission. He is touring the north in search of high wages. In his tour he is going to tell the facts about all classes of wages, high and low. He is going to show how low wages reduce the morale of the community. He is going to show how courageous spending meets favourably throughout the community.

In his first dispatch it is revealed that the steel masters of Middlesbrough have been able to make a gesture in the cause of high wages.

Fitters in steel works have recently received an increase of 3s. a week. Their wages are not high, but this is a step in the right direction. It is a bright spot in an otherwise black town—brought low by wage cuts and false economy.

CUTS THAT KILL

The Powell Duffryn Steam Coal Company of South Wales increased its profits by £100,000 last year.

Mr. E. L. Hann, the chairman and managing director, told shareholders at yesterday's annual meeting that this had not been achieved by a policy of wage reductions.

In fact he revealed that the earnings of their employees were higher than the average of the coal-field.

New "Sweep" Bill In The Dail

MR. DE VALERA AND STATE CONTROL

SECRET MEASURE

"Daily Express" Special Correspondent.

DUBLIN, Thursday.

WHILE Mr. Jack O'Sheehan, the official announcer of the Irish sweepstake draw, was dividing the remnants of the latest Irish lottery prize fund between lucky ticket-holders in the Plaza Ballroom to-day I was listening to an Under-Secretary of the Free State Government introducing a Bill which may change the whole future, scope, and purpose of these Irish hospitals lotteries.

NO CLUE

A mysterious measure entitled "The Public Hospitals Bill, 1933," appeared on the agenda for this afternoon's meeting of the Dail.

The Bill was set down for formal introduction, and a motion to this effect was moved in a six or seven-words speech by Mr. Sean O'Kelly, the Minister of Local Government.

His brief remarks gave no clue to the

MR. MacDONALD EXPLAINS HIS TOUR

"PEACE CAN BE KEPT IN EUROPE"

MR. RAMSAY MACDONALD faced a crowded House of Commons yesterday afternoon when he spoke about his visits to Paris, Geneva, and Rome, and his talks with Signor Mussolini.

The German, French, Italian, Polish, and Belgian Ambassadors were among a crowd of diplomatic representatives in their special gallery, and Mr. Bruce, of Australia, and other representatives of the Dominions also listened.

Mr. MacDonald, who was greeted with cheers, said:—

"I take the earliest available opportunity to inform the House regarding what happened while the Foreign Secretary and myself were at Geneva and Rome.

"We were asked by our colleagues here to go to Geneva, first because every report showed that there was a grave danger of immediate collapse of the Disarmament Conference, and we felt strongly that, if it might be possible to avert such a collapse, because we were not, at all events, sure that it was necessary that it should take place.

Mr. MacDonald.

"Therefore, the Foreign Secretary and I went to Geneva. Our first business was to discover how matters exactly stood, and for that purpose we interviewed for some two or three days the leading delegates from various nations.

"It is perfectly plain that the conference work had somewhat lost itself in details.

"While we must have the

LATE NEWS

LABOUR LEADER BEREAVED

Broadcasting Programmes on Page Fifteen.

Geneva and discussed candidly and personally an investigation of the condition of affairs there.

"Progress has come far beyond plans for merely reducing things by a third or

As had been the case in World War One, Great Britain and other nations threatened by Germany's economic power looked for an excuse to go to war, and as public anger in Germany grew over the boycott, Hitler foolishly gave them that excuse. Years later, in a spirit of candour, the real reasons for that war were made clear:

"The war wasn't only about abolishing fascism, but to conquer sales markets. We could have, if we had intended to, prevented this war from breaking out without firing one shot, but we didn't want to". **- Winston Churchill to Truman (Fultun, USA March 1946).**

"Germany's unforgivable crime before WW2 was its attempt to loosen its economy out of the world trade system and to build up an independent exchange system from which the world-finance couldn't profit any more. ...We butchered the wrong pig". **- Winston Churchill (The Second World War - Bern, 1960).**

As a side note, we need to step back before WW2 and recall Marine Major General Smedley Butler. In 1933, Wall Street bankers and financiers had bankrolled the successful coups by both Hitler and Mussolini. Brown Brothers Harriman in New York was financing Hitler right up to the day war was declared with Germany. And they decided that a fascist dictatorship in the United States based on the one on Italy would be far better for their business interests than Roosevelt's "New Deal" which threatened massive wealth re-distribution to recapitalise the working and middle class of America. So the Wall Street tycoons recruited General Butler to lead the overthrow of the US Government and install a "Secretary of General Affairs" who would be answerable to Wall Street and not the people, would crush social unrest and shut down all labour unions.

General Butler pretended to go along with the scheme but then exposed the plot to Congress. Congress, then as now in the pocket of the Wall Street bankers, refused to act. When Roosevelt learned of the planned coup he demanded the arrest of the plotters, but the plotters simply reminded Roosevelt that if any one of them were sent to prison, their friends on Wall Street would deliberately collapse the still-fragile economy and blame Roosevelt for it. Roosevelt was thus unable to act until the start of WW2, at which time he prosecuted many of the plotters under the Trading With The Enemy Act. The Congressional minutes into the coup were finally released in 1967 and became the inspiration for the movie, "Seven Days in May" but with the true financial villains erased from the script.

"I spent 33 years and four months in active military service as a member of our country's most agile military force -- the Marine Corps. I served in all commissioned ranks from second lieutenant to Major General. During that

period I spent more of my time being a high-class muscle man for Big Business, for Wall Street and for the bankers. In short, I was a racketeer, a gangster for capitalism. I suspected that I was just a part of a racket at the time. Now I am sure of it. Like all members of the military profession I never had an original thought until I left the service. My mental faculties remained in suspended animation while I obeyed the orders of the higher-ups. This is typical with everyone in the military service. Thus I helped make Mexico and especially Tampico safe for American oil interests in 1914. I helped make Haiti and Cuba a decent place for the National City Bank boys to collect revenues in. I helped in the raping of half a dozen Central American republics for the benefit of Wall Street. The record of racketeering is long. I helped purify Nicaragua for the international banking house of Brown Brothers in 1909-12. I brought light to the Dominican Republic for American sugar interests in 1916. In China in 1927 I helped see to it that Standard Oil went on its way unmolested. During those years, I had, as the boys in the back room would say, a swell racket. I was rewarded with honours, medals and promotion. Looking back on it, I feel I might have given Al Capone a few hints. The best he could do was to operate his racket in three city districts. I operated on three continents". -- General Smedley Butler, former US Marine Corps Commandant,1935

As President, John F. Kennedy understood the predatory nature of private central banking, he understood why Andrew Jackson fought so hard to end the Second Bank of the United States. So Kennedy wrote and signed Executive Order 11110 which ordered the US Treasury to issue a new public currency, the United States Note.

Kennedy's United States Notes were not borrowed form the Federal Reserve but created by the US Government and backed by the silver stockpiles held by the US Government. It represented a return to the system of economics on which the United States had been founded, and this was perfectly legal for Kennedy to do. All told, some four and a half billion dollars went into public circulation, eroding interest payments to the Federal Reserve and loosening their control over the nation. Five months later John F. Kennedy was assassinated in Dallas Texas, and the United States Notes pulled from circulation and destroyed (except for samples held by collectors). John J. McCloy, President of the Chase Manhattan Bank, and President of the World Bank, was named as part of the Warren Commission, presumably to make certain that the banking dimensions behind the assassination were concealed from the public.

As we enter the eleventh year of what future history will most certainly describe as World War Three, we need to examine the financial dimensions behind the wars.

Towards the end of World War Two, when it became obvious that the allies were going to win and dictate the post war environment, the major world economic powers met at Bretton Woods, a luxury resort in New Hampshire in July 1944, and hammered out the Bretton Woods agreement for international finance. The British Pound lost its position as the global trade and reserve currency and its place was taken by the US dollar (part of the price demanded by Roosevelt in exchange for the US entry into the war). Without the economic advantages of being the world's central currency, Britain was forced to nationalise the Bank of England in 1946. The Bretton Woods agreement, was ratified in 1945, and in addition to making the US dollar the global reserve and trade currency, obliged the signatory nations to tie their currencies to the dollar. The nations which ratified Bretton Woods did so on two conditions. The first was that the Federal Reserve would refrain from over-printing the dollar as a means to loot real products and produce from other nations in exchange for ink and paper; basically an imperial tax. That assurance was backed up by the second requirement, which was that the US dollar would always be convertible to gold at $35 per ounce.

Of course, the Federal Reserve, being a private bank and not answerable to the US Government, did start overprinting paper dollars, and much of the perceived prosperity of the 1950s and 1960s was the result of foreign nations' obligations to accept the paper notes as being worth gold at the rate of $35 an ounce. Then in 1970, France looked at the huge pile of paper notes sitting in their vaults, for which real French products like wine and cheese had been traded, and notified the United States government that they would exercise their option under Bretton Woods to return the paper notes for gold at the $35 per ounce exchange rate. Of course, the United

States had nowhere near the gold to redeem the paper notes, so on August 15th, 1971, Richard Nixon "temporarily" suspended the gold convertibility of the US Federal Reserve Notes.

This "Nixon shock" effectively ended Bretton Woods and many global currencies started to disengage from the US dollar. Worse still, since the United States had collateralised their loans with the nation's gold reserves, it quickly became apparent that the US Government did not in fact have enough gold to cover the outstanding debts. Foreign nations began to get very nervous about their loans to the US and understandably were reluctant to lend any additional money to the United States without some form of collateral. So Richard Nixon started the environmental movement, with the EPA and its various programs such as "wilderness zones", Road-less areas", Heritage rivers", "Wetlands", all of which took vast areas of public lands and made them off limits to the American people who were technically the owners of those lands. But Nixon had little concern for the environment and the real purpose of this land grab under the guise of the environment was to pledge those pristine lands and their vast mineral resources as collateral on the national debt. The plethora of different programs was simply to conceal the true scale of how much American land was being pledged to foreign lenders as collateral on the government's debts; eventually almost 25% of the nation itself.

With open lands for collateral already in short supply, the US Government embarked on a new program to shore up sagging international demand for the dollar. The United States approached the world's oil producing nations, mostly in the Middle East, and offered them a deal. In exchange for only selling their oil for dollars, the United States would guarantee the military safety of those oil-rich nations. The oil rich nations would agree to spend and invest their US paper dollars inside the United States, in particular in US Treasury Bonds, redeemable through future generations of US taxpayers. The concept was labelled the "petrodollar". In effect, the US, no longer able to back the dollar with gold, was now backing it with oil. Other peoples' oil, and the necessity to keep control over those oil nations in order to prop up the dollar has shaped America's foreign policy in the region ever since.

But as America's manufacturing and agriculture has declined, the oil producing nations faced a dilemma. Those piles of US Federal Reserve notes were not able to purchase much from the United States because the United States had little (other than real estate) which anyone wanted to buy. Europe's cars and aircraft were superior and less costly, while experiments with GMO food crops led to nations refusing to buy US food exports. Israel's constant belligerence against its neighbours caused them to wonder if the US could actually keep their end of the petrodollar arrangement. Oil-producing nations started to talk of selling their oil for whatever currency the purchasers chose to use. Iraq, already hostile to the United States following Desert Storm, demanded the right to sell their oil for Euros in 2000 and in 2002, the United Nations agreed to allow it under the "Oil for food" program instituted following Desert Storm. One year later the United States re-invaded Iraq, lynched Saddam Hussein, and placed Iraq's oil back on the world market only for US dollars.

Following 9-11, the clear US policy shift away from being an impartial broker of peace in the Middle East to one of unquestioned support for Israel's aggressions, only further eroded confidence in the Petrodollar deal and even more oil-producing nations started openly talking of oil trade for other global currencies.

Over in Libya, Muammar Gaddafi had instituted a state-owned central bank and a value-based trade currency, the Gold Dinar. Gaddafi announced that Libya's oil was for sale, but only for the Gold Dinar. Other African nations, seeing the rise of the Gold Dinar and the Euro, even as the US dollar continued its inflation-driven decline, flocked to the new Libyan currency for trade. This move had the potential to seriously undermine the global hegemony of the dollar. French President Nicolas Sarkozy reportedly went so far as to call Libya a “threat” to the financial security of the world.

So, the United States invaded Libya, brutally murdered Gaddafi (the object lesson of Saddam's lynching not being enough of a message, apparently), imposed a private central bank, and returned Libya's oil output to dollars only. The gold that was to have been made into the Gold Dinars is, as of last report, unaccounted for.

According to General Wesley Clark, the master plan for the "dollarification" of the world's oil nations included seven targets, Iraq, Syria, Lebanon, Libya, Somalia, Sudan, and Iran (Venezuela, which dared to sell their oil to China for the Yuan, is a late addition). What is notable about the original seven nations originally targeted by the US is that none of them are members of the Bank for International Settlements, the private central bankers private

central bank, located in Switzerland. This meant that these nations were deciding for themselves how to run their nations' economies, rather than submit to the international private banks.

Now the gun sights of the bankers are on Iran, which dares to have a government central bank and sell their oil for whatever currency they choose. The war agenda is, as always, to force Iran's oil to be sold only for dollars and to force them to accept a privately owned central bank.

The German government recently asked for the return of some of their gold bullion from the Bank of France and the New York Federal Reserve. France has said it will take 5 years to return Germany's gold. The United States has said they will need 8 years to return Germany's gold. This suggests strongly that the Bank of France and the NY Federal Reserve have used the deposited gold for other purposes, and they are scrambling to find new gold to cover the shortfall and prevent a gold run. So it is inevitable that suddenly France invades Mali, ostensibly to combat Al Qaeda, with the US joining in. Mali just happens to be one of the world's largest gold producers with gold accounting for 80% of Mali exports. War for the bankers does not get more obvious than that!

Americans have been raised by a public school system and media that constantly assures them that the reasons for all these wars and assassinations are many and varied. The US claims to bring democracy to the conquered lands (they haven't; the usual result of a US overthrow is the imposition of a dictatorship, such as the 1953 CIA overthrow of Iran's democratically elected government of Mohammad Mosaddegh and the imposition of the Shah, or the 1973 CIA overthrow of Chile's democratically elected government of President Salvador Allende, and the imposition of Agusto Pinochet), or to save a people from a cruel oppressor, revenge for 9-11, or that tired worn-out catch all excuse for invasion: “weapons of mass destruction”. Assassinations are always passed off as "crazed lone nuts" to obscure the real agenda.

The real agenda is simple. It is enslavement of the people by the creation of a false sense of obligation. That obligation is false because the Private Central Banking system, by design, always creates more debt than money with which to pay that debt. Private Central Banking is not science, it is a religion; a set of arbitrary rules created to benefit the priesthood, meaning the owners of the Private Central Bank. The fraud persists, with often lethal results, because the people are tricked into believing that this is the way life is supposed to be and no alternative exists or should be dreamt of. The same was true of two earlier systems of enslavement, ‘Rule by Divine Right’ and ‘Slavery’, both of which are systems designed to trick people into obedience, and both of which are now recognised by modern civilization as illegitimate. Now we are entering a time in human history where we will recognise that ‘rule by debt’, or rule by Private Central Bankers issuing the public currency as a loan at interest, is equally illegitimate. It only works as long as people allow themselves to believe that this is the way life is supposed to be.

But understand this above all else, Private Central Banks do not exist to serve the people, the community, or the nation. Private Central Banks exist to serve their owners, to make them rich beyond the dreams of Midas and all for the cost of ink, paper, and the right bribe to the right official.

Behind all these wars, all these assassinations, the hundred million horrible deaths from all the wars lies a single policy of dictatorship. The private central bankers allow rulers to rule only on the condition that the people of a nation remain enslaved to the private central banks. Failing that, any ruler will be killed, and their nation invaded by those other nations which are already enslaved to private central banks.

The so-called "clash of civilizations" we read about on the corporate media is really a war between banking systems, with the private central bankers forcing themselves on to the rest of the world, no matter how many millions must die for it. Indeed the constant hatemongering against Muslims lies in a simple fact. Like the ancient Christians (prior to the Knights Templar's private banking system), Muslims forbid usury (the lending of money at interest), and that is the reason why the American government and media insist that Muslims must be killed or converted. They refuse to submit to currencies issued at interest. They refuse to be debt slaves. So off to war American children must go, to spill their blood for the gold of the money-junkies. We barely survived the last two world wars. In the nuclear/bio weapon age, are the private central bankers willing to risk incinerating the whole planet just to feed their greed? Apparently so.

Flag waving and propaganda aside, all modern wars are wars by and for the private bankers, fought and bled for by third parties unaware of the true reason why they are expected to be killed and crippled. The process is quite simple. As soon as the Private Central Bank issues its currency as a loan at interest, the public is forced deeper and deeper into debt. When the people are reluctant to borrow any more, that is when the Keynesian economists demand the government borrow more to keep the pyramid scheme working. When both the people and government refuse to borrow any more, that is when wars are started, to plunge everyone even deeper into debt to pay for the war, then after the war to borrow more to rebuild. When the war is over, the people have about the same as they did before the war, except the graveyards are far larger and everyone is in debt to the private bankers for the next century. This is why Brown Brothers Harriman in New York was funding the rise of Adolf Hitler.

As long as Private Central Banks are allowed to exist, inevitably, as night follows day, there will be poverty, hopelessness, and millions of deaths in endless World Wars, until the Earth itself is sacrificed in flames to Mammon. The path to true peace on Earth lies in the abolishment of all private central banking everywhere, and a return to the state-issued value-based currencies that allow nations and people to become prosperous.

Michael Rivero

The Economic Hit-Men

It is one of the most ingenious scams for social manipulation ever invented, and at it's core, it is an invisible war waged against the population, and you will remember that the Emergency Banking Relief Act of 9th March 1933, **defined** the American people as "the **enemy"** of the United States Government. Debt is the weapon used to conquer and enslave societies and interest is its prime ammunition. While the majority of people are not aware of this reality, banks in collusion with governments and corporations continue to expand and perfect their tactics of economic warfare, spawning new bases such as the World Bank and the International Monetary Fund and introducing a new kind of soldier - the economic hit-man.

Here is the testimony of one of those American economic hit-men, John Perkins:

We, the economic hit men, are the ones really responsible for creating the first really global empire, and we work many different ways. Perhaps the most common is that we will identify a country which has resources which our corporations covet, such as oil, and then arrange a huge loan to that country from the World Bank or one of it's sister organisations. But the money never actually goes to the country, instead, it goes to our big construction corporations to build infrastructure projects in that country, like industrial power plants, things which benefit a few rich people in that country (in addition to our corporations), but really don't help the people involved, however those people and the whole country are left holding a huge debt so big that they can't repay it, which is the whole plan - that they can't repay it. So we economic hit men go back to them and say "listen, you owe us a lot of money, you can't repay it, so sell your oil real cheap to our oil companies, allow us to build a military base in your country or send troops to some place in the world to support our troops in somewhere like Iraq, or vote with us in the UN to have their electric utility company privatised, or their sewage system privatised and sold to US corporations or other multinational corporations, so there is a whole mushrooming thing and it's the way that the World Bank and the International Monetary Fund operate to put a whole country in debt with such a big debt that it can't pay it, so you then offer to refinance that debt and get them to pay even more interest and you demand this

'quid pro quo' or 'conditionality' or 'good governance' which means basically that they have to sell off their resources, including their social services and utility companies, their school systems sometimes, their penal systems, their insurance systems to foreign corporations. So, it's a double, triple, quadruple whammy.

Examples:

Iran 1953: The precedent for economic hit men began back in the early 1950s when the democratically elected Prime Minister of Iran, Dr Mohammed Mossadegh, was considered to be a hope for democracy - Time Magazine's Man of the Year. One of the policies which he ran on was the idea that the oil companies needed to pay the Iranian people a lot more for the oil which they were taking out of Iran. We didn't like that, of course, but we were afraid to do what we would normally do, which is to send in the military. Instead, we sent in one CIA agent, a relative of President Roosevelt. He went in with a few million dollars and was very efficient and effective and in a short amount of time he managed to get Mossadegh overthrown and brought in the Shah of Iran to replace him. It was extremely effective. So back here in the United States, people in Washington looked around and said "wow - that was easy and cheap !" So this established the whole new way of manipulating countries and creating empire. The only problem with Roosevelt was that he was a card-carrying CIA agent and if he had been caught, the ramifications would have been pretty serious. So, very quickly at that point a decision was made to use private 'consultants', to channel the money through the World Bank, the IMF or one of the other such agencies, bring in people like me who worked for private companies, so that if we got caught, there would be no government ramifications.

Guatemala 1954: When Arbenz (Jacobo Arbenz Guzman) became President of Guatemala, the country was very much under the thumb of the United Fruit Company, the big international corporation and Guzman ran under the strategy of giving the land back to the people. When he was elected he started implementing those policies of giving the land back to the people. United Fruit didn't like that so much, so they hired a Public Relations company in the United States mounted a huge campaign to convince the people of the US, the press of the US and the Congress of the US that Arbenz was a Soviet puppet and if we allowed him to stay in power the Soviets would have a foothold in this hemisphere, which, at that time, was a huge fear on everybody's minds, so to make a long story short, out of this public relations campaign came a commitment on the part of the CIA and the military to take this man out, and in fact, we did. We sent in planes, we sent in soldiers, we sent in jackals, we sent in everything

to take him out, and we did take him out. And as soon as he was removed, his successor reinstated the links to the big corporations including United Fruit.

Ecuador 1981: Ecuador for many, many years was ruled by pro-US dictators, often relatively brutal. Then it was decided that they were going to have a truly democratic election Jaime Roldos (Aguilera) ran for office and his main goal, he said, as President was to make sure that Ecuador's resources were used to help the people. And he won, overwhelmingly, by more votes than anyone has ever won anything in Ecuador and he began to implement these policies to make sure that the profits from oil went to help the people. Well, we didn't like that in the United States. I was sent down as one of several economic hit men to change Roldos, to corrupt him, to bring him around, to let him know "you know, you and your family can get very rich if you play our game but if you but if you continue to try and keep these policies you promised, then you are going to go". He wouldn't listen. He was assassinated "Ecuadorian Leader Dies in Plane Crash". As soon as the plane crashed, the whole area was cordoned off. The only people allowed in were the US military from a nearby base and some of the Ecuadorian military. When an investigation was launched, two of the key witnesses died in car accidents before they had a chance to testify. A lot of very, very strange things went on around the assassination of Roldos. I, like most of the people who really looked at this case had absolutely no doubt that it was an assassination and of course, in my position of an economic hit man, I was expecting something to happen to Jaime, whether a coup or an assassination I was sure that he would be taken down as he was not being corrupted, he would not allow himself to be corrupted the way we wanted to corrupt him.

Panama 1981: Omar Tarrijas President of Panama was one of my favourite people. I really, really liked him, he was very charismatic and he really wanted to help his country. When I tried to bribe him, to corrupt him, he said "look, John, he called me Juanita, look Juanita, you know, I don't need the money, what I really need is for my country to be treated fairly. I need the United States to replay the debts which you owe my people for all the

destruction you have done here. I need to be in a position to help other Latin American countries with their independence and be free of this terrible presence from the North. You people are exploiting us so badly. I need to have the Panama Canal back in the hands of the Panamanian people. That's what I want. So leave me alone, don't try to bribe me". It was in 1981 and in May Jaime Roldos was assassinated and Omar was very aware of this. He got his family together and said "I'm probably next, but it is ok because I've done what I came here to do, I've renegotiated the Canal, the Canal will now be in our hands". He had been renegotiating the treaty with Jimmy Carter:

In June of that same year, just a couple of months later, he also went down in an aeroplane crash which there is no question was executed by CIA jackals. There is a tremendous amount of evidence, one of Tarrijas' security guards handed him, at the last moment as he was getting on the plane, a small tape recorder which contained a bomb.

Venezuela 2002: It is interesting to me how this system has continued pretty much the same way for years and years and years except the economic hit men get better and better and better. Then we come up with, very recently, what happened in Venezuela in 1998 when (Hugo Rafael) Chavez gets elected, following a long line of Presidents who were very corrupt and who basically destroyed the economy of the country, and Chavez was elected following all of that. Chavez stood up to the United States demanding that Venezuela oil be used to help the Venezuelan people. Well we didn't like that in the United States, so in 2002, a coup was staged, and there is no question in my mind or in most other people's minds, that the CIA was behind that coup. The way that coup was fermented was very effective, like Colonel Roosevelt had done in Iran - paying people to go out into the streets to riot, to protest and say that Chavez is very unpopular. You know, if you can get a few thousand people to do that, television can make it look like the whole country and things start to mushroom, except in the case of Chavez, he was smart enough and the people were so strongly behind him that they overcame it, which was a phenomenal moment in the life of Latin America.

Iraq 2003: Iraq, actually, is a perfect example of the way the whole system works. We economic hit men are the first line of defence. We go in and try to corrupt governments and get them to accept these huge loans which we then use as leverage to basically own them. If we fail, as I failed in Panama with Omar Tarrijas, and in Ecuador with Jaime Roldos, men who refused to be corrupted, then the second line of defence is we send in the jackals and the jackals either overthrow governments or they assassinate and when that happens and a new government comes in, boy, it's going to toe the line as the new President knows what will happen if he doesn't. And in the case of Iraq, both of those things failed and the economic hit men were not able to get through to Saddam Hussein although they tried very hard. We tried very hard to get him to accept a deal but he wouldn't accept it and so the jackals went in to take him out but they couldn't do it as his security was very good. After all, he had at one time worked for the CIA and been hired to assassinate a former President of Iraq and failed, but he knew the system. So, in 1991 we send in the troops and we take out the Iraqi military, assuming that at that point that Saddam is going to come around. We could, of course, have taken him out at that time, but we didn't want to as he is the kind of strong man that we like, able to control his people and keep pumping oil for us. But when we took out his military he didn't come around so the economic hit men go back in the 90s without success. If they had had success, he'd still be running the country - we'd be selling him all the fighter jets he'd want, but they didn't have success. The jackals couldn't take him out again, so we sent the military in once again and took him out, and created for ourselves very lucrative reconstruction deals to rebuild a country which we had essentially destroyed - which is a very good deal if you own a construction company - a really big one. So, Iraq shows the three stages: the economic hit men - failed there, the jackals failed there, so as the final measure, the military goes in.

And in that way, we have really created an empire, but we have done it very, very subtly, it's clandestine. All the empires in the past were built by the military and everybody knew they were building them. The British knew they were building it, the French, the Germans, the Romans, the Greeks - they were all proud of it and they always had some excuse like "spreading civilisation", or spreading some religion, something like that, but they knew they were doing it. We don't. The majority of the people in the United States have no idea that we are living off the benefits of a clandestine empire, that today there is more slavery in the world than ever before.

You may have to ask yourself if it's an empire, then who's the Emperor? Obviously, our Presidents in the United States are not Emperors. An Emperor is someone who is not elected, does not serve a limited term, and who doesn't report to anyone. So you can't classify Presidents that way. But we do have what I consider to be the equivalent of the Emperor in what I call the "Corporatocracy". The Corporatocracy is this group of individuals who run our biggest corporations and they really act as the Emperor of this empire. They control our media (either by direct ownership or by advertising), they control most of our politicians because they finance their campaigns, either through corporations or through personal donations. They are not elected, they don't serve a limited term, they don't report to anybody, and at the very top of the Corporatocracy, we can't tell if the person is working for a corporation or for the government as they are always moving back and forth, so you get a guy who is one moment the president of a big company like Halliburton and the next moment he's the Vice President of the United States, or the President who is in the oil business and this is true whether you get Democrats or Republicans in the office, you have them moving back and forth through the revolving door, and in a way, our government is invisible a lot of the time as it's policies are carried out by a corporation on one level or another, and then again, the policies of the government are basically forged by the Corporatocracy and then presented to the government and become government policies. So, it is an incredibly cosy relationship. This isn't a "conspiracy theory" kind of thing, these people don't have to get together and plot to do things, they all basically work under one primary assumption, and that is that they must maximise profits regardless of the social and environmental costs.

This process of manipulation by the Corporatocracy, through the use of debt, bribery and political overthrow, is called "Globalisation". Just as the Federal Reserve keeps the American people in a position of indentured

servitude, through perpetual debt, inflation and interest, the World Bank and International Monetary Fund fulfil this role on a global scale.

Oh, and by the way, the lame excuse for going to war with Iraq which was never believable beforehand, is laid out in fair detail here by a high-ranking US Major General expert in gathering and interpreting military and terrorist intelligence: http://www.youtube.com/watch?v=W0xzsbSbVUE&feature=player_embedded.

Libya 2011: If the ruler of a country refuses to become the servant of the New World Order people, he is subjected to character assassination and physically attacked with troops. I don't for a moment suggest that Col. Gaddafi was a saint, but as the Bible so aptly remarks, you can tell a man's character by the things which he does, so what did Col. Gaddafi do? Well, here is some information passed to me and I leave you to decide for yourself:

1. In Libya, electricity is free for all of it's citizens.

2. There is no interest on loans, banks in Libya are state-owned and loans are given to all its citizens at 0% interest which is requirement of Libyan law.

3. Owning a home is considered a human right in Libya – Gaddafi vowed that his parents would not get a house until everyone else in Libya had a home. Gaddafi's father has died while he, his wife and his mother were still living in a tent.

4. All newlyweds in Libya receive $60,000 Dinar (US$50,000) from the government, to buy their first apartment to help them start a family.

5. Education and medical treatments are free in Libya. Before Gaddafi only 25% of Libyans were literate. Today the figure is 83%.

6. If any Libyan wants to take up farming, he receives farming land, a farm house, equipment, seeds and livestock to kick- start their farms – all for free.

7. If Libyans cannot find the education or medical facilities they need in Libya, the government funds them to go abroad for it – not just free, but they get US$2, 300 per month for accommodation and running a car.

8. If a Libyan buys a car, the government pays 50% of the price.

9. The price of petrol in Libya is US$0.14 per litre (in the UK, it is $2.17 per litre at this time, which is more than 15 times as much).

10. Libya has no external debt and its reserves amount to US$150 billion – now frozen globally.

11. If a Libyan is unable to get employment after graduation, the state pays the average salary of the profession as if he or she is employed, until employment is found.

12. A portion of Libyan oil sale revenues is credited directly to the bank accounts of all Libyan citizens.

13. A mother who gives birth to a child receives US$5 ,000.

14. 40 loaves of bread in Libya costs just US$ 0.15 (in the UK, it is 200 times more expensive).

15. 25% of Libyans have a university degree, while in the UK, Sir Clive Woodhead, former Chief Inspector of Schools, has recently recommended that children should leave school at age 14 and become apprentices.

16. Col. Gaddafi carried out the world's largest irrigation project, known as the Great Man-Made River project, to make water readily available throughout his desert country.

So, I will leave it up to you to decide if the attacks on Libya and the killing of Col. Gaddafi was an honest action designed to make the world a safer and better place, or if those actions were motivated by some other agenda.

The Technique:

The lawyers and bankers have the objective of taking all of your money and keeping you permanently in debt to them. In the UK, the currency was "the pound sterling" and a £1 note was a promise certifying that there was literally a one pound weight of sterling silver held in storage to back the value of that note and save the person carrying around the actual pound weight of silver.

That link was broken, and nowadays, the £1 note (or in the US, the $1 note) has no actual value as it is no longer a promise that an amount of valuable metal is being held in storage in order to give that piece of paper it's value. In reality, these paper currency notes have no more value than the paper notes in the board game "Monopoly" and that value is only the value of the paper on which the notes are physically printed, which is, almost nothing at all.

Take the process in America as an example: The "government" decides that it needs some money, so it calls up the (privately owned) Federal Reserve Bank and requests, say, $10,000,000,000. The Federal Reserve Bank replies saying "Sure, we will buy $10 billion in government Bonds from you". So the "government" takes some pieces of paper, paints some official-looking designs on them and calls them "Treasury Bonds". Then it puts a value on them of $10 billion and sends them over to the Federal Reserve Bank. In turn, the people in the Federal Reserve Bank draw up a series of impressive looking pieces of paper themselves, only, this time, calling them Federal Reserve notes, also designated with a value of $10 billion. The Federal Reserve Bank then takes these notes and trades them for the Treasury Bonds. Once this trade is complete, the government then takes the $10 billion in Federal Reserve notes and deposits them in a bank account, and on doing this, the $10 billion becomes legal tender money, adding $10 billion to the US money supply. And there it is: $10 billion in new money has been created out of nothing and with no intrinsic value whatsoever. Of course, this example is a generalisation because in reality, this transaction would occur electronically, with no paper used at all. In fact, only 3% of the US money supply exists in physical currency and the other 97% exists in computers alone.

The same style of operation happens in the UK as well with the (privately owned) Bank of England, producing valueless 'currency'. The banks deliberately compound the problem by lending this valueless money to individual people as a loan or mortgage, knowing full well that there is not enough money in circulation to pay off the loan and the interest. If, for example, the loan is £1000 on a credit card, then the credit card company will want about £1200 back as they charge interest. But as they do that with all of the "money" fraudulently created (actually ten times the amount in America), there is physically not enough to allow the interest to be paid unless the economy of the country keeps on expanding rapidly. The objective here is to take physical goods and possessions away from the borrower when he finds himself unable to pay the amount asked for.

The total lack of value of present day currency was demonstrated when a mortgage loan repossession was challenged in court in America. For any commercial contract anywhere in the world to be valid, something of worth has to be put forward by both parties to that contract. When a bank "lends" it's worthless money as a mortgage, it has put forward nothing of value and so the mortgage contract is not valid in law.

In 1969, there was a Minnesota court case involving a man named Jerome Daly who was challenging the foreclosure of his home by the bank which provided the loan to purchase it. His argument was that the mortgage contract required both parties (himself and the bank) to put up a legitimate form of property for the exchange. In legal language, this is called a "consideration" from both parties. Mr Daly explained that the money was, in fact, not the property of the bank as it was created out of nothing as soon as the loan agreement was signed. In other words, the money didn't come out of the bank's existing assets as the bank simply invented it, consequently, the bank put up nothing of value of it's own, except for a theoretical liability on paper. As the court case progressed, the bank's President, Mr Morgan, took the stand and the judges memorandum records that the "Plaintiff admitted that it, in combination with the Federal Reserve Bank ... did create the entire $14,000 in money in credit upon it's own books by bookkeeping entry ... the money and credit first came into existence when they created it. Mr Morgan admitted that no United States Law or Statute existed which gave him the right to do this. A lawful

consideration must exist and must be tendered to support the Note" ... The jury found that there was no lawful consideration, and on this finding, the court rejected the bank's application for foreclosure and Mr Daly kept his home.

The Consequences:

There does not seem to be much point in commenting on the situation in other countries around the world as you will find the same style of underhand operation in most of them. However, when you understand what is going on, it explains why it has been so difficult to get any free-energy device on to the market for sale to members of the public.

A really major strand of both income and control comes from access to energy. The average person wants energy for heating and cooling a home, running household equipment such as refrigeration, lighting, computing, etc. and for powering a vehicle, operating power tools, and so on. Under the present deceptive regime, we are told from an early age that we need to burn a fuel in order to get this power, and instead of doing it individually, it appears to be cheaper to buy the power from the local energy company.

The reality is that we live in a sea of energy which can supply our every need for ever and ever without the slightest need for burning any kind of fuel. Devices for tapping this energy have been produced by literally thousands of inventive people and yet, they are not for sale. Why do you think that is?

Well, one of the reasons is that inventors of these devices often feel that they should "protect" their invention by patenting it and so they make an application to a Patent Office. Unfortunately, the Patent Office is a commercial business owned by the same people who make massive amounts of money from selling you oil to burn and electricity to use. Because of that, they steal the best of these inventions, often by telling the inventor that his invention is "Of National Security Importance" and so he can't use it, sell it or even speak to other people about it, and if he did, he would be put in prison. That is, of course, an unlawful deception as the people telling him that have nothing whatsoever to do with anything National, except the use of the word in their company name. There certainly is no government around to make any kind of statement like that.

The wonderful people then use your money in unlimited amounts to develop these inventions as secret "black" projects and they have been doing this for at least a hundred years. People who have been employed in some of these projects report that they now have technology anything from 100 to 1,000 years in advance of what the public is told about. Almost everything shown in the fictional "Star Trek" series has already been produced covertly in real life. The internet has made it difficult to keep on hiding these advances, but they fight back by posting false information, contradicting people who present the real deal, and generally disrupting and flooding every useful forum with a combination of rubbish and inflammatory attacks on forum members.

A major strand of the attempt to suppress the truth is through universities and respected Engineering bodies and publications. This is done through financial pressure. To have a successful career, a university professor needs to have regular publication of research papers. For this, he needs researchers who are paid to do the research. This is done through grants from rich people and organisations. Those are the very people who will under no circumstances allow the publication of any research which demonstrates that free-energy exists and they make sure that no funding will ever be available for any such research.

If a scientist steps out of line and goes public with the facts, he is immediately attacked by all of his former colleagues who promptly say that what he says is "impossible" and "breaks the laws of physics", neither of which are actually true. For example, in 2010, Rosemary Ainslie and her colleagues demonstrated under strict laboratory conditions that a pulsed heating element produces 17 times more heat output than could be produced by the tiny electrical input driving the element. This was written up in standard scientific paper format and passed to the Institute of Electrical Engineers for examination and publication. As was expected, the paper was refused in spite of the fact (or perhaps because of the fact) that the laboratory results were conclusive. None of the present scientific institutions will publish any form of free-energy paper as that would not be good for the business of their owners.

If an inventor skips the Patent Office, avoids the scientific journals, and tries to get his invention manufactured, he needs substantial funding and that is usually blocked by the concerted efforts of the groups of companies whose businesses would be harmed by the invention becoming freely available to the public. Some times, the inventor is threatened and intimidated. I personally know several people who have had this done to them, but the most common block is financial. For example, Cal-Tech in America spent a million dollars developing, testing, proving and perfecting a carburettor which gave vastly improved mpg results and which cut pollution to near zero. They teamed up with the very large auto-parts supplier Arvin Meritor to put the carburettor into production vehicles. Then "One Equity Partners" bought out the Arvin Meritor division which did all the final work, and then they created a new company, "EMCON Technologies", which dropped the carburettor from their product line, not

because it did not work but because it did work. Oil sales were the only interest and cutting pollution was seen as irrelevant.

Web Links

If you want to research this fraudulent commercial domination over you, then here are some links which might be helpful:

http://www.youtube.com/user/ThriveMovement How To Thrive: A clear and comprehensive explanation of the present situation and what YOU can do about it (very highly recommended).

http://www.youtube.com/watch?v=AXAnba23qV4 Mortgages are Fake: A recent video demonstration that mortgages are fake and fraudulent.

http://www.truth-now.net/case1.htm Australian Fakes: A video demonstrating that Australia has no lawful Courts, Police, Customs, Parliament, etc. from www.truth-now.net

http://thecrowhouse.com/Documents/mary-book.pdf Mary Croft's eBook "How I Clobbered Every Bureaucratic Cash-Confiscatory Agency Known to Man"

http://www.documentarywire.com/john-harris-its-an-illusion A lecture by John Harris entitled "It's An Illusion"

http://www.tpuc.co.uk/ John Harris' web site "The People's United Community"

http://www.zshare.net/audio/5485964026c8ef48 An audio interview with Thomas Anderson – part 1

http://www.zshare.net/audio/548937360de2ee12 An audio interview with Thomas Anderson – part 2

http://educate-yourself.org/cn/strawmanillusion02apr10.shtml The Strawman Illusion Explained

http://panacea-bocaf.org/honordishonor.htm Panacea's strawman information web page

http://video.google.com.au/videoplay?docid=3296715122664269567# The "Hi-jacking Humanity" video

http://www.youtube.com/watch?v=YHZTjTmrgjI Interview with Mary Croft, part 1

http://www.youtube.com/watch?v=alqgpuAwpzM&NR=1 Interview with Mary Croft, part 2

http://www.youtube.com/watch?v=8hqnaHNw_fo&feature=related Interview with Mary Croft, part 3

http://www.youtube.com/watch?v=P9zxVMNmJsE&feature=related Interview with Mary Croft, part 4

http://www.youtube.com/watch?v=iWnv66ERXiw&feature=related Interview with Mary Croft, part 5

http://www.youtube.com/watch?v=BUD5f0MWnKo&feature=related Interview with Mary Croft, part 6

http://www.archive.org/details/Michael_Badnarik Seven hours of intensive training by Michael Badnarik

http://web.archive.org/web/20060407062015/http://www.worldnewsstand.net/law/PLAY_BALL.htm The Commerce Game Exposed

http://famguardian.org/TaxFreedom/Instructions/3.17QuitSocialSecurity.htm Tax Freedom in America

http://www.getoutofdebtfree.org/ Get Out of Debt Free website

http://video.google.com/videoplay?docid=-2550156453790090544# The "Money As Debt" video

http://video.google.com/videoplay?docid=-5153195602561839 36# "The Money Masters" video

http://www.thebcgroup.org.uk/video The British Constitution Group videos

http://web.archive.org/web/20060206085143/www.worldnewsstand.net/law/REDEMPTION3.htm The Application of Commercial Law

The Practical Situation

Some of this is quite difficult to absorb and understand, so perhaps a little extra detail may be helpful. Just to recap; there are two separate and totally different systems: the law and the legal system.

The law boils down to “don’t harm anyone or his property, steal his property or swindle him”. That is pretty straightforward and it is binding on everyone. You will notice that it does not say anything about keeping to a speed limit, paying taxes, getting married, serving in an army, voting in an election, or any one of thousands of other things. Those things are not required by the law. The requirements of the law are “don’t harm anyone, damage or steal his property or defraud him”, nothing else.

The legal system is an invention intended to do exactly the opposite, namely, harm you, steal your property, subject you to the will of someone else and to defraud you generally. The operating method is lies and deception and a prolonged attempt to prevent you finding out the facts. Thankfully, being subject to the legal system is optional – completely your own choice. So, by refusing to consent to being subject to it, you can shrug it off and live a life free from it.

One of the really difficult things to understand is finance. For example: Al Baker, decides to set up a finance company, lending money and charging interest on the loans. He registers “The Swindle Finance Company Limited”, which is a paper invention of the legal system and is not actually a real thing at all – just some pieces of paper and a name written into a list of paper companies.

His friend Charles Daly who lives across the street, decides to deposit £10,000 (or dollars, or Euros, or whatever the local currency is called) with Al because Al is offering 10% annual interest on all deposits as a special advertising promotion. Al puts the money in the bank account of Swindle Finance Company Limited and enters the amount as a credit in the company accounts ledger. Big smile – the business is starting to roll in.

Gary Hughes then decides that he could do with a round-the-world holiday cruise and so he asks Al for a loan of £10,000. Al is delighted and is about to lend him the £10,000 which Charles lodged when his financial director, Ed Foyle, says “hey, you can’t do that, there is a (statute) law against it - if Charles suddenly need money for something and wants to withdraw his money, you have to be able to repay him in full, so you are not allowed to lend his money to anyone else”.

Al is a bit stumped by that and asks “how can we lend money then – it seems impossible?”. It’s quite ok says Ed, we just invent the money. Here, get Gary Hughes to sign this form, but make quite sure that the form says “Mr Gary Hughes” or “MR GARY HUGHES” as that is very important. When the form is signed, we can enter it in our company accounts as a credit because his signature makes it a promissory note which has a value of £10,000 for us. Then, we start making real money by writing him a cheque instead of handing him money. He will sign that cheque in order to cash it and that is worth another £10,000 to us. This is great business as we can enter his signed cheque into our ledger as another £10,000 asset, and write the £10,000 which goes into his account as a debit. We have now got £20,000 coming in and only £10,000 going out, so we are already ahead by £10,000 and it isn’t even time for lunch on our first day. Also, if Gary pays back the loan and our 20% interest, then this transaction will have made us a net profit of £22,000. If he didn’t pay us back anything, we would still be £10,000 ahead, so let’s celebrate.

Poor Al is a bit confused, so he asks “but where does the £10,000 that we lend Patrick come from?”. Ed laughs and says, it doesn’t come from anywhere, we just write numbers into our ledger against the name of Mr Gary Hughes – it doesn’t cost us anything – do you think he would hold still long enough for us to stamp “sucker” on his forehead? After all, he is about to pay us more than twice, which means that we will make millions in profit each year. Poor Al is still confused, which is a pity, as Gary is the only person which Ed wants to keep confused about the transaction.

Ed knows all of the legal angles and when he is driving home he exceeds the speed limit. It is not dangerous in any way as there is no other traffic, and the road conditions are perfect. A police car pulls alongside him and indicates that they want him to stop, so he pulls over and stops. A police officer approaches and asks Ed if he is aware that he was driving at 40 mph in a 30 mph zone. Ed says that he was indeed aware of that.

The police officer then asks Ed for his name and address. But, Ed is aware of the fact that this is a fraudulent attempt to get him to agree to “Joinder”, that is, agreeing that he is subject to all legal statutes, giving the police officer authority over him. So, he promptly states that the law does not require him to provide any such information, he also mentions the fact that the law does not require him to drive at, or below, 30 mph in a 30 mph zone. The police officer tries this again several times, getting the same answer each time.

Being a bit stumped by this unexpected response, the police officer switches tactics and says that he may have to take Ed down to the police station for further questioning and asks if Ed understands. Being well up on legal

tactics and realising that asking “do you understand” is a fraudulent attempt to establish Joinder, he calmly states that under the law, he is not required to go to a police station and he does NOT “stand under” the police officer or his statement. He also remarks that if he is taken to a police station unlawfully and against his will and without his consent, then he will take action against the police officer personally as a man and not as a police officer, for his unlawful actions. He also asks the police officer if he is aware that the Police Authority which employs him is not a government agency but is a commercial for-profit company with no powers whatsoever to uphold the law, that the police officer is only empowered to uphold the law through the oath which he personally took when he became a police**man** and it has nothing whatsoever to do with any form of authority coming from his employers. In other words, he is on his own on this one.

Ed further points out that under the law, there has to be a victim (who has been injured, has his goods damaged or stolen, or who has been defrauded) and he asks the police officer to identify the victim in this case. He also remarks that legal statutes only have the force of law for those people who consent to be bound by them, and he does NOT consent to be bound by them, and so, driving at 40 mph in a 30 mph zone where nobody was endangered and no harm or damage occurred, is not an offence for him.

I’ll leave you to work out the likely outcome of the conversation between Ed and the police officer.

Six months later, Ivan Johnson, a friend of Gary Hughes, remarks that Gary is looking worried and asks him if anything is wrong. Gary, who is self-employed, says that business has been very poor for some time and he is having financial problems. His main difficulty is this loan that he has with the Swindle Finance Company Limited. He has not been able to keep up the payments and he has just has a letter from them saying that he has to pay £2,231.47 by Friday, or they will take him to court. Fortunately, Ivan is familiar with both finance and legal matters, so he tells Gary what to do. His advice is simple, he says; for your contract with the Swindle Finance Company Limited to be valid, it has to meet these requirements:

1. Full disclosure, which didn’t happen as they did not tell you that your signature created the credit.
2. Equal Consideration, which didn’t happen as they have nothing to lose as they didn’t bring anything to the table.
3. Lawful terms and conditions
4. Signatures of the Parties, which didn’t happen as companies can’t produce a “wet ink” signature.

So, you send them a letter agreeing to pay any financial obligation which you might lawfully owe, as soon as they:
1. Provide validation of the debt, that is, the actual accounting.
2. Verification of their claim against you, that is, a sworn affidavit or a signed Invoice.
3. A copy of the Contract binding both parties.

Sign the Letter “Gary Hughes” (**not** “Mr Gary Hughes”) and write “All Rights Reserved” after your signature. Send the letter by registered delivery so that there is a third-party record of it being delivered to them.

The financial institution can’t validate the debt as they never sustained a loss. They can’t verify any claim against you as their billing is to “Mr Gary Hughes” or “MR GARY HUGHES” or “Mr G. Hughes” or “Hughes, Mr G” or some such wording, and none of those is you, since you are the man Gary Hughes. They can’t produce a copy of the contract, because one doesn’t exist – what exists is an unenforceable unilateral contract with no standing in law.

If they send you a statement of account, then tell them that a statement is not an Invoice and that you need an Invoice. If they send you a photocopy of your loan application form, then tell them that it is not a contract, being signed by only one of the parties.

If they delay and delay, then write and tell them that if they do not provide the requested documentation within fourteen days, that you will consider any debt as being fully discharged.

If they are ill-advised enough to attempt to take you to court, send a copy of your letter agreeing to pay provided they produce the evidence that there is something outstanding to pay, to the court. They may well be penalised by the court for wasting court time as courts only adjudicate between parties who disagree and there is no disagreement in your case as you have agreed to pay if the Swindle Finance Company Limited shows you that you are liable to pay them some amount of other.

Having had this explained to him, I will leave it up to you to decide if Gary felt any better about his situation.

The legal system is very large and very complex. This is quite deliberate and it is intended to make sure that no ordinary person knows all of it. Fortunately, the law is much less complicated and if you are very careful not to walk into any of the traps set for you, you can effectively ignore the legal system. If required to sign anything and you decide to do so, then Gary Hughes should sign it “By: Gary Hughes, agent” or “By: Gary Hughes, principal” as that you are not taking on any responsibilities in the strawman’s name.

John Harris of the TPUC organisation, has found a very effective technique for dealing with the attempts of the legal people to extract money from you, is to never go to court (remember that a 'Summons' is only an invitation and you can decline it) and returning all correspondence marking it "No Contract – Return to Sender". This avoid you ever walking into a "Joinder" trap and becoming liable for the fraudulent claims being made against your strawman. I suggest that you research this subject further and become completely familiar with your rights under common law, and the fraudulent "Joinder" traps being set for you. What you decide to do after that, is entirely up to you.

The Situation in Australia

It appears that both the government and the legal system of Australia are null and void as the basis on which they were created no longer exists. Sir Harry Gibb has this to say about the situation:

I am a former member of the High Court and I wish to take this unusual method of informing you about a matter that is going to deeply affect us all. Unfortunately, a document such as this is too easily "lost" in the bureaucratic jungle in which we operate.

A group of Australian Citizens have taken it upon themselves to test the validity of our current political and judicial system. Like you, I have lived my entire legal career with the assumption that the basis for our legal and political system, state and federal, was written in stone. This group has undertaken to present this paper where they test the legal system.

The group is articulate, well educated and has some of our best legal minds amongst its members. One of Australia's best known barristers is one of the group's leading lights. It is far better informed with regard to international law than most members of the judiciary or for that matter, the legal academe. It has better international contacts than I would have thought possible.

After spending some time with the group leader, I was able to elicit it's primary intentions. It is the introduction of a totally democratic system of government devoid of party politics operated by the will of the people incorporating a system of debit taxation which should go a long way to eliminating the current unemployment problem and addressing other pressing social issues. An A.B.S. financial model supports the proposal.

The group has so far concentrated on matters relating to taxation, state and federal, minor industrial and motor traffic while undertaking not to present a criminal defence using their current presentation. I challenged the leader of this group to present any evidence which he had with regard to the above defence so that I could use my legal expertise to play the part of devil's advocate. It should be brought to your attention that the group has access to documentation of which we members of the judiciary have little knowledge. I refer to the British Parliamentary Papers for the Colony of Australia for the years 1860 through to 1922.

These are photocopies of all documents correspondence etc., between the states and later the Commonwealth of Australia, the British Crown and the British Government. They are very revealing documents and indicate the degree of chicanery in which the politicians of all shades were involved and as I can now see, at the expense of the legal academe and the judiciary. I present for your perusal the details of the group's presentation along with my comment on each major item. The group relies solely upon historical fact and rejects political rhetoric and legal opinion unless based upon historical fact.

1. "The Commonwealth of Australia Constitution Act 1900 (UK) is an act of the parliament of the United Kingdom. It did not contain any substance of sovereignty and was a colonial act centralising self-government of the six Australian Colonies. Australia remained a colony of the United Kingdom."

1a. Although the late Lionel Murphy attempted to show that there was an element of sovereignty in this act, he failed to do so. The international definition of sovereignty has been espoused at length and the above act although important in the development of Australia, did not have the authority of sovereignty. The historical evidence that Australia remained a British Colony post 1901 is overwhelming.

2. "Australia made an international declaration of its intention to become a sovereign nation when Prime Minister Hughes and his deputy; Sir Joseph Cook signed the Treaty of Versailles on June 28, 1919. On its cognisance of signing this treaty, Australia was granted a "C" class League of Nations mandate over former German territories in the Pacific. In effect, Papua New Guinea became a colony of Australia achieving its own independence on 16 September 1975. The League of Nations became part of International Law on 10 January 1920 with Article X of the Covenant of League of Nations guaranteeing the sovereignty of each member."

2A. The Significance of Australia joining the League of Nations as a foundation member has never been

addressed in Australia before. Strangely, only one book has ever examined the question of Australian independence. Written by W. J. Hudson and M. P. Sharp in 1988 "Australian Independence" printed by Melbourne University Press. As both were members of the Department of Foreign Affairs and Trade at the time of authorship and had access to the, British Parliamentary Papers, I find it most interesting they have avoided any mention of these papers in their book. Their conclusion that Australia became an independent nation via. the Statute of Westminster in 1931 flies in the face of contradictory evidence within the above mentioned papers and readily available historical fact.

Prime Minister Hughes address to the Commonwealth Parliament on 10 September 1919, “Australia has now entered into a family of nations on a footing of equality. Australia has been born in a blood sacrifice" demonstrates that the politicians of the day were only too well aware of the change of status from a colony to that of a sovereign nation, while attempting to remain within the Empire.

Prime Minister Bruce made this reply to the British Government in 1922 after a request for troops against Kernel Ataturk in the Chanak crisis. Bruce’s reply is contained in the British Parliamentary Papers: “We have to try to ensure there shall be an Empire foreign policy which if we are to be in anyway responsible for it, must be one to which we agree and to which we have assented. If we are to take any responsibility for the Empire's foreign policy, then there must be a better system, so that we may be consulted and have a better opportunity to express the views of the people of this country. We cannot blindly submit to any policy which may involve us in war." This is a far cry from the declaration of war against Germany made on behalf of the British Colony of Australia by George V of the United Kingdom in 1914.

I have reproduced Bruce's reply in full as I believe this reply contains clear historical evidence of a Prime Minister who was well aware of the change of status from a. colony to a sovereign nation. The later Statute of Westminster 1931 was an acknowledgment of that status.

3. "Paragraph 4 of the Statue of Westminster Act 1931 contravenes Article X of the Covenant of the League of Nations. Paragraph 1 of the Australia Act 1986 contravenes Article 2 paragraphs 1 and 4 of the Charter of the United Nations."

3A Paragraph 4 of the Statute of Westminster reads "No Act of Parliament of the United Kingdom, passed after the commencement of this Act, shall extend, or be deemed to extend, to a Dominion as part of the law of that Dominion, unless it is expressly declared in that Act, that the Dominion, has requested, and consented to the enactment thereof." Paragraph 1 of the Australia Act is very similar: “No Act of the Parliament of the United Kingdom passed after the commencement of this Act shall extend, or be deemed to extend, to the Commonwealth, to a State or Territory as part of the law of the Commonwealth, of the State or of the Territory."

I passed this one to the Federal Attorney General and asked him what was the source of this quite incredible authority that sought to overturn the authority legislated within the Covenant of the League of Nations in Article X and the Charter of the United Nations in Article 2 paragraphs 1 and 4. He is unable to provide any documentation to support these clauses, Article X of the Covenant of the League of Nations states: “The members of the League undertake to respect and preserve against external aggression the territorial integrity and existing political independence of all Members of the League. In case of any such aggression or in case of any threat or danger of such aggression, the Council shall advise upon the means by which this obligation shall be fulfilled."

It is appropriate that I now introduce a statement by Sir Geoffrey Butler KBE, MA and Fellow, Librarian and Lecturer in International Law and Diplomacy of Corpus Christi College, Cambridge author of “A Handbook to-the League of Nations" used as a reference to the League by virtually all nations at that time. He refers to Article 1 of the Covenant of the League of Nations:

"It is arguable that this article is the Covenant's most significant single measure. By it the British Dominions, namely New Zealand, Australia, South Africa, and Canada, have their independent nationhood established for the first time. There may be friction over small matters in giving effect to this internationally acknowledged fact but the Dominions will always look to the League of Nations Covenant as their Declaration of Independence.

Article 2 paragraph 1 of the United Nation’s Charter states "The Organisation is based on the principle of the sovereign equality of all of its Members."

Article 2 paragraph 4 of the Charter states ‘All members shall refrain in their international relations, from the threat or use of force against the territorial integrity or political independence of any state, or in any other manner inconsistent with the Purposes of the United Nations."

In view of the above, the historical evidence for Australian Independence by 10 January 1920 when the League of Nations became part of International Law is overwhelming. When this evidence is reinforced with the contents of

the Charter of the United Nations, the continued usage of any legislation that owes its very legitimacy to the parliament of an acknowledged foreign power cannot be supported by either legal opinion or indeed historical evidence.

I therefore have come to the conclusion that the current legal and political system in use in Australia and its States and Territories has no basis in law.

Following discussions with members of the British Government relating to the Letters Patent for the Governor General and State Governors I find that these documents no longer have any authority. Indeed, the Queen of the United Kingdom is excluded from any position of power in Australia by the United Nations Charter and is excluded under UK law from the issue of a Letters Patent to other than a British Subject. A Letters Patent must refer to an action to be taken with regard to British Citizens. In the Immigration Act 1972, the UK defines Australian Citizens as “aliens”.

The Governor General’s Letters Patent is a comedy of errors. We are greeted in the name of the Queen of Australia who suddenly becomes the Queen of the United Kingdom in the next paragraph of the Letters Patent. This Queen then gives instructions to the Governor General with reference to the Commonwealth of Australia Constitution Act 1900 UK. Here we have a clear breach of Article 2 paragraph 1 of the United Nation Charter. Under both UK and international law, the-Queen is a British Citizen.

State Governors are in a worse position since their authority comes from the late Queen Victoria of the United Kingdom. Regardless of the validity of the Commonwealth of Australia Constitution Act 1900 UK, if the authority of Governor General and the State Governors is invalid then so is the entire political and legal system of government.

When advised that the War Crimes Commission was taking an interest, I called them in Geneva. Under the 1947 Geneva Convention, they are empowered to look into cases here in Australia where it is alleged that the law of a foreign country was enforced against a citizen of a member state of the United Nations. As they perceive that only the judiciary can actually enforce the law, the judiciary becomes their target. The group has already placed cases before them which they are currently investigating. If found guilty, the penalties are horrific and include the death penalty!

I could go on with more relevant information however I think now is the time for a summary. The group leader, a QC, states the obvious when he asked me how could a colony now acknowledged by all world nations to be a sovereign Nation, retain exactly the same legal and political system which it enjoyed as a colony, without any change whatsoever to the basis for law. This point alone requires an answer.

The High Court has already answered with regard to the position held by treaties signed by the Commonwealth Government in the Teoh case of 1994. "Ordinary people have the right to expect government officials to consider Australia's international obligations even if those obligations are not reflected in specific Acts of Parliament: the rights recognised in international treaties are an implied limit on executive processes."

My advice is to adjourn any case "sine die" that challenges the authority of the Letters Patent. Under no circumstances hear a case that challenges the validity of a State or the Federal Constitution. It is the politicians who are using us as pawns without them having to face the music. If these matters are of concern to politicians, then let them sort out these problems and accept any inherent risks themselves! Article 36 of the Statute of the International Court of Justice is the correct reference for you to refuse to hear a matter when an international treaty is cited as a defence.

The Comments of Brian Collins

The following is some of what the very knowledgeable Brian Collins has to say on his web site. I merely quote what he says here and I have not verified any of this material, and so I recommend that you check out these things for yourself.

The Greatest Lie Ever Told

The government and legal system of the United States, Canada, Australia, New Zealand and of course, Britain, is totally controlled by the Crown. The Crown is **not** the British Monarch. The Crown is the Inner City of London, which is an independent State in London belonging to the Vatican system. It is a banking cartel which has a

massive system around and beneath it, which hides its true power. The City is in fact, the Knights Templar Church, also known as the ‘Crown Temple’ or ‘Crown Templar’, and is located between Fleet Street and Victoria Embankment. The Temple grounds are also home to the Crown Offices at Crown Office Row.

The Crown Temple controls the Global 'Legal' system, including those in the United States, Canada, Australia, and much more; this is because all Bar Associations are franchises of the International Bar Association at the Inns of Court at Crown Temple based at Chancery Lane in London. All Bar Associations are franchises of the Crown and all Bar Attorneys/ Barristers throughout the world pledge a solemn oath to the Temple, even though many may not be aware that this is what they are doing. Bar Association 'licensed' Solicitors / Barristers must keep to their Oath, Pledge and terms of allegiance to the Crown Temple if they are to be "called to the Bar" and work in the legal profession. The ruling Monarch is also subordinate to the Crown Temple, this as been so since the reign of King John in the 13th century when Royal Sovereignty was transferred to the Crown Temple and, through this, to the Roman Church. King John 1167-1216 is the key to this deception.

It was at the Chancel, or Chancery, of the Crown Inner Temple Court in January 1215 that King John was faced with the demands of the French/English Barons in England (mainly French), to confirm the rights enshrined in the Magna Carta. When he signed the Magna Carta in 1215 history records this as an event that extended human freedom, but the real affect was very different as we shall see. The governments of the USA, Canada, Australia and more, are subsidiaries of the Crown Temple and so is the US Central Bank the Federal Reserve, as are all Central Banks on the planet, including the IMF World Bank.

Bar Associations are awarded their franchises by the Four Inns of Court at Temple Bar. These are the Inner Temple, Middle Temple, Lincolns Inn and Greys Inn and they are nothing less than elite secret societies without charters or statutes. They are the law unto themselves. The Inner Temple controls the legal system franchise for Canada and Britain while the Middle Temple does the same for the United States. Queen Elizabeth II is a member of both Temples. At least five signatories to the American Declaration of Independence were Temple Bar Attorneys who had pledged allegiance to the Crown! Another Middle Temple operative during the formation of the USA was Alexander Hamilton who structured the American Banking system to fulfil the Crown Temples agenda for total control of the United States. So in truth, a State wherever it is on this planet, is a legal entity of the Temple Crown, or a Crown Colony. A man named Michael Edwards wrote:

"Americans were fooled into believing that the legal Crown Colonies comprising New England, were independent nation states, but they never were, nor are they today. They were, and still are, Colonies of the Crown Temple, through letters patent and charters, who have no legal authority to be independent from the rule and order of the Crown Temple. A legal state is a Crown Temple Colony".

"Neither the American people nor the monarch of Britain own America. The Crown Temple owns America through the deception of those who have sworn their allegiance by oath to the Middle Temple Bar. The Crown Bankers and their Middle Templar Attorneys rule America through unlawful contracts, unlawful Taxes, and contract documents of false equity through debt deceit, all strictly enforced by their completely unlawful, but 'legal', orders, rules and codes of the Crown Temple Courts, or the so called, 'Judiciary', in America. This is because the Crown Temple holds the land titles and estate deeds to all of North America".

Seven Middle Inn Templars who had pledged an oath of allegiance to the Crown Temple (including Alexander Hamilton) were among the members of the Constitutional Convention who signed the completed 'American Constitution'. How symbolic it is that copies of the American Constitution and the Declaration of Independence hang on the wall of the Middle Temple in London. It's not that surprising when you consider that this Temple controlled both sides in these shenanigans.

So while the Middle Bar Templars were orchestrating the illusion of freedom from the perceived rule by King George III, the King too, was a sworn member of the same Temple. Michael Edwards continues:

"1776 is the year that will truly live in infamy for all Americans. It is the year that the Crown Colonies became legal Crown States. The Declaration of Independence was a legal, not lawful, document. It was signed on both sides by representatives of the Crown Temple. Legally, it announced the status quo of the Crown Colonies to that of the new legal name called 'States' as directly possessed estates of the Crown.

"The American people were hoodwinked into thinking that they were declaring lawful independence from the Crown. The proof that the colonies are still a Crown possession is the use of the word 'State' to signify a 'legal estate of possession'. Had this been a document “of and by the people”, then both the Declaration of Independence and the US Constitution would have been written using the word 'States'. By the use of 'State', the significance of government of estate possessions was legally established. All of the North American States are Crown Templar possessions through their legal document, signed by their representation of both parties to the contract, known as the Constitution of the United States of America.

Who controls the Crown? On 15th May 1213, King John, being a puppet of the Knights Templar, effectively signed away the Kingdom of England and Ireland to Pope Innocent III and the Roman Church, witnessed by the Crown Templars. As King John said at the time: "I myself bearing witness in the house of the Knights Templars".

The King John Charter said:

" We wish it to be known to all of you, through this our charter, furnished with our seal… not induced by force or compelled by fear, but of our own good and spontaneous will and by common counsel of our Barons, do offer and freely concede to God and his Holy Apostles Peter and Paul and to our mother the holy Roman Church, and to our lord Pope Innocent and to his Catholic successors, the whole kingdom of England and the whole kingdom of Ireland, with all their rights and appurtenances… we perform and swear fealty for them to him are aforesaid lord pope Innocent, and his Catholic successors and the Roman Church…
binding our successors and our heirs by our wife forever, in a similar manner to perform fealty and show homage to him who shall be chief pontiff at the time, and to the Roman church without demur. As a sign… we will and establish perpetual obligation and concession… from the proper and especial revenues of our aforesaid kingdoms…
the Roman church shall receive yearly a thousand marks sterling… saving to us and to our heirs our rights, liberties and regalia; all of which things, as they have been described above, we wish to have perpetually valid and firm; and we bind ourselves and our successors not to act counter to them. And if we or any of our successors shall presume to attempt this, whoever he be, unless being duly warned he come to his kingdom, and his senses, he shall lose his right to the kingdom, and this charter of our obligation and concession shall always remain firm."

Now it gets interesting. Historical accounts concentrate on the fact that this charter obliged the Crown to pay money to the Roman Church, but it also states that if the terms of the charter are broken, the 'right to the kingdom' shall be lost. When King John signed the Magna Carter on 15th June 1215, he broke the terms of the charter with the Pope and therefore lost the right to his Kingdom. Pope Innocent III then declared the Magna Carter to be null and void. From that time, the 'Crown' passed from the Monarch to the Knights Templar who to this day, govern Britain on behalf of the Roman Illuminati Church.

The St Clair's now Sinclair are the initial directors of this link of Templar and Rome in the British Isles as they took hold in Scotland at the time of the 1066 Norsemen-Norman invasion. This means that ultimately, through the Crown Temple, the Roman system also owns the United States, Canada and all countries and Bar associations controlled both overtly and covertly by the Crown. This is why the Peace Treaty between the American Colonies and the British Crown in 1783, stated:

"It having pleased the Divine Providence to dispose the hearts of the most serene and most potent Prince George the third, by the grace of God, King of Great Britain, France and Ireland, defender of the faith, Duke of Brunswick and Luneburg, arch treasurer and prince elector of the Holy Roman Empire etc… and of the United States of America…"

And ironically, a Roman Catholic is barred from being the British Monarch while all along the Monarch and the Kingdom have been owned by the system that is Rome.

The Queen is the Grand Patroness of Freemasonry. She is served by the 390 members of the Privy Council which connects with it's equivalent in all other Commonwealth Countries. The Privy Council is 'LEGALLY' above Parliament because of its prerogative powers. It's members, are appointed for life and they include Prince Philip, Prince Charles, the Archbishop of Canterbury, the Prime Minister and many others. They hold nine official meetings each year and the government (Crown) minister's stand to attention while the Queen is told of the government measures they are asking the Queen to approve, which are nothing other than what the Crown dictates via the elected leader and via the Whips of whichever political party has power... After taking their bow to the Queen and shaking her hand they are sworn to conduct their business in the utmost secrecy.

Another network of this structure is the CROWN AGENTS. Formed in 1833 as 'Crown Agents for the Colonies', to run the day to day administration of the empire and to serve as private bankers to government officials, colonial authorities, and heads of state. It is also an older version of Brown & Root supplying a vast range of goods throughout the empire. Goods that included arms, drugs, viruses, especially to Africa to create the genocidal wars we are now witnessing and have witnessed in the past. The 'Crown Agents' is the network which links the Crown Temple to the organised crime operations throughout the empire, the Golden Dawn being one of the most powerful occult operations to be used in Britain amongst the crime syndicates to ensure that the right crime waves are in full swing when required.

We can see this in the current agenda for guns. The Golden Dawn also operates within legal organisations, but is the interconnecting network across all secret societies and religion. Remember that the Crown Agents work

directly for the Monarch who works for the Knights Templar Church, which operates for the Vatican, and to add insult to injury, the entire debt of the Crown Agents is guaranteed by the British Government, and that means you!

In the 1970's it was bailed out by the Bank of England costing hundreds of millions of pounds. For many years the Crown Agents managed the personal wealth of the Sultan of Brunei who just happens to be a great friend of the Queen, and who has funded many of Prince Philips', Prince Charles', and George Bush senior's private projects. The Sultan is also the financier of both the British and the American unofficial intelligence operations as he has backed many of Mohammed Al Fayed's operations, in fact the Sultan could be said to be Mohammed Al Fayed!

After the murder of Princess Diana, the role of Al Fayed was for him to hijack the conspiracy and thus lead investigators down the Al Fayed rap of misinformation as was done with Nick Brown over the top-up fees debate, and is currently being played by David Davis. Fortunately no one needed to go to him for information because we all in our hearts already know that she was murdered.

In 1996 the 'Crown Agents' operation was privatised and given the name 'Crown Agents for Overseas Government and Administration Ltd'. Of course, privatisation is nothing but transfer of Power from Crown Temple through Government agency to Crown Temple via direct ownership: they own everything in this nation!

Just to give an idea of What the Crown Agencies controls, here are a few little companies which it controls. It acts as a holding company for:

Barclays bank
Standard and Chartered bank
Unilever
Tate and Lyle
Securicor
British Telecom
The Prince of Wales Business Leaders Forum, and
Aga Khan Foundation

I think you get the picture; it's big to say the least.

It manages the customs services of Mozambique, and through a company called Europe SA, it is in charge of all of the economic construction procurement for Bosnia. It is also in a joint venture with the Monaco-based company ES-KO, to provide all the food for United Nations peace keeping forces in Angola and Bosnia. So, the more war the more profit for the Crown Agents, and if it were to fail financially, then the British Taxpayer will have to foot the bill !!!!

The queen of England owns the 40,000 acre Duchy of Lancaster, and the 44,000 acre Duchy of Cornwall which gives Charles his income. In truth, it is the Crown that owns these assets, which are assigned to the ruling monarch, so long as they do as they are told. The Crown, it transpires, owns land in the UK to the tune of 40 million acres. The Queen owns over 300 residences around the country, and invests heavily in corporations like Rio Tinto Zinc, Royal Dutch Shell, ICI, and General Electric. This is no surprise as these companies are 100% Illuminati operations.

Rio Tinto is the largest mining company in the world and was established in 1873 by Hugh Matheson of the Global drug running operation Jardine Matheson. Rio Tinto was in at the start of the North Sea Oil, along with Texaco. They used the refineries of BP in which the Queen also has massive investments, so now you know where all the revenues from the North Sea oil reserves went - to the Crown Temple. This shows insider dealing on an individual basis enabling the Queen to make massive profits.

Another blatant conflict of interest to have been identified, was Rio Tinto's involvement in a cartel formed in 1971 to fix the price of uranium. A Federal grand jury and the 1976 US Senate Foreign Relations Committee, chaired by Frank Church, exposed the sting. It also included a company called Mary Kathleen Uranium of Australia. This company had been secretly encouraging the Aborigines to occupy uranium-rich lands in Australia to take these lands out of production in order to raise the price of uranium on the world market. The manipulated shortage of uranium had a serious effect on the American Westinghouse company who sought to take legal action against Rio Tinto for price rigging.

This strategy has been used in the USA against the Native Americans and throughout Africa. An American court ordered that Rio Tinto officials answer questions, but this was quashed by the British Law Lords. The Australian government passed legislation to the same effect. This was after the Australian Prime Minister, Gough Whitlam, had been dismissed from office by the Queen's Governor General of Australia, Sir John Kerr. Whitlam was pursuing a policy of buying out the mining and raw material cartels, like Rio Tinto and Anglo American, to stop them raping Australia's resource base, while giving nothing in return.

The Queen having enormous investments in both companies, removed Whitlam by using some of her wide range of prerogative powers, which by the way, she has in all commonwealth countries including Britain, which she can instigate when the Crown deem necessary, but of course in the minds of the masses the Queen has no powers today! Well I suppose she doesn't in that she is a puppet of the Crown Temple and ultimately, of the Roman system. Sir John Kerr, a former high level operative of British Intelligence, was made a member of the Privy Council and the Royal Victorian Order for his loyal and most profitable service to the Queen's portfolio.

He was later murdered however, when the truth about the removal of Gough Whitlam was in danger of coming out - how they look after their own! He was obviously outer circle, and a good man and would have exposed the lies of the inner circle. The Crown Cartels continue through a House of Windsor extremely secretive society called the Club of the Isles. It was named after King Edward VII, Queen Victoria's son.

The title is held today by Prince Charles. King Edward VII was heavily involved with the Black Nobility barons of 'the square mile of London' Financial District and helped them to engineer the Crimean War, the Russia - Japan War, the preparations for the First World War, and the Opium Wars against China. Through the central organisation of the Club of the Isles comes the fantastic web of interlocking directorships which holds apparently 'independent' companies in a network of common control and common agenda. Part of that web includes:

The Bank of England
Anglo American Corporation of South Africa
Rio Tinto
Minorco Minerals and Resources Corp
De Beers Consolidated Mines
De Beers Centenary AG
N.M Rothschild's Bank
Barclays Bank
Lloyds Bank
Lloyds Insurance Market
HSBC Bank
National Westminster Bank
Barings Bank
Schroders bank
Standard Chartered Bank
Hambros Bank
S.G Warburg
Toronto Dominion Bank
Johnson Matthey
Kleinwort Benson Group
Lazard Brothers
Lonrho
J.P Morgan and CO
Morgan Grenfell Group
British Petroleum
Shell and Royal Dutch Petroleum
Cadbury - Schweppes
Bat Industries
Assicurazioni Generali SpA (VeniceItaly)
Courtaulds
General Electric
Cazeenove and Co
Grand Metropolitan
Hanson PLC
HSBS (Hong Kong and Shanghai Bank)
Imperial and Chemical Industries (ICI)
Inchscape PLC
Inco Ltd.
ING Group
Jardine Matheson
Peninsular and Oriental Steam Navigation & Co (P & O ferries)
Pilkington Glass
Reuters Holdings
Glaxo Wellcome
Smithkline Beecham
Unilever and UnileverNV

Vickers PLC.

This is just a few of them. In 1999, Lonrho had 640 subsidiary companies itself; today, it has many more, and all of these companies have themselves, staggering numbers of subsidiaries. In studying the subsidiaries of the multi-nationals, you arrive at the fact that everything is controlled by the Crown, and that all of them lead back to operatives of the Roman Illuminati Church via the Knights Templar Church.

A fact worth noting is that just because a company like Pilkington Glass is now perceived to be owned by a non-British company, it has only moved from the British economy, not from the control or economy of the Illuminati who own everything, this type of movement is only a movement within the global conglomerate that is the New World Order and still earning for, and controlled by, the bloodline elite, albeit via a different branch of the elite tree.

Just look at some of the TV companies owned and thus controlled by General Electric:

NBC network
CNBC
MSNBC
Bravo
MUN2 TV
Sci-Fi channel
Trio
WNBC - New York
KNBC - Los Angeles
WMAQ - Chicago
WCAU - Philadelphia
KNTV - San Jose - San Francisco
KXAS - Dallas - Fort Worth
WRAC Washington
WTVJ - Miami
KNSDF - San Diego
WVIT - Hartford
WNCN - Raleigh
WCMH - Columbus
WVTM - Birmingham (USA)
WJAR - Providence
KVEA / KWHY - Los Angeles
WNJU - New York
WSCV - Miami
KTMD - Houston
WSNS - Chicago
KXTX - Dallas
KVDA - San Antonio
KSTS - San Francisco
KDRX - Phoenix
KNSO - Fresno
KMAS - Denver
WNEU - Boston / Merrimack
KHRR - Tucson
WKAQ - Puerto Rico
Universal Studios
NBC Universal Television Studio
NBC Universal Television Distribution.

This is one company controlling a major chunk of the TV dream state given out to the whole of America. Is it any wonder that Americans are completely under the control of the American elite?

Another important part of the Windsor network which is part of the Crown Temple, are the City Livery Companies. These purport to represent the various groups of merchants like the gun makers, stationers, newspaper makers, goldsmiths and the like. They are in fact, very secret societies, fundamental to the control of the City institutions and much further afield. In the 1350's, in the wake of the plague known to history as the Black Death, the government of the City passed from the ward councils to the City Livery Companies. In 1979, the year when Margaret Thatcher took power, the Honourable Company of Freemen of the City of London and of North America, began to hold meetings in New York and Toronto, and on 21st October 1991, the Association of Liverymen of the City of London in Hong Kong was founded and all their members were Architects (Freemasons). The late author Peter Jones, researched some of the Livery Companies in the 1990's for his book,' The Obedience of Australia',

which exposed the manipulation which led to the removal by the Queen, of Australian Prime Minister, Gough Whitlam. These are some of the names which he found within these companies:

Engineers: The Duke of Edinburgh.

Airline Pilots and Navigators: The Duke of Edinburgh Prince Andrew

Butchers: Queen Mother, Lord Vesty (of the meat family and Lord Prior of the Order of St John of Jerusalem)

Merchant Tailors: Queen Mother, Lord Whitelaw

Glovers: Margaret Thatcher, Sir John Fieldhouse (Admiral of the Fleet during the Falklands conflict)

Poulterers: Margaret Thatcher, Duchess of Devonshire.

Fishmongers: Duke and Duchess of Devonshire, C.E.A

Hambro (Hambros Bank, Taylor Woodrow, and P&O), Lord Inchcape (Her Majesty's Lieutenant of London)

Goldsmiths: J.H Hambro.

Grocers: Edward Heath.

Salter's: Duke of Kent (Grand Master of English Freemasonry) Lord Armstrong.

Cloth workers: Sir Peter Gadson (a Grand Master of the United Grand Lodge), Lord Carrington (Thatcher's handler).

Another name prominent within these companies is McAlpine of the construction family.

You will of course have noted that Prince Philip heads the Pilots and Navigators, which gives indication as to whom is in control of the current attack on our right to travel by plane and the whole swath of legislation about to come out relative to air traffic, using the environment scam to force it through. Just as a matter of interest, the cheap-flight operator 'Easy Jet' is financed by the Warburg Bank which is part of the Rothschild empire which has been proved beyond doubt, to have funded the rise of Hitler's war machine on behalf of the Rothschilds.

The fact that they are now blaming these cheap flight operators for the rise in carbon emissions shows you why the Warburg bank has funded this airline, it is the catalyst for the introduction of the carbon emissions scam. Basically they will determine how much you, as a household, can travel. You have to look at this whole carbon emissions agenda as an attempt to instil a massive suppressing layer of control upon families and business via the introduction of quotas per household.

This will stifle people's movements to unimaginable levels. Don't forget that the House of Rothschild operates directly under the current British chief, Prince Philip. Prince Philip's whole family were supporters of the Nazi's as was he himself. In 1935, Prince Christoph, the husband of Philip's sister Sophie, was a colonel in the SS of Himmler's personal staff and head of the Forschungsamt, an elite intelligence operation controlled by Hermann Goering.

It was they who carried out the famous 'Night of the Long Knives' when Hitler removed his key opponents. Christoph and Sophie named their eldest child, Karl Adolf, after Hitler and Prince Philip was involved in his education. Christoph's brother, Philip of Hess, was related to the King of Italy and was the official liaison between the fascists of Italy and Germany. Also at the same time the British King Edward VIII, who was forced to abdicate in order the Queen Mother could be placed in the position of Chief Toad which meant she had to marry the lesser of the brothers, Bertie, the one she could obviously control.

King Edward VIII was well known to be a Nazi supporter. Prince Philip kept up very serious connections with Edward even after his abdication in 1936. Another of Edward's supporters, and mentor to Prince Philip, was the known paedophile and Satanist Lord Louis Mountbatten (Battenberg). Mountbatten was Philip's route into the House of Windsor. Mountbatten was the grandson of Queen Victoria and Prince Albert and was born at Windsor Castle in 1900. While Mountbatten was supposedly fighting on the British side during the war, he maintained serious communications with the Windsor's German Clan via his sister Louise, the Crown Princess of Sweden and wife of King Gustav. Louise was Prince Philip's aunt.

At the end of the war, in June 1945, the British King, George VI, the father of Queen Elizabeth and puppet of the Queen Mother, sent the former MI5 officer, Anthony Blunt, to the Kronberg Castle of Prince Philip's sister Sophie,

and her Nazi husband Prince Christoph of Hess, to recover correspondence between the British Royal Family and their Nazi relatives, for propaganda aimed at convincing the aristocrats of Britain they had not been in contact all along. Blunt was the 'surveyor of the Queens Pictures' and a world expert in the paintings of Poussin, the initiate who painted pictures called 'The Shepherds of Arcadia' which related to the Rennes-le-Chateau mysteries. Blunt was named as a member of a KGB unit inside British Intelligence along with Burgess, MacLean, and Philby, the fifth man was never named, but was in fact, Lord Victor Rothschild.

If you have seen the latest Bond movies, you hear the Character 'M' played by Dame Judy Dench state, ' in the cold war days intelligence operatives that made a big mistake had the common decency to defect', of course they would the intelligence services of the world are one and the same. When Blunt was finally caught in the 1980's, Queen Elizabeth II demanded he was not questioned on his clandestine mission to Kronberg Castle… So who controls the law?

Another important secret society within the Crown Temple is the Knights Grand Cross of the Order of the Bath, sounds pretty strange until you understand its symbolism. Both George Bush snr. and Ronald Reagan were given this title. The resurrection bath of alchemy, symbolises rebirth and purification or absolution. Baths are given to Masonic Knights of the Bath before they perform horrendous deeds, hence the term 'Blood Baths'.

The information in the last few paragraphs only hints at the true power of the Crown Temple. One has also to include all that the Vatican openly controls, to see the true extent of the power of the Roman Illuminati Church to which the Knights Templar Church is subordinate. Through all these networks and more, thirteen Royal Nefilim families control the world, and it is they who are carrying out all of the atrocities in the world at this present time, to carry out the agenda of the Elohim.

The Crown also owns massive parts of our coastline around this nation and 12 miles out to sea, in total about half of the coastline. Some of this is under the control of the Ministry of Defence which is itself controlled by the Crown. The current legislation, under which the National Trust will purchase some of this land from the MOD and coast not belonging to the MOD, changes nothing. The National Trust is 100% controlled by the Crown. We have the British Waterways agency, and the railways again controlling massive amounts of land, and we are effectively almost back to the middle ages now.

One of the greatest means to the Windsor land-grab is the:

CLUB OF THE ISLES

Controlling the World Wildlife Fund, WWF, this is one hell of a heinous operation and given the fact that Prince Philip inherited the position as the Queen's husband, becoming Commander of the Club of the Isles...then he is at the helm of genocide the likes of which the Holocaust could never imagine.

I see that the Royals are playing the good guys relative to the Sunni Muslims with the release of the Manchester man after 18 years in prison in Pakistan, this in itself shows you who operates President Musharaff of Pakistan, via such orders as the Order of the Mystic Shrine. This is important, given that Pakistan is now under attack from NATO forces, again controlled by the Inner City banking cartel which includes the Dutch and the Swiss.

So how does the Crown control all the MPs? Answer; via the whip system. All of the MPs which you vote in to represent you, are told how to vote each day by each party's leadership via the whips. So when any politician that has ever been and will be while this system exists, tells lies to the whole population the minute he or she opens his mouth and speaks. No matter what he speaks about, it is saying what he has been told to say, how to speak and when to speak. If it does not obey the whips, then his career is over. Thus the Crown need only control the people at leadership level or the cabinet of each political party and the Whips office, then through that clique of puppets, they demand obedience from all MPs down to the MPs sitting in each constituency, claiming to be acting on behalf of their constituents. The majority of MPs don't even get to see each bill before it goes to the vote, and indeed they don't need to as they are told how to vote on every Bill.

On Sunday 11th February 2007 on the political show 'Sunday AM', Giles Brandreth, a Conservative Party whip stated:

"As a whip, we understand each MP's marital problems, their drinking problems their financial problems, we know a lot about them. We then advise and guide them to making the right decisions".

Do you still feel that MPs act according to their own understanding and emotions? Of course they do not. This is the best I have witnessed from a politician handing you on a plate how the game of politics actually operates. They are all, through fear of their personal information being leaked; totally afraid of the whips, and so do the whips' bidding, which, of course, is the Crown's bidding. All MPs who speak at the box in the Commons do so under oath to the Bible positioned under the box. What does that tell you about those who operate as very

religious people (Blair) when they spew the bullshit? Blair is a high member of the Illuminati and that is why he can lie with ease as he speaks to the House and the people.

The Economic Techniques

The following is an excerpt from a document, dated May 1979, and which appears to be a statement of the methods and techniques used in the economic warfare which is being waged against all ordinary people at this time. Let me emphasise that these are not my words but that I am quoting the contents of an old and unverified document of unknown origin. Anonymous documents are a standard method of spreading disinformation, however, the statements quoted below are included because they have every appearance of being a factual New World Order strategy already being implemented.

Top Secret

Silent Weapons for Quiet Wars, An introductory programming manual, Operations Research Technical Manual TM-SW7905.1 This publication marks the 25th anniversary of the Third World War, called the "Quiet War", being conducted using subjective biological warfare, fought with "silent weapons." May 1979

Security

It is patently impossible to discuss social engineering or the automation of a society, i.e., the engineering of social automation systems (silent weapons) on a national or worldwide scale without implying extensive objectives of social control and destruction of human life, i.e., slavery and genocide. Whenever any person or group of persons in a position of great power and without full knowledge and consent of the public, uses such knowledge and methodologies for economic conquest - it must be understood that a state of domestic warfare exists between that person, or group of persons, and the public. The solution of today's problems requires an approach which is ruthlessly candid, with no agonizing over religious, moral or cultural values.

Historical Introduction

Silent weapon technology has evolved from Operations Research ("OR"), a strategic and tactical methodology developed under the Military Management in England during the Second World War. The original purpose of Operations Research was to study the strategic and tactical problems of air and land defence with the objective of effective use of limited military resources against foreign enemies (i.e., logistics). It was soon recognised by those in positions of power that the same methods might be useful for totally controlling a society, but better tools were needed.

Social engineering (the analysis and automation of a society) requires the correlation of great amounts of constantly changing economic information (data), so a high-speed computerised data-processing system was necessary which could predict when society would arrive at the point of capitulation. Mechanical computers were too slow, but the electronic computer fills the bill. The next breakthrough was the development of the simplex method of linear programming in 1947 by the mathematician George B. Dantzig. Then in 1948, the transistor, promised great expansion of the computer field by reducing space and power requirements.

With these three inventions, those in positions of power strongly suspected that it was possible for them to control the whole world. Immediately, the Rockefeller Foundation got in on the ground floor by making a four-year grant to Harvard College, funding the Harvard Economic Research Project for the study of the structure of the American Economy. One year later, in 1949, The United States Air Force joined in. In 1952 the grant period terminated, and a high-level meeting of the Elite was held to determine the next phase of social Operations Research. The Harvard project had been very fruitful, as is borne out by the publication of some of its results in 1953 suggesting the feasibility of economic (social) engineering. (Studies in the Structure of the American Economy - copyright 1953 by Wassily Leontief, International Science Press Inc., White Plains, New York).

Engineered during the last half of the 1940's, by 1954 the new Quiet-War machine was ready. With the creation of the maser in 1954, the promise of unlocking unlimited sources of fusion atomic energy from the heavy hydrogen in sea water and the consequent availability of unlimited social power was a possibility only decades away. The combination was irresistible. The Quiet War was quietly declared by the International Elite at a meeting held in 1954. Although the silent weapons system was nearly exposed 13 years later, the evolution of the new weapon-system has never suffered any major setbacks. This year of 1979 marks the 25th anniversary of the beginning of the Quiet War. Already this domestic war has had many victories on many fronts throughout the world.

Political Introduction

In 1954 it was well recognised by those in positions of authority, that it would be only a few decades, before the general public would be able to grasp and upset the cradle of power, for the elements of the new silent-weapon

technology were as accessible for use in providing a public utopia as they were for providing a private utopia. So, the issue of primary concern, namely that of dominance, revolved around the subject of the energy sciences.

Energy

Energy is recognised as the key to all activity on earth. Natural science is the study of the sources and control of natural energy, and social science (theoretically expressed as economics) is the study of the sources and control of social energy. Both are bookkeeping systems based on mathematics. Therefore, mathematics is the primary energy science and the bookkeeper can be king if the public can be kept ignorant of the methodology of the bookkeeping.

All science is merely a means to an end. The means is knowledge. The end is control. Beyond this remains only one issue: Who will be the beneficiary? In 1954 this was the issue of primary concern. Although the so-called "moral issues" were raised, in view of the law of natural selection it was agreed that a nation or world of people who will not use their intelligence are no better than animals who do not have intelligence. Such people are beasts of burden and steaks on the table by choice and consent.

Consequently, in the interest of future world order, peace, and tranquillity, it was decided to wage a private quiet war against the American public with an ultimate objective of permanently shifting the natural and social energy (wealth) of the undisciplined and irresponsible many into the hands of the self-disciplined, responsible, and "worthy" few.

In order to implement this objective, it was necessary to create, secure, and apply new weapons which, as it turned out, were a class of weapons so subtle and sophisticated in their principle of operation and public appearance as to earn for themselves the name of "silent weapons." In conclusion, the objective of economic research, as conducted by the magnates of capital (banking) and the industries of commodities (goods and services), is the establishment of an economy which is totally predictable and which can be manipulated.

In order to achieve this totally predictable economy, the low-class elements of society must be brought under total control, i.e., They must be housebroken, trained, and assigned a yoke and long-term social duties from a very early age, before they have an opportunity to question the propriety of the matter. In order to achieve such conformity, the lower-class family unit must be disintegrated by a process of increasing preoccupation of the parents and the establishment of government-operated day-care centres for the occupationally orphaned children.

The quality of education given to the lower class must be of the poorest sort, so that the moat of ignorance isolating the inferior class from the superior class is and always remains, incomprehensible to the inferior class. With such an initial handicap, even bright lower class individuals have little if any hope of extricating themselves from their assigned lot in life. This form of slavery is essential to maintain some measure of social order, peace, and tranquillity for the ruling upper class.

Descriptive Introduction to the Silent Weapon

Everything which can be expected from an ordinary weapon is also expected from a silent weapon by its creators, but only in its own manner of functioning. It shoots situations, instead of bullets; it's propellant is data processing, instead of the conventional chemical explosion. The power originates from bits of data in a computer, instead of grains of gunpowder inside a gun. The operator is a computer programmer, instead of a military marksman, and the orders are issued by a banking magnate, instead of a military general. It makes no obvious explosive noises, causes no obvious physical or mental injuries, and does not interfere in an obvious way with anyone's daily social life.

However, it makes an unmistakable "noise," causes unmistakable physical and mental damage, and interferes unmistakably with daily social life, that is, in ways which are unmistakable to a trained observer who knows what to look for. The public cannot comprehend this weapon, and therefore cannot believe that they are being attacked and subdued by a weapon. The public might instinctively feel that something is wrong, but because of the technical nature of the silent weapon, they cannot express their feeling in any rational way, or handle the problem with intelligence. Therefore, they do not know how to call for help, nor do they know how to associate with others to defend themselves against it.

When a silent weapon is applied gradually, the public adjusts and adapts to its presence and learns to tolerate its encroachment on their lives until the psychological pressure (applied via economic methods) becomes too great and they crack up. Therefore, in a way, the silent weapon can be considered to be a type of biological weapon. It attacks the vitality, options, and mobility of the individuals of a society by knowing, understanding, manipulating, and attacking their sources of natural and social energy, and their physical, mental, and emotional strengths and weaknesses.

Theoretical Introduction

"Give me control over a nation's currency, and I care not who makes its laws." - Mayer Amschel Rothschild (1743

- 1812). Today's silent weapons technology is an outgrowth of a simple idea discovered, succinctly expressed, and effectively applied by Mr. Mayer Amschel Rothschild. Mr. Rothschild discovered the missing passive component of economic theory known as economic inductance. He, of course, did not think of his discovery in these twentieth century terms, and, to be sure, mathematical analysis had to wait for the Second Industrial Revolution, the rise of the theory of mechanics and electronics, and finally, the invention of the electronic computer before it could be effectively applied in the control of the world economy.

General Energy Concepts

In the study of energy systems, there always appears three elementary concepts. These are:

Potential Energy,
Kinetic Energy, and
Energy Dissipation.

Corresponding to these concepts, there are three idealised, essentially pure physical counterparts called passive components.

In the science of physical mechanics, the phenomenon of Potential Energy is associated with a physical property called Elasticity or stiffness, and can be represented by a stretched spring. In electronic science, Potential energy is stored in a capacitor instead of a spring. This property is called Capacitance instead of Elasticity.

In the science of physical mechanics, the phenomenon of Kinetic Energy is associated with a physical property called Inertia or Mass, and can be represented by a mass or a flywheel in motion. In electronic science, Kinetic Energy is stored in an inductor (in a magnetic field) instead of a mass. This property is called Inductance instead of Inertia.

In the science of physical mechanics, the phenomenon of Energy Dissipation is associated with a physical property called Friction or Resistance, and can be represented by a dashpot or other device which converts energy into heat. In electronic science, Energy Dissipation is performed by an element called either a Resistor.

In economics the equivalents of these three energy concepts are:
Economic Capacitance - **Capital** (money, stock/inventory, investments in buildings and durables, etc.)
Economic Conductance - **Goods** (production flow coefficients)
Economic Inductance - **Services** (the influence of the industrial population on output)

All of the mathematical theory developed in the study of one energy system (e.g., mechanics, electronics, etc.) can be immediately applied in the study of any other energy system (e.g., economics).

Mr Rothschild's Energy Discovery

What Mr. Rothschild discovered was the basic principle of power, influence, and control over people as applied to economics. That principle is "when you assume the appearance of power, people soon give it to you." Mr. Rothschild had discovered that currency or deposit loan accounts had the required appearance of power that could be used to induce people (inductance, with people corresponding to a magnetic field) into surrendering their real wealth in exchange for a promise of greater wealth (instead of real compensation). They would put up real collateral in exchange for a loan of promissory notes. Mr. Rothschild found that he could issue more notes than he had backing for, provided he had someone's stock of gold to show his customers as a persuader.

Mr. Rothschild loaned his promissory notes to individual and to governments. These would create overconfidence. Then he would make money scarce, tighten control of the system, and collect the collateral through the obligation of contracts. The cycle was then repeated. These pressures could be used to ignite a war. Then he would control the availability of currency to determine who would win the war. Any government which agreed to give him control of its economic system got his support. Collection of debts was guaranteed by economic aid to the enemy of the debtor. The profit derived from this economic methodology made Mr. Rothschild all the more able to expand his wealth. He found that the public greed would allow currency to be printed by government order beyond the limits (inflation) of backing in precious metal or the production of goods and services.

Apparent Capital as "Paper" Inductor

In this structure, credit, presented as a pure element called "currency," has the appearance of capital, but is in effect, negative capital. Hence, it has the appearance of service, but it is, in fact, indebtedness or debt. It is therefore an economic inductance instead of an economic capacitance, and if balanced in no other way, will be balanced by the negation of population (war, genocide). The total sum of goods and services represents real capital and it is called the **Gross National Product**, and currency may be printed up to this level and still represent economic capacitance; but currency printed beyond this level is subtractive, represents the introduction of economic inductance, and so becomes notes of indebtedness.

War is therefore the balancing of the system by killing the true creditors (the public which has been taught to exchange true value for inflated currency) and falling back on whatever is left of the resources of nature and regeneration of those resources. Mr. Rothschild had discovered that currency gave him the power to rearrange the economic structure to his own advantage, to shift economic inductance to those economic positions which would encourage the greatest economic instability and oscillation.

The final key to economic control had to wait until there was sufficient data and high-speed computing equipment to keep close watch on the economic oscillations created by price shocking and excess paper energy credits - paper inductance/inflation.

Breakthrough

The aviation field provided the greatest evolution in economic engineering by way of the mathematical theory of shock testing. In this process, a projectile is fired from an airframe on the ground and the impulse of the recoil is monitored by vibration transducers connected to the airframe and wired to chart recorders. By studying the echoes or reflections of the recoil impulse in the airframe, it is possible to discover critical vibrations in the structure of the airframe which either vibrations of the engine or aeolian vibrations of the wings, or a combination of the two, might reinforce resulting in a resonant self-destruction of the airframe in flight. From the standpoint of engineering, this means that the strengths and weaknesses of the structure of the airframe in terms of vibrational energy can be discovered and manipulated.

Application in Economics

To use this method of airframe shock testing in economic engineering, the prices of commodities are shocked, and the public consumer reaction is monitored. The resulting echoes of the economic shock are interpreted theoretically by computers and the psycho-economic structure of the economy is thus discovered. It is by this process that partial differential and difference matrices are discovered that define the family household and make possible its evaluation as an economic industry (dissipative consumer structure). Then the response of the household to future shocks can be predicted and manipulated, and society becomes a well-regulated animal with its reins under the control of a sophisticated computer-regulated social energy bookkeeping system. Eventually every individual element of the structure comes under computer control through a knowledge of personal preferences, such knowledge guaranteed by computer association of consumer preferences (universal product code - the striped bar codes on packages) with identified consumers (initially identified through the use of a credit card and later through a permanent "tattooed" body number invisible under normal ambient illumination).

The Economic Model

The Harvard Economic Research Project (1948-1952) was an extension of World War II Operations Research. Its purpose was to discover the science of controlling an economy: at first the American economy, and then the world economy. It was felt that with sufficient mathematical foundation and data, it would be nearly as easy to predict and control the trend of an economy as to predict and control the trajectory of a projectile. Such has proven to be the case. Moreover, the economy has been transformed into an accurately targeted guided missile.

The immediate aim of the Harvard project was to discover the economic structure, what forces can change that structure, how the behaviour of the structure can be predicted, and how it can be manipulated. What was needed was a well-organised knowledge of the mathematical structures and interrelationships of investment, production, distribution, and consumption. Briefly, it was discovered that an economy obeyed the same laws as electricity and that all of the mathematical theory and practical and computer know-how developed for the electronic field could be directly applied in the study of economics. This discovery was not openly declared, and its more subtle implications were, and are, kept as a closely guarded secret, for example, in an economic model, human life is measured in dollars, and that the electric spark generated when opening a switch connected to an active inductor is mathematically the same as starting a war.

The greatest hurdle which theoretical economists faced was the accurate description of the household as an industry. This is a challenge because consumer purchases are a matter of choice which in turn is influenced by family income, purchase price, and other economic factors. This hurdle was cleared in an indirect and statistically approximate way by an application of shock testing to determine the current characteristics, called current technical coefficients, of a household industry. Finally, because problems in theoretical electronics can be translated very easily into problems of theoretical economics, and the solution translated back again, it follows that only a book of language translation and concept definition needed to be written for economics. The remainder could be got from standard works on mathematics and electronics. This makes the publication of books on advanced economics unnecessary, and greatly simplifies the silent war project security.

Industrial Diagrams

An ideal industry is defined as a device which receives value from other industries in several forms and converts them into one specific product for sales and distribution to other industries. It has several inputs and one output. What the public normally thinks of as one industry is really an industrial complex, where several industries under one roof produce one or more products.

Three Industrial Classes
Industries fall into three categories or classes by type of output:
Class 1 - **Capital** (resources)
Class 2 - **Goods** (commodities or use - dissipative)
Class 3 - **Services** (action of population).

Class 1 industries exist at three levels:
(a) Nature - sources of energy and raw materials.
(b) Government - printing of currency equal to the gross national product (GNP), and extension of currency in excess of GNP.
(c) Banking - loaning of money for interest, and extension (inflation/counterfeiting) of economic value through deposit loan accounts.

Class 2 industries exist as producers of tangible or consumer (dissipated) products. This sort of activity is usually recognised and labelled by the public as "industry."

Class 3 industries are those which have service rather than a tangible product as their output. These industries are called
(a) Households, and
(b) Governments. Their output is human activity of a mechanical sort, and their basis is population.

Aggregation
The whole economic system can be represented by a three-industry model if one allows the names of the outputs to be:
(1) Capital,
(2) Goods, and
(3) Services.

The problem with this representation is that it would not show the influence of, say, the textile industry on the steel industry. This is because both the textile industry and the steel industry would be contained within a single classification called the "goods industry" and by this process of combining or aggregating these two industries under one system block they would lose their economic individuality.

The E- Model
A national economy consists of simultaneous flows of production, distribution, consumption, and investment. If all of these elements including labour and human functions are assigned a numerical value using common units of measure, say, US dollars at their 1939 value, then this flow can be further represented by a current flow in an electronic circuit, and its behaviour can be predicted and manipulated with useful precision.

The three ideal passive energy components of electronics, the capacitor, the resistor, and the inductor correspond to the three ideal passive energy components of economics called the pure industries of capital, goods, and services.

Economic Capacitance represents the storage of capital in one form or another.

Economic Conductance represents the level of conductance of materials for the production of goods.

Economic Inductance represents the inertia of economic value in motion. This is a population phenomenon known as services.

Economic Inductance
An electrical inductor (e.g., a coil or wire) has an electric current as its primary phenomenon and a magnetic field as its secondary phenomenon (inertia). Corresponding to this, an economic inductor has a flow of economic value as its primary phenomenon and a population field as its secondary field phenomenon of inertia. When the flow of economic value (e.g., money) diminishes, the human population field collapses in order to keep the economic value (money) flowing (extreme case - war). This public inertia is a result of consumer buying habits, expected standard of living, etc., and is generally a phenomenon of self-preservation.

Inductive Factors to Consider
(1) Population
(2) Magnitude of the economic activities of the government
(3) The method of financing these government activities
(See Peter-Paul Principle - inflation of the currency.)

Translation
(a few examples will be given.)
Charge - coulombs - dollars (1939).
Current Flow - amperes (coulombs per second) - dollars of flow per year.
Motivating Force - volts - dollars (output) demand.
Conductance - amperes per volt - dollars of flow per year per dollar demand.
Capacitance - coulombs per volt - dollars of production inventory/stock per dollar demand.

Time-Flow Relationships and Self-Destructive Oscillations
An ideal industry may be symbolised electronically in various ways. The simplest way is to represent a demand by a voltage and a supply by a current. When this is done, the relationship between the two becomes what is called an admittance, which can result from three economic factors:

(1) **Hindsight flow**,
(2) **Present flow**, and
(3) **Foresight flow**.

Foresight flow is the result of that property of living entities to cause energy (food) to be stored for a period of low energy (e.g., a winter season). It consists of demands made upon an economic system for that period of low energy (winter season). In a production industry it takes several forms, one of which is known as production stock or inventory. In electronic symbology this specific industry demand (a pure capital industry) is represented by capacitance and the stock or resource is represented by a stored charge. Satisfaction of an industry demand suffers a lag because of the loading effect of inventory priorities.

Present flow ideally involves no delays. It is, so to speak, input today for output today, a "hand to mouth" flow. In electronic symbology, this specific industry demand is represented by a conductance which is then a simple economic valve (a dissipative element).

Hindsight flow is known as habit or inertia. In electronics this phenomenon is the characteristic of an inductor (economic analogue = a pure service industry) in which a current flow (economic analogue = flow of money) creates a magnetic field (economic analogue = active human population) which, if the current (money flow) begins to diminish, will collapse (war) to maintain the current (flow of money - energy).

Other large alternatives to war as economic inductors or economic flywheels are an open-ended social welfare program, or an enormous (but fruitful) open-ended space program. The problem with stabilising the economic system is that there is too much demand on account of:
(1) Too much greed and
(2) Too much population.
This creates excessive economic inductance which can only be balanced with economic capacitance (true resources or value - e.g., in goods or services).

The social welfare program is nothing more than an open-ended credit balance system which creates a false capital industry to give non-productive people a roof over their heads and food in their stomachs. This can be useful, however, because the recipients become state property in return for the "gift," and form a standing army for the Elite - as he who pays the piper picks the tune. Those who get hooked on the economic drug, must go to the Elite for a fix. In this, the method of introducing large amounts of stabilising capacitance is by borrowing on the future "credit" of the world. This is a fourth law of motion - **onset**, and consists of performing an action and leaving the system before the reflected reaction returns to the point of action - a delayed reaction.

The means of surviving the reaction is by changing the system before the reaction can return. By this means, politicians become more popular in their own time and the public pays later. In fact, the measure of such a politician is the delay time. The same thing is achieved by a government by printing money beyond the limit of the gross national product, and economic process called inflation. This puts a large quantity of money into the hands of the public and maintains a balance against their greed, creates a false self-confidence in them and, for a while, stays the wolf from the door.

They must eventually resort to war to balance the account, because war ultimately is merely the act of destroying the creditor, and the politicians are the publicly hired hit men that justify the act to keep the responsibility and blood off the public conscience. If the people really cared about their fellow man, they would control their appetites (greed, procreation, etc.) so that they would not have to operate on a credit or welfare social system which steals from the worker to satisfy the bum. Since most of the general public will not exercise restraint, there are only two alternatives to reduce the economic inductance of the system.

(1) Let the populace bludgeon each other to death in war, which will only result in a total destruction of the living earth.

(2) Take control of the world by the use of economic "silent weapons" in a form of "quiet warfare" and reduce the economic inductance of the world to a safe level by a process of benevolent slavery and genocide.

The latter option has been taken as the obviously better option. At this point it should be crystal clear to the reader why absolute secrecy about the silent weapons is necessary. The general public refuses to improve its own mentality and its faith in its fellow man. It has become a herd of proliferating barbarians, and, so to speak, a blight upon the face of the earth. They do not care enough about economic science to learn why they have not been able to avoid war despite religious morality, and their religious or self-gratifying refusal to deal with earthly problems, places the solution of the earthly problem beyond their reach.

It is left to those few who are truly willing to think and survive, as the fittest to survive, to solve the problem for themselves as the few who really care. Otherwise, exposure of the silent weapon would destroy our only hope of preserving the seed of the future true humanity.

The Household Industry

The industries of finance (banking), manufacturing, and government, real counterparts of the pure industries of capital, goods, and services, are easily defined because they are generally logically structured. Because of this their processes can be described mathematically and their technical coefficients can be easily deduced. This, however, is not the case with the service industry known as the household industry.

Household Models

The problem which a theoretical economist faces is that the consumer preferences of any household are not easily predictable and the technical coefficients of any one household tend to be a non-linear, very complex, and variable function of income, prices, etc. Computer information derived from the use of the universal product code in conjunction with credit-card purchase as an individual household identifier, could change this state of affairs, but the Universal Product Code method is not yet available on a national or even a significant regional scale. To compensate for this data deficiency, an alternate indirect approach of analysis has been adopted known as economic shock testing. This method, widely used in the aircraft manufacturing industry, develops an aggregate statistical sort of data.

Applied to economics, this means that all of the households in one region or in the whole nation are studied as a group or class rather than individually, and the mass behaviour rather than the individual behaviour is used to discover useful estimates of the technical coefficients governing the economic structure of the hypothetical single-household industry. One method of evaluating the technical coefficients of the household industry depends upon shocking the prices of a commodity and noting the changes in the sales of all of the commodities.

Economic Shock Testing

In recent times, the application of Operations Research to the study of the public economy has been obvious for anyone who understands the principles of shock testing. In the shock testing of an aircraft airframe, the recoil impulse of firing a gun mounted on that airframe causes shock waves in that structure which tell aviation engineers the conditions under which some parts of the airplane or the whole airplane or its wings will start to vibrate or flutter like a guitar string, a flute reed, or a tuning fork, and disintegrate or fall apart in flight. Economic engineers achieve the same result in studying the behaviour of the economy and the consumer public by carefully selecting a staple commodity such as beef, coffee, gasoline, or sugar, and then causing a sudden change or shock in its price or availability, thus kicking everybody's budget and buying habits out of shape. They then observe the shock waves which result by monitoring the changes in advertising, prices, and sales of that and other commodities.

The objective of such studies is to acquire the know-how to set the public economy into a predictable state of motion or change, even a controlled self-destructive state of motion which will convince the public that certain "expert" people should take control of the money system and re-establish security (rather than liberty and justice) for all. When the subject citizens are rendered unable to control their financial affairs, they, of course, become totally enslaved, and a source of cheap labour. Not only the prices of commodities, but also the availability of labour can be used as the means of shock testing. Labour strikes deliver excellent tests shocks to an economy, especially in the critical service areas of trucking (transportation) , communication, public utilities (energy, water, garbage collection), etc. By shock testing, it is found that there is a direct relationship between the availability of money flowing in an economy and the real psychological outlook and response of masses of people dependent upon that availability. For example, there is a measurable quantitative relationship between the price of gasoline and the probability that a person would experience a headache, feel a need to watch a violent movie, smoke a cigarette, or go to a tavern for a mug of beer.

It is most interesting that, by observing and measuring the economic models by which the public tries to run from their problems and escape from reality, and by applying the mathematical theory of Operations Research, it is possible to program computers to predict the most probable combination of created events (shocks) which will

bring about a complete control and subjugation of the public through a subversion of the public economy (by shaking the plum tree).

Introduction to Economic Amplifiers

Economic Amplifiers are the active components of economic engineering. The basic characteristic of any amplifier (mechanical, electrical, or economic) is that it receives an input control signal and delivers energy from an independent energy source to a specified output terminal in a predictable relationship to that input control signal. The simplest form of an economic amplifier is a device called advertising. If a person is spoken to by a T.V. advertiser as if he were a twelve-year-old, then, due to suggestibility, he will, with a certain probability, respond or react to that suggestion with the uncritical response of a twelve-year-old and will reach into his economic reservoir and deliver its energy to buy that product on impulse when he passes it in the store.

An Economic Amplifier may have several inputs and output. Its response might be instantaneous or delayed. Its circuit symbol might be a rotary switch if its options are exclusive, qualitative, "go" or "no-go", or it might have its parametric input/output relationships specified by a matrix with internal energy sources represented. Whatever its form might be, its purpose is to govern the flow of energy from a source to an output sink in direct relationship to an input control signal. For this reason, it is called an active circuit element or component. Economic Amplifiers fall into classes called strategies, and, in comparison with electronic amplifiers, the specific internal functions of an economic amplifier are called logistical instead of electrical. Therefore, economic amplifiers not only deliver power gain but also, in effect, are used to cause changes in the economic circuitry.

In the design of an economic amplifier we must have some idea of at least five functions ,which are:

(1) The available input signals,
(2) The desired output-control objectives,
(3) The strategic objective,
(4) The available economic power sources,
(5) The logistical options.

The process of defining and evaluating these factors and incorporating the economic amplifier into an economic system has been popularly called "game theory". The design of an economic amplifier begins with a specification of the power level of the output, which can range from personal to national. The second condition is accuracy of response, i.e., how accurately the output action is a function of the input commands. High gain combined with strong feedback helps to deliver the required precision. Most of the error will be in the input data signal. Personal input data tends to be specified, while national input data tends to be statistical.

Short List of Inputs

General sources of information:
(1) Telephone taps
(2) Surveillance
(3) Analysis of garbage
(4) Behaviour of children in school

Standard of living by:
(1) Food
(2) Clothing
(3) Shelter
(4) Transportation

Social contacts:
(1) Telephone - itemized record of calls
(2) Family - marriage certificates, birth certificates, etc.
(3) Friends, associates, etc.
(4) Memberships in organizations
(5) Political affiliation

The Personal Paper Trail

Personal buying habits, i.e., personal consumer preferences:
(1) Bank accounts
(2) Credit-card purchases

(3) "Tagged" credit-card purchases - those with a Universal Product Code

Assets:
(1) Bank accounts
(2) Savings accounts
(3) Property
(4) Business
(5) Vehicles, etc.
(6) Safety deposit at a bank
(7) Stock market purchases

Liabilities:
(1) Creditors
(2) Enemies (see - legal)
(3) Loans
(4) Consumer credit

Government sources (ploys)*:
(1) Welfare
(2) Social Security
(3) U.S.D.A. surplus food
(4) Dole
(5) Grants
(6) Subsidies

*Principle of this ploy - the citizen will almost always make the collection of information easy if he can operate on the "free sandwich principle" of "eat now, and pay later."

Government sources (via intimidation) :
(1) Internal Revenue Service
(2) OSHA
(3) Census
(4) etc.

Other government sources - surveillance of U.S. mail.

Habit Patterns - Programming

Strengths and weaknesses:
(1) Activities (sports, hobbies, etc.)
(2) See "legal" (fear, anger, etc. - crime record)
(3) Hospital records (drug sensitivities, reaction to pain, etc.)
(4) Psychiatric records (fears, angers, disgusts, adaptability, reactions to stimuli, violence, suggestibility or hypnosis, pain, pleasure, love, and sex)

Methods of coping - of adaptability - behaviour:
(1) Consumption of alcohol
(2) Consumption of drugs
(3) Entertainment
(4) Religious factors influencing behaviour
(5) Other methods of escaping from reality

Payment modus operandi (MO) - pay on time, etc.:
(1) Payment of telephone bills
(2) Energy purchases (electrical, gas,...)
(3) Water purchases
(4) Repayment of loans
(5) House payments
(6) Vehicle payments

(7) Payments on credit cards

Political sensitivity:
(1) Beliefs
(2) Contacts
(3) Position
(4) Strengths/weaknesses
(5) Projects/activities

Legal inputs - behavioural control (Excuses for investigation, search, arrest, or use of force to modify behaviour):
(1) Court records
(2) Police records - NCIC
(3) Driving record
(4) Reports made to police
(5) Insurance information
(6) Anti-establishment acquaintances

National Input Information

Business sources (via I.R.S., etc.):
(1) Prices of commodities
(2) Sales
(3) Investments in
  (a) stocks/inventory
  (b) production tools and machinery
  (c) buildings and improvements
  (d) the stock market

Banks and credit bureaus:
(1) Credit information
(2) Payment information

Miscellaneous sources:
(1) Polls and surveys
(2) Publications
(3) Telephone records
(4) Energy and utility purchases

Short List of Inputs
Outputs - create controlled situations - manipulation of the economy, hence society - control of compensation and income.

Sequence:
(1) Allocates opportunities.
(2) Destroys opportunities.
(3) Controls the economic environment.
(4) Controls the availability of raw materials.
(5) Controls capital.
(6) Controls bank rates.
(7) Controls the inflation of the currency.
(8) Controls the possession of property.
(9) Controls industrial capacity.
(10) Controls manufacturing.
(11) Controls the availability of goods (commodities) .
(12) Controls the prices of commodities.
(13) Controls services, the labour force, etc.
(14) Controls payments to government officials
(15) Controls the legal functions.
(16) Controls the personal data files - uncorrectable by the party slandered.
(17) Controls advertising.
(18) Controls media contact.

(19) Controls material available for T.V. viewing.
(20) Disengages attention from real issues.
(21) Engages emotions.
(22) Creates disorder, chaos, and insanity.
(23) Controls design of more probing tax forms.
(24) Controls surveillance.
(25) Controls the storage of information.
(26) Develops psychological analyses and profiles of individuals.
(27) Controls sociological factors.
(28) Controls health options.
(29) Preys on weakness.
(30) cripples strengths.
(31) leaches wealth and substance.

Table of Strategies

Do This	To Get This
Keep the public ignorant	Less public organisation
Maintain access control	Required reaction to output (prices, sales)
Create preoccupation	Lower defences
Attack the family unit	Control the education of the young
Give less cash and more credit and dole	More self-indulgence and more data
Attack the privacy of the church	Destroy faith in this sort of government
Social conformity	Computer programming simplicity
Minimise tax protests	Maximum economic data, minimum enforcement problems
Stabilise the consent	Simplicity coefficients
Tighten control of variables	Simpler computer input data - greater predictability
Establish boundary conditions	Problem simplicity / solutions of differential and difference equations
Proper timing	Less data shift and blurring
Maximise control	Minimum resistance to control
Collapse of currency	Destroy the faith of the American people in each other

Diversion, the Primary Strategy

Experience has shown that the simplest method of securing a silent weapon and gaining control of the public is to keep the public undisciplined and ignorant of the basic system principles on the one hand, while keeping them confused, disorganised, and distracted with matters of no real importance on the other hand. This is achieved by:

(1) Disengaging their minds; sabotaging their mental activities; providing a low-quality program of public education in mathematics, logic, systems design and economics; and discouraging technical creativity.
(2) Engaging their emotions, increasing their self-indulgence and their indulgence in emotional and physical activities, by:
(a) Unrelenting emotional affronts and attacks (mental and emotional rape) by way of constant barrage of sex, violence, and wars in the media - especially the T.V. and the newspapers.
(b) Giving them what they desire - in excess - "junk food for thought" - and depriving them of what they really need.
(3) Rewriting history and law and subjecting the public to the deviant creation, thus being able to shift their thinking from personal needs to highly fabricated outside priorities.

These preclude their interest in, and discovery of, the silent weapons of social automation technology. The general rule is that there is a profit in confusion; the more confusion, the more profit. Therefore, the best approach is to create problems and then offer solutions.

Diversion Summary

Media: Keep the adult public attention away from the real social issues, and captivated by matters of no real importance.

Schools: Keep the young public ignorant of real mathematics, real economics, real law, and real history.

Entertainment: Keep the public entertainment below a sixth-grade (12 year old) level.

Work: Keep the public busy, busy, busy, with no time to think; back on the farm with the other animals.

Consent, the Primary Victory

A silent weapon system operates upon data obtained from a docile public by legal (but not always lawful) force. Much information is made available to silent weapon systems programmers through the Internal Revenue Service. (See Studies in the Structure of the American Economy for an I.R.S. source list). This information consists of the enforced delivery of well-organised data contained in federal and state tax forms, collected, assembled, and submitted by slave labour provided by taxpayers and employers. Furthermore, the number of such forms submitted to the I.R.S. is a useful indicator of public consent, an important factor in strategic decision making. Other data sources are given in the Short List of Inputs.

Consent Coefficients - numerical feedback indicating victory status. Psychological basis: When the government is able to collect tax and seize private property without just compensation, it is an indication that the public is ripe for surrender and is consenting to enslavement and legal encroachment. A good and easily quantified indicator of harvest time is the number of public citizens who pay income tax despite an obvious lack of reciprocal or honest service from the government.

Amplification Energy Sources

The next step in the process of designing an Economic Amplifier is discovering the energy sources. The energy sources which support any primitive economic system are, of course, a supply of raw materials, and the consent of the people to labour and consequently assume a certain rank, position, level, or class in the social structure, i.e., to provide labour at various levels in the pecking order. Each class, in guaranteeing its own level of income, controls the class immediately below it, hence preserves the class structure. This provides stability and security, but also government from the top. As time goes on and communication and education improve, the lower-class elements of the social labour structure become knowledgeable and envious of the good things that the upper-class members have. They also begin to attain a knowledge of energy systems and the ability to enforce their rise through the class structure. This threatens the sovereignty of the Elite.

If this rise of the lower classes can be postponed long enough, the Elite can achieve energy dominance, and labour by consent no longer will hold a position of an essential energy source. Until such energy dominance is absolutely established, the consent of people to labour and let others handle their affairs must be taken into consideration, since failure to do so could cause the people to interfere in the final transfer of energy sources to the control of the Elite. It is essential to recognise that at this time, public consent is still an essential key to the release of energy in the process of economic amplification. Therefore, consent as an energy release mechanism will now be considered.

Logistics

The successful application of a strategy requires a careful study of inputs, outputs, the strategy connecting the inputs and the outputs, and the available energy sources to fuel the strategy. This study is called "Logistics". A logistical problem is studied at the elementary level first, and then levels of greater complexity are studied as a synthesis of elementary factors. This means that a given system is analysed, i.e., broken down into its sub-systems, and these in turn are analysed, until by this process, one arrives at the logistical "atom," the individual.

The Artificial Womb

From the time a person leaves his mother's womb, his every effort is directed towards building, maintaining, and withdrawing into artificial wombs, various sorts of substitute protective devices or shells. The objective of these artificial wombs is to provide a stable environment for both stable and unstable activity; to provide a shelter for the evolutionary processes of growth and maturity - i.e., survival; to provide security for freedom and to provide defensive protection for offensive activity. This is equally true of both the general public and the Elite. However, there is a definite difference in the way each of these classes go about the solution of problems.

The Political Structure of a Nation - Dependency

The primary reason why the individual citizens of a country create a political structure is a subconscious wish or desire to perpetuate their own dependency relationship of childhood. Simply put, they want a human god to eliminate all risk from their life, pat them on the head, kiss their bruises, put a chicken on every dinner table, clothe their bodies, tuck them into bed at night, and tell them that everything will be all right when they wake up in the morning. This public demand is incredible, so the human god, the politician, meets incredibility with incredibility by promising the world and delivering nothing. So who is the bigger liar? The public? or The "godfather"? This public behaviour is surrender born of fear, laziness, and expediency. It is the basis of the welfare state as a strategic weapon, useful against a disgusting public.

Action / Offence

Most people want to be able to subdue and/or kill other human beings who disturb their daily lives, but they do not want to have to cope with the moral and religious issues which such an overt act on their part might raise. Therefore, they assign the dirty work to others (including their own children) so as to keep the blood off their hands. They rave about the humane treatment of animals and then sit down to a delicious hamburger from a whitewashed slaughterhouse down the street and out of sight. But even more hypocritical, they pay taxes to finance a professional association of hit men collectively called politicians, and then complain about corruption in

government.

Responsibility
Again, most people want to be free to do the things (to explore, etc.) but they are afraid of failing. The fear of failure is manifested in irresponsibility, and especially in delegating those personal responsibilities to others where success is uncertain or carries possible or created liabilities (law) which the person is not prepared to accept. They want authority (root word - "author"), but they will not accept responsibility or liability. So they hire politicians to face reality for them.

Summary
The people hire the politicians so that the people can:
(1) Obtain security without managing it.
(2) Obtain action without thinking about it.
(3) Inflict theft, injury, and death upon others without having to contemplate either life or death.
(4) Avoid responsibility for their own intentions.
(5) Obtain the benefits of reality and science without exerting themselves or learning either of these things.

They give the politicians the power to create and manage a war machine:
(1) Provide for the survival of the nation/womb.
(2) Prevent encroachment of anything upon the nation/womb.
(3) Destroy the enemy who threatens the nation/womb.
(4) For the sake of stability of the nation/womb, destroy those citizens of their own country who do not conform.

Politicians hold many quasi-military jobs, the lowest being the police who are soldiers, the attorneys and C.P.A.s next who are spies and saboteurs (licensed), and the judges who shout orders and run the closed union military shop for whatever the market will bear. The generals are industrialists. The "presidential" level of commander-in-chief is shared by the international bankers. The people know that they have created this farce and financed it with their own taxes (consent), but they would rather knuckle under than be the hypocrite. Thus, a nation becomes divided into two very distinct parts, a docile sub-nation [great silent majority] and a political sub-nation. The political sub-nation remains attached to the docile sub-nation, tolerates it, and leaches its substance until it grows strong enough to detach itself and then devour its parent.

System Analysis
In order to make meaningful computerized economic decisions about war, the primary economic flywheel, it is necessary to assign concrete logistical values to each element of the war structure - personnel and material alike. This process begins with a clear and candid description of the subsystems of such a structure.

The Draft (military service)
Few efforts of human behaviour modification are more remarkable or more effective than that of the socio-military institution known as the draft. A primary purpose of a draft or other such institution is to instil, by intimidation, in the young males of a society the uncritical conviction that the government is omnipotent. He is soon taught that a prayer is slow to reverse what a bullet can do in an instant. Thus, a man trained in a religious environment for eighteen years of his life can, by this instrument of the government, be broken down, be purged of his fantasies and delusions in a matter of mere months. Once that conviction is instilled, all else becomes easy to instil.

Even more interesting is the process by which a young man's parents, who purportedly love him, can be induced to send him off to war to his death. Although the scope of this work will not allow this matter to be expanded in full detail, nevertheless, a coarse overview will be possible and can serve to reveal those factors which must be included in some numerical form in a computer analysis of social and war systems. We begin with a tentative definition of the draft. The draft (selective service, etc.) is an institution of compulsory collective sacrifice and slavery, devised by the middle-aged and elderly for the purpose of pressing the young into doing the public dirty work. It further serves to make the youth as guilty as the elders, thus making criticism of the elders by the youth less likely (Generational Stabilizer). It is marketed and sold to the public under the label of "patriotic = national" service.

Once a candid economic definition of the draft is achieved, that definition is used to outline the boundaries of a structure called a Human Value System, which in turn is translated into the terms of game theory. The value of such a slave labourer is given in a Table of Human Values, a table broken down into categories by intellect, experience, post-service job demand, etc. Some of these categories are ordinary and can be tentatively evaluated in terms of the value of certain jobs for which a known fee exists. Some jobs are harder to value because they are unique to the demands of social subversion, for an extreme example: the value of a mother's instruction to her daughter, causing that daughter to put certain behavioural demands upon a future husband ten or fifteen years hence; thus, by suppressing his resistance to a perversion of a government, making it easier for a banking cartel to buy the State of New York in, say, twenty years.

Such a problem leans heavily upon the observations and data of wartime espionage and many types of psychological testing. But crude mathematical models (algorithms, etc.) can be devised, if not to predict, at least to predetermine these events with maximum certainty. What does not exist by natural cooperation is thus enhanced by calculated compulsion. Human beings are machines, levers which may be grasped and turned, and there is little real difference between automating a society and automating a shoe factory. These derived values are variable. (It is necessary to use a current Table of Human Values for computer analysis). These values are given in true measure rather than U.S. dollars, since the latter is unstable, being presently inflated beyond the production of national goods and services so as to give the economy a false kinetic energy ("paper" inductance). The silver value is stable, it being possible to buy the same amount with a gram of silver today as it could be bought in 1920. Human value measured in silver units changes slightly due to changes in production technology.

Factor 1 - Enforcement
As in every social system approach, stability is achieved only by understanding and accounting for human nature (action/reaction patterns). A failure to do so can be, and usually is, disastrous. As in other human social schemes, one form or another of intimidation (or incentive) is essential to the success of the draft. Physical principles of action and reaction must be applied to both internal and external subsystems. To secure the draft, individual brainwashing/programming and both the family unit and the peer group must be engaged and brought under control.

Factor 2 - Father
The man of the household must be housebroken to ensure that junior will grow up with the right social training and attitudes. The advertising media, etc., are engaged to see to it that father-to-be is pussy-whipped before or by the time he is married. He is taught that he either conforms to the social notch cut out for him or his sex life will be hobbled and his tender companionship will be zero. He is made to see that women demand security more than logical, principled, or honourable behaviour. By the time his son must go to war, father (with jelly for a backbone) will slam a gun into junior's hand before father will risk the censure of his peers, or make a hypocrite of himself by crossing the investment he has in his own personal opinion or self-esteem. Junior will go to war or father will be embarrassed. So junior will go to war, the true purpose not withstanding.

Factor 3 - Mother
The female element of human society is ruled by emotion first and logic second. In the battle between logic and imagination, imagination always wins, fantasy prevails, maternal instinct dominates so that the child comes first and the future comes second. A woman with a newborn baby is too starry-eyed to see her child as a wealthy man's cannon fodder or a cheap source of slave labour. A woman must, however, be conditioned to accept the transition to "reality" when it comes, or even sooner. As the transition becomes more difficult to manage, the family unit must be carefully disintegrated, and state-controlled public education and state-operated child-care centres must be become more common and legally enforced so as to begin the detachment of the child from the mother and father at an earlier age. Inoculation of behavioural drugs [Ritalin] can speed the transition for the child (mandatory). Caution: A woman's impulsive anger can override her fear. An irate woman's power must never be underestimated, and her power over a pussy-whipped husband must likewise never be underestimated. It got women the vote in 1920.

Factor 4 - Junior
The emotional pressure for self-preservation during the time of war and the self-serving attitude of the common herd that have an option to avoid the battlefield - if junior can be persuaded to go - is all of the pressure finally necessary to propel Johnny off to war. Their quiet blackmailing of him are the threats: "No sacrifice, no friends; no glory, no girlfriends".

Factor 5 - Sister
And what about junior's sister? She is given all the good things of life by her father, and taught to expect the same from her future husband regardless of the price.

Factor 6 - Cattle
Those who will not use their brains are no better off than those who have no brains, and so this mindless school of jelly-fish, father, mother, son, and daughter, become useful beasts of burden or trainers of the same.

Please note that the section of text above, shown in blue, is not an expression of my own personal opinion, but comes from an anonymous document. However, sections of it describe clearly what is without question being applied to the people in many nations of the world today, so it is difficult to discount any of what is said. At the present time, economic warfare is most definitely being waged against normal, innocent people in most countries of the world, and there is every appearance that a major offensive against us is in progress, instigated and orchestrated by the few "Elite" New World Order people who have already caused so much unnecessary death and suffering.

So who are the New World Order people whose aims are to disadvantage and destroy ordinary people through their "Quiet War"? Well, there are several major branches of them, each presenting a benign and caring face to the world and most with large numbers of sincere adherents who have not the slightest idea of the aims and policies of the few people at the top who actually direct and control those organisations. While most operations are carried out in complete secrecy, the most visible people in the ruling bodies form the Bilderberg Group. This comprises 13 Freemasons, 13 Vatican representatives, and 13 Black Mobility people. These 39 people of the Bilderberg Group have offices in Switzerland (the only European country which was never invaded or bombed during World War One or World War Two) answer to the 13 people who form the Policy Group, which in turn answers to the 9 people of the Round Table. The individuals concerned are typically, powerful financiers, industrialists, statesmen and intellectuals.

The power and influence of these people should not be underestimated. They have infiltrated and subverted the US Congress and Senate, rendering ineffective the checks and balances intended when the Constitution was originally set up. At 3:30 am on Saturday 4th August 1990, a minority of US Senators, maybe ten at most, passed the Senate Intelligence Authorisation Act for Fiscal Year 1991 (S.B. 2834). This bill is not widely known. It was brought to a vote by Senator Sam Nunn in the dead of night when the opposition was gone. It effectively transfers most authority of the United States Government directly into the hands of the President. It gives him the power to initiate war, appropriate public funds, define foreign policy goals, and decide what is important to US national security. It gives the President the power to initiate covert actions (a power never before given to any President) and prevents Congress from stopping the President's initiation of covert actions. It allows the President to use any federal "departments, agencies or entities" to operate or finance a covert operation. It empowers the President to use any other nation or private contractor or person to fund or operate a covert action. It redefines covert actions as operations "necessary to support foreign policy objectives of the United States" which is a definition which is so broad and vague as to be essentially unlimited. It, for the first time ever, officially claims the right of the United States to interfere secretly in the internal "political, economic or military affairs" of other countries in direct and flagrant violation of international law. It requires the President to prepare and deliver a written finding to the Intelligence committees of Congress, but it allows the President to omit "extremely sensitive matters" and authorises the President to claim Executive Privilege if Congress asks too many questions.

Further, there are no penalties in the Bill for violating any of its provisions, including the provision for requiring a finding. That Bill effectively handed all the powers of government to the President, effectively making him the sole ruler of the USA without accountability to anyone.

I suggest that the New World Order people have sufficient influence to ensure that every candidate for President of the USA is a member of the New World Order. That way, they are assured that they have full control of the whole of the USA at all times no matter what the outcome of the "democratic" voting. It is said that Henry Ford stated that customers could have a new Model T car in any colour they wanted, provided that the colour was black. Well, the equivalent is now in place, where the American people can have any President they want, provided that the President is a New World Order man. This information is not even considered secret any more. For example:

"... Some even believe we (the Rockefellers) are part of a secret cabal working against the best interests of the United States, characterising my family and me as 'internationalists' and of conspiring with others around the world to build a more integrated global political and economic structure - one world, if you will. If that's the charge, I stand guilty, and I am proud of it. "

—David Rockefeller, Memoirs (2002, Random House publishers), page 405

It should not be imagined that the events in the USA have no effect elsewhere. The New World Order people are working on expanding the EU, blurring the identities of individual countries through legislation, reducing the effects of borders and generally moving towards a single entity with one central government. They are actually aimed at producing ten unified areas of the world which they then intend to amalgamate into a single world state governed by them. This is judged to be easier if there are fewer people, so one of their major aims is to reduce the number of people living at the present time. They also need the remaining people to be wholly dependent on them for the essentials of life, which is one reason why they oppose the introduction of any free-energy device, since having an independent source of power would put people outside their direct control and so will not be allowed.

If the video has not been wiped off the web, then the video from Walter Burien:
http://www.youtube.com/watch?v=r5XbvCRHYIk demonstrates the double-bookkeeping which is being operated by the US government where the US people are being told that they are in debt when in fact, “public” funds are massively in profit. In Britain, and probably in many other places in the world, nearly 80% of all earned income is taken away from ordinary people. Natural rights are being done away with and replaced by “privileges” which can be revoked. It has been remarked with considerable insight, that “in the old days, slaves were housed and fed, but today, those slaves are expected to house and feed themselves”. For a simple straightforward explanation of how this is done, try http://www.yourstrawman.com.

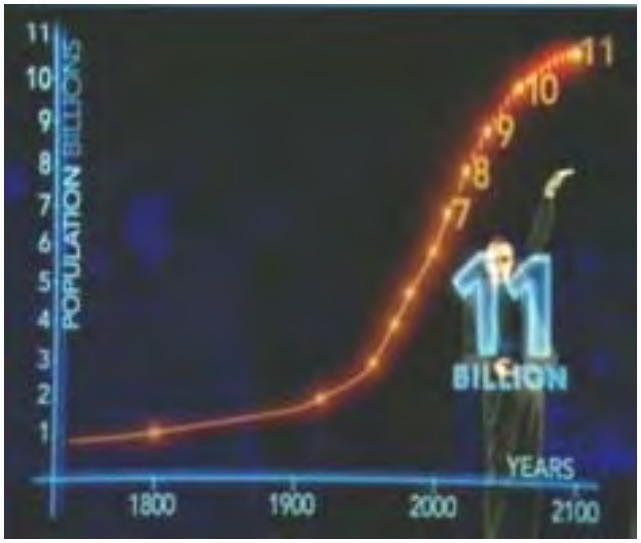
World Population

Recently, an excellent video presentation has been made by the very competent Swedish p Prof. Hans Rosling, examining population growth around the world, family size, economic status, personal aspirations and likely future trends around the world.

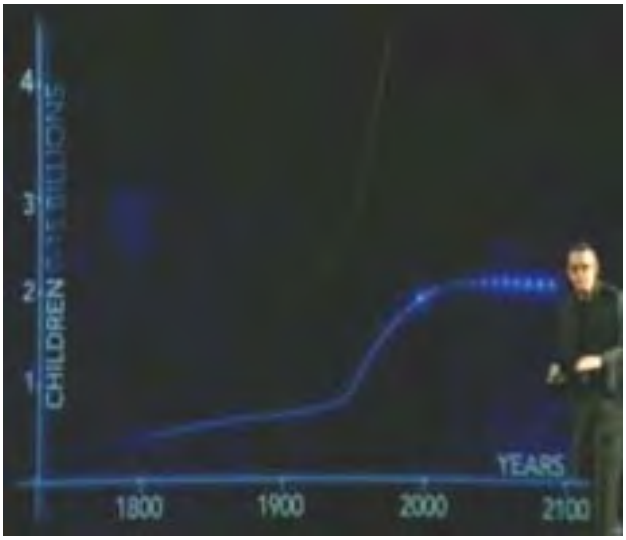
He points out that world population reached 1 billion in the year 1800, and population growth which had been minimal up to then, due to family deaths running at four people in a family of six, reinforced by the fact that average lifespans in those early years being very low. World population reached 2 billion in little more than 100 years afterwards due to the Industrial Revolution. He states that when he was at school, it was 3 billion and at that time, experts said that the planet just could not support so many people. However, contrary to those opinions, in his lifetime it has more than doubled to 7 billion, which he displays graphically like this:

He takes Bangladesh as a particularly severe case. There, during his lifetime, the population has tripled from 50 million to 150 million, but there has already been a major change in lifestyles and attitudes. In 1972 there were on average, 7 babies per woman and the baby's lifespan averaged at just 50 years. That has now dropped to just 2.2 children and the lifespan is now 70 years. In addition, the government now pays parents to keep their children in school longer so that they get a much better education and have far greater opportunity to get well-paid jobs. In just 40 years, Bangladesh has made spectacular progress, but it is not alone in that. In 1963 the average number of babies per woman world wide was 5 while in 2012 that had fallen to just 2.5 and is still decreasing. This change is also not related in any way to religion as every religion has 2-child families as the most common family unit. Child survival is a key factor because if it is almost certain that one or more children will die before the age of five, then parents have more children to allow for that near certainty. The rapid increase in population size which became very noticeable in 1963 was because the average number of children per woman was five, but only one of those five did not reach adulthood, while not long before that, the average number of children per woman was six but, tragically, four of those six died at an early age. Child mortality rate is a critical factor.

The best projections available for the future are that the world population will continue to increase, primarily due to better health care and will continue to rise steadily to about 11 billion, and around the end of this century, will level off around 11 billion. This increase is due to the present population growing older and staying healthy, rather than an increasing number of babies. Actually, the number of babies in the world stabilised around the year 2000 at about 2 billion and is staying steady at that level. Prof. Rosling, who knows a great deal more about this than I do, sees this as the key factor for population zero-growth rate. Personally, I would have thought that the ratio between births and deaths on average world wide was the key factor, but Prof. Rosling certainly seems to know exactly what he is talking about and he displays what he considers to be a reliable estimate of future population growth like this:



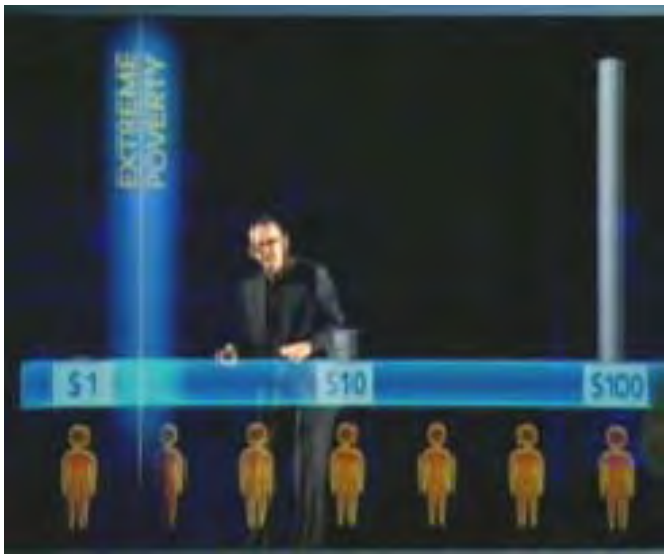
And showing the flattening of the graph for child population in more detail:



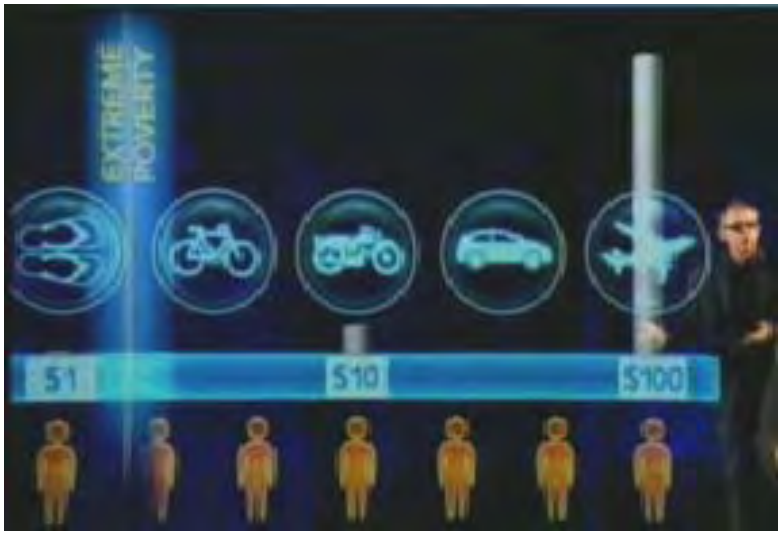
The prediction is for child numbers to remain around 2 billion right through to the end of this century. The overall increase in population growth is caused by young people living today and who survive to old age. The location of people in the world is also important. At the present time, in approximate figures, some 14% of world population is located in America (North and South combined). A further 14% are located in Europe and another 14% in Africa, while the remainder of 58% live in Asia (including Australia and New Zealand).

By the end of this century, it is expected that about 9% will be in America and 9% in Europe, while the population in Africa will have risen to about 36% and in Asia to about 46%. Prof. Rosling expresses the opinion that it is perfectly possible for those levels of population to have a good and prosperous life. He takes Mozambique in Africa as an example as he has personal knowledge of that location, having worked as a medical doctor in that country thirty years ago. The question is, can Africa support a population of 4 billion people and provide all of them with a good life? He states that things have improved immensely in Mozambique even though it is a very poor country. Those born there now have much better conditions with vastly improved medical care, a booming economy and new industries with many new jobs. In common with most of Africa, the biggest improvements are in the towns and cities while the rural areas where most people live have the least improvement, but things are changing there too. In rural areas, most people are farmers, relying on what they grow, for what they eat. These people have needs which you might not suspect. Their first need is for shoes, as they have to walk everywhere, which wastes a great deal of time and requires considerable effort. Their second need is for a bicycle as that allows them to avoid major waste of time in reaching the fields, collecting water and taking large loads to the market and it allows them to search for work and if they get sick, to reach a clinic in time. They have no electricity and farming methods are very basic, causing yields to be far lower than they could be. However, change and improvements are coming, even to the poorest regions of the poorest countries, although progress is slow.

Prof. Rosling then examines the income levels of the present 7 billion people, from the richest to the poorest, illustrating it like this:



Here, it is estimated that the richest billion people average US $100 per day and the poorest billion average just US $1 per day. The economists estimate extreme poverty at just over US $1 per day, defining 'extreme poverty' as not having enough to be sure that you can feed your family. That extreme poverty line divides the second least well-off billion people. To give us a feel for the desires of each billion people, Prof. Rosling illustrates it like this:



Here, the poorest billion people want to get shoes to help with having to walk everywhere. The second and third billion people would find life much easier if each family had a bicycle. The middle billion people want to have a motorcycle, the fifth and sixth billion want a car while the top billion people generally travel by aeroplane. He also points out that the richest people are so much richer than the $100 per day people that the difference is almost as great as the difference between the $1 per day people and the $100 per day people.

However, while we might well tend to think that both the people earning $1 per day and the people earning $10 per day as all of them being poor. But that is very much **not** the case. Earning $10 per day makes a massive difference in quality of life over those people earning just $1 per day. Having a bicycle allows people to reach literacy classes where they can learn to read and write which is a very important step forward. Interestingly, at the present time, more than 80% of people in the world can read and write. Things have improved enormously in the last fifty years. Most countries in Africa today are making steady progress, although we still have more than a billion people in extreme poverty. The United Nations as set the goal of eliminating extreme poverty within twenty years. That can be achieved, mainly by the hard work of those people in need. In that poorest billion people is almost all of the large families, and lack of literacy. Extreme poverty is almost self-creating and so it needs to be ended quickly.

Unfortunately, Prof. Rosling in his excellent presentation entitled "DON'T PANIC the Truth About World Population", now leaves his area of expertise and then discusses the global-warming scam which pretends that human made Carbon Dioxide is responsible for excessive weather conditions, and that is just not the case. The main "global warming gas" is actually water vapour and not carbon dioxide (which is needed by growing crops).

Also, of the carbon dioxide produced each year on planet Earth, humans come only a poor third as the vast bulk comes from volcano emissions, the next largest volume comes from insects and animals which outnumber humans by 200,000,000 to one. This notion of human-caused "global warming" is just part of a scam dreamt up to make money by fooling people into paying money when there is absolutely no justification at all. It is complete nonsense, especially so, since the other planets in the solar system are experiencing the same thing, not to mention the fact that global warming and cooling were taking place on Earth long before the Industrial Revolution which is supposed to be causing all this 'extra' carbon dioxide. These scams always rely on the ignorance of most people.

But hold on a minute, this skilful and very well presented programme is being broadcast by the British Broadcasting Corporation (the "BBC") which is a major propaganda outlet for the New World Order. This can be seen quite clearly from the evidence of the BBC's complicit role with the events of 9/11, the BBC reported the collapse of WTC Building 7 on live TV more than 20 minutes before the building actually collapsed. As BBC reporter Jane Standley reported that WTC Building 7 had collapsed, the actual building could still be seen standing intact behind her: http://www.youtube.com/watch?v=ltP2t9nq9fI.

Reporting an unexpected building collapse before it occurred, takes some explaining as to how prior knowledge of the event was in the hands of the BBC. The BBC operates a protection racket called the "TV Licence" scam, where people are threatened and coerced into paying an arbitrarily chosen amount of money to this private

commercial company. There is actually, not the slightest doubt about this, a fact which was highlighted by the court case of Mr Rooke who was charged with the "crime" of withholding his TV licence fees, and in February 2013, in Surrey in England, the BBC took Mr Rooke to court. However, once the judge saw the evidence presented by Mr Rooke and his team of experts against the BBC, he had no option but to discharge the conviction. Mr Rooke felt that the BBC's cover up of this evidence was supporting the true terrorists of 9/11 to get away with their crime. He refused to pay his TV licence fee on the basis of Section 15, Article 3 of the UK Terrorism Act.

This states that it is a crime to provide monetary support to any organisation that engages in or supports terrorist activity. In the opinion of Mr Rooke, by the BBC covering up the evidence about 9/11, they have been aiding and supporting the true perpetrators of 9/11, allowing them to escape justice, and that this is allowing a false version of 9/11 to continue to be used to justify further terrorist activity which is still continuing today. Therefore, in Mr Rooke's opinion, to provide funds to the BBC through his TV licence fees would in effect be funding an organisation which is supporting terrorist activity, and under Section 15, Article 3 of the Terrorism Act, he would be committing a serious crime if he did this.

So, the question faces us: why does the BBC broadcast the presentation by Prof. Rosling? Is it just to bolster the global warming caused by humans myth or is there some other reason? I'll leave that up to you to work out.

Patrick Kelly
http://www.free-energy-info.tuks.nl/

Chapter 16: Questions and Answers

This is a fast-moving field of research and development, and one where a single person working alone, can outperform research establishments with unlimited budgets and many staff members. If you believe that the answers given below are inadequate or incorrect, then please tell me as this is definitely a learning process for all of us.

Topic: What To Build

I am regularly asked to recommend a device for someone to build. This is an impossible task as the person asking never says in which part of the world he lives, how much money could be used in the project, what tools are to hand, what workshop space (if any) is available, what local supplies of materials is like, whether the home is in a city or an isolated remote place, local assets like a shoreline, a fast-flowing stream or lots of sunshine all year round, how much ground space is available, what his skills are, etc. etc.

What should be realised is that a small group of friends can achieve a great deal. If one person likes working with engines and another can fix television sets and a third can weld, or … whatever. By working together, they can achieve a great deal and benefit both themselves and their friends, not to mention having a good deal of fun in the process and ending up with a great sense of achievement.

We should not dismiss the various forms of renewable energy such as wave power, RAM pump, solar panels, wind power, biomass and the like. Granted, they do not always provide power all of the time, but if they provide most of the power you need for most of the time you need it, then that can be a considerable saving on your present power costs, so you should consider the devices covered in chapter 14. One possibility is to use a RAM pump with a fast-flowing stream, to pump water up into a water tower over night, and use the tank of water to power a turbine driving an electrical generator when needed.

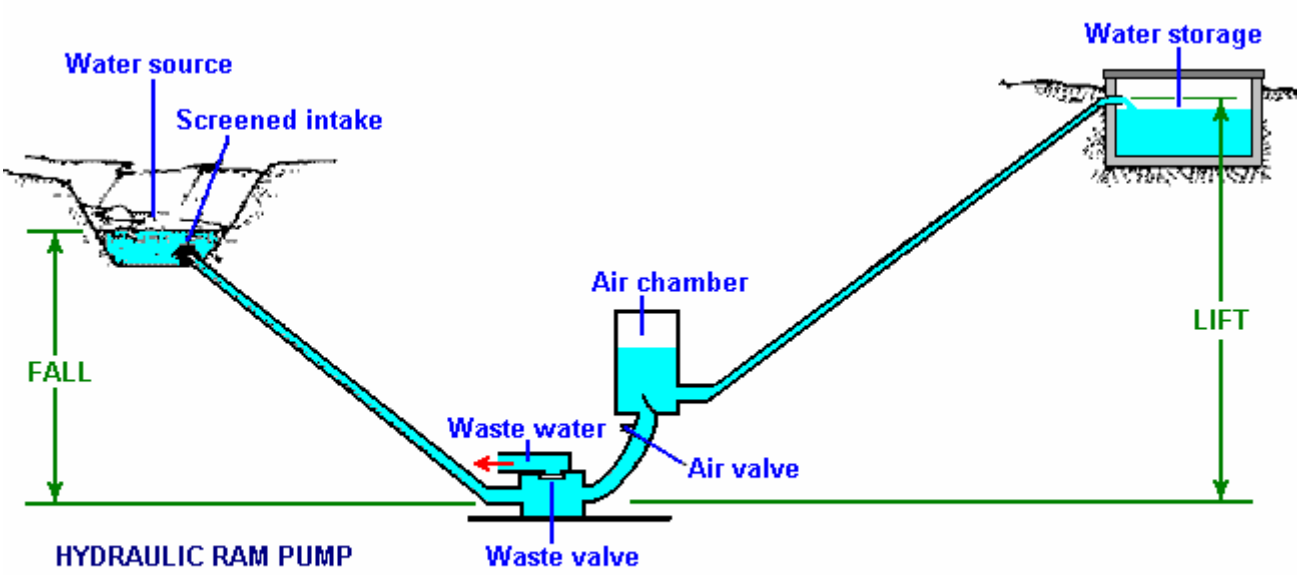
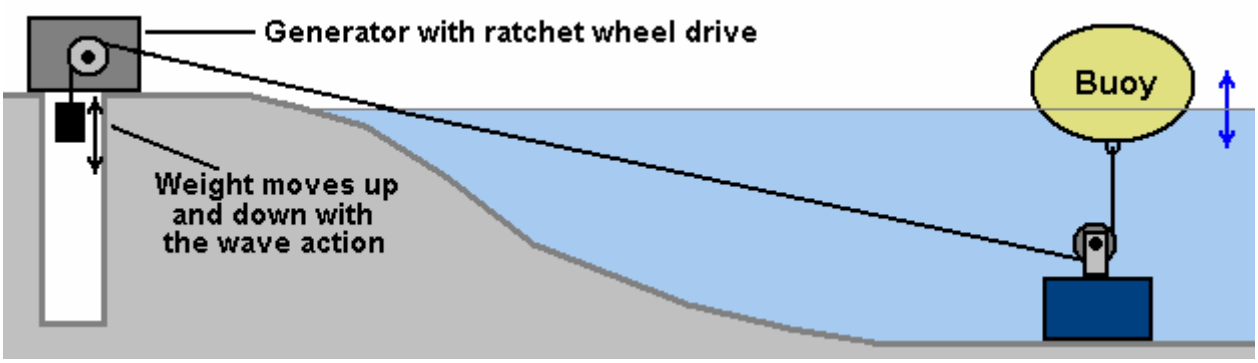


Diagram from: www.thefarm.org/charities/i4at/lib2/hydrpump.htm

If you have the space and the sunshine, then a parabolic mirror can focus the sunshine, produce steam and drive a generator, a mechanical drive, a pump, or whatever else is needed.

Above: Steven Vanek with his machine which uses solar thermal energy to make ice.

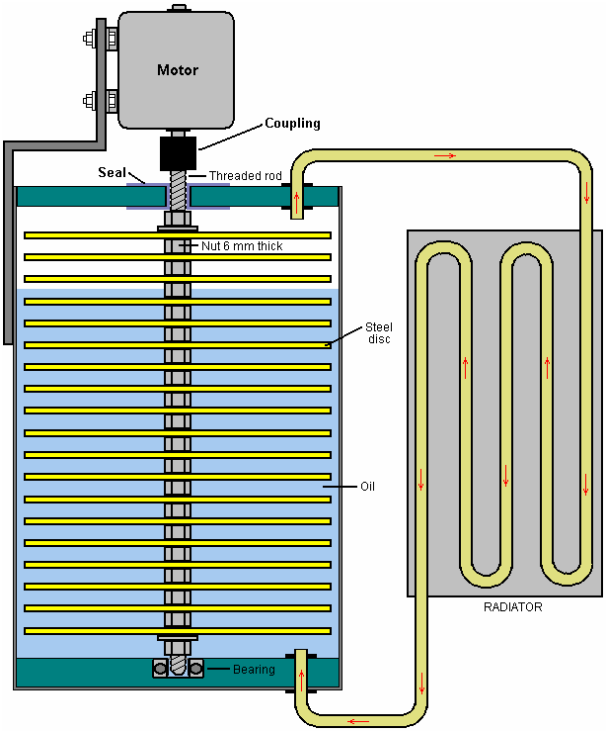
If you are lucky enough to have access to a small section of coastline, then using either wave power or tidal power gives a major amount of power each year with even a simple low-tech construction like this:



It is not necessary to "go for broke" and aim at a device which will completely do away with your power needs for evermore. Initially, it can be a major step forward to produce a device which reduces your energy bill by a good deal and later on, progress to an additional system which makes you fully independent of others. For example, it is perfectly possible to drive a RotoVerter (chapter 2) with a solar panel:

The output is much more powerful than the input from the solar panel and can be used to run power tools, charge batteries or perform other useful tasks at effectively no cost at all.

The Frenette heater (chapter 14) does give significant heat output for the modest power input needed to spin the central shaft. It is a simple construction well within the capabilities of most people, and it could be driven by a solar panel if you live in an area where it is cold and sunny. As the array of discs spin, the vegetable oil inside the cylinder gets heated and pushed upwards and outwards, flowing through the radiator tubing, heating the house:

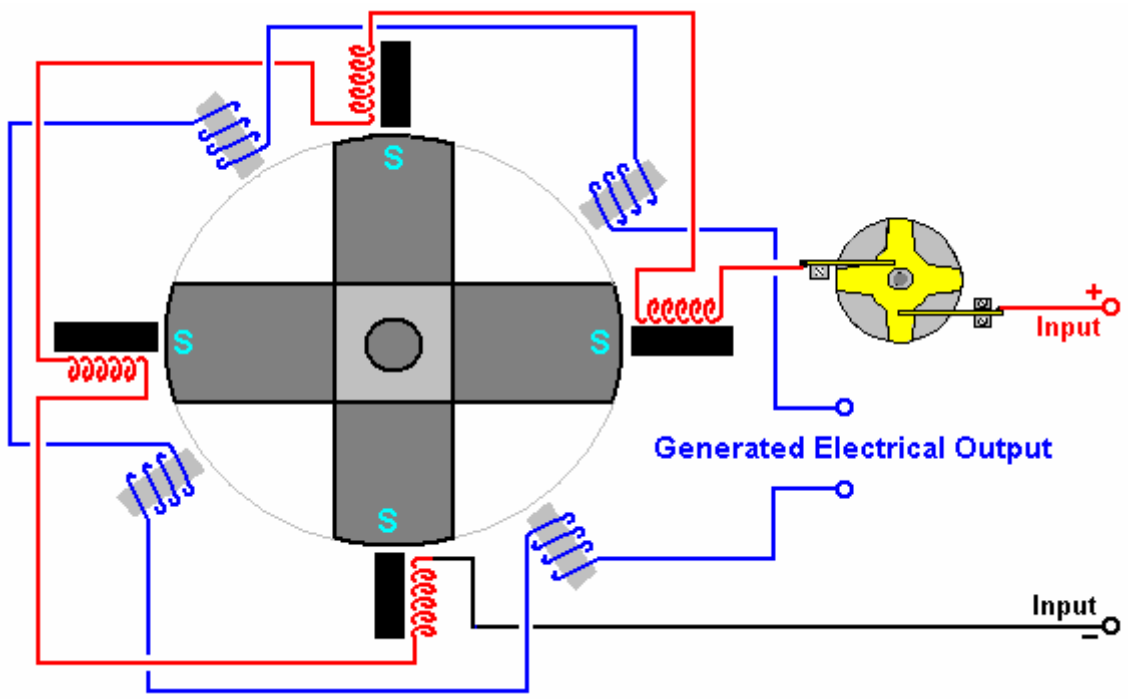


There are many high-performance alternatives which need special skills or equipment to build, but these can be found and constructed by those people who have above average skills and workshop facilities.

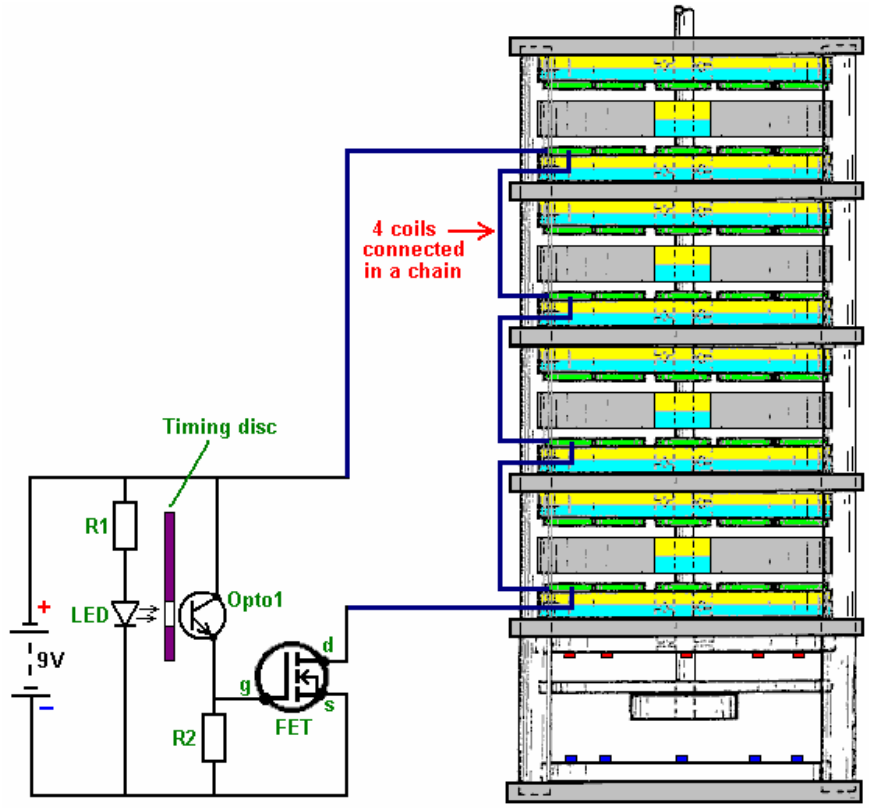
Devices which can provide power at any time, and at any location, include running a standard electrical generator with water as the only fuel (chapter 10). Strictly speaking, the generator runs on energy drawn from the environment and not on water.

While it is perfectly possible to run this type of generator with water as what appears to be the only fuel, it must be realised that a generator of this type produces noise which will not be acceptable for neighbours if the user lives in a congested city environment. Admittedly, a suitable housing with many carpet-covered baffles would allow good air flow and cooling while reducing the noise to very low levels, but generally speaking, this is a solution for people who like working with internal combustion engines and who live some distance away from other people.

The Adams motor (chapter 2) when accurately built, can produce eight times as much output power as the power needed to make it operate. This is a good energy gain and the device is not difficult to build:

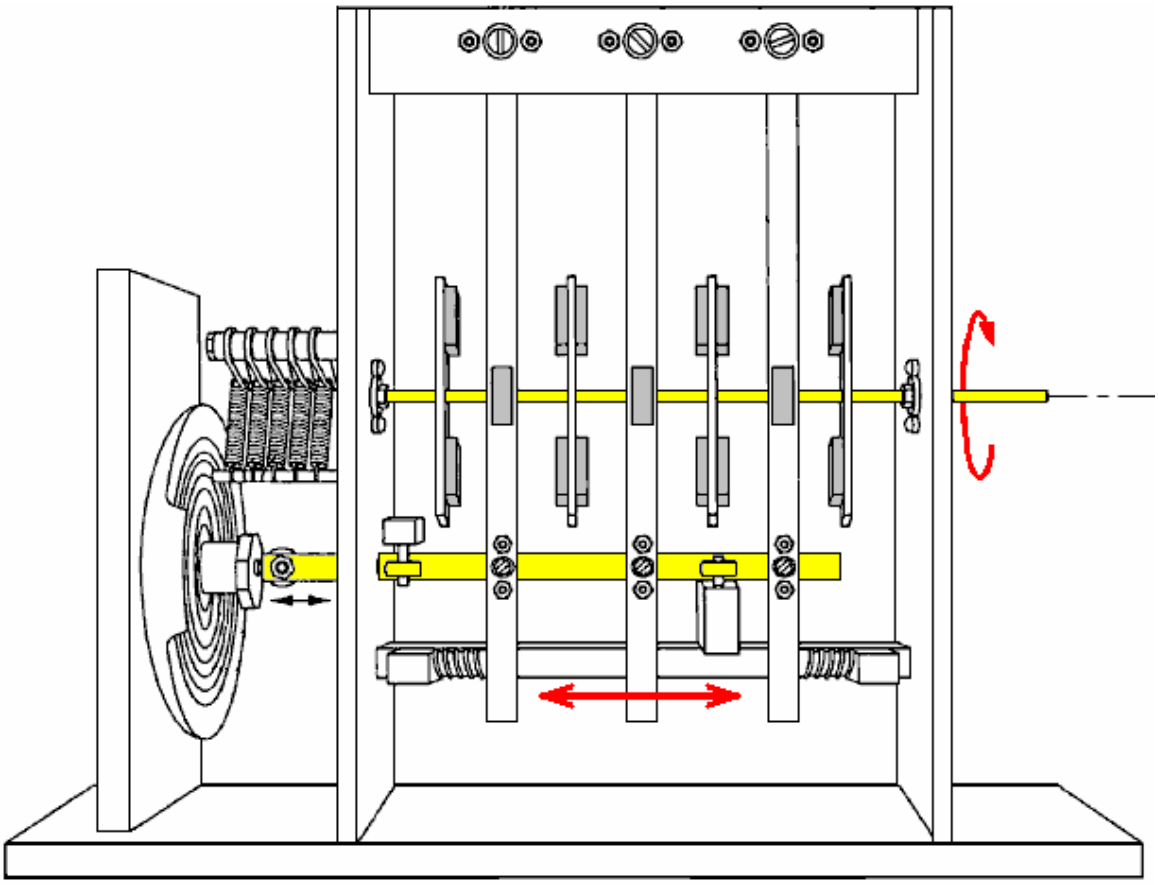


Another device which is not hard to build is the Charles Flynn magnet motor (chapter 1):



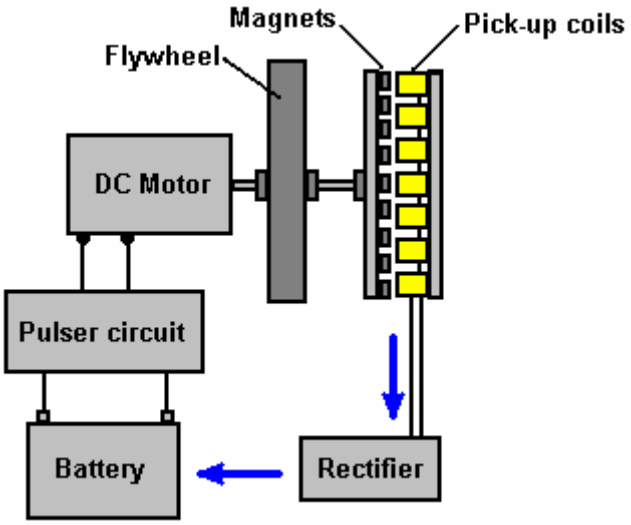
And if the electronics used to drive it is something which you have not come across before, then chapter 12 shows you how to understand and make these kinds of electronic circuits. A friend who is familiar with circuitry might help. A motor of this type can be made to produce any amount of power. Flynn's prototype ran at 20,000 rpm driven by just an ordinary 9-Volt dry cell battery. A big advantage with a motor like this is that you understand exactly how it works, and as you built it in the first place, if it ever breaks down, then you can fix it.

Another possible simple project is the Kundel magnet motor (chapter 1). Using just a simple rocking arm movement, powerful shaft rotation is produced. That rotation can be used as a mechanical drive, or to spin magnets past wire coils to produce electricity:



This motor operates by moving the magnets on the rocker arms, away from the rotor magnets when they would oppose the rotation and closer to the magnets on the next rotor which would promote the rotation. The rate at which the loudspeaker cone electromagnet is driven by an electronic circuit, controls the speed of the motor.

Another device which can supply extra power is John Bedini's pulsed flywheel (chapter 4). The objective is to apply a pulsed drive to a heavy flywheel. Each pulse acts as an impact, providing excess power to the flywheel. John had a small one of these running in his workshop for more than three years.



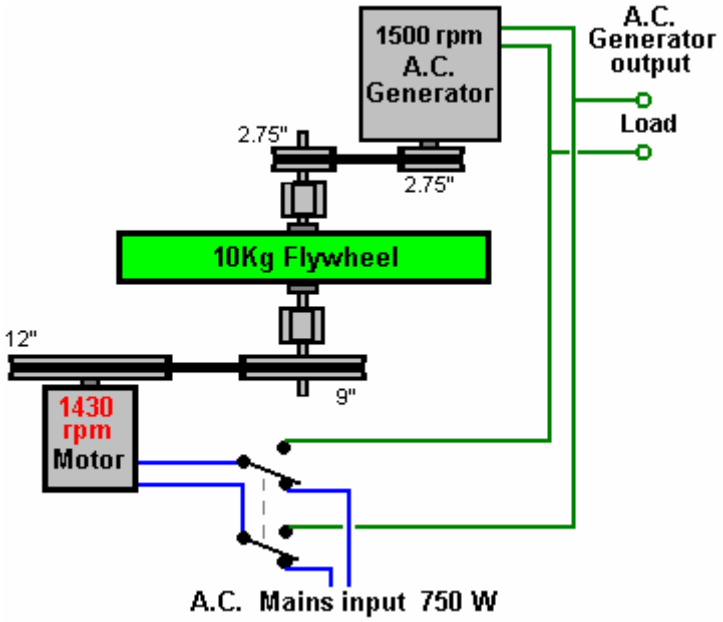
The DC motor is pulsed by a simple electronic circuit and the flywheel spins permanent magnets past a circle of coils of wire. The voltage generated in the coils is converted with four diodes to produce a pulsing DC voltage which can charge the battery and do additional work as well.

The output from a device of this nature depends on the size of the device. Jim Watson built one which was twelve feet high and he got 12 kilowatts of excess power from it. Obviously, I would not suggest that you build one as big as that, but perhaps one of say, three feet in diameter might have a very useful output. You have the option of wiring the coils in a chain in order to have a higher voltage output. If you do that, then have the same number of coils as permanent magnets so that they all pulse at the same moment. Alternatively, if you want to use magnets mounted with every second one having a different pole facing outwards and still have the coils in a chain, then have twice as many magnets as coils.

A similar wheel-pulsing method is used in the water-jet generator (chapter 4) which looks like a very simple thing to implement. Here, a jet of water provides a jerky drive to a wheel when the jet of water hits paddles attached to the rim of a wheel, geared through to an electrical generator.

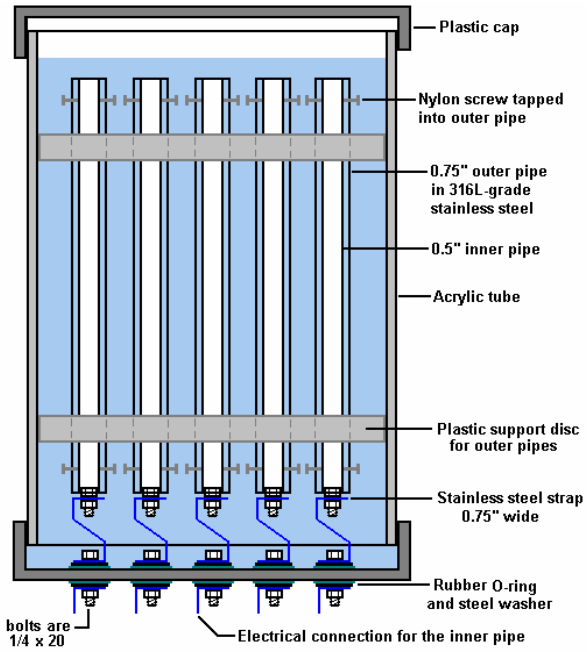
http://www.youtube.com/watch?gl=GB&hl=en-GB&v=zlinM1wAI5U has a video showing this arrangement in operation. If you decide to build it then please arrange for an external change-over switch located in a dry area outside the box for switching the pump over from the mains supply to the output from the generator as that is not something which should be done with a plug and socket in a wet area.

One of the most straightforward projects would be replicating Chas Campbell's flywheel system. Perhaps an arrangement like this:



Or if you prefer, an arrangement where there is more than one flywheel, again, keeping the flywheel shaft input drive pulley larger than the output pulley on that shaft.

Moving to a somewhat more difficult, but quite possible project. Dr Scott Cramton's version of Dave Lawton's electrolyser (chapter 10) is capable of generating a serious HHO gas output of 6 litres per minute of high quality gas for a very small power input.



This project needs patience as each of the tubes needs to be filed down so that they all resonate at the same frequency. Some minor electronics needs to be built and the pipes either conditioned or alternatively, insulated. The gas output can be used for heating, welding, cutting, cooking, boosting a vehicle or running a generator on water.

If you decide to start some project, then whatever project you pick, the most important thing is that it should be one which interests you and building it is fun. You will notice that the projects suggested here generally have moving parts which make it easy to see how the device is operating. The more difficult projects where there are no moving parts and meters need to be used can be left for a later time. If you decide to build something, then let me wish you good luck with your project.

A Matter of Opinion

This eBook is just a sharing of the information which I have come across during some years of investigating devices which are loosely described as being “free-energy” devices – a description which some people dislike because if costs money to construct or purchase a device which can produce useful work, even if the device appears to be self-powered. I am not an expert in the subject and so I have attempted to present the information in a direct and factual way.

However, I am regularly asked for my opinion on various devices and which one is best suited to some particular person, so, reluctantly, I will share some personal opinions on a selection of devices. Please understand that the following is just my opinion and could very well be completely wrong. Your level of interest in any particular device depends on your background and whether you are interested in a small device which proves that free-energy exists and can be harnessed, even though the output level is very small. Or if your objective is to obtain a device which can reduce your expenses. Or if you just want to investigate the subject for the sake of general interest. So, we have a wide range of different devices and a range of possible interests, so an attempt has been made to divide the devices into categories based on ease of construction and the output power levels produced by each device.

Please bear in mind that there has been a very high level of opposition to any of this type of information becoming well-known and so, information on each device tends to be limited, and while half a dozen devices may well be offered for sale in 2013, at the present time few are commercially available, leaving you to build your own or get a friend to join with you in the construction, which is actually a very good idea as small groups of people can have a very wide range of skills.

Devices Which are the Easiest to Build

The Magnetic Frame of Lawrence Tseung (chapter 3)
The Transformers of Thane Heins (chapter 3)
Stephan Leben's adaptation of the Flynn Magnetic Frame (chapter 3)

These three devices have no moving parts and their performance has to be determined by measurements of the input power and the output power. They all operate by passing a fluctuating magnetic field through a frame, which is typically made from thin strips of laminated steel. These devices are limited in their power levels by the amount of magnetic flux which can be carried by their magnetic frames, and to a much lesser extent, by the maximum current which can be carried by the coil windings placed on those frames. Once the core reaches magnetic saturation, applying a stronger magnetic field to it has no effect at all and just wastes the additional power. So, in order to get any form of powerful output, the magnetic frame or core needs to have a high magnetic capacity, which typically means that the frame has to be big, which in turn, means it will probably be heavy and it may be expensive. Using specialised materials for the core could make a major difference, but most constructors will not have access to any of these nanocrystalline materials.

When properly constructed and operated, these devices can certainly show you that COP>1 is definitely possible. However, as the results are based on measurements, it is generally difficult to convince anyone else as many people will presume that you are operating a hoax as they have always been taught that any such operation is "impossible" as there is no such thing as "a free lunch". If you do want to comment on the input and output power levels, then be sure to measure the DC voltage and the DC current on both the input and the output as it will be argued that any form of waveform will give measurement errors.

The most convincing thing is if the device supplies it's own input power as well as supplying some additional item, even a single LED. Where the COP is greater than 1, this should be possible, but to keep the device operating steadily, I suggest that you use a voltage-stabilised circuit to supply the input power as that prevents a runaway situation where increased output power produces increased input power and a positive feedback is created which could lead to the almost instant destruction of circuit components which have been driven well past their rated values.

Clemente Figuera's Split Transformer (chapter 3)
The Figuera split transformer is exceptional in that it is completely free of the usual crippling Lenz-Law effects and it has the capability of outputs of several kilowatts. The details have been lost for more than a century and only recently been made available, so considerable experimentation is needed but the possible results are spectacular. If electronics does not appeal to you, then using mechanical switching like Clemente did, is a perfectly feasible method of construction. If you want to supply the input power from part of the output, then I would recommend using a stabilised power supply to ensure that only your chosen voltage is supplied as input power.

Carlos Benitez's Generator (chapter 5)
This spectacular design is capable of kilowatts of continuous excess power through the use of just two batteries. The design documentation explains how to calculate the size of the components and the mechanical switching is only used very occasionally and so can be replaced with simple solid-state switching to avoid mechanical construction.

Lawrence Tseung's "FLEET" Toroid (chapter 5)
This is a very easy device to construct and use, but it's output power is very small. Demonstration kits are due to be offered for sale quite soon. Developments are in hand and I have been told of a version which can run at six kilowatts. The presently known versions appeal enormously to many people who are intrigued that an almost totally exhausted single dry cell battery with far less than one volt, can charge a 6-volt battery or light LEDs for long periods. While this is definitely a 'fun' device, it is unlikely to convince anyone that 'free-energy' exists. A one-kilowatt version would be a very different matter. I have confirmed COP>1 operation by using one 12V battery to charge another, swapping the batteries over and repeating the process. The result was that both batteries ended up with considerably more genuinely useable power.

The Aerial Plate System (chapter 7)
This is also very easy to construct and being a modular system it has the advantage that it can easily be expanded to give substantial power outputs. This device is based on a Nikola Tesla invention and it has nothing to do with radio transmissions as it picks up ambient energy from the environment. While it can be used to charge batteries, it might also be used for direct power supply if suitable control circuitry is constructed for it. It needs a very high aerial and an extremely good earth in order to perform properly.

John Bedini's Simple School Girl Battery-pulser (chapter 6)
Alexkor's Solid-state battery-pulsers (chapter 6)
These devices can be very effective in what they do but they are not likely to convince anyone of COP>1 operation. Ron Pugh of Canada has achieved COP=13 with his implementation of the Bedini device running on 24-volts. It is difficult to demonstrate the actual performance as batteries are capable of getting a false "surface charge" which shows a high voltage but does not indicate a genuinely high level of recharging for that battery as the voltage will drop rapidly when driving a load. The only reliable way to assess the state of charge of a battery is to make it drive a load and observe how long it can sustain that load. The most recent, very-simple, solid-state Alexkor circuit has a COP=12 performance.

The Bedini pulse-charger has an impressive spinning wheel or rotor, but it's usefulness is severely limited as it can only charge batteries which are not being used to power a load. That means that it's charged batteries are only occasionally available, or two sets of batteries are needed, one for use while the other set is being charged. Ordinary batteries are not all that useful as their life is severely shortened if they are discharged in less than a 20-hour period.

Joseph Newman's Motor (chapter 11)
This is a particularly easy device to build and one which had an experienced technician considerably bothered by it's performance (as shown in Joseph Newman's book which can be downloaded from http://www.free-energy-info.tuks.nl/) as he could not explain how it could possibly perform as it does. However, the amount of power from any reasonably-sized replication will be low.

Internal Combustion Engine Boosters (chapter 10)
Simple boosters such as the "Smack's Booster" can be built quite easily and they can improve the burn quality of the standard fuel used in an internal combustion engine very considerably, reducing the unwanted emissions and raising the mpg figures by a typical amount of 20% to 40%. While building and using a booster, the actual installing of one in a vehicle does require some basic vehicle knowledge and a few basic tools.

Dietmar Hohl's Magnet Motor (chapter 1)
This motor has been replicated and proven to work. Given a little care when constructing it and angling the stator magnets as shown gives a considerable improvement in the power available. However, as shown with just one V-set of magnets and one cylinder of stator magnets, the power is not that great although it demonstrates conclusively that permanent magnets can be used to provide power and demonstrate "perpetual motion" which is so much feared by "scientists" because it shows that their pet theories are, at best, incomplete and at worst, completely wrong. To get some power from this motor design, you need to use a larger diameter drum with two, three, four or five sets of V-magnets around it's circumference, and possibly more than one set of magnets along the length of the drum. The number of magnets involved in even the minimum arrangement is quite large.

Ben Teal's Motor (chapter 1)
This motor design is very, very simple and the first version was built using wood and it proved to be very effective, having a two-inch (50 mm) diameter drive shaft. Ben found that he was unable to slow or stop the motor by grasping the drive shaft as the motor was just too powerful. This motor has not been shown to be COP>1 but I wouldn't be surprised if is.

Don Kelly's Magnetic Motor (chapter 1)
I have never heard of anybody building this design. It has the potential for serious COP>1 operation as the driving power comes from the magnets and only a small amount of power is needed for the motors to move the stator magnets. I see synchronising the motors as being a potential problem and so would suggest that using only one motor to drive all of the stators as being a sensible option. The motor speed during start-up can be controlled manually with a DC Motor Speed Controller and then the setting just left when it is running at full speed. It could be used to drive sets of magnets between coils to generate electricity and that output boosted substantially further with the coil-shorting method described in chapter 2.

Stephen Kundel's Magnetic Motor (chapter 1)
This is a very simple and direct method of getting permanent magnets to provide continuous drive with just a very low electrical input used to change the location of the stator magnets in much the same way as in Don Kelly's motor design. The movement is very simple as it is just a rocking movement, driven by the mechanism used in a loudspeaker. As there can be as many sets of rotor and stator magnets on any drive shaft, the power of this design can be substantial.

Raoul Hatem (chapter 2)
This arrangement of magnets and generators is expensive as each generator uses thirty-six powerful magnets to give good magnetic coupling between the drive motor and every generator. Some people believe that this design cannot possibly be COP>1 as the motor drives each of the generators (although most are driven indirectly). I don't know if that is correct or not, but there is the possibility that as the motor creates a rotating magnetic field

whether or not there are any generators there at all, it may be that the fact that a nearby generator happens to use that rotating magnetic field to rotate its armature, has little effect on the drive motor. Certainly, Raoul didn't keep buying and adding extra generators if he was not getting any benefit from them. This arrangement has the potential of serious power gains, but it would be fairly expensive to investigate if that is indeed the case.

John Bedini's pulsed flywheel (chapter 4)
Lawrence Tseung's pulsed flywheel (chapter 4)
These are definite winners. Jim Watson built a monster size version of John's which was self-powered and had 12 kilowatts of excess power. Lawrence Tseung's prototype is only low power but it demonstrates COP>1 very clearly indeed on only one set of rotor magnets. The design performance should be dramatically increased by using several sets of rotor magnets as that would raise the power levels considerably. This is a very easy home-build project requiring very little in the way of constructional skills or specialised tools.

Lawrence Tseung's Impact Gravity Wheel (chapter 4)
This is a device which has been replicated and it too demonstrates the "impossible" perpetual motion so feared by our revered "scientists". It requires minimal skills to replicate and the components are all low-cost. Definitely and interesting project to build and investigate and one which has the potential of power generation without the need for any form of input power or fuel.

The Nudged-weights Design of Mikhail Dmitriev (chapter 4)
This is a design which I have only encountered very recently and it has major potential. It does with weights, what designs like Kundel and Kelly do with magnets. Energy is extracted from the gravitational effect (if you will pardon the expression) as weights are pushed downwards. The lever arm for those weights is longer than the lever arm when they are being lifted upwards again, and that gives a continuous turning force. The increased lever arm for the falling weights is caused by a small electric motor giving them a nudge away from the rotor axle. Before the weight has time to swing back into it's vertical position it's rotor pivot has reached the bottom of it's travel. In some of his designs, Mikhail just allows the weight to swing freely and in other designs he uses an angled ramp to control and limit the positioning. The brief presentation in chapter 4 shows tiny weights, but even for a first-time home-build replication, there is no reason why the weights should not be substantial. They can be suspended and given a ball or roller bearing suspension and the motor can have broader strips which press directly on the rectangular body of the weight. Mikhail is thinking of marketing ready-built units providing 6 kilowatts or more of electrical power, and plans and kits as well. The rotor turns slowly and powerfully and so need to be geared up to increase the speed for electrical generation.

The Gravity Wheel Design of Abdulsalam Al-Mayhi (chapter 4)
This device is very simple to build, being entirely mechanical and requiring no complicated parts. It is essentially, a wagon wheel with spokes which have sliding weights on them. If the style of construction shown in chapter 4 is used, then please be aware that a large number of roller bearings are needed. It has been suggested that spokes which have a square cross-section rather than round as shown, might be an advantage.

Chas Campbell's Pulsed Flywheel (chapter 4)
This is a good design for people who do not want to get involved in electronics as it uses standard mechanical components and an off-the-shelf motor and generator. I don't know of anyone who has replicated this design, but since nobody has any particular reason to tell me about what they have done, that does not mean anything much. The potential for excess output power is high although Chas has not made any output power measurements.

James Hardy's Water-jet Generator (chapter 2)
This device is both patented and shown on video on the web. It is very, very simple in concept and should be easy to build. Unlike the video demonstration, a proper working model is wholly enclosed and fairly quiet in operation. It is self-powered and has significant excess electrical power. Construction is very straightforward and uses readily available off-the-shelf components.

The Robert Adams Motor/Generator (chapter 2)
If built without paying any attention to the essential details, this design will operate but always be COP<1 in spite of the design having an extremely high potential. If care is taken with the construction and high voltage used and the timing adjusted to it's optimum settings, then an input of just 27 watts has been shown to produce 32 kilowatts of power, capable of running a household or office. Adjustments for optimum performance will take time and patience and this design is more suited to people who have some understanding of electronics.

Pyramids (chapter 9)
The usefulness of a pyramid depends to a considerable degree on how sunny the local environment is. The bulk of the energy which they concentrate flows from the sun, so sunny areas have a much higher pyramid effect than dull areas where direct sunlight rarely reaches the ground. A pyramid is not particularly difficult to construct if you have the dimensions, but size is very important with increased size having a disproportionate effect. In this area, bigger is definitely better. Covering the pyramid with gold-plated iron is the optimum, but many other much

cheaper materials can be used very effectively. Pyramids are probably better used to enhance growth and for health improvement purposes rather than for generating electrical power.

The Tesla Coil and Grounded Plate (chapter 5)
This method for extracting substantial power was disclosed by Nikola Tesla and has been demonstrated in broad outline on video by Don Smith, although he did not demonstrate equipment being powered by the electrical energy generated. A standard Tesla Coil (available commercially) directed on to a metal plate or two-plate capacitor which is connected to an effective earth wire, can supply serious amounts of electrical power. Nobody seems to be interested in using this method to generate excess energy and it is not at all obvious why that should be.

Art Porter's Magnet Motor (chapter 2)
This is a variation of the Charles Flynn motor where an electromagnet coil is used to modify the magnetic field of a permanent magnet. Art states that he gets a 2.9 times improvement in the field strength doing this and he uses it to operate a reciprocating motor. Instead of having mechanical movement, it seems as if it would be more effective to use that magnetic field alteration to generate electricity in a stationary coil as that seems likely to give a COP>1 performance and the currents used could easily be high, giving substantial excess output power. It should be perfectly possible to have a whole bank of these magnet/coil combinations in a single device.

Devices which are of Medium Difficulty

Charles Flynn's Magnet Motor (chapter 1)
This design is powered by permanent magnets which are intermittently shielded by electromagnet windings. As the construction of those windings depends on the voltage being used to drive them, the spacing between the magnets and the strength of the actual magnets used in the construction, experimentation needs to be undertaken before the final construction and assembly is undertaken. An understanding of basic electronics is needed. The power and speed of the motor are exceptional. The input power needed for the prototype described by Charles is very low with just an ordinary 9-volt dry-cell battery being able to power the motor at 20,000 rpm. If the motor is used to drive a generator (whether home-built or commercial), there does not appear to be any reason why the input power should not be provided by a voltage-stabilised power supply driven by the output, making this a self-powered system which can power other electrical equipment or a mechanical load.

Electrical Generator Running on Water (chapter 10)
As far as power output is concerned, this is a major winner as it is both self-powered and provides kilowatts of excess electrical power. As so much electrical power is generated, the design of the electrolyser used should not be critical. While the generator spark timing can be delayed so that just HHO and cold water mist can be used, the David Quirey method of bubbling the HHO through acetone, allows an unmodified generator to be used, making the project very much more simple. This project is particularly suited to people who like maintaining, adjusting and tuning vehicles.

The Pulsed Coil of 'UFOpolitics' (chapter 3)
This is a design which generates cold electricity which can safely light a bulb under water. With no moving parts, the only (minor) difficulty is constructing the electronics to drive the coil. Running off 36V, this is a very interesting design for people who want to experience the remarkable effects of cold electricity.

Bill Muller's Motor/Generator (chapter 2)
Ron Classen has replicated this design and reached COP=2.2 after very considerable expenditure of time and money. You need excellent mechanical skills to be successful with this design although the coil-shorting technique might well raise the electrical output very substantially.

The RotoVerter (chapter 2)
This design involves running a 3-phase motor from a single-phase power supply. The capacitor tuning needed to do this depends on the loading of the output unless Phil Wood's method of adding in a DC motor is used. This project mainly requires mechanical skills rather than out and out electronics. This arrangement is best suited to where an existing motor is used extensively and under those circumstances, the cost of running the original motor can be reduced substantially by the RotoVerter system and the system can be run on a solar panel used with an inverter.

Richard Willis' Motionless Generator (chapter 3)
Silverhealtheu's adaption of the Willis Generator (chapter 3)
Richard has shown his generator on the US version of The Dragon's Den television programme and he has offered his "Magnacoaster" generators for sale. I understand that the technical problems initially encountered have now been overcome and so units should be shipping at this time. The simplified 'silverhealtheu' version

should not have any problems when run at reasonable power. These devices are a starting point for experimentation and they are best suited to people who have some knowledge of electronics.

Dan Davidson's Acoustic Generator (Appendix)
Pavel Imris' Optical Amplifier (chapter 3)
These devices appear quite straightforward, but I have never heard of anyone actually attempting to replicate either of them. That being the case, they would need to be categorised as being designs which need further investigation if they are to be home-built.

Dale Simpson's Hinged Plate and Gravity Wheel (chapter 4)
Murilo Luciano's Gravity Chain (chapter 4)
These three devices have to be categorised as being untested designs and some people are of the opinion that they can't work. Until such time as they are built and tested, they have to remain as design ideas. The construction is not particularly difficult being mainly mechanical in nature, but there is no guarantee of success and it seems unlikely that the components could be easily used again in some other project.

Dave Lawton's Hydroxy Cell (chapter 10)
Dr Scott Cramton's Hydroxy Cell (chapter 10)
Bob Boyce's Electrolyser (chapter 10)
These three designs have very impressive performances, far outperforming Michael Faraday's efficiency results (which "scientists" believe are the maximum possible). For example, Bob Boyce's electrolyser design run on DC, can have efficiency results over 200% those of Faraday. Run on pulsed DC, the same electrolyser can reach 1,200% Faraday's maximum. The Lawton and Cramton cells, operating differently, also have outputs which are far in excess of Faraday's results. As more power is available from recombining the gas mix produced through burning when diluted with air and/or non-combustible gasses, or through use in an internal combustion engine, these devices represent considerable possibilities. However, the amount of work involved is not trivial and the use of the gas mix produced calls for ingenuity, so these are devices which have a high potential but they are not an instant-fix for anything.

The Rosemary Ainslie Heater (chapter 5)
This is a most interesting design which has COP=17. However, at the present time, power levels are so low that while it may be an interesting project and the principle is more than interesting, the output of this device is too low to provide any form of useful heating.

Robert Patterson's Ram Wing (chapter 10)
This is a simple physical device capable of improving the air flow around a moving vehicle to such an extent that it produces better mpg figures for the vehicle. Constructing one and mounting it on a vehicle is not a difficult task and the gains can be significant.

Co-Axial Cable Electrets (chapter 9)
This is a system which nobody seems willing to try. It involves baking a reel of co-axial (TV aerial) cable in an oven and allowing it to cool down very slowly. It is capable of supplying 10 mA at 10,000 volts which is a continuous power of 100 watts (better than most solar panels). The output energy from this passive device is drawn in from the surrounding environment. The high voltage of the output appears to put people off this device.

The Colman / Seddon-Gillespie 70-year Battery (chapter 3)
This is a device which has major potential and yet one which nobody seems interested in attempting to replicate. It is a small, compact, simple device which can provide one kilowatt of excess power for an estimated seventy years. Several of these in a single small housing could supply the electrical needs of a whole household. The materials used are not hazardous in any way and apart from some basic electronic circuitry and experimentation, there does not appear to be any problem with attempting to replicate it.

Robert Krupa's "Firestorm" Spark Plug (chapter 10)
This plug design is quite capable of giving improved fuel burn inside an internal combustion engine that the mpg performance can be increased very substantially and harmful emission drastically reduced. While Robert has a patent on his designs, he has not been able to find any manufacturer willing to make and sell Firestorm spark plugs. Prototypes have been made by getting a jeweller to modify a new set of standard spark plugs. This is expensive, but it is definitely something which is possible to do.

Jesse McQueen's Self-powered System (chapter 13)
On the surface, the patented proposal by Jesse appears to be impossible. However, as our understanding of how the various free-energy devices draw in environmental energy increases, some of the more impossible looking devices suddenly look as if they might actually be based on principles and methods which have been proven to work. Essentially, a battery powers a motor which spins a generator which keeps the battery charged and powers other devices. If both the motor and the generator are COP<1, then the system won't work. But, gearing is used

between the motor and the generator, and Chas Campbell has shown that with a flywheel in that linkage, additional power **is** actually available. It may be that Jesse does not mention the fact that his gearing linkage is heavy. It is interesting to note that the gearing that he specifies, has the correct ratio that Jacob Byzehr specifies in his patent for an energy gain through such linkage. It seems likely that Jesse has not disclosed all of the details in his patent. While constructing a test bed for this system should be easy if you have the components to hand, I suggest that you consider it to be an experimental investigation of his claims and don't expect immediate, spectacular results.

Robert Tracy's Permanent Magnet Motor (chapter 1)
This style of motor with magnetic shields being physically moved between pairs of magnets is probably one of the more difficult magnet motors to get operational. The main problem is getting an effective magnetic shield material. The expensive "mu-metal" material is supposed to be effective if it is fairly thick, but I am by no means sure that this is the case in actual practice. Screening powerful magnets with a passive material is not an easy thing to do and there are very few materials which seem able to do it. It might be possible to use a thin neodymium magnet as the 'shield' if it were orientated so that it was in repulsion mode for both the stator and rotor magnets as they approach each other. Because of the screening material problem, I could see the building of a Tracy prototype to be a lengthy experimental process in spite of the fact that the theory is perfectly sound.

Devices which require higher skill levels

Muammer Yildiz's Permanent Magnet Motor (chapter 1)
This motor has been demonstrated in a Dutch technical college and taken apart after the initial running was completed, showing that it's operation was wholly genuine. However, some of the magnets used are of a tapered type which are probably not readily available unless ordered as a specially constructed batch which is generally an expensive way to buy magnets. Magnetic fields and their interactions and distortions are not the easiest things to understand or observe in a practical situation, and that makes the construction of permanent magnet only motors a difficult job. Howard Johnson and Mike Brady managed to make successful prototypes but then found it almost impossible to make others to that same design. So, if you decide to replicate Muammer's motor, please understand that it is not likely to be an easy thing to do.

Michael Ognyanov's Self-Powered Power Pack (Appendix)
This is a low power device intended to operate the lights on road traffic cones without the need for a battery. There should be no particular problem in building it if it were not for the need to cast a semiconductor block from a mixture of materials such as silicon, germanium, neodymium, gallium, etc. It seems likely that the average home constructor of devices is not geared up for melting and casting such materials, and so that puts it into the 'higher skills' bracket.

Don Smith's High-frequency, High-voltage Devices (chapter 3)
These devices appeal to a high percentage of interested people and yet they seem to be some of the most difficult to replicate. Admittedly, Don does say that he has not revealed all of the details on any one of his designs, but independent people have confirmed that doubling the voltage quadruples the output power and that additional electrical outputs can be produced from the fluctuating magnetic field without requiring any additional input power. Considering the limited success encountered by several people who have tried to construct replications, these devices have to be rated as some of the most difficult to get operational. Having said that, success has been achieved by putting two Tesla Coils back-to-back, connecting their secondary coils with a wire and taking the output from the larger diameter "L1" coil of the second Tesla Coil.

Jerzy Zbikowski's Chain Drive (chapter 4)
This device looks impossible, and yet it is reported to have been laboratory tested at COP=1.47 which is quite remarkable. As it is a mechanical chain drive, it requires good mechanical skills to construct the chain links and the two sprocket wheels. The mechanical operation needs to be high-quality, so it is unlikely that beginners to mechanical construction would have immediate success.

The Magnetic Pendulum (chapter 4)
While I don't know the name of the builder, this pendulum is reported to have been swinging continuously for more than two years now and while it does not appear to generate any useful power output, it is certainly an amusing project, especially should you want to show it to those people who say that 'perpetual motion' is impossible. It might well be possible to adapt it so that the magnetic push provided by dropping permanent magnets downwards near the end of the swing were provided instead by a coil which extracts power and then, when the current is cut off, generates the necessary push from the magnetic field generated by the back EMF. While the output power would be low, it would be very nice to demonstrate power generation with no input power being provided – just pure energy extraction from gravity and magnets (both of which are powered by the zero-point energy field).

The Tesla Switch (chapter 5)
This is one of the most difficult devices to get operational, in spite of the fact that it appeals to a large number of people. There are three possible modes of operation. If the diodes are turned the wrong way round so that they can feed current from each battery, then the operation will definitely be COP<1 but it will be a good deal better than operating without the switch circuit in place.

The second way has only been achieved by John Bedini as far as I am aware. This is where the circuitry is the same but the circuit components and connecting wires are adjusted very carefully to produce circuit resonance. When that happens, the circuit becomes self-powering although there is little or no extra power for other devices.

The third way was developed and tested over three years by staff of the Electrodyne Corporation in America. In this version, the diodes are reversed and they only feed sharp voltage spikes back to the batteries, through the diodes which supposedly don't allow current to flow in that direction. This is a very different form of operation where the operating power flows into the circuit from the local environment. The batteries need to be 'conditioned' through long periods of being operated this way as the 'cold electricity' used in the circuit is the opposite of the 'hot electricity' which the batteries have been using up to now. This long conditioning period is generally enough to make the average builder give up and believe that the circuit just doesn't work. Dave Lawton was faced with exactly the same type of problem when he attempted to replicate Stan Meyer's "Water Fuel Cell". It appeared 'dead' and produced nothing during a whole month of testing, and then it suddenly burst into life, producing large amounts of HHO gas mix for almost no electrical input. Without his exceptional patience, Dave would never have succeeded. I believe that the same applies to the Tesla Switch when wired correctly with the diodes blocking current flow from the batteries – it is likely to take long-term and patient testing before the system swings into life.

One experimenter who did not believe the diodes could possibly work that way round, tested the arrangement and discovered that in spite of the theory, in practice, the reverse-biased diodes actually pass very sharp voltage spikes to the batteries, so the effect may well be like a slick version of John Bedini's battery pulsing circuits.

Hermann Plauson's High-Power Aerial Systems (chapter 7)
Hermann Plauson's patent which is in the Appendix, is so detailed and authoritative that it gives every indication that he is highly experienced in the actual design, construction and installation of serious aerial installations. He casually refers to "small" installations as being those of 100 kilowatt or lower capacity, so we are not talking about milliwatts here, although I understand that Hermann extracted somewhat under one kilowatt per aerial and used many aerials in any one installation. The difficulty with these types of system is that the input is very high voltage 'electrostatic' electricity, usually though of as being useless for power generation. It is therefore, important to consider carefully the mechanisms which Hermann has in his patent for converting that power to ordinary electricity at normal voltage and current levels. We need to remember that Paul Baumann's "Thestatika" devices operate off this type of input power and yet put out kilowatts of mains electricity, so we know that it is perfectly possible. We do not have step-by-step building instructions for these types of system, and so they have to be rated as requiring more than the lowest skill levels. Also, many people are not located where they can erect a tall aerial, but remember that the Thestatika uses a Wimshurst style electrostatic generator rather than an aerial, so experimenters can work with a generator input. A.D. Moore's "Dirod" electrostatic generator seems like the most suitable home-build generator. It is described in detail in his book "Electrostatics – Exploring, Controlling and Using Static Electricity" Second Edition ISBN 1-885540-04-3 which is the only source for construction information on his generator.

Frank Prentice's Ground Aerial System (chapter 5)
Frank's design uses a one mile long wire supported just a few inches above the ground. He inputs a modulation signal of 500 watts and extracts 3 kilowatts of power. The information in his patent is not exactly overwhelming in it's details, but the major obstacle is that very few people can install any such wire. It is not known if the wire could be run backwards and forwards over a small area, but if that were possible, then it would form a barrier against the normal use of that area. Frank didn't have that problem as he worked on railway tracks. The limited amount of information adds to the skill level needed to be successful with his design.

Alfred Hubbard's Electrical Generator (chapter 5)
This has a very high potential as it can put out substantial electrical power in spite of it's simple construction. As it is based on just one section of the magnetisation graph where a small additional current produces a major rise in the magnetisation of the iron core, it will require careful adjustment and attention paid to the magnetic saturation levels of the materials used. It is very interesting that Joseph Cater's analysis of the device has spacings which match the recently discovered optimum tube spacings for the Joe Cell. I am not aware of anyone who has managed to replicate Alfred's device.

Joseph Cater's Water-filled Generator (chapter 5)
This looks like a really good design but it has the problem that it needs a transducer or klaxon running at 600 kHz and there does not appear to be any such transducer readily available on the market. The highest frequency I

have located was just 300 kHz which is clearly not suitable, so this design appears to be restricted to people who have skills which can design and construct a klaxon which can operate at 600 kHz.

Joseph Cater's 'Orgone' Box (chapter 9)
The 'Joe Cell' (chapter 9)
The 'Nitro' Cell (chapter 13)
In theory, each of these is capable of running a car engine without the need for any form of combustible fuel. I am aware of people who have succeeded in doing this with the Joe Cell and the Nitro Cell, but they both live in much sunnier locations than I do. These devices are some of the most difficult to get operational and it could well take a week for the car to absorb the cell influence before it can run without fuel. It has to be said that the technology involved in the operation of these devices has not been fully documented at this time and so is not fully understood and so experimenting with them is effectively a research project.

Fuel-less Engines (chapter 8)
These engines are perfectly viable but few people have the knowledge, skills, tools, workspace, inclination and time to go building or modifying an engine. These are Engineering tasks and not generally suited to 'back-yard' projects.

The Searle-Effect Generator ("SEG") (chapter 11)
According to John Searle, constructing just one of the many magnets needed for his generator requires more than a day of work and a million pounds in finance. This is enough to put most people off even considering his design. However, John certainly didn't spend anything like that amount of money when he built his first successful prototype, so we can probably ignore the quoted costs. Investigation would need to take place through the enthusiasts on John's forum and web site, but unless something has changed recently, it will probably take considerable research and subsequent thought in order to come up with a viable design for the construction of a useful generator or a flying disc.

Michael Eskeli's 'No-Work' Turbine (chapter 11)
Michael is clearly a gifted and very competent man who has many patents. In common with most gifted inventors, he does not excel in explaining and documenting his inventions and so a certain amount of background Mechanical Engineering knowledge would be of major help in understanding how to home-build one of his designs. The No-Work heating turbine in chapter 11 is a heater which only needs the friction of the bearings to be overcome by the motor spinning the turbine as the actual heating effect is 'free'. This is, of course, spectacular. If you happen to be a gifted individual and replicate Michael's turbine, then I would be most grateful if you would fill me in on suitable construction methods so that I could expand the description into a more detailed version which would encourage people to build one for themselves.

William Hyde's 10 kW Electrostatic Generator (chapter 11)
This involves the construction of a rotary generator with metal bodywork and so metalworking skills are needed (or somebody else with those skills needs to help out or be paid to do the work). The completed electrostatic generator has been patented and it is claimed to have an output of ten kilowatts of ordinary electricity, which makes it of serious interest to anyone needing to power a household. It is reputed to be COP=10 which means that the input power will be substantial for a large electrical output.

Michael Faraday's Homopolar Generator (chapter 13)
This design appeals to most people as it looks so simple. However, the problem is that most versions produce a very low voltage at very high current – potentially a thousand amps or more. Trying to come up with brushes which can carry high current for long periods of time without appreciable wear has proven to be a major problem. The Borderlands Science people discovered how to get an AC output from the device which allows the voltage to be stepped up and the necessary current reduced substantially. To date, I am not aware of anyone managing to produce a homopolar generator with a practical power input and output, but it is definitely a device which is wide open to experimentation.

The Romag and Mini-Romag Generators (chapter 13)
These designs call for construction with a range of different materials, and generally speaking, the materials used in any prototype are normally of major importance. Although these designs have been available for many years now, the only prototype which I have heard of is the 25 watt version built by J.L. Naudin, although there is probably not particular reason why the Romag could not be scaled up to produce a much larger output.

Nikola Tesla's Generators (chapter 11)
Neglecting the version which uses radium as that is probably difficult to obtain at the present time, the spark gap version is said to produce very substantial amounts of electricity while only using very simple components. I have never heard of anyone attempting to build one of these devices but in spite of the fact that we have limited information on the device, it seems like a very powerful unit which could well be extremely useful.

The Jines Permanent Magnet Motor (chapter 1)
This design includes magnetic shields and finding a suitable material for them might be a considerable problem. This design has to be considered to be fairly complicated due to the amount of mechanical work needed to construct the various mechanisms required for the motor.

The Annis and Eberly Generator (chapter 3)
This is a very clever solid-state generator. However, from the point of view of the home-constructor, producing their solid-state reluctance-switching material and it's controller would appear to be a major obstacle. Given that material and it's controller, many of the other devices would be exceptionally easy to build successfully.

The "Carousel" Permanent Magnet Motor (chapter 1)
This is a compact and very effective looking device but the very compact nature of it's construction probably makes it somewhat tricky to replicate. Anyone with good metalworking skills should be able to do it ok.

Devices which are unlikely to be practical

The Ecklin-Brown Generator (chapter 1)
People have had difficulty in making replications of this to work well. The biggest problem area is coming up with suitable magnetic screening material. It might be possible to adapt the design to use some of the very thin neodymium magnets instead of material which just screens the existing magnetic field. The Clemente Figuera design of chapter 3 is a better option.

Howard Johnson's Permanent Magnet Motor (chapter 1)
It appears that Howard has not been able to build replications of his original working prototype and while the theory is perfectly valid, that failure has to put a question mark against almost all permanent magnet motors of that type, especially since Mike Brady experienced exactly the same thing and Australian developers with a successful "Lutec" magnet motor found that they could not replicate it unless they used parts from their original prototype. The V-magnet system of the Hohl Motor in chapter 1 is an exception to this problem.

Mike Brady's "Perendev" Permanent Magnet Motor (chapter 1)
This design is in exactly the same position as Howard Johnson's Permanent Magnet Motor. It seems that Mike has experienced considerable difficulty in constructing working copies of his prototype. If you want to try to build one, by all means feel free.

Floyd Sweet's "VTA" Magnetic Generator (chapter 3)
Shown on video, producing more than 500 watts of mains output power for 1.2 milliwatts of input power, this is clearly a device of major importance. Unfortunately, we do not have Floyd's method of producing isolated and easily moved magnetic poles in a block of magnetic material and that is a major obstacle for the would-be replicator. I am not aware of anyone who has successfully replicated Floyd's device.

Oleg Gritsevitch's Toroidal Generator (chapter 5)
This incredibly successful and powerful device which produced 1.5 megawatts continuously for two years, took major equipment to get it started. It is not known if it can be successfully constructed in a smaller size and the coating between the water in the toroid and the toroid casing may be difficult to achieve.

Tariel Kapanadze's Electrical Generator (chapter 3)
Unfortunately, Tariel has repeatedly refused to reveal how his designs operate. He has also been involved in several very questionable business deals where the other party felt cheated. It is most unlikely that he will ever reveal the details. Replicating it is not really a beginner's project and an overunity.com forum has been trying to replicate it for a long time now.

Hans Coler's Devices (chapter 9)
The passive design of Hans Coler can certainly be made to operate if you can persuade it to begin oscillating. The output is quite small but interestingly, it does demonstrate that we are indeed surrounded by energy which can be tapped for practical purposes. This device needs patience in order to get it going.

The Meyer-Mace Isotopic Generator (chapter 3)
I have never heard of this design being replicated. It calls for a 21 MHz generator which is quite a high frequency for the home constructor. The design itself is certainly simple but getting it to work if it doesn't function initially is not likely to be easy. The Nuclear Magnetic Resonance involved is explained by William McFreey in chapter 3.

Ed Gray Power Tube (chapter 5)
Ed Gray's power tube is unlikely to be COP>1. The genius behind his achievements was Marvin Cole whose ambition was to be a chef. Marvin built a motor which was certified as COP=237 by Cal-Tech's independent testing, but when he disappeared from the scene, Ed was left at a bit of a loss. Ed only produced patent applications in order to reassure his investors and he never wanted to disclose anything of worth. Mark McKay's detailed information in the Appendix provides extensive background and Mark has now written a book on the subject.

Roy Meyer's Aerial System (chapter 7)
I know of only one person who tried to replicate this design. The first time he tried it, it worked ok although it was not aimed in the 'right' direction, but he was never able to get it to work again. So it appears that while the design may well work, it is not easy to replicate and get functioning.

Karl Schappeller's Device (chapter 11)
To be perfectly honest, I have never managed to understand the work of either Karl Schappeller or that of Viktor Schauberger. Not properly understanding the design makes me believe that newcomers to free-energy construction work would have difficulty with this device. That may well be an unfair assessment and if you are able to build and use this design, then please explain it to me so that I can document it properly.

The "HydroStar" and "HydroGen" devices (chapter 13)
These designs have been around for a long time and to date I have never heard of anyone who has managed to build one and got it to work. In theory, they are supposed to allow a car to run with water as the fuel. On the surface, I can see no way that that could happen unless it operates as a version of the Joe Cell, which it very well may do. If that is the case, then great patience will be needed before the vehicle will operate. I would not recommend anyone to build either of these designs.

Francois Cornish's Aluminium-feed System (chapter 13)
This system is not economical in the broad sense in that it takes more energy to produce or reform the aluminium than is given out when it is converted to oxide in the process which produces hydrogen for combustion. It can be economical for the user in that the processed aluminium is less expensive than oil products. However, the oxide residue generated by the process limits the effective applications for the process as a whole. It is not a good constructional project.

The Motionless Electric Generator ("MEG") (chapter 13)
This is a device which most people find very difficult if not impossible to get to COP>1 operation. If the frame is made of Metglas nanocrystalline material, then it might be possible. However, the design has been around for a decade and it has not (yet) made it to commercial production, so I would suggest that this is a device which is not a good choice for a constructional project.

Devices which are not really possible

The ShenHe Wang Permanent Magnet Motor (chapter 1)
While there is a patent for this design, there is not enough information for there to be any reasonable degree of success if you were to try to replicate it. For that reason, it probably should not be considered as being a viable construction project at this time. It may well become available commercially at a later date.

Thomas Henry Moray's Aerial System (chapter 7)
While this device is, without question, a most impressive device which was demonstrated in public on many occasions, Thomas was intimidated and so never released full constructional details of how his device worked. That lack of specific information prevents this from being a realistic constructional project.

Steven Mark's Toroidal Power Unit ("TPU") (chapter 5)
While this unit has been demonstrated on video, showing impressive operation, I understand that it overheats after 22 minutes of operation. That is not a major problem and could be overcome, but the lack of specific constructional or design information makes this a research project rather than a normal constructional project.

Paul Baumann's "Thestatika" Generator (chapter 13)
The late Paul Baumann designed and built several self-powered generators which were demonstrated on many occasions. These are still in the possession of a Swiss religious commune who refuse to release the details as they think that "the world is not ready for this information". The lack of the vital design information makes replicating the device a non-starter which is a great pity since it has a substantial mains power output which could ease suffering around the world.

Muammer Yildiz's "Ocean Star" Generator (chapter 13)
While the demonstration of this device looks quite like a hoax, Muammer's credibility jumped massively with the production of his current permanent-magnet-only motor which has been demonstrated publicly. The "Ocean Star" device was supposed to be able to power a house, but the patent information is not nearly clear enough to deduce how to build the device. Perhaps somebody could ask Muammer to explain it in more detail and disclose the key design details in order to allow home-built replications.

Topic: Electrolysis

Question:
I was reading the D9.pdf file you put together on Bob Boyce's electrolyser and PWM3G circuit. The close up picture of the PWM3G circuit shows only 5 LED lights but below under the ***Testing the Completed Board*** section it talks about 8 LED lights. also on the animated picture (the one with all the components on the board) there are 8 led lights. did bob build another type of board with fewer LED's on it? If so could you explain how to test the different board? Also I was considering buying the board from the Hydrogen Garage website, would you recommend this?

Answer:
The D9 document was created long before the 3G version of the board was produced. The version shown in the document is only there to give an impression of what the board was like at the time. Bob has always been reluctant for any specific details of his circuits to be published which is why the circuit schematic is not shown in the document.

The supplier of your board will tell you how to tune it. I would not expect it to be a different method. Essentially, you start with the frequency variable resistors in their central position and the Mark/Space resistors set to about 10% On time. Then you tune the top frequency of 42.8 kHz to give the best gas output. You then adjust the middle frequency to give the best gas output. Finally, you adjust the lowest frequency channel to give the best gas output. The Hydrogen Garage in the US or Courtierstown Marine in the UK can advise you.

The Hydrogen Garage is run by a very honest and reliable electrolysis enthusiast with many years of practical experience behind him, and there is no reason why you should not buy from there if you want to.

Question:
OK, so if I want to build one of those boosters and I have 14.8 volts in my car how am I going to power 18 plates 5" x 4". That's the only thing that's confusing because there would not be enough voltage

Answer:
I'm not sure where the idea of 18 plates came from. It is likely that your electrics give 13.8 volts rather than 14.8, although that is not critical in any way.

You are aimed at seven cells in a chain so that each cell gets about 2 volts across it. If you put partitions in your box, then with your proposed current you would use just two 5" x 4" plates per cell - a total of 14 plates. If you wanted to provide more plate area, then you would use three plates per cell, making 21 plates.

However, if all your journeys are short and the cell will never get up to its full temperature due to the short time it operates, then you could go for just six cells and with three plates per cell and that would make 18 plates.

If you use separate plates, then you have the difficulty of making good-quality electrical connections between six pairs of plates, and that can be quite difficult to do. I would suggest that you consider the bending method shown at the start of Chapter 10 as that overcomes the need for making any electrical connections inside the cells.

The best plate performance is with two to four square inches per amp of current. As only bending is involved, it would pay off to have a generous plate area, especially if you are using mesh. Mind you, the mesh surface area is not much reduced by the holes as the surface around each hole as it passes through the thickness of the plate is also active. The hole edges are also very helpful in getting the bubbles to break away when the cell is new. Later on, when the cell is conditioned, the bubbles will not stick to the plates at all.

Question:
Is it possible to generate electricity using hydroxy gas?

Answer:
Yes, it is perfectly possible to generate electricity using hydroxy gas:

1. Take a standard electrical generator such as a Honda 6.6 kVA V twin or a Vanguard V twin.

2. Adapt it by filling in the timing key slot and cutting another which lets you set the spark timing to two degrees after TDC.

3. Build any electrolyser capable of producing 5 lpm of hydroxy gas.

4. Feed the generator air, hydroxy and cold water mist from a Venturi tube, pond fogger or miniature carburettor.

5. The generator can run well on that and produce kilowatts of excess electricity over and above the amount needed to power the electrolyser.

This has already been done on at least nine different generators and run 24 x 7 for two and a half years now. However, please be aware that generator engines wear out eventually and will need a rebuild or replacement, so the electricity produced is not "free".

Q: What mmw (efficiency) rating of hydroxy generation is need to achieve this? Also, is it necessary to be using some type of resonance system or will brute force/catalytic electrolysis work?

A: You can use any type of electrolyser you like. Unlike the alternator in a vehicle, a generator is designed to put out a major amount of electrical power on a continuous basis. So, if you take half a kilowatt or more generating the hydroxy, who cares? If the generator is called a 5.5 kW by the manufacturer and you wish to run it at 4.5 kW assuming that the seller is being generous with his rating in order to compete with other brands, then if you use 0.5 kilowatts (12V at 40 amps) for electrolysis you are definitely in business. The two makes of generator quoted, have proven very reliable in long-term use. The addition of very tiny cold water droplets is a key factor in making this work.

Question:
Do you know if the V-twin Vanguard and Honda generator you have mentioned, do not have the waste spark? I would like to buy one, but need to make sure before doing so.

Answer:
The people who have done these conversions say that retarding the spark to 2 degrees after Top Dead Centre was sufficient. Never having worked on internal combustion engines, I can't personally confirm it myself from my own experience.

If you decide to do it and are successful (as I would expect), then I would be happy to produce a document for you which would go out under your name. If you decide to do that, then please take photographs before, during and after your filling in of the timing key slot and the cutting of the new one as that is one item which would benefit from having detailed pictures. Supplier details for the cold water mist device which you pick, be it pond fogger, Venturi tube or miniature carburettor, would be helpful for others following after you.

Question:
Do you have any values for Peter Lowrie's coils, the ones mentioned in the doc are "from a commercial installation with copper sheet wrapped around them" as it is my intention to use a Marine alternator.

Answer:
There are two different techniques for producing hydroxy gas:

1. Pulsed DC signal which can reach 1,200% Faraday.

2. Brute force DC which is usually less than Faraday although with great care, it can reach 212% Faraday.

Peter Lowrie used a marine alternator solely because he wanted three separate 2-volt supplies at 900 amps each. It was because of his very high current that he used an industrial 3-phase power supply choke. Unless you are aimed at currents that high, then you definitely don't need chokes of that massive current capability.

If you are determined to use a similar item, then you need to look at the spares options for professional, commercial equipment. Peter only used a marine alternator for its very high current producing capability and he told me that a second-hand truck alternator would do just as well.

Question:
You said...

"When that gas is added to gasoline droplets in a gasoline engine, it acts on those fuel droplets during the compression stroke, breaking those long chains of hydrogen and carbon atoms into shorter, more active fragments of chains. This produces more active and better combustion of the fuel when it is ignited by the spark, and that extracts more energy from the gasoline, energy which would have been wasted in the catalytic converter and in producing unwanted excess heat in the engine".

This is the first time I have heard that the Hydroxy is doing something BEFORE it is ignited. I thought the Hydroxy ignition was what was breaking down the gasoline into short molecules? Can you explain this more?

Also, can you email me here or privately what you feel are the best Free Energy Devices these days. I think many may be interested in Free Energy Devices to drive their electrolysers. I'm interested in the devices themselves. Also, is anything near to coming to market?

Answer:
On 14th January, Bob Boyce's post to his Hydroxy forum included the statement: "When either hydroxy gas, or HHO, is compressed in an engine cylinder, it is after it has been diluted with intake air, as well as fuel in non-diesel engines. Not only does the dilution in air raise the auto-ignition pressure, but in non-diesel engines the orthohydrogen has already bound to the long hydrocarbon fuel chains and broken them prior to compression. As this bonding occurs, orthohydrogen decays to parahydrogen. It is the energy from this decay which breaks the hydrocarbon chains into simpler hydrocarbons. These simpler hydrocarbons combust much more efficiently and completely. With diesels, the process is slightly altered, as there is less time for bonding to occur, so improvement occurs much better at lower RPM than at higher RPM".

Gasoline molecules vary from chains of seven carbon atoms surrounded by sixteen hydrogen atoms, to chains of nine carbon surrounded by twenty hydrogen atoms. These chains need to be broken down in order for those hydrogen atoms to combine with oxygen atoms from the air in the cylinder, to produce the heat and expansion which powers the engine. That is why the spark is generally about eight degrees before Top Dead Centre, giving time for this to happen. The more broken up those chains are before the spark, the better.

With regards to free-energy devices, I get asked on a regular basis, what device is best for somebody to try building. The question always comes without the relevant details needed to even make a rough guess at a realistic answer. There is generally no information on country, skills, tools, finance, workspace, time available, main objective, local materials available, environmental space, etc. etc.

Obviously, if someone has converted a pick-up to electric operation and could put a water-fuel generator in the back, the kilowatts of excess electrical power would extend the battery range substantially and would recharge the batteries during a meal break or while away from the vehicle.

This chapter is an attempt to answer your question as to what devices you might choose to attempt to replicate. There is massive opposition to any free-energy device being brought to market. Shenhe Wang intended his permanent magnet motor/generator design to be given freely to all countries in the world. I always felt that he had to be completely unaware of how governments who get billions per year in oil taxes and utility company taxes, would view such a venture, not to mention the oil companies who would see their massive revenue stream being threatened by the introduction of a small device which can output kilowatts of electrical power without using any form of fuel. At this time, it appears that Shenhe's generator is being restricted to China, which hardly comes as a shock.

Chas Campbell in Australia has a viable, commercial machine which also produces significant power without the need for any fuel. He can't find anyone interested in manufacturing it and his local power company just doesn't want to know. However, there are several devices which are nearing being brought to market at this time, and I would expect to see them in use early in 2013.

Question:
I have bought a 7-cell, series-connected electrolyser but the terminals are not marked for Plus and Minus connections. Does it matter which way round it is connected?

Answer:
Any series-connected electrolyser will operate well when connected either way around. However, there is an effect if the unit has already been run. If you are aiming for the top level of efficiency of DC electrolysis, having prepared, cleansed and conditioned the plates with Bob Boyce's methods, then the plates will have been conditioned with the current flowing in one direction. If the direction of current flow is reversed on Meyer type cells the conditioning reverses, the performance drops off and then the conditioning rebuilds, bringing the performance back up again. I am reasonably sure that the same thing happens in a Boyce conditioned electrolyser, so if at all possible, try to identify which way round the unit was originally connected and keep to that direction of current flow.

Question:
In Bob Boyce's electrical diagram, it shows three heavy-duty chokes: one from the rectifier bridge positive, one on the negative and one from the positive line to the circuit board. My question is: "what is the value or specification for these chokes and where can you find them?"

Answer:
There is a document called "D9.pdf" which provides the details of Bob Boyce's resonant electrolyser design. On page 30 of that document, immediately under the circuit diagram it tells you the choke core, the supplier of that core, the wire used to wind the choke and the number of turns to be wound on the core.

Question:
What size hose would be required to run a V-8 on all hydroxy? Does anyone have a flow chart for gas, or know were to find/get one?

Answer:
It is not possible to answer your question directly as we don't know what rate of gas flow you think would be needed or how long the pipe has to be. There is a table of pipe sizes for various gas flow rates on the web site: http://www.engineeringtoolbox.com/natural-gas-pipe-sizing-d_826.html and as 1 cubic foot per hour is about 0.15 litres per minute, if you were feeding the engine 150 lpm, then the table entry would be 1000 cu. ft./hr. and you would be looking at the nominal pipe size of 1.25 inch inner diameter. My guess would be a 1.5 inch pipe as the less restriction to gas flow the better. If you want to get a feel for the capacity of any particular diameter of pipe, then try blowing through it. The results are surprising and a small pipe around a quarter of an inch in diameter (6 mm) is incredibly difficult to blow through in spite of looking like an easy gas-flow path.

The amount of gas needed for your engine is reduced massively if you add cold water mist to the intake air. Nobody can answer the question for you because engines vary so much. Even supposedly identical engines have very different gas requirements, so you really need to experiment with your engine to discover what it actually needs. There is some cold water mist information in Chapter 10.

Q: Thank you for that chart, it's just what I was looking for. Now that I know what size of pipe can carry any particular gas-flow rate, I guess the real question is "how many lpm would it take to run a 318 cubic inch Dodge motor?" I read somewhere that it takes 80 litres per minute per 100 cubic inches of engine capacity. I may have got that backwards. Any thoughts?

A: Unfortunately, there isn't any fixed answer to your question and you definitely can't use a rule of thumb to get a reliable answer. Using cold water fog from a pond fogger reduces the amount of hydroxy gas needed by a large percentage, so try reading Chapter 10 which explains this. You need to understand the operating principles if you are to be successful. Also, you have a much easier task if you start with an engine which is much smaller in size, so I suggest that you begin with an electric generator of about 6 kilowatts as described in Chapter 10 before trying to run a very large engine of 318 cubic inch capacity (5.21 litres), which is about twenty times more difficult.

Question:
I am trying to make different devices in order to improve gas mileage on my car. However, I have some difficulties in understanding the automotive circuits. Could you, please, help me?...

I found in Tad Johnson's experience this: "What I found frustrating is that the cell temperature would change and the system would stop making gas. In order to keep the system making gas you constantly have to keep the cell in resonance, and thus you really need the system to be controlled by a processor, that constantly checks frequency on both legs and then adjust inductance to keep the cell in resonance. This is why Stanley moved to the other patents where the spark plug type of electrolysis chamber was used instead of a large cell."

In this case you gave me a clue in this paragraph: “A John Bedini pulser circuit can be used very effectively with a cell of this type and it adjust automatically to the resonant frequency as the cell is part of the frequency-determining circuit.”
First, I didn’t find in your book a Bedini pulser circuit so please give me direction where I can find it.
Second, in Stan Meyer's large cell, I understand, we have two kinds of resonance. One resonance is in an LC circuit where the "C" is the cell itself and the coil is the "L". This is electronic resonance. Another resonance is mechanical resonance where the tube pairs vibrate at the same frequency as the electronic circuit (or a harmonic of that frequency). In this case if I make a device to check the current of the cell, this device can adjust the frequency to be identical with the frequency of the tubes, because the current is in direct relation with resonance. If I am wrong, please let me know.

In your book there is a quotation which says: “I plan to market a pre-programmed chip capable of making any hydroxy system work. While I plan to have a profitable venture with the professional pre-programmed architecture, I also believe in the open source do-it-yourself community, which is where I got started. The chips I plan to sell will be a ‘plug and play’ device. You send me the info on the type of vehicle you are modifying, and the efficiency data of your cell, and I’ll be able to send you a chip that will make your ECU work with those conditions. The Do-it-yourself version would be quite time consuming but, would work for less than one quarter of the price.” I may be interested in buying the chip if you let me know the price before I order it.

Answer:
Tad Johnson used 1,200 volts and was able to get 3 lpm of hydroxy gas at just 1 milliamp of current (1.2 watts) but he found that he could not scale it up for greater gas output.

Chapter 10 has a circuit which finds and holds a cell on its resonant frequency and that circuit will be available ready-made from Courtierstown Marine in the near future. The Stan Meyer "Water Fuel Cell" has been replicated by Dave Lawton of Wales and the full construction details are in Chapter 10. Further advances have been made by Dr Scott Cramton who is getting 6 lpm on just 12 watts of input energy (plus the waveform from the generator which he is boosting - probably 36 watts overall - 12 volts at 3 amps) and who can reduce the diesel requirement of any diesel engine or generator by 60%. Chapter 10 also shows how to run an electrical generator on water alone.

The ECU piggyback chip details are found in the http://www.free-energy-info.tuks.nl/D17.pdf document and it is Les Pearson of the USA (lespearson@hotmail.com) and not myself who is dealing with piggyback chips. I am located in the UK and European vehicles have a completely different computer control system to the ECUs of vehicles in the USA.

Question:
My project is to deliver HHO to a Mazda turbo rotary with 13 lbs boost. I'm not comfortable about injecting HHO into the low-pressure side of the turbo. Even thought it's mixed with incoming air, the turbo radial scroll is a wicked device with lots of sharp blades and turbulent flow. The discharge air at full power can reach over 300 degrees Fahrenheit and easily exceed the 13 lbs boost to compensate for losses. This pressurised air charge is sent to a double-pass intercooler before going on to the manifolds. A violent, hot, and long path. If the HHO survives the journey it has definitely reverted to diatomic.

I need to know what to expect if the HHO can leave the bubbler/dryer and move directly to a diaphragm compressor (premixed and diluted with air, if necessary) for direct port injection at 20 psi. What data are available to show auto-ignition at 15 psi? What temperatures, concentrations, method of compression, etc. If I can't up the delivery pressure for direct port injection, I might consider stopping development of my prototype. Please, there must be a safe way to do this???

Answer:
This subject has come up several times on the different forums and the experienced people recommend putting the hydroxy in on the low-pressure side of a turbo-charger. The hydroxy gets heavily diluted by being mixed with the incoming air and so its characteristics change before it reaches the potential trouble spots which you mention. I am not an automotive expert and so can't make any useful comments from my own experience.

One or two people have tried putting it in on the high-pressure side and have found it ok, but it is distinctly possible that their hydroxy gas was low-grade and already laced with hot water vapour, lowering it's energy level and potential hazard, so my money is definitely on the low-pressure side.

The self-ignition at 12 to 15 psi information comes from Bob Boyce who is easily the most experienced hydroxy user and investigator on all of the forums. That figure relates to Bob's hydroxy which is the highest possible quality and most people never manage to get anything as high-energy as Bob's gas output. The information

comes from repeated experiments by Bob who just lets the pressure build up and then notes the pressure at which it explodes inside his container which is strong enough to withstand the explosion.

Q: Thank you for your quick reply. I am due to meet with Bob Boyce next week. I hope he will have the time (and patience) to deal with my newbie questions but I intend to drill down on this turbocharger thing. For my project, it's mandatory that I run with a turbo, and mandatory that I can feed the maximum amount of high quality HHO. This engine will be developing 400 hp with 13 lbs boost. I think I'll be defeating my purpose if trying to push maximum amounts of HHO into the low pressure side of the intake. First, it will partially displace the necessary air for combustion. Second, it has a torturous route to travel. Third, it will no longer be monatomic HHO when it finally reaches the engine. There must be a better way.

That's why I was hoping to experiment with a medical diaphragm compressor (same as they use to store hydrogen). If I can blend some outside air with the HHO prior to compression, maybe Bob can tell me what ratio of air/HHO to offset/raise the self-ignition. I'm only looking for 20 psi so maybe the dilution won't be a big deal. Anyway, I'm trying to maintain the purity of the HHO prior to engine combustion.

A: Bob Boyce, who is a very highly experienced hydroxy researcher, states categorically, that you should use the low-pressure side of the turbocharger. I am reasonably certain that there is no need for you to pre-mix the hydroxy and air as that happens adequately during the normal intake process, but Bob can certainly give you much more specific detail on this than I ever could. The pressure for self-ignition is raised enormously when the hydroxy is mixed with air, so it should not be a problem in your case.

The objective is to extract a higher proportion of the energy contained in the normal hydrocarbon fuel, and not to add in energy produced by the combustion of hydroxy gas. That being the case, the amount of hydroxy does not have to be enormous for there to be a very significant gain in engine power and thrust quality. Again, Bob is the man to give you the specifics. Good luck with your project, and I hope your visit to Bob works out well.

Question:
I'm spending my day watching YouTube videos of Bob Boyce. It seems every minute I'm learning more and more. Big problem for me is (-40) cold weather and freezing of the water when not in use. (Canadian winters are brutal). The bubblers can be filled with KOH, no problem there. Same with the main cell if I use KOH. But the biggest problem is the water reservoir and supply lines. I had thought of 50/50 blend of methyl hydrate/water (windshield washer fluid) but, according to one of Bob's videos, the carbon will bond with the nickel in the stainless steel plates and eventually stop gas production (sigh). If I had a consistent and reliable electrical power source, I could just plug in a heater of some fashion, but I won't always have that luxury. Any thought on the subject??

Answer:
Please do NOT put KOH in your bubblers as a major function of the bubblers is to wash all traces of KOH out of the gas before it reaches your engine as traces of KOH are not good for your engine. You can use alcohol or paraffin ("kerosene" in the US) which many Canadians have found satisfactory in bubblers as the fumes are not harmful to your engine in any way. If you are boosting, then the water usage will be so low that you can manually top up with warm water from time to time.

Question:
I have watched the interview of Bob Boyce with Chris Patton countless times, even now, looking for signs of 'bullshit'. I have a good intuition which I trust and listening to Bob I can't help but like the guy and still believe every word he says. I'm a plumber and have no previous experience with hydrogen. I've built a 101 plate electrolyzer following your plans and I'm getting 4 lpm when using 190 volts DC and about 4 amps of current.

I've wound my toroid and just run out of money for a scope to try to get the resonance. I've noticed that no one seems to have ever replicated what is described on your website and that Bob is copping it on the forums. I believe him still, but I was wondering in your opinion why is it that no one has managed to get 50 lpm or anything mind blowing? Where could the problem be? I'll tell you one thing though, I'm gonna keep at it as soon as I can afford a scope.

Answer:
Resonance is very important in any COP>1 system, which is what Bob Boyce's hydroxy system is. I'm not sure about your need for an oscilloscope, but don't let me discourage you from getting one. If you have not seen them, the series of YouTube video clips on resonance http://www.youtube.com/user/gotoluc#p/u/15/L0AZkovLTT8 should be both interesting and helpful.

Dave Lawton has designed and tested a circuit which finds and locks on to the resonant frequency of his style of cell. That circuit is available as a pcb, a kit, or a ready-made unit from Courtierstown Marine in Aberdeen.

However, Bob Boyce's design calls for three drive channels instead of just one. Bob has designed, built and tested what he calls his "Hex Controller" which is probably a PIC chip programmed device. He has kept it secret while he applies for a patent on it, but it should become commercially available shortly. I understand that like Dave's design, it locates and locks on to the resonant frequency of the cell.

When not tuned to resonance, the pulsing system should give about three times the gas output of the DC cell performance. When tuned to resonance, the increase should be five or six times.

As you are probably aware, the electrolyser needs to have the plates cleansed and conditioned as described in the D9.pdf document on Bob's unit. Then the cell needs to be "run-in" on DC until it gives around twice the Faraday output. After that, the unit is filled with 28% by weight KOH electrolyte and fed the pulsing signal.

Faraday defined electrolysis as being 2.34 watts per litre per hour of gas output. That is 2.34 x 60 = 140.4 watts for 1 lpm of hydroxy. At this time, you are feeding in 4 amps at 190 volts which is 760 watts. For that 760 watts Faraday would expect 5.41 lpm, which means that your unit is running at 74% Faraday, or about one third of the expected output from Bob's design before it is ready to go to pulsing operation.

Admittedly, the 190 volts is fractionally low for 100 cells, but it is close enough not to be any problem. So, at this point in time, your cell needs mechanical attention in order to get up to the 11 lpm or so that can be achieved on DC alone. When it reaches that level of performance on a 10% NaOH solution, it should then achieve 33 lpm or more with untuned pulsing. But the first step is to get the cell to it's necessary DC performance level, which, unfortunately, probably will need the plates to be prepared to a higher degree.

Unfortunately, there is a difficulty with Bob's design and that is that the plates need to be positioned and sealed to a degree well beyond the capabilities of the average person to construct. If you bought your case from Ed Holdgate, then there will be no problem, but a home-made case can't generally be constructed to the necessary 1/3000 inch accuracy by the average person, as that is a skilled job for an expert machinist who needs to know the exact, actual electrode plate thickness before he starts making your box.

The need for that extreme accuracy is because the resonant operation sets up a magnetic waveform inside the plate array and that won't happen unless the plates are positioned VERY accurately. The gas bubbles form in the electrolyte between the plates rather than on the plates and when that happens it looks like the electrolyte is boiling although it is actually quite cool. I'm sorry that my reply has to sound so negative especially since you will already have put in a large amount of effort on your project. Good luck with your upgrades

Question:
You must be busy, but I'd appreciate any feedback about this link:
http://www.pureenergysystems.com/news/2005/03/17/6900069_Acetone/ Would you say engines in good condition get improvement, or the opposite? In you opinion, could I damage a diesel tractor engine, or car gasoline engine by trying small amounts of acetone added?

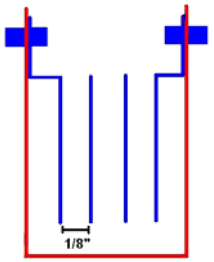
Answer:
There was a lot of discussion about acetone some years ago. Some people got around 20% mpg improvement while others got no improvement at all. It appears to depend on the engine and the condition that the engine is in when the test is run. You need to understand that I am not an automotive expert and so my opinion does not count for much. I would expect it to be the older and poor-condition engines which would benefit most from acetone - which, incidentally, is a fairly dangerous substance for humans and needs to be handled with considerable care as it can be absorbed through the skin and by inhalation and it can cause kidney damage.

I don't think that you will damage any engine by adding some to the fuel. Having said that, you can get up to 60% fuel reduction on a diesel engine by adding hydroxy gas in sufficient quantity and petrol engines respond even better than diesels. The GEET device allows a major reduction in fuel use and is very popular in France where many tractors use it. The http://www.panacea-bocaf.org web site has a good deal of information on it and are experimenting with it at this time. There are also two GEET enthusiast forums.

Question:
I have a Ford "Transit" diesel 70 hp manufactured in 1995. Is it possible to add hydrogen in order to let it run with less fuel consumption and be more environmentally friendly?

Answer:
The fuel consumption of a diesel engine can be reduced by anything up to 60% by adding hydroxy gas to the incoming air. As far as I am aware, no modification to the engine is needed. The improvement in fuel consumption is proportional to the amount of hydroxy gas added. I hope that this helps and good luck with your project.

Question:
Hello, we need your help. Based on your 6-cell model of Chapter 10, we have built such a generator, which is as follows:

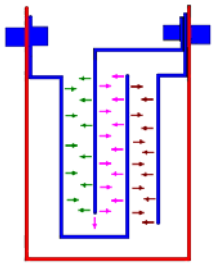


- Stainless steel type 304 (do not get the 316), calibre 0.08.

- Badges 4 "1 / 4 X 2", 4 plates per cell, total 24 plates (206.4 in^2).

NaOH was used as electrolyte, with a concentration of 20% (as described in chapter 10), we just apply 13.5v a current consumption of 1.5A. We use a higher concentration of electrolyte at 30%, 40% and gave us the same result with a current consumption always similar to 1.5A. We need to know because consumption is so low that, according to calculations based on Chapter 10, current consumption with 12V, it should be up to 40A. Or have we miscalculated? See attached image.

Answer:
You have invented your own plate arrangement which needs 6 volts per cell because you have 3 (very inefficient) voltage gaps inside each cell. At this time, you are applying only 0.7 volts between each pair of plates and that is why the current does not change much with stronger electrolyte. You need to change the plate arrangement to what is shown in this drawing. Good luck with your project.



Question:
In the WaterFuel1978 Yahoo forum, the contributor "s1r9a9m9" says that he runs his large capacity car which has 8-cylinders, using just water and an inverter. Do you know of anyone who has replicated this?

Answer:
I am reasonably sure that Nathren Armour is a professional con artist, paid to mislead people investigating how to run an engine on water. "s1r9a9m9" (Nathren Armour of the USA) came on the scene about four years ago. I didn't believe him at all, so I asked him some questions by direct e-mail. I was very surprised when his answers were quite convincing. This was in the old "egaspower" forum. So I wrote a document drawing together all that he had said about the design and adding in each relevant post by people in the egaspower forum who tried to understand and build his design.

After two years of intensive experimentation by several competent people, there was not the slightest success by anyone. Because of this, I withdrew the document "D13.pdf" (although there are probably copies still floating around). I no longer believe that anything Nathren says now, or said back then, is true. In the early days, he posed as an uneducated country car mechanic, who for some unknown reason, was asked by NASA to make some component for a mission to Mars. Why NASA should rely on a country car mechanic is not at all clear. At that time, he told me in a direct e-mail, that he had converted his own personal car, and the engine out of a truck which he used to power an electrical generator, and the cars (plural) of his friends.

His car has 8 cylinders, the truck at least 4 (though probably more), at least two friends with four cylinder cars, makes a minimum of 20 relays needed, plus one which he took apart to see how it was made, so say 21 although 27 would be a more likely number. When writing at that time, his posts were written in broken, uneducated English with a space before the full-stop at the end of each of his sentences.

Now, his posts are in much better English, the gap before the full-stop has gone and the relay count which he states, contradicts what he told me by a really major factor. He no longer allows for the relays used for his friend's cars, or the one which he supposedly took apart. Any time he was asked for specific information, he always made unconvincing excuses and supplied no worthwhile data.

He has posted videos of the Steorn device, showing bench test equipment of a type which no country car mechanic would ever own or understand. His videos have sound but he does not say a single word, which is

most unusual for an English speaker, and in this case is highly suspicious.

My opinion is that his information is wholly fake. It is distinctly possible that this con is now being carried on by a second fake "Nathren Armour". I have never heard of anyone managing to replicate any part of his supposed invention, and I doubt that anyone will ever manage it, although I would be delighted to be proved wrong.

It is perfectly possible to run an engine of any size on what appears to be water alone (the engine actually runs on stored energy drawn directly from the local environment, coming mainly from our Sun). I know of several "petrol-powered" electrical generators which have been converted to run on water in spite of the fact that the efficiency of their tiny internal combustion engines can be as low as 10%. Some of those generators have been running continuously for more than two years now.

Running a large engine on water alone is a much more difficult task but it can be achieved by the energising and injection systems of Stanley Meyer. Sorting out a system of that type has been quite difficult as Stan developed many different types and styles of adaption and it was not immediately clear to me, which of his descriptions apply to which of his systems. Additional information from the contributor whose ID is "H2Opower", has been most helpful, and Stan's design is now much more understandable, which means that replications are now a distinct possibility.

Question:
My daughter came to visit from California. She got 27 miles per (US) gallon average on the trip up here with her Buick Regal. For her return trip with a full car load of her stuff I installed an improved (single water bath) electrolyser, current control with a BASIC Pulse-Width Modulator, and I made her a MAP sensor enhancer override. She followed the same route back to California and averaged just under 40 miles per gallon (a 45% increase). No other changes were made. The electrolyser has eight 2.5" x 5" stainless steel plates with 1/8" spacing in a Mason jar. The plates were all cross-hatched with sanding score marks and all skin oils were removed with trichloretheleyne (Automotive Brake Parts Cleaner). I used the method Boyce recommended of polarizing the plates with a wire wrapped around the body of the plates and sparking across a 12 volt DC source. I set the Pulse-Width Modulator to supply 18 amps steady once warmed up. Previous to this project the best I had achieved is a 23% increase. Your book helped me to tweak a few things and waah-laaaah 39+ mpg.

I do have one problem which I've been trying to figure out. I am building a Don Smith type device. The unit he showed is rated at 8,000 volts at 20 amps (160 KW). I used a 10 foot section of #12 solid copper wire for the L2 coil wound around a 2" pvc pipe. I then released the hold on the wire allowing it to expand to approx 3"diameter coils. Then I cut four ABS plastic strips and hot glued them into the interior of the coil spaced evenly. The coils are spaced at 3/8" space per turn. I then used insulated #12 multi-stranded (maybe 50 strands) wire 2.5 feet long for the L1 coil wrapped around a 1.5" PVC pipe which made about 5 and 1/8 turns. the excess wire was run through holes into the centre of the pipe and carried through to near the lower end to exit the pipe and was then led to the terminal strip for the neon generator.

The neon generator throws a 2" spark until it is hooked to the coil. Once the L1 coil is hooked up I can't get the slightest spark! The spark gap will not arc even a 64th of an inch. I've tried rectifying the output of the neon transformer and tried direct hook-up. It seems odd to expect that a spark could be drawn with the L1 coil direct shorted across the transformer output. My understanding is that the spark gap is required to maintain the frequency as a trigger when DC is used, but that it is not absolutely necessary with AC. Is this your understanding? I know from Tesla's drawings that he commonly used them, but often it was after a capacitor.

When I was looking at the photo clip of the device I didn't see any diodes. Could they be in the black insulating tubing on the terminal strip? Do you know why he converts to DC before he goes into the coil, when he again rectifies with a bridge after the L1 / L2 coil assembly? Could he just as well be using AC until he completes the zero-point energy collection process like Tesla often did or are you positive he has diodes in there at the neon transformer?

One last thing... Do you know how I can read the frequency of the neon's High Voltage High Frequency output with an oscilloscope? I tried to contact the sellers of the devices and they have no idea what the frequency of the output is. They just read what the data plate says on input frequency. I am now trying to get the info from the manufacturer, so far unsuccessfully.

Answer:
May I suggest that you hold off on your assessment of the improvement in mpg got by your daughter. Some ECUs are programmed to adapt to changes such as the addition of hydrogen, and after a few days or weeks, revert to pumping in excess gasoline. Let's see what the effect is after some time has passed. Certainly, the improvement achieved so far is a good indication of what is definitely possible in spite of the opposition from the ECU.

I'm afraid that the suggestion on wrapping wire around the plates and pulsing it, came from me rather than Bob Boyce, who I understand thought that it was a ridiculous idea until one electrolyser builder told him that he had got a substantial percentage gas improvement by doing that. In theory, if the steel is perfect, then it would have no effect, but if, as can happen, the steel supplied is not perfect, then it could help, especially in the early stages of electrolysis.

The devices described by Don Smith are probably the most difficult you could try to replicate (with the possible exception of the Joe Cell). Each of Don's devices need to be tuned to resonance which is difficult to do unless you have a lot of patience, equipment and know-how. Your construction method sounds great and your choice of materials, very good - most people ignore what they are told and decide to do something else and then say that "it doesn't work", while in reality, they should be saying "I couldn't get my altered design to work".

In my opinion, the spark will not occur unless you have a capacitor across the L1 coil, forming an L-C combination which has high impedance at the frequency produced by the neon driver circuit. On it's own, I would expect the L1 coil to load the neon supply too heavily to get any serious voltage across it. As an L-C pair, the impedance at one particular frequency will be very much higher (which is how an AM crystal set radio receiver tunes in a particular station). The spark will only occur at resonance as the voltage will be pulled down too low at all other frequencies.

The diodes are hard to see in the video. They are long and black and don't show up well against the dark background. There are actually four of them on the output side and as you correctly remark, Don describes them as a bridge. They are very long because of their high voltage rating, perhaps ten times as long as they are wide.

I think that the neon driver which Don had, gives two separate high voltage outputs and Don combines them on the input side of L1 with two diodes to stop them interacting with each other. Your neon driver circuit is liable to be very different to Don's.

The frequency and shape of the neon driver output pulses can be read with an oscilloscope. You set the horizontal time adjustment so that you have two successive pulses widely spaced on the screen, and then the horizontal timing setting combined with the number of squares on the screen between the start of each pulse shows you how far apart in time the pulses are. If the time between them is 0.033 milliseconds, then the number in one second (which is the frequency) will be the number of 0.033 milliseconds which fit into one second (of 1000 milliseconds), i.e. 1000 / 0.033 = 30,000 pulses per second or 30 kHz.

Topic: Magnetic Motors

Question:
If a Flynn Magnet Motor has an even number of magnets and coils on the Stator, how does the motor self start when power is applied, if at all?

Would the assumption be correct that the position of the two magnets on the rotor would be slightly off the exact point of registration of their opposing coils in the direction of rotation, say by 1 or 2 degrees or less perhaps as power is applied to them? The next obvious question would be, how long does the energising pulse have to be? I realise it has to be long enough for the face of the rotating magnet to pass over the Stator magnet. After how many degrees will the rotating magnet be caught in the attraction flux of the next magnet? This period of the coil being energised must obviously be as short and effective as possible, since this will determine the input power.

You state in the literature that the two opposing coils, 180 degrees apart are driven in series. What's the reason for series instead of parallel? If current consumption is the problem, then an adjustment in coil design could easily make a parallel connection draw the same current as a series connection, no?

I know this is not a reasonable question but one I wish to ask anyway. If a motor with the magnets I propose to use, works and works well, using a single platter with the specifications as they are, what could one expect in the way of output power, and what would the total input power be? IF and that is a big IF, this design could attain, even 6000 rpm, it could generate considerable power. If speeds of 20 000 rpm as claimed by Flynn are possible, then I guess even a small motor like this one, could in fact yield a heck of a lot of power !!

Answer:
Self-starting requires a different number of stator magnets or coils to the number of rotor magnets. If the numbers are the same and the stator has separate magnets rather than being one large ring magnet, then the rotor will stop in a position with exact registration of the magnets. So if the build does not allow the user to give it a starting twist, then an odd number of magnets are necessary.

I would suggest that you use an electronic circuit to control the length of the pulse but if that is not convenient, then I would suggest a twin-rotor optical method of adjusting the pulse length as shown for the timing of the Robert Adams motor in Chapter 2. By adjusting to position of the two timing rotor slots relative to each other it allows you to control the unmasked slot length in the timing disc and so adjust the pulse length passed to the optical sensor.

Connecting coils in series or in parallel is just a choice governed by the windings and the user's choice of current, as you point out. The point being made in the text was that being exactly opposite each other, the pulses occur at exactly the same time. However, as you point out, that arrangement would have starting problems.

The torque provided by a single rotor depends directly on the radius of the rotor from the bearing out to the centre of the magnets. Another factor is the pull provided by the permanent magnets. The final factor (ignoring the bearing friction and wind resistance) is the speed of the rotor as the number of drive pulses experienced by the rotor depends directly on the rotational speed. I could see the power output of such a motor as being substantial, even with just one rotor, and by adding additional rotors on a single shaft, the power level should be almost unlimited and I can see no reason why a standard electrical generator could not be driven by such a magnet motor. Remember that only one coil is powered at any one time.

Question:
Here is a question which I hope you can answer. I really have no idea why the coils in the Flynn Magnet Motor drawing have such a large area inside the windings. Obviously the shape of the coil follows the shape of the magnet, but why is the centre cavity so large? If the magnets are circular would there still be a large centre hole? I have noticed this in coils designed for wind generators as well. Your explanation would be appreciated.

Answer:
When a current is passed through a coil of wire the strongest part of the magnetic field which is generated is inside the coil windings. The gap inside the coil windings is the most electronically shielded area when the coil is activated and that is why the gaps inside the coils cover the majority of the stator area. This would be the same if the stator was non-magnetic and had circular magnets mounted on it. In that instance, the gap inside the coil would be the same size and shape as the stator magnet and might well surround it.

Question:
I am interested in experimenting with the Phi transformer. I have been reading your free-energy e-Book for quite some time now. You mention tests with a COP of 8.5 , namely, 140 watts input and 1.2 kilowatt output. Would you assist by pointing me in the right direction in examining studies/literature on the Phi transformer.

Answer:
I'm afraid that I am just passing on the research information which I have come across, and I have not built and tested everything mentioned in the eBook as that would take a great deal of money and probably several lifetimes.

In the case of the Phi transformer, I only came across it twice. Once on a German website where there was a very nice photograph of an exceptionally well-built prototype and once on JL Naudin's website. Unfortunately, I lost track of the German site and did not take a copy of that nice photograph. I think that the performance figures came from that site. I tried e-mailing JL Naudin about the variation shown by him but he never replied, which I gather is not unusual for him.

I am not aware of anywhere else that carries information on the Phi transformer. Two people e-mailed me about their attempts to replicate it using very rough bent steel strips and spinning the rotor with a power drill. Their initial tests appeared to confirm that there was either very little, or no Lenz Law effect. They soon dropped me out of the loop and went on developing while communicating directly with each other. I don't recall ever hearing their final results but have the impression that they did not do very well. However, my impression of the quality of their construction and testing was that it was not great, and with free-energy devices, a high quality of construction can make a major difference.

Having said that, I would suggest that you pick on some other device with which to experiment and I would like to remind you that the eBook is upgraded about seventy times per year on average, so downloading another copy is probably not a bad idea.

One think you might care to consider is adapting a generator to run on water alone as the excess output is in the kilowatts range and not much construction work is needed. Alternatively, if internal combustion does not appeal, then perhaps the RotoVerter which can give a 90% saving in the running costs of electrical equipment.

Hope this helps, and good luck with whatever project you pick,

Question:
I've got this idea for a motor and want your feedback. Say you have a wheel, and around the edge you place magnets having the same poles pointed outward. Maybe 30 magnets and all have their north poles facing outward on this wheel. One magnet is then brought, south end first, at right angles to the magnetic field of one of the magnets. The idea is that this will push the magnet in one direction and force the wheel to move, and because all of the magnets have the same pole exposed, this process of continually interacting magnetic fields should keep on for a while. Very possibly a crackpot idea. I'm getting 40 magnets shortly, each with a pull force of 11 pounds, so I hope they're strong enough to do something.

Answer:
It is very difficult indeed to get a permanent-magnet-only motor going and the setup which you appear to suggest is one which is commonly tried by almost everyone. It generally doesn't work because while there is a repulsion on one magnet which drives the wheel around, as soon as the next magnet on the rotor approaches, it encounters a backward push of exactly the same size and that causes the rotor to oscillate and then come to a stop with the magnets at their shortest distance apart. Using electric shielding to block the unwanted retarding action, as in the Adams motor or the Flynn motor, makes a major difference and can produce great results and high rotational speeds. Good luck with your experimenting.

Question:
Regarding the Orbo details in Chapter 1: What Sean at Steorn is saying is that the permeability of the ferrite core decreases as the rotor magnet approaches (I have verified that this assertion is true). When the permeability is at its lowest point, the coil is pulsed with a very small current so that complete saturation occurs. The ensuing small magnetic field then allows the magnet to skate past the ferrite core. Once the magnet no longer influences the core, a large change of inductance occurs in the core, and as Sean states, it leads to an "energy gain". At this point, the flyback voltage can be captured, and indeed, it is greater than the energy it took to energise the coil. The higher the RPM, the more energy can be captured. There are numerous factors involved in all of this, all of which are quite critical in achieving maximum performance -- coil positions, core permeability, speed, etc.

Answer:
Thank you for the detail on the Orbo design. I find it quite hard to get excited about it due to the earlier Adams and Flynn motor designs. I am more than happy that the people in Dublin are genuine as that was my strong impression when I met them on the first day of the failed London demos.

We are only scratching the surface here. I have had person-to-person information on the Adams motor, information which I think has never been made public. Robert Adams was not willing to share his more advanced developments and died before he decided to share his findings. I am informed that he had got a 200 mm diameter rotor motor/generator up to megawatt level output, which is pretty spectacular. I wonder if we will ever replicate that performance.

I understand that Chas Campbell in Australia reached COP=10 before moving on to a more advanced design which he says has unlimited energy output. He says:

"**How 2 Fix**

To create a power source by using Gravity you must have a constant source of momentum and it doesn't have to be perpetual motion, My system uses an electric motor to generate centrifugal force this is achieved with flywheels. A balanced flywheel attached to an electric motor will actual reduce the power consumption of the electric motor while generating Torque, and you use that extra power to drive another flywheel and so on. Remember, to produce electricity you only have to spin a alternator at its designed speed once that speed is reached you use the torque (hp) to keep it spinning - sounds too easy doesn't it? I can spin a flywheel which weighs 80 kg, at 1,000 rpm and the power consumption of the 750 watt motor is less than it takes to spin the motor when it is attached to nothing.

As a matter of interest, the flywheel has a diameter of 600 mm. Imagine a steel wheel weighing 80 Kg travelling at 113 kilometres per hour. So the next time your electricity bill arrives, ask yourself "why does it cost so much when it can be produced by Free wind, Free sun, Free tides, Free water and Free Gravity which can be produced on site?" No more power lines or high insurance cover needed on your solar panels or wind towers.

Gravity is available everywhere even to the 2+ billion people who live without electricity. I am not smart enough to write a book or design a computer game so I decided to invent something that would make me rich it's been a lot of fun and very frustrating as everybody keeps saying "it can't be done". I am now 73, live in a unit, have a foreign landlord and have spent all my money on my invention. It's time that I told the world how to build the "Safe, Affordable, Helpful, & Clean" generator.

Sir/Madam if you, your company, or your Government would like to be associated with something that will make the world a better, safer place please contact me. I've had a film made that I'm sure you will find very interesting.

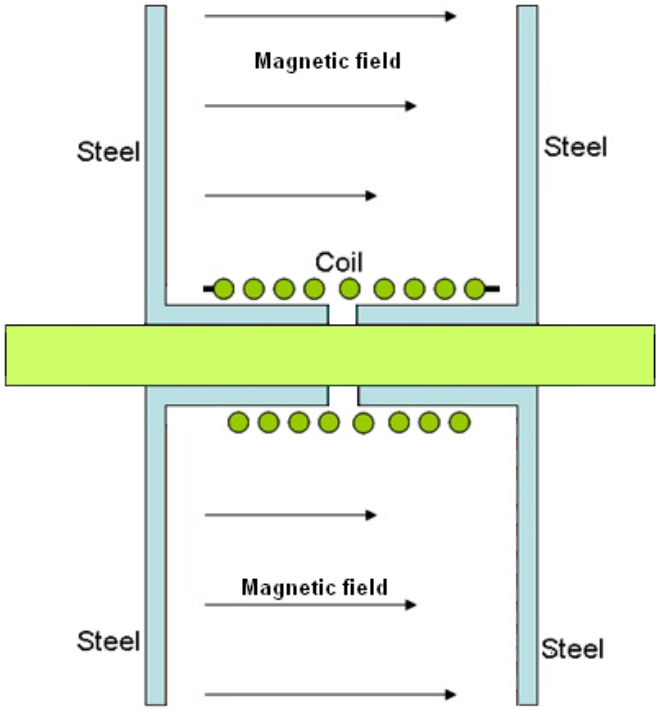
If you send your POSTAL ADDRESS with a small amount to cover my costs to my POSTAL ADDRESS, then a copy of the film in DVD mode will be posted to you.

Sincerely,
Chas Campbell
PO BOX 137, Sunnybank, Queensland, Australia 4109."

It looks like Chas has hit the same brick wall of "zero funds" that most other inventors have. He also seems to have been bitten by the "hey, this design is worth a lot of money, so I could get rich here if I play my cards right" bug. If he manages to make big money, then he will be the first inventor to ever do it.

Question:
How can I block the magnetism between two strong magnets for my new magnet motor design? I am using five layers of mu-metal at the moment but I need a complete magnetic block.

Answer:
Your question suggests that you do not fully understand the nature of magnetism, which is not too surprising as it is not really well taught anywhere at the present time.



If you place two circular steel flanges on a plastic shaft, wind a coil on the plastic shaft and power the coil. the resulting magnetic field is strongest between the outer edges of the flanges and weakest at the coil of wire. Unexpected, but true. The reason is that magnetic "lines of force" (for want of a better term) spread out along the paths of least resistance and the flow divides up in direct proportion to the resistance of the paths. Steel has 1000 times the magnetic conductivity when compared to air, so for every 1 unit flowing through the air, 1000 units will

flow through the steel if it provides an alternative path. As the flange diameter increases, so does the area of steel in the flange and hence the magnetic conductivity. The result is that the strongest magnetic field between the flanges is at the outer edge as can be seen from the attached diagram. There IS additional magnetic flow (not shown on the diagram) through the air outside the flanges, but it is very small compared to the flow directly between the flanges.

The bottom line is that it is not possible to **block** magnetic flow with a shield (and mu-metal is one of the very best materials for attempting that). The most you can do is to provide a more attractive flow path for the "lines" of magnetic force, and so persuade them to divert to another direction - they never just stop at a barrier.

So, for your problem, I would suggest just two ways of dealing with it:

1. Use an identical magnet as the "shield" but reversed so that the field oppose each other. That will work, but if your design requires the shield to be moved into position, then there will be a very marked overhead of the force needed to move the magnet against the existing magnetic fields.

2. You oppose the magnetic flow with a magnetic field from a coil, the field only being present when the coil is being powered. That is very effective and the motor power is very much greater than the mechanical power which could be generated from that same electric current. That arrangement is called a Flynn Motor and it works very well indeed.

One other option which does not match your query, is to physically move one of the two magnets relative to the other as is done in the Steven Kundel magnet motor.

Good luck with your project,

Topic: Don Smith's Devices

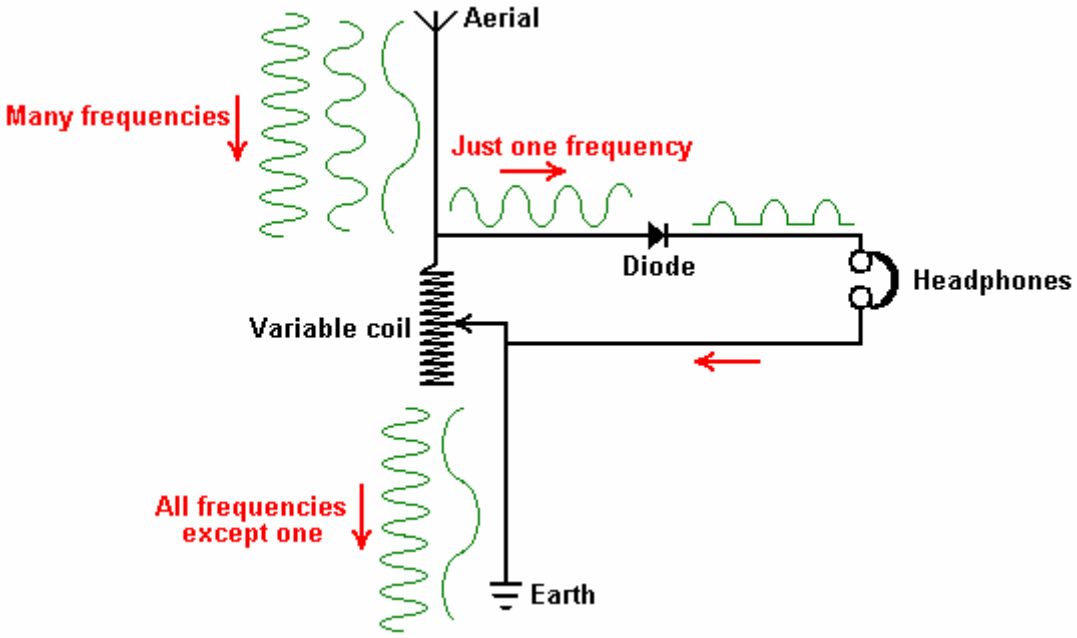
Question:
I am a bit confused though by the frequency conversions necessary to find the proper length of coil for an L1 coil. On page 338 of PJLBook.pdf it says...

Point 1, section d "If using one quarter wavelength, then divide 247 by the frequency in MHz."

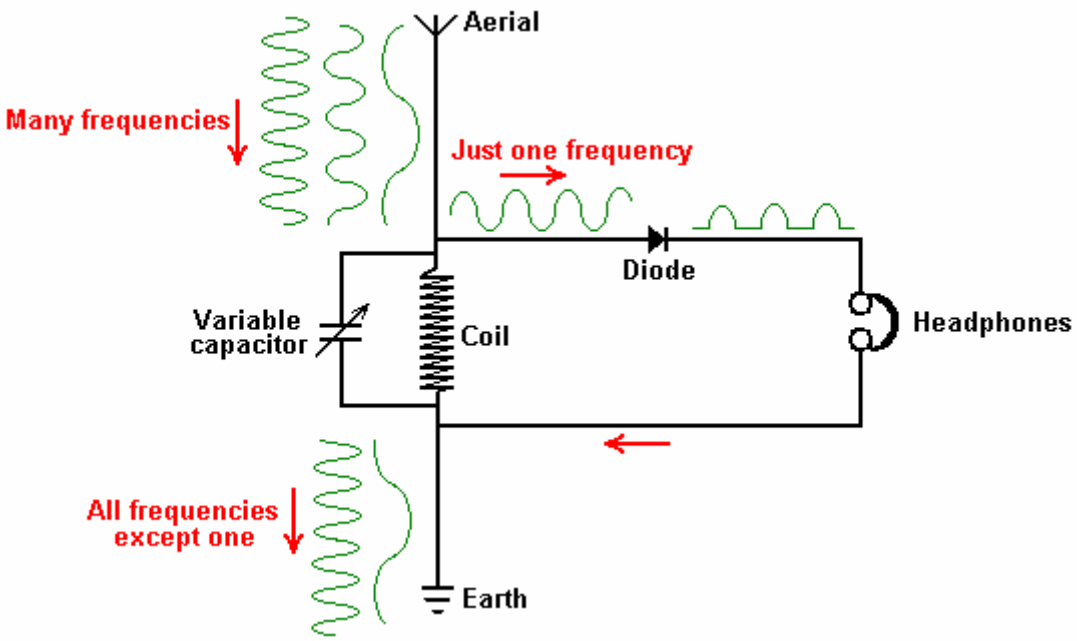
My frequency reads 35.4 KHz or 35400 Hz. That calculates out to .035400 MHz. If I divide 247 by .035400 I get 6,977.40112994 feet. In reading the examples given in Dons pictures and text no reference is given to use of anywhere near that length of wire. In fact the reference to length is in the ten foot range. Am I way off base here or what? Do you come up with the same figures? I have to be missing something, can you help me see it?

Answer:
You are not the first person to ask me that question. It needs to be understood that you are quoting Don Smith and not myself. You should also understand that Don Smith does not reveal everything about any of his designs and I, personally, do not by any means understand all that he says. The Don Smith devices are some of the most difficult to understand and replicate. Don uses various techniques with his different designs although most of those techniques appear to have a common background strategy for energy extraction from what he describes as "the ambient background" and others call "the local environment".

Don goes into the overall principle for estimating the resonant frequency of a coil, where a factor of 247 comes into play. However, a coil and capacitor combination resonate at a different frequency set by their combined characteristics. This is clearly seen in a "crystal set" radio receiver. These very simple receivers can be tuned to a radio station by adjusting the number of turns on a coil. The coil is generally built like a rheostat with a slider connecting to each coil wind in turn.



The tuning relies on the fact that a coil of any particular characteristics such as length, diameter, core material, turn spacing, etc., has one frequency at which it has a very high resistance to AC current flow through it. While we don't tend to think of it as such, a radio signal coming down an aerial wire heading for an earth connection, is actually an AC signal causing an AC current in the aerial and earth connections (and consequently, the coil connected between them). If there happens to be a radio signal which is at that exact frequency, then it finds it very hard to get through the coil and tries to find an easier way past the blockage. The radio builder provides that easier path by connecting a germanium diode and a pair of headphones across the coil. Only that one radio signal takes a detour through the headphones and so the listener only hears one radio station out of the many coming down his aerial.



An alternative construction is where the set is tuned by adjusting the setting of a variable capacitor connected across the coil which has a fixed number of turns. This adjustment causes the coil/capacitor combination resonate at different frequencies, the frequencies of interest are those used by different radio transmitters. The average person thinks of this as "tuning the radio in" to different radio stations, but the reality is that the user is altering the resonant frequency of a coil/capacitor pair. The higher the required frequency, the smaller the capacitor needed.

If a coil has its turns spaced out like the Barker & Williamson coils which Don uses, the coil's inherent self-capacitance is increased dramatically. Dealing with coils energised at high frequencies is a fairly tricky area as there is stray capacitance through the air between components on a board, and so, the physical layout of the

circuit becomes a factor in the design and in some instances, placing your hand near the tuning components can alter the stray capacitance of the tuning system.

Don admits freely that he does not reveal all of the details about any of his designs, but he does share enough for enthusiasts to have a reasonable chance of working out the missing details for themselves. To help with this, he explains a couple of his easiest designs in fair detail. The neon tube driver design is one of these. Here, the frequency of the commercial driver circuit is imposed on the small "L1" primary coil winding. If it is not the natural resonant frequency of the coil, then circuit forces the frequency on the coil by pulsing it at that frequency.

However, one thing which Don does not mention is the fact that the current which will flow through that "L1" coil depends on the applied voltage (which is very high and might be 6,000 volts) and on the coil's impedance at that frequency. The practical details of this radio-frequency work is something about which I am almost 100% ignorant, so you would need to read up on the subject or consult an experienced radio-frequency expert.

In the design which Don shows most frequently, he indicates that the performance of the first part of the circuit is governed by the resonant match between the L1 and L2 windings of the Tesla Coil which he uses to step up the voltage to a high level. To make these two coils resonate in step with each other, the L1 coil needs to be made from wire which is exactly one quarter of the length of the wire in the L2 coil. If the coils were the same diameter, that would give a 1:4 step up in voltage, but as the L1 coil has a larger diameter, the wire length makes fewer turns than one quarter of the number of turns in the L2 coil. Consequently, the step-up ratio is more than 1:4 and more than four time the voltage is generated in the L2 coil.

Don implies that if the coil winding is not absolutely exact, you might need a very small capacitor across the L2 coil in order to make the two coils resonate at the same frequency. If this were the only requirement then you would expect to see one capacitor across the "L2" coil to make the matching perfect. However, you will notice that Don uses a capacitor across **both** of the coils. So, why does he need two? This may be my ignorance showing here, but I would suggest that the capacitor across the "L1" coil has nothing directly to do with the "L2" coil at all but is there to tune the "L1" coil exactly to the output frequency of the neon-tube driver circuit, presenting it with a high impedance at the working frequency and so it will draw very little current from the driver circuit.

The frequency in L2 will always be the same as that in L1 which is driving it, but the energy performance is massively better if the L2 coil is operating at its own natural resonant frequency. This may well be achieved by the wire length ratio between L1 and L2 but as the L1 natural frequency has been manipulated slightly by putting a small capacitor across it, I suggest that the capacitor across the L2 coil is to match the capacitor-induced shift in resonance of L1.

In versions of this design, Don converts the AC coming out of L2 to DC with a high-voltage four-diode rectifier bridge and then feeds the power into a large high-voltage capacitor before stepping the voltage down and the current up, with an "isolation" output transformer.
I have seen it said by several different people that a capacitor used like this causes a change in the nature of the incoming energy, making it much more like conventional electricity.

It seems highly likely that what Dons says has a typing error and the "MHz" should read "kHz" and the wire length at your 35.4 kHz would then be 6.9774 feet (not including the straight connecting wires) or 6 feet, 11 inches, and 23 thirty-seconds or 2127 mm. Wound on a 3" former with wire of 0.25" diameter, that would be about 8.22 turns which does not seem unrealistic.

However, bottom line, you have to consider me to be a source of unreliable information here as you are asking me to interpret what Don means and there are probably not many people who can do that and I am definitely not one of them and so have to resort to guesswork here.

With regards to the Ecklin-Brown style motor-driven rotor device, I would suggest that although he does not show it or mention it, that Don tunes the power pick-up coils to the resonant frequency of the pulsing produced by the rotor arms passing by the magnet pairs. Remember that the title of Don's document is "Resonant" Energy Methods and he shows the Ecklin-Brown style device in that document.

Question:
Just to let you know I was at first impressed with Don Smith's research. Having checked other comments about him I now have my doubts. The question is: why did he not commercialise his Tesla system which is described in sufficient detail in your eBook? One of the sites I found did not have good things to say about him. Who does one believe. I sent him an e-mail recently and it just bounced back. It seems that Don has gone quiet.

Answer:
Don Smith is very elderly at this point in time and he has suffered several severe strokes. It is almost certain that he is in no fit state to respond to e-mails. There is a web site set up by his son who gives every appearance of not understanding anything much about his dad's technology.

Don methods are some of the most difficult to replicate as he freely admits that he withholds some vital information because he exploits his designs financially and so Non Disclosure Agreements get in the way. In his various video lectures he remarks that it is not in his financial interests to disclose all of the details but he says that he discloses enough to allow an experienced developer deduce the missing parts and fill in the gaps for himself.

Probably a key point is having the "L1" / capacitor combination reflect a high impedance at the frequency of the driving circuit (probably a neon tube driver). With a high impedance, there is little current draw. Another feature is the need to feed the energy into a capacitor before attempting to use it for "useful work" as the nature of the energy changes in the capacitor and aligns it to our everyday power supplies. The radio frequency experts have a head start in understanding Don's designs as they are nearly all based on high frequency and, for example, RF experts know that metal can be used to insulate between two wires if the spacers are the correct length for the frequency used.

Although Don does not say it, it is probable that his Ecklin-Brown style device has the output coils tuned to the frequency produced by the spinning rotor arms. Also, it is likely that he is using wire with a very large number of fine strands inside it and that seems to have quite an effect on the output.

Question:
I took off from the internet PJK book. I am interested in the device by Don Smith. Unfortunately, since my English is bad, and his logic is a bit different than my Croatian language, I have problems with understanding. The problem is on page 223, chapter 3-37. So I do not understand why I have to divide some numbers with frequency in order to get the length of the wire. For example, if I want to get the wire length in meters for one-quarter the length of the coil, why do I have to divide the number of 75.29 with the desired frequency. What does that number - 75.29? Also I am interested in one quarter wavelength? What do you mean by the term wavelength? Also, do you think that the term quarter means the length of the wire or it means the length of the coil? I counted the number of turns in the coil of Don and concluded that the coil L1 is a 1 / 4 of 1 / 2 of the coil L2. Can you help me, please, to understand it.

Answer:
Your question is not easy to answer, but I will try to answer it.

If you hit a church bell with a hammer, the bell will vibrate and produce a musical note which depends on the size of the bell. That bell always produces the same note when it is struck.

The bell produces a waveform in the air very much like the electronic waveform produced by most electronic devices.

The note (or "pitch") produced by the bell has a 'wavelength' which is determined by how many times it vibrates in one second and how fast the signal moves away from the bell during that second. The "wavelength" is the distance travelled by the sound during the length of time which it takes for a single vibration.

In electronics, a coil of wire has a resonant frequency, just like the bell has. At that frequency, the coil will vibrate with the minimum amount of input energy possible.

The wavelength of the resonant frequency of the coil is the distance which the electronic signal travels in the length of time needed for just one cycle of it's vibration. Electronic signals generally travel at what is called 'the speed of light' and believed to be about 299,792,458 metres per second. If the coil is vibrating 30,000 times per second (30kHz), then the wavelength will be about 299,792,458 / 30,000 metres or about 9993 metres.

As far as I know, the number quoted by Don Smith is a fixed relationship between the length of wire in a coil and the wavelength of the coil at it's resonant frequency. I am not expert in radio frequency technology, so my explanation of that number may not be fully correct.

A very good resonant ratio between the primary coil ("L1") and the secondary coil ("L2") of a Tesla Coil is 1:4 which means that the length of wire in the secondary coil should be exactly four times the length of wire in the primary coil. That automatically gives the quarter wavelength ratio without you needing to know what the actual wavelength is.

In Don's case, he takes the frequency generated by his neon-tube driver module, and uses the number which you are asking about, to calculate the length of wire in a primary coil which will require the lowest possible amount of current to make it vibrate. That is, the coil's resonant frequency of vibration will match the neon-tube driver module frequency exactly. Other wire lengths will work, but the very best length is the one which he calculates. I hope this helps in spite of the language difficulty.

Question:
in Don's air core coil builders guide he says for L2 wire length at 1/4 wave - divide 247 by frequency. in his example it works out nicely at 10ft. My transformer is from 30 to 40kHz. lets say 35kHz to be in the middle. 35kHz = 0.035MHz (I think). so 247 divided by 0.035= 7057 ft! Now that's a wee tad on the long side me thinks! So I'm guessing that I can keep dividing until I reach a realistic coil length? Here's the question, for L2 to be in harmony with the 35kHz transformer output should I keep dividing the 7057 by even numbers OR by 247, or am I barking up the wrong tree completely?

Answer:
I don't claim to be any kind of expert on this stuff, but this is what I understand:

The heart of a Don Smith design is a resonant transformer. That transformer runs at high frequency and so it needs to be air-core. It can be either step-up or step-down. The lower voltage side needs wire of twice the thickness of the high voltage side. It is this transformer which is the ONLY part of the design which must be resonant and the resonance is between the two windings of that transformer, because at resonance (and not at ANY other frequency) the secondary becomes a room-temperature superconductor with no resistance at all.

To get that exact resonant match, either the length of wire in the turns of the thick-wire coil has a length which is exactly one quarter of the length of wire in the turns of the thin-wire coil OR alternatively, either or both of those coils has a capacitor placed across it so that the coil/capacitor combination has exactly the same resonant frequency of the other coil (or coil/capacitor pair if the other coil also has a capacitor wired across it).

If the two sides of that transformer are matched for resonance, then it doesn't matter what frequency is fed to the primary as the secondary will resonate with it. It is like striking a bell with a hammer. The bell has a specific resonant frequency and it will resonate at that frequency no matter how fast or how slowly it is hit with a hammer.

So, if you let your high-frequency power supply fill up a small high-voltage capacitor to some high voltage until it discharges suddenly through a spark gap (or neon or gas-filled discharge tube) and through the primary, that sudden voltage spike will generate a massive output in the superconducting secondary winding, even though the frequency of the sparks is not the resonant frequency of the primary coil.

The main problem is now encountered and that is the fact that the voltage and current in the secondary coil are ninety degrees out of phase. That means that when the voltage is high, the current is very low and when the current is very high, the voltage is very low, and so the output power appears to be very low even though it has the potential of massive output power. As I understand it, one way to overcome this is to wind the second half of the output coil in the opposite direction to the turns in the first half of the coil. That is, if the first half has clockwise turns, then the second half has counter-clockwise turns. The two ends are then joined together either with a high-voltage diode on both ends or without any diode. The output is then between the centre of the two-section coil and the two combined ends. This joining of the ends of the coil, combines the current and voltage outputs and that composite signal is then ready to drive the primary of an air-core step-down transformer to give lower voltage at higher current, ready to be rectified to DC or to have it's frequency altered by pulsing the output at the local mains frequency.

Chapter 3 has an attempt to explain this in fair detail, starting at page 46. Some people are of the opinion that the grounding wire and/or the connection to the load will alter the frequency of the secondary coil. I have no idea if that is so or not and experimentation is needed to check it out.

Good luck with your project

Topic: General Items

Question:
I was doing a web search today and came across your comprehensive book on free-energy devices. Since I've recently been introduced to this concept of the Zero-Point Energy Field, and still remain sceptical of a true Over Unity machine... I would like to ask if you have seen any such models working in person, with your own eyes? Something that runs without external energy input for an extended period of time!

If I were to replicated some of the experiments, is it easy to get in contact with those inventors? Have you been able to contact them and have had fruitful communications in the past? My curiosity is high, and I would highly appreciate your assistance!

Answer:
Thank you for your e-mail. I fully understand your caution in accepting something which you can't see and about which conventional science appears to be doubtful - calling it "dark energy".

I have met with just a very few of the people mentioned in the eBook as the internet allows us to talk freely to anyone around the world while getting together is usually VERY expensive and often almost impossible. I have met with Bob Boyce and Ed Holdgate in the USA and Dave Lawton who is in Wales. I have corresponded with many people around the world, including John Bedini, Tom Bearden, Ravi Ravu, Lawrence Tseung, Scott Cramton, Tom Thayer, Ron Pugh, Bill Williams, Lawrence Rayburn and a host of other people who have achieved COP>1.

Because of the fact that people are so scattered around the world, I have only seen "with my own eyes" two COP>1 electrolysis systems and one battery-charging COP>1 system. After much correspondence, I have not the slightest doubt as to the complete honesty of many of the people claiming to have working devices, many of whom do not want their names published and several who do not want any details of what they have achieved, disclosed in any way.

Many people think that those who claim these things are just seeking publicity while the reality is just the reverse with most wanting to be left in peace and many quite reluctant to share the details of what they have achieved. One of my biggest problems is that most inventors are driven by curiosity and when they succeed, they lose interest and move on to something else about which they are curious, usually taking the successful device apart in order to use the components for something else, and never documenting what they did.

Dave Lawton's successful replication of Stan Meyer's "Water Fuel Cell" had reached his spares box for recycling when he happened to mention it to me when talking on the phone. I then persuaded him to let me document it, and dozens of people have, since then, made successful copies of Dave's cell. Dr Scott Cramton worked on from that document and has reached 6 litres per minute of hydroxy gas for about 36 watts of input power (which Faraday would have believed to be impossible, stating the maximum possible performance would be 842.4 watts to get 6 litres per minute).

There is no way that I could convince you of the reality of the energy field, although I might be able to persuade you on devices which have a greater output than your input (your refrigerator for example which is COP=3).

Question:
I was looking through the "bedini_monopole3" Yahoo forum files for performances achieved, and most people have reported COP results of between 0.85 and 1.2 with a couple of wild results. Is this all you can get from a Simple Schoolgirl battery pulser?

If the recapture of Back EMF is central to COP success, then should we simply set up a coil, pulse it with DC, capture the Back EMF, making sure that the pulsing frequency is low enough to allow the coil to charge up? It seems too easy and too straightforward.

Answer:
The Ron Pugh Bedini pulser shown with photographs in the eBook has run at COP=11 with a bank being charged from a single battery. Few people construct to that quality or tune accurately like they should. Batteries are not a great solution and a generator running on water has massively greater use to a person - you won't get any battery pulser gaining more than 4 kilowatts.

Picking up Back EMF can be relatively straightforward. Remember the man in South Africa who charges up his 10 Amp-Hour battery each night using the battery which is being charged to power the circuit which is doing the recharging. That system uses Bob Boyce's toroid plus three diodes and one small toroidal choke. At the present time, that circuit recharges at just 12 watts (i.e. 1 amp of current) which is not a massive rate of recharging, but the overall performance is spectacular as present-day science says that doing this is "impossible" and yet this man has done it more than 35 times in a row.

I have a problem with batteries. To get the sort of current supply which a household needs, an impossible number of large, expensive batteries are needed. Even if you keep the rate of discharge down to discharging over the recommended twenty hour period, those batteries will need to be replaced on a regular basis. Also, batteries are not 100% efficient and so part of the current which you feed into them will not be returned to you when you need it to power your equipment. What is really needed is some alternative system which can generate the needed

power at the time when you need it. For that reason, while I am not a fan of internal combustion engines, a generator which has been adapted to run off water alone and which provides several kilowatts of excess power, looks like a much more viable solution to off-the-grid power.

Question:
I would welcome your comments re certain passages in Chapter 5. It talks about the load being powered indefinitely. However there are certain references to discharged batteries being restored under a minute and also that the batteries would charge up to nearly 36 volts with no apparent ill effects. After the discharged battery is restored after a minute this would seem to indicate that after one minute all batteries would start being overcharged. Does this matter? Is there any danger of damage, fire or explosion? There was some mention of control circuitry to prevent over-charging. If the circuit behaved itself during trials why would control circuitry be needed? Also, would not this compromise the correct working of this system? At the moment I am only thinking of using rechargeable batteries.

Answer:
Rechargeable batteries should be ok. John Bedini's "cigar-box" demo Tesla Switch used rechargeable and it ran for six months solid until it was smashed by the opposition who then intimidated John. The section in Chapter 5 which you mention is a quotation of what the Electrodyne Corp. people reported as their experiences with their implementation of the Tesla Switch. They were using four ordinary car batteries. I don't think that lead-acid batteries have any problem with over-voltage when they have become conditioned through Tesla Switch use for a month or two. The over-voltage protection was a suggestion from me and I believe that any problems which they had with the higher voltage were most likely to do with the voltage rating of the components in their circuit rather than with the batteries themselves.

Please be aware that the Tesla Switch pulser physical layout on stripboard shown in that chapter was corrected for an error very recently and so you should download a new copy of chapter 5 (or the eBook) if you intend to build from that particular layout.

Question:
I may have missed any discussion in regards to patents, so I would like to know if these are public domain, or what, if any, restrictions are involved in making one for oneself, or selling plans, or products.

Also, do you have a recommended parts warehouse? I am in the midst of searching, and would like to buy capacitors, chokes, and diodes from the one place. Any help you can give along these lines would be appreciated.

Answer:
Patents are documents of Public Record and so may be reproduced by anybody for any purpose. You are able to make anything shown in a patent for your own use. A patent has extremely limited power and is only there as a low-grade tool against competing commercial companies trying to outsell the patent holder. They can be used by rich and powerful individuals to harass and oppose the individual inventor or very small business by wasting a massive amount of time in court - time which should have been spent in production, and wasting vast amounts of money in defending against court cases (which may well be frivolous). Even if you personally, hold a patent on some device, you would need a tremendous amount of money to fight a patent application from someone else, even though the new application is clearly based on your design. Patents are only useful for very rich people.

There are links to electronics outlets on my websites, but you don't say what country you are in. In the USA, mouser.com might be a good source, while in the UK, perhaps ESR, Maplins or RadioSpares might be your choice.

Question:
May I ask what excess energy devices are there which have no moving parts?

Answer:
There are many devices of that type, with Don Smith having about four dozen, Herman Plauson describing aerial systems of his as being "small" if their output is not over 100 kilowatts, Bob Boyce's toroid which self-charges batteries, Steven Mark's toroid, the Colman/Seddon-Gillespie 1 kW battery, and many, many more.

Mind you, the ones with no moving parts are more difficult to get working as they usually need tuning and most people don't have the necessary level of patience (and many don't have the know-how either).

Question:
It appears resonance is the key to success. Although I have a background as an electronic technician, I am a complete novice regarding "free energy". I have been reading everything I can get my hands on. If you have any other informative references please let me know.

Answer:
The problem with the information is that different people have different theories on what happens and how it happens in the free-energy arena. Some people believe that it is electron spin while others believe that it is the ambient background being nudged into imbalance and others believe that the interface is through rotating magnetic fields.

It is distinctly possible that each of these views is correct and that there are several ways of tapping this energy. However, I can't point you at any one definitive text which describes it all in detail and explains how to gather and use it easily and cheaply.

We have been held back to the 1903 level of flight where flying is just not available to almost anyone. For us now, free-energy is just not available to almost anyone. Worse still, the internet is being smothered with legislation to curtail its use. Unelected people can dictate that any site is excluded from people in their country. The countries which have already implemented this are China, Australia and the UK and I think that the US is likely to do the same in the near future. I would expect my website to be in the list of excluded sites.

So, to answer your question, if I knew of any really good material on the subject, I would either have it on my web site or I would point to it. So sorry, nothing additional springs to mind.

Question:
I recently read with interest what you had to say about colloidal silver, but have I think found a problem! You state that in order to produce colloidal silver that it is necessary to use distilled water and that the silver rods are placed on opposite sides of the beaker. Well even at 27 Volts it is impossible to produce Colloidal Silver in distilled water with a electrode spacing of say 6.5 cm. So I ask, is distilled water really used or if it is anything else added?

Answer:
Highly distilled water is indeed used, and if you don't stir it, you can see streamers moving across between the electrodes. I have found the resulting colloidal silver to be very effective. Ravi in India uses the 200 cc container as he treats many people. I use a much smaller "snifter" glass whose internal dimensions are 40 mm tapering to 28 mm over a depth of 45 mm. This is because I generally use a nebuliser (or apply it externally which is really excellent for skin damage) and the dose volume needed is then very small, perhaps 8 cc or so.

I generally prepare the liquid for about 15 minutes and clean the silver oxide off the negative electrode twice in that period as it builds up very fast. The positive electrode goes non-reflective grey. I could only get much smaller diameter silver wire and so I use about eight strands on each side, bent back on itself in a vertical bundle, in order to get greater surface area.

Question:
There are so, so many areas of possible research that the challenge, for me anyway, is just in making the decision on what to try. There are two that I think I'd like to take a stab at:

1. The Tesla Switch.

2. The Stephan Leben circuit (base upon the original of Alexander Meissner).

However, I have a question concerning each of them. For the Tesla Switch, I've read that an electro-mechanical switch is preferred (and maybe necessary for the highest potential for bringing in radiant energy). In searching a switch of this type I could not find much online to help direct me, so can you suggest a possible component here?

For the Leben/Meissner circuit, what would be a likely choice for the transistor? (That is the only component that was not named, of all the parts listed).

I would have preferred to just dig in and try these two on my own, but since my electronics background is practically nil (with the exception of assembling a Xenon strobe light circuit from a kit many years ago) I was hoping you might take some pity on this poor physics-challenged soul and offer a couple of pointers. Any information you may provide in this regard will be most appreciated. Thank you, in advance.

Answer:
Thank you for your e-mail. Chapter 12 is supposed to give you enough electronics background to allow you to understand and build circuits of this kind. If it has not done that for you, then please give me a steer as to what more is needed and I will expand the chapter accordingly. (This is a genuine request on my part and is not meant to have any kind of negative connotations.)

The phrase electro-mechanical switch does not ring an immediate bell with me but the term implies that it is a mechanical switch which is activated electrically. That would include relays and the mechanical motor-driven switch shown as being what the Electrodyne Corp. staff used (it seems likely that Don Kelly was one of those Electrodyne Corp. people). Transistors were used in John Bedini's circuit replication which ran solidly for six months before it was smashed and he was intimidated into silence.

The reason why mechanical switching is sometimes recommended is because many people do not know how to get fast switching with semiconductors. The Tesla Switch circuit needs to have all the components working together and so even the diameter and length of the connecting wires can make a difference. It is a circuit where heavy loads will work better than tiny test loads as cold electricity operates the reverse of hot electricity and so the heavier the load, the better it works. It is one of the most difficult circuits to get working well, so you might be better off picking a different project.

I think that it is unlikely that the transistor in the Leben circuit is at all critical. Technically, it needs to be able to handle the current, and that will depend on the wire size you use and the number of turns (and to a lesser degree, on the frequency which is determined by the size of the capacitor). An MJ11016 is a powerful, high-gain transistor which should work well. A BD243C transistor is popular in circuits of this type although it has a much lower gain. Hope this helps, and good luck with your project(s).

Question:
Always been a fan of yours ever since I built my first "D18" cell which may I add is still going strong and saving me loads of money. I have a question regarding the Tesla Switch: you mention that the contacts on the discs attached to the motor are set at angles. Can you please explain in more detail.

Answer:
You are the first person I have ever come into direct contact with who has ever had any success with the "D18" Nitro Cell and I have had to classify it as a "doubtful" device because of that. If you are getting a benefit from it, then I would be delighted to change it's classification and combine with you to do a better presentation of the cell including your experiences with making and using it - if you have no objection to this, then please get back to me.

I'm sorry that you have found the switching disc details difficult to understand and I will upgrade the description in the next issue of the eBook. What I was trying to say is that each disc has one sixth of it's circumference clad in copper while the following sixth of the circumference is a non-conducting material.

As there are 360 degrees in a complete circle, one sixth of the circle covers an angle of 60 degrees, which is why 60 degrees is mentioned at the present time. I will expand the description to explain this in more detail. There is no question of the actual copper cladding (or whatever method of conducting contact that you choose to use) being itself angled in any way - it is just a case of there being three sections of the disc being conducting and having an equal sized non-conducting section on each side of the conducting sections. as the contact pairs are exactly opposite each other, this causes one pair of contacts to conduct while the pair directly opposite it is not conducting. Hope this helps and my apologies for the poor description in the eBook.

Question:
I was starting to get really interested in this Free Energy stuff, and before I found your website, I had contacted CAT (Centre for Alternative Technology, based in Wales). I presume that you have heard of them?

Anyways, I was hoping to get some information from them regarding the best renewable source of energy (I was thinking of using a small hydro scheme). I thought I'd also ask their advice on Hydroxy production. As I should have expected, they lambasted and patronised my comments so thoroughly I gave up any idea of hydroxy. If an organisation such as CAT finds hydroxy ridiculous, I couldn't believe that there was any future for Hydroxy on

demand.

Do you know of anyone that has managed to heat their home using hydroxy? and if so how? Would be interested to hear any ideas.

Answer:
The 'Chapter14.pdf' document shows ways to use hydroxy for heating - which is one of the most difficult things to use it for.

Your problem with ‘CAT’ is due to your not understanding the overall situation. Please don't for a moment take what I am about to say as being condescending or negative in any way as it most definitely is not intended to be.

There is a foreign State which is not answerable to anyone, based in London but it is not part of England or Britain. They have chosen the deliberately confusing name of "The Crown" (which has nothing to do with any Monarch) or the "Inner City of London". Their objective is to enslave and rob the British population without them ever realising what is happening. They are a consortium of bankers and lawyers and their biggest attack weapon is a society which they set up and own. It is called "The Law Society" and they use it to mount attacks on people who do not understand the situation.

They have set up a large number of commercial companies, each with a name intended to confuse and distract. One of their companies is called "The Ministry of Justice" and the name is intended to give the impression that it has ruling powers, when in fact, it is no different to a company such as Ladbrokes betting shop. Other companies which it owns include every Magistrates Court, every Police Constabulary, every Local Council, "The Bank of England", etc. etc.

They bolster their illusion by having a 'parliament' where their employees appear to argue with each other and people in general are fooled into thinking that those employees are working for them, and every so often, they even get to choose which employees are hired. The reality is that all important decisions are made in accordance with whatever the owners want and those decisions bear no relationship whatsoever to the facts or the benefit of the population as a whole.

Vast amounts of money are stolen on a continuous basis, using names such as "National Debt", "Foreign Aid", "Income Tax", "Council Tax", "Value Added Tax", "Fuel Levy", "TV Licence", etc. etc. etc.

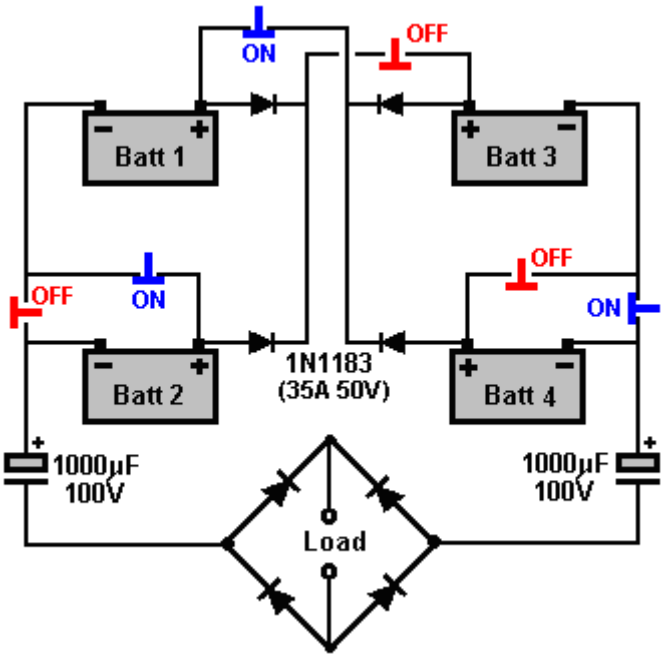
They own agencies such as CAT and under no circumstances would they ever approve the use of any fuel which you do not have to buy from them, so HHO is OUT !! I suspect, although I have no direct evidence for it, that any subsidies offered for "green" projects are arranged to funnel money into the coffers of their friends and colleagues as that would be wholly consistent with everything else that they do. They will only ever support projects which will be ineffective and expensive. Dave Lawton has suffered at their hands when he made a grant application and was then visited by two CIA operatives (and yes, the CIA does operate freely in Britain with the full approval of the government which is owned by the foreign independent State "The Crown").

So, your asking them about HHO was roughly equivalent to asking them if they would mind if you were to stab them with a large knife - not exactly a popular enquiry. They repeatedly undertake false public relations exercises to convince the population that free-energy is 'impossible' and that using an HHO booster on an engine never gives any benefits at all - both of those statements being wholly false.

So, I would suggest that you stay away from CAT and similar companies as they are, in fact, the opposition and most definitely are not there to help you.

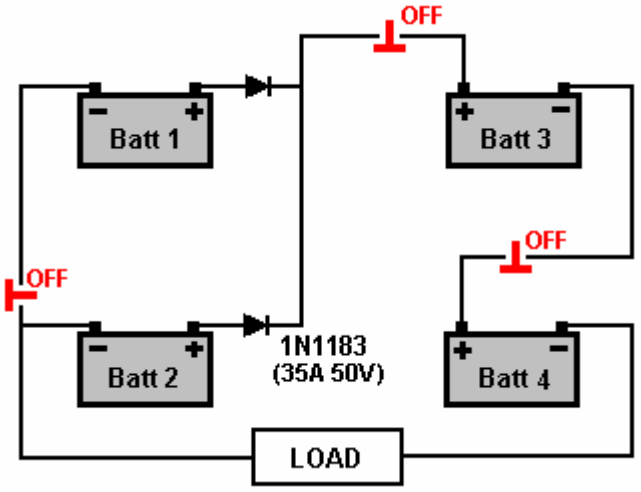
Question:
I have difficulty in understanding how electronic circuitry can be used to drive a Tesla Switch. Would you please show me how the components can be connected together.

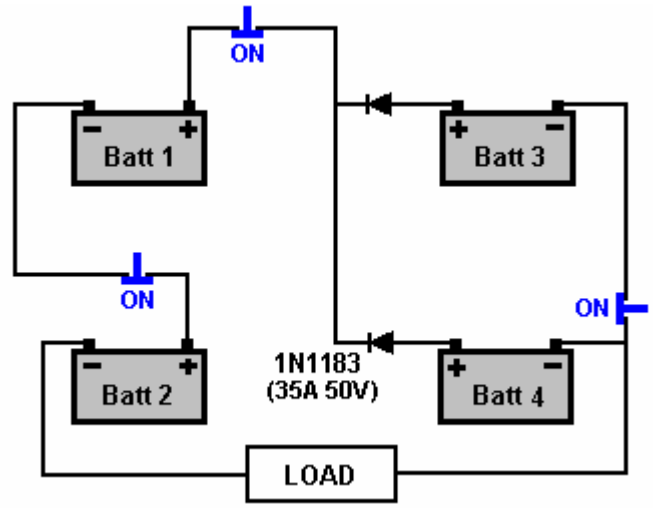
Answer:
The Tesla Switch circuit used by the Electrodyne Corp. staff is shown here:



It is distinctly possible that the circuit does not reveal all the information needed for this system, as it may be necessary that the frequency at which the circuit runs needs to be a resonant harmonic of the batteries, or possibly, the load need to be a motor which generates voltage spikes. I do not know if that is the case, but I do know that the 1N1183 diodes are connected in reverse because they break down suddenly, passing a voltage spike to the batteries. It is distinctly possible that the batteries need to be conditioned for this circuit to perform the way that the Electrodyne Corp. staff reported. It should be remembered that they experimented with this and similar circuits for something like three years and it is highly likely that by the time they recorded their spectacular performances, the batteries had been very well conditioned due to the tests which they ran during those three years.

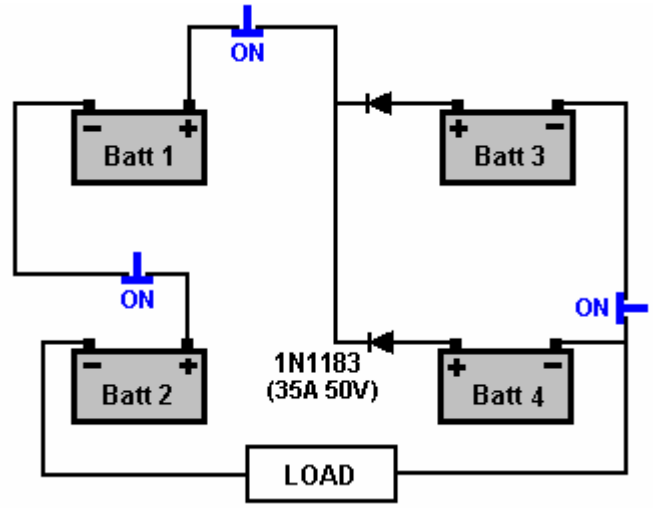
However, if I were to start experimenting with the circuit, I would show the two separate switching arrangements like this:



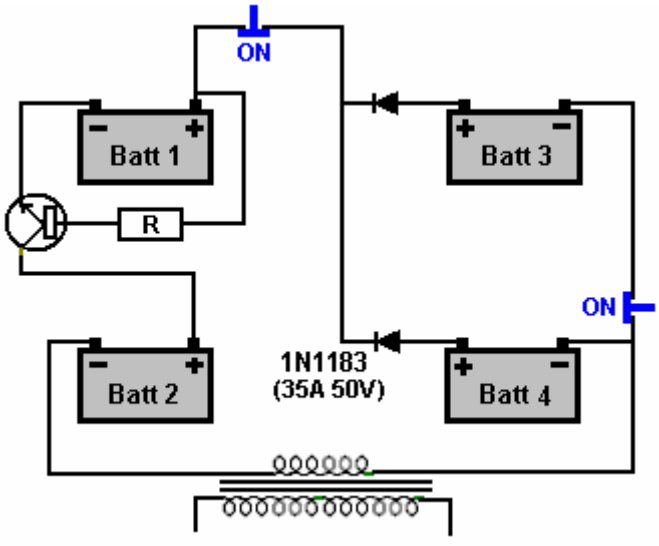


You will notice that each of these circuits are very direct and simple in themselves, but the “LOAD” is an essential part of the circuit and without it being in place (and able to carry current) neither circuit is completed. Personally, I would be inclined to use the 12V winding of a mains transformer as the load connection as that can handle both DC and AC current flow and the mains winding can have a bulb connected to it to show how the circuit is performing.

So, we are faced with two separate sets of three switches where each set needs to be closed when the other set is open. The question is, how can we use transistors to implement those switch operations? The difficulty is that there is no common negative voltage rail as the batteries keep getting swapped over as the circuit operates (which is the whole point of the circuit), and so, ideally, we want to switch each transistor with it connected permanently to just one battery. One circuit is:

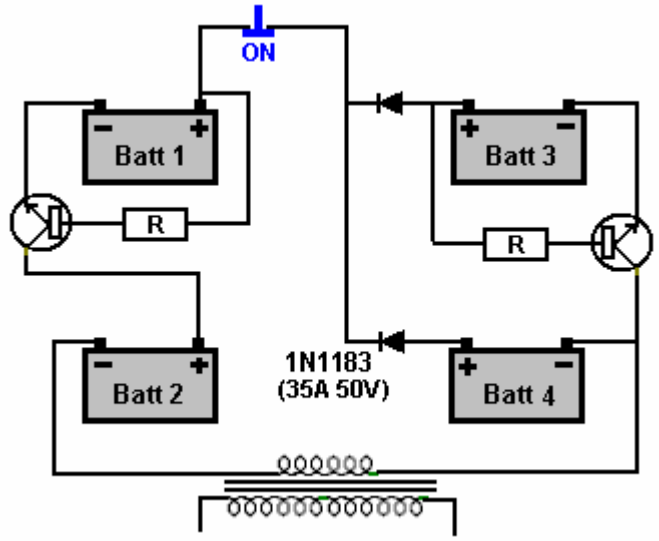


So if we were to use a transistor instead of one of the switches, we might make the circuit this:

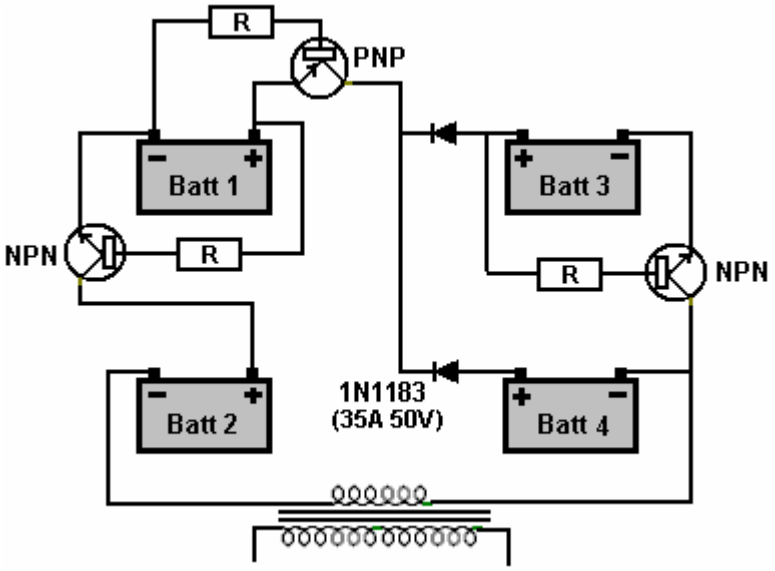


Here, the NPN transistor base is fed a current through resistor R, and as the transistor emitter is connected to the Battery 1 Minus and the resistor is connected to the Battery 1 Plus, there will always be a voltage difference able to feed current to the transistor base, switching it fully On. Obviously, we do not leave the resistor R connected to the Battery 1 Plus all the time, but will introduce some form of solid-state switching to switch the transistor on when we want it to be on and not at other times.

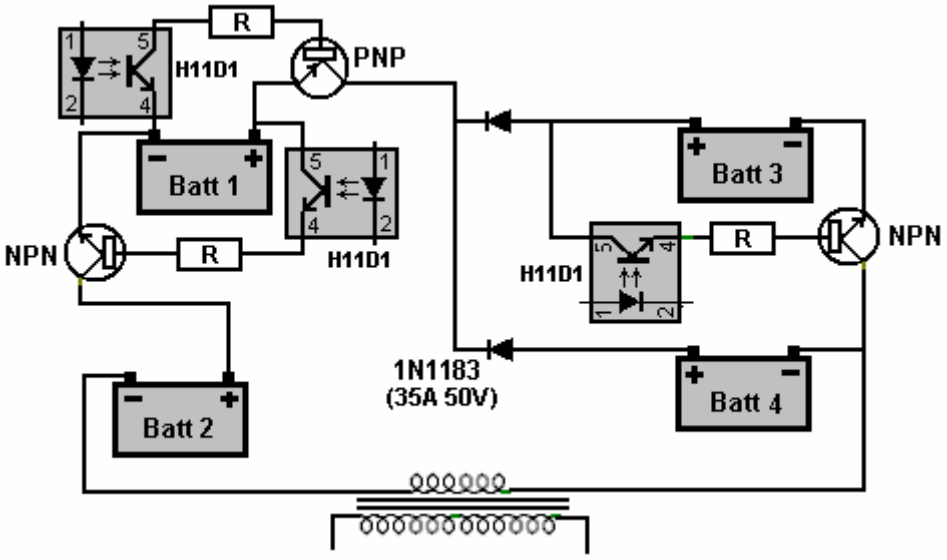
We can swap another of these switches with an NPN transistor like this:



But we now get stuck because the remaining switch does not connect to the negative side of a battery. We can get around this by using a PNP transistor as it is the mirror image of an NPN type. The circuit could then be:

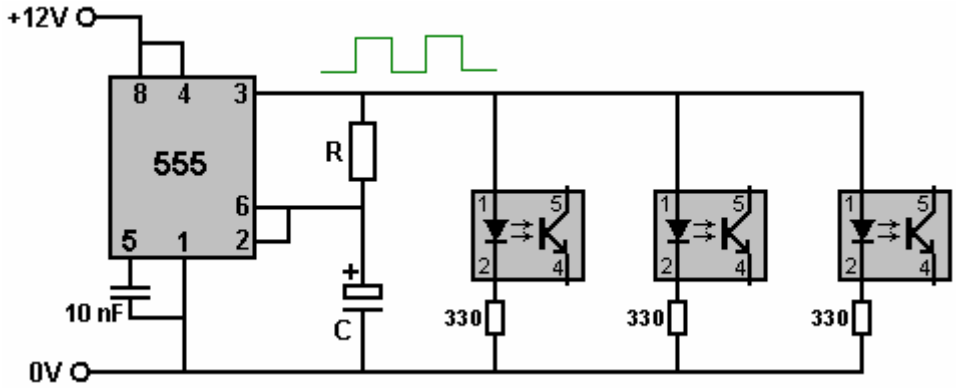


This circuit will have all three transistors switched On as heavily as possible at all times. But while that does appear to substitute transistors for mechanical switches, we want to be able to turn that set of three transistors On and Off when we want. That can be achieved by making the transistor base current flow through an Opto-Isolator which can be controlled by a completely separate circuit which is not connected to this Tesla Switch circuit in any way, perhaps like this:



The opto-isolators shown here happen to be H11D1 types but many others could be used. Many different components could be used. The resistors 'R' could be, perhaps, 4.7 K, the PNP transistors MJ11015 and the NPNs MJ11016 or any high-gain, high-power transistors.

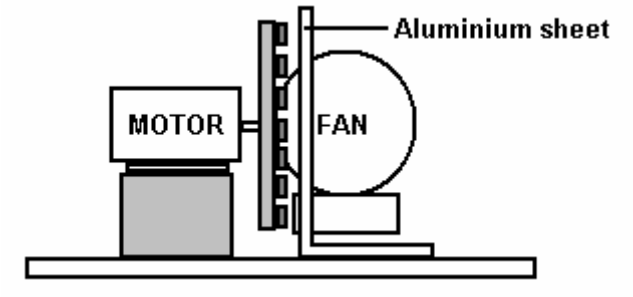
The three opto-isolators can be driven by a 555 timer signal generator which has a 50% Mark/Space ratio, perhaps like this:



The 555 timing resistor 'R' can be replaced with a fixed resistor and a variable resistor in series, so that the frequency of the switching can be adjustable. A rate of 800 Hz might be suited to this circuit. Chapter 12 has a table of the R and C values for different frequencies (perhaps 100nF and a 4.7K plus a 100K variable in this case). The other three opto isolators can be driven via an inverter IC or via a PNP fed by pin 3 of the 555 timer IC. The transistor substitution for switches for the second part of the circuit can be done in the same way.

Question:
In chapter 14 you talk about the Frenette heater, under that article you make the suggestion that one could spin magnets next to an Al plate and make heat. I do not quite understand this, is there any way you could explain a little further, or make some drawing with suggestions?

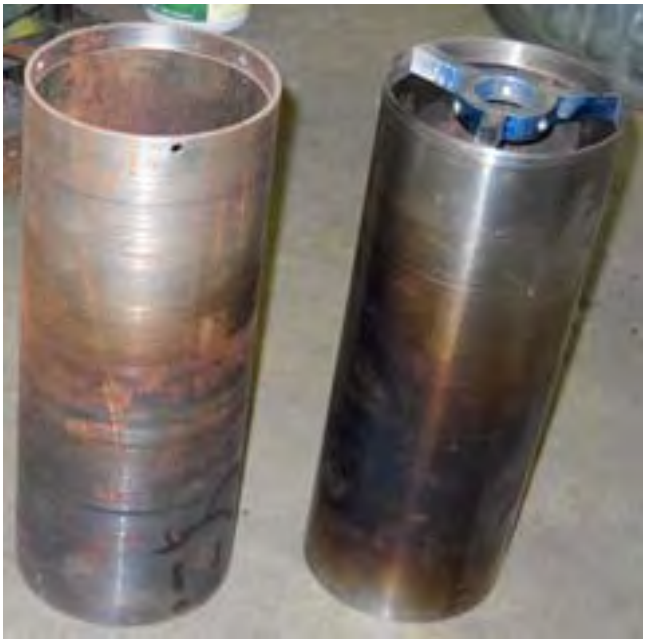
Answer:
A simple magnetic heater is just a motor with a disc mounted on it's shaft. That rotor disc has magnets mounted on it and an aluminium plates is placed beside it as shown in the diagram below. As the magnet disc spins, it causes rapidly changing magnetic currents in the aluminium which causes it to heat up. If you use an ordinary fan to blow air across the aluminium plate, you get a hot-air heater system. Having sideways "fin" plates on the aluminium sheet should improve the heat transfer to the air.



However, I have asked a man who has considerable experience in this area and he says:

Magnetic Heaters:

These are pictures of magnetic heaters which I have built and tested:

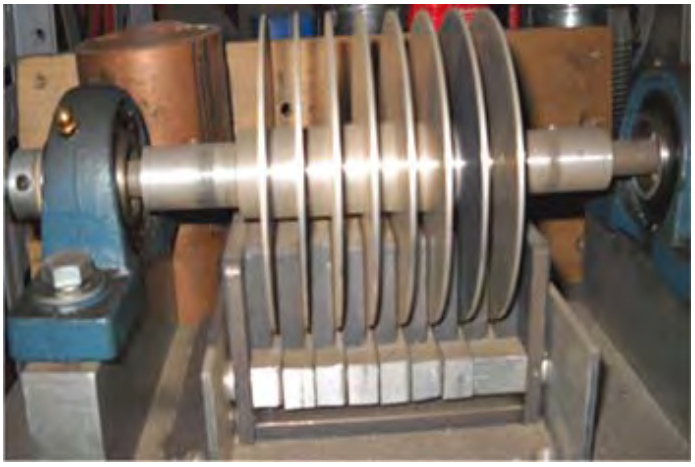


These pictures are of a magnetic heater driven by an induction motor. There are magnets placed between steel pole pieces which are adjustable to allow for the expansion caused by heating. The cylinders are made from A53 type steel. Cast iron is not suitable as the motor used spins at 3,600 rpm and cast iron has a different magnetic angle, which, while it causes a lot of resistance when being turned, slowly, does not heat up at the high speed used my me.

The steel cylinder needs to be electrically isolated from all of the other working parts of the heater and it needs to be allowed to float on the shaft so that it can increase in length as it heats up. The speed of rotation was 3,600 rpm and the cylinder was cut to a wall thickness of 0.108 inch (2.75 mm).

The colour distortion on the cylinders shown in the last photograph was caused by the heating of the cylinder when in use. These cylinders reached about 850^{O}F and with the volume of air blown through the cylinder it produced air exhaust temperatures of 148^{O}F. importantly, this exiting air was ionized which provided benefits in addition to heating. Technically speaking, this ionization is actually an inefficiency.

When performing magnetic heating, the magnetic field needs to go through the metal and not along it's surface as is the case in the heater shown above. Because of this, a different style of heater was built as shown here:



With this design, the magnets should not cover more than one quarter of the disc face and the magnetic field must pass all the way through each disc. Even with air being blown between the discs, the magnets get so hot that it becomes a long-term problem in that magnets lose their magnetism if raised to too high a temperature. The disc thickness is important.

There are some advantages of this type of heating mainly the motor type and power consumption allows for better than normal efficiency but that is not high enough for the construction cost to be recovered and that is why this

type of heater is not popular. However, it has the advantage that the ionization of the heated air can suppress virus and bacteria on surfaces.

It is important to note that a magnet moving over a plate heats only a small zone and the heated metal is only a narrow area within the magnetic field.

The magnetic cylinder heater shown was designed in 1989 by the late Frederick J. Galloway. Four videos have been made specifically for this query, showing the 4-kilowatt input power cylinder heater at:

http://www.youtube.com/watch?v=JN1BXx-zXLA Part 1 of 4

http://www.youtube.com/watch?v=aS3SSeFbyds Part 2 of 4

http://www.youtube.com/watch?v=vhmQP0UStcA Part 3 of 4

http://www.youtube.com/watch?v=cV4Pg61wcp0 Part 4 of 4

Question:
What drawing package do you use to produce the diagrams in your eBook?

Answer:
I use the "Paint" program which comes free with Windows. As that answer upsets people, I have produced three casual videos explaining how easy it is to use. As I don't have a video camera or a phone with video facilities, these videos were produced using the FoxMagic program "Screen Virtuoso".

http://www.youtube.com/watch?v=Q6inBL3f13E&context=C3f535afADOEgsToPDskI07u7HkyPbPJElEQhd5qpV

http://www.youtube.com/watch?v=HUfXT5U-xGg&feature=context&context=C3f535afADOEgsToPDskI07u7HkyPbPJElEQhd5qpV

http://www.youtube.com/watch?v=PBAm7LwTG_4&feature=context&context=C3f535afADOEgsToPDskI07u7HkyPbPJElEQhd5qpV

Patrick Kelly

http://www.free-energy-info.tuks.nl/

Appendix

TABLE OF WIRE SIZES:

The wire sizes specified for use in some designs are American Wire Gauge so a comparison table showing the UK Standard Wire Gauge (with lengths on a 500 gram reel of enamelled copper wire), and the American Wire Gauge is given here:

AWG	Dia mm	Area sq. mm	SWG	Dia mm	Area sq. mm	Max Amps	Ohms / metre	Metres Per 500g	Max Hz
1	7.35	42.40	2	7.01	38.60	119			325
2	6.54	33.60	3	6.40	32.18	94			410
3	5.88	27.15	4	5.89	27.27	75			500
4	5.19	21.20	6	4.88	18.68	60			650
5	4.62	16.80	7	4.47	15.70	47			810
6	4.11	13.30	8	4.06	12.97	37			1,100
7	3.67	10.60	9	3.66	10.51	30			1,300
8	3.26	8.35	10	3.25	8.30	24			1,650
9	2.91	6.62	11	2.95	6.82	19			2,050
10	2.59	5.27	12	2.64	5.48	15	0.0042		2,600
11	2.30	4.15	13	2.34	4.29	12	0.0047		3,200
12	2.05	3.31	14	2.03	3.49	9.3	0.0053	17.5 m	4,150
13	1.83	2.63	15	1.83	2.63	7.4	0.0068		5,300
14	1.63	2.08	16	1.63	2.08	5.9	0.0083	27 m	6,700
15	1.45	1.65	17	1.42	1.59	4.7	0.0135		8,250
16	1.29	1.31	18	1.219	1.17	3.7	0.0148	48 m	11 kHz
17	1.15	1.04				2.9	0.0214		13 kHz
18	1.024	0.823	19	1.016	0.811	2.3	0.027		17 kHz
19	0.912	0.653	20	0.914	0.657	1.8	0.026	85 m	21 kHz
20	0.812	0.519	21	0.813	0.519	1.5	0.036		27 kHz
21	0.723	0.412	22	0.711	0.397	1.2	0.043	140 m	33 kHz
22	0.644	0.325	23	0.610	0.292	0.92	0.056		42 kHz
23	0.573	0.259	24	0.559	0.245	0.729	0.070	225 m	53 kHz
24	0.511	0.205	25	0.508	0.203	0.577	0.087		68 kHz
25	0.455	0.163	26	0.457	0.164	0.457	0.105	340 m	85 kHz
26	0.405	0.128	27	0.417	0.136	0.361	0.130		107 kHz
27	0.361	0.102	28	0.376	0.111	0.288	0.155	500 m	130 kHz
28	0.321	0.0804	30	0.315	0.0779	0.226	0.221	700 m	170 kHz
29	0.286	0.0646	32	0.274	0.0591	0.182	0.292	950 m	210 kHz
30	0.255	0.0503	33	0.254	0.0506	0.142	0.347	1125 m	270 kHz
31	0.226	0.0401	34	0.234	0.0428	0.113	0.402	1300 m	340 kHz
32	0.203	0.0324	36	0.193	0.0293	0.091	0.589	1900 m	430 kHz
33	0.180	0.0255	37	0.173	0.0234	0.072	0.767	2450 m	540 kHz
34	0.160	0.0201	38	0.152	0.0182	0.056	0.945	3000 m	690 kHz
35	0.142	0.0159	39	0.132	0.0137	0.044	1.212	3700 m	870 kHz

Patent US 6,867,514 B2 **15th March 2005** **Inventor: Frank J. Fecera**

PERMANENT MAGNET MOTOR

This patent application shows the details of a permanent magnet motor. It should be noted that while in this text, Frank states that permanent magnets store a finite amount of magnetism, in actual fact, the magnet poles form a dipole which causes a continuous flow of energy drawn from the quantum foam of our universe, and that flow continues until such time as the dipole is destroyed. The energy which powers any permanent magnet motor comes directly from the zero-point energy field and not actually from the magnet itself. A piece of iron can be converted into a magnet by a single nanosecond magnetic pulse. It makes no sense that a pulse of that duration could provide months of continuous power from anything stored in the magnet itself, but it makes perfect sense if that brief pulse created a magnetic dipole which acts as a gateway for the inflow of zero-point energy from the environment.

ABSTRACT

A motor providing unidirectional rotational motive power is provided. The motor has a generally circular stator with a stator axis, an outer surface, and a circumferential line of demarcation at about a midpoint of the outer surface. The motor also includes one or more stator magnets attached to the outer surface of the stator. The stator magnets are arranged in a generally circular arrangement about the stator axis and generate a first magnetic field. An armature is attached to the stator so that it rotates with it, the armature having an axis parallel to the stator axis. One or more rotors, are spaced from the armature and coupled to it by an axle to allow each rotor to rotate around an axis, each rotor rotating in a plane generally aligned with the axis of the armature. Each rotor includes one or more rotor magnets, with each rotor magnet generating a second magnetic field. The second magnetic field generated by each rotor magnet interacts with the first magnetic field, to cause each rotor to rotate about the rotor axis. A linkage assembly drive connects each rotor to the stator to cause the armature to rotate about the armature axis thereby providing the unidirectional rotational motive power of the motor.

BACKGROUND OF THE INVENTION

This invention relates to dynamo electric motor structures and more particularly to rotary and linear permanent magnet motors. Conventional electric motors rely on the interaction of magnetic fields to produce a force which results in either rotary or linear motion. The magnetic fields in conventional electric motors providing rotary power, are generated by passing an externally provided electric current through conductors in either a stator (i.e. stationary portion of the motor), a rotor (i.e. rotary portion) or both the stator and the rotor. The rotary power of the motor arises from a rotating magnetic field which is created by commutating the electric current, either by a switching the current through different conductors, as in a direct current motor or by a polarity reversal of the electric current as in an alternating current motor.

It is well known that a class of materials known as ferromagnetic materials are also capable of generating a magnetic field having once been energised. Ferromagnetic materials with high coercivity are known as permanent magnets. Permanent magnets are capable of storing a finite amount of energy and retaining the ability to generate a substantial magnetic field until the stored energy is depleted.

There are electric motors which use permanent magnets in either the stator portion of the motor or the rotor portion of the motor. These motors achieve a small size for the amount of power delivered by the motor because the motors avoid having current carrying conductors to produce the magnetic field which is otherwise produced by the permanent magnets. However, these conventional permanent magnet motors still require a source of external power to produce a rotating magnetic field.

There have also been developed permanent magnet motors which use permanent magnets for both the stator and the rotor. For example, U.S. Pat. No. 4,598,221 discloses a permanent magnet motor which relies on an external source of power to rotate the magnetic fields of a rotor by ninety degrees with respect to the interacting stator magnetic fields to eliminate the counterproductive magnetic repulsion and attraction between the rotor and the stator magnets. In another example, U.S. Pat. No. 4,882,509 discloses a permanent magnet motor which relies on an external source of power to position a shield which does not permit coupling between the rotor and the stator magnets at times when attraction or repulsion would drag down the strength of the motor.

There are many instances where a motor action is required and no source of external power is available. Accordingly, a motor which relies solely on the energy stored in permanent magnets would be useful.

BRIEF SUMMARY OF THE INVENTION

Briefly stated, the present invention comprises a rotor for use in a permanent magnet motor and for providing motive power by rotation of the rotor about a rotor axis. The rotor comprises at least one first U-shaped magnet having a rear side and generating a first magnetic field. The rotation of the rotor about the rotor axis is caused by an interaction of a portion of the first magnetic field directly adjacent to the rear of the at least one U-shaped magnet with a stationary second magnetic field.

Another aspect of the present invention comprises a rotor providing motive power by a rotation of the rotor about the rotor axis and by a translation of the rotor in a direction of the rotor axis. The rotor comprises: a first U-shaped magnet having a north pole, a south pole and a rear side, the first U-shaped magnet generating a first magnetic field; a second U-shaped magnet having a north pole and a south pole, the south pole of the second U-shaped magnet abutting the north pole of the first U-shaped magnet; and a third U-shaped magnet having a north pole and a south pole, the north pole of the third U-shaped magnet abutting the south pole of the first U-shaped magnet. A portion of the first magnetic field generated by the first U-shaped magnet directly adjacent to the rear of the first U-shaped magnet interacts with a stationary fourth magnetic field to cause the rotor to rotate. A second magnetic field generated by the north pole of the second U-shaped magnet and a third magnetic field generated by the south pole of the third U-shaped magnet interact with the fourth magnetic field to cause the rotor to translate in the direction of the rotor axis.

A further aspect of the present invention comprises a rotor including a rotor axis, and a thruster axis in a plane of the rotor and intersecting the rotor axis. The rotor provides motive power by a rotation of the rotor about the rotor axis and by a translation of the rotor in a direction of the rotor axis. The rotor comprises: a first U-shaped magnet having a north pole and a south pole and a rear side, the north pole and the south pole being generally aligned with the thruster axis, the first U-shaped magnet generating a first magnetic field; a first thruster magnet having a direction of magnetisation generally aligned with the thruster magnet axis, the first thruster magnet being proximate to and spaced from the north pole of the first U-shaped magnet; and a second thruster magnet having a direction of magnetisation generally aligned with the thruster magnet axis, the second thruster magnet being near to and spaced from the south pole of the first U-shaped magnet, the first U-shaped magnet being interposed between the first and the second thruster magnets. A portion of the first magnetic field generated by the first U-shaped magnet directly adjacent to the rear side of the first U-shaped magnet interacts with a stationary fourth magnetic field to cause the rotor to rotate, a second magnetic field generated by the first thruster magnet and a third magnetic field generated by the second thruster magnet respectively interact with a stationary fifth magnetic field to cause the rotor to translate in the direction of the rotor axis.

Another aspect of the present invention comprises a rotor providing motive power by rotation of the rotor about a rotor axis and translation of the rotor in the direction of the rotor axis. The rotor has at least one rotor magnet generating a first magnetic field, the first magnetic field being generated by the rotor magnet interacting with at least one stationary U-shaped magnet, the U-shaped magnet having a rear side and generating a second magnetic field. The rotational and translational motive power of the rotor is provided by an interaction of a portion of the second magnetic field directly adjacent to the rear of the U-shaped magnet with the first magnetic field.

A further aspect of the present invention comprises a motor providing unidirectional rotational motive power. The motor includes a generally circular stator having a stator axis, an outer surface, and a circumferential line of demarcation at about a midpoint of the outer surface; at least one stator magnet attached to the outer surface of the stator, the at least one stator magnet being arranged in a generally circular arrangement about the stator axis and generating a first magnetic field; an armature attached to the stator for rotation with it; the armature having an axis parallel to the stator axis; at least one rotor, the rotor being spaced from the armature and coupled to it by an axle to allow rotation about an axis of the rotor, the rotor rotating in a plane generally aligned with the armature axis, the rotor, including at least one magnet generating a second magnetic field, where the second magnetic field generated by the rotor magnet interacts with the first magnetic field to cause the rotor to rotate about it's axis; and a drive linkage assembly connecting the rotor to the stator to cause the armature to rotate about it's axis as the rotor rotates about it's axis, thereby providing the unidirectional rotational motive power of the motor.

In another aspect, the present invention is directed to a motor providing unidirectional rotational motive power comprising: a generally circular stator having an axis, an outer surface, and a circumferential line of demarcation around the outer surface, the line of demarcation having a pre-determined direction around the stator axis and separating a first side of the outer surface and a second side of the outer surface, wherein at least one pair of stator magnets is attached to the outer surface generating a first magnetic field, the pair of magnets comprising a first stator magnet having a north pole and a south pole and a second stator magnet having a north pole and a south pole, the south pole of the first stator magnet being located on the first side of the outer surface and the north pole of the first stator magnet being closest to the line of demarcation, the north pole of the second stator magnet being located on the second side of the outer surface and the south pole of the second stator magnet being closest to the line of demarcation, wherein the at least one pair of stator magnets is spaced along the line of demarcation so that a first inter-magnet distance measured along the line of demarcation between the north pole of the first stator magnet and the south pole of the second stator magnet of an adjacent pair of the at least one pair of stator magnets is generally equal to a second inter-magnet distance measured along the line of

demarcation between the south pole of the first stator magnet and the north pole of the second stator magnet; an armature attached to the stator, the armature having an axis parallel to the stator axis and attached to the stator for rotation therewith; and at least one rotor attached to the armature, the at least one rotor being spaced from the armature and coupled to it by an axle for rotation about an axis of the rotor, the rotor rotating in a plane generally aligned with the armature axis, the rotor comprising at least one rotor magnet, the rotor magnet generating a second magnetic field which interacts with the first magnetic field to cause the rotor to rotationally oscillate about the axis of the rotor and to generate a force in a direction of the rotor axis, thereby causing the armature to rotate in the pre-determined direction around the armature axis to provide the unidirectional rotational motive power of the motor.

In a further aspect, the present invention is directed to a motor providing unidirectional linear motive power comprising: a linear stator having a generally curved cross-section and a longitudinal line of demarcation perpendicular to the cross-section extending on about a midpoint of a surface of the stator between a first end and a second end of the stator, the stator including at least one magnet arranged between the first end and the second end, the magnet having a direction of magnetisation at about a right angle to the line of demarcation and generating a first magnetic field, the magnitude of the first magnetic field being generally uniform along the line of demarcation except in a pre-determined number of null regions, wherein the first magnetic field is substantially zero a rail connected to the stator, the rail having a longitudinal axis generally parallel to the line of demarcation and a helical groove with a pre-determined pitch running around a periphery of the rail; at least one rotor having a rotor axis aligned with the axis of the rail, the rotor being connected to the rail so that the rotor is free to rotate about the axis of the rail and slide along the rail, the rotor including at least one U-shaped magnet having a rear side and generating a second magnetic field, where a portion of the second magnetic field directly adjacent to the rear of the U-shaped magnet interacts with the first magnetic field to cause the rotor to rotate about the axis of the rail; a bearing assembly connecting the rotor to the helical groove, the bearing assembly converting the rotary motion of the rotor about the axis of the rail to linear motion along the rail; and a cross-link connecting the bearing assembly of a first rotor to a second rotor, thereby adding together the linear motion along the rail of the first rotor and the second rotor to provide the unidirectional linear motive power.

In yet another aspect, the present invention is directed to a motor providing unidirectional motive power comprising: a rail having a longitudinal axis and at least one helical groove having a pre-determined pitch running around a periphery of the rail; at least one first helical stator concentrically surrounding the rail, the first helical stator having the pre-determined pitch of the groove and a longitudinal axis generally parallel to the axis of the rail, at least one first stator magnet being attached to the first helical stator, the first stator magnet generating a first magnetic field; at least one rotor having an axis generally aligned with the axis of the rail, the rotor being connected to the rail so that the rotor is free to rotate about the axis of the rail and slide along the rail, the rotor comprising at least one rotor magnet generating a second magnetic field, the second magnetic field interacting with the first magnetic field generated by the first stator magnet to cause the rotor to rotate about the axis of the rail; and a bearing assembly connecting the rotor to the helical groove around the periphery of the rail, the bearing assembly converting the rotational motion of the rotor about the rail to unidirectional linear motion along the rail.

A further aspect of the present invention is directed to a motor providing unidirectional motive force comprising: a rail having a longitudinal axis and a helical groove running around the rail, the groove having a predetermined pitch; at least one first helical stator comprising a plurality of discontinuous spaced apart first ribs, each first rib partially surrounding the rail at a generally uniform distance from the rail, the first helical stator having the pre-determined pitch of the groove and a longitudinal axis generally aligned with the rail, at least one first stator magnet being attached to each rib, each first stator magnet generating a first magnetic field; at least one rotor having an axis generally aligned with the axis of the rail, the rotor being connected to the rail so that the rotor is free to rotate about the axis of the rail and to slide along the rail, the rotor comprising at least one rotor magnet generating a second magnetic field, the second magnetic field interacting with the first magnetic field generated by the first stator magnet to cause the rotor to rotate about the axis of the rail; and a bearing assembly connecting the rotor to the helical groove around the rail, the bearing assembly converting the rotary motion of the rotor about the rail to linear motion along the rail.

The present invention is further directed to a motor providing unidirectional motive power comprising: a rail having a longitudinal axis and a generally sinusoidal groove running around a periphery of the rail, the sinusoidal groove having a pre-determined period; at least one stator having a generally curved cross-section and a longitudinal line of demarcation perpendicular to the cross-section located at about a midpoint of a surface of the stator, the surface of the stator being disposed generally equidistant from and parallel to the axis of the rail; at least one stator magnet attached to the surface of the stator generating a first magnetic field, the stator magnet having a magnetisation which is displaced sinusoidally from the line of demarcation, the sinusoid having a pre-determined period and a pre-determined maximum amplitude and being divided into a plurality of alternating first and second sectors, with a boundary between the alternating first and second sectors occurring at the maximum amplitude of the sinusoid, the direction of magnetisation of the stator magnet being opposite in direction in the first and second segments; at least one rotor having an axis aligned with the axis of the rail, the rotor being connected to the rail so that the rotor is free to rotate about the axis of the rail and slide along the rail, the rotor including at least one U-

shaped magnet having a rear side and generating a second magnetic field, the U-shaped magnet being positioned on the rotor so that the rear side of the U-shaped magnet is apposite to the first and the second segments of the stator as the rotor rotates about the rotor axis, wherein an interaction of a portion of the second magnetic field directly adjacent to the rear of the U-shaped magnet with the first magnetic field causes the rotor to rotationally oscillate about the axis of the rail; and a bearing assembly connecting the rotor to the sinusoidal groove around the rail, the bearing assembly converting the oscillatory motion of the rotor about the rail to unidirectional linear motion along the rail.

The present invention is also directed to a motor providing unidirectional motive power comprising: a rail having a longitudinal axis and a helical groove running around a periphery of the rail, the helical groove having a pre-determined pitch; at least one stator having a generally having a longitudinal line of demarcation located at about a midpoint of a surface of the stator, the surface of the stator being disposed generally equidistant from and parallel to the axis of the rail; at least one stator magnet attached to the surface of the stator, the stator magnet having a direction of magnetisation which rotates about a magnetic axis parallel to the line of demarcation with a predetermined pitch, thereby generating a first magnetic field having a substantially uniform magnitude along the magnetic axis and rotates around the magnetic axis with the pre-determined pitch of the stator magnet rotation; at least one rotor having an axis aligned with the axis of the rail, the rotor being connected to the rail so that the rotor is free to rotate about the axis of the rail and slide along the rail, the rotor including at least one U-shaped magnet generating a second magnetic field, the U-shaped magnet being positioned on the rotor so that a portion of the second magnetic field directly adjacent to the rear side of the U-shaped magnet interacts with the first magnetic field of the stator magnet to cause the rotor to rotate about it's axis; and a bearing assembly connecting the rotor to the helical groove, the bearing assembly converting the rotary motion of the rotor about the rail to unidirectional linear motion along the rail.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWINGS

The foregoing summary, as well as the following detailed description of preferred embodiments of the invention, will be better understood when read in conjunction with the appended drawings. For the purpose of illustrating the invention, there are shown in the drawings embodiments which are presently preferred. It should be understood, however, that the invention is not limited to the precise arrangements and instrumentalities shown. In the drawings:

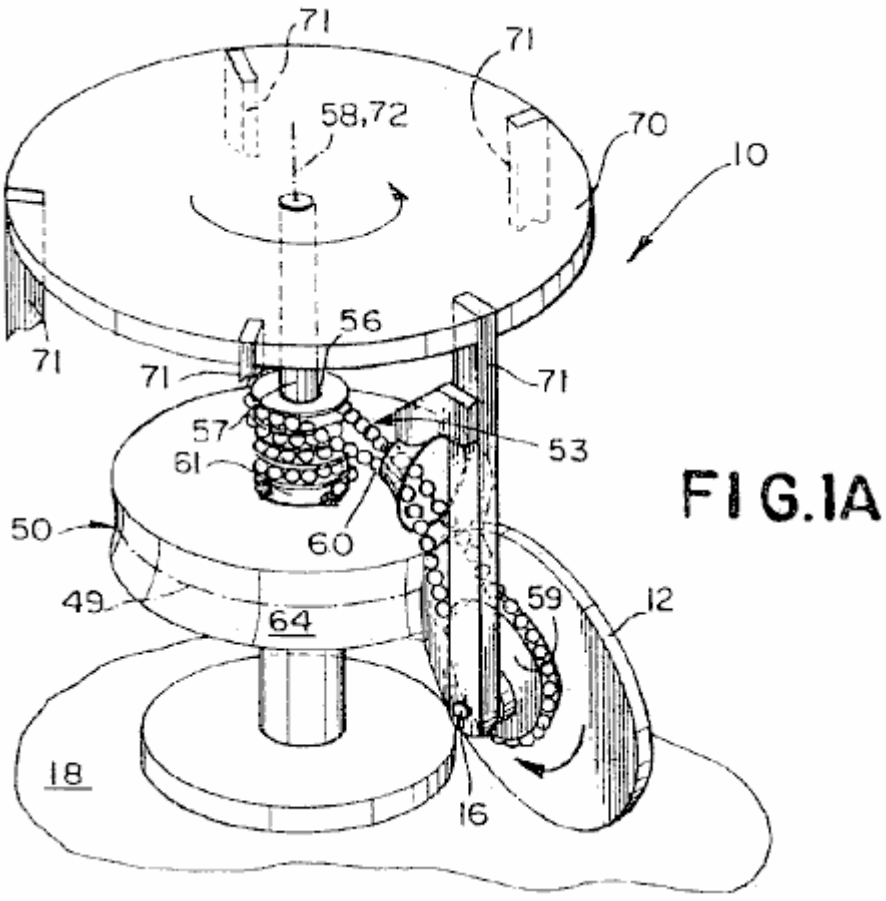


FIG.1A

Fig.1A is a schematic perspective drawing of a first preferred embodiment of a motor providing unidirectional motive power;

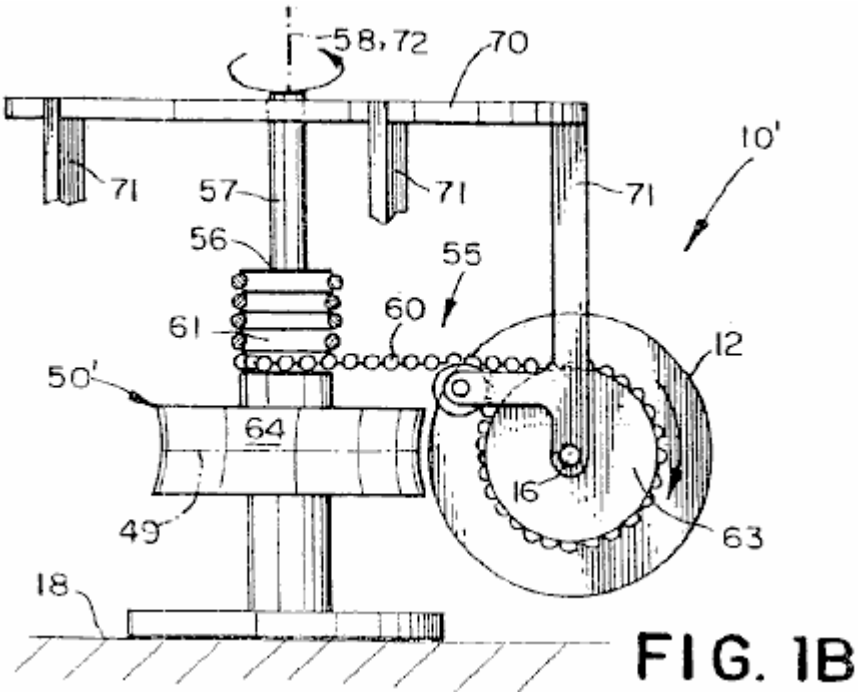


FIG. 1B

Fig.1B is a schematic perspective drawing of a second preferred embodiment of the motor;

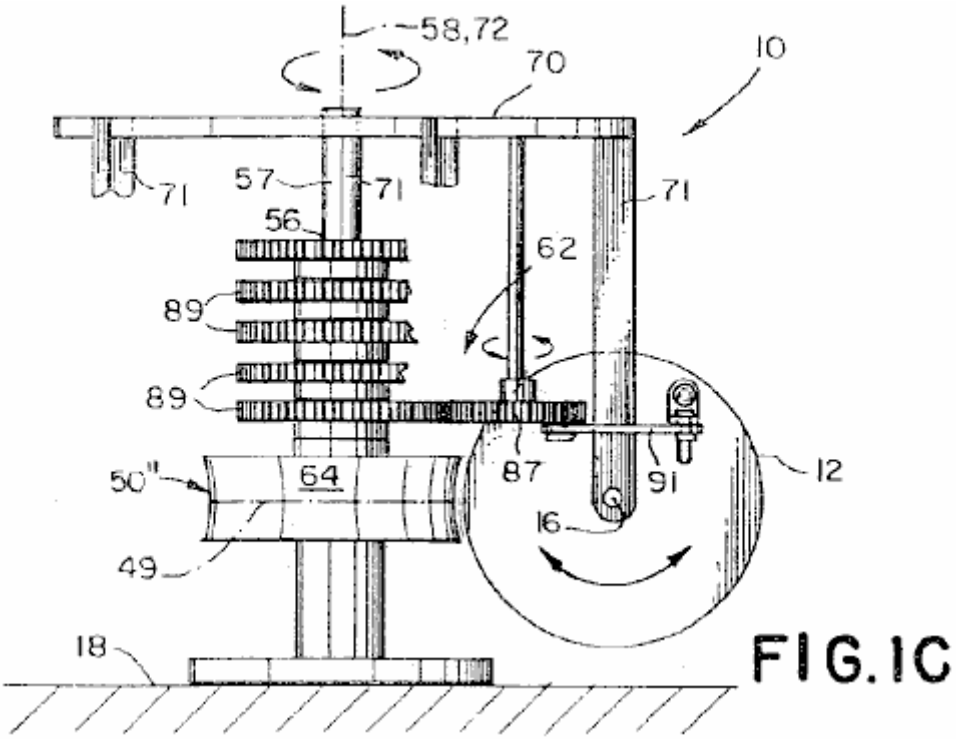


FIG. 1C

Fig.1C is a schematic perspective drawing of a third preferred embodiment of the motor;

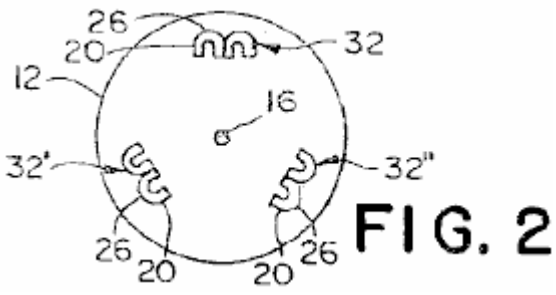


FIG. 2

Fig.2 is a schematic plan view of a rotor comprising three pair of U-shaped magnets;

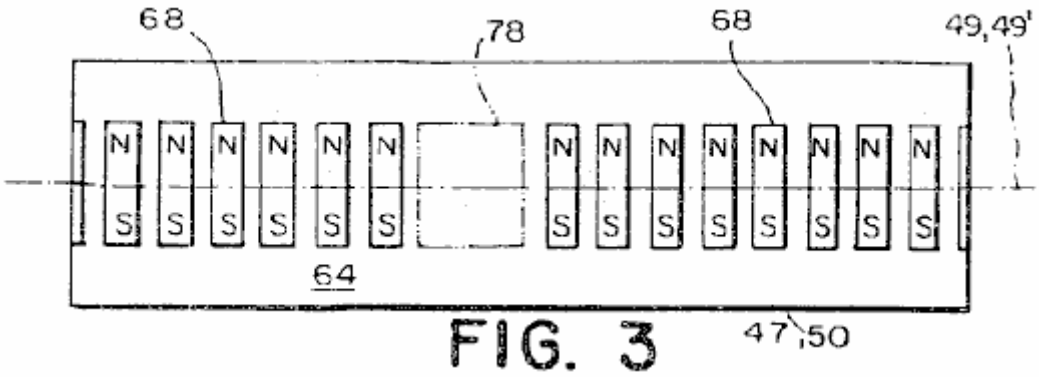


FIG. 3

Fig.3 is a schematic plan view of stator having a plurality of stator magnets generating a uniform magnetic field except in single null region, laid out flat for ease of illustration;

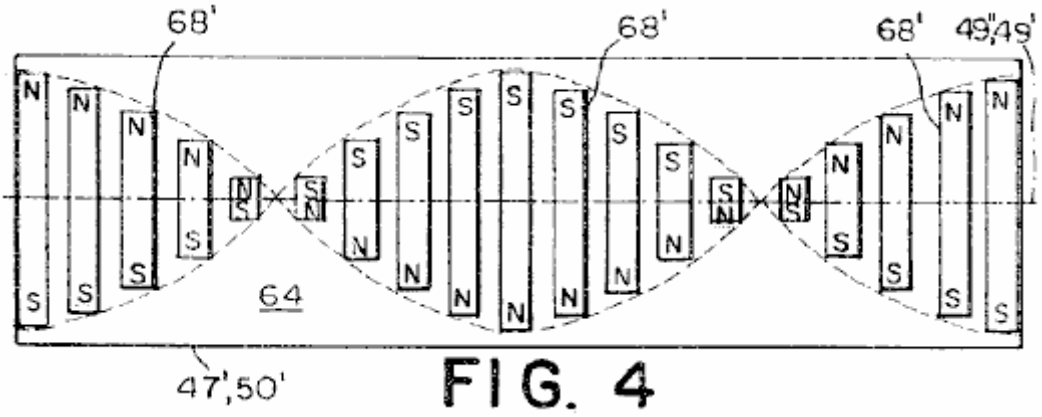


FIG. 4

Fig.4 is an schematic plan view of a stator having a plurality of stator magnets which rotate about a magnetic axis, laid out flat for ease of illustration;

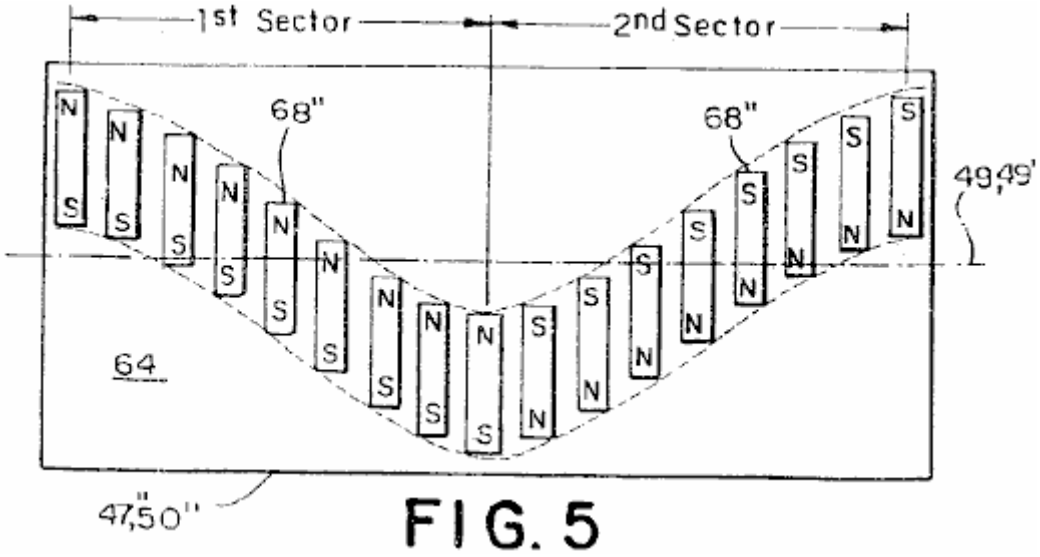


FIG. 5

Fig.5 is an schematic plan view of a stator having a plurality of stator magnets which are sinusoidally displaced from a line of demarcation, laid out flat for ease of illustration;

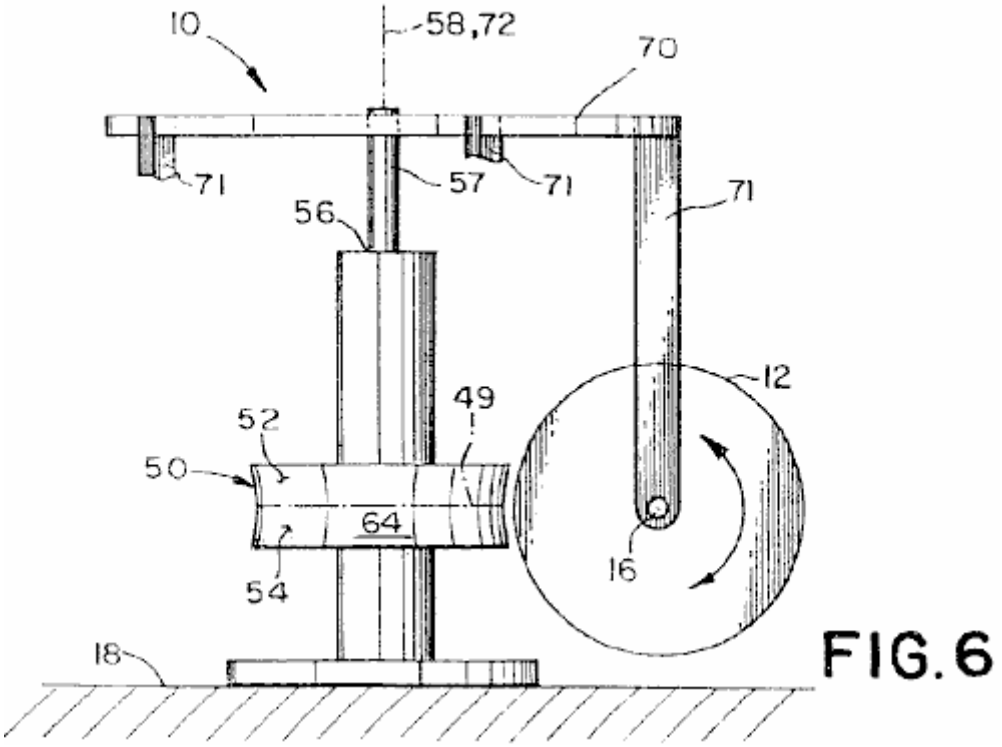


FIG.6

Fig.6 is a schematic perspective view of a fourth through a seventh preferred embodiment of the motor;

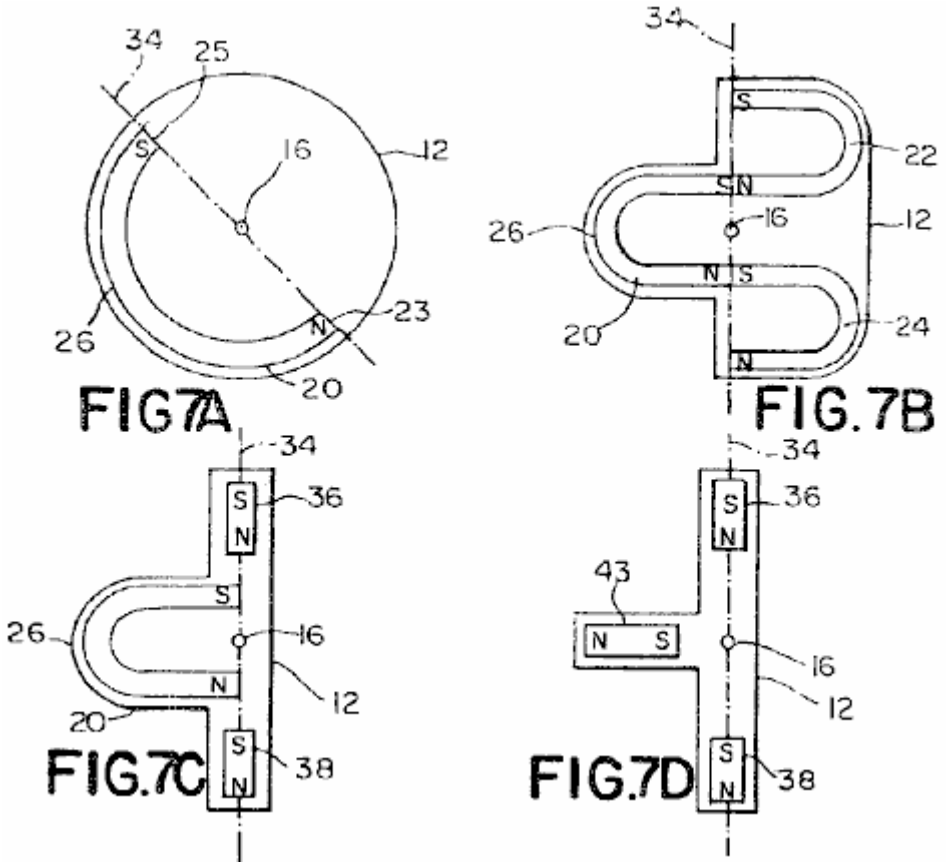


Fig.7A is a schematic plan view of a rotor used in the fourth preferred embodiment and in an eighth preferred embodiment of the motor;

Fig.7B is a schematic plan view of a rotor used in a fifth preferred embodiment and in a ninth preferred embodiment of the motor;

Fig.7C is a schematic plan view of a rotor used in a sixth preferred embodiment and in a tenth preferred embodiment of the motor;

Fig.7D is a schematic plan view of a rotor used in the seventh preferred embodiment and in an eleventh preferred embodiment of the motor;

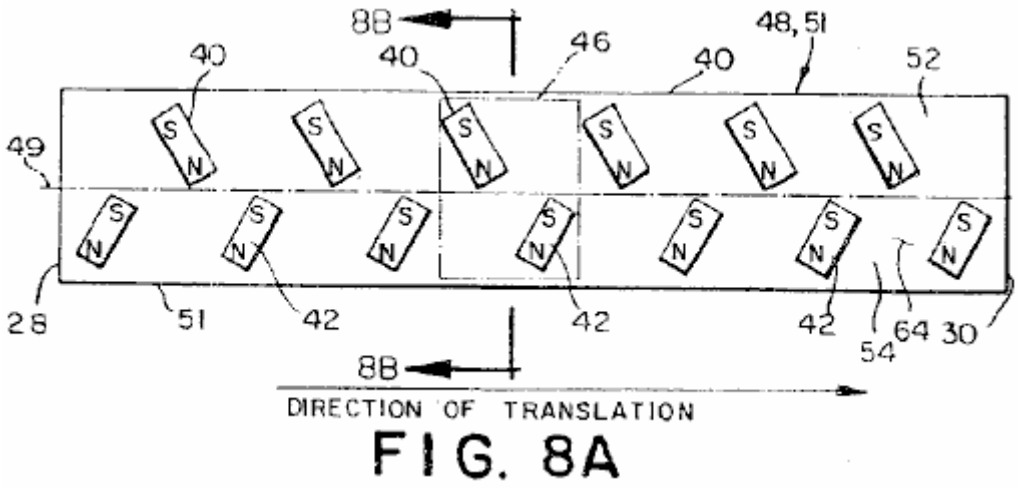


Fig.8A is a schematic plan view of a stator used in the fourth, fifth, eighth and ninth preferred embodiments of the motor;

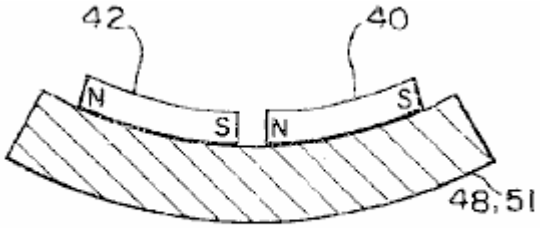


FIG. 8B

Fig.8B is a schematic sectional view of the stator shown in Fig.8A taken along the line 8B-8B;

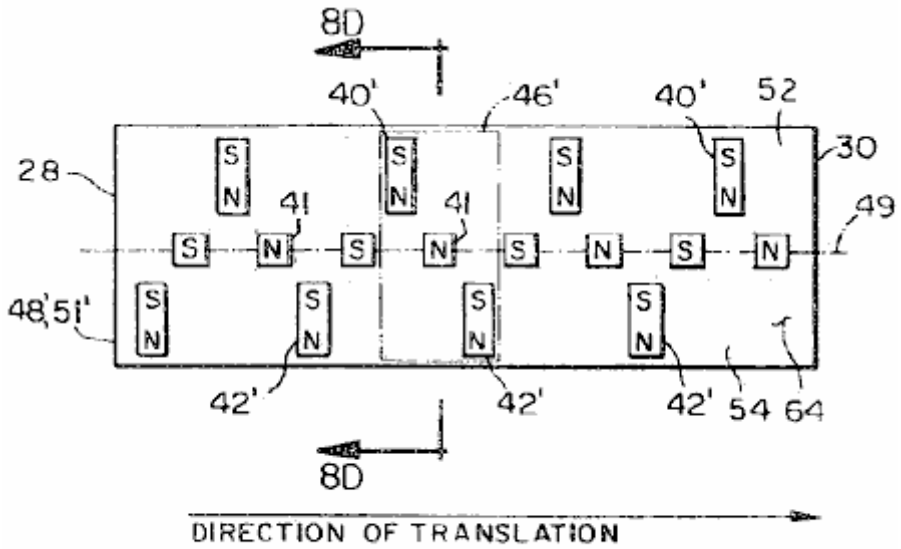


FIG. 8C

Fig.8C is a schematic plan view of a stator used in the sixth and in the tenth preferred embodiments of the motor;

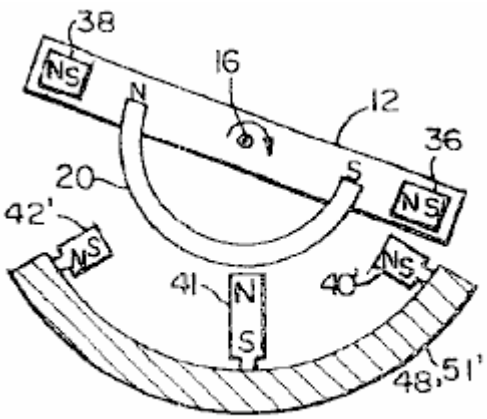


FIG. 8D

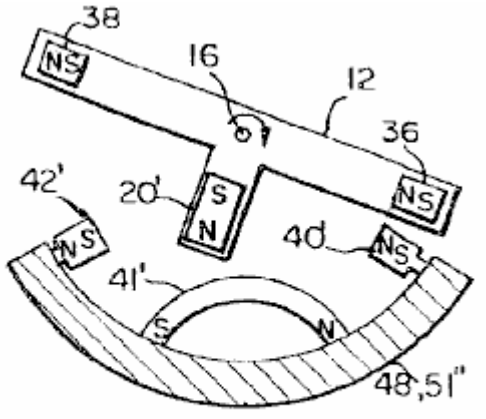


FIG. 8E

Fig.8D is a schematic elevational view of the stator shown in Fig.8C taken along the line 8D-8D shown with the rotor shown in Fig.7C;

Fig.8E is a schematic elevational view of an alternative stator shown with the rotor shown in Fig.7D;

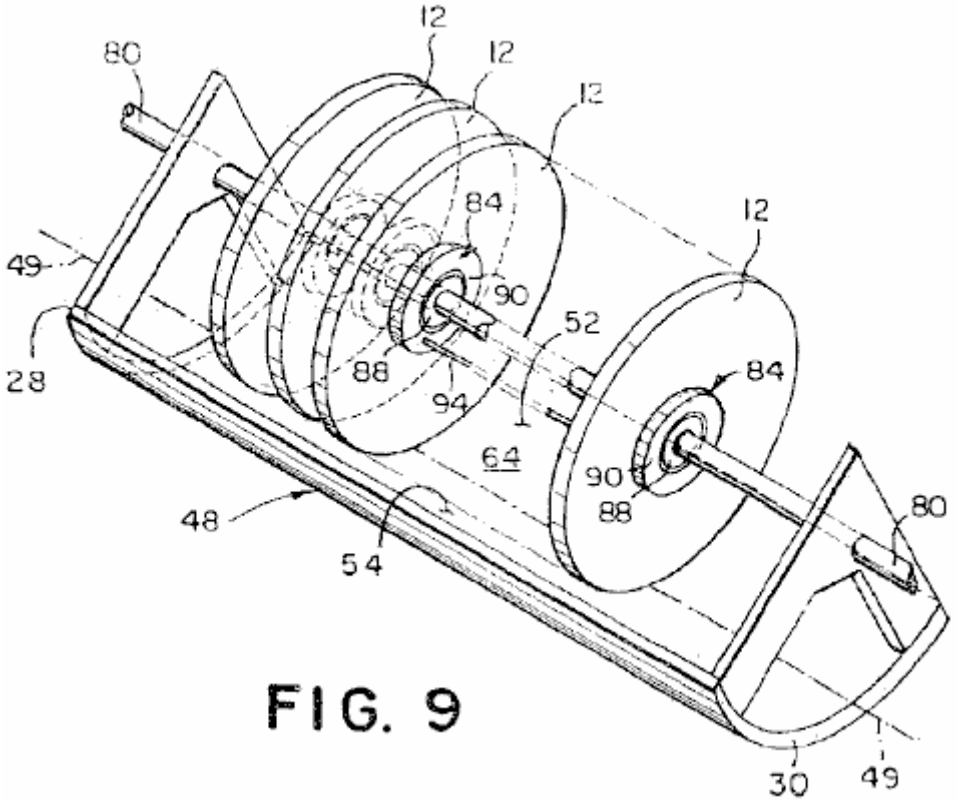


FIG. 9

Fig.9 is a schematic perspective view of the eighth through an eleventh preferred embodiment of the motor;

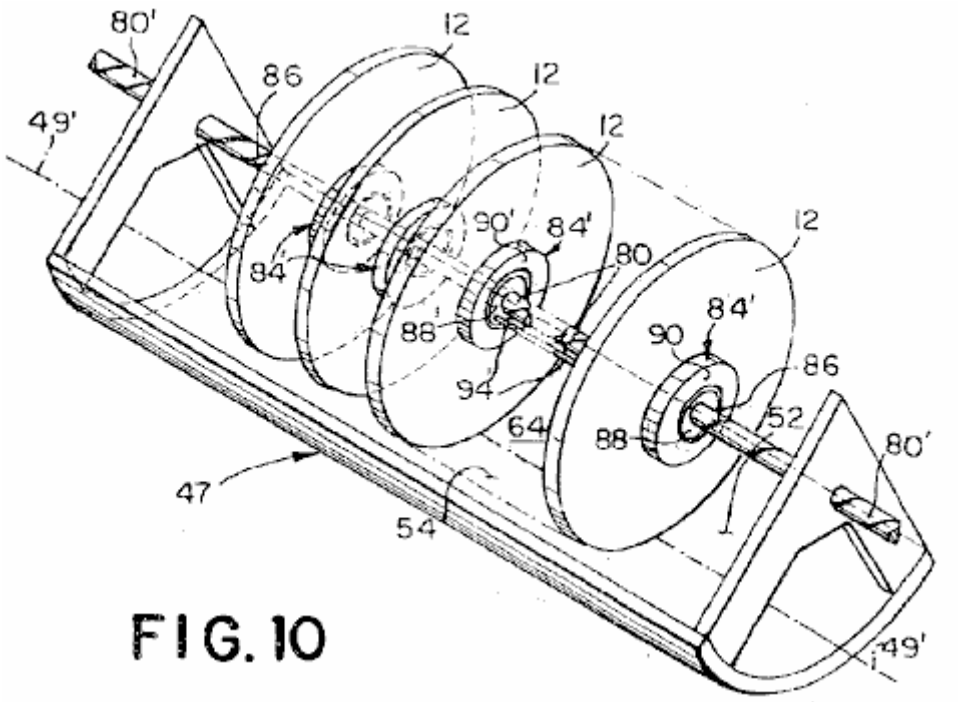


FIG. 10

Fig.10 is a schematic perspective view of a twelfth preferred embodiment of the motor;

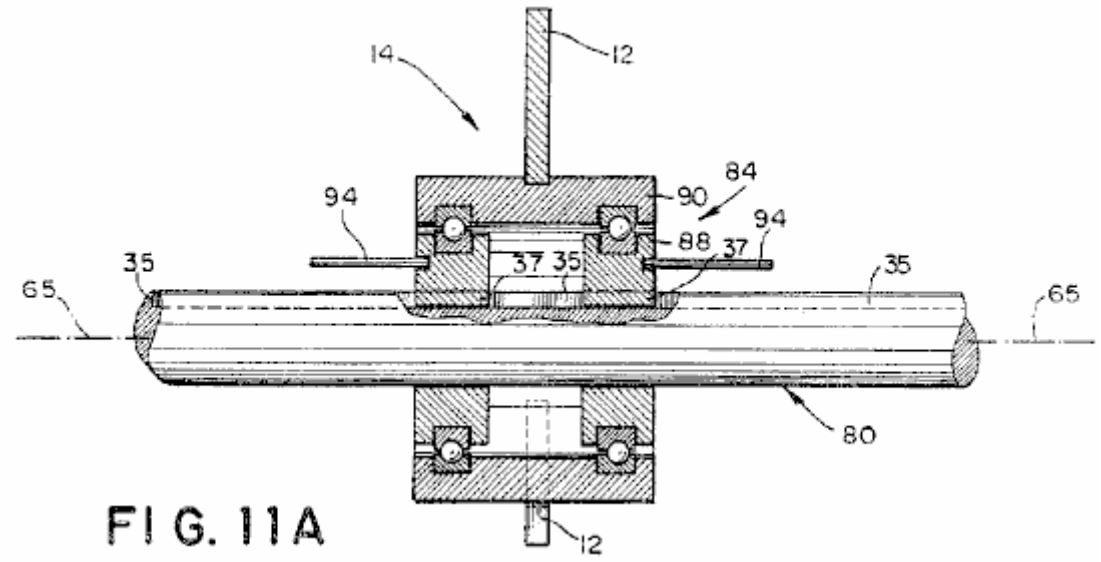


FIG. 11A

Fig.11A is a plan view of a rotor assembly used in the eighth through the eleventh preferred embodiments;

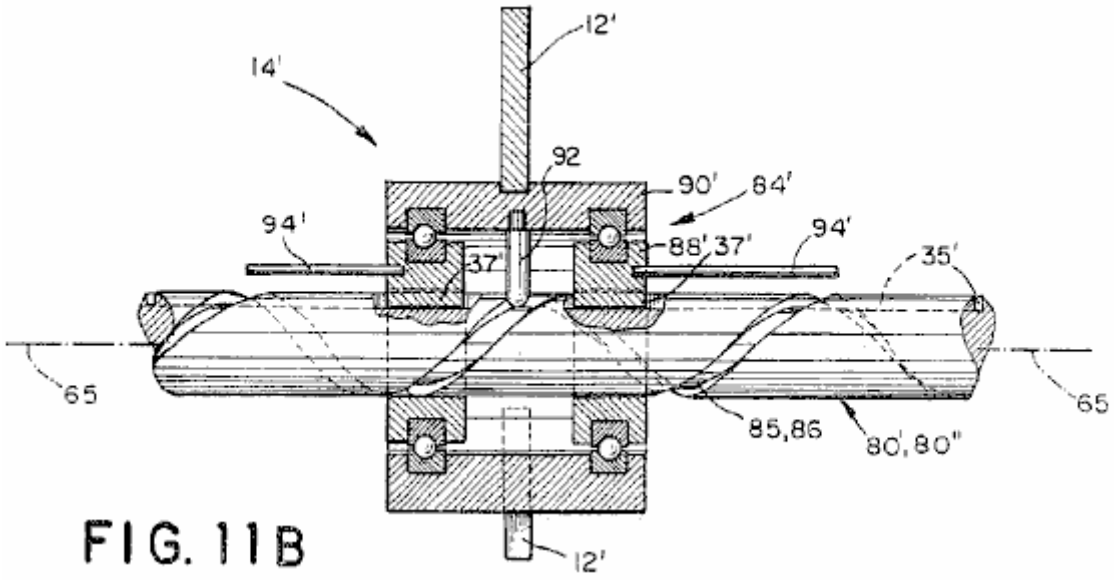


FIG. 11B

Fig.11B is a plan view of a rotor assembly used in the twelfth through a sixteenth preferred embodiment;

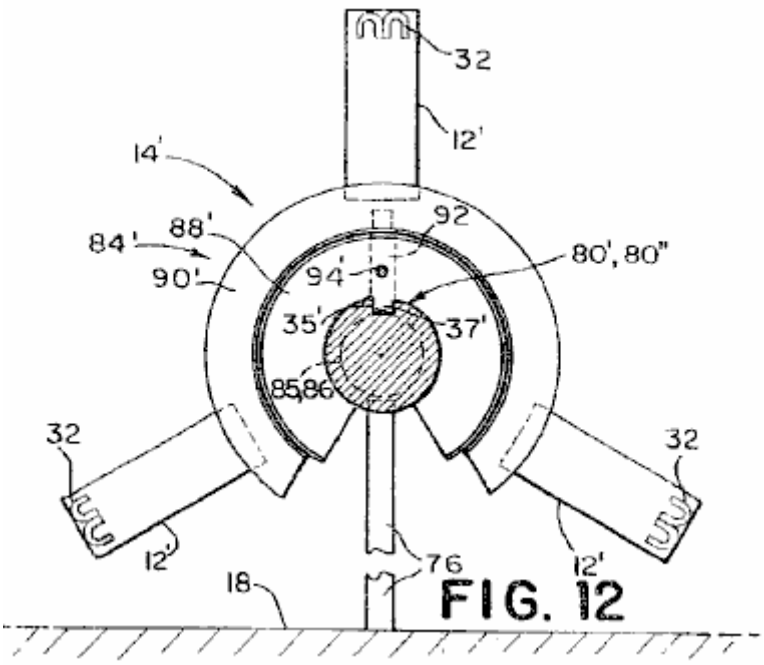


FIG. 12

Fig.12 is an end elevational view of the rotor assembly shown in Fig.11B, further including a rail mounting post;

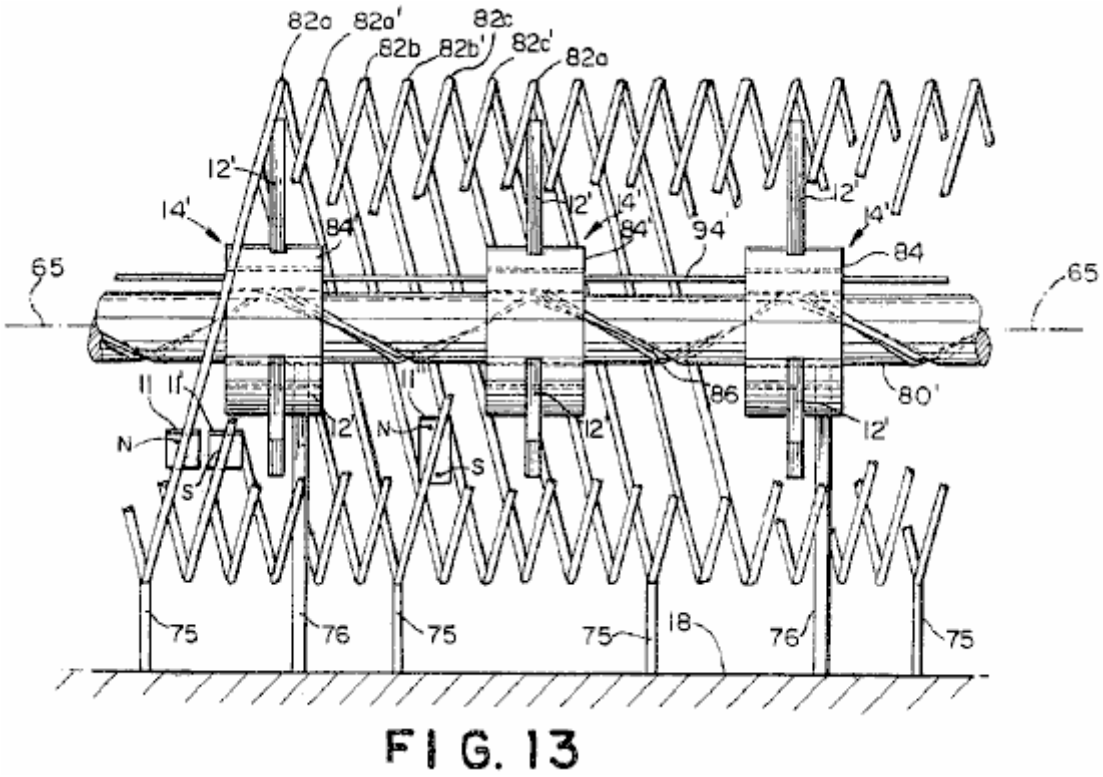


FIG. 13

Fig.13 is an elevational view of a thirteenth preferred embodiment of the motor;

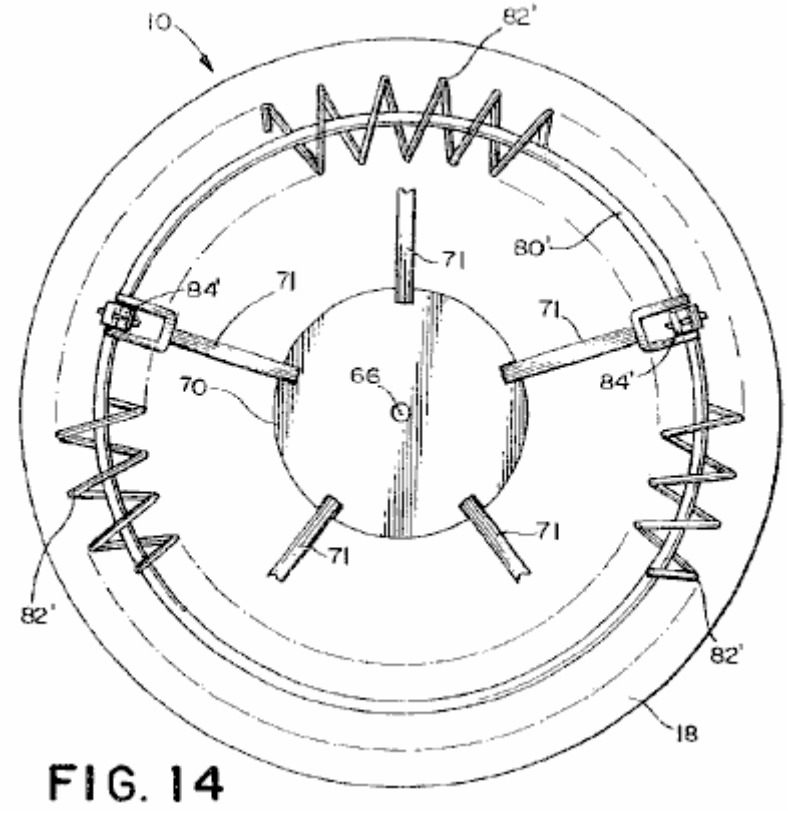


FIG. 14

Fig.14 is a plan view of a rotary configuration of the thirteenth preferred embodiment;

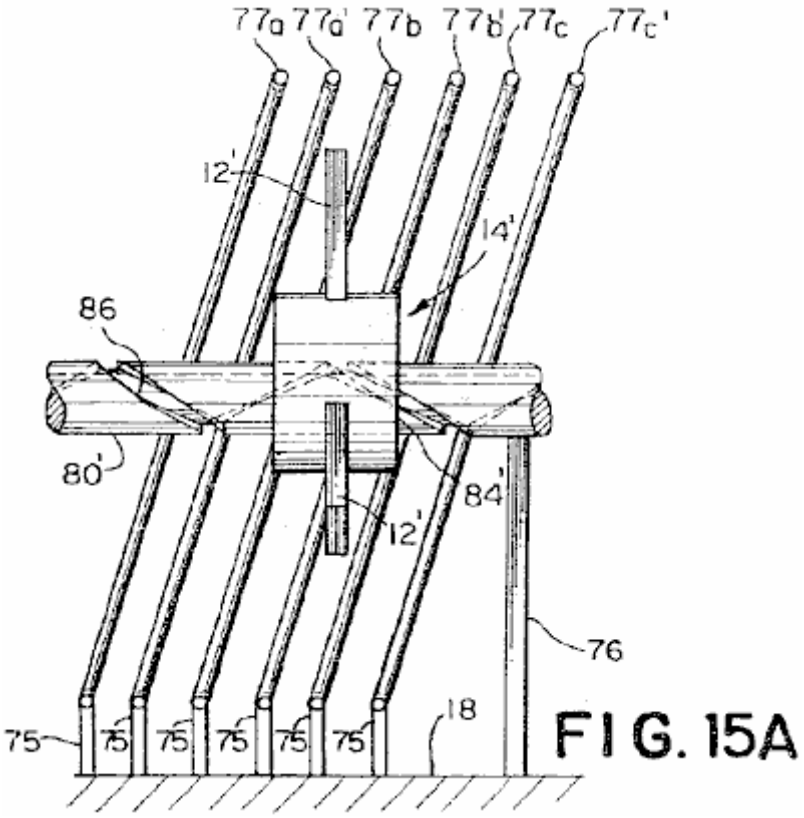


FIG. 15A

Fig.15A is an elevational view of a portion of a fourteenth preferred embodiment employing spaced apart ribs;

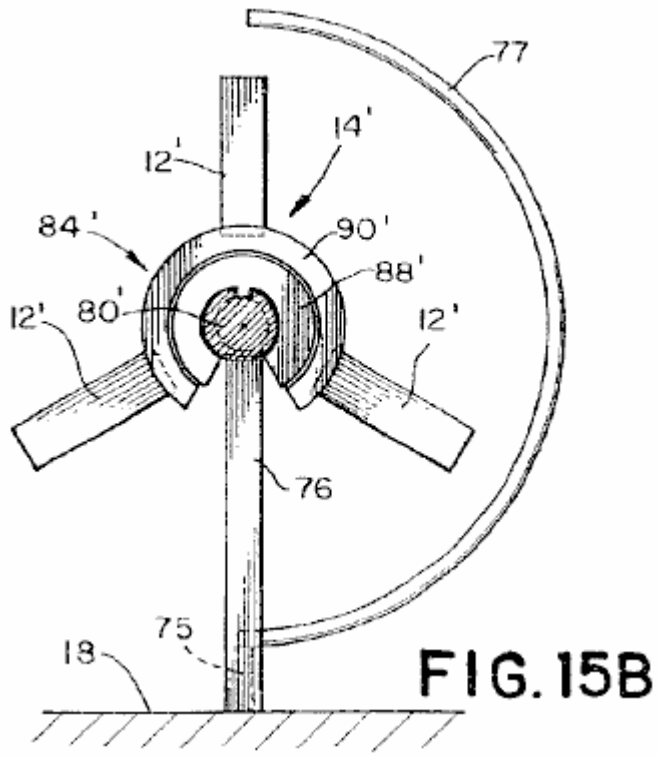


FIG. 15B

Fig.15B is an end elevational view of the fourteenth embodiment shown in Fig.15A;

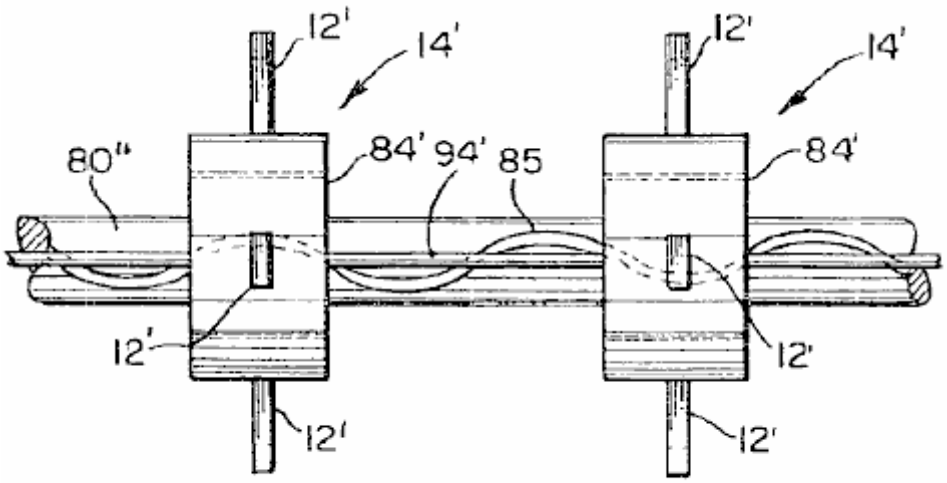


FIG. 16

Fig.16 is a top plan view of a portion of the fifteenth preferred embodiment of the motor;

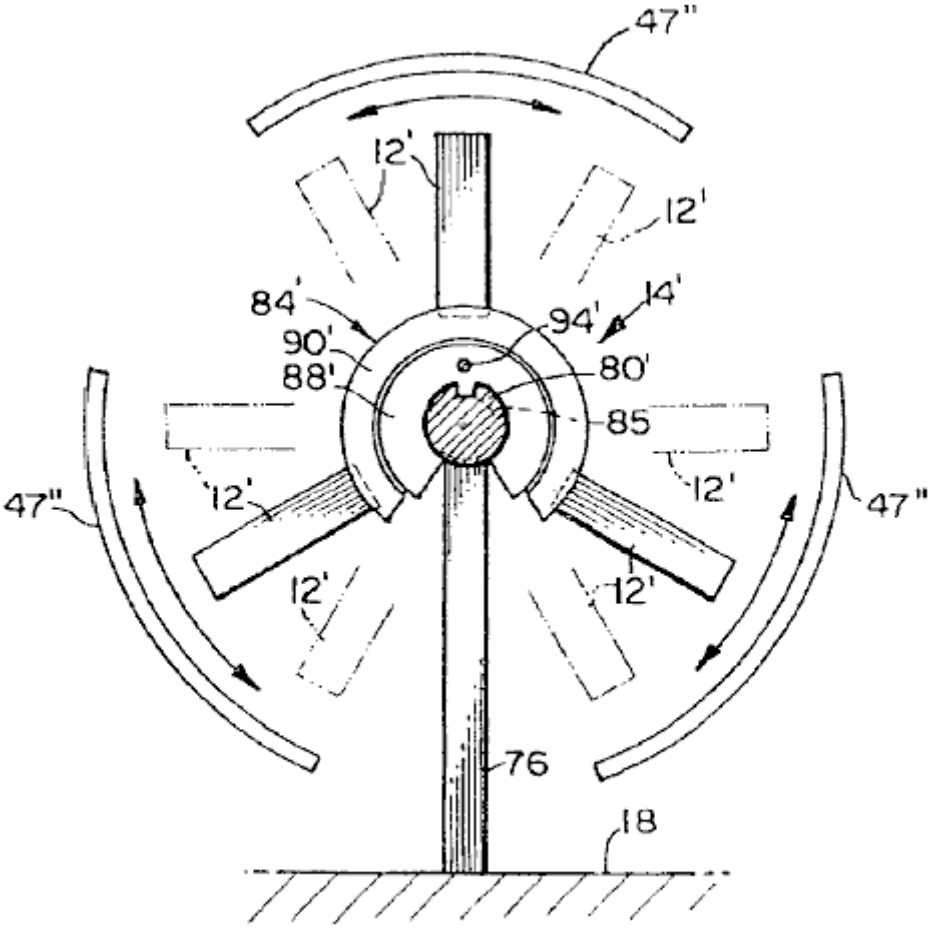


FIG. 17

Fig.17 is an elevational end view of the fifteenth preferred embodiment shown in Fig.16;

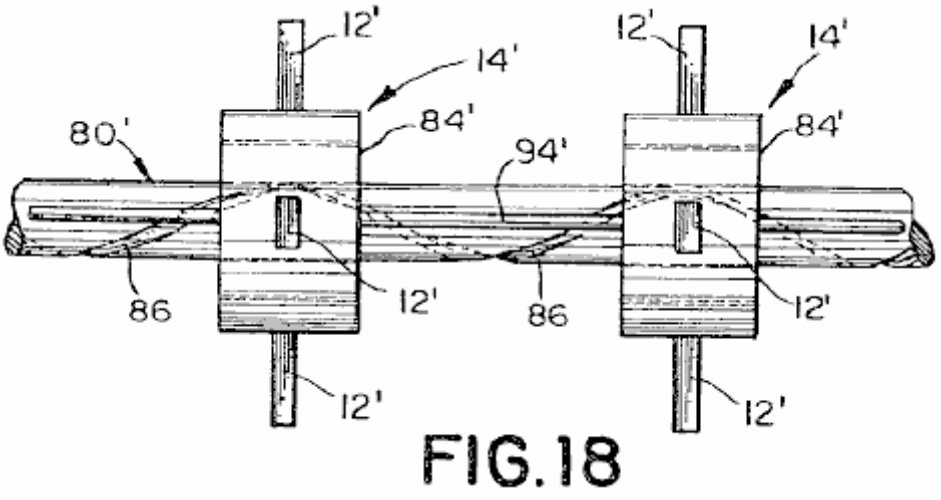


FIG.18

Fig.18 is a top plan view of a portion of the sixteenth preferred embodiment of the motor; and

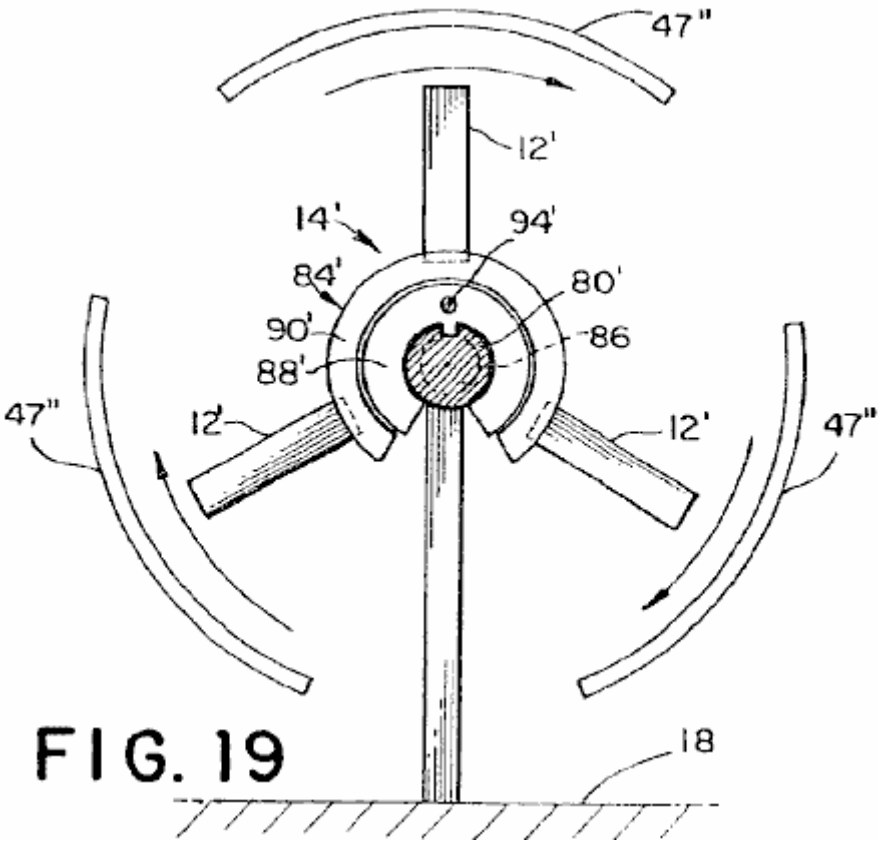


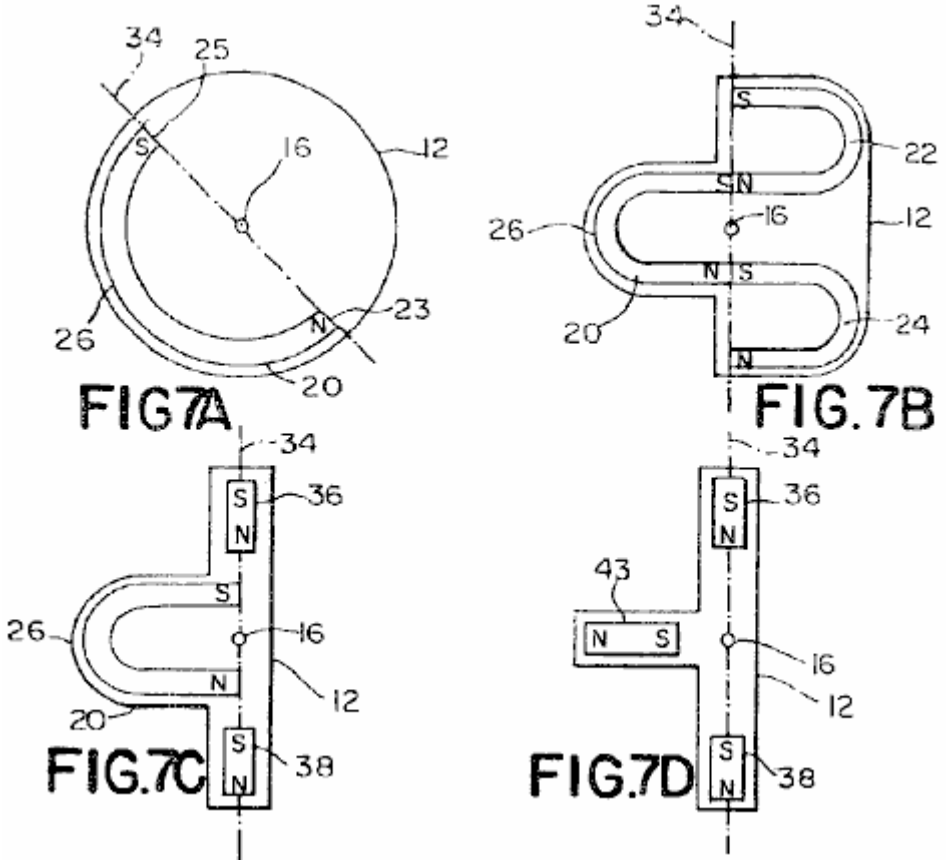
FIG. 19

Fig.19 is an elevational end view of the sixteenth preferred embodiment of the motor shown in Fig.18.

DETAILED DESCRIPTION OF THE INVENTION

It will be appreciated by those skilled in the art that changes could be made to the embodiments described above without departing from the broad inventive concept thereof. It is understood, therefore, that this invention is not limited to the particular embodiments disclosed, but it is intended to cover modifications within the spirit and scope of the present invention as defined by the appended claims. It should also be understood that the articles "a" and "the" used in the claims to define an element may refer to a single element or to a plurality of elements without a limit as to the number of elements.

Past attempts to construct a working permanent magnet motor have met with difficulties because of the simultaneous attractive and repulsive characteristics of a permanent magnet. A principle has been discovered where, by engaging a magnetic field at the rear of one or more U-shaped magnets mounted on a rotor with a second stationary magnetic field, a torque is created that rotates the rotor about a rotational axis of the rotor. Further, by properly shaping the second magnetic field, the rotor may be caused to also translate in the direction of the rotor axis.



Accordingly, using the aforementioned principle, and referring to **Fig.7A**, one aspect of the present invention is directed to a rotor **12** for use in a motor and which provides motive power by a rotation of the rotor **12** about a rotor axis **16** and by a translation of the rotor **12** in a direction of the rotor axis **16**. In one aspect, the rotor **12** comprises a first U-shaped magnet **20** in which the U-shaped magnet **20** generates a first magnetic field. A rotation of the rotor **12** about the rotor axis **16** is caused by an interaction of a portion of the first magnetic field directly adjacent to a rear **26** of the U-shaped magnet **20** with a stationary second magnetic field. A translation of the rotor **12** in the direction of the rotor axis **16** is caused by an interaction of the first magnetic field adjacent to a north pole **23** and a south pole **25** of the U-shaped magnet **20** with the stationary second magnetic field. As will be appreciated by those skilled in the art, the design of the rotor **12** is not limited to a single U-shaped magnet **12**. A plurality of U-shaped magnets **20**, arranged around a periphery of the rotor **12** is within the spirit and scope of the invention.

Another aspect of the present invention, shown in **Fig.7B** comprises a rotor **12** including a first U-shaped magnet having a north pole and a south pole generating a first magnetic field; a second U-shaped magnet **24** having a north pole and a south pole with the south pole of the second U-shaped magnet **24** abutting the north pole of the first U-shaped magnet **20**; and a third U-shaped magnet **22** having a north pole and a south pole with the north pole of the third U-shaped magnet **22** abutting the south pole of the first U-shaped magnet **20**. A portion of the first magnetic field generated by the first U-shaped magnet **20** directly adjacent to the rear **26** of the first U-shaped magnet **20** interacts with a stationary fourth magnetic field to cause the rotor **12** to rotate. A second magnetic field generated by the north pole of the second U-shaped magnet **24** and a third magnetic field generated by the south pole of the third U-shaped magnet **22** respectively interact with the fourth magnetic field to cause the rotor **12** to translate in the direction of the rotor axis **16**.

A further aspect of the present invention, shown in **Fig.7C**, comprises a first U-shaped magnet **20** having a north pole and a south pole generating a first magnetic field. The north pole and the south pole of the U-shaped magnet **20** are generally aligned with a thruster axis **34** which lies in the plane of the rotor **12** and intersects the rotor axis **16**. A first thruster magnet **36** is located proximate to and spaced from the north pole of the first U-shaped magnet with a direction of magnetisation being generally aligned with the thruster magnet axis **34**. A second thruster magnet **38** is located proximate to and spaced from the south pole of the first U-shaped magnet **20** with a direction of magnetisation also being generally aligned with the thruster magnet axis **34**. A portion of the first magnetic field generated by the first U-shaped magnet **20** directly adjacent to the rear side **26** of the first U-shaped magnet **20** interacts with a stationary fourth magnetic field to cause the rotor **12** to rotate. A second magnetic field generated by both the north pole and the south pole of the first thruster magnet **36** and a third magnetic field generated by both the north pole and the south pole of the second thruster magnet **38** respectively interact with a fifth magnetic field to cause the rotor **12** to translate in the direction of the rotor axis **16**. In one

further aspect of the rotor **12**, as shown in **Fig.7D**, a bar magnet **43** may be substituted for the U-shaped magnet **20** and the fourth magnetic field is formed by one or more U-shaped magnets, where the bar magnet **43** interacts with a portion of the stationary fourth magnetic field adjacent to the rear of a U-shaped magnet.

As will be appreciated by those skilled in the art, the polarities of the magnets shown in **Figs. 7A, 7B, 7C** and **7D** may be reversed and still be within the spirit and scope of the invention.

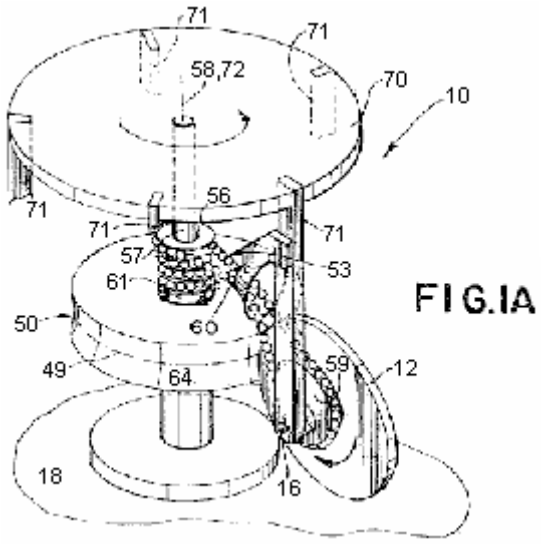


FIG.1A

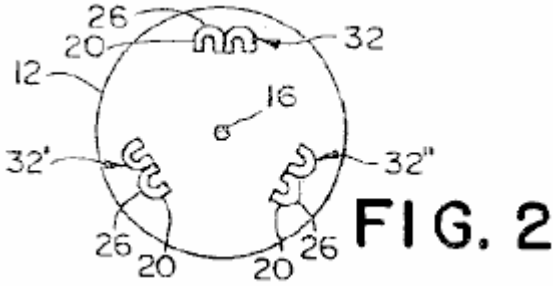


FIG. 2

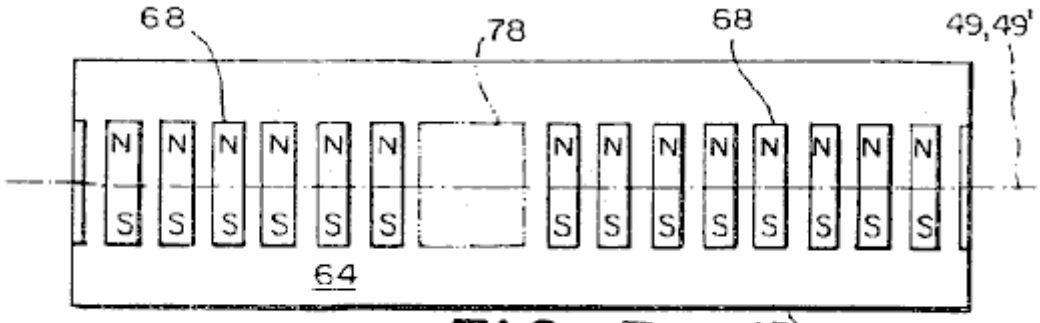


FIG. 3

Referring now to **Fig.1A**, **Fig.2** and **Fig.3** there is shown a first preferred embodiment of a motor **10** using the rotor **12** and providing unidirectional rotational motive power. The first preferred embodiment comprises a generally circular stator **50** having a stator axis **72** and a circumferential surface **64** mounted to a base **18**; an armature **70**, having an armature axis of rotation **58** coincident with the stator axis **72**, attached to the stator **50** by an armature axle **57** for rotation about the armature axis of rotation **58**; and five rotors **12** (only one of which is shown for clarity), the rotors **12** being spaced at intervals of about **72** degrees around the armature **70**. Each rotor **12** is spaced from the armature by an armature strut **71** and attached to the armature strut **71** by an axle, for rotation about an axis **16** of the rotor **12** in a plane generally aligned with the armature axis of rotation **58**. The motor **10** further includes a driving linkage assembly **53** connecting each rotor **12** and the stator **50** together, the linkage **53** urging the armature **70** to rotate about the armature axis of rotation **58** as each rotor **12** rotates about its respective rotor axis **16**. As will be appreciated by those skilled in the art the number of rotors **12** is not limited to the five rotors **12** disclosed in the first embodiment. Any number of rotors **12** from one to as many as there would be space for mounting on the armature **70** is within the spirit and scope of the invention.

Preferably, the surface **64** of the stator **50** is curved, having a curvature conforming to the arc of the rotors **12**. However, it will be appreciated by those skilled in the art that the surface **64** need not be curved but could be planar and still be within the spirit and scope of the invention. As will be appreciated by those skilled in the art the stator **50** is merely intended as a stationary supporting structure for stator magnets and, as such, the shape of the stator **50** is not intended to be controlling of the size and shape of the air gap between the magnets attached to the stator **50** and the magnets attached to the rotors **12**.

Preferably, the stator **50** is made of a material (or a combination of materials) having a magnetic susceptibility less than 10-3, i.e. a material displaying paramagnetic or diamagnetic properties. For example, the stator **50** could be

made of a non-magnetic metal such as aluminium or brass. Also, the rotor **12** could be made of a natural material such as wood, glass, a polymeric material or a combination of any of the aforementioned materials within the spirit and scope of the invention. Further, it should be understood that the aforementioned materials are preferred for the stators and all other parts of the motor **10** that could significantly disrupt the magnetic interaction between the stator and the rotor of all of the disclosed preferred embodiments of the motor **10**.

In the first preferred embodiment, the surface **64** of the stator **50** includes a circumferential line of demarcation **49** at about a midpoint of the surface **64** formed by an intersection with the surface **64** of a plane perpendicular to the armature axis of rotation **58**. As shown in **Fig.3**, the stator **50** includes a plurality of bar magnets **68** attached to the outer surface **64** along the line of demarcation **49**, except in a single null region **78** where the magnitude of the first magnetic field is substantially reduced. The bar magnets **68** have a direction of magnetisation at about a right angle to the line of demarcation **49** thereby creating a first magnetic field adjacent to the outer surface **64**, the magnitude and the direction of which is substantially uniform along the circumferential line of demarcation **49** around the axis **58** of the stator **50**, except within the null region **78**. As will be appreciated by those skilled in the art, the stator axis **72** need not be coincident with the armature axis of rotation **58**. Accordingly, a stator **50** arranged around the armature axis **58** at any location at which the stator axis **72** is parallel to the armature axis **58** and the surface **64** of the stator **50** faces the periphery of the rotors **12** thereby providing for the interaction between the first magnetic field and the second magnetic field around the armature axis **58**, is within the spirit and scope of the invention.

Preferably, as further shown in **Fig.3**, the bar magnets are attached to the surface **64** of the stator **50** so that the direction of magnetisation of the bar magnets **68** are about perpendicular to a radial line of the rotor **12**. However, the bar magnets **68** could also be attached to the surface **64** of the stator so that the direction of magnetisation of the bar magnets **68** is aligned with a radial line of the rotor **12**. The bar magnets **68** are preferably abutting so as to form the substantially uniform first magnetic field. However, it is not necessary for the bar magnets **68** to abut one another. Further, it is not necessary to use a plurality of bar magnets **68** to form the first magnetic field. A single magnet producing a uniform first magnetic field in the region in which the first magnetic field interacts with the second magnetic field of the rotors **12** would provide the required first magnetic field. Also, the number of null regions **78** may be more than one, depending upon the desired speed of the motor, as explained below.

Preferably, the stator magnets **68** are permanent magnets made of a neodymium-iron-boron material. However, as will be appreciated by those skilled in the art, any type of permanent magnet material displaying ferromagnetic properties could be used for the stator magnets **68**. For instance, stator magnets **68** made of samarium cobalt, barium ferrite or AlNiCo are within the spirit and scope of the invention. It should be understood that these permanent magnet materials or their equivalents are preferred for the stator magnets and the rotor magnets of all of the disclosed preferred embodiments of the motor **10**. Also, while the use of permanent magnets is preferred, the use of electro-magnets for some or all of the magnets is within the spirit and scope of the invention.

As discussed above, the stator **50** may include a pre-determined number of null regions **78** on the surface of the stator **64**. In the first preferred embodiment, the single null region **78** is formed by a shield of a ferromagnetic material, such as iron, placed adjacent to the surface **64**. However, as those skilled in the art will appreciate, the null region **78** can also be formed by an absence of the bar magnets **68** in the region coinciding with the null region **78**. The null region **78** of substantially reduced magnetic field magnitude may also be formed by an auxiliary magnetic field suitably generated by one or more permanent magnets or by one or more electromagnets powered by an electric current arranged so that the auxiliary magnetic field substantially cancels the first magnetic field in the null region **78**. In the case of the electromagnets, the electric current may be turned off in synchronism with the rotation of the rotors **12** passing through the null region **78**, in order to conserve power. Preferably, the first magnetic field is reduced to ten percent or less of the magnetic force outside of the null region. However, the motor **10** will operate with a reduction of only fifty percent. Accordingly, a motor **10** having a substantial reduction of the first magnetic field of fifty percent or less is within the spirit and scope of the invention.

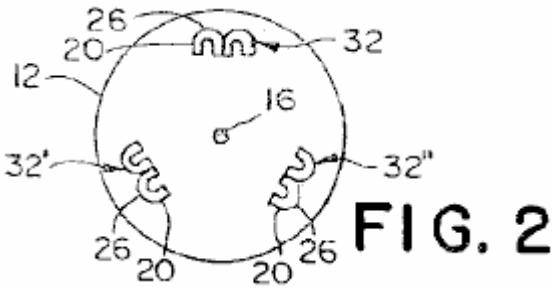


FIG. 2

As shown in **Fig.2**, the rotor **12** of the first preferred embodiment includes three pairs **32**, **32'**, **32''** of abutted U-shaped magnets **20** spaced apart at about 120 degree intervals around the periphery of the rotor **12**. Preferably,

the U-shaped magnets **20** having substantially identical magnetic properties and are arranged to have opposite poles of the abutting each other. The pairs **32**, **32'**, **32"** of abutted U-shaped magnets **20** are positioned so that the north pole and the south poles of each U-shaped magnet **20** face toward the axis of the rotor **16**, and the rear side **26** of each U-shaped magnet **20**, opposite to the north and the south pole of the U-shaped magnet **20**, faces out from the axis of the rotor **16** toward the surface **64** of the stator **50**. The pairs **32**, **32'**, **32"** of the U-shaped magnets **20** are situated on the rotor **12** so that a portion of the second magnetic field directly adjacent to the rear **26** of each U-shaped magnet **20** interacts with a first stationary magnetic field to cause the rotor **12** to rotate about its respective rotor axis **16**. Those skilled in the art will appreciate that it is not necessary to have exactly three pairs **32**, **32'**, **32"** of U-shaped magnets **20** on the rotor **12**. For instance, the number of U-shaped magnets **20** (or groups of abutted U-shaped magnets) spaced apart around the periphery of the rotor **12** may range from merely a single U-shaped magnet **20**, up to a number of magnets limited only by the physical space around the periphery of the rotor **12**. Further, the number of abutted U-shaped magnets **20** within each group of magnets **32** is not limited to two magnets but may also range from 1 up to a number of magnets limited only by the physical space around the periphery of the rotor **12**.

Preferably, the rotor **12** is made of a material (or a combination of materials) having a magnetic susceptibility less than 10-3. Accordingly, the rotor could be made of any of the same materials used to make the stator, such as for instance, a non-magnetic metal, wood, glass, a polymeric or a combination of any of the above as shown in **Fig.1A**, the rotor **12** is preferably disk shaped with the rear **26** of the U-shaped rotor magnets **20** being arranged on the periphery of the rotor **12** in such a way that the U-shaped magnets **20** pass in close proximity to the circumferential line of demarcation **49** on the outer surface **64** of the stator **50** as the rotor **12** rotates. However, as will be clear to those skilled in the art, the structure of the rotor **12** need not be disk shaped. The rotor **12** could be a structure of any shape capable of rotating around the rotor axis **16** and capable of supporting the U-shaped magnets **20** so that, as the rotor **12** rotates, the U-shaped magnets **20** come into close proximity with the outer surface **64** of the stator **50**. For example, a rotor **12** comprised of struts connected to a central bearing, where each strut holds one or more U-shaped magnets **20**, is within the spirit and scope of the invention.

In the first preferred embodiment, the linkage **53** connecting each rotor **12** and the stator **50** comprises a beaded chain drive **60** which meshes with a stator sprocket **61** on the stator **50**, and an eccentric rotor sprocket **59** on each rotor **12** so that, as each rotor **12** rotates about its respective rotor axis **16**, the armature **70** is forced to rotate about the armature axis of rotation **58**. The eccentric rotor sprocket **59** causes the instantaneous angular velocity of the rotor **12** about the rotor axis **16** to increase above the average angular velocity of the rotor **12** as each pair **32**, **32'**, **32"** of U-shaped magnets **20** passes through the null region **78**. As will be appreciated by those skilled in the art, the rotor sprocket **59** could be circular and the stator sprocket **61** eccentric and still cause the angular velocity of the rotor **12** to increase. Further, the beaded chain **60** in combination with the stator sprocket **61** and the eccentric rotor sprocket **59** are not the only means for connecting each rotor **12** to the stator **50**. For instance, the beaded chain **60** could also be a belt. Further, the linkage **53** could comprise a drive shaft between each rotor **12** and the stator **50**, the drive shaft having a bevel gear set at each end of the shaft mating with a bevel gear on the rotor **12** and the stator **50**. An automatic gear shift mechanism would shift gears as each U-shaped magnet pair **32**, **32'**, **32"** entered the null regions **78** to increase the instantaneous angular velocity of the rotor **12** as the pair **32**, **32'**, **32"** of rotor magnets **20** passed through the null region **78**. Alternatively the linkage **53** could comprise a transmission system employing elliptical gears.

While it is preferred that the instantaneous angular velocity of the rotor **12** to increase above the average angular velocity of the rotor **12** as each pair of U-shaped magnets **20** passes through the null region **78**, it is not necessary to provide the increased angular velocity of the rotor **12** to provide motive power from the motor **10**.

Preferably, the diameters of the rotor sprocket **59** and stator sprocket **61** are selected so that the rear **26** of each U-shaped magnet **20** passes through one and only one null region **78** for each full revolution of the rotor **12** about the respective rotor axis **16** as the armature **70** rotates about the armature axis of rotation **58**. Accordingly, the revolution rate of the armature **70** is related to the revolution rate of the rotor **12** by the expression:

Sa = (Nr / Ns) x Sr ..... (1)

Where:
Sa is the angular velocity of the armature **70** (RPM);
Nr is the number of the U-shaped magnets **20** (or groups of abutted U-shaped magnets **32**) on a rotor **12**;
Ns is the number of null regions **12** on the stator **50**; and
Sr is the angular velocity of the rotor **12** (RPM).

The timing of the rotation of the rotor **12** around its respective rotor axis **16**, and the armature **70** about the armature axis of rotation **58** is such that each U-shaped magnet **20** (or U-shaped magnet pair **32**, **32'**, **32"**) on each rotor **12** enters into a null region **78** at a point where the magnetic interaction between the first magnetic field

and the second magnetic field is substantially reduced, thus providing a commutation of the second magnetic field. As each rotor **12** continues to rotate about the rotor axis **16** and the armature **70** rotates about the armature axis of rotation **58**, the U-shaped magnet **20** traces a slanted path through the null region **78**. As the U-shaped magnet emerges from the null region **78**, the U-shaped magnet **20** encounters the strong first magnetic field, which urges the U-shaped magnet **20** to continue the rotation of the rotor **12** about the rotor axis **16**.

As previously discussed, the first preferred embodiment of the motor **10** comprises a single null region **78** and five rotors **12**, each rotor **12** having three pairs **32**, **32'**, **32"** of abutted U-shaped magnets **20**. Preferably, the rotors **12** are uniformly spaced around the armature axis of rotation **58** and the pairs **32**, **32'**, **32"** of U-shaped magnets **20** are uniformly spaced around the periphery of each respective rotor **12**. Further, the pairs **32**, **32'**, **32"** of U-shaped magnets **20** on each rotor **12** are phased with respect to each other by one-fifth of a revolution of the rotor **12** (i.e. the reciprocal of the number of rotors) so that the pairs **32**, **32'**, **32"** of U-shaped magnets **20** of all the rotors **12** enter the null region at substantially uniform intervals to provide a more or less continuous magnetic interaction between the first magnetic field of the stator **50** and the second magnetic field of the rotors **12**. As will be appreciated by those skilled in the art, the motive power provided by the motor is proportional to the number of rotors **12** and the number of magnets **20** on each rotor **12** as well as the strength of the rotor **12** magnets **20** and the stator **50** magnets **68**. Accordingly, the number of rotors **12** and the number of pairs **32**, **32'**, **32"** of U-shaped magnets **20** are not limited to five rotors **12** and three pairs of U-shaped magnets **32**. Similarly, the number of null regions **78** is not limited to one. The number of U-shaped magnets **20** and the number of null regions **78** are limited only by adherence to the rule established by **Equation (1)**.

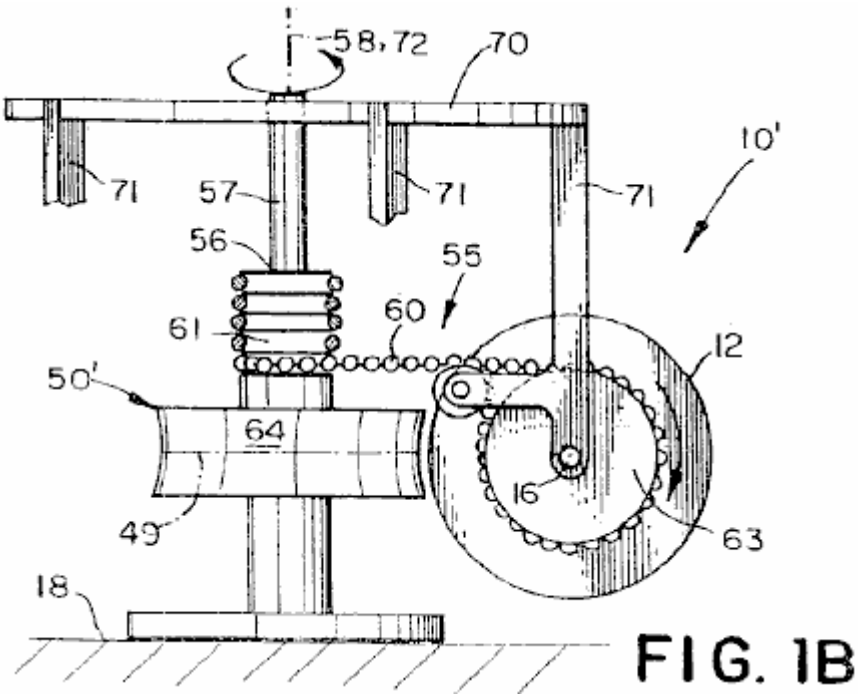


FIG. 1B

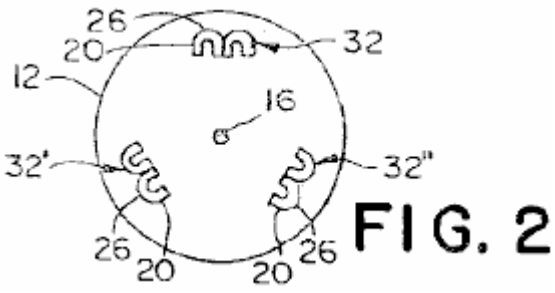


FIG. 2

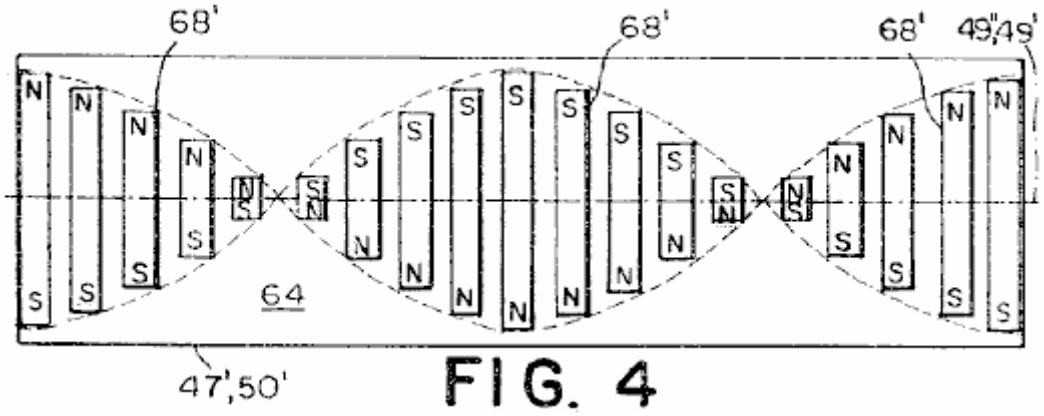


FIG. 4

Referring now to **Fig.1B**, **Fig.2 and** Fig.4 there is shown a second preferred embodiment of a motor **10** providing unidirectional rotational motive power. The second preferred embodiment comprises a generally circular stator **50'** having a stator axis **72** with magnets **68'** attached to a surface **64** of the stator **50'**; an armature **70** attached to the stator **50'** by an armature axle **57** for rotation about an armature axis of rotation **58** coincident with the stator axis **72**; and five rotors **12** (for clarity, only one of which is shown) having three pairs **32**, **32'**, **32''** of abutted U-shaped magnets **20**, the rotors **12** being spaced at intervals of about **72** degrees around the armature **70**. Each rotor **12** is spaced from the armature by a strut **71** and attached to the strut **71** by an axle for rotation in the plane of the armature axis of rotation **58** about a rotor **12** axis of rotation **16**. The motor **10** further includes a driving linkage **55** connecting each rotor **12** and the stator **50** together to cause the armature **70** to rotate about the armature axis of rotation **58** as each rotor **12** rotates about its respective rotor axis **16**.

The second preferred embodiment is identical to the first preferred embodiment except for two differences. First, instead of the first magnetic field being uniform in both magnitude and direction along the circumferential line of demarcation **49** (except in one or more null regions **78** as in the first preferred embodiment), the direction of the first magnetic field rotates about a magnetic axis parallel to the circumferential line of demarcation **49** with a pre-determined periodicity along the line of demarcation **49**. Preferably, the first magnetic field is formed from one or more stator magnets **68'** attached to the outer surface **64** of the stator **50'**, each magnet **68'** having a direction of magnetisation which causes the first magnetic field to rotate about the magnetic axis. In the second preferred embodiment, as shown in **Fig.4**, the stator magnets **68'** are equally sized bar magnets, attached to the stator **50'** so that the bar magnets **68'** spiral on the stator **50'** with the pre-determined periodicity. However, as would be apparent to those skilled in the art, the first magnetic field need not be formed by bar magnets but could be formed from a single magnet (or groups of magnets) such that the direction of magnetisation of the single magnet rotates around the magnetic axis.

The second difference between the first preferred embodiment and the second preferred embodiment is that the linkage **55** of the second preferred embodiment does not include a component for increasing the angular velocity of the rotor **12** above the average velocity of the rotor **12**. Accordingly, in the second preferred embodiment, a circular rotor sprocket **63** is used in place of the eccentric rotor socket **59**, thereby providing a constant rate of rotation of the rotor **12** about the rotor axis **16** as the armature **70** rotates about the stator **50'**.

As will be clear to those skilled in the art, the rotation of the direction of the first magnetic field around the circumferential line of demarcation **49** commutates the second magnetic field, overcoming the need for the null regions **78**. In all other respects, the operation of the second embodiment is the same as that of the first embodiment. That is, the revolution rate of each rotor **12** is related to the revolution rate of the armature **70** by **Equation (1)**, where the parameter **Ns** is the number of rotations around the line of demarcation **49** of the first magnetic field along the line of demarcation **49**. In the second preferred embodiment, as shown in **Fig.4**, the number of rotations of the first magnetic field is one. Accordingly, since there are three pairs **32**, **32'**, **32''** of U-shaped magnets **20**, each of the five rotors **12** makes one-third revolution for each full revolution of the armature **70** around the armature axis **58**. However, as will be appreciated by those skilled in the art, the motor **10** could be designed for the first magnetic field to have any number of whole periods of rotation about the armature axis **58** provided that the revolution rate of the rotors **12** was adjusted to conform to **Equation (1)**.

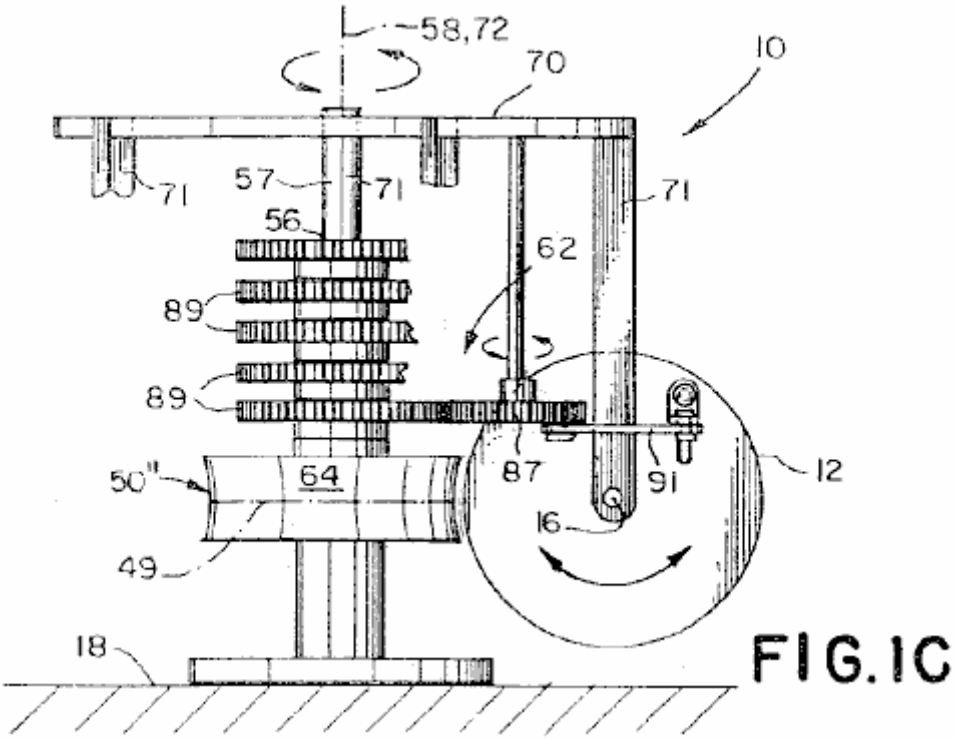


FIG. 1C

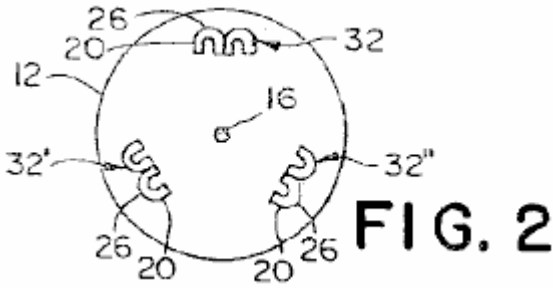


FIG. 2

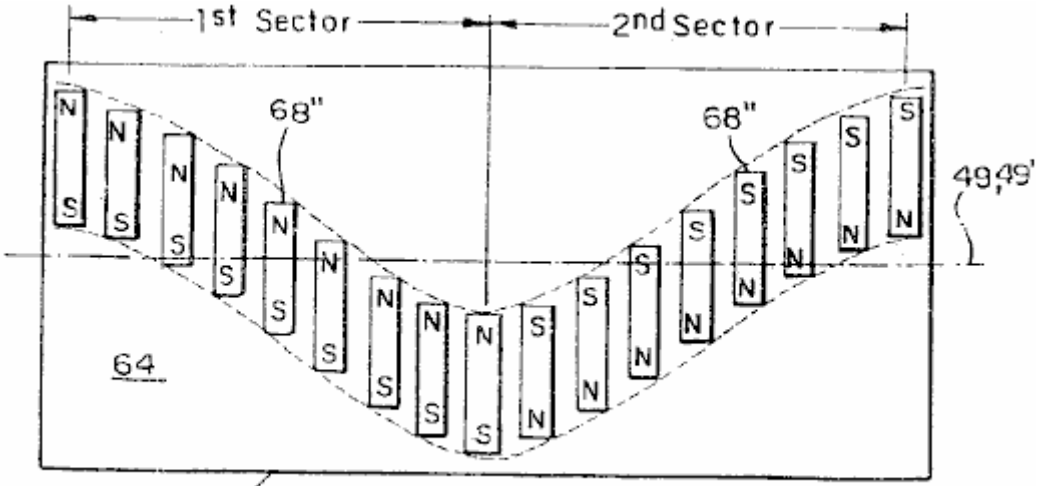


FIG. 5

Referring now to **Fig.1C**, **Fig.2** and **Fig.5** there is shown a third preferred embodiment of a motor **10** providing unidirectional rotational motive power. The third preferred embodiment comprises a generally circular stator **50"** mounted to a base **18** and having an axis **72**, with magnets **68"** attached to the surface **64** of the stator **50"**, an armature **70** attached to the stator **50"** by an axle **57** for rotation about an armature axis of rotation **58** coincident with the stator axis **12**, and five rotors **12** (for clarity, only one of which is shown) having three pairs **32**, **32'**, **32"** of abutted U-shaped magnets **20**, the rotors **12** being spaced at intervals of about **72** degrees around the armature **70**. Each rotor **12** is spaced from the armature by an armature strut **71** and attached to the armature strut **71** by an axle for rotation about an axis **16** of the rotor **12** in a plane generally aligned with the armature axis **58** about

an axis **16** of the rotor **12**. The motor **10** further includes a driving linkage **62** connecting each rotor **12** and the stator **50** together to cause the armature **70** to rotate about the armature axis of rotation **58** as each rotor **12** oscillates about its respective rotor axis **16**.

The third preferred embodiment is identical to the first preferred embodiment except for three differences. First, instead of the first magnetic field being uniform in both magnitude and direction around the circumferential line of demarcation **49** (except in the null zone **78**), the first magnetic field is displaced by a sinusoidal pattern having a pre-determined peak amplitude and a pre-determined period along the circumferential line of demarcation **49**, with the direction of the first magnetic field alternating in opposite directions along the line of demarcation **49** between each peak amplitude of the sinusoidal pattern.

Preferably, as shown in **Fig.5** the first magnetic field is formed by a plurality of bar magnets **68"** arranged on the surface **64** of the stator **50"** so that the magnetisation of the bar magnets **68"** is displaced in the sinusoidal pattern from the line of demarcation **49** around the circumferential line of demarcation **49**. The sinusoidal pattern of the bar magnets **68"** is divided into first and second sectors, the boundary of which occurs at the peaks of the sinusoidal pattern. The direction of magnetisation of the bar magnets **68"** is opposite in direction in the first and the second sectors providing a commutation of the second magnetic field and causing the rotors **12** to reverse in rotational direction as the rotor **12** oscillates around the rotor axis **16** and rotates around the armature axis of rotation **58**.

Preferably, the sinusoidal pattern of the magnets has a predetermined peak amplitude so that each rotor **12** oscillates approximately +/-thirty (30) degrees from a neutral position. However, the value of the peak amplitude is not critical to the design of the motor **10**. Further, the predetermined period of the sinusoidal pattern may be selected to be any value for which the number of cycles of the sinusoidal pattern around the surface **64** of the stator **50"** is an integer value.

As will be apparent to those skilled in the art, the first magnetic field need not be formed by the bar magnets **68"** but could be formed from a single magnet (or groups of magnets) so that the first magnetic field would be sinusoidally displaced around the armature axis of rotation **58** and would alternate in opposite directions between each peak of the sinusoidal pattern. Further, as will be appreciated by those skilled in the art, the displacement of the first magnetic field need not be precisely sinusoidal. For instance the displacement may be in a shape of a sawtooth or in a shape having a portion with constant plus and minus amplitude values, within the spirit and scope of the invention.

As a result of the first magnetic field being sinusoidally displaced and alternating each one-half period, each rotor **12** oscillates through an angle corresponding to approximately the peak amplitude of the sinusoid as the rotor **12** follows the stator magnets **68"**. Accordingly, a second difference between the third embodiment and the first embodiment is in the structure of the linkage **62**. In the third preferred embodiment, shown in **Fig.1C**, the linkage **62** comprises a reciprocating rod **91** connecting each rotor **12** to a respective first gear **87** rotationally attached to the armature **70**. The reciprocating rod **91** is pivotally mounted to each rotor **12** and to each first gear **87** so that the oscillating motion of the rotor **12** is converted to rotary motion of the first gear **87**. Each first gear **87** is coupled to a single second gear **89**, attached to the stator **50** in a fixed position. The rotary motion of each first gear **87** causes the armature **70** to rotate about the armature axis of rotation **58** as the rotors **12** oscillate about the rotor axis **16**. As will be appreciated by those skilled in the art, the speed of the motor **10** is fixed by the ratio of the first gear **87** to the second gear **89** in accordance with the expression:

Sa = (1 / Ns) x Sr ..... (2)

Where:
Ss is the angular velocity of the armature **70** (RPM);
Ns is the number of first magnetic field periods around the stator **50"**; and
Sr is the angular velocity of the rotor **12** (RPM).

Because each rotor **12** oscillates instead of continually rotating, only a single rotor magnet. (or group of magnets) on a given rotor **12** interacts with the single stator **50"**. Accordingly, a third difference between the third preferred embodiment and the first preferred embodiment arises because of the oscillatory motion of each rotor **12** whereby each rotor **12** of the third preferred embodiment has only a single pair of magnets **32**. However, as will be appreciated by those skilled in the art, additional stators **50"** may be added around the periphery of the rotors **12** and additional pairs of U-shaped magnets **20** may be included on each rotor 12 to interact magnetically with each additional stator **50"**, thus providing additional motive power.

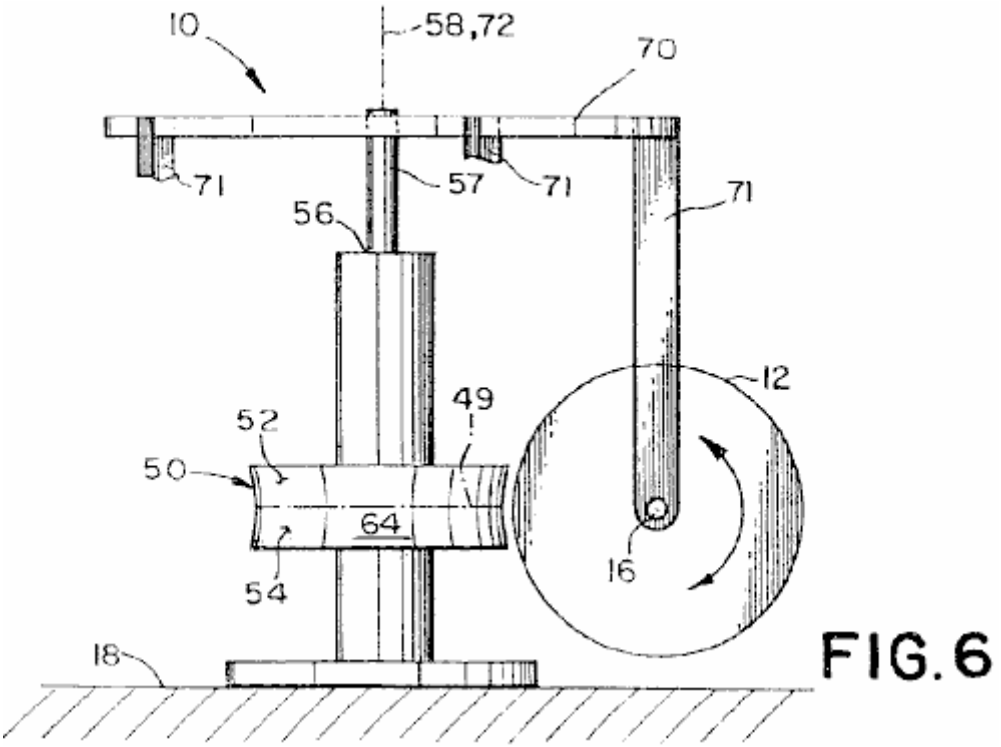


FIG. 6

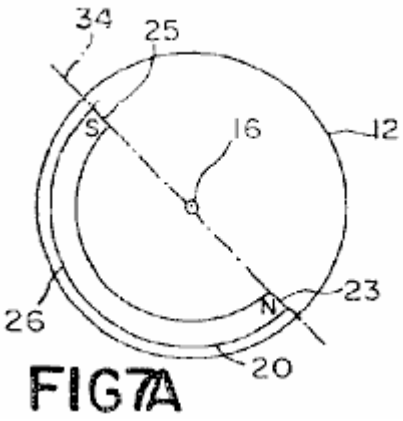


FIG 7A

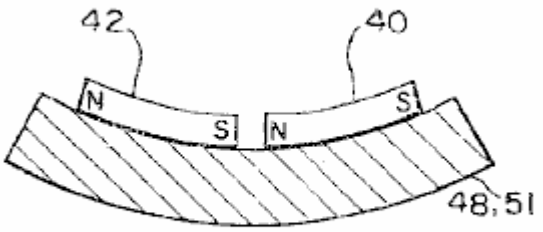


FIG. 8B

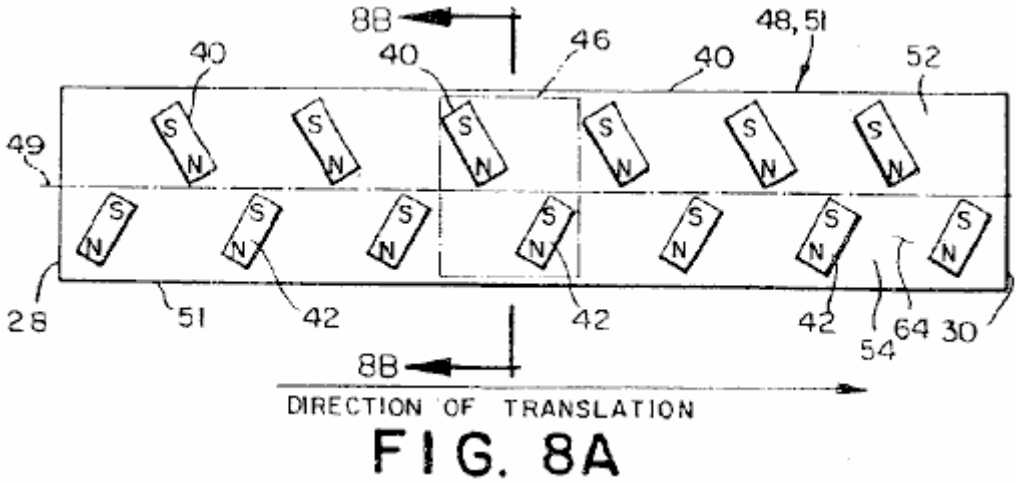


FIG. 8A

Referring now to **Figs. 6, 7A, 8A and 8B**, there is shown a fourth preferred embodiment of the permanent magnet motor **10** for providing unidirectional rotational motive power. The fourth preferred embodiment comprises a generally circular stator **51** having a stator axis **72**, attached to a base **18**. The stator **51** includes an outer surface **64** divided into a first side **52** and a second side **54** by a circumferential line of demarcation **49**, having a pre-determined direction around the stator axis **72**, at about a midpoint of the outer surface **64**.

Preferably, the surface **64** of the stator **51** is curved, having a curvature conforming to the arc of the rotors **12**. However, it will be appreciated by those skilled in the art that the surface **64** need not be curved but could be planar and still be within the spirit and scope of the invention. As will be appreciated by those skilled in the art the stator **51** is merely intended as a stationary supporting structure for stator magnets and, as such, the shape of the stator is not intended to be controlling of the size and shape of the air gap between the magnets attached to the stator and the magnets attached to the rotors.

As shown in **Fig.8A**, one or more pairs of stator magnets **46** are attached to the outer surface **64** spaced along the line of demarcation **49**. Each pair of stator magnets **46** comprises a first stator magnet **40** having a north pole and a south pole and a second stator magnet **42** having a north pole and a south pole. The south pole of each first stator magnet **40**, is located on the first side **52** of the outer surface **64**, and the north pole of the first stator magnet **40** is closest to the line of demarcation **49**. The north pole of each second stator magnet **42** is located on the second side **54** of the outer surface **64** and the south pole of each second stator magnet **42** being closest to the line of demarcation **49**. The first and the second stator magnets **40**, **42** are spaced along the line of demarcation **49** so that a first inter-magnet distance measured along the line of demarcation **49** between the north pole of the first stator magnet **40** and the south pole of the second stator magnet **42** of an adjacent pair of magnets **46** is generally equal to a second inter-magnet distance measured along the line of demarcation **49'** between the south pole of the first stator magnet **40** and the north pole of the second stator magnet **42**.

In the fourth preferred embodiment, the stator magnets **40**, **42** are bar magnets. Preferably, the north pole of each first stator magnet **40** and the south pole of each second stator magnet **42** are inclined toward the pre-determined direction. Also, the bar magnets are preferably oriented on the surface **64** of the stator **50** so that the south pole of each first magnet **40** and the north pole of each second magnet **42** are nearer to the periphery of each rotor **12** than the opposite polarity pole of each of the magnets **40**, **42**. As will be appreciated by those skilled in the art, the stator magnets **40**, **42** need not be bar magnets. For instance, each stator magnet **40**, **42** could be a U-shaped magnet, or could be made up of separate magnets, as long as the first magnetic field generated by the magnets was generally equivalent to that produced by the bar magnets.

In the fourth preferred embodiment, an armature **70** having an armature axis of rotation **58** coincident with the stator axis **72** is attached to the stator **51** by an armature axle **57**, which armature axle **57** allowing the armature **70** to freely rotate about the stator axis **72**. Each rotor **12** is spaced from the armature **70** by an armature strut **71** and is mounted to the armature strut **71** so as to be free to rotate about the rotor axis **16**. The rotor axis **16** is oriented so that the rotor **12** rotates in a plane generally aligned with the armature axis of rotation **58**. In the fourth preferred embodiment, five rotors **12** are attached to the armature **70**. Preferably, the rotors **12** are uniformly spaced around the circumference of the stator **50** with a spacing of the rotors **12** as measured at the surface **64** of the stator **51** about equal to an integer multiple of twice the inter-magnet distance. However, as those skilled in the art will appreciate, it is not necessary to have the rotors **12** uniformly spaced. Further, the number of rotors **12** can be as few as one and as large as size and space constraints allow. As will be appreciated by those skilled in the art, the stator axis **72** need not be coincident with the armature axis of rotation **58**. Accordingly, a stator **50** arranged around the armature axis **58** at any location at which the stator axis **72** is parallel to the armature axis **58** and the surface of the stator **50** faces the periphery of the rotors **12**, thereby providing for the interaction between the first magnetic field and the second magnetic field around the armature axis **58**, is within the spirit and scope of the invention.

Referring now to **Fig.7A**, each rotor **12** comprises a first U-shaped magnet **20** generating a second magnetic field. The first U-shaped magnet **20** is positioned on the rotor **12** so that the north pole and the south pole of the first U-shaped magnet **20** faces toward the axis **16** of the rotor **12**, and the rear side **26** of the first U-shaped magnet **20** faces the periphery of the rotor **12**. When the rear **26** of the first U-shaped magnet **20** is adjacent to the north pole of one of the first stator magnets **40** along the line of demarcation **49**, a portion of the second magnetic field directly adjacent to the rear **26** of the first U-shaped magnet **20** interacts with a portion of the first magnetic field generated by the north pole of the first stator magnet **40** to cause the rotor **12** to rotate in a counterclockwise direction. As the rotor **12** rotates in the counterclockwise direction, a portion of the second magnetic field associated with the south pole of the first U-shaped magnet **20** interacts with a portion of the first magnetic field associated with the south pole of the first stator magnet **40**, giving rise to a force in the direction of the rotor axis **16**, repelling the U-shaped magnet **20**, and causing the rotor **12** to translate in the pre-determined direction around the stator axis. As the rotor **12** moves away from first stator magnet **40** in the pre-direction the second magnetic field adjacent to the rear **26** of the U-shaped magnet **20** interacts with the portion of the first magnetic field associated with the south pole of the second stator magnet **42** of the pair of magnets **46**, causing the rotor **12** to reverse direction and rotate in the clockwise direction. The portion of the second magnetic field associated with the north pole of the U-shaped magnet **20** then interacts with the portion of the first magnetic field associated with the north pole of the second stator magnet **42**, again giving rise to a force in the direction of the rotor axis **16**, repelling the U-shaped magnet **20** and causing the rotor **12** to translate in the pre-determined direction. An

oscillation cycle is then repeated with the second magnetic field of the rotor **12** interacting with the first magnetic field of the adjacent pair of magnets **46**. Accordingly, the rotor **12** rotationally oscillates about the respective rotor axis **16** and generates a force in the direction of the rotor axis **16**, causing the armature **70** to rotate in the pre-determined direction around the armature axis of rotation **58** to provide the unidirectional rotational motive power of the motor. As would be appreciated by those skilled in the art, the fourth embodiment is not limited to a single stator **51** and a single U-shaped magnet **20**. Additional stators having first and second stator magnets **40**, **42** arranged identically to the stator **51** to interact with corresponding U-shaped magnets spaced around the periphery of each rotor are with in the spirit and scope of the invention.

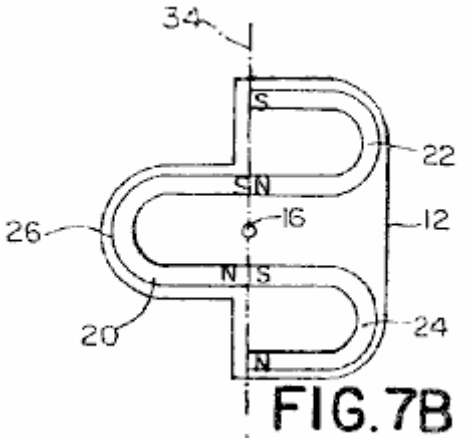


FIG.7B

Referring now to **Fig.6**, **Fig.7B** and **Fig.8A** there is shown a fifth preferred embodiment of the permanent magnet motor **10** for providing unidirectional rotary motive force. The structure and operation of the fifth preferred embodiment is similar to that of the fourth preferred embodiment except that each rotor **12** further includes a second U-shaped magnet **24** having a north pole and a south pole with the south pole of the second U-shaped magnet **24** abutting the north pole of the first U-shaped magnet **20**, and a third U-shaped magnet **22** having a north pole and a south pole, with the north pole of the third U-shaped magnet **22** abutting the south pole of the first U-shaped magnet **20**. As the rotor **12** rotates due to interaction of the portion of the second magnetic field adjacent to the rear of the U-shaped magnet **20** with the first magnetic field, a third magnetic field generated by the north pole of the second U-shaped magnet **24** and a fourth magnetic field generated by the south pole of the third U-shaped magnet **22** each interact with the first magnetic field generated by each stator magnet pair **46** to cause each rotor **12** to generate a force in the direction of the rotor axis **16**, thereby causing the armature **70** to rotate in the pre-determined direction around the axis **58** of the stator **51** to provide the unidirectional rotational motive power of the motor.

In the fifth preferred embodiment, the portion of the second magnetic field adjacent to the rear **26** of the first U-shaped magnet **20** serves to rotate the rotor **12** while the second and third U-shaped magnets **24**, **22** generate the magnetic fields providing the force in the direction of the rotor axis **16**. Accordingly, the fifth preferred embodiment is potentially more powerful than the fourth preferred embodiment. As will be appreciated by those skilled in the art, the stator magnets **40**, **42** need not be bar magnets. For instance, each stator magnet **40**, **42** could be replaced by a U-shaped magnet or could be made up of separate magnets, as long as the first magnetic field generated by the magnets was generally equivalent to that produced by the bar magnets.

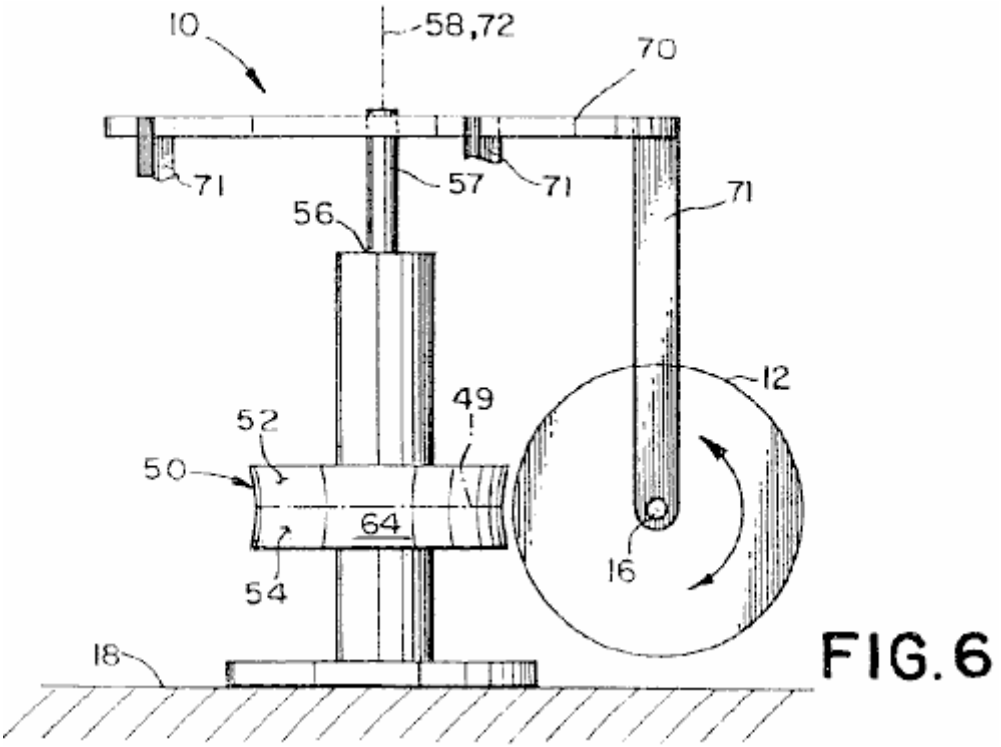

10
58,72
70
71
57
71
71
56
12
49
52
50
64
54
16
18

FIG.6

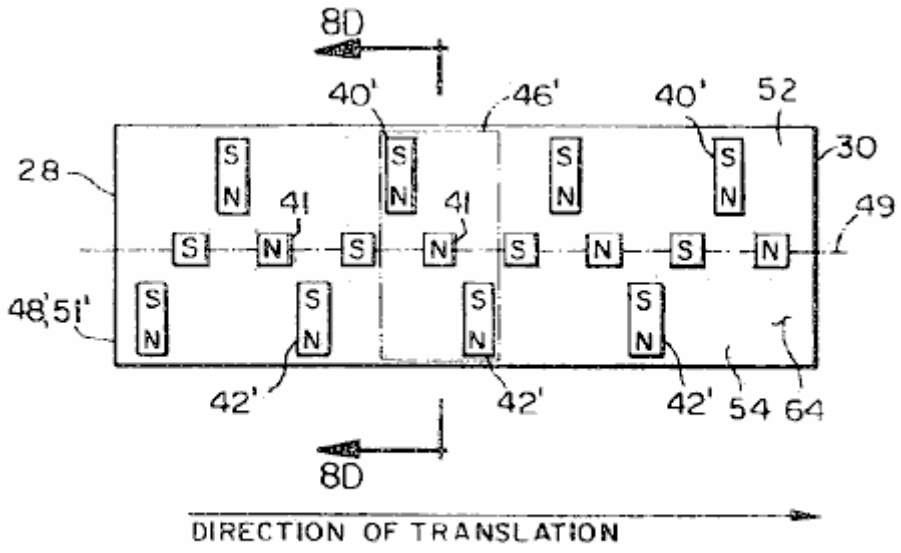

8D
40'
46'
40'
52
30
28
49
41
41
48'51'
42'
42'
42'
54
64
8D
DIRECTION OF TRANSLATION

FIG. 8C

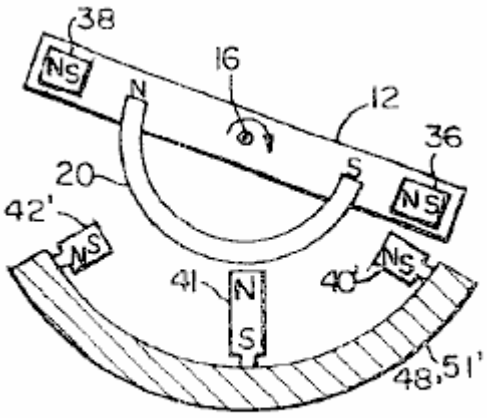


FIG. 8D

Referring now to **Fig.6** and **Fig.8C** and **Fig.8D** there is shown a sixth preferred embodiment of the motor **10**. The structure and operation of the sixth preferred embodiment is identical to that of the fifth preferred embodiment except that:
(1) The stator magnets **40'**, **42'** on the surface **64** of the stator **51'** are in a slightly different orientation;
(2) an additional stator magnet **41** is added to each pair of stator magnets **46** and
(3) the U-shaped magnets **22**, **24** attached to each rotor **12** are replaced with bar magnets **36**, **38**.

Specifically, and referring now to **Fig.8C**, the direction of magnetisation of each first stator magnet **40'** and each second stator magnet **42'** is aligned to be generally perpendicular to the line of demarcation **49** instead of being inclined in the pre-determined direction around the armature axis of rotation **58** as in the fifth embodiment. Also, the stator **51'** also includes a third stator magnet **41** mounted on the outer surface **64** along the line of demarcation **49** mid-way between each first stator magnet **40'** and each second stator magnet **42'**. As shown in **Fig.8C** and **Fig.8D**, the third stator magnet **41** is oriented so that the direction of magnetisation of the third magnet **41** is aligned with the axis **16** of the rotors **12**.

As shown in **Fig.8C** and **Fig.8D**, the rotor **12** used in the sixth preferred embodiment includes a first U-shaped magnet **20**, similar to that of the fifth preferred embodiment. However, in place of the second and the third U-shaped magnets **24**, **22** used in the fifth preferred embodiments, the sixth preferred embodiment includes a first thruster bar magnet **36**, spaced from and proximate to the south pole of the first U-shaped magnet **20** and generally aligned with a thruster magnet axis **34**, and a second thruster bar magnet **38**, spaced from and proximate to the north pole of the first U-shaped magnet **20** and also generally aligned with the thruster magnet axis **34**. The thruster axis **34** lies in the plane of the rotor **12** and intersects the rotor axis **16**. Similar to the fifth preferred embodiment, the interaction of the portion of the second magnetic field directly adjacent to the rear of the U-shaped magnet **20** with the first magnetic field provides the rotational force for the rotors **12**. As the rotor **12** rotates in the clockwise direction (viewed from the second end **30** of the stator **51'**), a third magnetic field generated by both the north pole and the south pole of the second thruster magnet **36** interacts with the first stator magnet **40'**, again generating a force in the direction of the rotor axis **16**. Similarly, when the rotor **12** rotates in the counterclockwise direction a fourth magnetic field generated by both the north pole and the south pole of the first thruster magnet **38** interacts with second stator magnet **42'**, generating a force in the direction of the rotor axis **16**. The result of the force in the direction of the rotor axis **16** is to cause the armature **70** to rotate in the predetermined direction around the armature axis of rotation **58** to provide the unidirectional rotational motive power of the motor **10**.

In the sixth preferred embodiment, the stator magnets **40'**, **41**, **42'** and the thruster magnets **36**, **38** are bar magnets. However, as will be appreciated by those skilled in the art, the stator magnets **40'**, **41 42'** and the thruster magnets **36**, **38** need not be bar magnets. For instance, each stator magnet **40'**, **42'** could be a U-shaped magnet or could be made up of separate magnets, as long as the first magnetic field generated by the magnets was generally equivalent to that produced by the bar magnets.

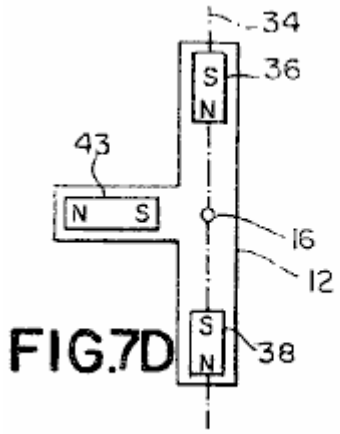


FIG.7D

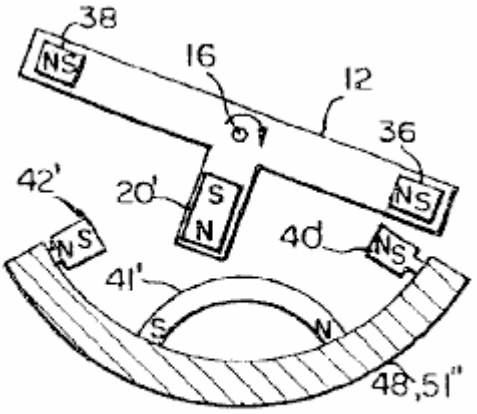


FIG. 8E

Referring now to **Fig.6**, **Fig.7D** and **Fig.8E** there is shown a seventh preferred embodiment of the motor **10**. The structure and operation of the seventh preferred embodiment is similar to the sixth preferred embodiment except that the third stator magnet **41'** located on the surface **64** of the stator **51"** along the line of demarcation **49** is a U-shaped magnet **41'** with the rear of the U-shaped magnet **41'** facing the rotor **12** and the direction of magnetisation being perpendicular to the line of demarcation **49**; and the U-shaped magnet **20** is replaced with a bar magnet **20'** oriented to have the direction of magnetisation aligned with a radial line of the rotor **12**. As in the sixth preferred embodiment, each stator magnet **40'**, **42'** could be a U-shaped magnet or could be made up of separate magnets, as long as the first magnetic field generated by the stator magnets **40'**, **42'** was generally equivalent to that produced by the bar magnets.

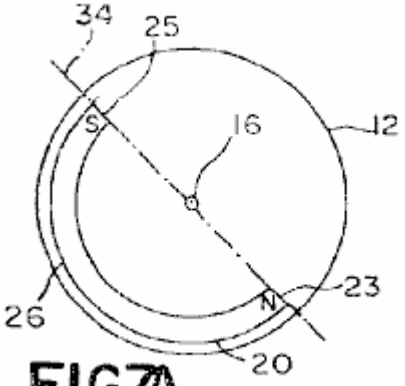


FIG7A

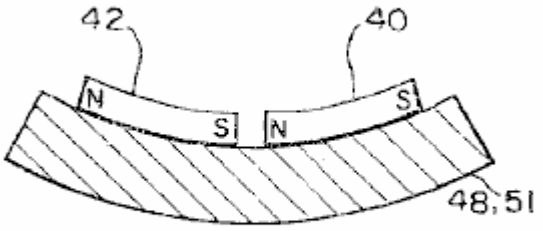


FIG. 8B

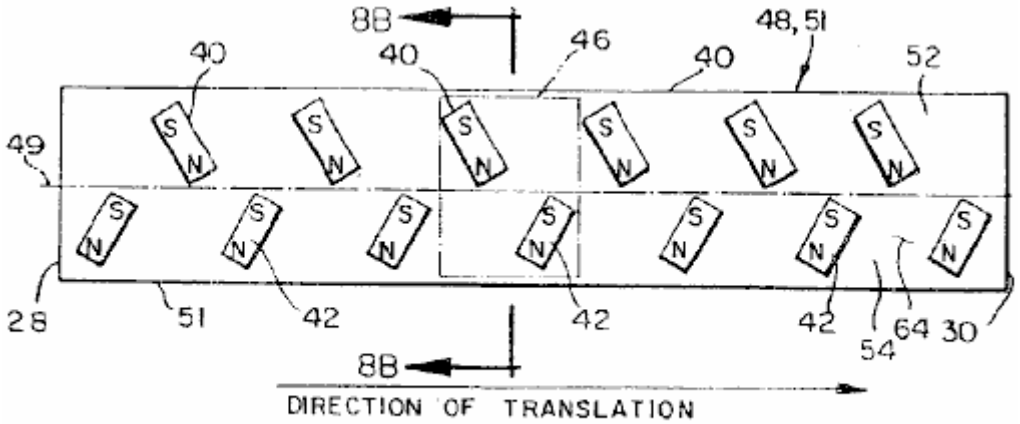


FIG. 8A

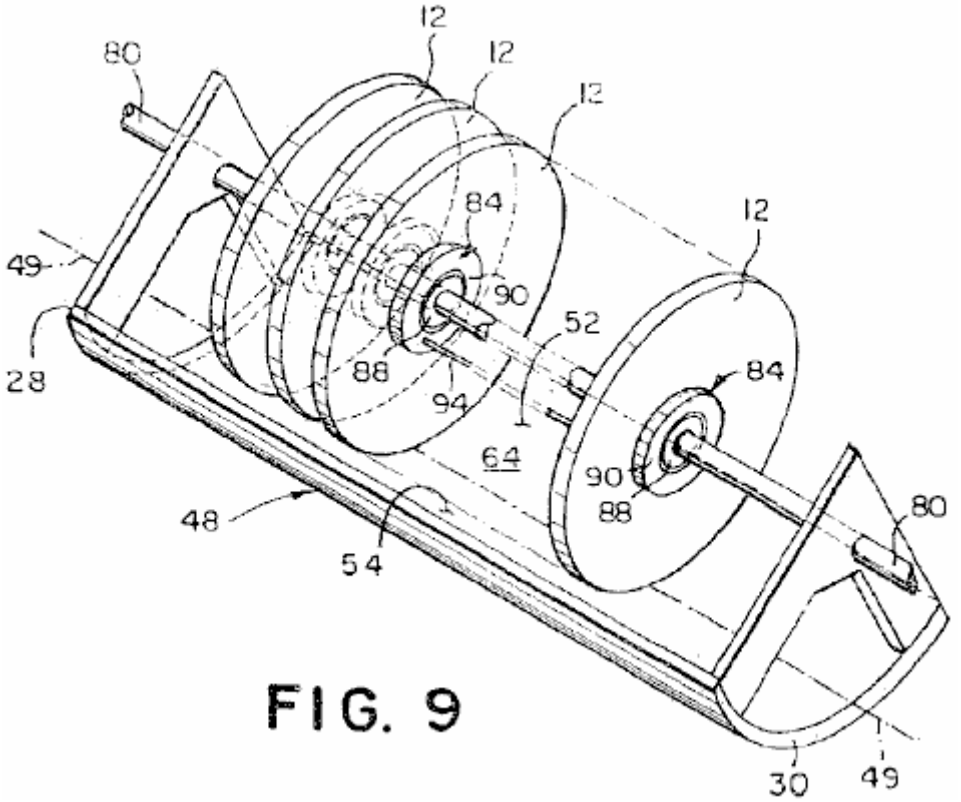


FIG. 9

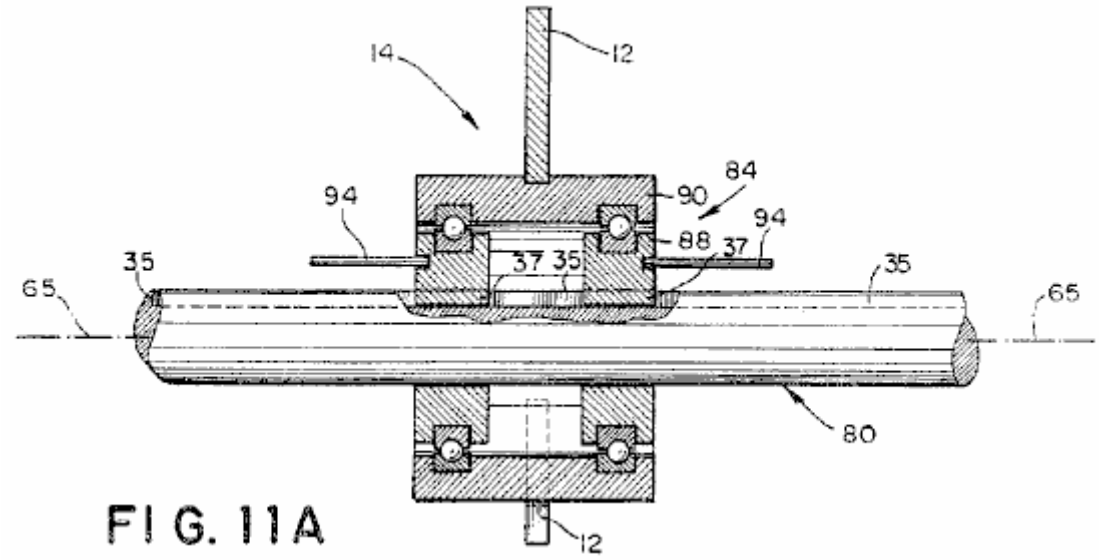


FIG. 11A

Referring now to **Fig.7A**, **Fig.8A**, **Fig.8B**, **Fig.9** and **Fig.11A**, there is shown an eighth preferred embodiment of the motor **10** for providing unidirectional linear motive power. The eighth preferred embodiment comprises a linear stator **48** having a generally curved cross-section perpendicular to a longitudinal line of demarcation **49** extending on a surface **64** of the stator between a first end **28** and a second end **30** and dividing the surface **64** of the stator **48** into a first side **52** and a second side **54**. Preferably, the generally curved cross-section of the stator **48** is concave. However, it will be appreciated by those skilled in the art that the cross-section need not be concave but could be planar or even convex and still be within the spirit and scope of the invention.

The linear stator **48** is identical to the generally circular stator **51** except for the surface **64** of the stator **48** being linear in the direction of the line of demarcation **49** instead of being circular in the direction of the line of demarcation **49**.

The eighth preferred embodiment includes the first and the second stator magnets **40**, **42** (see **Fig.8A**), the location and orientation of which are virtually identical to the orientation and location of the stator magnets **40**, **42** on the circular stator **51**. Accordingly, attached to the linear stator **48** is one or more pairs of magnets **46**, each pair of stator magnets **46** generating a first magnetic field and comprising a first stator magnet **40** having a north pole and a south pole and a second stator magnet **42** having a north pole and a south pole. The south pole of

each first stator magnet **40**, is located on the first side **52** of the outer surface **64**, with the north pole of the first stator magnet **40** being closest to the line of demarcation **49**. The north pole of each second stator magnet **42** is located on the second side **54** of the outer surface **64** with the south pole of each second stator magnet **42** being closest to the line of demarcation **49**. The first and the second stator magnets **40**, **42** are spaced along the line of demarcation **49** so that a first inter-magnet distance measured along the line of demarcation **49** between the north pole of the first stator magnet **40** and the south pole of the second stator magnet **42** of an adjacent pair of magnets **46** is generally equal to a second inter-magnet distance measured along the line of demarcation **49** between the south pole of the first stator magnet **40** and the north pole of the second stator magnet **42**.

In the eighth preferred embodiment, the stator magnets **40**, **42** are bar magnets, the north pole of each first stator magnet **40** and the south pole of each second stator magnet **42** being inclined toward the second end **30** of the linear stator **48**. Also, as shown in **Fig.8A**, the stator magnets **40**, **42** are oriented on the surface **64** of the stator **51** so that the south pole of each first magnet **40** and the north pole of each second magnet **42** are nearer to the periphery of each rotor **12** than the opposite polarity pole of each of the stator magnets **40**, **42**. As will be appreciated by those skilled in the art, the stator magnets **40**, **42** need not be bar magnets. For instance, each stator magnet **40**, **42** could be a U-shaped magnet or could be made up of separate magnets, as long as the first magnetic field generated by the magnets was generally equivalent to that produced by the bar magnets.

The eighth preferred embodiment also includes rail **80** having a longitudinal axis located generally parallel to the line of demarcation **49** of the stator **48**. Five rotor assemblies **14** comprising a rotor **12** and a bearing assembly **84** are slidably attached to the rail **80**.

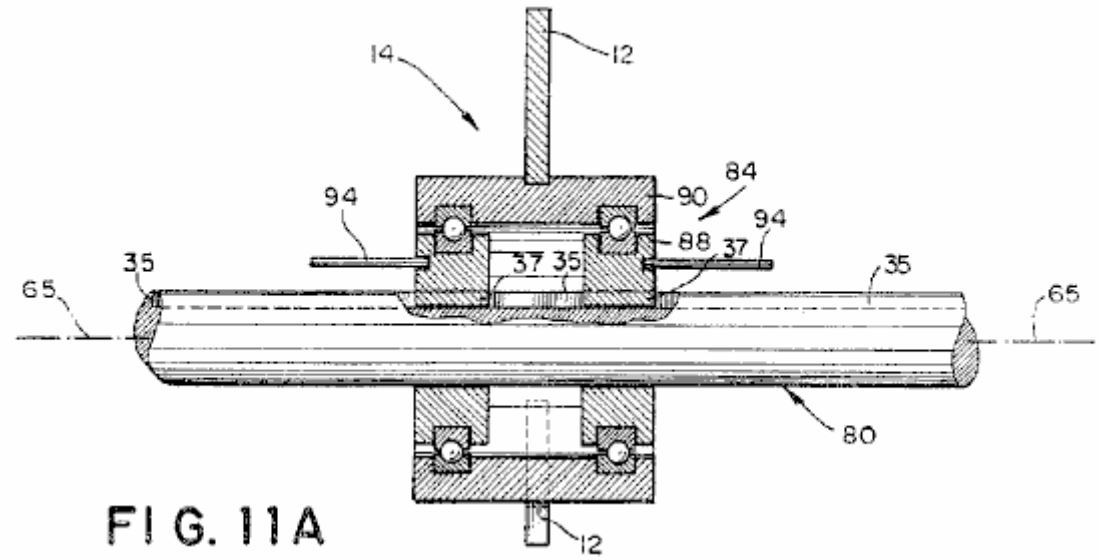


FIG. 11A

Preferably, the bearing assembly **84**, as shown in **Fig.11A**, includes a pair of first bearings **88** slidably mounted to the rail **80** and constrained to slide along the rail without any substantial rotation, by a boss **37** in each first bearing **88**, which is keyed to a longitudinal groove **35** on the rail **80**. A second bearing **90** is connected for rotation to the pair of first bearings **88** by ball bearings. The rotor **12** is attached to the second bearing **90**. Thus, the rotor **12** attached to each bearing assembly **84** is free to oscillate rotationally about the rail **80** and to generate a force along the rail **80** in the direction of the second end of the stator **30**.

Preferably, the eighth preferred embodiment includes a cross-link **94** which ties each bearing assembly **84** together by connecting together the first bearings **88** of each bearing assembly **84**, thereby adding together the linear motion along the rail **80** of each rotor **12**.

Preferably, each rotor **12** comprises one or more one rotor magnets **20**, each rotor magnet **20** generating a second magnetic field which interacts with the first magnetic field to cause the rotor **12** to oscillate rotationally about the axis of the rail **80** and to generate a force in the direction of the axis of the rail **80** to provide the unidirectional linear motive power of the motor. In the eighth preferred embodiment, each rotor **12** is substantially identical to the rotor **12** described for the fourth preferred embodiment. Accordingly, each rotor magnet comprises a first U-shaped magnet **20** having a north pole, a south pole and a rear side **26**, a first portion of the second magnetic field directly adjacent to the rear **26** of the U-shaped magnet **20** interacting with each first magnetic field to cause each rotor **12** to oscillate rotationally about the rail **80**. A second portion of the second magnetic field adjacent to the north and the south poles of the first U-shaped magnet **20** interacts with the first magnetic field to cause the rotor **12** to generate a force in the direction of the axis of the rail **80** thereby providing the unidirectional linear motive power of the motor. As would be clear to those skilled in the art, the operation of the eighth

preferred embodiment is identical to that of the fourth preferred embodiment except that the motion of the cross-linked rotors **12** is linear along the rail **80** instead of being rotational about the armature axis of rotation **58**. Accordingly, for the sake of brevity, a description of the operation of the eighth preferred embodiment is not repeated.

Referring now to **Fig.7B**, **Fig.8A**, **Fig.8B**, **Fig.9** and **Fig.11A** there is shown a ninth preferred embodiment of the motor **10** for providing unidirectional linear motive power. As would be apparent to those skilled in the art, the structure and the operation of the ninth preferred embodiment is virtually identical to that of the fifth preferred embodiment except that the motion of the cross-linked rotors **12** is linear instead of rotational about the armature axis of rotation **58**. Accordingly, for the sake of brevity, a description of the structure and the operation of the ninth preferred embodiment is not repeated.

Referring now to **Figs. 7C, 8C, 8D, 9 and 11A** there is shown a tenth preferred embodiment of the motor **10** for providing unidirectional linear motive power. As would be apparent to those skilled in the art, the structure and the operation of the tenth preferred embodiment is virtually identical to that of the sixth preferred embodiment except that the motion of the cross-linked rotors **12** is linear instead of rotational about the armature axis of rotation **58**. Accordingly, for the sake of brevity, the operation of the tenth preferred embodiment is not repeated.

Referring now to **Figs. 7D, 8C, 8E, 9 and 11A** there is shown an eleventh preferred embodiment of the motor **10** for providing unidirectional linear motive power. The structure and operation of the eleventh preferred embodiment is virtually identical to the seventh preferred embodiment except that the motion of the cross-lined rotors **12** is linear instead of rotational about the armature axis of rotation **58**. Accordingly, for the sake of brevity, the operation of the tenth preferred embodiment is not repeated.

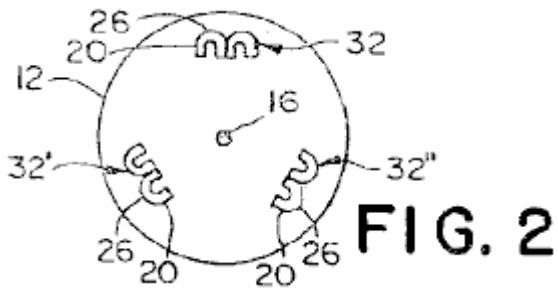


FIG. 2

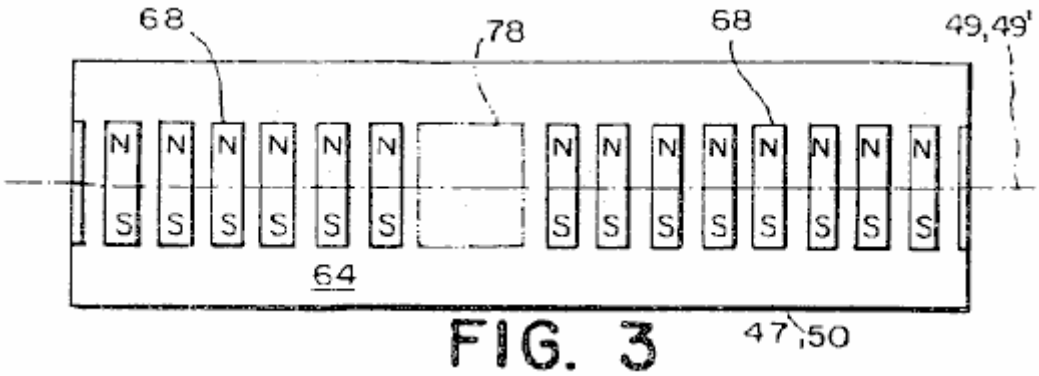


FIG. 3

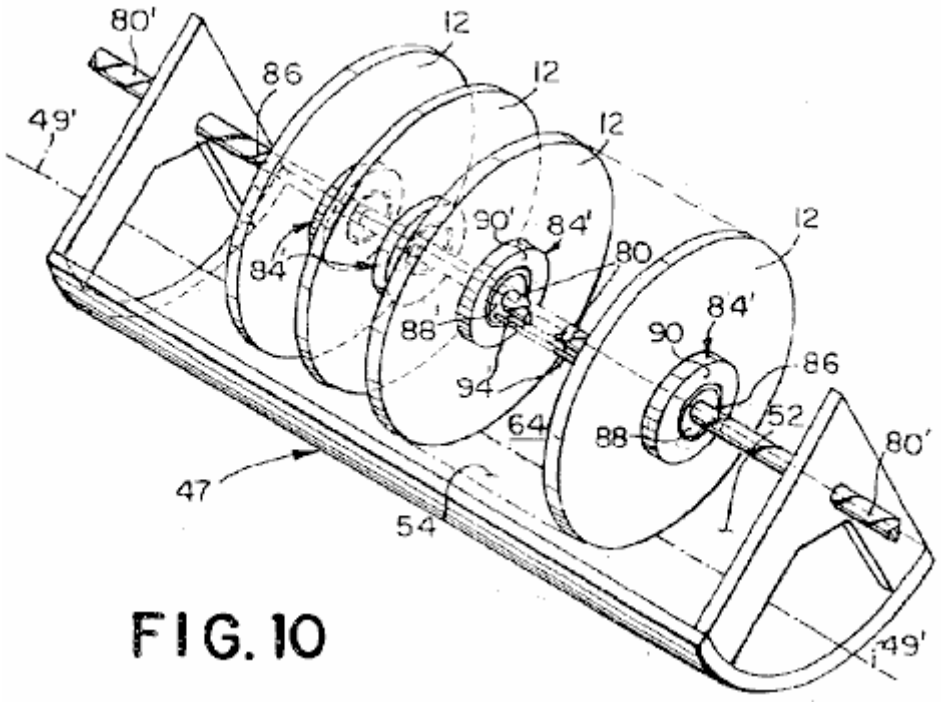


FIG. 10

Referring now to **Fig.2**, **Fig.3**, **Fig.10** and **Fig.11B**, there is shown a twelfth preferred embodiment of the motor **10** for providing linear motive power. As shown in **Fig.10**, the twelfth preferred embodiment comprises a linear stator **47** having a generally curved cross-section perpendicular to a line of demarcation **49'** extending along a midpoint of the stator **47** between a first end **28** and a second end **30** of the linear stator **47**, a rail **80'** connected to the linear stator **47** having an axis generally parallel to the line of demarcation **49'**, one or more rotor assemblies **14'** comprising rotors **12** connected to the rail **80'** by a bearing assembly **84'**, and a cross-link **94'** connecting together the linkages **84'** of adjacent rotors **12**. Preferably, the generally curved cross section of the stator **47** is concave, having a curvature conforming to the arc of the rotors **12**. However, it will be appreciated by those skilled in the art that the generally curved cross-section need not be concave but could be planar or even convex and still be within the spirit and scope of the invention.

As shown in **Fig.3**, the linear stator **47** includes one or more magnets **68** arranged on the surface **64** of the linear stator **47**, each magnet **68** having a direction of magnetisation directed at about a right angle to the line of demarcation **49'** and resulting in a first magnetic field directed generally at a right angle to the line of demarcation **49'**. The magnitude of the first magnetic field is generally uniform except in the null region **78**, in which the magnitude of the first magnetic field is substantially reduced. The linear stator **47** of the twelfth preferred embodiment is virtually identical to the circular stator **50** of the first preferred embodiment except the linear stator **50** is linear in the direction of the line of demarcation **49'** instead of being circular around the armature axis of rotation **58**. Also, the arrangement of the magnets **68** on the surface **64** of the stator **47** and the structure of the null region(s) **78** is the same as for the first preferred embodiment, as shown in **Fig.3** and as fully described in the discussion of the first embodiment. Accordingly, for the sake of brevity, a more detailed description of the structure of the linear stator **47** is not repeated.

The rotors **12** of the twelfth preferred embodiment each have an axis of rotation **16** which is aligned with an axis of the rail **80'**. The rotors **12** are connected to the rail **80'** by the bearing assembly **84'** so that each rotor **12** is free to rotate about the rail **80'** and to slide along the rail **80'.** Preferably, as shown in **Fig.2**, each rotor **12** includes three pairs of U-shaped magnets **32, 32, 32'**, each U-shaped magnet having a rear side **26** and generating a second magnetic field. A portion of the second magnetic field adjacent to the rear-side **26** of each U-shaped magnet **20** interacts with the first magnetic field to cause each rotor **12** to rotate about the axis of the rail **80**. The rotors **12** of the twelfth preferred embodiment are the same as the rotors in the first preferred embodiment, as described in **Fig.2** and fully discussed above. Accordingly, for the sake of brevity, the detailed description of the rotors **12** is not repeated.

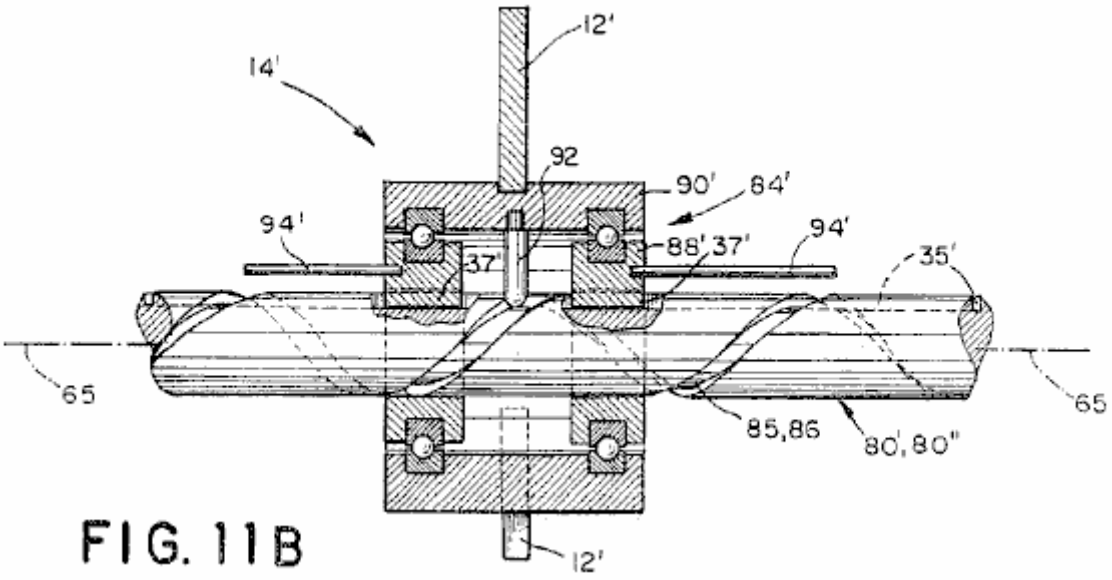


FIG. 11B

As shown in **Fig.11B**, the rail **80'** has a helical groove **86** with a pre-determined pitch (i.e., turns/unit length) running around a periphery of the rail **80'**. The bearing assembly **84'** connects each rotor **12** to the helical groove **86**, converting the rotational motion of each rotor **12** around the rail **80'** to the linear motion along the rail **80'**. As shown in **Fig.11B**, the bearing assembly **84'** comprises a pair of first bearings **88'** mounted to the rail **80'** and constrained to slide along the rail **80'** without any substantial rotation, and a second bearing **90'**, mounted to an outer surface the first bearing **88'** for receiving the rotor **12**. Preferably, each first bearing **88'** has a boss **37** which engages a longitudinal groove **35** so that each first bearing **88'** slides on the rail **80'** without rotation as the second bearing **90'** rotates on the first bearings **88'**. It will be appreciated by those skilled in the art, other methods for securing the first bearings **88'** to the rail **80'** could be employed, as for instance, by making the cross-section of the rail **80'** oblate (flattened at the poles). As in the first preferred embodiment, each rotor **12** must rotate at a rate which results in the rear of each U-shaped magnet **20** on the rotor **12** passing through one of the null regions **78** each full rotation of the rotor **12**. Accordingly, the pre-determined pitch of the helical groove **86** on the rail **80'** preferably equals:

Pg = (1 / Nr) x Pr ..... (3)

Where:
Pr = the pitch of the null regions **78** (null regions/unit length);
Nr = the number of U-shaped magnets (or groups of abutted U-shaped magnets) on a rotor **12**; and
Pg = the pitch of the helical groove **86** (revolutions/unit length).

Preferably, the portions of the helical groove **86** corresponding to each null region **78** have an instantaneous pitch which is greater than the pre-determined pitch of the groove **86** for increasing the angular velocity of the each rotor **12** as each one of the pairs **32, 32', 32"** of U-shaped magnets **20** passes through one of the null regions **78**. However, as will be appreciated by those skilled in the art, it is not necessary to provide the greater instantaneous pitch in order for the motor **10** to provide motive power.

As described above, the cross-link **94'** connects the bearing assembly **84'** of adjacent rotors **12** together. As shown in **Fig.10**, the cross-link **94'** connects the first bearings **88'** of each bearing assembly **84'** to the first bearing **88'** of the adjacent bearing assemblies **84'** so that the linear motion of all the rotor assemblies **14'** are added together to provide the unidirectional linear motive power of the motor **10**.

As previously stated, the first preferred embodiment of the motor **10** comprises a single null region **78** and five rotors **12**, each rotor **12** having three pairs **32, 32', 32"** of abutted U-shaped magnets **20**. Preferably, the rotors **12** are uniformly spaced along the rail **80'** and the pairs **32, 32', 32"** of U-shaped magnets **20** are uniformly spaced around the periphery of each respective rotor **12**. Further, the pairs **32, 32', 32"** of U-shaped magnets **20** are phased with respect to each rotor **12** by one-fifth of a revolution of the rotor **12** so that the pairs **32, 32', 32"** of U-shaped magnets **20** of all the rotors **12** pass through the null region **78** at a substantially uniform rate to provide a more or less continuous interaction between the first magnetic field and the second magnetic field of the rotors **12**, resulting in a more or less continuous urging of the rotor assemblies **14'** toward the second end of the stator **47**. As will be appreciated by those skilled in the art, the motive power provided by the motor **10** is proportional to the number of rotors **12** and the number of U-shaped magnets **20** on each rotor **12**. Accordingly, the number of rotors **12** and the number of pairs **32, 32', 32"** of magnets **20** of the present invention are not limited to five rotors

12 and three pairs **32** of U-shaped magnets **20**. Neither is the number of null regions limited to one. The number of U-shaped magnets **20** and null regions **78** are limited only by adherence to the rule established by **Equation 3**.

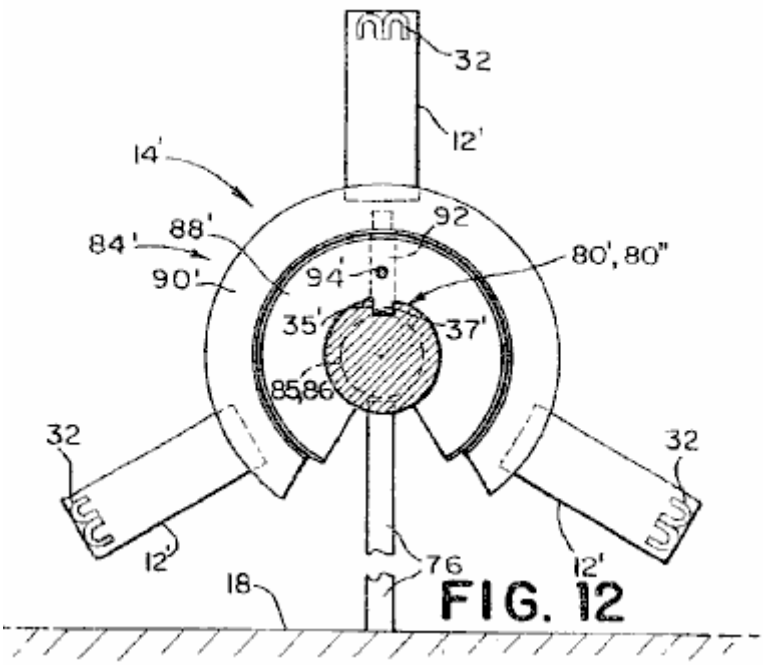


FIG. 12

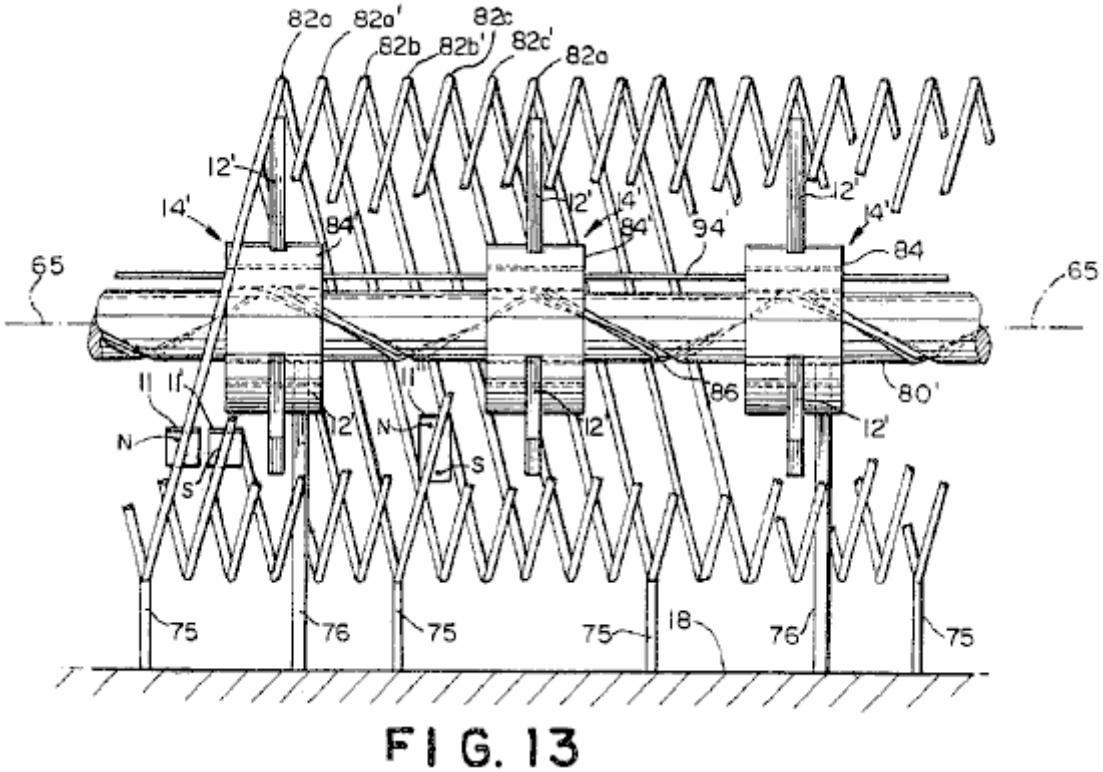


FIG. 13

Referring now to **Fig.2**, **Fig.11B**, **Fig.12** and **Fig.13** there is shown a thirteenth preferred embodiment of the motor **10** comprising a rail **80'** supported by rail mounting posts **76** and having a longitudinal axis **65**. A helical groove **86** having a pre-determined pitch runs around a periphery of the rail **80**.

The thirteenth preferred embodiment also includes three first helical stators **82a, 82b, 82c** (**82**) concentrically surrounding the rail **80'** corresponding to three pairs **32, 32' 32"** of U-shaped magnets **20** mounted on each of five rotors **12**. Preferably, the first helical stators **82** have the same pitch as the pre-determined pitch of the groove **86** and a longitudinal axis generally parallel to the axis **65** of the rail **80'**. A plurality of first stator magnets **11** having a direction of magnetisation aligned with a radial line of each rotor **12** are spaced along each first helical stator **82** with the first stator magnets **11** generating a first magnetic field.

The thirteenth preferred embodiment further includes plurality of second helical stators **82a', 82b', 82c'** (**82'**) alternating with the first helical stators **82'** along the axis **65** of the rail **80'**, and having the pre-determined pitch of the groove **86**. Each second helical stator **82'** has mounted upon it a plurality of second stator magnets **11'** having a direction of magnetisation aligned with a radial line of the rotor **12** and having a direction of magnetisation opposite in direction to the first stator magnets **11** mounted on each of the first helical stators **82**. As a consequence of the second helical stators **82'** being located midway between the first helical stators **82**, a point at about a midpoint between each rotor magnet pair **32, 32', 32''** is apposite to one of the second helical stators **82'** as each rotor **12** rotates about the axis **65** of the rail **80'** and slides along the rail **80'**.

The thirteenth preferred embodiment also includes five rotors **12**, (for clarity, only three are shown), having an axis of rotation **16** generally aligned with the longitudinal axis **65** of the rail **80'**. Each rotor **12** is connected to the rail **80'** by a bearing assembly **84'** so that the rotor **12** is free to rotate about the axis **65** of the rail **80'** and slide along the rail **80'**. Preferably, each rotor **12** includes three pairs **32, 32', 32''** of U-shaped magnets **20** wherein each U-shaped magnet **20** generates a second magnetic field, a portion of which adjacent to a rear **26** of the pair of U-shaped magnets **20** interacts with the first magnetic field of each first stator magnet to cause each rotor **12** to rotate about the axis **65** of the rail **80'**.

The bearing assembly **84'** (shown in detail in **Fig.11B** and **Fig.12**) connects each rotor **12** to the helical groove **86** around the periphery of the rail **80**. The bearing assembly **84'** is similar to the bearing assembly **84'** described in the twelfth preferred embodiment except for the openings in the first bearings **88'** and in the second bearing **90'** which allow the bearing assembly **84'** past the rail mounting posts **76** as the bearing assembly **84'** moves along the rail **80'**.

The thirteenth preferred embodiment may be constructed as either a linear motor or a rotary motor. In the case of the linear motor, the axes of the rail **80'** and of each helical stator **82** are substantially straight. The rail **80'** is supported on the base **18** by rail mounting posts **76** placed at intervals along the rail **80'**. The posts **76** are situated at locations along the rail **80'** at which the rotation of the rotor **12** orients the openings in the first and second bearings **88'**, **90'** to correspond to the mounting posts **76**. Each helical stator **82a, 82b, 82c** is supported on the base by stator mounting posts **75**. The rotors **12** are connected together by a cross-link **94'** which connects the first bearings **88'** of each bearing assembly **84'** to the first bearing **88'** of the bearing assembly **84'** of an adjacent rotor **12**. In this manner, the rotational motion of each rotor assembly **14'** is added together to provide the linear motive power of the linear motor.

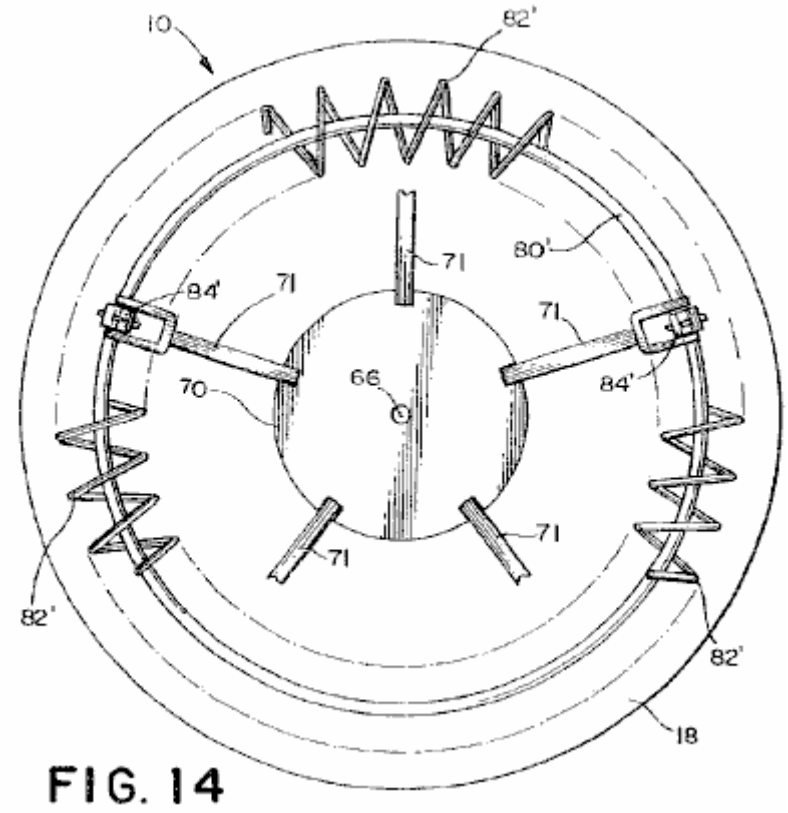


FIG. 14

The thirteenth preferred embodiment may also be constructed as a rotary motor **10** as shown in **Fig.14**. In this case, the axes of the rail **80'** and the helical stators **82** are configured to be circular. The circularly configured motor **10** includes an armature **70** centrally located within the perimeter of the rail **80'**. The armature **70** rotates

about an armature axis of rotation **58** connected for rotation within a motor base **18** to which the rail **80'** is also attached by mounting posts **76** (not shown). The pitch of the first and the second helical stators **82**, **82'**, measured at a radius of the rail **80**, preferably equals the predetermined pitch of the helical groove **86**. The armature **70** is fixedly attached to the first bearing **88** (see **Fig.11B**) of each bearing assembly **84'** by an armature strut **71** thereby adding together the rotational motive power of each rotor assembly **14**. In order that the armature strut **71** does not interfere with the first and second helical stators **82**, **82'**, the first and second helical stators **82**, **82'** are made to have an opening toward the armature axis of rotation **58**.

Preferably, each first helical stator **82a, 82b, 82c** has mounted upon it a plurality of first stator magnets **11** with each stator magnet **11** having a direction of magnetisation aligned with a radial line of the rotor **12**. Preferably, the first helical stators **82** are uniformly spaced along the longitudinal axis **65** of the rail **80'** with each first helical stator **82** corresponding to one of the plurality of magnet pairs **32, 32', 32"**. Preferably, each rotor **12** is positioned on the rail **80'** so that one of the rotor magnet pairs **32, 32', 32"** is apposite to one of the corresponding first helical stators **82** as the rotor **12** rotates about the axis **65** of the rail **80** and slides along the rail **80'**. However, as those skilled in the art will appreciate, the rotor magnet pairs **32, 32', 32"** need not be directly apposite to each helical stator **82** as the rotors **12** rotate in order to generate a rotational force.

Alternatively, as will be appreciated by those skilled in the art, the motor **10** can be constructed without the second helical stator **82'**. In the simplest case the motor **10** could comprise only a single first helical stator **82** and a single rotor **12** comprising a single U-shaped magnet **20** generating the second magnetic field. The single rotor **12** is preferably positioned in the groove **86** on the rail **80'** so that the U-shaped rotor magnet **20** is continually apposite to the single first helical stator **82**. Consequently, a portion of the second magnetic field directly adjacent to a rear **26** of the U-shaped magnet **20** interacts with the first magnetic field generated by each first stator magnet **11"** mounted on the helical stator **82** to cause the rotor **12** to rotate about the axis **65** of the rail **80** and to slide along the rail **80'**. Preferably, when only a single first stator **82** set of first stators **82** is used, each first stator magnet **11"** has a direction of magnetisation oriented to be in the plane of the rotor **12** and generally perpendicular to a radial line of the rotor **12**. The north pole and the south pole of the first stator magnet **11"** are preferably spaced apart so that when one pole of the first stator magnet **11** is directly apposite to the rotor magnet **20**, the pole of opposite polarity is equally spaced from the U-shaped magnet **20** of the rotor **12**. As one skilled in the art would appreciate, a plurality of U-shaped rotor magnets **20** and corresponding first helical stators could be used. Further, as those skilled in the art will appreciate, other configurations of the rotor magnet **20** and the stator magnet **11** are possible, all of which rely on the novel attributes of the magnetic field adjacent to the rear **26** of a U-shaped rotor magnet **20**. For example, the previously described stator magnet **11"** perpendicular to the radial line of the rotor **12** could be two separate bar magnets, spaced apart, with the magnetisation of each of the two magnets aligned with a radial line of the rotor and having opposite directions of magnetisation.

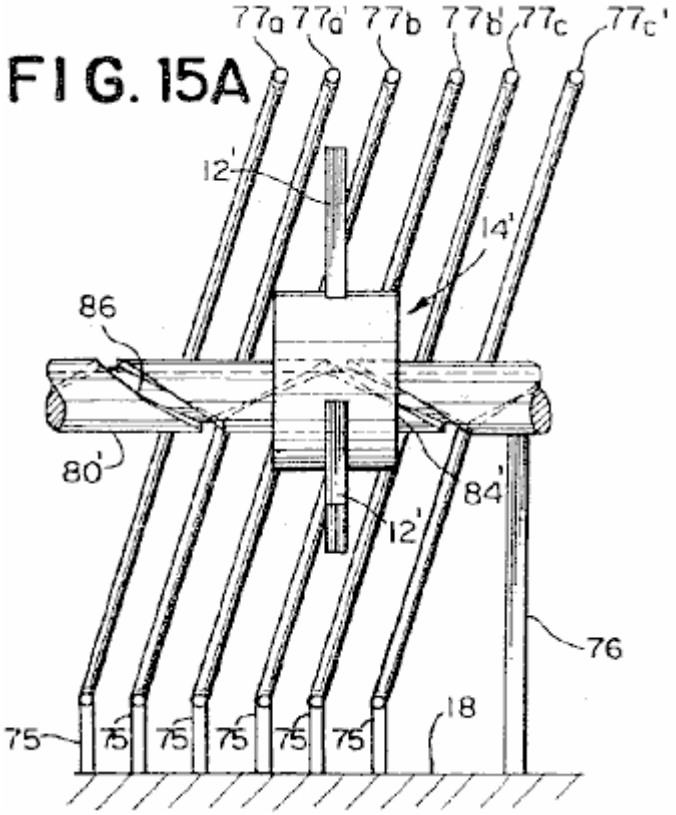


FIG. 15A

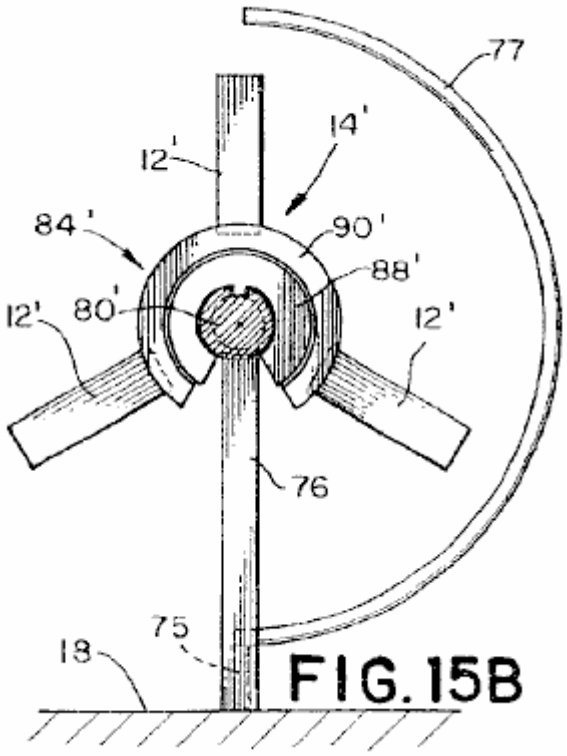


FIG. 15B

Referring now to **Fig.15A** and **Fig.15B** there is shown a fourteenth preferred embodiment of the motor **10**. The fourteenth embodiment is identical in structure to the thirteenth preferred embodiment except that the stator comprises a plurality of first ribs **77a, 77b, 77c** (**77**) and second ribs **77a', 77b', 77c'** (**77'**) in place of the first and the second helical stators **82, 82'** of the thirteenth embodiment. By substituting ribs **77, 77'** for the helical stators **82, 82'**, the attachment of the armature **70** to the rotors **12** is simplified. As those skilled in the art will appreciate, the length of the ribs **77, 77'** may vary from as little as 45 degrees to up to 265 degrees, with the motive power of the motor **10** being proportional to the length of the ribs.

Preferably, the first and the second ribs **77, 77'** have a pitch and a spacing that conforms to the pre-determined pitch of the rail **80'**. Further the orientation of the first and second stator magnets **11, 11'** and of the U-shaped rotor magnets **20** would be identical to the thirteenth embodiment. Accordingly, the operation of the fourteenth embodiment is identical to that of the thirteenth embodiment and is not repeated here for the sake of brevity.

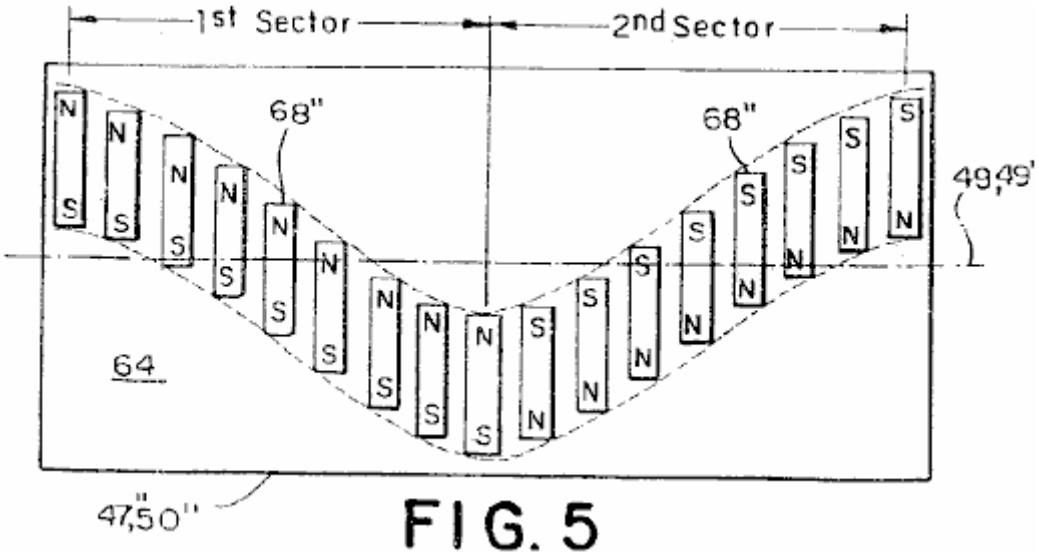


FIG. 5

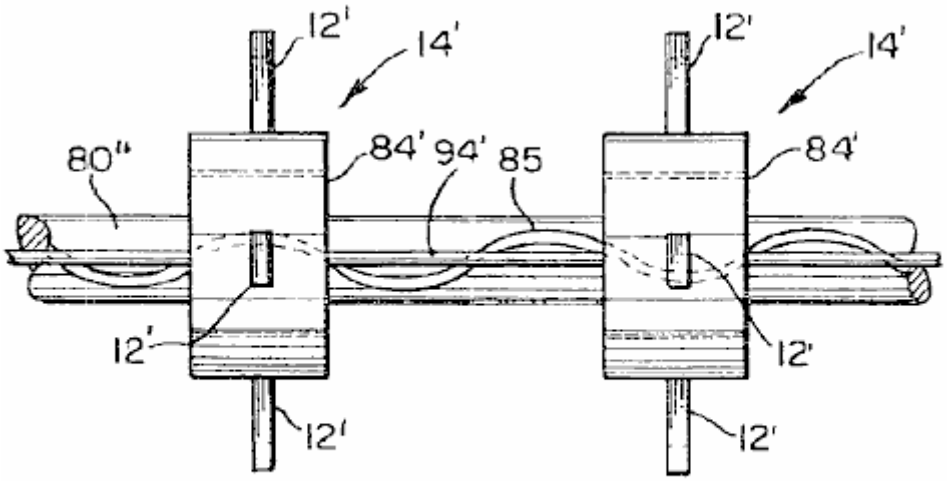


FIG. 16

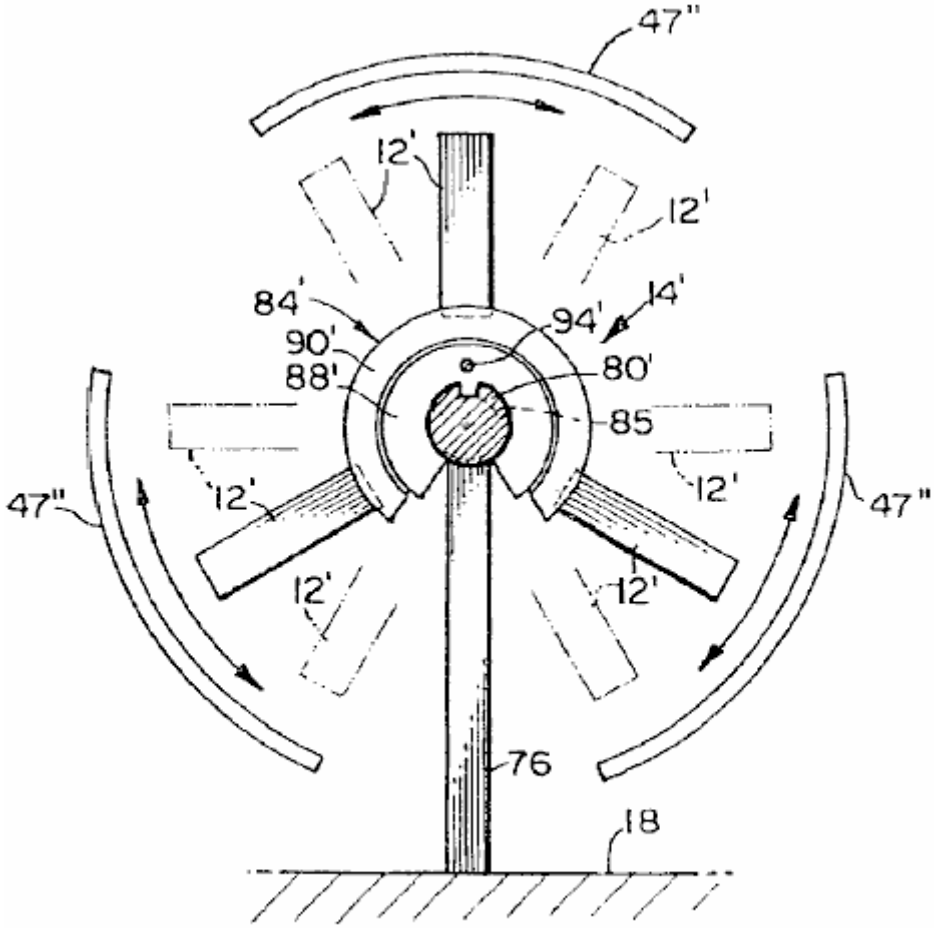


FIG. 17

Referring now to **Fig.5**, **Fig.16** and **Fig.17** there is shown a fifteenth preferred embodiment of the motor **10** comprising a rail **80"** having a longitudinal axis **65** and a generally sinusoidal groove **85** having a pre-determined period running around a periphery of the rail **80"**.

Preferably, the fifteenth preferred embodiment includes three generally identical stators **50"** arrayed in a circular fashion around the rail **80"**. Each stator **50"** has a surface **64** facing the rail **80"** and disposed generally equidistant from and parallel to the axis **65** of the rail **80"**. As shown in **Fig.5** and **Fig.17** each stator **50"** has a generally curved cross-section and a longitudinal line of demarcation **49** perpendicular to the cross-section and located about a midpoint of the surface **64**.

A plurality of stator magnets **68"** are attached to the surface **64** of the stator **50"** generating a first magnetic field. The stator magnets **68"** are displaced on the surface **64** in a sinusoidal pattern around the line of demarcation **49**. The sinusoidal pattern has a pre-determined period and a pre-determined maximum (peak) amplitude along the line of demarcation **49**. In the case where the rail **80"** and the longitudinal line of demarcation **49** of the stator **50"** are in a straight line, the period of the sinusoid is preferably equal to the period of the groove **85** on the rail **80**.

The sinusoidal pattern is also divided into a plurality of first and second alternating sectors with a boundary between the alternating sectors occurring at each maximum (peak) amplitude of the sinusoid. The direction of magnetisation of the stator magnets **68"** is opposite in the first and the second segments so that the direction of the first magnetic field in each first segment is opposite to the direction of the first magnetic field in each second segment. Preferably, the direction of magnetisation of the stator magnets **68"** is generally perpendicular to a radial line of the rotor **12**. Alternatively, the direction of magnetisation of the stator magnets **68"** could be generally aligned with a radial line of the rotor **12**. Further, as will be apparent to those skilled in the art, the first magnetic field need not be formed by a plurality of bar magnets but could be formed from a single magnet so that the first magnetic field would be sinusoidally displaced from the line of demarcation **49** and would alternate in opposite directions between the peaks of the sinusoid. Further, as will be appreciated by those skilled in the art, the displacement of the first magnetic field need not be precisely sinusoidal. For instance the displacement may be in a shape of a sawtooth or in a shape having a portion with constant plus and minus amplitude values, within the spirit and scope of the invention.

Preferably, the fifteenth preferred embodiment includes five rotors **12**, each rotor **12** having an axis **16** aligned with the axis of the rail **80''**. Each rotor **12** is connected to the rail **80''** by a bearing assembly **84'** so that the rotor **12** is free to rotate about the axis of the rail **65** and slide along the rail **80''**. Preferably, each rotor **12** includes three U-shaped magnet pairs **32, 32' 32''**, each pair comprising two U-shaped magnets **20**. Each U-shaped magnet **20** has a rear side and generates a second magnetic field. Each of the U-shaped magnet pairs **32, 32', 32''** is positioned on each rotor **12** so that the rear side **26** of each U-shaped magnet **20** is apposite to the first and the second segments of the sinusoidal pattern as the at least one rotor assembly **14** rotates about the rotor axis **16**, wherein an interaction of a portion of the second magnetic field directly adjacent to the rear **26** of each U-shaped magnet **20** with the first magnetic field of a corresponding stator **50''** causes the at least one rotor **12** to oscillate rotationally about the axis **65** of the rail **80''**. Those skilled in the art will appreciate that it is not necessary to have three pairs of U-shaped magnets **32, 32', 32''**. For instance, the number of U-shaped magnets **20** (or groups of abutted U-shaped magnets) spaced apart around the periphery of the rotor **12** may range from merely a single U-shaped magnet **20**, or may range in number up to a number of magnets limited only by the physical space around the periphery of the rotor **12**. Further the number of abutted U-shaped magnets **20** in a group of magnets **32** may also range from 1 up to a number of magnets limited only by the physical space around the periphery of the rotor **12**. Preferably, the number of stators **50''** equals the number of U-shaped magnet pairs **32, 32', 32''**. However, as will be appreciated by those skilled in the art, the number of stators **50''** is not limited to three but could be any number ranging upward from one, where the number of stators **50''** would preferably equal the number of U-shaped magnet pairs **32, 32', 32''**.

As shown in **Fig.16** the bearing assembly **84'** converts the oscillatory motion of the at least one rotor **12** about the rail to unidirectional linear motion along the rail **80'** by following the sinusoidal groove **85** in the rail **80'** with the boss **92** (shown in **Fig.11B**). A cross-link **94** connects the bearing assembly **84'** of adjacent rotors **12** together, thereby adding together the linear motion of each rotor assembly **14'** along the rail to provide the unidirectional linear motive power. The structure of the bearing assembly **84'** and the cross-link **94** is shown in **Fig.11B** and **Fig.12**, and the operation is identical to the linkage **84'** and the cross-link **94'** described for the twelfth embodiment. Accordingly, a detailed description of the linkage **84'** and the cross-link **94** is not repeated, for the sake of brevity.

In another aspect, the fifteenth preferred embodiment may also be configured in a circular arrangement similar to that of the fourteenth embodiment. In the fifteenth preferred embodiment, the helical stator **82'** shown in **Fig.14** is replaced with one or more curved stators **50''** spaced around the rotors **12**. In this case, the period of the sinusoidal pattern of the stator magnets is adjusted in accordance with the distance of the surface **64** of the respective stator **50''** from the armature axis of rotation **58** in order that the U-shaped magnets **20** on the rotors **12** remain apposite to the first and the second segments, as the rotors **12** slide along the rail **80''**. Accordingly, a description of those elements of circular arrangement of the fifteenth embodiment which are the same as for the linear embodiment are not repeated, for the sake of brevity.

Referring now to **Fig.4**, **Fig.18** and **Fig.19** there is shown a sixteenth preferred embodiment of the motor **10** for providing unidirectional motive power comprising a rail **80''** having a longitudinal axis **65** and a helical groove **86** having a pre-determined pitch, running around a periphery of the rail **80**.

Preferably, the sixteenth preferred embodiment further includes three generally identical stators **50'**, each stator **50'** having a surface **64** disposed generally equidistant from and parallel to the axis **65** of the rail **80**. Each stator **50'** has a longitudinal line of demarcation **49** located about a midpoint of the surface **64**. Preferably, a plurality of stator magnets **68'** are attached to the surface of the stator **50'** generating a first magnetic field. The plurality of stator magnets **68'** have a direction of magnetisation which rotates about a magnetic axis parallel to the line of demarcation **49**. In the case where the rail **80''** and the longitudinal line of demarcation **49** of the stator **50'** are in a straight line, the pitch of the rotation of the stator magnets **68'** is preferably equal to the pre-determined pitch of the helical groove **86** on the rail **80**.

The sixteenth embodiment further includes five rotors **12**, each rotor **12** having an axis of rotation **16** aligned with the axis **65** of the rail **80**. Each rotor **12** is connected to the rail **80** so that the rotor **12** is free to rotate about the axis **65** of the rail **80** and slide along the rail **80**. Each rotor **12** includes three pairs **32, 32', 32''** of U-shaped magnets **20** spaced around the periphery of the rotor **12**, each U-shaped magnet **20** generating a second magnetic field. The U-shaped magnets **20** are positioned on each rotor **12** so that a portion of the second magnetic field directly adjacent to the rear side **26** of the U-shaped magnet **20** interacts with the first magnetic field generated by the plurality of stator magnets **68'** to cause each rotor **12** to rotate about the rotor axis **16**. Those skilled in the art will appreciate that it is not necessary to have exactly three pairs of U-shaped magnets **32, 32', 32''**. For instance, the number of U-shaped magnets **20** (or groups of abutted U-shaped magnets) spaced apart around the periphery of the rotor **12** may range from merely a single U-shaped magnet **20**, or may range in number up to a number of U shaped magnets **20** limited only by the physical space around the periphery of the

rotor **12**. Further the number of abutted U-shaped magnets **20** in a group of magnets **32** may also range from 1 up to a number of magnets limited only by the physical space around the periphery of the rotor **12**.

The sixteenth embodiment also includes a bearing assembly 84' connecting each rotor **12** to the helical groove **86**, the bearing assembly **84'** converting the rotary motion of each rotor **12** about the rail **80'** to unidirectional linear motion along the rail **80'**. A cross-link **94** connects the bearing assembly **84'** of adjacent rotors **12** together, thereby adding together the linear motion of each rotor assembly **14'** along the rail **80'** to provide the unidirectional linear motive power. The structure of the bearing assembly **84'** and the cross-link **94** is shown in **Fig.11B** and **Fig.12**, is identical to the bearing assembly **84'** and cross-link **94** described for the twelfth embodiment. Accordingly, a description of the linkage **84** and the cross-link **94** is not repeated, for the sake of brevity.

In another aspect of the sixteenth preferred embodiment the motor **10** may be configured in a circular arrangement similar to that of the fourteenth embodiment, as shown in **Fig.14**, except that the helical stator **82'** shown in **Fig.14** is replaced with one or more stators **50'** spaced around the rotors **12**. In this case, the pitch of the rotation of the plurality of stator magnets **68'** is adjusted in accordance with the distance of the surface **64** of the respective stator **50'** from the armature axis of rotation **58** in order that the U-shaped magnets **20** on the rotors **12** remain aligned with the plurality of stator magnets **68'** as the rotors **12** rotate about the axis **65** of the rail **80'** and slide along the rail **80'**. Accordingly, a description of those elements of the circular arrangement of the sixteenth embodiment which are the same as for the straight line configuration are not repeated, for the sake of brevity.

It will be appreciated by those skilled in the art that changes could be made to the embodiments described above without departing from the broad inventive concept thereof. It is understood, therefore, that this invention is not limited to the particular embodiments disclosed, but it is intended to cover modifications within the spirit and scope of the present invention as defined by the appended claims.

CLAIMS

1. An apparatus (10, 10') characterised by:

at least one rotor (12) having a periphery and a rotor axis (16), the at least one rotor (12) comprising a first rotor magnet (20) producing a first magnetic field, said first rotor magnet being U-shaped and having a north pole (23), a south pole (25) and a rear side (26), the rear side (26) of the first rotor magnet being adjacent to the periphery;

an axle (80) to which the at least one rotor (12) is connected at the rotor axis (16) for rotation of the at least one rotor (12) about the rotor axis (16); and

a stationary stator (48, 51) comprising a generally curved cross-section, said stator (51) having a surface (64) opposing the periphery of the at least one rotor (12), and a longitudinal line of demarcation (49) perpendicular to the cross-section at about a midpoint of the surface, the line of demarcation (49) delineating a first side (52) of the surface from a second side (54) of the surface (64), wherein a plurality of pairs of stator magnets (40, 42) producing a second magnetic field are attached to the surface (64), each pair of stator magnets (40, 42) comprising a first stator magnet (40) having a north pole and a south pole and a direction of magnetisation substantially parallel to the surface (64), and a second stator magnet (42) having a north pole and a south pole and a direction of magnetisation substantially parallel to the surface (64), the first stator magnet (40) being on the first side of the surface (64) with the north pole of the first stator magnet being closest to the line of demarcation (49), the second stator magnet (42) being on the second side (52) of the surface with the south pole of the second stator magnet (42) being closest to the line of demarcation (49), wherein the plurality of pairs of stator magnets (40, 42) are spaced along the line of demarcation (49) so that a first inter-magnet distance measured along the line of demarcation (49) between the north pole of the first stator magnet (40) and the south pole of the second stator (42) magnet of an adjacent pair of stator magnets (40, 42) is about equal to a second inter-magnet distance measured along the line of demarcation between the south pole of the first stator magnet (40) and the north pole of the second stator magnet (42), and wherein the interaction of the first and the second magnetic fields cause the at least one rotor (12) to translate in a predetermined direction along the line of demarcation.

2. The apparatus (10, 10') of claim 1, characterised by the north pole of each first stator magnet (40) and the south pole of each second stator magnet (42) being inclined toward the predetermined direction.

3. The apparatus (10, 10') of claim 1, further characterised by the rotor (12) including a second rotor magnet (22), said second rotor magnet (24) being U-shaped and having a north pole, a south pole and a rear side, the south pole of the second rotor magnet (22) abutting the north pole of the first rotor magnet (26) and the north pole of the second rotor magnet being adjacent to the periphery, and a third rotor magnet (24), said third rotor

magnet (24) having a north pole, a south pole and a rear side, the north pole of the third rotor magnet (24) abutting the south pole of the first rotor magnet (26) and the south pole of the third rotor magnet (24) being adjacent to the periphery, said second magnet producing a third magnetic field and third magnet producing a fourth magnetic field.

4. The apparatus (10) of claim 1, characterised by the apparatus further including an armature (70) having an armature axis (58), the at least one rotor (12) being spaced from the armature (70) by an armature strut (71) and connected thereto by the axle (80) for rotation about the rotor axis (16), the at least one rotor (12) configured for rotation in a plane generally aligned with the armature axis (58), wherein the stator (51) is circular-cylindrical, with a stator axis (72) aligned with the armature axis (58).

5. The apparatus (10') of claim 1, further characterised by the stator (48) being linear, the stator (48) oriented so that the surface (64) of the stator (48) is generally parallel to the axle (80), each at least one rotor (12) being connected to the axle (80) by a bearing assembly (84) comprising a pair of first bearings (88) slidably attached to the axle (80), and a second bearing (90) connected to the pair of first bearings (88) for rotation about the first pair of bearings (88), said at least one rotor (12) being fixedly attached to the second bearing (90).

6. The apparatus (10') of claim 5, further characterised by a crosslink (94) which connects together the at least one rotors (12).

7. A apparatus (10, 10') characterised by:

at least one rotor (12) having a periphery, a rotor axis (16) and a thruster axis (34) perpendicular to the rotor axis (16) and intersecting the rotor axis (16), the at least one rotor (12) comprising spaced apart first and second rotor magnets (36, 38) having north and south poles aligned with the thruster axis (34), and a third rotor magnet (20, 20') located between the first and second rotor magnets (34, 38) on an axis generally perpendicular to the thruster axis (34), said first, second and third magnets producing a first magnetic field;

an axle (80) to which the at least one rotor (12) is connected at the rotor axis (16) for rotation of the at least one rotor (12) about the rotor axis (16); and

a stationary stator (48', 51') comprising a generally curved cross-section, said stator (48', 51') having a surface 64 opposing the periphery of the at least one rotor (12), and a longitudinal line of demarcation (49) perpendicular to the cross-section at about a midpoint of the surface (64), the line of demarcation (49) delineating a first side (52) of the surface from a second side (54) of the surface, wherein a plurality of sets of stator magnets (40', 42', 41) producing a second magnetic field are attached to the surface (64), each set of stator magnets (40', 42', 41) comprising a first stator magnet (40') having a north pole and a south pole and a direction of magnetisation substantially perpendicular to the surface (64), a second stator magnet (42') having a north pole and a south pole and a direction of magnetisation substantially perpendicular to the surface (64), and a third stator magnet (41), the third stator magnet (41) being attached to the stator (48', 51') along the line of demarcation (49) midway between the first stator magnet (40') and the second stator magnet (42'), the first stator magnet (40') being on the first side (52) of the surface with the south pole of the first stator magnet (40') being closest surface (64), the second stator magnet (42') being on the second side (54) of the surface (64) with the north pole of the second stator magnet (42') being closest to the surface (64), wherein the plurality of sets of stator magnets (40', 42', 41) are spaced along the line of demarcation (49) so that a first inter-magnet distance measured along the line of demarcation (49) between the north pole of the first stator magnet (40') and the south pole of the second stator magnet (42') of an adjacent pair of stator magnets (40', 42', 41) is about equal to a second inter-magnet distance measured along the line of demarcation (49) between the south pole of the first stator magnet (40') and the north pole of the second stator magnet (42'), wherein the interaction of the first and the second magnetic fields cause the at least one rotor (12) to translate in a predetermined direction along the line of demarcation.

8. The apparatus (10, 10') of claim 7, characterised by the third rotor magnet (20) being a U-shaped magnet and the third stator magnet (41) being a bar magnet.

9. The apparatus (10, 10') of claim 7, characterised by the third rotor magnet (20') being a bar magnet and the third stator magnet (41') being a U-shaped magnet.

10. The apparatus (10) of claim 7, characterised by the apparatus further including an armature (70) having an armature axis (58), the at least one rotor (12) being spaced from the armature (70) by an armature strut (71) and connected thereto by the axle (80) for rotation about the rotor axis (16), the at least one rotor (12) being configured for rotation in a plane generally aligned with the armature axis (58), wherein the stator (51') is circular, with a stator axis (72) aligned with the armature axis (58).

11. The apparatus (10') of claim 7, further characterised by the stator (48') being linear, the stator (48') oriented so that the surface (64) of the stator (48') is generally parallel to the axle (80), each at least one rotor (12) being connected to the axle (80) by a bearing assembly (84) comprising a pair of first bearings (88) slidably attached to the axle (80), and a second bearing (90) connected to the pair of first bearings (88) for rotation about the pair of first bearings (88), said at least one rotor (12) being fixedly attached to the second bearing (90).

12. The apparatus (10') of claim 11, further characterised by a crosslink (94) which connects together the at least one rotors (12).

13. An apparatus (10) for providing motion characterised by:

a stationary, generally circular, stator (50, 50', 50'') having a stator axis (58), an outer surface (64), and a circumferential line of demarcation (49) in a plane perpendicular to the stator axis (58) at about a midpoint of the outer surface (64);

at least one stator magnet (68, 68', 68'') attached to the outer surface (64) of the stator (50, 50', 50''), the at least one stator magnet (68, 68', 68'') being arranged in a generally circular arrangement about the stator axis (58);
an armature (70) attached to the stator (50, 50'', 50'') for rotation therewith, the armature (70) having an axis parallel to the stator axis (58);

at least one rotor (12) including at least one rotor magnet (20), the at least one rotor (12) being spaced from the armature (70) by an armature strut (71) and connected thereto by an axle (80) for rotation about a rotor axis (16), the at least one rotor (12) being configured for rotation in a plane generally aligned with the stator axis (58); and

a driving linkage assembly (53, 55, 62) connecting the at least one rotor to the stator, the linkage assembly (53, 55, 62) configured to cause the armature (70) to rotate about the stator axis (58) when the at least one rotor (12) rotates about the rotor axis (16).

14. The apparatus according to claim 13 wherein a direction of magnetisation of the at least one stator magnet (68) is generally perpendicular to a radial line of the at least one rotor (12).

15. The apparatus according to claim 13 wherein a direction of magnetisation of the at least one stator magnet (68) is generally aligned with a radial line of the at least one rotor (12).

16. The apparatus according to claim 13 wherein the at least one rotor magnet (20) comprises a U-shaped magnet.

17. The apparatus according to claim 13 wherein the at least one rotor magnet (20) comprises a bar magnet and the at least one stator magnet (68) is a U-shaped magnet.

18. The apparatus according to claim 13, the at least one stator magnet (68') having a direction of magnetisation which rotates about the circumferential line of demarcation (49) with a predetermined periodicity.

19. The apparatus according to claim 13, the at least one stator magnet (68'') having a direction of magnetisation in a plane of the stator (50'') and which is displaced in a sinusoidal pattern from the line of demarcation (49), the sinusoidal pattern having a pre-determined period and a pre-determined maximum amplitude and divided into a plurality of alternating first and second sectors with a boundary between the alternating first and second sectors occurring at peak amplitudes of the sinusoid, the direction of magnetisation of the at least one magnet (68'') being opposite in direction in the first and the second segments.

HOWARD JOHNSON: PERMANENT MAGNET MOTOR

Patent US 4,151,431 **24th April 1979** **Inventor: Howard R. Johnson**

PERMANENT MAGNET MOTOR

This is a re-worded extract from this Patent. It describes a motor powered solely by permanent magnets and which it is claimed can power an electrical generator.

ABSTRACT

The invention is directed to the method of utilising the unpaired electron spins in ferromagnetic and other materials as a source of magnetic fields for producing power without any electron flow as occurs in normal conductors, and to permanent magnet motors for utilising this method to produce a power source. In the practice of the invention the unpaired electron spins occurring within permanent magnets are utilised to produce a motive power source solely through the superconducting characteristics of a permanent magnet, and the magnetic flux created by the magnets is controlled and concentrated to orientate the magnetic forces generated in such a manner to produce useful continuous work, such as the displacement of a rotor with respect to a stator. The timing and orientation of magnetic forces at the rotor and stator components produced by the permanent magnets is accomplished by the proper geometrical relationship of these components.

BACKGROUND OF THE INVENTION:

Conventional electric motors employ magnetic forces to produce either rotational or linear motion. Electric motors operate on the principal that when a conductor which carries a current is located in a magnetic field, a magnetic force is exerted upon it. Normally, in a conventional electric motor, the rotor, or stator, or both, are so wired that magnetic fields created by electromagnets use attraction, repulsion, or both types of magnetic forces, to impose a force upon the armature causing rotation, or linear displacement of the armature. Conventional electric motors may employ permanent magnets either in the armature or stator components, but to date they require the creation of an electromagnetic field to act upon the permanent magnets. Also, switching gear is needed to control the energising of the electromagnets and the orientation of the magnetic fields producing the motive power.

It is my belief that the full potential of magnetic forces existing in permanent magnets has not been recognised or utilised because of incomplete information and theory with respect to atomic motion occurring within a permanent magnet. It is my belief that a presently unnamed atomic particle is associated with the electron movement of a superconducting electromagnet and the loss-less flow of currents in permanent magnets. The unpaired electron flow is similar in both situations. This small particle is believed to be opposite in charge to an electron and to be located at right angles to the moving electron. This particle must be very small to penetrate all known elements in their various states as well as their known compounds (unless they have unpaired electrons which capture these particles as they endeavour to pass through).

The electrons in ferrous materials differ from those found in most elements in that they are unpaired, and being unpaired they spin around the nucleus in such a way that they respond to magnetic fields as well as creating a magnetic field themselves. If they were paired, their magnetic fields would cancel out. However, being unpaired they create a measurable magnetic field if their spins are orientated in one direction. The spins are at right angles to their magnetic fields.

In niobium superconductors, at a critical state, the magnetic lines of force cease to be at right angles. This change must be due to establishing the required conditions for unpaired electronic spins instead of electron flow in the conductor, and the fact that very powerful electromagnets can be formed with superconductors illustrates the tremendous advantage of producing the magnetic field by unpaired electron spins rather than conventional electron flow. In a superconducting metal, wherein the electrical resistance becomes greater in the metal than the proton resistance, the flow turns to electron spins and the positive particles flow parallel in the metal in the manner occurring in a permanent magnet where a powerful flow of magnetic positive particles or magnetic flux

causes the unpaired electrons to spin at right angles. Under cryogenic superconduction conditions the freezing of the crystals in place makes it possible for the spins to continue, and in a permanent magnet the grain orientation of the magnetised material allows these spins, permitting them to continue and causing the flux to flow parallel to the metal. In a superconductor, at first the electron is flowing and the positive particle is spinning; later, when critical, the reverse occurs, i.e., the electron is spinning and the positive particle is flowing at right angles. These positive particles will thread or work their way through the electron spins present in the metal.

In a sense, a permanent magnet may be considered a room-temperature superconductor. It is a superconductor because the electron flow does not cease, and this electron flow can be made to do work through the magnetic field which it creates. Previously, this source of power has not been used because it was not possible to modify the electron flow to accomplish the switching functions of the magnetic field. Such switching functions are common in a conventional electric motor where electrical current is employed to align the much greater electron current in the iron pole pieces and concentrate the magnetic field at the proper places to give the thrust necessary to move the motor armature. In a conventional electric motor, switching is accomplished by the use of brushes, commutators, alternating current, or other means.

In order to accomplish the switching function in a permanent magnet motor, it is necessary to shield the magnetic leakage so that it will not appear as too great a loss factor at the wrong places. The best method to accomplish this is to concentrate the magnetic flux in the place where it will be the most effective. Timing and switching can be achieved in a permanent magnet motor by concentrating the flux and using the proper geometry of the motor rotor and stator to make most effective use of the magnetic fields. By the proper combination of materials, geometry and magnetic concentration, it is possible to achieve a mechanical advantage of high ratio, greater than 100 to 1, capable of producing continuous motive force.

To my knowledge, previous work done with permanent magnets, and motive devices utilising permanent magnets, have not achieved the result desired in the practice of the inventive concept, and it is with the proper combination of materials, geometry and magnetic concentration that the presence of the magnetic spins within a permanent magnet may be utilised as a motive force.

SUMMARY OF THE INVENTION:

It is an object of the invention to utilise the magnetic spinning phenomenon of unpaired electrons occurring in ferromagnetic material to produce the movement of a mass in a unidirectional manner so as to permit a motor to be driven solely by the magnetic forces occurring within permanent magnets. Both linear and rotational types of motor may be produced. It is an object of the invention to provide the proper combination of materials, geometry and magnetic concentration to power a motor. Whether the motor is a linear type or a rotary type, in each instance the "stator" may consist of several permanent magnets fixed relative to each other, to create a track. This track is linear for a linear motor and circular for a rotary motor. An armature magnet is carefully positioned above this track so that an air gap exists between it and the track. The length of the armature magnet is defined by poles of opposite polarity, and the longer axis of the armature magnet is pointed in the direction of its movement.

The stator magnets are mounted so that all the same poles face the armature magnet. The armature magnet has poles which are both attracted to and repelled by the adjacent pole of the stator magnets, so both attractive and repulsive forces act upon the armature magnet to make it move.

The continuing motive force which acts on the armature magnet is caused by the relationship of the length of the armature magnet to the width and spacing of the stator magnets. This ratio of magnet and magnet spacings, and with an acceptable air gap spacing between the stator and armature magnets, produces a continuous force which causes the movement of the armature magnet.

In the practice of the invention, movement of the armature magnet relative to the stator magnets results from a combination of attractive and repulsive forces between the stator and armature magnets. By concentrating the magnetic fields of the stator and armature magnets the motive force imposed upon the armature magnet is intensified, and in the disclosed embodiments, the means for achieving this magnetic field concentration are shown.

This method comprises of a plate of high magnetic field permeability placed behind one side of the stator magnets and solidly engaged with them. The magnetic field of the armature magnet may be concentrated and directionally oriented by bowing the armature magnet, and the magnetic field may further be concentrated by shaping the pole ends of the armature magnet to concentrate the magnet field at a relatively limited surface at the armature magnet pole ends.

Preferably, several armature magnets are used and these are staggered relative to each other in the direction their movement. Such an offsetting or staggering of the armature magnets distributes the impulses of force

imposed upon the armature magnets and results in a smoother application of forces to the armature magnet producing a smoother and more uniform movement of the armature component.

In the rotary embodiment of the permanent magnet motor of the invention the stator magnets are arranged in a circle, and the armature magnets rotate about the stator magnets. A mechanism is shown which can move the armature relative to the stator and this controls the magnitude of the magnetic forces, altering the speed of rotation of the motor.

BRIEF DESCRIPTION OF THE DRAWINGS

The objects and advantages of the invention mentioned earlier, will be appreciated from the following description and accompanying drawings:

Fig. 1 is a schematic view of electron flow in a superconductor indicating the unpaired electron spins,
Fig. 2 is a cross-sectional view of a superconductor under a critical state illustrating the electron spins,
Fig. 3 is a view of a permanent magnet illustrating the flux movement through it,
Fig. 4 is a cross-sectional view illustrating the diameter of the magnet of Fig.3,
Fig. 5 is an elevational representation of a linear motor embodiment of the permanent magnet motor of the invention illustrating one position of the armature magnet relative to the stator magnets, and indicating the magnetic forces imposed upon the armature magnet,
Fig. 6 is a view similar to Fig.5 illustrating displacement of the armature magnet relative to the stator magnets, and the influence of magnetic forces thereon at this location,
Fig. 7 is a further elevational view similar to Fig.5 and Fig.6 illustrating further displacement of the armature magnet to the left, and the influence of the magnetic forces thereon,
Fig. 8 is a top plan view of a linear embodiment of the inventive concept illustrating a pair of armature magnets in linked relationship disposed above the stator magnets,
Fig. 9 is a diametrical, elevational, sectional view of a rotary motor embodiment in accord with the invention as taken along section IX-IX of Fig.10, and
Fig. 10 is an elevational view of the rotary motor embodiment as taken along X-X of Fig.9.

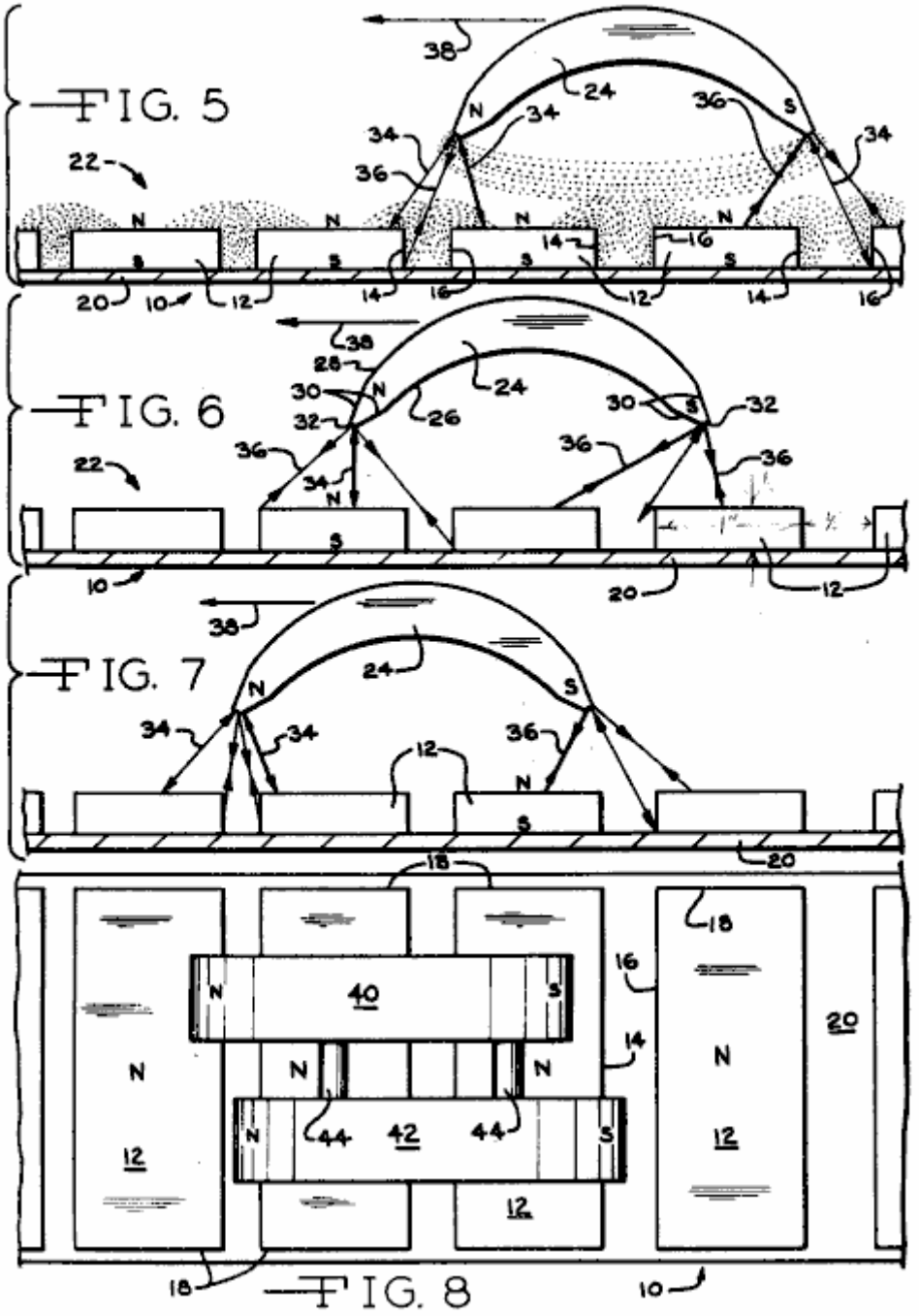


FIG. 5
FIG. 6
FIG. 7
FIG. 8
38
24
34
36
22
N
S
14
16
20
10
12
18
40
42
44

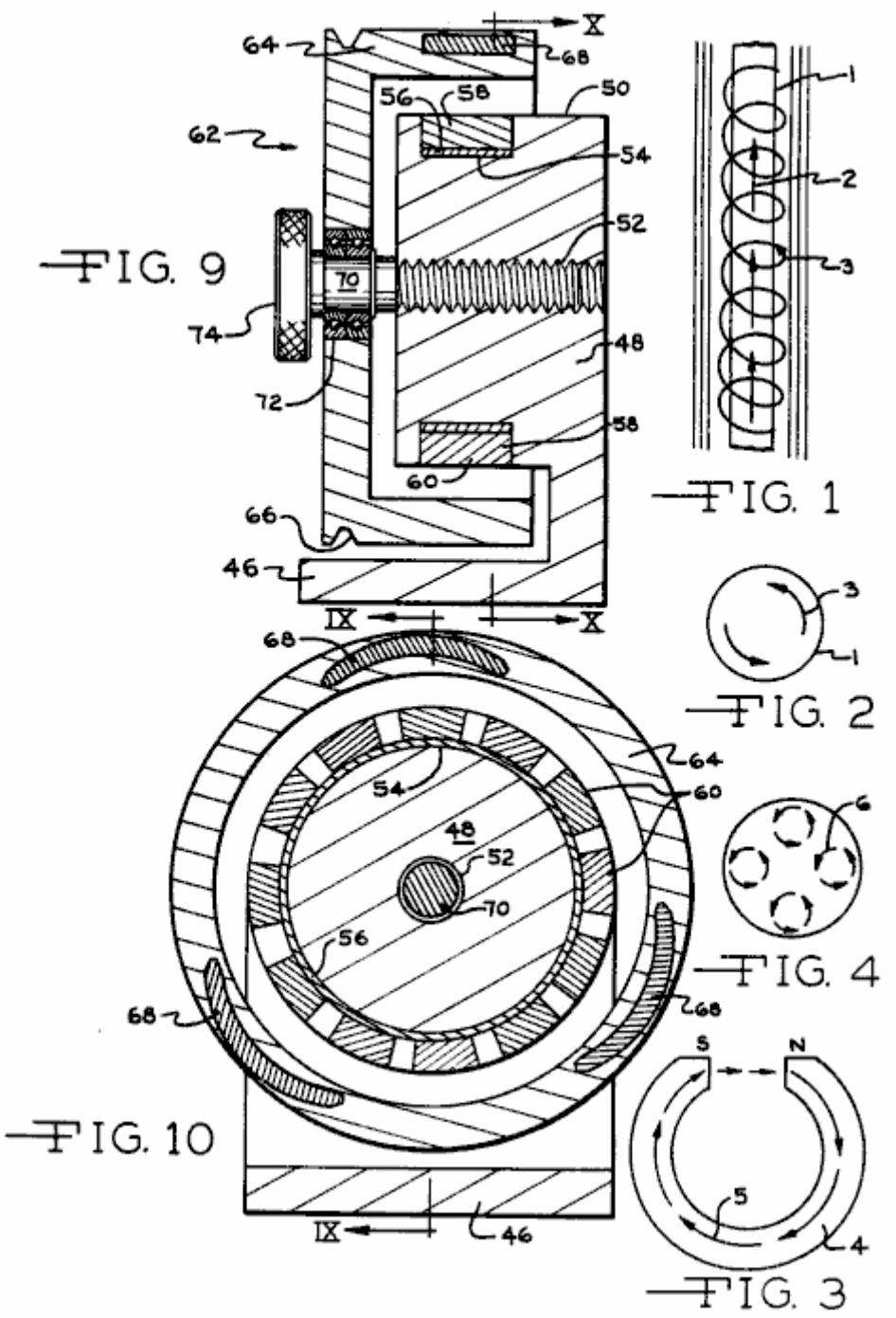


FIG. 9

FIG. 1

FIG. 2

FIG. 4

FIG. 10

FIG. 3

DESCRIPTION OF THE PREFERRED EMBODIMENTS

In order to better understand the theory of the inventive concept, reference is made to **Figs. 1 through 4**. In **Fig.1** a superconductor 1 is illustrated having a positive particle flow as represented by arrow **2**, the unpaired electrons of the ferrous conductor **1** spin at right angles to the proton flow in the conductor as represented by the spiral line and arrow **3**. In accord with the theory of the invention the spinning of the ferrous unpaired electrons results from the atomic structure of ferrous materials and this spinning atomic particle is believed to be opposite in charge and located at right angles to the moving electrons. It is assumed to be very small in size capable of penetrating other elements and their compounds unless they have unpaired electrons which capture these particles as they endeavour to pass through.

The lack of electrical resistance of conductors at a critical superconductor state has long been recognised, and superconductors have been utilised to produce very high magnetic flux density electromagnets. **Fig.2** represents a cross section of a critical superconductor and the electron spins are indicated by the arrows **3**. A permanent magnet may be considered a superconductor as the electron flow therein does not cease, and is without resistance, and unpaired electric spinning particles exist which, in the practice of the invention, are utilised to produce motor force. **Fig.3** illustrates a horseshoe shaped permanent magnet at **4** and the magnetic flux through it is indicated by arrows **5**, the magnetic flow being from the south pole to the north pole and through the magnetic material. The accumulated electron spins occurring about the diameter of the magnet **5** are represented at **6** in

Fig.4, and the spinning electron particles spin at right angles in the iron as the flux travels through the magnet material.

By utilising the electron spinning theory of ferrous material electrons, it is possible with the proper ferromagnetic materials, geometry and magnetic concentration to utilise the spinning electrons to produce a motive force in a continuous direction, thereby resulting in a motor capable of doing work.

It is appreciated that the embodiments of motors utilising the concepts of the invention may take many forms, and in the illustrated forms the basic relationships of components are illustrated in order to disclose the inventive concepts and principles. The relationships of the plurality of magnets defining the stator 10 are best appreciated from **Figs. 5 through 8**. The stator magnets **12** are preferably of a rectangular configuration, **Fig.8**, and so magnetised that the poles exist at the large surfaces of the magnets, as will be appreciated from the N (North) and S (South) designations. The stator magnets include side edges **14** and **16** and end edges **18**. The stator magnets are mounted upon a supporting plate **20**, which is preferably of a metal having a high permeability to magnetic fields and magnetic flux such as that available under the trademark Netic CoNetic sold by Perfection Mica Company of Chicago, Illinois. Thus, the plate **20** will be disposed toward the south pole of the stator magnets **12**, and preferably in direct engagement therewith, although a bonding material may be interposed between the magnets and the plate in order to accurately locate and fix the magnets on the plate, and position the stator magnets with respect to each other.

Preferably, the spacing between the stator magnets **12** slightly differs between adjacent stator magnets as such a variation in spacing varies the forces being imposed upon the armature magnet at its ends, at any given time, and thus results in a smoother movement of the armature magnet relative to the stator magnets. Thus, the stator magnets so positioned relative to each other define a track **22** having a longitudinal direction left to right as viewed in **Figs. 5 through 8**.

In **Figs. 5 through 7** only a single armature magnet **24** is disclosed, while in **Fig.8** a pair of armature magnets are shown. For purposes of understanding the concepts of the invention the description herein will be limited to the use of single armature magnet as shown in **Figs. 5 through 7**.

The armature magnet is of an elongated configuration wherein the length extends from left to right, **Fig.5**, and may be of a rectangular transverse cross-sectional shape. For magnetic field concentrating and orientation purposes the magnet **24** is formed in an arcuate bowed configuration as defined by concave surfaces **26** and convex surfaces **28**, and the poles are defined at the ends of the magnet as will be appreciated from **Fig.5**. For further magnetic field concentrating purposes the ends of the armature magnet are shaped by bevelled surfaces **30** to minimise the cross sectional area at the magnet ends **32**, and the magnetic flux existing between the poles of the armature magnet are as indicated by the light dotted lines. In like manner the magnetic fields of **6** the stator magnets **12** are indicated by the light dotted lines.

The armature magnet **24** is maintained in a spaced relationship above the stator track **22**. This spacing may be accomplished by mounting the armature magnet upon a slide, guide or track located above the stator magnets, or the armature magnet could be mounted upon a wheeled vehicle carriage or slide supported upon a non-magnetic surface or guideway disposed between the stator magnets and the armature magnet. To clarify the illustration, the means for supporting the armature magnet **24** is not illustrated and such means form no part of invention, and it is to be understood that the means supporting the armature magnet prevents the armature magnet from moving away from the stator magnets, or moving closer thereto, but permits free movement of the armature magnet to the left or right in a direction parallel to the track **22** defined by the stator magnets.

It will be noted that the length of the armature magnet **24** is slightly greater than the width of two of the stator magnets **12** and the spacing between them. The magnetic forces acting upon the armature magnet when in the position of **Fig.5** will be repulsion forces **34** due to the proximity of like polarity forces and attraction forces at **36** because of the opposite polarity of the south pole of the armature magnet, and the north pole field of the sector magnets. The relative strength of this force is represented by the thickness of the force line.

The resultant of the force vectors imposed upon the armature magnet as shown in **Fig.5** produce a primary force vector **38** toward the left, **Fig.5**, displacing the armature magnet **24** toward the left. In **Fig.6** the magnetic forces acting upon the armature magnet are represented by the same reference numerals as in **Fig.5**. While the forces **34** constitute repulsion forces tending to move the north pole of the armature magnet away from the stator magnets, the attraction forces imposed upon the south pole of the armature magnet and some of the repulsion forces, tend to move the armature magnet further to the left, and as the resultant force **38** continues to be toward the left the armature magnet continues to be forced to the left. **Fig.7** represents further displacement of the armature magnet **24** to the left with respect to the position of **Fig.6**, and the magnetic forces acting thereon are represented by the same reference numerals as in **Fig.5** and **Fig.6**, and the stator magnet will continue to move to the left, and such movement continues the length of the track **22** defined by the stator magnets **12**.

Upon the armature magnet being reversed such that the north pole is positioned at the right as viewed in **Fig.5**, and the south pole is positioned at the left, the direction of movement of the armature magnet relative to the stator magnets is toward the right, and the theory of movement is identical to that described above.

In **Fig.8** a plurality of armature magnets **40** and **42** are illustrated which are connected by links **44**. The armature magnets are of a shape and configuration identical to that of the embodiment of **Fig.5**, but the magnets are staggered with respect to each other in the direction of magnet movement, i.e., the direction of the track **22** defined by the stator magnets **12**. By so staggering a plurality of armature magnets a smoother movement of the interconnected armature magnets is produced as compared when using a single armature magnet as there is variation in the forces acting upon each armature magnet as it moves above the track **22** due to the change in magnetic forces imposed thereon. The use of several armature magnets tends to "smooth out" the application of forces imposed upon linked armature magnets, resulting in a smoother movement of the armature magnet assembly. Of course, any number of armature magnets may be interconnected, limited only by the width of the stator magnet track **22**.

In **Fig.9** and **Fig.10** a rotary embodiment embracing the inventive concepts is illustrated. In this embodiment the principle of operation is identical to that described above, but the orientation of the stator and armature magnets is such that rotation of the armature magnets is produced about an axis, rather than a linear movement being achieved.

In **Fig.9** and **Fig.10** a base is represented at **46** serving as a support for a stator member **48**. The stator member **48** is made of a non-magnetic material, such as synthetic plastic, aluminium, or the like. The stator includes a cylindrical surface **50** having an axis, and a threaded bore **52** is concentrically defined in the stator. The stator includes an annular groove **54** receiving an annular sleeve **56** of high magnetic field permeability material such as Netic Co-Netic and a plurality of stator magnets **58** are affixed upon the sleeve **56** in spaced circumferential relationship as will be apparent in **Fig.10**. Preferably, the stator magnets **58** are formed with converging radial sides as to be of a wedge configuration having a curved inner surface engaging sleeve **56**, and a convex pole surface **60**.

The armature **62**, in the illustrated embodiment, is of a dished configuration having a radial web portion, and an axially extending portion **64**. The armature **62** is formed of a non-magnetic material, and an annular belt receiving groove **66** is defined therein for receiving a belt for transmitting power from the armature to a generator, or other power consuming device. Three armature magnets **68** are mounted on the armature portion **64**, and such magnets are of a configuration similar to the armature magnet configuration of **Figs. 5 through 7**.

The magnets **68** are staggered with respect to each other in a circumferential direction wherein the magnets are not placed exactly 120 degrees apart but instead, a slight angular staggering of the armature magnets is desirable to "smooth out" the magnetic forces being imposed upon the armature as a result of the magnetic forces being simultaneously imposed upon each of the armature magnets. The staggering of the armature magnets **68** in a circumferential direction produces the same effect as the staggering of the armature magnets **40** and **42** as shown in **Fig.8**.

The armature **62** is mounted upon a threaded shaft **70** by anti-friction bearings **72**, and the shaft **70** is threaded into the stator threaded bore **52**, and may be rotated by the knob **74**. In this manner rotation of the knob **74**, and shaft **70**, axially displaces the armature **62** with respect to the stator magnets **58**, and such axial displacement will very the magnitude of the magnetic forces imposed upon the armature magnets **68** by the stator magnets thereby controlling the speed of rotation of the armature. As will be noted from **Figs. 4 to 7, 9 and 10**, an air gap exists between the armature magnets and the stator magnets and the dimension of this spacing, effects the magnitude of the forces imposed upon the armature magnet or magnets. If the distance between the armature magnets and the stator magnets is reduced the forces imposed upon the armature magnets by the stator magnets are increased, and the resultant force **8** vector tending to displace the armature magnets in their path of movement increases. However, the decreasing of the spacing between the armature and stator magnets creates a "pulsation" in the movement of the armature magnets which is objectionable, but can be, to some extent, minimised by using a plurality of armature magnets. Increasing the distance between the armature and stator magnets reduces the pulsation tendency of the armature magnet, but also reduces the magnitude of the magnetic forces imposed upon the armature magnets. Thus, the most effective spacing between the armature and stator magnets is that spacing which produces the maximum force vector in the direction of armature magnet movement, with a minimum creation of objectionable pulsation.

In the disclosed embodiments the high permeability plate **20** and sleeve **56** are disclosed for concentrating the magnetic field of the stator magnets, and the armature magnets are bowed and have shaped ends for magnetic field concentration purposes. While such magnetic field concentration means result in higher forces imposed upon

the armature magnets for given magnet intensities, it is not intended that the inventive concepts be limited to the use of such magnetic field concentrating means.

As will be appreciated from the above description of the invention, the movement of the armature magnet or magnets results from the described relationship of components. The length of the armature magnets as related to the width of the stator magnets and spacing between them, the dimension of the air gap and the configuration of the magnetic field, combined, produce the desired result and motion. The inventive concepts may be practised even though these relationships may be varied within limits not yet defined and the invention is intended to encompass all dimensional relationships which achieve the desired goal of armature movement. By way of example, with respect to **Figs. to 7**, the following dimensions were used in an operating prototype:

The length of armature magnet **24** is 3.125", the stator magnets **12** are 1" wide, .25" thick and 4" long and grain oriented. The air gap between the poles of the armature magnet and the stator magnets is approximately 1.5" and the spacing between the stator magnets is approximately .5" inch.

In effect, the stator magnets define a magnetic field track of a single polarity transversely interrupted at spaced locations by the magnetic fields produced by the lines of force existing between the poles of the stator magnets and the unidirectional force exerted on the armature magnet is a result of the repulsion and attraction forces existing as the armature magnet traverses this magnetic field track.

It is to be understood that the inventive concept embraces an arrangement wherein the armature magnet component is stationary and the stator assembly is supported for movement and constitutes the moving component, and other variations of the inventive concept will be apparent to those skilled in the art without departing from the scope thereof. As used herein the term "track" is intended to include both linear and circular arrangements of the static magnets, and the "direction" or "length" of the track is that direction parallel or concentric to the intended direction of armature magnet movement.

CLAIMS

1. A permanent magnet motor comprising, in combination, a stator track defining a track direction and having first and second sides and composed of a plurality of track permanent magnets each having first and second poles of opposite polarity, said magnets being disposed in side-by-side relationship having a spacing between adjacent magnets and like poles defining said track sides, an elongated armature permanent magnet located on one of said track sides for relative movement thereto and in spaced relationship to said track side wherein an air gap exists between said armature magnet and said track magnets, said armature magnet having first and second poles of opposite polarity located at the opposite ends of said armature magnet deeming the length thereof, the length of said armature magnet being disposed in a direction in general alignment with the direction of said track, the spacing of said armature magnet poles from said track associated side and the length of said armature magnet as related to the width and spacing of said track magnets in the direction of said track being such as to impose a continuous force on said armature magnet in said general direction of said track.

2. In a permanent magnet motor as in claim 1 wherein the spacing between said poles of said armature and the adjacent stator track side are substantially equal.

3. In a permanent magnet motor as in claim 1 wherein the spacing between adjacent track magnets varies.

4. In a permanent magnet motor as in claim 1 wherein a plurality of armature magnets are disposed on a common side of said stator track, said armature magnets being mechanically interconnected.

5. In a permanent magnet motor as in claim 4 wherein said armature magnets are staggered with respect to each other in the direction of said track.

6. In a permanent magnet motor as in claim 1 wherein magnetic field concentrating means are associated with said track magnets.

7. In a permanent magnet motor as in claim 6 wherein said field concentrating means comprises a sheet of magnetic material of high field permeability engaging side and pole of said track opposite to that side and pole disposed toward said armature magnet.

8. In a permanent magnet as in claim 1 wherein said armature magnet is of an arcuate configuration in its longitudinal direction bowed toward said track, said armature magnet having ends shaped to concentrate the magnetic field at said ends.

9. In a permanent magnet motor as in claim 1 wherein said stator track is of a generally linear configuration, and means supporting said armature magnet relative to said track for generally linear movement of said armature magnet.

10. In a permanent magnet motor as in claim 1 wherein said stator track magnets define a circle having an axis, an armature rotatably mounted with respect to said track and concentric and coaxial thereto, said armature magnet being mounted upon said armature.

11. In a permanent magnet motor as in claim 10, means axially adjusting said armature relative to said track whereby the axial relationship of said armature magnet and said stator magnets may be varied to adjust the rate of rotation of said armature.

12. In a permanent magnet motor as in claim 10 wherein a plurality of armature magnets are mounted on said armature.

13. In a permanent magnet motor as in claim 12 wherein said armature magnets are circumferentially non-uniformly spaced on said armature.

14. A permanent magnet motor comprising, in combination, a stator comprising a plurality of circumferentially spaced stator permanent magnets having poles of opposite polarity, said magnets being arranged to substantially define a circle having an axis, the poles of said magnets facing in a radial direction with respect to said axis and poles of the same polarity facing away from said axis and the poles of opposite polarity facing toward said axis, an armature mounted for rotation about said axis and disposed adjacent said stator, at least one armature permanent magnet having poles of opposite polarity mounted on said armature and in radial spaced relationship to said circle of stator magnets, said armature magnet poles extending in the circumferential direction of armature rotation, the spacing of said armature magnet poles from said stator magnets and the circumferential length of said armature magnet and the spacing of said stator magnets being such as to impose a continuing circumferential force on said armature magnet to rotate said armature.

15. In a permanent magnet motor as in claim 14 wherein a plurality of armature magnets are mounted upon said armature.

16. In a permanent magnet motor as in claim 14 wherein said armature magnets are asymmetrically circumferentially spaced on said armature.

17. In a permanent magnet motor as in claim 14 wherein the poles of said armature magnet are shaped to concentrate the magnetic field thereof.

18. In a permanent magnet motor as in claim 14, magnetic field concentrating means associated with said stator magnets concentrating the magnetic fields thereof at the spacings between adjacent stator magnets.

19. In a permanent magnet motor as in claim 18 wherein said magnet field concentrating means comprises an annular ring of high magnetic field permeability material concentric with said axis and in substantial engagement with poles of like polarity of said stator magnets.

20. In a permanent magnet motor as in claim 14 wherein said armature magnet is of an arcuate bowed configuration in the direction of said poles thereof defining a concave side and a convex side, said concave side being disposed toward said axis, and said poles of said armature magnet being shaped to concentrate the magnetic field between said poles thereof.

21. In a permanent magnet motor as in claim 14, means for axially displacing said stator and armature relative to each other to adjust the axial alignment of said stator and armature magnets.

22. The method of producing a unidirectional motive force by permanent magnets using a plurality of spaced stator permanent magnets having opposite polarity poles defining a track having a predetermined direction, and an armature magnet having a length defined by poles of opposite polarity movably mounted for movement relative to the track in the direction thereof, and of a predetermined length determined by the width and dimensions of said stator magnets comprising forming a magnetic field track by said stator magnets having a magnetic field of common polarity interrupted at spaced locations in a direction transverse to the direction of said magnetic field track by magnetic fields created by magnetic lines of force existing between the poles of the stator magnets and positioning the armature magnet in spaced relation to said magnetic field track longitudinally related to the direction of the magnetic field track such a distance that the repulsion and attraction forces imposed on the armature magnet by said magnetic field track imposes a continuing unidirectional force on the armature magnet in the direction of the magnetic field track.

23. The method of producing a unidirectional motive force as in claim 22 including concentrating the magnetic fields created by magnetic lines of force between the poles of the stator magnets.

24. The method of producing a unidirectional motive force as in claim 22 including concentrating the magnetic field existing between the poles of the armature magnet.

25. The method of producing a unidirectional motive force as in claim 22 including concentrating the magnetic fields created by magnetic lines of force between the poles of the stator magnets and concentrating the 12 magnetic field existing between the poles of the armature magnet.

26. The method of producing a motive force by permanent magnets wherein the unpaired electron spinning particles existing within a permanent magnet are utilised for producing a motive force comprising forming a stator magnetic field track by means of at least one permanent magnet, producing an armature magnetic field by means of a permanent magnet and shaping and locating said magnetic fields in such a manner as to produce relative continuous unidirectional motion between said stator and armature field producing magnets.

27. The method of producing a motive force by permanent magnets as in claim 26 wherein said stator magnetic field is substantially of a single polarity.

28. The method of producing a motive force by permanent magnets as in claim 26 including concentrating the magnetic field of said stator field track and armature magnetic field.

HAROLD EWING: THE CAROUSEL ELECTRIC GENERATOR

US Patent 5,625,241 **29th April 1997** **Inventor: Harold E. Ewing et al.**

CAROUSEL ELECTRIC GENERATOR

This is a reworded excerpt form this patent which shows a compact, self-powered, combined permanent magnet motor and electrical generator. There is a little extra information at the end of this document.

ABSTRACT

A permanent magnet generator or motor having stationary coils positioned in a circle, a rotor on which are mounted permanent magnets grouped in sectors and positioned to move adjacent to the coils, and a carousel carrying corresponding groups of permanent magnets through the centres of the coils, the carousel movies with the rotor by virtue of its being magnetically coupled to it.

Inventors:
Ewing, Harold E. (Chandler, AZ, US)
Chapman, Russell R. (Mesa, AZ, US)
Porter, David R. (Mesa, AZ, US)

Assignee:
Energy Research Corporation (Mesa, AZ)

US Patent References:

3610974	Oct, 1971	Kenyon	310/49.
4547713	Oct, 1985	Langley et al.	318/254.
5117142	May, 1992	Von Zweygbergk	310/156.
5289072	Feb, 1994	Lange	310/266.
5293093	Mar, 1994	Warner	310/254.
5304883	Apr, 1994	Denk	310/180.

BACKGROUND OF THE INVENTION

There are numerous applications for small electric generators in ratings of a few kilowatts or less. Examples include electric power sources for emergency lighting in commercial and residential buildings, power sources for remote locations such as mountain cabins, and portable power sources for motor homes, pleasure boats, etc.

In all of these applications, system reliability is a primary concern. Because the power system is likely to sit idle for long periods of time without the benefit of periodic maintenance, and because the owner-operator is often inexperienced in the maintenance and operation of such equipment, the desired level of reliability can only be achieved through system simplicity and the elimination of such components as batteries or other secondary power sources which are commonly employed for generator field excitation.

Another important feature for such generating equipment is miniaturisation particularly in the case of portable equipment. It is important to be able to produce the required level of power in a relatively small generator.

Both of these requirements are addressed in the present invention through a novel adaptation of the permanent magnet generator or magneto in a design that lends itself to high frequency operation as a means for maximising power output per unit volume.

DESCRIPTION OF THE PRIOR ART

Permanent magnet generators or magnetos have been employed widely for many years. Early applications of such generators include the supply of electric current for spark plugs in automobiles and aeroplanes. Early telephones used magnetos to obtain electrical energy for ringing. The Model T Ford automobile also used magnetos to power its electric lights.

The present invention differs from prior art magnetos in terms of its novel physical structure in which a multiplicity of permanent magnets and electrical windings are arranged in a fashion which permits high-speed/high-frequency operation as a means for meeting the miniaturisation requirement. In addition, the design is enhanced through the

use of a rotating carousel which carries a multiplicity of field source magnets through the centres of the stationary electric windings in which the generated voltage is thereby induced.

SUMMARY OF THE INVENTION

In accordance with the invention claimed, an improved permanent magnet electric generator is provided with a capability for delivering a relatively high level of output power from a small and compact structure. The incorporation of a rotating carousel for the transport of the primary field magnets through the electrical windings in which induction occurs enhances field strength in the locations critical to generation.

It is, therefore, one object of this invention to provide an improved permanent magnet generator or magneto for the generation of electrical power. Another object of this invention is to provide in such a generator a relatively high level of electrical power from a small and compact structure. A further object of this invention is to achieve such a high level of electrical power by virtue of the high rotational speed and high frequency operation of which the generator of the invention is capable.

A further object of this invention is to provide such a high frequency capability through the use of a novel field structure in which the primary permanent magnets are carried through the centres of the induction windings of the generator by a rotating carousel.

A still further object of this invention is to provide a means for driving the rotating carousel without the aid of mechanical coupling but rather by virtue of magnetic coupling between other mechanically driven magnets and those mounted on the carousel.

A still further object of this invention is to provide an enhanced capability for high speed/high frequency operation through the use of an air bearing as a support for the rotating carousel.

Yet another object of this invention is to provide in such an improved generator a sufficiently high magnetic field density in the locations critical to voltage generation without resort to the use of laminations or other media to channel the magnetic field.

Further objects an advantages of the invention will become apparent as the following description proceeds and the features of novelty which characterise the invention will be pointed out with particularity in the claims annexed to and forming a part of this specification.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention may be more readily described by reference to the accompanying drawings, in which:

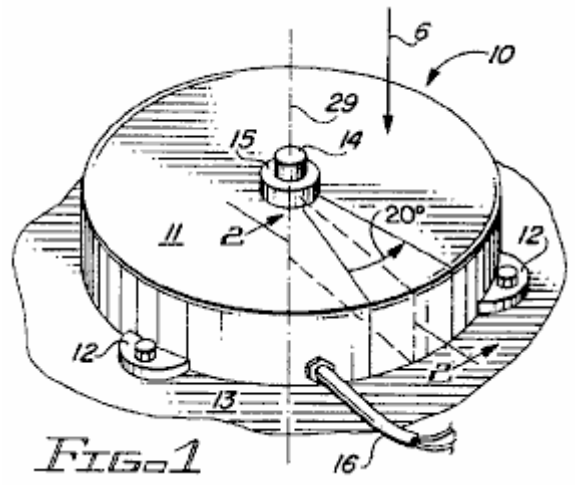


FIG-1

Fig.1 is a simplified perspective view of the carousel electric generator of the invention;

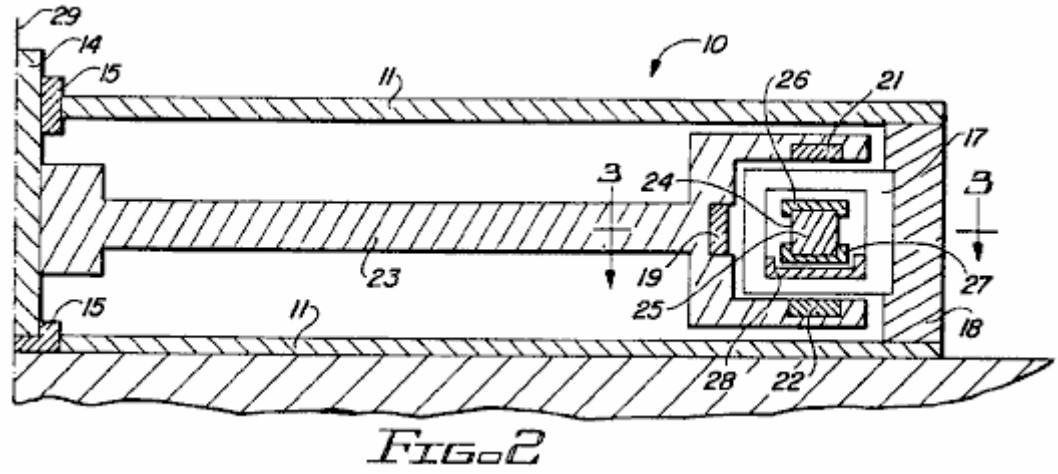


FIG. 2

Fig.2 is a cross-sectional view of **Fig.1** taken along line **2--2**;

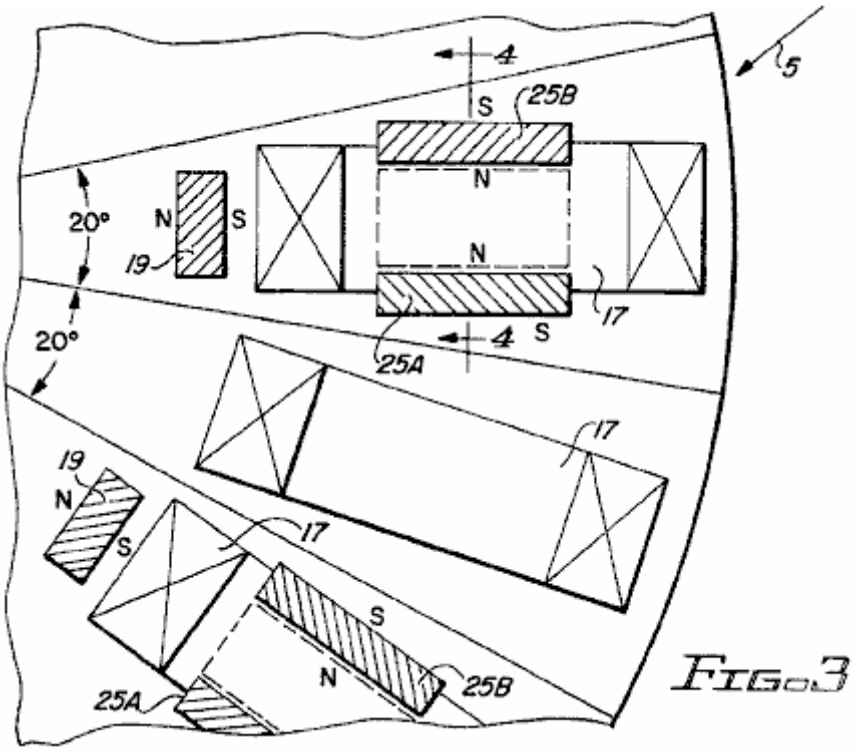


FIG. 3

Fig.3 is a cross-sectional view of the generator of **Fig.1** and **Fig.2** taken along line **3--3** of **Fig.2**;

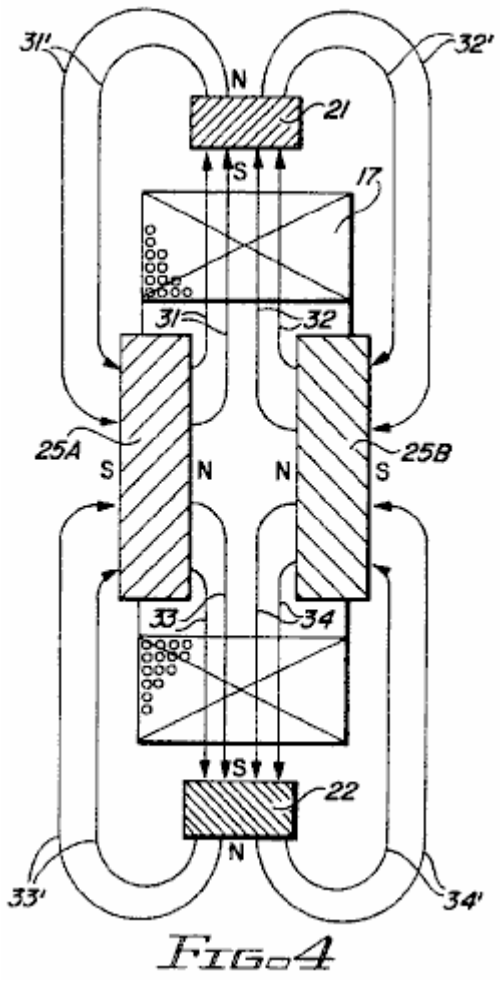


FIG-4

Fig.4 is a cross-sectional view of **Fig.3** taken along line **4--4**;

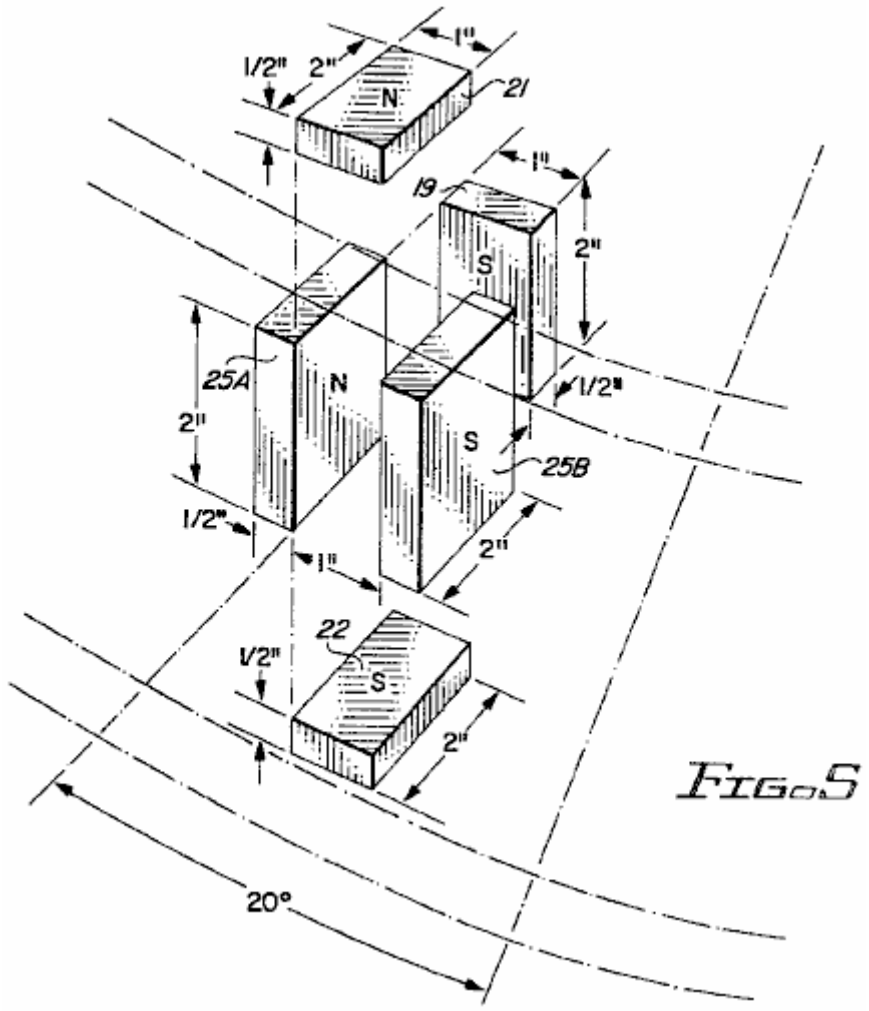


FIG. 5

Fig.5 is a partial perspective view showing the orientation of a group of permanent magnets within a twenty degree sector of the generator of the invention as viewed in the direction of arrow **5** of **Fig.3**;

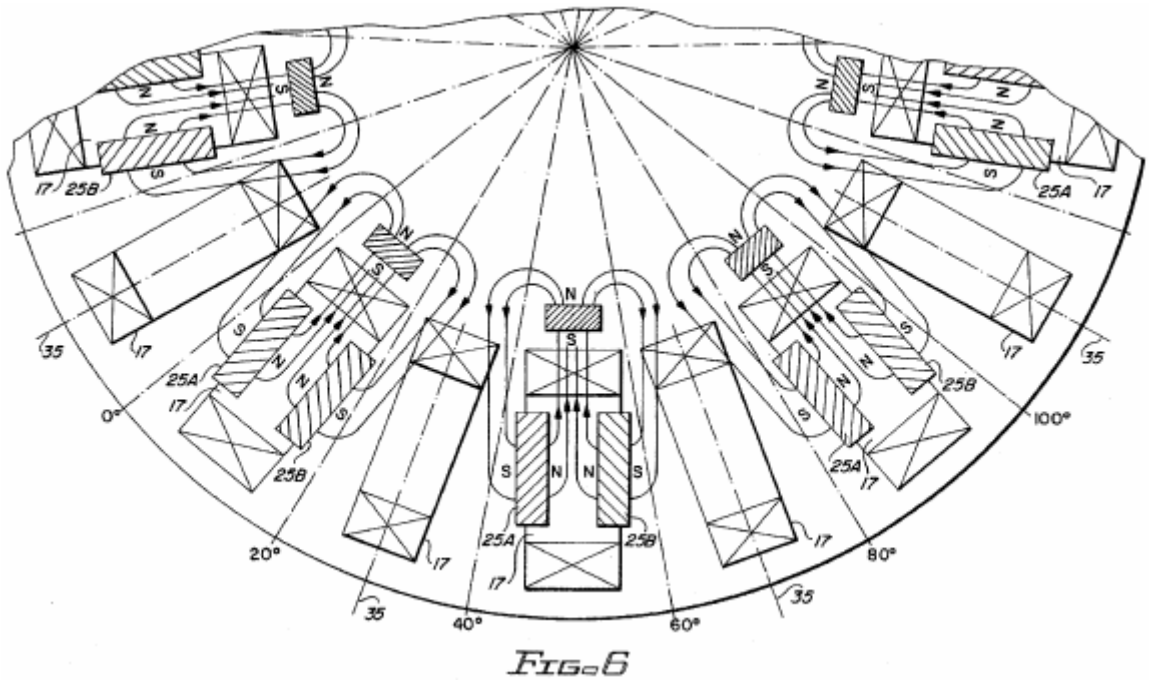


FIG. 6

Fig.6 is an illustration of the physical arrangement of electrical windings and permanent magnets within the generator of the invention as viewed in the direction of arrow **6** in **Fig.1**;

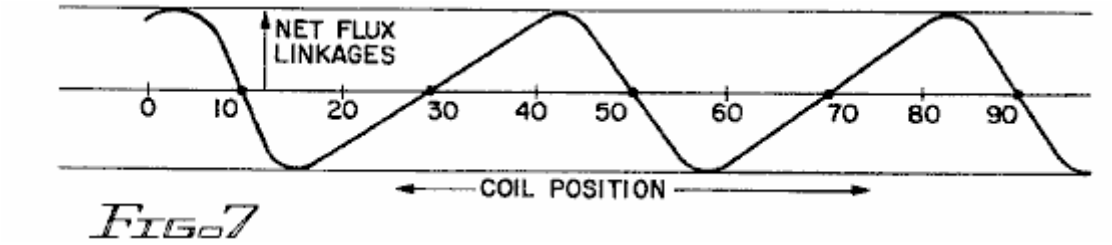


FIG. 7

Fig.7 is a wave form showing flux linkages for a given winding as a function of rotational position of the winding relative to the permanent magnets;

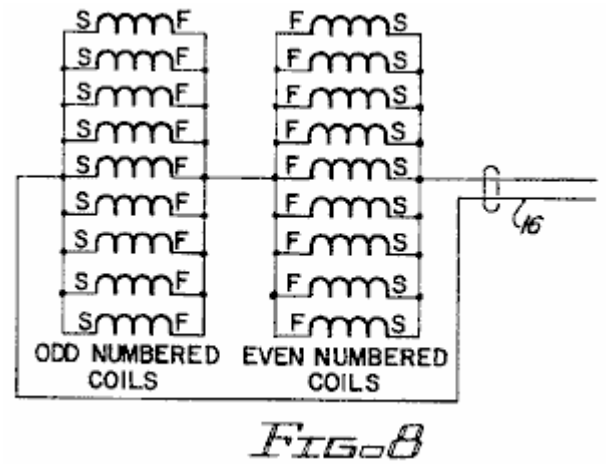


FIG. 8

Fig.8 is a schematic diagram showing the proper connection of the generator windings for a high current low voltage configuration of the generator;

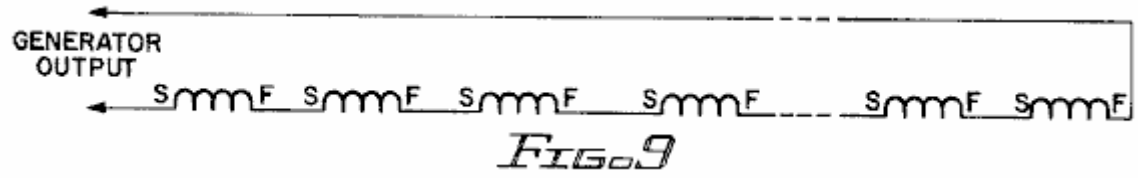


FIG. 9

Fig.9 is a schematic diagram showing a series connection of generator coils for a low current, high voltage configuration;

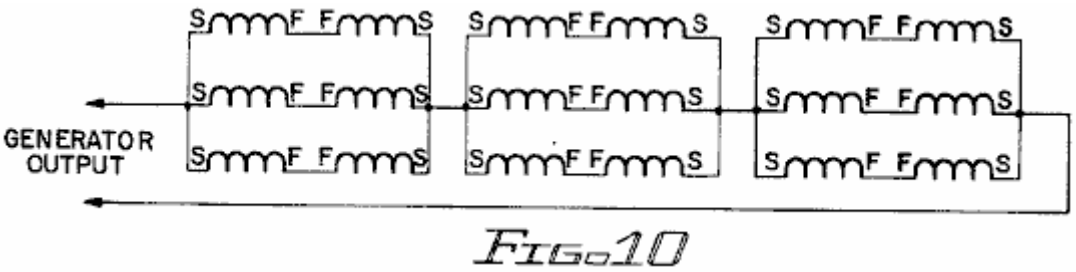


FIG. 10

Fig.10 is a schematic diagram showing a series/parallel connection of generator windings for intermediate current and voltage operation;

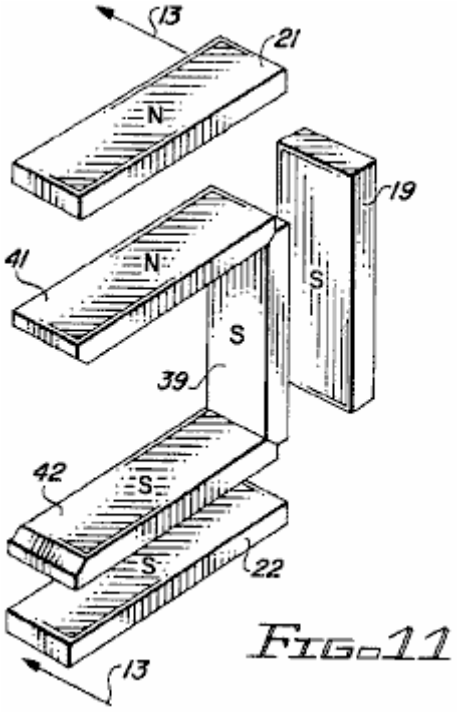


FIG. 11

Fig.11 is a perspective presentation of a modified carousel magnet configuration employed in a second embodiment of the invention;

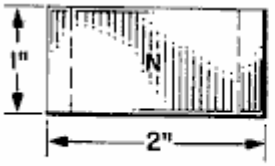


FIG. 12A

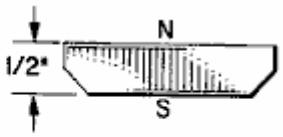


FIG. 12B

Fig.12A and **Fig.12B** show upper and lower views of the carousel magnets of **Fig.11**;

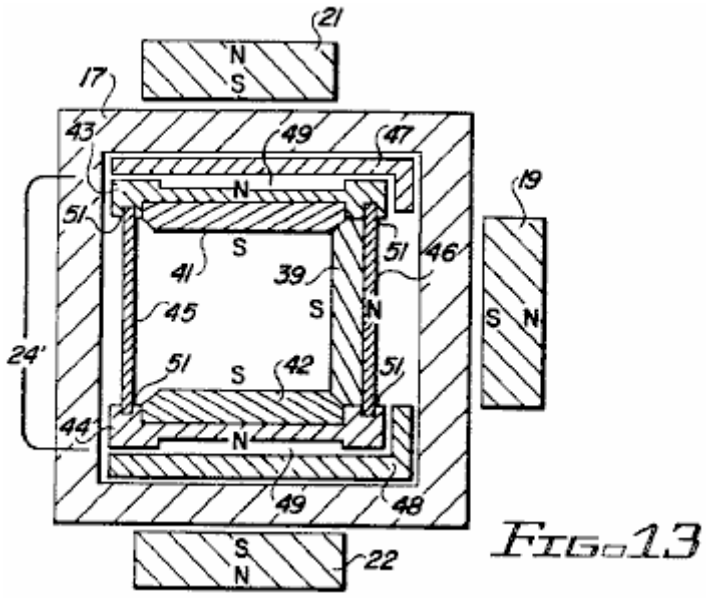


FIG. 13

Fig.13 is a cross-sectional view of the modified magnet configuration of **Fig.11** taken along line **13--13** with other features of the modified carousel structure also shown;

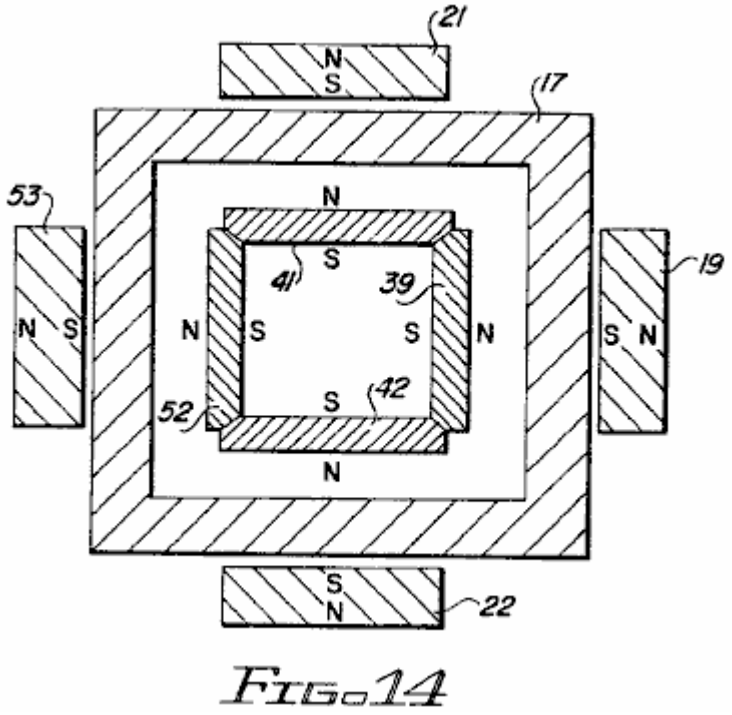


FIG. 14

Fig.14 is a modification of the carousel structure shown in **Figs. 1-13** wherein a fourth carousel magnet is positioned at each station; and

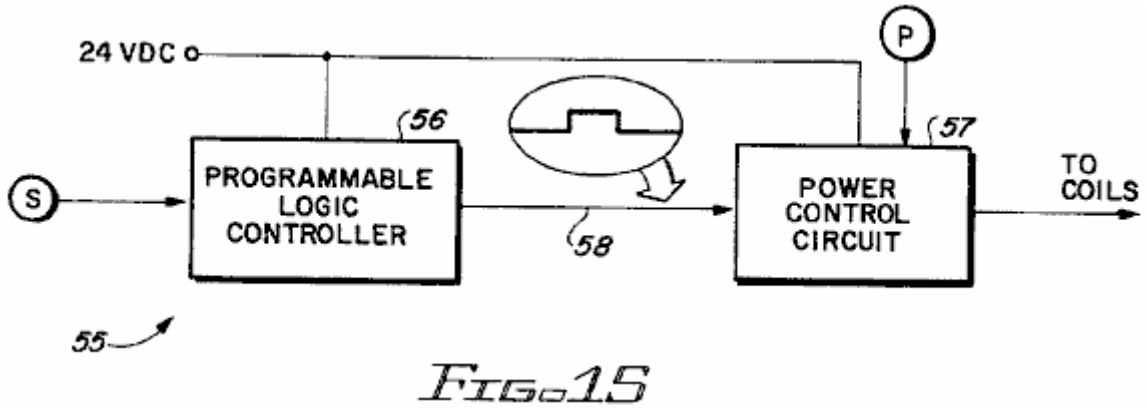


FIG. 15

Fig.15 illustrates the use of the claimed device as a pulsed direct current power source.

DESCRIPTION OF THE PREFERRED EMBODIMENT

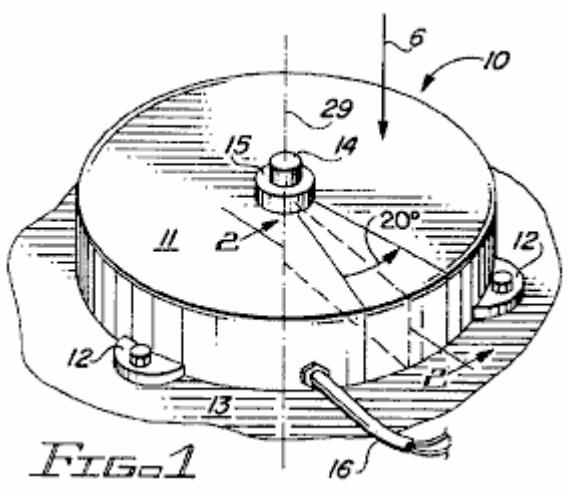


FIG. 1

Referring more particularly to the drawings by characters of reference, **Fig.1** shows the external proportions of a carousel electric generator **10** of the invention. As shown in **Fig.1**, generator **10** is enclosed by a housing **11** with mounting feet **12** suitable for securing the generator to a flat surface **13**. The surface **13** is preferably horizontal, as shown in **Fig.1**.

Housing **11** has the proportions of a short cylinder. A drive shaft **14** extends axially from housing **11** through a bearing **15**. The electrical output of the generator is brought out through a cable **16**.

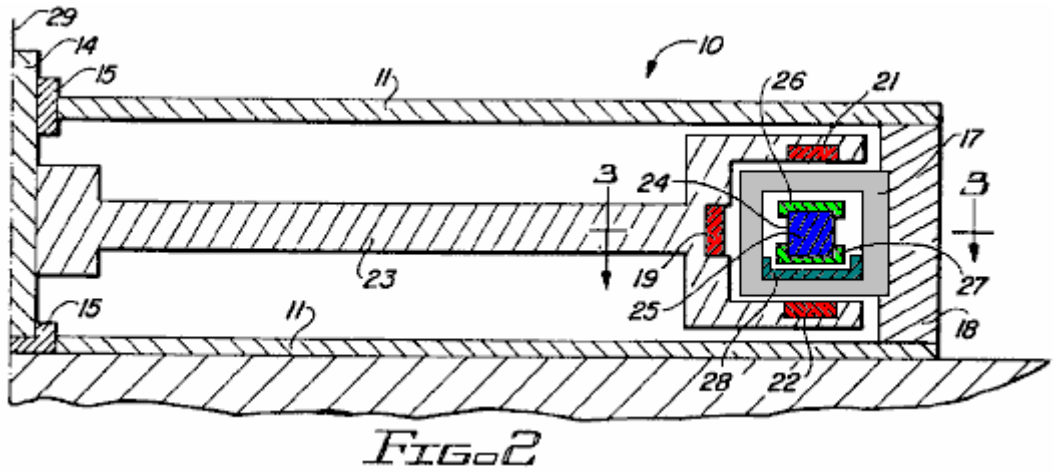


FIG. 2

The cross-sectional view of **Fig.2** shows the active elements incorporated in one twenty degree sector of the stator and in one twenty degree sector of the rotor.

In the first implementation of the invention, there are eighteen identical stator sectors, each incorporating a winding or coil **17** wound about a rectangular coil frame or bobbin. Coil **17** is held by a stator frame **18** which may also serve as an outer wall of frame **11**.

The rotor is also divided into eighteen sectors, nine of which incorporate three permanent magnets each, including an inboard rotor magnet **19**, an upper rotor magnet **21** and a lower rotor magnet **22**. All three of these magnets have their south poles facing coil **17**, and all three are mounted directly on rotor frame **23** which is secured directly to drive shaft **14**.

The other nine sectors of the rotor are empty, i.e. they are not populated with magnets. The unpopulated sectors are alternated with the populated sectors so that adjacent populated sectors are separated by an unpopulated sector as shown in **Fig.3** and **Fig.6**.

With reference again to **Fig.2**, generator **10** also incorporates a carousel **24**. The carousel comprises nine pairs of carousel magnets **25** clamped between upper and lower retainer rings **26** and **27**, respectively. The lower retainer ring **27** rests inside an air bearing channel **28** which is secured to stator **18** inside the bobbin of coil **17**. Air passages (not shown) admit air into the space between the lower surface of ring **27** and the upper or inside

surface of channel **28**. This arrangement comprises an air bearing which permits carousel **24** to rotate freely within the coils **17** about rotational axis **29** of rotor frame **23**.

Carousel **24** is also divided into **18** twenty-degree sectors, including nine populated sectors interspersed with nine unpopulated sectors in an alternating sequence. Each of the nine populated sectors incorporates a pair of carousel magnets as described in the preceding paragraph.

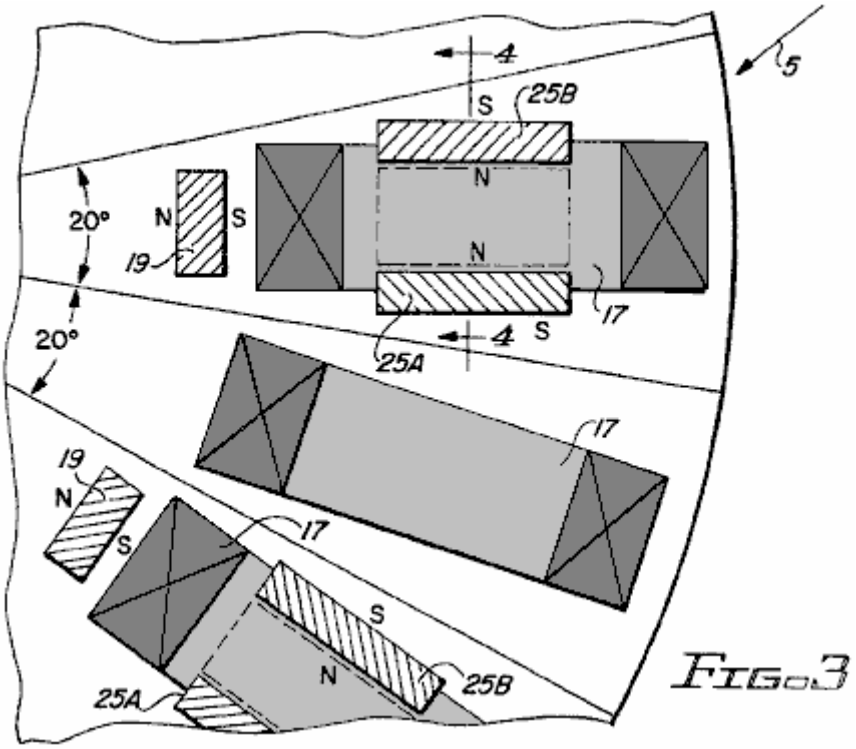


FIG. 3

The geometrical relationship between the rotor magnets, the carousel magnets and the coils, is further clarified by **Fig.3**, **Fig.4** and **Fig.5**. In each of the three figures, the centre of each populated rotor sector is shown aligned with the centre of a coil **17**. Each populated carousel sector, which is magnetically locked into position with a populated rotor sector, is thus also aligned with a coil **17**.

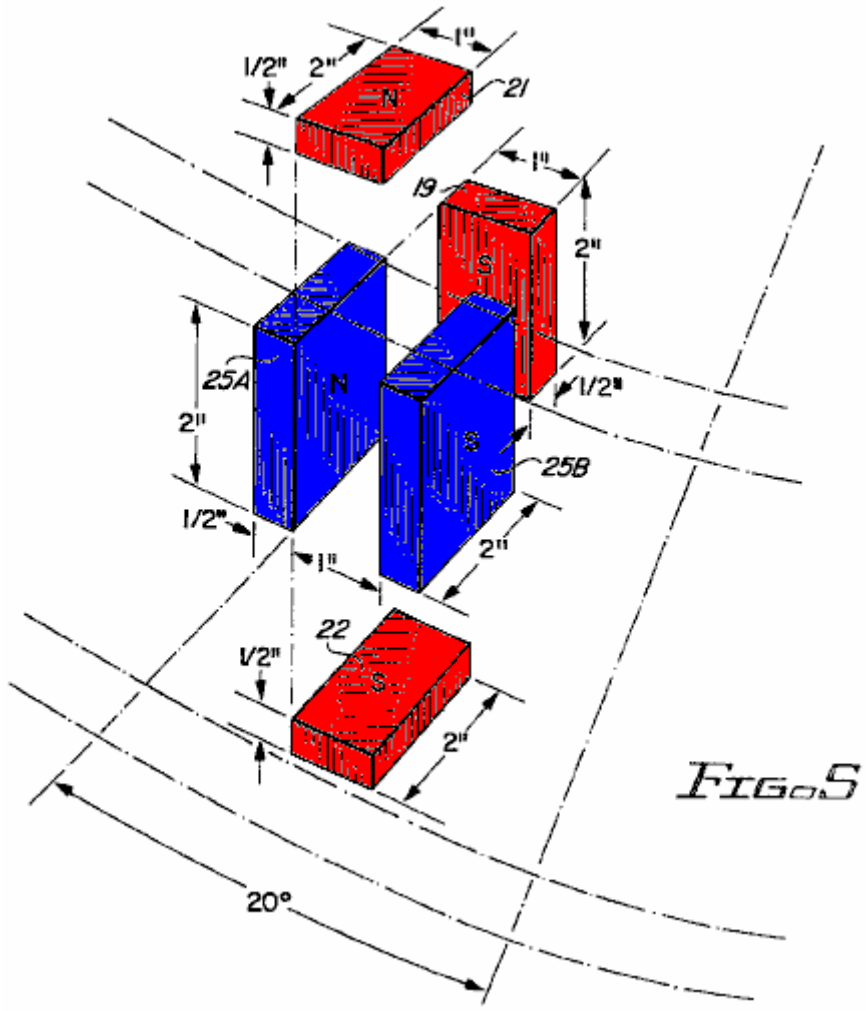


Fig.5

In an early implementation of the invention, the dimensions and spacings of the rotor magnets **19**, **21** and **22** and carousel magnets, **25A** and **25B** of carousel magnet pairs **25** were as shown in **Fig.5**. Each of the rotor magnets **19**, **21** and **22** measured one inch by two inches by one-half inch with north and south poles at opposite one-inch by two-inch faces. Each of the carousel magnets **25A** and **25B** measured two inches by two inches by one-half inch with north and south poles at opposite two-inch by two-inch faces. The magnets were obtained from Magnet Sales and Manufacturing, Culver City, Calif. The carousel magnets were part No.35NE2812832; the rotor magnets were custom parts of equivalent strength (MMF) but half the cross section of the carousel magnets.

Coil supports and other stationary members located within magnetic field patterns are fabricated from Delrin or Teflon plastic or equivalent materials. The use of aluminium or other metals introduce eddy current losses and in some cases excessive friction.

As shown in **Fig.5**, carousel magnets **25A** and **25B** stand on edge, parallel with each other, their north poles facing each other, and spaced one inch apart. When viewed from directly above the carousel magnets, the space between the two magnets **25A** and **25B** appears as a one-inch by two-inch rectangle. When the carousel magnet pair **25** is perfectly locked into position magnetically, upper rotor magnet **21** is directly above this one-inch by two-inch rectangle, lower rotor magnet **22** is directly below it, and their one-inch by two-inch faces are directly aligned with it, the south poles of the two magnets **21** and **22** facing each other.

In like manner, when viewed from the axis of rotation of generator **10**, the space between carousel magnets **25A** and **25B** again appears as a one-inch by two-inch rectangle, and this rectangle is aligned with the one-inch by two-inch face of magnet **19**, the south pole of magnet **19** facing the carousel magnet pair **25**.

Rotor magnets **19**, **21** and **22** are positioned as near as possible to carousel magnets **25A** and **25B** while still allowing passage for coil **17** over and around the carousel magnets and through the space between the carousel magnets and the rotor magnets.

In an electric generator, the voltage induced in the generator windings is proportional to the product of the number of turns in the winding and the rate of change of flux linkages that is produced as the winding is rotated through

the magnetic field. An examination of magnetic field patterns is therefore essential to an understanding of generator operation.

In generator **10**, magnetic flux emanating from the north poles of carousel magnets **25A** and **25B** pass through the rotor magnets and then return to the south poles of the carousel magnets. The total flux field is thus driven by the combined MMF (magnetomotive force) of the carousel and field magnets while the flux patterns are determined by the orientation of the rotor and carousel magnets.

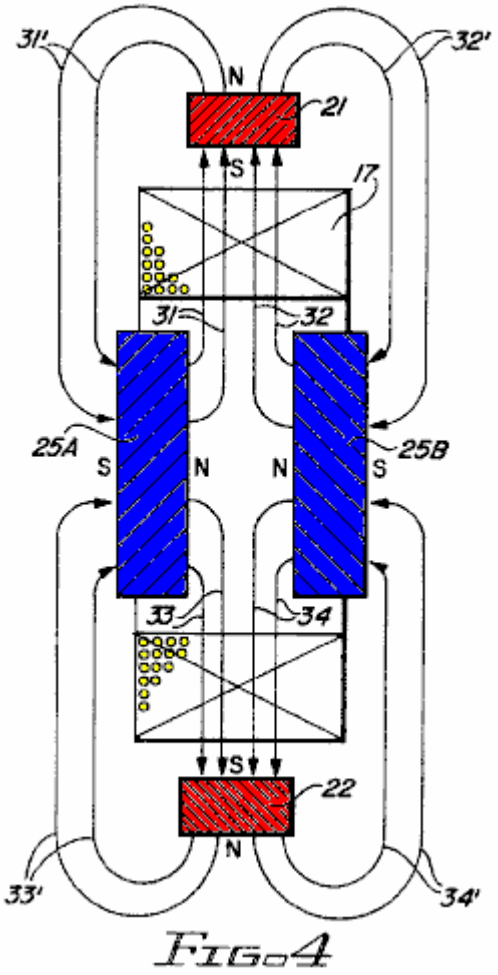


FIG. 4

The flux pattern between carousel magnets **25A** and **25B** and the upper and lower rotor magnets **21** and **22** is illustrated in **Fig.4**. Magnetic flux lines **31** from the north pole of carousel magnet **25A** extend to the south pole of upper rotor magnet **21**, pass through magnet **21** and return as lines **31'** to the south pole of magnet **25A**. Lines **33**, also from the north pole of magnet **25A** extend to the south pole of lower rotor magnet **22**, pass through magnet **22** and return to the south pole of magnet **25A** as lines **33'**. Similarly, lines **32** and **34** from the north pole of magnet **25B** pass through magnets **21** and **22**, respectively, and return as lines **32'** and **34'** to the south pole of magnet **25B**. Flux linkages produced in coil **17** by lines emanating from carousel magnet **25A** are of opposite sense from those emanating from carousel magnet **25B**. Because induced voltage is a function of the rate of change in net flux linkages, it is important to recognise this difference in sense.

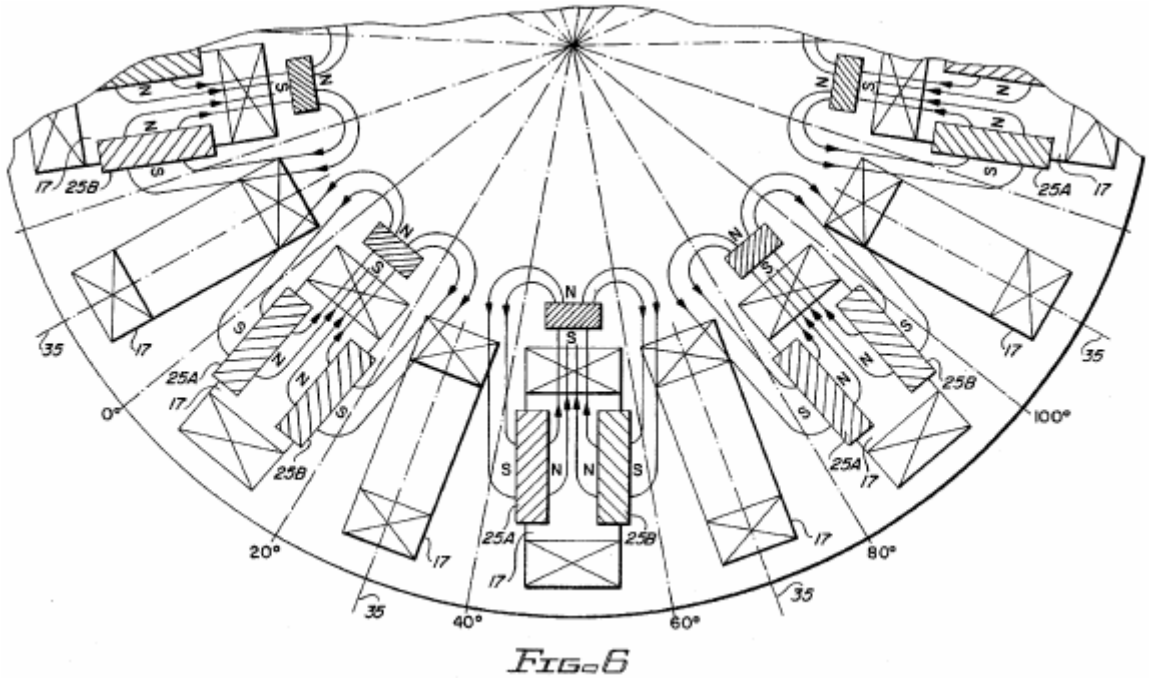


FIG. 6

Fig.6 shows a similar flux pattern for flux between carousel magnets **25A** and **25B** and inboard rotor magnet **19**. Again the lines emanating from carousel magnet **25A** and passing through rotor magnet **19** produce flux linkages in coil **17** that are opposite in sense from those produced by lines from magnet **25B**.

The arrangement of the carousel magnets with the north poles facing each other tends to confine and channel the flux into the desired path. This arrangement replaces the function of magnetic yokes or laminations of more conventional generators.

The flux linkages produced by magnets **25A** and **25B** are opposite in sense regardless of the rotational position of coil **17** including the case where coil **17** is aligned with the carousel and rotor magnets as well as for the same coils when they are aligned with an unpopulated rotor sector.

Taking into account the flux patterns of **Fig.4** and **Fig.6** and recognising the opposing sense conditions just described, net flux linkages for a given coil **17** are deduced as shown in **Fig.7**.

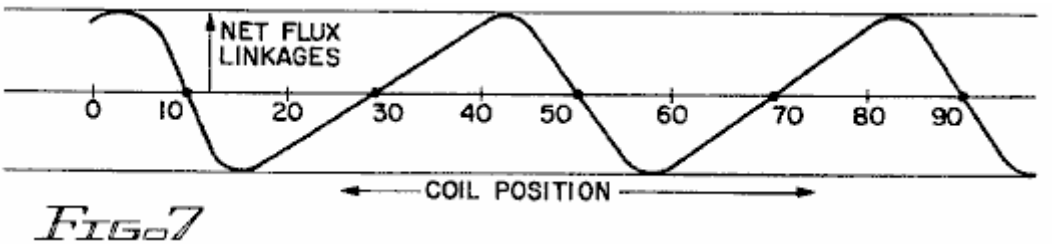


FIG. 7

In **Fig.7**, net flux linkages (coil-turns x lines) are plotted as a function of coil position in degrees. Coil position is here defined as the position of the centreline **35** of coil **17** relative to the angular scale shown in degrees in **Fig.6**. (Note that the coil is stationary and the scale is fixed to the rotor. As the rotor turns in a clockwise direction, the relative position of coil **17** progresses from zero to ten to twenty degrees etc.).

At a relative coil position of ten degrees, the coil is centred between magnets **25A** and **25B**. Assuming symmetrical flux patterns for the two magnets, the flux linkages from one magnet exactly cancel the flux linkages from the other so that net flux linkages are zero. As the relative coil position moves to the right, linkages from magnet **25A** decrease and those from magnet **25B** increase so that net flux linkages build up from zero and passes through a maximum negative value at some point between ten and twenty degrees. After reaching the negative maximum, flux linkages decrease, passing through zero at 30 degrees (where coil **17** is at the centre of an unpopulated rotor sector) and then rising to a positive maximum at some point just beyond 60 degrees. This cyclic variation repeats as the coil is subjected successively to fields from populated and unpopulated rotor sectors.

As the rotor is driven rotationally, net flux linkages for all eighteen coils are altered at a rate that is determined by the flux pattern just described in combination with the rotational velocity of the rotor. Instantaneous voltage

induced in coil **17** is a function of the slope of the curve shown in **Fig.7** and rotor velocity, and voltage polarity changes as the slope of the curve alternates between positive and negative.

It is important to note here that a coil positioned at ten degrees is exposed to a negative slope while the adjacent coil is exposed to a positive slope. The polarities of the voltages induced in the two adjacent coils are therefore opposite. For series or parallel connections of odd and even-numbered coils, this polarity discrepancy can be corrected by installing the odd and even numbered coils oppositely (odds rotated end for end relative to evens) or by reversing start and finish connections of odd relative to even numbered coils. Either of these measures will render all coil voltages additive as needed for series or parallel connections. Unless the field patterns for populated and unpopulated sectors are very nearly symmetrical, however, the voltages induced in odd and even numbered coils will have different waveforms. This difference will not be corrected by the coil reversals or reverse connections discussed in the previous paragraph. Unless the voltage waveforms are very nearly the same, circulating currents will flow between even and odd-numbered coils. These circulating currents will reduce generator efficiency.

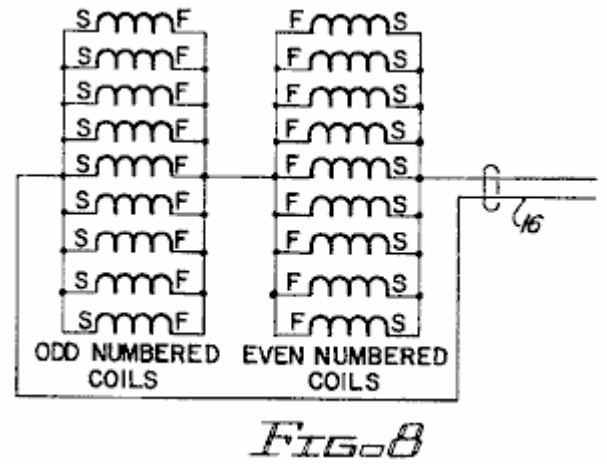


FIG. 8

To prevent such circulating currents and the attendant loss in operating efficiency for non symmetrical field patterns and unmatched voltage waveforms, the series-parallel connections of **Fig.8** may be employed in a high-current, low-voltage configuration of the generator. If the eighteen coils are numbered in sequence from one to eighteen according to position about the stator, all even-numbered coils are connected in parallel, all odd-numbered coils are connected in parallel, and the two parallel coil groups are connected in series as shown with reversed polarity for one group so that voltages will be in phase relative to output cable **16**.

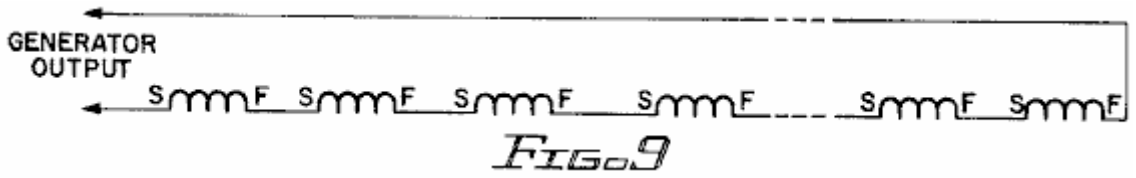


FIG. 9

For a low-current, high voltage configuration, the series connection of all coils may be employed as shown in **Fig.9**. In this case, it is only necessary to correct the polarity difference between even and odd numbered coils. As mentioned earlier, this can be accomplished by means of opposite start and finish connections for odd and even coils or by installing alternate coils reversed, end for end.

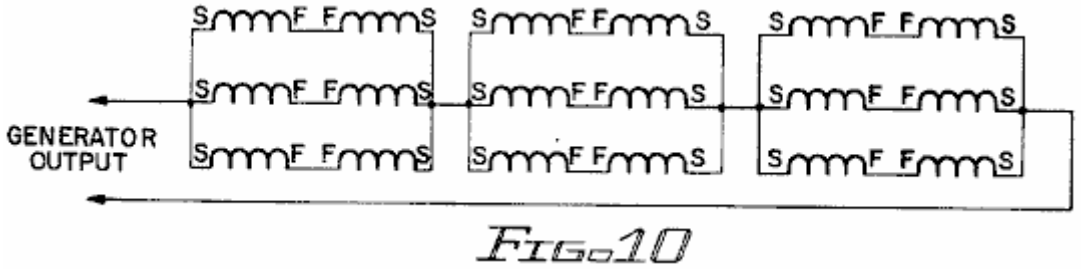


FIG. 10

For intermediate current and voltage configurations, various series-parallel connections may be employed. **Fig.10**, for example, shows three groups of six coils each connected in series. Circulating currents will be avoided so long as even-numbered coils are not connected in parallel with odd-numbered coils. Parallel connection of

series-connected odd/even pairs as shown is permissible because the waveforms of the series pairs should be very neatly matched.

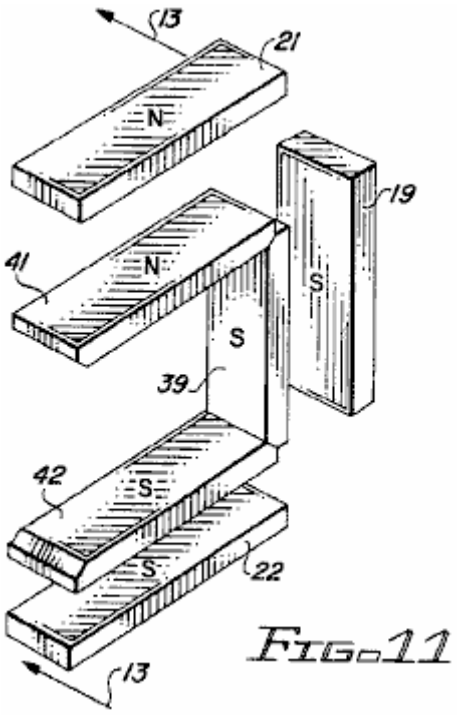


FIG. 11

In another embodiment of the invention, the two large (two-inch by two-inch) carousel magnets are replaced by three smaller magnets as shown in **Fig.11**, **Fig.12** and **Fig.13**. The three carousel magnets comprise an inboard carousel magnet **39**, an upper carousel magnet **41** and a lower carousel magnet **42** arranged in a U-shaped configuration that matches the U-shaped configuration of the rotor magnets **19**, **21** and **22**. As in the case of the first embodiment, the rotor and carousel magnets are present only in alternate sectors of the generator.

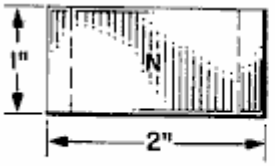


FIG. 12A

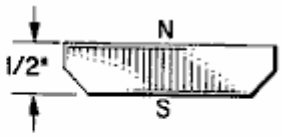


FIG. 12B

The ends of the carousel magnets are bevelled to permit a more compact arrangement of the three magnets. As shown in **Fig.12**, each magnet measures one inch by two inches by one half inch thick. The south pole occupies the bevelled one-inch by two-inch face and the north pole is at the opposite face.

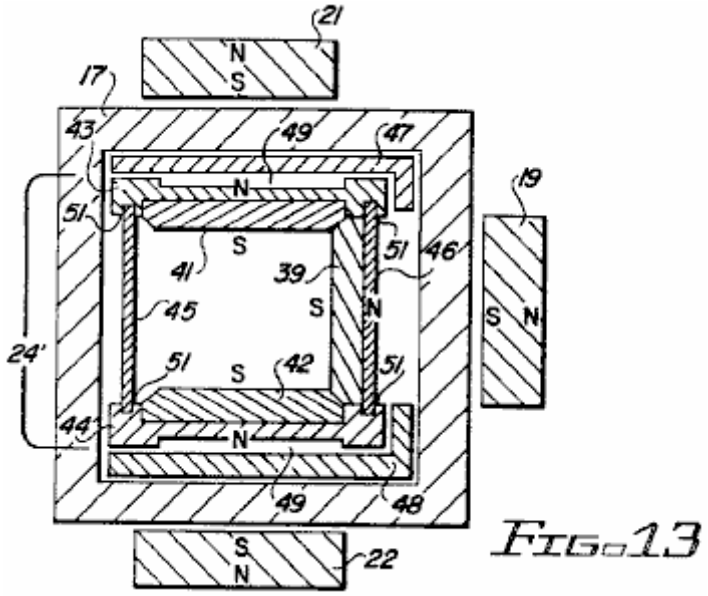


FIG. 13

The modified carousel structure **24'** as shown in **Fig.13** comprises an upper carousel bearing plate **43**, a lower carousel bearing plate **44**, an outer cylindrical wall **45** and an inner cylindrical wall **46**. The upper and lower bearing plates **43** and **44** mate with the upper and lower bearing members **47** and **48**, respectively, which are stationary and secured inside the forms of the coils **17**. Bearing plates **43** and **44** are shaped to provide air channels **49** which serve as air bearings for rotational support of the carousel **24'**. The bearing plates are also slotted to receive the upper and lower edges **51** of cylindrical walls **45** and **46**.

The modified carousel structure **24'** offers a number of advantages over the first embodiment. The matched magnet configuration of the carousel and the rotor provides tighter and more secure coupling between the carousel and the rotor. The smaller carousel magnets also provide a significant reduction in carousel weight. This was found beneficial relative to the smooth and efficient rotational support of the carousel.

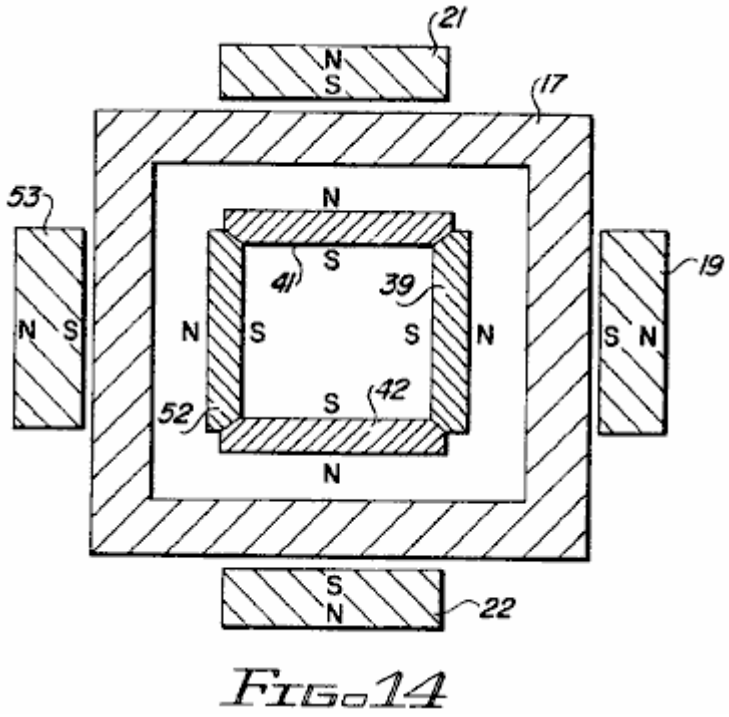


FIG.14

The modification of the carousel structure as described in the foregoing paragraphs can be taken one step further with the addition of a fourth carousel magnet **52** at each station as shown in **Fig.14**. The four carousel magnets **39**, **41**, **42** and **52** now form a square frame with each of the magnet faces (north poles) facing a corresponding inside face of the coil **17**. Carousel magnets for this modification may again be as shown in **Fig.12**. An additional rotor magnet **53** may also be added as shown, in alignment with carousel magnet **52**. These additional modifications further enhance the field pattern and the degree of coupling between the rotor and the carousel.

The carousel electric generator of the invention is particularly well suited to high speed, high frequency operation where the high speed compensates for lower flux densities than might be achieved with a magnetic medium for routing the field through the generator coils. For many applications, such as emergency lighting, the high frequency is also advantageous. Fluorescent lighting, for example, is more efficient in terms of lumens per watt and the ballasts are smaller at high frequencies.

While the present invention has been directed toward the provision of a compact generator for specialised generator applications, it is also possible to operate the device as a motor by applying an appropriate alternating voltage source to cable **16** and coupling drive shaft **14** to a load.

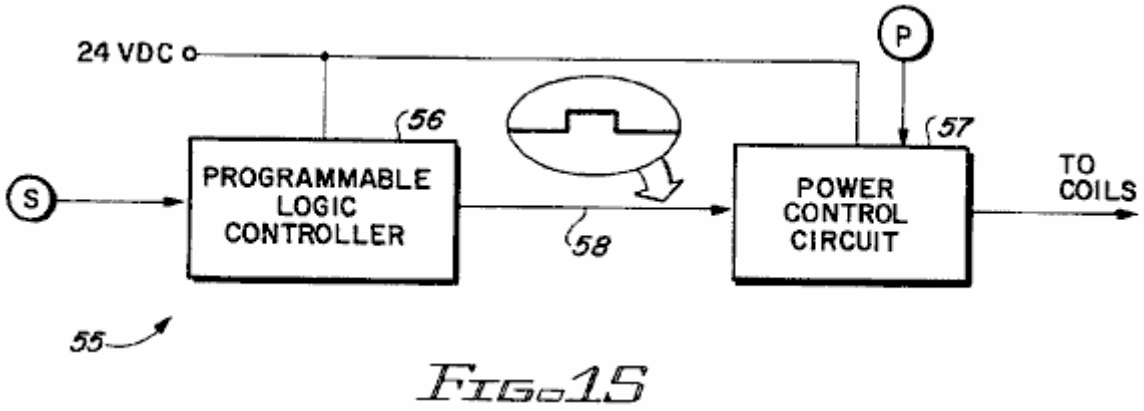


FIG. 15

It is also possible to operate the device of the invention as a motor using a pulsed direct-current power source. A control system **55** for providing such operation is illustrated in **Fig.15**. Incorporated in the control system **55** are a rotor position sensor **S**, a programmable logic controller **56**, a power control circuit **57** and a potentiometer **P**.

Based on signals received from sensor **S**, controller **56** determines the appropriate timing for coil excitation to assure maximum torque and smooth operation. This entails the determination of the optimum positions of the rotor and the carousel at the initiation and at the termination of coil excitation. For smooth operation and maximum torque, the force developed by the interacting fields of the magnets and the excited coils should be unidirectional to the maximum possible extent.

Typically, the coil is excited for only 17.5 degrees or less during each 40 degrees of rotor rotation.

The output signal **58** of controller **56** is a binary signal (high or low) that is interpreted as an ON and OFF command for coil excitation.

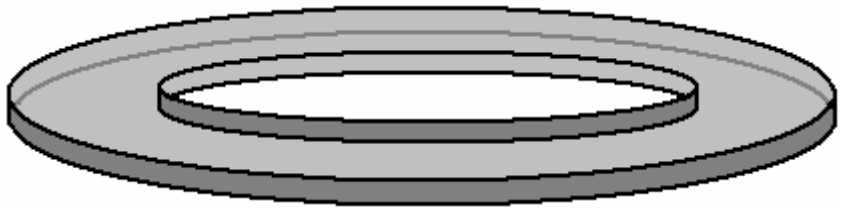
The power control circuit incorporates a solid state switch in the form of a power transistor or a MOSFET. It responds to the control signal **58** by turning the solid state switch ON and OFF to initiate and terminate coil excitation. Instantaneous voltage amplitude supplied to the coils during excitation is controlled by means of potentiometer **P**. Motor speed and torque are thus responsive to potentiometer adjustments.

The device is also adaptable for operation as a motor using a commutator and brushes for control of coil excitation. In this case, the commutator and brushes replace the programmable logic controller and the power control circuit as the means for providing pulsed DC excitation. This approach is less flexible but perhaps more efficient than the programmable control system described earlier.

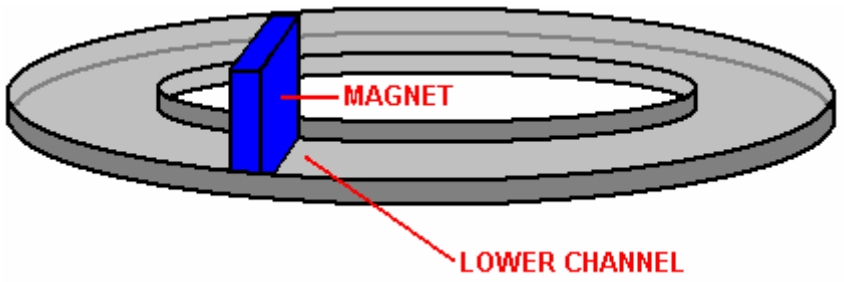
It will now be recognised that a novel and useful generator has been provided in accordance with the stated objects of the invention, and while but a few embodiments of the invention have been illustrated and described it will be apparent to those skilled in the art that various changes and modifications may be made without departing from the spirit of the invention or from the scope of the appended claims.

Notes:

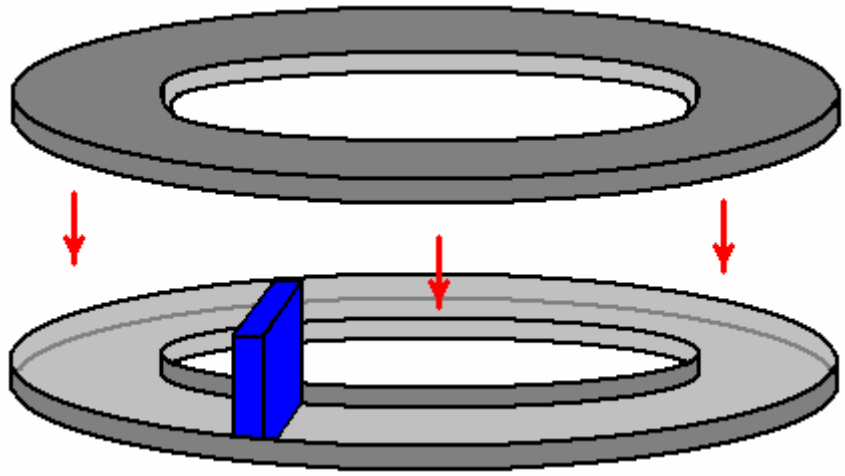
I found it a little difficult to visualise the carousel part, so the following may be helpful for some people. The "carousel" is formed from two circular plastic channels like this:



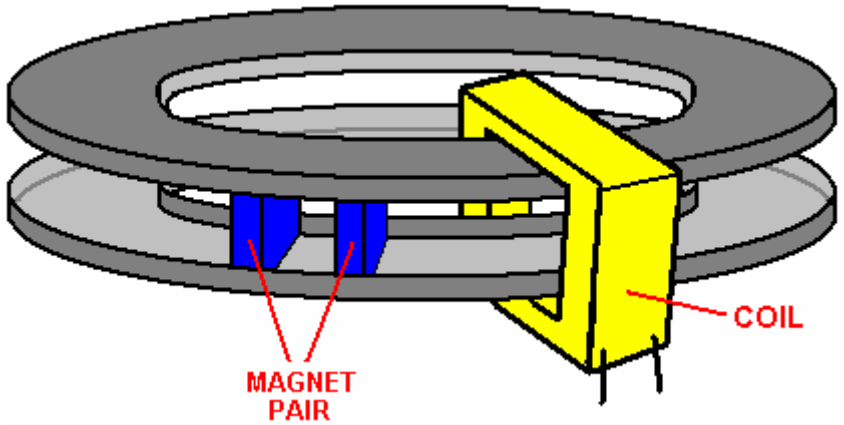
These channels are placed, one below and one above, nine pairs of carousel magnets (coloured blue in some of the patent diagrams shown above. Each carousel magnet sits in the lower channel:



And these magnets are secured as a unit by an identical plastic channel inverted and placed on top of the magnet set:



And this ring assembly of magnets spins inside the wire coils used to generate the electrical output. The ring spins inside the coils because the nine pairs of magnets in the ring, lock in place opposite the matching nine pairs of magnets in the rotor and the magnetic force and rotor rotation causes the ring to spin inside the coils.



PAVEL IMRIS: OPTICAL GENERATOR

US Patent 3,781,601 25th December 1973 Inventor: Pavel Imris

OPTICAL GENERATOR OF AN ELECTROSTATIC FIELD HAVING LONGITUDINAL OSCILLATION AT LIGHT FREQUENCIES FOR USE IN AN ELECTRICAL CIRCUIT

Please note that this is a re-worded excerpt from this patent. It describes a gas-filled tube which allows many standard 40-watt fluorescent tubes to be powered using less than 1-watt of power each.

ABSTRACT

An Optical generator of an electrostatic field at light frequencies for use in an electrical circuit, the generator having a pair of spaced-apart electrodes in a gas-filled tube of quartz glass or similar material with at least one capacitor cap or plate adjacent to one electrode and a dielectric filled container enclosing the tube, the generator substantially increasing the electrical efficiency of the electrical circuit.

BACKGROUND OF THE INVENTION

This invention relates to improved electrical circuits, and more particularly to circuits utilising an optical generator of an electrostatic field at light frequencies.

The measure of the efficiency of an electrical circuit may broadly be defined as the ratio of the output energy in the desired form (such as light in a lighting circuit) to the input electrical energy. Up to now, the efficiency of many circuits has not been very high. For example, in a lighting circuit using 40 watt fluorescent lamps, only about 8.8 watts of the input energy per lamp is actually converted to visible light, thus representing an efficiency of only about 22%. The remaining 31.2 watts is dissipated primarily in the form of heat.

It has been suggested that with lighting circuits having fluorescent lamps, increasing the frequency of the applied current will raise the overall circuit efficiency. While at an operating frequency of 60 Hz, the efficiency is 22%, if the frequency is raised to 1 Mhz, the circuit efficiency would only rise to some 25.5%. Also, if the input frequency were raised to 10 Ghz, the overall circuit efficiency would only be 35%.

SUMMARY OF THE PRESENT INVENTION

The present invention utilises an optical electrostatic generator which is effective for producing high frequencies in the visible light range of about 10^{14} to 10^{23} Hz. The operation and theory of the optical electrostatic generator has been described and discussed in my co-pending application serial No. 5,248, filed on 23rd January 1970. As stated in my co-pending application, the present optical electrostatic generator does not perform in accordance with the accepted norms and standards of ordinary electromagnetic frequencies.

The optical electrostatic generator as utilised in the present invention can generate a wide range of frequencies between several Hertz and those in the light frequency. Accordingly, it is an object of the present invention to provide improved electrical energy circuits utilising my optical electrostatic generator, whereby the output energy in the desired form will be substantially more efficient than possible to date, using standard circuit techniques and equipment. It is a further object of the present invention to provide such a circuit for use in fluorescent lighting or other lighting circuits. It is also an object of the present invention to provide a circuit with may be used in conjunction with electrostatic precipitators for dust and particle collection and removal, as well as many other purposes.

DESCRIPTION OF THE DRAWINGS

Fig.1 is a schematic layout showing an optical electrostatic generator of the present invention, utilised in a lighting circuit for fluorescent lamps:

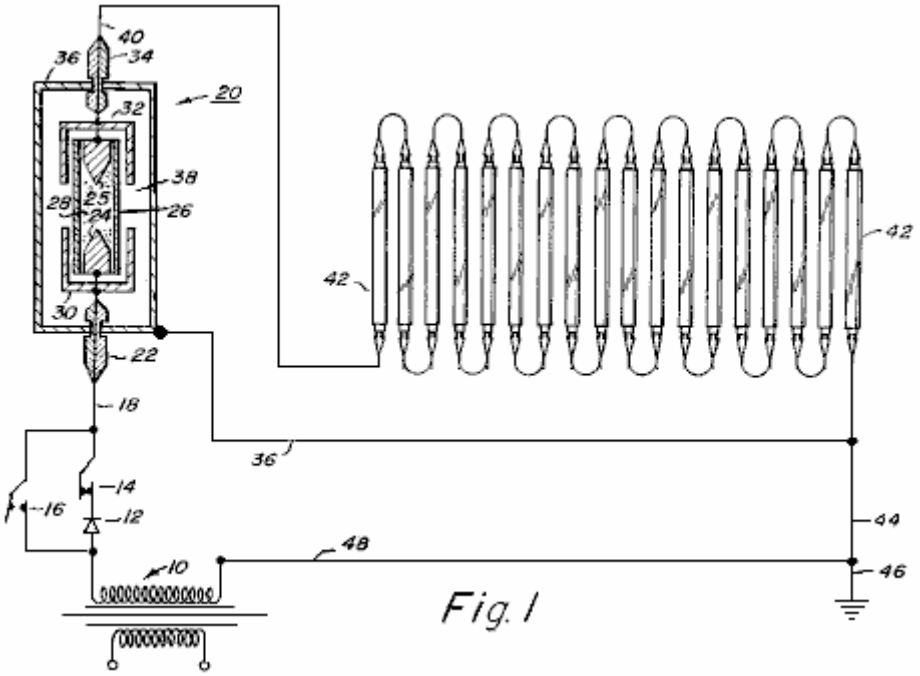


Fig. 1

Fig.2 is a schematic layout of a high-voltage circuit incorporating an optical electrostatic generator:

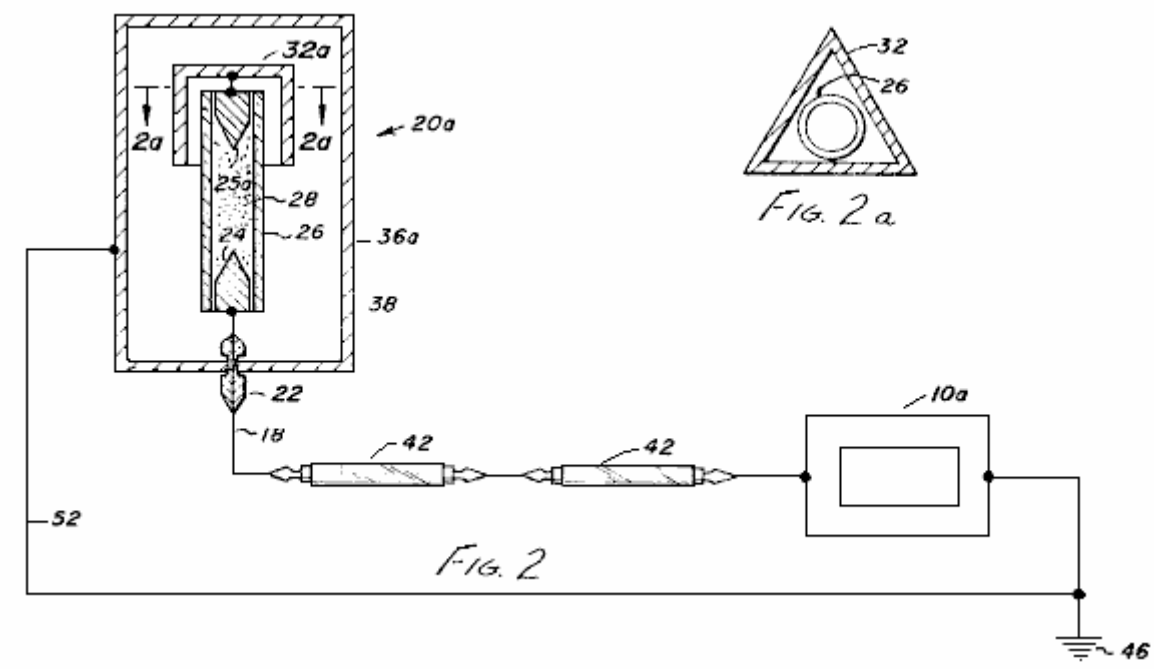


Fig. 2a

Fig. 2

Fig.2A is a sectional view through a portion of the generator and

Fig.3 is a schematic sectional view showing an optical electrostatic generator in accordance with the present invention, particularly for use in alternating current circuits, although it may also be used in direct current circuits:

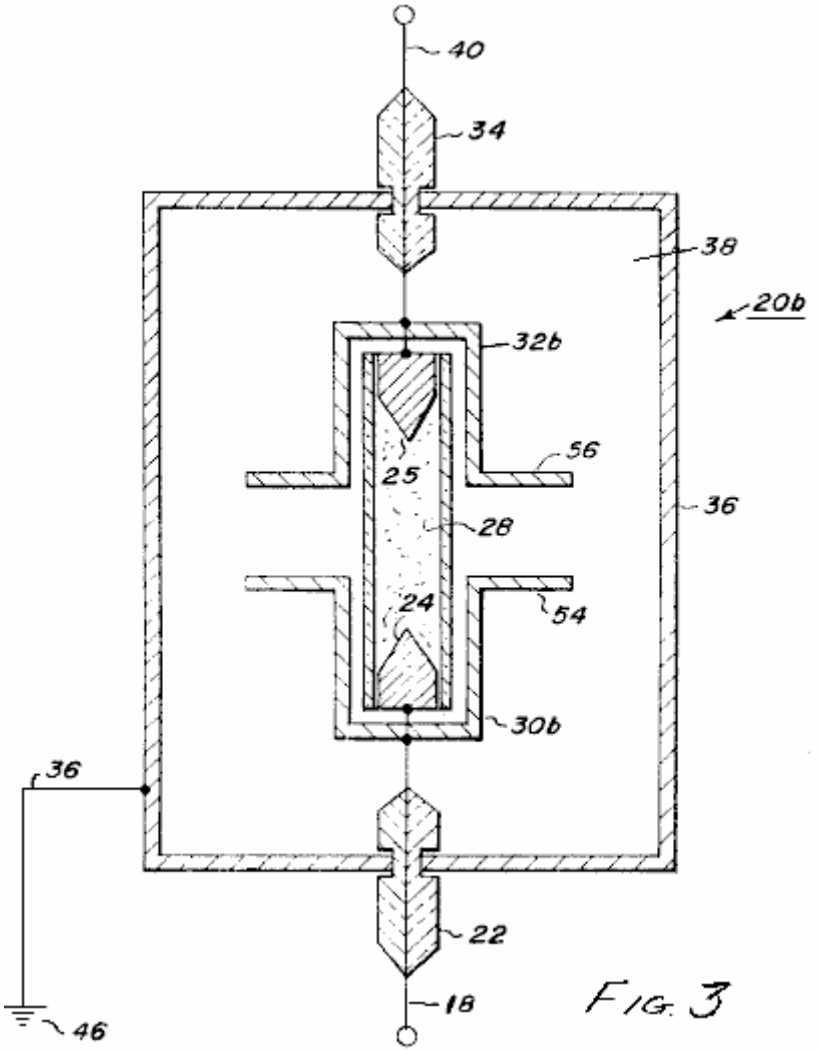


FIG. 3

DESCRIPTION OF THE ILLUSTRATED EMBODIMENTS

Referring to the drawings and to **Fig.1** in particular, a low voltage circuit utilising an optical electrostatic generator is shown. As shown in **Fig.1**, a source of alternating current electrical energy **10**, is connected to a lighting circuit. Connected to one tap of the power source **10** is a rectifier **12** for utilisation when direct current is required. The illustrated circuit is provided with a switch **14** which may be opened or closed depending on whether AC or DC power is used. Switch **14** is opened and a switch **16** is closed when AC is used. With switch **14** closed and switch **16** open, the circuit operates as a DC circuit.

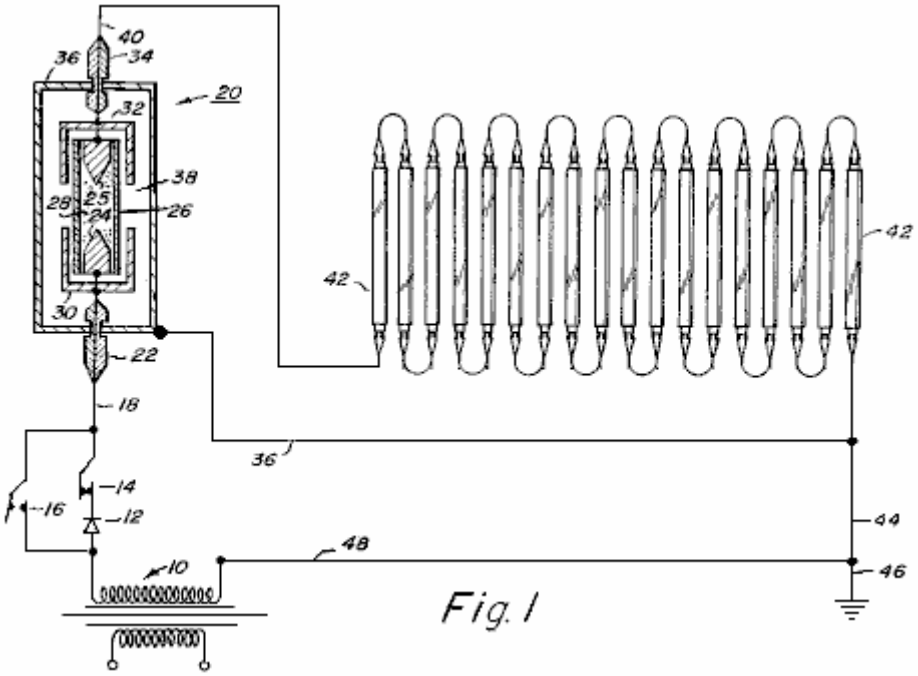


Fig. 1

Extending from switches **14** and **16** is conductor **18** which is connected to an optical electrostatic generator **20**. Conductor **18** is passed through an insulator **22** and connected to an electrode **24**. Spaced from electrode **24** is a second electrode **25**. Enclosing electrodes **24** and **25**, which preferably are made of tungsten or similar material, is a quartz glass tube **26** which is filled with an ionisable gas **28** such as xenon or any other suitable ionisable gas such as argon, krypton, neon, nitrogen or hydrogen, as well as the vapour of metals such as mercury or sodium.

Surrounding each end of tube **26** and adjacent to electrodes **24** and **25**, are capacitor plates **30** and **32** in the form of caps. A conductor is connected to electrode **25** and passed through a second insulator **34**. Surrounding the tube, electrodes and capacitor caps is a metal envelope in the form of a thin sheet of copper or other metal such as aluminium. Envelope **36** is spaced from the conductors leading into and out of the generator by means of insulators **22** and **34**. Envelope **36** is filled with a dielectric material such as transformer oil, highly purified distilled water, nitro-benzene or any other suitable liquid dielectric. In addition, the dielectric may be a solid such as ceramic material with relatively small molecules.

A conductor **40** is connected to electrode **25**, passed through insulator **24** and then connected to a series of fluorescent lamps **42** which are connected in series. It is the lamps **42** which will be the measure of the efficiency of the circuit containing the optical electrostatic generator **20**. A conductor **44** completes the circuit from the fluorescent lamps to the tap of the source of electrical energy **10**. In addition, the circuit is connected to a ground **46** by another conductor **48**. Envelope **36** is also grounded by lead **50** and in the illustrated diagram, lead **50** is connected to the conductor **44**.

The capacitor caps or plates **30** and **32**, form a relative capacitor with the discharge tube. When a high voltage is applied to the electrode of the discharge tube, the ions of gas are excited and brought to a higher potential than their environment, i.e. the envelope and the dielectric surrounding it. At this point, the ionised gas in effect becomes one plate of a relative capacitor in co-operation with the capacitor caps or plates **30** and **32**.

When this relative capacitor is discharged, the electric current does not decrease as would normally be expected. Instead, it remains substantially constant due to the relationship between the relative capacitor and an absolute capacitor which is formed between the ionised gas and the spaced metal envelope **36**. An oscillation effect occurs in the relative capacitor, but the electrical condition in the absolute capacitor remains substantially constant.

As also described in the co-pending application serial No. 5,248, there is an oscillation effect between the ionised gas in the discharge lamp and the metallic envelope **36** will be present if the capacitor caps are eliminated, but the efficiency of the electrostatic generator will be substantially decreased.

The face of the electrode can be any desired shape. However, a conical point of 60^0 has been found to be satisfactory and it is believed to have an influence on the efficiency of the generator.

In addition, the type of gas selected for use in tube **26**, as well as the pressure of the gas in the tube, also affect the efficiency of the generator, and in turn, the efficiency of the electrical circuit.

To demonstrate the increased efficiency of an electrical circuit utilising the optical electrostatic generator of the present invention as well as the relationship between gas pressure and electrical efficiency, a circuit similar to that shown in **Fig.1** may be used with 100 standard 40 watt, cool-white fluorescent lamps connected in series. The optical electrostatic generator includes a quartz glass tube filled with xenon, with a series of different tubes being used because of the different gas pressures being tested.

Table 1 shows the data to be obtained relating to the optical electrostatic generator. **Table 2** shows the lamp performance and efficiency for each of the tests shown in **Table 1**. The following is a description of the data in each of the columns of **Tables 1 and 2**.

Column	Description
B	**Gas used in discharge tube**
C	**Gas pressure in tube** (in torrs)
D	**Field strength across the tube** (measured in volts per cm. of length between the electrodes)
E	**Current density** (measured in microamps per sq. mm. of tube cross-sectional area)
F	**Current** (measured in amps)
G	**Power across the tube** (calculated in watts per cm. of length between the electrodes)
H	**Voltage per lamp** (measured in volts)
K	**Current** (measured in amps)
L	**Resistance** (calculated in ohms)
M	**Input power per lamp** (calculated in watts)
N	**Light output** (measured in lumens)

Table 1

		Optical	Generator	Section		
A	**B**	**C**	**D**	**E**	**F**	**G**
Test No.	**Type of discharge lamp**	**Pressure of Xenon**	**Field strength across lamp**	**Current density**	**Current**	**Power str. across lamp**
		(Torr)	**(V/cm)**	**(A/sq.mm)**	**(A)**	**(W/cm.)**
1	Mo elec	-	-	-	-	-
2	Xe	0.01	11.8	353	0.1818	2.14
3	Xe	0.10	19.6	353	0.1818	3.57
4	Xe	1.00	31.4	353	0.1818	5.72
5	Xe	10.00	47.2	353	0.1818	8.58
6	Xe	20.00	55.1	353	0.1818	10.02
7	Xe	30.00	62.9	353	0.1818	11.45
8	Xe	40.00	66.9	353	0.1818	12.16
9	Xe	60.00	70.8	353	0.1818	12.88
10	Xe	80.00	76.7	353	0.1818	13.95
11	Xe	100.00	78.7	353	0.1818	14.31
12	Xe	200.00	90.5	353	0.1818	16.46
13	Xe	300.00	100.4	353	0.1818	18.25
14	Xe	400.00	106.3	353	0.1818	19.32
15	Xe	500.00	110.2	353	0.1818	20.04
16	Xe	600.00	118.1	353	0.1818	21.47
17	Xe	700.00	120.0	353	0.1818	21.83
18	Xe	800.00	122.8	353	0.1818	22.33
19	Xe	900.00	125.9	353	0.1818	22.90
20	Xe	1,000.00	127.9	353	0.1818	23.26
21	Xe	2,000.00	149.6	353	0.1818	27.19
22	Xe	3,000.00	161.4	353	0.1818	29.35
23	Xe	4,000.00	173.2	353	0.1818	31.49
24	Xe	5,000.00	179.1	353	0.1818	32.56

Table 2

		Fluorescent	Lamp	Section	
A	**H**	**K**	**L**	**M**	**N**
Test No.	**Voltage**	**Current**	**Resistance**	**Input Energy**	**Light Output**
	(Volts)	**(Amps)**	**(Ohms)**	**(Watts)**	**(Lumen)**
1	220	0.1818	1,210	40.00	3,200
2	218	0.1818	1,199	39.63	3,200
3	215	0.1818	1,182	39.08	3,200
4	210	0.1818	1,155	38.17	3,200
5	200	0.1818	1,100	36.36	3,200
6	195	0.1818	1,072	35.45	3,200
7	190	0.1818	1,045	34.54	3,200
8	182	0.1818	1,001	33.08	3,200
9	175	0.1818	962	31.81	3,200
10	162	0.1818	891	29.45	3,200
11	155	0.1818	852	28.17	3,200
12	130	0.1818	715	23.63	3,200
13	112	0.1818	616	20.36	3,200
14	100	0.1818	550	18.18	3,200
15	85	0.1818	467	15.45	3,200
16	75	0.1818	412	13.63	3,200
17	67	0.1818	368	12.18	3,200
18	60	0.1818	330	10.90	3,200
19	53	0.1818	291	9.63	3,200
20	50	0.1818	275	9.09	3,200
21	23	0.1818	126	4.18	3,200
22	13	0.1818	71	2.35	3,200
23	8	0.1818	44	1.45	3,200
24	5	0.1818	27	0.90	3,200

The design of a tube construction for use in the optical electrostatic generator of the type used in **Fig.1**, may be accomplished by considering the radius of the tube, the length between the electrodes in the tube and the power across the tube.

If **R** is the minimum inside radius of the tube in centimetres, **L** the minimum length in centimetres between the electrodes, and **W** the power in watts across the lamp, the following formula can be obtained from **Table 1:**

R = (Current [A] / Current Density [A/sq.mm]) / pi

L = 8R

W = L[V/cm] x A

For example, for Test No. 18 in Table 1:
The current is 0.1818 A,
The current density 0.000353 A/sq.mm and
The Voltage Distribution is 122.8 V/cm; therefore

R = $(0.1818 / 0.000353)^2$ /3.14 = 12.80 mm.

L = 8 x R = 8 * 12.8 = 102.4 mm (10.2 cm.)

W = 10.2 x 122.8 x 0.1818 = 227.7 VA or 227.7 watts

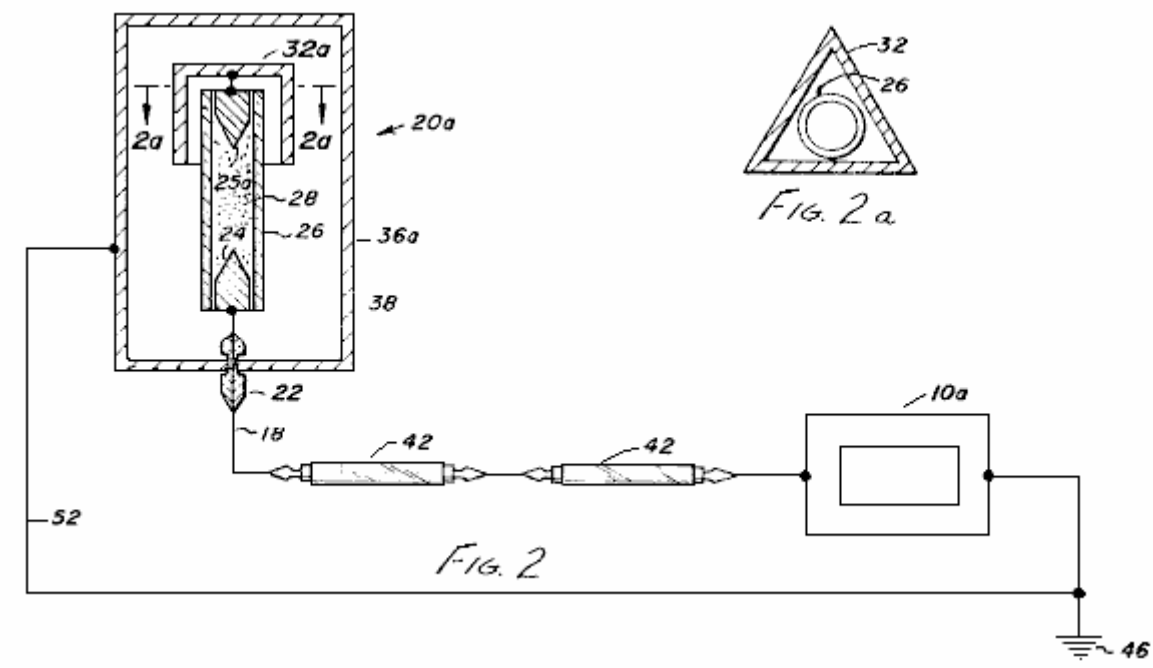
The percent efficiency of operation of the fluorescent lamps in Test No. 18 can be calculated from the following equation:

% Efficiency = (Output Energy/Input energy) x 100

Across a single fluorescent lamp, the voltage is 60 volts and the current is 0.1818 amps therefore the input energy to the lamp **42** is 10.90 Watts. The output of the fluorescent lamp is 3,200 lumens which represents 8.8 Watts power of light energy. Thus, the one fluorescent lamp is operating at 80.7% efficiency under these conditions.

However, when the optical generator is the same as described for Test No. 18 and there are 100 fluorescent lamps in series in the circuit, the total power input is 227.7 watts for the optical generator and 1,090 watts for 100 fluorescent lamps, or a total of 1,318 watts. The total power input normally required to operate the 100 fluorescent lamps in a normal circuit would be 100 x 40 = 4,000 watts. So by using the optical generator in the circuit, about 2,680 watts of energy is saved.

Table 1 is an example of the functioning of this invention for a particular fluorescent lamp (40 watt cool white). However, similar data can be obtained for other lighting applications, by those skilled in the art.



In **Fig.2**, a circuit is shown which uses an optical electrostatic generator **20a**, similar to generator **20** of **Fig.1**. In generator **20**, only one capacitor cap **32a** is used and it is preferably of triangular cross-sectional design. In addition, the second electrode **25a** is connected directly back into the return conductor **52**, similar to the arrangement shown in my co-pending application serial No. 5,248, filed 23rd January 1970.

This arrangement is preferably for very high voltage circuits and the generator is particularly suited for DC usage.

In **Fig.2**, common elements have received the same numbers which were used in **Fig.1**.

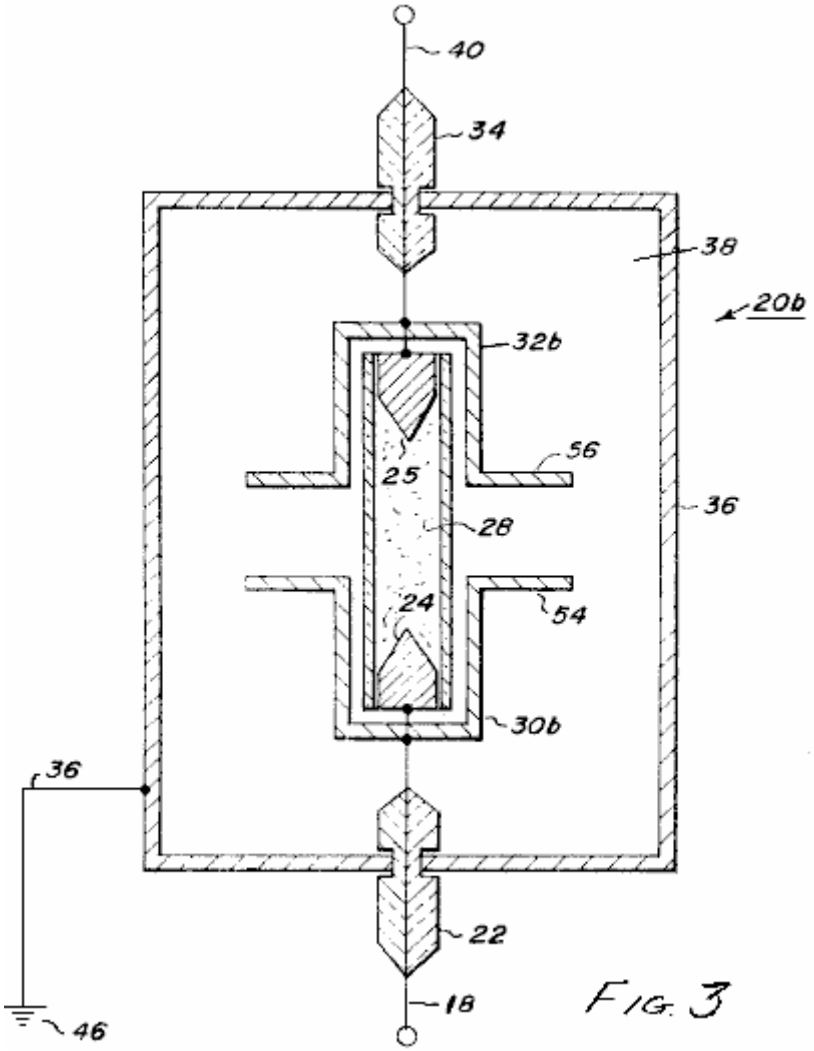


FIG. 3

In **Fig.3**, still another embodiment of an optical electrostatic generator **20b** is shown. This generator is particularly suited for use with AC circuits. In this embodiment, the capacitor plates **30b** and **32b** have flanges **54** and **56** which extend outwards towards the envelope **36**. While the utilisation of the optical electrostatic generator has been described in use in a fluorescent lighting circuit, it is to be understood that many other types of circuits may be used. For example, the high-voltage embodiment may be used in a variety of circuits such as flash lamps, high-speed controls, laser beams and high-energy pulses. The generator is also particularly usable in a circuit including electrostatic particle precipitation in air pollution control devices, chemical synthesis in electrical discharge systems such as ozone generators and charging means for high-voltage generators of the Van de Graff type, as well as particle accelerators. To those skilled in the art, many other uses and circuits will be apparent.

Patent GB 763,062 5th December 1956 Inventors: Harold Colman and Ronald Seddon-Gillespie

APPARATUS FOR PRODUCING AN ELECTRIC CURRENT

This patent shows the details of a lightweight device which can produce electricity using a self-powered electromagnet and chemical salts. The working life of the device before needing a recharge is estimated at some seventy years. The operation is controlled by a transmitter which bombards the chemical sample with 300 MHz radio waves. This produces radioactive emissions from the chemical mixture for a period of one hour maximum, so the transmitter needs to be run for fifteen to thirty seconds once every hour. The chemical mixture is shielded by a lead screen to prevent harmful radiation reaching the user. The output from the tiny device described is estimated to be some 10 amps at 100 to 110 volts DC.

DESCRIPTION

This invention relates to a new apparatus for producing electric current the apparatus being in the form of a completely novel secondary battery. The object of this invention is to provide apparatus of the above kind which is considerably lighter in weight than, and has an infinitely greater life than a known battery or similar characteristics and which can be re-activated as and when required in a minimum of time.

According to the present invention we provide apparatus comprising a generator unit which includes a magnet, a means for suspending a chemical mixture in the magnetic field, the mixture being composed of elements whose nuclei becomes unstable as a result of bombardment by short waves so that the elements become radio-active and release electrical energy, the mixture being mounted between, and in contact with, a pair of different metals such as copper and zinc, a capacitor mounted between those metals, a terminal electrically connected to each of the metals, means for conveying the waves to the mixture and a lead shield surrounding the mixture to prevent harmful radiation from the mixture.

The mixture is preferably composed of the elements Cadmium, Phosphorus and Cobalt having Atomic Weights of 112, 31 and 59 respectively. The mixture, which may be of powdered form, is mounted in a tube of non-conducting, high heat resistivity material and is compressed between granulated zinc at one end of the tube and granulated copper at the other end, the ends of the tube being closed by brass caps and the tube being carried in a suitable cradle so that it is located between the poles of the magnet. The magnet is preferably an electro-magnet and is energised by the current produced by the unit.

The means for conveying the waves to the mixture may be a pair of antennae which are exactly similar to the antennae of the transmitter unit for producing the waves, each antenna projecting from and being secured to the brass cap at each end of the tube.

The transmitter unit which is used for activating the generator unit may be of any conventional type operating on ultra-shortwave and is preferably crystal controlled at the desired frequency.

DESCRIPTION OF THE DRAWINGS

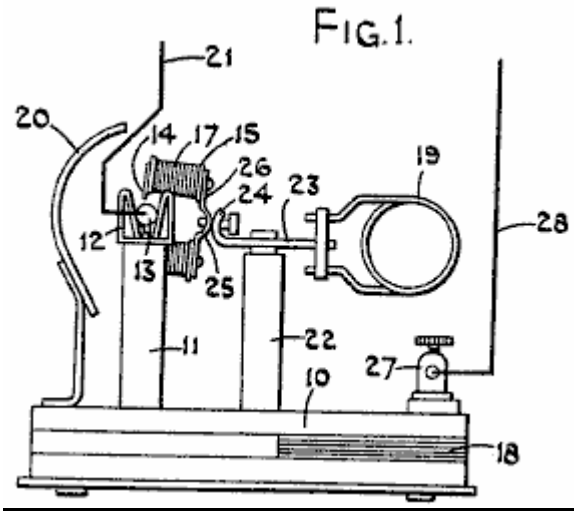


Fig.1 is a side elevation of one form of the apparatus.

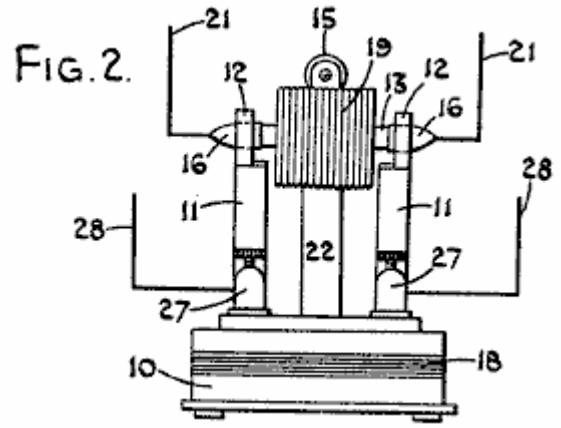


Fig.2 is a view is an end elevation

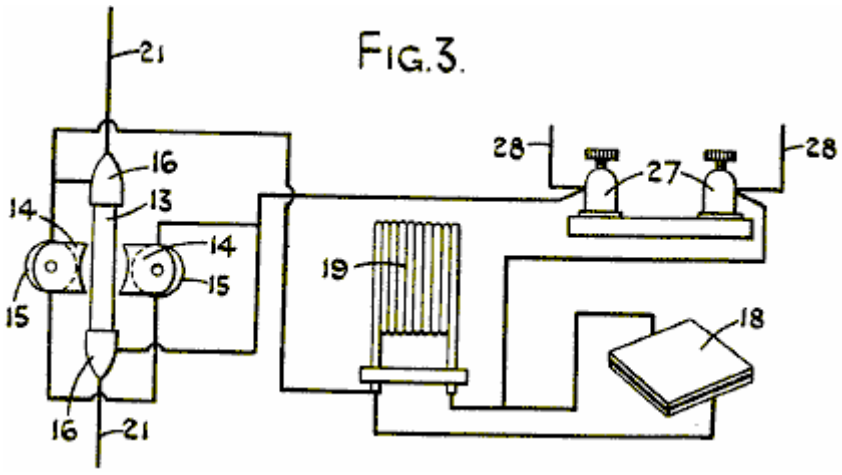


Fig.3 is a schematic circuit diagram.

In the form of our invention illustrated, the generator unit comprises a base **10** upon which the various components are mounted. This base **10**, having projecting upwards from it a pair of arms **11**, which form a cradle housing **12** for a quartz tube **13**, the cradle **12** preferably being made of spring material so that the tube **13** is firmly, yet removably held in position. The arms **11** are positioned relative to the poles **14** of an electromagnet **15** so that the tube **13** is located immediately between the poles of the magnet so as to be in the strongest magnetic field created by the electromagnet. The magnet serves to control the alpha and beta rays emitted by the cartridge when it is in operation.

The ends of the quartz tube **13** are each provided with a brass cap **16**, and these caps **16** are adapted to engage within the spring cradles **12** and the coils **17** associated with the magnet being so arranged that if the base **10** of the unit is in a horizontal plane, the poles **14** of the magnet are in a substantially vertical plane.

Also connected across the cradles is a lead capacitor **18** which may conveniently be housed in the base **10** of the unit and connected in parallel with this capacitor **18** is a suitable high frequency inductance coil **19**. The unit is provided with a lead shield **20** so as to prevent harmful radiation from the quartz tube as will be described later.

The quartz tube **13** has mounted in it, at one end, a quantity of granulated copper which is in electrical contact with the brass cap **16** at that end of the tube. Also mounted within the tube and in contact with the granulated copper is a chemical mixture which is in powdered form and which is capable of releasing electrical energy and which becomes radioactive when subjected to bombardment by ultra-short radio waves.

Mounted in the other end of the tube, and in contact with the other end of the powdered chemical mixture is a quantity of granulated zinc which is itself in contact with the brass cap on this end of the tube, the arrangement being that the chemical mixture is compressed between the granulated copper and the granulated zinc.

Projecting outwards from each brass cap **16**, and electrically connected to them, is an antenna **21**. Each antenna **21** corresponding exactly in dimension, shape and electrical characteristics to the antenna associated with a transmitter unit which is to produce the ultra shortwaves mentioned earlier.

The electromagnet **15** is conveniently carried by a centrally positioned pillar **22** which is secured to the base **10**. At the upper end of pillar **22** there is a cross-bar **23**, which has the high frequency coil **19** attached to one end of it. The other end of the cross-bar **23** is bent around into the curved shape as shown at **24** and is adapted to bear against a curved portion **25** of the base **26** of the electromagnet **15**. A suitable locking device is provided for holding the curved portions **24** and **25** in the desired angular position, so that the position of the poles **14** of the electromagnet can be adjusted about the axis of the quartz tube **13**.

The transmitter unit is of any suitable conventional type for producing ultra shortwaves and may be crystal controlled to ensure that it operates at the desired frequency with the necessity of tuning. If the transmitter is only required to operate over a short range, it may conveniently be battery powered but if it is to operate over a greater range, then it may be operated from a suitable electrical supply such as the mains. If the transmitter is to be tuned, then the tuning may be operated by a dial provided with a micrometer vernier scale so that the necessary tuning accuracy may be achieved.

The mixture which is contained within the quartz tube is composed of the elements Cadmium, Phosphorus and Cobalt, having atomic weights 112, 31 and 59 respectively. Conveniently, these elements may be present in the following compounds, and where the tube is to contain thirty milligrams of the mixture, the compounds and their proportions by weight are:

1 Part of Co (No_3) 2 $6H_2O$
2 Parts of $CdCl_2$
3 Parts of 3Ca (Po_3) 2 + 10C.

The cartridge which consists of the tube **13** with the chemical mixture in it is preferably composed of a number of small cells built up in series. In other words, considering the cartridge from one end to the other, at one end and in contact with the brass cap, there would be a layer of powdered copper, then a layer of the chemical mixture, then a layer of powdered zinc, a layer of powdered copper, etc. with a layer of powdered zinc in contact with the brass cap at the other end of the cartridge. With a cartridge some forty five millimetres long and five millimetres diameter, some fourteen cells may be included.

The cradles **12** in which the brass caps **16** engage, may themselves form terminals from which the output of the unit may be taken. Alternatively, a pair of terminals **27** may be connected across the cradles **12**, these terminals **27** being themselves provided with suitable antennae **28**, which correspond exactly in dimensions, shape and electrical characteristics to the antennae associated with the transmitter, these antennae **28**, replacing the antennae **21**.

In operation with the quartz tube containing the above mixture located between the granulated copper and the granulated zinc and with the tube itself in position between the poles of the magnet, the transmitter is switched on and the ultra shortwaves coming from it are received by the antennae mounted at each end of the tube and in contact with the copper and zinc respectively, the waves being thus passed through the copper and zinc and through the mixture so that the mixture is bombarded by the short waves and the Cadmium, Phosphorus and Cobalt associated with the mixture become radioactive and release electrical energy which is transmitted to the granulated copper and granulated zinc, causing a current to flow between them in a similar manner to the current

flow produced by a thermo couple. It has been established that with a mixture having the above composition, the optimum release of energy is obtained when the transmitter is operating at a frequency of 300 MHz.

The provision of a quartz tube is necessary for the mixture evolves a considerable amount of heat while it is reacting to the bombardment of the short waves. It is found that the tube will only last for one hour and that the tube will become discharged after an hours operation, that is to say, the radioactiveness of the tube will only last for one hour and it is therefore necessary, if the unit is to be run continuously, for the transmitter to be operated for a period of some fifteen to thirty seconds duration once every hour.

With a quartz tube having an overall length of some forty five millimetres and an inside diameter of five millimetres and containing thirty milligrams of the chemical mixture, the estimated energy which will be given off from the tube for a discharge of one hour, is 10 amps at between 100 and 110 volts. To enable the tube to give off this discharge, it is only necessary to operate the transmitter at the desired frequency for a period of some fifteen to thirty seconds duration.

The current which is given off by the tube during its discharge is in the form of direct current. During the discharge from the tube, harmful radiations are emitted in the form of gamma rays, alpha rays and beta rays and it is therefore necessary to mount the unit within a lead shield to prevent the harmful radiations from affecting personnel and objects in the vicinity of the unit. The alpha and beta rays which are emitted from the cartridge when it is in operation are controlled by the magnet.

When the unit is connected up to some apparatus which is to be powered by it, it is necessary to provide suitable fuses to guard against the cartridge being short-circuited which could cause the cartridge to explode.

The estimated weight of such a unit including the necessary shielding, per kilowatt hour output, is approximately 25% of any known standard type of accumulator which is in use today and it is estimated that the life of the chemical mixture is probably in the region of seventy to eighty years when under constant use.

It will thus be seen that we have provided a novel form of apparatus for producing an electric current, which is considerably lighter than the standard type of accumulator at present known, and which has an infinitely greater life than the standard type of accumulator, and which can be recharged or reactivated as and when desired and from a remote position depending on the power output of the transmitter. Such form of battery has many applications.

Patent US 6,208,061 27th March 2001 Inventor: Jong-Sok An

NO-LOAD GENERATOR

Electrical power is frequently generated by spinning the shaft of a generator which has some arrangement of coils and magnets contained within it. The problem is that when current is drawn from the take-off coils of a typical generator, it becomes much more difficult to spin the generator shaft. The cunning design shown in this patent overcomes this problem with a simple design in which the effort required to turn the shaft is not altered by the current drawn from the generator.

ABSTRACT

A generator of the present invention is formed of ring permanent magnet trains 2 and 2' attached and fixed on to two orbits 1 and 1' about a rotational axis 3, magnetic induction primary cores 4 and 4' attached and fixed above outer peripheral surfaces of the ring permanent magnet trains 2 and 2' at a predetermined distance from the outer peripheral surfaces, magnetic induction secondary cores 5 and 5' attached and fixed on to the magnetic induction primary cores 4 and 4' and each having two coupling, holes 6 and 6' formed therein, tertiary cores 8 and 8' inserted for coupling respectively into two coupling holes 6 and 6' of each of the associated magnetic induction secondary cores 5 and 5' opposite to each other, and responsive coils 7 and 7'. The ring permanent magnetic trains 2 and 2' are formed of 8 sets of magnets with alternating N and S poles, and magnets associated with each other in the axial direction have opposite polarities respectively and form a pair.

DESCRIPTION

TECHNICAL FIELD

The present invention relates to generators, and particularly to a load-free generator which can maximise the generator efficiency by erasing or eliminating the secondary repulsive load exerted on the rotor during electric power generation.

BACKGROUND ART

The generator is a machine which converts mechanical energy obtained from sources of various types of energy such as physical, chemical or nuclear power energy, for example, into electric energy. Generators based on linear motion have recently been developed while most generators are structured as rotational type generators. Generation of electromotive force by electromagnetic induction is a common principle to generators regardless of their size or whether the generator is AC or DC generator.

The generator requires a strong magnet such as permanent magnet and electromagnet for generating magnetic field as well as a conductor for generating the electromotive force, and the generator is structured to enable one of them to rotate relative to the other. Depending on which of the magnet and the conductor rotates, generators can be classified into rotating-field type generators in which the magnetic field rotates and rotating-armature type generators in which the conductor rotates.

Although the permanent magnet can be used for generating the magnetic field, the electromagnet is generally employed which is formed of a magnetic field coil wound around a core to allow direct current to flow through them. Even if a strong magnet is used to enhance the rotational speed, usually the electromotive force produced from one conductor is not so great. Thus, in a generally employed system, a large number of conductors are provided in the generator and the electromotive forces generated from respective conductare serially added up so as to achieve a high electric power.

As discussed above, a usual generator produces electricity by mechanically rotating a magnet (or permanent magnet) or a conductor (electromagnet, electrically responsive coil and the like) while reverse current generated at this time by magnetic induction (electromagnetic induction) and flowing through the coil causes magnetic force which pulls the rotor so that the rotor itself is subjected to unnecessary load which reaches at least twice the electric power production.

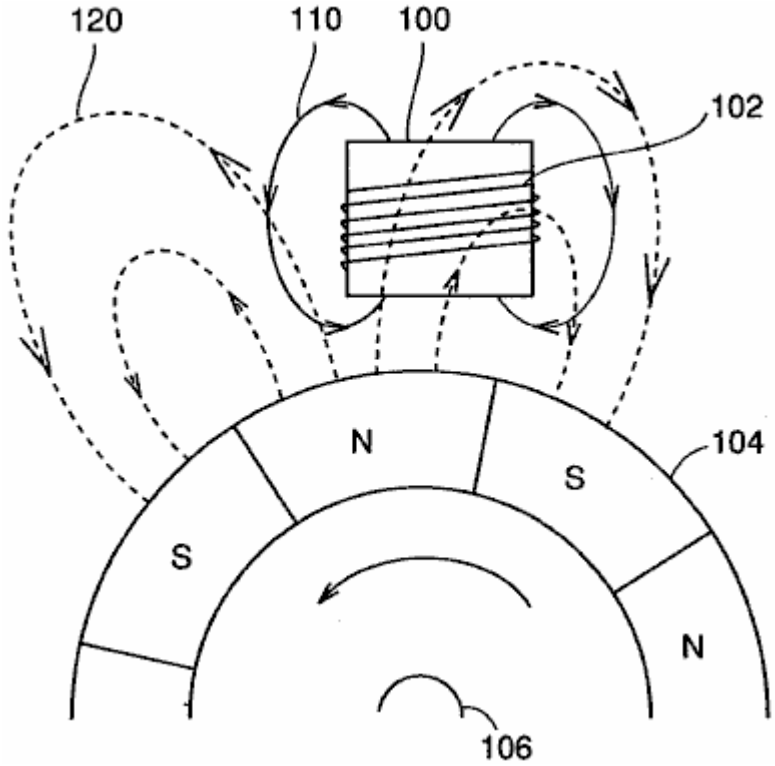


Fig.6 illustrates that the load as discussed above is exerted on a rotor in a rotating-field type generator mentioned above.

Referring to **Fig.6**, a permanent magnet train **104** is arranged about an axis of rotation **106** such that N poles and S poles are alternately located on the outer peripheral surface of the train. At a certain distance outward from the outer periphery of permanent magnet train **104**, a magnetic induction core **100** is arranged and a coil **102** is wound around magnetic induction core **100**.

As permanent magnet train **104** rotates, the magnetic field produced in the coil by permanent magnet train **104** changes to cause induced current to flow through coil **102**. This induced current allows coil **102** to generate a magnetic field **110** which causes a repulsive force exerted on permanent magnet train **104** in the direction which interferes the rotation of the magnet train.

For example, in the example shown in **Fig.6**, the S pole of magnetic field **110** faces permanent magnet train **104**. The S pole of permanent magnet train **104** approaches coil **102** because of rotation of permanent magnet train **104**, resulting in the repulsive force as described above.

If reverse current flows in a responsive coil of an armature wound around a magnetic induction core of a generator so that the resulting load hinders the rotor from rotating, reverse magnetic field of the armature responsive coil becomes stronger in proportion to the electricity output and accordingly a load corresponding to at least twice the instantaneous consumption could occur.

If electric power of 100W is used, for example, reverse magnetic field of at least 200W is generated so that an enormous amount of load affects the rotor to interfere the rotation of the rotor.

All of the conventional generators are subjected to not only a mechanical primary load, i.e. the load when the electric power is not consumed but a secondary load due to reverse current which is proportional to electric power consumption and consequently subjected to a load of at least twice the instantaneous consumption.

Such an amount of the load is a main factor of reduction of the electric power production efficiency, and solution of the problem above has been needed.

DISCLOSURE OF THE INVENTION

One object of the present invention is to provide a generator capable of generating electric power with high efficiency by cancelling out the secondary load except the mechanical load of the generator, i.e. cancelling out the load which is generated due to reverse current of a responsive coil of an armature wound around a magnetic induction core, so as to entirely prevent the secondary load from being exerted.

In short, the present invention is applied to a load-free generator including a rotational axis, a first ring magnet train, a second ring magnet train, a first plurality of first magnetic induction primary cores, a first plurality of second magnetic induction primary cores, a first responsive coil, and a second responsive coil.

The first ring magnet train has N poles and S poles successively arranged on an outer periphery of a first rotational orbit about the rotational axis. The second ring magnet train has magnets successively arranged on an outer periphery of a second rotational orbit about the rotational axis at a predetermined distance from the first rotational orbit such that the polarities of the magnets on the second rotational orbit are opposite to the polarities at opposite locations on the first rotational orbit respectively. The first plurality of first magnetic induction primary cores are fixed along a first peripheral surface of the first ring magnet train at a predetermined distance from the first peripheral surface. The first plurality of second magnetic induction primary cores are fixed along a second peripheral surface of the second ring magnet train at a predetermined distance from the second peripheral surface. A first plurality of first coupling magnetic induction cores and a first plurality of second coupling magnetic induction cores are provided in pairs to form a closed magnetic circuit between the first and second magnetic induction primary cores opposite to each other in the direction of the rotational axis. The first responsive coil is wound around the first coupling magnetic induction core. The second responsive coil is wound around the second coupling magnetic induction core, the direction of winding of the second responsive coil being reversed relative to the first responsive coil.

Preferably, in the load-free generator of the invention, the first ring magnet train includes a permanent magnet train arranged along the outer periphery of the first rotational orbit, and the second ring magnet train includes a permanent magnet train arranged along the outer periphery of the second rotational orbit.

Still preferably, the load-free generator of the present invention further includes a first plurality of first magnetic induction secondary cores provided on respective outer peripheries of the first magnetic induction primary cores and each having first and second coupling holes, and a first plurality of second magnetic induction secondary cores provided on respective outer peripheries of the second magnetic induction primary cores and each having third and fourth coupling holes. The first coupling magnetic induction cores are inserted into the first and third coupling holes to couple the first and second magnetic induction secondary cores, and the second coupling magnetic induction cores are inserted into the second and fourth coupling holes to couple the first and second magnetic induction secondary cores.

Alternatively, the load-free generator of the present invention preferably has a first plurality of first responsive coils arranged in the rotational direction about the rotational aids that are connected zigzag to each other and a first plurality of second responsive coils arranged in the rotational direction about the rotational axis that are connected zigzag to each other.

Alternatively, in the load-free generator of the present invention, preferably the first plurality is equal to 8, and the 8 first responsive coils arranged in the rotational direction about the rotational axis are connected zigzag to each other, and the 8 second responsive coils arranged in the rotational direction about the rotational axis are connected zigzag to each other.

Accordingly, a main advantage of the present invention is that two responsive coils wound respectively in opposite directions around a paired iron cores are connected to cancel reverse magnetic forces generated by reverse currents (induced currents) flowing through the two responsive coils, so that the secondary load which interferes the rotation of the rotor is totally prevented and thus a load-free generator can be provided which is subjected to just a load which is equal to or less than mechanical load when electric power production is not done, i.e. the rotational load even when the generator is operated to the maximum.

Another advantage of the present invention is that the reverse magnetic force, as found in the conventional generators, due to reverse current occurring when the rotor rotates is not generated, and accordingly load of energy except the primary gravity of the rotor and dynamic energy of the rotor is eliminated to increase the amount of electricity output relative to the conventional electric power generation system and thus enhance the electric power production and economic efficiency.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig.1 is a cross sectional view of a rotating-field type generator according to an embodiment of the present invention illustrating an arrangement a permanent magnet, magnetic induction cores and coils.

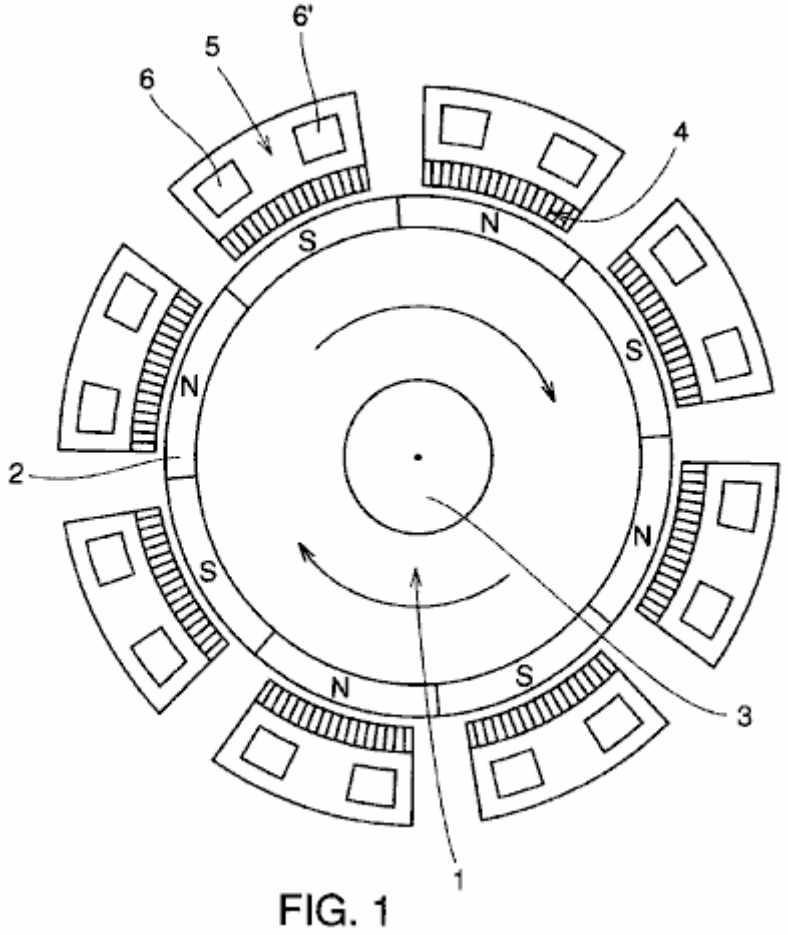


FIG. 1

Fig.2 is a partial schematic view illustrating a magnetic array of the permanent magnet rotor and an arrangement of one of magnetically responsive coils placed around that rotor in an embodiment of the present invention.

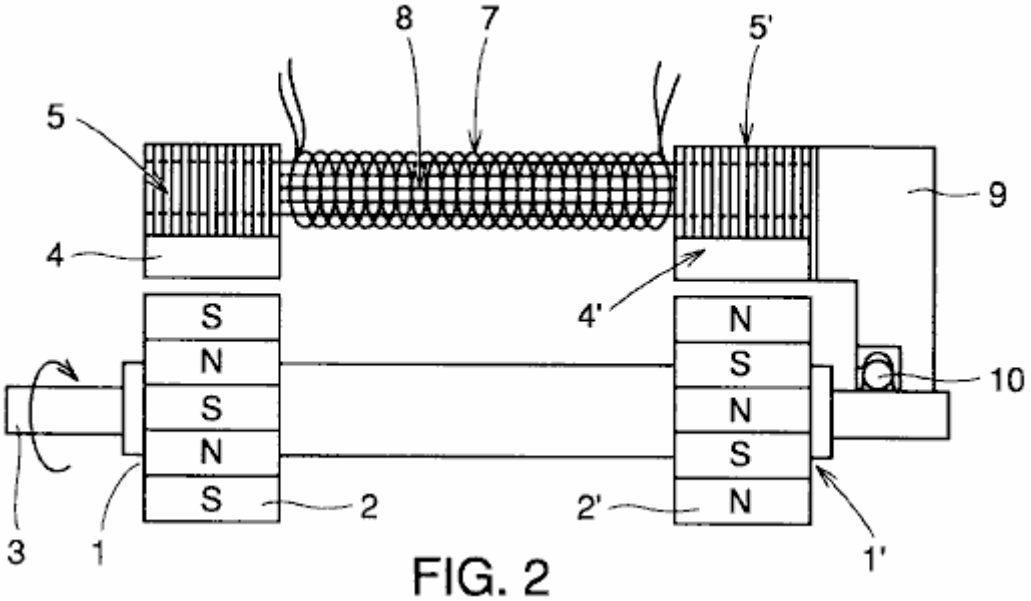


FIG. 2

Fig.3 illustrates a structure of the magnetically responsive coils and cores in the embodiment of the present invention.

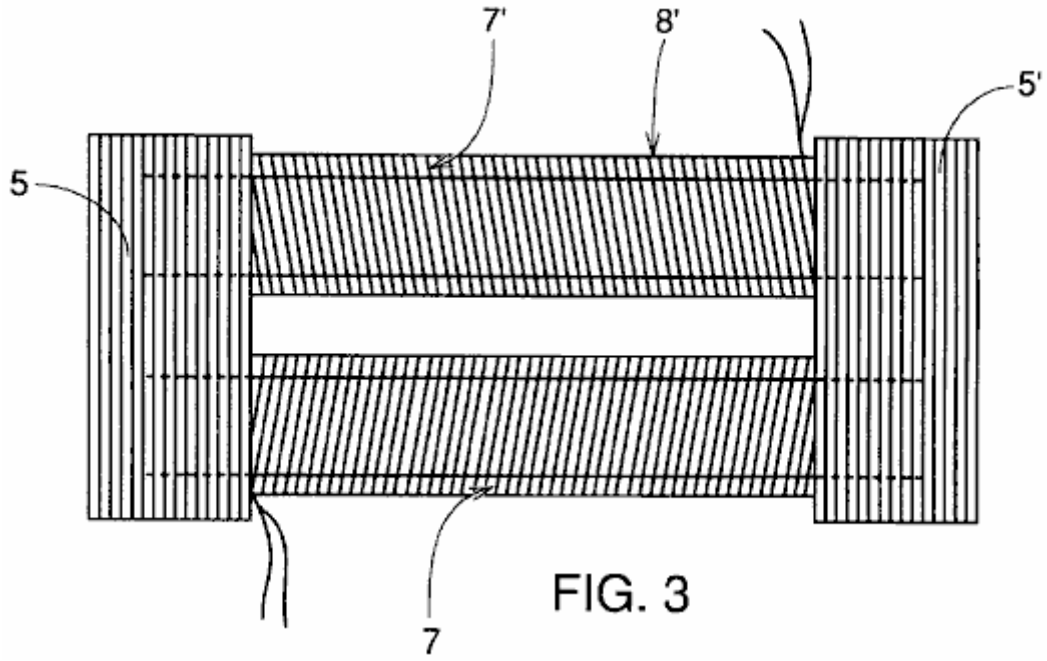


FIG. 3

Fig.4 is an enlarged plan view of magnetically sensitive cores and coil portions of the load-free generator of the present invention illustrating magnetic flow therethrough.

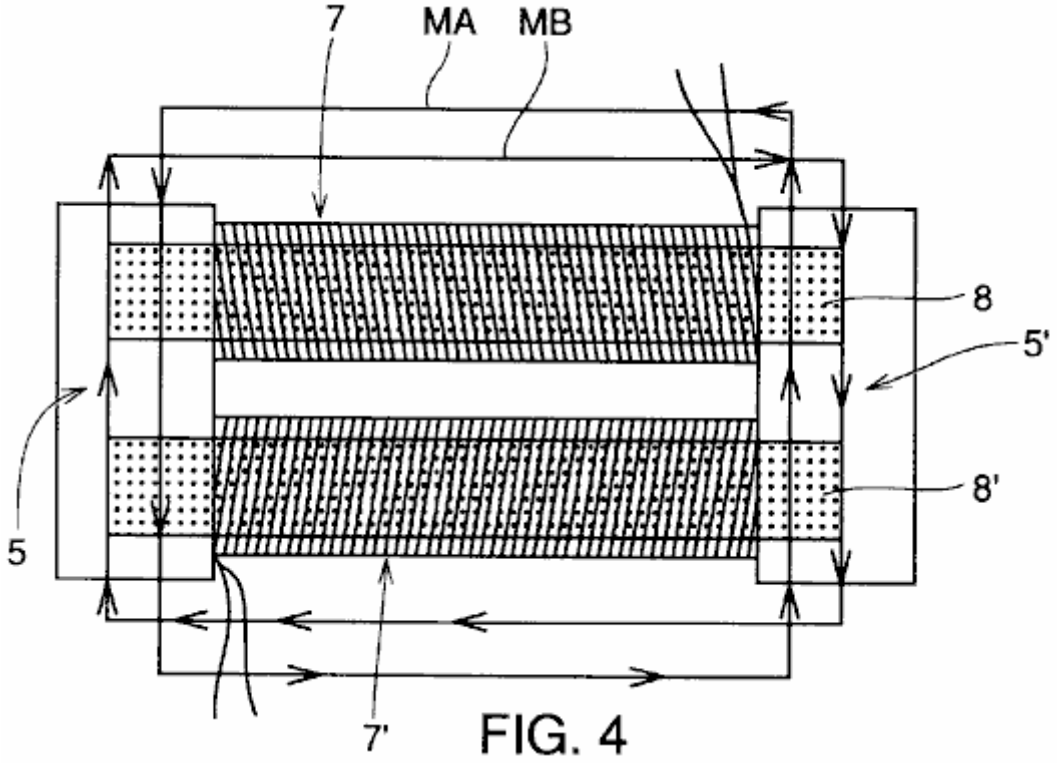


FIG. 4

Fig.5 is an exploded view about a central axis showing the interconnection of magnetic field coils which are respectively wound around tertiary cores surrounding the permanent magnet rotor in FIG. 1 according to the present invention.

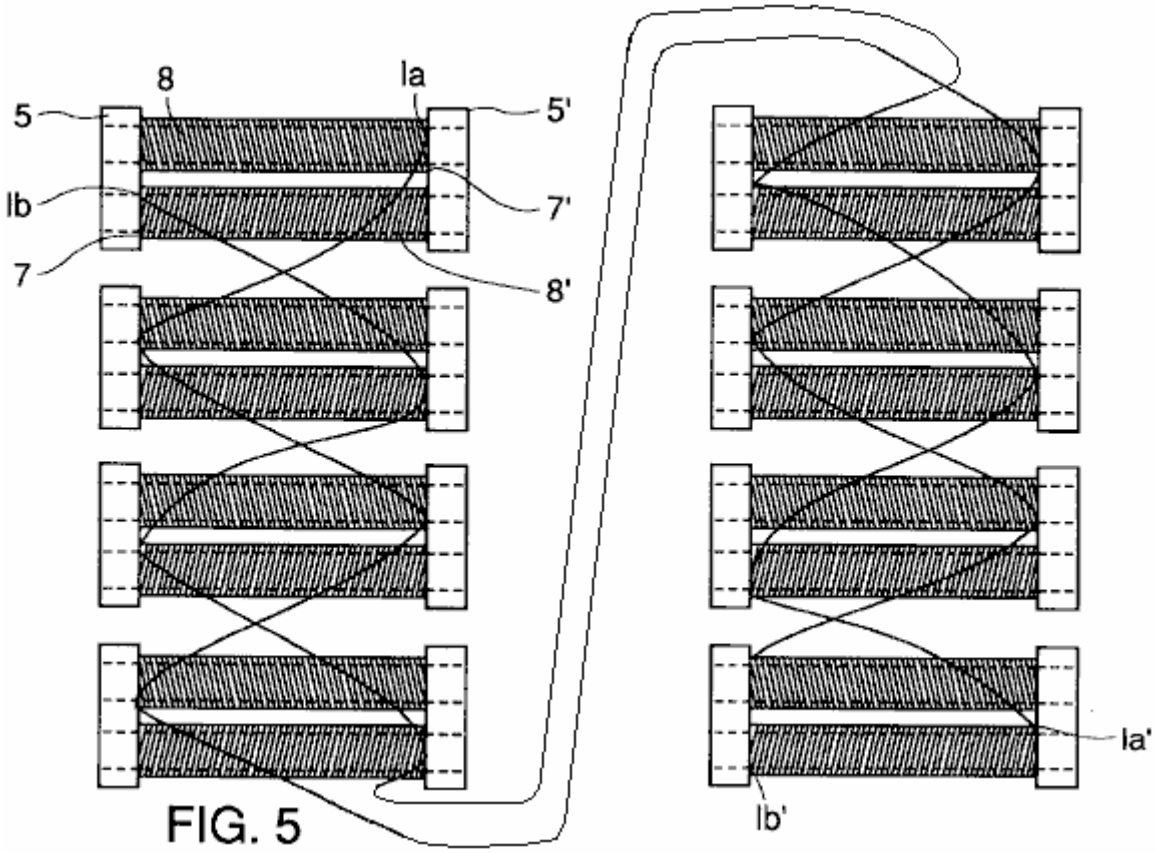
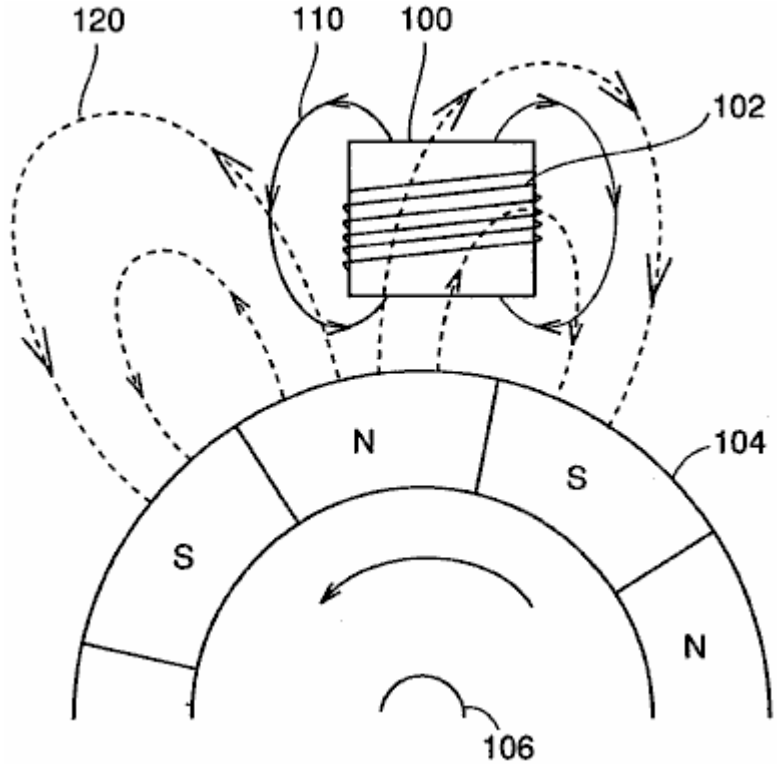


FIG. 5

Fig.6 illustrates generation of the secondary load in a conventional generator.



BEST MODES FOR CARRYING OUT THE INVENTION

The structure and operation of a load-free generator according to the present invention are now described in conjunction with the drawings.

Fig.1 illustrates a cross sectional structure of the load-free generator of the invention perpendicular to a rotational axis **3**.

Fig.2 partially illustrates a cross sectional structure of the load-free generator of the invention in parallel to rotational axis **3**. Specifically, in **Fig.2**, only one of eight sets of magnetic induction primary cores **4** and **4'** arranged around rotational axis **3** as described below is representatively shown.

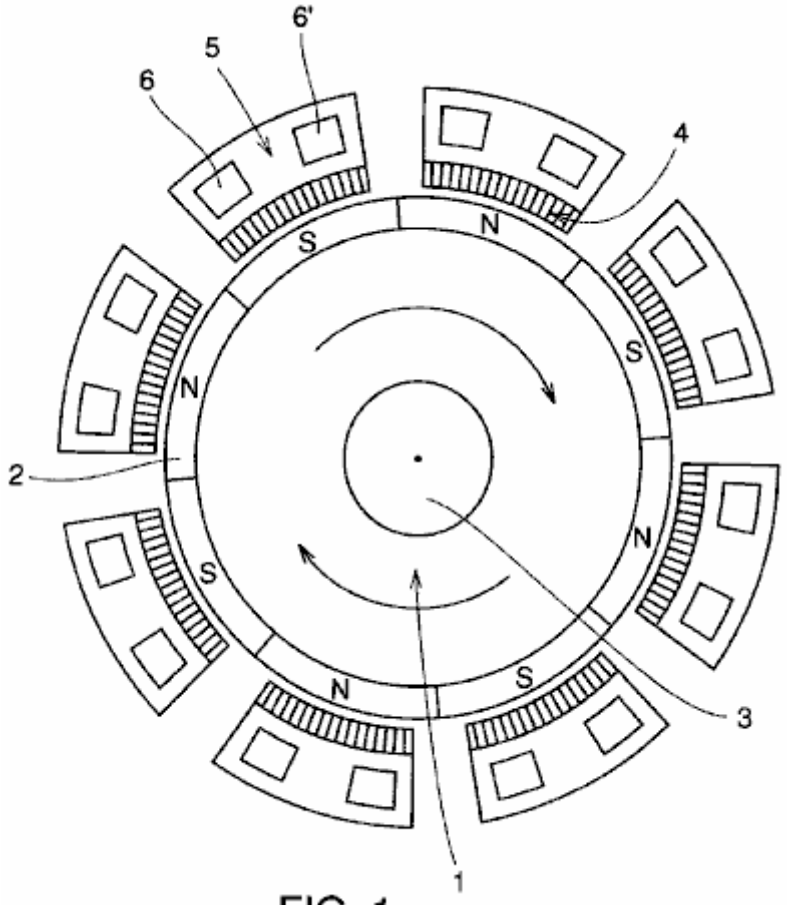


FIG. 1

Referring to **Fig.1** and **Fig.2**, the structure of the load-free generator of the invention is now described. Permanent magnet trains **2** and **2'** in ring forms are attached and fixed to respective left and right orbits **1** and **1'** provided relative to rotational axis **3** with a certain interval between them. Permanent magnet trains **2** and **2'** are fixed onto left and right orbits **1** and **1'** respectively such that the polarities on the outer peripheral surface of each magnet train relative to the rotational axis are alternately N poles and S poles. The permanent magnet trains are rotatable about the axis. Further, the facing polarities of respective permanent magnet train **2** and permanent magnet train **2'** relative to the direction of rotational axis **3** are arranged to be opposite.

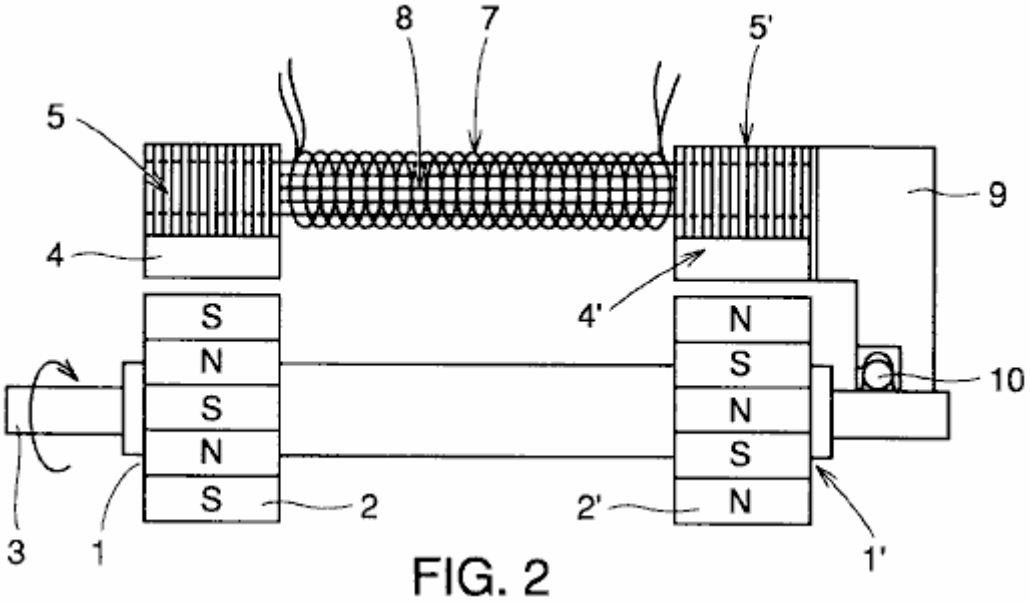


FIG. 2

As shown in **Fig.2**, rotational axis **3** and a case **9** are joined by a bearing **10** at a certain distance from the permanent magnet trains **2** and **2'**.

At a predetermined distance from permanent magnet trains **2** and **2'**, magnetic induction primary cores **4** and **4'** with respective coils wound around them are fixed to case **9**.

In addition, magnetic induction secondary cores **5** and **5'** each having two coupling holes **6** and **6'** formed therein are structured by stacking and coupling a plurality of thin cores attached and fixed to magnetic induction primary cores **4** and **4'** respectively and the secondary cores are attached and fixed to case **9**.

Magnetic induction tertiary cores **8** and **8'** are inserted respectively into coupling holes **6** and **6'** of magnetic induction secondary cores **5** and **5'** so as to couple magnetic induction secondary cores **5** and **5'** of each other.

Responsive coils **7** and **7'** are wound in opposite directions to each other around respective magnetic induction cores **8** and **8'**.

Fig.3 illustrates a structure formed of magnetic induction secondary cores **5** and **5'**, magnetic induction cores **8** and **8'** and responsive coils **7** and **7'** viewed in the direction perpendicular to rotational axis **3**.

As explained above, the directions of windings of responsive coils **7** and **7'** are respectively opposite to each other around magnetic induction cores **8** and **8'** which couple magnetic induction secondary cores **5** and **5'**.

In the structure described in conjunction with **Fig.1**, **Fig.2** and **Fig.3**, when rotational axis **3** of the generator rotates, permanent magnetic trains **2** and **2'** accordingly rotate to generate magnetically sensitive currents (electromagnetically induced current) in responsive coils **7** and **7'** and the current thus produced can be drawn out for use.

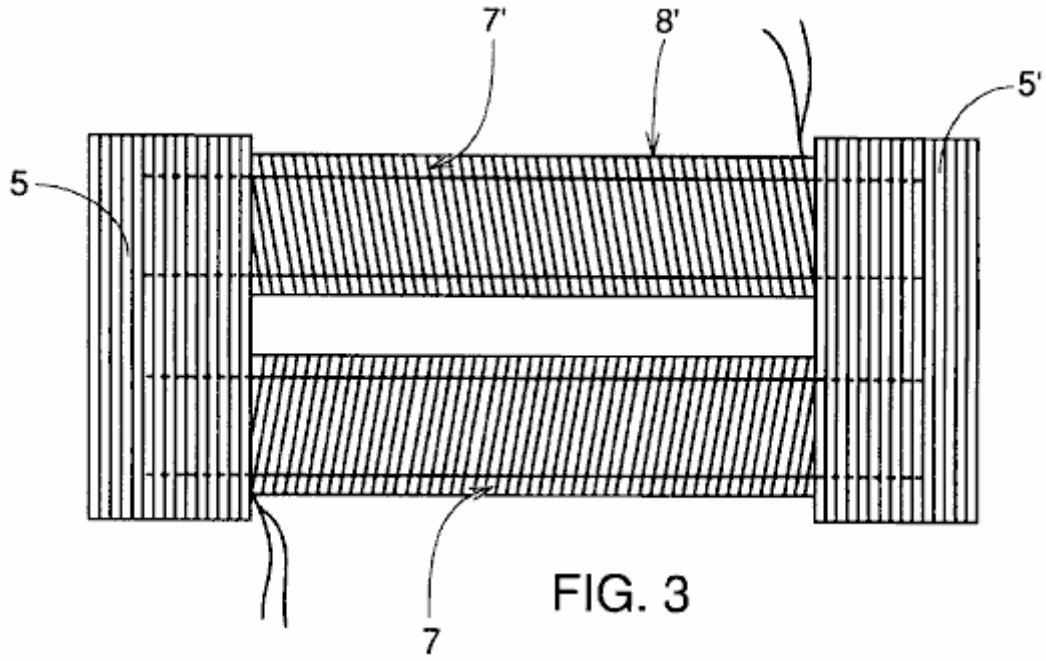


FIG. 3

As shown in **Fig.3**, the coils are wound about magnetic induction cores **8** and **8'** respectively in the opposite directions in the generator of the present invention, and the directions of the magnetic fields generated by the flow of the induced currents are arranged such that the N pole and S pole alternately occurs around rotational axis **3**.

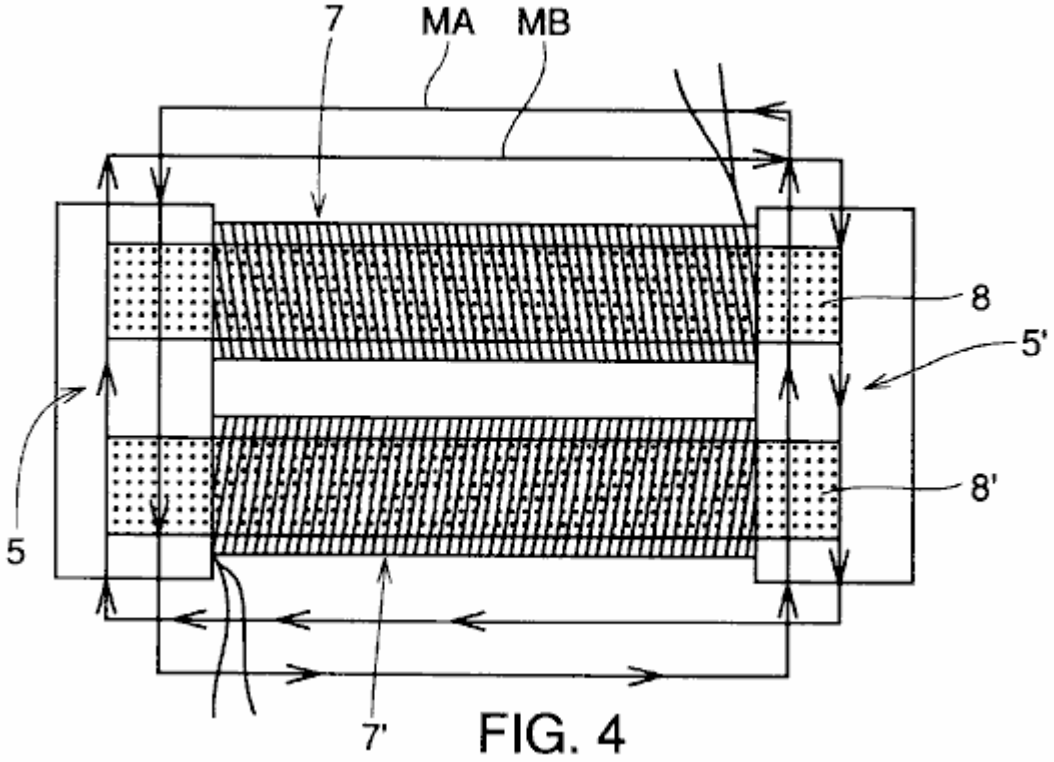


FIG. 4

Fig.4 illustrates magnetic fields induced in a set of magnetic induction secondary cores **5** and **5'**, magnetic induction cores **8** and **8'** and responsive coils **7** and **7'**.

At iron strips on both ends of respective magnetic induction secondary cores **5** and **5'**, a reverse current magnetic field is generated by responsive coil **7** upon the rotation of N and S poles of permanent magnet trains **2** and **2'** is in the direction of **MA** shown in **Fig.4**, for example, while a reverse current magnetic field generated by responsive coil **7** is in the direction of **MB** in **Fig.4**. Consequently, the reverse magnetic fields generated by the flow of currents cancel each other. The cores are formed of a plurality of iron strips in order to eliminate heat generated by eddy currents.

The magnetic field of the rotor thus has no dependence on the flow of currents, the load caused by the induced magnetisation phenomenon disappears, and energy of movement necessary for rotation against the mechanical primary load of the rotor itself is applied to the rotor.

At this time, a magnetic circuit including magnetic induction secondary cores **5** and **5'** and magnetic induction tertiary cores **8** and **8'** should be shaped into ".quadrature." form. If the circuit does not structured as ".quadrature." form, a part of the reverse magnetic field functions as electrical force which hinders the rotational force of the rotor.

Further, permanent magnet trains **2** and **2'** of the rotor are arranged to have opposite poles to each other on the left and right sides as shown in **Fig.2** so as to constitute the flow of magnetic flux. Each rotor has alternately arranged magnets, for example, eight poles are provided to enhance the generator efficiency.

More detailed description of the operational principle is given now. When the rotor in **Fig.1** rotates once, S and N poles of permanent magnets **2** and **2'** attached to the periphery of the rotor successively supply magnetic fields to induction primary cores **4** above, and magnetic field is accordingly generated in a path from one orbit of the rotor along induction primary core **4**, induction secondary core **5**, induction tertiary core **8**, induction secondary core **5'**, induction primary core **4'** to the other orbit of the rotor as shown in **Fig.2**.

Accordingly, current flows in the coils affected by this electric field to generate electric power. For example, if the generated power is used as generated output for switching on an electric light or for using it as motive energy, the current flowing through the coils generates the reverse magnetic fields. However, this reverse magnetic fields do not influence permanent magnets **2** and **2'** attached to the rotor in **Fig.2** since the reverse magnetic fields of the same magnitude respectively of S and N or N and S on both ends of magnetic induction secondary cores **5** and **5'** cancel out each other as shown in **Fig.4**. Because of this, the rotor is in a no-load state in which any resistance except the weight of the rotor itself and dynamic resistance is not exerted on the rotor.

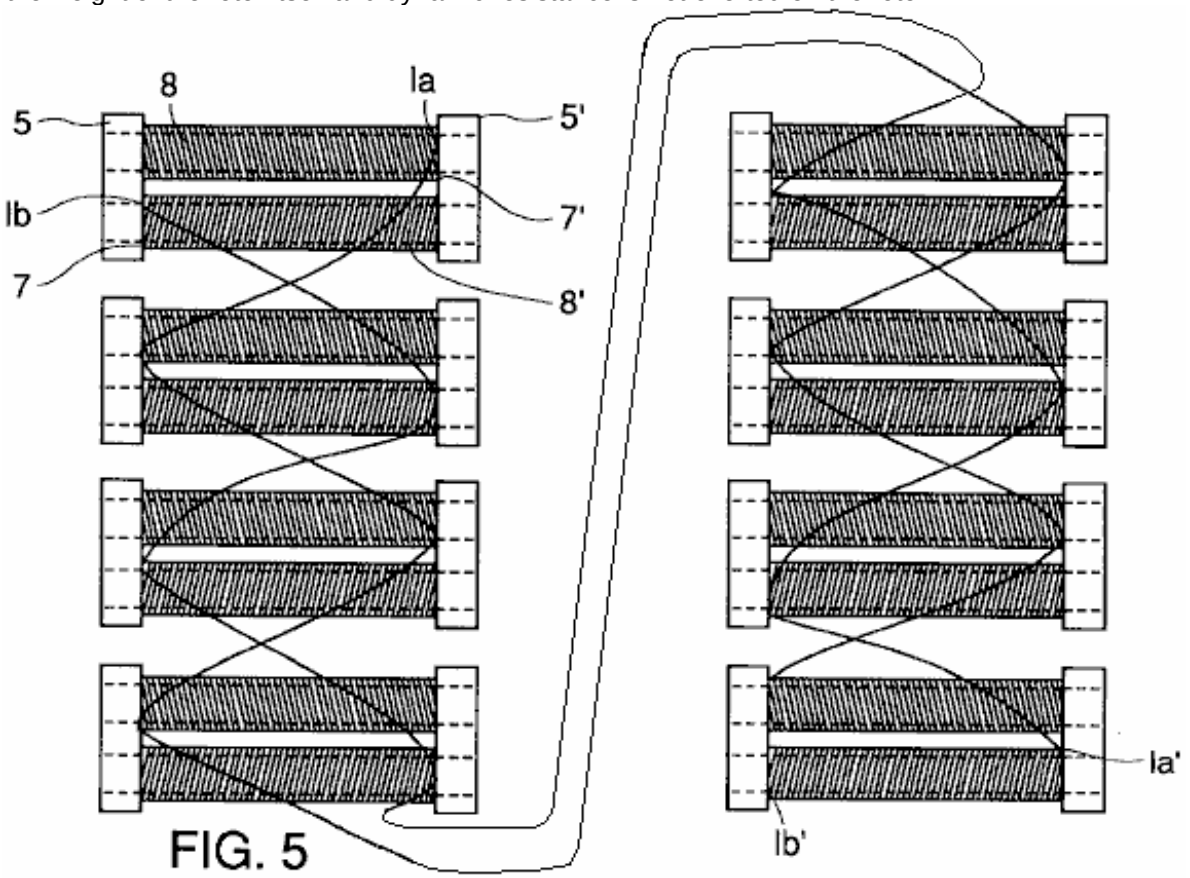


FIG. 5

Fig.5 illustrates a manner of connecting magnetically responsive coils **7** and **7'** wound around magnetic induction tertiary cores **8** and **8'** with eight poles.

Referring to **Fig.5**, according to a method of connecting magnetically responsive coils **7** and **7'** , line **1a1** of responsive coil **7'** (one drawn-out line of the wire coiled around a first magnetic induction core **8**) is connected to line **1a2'** (one drawn-out line of the wire coiled around a second magnetic induction core **8**), and then line **1a2** (the other drawn-out line of the wire coiled around a second magnetic induction core **8**) is connected to line **1a3'**, and subsequently lines **1a** and **1a'** are connected successively in zigzag manner to allow current to flow. Further, responsive coil **7** is arranged to connect lines represented by **1b1** in zigzag manner such that lines **1b** and **1b'** are successively connected. In this way, lines **1b**, **1b'** and lines **1a** and **1a'** of respective magnetically responsive coils **7** and **7'** are connected. As a whole, total four electric wires are drawn out for use.

When electric power is to be generated according to the present invention as described above, specifically, a closed circuit is formed by responsive coils **7** and **7'**, electric currents are induced in responsive coils **7** and **7'** wound around the magnetic induction cores of the generator, and the induced magnetic fields produced respectively by responsive coils **7** and **7'** could cause a great load which interferes the rotational force of the rotor. However, as shown in **Fig.4**, the direction of convolution of one coil **7** is opposite to that of the other coil **7'** so that the magnetic force generated by the reverse currents (induced currents) in responsive coils **7** and **7'** wound around magnetic induction core **4** is not transmitted to magnetic induction cores **8** and **8** accordingly no reverse magnetic force is transmitted to permanent magnets **2** and **2'**.

Therefore, each time the N poles and S poles alternate with each other because of the alternation of permanent magnets **2** and **2'** shown in **Fig.2**, the reverse magnetic forces in the right and left direction opposite to the direction of arrows denoted by **MA** and **MB** completely disappear as shown in **Fig.4**. Consequently, the reverse magnetic forces caused by the reverse currents are not influenced by permanent magnets **2** and **2'** and accordingly no load except the mechanical primary load is exerted on the generator of the invention.

As discussed above, the load-free generator of the present invention, secondary load except mechanical load of the generator, i.e. the load caused by the reverse currents flowing through the responsive coils can be nulled. With regard to this load-free generator, even if 100% of the current generated by magnetic induction (electromagnetic induction) is used, the magnetic secondary load due to the reverse currents except the mechanical primary load does not serve as load.

Although the number of poles of the rotor is described as **8** in the above description, the present invention is not limited to such a structure, and the invention can exhibit its effect when the smaller or greater number of poles is applied.

Further, although the magnet of the rotor is described as the permanent magnet in the above structure, the invention is not limited to such a case and the magnet of the rotor may be an electromagnet, for example.

In addition, although the description above is applied to the structure of the rotating-field type generator, the generator may be of the rotating-armature type.

EXPERIMENTAL EXAMPLE

More detailed description of the generator of the present invention is hereinafter given based on specific experimental examples of the invention.

The generator of the present invention and a conventional generator were used to measure the electric power production efficiency and the amount of load and compare the resultant measurements.

EXPERIMENTAL EXAMPLE 1

A 12-pole alternating current (AC) generator for battery charging was used, and the electricity output and the load when 50% of the electricity output was used as well as those when 100% of the electricity output was used were measured. The generator above is a single-phase AC motor and the employed power source was 220V, with 1750 rpm and the efficiency of 60%. The result of measurement using power of a motor of 0.5HP and ampere .times.volt gauge is shown in **Table 1**.

EXPERIMENTAL EXAMPLE 2

Measurement was done under the same conditions as those of experimental example 1 and a generator used was the one which was made according to the present invention to have the same conditions as those of the product of the existing model above. The result of measurement using ampere x volt gauge is shown in **Table 1**.

Table 1

	50% Electricity Used		100% Electricity Used	
Type of Generator	Electricity Output (Watts)	Amount of Load (Watts)	Electricity Output (Watts)	Amount of Load (Watts)
Conventional:	100	221	14	347
This invention:	100	220	183	200

(electricity output and load amount of the alternating current generators when 50% and 100% of the electricity were used)

From the result of Experimental Example 1 above, the reason for the remarkable reduction of the electricity output when the electricity consumption was 100% relative to the electricity consumption of 50% in the conventional generator is considered to be the significant increase of the repulsive load exerted on the generator when 100% of the electricity is used.

On the other hand, in the generator of the present invention, there was no appreciable difference in the amount of load between those cases in which 50% of the electricity was used and 100% thereof was used respectively. Rather, the amount of load slightly decreased (approximately 20W) when 100% of the electricity was used. In view of this, it can be understood that the amount of generated electric power of the generator of the present invention is approximately doubled as the electricity consumption increases, which is different from the conventional generator producing electric power which sharply decreases when the electricity consumption increases.

In conclusion, the amount of load above is supposed to be numerical value relative to the mechanical load of the generator as described above. Any secondary load except this, i.e. load due to the reverse currents generated in the armature responsive coils can be confirmed as zero.

EXPERIMENTAL EXAMPLE 3

12V direct current (DC) generators having similar conditions to those in experimental example 1 were used to make measurement under the same conditions (efficiency 80%). The result of the measurement is presented below.

Table 2

	50% Electricity Used		100% Electricity Used	
Type of Generator	Electricity Output (Watts)	Amount of Load (Watts)	Electricity Output (Watts)	Amount of Load (Watts)
Conventional:	103	290	21	298
This invention:	107	282	236	272

(electricity output and load amount of the alternating current generators when 50% and 100% of the electricity were used)

The DC generator has higher efficiency (80%) than that of the AC generator, while use of the brush increases the cost of the DC generator. When 100% of the electricity was used, the amount of load slightly decreased which was similar to the result shown in Table 1 and the electricity output was approximately at least 2.2 times that when 50% of the electricity was used.

EXPERIMENTAL EXAMPLE 4

A 220V single-phase alternating current (AC) generator (0.5HP) having similar conditions to those in experimental example 1 was used, and the rotation per minute (rpm) was changed to make measurement under the condition of 100% consumption of the generated electricity. The result of measurement is illustrated in the following **Table 3**.

Table 3

1750 rpm		3600 rpm		5100 rpm	
Electricity Output (Watts)	Amount of Load (Watts)	Electricity Output (Watts)	Amount of Load (Watts)	Electricity Output (Watts)	Amount of Load (Watts)
130	160	210	228	307	342

(amounts of generated electric power and load when the rotation per minute of the generator of the present invention was varied)

As shown in **Table 3** above, as the rotation per minute (rpm) increases as from 1750, 3600 to 5100, the amount of electric power increases respectively from 130, 210 to 307W and consequently the difference between the amount of generated electric power and the amount of load decreases to cause relative decrease of the amount of load as the rotation per minute (rpm) increases.

EXPERIMENTAL EXAMPLE 5

Measurement was done by changing the number of N and S poles of the permanent magnets of the invention under the same conditions as those of experimental example 1 and under the condition that 100% of the generated electricity was used.

The result of the measurement is illustrated below.

Table 4

2 poles		4 poles		8 poles	
Electricity Output (Watts)	Amount of Load (Watts)	Electricity Output (Watts)	Amount of Load (Watts)	Electricity Output (Watts)	Amount of Load (Watts)
80	152	130	200	265	296

(amounts of generated electric power and load when the number of poles of the permanent magnets of the generator of the invention was changed)

From **Table 4** above, it can be understood that as the number of poles increases, both of the amounts of generated electric power and load increase. However, the ratio of the amount of generated electric power to the amount of load monotonously increases. In the table above, in terms of the amount of load, only the mechanical primary load is exerted and electrical secondary is not exerted.

The increase of the number of poles causes increase, by the number of increased poles, in the number of lines of magnetic flux which coils traverse, and accordingly the electromotive force increases to increase the amount of generated electric power. On the other hand, the amount of mechanical load has a constant value regardless of the increase of the number of poles, so that the mechanical load amount relatively decreases to reduce the difference between the amount of load and the amount of generated electric power.

Detailed description of the present invention which has been given above is just for the purpose of presenting example and illustration, not for limitation. It will dearly be appreciated that the spirit and scope of the invention will be limited only by the attached scope of claims.

ALBERTO MOLINA-MARTINEZ: ELECTRICAL GENERATOR

Patent Application US 20020125774 6th March 2002 Inventor: Alberto Molina-Martinez

CONTINUOUS ELECTRICAL GENERATOR

This patent application shows the details of a device which it is claimed, can produce sufficient electricity to power both itself and external loads. It also has no moving parts.

ABSTRACT

A stationary cylindrical electromagnetic core, made of one piece thin laminations stacked to desired height, having closed slots radially distributed, where two three-phase winding arrangements are placed together in the same slots, one to the centre, one to the exterior, for the purpose of creating a rotational electromagnetic field by temporarily applying a three-phase current to one of the windings, and by this means, inducting a voltage on the second one, in such a way that the outgoing energy is a lot greater than the input. A return will feedback the system and the temporary source is then disconnected. The generator will run by itself indefinitely, permanently generating a great excess of energy.

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention relates generally to electrical power generating systems. More specifically, the present invention relates to self-feeding electrical power generating units.

2. Description of Related Art

Since Nikola Tesla invented and patented his Polyphase System for Generators, Induction Motors and Transformers, no essential improvement has been made in the field. The generators would produce the polyphase voltages and currents by means of mechanical rotational movement in order to force a magnetic field to rotate across the generator's radially spaced windings. The basis of the induction motor system was to create an electro-magnetically rotating field, instead of a mechanically rotated magnetic field, which would induce voltages and currents to generate electromotive forces usable as mechanical energy or power. Finally, the transformers would manipulate the voltages and currents to make them feasible for their use and transmission for long distances.

In all present Electric Generators a small amount of energy, normally less than one percent of the outgoing power in big generators, is used to excite the mechanically rotated electromagnetic poles that will induce voltages and currents in conductors having a relative speed or movement between them and the polar masses.

The rest of the energy used in the process of obtaining electricity, is needed to move the masses and to overcome the losses of the system: mechanical losses; friction losses; brushes losses, windage losses; armature reaction losses; air gap losses; synchronous reactance losses; eddy current losses; hysteresis losses, all of which, in conjunction, are responsible for the excess in power input (mechanical power) required to generate always smaller amounts of electric power.

SUMMARY OF THE INVENTION

The **C**ontinuous **E**lectrical **G**enerator consists of a stationary cylindrical electromagnetic core made of one piece thin laminations stacked together to form a cylinder, where two three-phase windings arrangements are placed in the same slots not having any physical relative speed or displacement between them. When one of the windings is connected to a temporary three-phase source, an electromagnetic rotating field is created, and the field this way created will cut the stationary coils of the second winding, inducting voltages and currents. In the same way and extent as in common generators, about one percent or less of the outgoing power will be needed to keep the rotational magnetic field excited.

In the **C**ontinuous **E**lectrical **G**enerator there are no mechanical losses; friction losses; brush losses; windage losses; armature reaction losses; or air gap losses, because there is not any movement of any kind. There are: synchronous reactance losses, eddy current losses and hysteresis losses, which are inherent to the design, construction and the materials of the generator, but in the same extent as in common generators.

One percent or less of the total energy produced by present electric generators goes to create their own magnetic field; a mechanical energy that exceeds the total output of present generators is used to make them rotate in the process of extracting electrical currents from them. In the **C**ontinuous **E**lectrical **G**enerator there is no need for movement since the field is in fact already rotating electro-magnetically, so all that mechanical energy will not be needed. Under similar conditions of exciting currents, core mass and windings design, the **C**ontinuous **E**lectrical **G**enerator is significantly more efficient than present generators, which also means that it can produce significantly more than the energy it needs to operate. The **C**ontinuous **E**lectrical **G**enerator can feedback the system, the temporary source may be disconnected and the Generator will run indefinitely.

As with any other generator, the **C**ontinuous **E**lectrical **G**enerator may excite its own electromagnetic field with a minimum part of the electrical energy produced. The **C**ontinuous **E**lectrical **G**enerator only needs to be started up by connecting its inducting three-phase windings to a three-phase external source for an instant, and then to be disconnected, to start the system as described herein. Then, disconnected, it will run indefinitely generating a great excess of electric power to the extent of its design.

The **C**ontinuous **E**lectrical **G**enerator can be designed and calculated with all mathematical formulas in use today to design and calculate electrical generators and motors. It complies with all of the laws and parameters used to calculate electrical induction and generation of electricity today.

Except for the Law of Conservation of Energy, which, by itself, is not a mathematical equation but a theoretical concept and by the same reason does not have any role in the mathematical calculation of an electrical generator of any type, the **C**ontinuous **E**lectrical **G**enerator complies with all the Laws of Physics and Electrical Engineering. The **C**ontinuous **E**lectrical **G**enerator obligates us to review the Law of Conservation of Energy. In my personal belief, the electricity has never come from the mechanical energy that we put into a machine to move the masses against all oppositions. The mechanical system is actually providing the path for the condensation of electricity. The **C**ontinuous **E**lectrical **G**enerator provides a more efficient path for the electricity.

DESCRIPTION OF DRAWINGS

Fig.1 shows one embodiment of the present invention.

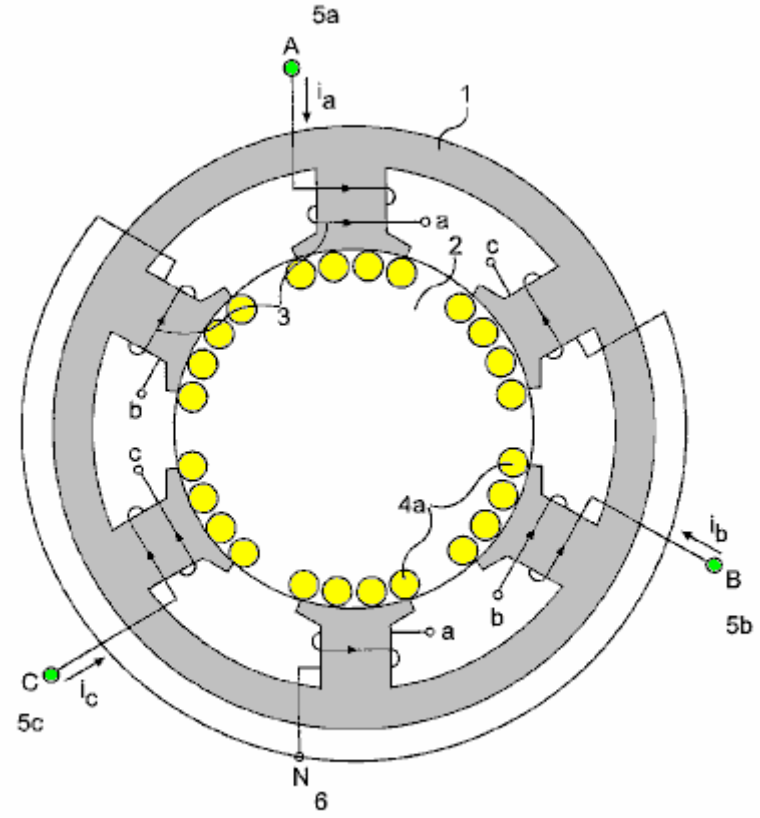


FIG. 1

Fig.2 shows an internal wiring diagram for the embodiment of the present invention shown in **Fig.1**.

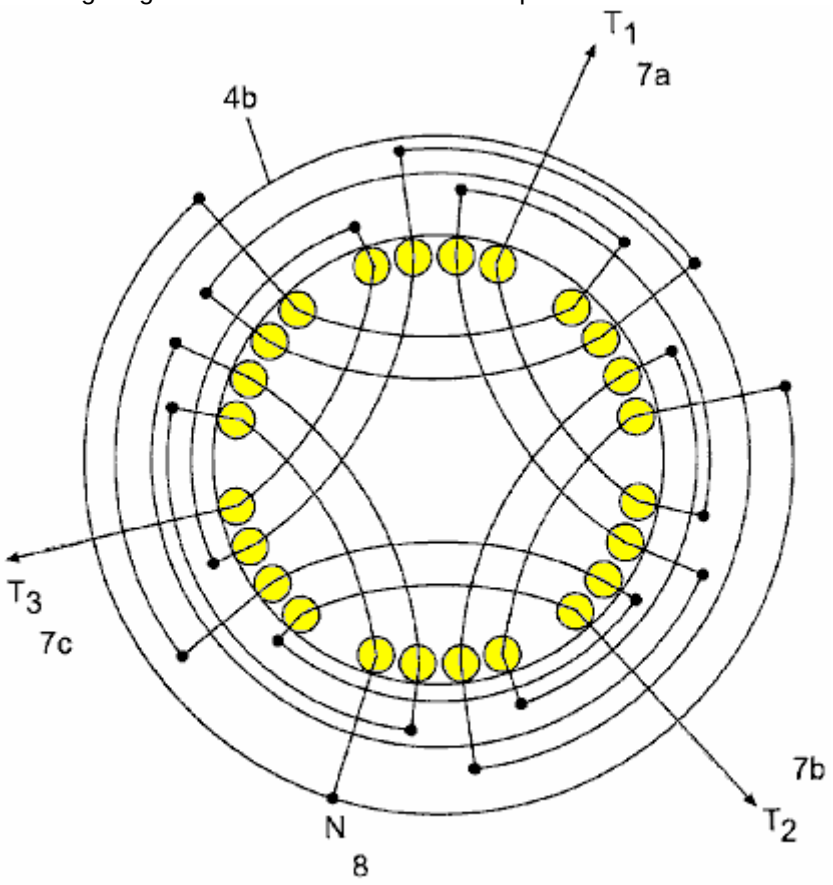


FIG. 2

Fig.3 shows a single laminate for an alternate embodiment of the present invention.

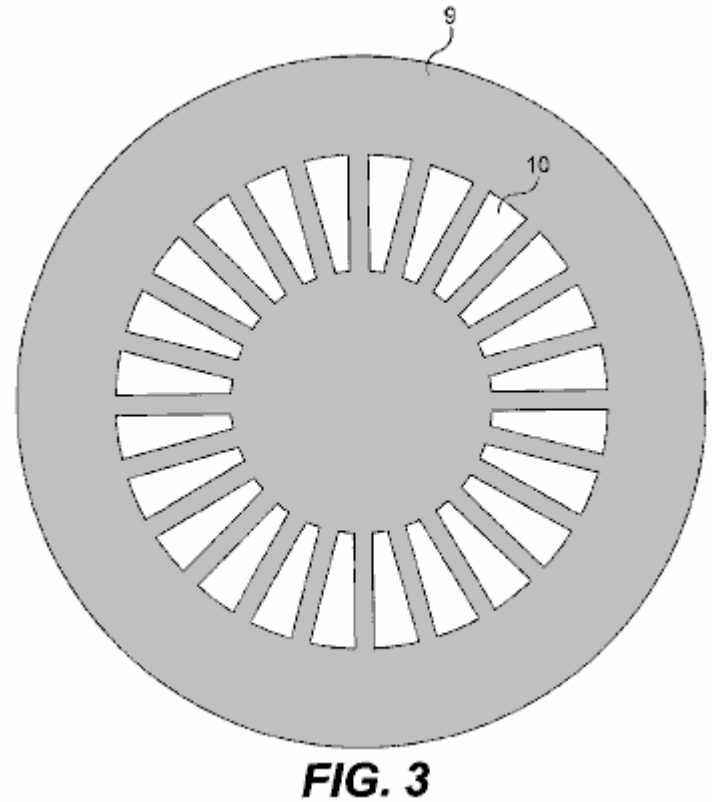


FIG. 3

Fig.4 shows a two-piece single laminate for another alternate embodiment of the present invention.

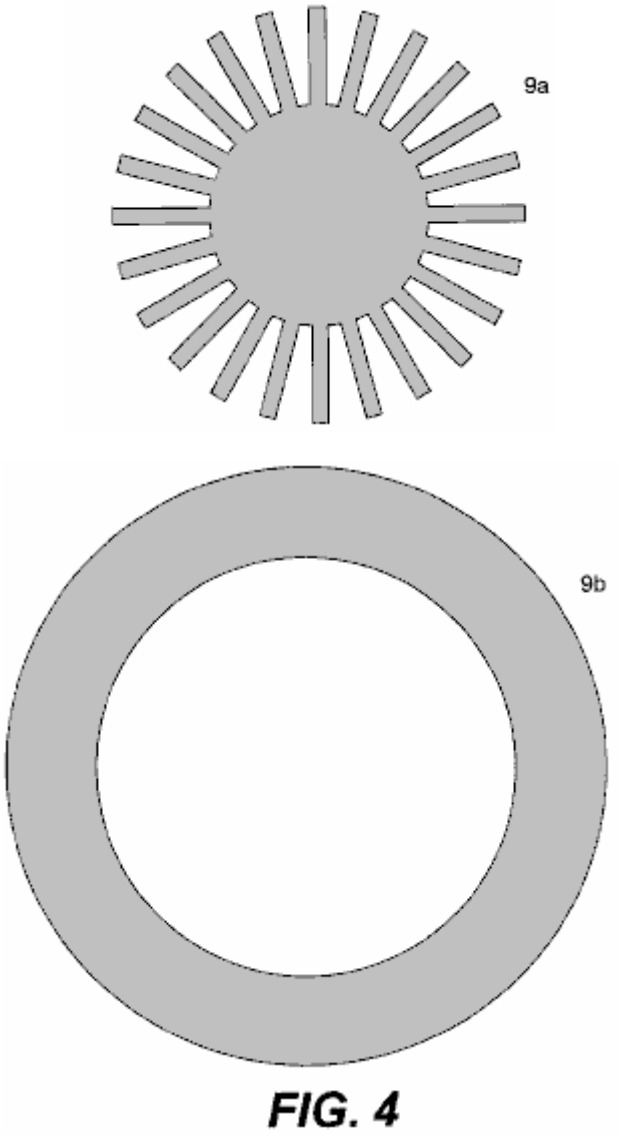


FIG. 4

Fig.5 shows a wiring diagram for an embodiment of the present invention constructed from the laminate shown in **Fig.3** or **Fig.4**.

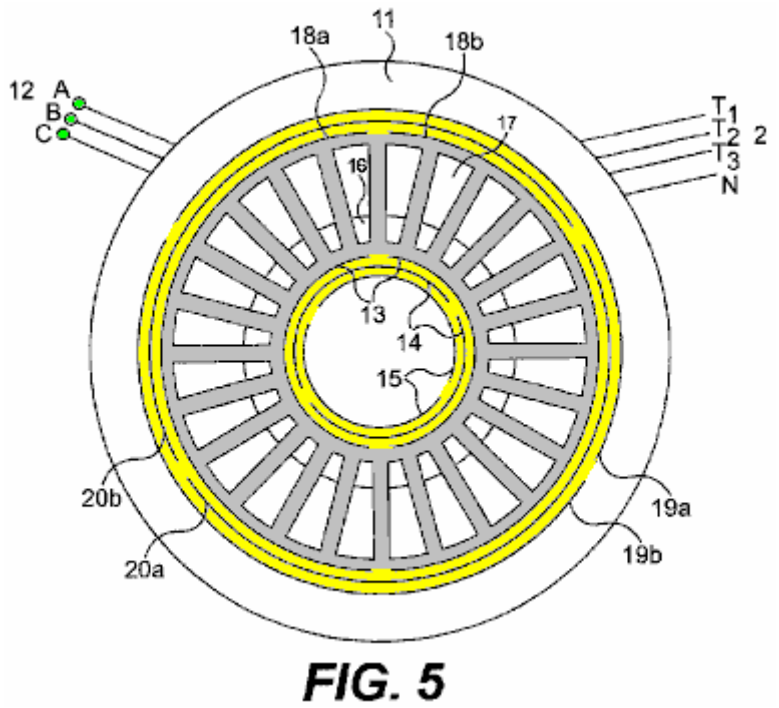


FIG. 5

Fig.6 shows the magnetic flux pattern produced by the present invention.

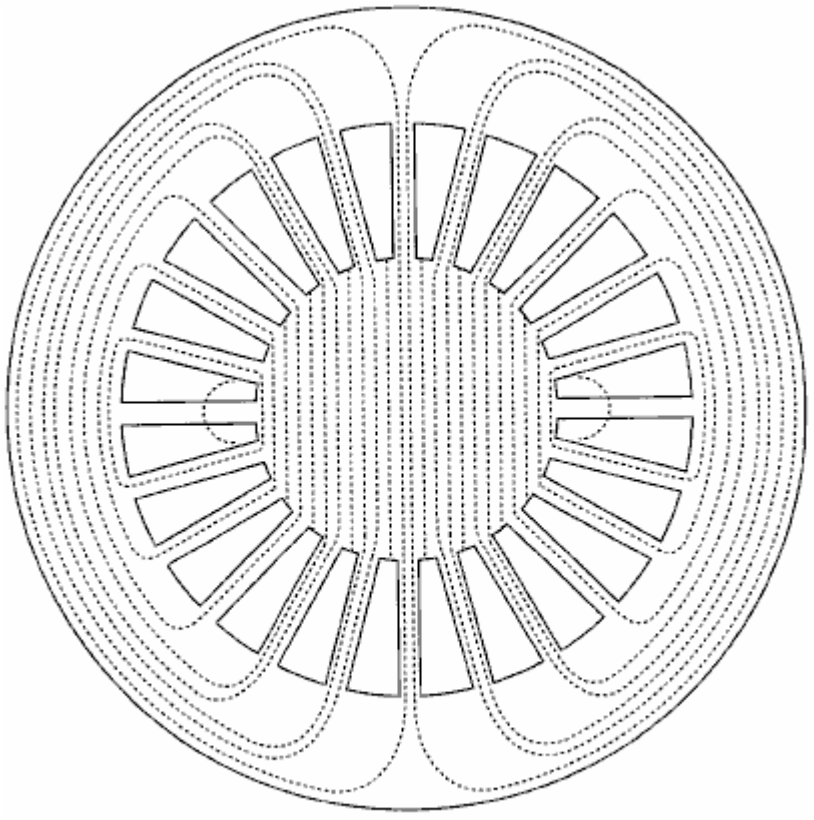


FIG. 6

Fig.7 shows the rotational magnetic field patterns produced by the present invention.

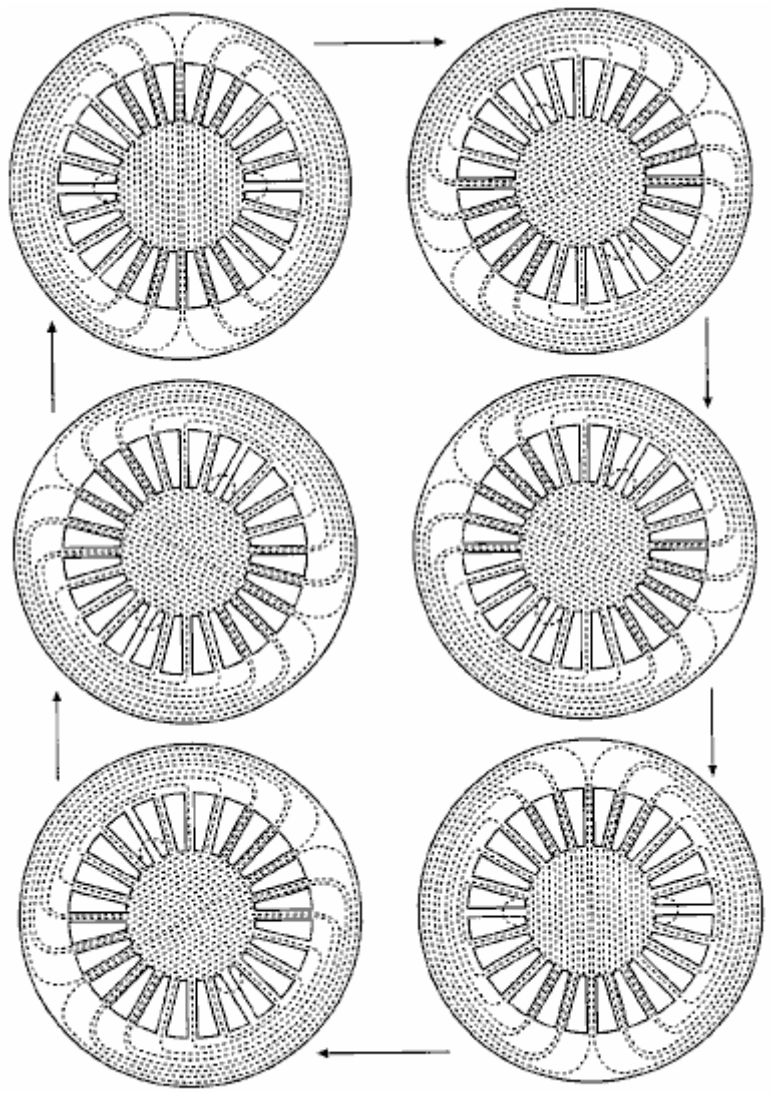


FIG. 7

Fig.8 shows the complete system of the present invention.

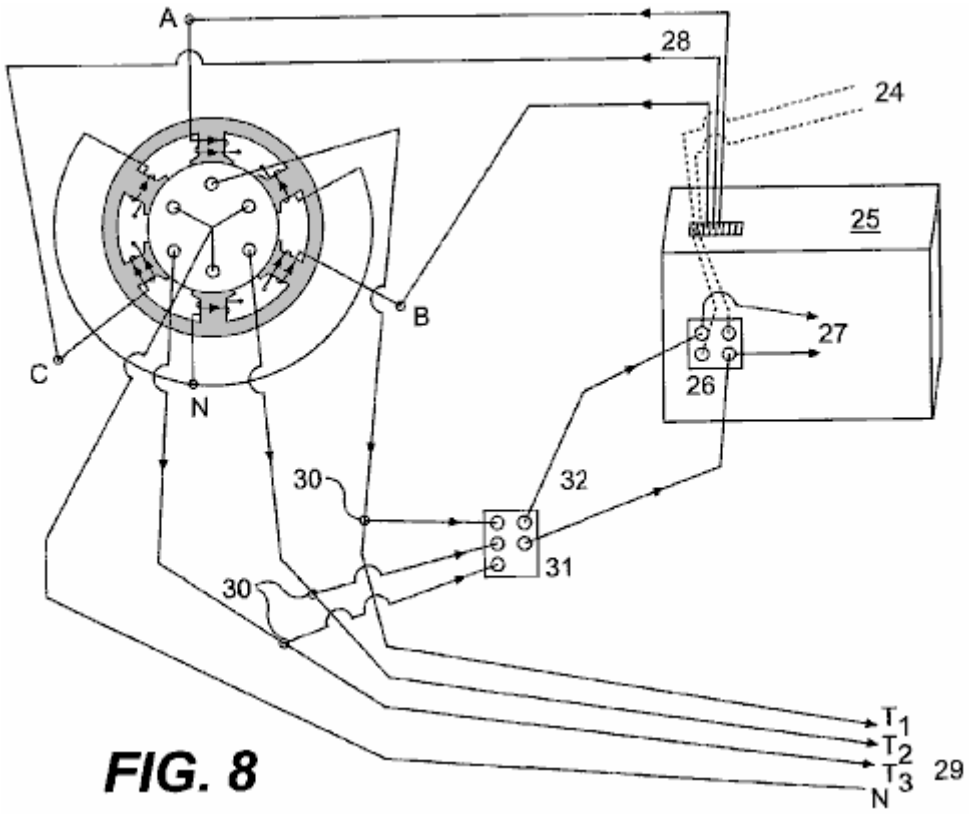


FIG. 8

Fig.9 is an expanded view of the alternate embodiment of the present invention shown in **Fig.3** or **Fig.4**.

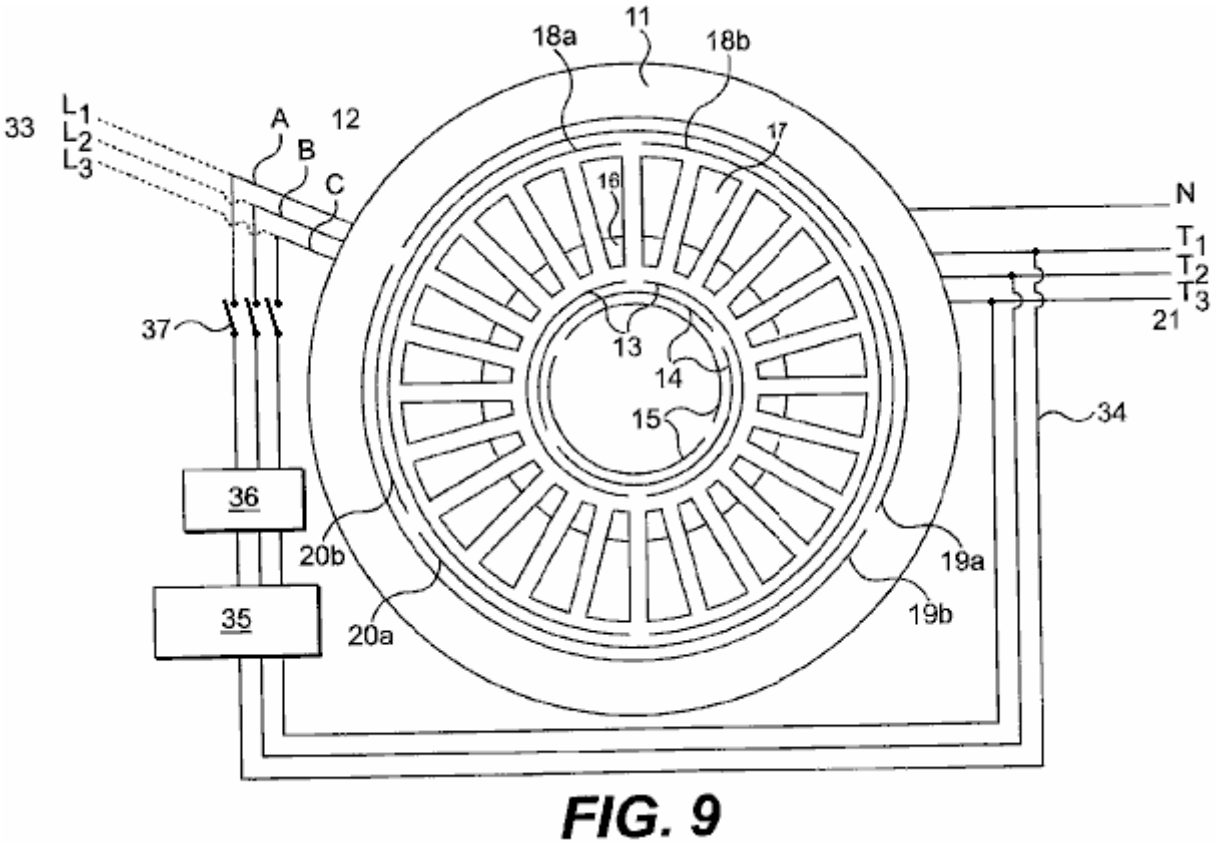


FIG. 9

DETAILED DESCRIPTION OF THE INVENTION

The present invention is a Continuous and Autonomous Electrical Generator, capable of producing more energy than it needs to operate, and which provides itself the energy needed to operate. The basic idea consists in the induction of electric voltages and currents without any physical movement by the use of a rotational magnetic field created by a three-phase stator connected temporarily to a three-phase source, and placing stationary conductors on the path of said rotational magnetic field, eliminating the need of mechanical forces.

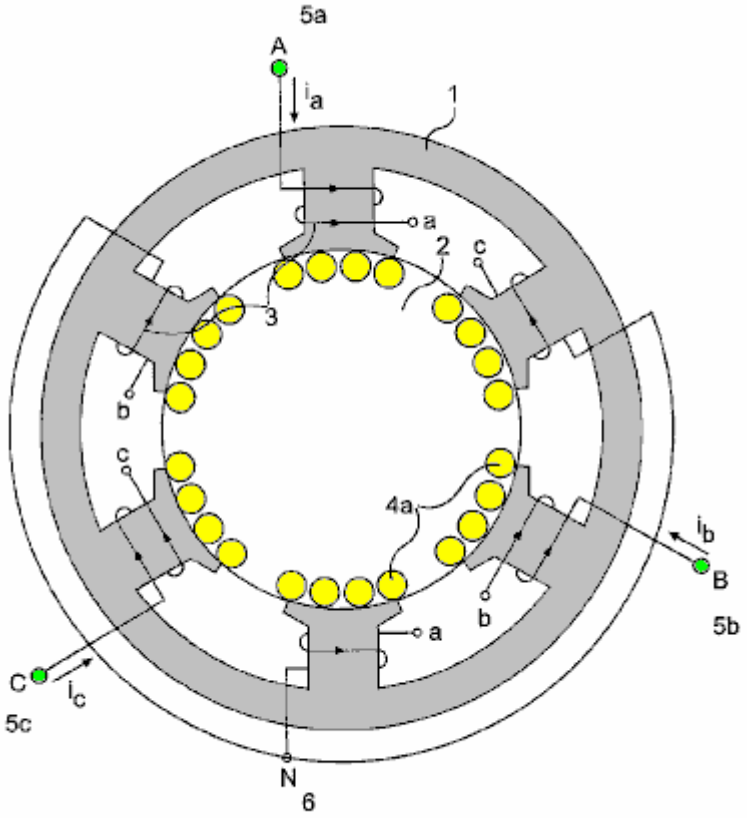


FIG. 1

The basic system can be observed in **Fig.1**, which shows one embodiment of the present invention. There is a stationary ferromagnetic core **1** with a three-phase inducting windings **3**, spaced 120 degrees and connected in Y **6** in order to provide a rotating electromagnetic field, when a three-phase voltage is applied; for the case, a two-pole arrangement. Inside this core **1** there is a second stationary ferromagnetic core **2**, with no space between them, this is, with no air-gap. This second core **2** has also a three-phase stationary winding arrangement (**4a** in **Fig.4b** and **4b** in **Fig.2**), aligned as shown in **Fig.1** and **Fig.2** with the external core inducting windings **3**. There is not any movement between the two cores, since there is no air-gap between them.

There is no shaft on either core since these are not rotating cores. The two cores can be made of stacked insulated laminations or of insulated compressed and bonded ferromagnetic powder. The system works either way, inducting three-phase voltages and currents on the stationary conductors **4a** of the internal windings **4b**, applying three-phase currents to terminals **A 5a**, **B 5b** and **C 5c** of the external windings **3**; or inducting three-phase voltages and currents on the external windings **3**, by applying three-phase currents to the terminals **T1 7a**, **T2 7b** and **T3 7c**, of the internal windings **4b**. When a three-phase voltage is applied to terminals **A 5a**, **B 5b** and **C 5c**, the currents will have the same magnitude, but will be displaced in time by an angle of 120 degrees. These currents produce magneto motive-forces, which, in turn, create a rotational magnetic flux. The arrangements may vary widely as they occur with present alternators and three-phase motors, but the basics remain the same, a stationary but electro-magnetically rotating magnetic field, inducting voltages and currents on the stationary conductors placed on the path of said rotating magnetic field. The diagram is showing a two-pole arrangement for both windings, but many other arrangements may be used, as in common generators and motors.

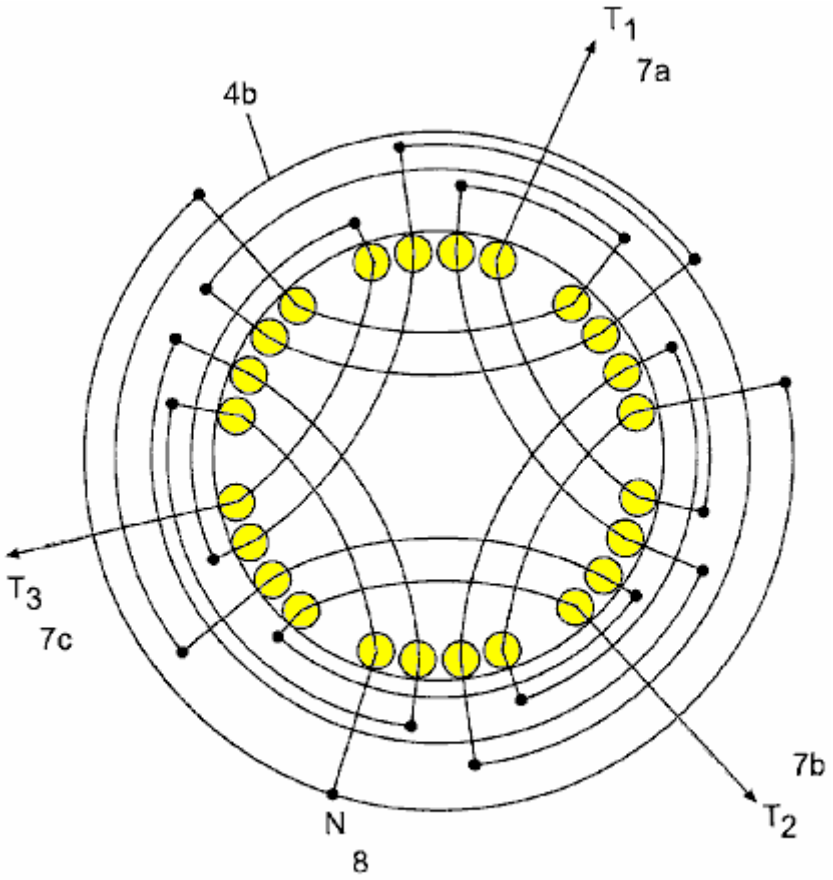


FIG. 2

Fig.2 shows the three-phase arrangement of the internal winding **4b** which has provided, in practice, symmetrical voltages and currents, due to a space angle of 120 degrees. It is similar to a two-pole arrangement. Many other three-phase or poly-phase arrangements may be used. Wherever a conductor is crossed by a rotational magnetic field, a voltage will be induced across its terminals. The interconnections depend on the use that we will give to the system. In this case, we will have a three-phase voltage in terminals **T1 7a, T2 7b** and **T3 7c** and a neutral **8**. The outgoing voltage depends on the density of the rotational magnetic flux, the number of turns of the conductor, the frequency (instead of the speed) and the length of the conductor crossed by the field, as in any other generator.

Fig.3 shows an alternate embodiment of the present invention in which the generator is made from multiple one-piece laminations **9**, stacked as a cylinder to the desired height. This embodiment can also be made of a one-piece block of compressed and bonded insulated ferromagnetic powder. The same slot **10** will accommodate the internal **4a/4b** and the external windings **3**, that is, the inducting and the induced windings (see **Fig.5**). In this case, a 24-slot laminate is shown, but the number of slots may vary widely according to the design and needs.

Fig.4 shows a two-piece single laminate for another alternate embodiment of the present invention. For practical effects the lamination can be divided into two pieces **9a, 9b**, as shown, to facilitate the insertion of the coils. Then, they are solidly assembled without separation between them, as if they were only one piece.

The laminates described above may be constructed with thin (0.15 mm thick or less) insulated laminations **9** or **9a** and **9b** of a high magnetic permeability material and low hysteresis losses such as Hiperco 50A, or similar, to reduce losses or with compressed electrically isolated ferromagnetic powder, which has lower eddy current losses and also may have low hysteresis losses, which can make the generator highly efficient.

OPERATING THE GENERATOR

The **C**ontinuous **E**lectrical **G**enerator as described and shown in the following drawings is designed and calculated to produce a strong rotating electromagnetic field with low exciting currents. By using a laminated material, such as the said Hiperco 50A, we can achieve rotating magnetic fields above two Teslas, since there are no air gap losses, mechanical losses, windage losses, armature reaction losses, etc. as said before. This may be obtained by applying a temporary three-phase current to the terminals **A**, **B** and **C 12** of the inducting coils **13, 14** and **15** (**5a, 5b** and **5c** in **Fig.1**), spaced 120 degrees from each other (see **Fig.5**).

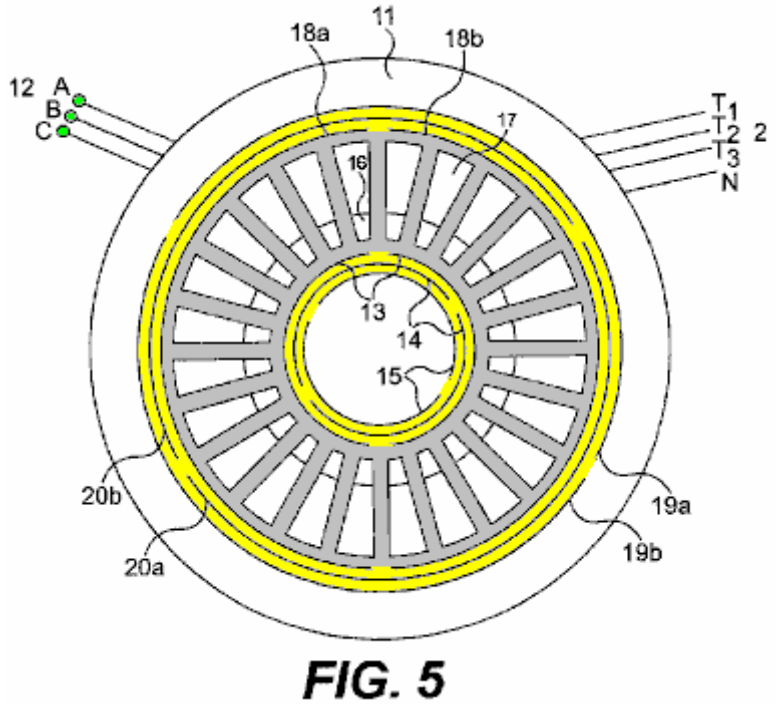


FIG. 5

Fig.5 shows the spatial distribution of the inducting windings **13, 14** and **15**, as well as the induced windings **18a, 18b, 19a, 19b, 20a** and **20b**. Both, the inducting and the induced windings are placed in the same slots **10** or **16** and **17**, with similar arrangements. Even though the system works in both directions, the better configuration seems to be to place the inducting windings **13, 14** and **15**, to the centre and the induced windings **18a, 18b, 19a, 19b, 20a** and **20b**, to the exterior, since small windings will be needed to induce a very strong rotational magnetic field, due to the small losses involved in the process, and in exchange, bigger and powerful windings will be needed to extract all the energy that the system will provide. Both windings are connected in Y (not shown), but they can be connected in different ways, as any other generator. These arrangements are equivalent to the arrangements shown for the embodiment in **Fig.1** and **Fig.2**.

The inducting coils **13, 14** and **15** are designed and calculated so that the generator may be started with common three-phase lines voltages (230 Volts 60 Hz per phase, for example). If the local lines voltages are not appropriate, we can control the voltage to the designed level by means of a three-phase variable transformer, an electronic variator or inverter etc. Once we have such strong magnetic field rotating and crossing the stationary induced coils **18a, 18b, 19a, 19b, 20a** and **20b**, a three-phase voltage will be induced across terminals **T1, T2, T3** and **N 21** in proportion to the magnetic flux density, the number of turns in the coils, the frequency used (instead of the speed), the length of the conductors cut by the rotating field, as in any other alternator. We can connect, as we desire in Y or delta, etc., as in any other alternator or generator. The outgoing currents will be three-phase currents (or poly-phase currents depending on the arrangement) and we can have a neutral **21** if we are using a Y connection, as in any other alternator.

The outgoing alternate voltages and currents are perfect sinusoidal waves, perfectly spaced in time, and totally symmetrical. The voltages and currents obtained by this method are usable in any conventional manner. Any voltage can be produced, depending on the design.

Fig.6 shows the magnetic flux pattern produced by the three-phase inducting windings **13, 14** and **15**. This pattern is similar to the pattern of an induction motor's stators. Since there is no air gap; the whole path for the magnetic flux is homogeneous with no change in materials. The core is made of thin insulated laminations of a high magnetic permeability and low hysteresis loss material; eddy current losses are minimal due to the thin lamination. There are no counter fluxes or armature reactions thus the magnetic flux may be near to saturation with a small exciting current or input energy. Due to the time differential between the three phases and the spatial distribution of the inducting windings, a rotational magnetic field will be created in the core, as shown in **Fig.7**.

Once the generator is started, a small part of the energy obtained is sent back (**Fig.8** and **Fig.9**) to feed the inducting coils **3** (in **Fig.1**) or **13, 14** and **15** (in **Fig.5**), as in any other auto-excited alternator or generator. Of course voltages and phases should be perfectly identical and aligned, and if necessary the feedback voltages should be controlled and handled by means of variable transformers, electronic variators, phase shifters (to align phases) or other type of voltage or phase controllers.

One possible method consists of the use of an electronic converter or variator **25** which initially converts two or three lines of alternating current **24** to direct current by an electronic rectifier **26** and then, electronically, converts the direct current **27** to three-phase current **28** to supply three-phase currents spaced in time 120 degrees for the electromagnetic fields **A, B** and **C 3**. Some variators or converters can accept two lines of voltage, while others will accept only a three-phase line voltage. This embodiment uses a variator of 3 kVA that accepts two 220-volt lines.

The rotational magnetic field created by the currents going through the inducting three-phase windings **13**, **14** and **15**, will induce a voltage across the terminals **T1, T2, T3, N, 29** (**7a, 7b, 7c, 8** in **Fig.2**). Then, from the outgoing current lines **29**, a derivation is made **30** to feed back the system, converting the feed back alternate currents, by means of electronic diode rectifiers **31**, to direct current **32** and then feed back the electronic converter or variator **25** to the DC terminals of the electronic rectifier **26** (See **Fig.8**). Once the feedback is connected, the **C**ontinuous **E**lectrical **G**enerator may be disconnected from the temporary source **24**, and will continue generating electric energy indefinitely.

In **Fig.9**, an alternate embodiment of the **C**ontinuous **E**lectrical **G**enerator can be observed. The basic principles remain the same as for the embodiment described above and shown in **Fig.1** and **Fig.2**. The basic differences are in the shape of the laminations and the physical distribution of the windings, as discussed and shown previously. A variation of the feedback, using a variable and shifting transformers is also shown.

The ferromagnetic core **11** is made of one-piece laminates **9** as shown in **Fig.3** (or two for convenience **9a**, **9b** as shown in **Fig.4**) stacked to the desired height. The slots **10**, as indicated before, will accommodate both the inducting **13, 14** and **15** and the induced **18a-b, 19a-b** and **20a-b** windings in the same slot **10** or **16** and **17**. The incoming three phase lines **12** feed the inducting three-phase windings **13, 14** and **15**. They are fed, initially by the temporary source **33** in the first instance, and by the three-phase return **34** once the generator is running by itself.

The inducting windings **13, 14** and **15** have a two-pole arrangement, but many other three-phase or poly-phase arrangements can be made to obtain an electromagnetic rotating field. These windings are connected in Y (not shown) in the same way shown for the embodiment shown in **Fig.1**, **Fig.2** and **Fig.8**, but may be connected in many different ways. The inducting windings **13, 14** and **15** are located in the internal portion **16** of the slot **10** (**Fig.5**).

The induced windings **18a-b**, **19a-b** and **20a-b** have a two-pole arrangement, exactly equal to the arrangement for the inducting windings **13, 14** and **15**, but many other arrangements can be made depending on the design and the needs. The induced windings must be calculated in a way that the generator will have the lowest possible synchronous reactance and resistance. In this way, most of the outgoing power will go to the charge instead of staying to overcome the internal impedance. These windings are connected in Y to generate a neutral **21**, in the same way shown in the embodiment of the present invention shown in **Fig.2**, but may be connected in different ways according to the needs. The induced windings **18a-b, 19a-b** and **20a-b** are located in the external portion **17** of the slot **10**.

The outgoing three-phase and neutral lines **21** come from the induced windings **18a-b, 19a-b** and **20a-b**. The rotational magnetic field created in the core (see **Fig.6** & **Fig.7**) by the inducting windings **13, 14** and **15**, induces a voltage across the terminals **T1, T2** and **T3**, plus a neutral, **29**. From each of the three-phase outgoing lines **21**, a return derivation **34** is made to feedback the system.

The temporary three-phase source **33** is temporarily connected to terminals **A, B** and **C 12**. The **C**ontinuous **E**lectrical **G**enerator must be started with an external three-phase source for an instant, and then disconnected.

Even though the return lines voltage can be calculated and obtained precisely by tabbing the induced windings at the voltage required by the inducting windings (according to the design), it may be convenient to place a three-phase variable transformer or other type of voltage controller **35** in the middle for more precise adjustment of the return voltage.

Placed after the variable transformer **35**, the three-phase shifting transformer **36** will correct and align any phase shift in the voltage and currents angles, before the return is connected. This system functions similarly to the system shown in **Fig.8** which uses a variator or a converter **25**.

Once the voltage and phases are aligned with the temporary source **33**, the return lines **34** are connected to the incoming lines **A, B** and **C 12** at feedback connection **37** and the temporary source **33** is then disconnected. The **C**ontinuous **E**lectrical **G**enerator will remain working indefinitely without any external source of energy, providing a great excess of energy permanently.

The outgoing electric energy provided by this system has been used to produce light and heat, run poly-phase motors, generate usable mono-phase and poly-phase voltages and currents, transform voltages and currents by means of transformers, convert the alternate outgoing poly-phase currents to direct current, as well as for other uses. The electricity obtained by the means described is as versatile and perfect as the electricity obtained today with common electric generators. But the **C**ontinuous **E**lectrical **G**enerator is autonomous and does not depend on any other source of energy but itself once it is running; may be carried anywhere with no limitations; it can be constructed in any size and provides any amount of electricity indefinitely, according to the design.

The **C**ontinuous **E**lectrical **G**enerator is and will be a very simple machine. The keystones of the systems reside in the ultra-low losses of a non-movement generation system, and in a very low synchronous reactance design.

The induced windings must be calculated in a way that the generator may have the lowest possible synchronous reactance and resistance. In this way, most of the outgoing power will go to the charge instead of staying to overcome the internal impedance.

Patent Application US 3,766,094 20th September 1971 Inventor: Michael Ognyanov

SEMICONDUCTOR COMPOSITIONS

This patent application shows the details of a device which it is claimed, can produce electricity via a solid-state oscillator. It should be noted that while construction details are provided which imply that the inventor constructed and tested several of these devices, this is only an application and not a granted patent.

ABSTRACT

A resonance oscillator electric power pack for operating a flash lamp, for example, or other electrically operated device, operates without moving mechanical parts or electrolytic action. The power pack is contained in a cylindrical metal envelope and in a preferred embodiment, is coupled to a relaxation oscillator and an incandescent lamp. Within the envelope, and insulated from it, is a semiconductor tablet having a metal base connected to the external circuit. A metal probe makes contact with a point on the semiconductor tablet and with a cylindrical ferrite rod, axially aligned with the envelope. Wound about the ferrite rod, are concentric helical coils designated as a 'primary' with many turns, and a 'secondary' with fewer turns than the primary.

One end of the primary coil is connected to the probe and the other end is connected to the secondary coil. the leads from the secondary coil are connected to the relaxation oscillator via an adjustable capacitor. Oscillation within the envelope is resonance amplified , and the induced voltage in the secondary coil is rectified for application to the relaxation oscillator and lamp. Selenium and germanium base semiconductor compositions including Te, Nd, Rb and Ga in varying proportions area used for the tablet.

BACKGROUND OF THE INVENTION

This is a continuation-in-part of my co-pending patent application Serial No. 77,452, filed 2nd October 1970, entitled "Electric Power Pack" now abandoned.

In many situations it is desirable to have a source of electric power which is not dependent on wires from a central generating station, and therefore, portable power supplies having no moving parts have been employed. typically, such portable power packs have been primary or secondary electrolytic cells which generate or store electrical energy for release by chemical action. Such batteries have a limited amount of contained energy and must often be replaced at frequent intervals to maintain equipment in operation.

Thus, as one example, flashing lights are commonly used along highways and other locations to warn of dangerous conditions. These flashing lights in remote locations are typically incandescent or gas-discharge lamps connected to some type of relaxation oscillator powered by a battery. The batteries employed in such blinking lights have a limited lifetime and must be periodically replaced, typically each 250 to 300 hours of operation. This involves a rather large labour cost in replacing the expended batteries with fresh ones and additional cost for primary cells or for recharging secondary cells. It is desirable to provide an electric power pack capable of providing a sufficient quantity of electrical energy over a prolonged period of time so that the requirement for periodic replacement of the electrolytic cells can be avoided. Such a power pack is valuable even if appreciably more expensive than batteries because of the greatly reduced labour costs required for periodic replacements.

BRIEF SUMMARY OF THE INVENTION

There is provided in practice of this invention according to a preferred embodiment, semiconductive compositions selected from the Group consisting of:

Selenium with, from 4.85% to 5.5% Tellurium, from 3.95% to 4.2% Germanium, from 2.85% to 3.2% Neodymium, and from 2.0% to 2.5% Gallium.

Selenium with, from 4.8% to 5.5% Tellurium, from 3.9% to 4.5% Germanium, from 2.9% to 3.5% Neodymium and from 4.5% to 5% Rubidium, and

Germanium with, from 4.75% to 5.5% Tellurium, from 4.0% to 4.5% Neodymium and from 5.5% to 7.0% Rubidium.

DRAWINGS

These and other features and advantages of the invention will be appreciated and better understood by reference to the following detailed description of a preferred embodiment when considered in conjunction with the following drawings:

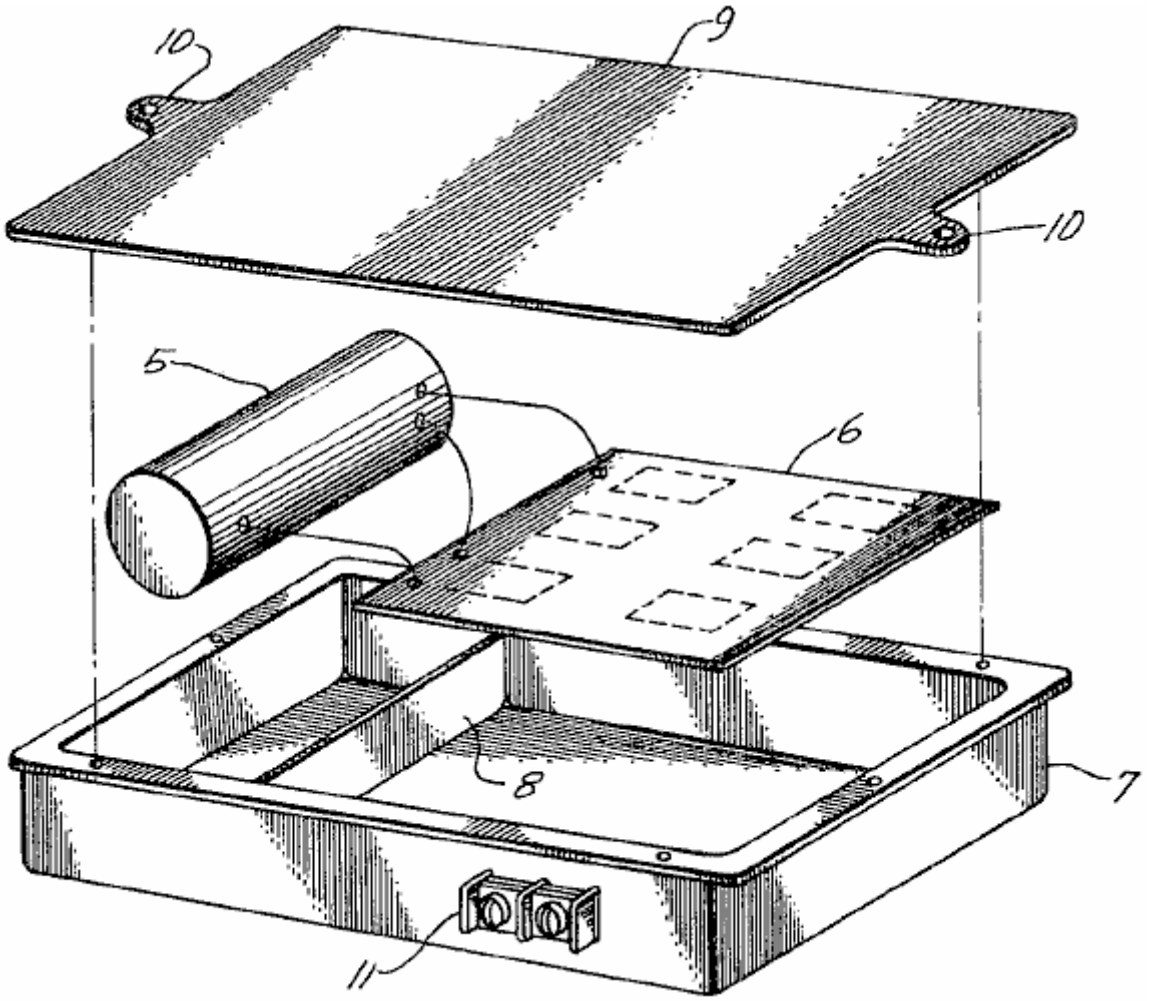


Fig.1 illustrates in exploded schematic, a flashing lamp connected to an electric power supply constructed according to the principles of this invention.

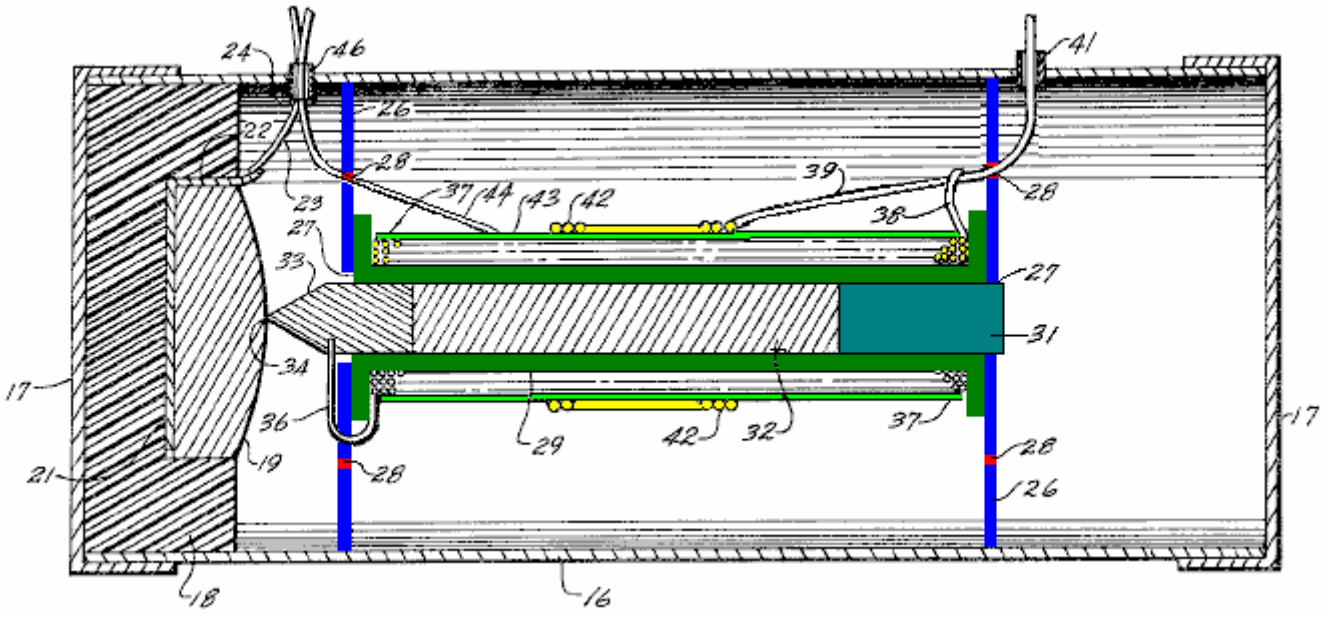


Fig.2 illustrates in longitudinal cross-section, the power pack of Fig.1

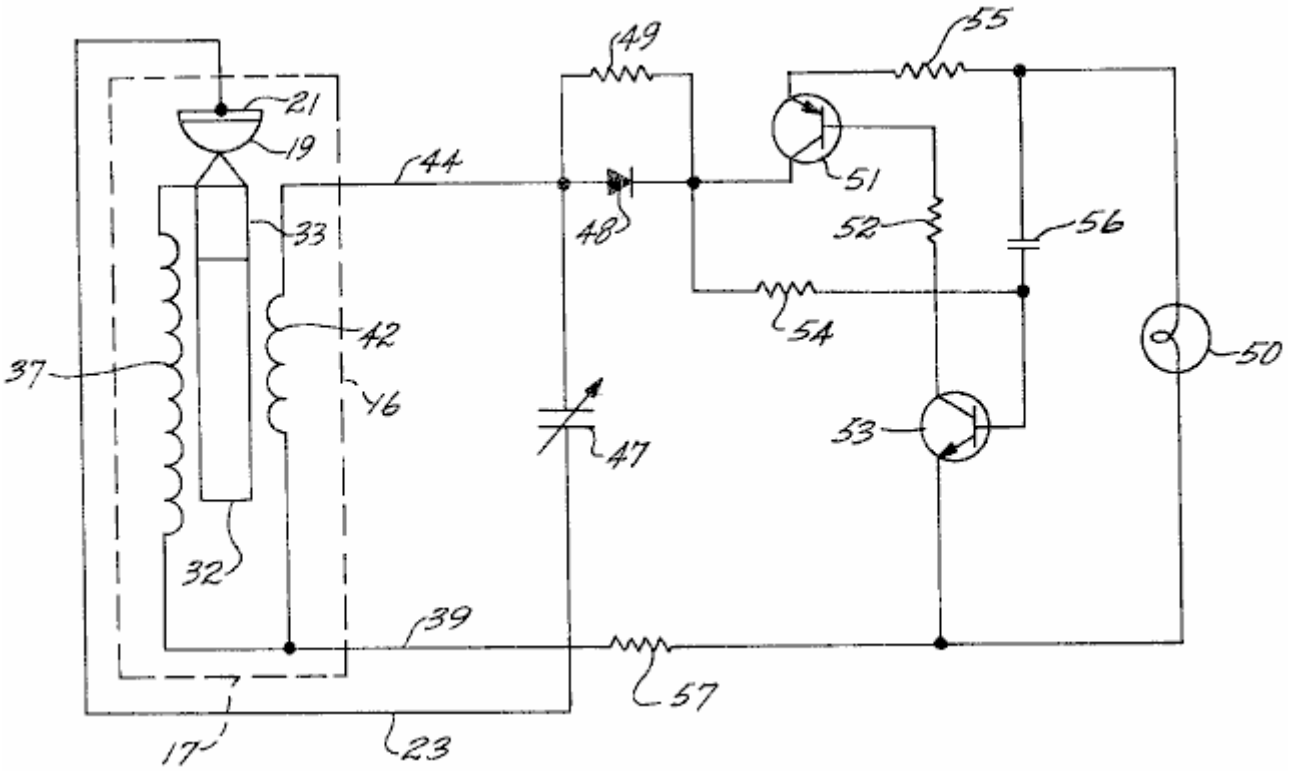


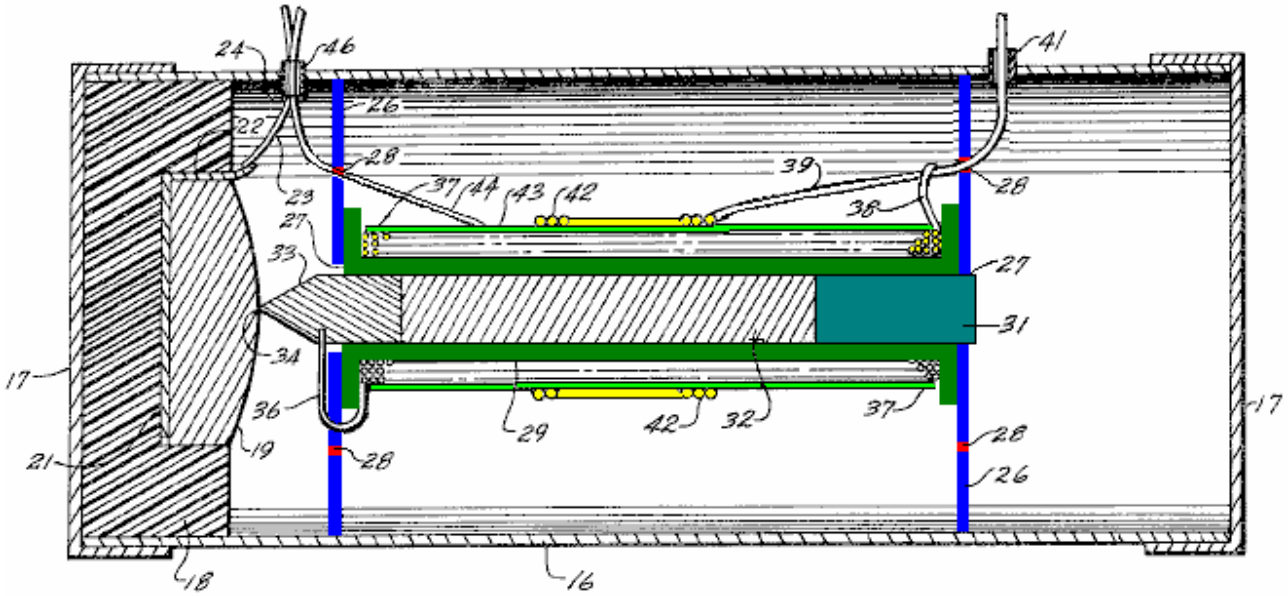
Fig.3 is an electric circuit diagram of the system.

DESCRIPTION

Fig.1 illustrates schematically, a typical flashing lamp having a power supply constructed according to the principles of this invention. As illustrated in this preferred embodiment, an electric power pack **5**, is connected electrically to a relaxation oscillator circuit (shown only schematically) on a conventional printed-circuit board **6**.

The power pack **5** and the printed-circuit board are mounted in a metal box **7**, which has a transverse partial partition **8**, which creates two spaces, one for the power pack and the other for the printed-circuit board which is prevented from contacting the metal box by any convenient insulating mounting. Preferably, these components are potted in place in a conventional manner.

A cover **9**, having mounting lugs **10**, is riveted on to the box after assembly. A small terminal strip **11**, mounted on one side of the box **7**, provides electrical contacts for connection to a load such as an incandescent lamp (not shown in Fig.1). the lamp provides a flash of light when the relaxation oscillator switches. Although the described system is employed for a flashing lamp, it will be apparent that other loads may be powered by the invention.



In **Fig.2**, the electric power pack **10**, is illustrated in longitudinal cross-section and has dimensions as follows: These dimensions are provided by way of example for powering a conventional flashing lamp and it will be clear that other dimensions may be used for other applications. In particular, the dimensions may be enlarged in order to obtain higher power levels and different voltage or current levels. The power pack is comprised of a cylindrical metal tube **16**, having closely fitting metal caps **17** at each end, which are preferably sealed to the tube after the internal elements are inserted in place. The metal tube **16** and caps **17**, which are preferably of aluminium, thus form a closed conductive envelope, which in a typical embodiment, has an inside diameter of about 0.8 inch and a length of about 2.25 inches.

Mounted within one end of the envelope is a plastic cup **18**, the dimensions of which are not critical, however, a wall thickness of at least 1/16 inch is preferred. Mounted within the plastic cup **18** is a semiconductor tablet **19** having a flat base and somewhat domed opposite side. The composition of the semiconductor tablet **19** is set out in greater detail below. Typically, the semiconductor tablet has a mass of about 3.8 grams. A metal disc **21** is positioned beneath the base of the tablet **19** in the cup **18**, and is preferably adhesively bonded inside the cup. The metal disc is tightly fitted to the base of the tablet so that good electrical contact is obtained over a substantial area of the semiconductor.

An ear **22** on one edge of the disc is soldered to a wire **23**, which extends through a short insulating sleeve **24** which passes through a hole in the side of the metal envelope. The insulating sleeve **24** acts as a grommet and ensures that there is no damage to the insulation of wire **23** and subsequent accidental short circuiting between the wire and the metal envelope. Preferably, the insulating sleeve **24** is sealed with a small amount of plastic cement or the like, in order to maintain clean air within the cylindrical envelope. Two other openings for leads through the tube **16**, as mentioned below, are also preferably sealed to maintain cleanliness within the envelope.

A pair of circular metal discs **26**, are fitted inside tube **16** and are preferably cemented in place to prevent shifting. The two discs **26**, are equally spaced from the opposite ends of the envelope and are spaced apart by slightly more than 1.15 inches. Each of the discs has a central aperture **27**, and there is a plurality of holes **28**, extending through the disc in a circular array midway between the centre of the disc and it's periphery. The holes **28** are preferably in the size range of about 0.01 to 0.06 inch in diameter and there are 12 on each disc located at 30^0 intervals around the circle.

The two discs **26** divide the interior of the cylindrical envelope into three chambers, and the pattern of holes **28** provides communication between the chambers and affects the electrical properties of the cavity. It is believed that the pattern of holes affects the inductive coupling between the cavities inside the envelope and influences the oscillations in them.

Although an arrangement of 12 holes at 30^0 centres has been found particularly advantageous in the illustrated embodiment, it is found in other arrangements that a pattern of 20 holes at 18^0 centres or a pattern of 8 holes at 45^0 centres, provides optimum operation. In either case, the circle of holes **28** is midway between the centre and the periphery of the disc.

Mounted between the discs **26** is a plastic spool **29** which has an inside distance of 1.1 inches between its flanges. The plastic spool **29** preferably has relatively thin walls and an internal bore diameter of 1/8 inch. A plastic mounting plug **31**, is inserted through the central aperture **27** of the disc **26** farthest from the semiconductor table **19**, and into the bore of the spool **29**. The plastic plug **31** is preferably cemented to the disc **26** in order to hold the assembly together.

Also mounted inside the bore of spool **29** is a cylindrical ferrite core **32**, about 1/8 inch diameter and 3/4 inch long. Although a core of any magnetic ferrite is preferred, other ferromagnetic materials having similar properties can be used if desired. The core **32**, is in electrical contact with a metal probe **33** about 1/4 inch long. half of the length of the probe **33** is in the form of a cylinder positioned within the spool **29**, and the other half is in the form of a cone ending in a point **34** in contact with the domed surface of the semiconductor tablet **19** where it makes an electrical contact with the semiconductor in a relatively small point.

Electrical contact is also made with the probe **33** by a lead **36**, which passes through one of the holes **28** in the disc **26** nearer to the semiconductor tablet and thence to a primary coil **37**, wound on the plastic spool **29**. The primary coil **37** is in the form of 800 to 1000 turns wound along the length of the spool, and the lead **38** at the opposite end of the coil **37** is soldered to one of the external leads **39** of the power pack. This lead **39** proceeds through one of the holes **28** in the disc farthest from the semiconductor tablet **19**, and through an insulating sleeve **41** in the metal tube **16**.

The lead **39** is also connected to one end of a secondary coil **42** which is composed of 8 to 10 turns around the centre portion of the primary coil **37**. A thin insulating sheet **43** is provided between the primary and secondary coils. The other lead **44** from the secondary coil passes through one of the holes **28** in the disk nearer the semiconductor tablet and thence through an insulating sleeve **46** through the wall of the tube **16**.

Fig.3 illustrates schematically, the electrical circuit employing an electric power pack constructed according to the principles of this invention. At the left hand side of **Fig.3**, the arrangement of elements is illustrated in a combination of electrical schematic and mechanical position inside tube **16** for ready correlation with the embodiment illustrated in **Fig.2**. Thus, the semiconductor tablet **19**, probe **33** and ferrite core **32** are shown in both their mechanical and electrical arrangement, the core being inductively coupled to the coils **37** and **42**. The lead **23** from the metal base of the semiconductor tablet **19**, is connected to a variable capacitor **47**, the other side of which is connected to the lead **44** from the secondary coil **42**. The lead **44** is also connected to a rectifying diode **48** shunted by a high value resistor **49**.

It will be seen that the variable capacitor **47** is in a tank circuit with the inductive coils **37** and **42** which are coupled by the ferrite core **32**, and this circuit also includes the semiconductor tablet **19** to which point contact is made by the probe **33**. The mechanical and electrical arrangement of these elements provides a resonant cavity in which resonance occurs when the capacitor 47 is properly trimmed. The diode **48**, rectifies the oscillations in this circuit to provide a suitable DC for operating an incandescent lamp **50** or similar load.

The rectifying diode **48** is connected to a complementary-symmetry relaxation circuit for switching power to the load **50**. The diode is connected directly to the collector of a PNP transistor **51** which is in an inverted connection. the emitter of the PNP transistor is connected to one side of the load **50** by way of a timing resistor **55**. The base of the transistor **51** is connected by way of a resistor **52** and a capacitor **56** to the collector of an NPN transistor **53**, the emitter of which is connected to the other side of the load **50**. The base of the NPN transistor **53** is coupled to the diode by a resistor **54**. The emitter of the PNP transistor **51** is fed back to the base of the NPN transistor **53** by the resistor **55**. Current flow through the lamp **50** is also limited by a resistor **57** which couples one side of the lamp and the emitter of the NPN transistor **53** to the two coils **37** and **42** by way of the common lead **39**.

The electrical power pack is believed to operate due to a resonance amplification once an oscillation has been initiated in the cavity, particularly the central cavity between the discs **26**. This oscillation, which apparently rapidly reaches amplitudes sufficient for useful power, is then half-wave rectified for use by the diode **48**. With such an arrangement, a voltage level of several volts has been obtained, and power sufficient for intermittent operation of a lamp requiring about 170 to 250 milliwatts has been demonstrated. The resonant amplification is apparently due to the geometrical and electrical combination of the elements, which provide inductive coupling of components in a suitable resonant circuit. This amplification is also, at least in part, due to unique semiconductor properties in the tablet **19**, which has electronic properties due to a composition giving a unique atomic arrangement, the exact nature of which has not been measured.

The semiconductor tablet has electronic properties which are determined by it's composition and three such semiconductors satisfactory for use in the combination have been identified. In two of these, the base semiconductor material is selenium provided with suitable dopant elements, and in the third, the base element is germanium, also suitably doped. The semiconductor tablets are made by melting and casting in an arrangement which gives a large crystal structure. It has not been found necessary to provide a selected crystal orientation in order to obtain the desired effects.

A preferred composition of the semiconductor includes about 5% by weight of tellurium, about 4% by weight of germanium, about 3% by weight of neodymium and about 4.7% by weight of rubidium, with the balance of the composition being selenium. Such a composition can be made by melting these materials together or by dissolving the materials in molten selenium.

Another highly advantageous composition has about 5% by weight of tellurium, about 4% by weight of germanium, about 3% by weight of neodymium, and about 2.24% by weight of gallium, with the balance being selenium. In order to make this composition, it is found desirable to add the very low melting point gallium in the form of gallium selenide rather than elemental gallium.

A third suitable composition has about 5% by weight of tellurium, about 4% by weight of neodymium, about 6% by weight of rubidium, with the balance being germanium. These preferred compositions are not absolute and it has been found that the level of dopant in the compositions can be varied within limits without significant loss of performance. Thus, it is found that the proportion of tellurium in the preferred composition can range from about 4.8% to about 5.5% by weight; the germanium can range from about 3.9% to 4.5% by weight; neodymium can range from about 2.9% to 3.5% by weight, and rubidium can vary from about 4.5% to 5.0% by weight. The

balance of the preferred composition is selenium although it has also been found that nominal impurity levels can be tolerated and no great care is required in preventing minor contamination.

The other selenium base composition useful in practice of this invention can have a tellurium concentration in the range of from about 4.85% to 5.5% by weight, germanium in the range of from about 3.95% to 4.2% by weight, neodymium in the range of from about 2.85% to 3.2% by weight, and gallium in the range of from about 2.0% to 2.5% by weight. As in the preferred composition, the balance is selenium and nominal impurity levels can be tolerated. It is preferred to add the gallium in the form of gallium selenide rather than as elemental gallium with a corresponding decrease in the selenium used to make up the composition.

The above selenium base compositions are easier to make and less expensive than the germanium base composition and are therefore preferable for most applications. It is found that these are particularly suited for relatively small semiconductor tablets up to about 1 inch or a little less. For relatively large tablets, it is preferred to use the germanium base composition.

The germanium base composition has a tellurium level in the range of from about 4.75% to 5.5% by weight, neodymium in the range of from about 4.0% to 4.5% by weight, and rubidium in the range of from about 5.5% to 7.4% by weight. It is also found that it is of greater importance to maintain purity of the germanium base compositions than the selenium base compositions. Although the exact purity levels have not been ascertained, it is in excess of 99%.

It has been found that it is not necessary to have single crystals in the semiconductor tablets and any convenient grain size in excess of about 1 millimetre appears satisfactory. In the above compositions, when the recited ranges are exceeded, oscillation in the power pack drops off rapidly and may cease altogether.

The reasons that these compositions are satisfactory in the arrangement providing resonance amplification has not been determined with certainty. It is possible that the semiconductor serves as a source of electrons for providing an oscillating current in the circuit. This is, of course, combined with a relatively large area contact to one side of the semiconductor tablet, and a point contact on another area. Any resonant current in the coils wound on the ferrite rod, induces a varying magnetic field in the resonant cavity, and the electrical connection between the ferrite rod and the metal probe, provides a feedback of this oscillation to the semiconductor tablet.

it should particularly be noted that the oscillation in the circuit does not commence until it is initiated by an oscillating signal. In order to accomplish this, it is only necessary to apply a few millivolts of AC for a few seconds to the semiconductor tablet and the associated coils coupled to it. The initial signal applied to the base of the semiconductor tablet and the lead **39** is preferably in the frequency range of 5.8 to 18 Mhz and can be as high as 150 Mhz. Such a signal can be applied from any conventional source and no great care appears necessary to provide a single frequency signal or to eliminate noise. Once such energisation has been applied to the circuit and oscillations initiated, it does not appear to be necessary to apply such a signal again. This is apparently due to the feedback provided by the ferrite rod to the probe which makes contact with the semiconductor tablet.

Energy is, of course, dissipated in the lamp, or other utilisation device, as the combination operates. Such energy may come from deterioration of the semiconductor tablet as oscillations continue; however, if there is any such deterioration, it is sufficiently slow that a power source may be operated for many months without attendance. Such a source of energy may be augmented by ambient Radio Frequency radiation, coupled into the resonant cavity by the external leads. This is a surprising phenomenon because the leads are small compared to what would normally be considered an adequate antenna, and it is therefore postulated that stimulated amplification may also be a consequence of the unique electronic configuration of the semiconductors having the compositions specified above.

Although only one embodiment of electric power pack constructed according to principles of this invention has been described and illustrated here, many modifications and variations will be apparent to one skilled in the art. Thus, for example, a larger power pack may be axially arranged in a cylindrical container with various electronic elements arranged in the annular space. It is therefore to be understood that other configurations are included within the scope of the invention.

EDWIN GRAY: ELECTRIC MOTOR

US Patent 3,890,548 June 17, 1975 Inventor: Edwin V. Gray snr.

PULSED CAPACITOR DISCHARGE ELECTRIC ENGINE

Please note that this is a re-worded extract from Edwin Gray's Patent 3,890,548. It describes his high voltage motor and the circuitry used to drive it. Please be aware that the underlying technology was developed by Marvin Cole and Edwin Gray did not understand it. Also, Edwin wanted at all costs to conceal any useful technology while getting patents to encourage investors, so please understand that this patent is not intended to tell you how to make a working system of this type.

SUMMARY OF THE INVENTION:

This invention relates to electric motors or engines, and more particularly to a new electric machine including electromagnetic poles in a stator configuration and electromagnetic poles in a rotor configuration, wherein in one form thereof, the rotor is rotatable within the stator configuration and where both are energised by capacitor discharges through rotor and stator electromagnets at the instant of the alignment of a rotor electromagnet with a stator electromagnet. The rotor electromagnet is repelled from the stator electromagnet by the discharge of the capacitor through the coils of both the rotor and stator electromagnets at the same instant.

In an exemplary rotary engine according to this invention, rotor electromagnets may be disposed 120 degrees apart on a central shaft and major stator electromagnets may be disposed 40 degrees apart in the motor housing about the stator periphery. Other combinations of rotor elements and stator elements may be utilised to increase torque or rate of rotation.

In another form, a second electromagnet is positioned to one side of each of the major stator electromagnets on a centreline 13.5 degrees from the centreline of the stator magnet, and these are excited in a predetermined pattern or sequence. Similarly, to one side of each rotor electromagnet, is a second electromagnet spaced on a 13.5 degree centreline from the major rotor electromagnet. Electromagnets in both the rotor and stator assemblies are identical, the individual electromagnets of each being aligned axially and the coils of each being wired so that each rotor electromagnetic pole will have the same magnetic polarity as the electromagnet in the stator with which it is aligned and which it is confronting at the time of discharge of the capacitor.

Charging of the discharge capacitor or capacitors is accomplished by an electrical switching circuit wherein electrical energy from a battery or other source of d-c potential is derived through rectification by diodes.

The capacitor charging circuit comprises a pair of high frequency switchers which feed respective automotive-type ignition coils employed as step-up transformers. The "secondary" of each of the ignition coils provides a high voltage square wave to a half-wave rectifier to generate a high voltage output pulse of d-c energy with each switching alternation of the high frequency switcher. Only one polarity is used so that a unidirectional pulse is applied to the capacitor bank being charged.

Successive unidirectional pulses are accumulated on the capacitor or capacitor bank until discharged. Discharge of the bank of capacitors occurs across a spark gap by arc-over. The gap spacing determines the voltage at which discharge or arc-over occurs. An array of gaps is created by fixed elements in the engine housing and moving elements positioned on the rotor shaft. At the instant when the moving gap elements are positioned

opposite fixed elements during the rotor rotation, a discharge occurs through the coils of the aligned rotor and stator electromagnets to produce the repulsion action between the stator and rotor electromagnet cores.

A plurality of fixed gap elements are arrayed in a motor housing to correspond to the locations of the stator electromagnets in the housing. The rotor gap elements correspond to the positions of the rotor electromagnets on the rotor so that at the instant of correct alignment of the gaps, the capacitors are discharged to produce the necessary current through the stator and rotor coils to cause the electromagnets to repel one another.

The charging circuits are arranged in pairs, and are such that the discharge occurs through both rotor and stator windings of the electromagnets, which are opposite one another when the spark gap elements are aligned and arc-over.

The speed of the rotor can be changed by means of a clutch mechanism associated with the rotor. The clutch shifts the position of the rotor gap elements so that the discharge will energise the stator coils in a manner to advance or retard the time of discharge with respect to the normal rotor/stator alignment positions. The discharge through the rotor and stator then occurs when the rotor has passed the stator by 6.66 degrees for speed advance.

By causing the discharge to occur when the rotor position is approaching the stator, the repulsion pulse occurs 6.66 degrees before the alignment position of the rotor and stator electromagnets, thus reducing the engine speed.

The clutch mechanism for aligning capacitor discharge gaps for discharge is described as a control head. It may be likened to a firing control mechanism in an internal combustion engine in that it “fires” the electromagnets and provides a return of any discharge overshoot potential back to the battery or other energy source.

The action of the control head is extremely fast. From the foregoing description, it can be anticipated that an increase in speed or a decrease in speed of rotation can occur within the period in which the rotor electromagnet moves between any pair of adjacent electromagnets in the stator assembly. These are 40 degrees apart so speed changes can be effected in a maximum of one-ninth of a revolution.

The rotor speed-changing action of the control head and its structure are believed to be further novel features of the invention, in that they maintain normal 120 degree firing positions during uniform speed of rotation conditions, but shift to 6.66 degree longer or shorter intervals for speed change by the novel shift mechanism in the rotor clutch assembly.

Accordingly, the preferred embodiment of this invention is an electric rotary engine wherein motor torque is developed by discharge of high potential from a bank of capacitors, through stator and rotor electromagnet coils when the electromagnets are in alignment. The capacitors are charged from batteries by a switching mechanism, and are discharged across spark gaps set to achieve the discharge of the capacitor charge voltage through the electromagnet coils when the gaps and predetermined rotor and stator electromagnet pairs are in alignment.

Exemplary embodiments of the invention are herein illustrated and described. These exemplary illustrations and description should not be construed as limiting the invention to the embodiments shown, because those skilled in the arts appertaining to the invention may conceive of other embodiments in the light of the description within the ambit of the appended claims.

BRIEF DESCRIPTION OF THE DRAWINGS:

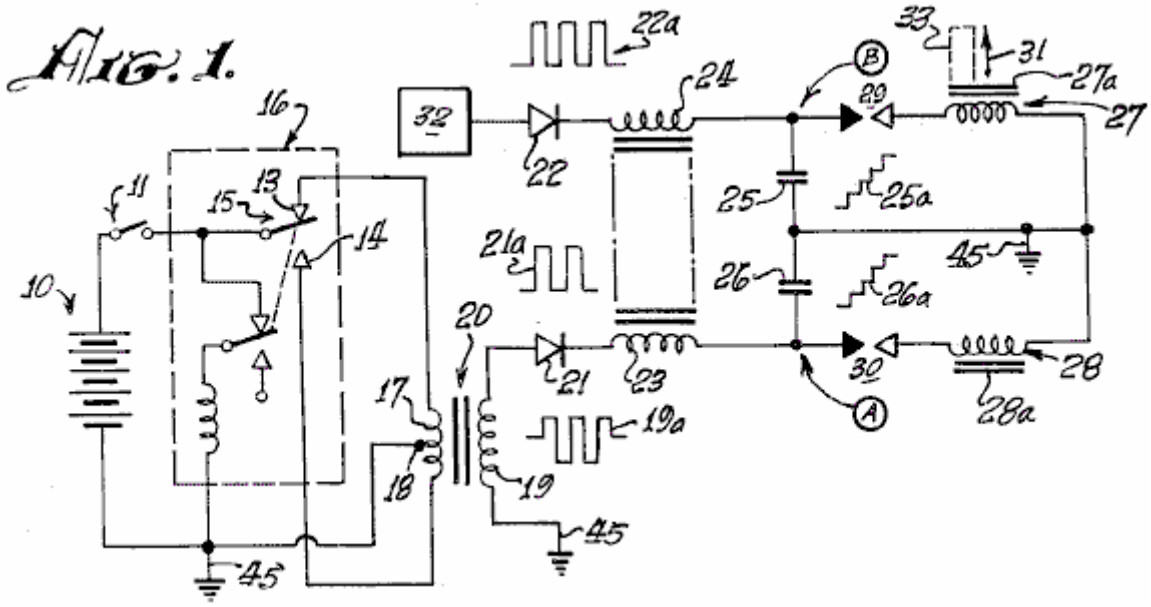


Fig.1 is an explanatory schematic diagram of a capacitor charging and discharging circuit utilised in the present invention.

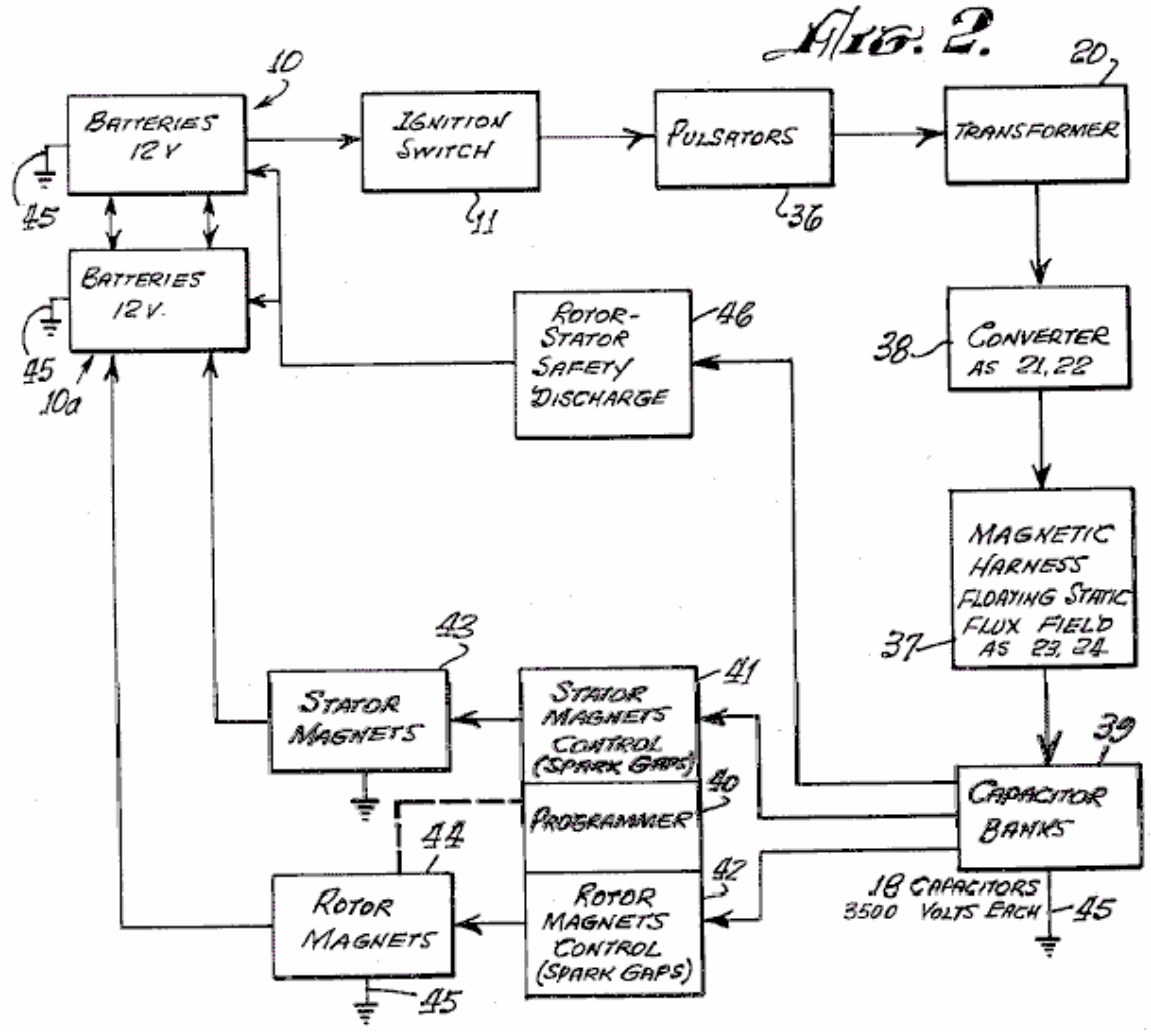


Fig.2 is a block diagram of an exemplary engine system according to the invention.

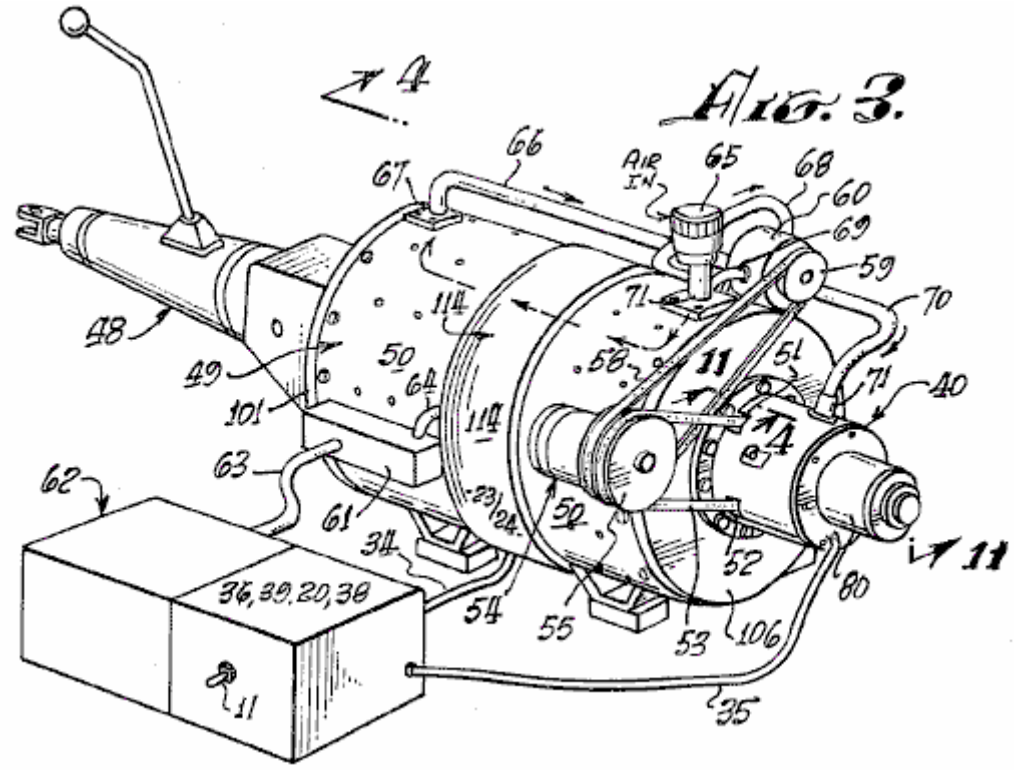


Fig.3 is a perspective view of a typical engine system according to the invention, coupled to an automotive transmission.

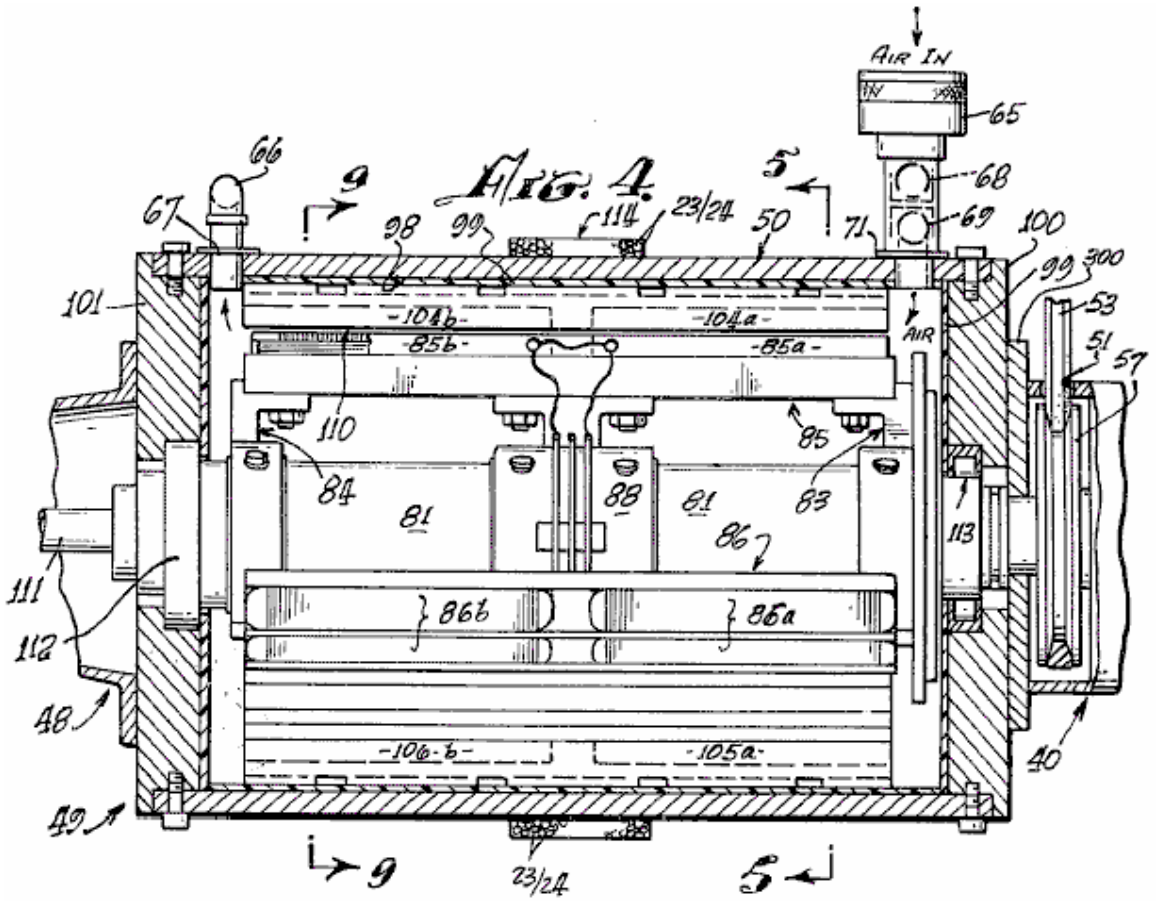


Fig.4 is an axial sectional view taken at line 4---4 in Fig.3

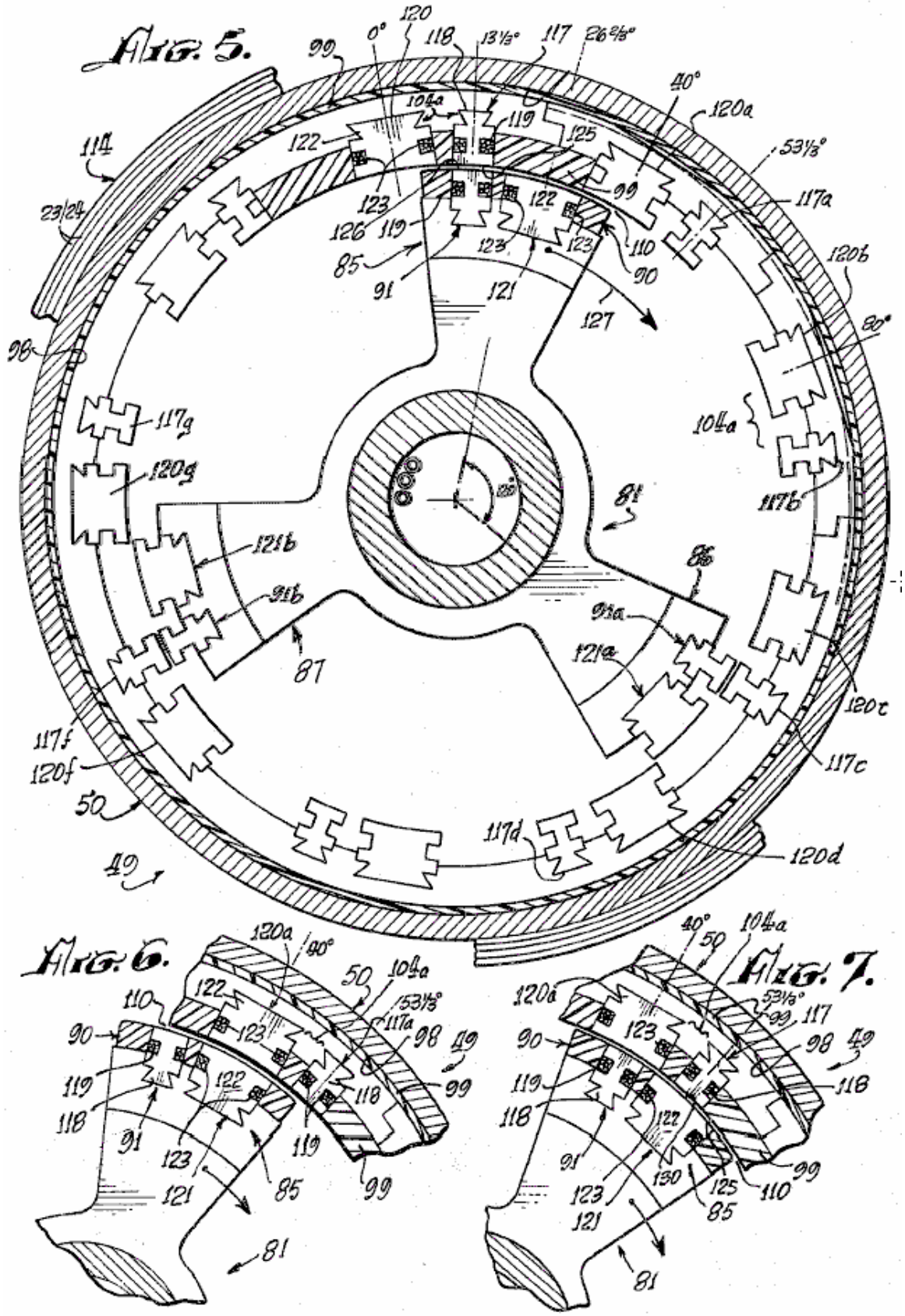


Fig.5 is a sectional view taken at line 5---5 in Fig.4

Fig.6 and **Fig.7** are fragmentary sectional views, corresponding to a portion of Fig.5, illustrating successive advanced positions of the engine rotor therein.

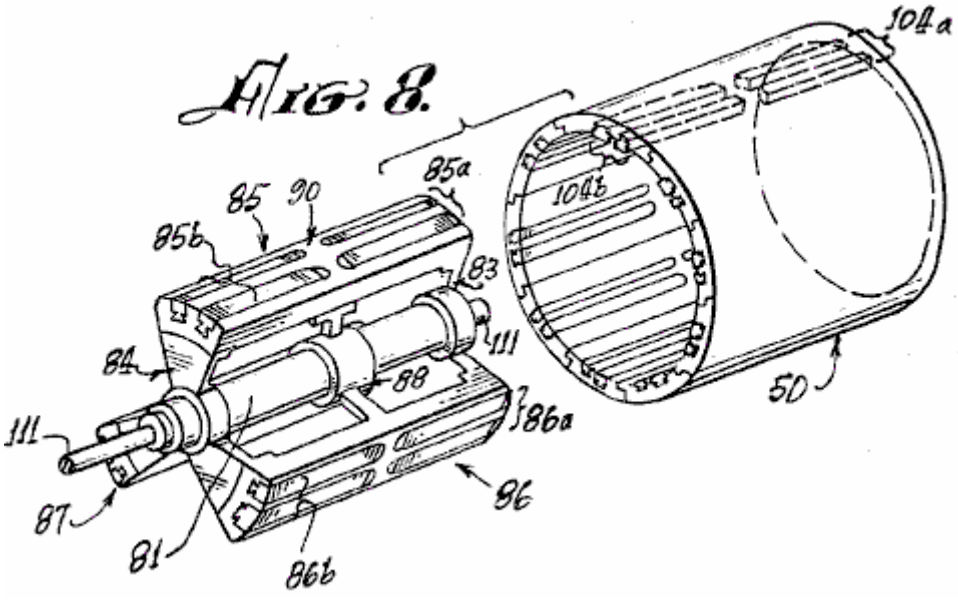


Fig.8 is an exploded perspective view of the rotor and stator of the engine of Fig.3 and Fig.4

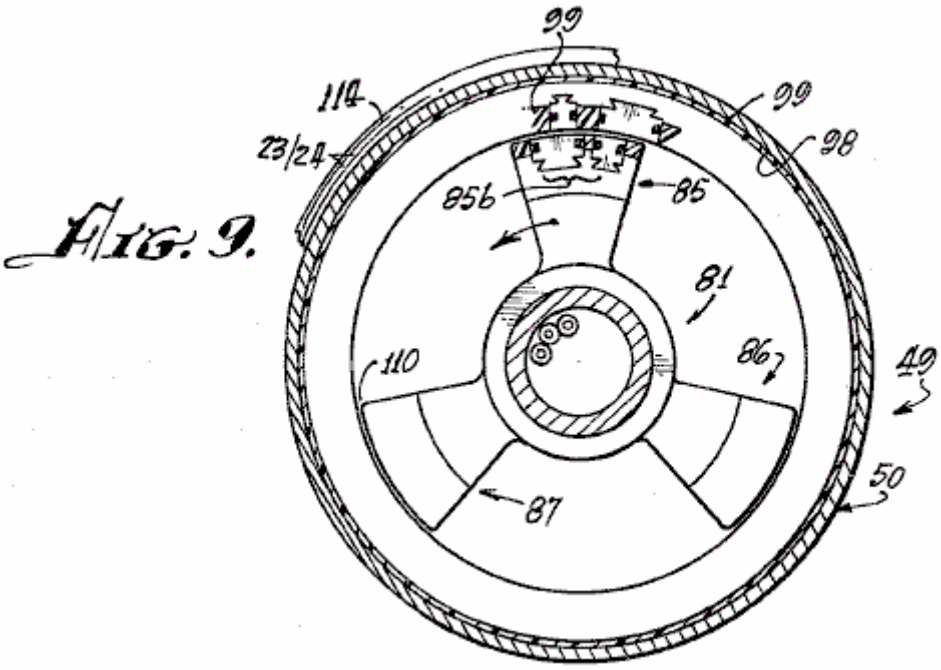


Fig.9 is a cross-sectional view taken at line 9---9 of Fig.4

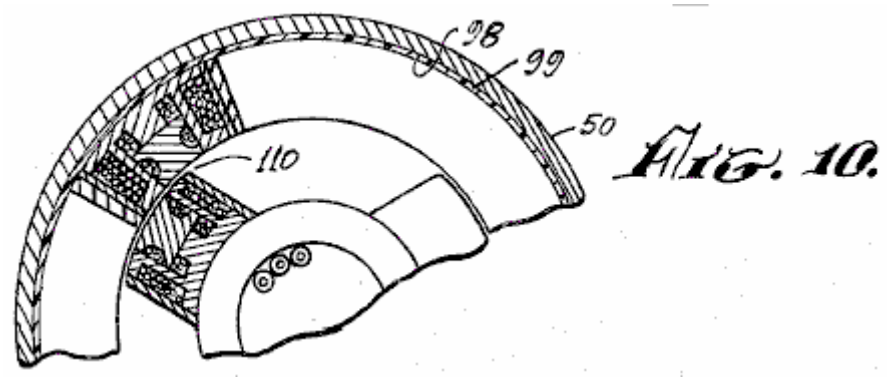


Fig.10 is a partial sectional view, similar to the view of Fig.9, illustrating a different configuration of electromagnets in another engine embodiment of the invention.

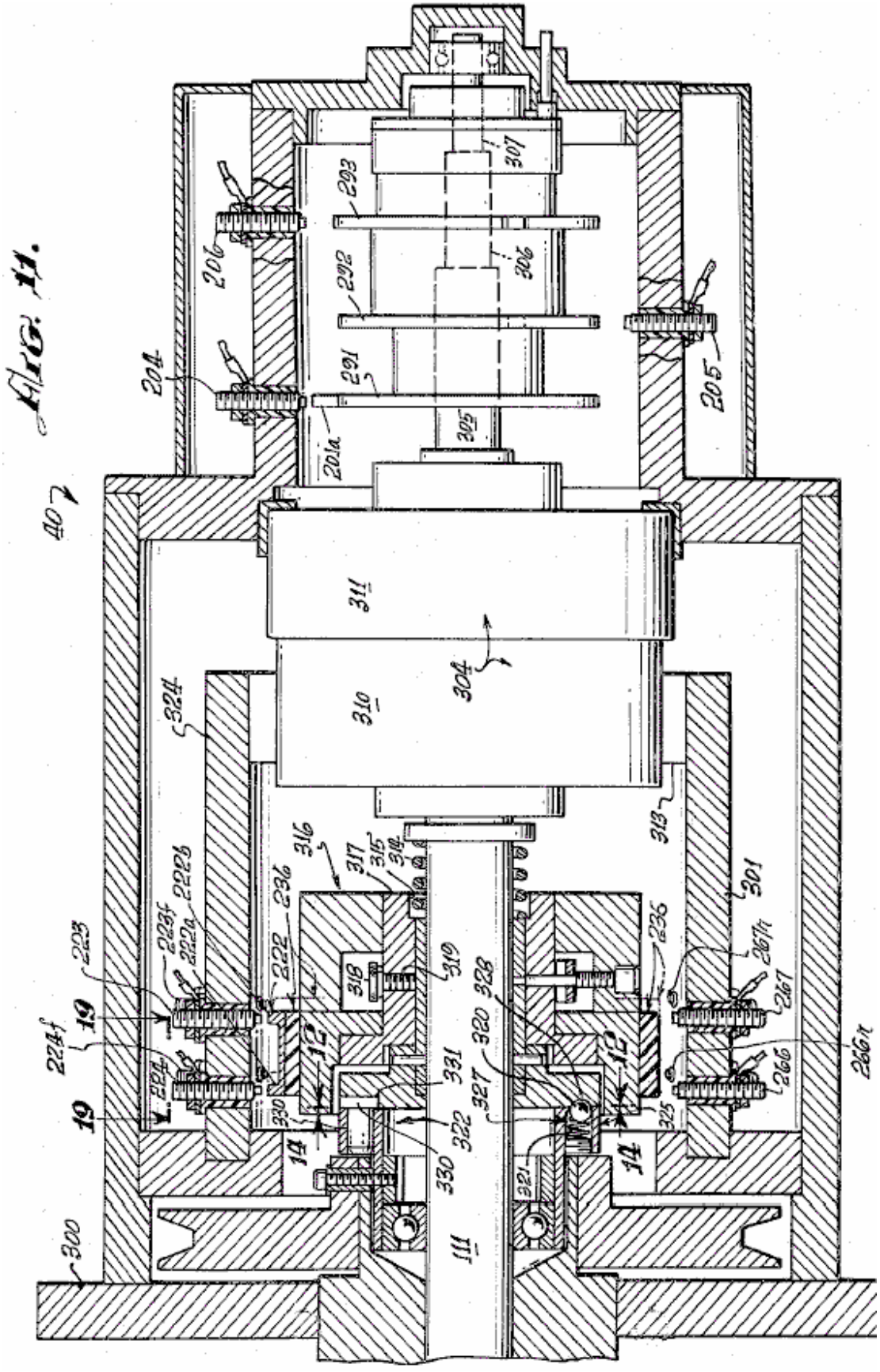


Fig.11 is a sectional view taken at line 11---11 in Fig.3, illustrating the control head or novel speed change controlling system of the engine.

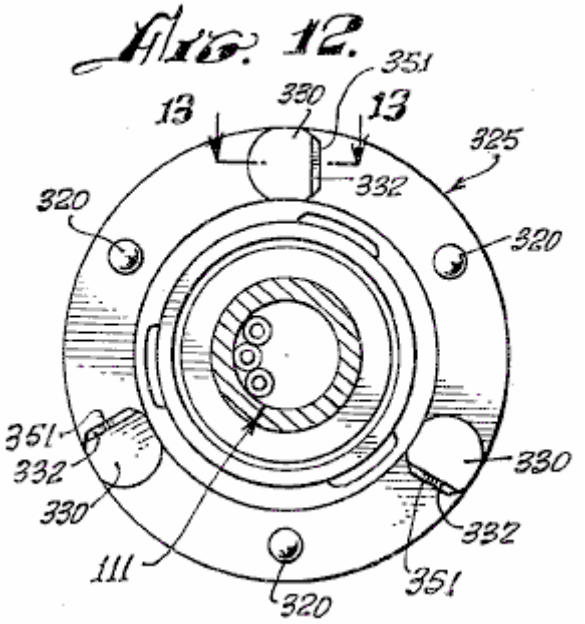


Fig.12 is a sectional view, taken at line 12---12 in Fig.11, showing a clutch plate utilised in the speed change control system of Fig.11

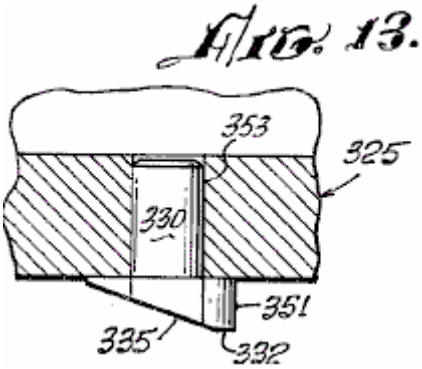


Fig.13 is a fragmentary view, taken at line 13---13 in Fig.12

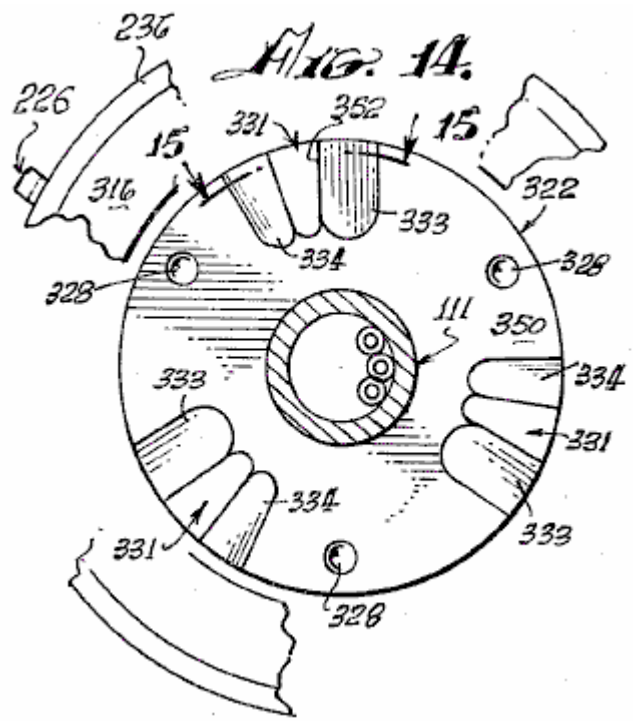


Fig.14 is a sectional view, taken at line 14---14 in Fig.11, showing a clutch plate which co-operates with the clutch plate of Fig.12

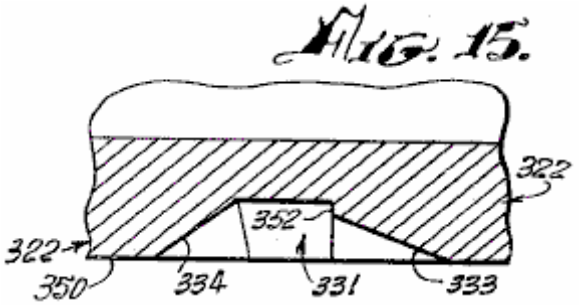


Fig.15 is a fragmentary sectional view taken at line 15---15 of Fig.13

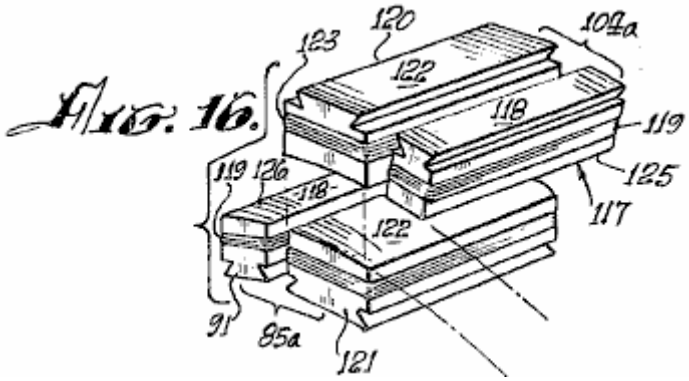


Fig.16 is a perspective view of electromagnets utilised in the present invention.

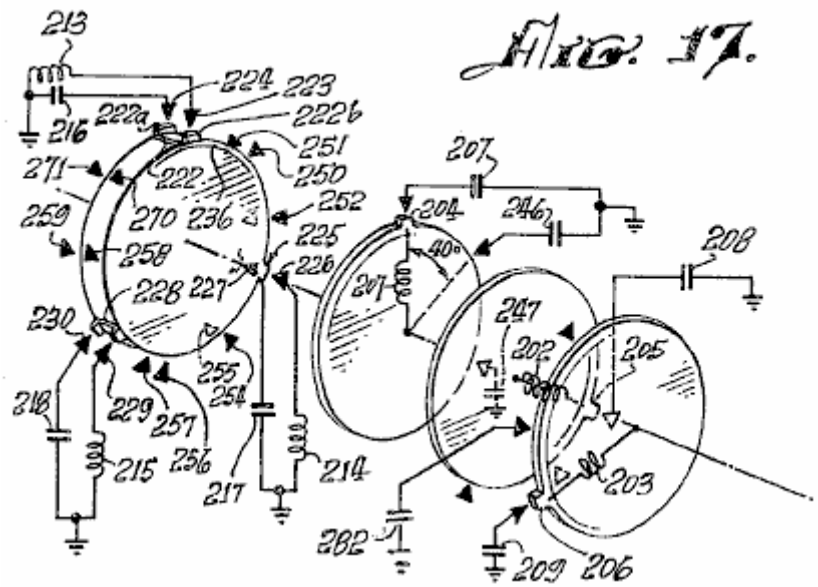


Fig.17 is a schematic diagram showing co-operating mechanical and electrical features of the programmer portion of the invention.

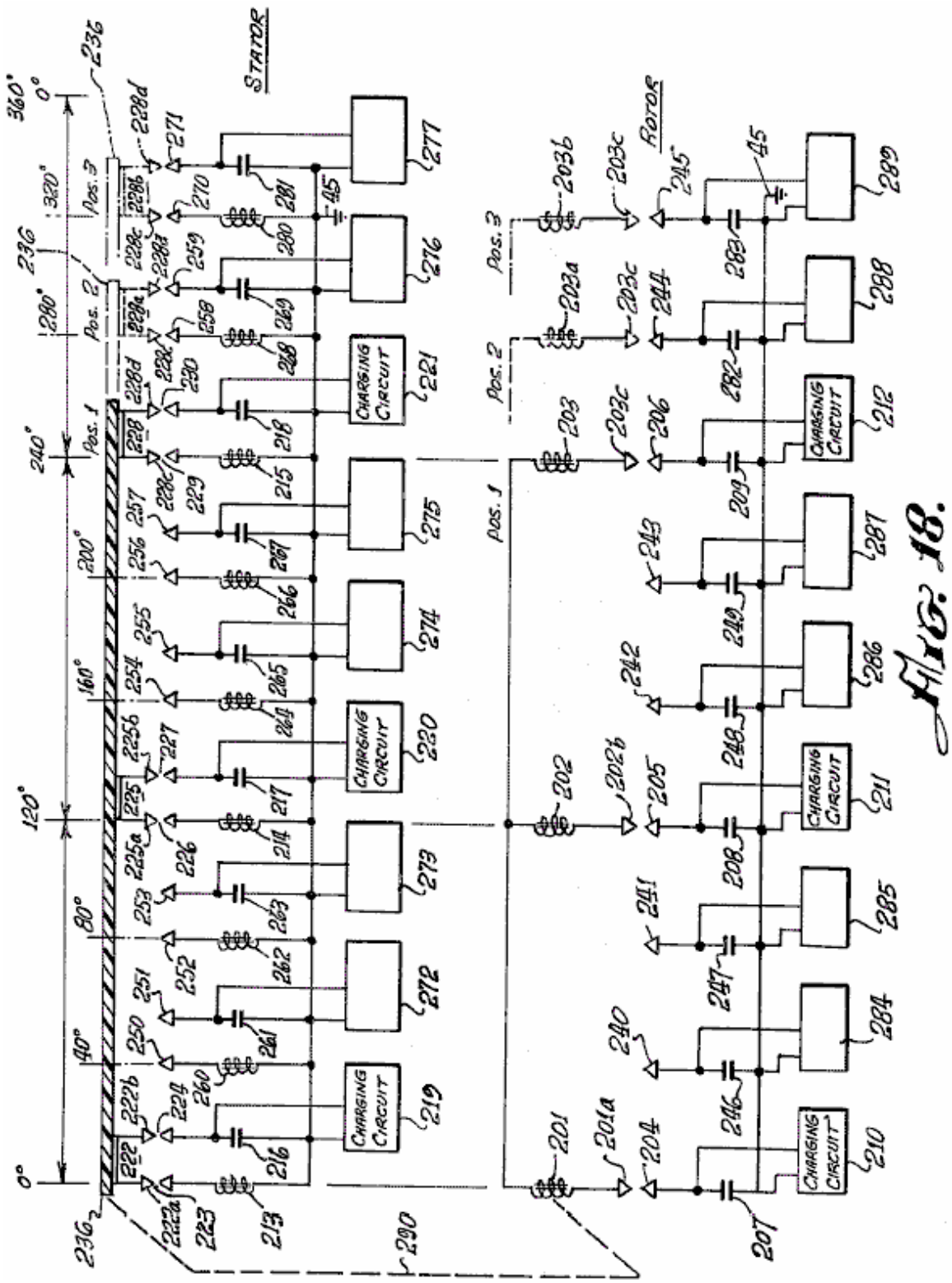


Fig.18 is an electrical schematic diagram of an engine according to the invention, showing the electrical relationships of the electromagnetic components embodying a new principle of the invention, and

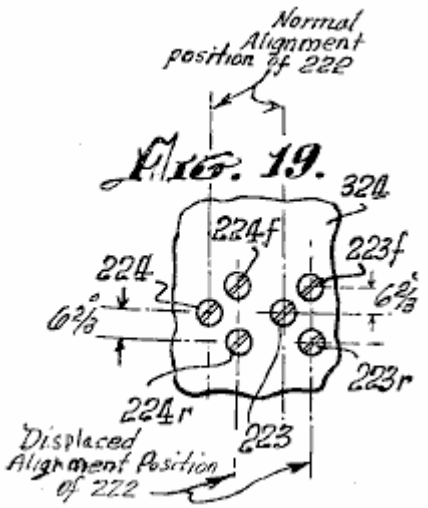
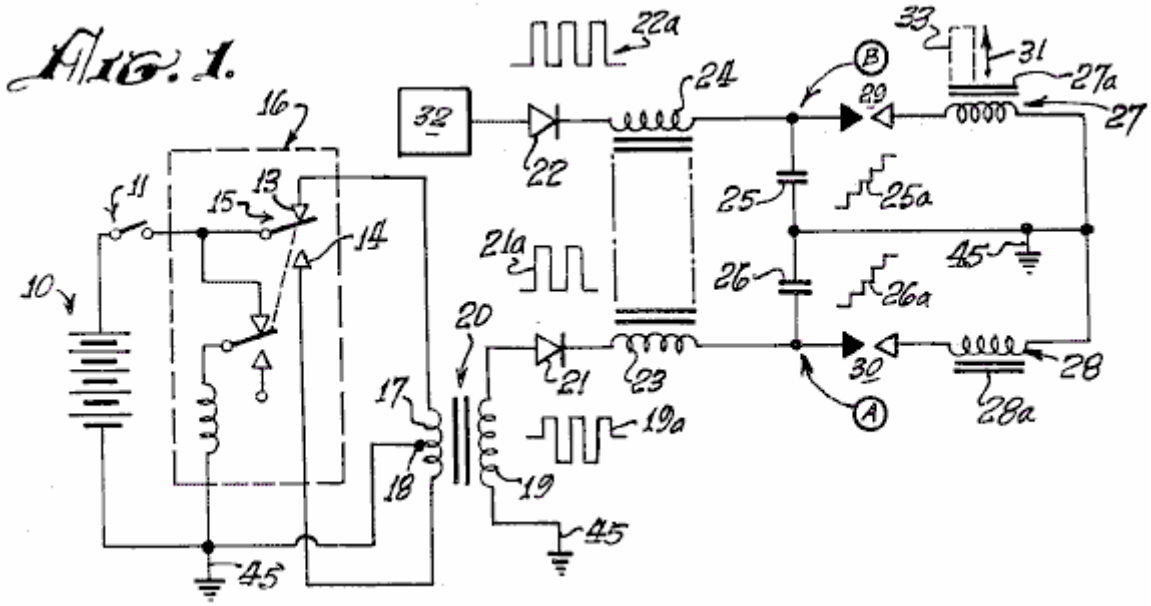


Fig.19 is a developed view, taken at line 19---19 of Fig.11, showing the locations of displaced spark gap elements of the speed changing mechanism of an engine according to the invention.

DESCRIPTION OF THE PREFERRED EMBODIMENT

As mentioned earlier, the basic principle of operation of the engine of the invention, is the discharge of a capacitor across a spark gap and through an inductor. When a pair of inductors is used, and the respective magnetic cores thereof are arranged opposite one another and arranged in opposing magnetic polarity, the discharge through them causes the cores to repel each other with considerable force.



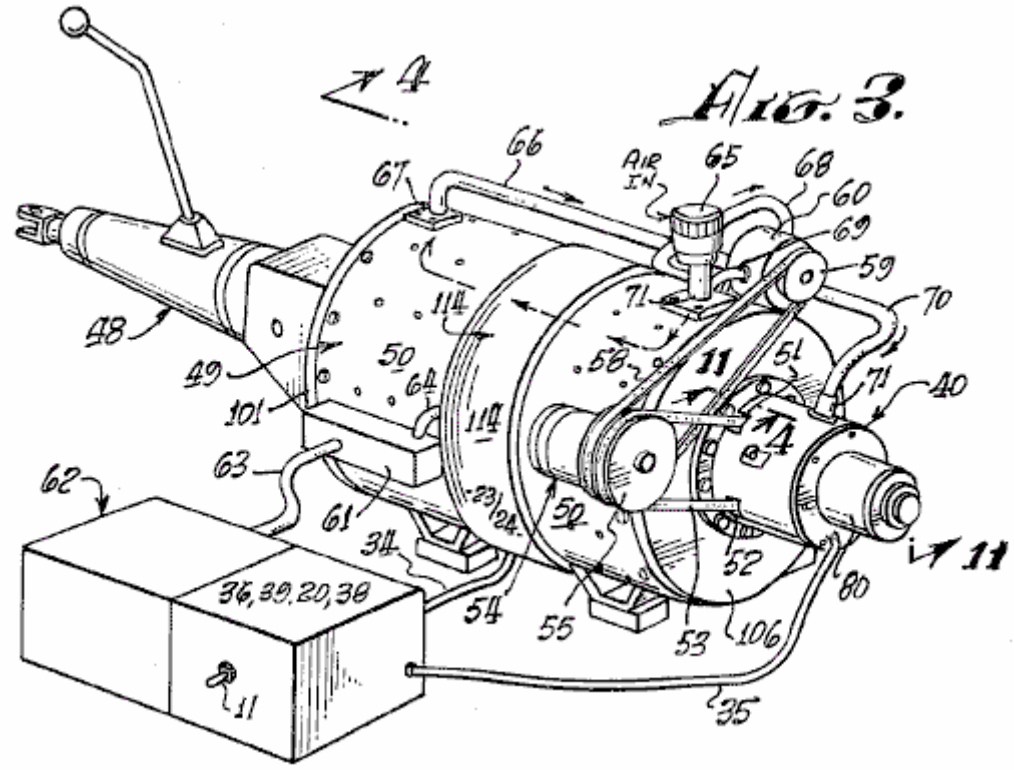
Referring to the electrical schematic diagram of **Fig.1**, a battery **10** energises a pulse-producing vibrator mechanism **16**, which may be of the magnetic type, incorporating an armature **15** moving between contacts **13** and **14**, or of the transistor type (not shown) with which a high frequency bipolar pulsed output is produced in primary **17** of transformer **20**. The pulse amplitude is stepped up in secondary **19** of transformer **20**. Wave form **19a** represents the bi-directional or bi-polar pulsed output. A diode rectifier **21** produces a unidirectional pulse train, as indicated at **21a**, to charge capacitor **26**. Successive unidirectional pulses of wave **21a** charge capacitor **26** to high level, as indicated at **26a**, until the voltage at point **A** rises high enough to cause a spark across the spark gap **30**. Capacitor **26** discharges via the spark gap, through the electromagnet coil **28**. A current pulse is produced which magnetises core **28a**. Simultaneously, another substantially identical charging system **32** produces a discharge through inductor **27** across spark gap **29**, to magnetise core **27a**. Cores **27a** and **28a** are wound with coils **27** and **28** respectively, so that their magnetic polarities are the same. As the cores **27a** and **28a** confront one another, they tend to fly apart when the discharge occurs through coils **27** and **28** because of repulsion of identical magnetic poles, as indicated by arrow **31**. If core **28a** is fixed or stationary, and core **27a** is moveable, then core **27a** may have tools **33** attached to it to perform work when the capacitor discharges.

Referring to **Fig.1** and **Fig.2**, a d-c electrical source or battery **10**, energises pulsators **36** (including at least two vibrators **16** as previously described) when switch 11 between the battery **10** and pulsator **36** is closed, to apply

relatively high frequency pulses to the primaries of transformers **20**. The secondaries of transformers **20** are step-up windings which apply bipolar pulses, such as pulses **19a** (Fig.1) to the diodes in converter **38**. The rectified unidirectional pulsating output of each of the diodes in converter **38** is passed through delay coils **23** and **24**, thus forming a harness **37**, wound about the case of the engine, as herein after described, which is believed to provide a static floating flux field. The outputs from delay lines **37**, drive respective capacitors in banks **39**, to charge the capacitors therein, to a relatively high charge potential. A programmer and rotor and stator magnet control array **40**, **41**, **42**, is formed by spark gaps positioned, as hereinafter described, so that at predetermined positions of the rotor during rotation of the engine, as hereinafter described, selected capacitors of the capacitor banks **39** will discharge across the spark gaps through the rotor and stator electromagnets **43** and **44**. The converters **38**, programmer **40**, and controls **41** and **42**, form a series circuit path across the secondaries of transformers **20** to the ground, or point of reference potential, **45**. The capacitor banks **39** are discharged across the spark gaps of programmer **40** (the rotor and stator magnet controls **41** and **42**). The discharge occurs through the coils of stator and rotor electromagnets **43** and **44** to ground **45**. Stator and rotor electromagnets are similar to those shown at **27**, **27a**, **28** and **28a** in **Fig.1**.

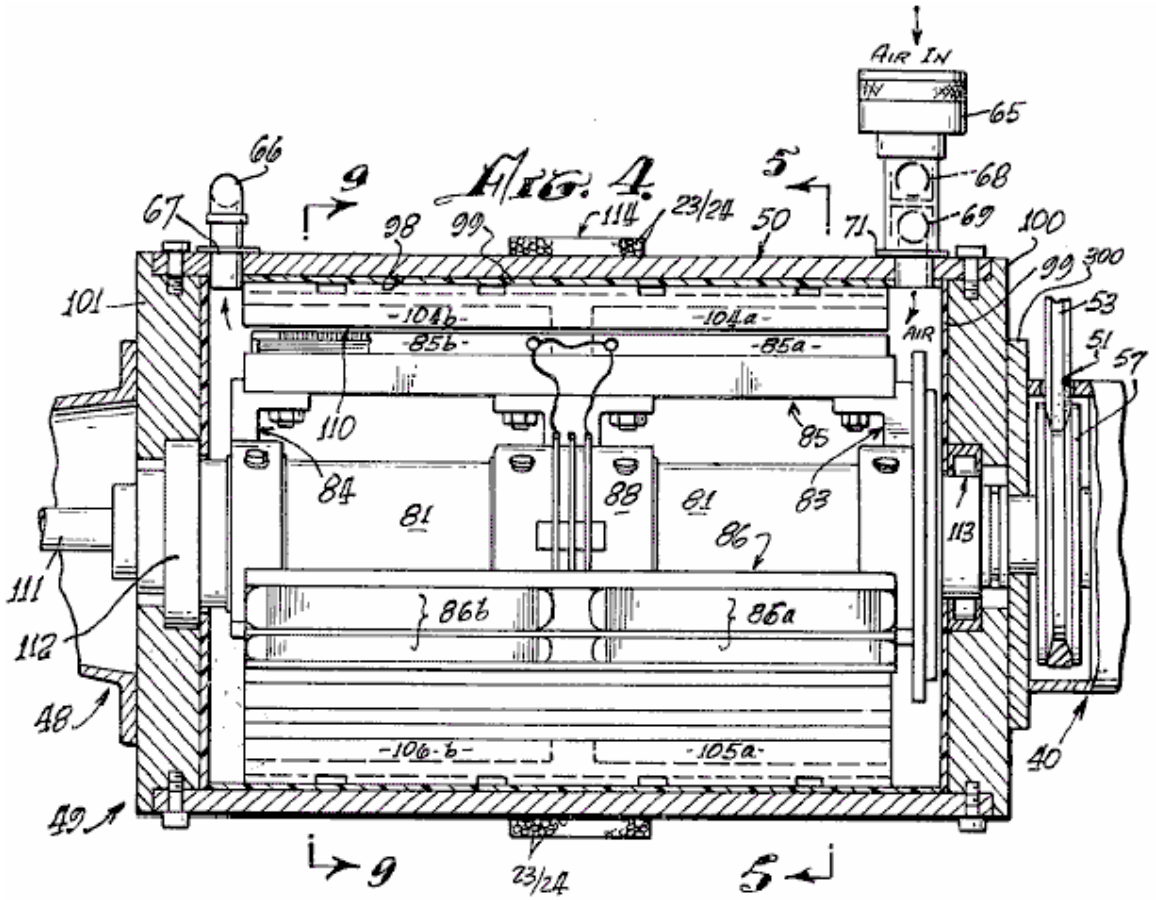
The discharge through the coils of stator and rotor electromagnets **43** and **44** is accompanied by a discharge overshoot or return pulse, which is applied to a secondary battery **10a** to store this excess energy. The overshoot pulse returns to battery **10a** because, after discharge, the only path open to it is that to the battery **10a**, since the gaps in **40**, **41** and **42** have broken down, because the capacitors in banks **39** are discharged and have not yet recovered the high voltage charge from the high frequency pulsers **36** and the converter rectifier units **38**.

In the event of a misfire in the programmer control circuits **40**, **41** and **42**, the capacitors are discharged through a rotor safety discharge circuit **46** and returned to batteries **10-10a**, adding to their capacity. The circuit **46** is connected between the capacitor banks **39** and batteries **10**, **10a**.

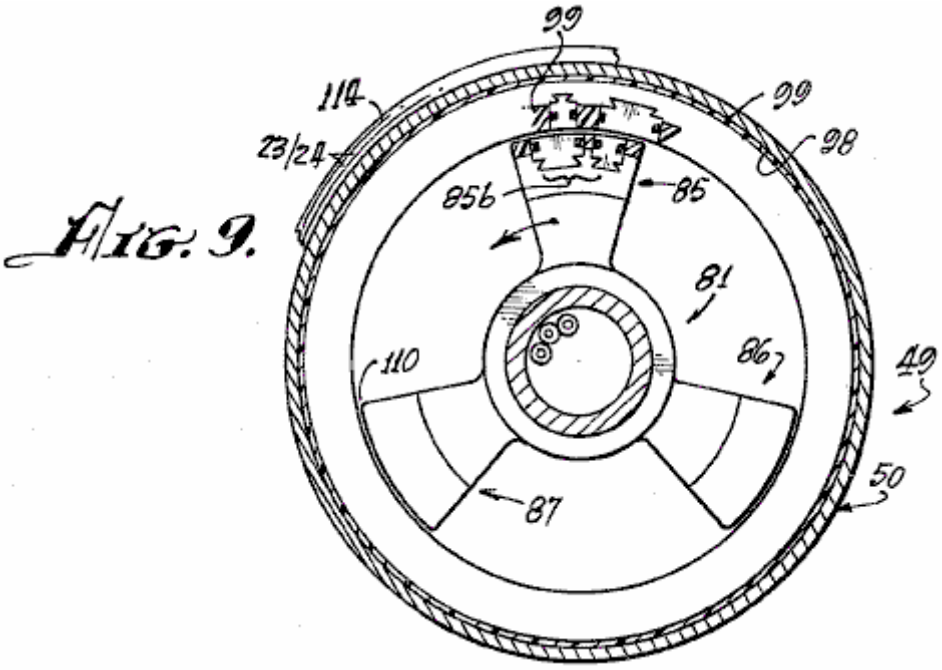


Referring to **Fig.3**, a motor or engine **49** according to the present invention is shown connected with an automotive transmission **48**. The transmission **48**, represents one of many forms of loads to which the engine may be applied. A motor housing **50**, encases the operating mechanism hereinafter described. The programmer 40 is axially mounted at one end of the housing. Through apertures **51** and **52**, a belt **53** couples to a pulley **57** (not shown in this view) and to an alternator **54** attached to housing **50**. A pulley **55** on the alternator, has two grooves, one for belt **53** to the drive pulley **58** on the shaft (not shown) of the engine **49**, and the other for a belt **58** coupled to a pulley **59** on a pump **60** attached to housing **50**, A terminal box **61** on the housing, interconnects between the battery assembly **62** and motor **49** via cables **63** and **64**.

An intake **65** for air, is coupled to pump **60** via piping **68** and **69** and from pump **60** via tubing or piping **66** and **70** to the interior of housing **50** via coupling flanges **67** and **71**. The air flow tends to cool the engine and the air may preferably be maintained at a constant temperature and humidity so that a constant spark gap discharge condition is maintained. A clutch mechanism **80** is provided on programmer **40**.



Referring to **Fig.4**, **Fig.5** and **Fig.9**, rotor **81** has spider assemblies **83** and **84** with three electromagnet coil assembly sets mounted thereon, two of which are shown in **Fig.4**, on **85**, at **85a** and **85b** and on **86** at **86a** and **86b**. One of the third electromagnet coil assemblies, designated **87a**, is shown in **Fig.5**, viewed from the shaft end. As more clearly shown in the perspective view of **Fig.8**, a third spider assembly **88** provides added rigidity and a central support for the rotor mechanism on shaft **81**.

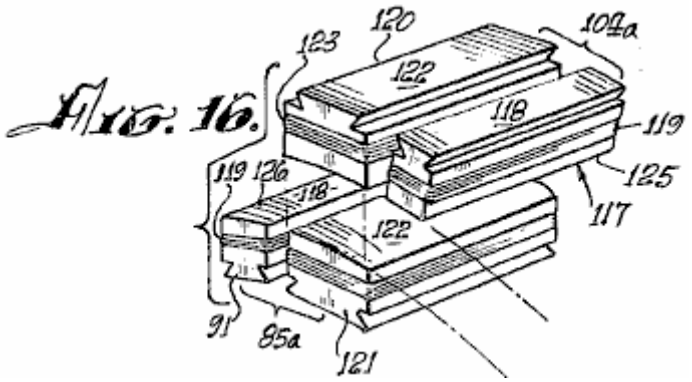


The electromagnet sets **85a**, **85b**, **86a**, **86b**, **87a** and **87b**, disposed on rotor **81** and spiders **83**, **84** and **88**, each comprise pairs of front units **85a**, **86a** and **87a** and pairs of rear units **85b**, **86b** and **87b**. Each pair consists of a major electromagnet and a minor electromagnet, as hereinafter described, which are imbedded in an insulating

material **90**, which insulates the electromagnet coil assemblies from one another and secures the electromagnets rigidly in place on the spider/rotor cage **81**, **83**, **84** and **88**.

The interior wall **98**, of housing **50**, is coated with an electrically insulating material **99** in which are imbedded electromagnet coils, as hereinafter described, and the interiors of end plates **100** and **101** of the housing **50**. On the insulating surface 98 of housing **50** is mounted a series of stator electromagnet pairs **104a**, identical with electromagnet pairs **85a**, **86a**, **87a**, etc. Electromagnet pairs such as **104a** or **105a** are disposed every **40** degrees about the interior of housing **50** to form a stator which co-operates with the rotor **81-88**. An air gap **110** of very close tolerance is defined between the rotor and stator electromagnets and air from pump **65** flows through this gap.

As shown in **Fig.8**, the electromagnet assemblies, such as **85** through **87**, of the rotor and magnet assemblies, such as **104a** in the stator, are so embedded in their respective insulating plastic carriers (rotor and stator) that they are smoothly rounded in a concave contour on the rotor to permit smooth and continuous rotation of rotor **81** in stator housing **50**. The air gap **110** is uniform at all positions of any rotor element within the stator assembly, as is clearly shown in **Fig.16**.

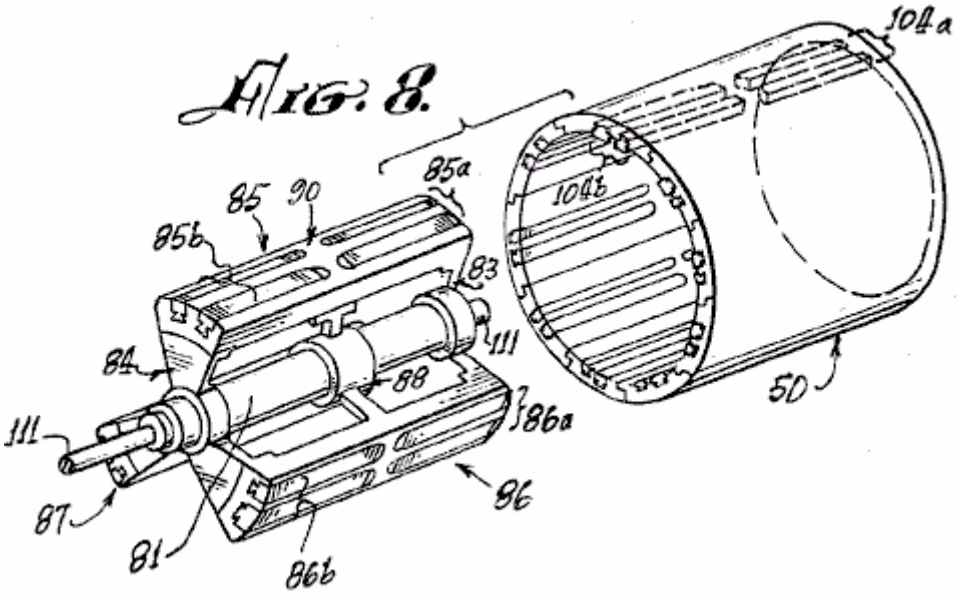


The rotor **81** and spiders **83**, **84** and **88** are rigidly mounted on shaft **111** journaled in bearing assemblies **112** and **113** which are of conventional type, for easy rotation of the rotor shaft **111** within housing **50**.

Around the central outer surface of housing **50**, are wound a number of turns of wire **23** and **24** to provide a static flux coil **114** which is a delay line, as previously described. **Figs. 5, 6, 7** and **9** are cross-sectional views of the rotor assembly **81-88**, arranged to show the positioning and alignment of the rotor and stator electromagnet coil assemblies at successive stages of the rotation of the rotor **81-88** through a portion of a cycle of operation thereof. For example, in **Fig.5** the rotor assembly **81-88** is shown so positioned that a minor rotor electromagnet assembly **91** is aligned with a minor stator electromagnet assembly **117**.

As shown in further detail in **Fig.16**, minor electromagnet assembly **117** consists of an iron core **118**, grooved so that a coil of wire **119** may be wound around it. Core **118** is the same in stator electromagnet **117** as it is in rotor electromagnet **91**.

As a position 13.33 degrees to the right of rotor electromagnet **91**, as viewed in **Fig.5** and **Fig.16**, there is a second or major rotor electromagnet **121** which has a winding **123** about its core **122**. The electromagnets **91** and **121** are the pair **85a** of **Fig.4** and **Fig.8**.



At a position 13.33 degrees to the left of stator electromagnet **117**, as viewed in **Fig.5**, there is a second or major stator electromagnet **120** whose core **122** is of the same configuration as core **122** of rotor electromagnet **121**. A winding **123** about core **122** of electromagnet **120** is of the same character as winding **123** on electromagnet **121**.

Electromagnet assembly pair **85a** on the rotor is identical in configuration to that of the electromagnet stator assembly pair **104a** except for the position reversal of the elements **117-120** and **91-121** of the respective pairs.

There are none pairs of electromagnets **120-117** (**104a**) located at 40 degree intervals about the interior of housing **50**. The centreline of core **122** of electromagnet **120** is positioned 13.33 degrees to the left of the centreline of the core **118** of electromagnet **117**. Three pairs of electromagnets **85a**, **86a** and **87a** are provided on rotor assembly **81-88** as shown in **Fig.5**.

Other combinations are possible, but the number of electromagnets in the rotor should always be in integral fraction of the number of electromagnets in the stator. As shown in **Fig.8**, for the rotor assembly **85a** and **85b**, there are three of each of the front and back pairs of electromagnetic assemblies. Similarly, as shown in **Fig.4** and **Fig.8**, there are nine front and back pairs of electromagnets in the stator such as **104a** and **104b**.

In order to best understand the operation of the rotor **81-88** rotating within the stator housing **50** of an engine according to this invention, the positions of rotor electromagnets **91** and stator electromagnets **117** are initially exactly in line at the 13.33 degree peripheral starting position marked on the vertical centreline of **Fig.5**. The winding direction of the coils of these magnets is such that a d-c current through the coils **119** will produce a particular identical magnet polarity on each of the juxtaposed surfaces **125** of magnet **117** and **126** of magnet **91** (**Fig.5**). **Fig.16** and **Fig.6** illustrate the next step in the motion wherein the two major electromagnets, **120** in the stator and **121** in the rotor, are in alignment.

When the d-c discharges from the appropriate capacitors in banks **39** occur simultaneously across spark gaps through the coils **119** of electromagnets **117** and **91**, at the instant of their alignment, their cores **118**, will repel one another to cause rotor assembly **81-88** to rotate clockwise in the direction indicated by arrow **127**. The system does not move in the reverse direction because it has been started in the clockwise direction by the alternator motor **54** shown in **Fig.3**, or by some other starter means. If started counterclockwise, the motor will continue to rotate counterclockwise.

As noted earlier, the discharge of any capacitor occurs over a very short interval via its associated spark gap and the resulting magnetic repulsion action imparts motion to the rotor. The discharge event occurs when electromagnets **117** and **91** are in alignment. As shown in **Fig.5**, rotor electromagnet **91a** is aligned with stator electromagnet **117c**, and rotor electromagnet **91b** is aligned with stator electromagnet **117e** at the same time that similar electromagnets **117** and **91** are aligned. A discharge occurs through all six of these electromagnets simultaneously (that is, **117**, **91**, **117c**, **91a**, **117e** and **91b**). A capacitor and a spark gap are required for each coil of each electromagnet. Where, as in the assembly shown in **Fig.8**, front and back pairs are used, both the axial in-line front and back coils are energised simultaneously by the discharge from a single capacitor or from a bank of paralleled capacitors such as **25** and **26** (**Fig.1**). Although **Fig.4** and **Fig.8** indicate the use of front and back electromagnets, it should be evident that only a single electromagnet in any stator position and a corresponding single electromagnet in the rotor position, may be utilised to accomplish the repulsion action of the rotor with respect to the stator. As stated, each electromagnet requires a discharge from a single capacitor or

capacitor bank across a spark gap for it to be energised, and the magnetic polarity of the juxtaposed magnetic core faces must be the same, in order to effect the repulsive action required to produce the rotary motion.

Referring to **Fig.5** and **Fig.6**, the repulsion action causes the rotor to move 13.33 degrees clockwise, while electromagnets **91**, **91a** and **91b** move away from electromagnets **117**, **117c** and **117e** to bring electromagnets **121**, **121a** and **121b** into respective alignment with electromagnets **120a**, **120d** and **120f**. At this time, a capacitor discharge across a spark-gap into their coils **123** occurs, thus moving the rotor. Another 13.33 degrees ahead, as shown in **Fig.7**, major electromagnets 121, 121a and 121b come into alignment with minor electromagnets **117a**, **117d** and **117f**, at which time a discharge occurs to repeat the repulsion action, this action continuing as long as d-c power is applied to the system to charge the capacitor banks.

Fig.18 further illustrates the sequencing of the capacitor discharges across appropriate spark gap terminal pairs. Nine single stator coils and three single rotor coils are shown with their respective interconnections with the spark gaps and capacitors with which they are associated for discharge. When the appropriate spark gap terminals are aligned, at the points in the positioning of the rotor assembly for most effective repulsion action of juxtaposed electromagnet cores, the discharge of the appropriate charged capacitors across the associated spark gap occurs through the respective coils. The capacitors are discharged is sets of three, through sets of three coils at each discharge position, as the rotor moves through the rotor positions. In **Fig.18**, the rotor electromagnets are positioned linearly, rather than on a circular base, to show the electrical action of an electric engine according to the invention. These motor electromagnets **201**, **202** and **203** are aligned with stator electromagnets **213**, **214** and **215** at 0 degrees, 120 degrees and 240 degrees respectively. The stator electromagnets are correspondingly shown in a linear schematic as if rolled out of the stator assembly and laid side by side. For clarity of description, the capacitors associated with the rotor operation **207**, **208**, **209** and **246**, **247**, **248**, **249**, **282** and **283**, are arranged in vertical alignment with the respective positions of the rotor coils **201**, **202** and **203** as they move from left to right, this corresponding to clockwise rotation of the rotor. The stator coils **213, 214, 215, 260, 261, 262, 263, 264, 265, 266,** etc. and capacitor combinations are arranged side by side, again to facilitate description.

An insulative disc **236** (shown in **Fig.17** as a disc but opened out linearly in **Fig.18**) has mounted thereon, three gap terminal blocks **222**, **225** and **228**. Each block is rectangularly U-shaped, and each interconnects two terminals with the base of the U. Block **222** has terminals **222a** and **222b**. Block **225** has terminals **225a** and **225b**. Block **228** has terminals **228c** and **228d**. When insulative disc **230** is part of the rotor as indicated by mechanical linkage **290**, it can be seen that terminal U **222** creates a pair of gaps with gap terminals **223** and **224** respectively. Thus, when the voltage on capacitor **216** from charging unit **219**, is of a value which will arc over the air spaces between **222a** and **223**, and between **222b** and **224**, the capacitor **216** will discharge through the coil of electromagnet **213** to ground. Similarly, gap terminal U 225 forms a dual spark gap with gap terminals **226** and **227** to result in arc-over when the voltage on capacitor **217**, charged by charging circuit **220**, discharges into the coil of electromagnet **214**. Also, U-gap terminal **228** with terminals **228c** and **228d**, creates a spark gap with terminals **229** and **230** to discharge capacitor **218**, charged by charging circuit **221**, into coil **215**. At the same time, rotor coils, **201**, **202** and **203** across gaps **201a - 204**, **202b - 205** and **203c - 206** each receives a discharge from respective capacitors **207**, **208** and **209**.

When the electromagnet coils **213**, **214** and **215** and **201**, **202** and **203** are energised, the repulsion action causes the rotor assembly to move to position 2 where a new simultaneous group of discharges occurs into rotor coils **201**, **202** and **203** from capacitors **246**, **248** and **282** across gaps **201a - 240**, **202b - 242** and **203c - 244**. Simultaneously, because gap-U-elements **222**, **225** and **228** have also moved to position 2 with the rotor assembly, capacitor **261** is discharged through electromagnet coil **260**, capacitor **265** is discharged through electromagnet coil **264**, and capacitor **269** is discharged through electromagnet coil **268** in alignment with position 2 of the rotor electromagnet coils, thus to cause the rotor electromagnets to move to position 3 where the discharge pattern is repeated now with capacitors **247**, **249** and **283** discharging through the rotor electromagnet coils **201**, **202** and **203**, and the capacitors **263**, **267** and **281** discharging respectively through stator electromagnet coils **262**, **266** and **280**.

After each discharge, the charging circuits **219 - 221** and **272 - 277** for the stator capacitors, and **210 - 212** and **284 - 289** for the rotor capacitors, are operated continuously from a battery source as described earlier with reference to **Fig.1**, to constantly recharge the capacitors to which each is connected. Those versed in the art will appreciate that, as each capacitor discharges across an associated spark gap, the resulting drop in potential across the gap renders the gap an open circuit until such time as the capacitor can recharge to the arc-over level for the gap. This recharge occurs before a rotor element arrives at the next position in the rotation.

The mechanical schematic diagram of **Fig.17**, further clarifies the operation of the spark-gap discharge programming system. A forward disc **236** of an electrically insulative material, has thereon the set of U-shaped gap terminal connectors previously described. These are positioned at 0 degrees, 120 degrees and 240 degrees respectively. In **Fig.17**, schematic representations of the position of the coil and capacitor arrangements at the

start of a cycle are shown to correspond to the above description with reference to **Fig.18**. Accordingly, the coil and capacitor combinations **213/216, 214/217** and **215/218** are shown connected with their gap terminals, respectively, **223/224, 226/227** and **229/230**. On the rotor coil and capacitor connection, three separate discs **291, 292** and **293** are shown, each with a single gap terminal. The discs **291 - 293** are rotated so as to position their respective gap terminals 201a, 201b and 201c, at 120 degree increments, with the 0 degrees position corresponding to the 0 degrees position of U-gap terminal **222** on disc **230**.

Representative gap terminals are shown about the peripheries of discs **230, 291 - 293** to indicate clearly how, as the discs turn in unison, the gap alignments correspond so that three rotor coils always line up with three stator coils at 120 degree intervals about the rotary path, producing an alignment every 40 degrees, there being nine stator coils. Thus, there are three simultaneous discharges into stator coils and three into rotor coils at each 40 degree position. Nine positions displaced 40 degrees apart provide a total of 27 discharge points for capacitors into the rotor coils and 27 discharge points for capacitors into the stator coils in one revolution of the rotor.

It will be understood that, as illustrated in **Fig.17** and **Fig.18**, nine individual electromagnet coils are shown in the stator and three in the rotor, in order to show in its simplest form, how the three rotor electromagnets are stepped forward from alignment with three of the stator electromagnets, when the appropriate spark gaps are in alignment, to effect the discharge of capacitors through juxtaposed pairs of rotor/stator electromagnets. The repulsion moves the rotor electromagnet from the stator electromagnet to the next alignment position 40 degrees further on. In the interval, until another rotor electromagnet, 120 degrees removed, is aligned with the stator electromagnet which had just been pulsed, the associated capacitor is recharged. Thus, the rotor moves from one position to the next, with capacitor discharges occurring each 40 degrees of rotation, a total of nine per revolution. It should be obvious that, with other rotor/stator combinations, the number of electromagnet coincidences and spark-gap discharges will vary. For example, with the coil pairs shown in **Figs 4 through 8**, a total of 27 discharges will occur. Although there are 18 stator electromagnets and 3 rotor electromagnets, the discharge pattern is determined by the specific spark gap arrangement.

The rotor/stator configuration of **Fig.5** and **Fig.8**, involving the major and minor pairs of electromagnets, such as **85a** and **104a** (the terms “minor” and “major” referring to the difference in size of the elements), include nine pairs of electromagnets in the stator, such as **104a**, with three electromagnet pairs of the rotor, such as **85a**. Because of the 13.33 degree separation between the major and minor electromagnets in the rotor pair **85a**, with the same separation of minor and major electromagnets of the stator pair **104a**, the sequence of rotation and discharge described above, with respect to the illustrative example of **Fig.5**, involves the following:
1. A minor element **117** of stator pair **104a** is aligned with the minor element **91** of rotor pair **85a**. On the discharge, this moves the rotor ahead 13.33 degrees.
2. the major rotor element **122** of the pair **85a**, now is aligned with the major stator element **120b** of the next stator electromagnet pair, in the stator array as shown in **Fig.6**. On the discharge, the rotor moves ahead 13.33 degrees.
3. This brings the minor rotor electromagnet **91** into alignment with the major stator electromagnet **120b** of pair **104d**, and the major electromagnet **122** (just discharged) of pair **85a** into alignment with minor electromagnet **117b** of pair **104d**, and the rotor spark gap elements into alignment with a different position of gap elements connected with capacitors not discharged in the previous position of the rotor. It should be remembered at this point that it is the positioning of a rotatable spark gap array, similar to that illustrated in **Fig.17** and **Fig.18**, which controls the time of discharge of capacitors connected to these gap terminals. Therefore, any electromagnet can be energised twice, successively, from separate capacitors as the rotor brings appropriate gap terminals into alignment with the coil terminals of a particular electromagnet.

Thus, although major electromagnet **120b** of pair **104d** has just been energised as described above, it can now be energised again along with minor rotor electromagnet **91** in step 3, because the rotor moved to a new set of terminals of the spark gap arrays connected to capacitors which have not yet been discharged. These capacitors now discharge through rotor electromagnet **91** and stator electromagnet **120b**, causing the rotor to move ahead another 13.33 degrees, thus again aligning two minor electromagnets again, these being **117b** of stator pair **104d** and **91** of rotor pair **85a**. The rotor has now moved 40 degrees since step 1 above. The sequence is now repeated indefinitely. It is to be noted that at each 13.33 degree step, the discharges drive the rotor another 13.33 degrees. There are 27 steps per revolution with nine stator coil pairs. The discharge sequence is not uniform, as is shown in Table 1. In the stator, three major electromagnets 120 degrees apart are energised twice in sequence, followed by a hiatus of one step while three minor electromagnets of the stator, 120 degrees apart, are energised during the hiatus. In the rotor the major electromagnets are energised during a hiatus step following two minor electromagnet energisation steps. A total of 27 energisations are this accomplished in the nine pairs of coils of the stator.

In Table 1, the leftmost column shows the location of each rotor arm **85, 86** and **87** at an arbitrarily selected step No. 1 position. For example, in step 1, rotor arm **85** has a minor stator and minor rotor electromagnet in alignment for capacitors to discharge through them simultaneously at the 13.33 degree position.

TABLE I

CAPACITOR DISCHARGE SEQUENCE FOR ONE REVOLUTION OF ROTOR IN SYSTEM OF FIGS. 5 AND 8

Step No. Rotor Arm 87	85	86	Rotor Angle	Rotor Electro-magnet Minor	Rotor Electro-magnet Major	Stator Electro-magnet Minor	Stator Electro-magnet Major
10	1	19	13 1/3°	x		x	
11	2	20	26 2/3°		x		x
12	3	21	40°	x			x
13	4	22	53 1/3°	x		x	
14	5	23	66 2/3°		x		x
15	6	24	80°	x			x
16	7	25	93 1/3°	x		x	
17	8	26	106 2/3°		x		x
18	9	27	120°	x			x
19	10	1	133 1/3°	x		x	
20	11	2	146 2/3°		x		x
21	12	3	160°	x			x
22	13	4	173 1/3°	x		x	
23	14	5	186 2/3°		x		x
24	15	6	200°	x			x
25	16	7	213 1/3°	x		x	
26	17	8	226 2/3°		x		x
27	18	9	240°	x			x
1	19	10	253 1/3°	x		x	
2	20	11	266 2/3°		x		x
3	21	12	280°	x			x
4	22	13	293 1/3°	x		x	
5	23	14	306 2/3°		x		x
6	24	15	320°	x			x
7	25	16	333 1/3°	x		x	
8	26	17	346 2/3°		x		x
9	27	18	360°	x			x

Similarly, in step 1, rotor arm **86** is at the 133.33 degree position which has two minor electromagnets in alignment, ready for discharge. Simultaneously, rotor arm 87 is at the 253.33 degree position with two minor electromagnets aligned for capacitor discharge. The other steps of the sequence are apparent from Table 1, for each position of the three rotor arms at any step and the juxtapositions of respective stator and rotor electromagnet elements at that position.

In the simplified motor arrangement shown in schematic form in **Fig.18**, with single electromagnet configuration, the alignment is uniform and the discharge sequences follow sequentially.

As mentioned before, a change in speed is effected by displacing the stator spark gap terminals on the rotor (shown at **236** in **Fig.17** and **Fig.18**) either counterclockwise or clockwise 6.66 degrees so that the discharge position of the stator electromagnets is displaced. Referring to **Figs. 11 to 15**, the simultaneous discharge of selected capacitors into the displaced electromagnets results in a deceleration if the rotor electromagnet is approaching the stator electromagnet at the time of discharge, or an acceleration if the rotor electromagnet is leaving the stator electromagnet at the time of the discharge pulse. In each event, there is a repulsive reaction between the stator and rotor electromagnets which effects this change in speed.

Referring to **Fig.11**, clutch mechanism **304** about shaft **111** is operated electromagnetically in conventional manner, to displace the spark-gap mechanism **236** which is operated normally in appropriate matching alignment with the rotor spark-gap discs **291, 292** and **293**. Clutch **304** has a fixed drive element **311**, containing an electromagnetic drive coil (not shown) and a motor element **310** which, when the electromagnetic drive coil is energised, can be operated by a direct current. The operation of motor element 310, brings into operation, spark gap elements **224r, 223r** or **223f, 224f** of the system shown in **Figs. 4, 5 and 8,** as illustrated in **Fig.19**.

The fixed stator coil spark gap terminal pairs **223, 224** and **266, 267** are arrayed about a cylindrical frame **322** which is fabricated in insulative material. In the illustrative example of **Fig.17** and **Fig.18**, there are nine such spark gap terminal pairs positioned around the periphery of the cylinder frame **324**. In the engine of **Figs. 4 to 8**, a total of 27 such spark gap pairs are involved. In addition, although not shown in the drawing, there are also pairs of terminals, such as **223r or 223f**, **224r or 224f** and **226r or 226f**, **267r or 267f**, displaced 6.66 degrees on either side of the pairs **223, 224** or **266, 267** and all other pairs in the spark gap array, the letters “r” and “f” denoting “retard” or “faster”. The latter displaced pairs are used in controlling the speed of the engine rotor. The displaced pairs not shown are involved in the operation of the clutch **304**, the speed-changing control element.

Clutch **304** is associated with shaft **111** in that the movable element **310** draws clutch disc element **316** on shaft **111**, away from clutch disc element **322** when energised by a voltage of appropriate polarity applied to its motor electromagnet **311**. Such clutch drives are well known in the art.

The clutch mechanism **304** of **Fig.11** and **Fig.19**, when not energised, is in the configuration shown in **Fig.11**. The energised configuration of clutch **304** is not specifically illustrated. Upon energisation, spark-gap element **222** on disc **236** is displaced rightward, as viewed in **Fig.11**, by broken lines **236X**, into alignment with the positions of fixed spark-gap terminals **223f, 224f** and **267r, 266r**. When the disc is in position **236X**, the flattened edge **332** of pin **330** in disc **325** rides on surface **350** of disc **322**. Normally, the flattened edges **351** of pins **330** are engaged against the flat edge **352** in recess **331** of disc **322**. The displacement of disc **322** on shaft **111** is effected by the action of clutch **304** against spring **314** (Fig.11). An electric switch (not shown) of clutch mechanism **304** energises it from a d-c power source, and has two positions, one for deceleration and one for acceleration. In either position, clutch **304** is engaged to pull clutch disc **322** from clutch disc **325**, momentarily. For the decelerate or the accelerate position, the displaced alignment of spark gap elements **222** is with the **224f, 223f** and the **224r, 223r** spark-gap terminal elements. However, only the **224f, 223f** spark-gap elements are switched into operation with appropriate capacitors for the accelerate position, while in the decelerate position, only the **223r and 224r** spark-gap elements are switched into the circuit with their associated capacitors.

Of course, when insulative disc **236** is displaced by clutch **304**, its gap terminals **222, 225** and **228** (**Fig.14** and **Fig.18**) are all displaced into the alignment position of **236X** so as to engage the "r" and "f" lines of fixed spark gap elements. Although the accelerate and decelerate positions of disc **236** are the same, it is the switching into operation of the **223, 224** or **266, 267** exemplary "r" or "f" pairs of terminals which determines whether the rotor will speed up or slow down.

The momentary displacement of clutch disc **322** from clutch disc **325** results in rotation of disc **325** about disc **322** through an angle of 120 degrees. The detent ball and spring mechanism **320, 321** in disc **325**, positions itself between one detent dimple **328** and a succeeding one **328** at a position 120 degrees away on disc **325**.

As stated, flat **332** of pin **330** rides on surface **350** of disc **322**, and pin **330** leaves the pin-holding groove **331/352** along ramp **333** in disc **322** during the momentary lifting of disc **322** by clutch **304**. Pin **330** falls back into the next groove **331** at a point 120 degrees further on about disc **322**. Pin **330** falls into place in groove **331** on ramp **334**. Pins **330** are rotatable in their sockets **353**, so that for either clockwise or counterclockwise rotation, the flat **351** will engage the flat **352** by the particular ramp it encounters.

The deceleration or acceleration due to the action of clutch **304** thus occurs within a 120 degree interval of rotation of disc **325**. During this interval, disc **322** may only move a fraction of this arc.

There has been described earlier, an electromotive engine system wherein at least one electromagnet is in a fixed position and a second electromagnet of similar configuration is juxtaposed with it in a magnetic polarity relationship such that, when the cores of the electromagnets are energised, the juxtaposed core faces repel each other. One core being fixed, and the second core being free to move, any attachments to the second electromagnet core will move with it. Hence, if a plurality of fixed cores are positioned about a circular confining housing, and, within the housing, cores on a shaft are free to move, the shaft is urged rotationally each time the juxtaposed fixed and rotatable cores are in alignment and energised. Both the fixed and the movable cores are connected to spark gap terminal elements and the associated other terminal elements of the spark gaps are connected to capacitors which are charged to high voltage from pulsed unipolar signal generators. These capacitors are discharged through the electromagnets across the spark gaps. By switching selected groups of capacitors into selected pairs of spark gap elements for discharge through the electromagnets, the rotor of the circular array systems is accelerated and decelerated.

By confining a fixed electromagnet array in a linear configuration, with a linearly movable electromagnet to which a working tool is attached, exciting the juxtaposed pairs of electromagnets by capacitor discharge, results in the generation of linear force for such tools as punch presses, or for discharging projectiles with considerable energy.

CLAIMS:

1. An electric engine comprising:

A housing;

An array of electromagnets uniformly spaced in said housing to form a stator;

A rotor cage on a shaft journaled in and rotatable within said stator, said rotor cage having thereon a spaced array of electromagnets similar to said stator electromagnets and in number, comprising an integral fraction of the number of electromagnets in said stator array;

Each of the electromagnets of said stator and of said rotor, having a core which can be magnetised and of a particular configuration and each being wound with a coil such that a pulses of unidirectional electric current through said coil, magnetises the respective core thereof to a particular magnetic polarity, and the faces of rotor cores juxtaposing selected stator cores are magnetised to the same polarity, the juxtaposed cores thereby tending to repel one another, one lead of each of the stator and rotor coils being connected to a common terminal, the other lead of each of said coils being connected to a gap terminal, the gap terminals of said rotor coils being on the rotor and equal in number to the number of coils thereon and matching the positions of said rotor electromagnets thereon, the gap terminals of said stator being equal in number to the number of coils on the stator and disposed uniformly about said stator to match the positions of said stator electromagnets within said housing;

A first array of capacitors, each having a terminal in common with the common coil terminal of said stator electromagnets, and each capacitor having its other terminal connected to a gap terminal arrayed adjacent the gap terminal of an electromagnet associated therewith;

A second array of capacitors, each having a terminal in common with said common terminal of said rotor electromagnet coils but equal in number to the number of capacitors in said stator array, the other terminals of said capacitors in said second array being connected to gap terminals arrayed about said housing so as to be in axial alignment with said stator gap terminal positions and being alignable with said rotor gap terminals as said rotor is rotated in said housing and respective gap terminals of said rotor coils pass each second array capacitor gap terminals at a predetermined gap distance;

Gap coupling terminals on said rotor equal in number to the number of rotor electromagnet coils and positioned to match the rotor electromagnet positions on said rotor, the gap coupling terminals being rotatable with said rotor so as to pass said adjacent stator coil and associated stator capacitor gap terminal at a predetermined distance therefrom;

A plurality of capacitor charging circuits connected respectively across each of said capacitors in both said first and said second arrays of capacitors for charging each of said capacitors to a predetermined high d-c potential;

A first source of unidirectional electric potential connected to each of said capacitor charging circuits for energising said charging circuits; and

A second unidirectional electric potential source connected to said electromagnets of said rotor and said stator of such polarity as to receive a charge from the inverse inductive discharge of the electromagnet coils as their fields collapse following the discharge of each capacitor through a rotor or stator electromagnet coil,

Whereby, whenever a rotor electromagnet is aligned opposite a stator electromagnet, the rotor coil gap terminal of that electromagnet is opposite an associated second capacitor array gap terminal, and a gap coupling terminal of said rotor is aligned opposite the stator electromagnet coil gap terminal and associated first capacitor gap terminal, the capacitors discharge the charge thereon across the gaps through their associated electromagnet coils to magnetise their respective juxtaposed electromagnet cores to cause them to repel one another, thus aligning a succeeding pair of rotor and stator electromagnets for capacitor discharge across their respective gaps, to cause them to repel one another, alignments rotor rotation within the housing continuously bringing successive rotor-stator electromagnets into alignment for discharge of the capacitors through them to produce continuous rotary motion of the rotor on said rotor shaft, so long as energy is applied to said charging circuits to recharge said capacitors after each discharge.

2. In an electric engine having a rotor comprising electromagnetic coil means roatatable within a stator comprising similar electromagnetic coil means, said electromagnetic coil means being polarised for magnetic repulsion;

Capacitor means electrically coupled across successive spark gaps to selected ones of said stator and all of the coils of said rotor;

Charging means connected to said capacitor means for charging said capacitor means to an electrical charge potential sufficient to cause arcing across said spark gaps to result in the discharge of said capacitor means through the electromagnetic coil means repel one another; and

A unidirectional electric power source connected to said charging means to energise said charging means to continue charging said capacitor means following each discharge whereby the rotor of said engine is maintained

in rotation by the successive discharges of said capacitor means across successive spark gaps into said electromagnetic coil means.

3. An electric engine according to claim 2, wherein:

The charging means includes electronic square core oscillators connected to said unidirectional electric power source and includes step-up means and a rectifier to produce a substantial voltage step up from the voltage of said power source.

4. An electric engine according to claim 2, wherein:

The charging means includes a vibrator connected to said power source, and step-up transformer and rectifier means to provide a high voltage for charging said capacitor means.

5. A motive force-producing means comprising:

At least a first electromagnet means including at least one coil wound about a core,

At least a second electromagnet means including at least one coil wound about a core similar to said first core,

The respective cores being positioned adjacent to one another so that the magnetic polarities of the adjacent core surfaces are the same when a unidirectional electric current is passed through the coils,

At least one capacitor means having one terminal thereof connected to one terminal of both of said electromagnet coils,

The other terminal of said capacitor means being connected to one terminal of a spark gap means, the other terminals of the coils of both said first and said second electromagnet means being connected to the other terminal of said spark gap means,

At least one unidirectional pulse charging means connected to said capacitor means to charge said capacitor means to a relatively high potential sufficient to arc across said spark gap means at predetermined spacing of said gap terminals, and

A source of unidirectional potential connected to said charging circuit to energise said charging means,

Whereby upon application of current from said potential source to said charging means the successive pulses generated thereby charge said capacitor means to a voltage level sufficient to arc across said spark gap means to produce a discharge path for said capacitor means through said coils to cause said electromagnet means to repel one another with a substantial force.

6. A motive force-producing means according to claim 5, wherein:

Said first electromagnet means is secured in a relatively stable housing, and said second electromagnet means is connected with and freely movable relative to said stable housing, and has utilisation means connected thereto for performing work therewith when said capacitor means discharges through said coils of said electromagnet means.

7. A motive force-producing means according to claim 6, wherein said utilisation means is a motor rotor coupled with said second electromagnet means and said first electromagnet means is a stator.

8. A motive force-producing means according to claim 6, wherein said utilisation means is a piston attached to said second electromagnet means and is movable therewith to produce hammer-like blows when said capacitor means discharges through said electromagnet means.

9. In an electromotive force-generating system as disclosed, means for accelerating or decelerating the motion of a force-generating system, said means comprising:

At least two juxtaposed electromagnetic core elements, one fixed and one movable, including coils wound around it to provide a repulsion tendency when said cores are energised,

Spark gap terminals connected with said coils,

Capacitor means connected with said spark gap terminals to discharge across said spark gap terminals through said coils when a charge of sufficient voltage level appears across said capacitor means, thus to energise said juxtaposed electromagnets to induce said juxtaposed electromagnet cores to repel one another,

Charging means connected to said capacitors for charging them to said sufficient voltage level, and selective positioning means coupled with said spark gap terminals and with at least said movable electromagnet core to cause selective displacement of said movable core with respect to said fixed core.

10. An electromotive force-generating system according to claim 9, wherein:

Said juxtaposed electromagnetic cores include a plurality of fixed cores and a smaller number of movable cores, said smaller number being an integral fraction of the number of fixed cores, and

Said selective positioning means is an electromagnetic clutch coupled with said smaller number of movable cores for movement therewith, and includes selective displacement means coupled with said spark gap terminals connected with said capacitors in said capacitor means and selected combinations of coils in said plurality of fixed electromagnets.

11. The method of generating motive power comprising the steps of:

a. positioning similar electromagnets in juxtaposed relationship with their respective cores arranged for repulsion when said electromagnets are energised,

b. charging capacitors to a relatively high potential, and

c. discharging said capacitors simultaneously through said electromagnets across spark gaps set to break down at said relatively high potential, thereby to cause said similar electromagnets to repel one another with considerable force.

12. The method of generating motive power defined in claim 11, wherein, in said positioning step at least one of said electromagnets is maintained in a fixed position and another electromagnet is free to move relative to said fixed electromagnet.

13. The method of generating motive power according to claim 11, wherein:

The charging step includes the charging of capacitors to a relatively high potential from a pulsed unipolar source of electrical energy.

14. in an electromagnetic capacitor discharge engine including movable electromagnets and fixed electromagnets, said movable electromagnets being movable into polar alignment with said fixed electromagnets, capacitor means, means for charging said capacitor means, and means for discharging said charged capacitor means through said fixed and movable electromagnets to polarise aligned fixed and movable electromagnets for magnetic repulsion, an acceleration and deceleration control means comprising:

First selective means for momentarily delaying the discharge of the capacitors until the movable electromagnets in said engine have begun to recede from the fixed electromagnets, in order to accelerate the motion of said movable electromagnets by the added impetus of the repulsion, and

Second selective means for momentarily accelerating the discharge of the capacitors to occur at a point in the motion of the movable electromagnets where said movable electromagnets are approaching said fixed electromagnets to decelerate the motion of said movable electromagnets by the tendency to repel the approaching electromagnets by the fixed electromagnets.

15. An electric engine, comprising:

Fixed electromagnets;

Movable electromagnets, movable into alignment with said fixed electromagnets;

Capacitor means;

Means for charging said capacitor means, and

Means for discharging said charged capacitor means through said fixed and movable electromagnets to polarise said aligned fixed and movable electromagnets for magnetic repulsion.

16. An electric engine as recited in claim 15, wherein: said means for discharging said charged capacitor means comprises voltage breakdown switch means.

17. An electric engine as recited in claim 16, wherein:

Said voltage breakdown switch means includes at least one terminal movable with at least one of said movable electromagnets for breaking down when said at least one of said movable electromagnets is in alignment with a said fixed electromagnet.

18. An electric engine as recited in claim 17, wherein:

Said voltage breakdown switch means comprises a spark gap means.

US Patent 4,595,975 June 17, 1986 Inventor: Edwin V. Gray snr.

EFFICIENT POWER SUPPLY SUITABLE FOR INDUCTIVE LOADS

Please note that this is a re-worded excerpt from this patent. It describes the circuitry used with Edwin Gray's power tube. Please be aware Edwin wanted at all costs, to conceal any useful technology while getting patents to encourage investors, so please understand that this patent is not intended to tell you how to make a working system of this type.

Fig.1 is a schematic circuit diagram of the electrical driving system.
Fig.2 is an elevational sectional view of the electrical conversion element.
Fig.3 is a plan sectional view taken along line 3--3 of Fig.2.
Fig.4 is a plan sectional view taken along line 4--4 of Fig.2.
Fig.5 is a schematic circuit diagram of the alternating-current input circuit.

SUMMARY OF THE INVENTION

The present invention provides a more efficient driving system comprising a source of electrical voltage; a vibrator connected to the low-voltage source for forming a pulsating signal; a transformer connected to the vibrator for receiving the pulsating signal; a high-voltage source, where available, connected to a bridge-type rectifier; or the bridge-type rectifier connected to the high voltage pulse output of the transformer; a capacitor for receiving the voltage pulse output; a conversion element having first and second anodes, electrically conductive means for receiving a charge positioned about the second anode and an output terminal connected to the charge receiving means, the second anode being connected to the capacitor; a commutator connected to the source of electrical voltage and to the first anode; and an inductive load connected to the output terminal whereby a high energy discharge between the first and second anodes is transferred to the charge receiving means and then to the inductive load.

As a sub-combination, the present invention also includes a conversion element comprising a housing; a first low voltage anode mounted to the housing, the first anode adapted to be connected to a voltage source; a second high voltage anode mounted to the housing, the second anode adapted to be connected to a voltage source; electrically conductive means positioned about the second anode and spaced therefrom for receiving a charge, the charge receiving means being mounted to the housing; and an output terminal communicating with the charge receiving means, said terminal adapted to be connected to an inductive load.

The invention also includes a method for providing power to an inductive load comprising the steps of providing a voltage source, pulsating a signal from said source; increasing the voltage of said signal; rectifying said signal; storing and increasing the signal; conducting said signal to a high voltage anode; providing a low voltage to a second anode to form a high energy discharge; electrostatically coupling the discharge to a charge receiving element; conducting the discharge to an inductive load; coupling a second capacitor to the load; and coupling the second capacitor to the source.

It is an aim of the present invention to provide a system for driving an inductive load which system is substantially more efficient than any now existing. Another object of the present invention is to provide a system for driving an inductive load which is reliable, is inexpensive and simply constructed.

The foregoing objects of the present invention together with various other objects, advantages, features and results thereof which will be evident to those skilled in the art in light of this disclosure may be achieved with the exemplary embodiment of the invention described in detail hereinafter and illustrated in the accompanying drawings.

DESCRIPTION OF THE PREFERRED EMBODIMENT

While the present invention is susceptible of various modifications and alternative constructions, an embodiment is shown in the drawings and will herein be described in detail. It should be understood however that it is not the intention to limit the invention to the particular form disclosed; but on the contrary, the invention is to cover all modifications, equivalents and alternative constructions falling within the spirit and scope of the invention as expressed in the appended claims.

There is disclosed herein an electrical driving system which, on theory, will convert low voltage electric energy from a source such as an electric storage battery to a high potential, high current energy pulse that is capable of developing a working force at the inductive output of the device that is more efficient than that which is capable of being developed directly from the energy source. The improvement in efficiency is further enhanced by the capability of the device to return that portion of the initial energy developed, and not used by the inductive load in the production of mechanical energy, to the same or second energy reservoir or source for use elsewhere, or for storage.

This system accomplishes the results stated above by harnessing the “electrostatic” or “impulse” energy created by a high-intensity spark generated within a specially constructed electrical conversion switching element tube. This element utilises a low-voltage anode, a high-voltage anode, and one or more “electrostatic” or charge receiving grids. These grids are of a physical size, and appropriately positioned, as to be compatible with the size of the tube, and therefore, directly related to the amount of energy to be anticipated when the device is operating.

The low-voltage anode may incorporate a resistive device to aid in controlling the amount of current drawn from the energy source. This low-voltage anode is connected to the energy source through a mechanical commutator or a solid-state pulser that controls the timing and duration of the energy spark within the element. The high-voltage anode is connected to a high- voltage potential developed by the associated circuits. An energy discharge occurs within the element when the external control circuits permit. This short duration, high-voltage, high-current energy pulse is captured by the “electrostatic” grids within the tube, stored momentarily, then transferred to the inductive output load.

The increase in efficiency anticipated in converting the electrical energy to mechanical energy within the inductive load is attributed to the utilisation of the most optimum timing in introducing the electrical energy to the load device, for the optimum period of time.

Further enhancement of energy conservation is accomplished by capturing a significant portion of the energy generated by the inductive load when the useful energy field is collapsing. This energy is normally dissipated in load losses that are contrary to the desired energy utilisation, and have heretofore been accepted because no suitable means had been developed to harness this energy and restore it to a suitable energy storage device.

The present invention is concerned with two concepts or characteristics. The first of these characteristics is observed with the introduction of an energising cur- rent through the inductor. The inductor creates a contrary force (counter-electromotive force or CEMP) that opposes the energy introduced into the inductor. This CEMF increases throughout the time the introduced energy is increasing.

In normal applications of an alternating-current to an inductive load for mechanical applications, the useful work of the inductor is accomplished prior to terminating the application of energy. The excess energy applied is thereby wasted.

Previous attempts to provide energy inputs to an inductor of time durations limited to that period when the optimum transfer of inductive energy to mechanical energy is occurring, have been limited by the ability of any such device to handle the high current required to optimise the energy transfer.

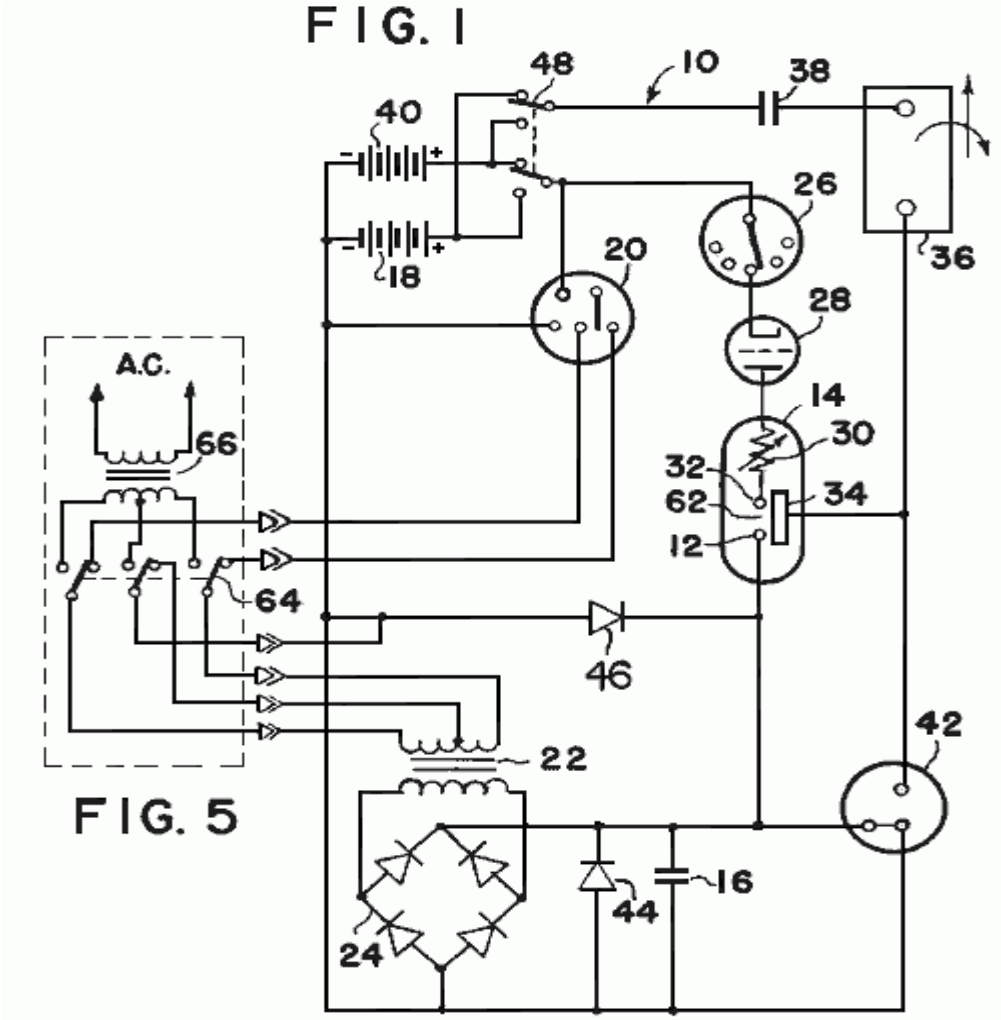
The second characteristic is observed when the energising current is removed from the inductor, As the current is decreased, the inductor generates an EMF that opposes the removal of current or, in other words, produces an energy source at the output of the inductor that simulates the original energy source, reduced by the actual energy removed from the circuit by the mechanical load. This “regenerated”, or excess, energy has previously been lost due to a failure to provide a storage capability for this energy.

In this invention, a high-voltage, high-current, short duration energy pulse is applied to the inductive load by the conversion element. This element makes possible the use of certain of that energy impressed within an arc across a spark-gap, without the resultant deterioration of circuit elements normally associated with high energy electrical arcs.

This invention also provides for capture of a certain portion of the energy induced by the high inductive kick

produced by the abrupt withdrawal of the introduced current. This abrupt withdrawal of current is attendant upon the termination of the stimulating arc. The voltage spike so created is imposed upon a capacitor that couples the attendant current to a secondary energy storage device.

A novel, but not essential, circuit arrangement provides for switching the energy source and the energy storage device. This switching may be so arranged as to actuate automatically at predetermined times. The switching may be at specified periods determined by experimentation with a particular device, or may be actuated by some control device that measures the relative energy content of the two energy reservoirs.



Referring now to **Fig.1**, the system **10** will be described in additional detail. The potential for the high- voltage anode, **12** of the conversion element **14** is developed across the capacitor **16**. This voltage is produced by drawing a low current from a battery source **18** through the vibrator **20**. The effect of the vibrator is to create a pulsating input to the transformer **22**. The turns ratio of the transformer is chosen to optimise the volt- age applied to a bridge-type rectifier **24**. The output of the rectifier is then a series of high-voltage pulses of modest current. When the available source is already of the high voltage, AC type, it may be coupled directly to the bridge-type rectifier.

By repetitious application of these output pulses from the bridge-type rectifier to the capacitor **16**, a high-voltage, high-level charge is built up on the capacitor.

Control of the conversion switching element tube is maintained by a commutator **26**. A series of contacts mounted radially about a shafts or a solid-state switching device sensitive to time or other variable may be used for this control element. A switching element tube type one-way energy path **28** is introduced between the commutator device and the conversion switching element tube to prevent high energy arcing at the commutator current path. When the switching element tube is closed, current from the voltage source **18** is routed through a resistive element **30** and a low voltage anode **32**. This causes a high energy discharge between the anodes within the conversion switching element tube **14**.

The energy content of the high energy pulse is electrostatically coupled to the conversion grids **34** of the conversion element. This electrostatic charge is applied through an output terminal **60** (**Fig.2**) across the load

inductance **36**, inducing a strong electromagnetic field about the inductive load. The intensity of this electromagnetic field is determined by the high electromotive potential developed upon the electrostatic grids and the very short time duration required to develop the energy pulse.

If the inductive load is coupled magnetically to a mechanical load, a strong initial torque is developed that may be efficiently utilised to produce physical work

Upon cessation of the energy pulse (arc) within the conversion switching element tube the inductive load is decoupled, allowing the electromagnetic field about the inductive load to collapse. The collapse of this energy field induces within the inductive load a counter EMF. This counter EMF creates a high positive potential across a second capacitor which, in turn, is induced into the second energy storage device or battery **40** as a charging current. The amount of charging current available to the battery 40 is dependent upon the initial conditions within the circuit at the time of discharge within the conversion switching element tube and the amount of mechanical energy consumed by the workload.

A spark-gap protection device **42** is included in the circuit to protect the inductive load and the rectifier elements from unduly large discharge currents. Should the potentials within the circuit exceed predetermined values, fixed by the mechanical size and spacing of the elements within the protective device, the excess energy is dissipated (bypassed) by the protective device to the circuit common (electrical ground).

Diodes **44** and **46** bypass the excess overshoot generated when the “Energy Conversion Switching Element Tube” is triggered. A switching element U allows either energy storage source to be used as the primary energy source, while the other battery is used as the energy retrieval unit. The switch facilitates interchanging the source and the retrieval unit at optimum intervals to be determined by the utilisation of the conversion switching element tube. This switching may be accomplished manually or automatically, as determined by the choice of switching element from among a large variety readily available for the purpose.

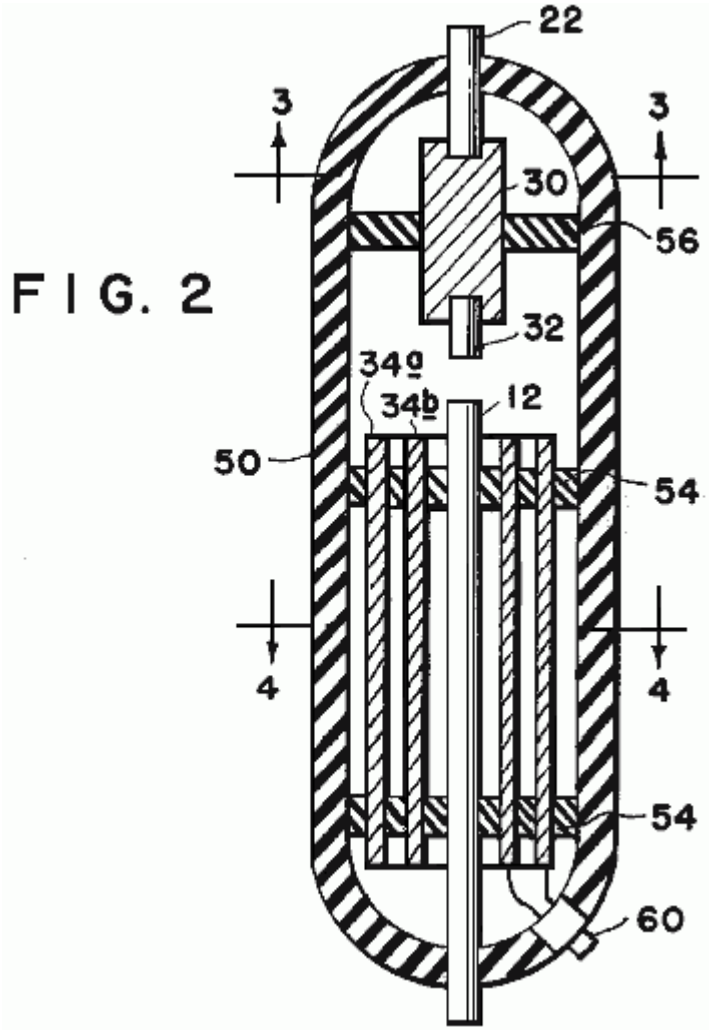


FIG. 2

FIG. 3

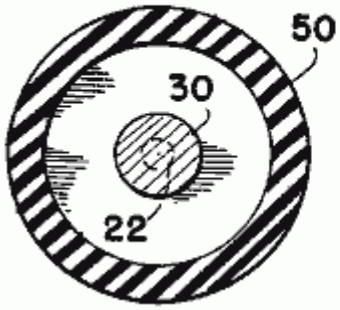


FIG. 4

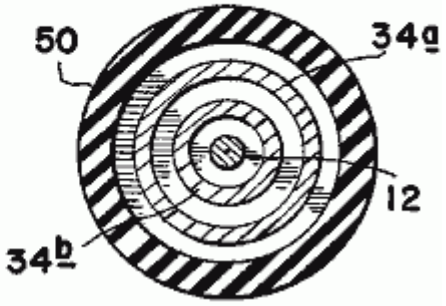


Fig.2, Fig.3, and **Fig.4** show the mechanical structure of the conversion switching element tube **14**. An outer housing **50** may be of any insulative material such as glass. The anodes **12** and **22** and grids **34a** and **34b** are firmly secured by nonconductive spacer material **54**, and **56**. The resistive element **30** may be introduced into the low-voltage anode path to control the peak currents through the conversion switching element tube. The resistive element may be of a piece, or it may be built of one or more resistive elements to achieve the desired result.

The anode material may be identical for each anode, or may be of differing materials for each anode, as dictated by the most efficient utilisation of the device, as determined by appropriate research at the time of production for the intended use. The shape and spacing of the electrostatic grids is also susceptible to variation with application (voltage, current, and energy requirements).

It is the contention of the inventor that by judicious mating of the elements of the conversion switching element tube, and the proper selection of the components of the circuit elements of the system, the desired theoretical results may be achieved. It is the inventor's contention that this mating and selection process is well within the capabilities of intensive research and development technique.

Let it be stated here that substituting a source of electric alternating-current subject to the required cur- rent and/or voltage shaping and/or timing, either prior to being considered a primary energy source, or there- after, should not be construed to change the described utilisation or application of primary energy in any way. Such energy conversion is readily achieved by any of a multitude of well established principles. The preferred embodiment of this invention merely assumes optimum utilisation and optimum benefit from this invention when used with portable energy devices similar in principle to the wet-cell or dry-cell battery.

This invention proposes to utilise the energy contained in an internally generated high-voltage electric spike (energy pulse) to electrically energise an inductive load.: this inductive load being then capable of converting the energy so supplied into a useful electrical or mechanical output.

In operation the high-voltage, short-duration electric spike is generated by discharging the capacitor **16** across the spark-gap in the conversion switching element tube. The necessary high-voltage potential is stored on the capacitor in incremental, additive steps from the bridge-type rectifier **24**. When the energy source is a direct-current electric energy storage device, such as the battery **12**, the input to the bridge rectifier is provided by the voltage step-up transformer **22**, that is in turn energised from the vibrator **20**, or solid-state chopper, or similar device to properly drive the transformer and rectifier circuits.

When the energy source is an alternating-current, switches **64** disconnect transformer **22** and the input to the bridge-type rectifier **24** is provided by the voltage step-up transformer **66**, that is in turn energised from the vibrator **20**, or solid-state chopper, or similar device to properly drive the transformer and rectifier circuits.

The repetitions output of the bridge rectifier incrementally increases the capacitor charge toward its maximum. This charge is electrically connected directly to the high-voltage anode **12** of the conversion switching element tube. When the low-voltage anode **32** is connected to a source of current, an arc is created in the spark-gap designated **62** of the conversion switching element tube equivalent to the potential stored on the high-voltage anode, and the current available from the low-voltage anode.

Because the duration of the arc is very short, the instantaneous voltage, and instantaneous current may both be very high. The instantaneous peak apparent power is therefore, also very high. Within the conversion switching element tube, this energy is absorbed by the grids **34a** and **34b** mounted circumferentially about the interior of the tube.

Control of the energy spike within the conversion switching element tube is accomplished by a mechanical, or

solid-state commutator, that closes the circuit path from the low-voltage anode to the current source at that moment when the delivery of energy to the output load is most auspicious. Any number of standard high-accuracy, variable setting devices are available for this purpose. When control of the repetitive rate of the system's output is required, it is accomplished by controlling the time of connection at the low-voltage anode.

Thus there can be provided an electrical driving system having a low-voltage source coupled to a vibrator, a transformer and a bridge-type rectifier to provide a high voltage pulsating signal to a first capacitor. Where a high-voltage source is otherwise available, it may be coupled direct to a bridge-type rectifier, causing a pulsating signal to a first capacitor. The capacitor in turn is coupled to a high-voltage anode of an electrical conversion switching element tube. The element also includes a low-voltage anode which in turn is connected to a voltage source by a commutator, a switching element tube, and a variable resistor. Mounted around the high-voltage anode is a charge receiving plate which in turn is coupled to an inductive load to transmit a high-voltage discharge from the element to the load. Also coupled to the load is a second capacitor for storing the back EMF created by the collapsing electrical field of the load when the current to the load is blocked. The second capacitor in turn is coupled to the voltage source.

Patent GB 2,282,708 12th April 1995 Inventors: Harold Aspden (UK) and Robert George Adams (NZ)

ELECTRICAL MOTOR / GENERATOR

This version of the patent has been re-worded in an attempt to make it easier to read and understand. It describes the design of a pulsed electromagnet / permanent magnet motor which is capable of a higher power output than it's own power input.

ABSTRACT

An electrodynamic motor-generator has a salient pole permanent magnet rotor interacting with salient stator poles to form a machine operating on the magnetic reluctance principle. The intrinsic ferromagnetic power of the magnets provides the drive torque by bringing the poles into register whilst current pulses demagnetise the stator poles as the poles separate. In as much as less power is needed for stator demagnetisation than is fed into the reluctance drive by the thermodynamic system powering the ferromagnetic state, the machine operates regeneratively by virtue of stator winding interconnection with unequal number of rotor and stator poles. A rotor construction is disclosed (**Fig.6** and **Fig.7**). The current pulse may be such as to cause repulsion of the rotor poles.

FIELD OF THE INVENTION

This invention relates to a form of electric motor which serves a generating function in that the machine can act regeneratively to develop output electrical power or can generate mechanical drive torque with unusually high efficiency in relation to electrical power input.

The field of invention is that of switched reluctance motors, meaning machines which have salient poles and operate by virtue of the mutual magnetic attraction and/or repulsion as between magnetised poles.

The invention particularly concerns a form of reluctance motor which incorporates permanent magnets to establish magnetic polarisation.

BACKGROUND OF THE INVENTION

There have been proposals in the past for machines in which the relative motion of magnets can in some way develop unusually strong force actions which are said to result in more power output than is supplied as electrical input.

By orthodox electrical engineering principles such suggestions have seemed to contradict accepted principles of physics, but it is becoming increasingly evident that conformity with the first law of thermodynamics allows a gain in the electromechanical power balance provided it is matched by a thermal cooling.

In this sense, one needs to extend the physical background of the cooling medium to include, not just the machine structure and the immediate ambient environment, but also the sub-quantum level of what is termed, in modern physics, the zero-point field. This is the field activity of the vacuum medium which exists in the space between atomic nuclei and atomic electrons and is the seat of the action which is that associated with the Planck constant. Energy is constantly being exchanged as between that activity and coextensive matter forms but normally these energy fluctuations preserve, on balance, an equilibrium condition so that this action passes unnoticed at the technology level.

Physicists are becoming more and more aware of the fact that, as with gravitation, so magnetism is a route by which we can gain access to the sea of energy that pervades the vacuum. Historically, the energy balance has been written in mathematical terms by assigning 'negative' potential to gravitation or magnetism. However, this is only a disguised way of saying that the vacuum field, suitably influenced by the gravitating mass of a body in the locality or by magnetism in a ferromagnet has both the capacity and an urge to shed energy.

Now, however, there is growing awareness of the technological energy generating potential of this field background and interest is developing in techniques for 'pumping' the coupling between matter and vacuum field to derive power from that hidden energy source. Such research may establish that this action will draw on the

2.7K cosmic background temperature of the space medium through which the Earth travels at some 400 km/s. The effect contemplated could well leave a cool 'vapour trail' in space as a machine delivering heat, or delivering a more useful electrical form of energy that will revert to heat, travels with body Earth through that space.

In pure physics terms, relevant background is of recent record in the August 1993 issue of Physical Review E, vol. 48, pp. 1562-1565 under the title: 'Extracting energy and heat from the vacuum', authored by D. C. Cole and H. E. Puthoff. Though the connection is not referenced in that paper, one of its author's presented experimental evidence on that theme at an April 1993 conference held in Denver USA. The plasma power generating device discussed at that conference was the subject of U. S. Patent No. 5,018,180, the inventor of record being K. R. Shoulders.

The invention, to be described below, operates by extracting energy from a magnetic system in a motor and the relevant scientific background to this technology can be appreciated from the teachings of E. B. Moullin, a Cambridge Professor of Electrical Engineering who was a President of the Institution of Electrical Engineers in U. K. That prior art will be described below as part of the explanation of the operation of the invention.

The invention presented here concerns specific structural design features of a machine adapted for robust operation, but these also have novelty and special merit in a functional operation. What is described is quite distinct from prior art proposals, one being a novel kind of motor proposed by Gareth Jones at a 1988 symposium held in Hull, Canada under the auspices of the Planetary Association for Clean Energy. Jones suggested the adaptation of an automobile alternator which generates three-phase AC for rectification and use as a power supply for the electrics in the automobile. This alternator has a permanent magnet rotor and Jones suggested that it could be used, with high efficiency gain and torque performance, by operating it as a motor with the three-phase winding circuit excited so as to promote strong repulsion between the magnet poles and the stator poles after the poles had come into register.

However, the Jones machine is not one exploiting the advantages of the invention to be described, because it is not strictly a reluctance motor having salient poles on both stator and rotor. The stator poles in the Jones machine are formed by the winding configuration in a slotted stator form, the many slots being uniformly distributed around the inner circumference of the stator and not constituting a pole system which lends itself to the magnetic flux actions to be described by reference to the E. B. Moullin experiment.

The Jones machine operates by generating a rotating stator field which, in a sense, pushes the rotor poles forward rather than pulling them in the manner seen in the normal synchronous motor. Accordingly, the Jones machine relies on the electric current excitation of the motor producing a field system which rotates smoothly but has a polarity pattern which is forced by the commutation control to keep behind the rotor poles in asserting a continuous repulsive drive.

Another prior art proposal which is distinguished from this invention is that of one of the applicants, H. Aspden, namely the subject of U.K. Patent No. 2,234,863 (counterpart U.S. Patent Serial No.4,975,608). Although this latter invention is concerned with extracting energy from the field by the same physical process as the subject invention, the technique for accessing that energy is not optimum in respect of the structure or method used. Whereas in this earlier disclosure, the switching of the reluctance drive excited the poles in their approach phase, the subject invention, in one of its aspects, offers distinct advantages by demagnetisation or reversal of magnetisation in the pole separation phase of operation.

There are unexpected advantages in the implementation proposed by the subject invention, inasmuch as recent research has confirmed that it requires less input power to switch off the mutual attraction across an air gap between a magnet and an electromagnet than it does to switch it on. Usually, in electromagnetism, a reversal symmetry is expected, arising from conventional teaching of the way forward and back magnetomotive forces govern the resulting flux in a magnetic circuit.

This will be further explained after describing the scope of the invention.

BRIEF DESCRIPTION OF THE INVENTION

According to one aspect of the invention, an electrodynamic motor/generator machine comprises a stator configured to provide a set of stator poles, a corresponding set of magnetising windings mounted on the stator pole set, a rotor having two sections each of which has a set of salient pole pieces, the rotor sections being axially spaced along the axis of rotation of the rotor, rotor magnetisation means disposed between the two rotor sections arranged to produce a unidirectional magnetic field which magnetically polarises the rotor poles, whereby the pole faces of one rotor section all have a north polarity and the pole faces of the other rotor section all have a south polarity and electric circuit connections between an electric current source and the stator magnetising windings arranged to regulate the operation of the machine by admitting current pulses for a duration determined according

to the angular position of the rotor, which pulses have a direction tending to oppose the polarisation induced in the stator by the rotor polarisation as stator and rotor poles separate from an in-register position, whereby the action of the rotor magnetisation means provides a reluctance motor drive force to bring stator and rotor poles into register and the action of the stator magnetisation windings opposes the counterpart reluctance braking effect as the poles separate.

According to a feature of the invention, the circuit connecting the electric current source and the stator magnetising windings is designed to deliver current pulses which are of sufficient strength and duration to provide demagnetisation of the stator poles as the stator and rotor poles separate from an in-register position.

In this regard it is noted that in order to suppress the reluctance drive torque or brake torque, depending upon whether poles are converging or separating, a certain amount of electrical power must be fed to the magnetising windings on the stator. In a sense these windings are really 'demagnetising windings' because the polarity of the circuit connections admit the pulse current in the demagnetising direction.

However, it is more usual to refer to windings on magnetic cores as 'magnetising windings' even though they can function as primary windings or secondary windings, the former serving the magnetisation function with input power and the latter serving a demagnetising function with return of power.

According to another feature of the invention, the circuit connecting the electric current source and the stator magnetising windings is designed to deliver current pulses which are of sufficient strength and duration to provide a reversal of magnetic flux direction in the stator poles as the stator and rotor poles separate from an in-register position, whereby to draw on power supplied from the electric current source to provide additional forward drive torque.

According to a further feature of the invention, the electric current source connected to a stator magnetising winding of a first stator pole comprises, at least partially, the electrical pulses induced in the stator magnetising winding of a different second stator pole, the stator pole set configuration in relation to the rotor pole set configuration being such that the first stator pole is coming into register with a rotor pole as the second stator pole separates from its in-register position with a rotor pole.

This means that the magnetising windings of two stator poles are connected so that both serve a 'demagnetising' function, one in resisting the magnetic action of the mutual attraction in pulling poles into register, an action which develops a current pulse output and one in absorbing this current pulse, again by resisting the magnetic inter-pole action to demagnetise the stator pole as its associated rotor pole separates.

In order to facilitate the function governed by this circuit connection between stator magnetising windings, a phase difference is needed and this is introduced by designing the machine to have a different number of poles in a set of stator poles from the number of rotor poles in each rotor section. Together with the dual rotor section feature, this has the additional merit of assuring a smoother torque action and reducing magnetic flux fluctuations and leakage effects which contribute substantially to machine efficiency.

Thus, according to another feature of the invention, the stator configuration provides pole pieces which are common to both rotor sections in the sense that when stator and rotor poles are in-register the stator pole pieces constitute bridging members for magnetic flux closure in a magnetic circuit including that of the rotor magnetisation means disposed between the two rotor sections.

Preferably, the number of poles in a set of stator poles and the number of rotor poles in each section do not share a common integer factor, the number of rotor poles in one rotor section is the same as that in the other rotor section and the number of poles in a stator set and the number of poles in a rotor section differs by one, with the pole faces being of sufficient angular width to assure that the magnetic flux produced by the rotor magnetisation means can find a circular magnetic flux closure route through the bridging path of a stator pole and through corresponding rotor poles for any angular position of the rotor.

It is also preferable from a design viewpoint for the stator pole faces of this invention to have an angular width that is no greater than half the angular width of a rotor pole and for the rotor sections to comprise circular steel laminations in which the rotor poles are formed as large teeth at the perimeter with the rotor magnetisation means comprising a magnetic core structure the end faces of which abut two assemblies of such laminations forming the two rotor sections.

According to a further feature of the invention, the rotor magnetisation means comprises at least one permanent magnet located with its polarisation axis parallel with the rotor axis. The motor-generator may include an apertured metal disc that is of a non-magnetisable substance mounted on a rotor shaft and positioned intermediate the two rotor sections, each aperture providing location for a permanent magnet, whereby the centrifugal forces acting on the permanent magnet as the rotor rotates are absorbed by the stresses set up in the

disc. Also, the rotor may be mounted on a shaft that is of a non-magnetisable substance, whereby to minimise magnetic leakage from the rotor magnetising means through that shaft.

According to another aspect of the invention, an electrodynamic motor-generator machine comprises a stator configured to provide a set of stator poles, a corresponding set of magnetising windings mounted on the stator pole set, a rotor having two sections each of which has a set of salient pole pieces, the rotor sections being axially spaced along the axis of rotation of the rotor, rotor magnetisation means incorporated in the rotor structure and arranged to polarise the rotor poles, whereby the pole faces of one rotor section all have a north polarity and the pole faces of the other rotor section all have a south polarity and electric circuit connections between an electric current source and the stator magnetising windings arranged to regulate the operation of the machine by admitting current pulses for a duration determined according to the angular position of the rotor, which pulses have a direction tending to oppose the polarisation induced in the stator by the rotor polarisation as stator and rotor poles separate from an in-register position, whereby the action of the rotor magnetisation means provides a reluctance motor drive force to bring stator and rotor poles into register and the action of the stator magnetisation windings opposes the counterpart reluctance braking effect as the poles separate.

According to a feature of this latter aspect of the invention, the electric current source connected to a stator magnetising winding of a first stator pole comprises, at least partially, the electrical pulses induced in the stator magnetising winding of a different second stator pole, the stator pole set configuration in relation to the rotor pole set configuration being such that the first stator pole is coming into register with a rotor pole as the second stator pole separates from its in-register position with a rotor pole.

BRIEF DESCRIPTION OF THE DRAWINGS

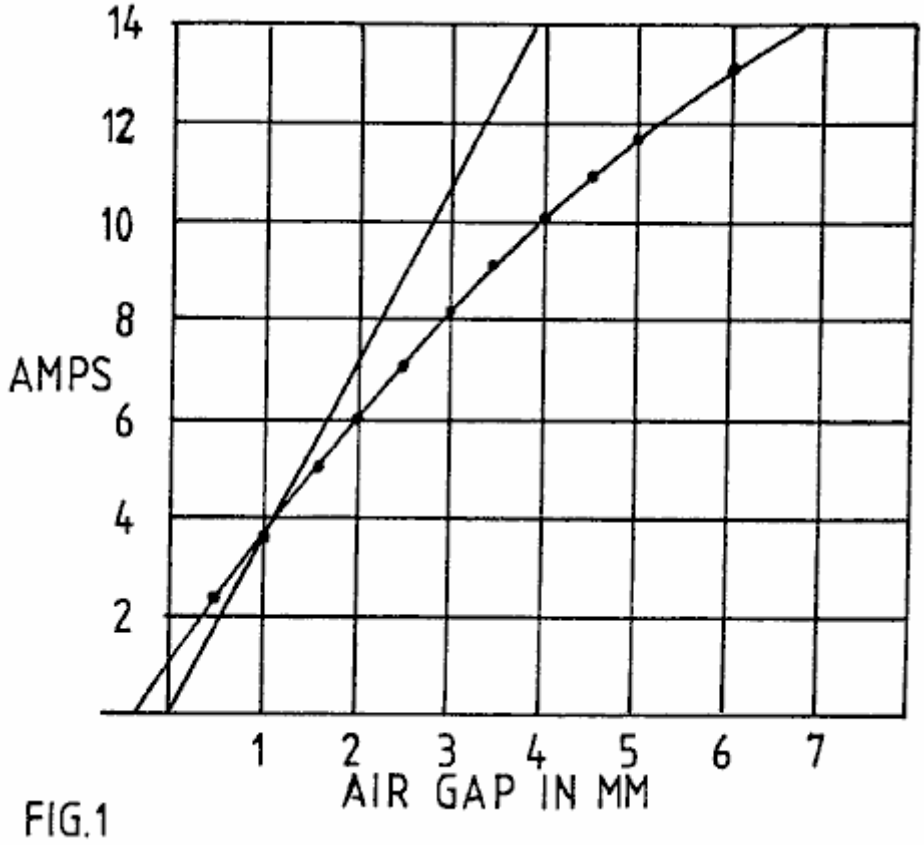


FIG.1

Fig.1 presents magnetic core test data showing how the volt-amp reactance power required to set up a constant magnetic flux action in an air gap, as assured by constant AC voltage excitation of a magnetising winding, falls short of the associated power of the potential implicit in the force action across that air gap.

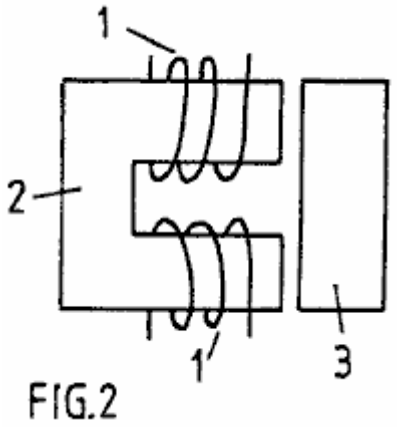


FIG.2

Fig.2 depicts the test structure to which Fig. 1 data applies.

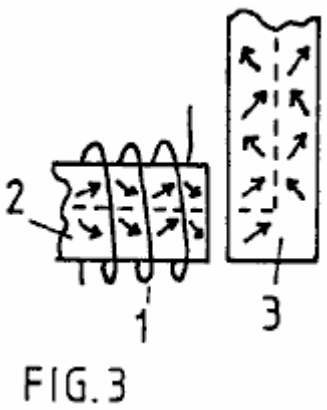


FIG.3

Fig.3 depicts the magnetisation action at work in causing magnetic flux to traverse an air-gap and turn a corner in a circuit through a magnetic core.

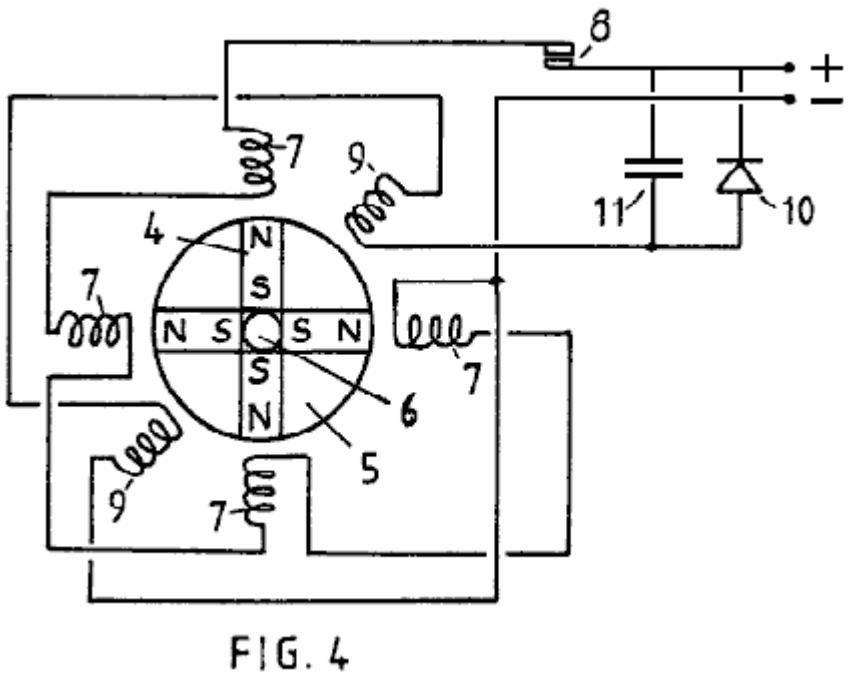


FIG. 4

Fig.4 shows the configuration of a test device used to prove the operating principles of the invention described.

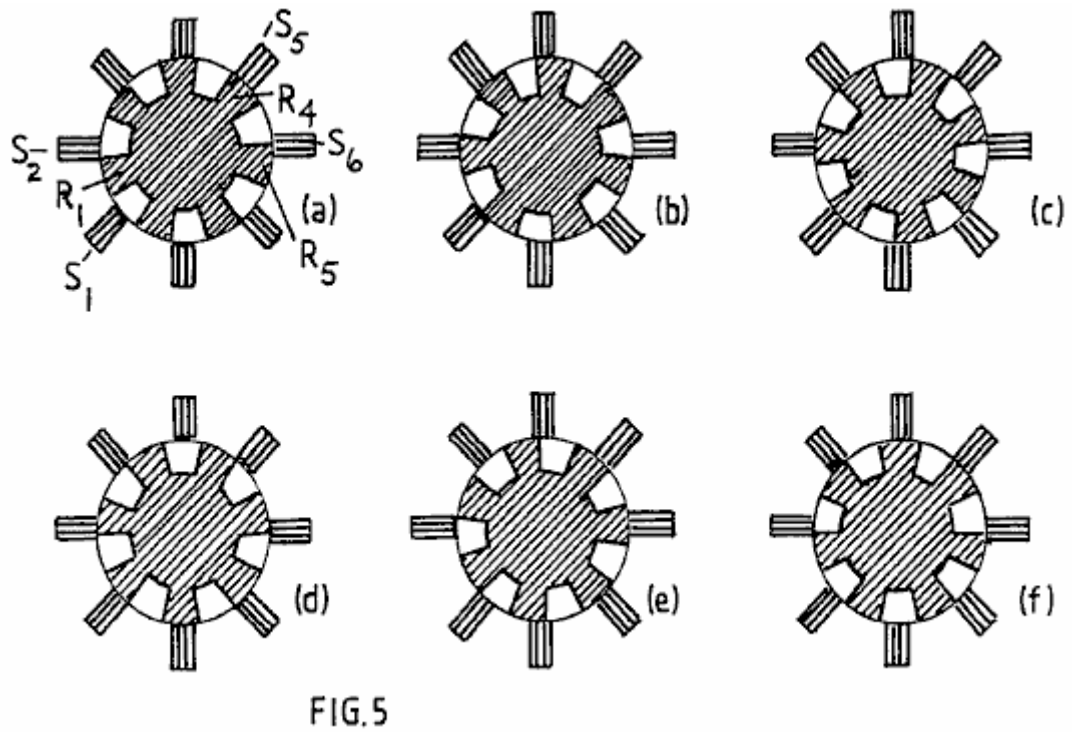


FIG.5

Fig.5 in its several illustrations depicts the progressive rotor pole to stator pole relationship as a rotor turns through a range of angular positions in a preferred embodiment of a machine according to the invention.

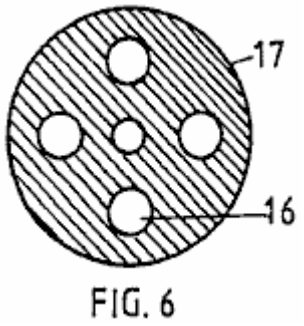


FIG. 6

Fig.6 shows the form of a disc member which provides location for four permanent magnets in the machine described.

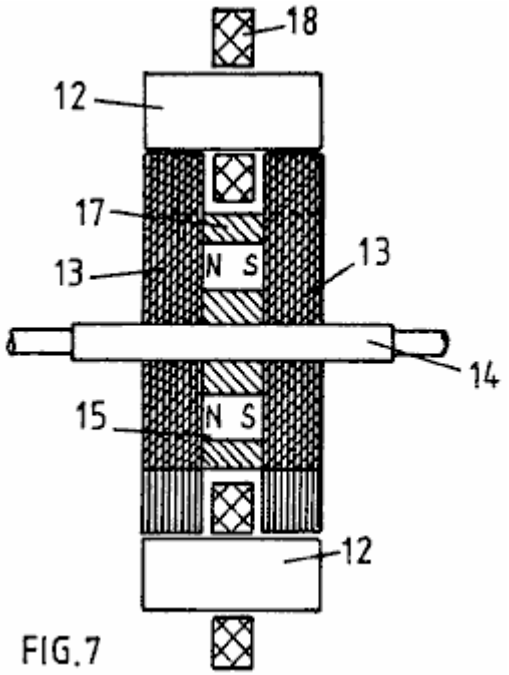


FIG.7

Fig.7 shows a cross-section of the magnetic circuit structure of a machine embodying the invention.

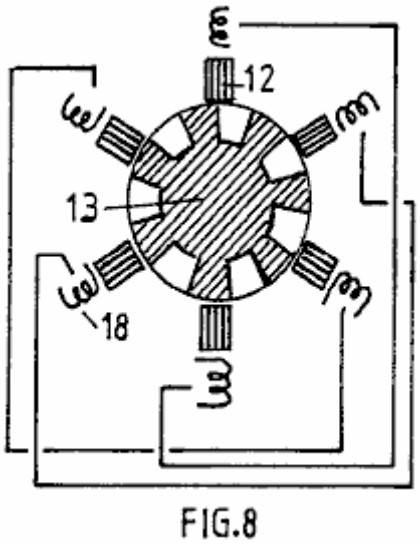


FIG.8

Fig.8 shows a six stator pole configuration with a seven pole rotor and depicts a schematic series connected linking of the magnetising windings of diametrically opposite stator poles.

DETAILED DESCRIPTION OF THE INVENTION

The fact that one can extract energy from the source which powers the intrinsic ferromagnetic state is not explicitly evident from existing textbooks, but it is implicit and, indeed, does become explicit once pointed out, in one textbook authored by E. B. Moullin. His book 'The Principles of Electromagnetism' published by Clarendon Press, Oxford (3rd Edition, 1955) describes on pages 168-174 an experiment concerned with the effect of air gaps between poles in a magnetic circuit. The data obtained are reproduced in **Fig.1**, where Professor Moullin shows a curve representing AC current input for different air gaps, given that the voltage supplied is constant. In the same figure, Moullin presents the theoretical current that would need to be applied to sustain the same voltage, and so the related pole forces across the air gap, assuming (a) no flux leakage and (b) that there is complete equality between inductive energy input and the mechanical energy potential for the magnetisation that is established in the air gap in a quarter-cycle period at the AC power excitation frequency.

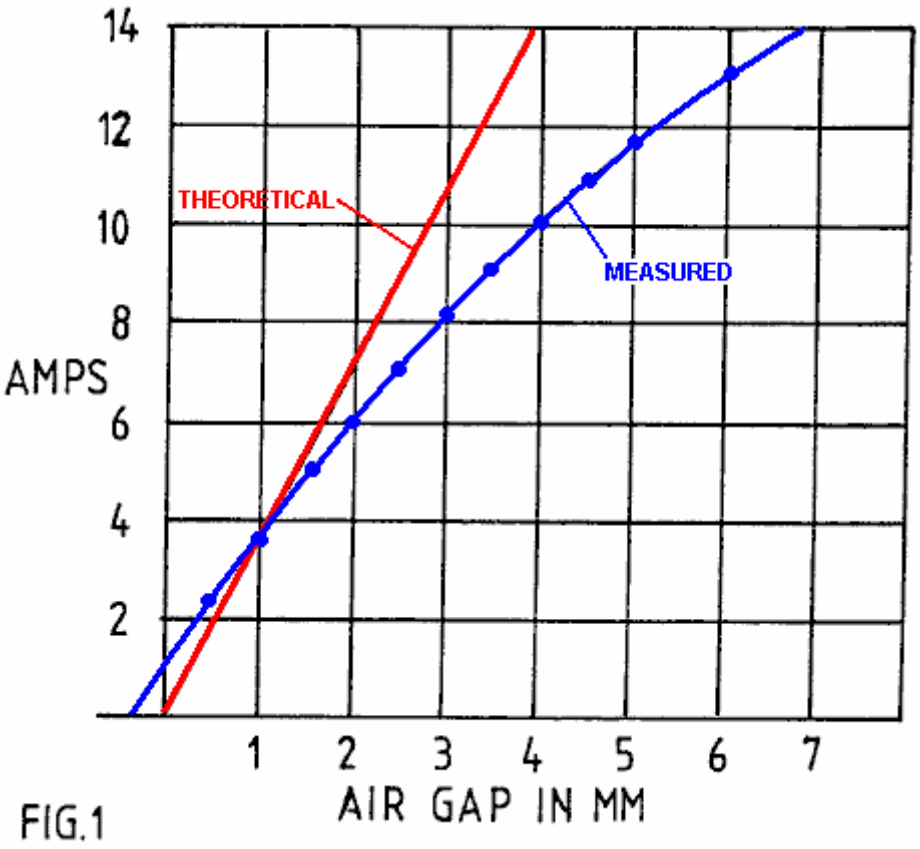


FIG.1

The data show that, even though the level of magnetic polarisation is well below the saturation value, being confined to a range that is regarded as the linear permeability range in transformer design, there is a clear drop-off of current, and so the volt-amp reactive power input needed, as current increases, compared with that predicted by the mechanical potential built up in the air gaps. Unless leakage flux is excessive, here was clear evidence of anomalous energy activity.

Moullin discusses the leakage flux inferred by this experiment but points out that there is considerable mystery in why the effect of a small gap, which should certainly not result in much flux leakage in the gap region, nevertheless has an enormous effect in causing what has to be substantial leakage in the light of the energy discrepancy. Moullin did not contemplate that energy had been fed in from the zero-point field system and so he left the issue with the statement that it was virtually impossible to predict leakage flux by calculation.

He was, of course, aware of magnetic domain structure and his argument was that the leakage flux problem was connected with what he termed a 'yawing' action of the flux as it passes around the magnetic circuit. Normally, provided the level of polarisation is below the knee of the B-H curve, which occurs at about 70% of saturation in iron cores of general crystal composition, it requires very little magnetising field to change the magnetic flux density. This is assuming that every effort is made to avoid air gaps. The action involves domain wall movements so that the magnetic states of adjacent domains switch to different crystal axes of easy magnetisation and this involves very little energy change.

However, if there is an air gap ahead in the flux circuit and the magnetising winding is not sitting on that air gap, the iron core itself has to be the seat of a progressive field source linking the winding and the gap. It can only serve in that sense by virtue of the lines of flux in the domains being forced to rotate somewhat from the preferred easy axes of magnetisation, with the help of the boundary surfaces around the whole core. This action means that, forcibly, and consequential upon the existence of the air gap, the flux must be carried through the core by

that 'yawing' action. It means that substantial energy is needed to force the establishment of those fields within the iron core. More important, however, from the point of view of this invention, it means that the intrinsic magnetic polarisation effects in adjacent magnetic domains in the iron cease to be mutually parallel or orthogonal so as to stay directed along axes of easy magnetisation. Then, in effect, the magnetising action is not just that of the magnetising winding wrapped around the core but becomes also that of adjacent ferromagnetic polarisation as the latter act in concert as vacuum-energy powered solenoids and are deflected into one another to develop the additional forward magnetomotive forces.

The consequences of this are that the intrinsic ferromagnetic power source with its thermodynamic ordering action contributes to doing work in building up forces across the air gap. The task, in technological terms, is then to harness that energy as the gap is closed, as by poles coming together in a reluctance motor, and avoid returning that energy as the poles separate, this being possible if the controlling source of primary magnetisation is well removed from the pole gap and the demagnetisation occurs when the poles are at the closest position.

This energy situation is evident in the Moullin data, because the constant AC voltage implies a constant flux amplitude across the air gap if there is no flux leakage in the gap region. A constant flux amplitude implies a constant force between the poles and so the gap width in relation to this force is a measure of the mechanical energy potential of the air gap. The reactive volt-amp power assessment over the quarter-cycle period representing the polarisation demand can then be compared with the mechanical energy so made available. As already stated, this is how Moullin deduced the theoretical current curve. In fact, as his data show, he needed less current than the mechanical energy suggested and so he had in his experiment evidence of the vacuum energy source that passed unnoticed and is only now revealing itself in machines that can serve our energy needs.

In the research leading to this patent application the Moullin experiment has been repeated to verify a condition where a single magnetising winding serves three air gaps. The Moullin test configuration is shown in **Fig.2**, but in repeating the experiment in the research leading to this invention, a search coil was mounted on the bridging member and this was used to compare the ratio of the voltage applied to the magnetising winding and that induced in the search coil.

The same fall-off feature in current demand was observed, and there was clear evidence of substantial excess energy in the air gap. This was in addition to the inductive energy that necessarily had to be locked into the magnetic core to sustain the 'yawing' action of the magnetic flux already mentioned.

It is therefore emphasised that, in priming the flux 'yawing' action, energy is stored inductively in the magnetic core, even though this has been deemed to be the energy of flux leakage outside the core. The air gap energy is also induction energy. Both energies are returned to the source winding when the system is demagnetised, given a fixed air gap.

If, however, the air gap closes after or during magnetisation, much of that inductive energy goes into the mechanical work output. Note then that the energy released as mechanical work is not just that stored in the air gap but is that stored in sustaining the 'yaw'. Here, then is reason to expect an even stronger contribution to the dynamic machine performance, one that was not embraced by the calculation of the steady-state situation.

Given the above explanation of the energy source, the structural features which are the subject of this invention will now be described.

The 'yawing' action is depicted in **Fig.3**, which depicts how magnetic flux navigates a right-angled bend in a magnetic core upon passage through an air gap. By over-simplification it is assumed that the core has a crystal structure that has a preferred axis of magnetisation along the broken line path. With no air gap, the current needed by a magnetising winding has only to provide enough magnetomotive force to overcome the effects of non-magnetic inclusions and impurities in the core substance and very high magnetic permeabilities can apply. However, as soon as the air gap develops, this core substance has to find a way of setting up magnetomotive force in regions extending away from the locality of the magnetising winding. It cannot do this unless its effect is so powerful that the magnetic flux throughout the magnetic circuit through the core substance is everywhere deflected from alignment with a preferred easy axis of magnetisation. Hence the flux vectors depicted by the arrows move out of alignment with the broken line shown.

There is a 'knock-on' effect progressing all the way around the core from the seat of the magnetising winding and, as already stated, this harnesses the intrinsic ferromagnetic power that, in a system with no air gap, could only be affected by magnetisation above the knee of the B-H curve. Magnetic flux rotation occurs above that knee, whereas in an ideal core the magnetism develops with very high permeability over a range up to that knee, because it needs very little power to displace a magnetic domain wall sideways and promote a 900 or a1800 flux reversal. Indeed, one can have a magnetic permeability of 10,000 below the knee and 100 above the knee, the latter reducing progressively until the substance saturates magnetically.

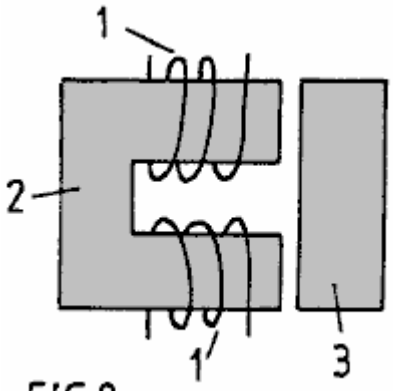


FIG.2

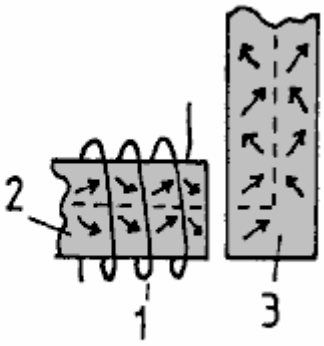


FIG.3

In the situation depicted in **Fig.2** and **Fig.3** the field strength developed by the magnetising windings **1** on magnetic core **2** has to be higher, the greater the air gap, in order to achieve the same amount of magnetisation as measured by the voltage induced in a winding (not shown) on the bridging member **3**. However, by virtue of that air gap there is potential for harnessing energy supplied to that air gap by the intrinsic zero-point field that accounts for the magnetic permeability being over unity and here one can contemplate very substantial excess energy potential, given incorporation in a machine design which departs from convention.

One of the applicants has built an operative test machine which is configured as depicted schematically in **Fig.4**. The machine has been proved to deliver substantially more mechanical power output than is supplied as electrical input, as much as a ratio of 7:1 in one version, and it can act regeneratively to produce electrical power.

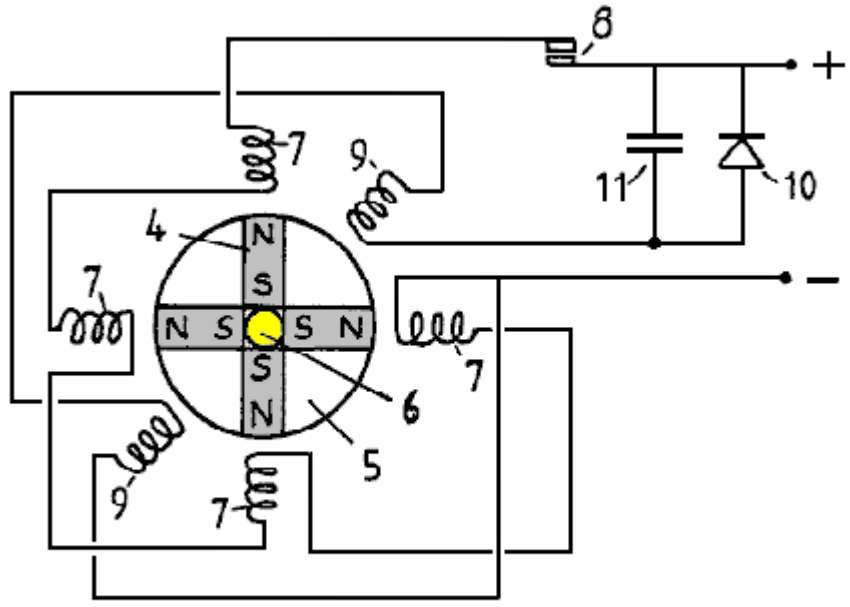


FIG. 4

What is shown in **Fig.4** is a simple model designed to demonstrate the principle of operation. It comprises a rotor in which four permanent magnets **4** are arrayed to form four poles. The magnets are bonded into four sectors of a non-magnetic disc **5** using a high density polyurethane foam filler and the composite disc is then assembled on a brass spindle **6** between a split flange coupling. Not shown in the figure is the structure holding the spindle vertically in bearings or the star wheel commutator assembly attached to the upper shaft of the spindle.

Note that the magnets present north poles at the perimeter of the rotor disc and that the south poles are held together by being firmly set in the bonding material. A series of four stator poles were formed using magnetic cores from standard electromagnetic relays are were positioned around the rotor disc as shown. The magnetising windings **7** on these cores are shown to be connected in series and powered through commutator contacts **8** by a DC power supply. Two further stator cores formed by similar electromagnetic relay components are depicted by their windings **9** in the intermediate angle positions shown and these are connected in series and connected to a rectifier **10** bridged by a capacitor **11**.

The rotor spindle **6** is coupled with a mechanical drive (not shown) which harnesses the torque developed by the motor thus formed and serves as a means for measuring output mechanical power delivered by the machine.

In operation, assuming that the rotor poles are held initially off-register with the corresponding stator poles and the hold is then released, the strong magnetic field action of the permanent magnets will turn the rotor to bring the stator and rotor poles into register. A permanent magnet has a strong attraction for soft iron and so this initial impulse of rotation is powered by the potential energy of the magnets.

Now, with the rotor acting as a flywheel and having inertia it will have a tendency to over-shoot the in-register pole position and that will involve a reverse attraction with the result that the rotor will oscillate until damping action brings it to rest. However, if the contacts of the commutating switch are closed as the poles come

The commutating switch **8** needs only to be closed for a limited period of angular travel following the top dead centre in-register position of the stator and rotor poles. The power supplied through that switch by those pulses will cause the rotor to continue rotating and high speeds will be achieved as the machine develops its full motor function.

Tests on such a machine have shown that more mechanical power can be delivered than is supplied electrically by the source powering the action through the commutating switch. The reason for this is that, whereas the energy in the air gap between rotor and stator poles which is tapped mechanically as the poles come into register is provided by the intrinsic power of the ferromagnet, a demagnetising winding on the part of the core system coupled across that air gap needs very little power to eliminate the mechanical force acting across that air gap. Imagine such a winding on the bridging member shown in **Fig.2**. The action of current in that winding, which sits astride the 'yawing' flux in that bridging member well removed from the source action of the magnetising windings **1**, is placed to be extremely effective in resisting the magnetising influence communicated from a distance. Hence very little power is needed to overcome the magnetic coupling transmitted across the air gap.

Although the mutual inductance between two spaced-apart magnetising windings has a reciprocal action, regardless of which winding is primary and which is secondary, the action in the particular machine situation being described involves the 'solenoidal' contribution represented by the 'yawing' ferromagnetic flux action. The latter is not reciprocal inasmuch as the flux ‘yaw' depends on the geometry of the system. A magnetising winding directing flux directly across an air gap has a different influence on the action in the ferromagnetic core from one directing flux lateral to the air gap and there is no reciprocity in this action.

In any event, the facts of experiment do reveal that, owing to a significant discrepancy in such mutual interaction, more mechanical power is fed into the rotor than is supplied as input from the electrical source.

This has been further demonstrated by using the two stator windings **9** to respond in a generator sense to the passage of the rotor poles. An electrical pulse is induced in each winding by the passage of a rotor pole and this is powered by the inertia of the rotor disc **5**. By connecting the power so generated, to charge the capacitor **11,** the DC power supply can be augmented to enhance the efficiency even further.

Indeed, the machine is able to demonstrate the excess power delivery from the ferromagnetic system by virtue of electrical power generation charging a battery at a greater rate than a supply battery is discharged.

This invention is concerned with a practical embodiment of the motor-generator principles just described and aims, in its preferred aspect, to provide a robust and reliable machine in which the tooth stresses in the rotor poles, which are fluctuating stresses communicating high reluctance drive torque, are not absorbed by a ceramic permanent magnet liable to rupture owing to its brittle composition.

Another object is to provide a structure which can be dismantled and reassembled easily to replace the permanent magnets, but an even more important object is that of minimising the stray leakage flux oscillations from the powerful permanent magnets. Their rotation in the device depicted in **Fig.4** would cause excessive eddy-current induction in nearby metal, including that of the machine itself, and such effects are minimised if the flux changes are confined to paths through steel laminations and if the source flux from the magnets has a symmetry or near symmetry about the axis of rotation.

Thus, the ideal design with this in mind is one where the permanent magnet is a hollow cylinder located on a non-magnetic rotor shaft, but, though that structure is within the scope of this invention, the machine described will utilise several separate permanent magnets approximating, in function, such a cylindrical configuration.

Referring to **Fig.4**, it will further be noted that the magnetic flux emerging from the north poles will have to find its way along leakage paths through air to re-enter the south poles. For periods in each cycle of machine operation

the flux will be attracted through the stator cores, but the passage through air is essential and so the power of the magnets is not used to full advantage and there are those unwanted eddy-current effects.

To overcome this problem the invention provides for two separate rotor sections and the stator poles become bridging members, which with optimum design, allow the flux from the magnets to find a route around a magnetic circuit with minimal leakage through air as the flux is directed through one or other pairs of air gaps where the torque action is developed.

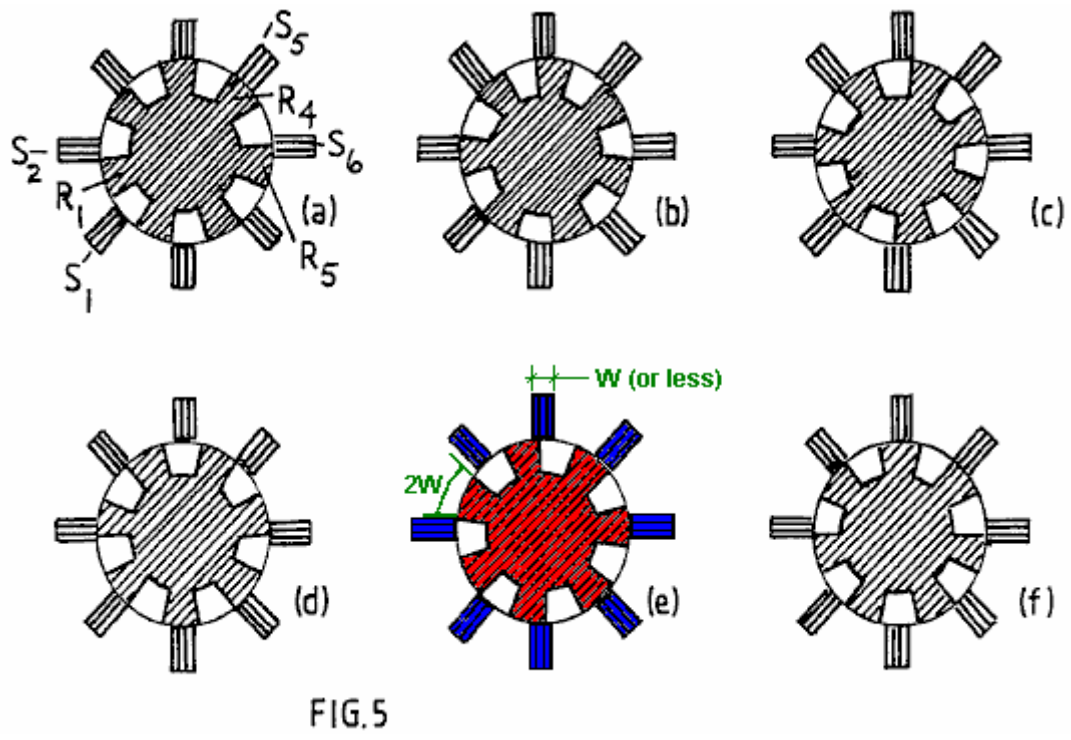


FIG.5

Reference is now made to **Fig.5** and the sequence of rotor positions shown. Note that the stator pole width can be significantly smaller that that of the rotor poles. Indeed, for operation using the principles of this invention, it is advantageous for the stator to have a much smaller pole width so as to concentrate the effective pole region. A stator pole width of half that of the rotor is appropriate but it may be even smaller and this has the secondary advantage of requiring smaller magnetising windings and so saving on the loss associated with the current circuit.

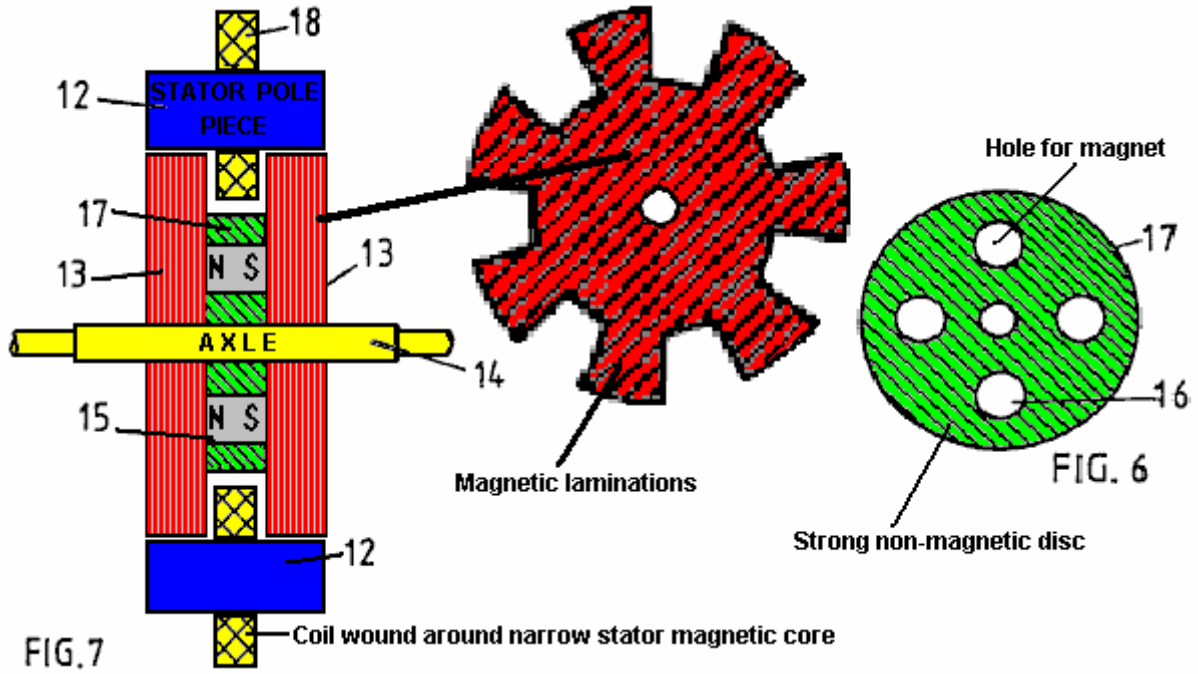


FIG.7

FIG. 6

The stator has eight pole pieces formed as bridging members **12**, more clearly represented in **Fig.7**, which shows a sectional side view through two rotor sections **13** axially spaced on a rotor shaft **14**. There are four permanent magnets **15** positioned between these rotor sections and located in apertures **16** in a disc **17** of a non-magnetic

substance of high tensile strength, the latter being shown in **Fig.6**. The rotor sections are formed from disc laminations of electrical steel which has seven large teeth, the salient poles. Magnetising windings **18** mounted on the bridging members **12** constitute the system governing the action of the motor-generator being described.

The control circuitry is not described as design of such circuitry involves ordinary skill possessed by those involved in the electrical engineering art.

It suffices, therefore, to describe the merits of the structural design configuration of the core elements of the machine. These concern principally the magnetic action and, as can be imagined from **Fig.7**, the magnetic flux from the magnets enters the rotor laminations by traversing the planar faces of the laminations and being deflected into the plane of the laminations to pass through one or other of the stator pole bridging members, returning by a similar route through the other rotor.

By using eight stator poles and seven rotor poles, the latter having a pole width equal to half the pole pitch in an angular sense, it will be seen from **Fig.5**, that there is always a flux passage across the small air gap between stator and rotor poles. However, as one pole combination is in-register the diametrically-opposed pole combinations are out-of register.

As described by reference to **Fig.4** the operation of the machine involves allowing the magnet to pull stator and rotor poles into register and then, as they separate, pulsing the winding on the relevant stator member to demagnetise that member. In the **Fig.4** system, all the stator magnetising windings were pulsed together, which is not an optimum way in which to drive a multi-pole machine.

In the machine having the pole structure with one less rotor pole than stator poles (or an equivalent design in which there is one less stator pole than rotor poles) this pulsing action can be distributed in its demand on the power supply, and though this makes the commutation switch circuit more expensive the resulting benefit outweighs that cost. However, there is a feature of this invention by which that problem can be alleviated if not eliminated.

Suppose that the rotor has the position shown in **Fig.5(a)** with the rotor pole denoted **R1** midway between stator poles **S1** and **S2** and imagine that this is attracted towards the in-register position with stator pole **S2**. Upon reaching that in-register position, as shown in **Fig.5(c)**, suppose that the magnetising winding of stator pole **S2** is excited by a current pulse which is sustained until the rotor reaches the **Fig.5(e)** position.

The combination of these two actions will have imparted a forward drive impulse powered by the permanent magnet in the rotor structure and the current pulse which suppresses braking action will have drawn a smaller amount of energy from the electrical power source which supplies it. This is the same process as was described by reference to **Fig.4**.

However, now consider the events occurring in the rotor action diametrically opposite that just described. In the **Fig.5(a)** position rotor pole **R4** has come fully into register with stator pole **S5** and so stator pole **S5** is ready to be demagnetised. However, the magnetic coupling between the rotor and stator poles is then at its strongest. Note, however, that in that **Fig.5(a)** position **R5** is beginning its separation from stator poles and the magnetising winding of stator pole **S6** must then begin draw power to initiate demagnetisation. During that following period of pole separation the power from the magnet is pulling **R1** and **S2** together with much more action than is needed to generate that current pulse needed to demagnetise **S6**. It follows, therefore, that, based on the research findings of the regenerative excitation in the test system of **Fig.4**, the series connection of the magnetising windings on stators **S2** and **S6** will, without needing any commutative switching, provide the regenerative power needed for machine operation.

The complementary action of the two magnetising windings during the pole closure and pole separation allows the construction of a machine which, given that the zero-point vacuum energy powering the ferromagnet is feeding input power, will run on that source of energy and thereby cool the sustaining field system.

There are various design options in implementing what has just been proposed. Much depends upon the intended use of the machine. If it is intended to deliver mechanical power output the regenerative electrical power action can all be used to power the demagnetisation with any surplus contributing to a stronger drive torque by reversing the polarity of the stator poles during pole separation.

If the object is to generate electricity by operating in generator mode then one could design a machine having additional windings on the stator for delivering electrical power output. However, it seems preferable to regard the machine as a motor and maximise its efficiency in that capacity whilst using a mechanical coupling to an alternator of conventional design for the electrical power generation function.

In the latter case it would still seem preferable to use the self-excitation feature already described to reduce commutation switching problems.

The question of providing for machine start-up can be addressed by using a separate starter motor powered from an external supply or by providing for current pulsing limited to, say, two stator poles. Thus, for example, with the eight stator pole configuration, the cross-connected magnetising windings could be limited to three stator pairs, with two stator magnetising windings left free for connection to a pulsed external supply source.

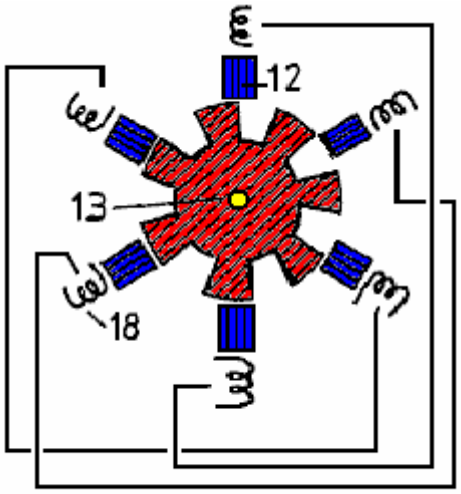


FIG.8

If the latter feature were not required, then the stator magnetising windings would all be connected in pairs on a truly diametrically opposite basis. Thus **Fig.8** shows a rotor-stator configuration having six stator poles interacting with seven rotor poles and stator magnetising windings linked together in pairs.

The invention, therefore, offers a wide range of implementation possibilities, which, in the light of this disclosure will become obvious to persons skilled in the electrical engineering art, all based, however, on the essential but simple principle that a rotor has a set of poles of common polarity which are attracted into register with a set of stator poles that are suppressed or reversed in polarity magnetically during pole separation. The invention, however, also offers the important feature of minimising commutation and providing further for a magnetic flux closure that minimises the leakage flux and fluctuations of leakage flux and so contributes to efficiency and high torque performance as well as durability and reliability of a machine incorporating the invention.

It is noted that although a machine has been described which uses two rotor sections it is possible to build a composite version of the machine having several rotor sections. In the eventuality that the invention finds use in very large motor-generator machines the problem of providing very large magnets can be overcome by a design in which numerous small magnets are assembled. The structural concept described by reference to **Fig.6** in providing locating apertures to house the magnets makes this proposal highly feasible. Furthermore, it is possible to replace the magnets by a steel cylinder and provide a solenoid as part of the stator structure and located between the rotor sections. This would set up an axial magnetic field magnetising the steel cylinder and so polarising the rotor. However, the power supplied to that solenoid would detract from the power generated and so such a machine would not be as effective as the use of permanent magnets such as are now available.

Nevertheless, should one see significant progress in the development of warm superconductor materials, it may become feasible to harness the self-generating motor-generator features of the invention, with its self-cooling properties, by operating the device in an enclosure at low temperatures and replacing the magnets by a superconductive stator supported solenoid.

CLAIMS

1. An electrodynamic motor-generator machine comprising a stator configured to provide a set of stator poles, a corresponding set of magnetising windings mounted on the stator pole set, a rotor having two sections each of which has a set of salient pole pieces, the rotor sections being axially spaced along the axis of rotation of the rotor, rotor magnetisation means disposed between the two rotor sections arranged to produce a unidirectional magnetic field which magnetically polarises the rotor poles, whereby the pole faces of one rotor section all have a north polarity and the pole faces of the other rotor section all have a south polarity and electric circuit connections between an electric current source and the stator magnetising windings arranged to regulate the operation of the machine by admitting current pulses for a duration determined according to the angular position of the rotor, which pulses have a direction tending to oppose the polarisation induced in the stator by

the rotor polarisation as stator and rotor poles separate from an in-register position, whereby the action of the rotor magnetisation means provides a reluctance motor drive force to bring stator and rotor poles into register and the action of the stator magnetisation windings opposes the counterpart reluctance braking effect as the poles separate.

2. A motor-generator according to claim 1, wherein the circuit connecting the electric current source and the stator magnetising windings is designed to deliver current pulses which are of sufficient strength and duration to provide demagnetisation of the stator poles as the stator and rotor poles separate from an in-register position.

3. A motor-generator according to claim 1, wherein the circuit connecting the electric current source and the stator magnetising windings is designed to deliver current pulses which are of sufficient strength and duration to provide a reversal of magnetic flux direction in the stator poles as the stator and rotor poles separate from an in-register position, whereby to draw on power supplied from the electric current source to provide additional forward drive torque.

4. A motor-generator according to claim 1, wherein the electric current source connected to a stator magnetising winding of a first stator pole comprises, at least partially, the electrical pulses induced in the stator magnetising winding of a different second stator pole, the stator pole set configuration in relation to the rotor pole set configuration being such that the first stator pole is coming into register with a rotor pole as the second stator pole separates from its in-register position with a rotor pole.

5. A motor-generator according to claim 1, wherein the number of poles in a set of stator poles is different from the number of rotor poles in each rotor section.

6. A motor-generator according to claim 1, wherein the stator configuration provides pole pieces which are common to both rotor sections in the sense that when stator and rotor poles are in-register the stator pole pieces constitute bridging members for magnetic flux closure in a magnetic circuit including that of the rotor magnetisation means disposed between the two rotor sections.

7. A motor-generator according to claim 6, wherein the number of poles in a set of stator poles and the number of rotor poles in each section do not share a common integer factor and the number of rotor poles in one rotor section is the same as that in the other rotor section.

8. A motor-generator according to claim 7, wherein the number of poles in a stator set and the number of poles in a rotor section differs by one and the pole faces are of sufficient angular width to assure that the magnetic flux produced by the rotor magnetisation means can find a circuital magnetic flux closure route through the bridging path of a stator pole and through corresponding rotor poles for any angular position of the rotor.

9. A motor-generator according to claim 8, wherein each rotor section comprises seven poles.

10. A motor-generator according to claim 7, wherein there are N rotor poles in each rotor section and each has an angular width that is 180/N degree of angle.

11. A motor-generator according to claim 7, wherein the stator pole faces have an angular width that is no greater than half the angular width of a rotor pole.

12. A motor-generator according to claim 1, wherein the rotor sections comprise circular steel laminations in which the rotor poles are formed as large teeth at the perimeter, and the rotor magnetisation means comprise a magnetic core structure the end faces of which abut two assemblies of such laminations forming the two rotor sections.

13. A motor-generator according to claim 1 in which the rotor magnetisation means comprises at least one permanent magnet located with its polarisation axis parallel with the rotor axis.

14. A motor-generator according to claim 13, wherein an apertured metal disc that is of a non-magnetisable substance is mounted on a rotor shaft and positioned intermediate the two rotor sections and each aperture provides location for a permanent magnet, whereby the centrifugal forces acting on the permanent magnet as the rotor rotates are absorbed by the stresses set up in the disc.

15. A motor-generator according to claim 1, having a rotor mounted on a shaft that is of a non-magnetisable substance, whereby to minimise magnetic leakage from the rotor magnetising means.

16. An electrodynamic motor-generator machine comprising a stator configured to provide a set of stator poles, a corresponding set of magnetising windings mounted on the stator pole set, a rotor having two sections each of which has a set of salient pole pieces, the rotor sections being axially spaced along the axis of rotation of the

rotor, rotor magnetisation means incorporated in the rotor structure and arranged to polarise the rotor poles, whereby the pole faces of one rotor section all have a north polarity and the pole faces of the other rotor section all have a south polarity and electric circuit connections between an electric current source and the stator magnetising windings arranged to regulate the operation of the machine by admitting current pulses for a duration determined according to the angular position of the rotor, which pulses have a direction tending to oppose the polarisation induced in the stator by the rotor polarisation as stator and rotor poles separate from an in-register position, whereby the action of the rotor magnetisation means provides a reluctance motor drive force to bring stator and rotor poles into register and the action of the stator magnetisation windings opposes the counterpart reluctance braking effect as the poles separate.

17. A motor-generator according to claim 16, wherein the electric current source connected to a stator magnetising winding of a first stator pole comprises, at least partially, the electrical pulses induced in the stator magnetising winding of a different second stator pole, the stator pole set configuration in relation to the rotor pole set configuration being such that the first stator pole is coming into register with a rotor pole as the second stator pole separates from its in-register position with a rotor pole.

Amendments to the claims have been filed as follows 1. An electrodynamic motor-generator machine comprising a stator configured to provide a set of stator poles, a corresponding set of magnetising windings mounted on the stator pole set, a rotor having two sections each of which has a set of salient pole pieces, the rotor sections being axially spaced along the axis of rotation of the rotor, rotor magnetisation means disposed between the two rotor sections arranged to produce a unidirectional magnetic field which magnetically polarises the rotor poles, whereby the pole faces of one rotor section all have a north polarity and the pole faces of the other rotor section all have a south polarity and electric circuit connections between an electric current source and the stator magnetising windings arranged to regulate the operation of the machine by admitting current pulses for a duration determined according to the angular position of the rotor, which pulses have a direction tending to oppose the polarisation induced in the stator by the rotor polarisation as stator and rotor poles separate from an in-register position, whereby the action of the rotor magnetisation means provides a reluctance motor drive force to bring stator and rotor poles into register and the action of the stator magnetisation windings opposes the counterpart reluctance braking effect as the poles separate, the machine being characterised in that the stator comprises separate ferromagnetic bridging members mounted parallel with the rotor axis, the ends of which constitute stator poles and the core sections of which provide closure paths operative when the stator and rotor poles are in register to confine magnetic flux developed by the rotor magnetisation means to a stator flux path of restricted cross-section disposed anti-parallel with the unidirectional magnetic field polarisation axis of the rotor magnetising means 2. A motor-generator according to claim 1, wherein the circuit connecting the electric current source and the stator magnetising windings is designed to deliver current pulses which are of sufficient strength and duration to provide demagnetisation of the stator poles as the stator and rotor poles separate from an in-register position.

3. A motor-generator according to claim 1, wherein the circuit connecting the electric current source and the stator magnetising windings is designed to deliver current pulses which are of sufficient strength and duration to provide a reversal of magnetic flux direction in the stator poles as the stator and rotor poles separate from an in-register position, whereby to draw on power supplied from the electric current source to provide additional forward drive torque.

4. A motor-generator according to claim 1, wherein the electric current source connected to a stator magnetising winding of a first stator pole comprises, at least partially, the electrical pulses induced in the stator magnetising winding of a different second stator pole, the stator pole set configuration in relation to the rotor pole set configuration being such that the first stator pole is coming into register with a rotor pole as the second stator pole separates from its in-register position with a rotor pole.

5. A motor-generator according to claim 1, wherein the number of poles in a set of stator poles is different from the number of rotor poles in each rotor section.

6. A motor-generator according to claim 1, wherein the stator configuration provides pole pieces which are common to both rotor sections in the sense that when stator and rotor poles are in-register the stator pole pieces constitute bridging members for magnetic flux closure in a magnetic circuit including that of the rotor magnetisation means disposed between the two rotor sections.

7. A motor-generator according to claim 6, wherein the number of poles in a set of stator poles and the number of rotor poles in each section do not share a common integer factor and the number of rotor poles in one rotor section is the same as that in the other rotor section.

WILLIAM BARBAT: POWER GENERATOR

Patent Application US 2007/0007844 A1 11th January 2007 Inventor: William N. Barbat

SELF-SUSTAINING ELECTRIC POWER GENERATOR UTILISING ELECTRONS OF LOW INERTIAL MASS TO MAGNIFY INDUCTIVE ENERGY

This patent application shows a very neat, self-powered electrical generator with a theoretical output of anything up to a COP of 59 when using cadmium selenide. The discussion of the theoretical aspects of the design includes a large amount of historical information and it covers the origin of the "law" of Conservation of Energy which, in spite of being incorrect, has been for decades, a major obstacle to the scientific development of free-energy devices.

Filed: 6th March 2006
Assignee: Levitronics, Inc.
Provisional application No. 60/697,729 filed on 8th July 2005

ABSTRACT

Electrical oscillations in a metallic "sending coil" radiate inductive photons toward one or more "energy-magnifying coils" comprised of a photoconductor or doped semiconductor coating a metallic conductor, or comprised of a superconductor. Electrons of low inertial mass in the energy-magnifying coil(s) receive from the sending coil, a transverse force having no in-line backforce, which exempts this force from the energy-conservation rule. The low-mass electrons in the energy-magnifying coil(s) receive increased acceleration proportional to normal electron mass divided by the lesser mass. Secondarily radiated inductive-photon energy is magnified proportionally to the electrons' greater acceleration, squared, e.g., the inductive-energy-magnification factor of CdSe photoelectrons with 0.13 x normal electron mass is 59 times. Magnified inductive-photon energy from the energy-magnifying coil(s) induces oscillating electric energy in one or more metallic "output coil(s)". The electric energy output exceeds the energy input if more of the magnified photon induction energy is directed toward the output coil(s) than is directed as a counter force to the sending coil. After an external energy source initiates the oscillations, feedback from the generated surplus energy makes the device a self-sustaining generator of electric power for useful purposes.

CROSS REFERENCE TO RELATED APPLICATION

This application corresponds to, and claims the benefit under 35 U.S.C. 119(e), of U.S. provisional application No. 60/697,729, filed on 8th July 2005, incorporated herein by reference in its entirety.

FIELD

This disclosure introduces a technical field in which practical electrical energy is created in accordance with the overlooked exception to the energy-conservation rule that Herman von Helmholtz described in his 1847 doctrine on energy conservation: "If . . . bodies possess forces which depend upon time and velocity, or which act in directions other than lines which unite each pair of material points, . . . then combinations of such bodies are possible in which force may be either lost or gained as infinitum". A transverse inductive force qualifies for Helmholtz's ad infinitum rule, but this force is not sufficient of itself to cause a greater energy output than input when applied to electrons of normal mass due to their unique charge-to-mass ratio. However, the increased acceleration of conduction electrons of less-then-normal inertial mass, as occurs in photoconductors, doped semiconductors, and superconductors, is proportional to the normal electron mass divided by the low electron mass, and the magnification of harnessable inductive energy is proportional to the square of the greater relative acceleration.

BACKGROUND

Magnetic force also satisfies Helmholtz's exemption to the energy-conservation rule because magnetic force is transverse to the force that causes it, and magnetic force is determined by the "relative velocity" (i.e. perpendicular to the connecting line) between electric charges. Magnification of magnetic force and energy was demonstrated by E. Leimer (1915) in the coil of a speaker phone and in the coil of a galvanometer when he irradiated a radio antenna-wire with radium. A 10 milligram, linear radium source produced a measured 2.6 fold increase in electrical current in the antenna wire in comparing inaudible radio reception without radium to audible reception with radium. This represented a $(2.6)^2$ = 7 times increase in electrical energy flowing through the respective wire coils. The possibility of this enhanced reception being attributed to a person's body holding the

unit of radium to the wire was eliminated by Leimer's additional observation that whenever the orientation of the small radium unit was changed to approximately 30 degrees relative to the wire, the energy enhancement ceased.

Applicant has deduced that Leimer's energy magnification was most likely due to low-mass electrons that were liberated and made conductive in the antenna by alpha radiation, which allowed these special electrons to be given a greater than normal acceleration by the received radio broadcast photons. Applicant has further deduced that such low-mass electrons must have originated in a thin-film coating of cupric oxide (CuO) on the antenna wire. CuO is a dull black polycrystalline semiconducting compound that develops in situ on copper and bronze wire in the course of annealing the wire in the presence of air. Such CuO coatings have been observed by Applicant on historical laboratory wire at the Science Museum at oxford University, U.K. and on copper house wire of that era in the U.S., indicating that CuO coatings were commonplace. In later years, annealing has taken place under conditions that prevent most oxidation. This is followed by acid treatment to remove any remaining oxides, leaving shiny wire.

The same year that the English translation of Leimer's paper appeared in *Scientific American*, 16-year old Alfred M. Hubbard of Seattle, Washington, reportedly invented a fuelless generator, which he later admitted, employed radium. Applicant interprets this as implying that Leimer's energy-magnification was utilised by Hubbard with feedback to make it self-sustaining. Three years later, Hubbard publicly demonstrated a relatively advanced fuelless generator that illuminated a 20-watt incandescent bulb (Anon. 1919a). A reputable physics professor from Seattle College, who was intimately familiar with Hubbard's device (but not at liberty to disclose its construction details), vouched for the integrity of the fuelless generator and declared that it was not a storage device, but he did not know why it worked (Anon. 1919b). Because Hubbard initially had no financial means of his own, it is likely that the professor had provided Hubbard with the use of the expensive radium initially and thereby witnessed the inventing process in his own laboratory.

Newspaper photos (Anon. 1920a) of a more impressive demonstration of Hubbard's fuelless generator, show a device described as 14 inches (36 cm) long and 11 inches (28 cm) in diameter, connected by four heavy electrical cables to a 35 horsepower (26 kW) electric motor. The motor reportedly propelled an 18-foot open launch around a like at a speed of 8 to 10 knots (Anon. 1920b). The event was witnessed by a cautious news reporter who claims to have checked thoroughly for any wires that might have been connected to hidden batteries, by lifting the device and motor from the boat. Radioactive-decay energy can be eliminated as the main power source because about 10^8 times more radium than the entire world's supply would have been needed to equal Hubbard's reported electric energy output of 330 amperes and 124 volts.

Lester J. Hendershott of Pittsburgh, Pa., reportedly demonstrated a fuelless generator in 1928 that was claimed by Hubbard to be a copy of his own device (1928h). The president of Stout Air services, William B. Stout, who also designed the Ford Trimotor aeroplane, reported (1928b): "The demonstration was very impressive. It was actually uncanny.... The small model appeared to operate exactly as Hendershot explained it did". Also reportedly attesting to the operability of Hendershott's fuelless generator were Colonel Charles A. Lindbergh and Major Thomas Lanphier of the U.S. Air Corps (1928a, et seq.), and Lanphier's troops reportedly assembled a working model of his device.

To the Applicant's best knowledge, the only depiction that was made public of the interior components of any of these reported generators consists of a sketchy drawing (Bermann 1928h) of Hubbard's apparatus similar in size to the device shown in his 1919 demonstration. It depicts a complex set of parallel coils measuring 6 inches (15 cm) in length and 4.5 inches (11.4 cm) in overall diameter. Four leads of insulated wire, with the insulation peeled back, are shown coming out of the end of the device. What those four wires were connected to internally was not shown. Hubbard's description of the internal arrangement of coils in the device generally matches the drawing (Anon. 1920a): "It is made up of a group of eight electromagnets, each with primary and secondary windings of copper wire, which are arranged around a large steel core. The core likewise has a single winding. About the entire group of cells is a secondary winding". Nothing was reported or depicted about how components functioned with each other, or how much radium was used and where the radium was positioned. The only connectors visible on the drawing were between the outer windings of the eight electromagnet coils. These connectors show that the direction of the windings alternated between clockwise and counterclockwise on adjacent coils, so that the polarity of each electromagnet would have been opposite to that of it's adjacent neighbours.

If the Hubbard and Hendershot devices actually operated as reported, they apparently never attained acceptance or commercial success. Assuming the devices actually worked, their lack of success may have been largely financially or supply based, or both, compounded with scepticism from believers in the energy-conservation doctrine. How much radium was employed by Hubbard in his larger generator can only be guessed at, but assuming a typical laboratory radium needle containing 10 milligrams of radium was used, that amount would have cost $900 in 1920, dropping to $500 in 1929. That much radium in a fuelless generator would have cost as much as an inexpensive automobile in the 1920s. Possibly much more radium was used than 10 milligrams.

In 1922, when the Radium Company of America of Pittsburgh, Pa., reportedly discontinued its work with Hubbard on his invention (1928h), the entire world's supply of radium was only about 250 grams. With the extreme assumption that only 1 milligram of radium was needed per generator, less than 10% of a single year's production

of autos in the US in the mid-1920s could have been supplied with such generators. Apparently Hendershott had tried to revive the technology by showing that the fuelless generator could extend the range of air flight indefinitely, but his technology never attracted a sponsor from any private, public or philanthropic entity.

U.S. Pat. No. 4,835,433 to Brown, superficially resembles the drawing of Hubbard's device. Brown's device appears to have the same number and essentially the same general arrangement of wire coils as Hubbard's generator, as nearly as can be understood from the newspaper articles depicting that device. Apparently, no information concerning either the Hubbard or Hendershot devices was considered during the prosecution of the '433 patent. Brown discusses the conversion of energy of radioactive decay products, principally alpha emissions, to electrical energy by amplifying electrical oscillations in a high-Q L-C circuit irradiated by radioactive materials. "During the absorption process, each alpha particle will collide with one or more atoms in the conductor, knocking electrons from their orbits and imparting some kinetic energy to the electrons in the conductor, thereby increasing its conductivity". (Col. 3, Line 68 to Col. 4, line 5). No claim was made by Brown, that the device employed a semiconductor or photoconductor that could have provided low-mass electrons for energy magnification.

Brown claimed an output of 23 amps at 400 volts, which is vastly greater than all the decay energy represented by his reported radioactive content of 1 milligram of radium that was surrounded by weakly radioactive uranium rods and thorium powder. Powered thorium is highly pyrophoric, so it is typically sealed in a nitrogen atmosphere to prevent spontaneous combustion. In his device, Brown reportedly confined the thorium in cardboard without any mention of sealing out air. This condition would have invited a meltdown that could have been interpreted as massive out-of-control electrical production.

To the best of the Applicant's knowledge, no person other than the Applicant has ever indicated that the presence of cupric oxide on their wires could have provided energy magnification. If Hubbard's device actually did work, certain characteristics of its design are unexplainable by the Applicant, namely the use of four rather than two large electrical cables to connect his device to an electrical motor, and the use of alternating polarity instead of single-direction polarity in the orientation of the multiple coils surrounding a central coil. Applicant therefore believes that the specification herein sets forth original configurations of electrical-energy generators that have no known precedent.

SUMMARY

To address the needs for electrical generators which are capable of self-generating substantial amounts of electrical power in various environments, and which are portable as well as stationary, apparatus and methods are provided for magnifying an electrical input, and (with feedback) for generating usable electrical power indefinitely without fuel or other external energy source, except for starting. The apparatus utilises electrons of low effective mass, which receive greater acceleration than normal electrons in an amount that is inversely proportional to the effective mass. Applicant has determined that effective mass is the same as the electron's true inertial mass. The photon energy that is radiated when an electron is accelerated is proportional to the square of the acceleration, so the increase in radiated photon energy from an accelerated low-mass electron over the energy from a normal electron is equal to the inverse square of the effective mass, e.g. the calculated energy magnification provided by photoconducting electrons in cadmium selenide, with an electron effective mass of 0.13, is 59 times. The use of a transverse force, that lacks a direct back-force, to accelerate low-mass electrons in an oscillating manner, circumvents any equal-and-opposite force that would invoke the application of the energy-conservation law of kinetics and thermodynamics.

The various embodiments of the apparatus, which are configured either to continuously magnify an input of oscillating electric energy, or to serve as a self-sustaining electric generator, employ three principal components:

At least one sending coil

At least one energy-magnification coil, comprising a material that produces , in a "condition" low-mass electrons, and

At least one output coil.

It is desirable that the apparatus also includes a means for establishing the condition with respect to the energy-magnifying coil(s). Except where otherwise indicated in the remainder of this text, where the number of coils of a particular type is referred to in the singular, it will be understood that a plurality of coils of the respective type can alternatively be utilised.

Electrical oscillation in the sending coil, which is comprised of a metallic conductor, causes radiation of inductive photons from the sending coil. The energy-magnifying coil is situated in a position relative to the sending coil so as to receive inductive photons from the sending coil. The inductive photons radiating from electrical oscillations in the sending coil, convey a transverse force to the low-mass electrons in the energy-magnification coil with no back-force on the sending coil. The greater-than-normal accelerations which are produced in the low-mass electrons of the energy-magnifying coil, produce greater irradiation energy of inductive photons than normal.

The output coil is positioned so as to receive the magnified inductive-photon energy from the energy-magnifying coil. The inductive-photon energy received by the output coil, which is comprised of a metallic conductor, is converted into an oscillating electrical current of normal electrons. In order for the electrical output to exceed the electrical input, the output coil is situated in such a manner that it receives more of the magnified inductive-photon energy than that which is directed back against the sending coil to act as a back-force. This "energy leverage" causes the electrical energy output to exceed the electrical energy input.

By way of example, the energy-magnifying coil can comprise a superconducting material, wherein the "condition" is a temperature (e.g. a cryogenic temperature) at which the superconducting material exhibits superconducting behaviour characterised by production of low-mass electrons.

By way of another example, the energy-magnifying coil can comprise a photoconductive material, wherein the "condition" is a situation in which the photoconductive material is illuminated by a wavelength of photon radiation sufficient to cause the photoconductive material of the energy-magnifying coil to produce conduction electrons having low effective mass. In this latter example, the means for establishing the condition can comprise a photoconductor exciter (e.g. one or more LEDs) situated and configured to illuminate the photoconductive material of the energy-magnifying coil with the wavelength of photon radiation.

By way of yet another example, the "condition" is the presence of a particular dopant in a semiconductor that provides a low-mass electron as a charge carrier. Also, by way of example, the energy-magnifying coil can comprise a semiconductive element or compound that has been doped with a particular element or compound that makes it conductive of low-mass electrons without illumination by photon radiation other than by ambient photons.

Various apparatus embodiments comprise different respective numbers and arrangements of the principal components. The various embodiments additionally can comprise one or more of circuitry, energisers, shielding and other components to fulfill the object of providing a self-sustaining source of electrical power for useful purposes.

Also provided, are methods for generating an electrical current. In an embodiment of such a method, a first coil is energised with an electrical oscillation sufficient to cause the first coil to radiate inductive photons. At least some of the radiated inductive photons from the first coil are received by a second coil, called "the energy-magnifying coil", comprising a material that produces low-mass electrons. The received inductive photons impart respective transverse forces to the low-mass electrons that cause the low-mass electrons to experience accelerations in the material which are greater than accelerations that otherwise would be experienced by normal free electrons experiencing the transverse forces.

Conduction of the accelerated low-mass electrons in the second coil, causes the second coil to produce a magnified inductive force. The magnified inductive force is received by a third coil which causes the third coil to produce an oscillating electrical output of normal conduction electrons which has greater energy than the initial oscillation. A portion of the oscillating electrical output is directed as feed-back from the third coil to the sending coil, so as to provide the electrical oscillation to the sending coil. This portion of the oscillating electrical current directed to the sending coil, desirably is sufficient to cause self-sustaining generation of inductive photons by the first coil without the need for any external energy source. The surplus oscillating electrical output from the third coil can be directed to a work loop.

The method can further comprise the step of starting the energisation of the first coil to commence generation of the oscillating electrical output. This "starting" step can comprise momentarily exposing the first coil to an external oscillating inductive force or for example, to an external magnetic force which initiates an electrical pulse.

The foregoing and additional features and advantages of the invention will be more readily apparent from the following detailed description, which proceeds with reference to the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

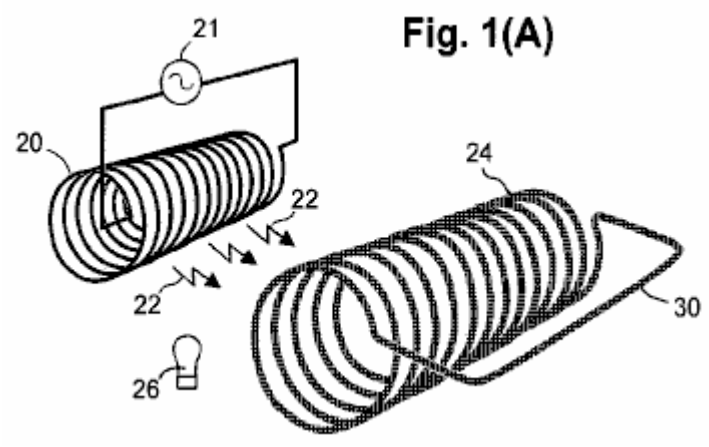


Fig.1A is a perspective view schematically depicting a sending coil in relationship to an energy-magnifying coil such that inductive photons from the sending coil, propagate to the energy-magnifying coil.

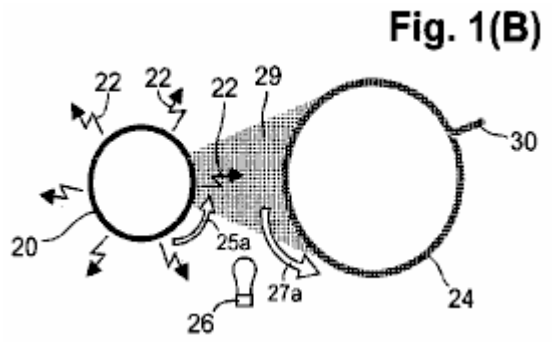


Fig.1B is a schematic end-view of the sending coil and energy-magnifying coil of **Fig.1A**, further depicting radiation of inductive photons from the sending coil and the respective directions of electron flow in the coils.

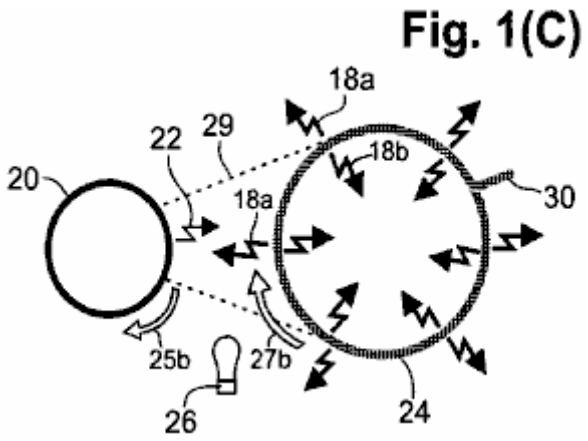


Fig.1C is a schematic end-view of the sending coil and energy-magnifying coil of **Fig.1A**, further depicting the production of inwardly-radiating and outwardly-radiating magnified inductive photons from the energy-magnifying coil.

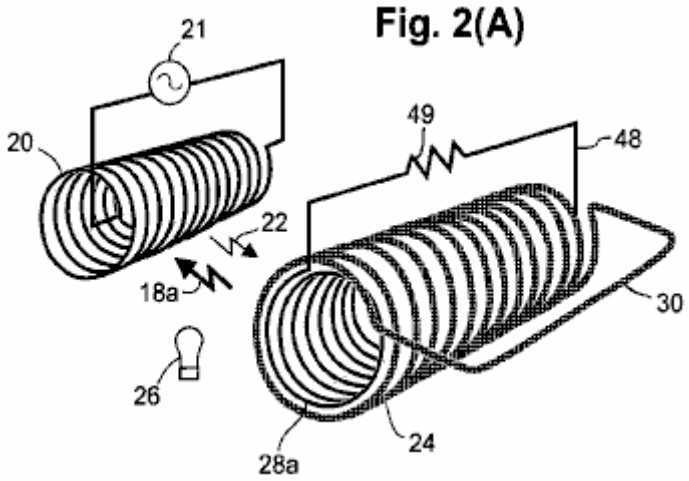


Fig.2A is a perspective view schematically showing an internal output coil, coaxially nested inside the energy-magnifying coil to allow efficient induction of the internal output coil by the energy-magnifying coil, wherein the induction current established in the internal output coil is used to power a load connected across the internal output coil.

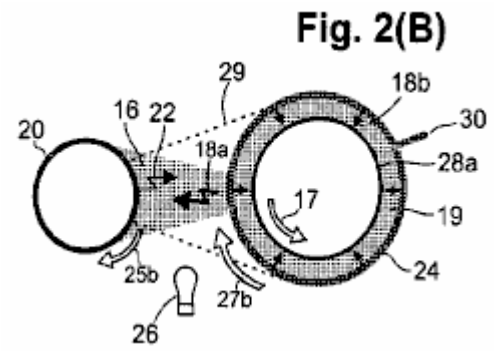


Fig.2B is a schematic end-view of the coils shown in **Fig.2A**, further depicting the greater amount of magnified inductive-photon radiation that is received by the external output coil in comparison to the lesser amount that is directed toward the sending coil to act as a back-force.

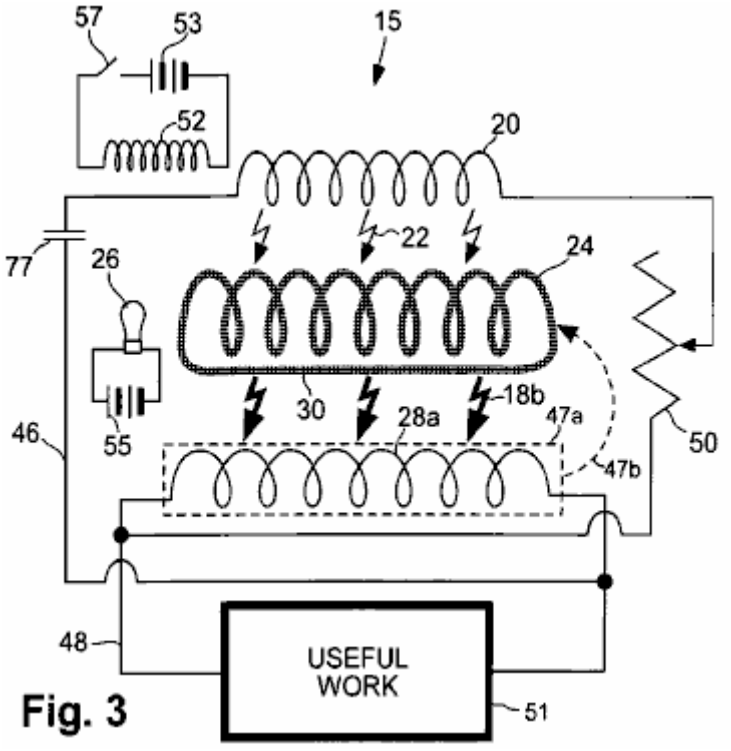


Fig.3 is an electrical schematic diagram of a representative embodiment of a generating apparatus.

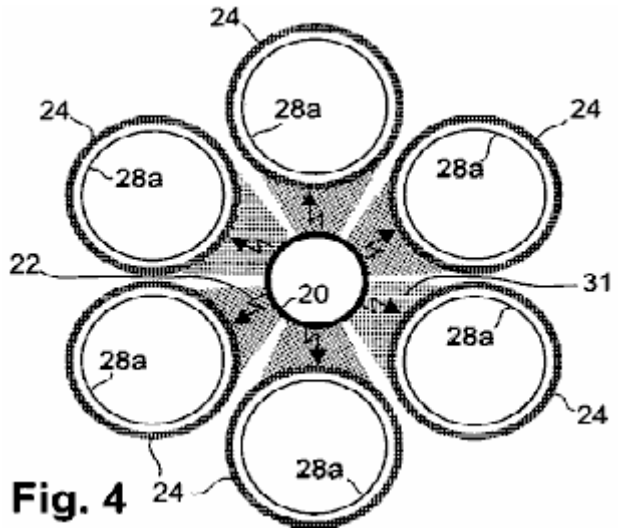


Fig. 4

Fig.4 is a schematic end-view of a representative embodiment, comprising a centrally disposed sending coil surrounded by six energy-magnifying coils, each having and axis which is substantially parallel to the axis of the sending coil. A respective internal output coil is coaxially nested inside each energy-magnifying coil, and the energy-magnifying coils are arranged so as to capture substantially all the inductive photons radiating from the sending coil.

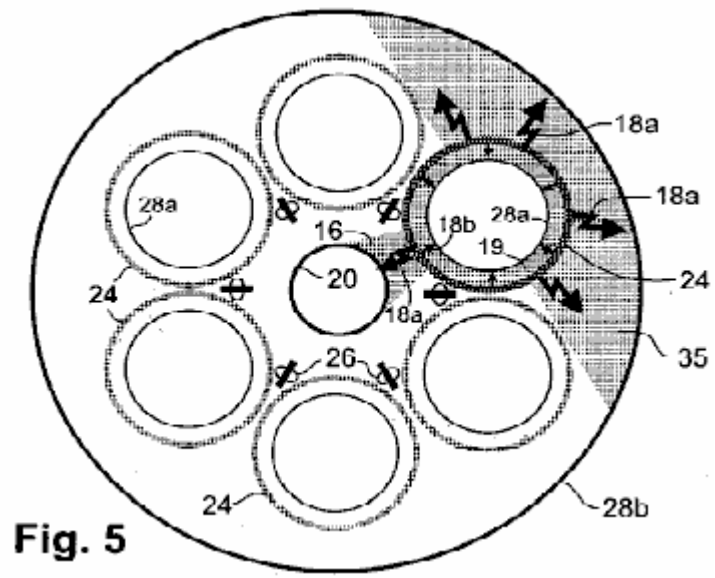


Fig. 5

Fig.5 is a schematic end-view of the embodiment of **Fig.4**, further including an external output coil situated coaxially with the sending coil and configured to surround all six energy-magnifying coils so as to capture outwardly-radiating inductive photons from the energy-magnifying coils. Also depicted is the greater amount of magnified inductive-photon radiation that is received by the internal output coils and the external output coil in comparison to the lesser amount of inductive-photon radiation that is directed towards the sending coil to act as a back-force. Also shown are the arrays of LEDs used for exciting the energy-magnifying coils to become photoconductive.

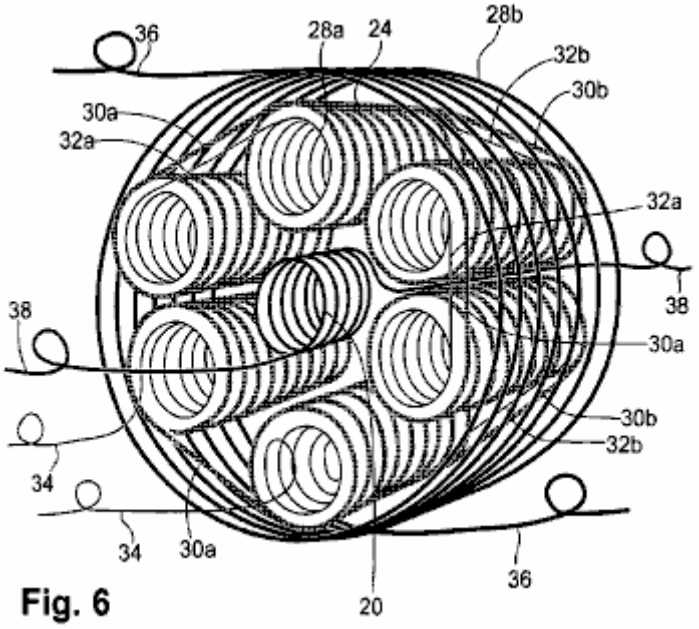


Fig. 6

Fig.6 is a perspective view of the embodiment of **Fig.4** and **Fig.5** but further depicting respective inter-coil connections for the energy-magnifying and internal output coils, as well as respective leads for the sending coil, internal output coils and external output coil.

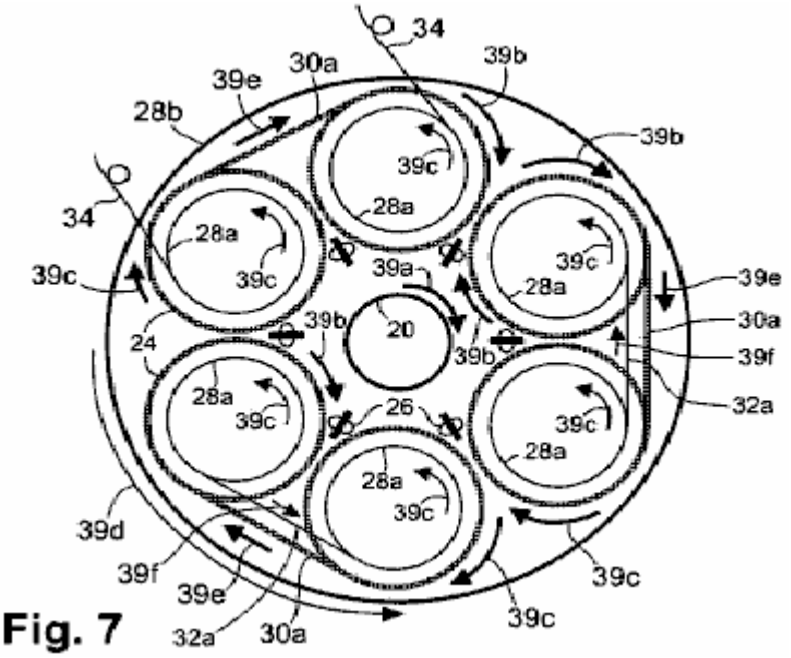


Fig. 7

Fig.7 is a head-end view schematically depicting exemplary current-flow directions in the sending coil, energy-magnifying coils, internal output coils, and external output coils, as well as in the various inter-coil connections of the embodiment of **Fig.4**.

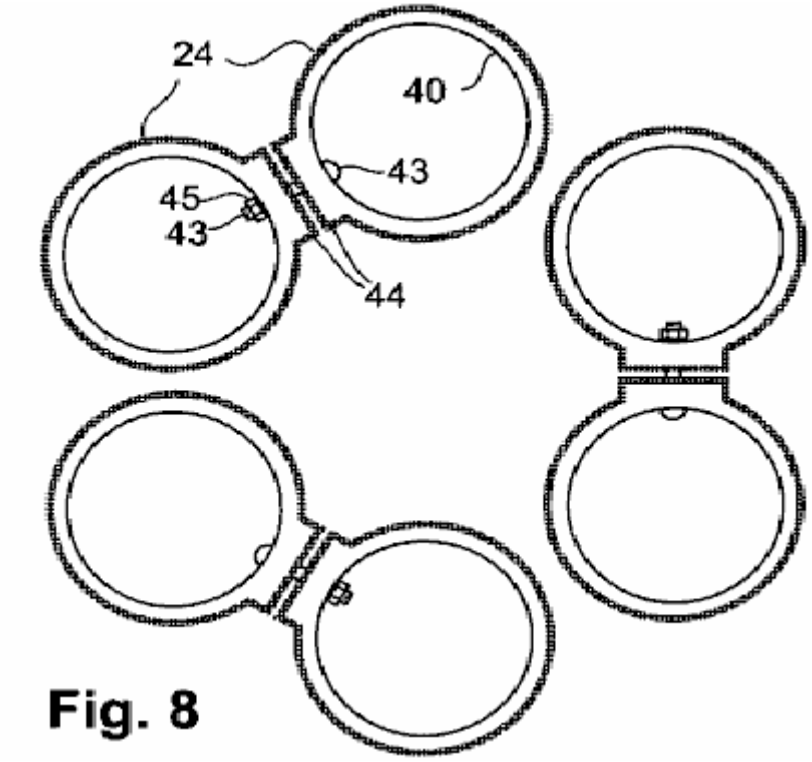


Fig. 8

Fig.8 is a schematic end-view showing an embodiment of the manner in which inter-coil connections can be made between adjacent energy-magnifying coils.

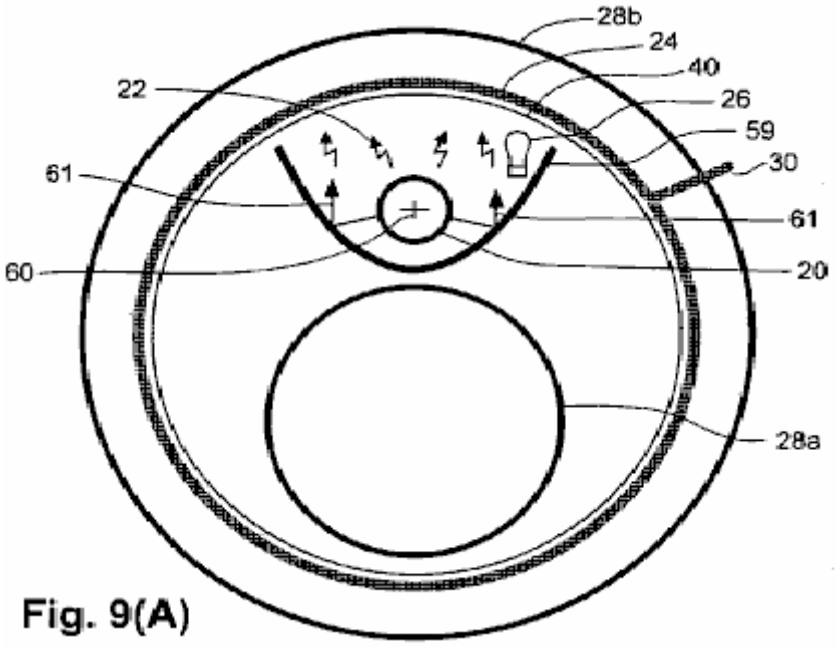


Fig. 9(A)

Fig.9A is a schematic end-view depicting the coil configuration of an embodiment in which a sending coil and an internal output coil are nested inside an energy-magnifying coil, which in turn is nested inside an exterior output coil. A metallic separator, having a substantially parabolic shape, and being situated between the sending coil and the internal output coil, reflects some of the otherwise unused inductive-photon radiation to maximise the effective radiation received by the energy-magnifying coil. Also, the metallic shield prevents the internal output coil from receiving radiation sent from the sending coil.

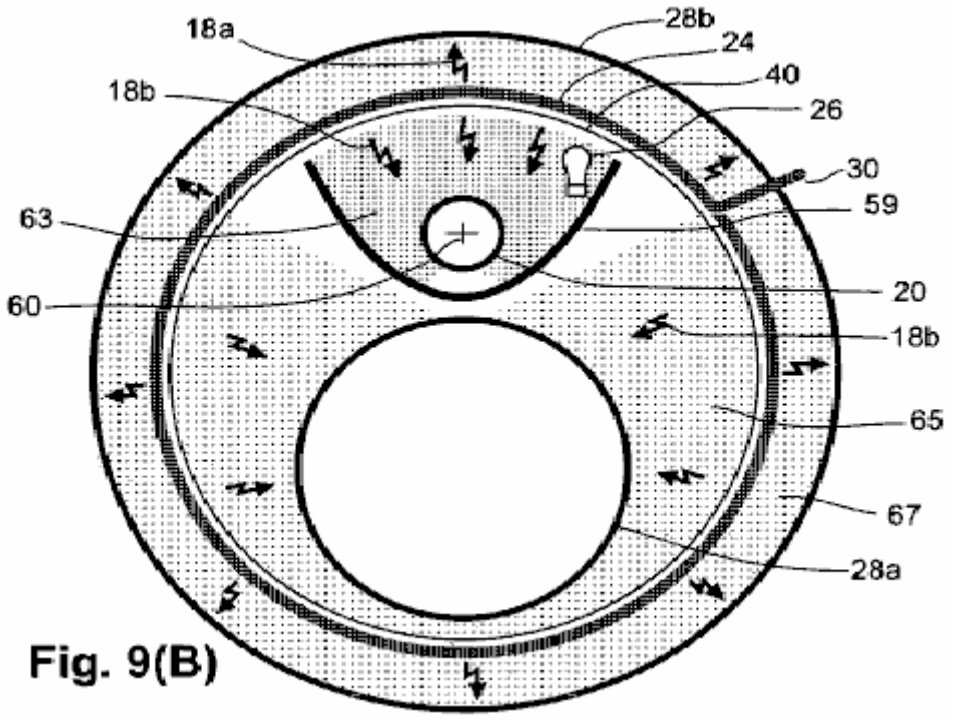


Fig. 9(B)

Fig.9B is a schematic end-view of the coil configuration of **Fig.9A**, further depicting the metallic separator acting as a shield to restrict the back-force radiation reaching the sending coil while allowing the internal output coil to receive a substantial portion of the magnified radiation from the energy-magnifying coil. Also depicted is the greater amount of magnified inductive-photon radiation that is received by the internal output coil and the external output coil in comparison to the lesser amount that is received by the sending coil to act as a back-force.

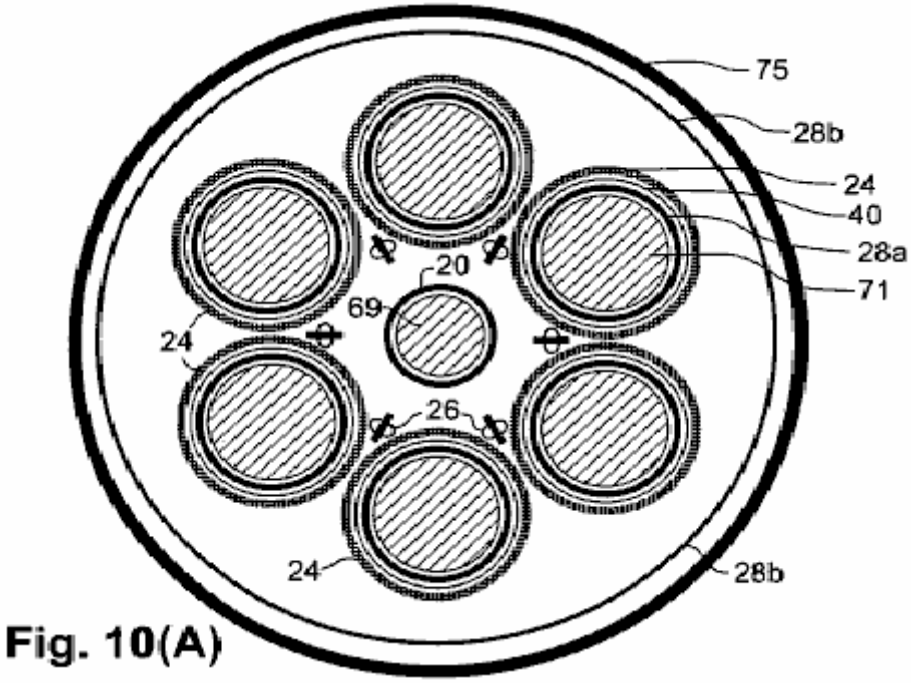


Fig. 10(A)

Fig10A is a schematic end-view depicting the coil configuration of yet another embodiment that is similar in some respects to the embodiment of **Fig.4**, but also including respective ferromagnetic cores inside the sending coil and internal output coils. Also depicted is a metallic shield surrounding the entire apparatus.

Fig. 10(B)

Fig.10B is a schematic end-view of a sending coil of yet another embodiment in which a ferromagnetic sleeve is disposed coaxially around the sending coil.

DETAILED DESCRIPTION

General Technical Considerations

An understanding of how “infinite energy” mistakenly came to be rejected by the scientific community, clarifies the basis of this invention. The electrodynamic function described in the embodiments described below, conforms to Helmholtz’s alternate energy rule, which states that a force which is not in line with it’s causative force “may be lost or gained ad infinitum”. This rule was included in “Uber die Erhaltung der Kraft” (“On the Conservation of Force”) that Hermann Helmholtz delivered to the Physical Society of Berlin in 1847. But, Helmholtz mistakenly believed that “all actions in nature are reducible to forces of attraction and repulsion, the intensity of the forces depending solely upon the distances between the points involved .... so it is impossible to obtain an unlimited amount of force capable of doing work as the result of any combination whatsoever of natural objects”.

Helmholtz refused to accept the idea that magnetic energy qualifies for ad infinitum status despite the fact that Ampere’s (1820) magnetic force on parallel straight conductors is obviously transverse to the direction of the electric currents rather than being in line with the currents. He omitted mention that the magnetic force in Ampere’s (1825) important invention, the solenoidal electromagnet, is caused by currents in the loops of his coils, which are transverse to the direction of magnetic force. Also, he failed to mention that Ampere considered the magnetic force of a permanent magnet to be caused by minute transverse circular currents, which are now recognised as electrons that spin and orbit transversely.

Helmholtz, who was educated as a military medical doctor without any formal study of physics, relied instead on an obsolete metaphysical explanation of magnetic force: “Magnetic attraction may be deduced completely from the assumption of two fluids which attract or repel in the inverse ratio of the square of their distance....It is known that the external effects of a magnet can always be represented by a certain distribution of the magnetic fluids on its surface”. Without departing from this belief in magnetic fluids, Helmholtz cited Wilhelm Weber’s (1846) similarly wrong interpretation that magnetic and inductive forces are directed in the same line as that between the moving electric charges which cause the forces.

Weber had thought that he could unify Coulombic, magnetic, and inductive forces in a single, simple equation, but Weber’s flawed magnetic-force term leads to the absurd conclusion that a steady current in a straight wire induces a steady electric current in a parallel wire. Also, a changing current does not induce an electromotive force in line with the current, as Weber’s equation showed. The induced force is offset instead, which becomes more apparent the further that two nested, coaxial coils are separated. What appears to be a directly opposing back-force is actually a reciprocal inductive force.

Helmholtz’s assertion that the total sum of the energy in the universe is a fixed amount that is immutable in quantity from eternity to eternity appealed to his young friends. But, the elder scientists of the Physical Society of Berlin declared his paper to be “fantastical speculation” and a “hazardous leap into very speculative metaphysics”, so it was rejected for publication in Annalen der Physik. Rather than accept this rejection constructively, Helmholtz found a printer willing to help him self-publish his work. Helmholtz headed the publication with a statement that his paper had been read before the Society, but he disingenuously withheld mention of its outright rejection. Unwary readers have since received the wrong impression that his universal energy-conservation rule had received the Society’s endorsement rather than its censure.

Helmholtz (1862, 1863) publicised his concept thus: “We have been led up to a universal natural law, which ... expresses a perfectly general and particularly characteristic property of all natural forces, and which ... is to be placed by the side of the laws of the unalterability of mass and the unalterability of the chemical elements”. Helmholtz (1881) declared that any force that did not conserve energy would be “in contradiction to Newton’s axiom, which established the equality of action and reaction for all natural forces” (sic). With this deceitful misrepresentation of Newton’s strictly mechanical principle, Helmholtz had craftily succeeded in commuting the profound respect for Newton’s laws to his unscientific doctrine. Subsequently, the Grand Cross was conferred on Helmholtz by the kings of Sweden and Italy and the President of the French Republic, and he was welcomed by the German Emperor into nobility with the title of “von” added to his name. These prestigious awards made his doctrine virtually unassailable in the scientific community.

Ampere’s principle of transverse magnetic attraction and repulsion between electric currents had been made into an equation for the magnetic force between moving electric charges by Carl Fredrick Gauss (written in 1835, published posthumously in 1865). The critical part of the Gauss equation shows, and modern physics texts agree, that magnetic force is transverse to the force that imparts a relative velocity (i.e. perpendicular to a connecting line) between charges. Lacking a direct back-force, a transverse magnetic force can produce a greater force than the force that causes it.

The only physicist to recognise in print, the profound significance of the work of Gauss, was James Clerk Maxwell (1873), who stated "(If Gauss's formula is correct), energy might be generated indefinitely in a finite system by physical means". Prepossessed with Helmholtz's "law", Maxwell chose not to believe Gauss's transverse magnetic-force equation and accepted Wilhelm Weber's (1846) erroneous in-line formula instead. Maxwell even admitted knowing of Gauss's (1845) rebuke of Weber for his mistaken direction of magnetic force as "a complete overthrow of Ampere's fundamental formula and the adoption of essential a different one".

In 1893, the critical part of Ampere's formula for magnetic force, which Weber and Maxwell rejected, and which Helmholtz had replaced with his contrary metaphysical explanation, was proposed for the basis for the international measure of electric current, the Ampere (or amp), to be defined in terms of the transverse magnetic force which the current produces. But Helmholtz's doctrine had become so impervious to facts that anyone who challenged this "law" faced defamation and ridicule.

The first recognition of unlimited energy came from Sir Joseph Larmor who reported in 1897, "A single ion e, describing an elliptic orbit under an attraction to a fixed centre ... must rapidly lose its energy by radiation ... but in the cases of steady motion, it is just this amount that is needed to maintain the permanency of motion in the aether". Apparently to mollify critics of his heretical concept, Larmor offered a half-hearted recantation in 1900: "The energy of orbital groups ... would be through time, sensibly dissipated by radiation, so that such groups could not be permanent".

In 1911, Rutherford found that an atom resembles a small solar system with negative ions moving like planets around a small, positively charged nucleus. These endlessly orbiting electrons were a source of the perpetual radiation that had aptly been described by Larmor, and these orbiting electrons were also Planck's (1911) "harmonic oscillators" which he used to explain Zero-point Energy (ZPE). ZPE was shown by the fact that helium remains liquid under atmospheric pressure at absolute zero, so that helium must be pressurised to become solid at that temperature. Planck believed that harmonic oscillators derived "dark energy" from the aether to sustain their oscillations, thereby admitting that an infinite source of energy exists. However, he assigned an occult origin to this infinite energy, rather than a conventional source that had not met with Helmholtz's approval.

Niels Bohr (1924) was bothered by the notion that radiation from an orbiting electron would quickly drain its energy so that the electron should spiral into the nucleus. Whittaker (1951) states, "Bohr and associates abandoned the principle ... that an atom which is emitting or absorbing radiation must be losing or gaining energy. In its place, they introduced the notion or virtual radiation, which was propagated in ... waves but which does not transmit energy or momentum". Subsequently, the entire scientific community dismissed Larmor radiation as a source of real energy because it failed to conform to Helmholtz's universally accepted doctrine.

Helmholtz's constraining idea that the vast amount of light and heat radiating from the many billions of stars in the universe can only come from previously stored energy, has led scientists to concur that fusion of pre-existing hydrogen to helium, supplies nearly all the energy that causes light and heat to radiate from the sun and other starts. If so, then the entire universe will become completely dark after the present hydrogen supply in stars is consumed in about 20 billion years. William A. Fowler (1965) believed that essentially all the hydrogen in the universe "emerged from the first few minutes of the early high-temperature, high-density stage of the expanding Universe, the so-called 'big bang' ..." Moreover, the background energy of the universe was thought by some to be "relic" radiation from the "Big Bang".

To accept the Big Bang idea that all the stars in the universe originated at the same time, it was necessary to disregard the fact that most stars are much younger or much older than the supposed age of the one-time event, which indicates that their energy must have come from a recurring source. The Big Bang is entirely dependent on the idea that the whole universe is expanding, which stemmed from the interpretation that Hubble's red-shift with distance from the light source, represents a Doppler shift of receding stars and galaxies. This expanding-universe interpretation was shattered by William G. Tifft (1976, 1977), who found that observed red-shifts are not spread randomly and smoothly over a range of values, as would be expected from the Doppler shifts of a vast number of receding stars and galaxies. Instead, the observed red-shifts all fall on evenly spaced, quantised values.

Moreover, Shpenkov and Kreidik (2002) determined that the radiation temperature corresponding to the fundamental period of the orbital electron motion in the hydrogen atom of 2.7289^0K matches the measured temperature of cosmic background radiation of 2.725^0K plus or minus 0.002^0K. This represents perpetual zero-level Larmor radiation from interstellar hydrogen atoms dispersed in the universe. So, Helmholtz's idea that "the energy in the universe is a fixed amount immutable in quantity from eternity to eternity" does not stand up to known facts.

The large aggregate quantity of heat-photons which is generated continually by Larmor radiation can account for the illumination of stars and for the enormous heat and pressure in active galactic centres. Based on the fact that photons exhibit momentum, photons must posses mass, because, as Newton explained, momentum is mass times velocity, which in this case is "c". Consequently, the creation of photons by induction or by Larmor radiation, also creates new mass. The conditions that Fowler was seeking for hydrogen nucleosynthesis, are apparently being supplied indefinitely in active galaxies and possibly in the sun and other stars above a certain size. This invention utilises a similar unlimited energy source.

Another principle that is important to this specification, is that the transfer of energy by electrical induction was found by the Applicant to work in the same manner as the transfer of energy by broadcast and reception of oscillating radio signals. A transverse force is communicated in both cases, the force declines similarly with distance, and the effects of shielding and reflection are identical. Since radio signals are communicated by photons, Applicant considers that inductive force is also communicated by photons. The radiation of newly formed inductive photons results when an accelerated charge experiences a change in direction of acceleration. Inductive radiation occurs when the acceleration of electric charges is reversed, as in Rontgen's bremsstrahlung, in Hertz's linear oscillator (plus all other radio-broadcasting antennas), and in coils which carry an alternating current.

In a similar case, when electric charges move in a curving motion due to a continually changing centripetal acceleration, inductive photons are radiated steadily. This includes the radiation from electrons orbiting atomic nuclei (Larmor radiation) and from conduction electrons flowing in a wire coil, whether the current is steady or not. Circularly produced inductive photons induce a circular motion (diamagnetism) in mobile electrons located near the axis of the electron's circular movement.

In both the reverse-acceleration and centripetal-acceleration cases, inductive photons convey a force to mobile electrons that is transverse to the photon's propagation path. As Lapp and Andrews (1954) reported, "Low-energy photons produce photoelectrons at right angles to their path ...". This same right-angle force without a direct back-force, applies as well, to all conduction electrons which are accelerated by low-energy photons. Hence, inductive energy qualifies for exemption from the energy-conservation law by Helmholtz's same ad infinitum principle which exempts magnetic energy.

The transverse force that inductively produced photons delivered to mobile electrons, is opposite in direction to the simultaneous movement of the primary charge which produces the radiation. This is shown by Faraday's induced current opposite to the inducing current and by the diamagnetically-induced circular motion which, in a rotational sense, is opposite to the circular electron motion in the coil producing it. An oscillating flow of electrons within a loop of a wire coil, induces a force on the conduction electrons which is in the opposite direction in adjacent loops of the same wire. This results in self-induction.

Important to this specification is the realisation that the energy transmitted by photons is kinetic rather than electromagnetic. Inductively radiated photons of low energy, light rays and X-rays cannot be deflected by and electric or magnetic field due to the photons' neutral charge. Neither do neutral photons carry an electric or magnetic field with them. Photon radiation is produced by a change in the acceleration of an electric charge, so only in special cases does it have an electrokinetic origin which involves a magnetic force. To honour these facts, Applicant uses the term "electrokinetic spectrum" in place of "electromagnetic spectrum".

Another principle which is important to this specification is the realisation that, although the charge on the electron has a constant value under all conditions, the mass of an electron is not a fixed, unchanging amount. All free electrons, as in cathode rays, have exactly the same amount of mass at sub-relativistic velocities. This is called "normal" mass and is denoted by m_e. Free electrons have a unique charge to mass ratio that makes the magnetic force resulting from a sub-relativistic velocity imparted to such an electron, exactly equal to the energy input with "normal" electrons.

Also, when a normal electron is given a sub-relativistic acceleration, the inductive force it produces is equal to the force it receives. The mass of highly conductive electrons of metals is apparently very close to normal, but any very slight inductive-energy gains would be masked by inefficiencies. The ubiquity of free electrons and the conduction electrons of metals has led to the view that electron mass is a never-varying figure that would allow the energy conservation law to apply to magnetic energy and inductive energy.

Accurate determinations of electron mass in solid materials have been made possible by cyclotron resonance, which is also called diamagnetic resonance. The diamagnetic force produced by the steady flow of electrons in a wire coil, induces the mobile electrons of a semiconductor to move in a circular orbit of indefinite radius but at a definite angular frequency. This frequency is only related to the inductive force and the mass of the electron. At the same time, a repulsive magnetic force is developed by the relative velocity between the electron flow in the coil and the conduction electrons, causing the mobile electrons of the semiconductor to move in a helical path away from the coil rather than in planar circles. Only two measurements are needed to determine the mass of such an electron: the cyclotron frequency which resonates with the frequency of the electron's circular motion, and the strength of the inductive force, which is determined by the current and dimensions of the coil. Since the co-produced magnetic field is related to the same parameters, its measurement serves as a surrogate for inductive force.

Because the measured mass of conduction electrons in semiconductors is less than normal, a complicated explanation has been adopted to defend the constancy of electron mass in order to support Helmholtz's energy doctrine. An extra force is supposedly received from the vibrational lattice-wave energy of the crystal (in what would have to be an act of self-refrigeration) to make normal-mass electrons move faster than expected around a circular path, thereby giving the appearance that the electron has less mass than normal. In this explanation, the electron is considered to be a smeared-out wave rather than a particle, which is contradicted by the billiard-ball-

like recoil of an electron when it is bumped by a quantum of radiation, as described by Arthur Crompton and Samuel Allison (1935).

The fallacy that borrowed energy can provide a boost in velocity to an electron, is more apparent in the case of linear motion. The effective-mass theory considers that the greater linear velocity is caused by a boost given to normal-mass electrons by a "longitudinal wave" imparted by an externally applied force in the same direction as the electron motion. Since this longitudinal wave is also considered to have a source in crystal-lattice vibrations, the effective-mass theory relies on a reversal of entropy in violation of the second Law of Thermodynamics.

No reasonable contribution of direct directional energy can be invoked from any source to impart abnormally great velocity to the conduction electrons in semiconductors. So, the operation of apparatus embodiments described herein, relies on electrons having particle properties and on electrons having less-then-normal inertial mass without invoking any special forces. This is supported by Brennan's (1999) statement that "the complicated problem of an electron moving within a crystal under the interaction of a periodic but complicated potential, can be reduced to that of a simple free particle, but with a modified mass". The term "effective" is herein considered redundant in referring to truly inertial mass, but "effective mass" still has relevance in referring to the net movement of orbital vacancies or "holes" in the opposite direction of low-mass electrons.

By F = ma, a low-mass electron receives greater acceleration and greater velocity from a given force than an electron of normal mass. The velocity and kinetic energy imparted to an electrically charged body by a force, are determined by the electric charge without regard to the body's mass. Having a smaller amount of mass, allows a body to attain a greater velocity with any given force. Hence, the magnetic force produced by the charge at this higher velocity will be greater than it would normally be for that same amount of force. This allows low-mass electrons to produce a magnetic force that is greater than the applied force.

Also, the amount of inductive radiation energy from accelerated electrons is related to an electron's charge without regard to its mass. The energy of inductive radiation increases with the square of the electron's acceleration according to Larmor's (1900) equation, while the acceleration is inversely proportional to the lesser electron mass relative to normal electron mass. Therefore, the greater-than-normal acceleration of low-mass electrons, allows the re-radiation of magnified inductive-photon energy at a magnification factor which is proportional to the inverse square of the electron's mass, e.g., the inductive-energy magnification factor of cadmium selenide photoelectrons with 0.13 of the normal electron mass is $(0.13)^2$ which is 59 times.

Electrons appear to acquire or shed mass from photons in order to fit the constraints of particular orbits around nuclei, because each orbit dictates a very specific electron mass. In metals, where the conduction electrons seem to move as would a gas, one might think that they would assume the normal mass of free electrons. But the largest mean free path of electrons in the most conductive metals is reportedly about 100 atomic spacings between collisions (Pops, 1997), so the conduction electrons apparently fall back into orbit from time to time and thereby regain their metal-specific mass values.

As conduction electrons pass from one metal type to another, they either lose or gain heat-photons to adjust their mass to different orbital constraints. In a circuit comprising two different metallic conductors placed in series contact with each other, the flow of conduction electrons in one direction will cause the emission of heat-photons at the junction, while an electron flow in the reverse direction causes cooling as the result of ambient heat-photons being absorbed by the conduction electrons at the junction (Peltier cooling effect). When a metal is joined with a semiconductor whose conductive electrons have much lower mass than in metals, much greater heating or cooling occurs at their junction.

John Bardeen (1941) reported that the (effective) mass of superconducting electrons in low-temperature superconductors is only 10^{-4} as great as the mass of normal electrons. This is demonstrated when superconducting electrons are accelerated to a much higher circular velocity than normal in diamagnetically induced eddy currents, which results in enormous magnetic forces which are capable of levitating heavy magnetic objects. Electrons with 10^{-4} times normal mass are apparently devoid, (or nearly devoid) of included photon mass, so normal electrons are deduced to posses about 10^4 times more included photon mass than the bare electron's own mass.

The means by which photon mass may be incorporated within, or ejected from electrons, can be deduced from known information. Based on the Thompson scattering cross-section, the classical radius of a normal electron is 2×10^{-15} cm. If the electron has uniform charge throughout a sphere of that radius, the peripheral velocity would greatly exceed the velocity of light in order to provide the observed magnetic moment. Dehmelt (1989) determined that the radius of the spinning charge which creates an electron's magnetism, is approximately 10^{-20} cm. This apparent incongruity can be explained if the electron is considered to be a hollow shell (which is commensurate with the bare electron's tiny mass in comparison to the very large radius) and if the negative charge of the shell is not the source of the magnetic moment.

It has long been known that a photon can be split into an negative ion (electron) and a positive ion (positron), each having the same amount of charge but of opposite sign. Electrons and positrons can recombine into electrically neutral photons, so it is apparent that photons are composed of a positive and a negative ion. Two

ions spinning around each other could produce the photon's wave nature. The only size of photon ion that can exist as a separate entity has a charge of exactly plus one or minus one, whereas the ions can have a very much larger or very much smaller charge and mass when combined in photons, as long as the two ions are equal in charge and mass. Combined in a photon, the two ions are apparently attracted together so strongly that their individual volumes are very much smaller than as separate entities.

When a dipole photon enters an electron shell, its negative-ion portion is expected to be forced towards the shell's centre by Coulombic repulsion, while the photon's positive ion would be attracted by the negative charge of the shell equally in all directions. The negative photon ions would likely merge into a single body at the electron's centre, while the positive-ion portion would orbit around the centralised negative ion to retain the photon's angular momentum. The high peripheral velocity of this orbiting photon mass would enable portions of photon material to spin off and exit the electron shell at the same velocity at which they entered the electron, i.e., the speed of light. The orbiting of the positive photon charge at Dehmelt's small radius, most likely accounts for the magnetic moment that is observed in electrons of normal mass.

Liberated low-mass conduction electrons within intrinsic semiconductors (which are also photoconductors by their nature) and within doped semiconductors, are mostly protected against acquiring mass from ambient-heat photons by the heat-insulative properties of the semiconductors. In contrast, low-mass electrons injected into heat-conducting metals, rapidly acquire mass from ambient-heat photons by the existence of cryogenic conditions, but they are vulnerable to internal heat-photons created by excessive induction.

Conduction electrons of metals, typically move as a group at drift velocities of less than one millimetre per second, although the velocity of the electrical effects approaches the velocity of light. (Photons are probably involved in the movement of electrical energy in metallic conductors.) In contrast, conductive low-mass electrons can move individually at great velocities in superconductors and semiconductors. Brennan (1999, p. 631) reports the drift velocity of a particular electron moving in a semiconductor, to be one micrometer in about 10 picoseconds, which is equivalent to 100 kilometers per second.

The concentration of the conduction electrons in metals is the same as the number of atoms, whereas in semiconductors, the mobile low-mass electrons which are free to move, can vary greatly with the amount of certain photon radiation received. Since the magnitude of an electric current is a summation of the number of electrons involved, times their respective drift velocities, the current developed by a small ensemble of photoconducting electrons moving at high speed, can exceed the current of a much greater number of conduction electrons moving at a very low speed in a metal.

A general feature of intrinsic semiconductors is that they become photoconductive in proportion to the amount of bombardment by some particular electron-liberating frequency (or band of frequencies) of photon energy, up to some limit. The amount of bombardment by the particular wavelength (or, equivalently, the frequency), increases along with all other photon wavelengths as the ambient temperature rises, that is, as the area under Planck's black-body radiation curve increases. Consequently, the conductivity of semiconductors continues to increase with temperature, while the conductivity drops to almost zero at low temperature unless superconductivity occurs.

A single high-energy alpha particle can liberate a great number of low-mass electrons in a thin-film semiconductor, as Leimer's (1915) energy-magnifying experiment appears to show. Leimer's alpha radiation was situated near the distant end of a suspended antenna wire of unreported length, when he experienced the maximum magnetic energy increase in the coil of the ammeter in the receiver. The low-mass electrons had to have travelled the entire length of the suspended antenna and the connecting line to his receiving apparatus without encountering any trapping holes. Assuming these electrons traversed a distance of 1 to 10 metres in less than one half-cycle of the radio frequency, (that is, less than 4 microseconds at 128 kHz) at which time the direction of the low-mass electron would have been reversed, this would be equivalent to velocities of 25 to 250 km/sec.

A great number of superconducting electrons can be set in motion by inductive photon radiation. In contrast, inductive photon radiation can pass mostly through photoconductors that have low concentrations of mobile, low-mass electrons. Applicant's interpretation of Leimer's experiment is that the liberated low-mass electrons of the semiconductor coating of the antenna wire, were not directly accelerated by the inductive photons of the radio signal, but rather were accelerated to high velocities by an oscillating electric field created in the metallic wire by the radio photons.

A review of an experiment performed by File and Mills (1963), shows that the very low mass of superconducting electrons is responsible for causing supercurrents to differ from normal electric currents. A superconducting solenoidal coil (comprising a Nb-25% Zr alloy wire below 4.3^0 K.) with the terminals spot-welded together to make a continuous conductor, was employed. Extremely slow declines of induced supercurrents were observed, which can be attributed to an enormous increase in the coil's self-induction. Because a supercurrent approaches its maximum charge asymptotically when discharging, a convenient measure of the coil's charging or discharging rate is the "time-constant". The time-constant has the same value for both charging and discharging, and it is defined as (a) the time needed for charging the coil to 63% of the maximum amount of current inducible in the coil by a given diamagnetic force, or (b) the time needed to discharge 63% of the coil's induced current.

In normal conductors, the inductive time-constant is calculated by the inductance of the coil, divided by the resistance of the coil. By use of an empirical equation, the inductance of the coil in its non-superconducting state is calculated to be 0,34 Henry, based on a double-layered solenoid of 384 turns that measured 4 inches (10 cm) diameter and 10 inches (25 cm) long. The resistance of the 0.020 inch (0.51 mm) diameter wire at a temperature of 5^0 K. (just above T_c) is estimated by using data for Zr alone, to be 4×10^2 ohms. (Resistivity data were not available for Nb or the subject alloy). Under non-superconducting conditions, the time-constant for charging and discharging this coil is thereby calculated to be approximately 8×10^{-5} sec.

The time it took to charge up a supercurrent in the coil in the experiment was not reported. But, based on the reported 50 re-energisings and magnetic determinations performed in 200 hours, the measured charging time in the superconducting state is computed to be no more than 4 hours on average.

Using Bardeen's (1941) formula of m is approximately equal to m_e times 10^{-4} for the order of magnitude of the low T_c superconducting electron's mass, and using Larmor's equation (1900) which relates inductive radiation power to the square of the acceleration of the charge, the inductance of the coil is expected to increase by $(10^4)^2 = 10^8$ times in the superconducting state. Thus, the calculated increase in the time-constant of charging up the supercurrent is $8 \times 10^{-5} \times 10^8$ which equals 8×10^3 seconds, or 2.2 hours, which is the same order of magnitude as the maximum actual charging time. The self-induction increased by that amount because the low-mass electrons are accelerated 10^4 times faster.

In the case of discharging, the time constant of the supercurrent was projected by File and Mills from measured declines observed over periods of 21 and 37 days. The projections of the two 63% declines agreed closely at 4×10^{12} seconds ($= 1.3 \times 10^5$ years). Therefore, the time-constant of supercurrent discharge, based on projecting actual measurements, had increased by 5×10^{16} times over the time-constant for electrons of normal mass.

The driving force during charging, had been the applied inductive force, whereas the driving force during discharging was the supercurrent that had been magnified 10^8 times. Therefore, during the discharging of the supercurrent, the time-constant is increased again by 10^8 times, so the calculated total increase in the time-constant of discharge is $10^8 \times 10^8 = 10^{16}$ times greater than the normal time-constant. This calculated value of the non-superconducting time-constant, based solely on the increase of inductive radiation due to extremely low electron mass, compares favourably in magnitude with the actually observed value of 5×10^{16} times the normal time-constant.

The superconducting coil required no more than four hours to charge up the supercurrent, yet during subsequent discharge, the superconducting coil was projected to radiate inductive photon energy from the centripetal acceleration of the superconducting electrons for 130,000 years before declining by 63%. If this experiment could take place where no energy would needed to sustain critical cryogenic conditions, as in outer space, the lengthy discharge of this energised coil would clearly demonstrate the creation of energy in the form of newly-created photons inductively radiating from the superconducting low-mass electrons that circulate around the coil's loops. Applicant interprets this as showing that low-mass electrons are capable of inductive-energy-magnification based solely on their mass relative to that of normal electrons.

In the embodiments described below, the magnified inductive energy of low-mass electrons is utilised in coils for electric-energy generation by employing a flow of inductively accelerated photons that alternates in direction. This, in turn, drives low-mass electrons in an oscillating manner, so this forced reversal involves only a single stage of inductive-energy magnification, rather than the two stages (charging and naturally discharging) in the foregoing experiment.

Mode of Operation

Inductive photons radiating from an oscillating electric current in a sending conductor (e.g. from a radio-wave broadcasting antenna) convey a force, on conduction electrons in a receiving conductor, that is transverse to the incidence direction of the incident inductive photons on the receiving conductor. As a result, no back-force is transferred directly back to the sending conductor. Applicant has discovered that the action of this transverse force on low-mass electrons in a receiving conductor is analogous to the action of Gauss's transverse magnetic force on free electrons in a conductor, which is not subject to the kinetics law of conservation of energy. If the receiving conductor has low-mass conduction electrons, then this transverse force would impart greater acceleration to the low-mass electrons than that it would impart to normal free electrons. The resulting greater drift velocities of low-mass electrons than normal free electrons in the receiving conductor, would yield an increased magnitude of inductive force produced by the low-mass electrons in the receiving conductor and hence produce a magnification of the irradiation energy of inductive photons.

The direction of the transverse force imparted by the radiated inductive photons on conduction electrons in the receiving conductor is opposite to the direction of the corresponding electron flow in the sending conductor. This

relationship is similar to the inductive force on electrons in the secondary coil of a transformer, which also is opposite to the direction of flow of electrons in the primary coil.

Various embodiments of Applicant's electrical generator employ inductive photons radiated from electrical oscillations in a "sending coil". Inductive photons are radiated from the sending coil toward and inductive-photon receiving coil, termed an "energy-magnifying coil", which comprises a photoconductive or superconductive material, or other suitable material as described below. The energy-magnifying coil is placed in a condition favourable for the production of low-mass electrons that participate in electrical conduction in the energy-magnifying coil. For example, if the energy-magnifying coil is made of photoconductive material, the coil is provided with a photoconduction exciter. Alternatively, if the energy-magnifying coil is made of a superconductive material, the coil is placed in an environment at a temperature (T) no greater than the critical temperature (T_C); i.e., $T < T_C$. In the former example, the photoconduction exciter can be a source of illumination which provides an appropriate wavelength of excitive electrokinetic radiation. If the energy-magnifying coil is comprised of a doped semiconductor, the condition that provides mobile low-mass electrons already exists.

In the energy-magnifying coil, the greater-than-normal acceleration of the low-mass electrons produces greater-than-normal inductive forces in the form of greater-than-normal radiation of inductive photons from the coil. The resulting increased inductive-photon energy from the photoconductor or superconductor is converted into useful electrical energy in an output coil inductively coupled to the energy-magnifying coil. The output coil can be made of insulated metallic wire. An exemplary output coil is situated coaxially with, and nested within, the energy-magnification coil. A coil of this type is termed herein, an "internal output coil".

The ability of the subject apparatus to produce more energy output than energy input, is based on the output coil receiving more of the magnified energy from the energy-magnifying coil than is returned as a back-force from the output coil to the energy-magnifying coil. This principle is termed herein "energy leverage".

The oscillations in the energy-magnifying coil are initiated by an external energy-input source that provides an initiating impulse of electron flow in the sending coil. For example, the external energy-input source can be an adjacent independent electromagnet or an adjacent permanent magnet moved rapidly relative to the sending coil. The initiating impulse starts an oscillation in the sending coil that stimulates radiation of inductive photons from the sending coil to the energy-magnifying coil. Energy from the external energy-input source is magnified by the apparatus so long as the energy-magnifying coil does not act as an independent oscillator at a different frequency. Independent oscillation is desirably avoided by connecting the ends or terminals of the energy-magnifying coil to each other in such a way that it results in one continuous coil, or a continuous multiple-coil system or systems, connected together in such a way that continuity exists for the conduction of low-mass electrons throughout the entire coil system. The energy-magnifying coil inductively creates more energy in the output coil than the energy of the initial impulse. The resulting magnified output of electrical energy produced by the apparatus is available for useful purposes in a work loop.

After initiation, the apparatus is made self-sustaining using a feed-back loop arranged in parallel with the work loop that includes the sending coil, and with a capacitor located in the feed-back loop to make it an L-C circuit, i.e., after start-up of the apparatus using the external energy-input source, the apparatus becomes self-resonating, which allows the external energy-input source to be decoupled from the apparatus without causing the apparatus to cease production of electrical energy.

During normal self-sustained operation, a portion of the output electrical energy is returned to the sending coil by the feed-back loop, thereby overcoming the need to use the external energy-input source for sustaining the oscillations in the sending coil. In other words, after startup, the external energy which was used by the sending coil to excite the photoconductive material or the superconducting material in the energy-magnifying coil is replaced by a portion of the output energy produced by the apparatus itself. The remainder of the output electrical energy is available in the work loop for useful purposes.

Initiating the generation of electrical energy by the apparatus, takes advantage of the fact that the inductive back-force sent from the output coil to the energy-magnifying coil (and hence ultimately, back to the sending coil), arrives at the sending coil one cycle behind the corresponding pulse that initiated the flow of electrons. This one-cycle lag of the back-force, as well as a corresponding one-cycle lag in the feed-back, enables small starting pulses produced in the sending coil to produce progressively greater electrical outputs each successive cycle. Consequently, assuming that the electrical load is not excessive during start-up, only a relatively few initiating cycles from the external energy-input source typically are needed for achieving production by the apparatus of an amount of output power sufficient to drive the load as well as providing sufficient energy feed-back to the sending coil in a sustained manner.

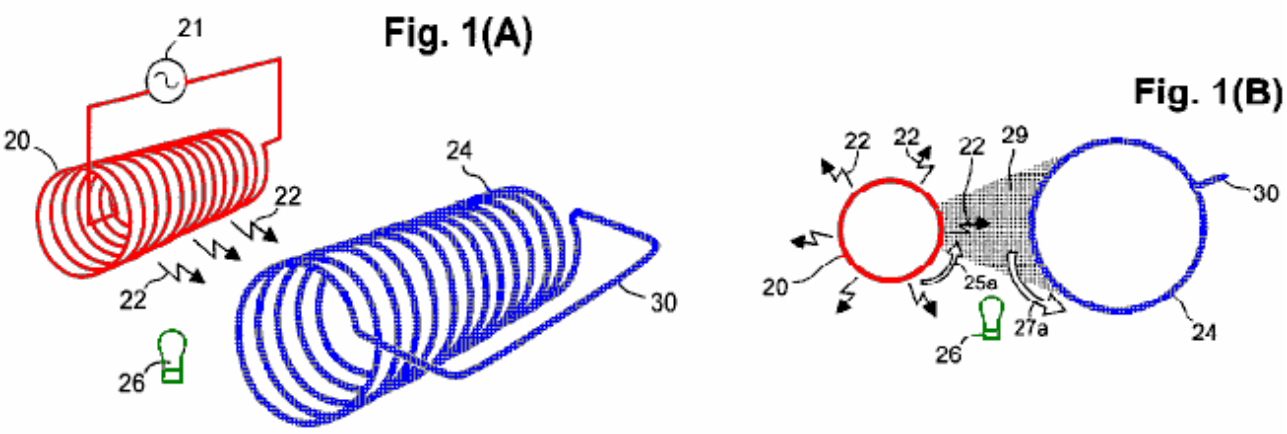
A half-cycle of the one-cycle lag occurs between an initial acceleration of electrons in the sending coil and a corresponding initial oscillation in the energy-magnifying coil. This half-cycle lag occurs because induction photons are not radiated from the initial acceleration of electrons in the sending coil, but rather are radiated when the electrons are reverse-accelerated. (Kramers, 1923, and Compton and Allison, 1935, p.106). As the newly formed photons are being radiated by the respective deceleration of electrons in the sending coil, even more new photons are simultaneously being formed by the new direction (i.e. reverse direction) of acceleration under

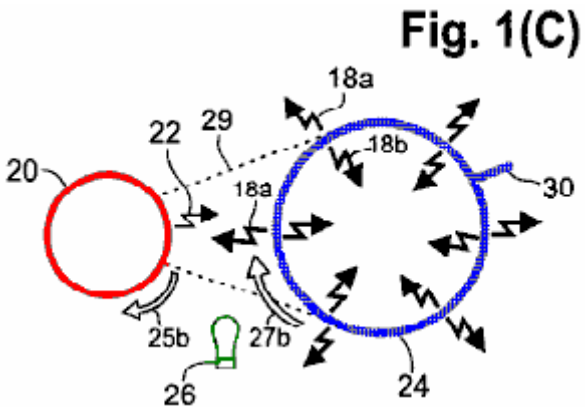
oscillating conditions. Thus, the radiation of photons from electrons alternatingly accelerated in the opposite direction from the conveyed force, continues each half-cycle after the initial half-cycle.

Applicant also discovered that a half-cycle lag also occurs between the initial flow of electrons in the primary coil of a certain type of transformer, which is simply comprised of coils nested coaxially rather than being inductively coupled by an iron core, and the resulting electron flow induced in the secondary coil. When applied to this apparatus, these finding indicate that a second half-cycle lag occurs between the acceleration of low-mass electrons in the energy-magnifying coil and the corresponding electron flow induced in the output coil. The feed-back from the output coil boosts the electron flow in the sending coil one whole cycle after the initial pulse.

As discussed above, the energy-magnifying coil comprises either a photoconductor, a doped semiconductor or a superconductor as a source of, and as a conductor of, low-mass electrons. The general configuration of the coil is similar in either case. The coil including a photoconductor or doped semiconductor, has an operational advantage at normal temperatures, and the coil including a superconductor has an operational advantage at sub-critical temperatures ($T < T_c$), such as in outer space.

Representative Embodiments





Reference is now made to **Fig.1A** to **Fig.1C** and **Fig.2A** and **Fig.2B** which depict a sending coil **20** connected to a source of alternating current **21**. The sending coil is shown having a desirable cylindrical profile, desirably with a circular cross-section as the most efficient configuration. In **Fig.1A** and **Fig.1B**, electrical oscillations from the source **21** are conducted to the sending coil **20** where they cause inductive photons **22** to radiate from the sending coil. The radiated photons **22** convey transverse forces in the same manner that a radio-broadcasting antenna transmits oscillating energy. The sending coil **20** can be a single layer or multiple layers of insulated metal wire (e.g. insulated copper wire). One layer is sufficient, but an additional layer or layers may increase operational efficiency. If necessary, or desired, the turns of wire can be formed on a cylindrical substrate made of a suitable dielectric.

The inductive photons **22** radiating from the sending coil **20**, propagate to an energy-magnifying coil **24** that desirably has a cylindrical profile extending parallel to the sending coil. In the embodiment shown in **Fig.1A** and **Fig.1B**, the energy-magnifying coil **24** does not terminate at the ends, but rather, it is constructed with a connector **30** to form a continuous conductor. The energy-magnifying coil **24** desirably is a helical coil made of a material comprising a photoconductive or superconductive material, or other suitable material. If necessary or desired, the energy-magnifying coil can be formed on a substrate which, if used, desirably is transmissive to the inductive-photon radiation produced by the coil.

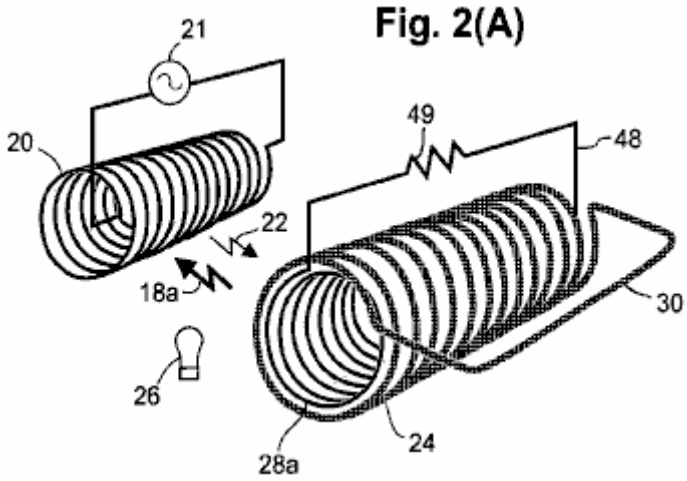
In an energy-magnifying coil **24** made of a superconducting material, a large population of conductive low-mass electrons is produced in the coil by lowering the temperature of the coil to a point below the critical temperature for that material. By way of an example, sub-critical temperatures are readily available in outer space or are produced under cryogenic conditions.

In an energy-magnifying coil **24** made of a photoconductor material, a large population of conductive low-mass electrons is produced in the coil by illuminating the coil with photons of an appropriate wavelength, such as photons produced by a photoconduction exciter **26**. The photoconductor exciter **26** desirably is situated and configured so as to illuminate substantially at least the same side of the energy-magnifying coil **24** that receives inductive photons **22** radiating directly from the sending coil **20**. Alternatively, the photoconduction exciter **26** can be situated and configured so as to illuminate all sides of the energy-magnifying coil **24**. In the depicted embodiment, the photoconduction exciter **26** can be at least one incandescent lamp (as shown) energised by conventional circuitry (not shown). Alternatively, the photoconduction exciter **26** can be at least one gas-discharge lamp or one or more Light Emitting Diodes. The wavelength produced by the photoconduction exciter **26** can be, for example, in the infrared (IR), visible, ultraviolet (UV), or X-ray range as required by the particular photoconductor material in the energy-magnifying coil **24**. Another possible form of the photoconduction exciter **26**, is a source of photons in the gigahertz or the terahertz portion of the electrokinetic spectrum. Other photoconduction exciters are configured, as required, to produce a suitable wavelength from the radio-wave portion of the electrokinetic spectrum. The illumination can be either direct from the photoconduction exciter **26** to the energy-magnifying coil **24** or conveyed from a remotely located photoconduction exciter to the energy-magnifying coil via optical fibres, light pipes, or the like.

Fig.1B and **Fig.1C** are respective orthogonal end views of the sending coil **20** and energy-magnifying coil **24** shown in **Fig.1A**. The radiation of inductive photons **22** from the sending coil **20**, is indicated schematically in **Fig.1A**, **Fig.1B** and **Fig.1C** by small, jagged arrows. The forces delivered by the photons **22** to the conductive low-mass electrons in the energy-magnifying coil **24**, alternate in directions which are opposite to the respective directions of simultaneous electron flow in the sending coil **20**. Whenever the particular oscillation phase of electron flow in the sending coil **20** is in the direction of the curved arrow **25a** adjacent to the sending coil **20** in **Fig.1B**, the resulting transverse photon force causes a flow of low-mass electrons in the energy-magnifying coil **24**, depicted by the curved arrow **27a** adjacent to the energy-magnifying coil **24**.

The shaded sector **29**, shown in **Fig.1B**, denotes the proportion of inductive-photon radiation **22** from the sending coil **20**, actually received by the single energy-magnifying coil **24** shown, compared to the entire 360-degree radiation of inductive photons **22** from the sending coil **20**. Aside from a small amount of inductive-photon radiation lost from the ends of the sending coil **20**, the relative amount of the total energy of inductive-photon radiation received by the energy-magnifying coil **24** is determined by the angle subtended by the energy-magnifying coil **24**, relative to the entire 360 degrees of inductive-photon radiation from the sending coil **20**.

In **Fig.1C**, the low-mass conduction electrons of the energy-magnifying coil **24** are accelerated to a higher drift velocity than normal free electrons in the energy-magnifying coil **24** would be. As noted above, the sending coil **20** is energised by alternating electron flow, which causes a periodic reversal of direction of electron flow in the sending coil **20** (compare the direction of the arrow **25b** in **Fig.1C** with the direction of the arrow **25a** in **Fig.1B**). Each reversal of direction of electron flow in the sending coil **20**, causes a corresponding reversal in the direction of acceleration of the low-mass electrons in the energy-magnifying coil **24** (compare the direction of the arrow **27b** in **Fig.1C** with the direction of arrow **27a** in **Fig.1B**). Each such reversal in direction of acceleration causes a corresponding radiation of inductive photons (jagged arrows **18a**, **18b**) radially outwards and radially inwards, respectively, from the energy-magnifying coil **24**. Note that the arrows **18a** and **18b** are larger than the arrows denoting the inductive photons (arrows **22**) from the sending coil **20**. This symbolically denotes energy magnification. Note also that, of the magnified inductive-photon energy radiating from the energy-magnifying coil **24**, substantially half is directed inwards (arrows **18b**), and substantially the other half is radiated outwards (arrows **18a**).



Turning now to **Fig.2A**, the sending coil **20**, and the energy-magnifying coil **24**, are shown. The energy-magnifying coil **24** in **Fig.2A** includes an internal output coil **28a**, that desirably is situated co-axially inside and is of the same length as the energy-magnifying coil **24**. A work loop **48** can be connected to the ends of the internal output coil **28a**, thereby forming an electrical circuit in which a load **49** is indicated symbolically as a resistor. The internal output coil **28a** and the conductors of the work loop **48**, desirably are made of insulated metallic (e.g. copper) wire.

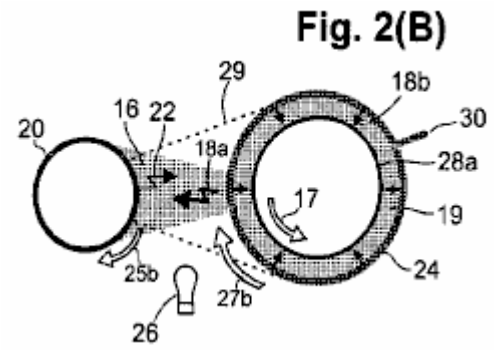


Fig.2B depicts a transverse section of the coils shown in **Fig.2A**. In **Fig.2B**, the magnified inductive-photon energy (shaded area **19**) produced by the energy-magnifying coil **24** and directed radially inwards towards the internal output coil **28a**, induces a corresponding oscillating electron flow in the internal output coil **28a**. Thus, the work loop **48** connected across the internal output coil **28a**, is provided with greater energy than was received by the energy-magnifying coil **24** from the sending coil **20**. The direction of the electron flow (arrow **17**) in the internal output coil **28a**, is opposite to the direction of flow (arrow **27b**) in the energy-magnifying coil **24**, which in turn is opposite to the direction of electron flow **25b** in the sending coil **20**.

In **Fig.2B**, the annular-shaped shaded area 19 between the energy-magnifying coil **24** and the internal output coil **28a**, indicates that substantially all of the internally-directed magnified inductive-photon energy (i.e. approximately half of the total radiation energy) from the energy-magnifying coil **24**, is directed to, and captured by, the internal output coil **28a**. In contrast, the shaded sector **16** extending from the energy-magnifying coil **24** to the sending coil **20**, indicates that a relatively small proportion of the outwardly directed magnified radiation **18a** from the energy-magnifying coil **24** is directed to the sending coil **20** where the radiation provides a corresponding back-force. Aside from the small amount of inductive-photon radiation lost from the ends of the energy-magnifying coil **24**, the relative amount of the magnified inductive-photon radiation (sector **16**) providing the back-force on the sending coil **20**, is a function of the angle subtended by the sector **16**, compared to the 360-degree radiation from the energy-magnifying coil **24**.

The ratio of magnified energy **18b** from the energy-magnifying coil **24** and received by the internal output coil **28a**, to the magnified energy **18a** received as a back-force by the sending coil **20**, denotes the energy “leverage” achieved by the subject apparatus. If this ratio is greater than unity, then the energy output from the internal output coil **28a** exceeds the energy input to the energy-magnifying coil **24**. This energy leverage is key to the self-sustained operation of the apparatus, especially whenever the apparatus is being used to drive a load. In other words, , with a sufficiently large energy-magnification factor achieved by the energy-magnifying coil **24**, the electrical energy available in the work loop **48**, exceeds the input energy that produces the oscillations in the sending coil **20**. The electric power input to the sending coil **20** thereby produces magnified electric power in the internal output coil **28a** that can perform useful work in the work loop **48** while self-powering the continued operation of the apparatus.

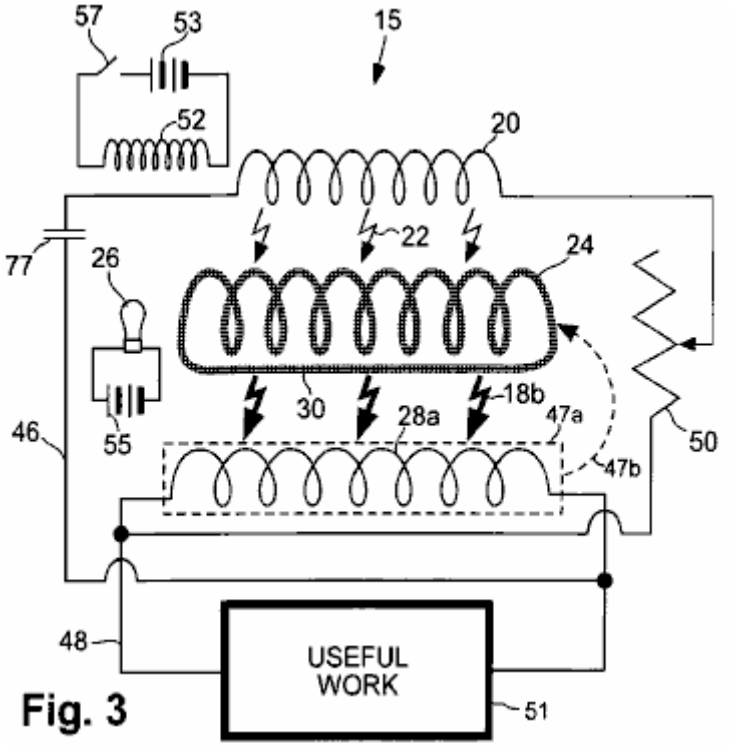


Fig. 3

Reference is now made to **Fig.3**, which schematically depicts aspects of the apparatus **15**, responsible for self-generation of electric power by employing a feed-back loop **46**. The conductors of the feed-back loop **46** can be made of insulated metallic wire. (In **Fig.3**, the dotted lines **47a** and dotted arrow **47b**, indicate that the internal output coil **28a** is actually positioned co-axially inside the energy-magnifying coil **24**, as described above, but is depicted in the figure as being outside the energy-magnifying coil for ease of illustration). The feed-back loop **46**, conducts a portion of the electric power from the internal output coil **28a**, back to the sending coil **20**. The remaining portion of the electric power from the internal output coil **28a** is directed to the work loop **48** where the power is utilised for useful work **51**. The relative proportions of output power delivered to the feed-back loop **46** and to the work loop **48**, can be varied by adjusting a variable resistor **50**.

As noted above, an initial source of electrical energy is used for "starting" the apparatus **15** by initiating an oscillation in the sending coil **20**. After starting, under usual operating conditions, the apparatus **15** is self-resonant and no longer requires the input of energy from the initial source. The particular inductance and distributed capacitance of the sending coil **20**, plus all other capacitances and inductances in the apparatus, provide a certain corresponding frequency of self-resonating oscillation. In the feed-back loop **46** is a capacitor **77** that makes the apparatus an L-C circuit which oscillates at its own frequency. the frequency can be changed by altering the capacitance or inductance of the apparatus, or both. the capacitor **77** can be a variable capacitor by which the frequency can be adjusted.

As shown in **Fig.3**, the initial source of oscillating electrical energy can be an impulse from an external electromagnet **52** powered by its own energy source (e.g. a battery **53** as shown, or other DC or AC source). For example, the electromagnet **52** can be placed near the sending coil **20** or other portion of the feed-back loop **46**, and energised by a momentary discharge delivered from the battery **53** via a switch **57**. The resulting pulse generated in the electromagnet **52**, initiates a corresponding electrical pulse in the sending coil **20** that initiates self-sustaining oscillations in the apparatus **15**. In another embodiment, the electromagnet **52** can be energised briefly by an AC source (not shown). In yet another embodiment, the initial source can be a permanent magnet which is moved rapidly (either mechanically or manually) near the sending coil **20** or other portion of the feed-back circuitry. In any event, the pulse provided by the initial source initiates electrical oscillations in the sending coil **20** that produce corresponding oscillating inductive-photon radiation **22** from the sending coil **20**, as shown schematically in **Fig.3** by thin jagged arrows. The inductive-photon radiation **22** from the sending coil **20** causes, in turn, re-radiation of magnified inductive-photon energy **18b** from low-mass electrons in the energy-magnifying coil **24**, as shown schematically in **Fig.3** by thick jagged arrows. **Fig.3** depicts a photoconductive energy-magnifying coil **24** which is illuminated by an incandescent photoconduction exciter **26** energised by its own power source **55** (e.g., an externally connected battery as shown).

A sufficiently high energy-magnification factor of the apparatus **15** allows the magnified energy from the energy-magnifying coil **24** to induce greater energy in the internal output coil **28a** than the energy of the corresponding initial pulse. A portion of the magnified electrical energy is returned to the sending coil **20** via the feed-back loop **46** to sustain the oscillations.

The remaining surplus energy from the internal output coil **28a** is available for application to useful work via the work loop **48**. In one embodiment, some of this useful work can be used for illuminating the photoconduction exciter **26** (circuitry not shown) in an apparatus configuration in which the energy-magnifying coil **24** comprises a photoconductor. In another embodiment, some of this useful work can be used for maintaining cryogenic ($T < T_c$) conditions for an apparatus configuration in which the energy-magnifying coil **24** comprises a semiconductor.

After starting oscillations in the apparatus **15**, electron flow builds up rapidly, so long as the load **49** does not draw off too much of the output energy during start-up. Upon reaching operating equilibrium, the output of electrical power from the apparatus **15** is a rapidly alternating current (AC). The AC output can be rectified by conventional means to produce direct current (DC), and the output can be regulated as required, using conventional means. Many variations of conventional circuitry are possible, such as, but not limited to, automatic voltage controllers, current controllers, solenoidal switches, transformers, and rectifiers,

Regarding the energy-magnifying coil **24**, an exemplary embodiment can be made from a low - T_c superconductor such as commercially available, flexible, niobium-zirconium wire which can be readily formed into a coil.. Other embodiments, as noted above, of the energy-magnifying coil **24** can be made using a photoconductive material or a high - T_c superconductor. Most high - T_c superconductors (and some photoconductors) have ceramic-like properties and thus require the application of special methods for forming the material into a cylindrical coil having electrical continuity throughout. Some commercially available high - T_c superconductors are available in ribbon or tape form. The energy-magnifying coil **24** can be free-standing or supported on a rigid substrate.

By way of example, an energy-magnifying coil **24** can be made from a ribbon of flexible photoconductive material such as the material discussed in patent US 6,310,281, incorporated herein for reference. Briefly, a layer of stress-compliant metal is placed on a plastic ribbon. Then the photoconductive material is deposited on both sides of the metal-covered ribbon and the edges of the ribbon so that the ribbon is coated all the way around. Such a configuration allows low-mass electrons in the photoconductive material, to receive energy from inductive-photons emitted from the sending coil **20** on one side of the ribbon while re-radiating magnified energy from both sides of the ribbon.

In another example, a flexible photoconductor ribbon is made from flexible organic polymer having photoconductive properties. (High electrical conductivity observed in photoconductive polymers is attributed to the presence of low-mass electrons in the material). The flexible photoconductive ribbon can be wound on a dielectric tubular support, to form the energy-magnifying coil 24.

In yet another example, a thick-film coating of photoconductive cadmium sulphide (CdS) or cadmium selenide (CdSe) is formed on a wire coil by sintering as paste, which comprises a powder of finely ground CdS or CdSe crystals mixed with water and at least a fluidiser such as cadmium chloride, at a temperature of 550^0 C. to 600^0 C. in a controlled atmosphere. During sintering, the boundaries of the small crystals become melted with the heated fluidiser, allowing the crystals to regrow together and solidify when the fluidiser evaporates and the sintered coating is cooled. Alternatively, copper oxides are formed in place on bare copper or bronze wire by heating the wire above about 260^0 C. in an oxygen atmosphere, or by application of chemical oxidants.

In yet another example, a coil of ceramic-like superconductor or photoconductor is made by tape-casting, extruding, slip-casting, cold or hot-pressing, or coating of the material as a thin film arranged helically on a tubular dielectric substrate. The assembly is heat-treated in a controlled atmosphere furnace to increase inter-crystalline contacts. Alternatively, the thin film of superconductor or photoconductor is formed over the entire exterior of the dielectric substrate, followed by removal of selected portions of the superconductor or photoconductor to form the desired helical coil.

[121]In some photoconductors and doped semiconductors, only a small portion of a population of inductive photons irradiated on the material, impact with, and yield acceleration of, low-mass electrons in the material. This is due to a low density of photoconductive low-mass electrons in the material. In such as case, inductive-photon radiation passing through the material can be captured efficiently by normal free conduction electrons in a metallic strip that desirably is in immediate contact with, or embedded in, the material. The acceleration of normal free electrons in the metallic conductor, sets up an electric field that assists in accelerating the low-mass photoelectrons. In this configuration, it is desirable that the photoconductive material be disposed completely over and around the metallic strip so that the photoconductor faces both outwards and inwards, with both sides of the photoconductor or doped semiconductor being in electrical contact with each other.

One factor in the choice of photoconductor material to use in forming the energy-magnifying coil **24** is the potential magnification of energy that can be realised by low-mass electrons of an n-type or p-type photoconductive material. Other important factors are the quantity of low-mass electrons that are available in the photoconductive material for a given amount of illumination and the actual electrical conductance of the material. Standard illumination-sensitivity measurements provide a general overall index of the ability of a photoconductor to serve effectively in magnifying energy.

Cadmium sulphide and cadmium selenide, the most common photoconductive compounds which are available commercially, have calculated magnification factors of 37 and 59, respectively. The peak response wavelength of cadmium sulphide is 515 nanometers (in the green part of the visible spectrum) and of cadmium selenide is 730 nanometers (in the near-infrared part of the spectrum). Cadmium sulphide can be mixed with caesium selenide under certain conditions, so the resulting mixture assumes photoconductive characteristics between those two values. Mixtures can be produced having peak wavelengths which are matched to the wavelengths of commercially available LEDs of many sizes and illumination intensities. Some semiconductors which become photoconductive at a wavelength smaller than the wavelength produced by currently available LEDs can be made conductive of low-mass electrons merely by heating.

Applicant has found that gallium arsenide develops considerably higher conductivity than copper or silver at a temperature of 100^0 C. and that the conductive electrons are low-mass. Also, alpha radiation is capable of liberating many low-mass electrons in some semiconductors. A second electron of comparatively low mass may have been liberated from cupric oxide by alpha radiation along with the outer copper electron in Leimer's (1915) experiments, since the measured energy magnification exceeded the magnification calculated from cyclotron resonance of CuO, which most likely pertains only to the mass of the outer electron.

Dopants can be added to a semiconductor to make it more conductive of low-mass electrons without illumination. Also, the illumination-sensitivity and conductivity of cadmium sulphide are increased by adding small amounts of donor-type dopants such as, but not limited to, sulphides, selenides, tellurides, arsenides, antimonides and phosphides of the Type-IIIa elements: aluminium, gallium, indium and thallium. In this regard, the photoconductors of high-sensitivity photovoltaic cells may comprise as many as five different compounds. The actual mixtures of photoconductive compounds and dopants used in commercially available photovoltaic cells often are trade secrets. But, the sensitivity and conductances of the cells are usually given or are measurable, and these data can be used advantageously in selecting a particular photoconductive compound for use in the apparatus.

Other photoconductive compounds or elements can be employed in energy-magnifying coils. For example, the conduction electrons of silicon have an energy-magnification factor of 15 times. Photoconductors having very high magnification factors include, but are not limited to, gallium arsenide, indium phosphide, gallium antimonide, cadmium-tin arsenide, and cadmium arsenide, which have calculated energy-magnification factors ranging between 200 times and 500 times, and mercury selenide (1100 times), indium arsenide (2000 times), mercury telluride (3400 times) and indium antimonide (5100 times).

The depth of optical transmission largely determines the optimum thickness of photoconductive films for energy-magnifying coils. For example, the highest optical transmission of sintered CdS is reported to be 20 micrometers, but since the average grain size increases (and the average porosity decreases) with an increase in film thickness, the maximum conductivity of a sintered film is at a thickness of 35 micrometers (J. S, Lee et al., 1987).

The metal chosen to be embedded must not react chemically with the photoconductor. For example, aluminium reacts with gallium arsenide (GaAs) in an electrical environment, to change the conductive character of both the GaAs and the aluminium. Gold, platinum, and palladium can serve in many cases because these materials are relatively inert chemically. Gold combines chemically with tellurium, however, so gold is not suitable for embedding in mercury telluride. Cadmium plating over a common metal serves to alleviate the reactivity in cases where cadmium sulphide or cadmium selenide is used as the photoconductor.

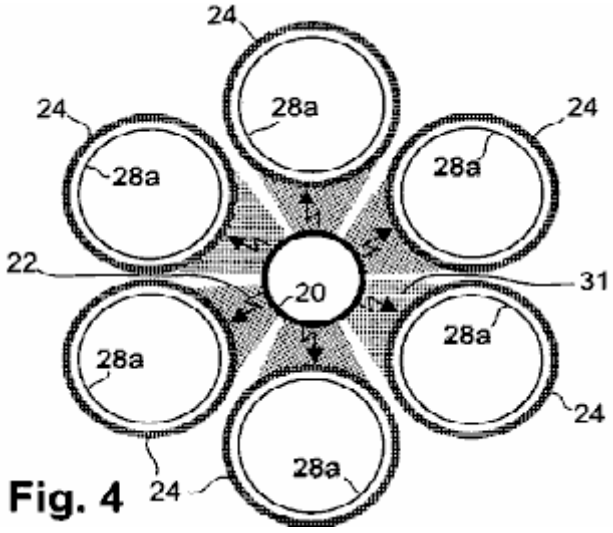


Fig. 4

The discussion above has been, for ease of explanation, in the context of the apparatus including one energy-magnifying coil **24**. However, as discussed, use of a single energy-magnifying coil **24** to capture inductive photons from the sending coil **20**, results in loss (by non-capture) of most of the inductive photons from the

sending coil **20**. This proportion of captured inductive photons can be increased greatly in an embodiment in which multiple energy-magnifying coils **24** substantially completely surround the sending coil **20**, such as shown in **Fig.4**. In this embodiment, the energy-magnifying coils **24** substantially completely surround the sending coil **20**, and (although six energy-magnifying coils **24** are shown) as few as three energy-magnifying coils **24** of adequate diameter, still could substantially completely surround the sending coil **20**. There is no limit, except as possibly related to packaging concerns, to the maximum number of energy-magnifying coils **24** which could be used. The depicted configuration of **Fig.4**, has a desirable number of six energy-magnifying coils **24**. In **Fig.4**, the shaded sectors **31**, considered collectively, illustrate that nearly all 360 degrees of inductive-photon radiation **22** from the sending coil **20**, are received by the energy-magnifying coils **24**. Not shown in **Fig.4** are photoconduction exciters (items **26** in **Fig.3**) used for illuminating respective portions of the energy-magnifying coils **24** in a photoconductive form of the apparatus **15**.

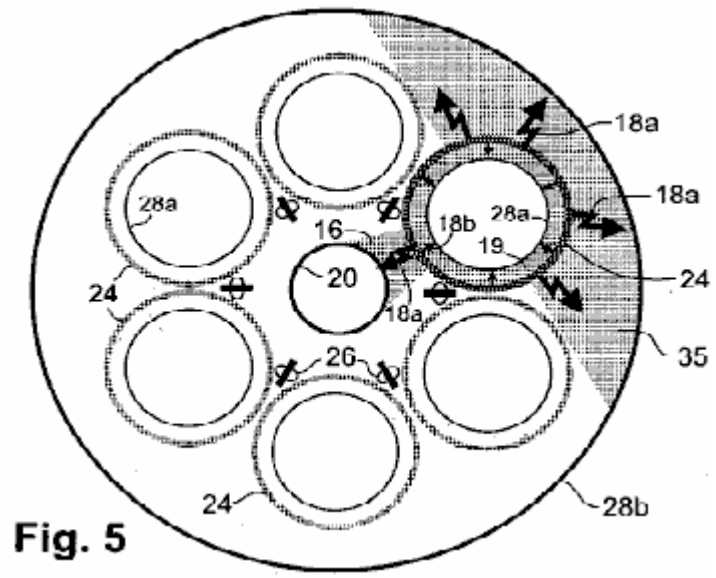


Fig. 5

Fig.4 also depicts respective internal output coils **28a** nested co-axially and co-extensively inside each of the energy-magnifying coils **24**. As discussed earlier, each internal output coil **28a** receives nearly all the inductive-photon radiation propagating radially inwards from the respective energy-magnifying coil **24**. Desirably, the overall energy output of the embodiment of **Fig.4**, can be increased by surrounding the array of energy-magnifying coils **24** with an external output coil **28b**, of which the conductors desirably are made of insulated metallic wire (**Fig.5**). In this embodiment, approximately half of the outwardly propagating magnified inductive-photon radiation (large arrows **18**) from each energy-magnifying coil **24** (one such coil is highlighted in **Fig.5**) is received by the external output coil **28b**. This captured radiation is denoted by the shaded sector **35**. When this externally directed inductive radiation captured from all the energy-magnifying coils **24** is added to all the inwardly directed radiation captured from the energy-magnifying coils **24** by their internal output coils **28a**, **28b**, greatly exceeds the back-force energy directed by the energy-magnifying coils **24** towards the sending coil **20** (the back-force energy from one energy-magnifying coil **24** is shown as the shaded sector **16**). Thus, the resulting energy "leverage" exhibited by the apparatus is increased substantially by including the external output coil **28b**.

The embodiment of **Fig.5** also includes respective arrays (viewed endwise) of light-emitting diodes (LEDs) collectively serving as photoconductor exciters **26** for the energy-magnifying coils **24**. The LED arrays are arranged back-to-back and disposed between adjacent energy-magnifying coils **24**. Each array in **Fig.5** can comprise multiple LEDs or as few as one LED.

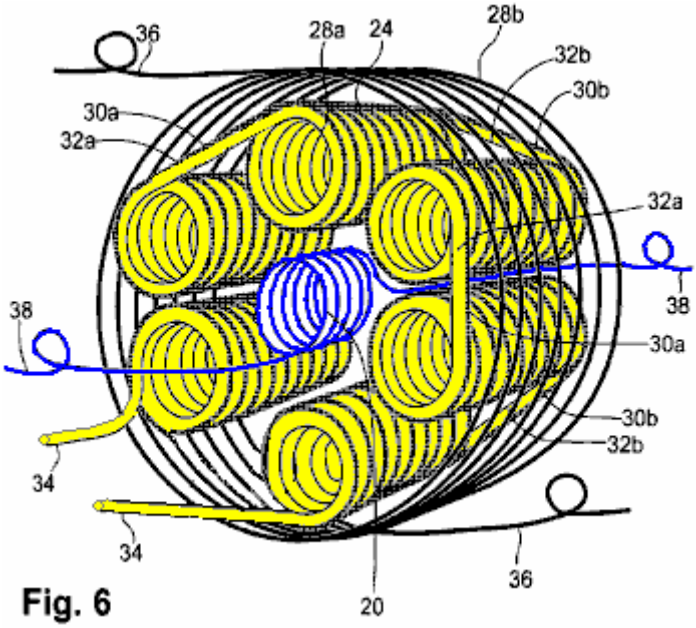


Fig. 6

Fig.6 provides a perspective view of an apparatus **15** having an arrangement of coils similar to the arrangement shown in **Fig.5**. In **Fig.6**, each energy-magnifying coil **24** comprises a helical coil of superconductive or photoconductive material in wire or ribbon (tape-like) form.

Whenever multiple energy-magnifying coils **24** are used, the respective directions of electron flow in them desirably occur in the same circular direction as viewed endwise. Thus, the flow of electrons in all the energy-magnifying coils **24** is clockwise during one phase of an oscillation cycle and counterclockwise during the other phase. The same principle applies to the flow of electrons in the output coils **28a**, **28b**. (But, in such an embodiment, the flow of electrons in the output coils **28a**, **28b**, is in the opposite direction to the electron flow in the energy-magnifying coils **24**). These relationships of electron flow in the coils during a particular phase of an oscillation cycle, are shown in **Fig.7**.

The energy-magnifying coils **24** desirably are connected together in series, using inter-coil connectors **30a**, **30b**, to maintain the same direction of electron flow, which can be clockwise or counterclockwise (as viewed from one end of such a coil). This direction of electron flow in a coil is termed the “handedness” of the coil. If the energy-magnifying coils **24** all have the same handedness, then the ends of adjacent energy-magnifying coils **24** are connected together in a head-to-foot manner progressively in one direction around the group of coils (not shown). (“Head” refers to the forward-facing end, and “foot” refers to the rearward-facing end of the apparatus in relation to the viewer). In this case, the inter-coil connectors **30a**, **30b**, must pass either completely through the apparatus or around the outside of the apparatus for its entire length, which reduces efficiency and can cause undesirable wear if the connectors are subjected to vibrations. A more desirable arrangement is depicted in **Fig.6**, in which short inter-coil connectors **30a**, cross directly head-to-head between one energy-magnifying coil **24** and an adjacent energy-magnifying coil **24**, and short inter-coil connectors **30b** cross over directly foot-to-foot in the next energy-magnifying coils **24**. In this configuration, the handedness of turns of the energy-magnifying coils **24** alternates from right-to-left to left-to-right in adjacent energy-magnifying coils **24**. In the same manner as a right-handed screw advances from head to foot as it is turned clockwise, and a left-handed screw moves in the opposite direction, clockwise electron flow in a right-handed coil advances from head to foot, and clockwise electron flow in a left-handed coil advances from foot to head.

The single-layered internal output coils **28a** in **Fig.6**, present the same situation in which these coils are connected in series. Desirably, the inter-coil connectors **32a** cross over directly from one internal output coil **28a** to the adjacent internal output coil **28a**, head-to-head and the inter-coil connectors **32b** cross over directly foot-to-foot from one internal output coil **28a** to the adjacent internal output coil **28a**. This same handedness convention generally applies to all series-connected internal output coils **28a** connected in this manner. The head-to-head inter-coil connectors **32a** and foot-to-foot inter-coil connectors **32b** for the internal output coils **28a**, need not coincide with the same respective connectors **30a**, **30b** for the energy-magnifying coils **24**.

In another embodiment (not shown), each internal output coil is two-layered, with both leads at either the head or foot. Such a configuration allows for short and direct connections between adjacent internal output coils. Multiple-layered internal output coils may be more efficient, but the extra layers of coiled wire increase the mass of the apparatus, which may be a concern in mobile applications. Multiple wire layers carrying high current may also result in overheating, which may require that some space be left between each internal output coil **28a** and its surrounding energy-magnifying coil **24** to accommodate one or more conduits of a coolant through the

apparatus (at a sacrifice of some efficiency). The coolant can be, for example, forced air (in the case of photoconductors or doped semiconductors) or liquefied cryogenic gas (in the case of superconductors).

Fig.6 also shows two external conductors **34** connected to respective internal output coils **28a**. Electrons flow through the conductors **34** and the internal output coils **28a** in series. In addition, two external conductors **36** are connected to respective ends of the external output coil **28b**, and two external conductors **38** are connected to respective ends of the sending coil **20**.

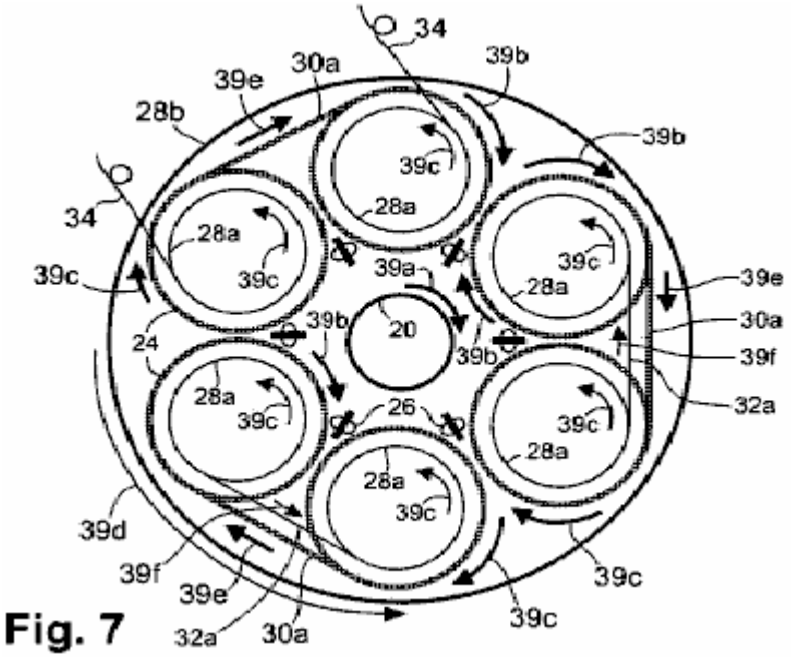


Fig. 7

Fig.7 is a schematic end view of the apparatus of **Fig.6**, showing the relative direction of electron flow in the various coils and in the inter-coil connections described for single-layer coils. At a particular oscillation phase, the clockwise electron flow denoted by the arrow **39a** in the sending coil **20** induces clockwise electron flow **39b** in all the energy-magnifying coils **24**. The magnified radiation from the clockwise electron flow in the energy-magnifying coils **24**, induces counterclockwise electron flow in all the internal output coils **28a**, as indicated by the arrows **39c**. The counterclockwise electron flow, denoted by the arrow **39d**, in the external output coil **28b** is opposite in direction to the electron flow in the energy-magnifying coils **24**.

The electron flow in the inter-coil connectors **30a** extending between adjacent energy-magnifying coils **24** is indicated by the arrows **39e**, and the electron flow in the inter-coil connectors **32a** extending between adjacent internal output coils **28a** is indicated by the arrows **39f**. During the next oscillation phase, all the direction arrows shown in **Fig.7** reverse themselves.

Connecting the internal output coils **28a** together in series is advantages if it is desired to maximise the output voltage from the apparatus **15**. Alternatively, the internal output coils **28a** can be connected together in parallel if it is desired to maximise the output electrical current from the apparatus **15** while minimising the output voltage. In this alternative configuration, all the internal output coils **28a** desirably are wound with the same handedness, with each coil **28a** having two respective leads. The leads at one end (e.g., the foot end) of the coils **28a** are connected to each other, and the leads at the other end (the head end) of the coils **28a** are connected to each other. The resulting parallel-coil system is connected in a conventional manner in other circuitry of the apparatus (not shown).

Further alternatively, the internal output coils **28a** can be connected together so as to provide more than one output circuit (so long as sufficient energy is produced for use as feedback to the sending coil **20** and for use in establishing favourable conditions for producing abundant low-mass electrons). Alternatively, the relative voltage(s) and current(s) of output power can be varied by changing the ratio of the number of turns in the energy-magnifying coils **24** to the number of turns in the internal output coils **28a**. Alternatively again, the energy-magnifying coils **24** can be employed in a separate manner to provide more than one energy-magnifying unit. Each unit can comprise one or more energy-magnifying coils that can serve its respective circuit of internal output coils.

The two conductors **36** connected to the external output coil **28b**, can be connected to the internal output coils **28a** or can be used (without being connected to the internal output coils **28a**) with only the external output coil **28b** to provide an independent output circuit (not shown). The two conductors **38** connected to the sending coil **20**,

are connected in the feed-back loop **46** such that electron flow in the sending coil **20** is in the same circular direction as in the internal output coils **28a**.

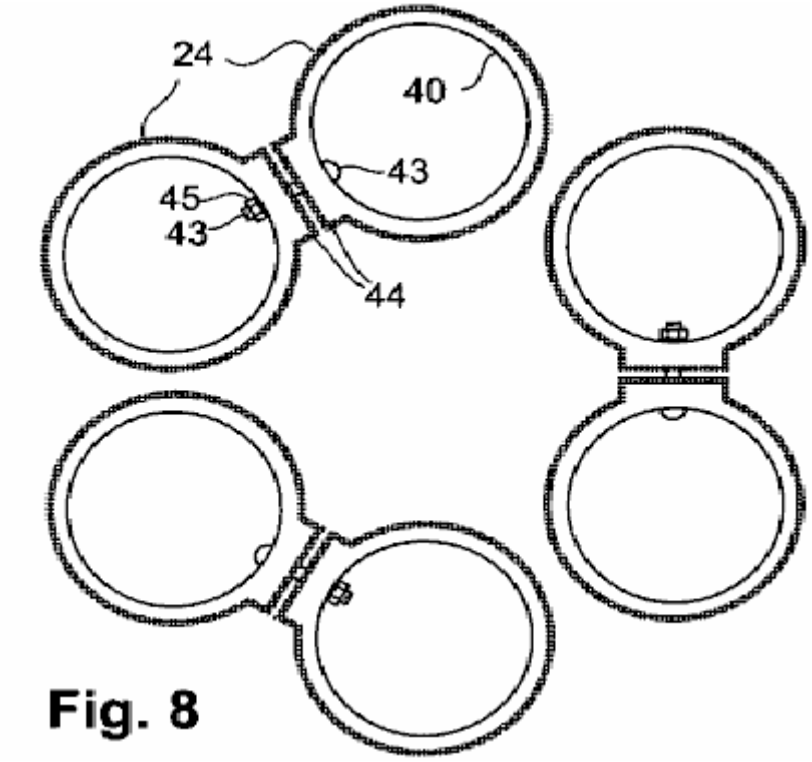


Fig. 8

Fig.8 depicts yet another embodiment of the apparatus **15**, in which each energy magnifying coil **24** comprises a thin film or thick film of a polycrystalline or other suitable photoconductor deposited in a helical manner directly on to a tubular substrate **40** desirably made of ceramic or other suitable dielectric material. on each energy magnifying coil **24**, the polycrystalline photoconductor is formed as a helical band on the outside of the tubular substrate **40**. The helical band of photoconductor can include a thin film of metal embedded within it. In certain cases, inter-coil connections between adjacent energy magnifying coils **24** can be made by extending the deposited photoconductor from the helices to contact areas **44** situated at the ends of the tubular substrates **40** and extending toward contact areas **44** on adjacent tubular substrates **40**. Electrical contact between adjacent energy magnifying coils **24** is made under moderate pressure via the contact areas **44**, which are shown in **Fig.8**. To distinguish the individual contact areas **44**, they are shown in a separated position before being pressed together to make contact. To maintain the integrity of the contact areas **44**, the energy magnifying coils **24** can be held together in mutual proximity by any of various non-metallic fasteners to make continuous electrical contact between all of the photoconductive portions. For example, bolts **43** and nuts **45** made of a plastic such as nylon, or other dielectric material, can be used. Another variation is to maintain contact pressure of one coil to the next by means of spring clips. Thus, in one embodiment, the energy magnifying coils **24** are connected so as to be in endless contact with each other, with no capacitative break between them. The remainder of the apparatus can be constructed in the same manner as the photoconductor or doped-semiconductor embodiment described above, wherein the same attention to the direction of electron flow in respective coils is observed.

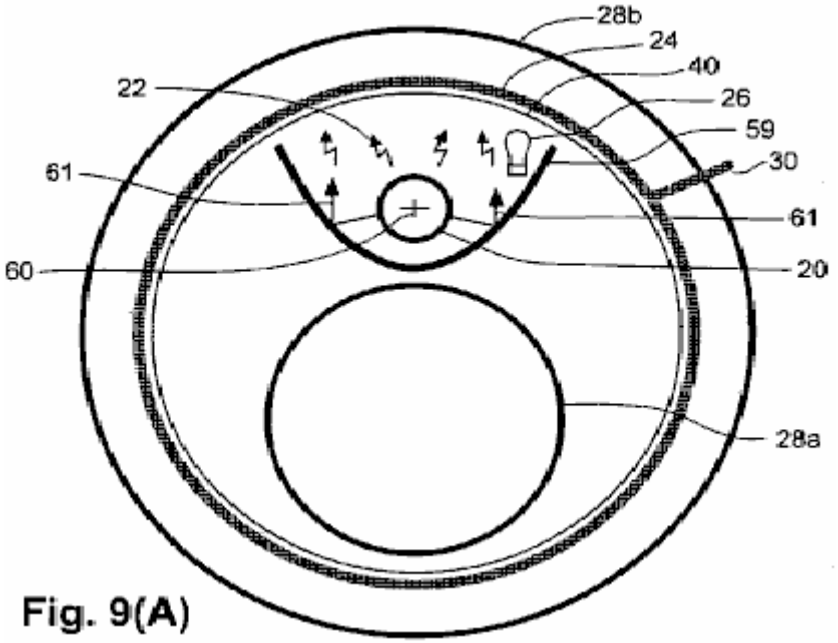


Fig. 9(A)

The coil configuration of yet another embodiment is shown in **Fig.9A** and **Fig.9B**. A tubular substrate **40**, supports a helical, thin film or thick film, dipole-type of energy-magnifying coil **24** that is nested inside and coaxial with a single external output coil **28b**. Nested inside the tubular substrate **40**, and with respective axes parallel to the axis of the tubular substrate **40**, are a sending coil **20** and an internal output coil **28a**. The sending coil **20** and the internal output coil **28a**, are positioned on opposite sides of a reflective metallic separator **59**. The separator **59** is substantially parabolic in cross-section throughout its axial extent, and is positioned so that the longitudinal edges are touching, or nearly touching, the tubular substrate **40**. The separator **59** can be composed of common, non-magnetic metal such as aluminium or magnesium. The sending coil **20** is positioned on the concave side of the separator **59**, with the axis of the sending coil **20** being positioned at the geometric focus **60** of the parabola and disposed parallel to the axis of the energy-magnifying coil **24**. The energy-magnifying coil **24** in this embodiment, comprises a thin film or thick film photoconductor formed helically on the tubular substrate **40**. A photoconduction exciter **26** is positioned inside the separator **59**. (The tubular substrate **40**, is made of a rigid material that is transparent to radiation produced by the photoconduction exciter **26**). All the other forms of the energy-magnifying coil **24** as described herein, including the superconducting form, can be employed in this embodiment.

The separator **59**, serves a double purpose. One purpose is to direct towards the energy-magnifying coil **24** the portion of the inductive-photon radiation **22** which is not otherwise directed towards the separator, as shown by the reflected-photon rays **61** in **Fig.6A**. (Reflection of these radiated photons does not change the directionality of the transverse force which these photons convey). Another purpose of the separator **59** is to serve as a shield to restrict the amount of inward radiation **18b** from the energy-magnifying coil **24** which is returned as a back-force to the sending coil **20**. The restricted back-force radiation is shown by the shaded area **63** in **Fig.9B**.

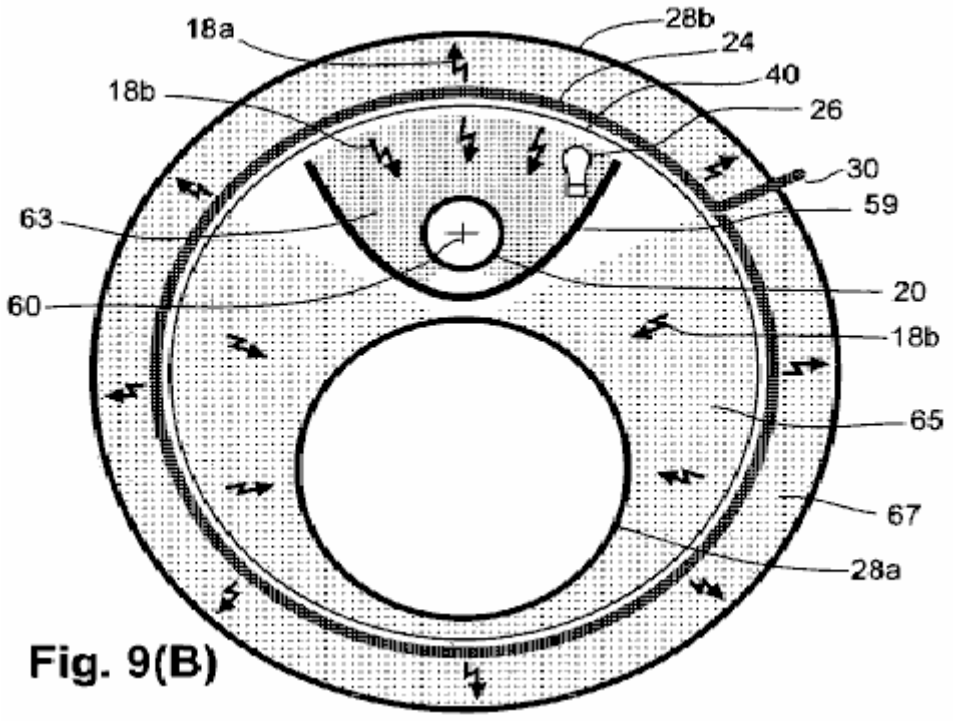


Fig. 9(B)

The portion of the inwardly directed, magnified inductive-photon radiation **18b** which is received by the internal output coil **28a**, is denoted by the shaded area **65**. The proportional amount of outwardly directed magnified radiation **18a** from the energy-magnifying coil **24** which is received by the external output coil **28b**, is shown by the shaded area **67**. The sum of the magnified radiation in the area **65** which reaches the external output coil **28b**, substantially exceeds the magnified radiation in the area **63** (the latter serving as a back-force on the sending coil **20**). This excess of utilised energy over the back-force energy, provides energy leverage. This embodiment also includes a starting mechanism, and initial power source for the photoconduction exciter, a work loop, and a feedback loop (not shown) as provided in the other embodiments described herein.

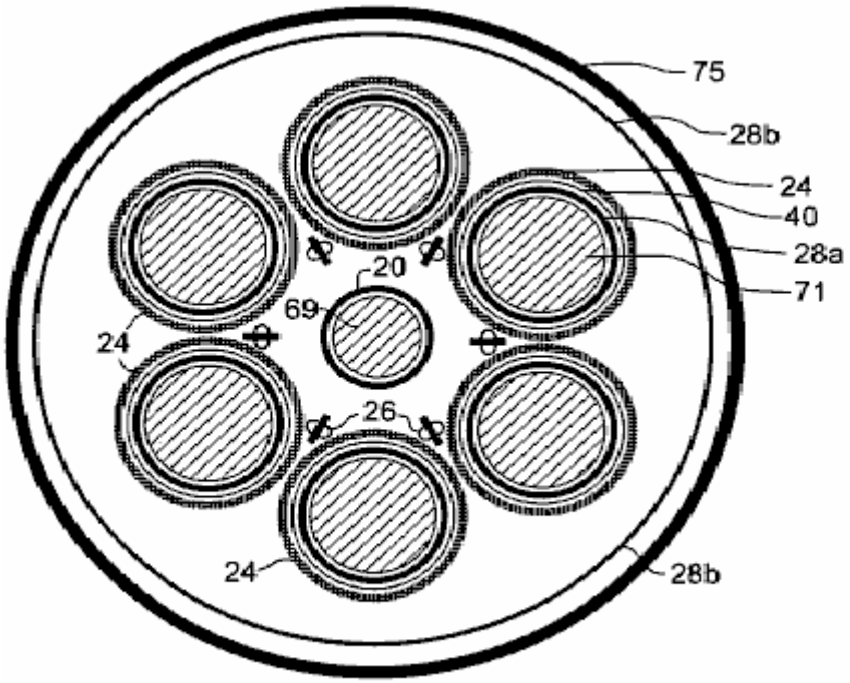


Fig. 10(A)

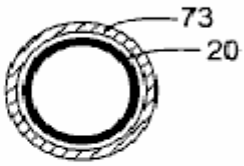


Fig. 10(B)

Certain features can be incorporated with any of the embodiments described herein, to add functional practicality. For example, referring to the schematic representation of a coil configuration shown in end view in **Fig.10A**, a ferromagnetic core **69** can be placed inside the sending coil **20**, and ferromagnetic cores **71** can be placed inside respective internal output coils **28a**. These cores increase the inductance of the apparatus, which lowers the frequency of the electrical oscillations produced by the apparatus. Although increases in inductance can cause

the output voltage and current to be out of phase, the phase difference can be corrected by adding capacitance to the circuitry by conventional means. Also shown, is an external metal shield which completely surrounds the apparatus to block any radiation from the device that could interfere with radios, televisions, telephones, computers and other electronic devices. the shield can be comprised of any of various non-magnetic metals such as aluminium or magnesium.

An alternative means of increasing the inductance of the apparatus is shown in **Fig.10B**, which is a variation of the end view of just the sending coil **20** that is depicted in **Fig.10A**. In **Fig.10B**, a ferromagnetic sleeve **73** is placed coaxially around the sending coil **20**.

The respective dimensional ratios of various components generally remain similar with respect to each other for different apparatus sizes, except for the longitudinal dimension, which generally can be as short or as long as desired, up to some practical limit. The respective gauges of wires used in the sending coil **20** and the output coils **28a** and **28b**, are commensurate with the electric current carried by these wires, and the respective thickness of insulation (if used) on the wires is commensurate with the voltage.

The outside diameter of the internal output coils **28a** desirably is only slightly less than the inside diameter of the respective energy-magnifying coils **24**, as shown in **Fig.6**, **Fig.7** and **Fig.8**, thereby ensuring close proximity of each internal output coil **28a** with its respective energy-magnifying coil **24**. At a sacrifice in efficiency, the outside diameter of the internal output coils **28a** can be made smaller, to allow space for heat from the current-carrying wires to escape or be removed by a coolant such as forced air in the case of a photoconductor type or doped semiconductor type apparatus, or by a cryogenic liquefied gas in the case of a superconductor type apparatus.

Also, desirably, the external output coil **28b** is connected in series with the internal output coils **28a** to maximise the output voltage from the apparatus **15** and to minimise heat produced by electric currents in the apparatus. The output voltage can be stepped down and the output electric current stepped up to normal operating ranges by using a transformer, wherein the primary of the transformer would comprise the load in the work loop **48**.

As discussed above, each energy-magnifying coil **24** can comprise a photoconductor or doped semiconductor formed as a helical pattern on a respective thin-walled tubular substrate provided with extended, raised contact surfaces at each end. The energy-magnifying coils **24** desirably are connected electrically (rather than capacitatively) to each other in series at the raised contact surfaces. The photoconductive coils desirably are coated using clear varnish or enamel to provide electrical insulation and to protect the photoconductors from oxidation and weathering.

Where the low-mass photoconducting electrons in the energy-magnifying coils **24** are present in a concentration which is insufficient for capturing most of the inductive-photon radiation from the sending coil **20**, each energy-magnifying coil desirably includes a thin metallic band. The metal desirably is in intimate contact with the low-mass-electron carrier. the metal can be on the exterior of a doped semiconductor, or it can be embedded in a photoconductor band of the coil to capture the inductive radiation and set up an electric field which, in turn, assists in accelerating the low-mass electrons. In the photoconductive embodiment, the photoconductive material desirably is disposed all around the metallic band so that the low-mass electrons are conducted on the outer side as well as the inner side and edges of the photoconductive band on the portion or portions which are exposed to illumination on the outside. The width of the metal band desirably is sufficient to capture as much of the inductive-photon radiation from the sending coil as is practical, since gaps between turns of the metal band in the energy-magnifying coil permit the sending coil's inductive radiation to pass through to the internal output coil. Since the sending coil's radiation is a half-cycle out of phase with the inductive radiation from the low-mass electrons, all the sending coil radiation which reaches the output coil, reduces the output efficiency of the apparatus.

Appropriate photoconductive materials (e.g. cadmium sulphide, cadmium selenide) for forming the energy-magnifying coils **24** are commercially available. The photoconductive material can be a single material or a mixture of materials, and can be formed by, for example, sputtering. A mixture of cadmium sulphide and cadmium selenide can be adjusted optimally to yield energy-magnifying coils exhibiting maximal energy-magnifying factors at a peak wavelength matching the brightest photoconduction exciters **26** which are available.

With respect to the photoconduction exciters **26**, photo-excitation of the energy-magnifying coils **24** can be provided by one or more LEDs, either surface-emitting or edge-emitting, for example, selected to produce an output wavelength matched to the peak photoconduction wavelength of the energy-magnifying coils **24**. In the embodiment of **Fig.7** and **Fig.10A**, individual LEDs **26** are positioned in linear arrays mounted back-to-back on respective mounting bars. The assembled mounting bars with LEDs are placed in the gaps between adjacent energy-magnifying coils **24** to illuminate at least the sides of the respective energy-magnifying coils **24** which receive inductive-photon radiation from the sending coil **20**. LEDs are advantageous compared to incandescent lamps because LEDs produce more light with less heat and have a much longer operational lifetime than

incandescent lamps. LEDs are also preferred because of their small size which facilitates fitting a large number of them into the relatively small space between adjacent energy-magnifying coils **24**.

Whereas the invention has been described in connection with several representative embodiments, the invention is not limited to those embodiments. On the contrary, the invention is intended to encompass all modifications, alternatives and equivalents as may be included within the spirit and scope of the invention, as defined by the appended claims.

References Cited in the Text

Helmholtz, H., 1847, *Uber die Erhaltung der Kraft*, G.A.Reimer, Berlin. *On the Conservation of Force*, English translation with Appendices of 1881, Wesleyan university Press, Middletown Conn., 1971.

Leimer, E., 1915, *"Uber Radiumantennen", Elektrotechnische Zeitschrift*, Heft 8, Feb. 25, 1915. English translation: "On Radium Antannae" *The Electrician*, Apr. 21 1916, and "Radium and Aerials" *Scientific American,* Supplement No. 2127, Oct. 7, 1916.

Anon, 1919a, "Youthful Seattle Inventor and his Invention" (photos and captions), *Seattle Post-Intelligencer*, Dec. 17, 1919.

Anon., 1919b, "Youth's Revolutionary invention is Backed by Professor - Hubbard's New Energy No Fake, Says Seattle College Man", *Seattle Post-Intelligencer*, Dec. 17, 1919,

Anon. 1920a, "Hubbard Coil Runs Boat on Portage Bay Ten Knots an Hour, Auto Test Next", Seattle Post-Intelligencer, Jul. 29, 1920.

Anon., 1920b, "Drives Boat with New Electric Generator" (photos and captions), *Seattle Post-Intelligencer*, Jul. 29, 1920.

Anon., 1928a, "Noted Flyers Try Out New Motor at Selfridge Field" (photo), *Detroit Free Press*, Fen. 25, 1928.

Anon., 1928b, "Fuelless Motor Shown; Gets Current From Air", *Detroit Free Press*, Feb. 25, 1928.

White, H. C., 1928c, "Lindbergh Tries Motor Earth Runs", *Detroit Free Press*, Feb. 25, 1928.

Anon., 1928d, "Fuelless Motor Impresses Experts", *New York Times*, Feb. 26, 1928.

Anon., 1928e, "No-Fuel Motor Finds Backers", *Detroit News*, Feb. 26, 1928.

Anon., 1928f, "Fuelless Motor Is A Generator", *New York Times*, Feb. 27, 1928.

Anon., 1928g, "Fuelless Motor Ends 2-year Task", *Seattle Post-Intelligencer*, Feb. 27, 1928.

Bermann, R. B., 1928 h, "Hubbard Believes Mystery Motor Based Upon His Own Invention", *Seattle Post-Intelligencer*, Feb. 27, 1928.

Ampere, A. M., 1820, "Memoire", *Annales de Chimie et de Physique*.

Ampere, A. M., 1825, "Memoire", *Annales de Chimie et de Physique*.

Weber, W., 1846, *Elektrodynamische Maassbestimmungen*, Weidmann'sche Buchhandlung, Leipzig.

Gauss, C. F., 1835, unpublished notes originally published posthumously in 1865 as "Grundgesetz Fur Alle Wechselwirkungen Galvanisvcher Strome" in Werke, Vol. 5, pp. 616-620; the correspondence from Gauss to Wilhelm Weber (1845) is on pp. 627-629.

Maxwell, J. C., 1873, *A Treatise on Electricity and Magnetism*, Clarendon Press, Oxford.

Larmor, J., 1897, "On the Theory of Magnetic Influence of Spectra; and on the Radiation From Moving Ions", *Phil. Mag* LXIII:503-512.

Larmor, J., 1900, *Aether and Matter*, Cambridge University Press.

Rutherford, E., 1911, *Phil. Mag.* XXI:669.

Planck, M., 1911, *Verh. D. Deutsch. Phys. Ges.* XIII:138

Bohr, N., Kramers, H. A., and Slater, J. C., 1924, *Phil. Mag.* XLVII:785

Whittaker, E., 1951, *A History of the Theories of Aether and Electricity*, Thos. Nelson & Sons, London; reprinted by Dover, New York, 1989.

Shpenkov, G. P., and Kreidik, L. G., "Microwave Background Radiation of Hydrogen Atoms", *Revista Ceencias Exastas e Naturais* 4, n.1, pp. 9-18, January/June 2002.

Fowler, W. A., 1983, "Experimental and Theoretical Nuclear Astrophysics; the Quest for the Origin of the Elements", Nobel Lecture, Dec. 8, 1983.

Tifft, W. G., 1976, "Discrete States of Redshift and Galaxy Dynamics I", *Astrophysical Journal* 206:38-56.

Tifft, W. G., 1977, "Discrete States of Redshift and Galaxy Dynamics I I: Systems of Galaxies", *Astrophysical Journal* 211:31-46; and "Discrete States of Redshift and Galaxy Dynamics I I I: Abnormal Galaxies", *Astrophysical Journal* 211:377-391.

Compton, A. H. and Allison, S. K., 1935, "X-Rays in Theory and Experiment" 2nd Ed., Van Nostrand, New York.

Pops, H., 1997, "The Metallurgy of Copper Wire", *Innovations*, Dec. 1997 (Copper Development Assoc.).

Lapp, R. E. and Andrews, H. L., 1954, *Nuclear Radiation Physics*, Prentice-Hall, Englewood Cliffs, N.J.

Bardeen, J., 1942, "Theory of Superconductivity" (abstr.), *Phys. Rev.* 59:298.

Dehmelt, H. G., 1989, "Experiments with an Isolated Subatomic Particle at Rest", Nobel Lecture, Dec. 8, 1989.

File, J. and Mills, R. G., 1963, "Observation of Persistent Current in a Superconducting Solenoid", *Phys. Rev. Lett.* 10(3), Feb. 1, 1963.

JOHN REARDON: AC GENERATOR

US Patent 6,946,767 **20th September 2005** **Inventor: John Reardon**

ALTERNATING CURRENT GENERATOR

This is a reworded excerpt form this patent which shows a high-efficiency electrical generator of alternating current. It is stated that this generator design is not affected by Lenz's law and the experimental results showed a 13,713% improvement over conventional power output.

ABSTRACT

An alternating current electrical generator creates three different and distinct magnetic fields between wound coil elements and rotating magnets, two fields of which are induced fields caused by magnet rotation. A plurality of magnets are positioned such that they extend outwardly from a rotating shaft. The magnets are circumferentially spaced around the shaft such that the north polar end of one magnet follows the south polar end of the next magnet or such that the polar end of one magnet follows a magnet with the same polar end. A plurality of stationary coil elements are positioned in spaced relation to the magnets. The coil elements each have electrical windings and metal cores which extend the lengths of the coil elements. The magnets rotate in spaced relation to the ends of the coil elements in such a way that the magnets' flux lines cut the cores located at the centre of each of the coil elements. This induces alternating electric current that oscillates back and forth along the lengths of the cores. This oscillating current creates an expanding and collapsing set of magnetic flux lines which expand and contract through every inch of the coil element's windings. This expanding and collapsing magnetic field induces an expanding and collapsing magnetic field and an alternating electric field in the coil elements.

US Patent References:

4009406	Feb, 1977	Inariba.
4823038	Apr, 1989	Mizutani et al.
5696419	Dec, 1997	Rakestraw et al.
5821710	Oct, 1998	Masuzawa et al.
5973436	Oct, 1999	Mitcham.
6069431	May, 2000	Satoh et al.
6373161	Apr, 2002	Khalaf.
6462451	Oct, 2002	Kimura et al.
6541877	Apr, 2003	Kim et al.
6717313	Apr, 2004	Bae.

BACKGROUND OF THE INVENTION

Alternating current generators are rotating devices which convert mechanical energy into electrical energy. To generate an electromotive force by mechanical motion, there must be movement between an electric coil and a magnetic field in a manner that will cause a change in the flux that passes through the coil. Fundamentally, the induced electromotive force is brought about by a change in the flux passing through the coil.

The use of electromagnets, magnets and magnet components in generators to create the magnetic field and its subsequent effect on electric coils to ultimately generate electric current is well known. Such magnetic generators operate by using the repelling forces created by the effect of changing polarities of both permanent and electromagnets. For instance, there are electrical generating devices which employ electromagnets which are fixed in position and which induce current by being selectively energised, as iron or other magnetic metal discs, bars, or similar elements are rotated at or around the magnets. Other systems employ electromagnet or permanent magnets which are rotated, by various means, in relation to iron cores or coils, inducing an alternating electrical current within the coils.

However, prior alternating current generators which employ rotating magnet systems are inefficient and generally fail to deliver adequate current, in relation to the mechanical effort applied.

SUMMARY OF THE INVENTION

It is thus an object of the present invention to address the limitations and disadvantages of prior alternating electric current generators.

It is an object of the present invention to provide an alternating current generator which generates a substantial amount of electrical current efficiently and effectively.

It is a further object of the present invention to provide an alternating current generator which employs rotating magnets to induce increased alternating electrical current within the iron cores of electrical coils.

It is still another object of the present invention to provide an alternating current generator which can be simply and readily manufactured and be operated with high efficiency.

These and other objects are obtained by the present invention, an alternating current electrical generator which creates three different and distinct magnetic fields between wound coil elements and rotating magnets, two fields of which are induced fields caused by magnet rotation. A plurality of magnets are positioned such that they extend outwardly from a rotating shaft. The magnets are circumferentially spaced around the shaft such that the north polar end of one magnet follows the south polar end of the next magnet or such that the polar end of one magnet follows a magnet with the same polar end. A plurality of stationary coil elements are positioned in spaced relation to the magnets. The coil elements each have electrical windings and metal cores which extend the lengths of the coil elements. The magnets rotate in spaced relation to the ends of the coil elements in such a way that the magnets' flux lines cut the cores located at the centre of each of the coil elements. This induces alternating electric current that oscillates back and forth along the lengths of the cores. This oscillating current creates an expanding and collapsing set of magnetic flux lines which expand and contract through every inch of the coil element's windings. This expanding and collapsing magnetic field induces an expanding and collapsing magnetic field and an alternating electric field in the coil elements.

The novel features which are considered as characteristic of the invention are set forth in particular in the appended claims. The invention itself, however, both as to its design, construction, and use, together with additional features and advantages thereof, are best understood upon review of the following detailed description with reference to the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

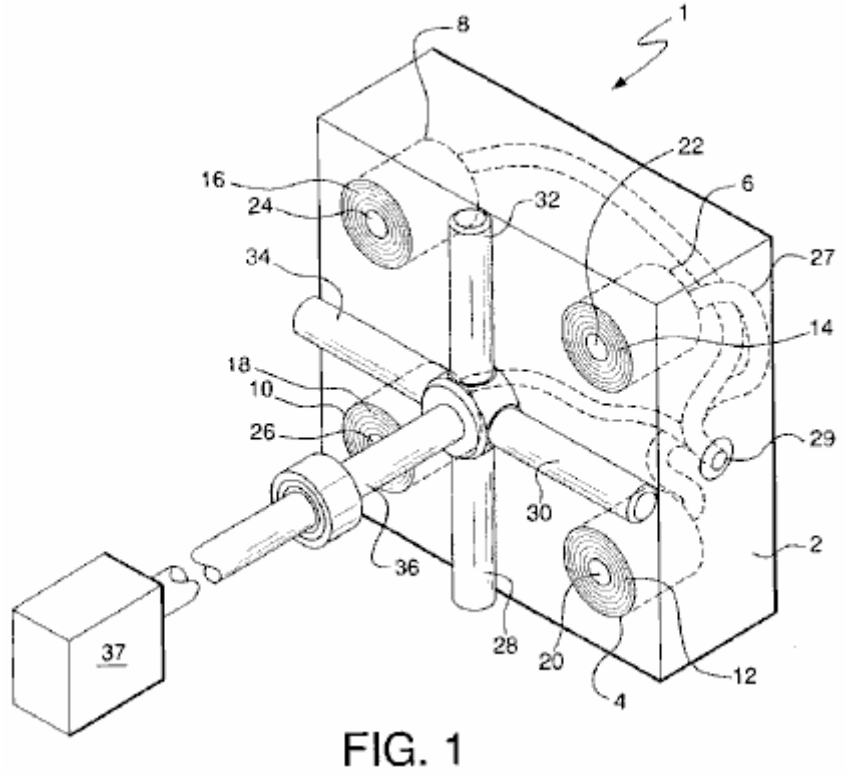


FIG. 1

Fig.1 is an isometric representation of keys components of the present invention.

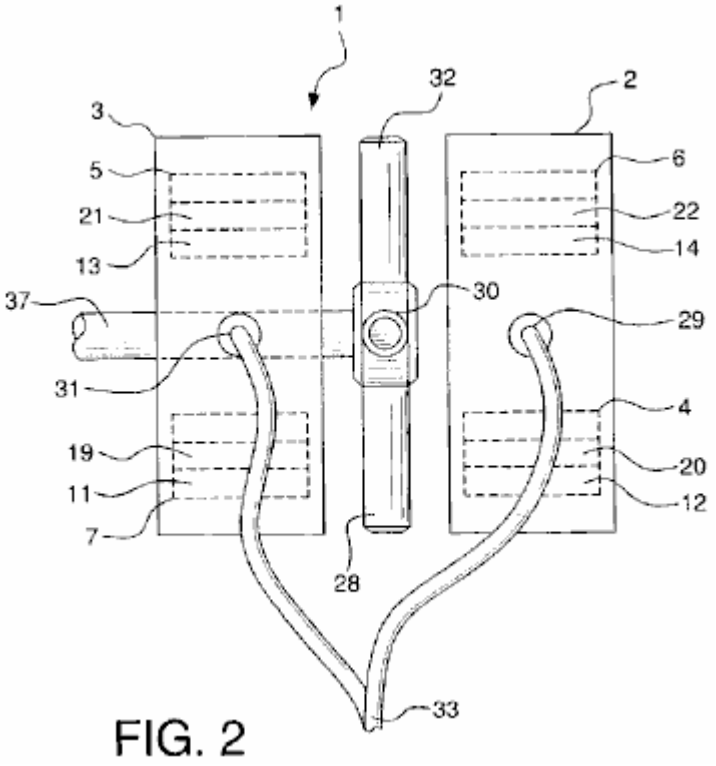


FIG. 2

Fig.2 is a side view representation of the present invention showing the two housed sets of coil elements and their relationship with the magnets.

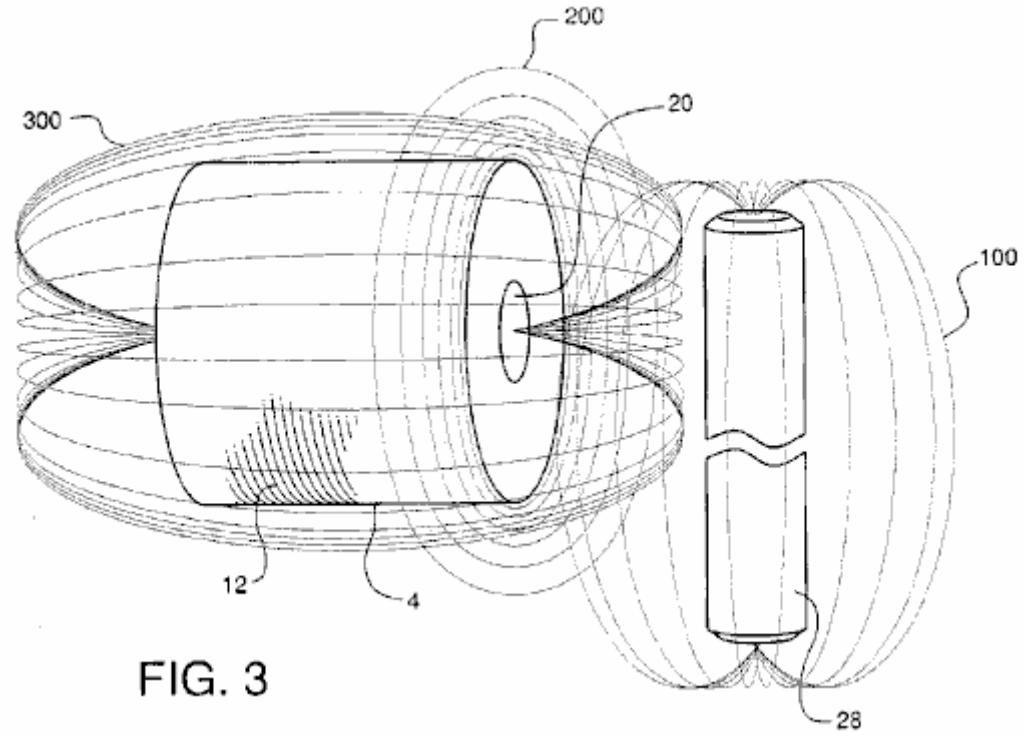


FIG. 3

Fig.3 is an explanatory view, showing the generation of flux lines which forms the basis for the operation of the present invention.

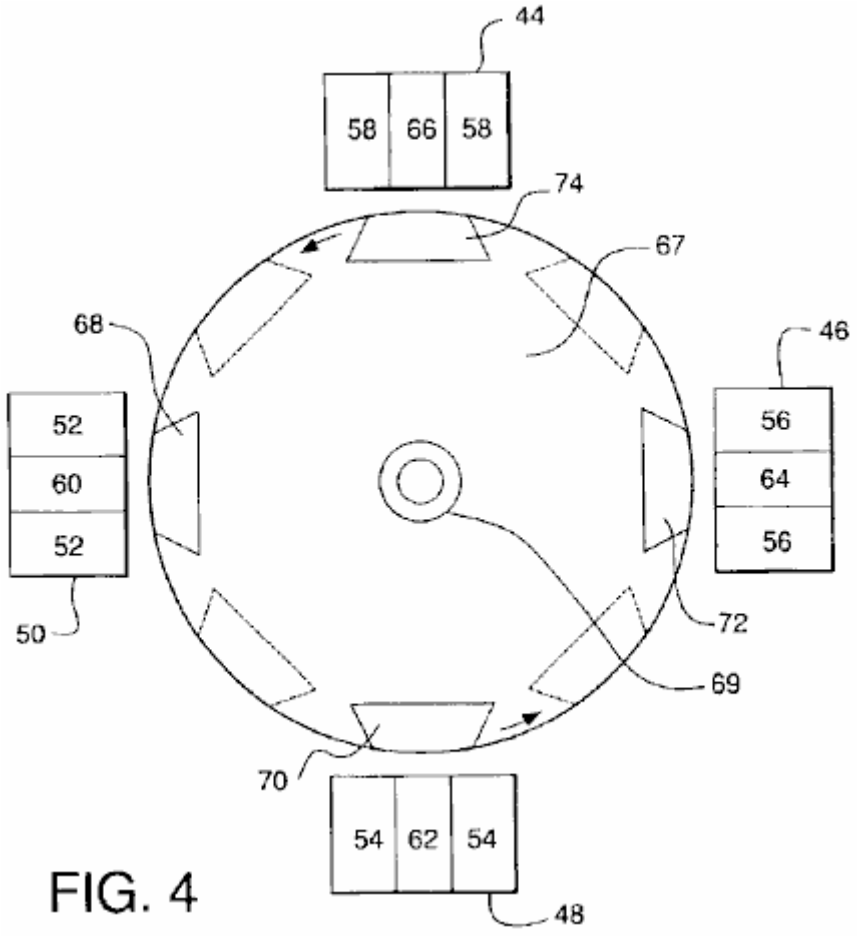


FIG. 4

Fig.4 is an alternate embodiment of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

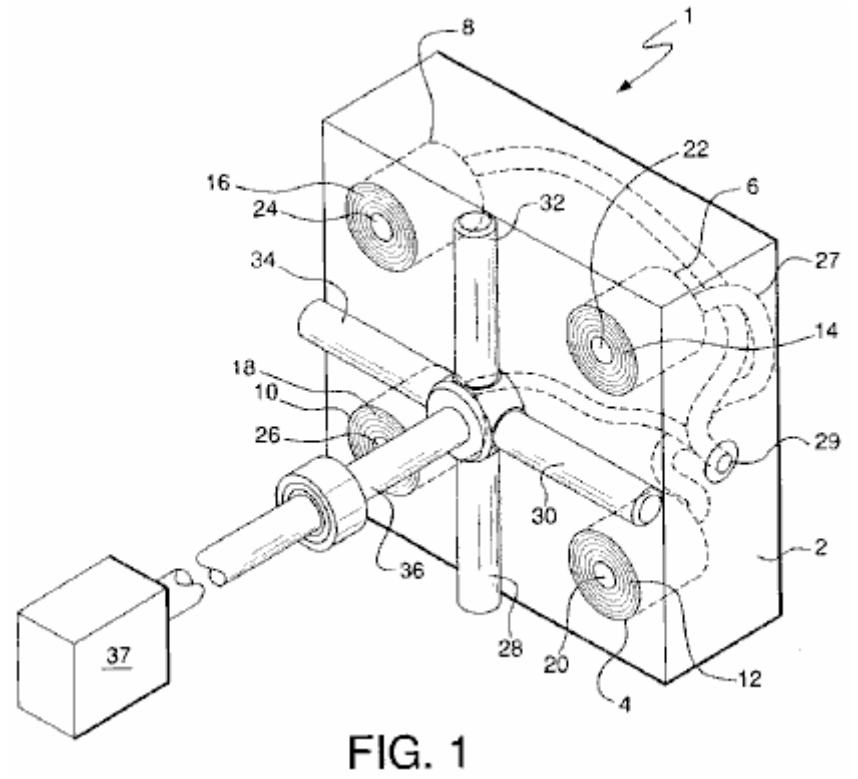


FIG. 1

Fig.1 and **Fig.2** show a clear depiction of the components of alternating current generator 1 of the subject invention. Generator 1 comprises housings **2** and **3**. For simplicity purposes and ease of understanding, only housing 2 is shown in **Fig.1**. It must be understood, however, that generator 1 of the present invention is configured for use with both housings **2** and **3**. Housing **2** contains coil elements **4**, **6**, **8** and **10**. Each coil element comprises multiple windings **12**, **14**, **16**, and **18**, respectively, wound around inner steel or similar metal cores **20**, **22**, **24**, and **26**, respectively. Each steel core extends the full length and directly through each of the coil elements. Coil elements **4**, **6**, **8**, and **10** are mounted within housing **2**, such that the end surfaces of the coil elements and the ends of cores **20**, **22**, **24**, and **26** are positioned flush with the external surface of housing **2**.

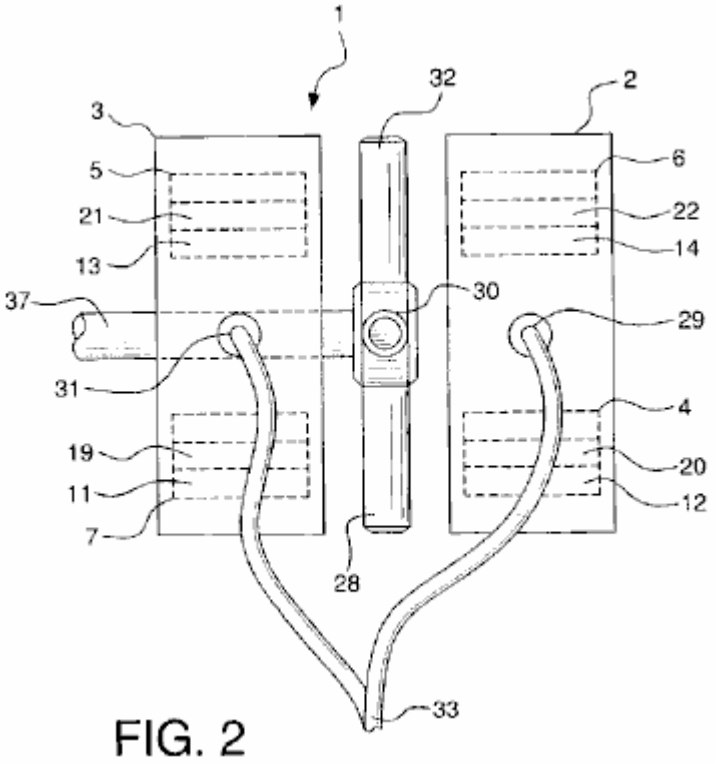


FIG. 2

Housing **3** also contains four coil elements positioned identically as has been described with regard to housing **2**. Two of these coil elements **5** and **7** are shown in **Fig.2**. Coil element **5** has multiple windings **13** and centre core **21** and coil element **7** has multiple windings **11** and centre core **21**.

Magnets **28**, **30**, **32**, and **34** are secured to shaft **36**, which is configured to be rotated by conventional power source **37**, such as a diesel engine, turbine, etc. Magnets **28**, **30**, **31**, and **32** all have ends with outwardly extending polarities. Magnets **28**, **30**, **32**, and **34** are positioned in spaced relation to the ends of exposed cores **20**, **22**, **24** and **26** of coil elements **4**, **6**, **8**, and **10** and in spaced relation to the ends of the four exposed cores in the four coil elements located in housing **3**, cores **19** and **21** being shown in **Fig.2**. All magnets are equidistantly spaced on and around shaft **36**, such that the outwardly extending pole of one magnet circumferentially follows the outwardly extending pole of the next magnet. The north polar end of one magnet may follow the south polar end of the next magnet or the polar end of one magnet may follow a magnet with the same polar end.

While four magnets and four cores are shown, it is contemplated that additional magnets and cores could be employed in the generator. Also, while permanent magnets are shown in the drawings, electromagnets could also be used, as they produce the same magnetic flux.

Alternating electrical current is generated when power source **37** rotates shaft **36**, thus causing rotation of magnets **28**, **30**, **32**, and **34** in spaced, adjacent relation to the ends of cores **20**, **22**, 24, and **26** of coil elements **4**, **6**, **8**, and **10**, and in spaced, adjacent relation to the ends of cores **19** and **21** of coil elements **7** and **5** and the ends of the cores of the other two similarly aligned coil elements in housing **3**. The current which is generated is transmitted through electrical conductive wiring **27**, which merges at connection points **29** in housing **2** and **31** in housing **3**, for the consolidated transmission at connection point **33** of the electricity produced.

As best represented in **Fig.2**, when magnet **28** is rotated in space relation to the end of core **20** of coil element **4**, flux lines **100** of the magnet cut the core at the centre of the coil element. This induces an alternating electrical current that oscillates back and forth along the length of core **20**. This oscillating current creates an expanding and collapsing set of magnetic flux lines **200** which expand and contract through every inch of coil windings **12**. Expanding and collapsing field **200** induces an alternating electric field in coil element **4** which is accompanied by

an expanding and collapsing magnetic field **300**. It is noted, significantly, that none of the magnetic field lines **100**, **200** and **300**, act in a negative fashion or in an opposing action. This allows the subject invention to overcome the limitations of Lenz's law, which states that whenever there is a change in magnetic flux in a circuit, an induced electromotive force is set-up tending to produce a current in a direction which will oppose the flux change.

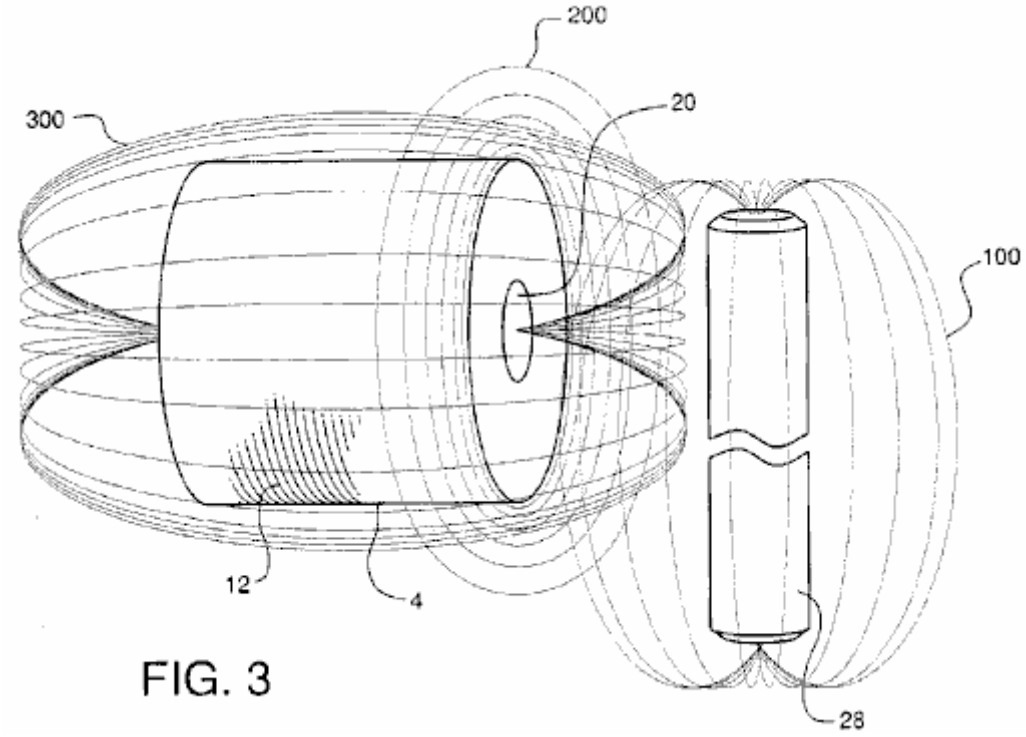


FIG. 3

Fig.3 illustrates an alternate embodiment of the invention to that which is shown in **Fig.1**. As shown in **Fig.3**, coil element **44** with outer windings **58** and inner steel core **66**, coil element **46** with windings **56** and core **64**, coil element **48** with windings **54** and core **62**, and coil element **50** with outer windings **52** and core **60** are positioned adjacent to rotor **67**, which is mounted on shaft **69**. Magnets **68** and **72** are mounted on rotor **67** such that the north poles of the magnets are positioned in spaced relation to coil elements **44**, **46**, **48** and **50**. Magnets **70** and **74** are mounted on rotor **67** such that the south poles of the magnets are also positioned in spaced relation to coil elements **44**, **46**, **48**, and **50**. All magnets are fixedly mounted on rotor **67** such that a north pole of one magnet circumferentially follows a south pole of the next magnet in line. The contemplated gap between the magnets and coil element cones is approximately 0.0001 of an inch, although the scope and use of the invention should not be deemed restricted to this distance.

As in the prior embodiment, rotation of magnets **68**, **70**, **72**, and **74**, by rotation of shaft **69** and hence rotor **67**, causes the flux lines of the magnets to cut cores **60**, **62**, **64**, and **66** of coil elements **44**, **46**, **48**, and **50**, eventually resulting in the output of electrical current as previously described.

It is noted that the larger the diameter of rotor **67**, the more coil elements can be positioned around the rotor. The greater the number of coil elements, the slower rotor **67** needs to rotate; however, there is a power loss in so doing. In addition, while rotor **67** is shown as being circular, it may be as square in shape or formed of as other appropriate multi-sided configurations.

This unique way of generating electricity allows generation of more electrical power, e.g. anywhere in the range of 4 to 137 times more power, than prior, conventional means. It also has the advantage of obtaining unity power with very little effort.

As evidence of such power gains, reference is made to the below outlined experimental outputs from coils and magnets which produced electric power the conventional way compared with the subject invention. The conventional way of generating power, for purpose of the following experimental outputs, as referenced herein, is accomplished by cutting the wires, not the cores, of the coil's windings with the magnet’s flux.

In this regard, proof is also provided that the herein described method of generating electrical power is not affected by Lenz's Law, by reference to the readings obtained by the conventional methods as the rpm and size of the coil increase. With conventional methods, the values do not change linearly, but are less because Lenz's Law restricts the outputs from increasing proportionally to the speed and size of the coil. In comparison, however,

in the method of producing power of the subject invention, there is an increase in the readings of V (voltage), I (current), and P (power) which are actually larger than anticipated.

It is also noted that, just like a transformer, when the number of turns ratio is increased, V increases and I decreases, which is exactly what is seen at the various rpm readings for the different size coils. However, they do not increase or decrease proportionally.

Thus, this presents the ideal model for producing electrical power that corresponds to the general law that states that as the speed increases, the voltage will increase proportionally, through the equation:
V = q (charge) × v (velocity) × B (magnetic field strength). This also holds true for a coil, in that transformers increase proportionally to the turns ratio.

With reference to the voltage outputs for each of the coils, 1100T, 2200T and 5500T, it is seen that they are consistent with the types of voltage outputs for a transformer action. That is to say, as the turns ratio goes up in a transformer so does the voltage. Since the increases in voltage between the number of turns is not exactly 2 to 5 times, one can pick any one of the coils and assume it is accurate and adjust the other coils accordingly. Thus, by fixing the 1100T coil, the other coils become 2837T and 5896T respectively. By fixing the 2200T coil, the other coils become 853T and 4572T respectively. And by fixing the 5500T coil, the other coils become 1026T and 2646T respectively. Also, if the adjustments are made as described here, i.e. that the coils are bigger than originally thought, and they are applied to the voltages for the conventional method of generating power, the voltages do not increase proportionally but are actually smaller than they are supposed to be, additional proof that Lenz's Law has application to conventional generators, but not to this invention.

The proportional changes in the voltage relative to speed can also be seen. Thus, considering the 350 RPM speed as accurate, the 1200 RPM and 1300 RPM speeds will adjust to 906 RPM and 1379 RPM respectively. Considering the 1200 RPM speed as accurate, the 350 RPM and 1300 RPM speed becomes 464 RPM and 1826 RPM respectively. And finally, considering the 1300 RPM speed as accurate, the 350 RPM and 1200 RPM speeds become 330 RPM and 854 RPM respectively.

It is noted that in using the various RPM readings based upon the above, it is seen that, in the conventional way of generating power, there are losses associated with the measured values. The calculated values again show the application of Lenz's Law in the conventional way of generating power, but not to this invention. In fact, whether or not there is an adjustment of RPM speed or coil size, the power generation of this invention is in no way affected by Lenz's Law.

Since Lenz's Law has no effect in this generator, it can be assumed that the voltages increase proportionally to the speed of the magnets rotation. Therefore, one can extrapolate the expected voltages at 1800 RPM, the speed necessary to create 60 Hz. With regard to this generator, for each of the three coils from the 350 RPM, 1200 RPM and 1300 RPM speeds, the following results (values are based on one coil/magnet.):

1. At assumed 350 RPM the voltages range as follows:

A. 5.863v @1100T
B. 15.12v @2200T
C. 31.42v @5500T

2. At assumed 1200 RPM the voltages range as follows:

A. 4.425v @1100T
B. 11.295v @2200T
C. 16.845v @5500T

3. At assumed 1300 RPM the voltages range as follows:

A. 6.217v @1100T
B. 10.716v @2200T
C. 17.668v @5500T

The reason the current is not changing linearly as the laws of physics imply from transformers, i.e. as voltage goes up based on the number of turns, the current goes down proportionally to the voltage gain, is due to the fact

that the inductive reactance is also going up. See the following chart for the inductive reactances for each coil at each speed.

Impedance (Z) or inductive reactance (X(L)) for a circuit with only a coil in it is the AC voltage divided by the AC current, and the inductance (L) is Z/2 × pi × F (frequency). For a circuit with a resistor and a coil
Z = square root of (R (resistance) squared + X(L) squared)).

The following is the chart of impedance Z for all coil sizes at all speeds for the conventional method of generating power and the method of generating power with this invention:

Where:
"T" stands for Turns,
"CM" stands for Conventional Method and
"SI" stands for Subject Invention:

(1) For 350 RPM for 1100T, 2200T and 5500T coils,

1. (a) CM: 0.57v / 56.6 mA = **10.021** ohms = Z
(b) SI: 1.14v / 106.6 mA = **10.694** ohms = Z

2. (a) CM: 0.93v / 32.4 mA = **28.704** ohms = Z
(b) SI: 2.94v / 70.1 mA = **41.94** ohms = Z

3. (a) CM: 2.09v / 17.3 mA = **120.81** ohms = Z
(b) SI: 6.11v / 37.9 mA = **161.21** ohms = Z

(2) For 1200 RPM for 1100T, 2200T and 5500T coils:

1. (a) CM: 1.45v / 60.2 mA = **23.387** ohms = Z
(b) SI: 2.95v / 141 mA = **20.922** ohms = Z

2. (a) CM: 3.225v / 36.2 mA = **89.088** ohms = Z
(b) SI: 7.53v / 73.5 mA = **102.449** ohms = Z

3. (a) CM: 4.81v / 17 mA = **282.941** ohms = Z
(b) SI: 11.23v / 31.4 mA = **357.643** ohms = Z

(3) For 1300 RPM for 1100T, 2200T and 5500T coils:

1. (a) CM: 1.6v / 83 mA = **19.27** ohms = Z
(b) SI: 4.59v / 157 mA = **29.236** ohms = Z

2. (a) CM: 2.75v / 50.4 mA = **54.455** ohms = Z
(b) SI: 7.74v / 88.5 mA = **87.458** ohms = Z

3. (a) CM: 5.061v / 17.3 mA = **292.543** ohms = Z
(b) SI: 12.76v / 36.4 mA = **350.549** ohms = Z

(4) For 400 RPM for 2300T coil with 24 gauge wire and 0.5" core:

(a) CM: 0.15v / 3.7 mA = **40.541** ohms = Z
(b) SI: 2.45v / 26.2 mA = **93.511** ohms = Z

(5) For 1200 RPM for 2300T coil with 24 gauge wire and 0.5" core:

(a) CM: 0.37v / 2.7 mA = **137.037** ohms = Z
(b) SI: 4.1v / 10.3 mA = **398.058** ohms = Z

(**6**) For 1400 RPM for 2300T coil with 24 gauge wire and 0.5" core:

(a) CM: 0.58v / 2.4 mA = **241.667** ohms = Z
(b) SI: 8.3v / 7.8 mA = **1065.385** ohms = Z

(**7**) For 400 RPM for 2300T coil with 24 gauge wire and 0.75" core:

(a) CM: 0.23v / 4.2 mA = **54.762** ohms = Z
(b) SI: 0.37v / 7.2 mA = **51.389** ohms = Z

(**8**) For 1200 RPM for 2300T coil with 24 gauge wire and 0.75" core:

(a) CM: 0.79v / 3.4 mA = **232.353** ohms = Z
(b) SI: 0.43v / 6.9 mA = **207.246** ohms = Z

(**9**) For 1400 RPM for 2300T coil with 24 gauge wire and 0.75" core:

(a) CM: 0.79v / 3.21 A = **246.875** ohms = Z
(b) SI: 2.1v / 2.7 mA = **777.778** ohms = Z

(**10**) For 400 RPM for 6000T coil with 28 gauge wire and 0.5" core:

(a) CM: 0.49v / 2 mA = **245** ohms = Z
(b) SI: 5.48v / 0.13 mA = **421.538** ohms = Z

(**11**) For 1200 RPM for 6000T coil with 28 gauge wire and 0.5" core:

(a) CM: 1.25v / 1.5 mA = **833.333** ohms = Z
(b) SI: 15.04v / 4.1 mA = **3668.293** ohms = Z

(**12**) For 1400 RPM for 6000T coil with 28 gauge wire and 0.5" core:

(a) CM: 2.08v / 1.1 mA = **1890.909** ohms = Z
(b) SI: 18.76v / 2.5 mA = **7504** ohms = Z

(**13**) For 400 RPM for 6000T coil with 28 gauge wire and 0.75" core:

(a) CM: 0.64v / 1.7 mA = **376.471** ohms = Z
(b) SI: 7.97v / 7.4 mA = **1077.027** ohms = Z

(**14**) For 1200 RPM for 6000T coil with 28 gauge wire and 0.75" core:

(a) CM: 2.08v / 1.3 mA = **1600** ohms = Z
(b) SI: 20.4v / 5.6 mA = **3642.857** ohms = Z

(**15**) For 1400 RPM for 6000T coil with 28 gauge wire and 0.75" core:

(a) CM: 2.28v / 1.2 mA = **1900** ohms = Z
(b) SI: 28.4v / 2.1 mA = **13523.81** ohms = Z

It is noted that, based upon the variations of wire size, core size and number of turns, the following effects take place:
(a) the smaller the wire size the higher the gains regardless of speed;
(b) the greater the number of turns, generally the higher the gains; and
(c) the smaller the core size the higher the gains.
However, when comparing coils with smaller cores but a higher number of turns, the effects stay about the same.

Finally, the magnets are placed in the rotor so that they are all north or south poles up or out. A pure half-wave generator is created without rectifying the AC signal, which otherwise must be accomplished in a normal AC generator with electronic components in an electronic circuit.

Experimental Values for Producing Power the Conventional Way and with the Subject Invention:

The results were achieved using a small 3" magnet with a diameter of ±2" on a 1.25" high coil of 1" diameter and 3/8" centre/core of steel. (Unknown wire gauge size.)

(a) Conventional method of generating electricity:

1. 0.324 volts
2. 2.782 mA (milli-amps)
3. 0.9014 mW (milli-watts)

(b) Subject invention method of generating electricity:

1. 7.12 volts
2. 17.35 mA
3. 100.87 mW

(c) Associated gains of Volts, Current and Watts:

1. 2,198% over conventional voltage output.
2. 624% over conventional current output.
3. 13,713% over conventional power output.

The following results show the voltage, current and power outputs for an 1100, 2200 and 5500 turn coil of 20 gauge copper wire, 6" in length, 3" in diameter with a 0.75" core of steel. The results are those taken at 350 rpm, 1200 rpm and 1300 rpm.

(A) 350 RPM for an 1100 turn coil

	Volts	**mA**	**mW**
(a) Conventional method:	0.57	56.6	32.3
(b) Subject invention method:	1.14	106.6	121.5
(c) Associated gains	200%	188.3%	376.6%

(B) 350 RPM for a 2200 turn coil

	Volts	**mA**	**mW**
(a) Conventional method:	0.93	32.4	30.1
(b) Subject invention method:	2.94	70.1	206.1
(c) Associated gains	316.1%	216.4%	684%

(C) 350 RPM for a 5500 turn coil

	Volts	**mA**	**mW**
(a) Conventional method:	2.09	17.3	36.2
(b) Subject invention method:	6.11	37.9	231.6
(c) Associated gains	292.3%	219.1%	640%

(D) 1200 RPM for an 1100 turn coil

	Volts	**mA**	**mW**
(a) Conventional method:	1.45	60.2	87.3
(b) Subject invention method:	2.95	141	416
(c) Associated gains	203.4%	234.2%	476%

(E) 1200 RPM for a 2200 turn coil

	Volts	mA	mW
(a) Conventional method:	3.225	36.2	116.75
(b) Subject invention method:	7.53	73.5	553.5
(c) Associated gains	233.5%	203%	474%

(F) 1200 RPM on a 5500 turn coil

	Volts	mA	mW
(a) Conventional method:	4.81	17	81.77
(b) Subject invention method:	11.23	31.4	352.6
(c) Associated gains	235.5%	184.7%	431.3%

(G) 1300 RPM on an 1100 turn coil

	Volts	mA	mW
(a) Conventional method:	1.6	83	132.8
(b) Subject invention method:	4.59	157	704.9
(c) Associated gains	280.6%	189.2%	530.8%

(H) 1300 RPM on a 2200 turn coil

	Volts	mA	mW
(a) Conventional method:	2.75	50.5	138.9
(b) Subject invention method:	7.74	88.5	685
(c) Associated gains	281.5%	175.2%	493.3%

(I) 1300 RPM on a 5500 turn coil

	Volts	mA	mW
(a) Conventional method:	5.061	17.3	87.56
(b) Subject invention method:	12.76	36.4	464.5
(c) Associated gains	252%	210%	530%

The following readings are taken from a coil with 24 gauge wire, 0.5" centre/core of steel and 2300T.

(A) 400 rpm

	Volts	mA	mW
(a) Conventional method:	0.15	3.7	0.56
(b) Subject invention method:	2.45	26.2	64.2
(c) Associated gains	1,633%	708%	11,563%

(B) 1200 rpm

	Volts	mA	mW
(a) Conventional method:	0.37	2.7	1
(b) Subject invention method:	4.1	10.3	42.2
(c) Associated gains	1,108%	381%	4,227%

(C) 1400 rpm

	Volts	mA	mW
(a) Conventional method:	0.58	2.4	1.39
(b) Subject invention method:	8.31	7.8	64.82
(c) Associated gains	1,433%	325%	4,657%

The following readings are taken from a coil made with 24 gauge wire, 0.75" centre/core of copper, 2300T.

(A) 400 rpm

	Volts	mA	mW
(a) Conventional method:	0.23	4.2	0.97
(b) Subject invention method:	0.37	7.2	2.66
(c) Associated gains	137%	171%	235%

(B) 1200 rpm

	Volts	mA	mW
(a) Conventional method:	0.79	3.4	2.69
(b) Subject invention method:	1.43	6.9	9.87
(c) Associated gains	181%	203%	367%

(C) 1400 rpm

	Volts	mA	mW
(a) Conventional method:	0.79	3.2	2.53
(b) Subject invention method:	2.1	2.7	5.67
(c) Associated gains	266%	84%	224%

The following readings were taken from a coil made of 28 gauge wire, 0.5" centre/core of steel and 6000T.

(A) 400 rpm

	Volts	mA	mW
(a) Conventional method:	0.49	2	0.98
(b) Subject invention method:	5.48	13	71.24
(c) Associated gains	1,118%	65%	7,269%

(B) 1200 rpm

	Volts	mA	mW
(a) Conventional method:	1.25	1.5	1.88
(b) Subject invention method:	15.04	4.1	61.66
(c) Associated gains	1,203%	273%	3,289%

(C) 1400 rpm

	Volts	mA	mW
(a) Conventional method:	2.08	1.1	2.29
(b) Subject invention method:	18.76	2.5	46.9
(c) Associated gains	902%	227%	2,050%

The following readings were taken from a coil made of 28 gauge wire, 0.75" steel centre/core and 6000T.

(A) 400 rpm

	Volts	mA	mW
(a) Conventional method:	0.64	1.7	1.09
(b) Subject invention method:	7.97	7.4	58.98
(c) Associated gains	1,245%	435%	5,421%

(B) 1200 rpm

	Volts	mA	mW
(a) Conventional method:	2.08	1.3	2.7
(b) Subject invention method:	20.4	5.6	114.24
(c) Associated gains	981%	431%	4,225%

(C) 1400 rpm

	Volts	mA	mW
(a) Conventional method:	2.28	1.2	2.74
(b) Subject invention method:	28.4	2.1	88.04
(c) Associated gains	1,246%	175%	2,180%

The extrapolated voltages for the items immediately above at the 1800 RPM speed for the method of the subject invention are as follows:

(A) 400-1400 RPM, 0.5" core, 2300T:
(1) 11.025v
(2) 6.15v
(3) 10.68v

(B) 400-1400 RPM, 0.75" core, 2300T:
(1) 1.665v
(2) 2.145v
(3) 2.7v

(C) 400-1400 RPM, 0.5" core, 6000T:
(1) 24.66v
(2) 22.56v
(3)24.12

(D) 400-1400 RPM, 0.75" core, 6000T:
(1) 10.25v
(2) 30.6v
(3) 36.51v

Some of the readings above do not seem consistent with others. This is attributed to the possibility that the wire connections may have been faulty or the proximity of the magnet relative to the core or coil may not have been the same. This was not taken into account at the time the tests were done.

The following figures are derived based on the premise that the subject invention has characteristics of a transformer when the number of turns on the coils change. In nearly all these situations, the subject invention acts exactly like a transformer, while the conventional way of producing electricity does not.

CM = conventional method;
SI = subject invention;

350 RPM		1100 to 2200 Turns	1100 to 5500 Turns	2200 to 5500 Turns
CM:	expected voltage:	1.14 volts	2.85 volts	2.325 volts
	actual voltage:	0.93 volts	2.09 volts	2.09 volts
	expected current:	28.3 mA	11.32 mA	12.96 mA
	actual current:	32.4 mA	17.3 mA	17.3 mA
	expected power:	32.3 mW	32.3 mW	30.1 mW
	actual power:	30.1 mW	36.2 mW	36.2 mW
	expected voltage gain:	2	5	2.5
	actual voltage gain:	1.636	3.667	2.247
	expected current gain:	0.5	0.2	0.4
	actual current gain:	0.572	0.306	0.534
	expected power gain:	1	1	1
	actual power gain:	0.932	1.12	1.203
SI:	expected voltage:	2.28 volts	5.70 volts	7.35 volts
	actual voltage:	2.94 volts	6.11 volts	6.11 volts
	expected current:	53.30 mA	42.64 mA	28.04 mA
	actual current:	70.10 mA	37.90 mA	37.90 mA
	expected power:	121.74 mW	243.05 mW	206.09 mW
	actual power:	206.10 mW	231.60 mW	231.60 mW
	expected voltage gain:	2	5	2.5
	actual voltage gain:	2.579	5.36	2.078
	expected current gain:	0.5	0.2	0.4
	actual current gain:	0.658	0.356	0.5407
	expected power gain:	1	1	1
	actual power gain:	1.696	1.906	1.124

1200 RPM		1100 to 2200 Turns	1100 to 5500 Turns	2200 to 5500 Turns
CM:	expected voltage:	2.90 volts	7.25 volts	8.06 volts
	actual voltage:	3.225 volts	4.81 volts	4.81 volts
	expected current:	30.10 mA	12.04 mA	14.48 mA
	actual current:	36.2 mA	17.0 mA	17.0 mA
	expected power:	87.29 mW	87.29 mW	116.71 mW
	expected voltage gain:	2	5	2.5
	actual voltage gain:	2.22	3.32	1.49
	expected current gain:	0.5	0.2	0.4
	actual current gain:	0.6	0.28	0.47
	expected power gain:	1	1	1
	actual power gain:	1.34	0.94	0.70
SI:	expected voltage:	5.9 volts	14.75 volts	18.83 volts
	actual voltage:	7.53 volts	11.23 volts	11.23 volts
	expected current:	70.50 mA	28.20 mA	29.40 mA
	actual current:	73.50 mA	31.40 mA	31.40 mA
	expected power:	415.95 mW	415.95 mW	553.60 mW
	actual power:	553.50 mW	352.60 mW	352.60 mW
	expected voltage gain:	2	5	2.5
	actual voltage gain:	2.55	3.81	1.49
	expected current gain:	0.5	0.2	0.4
	actual current gain:	0.52	0.22	0.43
	expected power gain:	1	1	1
	actual power gain:	1.33	0.85	0.64

1300 RPM		1100 to 2200 Turns	1100 to 5500 Turns	2200 to 5500 Turns
CM:	expected voltage:	3.20 volts	8.00 volts	6.88 volts
	actual voltage:	2.75 volts	5.06 volts	5.06 volts
	expected current:	41.50 mA	16.60 mA	20.20 mA
	actual current:	50.50 mA	17.3 mA	17.3 mA
	expected power:	132.8 mW	132.8 mW	138.98 mW
	actual power:	138.9 mW	87.56 mW	87.56 mW
	expected voltage gain:	2	5	2.5
	actual voltage gain:	1.72	3.16	1.84
	expected current gain:	0.5	0.2	0.4
	actual current gain:	0.61	0.21	0.34
	expected power gain:	1	1	1
	actual power gain:	1.05	0.66	0.63
SI:	expected voltage:	9.18 volts	22.95 volts	19.35 volts
	actual voltage:	7.74 volts	12.76 volts	12.76 volts
	expected current:	78.50 mA	31.40 mA	35.40 mA
	actual current:	88.50 mA	36.40 mA	36.40 mA
	expected power:	720.63 mW	720.63 mW	685.0 mW
	actual power:	685.0 mW	464.50 mW	464.50 mW
	expected voltage gain:	2	5	2.5
	actual voltage gain:	1.69	2.78	1.65
	expected current gain:	0.5	0.2	0.4
	actual current gain:	0.56	0.23	0.41
	expected power gain:	1	1	1
	actual power gain:	0.95	0.64	0.68

The following data represents the expected and actual voltage readings for the conventional method of producing voltage and the method of the subject invention. In virtually all circumstances, the herein invention produced more voltage than the conventional method and has gains that are higher than anticipated.

1100 Turns		350 to 1200 RPM	350 to 1300 RPM	1200 to 1399 RPM
CM:	expected voltage:	1.954 volts	2.117 volts	1.571 volts
	actual voltage:	1.45 volts	1.60 volts	1.60 volts
	expected voltage gain:	3.429	3.714	1.083
	actual voltage gain:	2.544	2.807	1.103
SI:	expected voltage:	3.909 volts	4.234 volts	3.196 volts
	actual voltage:	2.95 volts	4.59 volts	4.59 volts
	expected voltage gain:	3.429	3.714	1.083
	actual voltage gain:	2.579	4.026	1.556

2200 Turns		350 to 1200 RPM	350 to 1300 RPM	1200 to 1399 RPM
CM:	expected voltage:	3.189 volts	3.454 volts	3.494 volts
	actual voltage:	3.225 volts	5.061 volts	5.061 volts
	expected voltage gain:	3.429	3.714	1.083
	actual voltage gain:	3.468	2.957	0.853
SI:	expected voltage:	10.081 volts	10.919 volts	8.157 volts
	actual voltage:	7.53 volts	7.74 volts	7.74 volts
	expected voltage gain:	3.429	3.714	1.083
	actual voltage gain:	2.561	2.633	1.028

5500 Turns		350 to 1200 RPM	350 to 1300 RPM	1200 to 1399 RPM
CM:	expected voltage:	7.167 volts	7.62 volts	5.211 volts
	actual voltage:	4.81 volts	5.061 volts	5.061 volts
	expected voltage gain:	3.429	3.714	1.083
	actual voltage gain:	2.301	2.422	1.052
SI:	expected voltage:	20.951 volts	22.693 volts	12.166 volts
	actual voltage:	11.23 volts	12.76 volts	12.76 volts
	expected voltage gain:	3.429	3.714	1.083
	actual voltage gain:	1.838	2.088	1.049

2300 Turns	(0.5” core 24 gauge wire)	400 to 1200 RPM	400 to 1400 RPM	1200 to 1400 RPM
CM:	expected voltage:	0.45 volts	0.525 volts	0.432 volts
	actual voltage:	0.37 volts	0.58 volts	0.58 volts
	expected voltage gain:	3.00	3.50	1.167
	actual voltage gain:	2.467	3.867	1.568
SI:	expected voltage:	7.35 volts	8.57 volts	4.785 volts
	actual voltage:	4.10 volts	8.31 volts	8.31 volts
	expected voltage gain:	3.00	3.50	1.167
	actual voltage gain:	1.673	3.392	2.027

2300 Turns	(0.75” core 24 gauge wire)	400 to 1200 RPM	400 to 1400 RPM	1200 to 1400 RPM
CM:	expected voltage:	0.69 volts	0.805 volts	0.922 volts
	actual voltage:	0.79 volts	0.79 volts	0.79 volts
	expected voltage gain:	3.00	3.50	1.167
	actual voltage gain:	3.435	3.435	1.00
SI:	expected voltage:	1.11 volts	1.295 volts	1.688 volts
	actual voltage:	1.43 volts	2.10 volts	2.10 volts
	expected voltage gain:	3.00	3.50	1.167
	actual voltage gain:	3.865	5.676	1.469

6000 Turns	(0.5” core 28 gauge wire)	400 to 1200 RPM	400 to 1400 RPM	1200 to 1400 RPM
CM:	expected voltage:	1.47 volts	1.715 volts	1.459 volts
	actual voltage:	1.25 volts	2.08 volts	2.08 volts
	expected voltage gain:	3.00	3.50	1.167
	actual voltage gain:	2.551	4.245	1.664
SI:	expected voltage:	16.44 volts	19.18 volts	17.668 volts
	actual voltage:	15.04 volts	18.76 volts	18.76 volts
	expected voltage gain:	3.00	3.50	1.167
	actual voltage gain:	2,745	3.423	11.247

6000 Turns	(0.75” core 28 gauge wire)	400 to 1200 RPM	400 to 1400 RPM	1200 to 1400 RPM
CM:	expected voltage:	1.92 volts	2.24 volts	2.427 volts
	actual voltage:	2.08 volts	2.28 volts	2.28 volts
	expected voltage gain:	3.00	3.50	1.167
	actual voltage gain:	3.25	3.563	2.427
SI:	expected voltage:	23.91 volts	27.895 volts	23.80 volts
	actual voltage:	20.40 volts	28.40 volts	28.40 volts
	expected voltage gain:	3.00	3.50	1.167
	actual voltage gain:	2.56	3.563	1.392

CLAIMS

1. A generator for providing alternating electrical current comprising:

(a) an independently supported rotating drive shaft;

(b) a plurality of spaced apart magnets extending outwardly from the shaft, the magnets each creating magnetic flux and having a polar end with a particular north or south polarity, said magnets being circumferentially spaced and mounted around the shaft, such that the polar ends of the magnets extend away from and circumferentially around the shaft;

(c) a plurality of stationary coil elements, each said coil element comprising electrical windings wound about substantially the entire coil element, each of said coil elements further comprising a solid metal core with two ends extending substantially through the coil element at the centre of the coil element, each element being positioned such that one end of each of the cores is located in spaced, adjacent relation to the magnets, whereby rotation of the shaft causes rotation of the magnets around the shaft and in spaced, adjacent relation to the cores of the coil elements, the magnetic flux of the magnetics cutting the cores of the coil elements, creating alternating current in the coil elements; and

(d) a first housing in which some of the plurality of coil elements are mounted and a second housing in which the remainder of the plurality of coil elements are mounted.

2. The generator as in claim 1 wherein the magnets are spaced 90° apart around the shaft.

3. The generator as in claim 1 wherein magnets with north polar ends alternate with the magnets with south polar ends in spaced, circumferential relation around the shaft.

4. The generator as in claim 1 wherein all the plurality of magnets are magnets with the same polar ends.

5. The generator as in claim 1 wherein the magnets are equidistantly spaced around the shaft.

6. The generator as in claim 1 wherein the plurality of magnets is rotated by the drive shaft between and in spaced apart relation with the housings.

7. The generator as in claim 1 further comprising four magnets extending from the shaft, adjacent magnets being positioned perpendicular to each other, each magnet having either an outwardly extending north or south polar end, and said magnets being positioned such that a north polar end magnet follows a south polar end magnet, in spaced, circumferential relation around the shaft.

8. The generator as in claim 1 further comprising multiple north polar end magnets and multiple south polar end magnets extending from the shaft, said magnets being positioned in spaced, circumferentially relation around the shaft.

9. The generator as in claim 1 in which the shaft is positioned within a rotor and the magnets are circumferentially mounted on the rotor.

10. The generator as in claim 1 in which the shaft is connected to power means for rotating the shaft, whereby upon rotation of the shaft, the magnets are rotated around the shaft in spaced relation to the cores of the coil elements, thereby inducing an alternating electrical field along the length of each of the cores, thereby producing an alternating electric current in the windings of the coil elements.

11. The generator as in claim 10 further comprising means to transmit the alternating electrical current for electrical power usage.

GEOFFREY SPENCE: ENERGY CONVERSION SYSTEM

US Patent 4,772,816 20th September 1988 Inventor: Geoffrey M. Spence

ENERGY CONVERSION SYSTEM

This is a slightly reworded excerpt from this patent which has a substantial electrical output capable of providing it's own electrical input to be self-powering as well as generating kilowatts of excess power. The highly-respected Dr. Harold Aspden comments:

"In my Energy Science Report No. 8I, I also mentioned the apparatus designed by Geoffrey Spence, an inventor based in U.K. This is the subject of his U.S. Patent No. 4,772,816. Electrons injected into a chamber formed between two concentric electrodes are deflected into the inner electrode by a pair of magnets that provide a magnetic field along the central axis of the concentric electrodes. Of itself, this should add no excess energy, because the energy fed into accelerating the electrons is merely absorbed by electrostatic repulsion in charging the central electrode and so the capacitor. However, if that electron flow pulsates and there are connections to draw electron current from that central electrode then the pulsation implies a recurring sequence of charge and discharge. That 'magic capacitor' function is then harnessed.

The questions then are whether the Spence invention really works and whether it is commercially viable? Well, I wrote that Energy Science Report back in 1996, six years ago, and it is only a few months ago that I heard any more of that project. Geoffrey Spence has developed the prototype product to the stage where he has closed the loop in the sense that a portion of the output power was fed back to impart the energy needed to sustain the electron beams. He has a self-sustaining unit that can deliver kilowatts of useful electrical power with no visible energy input."

ABSTRACT

The apparatus uses a magnetic field (**80**) to accelerate a charged particle radially towards a target electrode (**10**). The increased kinetic energy of the particles enables the particle to give up more electrical energy to the target electrode (**10**) than was initially given to it. This charges the target electrode (**10**), and the increased energy is extracted from the apparatus by connecting an electrical load between the target electrode and a point of lower or higher potential.

US Patent References:

1717413	Jun, 1929	Rudenberg	310/306.
3202844	Aug, 1965	Hatch	310/306.
3234411	Feb, 1966	Klein	310/306.
3312840	Apr, 1967	Gabor	310/306.
3393330	Jul, 1968	Vary	310/306.
3899696	Aug, 1975	Shimadu	310/306.

DESCRIPTION

This invention relates to a process and apparatus for generating a potential difference between two or more electrodes and using charged particles as energy carriers.

Electrical power is usually generated by burning a fossil fuel and converting the energy released into rotary motion which drives electrical generators. This is cost-effective only if carried out on a large scale, the conversion process being inefficient; utilising natural resources, and producing waste products which can cause serious environmental pollution. An additional disadvantage is that the electrical power cannot be supplied directly to road vehicles or ships.

The energy-conversion process of this invention involves no health or pollution hazard and generates electrical power directly by a single-stage process without waste products. The overall energy-conversion factor and power-to-weight ratio are both high, making the apparatus suitable for most fixed and mobile applications.

One known apparatus for doing useful work by operating on electrons with a magnetic field is called the "betatron". This includes a doughnut-shaped vacuum chamber between the poles of a specially-shaped electromagnet. Thermionically-produced electrons are injected into the chamber with an initial electrostatic energy of about 50 keV. As the magnetic field builds up during its positive-going half-cycle, it induces an electromotive force within the doughnut, which force accelerates the electrons and forces them to move in an curved path, by

interaction with the magnetic field. An important distinction between the betatron and the energy converter of this invention is that in the former the magnetic field has got to be able to increase over a very short period, in order to accelerate the electrons sufficiently, whereas in the latter the magnetic field is virtually constant and the electrons fall inwardly to give up both their kinetic energy and electric charge to a central electrode.

The present invention aims at providing an energy converter which may be mobile and which has a permanent magnet or an energised source of magnetic radiation associated with it in order to amplify the electrical energy initially imparted to charge particles fed to, or produced in, a so-called "vacuum" chamber forming part of the generator, which increased energy is extracted from the target electrode on which the particles are incident.

Accordingly the present invention provides an energy converter as per the appended claims.

While the invention is not to be limited to any particular theory of operation, it is based on the fact that, when a charged particle is constrained to move through a radial distance d (irrespective of the path which it actually follows) through a magnetic field of intensity **H**, the work done on the particle is **H** x **d**. For an electron carrying a charge **e**, and moving at a speed **v** over distance **d**, the total force on the electron is the centripetal force the sum of **H** x **e** x **v**, less the force exerted on the electron in the opposite direction by the centrifugal force, which is the sum of (**m** x $\mathbf{v}^2$)/**r**. By making the radius of the centre electrode appreciably greater than the orbit of equilibrium, the centrifugal force can be minimised, thus maximising the centripetal force, and hence the work done in bringing the charge to the electrode.

The process by which the converter of this invention works uses, as a source of charge, electrically-charged particles, for example electrons and/or ions. Two or more electrodes are housed in a low-pressure chamber. A magnetic field as specified below traverses the chamber: it emanates from a permanent magnet, electromagnet or a source of magnetic radiation. An external source of energy is used to give the charge particles initial kinetic energy, for example by heating, acceleration through an electric field, or from nuclear radiation. The energy-conversion process uses the magnetic field to transfer the charged particles along a desired orbit until they impinge on a central electrode (cathode). The work done on the particles (therefore the electrical potential attained by the cathode) is proportional to the resultant magnetic force times the distance over which the force acts. As the particles move within the chamber they cross the magnetic field. This produces a force acting on the particles, the force being proportional to the field strength, speed and electrical charge of the particles, and the sine of the angle of incidence between the path of the particle and the magnetic lines of force. This force has an angular component and a centripetal one, which forces the particles to travel along a spiral orbit.

An opposing centrifugal force also acts on the particles in opposition to the centripetal magnetic force. The electrode potential is proportional to the work required to be done on the charged particles to overcome both the centrifugal force and the electric field around the cathode as the charges accumulate and the potential difference between the electrodes increases. Maximum electrode potential is reached when the centrifugal and repulsive forces are equal to the centripetal force, after which no further charged particles reach the electrode. The radius of the electrode determines the minimal value voltage between the central and an outer electrode: as the central electrode radius is reduced (by sputtering or erosion) the centrifugal force increases, reducing the number of charged particles which can reach the central electrode and therefore the electrode potential, for a given field strength and particle speed. The difference in mass between ions and lighter charged particles, such as electrons, results in different centrifugal forces for given particle kinetic energies. The generator output and efficiency are optimised when the generator uses the maximum magnetic field to minimise the centrifugal force and to maximise the radial distance over which the force acts for a given field strength. Particles having the highest charge-to-mass ratio should be used.

Low pressure gases can be used as a charge source when ionised by particle collision and excitation within the chamber. Doped gases can minimise the energy level for ionising gas atoms/molecules thereby improving efficiency. However, the resultant magnetic force is lower for the heavier ions due to their lower velocity so that the electric field radiated by the high voltage electrode (cathode) can attract oppositely charged particles (+ ions) and subsequently discharge the electrode reducing the output voltage. Various methods can be used to overcome or reduce this effect. For example one method would be to separate the opposite charges and/or to use electrical biased grids to control the flow of opposite charges to the high voltage electrode.

Gaseous systems are generally more complex than single charge systems, providing higher currents at lower voltages, whereas single charge systems, for example electrons used in high vacuum chambers, can generate higher voltages.

The magnetic field can be from one or more permanent magnets and/or from one or more electromagnets; a static magnetic field produces a constant output voltage, while a varying field produces a varying voltage for particles with equal mass and velocity.

An external source is used to accelerate the charged particles to give them initial kinetic energy, which is released as heat when the particles collide with the electrode. When the energy represented by the increased voltage between the electrodes is greater than the energy required to provide the charged particles; and accelerate them, the conversion process is self-sustaining, the output energy being the difference between the sum of the kinetic energy lost and the energy generated. Charge flows from the central electrode via an external load to another electrode. The electrical energy (work) released is a function of the current (sum of charges that flow per second) times the potential difference. Electrical and thermal output can be controlled by varying: the field strength; the particle speed; the particle density (mean free path), and/or by incorporating a grid to control the rate at which particles reach the central electrode. The output is also proportional to the heat lost or gained, since the translational energy of the particle is proportional to its temperature. Heat liberated at the electrode can be returned to the particles to maintain their energy, or be utilised in a heat exchanger for external use. The generator normally uses non-reacting conductive material to prevent chemical reaction by gases, coolants etc. with the electrodes, container walls or other components. Various particle trajectories, directional movements and positioning of the orbiting particles can be used with appropriate magnetic fields. The low-pressure gas can be ionised by any suitable means: one method would be to use an electron/ion gun where the plane and direction of the injected particles is correct for the applied magnetic field. In gas apparatus, the electrons flowing through the external circuit, on reaching the anode, recombine with a gaseous ion to form a neutral gas atom/molecule. This atomic particle is duly re-ionised by collision and/or the electric fields, the energy being directly or indirectly derived from the work done by the resultant force acting on the charged particles.

In order that the invention may be better understood, it will now be described with reference to the accompanying schematic drawings, which are given by way of example, and in which:

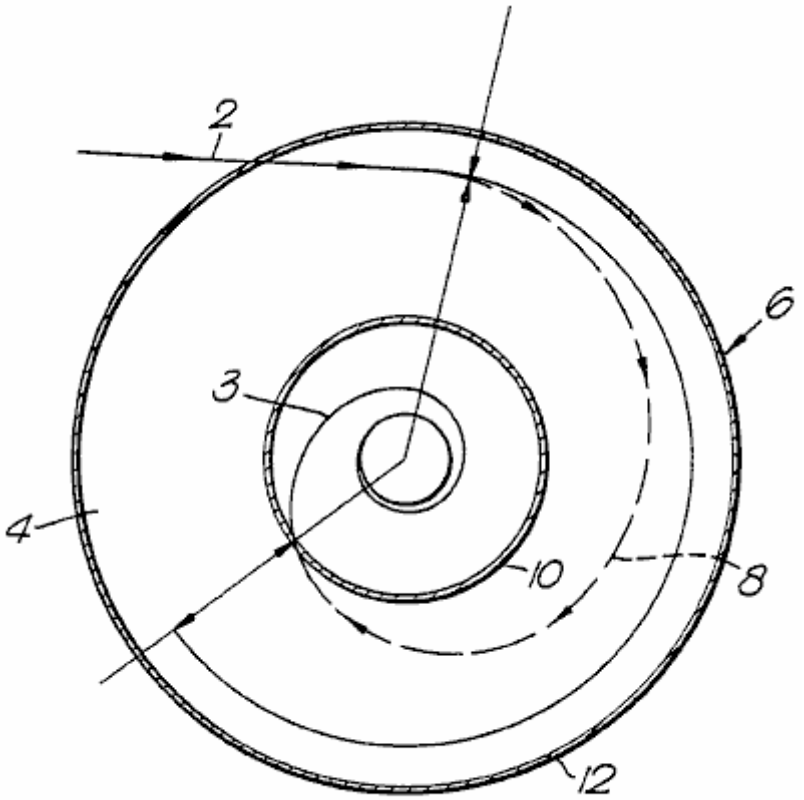


Fig.1.

Fig.1 shows schematically a cross-section of the generator; and the path followed by a particle during the energy-conversion process;

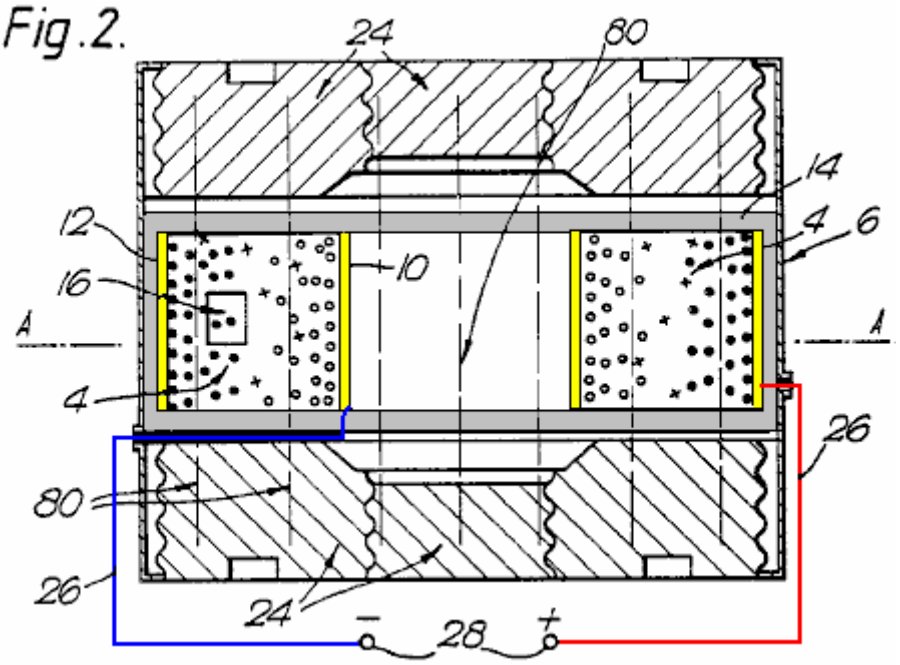


Fig.2 shows an axial cross-section of one type of apparatus for the invention, using permanent magnets; and a grid controlling ion migration to the cathode.

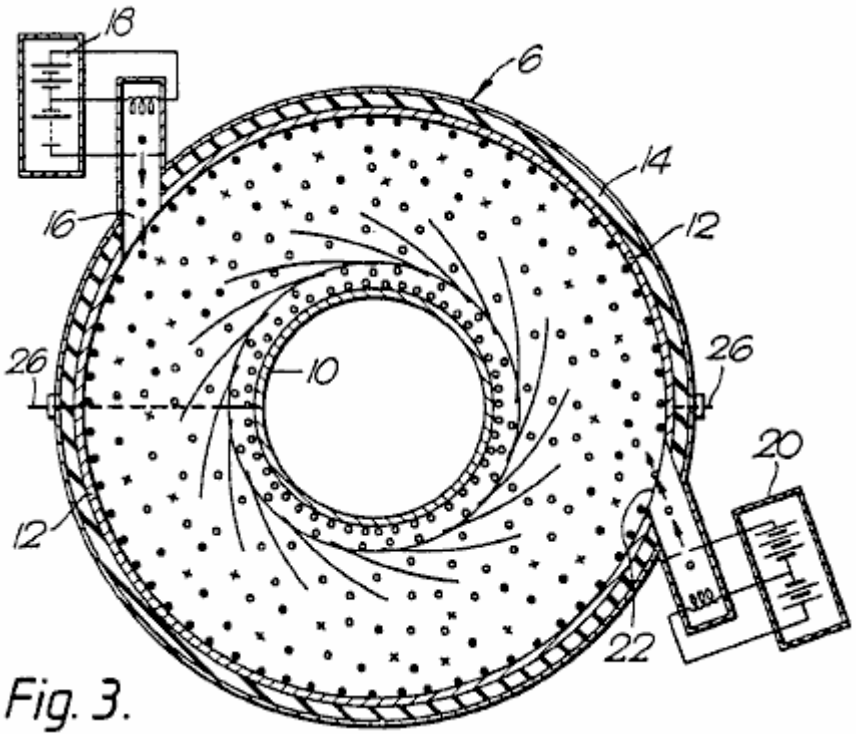


Fig.3 shows a cross-section of the apparatus of **Fig.2** along the line **A--A**;

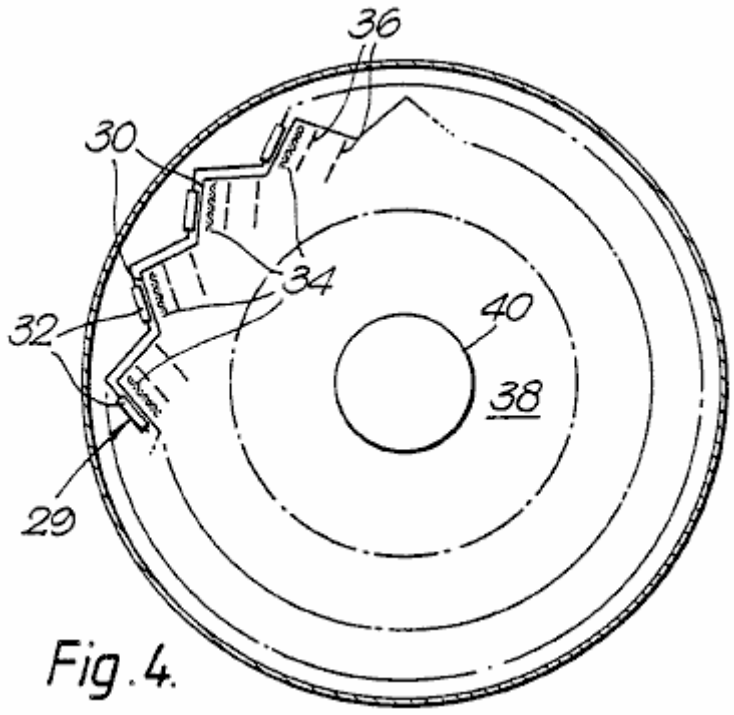


Fig.4 is a diagrammatic section through one form of converter using electrons, showing a circular series of electron sources;

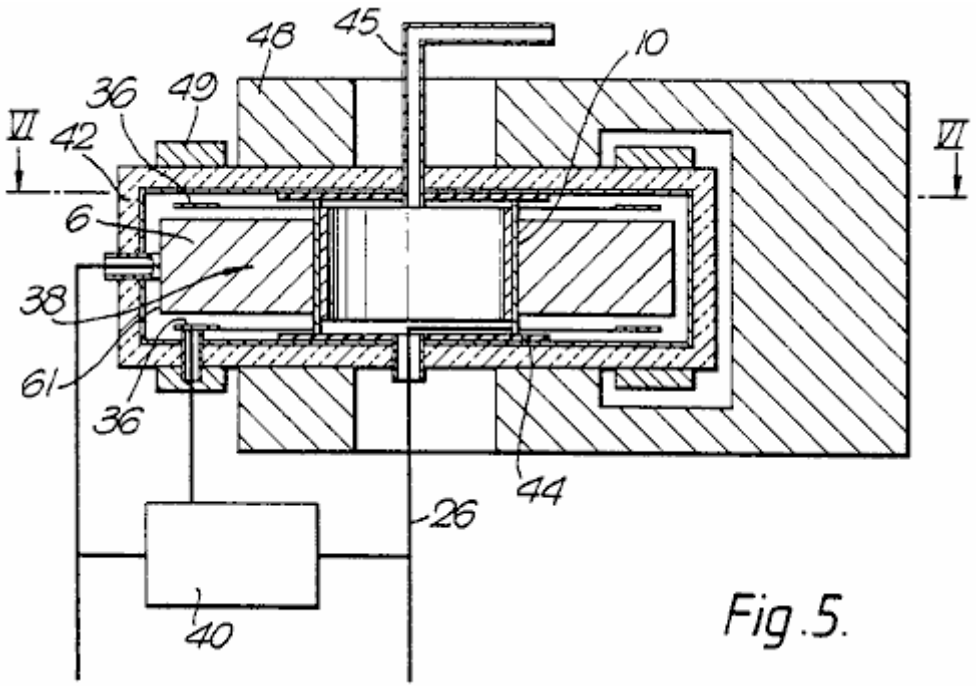


Fig.5 is an axial cross-section through a more practical embodiment of the **Fig.4** converter;

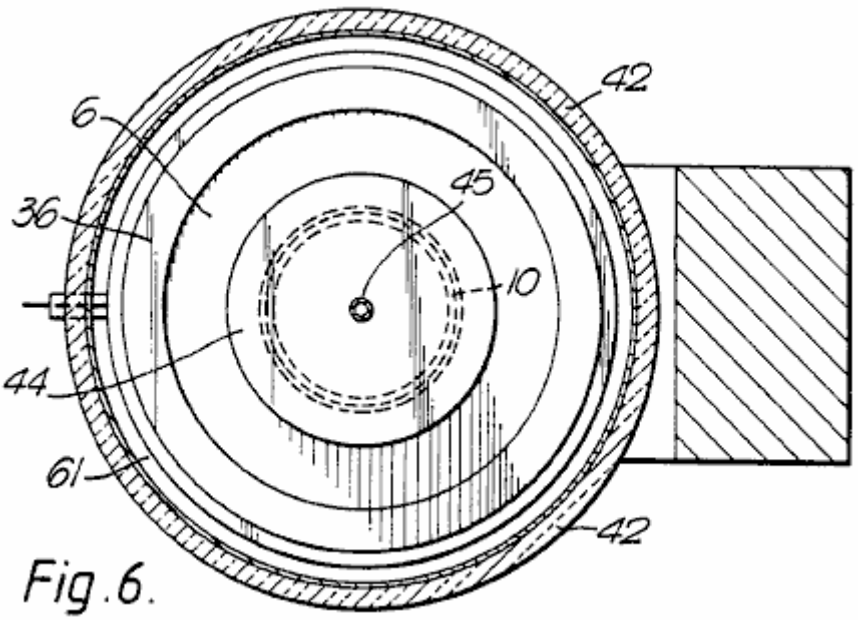


Fig. 6.

Fig.6 is a section along the line **VI--VI** of **Fig.5**;

Fig. 7.

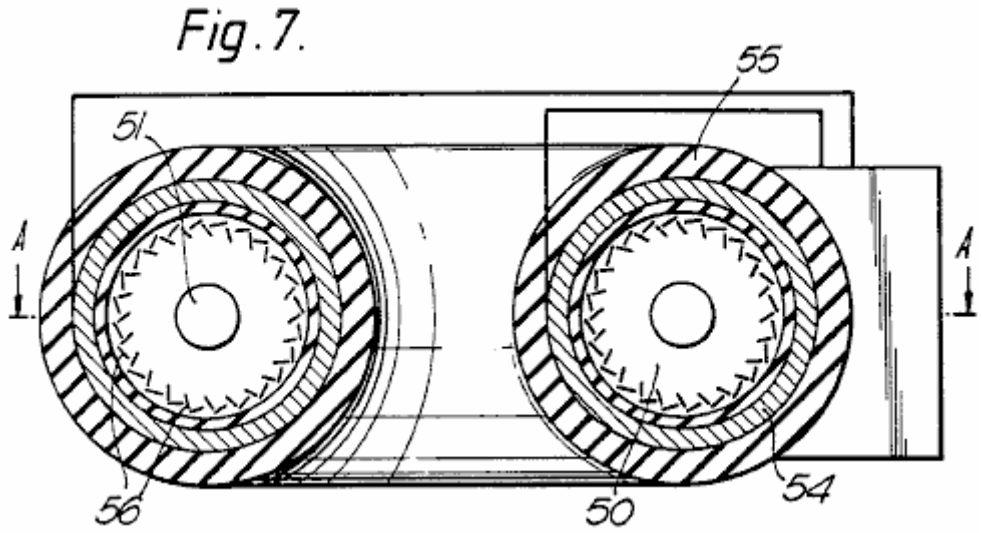


Fig.7 is a cross-section along a diameter of a doughnut-shaped (toroidal) high-power converter;

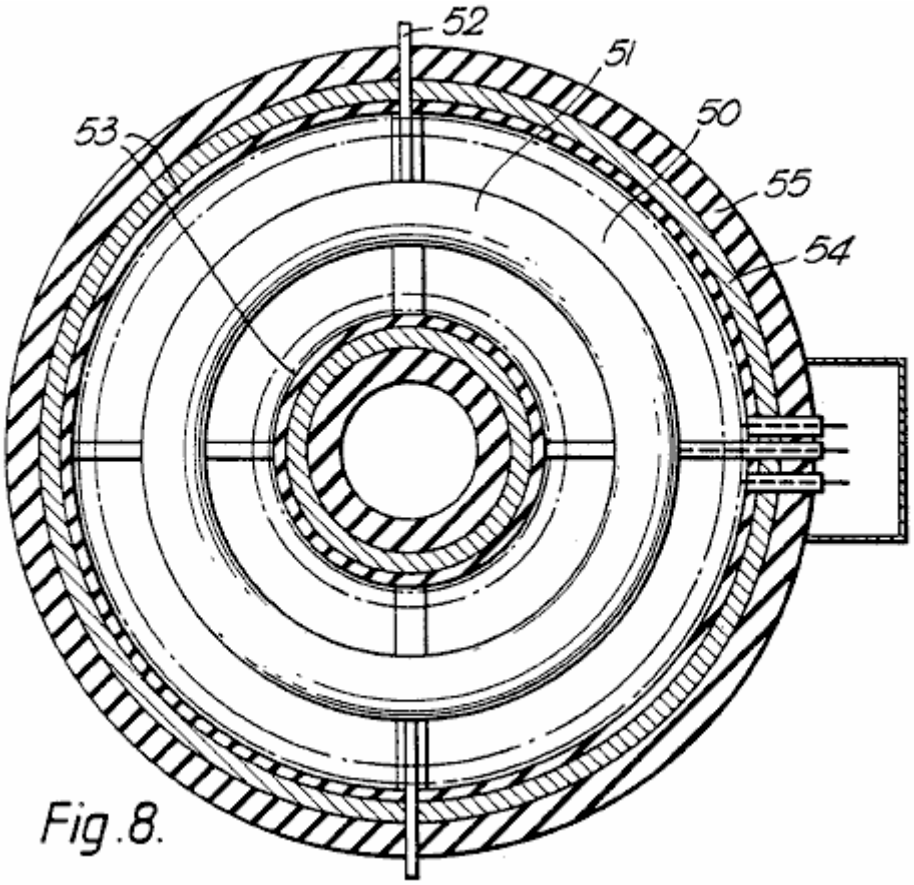


Fig.8.

Fig.8 is a section on line **A--A** of **Fig.7**, and

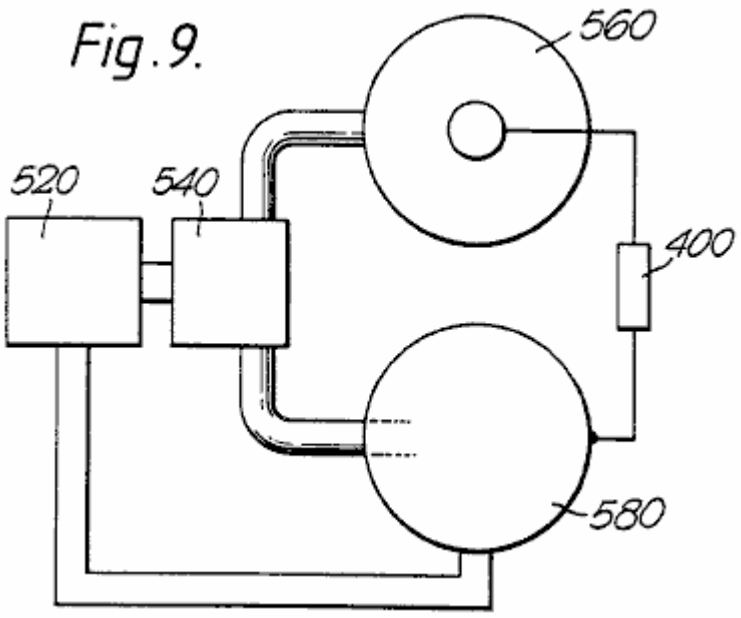


Fig.9.

Fig.9 is a scheme of a two-stage converter, using both forms of charged particles concurrently.

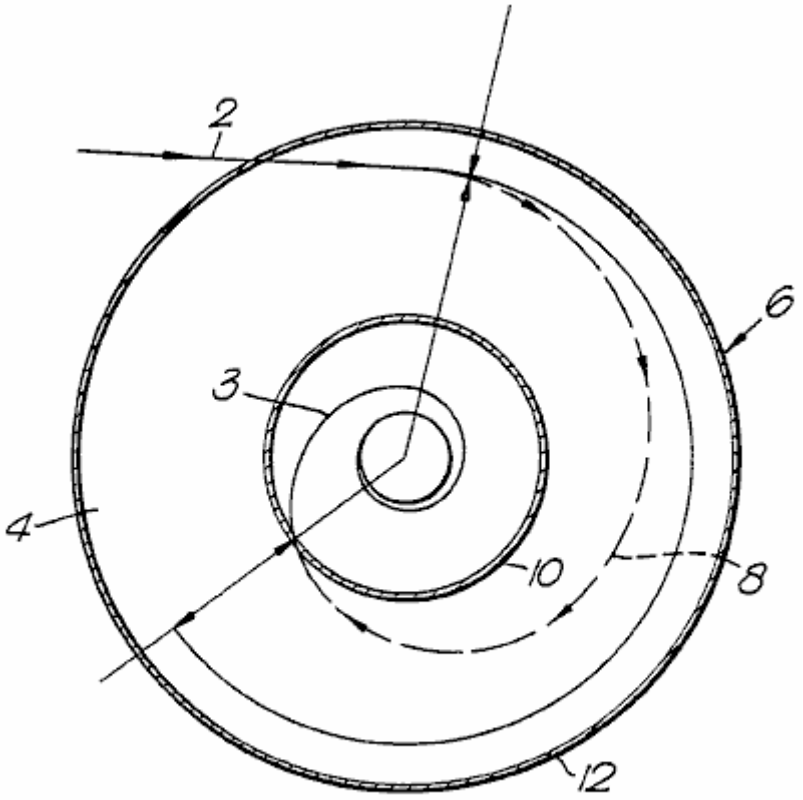
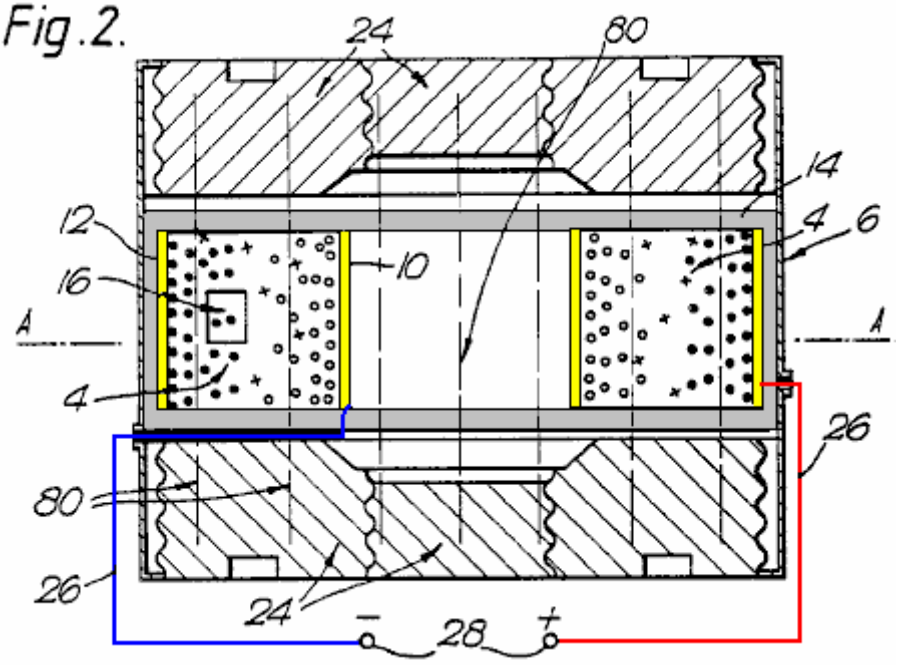


Fig.1.

As shown in **Fig.1**, a charged particle is injected along a trajectory **2** into a magnetic field extending normal to the plane of the drawing. The field permeates the space **4** of the annular cross-section within a cylindrical chamber **6**. The magnetic field produces a force on the particle, extending at right angles to both it's direction of motion and the magnetic field. The resultant centripetal force causes the particle to follow a spiral path **8** ending on the central electrode **10** spaced radially inwards from the outer cylindrical electrode **12**. The extra energy acquired by the particle is a function of the radial distance travelled and the strength of the magnetic field between the electrodes. This energy is given up on impact with the central electrode, in the form of heat and/or work done in bringing the charge against the opposing electric field to the electrode. In the absence of the central electrode **10**, the electrons would follow the orbit of equilibrium **3**, this being the orbit followed by a particle when the centrifugal and centripetal forces balance, resulting in no work being done on the particle.

Fig.2.



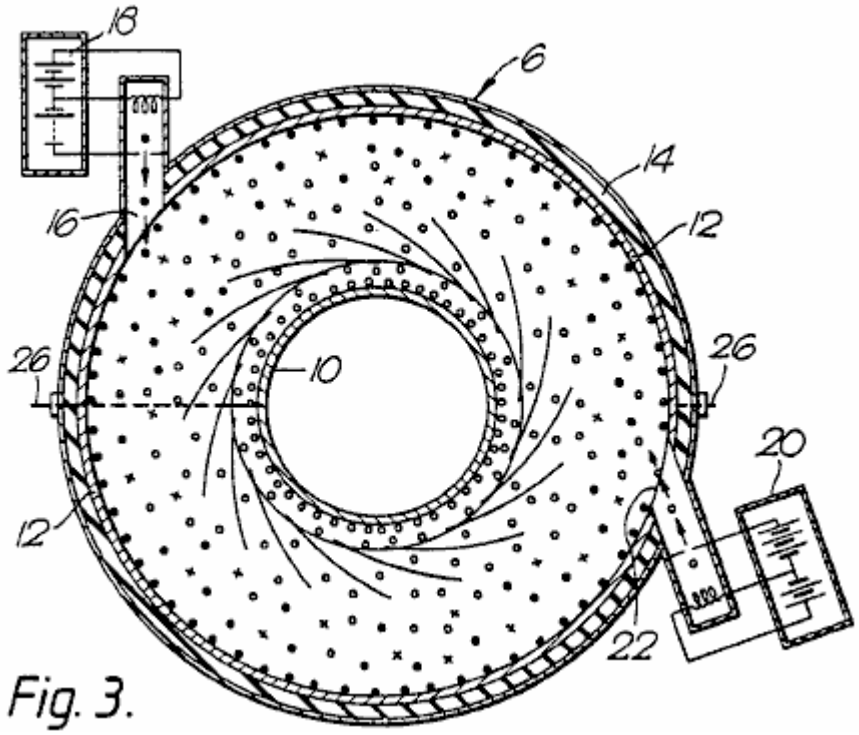


Fig. 3.

As shown more particularly in **Fig.2** and **Fig.3**, the energy converter **1** consists basically of a annular chamber **6** having an outer cylindrical electrode **12**; an inner cylindrical electrode **10**, and two gas-tight walls **14** of electrical insulation material. In the electrode **12** is a port **22** through which an electron gun **20** can inject electrons into space **4**. Additionally or alternatively, an ion gun **18** can inject positively-charged particles through port **16**.

Seated on the major flat surfaces of chamber **6** are magnetic pole-pieces **24** giving rise to a uniform magnetic field **80** which traverses the space **4** parallel with the axis of chamber **6**. The magnets may be ceramic permanent magnets, or they may be electromagnets. In either case, means (not shown) may be provided for adjusting the magnetic field strength.

Heavy conductors **26** connect the two electrodes to terminals **28** across which a resistive load can be placed to dissipate the generator output.

A vacuum pump (not shown) has its inlet in communication with the interior of chamber **6** so that the gas pressure in the generator can be reduced to, and kept at, a desired sub-atmospheric value. Associated with the pump, or separate from it, may be means for ensuring that the gas in the generator is of a desired composition, for instance, one which enhances the possibility of ionising collisions between the charged particles and gas atoms or molecules. One such suitable gas would be neon containing 0.1% argon by volume.

In order to cause the generator to start working, it is necessary to start the vacuum pump and to energise the electron gun or each particle source. The latter involves heating a filament from an external source of power until the required internal energy level (temperature) is reached which in turn causes a piece of thermo-emissive material to emit electrons. If the electrons are to be the charge carriers, they are accelerated by a suitable electric field and projected into the space **4**. Here they are further accelerated by the radial electric field between the electrodes, and at the same time have a deflecting force applied to them by the axial magnetic field through which they pass.

For an ion source, the electrons are accelerated until they impact some atoms or molecules, to produce a stream of ions which likewise pass into the space **4**. With the polarities shown, the electrons are attracted to the central electrode, while the ions are pulled towards the outer electrode, which accounts for the different orientations of sources **18** and **20**.

Any gas molecules which pass close to, or between, the electrodes are ionised by collision and/or the electrostatic field. Output current can then be taken through a load impedance connected across terminals **28**. The impedance is matched to prevent the internal process energy dropping below a value which would prevent the re-ionisation of the gaseous atoms. As each ion is deionised at the anode, the gas atoms will tend to continue to circulate until re-ionised, the resultant force drawing both the ions (shown by solid circles) and electrons (shown by hollow circles) back into their respective orbits.

It is envisaged that, in the case of a converter using electrons, the chamber could be evacuated to a chosen sub-atmospheric pressure and sealed.

In that form of the invention shown in **Fig.4**, each electron source forming one of a circular series **29** of sources has a body **30** of electro-emissive material, such as molybdenum coated by caesium, heated by an electric filament **32** connected in series or parallel across a source of electric power (not shown). Immediately in front of each emitter **30** is a grid **34** of fine wires, all the grids being connected with a source of adjustable voltage so as to control the flow of electrons from the emitter. These electrons are projected through one or more acceleration electrodes **36** across which a potential difference is established along the electron path, so that each incremental electron source injects a stream of electrons having known kinetic energy into a space **38**, indicated by the circle shown in a broken line, traversed by the deflection magnetic field, within which is the central, target, electrode **40**. The stream of electrons injected into the magnetic field may be focused by electric and/or magnetic fields.

In the remaining Figs, those parts already referred to will retain the same references.

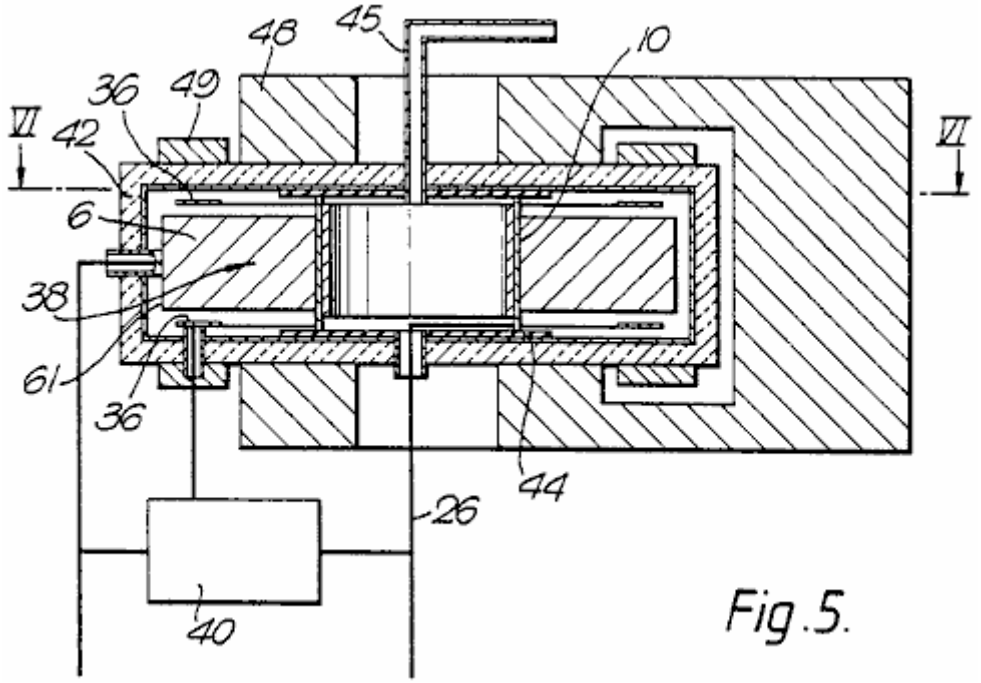


Fig.5.

In the "flat disc" configuration shown in **Fig.5**, the annular chamber **6** is enclosed in a body **42** of thermal insulation material. The central electrode **10** is seated on insulators **44** which are pierced by conduits **45** for the passage of a coolant fluid and by an output lead **26**, which may extend along the conduit so that it too is cooled.

Fig.5 shows how the deflection magnet is generally U-shaped, and has two annular pole-pieces **48**, so that the magnetic field is uniform between the surface of electrode **10** and the region **38** radially innermost of the circular electron source, the electric field between the electrode **36** and emission surface **61** providing the electrons initial accelerations (kinetic energy). **Fig.5** also shows how a voltage is tapped off the resistive load **40** (which thus functions as a potentiometer) and is fed through to the acceleration electrode **36**.

Chamber **6** is also provided with two annular magnets **49** (or a circular series of incremental magnets) designed to influence the direction along which the electrons pass into space **38**. The magnets provide local magnetic fields to ensure that the electrons meet the boundary of space **38** tangentially, i.e. with zero radial velocity.

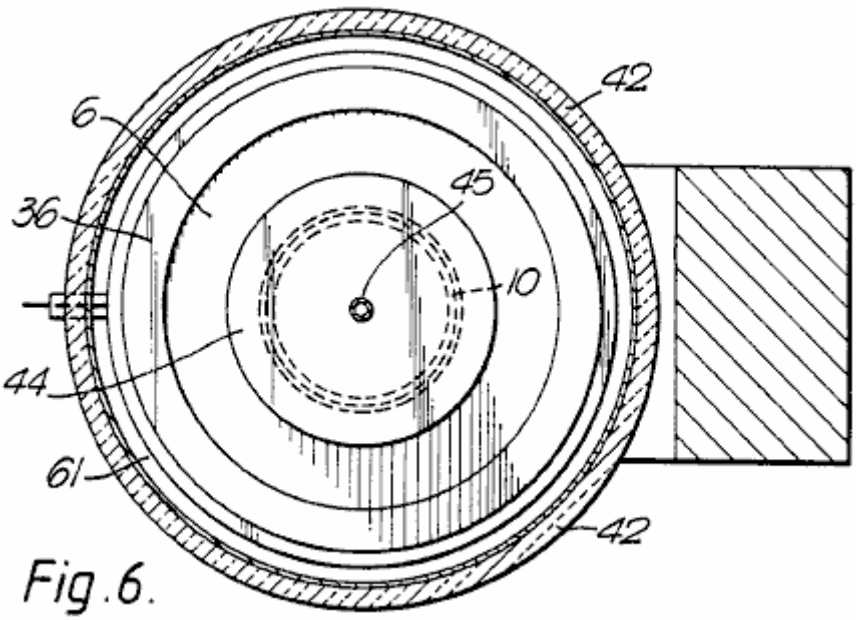
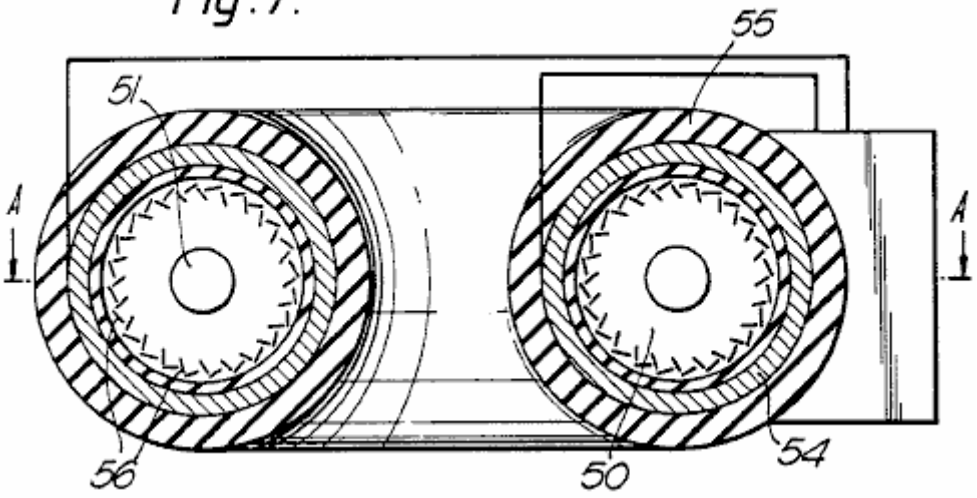


Fig. 6.

Fig. 7.



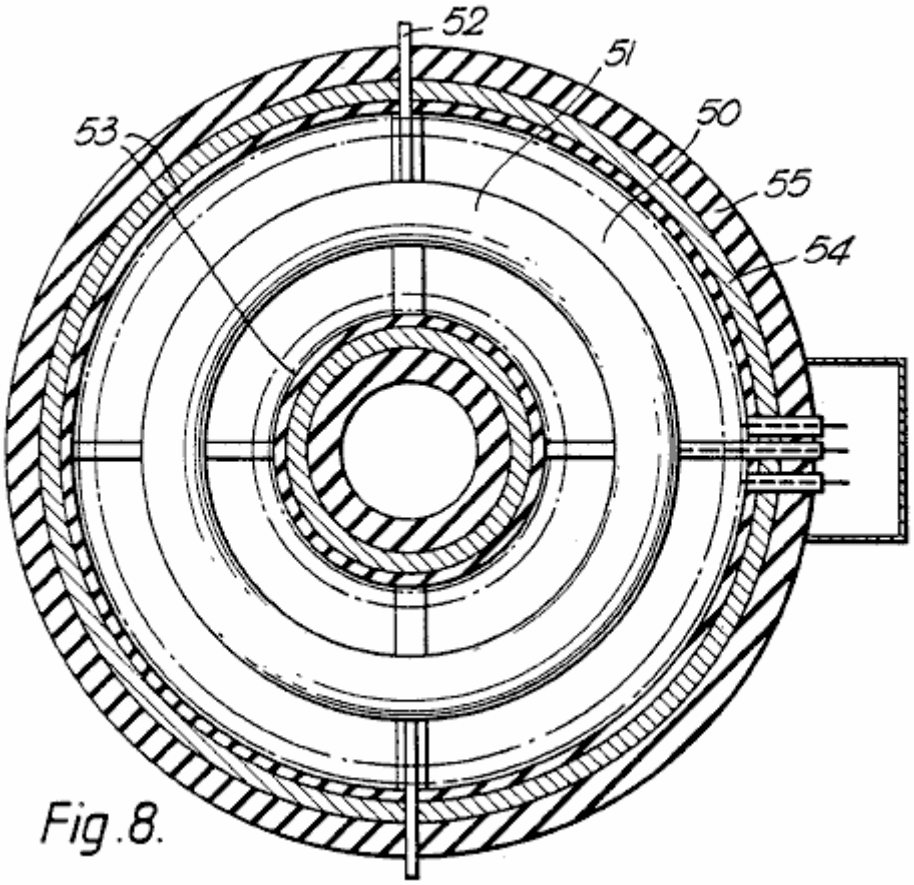
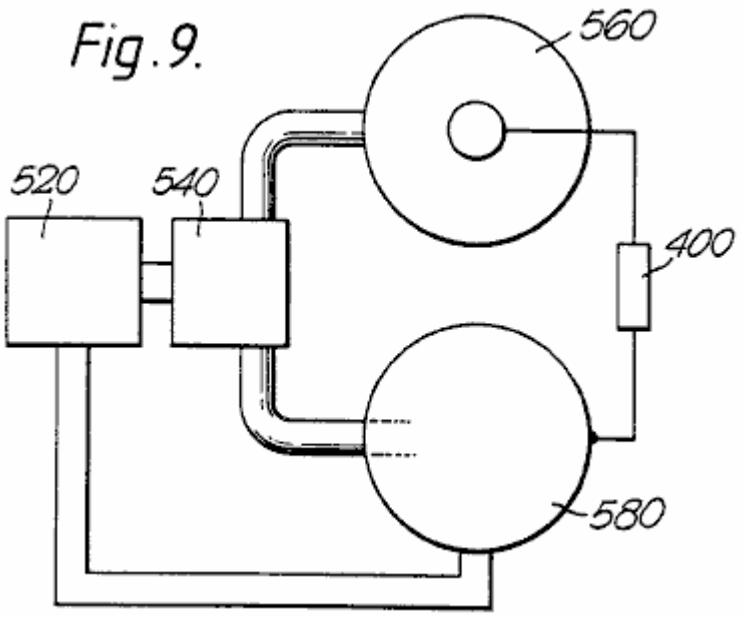


Fig.8.

In that form of the invention shown in **Fig.7** and **Fig.8**, the individual "flat disc" converters of **Fig.5** and **Fig.6** are arranged in a type of "circular" construction, such that the magnetic fields extend along the axis of the resulting toroidal space **50** penetrated by a single toroidal target electrode **51** through which a coolant fluid may pass, along conduits **52**. The cross-section of **Fig.8** shows that the magnetic fields are supplemented by an electric field produced by windings **53** wound on a magnetic core **54** bounded by insulation **55**.

Apart from the fact that the electrodes are common to all converters, each functions individually as described above. Obviously the power source driving the heaters for the electron guns **56**; the electromagnets (if any); the acceleration electrodes and the control grids, have to be of sufficient capacity to supply the greater power needed to drive this "toroidal" configuration. Some changes would need to be made to the physical dimensioning and positioning of the relatively-complex construction, but as all these are readily understood by a competent engineer, they are not further described in this specification.



As already mentioned, the converters of this invention are of two types, i.e. electronic and ionic. **Fig.9** shows diagrammatically how they may be combined to take advantage of their differences. In the two-stage power generation apparatus shown in **Fig.9** the first stage consists of an ioniser **520** supplying a mixture of charged particles, i.e. ions and electrons, to a separator **540**, which supplies electrons to a second stage consisting of a sealed electronic converter **560** in parallel with a gaseous ionic converter **580**.

The separator **540** may use the different particle masses to separate them centrifugally using, for example, the energy conversion system of **Fig.1** (without the target electrode), or it may use electromagnetic deflection fields, or a physical diffusion process, either alone or in combination. As this is not part of the subject-matter of this invention, it will not be described herein in any further detail.

In the generators of **Fig.6** and **Fig.8**, the respective particles are deflected magnetically and accelerated radially, to function as already described above.

Because each generator is designed to operate most effectively with its particular form of charge carrier, it can be designed optimally, thus reducing the energy absorption caused by ions and electrons recombining before each has fallen on its respective target electrode. Because the electronic converter would finish up with a negatively-charged electrode, whereas the converse is true for the ionic converter, the load **400** extracting energy from the apparatus is connected across the two target electrodes. The other two electrodes of the converters may be held at the same potential, as by being connected together, or their potentials may float.

The generator can be designed to produce a wide range of output voltages and currents. The lower-energy generators are light enough to be mobile, so that they can power vehicles or act as stand-by generators. Various electrode and magnet configurations can be used, and the generators can be connected in series or parallel. Cooling jackets are fitted to prevent overheating in high-powered apparatus, and the generator is enclosed within a thermally-insulating jacket to reduce heat losses thereby increasing particle velocities. For high-energy generators, it may be necessary to provide for forced cooling of the inner electrode, as by fins projecting therefrom into a high-speed stream of suitable coolant.

Although the process according to this invention is particularly suited to using external electrical energy, it must be understood that other sources can be used to provide the initial energy input, e.g. solar and waste process heat are some of the varied energy sources which could be utilised. Control of the charge-generation process can be achieved by other means, including one or more electrically-biased grids, as used in thermionic valves.

CLAIMS

1. An energy conversion process for generating an electric potential, the process comprising; providing a source of electric charge carriers of predetermined polarity, accelerating the carriers away from the source, introducing the carriers into a magnetic field transverse to the path of the carriers in a process chamber, the field bounding an inner electrode within the chamber such that the carriers orbit the electrode while accelerating radially toward the electrode; and converting the resulting increased kinetic energy of the carriers

into an electric potential at the electrode before the carriers reach an orbit of equilibrium in which the centripetal force is balanced by the centrifugal force on the carriers.

2. A process according to claim 1 in which the electric potential is created between the inner electrode and an outer electrode radially spaced from the inner electrode.

3. A process according to claim 2 in which the outer electrode provides the said source of the charge carriers.

4. A process according to claim 1 or claim 2 in which the chamber is maintained at a sub-atmospheric pressure.

5. A process according to claim 1 in which the electric potential drives a load connected between the inner electrode and a point remote from the electrode.

6. A process according to claim 1 or claim 2 in which the electric charge carriers comprise electrons or ions.

7. A process according to claim 1 in which further charge carriers of the opposite polarity traverse the magnetic field and accumulate at a second electrode to increase the potential difference between the two electrodes.

8. A process according to claim 1 in which electrically biased grids control the flow of the charge carriers from the source.

9. A process according to claim 1 in which the charge carriers are separated from charge carriers of the opposite polarity before being introduced into the magnetic field.

10. A process according to claim 9 in which the charge carriers of opposite polarity are introduced into a corresponding second magnetic field, whereby a potential difference is produced between respective electrodes in each field.

11. A process according to claim 1 in which the carriers are injected into the magnetic field.

12. A process according to claim 11 in which the injection energy is produced by accelerating the carriers through an electric field.

13. A process according to claim 11 in which the injection energy is produced by accelerating the carriers through a magnetic field.

14. A process according to claim 1 in which the injection energy of the carriers is produced by nuclear emission.

15. A process according to claim 1 in which the injection energy of the carriers is produced by heat.

16. A process according to claim 1 in which the generated electric potential is directly or indirectly used to maintain the generation of charge carriers or the internal temperature of the space traversed by the magnetic field, or the applied magnetic field.

17. A process according to claim 1 in which the generated electric potential is directly or indirectly used to maintain the generation of charge carriers and the internal temperature of the space traversed by the magnetic field and the applied magnetic field.

18. An energy converter including a source of electric charge carriers of a predetermined polarity, a process chamber having an inner electrode, means for accelerating the carriers away from the source and for introducing the carriers into the chamber, means for applying a magnetic field transverse to the path of the carriers and bounding the inner electrode of the chamber such that the carriers orbit the electrode while accelerating radially toward the electrode, the electrode being located at a radius which exceeds the equilibrium radius for the carrier mean velocity and applied field strength and intercepting the carriers such that the increased kinetic energy of the carriers due to centripetal acceleration is converted to an electric potential at the electrode.

19. An energy converter according to claim 18 in which the chamber includes an outer electrode spaced radially from the inner electrode, and means for injecting the charge carriers into the space between the electrodes.

20. An energy converter according to claim 19 in which the outer electrode provides the said source of charge carriers.

21. An energy converter according to claim 19 further comprising an insulating wall bounding the outer electrode.

22. A converter according to claim 18 further comprising means for maintaining the chamber at a predetermined sub-atmospheric pressure.

23. A converter according to claim 19 in which the outer electrode has at least one port through which the charge carriers can be injected into the chamber along a desired trajectory.

24. A converter according to claim 23 in which the outer electrode has plural ports and each port communicates with a thermionic source of the respective carriers.

25. A converter according to claim 18 in which the chamber is a vacuum chamber.

26. A converter according to claim 18 further comprising electrically biased grids for controlling the flow of charge carriers from the source.

27. A converter according to claim 22 or claim 25 in which the evacuated chamber comprises a sealed unit.

28. A converter according to claim 18 further comprising means for adjusting the strength of the applied magnetic field.

29. A converter according to claim 18 in which the chamber is filled with low pressure gas.

US Patent 3,913,004 **14th October 1975** **Inventor: Robert W. Alexander**

METHOD AND APPARATUS FOR INCREASING ELECTRICAL POWER

This patent shows a method of altering a standard electrical generator intended to be driven by a separate motor, so that it operates without the motor. In an example quoted, a DC input of 48 volts at 25 amps of current (1.2 kW) produces a 110 volt 60Hz AC output of 3.52 kW. That is a Coefficient Of Performance of 2.93 at an output level suited to Off-The-Grid operation of a house.

ABSTRACT

A form of rotating machine arranged in such a way as to convert a substantially constant input voltage into a substantially constant output voltage; involving generally, a rotor that revolves at a substantially constant speed within a stator, and which comprises a transformer core subjected to and having a primary motor-transformer winding and a secondary transformer-generator winding; whereby transformed and generated power are synchronously combined as increased output power.

BACKGROUND

Electrical power is frequently changed in voltage, phase, frequency, and the current is changed from alternating to direct or from direct to alternating. Voltage conversion in AC circuits is usually by means of transformers, and in DC circuits is usually by means of motor-generators. Phase conversion is also accomplished by either transformers or motor-generators, and frequency conversion is most simply done by motor-generators.

Motor-generators have various classifications of use, as follows:

(1) DC to DC, used to charge batteries and to boost voltage.
(2) AC to AC, used for frequency and phase conversion
(3) AC to DC used for all types of service, such as battery charging, generator and motor field excitation, railways, electrolysis, and speed control etc. and
(4) DC to AC used to limited extent for special applications.

To these ends combination motor-generators have been built, such as dynamotors stepping up DC voltage for radio equipment and amplidynes for reproducing a weak signal at a higher power level. When a particular variable frequency A.C. is required of a motor-generator set and the power supply is DC, the equipment will include a DC motor for variable speed and a separate alternator driven by it. Such equipment is special in nature and characterised by separation of the motor and generator and by polyphase (usually three-phase) generator windings and with auto transformers having suitable taps for obtaining the required voltages; and a DC speed controller for the motor. The phase output of such equipment is selective and its single phase capacity necessarily restricted (66%) as compared with its three-phase capacity, in which case transmission efficiency for single phase is poor.

When a higher level power output is desired, the amplidyne is employed with field windings and brushes equipped for the purpose, and in some instances to give a constant current output from a constant voltage input, for example, in inverted rotary converter provided to convert DC to A.C. However, the present invention is concerned with method and apparatus for increasing electrical power and provides a dynamo-electric converter that operates from an electrical energy supply to produce A.C. most efficiently for a useful load.

The method involves simultaneous motor-transformer-generator steps and the preferred embodiment of the apparatus involves a dynamo-electric converter (DEC) in the form of a rotary machine combined in a single rotor revolving within a stator, the rotor being comprised of a transformer core having both a primary motor-transformer winding and a secondary transformer-generator winding, and the stator being comprised of magnetic field poles.

Synchronous converters have been combined in single rotor machines to produce DC from A.C., but that effect is quite different from the effect of the present invention when A.C. is to be produced from DC in a single rotor having primary and secondary armature windings as distinguished from armature windings common to both A.C. and DC circuits. With the present invention, both a transforming and a generating effect are produced in the rotor, all of which is inherently synchronised and delivered through the A.C. outlet leads. A.C. motors and DC generators have been combined in one machine, that is in one rotor, and referred to as synchronous converters.

However, synchronous converters are lacking in their ability to change DC into A.C. when operating from the former as a prime mover to drive a generator simultaneously, and more specifically to drive an alternator synchronously.

SUMMARY OF INVENTION

This method involves the placement of a primary winding in a field to both motor the same and to have a transformer effect with respect to a secondary winding also in a field to have a generator effect. In its preferred embodiment, this dynamo-electric converter is comprised of primary and secondary windings combined in a rotor commutated to alternate a DC energy supply in and thereby motivate the rotor within a stator field. The primary winding is advantageously of fewer turns than the secondary and by means of electromotive force drives the secondary windings of more turns to cut the magnetic lines of force for the generation of electrical energy at a higher voltage level than the DC supply. This DC operated motor is shunt wound with the stator field poles fully energised by the DC energy supply, or is provided with permanent magnet field poles, to efficiently motivate the rotor and efficiently generate electrical energy in the secondary windings. The A.C. output of the secondary windings is inherently synchronised with the transformer function of the primary windings combined in the common slots of the single rotor; and by adding the transformer and generator voltages and amperages the wattage is correspondingly increased at the output.

DRAWINGS

The various objects and features of this invention will be fully understood from the following detailed description of the typical preferred form and application, which is made in the accompanying drawings, in which:

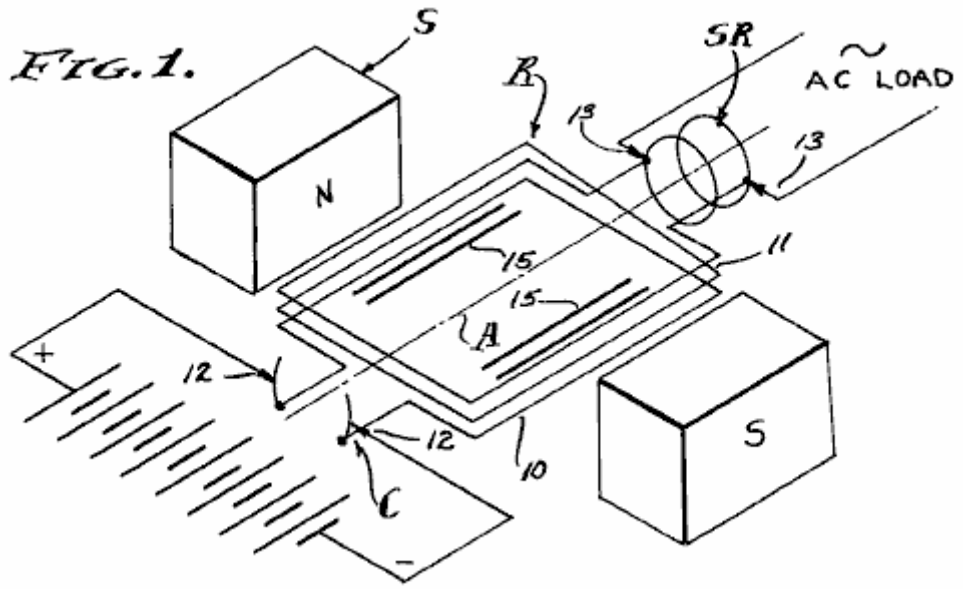


Fig.1 is a diagrammatic schematic view of the dynamo-electric converter components comprising the present invention.

FIG. 2.

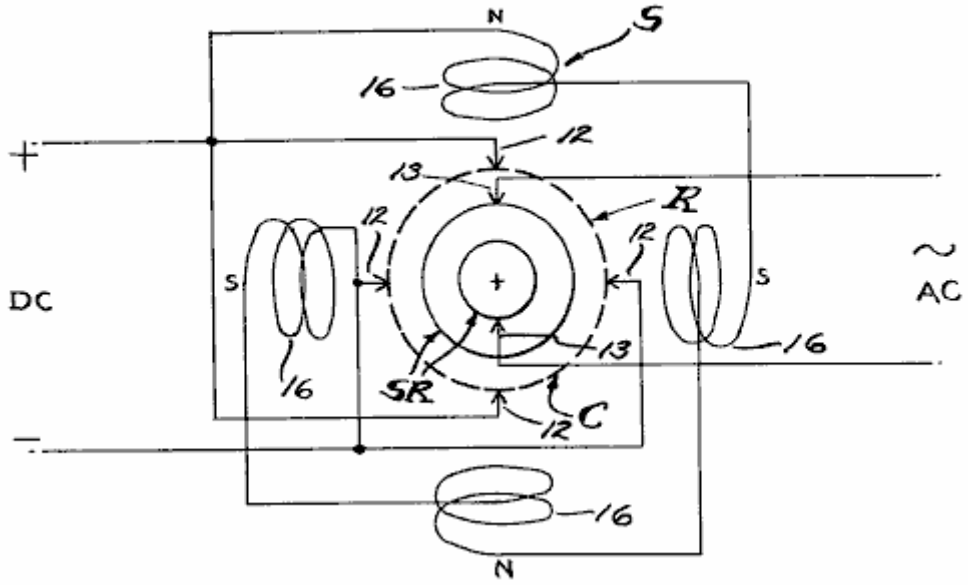


Fig.2 is a diagram of a typical commutator brush, slip ring brush and field pole arrangement which is utilised.

FIG. 3.

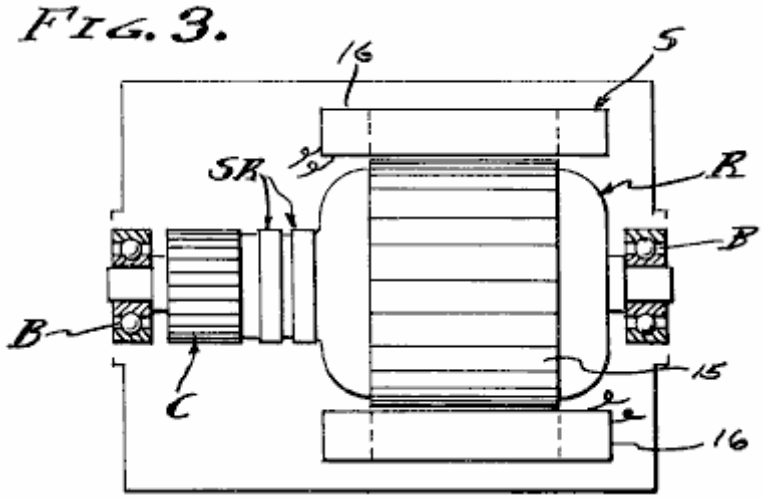


Fig.3 is a longitudinal section through a machine embodying the stator and rotor on bearings with the frame and brushes removed.

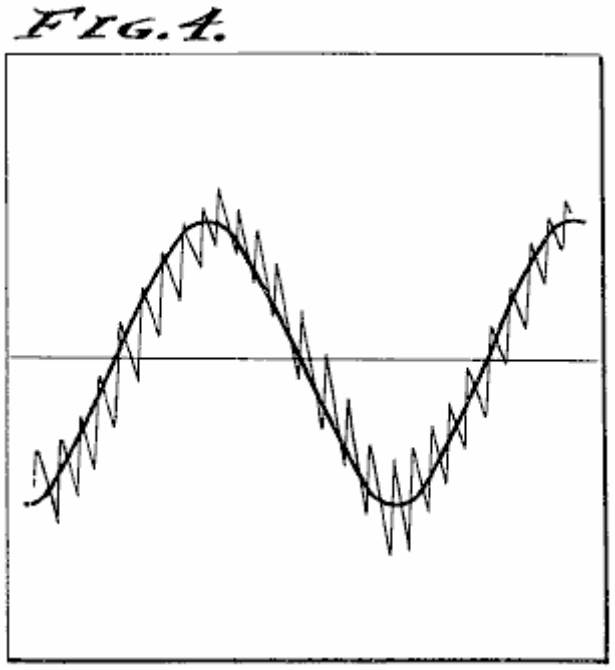
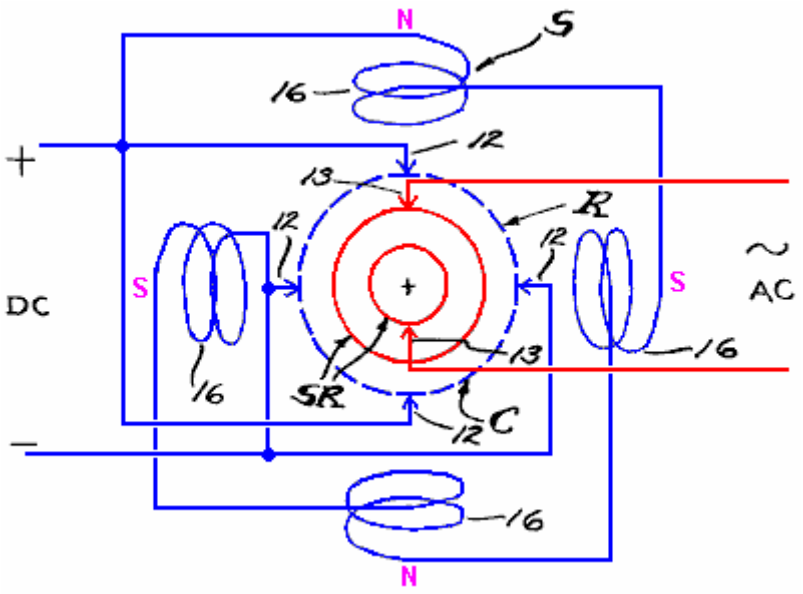


Fig.4 is a typical duplicate of an oscilloscope diagram showing the power output of the dynamo-electric converter.

PREFERRED EMBODIMENT



The dynamo-electric converter is illustrated diagrammatically in the drawings and involves, generally, a rotor **R** carried upon spaced bearings **B** so as to rotate on an axis **A** concentric within a stator **S**. The rotor **R** comprises the armature, while the stator **S** comprises the field, there being a commutator **C** associated with primary windings **10** on the rotor and slip rings **SR** associated with secondary windings **11** on the rotor. Brushes **12** and **13** are engaged slideably with the commutator and slip rings respectively, by conventional means, to conduct DC through the commutator **C** and to conduct AC through the slip rings **SR**. The brushes **12** and interconnected primary windings **10** comprise a motor while the brushes **13** and interconnected secondary windings **11** comprise a generator or alternator.

In practice, the field windings **16** can be separately energised or connected in parallel with the brushes **12** or shunted with respect to the primary motor winding **10**. Motorisation of the armature rotor **R**, or motoring thereof, causes continued polarity reversals on a cycle basis as determined by the speed of rotation, and this of course results in magnetic reversals in the rotor core **15** and a consequent induction in the secondary windings **11**. A feature of this invention is the combining and co-operative relationship of the primary and secondary windings which occupy common slots in and embrace a common portion of the core **15** of the rotor **R**, thereby to have a

transformer function as well as a generator function as the lines of magnetic force are cut by the secondary windings. The stator **S** has field poles of opposite magnetic polarity, excited independently from the armature, or as permanent magnets, and preferably shunted across the DC input. As shown, there are four equally spaced field poles in a circumferentially disposed series.

In practice, the primary DC motor windings are of fewer turns in the rotor slots than the secondary AC generator windings. For example, the primary motor windings **10** are flat wound between north to south poles of the field while the secondary generator windings are flat wound in the same or common slots of the rotor armature. In a typical unit having a four brush commutator with 20 bars and having a 20-slot armature, the primary windings **10** are comprised of a number of turns of conductor efficiently to draw 48 volts DC at 25 amperes or 1,200 watts to rotate at 1,750 rpm, while the secondary windings **11** are comprised of a number of turns of conductor efficiently to deliver 60 cycle (by transforming and generating) 110 volts AC at 32 amperes or 3,520 watts, the volt meter used to read these values upon an actual reduction to practice being calibrated to read the root-mean-square (RMS) value of the pure sine wave, which is 70.7% of the peak voltage.

The reduction to practice previously referred to as a "typical unit" was constructed of a machine originally designed as a self-exciting 60 cps 110 volt 2.5 kVA generator to be shaft driven by a separate prime mover. Firstly, the prime mover was eliminated. The exciter windings were intended to excite the field at 45 volts DC delivered through the commutator, while the generator windings were intended to independently deliver 110-120 volts AC through the slip rings. The winding ratio between the exciter and generator windings was approximately one to three, and these are the values which determined the values employed in the present reduction to practice. However, it is to be understood that other values can be employed by design, for operation at the desired input and output voltages and amperages. It is also to be understood that the example reduction to practice disclosed herein is not necessarily the optimum design, in that other input-output power balances are contemplated, such as a DC battery input voltage substantially equal to the AC power voltage. In any case, an unexpected increase in power is realised by practising this invention.

This dynamo electric converter inherently operates at a substantially constant angular velocity with the result that the alternating cycles of the output are substantially constant. Also, the DC input voltage can be maintained at a substantially constant level with the result that the AC output voltage is also substantially constant. As shown, the output is single phase AC in which case the effective power in watts delivered is the product of current, voltage and power factor. Since the voltage is substantially constant, the current varies with load applied to the output as it is affected by the power factor. It will be seen therefore, that the apparent power represented by voltage times amperage is drawn directly from the DC input and applied to the primary motor winding **10** to motivate the rotor **R** for the functions previously described. It will also be seen therefore, that the DC input is commutated into AC and transformed by induction from windings **10** into windings **11**.

It will also be seen therefore, that the AC generated by motorisation of the motor is synchronously imposed upon the windings **11**, and all to the end that the two alternating currents are complementary and one added to the other. It will be observed that the output wattage is approximately triple the input wattage, by virtue of the synchronous superimposing of transformed input voltage and generated voltage while utilising the former to operate the rotor in order to generate the latter. A feature of this invention is the separation of the primary and secondary circuits and the consequent isolation of the inverted input DC from the outlet AC and the utilisation of input energy commensurate with output load according to amperage required for the operations to which this DEC machine is applied.

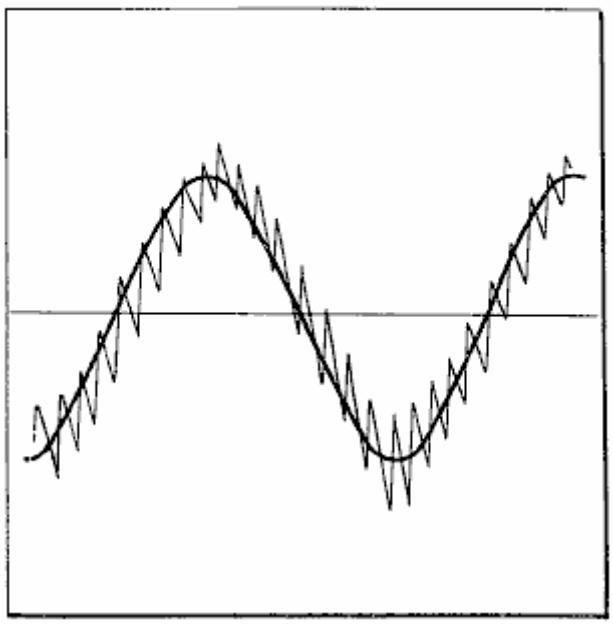
In carrying out this invention, the dynamo electric machine is conventional in design and the primary and secondary windings **10-11** are wound into the common slots of the armature as they are in self exciting generators. However, the primary windings **10** are motor-transformer windings and function totally as such. Similarly, the secondary windings **11** are wound into the armature slots together with the primary windings **10** and are powered with current that is alternated by virtue of the commutation and rotation of the armature, and consequently there is a transformer action between the primary windings **10** and secondary windings **11**, and this transformer function is supplemented by generation of a superimposed current by virtue of the secondary windings **11** cutting the magnetic lines of force provided by the surrounding stator field. Consequently, there is a multiplying of power synchronously applied through the slip rings **SR** to the output brushes **13**, and this increased output power is measurable as previously described and double or almost triple that of the input power.

METHOD

Referring now to this method of increasing electrical power, input alternating current is applied to a primary winding to both motor and alternately magnetise a core. The said primary winding is immersed in a field and consequently is caused to motor and simultaneously to perform the first stage of transforming. A second stage of transforming is then performed by a secondary winding associated with the core to function as both a transformer and a generator winding, and the output current is drawn from it at an increased power value as compared with the input power, since the current induced by transformer action is superimposed upon the current generated in

cutting the magnetic lines of force by motoring the secondary winding through the magnetic field. The direct application of AC power to the primary winding is contemplated, however the present and preferred embodiment employs commutation of DC power which is thereby inverted to AC power in the process of motoring the windings and the core in which they are carried together with the secondary winding. The net result is three fold, in that there is a motoring function, a transforming function, and a generating function, all of which are inherently synchronised to increase the output power with respect to the input power.

FIG. 1.



From the foregoing it will be seen that this method, and the dynamo-electric converter termed a DEC, synchronously superimposes transformed electrical energy and mechanically generated electrical energy when inverting DC to AC as is shown by observing the oscilloscope diagram duplicated in **Fig.4** of the drawings. The DC motor section of the rotor-stator unit will operate at its designed speed well within a small tolerance, by applying known engineering principles, and consequently, the AC generator-alternator section will operate at a substantially uniform frequency of, for example, 60 cycles per second. Thus, the output voltage potential is kept to a maximum while current is drawn as required, within the design capacity of the unit.

Having described only a typical preferred form and application of my invention, I do not wish to be limited or restricted to the specific details herein set forth, but wish to reserve to myself any modifications or variations that may appear to those skilled in the art:

CLAIMS

1. A dynamo-electric converter for inverting direct current voltage to alternating current voltage and including; a magnetic field having poles of opposite polarity, an armature coaxial with the field and having a core with means to receive windings, coaxial bearing means between the field and the armature, a primary motor-transformer winding in said means of the armature core and a commutator connected therewith, direct current input brushes which can be engaged with the said commutator, a secondary transformer-generator winding in said means of the armature core and slip rings connected therewith, and alternating current output brushes which can be engaged with the said slip rings, whereby direct current input power is both transformed and regenerated as alternating output power.

2. The dynamo-electric converter as set forth in claim 1, wherein the magnetic field is a stator comprised of said poles of opposite polarity, and wherein the armature is a rotor supported upon said bearing means coaxially within said field.

3. The dynamo-electric converter as set forth in claim 1, wherein the means to receive windings is a pair of slots in the armature core, said primary and secondary windings being carried in the slots and subjected to the magnetic capabilities of the core.

4. The dynamo-electric converter as set forth in claim 1, wherein the means to receive windings is a multiplicity of slots disposed in a circumferential series about the armature core, said primary and secondary windings being

circumferentially progressive windings respectively and carried in common slots respectively and subjected to the magnetic capabilities of the core.

5. The dynamo-electric converter as set forth in claim 1, wherein the magnetic field poles are permanent magnets.

6. The dynamo-electric converter as set forth in claim 1, wherein the magnetic field poles are electro magnets energised separately from the said primary motor winding.

7. The dynamo-electric converter as set forth in claim 1, wherein the field poles are electro magnets energised in parallel with the direct current input brushes which can be engaged with the commutator.

8. The dynamo-electric converter as set forth in claim 1, wherein the magnetic field is a stator comprised of said poles of opposite polarity, wherein the armature is a rotor supported on said bearing means coaxially within said field, and wherein the means to receive windings is a pair of slots in the armature core, said primary and secondary windings being carried in the slots and subjected to the magnetic capabilities of the core.

9. The dynamo-electric converter as set forth in claim 1, wherein the magnetic field is a stator comprised of permanent magnet poles of opposite polarity, wherein the armature is a rotor supported on said bearing means coaxially within said field, and wherein the means to receive windings is a pair of slots in the armature core, said primary and secondary windings being carried in the slots and subjected to the magnetic capabilities of the core.

10. The dynamo-electric converter as set forth in claim 1, wherein the magnetic field is a stator comprised of permanent magnet poles of opposite polarity, wherein the armature is a rotor supported on said bearing means coaxially within said field, and wherein the means to receive windings is a multiplicity of slots disposed in a circumferential series about the armature core, said primary and secondary windings being circumferentially progressive windings and carried in common slots respectively and subjected to the magnetic capabilities of the core.

11. The dynamo-electric converter is set forth in claim 1, wherein the magnetic field poles are electro magnets of opposite polarity energised in parallel with the direct current input brushes which can be engaged with the commutator, wherein the means to receive windings is a multiplicity of slots disposed in a circumferential series about the armature core, said primary and secondary windings being circumferentially progressive windings respectively and carried in common slots respectively and subjected to the magnetic capabilities of the core.

12. A method for increasing electrical power and comprised of; placing a primary winding within the flux of a magnetic field and applying alternating current therethrough while motoring the same to revolve, simultaneously revolving a secondary winding with the primary winding and through a flux of a magnetic field, and simultaneously transforming the first mentioned alternating current from the primary winding and into the secondary winding while synchronously generating alternating current in the secondary winding.

13. The method of increasing electrical power as set forth in claim 12 wherein the magnetic field is held stationary and the primary and secondary windings revolved together.

14. The method of increasing electrical power as set forth in claim 12 wherein the primary and secondary windings are related to a common armature synchronously inducing into and generating electrical power through the secondary winding.

15. The method of increasing electrical power as set forth in claim 12 wherein the first mentioned alternating current is commutated from direct current to alternating current by revolvement of said primary winding.

16. The method of increasing electrical power as set forth in claim 12 wherein the magnetic field is held stationary and the primary and secondary windings revolved together and related to a common armature synchronously inducing into and generating electrical power through the secondary winding.

17. The method of increasing electrical power as set forth in claim 12 wherein the first mentioned alternating current is commutated from direct current to alternating current by revolvement of said primary winding and the primary and secondary windings related to a common armature synchronously inducing into and generating electrical power through the secondary winding.

18. The method of increasing electrical power as set forth in claim 12 wherein the first mentioned alternating current is commutated from direct current to alternating current by revolvement of said primary winding and wherein the magnetic field is held stationary and the primary and secondary windings revolved together and

related to a common armature synchronously inducing into and generating electrical power through the secondary winding.

19. A dynamo-electric machine including; a first means applying a first alternating current into a primary motor-transformer winding, and a second means inducing a second alternating current into a secondary transformer-generator winding, said secondary winding being carried by said second means to operate through a flux of a field and thereby generating a third alternating current, whereby said second and third alternating currents are synchronously superimposed one upon the other.

20. The dynamo-electric machine as set forth in claim 19 wherein the field is stationary and the primary and secondary windings are rotary.

21. The dynamo-electric machine as set forth in claim 19 wherein the field is stationary and the primary and secondary windings are rotary with commutator bars synchronously applying a direct current to motorise the armature and to apply said first alternating current thereto.

22. The dynamo-electric machine as set forth in claim 19 wherein the transformer means comprises magnetic core means common to the primary and secondary windings.

23. The dynamo-electric machine as set forth in claim 19, wherein the field is stationary and the primary and secondary windings are rotary with commutator bars synchronously applying a direct current to motorise the armature and to apply said first alternating current thereto, and wherein the transformer means comprises magnetic core means common to the primary and secondary windings.

24. A rotary dynamo-electric machine including: means applying alternating current through a primary motor-transformer winding carried by an armature core carrying a secondary transformer-generator winding, a field, and bearing means for rotation of the armature core relative to the field, whereby the alternating current applied to the primary winding motors the armature and is transformed and an alternating current generated and superimposed thereon through the secondary winding for increased output power.

25. The rotary dynamo-electric machine as set forth in claim 24 wherein the primary and secondary windings are each comprised of a number of turns of conductor to transform the first mentioned applied alternating current to the voltage of the alternating current generated through the secondary winding.

26. The rotary dynamo-electric motor as set forth in claim 24 wherein the first mentioned applied alternating current is of different voltage than the increased output power and wherein the primary and secondary windings are each comprised of a number of turns of conductor to transform the first mentioned applied alternating current to the voltage of the alternating current generated through the secondary winding.

27. The rotary dynamo-electric machine as set forth in claim 24 wherein the first mentioned applied alternating current is of lower voltage than the increased output power and wherein the primary and secondary windings are each comprised of a number of turns of conductor to transform the first mentioned applied alternating current to the voltage of the alternating current generated through the secondary winding.

SHIGEAKI HAYASAKA: INDUCTION GENERATOR

Patent US 5,892,311 6th April 1996 Inventor: Shigeaki Hayasaka

INDUCTION GENERATOR HAVING A PAIR OF MAGNETIC POLES OF THE SAME POLARITY OPPOSED TO EACH OTHER WITH RESPECT TO A ROTATION SHAFT

This patent covers a device which is claimed to have a greater output power than the input power required to run it.

ABSTRACT

An induction generator having a pair of magnetic poles of the same polarity opposed to each other with respect to a rotation shaft is characterised by a high energy conversion efficiency. The induction generation has a rotation shaft driven by an external means; an even number of (more than three) stator cores provided to encircle the rotation shaft, predetermined gaps being provided between the adjacent stator cores; a first monopole rotor provided in the rotation shaft, surrounded by the even number of stator cores, and having first and second magnetic poles of the same polarity, the first and second magnetic poles being opposed to each other with respect to the rotation shaft in a cross section; a second monopole rotor provided in the rotation shaft so as to face the first monopole rotor at a predetermined distance along the rotation shaft, surrounded by the even number of stator cores, and having third and fourth magnetic poles of the same polarity opposite to the polarity of the first and second magnetic poles, the third and fourth magnetic poles being opposite to each other with respect to the rotation shaft; a plurality of windings provided in the even number of stator cores and connected according to a predetermined configuration.

US Patent References:

282472	Jan., 1883	Delaporte	318 / 197.
2982872	May., 1961	Fredrickson	310 / 163.
3858308	Jan., 1975	Peterson	29 / 598.
4780635	Oct., 1988	Neumann	310 / 216.
5030867	Jul., 1991	Yamada et al.	310 / 156.
5111095	May., 1992	Hendershot	310 / 168.
5402025	Mar., 1995	Saito et al.	310 / 156.

Other References:

Publication "Novel Reluctance Machine Concepts for Variable Speed Drives", Lipo, T.A., Proceedings from the Mediterranean Electrochemical Conference, Ljubljana, May 22-24, 1991, pp. 34-43.

DESCRIPTION

TECHNICAL FIELD

The present invention relates to an induction generator having a pair of magnetic poles of the same polarity opposed to each other with respect to a rotation shaft.

Induction generators have been known as one type of electrical appliance from relatively old days and embodied in various forms adapted for individual applications. In addition to applications in power plants, ships and aircraft, induction generators convenient for household or leisure purposes have also been developed and used extensively.

An induction generator converts kinetic energy into electric energy. Due to a necessity for improving efficiency of energy utilisation, there is a demand for a highly efficient energy conversion.

BACKGROUND ART

As is well known, an induction generator is operated on the principle that an electromotive force is induced in a coil, in proportion to the rate at which magnetic flux crosses that coil (Faraday's law of electromagnetic induction). According to Lenz's law, an induced electromotive force is generated in a direction in which a current that acts against a change in the magnetic flux is generated.

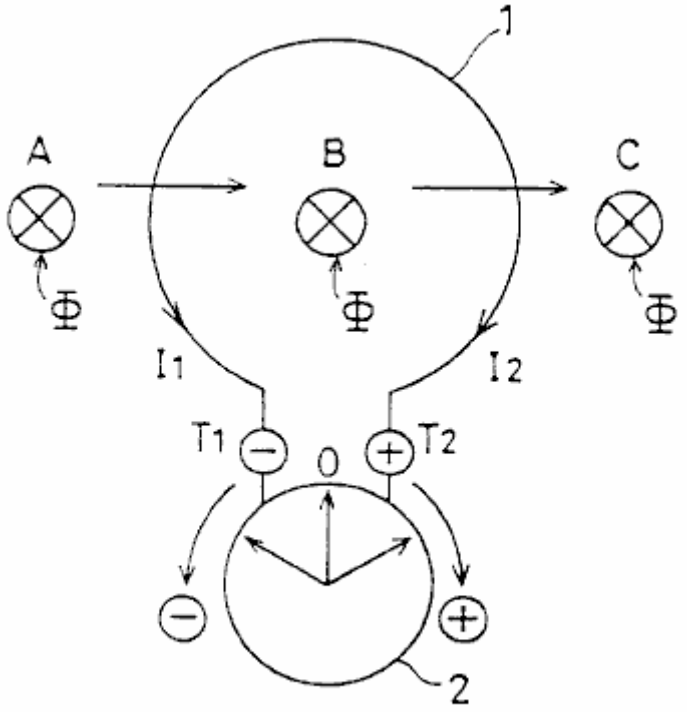


FIG. 1A

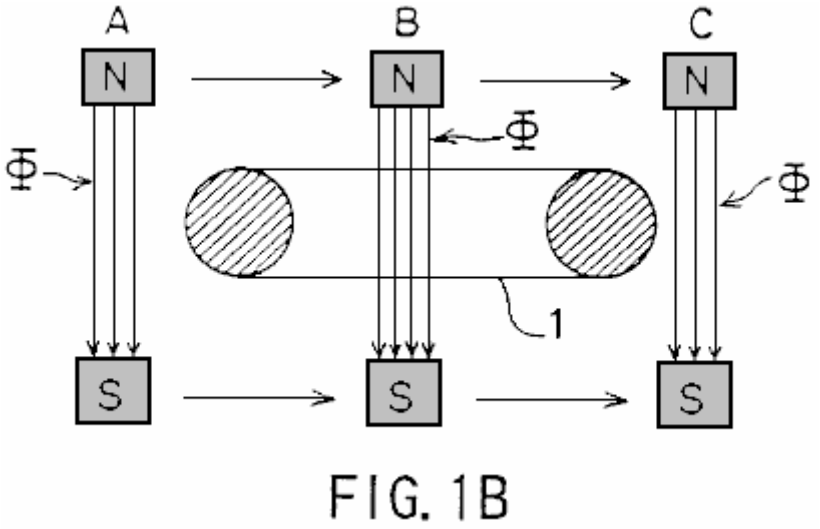


FIG. 1B

For example, as shown in **Fig.1A** and **Fig.1B**, assuming that the magnetic flux **phi** crossing a circular coil **1** at a perpendicular direction moves in the **A to B** direction as indicated by the arrow, a current $\mathbf{I_1}$ flows in accordance with Faraday's law of electromagnetic induction so that the pointer of a galvanometer **2** swings clockwise (+ direction) and then returns to the zero position. When the magnetic flux **phi** moves in the direction **B to C**, a current $\mathbf{I_2}$ flows so that the indicator of the galvanometer **2** swings counterclockwise (- direction) and then returns to the zero position.

Generally, an induction generator is constructed in such a way that an electromotive force is induced according to Flemming's right-hand rule by a conductor cutting magnetic flux lines (**Fig.1A**) or by the magnetic flux lines crossing the conductor (**Fig.1B**).

A rotor in an induction generator is usually constructed as a one-piece body having alternately disposed North poles and South poles. When there are two magnetic poles, the N-pole and the S-pole are opposite to each other. When there are more than two magnetic poles (for example, four magnetic poles or six magnetic poles etc.), the N-pole and the S-pole alternate, resulting in a N-S-N-S- . . . succession.

In this background, a unipolar induction generator is a special case wherein an electromotive force is generated by a conductor cutting the magnetic flux while moving or rotating, and a direct current is supplied through a slip

ring. In other words, a unipolar induction motor is unique in its construction characterised by a non-alternating magnetic field travelling in the same direction.

In the conventional induction generator such as the one described above, improvement in energy conversion efficiency is attained such that the rotor is constructed of a ferrite, or rare-earth, magnet characterised by a high energy product and a small reversing permeability (recoil permeability). Alternatively, the extent of demagnetisation due to generation of a counter magnetic field in an induction coil is reduced allowing the single polarity of the rotor to interact with the stator in forming a magnetic circuit. However, despite these measures, reduction in energy conversion efficiency due to a counter magnetic field of the rotor core, more specifically, due to demagnetisation resulting from the counter magnetic field caused by armature reaction presents a serious problem.

The present invention has been developed in view of the above points, and its object is to provide an induction generator having a pair of magnetic poles of the same polarity opposed to each other with respect to a rotation shaft, wherein a high energy conversion efficiency is attained.

DISCLOSURE OF THE INVENTION

The present invention provides an induction generator having a pair of magnetic poles of the same polarity opposed to each other with respect to a rotation shaft, characterised by comprising:

A rotation shaft driven by external means;

An even number of (more than three) stator cores provided to encircle the rotation shaft, predetermined gaps being provided between the adjacent stator cores;

A first single-opposed polarity rotor provided in the rotation shaft, surrounded by the even number of stator cores, and having first and second magnets magnetised such that the even number of stator cores remain facing a first polarity, the first and second magnets being opposed to each other with respect to the rotation shaft in a cross section;

A second single-opposed-polarity rotor provided in the rotation shaft so as to face the first single-opposed-polarity rotor at a predetermined distance along the rotation shaft, surrounded by the even number of stator cores, and having third and fourth magnets magnetised such that the even number of stator cores remain facing a second polarity which is opposite to the polarity of the first polarity, the third and fourth magnets being disposed opposite to each other with respect to the rotation shaft;

A plurality of windings provided in the even number of stator cores and connected according to a predetermined configuration, characterised in that:

A rotating magnetic field which causes electromagnetic induction in the even number of stator cores successively is created by the first, second, third and fourth magnets when the first and second single-opposed-polarity rotors are rotated; and

Periodic increase and decrease in the number of magnetic flux lines crossing a given winding and associated periodic decrease and increase crossing an adjacent winding causes a periodic electromotive force having a rectangular waveform to be output.

In one aspect of the present invention, the plurality of windings connected according to the predetermined configuration form first and second serial circuits:

The first serial circuit outputs a periodic first electromotive force having a rectangular waveform when a rotating magnetic field which causes electromagnetic induction in the even number of stator cores successively is created by the first, second, third and fourth magnets when the first and second single-opposed-polarity rotors are rotated; and

The second serial circuit outputs a periodic second electromotive force of a rectangular waveform 180^0. out of phase with the first electromotive force and having the same period as the first electromotive force, when a rotating magnetic field which causes electromagnetic induction in the even number of stator cores successively is created by the first and second single-opposed-polarity rotors are rotated.

The induction generator of the present invention may also comprise:

Rotation position detecting means for detecting a position of the first and second single-opposed-polarity rotors during their rotation; and

Switching means which alternately causes positive components of the first electromotive force having a rectangular waveform and provided by the first serial circuit, or positive components of the second electromotive force having a rectangular waveform and provided by the second serial circuit to be output at intervals of an electrical angle of 180^0.

In another aspect of the present invention, the plurality of windings comprise a first winding provided in a first stator core of the even number of stator cores, a second winding provided in a second stator core adjacent to the first stator core so as to wind in a direction opposite to a direction in which the first winding is provided, a third winding provided in a third stator core adjacent to the second stator core so as to wind in the same direction as the first winding, a fourth winding provided in a fourth stator core adjacent to the third stator core so as to wind in a direction opposite to a direction in which the third winding is provided, the first through fourth windings being connected with each other according to a predetermined configuration.

In still another aspect of the present invention, the first serial circuit comprises a first winding provided to wind in a first direction in a first stator core of the even number of stator cores, a second winding serially connected to the first winding and provided in a second stator core adjacent to the first stator core so as to wind in a second direction opposite to the first direction, a third winding serially connected with the second winding and provided in a third stator core adjacent to the second stator core so as to wind in the first direction, a fourth winding serially connected to the third winding and provided in a fourth stator core adjacent to the third stator core so as to wind in the second direction; and

The second serial circuit comprises a fifth winding provided to wind in the second direction in the first stator core, a sixth winding serially connected to the fifth winding and provided in the second stator core so as to wind in the first direction, a seventh winding serially connected with the sixth winding and provided in the third stator core so as to wind in the second direction, an eighth winding serially connected to the seventh winding and provided in the fourth stator core so as to wind in the first direction.

In yet another aspect of the present invention, the first through fourth magnets are arc-shaped; and

the even number of stator cores have arc-shaped cross sections.

In still another aspect of the present invention, the arc-shaped first through fourth magnets and the stator cores which have arc-like cross sections have an almost identical circumferential length.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig.1A and **Fig.1B** are diagrams explaining the principle of an induction generator;

Fig.2A and **Fig.2B** are diagrams showing a first embodiment of the present invention;

Fig.3A and **Fig.3B** are diagrams showing a single-opposed-polarity rotor **11N** according to the first embodiment of the present invention;

Fig.4A and **Fig.4B** are diagrams showing a single-opposed-polarity rotor **11S** according to the first embodiment of the present invention;

Fig.5A, **Fig.5B** and **Fig.5C** are diagrams showing how wirings are connected with each other according to the first embodiment of the present invention;

Fig.6A is a diagram schematically showing how a rotating magnetic field according to the first embodiment crosses windings **7c-10c**;

Fig.6B shows a magnetic path;

Fig.7 is a diagram showing a waveform of an output voltage according to the first embodiment;

Fig.8A and **Fig.8B** are diagrams showing a second embodiment of the present invention;

Fig.9 is a diagram showing how wirings are connected with each other according to a second embodiment; and

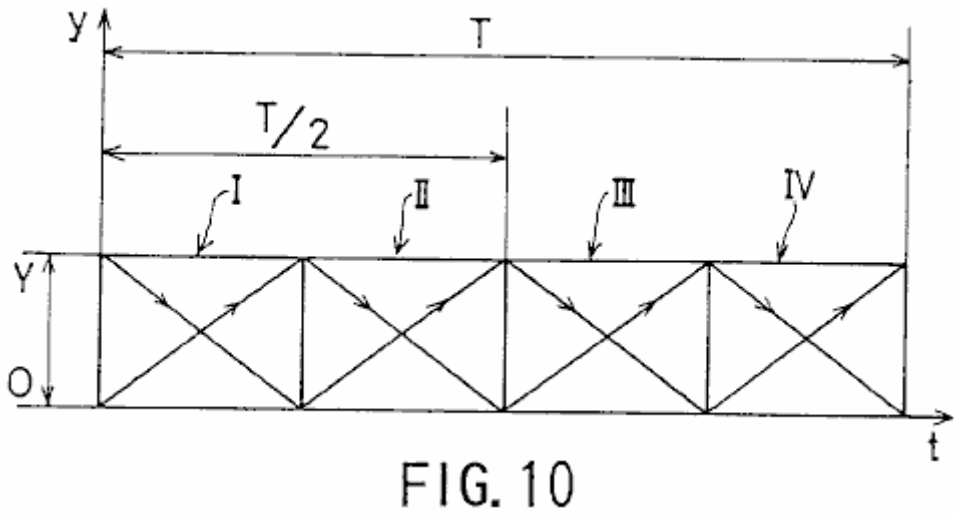


FIG. 10

Fig.10 is a diagram showing a waveform of an output voltage according to the second embodiment.

BEST MODE FOR CARRYING OUT THE INVENTION

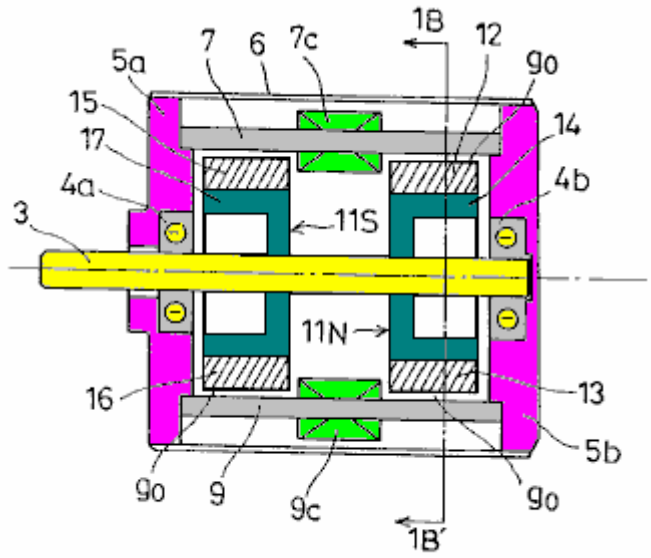


FIG. 2A

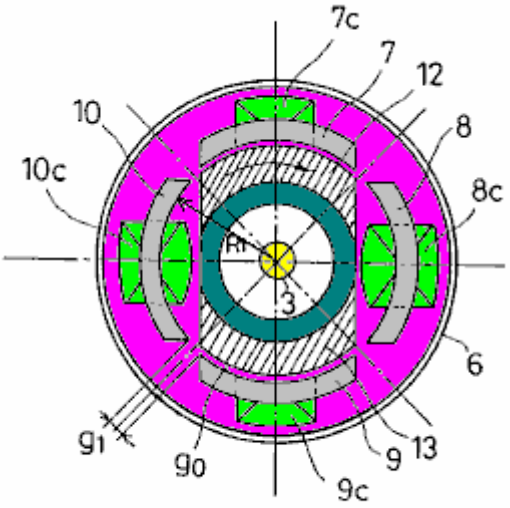


FIG. 2B

Fig.2A and **Fig.2B** show a first embodiment of the present invention. Specifically, **Fig.2A** is a longitudinal sectional view and **Fig.2B** is a cross-sectional view taken in the line **1B-1B'** of **Fig.2A**.

Referring to **Fig.2A** and **Fig.2B**, **3** indicates a rotation shaft formed of a non-magnetic material and driven by an external means; **4a** and **4b** bearings for supporting the rotation shaft **3**; **5a** and **5b** are flanges provided with the bearings **4a** and **4b**, respectively; and **6** is a cylindrical case cover for accommodating the flanges **5a** and **5b**.

Stator cores **7**, **8**, **9** and **10** are arranged so as to encircle the rotation shaft **3**, equidistant gaps g_1 being provided between the adjacent stator cores. Each of the stator cores **7**, **8**, **9** and **10** has the same arc-like cross section.

A single-opposed-polarity N-pole rotor **11N** and a single-opposed-polarity S-pole rotor **11S** are provided on the rotation shaft **3** so as to be opposite to each other. The single-opposed-polarity rotors **11N** and **11S** are surrounded by the stator cores **7**, **8**, **9** and **10**, a small rotation gap g_0 being provided between the single-opposed-polarity rotor and the stator core.

Referring to **Fig.2B**, windings **7c** and **9c** are provided clockwise around the stator cores **7** and **9**, respectively. Windings **8c** and **10c** are wound counterclockwise around the stator cores **8** and **10**, respectively. The windings **7c**, **8c**, **9c** and **10c** are connected with each other in a configuration described later.

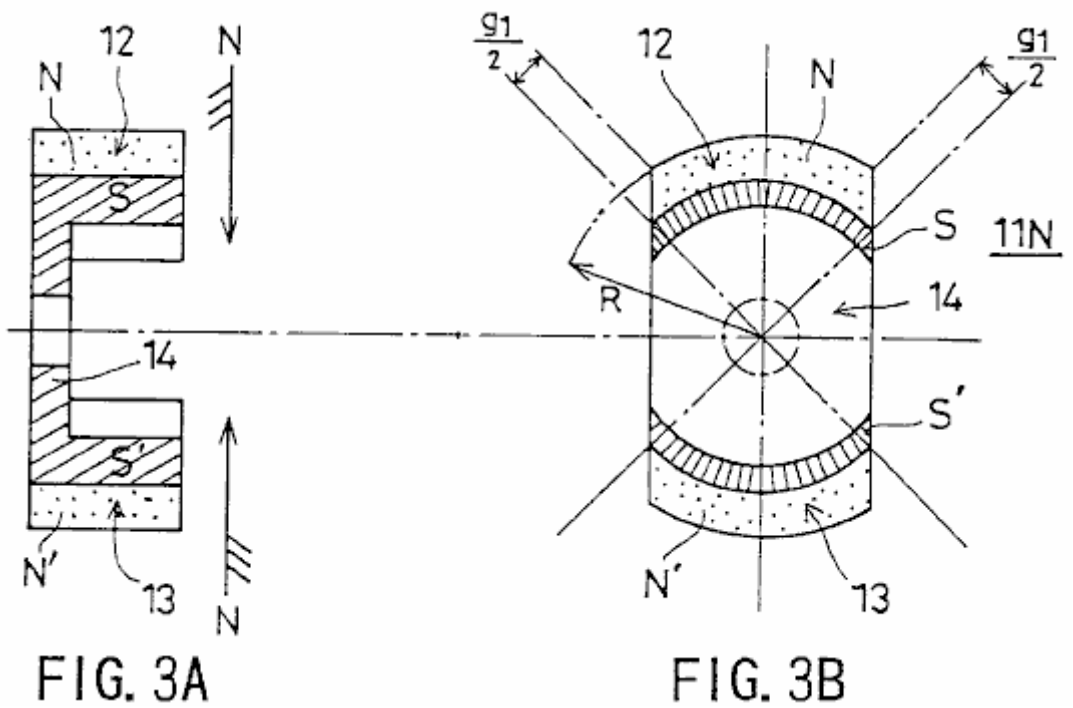


FIG. 3A FIG. 3B

Fig.3A and **Fig.3B** show the single-opposed-polarity rotor **11N**. Specifically, **Fig.3A** is a longitudinal sectional view, and **Fig.3B** is a cross-sectional view. The single-opposed-polarity rotor **11N** has arc-shaped magnets **12** and **13** which are 180^0 displaced from each other and are magnetised such that their surfaces which face the stator cores **7-10** are N-poles while their inner surfaces are S-poles. The arc-shaped magnets **12** and **13** are configured to match the outline of the stator cores **7**, **8**, **9** and **10**. Referring to **Fig.3B**, the symbols **N** and **N'** are used so as to differentiate between the magnets **12** and **13**.

A rotor piece **14** is positioned so as to connect the arc-shaped magnets **12** and **13**. The rotor piece **14** is magnetised by the arc-shaped magnets **12** and **13** so that it's surfaces which face the arc-shaped magnets **12** and **13** are S-poles and is formed of a substance (for example, a silicon steel) constructed of a low carbon steel having mixed therein several percent of non-ferrous metal subjected to a forging-cast process. The iron core embodied by the rotor piece **14** thus constructed is characterised by a well-balanced magnetic field where the permeability approximates a peak value in a unipolar magnetic field that the iron core presents to its surroundings.

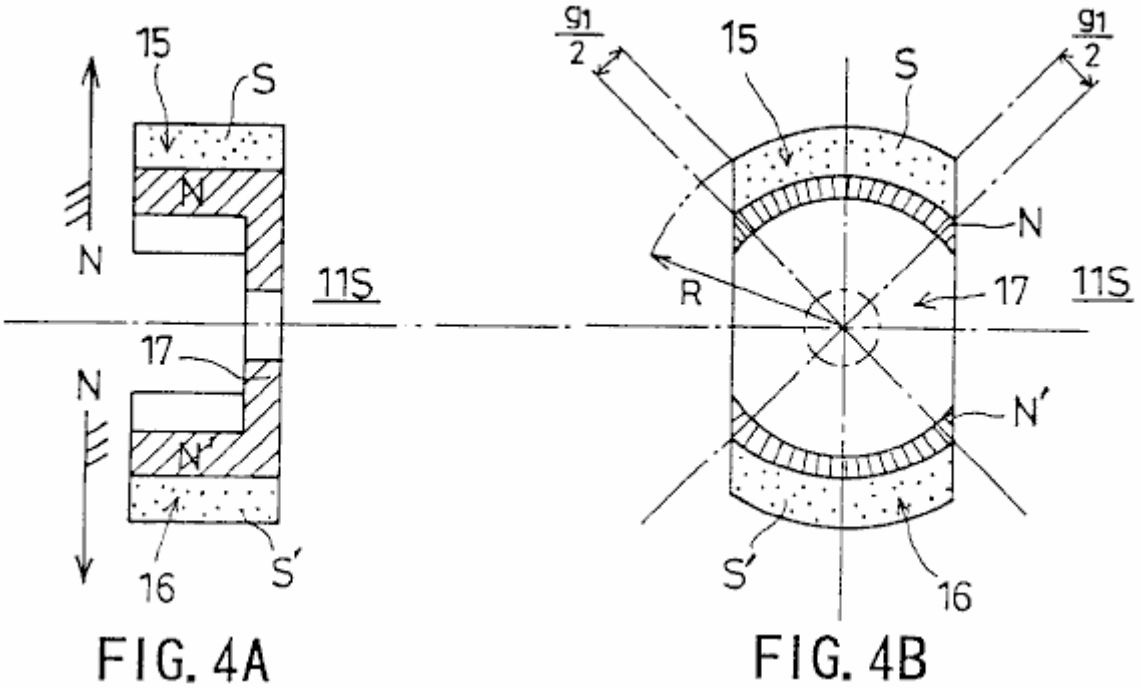


FIG. 4A FIG. 4B

Fig.4A and **Fig.4B** show the single-opposed-polarity rotor **11S**. Specifically, **Fig.4A** is a longitudinal sectional view, and **Fig.4B** is a cross-sectional view.

The single-opposed-polarity rotor **11S** has arc-shaped magnets **15** and **16** which are 180^0 displaced from each other and are magnetised such that the surfaces thereof facing the stator cores **7-10** are S-poles while their inner surfaces are N-poles. The arc-shaped magnets **15** and **16** are configured to match the outline of the stator cores **7**, **8**, **9** and **10**.

A rotor piece 17 is positioned so as to connect the arc-shaped magnets **15** and **16**. The rotor piece **17** is magnetised by the arc-shaped magnets **15** and **16** so that it's surfaces which face the arc-shaped magnets **15** and **16** are N-poles. The rotor piece is made from a substance constructed from a low carbon steel having mixed in it, several percent of non-ferrous metal subjected to a forging-cast process. The iron core embodied by the rotor piece **17** thus constructed is characterised by a well-balanced magnetic field where the permeability approximates a peak value in a unipolar magnetic field that the iron core presents to its surroundings.

The arc-shaped magnets **12**, **13**, **15** and **16** have the same circumferential length, which is also equal to the length of the arc formed by the circumference of the stator cores **7**, **8**, **9** and **10**. More specifically, this length is obtained by dividing the entire hypothetical circumference minus the four $\mathbf{g_1}$ gaps by four. Referring to **Fig.2A** and **Fig.2B**, the rotation gap $\mathbf{g_0}$ is equal to $\mathbf{R_1}$ - **R**, where $\mathbf{R_1}$ is a distance between the centre of the rotation shaft **3** and the inner surface of the stator cores **7-10**, and **R** is a distance between the centre of the rotation shaft **3** and the outer surface of the single-opposed-polarity rotors **11N** and **11S**, as indicated in **Fig.3B** and **Fig.4B**.

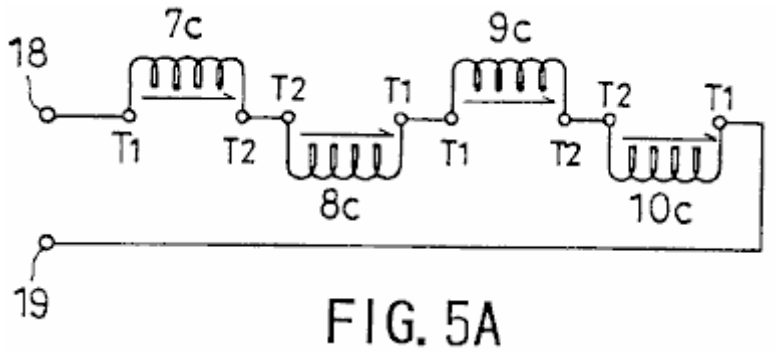


FIG. 5A

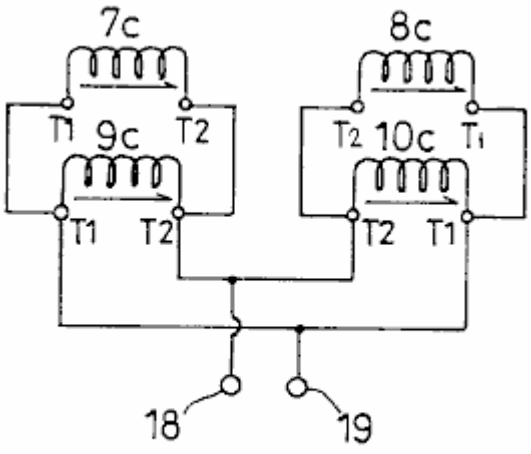


FIG. 5B

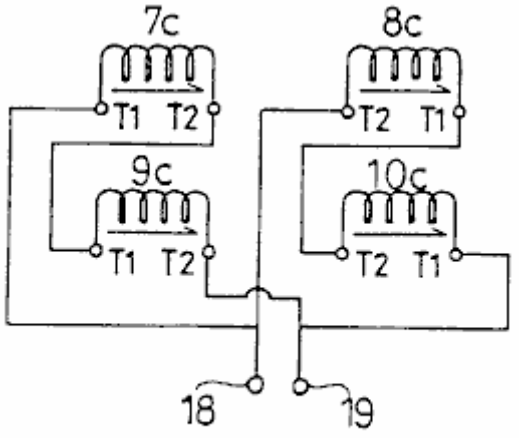


FIG. 5C

Fig.5A, **Fig.5B** and **Fig.5C**, show how the wirings are connected with each other. **T_1** indicates the beginning of a winding, **T_2** the end of a winding, and **18** and **19** output terminals. More specifically, **Fig.5A** shows a serial connection configuration, **Fig.5B** a serial-parallel connection configuration, and **Fig.5C** a parallel connection configuration. The serial connection configuration allows the electromotive force induced in the windings to be added together and provides a high-voltage output. The parallel connection configuration allows currents resulting from the electromotive force induced in the windings to be added together and provides a large-current output.

A description will now be given, with reference to **Fig.6A**, **Fig.6B** and **Fig.7**, of power generation operation of the serial connection configuration.

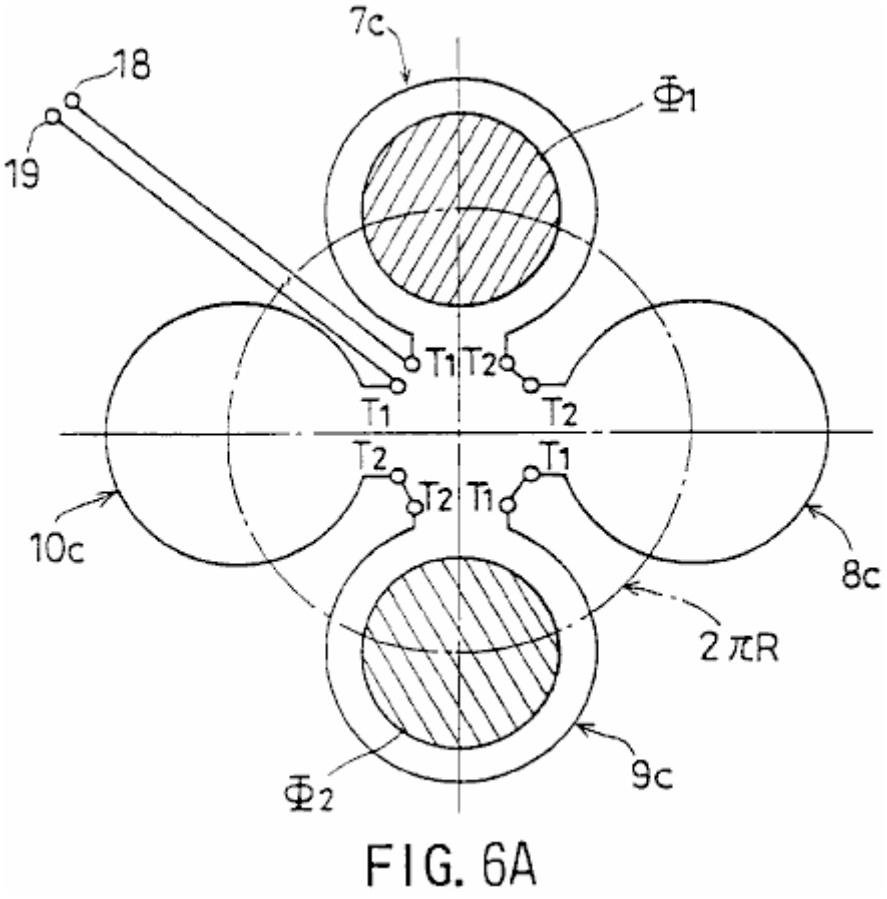


FIG. 6A

Fig.6A is a diagram showing schematically how the rotating magnetic field provided by the single-opposed-polarity rotors **11S** and **11N** crosses windings **7c-10c**. **Fig.6B** shows a magnetic path.

Referring to **Fig.6A**, **Phi_1** and **Phi_2** indicate rotating magnetic flux rotating along the circumference 2Pi x R. **Fig.6B** shows the arc-shaped magnets **12** and **15** directly opposite the stator core **7** over their entire length, and the arc-shaped magnets **13** and **16** directly opposite the stator core **9** over their entire length.

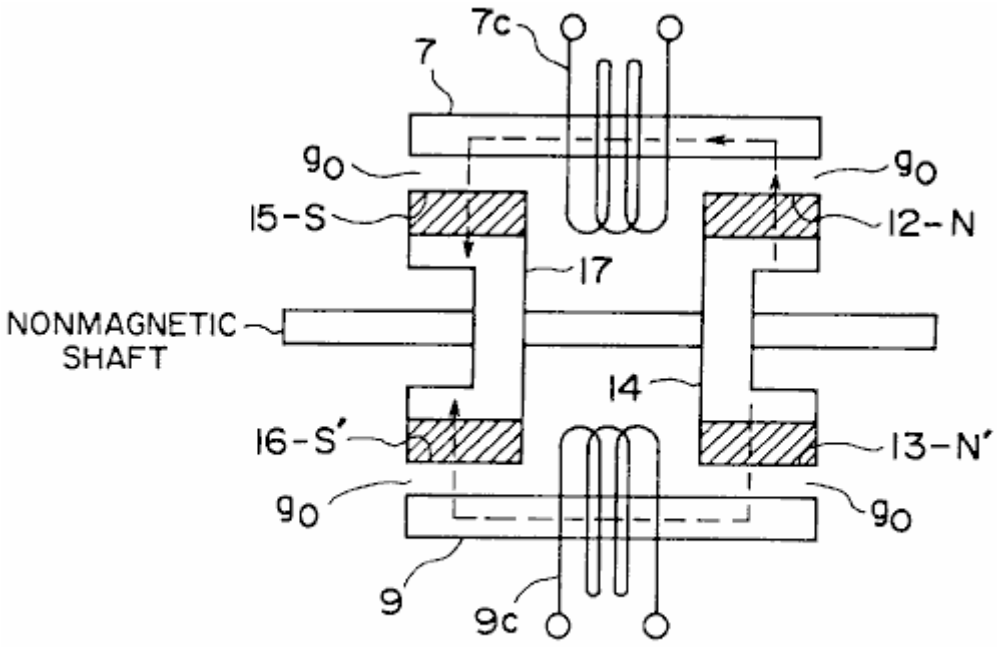


FIG. 6B

As shown in **Fig.6B**, the magnetic flux **Phi_1** forms a magnetic path as follows:

The rotor piece **14** (S) - the arc-shaped magnet pole **12** (N) - stator core **7** - the rotation gap **g_0** - the arc-shaped magnet **15** (S) - the rotor piece **17** (N).

The magnetic flux **Phi_2** forms a magnetic path as follows:

The rotor piece **14** (S) - the arc-shaped magnet **13** (N) - the rotation gap **g_0** - the stator core **9** - the rotation gap **g_0** - the arc-shaped magnet **16** (S) - the rotor piece **17** (N).

Thus, a parallel magnetic path is formed. In this state, the magnetic flux **Phi_1** crosses the winding **7c**, and the magnetic flux **Phi_2** crosses the winding **9c**.

A description focused on the rotation of the magnetic flux **Phi_1** is given. Specifically, a description will be given of a change in the way the magnetic flux **Phi_1** crosses the windings.

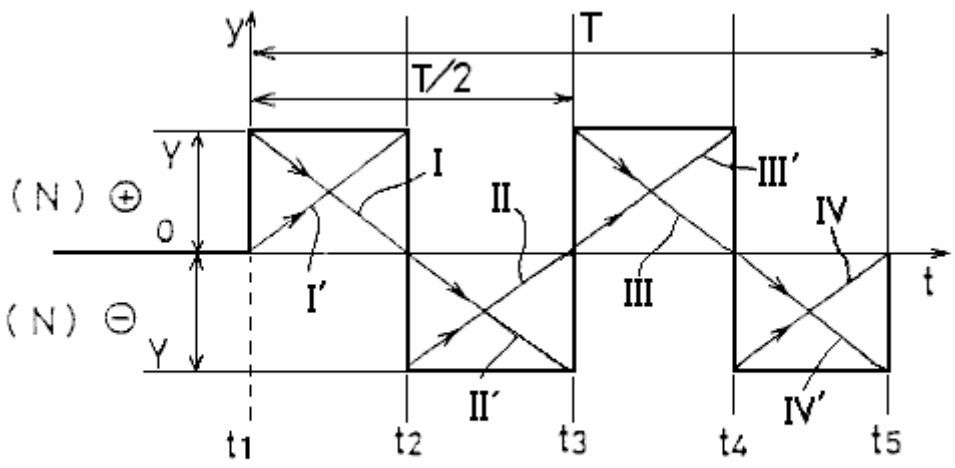


FIG. 7

Referring to a waveform of an output voltage shown in **Fig.7**, the entirety of the magnetic flux **Phi_1** crosses the winding **10c** at a time **t_1**. At a time **t_2**, the entirety of the magnetic flux **Phi_1** crosses the winding **7c**. At a time **t_3**, the entirety of the magnetic flux **Phi_1** crosses the winding **8c**. At a time **t_4**, the entirety of the magnetic flux **Phi_1** crosses the winding **9c**. At a time **t_5**, the entirety of the magnetic flux **Phi_1** crosses the winding **10c**. In this way, the magnetic flux **Phi_1** rotates at a constant speed during a time **T**, in a clockwise direction in **Fig.6A**.

Between the time **t_1** and the time **t_2**, an electromotive force having a descending triangular waveform, indicated by **I** in **Fig.7**, is generated in the winding **10c** due to a decrease in the number of magnetic flux lines of the magnetic flux **Phi** crossing the winding **10c**. An electromotive force having an ascending triangular waveform, indicated by **I'** in **Fig.6**, is generated in the winding **7c** due to an increase in the number of magnetic flux lines of the magnetic flux **Phi** crossing the winding **7c**. Accordingly, a positive rectangular waveform obtained by the sum of these triangular waveforms is output to the output terminals **18** and **19**.

Between the time **t_2** and the time **t_3**, an electromotive force having an ascending triangular waveform, indicated by **II** in **Fig.7**, is generated in the winding **7c** due to a decrease in the number of magnetic flux lines of the magnetic flux **Phi** crossing the winding **7c**. An electromotive force having a descending triangular waveform, indicated by **II'** in **Fig.7**, is generated in the winding **8c** due to an increase in the number of magnetic flux lines of the magnetic flux **Phi** crossing the winding **8c**. Accordingly, a negative rectangular waveform obtained by the sum of these triangular waveforms is output to the output terminals **18** and **19**.

Between the time **t_3** and the time **t_4**, an electromotive force having a descending triangular waveform, indicated by **III** in **Fig.7**, is generated in the winding **8c** due to a decrease in the number of magnetic flux lines of the magnetic flux **Phi** of the magnetic flux **Phi** crossing the winding **8c**. An electromotive force having an ascending triangular waveform, indicated by **III'** in **Fig.7**, is generated in the winding **9c** due to an increase in the number of magnetic flux lines of the magnetic flux **Phi** crossing the winding **9c**. Accordingly, a positive rectangular waveform obtained by the sum of these triangular waveforms is output to the output terminals **18** and **19**.

Between the time **t_4** and the time **t_5**, an electromotive force having an ascending triangular waveform, indicated by **IV** in **Fig.7**, is generated in the winding **9c** due to a decrease in the number of magnetic flux lines of the magnetic flux **Phi** crossing the winding **9c**. An electromotive force having a descending triangular waveform, indicated by **IV'** in **Fig.7**, is generated in the winding **10c** due to an increase in the number of magnetic flux lines of the magnetic flux **Phi** crossing the winding **10c**. Accordingly, a negative rectangular waveform obtained by the sum of these triangular waveforms is output to the output terminals **18** and **19**.

While the magnetic flux **Phi_1** makes one rotation, an electromotive force having a synthesised rectangular waveform and a period of **T/2** is output, as shown in **Fig.7**. Since the magnetic flux **Phi_2** also makes one rotation while the magnetic flux **Phi_1** makes one rotation and produces an output of an electromotive force having a similar rectangular waveform, the magnitude of the electromotive force obtained between the terminals **18** and **19** is actually double that indicated in **Fig.7**.

In this way, this embodiment makes it possible to cancel a counter magnetic field and provide an induction generator having a pair of magnetic poles of the same polarity opposed to each other with respect to a rotation shaft and characterised by a high energy conversion efficiency. Our operating practice has confirmed that the generator having the construction of this embodiment provides an energy conversion efficiency which is high enough to require only 1/5.2 of the driving torque for the conventional generator.

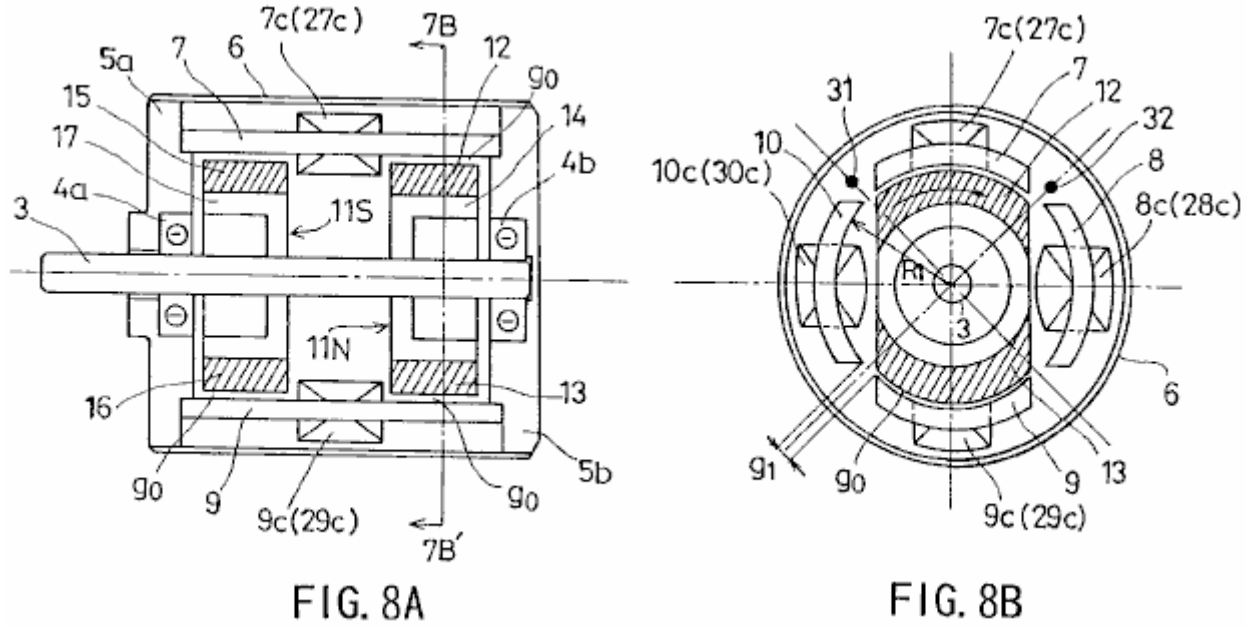


FIG. 8A FIG. 8B

Fig.8A and **Fig.8B** show a second embodiment of the present invention. Specifically, **Fig.8A** is a longitudinal sectional view, and **Fig.8B** is a cross-sectional view taken in the line **7B-7B'** of **Fig.8A**.

Referring to **Fig.8A** and **Fig.8B**, **3** indicates a rotation shaft formed of a non-magnetic material and driven by an external source; **4a** and **4b** are bearings which support the rotation shaft **3**, **5a** and **5b** are flanges housing the bearings **4a** and **4b**, and **6** is a cylindrical case cover for accommodating the flanges **5a** and **5b**.

Stator cores **7**, **8**, **9** and **10** are arranged so as to encircle the rotation shaft **3**, equidistant gaps $\mathbf{g_1}$ being provided between the adjacent stator cores. Each of the stator cores **7**, **8**, **9** and **10** has a same arc-like cross section.

A single-opposed-polarity N-pole rotor **11N** and a single-opposed-polarity S-pole rotor **11S** are provided on the rotation shaft **3** so as to be opposite to each other. The single-opposed-polarity rotors **11N** and **11S** are surrounded by the stator cores **7**, **8**, **9** and **10** a small rotation gap $\mathbf{g_0}$ being provided between the single-opposed-polarity rotor and the stator core.

Referring to **Fig.8B**, windings **7c** and **9c** are provided clockwise around the stator cores **7** and **9**, respectively. Windings **27c** and **29c** are provided counterclockwise around the stator cores **7** and **9**, respectively. Windings **8c** and **10c** are provided counterclockwise in the stator cores **8** and **10**, respectively. Windings **28c** and 30c are wound clockwise around the stator cores 8 and 10, respectively. The windings 7c, 8c, 9c, 10c, **27c**, **28c**, **29c** and **30c** are connected with each other according to a configuration described later.

A magnetic sensor (for rotation position detection) **31** is provided between the stator cores **7** and **10**, and a magnetic sensor (for rotation position detection) **32** is provided between the stator cores **7** and **8**. The magnetic sensors **31** and **32** detect the magnetic field so as to determine the position of the single-opposed-polarity rotors **11N** and **11S** during their rotation.

The single-opposed-polarity rotors **11N** has a configuration as shown in **Fig.3A** and **Fig.3B**, and the monopole rotor **11S** has a configuration as shown in **Fig.4A** and **Fig.4B**.

The single-opposed-polarity rotor **11N** has arc-shaped magnets **12** and **13** which are 180^0 displaced from each other and are magnetised such that their surfaces facing the stator cores are N-poles while their respective inner surfaces are S-poles. The arc-shaped magnets **12** and **13** are configured to match the outline of the stator cores **7**, **8**, **9** and **10**.

A rotor piece **14** is positioned so as to connect the arc-shaped magnets **12** and **13**. The rotor piece **14** is constructed from a low-carbon steel having several percent of non-ferrous metal, using a forging-cast process. The iron core rotor piece **14** constructed by this means, has a well-balanced magnetic field where the permeability approximates a peak value in a unipolar magnetic field that the iron core presents to its surroundings.

The single-opposed-polarity rotor **11S** has arc-shaped magnets **15** and **16** which are positioned 180^0 apart from each other and are magnetised so that their surfaces which face the stator cores are S-poles while their inner

surfaces are N-poles. The arc-shaped magnets **15** and **16** are shaped and positioned so as to match the outline of the stator cores **7**, **8**, **9** and **10**.

A rotor piece **17** is positioned so as to connect the arc-shaped magnets **15** and **16**. The rotor piece **17** is constructed from a low-carbon steel having several percent of non-ferrous metal, using a forging-cast process. The iron core rotor piece **17** constructed by this means, has a well-balanced magnetic field where the permeability approximates a peak value in a unipolar magnetic field which the iron core presents to its surroundings.

The arc-shaped magnets **12**, **13**, **15** and **16** have the same circumferential lengths, which is equal to the length of the arc formed by the circumference of the stator cores **7**, **8**, **9** and **10**. More specifically, this length is obtained by dividing by four, the entire hypothetical circumference minus the four gaps $\mathbf{g_1}$. Referring to **Figs. 3A, 3B, 4A, 4B** and **8**, the rotation gap $\mathbf{g_0}$ is equal to $\mathbf{R_1}$ -**R**.

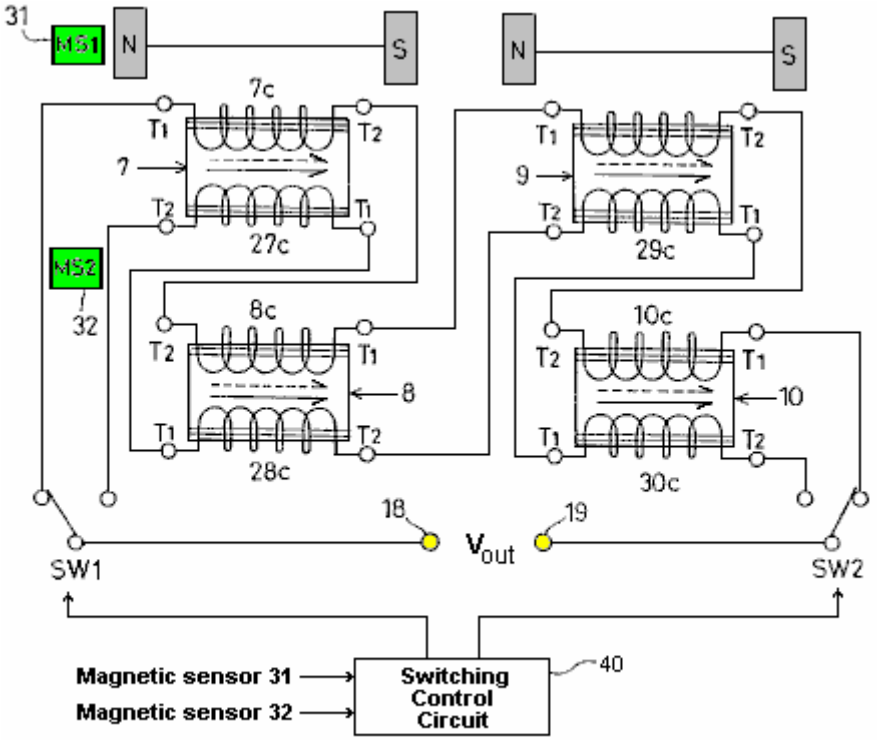


FIG. 9

Fig.9 shows how the wirings are connected with each other. $\mathbf{T_1}$ indicates the beginning of a winding, $\mathbf{T_2}$ the end of a winding, and **18** and **19** are the output terminals.

Two serial circuits are formed out of the windings. Switches **SW1** and **SW2** are used for selection of the respective serial circuits. A switching control circuit **40**, which processes a detection signal from the magnetic sensors **31** and **32**, drives the switches SW1 and SW2 selectively in accordance with the detection signal.

As shown in **Fig.9**, the first serial circuit comprises the winding **7c** provided clockwise in the stator core **7**, the winding **8c** serially connected with the winding **7c** and provided counterclockwise in the stator core **8** adjacent to the stator core **7**; the winding **9c** serially connected with the winding **8c** and provided clockwise in the stator core **9**; and the winding **10c** serially connected with the winding **9c** and provided counterclockwise in the stator core **10** adjacent to the stator core **9**.

As shown in **Fig.9**, the second serial circuit comprises the winding **27c** provided counterclockwise in the stator core **7**; the winding **28c** serially connected with the winding **27c** and provided clockwise in the stator core **8**; the winding **29c** serially connected with the winding **28c** and provided counterclockwise in the stator core **9**; and the winding **30c** serially connected with the winding **29c** and provided clockwise in the stator core **10**.

According to the construction described above, a rotating magnetic field which causes electromagnetic induction in the stator cores **7-10** successively is created by the arc-shaped magnets **12**, **13**, **15** and **16** when the single-opposed-polarity rotors **11N** and **11S** are rotated. As has been already explained with reference to **Fig.6A**, **Fig.6B** and **Fig.7**, as the magnetic flux lines crossing one of the windings **7c-10c** increase in number, the magnetic flux lines crossing the adjacent one of the windings **7c-10c** decrease in number. That is, the magnetic

flux lines periodically increase and decrease with respect to a given winding so that a first electromotive force, having a rectangular waveform similar to the one shown in **Fig.7** and a period that is 1/2 the period of the rotation, is output from the first serial circuit (**7c-10c**).

As the magnetic flux lines crossing one of the windings **27c-30c** increase in number, the magnetic flux lines crossing the adjacent one of the windings **27c-30c** decrease in number. That is, the magnetic flux lines periodically increase and decrease with respect to a given winding so that a second electromotive force of a rectangular waveform 180^0 out of phase with the first electromotive force and having the same period as the first electromotive force is output from the second serial circuit (**27c-30c**). That is, the second electromotive force is 180^0 out of phase with the electromotive force shown in **Fig.7**.

Referring to **Fig.10**, in accordance with the detection signal from the magnetic sensors **31** and **32**, the switches **SW1** and **SW2** effect switching at 90^0 intervals. By that means, the positive components **I** and **III** of the first electromotive force having a rectangular waveform and provided from the first serial circuit, and the positive components **II** and **IV** of the second electromotive force having a rectangular waveform and provided from the second serial circuit are alternately selected at 180^0 intervals and output to the output terminals **18** and **19**.

This means that, this embodiment ensures a high-efficiency energy conversion wherein a counter magnetic field is cancelled, and a DC electromotive force having a positive level is properly synthesised and output. It is of course possible to synthesise and output a negative DC electromotive force by shifting the switching timing by 180^0.

INDUSTRIAL APPLICABILITY

As has been described, according to the present invention, the rotation of the first and second single-opposed-polarity rotors generates a rotating magnetic field which causes an induction in an even number of stator cores successively. As the magnetic flux lines crossing one of the first-through-fourth windings increase in number, the magnetic flux lines crossing the adjacent one of the first-through-fourth windings decrease in number. That is, the magnetic flux lines periodically increase and decrease with respect to a given winding. The electromotive force generated as the magnetic flux lines crossing a winding increase in number and the electromotive force generated as the magnetic flux lines crossing an adjacent winding decrease in number are synthesised so that a periodic AC electromotive force having a rectangular waveform is generated out of the synthesis and output. In this way, a high-efficiency energy conversion wherein a counter magnetic field is cancelled is provided.

According to the first serial circuit of the present invention, the rotation of the first and second single-opposed-polarity rotors generates a rotating magnetic field which causes an induction in an even number of stator cores successively. As the magnetic flux lines crossing one of the first through fourth windings increase in number, the magnetic flux lines crossing the adjacent one of the first through fourth windings decrease in number. That is, the magnetic flux lines periodically increase and decrease in a given winding. Accordingly, the first electromotive force having a rectangular waveform is output. According to the second serial circuit, as the magnetic flux lines crossing one of the fifth-through-eighth windings increase in number, the magnetic flux lines crossing the adjacent one of the fifth-through-eighth windings decrease in number. That is, the magnetic flux lines periodically increase and decrease in a given winding. Accordingly, the second electromotive force 180^0 out of phase with the first electromotive force and having the same period as the first electromotive force is output. In accordance with the detection signal from the rotation position detecting means, the switching means selectively causes the positive components of the first electromotive force provided by the first serial circuit, or the positive components of the second electromotive force provided by the second serial circuit to be output at 180^0 intervals. In this way the DC electromotive force is synthesised and output. This results in a high-efficiency energy conversion where a counter magnetic field is cancelled.

In addition to extensive applications in power plants, ships, aircraft etc., the present invention may find household applications or may be conveniently adapted for leisure uses.

CLAIMS

1. An induction generator having a pair of magnetic poles of the same polarity opposed to each other with respect to a rotation shaft, characterised by comprising:

a rotation shaft driven by external means;

an even number of (more than three) stator cores provided to encircle said rotation shaft, predetermined gaps being provided between the adjacent stator cores;

a first single-opposed-polarity rotor provided on said rotation shaft, surrounded by said even number of stator cores, and having first and second magnets magnetised such that said even number of stator cores remain facing a first polarity, said first and second magnets sandwiching a magnetic body between them and being opposed to each other with respect to said rotation shaft in a cross section;

a second single-opposed-polarity rotor provided on said rotation shaft so as to face said first single-opposed-polarity rotor at a predetermined distance along the rotation shaft, surrounded by said even number of stator cores, and having third and fourth magnets magnetised such that said even number of stator cores remain facing a second polarity which is opposite to the polarity of said first polarity, said third and fourth magnets sandwiching a magnetic body between them and being disposed opposite to each other with respect to said rotation shaft;

a plurality of windings provided in said even number of stator cores and connected according to a predetermined configuration, characterised in that: a rotating magnetic field which causes electromagnetic induction in said even number of stator cores successively is created by the first, second, third and fourth magnets when said first and second single-opposed-polarity rotors are rotated; and

two windings adjacent to each other are wound in opposite directions and connected in series so that a rectangular waveform is formed by synthesising the electromotive forces generated by the two windings, so that an electromotive force having a triangular waveform caused by periodic increase and decrease in the number of magnetic flux lines crossing one of the two windings and another electromotive force having a triangular waveform caused by associated periodic decrease and increase in the number of magnetic flux lines crossing the other one of the windings are synthesised so as to generate a periodic voltage having a rectangular waveform.

2. The induction generator having a pair of magnets of the same polarity opposed to each other with respect to a rotation shaft as claimed in claim 1, characterised in that:

said plurality of windings connected according to the predetermined configuration form first and second serial circuits;

said first serial circuit outputs a periodic first electromotive force having a rectangular waveform when a rotating magnetic field which causes electromagnetic induction in said even number of stator cores successively is created by said first, second, third and fourth magnets when said first and second single-opposed-polarity rotors are rotated; and

said second serial circuit outputs a periodic second electromotive force of a rectangular waveform 180^0 out of phase with the first electromotive force and having the same period as the first electromotive force, when a rotating magnetic field which causes electromagnetic induction in said even number of stator cores successively is created by said first and second single-opposed-polarity rotors are rotated.

3. The induction generator having a pair of magnets of the same polarity opposed to each other with respect to a rotation shaft as claimed in claim 2, further comprising:

rotation position detecting means for detecting a position of said first and second single-opposed-polarity rotors during their rotation; and

switching means which alternately causes positive components of said first electromotive force having a rectangular waveform and provided by said first serial circuit, or positive components of said second electromotive force having a rectangular waveform and provided by said second serial circuit to be output at intervals of an electrical angle of 180^0 to thereby produce a DC output.

4. The induction generator having a pair of magnets of the same polarity opposed to each other with respect to a rotation shaft as claimed in claim 1, characterised in that:

said plurality of windings comprise a first winding provided in a first stator core of said even number of stator cores, a second winding provided in a second stator core adjacent to the first stator core so as to wind in a direction opposite to a direction in which the first winding is provided, a third winding provided in a third stator core adjacent to the second stator core so as to wind in the same direction as the first winding, a fourth winding provided in a fourth stator core adjacent to the third stator core so as to wind in a direction opposite to a direction in which the third winding is provided, the first through fourth windings being connected with each other according to a predetermined configuration.

5. The induction generator having a pair of magnets of the same polarity opposed to each other with respect to a rotation shaft as claimed in claim 2, characterised in that:

said first serial circuit comprise a first winding provided to wind in a first direction in a first stator core of said even number of stator cores, a second winding serially connected to said first winding and provided in a second stator core adjacent to the first stator core so as to wind in a second direction opposite to the first direction, a third winding serially connected with said second winding and provided in a third stator core adjacent to the second stator core so as to wind in the first direction, a fourth winding serially connected to said third winding and provided in a fourth stator core adjacent to the third stator core so as to wind in the second direction; and

said second serial circuit comprises a fifth winding provided to wind in the second direction in said first stator core, a sixth winding serially connected to said fifth winding and provided in said second stator core so as to wind in said first direction, a seventh winding serially connected with said sixth winding and provided in said third stator core so as to wind in said second direction, an eighth winding serially connected to said seventh winding and provided in said fourth stator core so as to wind in said first direction.

6. The induction generator having a pair of magnets of the same polarity opposed to each other with respect to a rotation shaft as claimed in claim 1, characterised in that:

said first through fourth magnets are arc-shaped; and

said even number of stator cores have arc-shaped cross sections.

7. The induction generator having a pair of magnets of the same polarity opposed to each other with respect to a rotation shaft as claimed in claim 6, characterised in that said arc-shaped first through fourth magnets and said stator cores which have arc-shaped cross sections have an almost identical circumferential length.

8. The induction generator having a pair of magnets of the same polarity opposed to each other with respect to a rotation shaft as claimed in claim 2, characterised in that:

said first through fourth magnets are arc-shaped; and

said even number of stator cores have arc-shaped cross sections.

9. The induction generator having a pair of magnets of the same polarity opposed to each other with respect to a rotation shaft as claimed in claim 3, characterised in that:

said first through fourth magnets are arc-shaped; and

said even number of stator cores have arc-shaped cross sections.

10. The induction generator having a pair of magnets of the same polarity opposed to each other with respect to a rotation shaft as claimed in claim 4, characterised in that:

said first through fourth magnets are arc-shaped; and

said even number of stator cores have arc-shaped cross sections.

11. The induction generator having a pair of magnetic poles of the same polarity opposed to each other with respect to a rotation shaft as claimed in claim 5, characterised in that:

said first through fourth magnets are arc-shaped; and

said even number of stator cores have arc-shaped cross sections.

12. The induction generator having a pair of magnetic poles of the same polarity opposed to each other with respect to a rotation shaft as claimed in claim 8, characterised in that said arc-shaped first through fourth magnets and said stator cores which have arc-shaped cross sections have an almost identical circumferential length.

13. The induction generator having a pair of magnetic poles of the same polarity opposed to each other with respect to a rotation shaft as claimed in claim 9, characterised in that said arc-shaped first through fourth magnets and said stator cores which have arc-shaped cross sections have an almost identical circumferential length.

14. The induction generator having a pair of magnetic poles of the same polarity opposed to each other with respect to a rotation shaft as claimed in claim 10, characterised in that said arc-shaped first through fourth magnets and said stator cores which have arc-shaped cross sections have an almost identical circumferential length.

15. The induction generator having a pair of magnetic poles of the same polarity opposed to each other with respect to a rotation shaft as claimed in claim 11, characterised in that said arc-shaped first through fourth magnets and said stator cores which have arc-shaped cross sections have an almost identical circumferential length.

LARRY JAMISON: ELECTRICAL ENERGISER

Patent Application EU 82,400,992.2 22nd December 1982 Inventor: Larry T. Jamison

ENERGY SOURCE EMPLOYING ELECTRICAL ENERGISER

This patent application shows the details of a device which it is claimed, can produce electricity without the need for any fuel. It should be noted that while construction details are provided which imply that the inventor constructed and tested several of these devices, this is only an application and not a granted patent.

ABSTRACT

An energy producing system is provided which produces energy for use, for example, in an electric vehicle or in a home power plant. The system includes an electrical energiser (**60**) including a double-wound rotor and a double-wound stator, for producing electrical energy which is stored in the system, e.g. in a battery (**66**) storage arrangement, which provides initial energisation of the system. the stored energy is supplied to an electric motor (**68**) which drives the energiser (**60**) to thereby create additional energy. the energiser is able to supply the needs of the system as well as to power a load.

BACKGROUND OF THE INVENTION

The present invention relates to energy producing systems and, more particularly, to an electrical energiser-motor system for providing energy, e.g., for an automotive vehicle or as part of a home energy plant.

With the advent of the so-called “energy crisis” and the consequent search for alternative energy sources to substitute for oil, considerable attention has been focused on automotive vehicles as chief users of oil products. One aspect of this search has fostered renewed interest in electrically driven vehicles such as electric cars and the like. A principal shortcoming of prior-art electrical vehicles has been the need to recharge the batteries which provide the power for the electrical motor drive system.

The present invention overcomes this problem through the provision of an electrical energiser-motor system which produces more energy than is expended, thereby enabling the excess energy to be stored in the battery system, to be drawn upon as required. Thus, the need for recharging of the batteries associated with conventional electrical vehicles is eliminated with the system of this invention. It should be noted that while the system of the invention has enormous potential in connection with its use in electrical vehicles, the system is clearly not limited to such use and would obviously be advantageous when used, for example, as the energy source for a home energy plant, as well as in many other applications.

In accordance with the invention, and energy producing system of the type described above is provided which comprises and electrical “energiser” comprising at least one double-wound stator and at least one double-wound shaft-mounted rotor located within a housing, electrical energy being collected from the rotor through a suitable electrical take-off device and being available for utilisation by the system, and an electric motor, powered by the energiser for driving the rotor shaft of the energiser. A battery arrangement is initially used to supply energy to the system and, as stated above, the excess energy generated by the energiser over and above that required by the system and the system load, is stored through charging of the batteries. The motor includes an armature with a plurality of winding slots in it and a plurality of windings being wound into two circumferentially spaced slots in the armature, i.e. such a winding is wound through a first slot (e.g. slot 1) and returned through a second spaced slot (e.g. slot 5). depending on the energy demands, the energiser may include a pair of stators and rotors, with the rotors being mounted on a common shaft. The motor is preferably energised through an arrangement of a commutator and plural brushes, while a slip ring and associated brushes connected to an output bridge circuit form the energy take-off for the energiser.

Other features and advantages of the invention will be shown in the detailed description of the preferred embodiments which follows.

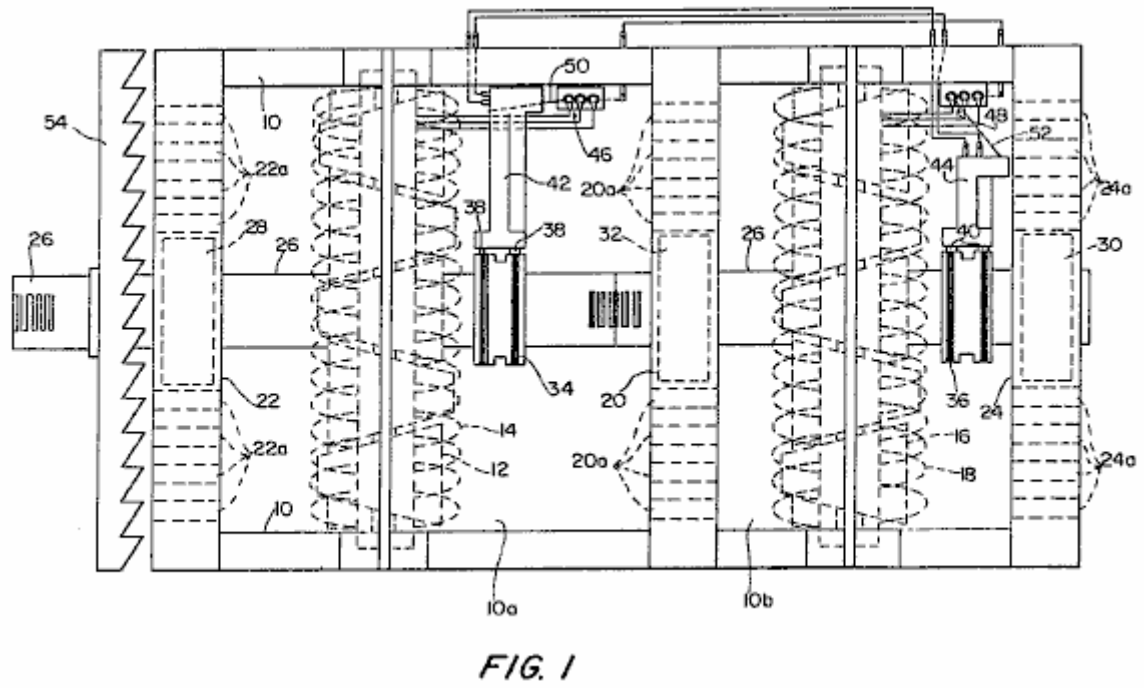


FIG. 1

Fig.1 is a partially sectioned elevational view of the electrical “energiser” of the invention.

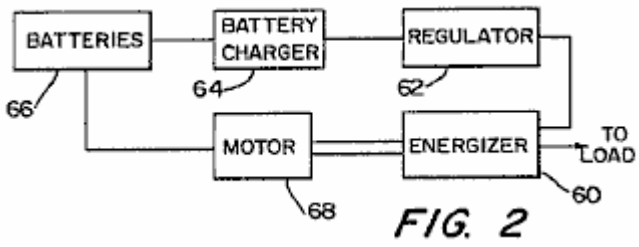


FIG. 2

Fig.2 is a block diagram of the overall energy-producing system of the invention

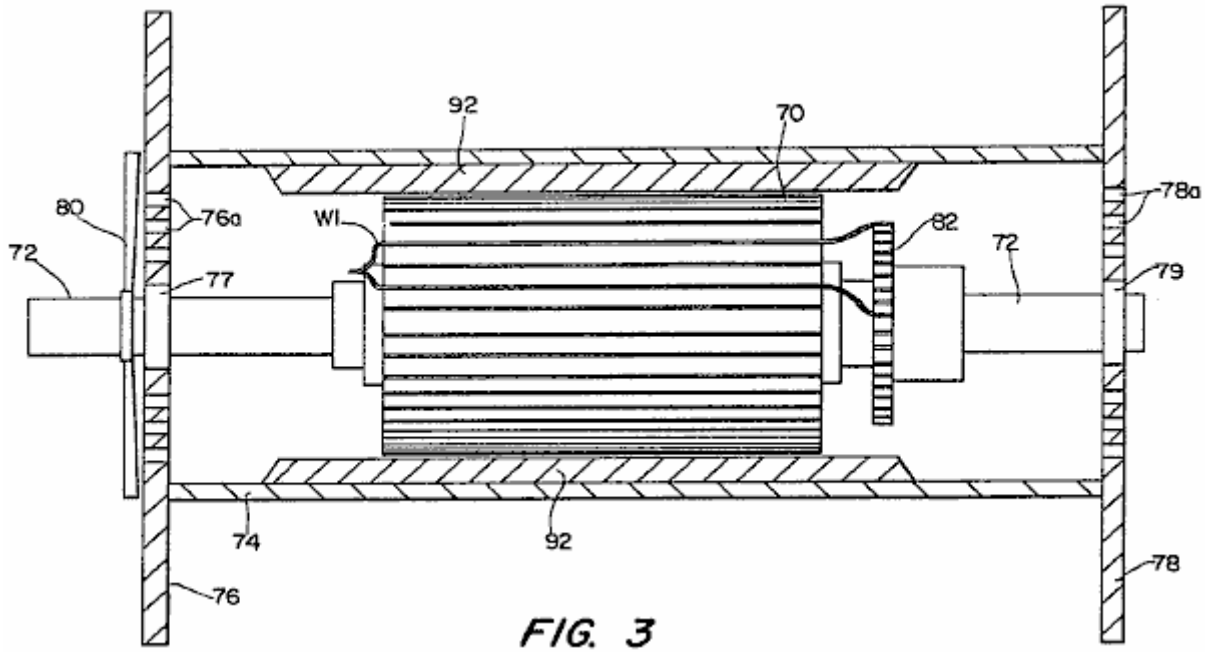


FIG. 3

Fig.3 is a partially sectioned side elevational view of a modified electrical motor constructed in accordance with the invention.

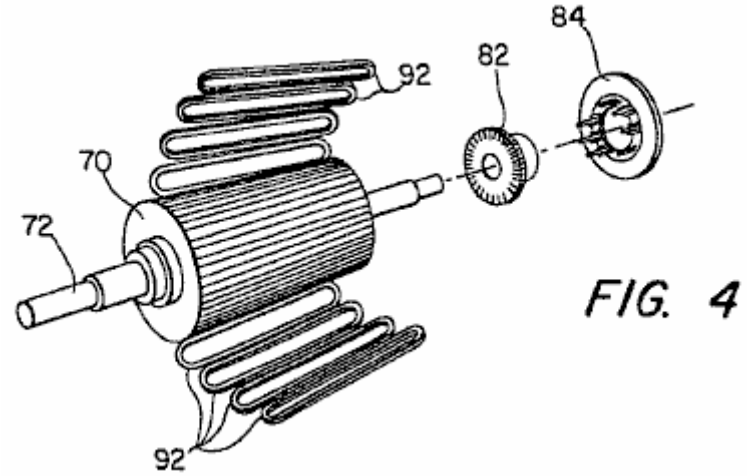


FIG. 4

Fig.4 is an exploded perspective view of the basic components of the motor of **Fig.3**.

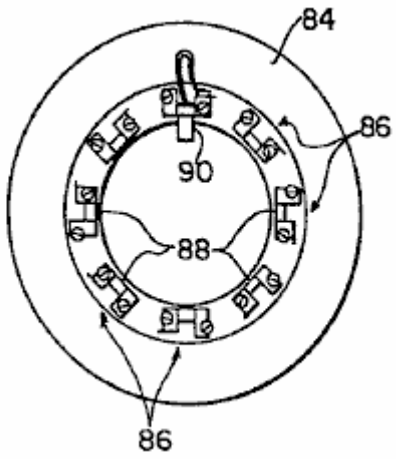


FIG. 5

Fig.5 is an end view of the brush holder also illustrated in **Fig.4**.

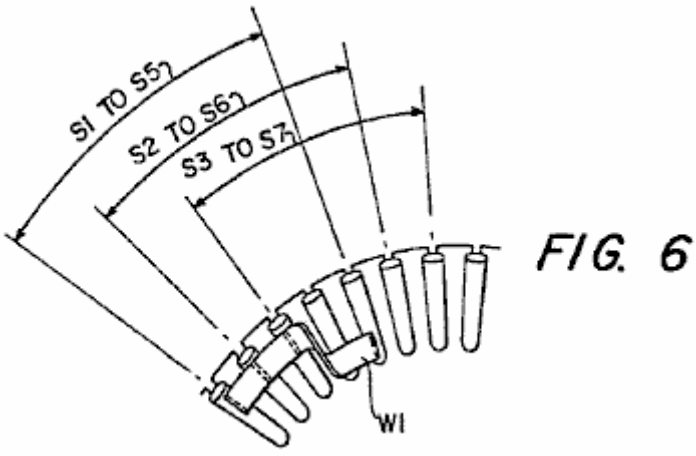


FIG. 6

Fig.6 and **Fig.7** show details of the winding pattern of the motor of **Fig.3**.

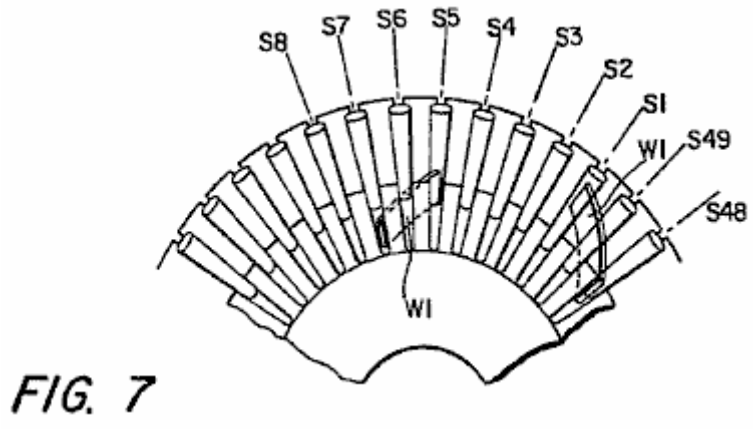


FIG. 7

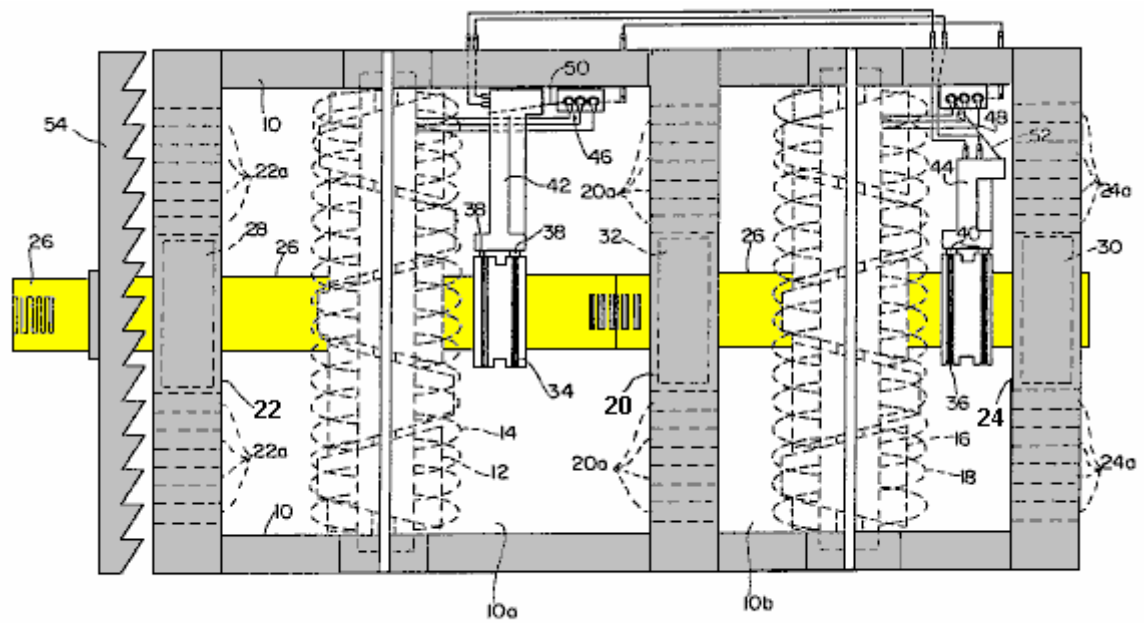


FIG. 1

Referring to **Fig.1**, a preferred embodiment of the "energiser" device of the invention is shown. The device includes a housing **10**, in which are located, in a first chamber or compartment **10a**, a first rotor **12** and a first stator **14** and, in a second compartment **10b**, a second rotor **16**, and a second stator **18**. It should be noted that although two stator-rotor combinations are used in this embodiment, a single stator-rotor combination can be used for some applications. Housing **10** is divided into the compartments **10a** and **10b**, by a centre plate **20** and it includes a pair of end plates **22** and **24**. Both the rotors **12**, **16** and the stators **14**, **18** are double wound and the rotors **12**, **16** are nested inside their respective stators **14** and **18** and mounted for rotation on a common shaft **26**. Shaft **26** extends longitudinally through housing **10** and is mounted on bearings **28** and **30**, supported by end plates **22** and **24**, and a further bearing **32** which is supported by central plate **20**.

A pair of slip rings **34** and **36**, are mounted on shaft **26** and connect with their corresponding brush pairs **38** and **40**.
Slip rings **34** and **36** are connected to rotors **12** and **16** respectively, and permit the current flowing in the rotor windings to be collected through the associated pairs of brushes **38** and **40**. Brush pairs **38** and **40** are mounted on respective brush holders **42** and **44**. The terminals of respective bridge circuits **46** and **48** are connected to stators **14** and **18**, while conversion bars 50 and 52 are connected to brush holders **42** and **44**, as indicated.

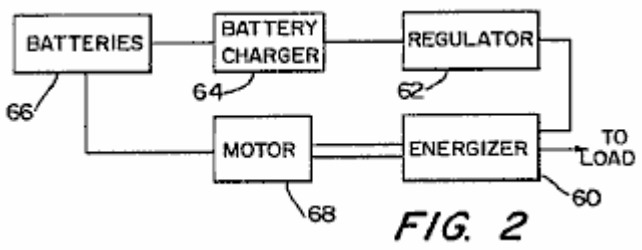


FIG. 2

A cooling fan **54**, is also mounted on shaft **26** and a plurality of apertures **201**, **22a** and **24a** are provided in centre plate **20** and end plates **22** and **24**, to promote cooling of the device. The energiser of **Fig.1** is preferably incorporated in a system such as shown in a highly schematic manner in **Fig.2** where the output of the energiser is used to supply the energy for driving a motor. To this end, the energiser, which is denoted by **60** in **Fig.2**, is connected through a regulator **62**, to battery charger **64** for batteries **66** connected to a motor **68**. These batteries **66** are used to provide the initial energisation of the system as well as to store energy produced by the energiser **60**. It will be understood that the energiser **60** provides energy enough to power motor **68** (which, in turn, drives energiser **60** through rotation of shaft **26**) as well as to provide storage for energy in the system. It will also be appreciated that the system illustrated schematically in **Fig.2** includes suitable controls (switches, rheostats, sensors, etc.) to provide initial energisation as well as appropriate operational control of the system.

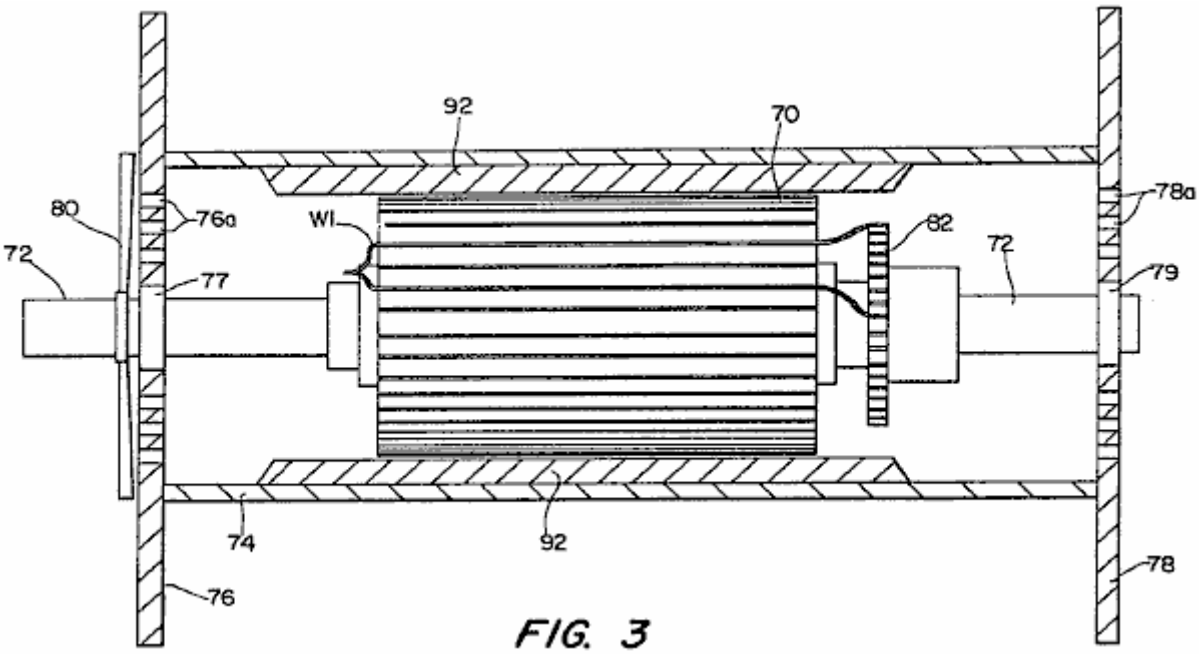


FIG. 3

In a preferred embodiment, motor **68** is of the form shown in **Fig.3**. As illustrated, the motor is of a generally conventional form (with exceptions noted below) and comprises an armature **70**, mounted on a shaft **72** within housing **74**. Housing **74** includes a pair of end plates **76** and **78**, which mount shaft bearings **77** and **79**. Apertures **76a** and **78a** are provided in end plates **76** and **78** and a cooling fan **80** is mounted on shaft **72** to provide cooling.

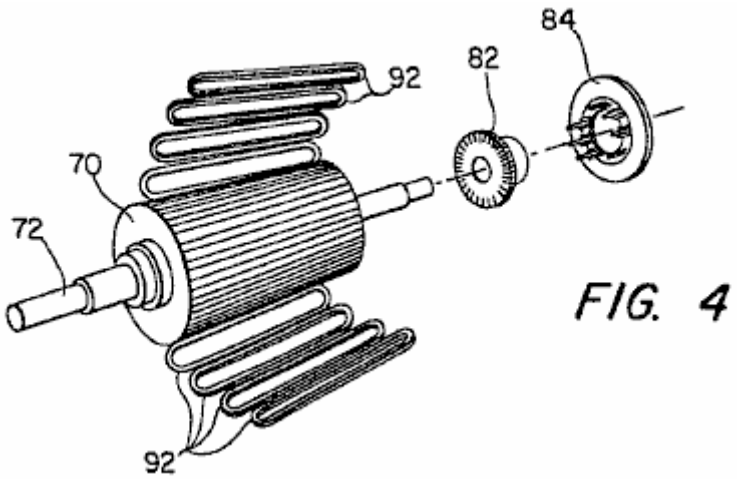
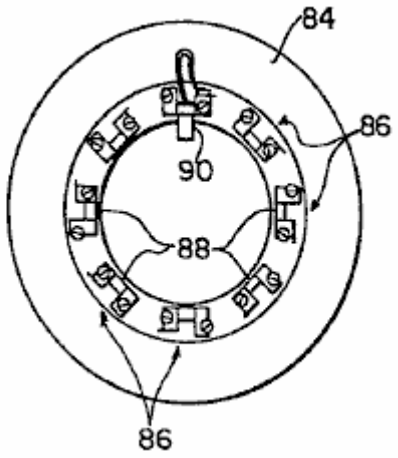


FIG. 4

A commutator **82** is also mounted on shaft **72**, and co-operates with associated brushes (not shown in **Fig.1**), to conduct current to the windings of armature **70**. This co-operation is shown best in **Fig.4** which is an exploded view, illustrating the armature **70**, commutator **82** and a brush holder **84**.

FIG. 5



As shown in **Fig.5**, the brush holder **84** includes eight brush mounts **86**, each of which defines a slot **88** in which a pair of brushes is mounted. One brush **90** is shown in **Fig.5**, it being understood that two such brushes are mounted in each slot **88** so that sixteen brushes are required.

The motor of **Fig.3** to **Fig.6** includes eight pole shoes (not shown) which are secured to housing **74** and which serve to mount eight field coils or windings **92** (see **Fig.3** and **Fig.4**) spaced out around the periphery of armature **72**.

An important feature of the motor of **Fig.3** to **Fig.6** concerns the manner in which the windings for armature **70** are wound. As illustrated in **Fig.3**, **Fig.6** and **Fig.7**, a typical winding **W1** is wound in two slots, with the illustrated winding being doubled back and continuing from armature slot **S1** to armature slot **S5** (see **Fig.3** and **Fig.6**). Similarly, the winding in slot **S2** continues to slot **S6**, the winding of slot **S3** continues to slot **S7**, and so on for the forty-nine windings.

In a specific preferred embodiment, the motor described above is a 48-volt, 412 horsepower motor having a top operating speed of 7,000 rpm. A rheostat control (not shown) is used to control the input voltage and, as discussed above, the motor is powered from the energiser of **Fig.1**. It will be appreciated that the energy take-off from the system is preferably from the output shaft of the motor, although the electrical energy may also be tapped off from the energiser output.

Although the invention has been described in relation to exemplary embodiments, it will be understood by those skilled in the art, that variations and modifications can be effected in these embodiments without departing from the scope and spirit of the invention.

CLAIMS

1. An energy-producing system providing an output for utilisation by a utilising device, the system comprising:

An electrical energising means comprising a housing (**10**); at least one double-wound stator (**14** or **18**) located within the housing; at least one double-wound rotor (**12** or **16**) located within the housing; a rotor shaft (**26**), supported in the housing, and on which the double-wound rotor is mounted; and an energy take-off mechanism (**34** or **36**) including a mechanism for collecting electrical energy from the rotor, mounted on the shaft and connected to the rotor, the mechanism having at least one stationary output.

A motor (**68**), including a connection to the electrical energiser through which to draw the power to operate the motor and drive the rotor shaft of the energiser, the motor having an armature (**70**) with a plurality of winding slots (**S1** to **S49**) in it, and a plurality of windings (**W1**) wound in those slots, at least some of the windings being wound in two slots spaced out around the circumference of the armature (for example, **S1** and **S5**), and an energy supply mechanism (**66**) for supplying electrical energy to the motor at least during initial energisation of the motor, and connected to the energiser for supplying energy to the motor during its operation.

2. A system as in Claim 1, where the energiser includes a pair of these rotors (**12**, **16**) and a pair of stators (**14**, **18**), the rotors being mounted on a common shaft (**26**).

3. A system as in Claim 1, where the energy take-off includes a slip ring (**34** or **36**) and at least one brush (**38** or **40**) for collecting electrical current from the rotor windings, the brush being connected to a bridge circuit (**46** or **48**).

4. A system as in Claim 1, where the motor contains a commutator (**82**) through which energy is supplied to the armature windings.

5. A system as in Claim 4, where the same winding (W1) is wound in the first and fifth slot positions of the motor armature, and the ends of that winding are connected to two positions spaced out around the circumference of the commutator (see **Fig.3**).

United States Patent 5,436,518 25th July 1995 Inventor: Teruo Kawai

MOTIVE POWER-GENERATING DEVICE

Please note that this is a re-worded excerpt from this patent. It describes a motor which has an output power greater than its input power.

SUMMARY OF THE INVENTION

It is an object of the invention to provide a motive power generation device in which the occurrence of a force acting in a direction opposite to the direction of movement of a rotor and/or a stator is prevented, so as to permit efficient use of electric energy to be applied to electromagnets, as well as magnetic energy generated by a permanent magnet.

In order to achieve the above object, the first invention comprises a permanent magnet disposed around a rotational output shaft which is mounted on a bearing, a magnetic body positioned concentrically with the permanent magnet for rotation with the output shaft, the magnetic body being subjected to the magnetic flux of the permanent magnet, a plurality of electromagnets permanently mounted on the support member so that they are spaced a predetermined distance around the periphery of the magnetic material, each magnetic circuit of the electromagnets being independent of one another and the excitation change-over mechanism of the electromagnets which can sequentially magnetise one of the electromagnets which is positioned forward, with regard to a rotational direction, of the output shaft, so as to impart to the electromagnet a magnetic polarity magnetically opposite to that of the magnetic pole of the permanent magnet, whereby a magnetic flux passing through the magnetic body converges in one direction thereby applying a rotational torque to the output shaft.

According to the first invention, when one of the electromagnets which is positioned ahead in the rotational direction of the rotational output shaft, a magnetic field created by the excited electromagnet and a magnetic field created by the permanent magnet interact with each other. Thus, the magnetic flux passing through the magnetic body converges toward the exited electromagnet, so as to rotate the rotational output shaft by a predetermined angle toward the excited electromagnet. When the rotational output shaft has been rotated by the predetermined angle, the above excited electromagnet is de-magnetised, and another electromagnet currently positioned ahead with respect to the rotational direction of the rotor output shaft is excited or magnetised. Sequential excitation of the electromagnets in the above manner permits rotation of the output shaft in a predetermined direction. In this regard, it should be noted that the electromagnets are excited so as to have a magnetic polarity opposite to that of the magnetic pole of the permanent magnet and that the magnetic circuit of the excited electromagnets is independent from those of adjacent electromagnets. Thus, the magnetic flux generated by the excited electromagnet is prevented from passing through magnetic circuits of adjacent electromagnets, which, if it occurs, might cause the electromagnets to be magnetised to have the same polarity as that of the magnetic pole of the permanent magnet. Accordingly, no objectionable force will be generated which might interfere with rotation of the output shaft.

In order to achieve the above object, the second invention comprises a permanent magnet mounted on a movable body arranged movably along a linear track, a magnetic body mounted on the permanent magnet, the magnetic body being subjected to a magnetic flux of the permanent magnet, a plurality of electromagnets spaced an appropriate distance along the linear track, the electromagnets having magnetic circuits which are independent of one another and the excitation mechanism arranged to magnetise each of the electromagnets sequentially when each is positioned forward of the movable body, (with respect to the direction of movement) so as to impart to the excited electromagnet a magnetic polarity opposite to that of the magnetic pole of the permanent magnet, whereby a magnetic flux passing through the magnetic body converges in a predetermined direction so as to cause linear movement of the movable body.

According to the second invention, when the electromagnet positioned ahead of the forward end of the movable body with regard to the direction of the movement of the movable body is excited, a magnetic field generated by the excited electromagnet and magnetic field generated by the permanent magnet interact with each other. Thus, a magnetic flux passing through the magnetic body converges toward the excited electromagnet, so as to displace the movable body a predetermined distance toward the excited electromagnet. When the movable body has been moved the predetermined distance, the movable body is positioned below the above excited electromagnet, and another electromagnet is positioned ahead of the forward end of the movable body. When this occurs, excitation of the electromagnet positioned above the movable body is interrupted, and excitation of the electromagnet now positioned ahead of the forward end of the movable body is initiated. Sequential excitation of

the electromagnets in the above manner permits movement of the movable body in a predetermined direction It should be noted that no objectionable force which would interfere with movement of the movable body is created for the same reason as that explained in relation to the first invention.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig.1 is a front elevational view, partly in section and partly omitted, of a motor according to a first embodiment of the invention;

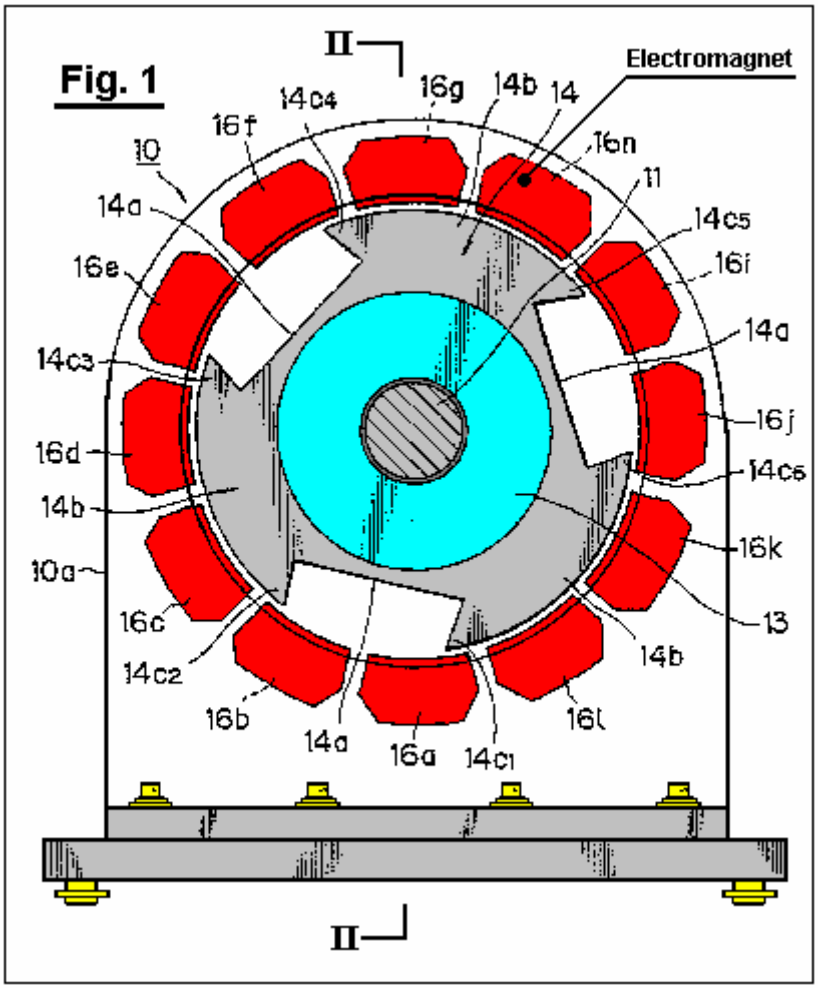


Fig.2 is a sectional view along line II--II in **Fig.1**;

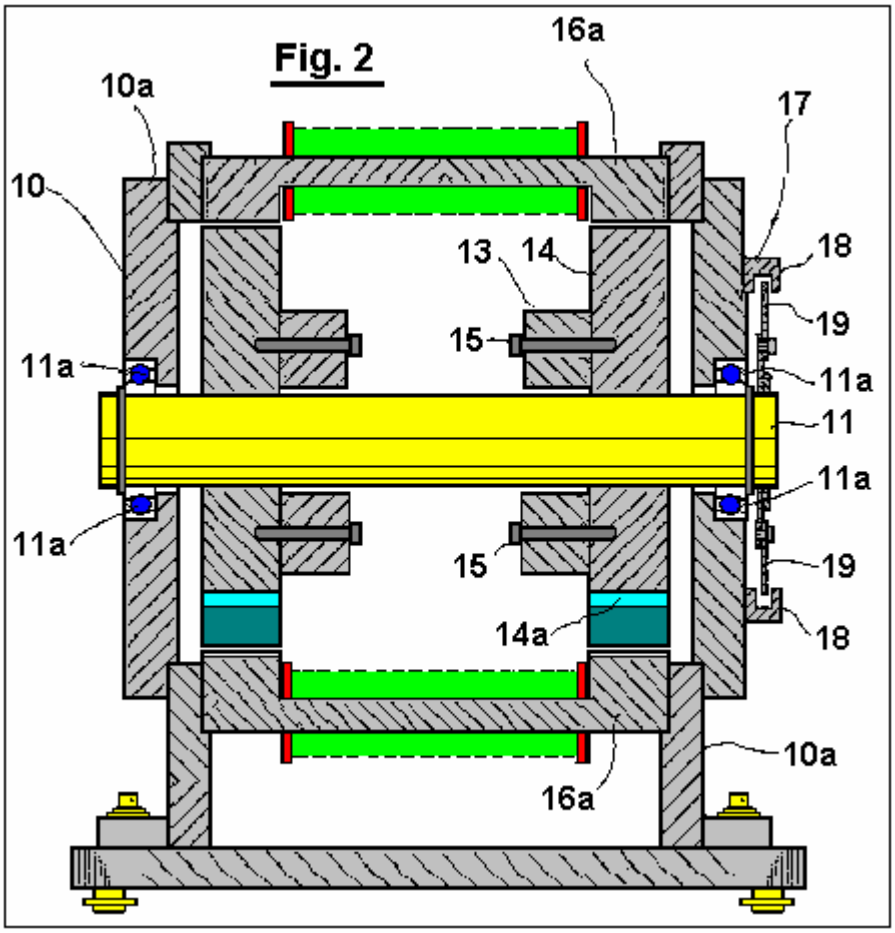


FIG. 3 is a rear elevational view of the motor provided with a light shield plate thereon;

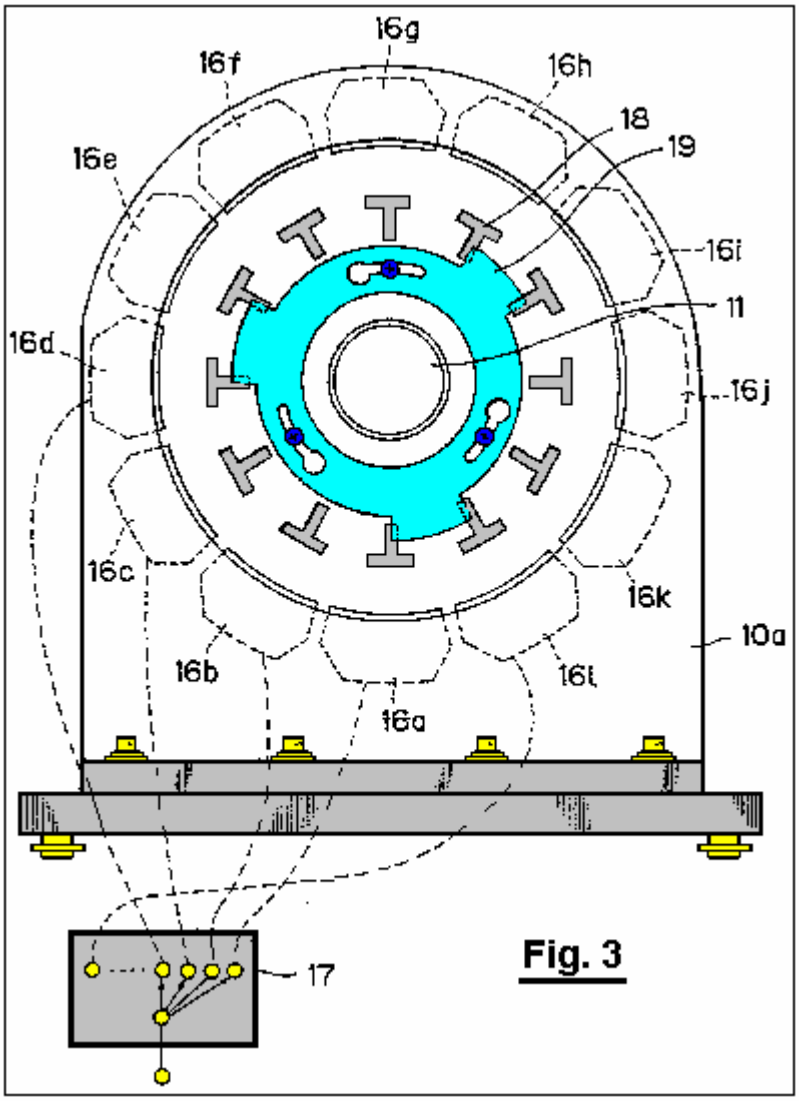


Fig. 3

Fig.4A through **Fig.4H** illustrate operation of the motor when the electromagnets are excited or magnetised;

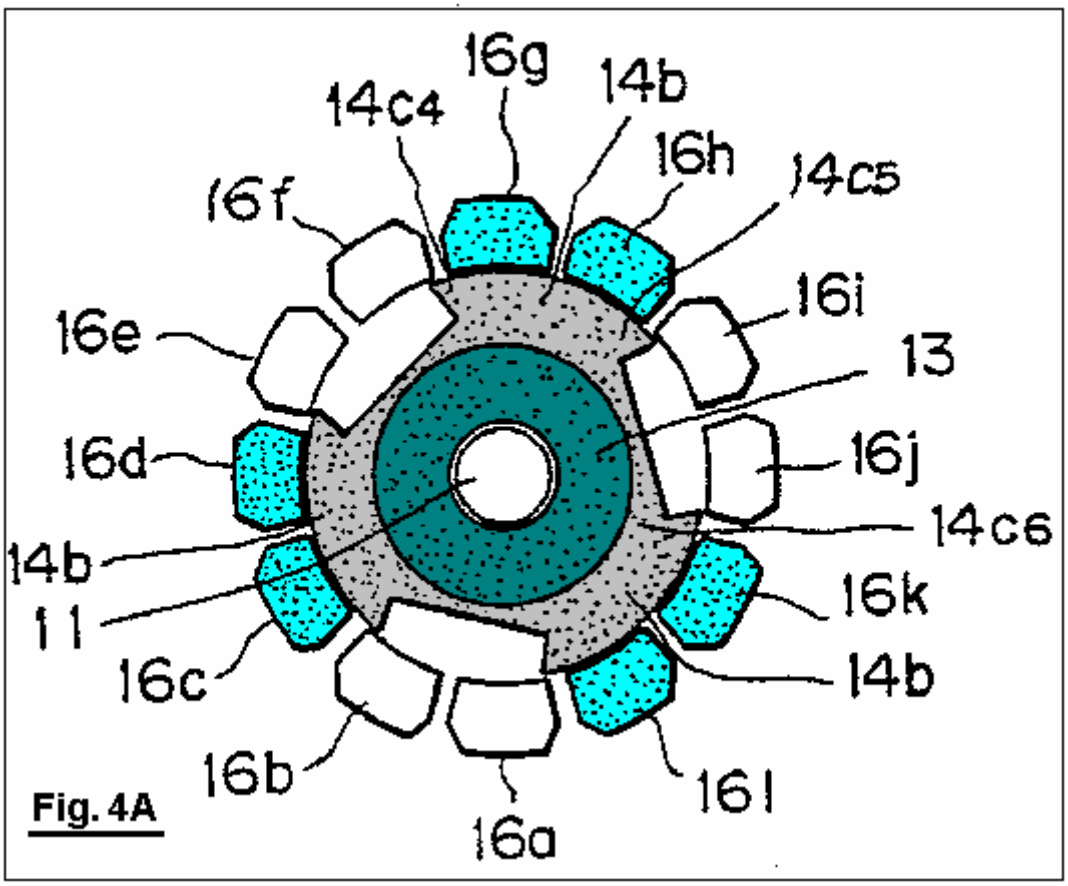


Fig. 4A

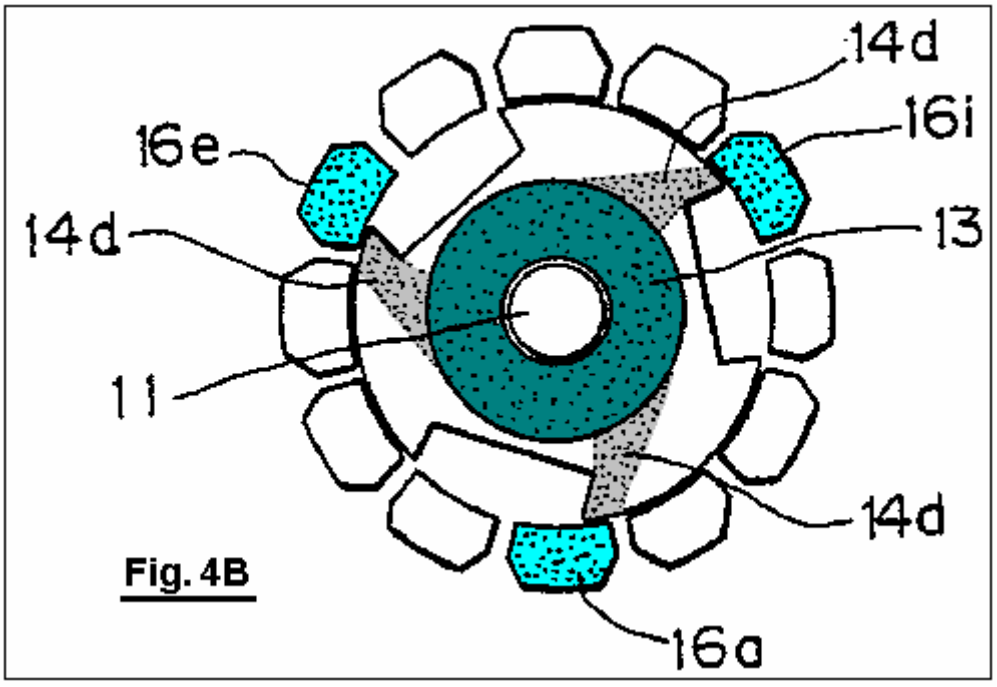


Fig. 4B

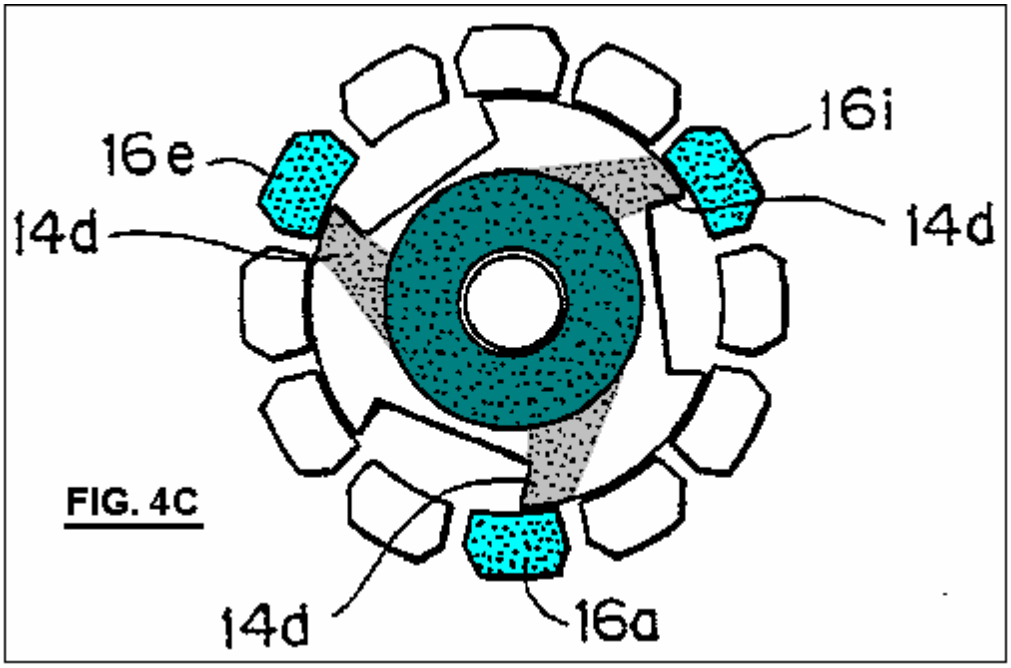


FIG. 4C

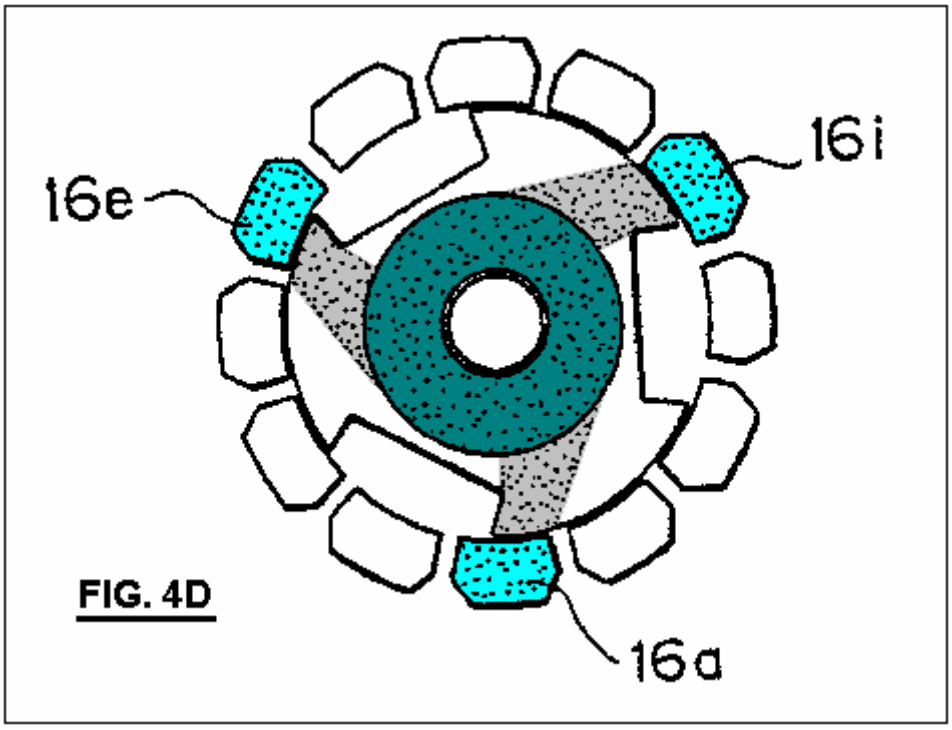


FIG. 4D

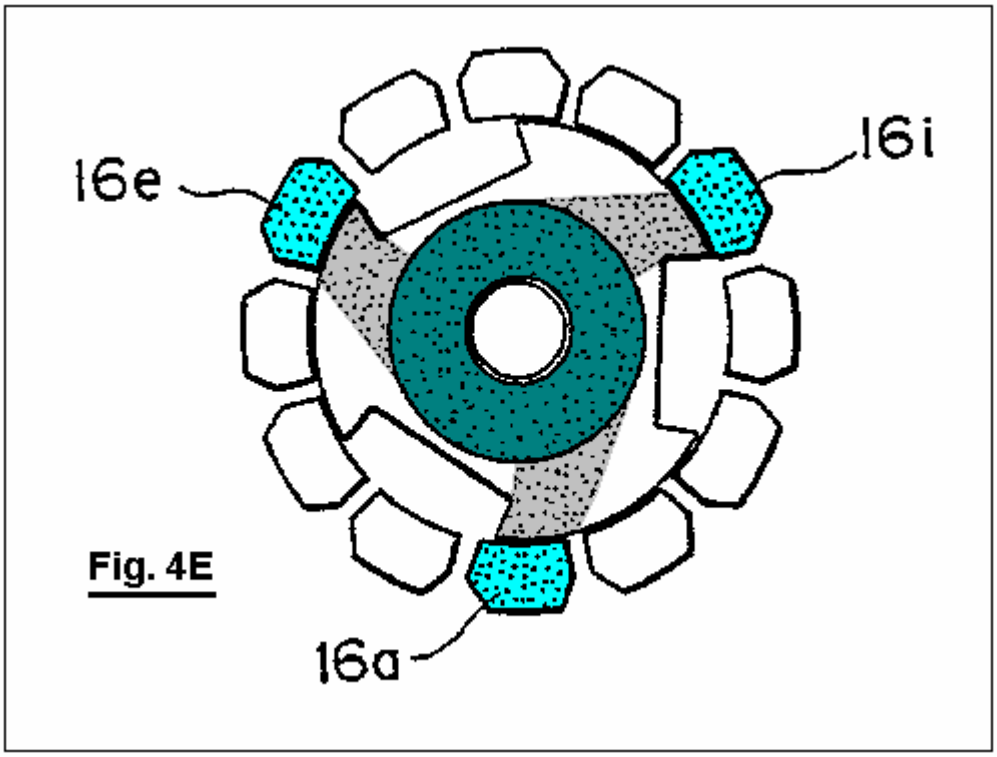


Fig. 4E

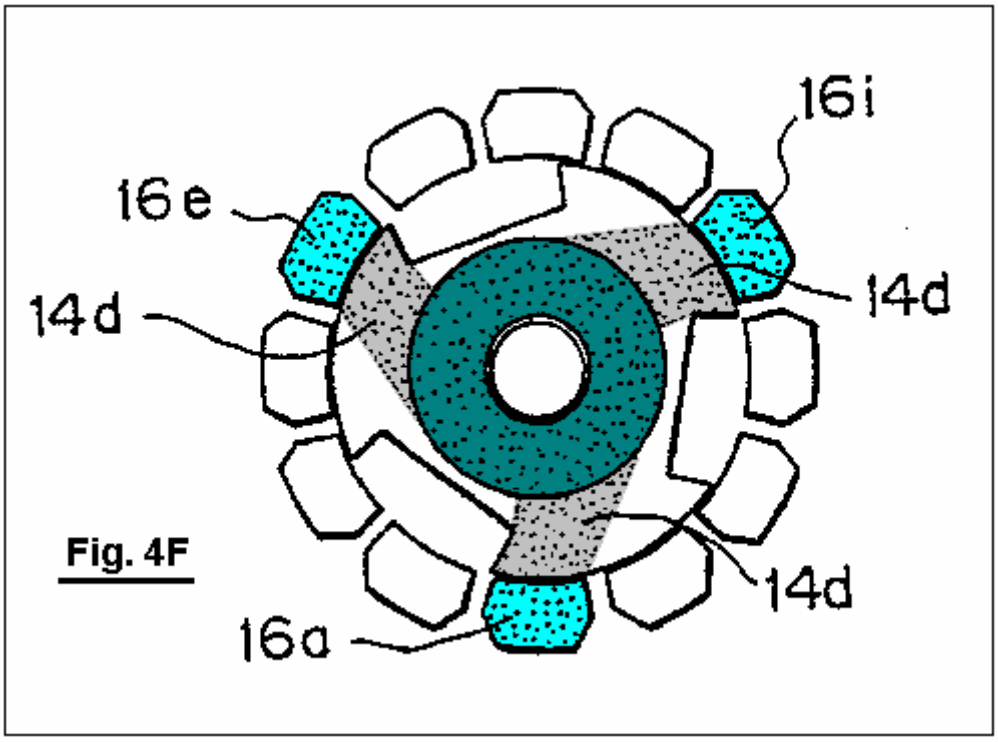


Fig. 4F

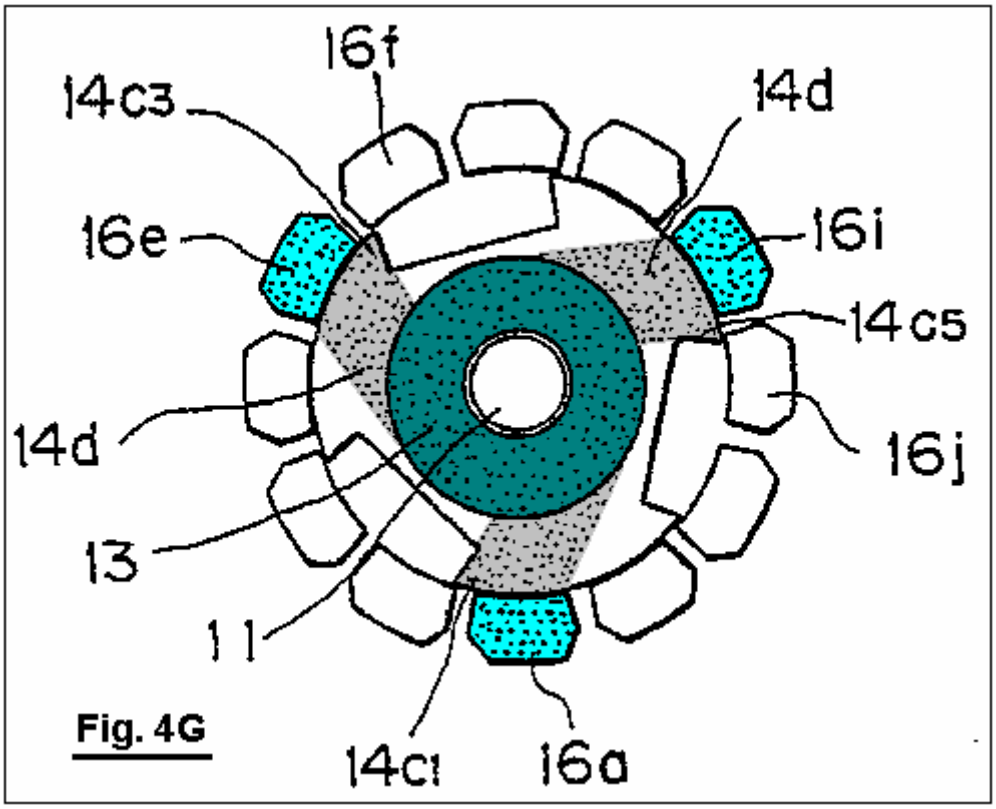

16f
14c3
14d
16e
16i
14c5
14d
16j
13
11
14c1
16a

Fig. 4G

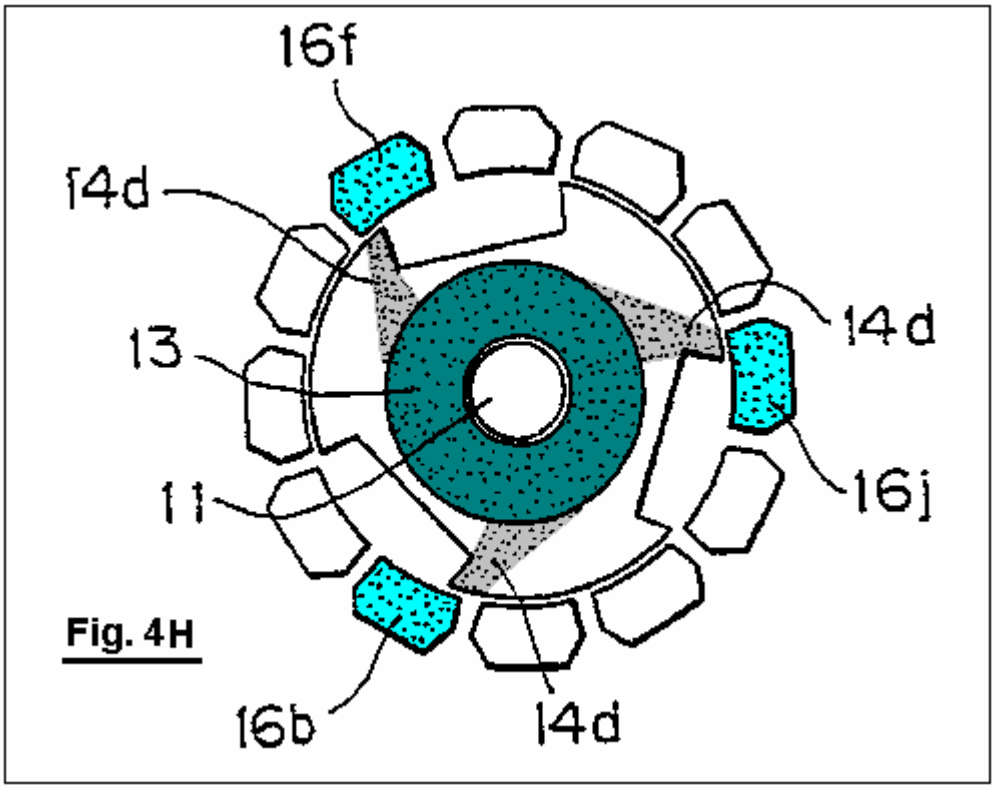

16f
14d
14d
13
11
16j
16b
14d

Fig. 4H

Fig.5A is an illustrative view showing a magnetic path of magnetic flux created by a permanent magnet of the motor when the electromagnets are not magnetised;

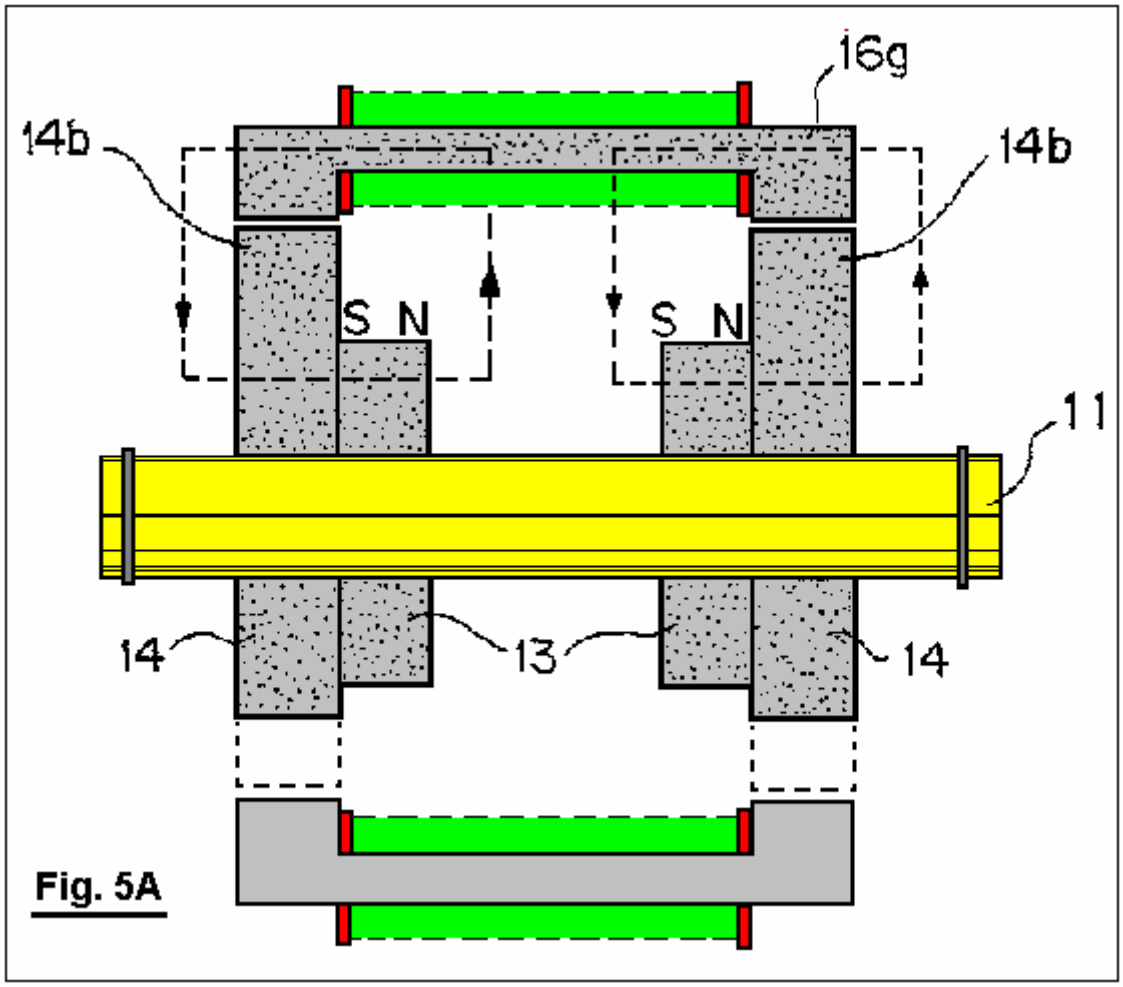


Fig. 5A

Fig.5B is an illustrative view showing a magnetic path of magnetic flux created by the permanent magnet of the motor, as well as magnetic path of magnetic flux created by the electromagnets;

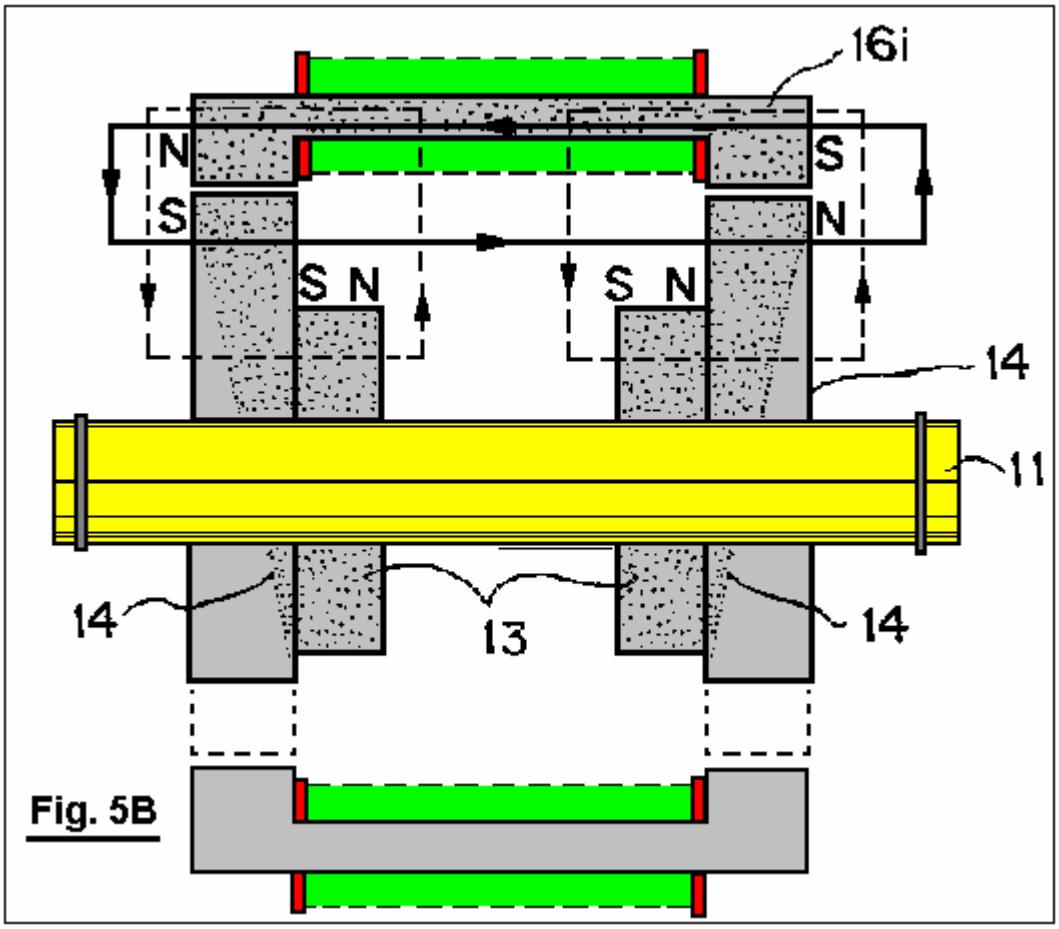


Fig. 5B

FIGS. 6 through 9 are cross-sectional view illustrating a modified form the motor;

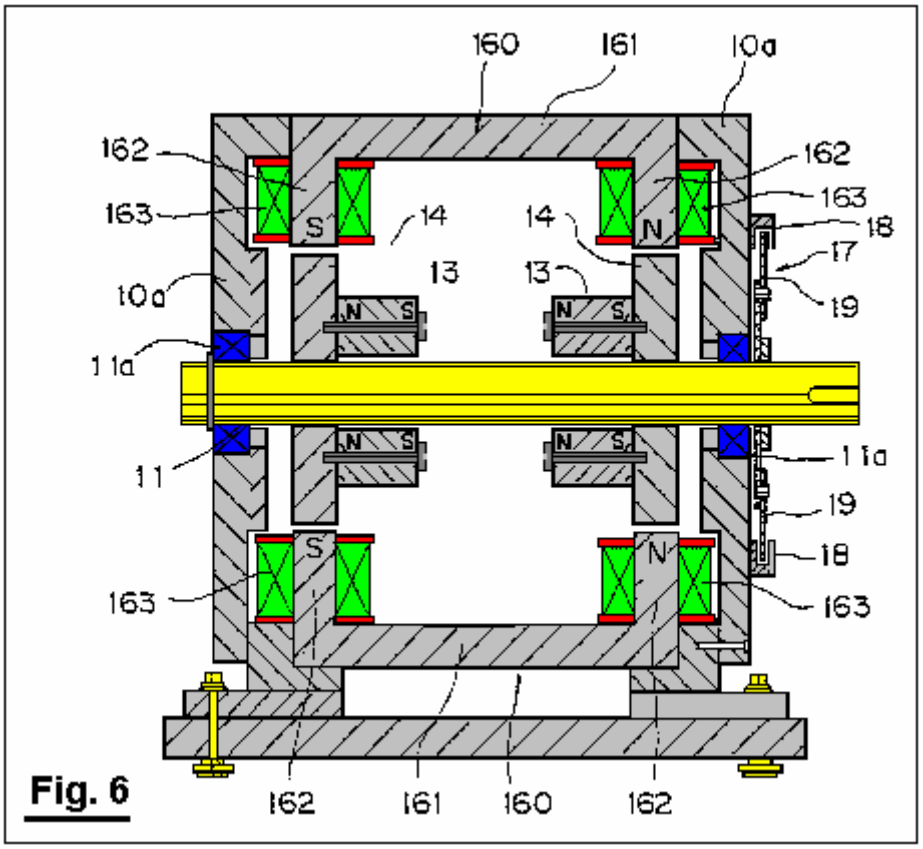


Fig. 6

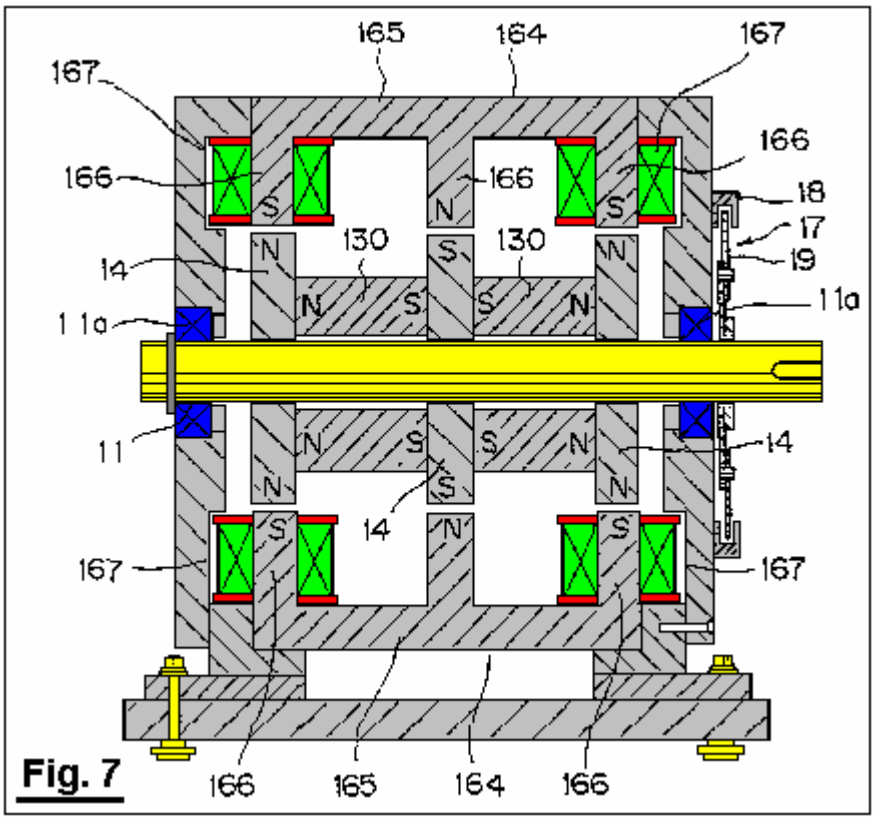


Fig. 7

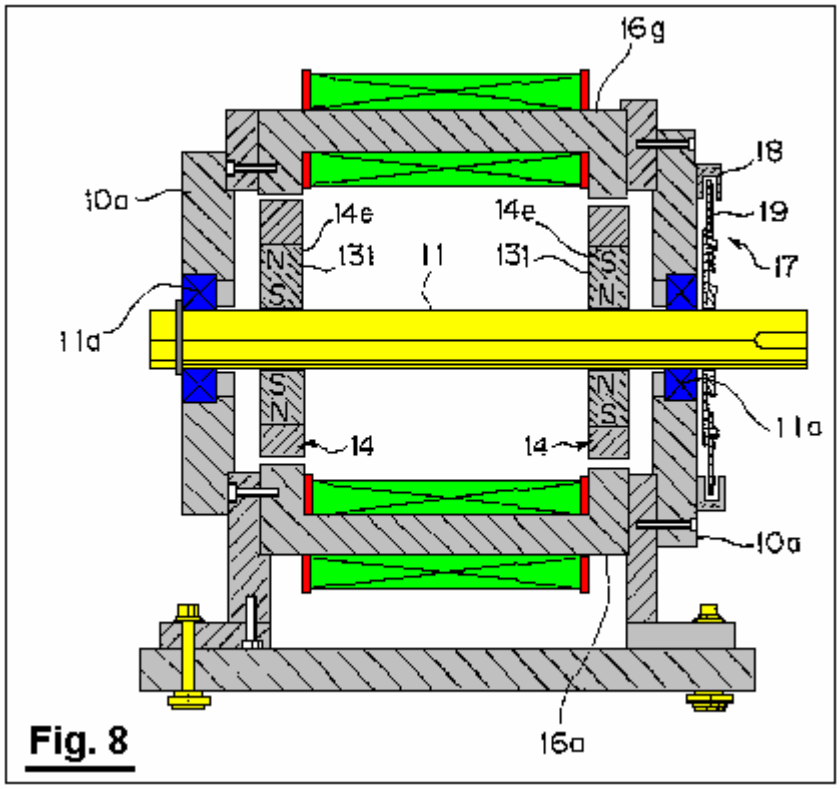

16g
18
10a
14e
14e
19
131
11
131
17
11a
S
N
N
S
S
N
N
S
14
14
11a
10a
16a

Fig. 8

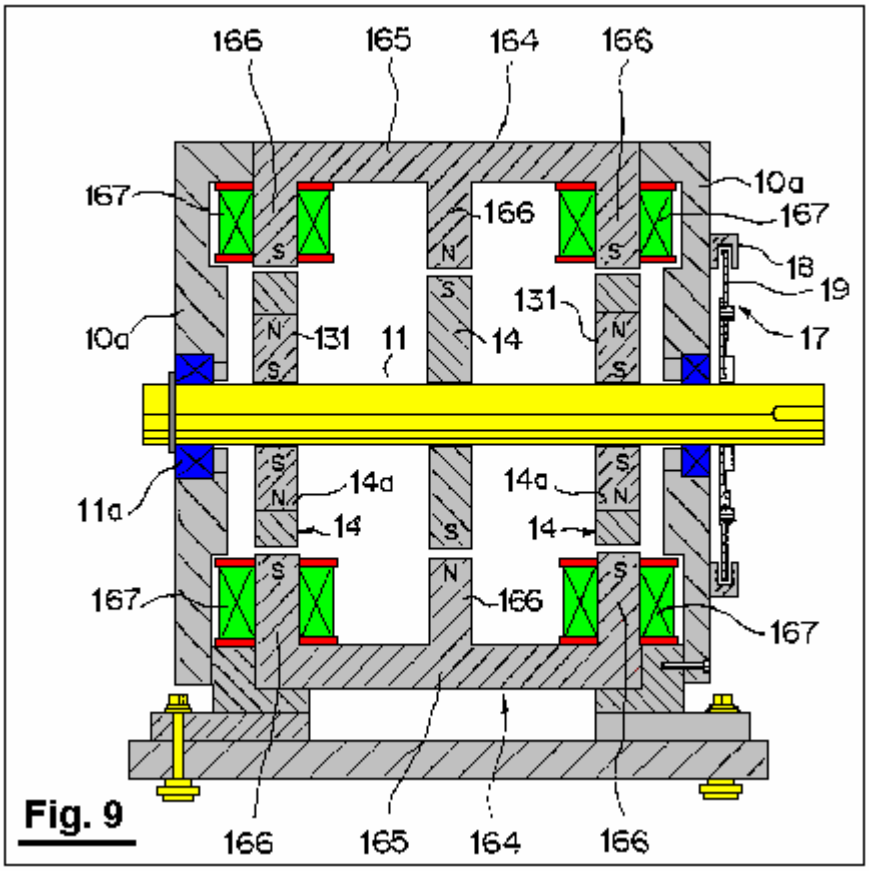

166
165
164
166
167
10a
166
167
18
19
131
10a
131
11
14
17
11a
14a
14a
14
14
167
166
167
166
165
164
166

Fig. 9

FIGS. 10A through 10C are cross-sectional views illustrating operation of the modified motor;

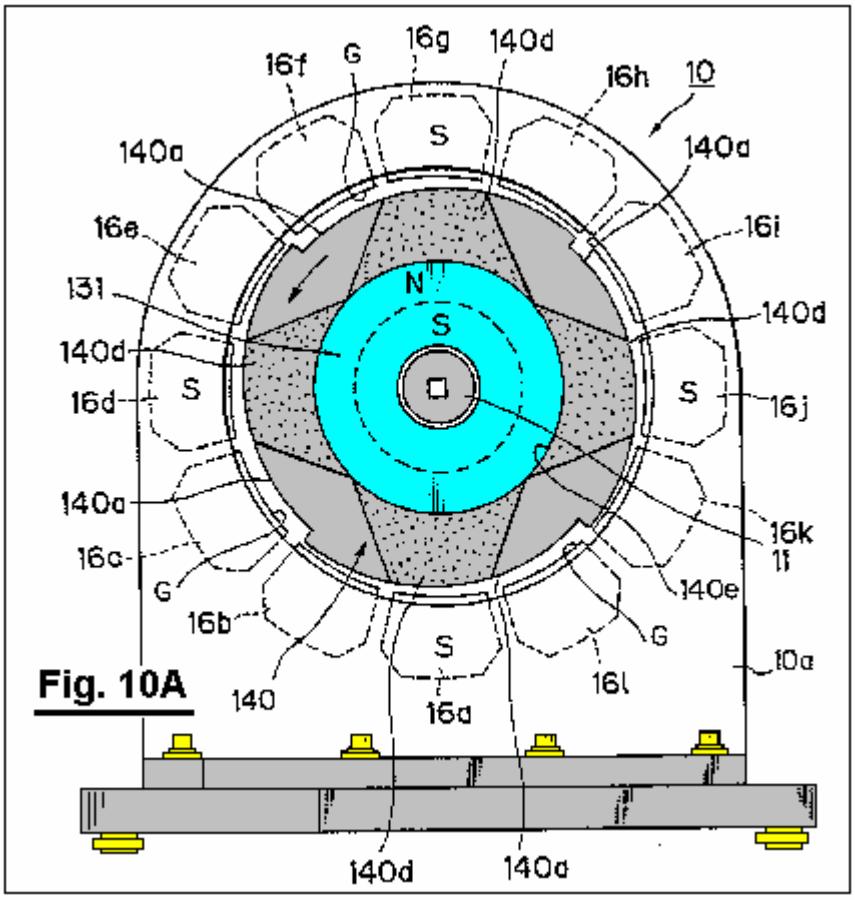


Fig. 10A

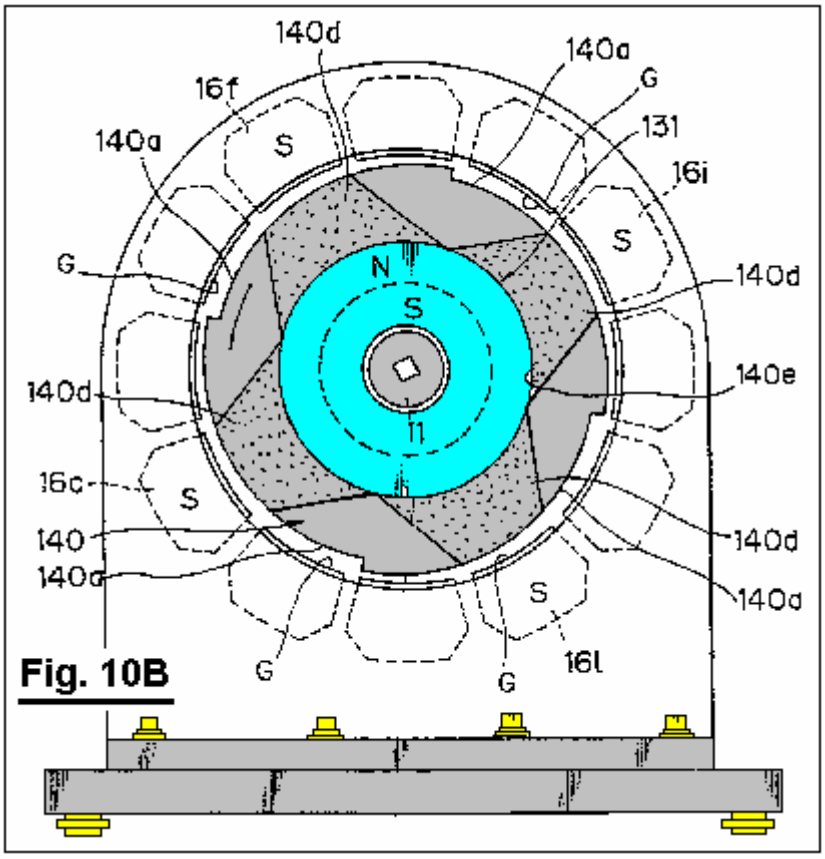


Fig. 10B

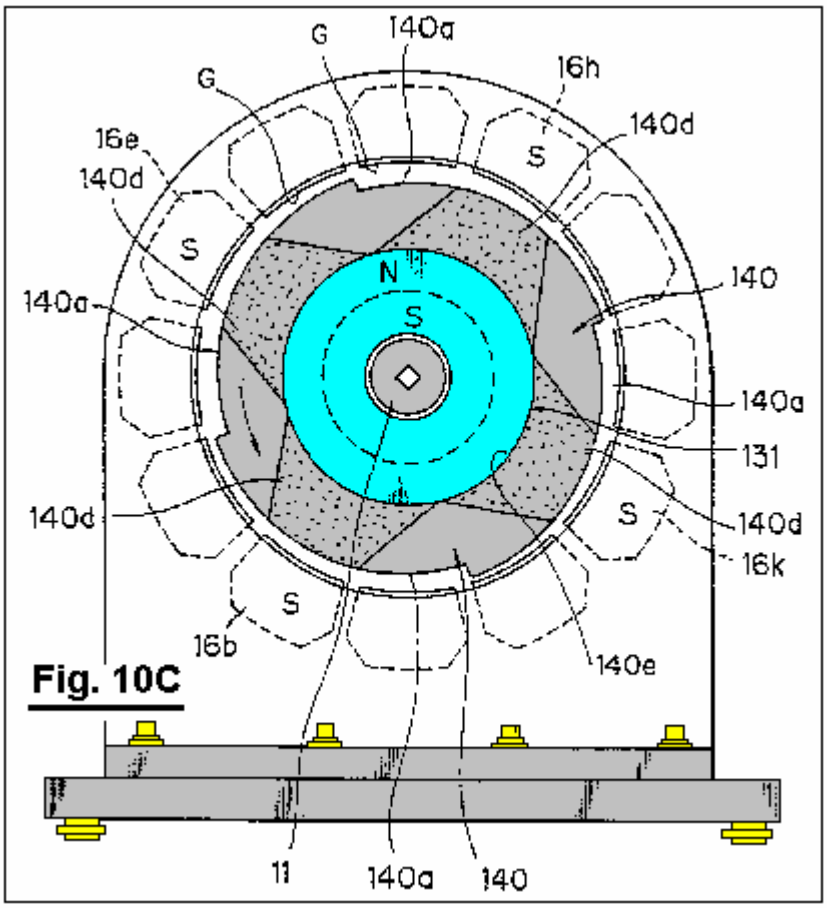


Fig. 10C

FIGS. 11A through 11H are illustrative diagrams showing operation of a motor in a form of a linear motor according to a second embodiment of the invention;

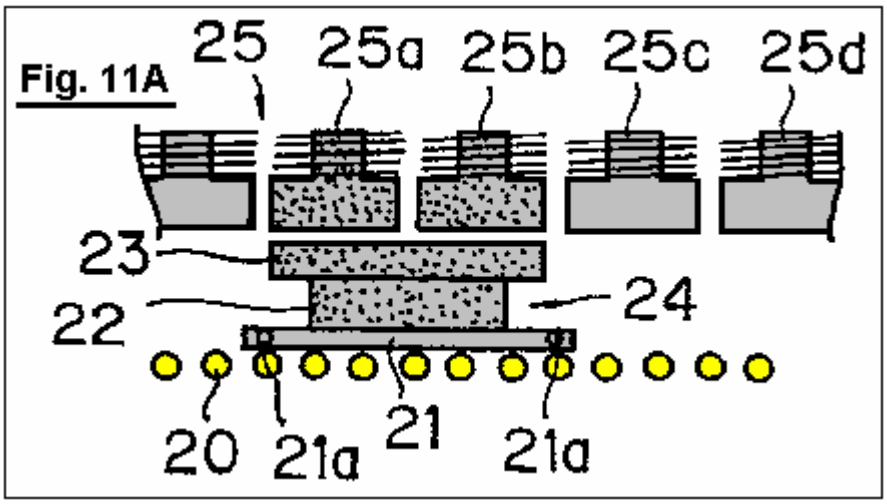


Fig. 11A

Fig. 11B

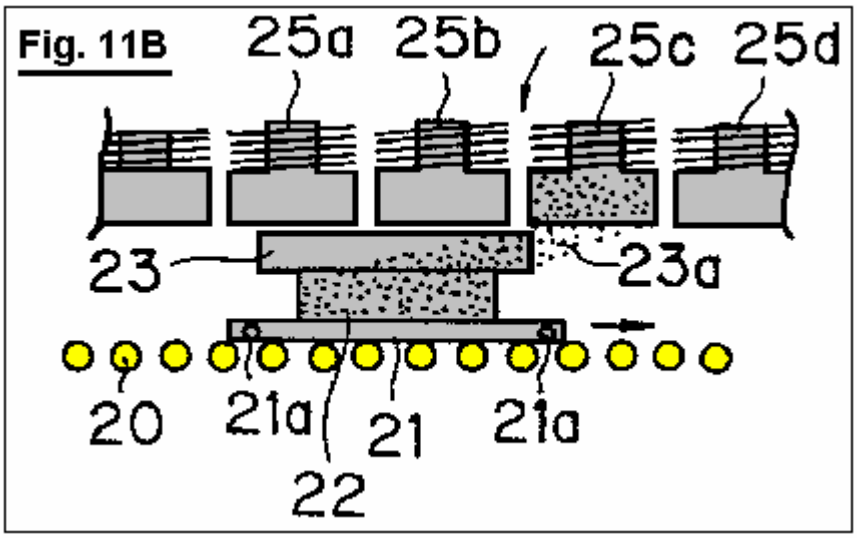


Fig. 11C

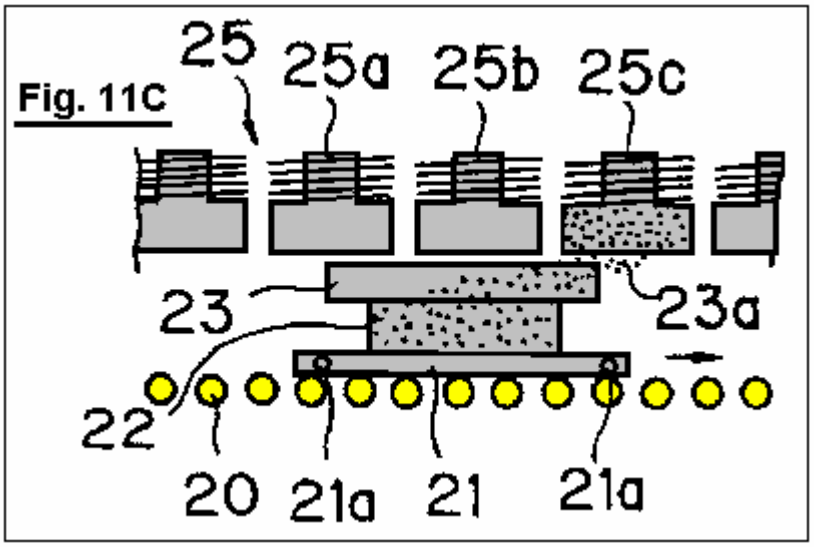
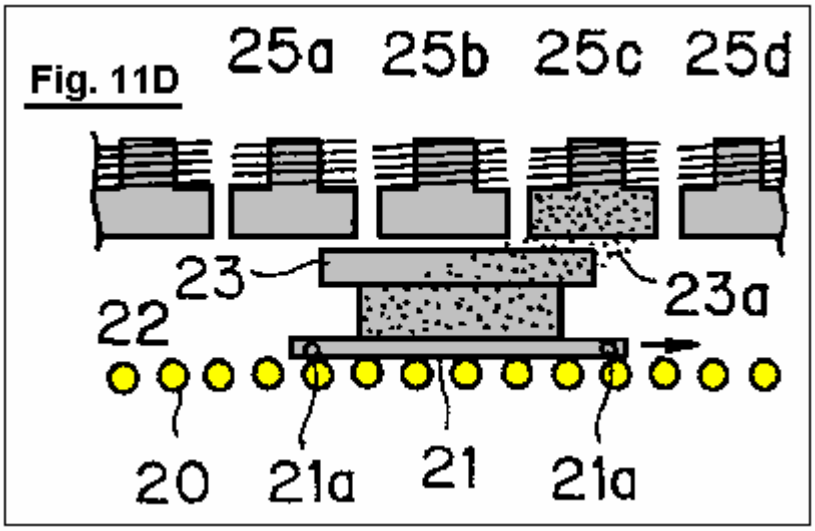
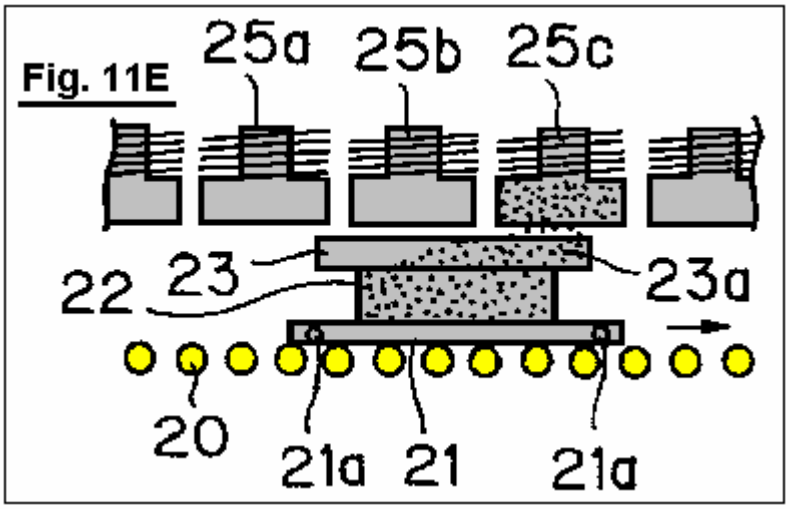
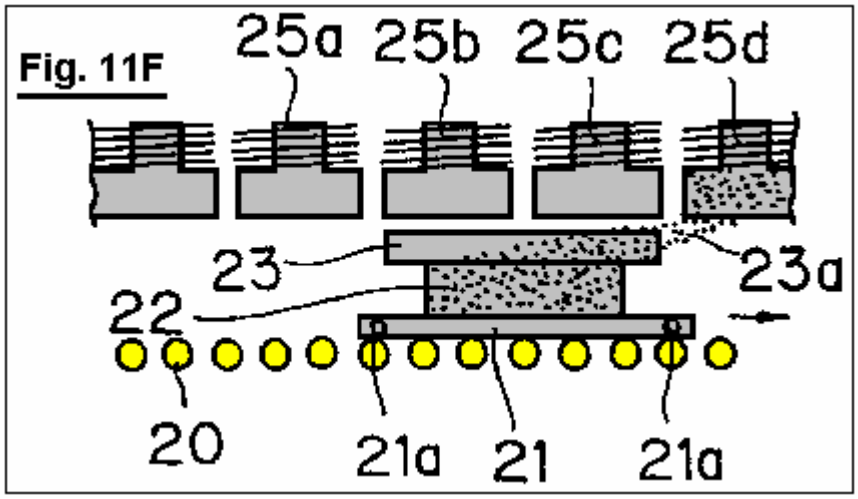


Fig. 11D







DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

Preferred embodiments of the invention will be explained in detail below with reference to the attached drawings.

According to a first embodiment of the invention, a rotational output shaft **11** is mounted in a bearing between front and rear side plates **10a** of a support member **10** through bearings **11a**, as shown in **Fig.1** and **Fig.2**. A ring of permanent magnets **13** are fitted over the opposite ends of the output shaft, inside the side plates **10a** and these move with the rotor shaft **11**. The permanent magnets are magnetised in the axial direction. A magnetic body **14** is rigidly mounted between each of the side plates **10a** of the rotor shaft **11** and the permanent magnets **13**. Each of these magnetic bodies **14** has alternate notches **14a** and magnetic teeth **14b**. It should be noted that the flux of the permanent magnets **13** passes through the respective magnetic bodies **14**. For example, **Fig.1** shows the magnetic body **14** with three notches **14a** and three magnetic teeth **14b**. The permanent magnets **13** and magnetic bodies **14** are positioned co-axially with the rotor output shaft **11**. The corresponding permanent magnets **13** and magnetic bodies **14** are shown connected together by bolts **15** so as to form a rotor **12** which is attached to the rotational output shaft **11**.

It should be noted that the support member **10** and rotational output shaft are both made from a non-magnetic material. The support member **10** may be formed, for example, from stainless steel, aluminium alloys, or synthetic resins, while the rotational output shaft **11** may be formed from stainless steel, for example. Thus, the magnetic circuit formed by the permanent magnet **13** and magnetic body at one axial end of the rotational output shaft **11** and the magnetic circuit formed by the permanent magnet **13** and magnetic body at the opposite axial end of the output shaft, are independent of one another. The magnetic bodies **14** may be formed from magnetic materials having a high magnetic permeability, such as various kinds of steel materials, silicon steel plate, permalloys, or the like.

The stator contains electromagnets **16a** through **16l**, which are positioned between the side plates **10a**. The electromagnets are evenly spaced around the magnetic pieces **14** so that they surround the magnetic bodies. As shown in **Fig.1**, twelve electromagnets may be used. The magnetic circuit of each of the electromagnets **16a**

through **16l** is arranged so as to be independent of each other, so that no flux of a magnetised electromagnet passes through the iron cores of the adjacent electromagnets.

The iron cores of the electromagnets **16a** through **16l** are positioned parallel to the rotor axis shaft **11**, and positioned with only a slight gap between them and the magnetic bodies **14**.

Some of the electromagnets **16a** through **16l** are located at a position corresponding to boundary portions **14c1** through **14c6** between the notch **14a** and the magnetic tooth **14b**. For example, as shown in **Fig.1**, electromagnets **16a, 16b, 16e, 16f, 16i** and **16j** are positioned opposite the boundary portions **14c1, 14c2, 14c3, 14c4, 14c5,** and **14c6**, respectively.

Fig.5A shows a path of magnetic flux created by the permanent magnet **13** when the electromagnets are not excited or magnetised, while, **Fig.5B** shows a path of magnetic flux created by the permanent magnet **13** and a path of magnetic flux created by the windings of the electromagnets when the electromagnets are magnetised. As will be clear from **Fig.5A** and **Fig.5B**, both paths of magnetic flux represent a uni-polar distribution in which N-pole or S-pole evenly appears at the opposite axial ends. When the electromagnets are magnetised, the magnetic fields of the permanent magnet and electromagnets co-operate or interact with each other so as to generate a rotational torque.

Excitation change-over mechanism **17** for sequentially exciting or magnetising the electromagnets **16a** through **16l** is basically consisted of a conventional excitation circuit for supplying direct current to each windings of the electromagnets **16a** through **16l**. In this embodiment, the change-over portion for changing electric feed to the electromagnets **16a** through **16l** includes a plurality of optical sensors **18** and a light shield plate **19** for turning the optical sensors ON and OFF as shown in **Fig.6**.

The optical sensors **18** are spaced apart from one another with a space between them for permitting the light shield plate **19** to pass through a light emitting element and a light receiving element. The optical sensors **18** are disposed in the outer surface of one of the side plates **10a** equally spaced apart along the circumference, so that they are positioned to correspond to the electromagnets **16a** through **16l** (for example, the optical sensor **18** is shown to be disposed in the outer surface of the rear side plate). The light shielding plate **19** is fixed to the rotational output shaft **11** at the end thereof, the light shielding plate protruding from the rear side plate **10a** on which the optical sensors are mounted.

According to the illustrated embodiment, when a particular optical sensor **18** is blocked by the light shielding plate **19**, the electromagnet corresponding to such optical sensor **18** is supplied with electricity.

The operation of the first embodiment described above will be explained with reference to **Fig.4A** through **Fig.4H**.

When the electromagnets **16a** through **16l** are not supplied with electricity by means of the excitation changeover mechanism **17**, the electromagnets **16c, 16d, 16g, 16h, 16k** and **16l** opposed to the magnetic teeth **14b** with a small gap between them merely serve as a magnetic material disposed within the magnetic field of the permanent magnet **13** (refer to shaded portion in **Fig.4A**), so as to absorb the magnetic teeth **14b**, and the rotor **12** remains stationary.

When the electromagnets **16a, 16e** and **16i** positioned adjacent to the boundary portion **14c1, 14c3** and **14c5** formed between the respective notches **14a** and the magnetic teeth **14b** are magnetised or excited simultaneously by means of the excitation change-over mechanism, as shown in **Fig.4B**, the magnetic field of the permanent magnet **13** and the magnetic fields of the electromagnets **16a, 16e** and **16i** interact with each other, so that a magnetic flux **14d** passing through the magnetic body **14** instantaneously converges to the electromagnets **16a, 16e,** and **16i**. In this way, the rotor **12** is imparted with a rotational torque in a direction in which the magnetic flux **14d** will be widened, i.e., counterclockwise direction as viewed in **Fig.4B**.

Fig.4C through **Fig.4G** illustrate change in the width of the magnetic flux **14d** in accordance with rotation of the rotor **12**. When the width of the magnetic flux becomes maximised, i.e., when only the magnetic teeth **14b** are opposed to the electromagnets **16a, 16e** and **16i**, while the notches **14a** are displaced completely away from the electromagnets **16a, 16e** and **16i**, the width of the magnetic flux **14d** is maximised. Thus, an absorption force acting between the permanent magnet **13** and the electromagnets **16a, 16e** and **16i** is maximised. On the other hand, the rotational torque acting on the rotor **12** becomes zero.

Before the rotational torque acting on the rotor **12** becomes zero, i.e., as the boundary portion **14c1, 14c3** and **14c5** approach another electromagnets **16b, 16f** and **16j** positioned ahead of (with regard to the rotational direction), respectively, the electromagnets **16a, 16e** and **16i** are demagnetised and the electromagnets **16b, 16f** and **16j** are excited or magnetised by means of the excitation change-over mechanism **17**. Thus, the magnetic

flux **14d** converges toward the electromagnets **16b, 16f** and **16j**, as shown in **Fig.4H**, so that a rotational torque acts upon the rotor, as described above.

Then, the electromagnets **16c, 16g** and **16k** are excited. When the boundary portion **14c1, 14c3** and **14c5** approach another electromagnets **16d, 16h** and **16l** positioned ahead with respect to the rotational direction, in response to rotation of the rotor **12**, the electromagnets **16c, 16g** and **16k** are de-magnetised and the electromagnets **16d, 16h** and **16l** are energised or excited.

As explained above, sequential excitation or energising of the electromagnets **16a** through **16l** causes interaction between the magnetic flux of the permanent magnet **13** and the electromagnets **16a** through **16l**, whereby a rotational torque is applied to the rotor **12**.

When this occurs, a rotational torque is generated between one of the magnetic poles of the permanent magnet **13** (for example, N-pole) and the magnetic poles (for example, S-poles) of the electromagnets **16a** through **16l** positioned at their respective axial ends. A rotational torque is also generated between the other magnetic pole (for example, S-pole) of the permanent magnet **13** and the other magnetic pole (for example, N-pole) of each of the electromagnets **16a** through **16l** positioned at the other axial end.

It should be noted that, at one magnetic pole, for example N-pole, of the permanent magnet **13**, certain of the electromagnets **16a** through **16l** are magnetised only to S-pole, thus preventing formation of a magnetic circuit, due to passage of magnetic flux from the excited electromagnets through either of the adjacent electromagnets, which tends to bring about N-poles magnetically similar to the permanent magnet **13**. It is also noted that, at the other magnetic pole, for example S-pole, of the permanent magnet **13**, certain of the electromagnets are magnetised only to N-pole, thus preventing formation of a magnetic circuit, due to passage of magnetic flux from the excited electromagnets through adjacent electromagnets, which tends to bring about S-poles magnetically similar to the permanent magnet **13**. The magnetic flux of the permanent magnet **13** passes through the magnetic bodies **14** so as to be converged to the excited electromagnets (refer to the magnetic flux **14d** shown in **Fig.4** through **Fig.4H**), thus forming dead zones, through which no magnetic flux passes, in the magnetic bodies **14** at a position opposite to the un-excited electromagnets. Accordingly, no force is generated which would tend to prevent rotation of the rotor **12**.

In view of electric energy applied to the electromagnets **16a** through **16l**, substantially all the electric energy applied is used to contribute to the rotation of the rotor **12**. On the other hand, and in view of magnetic energy of the permanent magnet **18**, all the magnetic energy contributes to the rotation of the rotor **12**.

It is also noted that, since the notches **14a** and the magnetic teeth **14b** are alternately disposed in the outer periphery of the magnetic materials **14** in an acute angle configuration seen in **Fig.4A** to **Fig.4H**, and the electromagnets are disposed at a position each corresponding to the boundary portions between the notches and the magnetic teeth, it is possible for the line of the magnetic force, generated in each gap between the boundary portions and the electromagnets when the electromagnets are excited, to be inclined to a substantial degree, so that a sufficient degree of rotational torque may be obtained upon initial excitation of the electromagnets.

The result obtained during an actual running test of the motor according to the first embodiment is shown in **Fig.1** to **Fig.3**.

Pure steel was used as a magnetic material. The magnetic material was 30 mm in thickness and formed to have magnetic teeth of 218 mm diameter and notches of 158 mm diameter. A ferrite magnet was used as a permanent magnet. The magnetic force of the magnet was 1,000 gauss. Electric power of 19.55 watts was applied to the electromagnets at 17 volts and 1.15 amperes. The above conditions produced a rotational speed of 100 rpm, with a torque of 60.52 Kg-cm and an output of 62.16 watts.

Alternative embodiments will be explained below with reference to **Fig.6** through **Fig.9**.

The modified embodiment shown in **Fig.6** is similar to the motor presented as the first embodiment as shown in **Fig.1** through **Fig.3**, with the exception that each electromagnet **160** used as part of the stator, comprises an iron core **161** having a pair of legs **162** which extend towards the outer periphery of the magnetic bodies (outer periphery of the magnetic teeth **14b**), each of the legs being wound with coils **163**. The remaining components are basically identical to those in the motor shown in **Fig.1** through **Fig.3**. In **Fig.6**, the components similar to those in **Fig.1** through **Fig.6** are denoted by like reference numerals. It should be noted that each coil **163** is supplied with electricity so that one leg **162** (left-hand side in **Fig.6**) of each of the iron cores **161** is magnetised to be S-pole which is magnetically opposite to the magnetic pole (N-pole) of the confronting magnetic body **14**, while the leg **162** disposed at the other end of each of the iron cores is magnetised to be N-pole which is magnetically opposite to the magnetic pole (S-pole) of the confronting magnetic body **14**.

According to this modified embodiment, it is possible to significantly reduce leakage of the magnetic flux created by the electromagnets **160** in gaps each defined between the surfaces of the magnetic poles of the electromagnets **160** and the outer peripheries of the magnetic teeth **14b** of the magnetic bodies **14**.

An alternative embodiment shown in **Fig.7** is similar to the motor shown in **Fig.1** through **Fig.8**, with the exception that: an additional magnetic body **14** is mounted on the rotational output shaft **11** at it's axial midpoint; two permanent magnets **130** are freely mounted on the output shaft **11** in the manner shown in **Fig.6**; and each iron core **165** is provided with three legs **166** positioned at the opposite axial ends and midpoint thereof and extending toward the respective outer periphery of the magnetic bodies, with the legs **166** positioned at axial opposite ends of the respective iron cores **165** being wound with a coil **167**, which form electromagnets **164**. The remaining components are substantially the same as those in the motor shown in **Fig.1** through **Fig.3**. It should be noted here, that the rotational output shaft **11** may be formed from either magnetic materials or non-magnetic materials.

As shown in **Fig.7**, each of the coils **167** is supplied with electricity so that the legs **166** positioned at the opposite axial ends of each of the iron cores **164** is magnetised to be S-pole which is magnetically opposite to the magnetic pole (N-pole) of the confronting magnetic body **14**. By this, the leg **166** positioned at the midpoint of the iron core **165** is magnetised to be N-pole which is magnetically opposite to the magnetic pole (S-pole) of the confronting magnetic body **14**.

In this embodiment, it is also possible, as in the modified embodiment shown in **Fig.6**, to significantly reduce the leakage of the magnetic flux generated by the electromagnets **164**. In addition to this, it is also possible to obtain a rotational torque between the leg **166** positioned at the midpoint of the iron core and the magnetic body **14** positioned at the axial midpoint of the rotational output shaft **11**. Accordingly, a higher rotational torque may be obtained with the same amount of electrical consumption, in comparison with the embodiment shown in **Fig.6**.

A further embodiment shown in **Fig.8** is similar to the motor shown in **Fig.1** though **Fig.3**, with the exception that a permanent magnet magnetised in the radial direction, rather than in the axial direction is employed. The permanent magnet **131** of an annular configuration has, for example, N-pole in the outer periphery and S-pole in the inner periphery. The permanent magnet **131** is received within a cavity **14e** provided in the respective magnetic body **14** at the intermediate portion thereof as disposed at the opposite axial ends of the rotational output shaft **11**. The remaining components are identical to those in the motor shown in **Fig.1** though **Fig.3**. The components identical to those in the motor shown in **Fig.1** though **Fig.3** are denoted by the same reference numerals. It should be noted that this embodiment may also employ the electromagnets **160** shown in **Fig.6**.

In this embodiment, the rotational output shaft **11** may be formed from magnetic materials, rather than non-magnetic materials.

Further embodiment shown in **Fig.9** is similar to the motor shown in **Fig.1** though **Fig.3**, with three exceptions. The first exception is that a permanent magnet magnetised in the radial direction, rather than in the axial direction is employed. The permanent magnet **131** having an annular configuration has, for example, N-pole in the outer periphery and S-pole in the inner periphery. The permanent magnet **131** is received within a cavity **14e** provided in the respective magnetic body **14** at the intermediate portion thereof as disposed at the axial opposite ends of the rotational output shaft **11**. The second exception is that an additional magnetic body **14** is disposed at the axial midpoint of the rotational output shaft **11**. Finally, the third exception is that the iron core **165** is provided with three legs **166** disposed at the axial opposite ends and the midpoint thereof, respectively, and extending toward the outer periphery of the magnetic body **14**, with the legs positioned at the opposite axial ends being wound with respective coils so as to form an electromagnet **164**. The remaining components are identical to those in the motor shown in **Fig.1** though **Fig.3**. The components identical to those in the motor shown in **Fig.1** though **Fig.3** are denoted by the same reference numerals.

As shown in **Fig.9**, each coil is supplied with electricity so that the legs **166** disposed at opposite axial ends of the iron core **165** are magnetised to be S-pole which is magnetically opposite to the magnetic pole (N-pole) of the confronting magnetic body **14**. By this, the leg **166** disposed at the midpoint of the iron core **165** is magnetised to be N-pole which is magnetically opposite to the magnetic pole (S-pole) of the confronting magnetic body **14**.

According to the embodiment described above, the rotational output shaft **11** may be formed from magnetic materials rather than non-magnetic materials. With this embodiment, it is possible to obtain the same effect as that obtained with the embodiment shown in **Fig.7**.

Further the alternative embodiments shown in **Fig.10A** to **Fig.10C** are similar to the motor shown in **Fig.1** though **Fig.3**, with the exception that: like the embodiments shown in **Fig.8** and **Fig.9**, an annular permanent magnet **131** is employed which is received in a cavity **140e** provided in the central portion **140** of the magnetic body **140**; the magnetic body **140** is provided with notches **140a** in the outer peripheral portion thereof, so that the gap G

between the magnetic body **140** and the electromagnet becomes gradually broader in the rotational direction of the rotor; and the electromagnets confronting to the gap G with an intermediate width as positioned between the electromagnets confronting to the gap G with a narrower width and the electromagnets confronting to the gap G with a broader width are excited or magnetised in a sequential manner. The remaining components are identical to those in the motor shown in **Fig.1** though **Fig.3**. In **Fig.10A** to **Fig.10C**, the components identical to those in **Fig.1** though **Fig.3** are denoted by the same reference numerals. In this regard, it should be noted that reference numeral **140d** indicates magnetic flux passing through the magnetic body **140**, so as to illustrate converged condition of such magnetic flux upon excitation of the electromagnets.

In the embodiment Just described above, it is possible to rotate the rotor in the counter clockwise direction as viewed in **Fig.10A**, for example, by exciting the electromagnets **16a, 16d, 16g** and **16j**, as shown in **Fig.10A**, then, the electromagnets **16c, 16f, 16i** and **16l**, as shown in **Fig.10B**, and then the electromagnets **16b, 16e, 16h** and **16k**. According to this embodiment, it is possible to obtain a stable rotational force, as well as a higher rotational torque, even though number of rotations is reduced in comparison with the above embodiment.

As shown in **Fig.10A**, four notches **140a** are provided. It should be noted, however, that two or three notches may be provided. It is also possible to attach the magnetic material **140** to the rotational output shaft **11** in an eccentric manner in its entirety, without providing notches **140a**.

Fig.11A through **Fig.11H** are illustrative diagrams showing the operation of the second embodiment of the invention when developed into a linear motor type.

According to this embodiment, a movable body **21** is adapted to be moved along a linear track **20** of a roller conveyor type. The track includes a frame on which a plurality of rollers are positioned in parallel relative to one another. A permanent magnet **22** is mounted on the movable body **21**. A magnetic body **23** of a plate-like configuration is fixed to the permanent magnet **22** in the upper surface, so as to form a movable element. It should be noted that magnetic flux from the permanent magnet **22** passes through the magnetic body **23**. A plurality of electromagnets **25a, 25b, 25c, 25d** and so on are disposed above the movable element **24** along the linear track positioned parallel to each other. These electromagnets constitute a stator **25**. Magnetic circuits of the electromagnets **25a, 25b, 25c, 25d**, and so on, are independent from one another, so that the electromagnets are magnetised in a sequential manner by means of excitation change-over mechanism (not shown), so as to have a magnetic polarity opposite to the magnetic pole of the permanent magnet **22**. Power output shafts **21a** are attached to a side surface of the movable body **21**.

Operation of the above second embodiment will be explained below.

As shown in **Fig.11A**, and when no electricity is supplied to the electromagnets, the electromagnets **25a** and **25b** positioned Just above the movable element **24** are subjected to magnetic field of the permanent magnet **22** (refer to shaded portion in **Fig.11A**). Thus, such electromagnets magnetically absorb the magnetic body **23**, so that the movable element **24** remains to be stopped.

As shown in **Fig.11B**, and when the electromagnet **25c**, positioned ahead with respect to the direction in which the movable element **24** moves, is excited, the magnetic field of the permanent magnet **22** and the magnetic field of the electromagnet **25c** interact with each other, so that magnetic flux **23a** passing through the magnetic body **23** converges instantaneously toward the electromagnet **25c**. By this, the movable element **24** is magnetically absorbed to the electromagnet **25c**, so that it is moved along the linear track **20** under the propulsive force acting in the direction in which the width of the magnetic flux **23a** becomes broader, i.e., in the direction of an arrow mark shown in **Fig.11B**.

Fig.11C through **Fig.11E** illustrate a change in width of the magnetic flux **23a** in response to movement of the movable element **24**. At the point at which the width of the magnetic flux **23a** becomes maximised, i.e., when the forward end of the magnetic material **23** of the movable element **24** is positioned just before passing by the electromagnet **25c**, the width of the flux **23** becomes maximised. At this time, magnetic absorption acting between the permanent magnet **22** and the electromagnet **25c** becomes maximised, but the propulsive force acting on the movable element becomes zero.

Before the propulsive force acting on the movable element **24** becomes completely zero, i.e., when the forward end of the magnetic body **23** of the movable element **24** is about to pass the electromagnet **25d**, the excitation changeover mechanism is actuated so as to stop excitation of the electromagnet **25c** and so as to initiate excitation of the electromagnet **25d**. Thus, the magnetic flux **23a** converges to the electromagnet **25d**, as shown in **Fig.11F**, so that a propulsive force acts on the movable element **24**, as in the previous stage.

Subsequently, and in response to further movement of the movable element **24**, the width of the magnetic flux **23a** is reduced as shown in **Fig.11G** and **Fig.11H**, and thus a similar operation will be repeated.

The sequential excitation of the electromagnets, as explained above, causes interaction between the magnetic fields of permanent magnet **22** and electromagnets, whereby a propulsive force is applied to the movable element **24**.

It should be noted that, when the magnetic polarity of the permanent magnet **22** confronting the electromagnets is assumed to be N-pole, the electromagnet **25c** is magnetised solely to be S-pole, so as to prevent formation of a magnetic circuit by virtue of passage of magnetic flux from the electromagnet **25c** through to the adjacent electromagnets **25b** and **25d**, which formation, if it occurs, tends to cause the polarity of the electromagnets to be N-pole identical to the magnetic pole of the permanent magnet **22**. Accordingly, and in a manner similar to that in the first embodiment, no force is generated which tends to interfere with movement of the movable element **24**.

In the present invention, a plurality of electromagnets serving as a stator are so arranged that their respective magnetic circuits become independent from one another. The electromagnets are also arranged so that they are solely magnetised or excited to have a magnetic polarity opposite to the magnetic pole of the confronting permanent magnet. Thus, each electromagnet is prevented from becoming magnetised to the same polarity as that of the permanent magnet, which may occur when magnetic flux from a particular electromagnet passes through to adjacent electromagnets. Accordingly, no force will be exerted which tends to interfere with the intended movement of a rotor or a movable element. As a result, electric energy applied to the electromagnets may be efficiently utilised, while, at the same time, magnetic energy contained in the permanent magnet may-also be efficiently utilised.

The coils constituting the electromagnets are consistently supplied with electric current with the same polarity, without any change, so that heating of coils may be prevented. Further, it is possible to obviate the problems of vibration and noise which might occur due to a repulsive force being generated when polarity of an electric current supplied to the coils is changed.

JOSEPH NEWMAN: ENERGY GENERATOR

Patent WO8300963 **17th March 1983** **Inventor: Joseph W. Newman**

ENERGY GENERATION SYSTEM HAVING HIGHER ENERGY OUTPUT THAN INPUT

This patent covers a device which is claimed to have a greater output power than the input power required to run it.

ABSTRACT

A system for generating obvious work motion, or electromagnetic energy (fields of force) or electric current utilising the electromagnetic energy which makes up a matter and results in a greater output of energy, than the initial input of conventional energy means and teachings. A first exemplary embodiment (**Fig.1**) of the generator uses a contained fluid (**117**) surrounding a series of aligned magnets (**120**); while a second exemplary embodiment (**Fig.3**) uses a special material (**201**) held stationary between two static magnets (**202**, **203**), the special material having its atoms aligned but maintaining the resulting magnetic field at least substantially within its boundary surface; while third and fourth exemplary embodiments (**Fig.5** and **Fig.6**) utilise a relatively heavy coil (**205**) made up of relatively large diameter wire of relatively great length and number of loops and length and a relatively small energising current to drive a rotatable permanent magnet (**200**).

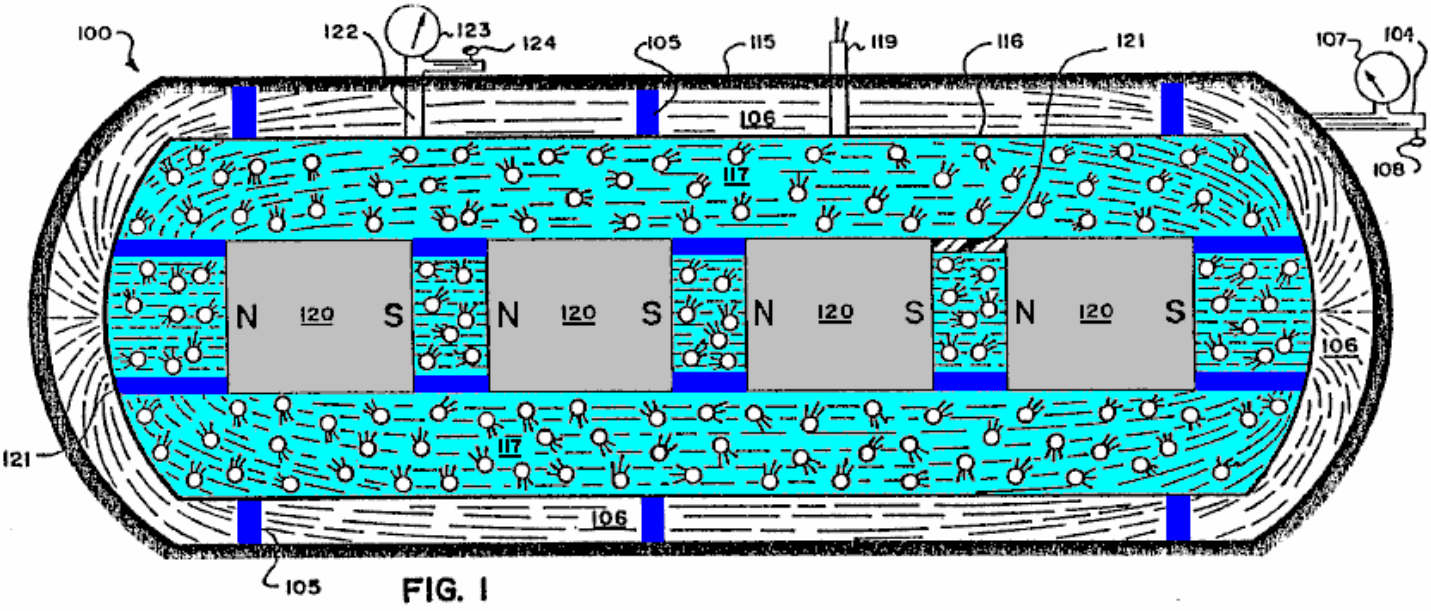


FIG. 1

DESCRIPTION

BACKGROUND OF THE INVENTION

1. Field of the Invention:
The present invention relates generally to devices or systems (including methods) for generating usable energy such as for example electrical energy from electromagnetic fields, electrical energy or electromagnetic fields from matter, and more particularly to devices or systems (including methods) for producing electrical current flow for use as electrical power, and magnetic fields of force which cause motion (obvious work) or electrical current flow or for increasing electromagnetic potential energy available for use or mechanical energy available for use.

2. Prior Art:
There have been many devices proposed over the years for producing electrical-energy, with mechanical friction, thermo-electricity, photoelectricity, piezoelectricity, electrochemistry and electromagnetic induction being the chief forms of primary energy capable of producing electricity. Of these, the only significant source of commercial electrical power has been the mechanical actions of electric generators, and for mobile electric power the chemical action of batteries has been important. Usable motion has resulted from the interactions between the input of electrical energy and the magnetic and/or electromagnetic fields of force (electric motors) and heat or light as a result of input of electrical current through conventional mechanical systems, heaters, lightbulbs, etc.

All of the prior art systems are designed accordingly to rigid mathematical laws taught both in physics and electrical engineering which coincide with the hypothesis rigidly accepted by the industrial and scientific communities concerning the Second Law of Thermodynamics (1850).

From the foregoing generally accepted hypothesis it has also been generally accepted and rigidly taught in physics and electrical engineering that the electric current flowing in a closed circuit from a battery, electric generator, etc.

is used up in the mechanical device being operated by this flow of electric current, and that all such electric current producing systems would only put out at most work equal to the work initially put into the system, or in accordance with generally accepted laws stating that a particular electrical generating system was only capable of a given output of energy and no more.

These beliefs have till this date still remained rigid in both the industrial and scientific communities in spite of proof of Einstein's equation $E=mC^2$ Nuclear reactors convert matter into usable electromagnetic energy in the form of heat, which converts water into steam to turn conventional turbines for production of electric current by conventional electrical generating means. This system is extremely inefficient using less than 1% of the energy of the atom and producing a deluge of contaminated materials which has caused a serious problem as to safe disposal.

Additionally, the basic electrical generators is use throughout the world today utilise the principle of causing relative movement between an electrical conductor (for example a rotor) and a magnetic field produced by a magnet or an electromagnet (for example a stator), all using the generally accepted hypothesis that the greater the relative speed or movement between the two are concerned and the more normal or perpendicular the relative movement of the conductive material to the lines of force of the electromagnetic field, the greater will be the efficiency of the prior art electrical generator. Additionally, all of the prior art systems are based on the generally accepted hypothesis that the greater the electrical conductivity of the material being moved through the field, the more efficient will be the electrical generation.

From the foregoing generally accepted hypotheses, it also has been generally accepted that there should always be movement between, for example, the rotor and stator elements, and that only generally accepted electrical conductors, that is materials with high electrical conductivity, will effectively serve in an electrical generation system.

However, in one of the systems (**Fig.3**) of the present invention, electrical generation can occur with relatively static elements and with materials that are not generally considered to be of high electrical conductivity, although, of course, the present invention likewise can utilise relatively moving elements as well as materials of generally accepted high electrical conductivity, if so desired, as occurs in the systems of the present invention illustrated in **Fig.5** and **Fig.6**.

The prior art has failed to understand certain physical aspects of matter and the makeup of electromagnetic fields, which failure is corrected by the present invention.

BRIEF DESCRIPTION OF DRAWINGS

For a further understanding of the nature and objects of the present invention, reference should be had to the following detailed description, taken in conjunction with the accompanying drawings, in which like parts are given like reference numerals and wherein:

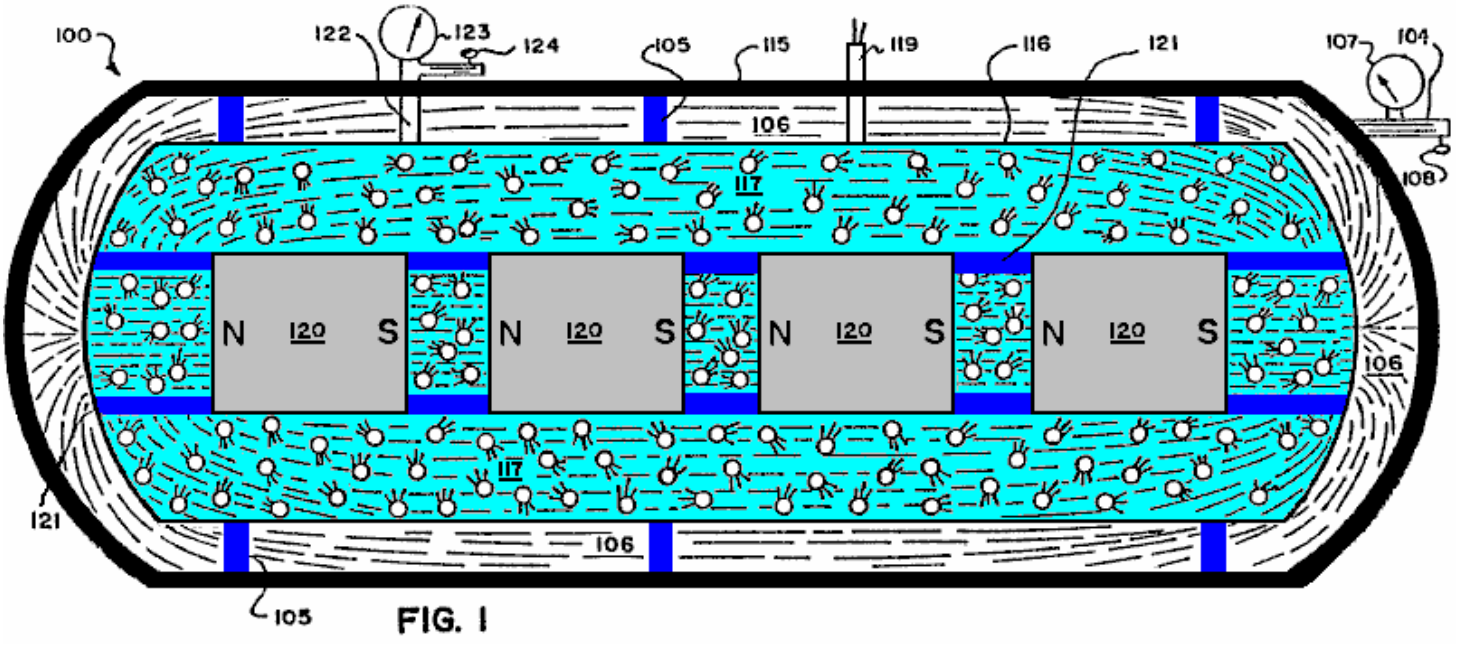


FIG. 1

Fig.1 is a schematic, side view in generalised, representational form of a first embodiment of an electrical generator based on the principles and guidelines of the present invention.

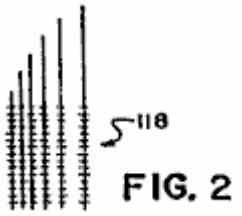


FIG. 2

Fig.2 is a close-up view in general form of an electrical charge pick-up element which can be used in the generator illustrated in **Fig.1**.

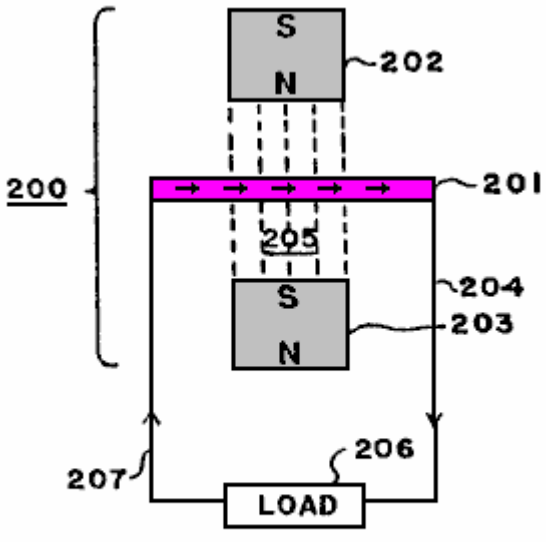


FIG. 3

Fig.3 is a schematic view in generalised, representational form of a second embodiment of an electrical generator based on the principles and guidelines of the present invention.

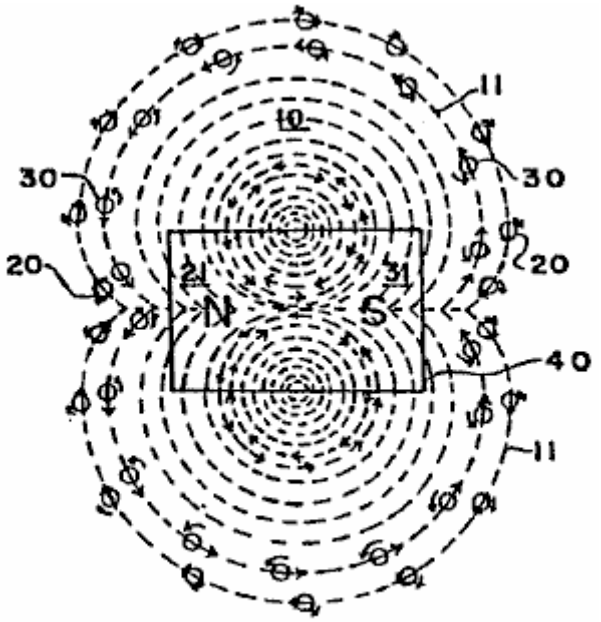


FIG. 4

Fig.4 is a schematic view in generalised, representational form of the negative and positive particles exhibiting gyroscopic actions which emanate from a magnet to form an electromagnetic field.

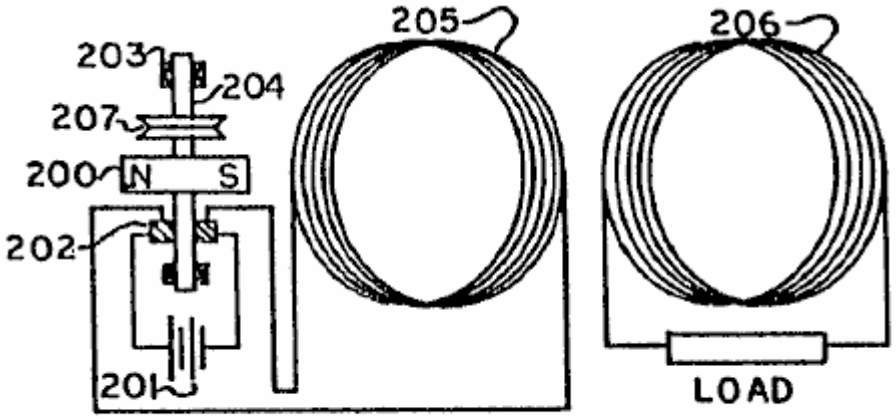


FIG 5

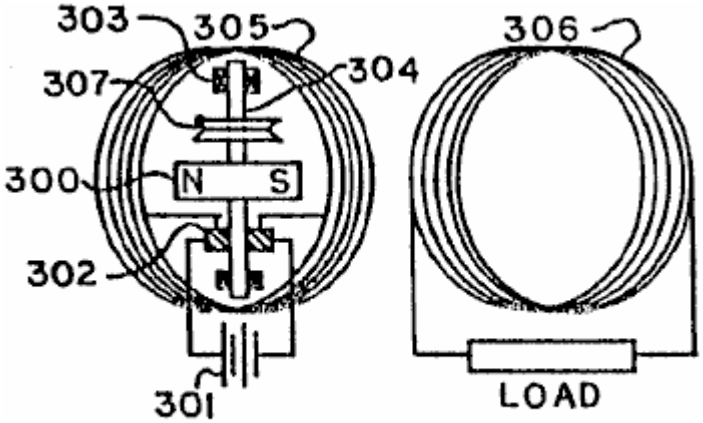


FIG 6

Fig.5 and **Fig.6** are schematic views in generalised, representational form of third and fourth embodiments of a combined electrical generator and motor utilising a static, relatively large coil energised by a relatively low current

driving a rotatable magnet, wherein in the embodiment of **Fig.5** the rotatable magnet is positioned along side of the coil and in the embodiment of **Fig.6** the rotatable magnet is positioned within the open core of the coil.

DETAILED DESCRIPTION OF-PREFERRED EMBODIMENTS:

Basic Principles and Guidelines

In accordance with the principles of the present invention and as generally illustrated in **Fig.3**, an electromagnetic field **10** comprises flows of quanta or particles **20**, **30** of electrical energy flowing from each of the poles **21**, **31** of a magnet (or electromagnet) **40** to the other pole, following the "lines of force" **11** of the electromagnetic field. These particles **20**, **30**, believed to be travelling at the speed of light, are always coming out of one end **21**, **31**, respectively, of the magnet **40** and going into the other pole **31**, **21**, respectively, flowing from a relatively high energy source to a low energy source.

These particles **20**, **30** are, it is believed, negative and positive charges and have a spin producing a gyroscopic motion and follow the mechanical laws of gyroscopic action.

The mass of each of the particles **20**, **30** equals the energy of the particle divided by the speed of light squared. The peripheral speed of the gyroscopic spin of the particles is believed to be the speed of light.

For purposes of illustration only and as a matter of nomenclature, the positive charge particle **20** is going in one direction ('-'N" to "S") with a clockwise spin, and the negative charge particle **30** is going in the opposite direction with a counter-clockwise spin. Of course, if a particle such as **20** or **30** is flipped around one-hundred-and-eighty degrees, it becomes the opposite charge or type of particle.

The electromagnetic field **10** is thus the orderly flow of the positive and negative charges **20**, **30** moving at the speed of light from the north and south poles **21**, **31**, to the south and north poles **31**, **21**, respectively, and follow the paths of what is termed in the art as the "lines of force" **11** of the electromagnetic field **10**.

As is known from the laws of gyroscopes, a gyroscopic particle or body moves at right angles to the direction of an applied force. Therefore, when a force is applied to the electrical energy particles **20**, **30**, they will move at right angles to that force.

It should also be noted from known gyroscopic laws that the electrical energy particles **20**, **30**, when they move with their gyroscopic axis straight into an object, tend to knock that object straight, but, if that object hits the particles at an angle to the axis other than at zero or one-hundred-and-eighty degrees, the particles are moved off at an angle from the straight.

Additionally, it is noted that a magnetic field caused by a current flowing through a wire comes from negative and positive particles, such as **20**, **30**, with a net flow of such particles going in the same direction but with opposite spin.

In the system and method of the present invention, the foregoing principles serve as guidelines in the present invention.

Reference is further had to pages DD23 through DD27 of the Disclosure Document and to page 8, line 26 through page 11, line 23 of the prior application Serial number 25,907 and its Figures 7 - 10.

From the foregoing disclosures, many different devices, structures, and methods are possible to embody the principles and guidelines of the system of the present invention, which will in general utilise a material or substance or structure to place a force at the proper angle to the gyroscopic particles **20**, **30** wherein the particles **20**, **30** follow a path or paths which do not cancel one another out, thereby producing electrical current at appropriate outputs for further use or for increasing available potential electrical energy for ultimate use.

-First Embodiment (**Fig.1**)

One possible, exemplary embodiment using the principles of the system of the present invention is schematically shown in the generalised illustration of **Fig.1**.

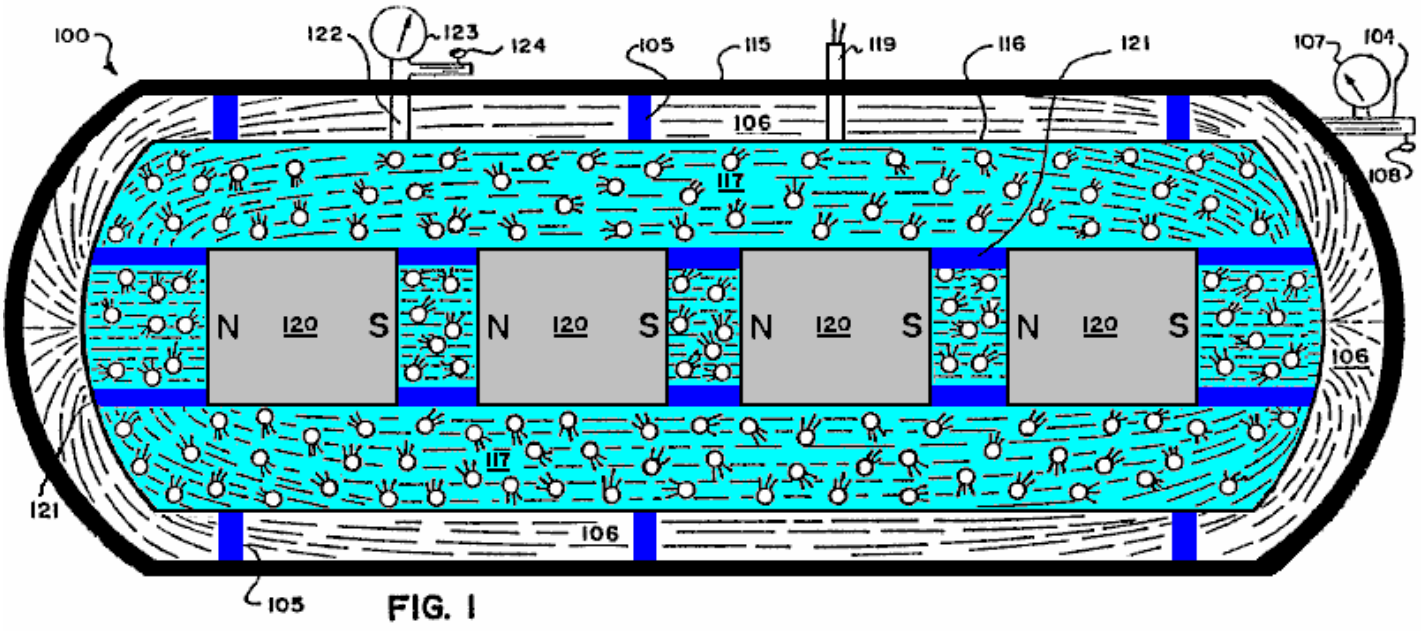


FIG. 1

As illustrated in **Fig.1**, there is provided an electrical current generator **100** comprising an outer keeper housing **115** and an inner, pressure containing, closed housing **116** supported therein by insulating supports **105**. A vacuum exists in the area **106** between the two housings **115**, **116**, which vacuum is regulated and induced by means of the vacuum line **104** with its gauge **107** and its control valve **108**. The outer housing **115** acts as a keeper for magnetic fields of force, and can be made for example of soft iron, while the vacuum in area **106** prevents the leakage or discharge of static electrical charges which might build up on the exterior of the inner housing **116**.

A gas or gas-liquid mixture **117** which may also include solid particles such as for example lead or brass filings, is included within the inner housing **116** surrounding a series of aligned magnets **120** carried by insulating braces or supports **121** and producing a high, combined electromagnetic field. The magnets **120**, which can for example be cryogenic magnets, have their "north" and "south" poles aligned (as illustrated by the "**Ns**" and "**Ss**") so that their magnetic fields reinforce one another.

The level of the gas or gas-liquid mixture **117** in the housing **116** is regulated by means of the line **122** with its gauge **123** and control valve **124**. Electric current output wires **119** are provided and extend down to electrically connect with a wire pick-up system **118** (shown in close-up in **Fig.2**), which can be for example in the form of very small wires forming a closely spaced network or mesh or of a porous conducting metal body or sheet, located in and extended throughout the fluid **117** in the housing **116**.

It is noted that a thimbleful of gas contains a fantastically large number of extremely tiny bodies which are in continuous, random motion moving at extremely high speeds. Hence, the fluid **117** continuously applies a force to the gyroscopic particles (analogous to particles **20**, **30** of **Fig.3**) moving at the speed of light ' in the high electromagnetic field (produced by the magnets **120**) as they continuously collide with each other, which results in the fluid **117** becoming electrically charged. The charged fluid **117** discharges its electrical charge to the pick-up wire network **118** positioned in the fluid, and the electric current so produced and generated is taken off for use via the electrical output wires **119**.

As an alternative to having internally contained magnets **120**, the electromagnetic field needed in the fluid **117** could be produced by a source located outside of the confines of the fluid **117** as long as a significant field was produced within the fluid **117**.

-Second Embodiment (Fig.3)

A further exemplary, generalised embodiment utilising the principles of the system of the present invention is shown in schematic form in **Fig.3**.

The electrical current generator **200** of **Fig.3** comprises an extended member **201** of a special material having its atoms especially aligned to produce electric current when positioned in an electromagnetic field but which does not on its own exhibit any substantial magnetic field outside of its boundary surfaces but substantially contains the field within itself. This is in contrast to "magnetic" materials which likewise have atom alignment but which also exhibit or produce a substantial magnetic field in the area surrounding it.

The generator **200** further comprises for example two magnets **202**, **203**, with their north and south poles facing each other, with the member **201** positioned between them, and with the three elements **201-203** held static with respect to each other. Because of the special nature of the material of the member **201** and its special atom alignment, it will produce a direct current through output line **204** as a result of the gyroscopic actions of the particles of the electromagnetic field **205** produced by the facing magnets **202**, **203**, on the especially aligned atoms in member **201**, which phenomenon occurs even when and even though the member **201** is completely static with respect to the magnets **212**, **203**.

However, it may be desirable in some applications to allow or produce some relative movement between the generator elements **201-203**. The output line **204** extends to an appropriate "load" **206** for using the electrical current generated by the generator **200**. A return line **207** completes the circuit back to the member **201**.

Based on experiments to date, it is believed that brass and lead are materials which can have their atoms especially aligned to interact with the gyroscopic particles (analogous to particles 20, 30) flowing between the magnets 202, 203 and will substantially contain within their surface boundaries the magnetic field produced by the aligned atoms or molecules.

With respect to producing the proper material with atom alignment for the member **201**, it is noted that most materials seem to align their atoms in random directions when formed by conventional methods of production. However, it can be observed that certain materials can be made magnetic by putting the material in an electromagnetic field while cooling from a temperature of around a thousand degrees Centigrade. The magnetism is the result of atom alignment of the material in a given direction (see pages DD19 through DD21 of the Disclosure Document). All materials are affected so as to align parallel or across lines of force when in a powerful electromagnetic-field. Accordingly, if a material while being formed is cooled in an extremely powerful electromagnetic field, the atoms of the material will take a particular alignment. The atom alignment direction could be varied depending on whether the electromagnetic field was aligned with the material or at a ninety degree angle to the material. This would result in the atoms of a material having their particular electromagnetic spin direction primarily along the same axis.

However, merely having atom alignment is not sufficient. Additionally the material for the invention should be such that it exhibits very little if any magnetic field in the area surrounding it. Thus it should be noted that the exterior electromagnetic field that occurs from the atom alignment of the conventional magnet is not duplicated in the material of the invention, because the electromagnetic energy resulting from atom alignment in the material of the invention will be primarily contained within the boundaries of the material It is believed that lead, made superconductive by immersion in a bath of for example liquid helium, is such a special material and could for example serve as the material for member **201**.

This then results in having a material which would place a force at the proper angle on the gyroscopic type particles moving in the electromagnetic field so as to cause an EMF to be produced even when the material was sitting still. (See also first paragraph of page DD23 and paragraphs four, A through E, of page DDl9 of the Disclosure Document).

It is believed that high, contained pressures, as well as other methods, can also probably produce atom alignment as the atoms of a conductor or any material will react to sufficient external force. (See first paragraph of page DD35 of the Disclosure Document). This possibility is also indicated by the fact that hard knocks or impacts will demagnetise a magnet.

The proper procedure of material production in achieving atom alignment with internally contained fields of force will cause the controlled release of electrical energy in electromagnetic fields of force when the material of the invention is placed in the lines of force of the electromagnetic field.

-Third and Fourth Embodiments (Fig.5 and Fig.6)

A. Related Principles

1. Numerous scientific tests and experiments made by the inventor indicate that the magnetic field resulting from an electrical current flowing through a conductor is the result of atom alignment within that conductor at an extremely high speed with an ability to reverse atom alignment just as rapidly without the magnetic hysteresis associated with conventional materials considered "magnetic." Prior to this time it has been believed and taught by the scientific community that the magnetic field associated with an electric current carrying conductor was the result of the electric current itself and not of the conductor material, for example copper, which was considered to be "nonmagnetic." Even the inventor was influenced and mislead by these teachings and attempted to mechanically explain and justify the prior teachings, as is seen on page DD-27 of the Disclosure Document which is an important part of this patent application.

However, as taught in the present invention, what mechanically happens is that the gyroscopic particles making up the electric current moving in a conductor interact with the electromagnetic makeup of the atoms of the conductor, causing them to align extremely rapidly, thereby then releasing some of their electromagnetic make-up in the form of a magnetic field exactly as explained in great detail for conventional magnetic materials in the Disclosure Document.

This is easily proven and understood by taking for example, a size 14-gauge conductor one foot long, winding it into a coil and connecting the coil to a meter and a 1.5 volt battery. The total current registered on the meter will be 1.5 amps and the strength of the magnetic field created from the short conductor will be extremely small. Next, the same type of test is run again but with the length of the conductor increased to for example two thousand feet, but still in a coil. The total current registered on the meter will now be considerably less, but the strength of the magnetic field given off from the conductor will now be extremely large!

This shows that the magnetic field is not from the electric current flow, but is the result of the interactions of the gyroscopic particles which make up the electric current interacting with the atoms of the conductor! This causes the gyroscopic particles of the electric current not to be able to make the circuit back to the battery so quickly, and therefore the meter shows less current used.

The magnetic field is the result of the atom alignment of the conductor. The more atoms in a conductor (up to a point), the stronger the magnetic field produced from a given amount of electric current input. Again, this is proven by changing the diameter of the conducting wires, and, with the lengths being the same, the strongest magnetic field will result from the conductor with the largest diameter. The reason for this is that there are more conducting atoms to interact with the gyroscopic particles of the electric current moving through the conductor, which results in a greater number of conducting atoms being aligned, thereby then releasing some of their electromagnetic make-up, exactly as has been explained in great detail in the Disclosure Document as being possible for all matter.

If the magnetic field produced was strictly based on the amount of current going through a conductor, as taught in the prior art, then the strongest magnetic field would result when current went through a large diameter and short length conductor, because the current flow through the entire circuit is greatest at that time. However, experiments prove that the shorter a conductor is made, the greater the current flow through the entire circuit and the less strength of the magnetic field surrounding that conductor. The longer that same conductor is made (up to a point), the greater the magnetic field surrounding the total mass of the conductor and the less current that makes the complete circuit of the entire system. Reason: more atoms!

2. Numerous scientific tests and experiments made by the inventor also indicate that the magnetic field created when an electric current moves in a conductor does not use up measurable energy when performing obvious or non-obvious work, force or power. This is true no matter how strong or how immense the power of the motor or electromagnets is.

Reason: the magnetic field coming from the conductor is the result of extremely quick atom alignment within that conductor. Therefore the energy in the magnetic field is the energy that makes up the atoms of the conductor! This energy is literally Einstein's equation of E=MC2, and therefore the energy is believed to be moving at the speed of light.

This energy use cannot be measured by today's measuring instruments. This has been explained in great detail in the
Disclosure Document and is believed to be true of all matter!

3. The same is true for the electric current that comes from a conventional battery. The electromagnetic energy coming from the battery is the energy that makes up the atoms of the material of the battery! Again this energy use is not measurable by today's measuring instruments. Electric meters of all types are simply mechanical

devices which measure the amount of electric current that comes into that instrument. They do not measure the amount of mass that has been converted into electromagnetic energy.

Present teachings in science state that the electric energy flowing from a battery is used up in the device operated by that flow of electric current. This is not true at all! The electromagnetic energy released from the atom make-up of a battery has a relatively infinite capacity to do obvious work, force, or power.

This is easily proven even with a small motor and a 1.5 volt battery. With a battery connected to motor to operate it and with a meter to take readings, the motor is then physically stopped from turning by physically holding or restraining the shaft. At that moment the motor is performing no obvious work, force or power, but the meter will register a greater flow of current. The magnets of the motor can be taken out and the reading will still be the same. If the electric current was being used to operate the motor, the meter would register more current when the motor was running.

The electric current not only will operate the motor but, once it flows through the complete circuit back to the battery, it also does additional work based on Faraday's Laws of Electrolysis within the battery itself. What has happened is that the electromagnetic energy released from the atoms of the material of the battery once they have completed the circuit, then take a "short cut" and move large pieces of the mass of one material of the battery over to the other material of the battery. The inventor has stated and shown throughout the Disclosure Document that the effect of gravity was the non-obvious effect of electromagnetic energy. Once the materials of the battery have combined, the extreme desire for the two materials to merge is physically reduced. These materials will attempt this merger anyway possible and, if the electric current initially released from a battery is not allowed by mechanical means to complete the circuit back within itself, the electromagnetic energy then in the mechanical means will perpetually (in a relative, theoretical sense) perform obvious work, force or power. The reason: the force which initiated this flow of current (electromagnetic make-up of atoms of material) is constant, similar to hydraulic pressure, with the noticeable exception that it is moving it is believed at the speed of light and will interact with the electromagnetic make-up of the atoms of other materials, causing them to release some of their electromagnetic make-up in the form of a magnetic field. This then multiples the capacity for doing obvious or non-obvious work, force or power, which can then react with another conducting coil or with the electromagnetic energy within the magnetic field of a conventional magnet and multiply this effect even further, and on and on and on for a relatively unlimited source of energy.

The same is true in not letting the current get back to a conventional generator. If a mechanical means is set up so that the electric current is "trapped," without completing a circuit, the gyroscopic particles of the current have a capacity for continuous work without increasing the power input into the generator system. However, if the circuit is complete and the electric current moving in the system does absolutely no obvious work, power or force, the gyroscopic particles making up the current on getting back to the generator will then increase the need for more power input into the system. Reason: the opposing effect of magnetic fields as defined in Lentz's Law. This law is simply an observation of this effect, which before now has never been fully understood.

4. Numerous scientific tests and experiments made by the inventor also indicate that there is a correlation between the electromagnetic spin orientation of the atoms of non-conductors, semi-conductors, and conductors, and the varying results achieved with an electric current in attempting to move through these materials, or when moving these materials through a magnetic field attempting to induce electric current. The property of resistance to electric current movement is generally speaking the same type factor already explained above for electric current producing a magnetic field when moving in a conductor.

The gyroscopic particles in a moving electric current interact with the atoms of the material through which the current is moving. Each atom can efficiently only interact with sun exact maximum amount of electric current, and, if exceeded, there is an interruption of orderly movement. Then the angle of release of the gyroscopic particles from the atoms are such that the electromagnetic release from those atoms are in the form of heat, exactly as explained in great detail in the Disclosure Document. This effect is easily observed by the fact that resistance decreases relative to an increase of the cross-section of the material. Reason: simply, more atoms within that given area, and, for a fixed input of electric current, there are more atoms to receive and interact efficiently with the gyroparticles making up the electric current.

Again the same is true for resistors designed for deliberately producing heat. Such resistors are not materials which are considered good conductors of electric current. It is stated and shown in great detail in the Disclosure Document that the electromagnetic spin orientation of the atoms of a non-conductor are different from that of conductor atoms, and therefore different results will occur from the same inputs of electromagnetic energy.

This is easily seen by the fact that, in a resistor, for a given amount of electric current input, the heat release increases as the diameter increases. What that means is that the property of resistance has decreased. On a conductor it is just the opposite. If the diameter is increased the resistance is decreased, but so is heat release. Again, this is an indication that the gyroparticles in the electric current movement interact with each atom of the

material. This same effect shows up again in conventional electrical induction from a conductor interacting with a magnetic field. Experiments by the inventor have indicated that the property of conventional induction is the result of the same property of resistance.

If one increases the diameter of a conductor, lengths staying the same, one decreases the amount of electric current produced relative to the total number of atoms within the conductors under consideration. Or, if one takes a given number of wires of the same diameter and length, and moves a magnet across them, the current produced will be considerably less, than if one takes the same diameter wire, but only one wire, and increases its length considerably and then forms it into a coil forming the same number of wires on any one side and then moves the same magnet across only one side of that coil, the electric current generated will then be considerably greater. Reason: the property of resistance. This is the mechanical effect within the gyroscopic electromagnetic make-up and orientation of the atoms of all materials which have the mechanical ability to perform a given task efficiently up to a point concerning input of additional electromagnetic energy and then mechanically causes varying results once this threshold is exceeded.

This and all the other thoughts and innovations in this and the previous disclosures of the previous applications and the Disclosure Document previously put forth show that there are many different mechanical ways to release a relatively unlimited source of energy from electromagnetic energy which makes up all matter and which results from this invention.

B. Working Prototypes

Fig.5 and **Fig.6** illustrate rough, working prototypes of this aspect of the invention. These embodiments are only relatively inefficient prototypes built by hand for the purpose of demonstrating the invention. It should be self-evident that the prototypes, by various mechanical means and designs, can easily be made extremely efficient and the illustrated embodiments are being presented only for general, representational purposes.

As is illustrated in **Fig.5**, there is provided a combined electrical current generator and an electromagnetic motor comprising a rotatably mounted, permanent magnet **200**, a battery **201**, brushes and commutator **202**, bearings **203** and power, mounting shaft **204**, and a first, primary, magnetic producing coil **205** and a second, secondary electric producing coil **206**. The two coils **205**, **206** are juxtaposed together in parallel disposition with concurrent core centre-lines, with the magnet 200 positioned alongside of coil **205** at or near its core centre-line with the rotational axis of the shaft 204 positioned orthogonally to the centre-line.

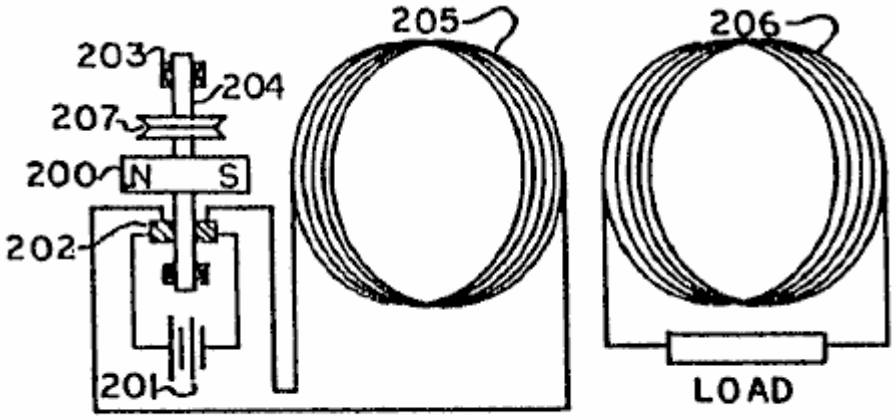


FIG 5

In the prototypes a very small battery **201**, for example, size "N", of 1.5 volts is used. When the circuit is completed, the battery **201** converts an immeasurable amount of its mass into electrical current (gyroscopic particles moving at the speed of light) which goes out through the communicator and brushes **202**, and then enters magnetic producing conductor coil **205** made, for example, from insulated 14-gauge or 15-gauge copper wire, with the total weight of the coil **205** being for example seventy to ninety pounds. This causes the atoms of coil **205** to align extremely fast then releasing some of their electromagnetic make-up (gyroscopic particles) in the form of a magnetic field. This field then interacts with the gyroscopic particles making up the magnetic field coming from the atoms of the material of the permanent magnet **200**.

This causes magnet 200 to attempt to align its magnetic field movement with the magnetic field movement coming from the atoms of coil 205, resulting in rotation of magnet 200 and the shaft 204 to which it is attached. This then changes the position of the commutator and brushes 202 relative to each other's initial positions, which then causes the electric current coming from battery 201 to be going in the opposite direction into coil 205, causing the

atoms of coil 205 to extremely quickly reverse their alignment and the polarity of their magnetic field which they are emitting. The reversed field then interacts again with the magnetic field of permanent magnet **200**, causing it to further rotate.

This process is then continuously repeated, producing continuous rotation of the shaft **204** which can be used as a source of motive power in many different ways. A power belt wheel **207** for example using a continuous "V" belt is illustrated as a general representation of this motive power source for producing useful, obvious work. In a prototype test run with a small 1.5 volt, type "N" battery, the shaft **204** and the magnet **200** - rotated at a high speed for approximately twelve hours before running down. By improving the particular design features of the prototype and by using longer lasting batteries, the rotation time of the shaft **204** can be greatly increased to a theoretical point approaching "perpetual" for all practical purposes. At the same time the alternating magnetic field produced by the coil **205** induces into coil **206** electrical induction, which then causes coil **206** to produce an alternating current across its "load," which current can be made to exceed the conventional output of the battery **201**. The battery source **201** can be replaced when needed.

It is very important to understand that, the longer the length of the conducting wire in coil **205**, the stronger will be the magnetic field produced and the less electric current that will complete the circuit and get back into the battery and destroy the mechanical source of the electrical current. This effect can be increased further by increasing the diameter of the conducting wire in coil **205** and then greatly increasing its length still further in the coil.

Reason: The gyroscopic particles making up the electric current interact with the atoms of coil **205**. The more atoms in coil **205**, relative to it's length, the longer it takes the gyroparticles of the electric current to influence them and exit from the other end of the coil. It is then easily seen that if the direction of the current flowing into coil **205** is then reversed, this then further increases the lag time. Reason: The gyroscopic particles have inertia and are believed to be moving at the speed of light and they are interacting with the gyroscopic particles making up the atoms of the conducting coil **205**. These atoms also have inertia, and when the direction of current in coil **205** is reversed, the incoming current then collides with the current already in coil **205** going in the opposite direction.

This causes a brief hesitation during the time the current already in the coil is being forced to reverse its direction, thereby then reversing the direction of the atoms within coil **205** which have already been influenced to become aligned. This causes a constant force throughout the circuit, but does not allow very much current to get back into the battery **201** to destroy the mechanical means which initiated the release of electric current in the first place.

Therefore, it should be further understood that, the faster the current direction reverses into the coil **205**, the more efficiently the matter of battery **201** is converted into 2 pure electrical energy ($E=MC^2$), without destruction of the mechanical situation that initiates the electrical current release.

It is also important to understand that, the stronger the magnetic field coming from the mass of magnet **200**, the greater will be its rotational speed. Additionally, the greater the magnetic field coming from the mass of coil **205**, the greater will be the rotational speed of magnet **200**, and, up to a point, the greater the electric current input from battery **201**, the greater the rotational speed of magnet **200**.

Reason: the greater the electric current flow into coil **205**, the greater will be the percentage of the atoms making up coil **205** that are aligned. This probably has the same relationship as does achieving atom alignment in conventional magnetic materials. Once complete atom alignment is reached in coil **205**, no amount of current will cause those atoms to increase the strength of the magnetic field emitting from those atoms.

Therefore, it should be clear that, for a given input of electric current from battery **201**, the most efficient design is one in which the most atoms of coil **205** are influenced to atom alignment by that given electric current, which means increasing the diameter and the length of the conducting wire of coil **205** to the point that the strength of the magnetic field produced is sufficient to cause rotation of the magnet **200** to a speed that allows none or at least very little of the electric current which initially comes from the battery **201** to complete the circuit and get back into battery **201** and destroy or reduce the mechanical effect which induced the conversion of the matter of battery **201** in electric current in the first place. Again this desired effect can be increased by increasing the strength of the magnetic field given off by the atoms of the permanent magnet **200**.

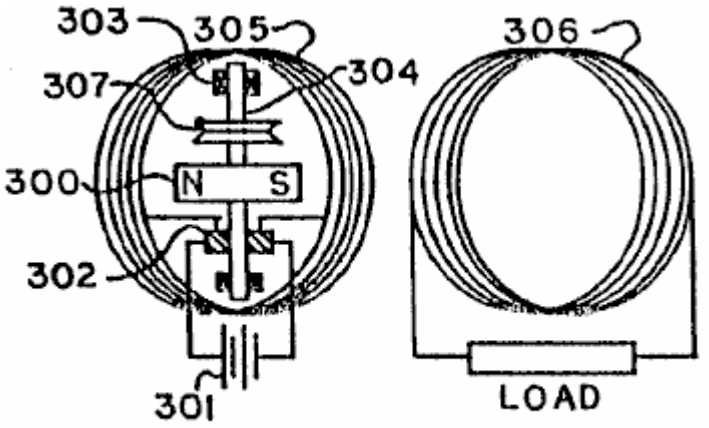


FIG 6

In the second prototype embodiment of **Fig.6**, the structure and operation of the prototype is substantially identical to that of **Fig.5** with the major exception being that the magnet **300**/shaft **304** elements (and related sub-elements **302**, **303** and **307**) are positioned inside of and within the core of the primary coil **305**, as compared to the placement of the magnet **200**/shaft **204** elements next to and along side of the coil **205** of **Fig.5**. Therefore, for brevity, a detailed description of the elements of **Fig.6** will not be repeated, but it is noted that the corresponding and analogous elements and sub-elements are similarly numbered in **Fig.5** and **Fig.6**.

It is also important to again stress the fact that the prototype designs shown are presented simply to prove the correctness of the invention, and it should be clear that the invention can be made extremely more efficient by utilising all of the magnetic field produced by coil **205** and designing the magnet **200** of a shape and strength that efficiently interacts with the majority of the magnetic fields from coil **205**. The illustrated prototypes is relatively highly inefficient in this regard, but even so, the results of the invention itself greatly exceed the prior art as to use of electric current from whatever source and interaction with an electric motor or whatever work was conventionally performed.

The applicant feels it is very important to again stress, in building many varying designs of this invention, consideration must be given to the fact that the Energy in the field of force of any type magnet is the Energy that makes up the Atoms of the material from which it comes! This Energy is a real Entity with, it is believed, a gyroscopic action. It is literally Einstein's Equation of $E=MC^2$ and it is believed that this Energy moves at the speed of light and makes up all Matter. And that this Energy has a constant pressure effect back to the Atoms of the material from which it came, similar to hydraulic pressure. This effect is additionally more fully understood by stating the following results obtained from experimentation by the applicant in the process of this invention.

a) When the system is initially attached to a 1.5 volt size N Battery **201** or **301** and the magnet **200** or **300** and related rotation entities are placed close to or in the centre of coil **205** or **305**, the following results are observed:

If the electric current produced in coil **206** (**306**) is then fed back into coil **205** (**305**) in accordance with proper polarity, the rotation speed of magnet **200** or **300** will then accelerate. If fed back into coil **205** (**305**) in wrong polarity, the rotation speed of magnet **200** (**300**) will slow down.

This proves that the total force from coil **205** (**305**) interacting with the magnet **200** (**300**) is greater when the electrical energy from coil **206**(**306**) is fed back into coil **205** (**305**), then when only the initial electric energy from battery **201** (**301**) is fed into coil **205** (**305**)! When two or three batteries are electrically connected together in series, so as to create for example three or four and a half volts of electrical input, this effect is multiplied. Remember, up to a point, the greater the electrical input, the greater the percentage of atom alignment within coil **205** (**305**).

This further proves that the electric current produced in coil **206** (**306**) is a result of the gyroscopic particles of Energy released from the magnetic fields which came from the Electromagnetic make-up of the atoms of coil **205** (**305**), and is not part of the initial Electrical Energy released from the atoms making up the materials of battery **201** (**301**)! The coil **206** (**306**) can be taken out of the system, or its electrical current fed away from the system, and the rotational speed of the magnet **200** (**300**) will not observably change. However, the rotational speed of magnet **200** (**300**) will noticeably change when the electric current from coil **206** (**306**) is fed back into coil **205** (**305**)!

Now a different result:

b) When the electric current from battery **201** (**301**) becomes weaker to the point that the magnetic field coming from coil **205** (**305**) has weakened and shrunk allowing the magnetic field of the rotating magnet **200** (**300**) to expand and then noticeably induce electric current into coil **206** (**306**) and into coil **205** (**305**), then reverse results are observed. When the magnetic field from the coil **205** (**305**) is large, then the magnetic field from magnet **200** (**300**) is retained! If coil **206** (**306**) is then short circuited, the rotation of magnet **200** (**300**) will noticeably slow down.

If electric current from coil **206** (**306**) is fed back into coil **205** (**305**) in wrong polarity, the rotation of the magnet **200** (**300**) will stop. If fed back into coil **205** (**305**) in correct polarity, the rotation of the magnet **200** (**300**) will slow down. At that point, the rotation of the magnet **200** (**300**) will not accelerate, no matter how connected!

These results show that, at this time, the magnetic field from magnet **200** (**300**) noticeably induces a current in coils **206** (**306**) and **205** (**305**) which opposes the rotation of the magnet **200** (**300**). This effect has already been mechanically explained, and it has been shown that Lenz's Law was simply an observation of that mechanical explanation. These results further demonstrate that the expanding and collapsing magnetic fields from coil **205** (**305**) and **206** (**306**) do not noticeably effect each other detrimentally.

Because the resulting magnetic fields from all the coils are the results of fluctuating atom alignment within the coils! Remember, the gyroscopic energy particles making up the magnetic fields have a hydraulic pressure effect back to the atoms from which they came. Also remember that the atoms making up the material of the permanent magnet **200** (**300**) are stationary as to atom alignment direction! Therefore, the pressure effect resulting from an opposing field which the magnet **200** (**300**) induced, is immediate. As is Hydraulic Pressure.

However, the magnetic field emitted from the atoms of coil **205** (**305**) relative to induction into the atoms of coil **206** (**306**) are fluctuating and out of step, so to speak, and therefore, in harmony with each other. The pressure effect from the induction of coil **205** (**305**) into coil **206** (**306**) is an action and reaction effect which reinforces the flipping action of the atoms of coil **205** (**305**) and back into the atoms of coil **206** (**306**).

This action is again seen when the invention is hooked into one-hundred-fifteen volt alternating current, and battery **201** (**301**) is not used. The magnet **200** (**300**) will not rotate even though the magnetic field from coil **205** (**305**) is strong and is alternating. Reason: The fluctuating magnetic field is so fast, that the inertia mass of magnet **200** (**300**) can not get started in one direction before the magnetic field from coil **205** (**306**) has reversed, thereby, causing magnet **200** (**300**) to vibrate only microscopically at sixty cycles per second. And, if a sixty watt bulb is hooked into the system of coil **205** (**306**), it will only light dimly. And there is a lag time of two to three seconds before it lights even dimly.

If then coil **206** (**306**) is hooked to a meter, there is a reading of forty-nine volts, and if the meter is replaced by another sixty watt bulb it will light only extremely dimly. However, the sixty watt bulb hooked to coil **205** (**305**) will now become noticeably brighter! This again shows that the action and reaction results of the atoms of the coils are not noticeably detrimental to each other. Because of the lag time (out of step, so to speak), resulting in reinforcing the flipping atom alignment of the coils.

From this further explanation of the invention it is seen that desirable results may be obtained by the following:

For example, in **Fig.6** the magnet **300** may be of a design and/or be located at a distance from the inside diameter of coil **305** and coil **306**, whereby the majority of the magnetic field from the magnet **300** does not cut the conducting loops of coil **305** or **306**. Yet the alternating magnetic field produced by coil **305** should efficiently have the majority of its gyroscopic particles interacting with the majority of the gyroscopic particles making up the magnetic field of the permanent magnet **300**, but not directly reacting with the atoms making up coil **305**, or magnet **300**!

When the magnetic lines of force of the magnet **200** (**300**) cross at right angles with the conducting wires of coil **205** (**305**), **206** (**306**), a braking action is incurred. It should be noted that, as the inner diameter of coil **205** (**305**) increases, the percentage of time of braking effect decreases.

Along this same line of instruction, the commutator segments **202** (**302**) can be made of a large diameter and the area of brushes made small, whereby, when the brushes cross over the gaps in the commutator segments, there will be no short circuit at any time directly back to the battery **201** (**301**).

By combining the slip rings and brushes (the slip rings can be made of a small diameter) to the side or sides of the brushes and commutator segments **202** (**302**), then battery **201** (**301**) does not have to rotate with magnet **200** (**300**).

The 14-gauge and 15-gauge insulated copper wire weighing seventy and ninety pounds respectively (31.5 kilograms and 40.5 kilograms) used for the motor coil **205** (**305**) and the generator coil **206** (**306**), respectively, in the first hand-made prototypes of the embodiments of **Fig.5** and **Fig.6**, for demonstration purposes only, come in standard buckets of varying weights from wholesale outlets.

It was then wound in coils as shown, and, as taught, the more conducting wire used, the better the results. The magnets **200** and **300** were each initially about a 2.5 inch (6.25 centimetre) cube and can be any size and strength desired.

In a further, rough, hand-built, demonstration, working prototype of the invention of the type illustrated in **Fig.6**, the primary or motor coil **305** was made of 5-gauge copper wire in a single, continuous wire, weighing approximately 4,100 pounds (1,845 kilograms) with a coil loop diameter of 4.5 feet (135 centimetres), while the secondary or generator coil **306** was made of 24-gauge copper wire in a single continuous wire weighing approximately 300 pounds (135 kilograms) with the same, approximate coil loop diameter of 4.5 feet (135 centimetres), with both coils **305**, **306** coincidentally forming a cylinder of approximately 30 inches (75 centimetres) in length. The coils **305**, **306** were built around a cylindrical, fiberglass core body of approximately 200 pounds (90 kilograms) having a vertical, longitudinal centre-line axis.

The rotating magnet **300** was made up of six, separate, parallel cylindrical magnetic columns spaced and disposed about the periphery of a hollow cylindrical fiberglass surface of approximately twenty inch (fifty centimetres) in diameter. Each column was 30 inches 75 cm.) long and was composed of a stack of 70, individual ceramic ring magnets in disc form as made by Jobmaster Magnets of Randallstown, Maryland, 21133, U.S.A. Each disc had a thickness of seven-sixteenths of an inch (1.09375 centimetres), an inner diameter of 1 inch (2.5 cm.) and an outer diameter of 4 inches (10 cm.). The discs were stacked and secured together in 4 inch (10 cm.) diameter fiberglass tubes longitudinally mounted on the inner surface of the twenty inch (fifty cm.) diameter fiberglass cylinder.

The composite magnet **300** had a total weight of approximately 400 pounds (180 kilograms) and a total length of 30 inches (75 cm.) and an approximate diameter of 20 inches (50 cm.).

The magnet **300** was mounted for rotation on a horizontal shaft **304** extending across the hollow core of the coils **305**, **306** crossing through the centre point of the longitudinal centre-line of the cylinder and orthogonally to the longitudinal centre-line of the magnet **300** for rotation within the open centre area of the cylindrically disposed coils **305**, **306** with the longitudinal centre-lines of the coils being vertically disposed.

With a D.C. battery source **301** of two 12 volt lantern batteries and seventeen 6 volt lantern batteries all in series (totalling 126 volts), a measured voltage of 126 volts and a measured current of 99 milliamps in the primary coil **305** were noted. Concurrently a voltage reading of 640 volts and an amperage measurement in excess of 20 milliamps were noted in the secondary or generating coil **306**, with the magnet **300** rotating at a speed of 120 revolutions per minute (rpm). Thus the system was outputting and producing in the generating coil **306** usable electrical energy in excess of 102% of that being inputted in the motor coil **305**! This excess useful electrical energy, of course, is in addition to the further useful mechanical energy available at the exemplary drive take-off **307** on the rotating shaft **304**, on which the 400 pound, 30 inch long magnet **300** was rotating at 120 rpm!

Thus the invention, by utilising the energy of the gyroscopic particles in the magnetic field, produces a greater energy output than the energy input into the system, thus producing results beyond presently accepted scientific teachings of the world.

This prototype achieves exactly what has already been described in great detail in applicant's prior patent applications. There was simply used in this prototype a stronger magnet and a larger diameter conducting wire of great length, that has a considerably greater number of atoms aligned when current is put into the system, and used a greater number of atoms in the generator coil of fine diameter conducting wire.

While the results of the energy released from this particular prototype is highly impressive to others, the applicant still has only scratched the surface of the energy that can be released using the principles of the present invention.

Again, as has already been stressed, the most efficient design, is one in which the least amount of input of current causes the greatest amount of atom alignment.

These data do not constitute any departure from applicant's previous work, but is only to further document that which has already been stressed in the prior patent applications.

Varying the D.C. voltage for the battery source **301** shows that obvious efficiency will continue to rise as the voltage input goes up! Also, the leverage factor advantage of the invention, combined with the inertia of the 400 pound magnet **300** rotating at 120 rpm (even while causing the electrical generator to put out over 100% of energy input) proves the invention to be greatly over 100% efficient even at this slow rpm.

It is contemplated that the next prototype will use super-conducting type material for the coil **305** with a magnet **300** having a magnetic field strength comparable to that of crycogenic-type magnet relative to percentage of atom alignment or size. This will result in the size of the device being much smaller and yet with the available work output being much greater than the prototype just described. Reason: The most efficient type design is one whereby the least amount of current input into the motor coil produces the greatest atom alignment of said motor coil and having rotatable magnet also comparable in strength, relative to size.

The invention can be made without using the coil **206** (**306**) and producing just useful mechanical energy.

Coil **206** (**306**) can be merged or wound with coil **205** (**305**).

The magnet **200** (**300**) can be an electromagnet, a permanent magnet, a cryogenic magnet or any magnet.

The design of magnet **200** (**300**) can create a strong but retained magnetic field.

The design of coil **205** (**305**) can be used to further retain the magnetic field of magnet **200** (**300**).

Alternating current (A.C.) can be used in place of the direct current (D.C.) battery **201** (**301**), if the magnet **200** (**300**) is designed accordingly.

The coils **205** (**305**) and **206** (**306**) may be made up of several coils rather than a single coil.

The magnet **200** (**300**) may be made up of several individual magnets rather than from just a single magnet.

From the foregoing it should be understood that, unlike the teachings of the prior art, the following is desired in the design of the coil **205/305** under the principles of the present invention:
a) Current initially flowing into and through the coil should be small compared to the energy output of the system;
b) A relatively large diameter wire or its equivalent is used for the coil;
c) A relatively large number of coil loops or coils is used;
d) A relative long, continuous length of coil wire or its equivalent is used; and
e) The greatest magnetism for a given mass of the magnet **200**/**300** is desired but may be designed so that the magnetic lines of force will not cut the coils at a right angle.

The present invention applies to any mechanical device which is operated by electrical energy. In accordance with the principles of the present invention, the mechanical device should be designed wherein the electric current as much as is feasible cannot get back to its source, but the circuit is completed whereby the "pressure force" is constant throughout the system.

What has been invented, built and disclosed is an invention of immense importance to the well-being of the entire world. There will be many devices built from what has been shown and taught. It should now be known that all matter is made up of electromagnetic energy and that there are many mechanical ways to release this energy, as has been stated throughout the five prior, related patent applications hereof and the Disclosure Document. All of these future developments will be as a result of the present invention which - releases energy above and beyond conventional energy release mechanisms, prior to this invention.

Some of the basic approaches of the invention are outlined below:

I. Any device which utilises a means by which the electric current (electromagnetic energy) is retained within a member or members outside of the source of said original electric current and then, as a result thereof, is capable of producing a continuous electromagnetic motion or current if so desired beyond present scientific teachings.

2. Any device which releases the electromagnetic energy make up of matter to such an impressive degree as does this invention that it defies several of the present accepted laws of physics and electrical engineering as of this time.

3. That the energy release is noticeably higher and in some cases more controllable than the conventional means of energy release of this time.

Because many varying and different embodiments may be made within the scope of the inventive concept taught here, and because many modifications may be made in the embodiments detailed here in accordance with the descriptive requirements of the law, it is to be understood that the details given above are to be interpreted as illustrative and not in any limiting sense.

CLAIMS

1. A usable energy generation system, comprising:

usable energy output means for making available for use the usable energy generated in the system; and
usable energy generation means associated with said output means and designed to take into account the reaction to a force of the gyroscopic type energy particles and to utilise the gyroscopic type energy particles moving in a magnetic field for producing usable energy of an amount greater than the amount of energy input.

2. The system of Claim 1, wherein said generation means includes structural means for placing a force at an angle to the gyroscopic particles causing the particles to follow paths having a net directional effect, producing electric current flow.

3. The system of Claim 2, wherein said structural means comprises magnetic means and a closed housing associated therewith containing a fluid in the magnetic field produced by said magnetic means, said fluid becoming charged as a result of its interaction with the gyroscopic type energy particles making up said magnetic field.

4. The system of Claim 3, wherein said magnetic means is a series of aligned magnets positioned centrally within said housing but electrically insulated therefrom.

5. The system of Claim 3, wherein there is included a further, keeper housing completely surrounding said closed housing and electrically insulated therefrom, said keeper housing tending to keep and concentrate the magnetic field produced by said magnetic fields within it.

6. The system of Claim 3, wherein said output means includes a network of metallic surfaces immersed in said fluid to pick up the electrical charges on said fluid.

7. The system of Claim 2, wherein said structural means comprises a member having its atoms aligned to produce a net magnetic field which is at least substantially contained within the surface boundaries of said member.

8. The system of Claim 7, wherein said member is positioned in operative association to at least one magnet, and said member and said magnet are held static with respect to one another.

9. The system of Claim 1, wherein:
The usable energy generation system comprises an electrical energy generation system; said usable energy output means comprises an electrical power output means; and said usable energy generation means comprises electrical energy generation means.

10. The system of Claim 1, wherein:
the usable energy generation system comprises usable motion generation system; said usable energy output means comprises usable motion output means; and said usable energy generation means comprises usable motion generation means.

11. The system of Claim 1, wherein said generation means includes:
a magnetic device;
a source of electrical energy;
complete electrical circuit means between said magnetic device and said source of electrical energy for producing an alternating electrical current potential; and
current retarding means for retarding the flow of current through said device back to said source to the greatest extent practical, producing a relatively small and preferably negligible current flow through said source.

12. The system of Claim 11, wherein said magnetic device includes at least one relatively large coil of wire having a relatively large number of turns of wire of a relatively large diameter and a relatively great length.

13. The method of producing usable energy utilising a magnetic field system, comprising the following steps:
a. providing a structure interacting with a magnetic field; and

b. arranging said structure to utilise the energy of the gyroscopic type particles in the magnetic field to generate an electrical current in said structure, or usable motion from said system, or both, and results in producing a greater energy output than energy input into the system.

14. The method of Claim 13, wherein there is included in step "b" the further step of arranging said structure to place a force at an angle to the gyroscopic particles cause the particles to follow paths having a net directional effect, producing electric current flow, without any normal, visible movement taking place in the system.

15. The method of increasing the availability of usable electrical energy or usable motion, or both, comprising the steps of:

a. providing a magnetic device for producing usable electrical energy or usable motion, which device includes a material through which electrical current can interact producing a magnetic field which interacts with a separate mass having a magnetic field, and further providing a source of electrical energy such as for example a battery, generator, or any other;

b. providing a complete electrical circuit between said magnetic device and said source of electrical energy and producing from said source to said device an alternating electrical current potential; and

c. retarding the flow of current through said device back to said source to the greatest extent practical, producing a relatively small and preferably negligible current flow through said source and resulting in electrical energy output, or usable motion output, being a greater energy output than energy input into the device.

16. The method of Claim 15, wherein step "c" is achieved at least in part by the step of providing in said device a relatively large coil or coils of wire having a relatively large number of turns of wire of a relatively large diameter and a relatively great length.

17. The method of Claim 15, wherein step "c" is achieved at least in part by the step of utilising a means by which the electric current is retained within at least one member outside of the source of said original electric current and then, as a result thereof, is capable of producing a continuous electromagnetic motion or current.

18. The method of claim 15, wherein there is included the step of providing a separate magnetic source positioned so that its magnetic lines of force avoid significantly cutting the material through which the electrical energy flows avoiding a braking effect which would retard the desired motion of said magnetic source.

19. The method of Claim 15, wherein step "a" is achieved by said material being a super conducting material and said separate magnetic mass is at least equivalent to a cryogenic magnet.

20. The method of Claim 15, wherein step "a" is achieved by the step of having said material a conducting material and said separate magnetic mass of any desired configuration or strength or type.

21. The method of increasing the availability of usable electrical energy, or usable motion, or both, comprising the steps of:

a. providing a magnetic device which has a material mass into which an electrical current is introduced, by any desired means, which results in causing pertinent atom alignment, within said material mass, thereby releasing some of the electromagnetic energy making up the atoms of said material mass, in the form of a magnetic field, which then causes the gyroscopic type energy particles of said magnetic field to then interact with the gyroscopic type energy particles making up a magnetic field coming from the atoms of a different material mass; and

b. having the magnetic device then cause a release of electrical current or usable motion or both through at least one power outlet and resulting in producing a greater energy output than energy input into the device.

22. The method of Claim 21, wherein the material mass or masses are made of a material or substance that allows for extremely fast atom alignment, without the delay, or conventional degree of hysteresis losses normally associated with conventional iron atom alignment.

23. A device which increases the availability of usable electrical energy or usable motion, or both, by causing the atoms of a material or materials to release some of their magnetic energy makeup in the form of a magnetic field, consisting of gyroscopic type energy particles which make up the atoms of the material from which the magnetic field comes; and

a properly designed mechanism, or power output arrangement being place to utilise the energy of said gyroscopic type energy particles, causing a release of energy output greater than energy input without producing radioactive material.

24. A device which increases the availability of usable electrical energy or usable motion, or both, from a given mass or masses by a device causing a controlled release of, or reaction to, the gyroscopic type energy

particles making up or coming from the atoms of the mass or masses, which in turn, by any properly designed system, causes an energy output greater than the energy input.

25. A system including an energy generator, motor, etc.

of any design or mechanism that takes into account- the reaction to a force of the gyroscopic type energy particles moving in a magnetic field at tremendous speeds which releases greater output than energy input.

26. The system of Claim 25, wherein a small input of electrical current into the mechanism causes extremely quick and high atom alignment, resulting from using a super conducting material, thereby creating a powerful magnetic field, whereby its gyroscopic type energy particles then interact with the gyroscopic type energy particles coming from a second powerful magnetic field and results in producing a greater output of energy than input of energy into the mechanism.

27. The system of Claim 25, wherein a structure is arranged, whereby, there is, or will be, a pressure, or force, exerted on or in said structure, thereby causing the atoms of said structure to react to said pressure or force, and as a result take a pertinent atom alignment direction that results in said atoms of said structure then causing the gyroscopic type energy particles moving in the magnetic field to be generally deflected in the same direction through said structure, which results in usable electric current flow, producing a greater energy output, than energy input into the system.

28. The system of Claim 25, wherein a rotary magnetic mass is designed to react to a reversing magnetic field of another mass, and said reversing magnetic field can not reverse any faster than the atoms of said other mass can flip and realign; said rotary magnet mass being made as long as is practical to adjust to this requirement, wherein the distance of the arc of circle travelled by the ends of said rotary magnet mass is great; great leverage from said reversible magnetic field of other mass being applied to said rotary magnet, and in addition the increased distance of arc travelled by the ends of rotary magnet before the magnetic field of said other mass reverses, greatly increasing the time in which a maximum force is exerted by the gyroscopic type energy particles moving in the magnetic field coming from the maximum number of atoms aligned in said other mass, thereby causing a longer time of acceleration of said rotatable magnet mass before the atoms of said other mass are required to reverse.

29. The method of producing usable energy, comprising the following steps:
- **a.** imputing energy into a device from an external source;
- **b.** having electrical current flow within said device; and
- **c.** utilising the internal electromagnetic energy of at least some of the matter in the device to add to the energy being imputed into the device from the external source to produce useful energy for use outside of the device having an amount greater than the energy being imputed to the device.

PHILIP BRODY: FERROELECTRIC CERAMIC DEVICES

Patent US 4,041,465 **27th September 1977** **Inventor: Philip S. Brody**

Assignee: The Unites States of America as represented by the Secretary of the Army

FERROELECTRIC CERAMIC DEVICES

This version of the patent has been re-worded in an attempt to make it easier to read and understand. The original can be examined at www.freepatentsonline.com and downloaded without charge. This patent covers several different applications, namely; a high-voltage very high-efficiency solar electric device, a photovoltaic memory device, an optical display device and a high-voltage battery, to name just a few. It should be noted that this patent is assigned to the US Army. In my opinion, that lends weight and credibility to this patent. It is claimed that a one centimetre square piece of this material can produce 1,500 volts as opposed to less than one volt using conventional solar cell materials.

ABSTRACT

A method and apparatus is disclosed by which high voltage and current can produced by a polycrystalline ferroelectric ceramic material in response to incident light. Numerous applications of the ferroelectric ceramic material taking advantage of such properties thereof are further disclosed. The polycrystalline ferroelectric ceramic material is initially poled by the application of a pulse of voltage of predetermined magnitude and direction. After being poled in such fashion, light shining on the various surfaces of the ferroelectric ceramic material will generate a consistent high voltage between the surfaces of the ferroelectric ceramic material. If electrodes are attached to the material, then a current will be generated and a load can then be powered by it. Importantly, the magnitude of the voltage produced by the light is directly proportional to the remanent polarisation of the ferroelectric ceramic material, and is further directly proportional to the length of the material, the polarity of the high voltage being dependent upon the polarity of the remanent polarisation and being capable of being reversed when the remanent polarisation is reversed. The open circuit voltages produced by the ferroelectric ceramic material are orders of magnitude higher than those which typically have been produced in the prior-art through the utilisation of standard photovoltaic materials.

DESCRIPTION

BACKGROUND OF THE INVENTION

This invention generally relates to solid state devices exhibiting photovoltaic effects and is particularly directed to the provision of a device consisting of a class of polycrystalline ferroelectric ceramic materials which have been discovered to produce voltages upon the application of light. These voltages are many orders of magnitude higher than voltages typically produced by conventional photovoltaic materials.

Initially, and as background, the instant inventive apparatus and techniques to be discussed below are to be clearly distinguished from the photovoltaic effect now know in the prior-art in that the mechanism for the effect to be discussed herein appears to be unique and different from photovoltaic mechanisms previously described.

SUMMARY OF THE INVENTION

It is the primary objective of the instant invention to provide a device and technique by which extremely high voltage can be generated utilising a solid state polycrystalline class of materials upon the application to such materials of incident light, the voltage generated exhibiting properties entirely unlike the well-known photovoltaic effect of the prior art and of orders of magnitude higher than voltages previously obtainable.

Another equally important objective of the instant invention is the provision of apparatus utilising ferroelectric ceramic materials of the type to be described below, such apparatus taking advantage of the unique properties as discovered to be existent in the class of materials to which the instant invention relates.

These broad objectives, as well as others which will become apparent as the following description proceeds, are implemented by the subject invention which utilises at its heart a class of materials known as ferroelectric

ceramics, and which take advantage of the unique photovoltaic properties discovered to be existent in such class of materials.

Specifically, by illuminating the surfaces of these materials, a steady voltage results across conducting electrodes placed in contact therewith. Currents can then be drawn through loads placed across these electrodes. It has been discovered that an arrangement of an initially polarised ceramic material with electrodes attached thereto as is shown in **Fig.1** of the application drawings produces steady high voltages from a steady illuminating source such as the sun, an incandescent bulb, a fluorescent tube, etc. and that the magnitude of these voltages is high and directly proportional to the length, **l** of the sheet of material provided. In **Fig.1**, the shaded area represents an electrode, and $\mathbf{P_r}$ is the remanent polarisation. In another basic arrangement of the invention, light enters through transparent electrodes and the material is poled in the direction of the light, and the photo-emf up to a certain limiting thickness is proportional to the thickness of the slab.

It has further been discovered that the magnitude of the photo-voltages produced is directly proportional to the remanent polarisation of the material. The polarity of the photo-voltage is dependent on the polarity of the remanent polarisation and reverses when the remanent polarisation is reversed. The magnitude of the voltages that are produced can also be varied by varying the sizes of the grains of which the ceramic is composed, the voltage having a generally proportional relation to the number of grains per unit length. Grain size can be controlled by well-known fabrication techniques involving compositional additives and firing rates, which techniques do not form a part of the present disclosure.

When illuminated at intensity levels such as that produced by direct sunlight or at lesser levels such as that produced by a fluorescent lamp, the materials will behave as voltage sources in series with a high output resistance. The output resistance will decrease an the intensity of illumination increases and also varies with wavelength.

The open circuit voltages produced by the materials of the invention are much higher than those that are typical of other photovoltaic materials. These high open circuit photo-voltages have been observed to some extent in virtually all materials examined which can generally be described or classified by the term ferroelectric ceramic, provided that the material was characterised by a net remanent polarisation. Such high photo-voltages are to be expected in virtually all polarised ferroelectric ceramic materials properly doped, the class including thousands of different known materials of this kind with numerous variations possible in each kind. Such variations are produced by additives, varying grain size, and by changing compositional blends, in those formed from mixtures. Any of these are expected to have application as photovoltaic materials.

From the viewpoint of application, the novel photovoltaic effect seen in ferroelectrics in accordance with the teachings herein differs in two important respects from the well known junction photovoltaic effect which is the mechanism in prior-art devices such as solar cells, and photo-diodes.

First, the prior-art junction photo-emf is independent of the length or thickness of the unit and is low, less than one volt. To obtain high voltages, many cells have to be connected in series. The photovoltaic effect in ferroelectrics, on the other hand, can be used to directly produce high voltages. The photo-emf is proportional to length, and the photo-emf per unit length can be very high. For example, the composition $Pb(Zr_{65}Ti_{35})O_3$ with 7% of the lead substituted by lanthanum, when composed of 2-4 microns grains produces, when illuminated as shown in **Fig.1**, 1500 volts for every centimetre of length between the electrodes. A single one cm square unit thus directly produces 1500 volts.

In this case, it is also clear that the voltage per unit length will be further increased by the development of a composition in which the average grain size is further decreased.

These voltages are so high that applications have been contemplated which are alternatives to the devices presently used for the generation of extremely high DC voltages at low currents -- such as belt machines (the Van de Graaf), in which high voltages are produced by mechanically moving electric charges.

Second, and perhaps even more important, is the fact that the direction of the photo-current and photo-voltage can be reversed simply by reversing the direction of its remanent polarisation. The magnitude of these quantities can be changed by changing that of the remanent polarisation, which in turn can be done (for example) by applying the proper polarity electrical voltage (poling voltage) to the same terminals across which the photo-voltages appear. The reversibility and control provided make immediately possible applications to use in computer memories of a new type -- in which information is stored as remanent polarisation and read out as the polarity and magnitude of a photo-current or photo-emf, such typical applications are disclosed here.

Application to the generation of electrical power from solar radiation, for example, to solar battery type devices and to electrical power generating stations operating on the basis of solar to electrical energy conversion also is

possible and contemplated but would require, to be practical, (except in special cases) considerably larger conversion efficiency than has been observed so far in the materials examined. A calculation of theoretical maximum efficiency, however, yields results which are large enough to suggest eventual practical use in this manner. A conversion system based on these high voltage materials would have the particular advantage of producing its electricity directly at high voltage which is advantageous for power transmission purposes.

The mechanism for the discovered effect appears to be unique and different from photovoltaic mechanisms previously described. Description will be provided explaining the mechanism and developing a theory for it. From this, it will be clear that the entire class of polycrystalline ferroelectrics are expected to exhibit high photo-emf's to at least some extent.

BRIEF DESCRIPTION OF THE DRAWINGS

The invention itself will be better understood and further features and advantages of it will become apparent from the following detailed description which makes reference to the drawings, where:

Fig.1 is a schematic diagram illustrating the basic arrangement by which photovoltaic voltages are generated by the application of light to a ferroelectric ceramic material as shown by this invention;

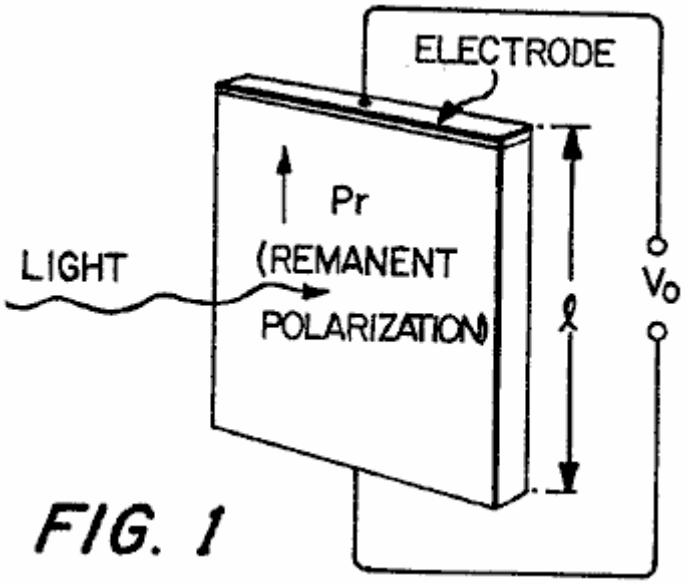


FIG. 1

Fig.2 is an electrical schematic diagram depicting an equivalent circuit to the basic apparatus of **Fig.1**, where C_0 is the capacitance of the sample measured utilising a capacitance meter connected between the electrodes and C_1 is the parallel capacitance of a load coupled to the electrodes, and R_1 is the resistive value of that load;

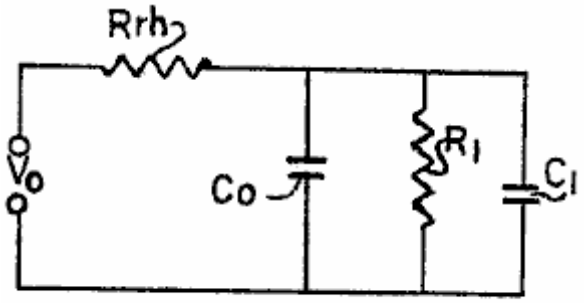


FIG. 2

Fig.3 is a graphical illustration of current vs. applied voltage to an illuminated ferroelectric wafer of the basic form depicted in **Fig.1**;

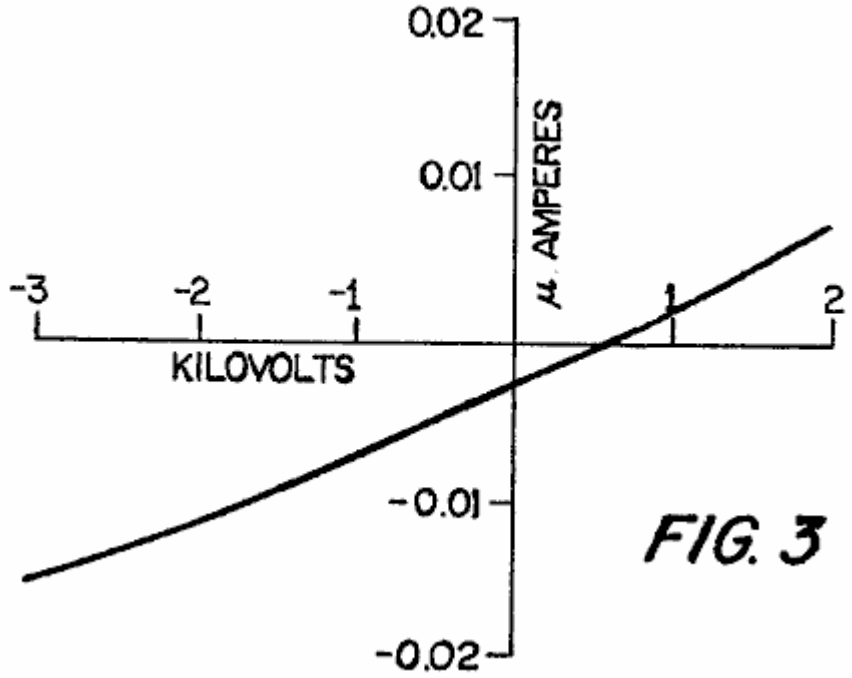


FIG. 3

Fig.4 is a graphical illustration of the photo-emf and photo-current as a function of intensity of illumination, with the particular graphical results being for a solid solution $Pb(Zr_{0.53}Ti_{0.47})0_3$ with about 1% by weight of Nb_20_5 added;

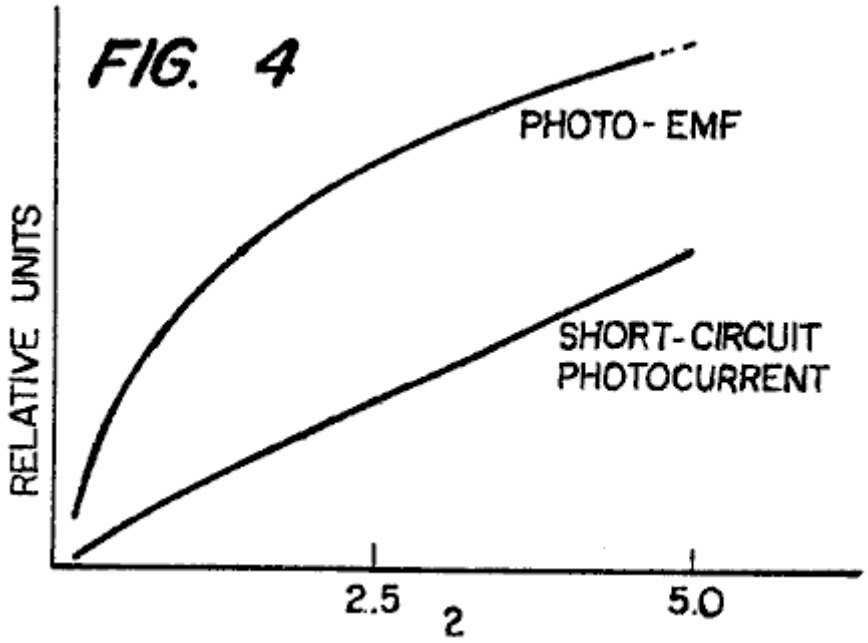


Fig.5 is a graphical illustration of photo-emf vs. grains per unit length (inverse median grain size) for two different materials;

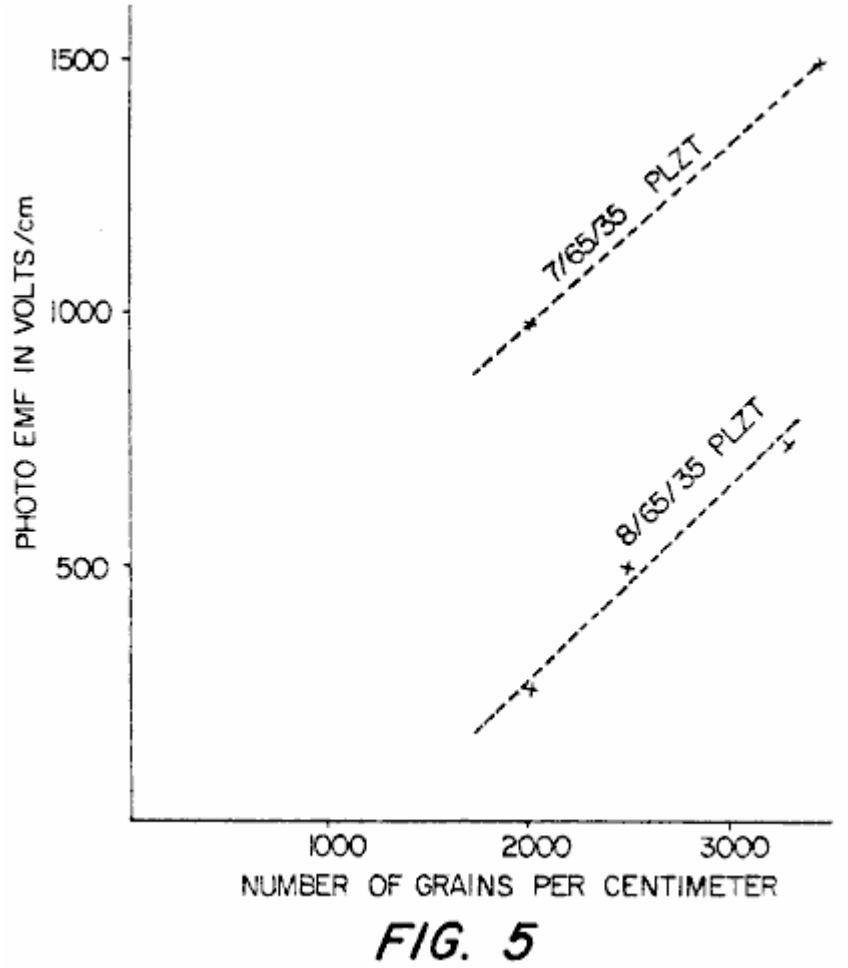


FIG. 5

Fig.6 is a graphical illustration of photo-voltage vs. remanent polarisation for ceramic $BaTi0_3$ + 5% by weight of $CaTi0_3$;

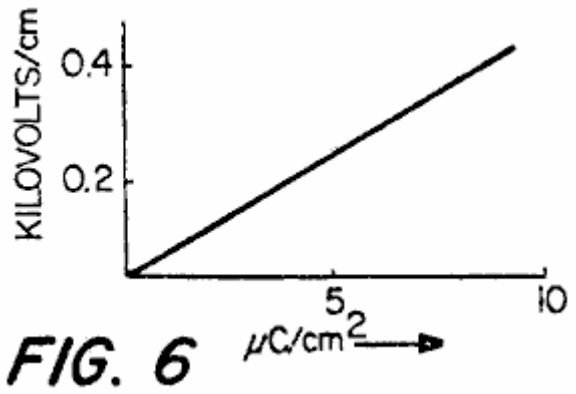


FIG. 6

Fig.7 is a diagram illustrating the short-circuit photo-current as a function of wavelength for the solid solution $Pb(Zr_{0.53}Ti_{0.47})0_3$;

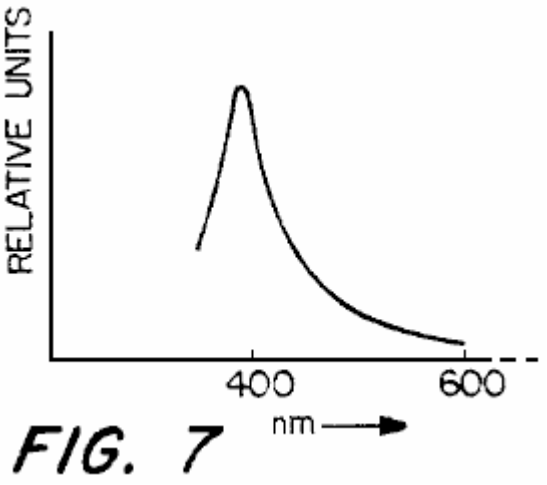


FIG. 7

Fig.8 is a diagram illustrating the short circuit photo-current as a function of wave length for ceramic $BaTi0_3$ + 5% by weight of $CaTi0_3$;

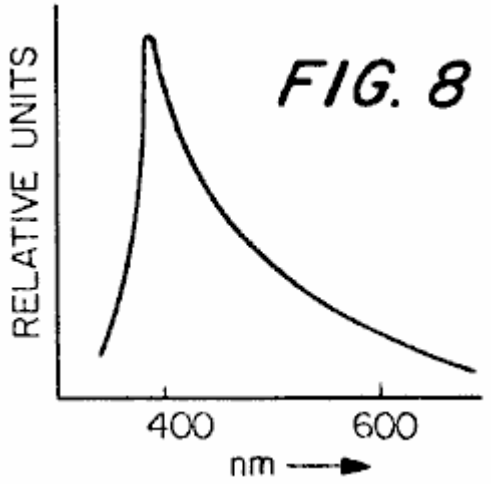


FIG. 8

Fig.9 is a diagram illustrating the short-circuit photo-current as a function of wavelength for the solid solution $Pb(Zr_{0.65}\ Ti_{.0.35})0_3$ with 7% of the lead substituted for by lanthanum;

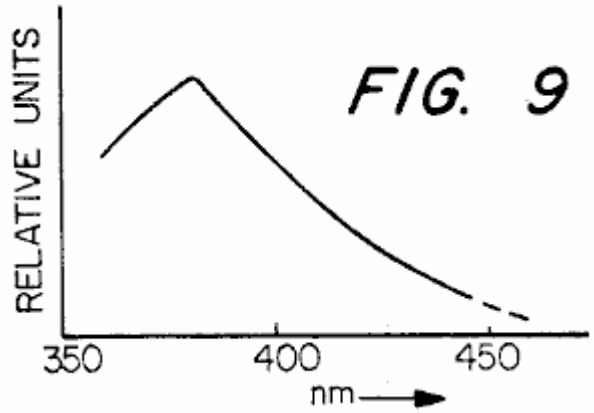


FIG. 9

Fig.10 is a diagram illustrating the photo-emf vs. wavelength for the solid solution $Pb(Zr_{.0.35}Ti_{0.47}0_3$ with 1% by weight of Nb_20_5 added;

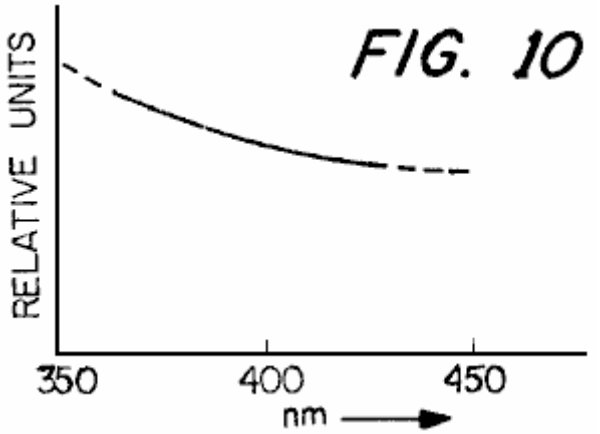


FIG. 10

Fig.11 is a diagram illustrating the photo-current divided by intensity vs. cut-off wave length of long wave length cut-off dichroic filters, with the materials being $Pb(Zr_{0.53}\ Ti_{.0.47})0_3$ with 1% by weight of Nb_20_5 added and utilising a high-pressure mercury arc as the illumination source;

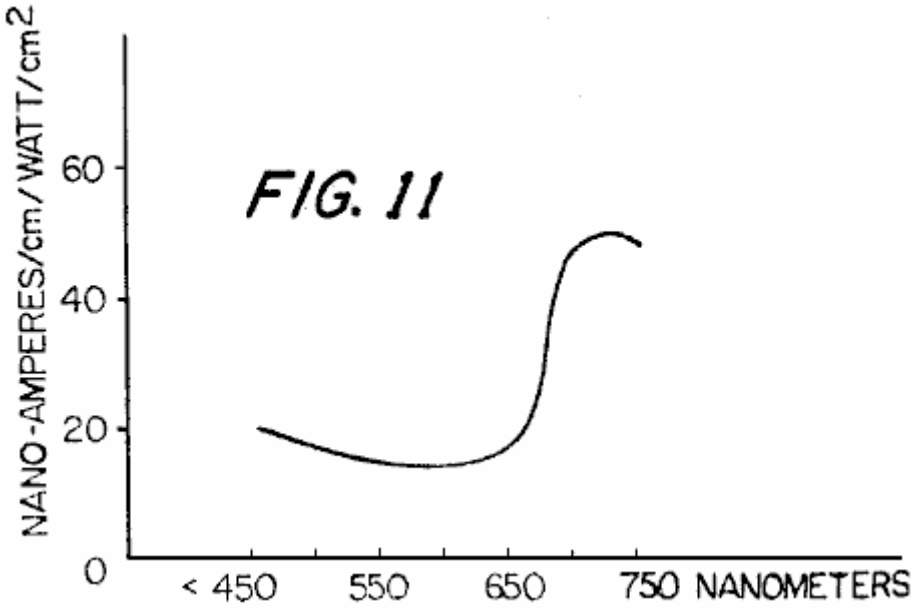


FIG. 11

Fig.12 is a diagram illustrating the photo-current divided by intensity vs. cut-off wavelength of short wave length cut-off filters, with the material being $Pb(Zr_{0.53}\ Ti_{.0.47})0_3$ with 1% by weight of Nb_20_5 added;

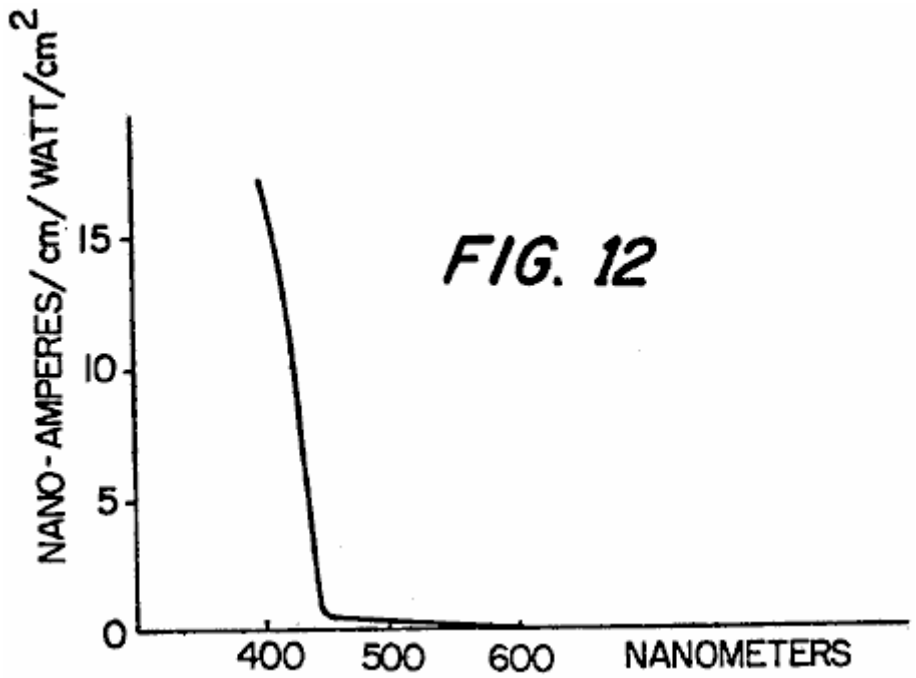


FIG. 12

Fig.13, is diagram illustrating the photo-emf vs. wave length of short wavelength cut-off filters, with the material being $Pb(Zr_{0.53}\ Ti_{.0.47})0_3$ with 1% by weight of Nb_20_5 added;

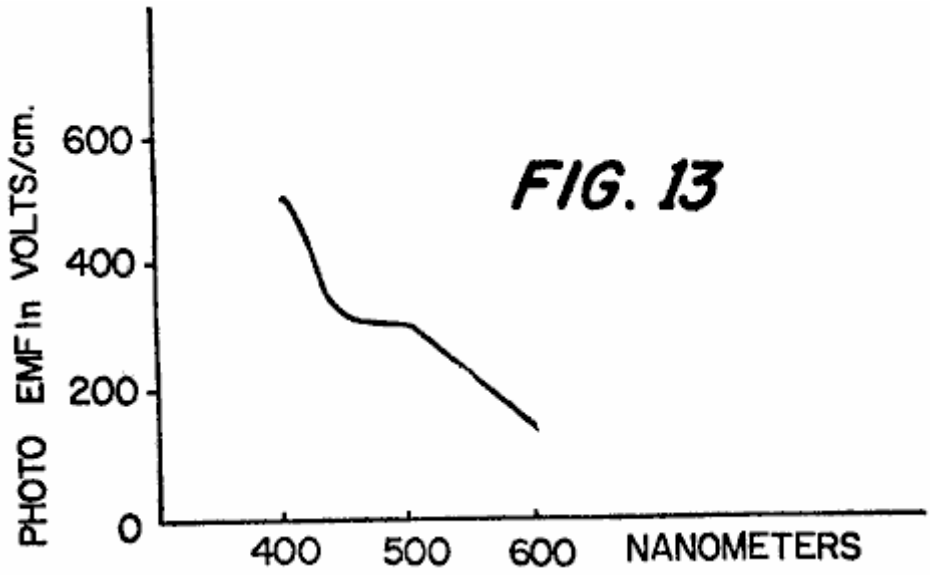


FIG. 13

Fig.14 is a pictorial illustration of the manner in which a single crystal produces a photo-emf, with the polarisation $\mathbf{P_s}$ being normal to the electrodes, which electrodes are illustrated by the shaded area;

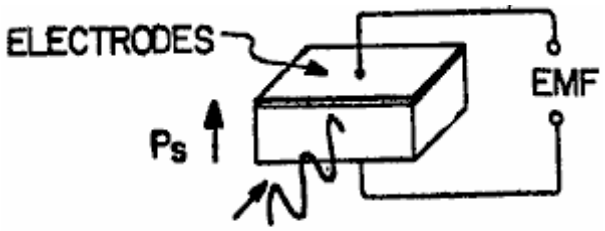


FIG. 14

Fig.15 is a diagram illustrating photo-current vs. wave length of the single crystal $BaTi0_3$;

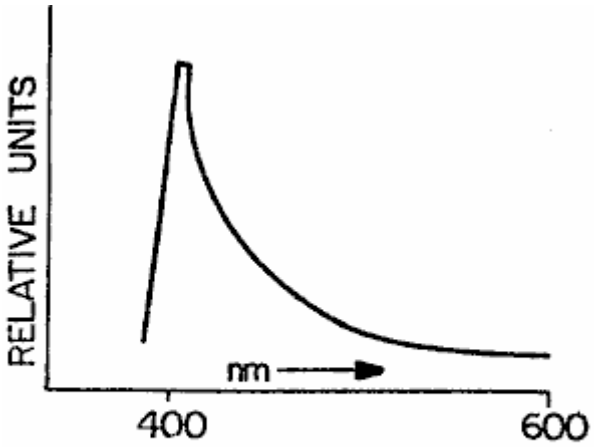


FIG. 15

Fig.16 is a diagram illustrating the photo-voltage vs. temperature for $BaTi0_3$ +5% by weight of $CaTi0_3$;

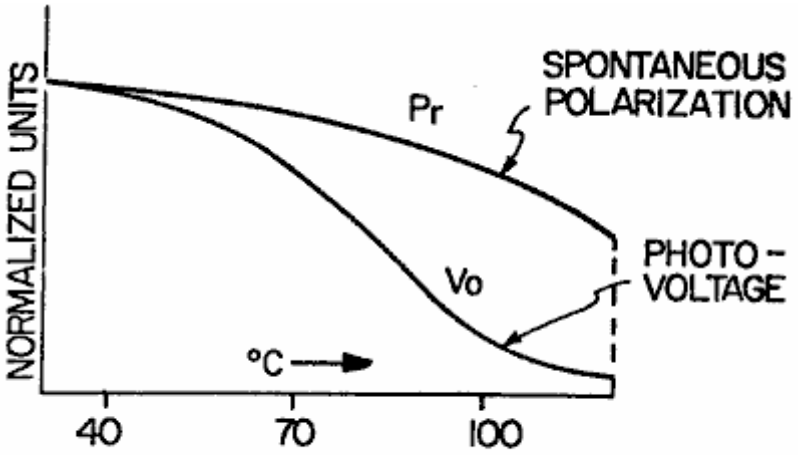


FIG. 16

Fig.17 is a diagram illustrating the photo-voltage vs. temperature of single crystal $BaTi0_3$;

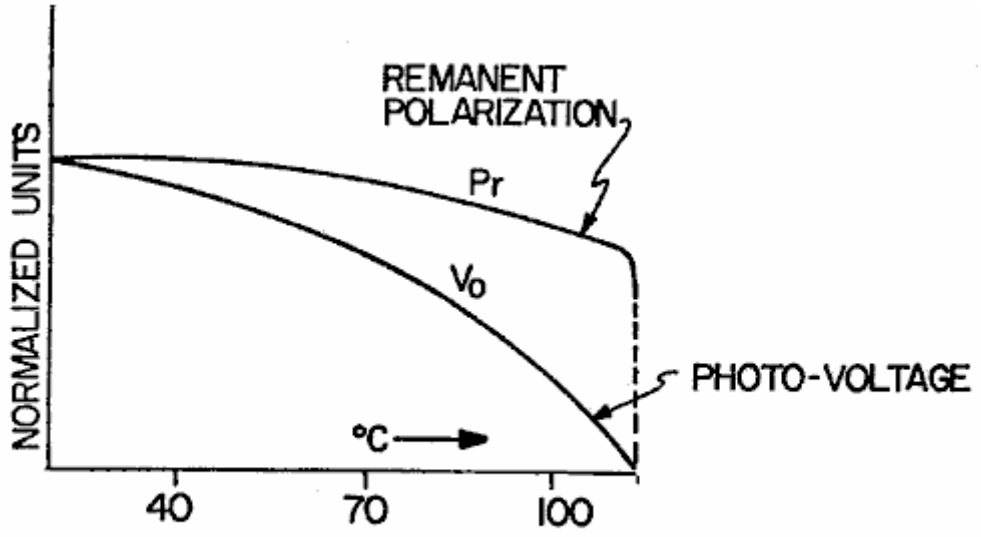


FIG. 17

Fig.18 is a diagram illustrating photo-current vs. temperature for $BaTi0_3$ + 5% by weight of $CaTi0_3$;

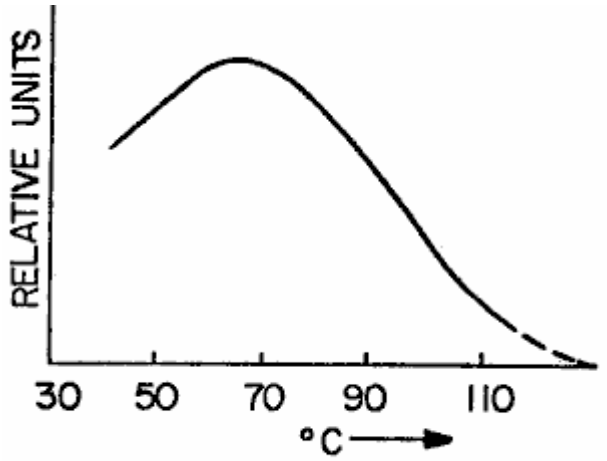


FIG. 18

Fig.19 is a cross-sectional, elevational view schematically depicting the ceramic slab of **Fig.1**, with the photo-emf appearing across the electrodes on the edge, and with most of the photo-current flow being found in the shaded region near the surface;

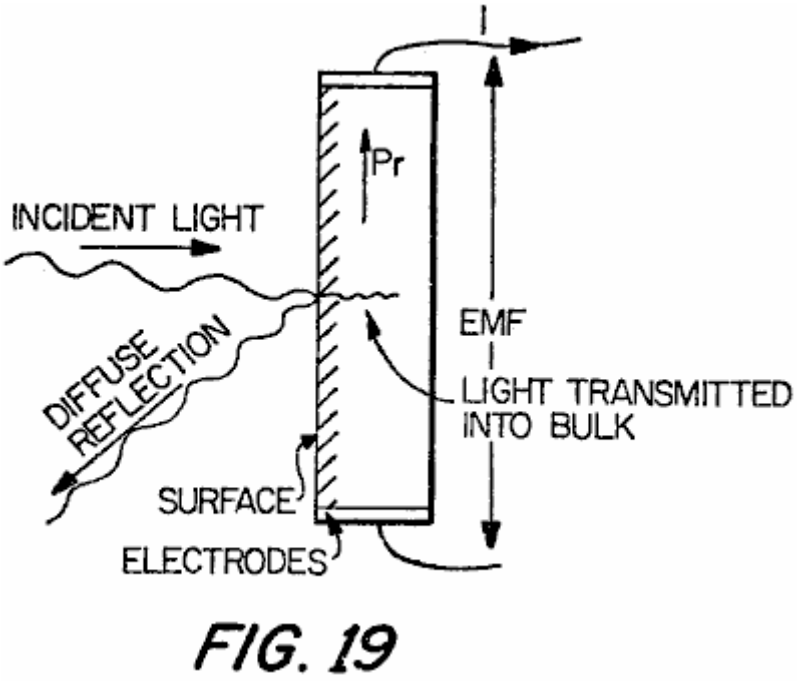


FIG. 19

Fig.20 is a cross-sectional, elevational view of a slab of ferroelectric ceramic material utilising transparent electrodes and depicting light incident through the transparent electrodes into the slab with the slab being polarised in the thickness direction;

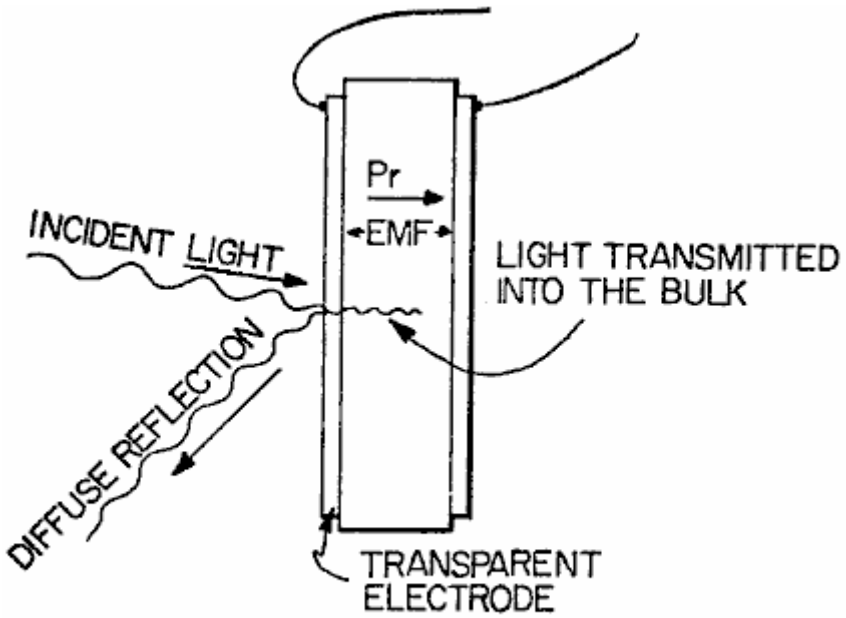


FIG. 20

Fig.21 is a cross-sectional diagrammatic illustration of a single layer of grains depicting the manner in which photo-emf's are produced across the grains in an additive fashion to produce a length dependent effect in the ceramic material, the illumination being incident from the left-hand portion of the drawing and being typically quickly absorbed as it penetrates the material;

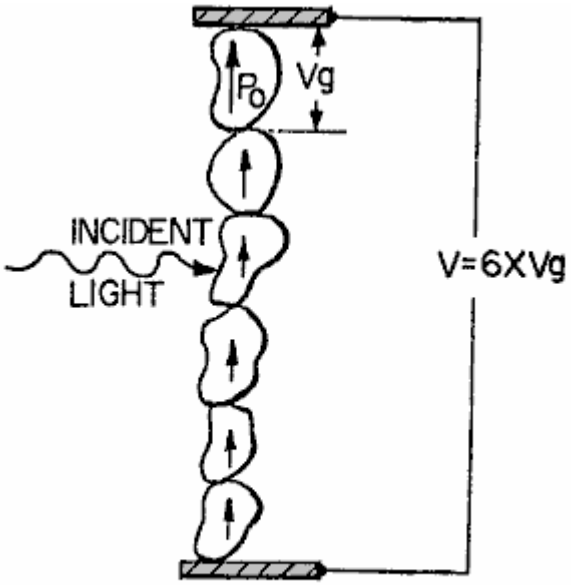


FIG. 21

Fig.22 is a diagram illustrating idealised two dimensional crystals of length **l** with spontaneous polarisation $\mathbf{P_s}$, dielectric constant Epsilon $_b$ compensating surface charge per unit area of Sigma = P_s ;

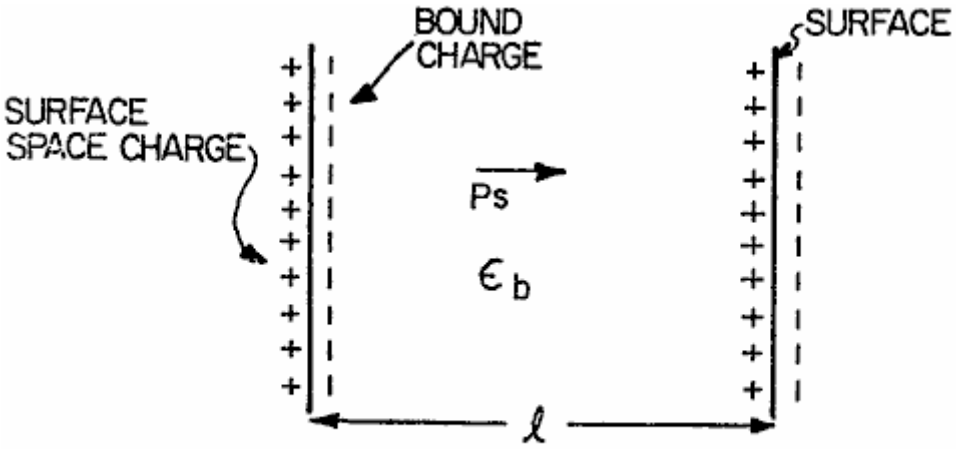


FIG. 22

Fig.23 is an illustration depicting the structure of a typical ferroelectric grain or crystallite;

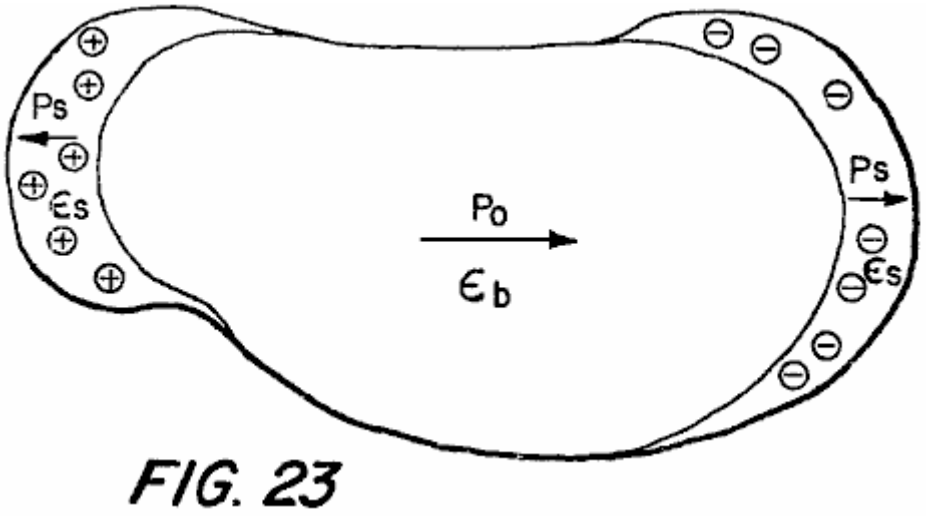


FIG. 23

Fig.24 is an illustration depicting a model of a crystal of length **l**;

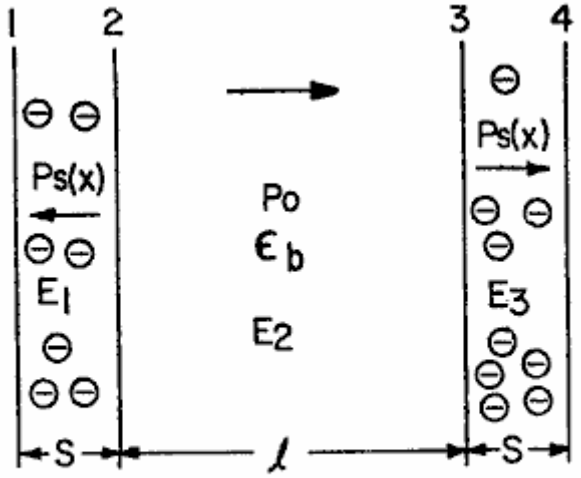


FIG. 24

Fig.25 is a diagram illustrating the potential distribution in an illuminated crystal;

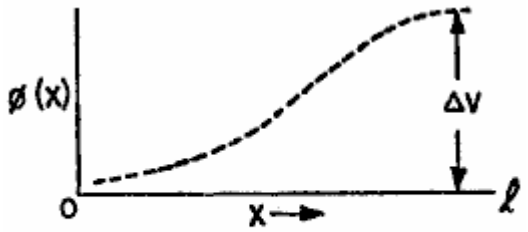


FIG. 25

Fig.26 is a schematic representation of the instant inventive ferroelectric ceramic substrate utilised as a photovoltaic memory device with optical scanning;

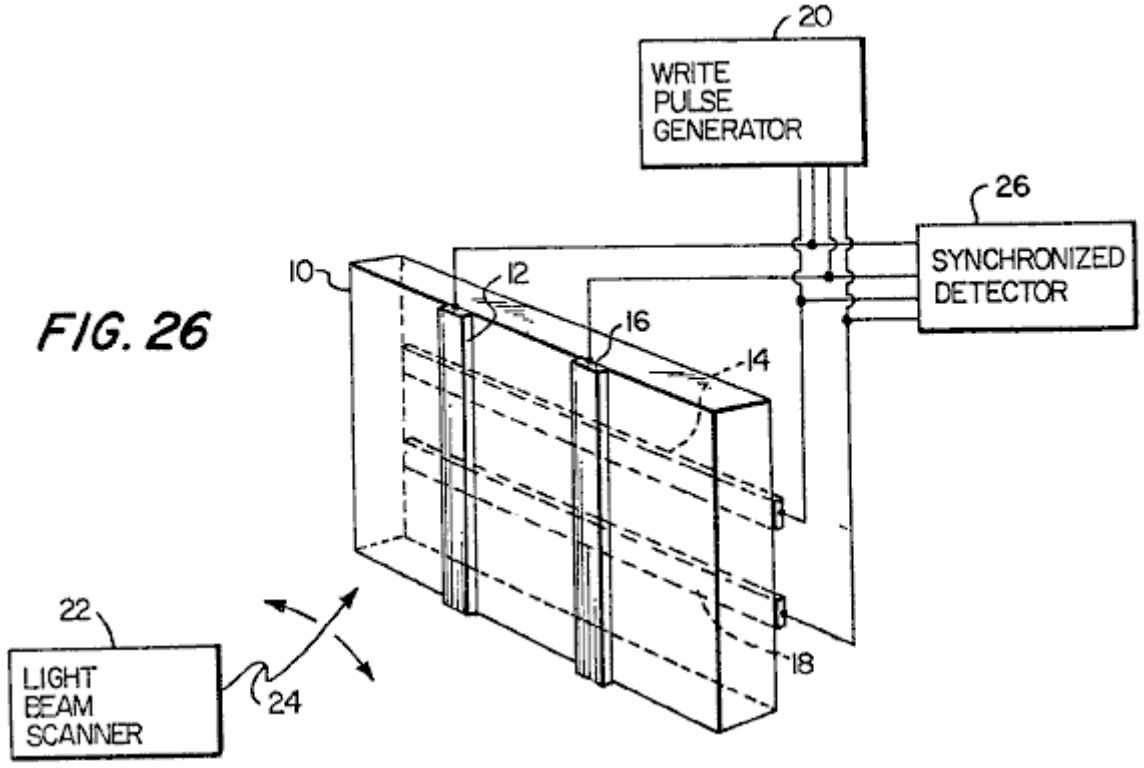


FIG. 26

Fig.27 is a schematic illustration of an optical display apparatus utilising a ferroeletric ceramic material in accordance with the general teachings of the instant invention;

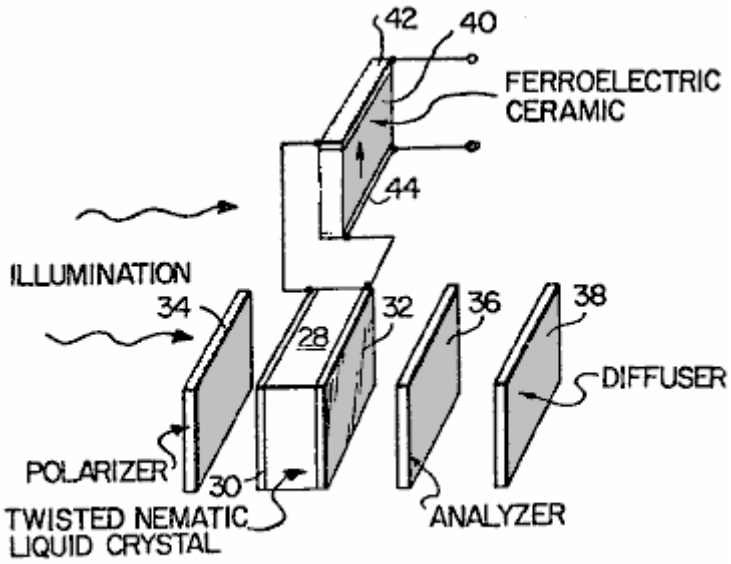


FIG. 27

Fig.28 is a schematic illustration depicting an optical display apparatus constructed in accordance with the teachings of the instant invention in monolithic form utilising a colour switching liquid crystal;

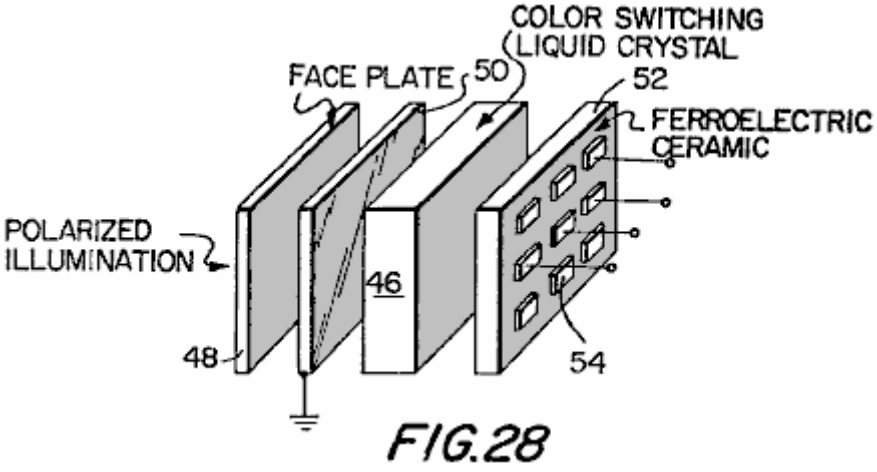


FIG.28

Fig.29 is a schematic illustration of the display apparatus of **Fig.28**, modified to make utilisation of a twisted nematic liquid crystal;

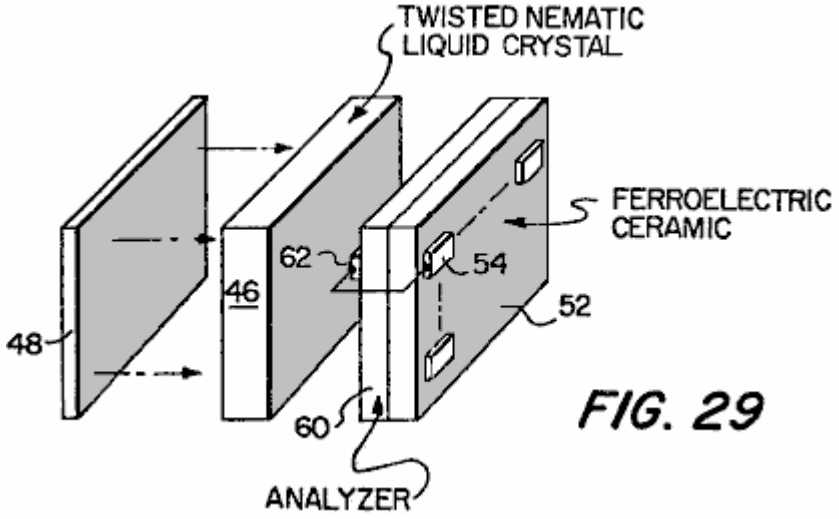


FIG. 29

Fig.30 is a cross-sectional elevational view depicting an optical display apparatus utilising a colour switching liquid crystal in conjunction with a ferroelectric ceramic substrate of the instant invention, and which display apparatus exhibits permanent memory capabilities;

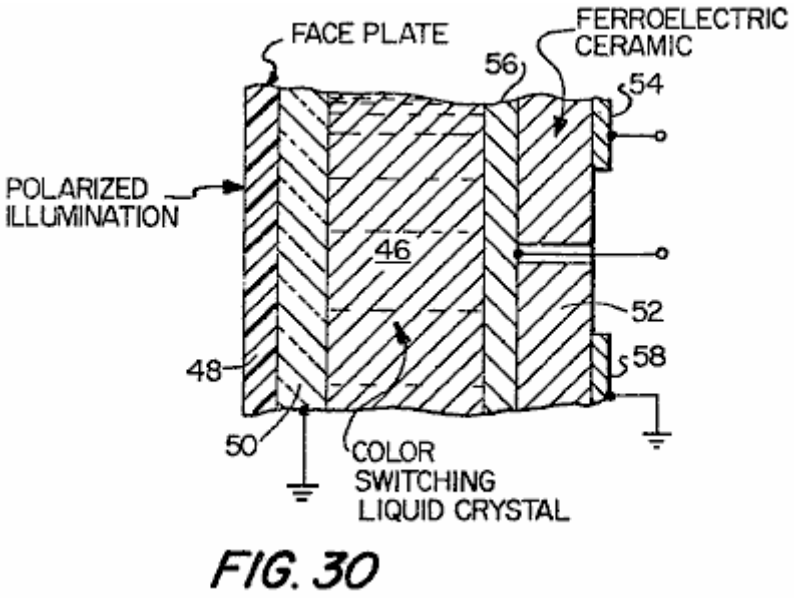


FIG. 30

Fig.31 is a cross-sectional elevational view of a further form of an optical display apparatus constructed in accordance with the teachings of the instant invention, said apparatus utilising a colour switching liquid crystal and further utilising length-wise polarisation of the ceramic substrate;

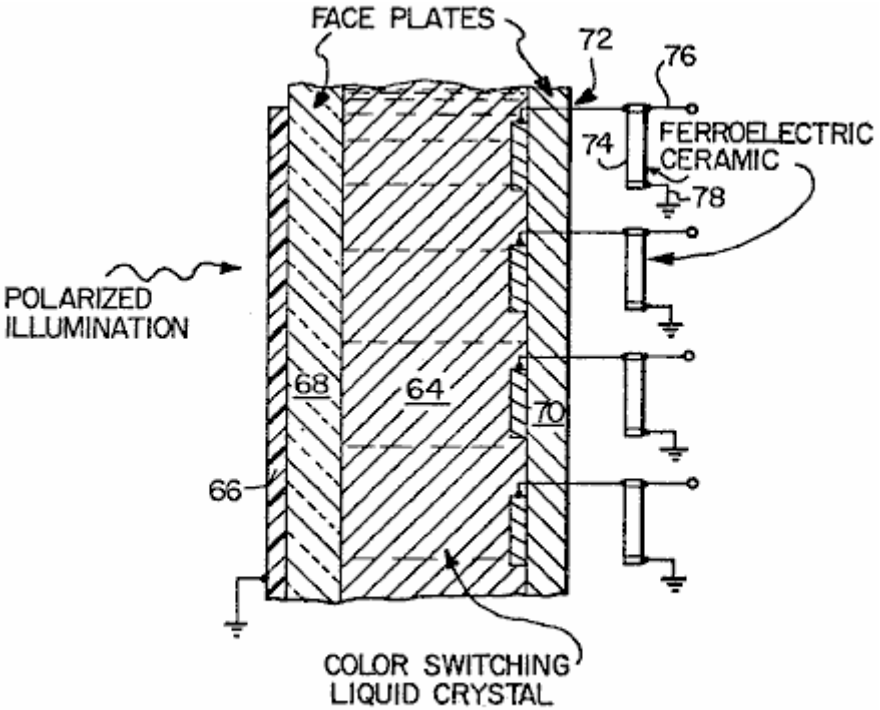


FIG. 31

Fig.32 is an elevational view, in section, of a further form of an optical display apparatus constructed in accordance with the teachings of the instant invention, this apparatus being similar to that depicted in **Fig.31** of the application drawings but utilising a liquid crystal of the twisted nematic type; and

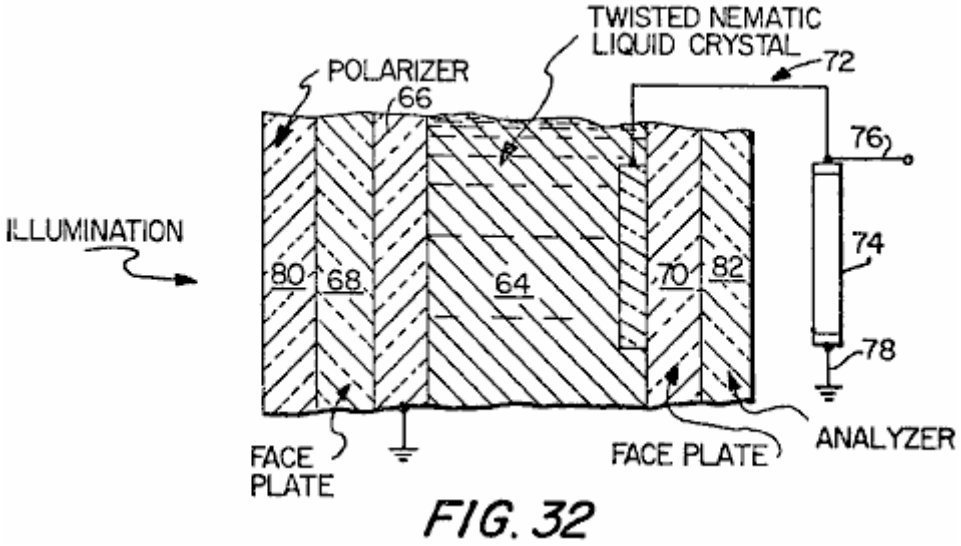


FIG. 32

Fig.33 is a schematic illustration of a further form of optical display and storage utilising the photoconductive as well as photovoltaic properties of the ferroelectric ceramics.

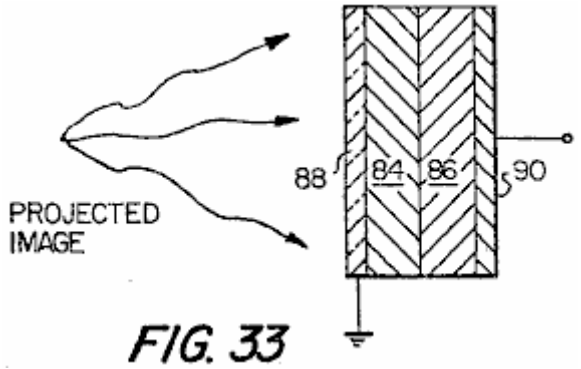


FIG. 33

Fig.34 illustrates how the image stored in a substrate is displayed.

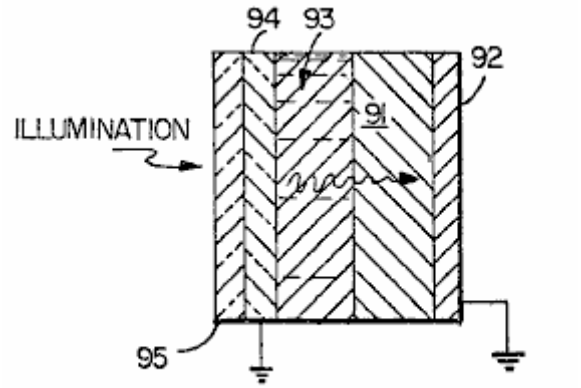


FIG. 34

DETAILED DESCRIPTION OF THE PREFERRED INVENTIVE EMBODIMENTS

With reference now initially to **Fig.1** of the application drawings, a discussion of the novel phenomena of the instant invention will ensue. Upon the application of incident illumination to the ferroelectric ceramic, a steady voltage is produced which is proportional to the length **l** between the electrodes. By dividing the sample into two equal segments along a line perpendicular to the direction of the remanent polarisation and by placing new electrodes on the cut edges, new samples would result each producing photo-emf's which is one half the original photo-emf.

An arrangement such as that shown in **Fig.1** can be described roughly by the equivalent circuit as shown in **Fig.2**. This has a saturation photo-emf **Vo**, in series with the photo resistance of the illuminated sample. **Fig.3** is a current-voltage characteristic of a typical illuminated ferroelectric slab, and has the form expected from the equivalent circuit in **Fig.2** except for the slight tendency towards saturation in the lower left quadrant. As a function of intensity, the photo-emf saturates at relatively low levels of illumination. The short circuit photo-current is, however, linear with light intensity. Results for the material $Pb(Zr_{0.53}Ti_{0.47})0_3$ with 1% by weight of Nb_20_5 are shown in **Fig.4**. The implication of these results and the equivalent circuit in **Fig.2** is that the photo-resistance R_{ph} is inversely proportional to intensity.

A saturation photo-emf and a short circuit current proportional to intensity has been measured in several poled ferroelectric materials. These are shown in Table I:

Table I

Photovoltaic outputs at room temperature for several ceramic compositions. The wafers were fully poled, to their maximum remanent polarization. Filtered illumination had a half bandwidth of about 10 nm. The photo-emf is a saturation value reached at relatively low value of intensity.

Sample	Illumination Wave length (nm)	Saturation Photo-emf (Volts/cm)	Short Circuit Photocurrent ($\frac{\mu Amperes/cm}{watts/cm^2}$)
$Pb(Zr_{.53}Ti_{.47})O_3$ +1 wt% Nb_2O_5	373	610	.31
$BaTiO_3$ + 5 wt% $CaTiO_3$	403	360	.020
$Pb(Zr_{.65}Ti_{.35})O_3$ with 7% lanthanum-lead substitution	382	1500	.030
$Pb(Zr_{.65}Ti_{.35})O_3$ with 8% lanthanum-lead substitution	382	750	.015
$BaTiO_3$ + 5wt% $CaTiO_3$	403	355	.02
$Pb(Zr_{.53}Ti_{.47})O_3$ + 1wt% Nb_2O_3 with polished surfaces	382	610	~.61

For a given composition the photo-emf is also a function of grain size. These results are shown in Table II.

Table II

Photo-emf for different grain size and percent lanthanum substituted for lead. The materials are $Pb(Zr_{.65}Ti_{.35})O_3$ with 7% lanthanum substitution for lead and the same material with an 8% lanthanum substitution for lead.

Grain Size (microns)	Percent Lanthanum-Lead Substitution (percent)	Saturation Photo-emf (Volts/cm)
2-4	7	1500
4-6	7	980
greater than 6	7	560
2-4	8	750
3-5	8	510
4-6	8	330
greater than 6	8	250

the photo-voltage v. number of grains per unit length is plotted in **Fig.5** for two different compositions. The plot clearly shows a relationship between the two quantities.

The fact that the photo-emf of a particular sample depends on the remanent polarisation is shown by the results for a typical ferroelectric material, barium titanate + 5% by weight of $CaTi0_3$,as plotted in **Fig.6**.

The short circuit photo-current depends strongly on the wave length of the impinging illumination. It is a maximum at a wavelength resulting in a photon energy equal to the band gap energy of the material. Other wavelengths can, however, contribute strongly to the current.

Results for typical materials are shown in **Fig.7**, **Fig.8**, and **Fig.9**. The current (ordinate) is that produced by illumination contained in a small band, of about +10 nm about a wavelength indicated on the abscissa. A mercury source and notch type dichroic filters were used. The total intensity within each band was only roughly constant.

The current that has been plotted has been therefore normalised to constant intensity by assuming the linear relation between the two.

The photo-emf is less strongly dependent on wave length. Results for a particular material, using notch dichroic filters is shown in **Fig.10**. These values are saturation values, roughly independent of intensity.

An important additional phenomena shows a dependence of current produced in the red and infrared regions in the presence of simultaneous blue band gap radiation. These results are shown in **Fig.11** and **Fig.12**. The ordinate (**Fig.11**) is the current produced by the light from a mercury arc shining through dichroic long wavelength cut off filters, the abscissa the wavelengths above which no light illuminates the sample. Note the step at 650 nm. Using short wavelength cut off filters which eliminate the band gap light results in no current until the cut off wavelength is below the band gap. These results are shown in **Fig.12**. The amount of output in the red actually depends on the intensity of simultaneous band gap radiation, thus the energy efficiency of these materials for a broad band source is not simply the intensity weighted average of the efficiencies for individual wavelengths as produced by notch filter. The actual value is larger.

Photo-emf vs. cut-off wavelength for $Pb(Zn_{0.53}Ti_{0.47})0_3$ +1% by weight of Nb_20_3 is shown in **Fig.13**. A substantial photo-emf appears at long wavelengths but no current can flow. In other words, the internal resistance R_{ph} is extremely high unless band gap is incident.

Single Crystal Results

The ceramic results imply a small photo-emf from a single crystal illuminated as shown in **Fig.14**. Such emf = 0.55V at room temperature was indeed observed.

The short circuit current is, as for the ceramic material, a strong function of wavelength. These results are shown in **Fig.15**.

Temperature Dependence

Ceramic photo-emf is a function of temperature. Results for barium titanate ceramic with 5% by weight of $CaTi0_3$ are shown in **Fig.16**. For both $Pb(Zn_{0.53}Ti_{0.47})0_3$ with 1% by weight of Nb_20_5 added and barium titanate the photo-emf decreases with increasing temperature. In these measurements, the temperature ranged to the transition temperature, the photo-emf vanishing at the temperature at which the remanent polarisation also vanishes. The remanent polarisation vs. temperature for this material is also shown in **Fig.16**. Similar results for single crystal barium titanate are shown in **Fig.17**. The single crystal photo-emf are, of course, much smaller. Short circuit was measured as a function of temperature. Results for barium titanate +5% by weight of $CaTi0_3$ are shown in **Fig.18**. Similar results over the same temperature range were obtained for $Pb(Zn_{0.53}Ti_{0.47})0_3$ + 1% by weight of Nb_20_5 material. In that case there was no maximum, the photo-current still increasing with increasing temperature at 130^0 C.

Effects of Optical Properties

In the arrangement shown in **Fig.1**, the direction of polarisation, and consequently the direction of the photo-emf is perpendicular to the direction of incidence of the light which is also the direction in which the light is strongly absorbed. The light only enters into a region near the surface of the material. The rapidity of the absorption depends strongly on the wavelength of the light, the light becoming fully absorbed in a region closer and closer to the surface as one decreases the wavelength of the light and approaches the band gap wavelength. For shorter wavelengths, the light no longer enters the material and thus for these wave lengths the light-induced effects decrease rapidly with decreasing wavelength.

Ceramic materials which exhibit these photo-emf's can appear transparent, translucent, and apparently opaque when viewed with white light. Light, however, obviously enters even the opaque materials to produce the photo-emf's. The apparent opacity is produced by diffuse reflection at granular boundaries. It is of course desirable to minimise the degree to which diffuse reflectivity prevents light from entering the material. Nevertheless, the largest photo-currents and greatest photovoltaic efficiency has been originally observed in a material which appears opaque in thickness more than a few thousandths of an inch. The cross sectional drawing **Fig.19** depicts the way light enters the material with the arrangement as originally shown in **Fig.1**.

When a circuit connects the electrodes, the maximum density of current occurs near the surface, the current density decreasing in regions deeper within the thickness.

Polishing the surfaces of these materials, however, increases the transparency and, as expected, the magnitude of the photo-current and the photovoltaic conversion efficiency. An emf will also be produced by the arrangement shown in **Fig.20** provided, of course, that the electrodes are of a nature to allow light to enter the material. Normal thick metal electrodes are opaque to light. When metal electrodes are thin enough, they permit light to be transmitted and yet are sufficiently conductive to function as electrodes. Other conducting transparent electrodes include indium oxide. The emf now will be seen to appear across the thickness of the material, in the direction of the remanent polarisation.

In this arrangement the high dark resistance of any un-illuminated bulk portion of the material is in series with the circuit connecting the electrodes. The current that can be drawn is limited. Maximum currents can be drawn when the thickness between the electrodes is equal to or less than the absorption depth of the radiation. However, since the saturation photo-emf is not a strong function of intensity, vanishing only for extremely low intensities, the full photo-emf per unit length v_o can usually be observed for this samples.

Proposed Mechanism for the High Voltage Photovoltaic Effect in Ferroelectrics

Briefly, it is proposed that the photo-emf results from the action of an internal field within the bulk of an individual ceramic grain on non-equilibrium carriers generated by illumination. These carriers move to screen the internal field. The photo-emf that appears is the open circuit result of such screening. A change in charge distribution upon illumination changes the voltage across a grain from an initial value of zero to the photo-voltages which are observed.

These photo-emf's appears across individual ceramic grains. What is observed as a length dependent high photo-voltage is the series sum of the photo-emf's appearing across grains, each of which is characterised by saturation remanent polarisation P_o. The situation is shown schematically in **Fig.21**. Individual grains typically are small, of the order of 10 microns in diameter. To produce a high photo-voltage per unit length in the ceramic the voltage across an individual grain need not be large. For example the results in **Table II** for $Pb(Zn_{0.65}Ti_{0.47})0_3$ with 7% Lator Pb can be explained by individual grain photo-voltage of only about 0.5 volts per grain. The clear implication of the experimental results (**Table II** and **Fig.5**) is that for the range of grain sizes investigated, the photo-emf across a grain is more or less independent of the size of the grain. This is supported also by the single crystal results.

Ferroelectric crystals are characterised by large spontaneous polarisation which would be expected to produce large emf's even in the dark. Such emf's are not observed even across highly insulating materials. This is presumed to be the result of space charge within the volume or on the surface of a ferroelectric crystal (which, in ceramics, are the individual grains or crystallites). The space charge produces a potential across a crystal cancelling the potential produced by the net polarisation within they crystal. It is obvious that as long as there are sufficient charges within the crystal which are free to move, any potential produced by an internal polarisation will eventually vanish.

This dark zero potential state is the initial state of a crystal crystallite, grain, and of the ceramic body composed of these grains. The absence of a net potential in the dark does not however mean the absence of internal fields. Internal fields can be expected to exist and are the consequence of the spatial distribution of the charges which bring the net potentials across grains to zero. These spatial distributions can not be arbitrarily assigned, but are subjected to constraints of a basic physical nature.

In the idealised two dimensional crystal shown in **Fig.22**, the surface charge density Upsilon = P_s reduces the potential between the surfaces to zero. If the surface charge density (in actuality this does not occur) is completely juxtaposed upon the bound polarisation surface charge, which has a value P_s, then there are no internal fields. Were there no charge, the crystal would show an internal field P_s/ϵ_b and a potential between the surfaces of $P_s l / \epsilon b$.

Such a field would be well above the dielectric breakdown strength of a real dielectric. For a single domain typical ferroelectric barium titanate $P_s = 26 \times 10^{-2}$ C/m, and the relative dielectric constant $Epsilon_r$ in the direction of polarisation is 137. The field that would have to exist in the absence of compensation charge is over 2×10^6 volts/cm which is well above the dielectric strengths typical of these materials. If such a field could momentarily exist within a ferroelectric crystal it would not exist for long but be reduced from its maximum value to some value below the dielectric strength of the material. The strong field would break down the material and a charge flow would produce a space charge distribution resulting in a new lower value for the internal fields within the crystal.

Such a space charge distribution must exist in an actual crystal. The space charge serves to reduce the potential across a crystal to zero. Such charges have limited mobility and the materials continue to behave as insulators for ordinary strength applied fields.

Such a space charge cannot occupy a delta function-like region as in the idealised situation shown in **Fig.22**, but must occupy instead a finite volume. If these are localised near the surface of the crystal, then an internal field $Epsilon_b$ exists within the bulk of the material and additional fields E_s exist within the space charge regions near the surface.

It is hypothesised that these space charge regions are near the surface of real crystals with the charge distributed within a surface layer thickness **s**. The reasons for same are as follows:

(1) The surface regions of ferroelectric crystals are characterised by regions whose dielectric, ferroelectric, and thermodynamic properties differ markedly from that of the bulk. These differences are best explained by the existence of strong fields in this region that would be produced by space charge. There is a considerable body of information in the literature supporting the existence and delineating the properties of these layers;

(2) The interplay of space charge and the very non-linear dielectric constant of ferroelectric would be expected to localise space charge in a low dielectric constant layer near the surface. In ferroelectrics, unusually high, low field relative dielectric constants (of the order of 1000) can be expected to reduce in value with increasing field strength. Thus charge in a region reduces the dielectric constant of that region increasing the field strength of that region. This feedback mechanism can be shown to localise charge within a layer.

The experimental results supporting the existence of surface layers will not be reviewed here, nor the calculations which support the localisation of charge into layers as a result of a non-linear (saturable) dielectric constant. These may be reviewed by referring to the literature.

A schematic description of a typical grain, i.e. crystallite, with space charge regions of thickness **s**, and a bulk region of thickness **l**, is shown in **Fig.23**. The internal fields (in the two dimensional model) of such a charge distribution superimposed on that produced by the bound polarisation charge will be calculated and also the effect of these fields on carriers within the bulk produced as the result of an internal photo effect (photo-ionisation). Formulae for the photo emf that will be derived will have the correct sign, a linear dependence on remanent polarisation, and the kind of temperature dependence that has actually been observed. In addition there will result an estimate of a size independent grain photo-emf for a typical ferroelectric, barium titanate, which is consistent with that implied from the observed ceramic emf, and single grain emf. The grain has as shown in **Fig.23**:

(1) A bulk region with dielectric constant $Epsilon_b$ and uniform polarisation (at zero applied field) P_o ;

(2) Surface layers of dielectric constant $Epsilon_s$, considerably less than that of the bulk. There are also polarisation in the surface regions P_s (x) which exist at zero applied field. These will generally be parallel to the bulk polarisation at one end and anti-parallel at the other end;

(3) Space charges in these surface layers which serves to remove any potential across the grain. It is the space charge layers which produce high fields which reduce the highly non-linear dielectric constant of the bulk to the lesser value in the surface layers, and also produce the remanent polarisation, P_s (x) with the surfaces.

Such a structure also has an internal bulk field, and surface fields which can be calculated. For the purposes of this calculation we assume a simple two dimensional model shown in **Fig.24**.

The polarisation with the various regions are assumed only for simplicity to be uniform within these regions. Again, only for simplicity those in the surface layers and the bulk are assumed equal in magnitude (i.e. P_s (x) = P_o). The space charge densities .+-.n_oe are also assumed uniform and equal in magnitude. The polarisations are equivalent to four bound surface charge densities,

$$\sigma_1 = \vec{P}_o \quad \sigma_2 = -2\vec{P}_o$$

$$\sigma_3 = o \quad \sigma_4 - \vec{P}_o$$

There are, using Gauss's law, electric fields as shown in **Fig.24**.

$$E_1 = \frac{1}{\epsilon_s}[P_o + m_o ex]$$

$$E_2 = \frac{1}{\epsilon_b}[-P_o + m_o es]$$

$$E_3 = \frac{1}{\epsilon_s}[-P_o + n_o e(s-x)]$$

It has been assumed that the voltage across the crystal vanishes,

$$\int_0^{l+2s} E(x)dx = O$$

n_o and s, from this and the three preceding equations, must be related by the expression

$$m_o es = \frac{P_o}{1 + \frac{\epsilon_b}{\epsilon_s}\frac{s}{l}}$$

and the bulk field

$$E_2 = \frac{-P_o}{\epsilon_b}\left[\frac{\frac{s}{e}\frac{\epsilon_b}{\epsilon_s}}{1 + \frac{s}{l}\frac{\epsilon_b}{\epsilon_s}}\right]$$

Surface layers in barium titanate ceramic grains have been estimated at 10^{-6} cm (see for example Jona and Shirane Ferroelectric Crystals, Pergammon Press, 1962). The remanent polarisation typical of the ceramic material is about 8 x 10^{-2} C/m^2, the relative dielectric constant of the poled ceramic about 1300. The high field dielectric constant will be estimated at roughly 0.5 the bulk dielectric constant. These numbers yield a bulk field, for a typical 10^{-3} cm grain of,

E_2 = 350 volts/cm

The potential across the bulk would thus be approximately -0.35 volts. The remaining potential across the grain would be that across the surface layers. Illumination has the effect of producing charges which screen the internal field, E_2 causing it to vanish.

The negative voltage vanishes and a positive potential appears across the sample. The light makes the sample look more positive. This is exactly what happens as the result of a thermally-induced decrease in polarisation. Thus the pyro-electric voltage is in the same direction as the photo-voltage as is experimentally observed.

In the fully screened case, the photo-emf is also the emf across the two surface layers

$$\Delta V = \left[\frac{P_o}{\epsilon_s}\frac{1}{1 + \frac{\epsilon_b}{\epsilon_s}\frac{s}{l}}\right] S \simeq .35 \text{ volts}$$

The light generated free electrons sets up a counter field which tends to cancel the bulk field E_2 ; thus, the observed voltage drop is less than it would be in a perfectly insulating medium. This is what is meant by the term ‘screening’. The counter field approaches $-E_2$. Assuming the shielding occurs only in the bulk, the total voltage across the grain is now the sum of the voltages across the surface layers.

The photo-emf is in the opposite direction to the bulk polarisation. This fact predicted in the theory is what is always observed experimentally. The complete screening of the bulk field thus would, in barium titanate, be expected to result in a photo-emf of +0.35 volts per grain or 350 V/cm and about 0.35 volts across a macroscopic single crystal. These are roughly the values actually observed as seen in **Table I**, and with the single crystal results. The linear relation between remanent polarisation and saturation photo-emf as shown in **Fig.6** is also predicted by these equations. The dependence on temperature of the photo-emf as shown in **Fig.16** and **Fig.17** is predicted by the fact that as one approaches the curie temperature, not only is P_o decreasing but the dielectric .epsilon..sub.s is increasing. The bulk internal field, E_2, should therefore decrease with temperature more rapidly than the remanent polarisation.

Screening

Solving the general problem of screening in a ferroelectric is difficult. Many of the principles involved can be demonstrated by solving a special case. The special case is meant to be particularly applicable to the $Pb(Zr_{0.53}, Ti_{0.47})0_3$ + 1% by weight of Nb_2O_5 material.

Utilised, only for simplicity, is a two dimensional model, with photo-produced carriers limited to those of a single sign. It will be assumed that these are electrons generated from deep trapping levels midway in the band gap, and that the illumination empties all the traps leaving fixed positive charges to replace the original traps. The complete emptying of a deep trapping level would produce the long wave length photo-voltages and the phenomena of an intensity saturation of the photo-emf typical of the $Pb(Zr_{0.53}, Ti_{0.47})O_3$ + 1% by weight of Nb_2O_5.

Consider a two dimensional illuminated slab of length **l** within which is an internal field **Epsilon** and within which, light generates a uniform density of electrons $\mathbf{n_o}$ (n electrons per unit length). Schematically the situation is shown in **Fig.25**, where **Phi.(x)** is the potential at a point **x**.

The carriers respond to the internal field and occupy a Boltzman distribution

$$M = M_o e^{\vartheta / KT}$$

if the fields due to the electrons could be neglected, then

$$\vartheta(e) = -\mathrm{Ex}$$

This is, of course, too rough an approximation. With n(O) the density of electrons at x=0, and n_o, the density of the immobile donor ions $m(x) = m(0)e^{\vartheta(x)/kt}$ with Phi(x) is given by Poisson's equation,

$$\frac{d^2\vartheta}{dx^2} = \frac{e}{\epsilon}[m(x) - m(0)] = \frac{e}{\epsilon}[m(0)e^{e\vartheta(x)/kT} - mo]$$

Since for Phi = 0 n(0) $=n_o$, and since all traps are emptied, assuming electrical neutrality,

$$\int_0^l mo \int_0^l e^{e\vartheta/kT} dx = m_o l$$

or

$$m_o \int_0^l e^{e\vartheta/kT} dx = m_o l$$

then

$$\int_0^l \frac{d^2\vartheta}{dx^2} dx = \int_0^l \frac{e}{\epsilon}[m(x) - m_o]\, dx = 0$$

or

$$\left.\frac{d\vartheta}{dx}\right|_{x=l} = \left.\frac{d\vartheta}{dx}\right|_{x=o}$$

If the crystal is neutral there must be no electric field at the boundary except the applied field $-E_o$

$$\left.\frac{d\vartheta}{dx}\right|_{x=o} = E_o$$

$$\left.\frac{d\vartheta}{dx}\right|_{x=l} = E_o$$

These two boundary conditions allow the solution of Poisson's equation.

$$\frac{d^2\phi}{dx^2} = \frac{m_e e}{\epsilon}\left[\frac{m(o)}{m_o}\, e^{e\phi/kT} - 1\right]$$

or

$$\frac{d^2\phi}{dx^2}\left[\frac{e\phi}{kT}\right] = \frac{M_e e^2}{\epsilon kT}\left[\frac{M(o)}{M_o}\, e^{e\phi/kT} - 1\right]$$

substituting

$$yo = \ln\frac{M_o}{m(o)} \quad y = \frac{e\phi(x)}{kT}, \quad l_D^2 = \frac{\epsilon kT}{m_e e^2}$$

we obtain,

$$\frac{d^2y}{dx^2} = \frac{1}{l_D^2}\left[e^{y-yo} - 1\right]$$

in this new notation,

$$\left.\frac{dy}{dx}\right|_{x=o} = \left.\frac{dy}{dx}\right|_{x=l} = \frac{eE_o}{kT}$$

let

$$\frac{dy}{dx} = P$$

$$\frac{d^2y}{dx^2} = \frac{dP}{dx} = \frac{dy}{dx}\,\frac{dP}{dy} = P\frac{dP}{dy} = \frac{d}{dy}\left(\frac{1}{2}P^2\right)$$

so

$$\frac{d}{dy}\left(\frac{1}{2}P^2\right) = \frac{1}{l_D^2}\, e^{y-yo} - 1$$

setting $y(o)=0$ since the zero for a potential may be set arbitrarily

$$\frac{1}{2}P^2(l) - \frac{1}{2}P^2(o) = \frac{1}{l_D^2}\int_o^l (e^{y-yo} - 1)\, dy$$

$$= \frac{1}{l_D^2}\left[e^{y(l)-y} - y(l) - e^{-yo}\right]$$

$$y(\infty) = \frac{e\phi(x)}{kT}$$

$$y(l) = \frac{e\Delta v}{kT}$$

$$\left.\frac{dy}{dx}\right|_o = \left.\frac{dy}{dx}\right|_l = P(o) = P(l)$$

so

$$O = \frac{1}{l_D^2} \quad e^{-\mu}[(e^{\Delta} - 1) - \Delta]$$

$$\text{where } \Delta = \frac{e}{kT} \Delta v \quad \text{or} \quad e^{\mu v} = \frac{e^{\Delta} - 1}{\Delta},$$

$$\text{and thus } \frac{Mo}{M(O)} = \frac{e^{\Delta} - 1}{\Delta} = \frac{e^{\frac{e}{kT} \Delta v} - 1}{\frac{e}{kT} \Delta v}$$

substituting

$$\frac{d}{dy} \left(\tfrac{1}{2} p^2\right) = \frac{1}{l_D^2} \left[\frac{\Delta}{e^{\Delta} - 1} e^{y} - 1 \right]$$

so

$$\tfrac{1}{2} p^2(x) - \tfrac{1}{2} p^2(o) = \frac{1}{l_D^2} \left[\frac{\Delta}{e^{\Delta} - 1} (e^{y} - 1) - y \right]$$

$$= \frac{1}{l_D^2} \left[\Delta \frac{e^{y} - 1}{e^{\Delta} - 1} - y \right]$$

Thus

$$\left(\frac{dy}{dx} \right)^2 = \left(\frac{eF_o}{kT} \right)^2 + \frac{2}{l_D^2} \left\{ \Delta \left(\frac{e^{y} - 1}{e^{\Delta} - 1} \right) - y \right\}$$

or

$$\frac{dy}{dx} = \frac{\sqrt{2}}{l_D}\sqrt{\Gamma^2 + \Delta\left(\frac{e^y - 1}{e^\Delta - 1}\right) - y}$$

where

$$\Gamma^2 = \frac{1}{2}\frac{e}{kT}E_o l_p$$

integrating this equation from O to l yield

$$\int_o^{y(l)} \frac{dy}{\sqrt{\Gamma^2 + \Delta\left(\frac{e^y - 1}{e^\Delta - 1}\right) - y}} = \frac{\sqrt{2}}{l_D}\int_o^l dx$$

or,

$$\int_o^\Delta \frac{dy}{\sqrt{\Gamma^2 + \Delta\left(\frac{e^y - 1}{e^\Delta - 1}\right) - y}} = \frac{\sqrt{2}l}{l_D}$$

which is an implicit expression for ΔV in terms of E_o, l, and l_D.

For low n_o and/or large E_o, Γ is large

$$\int_o^\Delta \frac{dy}{\sqrt{\Gamma^2}} = \frac{\sqrt{2}\,l}{l_D}$$

$$\Delta = \sqrt{2}\,l\,\Gamma/l_D$$

or

$$\Delta = \sqrt{2}\,l \times \frac{1}{\sqrt{2}}\left[\frac{e}{KT}E_o l_D\right] l_D$$

$$= \frac{e}{KT}E_o l$$

or

$$\Delta V = E_o l$$

which is the original potential across the bulk of the crystal.

The situation of interest is however large n_o and small l_D and small Γ^2.

It is in this situation that

$$\Delta = \frac{e}{KT}\,\Delta V$$

can be expected to vanish.

Expanding the expression for Δ small, which is always the case, then

$$\Delta\,\frac{e^y - 1}{e^\Delta - 1} - y \simeq \tfrac{1}{2}y^2 - \tfrac{1}{2}y\Delta$$

Keeping only second order terms in y and Δ, then

$$\int_o^\Delta \frac{dy}{\Gamma^2 + \frac{1}{2}(y^2 - y\Delta)} = \frac{l\sqrt{2}}{l_D}$$

Let

$$\xi = y - \frac{\Delta}{2},$$

this becomes

$$\int_{-o/2}^{\Delta/2} \frac{d\xi}{\sqrt{(\Gamma^2 - \frac{1}{8}\Delta^2) + \frac{1}{2}\xi}} = \frac{l\sqrt{2}}{l_D},$$

Setting

$$\xi = \sqrt{2\Gamma^2 - \left(\frac{\Delta}{2}\right)^2}\ ,\ \sin h\ O$$

gives

$$2\sqrt{2}\,\sin h^{-1}\left[\frac{\Delta/2}{\sqrt{2\Gamma^2 - \left(\frac{\Delta}{2}\right)^2}}\right] = \frac{\sqrt{2}\,l}{l_D}$$

or

$$\Delta = 2\sqrt{2}\ \Gamma \tan h \frac{l}{2l_D}$$

or

$$\Delta V = E_o(2l_D) \tan h \frac{l}{2l_D},$$

clearly as

$$l_D \rightarrow O\ \Delta V \rightarrow O$$

This approximation for Δ V is good for all reasonable values of T.

Illumination thus reduces the dark bulk emf = $E_o l$, producing a net photovoltage

$$V_{PHOTO} = E_o l \left[1 - \frac{\tanh \frac{l}{2l_D}}{\frac{l}{2l_D}} \right]$$

where

$$l_D = \sqrt{\frac{EkT}{m_o e^2}}$$

A simplified expression occurs for small

$$\frac{l}{2l_D}$$

where, $\tanh x \simeq x - \frac{1}{3}x^3$

$$V_{PHOTO} = E_o l \left[\frac{1}{2} \left(\frac{l}{2l_D} \right)^2 \right]$$

Here, it is clear that the photovoltage becomes insignificant for

$$\frac{l}{l_D}$$

The implication is therefore that photovoltaic contributions from the bulk will be much larger than that from the surface layers, for surface layers are extremely small white $\mathbf{l_D}$ can be estimated as very roughly equal in the bulk and the surface.

Thus, illumination will result in the vanishing of the internal field within the bulk resulting in a maximum photo-emf. $\Delta V = 0\ E_2 l$ where E_2 is the bulk field.

For small intensities, we can assume n_o small, then

$$V_{PHOTO} = E_o l \frac{1}{4} \frac{n_o e^2}{ekT}$$

i.e., the photo-voltage is proportional to n_o which can be reasonably assumed proportional to intensity which is experimentally observed (see **Fig.4**).

The model just described explains the long wave length photo-emfs, in the material $Pb(_{0.53}Zr,_{0.47}Ti)0_3$ + 1% by weight of Nb_20_5. Such a deep trapping level is probably typical of the lead titanate-lead zirconate materials with characteristic lead vacancies. These bind electrons leaving holes (producing p type dark conductivity). The addition of common dopants -- for example niobium gives rise to free electrons which combine with holes or get trapped by the lead vacancies. The doping can thus be said to provide electrons which fill traps.

It is these trapped electrons which are photo-injected into the conduction band by the long wave length light providing near maximum photo-emfs in material illuminated at 500 nm and even longer wave lengths as shown in the results plotted in **Fig.13**. Full saturation, that is the complete shielding of the bulk internal field, requires however band gap carriers which occurs as one approaches the 373 nm band gap wave length. Solving this problem, that of band gap carriers in addition to electrons generated by deep traps, can be accomplished in a manner similar to that which was accomplished for the trapped electrons but is more complex for example because mobile holes are being produced in addition to electrons and one cannot necessarily fix the maximum number of carriers.

The photo-emfs are created by photo-induced carriers shielding the bulk field. Effectively, no photo-current can flow however unless band gap light is present as is clear from the results shown in **Fig.12** and **Fig.13**. Here it is clear the band gap light produces maximum photo-emf and maximum photo-currents, less than band gap light, maximum or almost maximum photo-emf but no photo-currents and that the output resistance under these circumstances appears extremely high. Addition of band gap light allows current to flow.

The tentative explanation is that the surface layers from high resistance barriers, the magnitude of which lowers with band gap light. The surface layers thus act as intrinsic photoconductors in series with an emf. This picture not only explains the rather unique dependence of photo-emf and short circuit photo-current on wave length as shown in **Fig.12** and **Fig.13** but also the equivalent circuit which is typical of all these materials as described in **Fig.2** and as indicated by the current-voltage results in **Fig.3**.

A possible explanation for the high resistance of the surface layers is that they include quantities of charged ions which have been localised there. These are immobile under normal applied voltages moving only under the action of high fields such as produced by the reversal of the remanent polarisation. Those ions not only will occupy trapping levels, eliminating the need for easily ionised trapped electrons and thus reducing the intrinsic conductivity but also form centres for coulomb scattering of conduction electrons which should contribute markedly to the resistivity.

Efficiency

Some insight into the possible maximum efficiency of the process can be obtained by considering carriers generated by band gap light. with potential energy

$$U = 2 \int_o^L e \circ(x) \delta M_2 dx$$

with $\circ(x) < Ex$
so that a maximum value of energy

$$U = 2e \int_o^L E \delta m x dx$$

$$= e \delta m_o E L^2$$

The energy required to produce δm_o electron hole pairs

$$\epsilon = \delta M_o L E_g$$

where Eg is the band gap energy.
The power into the crystal is

$$P_m = LEg \frac{\delta M_o}{\delta t}$$

while the power out (the rate of increase in internal potential energy) is

$$P_{out} = eEL^2 \frac{\delta M_2}{\delta T}$$

The efficiency

$$= \frac{eEL}{Eg}$$

For $Pb(Zr_{.53}\ Ti_{.47})O_3 + 1$ wt% Nb_2O_5 added E is roughly 600 v/cm and the grain size roughly 5 microns. The emf across a grain is thus about .3 volts. The band gap is about 3 eV. Thus the efficiency is

$$\simeq \frac{.3}{3} \simeq 10\%,$$

Which compares with an observed band gap efficiency of about 0.06%. The calculation, of course, depends on idealising assumptions, some of which may be practically obtainable**.**

PHOTOVOLTAIC MEMORY DEVICE

With the above background and general teachings of the unique discovery of the invention now firmly in mind, numerous and important applications of the properties of the ferroelectric ceramics above-discussed are readily possible as will be evident to those skilled in this art. For example, the device of the instant invention will be shown to exhibit particular utility as a memory apparatus, thus making use of the property of the ferroelectric ceramic defined as remanent polarisation or "memory" as previously explained.

With particular reference now to **Fig.26** of the application drawings, one such photovoltaic memory apparatus is disclosed, the memory apparatus being optically addressed. In this respect, a substrate or sheet of a ferroelectric ceramic material of the type above-discussed is indicated by reference numeral 10 as being "sandwiched" between at least one pair of electrodes such as electrodes **12** and **14** positioned on opposing sides of the substrate.

In the preferred embodiment as shown, an array of electrode pairs, such as pairs **12-14** and **16-18** are disposed on opposing sides of the substrate **10** as to define a matrix configuration. Information is put into the memory and particularly into the region of the substrate **10** lying between electrode pairs by temporarily applying a voltage pulse of a predetermined polarity between the electrode pairs, such pulse being provided by the Write Pulse Generator **20** coupled to the various electrodes and of typical construction. Specifically, if a positive voltage pulse was provided by the Write Pulse Generator **20** between electrode pairs **12-14**, with electrode **12** being presumed to be the positive electrode in this example, a remanent ferroelectric polarisation will take place in the region of substrate **10** lying between the crossed electrode pair, this remanent polarisation being in a direction and of a polarity dependent upon the polarity of the write pulse.

Similarly, if a negative voltage pulse was applied between electrode 16 on the one hand, and electrode 18 on the other hand, with electrode **16** in this instance being presumed to have the negative polarity, a remanent polarisation within the ferroelectric ceramic **10** will take place in the region disposed between the intersecting or crossed electrodes **16** and **18**. In a similar fashion, predetermined remanent polarisation can be produced individually in all of the regions of the ferroelectric ceramic **10** that are disposed between crossed electrode pairs of the matrix array in direct dependence upon the polarity of the write pulse voltage applied, this remanent ferroelectric polarisation constituting stored information in that such polarisation within the ceramic will remain until removed by the application of a write voltage pulse of opposing polarity.

In accordance with the teachings of the instant invention, these stored "bits" of information in the form of remanent ferroelectric polarisation within the various regions of the substrate **10** can be extracted or "read" by selectively illuminating the poled regions of the substrate with a beam of light, as preferably can be provided by a laser, for example. Upon illumination, the polarised regions of the ferroelectric ceramic will produce a photovoltaic current and voltage at an associated electrode pair, with the polarity of the photo-current and photo-voltage being

dependent upon the "stored" remanent ferroelectric polarisation or "information" within the particular region of the substrate.

In the preferred embodiment of the device wherein a so-called matrix configuration of the electrode pairs are provided, the entire ferroelectric ceramic substrate can be scanned by the illuminating beam which is contemplated to be continuously swept in the fashion of a "light pencil" by a light beam scanner of conventional construction as is designated by reference numeral **22**, for example, light beam scanner **22** providing the sweeping illuminating beam designated by reference numeral **24**. Further, and in this particular embodiment, the illumination from the light beam **24** would be transmitted into the associated poled regions of the ferroelectric ceramic **10** by passing through electrodes **12**, **16** etc. disposed on the surface of the ceramic facing the illuminating beam, electrodes **12**, **16**, etc. being constructed so as to be transparent.

The generated photovoltaic currents and voltages at the electrode array would be detected by a synchronised detector designated by reference numeral **26** coupled to each of the electrode pairs, detector **26** being of conventional construction and serving to monitor the polarity of the photovoltaic currents and voltages developed in time synchronism with the light beam scanner **22**. Such synchronism can be effected through a direct coupling of the detector **26** to the light beam scanner **22** in typical fashion, or through the utilisation of an external computer clock, all in accordance with standardised matrix memory addressing techniques.

Optical Display Apparatus

The discovered properties of the ferroelectric ceramic substrate of the instant invention can further be applied in conjunction with liquid crystals to fabricate a novel display apparatus and, in this respect, attention is generally directed to **Fig.27** to **Fig.32** of the appended application drawings.

The operational principle associated with the fabrication of such optical displays relies upon the utilisation of the photovoltaic currents and voltages generated by substrates of a ferroelectric ceramic material to effect switching of the opacity state of a liquid crystal operating in the field - effect mode. This generalised combination will be seen to provide a write-in read-out memory and optical display. Both the liquid crystal and the ferroelectric ceramic effectively function as a memory, either in a binary or bi-stable mode having two possible states designated as an "on" state or an "off" state wherein the liquid crystal is switched from a substantially transparent condition to a substantially opaque condition, or in a multi-state mode by which the transmission characteristics of the liquid crystal are varied through many states to effect a so-called gray scale display.

With particular reference to **Fig.27** of the application drawings, a typical optical display device following the general teachings of the instant invention is shown, such display device providing so-called dark spot display capabilities. As depicted in **Fig.27**, a twisted nematic liquid crystal is designated by reference numeral **28**, such crystal being sandwiched between two transparent electrodes **30** and **32**.

As is known, the twisted nematic liquid crystal **28** will vary its transmission characteristic to incident light dependent upon the polarity and magnitude of a voltage applied across electrodes **30** and **32**. Specifically, the twisted nematic liquid crystal **28** serves to transmit illumination through it as long as there is no voltage across electrodes **30** and **32**. In conjunction with the twisted nematic liquid crystal **28**, a linear polariser **34** is provided, as is an analyser **36** of conventional construction. The linear polariser **34** and the analyser **36** are crossed so that no light passes through the combination to a diffuse reflector **38** except for the fact that the twisted nematic liquid crystal cell interposed between them rotates the polarisation of the incident illumination by 90^0 so as to allow passage of light. Application of a voltage across the cell electrodes **30** and **32** destroys the ability of the liquid crystal cell **28** to rotate the plane of the polarisation of the illumination and the illumination is consequently absorbed in the analyser **36** rather than transmitted and reflected off the diffuse reflector **38**.

Accordingly, when voltage is applied across electrodes **30** and **32**, a dark colour of the liquid cell would be displayed in so-called dark spot display. The magnitude of the display is dependent upon the magnitude of the applied voltage, such that a voltage applied across cell electrode **30** and **32** less than a characteristic amount necessary to effect full plane rotation will only partially reduce the rotating ability of the liquid crystal **28** thereby resulting in only a partial extinction of illumination and the generation of a gray-scale display. The above discussion of the operation of a so-called twisted nematic liquid crystal is entirely conventional.

To obtain the switching voltage for application to the cell electrodes **30** and **32**, a substrate of a ferroelectric ceramic designated by reference numeral **40** it utilised, the substrate **40** being sandwiched between electrodes **42** and **44** as shown, ceramic substrate **40** being disposed such that the illustrated illumination impinges not only on the liquid crystal **28**, but also on the ceramic substrate. As illustrated, electrodes **42** and **44** of the ceramic substrate **40** are respectively coupled to the transparent electrodes **30** and **32** of the twisted nematic liquid crystal cell **28**.

Initially, a polarisation voltage is applied to the ferroelectric ceramic substrate **40** across the associated electrodes **42** and **44**, such voltage being in the form of a pulse and serving to produce a remanent polarisation in the direction of the arrow shown within the substrate. Subsequently, and in accordance with the teachings of the invention, when the substrate **40** is illuminated, a current will flow in a circuit connecting terminal **42** to terminal **30** of the liquid crystal cell **28**, through the cell **28** to electrodes **32**, and then to terminals **44** of the ceramic substrate **40**, this current being a photovoltaic current proportional to the magnitude of the remanent polarisation effected within the ferroelectric ceramic by the initial application of the polarisation voltage pulse.

The magnitude of the photovoltaic current can be varied in accordance with the generalised teachings of the instant invention discussed at the outset by simply varying the magnitude of the initial polarising pulse. The so-called gray-scale display capability of the light transmission characteristics of the liquid crystal **28** is provided simply through a pre-selection of the magnitude of the remanent polarisation produced and, of course, assuming a constant intensity illumination. The memory characteristics of the ferroelectric ceramic **40** are inherently brought about in that the value of the photovoltaic current can be changed only through the application of another polarising pulse. Thus, the generalised apparatus of **Fig.27** functionally constitutes an apparatus which effects an optical display of the state of the memory within ferroelectric ceramic substrate **40**.

In the embodiment as described in **Fig.27**, a so-called "dark spot display" was effected. In the event that a so-called "bright spot" is desired to appear during the "on" state of the liquid crystal in transmission or reflection, polariser **34** and analyser **36** would be disposed in a parallel relationship with respect to one another, rather than crossed. Further, and although the basic embodiment above-discussed refers to the utilisation of liquid crystals of the twisted nematic type, similar results can be obtained with so-called colour switching crystals which, in like fashion, alter their light transmission characteristics to incident polarised light in response to the application of a voltage across them.

In accordance with the generalised teachings of **Fig.27**, various other forms of optical displays can be constructed. For example, and with particular reference to **Fig.28** of the application drawings, a different form of combined memory and optical display apparatus is illustrated, this apparatus making use of a colour switching liquid crystal 46 instead of the twisted nematic liquid crystal **28** of **Fig.27**. As was explained above, the colour switching liquid crystal such as crystal **46** serves to alter its light transmission characteristics to incident polarised light, and it is for this reason that the light source illustrated in **Fig.28** is defined as being polarised illumination, although it is to be understood that in this embodiment, as well as in the following embodiments to be discussed which use colour switching liquid crystals, a non-polarised light source can be provided if a linear polariser is disposed within the apparatus on the side of the liquid crystal nearest the incoming illumination.

The display apparatus of **Fig.28** defines a so-called monolithic structure as opposed to the exemplary structure of **Fig.27** wherein the liquid crystal was physically spaced from the energising ferroelectric ceramic. In **Fig.28**, a "sandwich" construction is provided comprising a face plate **48**, a transparent electrode **50** coupled to ground, the colour switching liquid crystal **46**, a slab or substrate of a ferroelectric ceramic **52**, and a plurality of electrodes such as electrodes **54** coupled to the ferroelectric ceramic **52** in an array.

When a short voltage pulse is initially applied between the ground electrode **50** and one of the polarity of rear electrodes **54**, the region of the liquid crystal **46** immediately in front of the rear electrode **54** will become transparent resulting in a potential appearing between the semi-transparent ground electrode **50** and the rear electrode **54** due to the incident illumination. In this instance, the ferroelectric ceramic material **52** would preferably be a transparent ceramic, such as 0.020 inch disk of 8.5/65/35 PLZT with a grain size of 6 microns, polarised in the thickness direction and producing a photo-emf of about 30 volts and a short circuit current of 10^{-7} amperes/cm.sup.2 per watt per cm.sup.2 input at 388 nm, for example. Further, the rear electrodes **54** are contemplated to be of a transparent variety, such as indium oxide 50 that a display can be provided in transmission.

A further variant of the operation of the device of **Fig.28** is possible, eliminating the necessity for the initial application of a short voltage pulse between the ground electrode **50** and one of the plurality of rear electrodes **54** to commence the process of clearing of the liquid crystal **46**. In this respect, and in addition to the normally provided uniform polarised illumination, an additional intense source of light providing a thin beam such as a laser would be provided, the laser constituting a so-called "light pencil". Upon application of the intense pencil beam of light of the apparatus of **Fig.28**, such intense light would penetrate the liquid crystal even in its nominally closed state thus illuminating the ferroelectric ceramic **52**, such illumination causing a photo-voltage to be generated as above-discussed which would then appear across the liquid crystal in the region of the intense light beam causing that region to become transparent and allowing the uniform polarised illumination to penetrate into that region, such uniform illumination further clearing the crystal in a regenerative process. This would result in a clear region which looked bright under reflected light, and a current flowing from the associated rear electrode **54** to ground, for example, through a non-illustrated resistor that would be provided. With this modification, the intense beam of

light constituting the "light pencil" can be utilised to actually enter a line drawing into the display, with a point by point read-out being provided.

As opposed to obtaining a point-by-point electrical read-out, the image written-in by the "light pencil" can be externally projected. In this respect, and as explained, the "image" constitutes transparent sections of the liquid crystal. If a light source such as a tungsten-halogen lamp normally associated with projectors was additionally provided to illuminate the display apparatus from the "rear" thereof in a direction opposing the direction of the incident polarised illumination, such auxiliary light source would pass through the display apparatus at the transparent regions, much in the same manner as a photographic slide is projected, the projection image being displayed on a suitable screen. In this instance, of course, a ferroelectric ceramic material that is transparent would be required, such as the material known as PLZT 7/65/35.

As can further be appreciated, the memory characteristics of the optical display of **Fig.28** are not permanent. If domain switching and a permanent memory capability is desired, an alternative electrode configuration would be required in the fashion illustrated in **Fig.30** of the application drawings, components of the apparatus of **Fig.30** that are the same as those of **Fig.28** being represented by the same reference numerals. Specifically, an additional transparent electrode **56** would be disposed between the colour switching liquid crystal **46** and the ferroelectric ceramic **52** polarisation within the ferroelectric ceramic **52** being effected by the application of a voltage pulse across electrodes **54** and **56**, and with an additional grounding electrode **52** being provided on the ceramic **52** as is shown so as to couple one end of the ferroelectric ceramic **52** to the transparent electrode **50**.

If a twisted nematic liquid crystal were desired to be utilised in the generalised configuration of the optical display of **Fig.28**, a still further modification of the electrode arrangement would be needed and, in this respect, attention is directed to **Fig.29** of the application drawings. Like parts in this figure are again represented by the same reference numerals.

Initially, since a twisted nematic liquid crystals alters its light transmissions characteristics by rotating the plane of the polarisation of the illumination, a further polariser such as analyser **60** is required to be disposed between the ferroelectric ceramic **52** and the liquid crystal **46**, the crystal **46** thereby being properly responsive to incoming polarised illumination either provided directly by a polarised source, or provided through the utilisation of a non-polarised illumination source in conjunction with a polariser such as polariser **34** of the embodiment of **Fig.27**. Additionally, a light transmitting electrode **62** would be disposed on the surface of the analyser **60** immediately adjacent the liquid crystal **46**, transparent electrode 62 being coupled through the analyser and the ferroelectric ceramic substrate **52** to an associated rear electrode 54. Each of the rear electrodes **54** of the array would have associated therewith an additional transparent electrode **62** in similar manner.

If the analyser **60** was constructed to be crossed with the incoming polarised illumination, the liquid crystal **46** would normally transmit light through it and, upon the application of a voltage between electrode **54** and the front transparent electrode **50**, would cause the apparatus to provide a so-called "dark spot display." Alternatively, if the incoming polarised light has a plane of polarisation parallel to the polarisation plane of analyser **60**, a so-called "bright spot display" would result. It should further be appreciated that the embodiment of **Fig.29** can be utilised with a "light pencil" to provide a functional operation similar to that discussed with respect to **Fig.28**.

Attention is now directed to **Fig.31** of the application drawings wherein an illustration is provided of an optical display array utilising a liquid crystal **64** of the colour switching type. Each of the units shown is contemplated to represent one of the horizontal row in an overall array. The structure illustrates is in monolithic form and, as shown, constitutes a polarity of superposed layers. Specifically, a transparent electrode **66** is provided, behind which is the liquid crystal **64** disposed between two face plates **68** and **70**. A transparent electrode structure **72** is provided imbedded at one end with the liquid crystal **64** and coupled at the other end to one end of the ferroelectric ceramic substrate **74** as is shown. The other end of each ferroelectric ceramic slab **74** is commonly coupled to ground along with the front transparent electrode 66 as was discussed.

With the embodiment of **Fig.31**, each ferroelectric ceramic substrate **74** would be initially polarised by the application of a polarising voltage pulse between the representative terminals or electrodes **76** and **78**, for example. Now, upon the application of illumination to the ferroelectric ceramic, a photovoltaic voltage will be generated which appears between the front transparent electrode **66** and the rear transparent electrode **72** causing the liquid crystal **64** between these electrodes to become transparent.

Liquid crystal **64** would normally be in a nominally opaque state. However, sufficient light would be transmitted through the liquid crystal material so as to produce the photo-voltage in the ferroelectric ceramic **74**, which photo-voltage applied to the electrodes **66** and **72** in a positive feed-back arrangement serves to increase the transparency of the colour switching liquid crystal **64** in the region between the electrodes. This increased transparency, in turn, increases the voltage output of the ferroelectric material **74** which further increases the transparency of the liquid crystal **64** such that a transparent region would be formed appearing as a bright spot

with reflected light. The surface of the ferroelectric ceramic **74** would in this instance serve itself as a diffuse reflector which would be required by a display function in the reflection mode.

Further, it should be appreciated that a certain threshold light transmission of the liquid crystal **64** would be required to begin this process of creating a transparent region. If the liquid crystal is sufficiently thick, the transmitted light through the crystal in its normally opaque state would be insufficient to commence this clearing process and an applied voltage would be initially necessary across the crystal to commence the process, this voltage being used as a "read" signal.

As can be appreciated, the remanent polarisation of the ferroelectric ceramic material **74** in the embodiment depicted in **Fig.31** is along the length of the ceramic substrate. An alternate arrangement is possible wherein the memory writing is accomplished by altering the remanent polarisation of the ferroelectric ceramic in the thickness direction. In this respect, reference is once again made to **Fig.30** of the application drawings illustrating the disposition of a ferroelectric ceramic **52** in conjunction with the colour switching liquid crystal **46** such that the remanent polarisation of the ceramic is achieved in the thickness direction, and such that permanent memory characteristics are imparted. With this arrangement, the incident illumination would be quickly absorbed in the surface of the ferroelectric ceramic material but would still penetrate sufficiently so as to produce relatively large photovoltaic voltages.

Finally, the optical display device of **Fig.31** can be constructed with a twisted nematic liquid crystal as opposed to the colour switching liquid crystal of **Fig.31** and attention is herein directed to **Fig.32** of the application drawings. Again, components of the apparatus of **Fig.32** which are similar to those in **Fig.31** are represented by the same reference numeral.

In this embodiment, a polariser **80** would initially be provided so as to polarise the incoming illumination. In a fashion similar to the generalised embodiment of **Fig.27**, an analyser **82** would likewise be provided, polariser **80** and analyser **82** being assumed to be parallelly disposed. Incoming polarised light will not impinge on the ferroelectric ceramic material **74** because the twisted nematic crystal **64** would rotate the plane of the polarisation of the illumination by 90^0 and such illumination would thus be absorbed in analyser **82**. The display unit, accordingly, would initially be in an "off" or dark state and no voltage would exist across the terminals or electrodes **76** and **78** of the ferroelectric ceramic.

The "on" of the display apparatus would be bright under reflected illumination and would be indicated by the appearance of a DC voltage across terminals **76** and **78**. The unit would be switched to the "on" stage through the application of an initial polarising voltage pulse between electrodes **76** and **78**. The twisted nematic liquid crystal would now lose its ability to rotate the plane of polarisation of the illumination and light would fall on the surface of the now-polarised ferroelectric ceramic material **74** such that the ceramic would generate a steady, high photovoltaic voltage which would appear across the electrodes of the liquid crystal. This photovoltaic voltage would prevent the liquid crystal from returning to the twisted phase and the liquid crystal would thus remain transparent and a voltage potential would be maintained across the electrodes for the duration of the illumination.

The display apparatus can be returned to its dark state simply by shorting across terminals **76** and **78** and the crystal cell would return to its opaque condition with no voltage appearing across the electrodes. A new external voltage pulse would be required across electrodes **76** and **78** to again switch the unit on. It should be appreciated that only a momentary voltage pulse is required to turn the display unit on, and only a momentary short circuit is needed to turn the unit off.

If the incident illumination were interrupted, the display unit would likewise be put into an "off" state. The memory characteristics of the display apparatus thus are volatile in the sense that a removal of illumination will put the display unit into an "off" state. Permanent memory characteristics can be obtained by depoling the ferroelectric ceramic **74** with additional circuitry and the illumination could then be interrupted. When illumination is restored, a voltage pulse would switch "on" only those units of the array which were in an "on" state at the time of interruption of illumination, since only the polarised ferroelectric ceramic units will produce a photo-voltage. The depoled units can then be repoled without switching them "on", utilising a suitable circuit to apply a polarising voltage to the ceramic but not to the liquid cell to therefore retain the liquid crystal cell in its dark state as it was at the time the illumination was removed.

Many other different embodiments combining a liquid crystal display with the ferroelectric ceramic substrate of the instant invention can be fabricated along the generalised teachings referred to above. From the standpoint of materials selection, PLZT is desired when a transparent ferroelectric ceramic is required, and other ferroelectric ceramics such as $Pb(Zr_{0.53}Ti_{0.47})O_3$ + 1% by weight of Nb_2O_5 (i.e. PZT-5), a solid solution of lead titanate, and lead zirconate can be utilised when relatively cheap "opaque" materials are acceptable. With the display devices as above-discussed, typical thickness of the ferroelectric ceramic material are on the order of 0.020 inches. In accordance with the generalised teachings appearing at the outset of this specification, it is to be appreciated that

the photovoltaic output of the ferroelectric ceramic material is proportional to the material length and, the higher the photovoltaic output, the faster the switching time of the associated liquid crystal.

A further form of optical display apparatus is contemplated herein by which the previously discussed photoconductive properties of ferroelectric ceramic materials are utilised in the formation of display apparatus. As will be recalled and appreciated, the resistivity of typical ferroelectric ceramic materials varies as a function of the illumination incident thereon and thus, the voltage drop across illuminated regions of a ferroelectric ceramic substrate that has a polarising voltage applied thereto would be less than the voltage drop across non-illuminated or dark regions of the ceramic. Attention in this respect is directed to **Fig.33** of the application drawings.

The display device depicted in **Fig.33** is such that a photograph in the form of a projected image can be stored in a ferroelectric ceramic sheet or substrate **84** as a pattern of poled ferroelectric regions where the remanent polarisation of such regions is simply related to the intensity of the projected image at that point. The pattern of poled regions can be produced by the already discussed technique of a photoconductive ferroelectric sandwich, or by utilising the photoconductive properties of ferroelectric materials directly.

In the embodiment of **Fig.33**, an image is projected onto a ferroelectric-photoconductive substrate **84**, which substrate is backed by a sheet of resistive material **86** such as evaporated carbon, semiconductor material or the like. A transparent front electrode **88** forming a ground plane covers the surface of the ferroelectric material **84**, which material is of the type which would exhibit a sizable polarisation dependent photovoltaic effect. A further electrode **90**, covers the rear surface of the resistive material **86**, and a polarising voltage would be applied to the apparatus between electrodes **90** and **83**.

With such an arrangement the voltage drop will be seen to exist across those regions of the ferroelectric substrate **84** which are illuminated will be less than the voltage drop apparent across the non-illuminated or dark regions. As such the lower remanent polarisation within the ferroelectric material will be effected than in those regions of the ferroelectric material that are not illuminated by the projected image. Accordingly a "negative" of the projected image would thus be stored in the ferroelectric substrate or sheet **84** as regions of varying remanent polarisation. In that the ferroelectric **84** is photovoltaic having polarisation dependent photo-voltages as discussed this stored image is now read out electrically utilising the techniques already described with respect to the embodiments of the invention illustrated in **Fig.26** of the application drawings or **Fig.28** et. seq. of the application drawings. It is displayed by applying the photo-voltages from regions of polarisation in which the image is effectively stored to liquid crystal electrodes as for example is illustrated in **Fig.34** of the application drawings where illumination sufficiently strong penetrates the dark liquid crystal **93**, to in a regenerative fashion, apply the photo-voltage from polarised region **91**, to the liquid crystal region immediately adjacent varying in intensity depending on the value of the polarisation. A negative image is produced in reflection.

High Voltage Battery

The teaching in this patent may be applied toward the provision of a novel high voltage battery serving to convert radiation such as X-radiation in this instance, directly into electrical energy. In this respect, a block or substrate of ferroelectric ceramic material would again be provided to which electrodes are attached in the identical fashion as was discussed with respect to the basic physical configuration of the invention illustrated in **Fig.1** of the application drawings. An example of the constituent material of the ferroelectric ceramic in this instance is solid solution PZT-5A consisting of 53 mole percent $ZrTiO_3$ and 47 mole $PbTiO_3$ with 1 percent by weight of niobium added such as Nb_2O_5. This ferroelectric ceramic material would be poled in the usual fashion by the application of a high voltage applied across the electrodes.

To function as a battery, the ceramic material can contain a radioactive component and this can be all or a portion of any of the above-discussed constituent elements. For example, the material may be fabricated with a radioactive isotope of Zr,TiO, Nb, etc., or a radioactive additive can be added to the composition. Alternatively, the composition may be placed next to a strong radioactive source and, for example, could actually be coated with a radioactive material. The primary requirement is that a flux of gamma rays or X-rays within the material be produced, which radiation has the effect of ionising the ferroelectric ceramic material so as to produce non-equilibrium carriers.

Thus, in the instance of the application of a poled ferroelectric ceramic material as a high voltage battery, an external light source would not be required as the ionising source in that the non-equilibrium carriers would be produced by the internal ionisation of the ferroelectric ceramic material effected by the radiation and would result in an emf which would appear across the electrodes.

Accordingly, an open circuit voltage proportional to the length of the ferroelectric ceramic material between the electrodes and inversely proportional to average grain size, and the like as was discussed at the outset of this

specification would be produced by the gamma or X-radiation. Similarly, a short circuit current proportional to the electrode area and the net (steady state) increment of excess carriers introduced into the conduction band would likewise be produced, this being related to the intensity of the ionising radiation.

As can be appreciated, the emf would persist as long as the ionising radiation persisted and, extrapolating from the detailed photo-effect results, the emf produced by this high voltage battery would be relatively independent of the intensity of the radiation and thus not strongly dependent on the half-life of the radioactive material.

While there has been shown and described several preferred embodiments and applications of the basic invention hereof, those skilled in the art should appreciate that such embodiments are exemplary and not limiting and are to be construed within the scope of the following claims:

CLAIMS

1. A photovoltaic memory apparatus comprising: a substrate of a ferroelectric ceramic; means for selectively applying a voltage pulse of a predetermined polarity across a region of said substrate to thereby effect a remanent ferroelectric polarisation in said region of said substrate representative of the information to be stored; means for selectively illuminating said poled region of said substrate with a source of radiation, whereby a photovoltaic voltage is produced at said region of a polarity dependent upon said predetermined polarity of said polarising voltage pulse; and means for detecting said photovoltaic voltage whereby the stored information is retrieved.

2. A memory apparatus as defined in claim 1, wherein an array of electrode pairs are disposed on opposing sides of said substrate to define a matrix configuration of poled regions, said polarising voltage pulse being applied across selected electrode pairs, and wherein said information reading means scans said matrix configuration in accordance with a desired pattern, said detecting means being coupled to said array of electrode pairs and being synchronised with said information reading means.

3. A memory apparatus as defined in claim 1, wherein said substrate is sandwiched between at least one electrode pair and one electrode of said electrode pair is transparent such that said illumination from said information reading means passes through it into said respective poled region of said substrate.

4. A method of addressing and storing information utilising a substrate of a ferroelectric ceramic as a memory core, said method comprising the steps of initially effecting a remanent electrical polarisation in regions of the ferroelectric ceramic by the application of a voltage pulse across the regions of the substrate, the voltage pulse having at least one of a polarity and magnitude representative of the information to be stored; addressing the memory core while illuminating the polarised regions of the ferroelectric ceramic substrate with a source of radiation; and detecting at least one of the polarity and magnitude of the photovoltaic current and voltage produced by such illumination upon the polarised regions, the polarity and magnitude being dependent upon the polarity and magnitude of the initial polarising voltage pulse whereby the stored information is recovered.

5. An optical apparatus comprising in combination: an electro-optic means providing variable light transmission characteristics in response to the magnitude and polarity of an applied voltage; a substrate of a ferroelectric ceramic; means for applying a polarising voltage pulse of a predetermined magnitude and polarity across said substrate to effect a remanent electrical polarisation within said substrate; means for illuminating said electro-optic means and said ceramic substrate, illumination impinging upon said substrate effecting the generation by said substrate of a photovoltaic current and voltage having a polarity dependent upon the polarity of said polarising voltage pulse; and means for applying said generated photovoltaic voltage to said electro-optic means, whereby the transmission characteristics of said electro-optic means to the illumination impinging thereon is varied to effect a visual display.

6. A display apparatus as defined in claim 5, wherein the light transmission characteristics of said electro-optic means is switched from a relatively low opacity to a relatively high opacity upon application thereto of said generated photovoltaic voltage.

7. A display apparatus as defined in claim 5, wherein the light transmission characteristics of said electro-optic means is switched from a relatively high opacity to a relatively low opacity upon application thereto of said generated photovoltaic voltage.

8. A display apparatus as defined in claim 6, wherein said relatively low opacity is of a value such that said electro-optic means is substantially transparent, said relatively high opacity being of a value such that said electro-optic means is substantially opaque.

9. A display apparatus as defined in claim 7, wherein said relatively low opacity is of a value such that said electro-optic means is substantially transparent, said relatively high opacity being of a value such that said electro-optic means is substantially opaque.

10. A display apparatus as defined in claim 5, wherein the magnitude of said polarising voltage is selected such that the light transmission characteristics of said electro-optic means is switched between varying opacities to define a gray scale.

11. A display apparatus as defined in claim 33, wherein said electro-optic means is a liquid crystal of the twisted nematic type.

12. A display apparatus as defined in claim 5, wherein said electro-optic means is a liquid crystal of the colour switching type.

13. A display apparatus as defined in claim 11, wherein said liquid crystal is sandwiched between a light polariser and a light analyser.

14. A display apparatus as defined in claim 5, wherein said electro-optic means and said ferroelectric ceramic substrate are disposed in superposition to define a monolithic structure.

15. A display apparatus as defined in claim 14, wherein said electro-optic means is a colour switching liquid crystal disposed in superposition with said ceramic substrate to define a monolithic structure, and wherein said means for applying a polarising voltage to said substrate and said means for applying said photovoltaic voltage to said liquid crystal comprises a plurality of electrodes disposed on opposite faces of said structure with said structure being sandwiched between them, at least one electrode pair being in contact with said liquid crystal and with said ceramic substrate, respectively; said electrode of said pair which is in contact with said liquid crystal being transparent.

16. A display apparatus as defined in claim 14, wherein said monolithic structure constitutes a plurality of stacked superposed layers comprising a first transparent electrode, an electro-optic means, a second transparent electrode, said substrate of a ferroelectric ceramic, and a third electrode, said third electrode being coupled to said first electrode, said means for applying said polarising voltage being defined by said second and third electrodes, said means for applying said generated photovoltaic voltage being defined by said first and second electrodes, and wherein said means for illuminating said electro-optic means and said substrate comprises a light beam directed to impinge upon said first transparent electrode.

17. A display apparatus as defined in claim 16, wherein said electro-optic means is a liquid crystal of the colour switching type.

18. A display apparatus as defined in claim 16, wherein said plurality of stacked layers further includes a polariser disposed over said first transparent electrode, and an analyser disposed between said second transparent electrode and said ceramic substrate, said electro-optic means being a liquid crystal of the twisted nematic type.

19. A display apparatus as defined in claim 16, wherein said illumination means comprises a source of polarised light, said plurality of stacked layers including an analyser disposed between said second transparent electrode and said ceramic substrate, said electro-optic means being a liquid crystal of the twisted nematic type.

20. A display apparatus as defined in claim 19, wherein said analyser is disposed in a direction parallel to the plane of polarisation of the incident illumination.

21. A display apparatus as defined in claim 19, wherein said analyser is disposed so as to be crossed with respect to the plane of polarisation of the incident illumination.

22. A method of electrically storing optical information comprising the steps of: projecting an image constituting the optical information onto a sandwich of a ferroelectric ceramic backed by a layer of resistive material to form an illumination pattern thereon; applying a voltage pulse across the sandwich whereby varying remanent polarisations within the ferroelectric ceramic are produced in dependence upon the illumination pattern.

23. The method of claim 22, further including the step of reading out the remanent polarisations to thereby extract the stored optical information.

24. A display apparatus as defined in claim 5, wherein said variation of the transmission characteristics of the electro-optic means ensures that illumination continues to impinge upon said substrate to latch said electro-optical means and maintain said transmission variation thereof.

25. A method of electrically storing optical information comprising the steps of: projecting an image constituting the optical information onto a ferroelectric ceramic layer to form an illumination pattern thereon and thereby alter the resistivity of the ceramic layer in accordance with said pattern; applying a voltage pulse across the ceramic whereby varying remanent polarisations within the ferroelectric ceramic are produced in dependence upon the illumination pattern.

Patent US 6,246,561 **12th June 2001** **Inventor: Charles J. Flynn**

METHODS FOR CONTROLLING THE PATH OF MAGNETIC FLUX FROM A PERMANENT MAGNET AND DEVICES INCORPORATING THE SAME

This patent covers a device which is claimed to have a greater output power than the input power required to run it.

ABSTRACT

A permanent magnet device includes a permanent magnet having north and south pole faces with a first pole piece positioned adjacent one pole face thereof and a second pole piece positioned adjacent the other pole face thereof so as to create at least two potential magnetic flux paths. A first control coil is positioned along one flux path and a second control coil is positioned along the other flux path, each coil being connected to a control circuit for controlling the energisation thereof. The control coils may be energised in a variety of ways to achieved desirable motive and static devices, including linear reciprocating devices, linear motion devices, rotary motion devices and power conversion.

DESCRIPTION

FIELD OF THE INVENTION

This invention relates generally to permanent magnet devices and more particularly, to a permanent magnet control component in which the flow of flux from a permanent magnet is controlled between two or more flux paths by utilising timed delivery of electrical signals through one or more coils placed along at least one of the flux paths. Such permanent magnet control components may take on a variety of configurations facilitating use of such components in a variety of applications including applications involving the production of reciprocating, linear, and rotary motion and power conversion. Several novel permanent magnet rotary motion devices of motor constructions which operate by controlling the path of magnetic flux from one or more permanent magnets are described, such permanent magnet rotary motor constructions having increased efficiency and more desirable torque characteristics as compared to many currently used motors.

BACKGROUND OF THE INVENTION

Magnetic force of attraction is commonly used in a variety of types of permanent magnet devices including both linear and rotary motors. In the field of such permanent magnet devices there is a continuous pursuit of increased efficiency and reduced complexity.

Accordingly, an object of the present invention is to provide a permanent magnet control component in which the path of a given level of permanent magnet flux can be controlled by a lesser level of electromagnetic flux.

Another object of the present invention is to provide a permanent magnet control component in which substantially all of the flux from a permanent magnet can be switched between at least two different flux paths of the permanent magnet control component so as to enable useful work in the form of linear, reciprocating, and rotary motion.

Still another object of the present invention is to provide permanent magnet control components and motor constructions in which flux path control is provided by energising an 10 electromagnet to oppose the magnetic flux of one or more permanent magnets.

Another object of the present invention is to provide permanent magnet control components and motor constructions in which flux path control is provided by energising an electromagnet to aid the magnetic flux of one or more permanent magnets.

Yet another object of the present invention is to provide permanent magnet motor 15 constructions with improved operating characteristics.

SUMMARY OF THE INVENTION

These and other objects of the invention are attained by an apparatus which, in one aspect, is a permanent magnet device, comprising a permanent magnet having north and south pole faces, a first pole piece, a second pole piece, a first control coil, a second control coil, and circuit means, the first pole piece positioned adjacent the north pole face of the permanent magnet and including a first path portion, a second path portion and a third portion, the first path portion extending beyond a perimeter of the north pole face and the second path portion extending beyond the perimeter of the north pole face to define first and second flux paths for magnetic flux emanating from the north pole face of the permanent magnet, the first path portion of the first pole piece connected to the second path portion of the first pole piece by the third portion which extends across the north pole face of the permanent magnet, the second pole piece positioned adjacent the south pole face and including a first path portion and a second path portion, the first path portion extending beyond a perimeter of the south pole face and substantially aligned with the first path portion of the first pole piece, the second path portion extending beyond the perimeter of the south pole face and substantially aligned with the second path portion of the first pole piece, the first control coil positioned around the first path portion of the first pole piece, the second control coil positioned around the second path portion of the first pole piece, the circuit means connected to each of the first control coil and the second control coil to alternately energise the first coil and the second coil in a timed sequential manner.

Another aspect of the present invention provides a method for controlling the path of magnetic flux from a permanent magnet which involves placing a first pole piece adjacent a first pole face of the permanent magnet so as to have at least first and second path portions extending beyond a perimeter of the first pole face. A second pole piece is placed adjacent a second pole face of the permanent magnet so as to include at least one portion which substantially aligns with the first and second path portions of the first pole piece. A first control coil is placed along and around the first path portion of the first pole piece and a second control coil is placed along and around the second path portion of the first pole piece. The first control coil is repeatedly energised in a permanent magnet magnetic flux opposing manner so as to prevent magnetic flux of the permanent magnet from traversing the first path portion of the first pole piece, and the second control coil is repeatedly energised in a permanent magnet magnetic flux opposing manner so as to prevent magnetic flux of the permanent magnet from traversing the second path portion of the first pole piece.

Yet another aspect of the present invention provides a method for controlling the path of magnetic flux from a permanent magnet by placing a first pole piece adjacent a first pole face of the permanent magnet so as to have at least first and second path portions extending beyond a perimeter of the first pole face. A second pole piece is placed adjacent a second pole face of the permanent magnet so as to include at least one portion which substantially aligns with the first and second path portions of the first pole piece. A first control coil is placed along and around the first path portion of the first pole piece, and a second control coil is placed along and around the second path portion of the first pole piece. The following steps are alternately performed in a repeated manner:

(i) energising the first control coil in a permanent magnet magnetic flux aiding manner so as to couple with substantially all magnetic flux of the permanent magnet such that substantially no magnetic flux of the permanent magnet traverses the second path portion of the first pole piece when the first control coil is so energised; and

(ii) energising the second control coil in a permanent magnet magnetic flux opposing manner so as to couple with substantially all magnetic flux of the permanent magnet such that substantially no magnetic flux of the permanent magnet traverses the first path portion of the first pole piece when the second control coil is so energised.

A further aspect of the present invention provides method for controlling the path of magnetic flux from a permanent magnet by placing a first pole piece adjacent a first pole face of the permanent magnet so as to have at least first and second path portions extending beyond a perimeter of the first pole face, and placing a second pole piece adjacent a second pole face of the permanent magnet so as to include at least one portion which substantially aligns with the first and second path portions of the first pole piece. A first control coil is placed along and around the first path portion of the first pole piece, and a second control coil is placed along and around the second path portion of the first pole piece. The following steps are alternately performed in a repeated manner:

(i) energising the first control coil in a permanent magnet magnetic flux aiding manner so as to couple with substantially all magnetic flux of the permanent magnet such that substantially no magnetic flux of the permanent magnet traverses the second path portion of the first pole piece when the first control coil is so energised; and

(ii) energising the second control coil in a permanent magnet magnetic flux opposing manner so as to couple with substantially all magnetic flux of the permanent magnet such that substantially no magnetic flux of the permanent magnet traverses the first path portion of the first pole piece when the second control coil is so energised.

BRIEF DESCRIPTION OF THE INVENTION

For a better understanding of the present invention reference may be made to the accompanying drawings in which:

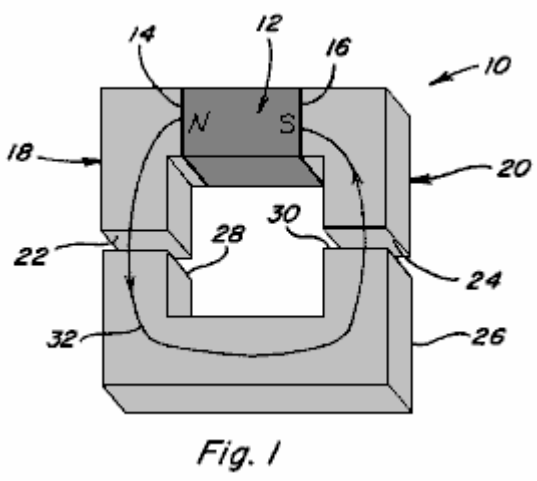


Fig. 1

Fig.1 is a perspective view of a magnetic device in which the magnetic flux from a magnetic member traverse a single path to produce a coupling force;

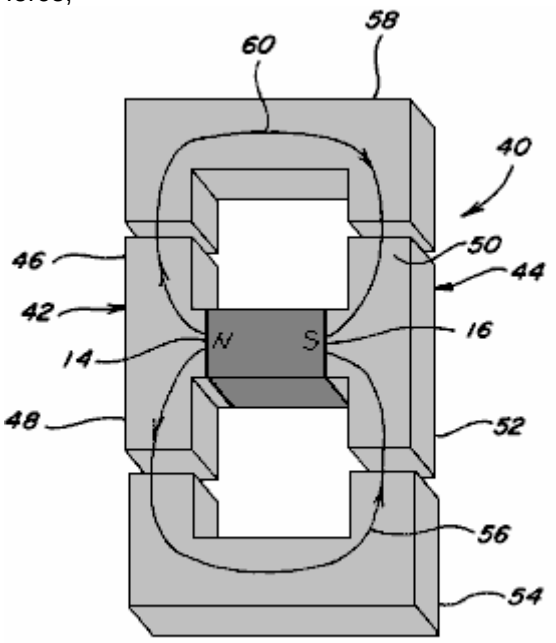


Fig. 2

Fig.2 is a perspective view of a magnetic device in which the magnetic flux from a magnetic member splits between two paths;

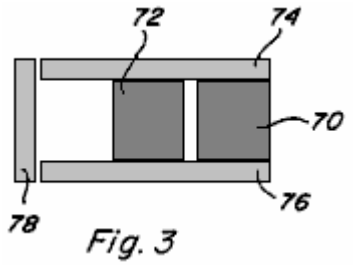


Fig. 3

Fig.3 is a side view of two magnetic members arrange in parallel between pole pieces;

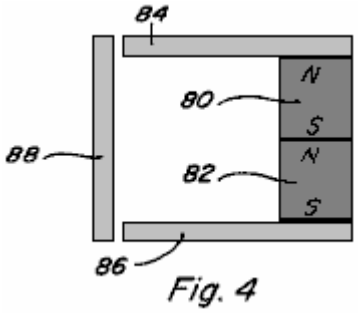


Fig. 4

Fig.4 is a side view of two magnetic members arranged in series between pole pieces;

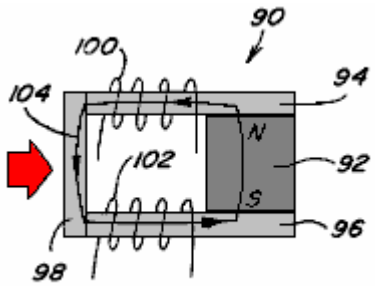


Fig. 5

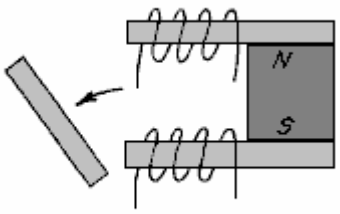


Fig. 6

Fig.5 and **Fig.6** are side views of a permanent magnet device including a permanent magnet having pole pieces positioned against the pole faces thereof and including a movable armature;

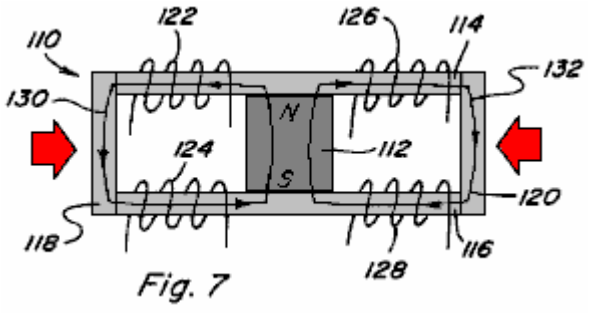


Fig. 7

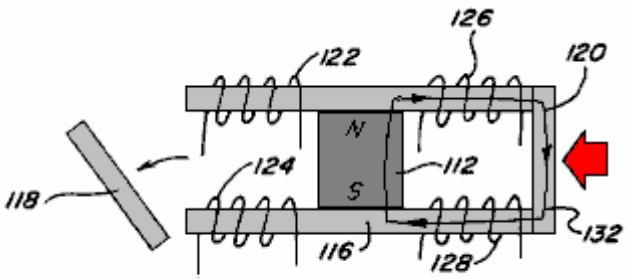


Fig. 8

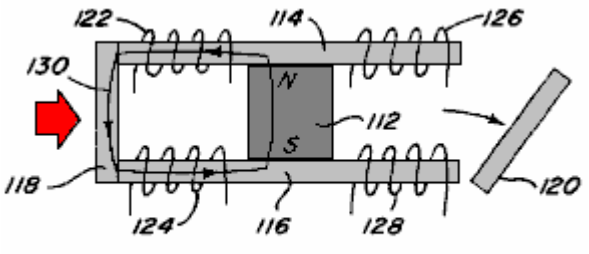


Fig. 9

Fig.7, **Fig.8** and **Fig.9** are side views of a permanent magnet device including a permanent magnet having pole pieces positioned against the pole faces thereof to provide two magnetic flux paths and including a movable armature which can be positioned along each magnetic flux path;

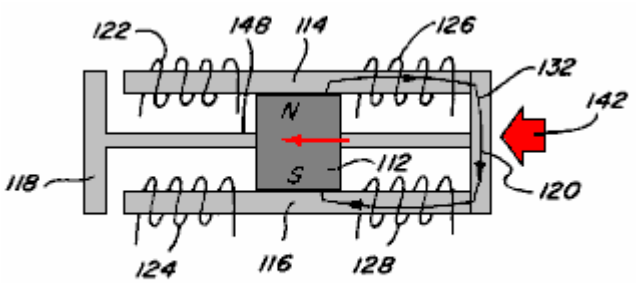


Fig. 10

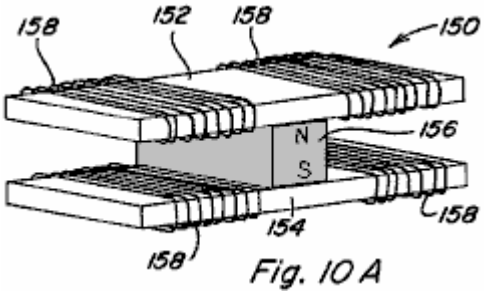


Fig. 10 A

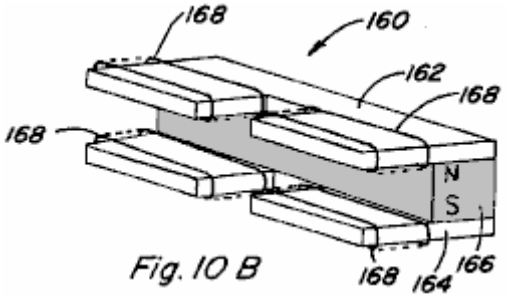


Fig. 10 B

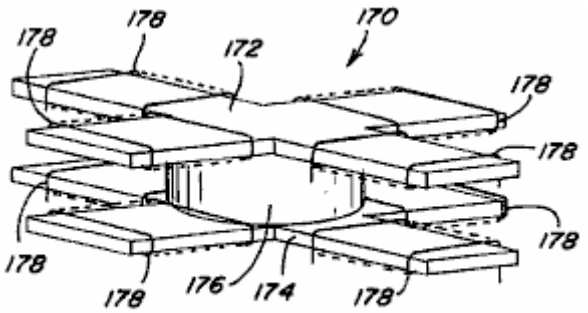


Fig. 10 C

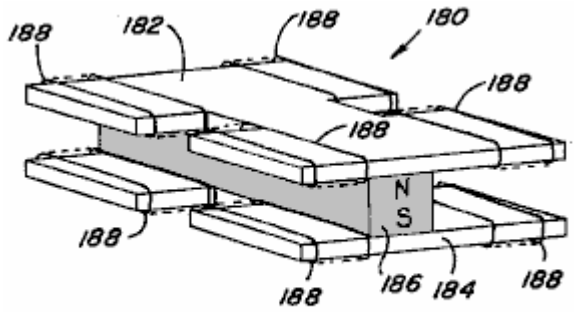


Fig. 10 D

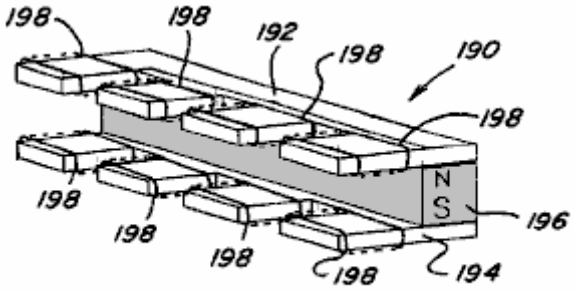


Fig. 10 E

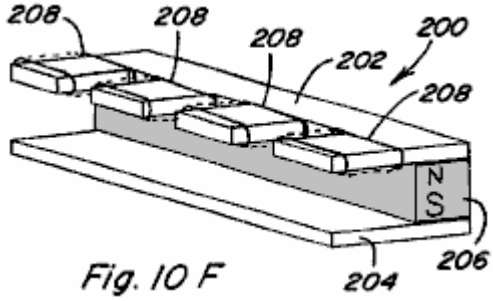


Fig. 10 F

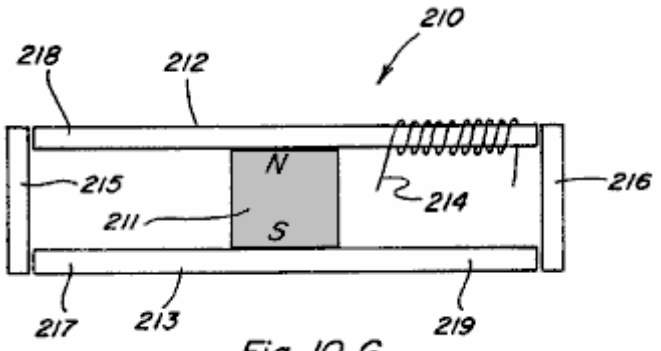


Fig. 10 G

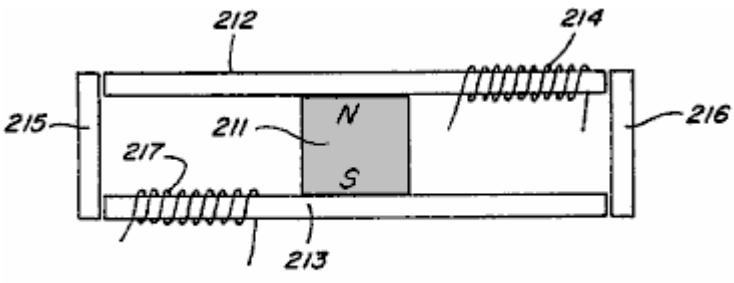


Fig. 10 H

Figs.10, **10A-10H** are perspective views of various embodiments of permanent magnet **5** control components which include two or more magnetic flux paths;

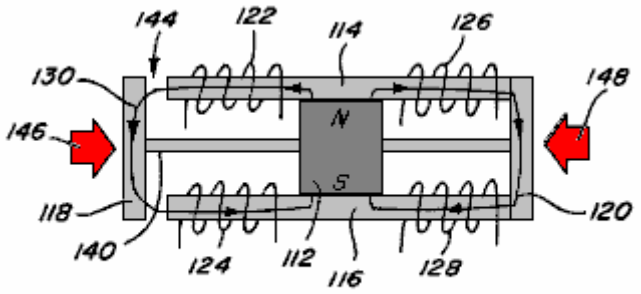


Fig. 11

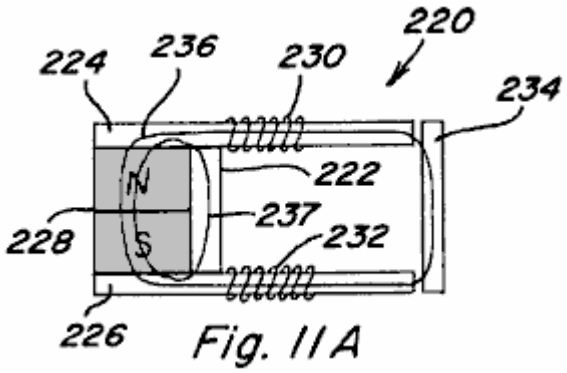


Fig. 11A

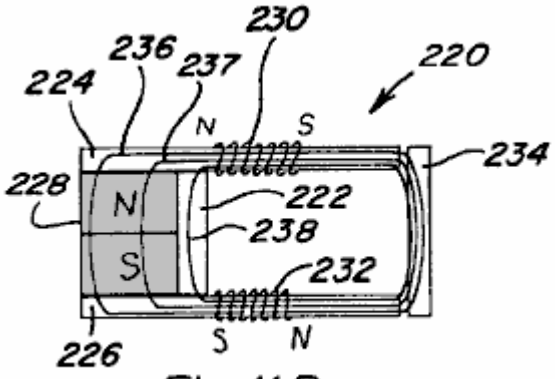


Fig. 11B

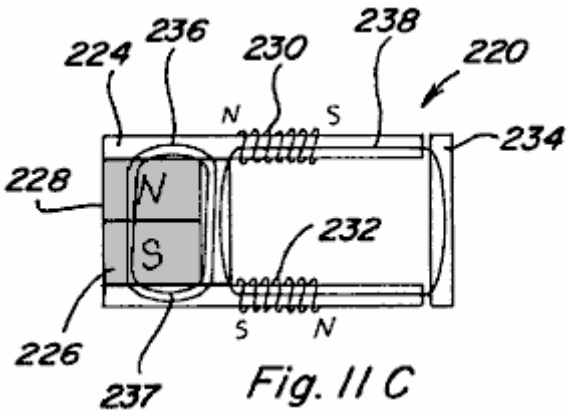


Fig. 11C

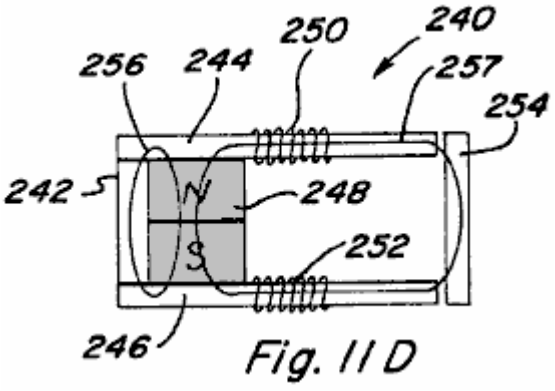


Fig. 11 D

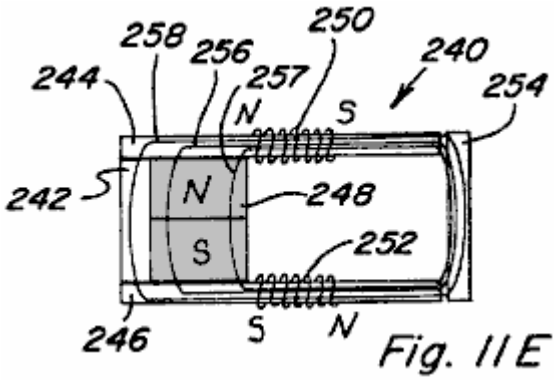


Fig. 11 E

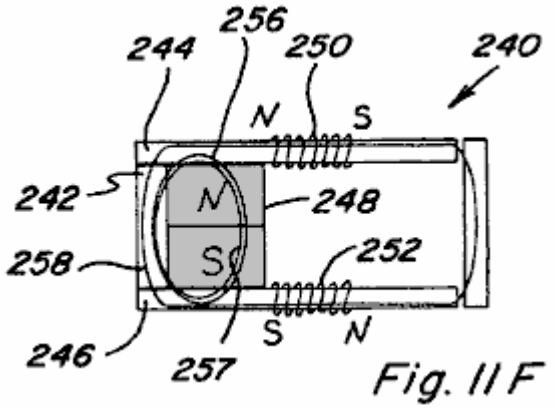


Fig. 11 F

Figs.11, **11A-11F** are side views of a permanent magnet device including a permanent magnet having pole pieces positioned against the pole faces thereof and including a movable armature and a permanent bypass extending between the pole pieces;

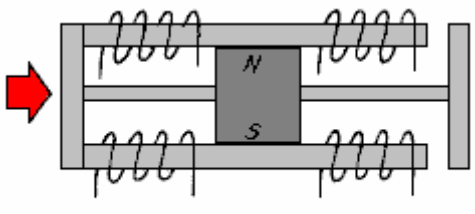


Fig. 12

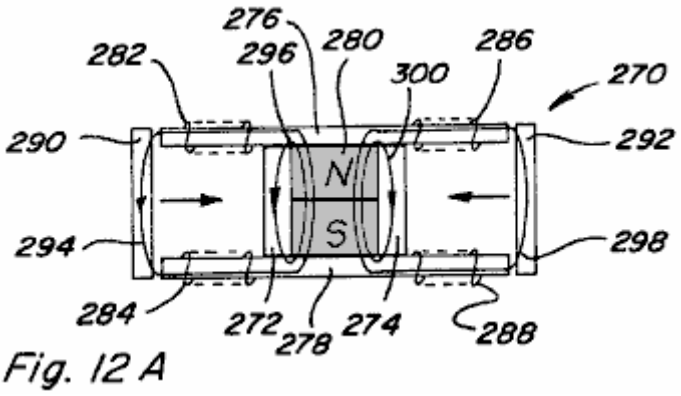


Fig. 12 A

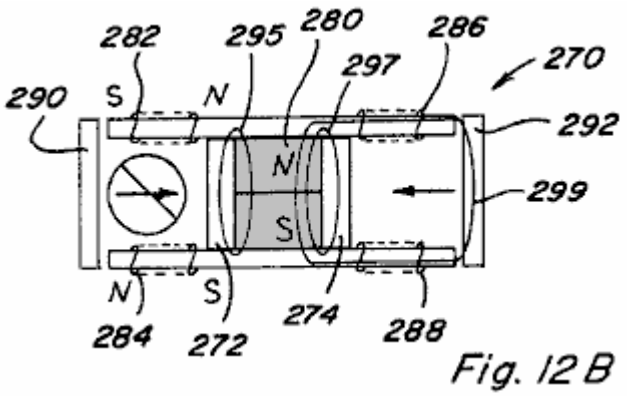


Fig. 12 B

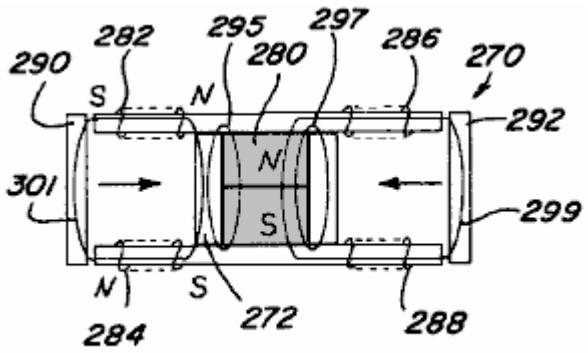


Fig. 12 C

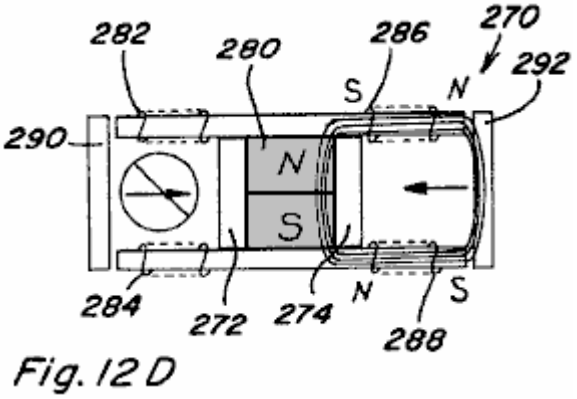


Fig. 12 D

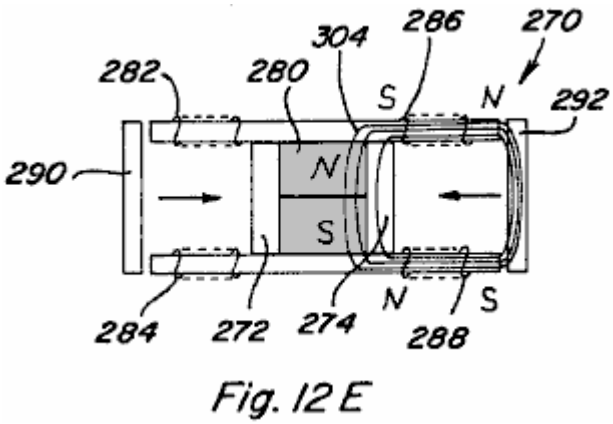


Fig. 12 E

Figs.12, **12A-12E** are side views of a two path permanent magnet device including two bypasses;

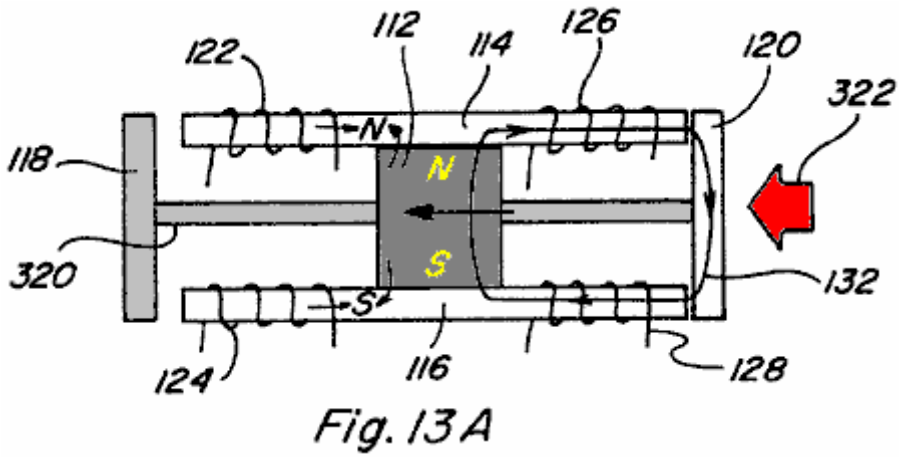


Fig. 13 A

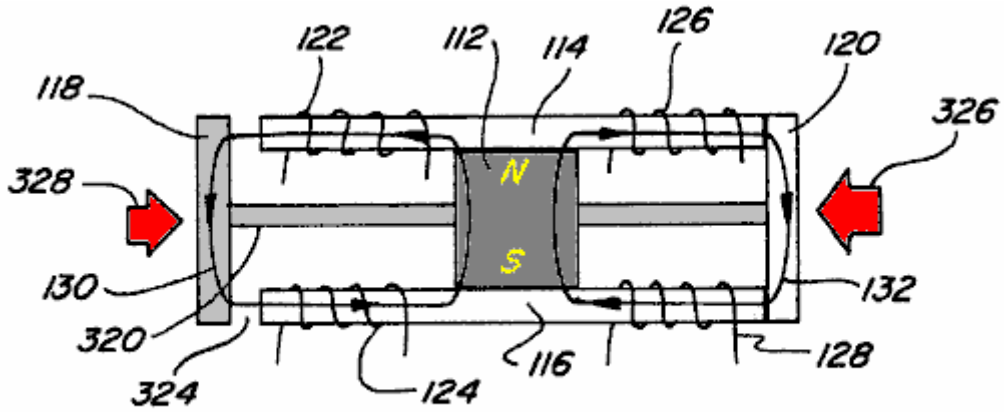


Fig. 13 B

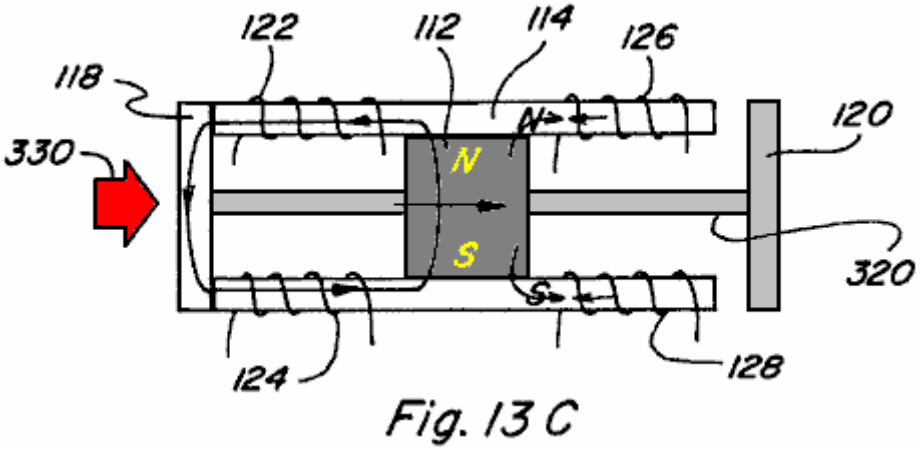


Fig. 13 C

Figs.13A-13C are side views of a permanent magnet linear reciprocating device;

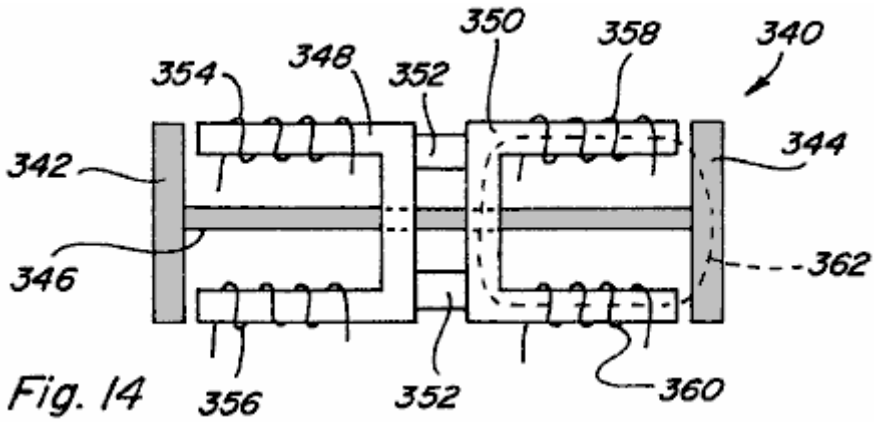


Fig. 14

Fig.14 is a side view of an electromagnetic linear reciprocating device;

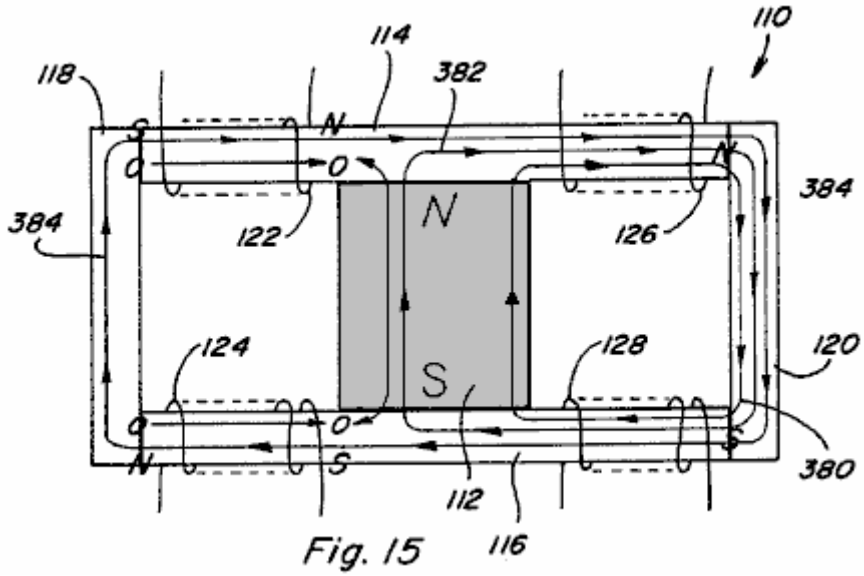


Fig. 15

Fig.15 is a side view of a two path permanent magnet device showing control coils energised in an exceeding manner;

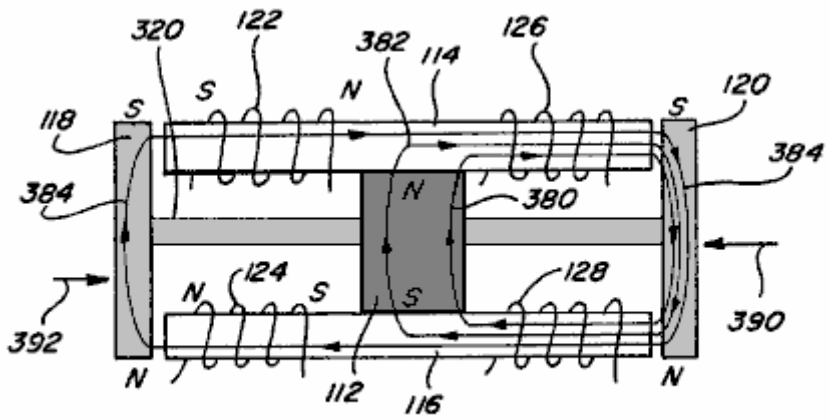


Fig. 16 A

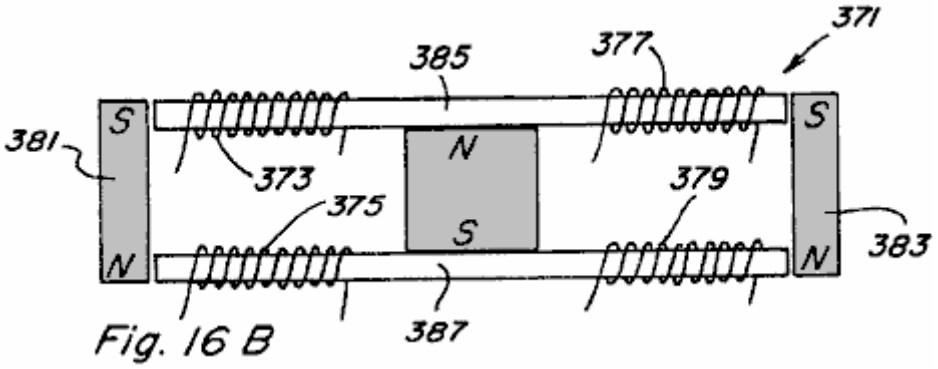


Fig. 16 B

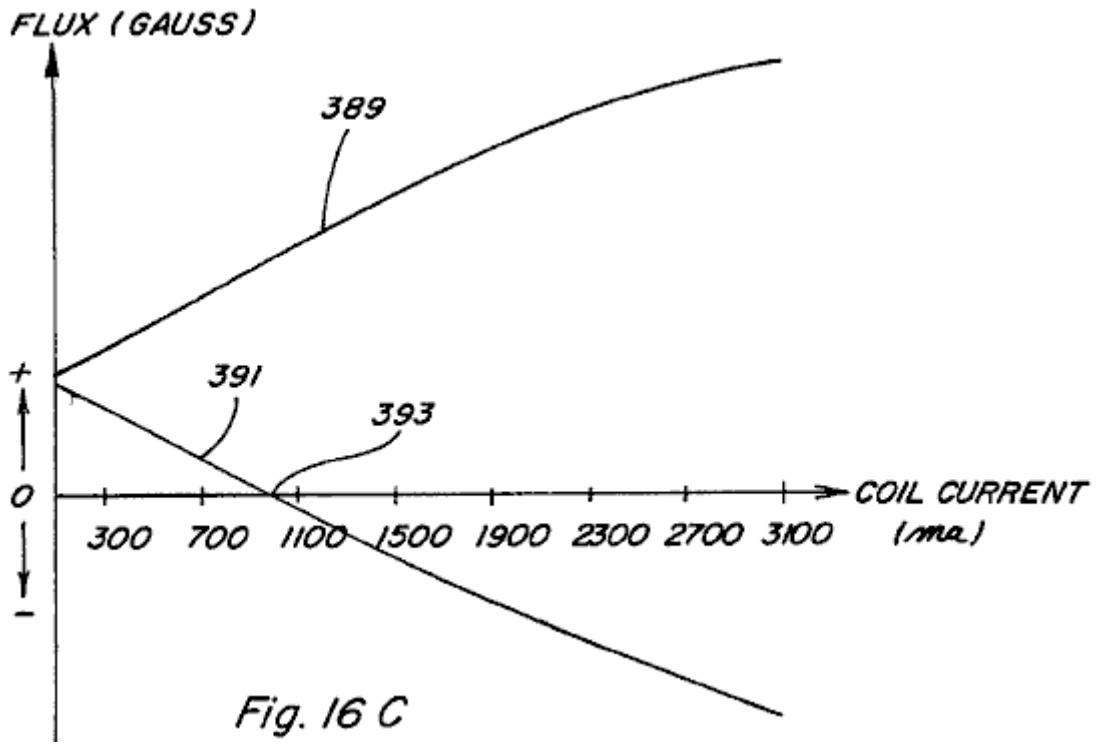


Fig. 16 C

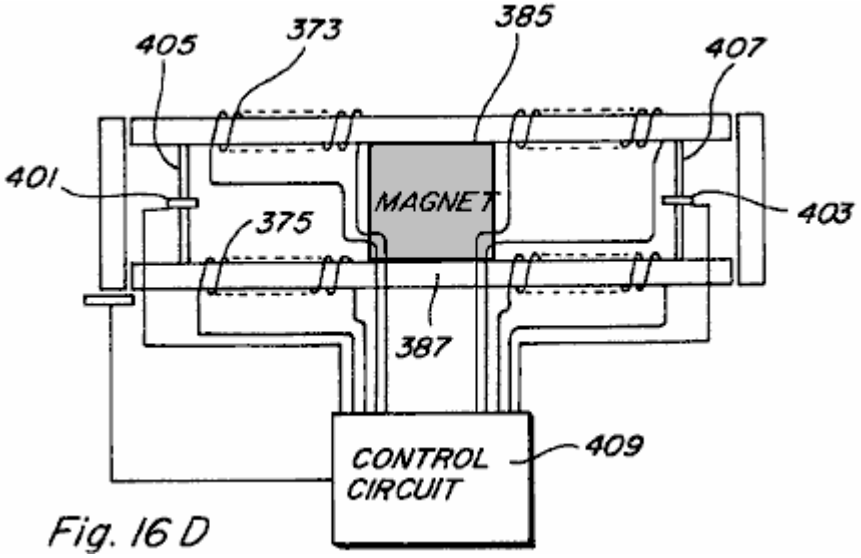


Fig. 16 D

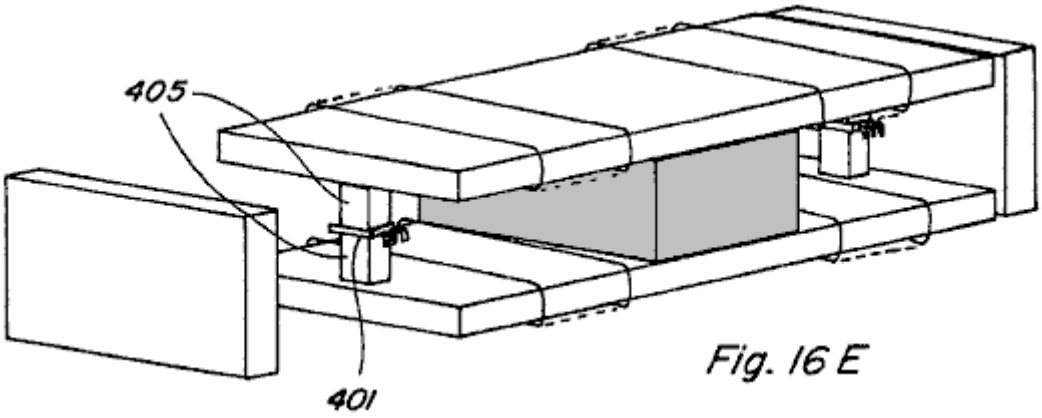


Fig. 16 E

Figs.16A-E are a side view of a linear reciprocating device with control coils energised in an exceeding manner;

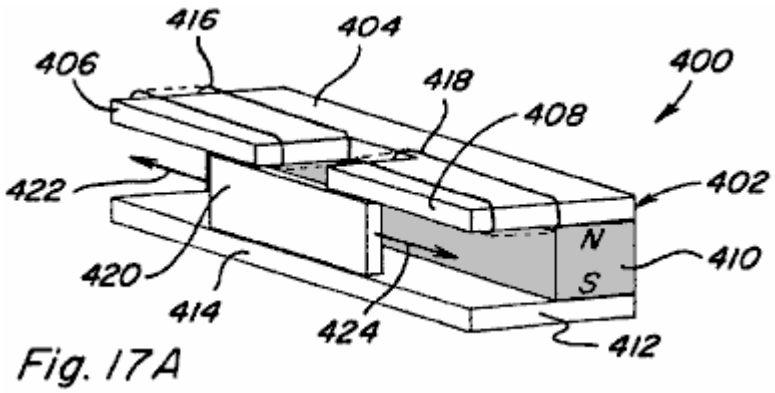


Fig. 17A

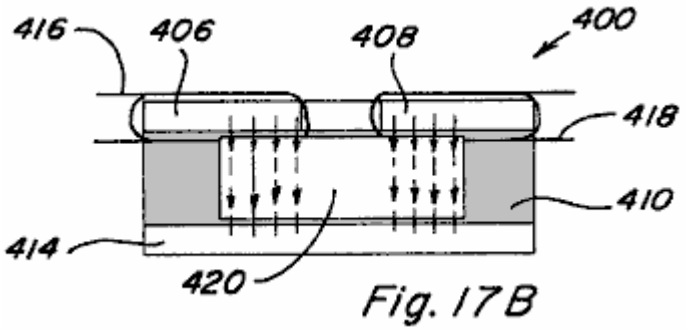


Fig. 17B

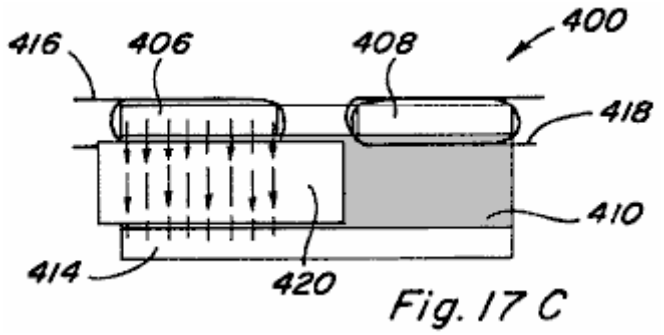


Fig. 17 C

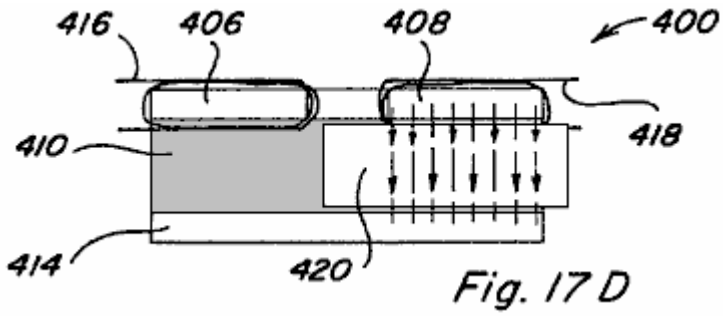


Fig. 17 D

Figs.17A-17D depict another embodiment of a linear reciprocating device;

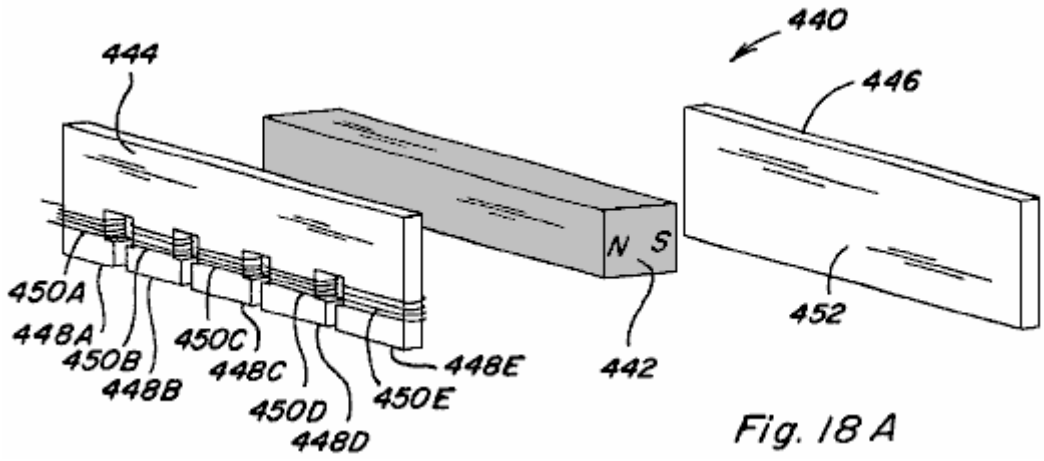


Fig. 18 A

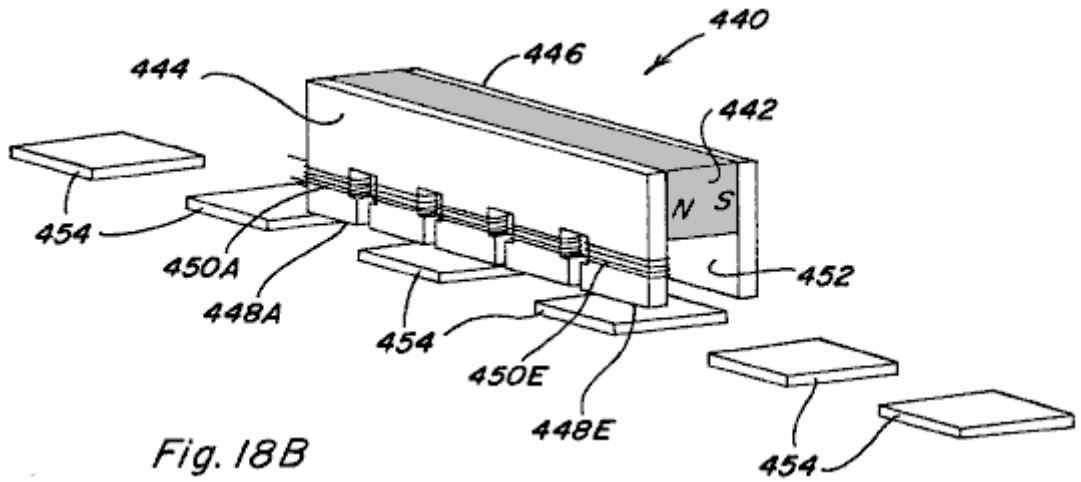
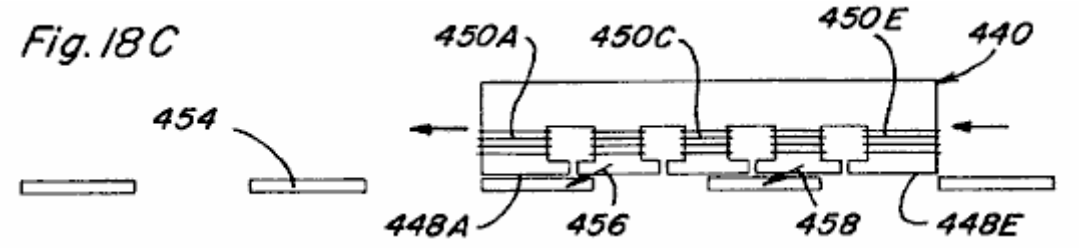
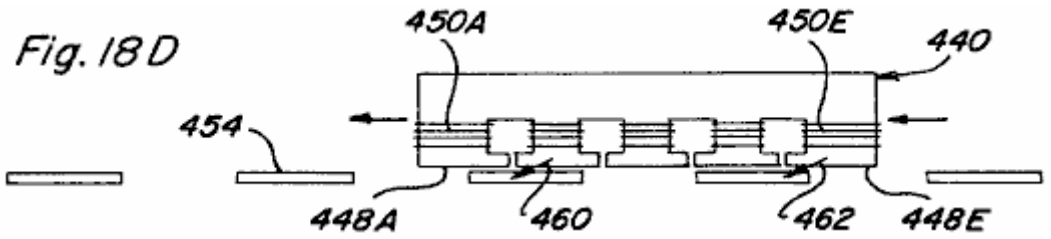
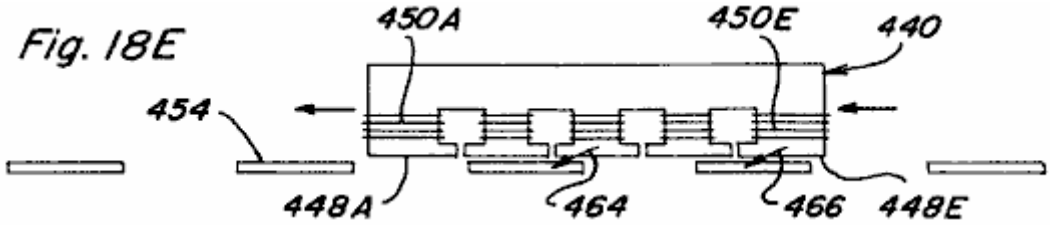


Fig. 18B







Figs.18A-18E show a linear motion device;

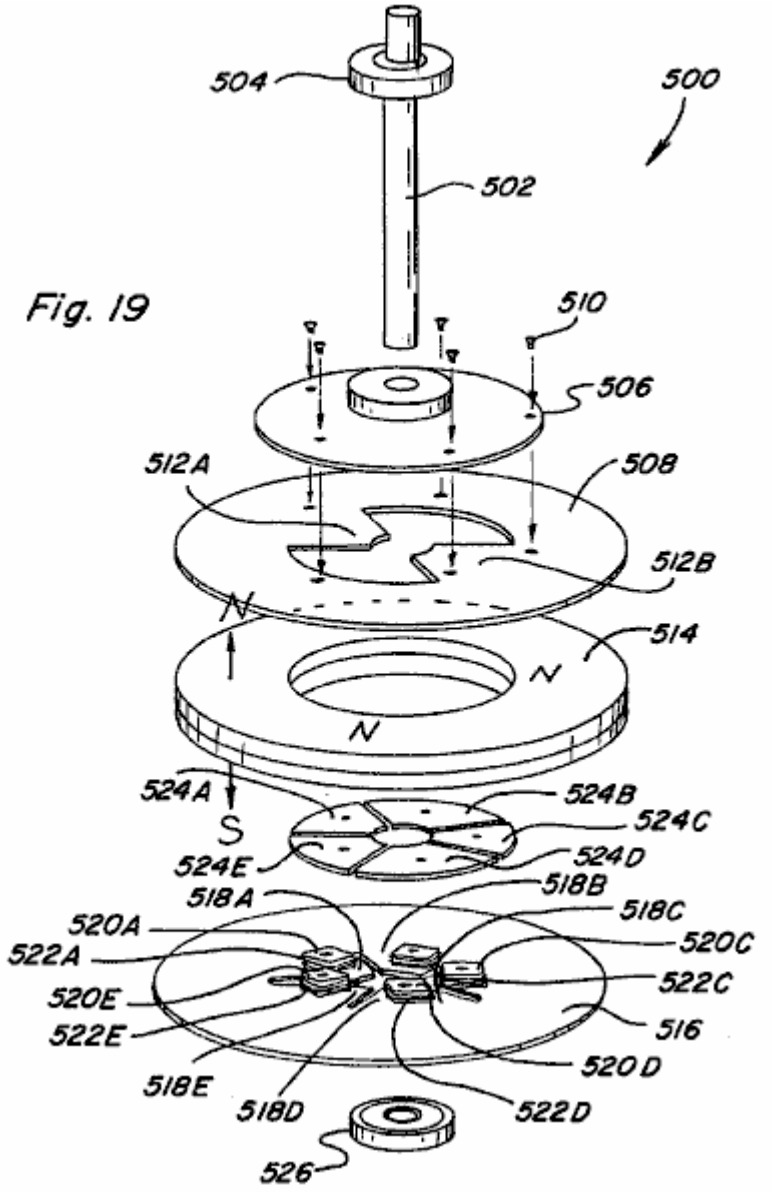


Fig.19 is an exploded perspective view of a rotary motion device;

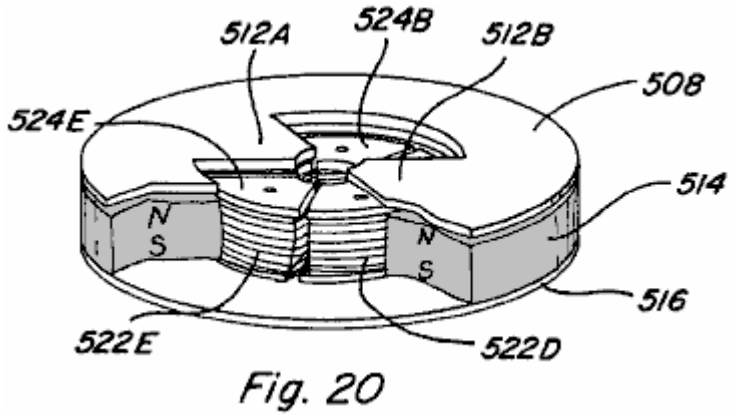


Fig.20 is a partial assembled and cut away view of the rotary motion device of Fig.19;

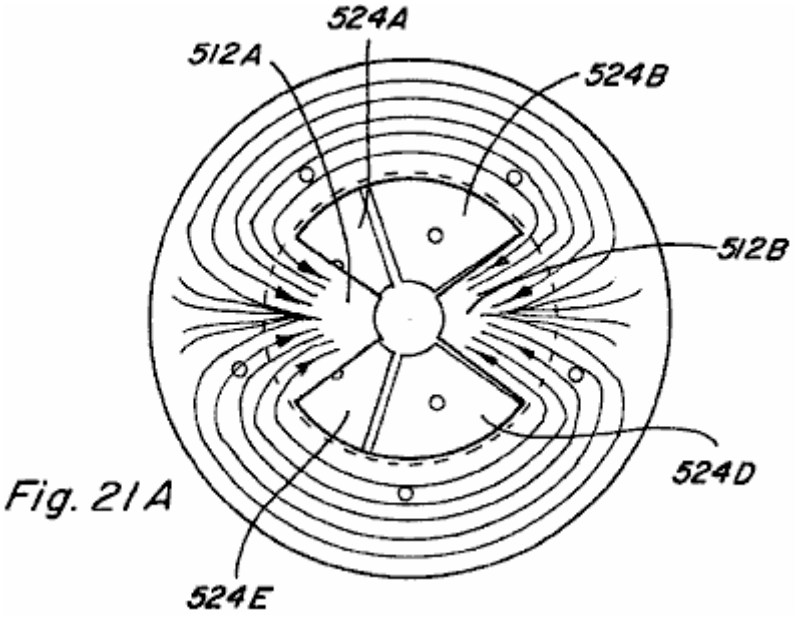


Fig. 21A

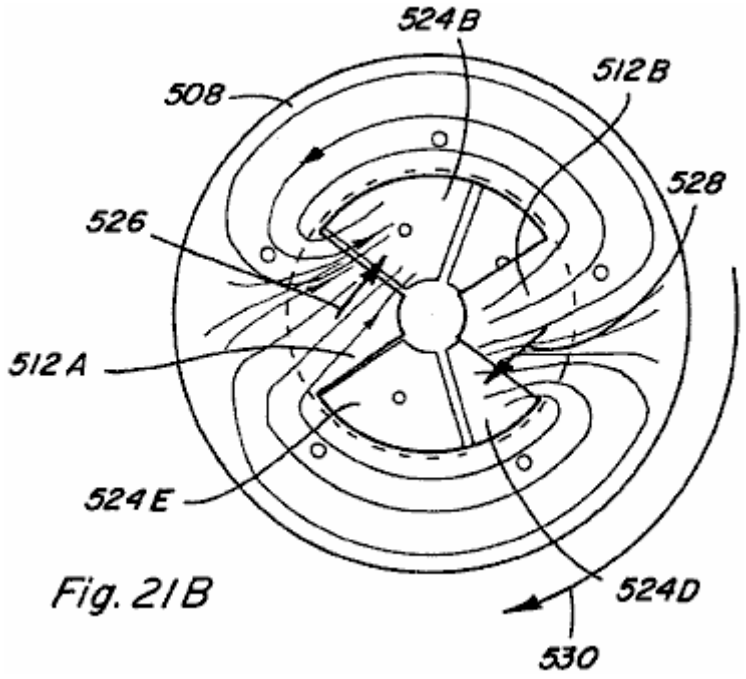


Fig. 21B

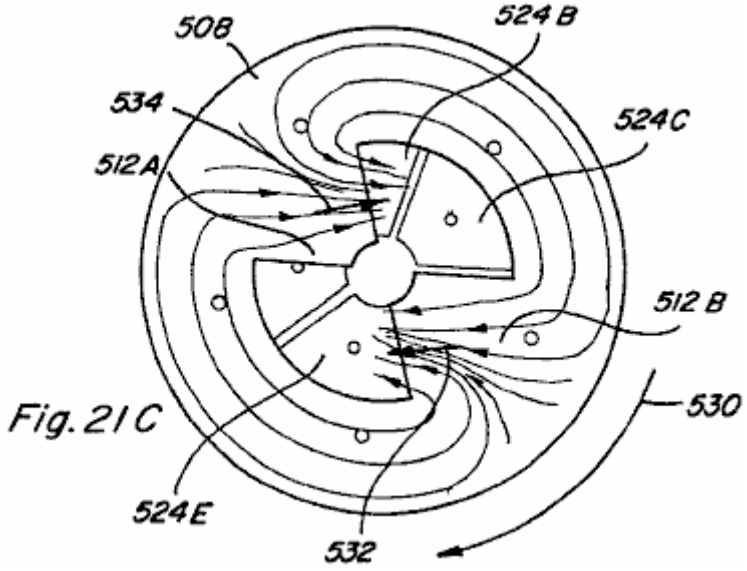


Fig. 21C

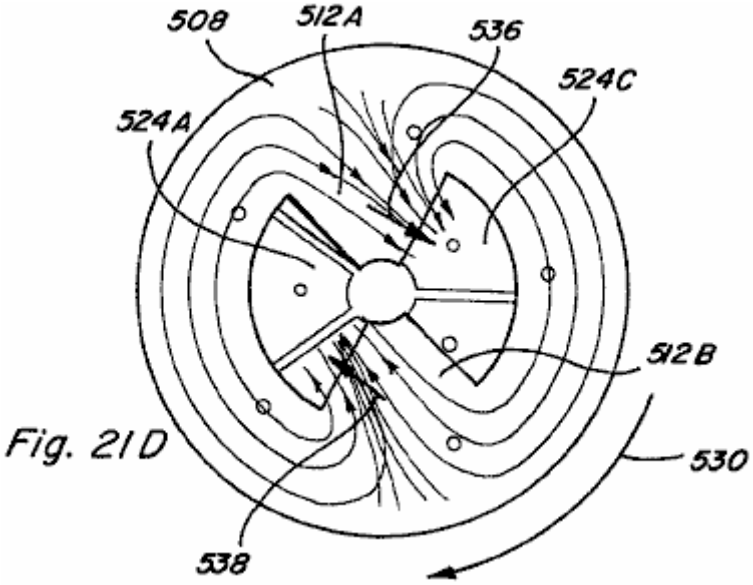


Fig. 21D

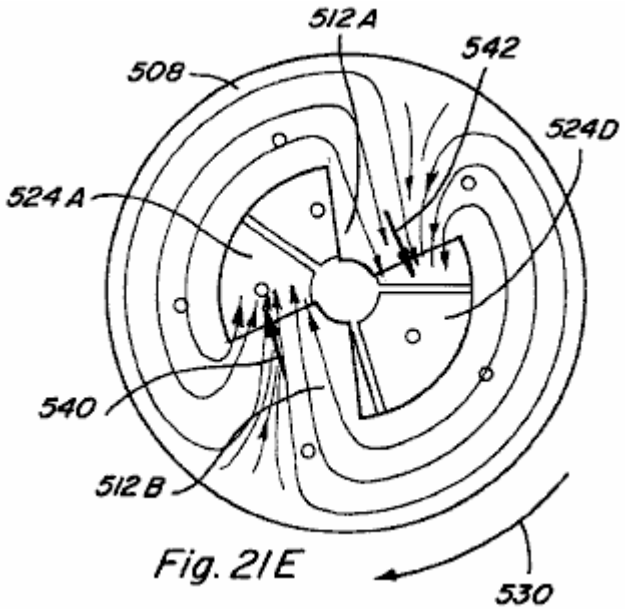


Fig. 21E

Figs.21A-21E are top views of the partial assembly of Fig.20, which views depict rotational motion thereof,

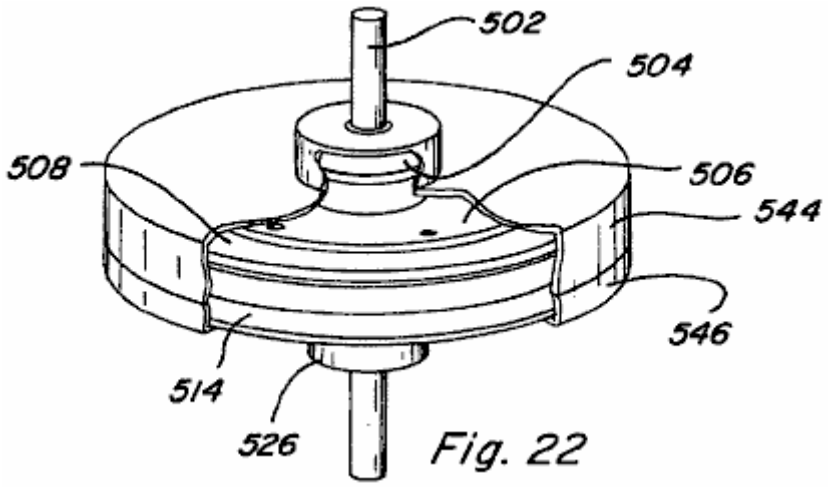


Fig. 22

Fig.22 is an assembled, cut-away view of the rotary motion device of Fig.19 including a housing;

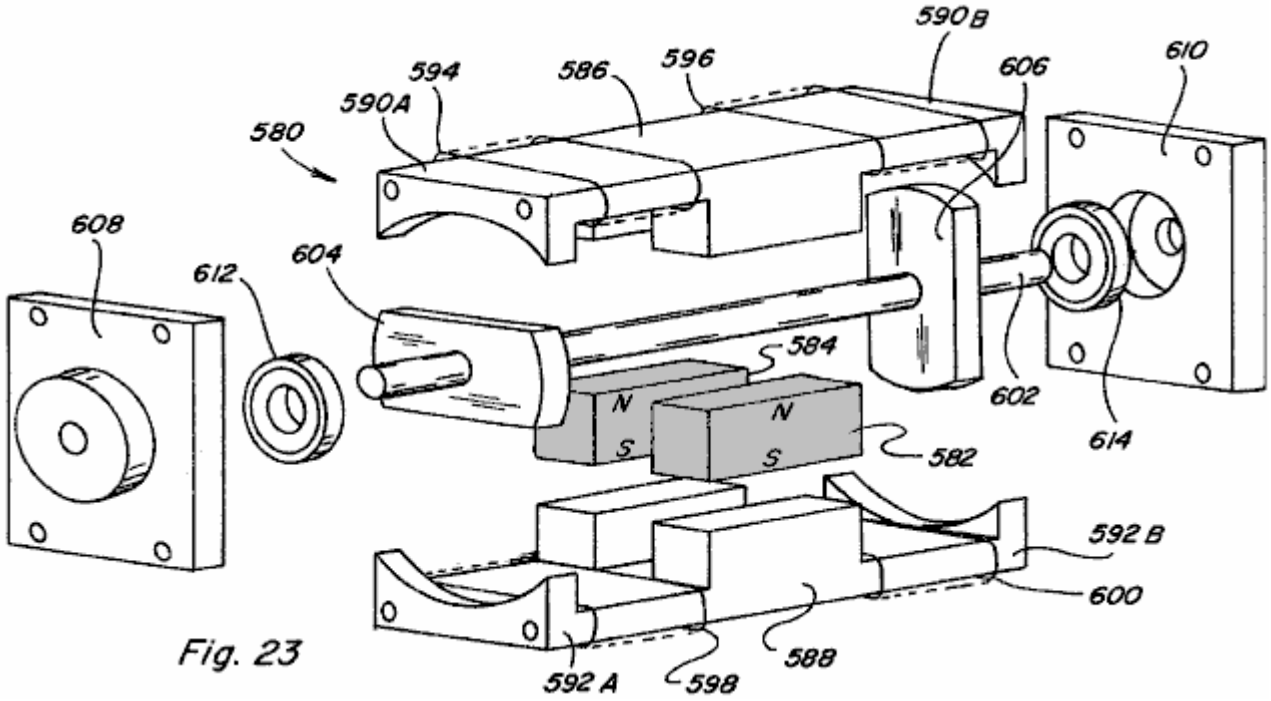


Fig. 23

Fig.23 is an exploded perspective view of another embodiment of a rotary motion device;

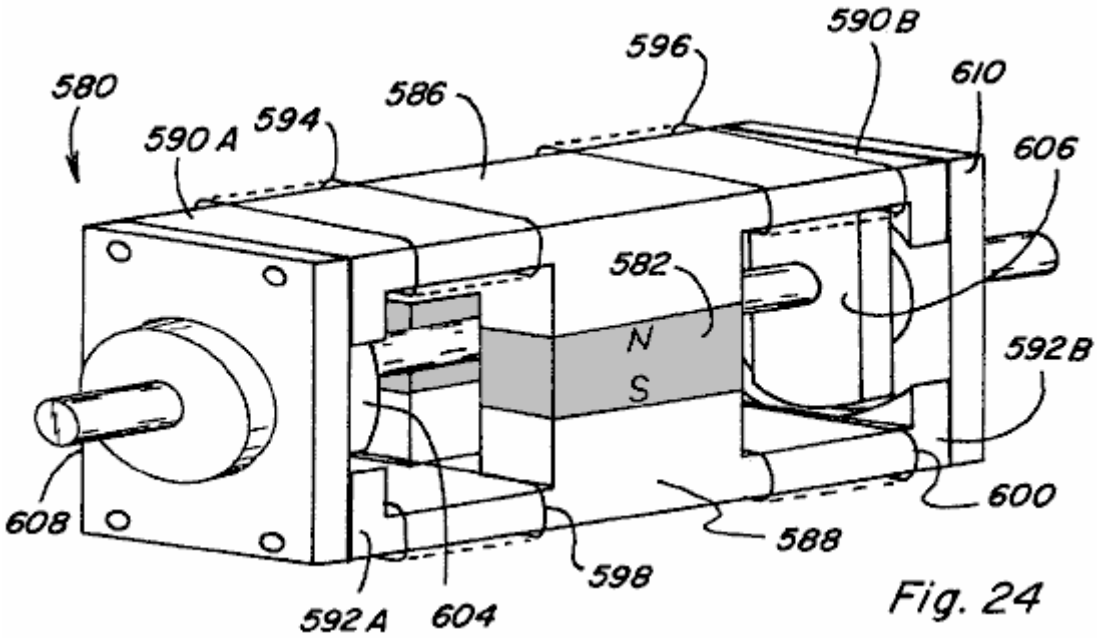
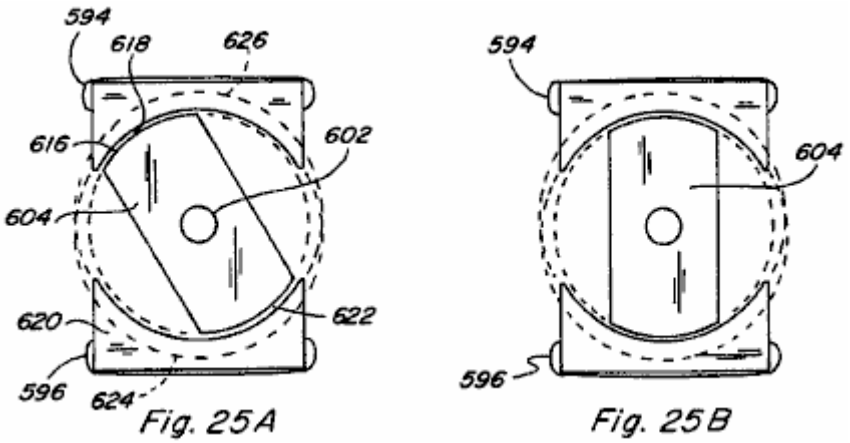
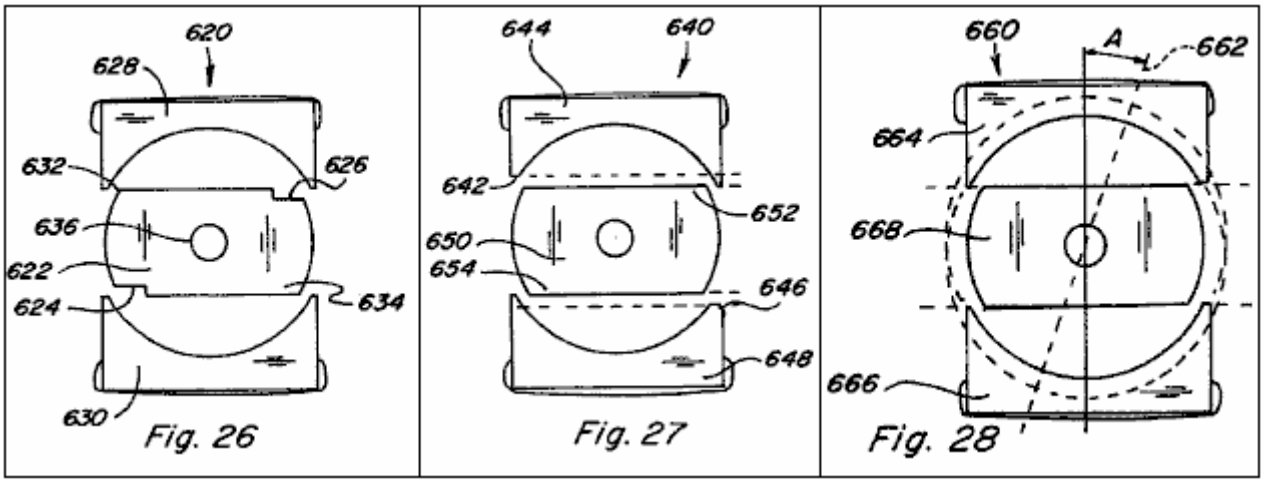


Fig.24 is a perspective view of the rotary motion device of Fig.23 as assembled;



Figs.25A-25B are end views of the rotary motion device of Fig.24 with the end cap removed to expose the rotor member;



Figs.26-28 show end views of various configurations for skewing the direction of rotation in the rotary motion device of Fig.24;

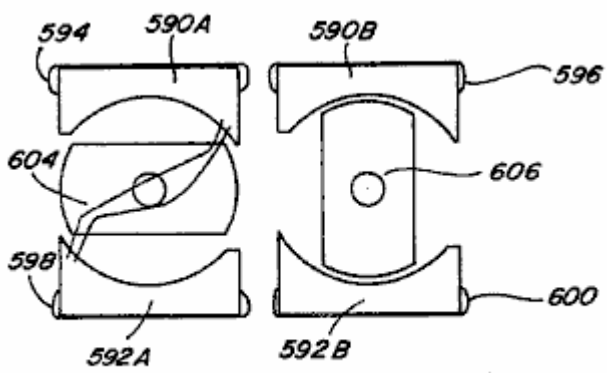


Fig. 29A

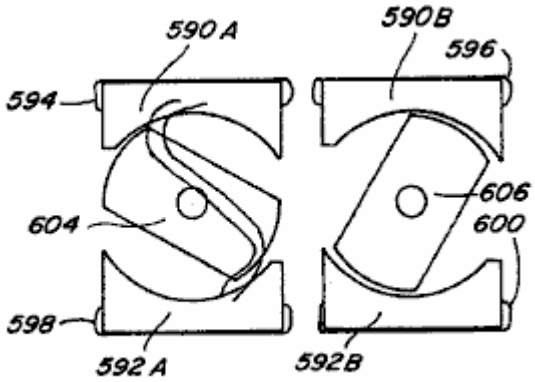


Fig. 29B

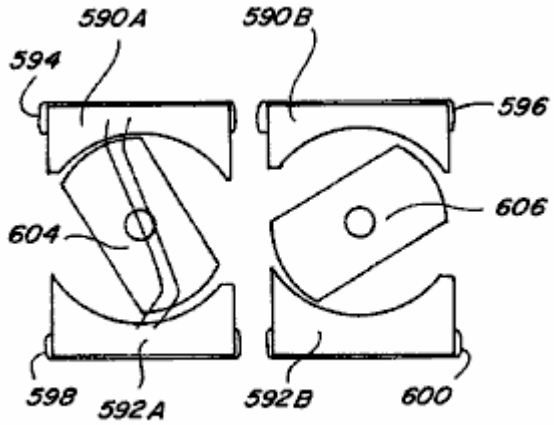


Fig. 29C

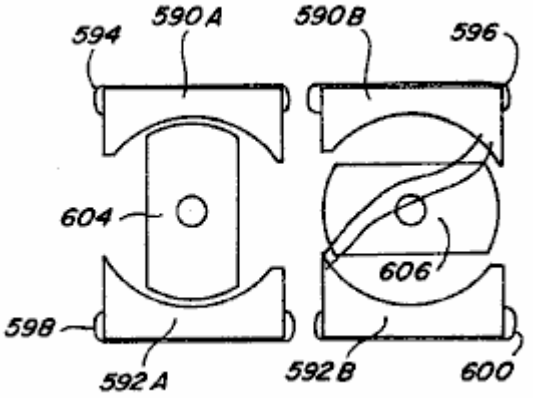


Fig. 29D

Figs.29A-29D are end views of the rotary motion device of Fig.24 illustrating a sequence of its rotational movements;

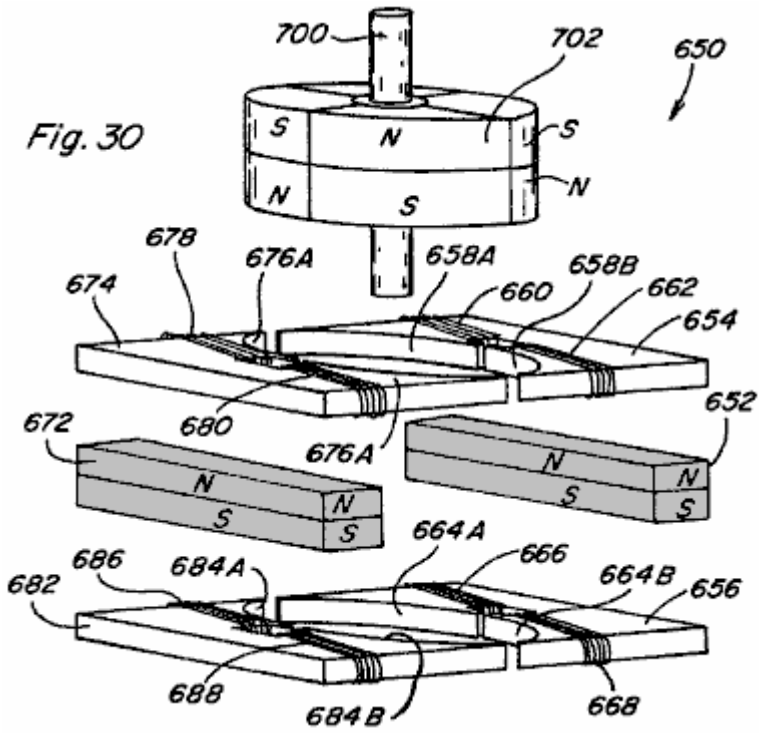


Fig.30 is an exploded partial perspective view of another embodiment of a rotary motion device;

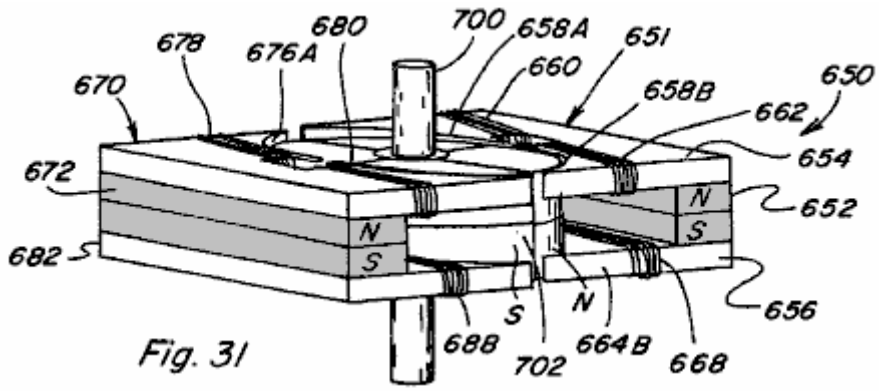
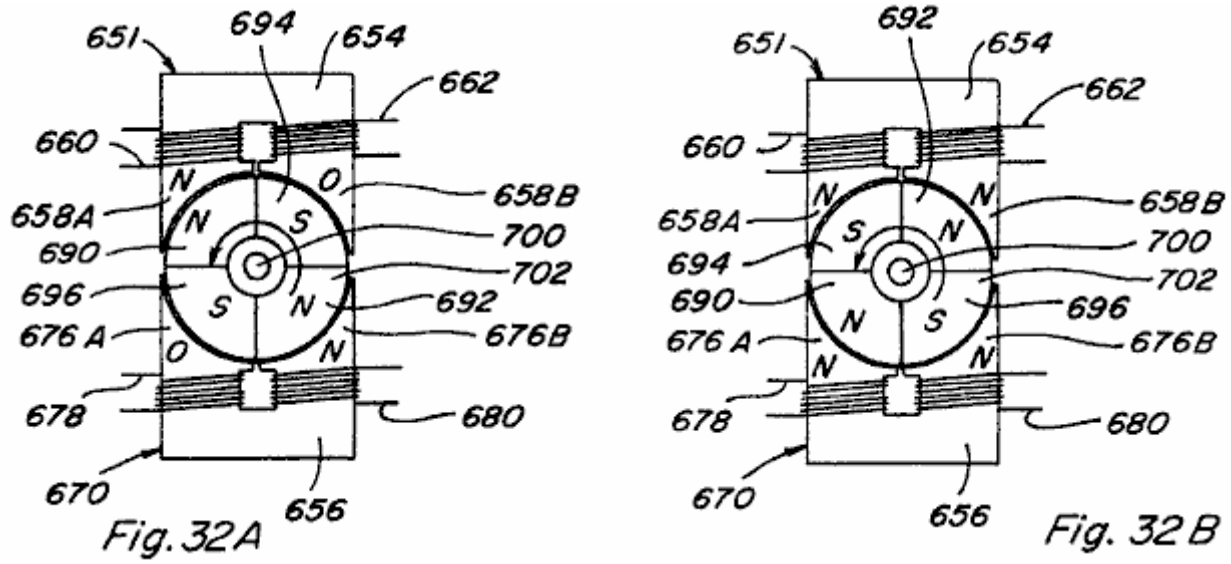
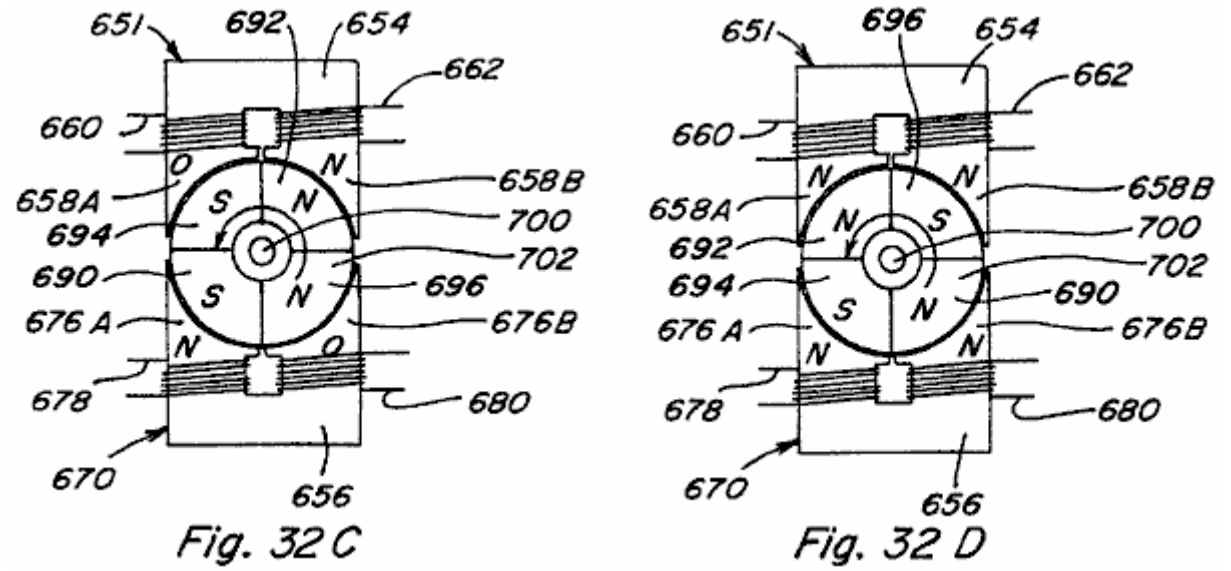


Fig.31 is a perspective view of the rotary motion device of Fig.30 as assembled





Figs.32A-32D are top views of the rotary motion device of Fig.31 illustrating it's rotational movement;

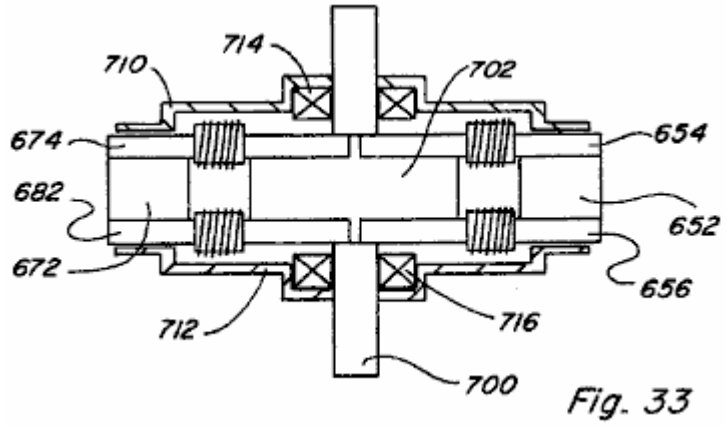


Fig. 33

Fig.33 is a side view of the rotary motion device of Fig.31 as assembled and including a housing;

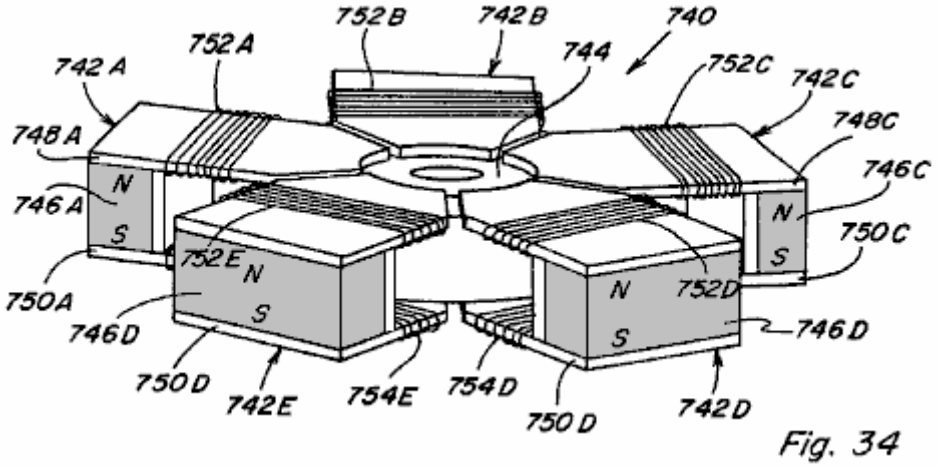


Fig. 34

Fig.34 is a perspective view of another embodiment of a rotary motion device;

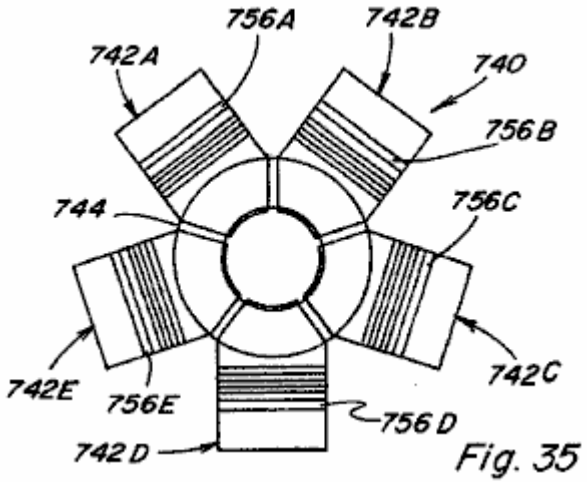


Fig. 35

Fig.35 is a top view of the rotary motion device of Fig.34;

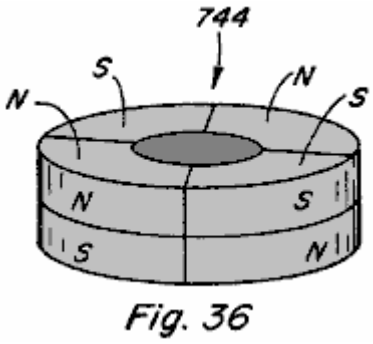


Fig. 36

Fig.36 is a perspective view of the permanent magnet rotor member of the rotary motion device of Fig.34;

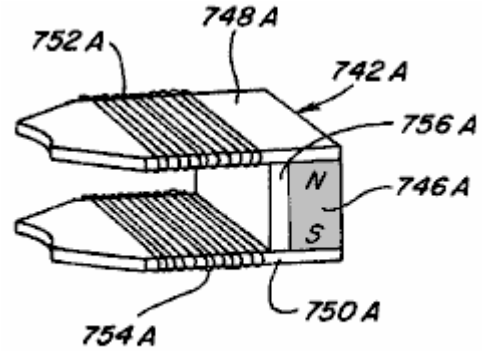


Fig. 37

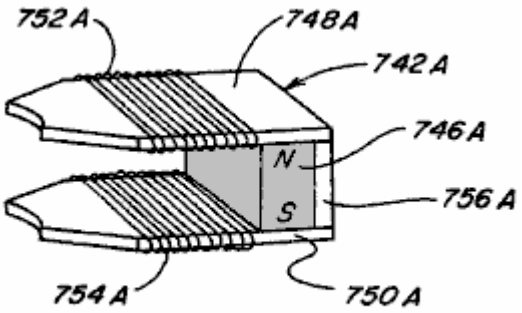


Fig. 38

Fig.37 and **Fig.38** show alternative configurations for the control component incorporated into the rotary motion device of Fig.34;

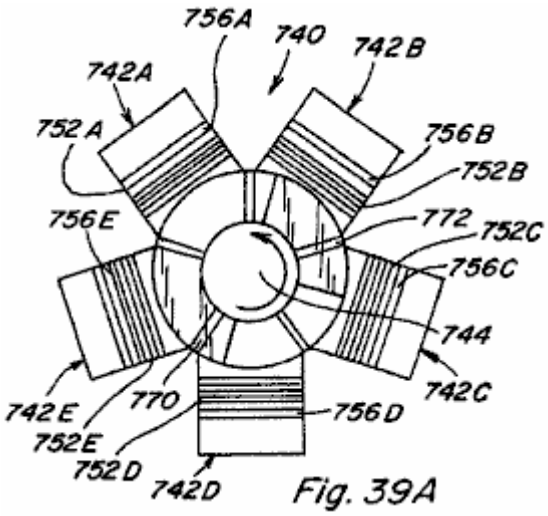


Fig. 39A

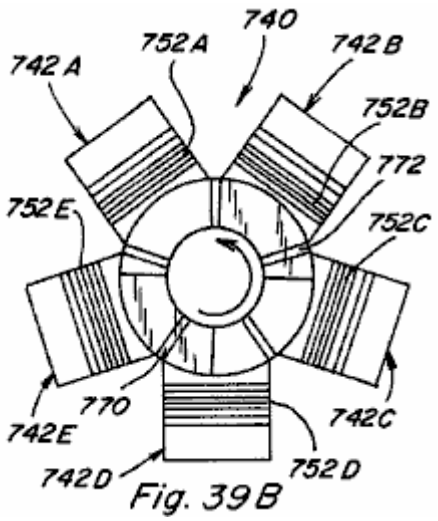


Fig. 39B

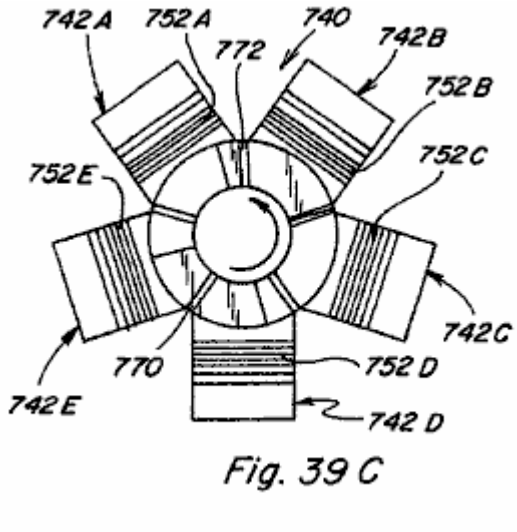


Fig. 39 C

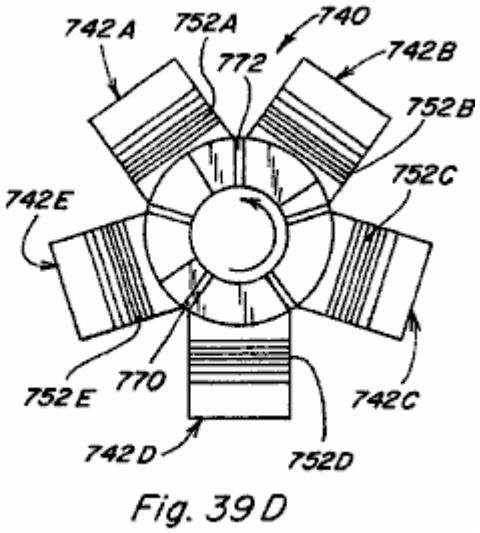


Fig. 39 D

Figs.39A-39D are top views of the rotary motion device of Fig.34 and depict its rotational movement;

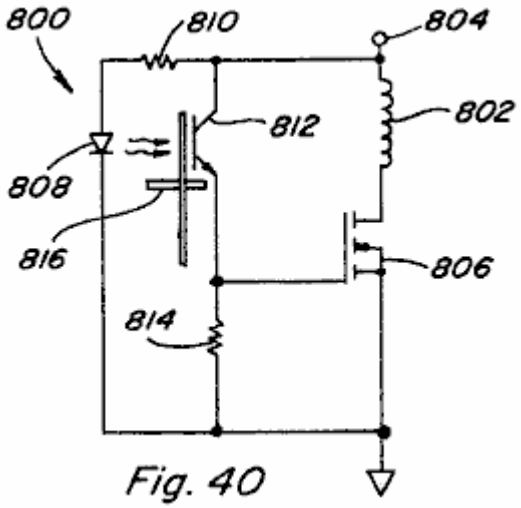


Fig. 40

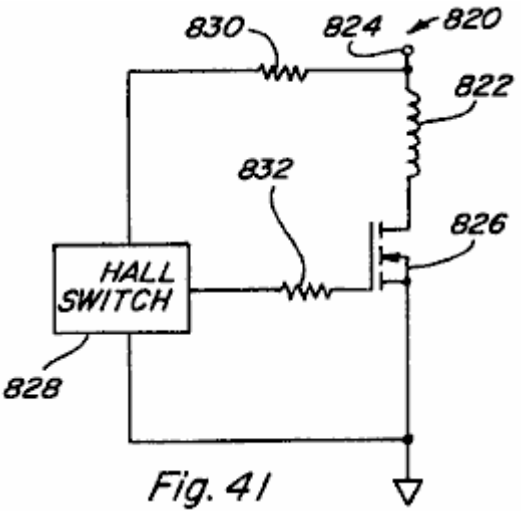


Fig. 41

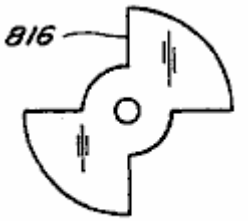


Fig. 42

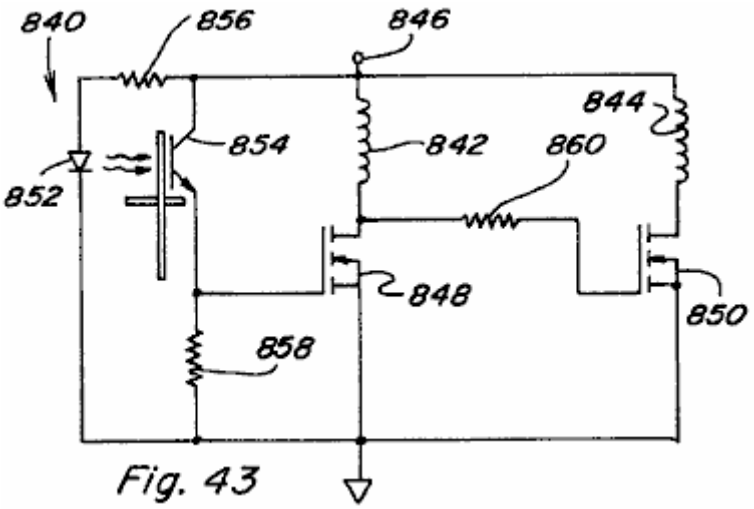


Fig. 43

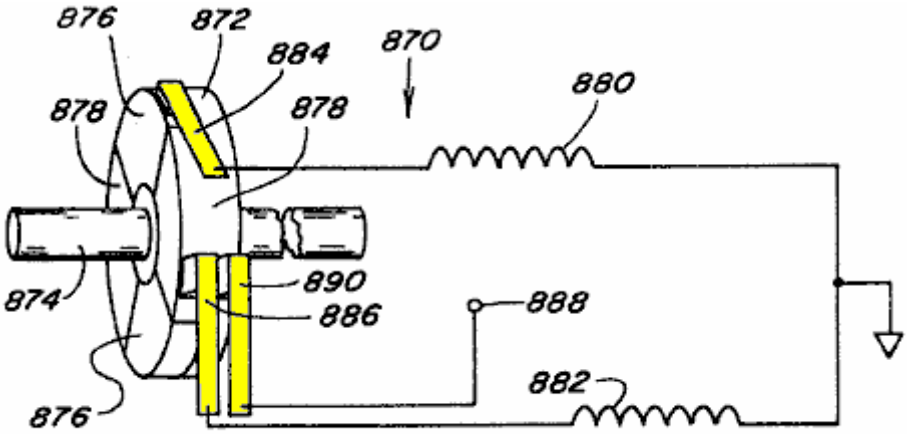


Fig. 44

Figs.40-44 are alternative variations of the circuit for controlling the timed energisation of control coils in the various devices of the present invention;

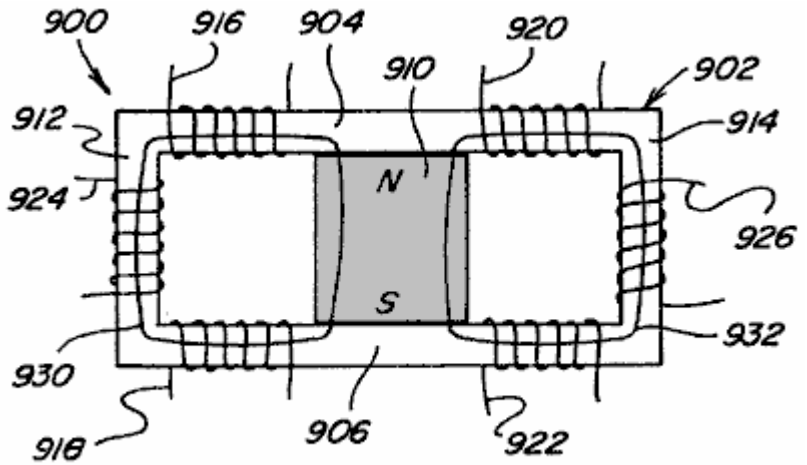


Fig. 45A

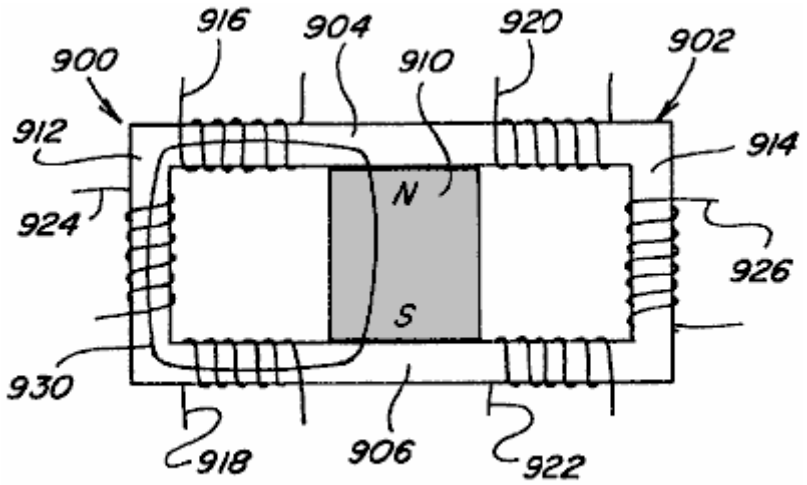


Fig. 45B

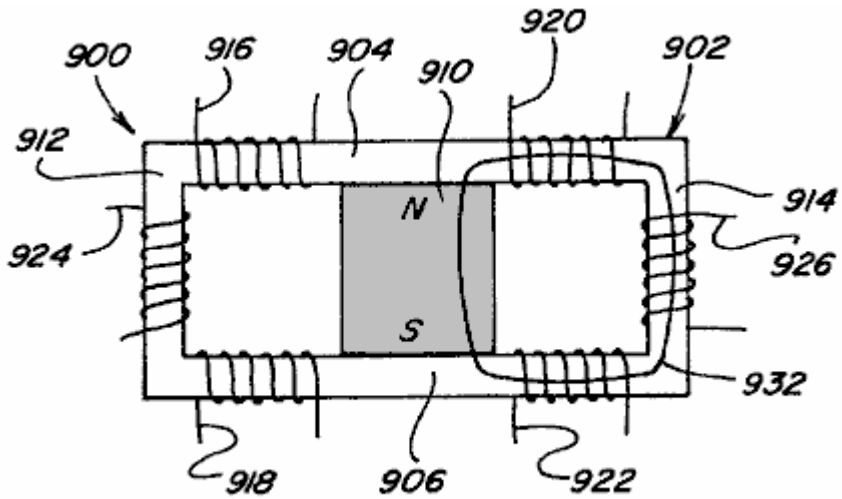


Fig. 45C

Figs.45A-45C and **Figs.45X-45Z** are side views of two path power conversion devices;

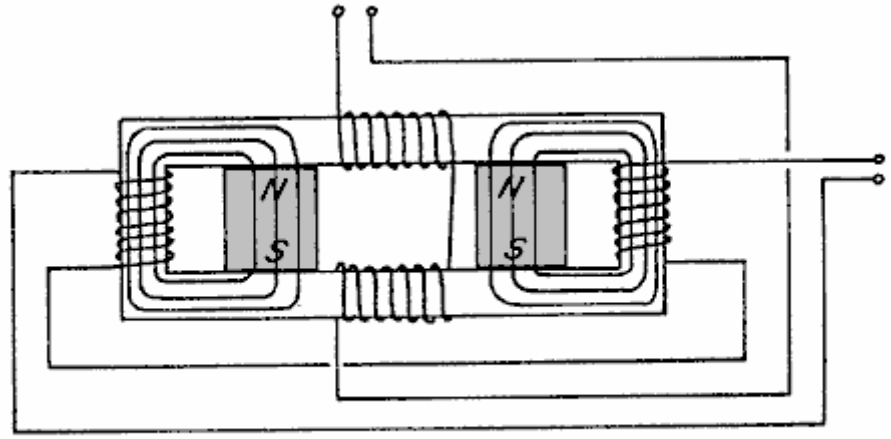


Fig. 45 X

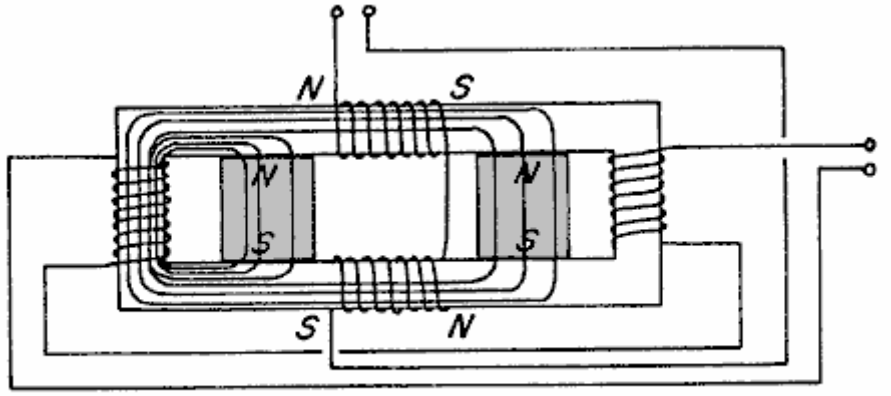


Fig. 45 Y

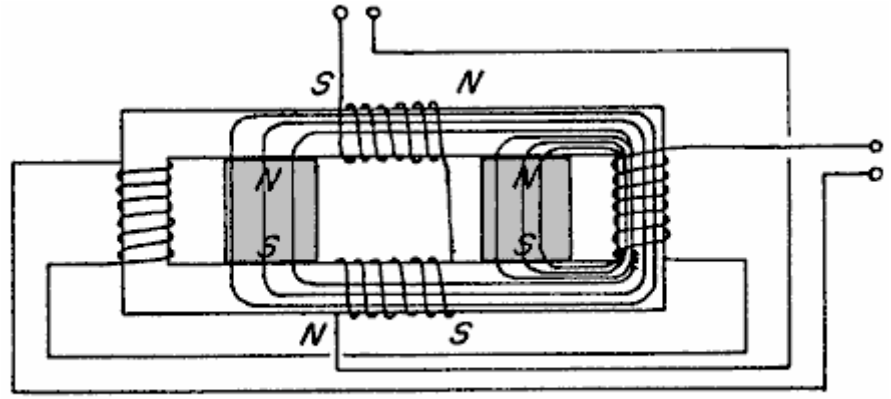


Fig. 45 Z

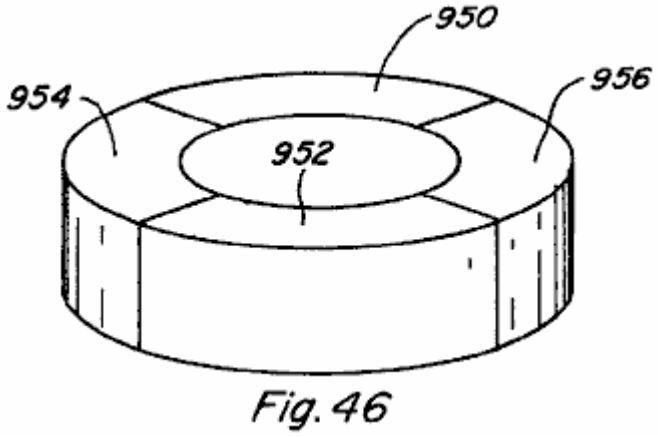


Fig. 46

Fig.46 is a schematic view of the permanent magnet portion of a rotor for use in some embodiments of the present device;

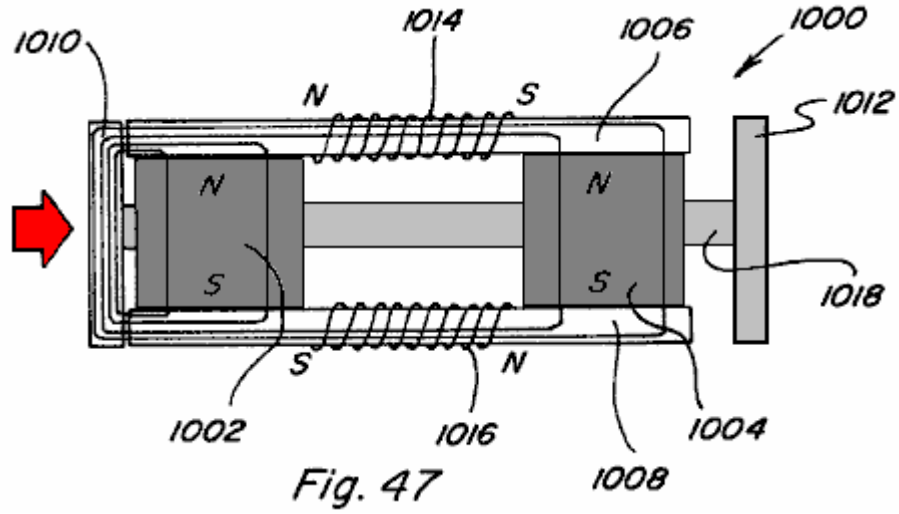


Fig. 47

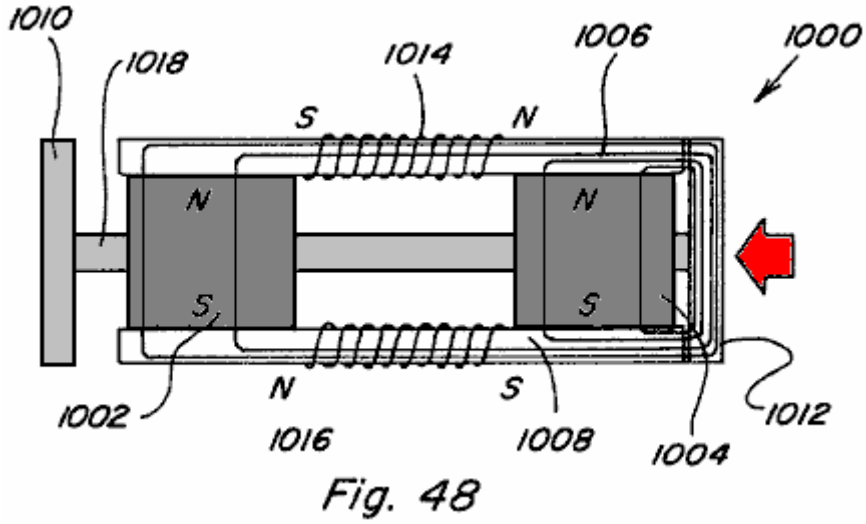


Fig. 48

Fig.47 and **Fig.48** show other embodiments of a linear motion device;

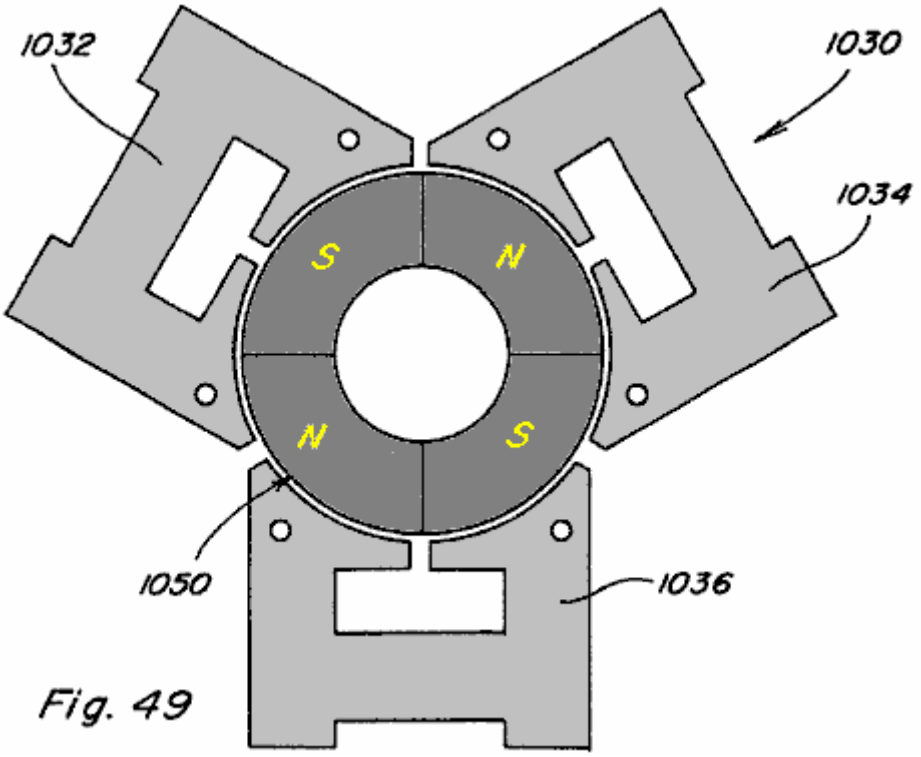


Fig. 49

Fig.49 is a top view of another embodiment of a rotating motor like construction; and

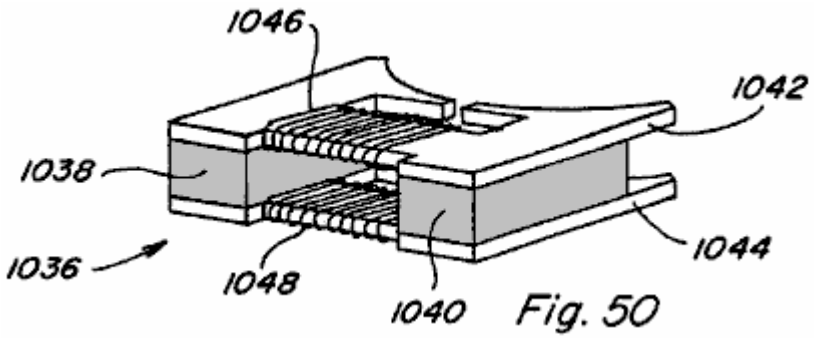


Fig. 50

Fig.50 is a schematic view of one of the three stator portions of the device shown in **Fig.49**.

DETAILED DESCRIPTION OF THE DRAWINGS

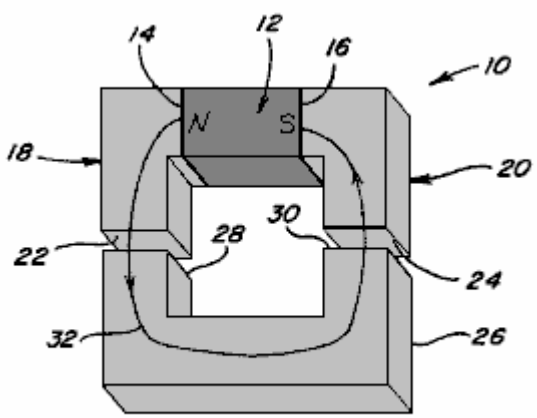


Fig. 1

Referring now to the drawings, Figs.1-4 are provided to facilitate an understanding of various aspects or features of the technology utilised in the present invention. **Fig.1** depicts a device **10** having a magnetic flux producing member **12** which may be a permanent magnet or electromagnet with magnetic poles **14** and **16** as shown. Pole pieces **18** and **20** are positioned adjacent respective poles **14** and **16** to provide a path for the magnetic flux of member **12**. Each pole piece **18** and **20** has a pole piece end face **22** and **24**. As used throughout this specification, it is understood that a pole piece, regardless of its shape or size, is preferably formed of soft iron, steel or some other magnetic material, with the preferred material being one which provides low reluctance, exhibits low hysterisis, and has a high magnetic flux density capability. Accordingly, the various pole pieces disclosed and described herein could likewise be of laminate type construction.

Referring again to **Fig.1** an armature **26**, also formed of magnetic material, is shown with end faces **28** and **30** which are positioned and sized for being placed adjacent pole piece end faces **22** and **24**, such that when so positioned a substantially continuous low reluctance path **32** is provided for magnetic flux from north pole **14**, through pole piece **18**, through armature **26**, through pole piece **16**, and to south pole **16**. The magnetic flux travelling along such path **32** results in a force which tends to hold armature **26** in position aligned with pole piece end faces **22** and **24**. The resulting magnetic coupling or holding force F provided between adjacent pole piece end face **22** and armature end face **28**, and between adjacent pole piece end face **24** and armature end face **30**, can be approximated by the following equation:

$$F=B^2A/2\mu_0,$$

where **B** is the magnetic flux density passing through the adjacent end faces and **A** is the surface area of the adjacent end faces. Assuming that if **B** is uniform throughout flux path **32** and that the area **A** of all end faces **22**, **24**, **28**, and **30** is the same, then the total holding force $\mathbf{F_{T26}}$ of armature **26** against pole pieces **18** and **20** will be:

$$F_{T26}=B^2A/\mu_0.$$

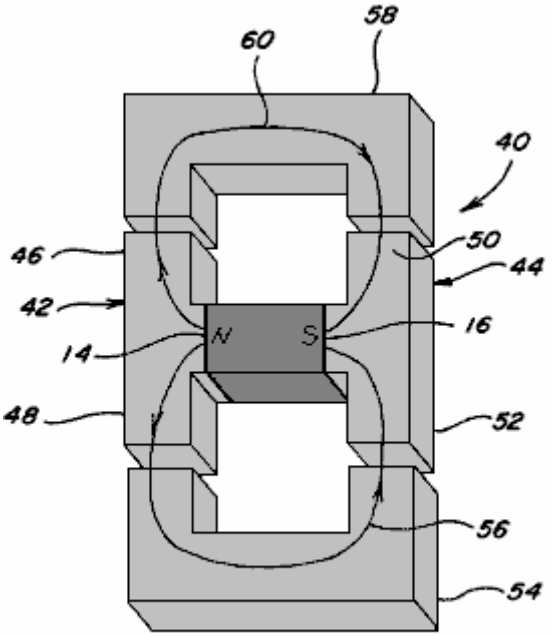


Fig. 2

In **Fig.2** a device **40** having the same magnetic flux producing member **12** with magnetic poles **14** and **16** is shown. Pole pieces **42** and **44** are positioned adjacent respective pole faces **14** and **16** to provide two paths, as opposed to one above, for the magnetic flux of member **12**. In particular, pole piece **42** includes a first path portion **46** extending beyond a perimeter of north pole face **14** in one direction and a second path portion **48** extending beyond the perimeter of north pole face **14** in another direction. Similarly, pole piece **44** includes a first path portion **50** extending beyond the perimeter of south pole face **16** in one direction and a second path portion **52** extending beyond the perimeter of south pole face **16** in another direction. Each pole piece path portion **46**, **48**, **50**, **52** includes a respective end face. A first armature **54** which can be positioned adjacent to the end faces of pole piece path components **48** and **52** provides a first magnetic flux path **56** and a second armature **58** is which can be positioned adjacent the end faces of pole piece path components **46** and **50** provides a second magnetic flux path **60**. If the flux carrying area along flux paths **56** and **60** is the same as the flux carrying area along flux path **32** of **Fig.1**, the magnetic flux density along each flux path **56** and **60** will be one-half the magnetic flux density along flux path **32** of **Fig.1** because the same amount of flux is split between two like paths. The effect of dividing a given amount of magnetic flux along two like flux paths instead of it passing along just one flux path can be seen by examining the holding force on armature **54** as compared to the holding force on armature **26** of **Fig.1**. As already noted the magnetic flux density along path **56** will be one-half that along flux path **32** and thus the total holding force F_{T54} can be determined as:

$$F_{T54}=(B/2)2A/,\mu_0=B^2A/4\mu_0=F_{T26}/4.$$

It is therefore seen that dividing the same amount of magnetic flux along two flux paths rather than along one flux path reduces the magnetic holding or coupling force on an armature to one-fourth rather than one-half as might have been expected. This unexpected magnetic holding or coupling force differential, resulting from multiple flux paths, can provide advantageous properties in linear, reciprocating, and rotary motion devices.

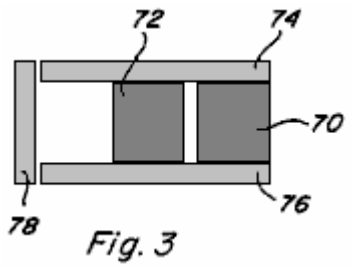


Fig. 3

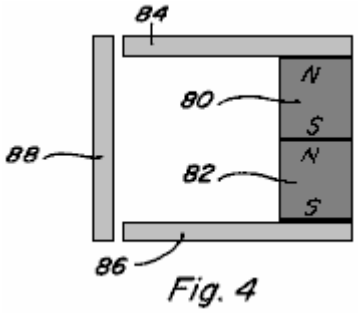


Fig. 4

Referring now to **Fig.3** and **Fig.4**, the behaviour of multiple magnetic flux sources arranged in parallel and series is described as compared to a single flux source. When identical flux sources or magnetic flux producing members **70** and **72** are positioned in parallel as shown in **Fig.3** with pole pieces **74** and **76** positioned adjacent the poles thereof to provide a flux path through armature **78**, the flux density B through armature **78** is double what the flux density would be if only one magnetic flux producing member were present. However, the field intensity H resulting from the two members **70** and **72** remains unchanged. This result holds true regardless of whether members **70** and **72** are both permanent magnets, are both electromagnets, or are a combination of one permanent magnet and one electromagnet. On the other hand, the properties resulting from magnetic flux producing members **80** and **82** arranged pole-to-pole in series between pole pieces **84** and **86**, with armature **88**, as shown in **Fig.4**, will vary depending on the nature of the members **80** and **82**.

In a first case, if both members **80** and **82** are permanent magnets, the magnetic field intensity H resulting from the two permanent magnets will be double that of one permanent magnet and the flux density B through armature **88** will be the same as what the flux density would be if only one permanent magnet type member were present.

In a second case, if both members **80** and **82** are electromagnets, the field intensity H again doubles and the flux density B increases according to the B/H curve or relationship of the pole piece **84**, **86** and armature **88** materials.

In a third case, if member **80** is a permanent magnet and member **82** is an electromagnet, the field intensity H again doubles, but, since the permanent magnet is near flux density saturation B_r the flux density can only be increased from B_r to B_{max} of the permanent magnet. At the point where electromagnet-type member **82** contacts permanent magnet-type member **80** the flux from the electromagnet-type member **82** couples with the flux of the permanent magnet-type member **82** until the flux density through permanent magnet-type member **80** reaches B_{max}. At that point additional flux from electromagnet-type member **82** does not contribute to the flux density along the flux path unless a bypass path around the permanent magnet-type member is provided. Use of such bypass paths will be described below.

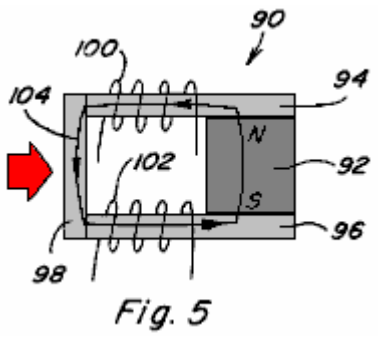


Fig. 5

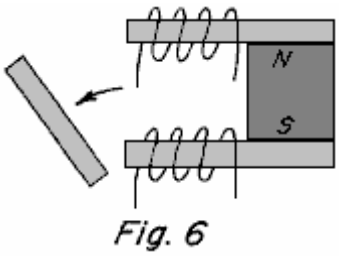


Fig. 6

Controlling the flow of flux along both one and multiple flux paths is best described with reference to **Figs.5-9**. In **Fig.5** and **Fig.6** a permanent magnet device **90** including a permanent magnet **92** having pole pieces **94** and **96** positioned adjacent to it's pole faces, and an armature **98** completing a low reluctance path **104** from pole to pole is shown. Control coils **100**, **102** are positioned along path **104**. When control coils **100**, **102** are not energised, the magnetic flux of permanent magnet **92** follows path **104** as shown and armature **98** is held in place against

pole pieces **94**, **96** due to the resulting magnetic coupling forces. However, if coils **100**, **102** are energised to provide an equal but opposing magnetic flux to that of permanent magnet **92**, the result is that the magnetic flux of permanent magnet **92** is blocked and no magnetic flux traverses the path which includes armature **98** and therefore no magnetic coupling forces act on armature **98** allowing it to fall away as shown in **Fig.6**. The permanent magnet device **90** is useful, although as will become apparent below, it is more advantageous to provide multiple flux paths rather than one.

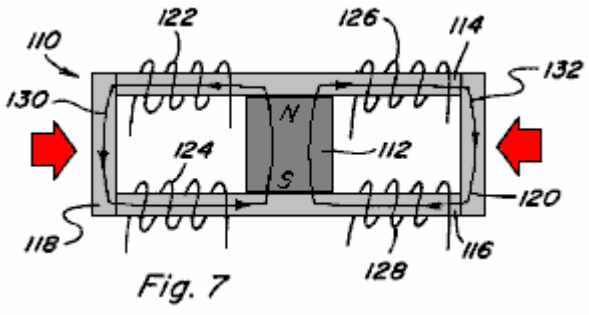


Fig. 7

In this regard, in **Fig.7** a permanent magnet device **110** includes a permanent magnet **112** having pole pieces **114**, **116** positioned adjacent the pole faces of it, with armatures **118**, **120** completing two low-reluctance paths **130**, **132** from pole to pole thereof. Control coils **122**, **124** are positioned along path **130** and control coils **126**, **128** are positioned along path **132**. The two paths provided are assumed to be of equal reluctance. With no coils energised, the magnetic flux of permanent magnet **112** divides equally along flux path **130** and flux path **132** such that both armatures **118**, **120** are subjected to a magnetic coupling force which holds them in place against pole pieces **114**, **116**.

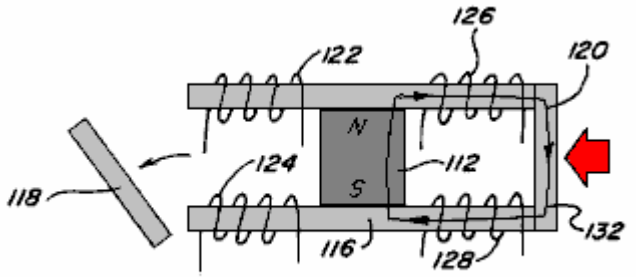


Fig. 8

If coils **122**, **124** are energised to provide a magnetic flux equal to but opposing the magnetic flux which travels along flux path **130** from permanent magnet **112** when no coils are energised, the result is that the magnetic flux of permanent magnet **112** is blocked and no magnetic flux traverses the path which includes armature **118** and therefore no magnetic coupling forces act on armature **118** allowing it to fall away as shown in **Fig.8**. Further, the magnetic flux traversing path **132** will be double that of when no coils are energised and therefore the magnetic coupling force on armature **120** will be about four (4) times that of when no coils are energised. By energising coils **126**, **128** in an opposing manner a similar result would be achieved such that armature **120** would fall away and such that the magnetic coupling force on armature **118** would be increased.

If coils **122**, **124** are energised to provide a magnetic flux equal to and aiding the magnetic flux which travels along flux path **130** when no coils are energised, the result is that the control coils couple completely with the magnetic flux of permanent magnet **112** and no magnetic flux traverses the path which includes armature **120** and therefore no magnetic coupling forces act on armature **120** allowing it to fall away as shown in **Fig.9**. Further, the magnetic flux traversing path **130** will be double that of when no coils are energised and therefore the magnetic coupling force on armature **118** will be about four (4) times that when no coils are energised. By energising coils **126**, **128** in an aiding manner a similar result would be achieved such that armature **118** would fall away and the magnetic coupling force on armature **120** would be increased.

Based on the foregoing, it is seen that the full magnetic coupling force available from the permanent magnet **112,** can be switched from one path to another path by the application of one half the power it would require for a coil alone to produce the same magnetic flux along one path. The ability to switch the full magnetic coupling force easily from one path to another, allows for efficient reciprocating, linear, and rotary motion and power conversion to be achieved.

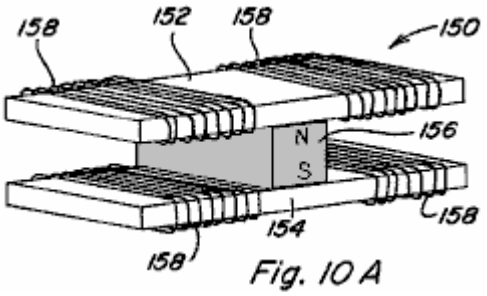


Fig. 10 A

The basic device utilised to achieve permanent magnet flux division and to control such permanent magnet flux division is defined herein as a "permanent magnet control component," various configurations of which are shown by way of example only, and not by way of limitation, in **Figs.10A-10F**. **Fig.10A** depicts a permanent magnet control component **150** in which pole pieces **152** and **154** are positioned adjacent to the pole faces of permanent magnet **156** to provide two magnetic flux paths extending from opposite sides of permanent magnet. Control coils **158** are positioned along each path.

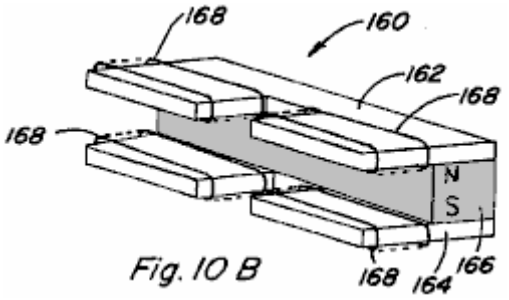


Fig. 10 B

Fig.10B depicts a permanent magnet control component **160** in which pole pieces **162** and **164** are positioned against the pole faces of permanent magnet **166** to provide two spaced, adjacent magnetic flux paths extending from the same side of permanent magnet **166**. Control coils **168** are positioned along each path.

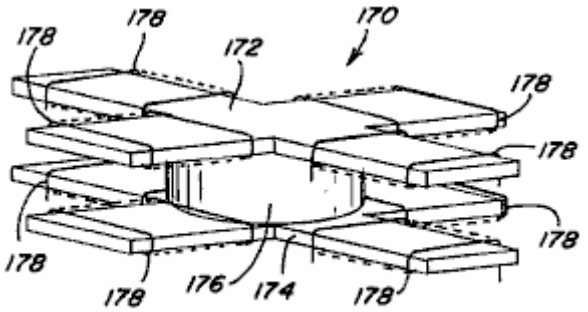


Fig. 10 C

Fig.10C depicts a permanent magnet control component **170** in which pole pieces **172** and **174** are configured so as to be positioned adjacent the pole faces of permanent magnet **176** so as to provide four flux paths, each flux path extending in a respective direction from permanent magnet **176**. Control coils **178** are also positioned along each path.

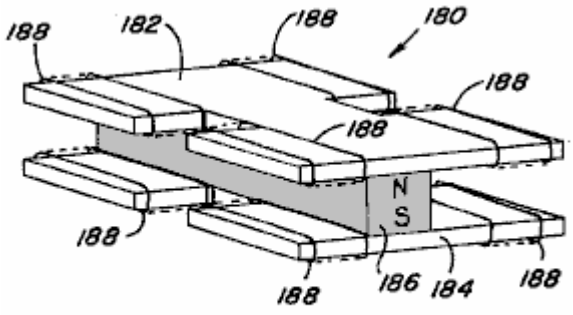


Fig. 10 D

Fig.10D depicts another four-path configuration of a permanent magnet control component **180** in which pole pieces **182**, **184** are configured and positioned to provide four flux paths for permanent magnet **186**, with a pair of spaced, adjacent flux paths extending from each side of permanent magnet **186**. Control coils **188** are positioned along each path.

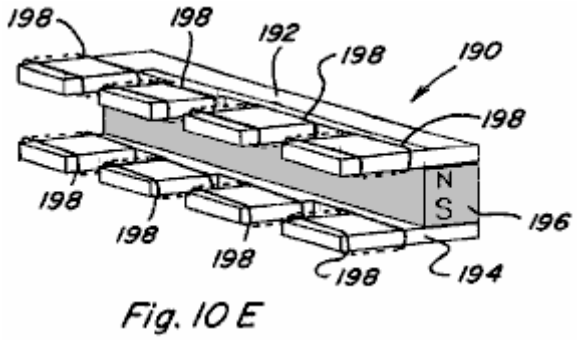


Fig. 10 E

Fig.10E depicts another four-path configuration of a permanent magnet control component **190** in which all four flux paths formed by pole pieces **192**, **194** extend from one side of permanent magnet **196**. Again, control coils **198** are positioned along each flux path.

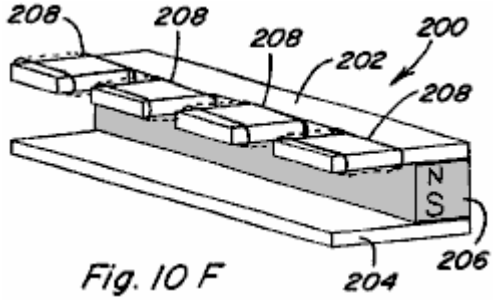


Fig. 10 F

Fig.10F still further depicts a four-path configuration of a permanent magnet control component **200** in which pole pieces **202**, **204** extend to one side of permanent magnet **206**, with pole piece **202** defining four flux paths and with pole piece **204** including a continuous return path. Control coils **208** are positioned along each path of pole piece **202**. Many other variations are possible.

Accordingly, it is seen that a variety of different configurations of permanent magnet control components are possible, in accordance with the present invention. The important considerations for division of permanent magnet flux in such permanent magnet control components include, extending each pole piece to, or beyond, the outer perimeter of the pole face of the permanent magnet in each region where a flux path is intended and assuring that the pole face of the permanent magnet intersects each of the flux paths. It is not necessary for each pole piece to include the same number of path portions extending beyond the perimeter of the respective permanent magnet pole face as noted with reference to permanent magnet control component **200**. Although two control coils are shown along each of the flux paths in **Figs.10A-10E**, it is apparent from component **200** in **Fig.10F** that one control coil positioned along a flux path is generally sufficient for purposes of the present invention. Further, although in the illustrated configurations each pole piece is positioned to contact a respective pole face of the permanent magnet, a small spacing between a pole piece and its adjacent permanent magnet pole face could be provided, particularly in applications where relative movement between the subject pole piece and the permanent magnet will occur.

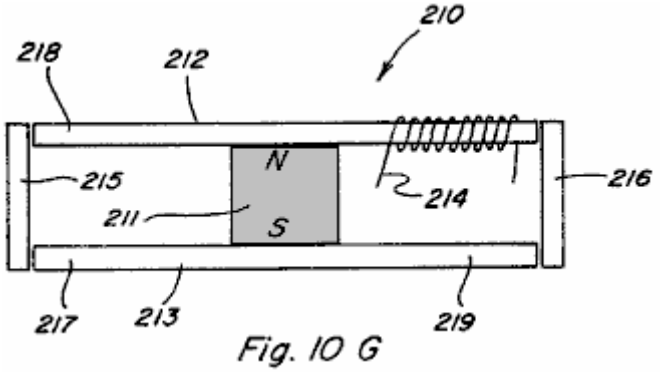


Fig. 10 G

In its simplest form a two path permanent magnet control component only requires one control coil positioned along one of the control paths to permit the magnetic flux of a permanent magnet to be switched between the two paths. In particular, a side view of such a two path component **210** is shown in **Fig.10G** and includes a permanent magnet **211** pole pieces **212** and **213**, and control coil **214** which may be connected to a suitable control circuit. By alternating energising control coil **214** in an opposing manner and an aiding manner the magnetic flux of permanent magnet can be switched between the path including armature **215** and the path including armature **216**. When control coil **214** is energised in an opposing manner the magnetic flux will traverse the path including armature **215** and when control coil **214** is energised in an aiding manner the magnetic flux will traverse the path including armature **216**. Control coil 214 could also be placed at any of the positions **217**, **218**, or **219** to achieve the flux path switching.

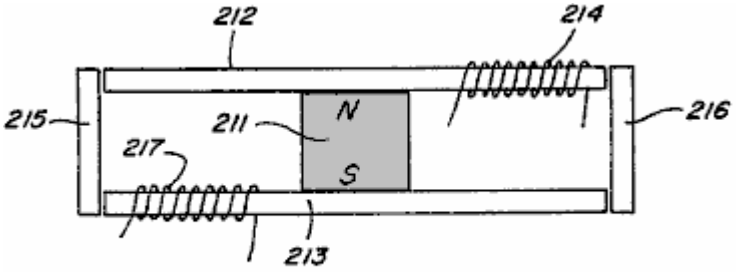


Fig. 10 H

Further, in the two coils embodiment shown in **Fig.10H** control coil **217** is added. In such a device, flux switching can be achieved by simultaneously energising control coil **214** in a flux aiding manner and control coil **217** in a flux opposing manner, and by then simultaneously reversing the energisation of the respective control coils **214** and **217**.

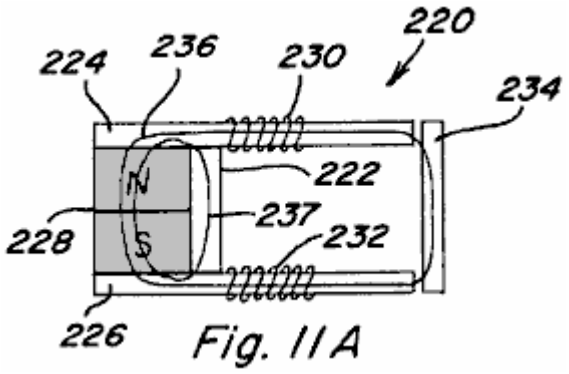


Fig. 11A

Reference is made to **Figs.11A-11F** which depict devices similar to that of **Figs.5-6** except that a bypass, formed of magnetic material, is provided in each case. In device **220** of **Figs.11A-11C** a bypass **222** is provided from pole piece **224** to pole piece **226** and is located between permanent magnet **228** and control coils **230**, **232**, with armature **234** located adjacent the ends of pole pieces **224**, **226**. In **Fig.11A** with no coil energisation, magnet flux components **236** and **237** travel as shown.

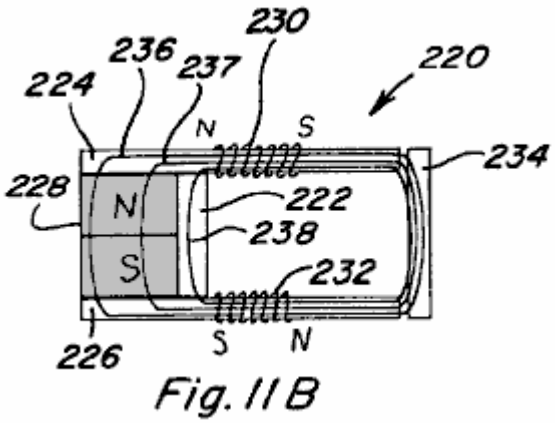


Fig. 11B

When coils **230** and **232** are energised in an aiding or adding manner as in **Fig.11B**, the result is permanent magnet magnetic flux components **236** and **237** travelling as shown, and with the added magnetic flux component

238 from coils **230** and **232** also travelling as shown. Thus, in device **220** energising the coils in an aiding manner results in an increased magnetic coupling force on armature **234**.

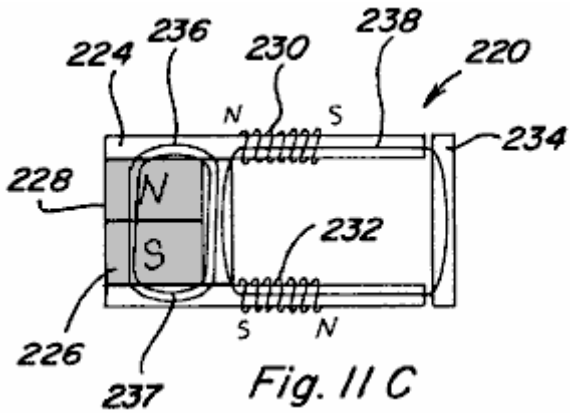


Fig. 11 C

In **Fig.11C** coils **230**, **232** are energised in an opposing exceeding manner which results in permanent magnetic flux components **236** and **237** travelling as shown and excess magnetic flux component **238** travelling as shown. Thus, in device **220** energising the coils in an opposing exceeding manner results in magnetic coupling force on armature **234**, albeit smaller than that in the aiding exceeding case.

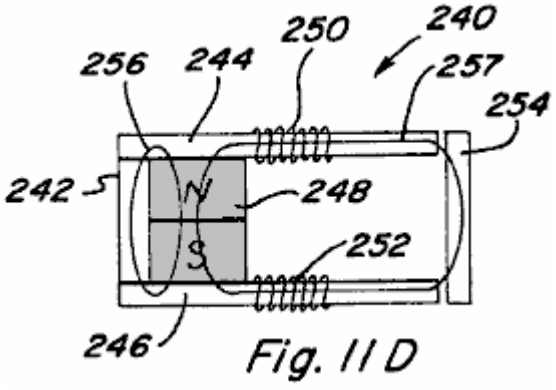


Fig. 11 D

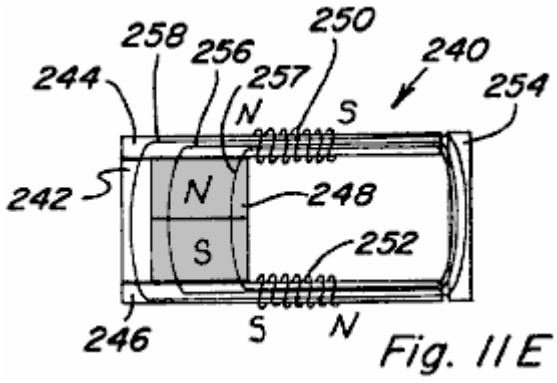


Fig. 11 E

In device **240** of **Figs.11D-11F** a bypass **242** is provided between pole piece **244** and pole piece **246** but is located on an opposite side of permanent magnet **248** as compared to control coils **250**, **252** and armature **254**. Permanent magnet flux components **256** and **257** are shown for no coil energisation in **Fig.11D**. In **Fig.11E** the paths of permanent magnet flux components **256** and **257**, as well as excess coil magnetic flux **258**, are shown when coils **250**, **252** are energised in an aiding exceeding manner.

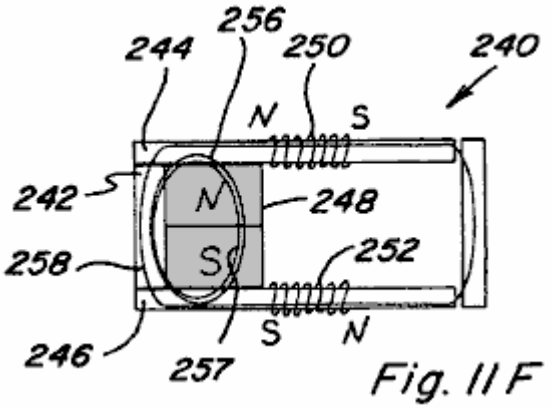


Fig. 11F

In **Fig.11F** the path of each magnetic flux component **256**, **257**, and **258** is shown when coils **230**, **232** are energised in an opposed exceeding manner.

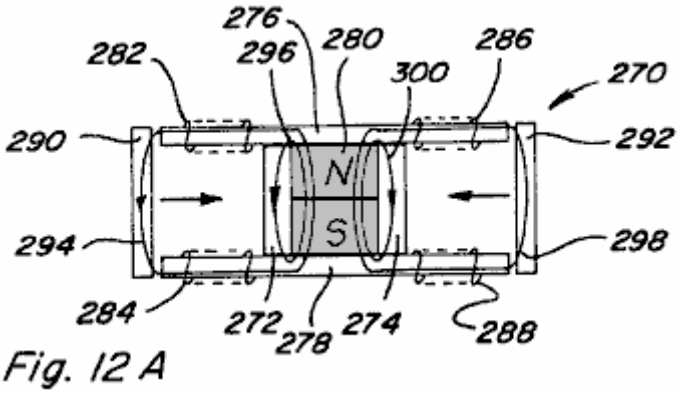


Fig. 12 A

Figs.12A-12E depict a device **270** similar to that shown in **Figs.7-9** except that bypasses **272** and **274** are provided from pole piece **276** to pole piece **278**. Bypass **272** is located between permanent magnet **280** and control coils **282**, **284** and bypass **274** is located between permanent magnet 280 and control coils **286**, **288**. Armatures **290** and **292** are also provided. When no coils are energised permanent magnet magnetic flux components **294**, **296**, **298**, and **300** travel as shown in **Fig.12A**.

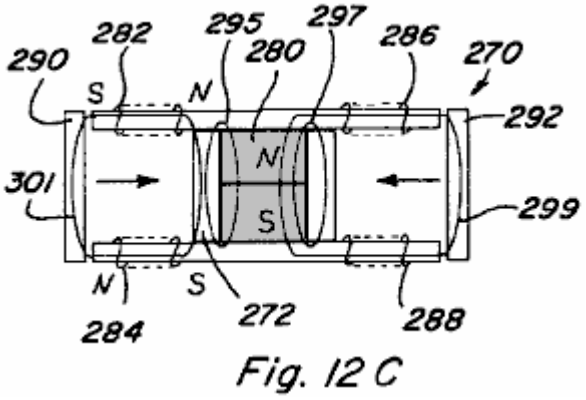


Fig. 12 C

If coils **282**, **284** are energised in an opposing manner permanent magnet flux components **295**, **297**, and **299** travel as shown, with no flux component traversing the path which includes armature **290** and therefore no magnetic coupling force acting thereon. This would be the case when coils **282**, **284** are energised to the level where the coils magnetic flux just blocks, but does not exceed, the magnetic flux component **294** (**Fig.12A**) from permanent magnet **280**. However, if coils **282**, **284** are energised in an opposed exceeding manner an excess coil magnetic flux component **301** is produced which travels a path including armature **290** and bypass **272** results as shown in **Fig.12C**.

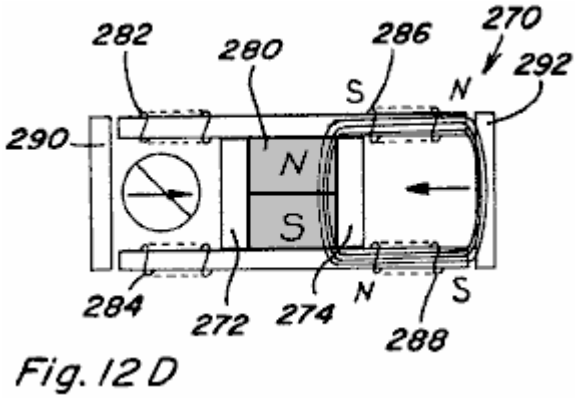


Fig. 12 D

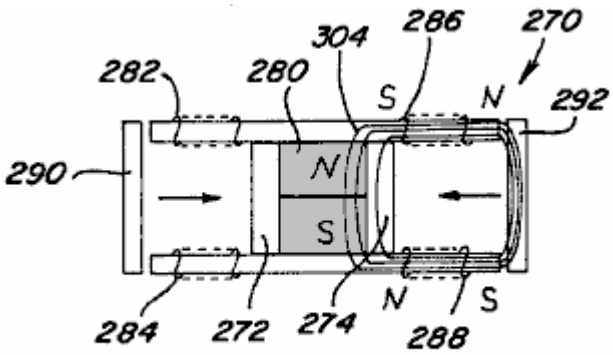


Fig. 12 E

Coils **286**, **288** may be energised in an aiding manner such that all permanent magnet magnetic flux travels along the path which includes armature **292** as shown in **Fig.12D**. If coils **286**, **288** are energised in excess of the level of **Fig.12D** then the excess magnetic flux component **304** traverses the path which includes armature **292** and bypass **274** as shown in **Fig.12E**, thereby increasing the magnetic coupling force on armature **292** as compared to **Fig.12D**. The advantage of incorporating such bypasses into permanent magnet control components in certain applications will become apparent below.

Reciprocating Motion

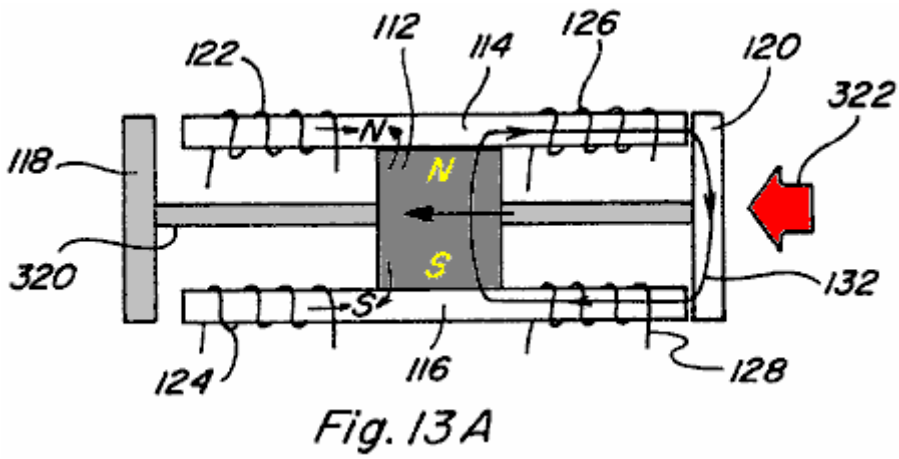


Fig. 13 A

As mentioned above, controlling the path of magnetic flux from a permanent magnet can be useful in a variety of applications such as achieving reciprocating motion. In this regard, if the device **110** of **Figs.7-9** is modified such that armatures **118** and **120** are fixed to a sliding shaft **320** as shown in **Figs.13A-13C**, and if the distance between the armatures is greater than the end to end length of pole pieces **114**, **116**, limited linear motion in two directions (left and right in **Figs.13A-13C**), and therefore linear reciprocating motion, can be achieved by the timed, alternate delivery of electrical signals to control coils **122**, **124** and control coils **126**, **128**. By way of example, **Fig.13A** represents the position of shaft connected armatures **118**, **120** when coils **122**, **124** are energised in an opposing manner to block the flux of permanent magnet **112** such that all magnetic flux traverses path **132** as shown and such that the resulting magnetic coupling force acts to the left as indicated by arrow **322**.

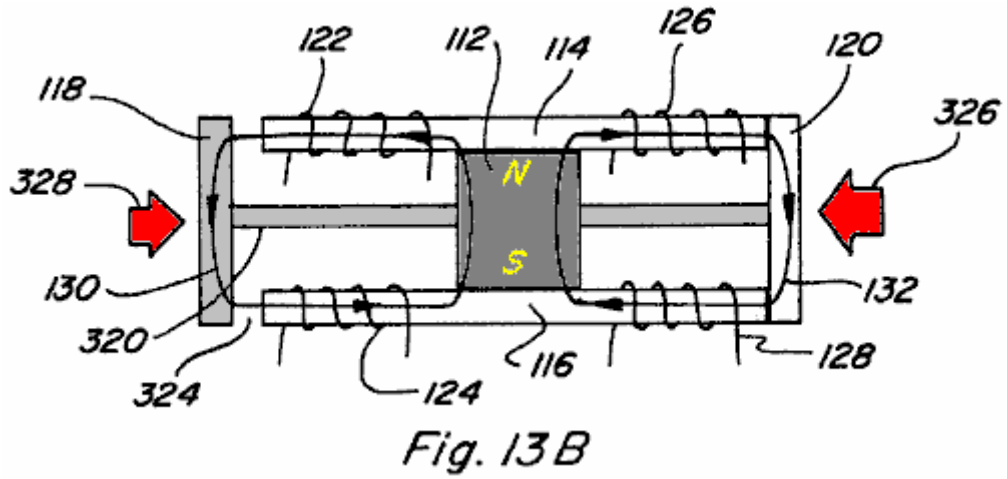


Fig. 13B

As shown in **Fig.13B** when coils **122**, **124** are de-energised the magnetic flux from permanent magnet **112** can again travel along path **130** through armature **118**. However, due to the air gap **324** between armature **118** and pole pieces **114**, **116** the reluctance along path **130** will be significantly greater than the reluctance along path **132**. Accordingly, the amount of magnetic flux which flows along path **130** will be less than the amount of magnetic flux which flows along path **132** such that the magnetic coupling force on armature **118** acting to the right will be significantly less than the magnetic coupling force on armature **120** acting to the left as shown by arrows **326** and **328**, which arrows are sized to represent the strength of the respective directional force.

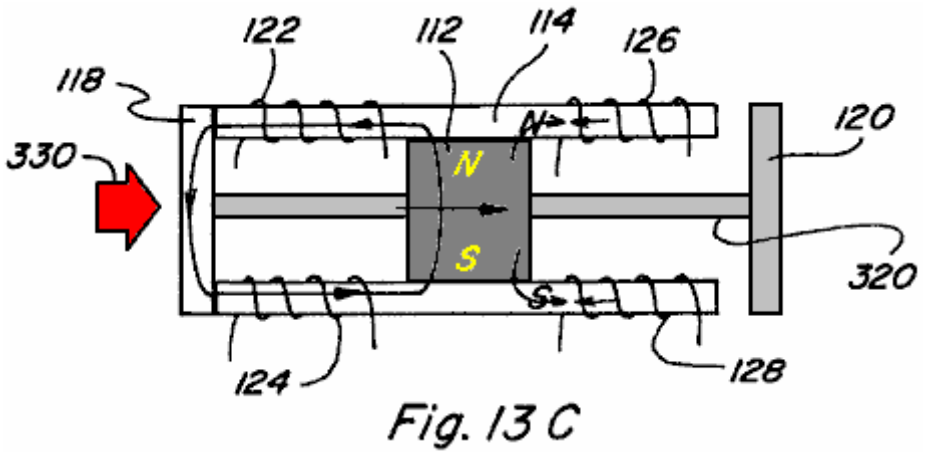


Fig. 13C

Fig.13C represents the position of shaft connected armatures **118**, **120** after coils **126**, **128** are energised in a manner to oppose the flux of permanent magnet **112** such that all flux traverses path **130** and the resulting magnetic coupling force on armature **118**, depicted by arrow **330**, moves the shaft **10** connected armatures **118**, **120** to the right.

Control coils **122**, **124** and **126**, **128** could also be energised in a flux aiding manner to achieve the same result. In such a device, **Fig.13A** would represent coils **126**, **128** energised to aid magnetic flux along path **132**, **Fig.13B** would again represent no coils energised, and **Fig.13C** would represent coils **122**, **124** energised to aid magnetic flux along path **130**.

Thus, by alternately energising and de-energising control coils **122**, **124** and **126**, **128** a linear reciprocating motion of shaft connected armatures **118**, **120** may be achieved. Further, such reciprocating motion may be achieved by energising the coils in either an opposing or aiding manner. The magnetic coupling force exerted on a given armature when **20** the control coils are energised to establish all magnetic flux along a single path which includes that armature is significantly greater than the magnetic coupling force which would be exerted on such armature by an identical energisation of the control coils in the absence of the permanent magnet.

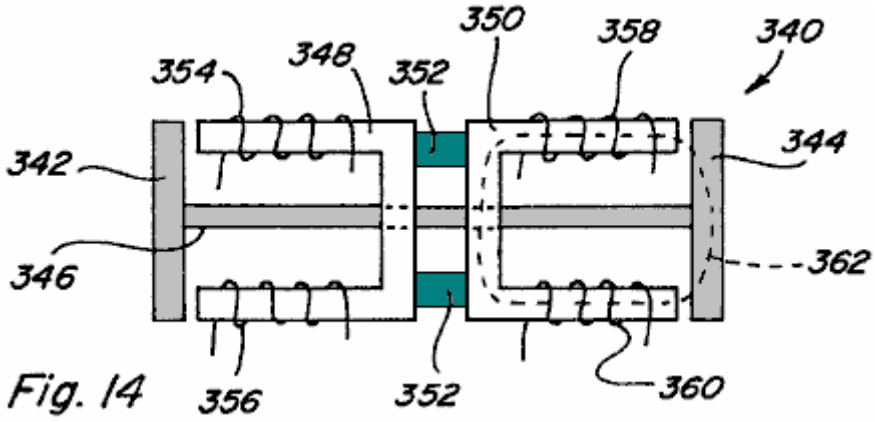


Fig. 14

This is demonstrated with reference to **Fig.14** which depicts a reciprocating device **340** in which only coils or electromagnets are utilised. As shown armatures **342** and **344** are connected by shaft **346**, and each armature **342**, **344** includes a respective U-shaped pole path piece **348**, **350** which pole path pieces are mechanically connected by a non-magnetic material **352**. Each pole path piece **348** and **350** has respective control coils **354**, **356** and **358**, **360** positioned along them. By comparison with the device of **Figs.13A-13C**, if coils **358**, **360** of device **340** are energised to cause magnetic flux flow in either direction, clockwise or counterclockwise, along path **362**, the amount of electrical energy which would be required in order to achieve the same magnetic coupling force on armature **344** as achieved on armature **120** above in **Fig.13A** would be twice that delivered to coils **122**, **124** or **126**, **128** in **Fig.13A**. It is therefore demonstrated, that by controlling or switching the flow of magnetic flux from a permanent magnet between at least two different paths results in greater coupling forces per unit of input electrical energy, and therefore that such control or switching will enable more work to be achieved per unit of input electrical energy.

As described above, if a coil is energised beyond the point where the magnetic flux produced by the coil aiding the amount of the permanent magnet's flux that is either opposed or aided, the extra magnetic flux needs a low reluctance path between the poles of the coil that produces the excess magnetic flux. If a complete low-reluctance path is not provided for the excess magnetic flux, there is little potential for taking advantage of the excess magnetic flux in terms of producing additional magnetic coupling forces. The path for such excess flux cannot be through a permanent magnet member. In assemblies which include an armature on each path, the armature will provide the necessary low-reluctance path.

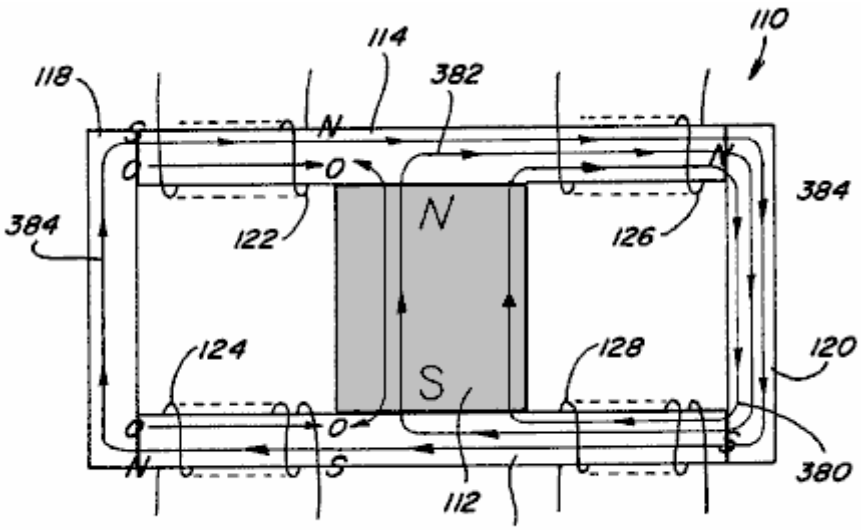


Fig. 15

Referring to **Fig.15**, various components of the magnetic flux in device **110** (**Figs.7-9**) are depicted by numerals **380**, **382**, and **384** for the case when coils **122**, **124** are energised to oppose the magnetic flux of permanent magnet **112** in an amount which exceeds the level of magnetic flux which permanent magnet **112** would cause to flow through armature **118** when no coils are energised. **Fig.15** is likewise representative of the case when coils **126**, **128** are energised to aid the magnetic flux of permanent magnet **112** by an amount which exceeds the level of magnetic flux which permanent magnet **112** would cause to flow through armature **118** when no coils are energised. In particular, magnetic flux component **380** represents the magnetic flux of permanent magnet **112** which normally flows through the path including armature **120**; magnetic flux component **382** represents the magnetic flux of permanent magnet **112** which is diverted by the opposing field of coils **122**, **124** so as to traverse

the path which includes armature **120**; and magnetic flux component **384** represents the magnetic flux produced by coils **122**, **124** which is in excess of the diverted magnetic flux **382**. As shown, the excess magnetic flux **384** produced by coils **122**, **124** traverses the path which includes armature **120** and bypasses permanent magnet **112** so as to also traverse the path which includes armature **118**. Thus, the excess magnetic flux produced by coils **122**, **124** adds to the permanent magnet flux traversing the path which includes armature **120**, thus increasing the magnetic coupling force on armature **120**, while at the same time providing a magnetic coupling force on armature **118**.

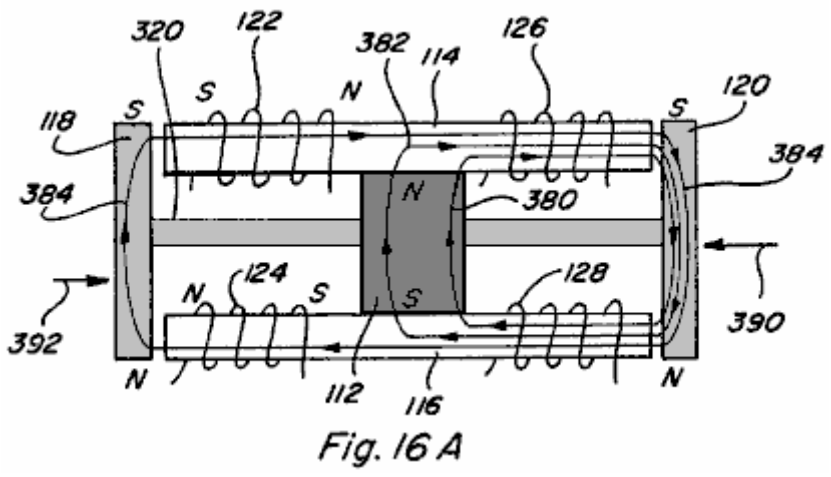


Fig. 16 A

In a reciprocating device where armatures **118** and **120** are connected by shaft **320** as shown in **Figs.13A-13C** and again in **Fig.16A**, excess magnetic flux **384** will increase magnetic coupling force **390** on armature **120** acting to the left. However, because such excess flux **384** also traverses the path which includes armature **118**, such excess magnetic flux **384** also results in a magnetic coupling force **392** on armature **118** which acts to the right. Even though excess magnetic flux **384** traversing the path which includes an armature **118** has an opposite polarity to that which would traverse the path due to permanent magnet **112**, the magnetic coupling force on armature **118** still acts to the right because armature **118** is not polarity sensitive, that is, armature **118** will be attracted regardless of the direction of the magnetic flux traversing the path. The overall effect is that a resultant force which is the difference between force **390** and force **392** will act on the shaft-connected armatures **118**, **120**. However, if armatures **118** and **120** were formed by permanent magnets having polarities as shown at the top and bottom of such armatures, the force acting on each armature would be in the same direction and therefore additive.

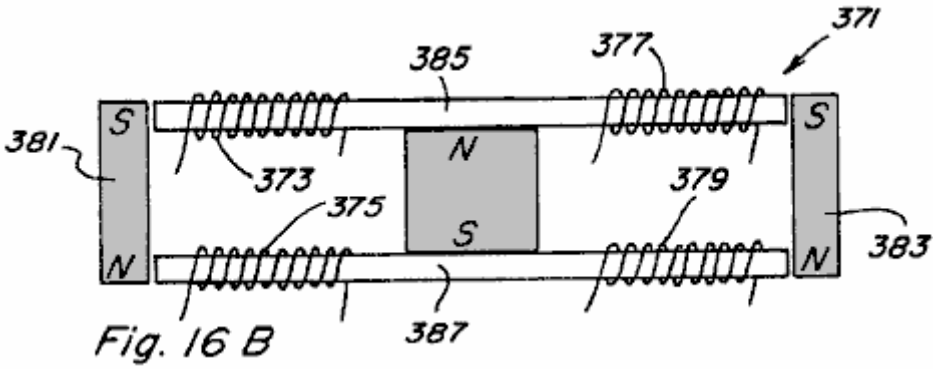


Fig. 16 B

In this regard reference is made to **Fig.16B** in which a two path device **371** having four control coils **373**, **375**, **377** and **379** is shown with the illustrated armatures being formed by permanent magnets **381** and **383** having polarities as shown. With no coils energised both permanent magnet armatures **381** and **383** are attracted to the ends of pole pieces **385** and **387**. With coils **373**, **375** energised in an opposing manner and coils **377**, **379** energised in an aiding manner, the attractive force on permanent magnet armature **383** will generally increase and the attractive force on permanent magnet armature **381** will generally decrease.

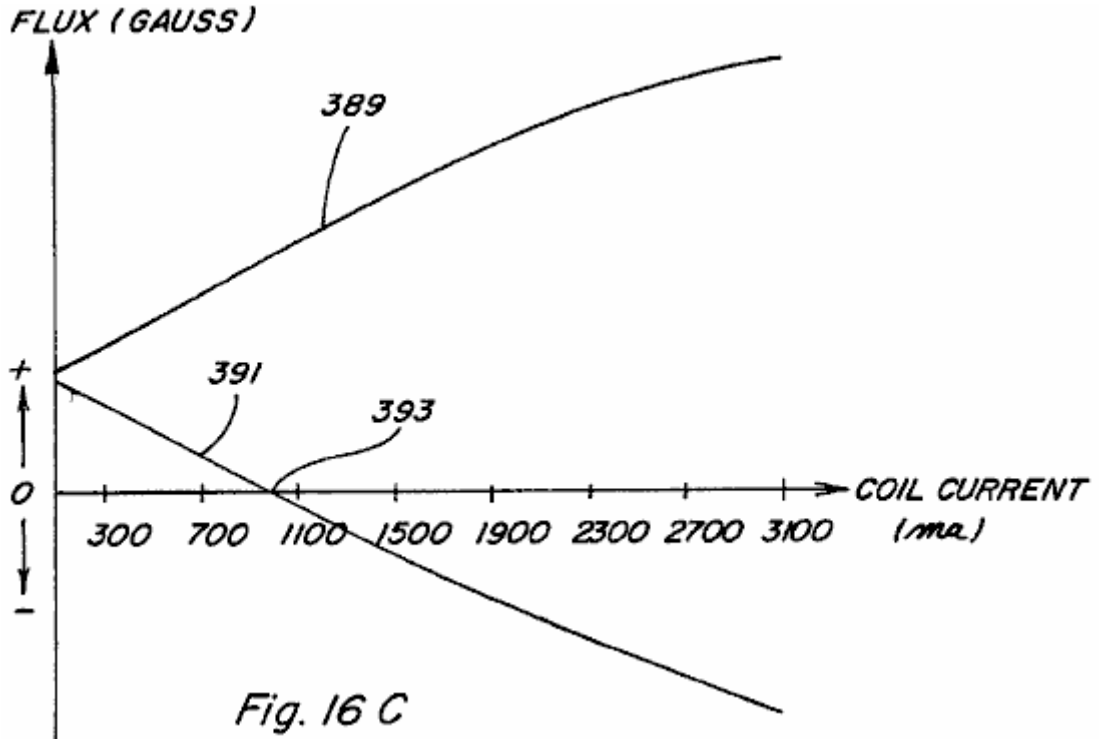


Fig. 16 C

This is demonstrated with reference to the graph of **Fig.16C** which depicts a graph of the current flowing in the control coils on the x-axis verses the magnetic flux in gauss on the y-axis with line **389** representing the flux along the aiding side of device **371** and line **391** representing the flux along the opposing side of device **371**. As shown, the magnetic flux on the coil opposing side decreases as the coil current increases and passes through zero at point **393**. After point **393**, reverse magnetic flux begins to be produced and would result in a repelling force on permanent magnet armature **381**. In some applications, particularly those where permanent magnet armatures and rotors are not utilised, it is critical to recognise point **393** so that reverse magnetic flux is not produced.

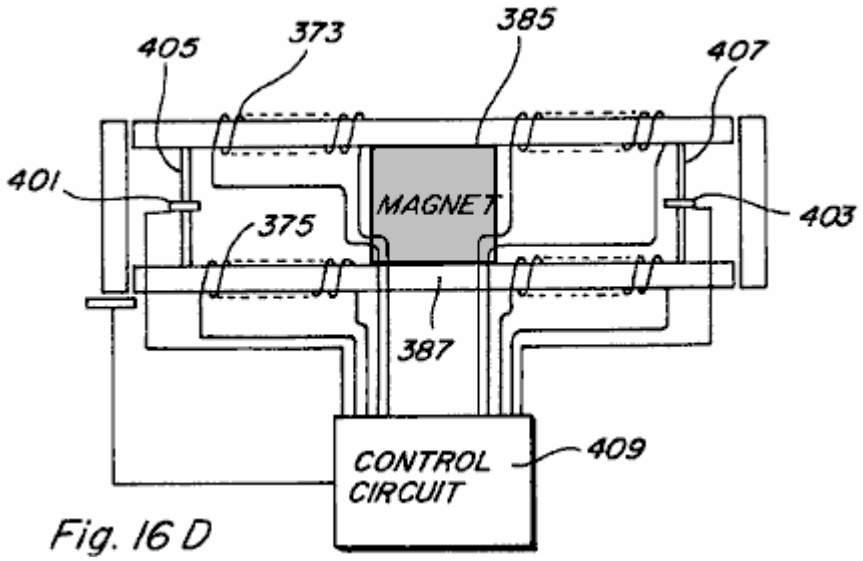


Fig. 16 D

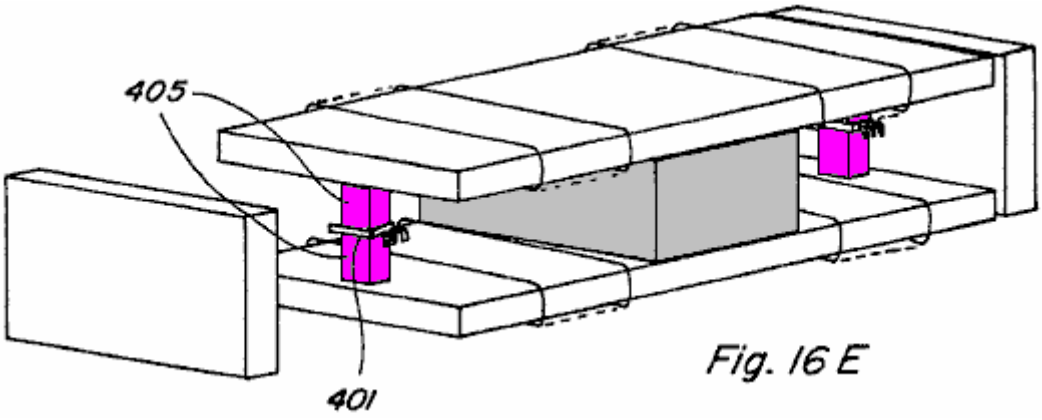


Fig. 16 E

In this regard, reference is made to **Fig.16D** and **Fig.16E**, in which use of Hall Effect switches **401** and **403** is made to enable control of the coil energising current in situations where it is desirable to prevent reverse magnetic flux. As shown, small bypasses **405** and **407** are provided with Hall Effect switches **401** and **403** positioned in gaps along them, the switches being connected to control circuit **409**. As the flux travelling along the bypass path falls to zero, the Hall Effect switch can be utilised to prevent further energisation of the control coils so that no reverse flux is created.

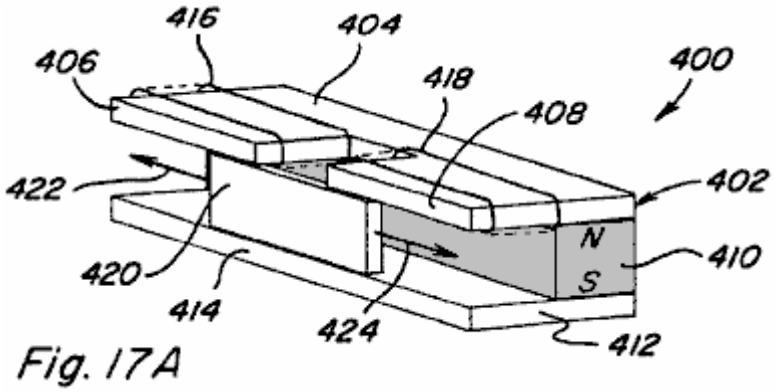


Fig. 17A

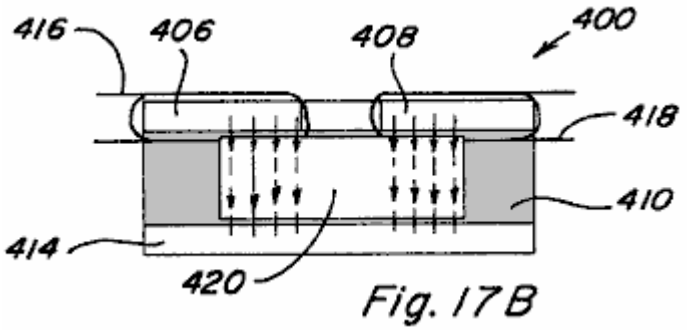


Fig. 17B

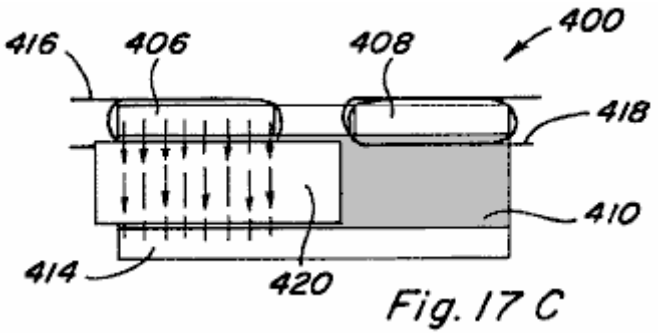


Fig. 17 C

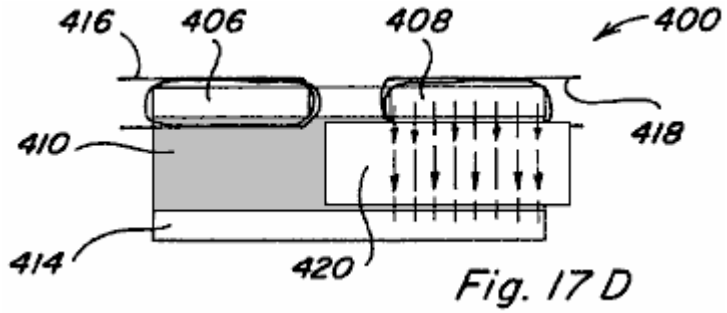


Fig. 17 D

Another embodiment of a device **400** which would provide reciprocating motion is shown in **Figs.17A-17D** in which a permanent magnet control component **402** having two flux paths may is provided. A first pole piece **404**, has two spaced, adjacent path portions **406** and **408** extending beyond the perimeter of the pole face of permanent magnet **410**, and a second pole piece **412** includes only one continuous portion **414** extending beyond the perimeter of the pole face of permanent magnet **410**, each path portion **406** and **408** of pole piece **404** being substantially aligned with at least a part of portion **414** of pole piece **412**. Control coil **416** is positioned along pole piece path portion **406** and control coil **418** is positioned along pole piece portion **408**. An armature **420** is positioned in the region between pole piece path portions **404**, **406** and pole piece portion **414** and is free to slide from side to side as shown by arrows **422** and **424**.

A front view of component device **400** with no coils energised and armature **420** at a mid-point depicts flux flowing from the north pole face of permanent magnet **410**, through each of pole piece path portions **406** and **408**, through armature **420**, and returning to the south pole face through pole piece portion **414**. Thus, the magnetic flux divides equally along two paths. If coil **416** is energised in an aiding manner, or if coil **418** is energised in an opposing manner, all or a majority of the magnetic flux of the permanent magnets can be made to flow through pole piece portion **406** so that a resulting magnetic coupling force on armature **420** causes it to move to the left as shown in **Fig.17C**.

Likewise, if control coil **416** is energised in an opposing manner, or if control coil **418** is energised in an aiding manner, all or a majority of the permanent magnet flux can be made to flow through pole piece path portion **408** such that a resulting magnetic coupling force on armature **420** causes it to move to the right as shown in **Fig.17D**. Accordingly, by alternately energising and de-energising coils **416** and **418** a reciprocating motion of armature **420** may be achieved.

Linear Motion

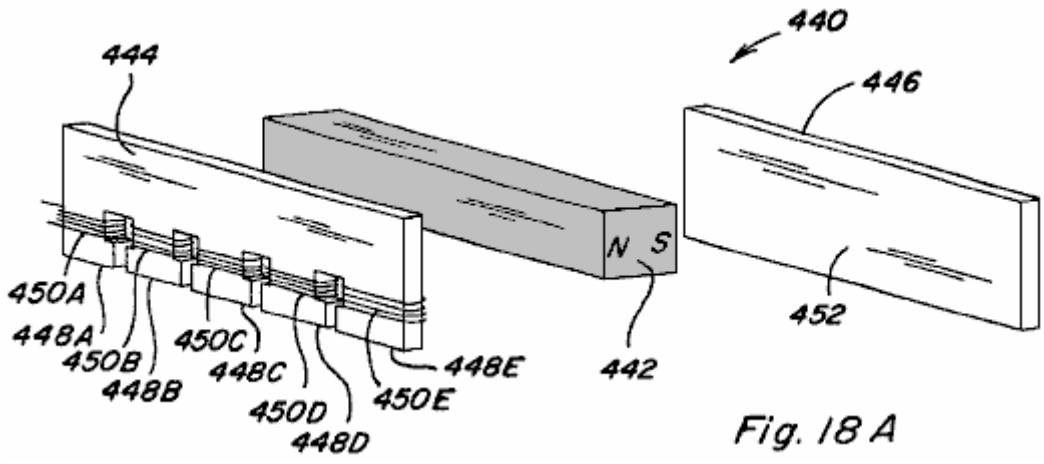


Fig. 18 A

Referring now to **Figs.18A-18E**, linear motion in accordance with the present invention is described. In particular, a permanent magnet control component **440** including a permanent magnet **442** with a pole piece **444** positioned against it's north pole face and a pole piece **446** positioned against it's south pole face is shown in an exploded view in **Fig.18A** and assembled in **Fig.18B**.

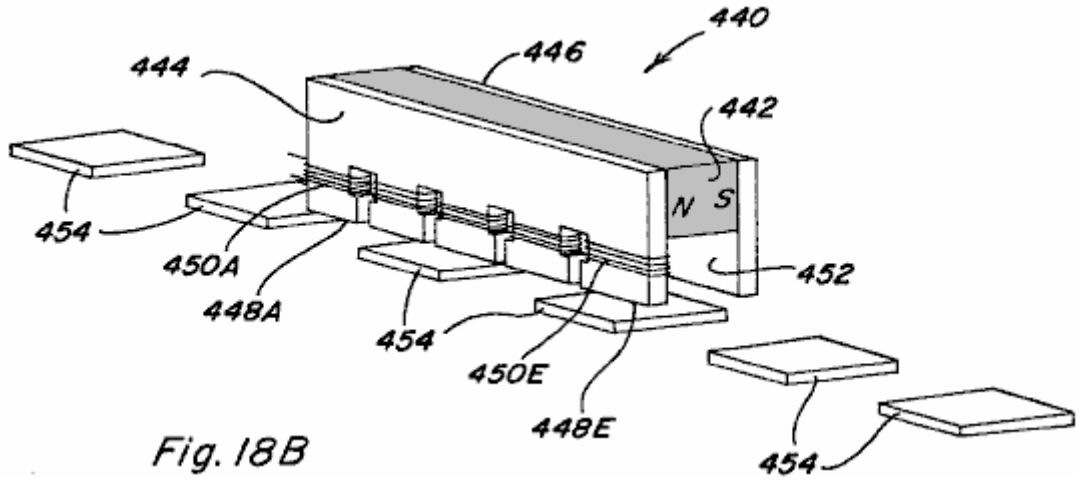


Fig. 18B

Pole piece **444** includes five path portions **448A-448E** which extend beyond the edge of the north pole face of permanent magnet **442** to one side of it and at respective positions along it's length, and it has path portion **448A-448E** each with a control coil **450A-450E** positioned around them. Pole piece **446** includes one portion **452** extending beyond the edge of the south pole face of permanent magnet **442** to the one side of it, and this portion **452** extends along the entire length of permanent magnet **442**. A number of armatures **454** define a path of relative movement between permanent magnet control component **440** and such armatures **454**, and by providing timed energisation of given control coils **450A-450E** such relative movement can be achieved.

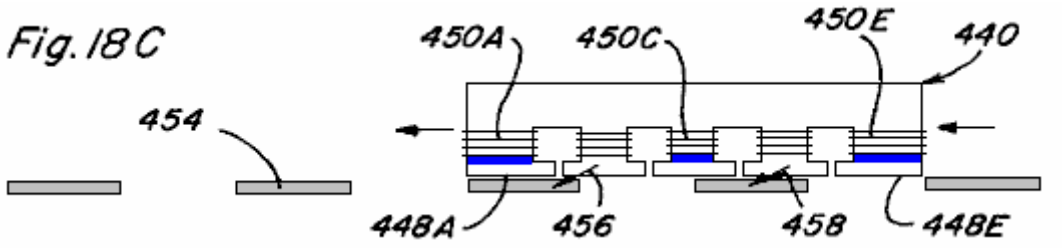


Fig. 18C

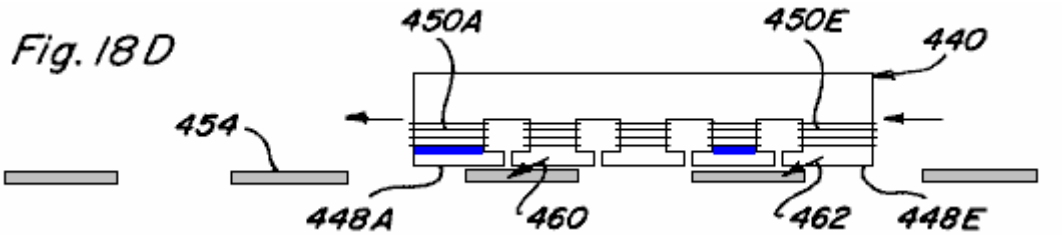


Fig. 18D

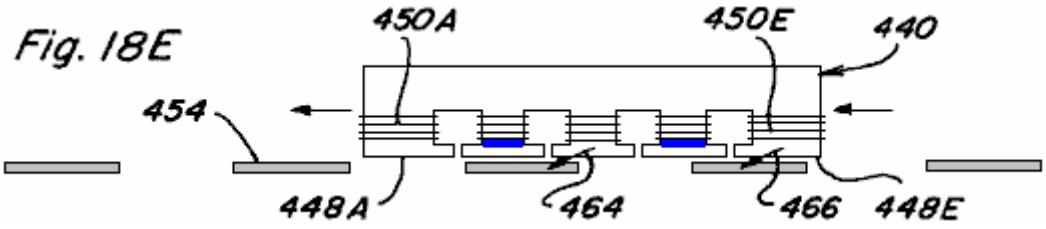


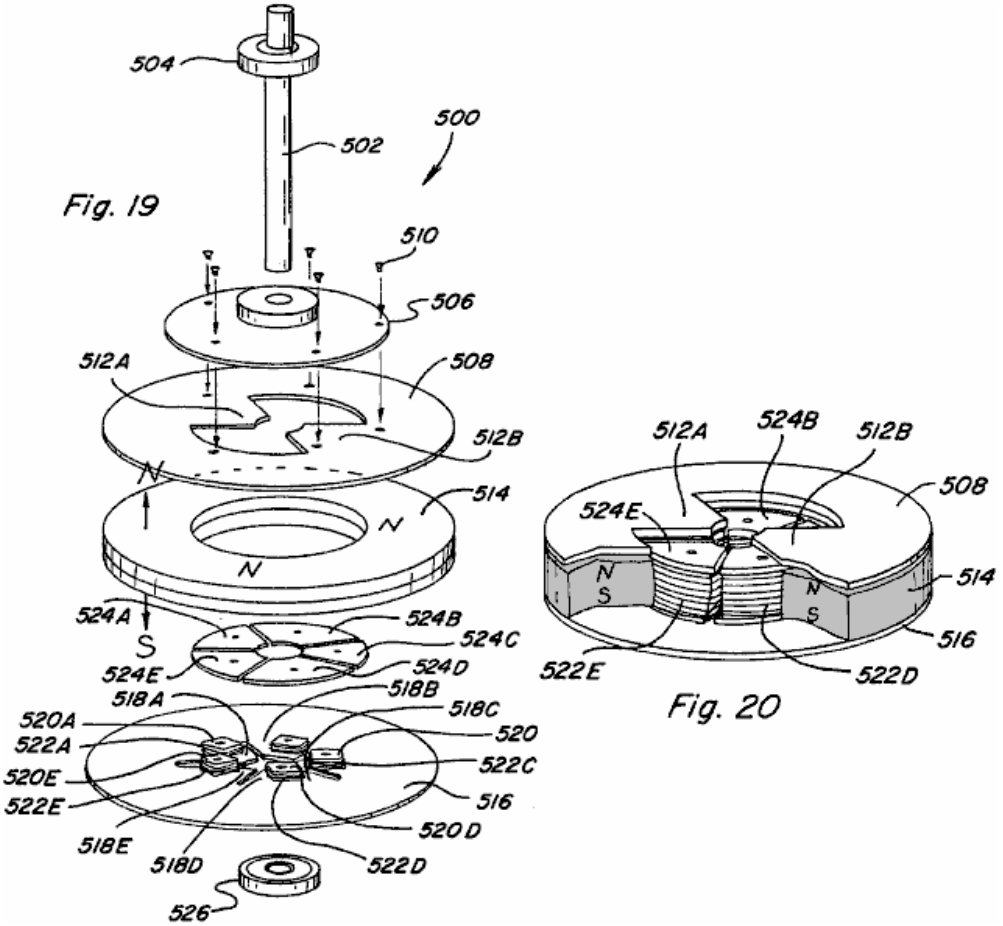
Fig. 18E

The sequence of side views depicted in **Figs.18C-18E** illustrate such relative movement,
with coils **450A**, **450C** and **450E** being **energised** in an opposing manner simultaneously in **Fig.18C**,
with coils **450A** and **450D** being **energised** simultaneously in an opposing manner in **Fig.18D**, and
with coils **450B** and **450D** being **energised** simultaneously in an opposing manner in **Fig.18E**.

In **Fig.18C**, magnetic flux will only flow along path portions **448B** and **448C** of pole piece **444** causing resultant magnetic coupling forces depicted by arrows **456**, **458** which act to move permanent magnet control component **440** to the left, assuming armatures **454** are fixed. Similarly, due to the timing of subsequent coil energisation resultant magnetic forces depicted by arrows **460**, **462** in **Fig.18D** and arrows **464**, **466** in **Fig.18E** act to continue movement of permanent magnet control component **440** to the left. Thus, if permanent magnet control component **440** were fixed to a device or structure, controlled movement of the device or structure along the path defined by armatures **454** could be achieved. Conversely, if permanent magnet control component **440** were fixed and armatures **454** were located on a device or structure, controlled movement of the device or structure

could also be achieved. It is also easily recognised that by varying the coil energisation sequence and timing relative movement in the opposite direction can be achieved. Further, if the permanent magnet was doughnut shaped and the armatures were arranged in a circumferential pattern, rotary motion would likewise be achievable.

Rotary Motion



One embodiment of a rotary motion device or motor **500** which incorporates various permanent magnet flux control aspects of the present invention is shown in the exploded view of **Fig.19** and in the partial assembled view of **Fig.20**. Motor **500** includes a rotor assembly which includes a shaft **502** and associated upper bearing **504**, a non-magnetic disk member **506** mounted for rotation with shaft **502**, and a rotor pole piece **508** which is mounted for rotation with disk member **506** such as by the use of screws **510**. Rotor pole piece **508** includes a ring-shaped portion having two inwardly extending magnetic flux path portions **512A** and **512B**. A stator assembly of motor **500** includes a doughnut or ring-shaped permanent magnet **514** having an upwardly directed north pole face positioned adjacent and in close proximity to rotor pole piece **508**, and a downward directed south pole face positioned adjacent and in contact with a stator pole piece **516**. Stator pole piece includes a ring-shaped portion having five inwardly projecting path portions **518A-518E**. Each path portion includes a respective winding post **520A-520E** extending therefrom and having a respective control coil **522A-522E** wound on it. Stator pole piece faces **524A-524E** are which can be positioned on respective winding posts **518A-518B** and, as shown in the partial assembly of **Fig.20**, are substantially aligned with the top surface of permanent magnet **514** so as to be which can be positioned adjacent rotor path portions **512A** and **512B** when aligned therewith. Each of winding posts **518A-518E** and stator pole piece faces are formed of magnetic material, and although shown as separate pieces, an integral, one piece stator could be formed with similar winding posts and pole piece faces machined on it. Lower bearing **526** is also shown.

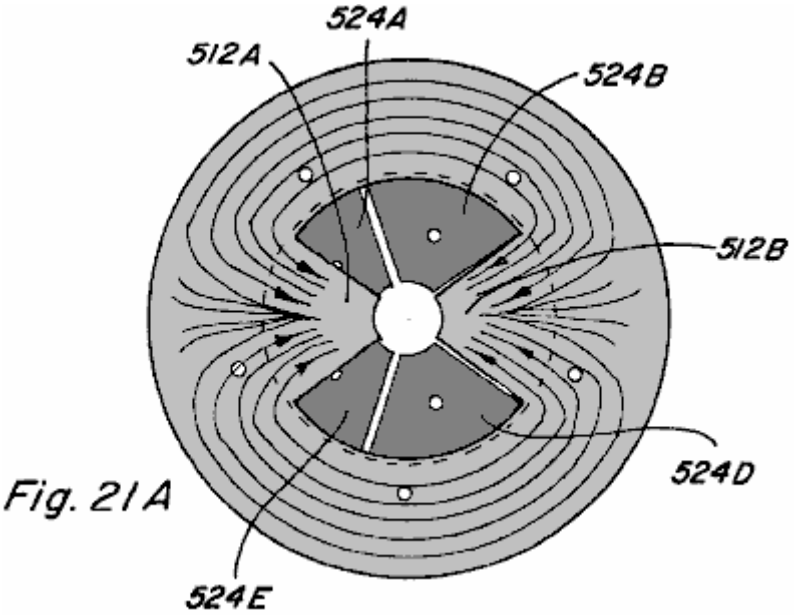


Fig. 21A

Figs.21A-21E illustrate top views of the partial assembly of **Fig.20** with magnetic flux shown. In **Fig.21A** magnetic flux travel when none of coils **522A-522E** are energised is depicted. Disregarding leakage flux, due to the low-reluctance path provided by rotor pole piece path portions **512A** and **512B**, the majority of magnetic flux from the north pole face of permanent magnet **514** will travel radially inward along one of such path portions before passing downward through the stator assembly and returning to the south pole face of permanent magnet **514**. It is noted that rotor pole piece **508** includes two path portions and stator pole piece **516** includes five path portions such that rotor pole piece path portions **512A** and **512B** will always be skewed relative to the stator pole piece faces **524A-524E**. Only one rotor pole piece path portion can directly align with a stator pole piece face at a given time. By alternately energising the control coils of each of the stator pole piece paths, rotary motion of the rotor may be achieved.

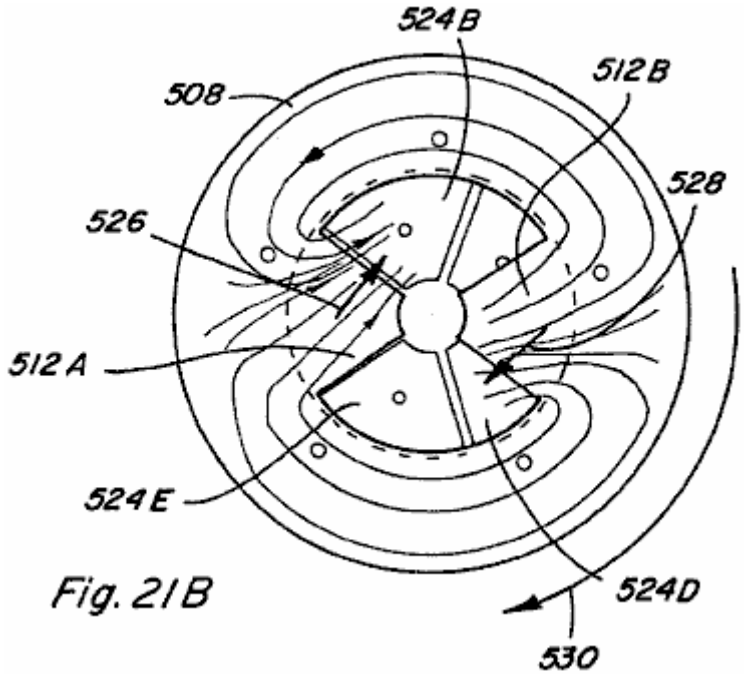


Fig. 21B

In particular, referring to **Figs.21B-21D**, an energising sequence which results in such rotary motion is described. In **Fig.21B**, control coils **522A** and **522C** are energised in a permanent magnet flux opposing manner. Permanent magnet magnetic flux travelling along rotor pole piece path portion **512A** tends to traverse to stator pole piece face **524B** causing a magnetic coupling force indicated by arrow **526**. Likewise, permanent magnet flux travelling along rotor pole piece path portion **512B** tends to traverse to stator pole piece face **524D** causing a magnetic coupling force indicated by arrow **528**. The result is rotation of rotor pole piece **508** in a clockwise direction as indicated by arrow **530**.

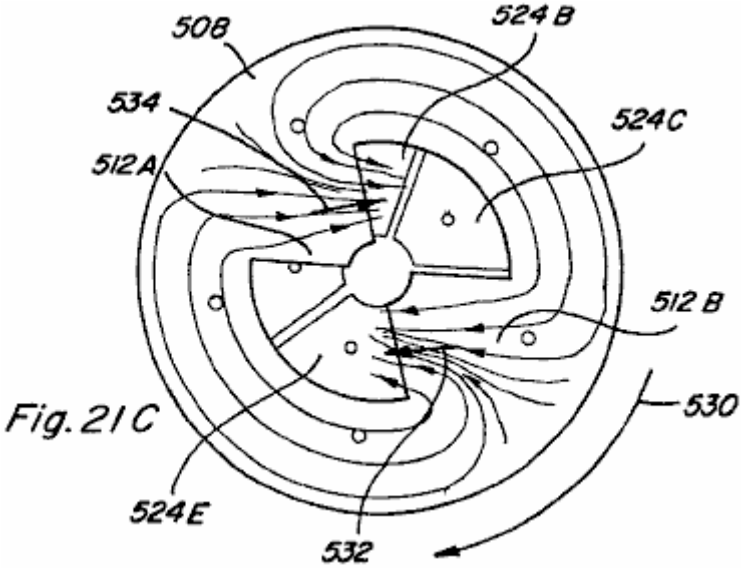


Fig. 21C

Referring to **Fig.21C**, just after rotor pole piece path portion **512B** is no longer aligned with stator pole piece face **524D**, control coil **522C** is de-energised and control coil **522D** is energised in an opposing manner such that the permanent magnet flux travelling along rotor pole piece path **512B** tends to traverse to stator pole piece face **524E** resulting in magnetic coupling force indicated by arrow **532**. Control coil **522A** remains energised such that a magnetic coupling force indicated by arrow **534** results. Accordingly, clockwise rotation of rotor pole piece **508** is continued.

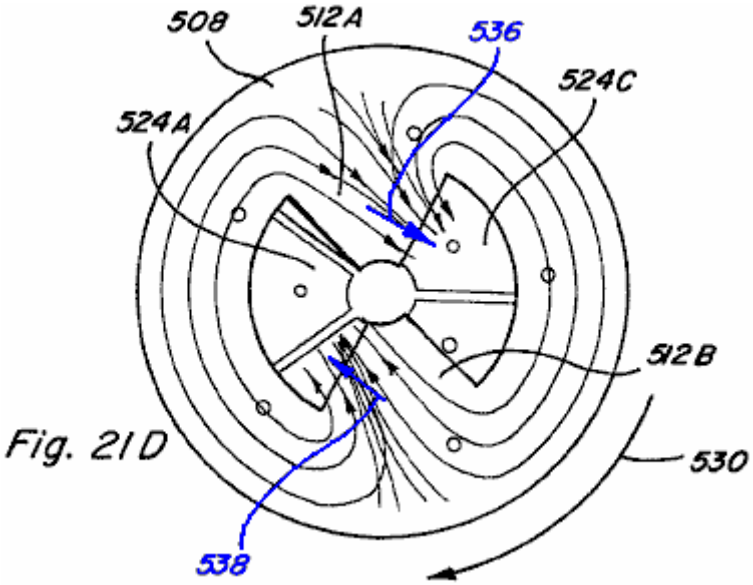


Fig. 21D

In **Fig.21D**, just after rotor pole piece path portion **512A** is no longer aligned with stator pole piece face **524B**, control coil **522A** is de-energised and control coil **522B** is energised in a permanent magnet magnetic flux opposing manner such that the permanent magnet magnetic flux travelling along rotor pole piece path **512A** tends to traverse to stator pole piece face **524C** such that a magnetic coupling force indicated by arrow **536** results. Control coil **522D** remains energised such that a magnetic coupling force indicated by arrow **538** results, and clockwise rotation of rotor pole piece **508** is continued.

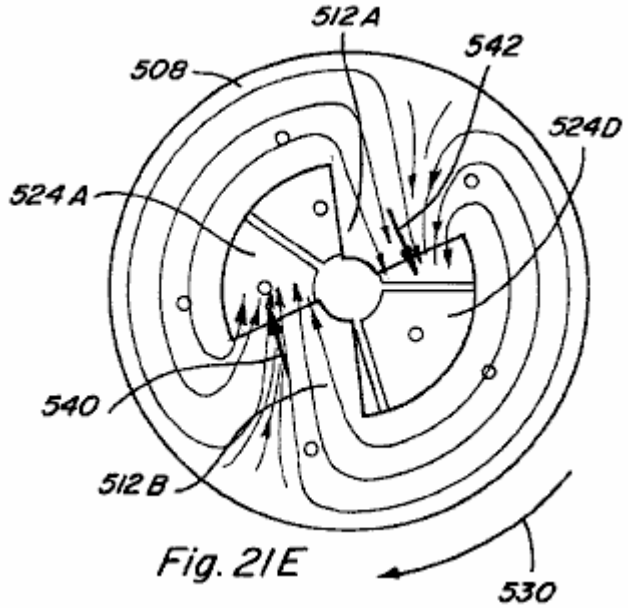


Fig. 21E

As shown in **Fig.21E**, just after rotor pole piece path portion **512B** is no longer aligned with stator pole piece face **524E**, control coil **522D** is de-energised and control coil **522E** is energised in a permanent magnet magnetic flux opposing manner such that the permanent magnet magnetic flux travelling along rotor pole piece path **512B** tends to traverse to stator pole piece face **524A** such that a magnetic coupling force indicated by arrow **540** results. Control coil **522B** remains energised such that a magnetic coupling force indicated by arrow **542** results, and clockwise rotation of rotor pole piece **508** is continued.

Thus, by alternating energising and de-energising control coils **522A-522E**, in a predetermined timed sequence based upon rotation of the rotor assembly, continued rotation movement of rotor pole piece **508** may be achieved. Such an energisation/de-energisation scheme can be achieved utilising circuitry common in the art, such as the control circuitry described in Applicant's U.S. Pat. Nos. 5,463,263 and 5,455,474, as well as various of the circuit configurations described below.

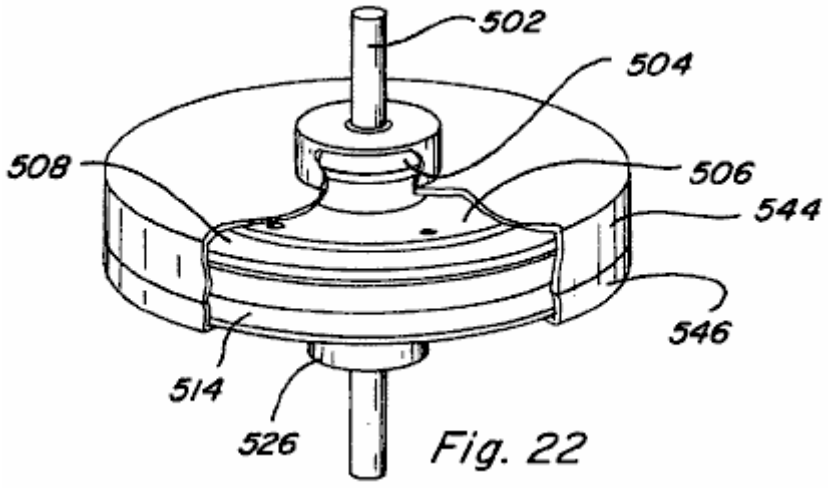


Fig. 22

Referring now to **Fig.22**, an assembled view of rotary motor **500** is shown including a housing or cover formed by an upper housing member **544** and a lower housing member **546**, with portions of each housing member cut away to expose motor structure described above. It is recognised that such housing members **544** and **546** should be constructed from a non-magnetic material, and likewise that motor shaft **502** and bearings **504**, **526** should be constructed from a non-magnetic material.

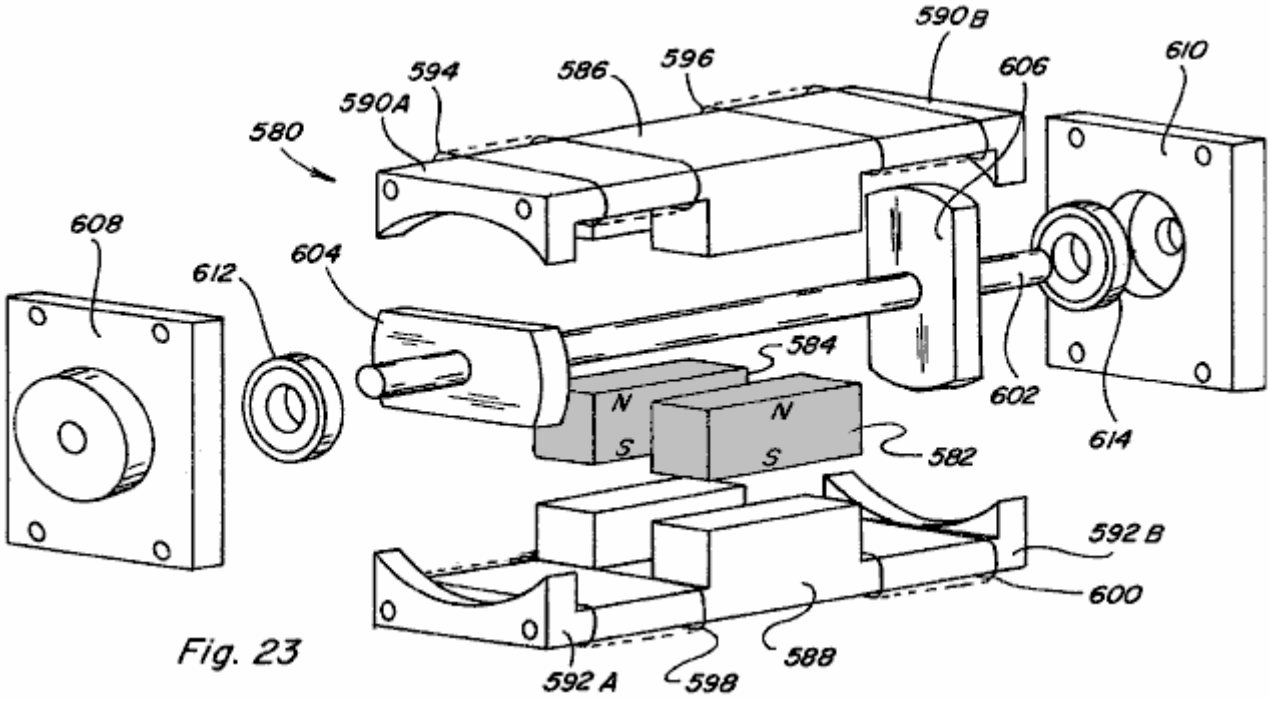


Fig. 23

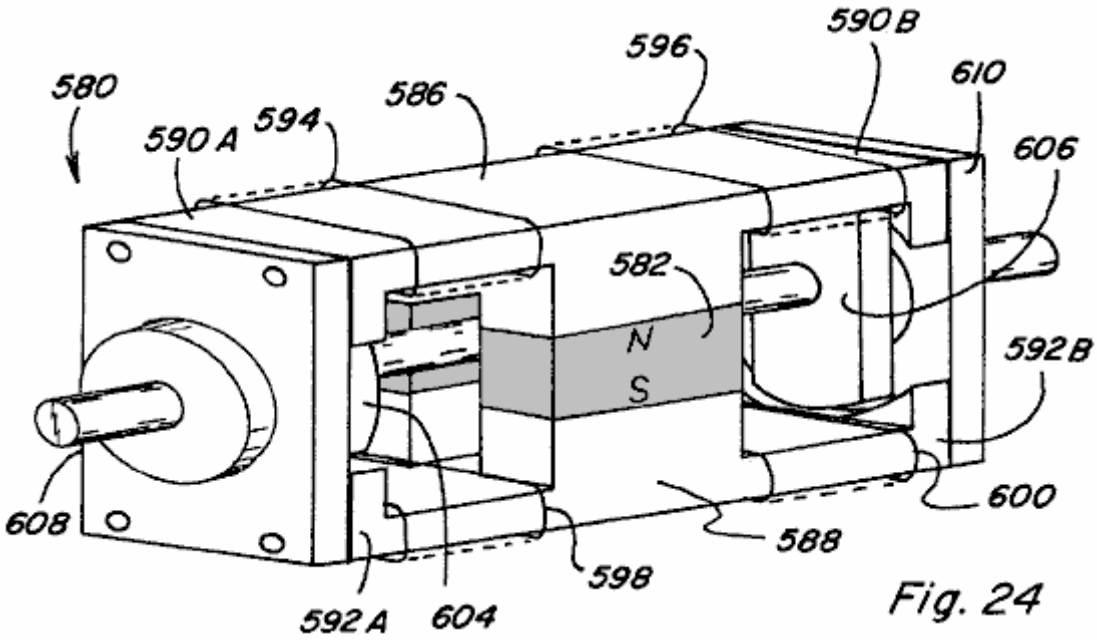
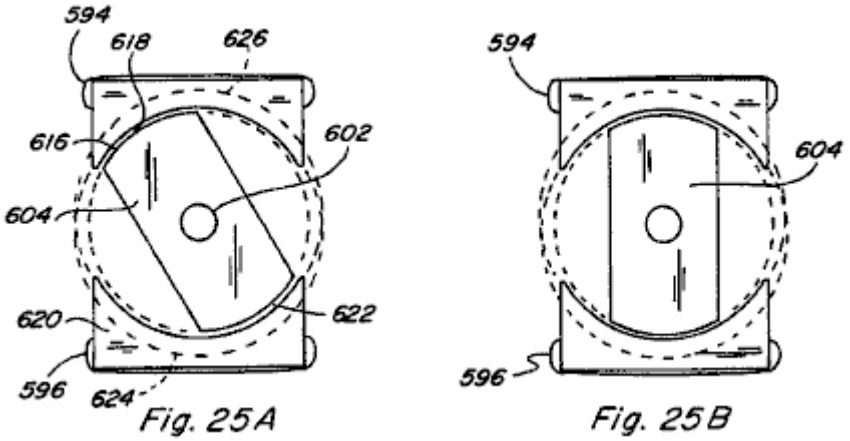


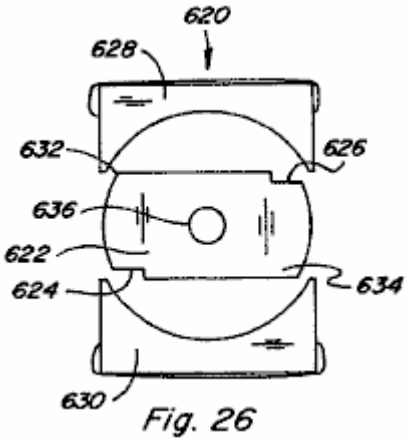
Fig. 24

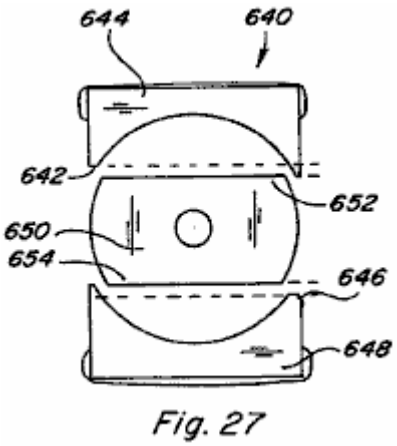
In another embodiment, a rotary motion device or motor **580** in accordance with the present invention is shown in an exploded perspective view in **Fig.23** and in an assembled perspective view in **Fig.24**. Two spaced permanent magnets **582** and **584** are positioned between stator pole pieces **586** and **588**. Stator pole piece **586** includes two path portions **590A** and **590B** extending away from permanent magnets **582**, **584** in opposite directions. Likewise, stator pole piece **588** includes two path portions **592A** and **592B** extending away from permanent magnets **582**, **584** in opposite directions and which can be aligned with stator pole piece path portions **590A** and **590B**. Control coils **594**, **596**, **598**, and **600** are each positioned along a stator pole piece path portion as shown. A non-magnetic shaft **602** includes a pair of matching elongated rotor members **604** and **606**, formed of magnetic material, mounted at spaced locations on the shaft and being set at an angle to each other, shaft **602** passing between spaced permanent magnets **582** and **584**. Two end cap members **608** and **610**, made from non-magnetic material, are attached to the ends of stator pole pieces **586** and **588** and are configured for receiving shaft **602** and respective bearings **612** and **614**.



The ends of the stator pole pieces **506** and **508** are configured for a given desired coupling relationship with rotor members **604** and **606**. For example, as shown in the exemplary end views of **Fig.25A** and **Fig.25B**, with end cap **608** removed, the end of stator pole piece **586** may include an curved portion **616** which is configured to create a variable-reluctance air gap **618** with elongated rotor member **604**. The end of stator pole piece **588** includes an curved portion **620** which is also configured to create a variable-reluctance air gap **622** with rotor member **604**.

In particular, portion **618** includes a circumferential curvature which has a centre point offset below the axis of rotation of shaft **602** and rotor member **604** as indicated by circle **624** shown in shadow. Similarly, portion **620** includes a circumferential radius of curvature which has a centre point offset above the axis of rotation of shaft **602** and rotor member **604**. When magnetic flux is passing along the path which includes a given end of the assembly, maximum coupling between the rotor member and stator pole pieces occurs when the rotor is positioned as shown in **Fig.25B**. Accordingly, the illustrated rotor member and stator pole piece configurations of themselves do not provide any skewing to the direction of rotation of the rotor assembly.





In this regard, various configurations for the rotor and ends of the stator pole piece are shown in the end views of **Figs.26-28**, which configurations provide skewing the direction of rotation. In particular, in device **620** of **Fig.26** a

rotor member **622** having notches **624** and **626**, which notches provide for greater magnetic coupling with the stator pole pieces **628** and **630** at corners **632** and **634** such that rotation is skewed in the clockwise direction. If notches were instead located at corners **632** and **634**, skewed rotation in the counterclockwise direction would be the result. In device **620** such counterclockwise rotation could also be achieved by removing rotor **622** from shaft **636**, flipping it end to end, and replacing it on shaft **636**.

In the device **640** of **Fig.27**, a portion **642** of the curved end portion of stator pole piece **644** is removed and a portion **646** of the curved end portion of stator pole piece **648** is removed. This configuration results in greater magnetic coupling between rotor member **650** and stator pole piece **644** at corner **652**, and greater magnetic coupling between rotor member **650** and stator pole piece **648** at corner **654**, such that rotation is skewed in the counterclockwise direction. Clockwise rotation could be achieved by instead modifying the opposite side of stator pole pieces **644** and **648**.

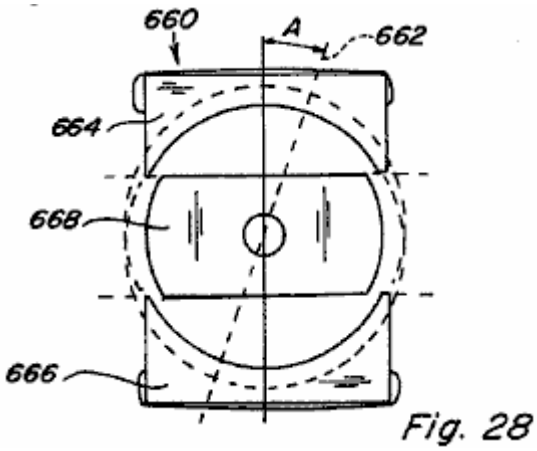


Fig. 28

Fig.28 depicts an end view of a device **660** in which the axis **662** of the curved end portion of upper stator pole piece **664** and lower stator pole piece **666** is placed at an angle **A** as shown. This configuration creates an unequal variable-reluctance air gap where opposite corners of rotor member **668** are closer to stator pole pieces **664** and **666**. Further, the angle at which maximum magnetic coupling between rotor member **668** and stator pole pieces **664** and **666** occurs is retarded by angle **A**. Rotation would be in the counterclockwise direction for the illustrated configuration.

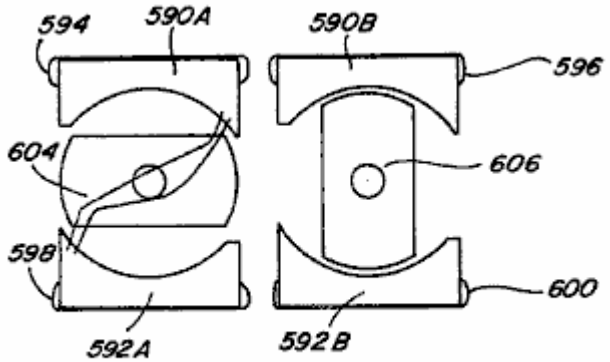


Fig. 29A

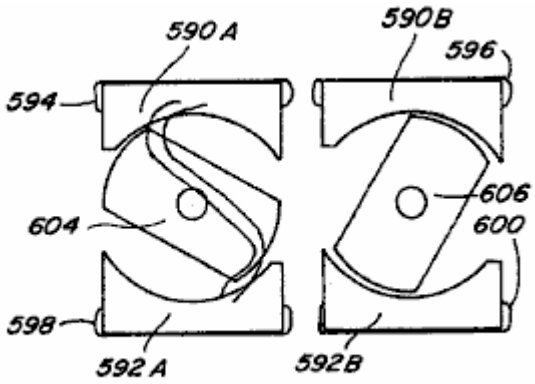


Fig. 29B

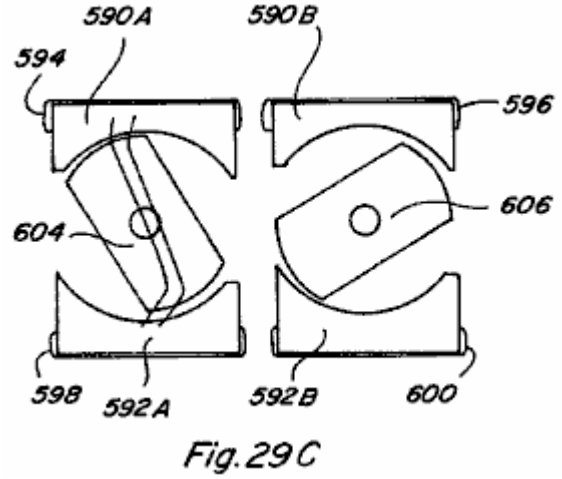


Fig. 29C

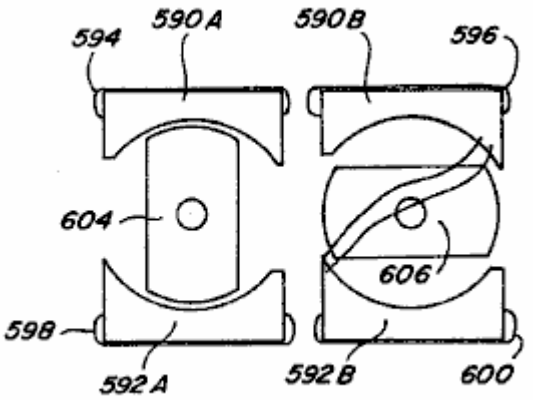
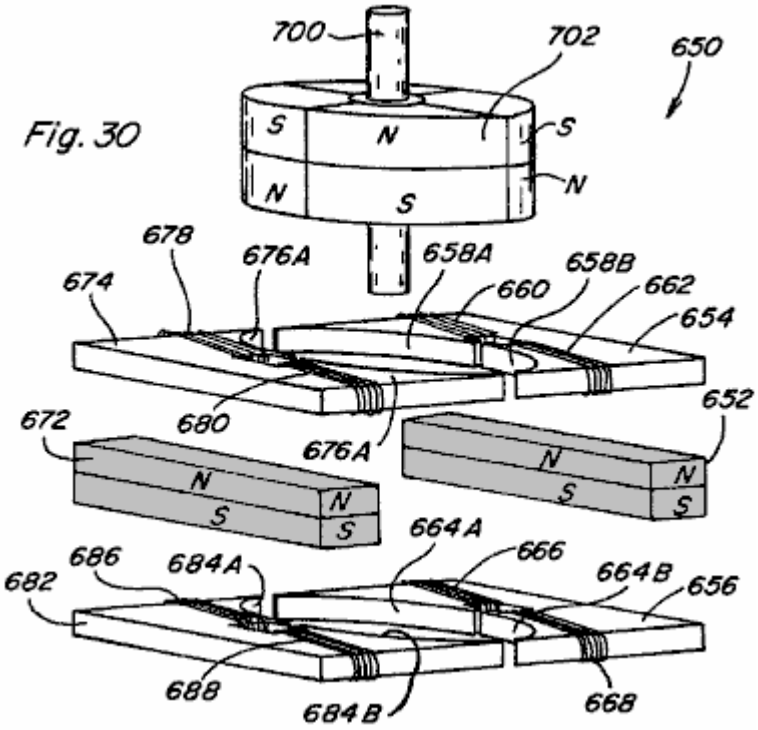


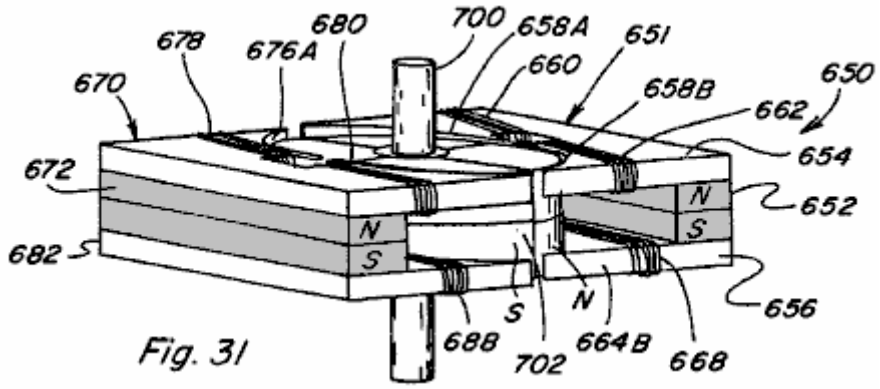
Fig. 29 D

Referring again to motor **580** of **Figs.23-25**, rotary motion of such device is depicted in the end views of **Figs.29A-29D**. In each end view the end cap has been removed to show rotation of the rotor members and in each of **Figs.29A-29D** an end view depicting rotor member **604** and an end view depicting rotor member **606** are shown side-by-side. In **Fig.29A**, rotor member **604** is defined as being at zero degrees and rotor member **606** is defined as being at ninety degrees. Control coils **594**, **598** are energised in a permanent magnet magnetic flux aiding manner such that no magnetic flux passes through stator pole piece path portions **590B** and **592B**. This allows rotor member **606** to move out of its ninety degree position and the magnetic coupling between rotor member **604** and stator pole piece path portions **590A** and **592A** will cause rotation to the position shown in **Fig.29B** and then **Fig.29C**. When rotor member **604** reaches the ninety degree position shown in **Fig.29D** control coils **594**, **598** are de-energised and control coils **596**, **600** are energised in a permanent magnet magnetic flux aiding manner causing rotation to continue due to the magnetic coupling between rotor member **606** and stator pole piece path portions **590B** and **592B**. Thus, by alternately energising the control coils of each path with every ninety degree rotation of rotor members **604** and **606**, continuous rotary motion is achieved.

The initial direction of rotation can be controlled by the circuit means used to energise control coils **594**, **598** and **596**, **600**, which circuit means includes circuitry for detecting the angular position of the rotor members. In particular, if rotor members **604** and **606** are at rest in the position shown in **Fig.29A**, and coils **594**, **598** are energised in an aiding manner, rotation may be clockwise or counterclockwise. If the desired direction is clockwise but upon energisation of coils **594**, **598** the rotor members begin to move counterclockwise, the detection circuitry will immediately de-energise coils **594**, **598** and energise coils **596**, **600** so that the clockwise direction is achieved.

Further, bypasses around permanent magnets **582** and **584** could be provided in rotary motion device **580**, such as those shown in **Fig.12**, and rotor members **604** and **606** could be formed by permanent magnets so as to take advantage of energising the control coils in an exceeding manner.





A third embodiment of a rotary motion device or motor **650** is shown in the exploded partial perspective view of **Fig.30** and in the assembled partial perspective view of **Fig.31**. In motor **650** the stator assembly includes a control component **651** including a permanent magnet **652** having a stator pole piece **654** positioned adjacent to one pole face of the magnet and a stator pole piece **656** positioned adjacent to the opposite pole face. Stator pole piece **654** includes a path portion **658A** extending to one side of permanent magnet **652** and a path portion **658B** extending to the one side thereof and spaced from first path portion **658A**. Control coils **660** and **662** are positioned along respective stator pole piece path portions **658A** and **658B**.

In the same way, stator pole piece **656** includes path portions **664A** and **664B** which extend in a similar manner from it so as to be aligned with stator path portions **658A** and **658B** respectively. Control coils **666** and **668** are positioned along respective stator pole piece path portions **664A** and **664B**. Positioned opposite, and facing control component **651**, is a similar control component **670** including permanent magnet **672** stator pole piece **674** with path portions **676A** and **676B** having the control coils **678** and **680**, and stator pole piece **682** with path portions **684A** and **684B** having their control coils **686** and **688**. The end of each of the pole piece path portions **658A, 658B, 664A, 664B, 676A, 676B, 684A**, and **684B** is of a generally curved configuration.

A rotor assembly of motor **650** includes a non-magnetic shaft **700** having a permanent magnet rotor member **702** mounted on it and which rotates with it. Permanent magnet rotor member **702** is generally ring-shaped and segmented to include distinct north and south pole faces which reverse about every ninety degrees around them. When assembled, the top and bottom surfaces of permanent magnet rotor member **702** align with pole pieces **654, 656, 674**, and **682** of the stator assembly and are preferably configured so that there is a minimal gap between the outer surface of permanent magnet rotor member **702** and the curved surfaces of the pole piece path portions.

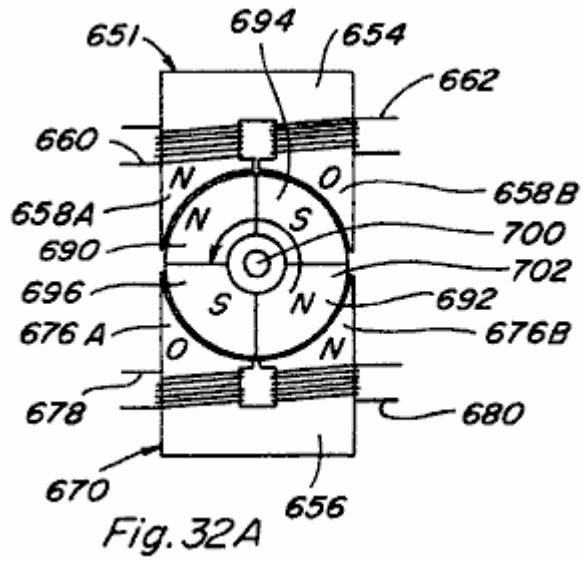


Fig. 32A

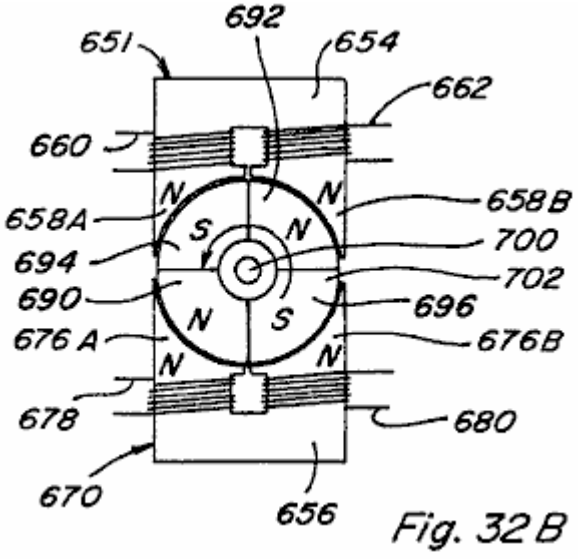


Fig. 32B

Rotation of device **650** can be achieved by controlled, timed energising and de-energising of control coils **660, 662, 666, 668, 678, 680, 686**, and **688**. Exemplary rotation is demonstrated with reference to the top views of **Figs.32A-32B** which depict counterclockwise rotation of permanent magnet rotor member **702** through one-hundred and eighty degrees. In **Fig.32A** stator pole piece path portion **658A** of component **651** is active and stator pole piece path portion **658B** is not active, which may be achieved by energising control coil **660** in a permanent magnet magnetic flux aiding manner or by energising control coil **662** in a permanent magnet magnetic flux opposing manner. Stator pole piece path portion **676B** of component **670** is active and stator pole piece path portion **676A** is not active, which may be achieved by energising control coil **680** in a permanent magnet magnetic flux aiding manner or by energising control coil **678** in a permanent magnet magnetic flux opposing manner.

Thus, portions **690** and **692** of permanent magnet rotor member **702**, which both have a north magnetic polarity, will be repelled by the north polarity of stator pole piece path portions **658A** and **676B** aligned with it. Portions **694** and **696** of permanent magnet rotor member **702**, both of which have a south magnetic polarity, will be attracted to the active path portions **658A** and **676B**. At the instant that rotor member portion **694** becomes aligned with stator pole piece path portion **658A**, as shown in **Fig.32B**, all coils are de-energised such that all pole piece path portions will be active as shown. Pole piece path portions **658B** and **676A** are then kept active while pole piece path portions **658A** and **676B** are made inactive. This is achieved by energising control coils **662** and **678** in a permanent magnet magnetic flux aiding manner or by energising control coils **660** and **680** in a permanent magnet magnetic flux opposing manner. Rotor member portions **690** and **692** will again be repelled by the north polarity of path portions **658B** and **676A** aligned with it so that rotation of permanent magnet rotor **702** is continued.

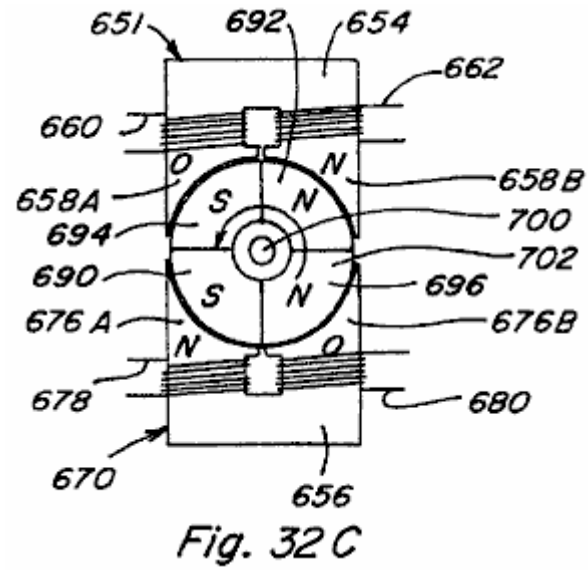


Fig. 32C

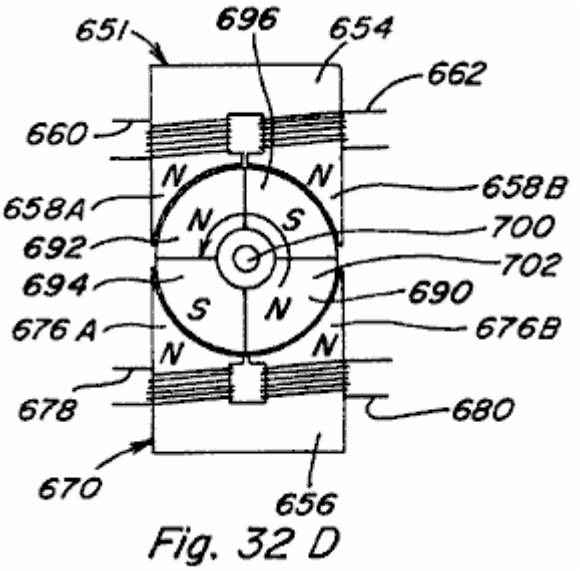


Fig. 32D

In **Fig.32D** all coils are shown de-energised when rotor portion **692** aligns with pole piece path portion **658A**. By continuing this timed sequence of energisation and de-energisation of the control coils, continued rotary movement is achieved. As explained above, the initial direction of rotation can be controlled by circuit means

which detects the initial direction of permanent magnet rotor **702** and immediately alters the coil energisation scheme if the initial direction is incorrect.

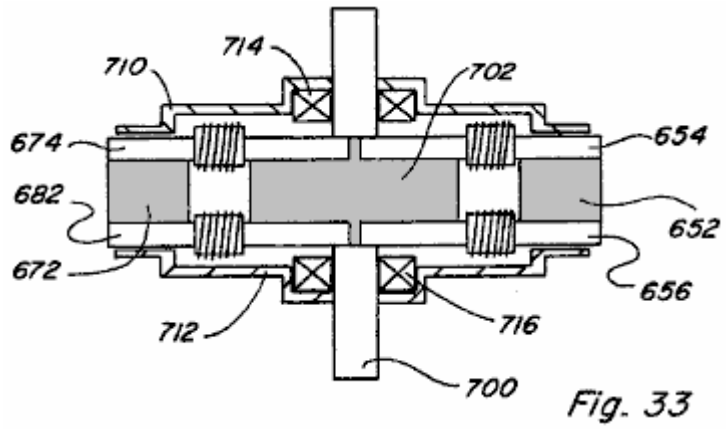


Fig. 33

A side view of assembled motor **650** is shown in **Fig.33** and includes an upper housing or enclosure portion **710**, a bottom housing portion **712**, upper bearing **714**, and a lower bearing **716**.

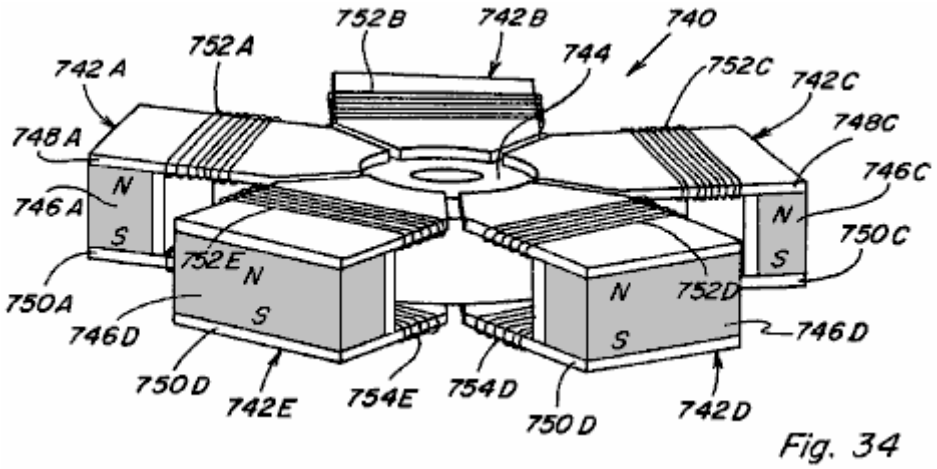


Fig. 34

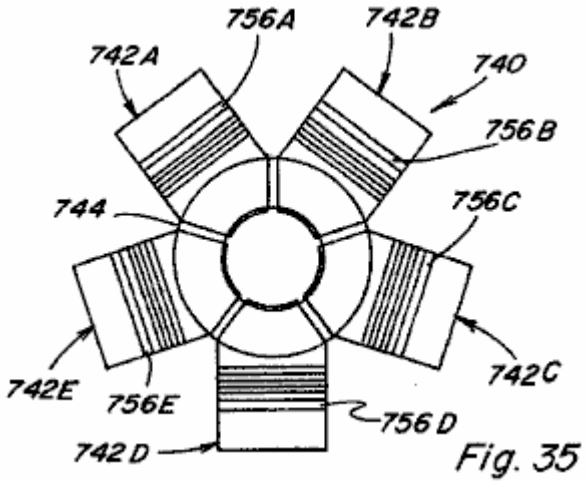


Fig. 35

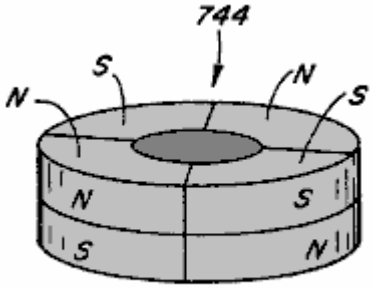


Fig. 36

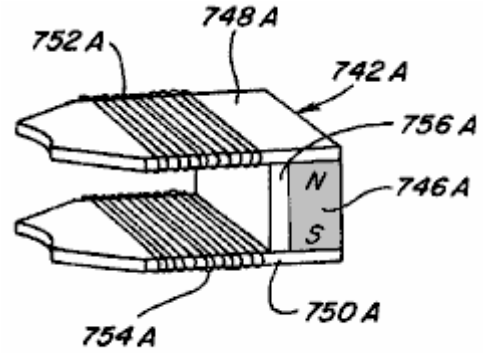


Fig. 37

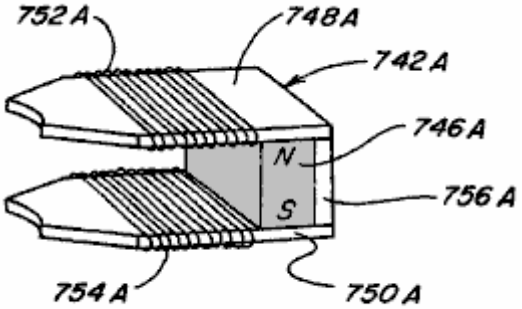


Fig. 38

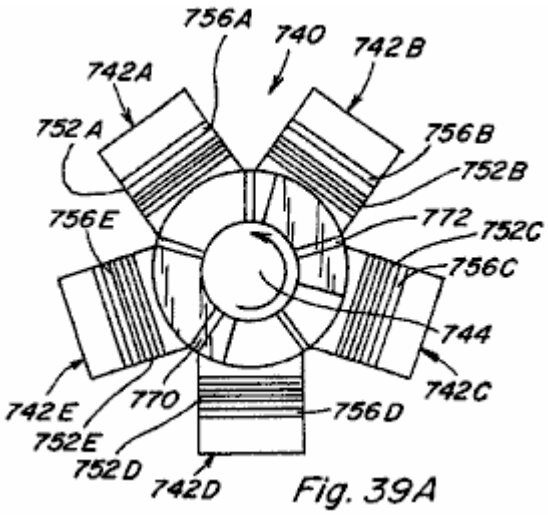


Fig. 39A

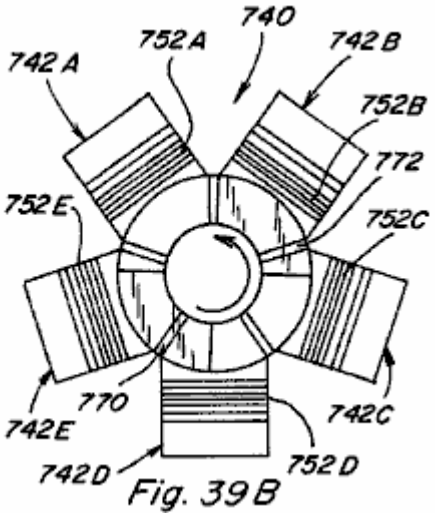


Fig. 39B

A fourth embodiment of a rotary motion device or motor **740** is illustrated in **Figs.34-39**. Motor **740** includes five stator control components **742A-742E** positioned around a ring shaped permanent magnet rotor member **744** (**Fig.36**). As shown with reference to component **742A** in **Fig.37**, each stator component **742A** includes a permanent magnet **746A** with an upper pole piece **748A** positioned adjacent to one pole face and a lower pole piece **750A** positioned adjacent to the opposite pole face. Control coils **752A**, **754A** are positioned along respective pole pieces **748A**, **750A**. A bypass **756A** extends from pole piece **748A** to pole piece **750A** and is positioned between permanent magnet **746A** and control coils **752A**, **754A**. Alternatively, bypass **756A** could be provided on the opposite side of permanent magnet **746A** as shown in **Fig.38**. Although not shown, it is anticipated that permanent magnet rotor member **744** would be mounted on an axis for rotation with it and that a motor housing or enclosure could be provided, such as shown in relation to motor **650** of **Fig.33**.

Referring to the top views of **Figs.39A-39D**, rotary motion of rotor member **744** is depicted by the sequence of views. Regions **770** and **772** in **Figs.39A-39D** represent the magnetic north regions of the top of permanent magnet rotor **744**. In **Fig.39A** control coils **752E** and **752C** are energised in a permanent magnet aiding and exceeding manner such that regions **770** and **772** of permanent magnet rotor **744** are repulsed by components **742E** and **742C** while permanent magnet motor regions **774** and **776** are attracted by components **742E** and **742C**. The resultant coupling forces act to move permanent magnet rotor in a counterclockwise direction to the location shown in **Fig.39B**. Just after permanent magnet rotor region **772** passes the point shown in **Fig.39C**, control coil **752B** is energised in a permanent magnet aiding and exceeding manner, while control coils **752E** and **752C** also remain energised, and counterclockwise rotation of permanent magnet rotor **744** is continued. Just

after permanent magnet rotor region **772** passes by control component **742C** control coil **752C** is de-energised, while control coils **752E** and **752B** remain energised, so as to continue counterclockwise rotation. Then, just after permanent magnet rotor region **770** reaches the location shown in **Fig.39D** control coil **752D** is energised in a permanent magnet flux aiding and exceeding manner, while coils **752E** and **752B** remain energised, so as to continue counterclockwise rotation. Thus, as in the other embodiments, repeated and timed energisation and de-energisation of the control coils produces the desired rotational movement.

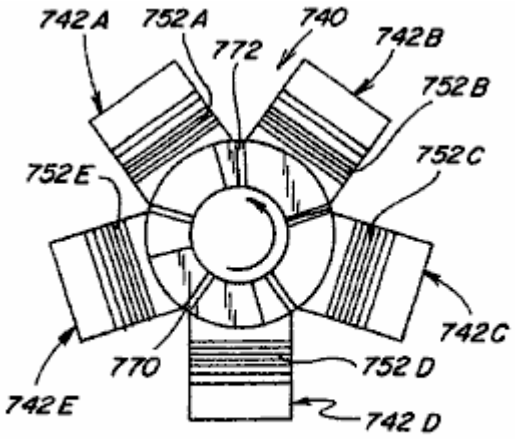


Fig. 39 C

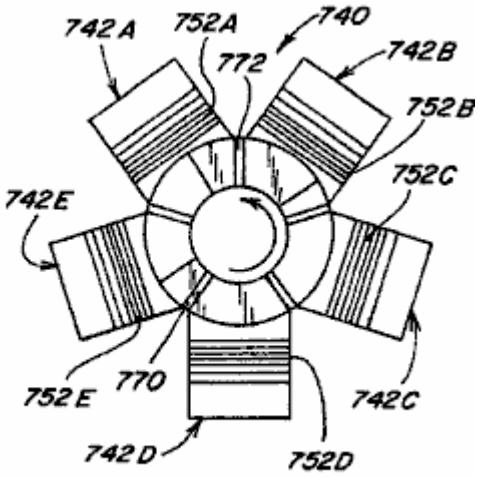


Fig. 39 D

In terms of controlling the energisation of coils in the devices described above, various electronic control circuit/switching means and electromechanical control circuit/switching machines are depicted in **Figs.40-44**. In circuit **800** of **Fig.40** a given coil **802** is placed in series between an electrical energy source **804** and a power MOSFET **806**. An LED **808** is connected to electrical energy source **804** through resistor **810** and is positioned to impinge upon a phototransistor **812** which is connected in series with resistor **814**. A control input of MOSFET **806** is connected between phototransistor **812** and resistor. Accordingly, when LED **808** activates phototransistor **812** the voltage drop across resistor **814** activates, or turns ON, MOSFET **806** and coil **802** is energised. Timed energisation of coil **802** is provided by mounting an interrupter **816**, such as shown in **Fig.42**, to the shaft **816** of the motor device to be controlled, such that as interrupter **814** rotates with shaft **816** coil **802** is alternately energised and de-energised. In a device with a plurality of coils a corresponding plurality of LED/photoresistor pairs may be provided.

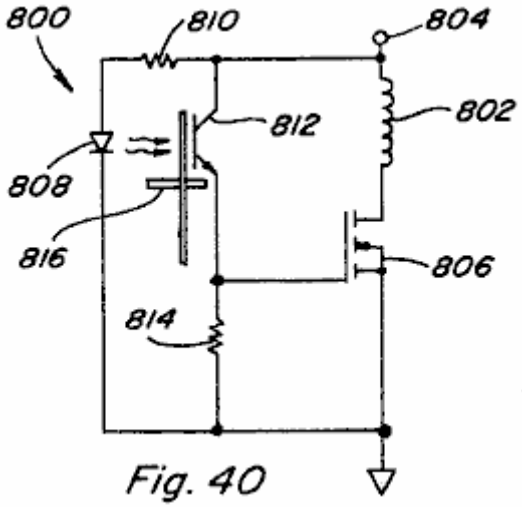


Fig. 40

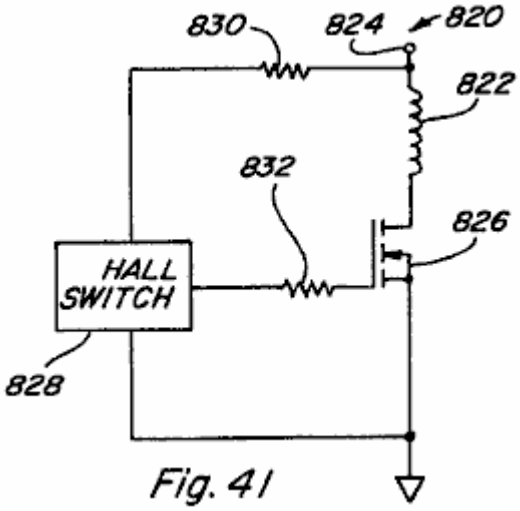


Fig. 41

In circuit **820** of **Fig.41** a coil **822** is positioned between electrical energy source **824** and power MOSFET **826**. A hall switch **828** is connected in series with resistor **830**. Hall switch **828** is also connected to the control input of MOSFET **826** through resistor **832**. In a given device hall switch **828** would be positioned to react to a change in magnetic flux so as to control the ON/OFF switching of MOSFET **826**, and thus the alternate energisation and de-energisation of coil **822**.

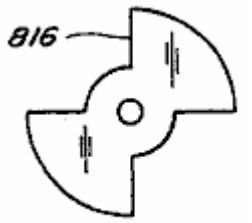


Fig. 42

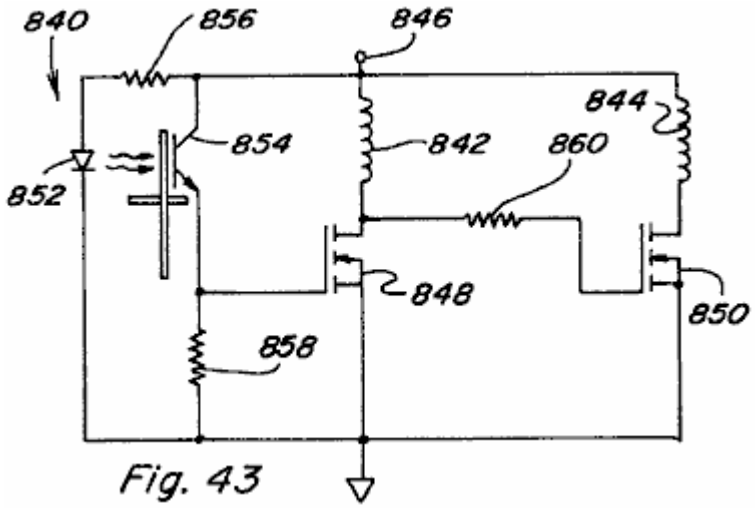


Fig. 43

In **Fig.43** a circuit **840** for controlling two coils in an opposite manner is provided such that when coil **842** is energised coil **844** is de-energised, and such that when coil **842** is de-energised coil **844** is energised. Both coils **842** and **844** are connected in series between electrical energy source **846** and respective power MOSFETs **848** and **850**. An LED **852** and phototransistor **854** arrangement is provided, LED connected in series with resistor **856** and phototransistor connected in series with resistor **858**. When LED **852** turns phototransistor **854** ON the voltage drop across resistor **858** turns MOSFET **848** ON and coil **842** is energised. At that time the voltage applied at the control input of MOSFET **850** will be low and therefore MOSFET **850** will be OFF and coil **844** will be de-energised. When interrupter **814** blocks LED 852, phototransistor **854** is turned OFF and MOSFET **848** is likewise turned OFF. The control input of MOSFET **850** is therefore pulled high through resistor **860** and MOSFET **850** is turned ON such that coil **844** is energised.

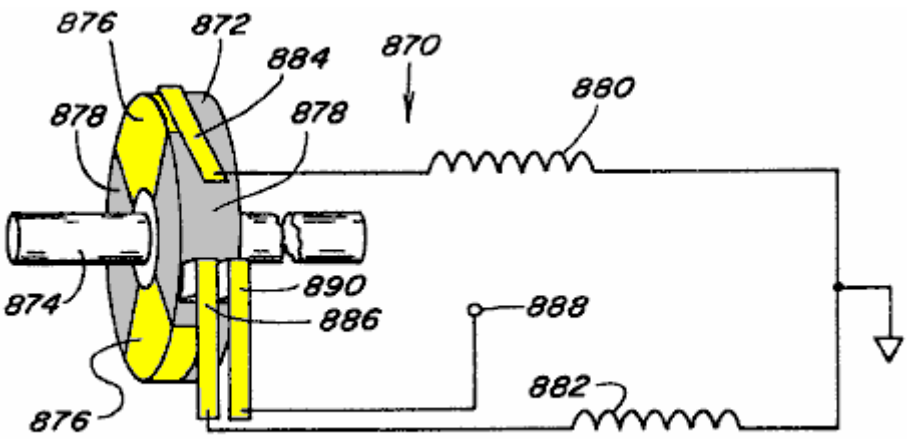


Fig. 44

In **Fig.44** a system **870** including member **872** mounted on rotating shaft **874** is provided, with the left side of member **872** being alternately conductive at **876** and non-conductive at **878**. Coils **880** and **882** are connected to respective brushes **884** and **886** which are positioned to contact member **872** during each rotation of the shaft. Member **872** is connected through brush **890** to power supply **888**. Thus, coils **880** and **882** will alternately be energised and de-energised as the respective brushes thereof contact the conductive and non-conductive portions of member **872**.

Any of such circuit means, variations thereof, or other circuit means may be used to provide the timed energisation of the control coils in the various embodiments of the present invention.

From the preceding description of the illustrated embodiments, it is evident that the objects of the invention are attained. Although the invention has been described and illustrated in detail, it is to be clearly understood that the same is intended by way of illustration and example only and is not to be taken by way of limitation.

For example, although the magnetic flux control techniques of the present invention have been discussed as applicable mainly to various motive applications, such magnetic flux control techniques are also useful in static applications.

Power Conversion

Referring to **Figs.45A-45C** there is shown the permanent magnet device **900** of **Figs.45A-45C** which has two magnetic flux paths provided by rectangular pole piece **902** which includes upper portion **904** and lower portion **906** each positioned against a respective pole face of permanent magnet **910**. Unlike the device of **Figs.7-9**, fall away armatures are not provided. Instead, fixed armatures in the form of integral pole piece portions **912** and **914** extend from upper portion **904** to lower portion **906** completing the two flux paths in a permanent manner. Control coils **916**, **918** are provided along one flux path and control coils **920**, **922** are provided along the other flux path, such control coils acting as primary windings in device **900**. One coil **924** is positioned around pole piece portion **912** and another coil **926** is positioned around pole piece portion **914**, such coils **924**, **926** acting as secondary windings in device **900**.

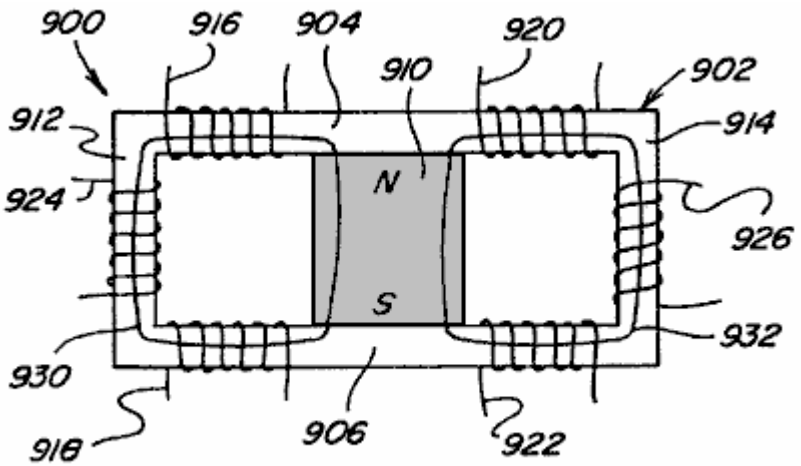


Fig. 45A

In **Fig.45A** no coils are energised and the permanent magnet magnetic flux splits evenly between paths **930** and **932**, coupling with both coil **924** and coil **926**.

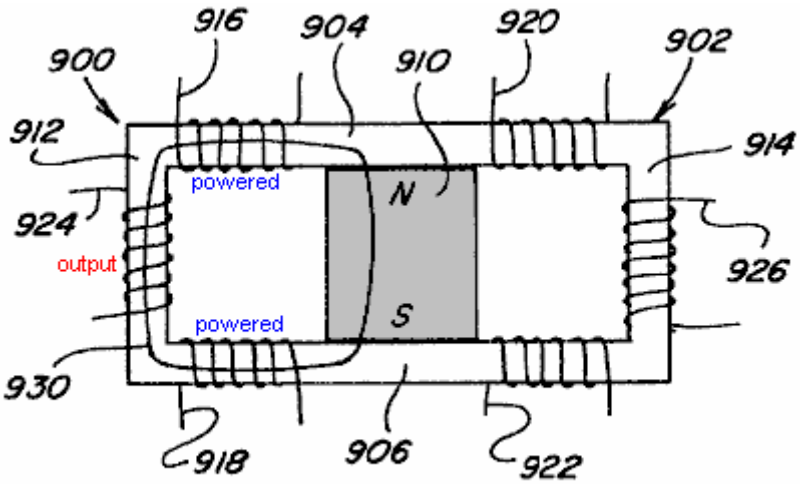


Fig. 45B

In **Fig.45B** coils **916**, **918** are energised in a permanent magnet magnetic flux aiding manner so as to couple with all the magnetic flux of permanent magnet **910**. All magnetic flux flows along path **930** as shown and thus couples with coil **924**.

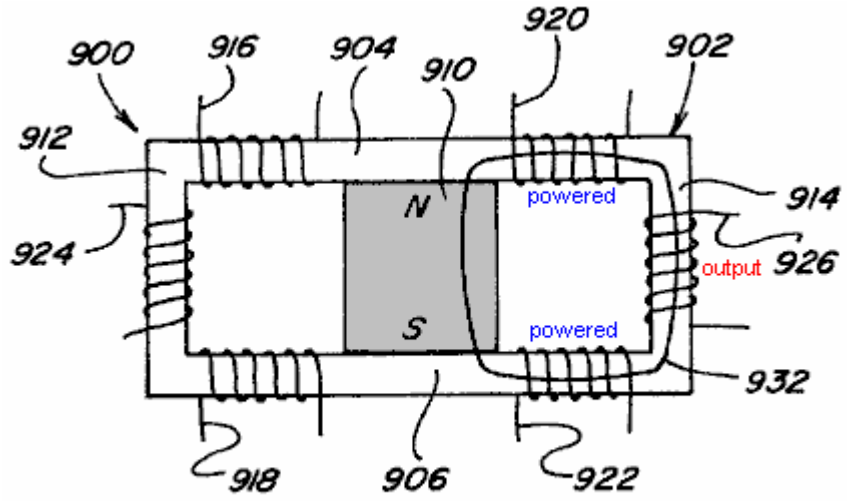


Fig. 45 C

In **Fig.45C** coils **920**, **922** are energised in a permanent magnet magnetic flux aiding manner such that all magnetic flux traverses path **932** and couples with coil **926**. By continuously alternately energising and de-energising coils **916**, **918** and **920**, **922** in such a manner energy conversion is achieved due to the coupling with coils **924** and **926**. The magnetic flux in the integral pole piece portions **912** and **914**, and thus the flux coupling with respective coils **924** and **926**, varies by a factor of twice the amount of magnetic flux generated by energising coils **916**, **918** and **920**, **922**.

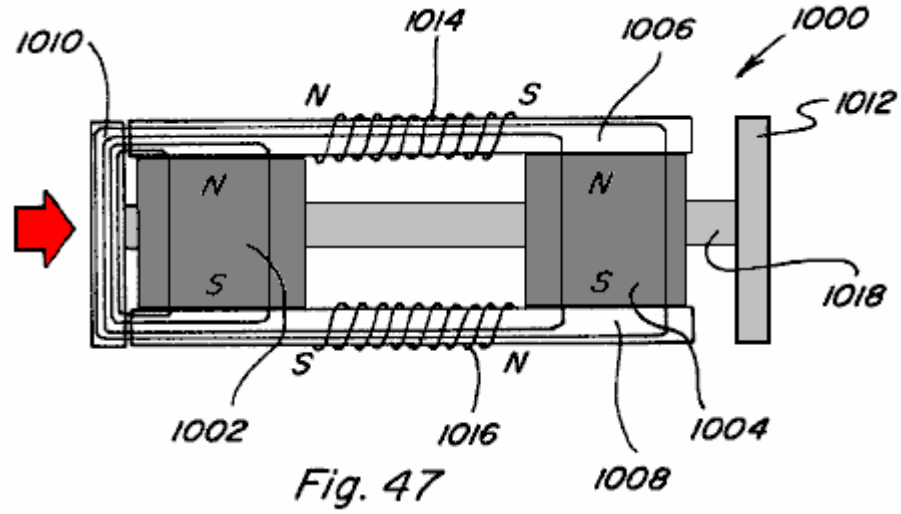


Fig. 47

The construction shown in **Fig.45A** and **Fig.45X** are similar to the construction shown in **Fig.7** and **Fig.47**. The difference in both cases relates to replacing the two flux paths and armatures with one continues flux path. The arrangement in **Fig.7** has one permanent magnet and four coils and the arrangement in **Fig.47** has two permanent magnets and two coils. Although the physical aspects of the two arrangements and the details of the flux control vary, the control method for varying the permanent magnets flux are similar and will be described simultaneously and only differences will be pointed out.

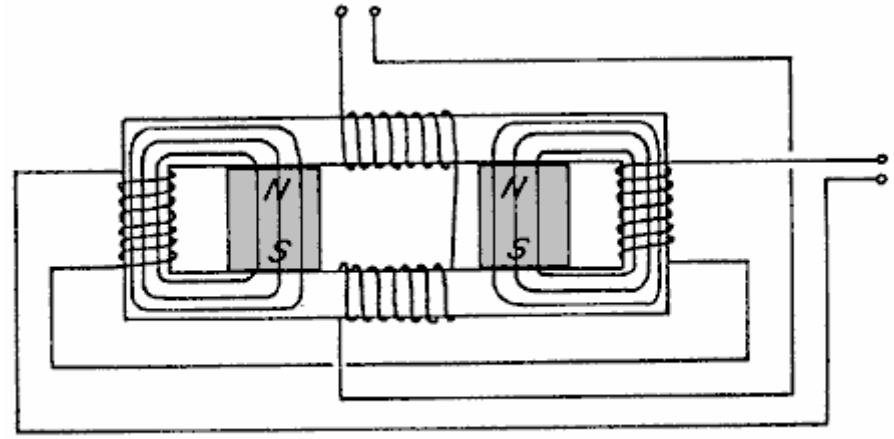


Fig. 45 X

With continuous flux paths the static flux from the permanent magnet or magnets is useless. However, if the static flux of the permanent magnet confined to the flux paths were modified to be time varying it would have utility for electromagnetic induction devices for power conversion like transformers and power inverters. However, the same basic method for controlling the flux of a permanent magnet to provide linear and rotary motion can also be applied to time varying the static flux from the permanent magnet. The construction shown in **Fig.45X** utilises four control coils and a single permanent magnet while the construction shown in **Fig.45A** uses two control coils and two permanent magnets. The flux that would normally be supplied by a primary winding is supplied by the static flux of the permanent magnet or magnets and the control coils convert this static flux into a time varying flux in a novel way. Both arrangements use two secondary coils, the secondary coils are placed in the region of the continuous flux path that would be occupied by an armature or rotor in the linear or rotary arrangements. The regions of the flux paths that perform work are the same in all cases.

In all cases the control coils can either be wired in series or parallel and the secondary coils can be either wound in series or parallel. More than one secondary coil or secondary coils with multiple taps can be placed in the working regions and further multiple flux paths can be utilised with one or more secondary coils placed in each of the working regions. This is made obvious by the disclosures of the linear and rotary devices herein and based on the fact that the working regions of the flux paths are identical.

Fig.45X and **Fig.45A** also show the paths of the static flux of the permanent magnet or magnets when no current is flowing in the control coils. In the arrangement shown in **Fig.45X** the flux from the single permanent magnet divides between the two working areas of the flux path. In the arrangement of **Fig.45A** all of the flux of one of the permanent magnets passes through one of the working regions and all of the flux of the second permanent magnet passes through the other working region. Each of the working regions in both cases are occupied by secondary coils.

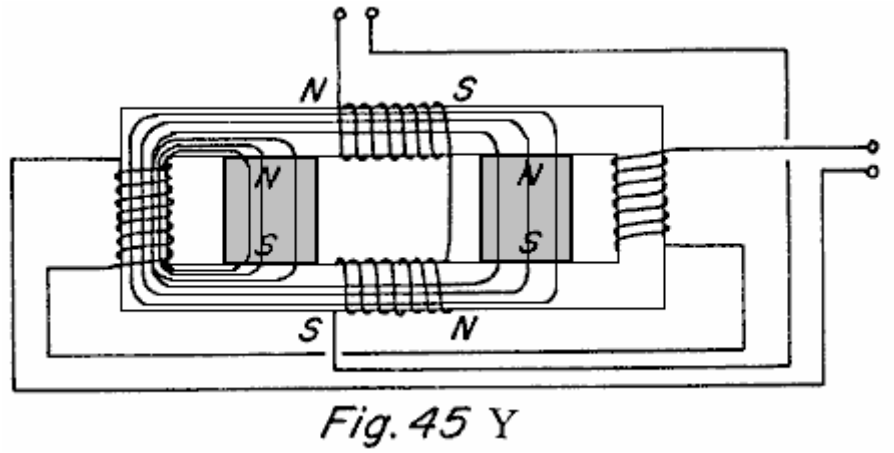


Fig. 45 Y

Fig.45Y and **Fig.45B** show the control coils energised with the polarity shown with respect to the polarity of the permanent magnet or magnets included. In **Fig.45Y** the opposing coil, blocks the passage of flux from the permanent magnet, and the aiding coil couples with the flux of the permanent magnet and therefore all of the flux of the permanent magnet passes through one working region as shown. In **Fig.45B** the opposing side of the coil blocks the passage of flux from the permanent magnet on the opposing side of the coil and the aiding side of the

coil couples with the flux of the other permanent magnet and therefore all of the flux of both the permanent magnets passes through the working region as shown.

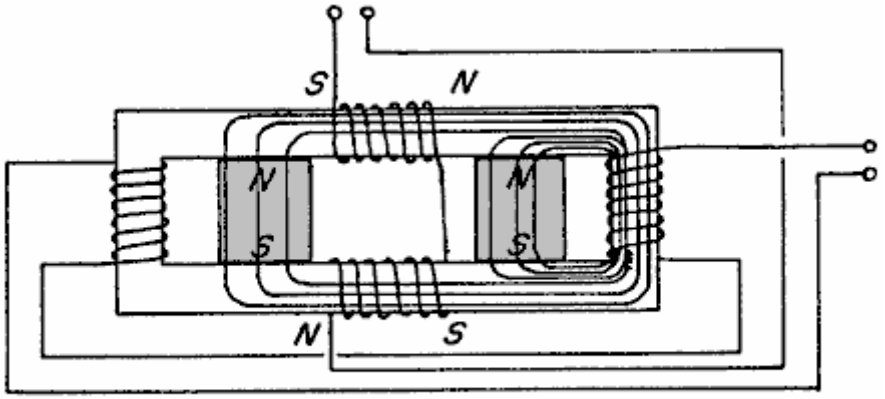


Fig. 45 Z

Fig.45Z and **Fig.45C** show the control coils energised with a polarity opposite of that shown in **Fig.45Y** and **Fig.45B**. The same action occurs and results in all of the permanent magnet or magnets path flux passing through the opposite working regions.

By alternating the polarity of the control coils during one cycle, one working region experiences an increasing flux and the opposite region experiences a decreasing flux and during the next cycle the opposite occurs. This results in the induction of a voltage in the secondary coils that is decided by the magnitude of the change in flux in the working region and the time in which this change occurs. The novelty of this discovery is that the primary flux inducing the voltage in the secondary coils is supplied by the permanent magnet or magnets and is far greater than the flux supplied by the control coils.

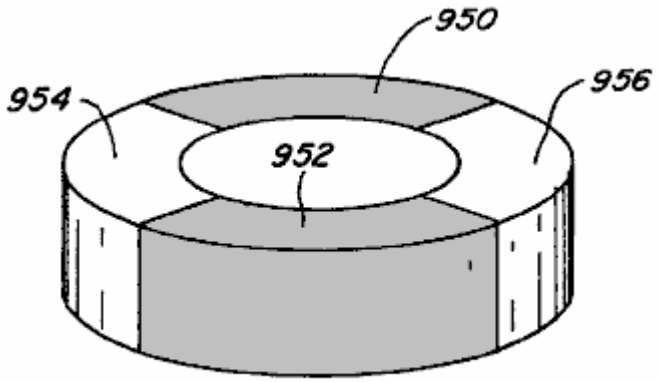


Fig. 46

Further, in the rotary motion devices of **Fig.31** and **Fig.34**, it is not necessary that respective rotor members **702** and **744** be formed of permanent magnets. Each could take the form shown in **Fig.46** where sections **950** and **952** are formed of magnetic material such as soft iron and sections **954** and **956** are formed by a non-magnetic filler material.

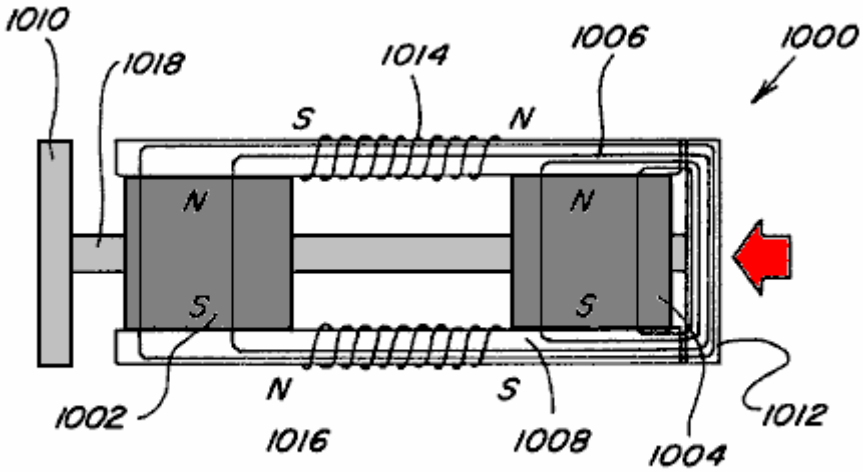


Fig. 48

Fig.47 and **Fig.48** show another embodiment **1000** of the subject device. The embodiment **1000** includes two spaced permanent magnets **1002** and **1004** each of which has its north pole adjacent to the upper surface and its south pole adjacent to the lower surface. A magnetisable bridging member **1006** extends across and makes contact with the north magnetic poles of the magnets **1002** and **1004** and another magnetisable bridging member **1008** makes contact with the south magnetic poles of the two permanent magnets **1002** and **1004**.

The members **1006** and **1008** extend slightly beyond the opposite sides of the respective permanent magnets **1002** and **1004** and a pair of spaced armature members **1010** and **1012** are positioned to move into and out of engagement with the ends of the members **1006** and **1008**. Coils **1014** and **1016** are mounted respectively on the members **1006** and **1008** in the space between the permanent magnets **1002** and **1004**, and the armatures **1010** and **1012** are shown connected together by a rod **1018** which enables them to move backwards and forwards into engagement with the respective members **1006** and **1008** when different voltages are applied to the respective coils **1014** and **1016**.

In **Fig.47**, the coils **1014** and **1016** are energised as shown with the coil **1014** having its north magnetic end to the left and its south magnetic end to the right and the opposite is true of the coil **1016**. In **Fig.48**, the voltage applied to the respective coils **1014** and **1016** is reversed so that the polarity of the left end of coil **1014** is south and the polarity of the opposite end of the same coil **1014** is a north magnetic pole. The reverse is true of the coil **1016**. In **Fig.47** and **Fig.48** it should be noted that the relationship of aiding and opposing is indicated on the figures to indicate the relationship when the coils are energised. For example, in **Fig.47** when the coils are energised as shown the relationship is opposing for the permanent magnet **1002** and is aiding with respect to the permanent magnet **1004**. The reverse is true when the voltage on the coils is reversed as shown in **Fig.48**. The movement of the armature is therefore controlled by the proper timing of the voltage on these coils. The same principles can be applied to produce rotating movement as shown in **Fig.42**.

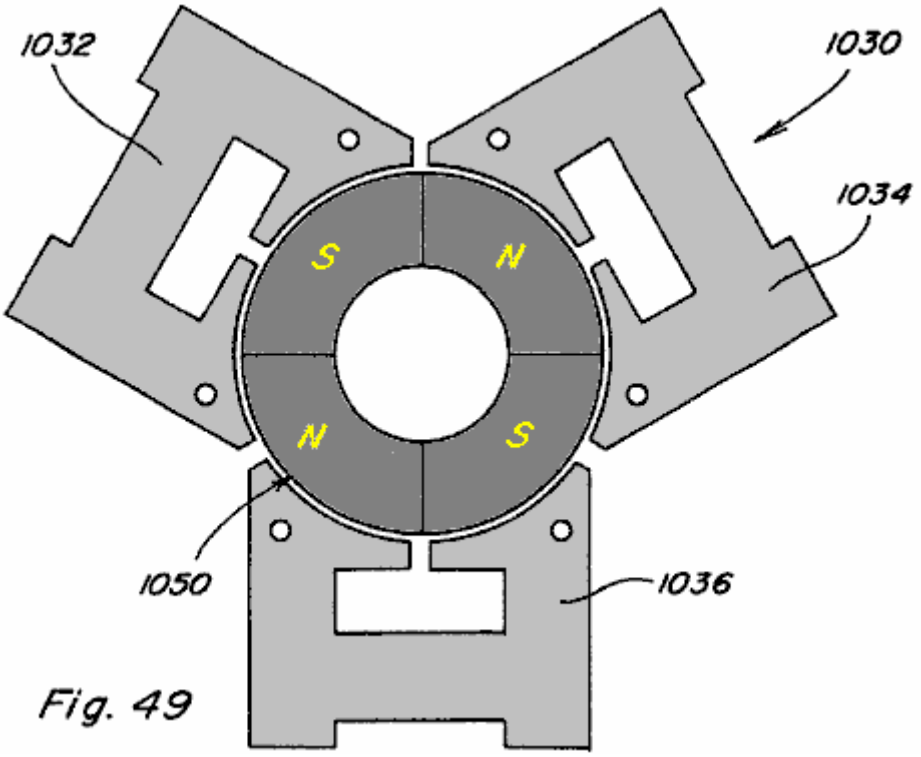


Fig. 49

Fig.49 shows another embodiment **1030** of the subject invention using principles similar to those described in connection with **Fig.47** and **Fig.48**. The embodiment **1030** includes a plurality, three being shown, of stationary members **1032**, **1034** and **1036**.

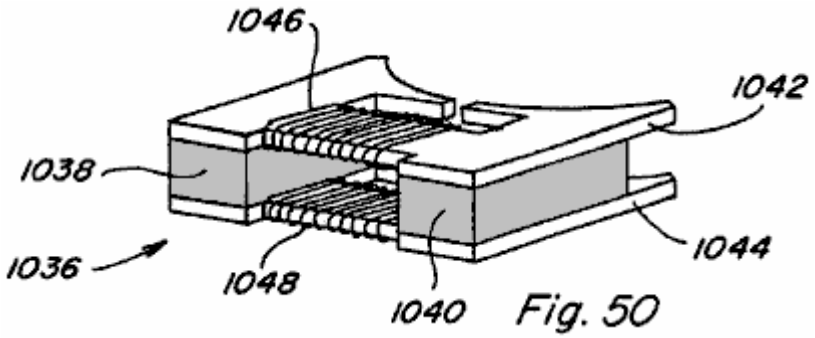


Fig. 50

The details of these members are better shown in **Fig.50** which shows the details of the member **1036**. This member includes a pair of permanent magnets **1038** and **1040**, each of which has magnetisable members mounted adjacent to it's opposite sides, as in the previous construction. The members **1042** and **1044** also have coils **1046** and **1048**, respectively, and the coils are energised as described in connection with **Fig.47** and **Fig.48** to produce aiding and opposing magnetism. The construction shown in **Fig.49** may have three stator portions as shown or it may have more stator portions as desired. The rotor **1050** is positioned in the space between the members **1032**, **1034** and **1036** and includes a permanent magnet portion part of which has its north magnetic pole on the surface as shown and the other parts has its south magnetic pole in the same surface as shown. The permanent magnets **1038** and **1040** on the stators interact with the permanent magnets on the rotor to produce the rotating motion and is controlled by the energising of the coils.

Other applications and advantages of the devices and methods of the present invention exist and various modifications are possible, and therefore the present invention is not intended to be limited to the specific examples disclosed herein. Accordingly, the spirit and scope of the invention are to be limited only by the terms of the appended claims.

CLAIMS

1. A permanent magnet device, comprising a permanent magnet having north and south pole faces, a first pole piece, a second pole piece, a first control coil, a second control coil, and circuit means, the first pole piece positioned adjacent the north pole face of the permanent magnet and including a first path portion, a second path portion and a third portion, the first path portion extending beyond a perimeter of the north pole face in

one direction and the second path portion extending beyond the perimeter of the north pole face in another direction to define first and second flux paths for magnetic flux emanating from the north pole face of the permanent magnet, the first path portion of the first pole piece connected to the second path portion of the first pole piece by the third portion which extends across the north pole face of the permanent magnet, the second pole piece positioned adjacent the south pole face and including a first path portion and a second path portion, the first path portion extending beyond a perimeter of the south pole face and substantially aligned with the first path portion of the first pole piece, the second path portion extending beyond the perimeter of the south pole face and substantially aligned with the second path portion of the first pole piece, the first control coil positioned around the first path portion of the first pole piece, the second control coil positioned around the second path portion of the first pole piece, the circuit means connected to each of the first control coil and the second control coil to alternately energise the first coil and the second coil in a timed sequential manner.

2. The permanent magnet device as set forth in claim 1, wherein the first control coil and the second control coil are alternately energised in a permanent magnet magnetic flux aiding manner.

3. The permanent magnet device as set forth in claim 1, wherein the first control coil and the second control coil are alternately energised in a permanent magnet magnetic flux opposing manner.

4. The permanent magnet device as set forth in claim 1, further comprising a rotor member mounted on a shaft for rotation therewith, the rotor member sized, shaped, and positioned to extend substantially from the first path portion of the first pole piece to the first path portion of the second pole piece during at least some part of its rotation.

5. The permanent magnet device as set forth in claim 4, wherein the rotor member is formed by at least one permanent magnet.

6. The permanent magnet device as set forth in claim 1, wherein the second path portion of the first pole piece and the second path portion of the second pole piece are positioned alongside the first path portion of the first pole piece and the first path portion of the first pole piece.

7. The permanent magnet device as set forth in claim 1, further comprising a first bypass extending from the first path portion of the first pole piece to the first path portion of the second pole piece, one end of the first bypass positioned adjacent the first path portion of the first pole piece and between the permanent magnet and the first control coil.

8. The permanent magnet device as set forth in claim 6, further comprising a second bypass extending from the second path portion of the first pole piece to the second path portion of the second pole piece, one end of the second bypass positioned adjacent the second path portion of the first pole piece and between the permanent magnet and the second control coil.

9. The permanent magnet device as set forth in claim 1, further comprising a plurality of armatures arranged to define a path of movement, wherein the second path portion of the first pole piece and the second path portion of the second pole piece are positioned alongside the first path portion of the first pole piece and the first path portion of the second pole piece, and wherein all of such pole piece path portions include an end face positioned adjacent the path of movement defined by the plurality of armatures.

10. The permanent magnet device as set forth in claim 1, wherein the first control coil and the second control coil are simultaneously energised one in a permanent magnet magnetic flux aiding manner and one in a permanent magnet magnetic flux opposing manner.

11. The permanent magnet device as set forth in claim 1, further comprising two shaft connected armatures which can be positioned adjacent the ends of the first and second pole pieces, wherein each of the armatures is formed by a permanent magnet.

12. The permanent magnet device of claim 1 further comprising a first fixed armature extending between the first path portion of the first pole piece to the first path portion of the second pole piece and a second fixed armature extending between the second path portion on the first pole piece to the second path portion of the second pole piece.

13. The permanent magnet device of claim 12 where a first secondary coil is wrapped around the first fixed armature and a second secondary coil is wrapped around the second fixed armature.

14. The permanent magnet device of claim 13 including circuit means connected to the control coils to control the energising thereof to produce a varying flux in the armatures and to induce voltage in the secondary coils.

15. The permanent magnet device of claim 1 wherein there are at least two permanent magnets each having north and south pole faces, the first pole piece being positioned extending between the north pole faces of the permanent magnets and the second pole piece positioned extending between adjacent south pole faces of the permanent magnets.

16. A method for controlling the path of magnetic flux from a permanent magnet, the method comprising the steps of:

(**a**) placing a first pole piece adjacent a first pole face of the permanent magnet so as to have at least first and second path portions extending beyond a perimeter of the first pole face;

(**b**) placing a second pole piece adjacent a second pole face of the permanent magnet so as to include at least one portion which substantially aligns with the first and second path portions of the first pole piece;

(**c**) placing a first control coil along and around the first path portion of the first pole piece;

(**d**) placing a second control coil along and around the second path portion of the first pole piece;

(**e**) repeatedly energising the first control coil in a permanent magnet magnetic flux opposing manner so as to prevent magnetic flux of the permanent magnet from traversing the first path portion of the first pole piece; and

(**f**) repeatedly energising the second control coil in a permanent magnet magnetic flux opposing manner so as to prevent magnetic flux of the permanent magnet from traversing the second path portion of the first pole piece.

17. The method as set forth in claim 16 wherein the energisation of steps (e) and (t) take place in a simultaneous manner.

18. A method for controlling the path of magnetic flux from a permanent magnet, the method comprising the steps of:

(**a**) placing a first pole piece adjacent a first pole face of the permanent magnet so as to have at least first and second path portions extending beyond a perimeter of the first pole face;

(**b**) placing a second pole piece adjacent a second pole face of the permanent magnet so as to include at least one portion which substantially aligns with the first and second path portions of the first pole piece;

(**c**) placing a first control coil along and around the first path portion of the first pole piece;

(**d**) placing a second control coil along and around the second path portion of the first pole piece; and

(**e**) alternately performing the following steps in a repeated manner:

(**i**) energising the first control coil in a permanent magnet magnetic flux aiding manner so as to couple with substantially all magnetic flux of the permanent magnet such that substantially no magnetic flux of the permanent magnet traverses the second path portion of the first pole piece when the first control coil is so energised; and

(**ii**) energising the second control coil in a permanent magnet magnetic flux opposing manner so as to couple with substantially all magnetic flux of the permanent magnet such that substantially no magnetic flux of the permanent magnet traverses the first path portion of the first pole piece when the second control coil is so energised.

19. A method for controlling the path of magnetic flux from a permanent magnet the method comprising the steps of:

(**a**) placing a first pole piece adjacent a first pole face of the permanent magnet so as to have at least first and second path portions extending beyond a perimeter of the first pole face;

(**b**) placing a second pole piece adjacent a second pole face of the permanent magnet so as to include at least one portion which substantially aligns with the first and second path portions of the first pole piece;

(**c**) placing a first control coil along and around the first path portion of the first pole piece;

(**d**) placing a second control coil along and around the second path portion of the first pole piece; and

(**e**) alternately performing the following steps in a repeated manner:

(**i**) simultaneously energising the first control coil in a permanent magnet magnetic flux aiding manner and the second control coil in a permanent magnet flux opposing manner; and

(**ii**) simultaneously energising the first control coil in a permanent magnet flux opposing manner and the second control coil in a permanent magnet magnetic flux aiding manner.

20. A rotary motion device, comprising a rotor assembly including a shaft which defines an axis of rotation of the assembly, a rotor pole piece mounted for rotation with the shaft, the rotor pole piece including an outer ring portion having at least two path portions extending inwardly from a periphery of the outer ring portion;

a stator assembly including a permanent magnet having a generally ring-shaped configuration, a first pole face of the permanent magnet positioned adjacent the outer ring portion of the rotor pole piece, the stator assembly further comprising a stator pole piece including an outer ring portion positioned adjacent a second pole face of the permanent magnet and having a plurality of path portions extending inwardly from the periphery, each path portion further including a respective portion which extends toward a plane defined by the first pole face of the permanent magnet and capable of being aligned with each of the rotor pole piece path portions at certain rotational positions of the rotor pole piece, each path portion including a control coil positioned along it;

and circuit means connected to each of the coils and including a source of electrical energy and switch means for energising respective ones of the control coils in a predetermined timed sequence based upon rotation of the rotor assembly.

21. A rotary motion device, comprising:

a rotor assembly including a shaft which defines an axis of rotation of the assembly, a pair of spaced elongated rotor members mounted on the shaft at spaced locations thereon and angularity oriented with respect to each other, each of the elongated rotor members formed of a magnetic material;

a stator assembly including a permanent magnet having opposed first and second pole faces, a first pole piece positioned adjacent the first pole face and a second pole piece positioned adjacent the second pole face, each pole piece including a respective first path portion extending beyond a perimeter of its adjacent pole face and having an curved shaped end portion, the first path portion of the first pole piece aligned with the first path portion of the second pole piece, each pole piece further including a respective second path portion extending beyond the perimeter of its adjacent pole face in a direction opposite to that of the first path portions and having an curved shaped end portion, the second path portion of the first pole piece aligned with the second path portion of the second pole piece, at least one of the first path portions of the first pole piece and the first path portion of the second pole piece including a control coil mounted on at least one of the pole pieces, at least one of the second path portions of the first pole piece and the second path portion of the second pole piece including a control coil mounted on at least one of the pole pieces,

wherein the rotor assembly extends from end to end of the stator assembly such that the elongate members are aligned with the curved shaped end portions of the path portions of the pole pieces;

and circuit means connected to each of the coils and including a source of electrical energy and switch means for energising respective ones of the control coils in a predetermined timed sequence based upon rotation of the rotor assembly.

22. A rotary motion device comprising:

a rotor assembly including a shaft which defines an axis of rotation of the assembly, a ring-shaped rotor member mounted for rotation with the shaft, the ring-shaped rotor member including a plurality of distinct circumferential regions;

a stator assembly including a first permanent magnet, a first pole piece positioned against a first pole face and a second pole piece positioned against a second pole face, the first pole piece including at least a first path portion extending beyond a perimeter of the first pole face, the second pole piece including at least a first path portion extending beyond a perimeter of the second pole face, the first path portion of the first pole piece aligned with the first path portion of the second pole piece, at least a portion of the ring-shaped rotor member

positioned between the first path portion of the first pole piece and the first path portion of the second pole piece, at least one of the first path portions of the first pole piece and the first path portion of the second pole piece including a first control coil positioned at a point intermediate the first permanent magnet and the ring-shaped rotor member;

and circuit means connected to the first control coil and including a source of electrical energy and switch means for energising the first control coil in a predetermined timed manner based upon rotation of the rotor assembly.

23. The rotary motion device as set forth in claim 22, wherein the ring-shaped rotor member is formed by a permanent magnet having distinct circumferential regions of opposite polarity.

24. The rotary motion device as set forth in claim 23, wherein the first pole piece includes a second path portion spaced from and extending adjacent to the first path portion, the second pole piece including a second path portion spaced from and extending adjacent to the first path portion such that the second path portion of the first pole piece is aligned with the second path portion of the second pole piece, at least a portion of the ring-shaped permanent magnet rotor member positioned between the second path portion of the first pole piece and the second path portion of the second pole piece, at least one of the second path portions of the first pole piece and the second path portion of the second pole piece having a second control coil mounted on at least one of the pole pieces at a point intermediate the first permanent magnet and the ring-shaped permanent magnet rotor member, the second control coil connected to the circuit means so as to be energised in a predetermined timed manner based upon rotation of the rotor assembly.

25. The rotary motion device as set forth in claim 22, wherein the stator assembly further comprises a second permanent magnet, a third pole piece positioned adjacent a first pole face of the second permanent magnet and a fourth pole piece positioned adjacent a second pole face of the second permanent magnet, the third pole piece including at least a first path portion extending beyond a perimeter of the second permanent magnet first pole face, the fourth pole piece including at least a first path portion extending beyond a perimeter of the second permanent magnet second pole face, the first path portion of the third pole face aligned with the first path portion of the fourth pole piece, at least a portion of the ring-shaped permanent magnet rotor member positioned between the first path portion of the third pole piece and the first path portion of the fourth pole piece, at least one of the first path portions of the third pole piece and the first path portion of the fourth pole piece including a third control coil mounted on at least one of the pole pieces at a point intermediate the second permanent magnet and the ring-shaped permanent magnet rotor member, the third pole piece including a second path portion spaced from and extending adjacent to the first path portion the fourth pole piece including a second path portion spaced from and extending adjacent to the first path portion thereof such that the second path portion of the third pole piece is aligned with the second path portion of the fourth pole piece, at least a portion of the ring-shaped permanent magnet rotor member positioned between the second path portion of the third pole piece and the second path portion of the fourth pole piece, at least one of the second path portions of the third pole piece and the second path portion of the fourth pole piece including a fourth control coil mounted on at least one of the pole pieces at a point intermediate the second permanent magnet and the ring-shaped permanent magnet rotor member, wherein each of the third and fourth control coils are connected to the circuit means so as to be energised in a predetermined timed manner based upon rotation of the rotor assembly.

26. A device for producing rotary motion comprising:

a rotor assembly including a shaft which defines an axis of rotation for the assembly, a ring-shaped rotor member mounted for rotation with the shaft, the ring-shaped rotor member having a plurality of distinct circumferentially positioned regions extending around the axis, a stator assembly including a first permanent magnet, a first pole piece positioned against the first pole face of the first pole piece and a second pole piece positioned against a second pole face of the first pole piece, the first pole piece including at least a first path portion extending beyond a perimeter of the first pole face, the second pole piece including at least a first path portion extending beyond the perimeter of the second pole face, the first path portion of the first pole piece aligned with the first path portion of the second pole piece, at least a portion of the ring-shaped rotor member positioned between the first path portion of the first pole piece and the first path portion of the second pole piece, at least one of the first path portions of the first pole piece and the first path portion of the second pole piece including a first control coil mounted on at least one of the pole pieces at a point intermediate the first permanent magnet and the ring-shaped rotor member; and circuit means connected to the first control coil and including a source of electrical energy and switch means for energising the first control coil in a predetermined timed manner based upon position of the rotor assembly during rotation of the rotor assembly.

27. The device for producing rotary motion of claim 26 wherein the circuit means includes means for timing the energising of the first control coil includes means for adjusting the timing thereof.

28. The device for producing rotor motion of claim 26 including means to vary the flux generated in the first and second pole pieces.

29. A device for handling the flux between two separate permanent magnets each of which has a north magnetic pole adjacent one side face and the south magnetic pole adjacent to the opposite side face, the north and south side pole faces respectively of both magnets being substantially in alignment, a first member in surface-to-surface contact with the north magnetic faces of the spaced permanent magnets, a second member in surface-to-surface contact with the south magnetic faces of the spaced permanent magnets, first and second armatures each positioned adjacent opposite ends of the first and second permanent magnets and adjacent to opposite ends of the spaced members, a coil mounted on each of the members in the space between the adjacent permanent magnets, and means for applying voltages of predetermined polarities across the respective coils to change the magnetic coupling between the permanent magnets and between the armatures.

30. A device for producing rotational movement comprising:

a rotor having a shaft rotatable about the axis thereof, a member constructed of permanent magnets mounted on the shaft, said member having circumferential portions some of which have a north magnetic pole and others a south magnetic pole adjacent to the same side thereof, the opposite surface of the permanent magnet member having north magnetic poles opposite the south magnetic poles and south magnetic poles opposite the north magnetic poles, a stator having a plurality of circumferentially spaced portions each of which includes at least one permanent magnet and a pair of members mounted adjacent opposite sides of the permanent magnets, the members being positioned adjacent to the periphery of the rotor permanent magnet member and means on the member adjacent each opposite side of the stator permanent magnet for mounting a coil, and means for energising the coil on each stator portion in sequence to produce magnetic coupling force between the stator and the rotor in a direction to produce rotating motion of the rotor.

31. A device including a rotating member and a stationary member, each having a permanent magnet portion positioned to produce magnetic coupling force between them in predetermined positions thereof, the rotor including a shaft rotatable about its axis and the permanent magnet extending around the shaft and formed by a plurality of adjacent portions of permanent magnet material whereby adjacent portions have their north and south magnetic pole faces on opposite sides of the rotor permanent magnet, a plurality of stator members each stator member having at least one permanent magnet having a north magnetic pole adjacent one side and a south magnetic pole adjacent to the opposite side, a pair of members positioned adjacent respective opposite sides of the stator permanent magnet in position to extend to adjacent the rotor permanent magnet whereby a flux path is formed between the members and the stator and rotor permanent magnets, a coil mounted on each member of the stator and means for applying a voltage of predetermined polarity to each of said coils to control the flux through a path between the permanent magnets and to control the coupling force between the permanent magnets on the stator and the permanent magnets on the rotor.

32. A motion producing device comprising at least one permanent magnet having a north pole opposite and spaced from a south pole, a pair of spaced substantially parallel members adjacent respectively the north and south poles of the at least one permanent magnet and extending outwardly to substantially aligned opposite edges, a flux supporting member positioned adjacent the respective opposite edges of each pair of parallel members, a coil on selected ones of the parallel members, and a source of electrical energy connected to each of the coils for energising the coils to change the flux in the parallel members and in the flux supporting members.

33. The motion producing device of claim 32 wherein there are at least two spaced permanent magnets extending between the parallel members.

34. The motion producing device of claim 32 wherein one of said pair of parallel members is subdivided into a plurality of sidewardly extending portions extending to one of said opposite side edges, at least one of said coils being positioned on at least one of said sidewardly extending portions.

35. The motion producing device of claim 34 wherein there are coils on a plurality of respective ones of the sidewardly extending portions.

36. The motion producing device of claim 32 wherein the permanent magnet and the parallel members are annular in shape.

37. The motion producing device of claim 32 including a by-pass member extending between the pair of spaced substantial parallel members adjacent one side of the permanent magnet.

38. A permanent magnet device comprising at least two permanent magnets each having north and south pole faces, a first pole piece, a second pole piece, a first control coil, a second control coil and circuit means, the first pole piece positioned adjacent the north pole faces of the at least two permanent magnets and including a first path portion, a second path portion and a third path portion, the first path portion extending beyond the perimeter of the north pole faces and the second path portion extending beyond the perimeter of the north pole faces to define first and second flux paths for magnetic flux emitting from the north pole faces of the at least two permanent magnets, the first path portion of the first pole piece connected to the second path portion of the first pole piece by a third portion which extends across the north pole face of the at least two permanent magnets, the second pole piece positioned adjacent to the south pole faces of the at least two permanent magnets and including a first path portion and a second path portion, the first path portion extending beyond a perimeter of the south pole faces and substantially aligned with the first path portion of the first pole piece, the second path portion extending beyond the perimeter of the south pole faces and substantially aligned with the second path portion of the first pole piece, the first control coil positioned around the first path portion of the first pole piece, the second control coil positioned around the second path portion of the first pole piece, and the circuit means connected to each of the first control coil and the second control coil to alternately energise the first coil and the second coil in a timed sequential manner.

39. The permanent magnet device of claim 38 further comprising a first fixed armature extending between the first path portion of the first pole piece to the first path portion of the second pole piece and a second fixed armature extending between the second path portion of the first pole piece to the second path portion of the second pole piece.

40. The permanent magnet device of claim 39 where a first secondary coil is wrapped around the first fixed armature and a second secondary coil is wrapped around the second fixed armature.

41. The permanent magnet device of claim 40 including circuit means connected to the control coils to control the energising thereof to produce a varying flux in the armatures and to induce voltage in the secondary coils.

42. The permanent magnet device of claim 38 wherein there are at least two permanent magnets each having north and south pole faces, the first pole piece being positioned extending between the north pole faces of the permanent magnets and the second pole piece positioned extending between the south pole faces of the permanent magnets.

United States Patent 6,362,718 **Dated: 26th March 2002**

Inventors: Stephen Patrick, Thomas Bearden, James Hayes, Kenneth Moore and James Kenny

MOTIONLESS ELECTROMAGNETIC GENERATOR

Please note that this is a re-worded excerpt from this patent. It describes an electrical device which both powers itself and supplies current to additional external equipment.

ABSTRACT

An electromagnetic generator without moving parts includes a permanent magnet and a magnetic core including first and second magnetic paths. A first input coil and a first output coil extend around portions of the first magnetic path, while a second input coil and a second output coil extend around portions of the second magnetic path. The input coils are alternatively pulsed to provide induced current pulses in the output coils. Driving electrical current through each of the input coils reduces a level of flux from the permanent magnet within the magnet path around which the input coil extends. In an alternative embodiment of an electromagnetic generator, the magnetic core includes annular spaced-apart plates, with posts and permanent magnets extending in an alternating fashion between the plates. An output coil extends around each of these posts. Input coils extending around portions of the plates are pulsed to cause the induction of current within the output coils.

DESCRIPTION

1. Field of the Invention: This invention relates to a magnetic generator without moving parts, used to produce electrical power, and more particularly, to such a device capable of powering itself.

2. Description of the Related Art: The patent literature describes a number of magnetic generators, each of which includes a permanent magnet, two magnetic paths external to the permanent magnet, each of which extends between the opposite poles of the permanent magnet, switching means for causing magnetic flux to flow alternately along each of the two magnetic paths, and one or more output coils in which current is induced to flow by means of changes in the magnetic field within the device. These devices operate in accordance with an extension of Faraday's Law, indicating that an electrical current is induced within a conductor within a changing magnetic field, even if the source of the magnetic field is stationary.

A method for switching magnetic flux to flow predominantly along either of two magnetic paths between opposite poles of a permanent magnet is described as a "flux transfer" principle by R. J. Radus in Engineer's Digest, Jul. 23, 1963. This principle is used to exert a powerful magnetic force at one end of both the north and south poles and a very low force at the other end, without being used in the construction of a magnetic generator. This effect can be caused mechanically, by keeper movement, or electrically, by driving electrical current through one or more control windings extending around elongated versions of the pole pieces **14**. Several devices using this effect are described in U.S. Patent Nos. 3,165,723, 3,228,013, and 3,316,514.

Another step toward the development of a magnetic generator is described in U.S. Patent No. 3,368,141, as a device including a permanent magnet in combination with a transformer having first and second windings about a core, with two paths for magnetic flux leading from each pole of the permanent magnet to either end of the core, so that, when an alternating current induces magnetic flux direction changes in the core, the magnetic flux from the permanent magnet is automatically directed through the path which corresponds with the direction taken by the magnetic flux through the core due to the current. In this way, the magnetic flux is intensified. This device can be used to improve the power factor of a typically inductively loaded alternating current circuit.

Other patents describe magnetic generators in which electrical current from one or more output coils is described as being made available to drive a load, in the more conventional manner of a generator. For example, U.S. Patent No. 4,006,401 describes an electromagnetic generator including a permanent magnet and a core member, in which the magnetic flux flowing from the magnet in the core member is rapidly alternated by switching to generate an alternating current in a winding on the core member. The device includes a permanent magnet and two separate magnetic flux circuit paths between the north and south poles of the magnet. Each of the circuit paths includes two switching means for alternately opening and closing the

circuit paths, generating an alternating current in a winding on the core member. Each of the switching means includes a switching magnetic circuit intersecting the circuit path, with the switching magnetic circuit having a coil through which current is driven to induce magnetic flux to saturate the circuit path extending to the permanent magnet. Power to drive these coils is derived directly from the output of a continuously applied alternating current source. What is needed is an electromagnetic generator not requiring the application of such a current source.

U.S. Patent No. 4,077,001 describes a magnetic generator, or dc/dc converter, comprising a permanent magnet having spaced-apart poles and a permanent magnetic field extending between the poles of the magnet. A variable-reluctance core is disposed in the field in fixed relation to the magnet and the reluctance of the core is varied to cause the pattern of lines of force of the magnetic field to shift. An output conductor is disposed in the field in fixed relation to the magnet and is positioned to be cut by the shifting lines of permanent magnetic force so that a voltage is induced in the conductor. The magnetic flux is switched between alternate paths by means of switching coils extending around portions of the core, with the flow of current being alternated between these switching coils by means of a pair of transistors driven by the outputs of a flip-flop. The input to the flip flop is driven by an adjustable frequency oscillator. Power for this drive circuit is supplied through an additional, separate power source. What is needed is a magnetic generator not requiring the application of such a power source.

U.S. Patent No. 4,904,926 describes another magnetic generator using the motion of a magnetic field. The device includes an electrical winding defining a magnetically conductive zone having bases at each end, the winding including elements for the removing of an induced current therefrom. The generator further includes two pole magnets, each having a first and a second pole, each first pole in magnetic communication with one base of the magnetically conductive zone. The generator further includes a third pole magnet, the third pole magnet oriented intermediately of the first poles of the two pole electromagnets, the third pole magnet having a magnetic axis substantially transverse to an axis of the magnetically conductive zone, the third magnet having a pole nearest to the conductive zone and in magnetic attractive relationship to the first poles of the two pole electromagnets, in which the first poles thereof are like poles. Also included in the generator are elements, in the form of windings, for cyclically reversing the magnetic polarities of the electromagnets. These reversing means, through a cyclical change in the magnetic polarities of the electromagnets, cause the magnetic flux lines associated with the magnetic attractive relationship between the first poles of the electromagnets and the nearest pole of the third magnet to correspondingly reverse, causing a wiping effect across the magnetically conductive zone, as lines of magnetic flux swing between respective first poles of the two electromagnets, thereby inducing electron movement within the output windings and thus generating a flow of current within the output windings.

U.S. Patent No. 5,221,892 describes a magnetic generator in the form of a direct current flux compression transformer including a magnetic envelope having poles defining a magnetic axis and characterised by a pattern of magnetic flux lines in polar symmetry about the axis. The magnetic flux lines are spatially displaced relative to the magnetic envelope using control elements which are mechanically stationary relative to the core. Further provided are inductive elements which are also mechanically stationary relative to the magnetic envelope. Spatial displacement of the flux relative to the inductive elements causes a flow of electrical current. Further provided are magnetic flux valves which provide for the varying of the magnetic reluctance to create a time domain pattern of respectively enhanced and decreased magnetic reluctance across the magnetic valves, and, thereby, across the inductive elements.

Other patents describe devices using superconductive elements to cause movement of the magnetic flux. These devices operate in accordance with the Meissner effect, which describes the expulsion of magnetic flux from the interior of a superconducting structure as the structure undergoes the transition to a superconducting phase. For example, U.S. Patent No. 5,011,821 describes an electric power generating device including a bundle of conductors which are placed in a magnetic field generated by north and south pole pieces of a permanent magnet. The magnetic field is shifted back and forth through the bundle of conductors by a pair of thin films of superconductive material. One of the thin films is placed in the superconducting state while the other thin film is in a non-superconducting state. As the states are cyclically reversed between the two films, the magnetic field is deflected back and forth through the bundle of conductors.

U.S. Patent No. 5,327,015 describes an apparatus for producing an electrical impulse comprising a tube made of superconducting material, a source of magnetic flux mounted about one end of the tube, a means, such as a coil, for intercepting the flux mounted along the tube, and a means for changing the temperature of the superconductor mounted about the tube. As the tube is progressively made superconducting, the magnetic field is trapped within the tube, creating an electrical impulse in the means for intercepting. A reversal of the superconducting state produces a second pulse.
None of the patented devices described above use a portion of the electrical power generated within the device to power the reversing means used to change the path of magnetic flux. Thus, like conventional rotary generators, these devices require a steady input of power, which may be in the form of electrical power

driving the reversing means of one of these magnetic generators or the torque driving the rotor of a conventional rotary generator. Yet, the essential function of the magnetic portion of an electrical generator is simply to switch magnetic fields in accordance with precise timing. In most conventional applications of magnetic generators, the voltage is switched across coils, creating magnetic fields in the coils which are used to override the fields of permanent magnets, so that a substantial amount of power must be furnished to the generator to power the switching means, reducing the efficiency of the generator.

Recent advances in magnetic material, which have particularly been described by Robert C. O'Handley in Modern Magnetic Materials, Principles and Applications, John Wiley & Sons, New York, pp. 456-468, provide nanocrystalline magnetic alloys, which are particularly well suited forth rapid switching of magnetic flux. These alloys are primarily composed of crystalline grains, or crystallites, each of which has at least one dimension of a few nanometres. Nanocrystalline materials may be made by heat-treating amorphous alloys which form precursors for the nanocrystalline materials, to which insoluble elements, such as copper, are added to promote massive nucleation, and to which stable, refractory alloying materials, such as niobium or tantalum carbide are added to inhibit grain growth. Most of the volume of nanocrystalline alloys is composed of randomly distributed crystallites having dimensions of about 2-40 nm. These crystallites are nucleated and grown from an amorphous phase, with insoluble elements being rejected during the process of crystallite growth. In magnetic terms, each crystallite is a single-domain particle. The remaining volume of nanocrystalline alloys is made up of an amorphous phase in the form of grain boundaries having a thickness of about 1 nm.

Magnetic materials having particularly useful properties are formed from an amorphous Co--Nb--B (cobalt-niobium-boron) alloy having near-zero magnetostriction and relatively strong magnetisation, as well as good mechanical strength and corrosion resistance. A process of annealing this material can be varied to change the size of crystallites formed in the material, with a resulting strong effect on DC coercivity. The precipitation of nanocrystallites also enhances AC performance of the otherwise amorphous alloys.

Other magnetic materials are formed using iron-rich amorphous and nanocrystalline alloys, which generally show larger magnetisation that the alloys based on cobalt. Such materials are, for example, Fe--B--Si--Nb--Cu (iron-boron-silicon-niobium-copper) alloys. While the permeability of iron-rich amorphous alloys is limited by their relatively large levels of magnetostriction, the formation of a nanocrystalline material from such an amorphous alloy dramatically reduces this level of magnetostriction, favouring easy magnetisation.

Advances have also been made in the development of materials for permanent magnets, particularly in the development of materials including rare earth elements. Such materials include samarium cobalt, SmCo.sub.5, which is used to form a permanent magnet material having the highest resistance to demagnetisation of any known material. Other magnetic materials are made, for example, using combinations of iron, neodymium, and boron.

SUMMARY OF THE INVENTION:

It is a first objective of the present invention, to provide a magnetic generator which eliminates the need for an external power source during operation of the generator.

It is a second objective of the present invention to provide a magnetic generator in which a magnetic flux path is changed without a need to overpower a magnetic field to change its direction.

It is a third objective of the present invention to provide a magnetic generator in which the generation of electricity is accomplished without moving parts.

In the apparatus of the present invention, the path of the magnetic flux from a permanent magnet is switched in a manner not requiring the overpowering of the magnetic fields. Furthermore, a process of self-initiated iterative switching is used to switch the magnetic flux from the permanent magnet between alternate magnetic paths within the apparatus, with the power to operate the iterative switching being provided through a control circuit consisting of components known to use low levels of power. With self-switching, a need for an external power source during operation of the generator is eliminated, with a separate power source, such as a battery, being used only for a very short time during start-up of the generator.

According to a first aspect of the present invention, an electromagnetic generator is provided, including a permanent magnet, a magnetic core, first and second input coils, first and second output coils, and a switching circuit. The permanent magnet has magnetic poles at opposite ends. The magnetic core includes a first magnetic path, around which the first input and output coils extend, and a second magnetic path, around which the second input and output coils extend, between opposite ends of the permanent magnet. The switching circuit drives electrical current alternately through the first and second input coils. The electrical

current driven through the first input oil causes the first input coil to produce a magnetic field opposing a concentration of magnetic flux from the permanent magnet within the first magnetic path. The electrical current driven through the second input coil causes the second input coil to produce a magnetic field opposing a concentration of magnetic flux from the permanent magnet within the second magnetic path.

According to another aspect of the present invention, an electromagnetic generator is provided, including a magnetic core, a plurality of permanent magnets, first and second pluralities of input coils, a plurality of output coils, and a switching circuit. The magnetic core includes a pair of spaced-apart plates, each of which has a central aperture, and first and second pluralities of posts extending between the spaced-apart plates. The permanent magnets each extend between the pair of spaced apart plates. Each permanent magnet has magnetic poles at opposite ends, with the magnetic fields of all the permanent magnets being aligned to extend in a common direction. Each input coil extends around a portion of a plate within the spaced-apart plates, between a post and a permanent magnet. An output coil extends around each post. The switching circuit drives electrical current alternately through the first and second input coils. Electrical current driven through each input coil in the first plurality of input coils causes an increase in magnetic flux within each post within the first plurality of posts from permanent magnets on each side of the post and a decrease in magnetic flux within each post within the second plurality of posts from permanent magnets on each side of the post. Electrical current driven through each input coil in the second plurality of input coils causes a decrease in magnetic flux within each post within the first plurality of posts from permanent magnets on each side of the post and an increase in magnetic flux within each post within the second plurality of posts from permanent magnets on each side of the post.

BRIEF DESCRIPTION OF THE DRAWINGS:

Figure 1 is a partly schematic front elevation of a magnetic generator and associated electrical circuits built in accordance with a first version of the first embodiment of the present invention:

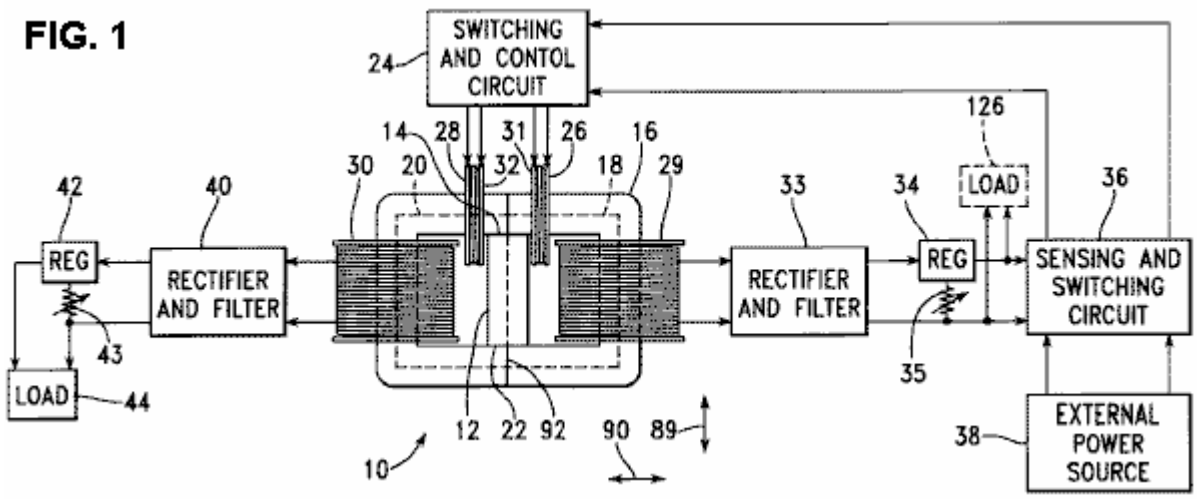


Figure 2 is a schematic view of a first version of a switching and control circuit within the associated electrical circuits of Figure 1:

FIG. 2

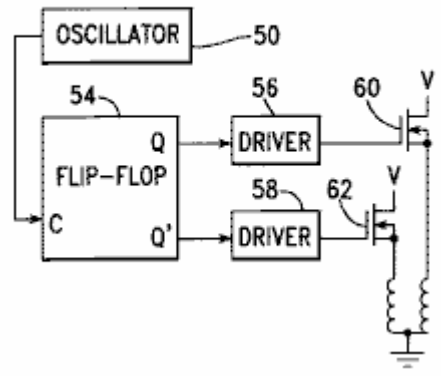


Figure 3 is a graphical view of drive signals produced within the circuit of Figure 2:

FIG. 3

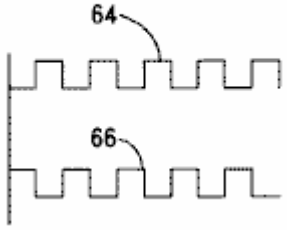


Figure 4 is a schematic view of a second version of a switching and control circuit within the associated electrical circuits of Figure 1:

FIG. 4

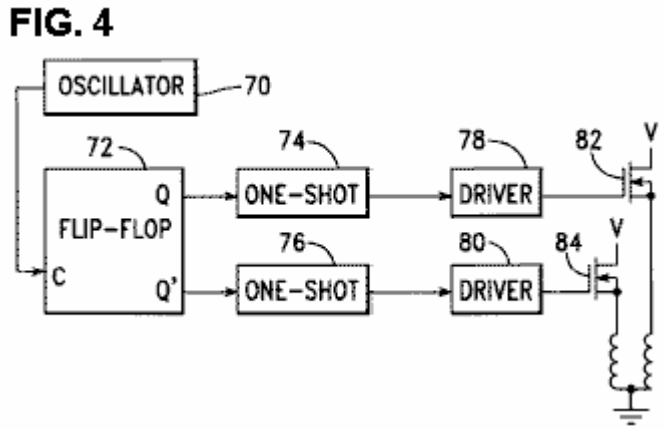


Figure 5 is a graphical view of drive signals produced within the circuit of Figure 3:

FIG. 5

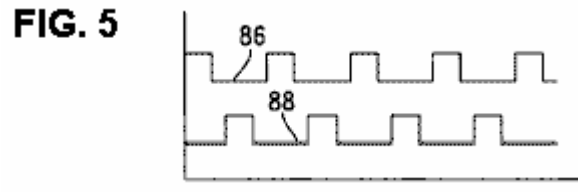


Figure 6A is a graphical view of a first drive signal within the apparatus of Figure 1,

Figure 6B is a graphical view of a second drive signal within the apparatus of Figure 1,

Figure 6C is a graphical view of an input voltage signal within the apparatus of Figure 1,

Figure 6D is a graphical view of an input current signal within the apparatus of Figure 1,

Figure 6E is a graphical view of a first output voltage signal within the apparatus of Figure 1,

Figure 6F is a graphical view of a second output voltage signal within the apparatus of Figure 1,

Figure 6G is a graphical view of a first output current signal within the apparatus of Figure 1,

Figure 6H is a graphical view of a second output current signal within the apparatus of Figure 1:

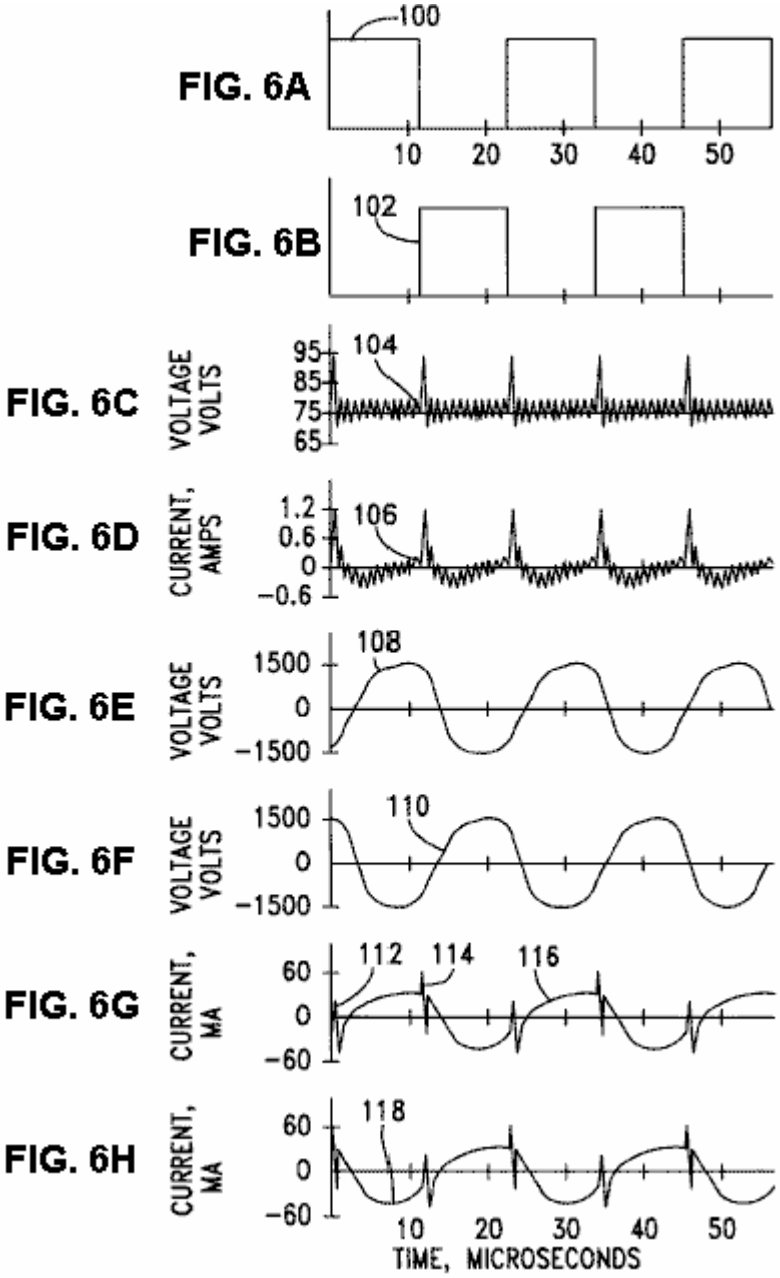


Figure 7 is a graphical view of output power measured within the apparatus of Figure 1, as a function of input voltage:

FIG. 7

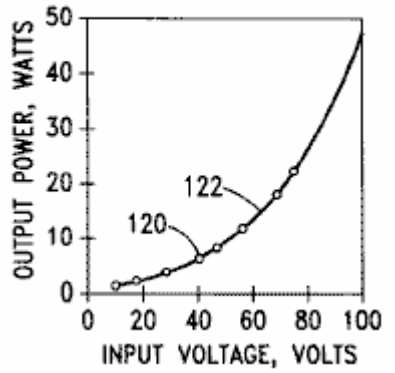


Figure 8 is a graphical view of a coefficient of performance, calculated from measurements within the apparatus of Figure 1, as a function of input voltage:

FIG. 8

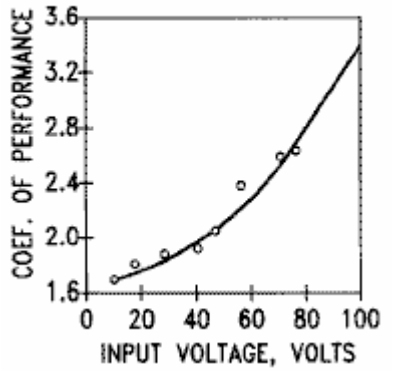


Figure 9 is a cross-sectional elevation of a second version of the first embodiment of the present invention:

FIG. 9

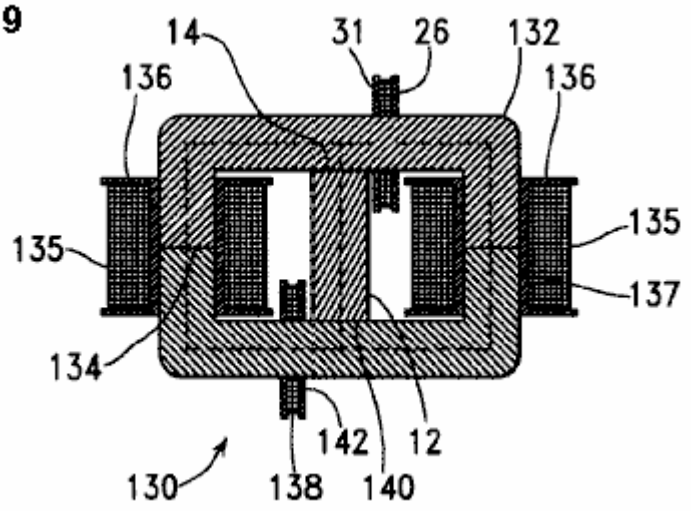


Figure 10 is a top view of a magnetic generator built in accordance with a first version of a second embodiment of the present invention:

FIG. 10

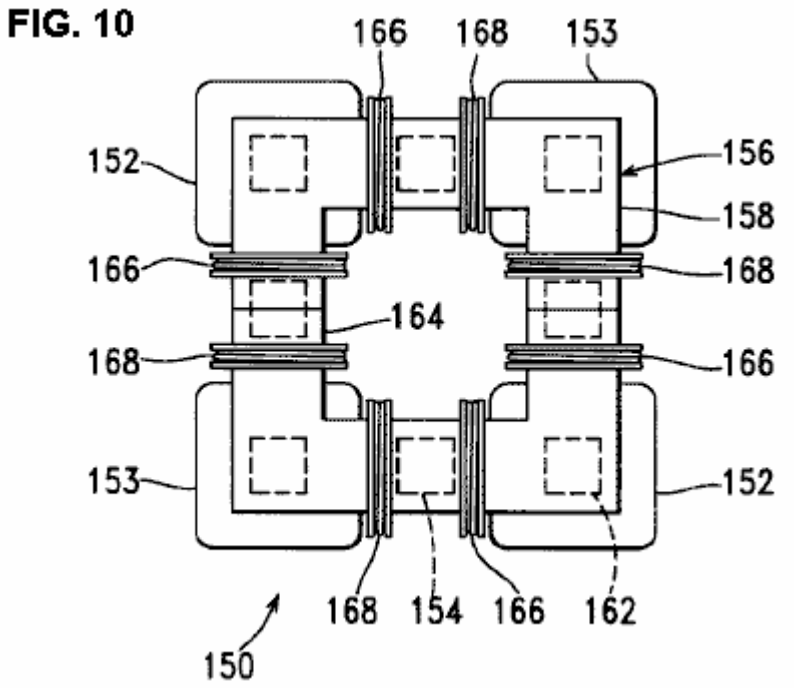


Figure 11 is a front elevation of the magnetic generator of Figure 10:

FIG. 11

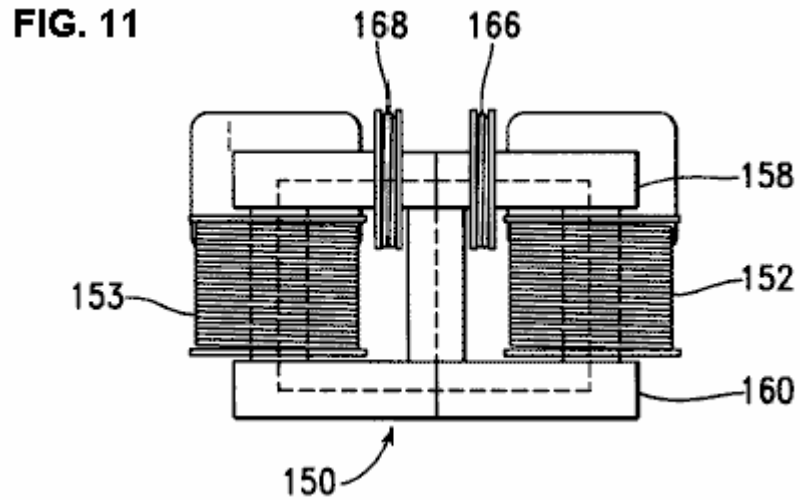
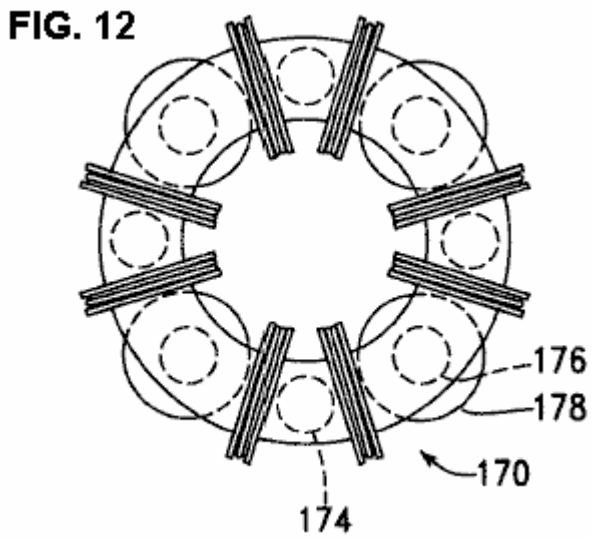


Figure 12 is a top view of a magnetic generator built in accordance with a second version of the second embodiment of the present invention:

FIG. 12



DETAILED DESCRIPTION OF THE INVENTION:

Fig.1 is a partly schematic front elevation, of an electromagnetic generator **10**, built in accordance with a first embodiment of the present invention, to include a permanent magnet **12** to supply input lines of magnetic flux moving from the north pole **14** of the magnet **12,** outward into magnetic flux path core material **16**.

The flux path core material **16** is configured to form a right magnetic path **18** and a left magnetic path **20**, both of which extend externally between the north pole **14** and the south pole **22** of the magnet **12**.

The electromagnetic generator **10** is driven by means of a switching and control circuit **24**, which alternately drives electrical current through a right input coil **26** and a left input coil **28**. These input coils each extend around a portion of the core material **16**, with the right input coil **26** surrounding a portion of the right magnetic path **18** and with the left input coil **28** surrounding a portion of the left magnetic path **20**. A right output coil **29** also surrounds a portion of the right magnetic path **18**, while a left output coil **30** surrounds a portion of the left magnetic path **20**.

In accordance with a preferred version of the present invention, the switching and control circuit **24** and the input coils **26, 28** are arranged so that, when the right input coil **26** is energised, a north magnetic pole is present at its left end **31**, the end closest to the north pole **14** of the permanent magnet **12**, and so that, when the left input coil **28** is energised, a north magnetic pole is present at its right end **32**, which is also the end closest to the north pole **14** of the permanent magnet **12**. Thus, when the right input coil **26** is magnetised, magnetic flux from the permanent magnet **12** is repelled from extending through the right input coil **26**. Similarly, when the left input coil **28** is magnetised, magnetic flux from the permanent magnet **12** is repelled from extending through the left input coil **28**.

Thus, it is seen that driving electrical current through the right input coil **26** opposes a concentration of flux from the permanent magnet **12** within the right magnetic path **18**, causing at least some of this flux to be transferred to the left magnetic path **20**. On the other hand, driving electrical current through the left input coil **28** opposes a concentration of flux from the permanent magnet **12** within the left magnetic path **20**, causing at least some of this flux to be transferred to the right magnetic path **18**.

While in the example of **Fig.1**, the input coils **26, 28** are placed on either side of the north pole of the permanent magnet **12**, being arranged along a portion of the core **16** extending from the north pole of the permanent magnet **12**, it is understood that the input coils **26, 28** could as easily be alternately placed on either side of the south pole of the permanent magnet **12**, being arranged along a portion of the core **16** extending from the south pole of the permanent magnet **12**, with the input coils **26, 28** being wired to form, when energised, magnetic fields having south poles directed toward the south pole of the permanent magnet **12**. In general, the input coils **26, 28** are arranged along the magnetic core on either side of an end of the permanent magnet forming a first pole, such as a north pole, with the input coils being arranged to produce magnetic fields of the polarity of the first pole directed toward the first pole of the permanent magnet.

Further in accordance with a preferred version of the present invention, the input coils **26, 28** are never driven with so much current that the core material **16** becomes saturated. Driving the core material **16** to saturation means that subsequent increases in input current can occur without effecting corresponding changes in magnetic flux, and therefore that input power can be wasted. In this way, the apparatus of the present

invention is provided with an advantage in terms of the efficient use of input power over the apparatus of U.S. Patent No. 4,000,401, in which a portion both ends of each magnetic path is driven to saturation to block flux flow.

In the electromagnetic generator **10**, the switching of current flow within the input coils **26, 28** does not need to be sufficient to stop the flow of flux in one of the magnetic paths **18, 20** while promoting the flow of magnetic flux in the other magnetic path. The electromagnetic generator **10** works by changing the flux pattern; it does not need to be completely switched from one side to another.

Experiments have determined that this configuration is superior, in terms of the efficiency of using power within the input coils **26, 28** to generate electrical power within the output coils **29, 30**, to the alternative of arranging input coils and the circuits driving them so that flux from the permanent magnet is driven through the input coils as they are energised. This arrangement of the present invention provides a significant advantage over the prior-art methods shown, for example, in U.S. Patent No. 4,077,001, in which the magnetic flux is driven through the energised coils.

The configuration of the present invention also has an advantage over the prior-art configurations of U.S. Patent Nos. 3,368,141 and 4,077,001 in that the magnetic flux is switched between two alternate magnetic paths **18, 20** with only a single input coil **26, 28** surrounding each of the alternate magnetic paths. The configurations of U.S. Patent Nos. 3,368,141 and 4,077,001 each require two input coils on each of the magnetic paths. This advantage of the present invention is significant both in the simplification of hardware and in increasing the efficiency of power conversion.

The right output coil **29** is electrically connected to a rectifier and filter **33**, having an output driven through a regulator **34**, which provides an output voltage adjustable through the use of a potentiometer **35**. The output of the linear regulator **34** is in turn provided as an input to a sensing and switching circuit **36**. Under start up conditions, the sensing and switching circuit **36** connects the switching and control circuit **24** to an external power source **38**, which is, for example, a starting battery. After the electromagnetic generator **10** is properly started, the sensing and switching circuit **36** senses that the voltage available from regulator **34** has reached a predetermined level, so that the power input to the switching and control circuit **24** is switched from the external power source **38** to the output of regulator **34**. After this switching occurs, the electromagnetic generator **10** continues to operate without an application of external power.

The left output coil **30** is electrically connected to a rectifier and filter **40**, the output of which is connected to a regulator **42**, the output voltage of which is adjusted by means of a potentiometer **43**. The output of the regulator **42** is in turn connected to an external load **44**.

Fig.2 is a schematic view of a first version of the switching and control circuit **24**. An oscillator **50** drives the clock input of a flip-flop **54**, with the Q and Q' outputs of the flip-flop **54** being connected through driver circuits **56, 58** to power FETs **60, 62** so that the input coils **26, 28** are driven alternately. In accordance with a preferred version of the present invention, the voltage V applied to the coils **26, 28** through the FETs **60, 62** is derived from the output of the sensing and switching circuit **36**.

Fig.3 is a graphical view of the signals driving the gates of FETs **60, 62** of **Fig.2**, with the voltage driving the gate of FET **60** being represented by line **64**, and with the voltage driving FET **62** being represented by line **66**. Both of the coils **26, 28** are driven with positive voltages.

Fig.4 is a schematic view of a second version of the switching and control circuit **24**. In this version, an oscillator **70** drives the clock input of a flip-flop **72**, with the Q and Q' outputs of the flip-flop **72** being connected to serve as triggers for one-shots **74, 76**. The outputs of the one-shots **74, 76** are in turn connected through driver circuits **78, 80** to drive FETs **82, 84**, so that the input coils **26, 28** are alternately driven with pulses shorter in duration than the Q and Q' outputs of the flip flop **72**.

Fig.5 is a graphical view of the signals driving the gates of FETs **82, 84** of **Fig.4**, with the voltage driving the gate of FET **82** being represented by line **86**, and with the voltage driving the gate of FET **84** being represented by line **88**.

Referring again to **Fig.1**, power is generated in the right output coil **29** only when the level of magnetic flux is changing in the right magnetic path **18**, and in the left output coil **30** only when the level of magnetic flux is changing in the left magnetic path **20**. It is therefore desirable to determine, for a specific magnetic generator configuration, the width of a pulse providing the most rapid practical change in magnetic flux, and then to provide this pulse width either by varying the frequency of the oscillator **50** of the apparatus of **Fig.2**, so that this pulse width is provided with the signals shown in **Fig.3**, or by varying the time constant of the one-shots **74, 76** of **Fig.4**, so that this pulse width is provided by the signals of **Fig.5** at a lower oscillator frequency. In this way, the input coils are not left on longer than necessary. When either of the input coils is left on for a

period of time longer than that necessary to produce the change in flux direction, power is being wasted through heating within the input coil without additional generation of power in the corresponding output coil.

A number of experiments have been conducted to determine the adequacy of an electromagnetic generator built as the generator **10** in **Fig.1**, to produce power both to drive the switching and control logic, providing power to the input coils **26, 28**, and to drive an external load **44**. In the configuration used in this experiment, the input coils **26, 28** had 40 turns of 18-gauge copper wire, and the output coils **29, 30** had 450 turns of 18-gauge copper wire. The permanent magnet **12** had a height of 40 mm (1.575 in. between its north and south poles, in the direction of arrow **89**, a width of 25.4 mm (1.00 in.), in the direction of arrow **90**, and in the other direction, a depth of 38.1 mm (1.50 in.). The core **16** had a height, in the direction of arrow **89**, of 90 mm (3.542 in.), a width, in the direction of arrow **90**, of 135 mm (5.315 in.) and a depth of 70 mm (2.756 in.). The core **16** had a central hole with a height, in the direction of arrow **89**, of 40 mm (1.575 mm) to accommodate the magnet **12**, and a width, in the direction of arrow **90**, of 85 mm (3.346 in.). The core **16** was fabricated of two "C"-shaped halves, joined at lines **92**, to accommodate the winding of output coils **29, 30** and input coils **26, 28** over the core material.

The core material was a laminated iron-based magnetic alloy sold by Honeywell as METGLAS Magnetic Alloy 2605SA1. The magnet material was a combination of iron, neodymium, and boron.

The input coils **26, 28** were driven at an oscillator frequency of 87.5 KHz, which was determined to produce optimum efficiency using a switching control circuit configured as shown in **Fig.2**. This frequency has a period of 11.45 microseconds. The flip flop **54** is arranged, for example, to be set and reset on rising edges of the clock signal input from the oscillator, so that each pulse driving one of the FETs **60, 62** has a duration of 11.45 microseconds, and so that sequential pulses are also separated to each FET are also separated by 11.45 microseconds.

Fig.6A to **Fig.6H**, are graphical views of signals which occurred simultaneously during the operation of the apparatus shown in **Fig.1** and **Fig.2**, when the input voltage applied was 75 volts. **Fig.6A** shows a first drive signal **100** driving FET **60**, which conducts to drive the right input coil **26**. **Fig.6B** shows a second drive signal **102**, driving FET **62**, which, when it conducts, provides the drive to the left input coil **28**.

Fig.6C and **Fig.6D** show voltage and current signals produced when the current driving the FETs **60, 62** is provided from a battery source. **Fig.6C** shows the level **104** of voltage V. While the nominal voltage of the battery was 75 volts, a decaying transient signal **106** is superimposed on this voltage each time one of the FETs **60, 62** is switched on. The specific pattern of this transient signal depends on the internal resistance of the battery, as well as on a number of characteristics of the magnetic generator **10**. Similarly, **Fig.6D** shows the current **106** flowing into FETs **60, 62** from the battery source. Since the signals **104, 106** show the effects of current flowing into both FETs **60, 62** the transient spikes are 11.45 microseconds apart.

Figs.6E to 6H, show the voltage and current levels measured at the output coils **29, 30**. **Fig.6E** shows a voltage output signal **108** of the right output coil **29**, while **Fig.6F** shows a voltage output signal **110** of the left output coil **30**. For example, the output current signal **116** of the right output coil **29** includes a first transient spike **112** caused when a pulse of current is generated in the left input coil **28** in order to boost the magnetic flux passing through the right magnetic path **18**, and a second transient spike **114** caused when the left input coil **28** is turned off as the right input coil **26** is being turned on. **Fig.6G** shows an output current signal **116** of the right output coil **29**, while **Fig.6H** shows an output current signal **118** of the left output coil **30**.

Fig.7 is a graphical view of output power measured using the electromagnetic generator **10** and eight levels of input voltage, varying from 10v to 75v. The oscillator frequency was retained at 87.5 KHz. The measured values are represented by points **120**, while the curve **122** is generated by polynomial regression, (a least squares fit).

Fig.8 is a graphical view of a coefficient of performance, defined as the ratio of the output power to the input power, for each of the measurement points shown in **Fig.7**. At each measurement point, the output power was substantially higher than the input power. Real power measurements were computed at each data point using measured voltage and current levels, with the results being averaged over the period of the signal. These measurements agree with RMS power measured using a Textronic THS730 digital oscilloscope.

While the electromagnetic generator **10** was capable of operation at much higher voltages and currents without saturation, the input voltage was limited to 75 volts because of voltage limitations of the switching circuits being used. Those familiar with electronics will understand that components for switching circuits capable of handling higher voltages are readily available for use in this application.

The experimentally-measured data were extrapolated to predict operation at an input voltage of 100 volts, with the input current being 140 mA, the input power being 14 watts, and with a resulting output power being

48 watts for each of the two output coils **29, 30**, at an average output current of 12 mA and an average output voltage of 4000 volts. This means that for each of the output coils **29, 30**, the coefficient of performance ("COP") would be 3.44.

While an output voltage of 4000 volts may be needed for some applications, the output voltage can also be varied through a simple change in the configuration of the electromagnetic generator **10**. The output voltage is readily reduced by reducing the number of turns in the output windings. If this number of turns is decreased from 450 to 12, the output voltage is dropped to 106.7, with a resulting increase in output current to 0.5 amps for each output coil **29, 30**, (i.e. 53 watts). In this way, the output current and voltage of the electromagnetic generator can be varied by varying the number of turns of the output coils **29, 30**, without making a substantial change in the output power, which is instead determined by the input current, which determines the amount of magnetic flux shuttled during the switching process.

All of the **C**oefficients **O**f **P**erformance were significantly greater than 1. These are plotted in **Fig.8** and they indicate that the output power levels measured in each of the output coils **29, 30** were substantially greater than the corresponding input power levels driving both of the input coils **26, 28**. Therefore, it is apparent that the electromagnetic generator **10** can be built in a self-powered form, as discussed above in reference to **Fig.1**. In the example of **Fig.1**, except for a brief application of power from the external power source **38** to start the process of power generation, the power required to drive the input coils **26, 28** is derived entirely from power developed within the right output coil **29**. If the power generated in the single output coil **29**, is more than sufficient to drive the input coils **26, 28**, an additional load **126** may be added to be driven with power generated in the output coil **29**. On the other hand, each of the output coils **29, 30** may be used to drive a portion of the input coil power requirements, for example, output coils 26 can provide the driving voltage V for FET **60** while output coil 28 can provide the driving voltage V for FET **62**.

Regarding thermodynamic considerations, it is noted that when the electromagnetic generator **10** is operating, it is an open system not in thermodynamic equilibrium. The system receives static energy from the magnetic flux of the permanent magnet. Because the electromagnetic generator **10** is self-switched without an additional energy input, the thermodynamic operation of the system is an open dissipative system, receiving, collecting, and dissipating energy from its environment; in this case, from the magnetic flux stored within the permanent magnet. Continued operation of the electromagnetic generator **10** causes demagnetisation of the permanent magnet. The use of a magnetic material including rare earth elements, such as a samarium cobalt material or a material including iron, neodymium, and boron is preferable within the present invention, since such a magnetic material has a relatively long life in this application.

Thus, an electromagnetic generator operating in accordance with the present invention should not be considered as a perpetual-motion machine, but rather as a system in which flux radiated from a permanent magnet is converted into electricity, which is used both to power the apparatus and to power an external load. This is analogous to a system including a nuclear reactor, in which a number of fuel rods radiate energy which is used to keep the chain reaction going and to heat water for the generation of electricity to drive external loads.

Fig.9 is a cross-sectional elevation of an electromagnetic generator **130** built in accordance with a second version of the first embodiment of the present invention. This electromagnetic generator **130** is generally similar in construction and operation to the electromagnetic generator **10** built in accordance with the first version of this embodiment, except that the magnetic core **132** of the electromagnetic generator **10** is built in two halves joined along lines **134**, allowing each of the output coils **135** to be wound on a plastic bobbin **136** before being placed over the legs **137** of the core **132**.

Fig.9 also shows an alternate placement of an input coil **138**. In the example of **Fig.1**, both of the input coils **26, 28** were placed on the upper portion of the magnetic core **16**, with these coils being configured to generate magnetic fields having north magnetic poles at the inner ends **31, 32** of the coils **26, 28**, with these north magnetic poles thus being closest to the end **14** of the permanent magnet **12** having its north magnetic pole. In the example of **Fig.9**, a first input coil **26** is as described above in reference to **Fig.1**, but the second input coil **138** is placed adjacent the south pole **140** of the permanent magnet **12**. This input coil **138** is configured to generate a south magnetic pole at its inner end **142**, so that, when input coil **138** is turned on, flux from the permanent magnet **12** is directed away from the left magnetic path **20** into the right magnetic path **18**.

Fig.10 and **Fig.11** show an electromagnetic generator **150** built in accordance with a first version of a second embodiment of the present invention, with **Fig.10** being a top view, and **Fig.11** being a front elevation. This electromagnetic generator **150** includes an output coil **152, 153** at each corner, and a permanent magnet **154** extending along each side between output coils. The magnetic core **156** includes an upper plate **158**, a lower plate **160**, and a square post **162** extending within each output coil **152, 153**. Both the upper plate **158** and the lower plate **160** include central apertures **164**.

Each of the permanent magnets **154** is oriented with a like pole, such as a north pole, against the upper plate **158**. Eight input coils **166, 168** are placed in positions around the upper plate **158** between an output coil **152, 153** and a permanent magnet **154**. Each input coil **166, 168** is arranged to form a magnetic pole at its end nearest to the adjacent permanent magnet **154** of the same polarity as the magnetic poles of the magnets **154** adjacent the upper plate **158**. Thus, the input coils **166** are switched on to divert the magnetic flux of the permanent magnets **154** from the adjacent output coils **152**, into magnetic paths through the output coils **153**. Then, the input coils **168** are switched on to divert magnetic flux of the permanent magnets **154** from the adjacent output coils **153**, with this flux being diverted into magnetic paths through the output coils **152**. Thus, the input coils form a first group of input coils **166** and a second group of input coils **168**, with these first and second groups of input coils being alternately energised in the manner described above in reference to **Fig.1** for the single input coils **26, 28**. The output coils produce current in a first train of pulses occurring simultaneously within coils **152** and in a second train of pulses occurring simultaneously within coils **153**.

Thus, driving current through input coils **166** causes an increase in flux from the permanent magnets **154** within the posts **162** extending through output coils **153** and a decrease in flux from the permanent magnets **154** within the posts **162** extending through output coils **152**. On the other hand, driving current through input coils **168** causes a decrease in flux from the permanent magnets **154** within the posts **162** extending through output coils **153** and an increase in flux from the permanent magnets **154** within the posts **162** extending through output coils **152**.

While the example of **Fig.10** and **Fig.11** shows all of the input coils **166,168** deployed along the upper plate **158**, it is understood that certain of these input coils **166, 168** could alternately be deployed around the lower plate **160**, in the manner generally shown in **Fig.9**, with one input coil **166**, **168** being within each magnetic circuit between a permanent magnet **154** and an adjacent post **162** extending within an output coil **152, 153**, and with each input coil **166, 168** being arranged to produce a magnetic field having a magnetic pole like the closest pole of the adjacent permanent magnet **154**.

Fig.12 is a top view of a second version **170** of the second embodiment of the present invention, which is similar to the first version thereof, which has been discussed in reference to **Fig.10** and **Fig.11**, except that an upper plate **172** and a similar lower plate (not shown) are annular in shape, while the permanent magnets **174** and posts **176** extending through the output coils **178** are cylindrical. The input coils **180** are oriented and switched as described above in reference to **Fig.9** and **Fig.10**.

While the example of **Fig.12** shows four permanent magnets, four output coils and eight input coils it is understood that the principles described above can be applied to electromagnetic generators having different numbers of elements. For example, such a device can be built to have two permanent magnets, two output coils, and four input coils, or to have six permanent magnets, six output coils, and twelve input coils.

In accordance with the present invention, material used for magnetic cores is preferably a nanocrystalline alloy, and alternately an amorphous alloy. The material is preferably in a laminated form. For example, the core material is a cobalt-niobium-boron alloy or an iron based magnetic alloy.

Also in accordance with the present invention, the permanent magnet material preferably includes a rare earth element. For example, the permanent magnet material is a samarium cobalt material or a combination of iron, neodymium, and boron.

While the invention has been described in its preferred versions and embodiments with some degree of particularity, it is understood that this description has been given only by way of example and that numerous changes in the details of construction, fabrication, and use, including the combination and arrangement of parts, may be made without departing from the spirit and scope of the invention.

CLAIMS:

1. An electromagnetic generator comprising: a permanent magnet having magnetic poles at opposite ends; a magnetic core including first and second magnetic paths between said opposite ends of said permanent magnet, wherein said magnetic core comprises a closed loop, said permanent magnet extends within said closed loop, and said opposite ends of said permanent magnet are disposed adjacent opposite sides of said closed loop and against internal surfaces of said magnetic core comprising said closed loop; a first input coil extending around a portion of said first magnetic path, a second input coil extending around a portion of said second magnetic path, a first output coil extending around a portion of said first magnetic path for providing a first electrical output; a second output coil extending around a portion of said second magnetic path for providing a second electrical output; and a switching circuit driving electrical current alternately through said first and second input coils, wherein said electrical current driven through said first input coil causes said first input coil to produce a magnetic field opposing a concentration of magnetic flux

from said permanent magnet within said first magnetic path, and said electrical current driven through said second input coil causes said second input coil to produce a magnetic field opposing a concentration of magnetic flux from said permanent magnet within said second magnetic path.

2. An electromagnetic generator comprising: a permanent magnet having magnetic poles at opposite ends; a magnetic core including first and second magnetic paths between said opposite ends of said permanent magnet, wherein said magnetic core comprises a closed loop, said permanent magnet extends within said closed loop, said opposite ends of said permanent magnet are disposed adjacent opposite sides of said closed loop, and a first type of pole of said permanent magnet is disposed adjacent a first side of said closed loop; a first input coil, disposed along said first side of said closed loop, extending around a portion of said first magnetic path, a second input coil, disposed along said first side of said closed loop, extending around a portion of said second magnetic path, a first output coil extending around a portion of said first magnetic path for providing a first electrical output; a second output coil extending around a portion of said second magnetic path for providing a second electrical output; and a switching circuit driving electrical current alternately through said first and second input coils, wherein said electrical current driven through said first input coil causes said first input coil to produce a magnetic field opposing a concentration of magnetic flux from said permanent magnet within said first magnetic path, and additionally causes said first input coil to produce a magnetic field having said first type of pole at an end of said first input coil adjacent said permanent magnet, and said electrical current driven through said second input coil causes said second input coil to produce a magnetic field opposing a concentration of magnetic flux from said permanent magnet within said second magnetic path, and additionally causes said second input coil to produce a magnetic field having said first type of pole at an end of said of said second input coil adjacent said permanent magnet.

3. An electromagnetic generator comprising: a permanent magnet having magnetic poles at opposite ends; a magnetic core including first and second magnetic paths between said opposite ends of said permanent magnet, wherein said magnetic core comprises a closed loop, said permanent magnet extends within said closed loop, and said opposite ends of said permanent magnet are disposed adjacent opposite sides of said closed loop, a first type of pole of said permanent magnet is disposed adjacent a first side of said closed loop, and a second type of pole, opposite said first type of pole, of said permanent magnet is disposed adjacent a second side of said closed loop; a first input coil extending around a portion of said first magnetic path, wherein said first input coil is disposed along said first side of said closed loop; a second input coil extending around a portion of said second magnetic path wherein said second input coil is disposed along said second side of said closed loop; a first output coil extending around a portion of said first magnetic path for providing a first electrical output; a second output coil extending around a portion of said second magnetic path for providing a second electrical output; and a switching circuit driving electrical current alternately through said first and second input coils, wherein said electrical current driven through said first input coil causes said first input coil to produce a magnetic field opposing a concentration of magnetic flux from said permanent magnet within said first magnetic path, and additionally causes said first input coil to produce a magnetic field having said first type of pole at an end of said first input coil adjacent said permanent magnet, and said electrical current driven through said second input coil causes said second input coil to produce a magnetic field opposing a concentration of magnetic flux from said permanent magnet within said second magnetic path, and additionally causes said second input coil to produce a magnetic field having said second type of pole at an end of said of said second input coil adjacent said permanent magnet.

4. An electromagnetic generator comprising: a permanent magnet having magnetic poles at opposite ends; a magnetic core including first and second magnetic paths between said opposite ends of said permanent magnet; a first input coil extending around a portion of said first magnetic path, a second input coil extending around a portion of said second magnetic path, a first output coil extending around a portion of said first magnetic path for providing a first electrical output; a second output coil extending around a portion of said second magnetic path for providing a second electrical output; and a switching circuit driving electrical current alternately through said first and second input coils, wherein said electrical current driven through said first input coil causes said first input coil to produce a magnetic field opposing a concentration of magnetic flux from said permanent magnet within said first magnetic path, and wherein said electrical current driven through said second input coil causes said second input coil to produce a magnetic field opposing a concentration of magnetic flux from said permanent magnet within said second magnetic path, wherein a portion of electrical power induced in said first output coil provides power to drive said switching circuit.

5. The electromagnetic generator of claim 4, wherein said switching circuit is driven by an external power source during a starting process and by power induced in said first output coil during operation after said starting process.

6. The electromagnetic generator of claim 2, wherein said magnetic core is composed of a nanocrystalline magnetic alloy.

7. The electromagnetic generator of claim 6, wherein said nanocrystalline magnetic alloy is a cobalt-niobium-boron alloy.

8. The electromagnetic generator of claim 6, wherein said nanocrystalline magnetic alloy is an iron-based alloy.

9. The electromagnetic generator of claim 2, wherein said changes in flux density within said magnetic core occur without driving said magnetic core to magnetic saturation.

10. The electromagnetic generator of claim 2, wherein said switching circuit drives said electrical current through said first input coil in response to a first train of pulses, said switching circuit drives said electrical current through said second input coil in response to a second train of pulses, alternating with pulses within said first train of pulses, and said pulses in said first and second trains of pulses are approximately 11.5 milliseconds in duration.

11. The electromagnetic generator of claim 2, wherein said permanent magnet is composed of a material including a rare earth element.

12. The electromagnetic generator of claim 11, wherein said permanent magnet is composed essentially of samarium cobalt.

13. The electromagnetic generator of claim 11, wherein said permanent magnet is composed essentially of iron, neodymium, and boron.

14. An electromagnetic generator comprising: a magnetic core including a pair of spaced-apart plates, wherein each of said spaced-apart plates includes a central aperture, and first and second pluralities of posts extending between said spaced-apart plates; a plurality of permanent magnets extending individually between said pair of spaced-apart plates and between adjacent posts within said plurality of posts, wherein each permanent magnet within said plurality of permanent magnets has magnetic poles at opposite ends, wherein all magnets within said plurality of magnets are oriented to produce magnetic fields having a common direction; first and second pluralities of input coils, wherein each input coil within said first and second pluralities of input coils extends around a portion of a plate within said spaced-apart plates between a post in said plurality of posts and a permanent magnet in said plurality of permanent magnets; an output coil extending around each post in said first and second pluralities of posts for providing an electrical output; a switching circuit driving electrical current alternatively through said first and second pluralities of input coils, wherein said electrical current driven through each input coil in said first plurality of input coils causes an increase in magnetic flux within each post within said first plurality of posts from permanent magnets on each side of said post and a decrease in magnetic flux within each post within said second plurality of posts from permanent magnets on each side of said post, and wherein said electrical current driven through input coil in said second plurality of input coils causes a decrease in magnetic flux within each post within said first plurality of posts from permanent magnets on each side of said post and an increase in magnetic flux within each post within said second plurality of posts from permanent magnets on each side of said post.

15. The electromagnetic generator of claim 14, wherein each input coil extends around a portion of a magnetic path through said magnetic core between said opposite ends a permanent magnet adjacent said input coil, said magnetic path extends through a post within said magnetic core adjacent said input coil, and driving electrical current through said input coil causes said input coil to produce a magnetic field opposing a concentration of magnetic flux within said magnetic path.

16. The electromagnetic generator of claim 14, wherein said switching circuit is driven by an external power source during a starting process and by power induced in said output coils during operation after said starting process.

17. The electromagnetic generator of claim 14, wherein said magnetic core is composed of a nanocrystalline magnetic alloy.

18. The electromagnetic generator of claim 2, wherein a portion of electrical power induced in said first output coil provides power to drive said switching circuit.

19. The electromagnetic generator of claim 18, wherein said switching circuit is driven by an external power source during a starting process and by power induced in said first output coil during operation after said starting process.

20. The electromagnetic generator of claim 3, wherein a portion of electrical power induced in said first output coil provides power to drive said switching circuit.

21. The electromagnetic generator of claim 20, wherein said switching circuit is driven by an external power source during a starting process and by power induced in said first output coil during operation after said starting process.

22. The electromagnetic generator of claim 3, wherein said magnetic core is composed of a nanocrystalline magnetic alloy.

23. The electromagnetic generator of claim 22, wherein said nanocrystalline magnetic alloy is a cobalt-niobium-boron alloy.

24. The electromagnetic generator of claim 22, wherein said nanocrystalline magnetic alloy is an iron-based alloy.

25. The electromagnetic generator of claim 3, wherein said changes in flux density within said magnetic core occur without driving said magnetic core to magnetic saturation.

26. The electromagnetic generator of claim 3, wherein said switching circuit drives said electrical current through said first input coil in response to a first train of pulses, said switching circuit drives said electrical current through said second input coil in response to a second train of pulses, alternating with pulses within said first train of pulses, and said pulses in said first and second trains of pulses are approximately 11.5 milliseconds in duration.

27. The electromagnetic generator of claim 3, wherein said permanent magnet is composed of a material including a rare earth element.

28. The electromagnetic generator of claim 27, wherein said permanent magnet is composed essentially of samarium cobalt.

29. The electromagnetic generator of claim 27, wherein said permanent magnet is composed essentially of iron, neodymium, and boron.

DAN DAVIDSON: ACOUSTIC-MAGNETIC POWER GENERATOR

US Patent 5,568,005 22nd October 1996 Inventor: Dan A. Davidson

ACOUSTIC-MAGNETIC POWER GENERATOR

Please note that this is a re-worded excerpt from this patent. If the content interests you, then you should obtain a full copy via the www.freepatentsonline.com web site. This patent describes an electrical device very similar to the MEG device, capable of powering itself while powering additional external items of equipment.

ABSTRACT

The Acoustic Magnetic Field Power Generator uses an acoustic signal focused into a permanent magnet to stimulate the nuclear structure of the magnet to cause the magnetic field of the permanent magnet to move or oscillate. This effect can be used to tap power from the oscillating magnetic field by putting a coil of wire in the oscillating field. When an alternating current signal generator is connected simultaneously to an acoustic transducer and a stimulating coil; whereby, both the acoustic transducer and the stimulating coil are located within the magnetic field of the magnet, the acoustic signal enhances the stimulating effect to the permanent magnet transformer. The acoustic transducer can be any acoustic generation device such as a piezoelectric, magnetostrictive, or other acoustic transducer. The combined effect of the acoustic signal and the stimulating coil increases the efficiency of permanent magnet induction transformers.

BACKGROUND OF THE INVENTION

The present invention relates to a solid state electrical generator having no moving parts. More particularly, the invention makes use of a new method of stimulating the nuclear material of a permanent magnet so that the electronic structure of the atom will vibrate and thus cause the magnetic field of the permanent magnet to oscillate. It is a well-known fact that an oscillating magnetic field will induce electrical current in a coil as was discovered by Michael Faraday in the last century. What is new in this invention, is the discovery of the ability of an acoustic field to stimulate the nuclear structure of a material to cause the electrons to wobble under the influence of the acoustic field. If the material is magnetic or temporarily magnetised by an external magnetic field then the magnetic field will vibrate under the stimulus of the acoustic field. If this effect is combined with a coil which is simultaneously stimulating the magnet then the efficiency of stimulating the permanent magnet's field is enhanced. If a pickup coil is placed in the oscillating magnetic field so as to create an induction transformer then the combination of the acoustic and magnetic stimulation will enhance the efficiency of the induction transformer.

The most relevant prior art known to the inventor comprises U.S. Pat. No. 4,904,926 (1990) to Mario Pasichinsky, entitled Magnet Motion Electrical Generator; and U.S. Pat. No. 4,077,001 (1978) to Frank Richardson, entitled Electromagnetic Converter With Stationary Variable-Reluctance Members; and U.S. Pat. No. 4,006,401 (1977) to de Rivas, entitled Electromagnetic Generator.

The above references to Pasichinsky, Richardson, and de Rivas, all use inductive methods to stimulate the motion of a permanent magnetic field. In the de Rivas invention, ‘Electromagnetic Generator’, the flux of the permanent magnet is "alternated by switching" using inductive coupling. In the Richardson disclosure an "energy conversion system" the flux of the permanent magnet is also "shifted" by inductive means. In the Pasichinsky disclosure, alternating magnetic coils induce flux changes in a closed magnetic circuit and output coils attached to the circuit are induced by the changing flux to produce a magnetic field. All of these devices are essentially variations of transformer design with permanent magnets as part of the transformer cores and all use magnetic induction. The transformer aspect of these references is the use of permanent magnets as the transformer core with coils wrapped around the magnetic core which are energised to produce oscillation or movement of the permanent magnet's field. The above references will, in this document, be called "permanent magnet transformers".

Other prior art relevant to the invention are U.S. Pat. No. 2,101,272 (1937) to H. J. Scott, entitled Combined Magnetostriction and Piezoelectric Selective Device; and U.S. Pat. No. 2,636,135 (1953) to R. L. Peek, Jr. entitled Stress Coupled Core and Crystal Transformer, and U.S. Pat. No. 2,834,943 (1958) to R. O. Grisdale, etal entitled Mechanically Coupled Electromechanical and Magnetomechanical Transducers, and U.S. Pat. No. 3,246,287 (1966) to H. F. Benson entitled Piezoelectric Transformer, and U.S. Pat. No. 3,261,339 (1966) to H. P. Quinn entitled Magnetostrictive Transformer, and U.S. Pat. No. 3,274,406 (1966) to H. S. Sommers, Jr. entitled Acoustic Electromagnetic Device, and U.S. Pat. No. 3,309,628 (1967) to F. A. Olson entitled YIG Variable Acoustic Delay Line, and U.S. Pat. No. 3,457,463 (1969) to L. Balamuth entitled Method and Apparatus for Generating Electric Currents of Small Magnitude, and U.S. Pat. No. 4,443,731 (1984) to Butler et al. entitled Hybrid Piezoelectric and

Magnetostrictive Acoustic Wave Transducer, and U.S. Pat. No. 4,552,071 (1985) to R. B. Thompson entitled Method and Apparatus for Measuring Stress.

The reference to Peek cited above, takes advantage of the difference in operation of piezoelectric and magnetostrictive crystals to produce a response in one when stimulated by the other. The Peek patent does not use an acoustic wave to stimulate a permanent magnet as in the present invention.

The reference to Sommers cited above, is a transducer which uses a conductive bar or tube, which supports relatively slow helicon waves, placed next to a piezoelectric or magnetostrictive crystal. The transducer is designed in such a way as to either enhance the acoustic wave or the electric wave by interaction of the two materials. The Sommers patent does not use an acoustic wave to stimulate a permanent magnet to enhance to oscillation of the magnetic field as the present invention does.

The reference to Balmuth cited above, uses mechanically resonant reeds, rods, or chambers which are coupled to transducers that are piezoelectric, magnetostrictive, or transistorised. The electrical output of the transducers stimulates an electrical circuit when the resonator receives acoustic energy and again does not use an acoustic wave to stimulate a permanent magnet to enhance to oscillation of the magnetic field as the present invention does.

The reference to Olson cited above, uses an acoustically responsive material such as a piezoelectric or a magnetostrictive to act as a delay line for microwave signals and again does not use an acoustic wave to stimulate a permanent magnet to enhance to oscillation of the magnetic field as the present invention does.

The references to Benson, Quinn, Grisdale, Scott, and Butler cited above, are all concerned with acoustic transducers which convert acoustic pressure to an electrical signal or vice versa using only the piezoelectric and/or the magnetostrictive effect. The Benson patent is an underwater acoustic transformer which converts acoustic waves hitting a transducer into an electromagnetic field which excites a transformer. The Quinn patent uses a magnetostrictive effect to stimulate piezoelectric crystals to output a high voltage which is a reverse of the Benson patent. The Grisdale patent uses mechanically stacked piezoelectric or magnetostrictive crystals to produce a more efficient mechanical gyrator. The Scott patent uses and electrical oscillator to stimulate magnetostrictive rods which put pressure on piezoelectric crystals to output a high voltage from the piezoelectric crystals. The Butler patent uses a combined effect of piezoelectric and magnetostrictive crystals to produce an enhanced acoustic energy detector.

The reference to Thompson cited above, uses a permanent magnetic transducer to induce eddy currents in metal which is in the field of the transducer or uses moving eddy currents in a piece of metal to stimulate a magnetic field. The induction of the eddy currents is the result of an oscillating magnetic field generated in the transducer.

None of the references cited above, use an acoustic wave to stimulate the atoms of a permanent magnet and hence are not related to this invention.

SUMMARY OF THE INVENTION

An object of this invention is to provide a power generator with no moving parts.

Another object of this invention is to use an acoustic field to stimulate the nuclear level of the magnetic material and provide a method of oscillating the magnetic field of permanent magnets.

Another object of this invention is to provide a simple method of generating electrical energy by including a piezoelectric transducer which is used to vibrate the magnetic field of a permanent magnet. When the nucleus of the atom is vibrated by the piezoelectric, it in turn, vibrates the electronic structure of all the atoms. Since the electronic structure is the basis of the magnetic field of the magnet then the entire magnetic field of the magnet is vibrated when the electronic structure is vibrated. Coils placed in the vibrating magnetic field will have voltage and current induced in them.

It is a well established fact, that when the magnetic field of a permanent magnet is vibrated, it is possible to generate an alternating current in a coil winding placed within the vibrating magnetic field. What is unique about this invention, is to increase the efficiency of permanent magnet transformers by using acoustic stimulation from piezoelectrics to further stimulate the permanent magnet so as to add to the inductive effects of permanent magnet transformers. This invention does this by stimulating the permanent magnet cores of permanent magnet transformers with an acoustic field generated by a piezoelectric or other acoustically active generator which is vibrated at the same frequency as the electrical induction of the permanent magnet transformers.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG.1 illustrates a frequency signal generator attached to and driving a piezoelectric transducer which is in the acoustic proximity of a bar type of permanent magnet with a output coil placed within the magnetic field of the permanent magnet.

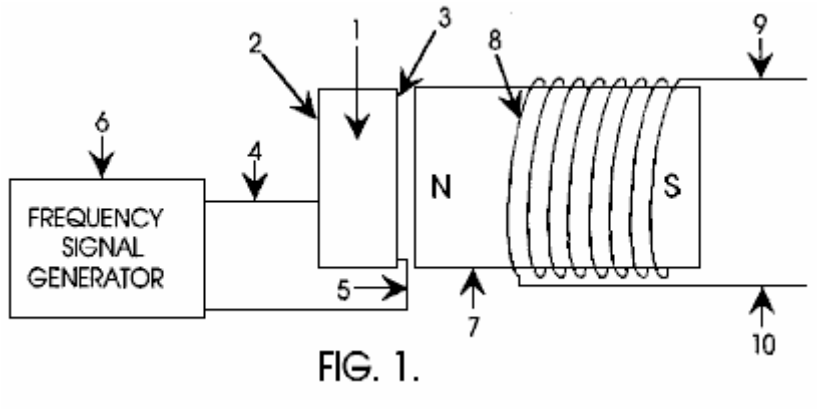


FIG. 1.

Fig.2 illustrates a frequency signal generator attached to and driving a piezoelectric transducer which is in the acoustic proximity of a torroidal type of permanent magnet with an output coil wrapped around the torroidal permanent magnet.

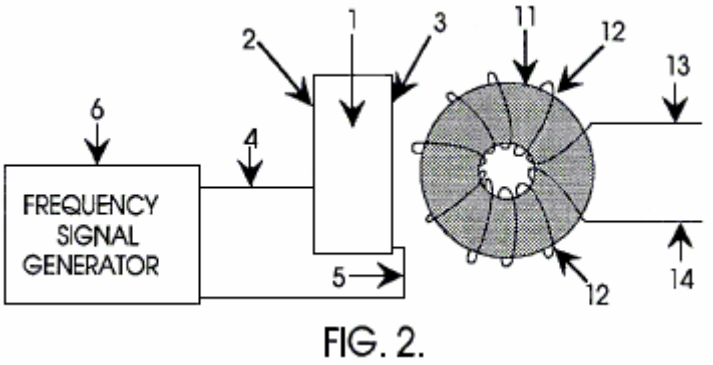


FIG. 2.

Fig.3 illustrates a frequency signal generator attached to and driving a piezoelectric transducer which is in the acoustic proximity of a torroidal type of permanent magnet transformer and the signal generator is also driving the input coil of the torroidal permanent magnet transformer.

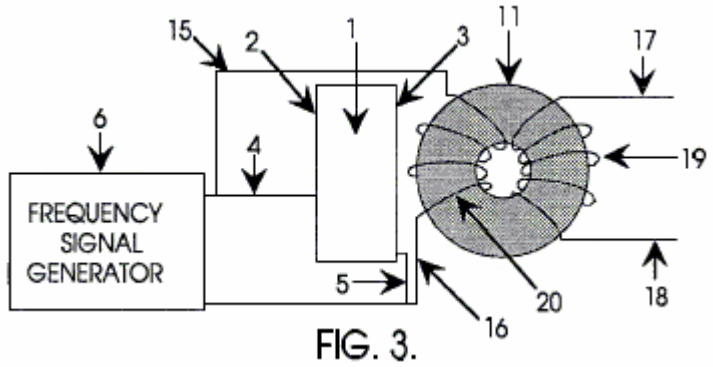


FIG. 3.

Fig.4 illustrates a frequency signal generator attached to and driving two torroidal core permanent magnet transformers as well as an acoustic transducer that is in acoustic proximity of the torroidal cores.

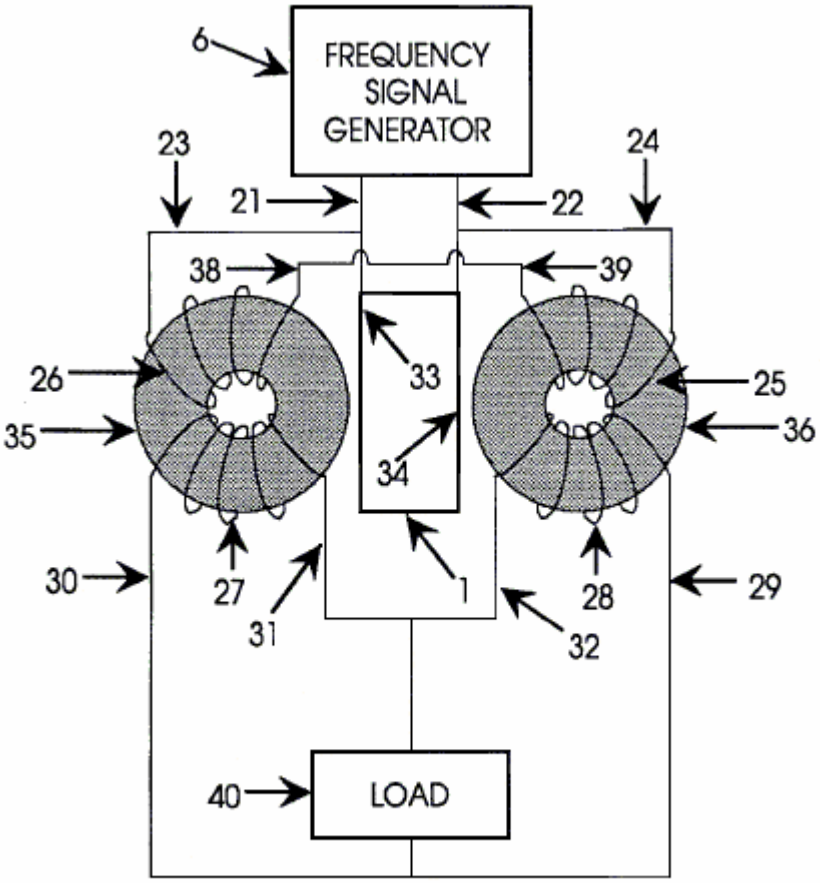

6
FREQUENCY SIGNAL GENERATOR
23
24
21
22
38
39
33
26
25
35
36
34
30
27
1
28
29
31
32
40
LOAD

FIG. 4.

DETAILED DESCRIPTION OF THE INVENTION

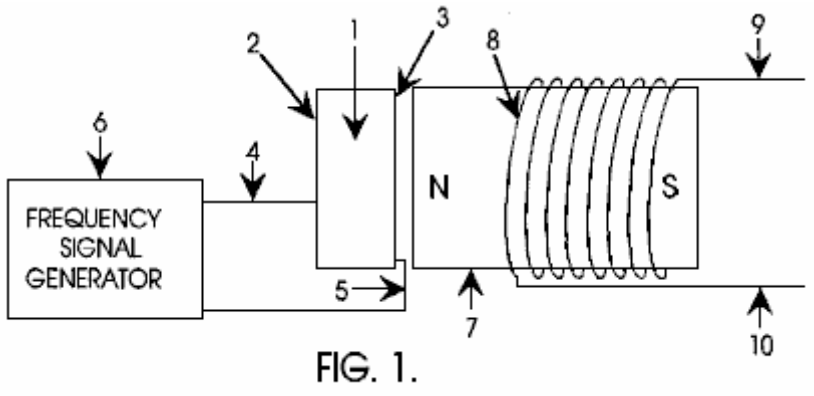


FIG. 1.

In **Fig.1**, a frequency signal generator **6** is connected to a piezoelectric transducer **1** via wires **4** and **5** connected to the electrode surfaces of the piezoelectric transducer **2** and **3** respectively. The piezoelectric transducer **1** is made from a high dielectric material such as barium titanate or lead zirconate titanate or any other acoustic transducer material suitable for sonic and ultrasonic generators. The piezoelectric transducer **1** is placed in close proximity to the permanent magnet **7** such that the acoustic field of the piezoelectric transducer **1** can radiate into the permanent magnet material. A permanent magnet transformer shown as coil **8** is positioned in the magnetic field of the permanent magnet **7**. When the piezoelectric transducer **1** is stimulated by the frequency generator **6** then a voltage and current is generated between the output leads **9** and **10** of the permanent magnet transformer.

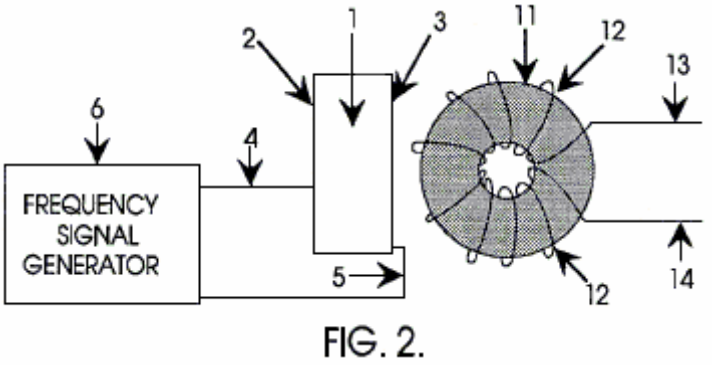


FIG. 2.

Another embodiment of this invention is shown in **Fig.2**. which is similar to **Fig.1**, with a similar frequency signal generator **6** connected to a piezoelectric material **1** via wires **4** and **5** connected to the electrode surfaces of the piezoelectric transducer **2** and **3**. The piezoelectric transducer **1** is as defined above, that is to say that it is constructed from a material suitable for sonic and ultrasonic generators. The piezoelectric transducer **1** is placed in close proximity to the permanent magnet **11** so that the acoustic field of the piezoelectric transducer **1** can radiate into the permanent magnet material. A permanent magnet transformer shown as coil **12** is placed in the magnetic field of the permanent magnet **11**. When the piezoelectric transducer **1** is stimulated by the frequency generator **6** then a voltage and current is generated between the output leads **13** and **14** of the above defined magnetic transformer.

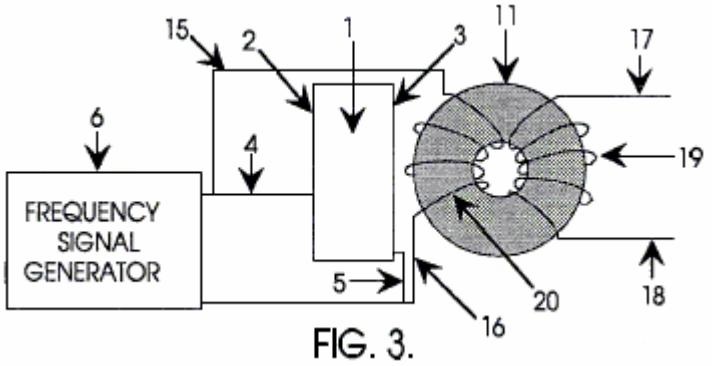


FIG. 3.

Fig.3 is similar to **Fig.1** and **Fig.2** with a frequency signal generator **6** connected to a piezoelectric transducer **1** via wires **4** and **5** connected to the electrode surfaces **2** and **3** of the piezoelectric transducer. The piezoelectric transducer **1** is as defined in the descriptions above. The signal generator **6** is also connected to the input coil **20**

of the permanent magnet transformer defined by the torroidal permanent magnet core **11**, input coil **20** and output coil **19**. The piezoelectric transducer **1** is placed in close proximity to the permanent magnet **11** so that the acoustic field of the piezoelectric transducer **1** can radiate into the permanent magnet material. The magnetic transformer defined by **11**, **19**, and **20** is in the magnetic field of the permanent magnet **11** and is connected to the frequency signal generator **6** via wires **15** and **16**. The frequency generator **6** stimulates the piezoelectric transducer **1** which stimulates the permanent magnet transformer via the acoustic field and at the same time the signal generator also stimulates the coil electromagnetically. A voltage and current is generated at the output coil **19** and power can be taken from the output wires **17** and **18** of the magnetic transformer.

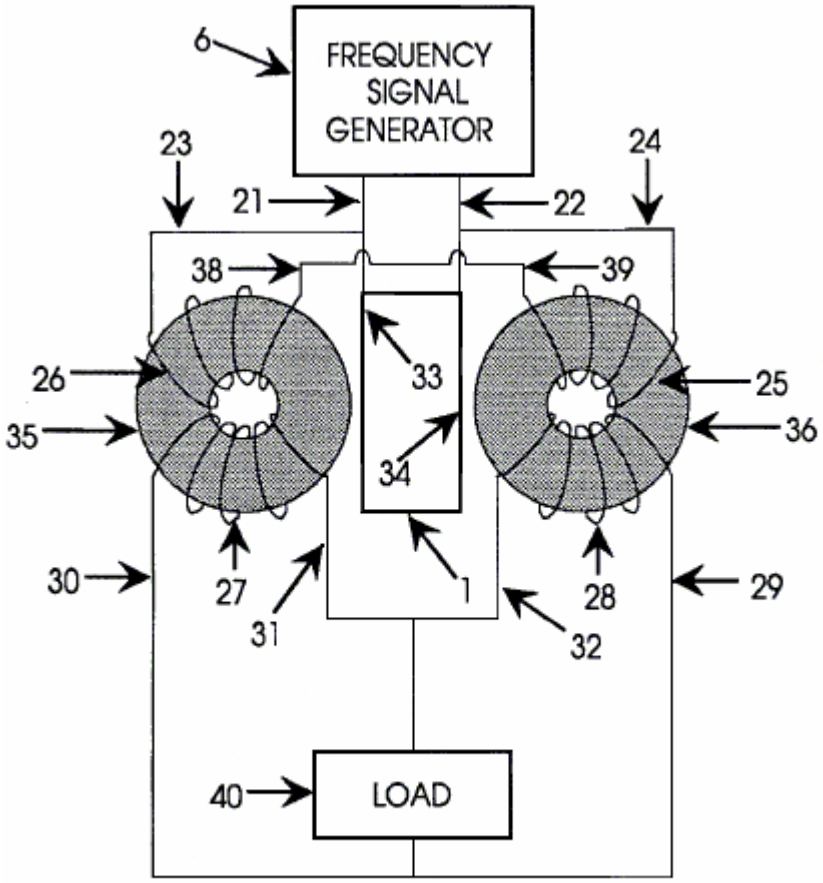


FIG. 4.

A further embodiment of this invention, shown in **Fig.4**, is a frequency signal generator **6** driving a pair of permanent magnet transformers defined by **26, 35, 27** and **25, 36, 28** respectively, also driving a piezoelectric transducer **1**. The piezoelectric transducer is as described above. The signal generator is connected via input wires **23** and **24** to the input coil **26** of the permanent magnet transformer on the left and to the input coil **25** of the transformer on the right respectively. The other input wire **38** of the left permanent magnet transformer is connected to the remaining input wire **39** of the right magnetic transformer. The output of the signal generator in also connected to the piezoelectric transducer **1** via connections **21** and **22** to the connector surface of the piezoelectric **33** and **34** respectively. The output of the permanent magnet transformer on the left is connected to a load **40** via wire **30** and the output of the permanent magnet transformer on the right is connected to the load via wire **29**. The remaining output wires **31** and **32** of the left and right permanent magnet transformers are also connected to the load. The load **40** can be anything such as a motor or electrical lights or any appliance.

This invention is not limited to the 4 different versions of the invention shown in **Figs. 1, 2, 3**, and **4** as there are any number of cascading and electrical hook-up techniques that can be accomplished to amplify power and to take advantage of the acoustic influence of the piezoelectric upon the magnetic material. Similarly, this invention is not limited to the torroidal core configuration as there can be many types of permanent magnet transformers with any number of magnetic core and coil configurations that can be enhanced with acoustic stimulation depending on power and output requirements according to the rules of electronics and those familiar with the state of the art in permanent magnet power transformers.

CLAIMS

1. An acoustic magnetic power generator composed of an alternating current signal generator connected to an acoustic transducer which stimulates the core of a permanent magnet such that the atoms of the magnet are caused to vibrate which in turn causes the magnetic field to vibrate and causes a current and voltage to be generated in an output coil wrapped around a permanent magnet or in the magnetic field of the permanent magnet which said current and voltage can be used for powering a load.

2. An acoustic magnetic power generator composed of an alternating signal generator connected to an acoustic transducer which stimulates the core of a permanent magnet and causes the core to vibrate; the signal generator further connected to a drive coil surrounding the permanent magnet, and an output coil within the field of the permanent magnet which by induction generates an electrical output.

3. A method of causing the magnetic field of permanent magnet transformers to oscillate by the application of an acoustic signal applied to the atomic structure of permanent magnet.

4. A method of maximising the efficiency of permanent magnet transformers by stimulating the core material of the permanent magnet transformers with both an acoustic vibration and an electromagnetic signal simultaneously.

JOHN BEDINI: BATTERY CHARGER

United States Patent 6,545,444 8th April 2003 Inventor: John C. Bedini

DEVICE AND METHOD FOR UTILISING A MONOPOLE MOTOR TO CREATE BACK-EMF TO CHARGE BATTERIES

Please note that this is a re-worded excerpt from this patent. It describes a self-contained device which can charge an external battery or battery bank.

ABSTRACT

A back EMF monopole motor and method using a rotor containing magnets all of the same polarity and in a monopole condition when in momentary apposition with a magnetised pole piece of a stator having the same polarity, the stator being comprised of a coil with three windings: a power-coil winding, a trigger-coil winding, and a recovery-coil winding. The back EMF energy is rectified using a high voltage bridge, which transfers the back EMF energy to a high voltage capacitor for storage in a recovery battery. The stored energy can then be discharged across the recovery battery through the means of a contact rotor switch for further storage.

DESCRIPTION

Technical Field:

The invention relates generally to the capturing of available electromagnetic energy using a device and method for creating an electromagnetic force ('EMF') and then using the available stored energy for recycling into the system as stored energy. The method of creating back EMF is the result of coupling/uncoupling a coil to a voltage source.

Background:

The operation of present day normal magnetic motors, has the rotor pole attracting the stator pole, resulting in the generation of mechanical power from the magnets to the rotor and flywheel. During this phase, energy flows from the magnetics to the rotor/flywheel and is stored as kinetic energy in the increased rotation. A rotor pole leaving a stator pole and creating a condition of "drag" results in power having to be put back into the magnetic section by the rotor and flywheel to forcibly overcome the drag. In a perfect, friction-free motor, the net force field is therefore referred to as "most conservative". A most conservative EMF motor has maximum efficiency. Without extra energy continually fed to the motor, no net work can be done by the magnetic field, since half the time the magnetic field adds energy to the load (the rotor and flywheel) and the other half of the time it subtracts energy from the load (the rotor and flywheel). Therefore, the total net energy output is zero in any such rotary process without additional energy input. To use a present day magnetic motor, continuous energy must be fed to the motor to overcome drag and to power the motor and its load.

Motors and generators presently in use, all use such conservative fields and therefore, have internal losses. Hence, it is necessary to continually input all of the energy that the motor outputs to the load, plus more energy to cover losses inside the motor itself. EMF motors are rated for efficiency and performance by how much energy "input" into the motor actually results in "output" energy to the load. Normally, the Coefficient of Performance ('COP') rating is used as a measure of efficiency. The COP is the actual output energy going into the load and powering it, divided by the energy that must be input into the device with its motor/load combination. If there were zero internal losses in a motor, that "perfect" motor would have a COP equal to 1.0. That is, all energy input into the motor would be output by the motor directly into the load, and none of the input energy would be lost or dissipated in the motor itself.

In magnetic motor generators presently in use, however, due to friction and design flaws, there are always internal losses and inefficiencies. Some of the energy input into the motor is dissipated in these internal losses. As a consequence, the energy that gets to the load is always less than the input energy. So a standard motor operates with a COP of less than 1.0, which is expressed as COP<1.0. An inefficient motor may have a COP of 0.4 or 0.45, while a specially designed and highly efficient motor may have a COP of 0.85.

The conservative field inside of a motor itself is divided into two phases. Producing a conservative field involves net symmetry between the "power out" phase from the magnetics to the rotor/flywheel and the "power back in" phase from the rotor/flywheel back to the magnetics. That is, the two flows of energy are identical in magnitude but opposite in direction. Each phase alone is said to be "asymmetrical", that is, it either has: 1) a net energy flow out to the rotor/flywheel; or 2) a net energy flow back into the magnetics from the rotor/flywheel. In simplified terms, it is referred to as "power out" and "power back in" phases with respect to the motor magnetics.

For the power-out phase, energy is derived from the EMF existing between the stator pole and incoming rotor pole in an attraction mode. In this phase, the rotary motion (angular momentum and kinetic energy) of the rotor and flywheel is increased. In short, power is added to the rotor/flywheel (and thus to the load) from the fields between stator pole and rotor pole (the electromagnetic aspects of the system).

For the "power back in" phase, energy must be fed back into the magnetics from the rotor and flywheel (and the load) to overcome the drag forces existing between stator pole and outgoing rotor pole. In this phase, energy is returned to the internal magnetic system from the rotary motion of the rotor and flywheel (the angular momentum, which is the rotational energy multiplied by time). As is well known in physics, a rotor/flywheel's angular momentum provides a convenient way to store energy with the spinning rotor/flywheel mass acting as an energy reservoir.

Most present day conventional magnetic motors use various methods for overcoming and partially reversing back EMF. Back EMF may be defined as the return pulse from the coil out of phase and is the result of re-gauging, which is the process of reversing the magnetics polarity, that is, form North to South, etc. The back EMF is shorted out and the rotor is attracted back in, therefore eliminating drag. This can be accomplished by pouring more energy in, which overpowers the back EMF, thereby producing a forward EMF in that region. The energy required for this method is furnished by the operator.

It is well known that changing the voltage alone creates a back EMF and requires no work. This is because to change the potential energy does not require changing the form of that potential energy, but only its magnitude. Work is the changing of the form of energy. Therefore, as long as the form of the potential energy is not changed, the magnitude can be changed without having to perform work in the process. The motor of the present invention takes advantage of this permissible operation to create back EMF asymmetrically, and thereby change its own usable available potential energy.

In an electric power system, the potential (voltage) is changed by inputting energy to do work on the internal charges of the generator or battery. This potential energy is expended within the generator (or battery) to force the internal charges apart, forming a source dipole. Then the external closed circuit system connected to that source dipole ineptly pumps the spent electrons in the ground line back through the back EMF of the source dipole, thereby scattering the charges and killing the dipole. This shuts off the energy flow from the source dipole to the external circuit. As a consequence of this conventional method, it is a requirement to input and replace additional energy to again restore the dipole. The circuits currently utilised in most electrical generators have been designed to keep on destroying the energy flow by continually scattering all of the dipole charges and terminating the dipole. Therefore, it is necessary to keep on inputting energy to the generator to keep restoring its source dipole.

A search of prior art failed to reveal any monopole motor devices and methods that recycle available energy from back EMF to charge a battery or provide electrical energy for other uses as described in the present invention. However, the following prior art patents were reviewed:
U.S. Pat. No. 4,055,789 to Lasater, Battery Operated Motor with Back EMF Charging.
U.S. Pat. No. 2,279,690 to Z. T. Lindsey, Combination Motor Generator.

SUMMARY OF THE INVENTION

An aspect of the device and method of the present invention is a new monopole electromagnetic motor that captures back EMF energy. The captured back EMF energy may be used to charge or store electrical energy in a recovery battery. The amount of energy recoverable, as expressed in watts, is dependent upon the configuration, circuitry, switching elements and the number and size of stators, rotors, magnets and coils which comprise the motor.

The motor uses a small amount of energy from a primary battery to "trigger" a larger input of available energy by supplying back EMF, thus increasing the potential energy of the system. The system then utilises this available potential energy to reduce, or reverse, the back EMF, thereby increasing the efficiency of the motor and, therefore, the COP.

If the energy in phase 1 (the power-out phase) is increased by additional available energy in the electromagnetics themselves, then the energy in phase 1 can be made greater than the energy in phase 2 (the power-back-in phase) without the operator furnishing the energy utilised. This produces a non-conservative nett field. Nett power can then be taken from the rotating stator and flywheel, because the available energy added into the stator and flywheel by the additional effects, is transformed by the rotor/flywheel into excess angular momentum and stored as such. Angular momentum is conserved at all times, but now, some of the angular momentum added to the flywheel, is evoked by additional effects in the electromagnetics, rather than being furnished by the operator.

That is, the motor is designed to deliberately create a back EMF itself, and thus increase its potential energy, thereby retaining each extra force for a period of time and applying it to increase the angular momentum and kinetic energy of the rotor and flywheel. Specifically, this back EMF energy with its nett force is deliberately applied in the motor of the present invention to overcome and even reverse the conventional drag-back (the back EMF). Hence, less energy needs to be taken from the rotor and flywheel to overcome the reduced back EMF, and in the ideal case, none is required since the back EMF has been overpowered and converted to forward EMF by the back EMF energy and force. In the motor, the conventional drag section of the magnetics becomes a forward-EMF section and now adds energy to the rotor/flywheel instead of reducing it. The important feature is that the operator only pays for the small amount of energy necessary to trigger the back EMF from the primary battery, and does not have to furnish the much larger back EMF energy itself.

Thus, when the desired energy in phase 1 (the power-out phase) is made greater than the undesired drag energy in phase 2, then part of the output power normally taken from the rotor and flywheel by the fields in phase 2, is not required. Hence, in comparison to a system without special back EMF mechanisms, additional power is available from the rotor/flywheel. The rotor therefore maintains additional angular momentum and kinetic energy, compared to a system which does not produce back EMF itself. Consequently, the excess angular momentum retained by the rotor and flywheel can be utilised as additional shaft power to power an external load.

In this motor, several known processes and methods are utilised. These allow the motor to operate periodically as an open dissipative system (receiving available excess energy from back EMF) far from thermodynamic equilibrium, whereby it produces and receives its excess energy from a known external source.

A method is utilised to temporarily produce a much larger source of available external energy around an energised coil. Design features of this new motor provide a device and method that can immediately produce a second increase in that energy concurrently as the energy flow is reversed. Therefore, the motor is capable of producing two asymmetrical back EMFs, one after the other, of the energy within a single coil, which dramatically increases the energy available and causes that available excess energy to then enter the circuit as impulses which are collected and utilised.

The motor utilises this available excess back EMF energy to overcome and even reverse the drag EMF between stator pole and rotor pole, while furnishing only a small trigger pulse of energy from a primary battery necessary to control and activate the direction of the back EMF energy flow.

By using a number of such dual asymmetrical self back EMFs for every revolution of the rotor, the rotor and flywheel collectively focus all the excess impulsive inputs into increased angular momentum (expressed as energy multiplied by time), shaft torque, and shaft power.

Further, some of the excess energy deliberately generated in the coil by the utilisation of the dual process manifests in the form of excess electrical energy in the circuit and can be utilised to charge a recovery battery or batteries. The excess energy can also be used to power electrical loads or to power the rotor and flywheel, with the rotor/flywheel also furnishing shaft horsepower for powering mechanical loads.

The motor utilises a means to furnish the relatively small amount of energy from a primary battery to initiate the impulsive asymmetrical self back EMF actions. Then part of the available excess electrical power drawn off from back EMF created energy is utilised to charge a recovery battery with dramatically increased over-voltage pulses.

Design features of this monopole motor utilise one magnetic pole of each rotor and stator magnet. The number of impulsive self-back EMF in a single rotation of the rotor is doubled. Advanced designs can increase the number of self-back EMFs in a single rotor rotation with the result that there is an increase in the number of impulses per rotation, which increase the power output of this new motor.

The sharp voltage spike produced in the coil of this monopole motor by the rapidly collapsing field in the back EMF coil is connected to a recovery battery(s) in charge mode and to an external electrical load. The nett result is that the coil asymmetrically creates back EMF itself in a manner which adds available energy and impulse to the circuit. The available energy collected in the coil is used to reverse the back-EMF phase of the stator-rotor fields to a forward EMF condition, with the impulses adding acceleration and angular momentum to the rotor and flywheel. The available back EMF energy collected in the coil is used to charge a battery. Loads can then be driven by the battery.

A device and method in which the monopole motor alters the reaction cross section of the coils in the circuit, which briefly changes the reaction cross section of the coil in which it is invoked. Thus, since this new motor uses only a small amount of current in the form of a triggering pulse, it is able to evoke and control the immediate change of the coil's reaction cross section to this normally wasted energy-flow component. As a result, the motor captures and directs some of this usually wasted available environmental energy, collecting the available excess energy in the coil and then releasing it for use in the motor. Through timing and switching, the innovative gate

design of this new motor directs the available excess energy so that it overcomes and reverses the return EMF of the rotor-stator pole combination during what would normally be the back EMF and demonstrates the creation of the second back EMF of the system. Now, instead of an "equal retardation" force being produced in the back EMF region, a forward EMF is produced which adds to the rotor/flywheel energy, rather than subtracting from it. In short, it further accelerates the rotor/flywheel.

This results in a non-conservative magnetic field along the rotor's path. The line integral of the field around that path (i.e., the net work on the rotor/flywheel to increase its energy and angular momentum) is not zero but a significant amount. Hence, the creation of an asymmetrical back EMF impulse magnetic motor:
1) Takes its available excess energy from a known external source, the huge usually non-intercepted portion of the energy flow around the coil;
2) Further increases the source dipolarity by this back EMF energy; and
3) Produces available excess energy flow directly from the source dipole's increased broken symmetry in its fierce energy exchange with the local vacuum.

By operating as an open dissipative system, not in thermodynamic equilibrium with the active vacuum, the system can permissibly receive available energy from a known environmental source and then output this energy to a load. As an open dissipative system not in thermodynamic equilibrium, this new and unique monopole motor can tap in on back EMF to energise itself, loads and losses simultaneously, fully complying with known laws of physics and thermodynamics.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig.1 is a perspective side view of a monopole back EMF motor with a single stator and a single rotor.

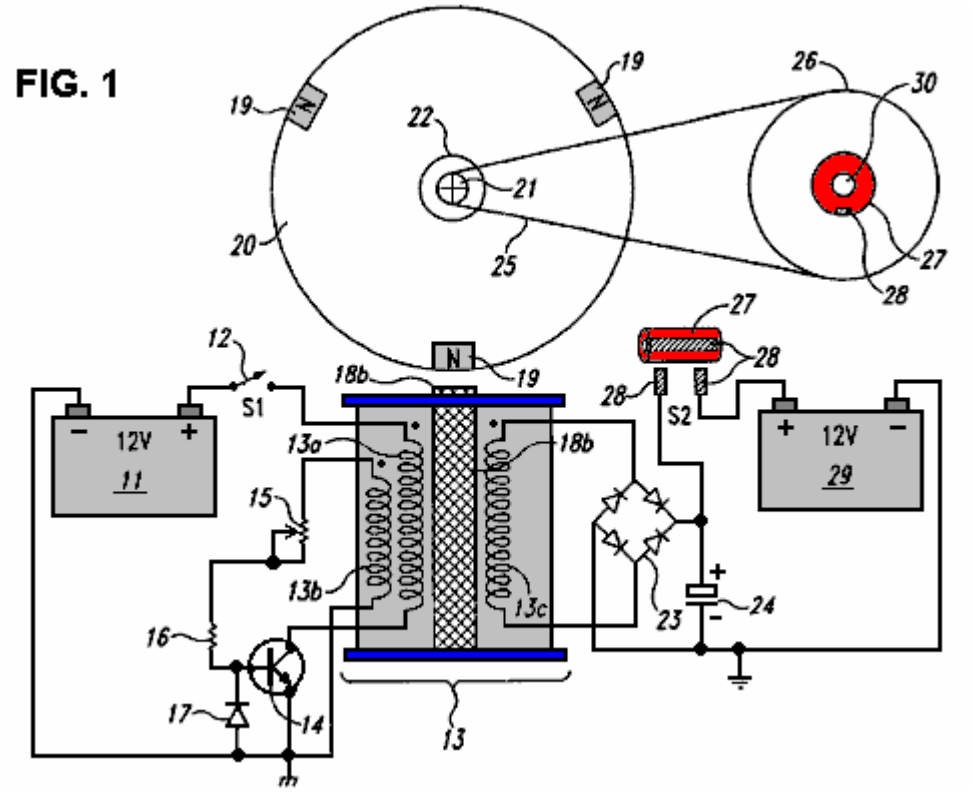


Fig.2 is a perspective top view of a monopole back EMF motor with a single stator and a single rotor.

FIG. 2

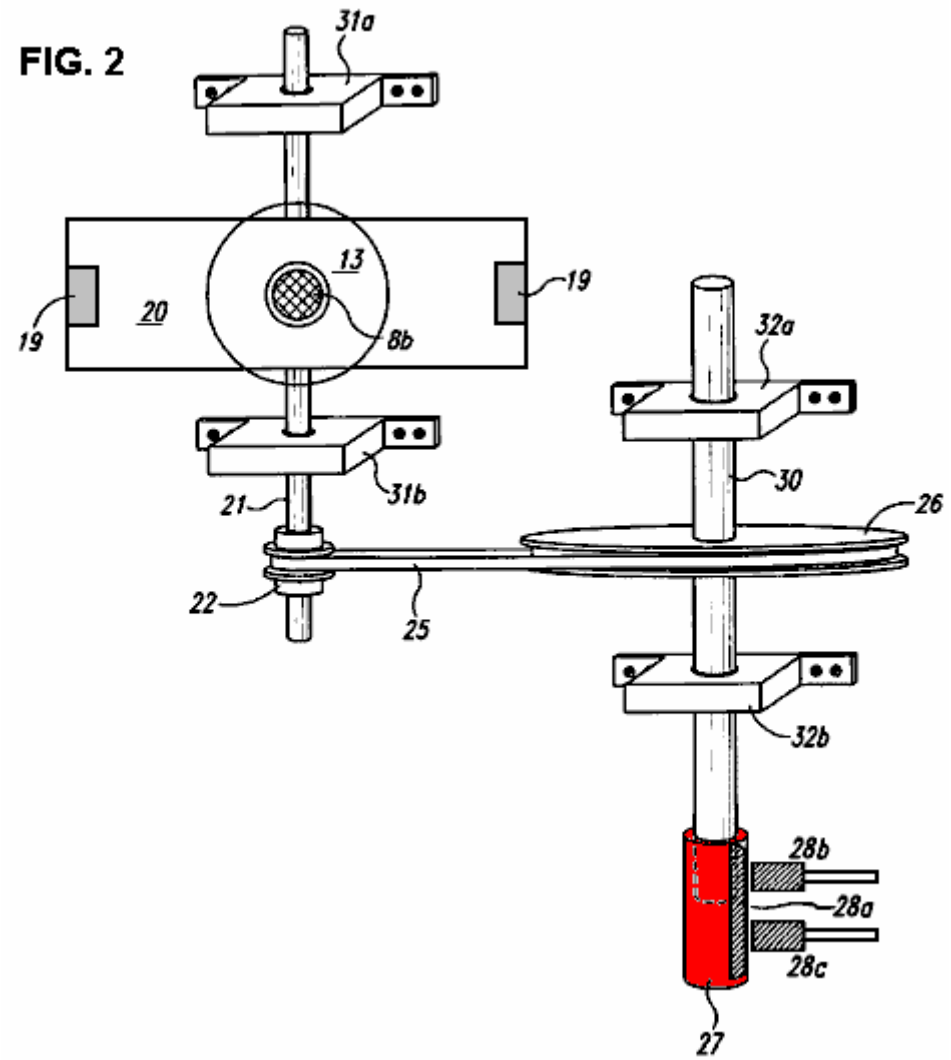
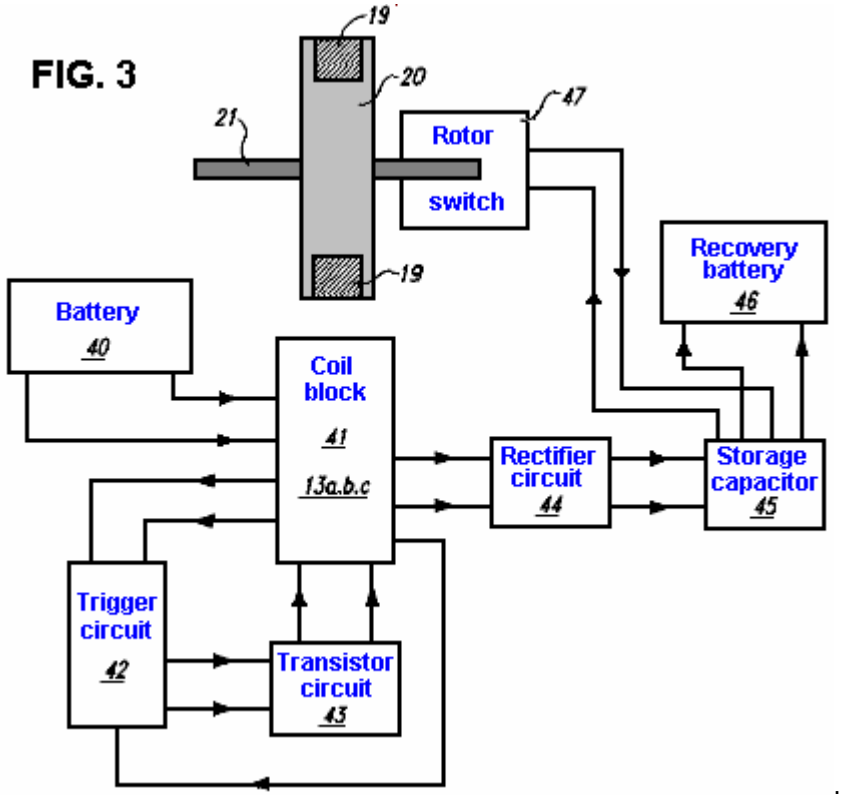


Fig.3 is a block diagram demonstrating the circuitry for a monopole back EMF motor.



DETAILED DESCRIPTION OF THE INVENTION

An embodiment of the present invention is a device and method for a monopole back EMF electromagnetic motor. As described in the Summary of the Invention, this monopole motor conforms to all applicable electrodynamic laws of physics and is in harmony with the law of the conservation of energy, the laws of electromagnetism and other related natural laws of physics.

The monopole back EMF electromagnetic motor comprises a combination of elements and circuitry to capture available energy (back EMF) in a recovery element, such as a capacitor, from output coils. The available stored energy in the recovery element is used to charge a recovery battery.

As a starting point, an arbitrary method in describing this device will be employed, namely, the flow of electrical energy and mechanical forces will be tracked from the energy's inception at the primary battery to its final storage in the recovery battery.

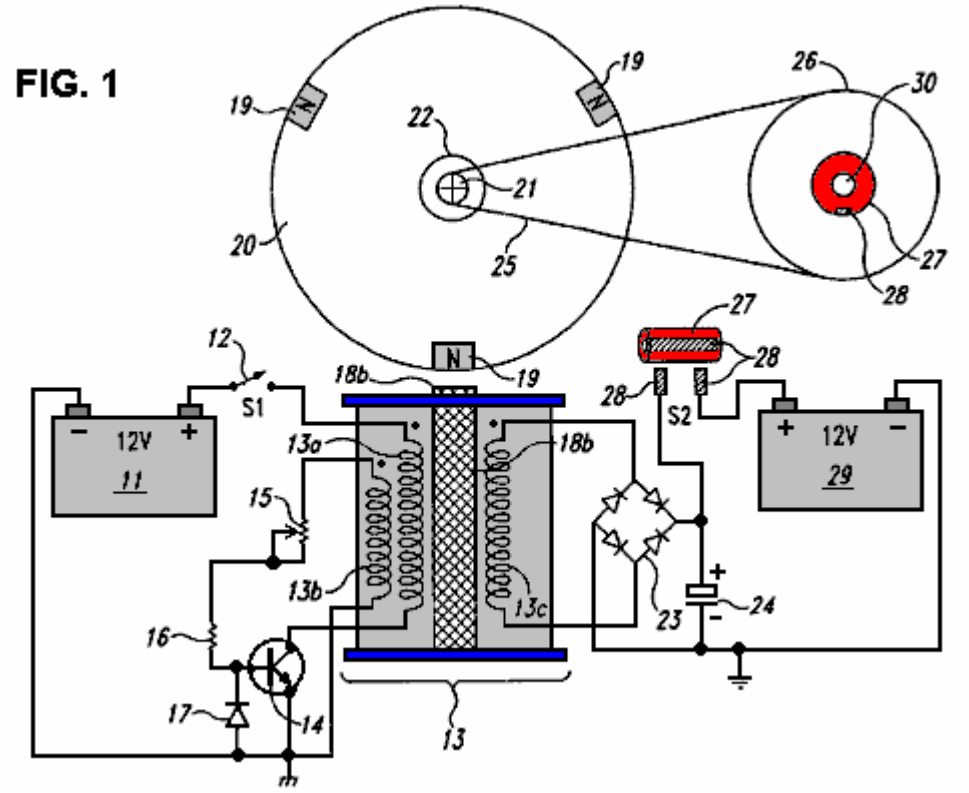


Fig.1 is a perspective side view of the monopole motor according to an embodiment of the invention. As shown in **Fig.1**, electrical energy from primary battery **11** periodically flows through power switch **12** and through power-coil wiring **13a**. In one embodiment, power switch **12** is merely an On-Off mechanical switch and is not electronic. However, the switch **12** may be a solid-state switching circuit, a magnetic Reed switch, a commutator, an optical switch, a Hall-effect switch, or any other conventional transistorised or mechanical switch. Coil **13** is comprised of three windings: power-coil winding **13a**, trigger-coil winding **13b**, and recovery-coil winding **13c**. However, the number of windings can be more or fewer than three, depending upon the size of the coil **13**, size of the motor and the amount of available energy to be captured, stored and used, as measured in watts. Electrical energy then periodically flows from power-coil winding **13a** and through transistor **14**.

Trigger energy also periodically flows through variable resistor **15** and resistor **16**. Clamping diode **17** clamps the reverse base-emitter voltage of transistor switch **14** at a safe reverse-bias level that does not damage the transistor. Energy flows to stator **18a** and pole piece **18b**, an extension of stator **18a**. Pole piece **18b** is electrically magnetised only when transistor switch **14** is on and maintains the same polarity as the rotor poles **19** - North pole in this instance - when electrically magnetised. The North rotor poles **19a**, **19b** and **19c**, which are attached to rotor **20**, come in momentary apposition with pole piece **18b** creating a momentary monopole interface. The poles **19a,b,c**, which are actually permanent magnets with their North poles facing outward from the rotor **20**, maintain the same polarity when in momentary alignment with pole piece **18b**.

Rotor **20** is attached to rotor shaft **21**, which has drive pulley **22**. Attached to rotor shaft **21** are rotor-shaft bearing blocks **31a** and **31b**, as seen in **Fig.2**. As rotor **20** begins to rotate, the poles **19a,b,c** respectively comes into alignment with magnetised pole piece **18b** in a momentary monopole interface with energy flowing through diode bridge rectifier **23** and capacitor **24**. The number of capacitors may be of a wide range, depending upon the amount of energy to be temporarily stored before being expelled or flash charged into recovery battery **29**. Timing belt **25** connects drive pulley **22** on timing shaft **21** to timing wheel **26**. Attached to timing wheel **26** is contact rotor **27**, a copper insulated switch that upon rotation, comes in contact with brushes on mechanical switch **28**. The means for counting the number of rotor revolutions may be a timing gear or a timing belt. Finally, the available energy derived from the back EMF that is stored in capacitor **24** is then discharged and stored in recovery battery **29**.

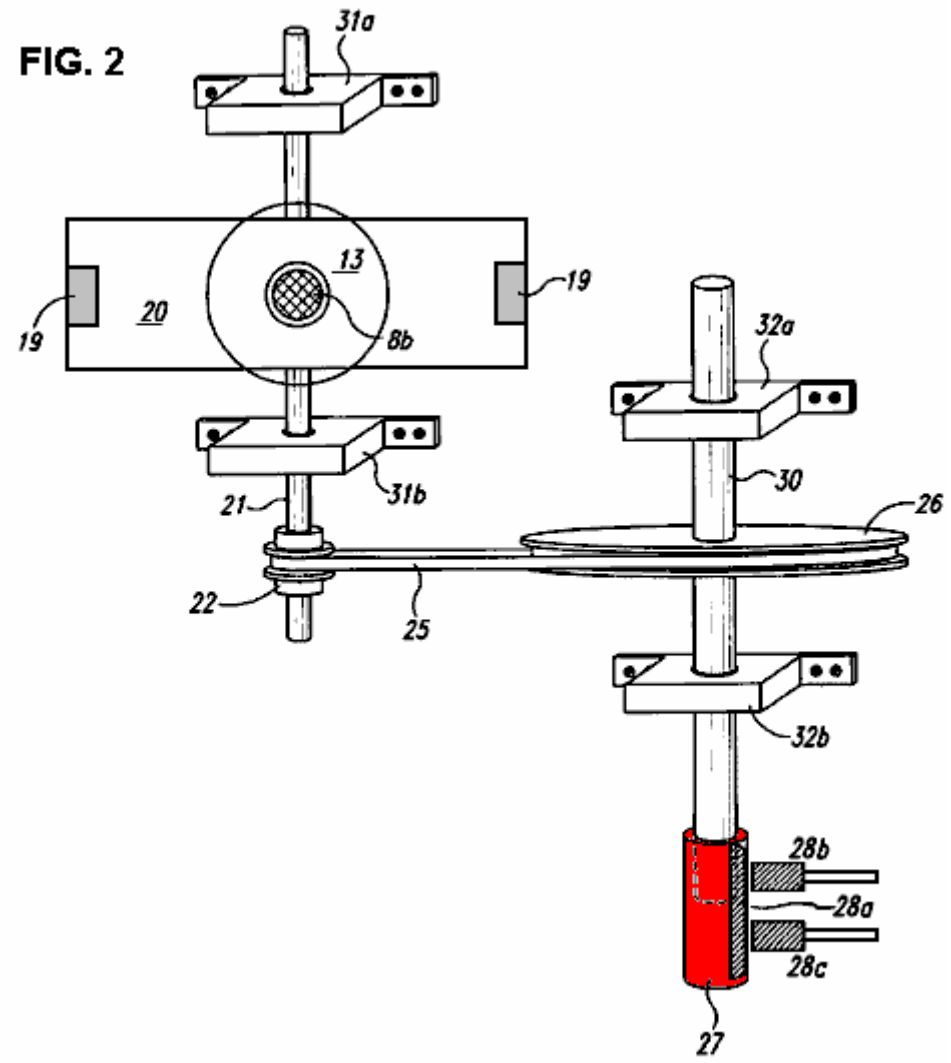
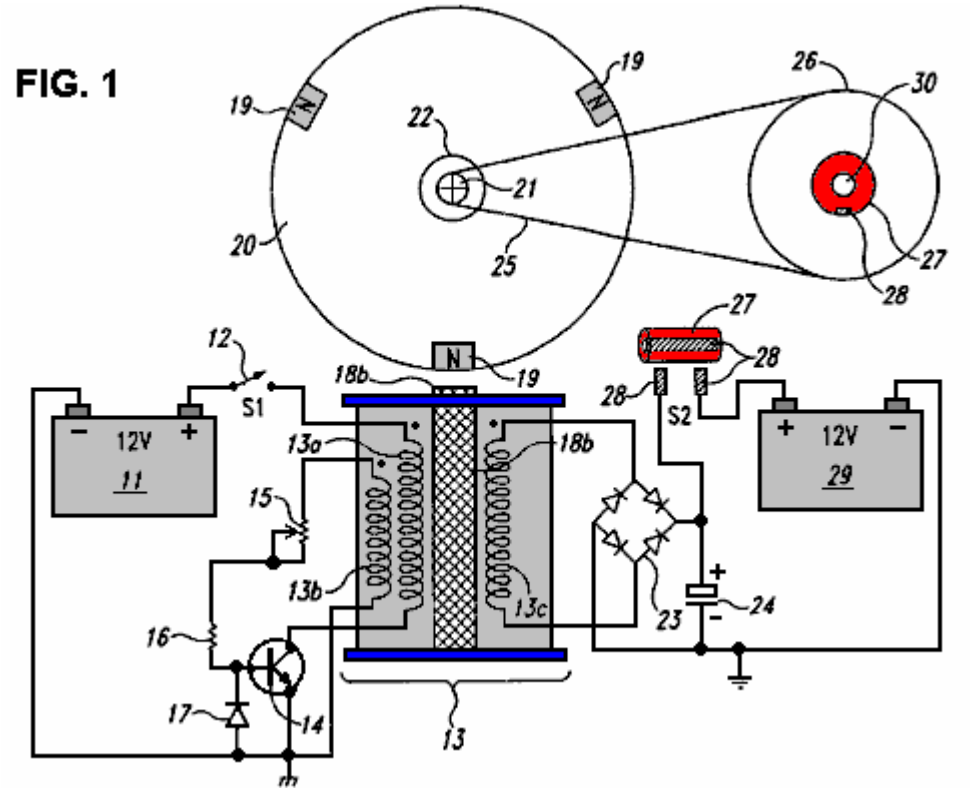


Fig.2 is a mechanical perspective top view of the monopole motor of the instant invention without electrical circuitry. Stator **18a** consists of coil **13**, which is comprised of three separate coil windings: power-coil winding **13a**, trigger-coil winding **13b** and recovery-coil winding **13c**. Pole piece **18b** is at the end of stator **18a**. As rotor **20**, (which is attached to rotor shaft **21**) rotates, each pole **19** respectively comes in a momentary monopole interface with pole piece **18b**. The polarity of pole piece **18b** is constant when electrically magnetised. Rotor shaft **21** has rotor shaft bearing blocks **31a,b** attached to it for stabilisation of rotor shaft **21**. Attached to rotor shaft **21** is drive pulley **22** with timing belt **25** engaged with it. Another means for timing may be a timing gear. Timing belt **25** engages with timing wheel **26** at its other end. Timing wheel **26** is attached to timing shaft **30**. Shaft **30** is stabilised with timing shaft bearing blocks **32a,b**. Attached to one end of timing shaft **30** is contact rotor **27** with brush **28a**, which, upon rotation of the timing shaft, comes into momentary contact with brushes **28b,c**.

Fig.3 is a block diagram detailing the circuitry of the monopole motor. Block **40** represents primary battery **11** with energy flowing to coil block **41**, which represents coil windings **13a,b,c**. From coil block **41** energy flows into three directions: to trigger-circuit block **42**, transistor-circuit block **43**, and rectifier-circuit block **44**. Energy flows from rectifier-block **44** to storage-capacitor block **45** with energy flowing from block **45** to both recovery-battery block **46** and rotor-switch block **47**.

Referring to **Fig.1**, the operation of the motor is described according to an embodiment of the invention. For purpose of explanation, assume that the rotor **20** is initially not moving, and one of the poles **19** is in the three o'clock position.



First, the switch **12** is closed. Because the transistor **14** is off, no current flows through the winding **13a**.

Next, the motor is started by rotating the rotor **20**, say, in a clockwise direction. The rotor may be rotated by hand, or by a conventional motor-starting device or circuit (not shown).

As the rotor **20** rotates, the pole **19** moves from the three o'clock position towards the pole piece **18b** and generates a magnetic flux in the windings **13a**,**13b** and**13c**. More specifically, the stator **18a** and the pole piece **18b** include a ferromagnetic material such as iron. Therefore, as the pole **19** moves nearer to the pole piece **18b**, it magnetises the pole piece **18b** to a polarity - South in this instance - that is opposite to the polarity of the pole **19** (which is North). This magnetisation of the pole piece **18b** generates a magnetic flux in the windings **13a-13c**. Furthermore, this magnetisation also causes a magnetic attraction between the pole **19** and the pole piece **18b**. This attraction pulls the pole **19** toward the pole piece **18b**, and thus reinforces the rotation of the rotor **20**.

The magnetic flux in the windings **13a-13c** generates voltages across their respective windings. More specifically, as the pole **19** rotates toward the pole piece **18b**, the magnetisation of the stator **18a** and the pole piece **18b**, and thus the magnetic flux in the windings **13a-13c**, increases. This increasing flux generates voltages across the windings **13a-13c** such that the dotted (top) end of each winding is more positive than the opposite end. These voltages are proportional to the rate at which the magnetic flux is increasing, and so, they are proportional to the velocity of the pole **19**.

At some point, the voltage across the winding **13b** becomes high enough to turn the transistor **14c** on. This turn-on, i.e., trigger, voltage depends on the combined serial resistance of the potentiometer **15** and the resistor **16**. The higher this combined resistance, the higher the trigger voltage, and vice-versa. Therefore, one can set the level of the trigger voltage by adjusting the potentiometer **15**.

In addition, depending on the level of voltage across the capacitor **24**, the voltage across the winding **13c** may be high enough to cause an energy recovery current to flow through the winding **13c**, the rectifier **23**, and the capacitor **24**. Thus, when the recovery current flows, the winding **13c** is converting magnetic energy from the rotating pole **19** into electrical energy, which is stored in the capacitor **24**.

Once turned on, the transistor **14** generates an opposing magnetic flux in the windings **13a-13c**. More specifically, the transistor **14** draws a current from the battery **11**, through the switch **12** and the winding **13b**. This current increases and generates an increasing magnetic flux that opposes the flux generated by the rotating pole **19**.

When the opposing magnetic flux exceeds the flux generated by the rotating pole **19**, the opposing flux reinforces the rotation of the rotor **20**. Specifically, when the opposing flux (which is generated by the increasing current

through winding **13a**) exceeds the flux generated by the pole **19**, the magnetisation of the pole piece **18** inverts to North pole. Therefore, the reverse-magnetic pole piece **18** repels the pole **19**, and thus imparts a rotating force to the rotor **20**. The pole piece **18** rotates the rotor **20** with maximum efficiency if the pole-piece magnetisation inverts to North when the centre of the pole **19** is aligned with the centre of the pole piece. Typically, the potentiometer **15** is adjusted to set the trigger voltage of the transistor **14** at a level which attains or approximates to this maximum efficiency.

The transistor **14** then turns off before the opposing flux can work against the rotation of the rotor **20**. Specifically, if the pole piece **18** remains magnetised to North pole, it will repel the next pole **19** in a direction (counterclockwise in this example) opposite to the rotational direction of the rotor **20**. Therefore, the motor turns transistor **14** off, and thus demagnetises the pole piece **18**, before this undesirable repulsion occurs. More specifically, when the opposing flux exceeds the flux generated by the pole **19**, the voltage across the winding **13b** reverses polarity such that the dotted end is less positive than the opposite end. The voltage across the winding **13b** decreases as the opposing flux increases. At some point, the voltage at the base of the transistor decreases to a level that turns transistor **14** off. This turn-off point depends on the combined resistance of potentiometer **15** and resistor **16** and the capacitance (not shown) at the transistor base. Therefore, potentiometer **15** can be adjusted, or other conventional techniques can be used to adjust the timing of this turn-off point.

The rectifier **23** and capacitor **24** recapture the energy that is released by the magnetic field (which energy would otherwise be lost) when the transistor **14** turns off. Specifically, turning transistor **14** off abruptly, cuts off the current flowing through winding **13a**. This generates voltage spikes across the windings **13a-13c** where the dotted ends are less positive than their respective opposite ends. These voltage spikes represent the energy released as the current-induced magnetisation of stator **18**a and pole piece **18b** collapses, and may have a magnitude of several hundred volts. But, as the voltage spike across the winding **13c** increases above the sum of the two diode drops of the rectifier **23**, it causes an energy-recovery current to flow through the rectifier **23** and the voltage across the capacitor **24** charge the capacitor **24**. Thus, a significant portion of the energy released upon collapse of the current-induced magnetic field is recaptured and stored as a voltage in the capacitor **24**. In addition, the diode **17** prevents damage to the transistor **14** by clamping the reverse base-emitter voltage caused by the voltage spike across the winding **13b**.

The recaptured energy can be used in a number of ways. For example, the energy can be used to charge a battery **29**. In one embodiment, the timing wheel **26** makes two revolutions for each revolution of the rotor **20**. The contact rotor **27** closes a switch **28**, and thus dumps the charge on the capacitor **24** into the battery **29**, once each revolution of the wheel **26**. Other energy-recapture devices and techniques may also be used. Rotor **20** may be stopped, either by applying a brake to it or by opening the switch **12**.

Other embodiments of the monopole motor are contemplated. For example, instead of remaining closed for the entire operation of the motor, the switch **12** may be a conventional optical switch or a Hall-effect switch that opens and closes automatically at the appropriate times. To increase the power of the motor, the number of stators **18a** and pole pieces **18b**, may be increased and/or the number of poles **19**. Furthermore, one can magnetise the stator **18a** and pole piece **18b** during the attraction of the pole **19** instead of or in addition to magnetising the stator and pole piece during the repulsion of the pole **19**.

Moreover, the stator **18a** may be omitted so that coil **13** becomes an air coil, or the stator **18a** and the pole piece **18b** may compose a permanent magnet. In addition, although the transistor **14** is described as being a bipolar transistor, a MOSFET transistor may also be used. Furthermore, the recaptured energy may be used to recharge the battery **11**. In addition, although described as rotating in a clockwise direction, the rotor **20** can rotate in a counterclockwise direction. Moreover, although described as attracting a rotor pole **19** when no current flows through winding **13a** and repelling the pole **19** when a current flows through winding **13a**, the pole piece **18b** may be constructed so that it attracts the pole **19** when a current flows through winding **13a** and repels the pole **19** when no current flows through winding **13a**.

In multiple stator/rotor systems, each individual stator may be energised one at a time or all of the stators may be energised simultaneously. Any number of stators and rotors may be incorporated into the design of such multiple stator/rotor monopole motor combinations. However, while there may be several stators per rotor, there can only be one rotor for a single stator. The number of stators and rotors that would comprise a particular motor is dependent upon the amount of power required in the form of watts. Any number of magnets, used in a monopole fashion, may comprise a single rotor. The number of magnets incorporated into a particular rotor is dependent upon the size of the rotor and power required of the motor. The desired size and horse power of the motor determines whether the stators will be in parallel or fired sequentially. Energy is made accessible through the capturing of available energy from the back EMF as a result of the unique circuitry and timing of the monopole motor. Individual motors may be connected in sequence with each motor having various combinations of stators and rotors or they may be connected in parallel. Each rotor may have any number of rotor magnets, all arranged without change of polarity. The number of stators for an individual motor may also be of a wide range.

One feature that distinguishes this motor from all others, is the use of monopole magnets in momentary apposition with the pole piece of the stator maintaining the same polarity when magnetised. In this particular embodiment, there are three magnets and one pole piece, the pole piece being an extension of a permanent-magnet stator. Finally, although the invention has been described with reference of particular means, materials and embodiments, it is to be understood that the invention is not limited to the particulars disclosed and extends to all equivalents within the scope of the claims.

CLAIMS

1. A back EMF monopole motor utilising a rotor wherein the magnets maintain a polarity when in apposition with a stator pole piece having the polarity, said motor to capture available back EMF energy for charging and storage in a recovery device, the motor comprising:
- **a.** A means for producing initial energy;
- **b.** A means for capturing energy in the form of back EMF, caused by a collapsing field in a coil comprised of multiple windings with a pole piece at one end of the stator of the coil, the pole piece having the correct polarity when magnetised and in aligned with the magnets of the rotor;
- **c.** A means for rectifying the back EMF energy, comprising of a voltage bridge for transferring the back EMF energy to a capacitor for storage;
- **d.** A means for discharging the stored voltage across a recovery battery; and
- **e**. A means for counting the revolutions of the rotor.

2. The back EMF monopole motor of Claim 1, where a battery is used to provide the initial energy.

3. The back EMF monopole motor of claim 1, where the rotor revolutions are counted by a timing gear.

4. The back EMF monopole motor of claim 1, where the rotor revolutions are counted by a timing belt.

5. The back EMF monopole motor of claim 1, where the means for discharging collected energy comprises a rotating switching commutator which discharges the collected energy into a recovery battery, the commutator switch having the same polarity as the recovery battery.

6. A back EMF monopole motor utilising a rotor in which the rotor magnets maintain a polarity when aligned with a magnetised stator pole piece, suited to capturing available back EMF energy for charging and storage in a recovery device, the motor comprising:
- **a.** A primary input battery and a means for switching the battery, namely, either a solid-state switching circuitry, a magnetic Reed switch, a commutator, an optical switch, or a Hall-effect switch;
- **b.** A means for capturing energy in the form of back EMF, created by a collapsing field in a coil comprised of multiple windings and a pole piece at one end of the stator coil;
- **c.** A means for rectifying the back EMF energy comprising a voltage bridge for transferring the energy to a capacitor for storage;
- **d.** A means for discharging the stored voltage across a recovery battery, the means being a rotating contact rotor switch;
- **e.** A means for counting the revolutions of the rotor via a timing gear or timing belt;

f. A rotating switching commutator for switching the rotating contact rotor switch.

US Patent 6,392,370 21st May 2002 Inventor: John C. Bedini

DEVICE AND METHOD OF A BACK EMF PERMANENT ELECTROMAGNETIC MOTOR GENERATOR

ABSTRACT

This invention is a back EMF permanent electromagnetic motor generator and method using a regauging process for capturing available electromagnetic energy in the system. The device comprises a rotor with magnets of the same polarity; a timing wheel in apposition to a magnetic Hall-effect pickup switch semiconductor; and a stator comprised of two bars connected by a permanent magnet with magnetised pole pieces at one end of each bar. There are input and output coils created by wrapping each bar with a conducting material such as copper wire. Energy from the output coils is transferred to a recovery rectifier or diode. The magnets of the rotor, which is located on a shaft along with the timing wheel, are in apposition to the magnetised pole pieces of the two bars. The invention works through a process of regauging, that is, the flux fields created by the coils is collapsed because of a reversal of the magnetic field in the magnetised pole pieces thus allowing the capture of available back EMF energy. Additional available energy may be captured and used to re-energise the battery, and/or sent in another direction to be used for work. As an alternative, the available back EMF energy may be dissipated into the system.

BACKGROUND OF THE INVENTION

1. Field of the Invention

The invention relates generally to the capturing of electromagnetic energy using a method and device to create back EMF (electromagnetic force) and re-phasing of the back EMF to recycle and capture the available back EMF energy. Back EMF is also referred to as regauging and may be defined as energy created by the magnetic field from coils, and only from coils, and not from magnets.

2. Background Information and Related Art

Operation of a normal magnetic motor has the rotor pole attracting the stator pole, resulting in the generation of power from the magnets to the rotor and flywheel. During this phase, energy flows from the magnetics to the rotor/flywheel and is stored in the increased rotation. A rotor pole leaving a stator pole and creating a condition of drag-back results in power having to be put back into the magnetic section by the rotor and flywheel to forcibly overcome the drag-back. In a perfect, friction-free motor, the nett force field is therefore referred to as most conservative. In other words, a most conservative EMF motor has maximum efficiency. Without extra energy continually fed to the motor, no nett work can be done by the magnetic field, since half the time the magnetic field adds energy to the load (the rotor and flywheel) and the other half of the time it subtracts energy from the load (the rotor and flywheel). Therefore the total nett energy output is zero in any such rotary process without additional energy input. To use a present day magnetic motor, continuous energy must be fed to the motor to overcome drag-back and to power the motor and its load.

Present EMF motors and generators all use such conservative fields and therefore, have internal losses. Hence, it is necessary to continually input all of the energy that the motor outputs to the load, plus more energy to cover losses inside the motor itself. EMF motors are rated for efficiency and performance by how much energy input into the motor actually results in output energy to the load. Normally, the Coefficient of Performance (COP) rating is used as a measure of efficiency. The COP is the actual output energy going into the load and powering it, divided by the energy that must be input into the device with its load. COP is the power out into the load, divided by the power input into the motor/load combination. If there were zero internal losses in a motor, that "perfect" motor would have a coefficient of performance (COP) equal to 1.0. That is, all energy fed into the motor would be output by the motor directly into the load, and none of the input energy would be lost or dissipated in the motor itself.

In magnetic motor generators presently in use, however, due to friction and design flaws, there are always internal losses and inefficiencies. Some of the energy input into the motor is dissipated in these internal losses. As a consequence, the energy that gets to the load is always less than the input energy. So a standard motor operates with a COP of less than 1.0 which is expressed as COP<1.0. An inefficient motor may have a COP of 0.4 or 0.45, while a specially designed, highly efficient motor may have a COP of 0.85.

The conservative field inside a motor itself can be divided into two phases. Producing a conservative field involves nett symmetry between the "power out" phase from the magnetics to the rotor/flywheel and the "power back in" phase from the rotor/flywheel back to the magnetics. That is, the two flows of energy (one from the

magnetics into the rotor and flywheel, and one from the rotor and flywheel back to the magnetics) are identical in magnitude but opposite in direction. Each phase alone is said to be "asymmetrical"; that is, it either has:
1) a nett energy flow out to the rotor/flywheel; or
2) a nett energy flow back into the magnetics from the rotor/flywheel.
In simplified terms, it is referred to as "power out" and "power back in" phases with respect to the motor magnetics. Hence, the two asymmetrical phases are:
1) the power-out phase; and
2) the "power back in" phase, with reference to the magnetics.

For the power-out phase, energy is derived from the EMF existing between the stator pole and incoming rotor pole in an attraction mode. In this phase, the rotary motion (angular momentum and kinetic energy) of the rotor and flywheel is increased. In short, power is added to the rotor/flywheel (and thus to the load) from the fields between stator pole and rotor pole (the electromagnetic aspects of the system).

For the "power back in" phase, energy must be fed back into the magnetics from the rotor and flywheel (and the load) to overcome the drag-back forces existing between stator pole and outgoing rotor pole. In this phase, energy is returned to the internal magnetic system from the rotary motion of the rotor and flywheel (the angular momentum, which is the rotational energy multiplied by time). As is well known in physics, a rotor/flywheel's angular momentum provides a convenient way to store energy with the spinning rotor/flywheel mass acting as an energy reservoir.

All present day conventional magnetic motors use various methods for overcoming, and partially reversing, back EMF. Back EMF is the out of phase return pulse from the coil and is also referred to as regauging. The back EMF is shorted out and the rotor is attracted back in, therefore eliminating back drag. This can be accomplished by pouring more energy in to overpower the back EMF, thereby producing a forward EMF in that region. The energy required for this method must be furnished by the operator.

The motor of the present invention uses only a small amount of energy to "trigger" a much larger input of available energy by supplying back EMF, thus increasing the potential energy of the system. It then utilises this excess potential energy to reduce or reverse back EMF, thereby increasing the efficiency of the motor and, therefore, the COP.

If the energy in phase 1 (the power-out phase) is increased by additional available energy in the electromagnetics themselves, then the energy in phase 1 can be made greater than the energy in phase 2 (the power-back-in phase) without the operator furnishing the energy utilised. This produces a non-conservative nett field. Nett power can then be taken from the rotating stator and flywheel, because the available energy transferred into the stator and flywheel by the additional effects, is transformed by the rotor/flywheel into additional angular momentum and stored as such. Angular momentum is conserved at all times; but now some of the angular momentum added to the flywheel is generated by additional effects in the electromagnetics rather than being provided by the operator.

Electrodynamicists assume that the potential available energy of any system can be changed at will and without cost. This is back EMF and is well-known in physics. It is also routinely employed by electrodynamicists in the theoretical aspects. However, to simplify the mathematics, electrodynamicists will create a back EMF twice simultaneously, each back EMF being carefully selected so that the two available forces which are produced, are equal and opposite and cancel each other "symmetrically". This is referred to as "symmetrical back EMF". A symmetrical back EMF system cannot produce a COP>1.0.
On the other hand, the motor of the present invention deliberately creates a back EMF itself and its potential energy only once at a time, thereby retaining each extra force for a period of time and applying it to increase the angular momentum and kinetic energy of the rotor and flywheel. Specifically, this back EMF energy with its nett force is deliberately applied in the motor of the present invention to overcome and even reverse the conventional drag-back (the back EMF). Hence less energy need be taken from the rotor and flywheel to overcome the reduced back EMF, and in the ideal case none is required since the back EMF has been overpowered and converted to forward EMF by the back EMF energy and force. In the motor of the present invention, the conventional back-drag section of the magnetics becomes a forward-EMF section and now adds energy to the rotor/flywheel instead of subtracting it. The important feature is that the operator only has to provide the small amount of energy necessary to trigger the back EMF, and does not have to furnish the much larger back EMF energy itself.

When the desired energy in phase 1 (the power out phase) is thus made greater than the undesired "drag-back" energy in phase 2, then part of the output power normally dragged back from the rotor and flywheel by the fields in phase 2 is not required. Hence, compared to a system without the special back EMF mechanisms, additional power is available from the rotor/flywheel. The rotor maintains additional angular momentum and kinetic energy, compared to a system which does not produce back EMF itself. Consequently, the excess angular momentum

retained by the rotor and flywheel can be utilised as additional shaft power to power an external load connected to the shaft.

A standard magnetic motor operates as the result of the motor being furnished with external energy input into the system by the operator to reduce phase 2 (power back into the magnetics from the rotor/flywheel) by any of several methods and mechanisms. The primary purpose of this external energy input into the system is to overcome the back EMF and also provide for the inevitable energy losses in the system. There is no input of energy separate from the operator input. Therefore, the COP of any standard magnetic motor is COP less than 1.0. The efficiency of a standard magnetic motor varies from less than 50% to a maximum of about 85%, and so has a COP<1.0. When nothing is done in the motor that will produce a reduction of the back EMF without the operator inputting all the energy for it, then for even a frictionless, ideal permanent magnet motor, the COP can never exceed 1.0.

Until the introduction of the motor of the present invention, it has been standard universal practice that the operator must furnish all energy used to reduce the back EMF, provide for the internal losses, and power the load. It is therefore a common belief by the scientific community that an ideal (loss-less) permanent magnet motor cannot exceed a COP of 1.0. That is true, so long as the operator himself must furnish all the energy. Furthermore, since real permanent magnetic motors have real internal losses, some of the input energy is always lost in the motor itself, and that lost energy is not available for powering the rotor/flywheel and load. Hence a real permanent magnetic motor of the conventional kind will always have a COP<1.0.

The common assumption that the COP of a motor is limited to less than 1.0 is not necessarily true, and that COP>1.0 is permitted without violating the laws of nature, physics, or thermodynamics. However, it can immediately be seen that any permanent magnet motor exhibiting a COP>1.0 must have some available energy input returning in the form of back EMF.

A problem relates to how back EMF energy can be obtained from a circuit's external environment for the specific task of reducing the back-drag EMF without the operator having to supply any input of that excess energy. In short, the ultimate challenge is to find a way to cause the system to:
1) become an open dissipative system, that is, a system receiving available excess energy from its environment, in other words, from an external source; and
2) use that available excess energy to reduce the drag-back EMF between stator and rotor poles as the rotor pole is leaving the stator pole.
If this objective can be accomplished, the system will be removed from thermodynamic equilibrium. Instead, it will be converted to a system out-of-thermodynamic equilibrium. Such a system is not obliged to obey classical equilibrium thermodynamics.

Instead, an out-of-equilibrium thermodynamic system must obey the thermodynamics of open systems far from the established and well-known parameters of thermodynamic equilibrium. As is well known in the physics of thermodynamics, such open systems can permissibly:
1) self-order;
2) self-oscillate;
3) output more back EMF energy than energy input by the operator (the available excess back EMF energy is received from an external source and some energy is input by the operator as well);
4) power itself as well as its loads and losses simultaneously (in that case, all the energy is received from the available external source and there is no input energy from the operator); and
5) exhibit negative entropy, that is, produce an increase of energy that is available in the system, and that is independent of the energy put into the system by the operator.

As a definition, entropy roughly corresponds to the energy of a system that has become unavailable for use. Negative entropy corresponds to additional energy of a system that has become available for use.

In the back EMF permanent magnet electromagnetic motor generator of the present invention, several known processes and methods are utilised which allow the invention to operate periodically as an open dissipative system (receiving available excess energy from back EMF) far from thermodynamic equilibrium, whereby it produces and receives its excess energy from a known external source.

A method is utilised to temporarily produce a much larger source of available external energy around an energised coil. Then the unique design features of this new motor provides a method and mechanism that can immediately produce a second increase in that energy, concurrently as the energy flow is reversed. Therefore, the motor is capable of producing two asymmetrical back EMFs, one after the other, of the energy within a single coil, which dramatically increases the energy available and causes that available excess energy to then enter the circuit as an impulse, being collected and utilised.

The present motor utilises this available excess back EMF energy to overcome and even reverse the back-drag EMF between stator pole and rotor pole, while furnishing only a small trigger pulse of energy necessary to control and activate the direction of the back EMF energy flow.

By using a number of such dual asymmetrical self back EMFs for every revolution of the rotor, the rotor and flywheel collectively focus all the excess impulsive inputs into increased angular momentum (expressed as energy multiplied by time), shaft torque, and shaft power.

Further, some of the excess energy deliberately generated in the coil by the utilisation of the dual process manifests itself in the form of excess electrical energy in the circuit and is utilised to power electrical loads, e.g., a lamp, fan, motor, or other electrical devices. The remainder of the excess energy generated in the coil can be used to power the rotor and flywheel, with the rotor/flywheel also furnishing shaft horsepower for powering mechanical loads.

This new and unique motor utilises a means to furnish the relatively small amount of energy to initiate the impulsive asymmetrical self back EMF actions. Then part of the available excess electrical power drawn off from the back EMFs is utilised to recharge the battery with dramatically increased over voltage pulses.

The unique design features of this motor utilise both north and south magnetic poles of each rotor and stator magnet. Therefore, the number of impulsive self back EMFs in a single rotation of the rotor is doubled. Advanced designs increase the number of self back EMFs in a single rotor rotation with the result that there is an increase in the number of impulses per rotation which increase the power output of this new motor.

The sharp voltage pulse produced in the coil of this new motor by the rapidly collapsing field in the back EMF coil is connected to a battery in charge mode and to an external electrical load. The nett result is that the coil asymmetrically creates back EMF itself in a manner adding available energy and impulse to the circuit. The excess available energy collected in the coil is used to reverse the back-EMF phase of the stator-rotor fields to a forward EMF condition, and through an impulse, adding acceleration and angular momentum to the rotor and flywheel. At the same time, a part of the excess energy collected in the coil is used to power electrical loads such as charging a battery and operating a lamp or such other device.

It is well known that changing the voltage alone, creates a back EMF and requires no work. This is because to change the potential energy does not require changing the form of that potential energy, but only its magnitude. Strictly speaking, work is the changing of the form of energy. Therefore, as long as the form of the potential energy is not changed, the magnitude can be changed without having to perform work in the process. The motor of the present invention takes advantage of this permissible operation to create back EMF asymmetrically, and thereby change its own usable available potential energy.

In an electric power system, the potential (voltage) is changed by inputting energy to do work on the internal charges of the generator or battery. This potential energy is expended within the generator (or battery) to force the internal charges apart, forming a source dipole. Then the external closed circuit system connected to that source dipole ineptly pumps the spent electrons in the ground line back through the back EMF of the source dipole, thereby scattering the charges and killing the dipole. This shuts off the energy flow from the source dipole to the external circuit. As a consequence of that conventional method, it is a requirement to input and replace additional energy to again restore the dipole. The circuits currently utilised in most electrical generators have been designed to keep on destroying the energy flow by continually scattering all of the dipole charges and terminating the dipole. Therefore, it is necessary to keep on inputting energy to the generator to keep restoring its source dipole.

An investigation of particle physics is required to see what furnishes the energy to the external circuit. Since neither a battery nor a generator furnishes energy to the external circuit, but only furnishes energy to form the source dipole, a better understanding of the electric power principle is required to fully understand how this new motor functions. A typical battery uses its stored chemical energy to form the source dipole. A generator utilises its input shaft energy of rotation to generate an internal magnetic field in which the positive charges are forced to move in one direction and the negative charges in the reverse direction, thereby forming the source dipole. In other words, the energy input into the generator does nothing except form the source dipole. None of the input energy goes to the external circuit. If increased current is drawn into the external load, there also is increased spent electron flow being rammed back through the source dipole, destroying it faster. Therefore, dipole-restoring-energy has to be inputted faster. The chemical energy of the battery also is expended only to separate its internal charges and form its source dipole. Again, if increased current and power is drawn into the external load, there is increased spent electron flow being rammed back through the source dipole, destroying it faster. This results in a depletion of the battery's stored energy faster, by forcing it to have to keep restoring the dipole faster.

Once the generator or battery source dipole is formed (the dipole is attached also to the external circuit), it is well known in particle physics that the dipole (same as any charge) is a broken symmetry in the vacuum energy flux.

By definition, this means that the source dipole extracts and orders part of that energy received from its vacuum interaction, and pours that energy out as the energy flowing through all space surrounding the external conductors in the attached circuit. Most of this enormous energy flow surging through space surrounding the external circuit does not strike the circuit at all, and does not get intercepted or utilised. Neither is it diverted into the circuit to power the electrons, but passes on out into space and is just "wasted". Only a small "sheath" of the energy flow along the surface of the conductors strikes the surface charges in those conductors and is thereby diverted into the circuit to power the electrons. Standard texts show the huge available but wasted energy flow component, but only calculate the small portion of the energy flow that strikes the circuit, is caught by it, and is utilised to power it.

In a typical circuit, the huge available but "wasted" component of the energy flow is about 10 to the power 13 times as large as the small component intercepted by the surface charges and diverted into the circuit to power it. Hence, around every circuit and circuit element such as a coil, there exists a huge non-intercepted, non-diverged energy flow that is far greater than the small energy flow being diverted and used by the circuit or element.

Thus there exists an enormous untapped energy flow immediately surrounding every EMF power circuit, from which available excess energy can be intercepted and collected by the circuit, if respective non-linear actions are initiated that sharply affect and increase the reaction cross section of the circuit (i.e., its ability to intercept this available but usually wasted energy flow).

The method in which the motor of the present invention alters the reaction cross section of the coils in the circuit, is by a novel use, which momentarily changes the reaction cross section of the coil in which it is invoked. Thus, by this new motor using only a small amount of current in the form of a triggering pulse, it is able to evoke and control the immediate change of the coil's reaction cross section to this normally wasted energy flow component. As a result, the motor captures and directs some of this usually wasted environmental energy, collecting the available excess energy in the coil and then releasing it for use in the motor. By timing and switching, the innovative gate design in this new motor directs the available excess energy so that it overcomes and reverses the return EMF of the rotor-stator pole combination during what would normally be the back EMF and demonstrates the creation of the second back EMF of the system. Now instead of an "equal retardation" force being produced in the back EMF region, a forward EMF is produced that is additive to the rotor/flywheel energy and not subtractive. In short, it further accelerates the rotor/flywheel.

This results in a non-conservative magnetic field along the rotor's path. The line integral of the field around that path (i.e., the nett work on the rotor/flywheel to increase its energy and angular momentum) is not zero but a significant amount. Hence, the creation of an asymmetrical back EMF impulse magnetic motor:
1) takes its available excess energy from a known external source, the huge usually non-intercepted portion of the energy flow around the coil;
2) further increases the source dipolarity by this back EMF energy; and
3) produces available excess energy flow directly from the source dipole's increased broken symmetry in its fierce energy exchange with the local vacuum.

No laws of physics or thermodynamics are violated in the method and device of the present invention, and conservation of energy rigorously applies at all times. Nonetheless, by operating as an open dissipative system not in thermodynamic equilibrium with the active vacuum, the system can permissibly receive available excess energy from a known environmental source and output more energy to a load than must be input by the operator alone. As an open system not in thermodynamic equilibrium, this new and unique motor can tap in to back EMF to energise itself, loads and losses simultaneously, fully complying with known laws of physics and thermodynamics.

A search of prior art failed to reveal any devices that recycle available energy from back EMF of a permanent electromagnetic motor generator as described in the present invention. However, the following prior art US patents were reviewed:

1. No. 5,532,532 to DeVault, et al., Hermetically Sealed Super-conducting Magnet Motor.
2. No. 5,508,575 to Elrod, Jr., Direct Drive Servovalve Having Magnetically Loaded Bearing.
3. No. 5,451,825 to Strohm, Voltage Homopolar Machine.
4. No. 5,371,426 to Nagate et al., Rotor For Brushless Motor.
5. No. 5,369,325 to Nagate et al., Rotor For Brushless Electromotor And Method For Making Same.
6. No. 5,356,534 to Zimmermann, deceased et al., Magnetic-Field Amplifier.
7. No. 5,350,958 to Ohnishi, Super-conducting Rotating Machine, A Super-conducting Coil, And A Super-conducting Generator For Use In A Lighting Equipment Using Solar Energy.
8. No. 5,334,894 to Nakagawa, Rotary Pulse Motor.
9. No. 5,177,054 to Lloyd, et al., Flux Trapped Superconductor Motor and Method.
10. No. 5,130,595 to Arora, Multiple Magnetic Paths Pulse Machine.
11. No. 4,980,595 to Arora, Multiple Magnetics Paths Machine.
12. No. 4,972,112 to Kim, Brushless D.C. Motor.

13. No. 4,916,346 to Kliman, Composite Rotor Lamination For Use In Reluctance Homopolar, And Permanent Magnet Machines.
14. No. 4,761,590 to Kaszman, Electric Motor.
15. No. 4,536,230 to Landa, et al., Anisotropic Permanent Magnets.
16. No. Re. 31,950 to Binns, Alternating Current Generators And Motors.
17. No. 4,488,075 to DeCesare, Alternator With Rotor Axial Flux Excitation.
18. No. 4,433,260 to Weisbord et al., Hysteresis Synchronous Motor Utilizing Polarized Rotor.
19. No. 4,429,263 to Muller, Low Magnetic Leakage Flux Brushless Pulse Controlled D-C Motor.
20. No. 4,423,343 to Field, II, Synchronous Motor System.
21. No. 4,417,167 to Ishii et al., DC Brushless Motor.
22. No. 4,265,754 to Menold, Water Treating Apparatus and Methods.
23. No. 4,265,746 to Zimmermann, Sr. et al. Water Treating Apparatus and Methods.
24. No. 4,222,021 to Bunker, Jr., Magnetic Apparatus Appearing To Possess a Single Pole.
25. No. 2,974,981 to Vervest et al., Arrester For Iron Particles.
26. No. 2,613,246 to Spodig, Magnetic System.
27. No. 2,560,260 to Sturtevant et al., Temperature Compensated Magnetic Suspension.

SUMMARY OF THE INVENTION

The device and method of the present invention is a new permanent electromagnetic motor generator that recycles back EMF energy (regauging) thus allowing the motor to produce an energy level of COP = 0.98, more or less, depending upon configuration, circuitry, switching elements and the number and size of stators, rotors and coils that comprise the motor. The rotor is fixed between two pole pieces of the stator. The motor generator is initially energised from a small starter battery means, analogous to a spark plug, that sends a small amount of energy to the motor, thus stimulating a rotating motion from the rotor. As the rotor rotates, energy is captured from the surrounding electromagnetic field containing an asymmetrical pulse wave of back EMF. The energy produced and captured can be directed in one of several directions, including returning energy to the initial starter battery, rotating a shaft for work and/or sending a current to energise a fan, light bulb or other such device.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG.1 is the top view of a back EMF permanent electromagnetic motor generator with a single stator and a single rotor.

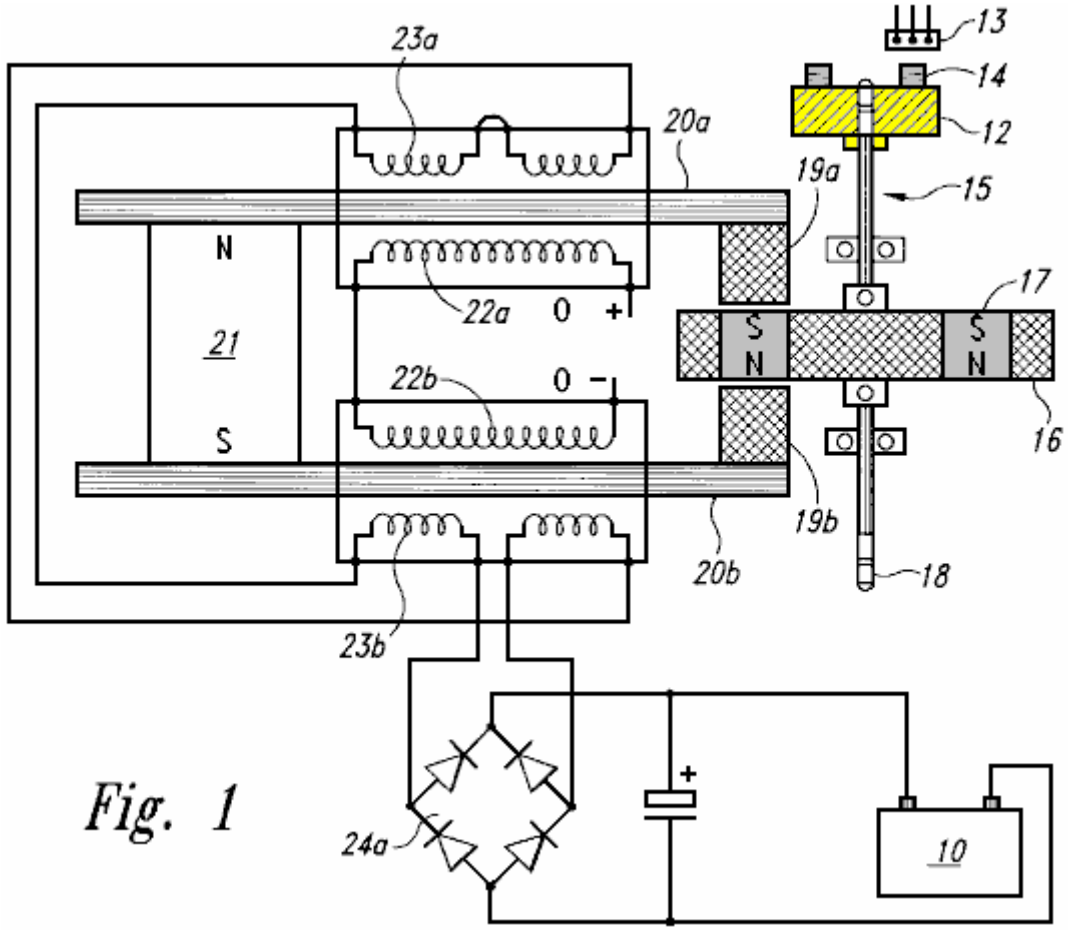


Fig. 1

Fig.1a is a side view of a timing wheel and magnetic Hall-effect sensor of the back EMF motor generator.

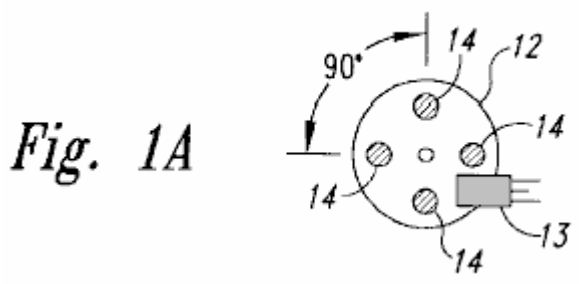


Fig.1b is a side view of the rotor of the back EMF motor generator.

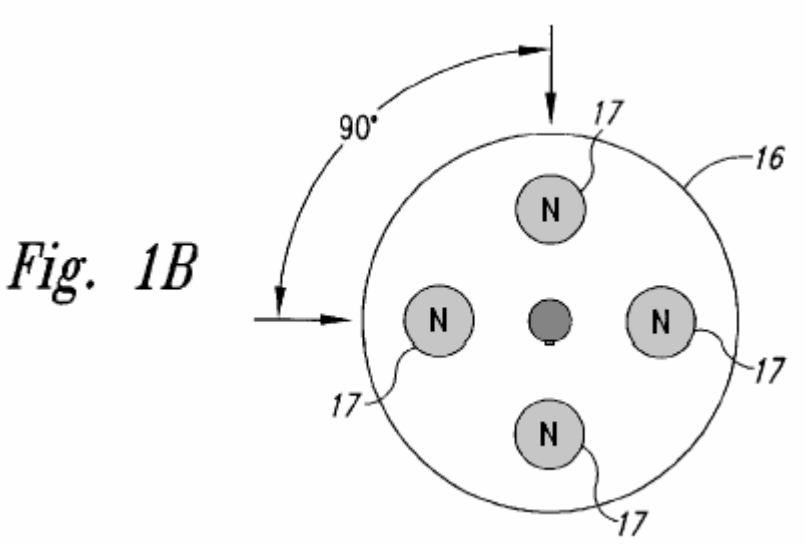


Fig.2 is a schematic drawing incorporating circuitry for the back EMF motor generator.

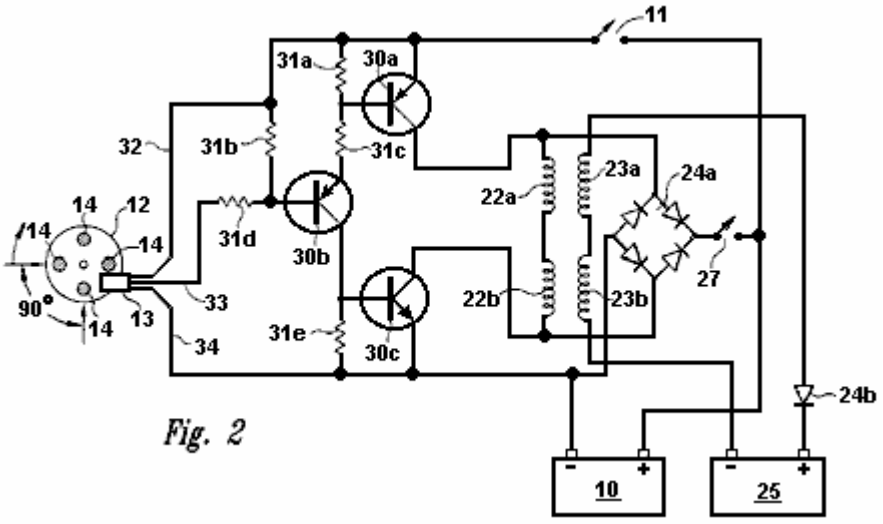


Fig.3 is a box diagram showing the relationships of the back EMF motor generator circuitry.

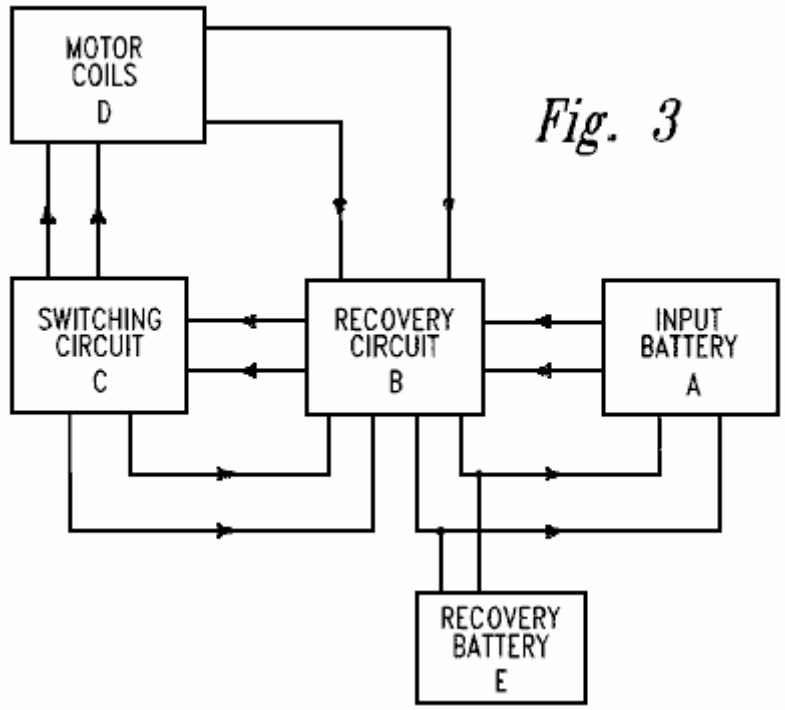


Fig. 3

DETAILED DESCRIPTION OF THE INVENTION

The present invention is a device and method for creating a back EMF permanent electromagnetic motor generator. As described in the Background Information, this new motor generator conforms to all applicable electrodynamic laws of physics and is in harmony with the law of the conservation of energy, the laws of electromagnetism and other related natural laws.

The back EMF permanent electromagnetic motor generator is comprised of a combination of electrical, material and magnetic elements, arranged to capture available electromagnetic energy (back EMF) in a recovery rectifier or single diode from output coils. The capturing of back EMF energy is also known as 'regauging'. As an arbitrary starting point in describing this invention, an input battery, as a means of energy, sends power through a power on-off switch and then to a timing mechanism, such as a magnetic timing switch (a semiconductor Hall-effect magnetic pickup switch) which is triggered by a magnet on a timing wheel. The timing wheel may contain any number of magnets (i.e. one or more), with the South poles facing outwards and aligned with the Hall-effect pickup switch.

The timing wheel is mounted at the end of a shaft which is located along the centreline of a rotor, which in turn, may contain any number of magnets (i.e. two or more). The rotor magnets are arranged so that they have the same polarity and are equidistant from each other. The shaft has the timing wheel mounted at one end, the rotor, and then some means for performing work, such as a power take off at the opposite end. However, there are other embodiments in which the position of the rotor, timing wheel and power take-off have other configurations. The rotor is mounted on a platform or housing which is fixed in a stationary position within a stator.

The stator is comprised of a permanent magnet connected to a means for conducting electromagnetic energy such as two parallel bars, each bar having a magnetised pole piece at one end. The conduction material of the bar may be ferrous, powdered iron, silicon steel, stainless magnetic steel, laminations of conductive material or any other magnetic conductive material. Each bar has an input coil placed around it. The coil may be constructed from copper, aluminium or any other suitable conductive material. The primary or input coil is connected to the switching circuit. A second coil on top of the input coil becomes a secondary or output coil. The secondary or output coil is connected to the recovery circuit. The rotor is located symmetrically between the pole pieces of the bars of the stator and it contains a series of magnets all having the same polarity, North or South, with each magnet in the rotor being in aligned with the pole piece as the rotor rotates.

When the rotor is energised from the battery of the switching circuit, there is an initial magnetic field that is instantly overcome as the magnetised pole pieces align with the rotor magnets. As the rotor begins to move, increasing electromagnetic energy is produced as a result of flux gaiting from the aligned magnets of the rotor and pole pieces. The coils surrounding the bars "buck" the permanent magnet connecting the bars. This is known as the "buck boosting" principle. When the permanent magnet is bucked by the coils, it reverses the polarity of the pole pieces which are aligned with the rotor magnets causing the rotor to increase its rate of rotation. The energy

available from the fields that are collapsing in the primary and secondary coils, (which creates the back EMF within the system), is now in non-equilibrium. Energy can now be put back into the system via the switching circuitry. Available energy captured from the back EMF, may be applied in different directions, including re-energising the input battery, storage in a capacitor, conversion by a recovery rectifier to be stored in the input battery, a capacitor or a secondary or recovery battery. Recovery rectifiers are used to convert this AC to DC. Available energy may be used to energise an electric bulb, fan or any other uses.

The shaft in the centre of the rotor can transfer energy in the form of work through a power take-off. The power take-off may be connected to any number of secondary shafts, wheels, gears and belts to increase or reduce torque.

This is a description of the basic invention, however, there are an innumerable number of combinations and embodiments of stators, rotors, Hall-effect magnetic pickup switches, coils, recovery rectifiers and electronic connecting modes that may be combined on a single shaft or several shafts connected in various combinations and sequences, and of various sizes. There may be any number of stators to one rotor, (however, there can be only one active rotor if there is a single stator). The number of Hall-effect pickup switches may vary, for example, in the case of multiple stators of high resistant coils, the coils may be parallel to form a low resistant coil so that one Hall-effect pickup with one circuit may fire all of the stators at the same time. The number of magnets in both the timing wheel and the rotor may also vary in number as well as the size and strength of the magnets. Any type of magnet may be used. The number of turns on both the input and output coils on each conducting bar may also vary in number and in conductive material.

The motor generator, as shown in **Fig.1**, a top perspective view of a single stator, single rotor back EMF motor and is comprised of a means of providing energy, such as input battery **10** connected to power switch **11** (shown in **Fig.2**) and Hall-effect magnetic pickup switch **13**. Magnetic pickup **13** interfaces with timing wheel **12** to form a timing switch. Timing wheel **12** contains four magnets **14** with the South pole of each said magnet facing outward towards magnetic pickup **13**. Timing wheel **12** is fixed at one end of shaft **15**. Located on shaft **15** is rotor **16**. Rotor **16** can be of any realistic size, and in this example the rotor contains four rotor magnets **17**. The rotor magnets **17** are arranged so all have the same polarity.

Opposite timing wheel **12** on shaft **15** is a means for performing work, such as a power take-off **18**. Rotor **16** is mounted in a fixed position with rotor magnets **17** in aligned with the magnetised pole pieces **19a** and **19b**. Each pole piece **19a** and **19b** is connected to iron bars **20a** and **20b**. These Iron bars are connected by a permanent magnet **21**. Wire is wrapped around iron bars **20a** and **20b** to form input coils **22a** and **22b**. Superimposed upon input coils **22a** and **22b** are output coils **23a** and **23b**. These output coils are connected to full wave bridge first recovery rectifier **24a** which then connects to battery **10**.

Fig.1a is a side view of the back EMF Motor Generator timing wheel **12** with Hall-effect magnetic pickup **13** positioned to be triggered by each of the four magnets **14** in turn as timing wheel **12** rotates. The magnets **14** have their South poles facing outward and they are spaced evenly with a 90 degree angular separation.

Fig.1b is a side view of rotor **16** with four rotor magnets **17** with 90 degree angular separation from each other and having the same polarity.

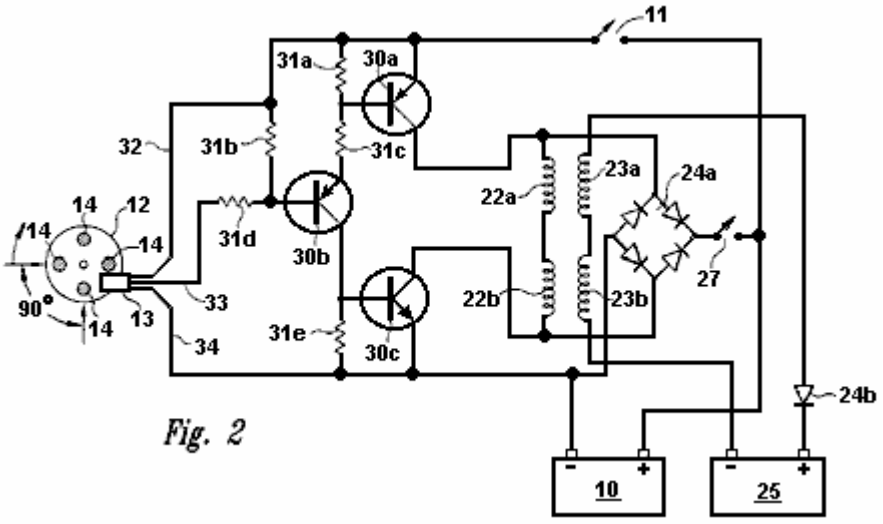


Fig. 2

Fig.2 is a schematic diagram of the motor generator circuitry showing input coil connections from input battery **10** through power switch **11**, transistors **30a,b,c** resistors **31a-e**, through power supply lead **32** ("VCC+") and to magnetic pickup **13**. Magnetic pickup **13** is in aligned with timing wheel magnets **14** located on timing wheel **12**.

Collector lead **33** and ground lead **34** carry the signals from magnetic pickup **13**. When current is reversed, it flows through resistor **31e** and transistor **30c** to input battery **10**. Input coils **22a,b** send power to full wave bridge first recovery rectifier **24a** which then sends power through switch recovery **27** back into the system, and/or to the input battery **10**. Output coils **23a** and **23b** send power through single diode second recovery rectifier **24b** to recovery battery **25**.

In this particular embodiment, the value and type number of the components are as follows:
Hall-effect magnetic pickup switch **13** is a No. 3020;
Transistor **30a** is a 2N2955;
Transistor **30b** is an **MPS8599**;
Transistor **30c** is a 2N3055;
Resistors **31a** and **31b** are 470 ohms
Resistor **31b** is 2.2 K ohms
Resistor **31c** is 220 ohms
Resistor **31d** is 1 K ohms
Recovery rectifier **24a** is a 10 Amp, 400 volts bridge rectifier.

Fig.3 is a box diagram showing the flow of voltage from input battery **A**, through recovery circuit **B**, switching circuit **C** and motor coils **D**. Motor coils **D** send available back EMF energy through recovery circuit **B**, and then on to recovery battery **E** and input battery **A**. Available back EMF energy can also flow from switching circuit **C** to recovery circuit **B**.

In multiple stator/rotor systems, each individual stator may be energised one at a time or all of the stators may be energised simultaneously. Any number of stators and rotors may be incorporated into the design of such multiple stator/rotor motor generator combinations. However, while there may be several stators per rotor, there can only be one rotor for a single stator. The number of stators and rotors that would comprise a particular motor generator is dependent upon the amount of power required in the form of watts. The desired size and horsepower of the motor determines whether the stators will be in parallel or fired sequentially by the magnetic Hall-effect pickup or pickups. The number of magnets incorporated into a particular rotor is dependent upon the size of the rotor and power required of the motor generator. In a multiple stator/rotor motor generator, the timing wheel may have one or more magnets, but must have one magnet Hall-effect pickup for each stator if the stators are not arranged in parallel. The back EMF energy is made available through the reversing of the polarity of the magnetised pole pieces thus collapsing the field around the coils and reversing the flow of energy to the recovery diodes, which are capturing the back EMF.

Individual motors may be connected in sequence, with each motor having various combinations of stators and rotors, or they may be connected in parallel. Each rotor may have any number of magnets ranging from a minimum of 2 to maximum of 60. The number of stators for an individual motor may range from 1 to 60 with the number of conducting bars ranging from 2 to 120.

What distinguishes this motor generator from all others is the presence of a permanent magnet connecting the two conducting bars which transfer magnetic energy through the pole pieces to the rotor, thereby attracting the rotor between the pole pieces. With the rotor attracted in between the two pole pieces, the coils switch the polarity of the magnetic field of the pole pieces so that the rotor is repelled out. Therefore there is no current and voltage being used to attract the rotor. The only current being used is the repulsion of the rotor between the two conductive bar pole pieces thereby requiring only a small amount of current to repel the rotor. This is known as ‘a regauging system’ and allows the capturing of available back EMF energy.

Finally, although the invention has been described with reference of particular means, materials and embodiments, it is to be understood that the invention is not limited to the particulars disclosed and extends to all equivalents within the scope of the claims.

JOHN BEDINI: BATTERY PULSE CHARGER

US Patent Application 2003/117111 26th June 2003 Inventor: John C. Bedini

DEVICE AND METHOD FOR PULSE-CHARGING A BATTERY AND FOR DRIVING OTHER DEVICES WITH A PULSE

This is a slightly reworded copy of this patent application which shows a method of pulse-charging a battery bank or powering a heater and/or a motor. John Bedini is an intuitive genius with very considerable practical ability, so any information coming from him should be considered most carefully. At the end of this document there is some additional information not found in the patent.

ABSTRACT

This two-phase solid-state battery charger can receive input energy from a variety of sources including AC current, a battery, a DC generator, a DC-to-DC inverter, solar cells or any other compatible source of input energy. Phase 1 is the charging phase and Phase 2 is the discharge phase, where a signal, or current, passes through a dual timing switch which independently controls two channels, thus producing the two phases.

The dual timing switch is controlled by a logic chip, or pulse width modulator. A potential charge is allowed to build up in a capacitor bank. The capacitor bank is then disconnected from the energy input source and then a high voltage pulse is fed into the battery which is there to receive the charge. The momentary disconnection of the capacitor from the input energy source allows a free-floating potential charge in the capacitor. Once the capacitor has completed discharging the potential charge into the battery, the capacitor disconnects from the charging battery and re-connects to the energy source, thus completing the two-phase cycle.

TECHNICAL FIELD

This invention relates generally to a battery pulse-charger using a solid-state device and method where the current going to the battery is not constant. The signal or current is momentarily switch-interrupted as it flows through either the first channel, (the charging phase), or the second channel, (the discharging phase). This two-phase cycle alternates the signal in the two channels thereby allowing a potential charge in a capacitor to disconnect from its power source an instant before the capacitor discharges its stored potential energy into a battery set up to receive the capacitor's stored energy. The capacitor is then disconnected from the battery and re-connected to the power source upon completion of the discharge phase, thereby completing the charge-discharge cycle. The battery pulse-charger can also drive devices, such as a motor and a heating element, with pulses.

BACKGROUND AND PRIOR ART

Present day battery chargers use a constant charge current in their operation with no momentary disconnection of the signal or current as it flows either: (1) from a primary energy source to the charger; or (2) from the charger itself into a battery for receiving the charge. Some chargers are regulated to a constant current by any of several methods, while others are constant and are not regulated. There are no battery chargers currently in the art or available wherein there is a momentary signal or current disconnection between the primary energy source and the charger capacitors an instant before the capacitors discharge the stored potential energy into a battery receiving the pulse charge. Nor are there any chargers in the art that disconnect the charger from the battery receiving the charge when the charger capacitors receive energy from the primary source. The momentary current interruption allows the battery a short "rest period" and requires less energy from the primary energy source while putting more energy into the battery receiving the charge while requiring a shorter period of time to do it.

SUMMARY OF THE INVENTION

One aspect of the invention relates to a solid-state device and method for creating a pulse current to pulse-charge a battery or a bank of batteries in which a new and unique method is used to increase and preserve, for a longer period of time, the energy stored in the battery, as compared to constant-current battery chargers. The device uses a timed pulse to create a DC pulse waveform to be discharged into the battery receiving the charge.

One embodiment of the Invention uses a means for dual switching such as a pulse-width modulator (PWM), for example, a logic chip SG3524N PWM, and a means for optical coupling to a bank of high-energy capacitors to

store a timed initial pulse charge. This is the charge phase, or phase 1. The charged capacitor bank then discharges the stored high energy into the battery receiving the charge in timed pulses. Just prior to discharging the stored energy into the battery, the capacitor bank is momentarily disconnected from the power source, thus completing the charge phase, and thereby leaving the capacitor bank as a free-floating potential charge disconnected from the primary energy source to then be discharged into the battery. The transfer of energy from the capacitor bank to the battery completes the discharge phase, or phase 2. The two-phase cycle now repeats itself.

This embodiment of the battery pulse-charger works by transferring energy from a source, such as an AC source, to an unfiltered DC source of high voltage to be stored in a capacitor or a capacitor bank. A switching regulator is set to a timed pulse, for example, a one second pulse that is 180 degrees out of phase for each set of switching functions. The first function is to build the charge in the capacitor bank from the primary energy source; the second function is to disconnect the power source from the capacitor bank; the third function is to discharge the stored high voltage to the battery with a high voltage spike in a timed pulse, for example, a one second pulse; and the fourth function is to re-connect the capacitor bank to the primary energy source.

The device operates through a two-channel on/off switching mechanism or a gauging/re-gauging function wherein the charger is disconnected from its primary energy source an instant before the pulse-charger discharges the high-energy pulse into the battery to be charged. As the primary charging switch closes, the secondary discharging switch opens, and vise-versa in timed pulses to complete the two phase cycle.

The means for a power supply is varied with several options available as the primary energy source. For example, primary input energy may come from an AC source connected into the proper voltage (transformer); from an AC generator; from a primary input battery; from solar cells; from a DC-to-DC inverter; or from any other adaptable source of energy. If a transformer is the source of primary input energy, then it can be a standard rectifying transformer used in power supply applications or any other transformer applicable to the desired function. For example, it can be a 120-volt to 45-volt AC step-down transformer, and the rectifier can be a full-wave bridge of 200 volts at 20 amps, which is unfiltered when connected to the output of the transformer. The positive output terminal of the bridge rectifier is connected to the drains of the parallel connected field-effect transistors, and the negative terminal is connected to the negative side of the capacitor bank.

The Field Effect Transistor (FET) switches can be IRF260 FETs, or any other FET needed to accomplish this function. All the FETs are connected in parallel to achieve the proper current handling capacity for the pulses. Each FET may be connected through a 7-watt, 0.05-ohm resistor with a common bus connection at the source. All the FET gates may be connected through a 240-ohm resistor to a common bus. There may also be a 2 K-ohm resistor wired between the FET gates and the drain bus.

A transistor, for example an MJE15024, can be used as a driver for the gates, driving the bus, and in turn, an optical coupler powers the driver transistor through the first channel. A first charging switch is used to charge the capacitor bank, which acts as a DC potential source to the battery. The capacitor bank is then disconnected from the power rectifier circuit. The pulse battery charger is then transferred to a second field effect switch through the second channel for the discharge phase. The discharge phase is driven by a transistor, and that transistor is driven via an optical coupler. When the second (discharge) switch is turned on, the capacitor bank potential charge is discharged into the battery waiting to receive the charge. The battery receiving the charge is then disconnected from the pulse-charger capacitor bank in order to repeat the cycle. The pulse-charger may have any suitable source of input power including:

(1) solar panels to raise the voltage to the capacitor bank;
(2) a wind generator;
(3) a DC-to-DC inverter;
(4) an alternator;
(5) an AC motor generator;
(6) a static source such as a high voltage spark; and
(7) other devices which can raise the potential of the capacitor bank.

In another embodiment of the invention, one can use the pulse-charger to drive a device such as a motor or heating element with pulses of energy.

BRIEF DESCRIPTION OF THE DRAWINGS

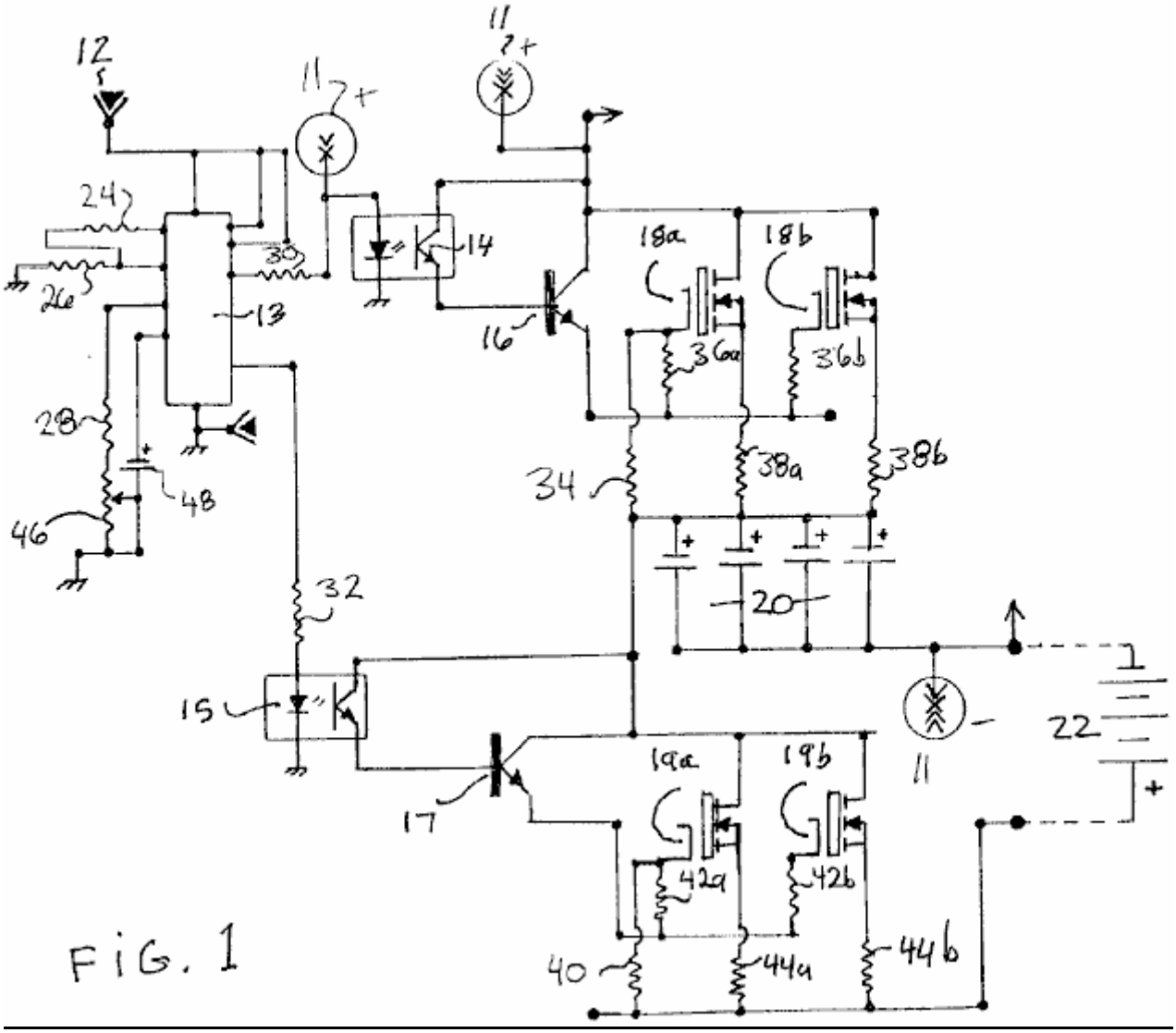


Fig.1 is a schematic drawing of a solid-state pulse-charger according to an embodiment of the invention.

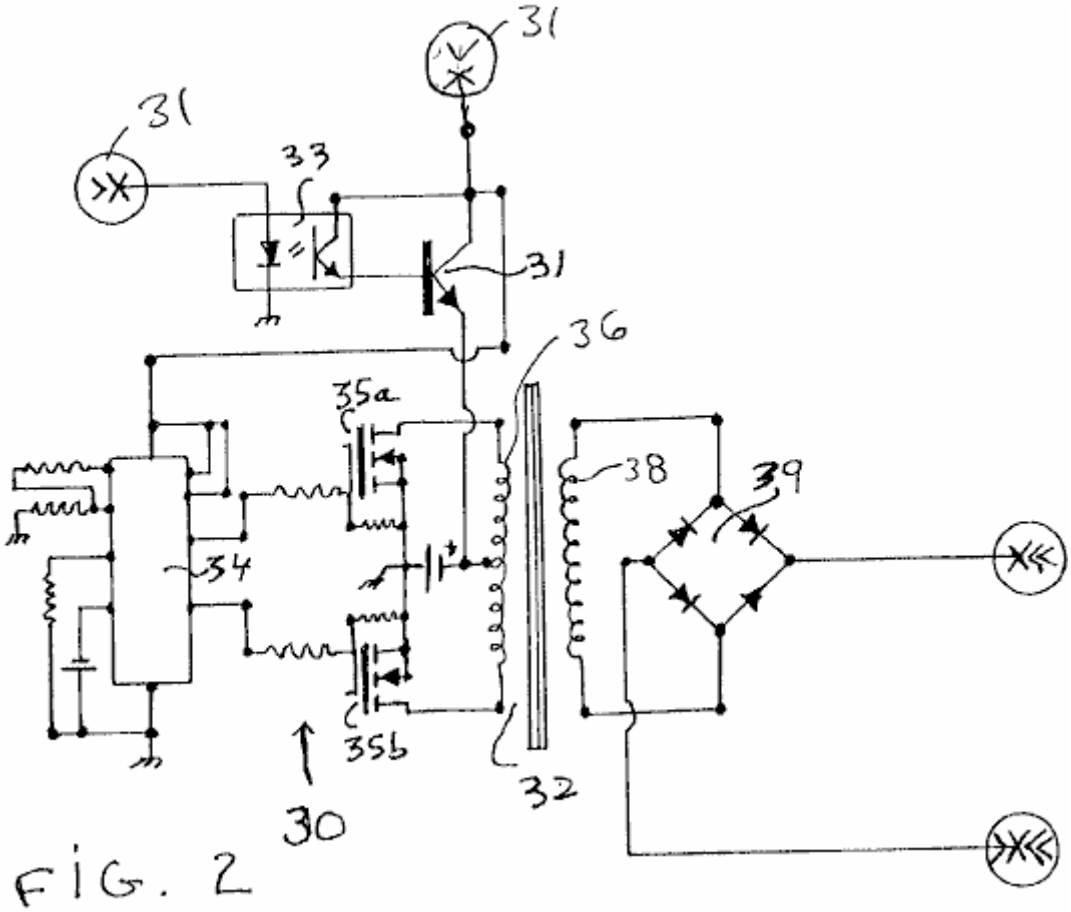


FIG. 2

Fig.2 is a schematic drawing of a conventional DC-to-DC converter that can be used to provide power to the pulse-charger of Fig.1 according to an embodiment of the invention.

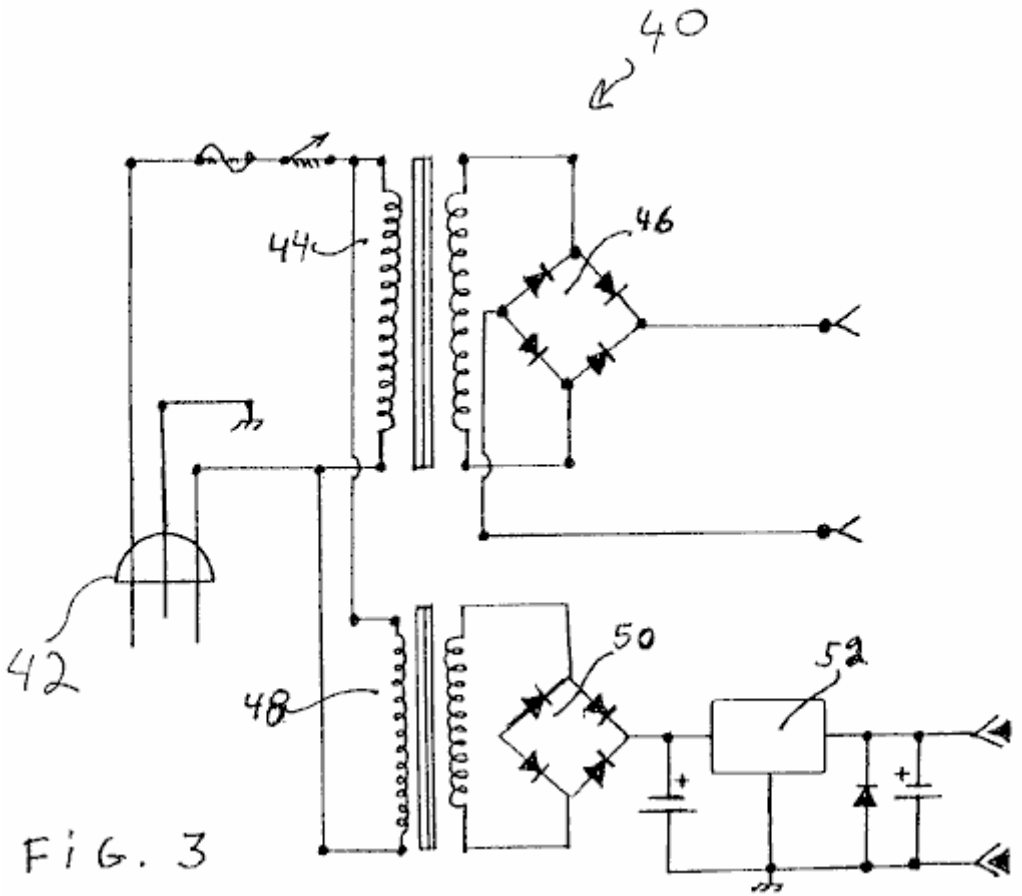


FIG. 3

Fig.3 is a schematic drawing of a conventional AC power supply that can be used to provide power to the pulse-charger of Fig.1 according to an embodiment of the invention.

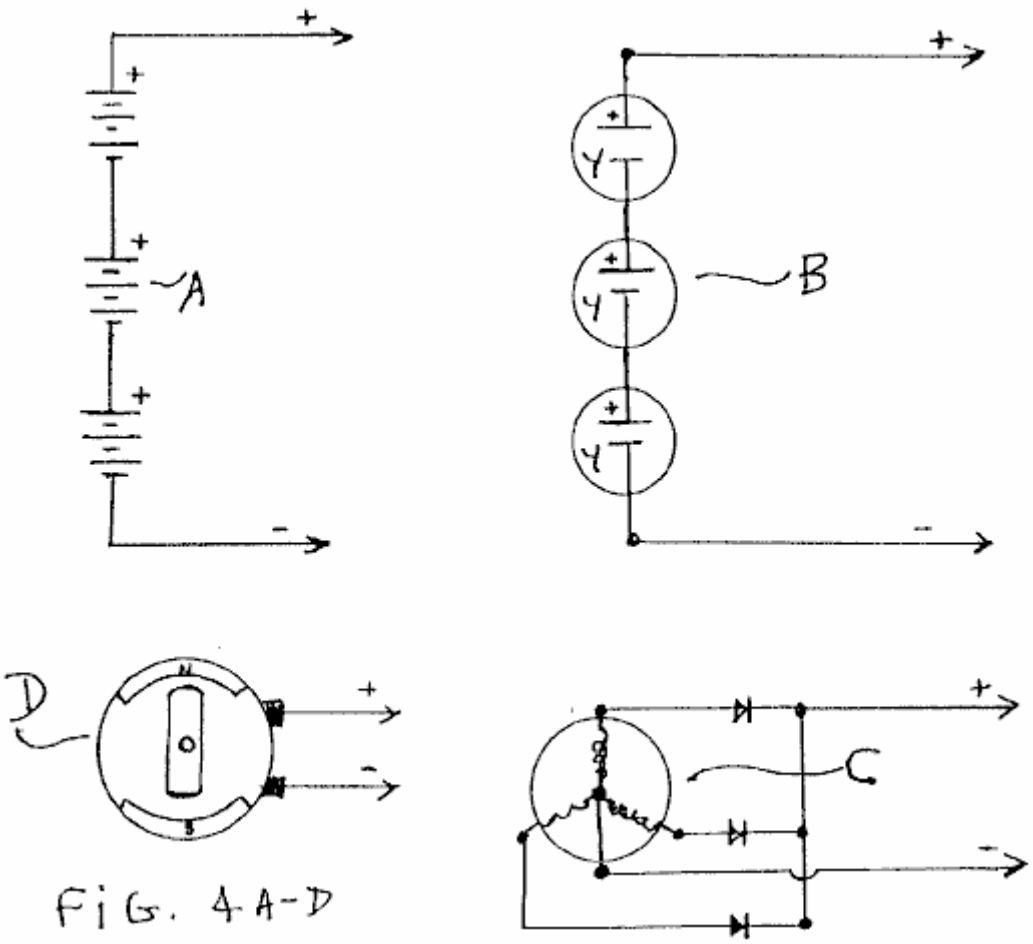


Fig.4A to Fig.4D are schematic drawings of other conventional power supplies that can be used to provide power to the pulse-charger of Fig.1 according to an embodiment of the invention.

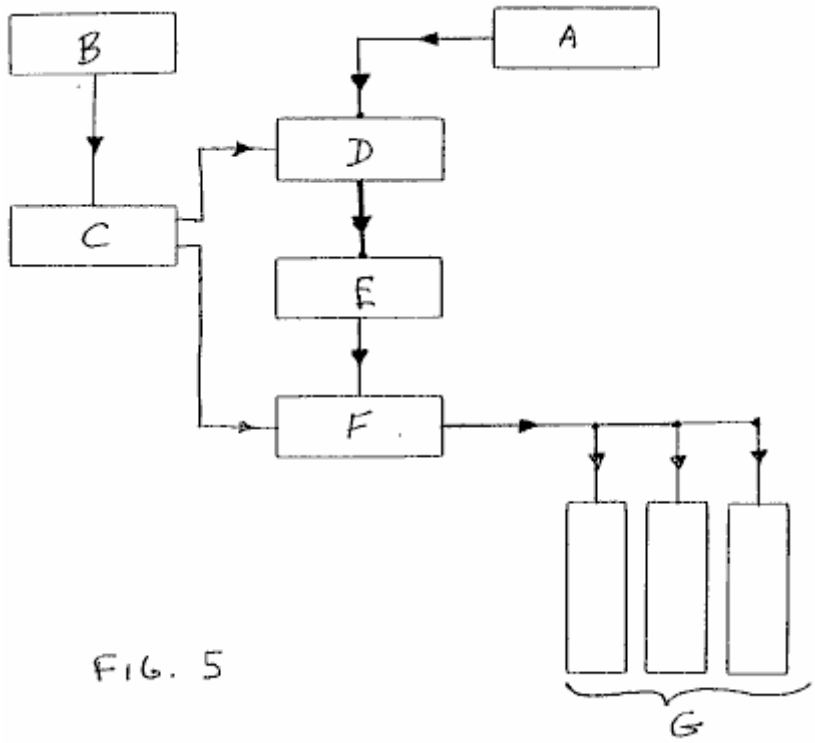


Fig.5 is a block diagram of the solid-state pulse-charger of Fig.1 according to an embodiment of the invention.

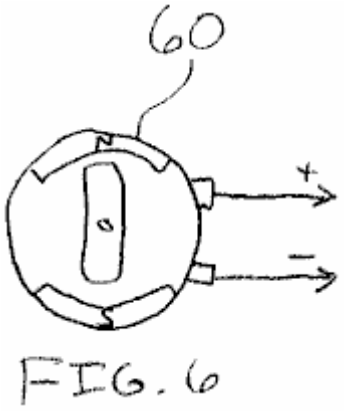


FIG. 6

Fig.6 is a diagram of a DC motor that the pulse-charger of Fig.1 can drive according to an embodiment of the invention.

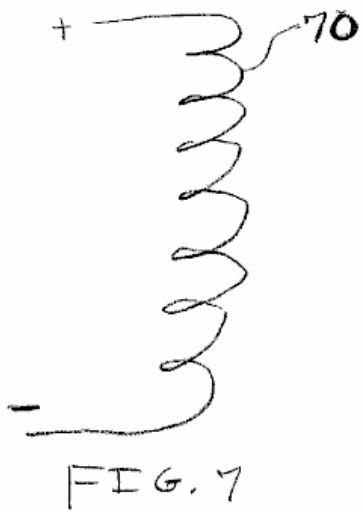


FIG. 7

Fig.7 is a diagram of a heating element that the pulse-charger of Fig.1 can drive according to an embodiment of the invention.

DETAILED DESCRIPTION OF THE INVENTION

An embodiment of the present invention is a device and method for a solid-state pulse-charger that uses a stored potential charge in a capacitor bank. The solid-state pulse-charger comprises a combination of elements and circuitry to capture and store available energy in a capacitor bank. The stored energy in the capacitors is then pulse-charged into the battery to be charged. In one version of this embodiment, there is a first momentary disconnection between the charger and the battery receiving the charge during the charge phase of the cycle, and a second momentary disconnection between the charger and the input energy source during the discharge phase of the cycle.

As a starting point, and an arbitrary method in describing this device and method, the flow of an electrical signal or current will be tracked from the primary input energy to final storage in the battery receiving the pulse charge.

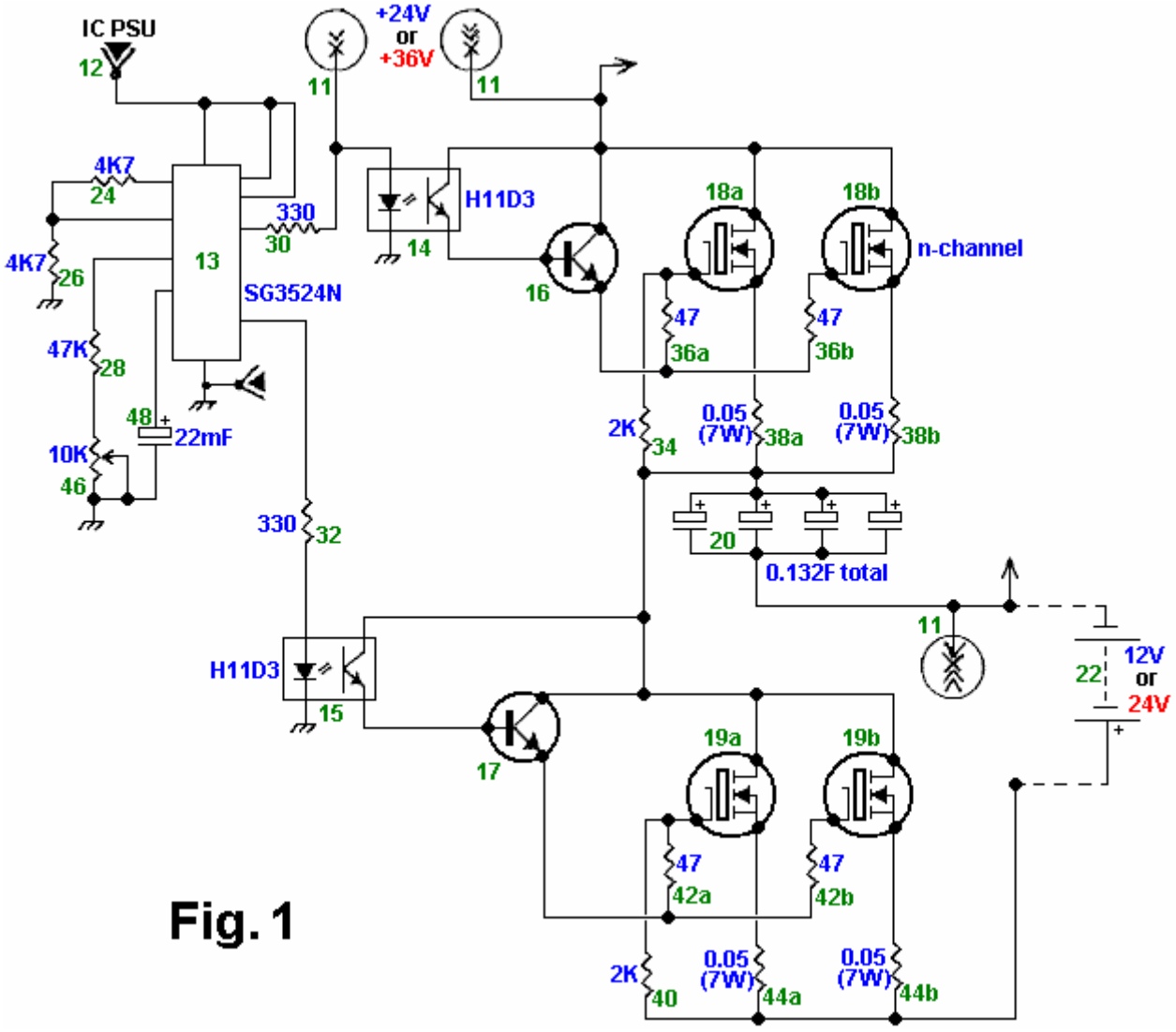


Fig. 1

Fig.1 is a schematic drawing of the solid-state pulse-charger according to an embodiment of the invention. As shown in **Fig.1**, the primary input energy source to the pulse-charger is a power supply **11**, examples of which are shown in **Fig.2**, **Fig.3**, and **Figs.4A-4D**. A 12-volt battery, as a low voltage energy source **12**, drives a dual switching means of control such as a logic chip or a pulse-width modulator (PWM) **13**.

Alternatively, the voltage from the power supply **11** may be converted to a voltage suitable to power the PWM **13**. The PWM **13** may be an SG3524N logic chip, and functions as an oscillator or timer to drive a 2-channel output with "on/off" switches that are connected when on to either a first optical isolator **14**, or alternatively, to a second optical isolator **15**. The first and second optical isolators **14** and **15** may be H11D3 optical isolators. When the logic chip **13** is connected to a first channel, it is disconnected from a second channel, thus resulting in two phases of signal direction; phase 1, a charge phase, and phase 2, a discharge phase.

When the logic chip **13** is switched to the charge phase, the signal flows to the first optical isolator **14**. From the optical isolator **14**, the signal continues its flow through a first NPN power transistor **16** that activates an N-channel MOSFET **18a** and an N-channel MOSFET **18b**. Current flowing through the MOSFETs **18a** and **18b** builds up a voltage across a capacitor bank **20**, thereby completing the charge phase of the switching activity.

The discharge phase begins when the logic chip **13** is switched to the second channel, with current flowing to the second optical isolator **15** and then through a second NPN power transistor **17**, which activates an N-channel MOSFET **19a** and an N-channel MOSFET **19b**. After the logic chip **13** closes the first channel and opens the second channel, the potential charge in the capacitor bank **20** is free floating between the power supply **11**, from which the capacitor bank **20** is now disconnected, and then connected to a battery **22** to receive the charge. It is at this point in time that the potential charge in the capacitor bank **20** is discharged through a high-energy pulse into the battery **22** or, a bank (not shown) of batteries. The discharge phase is completed once the battery **22** receives the charge. The logic chip **13** then switches the second channel closed and opens the first channel thus completing the charge-discharge cycle. The cycle is repetitive with the logic chip **13** controlling the signal

direction into either channel one to the capacitor bank, or to channel two to the battery **22** from the capacitor bank. The battery **22** is given a momentary rest period without a continuous current during the charge phase.

The component values for the described embodiment are as follows. The resistors **24**, **26**, . . . **44b** have the following respective values: 4.7K, 4.7K, 47K, 330, 330, 2K, 47, 47, 0.05(7W), 0.05(7W), 2K, 47, 47, 0.05(7 W), and 0.05(7W). The potentiometer **46** is 10K, the capacitor **48** is 22 mF, and the total capacitance of the capacitor bank **20** is 0.132F. The voltage of the battery **22** is between 12-24 V, and the voltage of the power supply **11** is 24-50 V such that the supply voltage is approximately 12-15 V higher than the battery voltage.

Other embodiments of the pulse-charger are contemplated. For example, the bipolar transistors **16** and **17** may be replaced with field-effect transistors, and the transistors **18a**, **18b**, **19a**, and **19b** may be replaced with bipolar or insulated-gate bipolar (IGBT) transistors. Furthermore, one can change the component values to change the cycle time, the peak pulse voltage, the amount of charge that the capacitor bank **20** delivers to the battery **22**, etc. In addition, the pulse-charger can have one or more than two transistors **18a** and **18b**, and one or more than two transistors **19a** and **19b**.

Still referring to **Fig.1**, the operation of the above-discussed embodiment of the pulse-charger is discussed. To begin the first phase of the cycle during which the capacitor bank **20** is charged, the logic circuit **13** deactivates the isolator **15** and activates the isolator **14**. Typically, the circuit **13** is configured to deactivate the isolator **15** before or at the same time that it activates the isolator **14**, although the circuit **13** may be configured to deactivate the isolator **15** after it activates the isolator **14**.

Next, the activated isolator **14** generates a base current that activates the transistor **16**, which in turn generates a current that activates the transistors **18a** and **18b**. The activated transistors **18a** and **18b** charge the capacitors in the bank **20** to a charge voltage equal or approximately equal to the voltage of the power supply **11** less the lowest threshold voltage of the transistors **18a** and **18b**. To begin the second phase of the cycle during which the capacitor bank **20** pulse charges the battery **22**, the logic circuit **13** deactivates the isolator **14** and activates the isolator **15**. Typically, the circuit **13** is configured to deactivate the isolator **14** before or at the same time that it activates the isolator **15**, although the circuit **13** may be configured to deactivate the isolator **14** after it activates the isolator **15**.

Next, the activated isolator **15** generates a base current that activates the transistor **17**, which in turn generates a current that activates the transistors **19a** and **19b**. The activated transistors **19a** and **19b** discharge the capacitors in the bank **20** into the battery **22** until the voltage across the bank **20** is or is approximately equal to the voltage across the battery **22** plus the lowest threshold voltage of the transistors **19a** and **19b**. Alternatively, the circuit **13** can deactivate the isolator **15** at a time before the bank **20** reaches this level of discharge. Because the resistances of the transistors **19a** and **19b**, the resistors **44a** and **44b**, and the battery **22** are relatively low, the capacitors in the bank **20** discharge rather rapidly, thus delivering a pulse of current to charge the battery **22**. For example, where the pulse-charger includes components having the values listed above, the bank **20** delivers a pulse of current having a duration of about 100 ms and a peak of about 250 A.

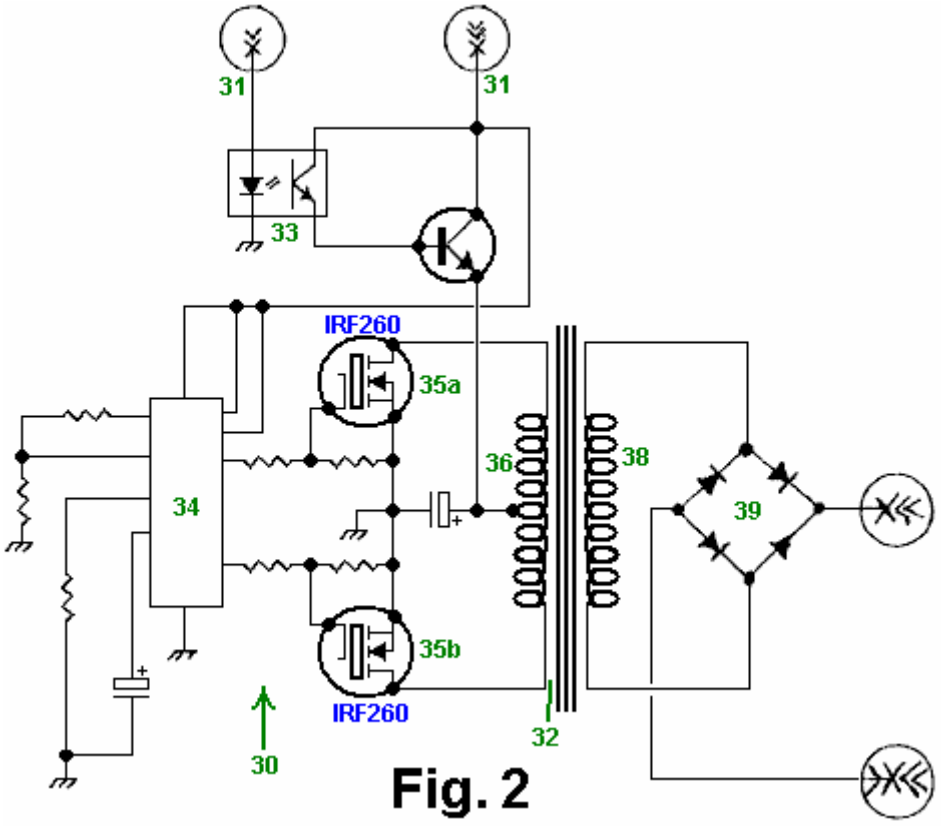


Fig. 2

Fig.2 is a schematic drawing of a conventional DC-to-DC converter **30** that can be used as the power supply **11** of **Fig.1** according to an embodiment of the invention. A DC-to-DC converter converts a low DC voltage to a higher DC voltage or vice-versa. Therefore, such a converter can convert a low voltage into a higher voltage that the pulse-charger of **Fig.1** can use to charge the capacitor bank **20** (**Fig.1**). More specifically, the converter **30** receives energy from a source **31** such as a 12-volt battery. An optical isolator sensor **33** controls an NPN power transistor which provides a current to a primary coil **36** of a power transformer **32**. A logic chip or pulse width modulator (PWM) **34** alternately switches on and off an IRF260 first N-channel MOSFET **35a** and an IRF260 second N-channel MOFSET **35b** such that when the MOSFET **35a** is on the MOSFET **35b** is off and vice-versa. Consequently, the switching MOSFETs **35a** and **35b** drive respective sections of the primary coil **36** to generate an output voltage across a secondary coil **38**. A full-wave bridge rectifier **39** rectifies the voltage across the secondary coil **38**, and this rectified voltage is provided to the pulse-charger of **Fig.1**. Furthermore, the secondary coil **38** can be tapped to provide a lower voltage for the PWM **13** of **Fig.1** such that the DC-to-DC converter **30** can be used as both the power supply **11** and the low-voltage supply **12** of **Fig.1**.

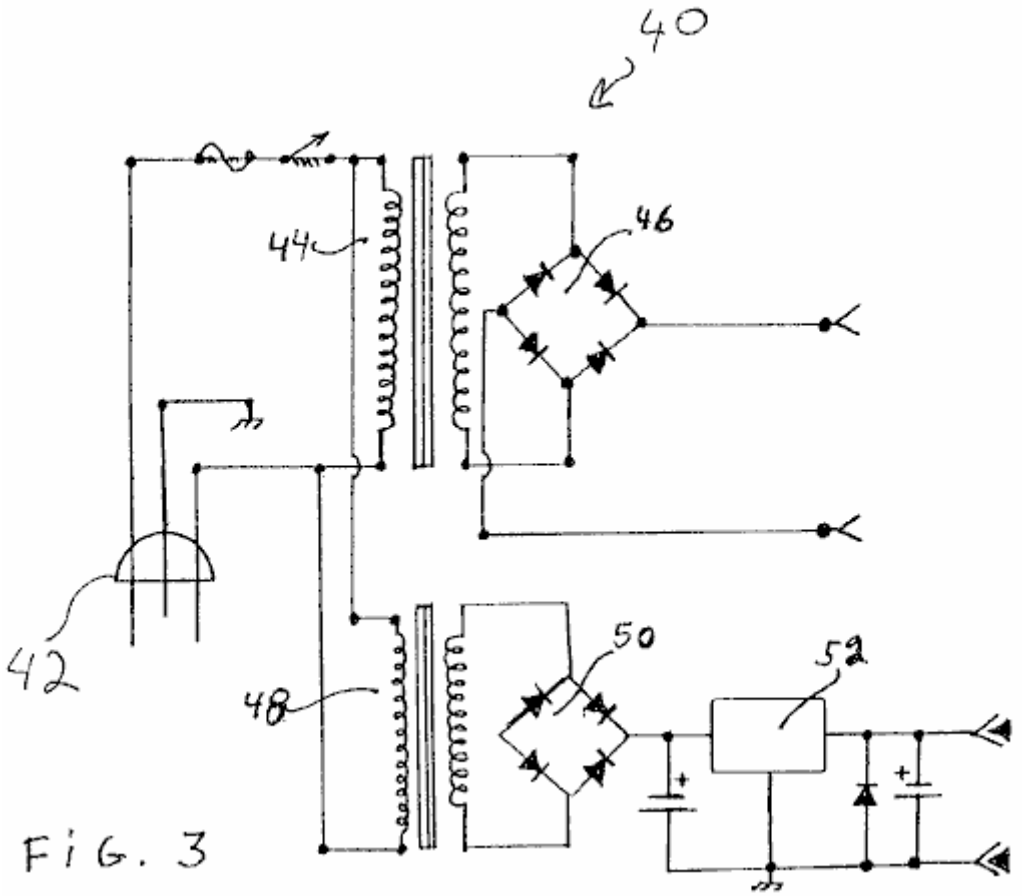


Fig.3 is a schematic drawing of an AC power supply **40** that can be used as both the power supply **11** and the power supply **12** of **Fig.1** according to an embodiment of the invention. The power input **42** to the supply **40** is 120V AC. A first transformer **44** and full-wave rectifier **46** compose the supply **11**, and a second transformer **48**, full-wave rectifier **50**, and voltage regulator **52** compose the supply **12**.

Fig.4A to **Fig.4D** are schematic drawings of various conventional primary energy input sources which can be used as the supply **11** and/or the supply **12** of **Fig.1** according to an embodiment of the invention. **Fig.4A** is a schematic drawing of serially coupled batteries. **Fig.4B** is a schematic drawing of serially-coupled solar cells. **Fig.4C** is a schematic drawing of an AC generator, and **Fig.4D** is a schematic drawing of a DC generator.

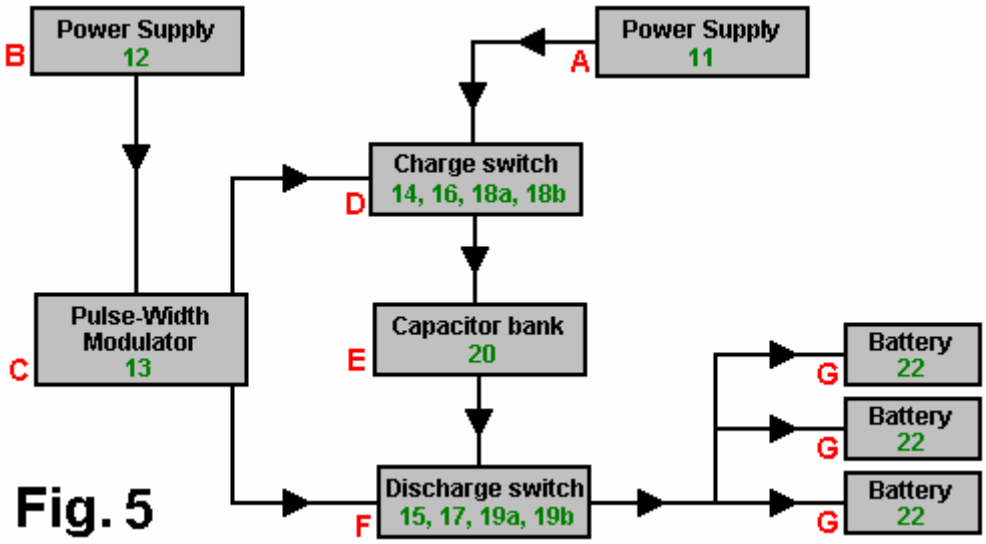


Fig. 5

Fig.5 is a block diagram of the solid-state pulse-charger of **Fig.1** according to an embodiment of the invention. Block **A** is the power supply **11**, which can be any suitable power supply such as those shown in **Fig.2**, **Fig.3**, **Figs.4A-4D**. Block **B** is the power supply **12**, which can be any suitable power supply such as a 12V DC supply or the supply shown in **Fig.3**. Block **C** is the PWM **13** and its peripheral components. Block **D** is the charge switch that includes the first optical isolator chip **14**, the first NPN power transistor **16**, the first set of two N-

channel MOSFETs **18a** and **18b**, and their peripheral resistors. Block **E** is the capacitor bank **20**. Block **F** is the discharge switch that includes the second optical isolator chip **15**, the second NPN power transistor **17**, the second set of two N-channel MOSFETs **19a** and **19b**, and their peripheral resistors. Block **G** is the battery or battery bank **22** which is being pulse-charged.

A unique feature that distinguishes one embodiment of the pulse-charger described above, from conventional chargers is the method charging the battery with pulses of current instead of with a continuous current. Consequently, the battery is given a reset period between pulses.

Fig.6 is a diagram of a DC motor **60** that the pulse-charger of **Fig.1** can drive according to an embodiment of the invention. Specifically, one can connect the motor **60** in place of the battery **22** (**Fig.1**) such that the pulse-charger drives the motor with pulses of current. Although one need not modify the pulse-charger to drive the motor **60**, one can modify it to make it more efficient for driving the motor. For example, one can modify the values of the resistors peripheral to the PWM **13** (**Fig.1**) to vary the width and peak of the drive pulses from the capacitor bank **20** (**Fig.1**).

Fig.7 is a diagram of a heating element **70**, such as a dryer or water-heating element, that the pulse-charger of **Fig.1** can drive according to an embodiment of the invention. Specifically, one can connect the heating element **70** in place of the battery **22** (**Fig.1**) such that the pulse-charger drives the element with pulses of current. Although one need not modify the pulse-charger to drive the element **70**, one can modify it to make it more efficient for driving the element. For example, one can modify the values of the resistors peripheral to the PWM **13** (**Fig.1**) to vary the width and peak of the drive pulses from the capacitor bank **20** (**Fig.1**).

In the embodiments discussed above, specific electronic elements and components are used. However, it is known that a variety of available transistors, resistors, capacitors, transformers, timing components, optical isolators, pulse width modulators, MOSFETs, and other electronic components may be used in a variety of combinations to achieve an equivalent result. Finally, although the invention has been described with reference of particular means, materials and embodiments, it is to be understood that the invention is not limited to the particulars disclosed and extends to all equivalents within the scope of the claims.

CLAIMS

1. A solid-state pulse battery charger wherein input power from a primary source is stored as a potential charge in a capacitor bank, said capacitor bank then disconnected from said input power source through a dual timing means, said capacitor then connected to a battery to receive the potential charge, the charge then discharged into said battery from said capacitor, said battery then disconnected from said capacitor through said dual timing means, said capacitor then re-connected to said input power source completing a two phase switching cycle comprising:

a. a means for providing input power;

b. a means for timing a signal and a current flow in two phases, a charge phase and a discharge phase, through either a first channel output for charging said capacitor bank, or a second channel output for discharging stored energy from said capacitor into said battery, the current flowing from said first channel output through a first optical isolator and through a first NPN power transistor, said first transistor activating a first pair of N-channel MOSFETs with voltage stored as the potential charge in said capacitor bank, said capacitor disconnecting from said input power means by said timing means;

c. said means for timing current flow connecting to said second channel output, current flowing from said second channel through a second optical isolator and through a second NPN power transistor, said second transistor activating a second pair of N-channel MOSFETs, said capacitor connecting to said battery, the potential charge discharging into said battery, said timing means disconnecting said capacitor from said battery, and connecting said capacitor to said power means.

2. The pulse-charger of claim 1 wherein the means for providing input power is an AC voltage current.

3. The pulse-charger of claim 1 wherein the means for providing input power is a battery.

4. The pulse-charger of claim 1 wherein the means for providing input power is a DC generator.

5. The pulse-charger of claim 1 wherein the means for providing input power is an AC generator.

6. The pulse-charger of claim 1 wherein the means for providing input power is a solar cell.

7. The pulse-charger of claim 1 wherein the means for providing input power is a DC-to-DC inverter.

8. The pulse-charger of claim 1 wherein the means for timing a signal is a pulse width modulator, said modulator an SG3524N logic chip.

9. The pulse-charger of claim 1 wherein the optical isolator is an H11D3 isolator.

10. The pulse-charger of claim 1 wherein the NPN power transistor is an MJE15024 transistor.

11. The pulse-charger of claim 1 wherein the N-channel MOSFET is a IRF260 MOSFET.

12. A solid-state pulsed battery charger wherein input power from a primary source is stored as a potential charge in a capacitor bank, said capacitor then disconnected from said input power source through a dual timing means, said capacitor then connected to a battery to receive the potential charge, the charge then discharged into said battery from said capacitor, said battery then disconnected from said capacitor through said dual timing means, said capacitor then reconnected to said input power source completing a two phase cycle comprising:

a. a means for providing said input power, said means either an AC voltage current, or a battery, or a DC generator, or an AC generator, or a solar cell, or a DC-to-DC inverter;

b. a means for timing a signal and a current flow, said timing means a pulse width modulator, logic chip SG3524N, the current flowing through either a first channel output, or a second channel output, the current flowing from said first channel output through a first optical isolator, said isolator an H11D3, and through a first NPN power transistor, said transistor an MJE15024, said first transistor activating a first pair of N-channel MOSFETs, said MOSFETs an IRF260, with current voltage stored as the potential charge in said capacitor bank, said capacitor disconnecting from said input power means by said logic chip;

c. said timing logic chip connecting to said second channel output, current flowing from said second channel through a second optical isolator, said isolator an H11D3, and through a second NPN power transistor, said second transistor an MJE15024, and activating a second pair of N-channel MOSFETs, said MOSFETs an IRF260, with current voltage stored as the potential charge in said capacitor bank, said capacitor disconnecting from said input power means by said logic chip, said capacitor connecting to said battery, the potential charge discharging into said battery, said timing means disconnecting said capacitor from said battery and connecting said capacitor to said power means.

13. A method of making a solid-state pulse battery charger wherein input power from a primary source is stored as a potential charge in a capacitor bank, said capacitor disconnected from said input power source through a dual timing means, said capacitor connected to a battery to receive the potential charge, said charge discharged into said battery from said capacitor, said battery disconnected from said capacitor through said dual timing means, said capacitor reconnected to said input power source completing a two phase cycle comprising the steps of:

a. providing a source of input power;

b. connecting a means for dual-timing said charger to control a signal or current flow through a first channel output comprising a first optical isolator, a first NPN power transistor and a first pair of N-channel MOSFETs;

c. capturing energy from said current and storing said energy in said capacitor bank thereby charging said capacitor;

d. switching the flow of said current using said timing device to a second channel comprising a second optical isolator, a second NPN power transistor and a second pair of N-channel MOSFETs, thus disconnecting said capacitor from said power source and connecting said capacitor to said battery;

e. discharging the potential charge into said battery;

f. switching the flow of the current using said timing device to said power source and said first channel to complete said cycle.

14. The pulse-charger of claim 13 wherein the means for providing input power is an AC voltage current.

15. The pulse-charger of claim 13 wherein the means for providing input power is a battery.

16. The pulse-charger of claim 13 wherein the means for providing input power is a DC generator.

17. The pulse-charger of claim 13 wherein the means for providing input power is an AC generator.

18. The pulse-charger of claim 13 wherein the means for providing input power is a solar cell.

19. The pulse-charger of claim 13 wherein the means for providing input power is a DC-to-DC inverter.

20. The pulse-charger of claim 13 wherein the means for timing a signal is a pulse width modulator, said modulator an SG3524N logic chip.

21. The pulse-charger of claim 13 wherein the optical isolator is an H11D3 isolator.

22. The pulse-charger of claim 13 wherein the NPN power transistor is an MJE15024 transistor.

23. The pulse-charger of claim 13 wherein the N-channel MOSFET is a IRF260 MOSFET.

24. A battery charger, comprising:
a supply node;
a charge node;
a switch circuit coupled to the supply and the charge nodes and operable to, allow a battery-charge current to flow into the charge node during a
battery-charge period, and prohibit the battery-charge current from flowing into the charge node during
a battery-rest period.

25. The battery charger of claim 24, further comprising:
a charge-storage device coupled to the switch circuit; and
wherein the switch circuit is operable to, allow the battery-charge current to flow from the charge-storage device into the charge node during the battery-charge period, and charge the charge-storage device during the battery-rest period.

26. The battery charger of claim 24, further comprising:
a capacitor coupled to the switch circuit; and
wherein the switch circuit is operable to, allow the battery-charge current to from the capacitor into the charge node
during the battery-charge period, and charge the capacitor during the battery-rest period.

27. A method, comprising:
charging a battery during a first period of a charge cycle; and
prohibiting the charging of the battery during a second period of the charge cycle.

28. The method of claim 27 wherein:
charging the battery comprises charging the battery with a charge current during the first period of the charge cycle; and
prohibiting the charging of the battery comprises prohibiting the charge current from flowing into the battery during the second period of the charge cycle.

29. The method of claim 27 wherein:
charging the battery comprises discharging a capacitor into the battery during the first period of the charge cycle; and
prohibiting the charging of the battery comprises uncoupling the capacitor from the battery during the second period of the charge cycle.

30. The method of claim 27, further comprising:
wherein charging the battery comprises discharging a capacitor into the battery during the first period of the charge cycle;
wherein prohibiting the charging of the battery comprises uncoupling the capacitor from the battery during the second period of the charge cycle; and
charging the capacitor during the second period of the charge cycle.

31. A method, comprising:
discharging a charge-storage device into a battery during a first period of a battery-charge cycle; and
uncoupling the charge-storage device from the battery and charging the charge-storage device during a second period of the battery-charge cycle.

32. The method of claim 31 wherein uncoupling the charge-storage device comprises uncoupling the charge-storage device from the battery before commencing charging of the charge-storage device.

33. The method of claim 31 wherein uncoupling the charge-storage device comprises uncoupling the charge-storage device from the battery after commencing charging of the charge-storage device.

34. The method of claim 31 wherein uncoupling the charge-storage device comprises simultaneously uncoupling the charge-storage device from the battery and commencing charging of the charge-storage device.

Notes:

The following information is NOT part of John's patent. It is information intended to be helpful, but as it is not coming from John it must be considered to be opinion and not fact. In the above diagrams, the SG3524N integrated circuit is likely to be unfamiliar to many readers, and an examination of the specification sheet does not make it obvious which pin connections are used in John's circuit. The following pin connections are believed to be correct, but cannot be guaranteed.

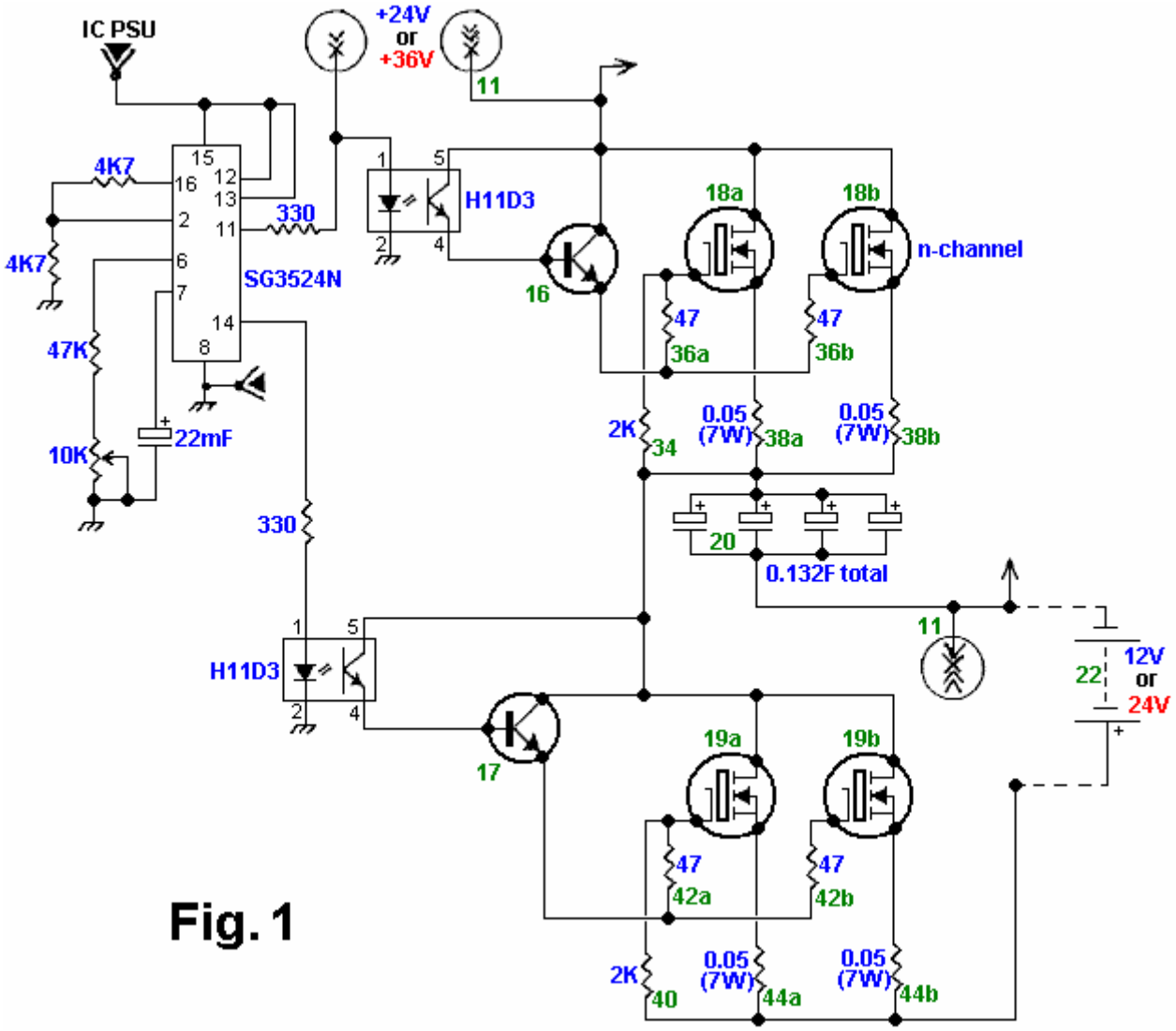
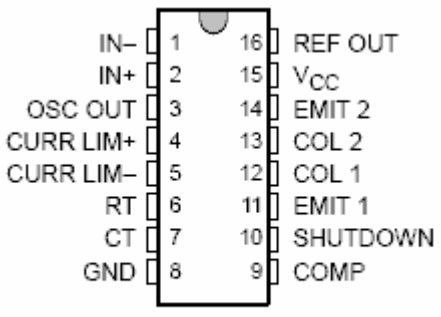


Fig. 1

In addition to these SG3524N pin connections, it is suggested that pins 1, 4 and 5 be connected to ground instead of just pin 8, and that a 100nF capacitor be connected from pin 9 to ground. Pins 3 and 10 are left unconnected. The pinouts for the chip are:



SG3524N

RICHARD WEIR and CARL NELSON

US Patent 7,033,406 25th April 2006 Inventors: Richard Weir and Carl Nelson

ELECTRICAL-ENERGY-STORAGE UNIT UTILISING CERAMIC AND INTEGRATED-CIRCUIT TECHNOLOGIES FOR REPLACEMENT OF ELECTROCHEMICAL BATTERIES

This patent shows an electrical storage method which is reputed to power an electric car for a 500 mile trip on a charge taking only five minutes to complete. This document is a very slightly re-worded copy of the original. It has been pointed out by Mike Furness that while a five minute recharge is feasible, it is not practical, calling for cables with a six-inch diameter. Also, the concept of recharging stations as suggested is also rather improbable as the electrical supply needed would rival that of a power station. However, if the charging time were extended to night time, then it would allow substantial driving range during the day time.

ABSTRACT

An Electrical-Energy-Storage Unit (EESU) has as a basis material a high-permittivity, composition-modified barium titanate ceramic powder. This powder is double coated with the first coating being aluminium oxide and the second coating calcium magnesium aluminosilicate glass. The components of the EESU are manufactured with the use of classical ceramic fabrication techniques which include screen printing alternating multi-layers of nickel electrodes and high-permittivity composition-modified barium titanate powder, sintering to a closed-pore porous body, followed by hot-isostatic pressing to a void-free body. The components are configured into a multi-layer array with the use of a solder-bump technique as the enabling technology so as to provide a parallel configuration of components that has the capability to store electrical energy in the range of 52 kWH. The total weight of an EESU with this range of electrical energy storage is about 336 pounds.

BACKGROUND OF THE INVENTION

1. Field of the Invention

This invention relates generally to energy-storage devices, and relates more particularly to high-permittivity ceramic components utilised in an array configuration for application in ultra high electrical-energy storage devices.

2. Description of the Relevant Art

The internal-combustion-engine (ICE) powered vehicles have as their electrical energy sources a generator and battery system. This electrical system powers the vehicle accessories, which include the radio, lights, heating, and air conditioning. The generator is driven by a belt and pulley system and some of its power is also used to recharge the battery when the ICE is in operation. The battery initially provides the required electrical power to operate an electrical motor that is used to turn the ICE during the starting operation and the ignition system.

The most common batteries in use today are:
Flooded lead-acid,
Sealed gel lead-acid,
Nickel-Cadmium (Ni-Cad),
Nickel Metal Hydride (NiMH), and
Nickel-Zinc (Ni-Z).

References on the subject of electrolchemical batteries include the following:

Guardian, Inc., "Product Specification": Feb. 2, 2001;
K. A. Nishimura, "NiCd Battery", Science Electronics FAQ V1.00: Nov. 20, 1996;
Ovonics, Inc., "Product Data Sheet": no date;
Evercel, Inc., "Battery Data Sheet—Model 100": no date;
S. R. Ovshinsky et al., "Ovonics NiMH Batteries: The Enabling Technology for Heavy-Duty Electrical and Hybrid Electric Vehicles", Ovonics publication 2000-01-3108: Nov. 5, 1999;
B. Dickinson et al., "Issues and Benefits with Fast Charging Industrial Batteries", AeroVeronment, Inc. article: no date.

Each specific type of battery has characteristics, which make it either more or less desirable to use in a specific application. Cost is always a major factor and the NiMH battery tops the list in price with the flooded lead-acid battery being the most inexpensive. Evercel manufactures the Ni-Z battery and by a patented process, with the

claim to have the highest power-per-pound ratio of any battery. See Table 1 below for comparisons among the various batteries. What is lost in the cost translation is the fact that NiMH batteries yield nearly twice the performance (energy density per weight of the battery) than do conventional lead-acid batteries. A major drawback to the NiMH battery is the very high self-discharge rate of approximately 5% to 10% per day. This would make the battery useless in a few weeks. The Ni-Cad battery and the lead-acid battery also have self-discharge but it is in the range of about 1% per day and both contain hazardous materials such as acid or highly toxic cadmium. The Ni-Z and the NiMH batteries contain potassium hydroxide and this electrolyte in moderate and high concentrations is very caustic and will cause severe burns to tissue and corrosion to many metals such as beryllium, magnesium, aluminium, zinc, and tin.

Another factor that must be considered when making a battery comparison is the recharge time. Lead-acid batteries require a very long recharge period, as long as 6 to 8 hours. Lead-acid batteries, because of their chemical makeup, cannot sustain high current or voltage continuously during charging. The lead plates within the battery heat rapidly and cool very slowly. Too much heat results in a condition known as "gassing" where hydrogen and oxygen gases are released from the battery's vent cap. Over time, gassing reduces the effectiveness of the battery and also increases the need for battery maintenance, i.e., requiring periodic de-ionised or distilled water addition. Batteries such as Ni-Cad and NiMH are not as susceptible to heat and can be recharged in less time, allowing for high current or voltage changes which can bring the battery from a 20% state of charge to an 80% state of charge in just 20 minutes. The time to fully recharge these batteries can be more than an hour. Common to all present day batteries is a finite life, and if they are fully discharged and recharged on a regular basis their life is reduced considerably.

SUMMARY OF THE INVENTION

In accordance with the illustrated preferred embodiment, the present invention provides a unique electrical-energy-storage unit that has the capability to store ultra high amounts of energy.

One aspect of the present invention is that the materials used to produce the energy-storage unit, EESU, are not explosive, corrosive, or hazardous. The basis material, a high-permittivity calcined composition-modified barium titanate powder is an inert powder and is described in the following references: S. A. Bruno, D. K. Swanson, and I. Burn, J. Am Ceram. Soc. 76, 1233 (1993); P. Hansen, U.S. Pat. No. 6,078,494, issued Jun. 20, 2000. The most cost-effective metal that can be used for the conduction paths is nickel. Nickel as a metal is not hazardous and only becomes a problem if it is in solution such as in deposition of electroless nickel. None of the EESU materials will explode when being recharged or impacted. Thus the EESU is a safe product when used in electric vehicles, buses, bicycles, tractors, or any device that is used for transportation or to perform work. It could also be used for storing electrical power generated from solar voltaic cells or other alternative sources for residential, commercial, or industrial applications. The EESU will also allow power averaging of power plants utilising SPVC or wind technology and will have the capability to provide this function by storing sufficient electrical energy so that when the sun is not shinning or the wind is not blowing they can meet the energy requirements of residential, commercial, and industrial sites.

Another aspect of the present invention is that the EESU initial specifications will not degrade due to being fully discharged or recharged. Deep cycling the EESU through the life of any commercial product that may use it will not cause the EESU specifications to be degraded. The EESU can also be rapidly charged without damaging the material or reducing its life. The cycle time to fully charge a 52 kWH EESU would be in the range of 4 to 6 minutes with sufficient cooling of the power cables and connections. This and the ability of a bank of EESUs to store sufficient energy to supply 400 electric vehicles or more with a single charge will allow electrical energy stations that have the same features as the present day gasoline stations for the ICE cars. The bank of EESUs will store the energy being delivered to it from the present day utility power grid during the night when demand is low and then deliver the energy when the demand hits a peak. The EESU energy bank will be charging during the peak times but at a rate that is sufficient to provide a full charge of the bank over a 24-hour period or less. This method of electrical power averaging would reduce the number of power generating stations required and the charging energy could also come from alternative sources. These electrical-energy-delivery stations will not have the hazards of the explosive gasoline.

Yet another aspect of the present invention is that the coating of aluminium oxide and calcium magnesium aluminosilicate glass on calcined composition-modified barium titanate powder provides many enhancement features and manufacturing capabilities to the basis material. These coating materials have exceptional high voltage breakdown and when coated on to the above material will increase the breakdown voltage of ceramics comprised of the coated particles from 3×10^6 V/cm of the uncoated basis material to around 5×10^6 V/cm or higher. The following reference indicates the dielectric breakdown strength in V/cm of such materials: J. Kuwata et al., "Electrical Properties of Perovskite-Type Oxide Thin-Films Prepared by RF Sputtering", Jpn. J. Appl. Phys., Part 1, 1985, 24(Suppl. 24-2, Proc. Int. Meet. Ferroelectr., 6th), 413-15. This very high voltage breakdown assists in allowing the ceramic EESU to store a large amount of energy due to the following: Stored energy $E = CV^2 / 2$,

Formula 1, as indicated in F. Sears et al., "Capacitance-Properties of Dielectrics", University Physics, Addison-Wesley Publishing Company, Inc.: Dec. 1957: pp 468-486, where C is the capacitance, V is the voltage across the EESU terminals, and E is the stored energy. This indicates that the energy of the EESU increases with the square of the voltage. **Fig.1** indicates that a double array of 2230 energy storage components 9 in a parallel configuration that contain the calcined composition-modified barium titanate powder. Fully densified ceramic components of this powder coated with 100 Angstrom units of aluminium oxide as the first coating 8 and a 100 Angstrom units of calcium magnesium aluminosilicate glass as the second coating 8 can be safely charged to 3500 V. The number of components used in the double array depends on the electrical energy storage requirements of the application. The components used in the array can vary from 2 to 10,000 or more. The total capacitance of this particular array 9 is 31 F which will allow 52,220 W·h of energy to be stored as derived by Formula 1.

These coatings also assist in significantly lowering the leakage and ageing of ceramic components comprised of the calcined composition-modified barium titanate powder to a point where they will not effect the performance of the EESU. In fact, the discharge rate of the ceramic EESU will be lower than 0.1% per 30 days which is approximately an order of magnitude lower than the best electrochemical battery.

A significant advantage of the present invention is that the calcium magnesium aluminosilicate glass coating assists in lowering the sintering and hot-isostatic-pressing temperatures to 800^{O}C. This lower temperature eliminates the need to use expensive platinum, palladium, or palladium-silver alloy as the terminal metal. In fact, this temperature is in a safe range that allows nickel to be used, providing a major cost saving in material expense and also power usage during the hot-isostatic-pressing process. Also, since the glass becomes easily deformable and flowable at these temperatures it will assist in removing the voids from the EESU material during the hot-isostatic-pressing process. The manufacturer of such systems is Flow Autoclave Systems, Inc. For this product to be successful it is mandatory that all voids be removed to assist in ensuring that the high voltage breakdown can be obtained. Also, the method described in this patent of coating the calcium magnesium aluminosilicate glass ensures that the hot-isostatic-pressed double-coated composition-modified barium titanate high-relative-permittivity layer is uniform and homogeneous.

Yet another aspect of the present invention is that each component of the EESU is produced by screen-printing multiple layers of nickel electrodes with screening ink from nickel powder. Interleaved between nickel electrodes are dielectric layers with screening ink from calcined double-coated high-permittivity calcined composition-modified barium titanate powder. A unique independent dual screen-printing and layer-drying system is used for this procedure. Each screening ink contains appropriate plastic resins, surfactants, lubricants, and solvents, resulting in a proper rheology (the study of the deformation and flow of matter) for screen printing. The number of these layers can vary depending on the electrical energy storage requirements. Each layer is dried before the next layer is screen printed. Each nickel electrode layer **12** is alternately preferentially aligned to each of two opposite sides of the component automatically during this process as indicated in **Fig.2**. These layers are screen printed on top of one another in a continuous manner. When the specified number of layers is achieved, the component layers are then baked to obtain by further drying sufficient handling strength of the green plastic body. Then the array is cut into individual components to the specified sizes.

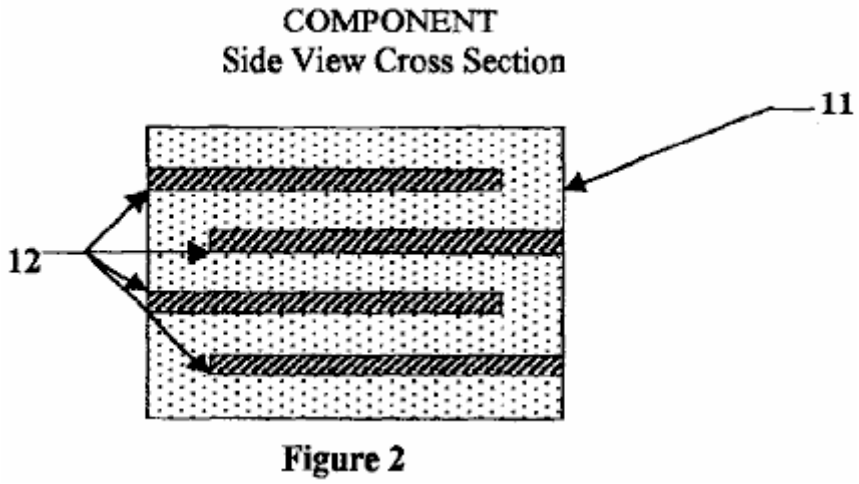


Figure 2

Alternatively, the dielectric powder is prepared by blending with plastic binders, surfactants, lubricants, and solvents to obtain a slurry with the proper rheology for tape casting. In tape casting, the powder-binder mixture is extruded by pressure through a narrow slit of appropriate aperture height for the thickness desired of the green plastic ceramic layer on to a moving plastic-tape carrier, known as a doctor-blade web coater. After drying, to develop sufficient handling strength of the green plastic ceramic layer, this layer is peeled away from the plastic-tape carrier. The green plastic ceramic layer is cut into sheets to fit the screen-printing frame in which the

electrode pattern is applied with nickel ink. After drying of the electrode pattern, the sheets are stacked and then pressed together to assure a well-bonded lamination. The laminate is then cut into components of the desired shape and size.

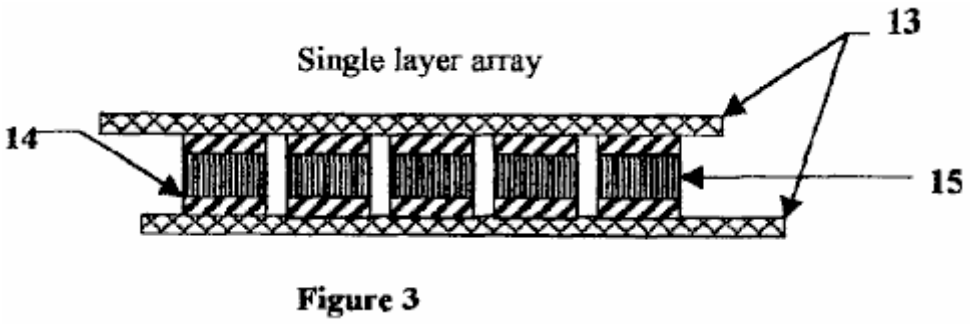


Figure 3

The components are treated for the binder-burnout and sintering steps. The furnace temperature is slowly ramped up to 350^{O}C and held for a specified length of time. This heating is accomplished over a period of several hours so as to avoid any cracking and delamination of the body. Then the temperature is ramped up to 850^{O}C and held for a specified length of time. After this process is completed the components are then properly prepared for the hot isostatic pressing at 700^{O}C and the specified pressure. This process will eliminate voids. After this process, the components are then side-lapped on the connection side to expose the preferentially aligned nickel electrodes **12**. Then these sides are dipped into ink from nickel powder that has been prepared to have the desired rheology. Then side conductors of nickel **14** are dipped into the same ink and then are clamped on to each side of the components **15** that have been dipped into the nickel powder ink. The components are then fired at 800^{O}C for 20 minutes to bond the nickel bars to the components as indicated in **Fig.3**. The components are then assembled into a first-level array, **Fig.3**, with the use of the proper tooling and solder-bump technology. Then the first-level arrays are assembled to form a second-level array, **Fig.4**, by stacking the first array layers on top of one another in a preferential mode. Then nickel bars **18** are attached on each side of the second array as indicated in **Fig.4**. Then the EESU is packaged to form its final assembly configuration.

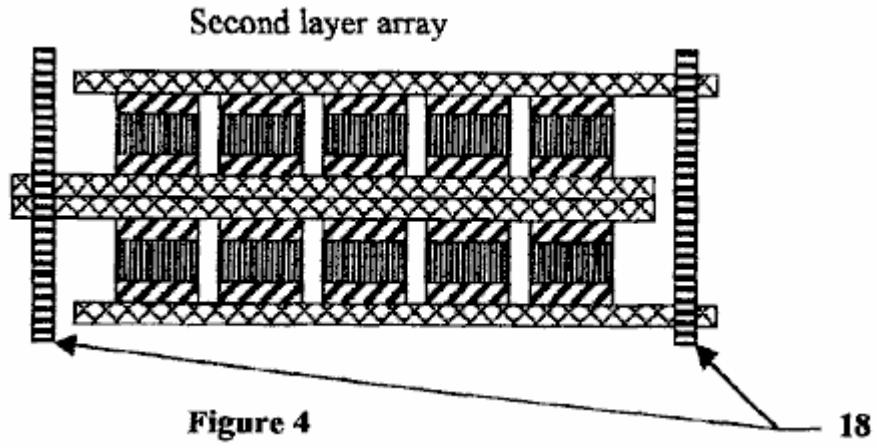


Figure 4

The features of this patent indicate that the ceramic EESU, as indicated in Table 1, outperforms the electrochemical battery in every parameter. This technology will provide mission-critical capability to many sections of the energy-storage industry.

TABLE 1
The parameters of each technology to store 52.2 kW · h of electrical energy
are indicated-(data as of February 2001 from manufacturer's specification sheets).

	NiMH	LA(Gel)	Ceramic EESU	Ni—Z
Weight (pounds)	1,716	3,646	336	1,920
Volume (cu. inch)	17,881	43,045	2,005	34,780
Discharge rate	5% in 30 days	1% in 30 days	0.1% in 30 days	1% in 30 days
Charging time (full)	1.5 hours	8.0 hours	3 to 6 minutes	1.5 hours
Life reduced with deep cycle use	moderate	high	none	moderate
Hazardous materials	Yes	Yes	None	Yes

This EESU will have the potential to revolutionise the electric vehicle (EV) industry, the storage and use of electrical energy generated from alternative sources with the present utility grid system as a backup source for residential, commercial, and industrial sites, and the electric energy point of sales to EVs. The EESU will replace the electrochemical battery in any of the applications that are associated with the above business areas or in any business area where its features are required.

The features and advantages described in the specifications are not all inclusive, and particularly, many additional features and advantages will be apparent to one of ordinary skill in the art in view of the description, specification and claims made here. Moreover, it should be noted that the language used in the specification has been principally selected for readability and instructional purposes, and may not have been selected to delineate or circumscribe the inventive subject matter, resort to the claims being necessary to determine such inventive subject matter.

BRIEF DESCRIPTION OF THE DRAWINGS

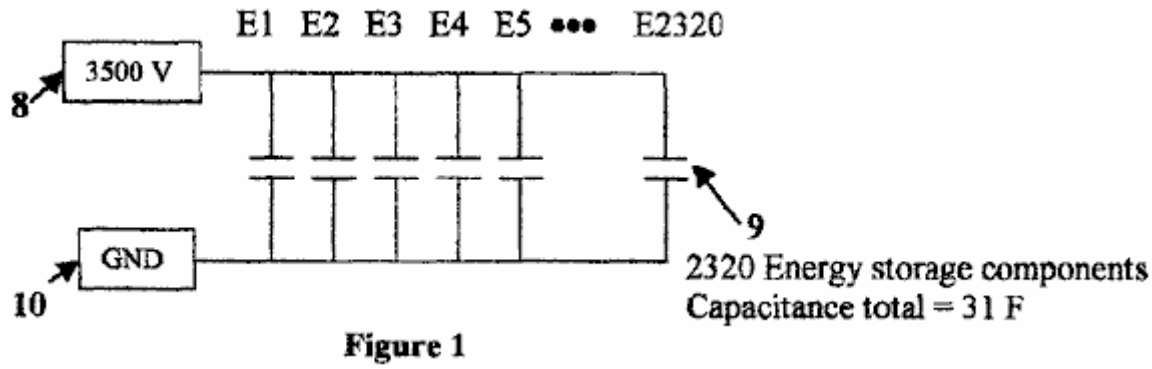


Figure 1

Fig.1 indicates a schematic of 2320 energy storage components **9** hooked up in parallel with a total capacitance of 31 Farads. The maximum charge voltage **8** of 3500 V is indicated with the cathode end of the energy storage components **9** hooked to system ground **10**.

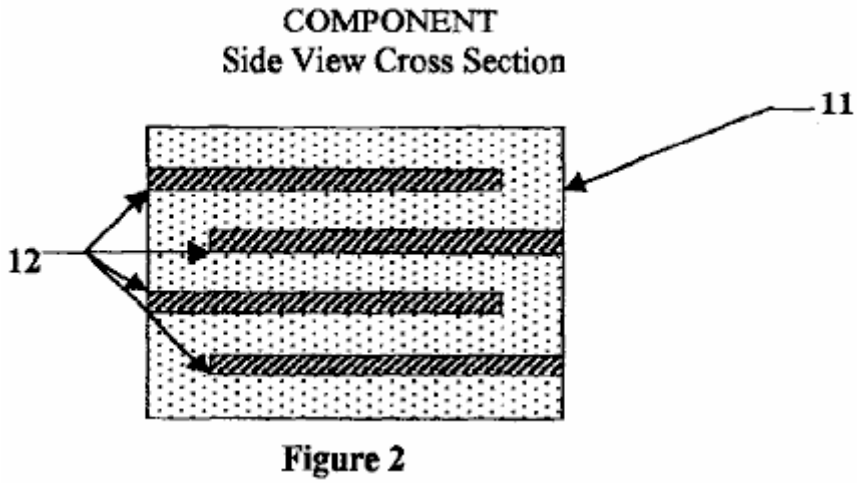


Figure 2

Fig.2 is a cross-section side view of the electrical-energy-storage unit component. This figure indicates the alternating layers of nickel electrode layers **12** and high-permittivity composition-modified barium titanate dielectric layers **11**. This figure also indicate the preferentially aligning concept of the nickel electrode layers **12** so that each storage layer can be hooked up in parallel.

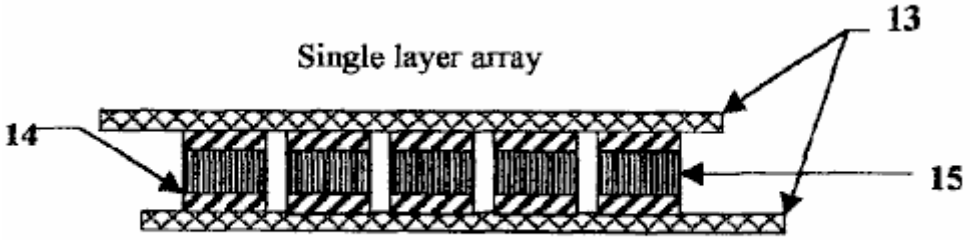


Figure 3

Fig.3 is side view of a single-layer array indicating the attachment of individual components **15** with the nickel side bars **14** attached to two preferentially aligned copper conducting sheets **13**.

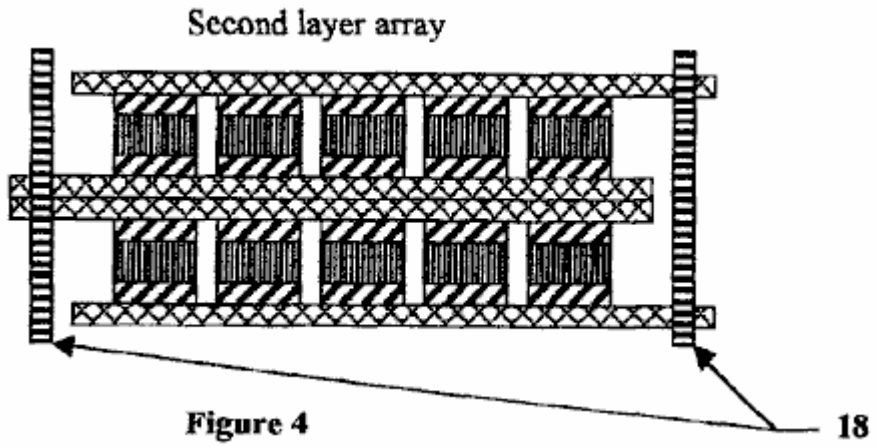


Figure 4

Fig.4 is a side view of a double-layer array with copper array connecting nickel bars **16** attaching the two arrays via the edges of the preferentially aligned copper conductor sheets **13**. This figure indicates the method of attaching the components in a multi-layer array to provide the required energy storage.

Reference No.	Refers to this in the drawings
8	System maximum voltage of 3500 V
9	2320 energy-storage components hooked up in parallel with a total capacitance of 31 Farad
10	System ground
11	High-permittivity calcined composition-modified barium titanate dielectric layers
12	Preferentially aligned nickel electrode layers
13	Copper conductor sheets
14	Nickel sidebars
15	Components
16	Copper array connecting nickel bars

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

Fig.1, **Fig.2**, **Fig.3**, and **Fig.4** of the drawings and the following description depict various preferred embodiments of the present invention for purposes of illustration only. One skilled in the art will readily recognise from the following discussion those alternative embodiments of the structures and methods illustrated herein may be employed without departing from the principles of the invention described here. While the invention will be described in conjunction with the preferred embodiments, it will be understood that they are not intended to limit the invention to those embodiments. On the contrary, the invention is intended to cover alternatives, modifications, and equivalents, which may be included within the spirit and scope of the invention as defined by the claims.

Preparation of the high-permittivity calcined composition-modified barium titanate powder that is used to fabricate the EESU is explained as follows. Wet-chemical-prepared powders of high-purity as well as composition-modified barium titanate with narrow particle-size distribution have been produced with clear advantages over those prepared by solid-state reaction of mechanically mixed, ball-milled, and calcined powdered ingredients. The

compositional and particle-size uniformity attained with a coprecipitated-prepared powder is vastly superior to that with a conventional-prepared powder. The microstructures of ceramics formed from these calcined wet-chemical-prepared powders are uniform in grain size and can also result in smaller grain size. Electrical properties are improved so that higher relative permittivities and increased dielectric breakdown strengths can be obtained. Further improvement can be obtained by the elimination of voids within the sintered ceramic body with subsequent hot isostatic pressing.

High-relative-permittivity dielectrics have inherent problems, namely ageing, fatigue, degradation, and decay of the electrical properties, which limit their application. The use of surface-coated powders in which the surface region is comprised of one or two materials different in composition from that of the powder overcomes these problems provided that the compositions are appropriately chosen.

Among ceramics, alumina [aluminium oxide (Al_2O_3)], and among glasses, calcium magnesium aluminosilicate ($CaO.MgO.Al_2O_3.SiO_2$) glasses are the best dielectrics in terms of having the highest dielectric breakdown strengths and to seal the high-relative-permittivity dielectric powder particles so as to eliminate or significantly reduce their inherent problems.

A glass with a given composition at temperatures below its glass transition temperature range, which is in the neighbourhood of its strain-point temperature, is in a fully rigid condition, but at temperatures above this range is in a viscous-flow condition, its viscosity decreasing with increasing temperature. The application of hot isostatic pressing to a sintered closed-pore porous ceramic body comprised of sufficient-thickness glass-coated powder will lead to void elimination provided the glass is in the viscous-flow condition where it is easily deformable and flowable.

The wet-chemical-prepared and calcined composition-modified barium titanate powder is accordingly coated with these layers of, first, alumina, and second, a calcium magnesium aluminosilicate glass. After the first layer has been applied by wet-chemical means, the powder is calcined at $1050^{O}C$ to convert the precursor, aluminium nitrate nonahydrate [$Al(NO_3)_3.9H_2O$] to aluminium oxide (corundum) [α-Al_2O_3]. Then the second layer is applied by wet-chemical means with the use of the precursors in the appropriate amounts of each, and in absolute ethanol (CH_3CH_2OH) as the solvent, shown in the accompanying table. After drying, the powder is calcined at $500^{O}C$ to convert the precursor mixture to a calcium magnesium aluminosilicate glass. It is important that the calcining temperature is not higher than the strain point of the selected glass composition to prevent sticking together of the powder. The glass coating has the further advantage of acting as a sintering aid and allowing a substantially lower firing temperature for densification of the ceramic body particularly during the hot-isostatic-pressing step.

Another significant advantage of the calcium magnesium aluminosilicate glass coating is that sintering and densification temperatures are sufficiently lowered to allow the use of nickel conductor electrodes in place of the conventional expensive platinum, palladium, or palladium-silver alloy ones.

Preparation of the Calcined Composition-Modified Barium Titanate Powder is Indicated by the Following Process Steps.

A solution of the precursors: $Ba(NO_3)_2$, $Ca(NO_3)_2.4H_2O$, $Nd(NO_3)_3.6H_2O$, $Y(NO_3)_3.4H_2O$, $Mn(CH_3COO)_2.4H_2O$, $ZrO(NO_3)_2$, and $[CH_3CH(O—)COONH_4]2Ti(OH)_2$, as selected from the reference; Sigma-Aldrich, Corp., "Handbook of Fine Chemicals and Laboratory Equipment", 2000-2001, in de-ionised water heated to $80^{O}C$ is made in the proportionate amount in weight percent for each of the seven precursors as shown in the most right-hand column of Table 3. A separate solution of $(CH_3)4NOH$ somewhat in excess amount than required, as shown in Table 4, is made in de-ionised water, free of dissolved carbon dioxide (CO_2) and heated to 80^{O}-$85^{O}C$. The two solutions are mixed by pumping the heated ingredient streams simultaneously through a coaxial fluid jet mixer. A slurry of the co-precipitated powder is produced and collected in a drown-out vessel. The co-precipitated powder is refluxed in the drown-out vessel at 90°-95° C. for 12 hr and then filtered, de-ionised-water washed, and dried. Alternatively, the powder may be collected by centrifugal sedimentation. An advantage of $(CH_3)4NOH$ as the strong base reactant is that there are no metal element ion residuals to wash away anyway. Any residual $(CH_3)4NOH$, like any residual anions from the precursors, is harmless, because removal by volatilisation and decomposition occurs during the calcining step. The powder contained in a silica glass tray or tube is calcined at $1050^{O}C$ in air. Alternatively, an alumina ceramic tray can be used as the container for the powder during calcining.

TABLE 2

Composition-modified barium titanate with metal element atom fractions given for an optimum result, as demonstrated in the reference: P. Hansen, U.S. Pat. No. 6,078,494, issued Jan. 20, 2000.

Composition-modified barium titanate with
metal element atom fractions as follows:

Metal Element	Atom Fraction	Atomic Weight	Product	Weight %
Ba	0.9575	137.327	131.49060	98.52855
Ca	0.0400	40.078	1.60312	1.20125
Nd	0.0025	144.240	0.36060	0.27020
Total:	**1.0000**			**100.00000**
Ti	0.8150	47.867	39.01161	69.92390
Zr	0.1800	91.224	16.42032	29.43157
Mn	0.0025	54.93085	0.13733	0.24614
Y	0.0025	88.90585	0.22226	0.39839
Total:	**1.0000**			**100.00000**

TABLE 3
Water-soluble precursors and reactant strong base for wet-chemical-prepared powder of a composition-modified barium titanate by a coprecipitation procedure

Precursor	Formula	FW	Mol fraction	Product	Weight %	Multiplier factor	Product	Weight %
Barium nitrate	$Ba(NO_3)_2$	261.34	0.9575	250.233050	95.95748	1.0	95.95748	48.09898
Calcium nitrate tetrahydrate	$Ca(NO_3)_2.4H_2O$	236.15	0.0400	9.446000	3.62228	1.0	3.62228	1.81568
Neodymium nitrate hexahydrate	$Nd(NO_3)_3.6H_2O$	438.35	0.0025	1.095875	0.42024	1.0	0.42024	0.21065
Yttrium nitrate tetrahydrate	$Y(NO_3)_3.4H_2O$	346.98	0.0025	0.86745	0.30676	0.995	0.30523	0.15300
Manganese(II) acetate tetrahydrate	$Mn(CH_3COO)_2.4H_2O$	245.08	0.0025	0.61270	0.21667	0.995	0.21559	0.10806
Oxozirconium(IV) nitrate	$ZrO(NO_3)_2$	231.23	0.1800	41.62140	14.71882	0.995	14.64523	7.34097
Bis(ammonium lactato) dihydroxotitanium(IV)	$[CH_3CH(O—)COONH_4]2Ti(OH)_2$	294.08	0.8150	239.67520	84.75775	0.995	84.33396	42.27266
							Total:	100.0000
Reactant strong base Tetramethylammonium hydroxide	$(CH_3)_4NOH$	91.15						

TABLE 4

Calculation of minimum amount of $(CH_3)_4NOH$
required for 100 g of the precursor mixture

Precursor	FW	Wt %	Wt %/FW	Reactant base multiplier	Mol of base required
$Ba(NO_3)_2$	261.34	48.09898	0.184048	2	0.368095
$Ca(NO_3)_2.4H_2O$	236.15	1.81568	0.007689	2	0.015377
$Nd(NO_3)_3.6H_2O$	438.35	0.21065	0.000481	3	0.001442
$Y(NO_3)_3.4H_2O$	346.98	0.15300	0.000441	3	0.001323
$Mn(CH_3COO)_2.4H_2O$	245.08	0.10806	0.000441	2	0.000882
$ZrO(NO_3)_2$	231.23	7.34097	0.031747	2	0.063495
$[CH_3CH(O—)COONH_4]_2Ti(OH)_2$	294.08	42.27266	0.143745	2	0.287491
	Total:	100.00000			0.738105
Reactant strong base					
$(CH_3)_4NOH$	91.15				

Note: The weight of (CH3)4NOH required is accordingly a minimum of
(0.738105 mol) (91.15 g/mol) = 67.278 g for 100 g of the precursor mixture.
Tetramethylammonium hydroxide (CH3)4NOH is a strong base.

Coating of Aluminium Oxide on Calcined Modified Barium Titanate Powder

Barium titanate $BaTiO_3$	FW 233.19	d 6.080 g/cm3
Aluminium oxide Al_2O_3	FW 101.96	d 3.980 g/cm3

Precursor, aluminium nitrate nonahydrate, as selected from the reference: Sigma-Aldrich Corp., "Handbook of Fine Chemicals and Laboratory Equipment", 2000-2001. $Al(NO_3)_3.9H_2O$ FW 3.75.13

For Calcined Aluminium Oxide (Al_2O_3) Coating of 100 Angstrom units Thickness on Calcined Modified Barium Titanate Powder 100 Angstrom units = 10-6 cm 1.0 m^2 = 104 cm^2

area thickness of Al_2O_3 coating volume $(10^4\ cm^2/g)(10^{-6}\ cm) = 10^{-2}\ cm^3/g$ - - - of calcined powder

$$\frac{(10^{-2}\ \text{cm}^3\ \text{volume}\ Al_2O_3\ \text{coating}) \times (3.98\ \text{g/cm}^3\ \text{density of}\ Al_2O_3)}{\text{g of calcined powder}} = \frac{39.8 \times 10^{-3}\ \text{g of}\ Al_2O_3\ \text{coating}}{\text{g of calcined powder}}\ \text{or}$$

$$= \frac{39.8\ \text{mg of}\ Al_2O_3\ \text{coating}}{\text{g of calcined powder}}$$

$Al(NO_3)_3.9H_2O$ (FW 375.13)(2)=750.26

Al_2O_3 FW 101.96=101.96

750.26/101.96=7.358

$$\frac{(7.358)(39.8\ \text{mg of}\ Al_2O_3\ \text{coating})}{\text{g of calcined powder}} = \frac{292.848\ \text{mg of}\ Al(NO_3)_3 \cdot 9H_2O}{\text{g of calcined powder}}$$

For an aluminium oxide (Al_2O_3) coating of 100 Angstrom units thickness on calcined modified barium titanate powder with particle volume of 1.0 μm^3, 39.8 mg of Al_2O_3 are required per g of this powder, corresponding to 292.848 mg of the aluminium nitrate nonahydrate [$Al(NO_3)_3.9H_2O$] precursor required per g of this powder.

Coating of Calcium Magnesium Aluminosilicate Glass on Aluminium Oxide Coated
Calcined Modified Barium Titanate Powder

	FW g/mol	d g/cm^3
Barium titanate $BaTiO_3$	233.19	6.080

Calcium magnesium aluminosilicate (CaO.MgO.Al2O3.SiO2) glass precursors, as selected from the reference: Sigma-Aldrich, Corp., "Handbook of Fine Chemicals and Laboratory Equipment", 2000-2001.

Calcium methoxide	$(CH_3O)2Ca$	101.15
Calcium isopropoxide	$[(CH_3)_2CHO]_2Ca$	158.25
Magnesium methoxide	$(CH_3O)_2Mg$	86.37
Magnesium ethoxide	$(CH_3CH_2O)_2Mg$	114.43
Aluminium ethoxide	$(CH_3CH_2O)_3Al$	162.16
Aluminium isopropoxide	$[(CH_3)_2CHO]_3Al$	204.25
Aluminium butoxide	$[CH_3(CH_2)_3O]_3Al$	246.33
Tetraethyl orthosilicate	$Si(OCH_2CH_3)_4$	208.33

Select glass composition, e.g.,

$CaO.MgO.2Al_2O_3.8SiO_2$ and accordingly the precursors:

1 mol	(158.25 g) calcium isopropoxide
1 mol	(114.43 g) magnesium ethoxide
4 mol	(817.00 g) aluminum isopropoxide
8 mol	(1666.64 g) tetraethyl orthosilicate
	2756.32 g for 1.0 mol glass

Prepare Mixture of these Precursors in Absolute Ethanol (to Avoid Hydrolysis) and in Dry-Air Environment (Dry Box) (also to Avoid Hydrolysis).

Glass Composition: $CaO.MgO.2Al_2O_3.8SiO_2$ or $CaMgAl_4Si_8O_{24}$

1 mol (56.08 g)	CaO
1 mol (40.30 g)	MgO
2 mol (101.96 g × 2 = 203.92 g)	Al_2O_3
8 mol (60.08 g × 8 = 480.64 g)	SiO_2

glass FW total 780.98 g/mol
Density of glass: about 2.50 g/cm^3

Calcined modified barium titanate powder
Particle volume: 1.0 μm^3 or $1.0(10^{-4}\ cm)^3 = 10^{-12}\ cm^3$;
so there are 10^{12} particles/cm^3 (assumption of no voids)
Particle area: 6 μm^2 or $(6)(10^{-4}\ cm)^2 = 6\times10^{-8}\ cm^3$;
Particle area/cm3 (no voids):
$(6\times10^{-8}\ cm^2/particle)(10^{12}\ particles/cm^3) = 6\times10^4\ cm^2/cm^3$ or 6 m^2/cm^3.

Then for density of 6 g/cm^3, the result is:

$$\frac{6\ m^2/cm^3}{6\ g/cm^3} = 1.0\ m^2/g$$

For Calcined Glass Coating of 100 Angstrom units Thickness on Calcined Powder:

100 Angstrom units = 10^{-6} cm 1.0 m^2 = 10^4 cm^2

(10^4 cm^2/g)(10^{-6} cm) = 10^{-2} cm^3/g of calcined powder of glass coating and then

$$\frac{(10^{-2}\ \text{cm}^3\ \text{of glass coating})}{\text{g of calcined powder}} \times (2.50\ \text{g/cm}^3\ \text{density of glass}) =$$

$$\frac{25.0\times10^{-3}\ \text{g of glass coating}}{\text{g of calcined powder}} \quad \text{or} \quad \frac{25.0\ \text{mg of glass coating}}{\text{g of calcined powder}}$$

Precursor mixture FW 2756.32 = 3.529
Glass FW 780.98

$$\frac{(3.529)(25.0\ \text{mg of glass coating})}{(\text{g of calcined powder})} = 88.228\ \text{mg of precursor mixture}$$

For a $CaMgAl_4Si_8O_{24}$ glass coating of 100 Angstrom units thickness on calcined modified barium titanate powder with particle volume of 1.0 μm3, 25.0 mg of this glass are required per g of this powder, corresponding to 88.228 mg of the precursor mixture required per g of this powder.

Particle Volume and Area

V particle = a^3 for cube
If a = 1.0 μm, V = 1.0 μm^3
A particle = $6a^2$ for cube
If a = 1.0 μm, A = 6 μm^2

Particle coating volume

($6\ a^2$)(t), if t = 100 Angstrom units = 10×10^3 μm, and 6 a^2=6.0 μm^2,
then (6.082 m^2)(10×10^{-3} μm) = 60×10^{-3} μm^3 = V coating

Ratio of particle coating volume to particle volume 60×10^{-3} μm^3/1.0 μm^3 = 60×10^{-3} = 0.06 or 6%

With the assumption of no voids and absolutely smooth surface, for an ideal cubic particle with volume of 1.0 μm3 and for a particle coating of 100 Angstrom units thickness, the coating volume is 60×10^{-3} μm3 or 6.0% that of the particle volume.

Calculations of the Electrical-Energy-Storage Unit's Weight, Stored Energy, Volume, and Configuration.

Assumptions:

The relative permittivity of the high-permittivity powder is nominally 33,500, as given in the reference: P. Hansen, U.S. Pat. No. 6,078,494, issued Jan. 20, 2000.

- The 100 ? coating of Al2O3 and 100 ? of calcium magnesium aluminosilicate glass will reduce the relative permittivity by 12%.
- K = 29,480

Energy stored by a capacitor: E = CV^2/(2×3600 s/h) = W·h

- C = capacitance in farads
- V = voltage across the terminals of the capacitor

It is estimated that is takes 14 hp, 746 watts per hp, to power an electric vehicle running at 60 mph with the lights, radio, and air conditioning on. The energy-storage unit must supply 52,220 W·h or 10,444 W for 5 hours to sustain this speed and energy usage and during this period the EV will have travelled 300 miles. Each energy-storage component has 1000 layers.

C = ε_oKA/t

* ε_o = permittivity of free space
* K = relative permittivity of the material
* A = area of the energy-storage component layers
* t = thickness of the energy-storage component layers

Voltage breakdown of the energy-storage components material after coating with Al_2O_3 and calcium magnesium aluminosilicate glass will be in the range of 1.0×10^6 V/cm to 5×10^6 V/cm or higher. Using the proper voltage breakdown selected from this range could allow the voltage of the energy-storage unit to be 3500 V or higher.
One hp = 746 W

EXAMPLE

Capacitance of one layer = 8.854×10^{-12} F / m × 2.948×10^4 × 6.45×10^{-4} m^2 / 12.7×10^{-6} m

C = 0.000013235 F

With 1000 layers:

C = 0.013235 F

The required energy storage is

E_t = 14 hp × 746 W /hp × 5 h = 52,220 W·h

The total required capacitance of the energy-storage unit:

CT = E_t × 2 × 3600 s/h / V^2 = 52,220 W·h × 2 × 3600 s/h/(3500 V)^2 C_T = 31 F

Number of capacitance components required:

N_c = 31 F / 0.013235 F = 2320

Volume and weight of energy-storage unit:

Volume of the dielectric material:

Volume = area x thickness x number of layers
= 6.45 cm^2 x 12.72 x 10^{-4} cm x 1000
= 8.2 cm^3

Total volume = 8.2 cm^3 × number of components (2320) = 19,024 cm^3
Density of the dielectric material = 6.5 g/cm^3
Weight of each component = density × volume = 53.3 g
Total weight of the dielectric material = 53.3 g × 2320 / 454 g per pound = 272 pounds

Volume of the nickel conductor layers:
Thickness of the nickel layer is 1×10-6 m
Volume of each layer = 6.45 cm2×1.0×10-4 cm × 1000 = 0.645 cm3
Density of nickel = 8.902 g/cm3
Weight of nickel layers for each component = 5.742 g
Total weight of nickel = 34 pounds

Total number of capacitance layers and volume of the EESU:
Area required for each component to solder bump = 1.1 $inch^2$
A 12 × 12 array will allow 144 components for each layer of the first array
19 layers of the second array will provide 2736 components which are more than enough to meet the required 2320 components. The distance between the components will be adjusted so that 2320 components will be in each EESU. The second array area will remain the same.

The total weight of the EESU (est.) = 336 pounds
The total volume of the EESU (est.) = 13.5 inches × 13.5 inches × 11 inches = 2005 $inches^3$ which includes the weight of the container and connecting material.
The total stored energy of the EESU = 52,220 W·h

From the above description, it will be apparent that the invention disclosed herein provides a novel and advantageous electrical-energy-storage unit composed of unique materials and processes. The foregoing discussion discloses and describes merely exemplary methods and embodiments of the present invention. As will be understood by those familiar with the art, the invention may be embodied in other specific forms and utilise other materials without departing from the spirit or essential characteristics thereof. Accordingly, the disclosure of the present invention is intended to be illustrative, but not limiting, of the scope of the invention, which is set forth in the following claims.

CLAIMS

1. A method for making an electrical-energy-storage unit comprising components fabricated by the method steps as follow;

a) preparing a wet-chemical-prepared calcined composition-modified barium titanate powder derived from a solution of precursors: $Ba(NO_3)_2$, $Ca(NO_3)_2.4H_2O$, $Nd(NO_3)_3.6H_2O$, $Y(NO_3)_3.4H_2O$, $Mn(CH_3COO)_2.4H_2O$, $ZrO(N_3O)_2$, and $[CH_3CH(O—)COONH_4]_2Ti(OH)_2$ in de-ionised water heated to 80^OC, and a separate solution of $(CH_3)_4NOH$ made in de-ionised water and heated to 80^O-85^OC, then mixing the solutions by pumping the heated ingredient streams simultaneously through a coaxial fluid mixer producing co-precipitated powder, then collecting the co-precipitated powder in a drown-out vessel and refluxing at a temperature of 90^O-95^OC for 12 hours, then filtering, washing with de-ionised water, drying, and then calcining 1050^OC in air;

b) fabricating an aluminium oxide (Al_2O_3) coating of 100 Angstrom units thickness on to the wet-chemical-prepared calcined composition-modified barium titanate powder, with the use of aluminium nitrate nonahydrate precursor applied by wet chemical means, then calcining at 1050^OC, resulting in a single-coated calcined composition-modified barium titanate powder;

c) fabricating on to the alumina-coated composition-modified barium titanate powder, a second uniform coating of 100 Angstrom units of calcium magnesium aluminosilicate glass derived from alcohol-soluble precursors: calcium methoxide or calcium isopropoxide, magnesium methoxide or magnesium ethoxide, aluminium ethoxide or aluminium isopropoxide or aluminium isopropoxide, and tetraethyl orthosilicate are applied by wet chemical means which upon calcining at 500^OC results in a double-coated composition-modified barium titanate powder;

d) blending, this double-coated composition-modified barium titanate powder with a screen-printing ink containing appropriate plastic resins surfactants, lubricants, and solvents to provide a suitable rheology for screen printing;

e) screen-printing into interleaved multi-layers of alternating offset nickel electrode layers **12** and double-coated calcined composition-modified barium titanate high-relative-permittivity layers **11** with the use of screening inks having the proper rheology for each of the layers;

f) drying and cutting the screen-punted multi-layer components **15** into a specified rectangular area;

g) sintering the screen-printed multi-layer components **15**, first at a temperature of 350^OC for a specified length of time, then at 850^OC for a specified length of time, to form closed-pore porous ceramic bodies; and

h) hot isostatically pressing the closed-pore porous ceramic bodies, at a temperature of 700^OC with a specified pressure, into a void-free condition;

i) grinding and each side of the component to expose the alternating offset interleaved nickel electrodes **12**;

j) connecting nickel side bars **14** to each side of the components **15**, that have the interleaved and alternating offset nickel electrodes **12** exposed, by applying nickel ink with the proper rheology to each side and clamping the combinations together;

k) heating the components and side nickel bar combination **14-15** 800^OC, and time duration of 20 minutes to bond them together;

l) wave soldering each side of the conducting bars;

m) assembling the components **15** with the connected nickel side bars **14** into the first array, utilising unique tooling and solder-bump technology;

n) assembling the first arrays into the second array;

o) assembling the second arrays into the EESU final assembly.

2. The method of claim 1 wherein a second coating of glass is provided on to the double-coated composition-modified barium titanate powder being in contact with the nickel electrodes and having an applied working voltage of 3500 V across the parallel electrodes.

3. The method of claim 1 wherein a dielectric voltage breakdown strength of 5.0×10^6 V/cm was achieved across the electrodes of the components.

4. The method of claim 1 wherein the method provides an ease of manufacturing due to the softening temperature of the calcium magnesium aluminosilicate glass allowing the relatively low hot-isostatic-pressing temperatures of 700^{O}C which in turn provides a void-free ceramic body.

5. The method of claim 1 wherein the method provides an ease of fabrication due to the softening temperature of the calcium magnesium aluminosilicate glass allowing the relatively low hot-isostatic-pressing temperatures of 700^{O}C which in turn allows the use of nickel for the conduction-path electrodes rather than expensive platinum, palladium, or palladium-silver alloy.

6. The method of claim 1 wherein the method provides an ease of fabrication due to the softening temperature of the calcium magnesium aluminosilicate lass allowing the relatively low hot-isostatic-pressing temperatures of 700^{O}C, which feature along with the coating method provided a uniform-thickness shell of the calcium magnesium aluminosilicate glass and in turn provides hot-isostatic-pressed double-coated composition-modified barium titanate high-relative-permittivity layers that are uniform and homogeneous in microstructure.

7. The method of claim 1 wherein the method provides the double coating of the basis particles of the composition-modified barium titanate powder thereby reducing the leakage and ageing of this material by an order of magnitude of the specification of this basis material, thus reducing the discharge rate to 0.1% per 30 days.

8. The method of claim 1 wherein the method provides a double coating of the composition-modified barium titanate powder, the hot-isostatic-pressing process, the high-density solder-bump packaging, and along with the double-layered array configuration stored 52,220 W·h of electrical energy in a 2005 $inches^3$ container.

9. The method of claim 1 wherein the method provides materials used: water-soluble precursors of barium (Ba), calcium (Ca), titanium (Ti), zirconium (Zr), manganese (Mn), yttrium (Y), neodymium (Nd), forming the composition-modified barium titanate powder, and the metals: nickel (Ni), and copper (Cu), which are not explosive, corrosive, or hazardous.

10. The method of claim 1 wherein the method provides an EESU that is not explosive, corrosive, or hazardous and therefore is a safe product when used in electrical vehicles, which include bicycles, tractors, buses, cars, or any device used for transportation or to perform work.

11. The method of claim 1 wherein the method provides an EESU which can store electrical energy generated from solar voltaic cells or other alternative sources for residential, commercial, or industrial applications.

12. The method of claim 1 wherein the method provides an EESU which can store electrical energy from the present utility grid during the night when the demand for electrical power is low and then deliver the electrical energy during the peak power demand times and thus provide an effective power averaging function.

13. The method of claim 1 wherein the method provides a double coating of the composition-modified barium titanate powder and a hot-isostatic-pressing process which together assists in allowing an applied voltage of 3500 V to a dielectric thickness of 12.76×10-6 m to be achieved.

14. The method of claim 1 wherein the method provides a EESU which when fully discharged and recharged, the EESU's initial specifications are not degraded.

15. The method of claim 1 wherein the method provides a EESU which can be safely charged to 3500 V and store at least 52.22 kW·h of electrical energy.

16. The method of claim 1 wherein the method provides a EESU at has a total capacitance of at least 31 F.

17. The method of claim 1 wherein the method provides a EESU that can be rapidly charged without damaging the material or reducing its life.

HERMANN PLAUSON

US Patent 1,540,998 9th June 1925 Inventor: Hermann Plauson

CONVERSION OF ATMOSPHERIC ELECTRIC ENERGY

Please note that this is a re-worded excerpt from this patent. It describes in considerable detail, different methods for abstracting useable electrical power from passive aerial systems. He describes a system with 100 kilowatt output as a “small” system.

Be it known that I, Hermann Plauson, Estonian subject, residing in Hamburg, Germany, have invented certain new and useful improvements in the Conversion of atmospheric Electric Energy, of which the following is a specification.

According to this invention, charges of atmospheric electricity are not directly converted into mechanical energy, and this forms the main difference from previous inventions, but the static electricity which runs to earth through aerial conductors in the form of direct current of very high voltage and low current strength is converted into electro-dynamic energy in the form of high frequency vibrations. Many advantages are thereby obtained and all disadvantages avoided.

The very high voltage of static electricity of a low current strength can be converted by this invention to voltages more suitable for technical purposes and of greater current strength. By the use of closed oscillatory circuits it is possible to obtain electromagnetic waves of various amplitudes and thereby to increase the degree of resonance of such current. Such resonance allows various values of inductance to be chosen which, by tuning the resonance between a motor and the transformer circuit, allows the control of machines driven by this system. Further, such currents have the property of being directly available for various uses, other than driving motors, including lighting, heating and use in electro-chemistry.

Further, with such currents, a series of apparatus may be fed without a direct current supply through conductors and the electro-magnetic high frequency currents may be converted by means of special motors, adapted for electro-magnetic oscillations, into alternating current of low frequency or even into high voltage direct current.

DESCRIPTION OF THE DRAWINGS

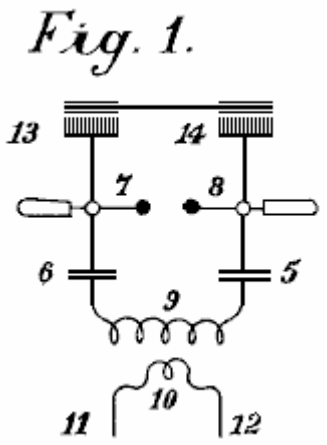


Fig.1 is an explanatory figure

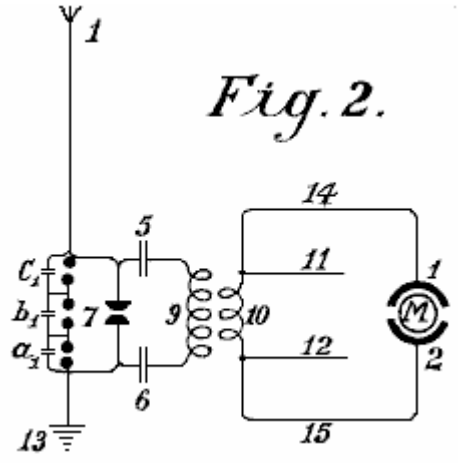


Fig.2 is a diagrammatic view of the most simple form.

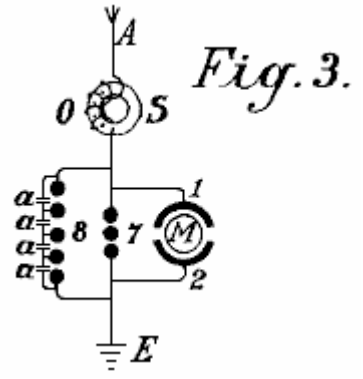


Fig.3 shows a method of converting atmospheric electrical energy into a form suitable for use with motors.

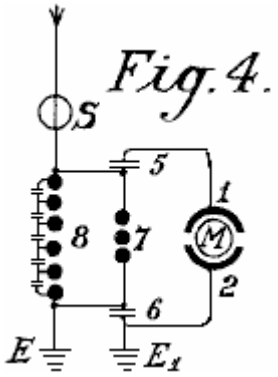


Fig.4 is a diagram showing the protective circuitry.

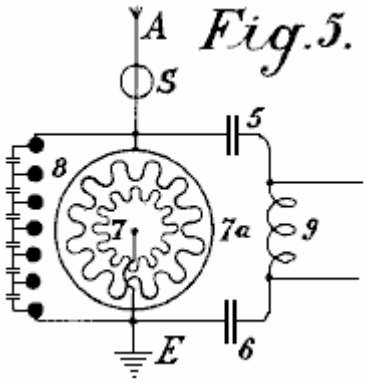


Fig.5.

Fig.5 is a diagram of an arrangement for providing control

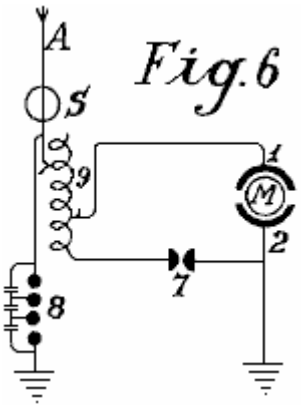


Fig.6

Fig.6 is an arrangement including a method of control

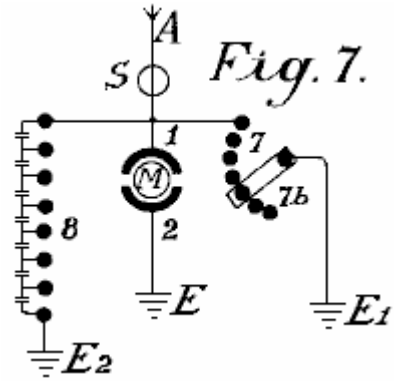


Fig. 7.

Fig.7 shows how the spark gap can be adjusted

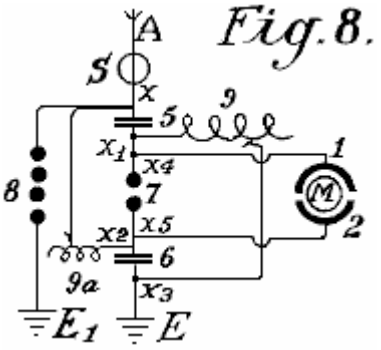


Fig.8 shows a unipolar connection for the motor

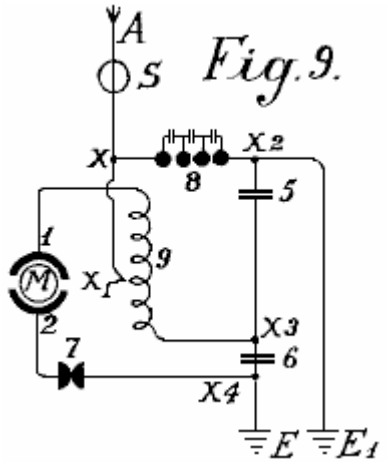


Fig.9 shows a weak coupled system suitable for use with small power motors

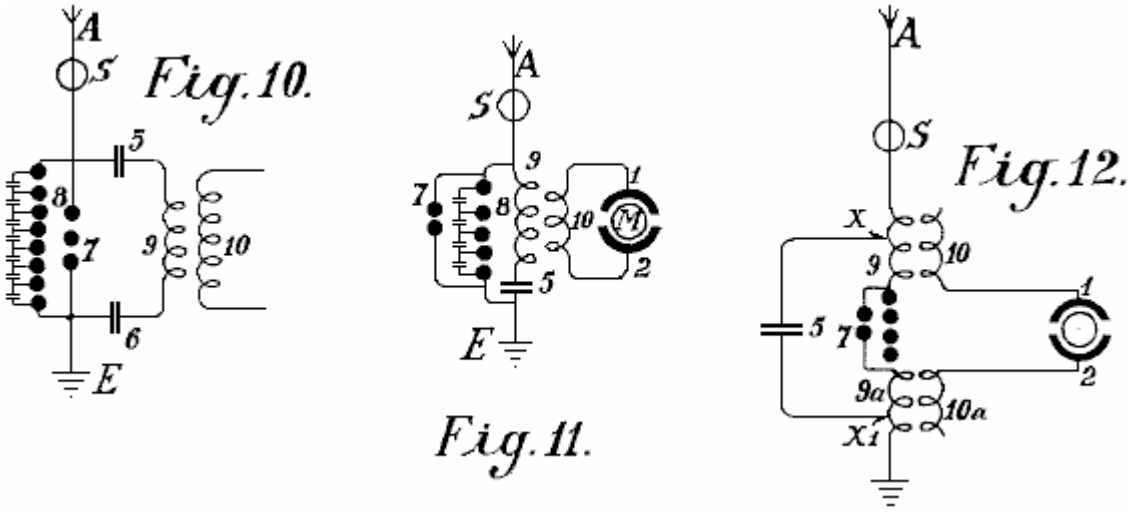


Fig.10, Fig.11 and Fig.12 show modified arrangements

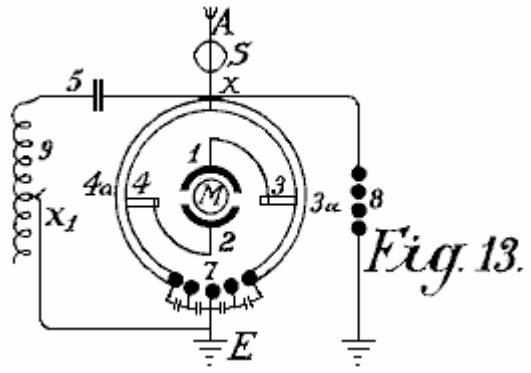


Fig.13 shows a form of inductive coupling for the motor circuit

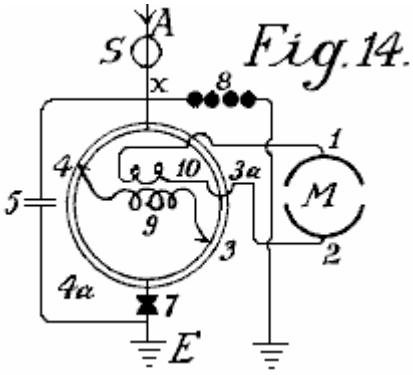


Fig.14 is a modified form of Fig.13 with inductive coupling.

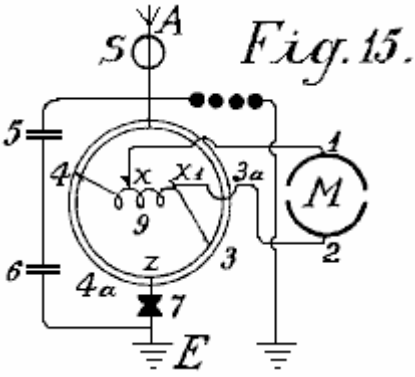


Fig.15 is an arrangement with non-inductive motor

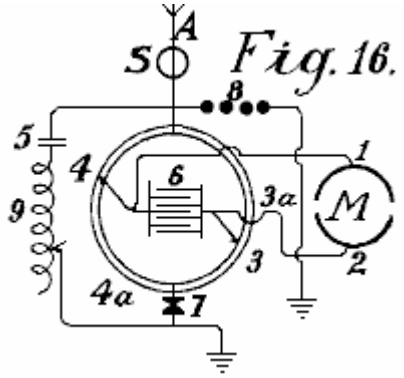


Fig.16 is an arrangement with coupling by capacitor.

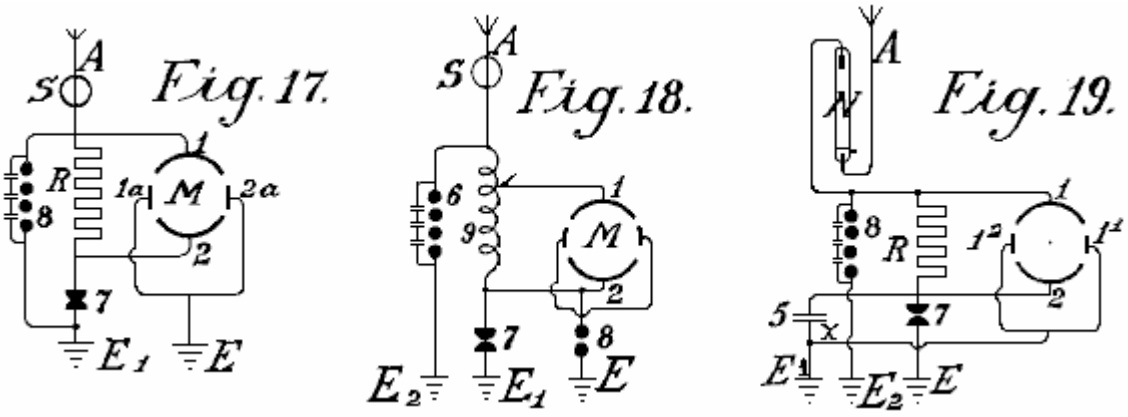


Fig.17, **Fig.18** and **Fig.19** are diagrams showing further modifications

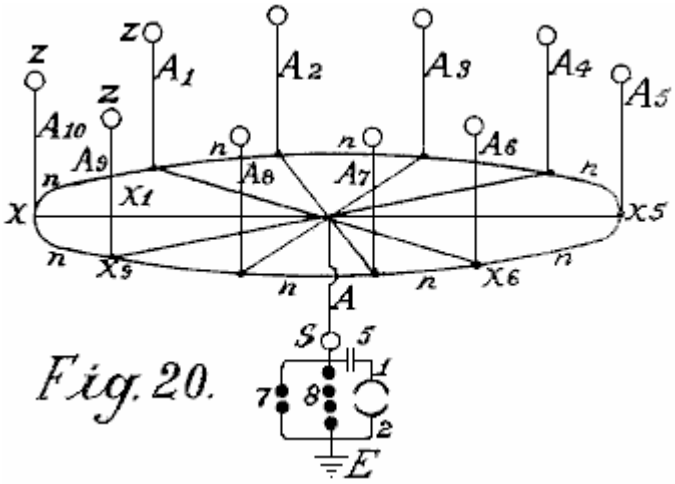


Fig.20 shows a simple form in which the aerial network is combined with special collectors

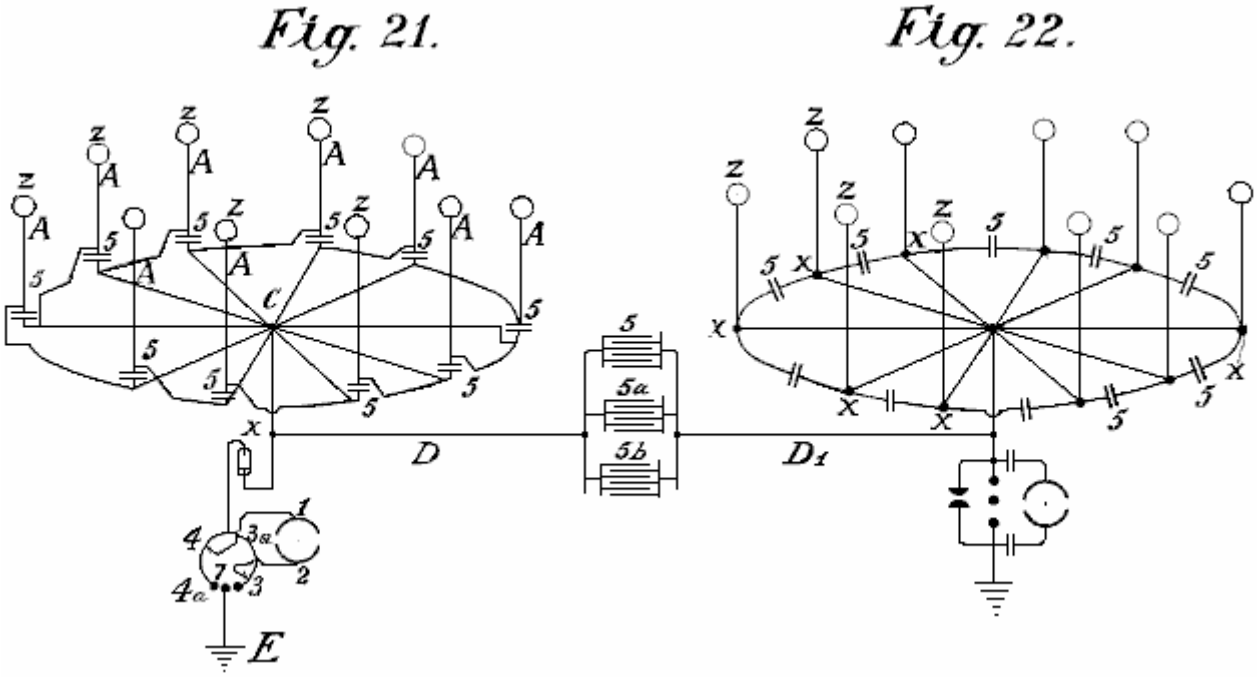


Fig.21 shows diagramatically, an arrangement suitable for collecting large quantities of energy.
Fig.22 is a modified arrangement having two rings of collectors

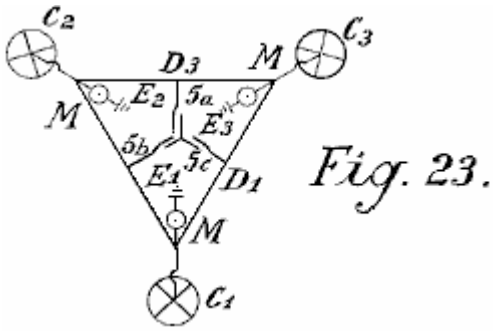


Fig. 23.

Fig.23 shows the connections for three rings of collectors

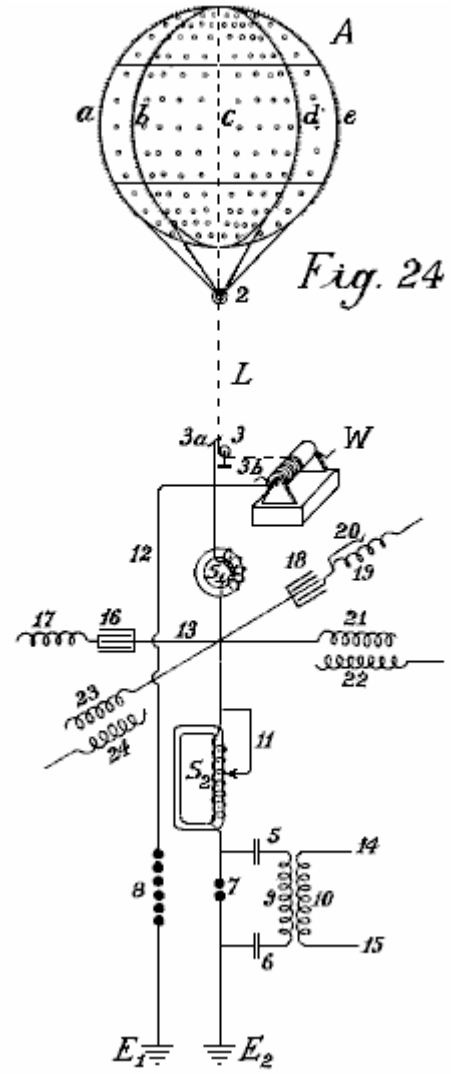


Fig. 24

Fig.24 shows a collecting balloon and diagram of its battery of capacitors

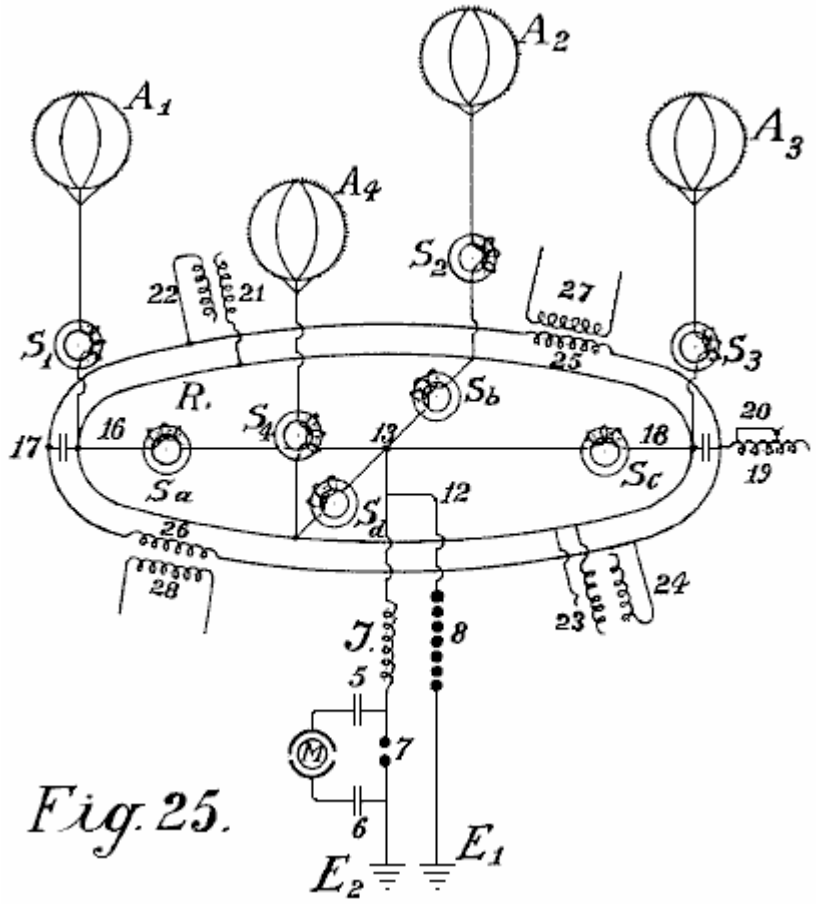


Fig. 25.

Fig.25 and Fig.26 show modified collector balloon arrangements.

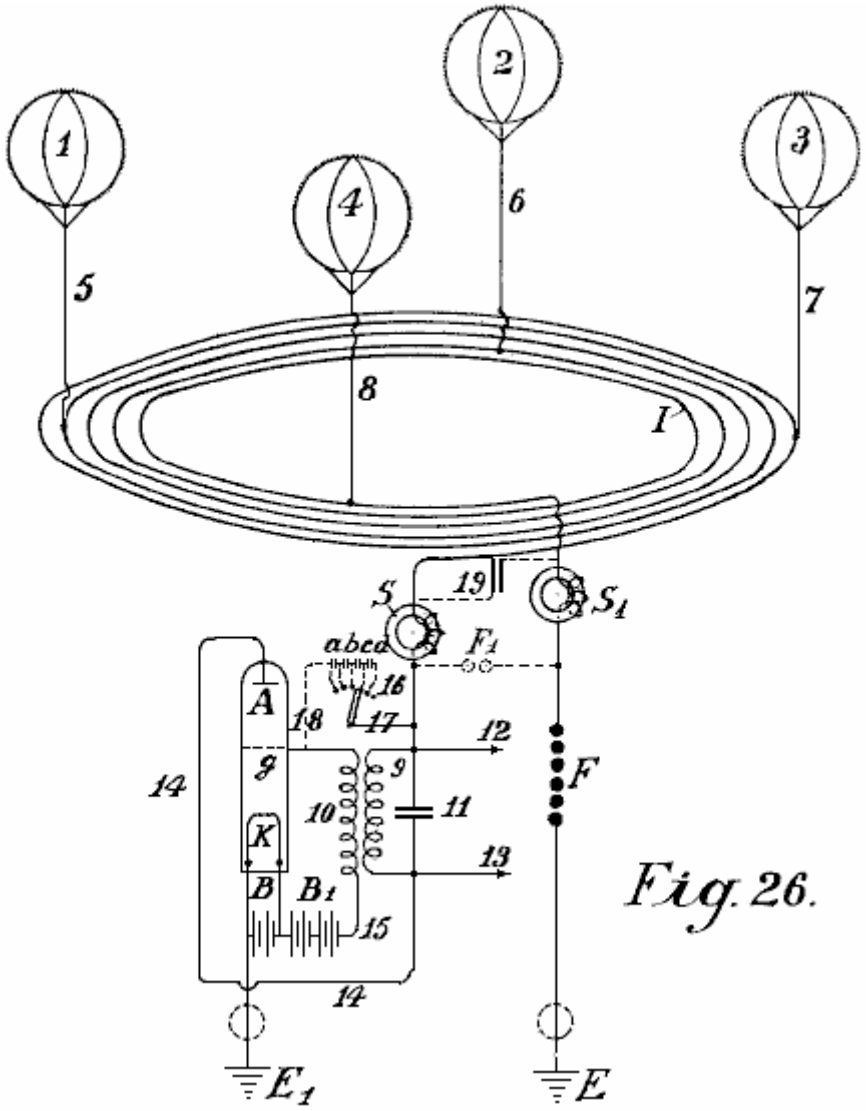

1
2
3
4
5
6
7
8
I
S
19
S1
abcd
F1
16
A
18
17
12
9
9
F
14
10
11
K
13
B
B1
15
14
E1
E

Fig. 26.

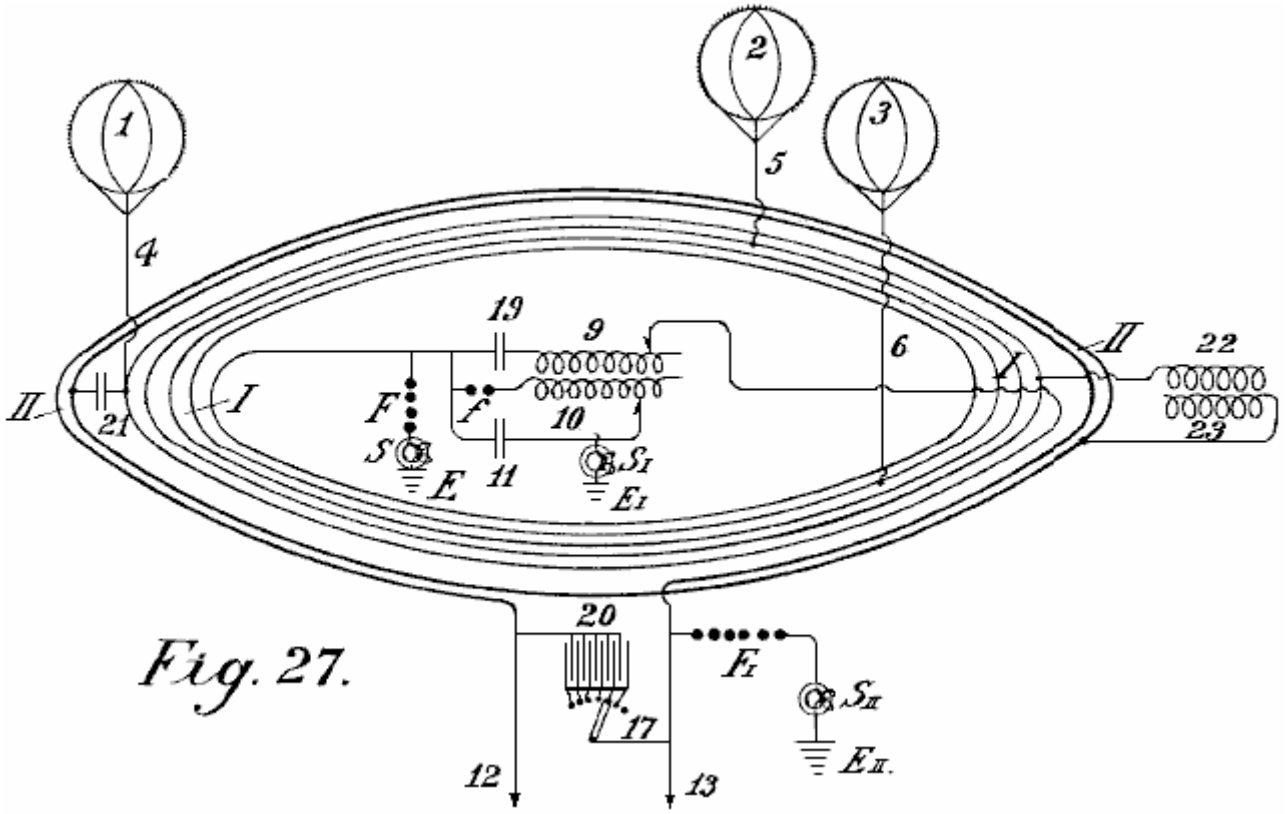


Fig. 27.

Fig.27 shows a second method of connecting conductors for the balloon aerials.

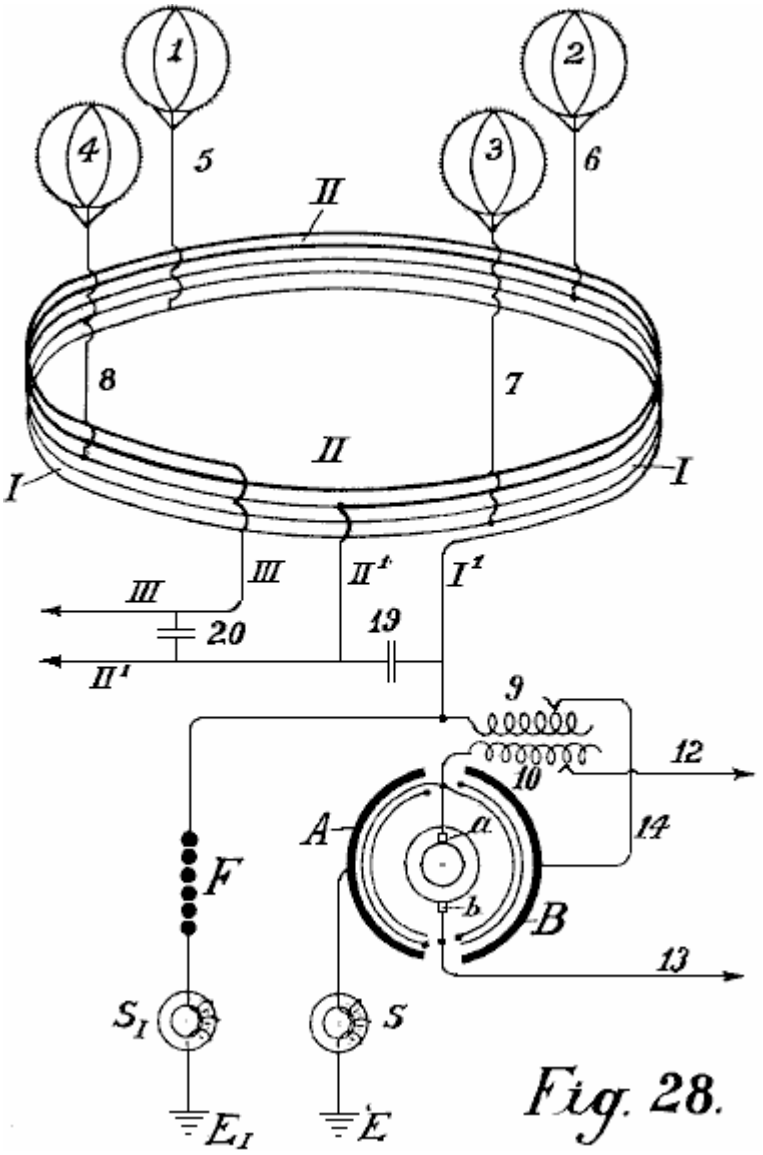


Fig. 28.

Fig.28 shows an auto-transformer method of connection.

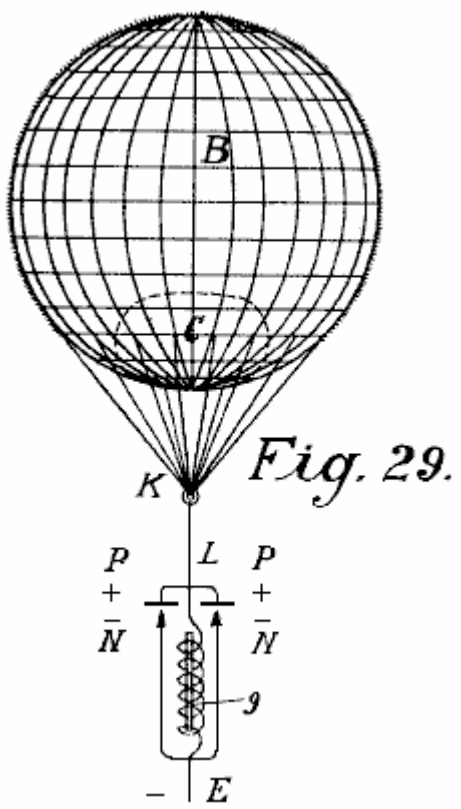


Fig. 29.

Fig.29 shows the simplest form of construction with incandescent cathode.

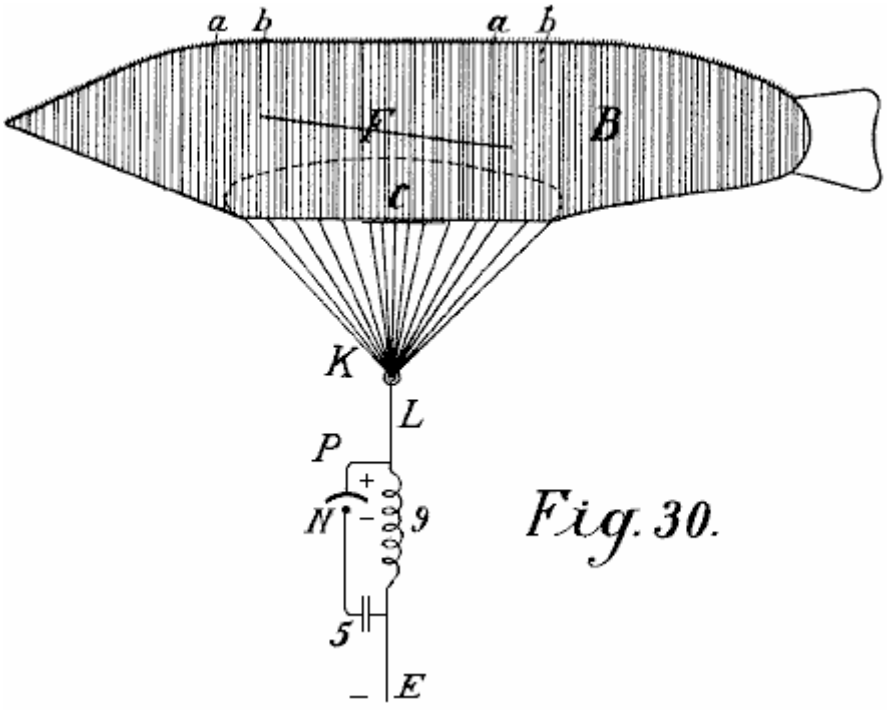


Fig. 30.

Fig.30 shows a form with a cigar-shaped balloon.

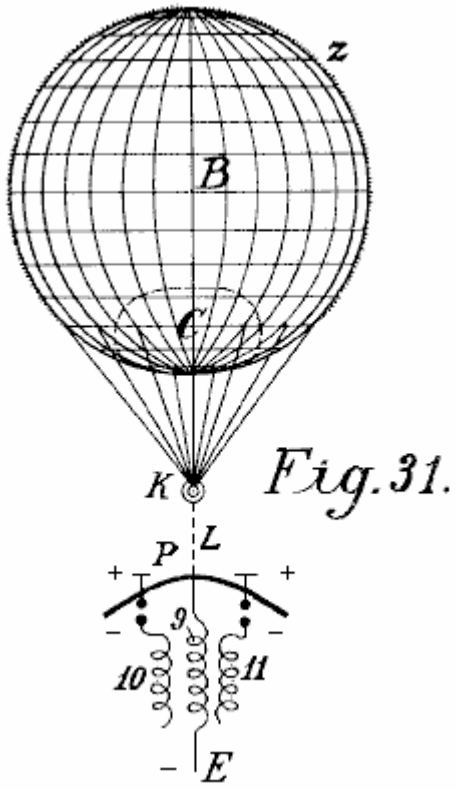


Fig. 31.

Fig.31 is a modified arrangement.

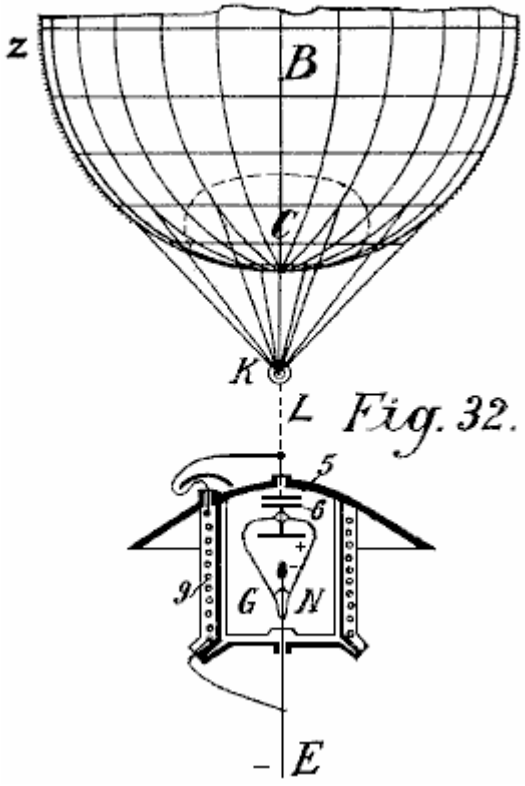


Fig. 32.

Fig.32 shows a form with cathode and electrode enclosed in a vacuum chamber.

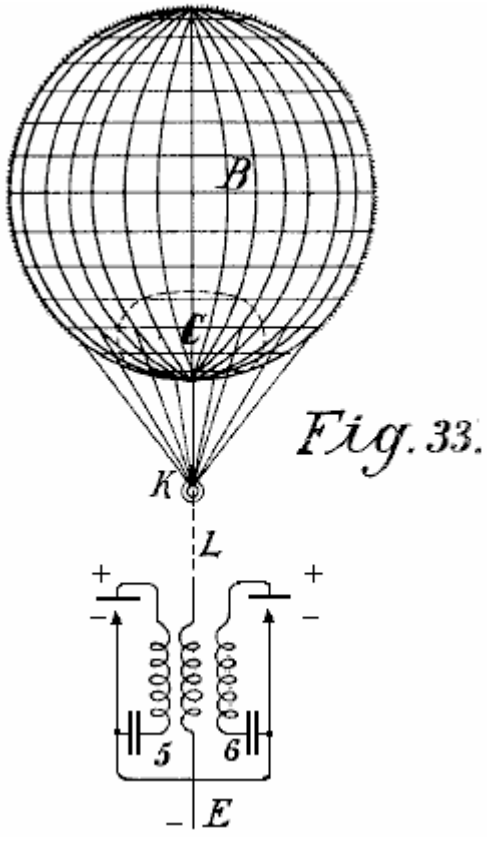


Fig. 33.

Fig.33 is a modified form of Fig.32

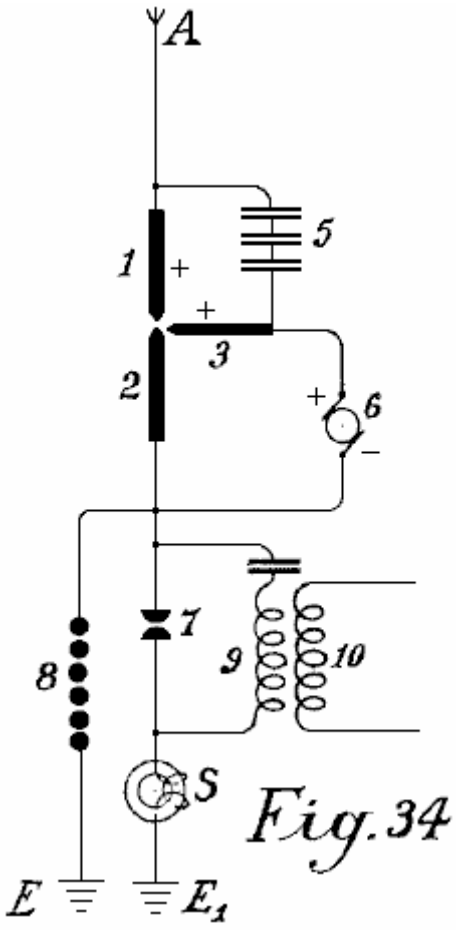


Fig. 34

Fig.34 shows an arc light collector.

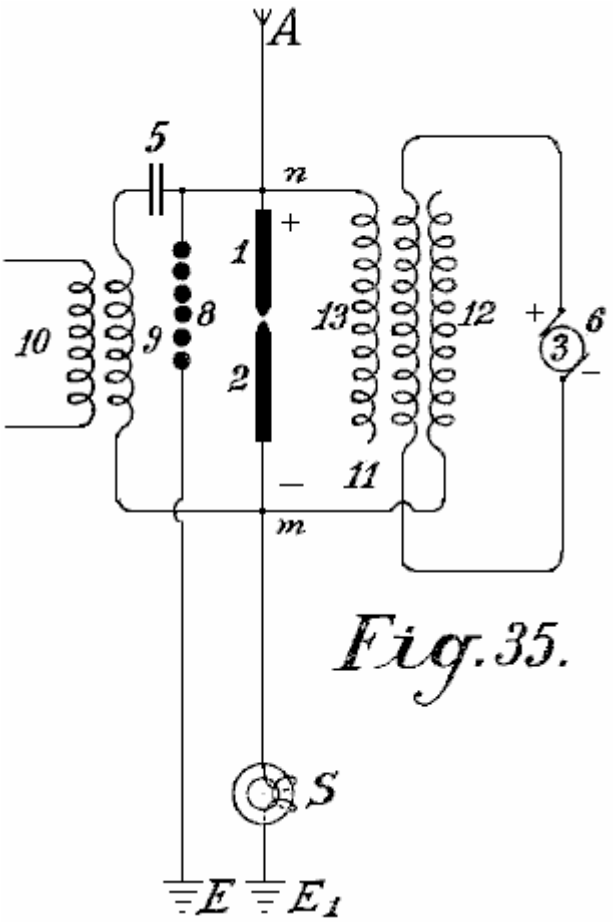


Fig. 35.

Fig.35 shows such an arrangement for alternating current

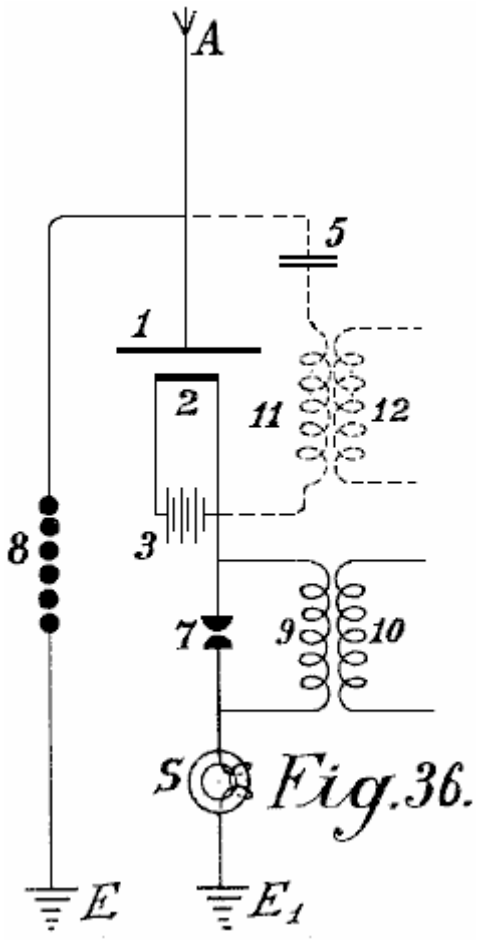


Fig.36.

Fig.36 shows an incandescent collector with Nernst lamp

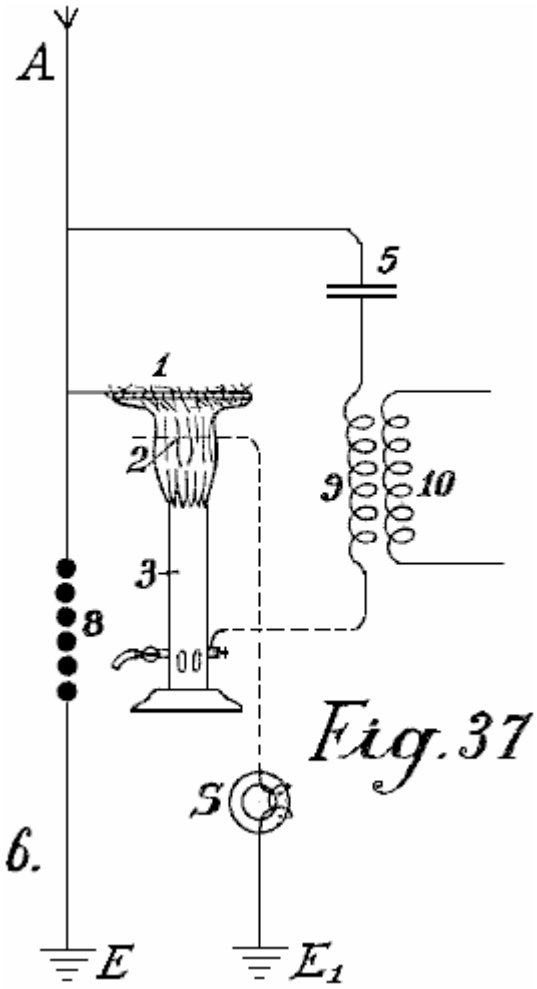


Fig.37 shows a form with a gas flame.

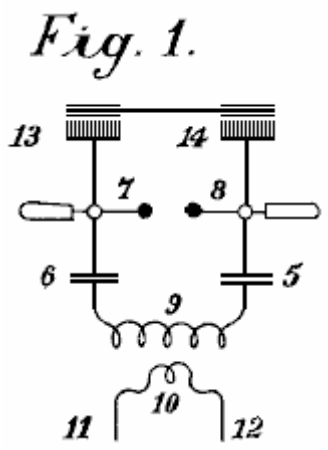
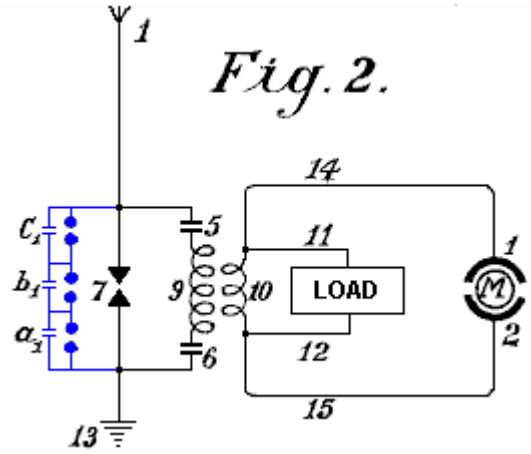


Fig.1 illustrates a simple diagram for converting static electricity into dynamic energy of a high number of oscillations. For the sake of clarity, a Wimshurst machine is assumed to be employed and not an aerial antenna. Items **13** and **14** are combs for collecting the static electricity of the influence machine. Items **7** and **8** are spark-discharging electrodes. Items **5** and **6** are capacitors, **9** is the primary winding of an inductive coil, **10** is the secondary winding whose ends are **11** and **12**. When the disc of the static influence machine is rotated by mechanical means, the combs collect the electric charges, one being positive and one negative and these charge the capacitors **5** and **6** until such a high voltage is developed across the spark gap **7-- 8** that the spark gap is jumped. As the spark gap forms a closed circuit with capacitors **5** and **6**, and inductive resistance **9**, as is well known, waves of high frequency electromagnetic oscillations will pass in this circuit.

The high frequency of the oscillations produced in the primary circuit induces waves of the same frequency in the secondary circuit. Thus, in the primary circuit, electromagnetic oscillations are formed by the spark and these oscillations are maintained by fresh charges of static electricity.

By suitably selecting the ratio between the number of turns in the primary and secondary windings, with regard to a correct application of the coefficients of resonance (capacitance, inductance and resistance) the high voltage of the primary circuit may be suitably converted into a low voltage high current output.

When the oscillatory discharges in the primary circuit become weaker or cease entirely, the capacitors are charged again by the static electricity until the accumulated charge again breaks down across the spark gap. All this is repeated as long as electricity is produced by the static machine through the application of mechanical energy to it.



An elementary form of the invention is shown in **Fig.2** in which two spark gaps in parallel are used, one of which may be termed the working gap **7** while the second serves as a safety device for excess voltage and consists of a larger number of spark gaps than the working section, the gaps being arranged in series and which are bridged by very small capacitors a_1, b_1, c_1, which allow uniform sparking in the safety section.

1 is the aerial antenna for collecting charges of atmospheric electricity, **13** is the earth connection of the second part of the spark gap, **5** and **6** are capacitors and **9** is the primary coil winding. When the positive atmospheric electricity seeks to combine with the negative earth charge via aerial **1**, this is prevented by the air gap between the spark gaps. The resistance of spark gap **7** is lower than that of the safety spark gap set of three spark gaps connected in series a which consequently has three times greater air resistance.

Therefore, so long as the resistance of spark gap **7** is not overloaded, discharges take place only through it. However, if the voltage is increased by any influence to such a level that it might be dangerous for charging the capacitors **5** and **6**, or for the coil insulation of windings 9 and 10, the safety spark gap set will, if correctly set, discharge the voltage directly to earth without endangering the machine. Without this second spark gap arrangement, it is impossible to collect and render available large quantities of electrical energy.

The action of this closed oscillation circuit consisting of spark gap **7**, two capacitors **5** and **6**, primary coil **9** and secondary coil **10**, is exactly the same as that of **Fig.1** which uses a Wimshurst machine, the only difference being the provision of the safety spark gap. The high frequency electromagnetic alternating current can be tapped off through the conductors **11** and **12** for lighting and heating purposes. Special motors adapted for working with static electricity or high frequency oscillations may be connected at **14** and **15**.

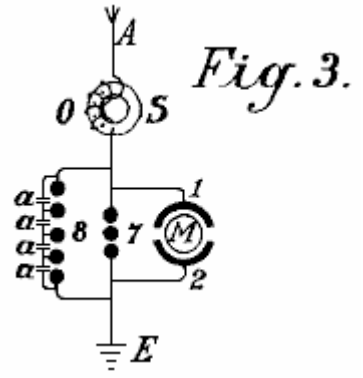


Fig. 3.

In addition to the use of spark gaps in parallel, a second measure of security is also necessary for taking the current from this circuit. This is the introduction of protective electromagnets or choking coils in the aerial circuit as shown by **S** in **Fig.3**. A single electromagnet having a core of the thinnest possible separate laminations is connected with the aerial. In the case of high voltages in the aerial network or at places where there are frequent thunderstorms, several such magnets may be connected in series.

In the case of large units, several such magnets can be employed in parallel or in series parallel. The windings of these electromagnets may be simply connected in series with the aerials. In this case, the winding preferably consists of several thin parallel wires, which together, make up the necessary cross-sectional area of wire. The winding may be made of primary and secondary windings in the form of a transformer. The primary winding will then be connected in series with the aerial network, and the secondary winding more or less short-circuited through a regulating resistor or an induction coil. In the latter case it is possible to regulate, to a certain extent, the effect of the choking coils. In the following circuit and constructional diagrams , the aerial electromagnet choke coil is indicated by a simple ring **S**.

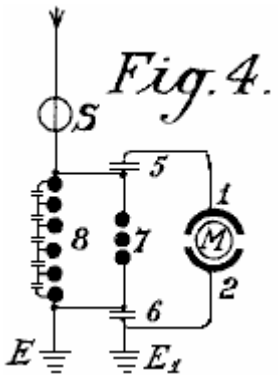
Fig.3 shows the most simple way of converting atmospheric electricity into electromagnetic wave energy by the use of special motors adapted for high oscillatory currents or static charges of electrical energy. Recent improvements in motors for working with static energy and motors working by resonance, that is to say, having groups of tuned electromagnetic co-operating circuits render this possible but such do not form part of the present invention.

A motor adapted to operate with static charges, will for the sake of simplicity, be shown in the diagrams as two semi-circles **1** and **2** and the rotor of the motor by a ring **M** (**Fig.3**). **A** is a vertical aerial or aerial network. **S** is the safety choke or electromagnet with coil **O** as may be seen is connected with the aerial **A**. Adjacent to the electromagnet **S**, the aerial conductor is divided into three circuits, circuit **8** containing the safety spark gap, circuit **7** containing the working spark gap, and then a circuit containing the stator terminal **1**, the rotor and stator terminal **2** at which a connection is made to the earth wire. The two spark gaps are also connected metallically with the earth wire. The method of working in these diagrams is as follows:

The positive atmospheric electric charge collected tends to combine with the negative electricity (or earth electricity) connected via the earth wire. It travels along the aerial **A** through the electromagnet **S** without being checked as it flows in the same direction as the direct current. Further, its progress is arrested by two spark gaps placed in the way and the stator capacitors. These capacitors charge until their voltage exceeds that needed to jump the spark gap **7** when a spark occurs and an oscillatory charge is obtained via the closed oscillation circuit containing motor **M**. The motor here forms the capacity and the necessary inductance and resistance, which as is well known, are necessary for converting static electricity into electromagnetic wave energy.

The discharges are converted into mechanical energy in special motors and cannot reach the aerial network because of the electromagnet or choke. If, however, when a spark occurs at spark gap **7**, a greater quantity of atmospheric electricity tends to flow to earth, then a counter voltage is induced in the electromagnet, which is greater the more rapidly and strongly the flow of current direct to earth is. This opposing voltage causes the circuit to exhibit a sufficiently high resistance to prevent a short circuit between the atmospheric electricity and the earth.

The circuit containing spark gap **8**, having a different wave length which is not in resonance with the natural frequency of the motor, does not endanger the motor and serves as security against excess voltage, which, as practical experiments have shown, may still arise in certain cases.



In **Fig.4**, spark gap **7** is shunted across capacitors **5** and **6** from the motor **M**. This arrangement provides improved over-voltage protection for the motor and it gives a uniform excitation through the spark gap **7**.

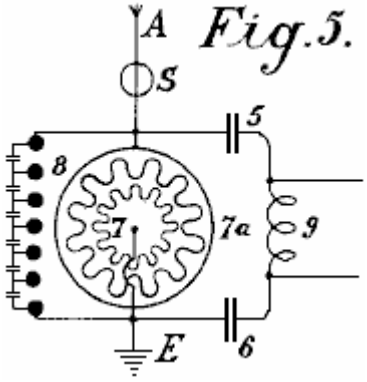


Fig.5 shows an arrangement for producing large currents which can be used direct without motors, to provide heating and lighting. The main difference here is that the spark gap consists of a star-shaped disc **7** which can rotate on its own axis and is rotated by a motor opposite similarly fitted electrodes **7a**. When separate points of starts face one another, discharges take place, thus forming an oscillation circuit with capacitors **5** and **6** and inductor **9**. It is evident that a motor may also be connected directly to the ends of inductor **9**.

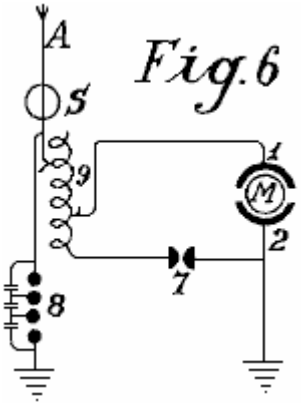
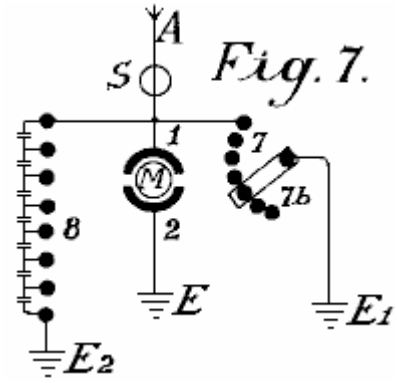


Fig.6 shows how the oscillation circuit may have a motor connected via a variable inductor which opposes any excess voltages which might be applied to the motor. By cutting the separate coils **9** (coupled inductively to the aerial) in or out, the inductive action on the motor may be more or less increased, or variable aerial action may be exerted on the oscillation circuit.



In **Fig.7** the oscillation circuit is closed through the earth (**E** and $\mathbf{E_1}$). The spark gap **7** may be increased or reduced by means of a contact arm **7b**.

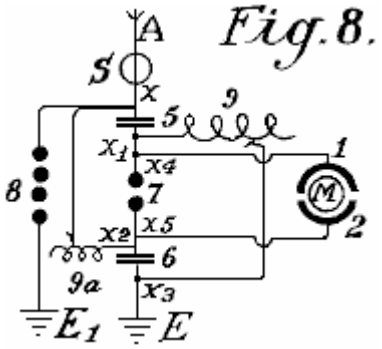


Fig.8 shows a unipolar connection of the motor with the aerial network. Here, two oscillation circuits are closed through the same motor. The first oscillation circuit passes from aerial **A** through electromagnet **S**, point **x**, inductance **9a** to the earth capacitor **6**, across spark gap **7** to the aerial capacitor **5** and back to point **x**. The second oscillation circuit starts from the aerial **5** at the point **x1** through inductor **9** to the earth capacitor **6** at the point **x3**, through capacitor **6**, across spark gap **7** back to point **x1**. The motor itself, is inserted between the two points of spark gap **7**. This arrangement produces slightly dampened oscillation wave currents.

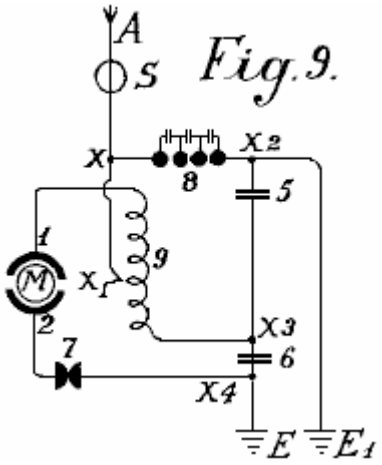


Fig.9 shows a loosely coupled system intended for small motors for measuring purposes. **A** is the serial, **S** is the electromagnet or aerial inductor, **9** the inductor, **7** the spark gap, **5** and **6** capacitors, **E** the earth, **M** the motor, and **1** and **2** the stator connections of the motor which is directly connected to the oscillator circuit.

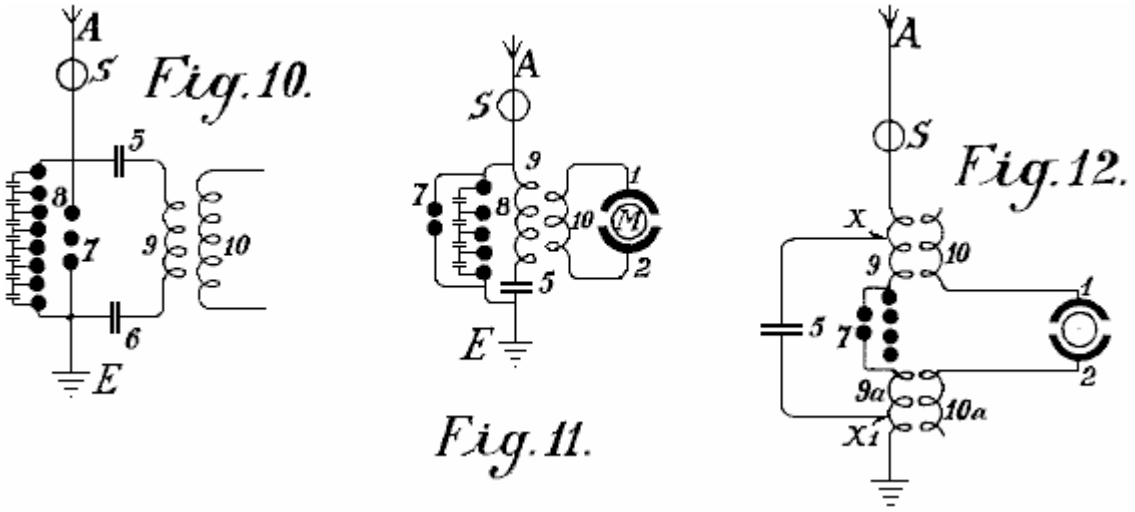
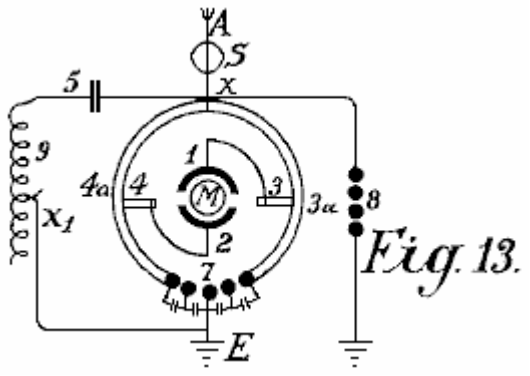


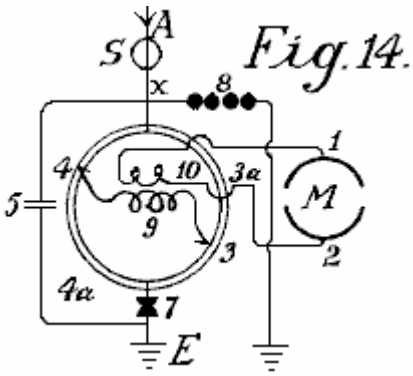
Fig.10 shows a motor circuit with purely inductive coupling. The motor is connected with the secondary wire **10** as may be seen in **Fig.11** in a somewhat modified circuit. The same applies to the circuit of **Fig.12**.

The circuit diagrams shown so far, allow motors of small to medium strength to be operated. For large aggregates, however, they are too inconvenient as the construction of two or more oscillation circuits for large amounts of energy is difficult; the governing is still more difficult and the danger in switching on or off is greater.

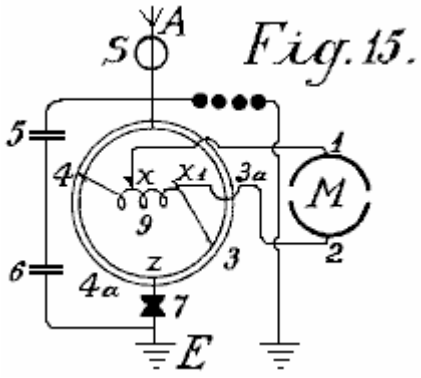


A means for overcoming such difficulties is shown in **Fig.13**. The oscillation circuit shown here, runs from point **x** over capacitor **5**, variable inductor **9**, spark gap **7** and the two segments **3a** and **3b** forming arms of a Wheatstone bridge, back to **x**. If the motor is connected by brushes **3** and **4** transversely to the two arms of the bridge as shown in the drawing, electromagnetic oscillations of equal sign are induced in the stator surfaces **1** and **2** and the motor does not revolve. If however, the brushes **3** and **4** are moved in common with the conducting wires **1** and **2** which connect the brushes with the stator poles, a certain alteration or displacement of the polarity is obtained and the motor commences to revolve.

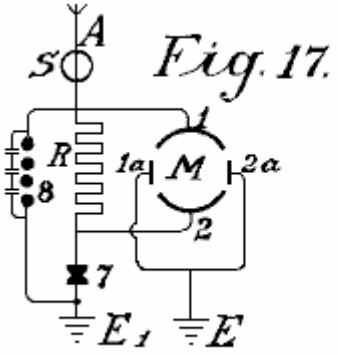
The maximum action will result if one brush **3** comes on the central sparking contact **7** and the other brush **4** on the part **x**. In practice however, they are usually brought on to the central contact **7** but only held in the path of the bridge segments **4a** and **3a** in order to avoid connecting the spark gaps with the motor oscillation circuit.

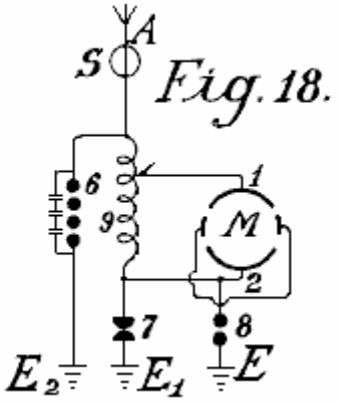


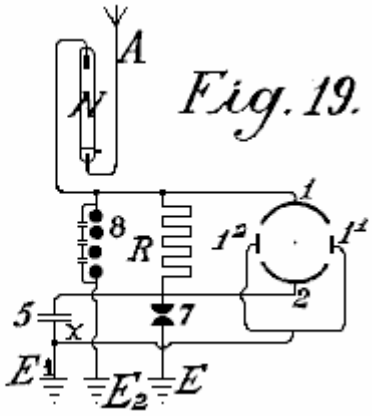
As this prevents the whole of the oscillation energy acting on the motor, it is better to adopt the modification shown in **Fig.14**. The only difference here is that the motor is not wired directly to the segments of the commutator, but instead it is wired to secondary coil **10** which receives induced current from primary coil **9**. This arrangement provides a good transforming action, a loose coupling and an oscillation circuit without a spark gap.



In **Fig.15**, the motor is wired directly to the primary coil at **x** and **x1** after the principle of the auto-transformer. In **Fig.16**, instead of an inductor, capacitor **6** replaces the inductance and is inserted between the segments **3a** and **4a**. This has the advantage that the segments **3a** and **4a** need not be made of solid metal, but may consist of spiral coils which allow a more exact regulation, and high inductance motors may be used.







The circuits shown in **Fig.17**, **Fig.18** and **Fig.19** may be used with resonance and particularly with induction capacitor motors; between the large stator induction capacitor surfaces, small reversing pole capacitors are connected which are lead together to earth. Such reversing poles have the advantage that, with large quantities of electrical energy, the spark formation between the separate oscillation circuits ceases.

Fig.19 shows another method which prevents high frequency electromagnetic oscillations formed in the oscillation circuit, feeding back to the aerial. It is based on the well known principle that a mercury lamp, one electrode of which is formed of mercury, the other of solid metal such as steel, allows an electric charge to pass in only one direction: from the mercury to the steel and not vice versa. The mercury electrode of the vacuum tube **N** is therefore connected with the aerial conductor and the steel electrode with the oscillation circuit. Charges can then only pass from the aerial through the vacuum tube to the oscillation circuit and no flow occurs in the opposite direction. In practice, these vacuum tubes must be connected behind an electromagnet as the latter alone provides no protection against the danger of lightning.

As regards the use of spark gaps, all arrangements as used for wireless telegraphy may be used. Of course, the spark gaps in large machines must have a sufficiently large surface. In very large stations they are cooled in liquid carbonic acid or better still, in liquid nitrogen or hydrogen; in most cases the cooling may also take place by means of liquefied low homologues of the metal series or by means of hydrocarbons, the freezing point of which lies between -90^{o}C and -40^{0}C. The spark gap casing must also be insulated and be of sufficient strength to be able to resist any pressure which may arise. Any undesirable excess super-pressure which may be formed must

be let off automatically. I have employed with very good results, mercury electrodes which were frozen in liquid carbonic acid, the cooling being maintained during the operation from the outside, through the walls.

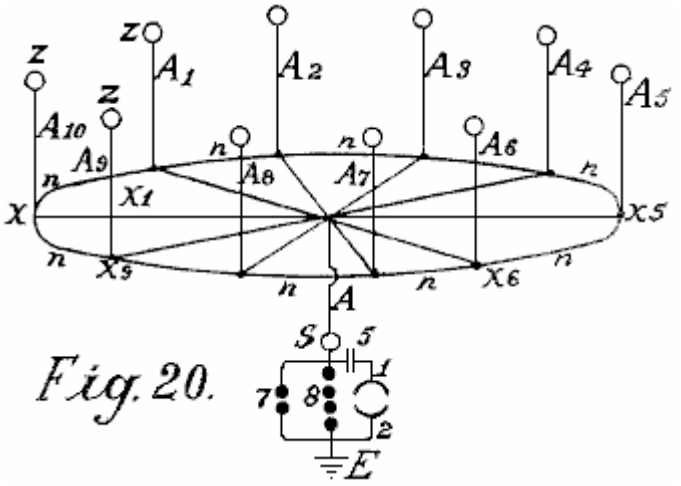


Fig. 20.

Fig.20 shows one of the most simple forms of construction of an aerial network in combination with collectors, transformers and the like. **E** is the earth wire, **8** the safety spark gap, **7** the working spark gap, **1** and **2** the stator surfaces of the motor, **5** a capacitor battery, **S** the protective magnet which is connected with the coil in the aerial conductor, $\mathbf{A^1}$ to $\mathbf{A^{10}}$ aerial antennae with collecting balloons, **N** horizontal collecting or connecting wires, from which, a number of connections run to the centre.

The actual collectors consist of metal sheaths, preferably made of an aluminium magnesium alloy, and are filled with hydrogen or helium, and are attached to copper-plated steel wires. The size of the balloon is selected so that the actual weight of the balloon and its conducting wire is supported by it. Aluminium spikes, made and gilded as described below, are arranged on top of the balloons in order to produce a conductor action. Small quantities of radium preparations, more particularly, polonium-ionium or mesothorium preparations, considerably increase the ionisation, and the performance of these collectors.

In addition to metal balloons, fabric balloons which are sprayed with a metallic coating according to Schoop's metal-spraying process may also be used. A metallic surface may also be produced by lacquering with metallic bronzes, preferably according to Schoop's spraying process, or lacquering with metallic bronze powders in two electrical series of widely different metals, because this produces a considerably increased collecting effect.

Instead of the ordinary round balloons, elongated cigar-shaped ones may be employed. In order also to utilise the frictional energy of the wind, patches or strips of non-conducting substances which produce electricity by friction, may be attached to the metallised balloon surfaces. The wind will impart a portion of its energy in the form of frictional electricity, to the balloon casing, thus substantially increasing the collection effect.

In practice however, very high towers of up to 300 metres may be employed as antennae. In these towers, copper tubes rise freely further above the top of the tower. A gas lamp secured against the wind is then lit at the point of the copper tube and a netting is secured to the copper tube over the flame of this lamp to form a collector. The gas is conveyed through the interior of the tube, up to the summit. The copper tube must be absolutely protected from moisture at the place where it enters the tower, and rain must be prevented from running down the walls of the tower, which might lead to a bad catastrophe. This is done by bell-shaped enlargements which expand downwards, being arranged in the tower in the form of high voltage insulators of Siamese pagodas.

Special attention must be devoted to the foundations of such towers. They must be well insulated from the ground, which may be achieved by first embedding a layer of concrete in a box form to a sufficient depth in the ground, and inserting in this, an asphalt lining and then glass bricks cast about 1 or 2 metres in thickness. Over this in turn, there is a ferro-concrete layer in which alone the metal foot of the tube is secured. This concrete block must be at least 2 metres from the ground and at the sides, be fully protected from moisture by a wooden covering. In the lower part of the tower, a wood or glass housing should be constructed to protect the capacitors and/or motors. In order to ensure that the ground lead connects to the water-table, a well insulated pit lined with vitreous bricks must be provided. Several such towers are erected at equal distances apart and connected with a horizontal conductor. The horizontal connecting wires may either run directly from tower to tower or be carried on bell-shaped insulators similar to those in use for high voltage electricity transmission lines. The width of the aerial tower network may be of any suitable size and the connection of the motors can take place at any convenient location.

Fig. 21.

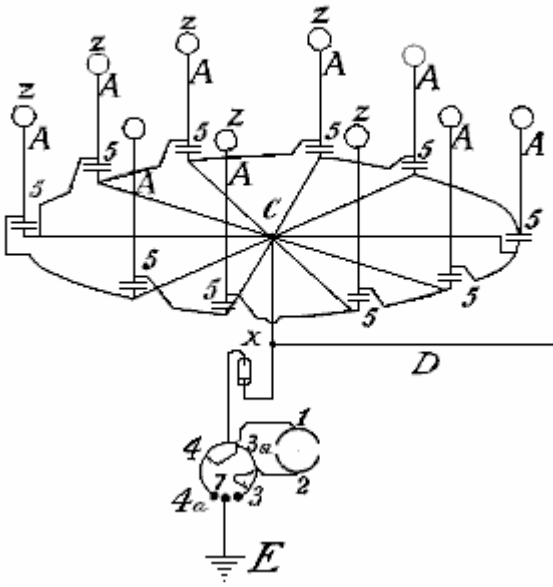
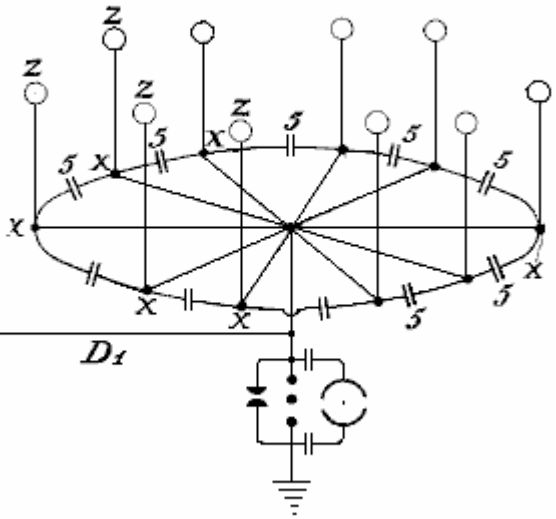


Fig. 22.



In order to collect large quantities of electricity with few aerials, it is as well to provide the aerial conductor with sets of capacitors as shown in the two methods of construction illustrated in **Fig.21** and **Fig.22**. In **Fig.21** the set of capacitors **5** is connected between the aerials **Z** via lead **A** and an annular conductor from which horizontal run to the connecting points **C** to which the earth wire is connected. **Fig.22** shows a similar arrangement.

Should two such series of antenna rings be shown by a voltmeter to have a large voltage difference (for example, one in the mountains and one on the plain) or even of a different polarity, these differences may be compensated for by connecting sufficiently large capacitor sets (**5, 5a, 5b**) by means of Maji star conductors **D** and **D[1]**. **Fig.23**, shows a connection of three such rings of collectors are positioned in a triangle with a central set of capacitors.

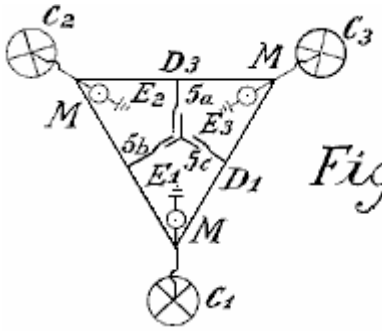


Fig. 23.

The capacitor sets of such large installations must be embedded in liquefied gasses or in liquids freezing at very low temperatures. In such cases, a portion of the atmospheric energy must be employed for liquefying these gasses. It is also preferable to employ pressure. By this means, the capacitor surfaces may be reduced in area and still allow the storage of large quantities of energy to be stored, secure against breakdown. For the smaller installations, the immersing of the capacitors in well insulated oil or the like, is sufficient. Solid substances, on the other hand, cannot be employed as insulators.

The arrangement in the diagrams shown earlier has always shown both poles of the capacitors connected to the aerial conductors. An improved method of connection has been found to be very advantageous. In this method, only one pole of each capacitor is connected to the collecting network. Such a method of connection is very important, as by means of it, a constant current and an increase in the normal working voltage is obtained. If, for example, a collecting balloon aerial which is allowed to rise to a height of 300 metres, shows 40,000 volts above earth voltage, in practice it has been found that the working voltage (with a withdrawal of the power as described earlier by means of oscillating spark gaps and the like) is only about 400 volts. If however, the capacity of the capacitor surfaces be increased, which capacity in the above mentioned case was equal to that of the collecting surface of the balloon aerials, to double the amount, by connecting the capacitors with only one pole, the voltage rises under an equal withdrawal of current up to and beyond 500 volts. This can only be ascribed to the favourable action of the connecting method.

In addition to this substantial improvement it has also been found preferable to insert double inductances with electromagnets and to place the capacitors preferably between two such electromagnets. It has also been found that the useful action of such capacitors can be further increased if an induction coil is connected as an inductive resistance to the unconnected pole of the capacitor, or still better if the capacitor itself be made as an induction capacitor. Such a capacitor may be compared to a spring, which when compressed, carries in itself accumulated force, which it gives off again when released. In charging, a charge with reversed sign is formed at the other free capacitor pole, and if a short circuit occurs through the spark gap, the accumulated energy is again given back since now new quantities of energy are induced at the capacitor pole connected to the conductor network, which in fact, charges with opposite sign to that at the free capacitor pole. The new induced charges have of course, the same sign as the collector network. The whole voltage energy in the aerial is thereby increased. In the same time interval, larger quantities of energy are accumulated than is the case without such capacitor sets being inserted.

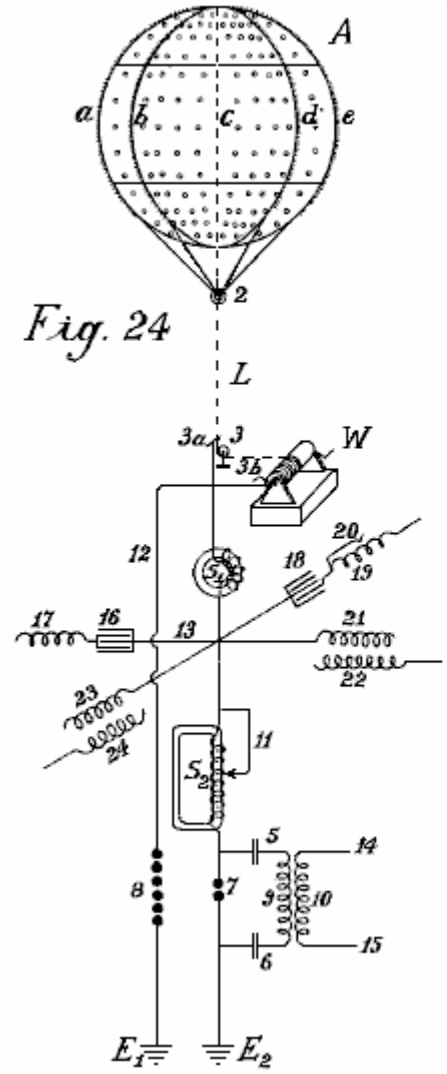


Fig. 24

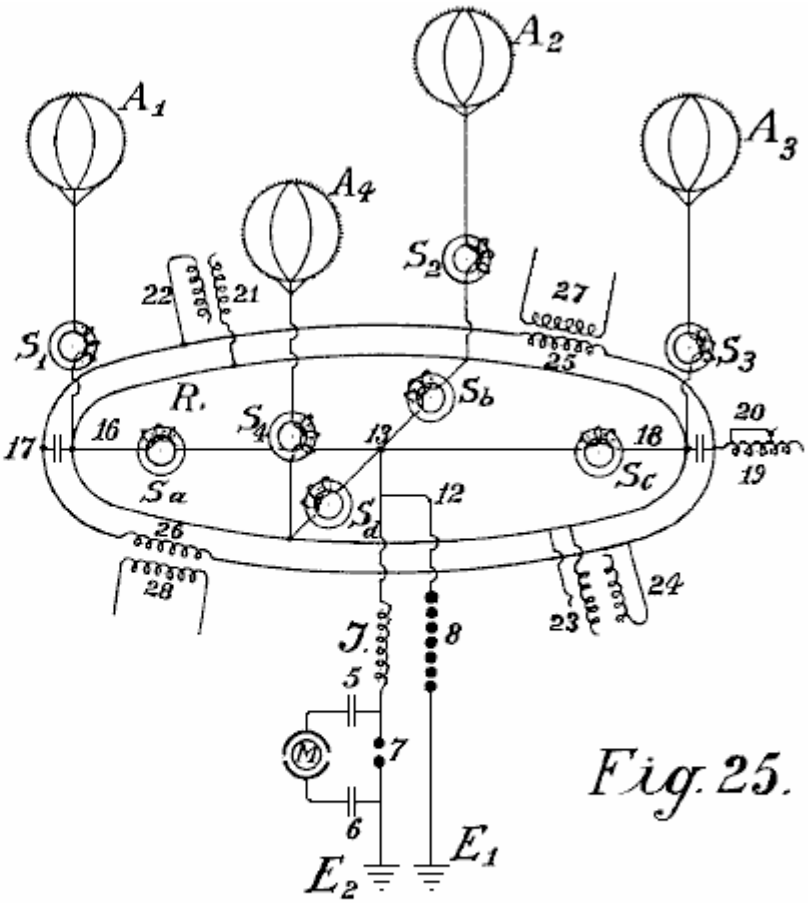


Fig. 25.

In **Fig.24** and **Fig.25**, two different connection diagrams are illustrated in more detail. **Fig.24** shows a collecting balloon along with its earth connections. **Fig.25** shows four collecting balloons and the parallel connection of their capacitor sets.

A is the collecting balloon made of an aluminium magnesium alloy (electron metal magnalium) of a specific gravity of 1.8 and a plate thickness of 0.1 mm to 0.2 mm. Inside, there are eight strong vertical ribs of T-shaped section of about 10 mm to 20 mm in height and about 3 mm in thickness, with the projecting part directed inwards (indicated by **a, b, c, d** and so forth). They are riveted together to form a firm skeleton and are stiffened in a horizontal direction by two cross ribs. The ribs are further connected to one another internally and transversely by means of thin steel wires, whereby the balloon obtains great strength and elasticity. Rolled plates of 0.1 mm to 0.2 mm in thickness made of magnalium alloy are then either soldered or riveted on to this skeleton so that a fully metallic casing with a smooth external surface is created. Well silvered or coppered aluminium plated steel wires run from each rib to the fastening ring **2**. Further, the coppered steel hawser **L**, preferably twisted out of separate thin wires (shown as dotted lines in **Fig.24**) and which must be long enough to allow the balloon to rise to the

desired height, leads to a metal roller or pulley **3** and on to a winch **W**, which must be well insulated from the earth. By means of this winch, the balloon which is filled with hydrogen or helium, can be allowed to rise to a suitable height of 300 to 5,000 metres, and brought to the ground for recharging or repairs.

The actual current is taken directly through a friction contact from the metal roller **3** or from the wire or even from the winch, or simultaneously from all three by means of brushes (**3**, **3a** and **3b**). Beyond the brushes, the conductor is divided, the paths being:- firstly, over **12** to the safety spark gap **8**, on to the earth conductor $\mathbf{E^1}$, and secondly over electromagnet $\mathbf{S^1}$, point **13**, to a second loose electromagnet having an adjustable coil $\mathbf{S^2}$, then to the spark gap **7** and to the second earth conductor $\mathbf{E^2}$. The actual working circuit is formed through the spark gap **7**, capacitors **5** and **6**, and through the primary coil **9**; here the static electricity formed by oscillatory discharges is accumulated and converted into high frequency electromagnetic oscillations. Between the electromagnets $\mathbf{S^1}$ and $\mathbf{S^2}$ at the crossing point **13**, four capacitor sets are introduced which are only indicated diagramatically in the drawings by a single capacitor. Two of these sets of capacitors (**16** and **18**) are made as plate capacitors and prolonged by regulating induction coils or spirals **17** and **19** while the two others (**21** and **23**) are induction capacitors. As may be seen from the drawings, each of the four capacitor sets, **16**, **18**, **21** and **23** is connected by only one pole to either the aerial or to the collector conductor. The second poles **17**, **19**, **22** and **24** are open. In the case of plate capacitors having no inductive resistance, an induction coil is inserted. The object of such a spiral or coil is the displacement of phase of the induction current by $^1/_4$ periods, whilst the charging current of the capacitor poles which lie free in the air, works back to the collector aerial. The consequence of this is that in discharges in the collector aerial, the back-inductive action of the free poles allows a higher voltage to be maintained in the aerial collecting conductor than would otherwise be the case. It has also been found that such a back action has an extremely favourable effect on the wear of the contacts. Of course, the inductive effect may be regulated at will within the limits of the size of the induction coil, the length of the coil in action being adjustable by means of wire connection without induction (see **Fig.24** No. **20**).

$\mathbf{S^1}$ and $\mathbf{S^2}$ may also be provided with such regulating devices, in the case of $\mathbf{S^2}$ illustrated by **11**. If excess voltage be formed, it is conducted to earth through wire **12** and spark gap **8**, or through any other suitable apparatus, since this voltage would be dangerous for the other components. The action of these capacitor sets has already been described.

The small circles on the collector balloon indicate places where small patches of extremely thin layers (0.01 to 0.05 mm thick) of zinc amalgam, gold amalgam or other photoelectric acting metals, are applied to the balloon casing of light metal. Such metallic patches may also be applied to the entire balloon as well as in greater thickness to the conducting network. The capacity of the collector is thereby considerably strengthened at the surface. The greatest possible effect in collecting may be obtained by polonium amalgams and the like. On the surface of the collector balloon, metal points or spikes are also fixed along the ribs. These spikes enhance the charge collection operation. Since it is well known that the sharper the spikes, the less the resistance of the spikes, it is therefore extremely important to use spikes which are as sharp as possible. Experiments have shown that the formation of the body of the spike or point also play a large part, for example, spikes made of bars or rollers with smooth surfaces, have point resistance many times greater than those with rough surfaces. Various kinds of spike bodies have been experimented with for the collector balloons and the best results were given with spikes which were made in the following way: Fine points made of steel, copper, nickel or copper and nickel alloys, were fastened together in bundles and then placed as anode with the points placed in a suitable electrolyte (preferably in hydrochloric acid or muriate of iron solutions) and so treated with weak current driven by 2 to 3 volts. After 2 to 3 hours, according to the thickness of the spikes, the points become extremely sharp and the bodies of the spikes have a rough surface. The bundle can then be removed and the acid washed off with water. The spikes are then placed as cathode in a bath containing a solution of gold, platinum, iridium, palladium or wolfram salts or their compounds, and coated at the cathode galvanically with a thin layer of precious metal, which mush however be sufficiently firm to protect them from atmospheric oxidation.

Such spikes act at a 20 fold lower voltage almost as well as the best and finest points made by mechanical means. Still better results are obtained if polonium or radium salts are added to the galvanic bath when forming the protective layer or coating. Such pins have low resistance at their points and have excellent collector action even at one volt or lower.

In **Fig.24**, the three unconnected poles are not connected with one another in parallel. That is quite possible in practice without altering the principle of the free pole. It is also preferable to interconnect a series of collecting aerials in parallel to a common collector network. **Fig.25** shows such an arrangement. $\mathbf{A^1, A^2, A^3, A^4}$ are four metal collector balloons with gold or platinum coated spikes which are electrolytically mad in the presence of polonium emanations or radium salts, the spikes being connected over four electromagnets $\mathbf{S^1, S^2, S^3, S^4}$, through an annular conductor **R**. From this annular conductor, four wires run over four further electromagnets $\mathbf{S^a, S^b, S^c, S^d}$, to the connecting point **13**. There, the conductor is divided, one branch passing over **12** and the safety spark gap **7** to the earth at $\mathbf{E^1}$, the other over inductive resistance **J** and working spark gap **7** to the earth at

E^2. The working circuit, consisting of the capacitors **5** and **6** and a resonance motor or a capacitor motor **M**, such as already described, is connected in proximity around the sparking gap section **7**. Of course, instead of connecting the capacitor motor directly, the primary circuit for high frequency oscillatory current may also be inserted.

The capacitor sets are connected by one pole to the annular conductor **R** and can be either inductionless (**16** and **18**) or made as induction capacitors as shown by **21** and **23**. The free poles of the inductionless capacitors are indicated by **17** and **19**, and those of the induction capacitors by **22** and **24**. As may be seen from the drawings, all of these poles **17, 22, 19** and **24** may be interconnected in parallel through a second annular conductor without any fear that thereby the principle of the free pole connection will be lost. In addition to the advantages already mentioned, the parallel connection also allows an equalisation of the working voltage in the entire collector network. Suitably calculated and constructed induction coils **25** and **26** may also be inserted in the annular conductor of the free poles, by means of which, a circuit may be formed in the secondary coils **27** and **28** which allows current produced in this annular conductor by fluctuations of the charges, to be measured or otherwise utilised.

According to what has already been stated, separate collector balloons may be connected at equidistant stations distributed over the whole country, either connected directly with one another metallically or by means of intermediate suitably connected capacitor sets through high voltage conductors insulated from earth. The static electricity is converted through a spark gap, into high frequency dynamic electricity which may be utilised as a source of energy by means of a suitable connection method, various precautions being observed, and with special regulations. The wires leading from the collector balloons, have up to now been connected through an annular conductor without this endless connection, which can be regarded as an endless induction coil, being able to exert any action on the whole conductor system.

It has now been found that if the network conductor connecting the aerial collector balloons with one another, is not made as a simple annular conductor, but preferably short-circuited in the form of coils over a capacitor set or spark gap or through thermionic valves, then the total collecting network exhibits quite new properties. The collection of atmospheric electricity is thereby not only increased but an alternating field may easily be produced in the collector network. Further, the atmospheric electrical forces showing themselves in the higher regions, may also be obtained directly by induction. In **Fig.26** and **Fig.28**, a form of construction is shown, on the basis of which, the further foundations of the method will be explained in more detail.

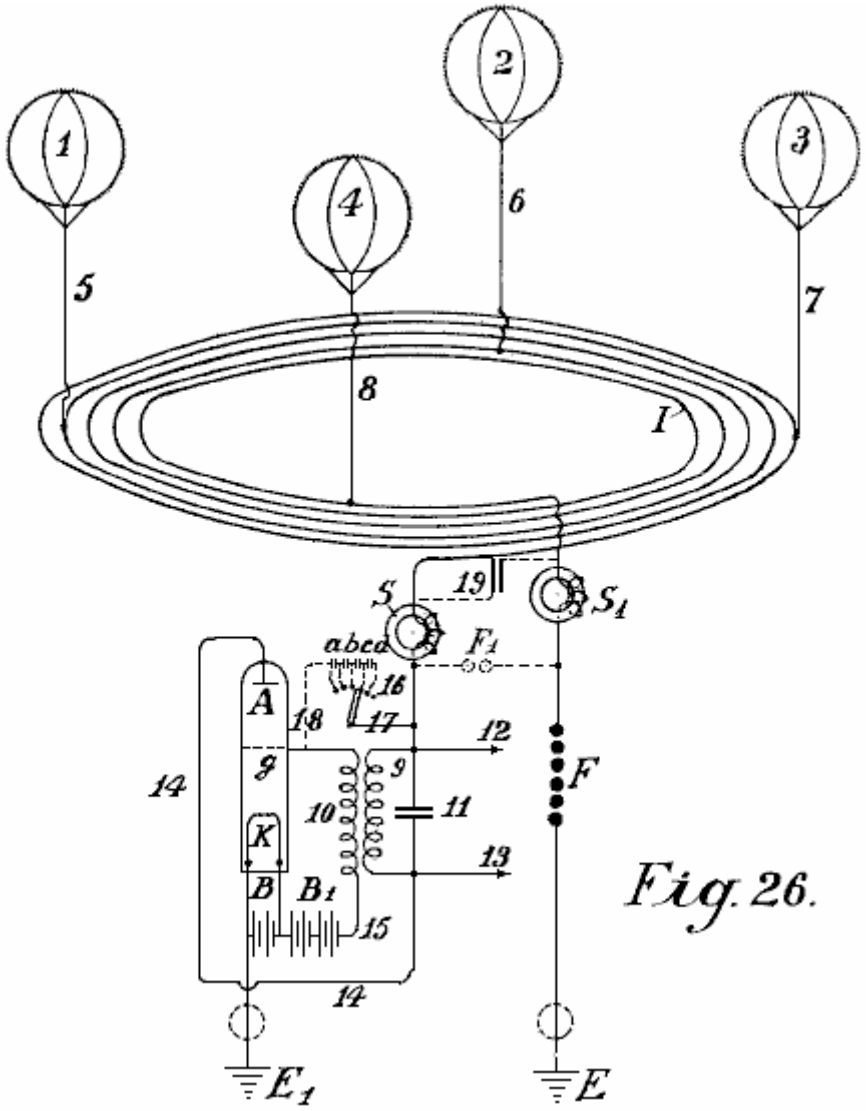


Fig. 26.

In **Fig.26, 1,2,3** and **4** are metallic collector balloons, with **5, 6, 7** and **8** their metallic aerial conductors and **I** the actual collector network. This consists of five coils and is mounted on high voltage insulators in the air, on high voltage masts (or with a suitable construction of cable, embedded in the earth). One coil has a diameter of 1 to 100 km. or more. **S** and **S^1** are two protective electromagnets, **F** is the second safety section against excess voltage, **E** its earth conductor and **E^1** the earth conductor of the working section. When an absorption of static atmospheric electricity is effected through the four balloon collectors, in order to reach the earth connection **E^1**, the current must flow spirally through the collector network, over the electromagnet **S**, primary induction coil **9**, conductor **14**, anode **A** of the audion tube, incandescent cathode **K**, as the way over the electromagnet and safety spark gap **F** offers considerably greater resistance. Owing to the fact that the accumulated current flows in one direction, an electromagnetic alternating field is produced in the interior of the collector network coil, whereby all of the free electrons are directed more or less into the interior of the coil. An increased ionisation of the atmosphere is therefore produced. Consequently, the points mounted on the collector balloon, show a considerably reduced resistance and therefore increased static charges are produced between the points on the balloon and the surrounding atmosphere. This results in a considerably increased collector effect.

A second effect, which could not be achieved in any other way, is obtained by the alternating electromagnetic field running parallel to the earth's surface, which acts more or less with a diminishing or increasing effect on the earth's magnetic field, whereby in the case of fluctuations in the current, a return induction current of reversed sign is always produced in the collector coil by earth magnetism. Now if a constantly pulsating, continuous alternating field is produced as stated in the collector network **I**, an alternating current of the same frequency is also produced in the collecting network coil. As the same alternating field is further transmitted to the aerial balloon, the resistance of its points is thereby considerably reduced, while the collector action is considerably increased. A further advantage is that positive charges which collect on the metal surfaces during the conversion into dynamic current, produce a so-called voltage drop in the collector area. As an alternating field is present, when discharge of the collector surfaces takes place, the negative ions surrounding the collector surfaces produce, by the law of induction, an induction of reversed sign on the collector surface - that is, a positive charge. In addition to the advantages already stated, the construction of connecting conductors in coil form, when of

sufficiently large diameter, allows a utilisation of energy arising in higher regions, also in the most simple way. As is well known, electric discharges frequently take place at very great elevations which may be observed, such as 'St. Elmo's fires' or 'northern lights'. These energy quantities have not been able to have been utilised before now. By this invention, all of these kinds of energy, as they are of electromagnetic nature and since the axis of the collector coils is at right angles to the earth's surface, can be absorbed in the same way as a radio absorbs distant radio signals. With a large diameter of the spiral, it is possible to connect large surfaces and thereby take up large quantities of energy.

It is well known that in the summer months and in the tropics, large radio stations are very frequently unable to receive signals due to interruptions caused by atmospheric electricity, and this takes place with vertical coils of only 40 to 100 metres in diameter. If, on the contrary, horizontal coils of 1 to 100 kilometres in diameter are used, very strong currents may be obtained through discharges which are constantly taking place in the atmosphere. Particularly in the tropics, or still better in the polar regions where the northern lights are constantly present, large quantities of energy may probably be obtained in this way. A coil with several windings should perform the best. In a similar manner, any alteration of the earth's magnetic field should act inductively on such a coil.

It is not at all unlikely that earthquakes and sunspots will also produce an induction in collector coils of that size. In similar manner, this collector conductor will react to earth currents more particularly when they are near the surface of the earth or even embedded in the earth. By combining the previous kind of current collectors, so far as they are adapted for the improved system with the improved possibilities of obtaining current, the quantities of free natural energy which are to be obtained in the form of electricity are considerably increased.

In order to produce uniform undamped current oscillations in the improved collector coil, so-called audion high vacuum or thermionic valves are used instead of the previous described spark gaps (**Fig.26, 9-18**). The main aerial current flows through electromagnet **S** (which in the case of a high number of alternations is not connected here but in the earth conductor $\mathbf{E^1}$) and may be conveyed over the primary coils in the induction winding through wire **14** to the anode **A** of the high vacuum grid valve. Parallel with the induction resistance **9**, a regulating capacity of suitable size, such as capacitor **11**, is inserted. In the lower part of the vacuum grid valve is the incandescent filament cathode **K** which is fed through a battery **B**. From the battery, two branches run, one to the earth conductor $\mathbf{E^1}$ and the other through battery $\mathbf{B^1}$ and secondary coil **10** to the grid anode **g** of the vacuum tube. By the method of connections shown in dotted lines, a desired voltage may also be produced at the grid electrode **g** through wire **17** which is branched off from the main current conductor through switches **16** and some small capacitors (**a, b, c, d**) connected in series, and conductor **18**, without the battery $\mathbf{B^1}$ being required. The action of the whole system is somewhat as follows:-

On the connecting conductor of the aerial collector network being short-circuited to earth, the capacitor pole **11** is charged, and slightly dampened oscillations are formed in the short-circuited oscillation circuit formed by capacitor **11** and self inductance **9**. Because of the coupling through coil **10**, voltage fluctuations of the same frequency take place in the grid circuit **15** and in turn, these fluctuations influence the strength of the electrode current passing through the high vacuum amplifying valve and thus produce current fluctuations of the same frequency in the anode circuit. A permanent supply of energy. Consequently, a permanent supply of energy is supplied to the oscillation circuits **9** and **10** takes place, until a balance is achieved where the oscillation energy consumed exactly matches the energy absorbed. This produces constant undamped oscillations in the oscillation circuits **9 - 11**.

For regular working of such oscillation producers, high vacuum strengthening tubes are necessary and it is also necessary that the grid and anode voltages shall have a phase difference of 180^0 so that if the grid is negatively charged, then the anode is positively charged and vice versa. This necessary difference of phase may be obtained by most varied connections, for example, by placing the oscillating circuit in the grid circuit or by separating the oscillation circuit and inductive coupling from the anodes and the grid circuit, and so forth.

A second important factor is that care must be taken that the grid and anode voltages have a certain relation to one another; the latter may be obtained by altering the coupling and a suitable selection of the self induction in the grid circuit, or as shown by the dotted lines **18, 17, 16** by means of a larger or smaller number of capacitors of suitable size connected in series; in this case, the battery $\mathbf{B^1}$ may be omitted. With a suitable selection of the grid potential, a glow discharge takes place between the grid **g** and the anode **A**, and accordingly at the grid there is a cathode drop and a dark space is formed. The size of this cathode drop is influenced by the ions which are emitted in the lower space in consequence of shock ionisation of the incandescent cathodes **K** and pass through the grid in the upper space. On the other hand, the number of the ions passing through the grid is dependent on the voltage between the grid and the cathode. Thus, if the grid voltage undergoes periodic fluctuations (as in the present case), the amount of the cathode drop at the grid fluctuates, and consequently, the internal resistance of the valve fluctuates correspondingly, so that when a back-coupling of the feed circuit with the grid circuit takes

place, the necessary means are in place for producing undamped oscillations and of taking current as required, from the collecting conductor.

With a suitably loose coupling, the frequency of the undamped oscillations produced is equal to the self-frequency of the oscillation circuits **9** and **10**. By selecting a suitable self-induction for coil **9** and capacitor **11**, it is possible to extend operation from frequencies which produce electromagnetic oscillations with a wavelength of only a few metres, down to the lowest practical alternating current frequency. For large installations, a suitable number of frequency producing tubes in the form of the well known high vacuum transmission tubes of 0.5 kW to 2 kW in size may be connected in parallel so that in this respect, no difficulty exists.

The use of such tubes for producing undamped oscillations, and the construction and method of inserting such transmission tubes in an accumulator or dynamo circuit is known, also, such oscillation producing tubes only work well at voltages of 1,000 volts up to 4,000 volts, so that on the contrary, their use at lower voltages is considerably more difficult. By the use of high voltage static electricity, this method of producing undamped oscillations as compared with that through spark gaps, must be regarded as an ideal solution, particularly for small installations with outputs from 1 kW to 100 kW.

By the application of safety spark gaps, with interpolation of electromagnets, not only is short-circuiting avoided but also the taking up of current is regulated. Oscillation producers inserted in the above way, form a constantly acting alternating electromagnetic field in the collector coil, whereby, as already stated, a considerable accumulating effect takes place. The withdrawal or 'working' wire is connected at **12** and **13**, but current may be taken by means of a secondary coil which is firmly or moveably mounted in any suitable way inside the large collector coil, i.e. in its alternating electromagnetic field, so long as the direction of its axis is parallel to that of the main current collecting coil.

In producing undamped oscillations of a high frequency (50 KHz and more) in the oscillation circuits **9** and **11**, electromagnets **S** and **S[1]** must be inserted if the high frequency oscillations are not to penetrate the collector coil, between the oscillation producers and the collector coil. In all other cases they are connected shortly before the earthing (as in **Fig.27** and **Fig.28**).

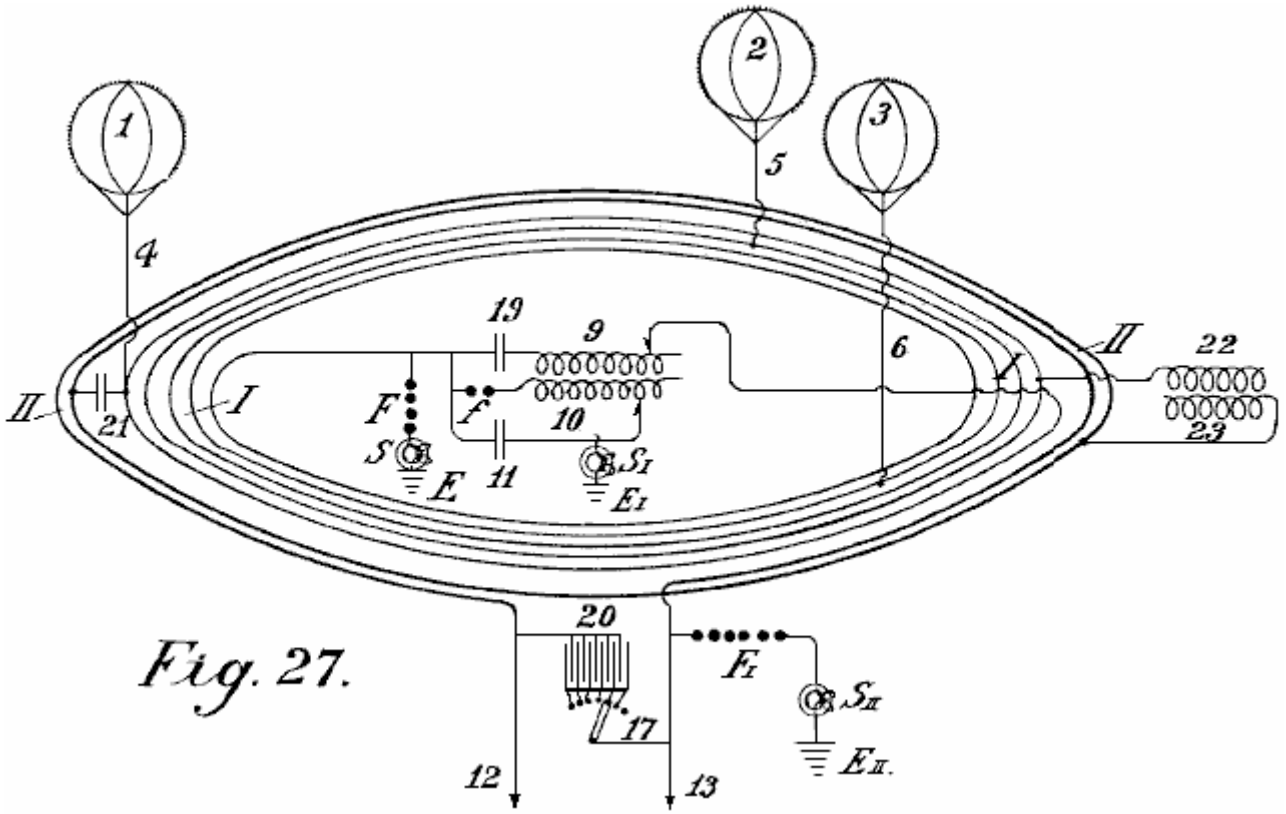


Fig. 27.

In **Fig.27** a second method of construction of the connecting conductor of the balloon aerials is illustrated in the form of a coil. The main difference is that in addition to the connecting conductor **I** another annular conductor **II** is inserted parallel to the former on the high voltage masts in the air (or embedded as a cable in the earth) but both in the form of a coil. The connecting wire of the balloon aerials is both a primary conductor and a current producing network while the coil is the consumption network and is not in unipolar connection with the current producing network.

In **Fig.27** the current producing network **I** is shown with three balloon collectors **1, 2, 3** and aerial conductors **4, 5, 6**; it is short-circuited through capacitor **19** and inductor **9**. The oscillation forming circuit consists of spark gap **f**, inductor **10** and capacitor **11**. The earth wire **E** is connected to earth through electromagnet S^1. F_I is the safety spark gap which is also connected to earth through a second electromagnet S_{II} at E_{II}. On connecting up the capacitor circuit **11** it is charged over the spark gap **f** and an oscillatory discharge is formed. This discharging current acts through inductor **10** on the inductively coupled secondary **9**, which causes a change in the producing network, by modifying the voltage on capacitor **19**. This causes oscillations in the coil-shaped producer network. These oscillations induce a current in the secondary circuit **II**, which has a smaller number of windings and lower resistance, consequently, this produces a lower voltage and higher current in it.

In order to convert the current thus obtained, into current of an undamped character, and to tune its wavelengths, a sufficiently large regulatable capacitor **20** is inserted between the ends **12** and **13** of the secondary conductor **II**. Here also, current may be taken without an earth conductor, but it is advisable to insert a safety spark gap E^1 and to connect this with the earth via electromagnet S^2. The producer network may be connected with the working network **II** over an inductionless capacitor **21** or over an induction capacitor **22, 23**. In this case, the secondary conductor is unipolarly connected with the energy conductor.

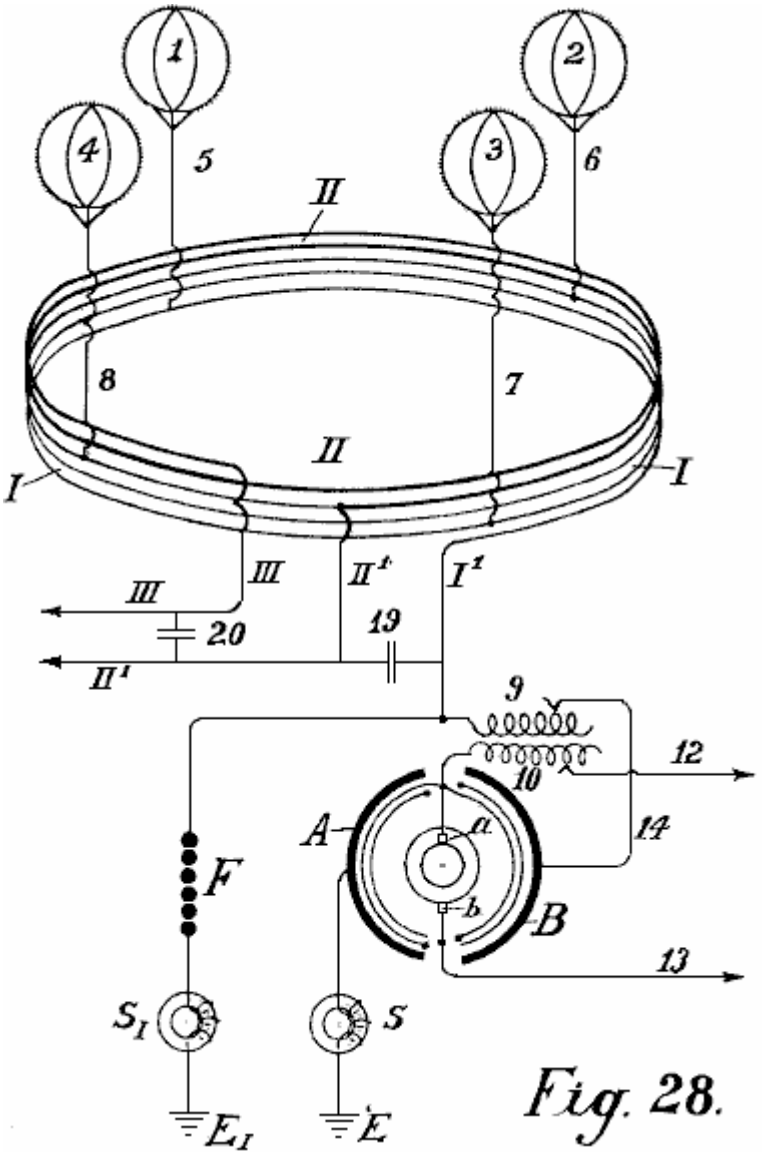


Fig. 28.

In **Fig.28**, the connecting conductor between the separate collecting balloons is carried out according to the autotransformer principle. The collecting coil connects four aerial balloons **1, 2, 3, 4**, the windings of which are not made side-by-side but one above the other. In **Fig.28**, the collector coil **I** is shown with a thin line and the metallically connected prolongation coils **II** with a thick line. Between the ends I^1 and II^1 of the energy network **I**, a regulating capacitor **19** is inserted. The wire I^1 is connected with the output wire and with the spark gap **F**.

As transformer of the atmospheric electricity, an arrangement is employed which consists of using rotary pairs of capacitors in which the stator surface **B** is connected with the main current, while the other **A** is connected to the earth pole. These pairs of short-circuited capacitors are caused to rotate and the converted current can be taken from them via two collector rings and brushes. This current is alternating current with a frequency dependent on the number of balloons and the rate of revolutions of the rotor. As the alternating current formed in the rotor can act through coils **10** on the inductor **9**, an increase or decrease of the feed current in **I** can be obtained according to the direction of the current by back-induction. Current oscillations of uniform rhythm are produced in the coil-shaped windings of the producer network.

As the ends of this conductor are short-circuited through the regulatable capacitor **19**, these rhythms produce short-circuited undamped oscillations in the energy conductor. The frequency of these oscillations can be altered at will by adjusting the capacitance of capacitor **19**. These currents may also be used as working current via the conductors **II**[1] and **III**. By inserting capacitor **20**, a connection between these conductors may also be made, whereby harmonic oscillations of desired wavelength are formed. By this means, quite new effects as regards current distribution are obtained. The withdrawal of current can even take place without direct wire connection if, at a suitable point in the interior of the producing network (quite immaterially whether this has a diameter of 1 or 100 km) a coil tuned to these wavelength and of the desired capacity, is firmly or moveably mounted in the aerial conductor in such a way that its axis is parallel with the axis of the collector coil. In this case, a current is induced in the producing network, the size of which is dependent on the total capacity and resistance and on the frequency selected. A future possibility is taking energy from the producer network by radio signals as in addition to atmospheric electricity, magnetic earth currents and energy from the upper atmosphere may be tapped.

Of course, vacuum tubes may be used to produce undamped oscillations anywhere spark gaps are shown in the circuits. The separate large-diameter coils of the producer network may be connected to one another through separate conductors all in parallel or all in series or in groups in series. By regulating the number of oscillations and the magnitude of the voltage, more or fewer large collector coils of this kind may be used. The coils may also be divided spirally over the entire section. The coils may be carried out in annular form or in triangular, quadrangular, hexagonal or octagonal form.

Of course, wires which form guides for the current waves, may be carried from a suitable place to the centre or also laterally. This is necessary when the currents have to be conducted over mountains and valleys and so forth. In all these cases, the current must be converted into a current of suitable frequency.

As already mentioned, separate collecting balloons may be directly metallically interconnected a equidistant stations distributed over the entire country, or may be connected by interpolation of suitable capacitor sets by means of high voltage conductors. The static electricity is converted through a spark gap into dynamic energy of high frequency and could then in that form be used as an energy source after special regulation.

According to this invention, in order to increase the collecting effect of the balloon in the aerial collector conductor or in the earth wire, radiating collectors are used. These consist of either incandescent metal or oxide electrodes in the form of vacuum grid valves, or electric arcs (mercury or similar electrodes), Nernst lamps, or flames of various kinds maybe simply connected with the respective conductor.

It is well known that energy can be drawn off from a cathode consisting of an incandescent body opposite an anode charged with positive electricity (vacuum grid tube). Hitherto however, a cathode was always first directly placed opposite an anode, and secondly, the system always consisted of a closed circuit.

Now if we dispense with the ordinary ideas in forming light or flame arcs in which a cathode must always stand directly opposite an anode charged to a high voltage or another body freely floating in the air, or consider the incandescent cathode to be only a source of unipolar discharge, (which represents group and point discharges in electro-static machines similar to unipolar discharges), it may be ascertained that incandescent cathodes and less perfectly, all incandescent radiators, flames and the like, have relatively large current densities and allow large quantities of electric energy to radiate into open space in the form of electron streams as transmitters.

The object of this invention is as described below, if such incandescent oxide electrodes or other incandescent radiators or flames are not freely suspended in space but instead are connected metallically with the earth so that they can be charged with negative terrestrial electricity, these radiators possess the property of absorbing the free positive electrical charges contained in the air space surrounding them (that is to say, of collecting them and conducting them to earth). They can therefore serve as collectors and have in comparison to the action of the spikes, a very large radius of action R; the effective capacity of these collectors is much greater than the geometrical capacity (R_0) calculated in an electro-static sense.

As is well known, our earth is surrounded with an electro-static field and the difference of potential dV/dh of the earth field according to the latest investigations, is in summer about 60 to 100 volts, and in winter, 300 to 500 volts

per metre difference in height, a simple calculation gives the result that when such a radiation collector or flame collector is arranged, for example, on the ground, and a second one is mounted vertically over it at a distance of 2,000 metres and both are connected by a conducting cable, there is a voltage difference in summer of about 2,000,000 volts and in winter 6,000,000 volts or more.

According to Stefan Boltzmann's law of radiation, the quantity of energy which an incandescent surface (temperature T) of 1 sq. cm. radiates in a unit of time into the open air (temperature T_0) is expressed by the following formula:

$$S = R\,(T^4 - T_0^4) \text{ watts per square centimetre}$$

and the universal radiation constant R, according to the latest researches of Ferry, is equal to 6.30×10^{-12} watts per square centimetre.

Now, if an incandescent surface of 1 sq. cm., as compared to the surrounding space, shows a periodic fall of potential dV, it radiates (independent of the direction of the current) in accordance with the above formula, for example at a temperature of 3715^0 C. an energy of 1.6 kW per square centimetre. As for the radiation, the same value can be calculated for the collection of energy, but reversed. Now, as carbon electrodes at the temperature of the electric arc, support a current density up to 60 to 65 amps per sq. cm., no difficulties will result in this direction in employing radiating collectors as accumulators.

If the earth be regarded as a cosmically insulated capacitor in the sense of geometrical electro-statics *x*, according to Chwolson, there results from the geometric capacity of the earth:

For negative charging 1.3×10^6 Coulomb For negative potential $V = 10 \times 10^8$ volts.

It follows from this that EJT is approximately equal to 24.7×10^{24} watts/sec. Now if it is desired to make a theoretical short circuit through an earthed flame collector, this would represent an electrical total work of about $79{,}500 \times 10^{10}$ kilowatt years. As the earth must be regarded as a rotating mechanism which is thermo-dynamically, electromagnetically and kinematically coupled with the sun and star system by cosmic radiation and gravitation, a reduction in the electric energy of the earth field is not to be feared. The energies which the incandescent collectors could withdraw from the earth field can only cause a lowering of the earth temperature. This is however, not the case as the earth does not represent a cosmically entirely insulated system. On the contrary, there is conveyed from the sun to the earth an energy of $18{,}500 \times 10^{10}$ kilowatts. Accordingly, any lowering of the earth temperature without a simultaneous lowering of the sun's temperature would contradict Stefan Boltzmann's law of radiation.

From this it must be concluded that if the earth temperature sinks, the total radiation absorbed by the earth increases, and further, the rate of cooling of the earth is directly dependent on that of the sun and the other radiators cosmically coupled with the sun.

The incandescent radiation collectors may, according to this invention, be used for collecting atmospheric electricity if they (1) are charged with the negative earth electricity (that is to say, when they are directly connected to the earth by means of a metallic conductor) and (2) if large capacities (metal surfaces) charged with electricity are mounted opposite them as positive poles in the air. This is regarded as the main feature of the present invention as without these inventive ideas it would not be possible to collect with an incandescent collector, sufficiently large quantities of the electrical charges contained in the atmosphere as technology requires; the radius of action of the flame collectors would also be too small, especially if it be considered that the very small surface density does not allow of large quantities of charge being absorbed from the atmosphere.

It has already been proposed to employ flame collectors for collecting atmospheric electricity and it is known that their collecting effect is substantially greater opposite the points. It is however, not known that the quantities of current which hitherto be obtained are too small for technical purposes. According to my experiments, the reason for this is to be found in the inadequate capacities of the collector conductor poles. If such flame or radiating collectors have no or only small positive surfaces, their radius of action for large technical purposes is too small. If the incandescent collectors be constantly kept in movement in the air, they may collect more according to the speed of the movement, but this is again not capable of being carried out in practice.

By this invention, the collector effect is considerably increased by a body charged with a positive potential and of the best possible capacity, being also held floating (without direct earth connection) opposite such an incandescent collector which is held floating in the air at a desired height. If, for example, a collecting balloon of sheet metal or metallised fabric, be caused to mount to 300 to 3,000 metres in the air, and as a positive pole it is brought opposite such a radiating collector connected by a conductor to earth, quite different results are obtained.

The metallic balloon shell which has a large surface area is charged to a high potential by atmospheric electricity. This potential is greater the higher the collecting balloon is above the incandescent collector. The positive electricity acts concentratedly on the anode floating in the air as it is attracted through the radiation shock ionisation, proceeding from the incandescent cathode. The consequence of this is that the radius of action of the incandescent cathode collector is considerably increased and so is the collecting effect of the balloon surface. Further, the large capacity of the anode floating in the air, plays therefore an important part because it allows the collection of large charges resulting in a more uniform current even when there is substantial current withdrawal - this cannot be the case with small surfaces.

In the present case, the metallic collecting balloon is a positive anode floating in the air and the end of the earth conductor of this balloon serves as positive pole surface opposite the surface of the radiating incandescent cathode, which in turn is charged with negative earth electricity as it is connected to the earth by a conductor. The process may be carried out by two such contacts (negative incandescent cathode and anode end of a capacity floating in the air) a capacitor and an inductive resistance being switched on in parallel, whereby simultaneously undamped oscillations may be formed.

In very large installations it is advisable to connect two such radiating collectors in series. Thus an arc light incandescent cathode may be placed below on the open ground and an incandescent cathode which is heated by special electro-magnetic currents, be located high in the air. Of course for this, the special vacuum Liebig tubes with or without grids may also be used. An ordinary arc lamp with oxide electrodes may be introduced on the ground and the positive pole is not directly connected with the collecting balloon, but through the upper incandescent cathode or over a capacitor. The method of connecting the incandescent cathode floating in the air may be seen in **Figs.29-33**.

B is the air balloon, **K** a Cardan ring (connection with the hawser) **C** the balloon, **L** a good conducting cable, **P** a positive pole, **N** negative incandescent cathode and **E** the earth conductor.

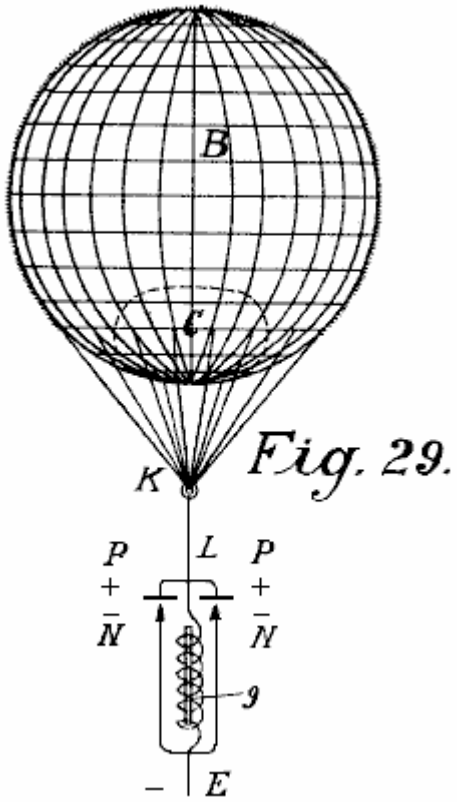


Fig. 29.

Fig.29 represents the simplest form of construction. If electric oscillations are produced below on the ground by means of a carbon arc lamp or in any other suitable way, a considerably greater electric resistance is opposed to that in the direct way by inserting an electrical inductive resistance **9**. Consequently, between **P** and **N**, a voltage is formed, and as, over **N** and **P** only an inductionless ohmic resistance is present, a spark will spring over so long as the separate induction coefficients and the like are correctly calculated. The consequence of this is that the oxide electrode (carbon or the like) is rendered incandescent and then shows as incandescent cathode, an increased collecting effect. The positive poles must be substantially larger than the negative in order that they may not also become incandescent. As they are further connected with the large balloon area which has a large capacity and is charged at high voltage, an incandescent body which is held floating in the air and a positive pole which can collect large capacities is thereby obtained in the simplest way. The incandescent cathode is first

caused to become incandescent by means of separate energy produced on the earth, and then maintained by the energy collected from the atmosphere.

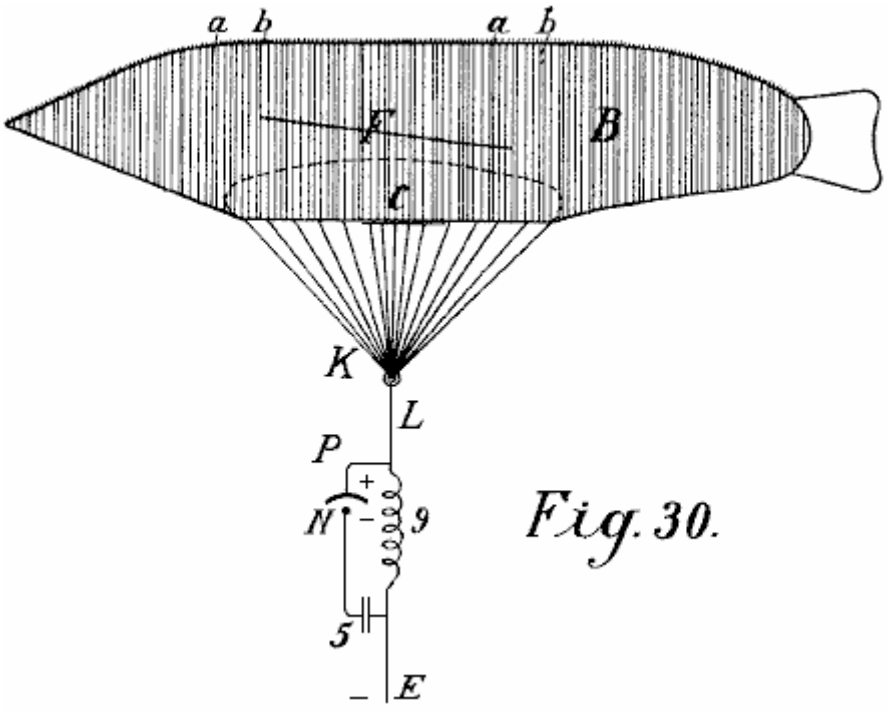


Fig. 30.

Fig.30 only shows the difference that instead of a round balloon, a cigar-shaped one may be used, also, a capacitor **5** is inserted between the incandescent cathode and the earth conductor so that a short-circuited oscillation circuit over **P N 5** and **9** is obtained. This has the advantage that quite small quantities of electricity cause the cathode to become incandescent and much larger cathode bodies may be made incandescent.

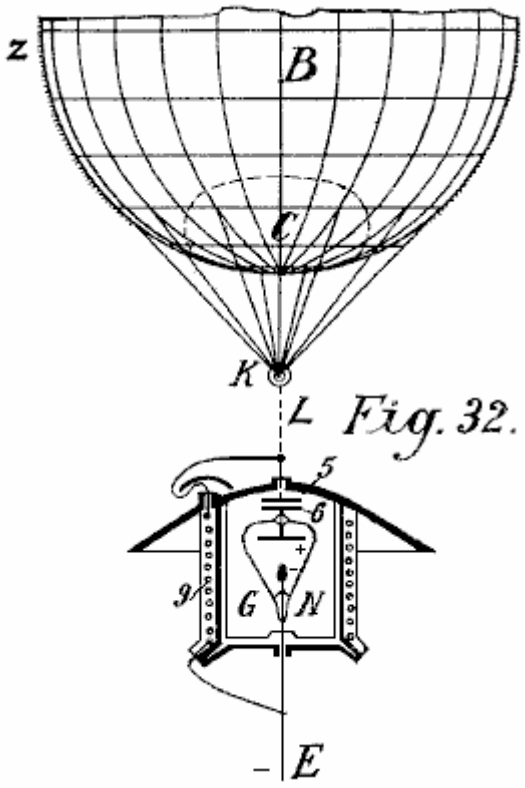


Fig. 32.

In this form of construction, both the incandescent cathode and the positive electrode may be enclosed in a vacuum chamber as shown in **Fig.32**. A cable **L** is carried well insulated through the cover of a vessel and ends in a capacitor disc **5**. The cover is arched in order to keep the rain off. The vessel is entirely or partially made of magnetic metal and well insulated inside and outside. Opposite disc **5** another disc **6** and on this again a metallic positive pole of the vacuum tube **g** with the incandescent cathode (oxide electrode) **N** is arranged. The negative electrode is on the one hand connected to the earth conductor **E**, and on the other hand with the inductive resistance **9** which is also connected with the cable **L** with the positive pole and wound around the vessel in coils.

The action is exactly the same as that in **Fig.29** only instead of an open incandescent cathode, one enclosed in vacuo is used. As in such collectors, only small bodies be brought to incandescence, in large installations a plurality of such vacuum tubes must be inserted in proximity to one another. According to the previous constructions **Fig.31** and **Fig.33** are quite self evident without further explanations.

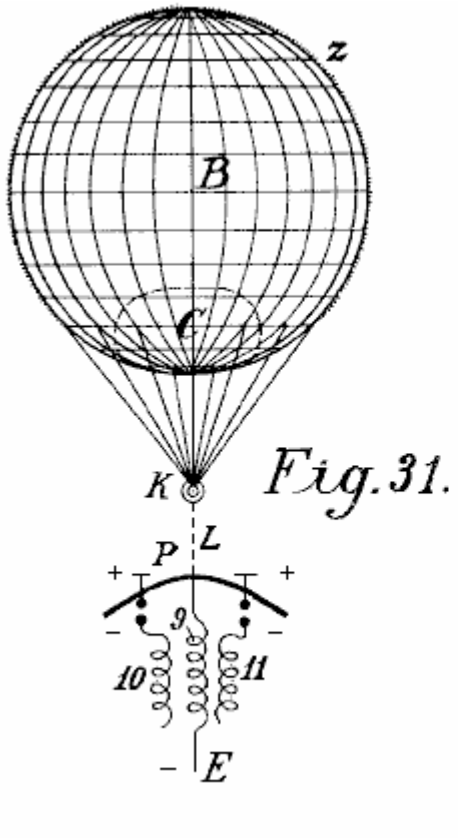


Fig. 31.

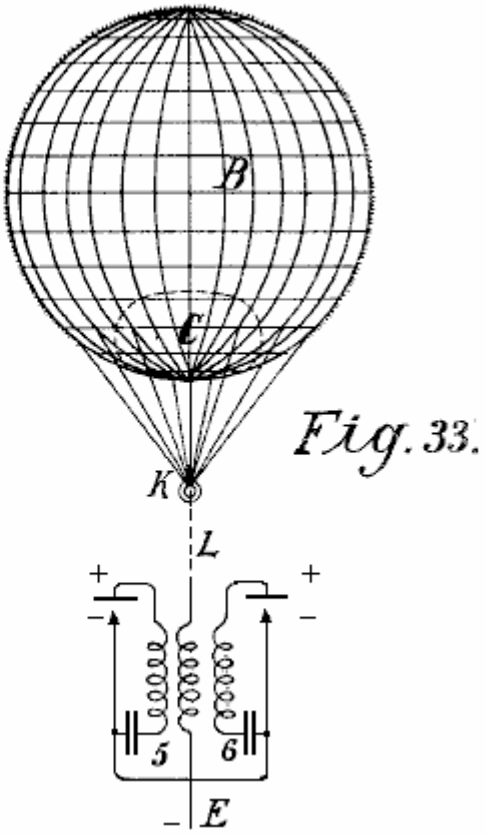


Fig. 33.

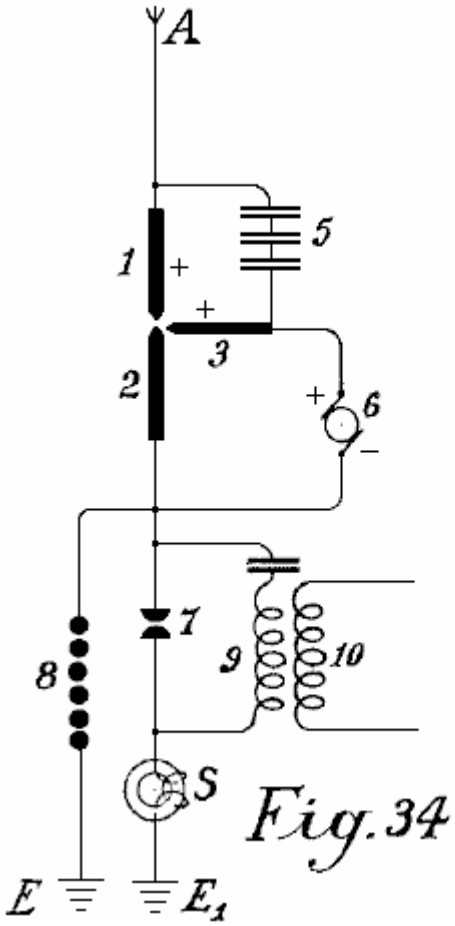


Fig. 34

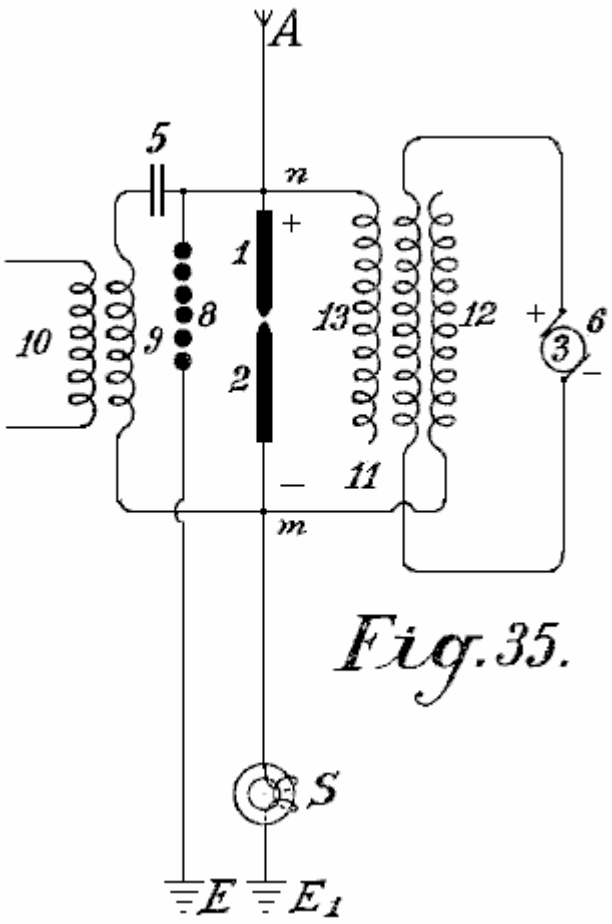


Fig. 35.

Figs.34-37 represent further diagrams of connections over radiating and flame collectors, and in fact, how they are to be arranged on the ground. **Fig.34** shows an arc light collector with oxide electrodes for direct current and its connection. **Fig.35** shows a similar one for alternating current. **Fig.36** an incandescent collector with a Nernst lamp and **Fig.37** a similar one with a gas flame.

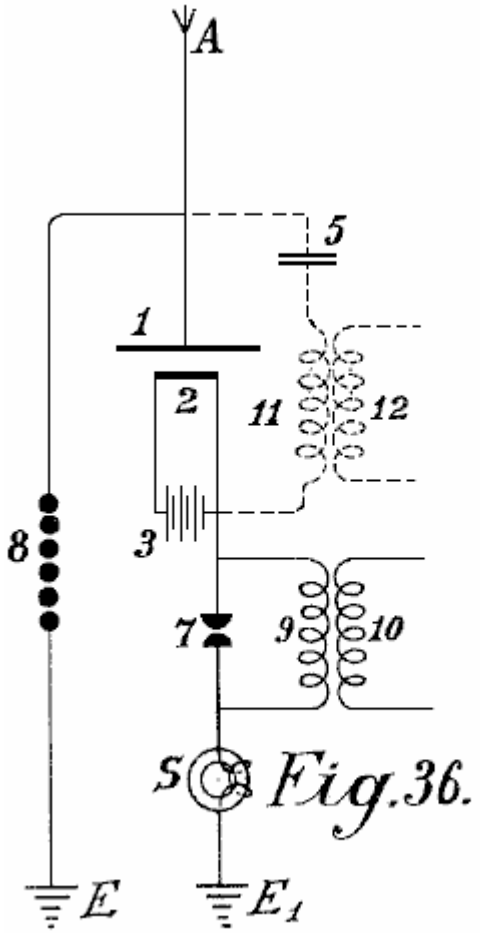


Fig.36.

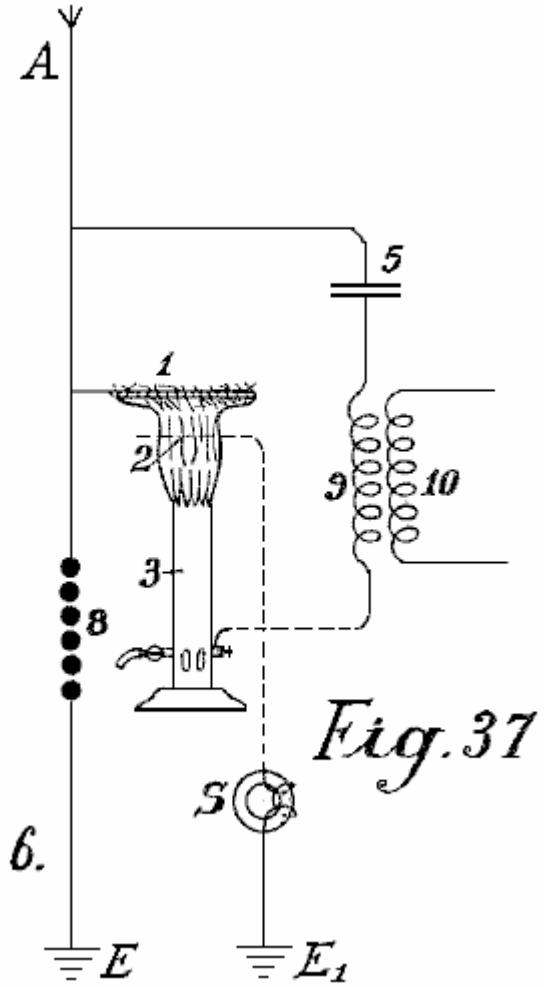


Fig.37

The positive pole **1** of the radiating collectors is always directly connected to the aerial collecting conductor **A**. In **Fig.34**, this is further connected over the capacitor set **5** with a second positive electrode **3**. The direct current dynamo **b** produces current which flows over between the electrodes **3** and **2** as an arc light. On the formation of an arc, the negative incandescent electrode **2** absorbs electricity from the positive poles standing opposite it and highly charged with atmospheric electricity which it conveys to the working circuit. The spark gap 7, inductive resistance **9** and induction coil **10** are like the ones previously described. The protective electromagnet **S** protects the installation from earth circuiting and the safety spark gap **8** from excess voltage or overcharging.

In **Fig.35**, the connection is so far altered that the alternating current dynamo feeds the excitation coil **11** of the induction capacitor. **12** is its negative and **13** its positive pole. If the coil **3** on the magnet core of the dynamo is correctly calculated and the frequency of the alternating current sufficiently high, then an arc light can be formed between poles **1** and **2**. As the cathode **2** is connected to the negatively charged earth, and therefore always acts as a negative pole, a form of rectification of the alternating current produced by the dynamo **3** is obtained, since the second half of the period is always suppressed. The working circuit may be carried out in the same way as in **Fig.34**; the working spark gap **7** may however be dispensed with, and instead of it, between the points **n** and **m**, a capacitor **5** and an induction resistance **9** may be inserted, from which, a current is taken inductively.

Fig.36 represents a form of construction similar to that shown in **Fig.34** except that here instead of an arc lamp, a Nernst incandescent body is used. The Nernst lamp is fed through the battery **3**. The working section is connected with the negative pole, the safety spark gap with the positive poles. The working spark gap **7** may also be dispensed with and the current for it taken at **12** over the oscillation circuit **5, 11** (shown in dotted lines).

Flame collectors (**Fig.37**) may also be employed according to this invention. The wire network **1** is connected with the aerial collector conductor **A** and the burner with the earth. At the upper end of the burner, long points are provided which project into the flame. The positive electrode is connected with the negative over a capacitor **5** and the induction coil **9** with the earth.

The novelty in this invention is:

(1) The use of incandescent cathodes opposite positive poles which are connected to large metallic capacities as automatic collecting surfaces.
(2) The connection of the incandescent cathodes to the earth whereby, in addition to the electricity conveyed to them from the battery of machine which causes the incandescing, also the negative charge of the earth potential is conveyed, and
(3) The connection of the positive and negative poles of the radiating collectors over a capacitor circuit alone or with the introduction of a suitable inductive resistance, whereby simultaneously an oscillatory oscillation circuit may be obtained. The collecting effect is by these methods quite considerably increased.

ROY MEYERS

Patent GB1913,01098 14th January 1914 Inventor: Roy J. Meyers

APPARATUS FOR PRODUCING ELECTRICITY

ABSTRACT

A rectifier for use with apparatus for producing electricity from the earth consists of mercury- vapour lamps constructed and arranged as shown in **Fig.4**. Each lamp comprises two wires **6<1>, 7<1>** wound around a steel tube **15** surrounding a mercury tube **11** preferably of copper. The coil **6<1>** is connected between the electrode **14** and the terminal **18**, and the coil **7<1>** between the terminals **19, 5**. The coils **6<1>, 7<1>** are preferably composed of soft iron.

DESCRIPTION

This invention relates to improvements in apparatus for the production of electrical currents, and the primary object in view is the production of a commercially serviceable electrical current without the employment of mechanical or chemical action. To this end the invention comprises means for producing what I believe to be dynamic electricity from the earth and its ambient elements.

I am, of course aware that it has been proposed to obtain static charges from upper strata of the atmosphere, but such charges are recognised as of widely variant potential and have thus far proved of no practical commercial value, and the present invention is distinguished from all such apparatus as has heretofore been employed for attracting static charges by the fact that this improved apparatus is not designed or employed to produce or generate irregular, fluctuating or other electrical charges which lack constancy, but on the other hand I have by actual test been able to produce from a very small apparatus at comparatively low elevation, say about 50 or 60 feet above the earth's surface, a substantially constant current at a commercially usable voltage and amperage.

This current I ascertained by repeated tests is capable of being readily increased by additions of the unit elements in the apparatus described below, and I am convinced from the constancy of the current obtained and its comparatively low potential that the current is dynamic and not static, although, of course, it is not impossible that certain static discharges occur and, in fact, I have found occasion to provide against the damage which might result from such discharge by the provision of lightning arresters and cut-out apparatus which assist in rendering the obtained current stable by eliminating sudden fluctuations which sometimes occur during conditions of high humidity from what I consider static discharges.

The nature of my invention is obviously such that I have been unable to establish authoritatively all of the principles involved, and some of the theories herein expressed may possibly prove erroneous, but I do know and am able to demonstrate that the apparatus which I have discovered does produce, generate, or otherwise acquire a difference of potential representing a current amperage as stated above.

The invention comprises the means for producing electrical currents of serviceable potential substantially without the employment of mechanical or chemical action, and in this connection I have been able to observe no chemical action whatever on the parts utilised although deterioration may possibly occur in some of the parts, but so far as I am able to determine such deterioration does not add to the current supply but is merely incidental to the effect of climatic action.

The invention more specifically comprises the employment of a magnet or magnets and a co-operating element, such as zinc positioned adjacent to the magnet or magnets and connected in such manner and arranged relative to the earth so as to produce current, my observation being that current is produced only when such magnets have their poles facing substantially to the north and south and the zincs are disposed substantially along the magnets.

The invention also comprehends other details of construction, combinations and arrangements of parts as will be fully set forth.

DESCRIPTION OF THE DRAWINGS

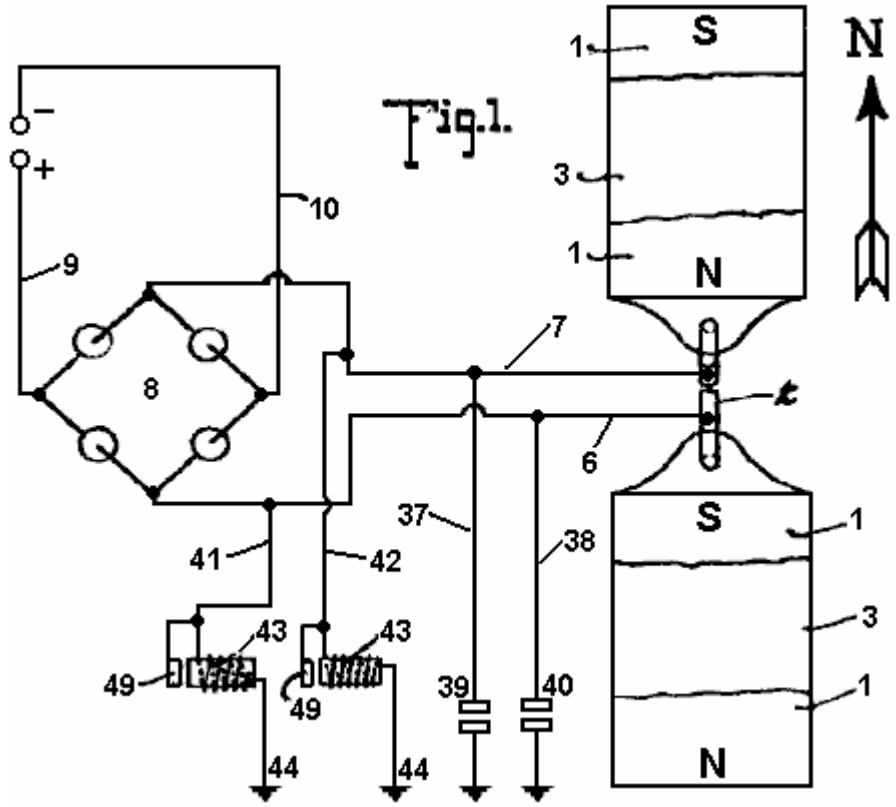


Fig.1 is a plan view of an apparatus embodying the features of the present invention, the arrow accompanying the figure indicating substantially the geographical north, parts of this figure are diagrammatic.

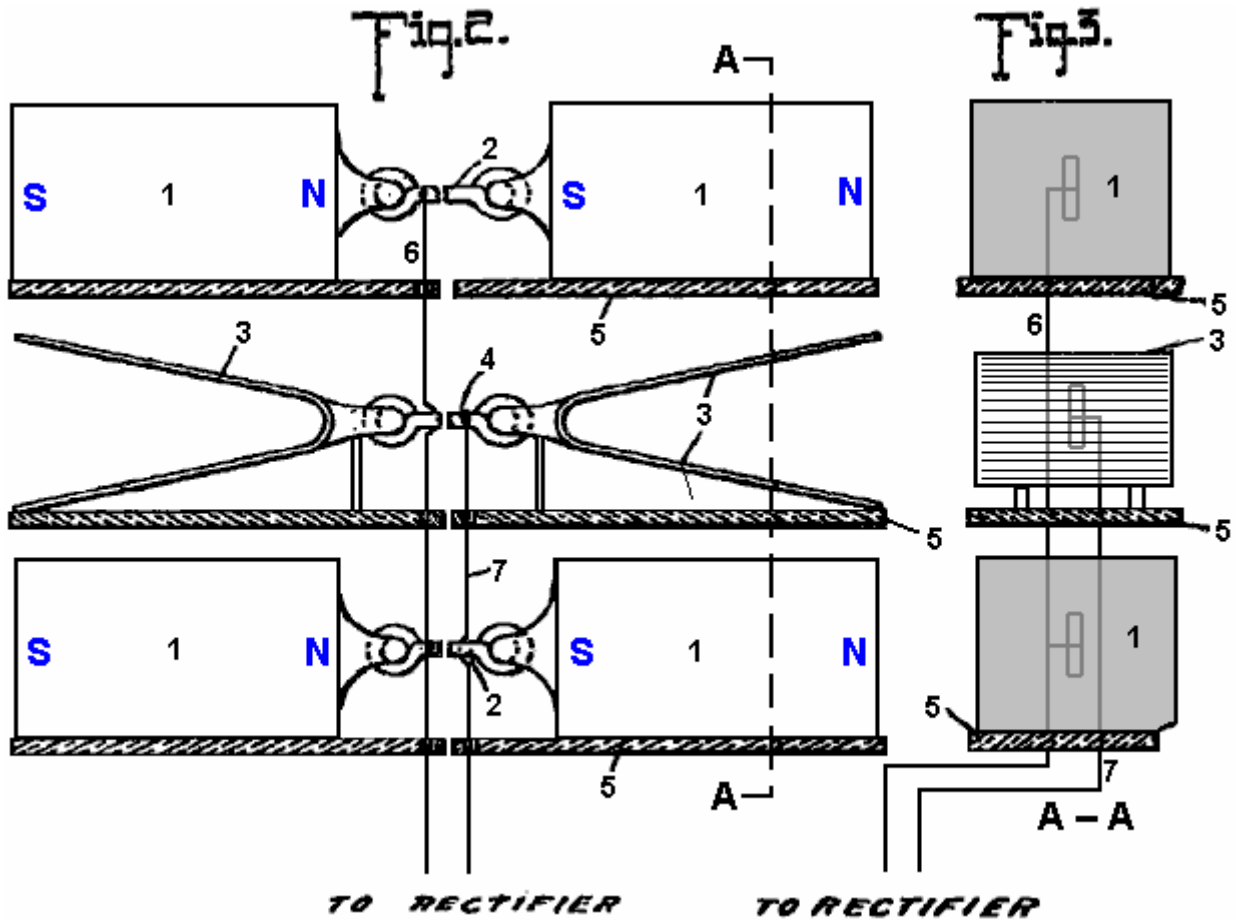


Fig.2 is a view is side elevation of the parts seen in plan in **Fig.1**
Fig.3 is a vertical section taken on the plane indicated by the line **A--A** of **Fig.2**.

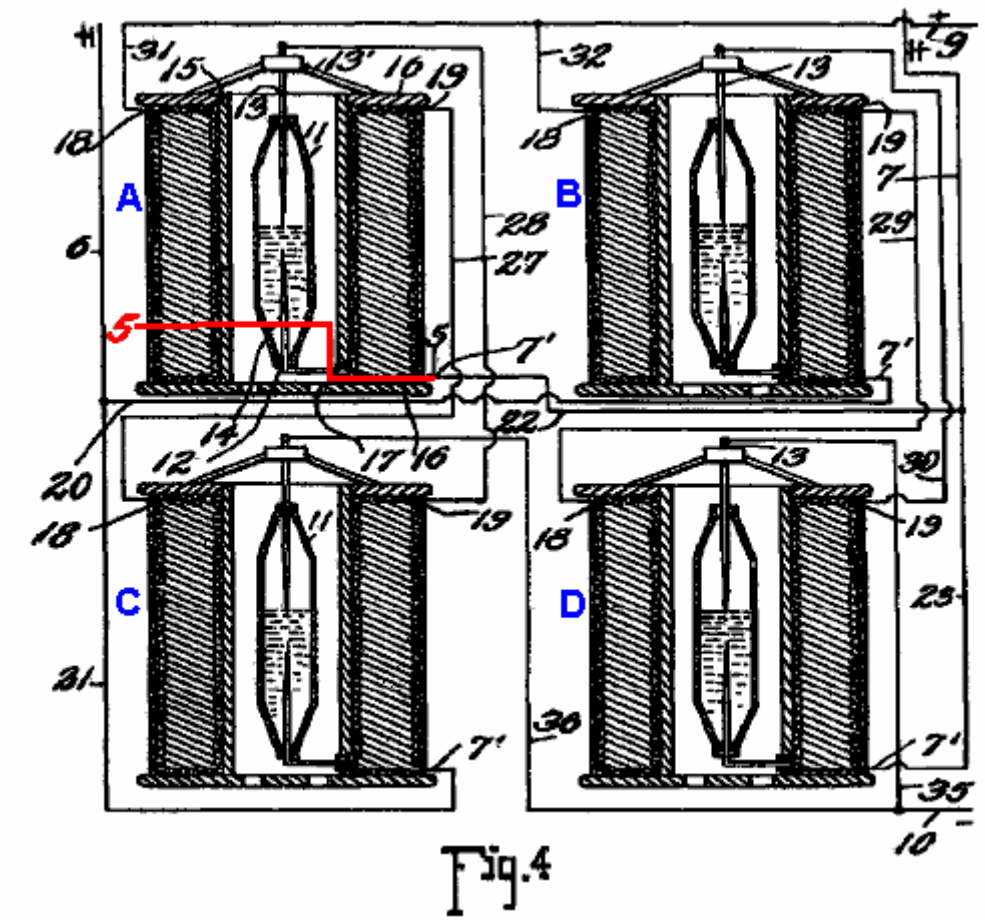


Fig.4 is a detail view, partly in elevation and partly in section, showing the connections of the converter and intensifier.

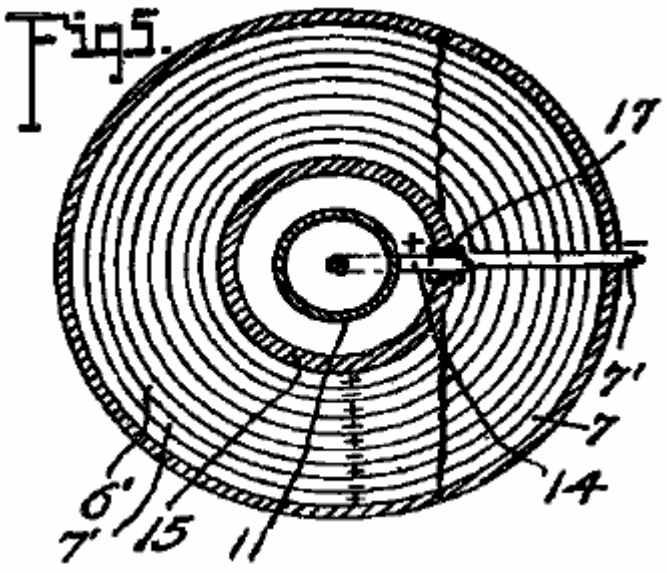


Fig.5 is a transverse section taken on the planes indicated by line **5-5** of **Fig.4**, looking downwards.

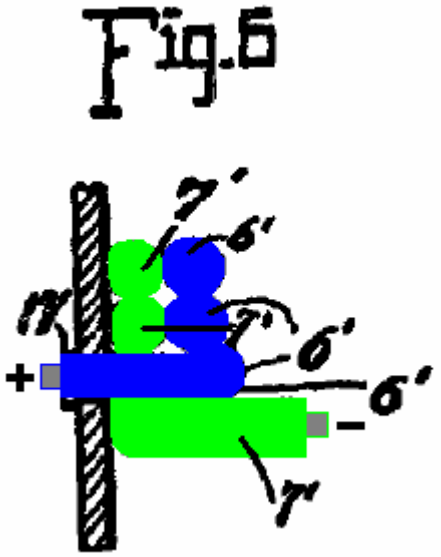


Fig.6 is an enlarged detail fragmentary section illustrating the parts at the junction of the conductors and one of the intensifiers.

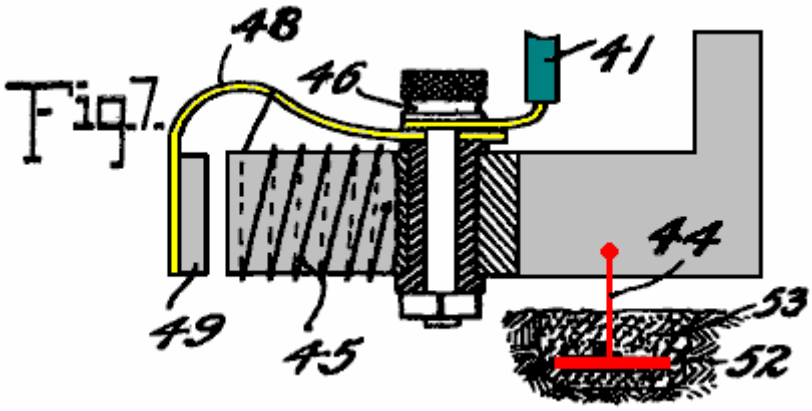


Fig.7 is an enlarged detail view partly in elevation and partly in section of one of the automatic cut-outs

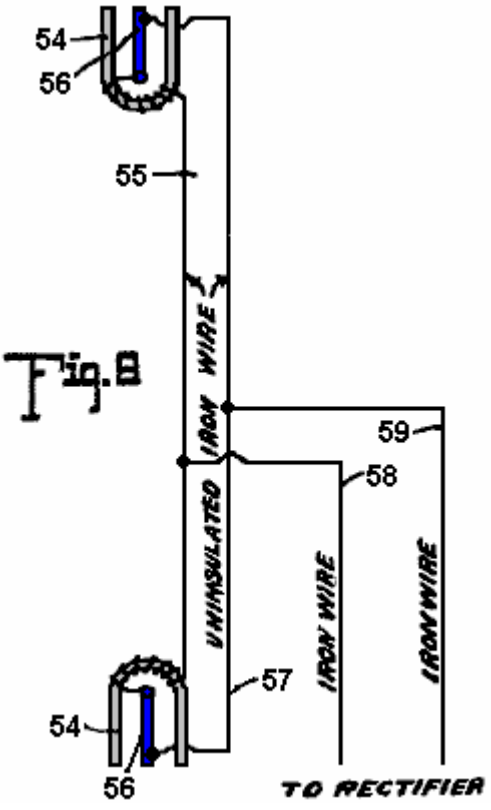
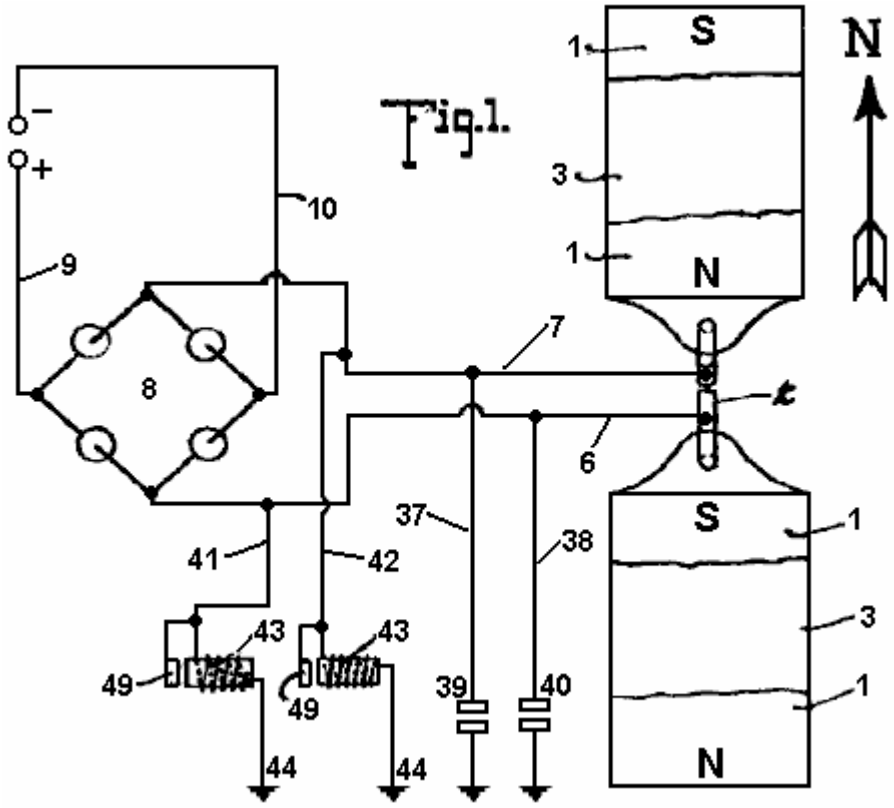


Fig.8 is a diagrammatic view of one of the simplest forms of embodiment of the invention.

Referring to the drawing by numerals, **1,1** indicates magnets connected by a magnetic substance **2**, preferably an iron wire. The magnets **1** are arranged in pairs, one pair being spaced beneath the other, and interposed between the magnets are zinc plates **3,3** connected by an iron wire conductor **4**. Suitable insulating supports **5** are arranged for sustaining the respective magnets **1** and plates **3,3**. Each plate **3** is preferably bent substantially into V form, as clearly seen in **Fig.1**, and the V of one of the plates opens or faces toward the North and the V of the other plate to the South. I have determined by experimentation that it is essential that the plates **3** be disposed substantially North and South with their flat faces approximately parallel to the adjacent faces of the co-operating magnets, although by experience I have not discovered any material difference in the current obtained when the plates are disposed slightly to one side of North and South, as for instance when the plates are disposed slightly to one side of North and South, as for instance when disposed in the line of the magnetic polarity of the earth. The same is true with respect to the magnets **1**, the said magnets being disposed substantially North and South for operative purposes, although I find that it is immaterial whether the North pole of one of the magnets is disposed to the North and the South pole to the South, or vice versa, and it is my conviction from experience that it is essential to have the magnets of each pair connected by magnetic material so that the magnets substantially become one with a pole exposed to the North and a pole exposed to the South.



In **Fig.1**, I have indicated in full lines by the letters **8** and **N** the respective polarities of the magnets **1**, and have indicated in dotted lines the other pole of those magnets when the connection **2** is severed. I have found that the magnets and zinc plates operate to produce, (whether by collection or generation I am not certain), electrical currents when disposed substantially North and South, but when disposed substantially East and West, no such currents are produced. I also find that the question of elevation is by no means vital, but it is true that more efficient results are obtained by placing the zincs and magnets on elevated supports. I furthermore find from tests, that it is possible to obtain currents from the apparatus with the zincs and magnets disposed in a building or otherwise enclosed, although more efficient results are obtained by having them located in the open.

While in **Figures 1, 2,** and **3**, I have shown the magnets and the zinc plates as superimposed, it will be apparent, as described in detail below, that these elements may be repositioned in horizontal planes, and substantially the same results will be secured. Furthermore, the magnets **1** with the interposed zincs **3**, as shown in **Figures 1, 2** and **3** merely represent a unit which may be repeated either horizontally or vertically for increasing the current supply, and when the unit is repeated the zinc plates are arranged alternating with the magnets throughout the entire series as indicated below.

A conductor **6** is connected in multiple with the conductors **2** and a conductor **7** is connected with conductor **4**, the conductor **6** extending to one terminal of a rectifier which I have indicated by the general reference character **8**, and the conductor **7** extending to the other terminal of the rectifier. The rectifier as seen in the diagram **Fig.1** may

assume any of several well known embodiments of the electrical valve type and may consist of four asymmetric cells or Cooper-Hewitt mercury vapour lamps connected as indicated in **Fig.1** for permitting communication of the positive impulses from the conductor **6** only to the line conductor **9** and the negative impulses from conductor **6** on only to the line conductor **10**. The current from this rectifier may be delivered through the conductors **9** and **10** to any suitable source for consumption.

While the said rectifier **8** may consist of any of the known types, as above outlined, it preferably consists of a specially constructed rectifier which also has the capacity of intensifying the current and comprises specifically the elements shown in detail in **Figures 4, 5,** and **6** wherein I have disclosed the detail wiring of the rectifier when composed of four of the rectifying and intensify in elements instead of asymmetric cells or simple mercury vapour valves. As each of these structures is an exact embodiment of all the others, one only will be described, and the description will apply to all. The rectifying element of each construction consists of a mercury tube **11** which is preferably formed of glass or other suitable material, and comprises a cylinder having its end portions tapered and each terminating in an insulating plug or stopper **12**. Through the upper stopper **12** is extended the electrode **13** which extends well into the tube and preferably about one-half its length, to a point adjacent the inner end of an opposing electrode **14** which latter electrode extends from there down through the insulation **12** at the lower end of the tube. The tube **11** is supplied with mercury and is adapted to operate on the principle of the mercury vapour lamp, serving to rectify current by checking back impulses of one sign and permitting passage of impulses of the other.

To avoid the necessity for utilising a starter, as is common with the lamp type of electrical valve, the supply of mercury within the tube may be sufficient to contact with the lower end of the electrode **13** when current is not being supplied, so that as soon as current is passed from one electrode to the other sufficiently for volatilising that portion of the mercury immediately adjacent the lower end of electrode **13**, the structure begins its operation as a rectifier. The tube **11** is surrounded by a tube **15** which is preferably spaced from tube **11** sufficiently for allowing atmospheric or other cooling circulation to pass the tube **11**. In some instances, it may be desirable to cool the tube **11** by a surrounding body of liquid, as mentioned below. The tube **15** may be of insulating material but I find efficient results attained by the employment of a steel tube, and fixed to the ends of the of the tube are insulating disks **16**, **16** forming a spool on which are wound twin wires **6'** and **7'**, the wire **6'** being connected at the inner helix of the coil with the outer end of the electrode **14**, the lower portion of said electrode being extended to one side of the tube **11** and passed through an insulating sleeve **17** extending through the tube **15**, and at its outer end merging into the adjacent end of the wire **6'**. The wire **7'** extends directly from the outer portion of the spool through the several helices to a point adjacent to the junction of the electrode **14** with wire **6'** and thence continues parallel to the wire throughout the coil, the wire **6'** ending in a terminal **18** and the wire **7'** ending in a terminal **19**.

For the sake of convenience of description and of tracing the circuits, each of the apparatus just above described and herein known as an intensifier and rectifier will be mentioned as **A**, **B**, **C** and **D**, respectively. Conductor **6** is formed with branches **20** and **21** and conductor **7** is formed with similar branches **22** and **23**. Branch **20** from conductor **6** connects with conductor **7'** of intensifier **B** and branch **21** of conductor **6** connects with the conductor **7'** of intensifier **C**, while branch **22** of conductor **7** of intensifier **C**, while branch **22** of conductor **7** connects with conductor **7'** of intensifier **D**. A conductor **27** is connected to terminal **19** of intensifier **A** and extends to and is connected with the terminal **18** of intensifier **C**, and a conductor **7** connects with conductor **7'** of intensifier **D**. A conductor **27** is connected to terminal **19** of intensifier **A**, and extends to and is connected to terminal **18** of intensifier **C**, and a conductor **28** is connected to the terminal **19** of intensifier **C** and extends from the terminal **19** of intensifier **B** to the terminal **18** of intensifier **D** to electrode **13** of intensifier **B**. Each electrode **13** is supported on a spider **13'** resting on the upper disk **16** of the respective intensifier. Conductors **31** and **32** are connected to the terminals **18** of intensifiers **A** and **B** and are united to form the positive line wire **9** which co-operates with the negative line wire **10** and extends to any suitable point of consumption. The line wire **10** is provided with branches **35** and **36** extending to the electrodes **13** of intensifiers **C** and **D** to complete the negative side of the circuit.

Thus it will be seen that alternating currents produced in the wires **6** and **7** will be rectified and delivered in the form of a direct current through the line wires **9** and **10**, and I find by experiment that the wires **6** and **7** should be of iron, preferably soft, and may of course be insulated, the other wiring not specified as iron being of copper or other suitable material.

In carrying out the operation as stated, the circuits may be traced as follows: A positive impulse starting at the zincs **3** is directed along conductor **7** to branch **23** to conductor **7'** and the winding of the rectifier of intensifier **B** through the rectifier to the conductor **6'**, through its winding to the contact **18**, conductor **32** and to the line wire **9**. The next, or negative, impulse directed along conductor **7** cannot find its way along branch **23** and the circuit just above traced because it cannot pass across the rectifier of intensifier **B** but instead the negative impulse passes along conductor **22** to conductor **7** of intensifier **A** and its winding to the contact **19** and to conductor **27** to contact **18** of intensifier **C**, to the winding of the wire **6'** thereof to the electrode **14** through the rectifier to the of the

electrode **13** and conductor of intensifier **A**, electrode **14** thereof and conductor **6'** to contact **18** and wire **31** to line wire **9**.

Obviously the positive impulse cannot pass along the wire **20** because of its inverse approach to the rectifier of intensifier **B**. The next impulse or negative impulse delivered to conductor **6** cannot pass along conductor **21** because of its connection with electrode **13** of the rectifier of intensifier **A**, but instead passes along conductor **20** to the wire **7'** and its winding forming part of intensifier **B** to the contact **19** and conductor **29** to contact **18** and the winding of wire **6'** of intensifier **D** to the electrode **14** and through the rectifier to the electrode **13** and conductor **35** to line wire **10**. Thus the current is rectified and all positive impulses directed along one line and all negative impulses along the other lie s that the potential difference between the two lines will be maximum for the given current of the alternating circuit. It is, of course, apparent that a less number of intensifiers with their accompanying rectifier elements may be employed with a sacrifice of the impulses which are checked back from a lack of ability to pass the respective rectifier elements, and in fact I have secured efficient results by the use of a single intensifier with its rectifier elements, as shown below.

Grounding conductors **37** and **38** are connected respectively with the conductors **6** and **7** and are provided with the ordinary lightning arresters **39** and **40** respectively for protecting the circuit against high tension static charges.

Conductors **41** and **42** are connected respectively with the conductors **6** and **7** and each connects with an automatic cut-out **43** which is grounded as at **4**. Each of the automatic cut-outs is exactly like the other and one of the these is shown in detail in **Fig.7** and comprises the inductive resistance **45** provided with an insulated binding post **46** with which the respective conductor **6** or **7** is connected, the post also supporting a spring **48** which sustains an armature **49** adjacent to the core of the resistance **45**. The helix of resistance **45** is connected preferably through the spring to the binding post at one end and at the other end is grounded on the core of the resistance, the core being grounded by ground conductor **44** which extends to the metallic plate **52** embedded in moist carbon or other inductive material buried in the earth. Each of the conductors **41**, **42** and **44** is of iron, and in this connection I wish it understood that where I state the specific substance I am able to verify the accuracy of the statement by the results of tests which I have made, but of course I wish to include along with such substances all equivalents, as for instance, where iron is mentioned its by-products, such as steel, and its equivalents such as nickel and other magnetic substances are intended to be understood.

The cut-out apparatus seen in detail in **Fig.7** is employed particularly for insuring against high voltage currents, it being obvious from the structure shown that when potential rises beyond the limit established by the tension of the spring sustaining the armature **40**, the armature will be moved to a position contacting with the core of the cut-out device and thereby directly close the ground connection for line wire **41** with conductor **44**, eliminating the resistance of winding **45** and allowing the high voltage current to be discharged to the ground. Immediately upon such discharge the winding **45** losing its current will allow the core to become demagnetised and release the armature **49** whereby the ground connection is substantially broken leaving only the connection through the winding **45** the resistance of which is sufficient for insuring against loss of low voltage current.

In **Fig.8** I have illustrated an apparatus which though apparently primitive in construction and arrangement shows the first successful embodiment which I produced in the course of discovery of the present invention, and it will be observed that the essential features of the invention are shown there. The structure shown in the figure consists of horseshoe magnets **54**, **55**, one facing North and the other South, that is, each opening in the respective directions indicated and the two being connected by an iron wire **55** which is uninsulated and wrapped about the respective magnets each end portion of the wire **55** being extended from the respective magnets to and connected with, as by being soldered to, a zinc plate **56**, there being a plate **56** for each magnet and each plate being arranged longitudinally substantially parallel with the legs of the magnet and with the faces of the plate exposed toward the respective legs of the magnet, the plate being thus arranged endwise toward the North and South. An iron wire **57** connects the plates **56**, the ends of the wire being preferably connected adjacent the outer ends of the plates but from experiment I find that the wire may be connected at practically any point to the plate. Wires **58** and **59** are connected respectively with the wires **55** and **57** and supply an alternating current at a comparatively low voltage, and to control such current the wires **58** and **59** may be extended to a rectifier or combined rectifier and intensifier, as discussed above.

The tests which I have found successful with the apparatus seen in **Fig.8** were carried out by the employment first of horseshoe magnets approximately 4 inches in length, the bar comprising the horseshoe being about one inch square, the zincs being dimensioned proportionately and from this apparatus with the employment of a single intensifier and rectifier, as above stated, I was able to obtain a constant output of 8 volts.

It should be obvious that the magnets forming one of the electrodes of this apparatus may be permanent or may be electromagnets, or a combination of the two.

While the magnets mentioned throughout the above may be formed of any magnetic substance, I find the best results obtained by the employment of the nickel chrome steel.

While the successful operation of the various devices which I have constructed embodying the present invention have not enabled me to arrive definitely and positively at fixed conclusion relative to the principles and theories of operation and the source from which current is supplied, I wish it to be understood that I consider myself as the first inventor of the general type described above, capable of producing commercially serviceable electricity, for which reason my claims hereinafter appended contemplate that I may utilise a wide range of equivalents so far as concerns details of construction suggested as preferably employed.

The current which I am able to obtain is dynamic in the sense that it is not static and its production is accomplished without chemical or mechanical action either incident to the actual chemical or mechanical motion or incident to changing caloric conditions so that the elimination of necessity for the use of chemical or mechanical action is to be considered as including the elimination of the necessity for the use of heat or varying degrees thereof.

PAULO and ALEXANDRA CORREA

Pat. Application US 2006/0082,334 20th April 2006 Inventors: Paulo & Alexandra Correa

ENERGY CONVERSION SYSTEMS

This patent application shows the details of devices which can produce ordinary electricity from Tesla longitudinal waves. If these claims are correct (and there does not appear to be the slightest reason for believing that they are not), then implementations of this patent application are capable of producing free electrical power and the importance of this information is enormous.

ABSTRACT

This invention relates to apparatus for the conversion of mass-free energy into electrical or kinetic energy, which uses in its preferred form a transmitter and a receiver both incorporating Tesla coils, the distal ends of whose secondary windings are co-resonant and connected to plates of a chamber, preferably evacuated or filled with water, such that energy radiated by the transmitter may be picked up by the receiver, the receiver preferably further including a pulsed plasma reactor driven by the receiver coil and a split phase motor driven by the reactor. Preferably the reactor operates in pulsed abnormal gas discharge mode, and the motor is an inertially dampened drag motor. The invention also extends to apparatus in which an otherwise driven plasma reactor operating in pulsed abnormal gas discharge mode in turn used to drive an inertially dampened drag motor.

DESCRIPTION

This is a continuation of application Ser. No. 09/907,823, filed Jul. 19, 2001.

FIELD OF THE INVENTION

This invention relates to systems for the conversion of energy, inter alia in the form of what we will refer to for convenience as Tesla waves (see below), to conventional electrical energy.

BACKGROUND OF THE INVENTION

Energy converters that are fed by local or environmental energy are usually explained by taking recourse to the notion that they convert zero point electromagnetic radiation (ZPE) to electric energy. The ZPE theories have gained a life of their own, as T. Kuhn has pointed out (in his "Black Body Theory and the Quantum"), after emerging from Planck's second theory, specifically from the term ½ hυ in the new formula for oscillator energy. In 1913, Einstein and Stern suggested that motional frequencies contributing to specific heat fell into two categories--those that were independent of temperature and those that were not (e.g. rotational energy), leading them to conclude that zero-point energy on the order of ½ hυ was most likely. In the second part of their paper, however, they provided a derivation of Planck's Law without taking recourse to discontinuity, by assuming that the value of the ZPE was simply ha. It is worth noting that Einstein had already in 1905 ("Erzeugung und Verwandlung des Lichtes betreffenden heuristichen Gesichtspunkt",Ann. d. Phys, 17, 132) framed the problem of discontinuity, even if only heuristically, as one of placing limits upon the infinite energy of the vacuum state raised by the Rayleigh-Jeans dispersion law. According to Einstein, the Rayleigh-Jeans law would result in an impossibility, the existence of infinite energy in the radiation field, and this was precisely incompatible with Planck's discovery - which suggested instead, that at high frequencies the entropy of waves was replaced by the entropy of particles. Einstein, therefore, could only hope for a stochastic validation of Maxwell's equations at high frequencies "by supposing that electromagnetic theory yields correct time-average values of field quantities", and went on to assert that the vibration-energy of high frequency resonators is exclusively discontinuous (integral multiples of hυ).

Since then, ZPE theories have gone on a course independent from Planck's second theory. The more recent root of modern ZPE theories stems from the work of H. Casimir who, in 1948, apparently showed the existence of a force acting between two uncharged parallel plates. Fundamentally the Casimir effect is predicated upon the existence of a background field of energy permeating even the “vacuum”, which exerts a radiation pressure, homogeneously and from all directions in space, on every body bathed in it. Given two bodies or particles in proximity, they shield one another from this background radiation spectrum along the axis (i.e. the shortest distance) of their coupling, such that the radiation pressure on the facing surfaces of the two objects would be less than the radiation pressure experienced by all other surfaces and coming from all other directions in space. Under these conditions, the two objects are effectively pushed towards one another as if by an attractive force. As the distance separating the two objects diminishes, the force pushing them together increases until they collapse one on to the other. In this sense, the Casimir effect would be the macroscopic analogy of the microscopic van der Waals forces of attraction responsible for such dipole-dipole interactions as hydrogen bonding. However, it is worth noting that the van der Waals force is said to tend to establish its normal radius, or

the optimal distance between dipoles, as the distance where the greatest attractive force is exerted, beyond which the van der Waals forces of nuclear and electronic repulsion overtake the attraction force.

Subsequently, another Dutch physicist, M. Sparnaay, demonstrated that the Casimir force did not arise from thermal radiation and, in 1958, went on to attribute this force to the differential of radiation pressure between the ZPE radiation from the vacuum state surrounding the plates and the ZPE radiation present in the space between them. Sparnaay's proposal is that a classical, non-quantal, isotropic and ubiquitous electromagnetic zero-point energy exists in the vacuum, and even at a temperature of absolute zero. It is further assumed that since the ZPE radiation is invariant with respect to the Lorentz transformations, it obeys the rule that the intensity of its radiation is proportional to the cube of the frequency, resulting in an infinite energy density for its radiation spectrum.

What appeared to be the virtue of this reformulated theory was the notion that the vacuum no longer figured as pure space empty of energy, but rather as a space exposed to constantly fluctuating “fields of electromagnetic energy”.

Puthoff has utilised the isomorphism between van der Waals and Casimir forces to put forth the zero-point (ZP) energy theory of gravity, based on the interpretation that the virtual electromagnetic ZP field spectrum predicted by quantum electrodynamics (QED) is functionally equivalent to an actual vacuum state defined as a background of classical or Maxwellian electromagnetic radiation of random phases, and thus can be treated by stochastic electrodynamics (SED). Whereas in QED, the quanta are taken as virtual entities and the infinite energy of the vacuum has no physical reality, for SED, the ZPE spectrum results from the distortion of a real physical field and does not require particle creation. Gravity then, could be seen as only the macroscopic manifestation of the Casimir force.

We do not dispute the fact that even in space-absent matter, there is radiant energy present which is not of a thermal nature. But we claim that this energy is not electromagnetic, nor is its energy spectrum-infinite. That this is so, stems not just from our opinion that it is high time that Einstein's heuristic hypothesis should be taken as literally factual - in the dual sense that all electromagnetic energy is photon energy and all photons are local productions, but above all from the fact that it is apparent, from the experiments of Wang and his colleagues (Wang, Li, Kuzmich, A & Dogariu, A. "Gain-assisted superluminal light propagation", Nature 406; #6793; 277), that the photon stimulus can propagate at supraluminal speeds and lies therefore well outside of any scope of electromagnetic theory, be this Maxwell's classical approach taken up by ZPE theories, or Einstein's special relativistic phenomenology of Maxwell's theory. The fact is, that if the light stimulus can propagate at speeds greater than those of light, then what propagates is not light at all, and thus not energy configured electromagnetically. Light is solely a local production of photons in response to the propagation of a stimulus that itself is not electromagnetic.

It is critical to understand that the implication from this, that - aside from local electromagnetic radiation and from thermal radiation associated with the motions of molecules (thermo-mechanical energy), there is at least one other form of energy radiation which is everywhere present, even in space-absent matter. Undoubtedly, it is that energy which prevents any attainment of absolute zero, for any possible local outpumping of heat is matched by an immediate local conversion of some of this energy into a minimum thermal radiation required by the manifolds of Space and Time. Undoubtedly also, this radiation is ubiquitous and not subject to relativistic transformations (i.e. it is Lorentz invariant). What it is not, is electromagnetic radiation consisting of randomistic phases of transverse waves.

To understand this properly, one must summarise the differences from existing ZPE theories - and all these differences come down to the fact that this energy, which is neither electromagnetic nor thermal per se, (and is certainly not merely thermo-mechanical), has nevertheless identifiable characteristics both distributed across sub-types or variants and also common to all of them.

Essentially, the first sub-type or variant consists of longitudinal mass-free waves which deploy electric energy. They could well be called Tesla waves, since Tesla-type transformers can indeed be shown experimentally to radiate mass-free electric energy, in the form of longitudinal magnetic and electric waves having properties not reducible to photon energy nor to “electromagnetic waves”, and having speeds of displacement which can be much greater than the limit c for all strictly electromagnetic interactions.

One may well denote the second sub-type by the designation of mass-free thermal radiation, since it contributes to temperature changes - and, as obviously indicated by the impossibility of reaching an absolute zero of temperature, this contribution occurs independently of the presence of matter, or mass-energy, in Space. In other words, not all thermal radiation can be reduced to vibration, rotation and translation (drift motion) of molecules, i.e. to thermomechanical energy, because the properties of pressure and volume which determine temperature and affect matter, appear indeed to a great extent to be independent from matter, a fact which itself is responsible for the observed catastrophic and unexpected phase changes of matter and has required to this day the insufficient explanation offered semi-empirically by the Van der Waals Force Law.

Finally, the third sub-type may be designated latent mass-free energy radiation - since it deploys neither charge, nor thermal or baroscopic effects, and yet it is responsible for "true latent heat" or for the "intrinsic potential energy" of a molecule. It is also responsible for the kineto-regenerative phenomenon whereby an electroscope performs a variable charge-mediated work against the local gravitational field.

The common characteristic of all three sub-types of mass-free energy radiation is that they share the same non-classical fine structure, written as follows for any energy unit, where **c** is any speed of light wave function, and the wavelength λ and wave function W are interconnected as a function of the physical quality of the energy field under consideration: $E=\lambda cW$

In the instance of longitudinal electric radiation, this takes on the directly quantifiable form:

$$E=(\lambda_q c)W_v=p_e W_v=(h/\lambda_x)W_v=\int=qV$$

where:

W_V is the voltage-equivalent wave function corresponding to V,

P_e constitutes the linear momentum corresponding to the conventional q or e,

h is the Planck constant,

λ_x is the Duane-Hunt constant expressed as a wavelength,

λ_q is a wavelength constant; and the sign

$=\int=$ signifies exact equality between an expression in the conventional dimensions of length, mass and time, and an expression in length and time dimensions alone.

In the instance of mass-free thermal radiation (contributing to temperature changes), the transformation obeys Boltzmann's rule (**k** is now Boltzmann's constant and T is Kelvin-scale temperature):

$$E=\lambda_{n1}cW_{n1}=\lambda_{n1}(\pi_V\zeta_p)(\lambda_{n1})\sim=kT$$

and in the third instance - of latent mass-free radiation, the transformation obeys the rule:

$$E=\lambda_{n1}cW_{n1} = \lambda_{n1}(\lambda_{n1}\zeta_{n1})(\lambda_{n1}f_{n1}) = \lambda_{n1}{}^3\zeta_{n1}f_{n1}$$

where ζ and f are frequency functions, f being a specific gravitational frequency term, and f_{n1} being defined as equal to $(\lambda_{n1})^{-0.5}\text{meter}^{0.5}\text{sec}^{-1}$ and ζ_{n1} has the value of c/λ_{n1}

If the electric variant of mass-free radiation has a direct quantum equivalence, via the Duane-Hunt Law, none of the three primary aether energy variants possess either the classic form of electromagnetic energy which requires square superimposition of speed of light wave functions **c**, as $\mathbf{c^2}$, or the quantum form of energy, requiring $E = hu$. The critical first step in the right direction may well be attributed to Dr. W. Reich, as it regards the fact that mass-free energy couples two unequal wave functions, only one of which is electromagnetic and abides by the limit **c**. We then unravelled the threefold structure described above, and further showed that, in the case of longitudinal electric waves, the postulated equivalence $(q=\lambda_q c)$ is merely phenomenological, as these waves are not restricted by the function **c** in their conveying of electric charge across space. It can further be demonstrated that all black-body photons are bound by an upper frequency limit (64×10^{14} Hz), above which only ionising photons are produced, and that all black-body photons arise precisely from the interaction of mass-free electric radiation with molecules of matter (including light leptons), whereby the energy of that radiation is locally converted into photon or electromagnetic radiation. In other words, all non-ionising electromagnetic energy appears to be secondary energy which results locally from the interaction of matter with mass-free electric energy. It cannot therefore consist of the primary energy that is present in the vacuum, an energy that is neither virtual nor electromagnetic, but actual and concrete in its electric, thermal and antigravitic manifestations. Lastly, gravitational energy, being either the potential or the kinetic energy responsible for the force of attraction between units of matter, is a manifestation that also requires, much as electromagnetic radiation does, coupling of mass-free energy to matter or to mass-energy.

The Tesla coil is a generator of a mass-free electric energy flux which it transmits both by conduction through the atmosphere and by conduction through the ground. Tesla thought it did just that, but it has been since regarded instead (because of Maxwell, Hertz and Marconi) as a transmitter of electromagnetic energy. The transmitter operates by a consumption of mass-bound electric power in the primary, and by induction it generates in the

coupled secondary two electric fluxes, one mass-bound in the coil conductor, and the other mass-free in the body of the solenoid. Tesla also proposed and demonstrated a receiver for the mass-free energy flux in the form of a second Tesla coil resonant with the first. The receiver coil must be identical and tuned to the transmitter coil; the capacitance of the antenna plate must match that of the transmitter plate; both transmitter and receiver coils must be grounded; and the receiver coil input and output must be unipolar, as if the coil were wired in series.

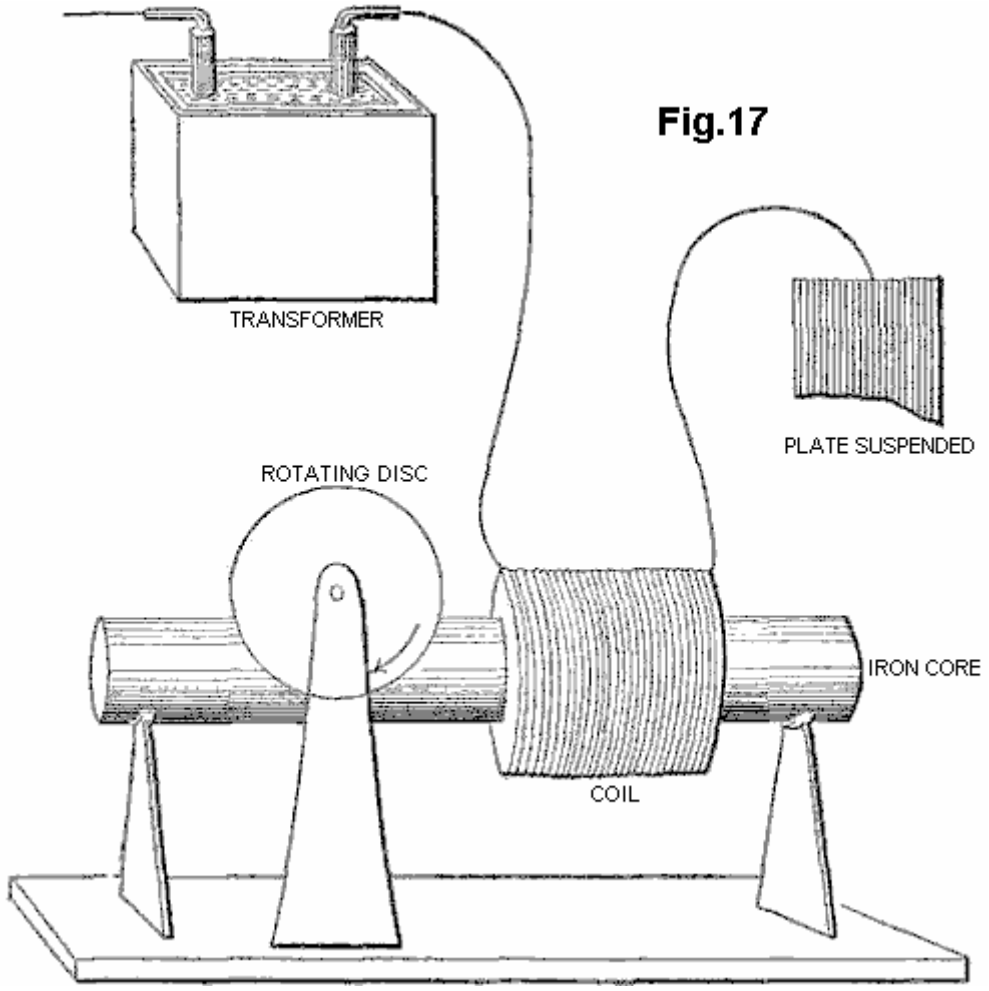
The generators of mass-free energy with which we are concerned, provide current pulses associated with a dampened wave (DW) oscillation of much higher frequency than the pulse repetition frequency. A particular problem in recovering the mass-free energy content of such pulses is provided by the dampened wave oscillations. Although in our U.S. Pat. No. 5,416,391 we describe arrangements incorporating split phase motors to recover such energy, their efficiency is a great deal less than what should theoretically be attainable. Other workers such as Tesla and Reich, have encountered the same problem to an even greater degree.

In nineteenth century motor engineering terminology, dynamos capable of producing direct current by continuous homopolar induction were known as “unipolar” generators. The term “unipolar induction” appears to have originated with W. Weber, to designate homopolar machines where the conductor moves continuously to cut the magnetic lines of one kind of magnetic pole only, and thus require sliding contacts to collect the generated current. Faraday's rotating copper disc apparatus was, in this sense, a homopolar generator when the disc was driven manually, or a homopolar motor when the current was provided to it. Where the rotating conductor continuously cuts the magnetic field of alternatingly opposite magnetic poles, the operation of a machine, whether a generator or a motor, is said to be “heteropolar”. Unipolar machines went on to have a life of their own in the form of low voltage and high current DC generators - from Faraday, through Plucker, Varley, Siemens, Ferraris, Hummel, to Lord Kelvin, Pancinoti, Tesla and others - almost exclusively in the form of disc dynamos, but some having wound rotors.

In Mordey's alternator, and in so-called “inductor alternators”, however, homopolar generators were employed to obtain alternating currents, with the use of rotors wound back and forth across the field. Use of smooth, unwound rotors in AC induction motors (as opposed to AC synchronous motors, such as hysteresis motors) was a later development than homopolar dynamos. By 1888, Tesla and Ferraris amongst still others, had independently produced rotating magnetic fields in a motor, by employing two separate alternate currents with the same frequency but different phase. Single phase alternate current motors were developed later, and split-phase motors were developed last. Ferraris (Ferraris, G (1888) "Rotazioni elettrodynamiche", Turin Acad, March issue.) proposed the elementary theory of the 2-phase motor, where the current induced in the rotor is proportional to the slip (the difference between-the angular velocity of the magnetic field and that of the rotating cylinder), and the power of the motor is proportional to both the slip and the velocity of the rotor.

If an iron rotor is placed within the rotating magnetic field of a 2-phase stator, it will be set in rotation, but not synchronously, given that it is always attracted to the moving magnetic poles with a lag. But if an aluminium or copper rotor is used instead, it gets “dragged” around by the rotating stator field because of the eddy currents induced in it. If the aluminium or copper rotor were to rotate synchronously with the stator magnetic field, there would be no induced eddy currents and thus no motor action would result. The motor action depends, in this instance, upon the presence of asynchronous slip, since the function of the latter is to sustain the induction of those currents in the rotor that are responsible for the motor action of the dragged rotor. This then is the origin of the term “AC drag motors”. Once the drag rotor evolved from a cylinder to a hollow cup, they earned the epithet of “drag-cup motors”. Later, already in the 20th century, the cups were fitted over a central stator member, and the sleeve rotor 2-phase servo motor was born.

Tesla knew that impulse currents as well as CW (constant wave) sinusoidal currents could be used to drive AC motors. Regarding his invention of a hysteresis motor (which he called a “magnetic lag motor”), he stated: " . . . pulsatory as well as an alternating current might be used to drive these motors . . . " (Martin, T C (1894) "The inventions, researches and writings of Nikola Tesla", Chapter XII, p. 68). In his search for efficient utilisation of the high frequency DW (dampened wave) impulse currents of his induction coils, Tesla began by employing an AC disc induction motor as shown in Fig.17 of his famous 1892 address (Tesla, N (1892) "Experiments with alternate currents of high potential and high frequency", in "Nikola Tesla Lectures", 1956, Beograd, pp. L-70-71). This consisted of a copper or aluminium disc mounted vertically along the longitudinal axis of an iron core on which was wound a single motor coil which was series wired to the distal terminal of an induction coil at one end, and to a large suspended and insulated metal plate at the other. What was new about this was the implementation of an AC disc induction motor drive, where the exciting current travelled directly through the winding with just a unipolar connection to the coil secondary (under certain conditions, even the series connection to the plate could be removed, or replaced with a direct connection to the experimenter's body): "What I wish to show you is that this motor rotates with one single connection between it and the generator" (Tesla, N. (1892), op. cit., L-70, Tesla's emphasis). Indeed, he had just made a critical discovery that, unlike in the case of mass-bound charge where current flow requires depolarisation of a bipolar tension, mass-free charge engages current flow unipolarly as a mere matter of proper phase synchronisation:



Tesla thought that his motor was particularly adequate to respond to windings which had “high-self-induction”, such as a single coil wound on an iron core. The basis of this self-induction is the magnetic reaction of a circuit, or an element of a circuit - an inductor - whereby it chokes, dims or dampens the amplitude of electric waves and retards their phase.

For the motor to respond to still higher frequencies, one needed to wind over the primary motor winding, a partial overlap secondary, closed through a capacitor, since "it is not at all easy to obtain rotation with excessive frequencies, as the secondary cuts off almost completely the lines of the primary" (Idem, L-71.).

Tesla stated that "an additional feature of interest about this motor" was that one could run it with a single connection to the earth ground, although in fact one end of the motor primary coil had to remain connected to the large, suspended metal plate, placed so as to receive or be bathed by "an alternating electrostatic field", while the other end was taken to ground. Thus Tesla had an ordinary induction coil that transmitted this "alternating electrostatic field", an untuned Tesla antenna receiving this "field", and a receiver circuit comprising his iron-core wound motor primary, a closely coupled, capacitatively closed secondary, and the coupled non-ferromagnetic disc rotor. Eventually, in his power transmission system, he would replace this transmitter with a Tesla coil, and place an identical receiving coil at the receiving end, to tune both systems and bring them into resonance. But his motor remained undeveloped, and so did the entire receiver system.

Tesla returned to this subject a year later, saying "on a former occasion I have described a simple form of motor comprising a single exciting coil, an iron core and disc" (Tesla, N (1893) "On light and other high frequency phenomena", in "Nikola Tesla Lectures", 1956, Beograd, pp. L-130, and L-131 with respect to Fig.16-II). He describes how he developed a variety of ways to operate such AC motors unipolarly from an induction transformer, and as well other arrangements for "operating a certain class of alternating motors founded on the action of currents of differing phase". Here, the connection to the induction transformer is altered so that the motor primary is driven from the coarse secondary of a transformer, whose finer primary is coupled, at one end, directly and with a single wire to the Tesla secondary, and at the other left unconnected. On this occasion, Tesla mentions that such a motor has been called a “magnetic lag motor”, but that this expression (which, incidentally, he had himself applied to his own invention of magnetic hysteresis motors) is objected to by "those who attribute the rotation of the disc to eddy currents when the core is finally subdivided" (Tesla, N (1893), op. cit., p. L-130).

In none of the other motor solutions, 2-phase or split-phase, that he suggests as unipolar couplings to the secondary of an induction coil, does the non-ferromagnetic disc rotor motor again figure. But he returns to it a page later, and indirectly so, by first addressing the disadvantages of ferromagnetic rotors: "Very high frequencies are of course not practicable with motors on account of the necessity of employing iron cores. But one may use sudden discharges of low frequency and thus obtain certain advantages of high-frequency currents-without rendering the iron core entirely incapable of following the changes and without entailing a very great expenditure of energy in the core. I have found it quite practicable to operate, with such low frequency disruptive discharges of condensers, alternating-current motors."

In other words--whereas his experiments with constant wave (CW) alternating currents, and as well with high-voltage dampened wave (DW) impulses from induction coils, indicated the existence of an upper frequency limit to iron core motor performance, one might employ instead high-current, DW impulses - of high DW frequencies but low impulse rates - to move these motors quite efficiently. Then he adds "A certain class of [AC] motors which I advanced a few years ago, that contain closed secondary circuits, will rotate quite vigorously when the discharges are directed through the exciting coils. One reason that such a motor operates so well with these discharges is that the difference of phase between the primary and secondary currents is 90 degrees, which is generally not the case with harmonically rising and falling currents of low frequency. It might not be without interest to show an experiment with a simple motor of this kind, inasmuch as it is commonly thought that disruptive discharges are unsuitable for such purposes."

What he proposes next, forms the basis of modern residential and industrial AC electric power meters, the AC copper disc motor whose rotor turns on the window of these meters, propelled forward by the supply frequency. But instead of employing any such Constant Wave input, Tesla uses the disruptive discharges of capacitors, incipiently operating as current rectifiers. With the proper conditions, e.g. correct voltage from the generator, adequate current from the capacitor, optimum capacitance for the firing rate, and tuned spark-gap, to mention a few, Tesla found that the non-ferromagnetic disc rotor turned but with considerable effort. But this hardly compared to the results obtained with a high-frequency CW alternator, which could drive the disc "with a much smaller effort". In summary then, Tesla went as far as being the first to devise a motor driven by Tesla waves, that employed a non-ferromagnetic rotor, and whose arrangement encompassed both transmitter and receiver circuits. For this purpose, he employed a single-phase method in which the signal is fed unipolarly to the winding, placed in series with a plate capacitance.

Tesla also later proposed driving a similar single-phase non-ferromagnetic disc motor from bipolar capacitative discharges through an atmospheric spark-gap now placed in parallel with the main motor winding, and again simulating a split-phase by a closely-wound secondary which was closed by a capacitance.

As Tesla admits, the results of all his AC eddy current motor solutions were meagre and limited by current and frequency problems. Likewise, the two-phase arrangements proposed by Reich for his OR motor, involving a superimposition of the Dampened Waves of a first phase on a fixed Continuous Wave second phase, require an external power source and a pulse amplifier circuit, and failed to meet Reich's own requirements.

We have previously proposed the use of squirrel cage motors with capacitative splitting of phase to convert the Dampened Wave output of plasma pulsers, but once a Squirrel Cage is introduced, the dampening effect which the non-ferromagnetic copper cage exerts in being dragged by the revolving stator field, is counteracted by the ferromagnetic cylinder of laminated iron, in which the copper cage is embedded, working to diminish the slip and bring the rotor to near synchronism. This is, in all likelihood, what limits Squirrel Cage motors responding to the DC component of the Dampened Wave impulse, and thus be limited to respond to fluxes of mass-bound charges. Historically, as we shall see, the obvious advantage of the Squirrel Cage servo motors lay in the fact that, in particular for 2-phase applications, they were far more efficient at performing work without evolution of heat. Indeed, if the eddy currents in the non-ferromagnetic rotor are permitted to circulate in non-ordered form, the rotor material and stator will heat up rapidly and consume much power in that heating. This is in fact considered to be a weakness of AC non-ferromagnetic-rotor induction motors.

SUMMARY OF THE INVENTION

The present invention is concerned with conversion to conventional electrical energy of the variants of mass-free energy radiation considered above, referred to for convenience as Tesla waves, mass-free thermal radiation and latent mass-free radiation. The first variant of such radiation was recognised, generated and at least partially disclosed by Tesla about a hundred years ago, although his work has been widely misinterpreted and also confused with his work on the transmission of radio or electromagnetic waves. The Tesla coil is a convenient generator of such radiation, and is used as such in many of the embodiments of our invention described below, but it should be clearly understood that our invention in its broadest sense is not restricted to the use of such a coil as a source of mass-free radiation and any natural or artificial source may be utilised. For example, the sun is

a natural source of such radiation, although interaction with the atmosphere means that it is largely unavailable at the earth's surface, limiting applications to locations outside of the earth's atmosphere.

According to the invention, a device for the conversion of mass-free radiation into electrical or mechanical energy comprises a transmitter of mass-free electrical radiation having a dampened wave component, a receiver of such radiation tuned to resonance with the dampened wave frequency of the transmitter, a co-resonant output circuit coupled into and extracting electrical or kinetic energy from the receiver, and at least one structure defining a transmission cavity between the transmitter and the receiver, a full-wave rectifier in the co-resonant output circuit, and an oscillatory pulsed plasma discharge device incorporated in the co-resonant output circuit. The output circuit preferably comprises a full-wave rectifier presenting a capacitance to the receiver, or an electric motor, preferably a split-phase motor, presenting inductance to the receiver. The transmitter and receiver each preferably comprise a Tesla coil and/or an autogenous pulsed abnormal glow discharge device. The transmission cavity is preferably at least partially evacuated, and comprises spaced plates connected respectively to the farthest out poles of the secondaries of Tesla coils incorporated in the transmitter and receiver respectively, the plates being parallel or concentric. The structure defining the cavity may be immersed in ion-containing water. The split-phase motor is preferably an inertially-dampened AC drag motor.

The invention, and experiments demonstrating its basis, are described further below with reference to the accompanying drawings.

SHORT DESCRIPTION OF THE DRAWINGS

Fig.1 is a schematic view of a Tesla coil connected to a full-wave rectifier to form an energy conversion device:

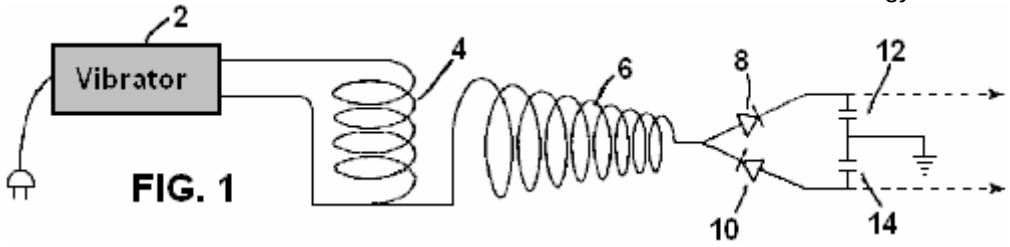


FIG. 1

Fig.2 is a schematic view of a Tesla coil connected to a gold leaf electrometer:

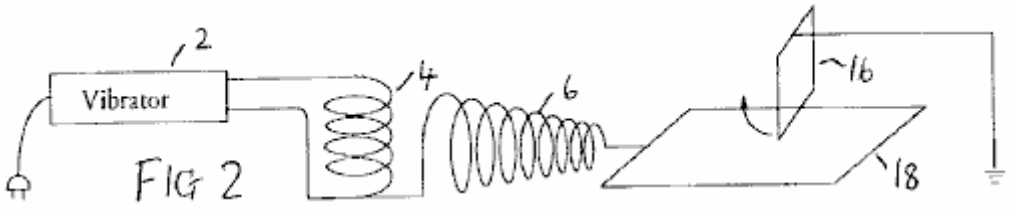


FIG 2

Fig.3 to Fig.6 show alternative electrometer configurations:

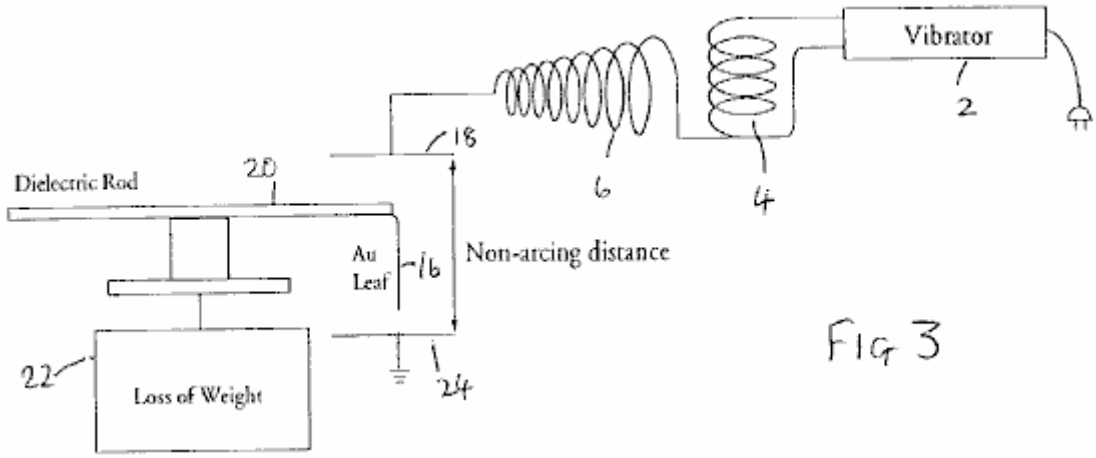


FIG 3

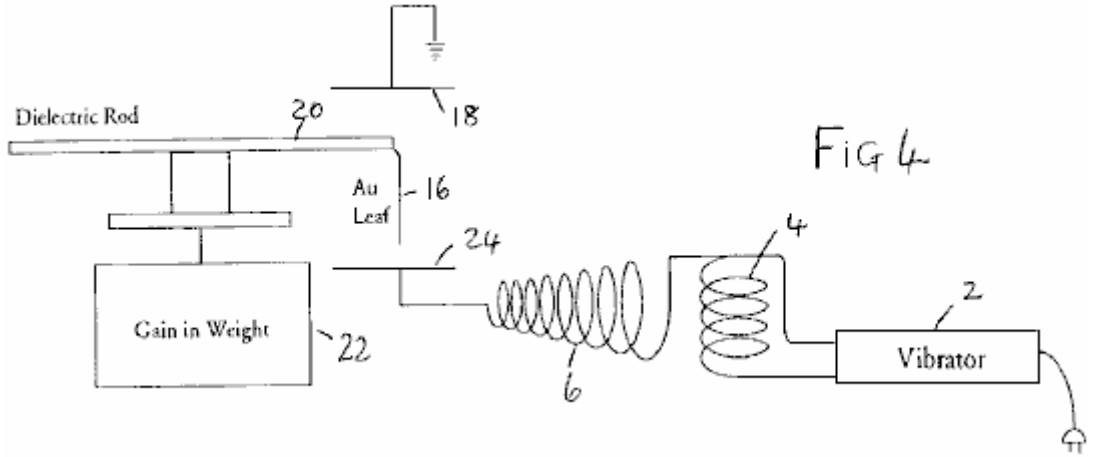


FIG 4

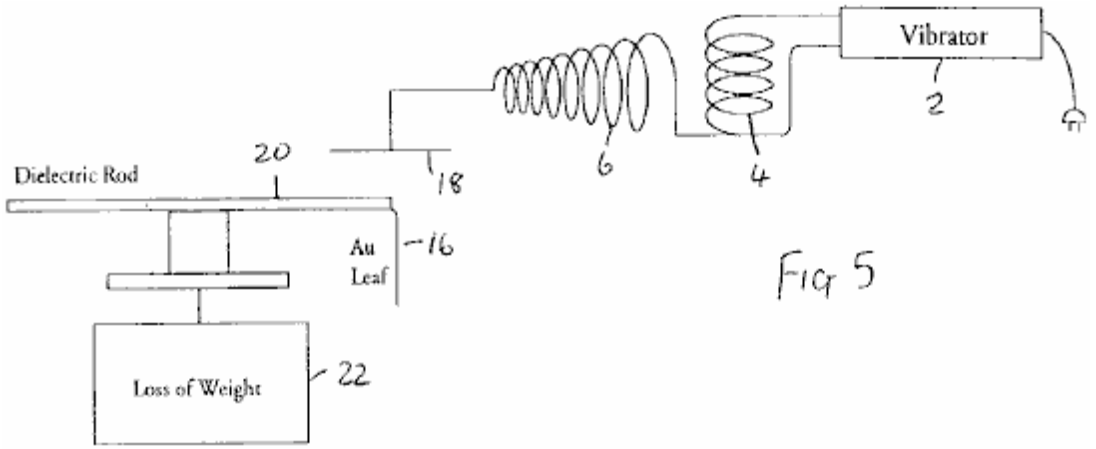


FIG 5

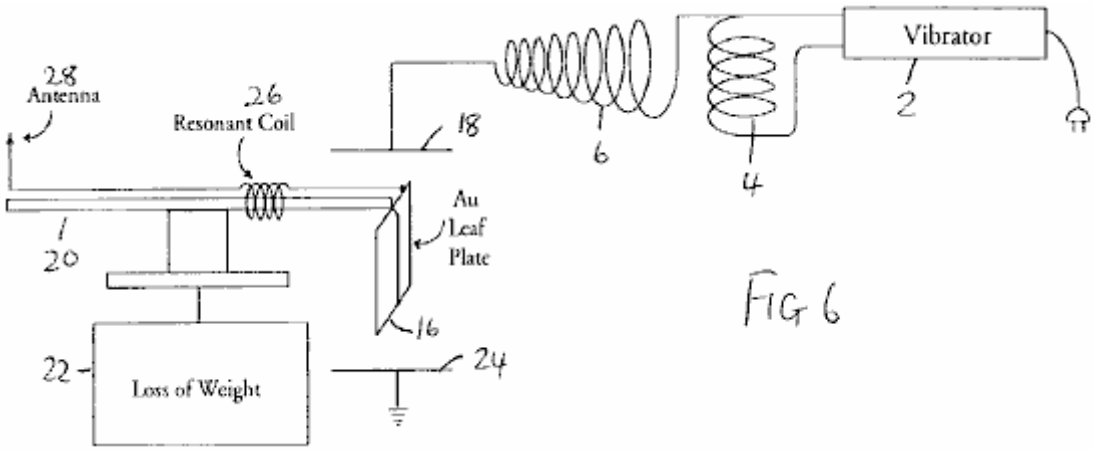


FIG 6

Fig.7 to Fig.11 show modifications of the circuit of **Fig.1**:

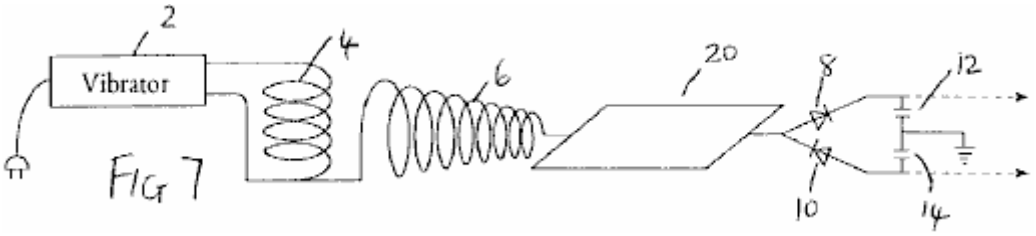


FIG 7

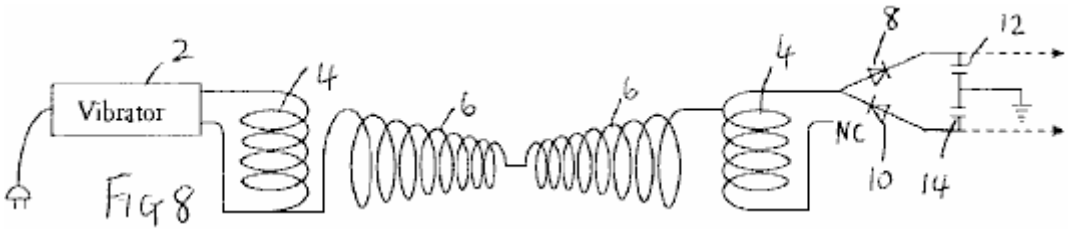


FIG 8

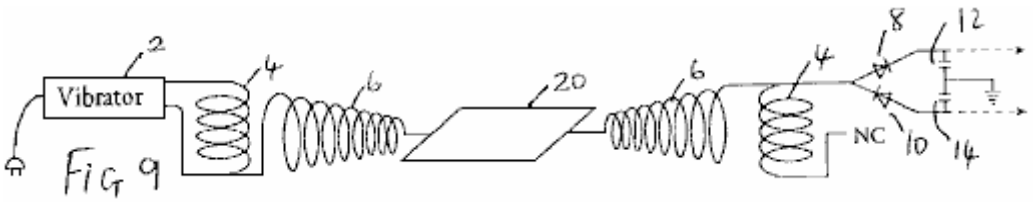


FIG 9

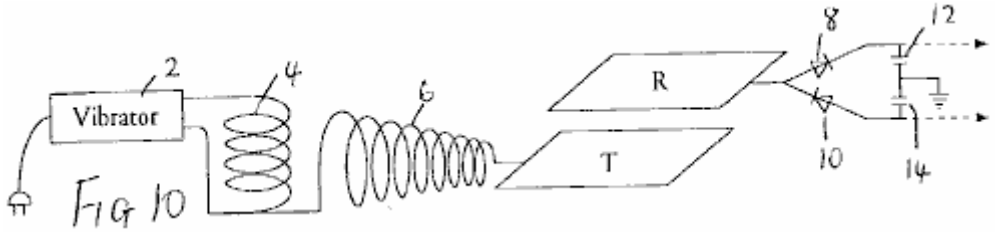


FIG 10

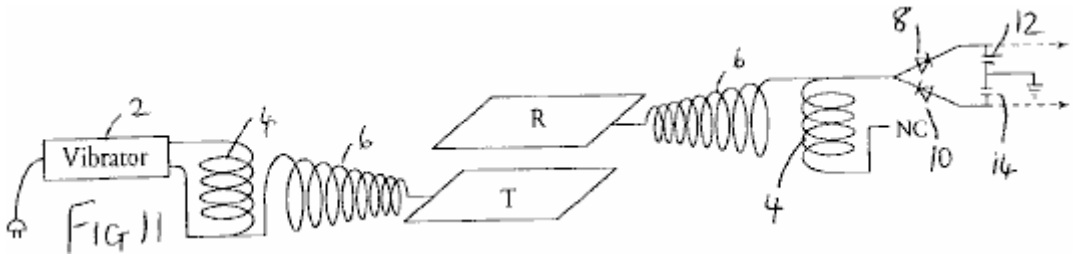


FIG 11

Fig.12 shows apparatus for investigating aspects of the experimental results obtained with the foregoing devices;

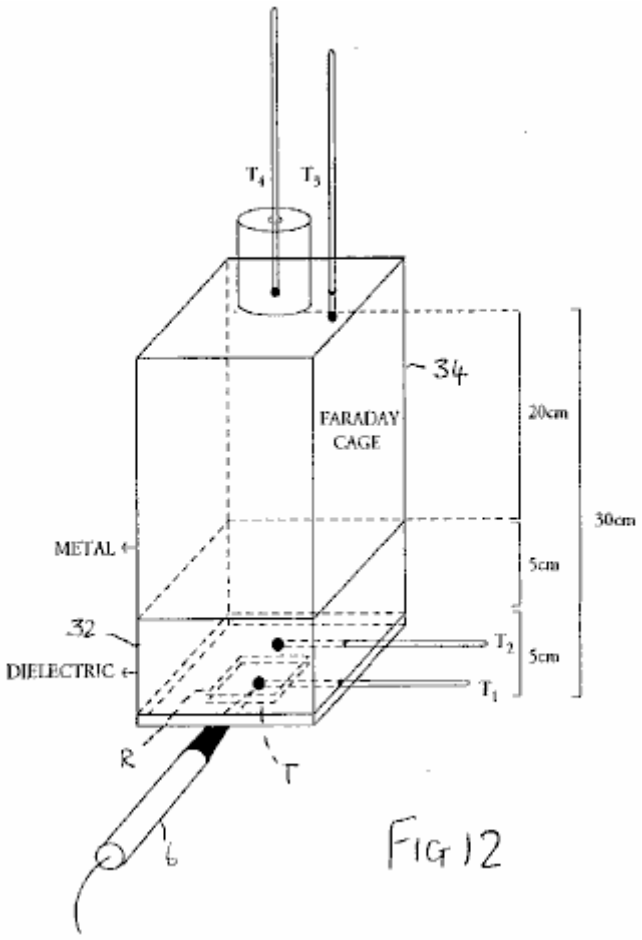


FIG 12

Fig.13 is a graph illustrating results obtained from the apparatus of **Fig.12**:

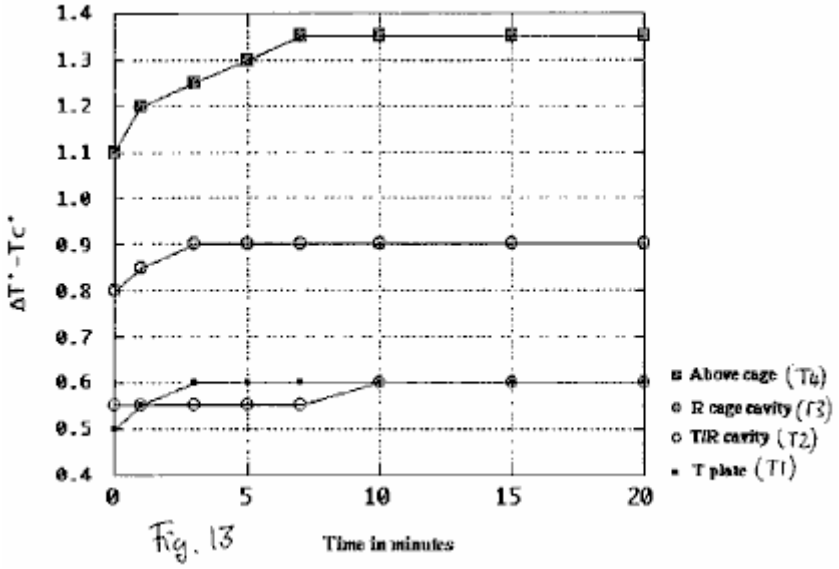


Fig. 13

Fig.14 to Fig.17 show schematic diagrams of embodiments of energy conversion devices:

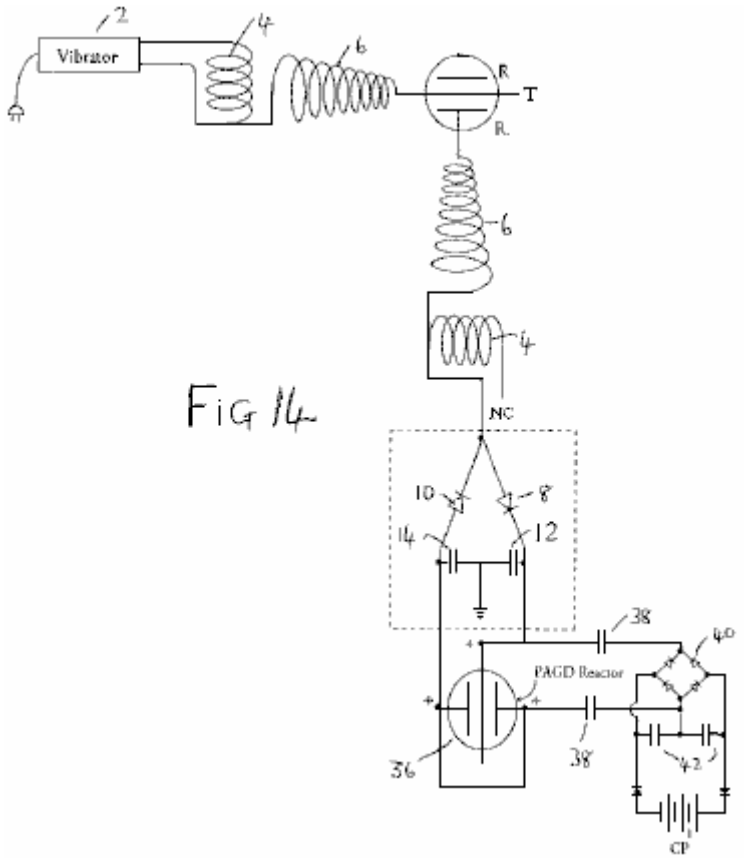


Fig 14

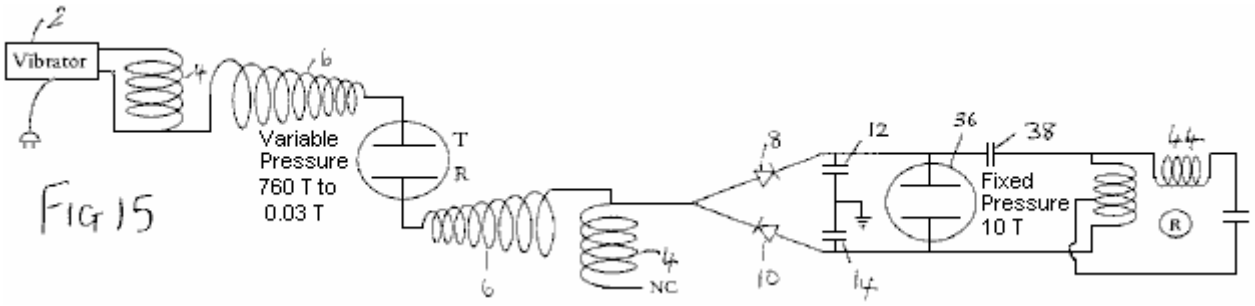


FIG 15

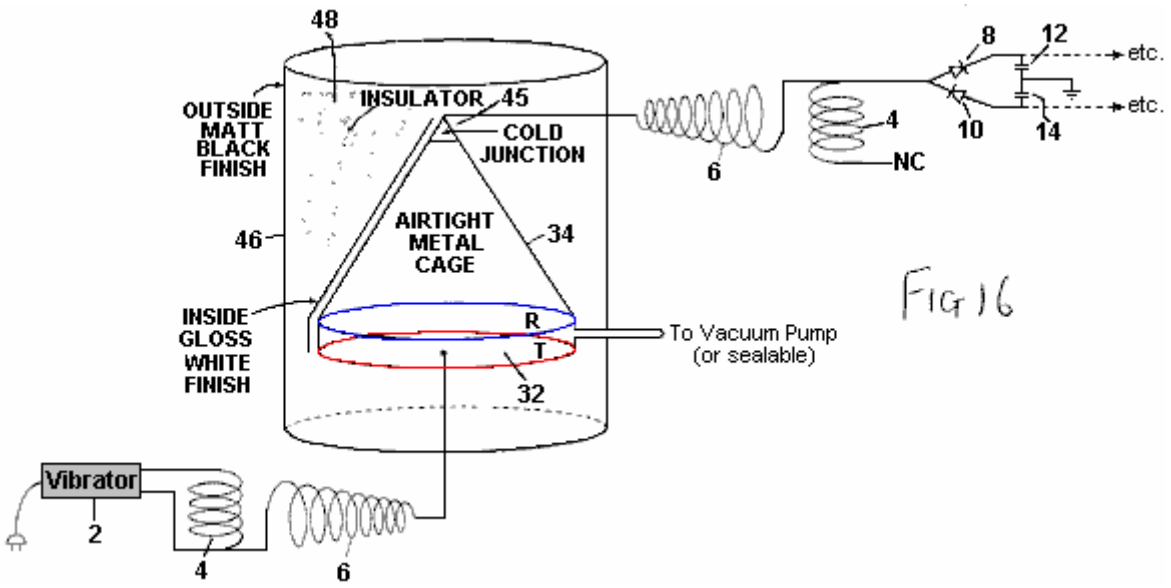


FIG 16

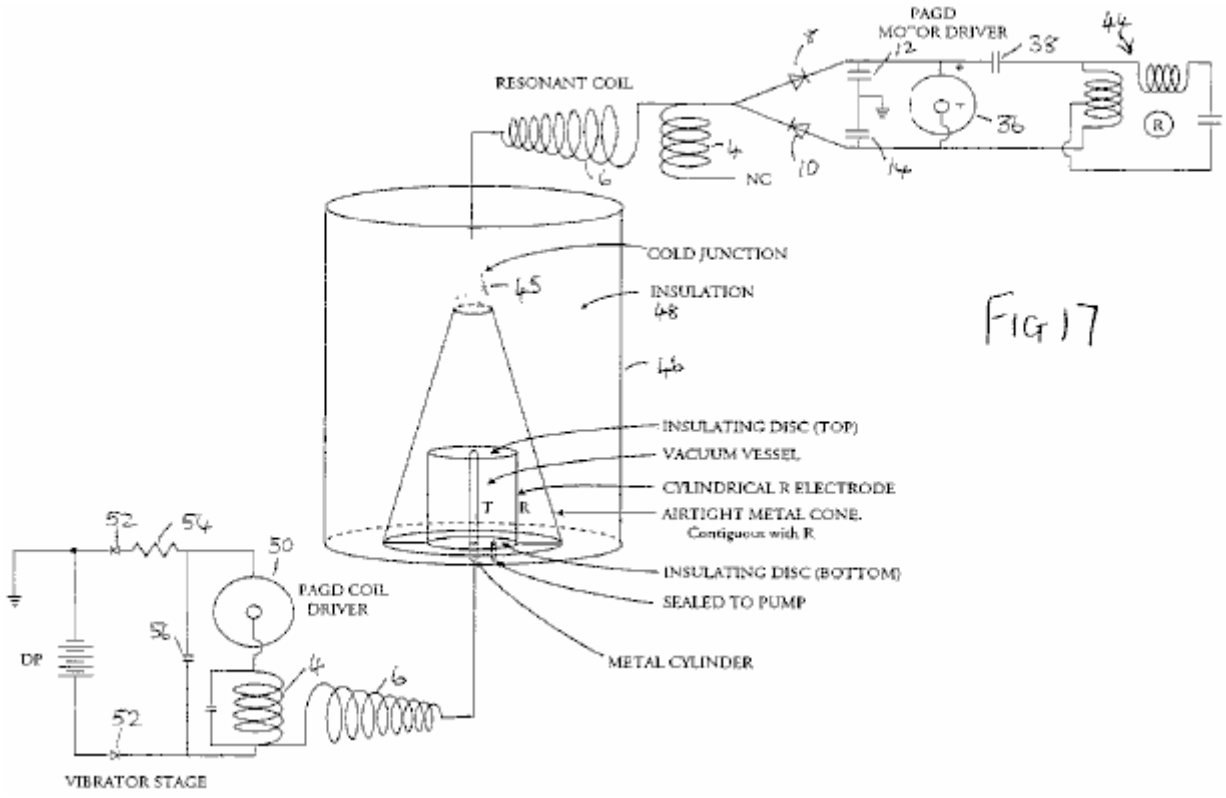


FIG 17

Fig.18 is a diagrammatic cross-section of an inertially dampened drag cup motor:

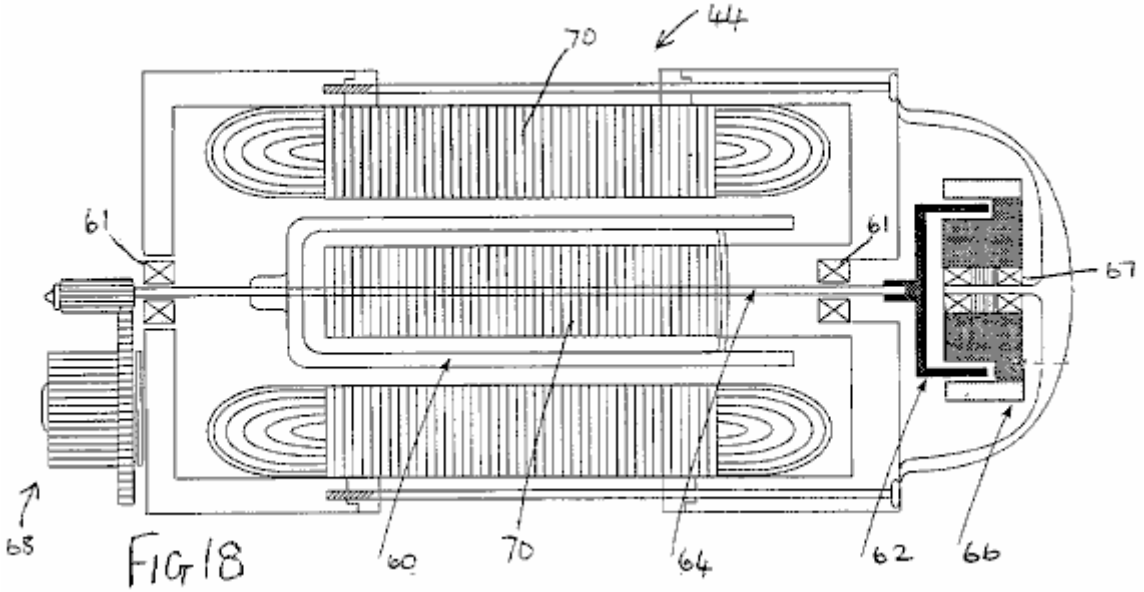


FIG 18

Fig.19 is a schematic diagram of a further embodiment of an energy conversion device incorporating such a motor:

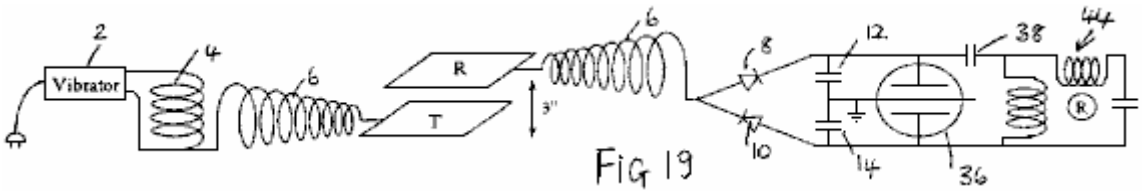


FIG 19

DESCRIPTION OF THE PREFERRED EMBODIMENTS

Based upon observations of weight loss in metallic matter as induced by exposure to high frequency alternating electric fields, we developed an experimental method to optimise this-weight loss, and from this a device that treats the forces causing weight loss as manifestations of intrinsic potential energy ΔU (or true "latent heat") of the molecules of matter, and converts both "true latent heat" energy present in the neighbourhood of a receiver, and "sensible" heat induced within that receiver, into electric energy which can be used to drive a motor, flywheel or charge batteries.

It is commonly believed that the output of the Tesla coil is ionising electromagnetic radiation. We have demonstrated that it is not, i.e. that it is neither electromagnetic radiation, nor ionising electromagnetic radiation. The output of an air-cored, sequentially-wound secondary, consists exclusively of electric energy: upon contact with the coil, a mass-bound AC current can be extracted at the resonant frequency, whilst across a non-sparking gap, mass-free AC-like electric wave radiation having the characteristics of longitudinal waves, can be intercepted anywhere in adjacent space. Accordingly, the radiation output from such coils is different to electromagnetic radiation.

The basic demonstration that the output of a Tesla coil does not consist of ionising radiation, is that it does not accelerate the spontaneous discharge rate of electroscopes, whether positively or negatively charged. In fact, in its immediate periphery, the coil only accelerates the spontaneous discharge rate of the negatively charged electroscope (i.e. the charge leakage rate), whereas it arrests the discharge of the positively charged electroscope (i.e. the charge seepage rate falls to zero). But this dual effect is not due to any emission of positive ions from the secondary, even if it can positively charge a discharged electroscope brought to its proximity. This charging effect is in fact an artifact, in that metals but not dielectrics are ready to lose their conduction and outer valence band electrons when exposed to the mass-free electric radiation of the coil.

This is simply demonstrated by the apparatus of **Fig.1**, in which the outer terminal of the secondary winding **6** of a Tesla coil having a primary winding **4** driven by a vibrator **2** is connected to the input of a full-wave voltage wave divider formed by diodes **8** and **10** and reservoir capacitors **12** and **14** (the same reference numerals are used for similar parts in subsequent diagrams). If the rectifiers employed are non-doped, then the coil appears to only charge the divider at the positive capacitance **10**, but if doped rectifiers are employed, the coil will be observed to charge both capacitances equally. Whereas positive ionises can charge either doped or un-doped dividers

positively, no positive ionise can charge a doped divider negatively, clearly demonstrating that the Tesla coil does not emit positive ions.

The basic demonstration that the output of a Tesla coil is not non-ionising electromagnetic radiation of high frequency, such as optical radiation, or of lower frequency, such as thermal photons, is also a simple one. Placement of a sensitive wide spectrum photoelectric cell (capable of detecting radiation to the limits of vacuum UV), wired in the traditional closed circuit manner from a battery supply, at any distance short of sparking from the outer terminal of the coil will show in the dark that the light output from the coil is negligible. This rules out optical radiation at high frequency. The demonstration that the sensible heat output from the Tesla coil is also negligible will be addressed below.

Our theory proposed the existence of physical processes whereby mass-free electric radiation can be converted into electromagnetic radiation. Such a process is at work whenever mass-free electric wave radiation interacts with electrons, such as those that remain in the valence bands of atoms. This mass-free electric energy interacts with charge carriers, such as electrons, to confer on them an electrokinetic energy which they shed in the form of light whenever that electrokinetic energy is dissociated from those carriers (e.g. by deceleration, collision or friction processes). Such a process is at work to a negligible extent in the coil itself and its usual terminal capacitance, hence the faint glow that can be seen to issue from it, but it can also be greatly amplified in the form of a corona discharge by connecting a large area plate to the output of the secondary, as Tesla himself did in his own experiments, and thus by increasing the capacitance of the coil system.

Now, what is interesting in this process is that, in the absence of virtually any I^2R losses at the plate, and if the plate thus introduced is bent at the edges so that it has no pointed edges, or if it is in the form of a bowl, or in any other manner that precludes sparking at edges and specially corners, and thus enhances the corona discharge, any electroscope, whether negatively or positively charged, now brought close to the plate will show a tendency to arrest its spontaneous discharge rate. One might say that this is simply the result obtained in a Faraday cage which disperses charge on its outside and electrically insulates its interior, and indeed if an electroscope is placed inside a Faraday cage no amount of Tesla radiation on the outside of that cage, save direct sparking, adversely affects the leakage or seepage rate of the electroscope. In fact, since the effect of such a cage can be shown to be that of, by itself, inducing arrest of either spontaneous electroscopic discharge, this effect simply remains or is magnified when the cage is bathed by Tesla radiation. However, a cage constitutes an electrically isolated environment, whereas a plate with or without curved or bent edges does not. Furthermore, the change observed in the properties of the output radiation from a Tesla coil when certain metal plates or surfaces are directly connected to the outer terminal of the secondary, takes place whilst the capacitance of the coil is increased by the connected plate, and thus the plate is an electrically active element of the circuit - and hence the opposite of an electrically isolated element.

For a long time, we believed that the anomalous cathode reaction forces observed in autoelectronic discharges (atmospheric sparks, autogenous PAGD (pulsed abnormal glow discharge) and vacuum arc discharges) were exclusive to an autoelectronic emission mechanism prompted by a direct potential between discharging electrodes. Sparking driven by AC potentials could sustain the same forces, but their mutual cancellation over time would not deploy a net force. In this sense, when a large gold leaf connected directly to the ground (via a water pipe or any other suitable connection) or to another large area plate suspended at some height above the ground, is vertically placed at a sparking distance above the surface of another plate connected to the secondary of a Tesla coil, one would not expect the AC spark to sustain any net force across the gap between the gold leaf and the plate. In terms of cathode reaction forces, one would expect their cancellation to be simply brought about by the high frequency of the current alternation in the coil, as both leaf and plate would alternate between being the emitting cathode or the receiving anode. However, this is not what is observed - instead, the gold leaf **16** lifts away from the plate **18** (**Fig.2**). If instead, the suspended gold leaf is connected to the coil terminal, and the bottom plate is connected to the ground in the same manner as described above, this also yields the same result.

Even more curious is the finding that this anomalous reaction force deployed by an alternate current of mass-bound charges in the arc, remains present when the sparking is prevented and instead the corona effect is enhanced (by employing a large plate connected to the outer pole of the secondary, and by employing a distance at which sparking ceases), as if the lift itself were the property of the corona underlying the spark channels and not the property per se of the autoelectronic emission mechanism.

By mounting the suspended leaf **16** (41 mg of hammered 99.9996% pure gold) directly at the end of a long dielectric rod **20** balanced at the centre and placed on a light stand over an electronic balance **22**, we sought to determine the observed lift of the leaf as weight lost. Surprisingly, and despite the most apparent lifting motion of the leaf, the balance registered a substantial weight gain, indicating the addition of 1 to 5 mg weight (with the same 14W input to the vibrator stage), independently of whether the leaf was connected to the terminal of the coil or instead to the earth ground via a water pipe. This suggested to us that, whether formed as a DC or AC spark channel, or whether in the form of a corona discharge, the electric gap develops an expansion force (exactly

opposite to a Casimir force) on both electrodes, independently of their polarity, which force is responsible for the observed repulsion. Yet, this expansion goes hand in hand with an increase in their weight such that some other process is at work in that electric gap.

To examine this problem further, we assembled a different experiment where the gold leaf **16** was suspended between two large metal plates **18** and **24** placed 20 cm apart, and the leaf was not electrically connected to them or to any other circuit, while attached to the dielectric rod employed to suspend it over the electronic balance. Given that the leaf is suitably and equally spaced from both plates, there is no arcing between it and either plate. The obvious expectation is that, since the electric field bathing the leaf alternates at high frequency (measured in hundreds of kilohertz), and the corona from both electrodes should equalise and balance any electric wind, no lift should be observed. In fact, no lift is apparent, but a most curious observation is made: depending upon which orientation is employed for the plates, the gold leaf either gains or loses 4-6% of its weight. This gain or loss is registered for as long as the coil is on. If the top plate is grounded and the bottom one connected to the different terminal of the secondary, a gain in weight is observed (**Fig.3**). If the connections are reversed, an equal weight loss is registered (**Fig.4**).

Furthermore, in this last instance, if the grounded plate **24** is entirely removed (**Fig.5**), and only the top plate remains connected to the outer terminal of the secondary, the observed loss of weight continues to occur such that in effect, this reaction can be obtained with unipolar electric fields of high frequency, and it provides a unidirectional force which, once exerted upon metallic objects bathed by its field, can be made to oppose or augment gravity.

Now, these effects can be greatly magnified, in the order of 10-fold, if the same gold leaf is made part of a simple series floating electric circuit where the leaf functions as a large area plate, and is wired in series with a coil **26** which, for best results, should be wound so as to be of a length resonant with the secondary of the Tesla-type coil employed; and this coil is connected in turn to a point antenna **28** upwardly oriented (**Fig.6**). The entire floating circuit is mounted on the rod **20** and this in turn, is mounted over the sensitive balance. If both plates are kept as in **Fig.3** and **Fig.4**, the observed weight loss and weight gain both vary between 30% and 95% of the total weight of the leaf. Again, the gain or loss of weight is registered for as long as the coil is on.

These anomalous findings suggested that, whatever is the nature of the energy responsible for the force observed in that high frequency alternating current gap, any metallic object placed in that gap will experience a force repelling it from the electric ground. This force will be maximised if the gap frequency is tuned to the elementary or molecular structure of the metallic object. If the electric ground is placed opposite the actual plane of the earth ground, that force will act in the direction of gravity. If, instead, the electric ground and the earth ground are made to coincide on the same plane, that force will act opposite the direction of gravity, i.e. will repel the metallic object from the ground.

No such weight alteration was observed with solid dielectrics, for instance with polyethylene and other thermoplastic sheets.

These facts rule out the possibility of a hidden electrostatic attraction force, acting between the plate connected to the different terminal of the secondary and the gold leaf. Firstly, such an attraction would be able to lift the gold leaf entirely, as is easily observed with the unipole of any electrostatic generator operating with a few milliwatts output with either negative or positive polarity; secondly, the same attraction, if it existed and were the product of an electric force, would surely be manifested independently from whether the experimental leaf was metallic or a dielectric (as again is observed with electrostatic generators).

The results suggest therefore, that whenever a large plate is connected to a Tesla-type coil, it induces in surrounding matter that is not part of its own circuit, a directional thrust which is oriented in a direction which is opposite to the electric ground and, if the electrical ground is on the same side as the surface of the Earth, then a thrust is produced which opposes gravity.

When this thrust is made to oppose gravity, we believe that its effect upon the gold leaf can be compared to the lifting power imparted to the water molecule when it transits from the liquid to the vapour state and which is associated with the increase in internal (or intrinsic) potential “thermal” energy ΔU (See Halliday D & ResnickR (1978) "Physics", Vol. 1, section 22-8, p. 489). The "specific latent heat" of water (m*L) contains indeed both an expression for the sensible radiant thermal work involving volume and pressure relations:
$W = P(V_V-V_L)$ where P = a pressure of 1 atmosphere, and V_V and V_L are the molar volumes in the vapour and liquid phases respectively, and an expression for a quantity of "latent" energy (ΔU) which is associated with the molecule in the more rarefied state. Hence, the relation for the latter with respect to water vapour is: $\Delta U = mL - P(V_V-V_L)$

We propose that likewise, if a very small portion of the energy of the mass-free electric waves is indirectly transformed by mass-bound charge carriers on that plate into blackbody photons (once those charge carriers shed their electrokinetic energy), the greater portion of those waves are directly transformed in the space adjacent to that plate into the latent energy equivalent to ΔU for the atoms of the surrounding air, and so on, until this process itself is also occurring for the atoms of that gold leaf, thus inducing their non-electrical weight loss and suggesting the existence of a non-thermal "antigravitokinetic" energy term previously unknown to mankind other than as "latent heat" or "internal potential energy".

From this viewpoint, the energy released by any Tesla-type coil to its surroundings, would be tantamount to a radiative injection of "internal potential energy" which would confer on local gas molecules a weight cancellation (a cancellation of gravitational mass occurring in the absence of any cancellation of inertial mass - a process which the inventors theorise is explained by the neutralisation of elementary gravitons), and the same process would be equally at work for metallic solids but not dielectric solids.

Gold vapour also deploys a substantial intrinsic potential energy. With an enthalpy of vaporisation on the order of H_V = 324 kJ mol^{-1}, the molar volumetric work performed by gold vapour at atmospheric pressure at the temperature of vaporisation T_V (2,856^{0}C., i.e. 3,129 degrees Kelvin) is:

$W = P\Delta V_{V\text{-}L}$ = 23.58 kJ $mol.^{-1}$ where $\Delta V_{V\text{-}L}$ = 0.2327m^3. The intrinsic potential energy of gold vapour is then given by:

$\Delta U = H_V - W$ = 300.4 kJ $mol.^{-1}$ i.e. 12.74 times greater than the volumetric work performed during the phase transition.

It is our contention that this intrinsic potential energy, associated with molecules as their "latent heat", has fine structure that in turn is altered if this energy is released from these molecules and fails to gain a "sensible" thermal form. What is suggested is that the fine structure of "latent heat" is not electromagnetic and obeys instead the molecular function:

$\Delta U / N_A = \lambda_{n2}^2 c f_{n2}$ where N_A is Avogadro's number, the wavelength denoted as λ_{n2} is the wavelength-equivalent of the mass of the molecule to which the "latent heat" is associated, obtained by a conversion method proposed in these inventors' theory, and the frequency term f is a non-electromagnetic frequency term, specifically in this case a gravitational frequency function.

Employing the conversion of Joules into m^3 sec^{-2} proposed by these inventors as being exactly:

1J = 10 N_A m^3 sec^{-2}, and putting the wavelength λ_{n2} down as the wavelength-equivalent of the mass of the gold atom, λ_{Au}, at 1.9698 m, that frequency term f_{n2} can be obtained as being equal to 2.6 x 10^{-3} sec^{-1}.

According to the present inventors' theory, the wave function c constitutive of the fine structure of "latent heat" associated with molecules of matter, carries the same wavelength λ_{Au} and its frequency is given in the usual manner by c/λ_{Au} = 1.52 x 103 sec^{-1}. The resultant frequency for the non-Planckian unit quantum of "latent energy" associated with each gold atom at the vaporisation temperature is then obtained by the geometric mean of the two synchronous frequency terms: $[(c/\lambda_{Au}) f_{n2}]^{0.5}$ = 624 Hz. However, this is the signature of that intrinsic potential energy when associated with that gold atom at its vaporisation temperature. It is not the signature of the energy quantum itself if it is released from that molecule, nor prior to being absorbed (i.e. in transit), at that same temperature.

The fine structure of the same non-Planckian "latent" energy quantum varies to encompass different determinations of the constituent wavelength and frequency functions. The basic relation for the determination of the wavelength of a "latent thermal" energy quantum not associated with matter, but corresponding to one that is, is:

$\lambda_{n1} = [(\Delta U / N_A) / c]^{0.666}$ $meters^{-0.333}$ $seconds^{0.666}$

which gives 0.046478 m for the unbound equivalent of the "latent heat" unit quantum of vaporisation associated with the gold atom at a pressure of one atmosphere. The fine structure of the free quantum is still parallel, as given by:

$\Delta U / N_A = \lambda_{n1}^2 c f_{n1}$

but now notice how the frequency terms have changed value, with the f_{n1} function having the value 4.65 sec^{-1} and c / λ_{n1} yielding 6.48 x 10^9 sec^{-1}. The geometric mean of the superimposition of the two frequencies is then:

$[(c / \lambda_{n1}^2) f_{n1}]^{0.5}$ = 173.7 KHz

We contend that it is at this frequency that the atoms of gold vapour absorb "latent heat".

However, this is just the overall scenario of what happens at the temperature of vaporisation of gold. But at room temperature (e.g. 293 degrees Kelvin), and with respect to processes where there is no sublimation of the atoms of that gold leaf under way (and indeed, once the coil is turned off, the leaf returns to its normal weight), one must infer to a different phase of matter what portion of "latent heat" energy, if any, do the atoms of gold hold in the solid phase lattice. Assuming the same proportionality between the "sensible" and "latent" thermal energy terms for atoms of gold at room temperature, where the unit thermal energy is $N_A kT$ = 2.436 kJ mol^{-1}, we speculate that the gold atom could absorb up to 12.74 times the value of this "sensible" thermal energy, and thus hold $N_A kT$ = 31.053 kJ more energy in its own micro-atmosphere.

If this speculation is correct, and employing the above novel methodology, then the mean geometric frequency of the maximal "latent heat" energy quantum of a gold atom at room temperature would be 538 KHz (versus 174 KHz at the vaporisation temperature), and once absorbed its mean frequency mode would reduce to 201.5 Hz (versus 630 Hz once the atom has vaporised).

To test this hypothesis, we employed two different Tesla-type coils having output frequencies of 200 KHz and 394 KHz. The circuit tested was that shown in **Fig.6**, and both coils were operated at 50 KV outputs. Whereas the former coil, closer to the 174 KHz marker, could only systematically produce 10mg to 11 mg of weight cancellation in the gold leaf of the floating circuit, the second coil, closer to the speculated 538 KHz marker, could produce 15mg to 35 mg of weight cancellation in the same gold leaf. The empirical results appear therefore to suggest that our speculation may well be a valid one.

The above-mentioned full wave divider (see **Fig.1**) can be easily coupled to our autogenous Pulsed Abnormal Glow Discharge technology as described in our U.S. Pat. No. 5,416,391 to form an alternative source of direct current, ultimately powered by Tesla waves, and such a drive can equally be applied to any other vacuum device that can sustain endogenous oscillatory discharges, whether in the PAGD regime or any other pulsatory regime. For the purposes of experimental and visual determination of power outputs from the divider in question, we have utilised either 2 Torr vacuum tubes operating in the high-current PAGD regime, or 20-100 Torr spark tubes requiring high voltages (2 to 10 KV) for their spark breakdown. As taught in the above US Patent, the output from the full wave voltage divider can be assessed by the energy spent in driving the tube and the motor, whose rotary speed is proportional, within the limits chosen, to the power input.

Two separate sets of experiments presented in Table 1 below, showed that direct connection of the wave divider to the outer terminal of the coil (set constantly at 6 clicks on the vibrator stage in **Fig.1**) or to the same terminal but across a large (2 or 3 square feet) plate **30** that increased the capacitance of the secondary (**Fig.7**), presented the same power output in either case (the effect of the plate is to lower the voltage of the output proportional to the increase in current). A substantial increase in power output through the divider is observed only when an identically wound Tesla coil is connected in reverse (**Fig.8**) with the non-common end of its winding **4** not connected, in order to obtain a condition of resonance, and this observed increase is further augmented by now interposing either of the metal plates **18**, **24** between the two chirally connected and identical coils (**Fig.9**). The increase in plate area appears to have the effect of increasing the output for as long as the plate is isolated between the two chiral image coils. Throughout these experiments, the input power to the vibrator was fixed at 14W (60 Hz AC). [Note: 'Chirality', or 'handedness', is a property of objects which are not symmetrical. Chiral objects have a unique three-dimensional shape and as a result a chiral object and its mirror image are not completely identical - PJK].

TABLE 1

Status	Pulse rate (PPS)	Motor rotation (RPM), M ± SEM
Expt A		
Tesla coil (TC) to divider	2.6	582.5 ± 3.9 (n = 4)
TC to inverted TC, to divider	4.4	621 7.6 (n = 4)
TC to 2 ft^2 plate, to inverted TC, to divider	5	775.25 ± 23.6 (n = 4)
Expt B		
Tesla coil (TC) to divider	2.2	613 ± 5.6 (n = 12)
TC to 3 ft^2 plate, to divider	2.3	605 ± 2.6 (n = 12)
TC to inverted TC, to divider	2.3	722 ± 5.7 (n = 12)
TC to 3 ft^2 plate, to inverted TC, to divider	4.2	877.6 ± 6.5 (n = 12)

In our loss of weight experiments described above, we noted that the phenomenon of weight loss by a metallic body placed in proximity of the coil output continued to be observed when only the plate connected to the distal pole of the secondary was retained. The leaf, although not part of the circuit of the secondary, could however be seen as part of a circuit for the capture of ambient radiant energy, specifically that generated by the coil and, as well, that also possibly picked up, in the process, from other ambient sources. To determine whether the last consideration is a possibility at all, or whether the energy picked up by an analogue of our metallic body or gold leaf in the experiments described above, is entirely a by-product of the energy transmitted by the plate connected to the outer pole of the secondary, we next determined what would happen if the pick-up for the full-wave divider were placed, not at the output from the secondary coil, but from an, in all respects identical, plate (the Receiver plate **R**, as opposed to the Transmitter plate **T**) placed a distance away from, and above, the first one. In other words, the gold leaf is replaced by a receiver plate, and this carries an attached test circuit identical to the test circuit employed to directly assess the coil output.

TABLE 2

Status	T R distance	Pulse rate (PPS)	Motor rotation (RPM), M ± SEM
2 ft^2 plates			
R plate to inverted TC, to divider	3"	6.7	882 ± 17.5 (n = 4)
	4"	8	906 ± 12.1 (n = 4)
	6"	10	936 ± 46.1 (n = 9)
3 ft^2 plates			
TC to T plate, to divider	0	2.3	605 ± 2.6 (n = 12)
R plate to divider	6"	3.3	890.1 ± 3.8 (n = 12)
R plate to inverted TC, to divider	6"	5.1	1009.2 ± 4 (n = 12)
R plate to divider	8"	4.0	783.1 ± 11.3 (n = 12)
R plate to inverted TC, to divider	8"	5.1	1005.7 ± 6 (n = 12)

As shown in Table 2 above, the results of the experiment show that there is no loss of energy picked up at the **R** plate (**Fig.10**) when compared to the most favourable situation involving the plate **30** (**Fig.9**) interposed between the chirally connected coils. This observation is however not always the case. For best results one should employ iron, gold or silver plates placed parallel to the horizon, with the **T** plate underneath the **R** plate. In fact, if one employs instead aluminium plates and suspends these vertically, one can consistently register a loss of output at the divider when changing the divider input from the **T** to the **R** plates.

If however the plate **R** is connected in turn to a second identical coil, also wired in reverse, and this second coil in turn serves as input to the full-wave divider (**Fig.11**), then a most curious occurrence takes place - the power output increases considerably (see Table 2), as if the divider circuit had undergone an energy injection not present at the source. Note that the circuits are in fact resonant, but the energy injection contributing nearly 60-66% (for both plate areas in the previous experiment) of the input that we refer to, is not caused by inductive resonance, since the effect of resonance can be ascribed to the set-up described in **Fig.9**. The distance between the plates, as well as their orientation with respect to the local horizon system of the observer also appear to matter, best results being achieved at optimal distances (e.g. for 2 square feet plates the best gap, at 43% RH and room temperature, was at least 6 inches).

We tested the possibility that environmental heat produced by operation of the coil might be the source of the injected energy, the plate of the second system acting possibly as collector for the heat present in the gap. As it turned out, experiments showed repeatedly that in the gap between the **T** and **R** plates there was no significant thermal radiation propagating between one and the other. The more illustrative experiments are those in which we identified where the sensible thermal energy appears, and which involved coupling two cavities: the Transmitter-Receiver gap between plates **T** and **R**, and a Faraday cage enclosure **34** (see **Fig.12**). The first cavity appears to be much like that of a capacitor: the two identical parallel plates are surrounded by a thick dielectric insulator **32**, and a thermometer **T2** is inserted half-way through it. A thermometer **T1** is also fixed to the **T** plate, to measure it's temperature. The second cavity is a simple insulated metal cage with a thermometer **T3** inserted 2 cm into its top. Some 2-4 cm above the top of the cage there is placed a fourth thermometer **T4**, inside an insulated cylinder.

If the Tesla Coil is a source of thermal energy (e.g. IR radiation, microwaves, etc.) we would expect the **T** plate to be the hottest element from which, by radiation, thermal energy would reach the middle of the first cavity making the next thermometer **T2** second hottest, and that the third thermometer **T3** inside the second cavity, even if it might initially be slightly warmer than the other two, would, over time, become comparatively cooler than either one of the other two thermometers, despite the fact that the rising heat would still be seen to warm it up over time. One would expect a similar outcome for the fourth thermometer **T4**, above the cage. As shown by **Fig.13**, where only the temperature differences (ΔT^0 - $T_C{}^0$) between the experimental thermometers and the control thermometer reading the air temperature $T_C{}^0$ of the laboratory are shown, the surface of the **T** plate warms up by 0.1^0C. at 3 minutes after initiation of the run (closed squares), whereas in the space of the T/R gap a diminutive warming, by 0.05^0C., is registered after 10 minutes (open circles). Conversely, the temperature inside the cage, at the top (shaded circles) rises by 0.1^0C. also by the third minute, and the temperature above the cage itself (shaded squares) rises by a much greater difference of 0.35^0C., which remains stable after the eighth minute.

These results show that it is not sensible heat that radiates from the **T** plate. Instead, some other form of radiation traverses these cavities to generate sensible heat at their metallic boundaries, such that more heat is generated above the **R** plate (inside the cage) and again above the third plate, i.e. above the top of the cage, than is generated in the T/R gap, i.e. near the **T** plate. This clearly shows that the Tesla coil is not a significant source of thermal radiation, and that sensible heat can be detected inside and on top of the Faraday cage only as a further transformation of the radiant energy transmitted across the T/R cavity.

The same experiment also illustrates that, whatever is the nature of the additional environmental energy being injected at the surface of **R** plate (as shown by Table 2 results above), it is most likely not thermal radiation, at least not energy in the form of sensible heat. And whatever is the nature of this ambient radiant energy being mobilised by the electric radiant energy transmitted from the **T** plate, it can produce significant heat inside an enclosure adjacent to plate **R**.

Since we also know experimentally, that this observation of an ambient energy injection at the **R** plate or **R** cage depends upon relative humidity, being most easily observable when the latter is low (<50% Relative Humidity), and being virtually impossible to observe when air is saturated with water vapour, we can infer that water vapour is a good absorber of the electric mass-free radiant energy emitted from the **T** plate. This strongly suggests that this absorption process is tantamount to increasing the potential intrinsic energy ΔU of the water vapour molecules adjacent to the **T** plate. In the absence of significant quantities of water vapour, when the atmosphere is dry, one may speculate that this absorption process is replaced by what one presumes is a parallel process involving the various gaseous molecules of air. However, either because the air molecules involve molecular species that readily give off this potential energy, as one might speculate is the case with molecular oxygen, hydrogen and nitrogen, or because the air molecules absorb far less "latent" energy (as appears to be the case with inert gases), and therefore there is more of it in the molecularly unbound state (as we explicitly propose as a possibility) and thus available for absorption by the appropriately tuned receiver, the increased ΔU of air molecules conferred by the absorption of the mass-free electric radiation in the T/R gap is transferred to the **R** conductor together with the latent energy which those molecules already possessed before entering that gap. Hence the

energy injection and its dependency upon the partial pressure of water vapour, which absconds instead with this "latent" energy and succeeds in withholding it from transmission to the **R** plate.

If the T/R gap can mobilise ambient energy which is neither electromagnetic nor thermal in nature, but which "latent" energy becomes injected into the divider circuit in electric form, the heat (i.e. sensible thermal energy) produced inside and on top of the cage, can also be mobilised electrically as input into the divider circuit. The obvious place to look for the positioning of the cool junction which could convert sensible heat into electrokinetic energy of mass-bound charges is at the top of the cage, where it is warmest (See top curve of **Fig.13** in shaded squares). This is clearly observed from the results shown in Table 3 below, where the initial temperature difference between the top of the box and the **T** plate surface was 0.5^0C., and the top of the box temperature rose by 0.2^0C. after 2.5 minutes when the divider was connected at the junction, versus 0.35^0C. when it was not (and the transmitter coil was on).

TABLE 3

Status	TR distance inches	Pulse rate PPS	Motor rotation RPM, M ± SEM (n = 12)
3 ft^2 plates			
TC to T plate, to divider	NA	4.2	877.6 ± 6.5
R plate to inverted TC, to divider	6"	5.1	1009.2 ± 4
Top of naked R plate/ cage to divider	6"	5.4	1047.1 ± 5.7
Top of insulated R plate/cage exposed to sun, to divider	6"	6.1	1072.4 ± 8.7

For the run performed with the naked **R** cage, the temperature directly above the top of the cage was 24.3^0C., at the outset, versus the control room temperature of 23.9^0C. For the run performed with the insulated **R** cage exposed directly to the sun at midday, on a cool and clear August day, the temperature directly above the top of the cage was 33^0C., versus the control air temperature of 18.4^0C. The temperature of the cool junction at the top of the cage was 31.9^0C. while the run was performed.

It is apparent from the data of Table 3, how a second injection of energy has occurred in the apparatus. If, within the T/R gap, the energy injected appears to be on the order of absorption of "latent heat", at the top of the cage cavity, at the cool junction, the injection is one of radiant "sensible" heat. Moreover, this secondary energy addition could be further enhanced by placing strong insulation around the whole apparatus or the cage itself, and further so, by exposing the whole apparatus to solar radiation.

We next turned our attention to the T/R gap cavity with the intention of determining whether atmospheric conditions or vacua yield the same or different results. We could not, of course, test the same large area plates as have been employed for the studies undertaken at atmospheric pressures. For the present purpose we employed instead large area electrodes (ca 0.2 ft^2) made of high grade stainless steel or even aluminium. Preliminary results showed that these T/R gap tubes, when coupled to the divider circuit, yielded faster pulse rates in the secondary circuit when evacuated than at atmospheric pressure. The strength of the corona discharge also intensified, as it eventually became replaced by a normal glow discharge. For purposes of improved spatial capture of (1) the electric mass-free energy radiated from the **T** electrode and (2) the non-radiant latent thermal energy mobilised by it to be collected electrically at the **R** plate, an axial cylindrical T electrode was inserted inside a larger concentric cylinder or between two common plates of large surface area (e.g. >100 cm^2) functioning as the **R** electrode(s), in a dielectric container suitable for evacuation (glass, polycarbonate), at a typical distance of at least 3 cm between electrodes, and the entire device was tested at different pressures.

The secondary circuit connected downstream from the full-wave divider was as shown in **Fig.14** (employing an autogenous pulsed abnormal glow discharge, or PAGD, converter circuit), with the PAGD reactor **36** set at 10 Torr (in light of the high-voltage input, which varied between 1,500V and 3,200V) and gave the results presented in Table 4 below. We should remark also that these pulses charged the charge pack **CP** through the coupling

capacitors **38**, bridge rectifier **40** and reservoir capacitors **42**, and blocking diodes **44**, as expected from the prior art represented by our patents related to PAGD devices.

TABLE 4

T/R tube Pressure (Torr)	Pulse rate (PPS)
760	0.376
0.025	0.513

The effect of the vacuum in the T/R gap tube seems to be dual. By transforming the corona discharge into a normal glow discharge, it increases the local production of photons (probably associated to the formation and discharge of metastable states in the plasma), and at the same time, increases the pulse rate in the output circuit and thus, in all probability, the energy injected in the T/R gap cavity. But this did not yet permit us to confirm whether or not it is "latent heat" energy of the plasma molecules which is being tapped at the receiver plate, even if it be plausible in principle that plasmas may effect more efficient transfer of "latent heat" to tuned receivers than atmospheric gases.

The vacuum dependency of the pulse rate of the PAGD reactor employed as example in the secondary circuit downstream from the divider is also rather well marked, with the fastest pulse rates being registered at 1 Torr for the sample run shown in Table 5 below.

TABLE 5

T/R tube Pressure (Torr)	Pulse rate (PPS)	PAGD Reactor Pressure (Torr)	Voltage (across divider)
0.025	0.115	90	4.5 kV
0.025	0.1553	75	3.5 kV
0.025	0.183	60	3.3 kV
0.025	0.291	30	
0.025	0.513	15	1.6 kV
0.025	0.602	10	1.4 kV
0.025	2.9	2	0.53 kV
0.025	4.1	1	0.45 kV

It is worth noting here that the illustrated polarity of the wiring of the PAGD reactor tube, as shown in **Fig.14**, is best for purposes of sustaining regular auto-electronic emission at high voltage. The reverse configuration, with the centre electrode negative and the plates positive favours instead heating of the cathode and a lapse into a normal glow discharge.

We tested a similar arrangement to that shown in **Fig.14** above, but with a PAGD motor circuit (see our U.S. Pat. No. 5,416,391). A split-phase motor **44** replaces the rectifier and charge pack, and the PAGD reactor is operated at the same pressure of 15 Torr, as shown in **Fig.15**. The T/R gap tube tested had a longer plate distance (2''), with one plate now functioning as Transmitter and the other as Receiver. Note also the different wiring of the PAGD reactor. The results, as shown below in Table 6, present pulse per second (PPS) and motor revolutions per minute (RPM) curve trends that appear to be analogous and parallel to the well known Paschen curves for breakdown voltage in vacuum - such that the T/R gap performs better either in the atmospheric corona discharge mode, or in the high vacuum normal glow discharge (NGD) mode, than in the low breakdown voltage range of the curve where the discharge forms a narrow channel and takes on the appearance of an "aurora" transitional region discharge (TRD).

TABLE 6

T/R tube Pressure (Torr)	Pulse rate (PPS)	Motor rotation (RPM), M ± SEM (n = 17)	Discharge Type
760	2.8	751.2 ± 7.1	Corona
100	2.1	611.5 ± 5.1	TRD
20	2.4	701.9 ± 4.6	TRD
0.006	2.8	748.4 ± 9.3	NGD
0.003	3.0	819.4 ± 6.3	NGD

These results suggest that plasmas with high lateral dispersion, i.e. formed over large electrode areas (e.g. corona and NGD plasmas) and thus devoid of pinch, are more likely to mobilise electrically, the intrinsic potential energy of the molecular charges than pinch plasmas appear to be able to do (e.g. TRD plasmas). Apparently also, the greater the vacuum drawn from the T/R gap cavity, the more efficient does the transfer of this intrinsic potential energy become, i.e. the mass-bound latent heat, to the electrokinetic energy of the charges circulating in the receiver circuit. At about 0.06 Torr, this transfer in vacuo is comparable to that observed under atmospheric conditions and thus for a much greater density of molecules.

We investigated whether it Is possible to tap the latent heat energy of water molecules. It is possible that in the vapour phase they can effectively hold on to their latent energy - but could they give off some of it once closely packed in liquid phase? To test this hypothesis we immersed the T/R gap in a glass water tank. The motor employed for these tests was a high-speed 2-phase drag-cup motor (see **Fig.18** and associated description), wired in split-phase with two identical phase windings capacitatively balanced, and the galvanised iron plates each had an area of one square foot. The results are shown in Table 7 below, and clearly indicate that it is possible to tap - within the T/R cavity - the `latent heat` of water in the liquid phase. As observed, immersion of the T/R cavity in water increased the motor output speed 22% (12,117 / 9,888) x 100). This corresponds to a 50% increase in power output, from 18W at 9,888 rpm to 27W at 12,117 rpm:

TABLE 7

	Pulse rate PPS	Motor rotation RPM M ± SEM	T/R distance cm
Direct from TC	0.3	8076 ± 89.3	NA
TC to T plate	0.5	9888 ± 78.7	NA
R plate	2.75	12117 ± 29.8	30
R plate	2.9	12203 ± 55.9	60

Thus the use of ion-containing water or other ion-containing aqueous liquid in the cavity promotes long distance propagation and a greater injection of latent and thermal energies in the receiver circuit. Such a result is not achieved if the cavity is filled with deionised water.

The preceding results lead therefore to the design of a presently preferred apparatus, based on these findings, for the conversion of mass-free electric energy, "latent heat" energy and "sensible" heat energy into conventional electric energy, as shown in **Fig.16**, which integrates all of the separate findings and improvements. The winding **6** of the Tesla coil at the bottom is driven in the usual manner employing a vibrator stage **2** to pulse the primary coil **4**. The outer pole of the secondary **6** is then connected to a circular metal plate **T** which is one end of an evacuated cylindrical cavity, connected to a vacuum pump or sealed at a desired pressure, or which forms a still containing water or other aqueous solution or liquid. This cavity constitutes the transmitter/receiver gap, and is therefore bounded by a dielectric envelope and wall structure **32**, with the circular receiver plate **R** as its top surface. In turn this plate **R** serves as the base of a conical Faraday cage **34**, preferably air-tight and at atmospheric pressure, but which could also be subject to evacuation, which conical structure carries at its apex provisions for a cold junction **45** and any possible enhancement of the same junction by surface application of different metallic conductors that may optimise the Peltier-Seebeck effect. The output from the cold junction where sensible thermal energy is added to the electrokinetic energy of charge carriers, is also the input to the distal end of the winding **6** of the chiral coil arrangement that sustains resonant capture of all three energy flows ((1) mass-free electric waves of a longitudinal nature, (2) true "latent heat" or the intrinsic (thermal) potential energy, and (3) the thermokinetic energy of molecules, (i.e. "sensible" heat) and, placed in series with the input of

the full wave divider **8**, **10**, feeds the circuit output from the series capacitors **12**, **14** grounded at their common tap. In the T/R gap, the transmitted electric longitudinal wave energy is captured along with any intrinsic potential energy shed by molecules caught in the field. Within the **R** element, expanded into an enclosure that guides "sensible" radiant heat, the latter is generated and then recaptured at the cold junction.

The apparatus consisting of the cylindrical T/R gap cavity and the contiguous conical cage is then preferably finished in gloss white and cylindrically enveloped within a matt black container **46** by effective thermal insulation **48**, the latter terminating at the height of the bottom disc **T**. Apparatus (not shown) may be provided to move the plate **T** vertically to adjust the T/R gap.

Another alternative embodiment of the apparatus is shown in **Fig.17**. Here the circuit driving the apparatus is as we have set forth in our prior patents, which employs an autogenous pulsed abnormal glow discharge tube **50** in the configuration shown, supplied by a battery pack **DP** through blocking diodes **52** and an RC circuit formed by resistor **54** and capacitor **56** to drive the primary **2** of a first Tesla coil to obtain at the distal pole of the secondary **6** the energy to be injected to plate **T** in the form of a central electrode of a coaxial vacuum chamber (sealed or not), of which the cylindrical metallic envelope forms the receiver plate **R**, the latter being placed centrally inside the conical cage **34** and contiguous with its walls and base. The top and bottom of the coaxial chamber carries suitable insulating discs, preferably with O-ring type fittings. Again, the apparatus is enclosed in insulation within a cylindrical container **46**, and the input into the capture circuit driven from the full wave divider is taken from the cold junction **45** at the apex of the air-tight cage. The output circuit is similar to that of **Fig.15**.

We have found however that even when the component values in the motor driver and motor circuits are carefully selected so that these circuits are co-resonant with the dampened wave (DW) component of the motor driver pulses, the motor power output falls well short of that which should theoretically be attainable. In an endeavour to meet this problem, we replaced the squirrel-cage type induction motor **44** by a drag cup motor of type KS 8624 from Western Electric in the expectation that the low-inertia non-magnetic rotor would allow better response to the Dampened Wave component. This motor is similar to one of the types used by Reich in his experiments. Although results were much improved they still fell short of expectations. Replacement of this motor by an inertially dampened motor of type KS 9303, also from Western Electric, provided much better results as discussed below.

Fundamentally, the difficulties we encountered stemmed from the inability of motor couplings to respond efficiently and smoothly, and at the same time, to the pulse and wave components of Dampened Wave impulses: that is, simultaneously to the high-intensity peak current pulses (the front end event), the DC-like component, and to the dampened wave trains these cause, i.e. the pulse tails (or back end event)-or AC-like component. This difficulty is present even when we just seek to run induction motors from the DW impulses of a Tesla coil, the very difficulty that led Tesla to abandon his project of driving a non-ferromagnetic disc rotor mounted on an iron core bar stator with dampened waves.

We believe that the key to the capture of the mass-free energy flux output in electric form by Tesla transmitters, including any injected latent or thermal energy that have undergone conversion into electrical energy is to employ the tuned, unipolar, Y-fed, PAGD-plasma pulser driven split-phase motor drive we have invented (U.S. Pat. No. 5,416,391) in conjunction with an inertially dampened AC servomotor-generator (see **Fig.18**): this has a motor shaft **64** which couples a drag-cup motor rotor **60**, preferably of aluminium, silver, gold or molybdenum, directly to a drag-cup generator rotor **62** that drives a permanent magnet (PM) flywheel **66**, freely rotatable in bearings **67**, that provides inertial damping. The shaft **64**, journalled by bearings **61** in the casing of the motor **44**, provides a power output through optional gearing **68**. The phase windings of the motor **44** are wound on a stator core **70** having concentric elements between which the rotor or cup **60** rotates. This structure makes it ideal for the capture of the DW impulses, whether sourced in the transmitter, amplified in the T/R cavity or sourced in the plasma pulser, all in synchrony. Effectively the motor couples the damping action of the drag-cup sleeve motor rotor, which action, as we have already found for the KS-8624 motors, is quite effective at absorbing the front-end DC-like event, with the inertial damping of the PM flywheel upon the drag-cup sleeve generator rotor, that in turn is quite efficient at absorbing the back-end AC-like wavetrain event.

The KS-9154 motor used by Reich was not an inertial dampened AC drag-cup servomotor-generator. Had Reich succeeded in overcoming the limitations of his 2-phase OR Motor solution, as we have now shown it is possible to do (by applying the Function Y circuit to the PAGD split-phase motor drive which we invented), his motor would have suffered the same limitations which we encountered with the KS 8624 motor.

Any motor, by itself, has an internal or inherent damping whereby the acceleration only vanishes when the rotor is running at constant speed. For motors which operate on the basis of the drag principle, where the asynchronous slip is actually constitutive of the motor action, by inducing eddy currents in the rotor, the inherent damping is always more pronounced than for other induction motors. The damping or braking torque is produced when a

constant current flows through a rotating drag disc or cup.

Aside from this inherent braking, dampers can also be applied to servo motors to further stabilise their rotation. They absorb energy, and the power output and torque of the motor is thereby reduced. Optimal operation of servo motors requires both rapid response on the part of the rotor to changes in the variable or control phase, and a stable response that is free from oscillation, cogging and overshooting. The rapid response is assured by employing low inertia rotors, such as drag-cups or cast alloy squirrel-cages, and the overshooting and oscillation are reduced to a minimum by damping or a retarding torque that increases with increasing motor speed. Typically, in a viscous-dampened servomotor, the damper is a drag-cup generator mounted rigidly on the shaft of the motor rotor, and the generator drag-cup rotates against the stator field of a static permanent magnet field. The generator develops a retarding torque directly proportional to speed, and the energy absorbed by the damper is proportional to speed squared. The damping can be adjusted and, as it increases, the same amount of input power yields lower torque and motor speeds. Inertial-dampened servo motors differ from viscous dampened motors in that the permanent magnet stator of the drag-cup generator is now mounted in its own bearings, either in the motor shaft or on a separate aligned shaft, forming a high-inertia flywheel.

This means that, whereas the motor rotor always experiences a viscous damping in viscous-dampened servo motors, in inertial-dampened servo motors the drag cup motor rotor only experiences a viscous damping while accelerating the flywheel, with the damping torque always opposing any change in rotor speed. Once the flywheel rotates synchronously with the rotor, all damping ceases. Note that this viscous damping is carried out via the coupling of the drag-cup generator rotor, rigidly affixed to the motor rotor, to the PM flywheel, so that their relative motion generates the viscous torque proportional to the relative velocity. Use of drag-cup sleeve rotors in inertially dampened servo motors was largely supplanted by squirrel-cage rotors once the latter became produced as cast alloy rotors. Since inertially dampened motors can be used in open and closed-loop servo applications, and present better stability - even in the presence of non-linearities - and higher velocity characteristics than other induction motors do (Diamond, A (1965) "Inertially dampened servo motors, performance analysis", Electro-Technology, 7:28-32.), they have been employed in antenna tracking systems, stable inertial-guidance platforms, analogue to digital converters, tachometers and torque tables.

The typical operation of an inertially dampened servomotor is as follows: with the reference phase fully excited, the motor rotor -fixedly linked to the generator rotor, as well as the flywheel - remain immobile; once power is applied to the control phase, the motor rotor immediately responds but the flywheel remains at rest. However, as the drag-cup generator **62** is forced to move through the permanent magnetic field of the flywheel, it creates a drag torque that slows down the attached motor rotor proportionally to the acceleration that it imparts to the flywheel that it now sets into motion, thus creating the viscous damper. As the flywheel accelerates, the relative speed of the motor with respect to the flywheel, as well as the damping torque, decrease until both motor and flywheel rotate synchronously and no damping torque is exercised - at which point the drag on the motor cup exerted by the generator cup is negligible.

The KS-9303 motor is an inertial dampened servomotor but is differentiated with respect to other inertially dampened motors, in that (1) it employs a drag-cup sleeve motor rotor made of aluminium, very much like that of the KS-8624, but with slightly altered dimensions and with a shaft extension for the drag-cup copper generator rotor, and (2) the moving flywheel structure was journalled on a separate, fixed shaft, as already described with reference to **Fig.18**. Now, in principle, even application of minimal damping decreases motor efficiency, resulting in diminished torque and speed. Whether the inertial-dampened motor has a drag-cup rotor, a sleeve rotor or a squirrel-cage rotor, the damping increases the rotor slip. Laithwaite considers drag-cup motors as being "dynamically inferior to their cage counterparts" (Laithwaite, E R (1957) "Induction machines for special purposes", London, England, p. 323). If we now add a viscous damping and retarding torque, we should not be able to get much more than a 55% efficiency in the best of conditions. On the other hand, the inertial damping arrangement described will only abstract or supply energy when the motor rotor is accelerating or decelerating relative to the flywheel.

These drag-cup motors, whether inertially dampened or not, develop a constant torque at constant rpm for a given supply frequency and a suitable phase shift capacitance. For each frequency the motors respond to, there is an optimum resonant split-phase capacitance, but other values nearby are still suited for operation, and for each value of capacitance, there is an optimum frequency to which the motors respond. For example the KS-8624 motor responds best at 450 Hz when a 1 microfarad capacitance is employed, responds best at 250 Hz when a capacitance of 10 microfarads is employed, and responds best at 60 Hz, when a capacitance of 100 microfarads is employed. As the capacitance increases, the resonant CW frequency of the motor is displaced to lower values. If we fix the capacitance at a value (e.g. 10 microfarads) suitable for testing the frequency response at a fixed voltage of 12 VAC, the observed result for both the KS-8624 and KS-9303 motors show a response distribution of the motor rotary velocity that has an identical peak at 250 Hz for both motors, with the response decreasing to zero smoothly on both sides of the peak.

These results indicate that, when wired as a split-phase motor, the motor rotary velocity varies not as a function of voltage or current, but as a function of frequency when the phase-splitting capacitance is fixed within a suitable range, there being an optimum frequency mode for each value of suitable capacitance, with lower values of capacitance favouring higher frequency modes. For a given frequency and capacitance, the motor rotary velocity remains essentially constant and independent from voltage and current input, and thus at a plateau. Torque, in the same circuit arrangement, follows exactly the same pattern as rotary velocity, as a function of input frequency at a fixed potential. Torque is linearly proportional to rpm in these motors when they are split-phase wired, and rpm linearly proportional to CW frequency, which makes them ideal for experimentation and determination of power output computations. Moreover, since these are drag machines, the slip itself determines the rotor currents and these are susceptible to tuning such that their retardation and relative position in the field can find resonant modes for varying CW frequency and capacitance.

In the circuit of **Fig.17** when using the KS 9303 motor, the inertial damping of the flywheel coupling retards the motor rotor currents sufficiently to allow them to build up torque, with the entire motor assembly serving as the preferred sink for all of the energy, mass-free and mass-bound, captured by the receiving coil circuit with a drawing action established by the motor on the circuit, and providing satisfactory absorption by an inertial damper of the combined, synchronised, dampened wave impulses, those occurring at a low frequency as a result of the firing of the PAGD reactor, and those occurring at a higher superimposed frequency -sourced in the transmitter circuit and picked-up by the receiver plate and coil. The action of each DW impulse train itself generates two different events: the DC-like auto-electronic-like discontinuity which sets the motor in motion and initiates the rotor currents, and the AC-like dampened wavetrain which supports the consistency of those rotors. The concentration of current required to kick-start the motor is provided by the DW impulses of the PAGD reactor, whereas, once the motor is in motion, and particularly, once it is stabilised by the flywheel, the cumulative action of the higher frequency DW impulses makes itself felt by accelerating the rotor to an optimum rotary velocity.

For the next series of tests we employed the basic circuit diagram of the improved motor shown in **Fig.19**. The transmission station is the typical Tesla transmitter with a line-fed, 60 Hz vibrator stage. At the line input to the first stage, we place a calibrated AC wattmeter (Weston Model 432), and a Beckman 330B rms ammeter in series with the hot lead, we set the vibrator stage for 41 clicks, consuming between 28.5W and 35W, depending upon circumstances yet to be described. This consumption was confirmed by driving the coil from an inverter powered by a 12 volt battery. The inverter consumes 2.16 watts, and is 90% efficient. The total consumption from the battery was 42 watts (12V at 3.5A); once the 2.16 watts is deducted and the efficiency taken into account, we obtain the same 36W (vibrator stage at max., i.e. 47 clicks, in this experiment). The T/R gap is adjusted to 3", and 2 square foot plates are used. Transmitter and receiver coils are tuned, and so are the plate capacitances, to 250 kHz, also the capacitances of the Function Y circuit connected at the output of the receiving coil.

The rectified voltage and current generated by the transmitter secondary and by the transmitter plate was ascertained with a coil-tuned wave-divider (Function Y) circuit by loading it with different resistive values. The results constitute a measure of the mass-bound electrical power output directly from the transmitter apparatus. The same method was employed to ascertain the voltage, current and power of the mass-bound charges circulating in the receiving plate and coil circuit. The results are shown in Table 8 below:

TABLE 8

Massbound currents rectified by Function Y at the output of the Tesla transmitter, transmitter plate and receiver plate, as a function of the bleeding resistance employed in each of the function Y arms

	VDC (kilovolts)	ADC (amp)	WDC (watts)	R/arm (Mohm)
Direct from 2°	42–50	$3 * 10^{-5}$	1.26–1.5	500
From 2° (T) plate	26	$2 * 10^{-5}$	0.52	500
From 2° (R) plate	15.1	$1.25 * 10^{-5}$	0.189	500
Direct from 2°	20.4	$3.4 * 10^{-4}$	6.936	50
From 2° (T) plate	15.2	$2.4 * 10^{-4}$	3.648	50
From 2° (R) plate	9	$1.2 * 10^{-4}$	1.08	50
Direct from 2°	3.3	$1.75 * 10^{-3}$	5.775	1
From 2° (T) plate	3.5	$2 * 10^{-3}$	7.0	1
From 2° (R) plate	2.95	$1.6 * 10^{-3}$	4.72	1

The results indicate that the highest mass-bound power assembled by the secondary transmitter circuit does not exceed 7 watts - and this is directly output from the secondary **26** when the load is 50 Megohm, or from the transmitter plate when the load is 1 Megohm. The mass-bound electric power emulated by the receiving circuit (plate, coil and Function Y without the plasma pulser circuitry) never exceeds the mass-bound electric power outputted directly by the transmitter, and peaks when the resistive load value (1 Megohm) approaches the pre-breakdown resistance range of the vacuum tube, at 4.72W. These findings then indicate that when the transmitter circuit is consuming a maximum of 35W, a typical output from the secondary of the transmitter is 7W, and at 3" of distance within the proximal field of the latter, the pick-up by a tuned receiver will be of the order of 5W of mass-bound current duplicated within the receiving coil. The loss in the first stage is therefore on the order of sevenfold.

Continuing with the description of the circuit of **Fig.19**, a 128 cm^2 plate area, 6 cm gap PAGD reactor is used, connected as described in our prior art to a high-vacuum rotary pump (Correa, P & Correa, A (1995) "Energy conversion system", U.S. Pat. No. 5,449,989). Pressure readings were obtained with a thermocouple gauge during the operational runs. The KS-9303 motors to be tested are then connected to the PAGD reactor in the usual capacitatively-coupled, inverter fashion described in our prior art (Correa, P & Correa, A (1995) "Electromechanical transduction of plasma pulses", U.S. Pat. No 5,416.391). Their rpm is detected by a stroboscopic tachometer and fed to a Mac Performa 6400 running a motor algorithm program calculating the power output. Motor measurements were made at five minutes into each run for the unloaded motors, and at ten minutes for the inertially dampened motors.

All experiments were carried out in the same work session. The experimental determination of the continuous rotary power output as a function of the reactor pulse rate confirmed that the improved circuit develops maximum rotary capture of the mass-free energy in the receiver circuit at the lowest rates of pulsation, just as we have previously found for the conversion system of U.S. Pat. No. 5,449,989. Furthermore, the data showed that even motors of type KS-8624 are able to output power mechanically in excess of the mass-bound power output by the transmitter (7W) or captured by the receiver (5 to a max. of 7W), once the PAGD rate decreases to 1.5 PPS. Such an anomaly can only be explained by the system having become able to begin capturing the mass-free energy flux in the receiver circuit that we know already is output by the transmitter circuit. But this excess mechanical power is still less than the power input into the transmitter, and clearly so. It represents a power gain with respect to the secondary, but a loss with respect to the primary. The full breadth of the capture of the mass-free electric energy flux circulating in the receiver circuit is not seen until the motors are resonantly loaded because they are inertially dampened.

The KS-9303 motors, once inertially dampened, and thus loaded, are able to recover enough power from the mass-free energy field to develop a mechanical power, not just greatly in excess of the mass-bound power of the secondary, but also greatly in excess of the mass-bound power input to the vibrator stage and the primary, at 28 to 35W. Once the pulse rate approaches the same 1.5 PPS marker, mechanical power in excess of the mass-bound electric power input to the primary becomes evident, peaking at nearly three times that input. In fact, the highest output recorded was also obtained with the lowest input to the transmitter circuit, the highest exact coefficient observed in this experiment being 100.8W / 28W = 3.6. Furthermore, with respect to the secondary mass-bound output, the same mechanical rotary output represents a much greater overunity coefficient of performance, on the order of 14.4 times greater. This is at least partly the result of the receiver and motor capture of the mass-free electric energy output by the transmitter, and may be partly the result of mass-free energy engrafted by the PAGD regime in the PAGD reactor.

Reviewing the mechanical power output results as a function of increasing vacuum in the PAGD reactor and at different output power levels, any motor performance below the 5-7W limit of the traditional mass-bound output power of the secondary represents an output mechanical power loss with respect to both the mass-bound secondary output and the mass-bound primary input. All the results for pressures down to 0.03 Torr fall into this category, and thus represent a very inefficient coupling to the PAGD regime. Any motor performance between 7W and 28-35W represent a loss with respect to the electrical power input to the transmitter system, but a net gain of power with respect to the mass-bound secondary power output. None of the non-inertially dampened motors tested were able to perform outside of this area, under the test conditions. With more efficient primary to secondary couplings in the transmitter station, however, one could advantageously employ these motors alone to extract some of the mass-free power of the secondary or to operate them in enclosed vessels without conventional external electrical connections.

To reach satisfactory levels of recovery of mass-free energy, one must dampen the superimposed DW impulses. Hence, all results showing outputs in excess of 35W were obtained using the inertially dampened KS-9303 motors, and represent a net overunity power gain over both the power input to the primary and the mass-bound power output by the secondary, or the mass-bound power emulated by the receiver circuitry. This happens when the PAGD pulse rate falls to 2 PPS, with the rotary power output steeply increasing as the rate falls to 1 PPS.

One of the interesting features of the motor circuitry we have proposed is that it can operate with pulsed plasmas in both the TRD and the AGD regions, the least efficient response occurring in the NGD region near the Paschen minimum. One might think that the voltage depression would allow increased current intensity supplied to the motors, but in fact that is not observed, with the flashing of the NGD yielding erratic oscillations and low values of current. In keeping with the notion that the TRD plasma is mainly composed of lagging positive ions, whereas the PAGD plasma is mostly an electron plasma, the observed direction of rotation of the motors is opposite in the TRD region to that of the AGD region. The NGD region therefore marks the depression where the velocity vectors change direction. In the second or PAGD region, motor operation is very quiet, unlike what is observed in the TRD region.

Part and parcel of the tuning of the circuit components is the selection of the optimum capacitances employed to couple the PAGD reactor to the motor circuit and split the phase to feed the auxiliary winding of the motor. We have experimented with capacitances ranging from 0.5 to 100 microfarads, and found that best results (for the specific circuit in question - including the characteristics of the transmission), were such that the optimum value of the PAGD coupling capacitance lay near 4 microfarads, and the phase splitting capacitance, near 1 to 4 microfarads, depending upon weather conditions. In good weather days lower capacitance values can be used, while in bad weather days higher capacitances are needed. For ease of comparison in demonstrating the need to tune the circuit by employing optimum capacitances in those two couplings (reactor to motor, and motor phase coupling), we employed the same capacitances in both circuit locations.

A comparison of tests using 1 and 4 microfarad values shows the difference caused by changing those capacitances from their optimum value: across all discharge regions of the pressure range that was examined, the four motors tested, operated with greater motor speeds when the capacitances are set to 4 microfarads rather than to 1 microfarad. The less efficient performance obtained with 1 microfarad capacitance fits the inverse correlation of pulse power with increasing pulse frequency, such as we have found for the PAGD regime. This is made evident by a comparison of rpm versus pulse rate for the two capacitance values being considered. They demonstrate the higher pulse rates observed with the lower capacitance, that correlate with the lower motor speeds, and result in lower efficiency of the motor response. The results equally indicate that low capacitance values increase the pulse rate, but if this increase is out of tune with the rest of the circuit values, it results in power waste because it imposes a rate that is not optimum.

We have also determined experimentally that the efficiency of the system is affected by external weather conditions, higher efficiencies being noted on a fine bright day than under poor weather conditions even though the apparatus is not exposed to such conditions. This may reflect a diminution under poor weather conditions of latent mass-free energy that can be taken up by the system.

The observed high efficiency of circuits including inertially dampened motors indicates that the phenomenon does not reduce to a mere optimum capture of, DC-like pulses produced by the reactor in what is essentially an AC motor circuit. Effectively, the pulsed plasma discharge deploys a front-end, DC-like pulse, or discontinuity, but this is followed by an AC-like dampened wave of a characteristic frequency (having a half-cycle periodicity identical to that of the front-end pulse) to which the motor circuit also responds. Moreover, the mass-free electric radiation from the transmitter circuit itself induces, in the receiver antenna, coil and circuit, and in the reactor discharge itself, the train of finer dampened wave impulses responsible, after conversion through the wave-divider, for the mass-bound rectified current which is employed to charge the plasma reactor to begin with. Serving as trigger of the plasma discharges in the reactor are the DW impulses circulating in the receiver circuit, such that the two different lines of DW impulses, in the receiver circuit (for example 120 PPS for the pulses and 154 kHz for the waves) and from the reactor, are synchronised by interpolated coincidences, since their pulse and wave frequencies are different. Ideally, these two superimposed DW frequencies are harmonics or made identical. The receiver stage involves capture of the mass-free electric energy received from the transmitter, duplication of the mass-bound current in the receiver coil, and injection of latent and sensible thermal energy in the T/R gap cavity which augments the emulated mass-bound current.

The mass-bound current is employed to charge the wave-divider capacitance bridge and therefore the reactor. In turn, the plasma pulses from the reactor are superimposed with the DW impulses from the receiving coil, and together they are coupled to the split-phase motor drive. Hence the first receiver stage employs the totality of the energy captured in the T/R gap cavity - mass-free electric energy transmitted by the **T** plate, latent and sensible thermal energy injected at the surface of the **R** plate - and produces in the receiving coil a mass-bound current comparable to that assembled in the transmitter coil by the action of the primary. The mass-bound current is stored in the wave-divider bridge and used to drive the plasma reactor in the PAGD region. Subsequently, the autogenous disruptive discharge that employs a substantial electron plasma generates both a concentrated, intense flux of mass-bound charges in the output circuit, and a mass-free oscillation of its own. The dampened motor is therefore fed directly with (1) the intense mass-bound current output from the reactor; (2) the pulse and wave components of the mass-free electric energy captured by the receiver plate and coil (and matched by conduction through the earth), and which are gated through the wave-divider and the reactor for the duration of the PAGD channel; and (3) any mass-free latent energy taken up from the vacuum by the PAGD event. Once the

motor is set into motion, and is resonantly loaded with an inertial damper, we believe that it will also respond to the much weaker DW impulses captured by the receiver, since these impulses encompass both a DC-like front end - further enhanced by analytic separation through the wave-divider - and a dampened wave at 154 kHz.

Essentially, the DW impulses that are ultimately sourced in the transmitter - and received unipolarly through the T/R gap - have sufficient DC-like potential (plus all the other requisite physical characteristics, such as frequency) to contribute directly to the motor response, once the motor has gained substantial speed (for they lack the current to set it into motion, one of the contributions from the plasma pulser). This is the case, provided that the motor itself is suited for absorption of both DC-like pulses and AC-like dampened waves, which is precisely the case with motors of the type shown in **Fig.18** since the inertia of the flywheel is overcome through homopolar absorption of the dampened oscillations simultaneously in the motor drag-cup rotor and in the generator drag-cup rotor.

We also tested these inertially dampened motors in the traditional DC power supply-driven PAGD circuit we have taught in our previous patents, that is, circuits with an overt HV DC power source, and thus in the absence of any Function Y circuit or transmitter circuit. Here then, only the DW impulses generated by the PAGD reactor can account for the motor response. The tube employed (A31) had an area of 256 cm^2, and a gap distance of 4 cm. Coupling capacitances employed were 4 microfarads for the inverter coupling, and 1 microfarad for the split phase motor coupling. The DC power supply delivered up to 1 ampere of current between 150 and 1,000 VDC, and the ballast resistor was adjusted to 215 ohms. Having determined the basic physical characteristics of the reactor's behaviour in the circuit under consideration, we conducted our experiment in the PAGD region. We chose a pressure of 0.6 Torr, just off from the Paschen minimum, as we intended to benefit from the lower sustaining voltage which it affords.

The experiment basically consisted of increasing the sustaining voltage at this fixed pressure in the PAGD regime, and measuring the diverse physical parameters of the circuit and motor response in order to ultimately ascertain the difference between the input electric DC power and the output mechanical rotary power. We first looked at how the motor rpm response varied as a function of the sustaining voltage (V_S): the results illustrate the importance of starting close to the Paschen minimum in the pressure scale, since the KS-9303 motors reach plateau response (at 17,000 rpm) when the reactor output voltage nears 450V. Any further increase in potential is simply wasted. Likewise, the same happened when we measured motor speed as a function of increasing peak DC current, plateau response being reached at 0.1 ADC. Again, any further increase in current is wasted. Essentially then, the optimal power input to the reactor when the output of the latter is coupled to the motor, lies around 45 watts. This is a typical expenditure in driving a PAGD reactor. As for pulse rate we once again find a motor response that is frequency proportional in the low frequency range, between 10 and 40 PPS (all pulse rates now refer solely to PAGDs per sec), but once rates of >40 PPS are reached, the response of the motor also reaches a plateau.

The observed increment in speed from 40 to 60 PPS translates only into an increase of 1,000 RPM, from 16,000 to 17,000 RPM. So, we can place the optimal PAGD rate at ca 40 PPS. The DC electric power input to drive the PAGD reactor was next compared to the rotary mechanical power output by the inertially loaded motor, driven in turn by the reactor. This comparison was first carried out with respect to the PAGD rates. The motor response far exceeds the conventional input power, indicating that the whole system can be tuned to resonance such that optimal power capture inside the reactor takes place, the critical limit rate lying at around 60 PPS, when the motor response is firmly within the pulse response plateau. At this juncture, the break-even efficiency for the measured rates of energy flux over time reach 700% (overunity coefficient of 7), in keeping with the observations and the values we have made in the PAGD conversion system. In the proportional part of the curve, before the plateau is reached, even greater rates of break-even efficiency - up to >1,000% were registered.

These results constitute the first time we have been able to confirm the presence of output energy in excess of break-even over conventional mass-bound energy input in the PAGD inverter system, and the results are comparable to what we have observed and previously reported for the PAGD converter system. At pulse rates greater than 60 PPS a greater input power results in decreased efficiency, also translated into a noticeable heating of the reactor and motor. And this is all the more remarkable as experiments we have conducted with inductive tuning of PAGD reactors, or employing PAGD reactors as replacements for the primaries of Tesla coil assemblies, and still, more recently, with the PAGD inverter circuit driving motors, have all shown that it is possible to operate these reactors with minimal mirroring and heating, preserving essentially the cold-cathode conditions and yet focusing the plasma column so that deposition on the insulator is negligible. It appears that above a certain threshold of optimal efficiency, surplus input energy is just dissipated thermally by both the reactor and the motors.

It should be understood that the above described embodiments are merely exemplary of our invention, and are, with the exception of the embodiments of **Figs. 16 to 19** designed primarily to verify aspects of the basis of the invention. It should also be understood that in each of these embodiments, the transmitter portion may be omitted

if an external or natural source of Tesla waves is available, provided that the receiver is tuned to the mass-free radiation mode of the source. For example if solar radiation is available in which the mass-free component has not interacted with the earth's atmosphere (as in space applications), the receiver is tuned to the voltage wave of the mass-free radiation sourced in the sun, e.g. by using a Tesla coil in the receiver constructed to have an appropriate voltage wave close to the 51.1 kV characteristic of such radiation.

CLAIMS

1. A device for the conversion of mass-free radiation into electrical or electrokinetic energy comprising a transmitter of mass-free electrical radiation having a dampened wave component, a receiver of such radiation tuned to resonance with the dampened wave frequency of the transmitter, a co-resonant output circuit coupled into and extracting electrical or electrokinetic energy from the receiver, and at least one of a transmission cavity between the transmitter and the receiver, a full-wave rectifier in the co-resonant output circuit, and an oscillatory pulsed glow discharge device incorporated in the co-resonant output circuit.

2. A device according to claim 1, wherein the output circuit comprises a full wave rectifier presenting a capacitance to the receiver.

3. A device according to claim 2, wherein the output circuit comprises an electric motor presenting inductance to the receiver.

4. A device according to claim 3, wherein the motor is a split phase motor.

5. A device according to claim 4, wherein the motor is a drag motor having a non-magnetic conductive rotor.

6. A device according to claim 5, wherein the motor has inertial damping.

7. A device according to claim 6, wherein the motor has a shaft, a drag cup rotor on the shaft, and inertial damping is provided by a further drag cup on the shaft.

8. A device according to claim 6, wherein the transmitter and receiver each comprise at least one of a Tesla coil and an autogenous pulsed abnormal glow discharge device.

9. A device according to claim 8, wherein the transmitter and receiver both comprise Tesla coils, and further including a transmission cavity which comprises spaced plates connected respectively to the distal poles of the secondaries of Tesla coils incorporated in the transmitter and receiver respectively.

10. A device according to claim 9, wherein the plates are parallel.

11. A device according to claim 9, wherein the plates are concentric.

12. A device according to claim 9, wherein at least the receiver comprises a Tesla coil driving a plasma reactor operating In PAGD (pulsed abnormal glow discharge) mode.

13. A device according to claim 1, wherein the transmitter and receiver each comprise at least one of a Tesla coil and an autogenous pulsed abnormal glow discharge device.

14. A device according to claim 12, wherein the transmitter and receiver both comprise Tesla coils, and further Including a transmission cavity which comprises spaced plates connected respectively to the distal poles of the secondaries of Tesla coils incorporated in the transmitter and receiver respectively.

15-17. (cancelled)

18. A device according to claim 1 wherein a transmitter/receiver cavity is present and filled with an aqueous liquid.

19. A device for the conversion of mass-free radiation into electrical or electrokinetic energy comprising a receiver of such radiation from a source of mass-free electrical radiation having a dampened wave component, the receiver being tuned to resonance with the dampened wave frequency of the source, a co-resonant output circuit coupled into and extracting electrical or electrokinetic energy from the receiver, and at least one of a transmission cavity between the source and the receiver, a full-wave rectifier in the co-resonant output circuit, and an oscillatory pulsed glow discharge device incorporated in the co-resonant output circuit.

PAULO and ALEXANDRA CORREA

US Patent 5,449,989 **12th September 1995** **Inventors: Correa, Paulo and Alexandra**

ENERGY CONVERSION SYSTEM

This patent shows a method of extracting environmental energy for practical use. In the extensive test runs, an input of 58 watts produced an output of 400 watts (COP = 6.9). This document is a very slightly re-worded copy of the original.

ABSTRACT

An energy conversion device includes a discharge tube which is operated in a pulsed abnormal glow discharge regime in a double ported circuit. A direct current source connected to an input port provides electrical energy to initiate emission pulses, and a current sink in the form of an electrical energy storage or utilisation device connected to the output port captures at least a substantial proportion of energy released by collapse of the emission pulses.

US Patent References:

3205162	Sep, 1965	MacLean.
3471316	Oct, 1969	Manuel.
3705329	Dec, 1972	Vogeli.
3801202	Apr, 1974	Breaux.
3864640	Feb, 1975	Bennett.
3878429	Apr, 1975	Iwata.
4009416	Feb, 1977	Lowther.
4128788	Dec, 1978	Lowther.
4194239	Mar, 1980	Jayaram et al.
4443739	Apr, 1984	Woldring.
4489269	Dec, 1984	Edling et al.
4527044	Jul, 1985	Bruel et al.
4772816	Sep, 1988	Spence.
4896076	Jan, 1990	Hunter et al.
5126638	Jun, 1992	Dethlefsen.

Other References:

Tanberg, R. "On the Cathode of an Arc Drawn in Vacuum", (1930), Phys. Rev., 35:1080.

Kobel, E. "Pressure & High Vapour Jets at the Cathodes of a Mercury Vacuum Arc", (1930), Phys. Rev., 36:1636.

Aspden, H. (1969) "The Law of Electrodynamics", J. Franklin Inst., 287:179.

Aspden, H. (1983) "Planar Boundaries of the Space-Time Lattice" Lettere Al Nuovo Cimento, vol. 38, No. 7, pp. 243-246.

Aspden, H. (1980) "Physics Unified", Sabberton Publications, pp. 14-17, 42-45, 88-89, 190-193.

Pappas, P. T. (1983) "The Original Ampere Force and Bio-Savart & Lorentz Forces", Il Nuovo Cimento, 76B:189.

Graham, G. M. & Lahoz, D. G. (1980) "Observation of Static Electromagnetic Angular Momentum in Vacuo", Nature, vol. 285, pp. 154 & 155.

Sethlan, J. D. et al., "Anomalous Electron-Ion Energy Transfer in a Relativistic-Electron-Beam-Plasma" Phys. Rev. Letters, vol. 40, No. 7, pp. 451-454 (1978).

REFERENCE TO RELATED APPLICATIONS

This application is a continuation-in-part of U.S. application Ser. No. 07/922,863, filed Jul. 31, 1992 (abandoned), and is also a continuation-in-part of U.S. patent application Ser. No. 07/961,531, filed Oct. 15, 1992, now U.S. Pat. No. 5,416,391.

BACKGROUND OF THE INVENTION

1. Field of the Invention:

This invention relates to energy conversion circuits utilising discharge tubes operating in the pulsed abnormal glow discharge (PAGD) regime.

2. Review of the Art:

Such discharge tubes and circuits incorporating them are described in our co-pending U.S. patent application Ser. Nos. 07/922,863 and 07/961,531. The first of these applications discloses discharge tube constructions particularly suited for PAGD operation, and the second discloses certain practical applications of such tubes, particularly in electric motor control circuits. The review of the art contained in those applications is incorporated here by reference, as is their disclosure and drawings.

It is known that there are anomalous cathode reaction forces associated with the cathodic emissions responsible for vacuum arc discharges, the origin and explanation of which have been the subject of extensive discussion in scientific literature, being related as it is to on-going discussion of the relative merits of the laws of electrodynamics as variedly formulated by Ampere, Biot-Savart and Lorentz. Examples of literature on the subject are referenced later in this application.

SUMMARY OF THE INVENTION

The particular conditions which prevail in a discharge tube operated in the PAGD regime, in which a plasma eruption from the cathode is self-limiting and collapses before completion of a plasma channel to the anode gives rise to transient conditions which favour the exploitation of anomalous cathode reaction forces.

We have found that apparatus utilising discharge tubes operated in a self-sustaining pulsed abnormal glow discharge regime, in a double ported circuit designed so that energy input to the tube utilised to initiate a glow discharge pulse is handled by an input circuit substantially separate from an output circuit receiving energy from the tube during collapse of a pulse, provides valuable energy conversion capabilities.

The invention extends to a method of energy conversion, comprising initiating plasma eruptions from the cathode of a discharge tube operating in a pulsed abnormal glow discharge regime utilising electrical energy from a source in a first circuit connected to said discharge tube, and capturing electrical energy generated by the collapse of such eruptions in a second circuit connected to the discharge tube.

BRIEF DESCRIPTION OF THE DRAWINGS

The invention is described further with reference to the accompanying drawings, in which:

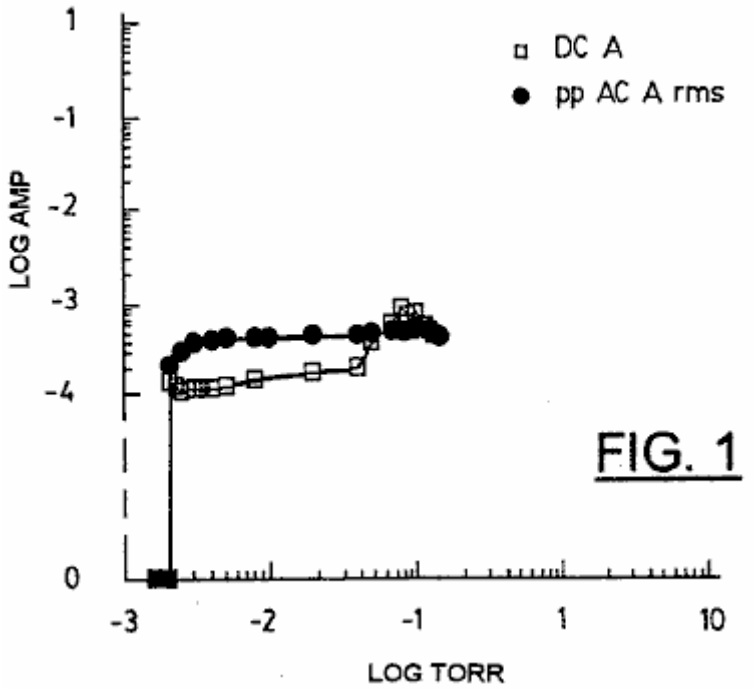


FIG. 1

Fig.1 shows variation of applied DC current and pulse AC rms currents characteristic of a low current PAGD regime, as a function of decreasing pressure, for a 128 cm^2 H34 aluminium plate pulse generator having a 5.5 cm gap length and being operated in the single or plate diode configuration of FIG. 11A, at about 600 V DC.

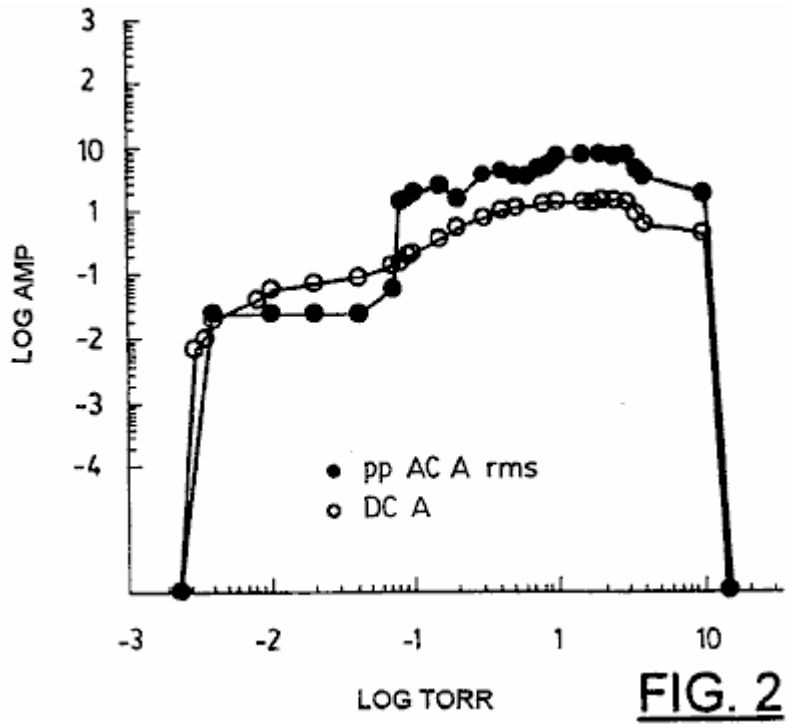


FIG. 2

Fig.2 shows variation of applied DC current and AC rms currents of a high current PAGD regime, as a function of the decreasing pressure, for a device identical to that of Fig.1, and operated at the same potential.

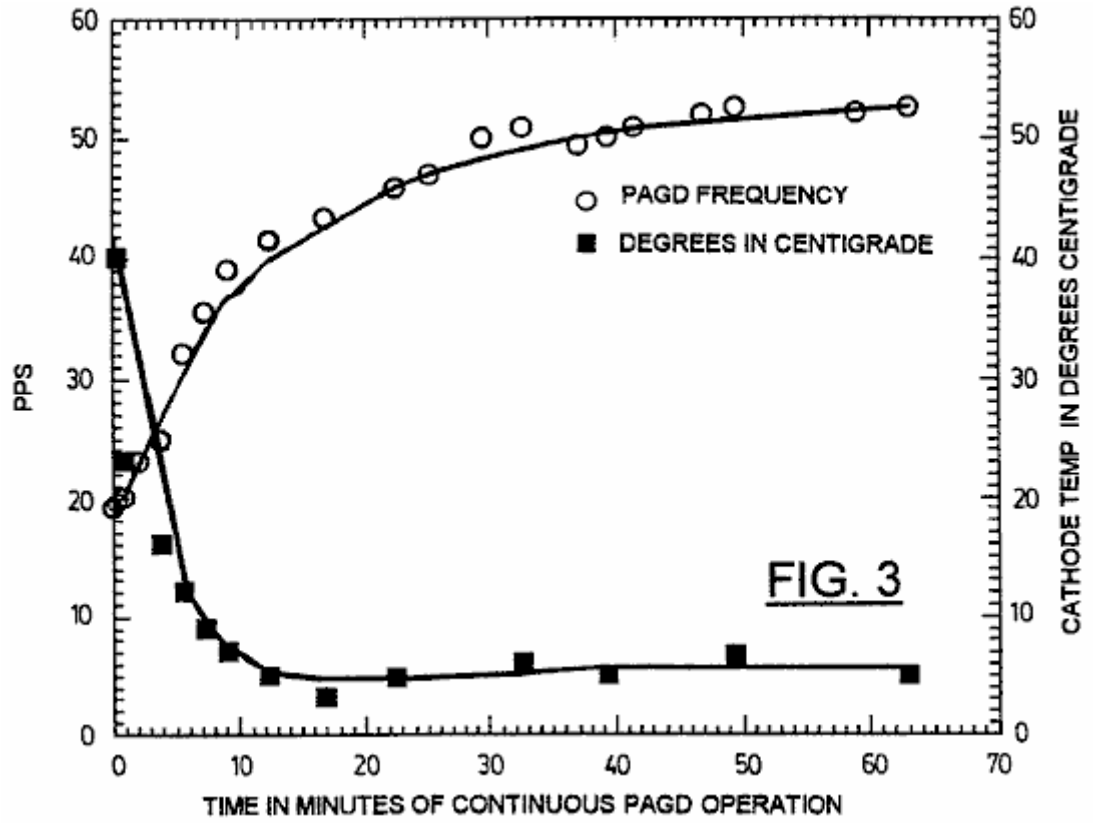


FIG. 3

Fig.3 shows PAGD rate vs pulse generator cathode temperature as a function of the time of continuous PAGD operation, for a pulse generator with 64 cm^2 plates having a 4 cm gap distance, operated at a DC voltage of 555 (av) and R1 = 600 ohms (see Fig.9).

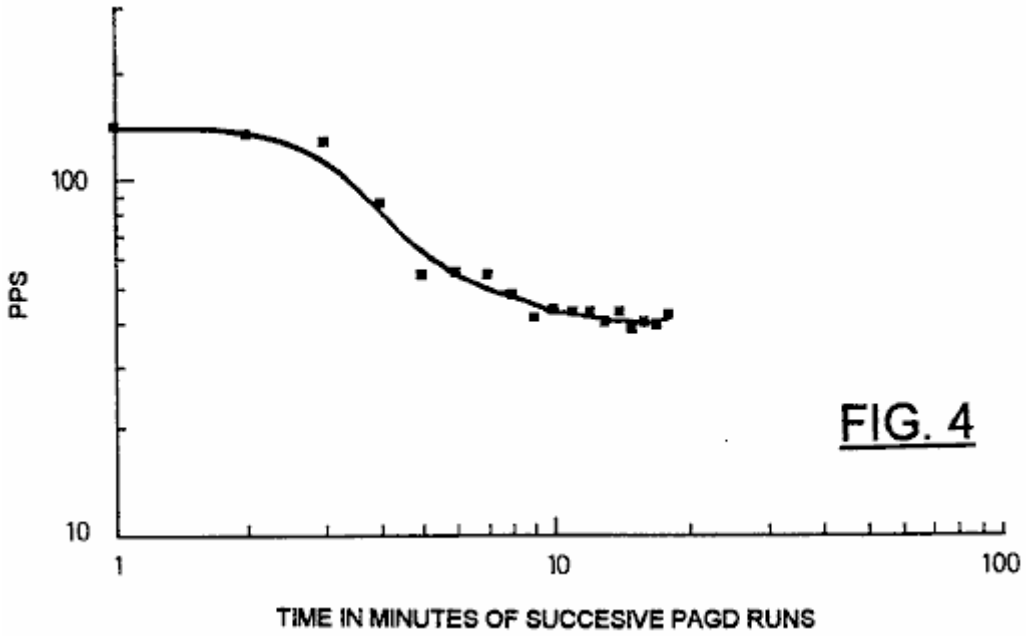


FIG. 4

Fig.4 shows PAGD frequency variation with time, for 18 successive spaced one-minute PAGD runs for a pulse generator with 128 cm^2 plates, and a 5.5 cm gap distance, operated at V DC = 560 (av) and R1 = 300 ohms.

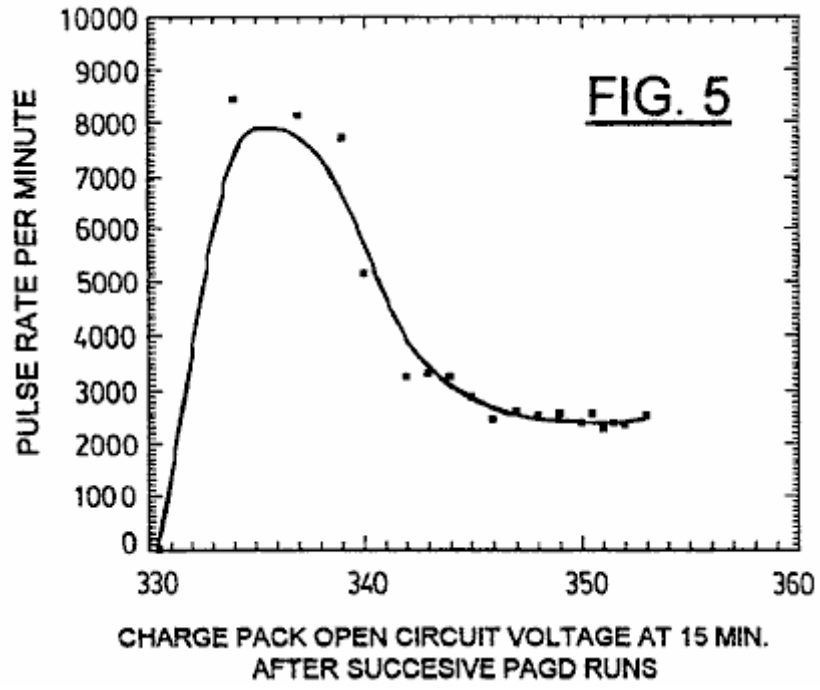


FIG. 5

Fig.5 shows variation of the PAGD frequency in pulses per minute (PPM) with increasing charge of a PAGD recovery charge pack (see Fig.9), as measured in terms of the open circuit voltage following 15 minutes of relaxation after each one minute long PAGD run, repeated 18 times in tandem, under similar conditions to Fig.4.

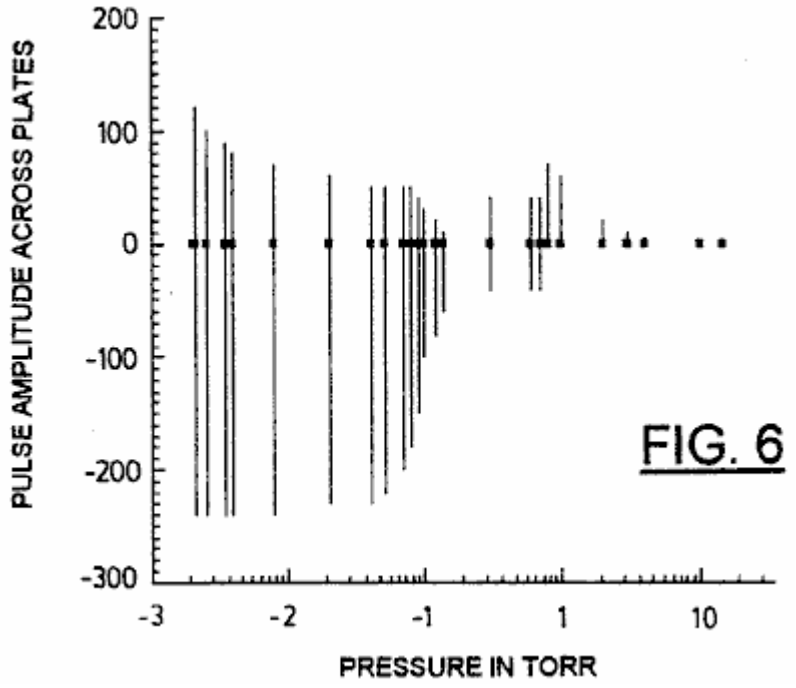


FIG. 6

Fig.6 shows volt amplitude variation of continuous PAGD at low applied current, as a function of decreasing air pressure, for a 128 cm^2 plate area device, gap length = 5 cm; (DC V at breakdown = 860).

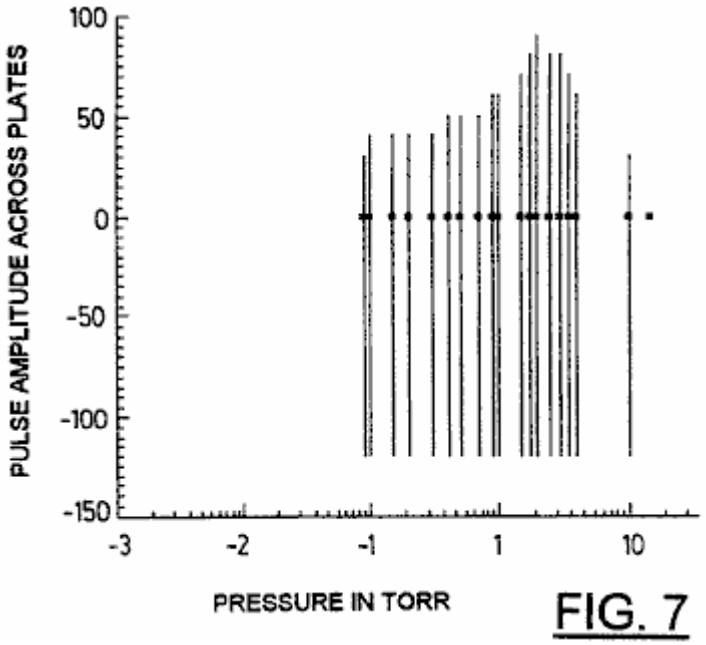


FIG. 7

Fig.7 shows volt amplitude variation of continuous PAGD at high applied current as a function of the decreasing air pressure, for a 128 cm^2 plate area device, gap length = 5 cm; (DC V at breakdown = 860).

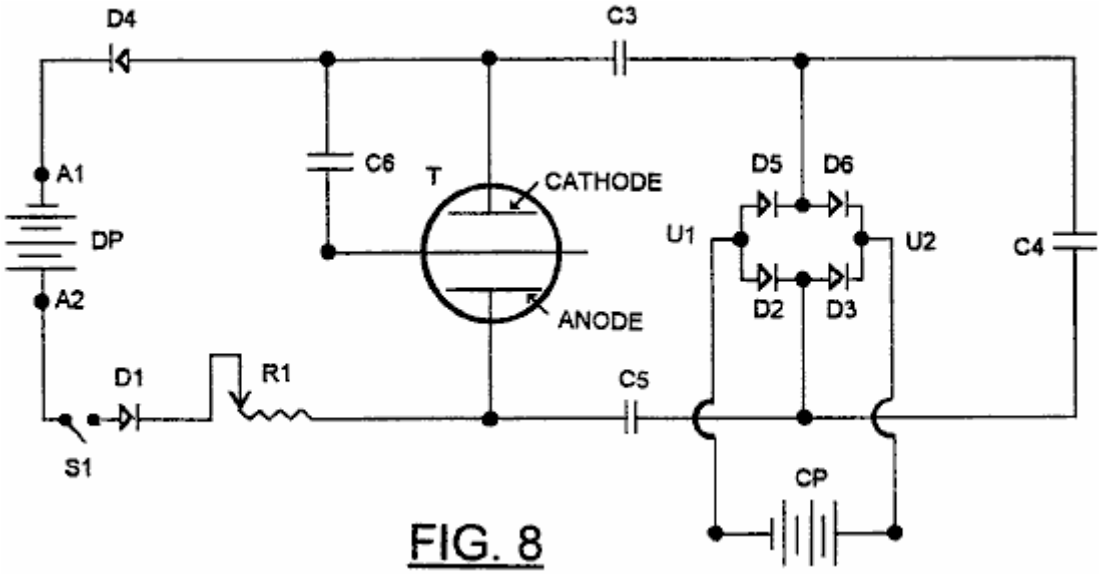


FIG. 8

Fig.8 is a schematic diagram of a first experimental diode (without C6) or triode PAGD circuit.

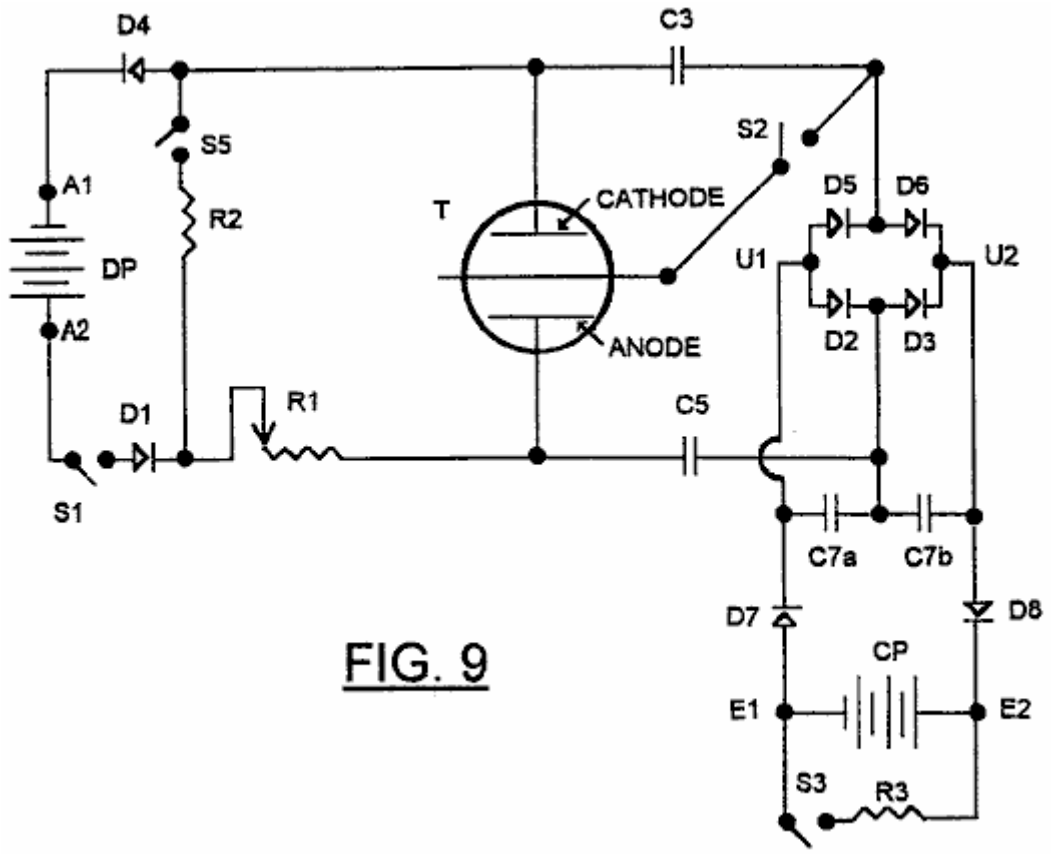


FIG. 9

Fig.9 is a schematic diagram of a preferred diode or triode PAGD circuit in accordance with the invention.

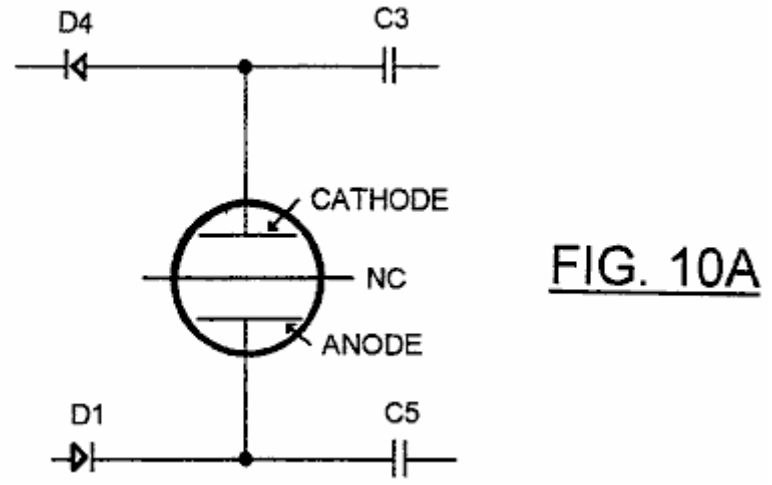


FIG. 10A

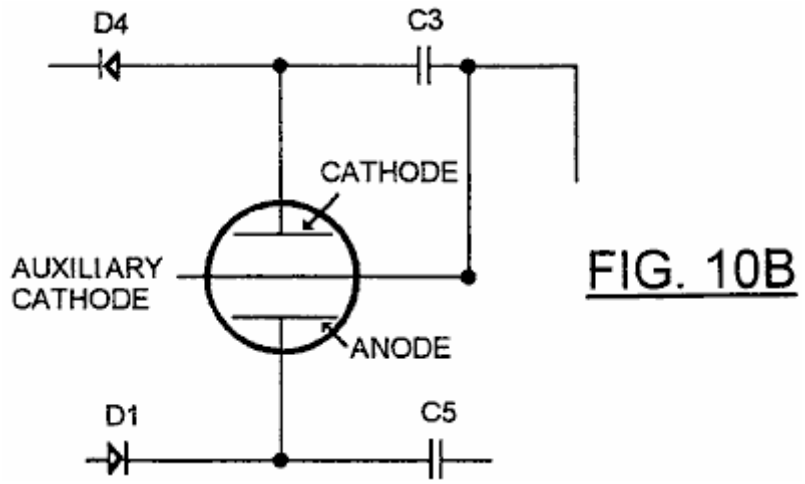


FIG. 10B

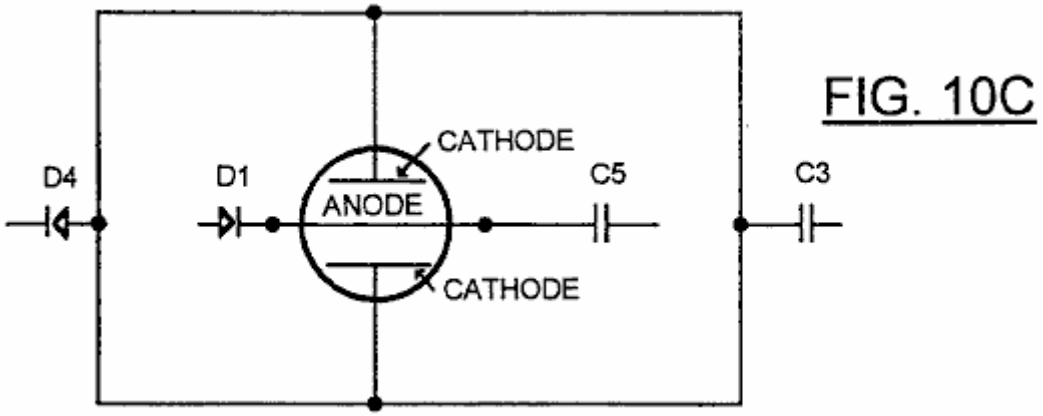


FIG. 10C

Fig.10A, **Fig.10B** and **Fig.10C** are fragmentary schematic diagrams showing variations in the configuration of the circuit of Fig.9.

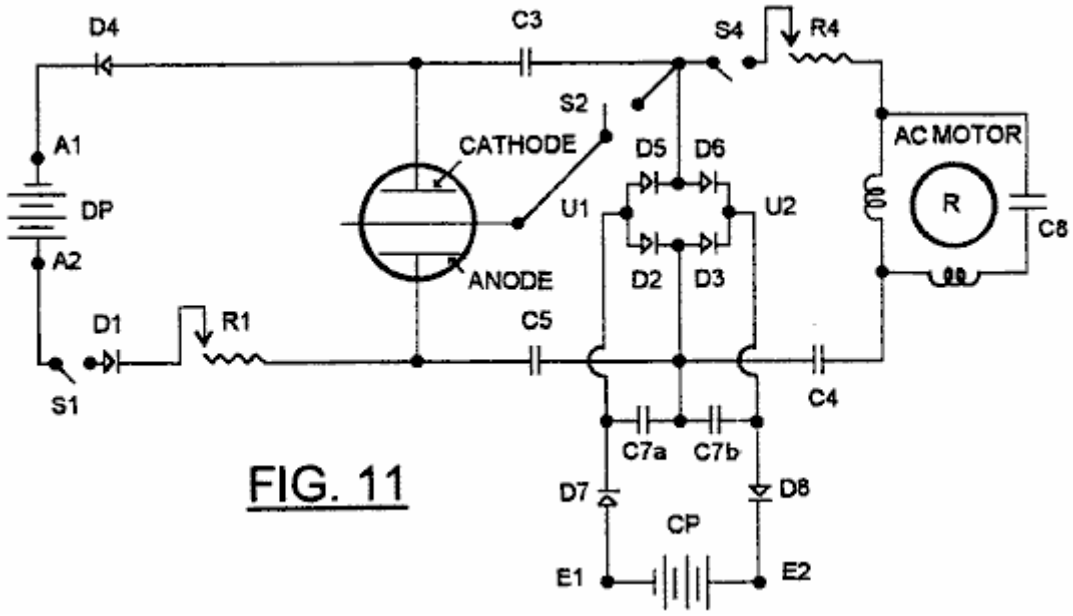


FIG. 11

Fig.11 is a modification of Fig.9, in which an electromagnetic machine, in the form of an electric motor, is connected into the circuit as an accessory electromechanical arm.

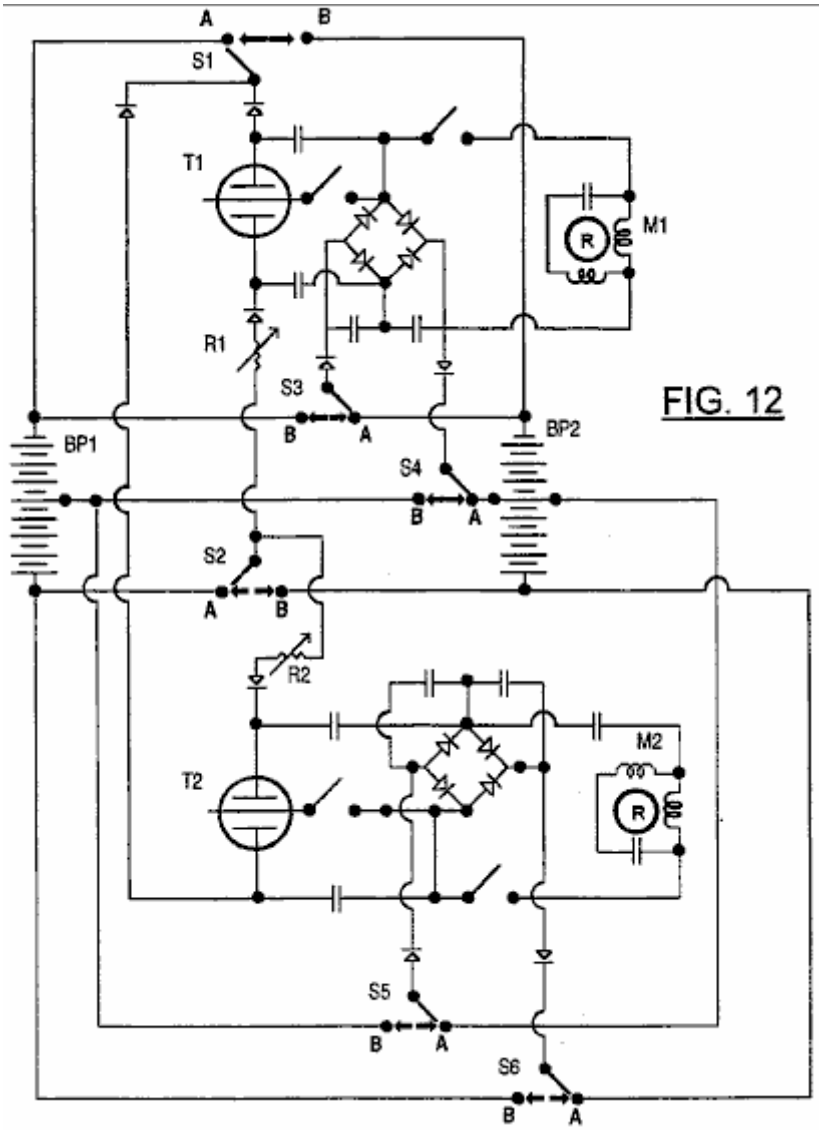


FIG. 12

Fig.12 shows a further development of the circuit of Fig.9, permitting interchange of driver pack and charge pack functions.

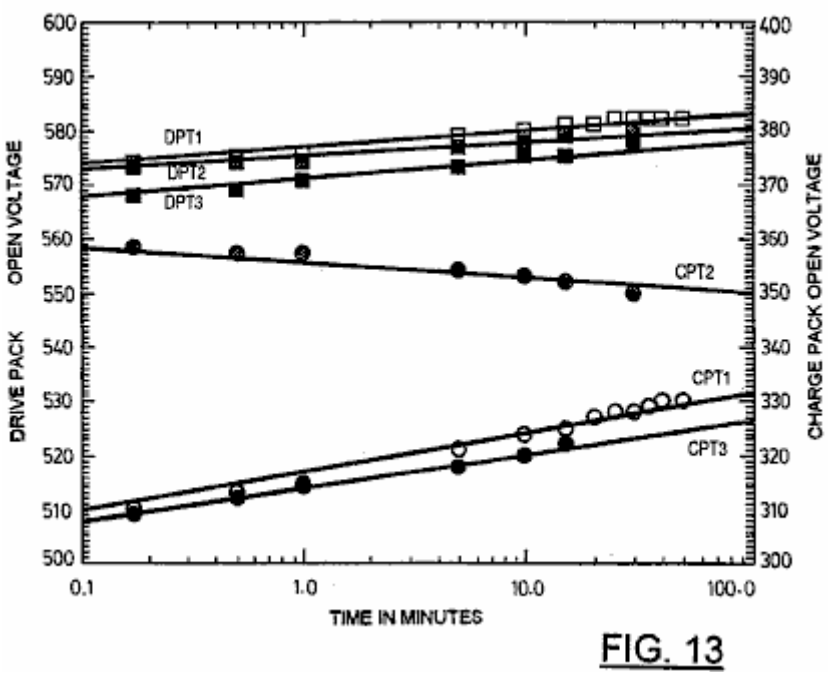


FIG. 13

Fig.13 shows open circuit voltage relaxation curves for battery packs employed in tests of the invention, respectively after pre-PAGD resistive discharge (DPT1 and CPT1), after a PAGD run (DPT2 and CPT2) and after post-PAGD resistive discharge (DPT3 and CPT3).

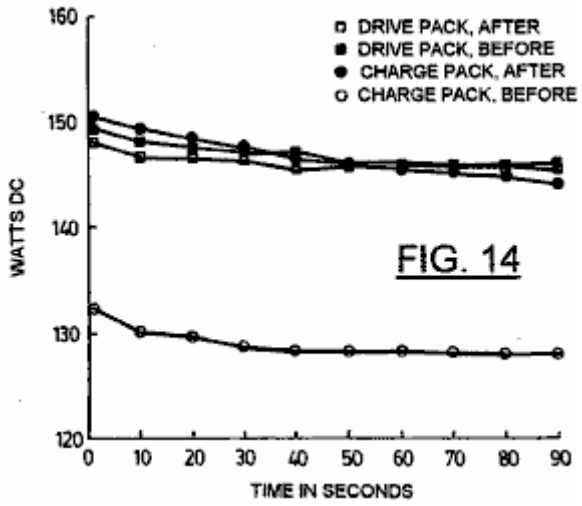
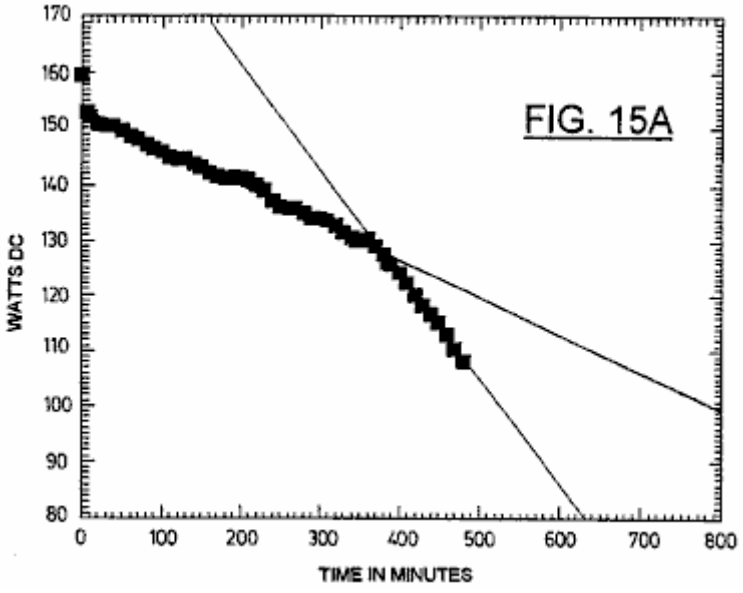


FIG. 14

Fig.14 shows an example of negligible actual power measurements taken immediately before or after a PAGD run, showing both the drive pack loss and the charge pack gain in DC Watts; DP resistance = 2083 ohms; CP resistance = 833 ohms.



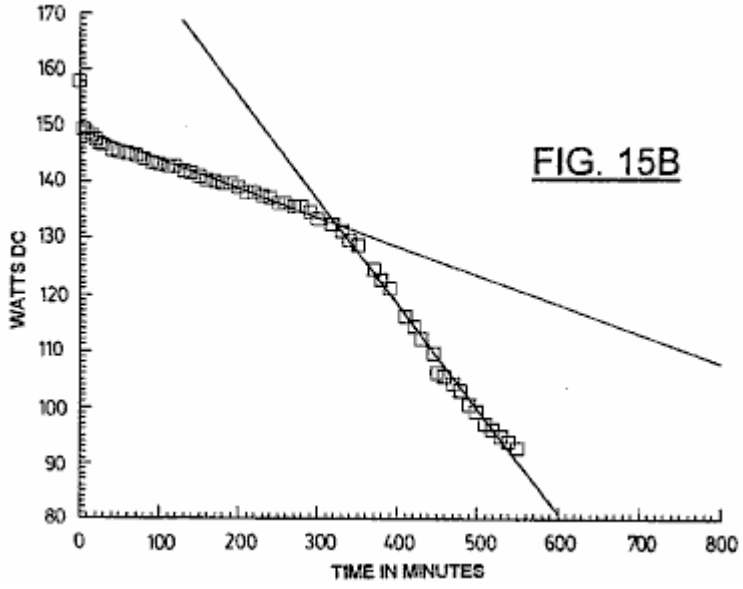


Fig.15A and **Fig.15B** show resistive voltage discharge curves for two separate lead-zero gel-cell packs utilised respectively as the drive and the charge packs; load resistances employed were 2083 ohms across the drive pack (Fig.15A) and 833 ohms across the charge pack (Fig.15B).

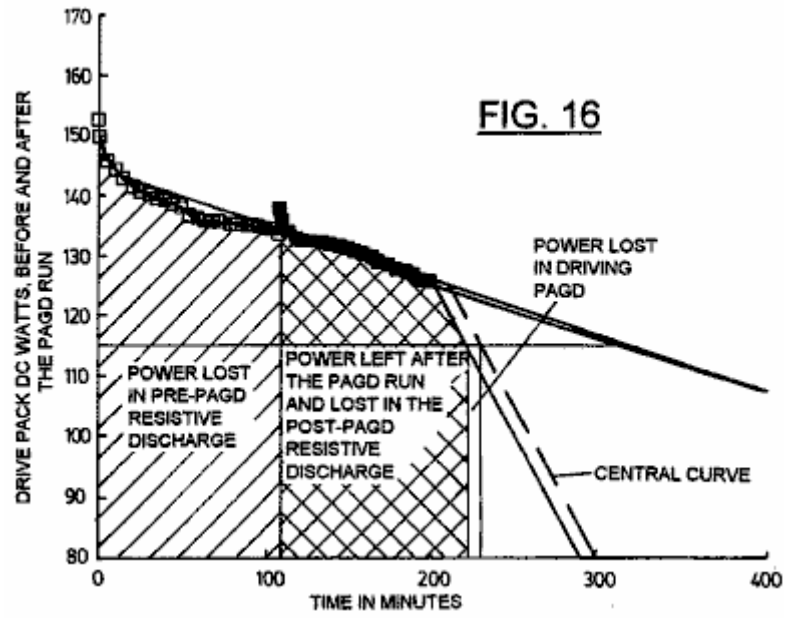


Fig.16 shows resistive discharge slopes for a drive pack before and after a very small expenditure of power in providing energy input to a PAGD run; R = 2083 ohms.

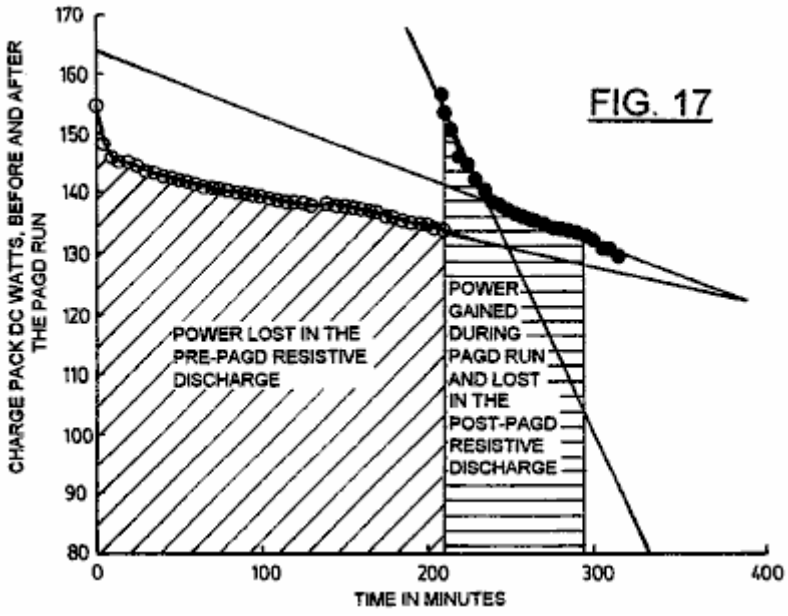


Fig.17 shows resistive discharge slopes for a charge pack before and after capturing energy from the collapse of PAGD pulses in the same test as Fig.15; R = 833 ohms.

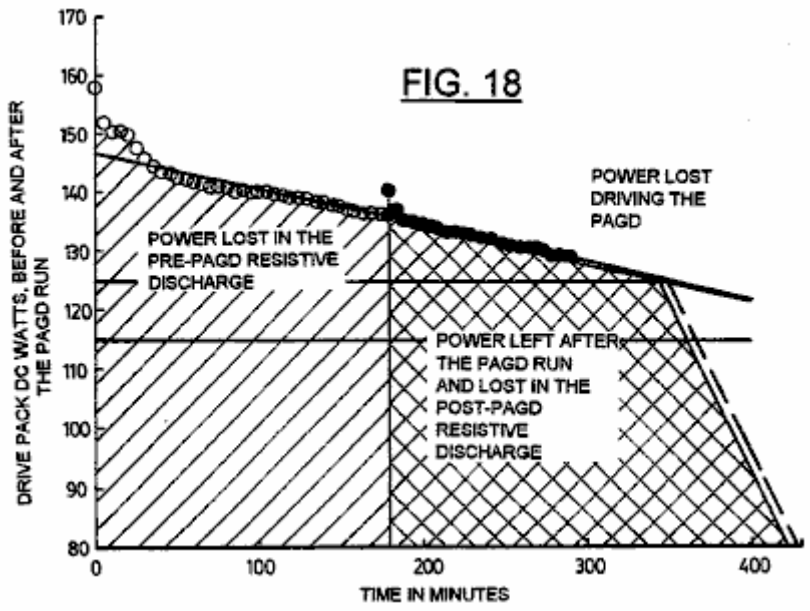


Fig.18 shows resistive discharge slopes for a drive pack before and after a very small expenditure of power in providing energy input to a PAGD run in a further experiment; R = 2083 ohms.

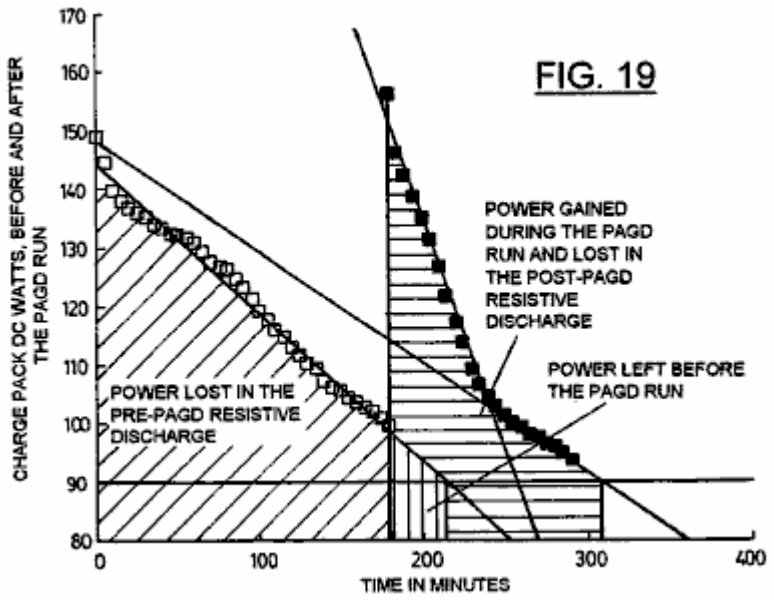


Fig.19 shows resistive discharge slopes for a charge pack before and after capturing energy from the PAGD run of Fig.18; R = 833 ohms.

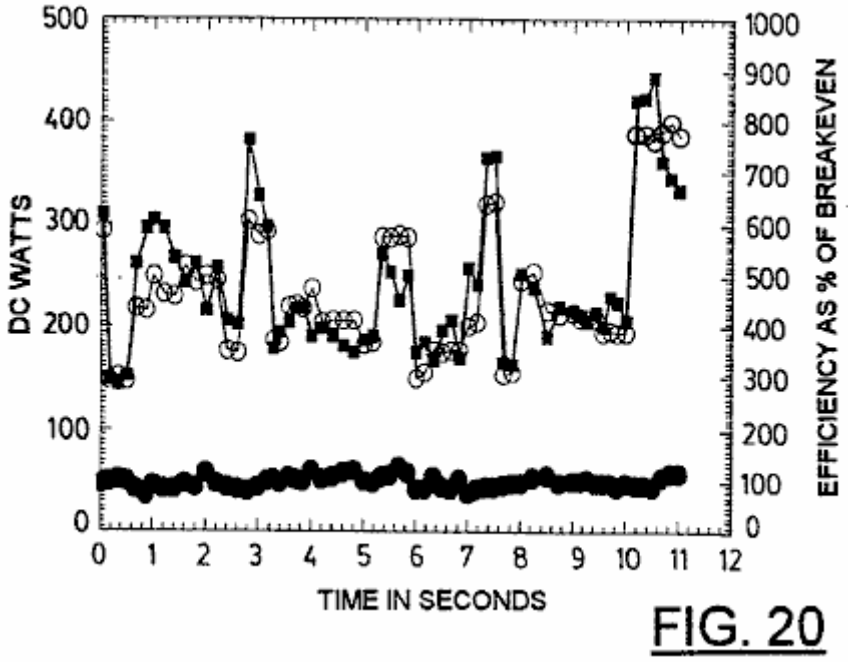


FIG. 20

Fig.20 shows an example of operational measurements taken videographically during a 10 second period for both the power consumption of the drive pack (PAGD input) and the power production captured by the charge pack (PAGD output); the two values are also related by the expression of percent break-even efficiency.

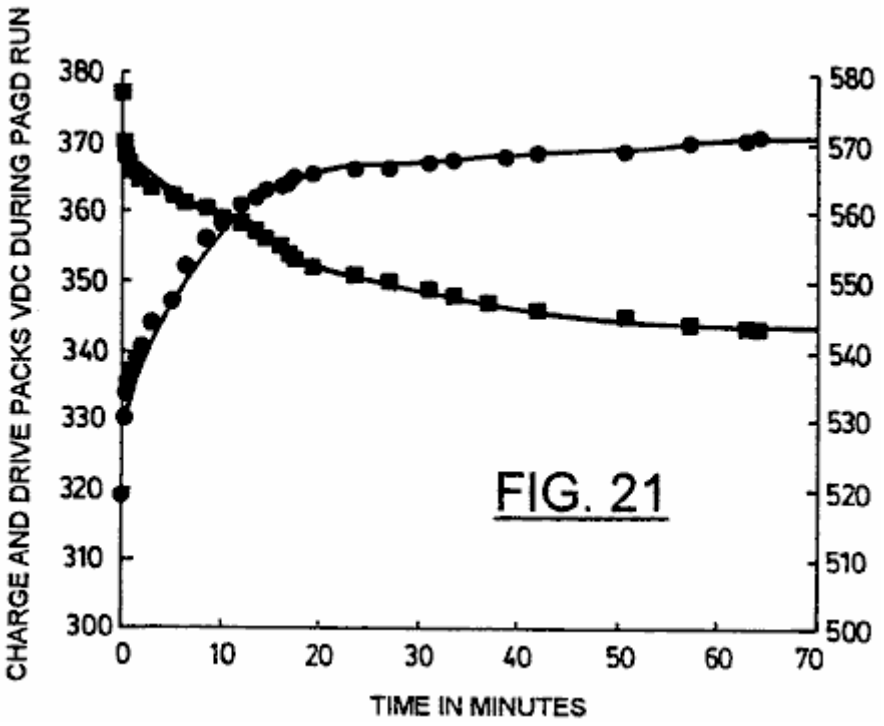


FIG. 21

Fig.21 shows variation of PAGD loaded voltage of a drive pack (in squares) compared with the PAGD charging voltage of the charge pack (in circles), during more than 1 hour of continuous PAGD operation.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

The basic PAGD function and the construction of discharge tubes specifically designed for PAGD operation are described in our corresponding co-pending applications Nos. 07/922,863 (the "863" application) and 07/961,531 (the "531" application). For purposes of the experiments described below four aluminium H34 plate devices (one with 64 and three with 128 cm^2 plate areas) and three aluminium (H200) plate devices (one with 64 and two with 128 cm^2 plate areas), with inter-electrode gap lengths of 3 cm to 5.5 cm, were utilised at the indicated vacua, under pump-down conditions and with either air or argon (ultra high purity, spectroscopic grade 99.9996% pure)

constituting the residual gas mixture. The pump-down conditions were as described in the “863” application. Some experiments were performed with the tubes under active evacuation, at steady-state conditions, while others utilised sealed devices enclosing the desired residual gas pressures.

The circuit designs utilised in the various experiments to be described are set out further below, and represent further developments and extensions of the circuits set forth in the “531” application.

Test equipment utilised was as follows:

An Edwards (trade mark) thermocouple gauge (TC-7) was employed for the determination of pressure down to 1 micron of mercury (0.001 Torr).

Banks of Beckman (trade mark) rms multimeters 225 and 330 (30 and 100 kHz bandwidths, respectively) were utilised for all current measurements.

Frequency meters capable of discriminating events up to 0.1 nanosecond apart, and having adjustable amplitude windows, were used. Direct analysis on a Tektronix (trade mark) dual-trace, storage scope (Model 549) was also carried out for both parameters.

Split-phase, single-phase and two-phase motors were employed, of the synchronous, induction and universal types, as previously described in the “531” application, in the accessory electromechanical arm that may be coupled to the power producing circuit described in the present application.

Large banks of 12 V, 6 Ah lead-acid gel cells (Sonnenschein (trade mark) A212/6S) were utilised either as power sources (designated as drive packs) or as accumulators of the energy (referred to as charge packs) captured by the test circuits. Charge packs made of rechargeable 9V NiCad or of nominally non-rechargeable C-Zn or alkaline batteries were also utilised.

PAGD emission areas were determined by metallographic examination of a series of craters produced by PAGDs in clean H34 cathodes, under a metallurgical Zeiss (trade mark) standard 18 microscope equipped with an epi-fluorescent condenser, very high power apochromatic objectives and a 100 W mercury lamp. For best results a focusable oblique source of light (12V halogen) was also added to the incident light.

Following our low and high applied current studies on PAGD production as set forth in the “863” application, we noticed that the AC rms value of the component associated with each abnormal glow discharge pulse varied non-linearly with the magnitude of the applied current. We originally noted the existence of a current induced shift of the entire PAGD region upward in the pressure scale: while the PAGD regime became more clearly defined as the applied constant DC was increased, the pressure required to observe the PAGD increased two to three orders of magnitude. In the course of these rarefaction studies we found that, at applied currents of 1mA or less, the rms value of the different AC waveforms associated with the consecutive regimes of the discharge (TRD --> NGDm --> AGD+PAGD) was, by more than half log, inferior to the value of the applied DC current, during the first two regimes (TRD and NGD) and reached a value equivalent to the applied current with the onset of spontaneous PAGD, at pressures < 0.1 Torr (see **Fig.1**); however, in the downward tail of the PAGD regime (down to 3×10^{-3} Torr), the AC rms current component of each PAGD again decreased to more than half log of the intensity of the applied DC value, in a manner proportional to the log of the decreasing pressure. In stark contrast, at high applied currents of about 500 mA, and aside from the high current-induced upward shift in pressure of the PAGD regime (to the point that the compression of the previous regimes on the pressure scale results in their suppressing, as was the case in the present example), the AC rms component associated with each pulse (see closed circles, **Fig.2**) is, from onset of the discharge at about 8 Torr, greater in magnitude than the value of the applied current (open circles, **Fig.2**). Under the conditions described, the distribution of the field current associated with each pulsed abnormal glow discharge approached (on a linear Y axis; not shown) an uni-modal gaussian distribution with the pressure peak at about 1 Torr, and a corresponding observed maximum of 7.5 times. higher AC rms values than the applied DC values.

We have previously described in the “863” application how the PAGD frequency is affected by several factors, namely:
the magnitude of the parallel discharge capacitance,
the value of the negative pressure for the relevant vacuum PAGD range,
the magnitude of the applied potential, the magnitude of the applied direct current,
the inter-electrode gap distance and
the area of the parallel plate electrodes.

In the “531” application we have also described how the wiring configuration (plate diode versus triode) affects the PAGD frequency by adding tungsten auto-electronic emissions from the axial electrode, to those emissions from

the plate. There are other factors which limit the PAGD regime of discharge and have also been discussed in the “863” application. The following data indicates their specific effect upon PAGD frequency.

In the data presented in Table 1, control of the frequency parameter for the circuit shown in **Fig.9** is by a ballast resistance **R1** within a specific range of interest (about 800-150 ohms, for Table 1 experimental conditions), and this in turn increases the applied current which, at "high current" values (i.e. >100 mA, as for Table 1 conditions), will drive the PAGD frequency up, as previously reported in the “863” application.

Table 2 shows the effect of the progressive displacement of a given frequency, chosen as 200 PPS, with the cumulative pulse count of the same device, in the plate diode configuration. This displacement of the same frequency (cf. group numbers 1-3 of Table 2) on to higher pressure regions is shown to be promoted by the alteration of the work function of the PAGD emitting cathode, such as this is caused by the cumulative pulse count and resultant crater formation on the electrode surface. After the first million pulses, the anode facing cathode surface is completely turned over by emission sites, and this corresponds well to the threshold crossed by group 2 of Table 2. Once the cathode surfaces are broken in, the rates shown in groups 3 and 4 of Table 2, tend to remain constant.

Originally we wondered whether this might be caused by the alteration of the electrostatic profile of the plasma sheaths at the periphery of the envelope, due to the mirroring deposits that result from the sputter of ions and trapped neutral atoms (from air gases or metallic vapour) associated with the auto-electronic emission mechanism (and from further emissions triggered in turn, by secondary ionic bombardment of the cathode with molecular species present in the plasma ball formed over the primary emission site). However, reversal of the plate polarity (firing the ex-anode as a crater-free cathode) for over a million counts, followed by re-reversal to the original polarity, the entire operation being performed in air as the residual gas substrate, led to the partial recovery of the original work function for as long as the test was run (1.5×10^4 pulses), as shown by a comparison of groups 2, 4 and 5, of Table 2. From a metallographic examination of the surfaces of plates used solely as anodes, we have also concluded that prolonged PAGD operation has the effect, not only of cleaning the anode surface from surface films and adsorbed gases, as ionic bombardment promoted by electromagnetic induction coils does, but it also does more: it polishes the target surface and smoothes it by a molecular erosive action. Observations of the surface of reversed cathodes, shows the same smoothing and polishing effects observed in exclusive anodes. Thus the recovery of the PAGD rates promoted by polarity reversal of the plates is not a function of the sputter-promoted mirroring deposits on the envelope wall, but a function of the actual work-function of the emitting cathode.

Another variable that interacts with the PAGD frequency is the molecular nature of the residual gas: Table 3 shows the differential frequency response of air with a halogen quencher, argon, for the same pulse generator employed in the tests of Table 2. It is apparent that argon obtains much higher rates of AGD pulsation for the same range of negative pressure, for the same "broken in" cathode, than does the air mixture. All these measurements were taken at cathode support-stem temperatures of 35^0C.

Time of operation is also a variable affecting the frequency and operating characteristics of the cathode, as it becomes expressed by the passive heating of the cathode, an effect which is all the more pronounced at the higher pressures and at the higher frequencies examined. Utilising the triode circuit discussed in the next section, the pulse rate of a PAGD generator with 64 cm^2 plates can be seen (see **Fig.3**) to decrease, at a negative pressure of 0.8 Torr, from 41 PPS to the operating plateau of 6 PPS within 15 minutes of continuous operation, as the temperature of the cathode support increased from 19^0C to about 44^0C. As the temperature plateaus at about 51^0C +/- 1^0C., so does the pulse rate at 6 PPS, for the remaining 48 minutes of continuous operation.

However, in order to confirm this time-dependent heating effect and threshold, we also performed the same experiment, utilising the same circuit and the same negative air pressure, with twice as large a cathode area (128 cm^2, which should take nearly twice as long to heat), being operated for 18 one-minute long continuous periods equally spaced apart by 15 minutes of passive cooling, with the cathode stem always at 19.7^0C to 21^0C., room temperature at the start of each period. The results surprised us, inasmuch as they showed that for a larger area tube which takes longer to heat to the same temperatures at comparable rates of PAGD triggering, one could observe a much earlier frequency reduction (by half, within the first 5 minutes or periods of interrupted functioning) in the absence of any significant heating effect (< 1.5^0C) of the cathode (see **Fig.4**). Repetition of these experiments has led us to conclude that, as shown in **Fig.5**, the variable responsible for this repeatedly observed reduction in the PAGD frequency, when the PAGD operation sequence is systematically interrupted, is the state of charge/discharge of the battery pack (the charge pack) at the output of the triode circuit in question: the PPM rates in **Fig.5** decrease rapidly with the steepest rate of charging of the charge pack and the fastest recovery rate of its open circuit voltage; above a given state of charge, when the open voltage of the charge pack climbs more slowly (> 340 V), in a log fashion, the PPM rate stabilises at its plateau values.

Confirmation of the importance of the charge pack in the PAGD function of the present circuitry here considered, comes from the fact that the size (the number of cells) and the intrinsic capacitance of the charge pack affect the PAGD frequency dramatically (see Table 4): increasing the charge pack size of 29 cells to 31, by 7% leads to a 10-fold reduction in frequency; further increases in the number of charge pack cells extinguishes the phenomenon. On the upper end of the scale, this effect appears to be tied in to restrictions that it places on the ability of the larger charge packs to accept the discharge power output once the charge pack voltage exceeds the PAGD amplitude potential. All of these measurements were conducted with the same 128 cm^2 plate PAGD generator, at a pressure of 0.8 Torr and in the triode configuration (see **Fig.9**).

Other factors can also affect the frequency: the motion of external permanent magnetic fields oriented longitudinally with the inter-electrode gap, external pulsed or alternating magnetic fields, external electrostatic or electromagnetic fields, specific connections of the earth ground, and the presence of a parallel capacitative, capacitative-inductive or self-inductive arm in the circuit, such as we have described for our electromechanical PAGD transduction method as described in the “531” application.

Analysis of the modulation of PAGD amplitude is simpler than that of its frequency, because fewer factors affect this parameter:
(1) magnitude of the applied potential,
(2) inter-electrode gap distance and
(3) the negative pressure, as shown in the “863” application, for "low" applied currents.

As the magnitude of the applied potential itself is limited by the gap and the pressure, to the desired conditions of breakdown, the important control parameter for the PAGD amplitude is the pressure factor. This is shown in **Fig.6** and **Fig.7**, respectively for "low" (5 mA) and "high" (about 500 mA) applied currents and for the same plate diode configuration of a H34 Al 128 cm^2 plate PAGD generator (5 cm gap), in the simple circuit described in the “863” application; it is apparent that both positive and negative components of the amplitude of these pulses in the oscillograph, are a function of the pressure, but the maximum cut-off limit of our equipment, for the negative component (at 240 volts for the "low" current experiment and at 120 volts for the "high" current), precluded us from measuring the peak negative voltage of these pulses.

However, rms measurements of the pulse amplitude at the plates and DC measurements at the circuit output to the charge pack indicate that the negative component increases with decreasing pressure to a maximum, for a given arrangement of potential and gap distance; no pressure-dependent bell shape variation of the pulse amplitude, as that seen for the positive component at "high" applied currents (**Fig.7**) is observed with the negative amplitude component. For the typical range of 0.8 to 0.5 Torr, the rms value for pulse amplitude varies from 320 to 480 volts, for a 5.5 cm gap distance and applied DC voltages of 540 to 580 volts. PAGD amplitude is a critical factor for the design of the proper size of the charge pack to be utilised in the optimal circuit.

The development of the circuits to be described stemmed from fundamental alterations to the principles implicit in our previous methods of electromechanical transduction of AGD plasma pulses as described in the “531” application. Whereas this electromechanical coupling (capacitative and self-inductive), utilised directly, energises the AGD pulses inverted from the DC input by the vacuum generator, the purpose of the development that led to the presently described experiments was to capture efficiently, in the simplest of ways, most of the pulse energy in a closed circuit, so that power measurements for the energy transduction efficiency of the observed endogenous pulsation could be carried out. Ideally, comparative DC power measurements would be performed at both the input and output of the system, taking into account the losses generated across the components; this would overcome the measurement problems posed by the myriad of transformations implicit in the variable frequency, amplitude, crest factor and duty-cycle values of the PAGD regime, and necessitated some form of rectification of the inverted tube output.

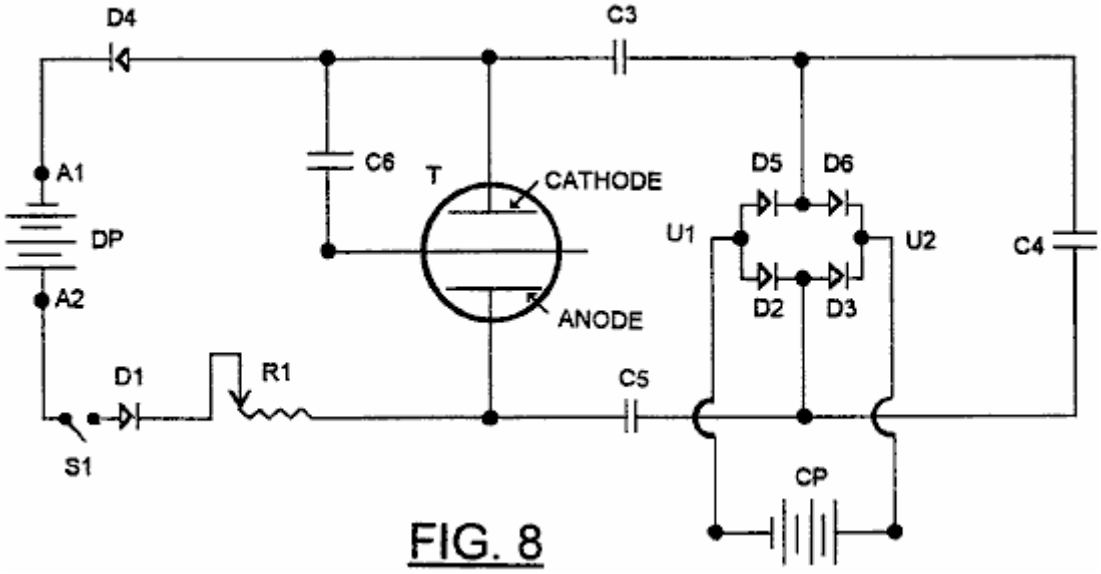


FIG. 8

From the start our objective was to do so as simply as possible. Early circuits utilising half-wave rectification methods coupled in series to a capacitative arm (for DC isolation of the two battery packs), with the charge pack also placed in series, showed marginal recoveries of the energy spent at the PAGD generator input. Attempts at inserting a polar full-wave rectification bridge led, as shown in **Fig.8**, to the splitting of the capacitor into capacitors **C3** and **C5**, at the rectification bridge input, and capacitor **C4** in series with both capacitors, all three being in a series string in parallel with the PAGD generator. Under these conditions a DC motor/generator could be run continuously in the same direction at the transversal output (**U1** and **U2**) of the bridge; but if this inductive load was replaced with a battery pack CP (charge recovery pack), either the parallel capacitor **C4** had to remain in the circuit, for the diode configuration or, less desirably, a further capacitor **C6** could replace **C4** and connect one electrode, preferably the cathode **C**, to the axial member of the discharge tube **T**, thus resulting in a first triode configuration as actually shown in **Fig.8**. Energy recovery efficiencies of the order of 15% to 60% were obtained utilising **C6** in this manner, but measurements of the potential and currents present at the output from the rectifier bridge were substantially lower than those obtained using optimal values of **C4**. Effectively, under these conditions, much of the power output from the tube was never captured by the output circuit formed by the second, right hand arm of the system and, being prevented from returning as counter-currents to the drive pack **DP** by diodes **D1** and **D4**, was dissipated and absorbed by the inter-electrode plasma, electrode heating and parasitic oscillations.

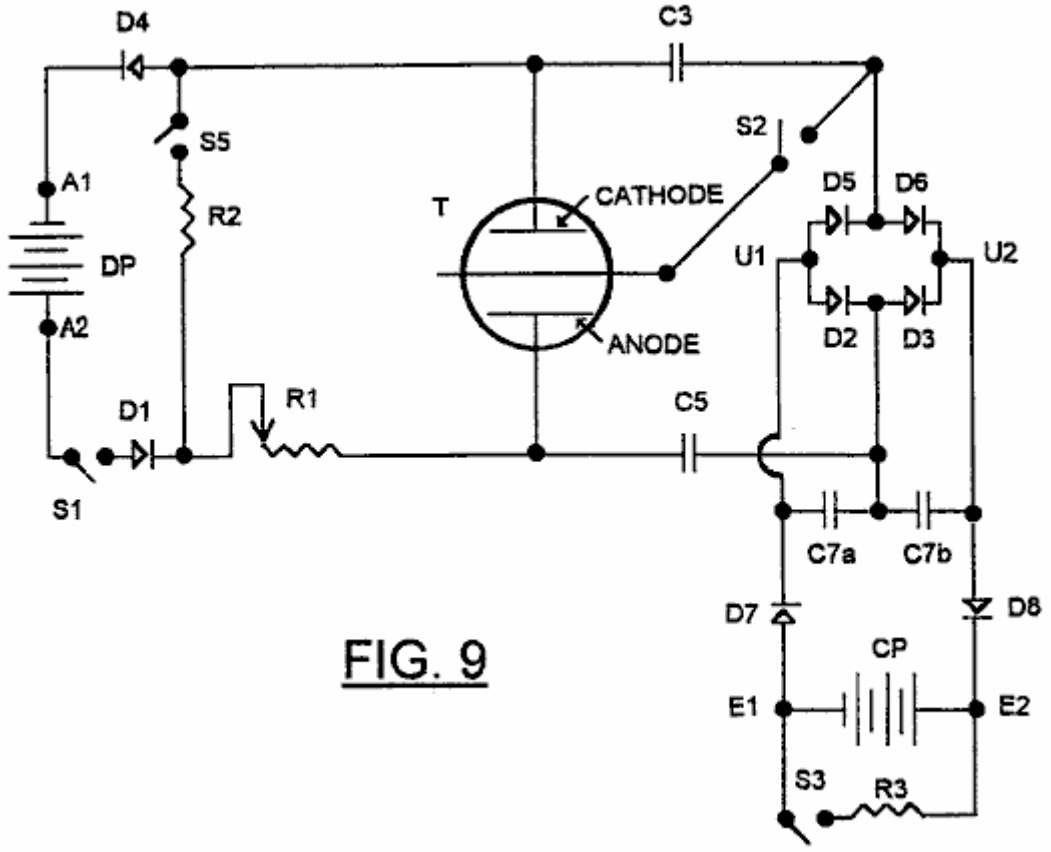


FIG. 9

Solutions to this problem were explored using the circuit shown in **Fig.9**, which still maintains the necessary communication link for the quasi-sinusoidal oscillation of the capacitatively stored charges at the input and outputs of the rectification bridge, but integrated the functions of capacitor **C4** into the single rectification circuit, in the form of an asymmetric capacitative bridge **C7a** and **C7b** placed transversally to the capacitative bridge formed by **C3** and **C5** and in parallel with the charge pack **CP** at the output from the rectification bridge **D5, D6, D2, D3**.

This second capacitative bridge is so disposed as to have its centre point connected to the anode **A** through capacitor **C5**. If the axial member of the Tube **T** were to connect to the junction of **D2** and **D3** instead of at the junction **D5-D6**, the function of bridge **C7a** and **C7b** would be connected to the cathode **C** through capacitor **C3**. The capacitative bridge is insulated from the charge pack whose voltage it stabilises, by rectifiers **D7** and **D8**, which also prevent leakage of charge across **C7a** and **C7b**.

The anode and cathode oscillations generated by the electrostatic charge transduction through **C3** and **C5** into the poles of the charge pack are trapped by the transversal transduction of the **C7** bridge, at the outputs from the rectification bridge, of which the oscillation has to become split between the bridge inputs into half-waves, for electrostatic transduction and full wave rectification to occur. In fact, under these conditions, removal of the **C7** bridge will suppress the PAGD phenomenon, unless other circuit variables are also altered. The transversal bridge is thus an essential piece of this novel circuit. Variations in the circuit as shown in **Fig.10** were then studied, the first two being selectable utilising switch **S2** (**Fig.9**).

The presence of the capacitative bridge effectively reduces the dynamic impedance of the charge pack **CP** so that the output circuit approximates to a characteristic in which it presents a very high impedance to the tube **T** at potentials below a certain level, and a very low impedance at potentials above that level.

With this modified circuit, more effective recovery of the energy produced by collapse of the PAGD pulses is possible, with more effective isolation from the input circuit utilised to trigger the pulses. Under these conditions, the energy captured by this circuit at the output, is not directly related to that utilised in triggering the pulses from the input. The attainment of this condition critically depends on the large capacitance of the transversal bridge being able to transfer the output energy from the tube **T** into the charge pack **CP**. Under these conditions, we have found, as will be shown below, that the large peak pulse currents released by collapse of the PAGD pulses released more energy than is used to trigger them, and these findings appeared to tally with other observations (abnormal volt-ampere characteristics and anomalous pulse currents, etc.) associated with the anomalous cathode reaction forces that accompany the auto-electronic emission-triggered PAGD regime. Experiments so far indicate that the power output can be increased proportionately to the series value of **C3**, **C5** and the two identical **C7** capacitors.

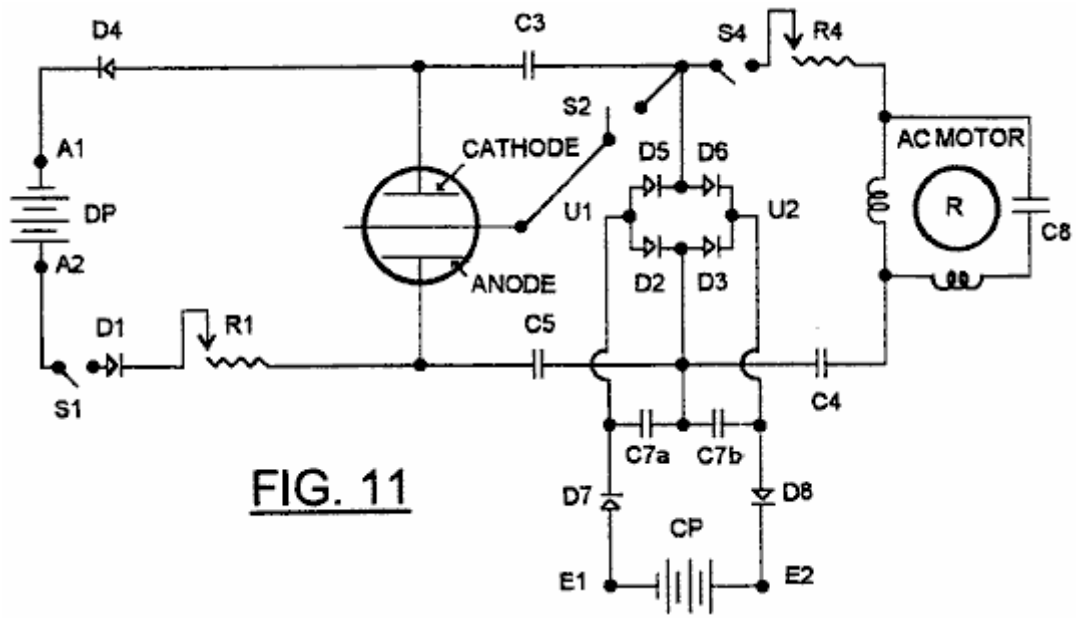


FIG. 11

The circuit of **Fig.10** can be integrated with a circuit such as that disclosed in the “863” application as shown in **Fig.11**, in which a part of the energy recovered can be shunted by the switch **S4** into an induction motor **M1** having rotor **R**, to a degree determined by the adjustment of potentiometer **R4** and the value selected for **C4**.

The circuit of **Fig.11** can be further developed as exemplified in **Fig.12** to include configurations which provide switching permitting interchange of the functions of charge packs and the drive packs, it being borne in mind that the nominal potential of the drive pack must be substantially higher than that of the charge pack, the former needing to exceed the breakdown potential of the tube at the beginning of a PAGD cycle, and the latter to be less than the extinction potential.

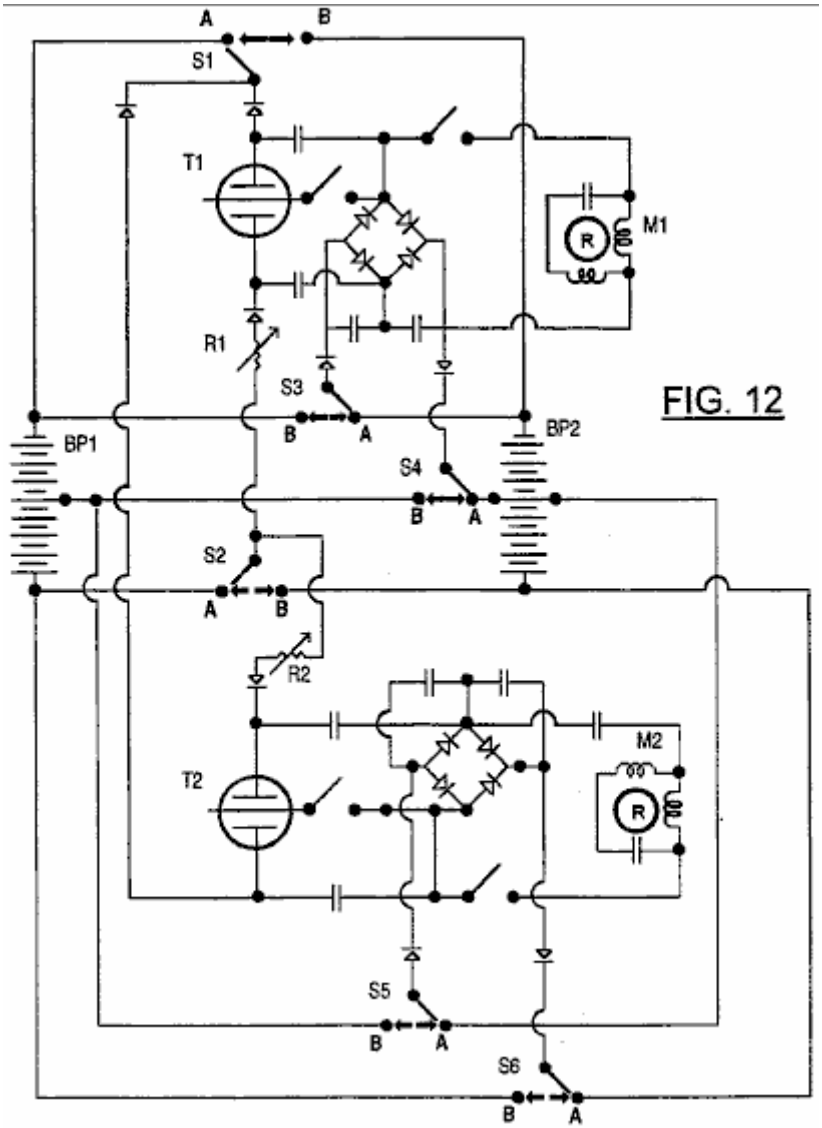


FIG. 12

Fig.12 essentially represents a duplication of the circuit of **Fig.11**, the two circuits however sharing two identical battery packs **BP1** and **BP2**, and being provided with a six pole two way switch, the contact sets of which are identified as **S1, S2, S3, S4, S5** and **S6**. When the contacts are in position **A** as shown, battery pack **BP1** acts as a drive pack for both circuits, with the upper half (as shown) of the battery pack **BP2** forming the charge pack for the upper circuit, and the lower half forming the charge pack for the lower circuit. When the pack **BP1** is at least partially discharged, the switch is thrown so that contacts move to position **B**, which reverses the function of the battery packs thus allowing extended operation of the motors in each circuit each time the switch is thrown.

Based on the manufacturer's data, and using current values within the range of our experimentation as discussed in the next sections, an optimal discharge cycle for a fully charged 6.0 AHr battery pack at 0.300 A draw is 20 hours, as claimed by the manufacturer, and this corresponds to a cycling between 100% (12.83 V/cell open circuit and load start voltage) and < 1% (10.3 V/cell load voltage) of the battery's absolute charge capacity. Even though the discharge mechanism is a time cumulative process with a log function, the discharge can, within 4 to 5 hour time segments (or periods with 20%-25% of the full range), be regarded as practically linear with time. This trait, or linearisation of the discharge slope, becomes more marked with advancing age and decreasing absolute storage capacity of the cells.

The proportionality between open circuit voltage and the percentage of residual relative capacity for these cells when new (uncycled and not yet aged) is uniform over 98% of the permissible charge capacity withdrawal. In practice this translates into a slope that becomes steeper with time, while the absolute storage capacity diminishes. In turn, this decreasing absolute capacity of the cells results in shorter load discharge times and their further linearisation.

A circuit in general accordance with **Fig.9**, employed in the studies reported in this and the following sections, utilises a drive pack of 46 12 V Lead acid gel-cells each with a 6.0 Ah rating, and a charge pack with 28 or 29 12 V identical cells. The charge pack was cycled anywhere from 11.2 V to 12.8 V/cell (open circuit voltages), within the proportional region of the relative capacity slope, to yield a capacity increment in the order of 50% (e.g. from

20% to 70%), anywhere within the range of 2% to 100% of its total charge capacity, assumed for now as invariant. The charging process, hereinafter referred to as a PAGD run, took about 20-30 minutes under optimal conditions. The drive pack typically consumed, in the same period of time, 4% to 11% of its initial total capacity , its open circuit voltage typically falling 0.1 V to 0.2 V per cell after a PAGD run, within the open circuit range of 12.8 V/cell (100% relative capacity) and 11.2 V/cell (about 2%). At the 100% capacity benchmark, the drive pack would theoretically have 20 h x 46 cells x 12.83 V/cell x 0.3 A = 3.5 kWh, and the charge pack, for example, 20 h x 29 x 12.83 V/cell x 0.3 A = 2.2 kWh. Since the capacity per cell is linear with the open circuit voltage within the proportional range, as claimed by the manufacturer, we projected the open circuit voltage intercepts on the manufacturer's proportional curve in order to determine the residual percentage of the total relative capacity and the standard hours of operation left, from any experimental open circuit voltage measurements.

Three pulse generators (one 64 cm^2 and two 128 cm^2 plate areas) were employed in these studies; they were operated in PAGD runs at 1-120 pulse/second rates, within a negative pressure range of 0.2 to 0.8 Torr and with applied direct currents of 0.2 to 0.6 A.

Both drive and charge packs utilised cells which were bought new at the same time and had initial charge values of 12.4 to 12.55 V/cell (open circuit). These batteries are capable of energy densities of 33-35 WHr/Kg. However, the experiments shown in Table 5 are selected from a series that spanned nearly 12 months, beginning 6 months after purchase; hence, loss of absolute storage capacity by the batteries had occurred in the intervening time, as a function of both age and charge/discharge cycle life.

Measurements of the open voltage of either drive (D) or charge (C) (see column 2, Table 5) packs for 8 separate experiments, all utilising the triode configuration, were performed before (b) and after (a) a PAGD run (see columns 3 and 4), at either 15 or 30 minutes (see column 26) of the open circuit voltage relaxation after a PAGD run was terminated. Corresponding open circuit voltages per cell are shown in column 5, and the percentages of the predicted total relative charge capacity resulting from the intercepts on the manufacturer's proportional curve are shown in column 6, Table 5. Equivalent maxima for the theoretical hours of operation left are shown in column 7, the percentage change in relative capacity arising as a consequence of either charge pack charge capture (capacity gained) or of drive pack output (capacity lost) is shown in column 8. Translating the intercepts into power units yields the values shown in column 9, Table 5, for total kWh left in each pack before and after PAGD production, those shown in column 10 for the actual power gained and lost during the periods of operation (presented in column 12) and those shown in column 13 for the power predicted to be gained or lost per hour of PAGD production.

On the basis of the experimental open voltage values and their intercepts, the predicted net kWh values per hour of PAGD energy production (after deduction of measured losses) and the corresponding experimental break-even efficiencies (where breakeven = 100%) are presented, respectively, in columns 14 and 15. The PAGD frequency per second is shown in column 11; the number of 12 V cells, in column 16; the tube ID, in column 17; the cathode (and anode) area (s), in column 18; the plate material, in column 19; the input ballast utilised (R1, FIG. 9), in column 20; the size of each capacitor (C3 or C5) of the tube output bridge, in column 21; the size of each capacitor (C7a or C7b) of the transversal capacitative bridge, in column 22; the status of **S4** and thus, of the parallel and auxiliary electromechanical arm (see **Fig.11**), in column 23; the negative air pressure in column 24; the gap distance between the plates, in column 25; and columns 27,28 and 29, show the status of the elements of the switched on parallel electromechanical arm of the circuit--the parallel **C4** capacitor, the motor input resistor **R4** and the motor revolutions per minute (measured stroboscopically), respectively.

From these figures of Table 5, and utilising the data for the two first examples shown, we calculated the predicted performance of the system based on the open voltage measurements. In the first example, where the system was run continuously without interruption, the charge pack increased the percentage of its total capacity by 43% (a two-fold increase in capacity) and, during the same period, the driver pack decreased the percentage of its total capacity by 7% (an approximately 10% decrease in capacity relative to the percentage of residual total capacity at the start, i.e. 77%) (cp. columns 6 and 8, Table 5). Subtracting the predicted initial total energy (0.835 kWh) available to the charge pack before the experimental run (first line of column 9, Table 5) from the predicted total energy (1.823 kWh, second line of column 9) available to the charge pack after the PAGD charge run, gives us the total energy gained by the charge pack: 0.988 kWh (column 10) in 21.5 minutes (column 12) of continuous PAGD performance.

Conversely, subtracting the predicted final total energy (2.4 kWh) available to the driver after the experimental run (fourth line of column 9, Table 5) from the predicted total energy (2.66 kWh, third line) available to the driver before the PAGD charge run, gives us the total energy lost by the drive pack: 0.26 kWh in 21.5 minutes. If we divide the total available energy gained by the charge pack, by the total energy lost by the drive pack, we obtain a surplus factor of 3.9., or 388% of the break-even point (column 15). The same values result from dividing the charge pack % of total capacity gain by the drive pack % of total capacity lost, and then down-scaling this value by multiplying it by the typical scale factor for the two packs, 29 / 46 = 0.63 times.

In an analogous fashion, we analysed the results for the second example shown in Table 5. Here, the charger increased the percentage of its total capacity by 45.5% (a 22.75 fold increase in estimated total relative capacity) and, during the same period, the driver decreased the percentage of its predicted total capacity by 7% (about a 17.5% decrease in capacity relative to the percentage of residual total capacity at the start, i.e. 40%). By dividing the predicted total available energy gained by the charge pack (0.962 kWh/18 minutes) by the expected total energy lost by the driver pack (0.246 kWh/18 minutes) we obtain a surplus factor of 3.9 times, or 391% of the break-even point. This corresponds to an interrupted, total sequential run of 18 minutes, each minute-long run being separated by a cooling and voltage relaxation period of 15 minutes before the next run is carried out, at an average PAGD frequency of 61 PPS.

Analysis of the remaining results illustrates how a number of PAGD controlling parameters interact to determine conditions for effective maintenance of a PAGD regime. The lower gain and higher loss per unit time registered for the third run of Table 5, which results in the lower break-even efficiency of 230% and a smaller net power production rate than before (power estimates of 1.396 kWh/h of PAGD operation vs 2.387 kWh/h, for the second run, Table 5) illustrate, for example, the combined effect of lowering the pressure (0.8 to 0.7 Torr) and running the PAGD continuously (the heating effect), both of which depress the PAGD frequency. The fourth run of Table 5 identifies the continuous performance of a "broken in" softer grade of aluminium (column 19), having a lower work-function (as determined from the higher PAGD frequency spectrum) than the harder H34 plates of the previous examples, and shows that, despite the series value of the total capacitance being higher (5,333 mF vs 4,030 mF for runs one through three), and despite the higher vacuum (0.2 Torr), the lower work-function results in a higher frequency; however, even though this run registers a predicted higher break-even efficiency (310%) than the previous experiments, these conditions result in a 4 / 5-fold lower estimate of net power produced, when compared to the previous three PAGD runs.

PAGD runs 5 and 6, Table 5, illustrate the effect of switching on the auxiliary electromechanical arm of the circuit shown in **Fig.11**. Increasing the amount of charge capacitatively shunted into the electromechanical arm by higher **C4** values (column 27), and increasing the current that feeds the squirrel cage induction motor utilised by lowering **R4** (column 28), results in a power capture by the charge pack that registers an energy loss (predicted to be 96% efficient, falling short 4% of break-even recovery), as most of the tube output power is spent in the electromechanical arm and its motor effect. Furthermore, under the conditions of maximum electromechanical action, the drain imposed on the drive pack becomes considerable (see loss in columns 10 and 13), even if the **C3** and **C5** values are reduced, column 21, Table 5). These runs also illustrate how the motor appears to function as an electrical induction generator having rpm values much higher than the synchronous values prescribed by the frequency of the PAGD (column 29, Table 5).

The extremely large break-even efficiency of PAGD run 5, Table 5, indicates that with selected values of **C4** and **R4**, it is possible to operate the motor in the auxiliary arm and still accumulate excess energy from the PAGD production in the charge pack.

Runs 7 and 8 illustrate results obtained for 64 cm^2 plates, and a shorter inter-electrode gap distance, for two pressures (0.8 and 0.5 Torr), the device being open to a rotary pump manifold in the first instance and sealed from the pump, in the second case. Despite the lower vacuum, the higher pulse frequency (32 vs 5 PPS) and break-even efficiency (906% vs 289%) registered by run 8 when compared to run 7, are a consequence of the method of run 8, which was interrupted systematically by 5 passive cooling periods, as in the case of run 2, whereas run 7 was continuous. This again resulted in higher average PAGD frequencies (at lower pressures), a predicted two-fold greater gain and a predicted two-fold smaller loss (columns 13 and 14) for run 8.

Fig.13 shows curves representing the slopes of the open circuit relaxation voltages, which are linear with the log of time elapsed from cessation of discharge, for both drive and charge packs, in the same run 8 set out in Table 5. The experiment in its entirety consisted of preliminary resistor-loaded measurement discharges and their corresponding open circuit voltages from the moment of cessation of the resistive discharge (illustrated, respectively, by the open squares of DPT1 for drive pack relaxation time 1, and by the open circles of CPT1 for charge pack relaxation time 1), followed by their relaxation rates in the wake of the PAGD production (the hatched squares of DPT2 for drive pack relaxation time 2, and the hatched circles of CPT2 for charge pack relaxation time 2), and finally, by the relaxation rates from the final resistor-loaded measurement discharges (the black squares of DPT3 for drive pack relaxation time 3, and the black circles of CPT3 for charge pack relaxation time 3). Discharge resistances were 833 ohms for the charge pack, and 2083 ohms for the drive pack in all cases, corresponding to resistors **R3** and **R2**, respectively, of **Fig.9**. This methodology will be examined in greater detail below. It is apparent that, after every load period, be this resistive (CPT1, DPT1, CPT3 and DPT3) or due to PAGD operation (DPT2), the relaxation slope is positive; as shown from slopes CPT1 and DPT1, the log time proportionality of the open circuit voltage relaxation, under these conditions, tends to plateau after about 30 minutes. The exception to this general behaviour lies in the voltage relaxation slope CPT2, which is negative and

reflects the charge accumulation occurring in the charge pack and obtained by capture of energy produced during PAGD operation, triggered by the energy drawn from the drive pack during load time 2.

As a first approximation of electrical power generated and consumed by the energy conversion system of the invention, the previous open circuit voltage method is of significance in showing the basic trends involved in interaction of the operating parameters. However, in all likelihood, it overestimates the actual values of electrical power consumed and generated, for a variety of reasons. First, it assumes that the relative capacity scale of the batteries in the drive and charge packs is an absolute charge capacity scale with an invariant maximal charge retention, which it is not; in fact, the absolute charge capacity is itself a variable subject to several factors, such as the cycle life, overcharging or undercharged conditions, cell age, residual memory and the rate of charge and discharge. Hence, the inference of a uniform time scale on the basis of the open circuit voltage/capacity intercepts may not be warranted. Finally, it does not integrate the open voltage decrease over time, and utilises the specification load current as the average current over time.

In order to obviate these problems, we resorted to a variety of other measurement methods. First, we proceeded to compare the closed circuit, preliminary, resistive-load discharge measurements for either charge or drive pack, under conditions of negligible loss of power, as these measurements were statistical means (n = 9) taken, at equal intervals, during the first 90 seconds of the load discharge, and obtained both just before the PAGD production runs (but separated from each PAGD run by an open circuit voltage relaxation of 30 minutes) and just after the runs (but equally separated by a relaxation of 30 minutes). As an example of the data generated by such an approach, **Fig.14** illustrates the shift of the slopes indicating marginal power loss for the drive pack (from the closed squares to the open squares) and those indicating gain of power for the charge pack (from the open circles to the closed circles), in actual total load power values.

Integration of these power measurements over the projected load discharge time, taken from the family of curves generated on the basis of the manufacturer's load voltage over discharge time specifications, led to a direct comparison of the new values, as shown in Table 6, with the values presented in Table 5, for the first three instances introduced. All values of Table 6 were obtained by resistive measurements of power that entailed a negligible power loss. Table 6 confirms the fundamental equivalence of runs 1 through 3, as already seen from their corresponding analysis using the open voltage method (see runs 1 to 3, Table 5). This new power estimation method also confirms the lower loss encountered in run 2 utilising interrupted PAGD operation. While the break-even efficiencies sensibly doubled using this method, the estimates of actual electrical power consumption recovery decreased by a 2 to 3-fold factor. Thus this direct load voltage/amperage measurement method of estimating actual power losses or gains, is a check upon the open voltage method previously utilised.

Direct, instantaneous measurements of the voltage and current characteristics of the PAGD production and capture phenomena being discussed, were also performed during PAGD runs for diverse sets of conditions, including all those described in the two previous sections. In Table 7 we show these results for two PAGD generators having an identical electrode area (128 cm^2) and connected to electrical energy capture circuits of three separate configurations as set forth in **Fig.10A, Fig.10B** and **Fig.10C** and column 2, Table 7. In the configuration of **Fig.10C**, or double diode configuration, both electrode plates act as cathodes and the axial member as the anode collector (experiments 1-4, for the H220 device and 13-14, Table 7, for the H34 device). In the configuration of **Fig.10B**, or triode configuration, one plate acts as the cathode, the axial member as an auxiliary cathode and the other plate as a collector (experiments 5-9, Table 7). In the configuration of **Fig.10A** or single (plate to plate) diode configuration, the axial member is disconnected, and the polarity of the plates remain as in the triode configuration (experiments 10-12). All measurements were taken after 1 minute of PAGD operation of the devices, which were, at the start of each run, at room temperature. All cathodes had been previously broken in with > 2×10^6 AGD pulses. The open circuit voltage of the charge pack was, for all cases, at 359 to 365 volts, before each test. The direct measurements of the PAGD input and output DC voltages and currents were obtained as statistical means of 10 second long measurements, and at no time did the standard error of the plate voltage mean exceed 35 volts.

The air pressure within the tube during these tests is shown in column 3, Table 7, the drive pack DC voltage (X), in column 5, the DC voltage across the plates (Y), in column 6, the drive pack output current (PAGD input current), in column 7, and the drive pack total watts output is shown in column 8. Columns 9 and 10 show the PAGD voltage (PAGD V = (X-Y) / I_{av}) and the value of the PAGD extinction potential in V/cm. The recovery co-ordinates (i.e. the PAGD output energy) found at the **U1-U2** output (**Fig.9**), are shown in columns 11 to 13, as the charge pack's E1-E2 input DC voltage, amperage and power watts, respectively. The calculated resistance of the entire circuit is given in column 14, the registered PAGD frequencies in column 16, and running conditions in columns 17 to 18. The break-even efficiency obtained by direct comparison of the electrical power figures for the drive and charge packs, respectively, is given in column 15. This assumes, for purposes of a generalisation of power production rates over time, that the quasi-instantaneous, direct measurements here obtained can be translated to outputs obtained per unit time, and thus into direct Watt-hour measurements.

Data from runs 1 through 4 demonstrate that, at these PAGD frequencies, there is no difference between using fast switching (32 nanoseconds) MUR 860 diodes, or regular 40HFR-120 silicon diodes, in the rectification bridge of the electrical energy capture circuit, and that the PAGD frequency varies as a function of decreasing air pressure.

Runs 5 to 14 show that, in general, for the same tube, the single and double diode configurations are the most efficient, for the same pressure, the diode configuration typically yields some 1.5 to 2 times larger break-even efficiencies (cp runs 10-11 and 13-14, with runs 5-9, Table 7). The largest accumulations of power are also registered in the diode mode(s). This trend appears to be a function of the much lower cathodic work-function of the aluminium plates, than of the tungsten of the axial member utilised as an auxiliary cathode in the triode configuration. A feature of the data from these 14 different runs is the consistent excess power outputs (column 15, Table 7) and their narrower range (218 to 563%), when compared to those observed with the previous two methods of experimental analysis.

Run 12, Table 7, shows that the switching on of the electromechanical arm can be performed without entailing a power loss in the PAGD capture circuit, as previously found for run 5, Table 5, utilising the open circuit voltage method. In fact, with **C4** = 8 microfarads and R4 = 500 ohms, the AC induction motor behaves as an electrical flywheel (e.g. 2800-3000 rpm for 10 PPS inputs), while the electrical energy capture circuit still registers a sizeable excess electrical power production (compare runs 11 and 12, Table 7). Runs 13 and 14 illustrate how the charge pack's state of charge and its inherent capacitance affects both the PAGD frequency and the power producing efficiency of the entire system: as the charge pack is reduced from 29 to 19 cells, the PAGD generator adjusts by reducing its frequency logarithmically and, while the charge pack input current is greater than before, the drive pack loss becomes still larger and the break-even efficiency much lower (by >1/2, from 563% to 228%). This is because the circuit must translate the naturally larger PAGD amplitude into a larger surplus of output current, and in this process becomes less efficient.

If the first measurement method employed (the open circuit method) had to make too many theoretical assumptions about the system's performance under load conditions and hence about its effective charge capacity, the second approach still had to suppose an invariant discharge time and thus an invariant absolute charge capacity on the part of the battery systems (charge packs) employed for capture which it approximated by an operation of integral calculus. With the third method described above, theoretical assumptions were avoided except that, in these measurements, the actual performance of a given battery in terms of time, time of delivery and time of capture, was also ignored; no account is taken of the time-dependent modulation of the PAGD frequency, as effected by certain of the parameters analysed, namely the charge pack state of charge, the method of sequencing the PAGD runs (continuous vs interrupted) and its concomitant heating effects, and the state of charge (load voltage and current capacity) of the drive pack. A simple, non-negligible, resistive measurement of power lost by the drive pack, and an identically non-negligible measurement of the power gained by the charge pack, for the same experiment and the same singular time of PAGD production, were performed repeatedly to corroborate the previous three approaches. For this purpose, all experiments were designed as a continuous series of sequential phases:

1) Before a PAGD run, a resistive discharge was measured across either pack over periods of 1 to 3 hours (utilising the DP and CP resistances previously reported in the open voltage section) and followed by a 15 to 30 minute open circuit voltage relaxation;

2) Then, the PAGD runs were performed, either continuously or as interrupted, composite sequences, and the corresponding open circuit relaxation voltage(s) were measured, after the cessation of the integral PAGD run;

3) Finally, resistive discharge measurements, obtained under identical conditions to those recorded before the PAGD run, were carried out for either pack, followed by concomitant battery voltage relaxation rate measurements.

Under these experimental conditions, exact power measurements could be taken from an analysis of the actual battery discharge curves before and after the PAGD run. Based on a comparison of the curve trends of the pre-run resistive discharge of the drive pack with those of the post-run resistive discharge, the effective power drawn ($DeltaE_c$) from the withdrawable power capacity of the drive pack incurred during a PAGD run, was ascertained. This represents the power consumption during the run, and the experimental value thus recorded constitutes the actual power figure that must be matched for break-even to occur. Hence, the break-even value equals, by definition, the electrical energy input to the system. Similarly, a comparison of the charge pack pre-run and post-run resistive discharge curve trends identified the effective power ($DeltaE_{rho}$) added to the withdrawable capacity of the charge pack. This quantity represents the electrical energy recovered during the run. The relation for the two quantities is expressed by the break-even efficiency equation:

$\% = DeltaE_{rho} / DeltaE_c \times 100$

If the break-even efficiency is less than 100%, then the apparatus registers a net loss in electrical energy in the CP with respect to the DP. Conversely, if the efficiency exceeds 100%, then there is a net gain in electrical energy in the CP, as compared to that lost in the DP. For purposes of this analysis, a limit to the minimum withdrawable capacity was placed, from experiment and in agreement with the load current curves of the manufacturer, at 115 W for the driver pack (average current of 0.250 A, minimum current of 0.230 A), and at 90 W for the charge pack (average current of 0.375 A, minimum current of 0.334 A), as a function of both their total cell size (respectively, 46:29) and the difference in the resistive loads employed for the discharge measurements. All cathodes had been broken in, as described before.

The results obtained with this fourth method, for six selected experiments with three diverse types of devices (using different electrode plate areas, gap lengths, and electrode work-functions), configured both in the triode or the (single) diode (e.g. **Fig.10B**) arrangements, at the indicated pressures, are presented in Table 8. In all cases, a net excess of combined battery pack charge, expressed as electrical watt hours, is registered (columns 8 and 10, Table 8) and the break-even efficiencies are all >100% (column 10). Experimental groups 1 and 2 again demonstrate that, for the same cathode, the interrupted PAGD sequence method of group 2 (1 minute of PAGD function, followed by a 15 minute relaxation, and so on) yields a higher break-even efficiency because of the lower losses registered with this minimal plate heating method (column 10, Table 8). Group 3 of Table 8, shows that the PAGD power production efficiency is also higher for a lower work-function cathode material (H220 vs H34), being subjected to PAGD auto-electronic conditions at a 4-fold lower pressure than the control groups 1 and 2; however, the lower pressure depresses the frequency and, together with the interrupted PAGD sequencing method, it also lowers the loss, causing an actually much larger break-even value than registered for the previous two groups. Groups 4 and 5 exemplify the dual effect of lowering both the plate area and the gap distance: the former affects the PAGD event frequency, whereas the latter affects the PAGD amplitude, and thus the capture efficiency of the charge pack. Despite a cathodic work-function practically and operationally identical to that of groups 1 and 2, these smaller plate area and shorter gap devices utilised in groups 4 and 5, yield 3- to 6-fold lower net power outputs, as well as lower break-even efficiencies, than the former groups, at the same pressure. Finally, group 6 exemplifies the results obtained for the plate diode configuration, where the frequency is lower (no triggering role for the axial member), and a higher loss leads to the lower break-even efficiency, comparable to that of the lower area and shorter gap groups 4 and 5.

In order to verify the discharge curve lengths employed in these analyses and experimentally establish the actual charge capacity of the battery packs, calibration resistive discharges, between the maximum charge state and the minimum limits chosen, were performed for each pack with their respective discharge resistances R2 and R3 (see **Fig.9**). These discharge calibration curves were plotted for half maximal charge values shown in **Fig.15A** and **Fig.15B**, and from the curve produced, we have determined the total half-charge capacities of each battery pack to be 1.033 kWh (100%=2.066 kWh) for the drive pack and 660 WHr (100%=1.320 kWh) for the charge pack. Based upon the corresponding maximal (100%) capacity values, we determined the actual percentages of the relative charge capacities shown in column 5, Table 8, which correspond to the experimental values obtained. We also noted that the curves plotted showed two quite distinct time linear slopes, the slope of the delivery of power per time unit steepening very markedly at the approach to the limits of the permissible withdrawable capacity, occurring at 115 W into **R2**, and 90 W into **R3**.

The pre-PAGD run and post-PAGD run, drive and charge pack discharge curves corresponding to groups 3 and 6, respectively for triode and plate diode configurations, in Table 8, are shown in **Fig.16** (drive pack) and 17 (charge pack), for group 3, and in **Fig.18** (drive pack) and **Fig.19** (charge pack), for group 6. In all cases, the open symbols represent the pre-PAGD run discharge curves, whereas the closed symbols represent the post-PAGD run discharge curves.

As a further check on these values, a videographic, millisecond analysis of the singular power simultaneities occurring at both ends of the system (drive and charge packs) was performed for various 10 second samples of diverse PAGD runs. A typical example is shown in **Fig.20**, which is a sample of the PAGD run designated as 6 in Table 8. While the drive pack DC wattage spent as input to PAGD production varied from 36.6 to 57.82 watts, by a factor of 1.6 times, the DC wattage entering the charge pack as captured PAGD output varied more pronouncedly by a factor of 2.7 times, from 146.4 to 399.6 watts (all meters were in the same selected ranges of voltage and current) with the semi-periodic, intermittent character of each singular emission, though within specific, ascertainable ranges for both amplitude and current outputs.

Assimilation of the singular behaviour of the PAGD in this sample, by a statistical treatment of its variation (with n = 64), indicates that the operational break-even efficiency observed during this sampled period lies at 485.2% +/- 18% with projected 48.3Wh drive pack loss and 221.7Wh charge pack gain. This matches rather closely the observed 483% break-even efficiency, and the 37.7Wh loss as well as the 182.2 kWh gain for the overall PAGD run reported in group 6 of Table 8, and indicates how close are the values obtained by the operational and extensive non-negligible resistive discharge power measurement methods employed.

Finally, an example of the correlation between the drive pack PAGD load voltage and the charge pack PAGD charging voltage, as a function of the duration of the intervening PAGD run between resistive discharge measurements, is shown in **Fig.21**, for the PAGD run corresponding to group 4 of Table 8.

Using the same pulse generator with H200 Al 128 cm^2 plates, in a double diode configuration, and the same circuit values (but with CP = 23 cells), three experiments were conducted at different PAGD frequencies, as a function of varying air pressure. Analysis of driver pack losses and charge pack gains by the extensive load discharge measurement method, as described before, led to the determination of the gross and net gains (respectively, without and with losses included) per pulse, in milliwatt-hour, for each frequency, as well as of the gross and net power gains per second of PAGD operation. The results are shown in Table 9. Even though the gross and net gains of power per pulse were observed to increase with decreasing frequency, the gross power gain per unit time increased with increasing frequency. However, this last trend does not necessarily translate into a higher net gain per unit time, because the losses in the driver pack (not shown) also increase significantly with PAGD frequency. These losses are in all probability related to more energy retention by the plasma at higher frequencies when plasma extinction becomes incomplete. We expect net gains to reach optimal thresholds for any given type of circuit configuration set of values and pulse generator dimensions.

Certain additional observations made during experiments with the double diode configuration of **Fig.10A** may assist in understanding of the invention.

1) Replacing residual air with argon gas leads to higher PAGD frequencies, as noted by us when utilising a 128 cm^2 H200 AC plate pulse generator in the double diode configuration (V = 575). At 1 Torr, the pulsation rate went from 20 PPS in air to 1300-1400 PPS in argon. With 29 12V cells in the charge pack, input currents ceased to flow into it. Under these conditions, the tube potential across the plates decreased and the drop across the input resistor increased. The value of E (= V/d) became smaller (gap size = 3 cm from plate to axial anode collector), as the extinction voltage decreased.

2) With frequencies of 400 PPS, the currents flowing into the charge pack fell to zero. Replacing a fast-recovery type HFR 120 (1200v, 40A) diode bridge by a type MUR 860 (600v, 8A) diode bridge had no effect. When the amplitude of plate potential oscillations falls below the potential of the charge pack, there is also a tendency to produce arc discharges. For output currents from the vacuum pulse generator to enter the charge pack, the number of cells must be reduced so that the potential of the charge pack is low enough to admit the transduced currents. A reduction from 29 to 23 cells allowed currents of 250 mA to enter the CP, and further reduction to 19 cells doubled these currents (per polarity arm).

3) Our observations show that it suffices under these conditions (CP of 19 cells) to increase the vacuum, so that the frequency decreases, and the plate potential and the charge pack input currents increase. At 0.1 Torr, the currents reached 1A DC per plate, and at 0.05 Torr, 2A DC

The interconnection between these factors indicates that the extinction voltage is a function of the PAGD frequency: the higher the PAGD frequency, the lower the extinction voltage, until empirical (in distinction from predicted) VAD field values are reached. As a consequence, the start voltage of the charge pack must be adjusted, by varying the number of cells composing it, so that it lies below the lowest extinction voltage of the PAGD, for any given geometry and gap distance.

Secondly, as the ion plasma is made more rarefied, the frequency of the emissions decreases, but the peak values of the output voltage and current per pulse increase. The slower the PAGD and the more rarefied the atmosphere, the higher is the output energy produced by the system relative to the input energy.

Autographic analysis of PAGD-induced cathode craters in H34 plates was performed, and their average inner diameter and maximum depth were determined. Similar studies were performed for PAGD-induced craters in Alzak (trade mark) plates. The secondary craters characteristically found in Alzak plates, along fracture lines irradiating from the main crater, are absent in H34 plates; instead, in H34 plates, one observes a roughened surface surrounding the emission crater, quite distinct from the original rough aspect of the pulled finish of these hardened aluminium plates. Also, unlike the Alzak main craters, the H34 craters often have a convex centre occupied by a cooled molten metal droplet, whereas the Alzak craters had a concave, hollowed out aspect. Eventually, as the pitting resulting from PAGD cathodic emissions covers the entire cathode, the metallic surface gains a very different rough aspect from its original appearance. In this process, craters from earlier metal layers become progressively covered and eroded by subsequent emissions from the same cathode. Altogether different is the surface deposition process occurring at the anode; here, the surface appears to become more uniform, through the mirroring and possibly abrasive actions of cathode jets. Macroscopically, with increased periods of PAGD operation, the anode surface appears cleaner and more polished.

With the data obtained by the metallographic method of crater measurement, we estimated the volume of metal ejected from the cathode, by assuming that the crater represents a concavity analogous to a spherical segment having a single base (1/6pi x H [$3r^2 + H^2$], where **H** is the height of the spherical segment and **r** the radius of the sphere), while disregarding the volume of the central droplet leftover from the emission. The following are mean +/- SEM crater diameters (D), crater depths (H) and maximum volumes (V) of extruded metallic material for two types of aluminium cathodes, Alzak and H34 hardened aluminium, subject to a high input current PAGD:

1. Alzak: D -0.028 cm +/- 0.003; H -0.002 cm +/- 0.0002; V - 6.2 x 10^{-7} cm^3

2. H34: D -0.0115 cm +/- 0.0004; H -0.0006 +/- 0.0001; V - 3.1 x 10^{-8} cm^3

Accordingly, utilising plates composed of either material with 3 mm of thickness, and thus with a volume of 38.4 cm^3 per plate and considering that only 2/3rds of the cathode shall be used (a 2 mm layer out of the 3 mm thickness), the total number of pulses per plate total (TLT) and partial (PLT) lifetimes is theoretically:

1. Alzak: TLT: 6.2 x 10^7 pulses; PLT: 4.1 x 10^7 pulses;

2. H34: TLT: 1.2 x 10^9 pulses; PLT: 8.1 x 10^8 pulses.

Typically, an H34 device can produce about 0.25 kWh per 10,000 pulses. The corresponding value for a PLT is thus a minimum of 1.0 MWh/Alzak cathode and of 20 MWh/H34 cathode. As the cathode for each combination is only 66.7% consumed, the vacuum pulse generator may continue to be used in a reverse configuration, by utilising the other plate in turn as the cathode; thus, the estimated minimal values become, respectively, 2.0 MWh/Alzak pulse generator and 40 MWh/H34 pulse generator. The same rationale applies for the double diode configuration of **Fig.10C**.

We have created a two-ported system for the production of the singular discharge events which we have previously identified in the “863” application as an endogenous pulsatory abnormal glow discharge regime where the plasma discharge is triggered by spontaneous electronic emissions from the cathode. We have examined the functioning of this two-ported system in order to determine what were the electrical power input and output characteristics of a sustained PAGD regime. Despite the wide (10-fold) variations in net power and break-even efficiencies measured by the four different methods employed (open voltage measurements, time integration of negligible power measurements, operational power measurements and real time non-negligible power measurements), all methods indicate the presence of an anomalous electrical transduction phenomenon within the vacuum pulse generator, such as can result in the production at the output port of electrical energy measured and directly captured which is greater than would be anticipated having regard to the electrical energy input at the input port. With the most accurate of the methods employed, we have found typical PAGD power production rates of 200 WHr/hour of PAGD operation, and these may reach >0.5 kWh/h values.

The discrepancies between the methods utilised have been extensively examined in the preceding section. Our systematic approach demonstrates that the most frequently employed method of measuring the charge capacity of batteries by the open voltage values is the least reliable approach for the determination of the actual net power lost or gained by the battery packs used in the system: when compared to all three other methods, it overestimates net power consumed and produced by up to 10 fold, as well as distorting the break-even efficiencies, particularly at the extremes of operation. All this results from the grossly diminished (50-60% of manufacturer's theoretical estimate) effective charge capacity of the lead acid gel cells employed, as determined experimentally from **Fig.18** and **Fig.19**, when compared to the theoretical maximal charge capacity values that serve as scale for the open voltage measurements. In other words, the effective energy density of the batteries during these experiments was in fact approximately half of the manufacturer's estimated 30 WHr/kg.

Under these actual conditions of battery performance, the third and fourth methods (respectively, operational and real-time non-negligible power measurements) of power consumption and production proved to be the best approach to measure both PAGD electrical power input and output, as the results of both methods matched each other closely, even though the former is a statistical treatment of simultaneous events and the latter is a real time integration of their cumulative effects. The second method is clearly less reliable than either the third or the fourth methods, and this stems from the fact that the power consumption slopes of negligible resistive discharges not only are very different from the quasi-steady state discharge slopes (beginning at >5 - 15 minutes) of extensive resistive discharges, but also their proportionality may not reflect the real time proportionality of equivalent prolonged resistive discharges.

The main advantage of the fourth method is that it effectively takes into account the actual time performance of the batteries comprised by the overall PAGD production and capture system we have described. As such, the method may have the main disadvantage of reflecting more the limitations of the batteries employed (their high

rate of degradation of the absolute value of total effective charge capacity, and limited efficiency in retaining charge derived from discontinuous input pulses) than indicating the actual power output. There are a number of possibilities for fine tuning of the system introduced by the present work, beginning with the utilisation of secondary batteries or other charge shortage or absorption devices that have less variable or more easily predictable actual charge capacity.

In this respect, there are two major shortcomings to the batteries used to form the drive and charge packs; (1) their significant memory effect and (
2) their design for constant, rather than discontinuous, DC charging.

Recently developed Nickel Hydride batteries are an example of an electrostatic charge-storage system that lacks a substantial charge memory effect, and their experimental batteries are being developed currently for higher efficiency intermittent charging methods. Electrostatic charge retention systems having better energy densities, better charge retentivities and insignificant memory effects will probably be more efficient at capturing and holding the energy output by the circuit. In practical embodiments of the invention, effectiveness in charge utilisation will be more important than measurability, and any device that will use the energy effectively whilst presenting an appropriate back EMF to the system may be utilised.

The effect of the performance characteristics of the drive and charge packs is only one amongst many parameters affecting operation of the invention. As shown by our extensive investigation of the diverse PAGD phenomenon the recovery of energy from it by electromechanical transduction as in the “531” application, or electrostatic capture as described above, the factors involved in modulating the frequency, amplitude and peak current characteristics of the PAGD regime are complex. Manipulation of these factors can improve electrical energy recovery, or reduce it or even suppress PAGD. We have so far noted numerous factors that affect PAGD frequency and some amongst those that also affect the PAGD amplitude. Aside from these factors, the circuit parameters of the output port portion of the circuit, in addition to the nature and chemical characteristics of the battery cells already discussed, the charge potential of the charge pack, the characteristics of the rectifiers in the recovery bridge in relation to the period of PAGD super-resonant frequencies, and the effective values of the parallel and transversal capacitance bridges can all influence the results achieved. Certain factors however have a radical effect on PAGD operation, such as the gap distance and the charge pack potential.

Too small a gap distance between the cold emitter (cathode) and the collector will result in an increasing reduction in energy recovery. The potential presented by the charge pack must be less than the voltage amplitude developed by the PAGD, as specified by a given gap distance at a given pressure. Too large a charge pack size with respect to PAGD amplitude and the gap length will preclude PAGD production or result in extremely low PAGD frequencies. In brief, the energy absorption rate and the counter potential presented by the charge pack or other energy utilisation device are important factors in the operation of the circuit as a whole, and should either be maintained reasonably constant, or changes should be compensated by changes in other operating parameters (as is typical of most power supply circuits).

Since our test results indicate that the electrical power output of the circuit can be greater than the electrical power input to the circuit, the circuit clearly draws on a further source of energy input. Whilst we do not wish to be confined to any particular theory of operation, the following discussion may be helpful in explaining our observations. These observations have been discussed in some detail so that the phenomenon observed can be reproduced, even if the principles involved are not fully understood.

In the “863” and “531” applications we have identified a novel, cold-cathode regime of vacuum electrical discharge, which we have termed the pulsed abnormal glow discharge (PAGD) regime. This regime, which occupies the abnormal glow discharge region of the volt-ampere curve of suitable discharge tubes, has the singular property of spontaneously pulsing the abnormal glow discharge in a fashion which is coming from the tube and its circuit environment that constitutes a vacuum pulse generator device, when it is operated under the conditions which we have identified. In fact, when stimulated with continuous direct current, in such conditions, such a circuit responds with spontaneous abnormal glow discharge pulses that enable effective segregation of input and output currents.

We have demonstrated electrically, metallographically, oscillographically and videographically, how the pulsed discontinuity results from a self-limiting, auto-electronic cathode emission that results in repeated plasma eruptions from the cathode under conditions of cathode saturated current input. The auto-electronic triggering of the PAGD regime is thus akin to that of the high-field emission mechanism thought to be responsible for vacuum arc discharges (VAD regime). However, under the PAGD conditions we have defined, this mechanism is found to operate in the pre-VAD region at very low field and low input average direct current values, with very large inter-electrode distances and in a self-limiting, repetitive fashion. In other words, the PAGD regime we have identified has mixed characteristics: its current versus potential (abnormal glow) discharge curve is not only distinct from that of a vacuum arc discharge, but the electrical cycle of the PAGD regime itself oscillates back and forth within the potential and current limits of the abnormal glow discharge region, as a function of the alternate

plasma generation and collapse introduced by the discontinuous sequencing of the auto-electronic emission process. Accordingly, the intermittent presence of the abnormal glow, as well as the observed segregation of the current flows, are due to the diachronic operation of these spontaneous cathode emission foci. The micro-crater and videographic analyses of the PAGD have demonstrated the presence of an emission jet at the origin of each pulse, a phenomenon which VAD theory and experiment has also identified. Metallic jets originating at the cathode spots of VADs have been known to present velocities up to, and greater than 1000 m/sec.

In light of the above, the energy graft phenomenon we have isolated would have to be operated, at the micro-event scale, by the interactions of the cathode emission jet with the vortex-formed impulse-transducing plasma in the inter-electrode space. Several aspects can be approached in terms of the complex series of events that constitute a complete cycle of operation, on a micro-scale. There are interactions within the cathode, interactions at the cathode surface, interactions between the emission jet and the plasma globule close to the cathode, and finally, interactions of the resulting electron and ion distributions in the inter-electrode plasma, within parallel boundaries.

In general, in the presence of an electrical field, the distribution of potential near the cathode forms a potential barrier to the flow of electronic charge, as this barrier is defined by the energy that the most energetic electrons within the metal (the Fermi energy electrons) must acquire before freeing themselves from the cathode surface potential, to originate an emission jet. Before any free electrons become available for conduction in the space adjoining the cathode, they must cross the boundary posed by the potential barrier. With a weak applied field, classical electron emission from a metal can only occur if an energy practically equal to the work-function of the metal is imparted in addition to the Fermi energy. Under thermionic conditions of emission, the heating of the cathode provides the needed energy input. However, the cold-cathode Fowler-Nordheim quantum-field emission theory predicted the existence of a finite probability for an electron to tunnel through the potential barrier, when the applied field is high. Cold-cathode electron emissions are thus possible, under these conditions, at practically Fermi energy levels, as the high field would catalyse the tunnelling through the potential barrier by narrowing the barrier width for the Fermi energy electrons. The exact localisation of the emission would then depend on the randomised fluctuations of high fields at the cathode, which were produced by positive space charges sweeping in proximity to it.

For most purposes, this theory has been the working hypothesis of the last 60 years of field emission studies, which have centred upon the VAD mechanism, despite the fact that observed field gradients are evidently inadequate to explain breakdown as a function of the theoretical high field mechanism. The Fowler-Nordheim theory has therefore suffered major revisions and additions, mostly to account for the fact that it postulates, as a condition for cold-cathode field emission in large area electrodes, the presence of enormous fields ($>10^9$ V/m) and extremely low work functions, neither of which are borne out by experimental VAD investigations. Some researchers have found that the breakdown responsible for the VAD field emission is promoted by Joule heating and vaporisation of microscopic emitter tips, and that this requires a critical current density (10^{12} A/cm^2), while others emphasised that this explanation and these thresholds did not hold for large area emitters and that a space charge effect of concentrating the ion distribution near the cathode promoted breakdown under these circumstances, when the field reached a critical value; large field enhancement factors (more than a thousand-fold) have been postulated to explain the discrepancy between theoretical predictions and experimental findings regarding the critical breakdown field values, and others have demonstrated how this critical field value effectively varies with work-function and electrode conditioning.

The PAGD regime and its self-extinguishing auto-electronic emission mechanism stands as an exception to the high field emission theory as it currently stands with all its modifications, especially given that in this phenomenon we are confronted with a cathode emission that spontaneously occurs across the large gaps in large plate area pulse generators, at very low field values (down to $<1 \times 10^4$ V/m), as shown above and in the “863” application. Moreover, a Fowler-Nordheim plot (in the form Log_{10} (I/V^2) versus 1/V) of the PAGD volt-ampere characteristic exhibits a positive slope, rather than the Fowler-Nordheim negative slope characteristic of VAD field emission. However, current density values obtained from correlations of autographic analysis of the cathode with an analysis of event-oscillogram (peak pulse currents), indicate that the PAGD current density J may reach values of 10^5 to 10^7 A/m^2 during the emission process (the larger Alzak craters have an associated lower J value), values which, at the upper end, do not reach the 10^9 A/m^2 current density threshold required by the Fowler-Nordheim theory. Considering these two distinct observations with regards to field strength and current density, we have to admit the existence of a low field, large area cold-cathode auto-electronic emission endowed with high current densities, which is not predicted by current field emission theory.

Unlike the typical VAD regime, the PAGD is neither a high frequency oscillation, nor does it occur in a random fashion. It constitutes a semi-regular, quasi-coherent, periodic energy transduction which cycles between cathode drop limits that are higher by a factor of 2 to 15 than typical vacuum arc cathode drops. The intermittent cathode emission responsible for the low frequency, pulsed behaviour of the abnormal glow, is also self

extinguishing and self-starting, under the conditions we have defined. Furthermore, we have also identified a novel and unexpected dependency of the periodic pulse rate upon the cathode area. This indicates the presence of field emission control parameters heretofore unsuspected. It is likely that field fluctuations of the polarised pre-breakdown field is responsible for eliciting the particular localisations of the auto-electronic emission foci, as well as what imparts, in a lens-like fashion, the distorted field energy needed for electron surface release. In this sense, external, electrical or magnetic field fluctuations (e.g. motion of static charges or of constant magnetic fields) induced by us at pre-breakdown potentials, provoked PAGD emissions and breakdown at these levels.

In general, VAD studies have shown that, for large area electrodes, microgeometry, adsorbed gas layers and gas impurity contents of the cathode play a role in modulating field emission. In our PAGD studies, the interactions at the cathode surface and across the cathode potential drop are clearly modulated by:
(1) the nature of residual gases, as shown by our air vs Argon studies;
(2) their pressure,
(3) electrode conditioning,
(4) work-function and
(5) cumulative pulse count, amongst others.

The plasma, in leak-controlled or low pressure PAGD devices, has both residual gas and metallic vapour substrates. In devices initially closed at high to very high vacua (diffusion pump pressures), the major residual substrate, whose presence increases with time of operation, is the metallic vapour released from the cathode and not impacted on to the envelope walls or the anode. It has been previously shown for externally (magnetically or electrostatically) pulsed plasma accelerators, that the amount of residual gas or vapour left in the inter-electrode space diminishes with increasing number of consecutive discharges and a growing amount of electrode-insulator absorption of gas. The effect of such removal of residual gas or vapour is to decrease the vacuum of a sealed envelope. With high vacuum sealed PAGD generators we have observed that prolonged operation and sputter-induced mirroring of the envelope causes a progressive disappearance of the discharge, as the voltage potential needed to trigger it also increases. At the thermocouple, low frequency pulsed abnormal glow discharges can also be seen to increase the vacuum significantly. These results suggest instead the presence of a pumping mechanism in the PAGD which is somewhat analogous to that of sputter ion pumps, where collision of ionised gas molecules with the cathode is responsible for the sputtering of cathode material that either combines with the gas substrate (`gettering` action) or `plasters over` the inert gas molecules on to the anode (a process known as `ion burial`). These are the two basic pressure reducing actions of sputtered getter atoms, in ion pumps.

However, in ion sputter pumps, the initiation of the cycle is a function of the presence of high velocity electrons in the high field plasma of the glow discharge, which are necessary to ionise the gas substrate molecules; also, the getter material typically has a high work-function for field emission. Hence, the sputtering is due to the secondary impact of plasma positive ions at the cathode, after plasma ionisation has occurred in the inter-electrode space. Altogether different is the mechanism of spontaneous, primary electron emission from the cathode, which is characteristic of the low field PAGD: here, the sputtering is caused by the electronic emission itself and attendant metallic vaporisation processes. By artificially confining the firing foci to a part of the cathode, we have shown in the single diode configuration how the PAGD induced sputtering is associated with the cathode auto-electronic emission mechanism, rather than with the abnormal cathode glow per se, given the localisation of sputtering on to the emission region of the plate, despite its overall cathode glow saturation.

These observations would thus seem to corroborate the hypothesis of a progressive vacuum increase with the cumulative number of emitted pulses, were it not for the fact that experiments performed with leak controlled devices (reported here and in previous studies) show that, when the negative pressure is maintained by balanced leak admission of air or argon, pulse rates still decrease with cumulative pulse count, and do so neither as a function of an increase in vacuum, nor as a function of envelope mirroring (unless this is so extensive as to establish envelope conduction), but rather as a function of processes (generally referred to as conditioning) inherent to the electrodes, specifically, to the cathode. We have further shown that, for such altered emitter states, the pressure of the vessel must be increased, not because of an increasing vacuum (precluded by the controlled gas leak), but because of the effect that residual gases may have in modulating the low field PAGD emission.

PAGD electrode conditioning is a cathode-dominant process resulting from the cumulative emission of high numbers of pulses by a cathode, and has been shown to be a factor independent of the nature and pressure of the residual gas and partially reversible only by operation with reversed plate polarity, unlike reports of copper cathode-dominant conditioning. It is thought that electrode conditioning and the accompanying increase in VAD breakdown potential are due to the progressive adsorption of residual gases, though cathode-dominant conditioning processes, such as subjecting the vacuum gap to consecutive discharges, have been shown to correlate the decrease in plasma impulse strength with electrode outgassing of absorbed or adsorbed gases. Moreover, given the pitting action of crater formation at the cathode by the PAGD regime, and, as we shall see below, the metallic plating of the anode, the PAGD cathode-dominant process of conditioning we have observed with respect to decreased pulse frequency and increase in potential, suggests that the apparent increase in

cathode work function is not due to gas adsorption or absorption. These processes are more likely to occur on the plated anode. It is likely that, given the observed PAGD pressure reducing effect caused by the cathodic jet, a certain outgassing of the cathode is in fact occurring during PAGD function.

One might also expect that the anode, if plated by sputtering atoms, would increase its gas content in the formed surface film. However, controlled leak experiments suggest instead that some other type of alteration of the cathode work function is occurring, which is, as we shall examine below, independent of the adsorbed gas state of the electrodes, as well as independent of the PAGD ion pump-like effect. Nonetheless, even at the level of the anode, the PAGD sputtering action may have contradictory effects: it may impact inter-electrode gap molecules on to the collector, as well as release, by ionic bombardment and vaporisation, gases adsorbed to, or contaminating the anode. If we assume that gas adsorption by impact on the collector is the predominant mechanism, one could explain the increase in the number of breakdown sites per unit time, as observed by us for a re-reversed cathode, if the number of PAGD breakdown sites depended on the quantity of adsorbed gases, e.g. oxygen, on the cathode being tested. Recovery of the cathode work-function would depend on the electronic charge recovery of the positively charged, adsorbed or occluded gas layer at the cathode- either by reversal or as a function of time of inactivity.

The surface film theory of “electrical double layer formation at the cathode” in fact contended that, low field flash over is a photocathodic effect dependent upon the presence of a glowingly positively polarised gaseous film at the cathode; this film would lower the cathode emissivity by decreasing the field between the cathode surface and the leading edge of the cathode glow, across the cathode drop. However, even though the surface film theory of “electrical double layer formation at the cathode” predicts the lowering of the emission breakdown potential and the increase in flash over rate when the electrodes are reversed - as the anode would have acquired a surface charge capable of affecting the breakdown potential, it acknowledges nevertheless, that the anodic surface charge hardly explains the observed intensity of the polarisation effects.

Moreover, non-reversed, conditioned cathodes retained their lower PAGD frequencies in a time-independent manner, for as long as reversal was avoided (excluding a PAGD frequency recovery effect due to plate cooling, which may be as short as 15 minutes). PAGD conditioning was independent of idle time and increased with cumulative pulse count. Moreover, the AGD pulses are not UV photocathodic Townsend discharges, liberating secondary electrons via positive ion impact at the cathode. Nor could photocathodic emissions generate currents of the magnitude observed in the PAGD. Lastly, the PAGD discharge and breakdown thresholds appear to be unaffected by UV, though they may be somewhat depressed by visible light, and the emission mechanism in the PAGD is the primary process.

Removal or flattening of protuberances and tips from the emitting cathode by the action of the discharge, is a process also thought to play a role in hardening the cathode or increasing its field emission work-function. However, this explanation may not be adequate for the PAGD emission process, if we consider our metallographic findings of a smoothing action of the discharge at the collector. In fact, it would appear that the flattened, smoother, plated, mirrored and cleaner surfaces subjected to PAGD bombardment are the explanation for the observed increased emission ability of re-reversed cathodes: mirrored Alzak surfaces emit at higher frequencies than do dull H34 and H220 surfaces; new, polished surfaces emit at a higher frequency than do pitted, broken-in surfaces; anode surfaces, never before utilised as cathodes but subjected to prolonged PAGD action, emit at higher frequencies when employed as cathodes, than do new, identical cathode surfaces; and ex-cathodes, employed for prolonged periods as anodes, regain a higher emission frequency upon re-use as cathodes. The better PAGD emission performance of smoother cathodes, compared with the worse VAD emission performance of the same, when pitted cathodes (lacking protuberances) are used, requires explanation.

Rakhovsky has put forth a VAD model for cathode spots, that distinguishes between Type I spots (quickly moving spots, far from steady state and responsible for crater formation), and Type II spots (quasi-stationary and near steady-state, but leaving an itinerant track with no sign of crater formation). Whereas the former would obey the Fowler-Nordheim requirement for high fields ($>10^9$ V/m), the latter could hardly be expected to do so with typical arc voltage drops in the order of 10 V. Once again, autographic analysis of the PAGD emission aspect indicates mixed characteristics: the PAGD cathode spot is a hybrid. It behaves as an intermittent instability that leaves single (e.g. in H34) or clustered (e.g. in Alzak) craters, which are both qualities of Type I cathode spots; and it exists under low field conditions ($<10^5$ V/m), with cathode drops of 20 to 150 V, in a quasi-coherent mode, leaving an itinerant track of successive craters when operating at the higher frequencies, all of which are properties approaching those of a VAD Type II cathode spot.

Furthermore, the macroscopically visible metal sputtering (due to the explosive action of the PAGD emission phenomenon) occurring at the upper end of the permissible DC current input scale, and the presence of large solidified molten metal droplets in and around the craters, suggest models which have been proposed for explosive electronic emission. Explosion models propose that the creation of a residual plasma ball in front of a microprotuberance provokes the large potential drop at the prospective emission focus and sufficiently high

resistive and Nottingham heating to reach $>10^7$ A/cm^2 current densities during the explosive consumption of these microemitters. Whether the explosive action associated with cathode spots is an auxiliary effect that applies solely to the vaporisation of the emitting microprotrusion, or an integral emission and vaporisation explosive process, it does not appear that it can be restricted to high-field VAD Type II cathode spots, given that it can be equally made to occur with the low field PAGD hybrid cathode spot, and be macroscopically observed. Indeed, in the plate diode configuration, it is easy to visualise the metallic particle explosions that surround and accompany the plasma jets, near to upper current limit conditions. However, if we are to assume that any of these models apply to the emission mechanism, we would, in all likelihood, have to conclude that the PAGD initial emission sites must be sub-microscopic (100 to 10 nm), rather than microscopic.

Resolution limits to our own metallographic examination of the smoothing action of the PAGD discharge on the collector would thus have precluded us from detecting formation of such sub-microscopic protrusions, as well as their presence in a “soft” cathode and thus infer their disappearance from a pitted, hardened cathode; but if the disappearance of such sub-microprotuberances were responsible for the observed alteration of cathode work function, one would also thereby have to postulate the existence of a mechanism for microroughness regeneration (e.g.. tip growth) at the anode, in order to explain the observed increased emission upon cathode re-reversal. Furthermore, this regeneration would have to be actively promoted by operation with reversed polarity, and this is problematic. Focusing of the distorted or magnified field upon alumina inclusions on pure iron electrodes has been demonstrated to degrade breakdown voltage for field emission, but the effect was greater for larger microscopic particles. If we were to apply this concept to our work, it would require the existence of unmistakably abundant microscopic heterogeneities in the quasi-homogeneous electrode surfaces employed, which we did not observe; on the contrary, their absence suggests that either the microroughness responsible for the low field PAGD emission is sub-microscopic, or that the field distortion responsible for eliciting the PAGD is independent of the presence of these protuberances. This last possibility must be taken all the more seriously, in light of the fact that PAGD functioning is able to cover the entire surface of an emitter with craters.

Whereas the discharge potentials observed in the PAGD have been shown to be relatively independent of the kind of gas present, there is a gas effect in the PAGD phenomenon, particularly in what concerns its frequency, observed when the same “run down” cathode was capable of much higher emission rates when exposed to argon, than to air. Utilising the technique of bias sputtering, it has been demonstrated that the number of charge symmetric collisions (dependent upon sheath thickness **d** and the ion mean free path) in the plasma sheath, which are responsible for lower energy secondary peaks in ion energy distribution **N**(E), at pressures of 0.2 Torr, is substantially greater in argon than in argon-nitrogen mixtures, and thus that, under these conditions, mostly Ar^+ and Ar^{++} ions impact the negatively biased electrode. In non-equilibrium RF discharges, greater ion densities have also been attained with argon, than with air. With respect to field emissions, one would expect a gas effect only with regards to changes on surface conditions, though such studies have shown contradictory effects of argon upon cathode work function.

In light of the foregoing, and given that the PAGD is an emission discharge and not a sputtering discharge per se, in the strict sense, we can conceive of the role of inert gas atoms in increasing, as compared to air or nitrogen, the ion energy density distribution at the PAGD cathode spot interface with the cathode surface emitter, and thus elicit increased emission rates from the cathode, by pulling electrons from the metal via the field effect. While this is consistent with the concept of focused distortions of space-charge field fluctuations inducing localisation of the emission foci, the argon effect can be observed in the PAGD regime over the entire range of the Paschen low vacuum curve, and into Cooke's mid to high vacuum curve, at low fields and without negative biasing. Thus, it is not simply a high pressure (nor a gas conditioning) effect, even if the gas effect in question applies to the description of a local pressure rise at the emission site/cathode spot interface, which may play a role in enhancing the local field.

Considered together, the PAGD emission-derived sputtering, the observed metallic plating of the anode and the explosive aspect of the discharge, suggest the presence of a jet of metallic vapour present in the discharge and running, contrary to the normal flow of positive ions, from the cathode to the anode. This jet appears to have properties similar to the high speed vapour ejected from the cathode in a VAD, as first detected by Tanberg with his field emission pendulum (Tanberg, R. (1930), "On the cathode of an arc drawn in vacuum", Phys. Rev., 35:1080) In fact, the VAD high field emission process is known to release, from the cathode spot, neutral atoms with energies much greater than the thermal energy of the emission discharge. This anomalous phenomenon brings into play the role of the reported cathode reaction forces detected in vacuum arc discharges (Tanberg, as above, also Kobel, E. (1930), "Pressure and high vapour jets at the cathodes of a mercury vacuum arc", Phys. Rev., 36:1636), which were thought to be due to the counterflow of neutral metallic atoms, from the cathode on to the anode (charged metallic ions are normally expected to target the cathode). In absolute units of current, this current quadrature phenomenon has been shown to reach, in the VAD regime, proportions of the order of $100 \times I^2$ (see also the Aspden papers referenced below).

Early interpretations attributed this to the cathode rebounding of <2% of gas substrate-derived plasma positive ions hitting the cathode and being charge-neutralised in the process, but having kept most of their thermal energy. Tanberg held instead that the counterflow of neutral particles responsible for the cathode reaction force was cathode derived, effectively, that it constituted a longitudinal interaction acting in the direction of the metallic arc jet. However, even though secondary high energy distributions of neutral atoms emanating from the cathode do not have thermal energies, their modal distribution does (Davis, W. D. and Miller, H. C. (1969) J. Appl. Phys., 40:2212) furthermore, the major anomalous atomic counterflow that accompanies the high-energy electron flow toward the anode, was shown mass spectrographically to consist predominantly of multiply ionised, positively charged ions of cathode metal, rather than neutral atoms. If this made it easier to abandon the primacy of the rebounding model, it was now more difficult for field emission theorists to accept and explain the observed high energies (ion voltages in excess of the discharge voltage drops) and the high ionisation multiplicity associated with these counter-flowing positive ions.

This field of investigation has indeed been one of the mounting sources of evidence suggesting that there is something amiss in the present laws of electrodynamics. The anomalous acceleration of counter-flowing ions, and the energy transfer mechanisms between high speed or “relativistic” electrons and ions in a plasma (Sethion, J. D. et al, "Anomalous Electron-Ion Energy Transfer in a Relativistic-Electron-Beam-Heated Plasma" Phys. Rev. Letters, Vol. 40, No. 7, pages 451-454), in these and other experiments, has been brilliantly addressed by the theory of the British physicist and mathematician, H. Aspden, who first proposed a novel formulation of the general law of electrodynamics capable of accounting for the effect of the mass ratio factor (M/m') in the parallel (and reverse) motion of charges with different masses, (Aspden, H. (1969) "The law of electrodynamics", J. Franklin Inst., 287:179; Aspden, H (1980) "Physics Unified", Sabberton Publications, Southampton, England). The anomalous forces acting on the counter-flowing metallic ions would stem from their out-of-balance interaction with the emitted high speed electrons, as predicated by the electrodynamic importance of their mass differential. This results in a fundamental asymmetry of the plasma flow between electrodes, localised on to the discontinuous interfaces of the plasma with the electrodes, namely, in the cathode dark space and in the anodic sheath: on the cathode side, electrons act upon ions, as the emitted electrons having less than zero initial velocities, drift against the incoming ion flux and in parallel with the ion and neutral counterflows; on the anode side of the discharge, positive ions flowing toward the cathode confront mainly the incoming counterflow of positive ions and neutral atoms, as the high speed electrons have abnormally transferred their energy to counter-flowing, high speed, cathodic metal ions. An out-of-balance reaction force thus results at the cathode, to which the leaving metallic atoms impart a force of equal momentum but opposite direction, a force which is added to the cathode momentum generated by impacting, normal flowing positive ions.

Moreover, Aspden confirmed theoretically the fundamental contention of Tanberg's experimental findings that an electrodynamic force will manifest itself along the direction of the discharge current flow, and thus, that the atomic counterflow is a metallic jet. Aspden further demonstrated that this asymmetry of plasma discharges does not imply any violation of the principles of conservation of energy and charge equivalence, given that there will be no out-of-balance force when such anomalous forces are considered in the context of the whole system of charge which must, perforce, include the local electromagnetic frame itself. Such discharges must be viewed as open-energy systems, in balance with their electromagnetic environment: their apparatuses may constitute materially closed or limited systems, but they are physically and energetically open systems. Current work on Aspden's formulation of Ampere's Law indicates that both classical electromagnetism and special relativity ignore precisely, in circuits or in plasma, the longitudinal interactions that coexist with transverse ones. Standing longitudinal pressure-waves, of a non-electromagnetic nature, have been previously shown in plasma electrons, which did not conform to the Bohm and Gross plasma oscillation mechanism (Pappas, P. T. (1983) "The original Ampere force and Bio-Savart and Lorentz forces", I1 Nuovo Cimento, 76B:189; Looney, D. H. and Brown, S. C. (1954) "The excitation of plasma oscillations" Phys. Rev. 93:965)

The present theoretical approach to the novel regime of electrical discharge which we have isolated in specially designed devices, and to its mixed glow-arc characteristics, suggests that a similar, out-of balance current quadrature phenomenon occurs in the discharge plasma during the low field, auto-electronic emission-triggered PAGD, and is responsible for the observed surplus of energy in the experimental system described in this report. Clearly, all the evidence we have adduced indicates that there is a powerful longitudinal component to the emission-triggered PAGD, i.e. that the discharge pulses characteristic of this pre-VAD regime are longitudinally propelled jets of cathode-ejected high speed electrons and high speed ions. We have performed experiments, in the PAGD regime of operation, with very thin axial members that bend easily when placed in the path of the discharge, or with Crooke radiometer-type paddle-wheels, and both show the presence of a net longitudinal force in the plasma discharge acting in the direction of the anode, which confirms the magnitude of the atomic counterflow (ionised and neutral) present during the PAGD, very much like Tanberg's pendulum did for the VAD.

These observations also tally with the explosive action of the emission mechanism, such as we have examined it above. In this context, two aspects of the PAGD are remarkable: the fact that a phenomenon akin to field emission occurs at low field values, for large area electrodes across large gaps, and the conclusion that the PAGD must deploy an excessively large counterflow of, in all probability, both ionised and neutral cathodic

particles. The observation of ion current contributions to the cathode current on the order of 8 to 10%, in VADs, can hardly apply to the PAGD mechanism responsible for the anomalous currents and counterflows observed. Hence, we should further expect that the characteristically intermittent, or chopped current regime of the PAGD, is a major factor in the generation of disproportionately high energy longitudinal pulses and in allowing our system to capture most of the electrical energy output from the device. In all probability, field collapse at the end of discharge favours the nearly integral collection of the plasma charge, and ensures the transduction of most of the plasma energy of the pulse (blocked, as it is, from flowing back through the input port to the drive pack) to the output port, through the parallel, asymmetric capacitance bridge that interfaces with the charge recovery reservoir (the charge pack). Collapse of the field of the discharge may also be a contributing factor to the anomalous acceleration of ions, and to the observed anode plating effect.

It is equally possible that such abnormally large longitudinal pulses may never be observable, for a given arrangement and scale, above threshold frequencies of the oscillation; we have, in this sense, presented data that indicates that for a given geometry, above specific PAGD frequencies, the capture of surplus energy decreases steadily in efficiency until it ceases altogether, for a given arrangement. The point at which this surplus begins to decrease coincides with the setting in of frequency-dependent irregularities in the discharge sequence and, most importantly, it coincides with a reduction of the peak pulse current for each PAGD pulse. We have further remarked that increasing the PAGD frequency above the zero surplus point, for a given arrangement, by manipulating any of the frequency control parameters, provokes the slippage of the PAGD into a full fledged VAD regime, while input currents greatly increase and output peak currents greatly decrease (to comparable peak input levels of 10 to 15A).

The transition between the two modes of emission-triggered discharge, PAGD and VAD, thus appears to be tied in to adjustable thresholds in the frequency of the emission discontinuities; in this sense, it is rather likely that the plasma field collapse plays a major role in regularising and optimising the anomalous energies of field emissions, as in the PAGD regime. At low frequencies of low field emission, the emission regime is highly discontinuous, diachronic and regular, for it has time to fully extinguish the discharge; hence the PAGD singularity, in which the phases of each discharge pulse are well defined and sequential. Above a given high frequency, when ion and electron recombination will happen more often, before each can be collected at the electrodes, the stream of emitted discontinuities merges into a noisy, randomised continuum, where simultaneous emissions become possible and the plasma field no longer has time to collapse and fully resolve the longitudinal pulses. Any anomalous energy generated is then minimised and trapped in the plasma body and, in these conditions, the VAD regime eventually sets in. Such model would easily explain why the high field VAD experiments performed to date have never detected such extraordinarily large anomalous forces.

On the other hand, the quasi-coherent aspect of the discharge suggests that the vacuum gap, in functioning during the PAGD regime both as an insulator and as a conductor with capacitative and self-inductive properties, is periodically altered by large and intense polarisations which are resolved by the discrete emission of longitudinal pulses from the cathode. It is possible that these non-linear oscillations resulting from sudden depolarisation of the vacuum gap by high-speed explosive emissions elicited at the convection focus of the distorted field, might be in resonance or near resonance with the external circuitry, but the most apparent effect of increasing the capacitance in all bridge members is to increase the jet current and the transduced current flowing into the charge pack. The PAGD amplitude variation also presents, after the large negative discontinuity, a growing oscillation at very high resonant frequencies, which are typical of inductive chopping currents in a VAD, before extinction occurs. Unlike the VAD inductive case, in the absence of any coils other than the wire wound resistors, the PAGD relaxation oscillations which follow each pulse only extinguish the discharge when the voltage potential of the amplitude curve rises above the applied voltage, just as the plasma potential drops the most.

Given the entirely non-inductive nature of the external circuit utilised in many instances, the inductive properties in evidence are those of the vacuum device itself. It also suggests that, in the absence of any need of an applied external magnetic field for the PAGD discharge to occur coherently, it is possible that the magnitude of the currents generated produces by itself a significant self-magnetic field. Thus, we cannot rule out the possibility of a self-organisation of the plasma discharge, which may, in Prigogine's sense, constitute a dissipative structure (Prigogine, I. and George, C. (1977), "New quantisation rules for dissipative systems", Int. J. Quantum Chem., 12 (Suppl.1):177). Such self-ordering of the PAGD plasma jet is suggested by the experimentally observed transition of these pulses from the current saturated limit of the normal glow discharge region, into the PAGD regime, as a function of increasing current: smaller foci of discharge can be seen to discontinuously agglutinate into larger emission cones, or into jets with a vortex-like appearance, when the input current reaches a given threshold.

It is possible that, under these conditions, the distribution of the charge carriers and their sudden fluctuations may render any steady-state plasma boundary conditions ineffective and provoke a singularity in the discharge mechanism; this non-linear behaviour, together with any self-magnetic effects, might provide radial coherence of the plasma flow along the longitudinal path of the discharge. This concept is akin to what has been proposed for periodically fading-away solution structures referred to as “instantons”, that represent self-organising transitions

between the two states of a system. The PAGD may well be an instance of an instanton type structure bridging the open, or conductive, and the closed, or insulating, states of the vacuum gap. An analytical formulation of the problem of the plasma flow from the cathode spot to the anode, which would take into account the self-magnetic and self-organising properties of the PAGD plasma channel, would be extremely difficult, given the out of balance longitudinal force, its abnormal energy transfer and associated counterflow, as well as the competition between collisional and inertial exchanges.

The plating observed at the anode most likely results from the impact of counter-flowing ions (and possibly neutral atoms), whereas the pitting of the (locally molten) cathode results from the emission of vaporised metallic material and electrons, as well as, secondarily, from bombardment by incident positive ions. The first action smoothes the surface by mirroring it (deposition of cathode-derived atoms) and abrading it, whereas the latter smoothes it in places by rounding concavities and by forming molten droplets upon local cooling, while simultaneously roughening it on the crater peripheries. One might think that this cathode roughening should lower the work function and facilitate the discharge, but the facts indicate that just the opposite must be happening in view of changes in the PAGD according to the nature and state of the cathode surface. The observed alterations of electrode work function for PAGD low field emission must thus be related to the molecular and charge effects of these different actions at the two electrodes. It appears that for large parallel plate electrodes, the PAGD low field emission is modulated by the nature and, most likely, by the molecular structure of the metallic surface layer of the emitter.

We have thus devised a system for the capture, as electricity, of the energy of anomalously energetic longitudinal pulses sequentially triggered by spontaneous emissions of high-speed electrons and ions generated from low work function cathodes, during the low field and singularly mixed PAGD regime of electrical discharge in vacuo. To confirm the above interpretation of the anomalous flux in the observed PAGD phenomenon, the cathode jet composition, as well as time-dependent and usage-dependent changes occurring in the tubes, with diverse sealed negative pressures and after submission to prolonged PAGD operation, must be analysed by mass-spectroscopy. In any event, the excess energy present in the anomalous counter-flowing force appears to stem from a discharge mechanism that effectively pulls high speed electrons and constituent atoms out of a metal surface, at low fields and with high current densities, and is modulated by a complex multiplicity of parameters.

The system described appears to transduce efficiently the observed non-linear longitudinal pulse discontinuities of the plasma field, under conditions of current saturation of the cathode, because the self-extinguishing and self-limiting properties of the discharge allows the energy from the collapse of the discharge to be captured. The particular design of the circuitry, which couples a rectification bridge to the asymmetric bridge quadrature of large capacitances, placed at the output of the PAGD generator, permits effective capture. Our findings constitute striking evidence for Aspden's contention of a need to revise our present electrodynamic concepts. The dual ported PAGD discharge tube circuits which we have described are the first electrical systems we know of which permit effective exploitation of anomalous cathode reaction forces and allow for the recovery of electrical energy from systems exhibiting this effect. Any apparent imbalance in the electrical energy input to the system and withdrawn from the system by its operator must be considered in the context of the entire continuum in which the system operates, within which it is anticipated that accepted principles of energy balance will be maintained.

Moreover, the energy conversion system of the invention has substantial utility as an electrical inverter accepting direct current, and providing one or more of a direct current output at lower voltage and higher current, variable frequency input to alternating current motors, and, by suitable combinations of discharge tube systems, more flexible DC-to-DC conversion systems.

As an alternative to the batteries used in the experiments described, a DC power supply may be utilised or, more advantageously from the viewpoint of entailing less transformation losses, a DC generator to provide the electrical energy input to the system. As a DC motor can be run directly from the rectified output of the circuit of **Fig.9** at **El-E2**, in place of a battery charge pack, DC motor/generator sets of suitable characteristics (in terms of back E.M.F. and circuit loading) can be used to charge the batteries of the drive pack, utilising the rectified PAGD output to drive the DC motor component of the set. This provides a simple, one battery pack solution, where the PAGD input and output circuits are electrically separated by the DC motor/generator interface: the drive pack is simultaneously being discharged to drive PAGD production, and charged by the DC generator output which, in turn, is being driven by the electromechanical transformation of the rectified PAGD output that would typically accrue to a charge pack in the experiments already described. The main limitations to such an arrangement lie in the efficiency of the motor and generator transformations utilised.

A pulsed DC source could be used to provide input to the circuit if suitably synchronised, but care is needed not to interfere unduly with the auto-electronic mechanism of the field induced cathode emissions.

TABLE 1

Results for the ballast resistance (and current) dependent PAGD frequency utilizing an H34 aluminum pulse generator with 128 cm^2 plates at 5.5 cm distance, in the triode configuration, at a pressure of 0.8 Torr. The circuit employed is that of the present invention, as described in the third Results Section. DCV = 560.

R in Ω	Regime of Discharge	Pulse Rate >100 V
5,000	NGD (Cold Cathode)	0
600	PAGD	10 PPS
300	PAGD	40 PPS
150	PAGD	180 PPS
100	VAD	0
50	VAD	0

TABLE 2

128 cm^2 H220 Al; 570 volts DC; 300 Ω = R1; Diode Configuration

	PPS	p(Torr)	Cumulative Pulse Count
1)	200	0.08	$\sim 2.4 \times 10^5$
2)	200	0.5	$\sim 1.5 \times 10^6$
3)	200	0.8–1	$\sim 2.5 \times 106$
4)	25	0.5	3×10^6 pulses
5)	200	0.5	1.5×10^6 (after first electrode reversal)

TABLE 3

pressure in Torr	RESIDUAL GAS EFFECT PPS in AIR	PPS in ARGON
0.45	ND	10
0.5	1.8 ± 0.3	ND
0.55	4.8 ± 0.9	16.7±1.8
1.0	11.4 ± 0.8	448 ± 27.4
1.25	214.5 ± 14.3	ND
2.0	36.2 ± 2.6	206 ± 19.6 158.7 ± 24
2.5	1.36 ± 0.3	0

TABLE 4

Charge pack No. of cells	PPS	PAGD
36	0	–
31	1	+
29	10	+
19	1	+
9	0	–

TABLE 5

1 Expt. No.	2 Battery Pack	3 Position	4 Open Voltage	5 V/cell	6 % total rel. cpty.	7 Max. hr. left	8 % rel. cpty gained	8 % rel. cpty lost	9 Total kWh	10 ΔkWh gain	10 ΔkWh loss	11 PAGD per sec
1	Charge	start	348	12.0	40	8			0.835			8
	Charge	end	366	12.62	83	16.6	43		1.823	0.988		
	Driver	start	576	12.52	77	15.4			2.660			
	Driver	end	572	12.43	70	14		7	2.402		0.258	
2	C	b	331	11.41	2	0.4			0.040			61
	C	a	351	12.1	47.5	9.5	45.5		1.002	0.962		
	D	b	553	12.02	40	8			1.327			
	D	a	546	11.9	33	6.6		7	1.081		0.246	
3	C	b	345	11.9	32.5	6.5			0.673			3
	C	a	361	12.45	72.5	14.4	40		1.559	0.886		
	D	b	559	12.15	51	10.2			1.710			
	D	a	552	12.0	40	8		11	1.324		0.386	
4	C	b	360	12.41	70	14			1.512			32
	C	a	373	12.86	103	>20	33		2.238	0.726		
	D	b	562	12.22	54.5	10.9			1.838			
	D	a	557	12.11	48	9.6		6.5	1.604		0.234	
5	C	b	340	11.7	20	4			0.408			2
	C	a	365	12.59	83	16.6	63		1.818	1.440		
	D	b	527	11.45	3.2	0.6			0.101			
	D	a	517	11.24	1.8	0.4		0.2	0.056		0.045	
6	C	b	340	11.72	21.5	4.3			0.438			8
	C	a	367	12.66	87.5	17.5	66		1.927	1.489		
	D	b	589	12.8	100	20			3.530			
	D	a	564	12.26	58.5	11.7		41.5	1.979		1.551	
7	C	b	318	10.97	1.2	0.24			0.023			5
	C	a	359	12.38	67.5	13.5	66.3		1.454	1.431		
	D	b	575	12.5	77	15.4			2.656			
	D	a	567	12.32	63.5	12.7		13.5	2.160		0.496	
8	C	b	328	11.71	20	4			0.393			32
	C	a	350	12.5	76.5	15.3	56.5		1.606	1.213		
	D	b	582	12.65	87.5	17.5			3.055			
	D	a	579.5	12.60	84	16.8		3.5	2.921		0.134	

1 Expt. No.	2 Battery Pack	3 Position	12 Exptl. time	13 rel. kWh/h gain	13 rel. kWh/h loss	14 net kWh/h production	15 Breakeven efficiency	16 Cell #/ pack	17 tube	18 Cathode Area	19 Plate
1	Charge	start	21.5'			2.071	388%	29	A26	128 cm^2	H34
	Charge	end		2.791							

TABLE 5-continued

	Driver	start						46				
	Driver	end			0.720							
2	C	b	18'			2.387	391%	29	A26	128 cm^2	H34	
	C	a		3.207								
	D	b						46				
	D	a			0.820							
3	C	b	21.5'			1.396	230%	29	A26	128 cm^2	H34	
	C	a		2.473								
	D	b						46				
	D	a			1.077							
4	C	b	63.5'			0.465	310%	29	A28	128 cm^2	H220	
	C	a		0.686								
	D	b						46				
	D	a			0.221							
5	C	b	80'			1.064	6,750%	29	A26	128 cm^2	H34	
	C	a		1.080								
	D	b						46				
	D	a			0.016							
6	C	b	21.5'			−0.173	96%	29	A26	128 cm^2	H34	
	C	a		4.155								
	D	b						46				
	D	a			4.328							
7	C	b	64.5'			0.870	289%	29	A45	64 cm^2	H34	
	C	a		1.331								
	D	b						46				
	D	a			0.461							
8	C	b	28.5'			2.272	906%	28	A45	64 cm^2	H34	
	C	a		2.554								
	D	b						46				
	D	a			0.282							

1 Expt. No.	2 Battery Pack	3 Position	20 R1 ohm	21 C3/C5 mfd	22 C7a/C7b mfd	23 Motor arm	24 Pressure	25 Gap cm	26 OV rlx. time	27 C4 mfd	28 R4 ohms	29 Motor rpm

TABLE 5-continued

1 Expt. No.	2 Battery Pack	3 Position	20 R1 ohm	21 C3/C5 mfd	22 C7a/C7b mfd	23 Motor arm	24 Pressure	25 Gap cm	26 OV rlx. time	27 C4 mfd	28 R4 ohms	29 Motor rpm
1	Charge	start	300	20,700	3,300	off	0.8 Torr	5.5	30'	NA	NA	NA
	Charge	end										
	Driver	start										
	Driver	end										
2	C	b	300	20,700	3,300	off	0.8 Torr	5.5	30'	NA	NA	NA
	C	a										
	D	b										
	D	a										
3	C	b	300	20,700	3,300	off	0.7 Torr	5.5	15'	NA	NA	NA
	C	a										
	D	b										
	D	a										
4	C	b	300	34,700	5,500	off	0.2 Torr	5.5	30'	NA	NA	NA
	C	a										
	D	b										
	D	a										
5	C	b	150	34,700	3,300	on	0.8 Torr	5.5	15'	8	500	1,200
	C	a										
	D	b										
	D	a										
6	C	b	300	20,700	3,300	on	0.8 Torr	5.5	15'	16	0	2,000
	C	a										
	D	b										
	D	a										
7	C	b	600	34,700	3,300	off	0.8 Torr	4	30'	NA	NA	NA
	C	a										
	D	b										
	D	a										
8	C	b	600	34,700	5,500	off	0.8 Torr	4	30'	NA	NA	NA
	C	a										
	D	b										
	D	a										

TABLE 6

Expt. No.	Battery Pack	Position	Load Voltage	Watts/ cell	Hr. left	Total kWh	ΔkWh gain	ΔkWh loss	rel. kWh/h gain	rel. kWh/h loss	net kWh/h	B. Eff.
1	C	s	335.7	4.445	4	0.516					3.014	776%
	C	e	357.5	5.05	12	1.757	1.241		3.46			
	D	s	568.0	3.20	13	1.766						
	D	e	564.6	3.175	11	1.606		0.16		0.446		
2	C	s	315.5	3.93	1	0.114					1.012	504%
	C	e	327.8	4.25	4.5	0.502	0.387		1.225			
	D	s	540.7	2.91	4	0.535						
	D	e	535.3	2.87	3.5	0.462		0.073		0.243		
3	C	s	328	4.23	2	0.245					1.175	703%
	C	e	351.7	4.91	7	0.737	0.492		1.370			
	D	s	546	2.95	5	0.680						
	D	s	545.5	2.90	4.5	0.610		0.070		0.195		

TABLE 7

1 Expt. No.	2 Config.	3 Pressure Torr	4 Tube	5 DP DCV	6 Plates DCV	7 DP DCA	8 DP Watts	9 PAGD Volts	10 PAGD V/cm	11 CP DCV
1	dd	0.8	A29	562	350	0.65	137.8	212	77.1	375
2	dd	0.09	A29	562	402	0.60	96	160	58.2	378
3	dd	0.8	A29	560	371	0.59	111.5	189	68.7	374
4	dd	0.09	A29	563	409	0.49	75.9	154	56	379
5	t	1.5	A28	561	439	0.41	49.9	122	22.2	377
6	t	1.5	A28	560	425	0.51	68.9	135	24.5	375
7	t	1.0	A28	556	398	0.48	75	158	28.7	376.5
8	t	0.5	A28	559.5	398	0.68	109.8	161.5	29.4	377.5
9	t	0.5	A28	563	390	0.75	112.45	173	31.5	373
10	sd	0.5	A28	565	422	0.47	67.2	143	26	376
11	sd	0.5	A28	561.5	415	0.50	73	146.5	26.6	380
12	sd	0.5	A28	562	413.5	0.55	81.7	148.5	27	380
13	dd	0.25	A28	553	438	0.35	40	115	41.8	381.5
14	dd	0.25	A28	549	325	0.70	156.8	224	81.5	263

1 Expt. No.	2 Config.	12 CP DCA	13 CP Watts	14 Total Resistance	15 Breakeven Efficiency	16 PPS	17 Bridge diode	18 Input diode	19 Motor status	20 FIG. 3
1	dd	1.25	468.8	326	340%	450	M860	HFR	off	+
2	dd	0.70	264.6	% 270	276%	92	M860	HFR	off	
3	dd	0.65	243.1	243	218%	500	HFR	HFR	off	
4	dd	0.76	288	314	379%	77	HFR	HFR	off	
5	t	0.58	219	298	439%	52	HFR	HFR	off	
6	t	0.69	259	265	376%	100	M860	HFR	off	
7	t	0.57	213.1	329	284%	355	M860	HFR	off	
8	t	0.67	252.9	238	230%	92	HFR	HFR	off	
9	t	0.65	280	266	249%	118	M860	HFR	off	+
10	sd	1.03	387.3	286	530%	25	M860	HFR	off	
11	sd	0.73	277.4	293	379%	11	HFR	HFR	off	+
12	sd	0.71	269.8	270	330%	10	HFR	HFR	on	+
13	dd	0.59	225.1	329	563%	10	HFR	HFR	off	
14	dd	1.36	257.7	320	228%	1	HFR	HFR	off	

TABLE 8

1 Expt. No.	2 Battery Pack	3 Position	4 Total Wh	5 Rel. Cap.	6 Torr	7 Limit in W	8 ΔkWh gain	8 ΔkWh loss	9 Exptl. time	10 abs. kWh/h gain	10 abs. kWh/h loss	net	11 BE
1	C	b	159	12%	0.8	90			21.5′			+664	846%
	C	a	428	32%			269			753			
	D	b	1764	85%		115							
	D	a	1732	84%				32			89		
2	C	b	118	9%	0.8	90			18′			+616	2,667%
	C	a	303.5	23%			192			640			
	D	b	542.3	26%		115							
	D	a	535	25.9%				7.3			24		
3	C	b	950.4	72%	0.2	90			70′			+186	3485%
	C	a	1,161	88%			210.9			191.7			
	D	b	660	32%		115							
	D	a	654	32%				6.5			5.6		
4	C	b	15.8	1.2%	0.8	90			64.5′			+53.7	406%
	C	a	81.9	6%			65			60			
	D	b	181	8.7%		115							
	D	a	165	8%				16			14.7		
5	C	b	34.5	2.6%	0.8	90			28.5′			+169.1	436%
	C	a	138.8	10.5%			104.3			219.6			
	D	b	1,114	54%		115							
	D	a	1,089	53%				24			50.5		
6	C	b	55.4	4.2%	0.8	90			74′			+117	483%
	C	a	237.6	18%			182.2			148			
	D	b	669.3	32%		115							
	D	a	631.7	30.6%				37.7			30.6		

1 Expt. No.	2 Battery Pack	3 Position	12 Config.	13 Tube	14 Cathode area	15 gap cm	16 PPS	17 PAGD seq. method	18 R1 ohms	19 Plate material	20	21 C3/C5 mfd	22 C7a/C7b mfd
1	C	b	Triode	A26	128 cm2	5.5	8	Continuous	300	H34		20,700	3.300
	C	a											
	D	b											
	D	a											
2	C	b	Triode	A26	128 cm2	5.5	61	Interrupted	300	H34		20,700	3,300
	C	a											
	D	b											
	D	a											
3	C	b	Triode	A28	128 cm2	5.5	32	Interrupted	300	H220		34,700	5,500
	C	a											
	D	b											
	D	a											
4	C	b	Triode	A46	64 cm2	4.0	5	Continuous	600	H34		34,700	5,500
	C	a											
	D	b											
	D	a											
5	C	b	Triode	A46	64 cm2	4.0	32	Interrupted	600	H34		34,700	5,500
	C	a											
	D	b											
	D	a											
6	C	b	Plate Diode	A29	128 cm2	5.5	8	Interrupted	300	H220		34,700	5,500
	C	a											
	D	b											
	D	a											

TABLE 9

Utilizing: Al H200, 128 cm^2 plates
DP = 46 cells
CP = 23 cells

	PPS	CP Gain per pulse in mWh	Net Gain per pulse mWh	CP Gain per second mWh	Net Gain per second mWh	Pressure in Torr
#1	1.5	22.3	11.7	33:45	17.55	0.2
#2	8	5.6	4.4	44.8	35.2	0.8
#3	110	0.78	0.27	85.8	29.7	2.0

CLAIMS

1. Apparatus comprising a discharge tube and an electrical circuit containing said discharge tube and configured to operate the latter to provide endogenous pulsatory cold cathode auto-electronic emissions, the circuit being double ported with an input port connected to a source of direct current at a potential sufficient to initiate said emissions, and an output port connected to a current sink effective to absorb at least a substantial portion of electrical energy released by collapse of said emissions.

2. Apparatus according to claim 1 configured so that the emissions occur in a pulsed abnormal glow discharge regime.

3. Apparatus according to claim 2, wherein the input port includes components ensuring that the flow of current therein is unidirectional, and incorporating impedance sufficient to limit the flow of current therein.

4. Apparatus according to claim 2, including capacitors connected to the discharge tube, the input port and the output port, which provide charge storage in the input port and direct current isolation between the input and output ports.

5. Apparatus according to claim 4, wherein the output port comprises a rectifier having an input connected to said capacitors, reservoir capacitance connected to the output of said rectifier, and reverse current blocking devices connected between said reservoir capacitance and the current sink.

6. Apparatus according to claim 5, wherein the rectifier is a bridge rectifier, and the reservoir capacitance is provided by a capacitor bridge having ends connected to outputs of the bridge rectifier, and an intermediate point connected to one input of the bridge rectifier.

7. Apparatus according to claim 4, further including an alternating current motor and a capacitor in series, connected between the connections of said capacitors to the output port.

8. Apparatus according to claim 2, wherein the current sink comprises a secondary battery.

9. Apparatus according to claim 2, wherein the current sink comprises an electric motor.

10. Apparatus according to claim 2, wherein the direct current source comprises a secondary battery.

11. Apparatus according to claim 2, wherein the direct current source is a DC generator.

12. Apparatus according to claim 9, wherein the motor is a DC motor.

13. Apparatus according to claim 10, including a circuit for charging from the output port a battery to be used as the direct current source.

14. Apparatus according to claim 2, wherein the direct current source is a rectified AC source.

15. Apparatus according to claim 2, wherein the discharge tube is connected as a single diode.

16. Apparatus according to claim 2, wherein the discharge tube is connected as a multiple diode with plates connected as cathodes and an intermediate electrode connected as an anode.

17. Apparatus according to claim 2, wherein the discharge tube is connected as a triode, with an intermediate electrode functioning as an auxiliary cathode.

18. Apparatus according to claim 2, wherein a first potential is applied to the input port by the source of direct current to induce emission, a back EMF is applied to the output port by the current sink, and an extinction potential of the emissions is greater than the back EMF.

19. A method of energy conversion, comprising initiating plasma eruptions from the cathode of a discharge tube operating in a pulsed abnormal glow discharge regime utilising electrical energy from a source in a first circuit connected to said discharge tube, and capturing electrical energy generated by the collapse of such eruptions in a second circuit connected to said discharge tube.

20. A method according to claim 19, wherein current flowing into the discharge tube during said eruptions is at least 50 ma.

21. A method according to claim 19, wherein current flowing into the discharge tube during said eruptions is at least 500 ma.

22. A method according to claim 19, in which charge carriers within plasma outputs are accelerated through at least one of an electric and magnetic field.

23. A method of energy conversion, comprising inducing endogenous pulsatory low-field, large-area cold-cathode auto-electronic emissions from the cathode of a discharge tube capable of sustaining such emissions, utilising electrical energy from a source in a first circuit connected to said discharge tube, and capturing electrical energy generated by the collapse of such emissions in a second circuit connected to said discharge tube.

FRANKLIN MEAD AND JACK NACHAMKIN

Patent US 5,590,031 31st December 1996 Inventors: Franklin Mead & Jack Nachamkin

SYSTEM FOR CONVERTING ELECTROMAGNETIC RADIATION ENERGY TO ELECTRICAL ENERGY

This patent shows a system for converting Zero-Point Energy into conventional electrical power.

ABSTRACT

A system is disclosed for converting high-frequency zero-point electromagnetic radiation energy to electrical energy. The system includes a pair of dielectric structures which are positioned near each other and which receive incident zero-point electromagnetic radiation. The volumetric sizes of the structures are selected so that they resonate at a frequency of the incident radiation. The volumetric sizes of the structures are also slightly different so that the secondary radiation emitted from them at resonance, interferes with each other producing a beat frequency radiation which is at a much lower frequency than that of the incident radiation and which is amenable to conversion to electrical energy. An antenna receives the beat frequency radiation. The beat frequency radiation from the antenna is transmitted to a converter via a conductor or waveguide and converted to electrical energy having a desired voltage and waveform.

US Patent References:

3882503	May., 1975	Gamara	343/100.
4725847	Feb., 1988	Poirier	343/840.
5008677	Apr., 1991	Trigon et al.	342/17.

DESCRIPTION

BACKGROUND OF THE INVENTION

The invention relates generally to conversion of electromagnetic radiation energy to electrical energy, and, more particularly, to conversion of high frequency bandwidths of the spectrum of a type of radiation known as ‘zero-point electromagnetic radiation’ to electrical energy.

The existence of zero-point electromagnetic radiation was discovered in 1958 by the Dutch physicist M. J. Sparnaay. Mr. Sparnaay continued the experiments carried out by Hendrik B. G. Casimir in 1948 which showed the existence of a force between two uncharged parallel plates which arose from electromagnetic radiation surrounding the plates in a vacuum. Mr. Sparnaay discovered that the forces acting on the plates arose from not only thermal radiation but also from another type of radiation now known as classical electromagnetic zero-point radiation. Mr. Sparnaay determined that not only did the zero-point electromagnetic radiation exist in a vacuum but also that it persisted even at a temperature of absolute zero. Because it exists in a vacuum, zero-point radiation is homogeneous and isotropic as well as ubiquitous. In addition, since zero-point radiation is also invariant with respect to Lorentz transformation, the zero-point radiation spectrum has the characteristic that the intensity of the radiation at any frequency is proportional to the cube of that frequency. Consequently, the intensity of the radiation increases without limit as the frequency increases resulting in an infinite energy density for the radiation spectrum. With the introduction of the zero-point radiation into the classical electron theory, a vacuum at a temperature of absolute zero is no longer considered empty of all electromagnetic fields. Instead, the vacuum is now considered as filled with randomly fluctuating fields having the zero-point radiation spectrum. The special characteristics of the zero-point radiation which are that it has a virtually infinite energy density and that it is ubiquitous (even present in outer space) make it very desirable as an energy source. However, because high energy densities exist at very high radiation frequencies and because conventional methods are only able to convert or extract energy effectively or efficiently only at lower frequencies at which zero-point radiation has relatively low energy densities, effectively tapping this energy source has been believed to be unavailable using conventional techniques for converting electromagnetic energy to electrical or other forms of easily usable energy. Consequently, zero-point electromagnetic radiation energy which may potentially be used to power interplanetary craft as well as provide for society's other needs has remained unharnessed.

There are many types of prior art systems which use a plurality of antennas to receive electromagnetic radiation and provide an electrical output from them. An example of such a prior art system is disclosed in U.S. Pat. No.

3,882,503 to Gamara. The Gamara system has two antenna structures which work in tandem and which oscillate by means of a motor attached to them in order to modulate the radiation reflected from the antenna surfaces. The reflecting surfaces of the antennas are also separated by a distance equal to a quarter wavelength of the incident radiation. However, the Gamara system does not convert the incident radiation to electrical current for the purpose of converting the incident electromagnetic radiation to another form of readily usable energy. In addition, the relatively large size of the Gamara system components make it unable to resonate at and modulate very high frequency radiation.

What is therefore needed is a system which is capable of converting high frequency electromagnetic radiation energy into another form of energy which can be more readily used to provide power for transportation, heating, cooling as well as various other needs of society. What is also needed is such a system which may be used to provide energy from any location on earth or in space.

SUMMARY OF THE INVENTION

It is a principle object of the present invention to provide a system for converting electromagnetic radiation energy to electrical energy.

It is another object of the present invention to provide a system for converting electromagnetic radiation energy having a high frequency to electrical energy.

It is another object of the present invention to provide a system for converting zero-point electromagnetic radiation energy to electrical energy.

It is another object of the present invention to provide a system for converting electromagnetic radiation energy to electrical energy which may used to provide such energy from any desired location on earth or in space.

It is another object of the present invention to provide a system for converting electromagnetic radiation energy to electrical energy having a desired waveform and voltage.

It is an object of the present invention to provide a miniaturised system for converting electromagnetic radiation energy to electrical energy in order to enhance effective utilisation of high energy densities of the electromagnetic radiation.

It is an object of the present invention to provide a system for converting electromagnetic radiation energy to electrical energy which is simple in construction for cost effectiveness and reliability of operation.

Essentially, the system of the present invention utilises a pair of structures for receiving incident electromagnetic radiation which may be propagating through a vacuum or any other medium in which the receiving structures may be suitably located. The system of the present invention is specifically designed to convert the energy of zero-point electromagnetic radiation; however, it may also be used to convert the energy of other types of electromagnetic radiation. The receiving structures are preferably composed of dielectric material in order to diffract and scatter the incident electromagnetic radiation. In addition, the receiving structures are of a volumetric size selected to enable the structures to resonate at a high frequency of the incident electromagnetic radiation based on the parameters of frequency of the incident radiation and propagation characteristics of the medium and of the receiving structures. Since zero-point radiation has the characteristic that its energy density increases as its frequency increases, greater amounts of electromagnetic energy are available at higher frequencies. Consequently, the size of the structures are preferably miniaturised in order to produce greater amounts of energy from a system located within a space or area of a given size. In this regard, the smaller the size of the receiving structures, the greater the amount of energy that can be produced by the system of the present invention.

At resonance, electromagnetically induced material deformations of the receiving structures produce secondary fields of electromagnetic energy therefrom which may have evanescent energy densities several times that of the incident radiation. The structures are of different sizes so that the secondary fields arising therefrom are of different frequencies. The difference in volumetric size is very small so that interference between the two emitted radiation fields, and the receiving structures at the two different frequencies produces a beat frequency radiation which has a much lower frequency than the incident radiation. The beat frequency radiation preferably is at a frequency which is sufficiently low that it may be relatively easily converted to usable electrical energy. In contrast, the incident zero-point radiation has its desirable high energy densities at frequencies which are so high that conventional systems for converting the radiation to electrical energy either cannot effectively or efficiently so convert the radiation energy or simply cannot be used to convert the radiation energy for other reasons.

The system of the present invention also includes an antenna which receives the beat frequency radiation. The antenna may be a conventional metallic antenna such as a loop or dipole type of antenna or a rf cavity structure

which partially encloses the receiving structures. The antenna feeds the radiation energy to an electrical conductor (in the case of a conventional dipole or comparable type of antenna) or to a waveguide (in the case of a rf cavity structure). The conductor or waveguide feeds the electrical current (in the case of the electrical conductor) or the electromagnetic radiation (in the case of the waveguide) to a converter which converts the received energy to useful electrical energy. The converter preferably includes a tuning circuit or comparable device so that it can effectively receive the beat frequency radiation. The converter may include a transformer to convert the energy to electrical current having a desired voltage. In addition, the converter may also include a rectifier to convert the energy to electrical current having a desired waveform.

BRIEF DESCRIPTION OF THE DRAWINGS

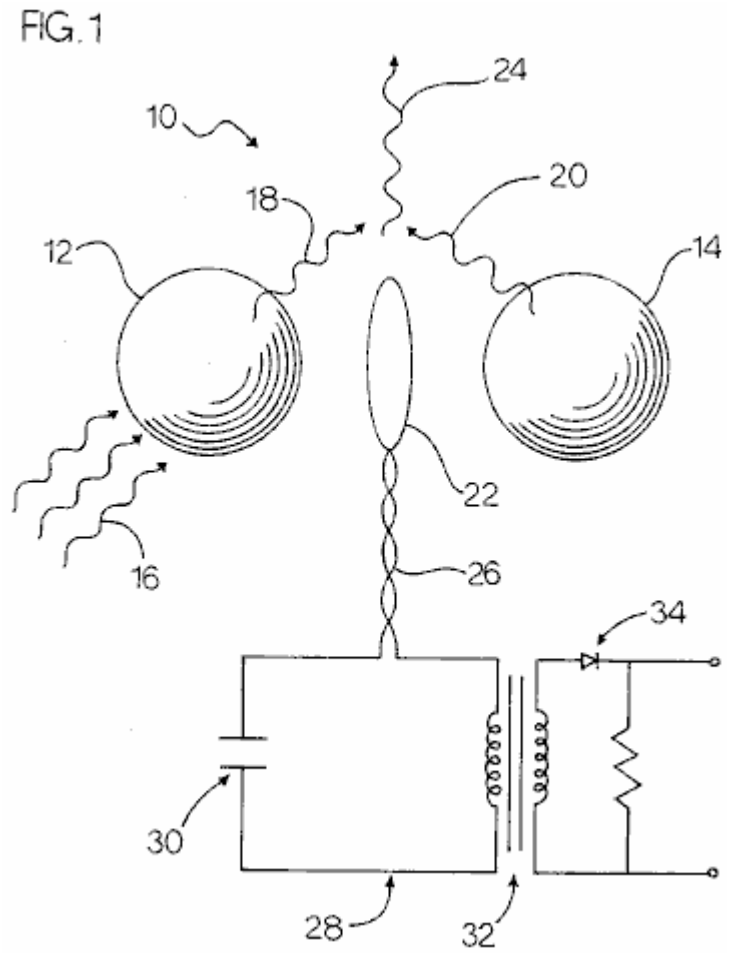


Fig.1 is a plan view of the receiving structures and antenna of a first embodiment of the system of the present invention with a schematic view of the conductor and converter thereof and also showing the incident primary and emitted secondary electromagnetic radiation.

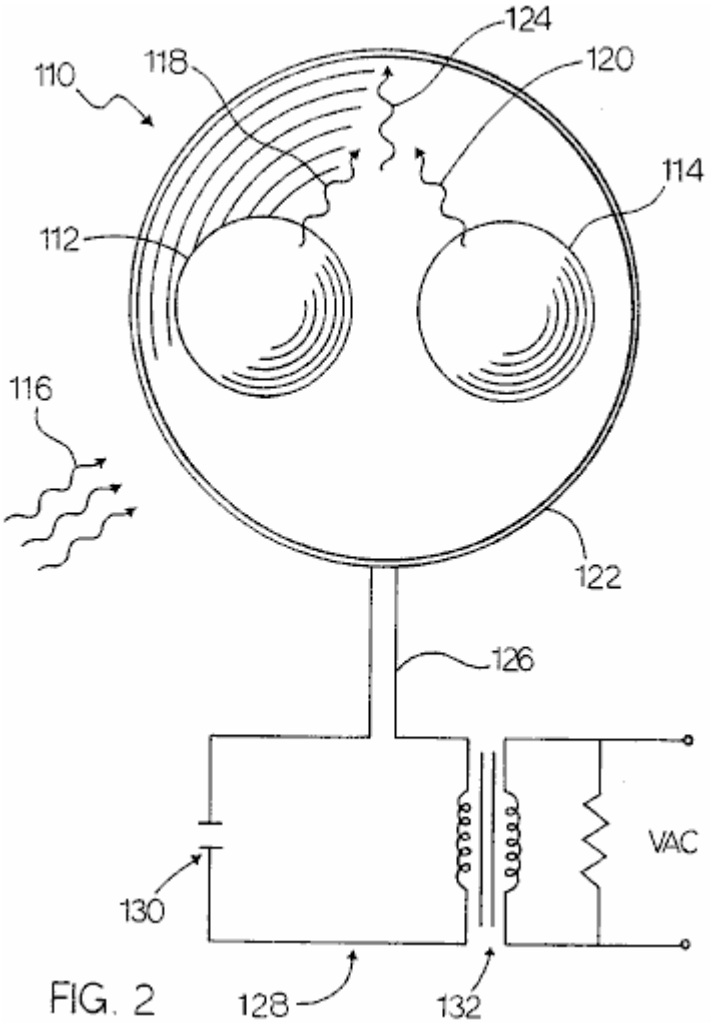


Fig.2 is a front view of the receiving structures, antenna and waveguide of a second embodiment of the system of the present invention with a schematic view of the converter thereof and also showing the incident primary and emitted secondary electromagnetic radiation.

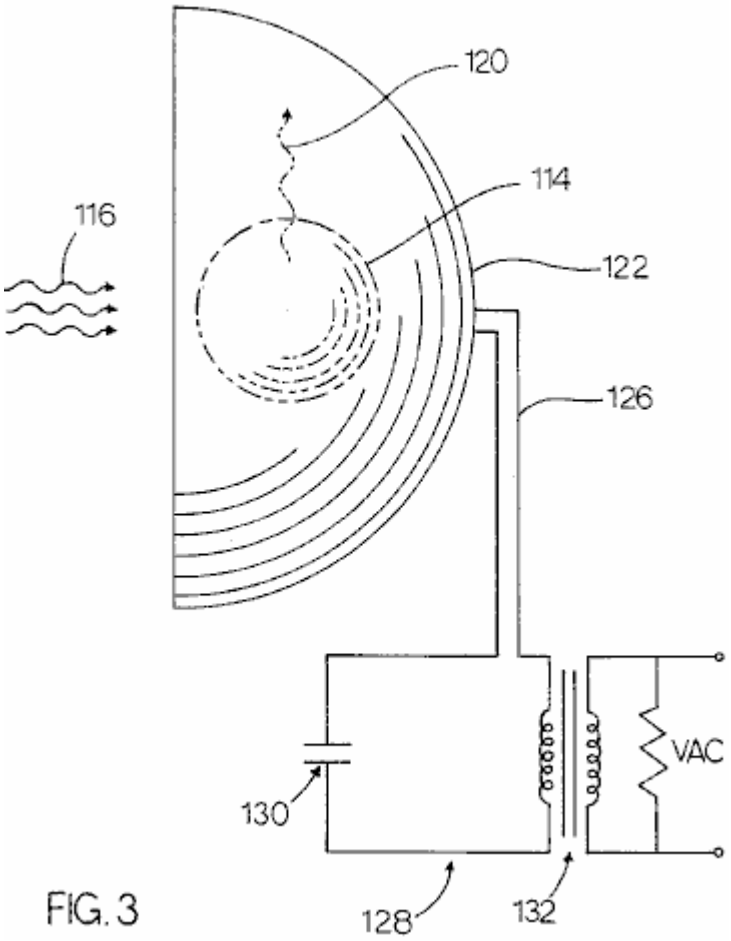


FIG. 3

Fig.3 is a perspective view of the receiving structures, antenna and waveguide of the second embodiment shown in **Fig.2** with a schematic view of the converter thereof and also showing the incident primary and emitted secondary electromagnetic radiation.

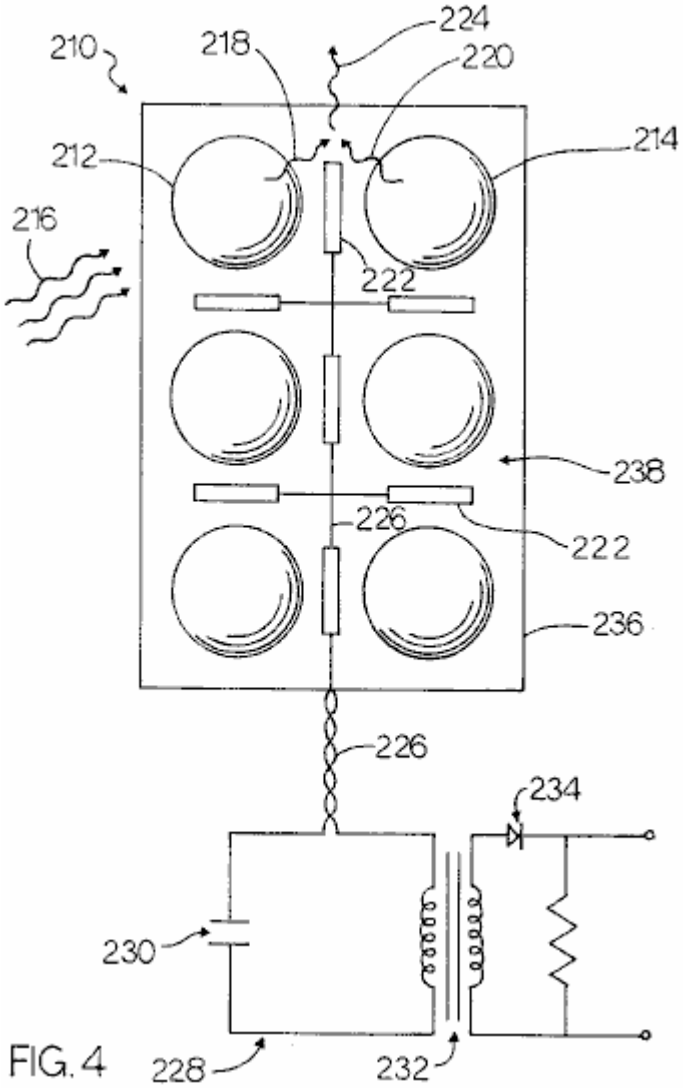


FIG. 4

Fig.4 is a front view of the substrate and a plurality of pairs of the receiving structures and a plurality of antennas of a third embodiment of the system of the present invention with a schematic view of the conductor and converter thereof and also showing the incident primary and emitted secondary electromagnetic radiation.

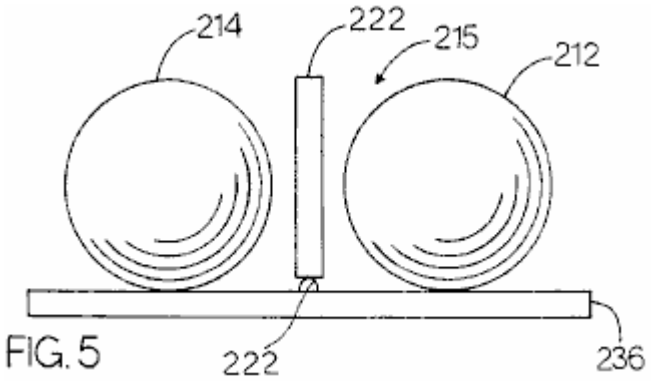


FIG. 5

Fig.5 is a top view of some of the components of the third embodiment of the system of the present invention showing two of the plurality of pairs of receiving structures and two of the plurality of antennas mounted on the substrate.

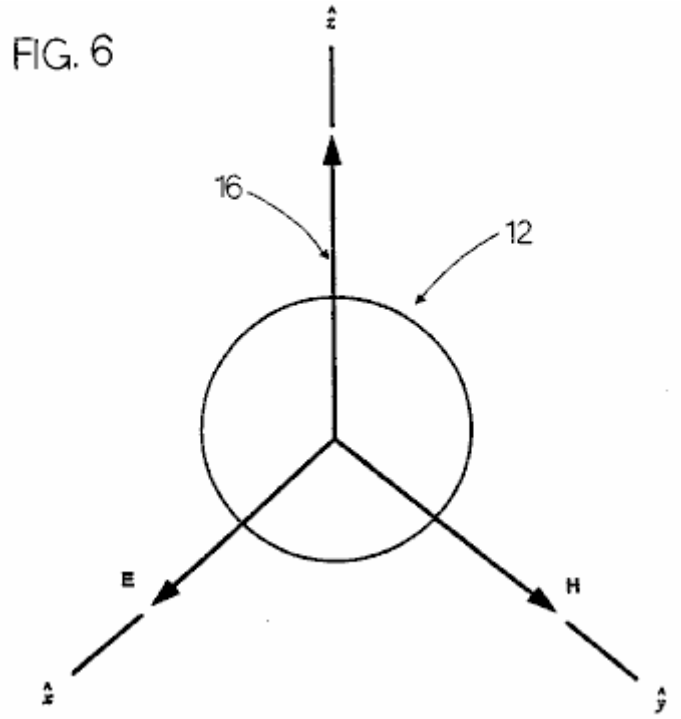


Fig.6 is a diagram of a receiving structure of the system of the present invention showing an incident electromagnetic plane wave impinging on the receiving structure and illustrating the directions of the electric and magnetic field vectors thereof.

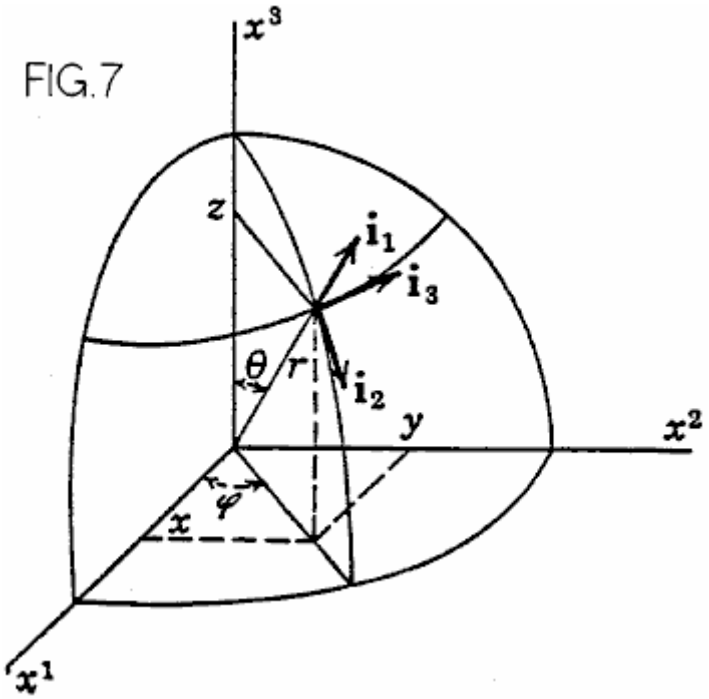


Fig.7 is a diagram of a spherical co-ordinate system as used in the formulas utilised in the system of the present invention.

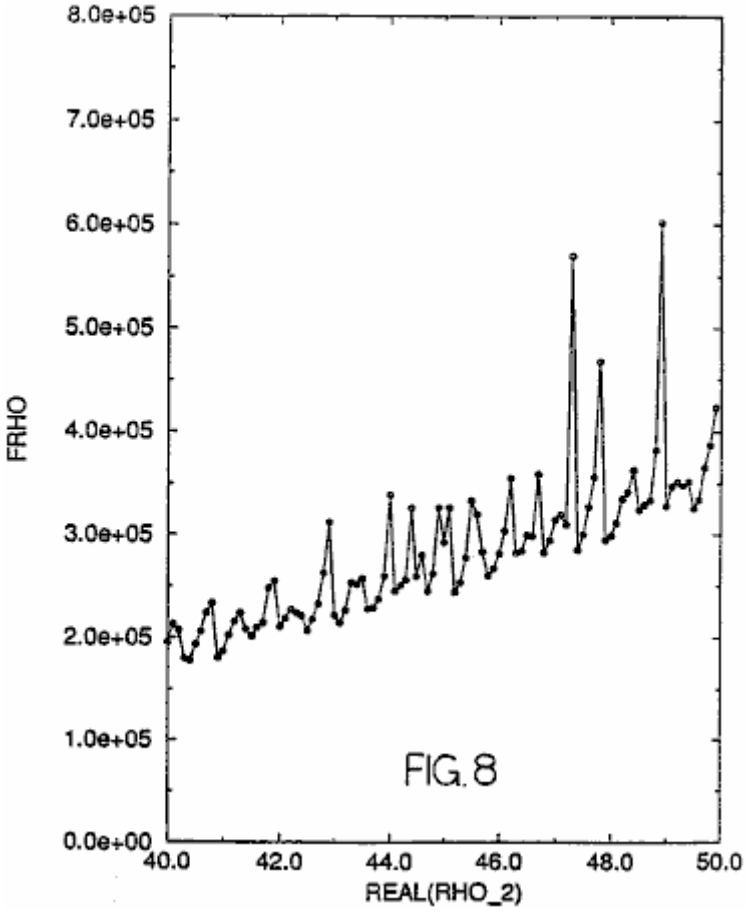


Fig.8 is a graph showing an imaginary rho parameter plotted against a real rho parameter illustrating the values thereof at resonance as well as values thereof at other than resonance.

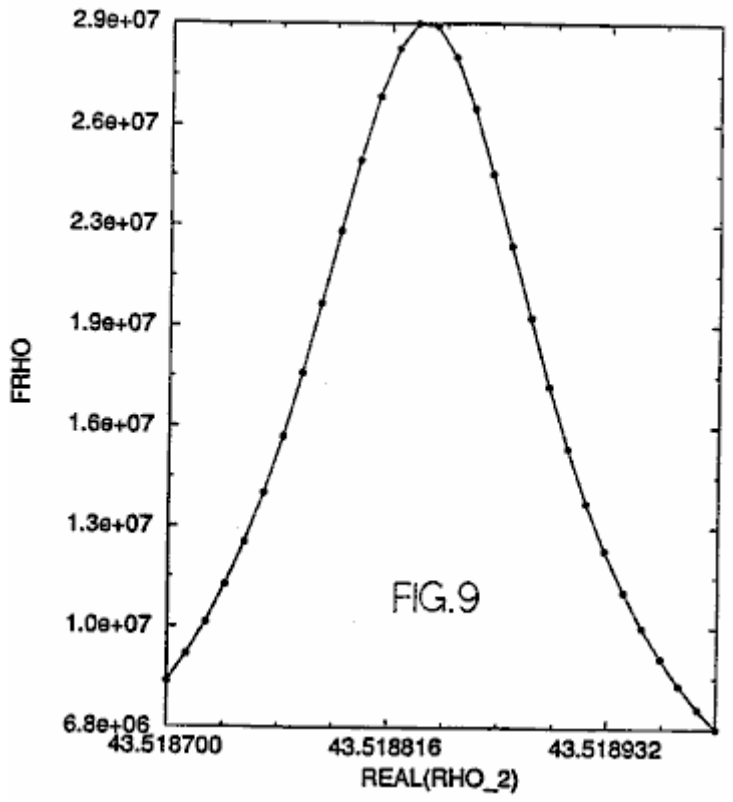


Fig.9 is a graph showing a portion of the graphical representation shown in **Fig.8** illustrating the real and imaginary rho values at or near a single resonance.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

Referring to the drawings, a first embodiment of the present invention is generally designated by the numeral **10**. The system **10** includes a first and second means for receiving **12** and **14** incident electromagnetic radiation **16**. The means for receiving **12** and **14** are preferably a pair of spherical structures **12** and **14** which are preferably composed of a dielectric material. Alternatively, the spheres **12** and **14** may be cubical structures or any other suitable shape. The spheres **12** and **14** may be mounted on a suitable foundation by any suitable mounting means (not shown), or spheres **12** and **14** may be suspended from a suitable foundation by any suitable suspension means (not shown). The spheres **12** and **14** are preferably composed of a dielectric material. The dielectric spheres **12** and **14** scatter and concentrate electromagnetic waves. At very sharply defined frequencies, the spheres **12** and **14** will have resonances wherein the internal energy densities can be five orders of magnitude larger than the energy density of the incident electromagnetic field driving the spheres **12** and **14**. At resonance, the electromagnetic stresses, equivalent to pressures proportional to the energy density, can cause material deformation of the spheres **12** and **14** which produce a secondary electromagnetic field. The spheres **12** and **14** are preferably positioned proximal to each other, as shown in **Fig.1**. Although the proximity of the spheres to each other will adversely affect the resonances, the very high "Q"s of the isolated-sphere resonances results in such adverse affect being relatively small. However, the proximity of the spheres **12** and **14** allows the spheres to interact electromechanically which increases the magnitude of the secondary radiation emitted from them.

The electromagnetic radiation incident upon the spheres **12** and **14** which drives the spheres to resonance is preferably zero-point radiation **16**. However, other types of electromagnetic radiation may also be used to drive the spheres **12** and **14**, if desired.

The effect of a dielectric sphere such as **12** or **14** on an incident electromagnetic radiation such as a plane wave thereof is shown in **Fig.6**. The plane wave propagates in the z axis direction and is diffracted by the sphere **12** resulting in scattering thereof. This scattering is commonly known as Mie scattering. The incident radiation wave has an electric vector component which is linearly polarised in the x axis direction and a magnetic vector component which is linearly polarised in the y axis direction.

An electromagnetic wave incident upon a structure produces a forced oscillation of free and bound charges in synch with the primary electromagnetic field of the incident electromagnetic wave. The movements of the charges produce a secondary electromagnetic field both inside and outside the structure. The secondary electromagnetic radiation comprising this secondary electromagnetic field is shown in **Fig.1** and designated by the numerals **18** and **20**. An antenna which is shown simply as a loop antenna but may also be a dipole or any other suitable type of antenna, is also shown in **Fig.1** and designated by the numeral **22**. The non-linear mutual interactions of the spheres produces interference between the secondary electromagnetic radiation **18** and **20** produces a beat frequency radiation **24** which is preferably at a much lower frequency than the primary radiation **16**. It is this beat frequency radiation **24** which is desired for conversion into electrical energy because it preferably is within the frequency range of rf radiation which may be converted into electrical energy by generally conventional systems. Thus, the radiation **24** received by the antenna **22** is fed via an electrical conductor **26** to a means for converting the beat frequency radiation **24** to electrical energy. This means for converting is designated by the numeral **28** and preferably includes a tuning capacitor **30** and a transformer **32** and a rectifier (preferably a diode) **34**. Instead of including the capacitor **30**, transformer **32** and rectifier **34**, the converter **28** may alternatively include an rf receiver of any suitable type.

The resultant field at any point is the vector sum of the primary and secondary fields. For the equations that follow, the structure receiving the incident plane wave is a sphere of radius a having a propagation constant k_1 positioned in an infinite, homogeneous medium having a propagation constant k_2. The incident plane wave propagates in the z axis direction and is as shown in **Fig.6**. The spherical co-ordinate system used for the vector spherical wave functions is shown in **Fig.7**.

Note: As this patent contains so many non-standard keyboard characters, the remainder of this document is produced using direct images of the original text.

Expansion of the incident field provides:

$$E_i = E_o e^{-i\omega t} \sum_{n=1}^{\infty} i^n \frac{2n+1}{n(n+1)} \left(m_{01n}^{(1)} - i n_{e1n}^{(1)}\right)$$

$$H_i = -\frac{k_2}{\omega \mu_2} E_o e^{-i\omega t} \sum_{n=1}^{\infty} i^n \frac{2n+1}{n(n+1)} \left(m_{e1n}^{(1)} + i n_{01n}^{(1)}\right)$$

where E is the electric field and H is the magnetic field; and

$$m^{(1)}_{\substack{01n\\e1n}} = \pm \frac{1}{\sin\theta} j_n(k_2R)P_n^1(\cos\theta) \begin{matrix}\cos\\ \sin\end{matrix} \phi i_2 - j_n(k_2R) \frac{\partial p_n^1}{\partial\theta} \begin{matrix}\sin\\ \cos\end{matrix} \phi i_3$$

$$n^{(1)}_{\substack{01n\\e1n}} = \frac{n(n+1)}{k_2R} j_n(k_2R)P_n^1(\cos\theta) \begin{matrix}\sin\\ \cos\end{matrix} \phi i_1 + \frac{1}{k_2R} [k_2Rj_n(k_2R)]' \times$$

$$\frac{\partial P_n^1}{\partial\theta} \begin{matrix}\sin\\ \cos\end{matrix} \phi i_2 \pm \frac{1}{k_2R\sin\theta} [k_2Rj_n(k_2R)]'P_n^1(\cos\theta) \begin{matrix}\cos\\ \sin\end{matrix} \phi i_3.$$

The electric and magnetic fields of the incident wave transmitted into the sphere i.e., R<a, can be similarly expanded:

$$E_t = E_0 e^{-i\omega t} \sum_{n=1}^{\infty} i^n \frac{2n+1}{n(n+1)} \left(a_n^m \, {}^{(1)}_{01n} - ib_n^n \, {}^{(1)}_{e1n} \right)$$

$$H_t = \frac{k_2}{\phi\mu_1} E_0 e^{-i\omega t} \sum_{n=1}^{\infty} i^m \frac{2n+1}{n(n+1)} \left(b_n^m \, {}^{(1)}_{e1n} - ia_n^n \, {}^{(1)}_{01n} \right)$$

If $j_n(k_2R)$ is replaced by $h_n^{(1)}(k_2R)$ in the previous equations, the functions $m^{(1)}$ and $n^{(1)}$ become $m^{(3)}$ and $n^{(3)}$. The outgoing fields i.e., R>a, are represented by:

$$E_R = E_0 e^{-i\omega t} \sum_{n=1}^{\infty} i^n \frac{2n+1}{n(n+1)} \left(a_n^{rm} \, {}^{(3)}_{01n} - ib_n^n \, {}^{(3)}_{e1n} \right)$$

$$H_r = \frac{k_2}{\phi\mu_1} E_0 e^{-i\omega t} \sum_{n=1}^{\infty} i^m \frac{2n+1}{n(n+1)} \left(b_n^{rm} \, {}^{(3)}_{e1n} - ia_n^n \, {}^{(3)}_{01n} \right)$$

where H_r represents the resultant wave in the medium surrounding the sphere. At resonance, the values of ρ at resonance require that the a_n^t and b_n^t coefficients be infinite. In order to determine these values of a_n^t and b_n^t, the boundary conditions at the sphere radius are needed. Since there must be continuity of the E and H values at the surface, the following equations are used:

$i_1\times(E_i+E_r)=i_1\times E_t$ and

$i_1\times(H_i+H_r)=i_1\times H_t$

which lead to two pairs of inhomogeneous equations:

$$a_n^t j_n(N\rho) - a_n^r h_n^{(1)}(\rho) = j_n(\rho)$$

$$\mu_2 a_n^r [N\rho j_n(N\rho)]' - \mu_1 a_n^r [\rho h_n^{(1)}(\rho)]' = \mu_1 [\rho j_n(\rho)]' \text{ and}$$

$$\mu_2 N b_n^t j_n(N\rho) - \mu_1 b_n^r h_n^{(1)}(\rho) = \mu_1 j_n(\rho)$$

$$b_n^t [N\rho j_n(N\rho)]' - N b_n^r [\rho h_n^{(1)}(\rho)]' = N[\rho j_n(\rho)]'$$

where $k_1=Nk_2$, $\rho=k_2a$, $k_1a=N\rho$. Spherical Bessel functions of the first kind are denoted by j_n, while those of the third kind are denoted by $h_n^{(1)}$. The resulting equations are:

$$a_n^t = \frac{\mu_1 j_n(\rho)[\rho h_n^{(1)}(\rho)]' - \mu_1 h_n^{(1)}\rho)[\rho j_n(\rho)]'}{\mu_1 j_n(N\rho)[\rho h_n^{(1)}(\rho)]' - \mu_2 h_n^{(1)}(\rho)[N\rho j_n(N\rho)]'}$$

and

$$b_n^t = \frac{\mu_1 N j_n(\rho)[\rho h_n^{(1)}(\rho)]' - \mu_1 N h_n^{(1)}\rho)[\rho j_n(\rho)]'}{\mu_2 N^2 j_n(N\rho)[\rho h_n^{(1)}(\rho)]' - \mu_1 h_n^{(1)}(\rho)[N\rho j_n(N\rho)]'}$$

At a resonance, the denominator of either a_n^t or b_n^t will be zero. Thus, ρ values are found using the above equations that correspond to a resonant combination of angular frequency (ω) and radius (a) for a given sphere material and given surrounding medium. In determining such values of ρ, the following equations are also specifically used:

$$\rho = ak_2 = a\omega\sqrt{\epsilon_2\mu_2} \quad \text{and}$$

$$\rho_1 = (k_1/k_2)\rho$$

where ρ_1 corresponds to the sphere material. An iterative method is preferably used to find the desired values of ρ at resonance. In calculating ρ utilizing the above equations for purposes of example, it was assumed that $\mu_1=\mu_2=\mu_0=4\pi\times 10^{-7}$ and $\epsilon_2=\epsilon_0=8.85419\times10^{-12}$.

One major root of ρ which was found has a value of:

Real (ρ)=+66.39752607619131

Imaginary (ρ)=−0.6347867071968998.

These particular values are not shown in FIG. **8**. However, other values of ρ found using the equations set forth herein are shown in FIG. **8**. The peaks in FIG. **8** are the resonances. One of these resonances shown in FIG. **8** is shown in detail in FIG. **9**. These resonance values are shown for purposes of example. Other resonances also exist which have not been determined; thus, not all possible resonance values are shown in FIGS. **8** and **9**.

Calculation of these values also allows the determination of a possible am combination which would have these root values. For ρ, ϵ (epsilon)=ϵ_0 and μ=μ_0, and

$$\rho = a\omega\sqrt{\epsilon_0\mu_0} = a\omega/c.$$

Expressed in SI units, the speed of light $c=2.99792458\times10^{14}$ m/s. If an a value of 10^{-6} m is assumed for the examples shown herein, then:

$$\omega=\rho c/a\approx 1.9919\times10^{16} - i1.9044\times10^{14} \text{ radians/s.}$$

This is an example of the angular frequency required within the impingent EM radiation in order to create a resonant situation. Examples of other resonances were indicated, and these are shown in FIG. **8**. No complex-frequency plane waves exist. Therefore, the calculations were made by considering only the real portion of the above root and setting the imaginary portion equal to zero. However, upon

doing this, the iterative calculation procedure becomes insensitive to any root in the vicinity of the root's real portion. In the iterative calculation procedure, initially a range of ρ values is input into the equations. These ρ values are in the neighborhood of the prospective root. A range of ρ values is subsequently studied to find any imaginary ρ i.e., fρ (a function of ρ), peaks in that range. Next, once a peak has been chosen, the function order n giving the dominant fρ is determined. This also gives a clue as to whether the peak is due to a magnetic resonance (a_n approaches infinity) or an electrical resonance (b_n approaches infinity). A large number of Newton-Raphson iterations is preferably performed in order to converge upon a root ρ value.

FIGS. **2** and **3** show a second embodiment of the present invention generally designated by the numeral **110**. Embodiment **110** is essentially the same as embodiment **10** except that the antenna is a rf cavity structure **122** which feeds the received beat frequency radiation **124** to a waveguide **126**. Embodiment **110** also preferably includes two spheres **112** and **114** which receive the primary incident electromagnetic radiation **116** and emit the secondary electromagnetic radiation **118** and **120**. As with the spheres **18** and **20** of embodiment **10**, spheres **118** and **120** are preferably composed of a dielectric material. Embodiment **110** also includes converter **128**, capacitor **130**, transformer **132** and rectifier **134** which are essentially identical to the correspondingly numbered elements of embodiment **10**. Therefore, a description of these components of embodiment **110** will not be repeated in order to promote brevity. In addition, the same equations and method of calculation set forth above with regard to embodiment **10** also apply to embodiment. Therefore, their description will not be repeated in order to promote brevity.

FIGS. **4** and **5** show a third embodiment of the present invention generally designated by numeral **210**. Embodiment **210** is essentially identical to the first embodiment **10** except that the embodiment **210** includes a plurality of pairs **215** of receiving means (spheres) **212** and **214** mounted on a substrate **236**. The spheres **212** and **214** are thus in the form of an array **238**. The pairs **215** of the array **238** are preferably positioned proximal to each other in order to maximize the amount of energy extracted from a particular area or space of a given size. Since, as set forth hereinabove, the energy density of the zero point radiation increases as the frequency of the radiation increases, it is desirable that the spheres resonate at as high a bandwidth of frequencies as possible. Because the spheres **212** and **214** must be small in direct proportion to the wavelength of the high frequencies of the incident electromagnetic radiation **216** at which resonance is desirably obtained, the spheres **212** and **214** are preferably microscopic in size. Current lithographic techniques are capable of manufacturing such microscopically small spheres mounted on a suitable substrate thereby providing a suitably miniaturized system **210**. A miniaturized system enhances the energy output capability of the system by

enabling it to resonate at higher frequencies at which there are correspondingly higher energy densities. Consequently, utilization of array **238** in the system **210** enhances the maximum amount of electrical energy provided by the system **210**.

Lithographic techniques may be more amenable to manufacturing microscopically small receiving structures **212** and **214** which may be disc shaped, semispherical or have another shape other than as shown in FIGS. **4** and **5**. Consequently, the receiving means **212** and **214** may accordingly have such alternative shapes rather than the spherical shape shown in FIGS. **4** and **5**. In addition, a large number of small spheres may be manufactured by bulk chemical reactions. Packing a volume with such spheres in close proximity could enhance the output of energy.

Embodiment **210** also includes a plurality of antennas **222** positioned preferably between the spheres **212** and **214** which receive the beat frequency radiation **224** produced by the interference between the secondary radiation **218** and **220**. The antennas **222** are shown as loop antennas **222** but may be any other suitable type of antennas as well.

Embodiment **210** has a plurality of electrical conductors **226** which preferably include traces mounted on the substrate **236** which occupies a finite volume. The electrical conductors **226** feed the electrical output from the antennas **222** to a suitable converter **228** which preferably includes tuning capacitor **230**, transformer **232** and rectifier **234**, as with embodiments **10** and **110**. Except as set forth above, the components of embodiment **210** are identical to embodiment **10** so the detailed description of these components will not be repeated in order to promote brevity. In addition, the same equations and method of calculation set forth above for embodiment **10** also apply to embodiment **210**. Therefore, the description of these equations and method of calculation will not be repeated in order to promote brevity.

Accordingly, there has been provided, in accordance with the invention, a system which converts high frequency zero point electromagnetic radiation into electrical energy effectively and efficiently and thus fully satisfies the objectives set forth above. It is to be understood that all terms used herein are descriptive rather than limiting. Although the invention has been specifically described with regard to the specific embodiments set forth herein, many alternative embodiments, modifications and variations will be apparent to those skilled in the art in light of the disclosure set forth herein. Accordingly, it is intended to include all such alternatives, embodiments, modifications and variations that fall within the spirit and scope of the invention as set forth in the claims hereinbelow.

What is claimed is:

1. A system for converting incident electromagnetic radiation energy to electrical energy, comprising:

a first means for receiving incident primary electromagnetic radiation, said means for receiving producing

emitted secondary electromagnetic radiation at a first frequency, said first means for receiving having a first volumetric size selected to resonate at a frequency within the frequency spectrum of the incident primary electromagnetic radiation in order to produce the secondary electromagnetic radiation at the first frequency at an enhanced energy density;

a second means for receiving the incident primary electromagnetic radiation, said means for receiving producing emitted secondary electromagnetic radiation at a second frequency, the secondary radiation at the first frequency and the secondary radiation at the second frequency interfering to produce secondary radiation at a lower frequency than that of the incident primary radiation, said second means for receiving having a second volumetric size selected to resonate at a frequency within the frequency spectrum of the incident primary electromagnetic radiation in order to produce the emitted secondary electromagnetic radiation at the second frequency at an enhanced energy density;

an antenna for receiving the emitted secondary electromagnetic radiation at the lower frequency, said antenna providing an electrical output responsive to the secondary electromagnetic radiation received;

a converter electrically connected to said antenna for receiving electrical current output from said antenna and converting the electrical current output to electrical current having a desired voltage and waveform.

2. The system of claim **1** wherein:

said first means for receiving is composed of a dielectric material; and

said second means for receiving is composed of a dielectric material.

3. The system of claim **1** wherein:

said first means for receiving is spherical; and

said second means for receiving is spherical.

4. A system for for converting incident zero point electromagnetic radiation energy to electrical energy, comprising:

a first means for receiving incident primary zero point electromagnetic radiation, said means for receiving producing emitted secondary electromagnetic radiation at a first frequency;

a second means for receiving the incident primary zero point electromagnetic radiation, said means for receiving producing emitted secondary electromagnetic radiation at a second frequency, the secondary radiation at the first frequency and the secondary radiation at the second frequency interfering to produce secondary radiation at a beat frequency which is lower than that of the incident primary radiation;

an antenna for receiving the emitted secondary electromagnetic radiation at the lower frequency, said antenna providing an electrical output responsive to the secondary electromagnetic radiation received;

means for transmitting the emitted secondary electromagnetic radiation at the beat frequency from said antenna, said means for transmitting connected to said antenna;

a converter connected to said means for transmitting for receiving the emitted secondary electromagnetic radiation at the beat frequency from said antenna and converting the same to electrical current having a desired voltage and waveform.

5. The system of claim **4** wherein:

said first means for receiving has a first volumetric spherical size selected to resonate in response to the incident primary electromagnetic radiation in order to produce the secondary electromagnetic radiation at the first frequency at an enhanced energy density; and

said second means for receiving has a second volumetric spherical size selected to resonate in response to the incident primary electromagnetic radiation in order to produce the emitted secondary electromagnetic radiation at the second frequency at an enhanced energy density, said first and second volumetric sizes selected based on parameters of propagation constant of said first and second means for receiving, propagation constant of medium in which said first and second means for receiving are located and frequency of the incident primary electromagnetic radiation.

6. The system of claim **5** wherein the first and second volumetric sizes are selected by utilizing the formulas:

$$a_n{}^t = \frac{\mu_1 j_n(\rho)[\rho h_n^{(1)}(\rho)]' - \mu_1 h_n^{(1)}\rho)[\rho j_n(\rho)]'}{\mu_1 j_n(N\rho)[\rho h_n^{(1)}(\rho)]' - \mu_2 h_n^{(1)}(\rho)[N\rho j_n(N\rho)]'}$$

$$b_n{}^t = \frac{\mu_1 N j_n(\rho)[\rho h_n^{(1)}(\rho)]' - \mu_1 N h_n^{(1)}\rho)[\rho j_n(\rho)]'}{\mu_2 N^2 j_n(N\rho)[\rho h_n^{(1)}(\rho)]' - \mu_1 h_n^{(1)}(\rho)[N\rho j_n(N\rho)]'}$$

$$\rho = a\omega\sqrt{\epsilon_2\mu_2}$$

wherein at a resonance, the denominator of either equation for $a_n{}^t$ or $b_n{}^t$ will be approximately zero and wherein k_1=propagation constant of the means for receiving, k_2=propagation constant of medium through which the incident electromagnetic radiation propagates, a is the radius of either means for receiving, $N=k_1/k_2$, $\rho=k_2a$, $k_1a=N\rho$, $a_n{}^t$=magnitude of oscillations of the electric field of the nth order, $b_n{}^t$=magnitude of oscillations of the magnetic field of the nth order, ω=angular frequency of the incident electromagnetic radiation, ϵ is the permittivity of the material or medium and μ is the permeability of the material or medium.

7. The system of claim **6** wherein the radius of the first means for receiving is different from the radius of the second means for receiving, difference between the radius of said first means for receiving and the radius of said second means for receiving selected so that the beat frequency resulting from the difference is a frequency which facilitates conversion of the beat frequency electromagnetic radiation to electrical energy.

8. The system of claim **4** wherein:

said first means for receiving is composed of a dielectric material; and

said second means for receiving is composed of a dielectric material.

9. The system of claim **4** wherein:

said first means for receiving is spherical; and

said second means for receiving is spherical.

10. The system of claim **4** wherein said antenna is positioned generally between said first and second means for receiving.

11. The system of claim **4** wherein said antenna is a loop antenna.

12. The system of claim **4** wherein said antenna is a generally concave shell partially enclosing said first and second means for receiving.

13. The system of claim **4** wherein said means for transmitting is a waveguide.

14. A system for for converting incident zero point electromagnetic radiation energy to electrical energy, comprising:

a substrate;

a plurality of pairs of first means for receiving incident primary zero point electromagnetic radiation and second means for receiving incident primary zero point electromagnetic radiation, said plurality of pairs of means for receiving mounted on said substrate, said first means for receiving producing emitted secondary electromagnetic radiation at a first frequency, said second means for receiving the incident primary zero point electromagnetic radiation producing emitted secondary electromagnetic radiation at a second frequency, the secondary radiation at the first frequency and the secondary radiation at the second frequency interfering to produce secondary radiation at a beat frequency which is lower than that of the incident primary radiation, said first means for receiving having a first volumetric size selected to resonate in response to the incident primary electromagnetic radiation in

order to produce the secondary electromagnetic radiation at the first frequency at an enhanced energy density, and said second means for receiving having a second volumetric size selected to resonate in response to the incident primary electromagnetic radiation in order to produce the emitted secondary electromagnetic radiation at the second frequency at an enhanced energy density, said first and second volumetric sizes selected based on parameters of propagation constant of said first and second means for receiving, propagation constant of medium in which said first and second means for receiving are located and frequency of the incident primary electromagnetic radiation, said first and second volumetric sizes being different from each other;

a plurality of antennas for receiving the emitted secondary electromagnetic radiation at the lower frequency, said antenna providing an output responsive to the secondary electromagnetic radiation received, said plurality of antennas mounted on said substrate, each of said plurality of antennas receiving the emitted secondary electromagnetic radiation of one of said pairs of first and second means for receiving;

means for transmitting the emitted secondary electromagnetic radiation at the beat frequency from said antenna, said means for transmitting connected to said plurality of antennas;

a converter connected to said means for transmitting for receiving the emitted secondary electromagnetic radiation at the beat frequency from said antenna and converting the same to electrical current having a desired voltage and waveform.

STANLEY MEYER

US Patent 4,936,961 June 26, 1990 Inventor: Stanley A. Meyer

METHOD FOR THE PRODUCTION OF A FUEL GAS

Please note that this is a re-worded excerpt from this patent. It describes one of the methods which Stan used to split water into hydrogen and oxygen using very low levels of input power.

OBJECTS OF THE INVENTION

It is an object of the invention to provide a fuel cell and a process in which molecules of water are broken down into hydrogen and oxygen gases, and other formerly dissolved within the water is produced. As used herein the term "fuel cell" refers to a single unit of the invention comprising a water capacitor cell, as hereinafter explained, that produces the fuel gas in accordance with the method of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS:

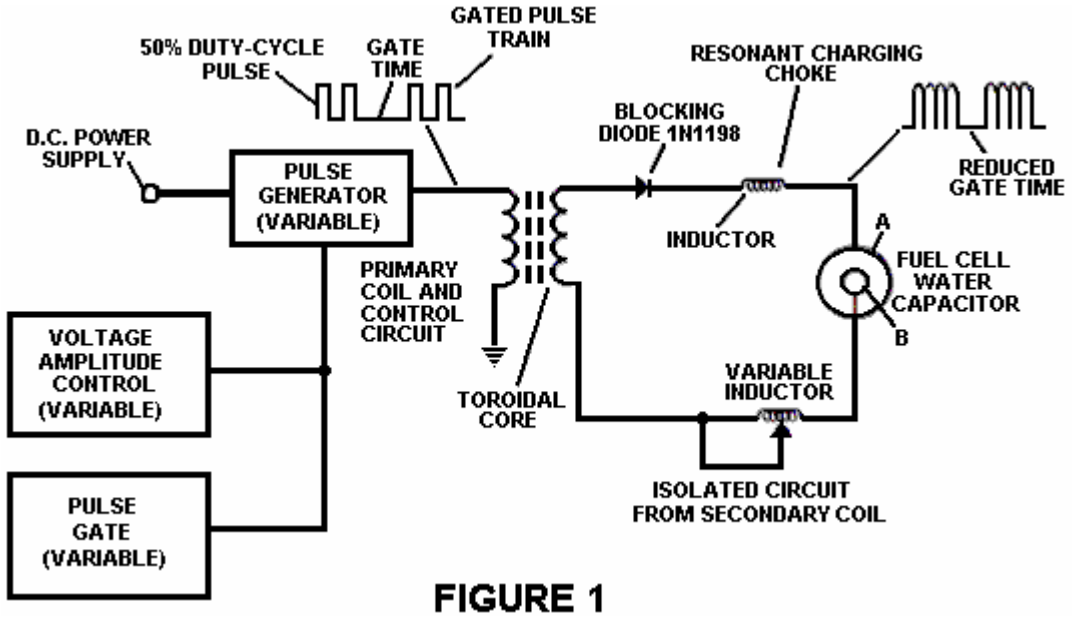


FIGURE 1

Fig.1 Illustrates a circuit useful in the process.

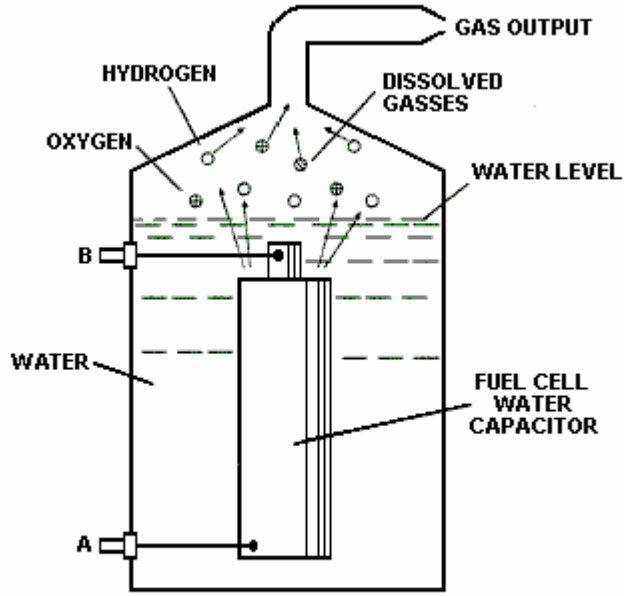


FIG 2

Fig.2 Shows a perspective of a "water capacitor" element used in the fuel cell circuit.

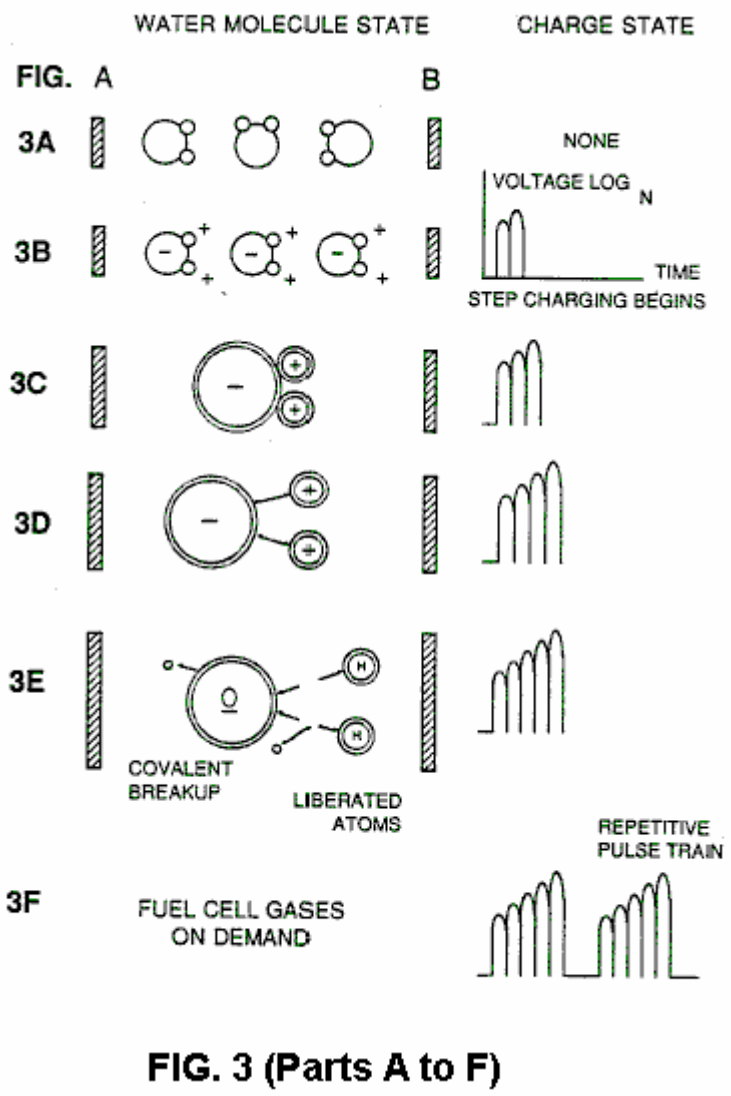


FIG. 3 (Parts A to F)

Figs. 3A through 3F are illustrations depicting the theoretical bases for the phenomena encountered during operation of the invention herein.

DESCRIPTION OF THE PREFERRED EMBODIEMENT

In brief, the invention is a method of obtaining the release of a gas mixture including hydrogen on oxygen and other dissolved gases formerly entrapped in water, from water consisting of:

(a) Providing a capacitor, in which the water is included as a dielectric liquid between capacitor plates, in a resonant charging choke circuit that includes an inductance in series with the capacitor;

(b) Subjecting the capacitor to a pulsating, unipolar electric voltage field in which the polarity does not pass beyond an arbitrary ground, whereby the water molecules within the capacitor are subjected to a charge of the same polarity and the water molecules are distended by their subjection to electrical polar forces;

(c) Further subjecting in said capacitor to said pulsating electric field to achieve a pulse frequency such that the pulsating electric field induces a resonance within the water molecule;

(d) Continuing the application of the pulsating frequency to the capacitor cell after resonance occurs so that the energy level within the molecule is increased in cascading incremental steps in proportion to the number of pulses;

(e) Maintaining the charge of said capacitor during the application of the pulsing field, whereby the co-valent electrical bonding of the hydrogen and oxygen atoms within said molecules is destabilised such that the force of the electrical field applied, as the force is effective within the molecule, exceeds the bonding force of the molecule, and hydrogen and oxygen atoms are liberated from the molecule as elemental gases; and

(f) Collecting said hydrogen and oxygen gases, and any other gases that were formerly dissolved within the water, and discharging the collected gases as a fuel gas mixture.

The process follows the sequence of steps shown in the following **Table 1** in which water molecules are subjected to increasing electrical forces. In an ambient state, randomly oriented water molecules are aligned with respect to a molecule polar orientation.

They are next, themselves polarised and "elongated" by the application of an electrical potential to the extent that covalent bonding of the water molecule is so weakened that the atoms dissociate and the molecule breaks down into hydrogen and oxygen elemental components.

Engineering design parameters based on known theoretical principles of electrical circuits determine the incremental levels of electrical and wave energy input required to produce resonance in the system whereby the fuel gas comprised of a mixture of hydrogen, oxygen, and other gases such as air were formerly dissolved within the water, is produced.

TABLE 1

Process Steps:

The sequence of the relative state of the water molecule and/or hydrogen/oxygen/other atoms:

A. (ambient state) random
B. Alignment of polar fields
C. Polarisation of molecule
D. Molecular elongation
E. Atom liberation by breakdown of covalent bond
F. Release of gases

In the process, the point of optimum gas release is reached at a circuit resonance. Water in the fuel cell is subjected to a pulsating, polar electric field produced by the electrical circuit whereby the water molecules are distended by reason of their subjection to electrical polar forces of the capacitor plates. The polar pulsating frequency applied is such that the pulsating electric field induces a resonance in the molecule. A cascade effect occurs and the overall energy level of specific water molecules is increased in cascading, incremental steps. The hydrogen and oxygen atomic gases, and other gas components formerly entrapped as dissolved gases in water, are released when the resonant energy exceeds the covalent bonding force of the water molecule. A preferred construction material for the capacitor plates is T304-grade stainless steel which is non-chemical reactive with water, hydrogen, or oxygen. An electrically conductive material which is inert in the fluid environment is a desirable material of construction for the electrical field plates of the "water capacitor" employed in the circuit.

Once triggered, the gas output is controllable by the attenuation of operational parameters. Thus, once the frequency of resonance is identified, by varying the applied pulse voltage to the water fuel cell assembly, gas output is varied. By varying the pulse shape and/or amplitude or pulse train sequence of the initial pulsing wave source, final gas output is varied. Attenuation of the voltage field frequency in the form of OFF and ON pulses likewise affects output.

The overall apparatus thus includes an electrical circuit in which a water capacitor having a known dielectric property is an element. The fuel gases are obtained from the water by the disassociation of the water molecule. The water molecules are split into component atomic elements (hydrogen and oxygen gases) by a voltage stimulation process called the electrical polarisation process which also releases dissolved gases entrapped in the water.

From the outline of physical phenomena associated with the process described in **Table 1**, the theoretical basis of the invention considers the respective states of molecules and gases and ions derived from liquid water. Before voltage stimulation, water molecules are randomly dispersed throughout water in a container. When a unipolar voltage pulse train such as shown in **Figs.3B** through **3F** is applied to positive and negative capacitor plates, an increasing voltage potential is induced in the molecules in a linear, step like charging effect. The electrical field of the particles within a volume of water including the electrical field plates increases from a low energy state to a high energy state successively is a step manner following each pulse-train as illustrated figuratively in the depictions of **Figs.3A** through **3F**. The increasing voltage potential is always positive in direct relationship to negative ground potential during each pulse. The voltage polarity on the plates which create the voltage fields remains constant although the voltage charge increases. Positive and negative voltage "zones" are thus formed simultaneously in the electrical field of the capacitor plates.

In the first stage of the process described in **Table 1**, because the water molecule naturally exhibits opposite electrical fields in a relatively polar configuration (the two hydrogen atoms are positively electrically charged relative to the negative electrically charged oxygen atom), the voltage pulse causes initially randomly oriented water molecules in the liquid state to spin and orient themselves with reference to positive and negative poles of the voltage fields applied. The positive electrically charged hydrogen atoms of said water molecule are attracted to a negative voltage field; while, at the same time, the negative electrically charged oxygen atoms of the same water molecule are attracted to a positive voltage field. Even a slight potential difference applied to inert, conductive plates of a containment chamber which forms a capacitor will initiate polar atomic orientation within the water molecule based on polarity differences.

When the potential difference applied causes the orientated water molecules to align themselves between the conductive plates, pulsing causes the voltage field intensity to be increased in accordance with **Fig.3B**. As further molecule alignment occurs, molecular movement is hindered. Because the positively charged hydrogen atoms of said aligned molecules are attracted in a direction opposite to the negatively charged oxygen atoms, a polar charge alignment or distribution occurs within the molecules between said voltage zones, as shown in **Fig.3B**. And as the energy level of the atoms subjected to resonant pulsing increases, the stationary water molecules become elongated as shown in **Fig.3C** and **Fig.3D**. Electrically charged nuclei and electrons are attracted toward opposite electrically charged equilibrium of the water molecule.

As the water molecule is further exposed to an increasing potential difference resulting from the step charging of the capacitor, the electrical force of attraction of the atoms within the molecule to the capacitor plates of the chamber also increase in strength. As a result, the covalent bonding between which form the molecule is weakened --- and ultimately terminated. The negatively charged electron is attracted toward the positively charged hydrogen atoms, while at the same time, the negatively charged oxygen atoms repel electrons.

In a more specific explanation of the "sub-atomic" action the occurs in the water fuel cell, it is known that natural water is a liquid which has a dielectric constant of 78.54 at 20 degrees C. and 1 atmosphere pressure. [Handbook of Chemistry & Physics, 68th ed., CRC Press(Boca Raton, Florida (1987-88)), Section E-50. H20(water)].

When a volume of water is isolated and electrically conductive plates, that are chemically inert in water and are separated by a distance, are immersed in water, a capacitor is formed, having a capacitance determined by the surface area of the plates, the distance of their separation and the dielectric constant of water.

When water molecules are exposed to voltage at a restricted current, water takes on an electrical charge. By the laws of electrical attraction, molecules align according to positive and negative polarity fields of the molecule and the alignment field. The plates of the capacitor constitute such as alignment field when a voltage is applied.

When a charge is applied to a capacitor, the electrical charge of the capacitor equals the applied voltage charge; in a water capacitor, the dielectric property of water resists the flow of amps in the circuit, and the water molecule itself, because it has polarity fields formed by the relationship of hydrogen and oxygen in the covalent bond, and intrinsic dielectric property, becomes part of the electrical circuit, analogous to a "microcapacitor" within the capacitor defined by the plates.

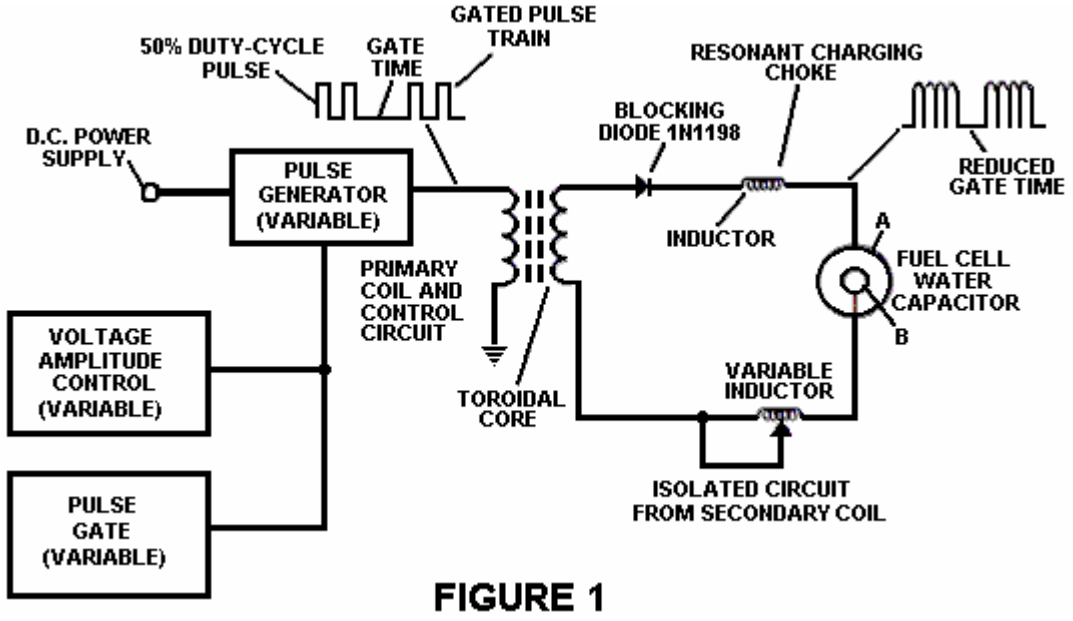


FIGURE 1

In the Example of a fuel cell circuit of **Fig.1**, a water capacitor is included. The step-up coil is formed on a conventional toroidal core formed of a compressed ferromagnetic powered material that will not itself become permanently magnetised, such as the trademarked "Ferramic 06# "Permag" powder as described in Siemens Ferrites Catalogue, CG-2000-002-121, (Cleveland, Ohio) No. F626-1205". The core is 1.50 inch in diameter and 0.25 inch in thickness. A primary coil of 200 turns of 24 gauge copper wire is provided and coil of 600 turns of 36 gauge wire comprises the secondary winding.

In the circuit of **Fig.1**, the diode is a 1N1198 diode which acts as a blocking diode and an electric switch that allows voltage flow in one direction only. Thus, the capacitor is never subjected to a pulse of reverse polarity.

The primary coil of the toroid is subject to a 50% duty cycle pulse. The toroidal pulsing coil provides a voltage step-up from the pulse generator in excess of five times, although the relative amount of step-up is determined by preselected criteria for a particular application. As the stepped-up pulse enters first inductor (formed from 100 turns of 24 gauge wire 1 inch in diameter), an electromagnetic field is formed around the inductor, voltage is switched off when the pulse ends, and the field collapses and produces another pulse of the same polarity i.e., another positive pulse is formed where the 50% duty cycle was terminated. Thus, a double pulse frequency is produced; however, in pulse train of unipolar pulses, there is a brief time when pulses are not present.

By being so subjected to electrical pulses in the circuit of **Fig.1**, water confined in the volume that includes the capacitor plates takes on an electrical charge that is increased by a step charging phenomenon occurring in the water capacitor. Voltage continually increases (to about 1000 volts and more) and the water molecules starts to elongate.

The pulse train is then switched off; the voltage across the water capacitor drops to the amount of the charge that the water molecules have taken on, i.e., voltage is maintained across the charged capacitor. The pulse train is the reapplied.

Because a voltage potential applied to a capacitor can perform work, the higher the voltage the higher the voltage potential, the more work is performed by a given capacitor. In an optimum capacitor that is wholly non-conductive, zero (0) current flow will occur across the capacitor. Thus, in view of an idealised capacitor circuit, the object of the water capacitor circuit is to prevent electron flow through the circuit, i.e. such as occurs by electron flow or leakage through a resistive element that produces heat. Electrical leakage in the water will occur, however, because of some residual conductivity and impurities or ions that may be otherwise present in the water. Thus, the water capacitor is preferably chemically inert. An electrolyte is not added to the water.

In the isolated water bath, the water molecule takes on charge, and the charge increases. The object of the process is to switch off the covalent bonding of the water molecule and interrupt the subatomic force, i.e. the electrical force or electromagnetic force, that binds the hydrogen and oxygen atoms to form a molecule so that the hydrogen and oxygen separate.

Because an electron will only occupy a certain electron shell (shells are well known) the voltage applied to the capacitor affects the electrical forces inherent in the covalent bond. As a result of the charge applied by the plates, the applied force becomes greater than the force of the covalent bonds between the atom of the water molecule;

and the water molecule becomes elongated. When this happens, the time share ratio of the electron shells is modified.

In the process, electrons are extracted from the water bath; electrons are not consumed nor are electrons introduced into the water bath by the circuit as electrons are conventionally introduced in as electrolysis process. There may nevertheless occur a leakage current through the water. Those hydrogen atoms missing electrons become neutralised; atoms are liberated from the water. The charged atoms and electrons are attracted to the opposite polarity voltage zones created between the capacitor plates. The electrons formerly shared by atoms in the water covalent bond are reallocated such that neutral elemental gases are liberated.

In the process, the electrical resonance may be reached at all levels of voltage potential. The overall circuit is characterised as a "resonant charging choke" circuit which is an inductor in series with a capacitor that produces a resonant circuit. [SAMS Modern Dictionary of Electronics, Rudolf Garff, copyright 1984, Howard W. Sams & Co. (Indianapolis, Ind.), page 859.] Such a resonant charging choke is on each side of the capacitor. In the circuit, the diode acts as a switch that allows the magnetic field produced in the inductor to collapse, thereby doubling the pulse frequency and preventing the capacitor from discharging. In this manner a continuous voltage is produced across the capacitor plates in the water bath; and the capacitor does not discharge. The water molecules are thus subjected to a continuously charged field until the breakdown of the covalent bond occurs.

As noted initially, the capacitance depends on the dielectric properties of the water and the size and separation of the conductive elements forming the water capacitor.

EXAMPLE 1

In an example of the circuit of **Fig.1** (in which other circuit element specifications are provided above), two concentric cylinders 4 inches long formed the water capacitor of the fuel cell in the volume of water. The outside cylinder was 0.75 inch in outside diameter; the inner cylinder was 0.5 inch in outside diameter. Spacing from the outside of the inner cylinder to the inner surface of the outside cylinder was 0.0625 inch. Resonance in the circuit was achieved at a 26 volt applied pulse to the primary coil of the toroid at 0 KHz (*suspected mis-typing for 10KHz*), and the water molecules disassociated into elemental hydrogen and oxygen and the gas released from the fuel cell comprised a mixture of hydrogen, oxygen from the water molecule, and gases formerly dissolved in the water such as the atmospheric gases or oxygen, nitrogen, and argon.

In achieving resonance in any circuit, as the pulse frequency is adjusted, the flow of amps is minimised and the voltage is maximised to a peak. Calculation of the resonance frequency of an overall circuit is determined by known means; different cavities have a different frequency of resonance dependant on parameters of the water dielectric, plate size, configuration and distance, circuit inductors, and the like. Control of the production of fuel gas is determined by variation of the period of time between a train of pulses, pulse amplitude and capacitor plate size and configuration, with corresponding value adjustments to other circuit components.

The wiper arm on the second conductor tunes the circuit and accommodates to contaminants in water so that the charge is always applied to the capacitor. The voltage applied determines the rate of breakdown of the molecule into its atomic components. As water in the cell is consumed, it is replaced by any appropriate means or control system.

Variations of the process and apparatus may be evident to those skilled in the art.

CLAIMS:

1. A method of obtaining the release of a gas mixture including hydrogen and oxygen and other dissolved gases formerly entrapped in water, from water, consisting of:

(a) Providing a capacitor in which water is included as a dielectric between capacitor plates, in a resonant charging choke circuit that includes an inductance in series with the capacitor;

(b) Subjecting the capacitor to a pulsating, unipolar electric charging voltage in which the polarity does not pass beyond an arbitrary ground, whereby the water molecules within the capacitor plates;

(c) Further subjecting the water in said capacitor to a pulsating electric field resulting from the subjection of the capacitor to the charging voltage such that the pulsating electric field induces a resonance within the water molecules;

(d) Continuing the application of the pulsating charging voltage to the capacitor after the resonance occurs so that the energy level within the molecules is increased in cascading incremental steps in proportion to the number of pulses;

(e) Maintaining the charge of said capacitor during the application of the pulsating charge voltage, whereby the covalent electrical bonding of the hydrogen and oxygen atoms within said molecules is destabilised, such

that the force of the electrical field applied to the molecules exceeds the bonding force within the molecules, and the hydrogen and oxygen atoms are liberated from the molecules as elemental gases.

2. The method of claim 1 including the further steps of collecting said liberated gases and any other gases that were formerly dissolved within the water and discharging said collected gases as a fuel gas mixture.

STANLEY MEYER

US Patent 4,389,981 28th June 1983 Inventor: Stanley A. Meyer

HYDROGEN GAS INJECTOR SYSTEM FOR INTERNAL COMBUSTION ENGINES

Please note that this is a re-worded excerpt from this patent. It describes one method for using hydrogen and oxygen gases to fuel a standard vehicle engine.

ABSTRACT

System and apparatus for the controlled intermixing of a volatile hydrogen gas with oxygen and other non-combustible gasses in a combustion system. In a preferred arrangement the source of volatile gas is a hydrogen source, and the non-combustible gasses are the exhaust gasses of the combustion system in a closed loop arrangement. Specific structure for the controlled mixing of the gasses, the fuel flow control, and safety are disclosed.

CROSS REFERENCES AND BACKGROUND

There is disclosed in my co-pending U.S. patent application Serial No. 802,807 filed Sept. 16, 1981 for a Hydrogen-Generator, a generating system converting water into hydrogen and oxygen gasses. In that system and method the hydrogen atoms are dissociated from a water molecule by the application of a non-regulated, non-filtered, low-power, direct current voltage electrical potential applied to two non-oxidising similar metal plates having water passing between them. The sub-atomic action is enhanced by pulsing this DC voltage. The apparatus comprises structural configurations in alternative embodiments for segregating the generated hydrogen gas from the oxygen gas.

In my co-pending patent application filed May 5, 1981, U.S. Serial No. 262,744 now abandoned for Hydrogen-Airdation Processor, non-volatile and non-combustible gasses are controlled in a mixing stage with a volatile gas. The hydrogen airdation processor system utilises a rotational mechanical gas displacement system to transfer, meter, mix, and pressurise the various gasses. In the gas transformation process, ambient air is passed through an open flame gas-burner system to eliminate gasses and other substances present. After that, the non-combustible gas-mixture is cooled, filtered to remove impurities, and mechanically mixed with a pre-determined amount of hydrogen gas. This results in a new synthetic gas.

This synthetic gas-formation stage also measures the volume and determines the proper gas-mixing ratio for establishing the desired burn-rate of hydrogen gas. The rotational mechanical gas displacement system in that process determines the volume of synthetic gas to be produced.

The above-noted hydrogen airdation processor, of my co-pending application, is a multi-stage system suited to special applications. Whereas the hydrogen generator system of my other mentioned co-pending application does disclose a very simple and unique hydrogen generator.

In my co-pending patent application Serial No. 315,945, filed Oct. 18, 1981 there is disclosed a combustion system incorporating a mechanical drive system. In one instance, this is designed to drive a piston in an automotive device. There is shown a hydrogen generator for developing hydrogen gas, and perhaps other non-volatile gasses such as oxygen and nitrogen. The hydrogen gas with the attendant non-volatile gasses is fed via a line to a controlled air intake system. The combined hydrogen, non-volatile gasses, and the air, after inter-mixing, are fed to a combustion chamber where they are ignited. The exhaust gasses of the combustion chamber are returned in a closed loop arrangement to the mixing chamber to be used again as the non-combustible gas component. Particular applications and structural embodiments of the system are disclosed.

SUMMARY OF THE INVENTION

The system of the present invention in its most preferred embodiment is for a combustion system utilising hydrogen gas; particularly to drive the pistons in an car engine. The system utilises a hydrogen generator for developing hydrogen gas. The hydrogen gas and other non-volatile gasses are then fed, along with oxygen, to a mixing chamber. The mixture is controlled in such a way as to lower the temperature of the combustion to bring it in line with that of the currently existing commercial fuels. The hydrogen gas feed line to the combustion chamber includes a fine linear control gas flow valve. An air intake is the source of oxygen and it also includes a variable

valve. The exhaust gasses from the combustion chamber are utilised in a controlled manner as the non-combustible gasses.

The hydrogen generator is improved by the inclusion of a holding tank which provides a source of start-up fuel. Also, the hydrogen gas generator includes a pressure-controlled safety switch on the combustion chamber which disconnects the input power if the gas pressure rises above the required level. The simplified structure includes a series of one-way valves, safety valves, and quenching apparatus. The result is an apparatus which comprises the complete assembly for converting a standard car engine from petrol (or other fuels) to use a hydrogen/gas mixture.

OBJECTS

It is accordingly a principal object of the present invention to provide a combustion system of gasses combined from a source of hydrogen and non-combustible gasses.

Another object of the invention is to provide such a combustion system that intermixes the hydrogen and non-combustible gasses in a controlled manner and thereby control the combustion temperature.

A further object of the invention is to provide such a combustion system that controls the fuel flow to the combustion chamber in s system and apparatus particularly adapted to hydrogen gas.

Still other objects and features of the present invention will become apparent from the following detailed description when taken in conjunction with the drawings in which:

BRIEF DESCRIPTION OF THE DRAWINGS

Fig.1 is a mechanical schematic illustration partly in block form of the present invention in its most preferred embodiment.

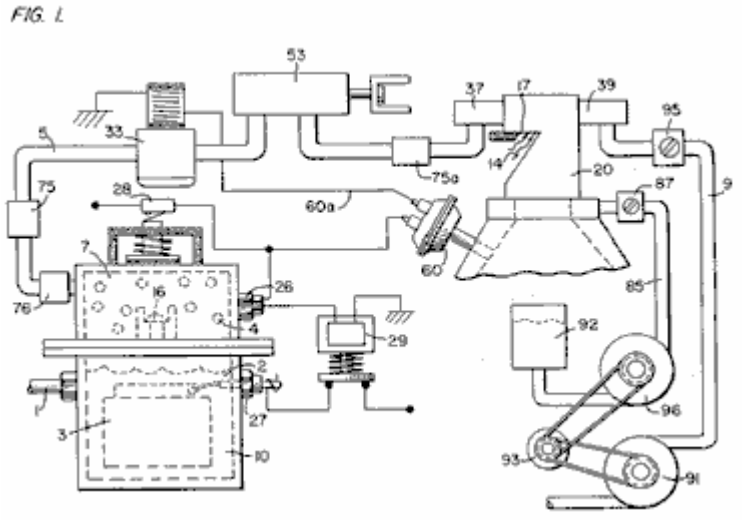


Fig.2 is a block schematic illustration of the preferred embodiment of the hydrogen injector system shown in Fig.1.

FIG. 2.

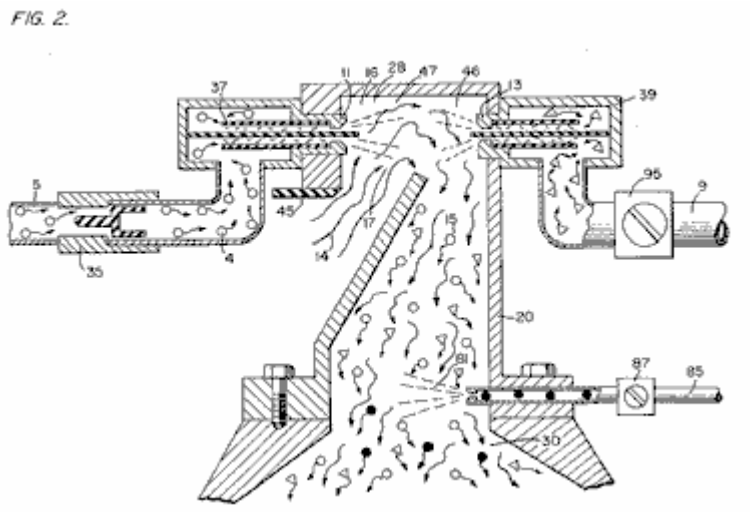


Fig.3 is the fine linear fuel flow control shown in Fig.1.

FIG. 3.

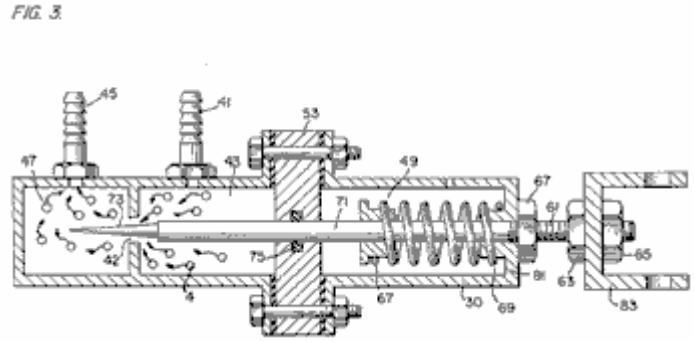


Fig.4 is cross-sectional illustration of the complete fuel injector system in an car utilising the concepts of the present invention.

FIG. 4.

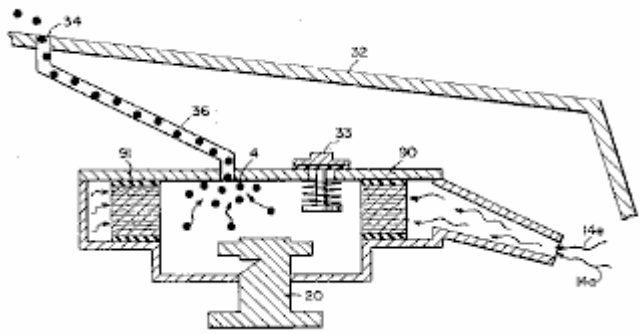


Fig.5 is a schematic drawing in a top view of the fuel injector system utilised in the preferred embodiment.

FIG. 5.

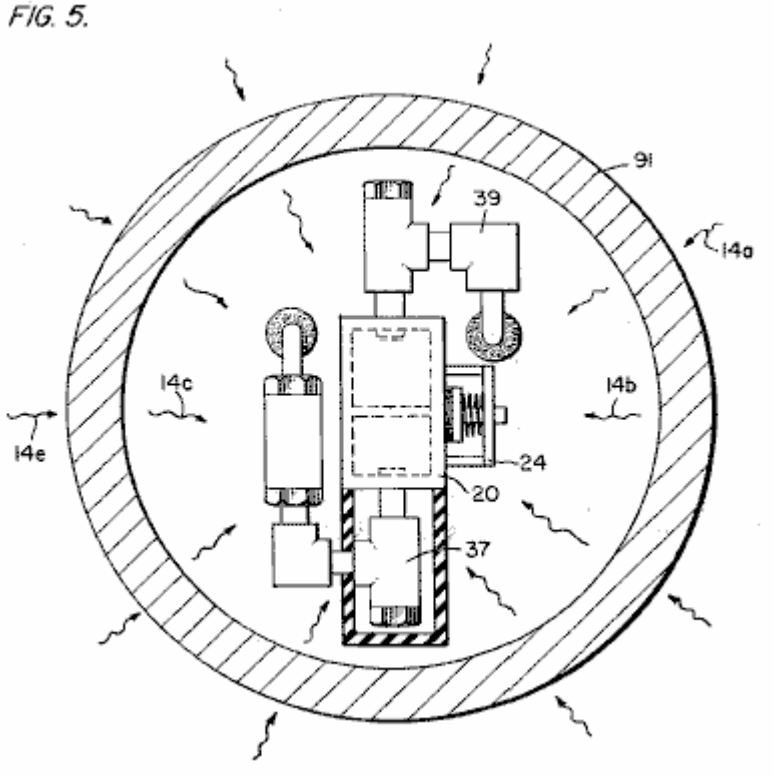


Fig.6 is a cross-sectional side view of the fuel injector system in the present invention.

FIG. 6.

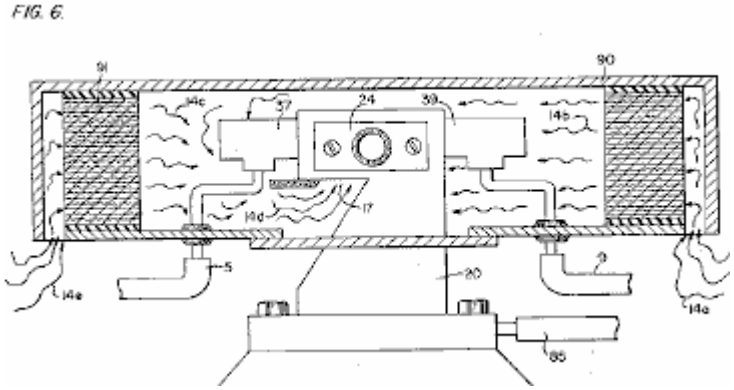


Fig.7 is a side view of the fuel mixing chamber.

FIG. 7.

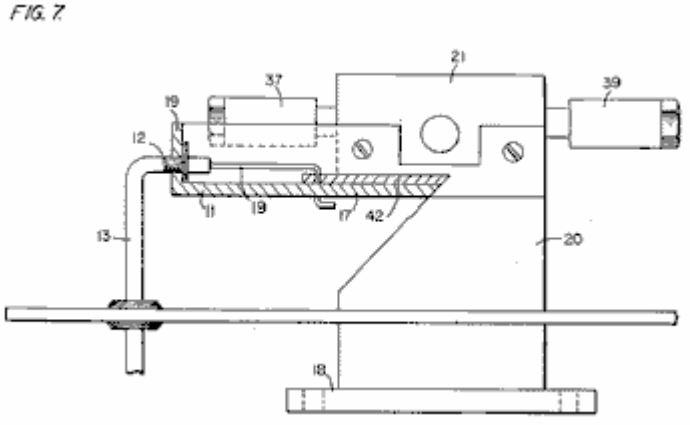


Fig.8 is a top view of the air intake valve to fuel mixing chamber.

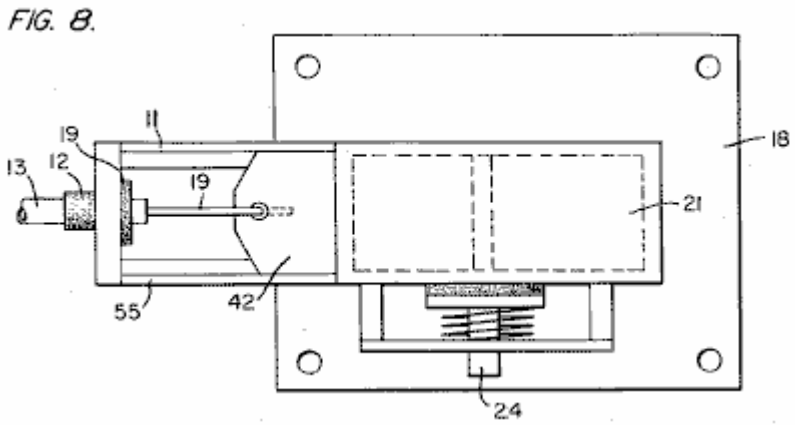
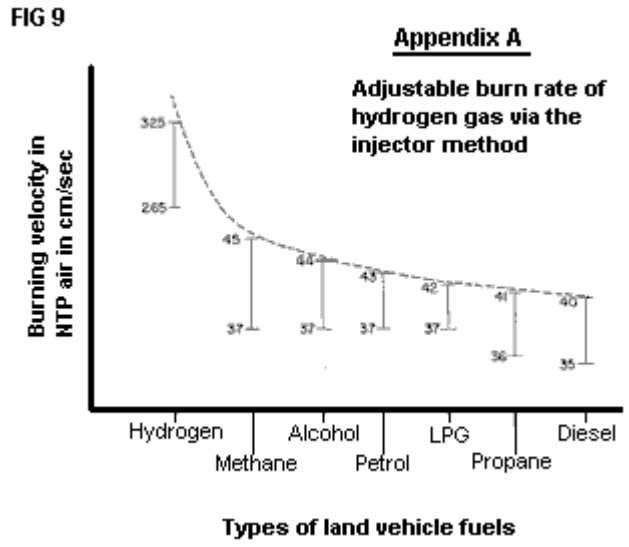
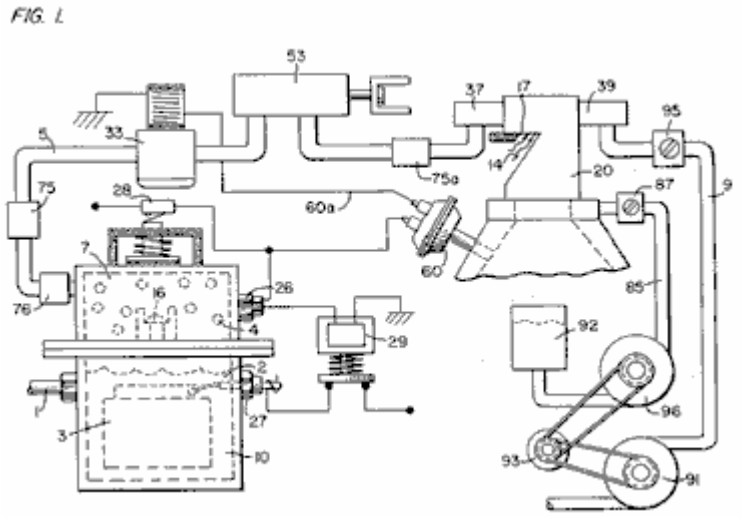


Fig.9 is a comparison of the burning velocity of hydrogen with respect to other fuels.



DETAILED DESCRIPTION OF INVENTION TAKEN WITH DRAWINGS:



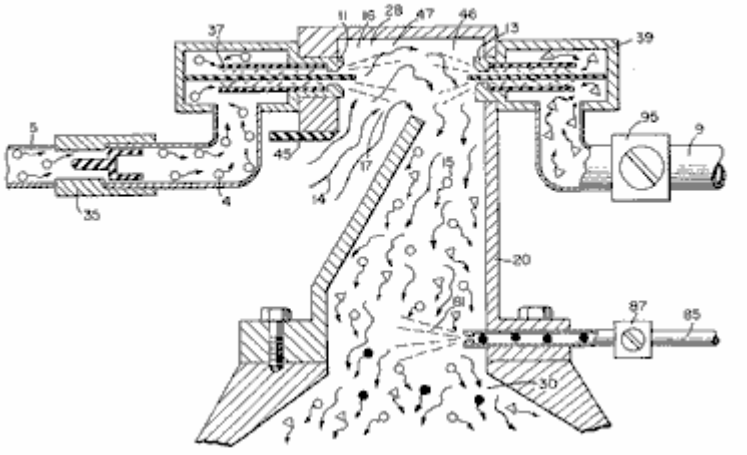
Referring to **Fig.1** the complete overall gas mixing and fuel flow system is illustrated together for utilisation in a combustion engine, particularly an engine in a car. With specific reference to **Fig.1**, the hydrogen source **10** is the hydrogen generator disclosed and described in my co-pending application, supra. The container **10** is an enclosure for a water bath **2**. Immersed in the water **2** is an array of plates **3** as further described in my co-pending application, supra. Applied to plates **3** is a source of direct current potential via electrical inlet **27**. The upper portion **7** of the container **10** is a hydrogen storage area maintaining a predetermined amount of pressure. In this way, there will be an immediate flow of hydrogen gas at start-up.

To replenish the expended water, the generator provides a continuous water source **1**. Thereafter, the generator is operable as described in the aforesaid patent application. The safety valve **28** is designed to rupture should there be an excessive build-up of gas. Switch **26** is a gas-pressure switch included to maintain a predetermined gas pressure level about a regulated low-volume.

The generated hydrogen gas **4** is fed from the one-way check valve **16** via pipe **5** to a gas-mixing chamber **20**, where the hydrogen gas is mixed with non-combustible gasses via pipe **9** from a source described later.

If the one-way valve **75** failed, there could be a return spark which could ignite the hydrogen gas **4** in the storage area **7** of the hydrogen generator **10**. To prevent this, the quenching assembly **76** has been included to prevent just such an ignition.

FIG. 2.



With particular reference to **Fig.2,** the hydrogen gas (via pipe **5**) and non-combustible gasses (via pipe **9**), are fed to a carburettor (air-mixture) system **20** also having an air intake **14** for ambient air.

The hydrogen gas **4** is fed via line **5** through nozzle **11** in a spray **16** in to the trap area **46** of the mixing chamber **20**. Nozzle **11** has an opening smaller than the plate openings in the quenching assembly **37**, thereby preventing flash-back in the event of sparking. The non-volatile gasses are injected into mixing chamber **20** trap area **47** in a jet spray **17** via nozzle **13**. Quenching assembly **39** is operable much in the same manner as quenching assembly **37**.

In the preferred arrangement, the ambient air is the source of oxygen necessary for the combustion of the hydrogen gas. Further, as disclosed in the aforesaid co-pending application, the non-volatile gasses are in fact, the exhaust gasses passed back via a closed loop system. It is to be understood that the oxygen and/or the non-combustible gasses might also be provided from an independent source.

With continued reference to **Fig.2** the gas trap area **47** is a predetermined size. As hydrogen is lighter than air, the hydrogen will rise and become trapped in area **47**. Area **47** is large enough to contain enough hydrogen gas to allow instant ignition upon the subsequent start-up of the combustion engine.

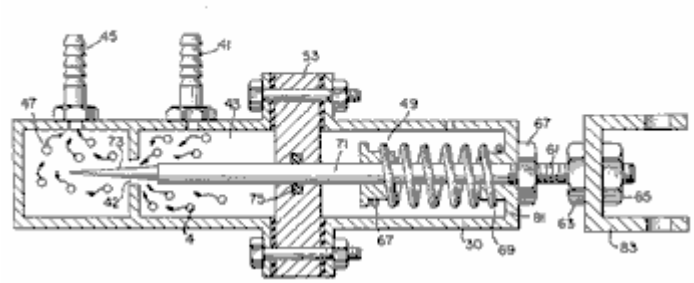
It will be noted that the hydrogen gas is injected in the uppermost region of the trap area **47**. Hydrogen rises at a much greater rate than oxygen or the non-combustible gasses; perhaps three times or greater. Therefore, if the hydrogen gas entered the trap area **47** (mixing area) at its lowermost region the hydrogen gas would rise so rapidly that the air could not mix with the oxygen. With the trap area **47** shown in **Fig.2**, the hydrogen is forced downwards into the air intake **15**. That is, the hydrogen gas is forced downwards into the upwardly forced air and this causes adequate mixing of the gasses.

The ratio of the ambient air (oxygen) **14** and the non-combustible gas via line **9** is a controlled ratio which is tailored to the particular engine. Once the proper combustion rate has been determined by the adjustment of valve **95** (for varying the amount of the non-combustible gas) and the adjustment of valve **45** (for varying the amount of the ambient air), the ratio is maintained thereafter.

In a system where the non-combustible gasses are the exhaust gasses of the engine itself, passed back through a closed loop-arrangement, and where the air intake is controlled by the engine, the flow velocity and hence the air/non-combustible mixture, is maintained by the acceleration of the engine.

The mixture of air with non-combustible gasses becomes the carrier for the hydrogen gas. That is, the hydrogen gas is mixed with the air/non-combustible gas mixture. By varying the amount of hydrogen gas added to the air/non-combustible mixture, the engine speed is controlled.

FIG. 3.



Reference is made to **Fig.3** which shows in a side view cross-section, the fine linear fuel flow control **53**. The hydrogen gas **4** enters chamber **43** via gas inlet **41**. The hydrogen gas passes from chamber **43** to chamber **47** via port or opening **42**. The amount of gas passing form chamber **43** to chamber **47** is dictated by the setting of the port opening **42**.

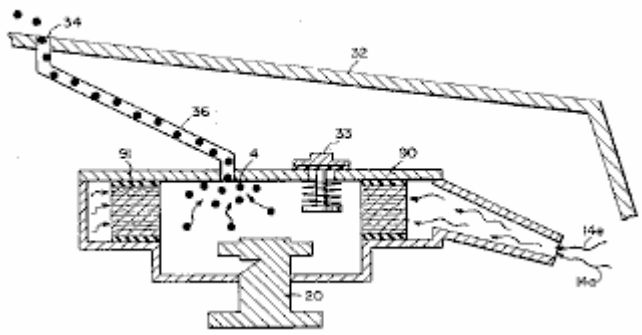
The port opening is controlled by inserting the linearly tapered pin **73** into it. The blunt end of pin **73** is fixed to rod **71**. Rod **71** is passed, (via supporting O-ring **75**), through opening **81** in housing **30**, to the manual adjustment mechanism **83**.

Spring **49** retains the rod **71** in a fixed position relative to pin **73** and opening **42**. When mechanism **83** is operated, pin **73** moves back from the opening **42**. As pin **73** is tapered, this backward movement increases the free area of opening **42**, thereby increasing the amount of gas passing from chamber **43** to chamber **47**.

The stops **67** and **69** maintain spring **49** in its stable position. The nuts **63** and **67** on threaded rod **61** are used to set the minimum open area of opening **42** by the correct positioning of pin **73**. This minimum opening setting, controls the idle speed of the engine, so pin **73** is locked in its correct position by nuts **63** and **67**. This adjustment controls the minimum rate of gas flow from chamber **43** to chamber **47** which will allow continuous operation of the combustion engine.

Referring now to **Fig.8** which illustrates the air adjustment control for manipulating the amount of air passing into the mixing chamber **20**. The closure **21** mounted on plate **18** has an opening **17** on end **11**. A plate-control **42** is mounted so as to slide over opening **17**. The position of this plate, relative to opening **17**, is controlled by the position of the control rod **19** which passes through grommet **12** to control line **13**. Release valve **24** is designed to rupture should any malfunction occur which causes the combustion of the gasses in mixing chamber **20**.

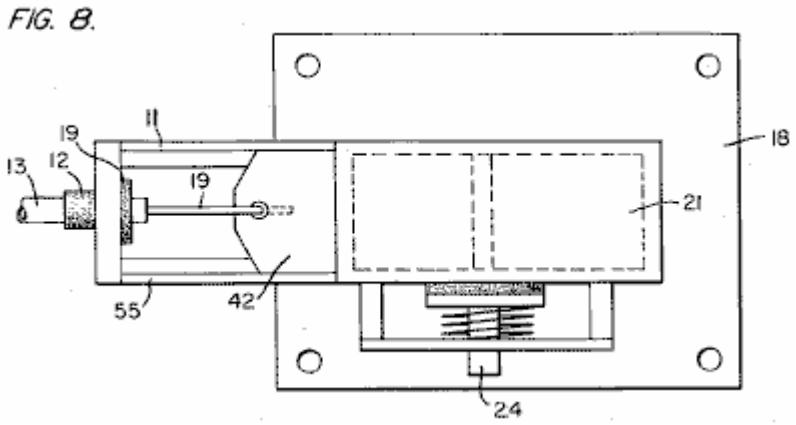
FIG. 4.



With reference now to **Fig.4**, if hydrogen gas **4** were to accumulate in mixing chamber **20** and reach an excessive pressure, the escape tube **36** which is connected to port **34** (located on the car bonnet **32**), permits the excess hydrogen gas to escape safely to the atmosphere. In the event of a malfunction which causes the combustion of the gasses in mixing chamber **20**, the pressure relief valve **33** will rupture, expelling the hydrogen gas without combustion.

In the constructed arrangement of **Fig.1**, there is illustrated a gas control system which may be fitted to an existing car's internal combustion engine without changing or modifying the car's design parameters or characteristics. The flow of the volatile hydrogen gas is, of course, critical; therefore, there is incorporated in line **5** a gas-flow valve **53**, and this is used to adjust the hydrogen flow. This gas-flow valve is shown in detail in **Fig.3**.

FIG. 8.

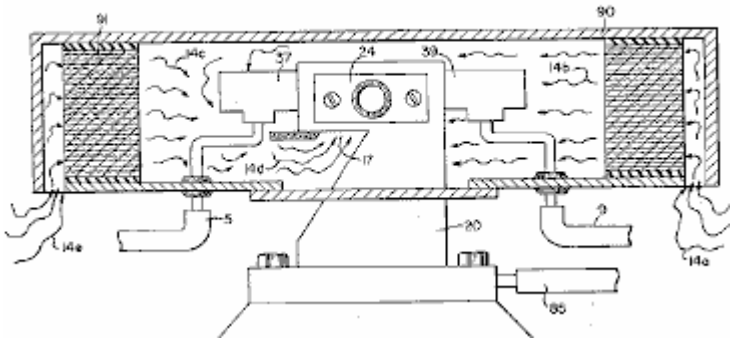


The intake air **14** may be in a carburettor arrangement with an intake adjustment **55** which adjusts the plate **42** opening. This is shown more fully in **Fig.8**. To maintain constant pressure in hydrogen gas storage **7** in the on-off operation of the engine, the gas flow control valve is responsive to the electrical shut-off control **33**. The constant pressure permits an abundant supply of gas on start-up and during certain periods of running time in re-supply.

The switch **33** is in turn responsive to the vacuum control switch **60**. During running of the engine vacuum will be built up which in turn leaves switch **33** open by contact with vacuum switch **60** through lead **60a**. When the engine is not running the vacuum will decrease to zero and through switch **60** will cause electrical switch **33** to shut off cutting off the flow of hydrogen gas to the control valve **53**.

As low-voltage direct current is applied to safety valve **28**, solenoid **29** is activated. The solenoid applies a control voltage to the hydrogen generator exciter **3** via terminal **27** through pressure switch **26**. As the electrical power activates solenoid **29**, hydrogen gas is caused to pass through flow adjustment valve **16** and then outlet pipe **5** for utilisation. The pressure differential hydrogen gas output to gas mixing chamber **20** is for example 30 lbs. to 15 lbs. Once hydrogen generator **10** reaches an optimum gas pressure level, pressure switch **26** shuts off the electrical power to the hydrogen excitors. If the chamber pressure exceeds a predetermined level, the safety release valve **28** is activated disconnecting the electrical current and thereby shutting down the entire system for safety inspection.

FIG. 6.



With particular reference now to **Fig.6** which illustrates the fuel injector system in a side cross-sectional view and to **Fig.5** the top view. The structural apparatus incorporated in the preferred embodiment comprises housing **90** which has air intakes **14a** and **14e**. The air passes through filter **91** around the components **14b** and **14c** and then to intake **14d** of the mixing chamber **20**. The hydrogen enters via line **5** via quenching plates **37** and into the mixing chamber **20**. The non-volatile gasses pass via line **9** to the quenching plates **39** and into the mixing chamber **20**.

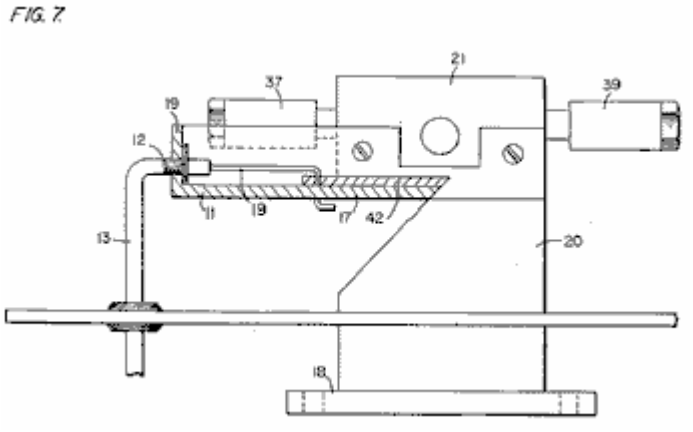
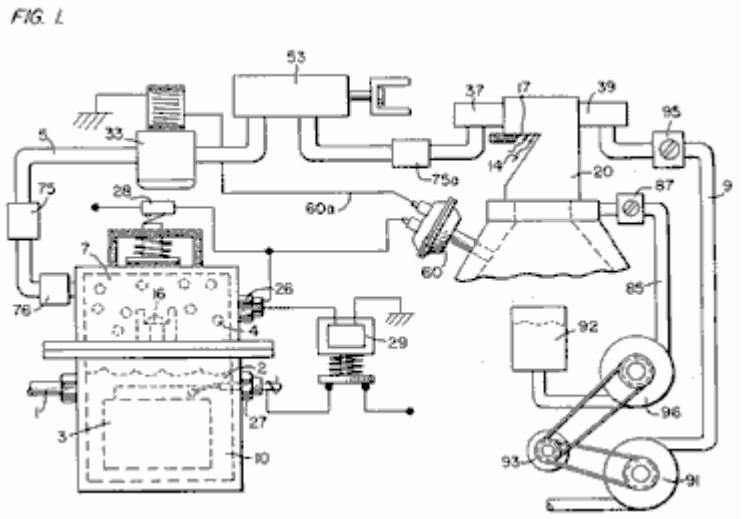


Fig.7 illustrates the mechanical arrangement of the components which make up the overall structure of mixing chamber **20** (shown independently in the other figures).



Returning to **Fig.1** there is illustrated the non-volatile gas line **9** passing through mixture pump **91** by engine pulley **93**. Valve **95** controls the rate of flow. Also driven by pulley **93** is pump **96** having line **85** connected to an oil reservoir **92** and valve **87** and finally to mixing chamber **20**. As a practical matter, such as in a non-oil lubricated engine, lubricating fluid such as oil **81** is sprayed in the chamber **20**, via oil supply line **85** for lubrication.

There have been several publications in the past year or so, delving into the properties of Hydrogen gas, its potential use, generating systems, and safety. One such publication is "Selected Properties of Hydrogen" (Engineering Design Data) issued February 1981 by the National Bureau of Standards.

These publications are primarily concerned with the elaborate and costly processes for generating hydrogen. Equally so, they are concerned with the very limited use of hydrogen gas because of its extremely high burning velocities. This in turn reflects the danger in the practical use of hydrogen.

With reference to the graph of the Appendix A, it is seen that the burning velocities of alcohol, propane, methane, petrol, Liquid Petroleum Gas, and diesel oil are in the range of minimum 35 to maximum 45. Further, the graph illustrates that the burning velocity of hydrogen gas is in the range of 265 minimum to 325 maximum. In simple terms, the burning velocity of hydrogen is of the order of 7.5 times the burning velocity of ordinary commercial fuels.

Because of the unusually high burning velocity of hydrogen gas, it has been ruled out as a substitute fuel, by these prior investigators. Further, even if an engine could be designed to accommodate such high burning velocities, the danger of explosion would eliminate any thoughts of commercial use.

The present invention, as above described, has resolved the above-noted criteria for the use of hydrogen gas in a standard commercial engine. Primarily, the cost in the generation of hydrogen gas, as noted in the aforementioned co-pending patent applications, is minimal. Water with no chemicals or metals is used. Also, as noted in the aforementioned co-pending patent applications, the reduction in the hydrogen gas burn velocity has

been achieved. These co-pending applications not only teach the reduction in velocity, but teach the control of the velocity of the hydrogen gas.

In the preferred embodiment, practical apparatus adapting the hydrogen generator to a combustion engine is described. The apparatus linearly controls the hydrogen gas flow to a mixing chamber mixing with a controlled amount of non-combustible gas oxygen, hence, the reduction in the hydrogen gas velocity. The reduction in the hydrogen gas velocity makes the use of hydrogen as safe as other fuels.

In more practical terms the ordinary internal combustion engine of any size or type of fuel, is retrofitted to be operable with only water as a fuel source. Hydrogen gas is generated from the water without the use of chemicals or metals and at a very low voltage. The burning velocity of the hydrogen gas has been reduced to that of conventional fuels. Finally, every component or step in the process has one or more safety valves or features thereby making the hydrogen gas system safer than that of conventional cars.

In the above description the terms ‘non-volatile’ and ‘non-combustible’ were used. It is to be understood they are intended to be the same; that is, simply, gas which will not burn.

Again, the term ‘storage’ has been used, primarily with respect to the hydrogen storage area 7. It is not intended that the term ‘storage’ be taken literally - in fact, it is not storage, but a temporary holding area. With respect to area 7, this area retains a sufficient amount of hydrogen for immediate start-up.

Other terms, features, apparatus, and the such have been described with reference to a preferred embodiment. It is to be understood modifications and alternatives can be had without departing from the spirit and scope of the invention.

STANLEY MEYER

US Patent 4,421,474 December 1983 Inventor: Stanley A. Meyer

HYDROGEN GAS BURNER

Please note that this is a re-worded excerpt from this patent. It describes how to burn the hydrogen and oxygen gas mix produced by electrolysis of water. Normally, the flame produced is too hot for practical use other than cutting metal or welding. This patent shows a method of reducing the flame temperature to levels suitable for general use in boilers, stoves, heaters, etc.

ABSTRACT

A hydrogen gas burner for the mixture of hydrogen gas with ambient air and non-combustible gasses. The mixture of gasses when ignited provides a flame of extremely high, but controlled intensity and temperature.

The structure comprises a housing and a hydrogen gas inlet directed to a combustion chamber positioned within the housing. Air intake ports are provided for adding ambient air to the combustion chamber for ignition of the hydrogen gas by an ignitor therein. At the other end of the housing there is positioned adjacent to the outlet of the burner (flame) a barrier/heating element. The heating element uniformly disperses the flame and in turn absorbs the heat. The opposite side to the flame, the heating element uniformly disperses the extremely hot air. A non-combustible gas trap adjacent to the heating element captures a small portion of the non-combustible gas (burned air). A return line from the trap returns the captured non-combustible gas in a controlled ratio to the burning chamber for mixture with the hydrogen gas and the ambient air.

CROSS REFERENCE

The hydrogen/oxygen generator utilised in the present invention is that disclosed and claimed in my co-pending patent application, Serial. No.: 302,807, filed: Sept. 16, 1981, for: HYDROGEN GENERATOR SYSTEM. In that process for separating hydrogen and oxygen atoms from water having impurities, the water is passed between two plates of similar non-oxidising metal. No electrolyte is added to the water. The one plate has placed thereon a positive potential and the other a negative potential from a very low amperage direct-current power source. The sub-atomic action of the direct current voltage on the non-electrolytic water causes the hydrogen and oxygen atoms to be separated--and similarly other gasses entrapped in the water such as nitrogen. The contaminants in the water that are not released are forced to disassociate themselves and may be collected or utilised and disposed of in a known manner.

The direct current acts as a static force on the water molecules; whereas the non-regulated rippling direct current acts as a dynamic force. Pulsating the direct current further enhances the release of the hydrogen and oxygen atoms from the water molecules.

In my co-pending patent application, Serial. No. 262,744, filed: May 11, 1981, for: HYDROGEN AERATION PROCESSOR, there is disclosed and claimed the utilisation of the hydrogen/oxygen gas generator. In that system, the burn rate of the hydrogen gas is controlled by the controlled addition of non-combustible gasses to the mixture of hydrogen and oxygen gasses.

SUMMARY OF INVENTION

The present invention is for a hydrogen gas burner and comprises a combustion chamber for the mixture of hydrogen gas, ambient air, and non-combustible gasses. The mixture of gasses is ignited and burns at a retarded velocity rate and temperature from that of hydrogen gas, but at a higher temperature rate than other gasses.

The extremely narrow hydrogen gas mixture flame of very high temperature is restricted from the utilisation means by a heat absorbing barrier. The flame strikes the barrier which in turn disperses the flame and absorbs the heat therefrom and thereafter radiates the heat as extremely hot air into the utilisation means.

Positioned on the opposite side of the heat radiator/barrier is a hot air trap. A small portion of the radiated heat is captured and returned to the combustion chamber as non-combustible gasses. Valve means in the return line regulates the return of the non-combustible gas in a controlled amount to control the mixture.

The present invention is principally intended for use with the hydrogen generator of my co-pending patent application, supra; but it is not to be so limited and may be utilised with any other source of hydrogen gas.

OBJECTS

It is accordingly a principal object of the present application to provide a hydrogen gas burner that has a temperature controlled flame and a heat radiator/barrier.

Another object of the present invention is to provide a hydrogen gas burner that is capable of utilising the heat from a confined high temperature flame.

Another object of the present invention is to provide a hydrogen gas burner that is retarded from that of hydrogen gas, but above that of other gasses.

Another object of the present invention is to provide a hydrogen gas burner that utilises the exhaust air as non-combustible gas for mixture with the hydrogen gas.

Another object of the present invention is to provide a hydrogen gas burner that is simple but rugged and most importantly safe for all intended purposes.

Other objects and features of the present invention will become apparent from the following detailed description when taken in conjunction with the drawings in which:

BRIEF DESCRIPTION OF THE DRAWINGS

Fig.1 is an overall cross-sectional view of the present invention in its most preferred embodiment.

Fig. 1

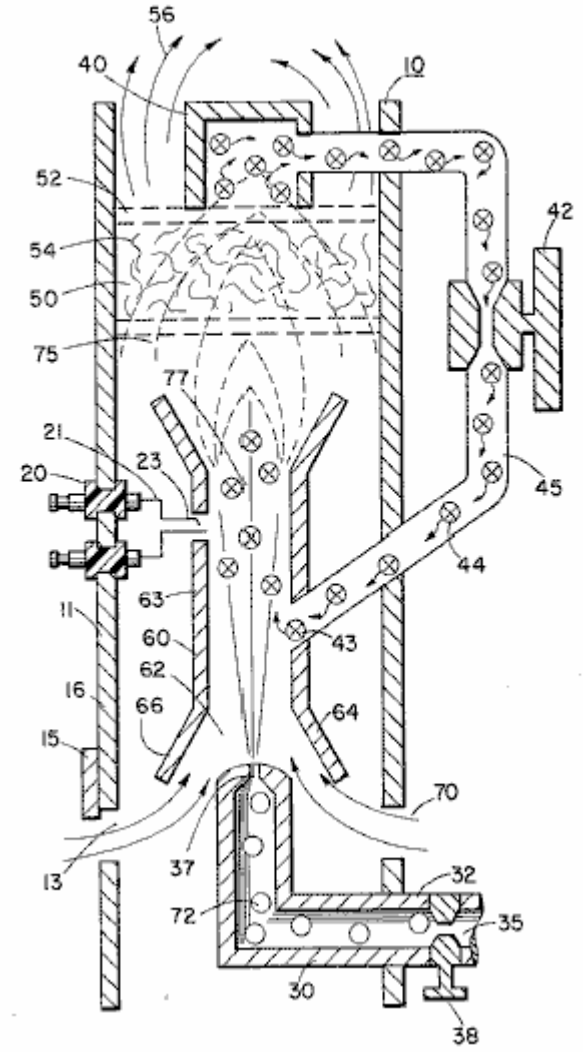
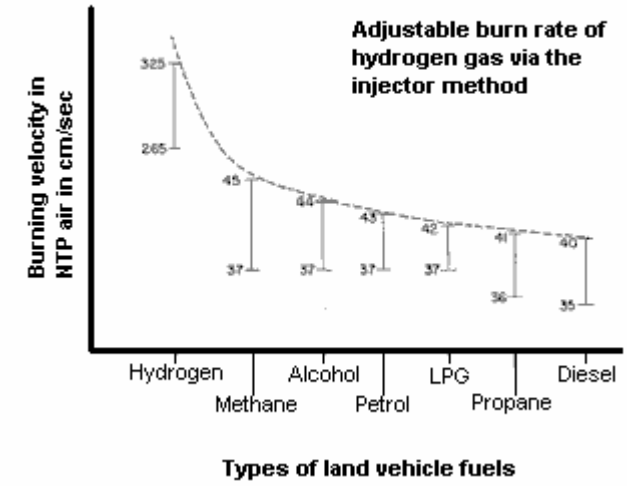


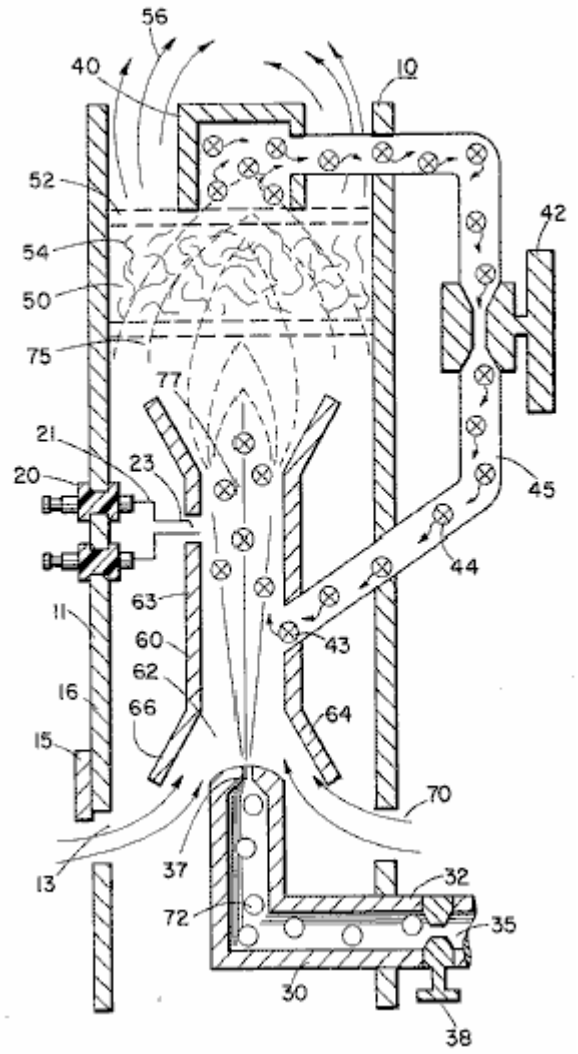
Fig.2 is a graphical illustration of the burning of various standard fuels with that of hydrogen velocities.

Fig. 2



DETAILED DESCRIPTION OF INVENTION

Fig. 1



With particular reference **Fig.1** there is illustrated in a schematic cross-section the principals of the present invention. The structure of the preferred embodiment comprises a housing **10**, having an igniter **20** extending through the wall **11** thereof. A combustion chamber **60** positioned within the housing **10** has a first open end **62**. A hydrogen gas **72** inlet **30** directs hydrogen gas via port **37** from a source **35** to the inlet **62** of the combustion chamber **68**. Also directed to the same inlet **62**, and assisted by flanges **64** and **66**, is ambient air **70** entering through ports **13** in the housing **10**.

Adjacent the opposite end of the combustion chamber **60** the gas mixture **75** is ignited by the ignitor **20** to produce flame **77**. The velocity of the flame **77** causes it to strike and penetrate the barrier/radiator **50**. The barrier **50** is of a material, such as metallic mesh or ceramic material, to disperse therein the flame and in turn become saturated with heat. The flame **77** is of a size sufficient to be dispersed throughout the barrier **50**, but yet, not penetrate through the barrier **50**.

Radiated from the surface **52** of the barrier **50** is superheated air **56** (gasses) to be passed on to a utilisation device. Adjacent to surface **52** of barrier/radiator **50** is a hot air trap **40** with closed loop line **45** returning non-combustible gas **44** to the combustion chamber **60**. Control valve **42** is intermediate the line **45**.

In operation of the preferred embodiment hydrogen gas, **72**, emitted from the nozzle **37** is directed to the combustion chamber **60**. The flanges **64** and **66** on the open end of housing **63** of the combustion chamber **60** enlarges the open end of **62**. In the enlargement ambient air from the opening **13** in the housing **10** is also directed to the combustion chamber **60**.

The ambient air and hydrogen traverses the opening **43** and further mixes with the non-combustible gas **44** from the closed loop line **45** with the hot air trap **40**. The mixture of hydrogen gas **72**, ambient air **70**, and non-combustible gas **44**, is ignited by the ignitor **20** having electrical electrodes **21** and **23**. Upon ignition flame **77** ensues. The mixture is controlled with each of three gasses. That is, the line **32** from the hydrogen source **35** has a valve **38** therein for controlling the amount of hydrogen **72** emitted from the nozzle **37**. The opening **13** has a

plate adjustment **15** for controlling the amount of ambient air **60** directed to the combustion chamber **60**, and the closed-loop line has valve **42**, as aforesaid, for controlling the amount of non-combustible gasses in the mixture.

It can be appreciated that the temperature of the flame **77** and the velocity of the flame **77** is a function of the percentage of the various gasses in the mixture. In a practical embodiment, the flame **70** temperature and velocity was substantially retarded from that of a hydrogen flame per se; but yet, much greater than the temperature and velocity of the flame from the gasses utilised in a conventional heating system.

To maintain a sufficient pressure for combustion of the hydrogen gas mixture with a minimum of pressure (for safety) and to limit blow-out, the nozzle **37** opening **39** is extremely small. As a consequence, if the hydrogen gas were burned directly from the nozzle **37**, the flame would be finite in diameter. Further, its velocity would be so great it is questionable whether a flame could be sustained. The mixing of ambient air and non-combustible gas does enlarge the flame size and reduce its velocity. However, to maintain a flame higher in temperature and velocity than the conventional gasses, the size and temperature of the flame is controlled by the mixture mentioned earlier.

Therefore, to utilise the flame **77** in a present day utilisation means, the flame is barred by the barrier **50**. The barrier **50** is of a material that can absorb safely the intense flame **77** and thereafter radiate heat from its entire surface **52**. The material **54** can be a ceramic, metallic mesh or other heat absorbing material known in the art. The radiated heat **56** is directed to the utilisation means.

As stated earlier, the mixture of gasses which are burned include non-combustible gasses. As indicated in the above-noted co-pending patent applications, an excellent source of non-combustible gasses is exhaust gasses. In this embodiment, the trap **50** entraps the hot air **74** and returns the same, through valve **42**, to the combustion chamber **60** as non-combustible gas.

With reference to **Fig.2** there is illustrated the burning velocity of various standard fuels. It can be seen the common type of fuel burns at a velocity substantially less than hydrogen gas. The ratio of hydrogen with non-combustible oxygen gasses is varied to obtain optimum burning velocity and temperature for the particular utilisation. Once this is attained, the ratio, under normal conditions, will not be altered. Other uses having different fuel burn temperature and velocity will be adjusted in ratio of hydrogen/oxygen to non-combustible gasses in the same manner as exemplified above.

Further, perhaps due to the hydrogen gas velocity, there will occur unburned gas at the flame **77** output. The barrier **50**, because of its material makeup will retard the movement and trap the unburned hydrogen gas. As the superheated air **77** is dispersed within the material **54**, the unburned hydrogen gas is ignited and burns therein. In this way the barrier **50** performs somewhat in the nature of an after-burner.

US Patent 5,149,407 22nd September 1992 Inventor: Stanley Meyer

PROCESS AND APPARATUS FOR THE PRODUCTION OF FUEL GAS AND THE ENHANCED RELEASE OF THERMAL ENERGY FROM SUCH GAS

Please note that this is a re-worded excerpt from this patent. It describes in considerable detail, one of Stan's methods for splitting water into hydrogen and oxygen gasses and the subsequent methods for using those gasses.

ABSTRACT

Water molecules are broken down into hydrogen and oxygen gas atoms in a capacitive cell by a polarisation and resonance process dependent on the dielectric properties of water and water molecules. The gas atoms are then ionised or otherwise energised and thermally combusted to release a degree of energy greater than that of combustion of the gas in air.

OBJECTS OF THE INVENTION

A first object of the invention is to provide a fuel cell and a process in which molecules of water are broken down into hydrogen and oxygen gasses, and a fuel gas mixture comprised of hydrogen, oxygen and other gasses formerly dissolved in the water, is produced. A further object of the invention is to realise significant energy-yield from a fuel gas derived from water molecules. Molecules of water are broken down into hydrogen and oxygen gasses. Electrically charged hydrogen and oxygen ions of opposite electrical polarity are activated by electromagnetic wave energy and exposed to a high temperature thermal zone. Significant amounts of thermal energy with explosive force beyond the gas burning stage are released.

An explosive thermal energy under a controlled state is produced. The process and apparatus provide a heat energy source useful for power generation, aircraft rocket engines or space stations.

BRIEF DESCRIPTION OF THE DRAWINGS

Figs.1A through **1F** are illustrations depicting the theoretical bases for phenomena encountered during operation of the fuel gas production stage of the invention.

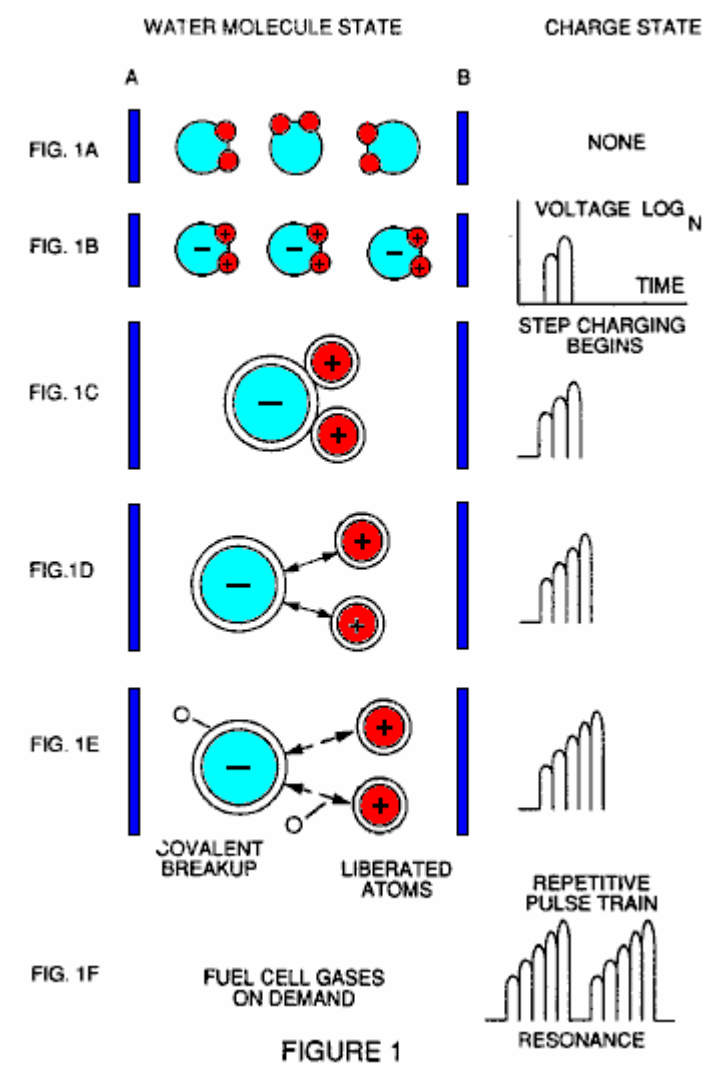


FIGURE 1

Fig.2 illustrates a circuit which is useful in the fuel gas generation process.

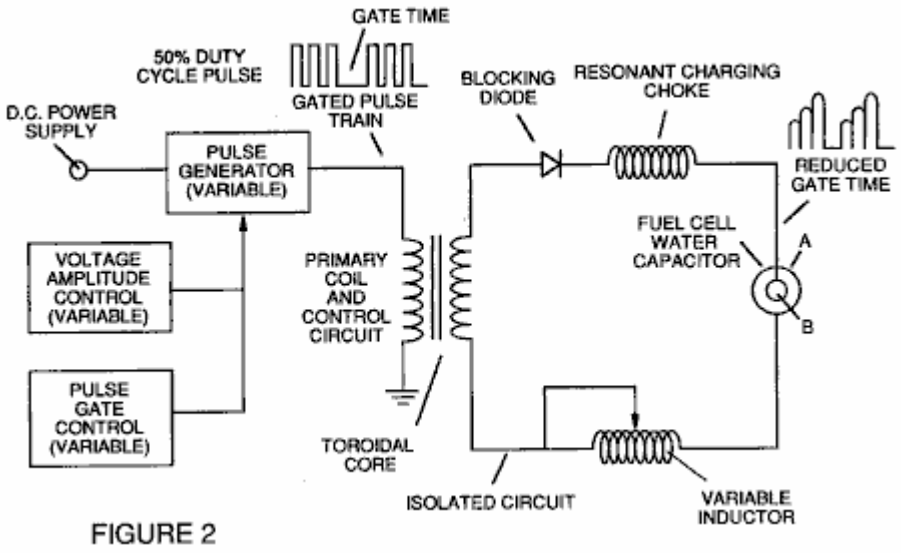


FIGURE 2

Fig.3 shows a perspective of a “water capacitor” element used in the fuel cell circuit.

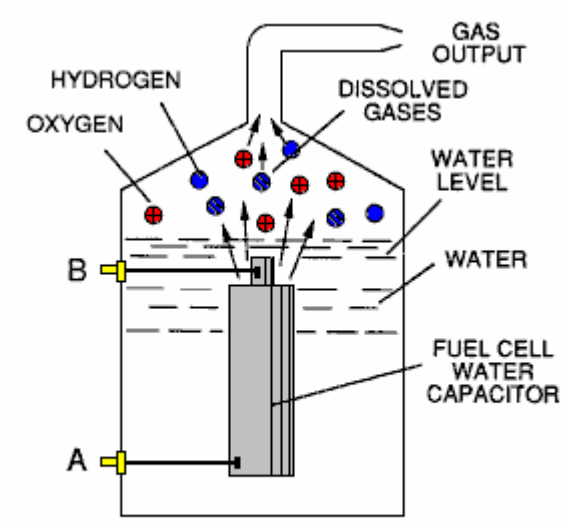


FIGURE 3

Fig.4 illustrates a staged arrangement of apparatus useful in the process, beginning with a water inlet and culminating in the production of thermal explosive energy.

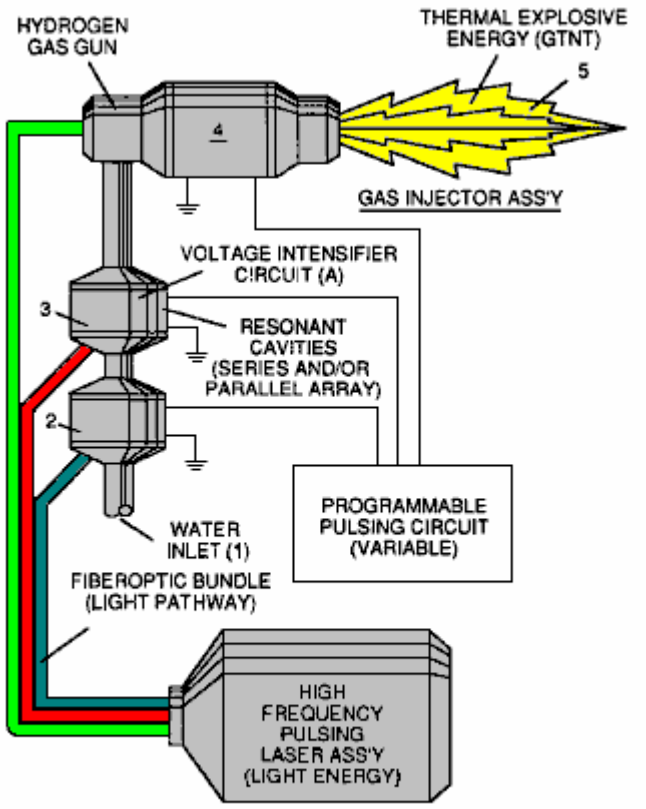


FIGURE 4

Fig.5A shows a cross-section of a circular gas resonant cavity used in the final stage assembly of **Fig.4**

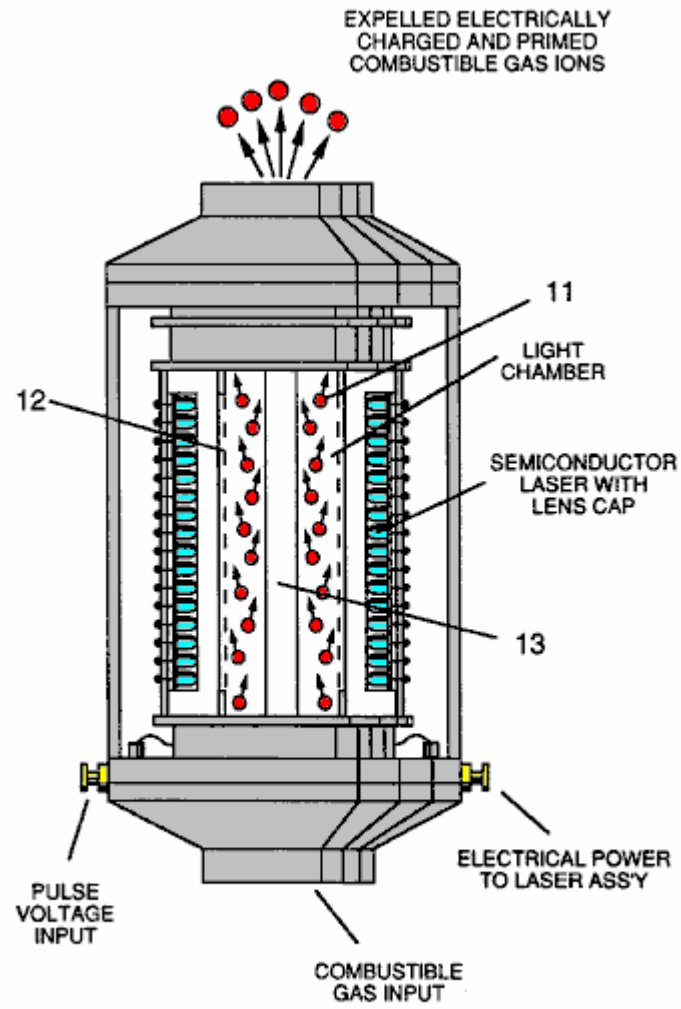


FIGURE 5A

Fig.5B shows an alternative final stage injection system useful in the apparatus of **Fig.4**

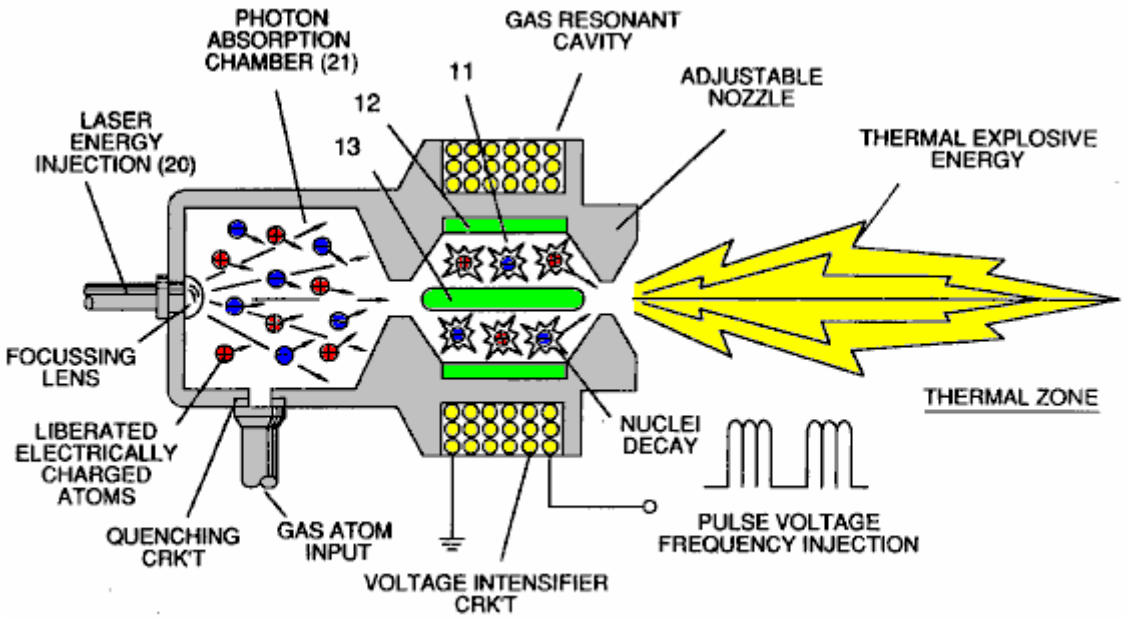


FIGURE 5B

Fig.5C shows an optical thermal lens assembly for use with either final stage of **Fig.5A** or **Fig.5B**.

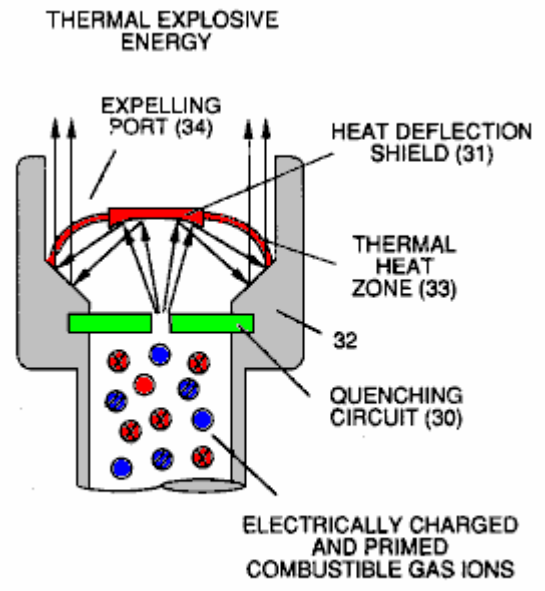


FIGURE 5C

Figs.6A, 6B, 6C and **6D** are illustrations depicting various theoretical bases for atomic phenomena expected to occur during operation of this invention.

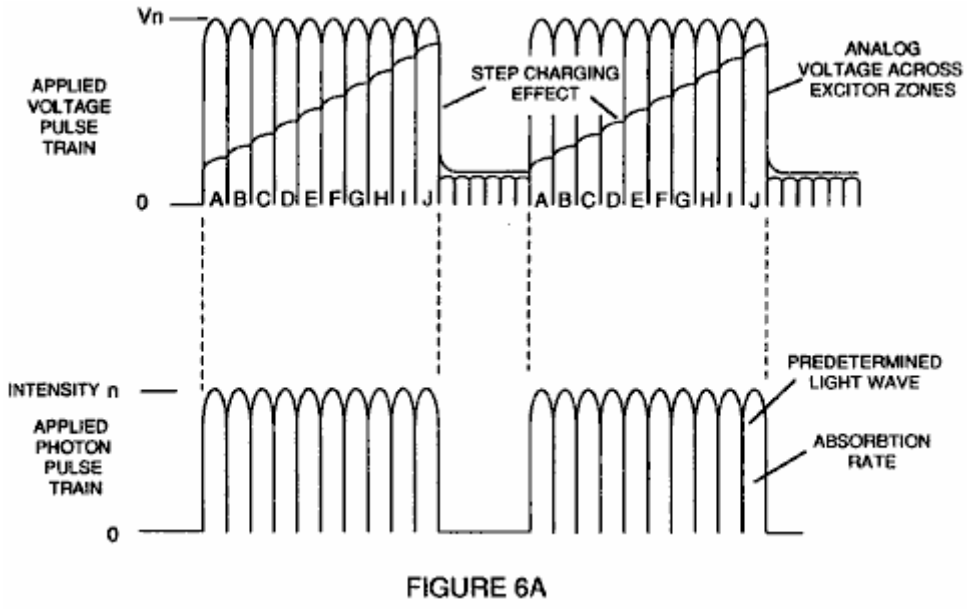


FIGURE 6A

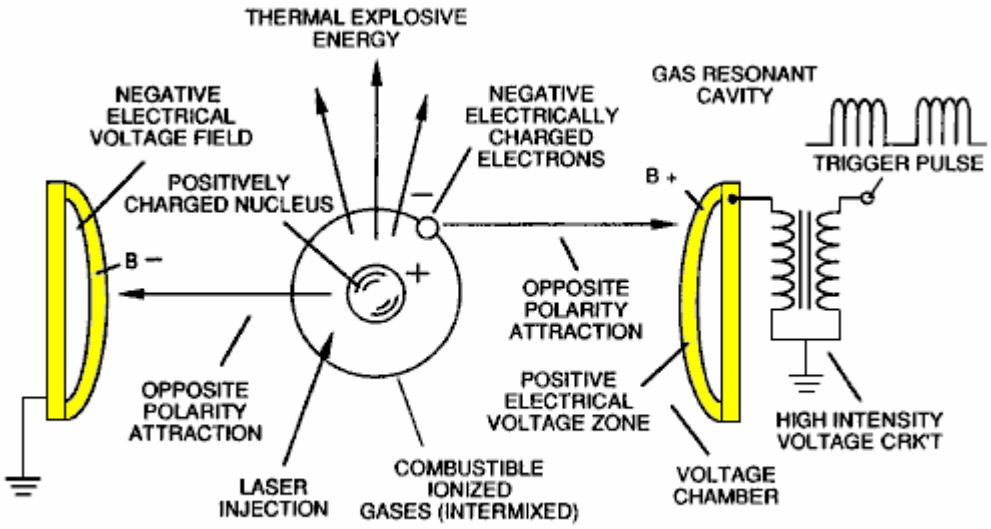


FIGURE 6B

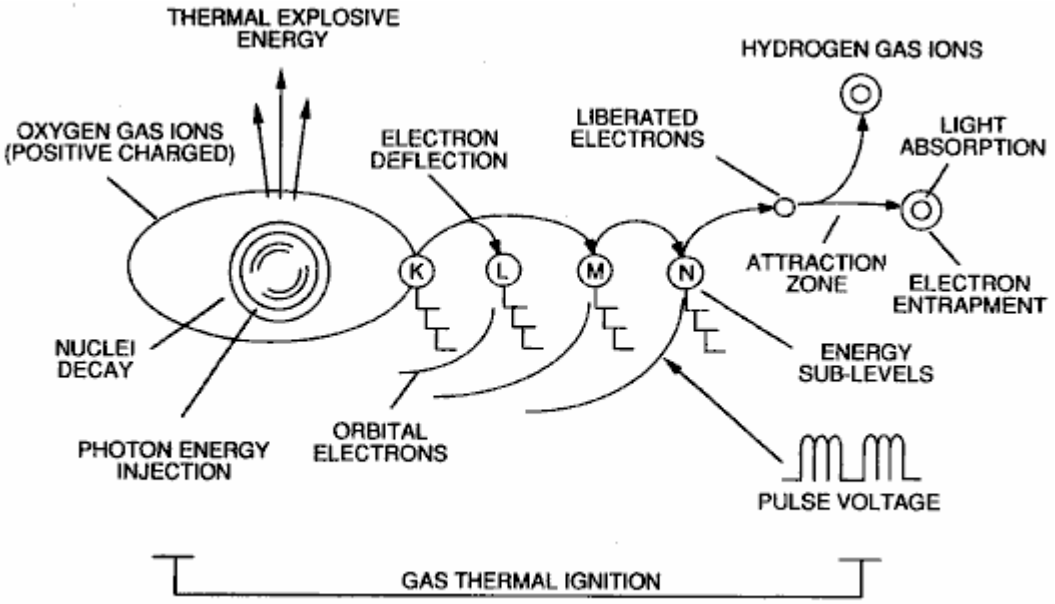


FIGURE 6C

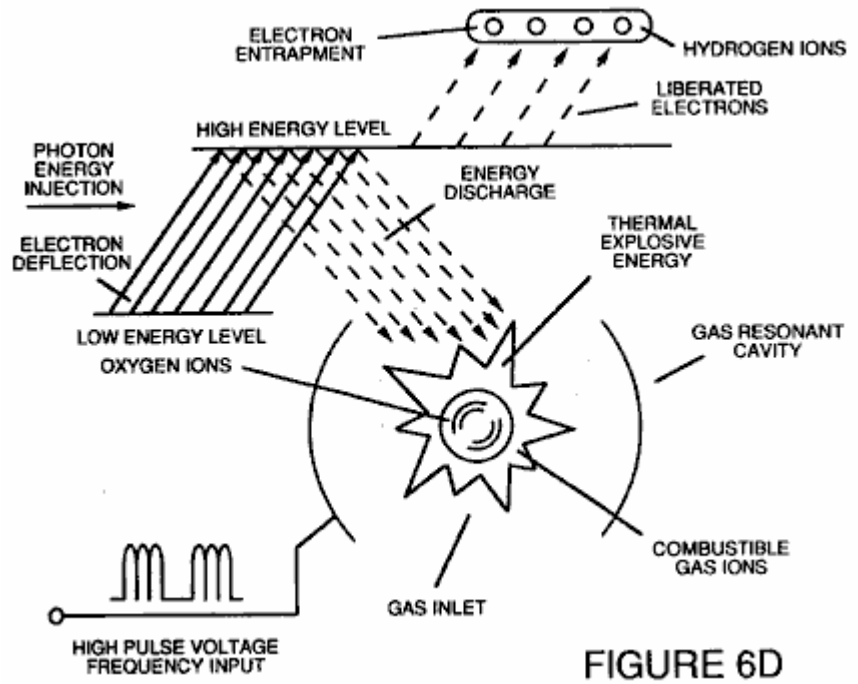


FIGURE 6D

Fig.7 is an electrical schematic of the voltage source for the gas resonant cavity.

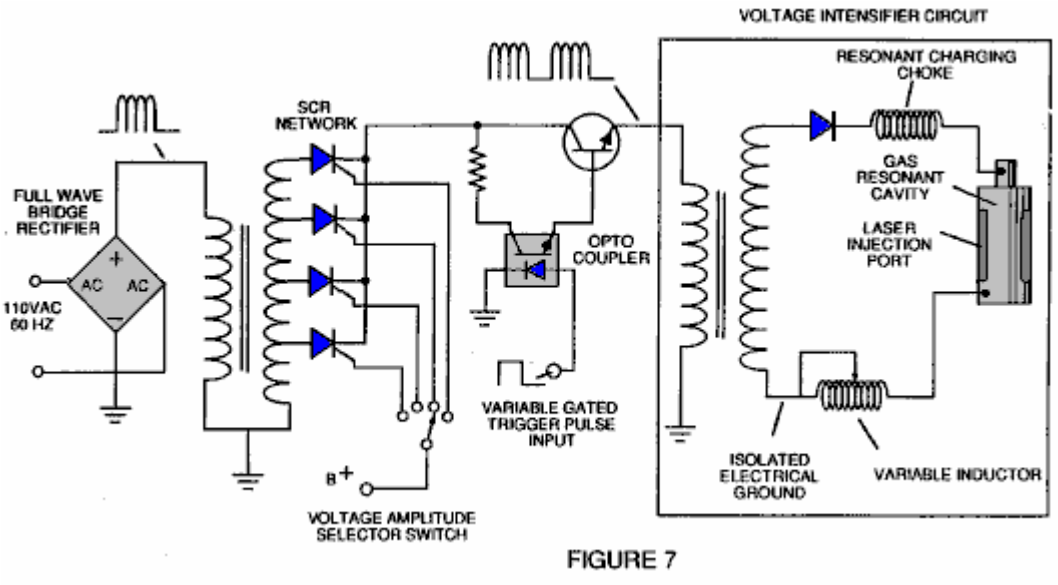


FIGURE 7

Figs.8A and **8B** respectively, show (A) an electron extractor grid used in the injector assemblies of **Fig.5A** and **Fig.5B**, and (B) the electronic control circuit for the extractor grid.

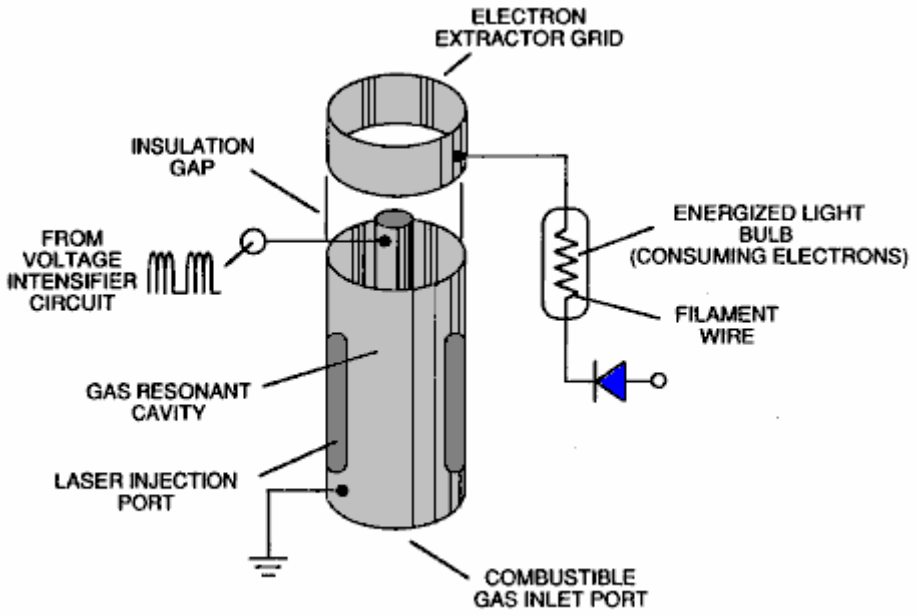


FIGURE 8A

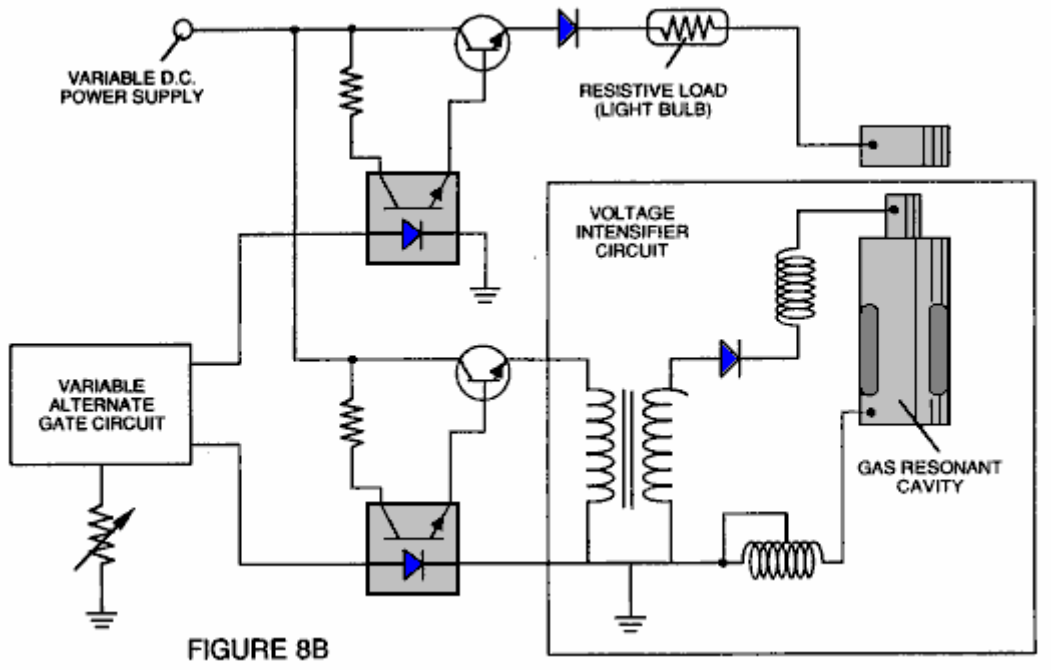


FIGURE 8B

Fig.9 shows an alternative electrical circuit useful in providing a pulsating waveform to the apparatus.

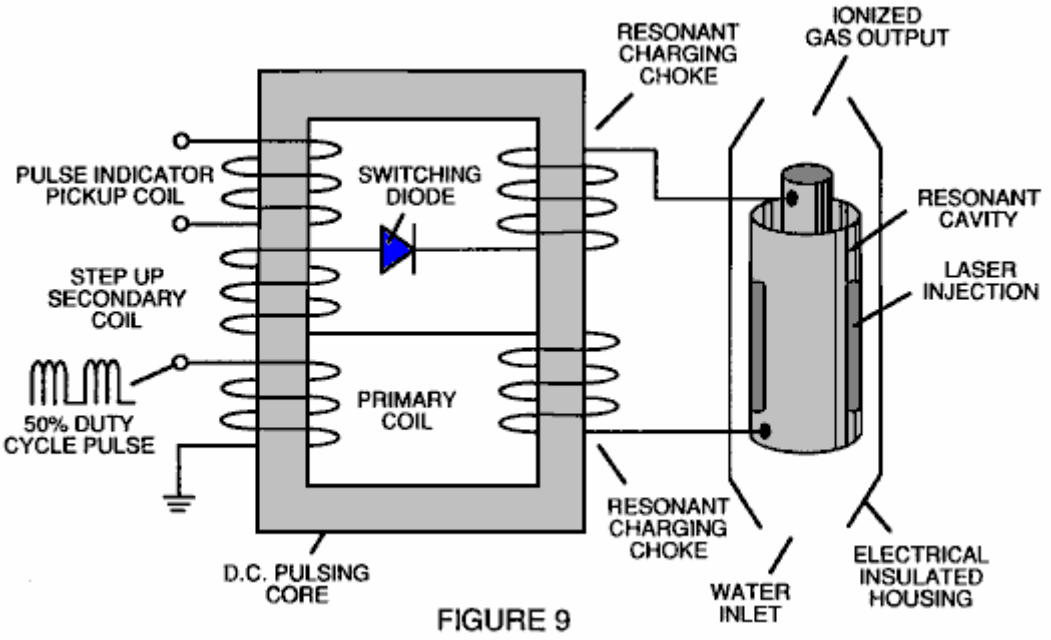


FIGURE 9

TABLE 1: PROCESS STEPS LEADING TO IGNITION

Relative State of Water Molecule and/or Hydrogen/Oxygen/Other Atoms	Stage
Random (ambient state) alignment of polar fields, polarisation of molecules. Molecular elongation. Atom liberation by breakdown of covalent bond	1st Stage: Water to Gas
Release of gasses, Liquid to gas ionisation, Electrical charging effect, Particle Impact	2nd Stage: Gas Ionisation
Electromagnetic Wave, Laser or photon injection, Electron extraction, Atomic destabilisation	3rd Stage: Priming
Thermal Ignition	Final Stage: Ignition

DESCRIPTION OF THE PREFERRED EMBODIMENT

A fuel gas is produced by a hydrogen fracturing process which follows the sequence of steps shown in Table 1. Beginning with water molecules, the molecule is subjected to successively increasing electrical wave energy and thermal forces. In the succession of forces, randomly orientated water molecules are aligned with respect to molecular polar orientation and themselves polarised and "elongated" by the application of an electric potential, to the extent that the co-valent bonding of the water molecules is so weakened that the atoms disassociate and the molecule breaks down into hydrogen and oxygen elemental components. Next, the released atomic gasses are ionised and electrically charged in a vessel while being subjected to a further energy source which promotes inter-particle impact in the gas at an increased overall energy level. Finally, the atomic particles in the excited gas, having achieved successively higher energy levels, are subjected to a laser or electromagnetic wave energy source which produces atomic destabilisation and the final release of thermal explosive energy.

Engineering design parameters based on known theoretical principles of atomic physics, determine the incremental levels of electrical and wave energy input required to produce resonance in each stage of the system. Instead of a dampening effect, a resonant energisation of the molecule, atom or ion provides a compounding energy interaction resulting in the final energy release.

In brief, in the first stage, a gas mixture including hydrogen, oxygen and other gasses formerly dissolved in the water, is obtained from water. In general, the method used in the first stage consists of:

(A) Providing a capacitor, in which the water is included as a dielectric liquid between capacitor plates, in a resonant charging choke circuit, which includes an inductor in series with the capacitor.
(B) Subjecting the capacitor to a pulsating, unipolar electric voltage field in which the polarity does not pass beyond an arbitrary ground, whereby the water molecules within the capacitor are subjected to a charge of the same polarity, and the water molecules are distended by the electrical polar forces.
(C) Further subjecting the water in the capacitor to the pulsating electric field to achieve a pulse frequency which induces a resonance within the water molecule.
(D) Continuing the application of the pulsing frequency to the capacitor cell after resonance occurs so that the energy level within the molecule is increased in cascading incremental steps in proportion to the number of pulses.
(E) Maintaining the charge of the capacitor during the application of the pulsating field, whereby the co-valent electrical bonding of the hydrogen and oxygen atoms within the water molecules is destabilised to such a degree that the force of the electrical field within the molecule exceeds the bonding force of the molecule, causing the molecule to break apart into the elemental gasses of hydrogen and oxygen.
(F) Collecting the hydrogen and oxygen gasses, along with any other gasses formerly dissolved in the water, and discharging the collected gasses as a fuel gas mixture.

The water molecules are subjected to increasing electrical forces. In an ambient state, randomly orientated water molecules are aligned with respect to a molecular polar orientation. Next, they themselves are polarised and "elongated" by the application of an electrical potential to the extent that co-valent bonding of the water molecules is so weakened that the atoms disassociate and the molecule breaks down into hydrogen and oxygen elemental components. In this process, the point of optimum gas release is reached when the circuit is at resonant frequency. Water in the cell is subjected to a pulsating, polar electric field produced by the electrical circuit, whereby the water molecules are distended by the electrical force on the plates of the capacitor. The polar pulsating frequency applied is such that the pulsating electric field induces a resonance in the molecules. A cascade effect occurs, and the overall energy of specific water molecules is increased in cascading incremental steps. The hydrogen and oxygen are released when the resonant energy exceeds the co-valent bonding force of the water molecules.

A preferred construction material for the capacitor plates is stainless steel T-304 which does not react chemically with water, hydrogen or oxygen. An electrically conductive material which is inert in the fluid environment, is a desirable material of construction for the electric field plates of the "water capacitor" employed in the circuit.

Once triggered, the gas output is controllable by the attenuation of operational parameters. Thus, once the frequency of resonance is identified, by varying the applied pulse voltage to the water fuel cell assembly, gas output is varied. By varying the pulse shape, pulse amplitude or pulse train sequence, the gas output can be varied. Attenuation of the voltage field's mark/space ratio of OFF/ON periods also affects the rate of gas production.

The overall apparatus thus includes and electrical circuit in which a water capacitor is an element. The water capacitor has a known dielectric property. The fuel gasses are obtained from the water by the disassociation of the water molecules. The water molecules are split into component atomic elements by a voltage stimulation process called the 'electrical Polarisation process' which also releases dissolved gasses trapped in the water.

From the outline of physical phenomena associated with the first stage of the process described in Table 1, the theoretical basis of the invention considers the respective states of molecules, gasses and ions derived from liquid water. Before voltage stimulation, water molecules are randomly dispersed throughout water in a container.

When a unipolar voltage pulse train such as that shown in **Figs.1B** through **1F** is applied to positive and negative capacitor plates, and increasing voltage potential is induced in the molecules in a linear, step-like charging effect. The electrical field of the particles within a volume of water including the electrical field plates, increases from a low energy state to a high energy state in a step manner following each pulse train as illustrated figuratively in **Figs.1A** through **1F**. The increasing voltage potential is always positive in direct relationship to negative ground potential during each pulse. The voltage polarity on the plates which create the voltage fields remains constant although the voltage charge increases. Positive and negative voltage "zones" are thus formed simultaneously in the electrical field of the capacitor plates.

In the first stage of the process described in Table 1, because the water molecule naturally exhibits opposite electrical fields in a relatively polar configuration (the two hydrogen atoms have a positive charge while the oxygen atom has a negative charge), the voltage pulse causes the water molecules which were initially orientated in random directions, to spin and align themselves with the electrical field applied to the cell. The positively charged hydrogen atoms are attracted to the negative field while the negatively charged oxygen atoms, of the same water molecule, are attracted to the positive voltage field. Even a slight potential difference between the plates of a containment chamber capacitor will initiate the alignment of each water molecule within the cell.

When the voltage applied to the plates causes the water molecules to align themselves, then the pulsing causes the voltage field intensity to be increased in accordance with **Fig.1B**. As further molecular alignment occurs, molecular movement is hindered. Because the positively charged hydrogen atoms of the aligned molecules are attracted in a direction opposite to the negatively charged oxygen atoms, a polar charge alignment or distribution occurs within the molecules between the voltage zones as shown in **Fig.1B**, and as the energy level of the atoms, subjected to resonant pulsing, increases, the stationary water molecules become elongated as shown in **Figs.1C** and **1D**. Electrically charged nuceli and electrons are attracted towards opposite electrically charged voltage zones - disrupting the mass and charge equilibrium of the water molecule.

As the water molecule is further exposed to an increasing potential difference resulting from the step charging of the capacitor, the electrical force of attraction of the atoms within the molecule to the capacitor plates of the chamber also increases in strength. As a result, the co-valent bonding between the atoms of the molecule is weakened and ultimately, terminated. The negatively charged electron is attracted toward the positively charged hydrogen atoms, while at the same time, the negatively charged oxygen atoms repel electrons.

In a more specific explanation of the "sub-atomic action which occurs in the water cell, it is known that natural water is a liquid which has a dielectric constant of 78.54 at 20 degrees Centigrade and 1 atmosphere of pressure [Handbook of Chemistry and Physics, Section E-50].

When a volume of water is isolated and electrically conductive plates that are chemically inert in water and which are separated by a distance, are immersed in the water, a capacitor is formed, having a capacitance determined by the surface area of the plates, the distance of their separation and the dielectric constant of the water.

When water molecules are exposed to voltage at a restricted current, water takes on an electrical charge. By the laws of electrical attraction, molecules align according to positive and negative polarity fields of the molecule and the alignment field. The plates of a capacitor constitute such an alignment field when a voltage is applied across them.

When a charge is applied to a capacitor, the electrical charge of the capacitor equals the applied voltage charge. In a water capacitor, the dielectric property of water resists the flow of current in the circuit, and the water molecule itself, because it has polarity fields formed by the relationship of hydrogen and oxygen in the co-valent bond, and an intrinsic dielectric property, becomes part of the electrical circuit, analogous to a "microcapacitor" within the capacitor defined by the plates.

In the Example of a fuel cell circuit of **Fig.2**, a water capacitor is included. The step-up coil is formed on a conventional torroidal core formed of a compressed ferromagnetic powered material that will not itself become permanently magnetised, such as the trademarked "Ferramic 06# 'Permag'" powder as described in *Siemens Ferrites Catalogue*, CG-2000-002-121, (Cleveland, Ohio) No. F626-1205. The core is 1.50 inch in diameter and 0.25 inch in thickness. A primary coil of 200 turns of 24 AWG gauge copper wire is provided and a coil of 600 turns of 36 AWG gauge wire comprises the secondary winding. Other primary/secondary coil winding ratios may be conveniently determined.

An alternate coil arrangement using a conventional M27 iron transformer core is shown in **Fig.9**. The coil wrap is always in one direction only.

In the circuit of **Fig.2**, the diode is a 1N1198 diode which acts as a blocking diode and an electric switch which allows current flow in one direction only. Thus, the capacitor is never subjected to a pulse of reverse polarity.

The primary coil of the torroid is subject to a 50% duty-cycle pulse. The torroidal pulsing coil provides a voltage step-up from the pulse generator in excess of five times, although the relative amount of step-up is determined by pre-selected criteria for a particular application. As the stepped-up pulse enters the first inductor (formed of 100 turns of 24 gauge wire, 1 inch in diameter), an electromagnetic field is formed around the inductor. Voltage is switched off when the pulse ends, and the field collapses and produces another pulse of the same polarity; i.e. another positive pulse is formed where the 50% duty-cycle was terminated. Thus, a double pulse frequency is produced; however, in a pulse train of unipolar pulses, there is a brief time when pulses are not present.

By being so subjected to electrical pulses in the circuit of **Fig.2**, the water between the capacitor plates takes on an electrical charge which is increased by a step-charging phenomenon occurring in the water capacitor.. Voltage continually increases (to about 1000 volts and more) and the water molecules start to elongate.

The pulse train is then switched off; the voltage across the water capacitor drops to the amount of charge that the water molecules have taken on, i.e. voltage is maintained across the charged capacitor. The pulse train is then applied again.

Because a voltage potential applied to a capacitor can perform work, the higher the voltage potential, the more work is performed by a given capacitor. In an optimum capacitor which is wholly non-conductive, zero current flow will occur across the capacitor. Thus, in view of an idealised capacitor circuit, the object of the water capacitor circuit is to prevent electron flow through the circuit, i.e. such as occurs by electron flow or leakage through a resistive element that produces heat. Electrical leakage in water will occur, however, because of some residual conductivity and impurities, or ions that may otherwise be present in the water. thus, the water capacitor is preferably chemically inert. An electrolyte is not added to the water.

In the isolated water bath, the water molecule takes on charge, and the charge increases. The object of the process is to switch off the co-valent bonding of the water molecule and interrupt the sub-atomic force that binds the hydrogen and oxygen atoms together to form a molecule, thus causing the hydrogen and oxygen to separate.

Because an electron will only occupy a certain electron shell, the voltage applied to the capacitor affects the electrical forces inherent in the co-valent bond. As a result of the charge applied by the plates, the applied force becomes greater than the force of the co-valent bonds between the atoms of the water molecule, and the water molecule becomes elongated. When this happens, the time share ratio of the electrons between the atoms and the electron shells, is modified.

In the process, electrons are extracted from the water bath; electrons are not consumed nor are electrons introduced into the water bath by the circuit, as electrons would be during conventional electrolysis. Nevertheless, a leakage current through the water may occur. Those hydrogen atoms missing electrons become neutralised and atoms are liberated from the water. The charged atoms and electrons are attracted to opposite polarity voltage zones created between the capacitor plates. The electrons formerly shared by atoms in the water co-valent bond are re-allocated so that neutral elemental gasses are liberated.

In the process, the electrical resonance may be reached at all levels of voltage potential. The overall circuit is characterised as a “resonant charging choke” circuit which is an inductor in series with a capacitor [*SAMS Modern Dictionary of Electronics*, 1984 p.859]. Such a resonant charging choke is on each side of the capacitor. In the circuit, the diode acts as a switch which allows the magnetic field produced in the inductor to collapse, thereby doubling the pulse frequency and preventing the capacitor from discharging. In this manner, a continuous voltage is produced across the capacitor plates in the water bath and the capacitor does not discharge. The water molecules are thus subjected to a continuously charged field until the breakdown of the co-valent bond occurs.

As noted initially, the capacitance depends on the dielectric properties of the water and the size and separation of the conductive elements forming the water capacitor.

Example 1

In an example of the circuit of **Fig.2** (in which other circuit element specifications are provided above), two concentric cylinders 4 inches long, formed the water capacitor of the fuel cell in the volume of water. The outside cylinder was o.75 in outside diameter; the inner cylinder was 0.5 inch in outside diameter. Spacing between the inside cylinder and the outside cylinder was 0.0625 inch (1.59 mm). Resonance in the circuit was achieved at a 26 volt pulse applied to the primary coil of the torroid at 10khz and a gas mixture of hydrogen, oxygen and dissolved gasses was given off. The additional gasses included nitrogen and argon from air dissolved in the water.

In achieving resonance in any circuit, as the pulse frequency is adjusted, the current flow is minimised and the voltage on the capacitor plates is maximised. Calculation of the resonant frequency of an overall circuit is determined by known means; different cavities have a different resonant frequency. The gas production rate is varied by the period of time between trains of pulses, pulse amplitude, capacitor plate size and plate separation.

The wiper arm on the second inductor tunes the circuit and allows for contaminants in the water so that the charge is always applied to the capacitor. The voltage applied , determines the rate of breakdown of the molecule into its atomic components. As water in the cell is consumed, it is replaced by any appropriate means or control system.

Thus, in the first stage, which is of itself independently useful, a fuel gas mixture is produced having, in general, the components of elemental hydrogen and oxygen and some additional atmospheric gasses. The fuel gas is itself combustible in a conventional manner.

After the first stage, the gas atoms become elongated during electron removal as the atoms are ionised. Laser or light wave energy of a predetermined frequency is injected into a containment vessel in a gas ionisation process. The light energy absorbed by voltage-stimulated gas nuclei, causes destabilisation of gas ions still further. The absorbed laser energy causes the gas nuclei to increase in energy state, which in turn, causes electron deflection to a higher orbital shell.

The electrically charged and laser-primed combustible gas ions from a gas resonant cavity, may be directed into a an optical thermal lens assembly for triggering. Before entry into the optimal thermal lens, electrons are stripped from the ions and the atom is destabilised. The destabilised gas ions which are electrically and mass unbalanced atoms having highly energised nuclei, are pressurised during spark ignition. The unbalanced, destabilised atomic components interact thermally; the energised and unstable hydrogen gas nuclei collide with highly energised and unstable oxygen gas nuclei, causing and producing thermal explosive energy beyond the gas burning stage. The ambient air gas components in the initial mixture aid the thermal explosive process under a controlled state.

In the process, the point of optimum energy yield is reached when the electron-deficient oxygen atoms (having less than a normal number of electrons) lock on to an capture a hydrogen atom electron, prior to, or during, thermal combustion of the hydrogen/oxygen mixture. Atomic decay results in the release of energy.

After the first stage, the gas mixture is subjected to a pulsating, polar electric field which causes the orbits of the electrons of the gas atoms to become distended . The pulsating electrical field is applied at a frequency which resonates with the electrons of the gas atoms. This results in the energy levels of the electrons increasing in cascading incremental steps.

Next, the gas atoms are ionised and subjected to electromagnetic wave energy of the correct frequency to induce further electron resonance in the ion, whereby the energy level of the electron is successively increased. Electrons are extracted from the resonating ions while they are in this increased energy state, and this destabilises the nuclear electron configuration of the ions. This gas mixture of destabilised ions is thermally ignited.

In the apparatus shown in **Fig.4**, water is introduced at inlet **1** into a first stage water fracturing module **2**, such as the water fuel cell described above, in which water molecules are broken down into hydrogen, oxygen and released gasses which were trapped in the water. These gasses may be introduced to a successive stage **3** or other number of like resonant cavities, which are arranged in either a series or parallel combined array. The successive energisation of the gas atoms, provides a cascading effect, successively increasing the voltage stimulation level of the released gasses as they pass sequentially through cavities **2**, **3**, etc. In a final stage, and injector system **4**, of a configuration of the type shown in **Fig.5A** or **Fig.5B**, receives energised atomic and gas particles where the particles are subjected to further energy input, electrical excitation and thermal stimulation, which produces thermal explosive energy **5**, which may be directed through a lens assembly of the type shown in **Fig.5C** to provide a controlled thermal energy output.

A single cell, or battery of cells such as shown in **Fig.3**, provides a fuel gas source for the stages following the first stage. The fuel gas is activated by electromagnetic waves, and electrically charged gas ions of hydrogen and oxygen (of opposite polarity) are expelled from the cascaded cells **2**, **3**, etc. shown in **Fig.4**. The circuit of **Fig.9** may be utilised as a source of ionising energy for the gasses. The effect of cascading, successively increases the voltage stimulation level of the released gasses, which are then directed to the final injector assembly **4**. In the injector assembly, gas ions are stimulated to an even greater energy level. The gasses are continually exposed to a pulsating laser or other electromagnetic wave energy source together with a high-intensity oscillating voltage field which occurs within the cell between electrodes or conductive plates of opposite electrical polarity. A preferred construction material for the plates is a stainless steel T-304 which is non-chemically reactive with water, hydrogen or oxygen. An electrically conductive material inserted in the fluid environment, is a desirable

material of construction for the electrical field producing plates, through which field, the stream of activated gas particles passes.

Gas ions of opposite electrical charges reach and maintain a critical energy level state. The gas ions have opposite electrical charges and are subjected to oscillating voltage fields of opposite polarity. They are also subjected to a pulsating electromagnetic wave energy source. Immediately after reaching critical energy, the excited gas ions are exposed to a high temperature thermal zone in the injection cell **4**, which causes the excited gas ions to undergo gas combustion. The gas ignition triggers atomic decay and releases thermal energy **5**, with explosive force.

Once triggered, the explosive thermal energy output is controllable by the attenuation of operational parameters. With reference to **Fig.6A**, for example, once the frequency of resonance is identified, by varying applied pulse voltage to the initial water fuel cell assemblies **2**, **3**, the ultimate explosive energy output is likewise varied. By varying the pulse shape and/or amplitude, or pulse train sequence of the electromagnetic wave energy source, final output is varied. Attenuation of the voltage field frequency in the form of OFF and ON pulses, likewise affects the output of the staged apparatus. Each control mechanism can be used separately, grouped in sections, or systematically arranged in a sequential manner.

A complete system in accordance with the present application thus includes:

1. A water fuel cell for providing a first fuel gas mixture consisting of at least a portion of hydrogen and oxygen gas.
2. An electrical circuit of the type shown in **Fig.7** providing a pulsating, polar electric field to the gas mixture as illustrated in **Fig.6A**, whereby electron orbits of the gas atoms are distended by being subjected to electrical polar forces, changing from the state shown conceptually in **Fig.6B** to that of **Fig.6C**, at a frequency such that the pulsating electric field induces a resonance with respect to electrons of the gas atoms. The energy level of the resonant electrons is thereby increased in cascading incremental steps.
3. A further electric field to ionise the gas atoms and
4. An electromagnetic wave energy source for subjecting the ionised gas atoms to wave energy of a predetermined frequency to induce further electron resonance in the ions, whereby the energy level of the electron is successively increased, as shown in **Fig.6D**.
5. An electron sink, which may be in the form of the grid element shown in **Fig.8A**, extracts further electrons from the resonating ions while such ions are in an increased energy state and destabilises the nuclear electron configuration of the ions. The “extraction” of electrons by the sink is co-ordinated with the pulsating electrical field of the resonant cavity produced by the circuit of **Fig.7**, by means of
6. An interconnected synchronisation circuit, such as shown in **Fig.8B**.
7. A nozzle, **10** in **Fig.5B**, or thermal lens assembly, **Fig.5C**, provides the means to direct the destabilised ions, and in which they are finally thermally ignited.

As previously noted, to reach and trigger the ultimate atomic decay of the fuel cell gasses at the final stage, sequential steps are taken. First, water molecules are slit into hydrogen and oxygen gasses by a voltage stimulation process. In the injector assembly, a laser produced coherent light wave is absorbed by the gasses. At this point, as shown in **Fig.6B**, the individual atoms are subjected to an electric field to begin an ionisation process. The laser energy is absorbed and causes gas atoms to lose electrons and form positively charged gas ions. The energised, positively charged hydrogen atoms now accept electrons liberated from the heavier gasses and attract other negatively charged gas ions as conceptually illustrated in **Fig.6C**. Positively and negatively charged gas ions are re-exposed to further pulsating energy sources to maintain random distribution of ionised gas particles.

The gas ions within the wave energy chamber are subjected to an oscillating high-intensity voltage field in a chamber **11** in **Fig.5A** and **Fig.5B** formed within electrodes **12** and **13** in **Fig.5A** and **Fig.5B** of opposite electrical polarity, to produce a resonant cavity. The gas ions reach a critical energy state at the point of resonance.

At this point, within the chamber, additional electrons are attracted to the positive electrode; while positively charged ions or atomic nuclei are attracted to the negative electrode. The positive and negative attraction forces are co-ordinated and act on the gas ions simultaneously; the attraction forces are non-reversible. The gas ions experience atomic component deflection approaching the point of electron separation. At this point electrons are extracted from the chamber by a grid system such as shown in **Fig.5A**. The extracted electrons are consumed and prevented from re-entering the chamber by a circuit such as shown in **Fig.8B**. The elongated gas ions are subjected to a thermal heat zone to cause gas ignition, releasing thermal energy with explosive force. During ionic gas combustion, highly energised and stimulated atoms and atom nuclei collide and explode during thermal excitation. The hydrogen fracturing process occurring, sustains and maintains a thermal zone, at a temperature in excess of normal oxygen/hydrogen combustion temperature, that is, in excess of 2,500 degrees Fahrenheit. To cause and maintain the atomic elongation depicted in **Fig.6C** before gas ignition, a voltage intensifier circuit such

as shown in **Fig.7** is utilised as a current-restricting voltage source to provide the excitation voltage applied to the resonant cavity. At the same time, the interconnected electron extractor circuit shown in **Fig.8B**, prevents the reintroduction of electrons back into the system. depending on calculated design parameters, a predetermined voltage and frequency range may be designed for any particular application or physical configuration of the apparatus.

In the operation of the assembly, the pulse train source for the gas resonant cavity shown at **2** and **3** in **Fig.4** may be derived from a circuit such as shown in Figs. **2**, **7** or **9**, and such cavity circuits may be in sequence to provide a cascading energy input. It is necessary in the final electron extraction, that the frequency with which electrons are removed from the system be sequenced and synchronised with the pulsing of the gas resonant cavity. In the circuit of **Fig.8B**, the co-ordination of synchronisation of the circuit with the circuit of **Fig.7** may be achieved by interconnecting point **"A"** of the gate circuit of **Fig.8B** to point **"A"** of the pulsing circuit of **Fig.7**.

The circuit shown in **Fig.9** enhances the voltage potential across the resonant charging choke coils during pulsing operations and restricts current flow by allowing an external electromagnetic pulsing field **F**, derived from the primary coil **A** being energised to traverse the coil windings **D** and **E** being energised by the incoming pulse train **Ha xxx Hn**, through switching diode **G**. The external pulse field **F**, and the incoming pulse train **Ha xxx Hn**, are sequentially the same, allowing resonant action to occur, restricting current flow while allowing voltage intensity to increase to stimulated the electrical polarisation process, the gas ionisation process and the electron extraction process. The voltage intensifier circuit of **Fig.9** prevents electrons from entering into those processes.

Together, the hydrogen injector assembly **4**, and the resonant cavity **2** and **3**, form a gas injector fuel cell which is compact, low in weight and whose design can be varied. For example, the hydrogen injector system is suited for cars and jet engines. Industrial applications require larger systems. For rocket engine applications, the hydrogen gas injector system is positioned at the top of each resonant cavity arranged in a parallel cluster array. If resonant cavities are sequentially combined in a parallel/series array, the hydrogen injection assembly is positioned after the exits of the resonant cavities have been combined.

From the outline of the physical phenomena associated with the process described in **Table 1**, the theoretical basis of the invention considers the respective states of molecules, gasses and ions derived from liquid water. Before voltage stimulation, water molecules are randomly dispersed throughout water within a container. When a unipolar voltage pulse train such as shown in **Fig.6A (53a xxx 53n)** is applied, an increasing voltage potential is induced in the molecules, gasses and/or ions in a linear, step-like charging effect. The electrical field of the particles within a chamber including the electrical field plates increases from a low-energy state (**A**) to a high-energy state (**J**) in a step manner, following each pulse train as illustrated in **Fig.6A**. The increasing voltage potential is always positive in direct relationship to negative ground potential during each pulse. The voltage polarity on the plates which create the voltage fields, remains constant. Positive and negative voltage "zones" are thus formed simultaneously.

In the first stage of the process described in **Table 1**, because the water molecule naturally exhibits opposite electric fields in a relatively polar configuration (the two hydrogen atoms are positively electrically charged relative to the negatively electrically charged oxygen atom), the voltage pulse causes initially randomly orientated water molecules in the liquid state to spin and orientate themselves with reference to the voltage fields applied.

When the potential difference applied causes the oriented water molecules to align themselves between the conductive plates, pulsing causes the voltage field intensity to be increased in accordance with **Fig.6A**. As further molecular alignment occurs, molecular movement is hindered. Because the positively charged hydrogen atoms are attracted in the opposite direction to the negatively charged oxygen atoms, a polar charge alignment or distribution occurs as shown in **Fig.6B**. As the energy level of the atoms subjected to resonant pulsing increases, the stationary water molecules become elongated as shown in **Fig.6C**. Electrically charged nuceli and electrons are attracted towards opposite voltage zones, disrupting the mass equilibrium of the water molecule.

In the first stage, as the water molecule is further exposed to a potential difference, the electrical force of attraction of the atoms to the chamber electrodes also increases in intensity. As a result, the co-valent bonding between the atoms is weakened and ultimately, terminated. The negatively charged electron is attracted towards the positively charged hydrogen atoms, while at the same time, the negatively charged oxygen atoms repel electrons.

Once the applied resonant energy caused by pulsation of the electrical field in the cavities reaches a threshold level, the disassociated water molecules, now in the form of liberated hydrogen, oxygen and ambient air gasses, begin to ionise and lose or gain electrons during the final stage in the injector assembly. Atom destabilisation occurs and the electrical and mass equilibrium of the atoms is disrupted. Again, the positive field produced within the chamber or cavity that the encompasses the gas stream, attracts negatively charged ions while the positively charged ions are attracted to the negative field. Atom stabilisation does not occur because the pulsing voltage

applied is repetitive without polarity change. A potential of approximately several thousand volts, triggers the ionisation state.

As the ionised particles accumulate within the chamber, the electrical charging effect is again an incremental stepping effect that produces an accumulative increased potential, while, at the same time, resonance occurs. The components of the atom begin to "vibrate" at a resonant frequency such that an atomic instability is created. As shown in **Fig.6D**, a high energy level is achieved, which then collapses, resulting in the release of thermal explosive energy. Particle impact occurs when liberated ions in a gas are subjected to further voltage. A longitudinal cross-section of a gas resonant cavity is shown in **Fig.5A**. To promote gas ionisation, electromagnetic wave energy such as a laser or photon energy source of a predetermined wavelength and pulse intensity is directed to, and absorbed by, the ions of the gas. In the device of **Fig.5A**, semiconductor optical lasers **20a - 20p, 20xxx** surround the gas flow path. In the device of **Fig.5B**, photo energy **20** is injected into a separate absorption chamber **21**. The incremental stimulation of nuclei to a more highly energised state by electromagnetic wave energy causes electron deflection to a higher orbital state. The pulse rate as well as intensity of the electromagnetic wave source is varied to match the absorption rate of ionised particles to produce the stepped incremental increase in energy. A single laser coupled by means of fibre optic light guides is an alternative to the plurality of lasers shown in **Fig.5B**. Continued exposure of the gas ions to different forms of wave energy during voltage stimulation, maintain individual atoms in a destabilised state and prevents atomic stabilisation.

The highly energised gas ions are thermally ignited when they pass from injector **4** and enter into and pass through a nozzle **10** in **Fig.5B**, or an optical thermal lens assembly as shown in **Fig.5C**. In **Fig.5C**, the combustible gas ions are expelled through and beyond a quenching circuit **30**, and reflected by lenses **31** and **32**, back and forth through a thermal heat zone **33**, prior to atomic breakdown and then exiting through a final port **34**. A quenching circuit is a restricted orifice through which the particle stream passes, such that flashback does not occur. The deflection shield or lens **31**, superheats beyond 3000 degrees Fahrenheit and the combustible gas ions passing through the exiting ports are regulated to allow a gas pressure to form inside the thermal zone. The energy yield is controlled by varying the applied voltage or pulse-train since the thermal-lens assembly is self-adjusting to the flow rate of the ionised and primed gasses. The combustible ionic gas mixture is composed of hydrogen, oxygen and ambient air gasses. The hydrogen gas provides the thermal explosive force, the oxygen atoms aid the gas thermal ignition, and the ambient air gasses retard the gas thermal ignition process to a controllable state.

As the combustible gas mixture is exposed to a voltage pulse train, the stepped increasing voltage potential causes the moving gas atoms to become ionised (losing or gaining electrons) and changes the electrical and mass equilibrium of the atoms. Gasses which do not undergo the gas ionisation process may accept the liberated electrons (electron entrapment) when exposed to light or photon stimulation. The electron extractor grid circuit shown in **Fig.8A** and **Fig.8B**, is applied to the assembly of **Fig.5A** or **Fig.5B**, and restricts electron replacement. The extractor grid **56**, is applied adjacent to electric field producing components **44** and **45**, within the resonant cavity. The gas ions incrementally reach a critical state which occurs after a high energy resonant state. At this point, the atoms no longer tolerate the missing electrons, the unbalanced electrical field and the energy stored in the nucleus. Immediate collapse of the system occurs and energy is released as the atoms decay into thermal explosive energy.

The repetitive application of a voltage pulse train (**A** through **J** of **Fig.6A**) incrementally achieves the critical state of the gas ions. As the gas atoms or ions (**1a xxx 1n**) shown in **Fig.6C**, become elongated during electron removal, electromagnetic wave energy of a predetermined frequency and intensity is injected. The wave energy absorbed by the stimulated gas nuclei and electrons, causes further destabilisation of the ionic gas. The absorbed energy from all sources, causes the gas nuclei to increase in energy state and induces the ejection of electrons from the nuclei.

To further stimulate the electron entrapment process beyond the atomic level (capturing the liberated electrons during the hydrogen fracturing process), the electron extractor grid (as shown in **Fig.8A**) is placed in spaced relationship to the gas resonant cavity structure shown in **Fig.5A**. The electron extractor grid is attached to an electrical circuit (such as that shown in **Fig.8B**) which allows electrons to flow to an electrical load **55**, when a positive electrical potential is placed on the opposite side of the electrical load. The electrical load may be a typical power-consuming device such as a light bulb or resistive heat-producing device. As the positive electrical potential is switched on, or pulse-applied, the negatively charged electrons liberated in the gas resonant cavity, are drawn away and enter into the resistive load where they are released as heat or light energy. The consuming electrical circuit may be connected directly to the gas resonant cavity positive electrical voltage zone. The incoming positive wave form applied to the resonant cavity voltage zone through a blocking diode, is synchronised with the pulse train applied to the gas resonant cavity by the circuit of **Fig.7** via an alternate gate circuit. As one pulse train is gated "ON", the other pulse train is switched "OFF". A blocking diode directs the electron flow to the electrical load, while resistive wire prevents voltage leakage during the pulse train "ON" time.

The electron extraction process is maintained during gas-flow change by varying the trigger pulse rate in relationship to the applied voltage. The electron extraction process also prevents spark-ignition of the combustible gasses travelling through the gas resonant cavity because electron build-up and potential sparking is prevented.

In an optical thermal lens assembly or thrust-nozzle, such as shown in **Fig.5C**, destabilised gas ions (electrically and mass unbalanced gas atoms having highly energised nuclei) can be pressurised during spark ignition. During thermal interaction, the highly energised and unstable hydrogen gas nuclei collide with the highly energised and unstable oxygen gas nuclei and produce thermal explosive energy beyond the gas-burning stage. Other ambient air gasses and ions not otherwise consumed, limit the thermal explosive process.

STANLEY MEYER

Canadian Patent 2,067,735 16th May 1991 Inventor: Stanley Meyer

WATER FUEL INJECTION SYSTEM

ABSTRACT

An injector system comprising an improved method and apparatus useful in the production of a hydrogen containing fuel gas from water in a process in which the dielectric property of water and/or a mixture of water and other components determines a resonant condition that produces a breakdown of the atomic bonding of atoms in the water molecule. The injector delivers a mixture of water mist, ionised gases and non-combustible gas to a zone within which the breakdown process leading to the release of elemental hydrogen from the water molecules occurs.

DESCRIPTION

This invention relates to a method and apparatus useful in producing thermal combustive energy from the hydrogen component of water.

In my patent no. 4,936,961 “Method for the Production of a Fuel Gas”, I describe a water fuel cell which produces a gas energy source by a method which utilises water as a dielectric component of a resonant electrical circuit.

In my patent no. 4,826,581 “Controlled Process for the Production of Thermal Energy From Gasses and Apparatus Useful Therefore”, I describe a method and apparatus for obtaining the enhanced release of thermal energy from a gas mixture including hydrogen and oxygen in which the gas is subjected to various electrical, ionising and electromagnetic fields.

In my co-pending application serial no. 07/460,859 “Process and Apparatus for the Production of Fuel Gas and the Enhanced Release of Thermal Energy from Fuel Gas”, I describe various means and methods for obtaining the release of thermal/combustive energy from the hydrogen (H) component of a fuel gas obtained from the disassociation of a water (H_2O) molecule by a process which utilises the dielectric properties of water in a resonant circuit; and in that application I more thoroughly describe the physical dynamics and chemical aspects of the water-to-fuel conversion process.

The invention of this present application represents generational improvement in methods and apparatus useful in the utilisation of water as a fuel source. In brief, the present invention is a microminiaturised water fuel cell which permits the direct injection of water, and its simultaneous transformation into a hydrogen-containing fuel, in a combustion zone, such as a cylinder in an internal combustion engine, a jet engine or a furnace. Alternatively, the injection system of the present invention may be utilised in any non-engine application in which a concentrated flame or heat source is desired, for example: welding.

The present injection system eliminates the need for an enclosed gas pressure vessel in a hydrogen fuel system and thereby reduces a potential physical hazard heretofore associated with the use of hydrogen-based fuels. The system produces fuel-on-demand in real-time operation and sets up an integrated environment of optimum parameters so that a water-to-fuel conversion process works at high efficiency.

The preferred embodiment of the invention is more fully explained below with reference to the drawings in which:

Fig.1 figuratively illustrates the sections and operating zones included in a single injector of the invention.

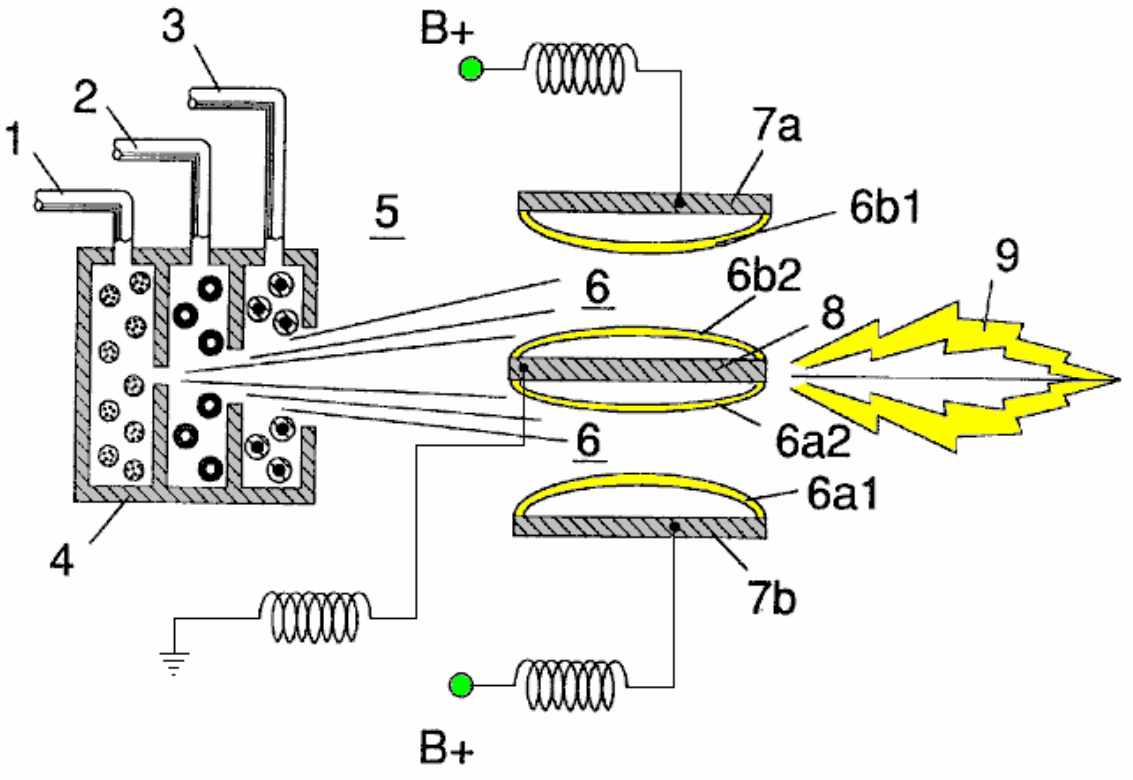


Fig.2A is a side cross-sectional view.

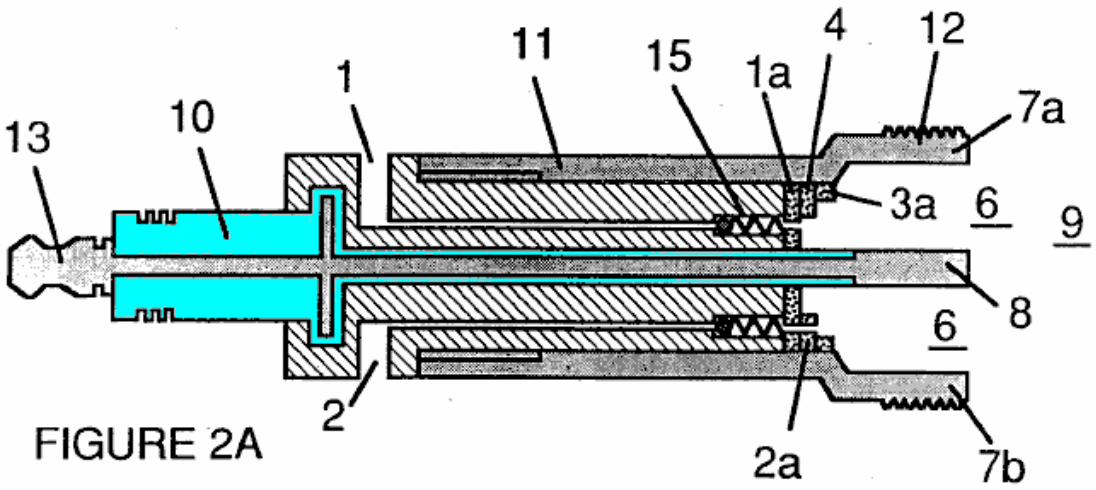


FIGURE 2A

Fig.2B is a frontal view from the operative end.

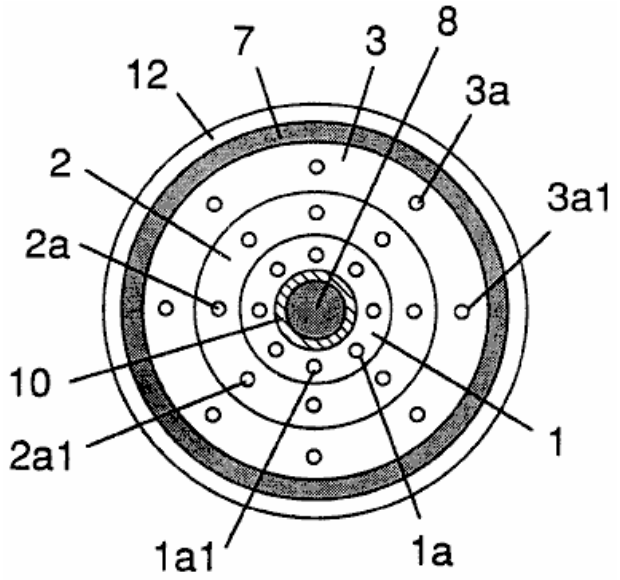


FIGURE 2B

Fig.2C is an exploded view of an individual injector.

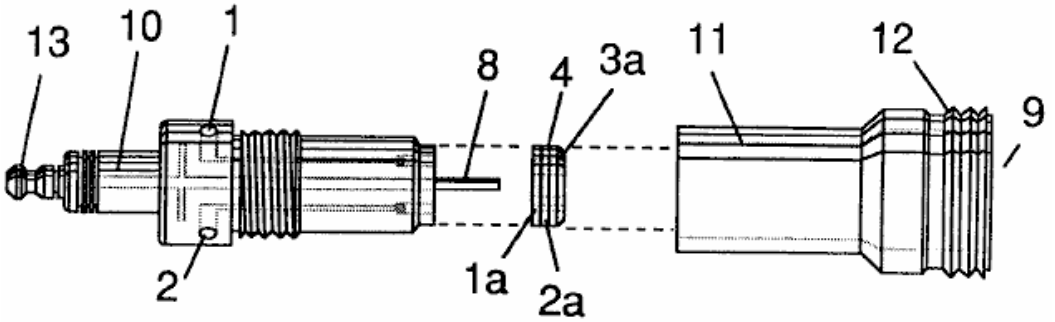


FIGURE 2C

Fig.3 and **Fig.3A** show the side and frontal cross-sectional views of an alternatively configured injector.

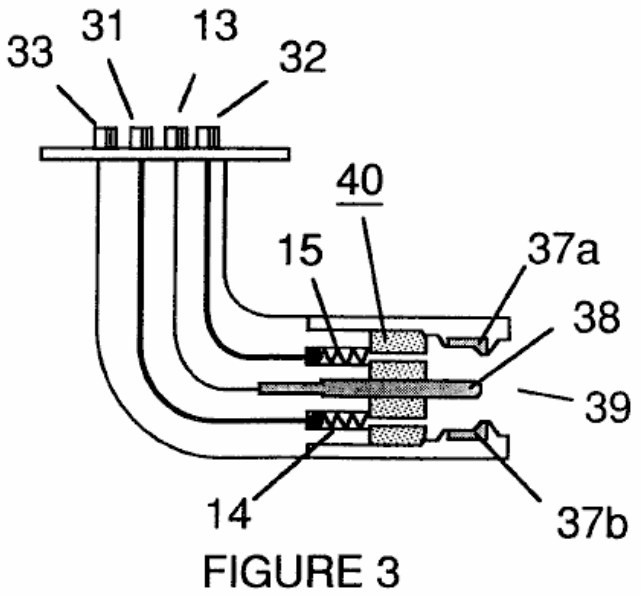


FIGURE 3

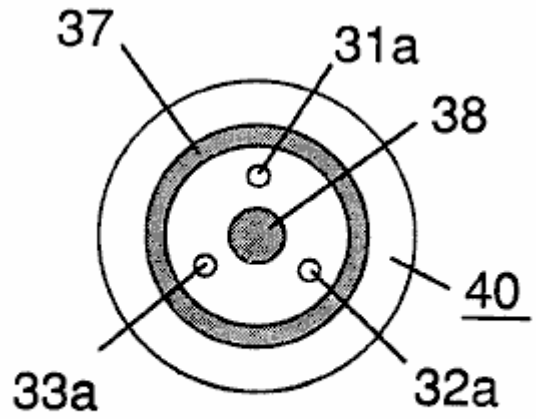


FIGURE 3A

Fig.4 shows a disk array of injectors.

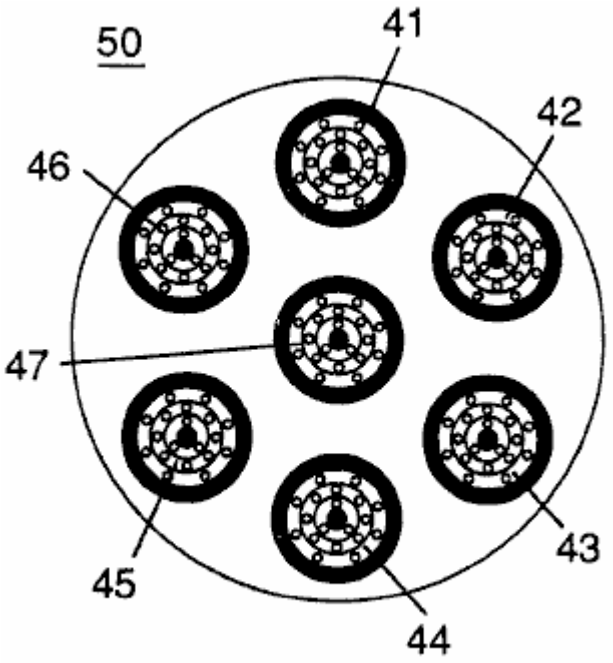


FIGURE 4

Fig.5 shows the resonance electrical circuit including the injector.

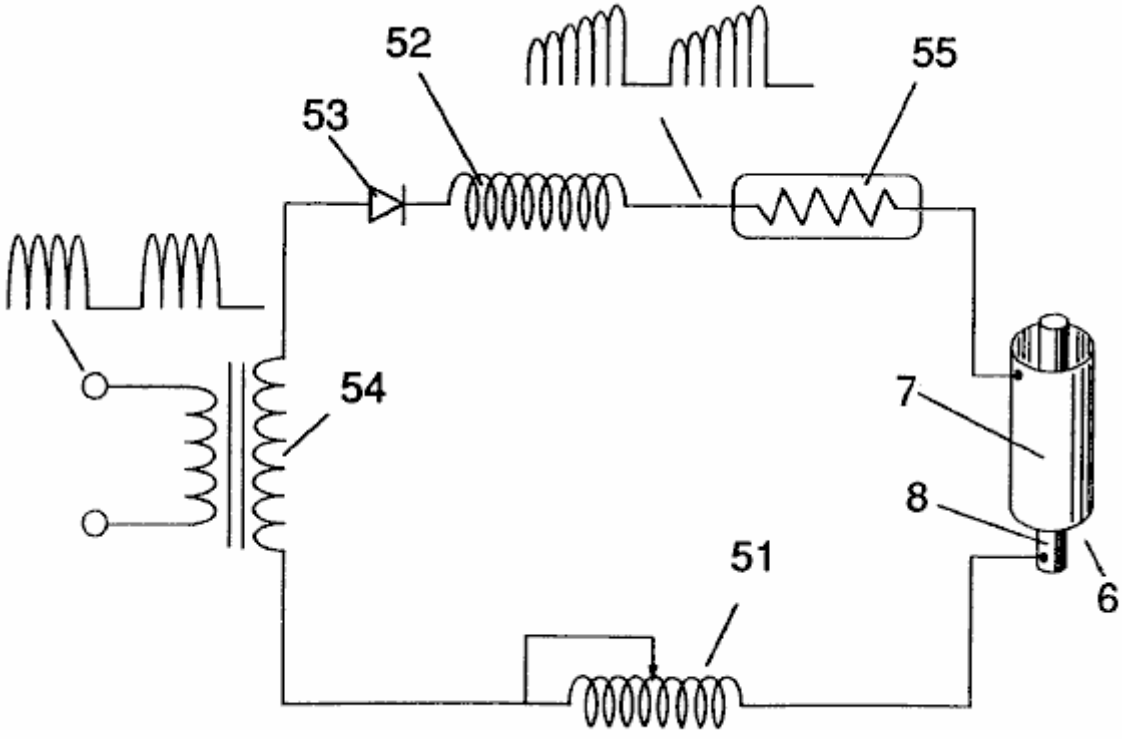


FIGURE 5

Fig.6 depicts the inter-relationship of the electrical and fuel distribution components of an injector system.

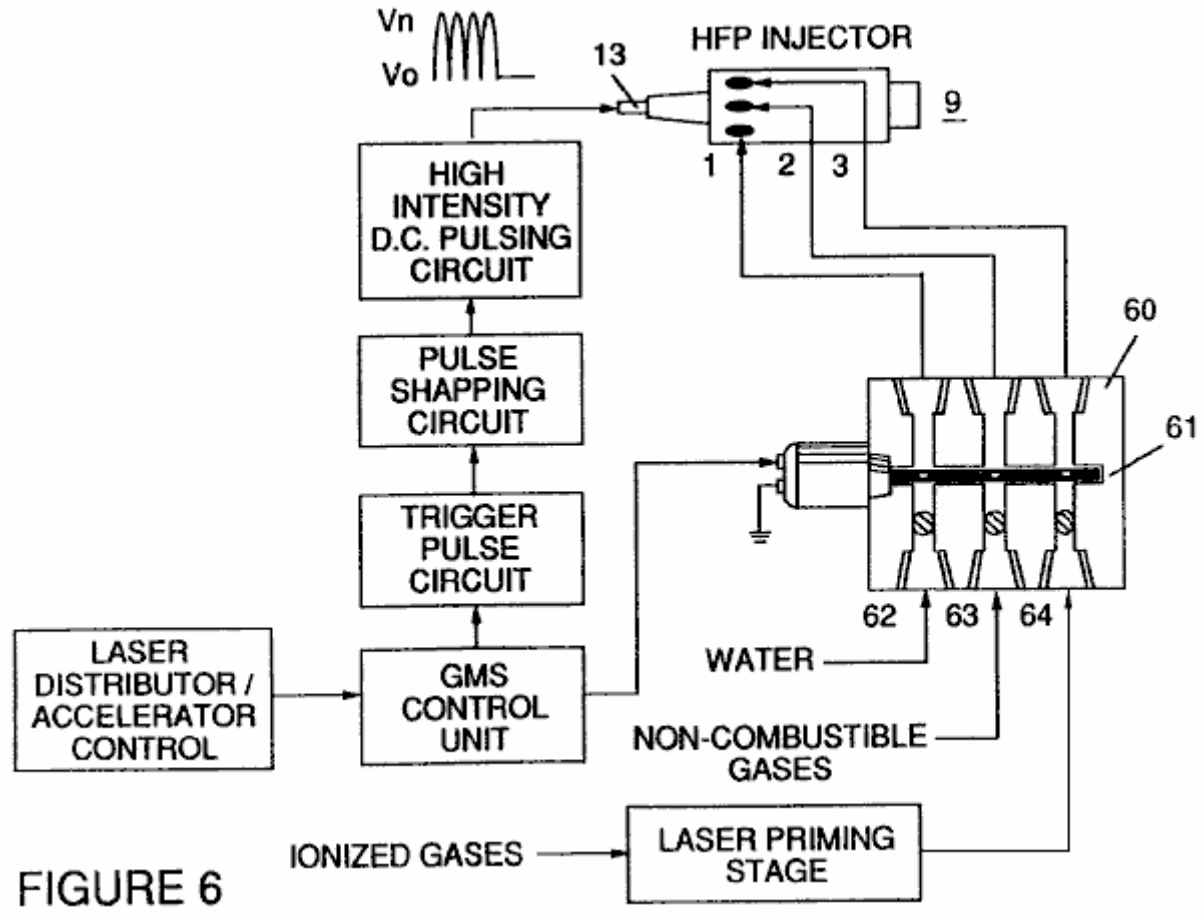


FIGURE 6

Although I refer to an "injector" in this document, the invention relates not only to the physical configuration of an injector apparatus, but also to the overall process and system parameters determined in the apparatus to achieve the release of thermal energy. In a basic outline, an injector regulates the introduction of process constituents into a combustion zone and sets up a fuel mixture condition permitting combustion. That combustion condition is triggered simultaneously with injector operation in real-time correspondence with control parameters for the process constituents.

In the fuel mixture condition which is created by the injector, water (H_20) is atomised into a fine spray and mixed with **1** ionised ambient air gasses and **2** other non-combustible gasses such as nitrogen, argon and other rare gasses, and water vapour. (Exhaust gas produced by the combustion of hydrogen with oxygen is a non-combustible water vapour. This water vapour and other inert gasses resulting from combustion may be recycled from an exhaust outlet in the injector system, back into the input mixture of non-combustible gasses.) The fuel mix is introduced at a consistent flow rate maintained under a predetermined pressure. In the triggering of the condition created by the injector, the conversion process described in my patent no. 4,936,961 and co-pending application serial no. 07/460,859 is set off spontaneously on a "micro" level in a predetermined reaction zone. The injector creates a mixture, under pressure in a defined zone of water, ionised gasses and non-combustible gasses. Pressure is an important factor in the maintenance of the reaction condition and causes the water/gas mixture to become intimately mixed, compressed and destabilised to produce combustion when activated under resonance conditions of ignition. In accordance with the earlier mentioned conversion process of my patent and application, when water is subjected to a resonance condition water molecules expand and distend; electrons are ejected from the water molecule and absorbed by ionised gasses and the water molecule, thus destabilised, breaks down into its elemental components of hydrogen (H_2) and oxygen (O) in the combustion zone. The hydrogen atoms released from the molecule provide the fuel source in the mixture for combustion with oxygen. The present invention is an application of that process and is outlined in Table 1:

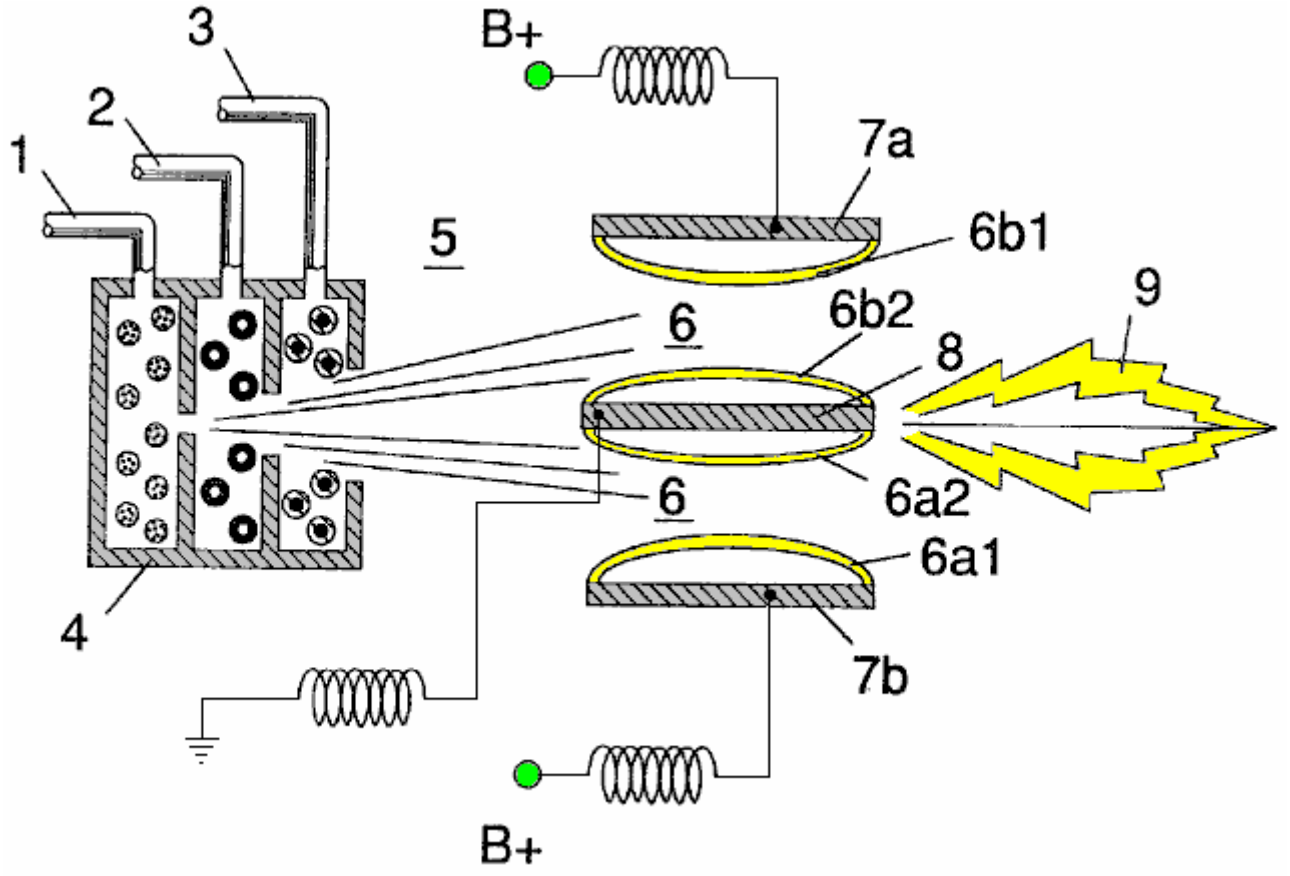
Table 1

Injector Mixture	**+**	**Process conditions**	**=**	**Thermal Energy**
(1) Water Mist		(1) Release Under pressure into Combustion Zone		(1) Heat
and		and		or
(2) Ionised Gas		(2) Resonance utilising the dielectric property of water as a capacitor		(2) Internal Combustion Engine (Explosive force)
and		and		or
(3) Non-combustible Gas		(3) Unipolar pulsing at high voltage		(3) Jet Engine
				or
				(4) Other application

The process occurs as water mist and gasses under pressure are injected into, and intimately mixed in the combustion zone and an electrically polarised zone. In the electrically polarised zone, the water mixture is subjected to a unipolar pulsed direct current voltage which is tuned to achieve resonance in accordance with the electrical, mass and other characteristics of the mixture as a dielectric in the environment of the combustion zone. The resonant frequency will vary according to the injector configuration and depends upon the physical characteristics, such as the mass and volume of the water and gasses in the zone. As my prior patents and application point out, the resonant condition in the capacitative circuit is determined by the dielectric properties of water: (1) as the dielectric in a capacitor formed by adjacent conductive surfaces, and (2) as the water molecule itself is a polar dielectric material. At resonance, current flow in the resonant electrical circuit will be minimised and voltage will peak.

The injector system provides a pressurised fuel mixture for subjection to the resonant environment of the voltage combustion zone as the mixture is injected into the zone. In a preferred embodiment, the injector includes concentrically nested serial orifices, one for each of the three constituent elements of the fuel mixture. (It may be feasible to combine and process non-combustible and ionised gasses in advance of the injector. In this event, only two orifices are required, one for the water and the other for the combined gasses.) The orifices disperse the water mist and gasses under pressure into a conically shaped activation and combustion zone.

Fig1A shows a transverse cross-section of an injector, in which, supply lines for water **1**, ionised gas **2**, and non-combustible gas **3**, feed into a distribution disk assembly **4** which has concentrically nested orifices. The fuel mixture passes through a mixing zone **5**, and a voltage zone **6**, created by electrodes **7a** and **7b** (positive) and **8** (negative or ground). Electrical field lines are shown as **6a1** and **6a2** and **6b1** and **6b2**. Combustion (i.e. the oxidation of hydrogen) occurs in the zone **9**. Ignition of the hydrogen can be primed by a spark or may occur spontaneously as a result of the exceptionally high volatility of hydrogen and its presence in a high-voltage field.



Although the mixing zone, the voltage zone and the combustion zone are mentioned separately in this explanation, they are not in fact physically separated, as can be seen from **Fig.1**. In the zone(s), there is produced an “excited” mixture of vaporised water mist, ionised gasses and other non-combustible gasses, all of which have been instantaneously released from under high pressure. Simultaneously, the released mixture in the zone, is exposed to a pulsed voltage at a frequency corresponding to electrical resonance. Under these conditions, outer-shell electrons of atoms in the water molecule are de-stabilised and molecular time-share is interrupted. Thus, the gas mixture in the injector zone is subjected to physical, electrical and chemical interactive forces which cause a breakdown of the atomic bonding forces of the water molecule.

Process parameters are determined, based on the size of a particular injector. In an injector sized appropriately for use to provide a fuel mixture to a conventional cylinder in a passenger vehicle car engine, the injector may resemble a conventional spark plug. In such an injector, the water orifice is 0.1 to 0.15 inch in diameter; the ionised gas orifice is 0.15 to 0.2 inch in diameter, and the non-combustible gas orifice is 0.2 to 0.25 inch in diameter. In such a configuration, the serial orifices increase in size from the innermost orifice, as appropriate in a concentric configuration. As noted above, it is desirable to maintain the introduction of the fuel components at a constant rate. Maintaining a back-pressure of about 125 pounds per square inch for each of the three fuel gas constituents appears to be satisfactory for a “spark-plug” injector. In the pressurised environment of the injector, spring-loaded one-way check valves in each supply line, such as **14** and **15**, maintain pressure during pulse off times.

Voltage zone **6** surrounds the pressurised fuel mixture and provides an electrically charged environment of pulsed direct current in the range from about 500 to 20,000 volts and more, at a frequency tuned into the resonant characteristic of the mixture. this frequency will typically lie within the range from about 20 KHz to about 50 KHz, dependent, as noted above, on the mass flow of the mixture from the injector and the dielectric property of the mixture. In a spark-plug sized injector, the voltage zone will typically extend longitudinally about 0.25 to 1.0 inch to permit sufficient dwell time of the water mist and gas mixture between the conductive surfaces **7** and **8** which form a capacitor so that resonance occurs at a high-voltage pulsed frequency, and combustion is triggered. In the zone, an energy wave which is related to the resonant pulse frequency, is formed. The wave continues to pulse through the flame in the combustion zone. The thermal energy produced is released as heat energy. In a confined zone such as a piston/cylinder engine, gas detonation under resonant conditions, produces explosive physical power.

In the voltage zone, the time-share ratio of the hydrogen and oxygen atoms comprising the individual water molecules in the water mist, is upset in accordance with the process explained in my patent no. 4,936,961 and application serial no. 07/460,859. Namely, the water molecule, which is itself a polar structure, is distended or

distorted in shape by being subjected to the polar electric field in the voltage zone. The resonant condition induced in the molecule by the unipolar pulses, upsets the molecular bonding of shell electrons such that the water molecule, at resonance, breaks apart into its constituent atoms. In the voltage zone, the water molecules are excited into an ionised state, and the pre-ionised gas component of the fuel mixture, captures the electrons released from the water molecule. In this manner, at the resonant condition, the water molecule is destabilised and the constituent atomic elements of the molecule 2H and O, are released and the released hydrogen atoms are available for combustion. the non-combustible gasses in the fuel mixture, reduce the burn rate of hydrogen to that of a hydrocarbon fuel such as gasoline (petrol) or kerosene (paraffin), from its normal burn rate which is about 2.5 times that of gasoline. Hence the presence of non-combustible gasses in the fuel mixture, moderates the energy release and the rate at which the free hydrogen and oxygen molecules combine in the combustion process.

The combustion process does not occur spontaneously so the conditions in the zone must be fine-tuned carefully to achieve an optimum input flow rate for water and the gasses corresponding to the maintenance of a resonant condition. The input water mist and gasses may likewise be injected into the zone in a physically pulsed (on/off) manner corresponding to the resonance achieved. In an internal combustion engine, the resonance of the electrical circuit and the physical pulsing of the input mixture may be required to be related to the combustion cycle of the reciprocating engine. In this regard, one or two conventional spark plugs may require a spark cycle tuned in correspondence to the conversion cycle resonance, so that combustion of the mixture will occur. Thus, the input flow, conversion rate and combustion rate are interrelated and optimally, each should be tuned in accordance with the circuit resonance at which conversion occurs.

The injection system of the present invention is suited to retrofit applications in conventionally fuelled gasoline and diesel internal combustion engines and conventionally fuelled jet aircraft engines.

Example 1

Figs 2A, 2B and **2C** illustrate a type of injector useful, among other things, as a fuel source for a conventional internal combustion engine. In the cross-section of **Fig.2A**, reference numerals corresponding to the identifying numerals used in **Fig.1** show a supply line for water **1**, leading to first distribution disc **1a** and supply line for ionised gas **2**, leading to second distribution disc **2a**. In the cross-section, the supply line for non-combustible gas **3** leading to distribution disc **3a**, is not illustrated, however, its location as a third line should be self evident. The three discs comprise distribution disc assembly 4. The supply lines are formed in an electrically insulating body **10**, surrounded by electrically conductive sheath/housing **11** having a threaded end segment **12**.

A central electrode **8**, extends the length of the injector. Conductive elements **7a** and **7b** (**7a** and **7b** depict opposite sides of the diameter in the cross-section of a circular body), adjacent threaded section **12** and electrode **8**, form the electrical polarisation zone **6** adjacent to combustion zone **9**. An electrical connector **13** may be provided at the other end of the injector. (In this document, the term “electrode” refers to the conductive surface of an element forming one side of a capacitor.) In the frontal view of **Fig.2B**, it is seen that each disc making up the distribution disc assembly **9**, includes a plurality of micro-nozzles **1a1, 2a1, 3a1**, etc. for the injection of the water and gasses into the polarisation/voltage and combustion zones. The exploded view of **Fig.2C** shows another view of the injector and additionally depicts two supply line inlets **1** and **2**, the third not being shown because of the inability of representing the uniform 120^0 separation of three lines in a two-dimensional drawing.

In the injector, water mist (forming droplets in the range, for example, of from 10 to 250 microns and above, with size being related to voltage intensity) is injected into the fuel-mixing and polarising zone by way of water spray nozzles **1a1**. The tendency of water to form a “bead” or droplet is a parameter related to droplet mist size and voltage intensity. ionised air gasses and non-combustible gasses, introduced through nozzles **2a1** and **3a1**, are intermixed with the expelling water mist to form a fuel-mixture which enters into voltage zone **6** where the mixture is exposed to a pulsating, unipolar, high-intensity voltage field (typically 20,000 volts at 50 Hz or above, at the resonant condition in which current flow in the circuit (amps) is reduced to a minimum) created between electrodes **7** and **8**.

Laser energy prevents discharge of the ionised gasses and provides additional energy input into the molecular destabilisation process which occurs at resonance. It is preferable that the ionised gasses be subjected to laser (photonic energy) activation prior to their introduction into the zone(s); although, for example, a fibre optic conduit may be useful to channel photonic energy directly into the zone. However, heat generated in the zone may affect the operability of such an alternate configuration. The electrical polarisation of the water molecule and a resonant condition occurs to destabilise the molecular bonding of the hydrogen and oxygen atoms. Combustion energy is then released by spark ignition.

To ensure proper flame projection and subsequent flame stability, pumps for the ambient air, non-combustible gas and water, introduce these components to the injector under static pressure up to and beyond 125 pounds per square inch.

Flame temperature is regulated by controlling the volume flow-rate of each fluid-media in direct relationship to applied voltage intensity. To elevate flame temperature, fluid displacement is increased while the volume flow rate of non-combustible gasses is maintained or reduced and the applied voltage amplitude is increased. To lower flame temperature, the fluid flow rate of non-combustible gasses is increased and pulse voltage amplitude is lowered. To establish a predetermined flame temperature, the fluid media and applied voltage are adjusted independently. The flame-pattern is further maintained as the ignited, compressed, and moving gasses are projected under pressure from the nozzle ports in distribution disc assembly **4** and the gas expands in the zone and is ignited.

In the voltage zone, several functions occur simultaneously to initiate and trigger thermal energy yield. Water mist droplets are exposed to high intensity pulsating voltage fields in accordance with an electrical polarisation process which separates the atoms of the water molecule and causes the atoms to experience electron ejection. The polar nature of the water molecule which facilitates the formation of minute droplets in the mist, appears to cause a relationship between the droplet size and the voltage required to effect the process, i.e. the greater the droplet size, the higher the voltage required. The liberated atoms of the water molecule interact with laser-primed ionised ambient air gasses to cause a highly energised and destabilised mass of combustible gas atoms to ignite thermally. Incoming ambient air gasses are laser primed and ionised when passing through a gas processor, and an electron extraction circuit (**Fig.5**) captures and consumes in sink **55**, ejected electrons, and prevents electron flow into the resonant circuit.

In terms of performance, reliability and safety, ionised air gasses and water fuel liquid do not become volatile until the fuel mixture reaches the voltage and combustion zones. Injected non-combustible gasses retard and control the combustion rate of hydrogen during gas ignition.

In alternate applications, laser-primed ionised liquid oxygen and laser-primed liquid hydrogen stored in separate fuel tanks, can be used in place of the fuel mixture, or liquefied ambient air gasses alone with water can be substituted as a fuel source.

The injector assembly is design variable and is retro-fittable to fossil fuel injector ports conventionally used in jet/rocket engines, grain dryers, blast furnaces, heating systems, internal combustion engines and the like.

Example 2

A flange-mounted injector is shown in cross-section in **Fig.3** which shows the fuel mixture inlets and illustrates an alternative three-nozzle configuration leading to the polarisation (voltage) and combustion zones in which one nozzle **31a**, **32a** and **33a** is provided for each of the three gas mixtures, and connected to supply lines **31** and **32** (**33** is not shown). Electrical polarisation zone **36** is formed between electrode **38** and surrounding conductive shell **37**. The capacitative element of the resonant circuit is formed when the fuel mixture, acting as a dielectric, is introduced between the conductive surfaces of **37** and **38**. **Fig.3A** is a frontal view of the operative end of the injector.

Example 3

Multiple injectors may be arranged in a gang as shown in **Fig.4** in which injectors **40, 41, 42, 43, 44, 45, 46, 47, 48** and **49** are arranged concentrically in an assembly **50**. Such a ganged array is useful in applications having intensive energy requirements such as jet aircraft engines and blast furnaces.

Example 4

The basic electrical system utilised in the invention is depicted in **Fig.5** showing the electrical polarisation zone **6** which receives and processes the water and gas mixture as a capacitive circuit element in a resonant charging circuit formed by inductors **51** and **52** connected in series with diode **53**, pulsed voltage source **54**, electron sink **55** and zone **6** formed from conductive elements **7** and **8**. In this manner, electrodes **7** and **8** in the injector, form a capacitor which has electrical characteristics dependent on the dielectric media (e.g. the water mist, ionised gasses and non-combustible gasses) introduced between the conductive elements. Within the macro-dielectric media, however, the water molecules themselves, because of their polar nature, can be considered micro-capacitors.

Example 5

Fuel distribution and management systems useful with the injector of this application are described in my co-pending applications for patent; PCT/US90/6513 and PCT/US90/6407.

A distribution block for the assembly is shown in **Fig.6**. In **Fig.6** the distribution block pulses and synchronises the input of the fuel components in sequence with the electrical pulsing circuit. The fuel components are injected into the injector ports in synchronisation with the resonant frequency, to enhance the energy wave pulse extending from the voltage zone through the flame. In the configuration of **Fig.6**, the electrical system is interrelated to distribution block **60**, gate valve **61** and separate passageways **62, 63** and **64** for fuel components. The distributor produces a trigger pulse which activates a pulse-shaping circuit that forms a pulse having a width and amplitude determined by resonance of the mixture and establishes a dwell time for the mixture in the zone to produce combustion..

As in my referenced application regarding control and management and distribution systems for a hydrogen-containing fuel gas produced from water, the production of hydrogen gas is related to pulse frequency on/off time. In the system shown in **Fig.6**, the distributor block pulses the fluid media introduced to the injector in relationship to the resonant pulse frequency of the circuit and to the operational on/off gate pulse frequency. In this manner, the rate of water conversion (i.e. the rate of fuel produced by the injector) can be regulated and the pattern of resonance in the flame controlled.

STANLEY MEYER

Patent WO 92/07861 2nd November 1990 Inventor: Stanley A. Meyer

CONTROL AND DRIVER CIRCUITS FOR A HYDROGEN GAS FUEL PRODUCING CELL

The major difficulty in using Stan's low-current Water Fuel Cell (recently reproduced by Dave Lawton and shown in Chapter 10) is the issue of keeping the cell continuously at the resonant frequency point. This patent application shows the Stan's circuitry for doing exactly that, and consequently, it is of major importance.

ABSTRACT

A control circuit for a capacitive resonant cavity water capacitor cell (**7**) for the production of a hydrogen containing fuel has a resonant scanning circuit co-operating with a resonance detector and PLL circuit to produce pulses. The pulses are fed into the primary transformer (**TX1**). The secondary transformer (**TX2**) is connected to the resonant cavity water capacitor cell (**7**) via a diode and resonant charging chokes (**TX4**, **TX5**).

This invention relates to electrical circuit systems useful in the operation of a Water Fuel Cell including a water capacitor/resonant cavity for the production of a hydrogen containing fuel gas, such as that described in my United States Letter Patent No. 4,936,961 "Method for the Production of a Fuel Gas" issued on 26th June 1990.

In my Letters Patent for a "Method for the Production of a Fuel Gas", voltage pulses applied to the plates of a water capacitor tune into the dielectric properties of the water and attenuate the electrical forces between the hydrogen and oxygen atoms of the molecule. The attenuation of the electrical forces results in a change in the molecular electrical field and the covalent atomic bonding forces of the hydrogen and oxygen atoms. When resonance is achieved, the atomic bond of the molecule is broken, and the atoms of the molecule disassociate. At resonance, the current (amp) draw from a power source to the water capacitor is minimised and voltage across the water capacitor increases. Electron flow is not permitted (except at the minimum, corresponding to leakage resulting from the residual conductive properties of water). For the process to continue, however, a resonant condition must be maintained.

Because of the electrical polarity of the water molecule, the fields produced in the water capacitor respectively attract and repel the opposite and like charges in the molecule, and the forces eventually achieved at resonance are such that the strength of the covalent bonding force in the water molecule (which are normally in an electron-sharing mode) disassociate. Upon disassociation, the formerly shared bonding electrons migrate to the hydrogen nuclei, and both the hydrogen and oxygen revert to net zero electrical charge. The atoms are released from the water as a gas mixture.

In the invention herein, a control circuit for a resonant cavity water capacitor cell utilised for the production of a hydrogen-containing fuel gas is provided.

The circuit includes an isolation means such as a transformer having a ferromagnetic, ceramic or other electromagnetic material core and having one side of a secondary coil connected in series with a high speed switching diode to one plate of the water capacitor of the resonant cavity and the other side of the secondary coil connected to the other plate of the water capacitor to form a closed loop electronic circuit utilising the dielectric properties of water as part of the electronic resonant circuit. The primary coil of the isolation transformer is connected to a pulse generation means. The secondary coil of the transformer may include segments which form resonant charging choke circuits in series with the water capacitor plates.

In the pulse generation means, an adjustable resonant frequency generator and a gated pulse frequency generator are provided. A gate pulse controls the number of the pulses produced by the resonant frequency generator sent to the primary coil during a period determined by the gate frequency of the second pulse generator.

The invention also includes a means for sensing the occurrence of a resonant condition in the water capacitor / resonant cavity, which when a ferromagnetic or electromagnetic core is used, may be a pickup coil on the transformer core. The sensing means is interconnected to a scanning circuit and a phase lock loop circuit, whereby the pulsing frequency to the primary coil of the transformer is maintained at a sensed frequency corresponding to a resonant condition in the water capacitor.

Control means are provided in the circuit for adjusting the amplitude of a pulsing cycle sent to the primary coil and for maintaining the frequency of the pulsing cycle at a constant frequency regardless of pulse amplitude. In

addition, the gated pulse frequency generator may be connected to a sensor which monitors the rate of gas production in the cell and controls the number of pulses from the resonant frequency generator sent to the cell in a gated frequency in correspondence with the rate of gas production. The sensor may be a gas pressure sensor in an enclosed water capacitor resonant cavity which also includes a gas outlet. The gas pressure sensor is connected to the circuit to determine the rate of gas production with respect to ambient gas pressure in the water capacitor enclosure.

Thus, a comprehensive control circuit and it's individual components for maintaining and controlling the resonance and other aspects of the release of gas from a resonant cavity water cell is described here and illustrated in the drawings which depict the following:

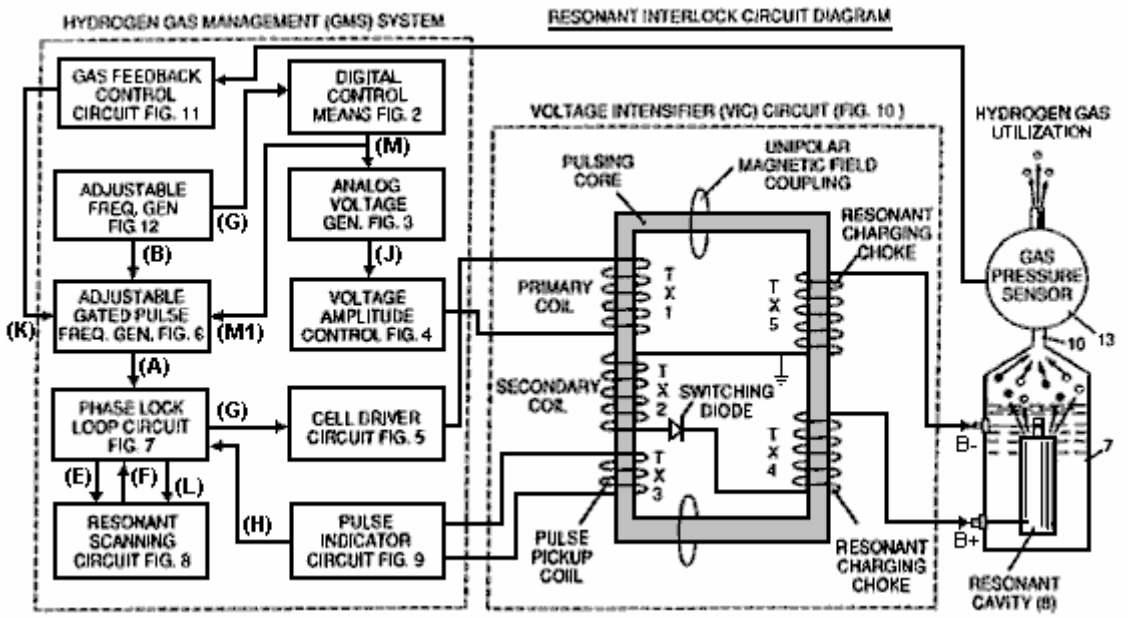


Fig.1 is a block diagram of an overall control circuit showing the interrelationship of sub-circuits, the pulsing core / resonant circuit and the water capacitor resonant cavity.

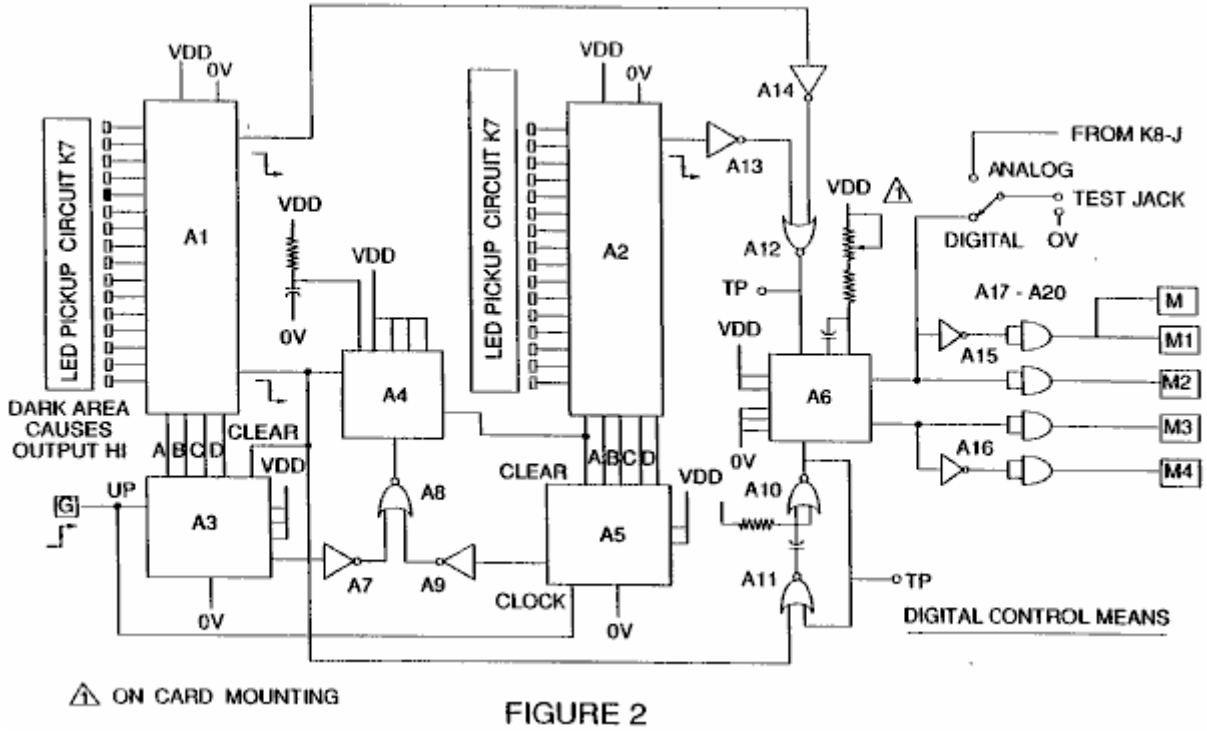


FIGURE 2

Fig.2 shows a type of digital control circuit for regulating the ultimate rate of gas production as determined by an external input. (Such a control circuit would correspond, for example, to the accelerator in a car, or the thermostat control in a building).

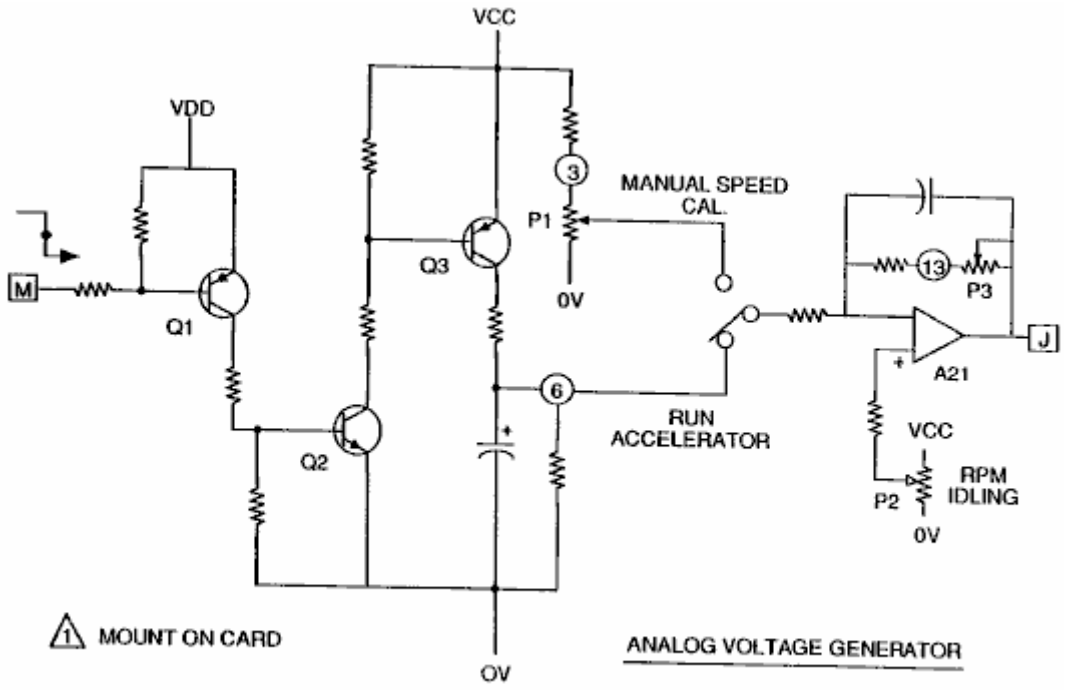


FIGURE 3

Fig.3 shows an analog voltage generator.

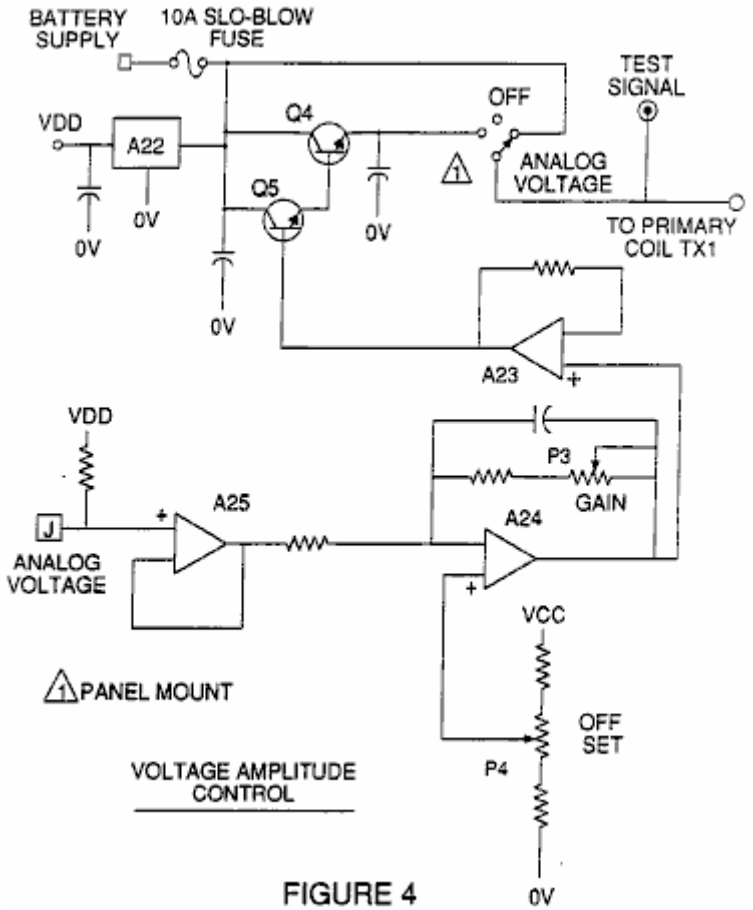


FIGURE 4

Fig.4 is a voltage amplitude control circuit interconnected with the voltage generator and one side of the primary coil of the pulsing core.

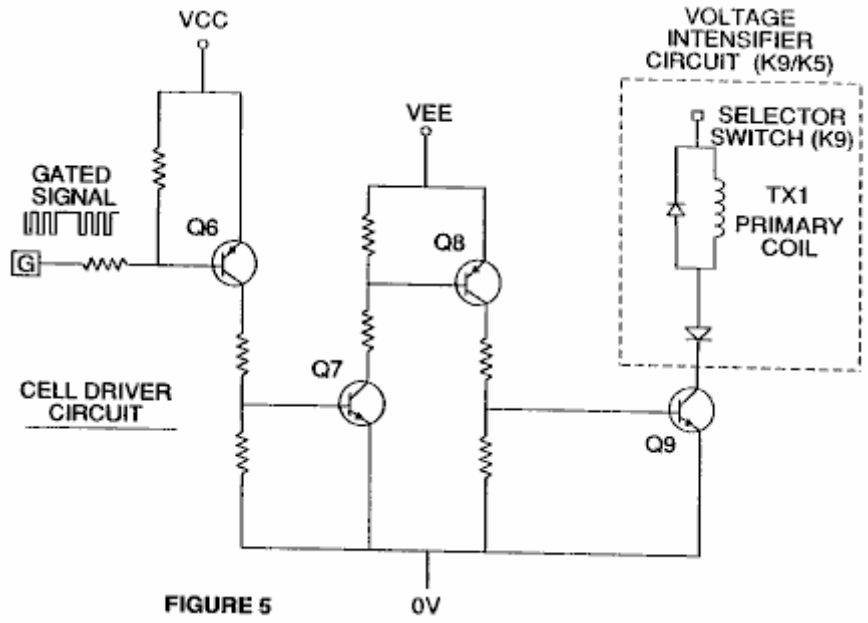


Fig.5 is the cell driver circuit that is connected with the opposite side of the primary coil of the pulsing core. **Figures 6 to 9** form the pulsing control circuitry:

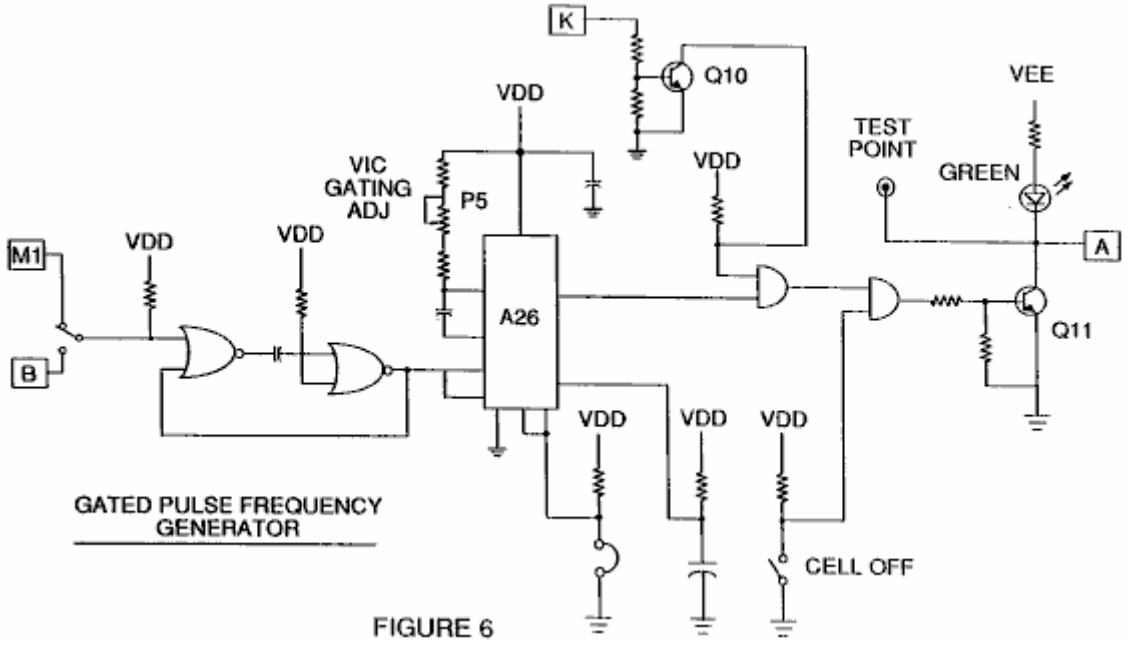


Fig.6 is a gated pulse frequency generator.

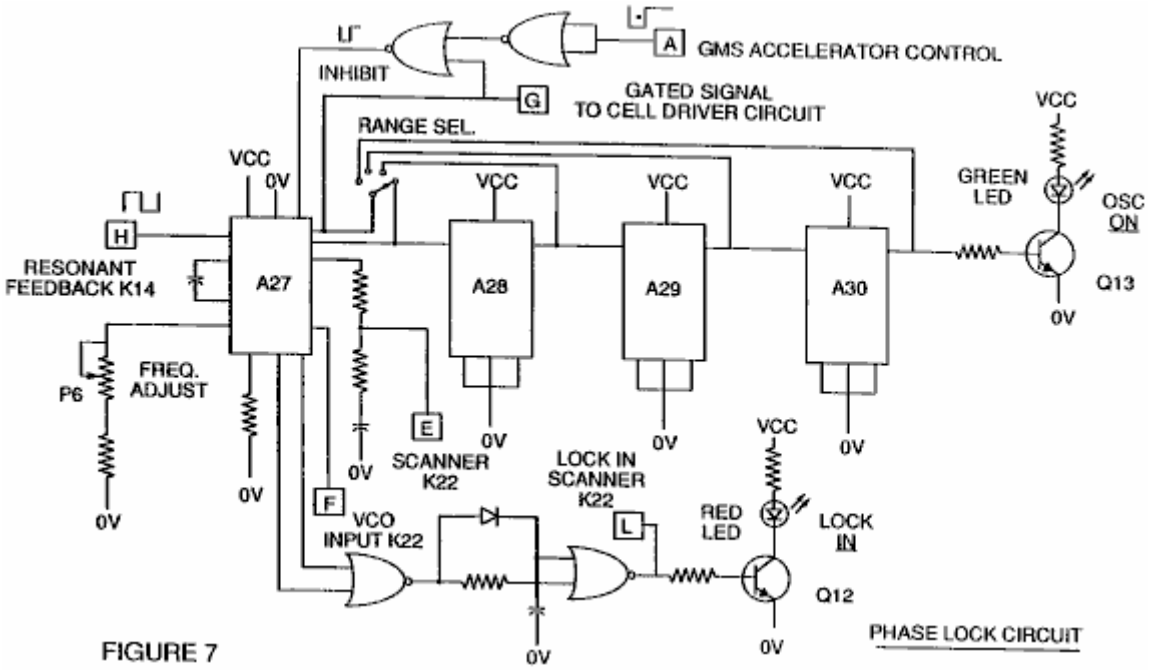


FIGURE 7

Fig.7 is a phase lock circuit.

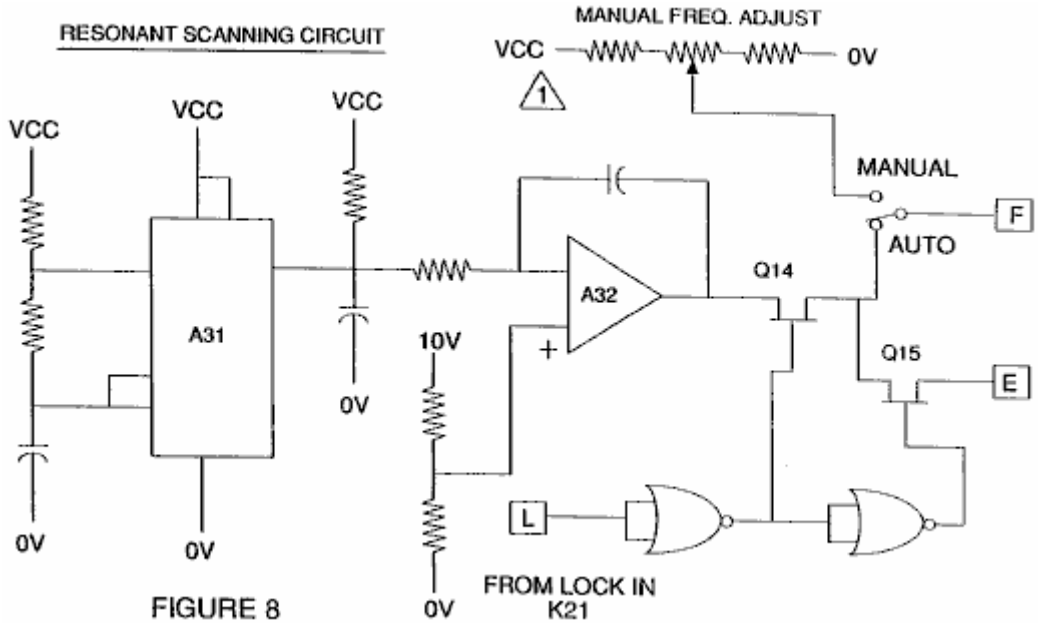


FIGURE 8

Fig.8 is a resonant scanning circuit

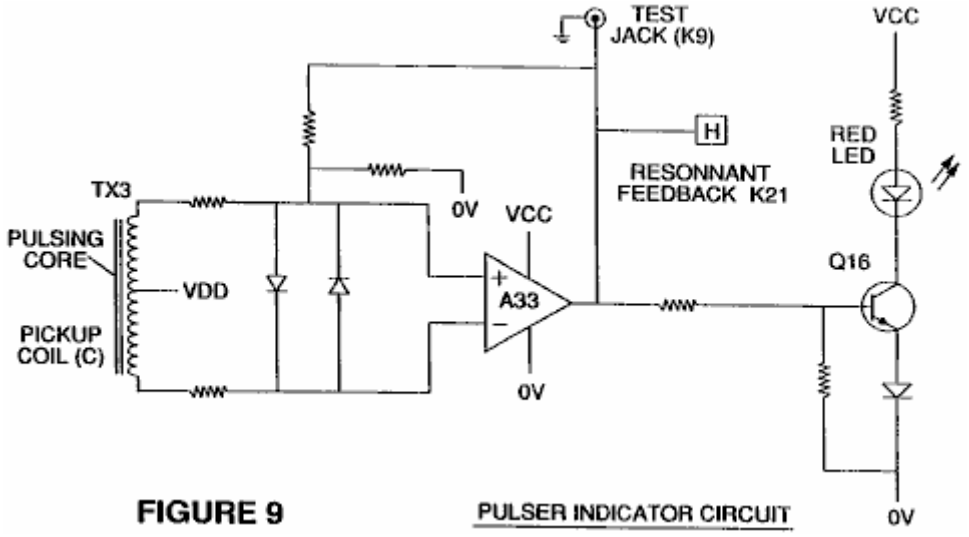


Fig.9 is the pulse indicator circuit.

These four circuits control the pulses transmitted to the resonant-cavity / Water Fuel Cell capacitor.

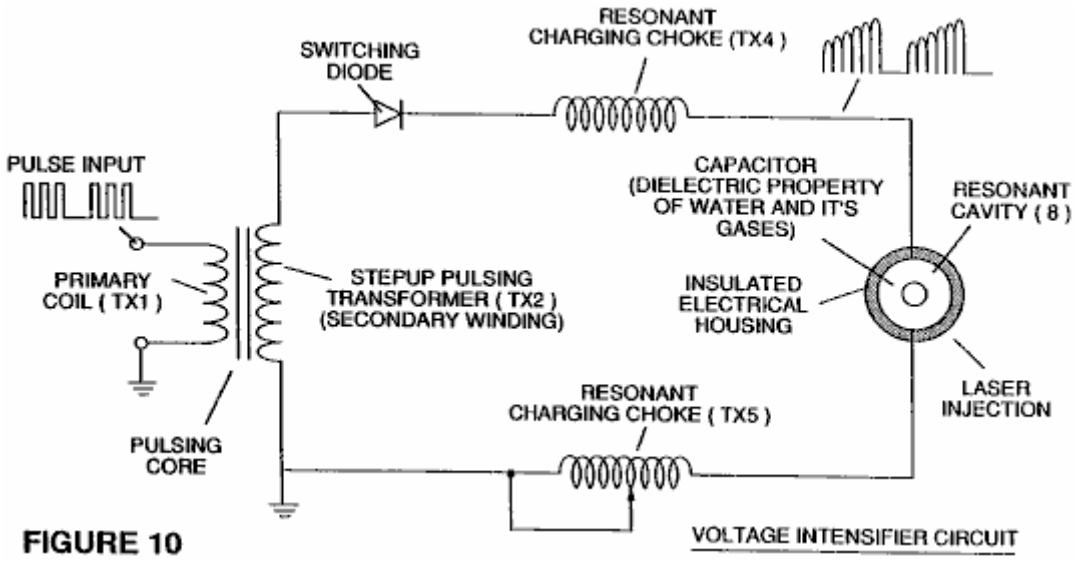


Fig.10 shows the pulsing core and the voltage intensifier circuit which forms the interface between the control circuit and the resonant cavity.

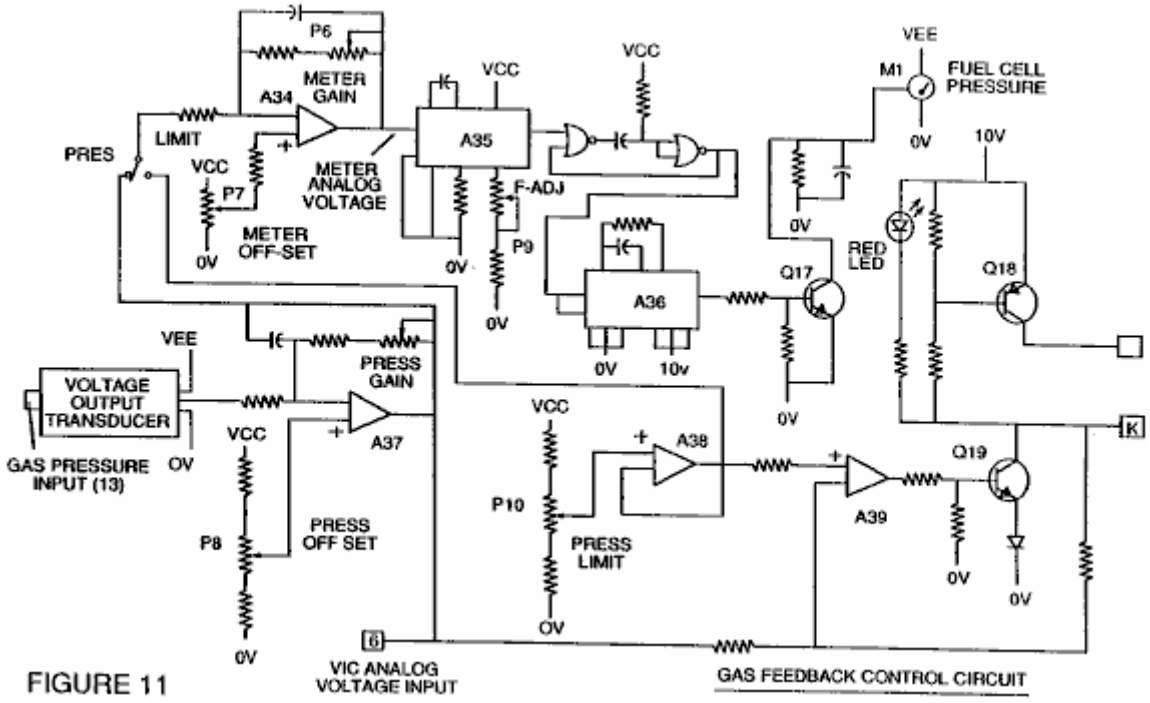


Fig.11 is a gas feedback control circuit

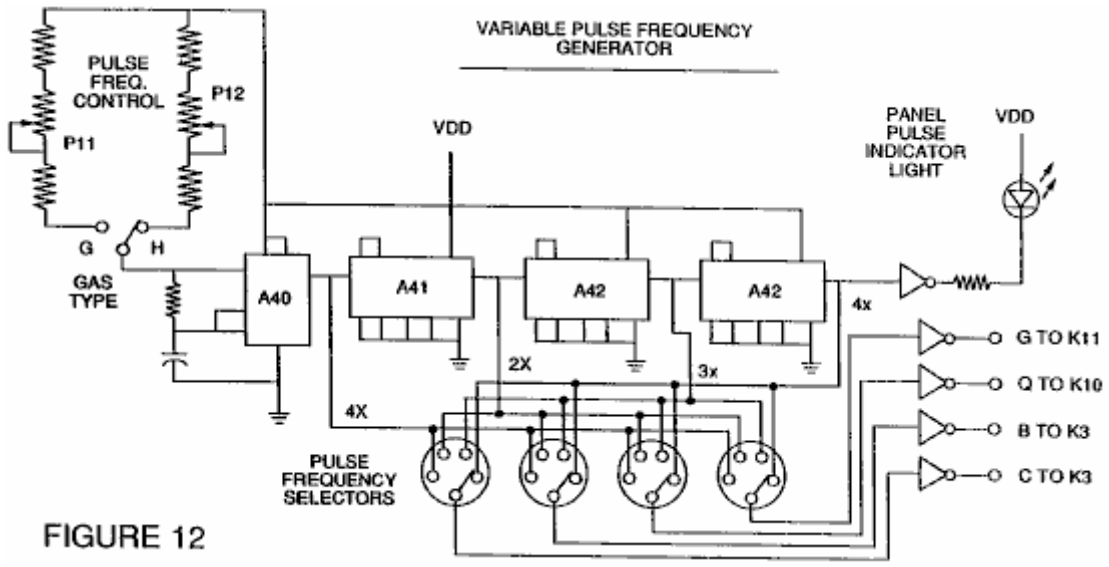
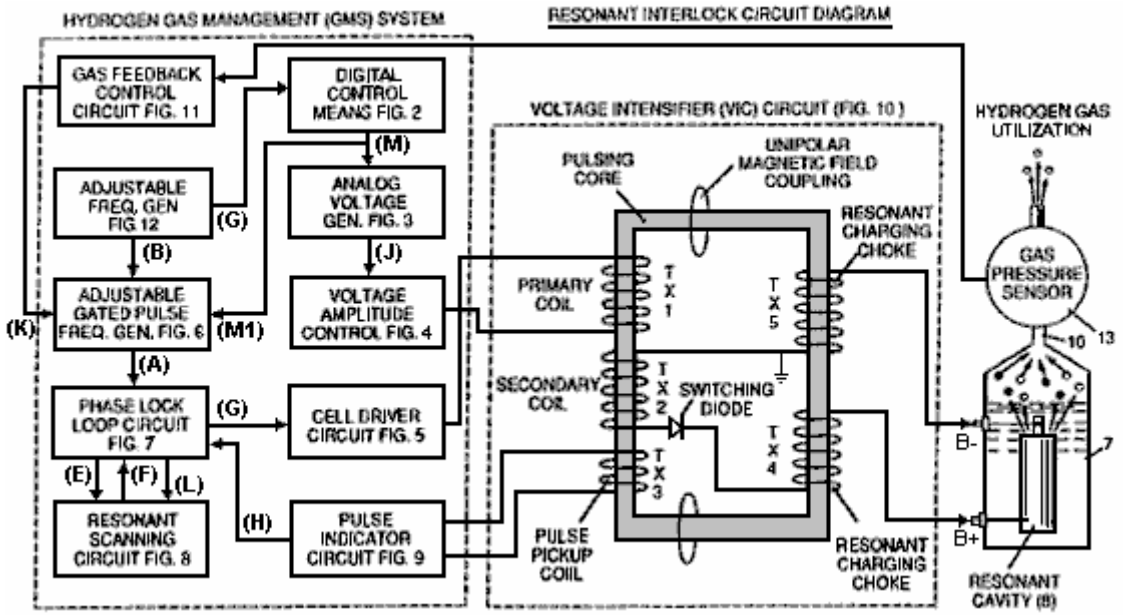


Fig.12 is an adjustable frequency generator circuit.



The circuits are interconnected as shown in **Fig.1** and to the pulsing core voltage intensifier circuit of **Fig.10**, which, among other things, isolates the water capacitor electrically so that it becomes an electrically isolated cavity for the processing of water in accordance with it's dielectric resonance properties. By reason of this isolation, power consumption in the control and driving circuits is minimised when resonance occurs, and current demand is minimised as voltage is maximised in the gas production mode of the water capacitor / Fuel Cell.

The reference letters "**A**" through "**M**" and "**M1**" show, with respect to each separate circuit shown, the point at which a connection in that circuit is made to another of the circuits shown.

In the invention, the water capacitor is subjected to a duty pulse which builds up in the resonant charging choke coil and then collapses. This occurrence allows a unipolar pulse to be applied to the Fuel Cell capacitor. When a resonant condition of the circuit is locked-in by the circuit, current leakage is held to a minimum as the voltage which creates the dielectric field tends to infinity. Thus, when high voltage is detected upon resonance, the phase-lock-loop circuit, which controls the cell driver circuit, maintains the resonance at the detected (or sensed) frequency.

The resonance of the water capacitor cell is affected by the volume of water in the cell. The resonance of any given volume of water contained in the water capacitor cell is also affected by "contaminants" in the water which act as a damper. For example, with a potential difference of 2,000 to 5,000 volts applied to the cell, a current spike or surge may be caused by inconsistencies in the water characteristics which cause an out-of-resonance condition which is remedied instantaneously by the control circuits.

In the invention, the adjustable frequency generator, shown in **Fig.12**, tunes in to the resonant condition of the circuit which includes the water cell and the water inside it. The generator has a frequency capability of 0 to 10 KHz and tunes into resonance typically at a frequency of 5 KHz in a typical 3-inch long water capacitor formed from a 0.5 inch rod inside a 0.75 inch inside-diameter cylinder. At start up, in this example, current draw through the water cell will measure about 25 milliamps; however, when the circuit finds a tuned resonant condition, the current drops down to a 1 to 2 milliamp leakage condition.

The voltage to the capacitor water cell increases according to the turns of the winding and the size of the coils, as in a typical transformer circuit. For example, if 12 volts is sent to the primary coil of the pulsing core and the secondary coil resonant charging choke ratio is 30 to 1, then 360 volts is sent to the capacitor water cell. The number of turns is a design variable which controls the voltage of the unipolar pulses sent to the capacitor.

The high-speed switching diode, shown in **Fig.10**, prevents charge leaking from the charged water in the water capacitor cavity, and the water capacitor as an overall capacitor circuit element, i.e. the pulse and charge status of the water/capacitor never pass through an arbitrary ground. the pulse to the water capacitor is always unipolar. The water capacitor is electrically isolated from the control, input and driver circuits by the electromagnetic coupling through the core. The switching diode in the Voltage Intensifier Circuit (**Fig.10**) performs several functions in the pulsing. The diode is an electronic switch which determines the generation and collapse of an electromagnetic field to permit the resonant charging choke(s) to double the applied frequency and it also allows the pulse to be sent to the resonant cavity without discharging the "capacitor" therein. The diode is, of course,

selected in accordance with the maximum voltage encountered in the pulsing circuit. A 600 PIV ("Peak Inverse Volts") fast switching diode, such as an NVR 1550, has been found to be useful in this circuit.

The Voltage Intensifier Circuit of **Fig.10** also includes a ferromagnetic or ceramic ferromagnetic pulsing core capable of producing electromagnetic flux lines in response to an electrical pulse input. The flux lines affect both the secondary coil and the resonant charging choke windings equally. Preferably, the core is of a closed loop construction. The effect of the core is to isolate the water capacitor and to prevent the pulsing signal from going below an arbitrary ground and to maintain the charge of the already charged water and water capacitor.

In the pulsing core, the coils are preferably wound in the same direction to maximise the additive effect of the electromagnetic field in them. The magnetic field of the pulsing core is synchronised with the pulse input to the primary coil. The potential from the secondary coil is introduced to the resonant charging choke(s) series circuit elements which are subjected to the same synchronous applied electromagnetic field, simultaneously with the primary pulse.

When resonance occurs, control of the gas output is achieved by varying the time of duty gate cycle. The transformer core is a pulse frequency doubler. In a figurative explanation of the workings of the fuel gas generator water capacitor cell, when a water molecule is "hit" by a pulse, electron time-share is effected and the molecule is charged. When the time of the duty cycle is changed, the number of pulses that "hit" the molecules in the fuel cell is modified correspondingly. More "hits" result in a greater rate of molecular disassociation.

With reference to the overall circuit of **Fig.1**, **Fig.3** receives a digital input signal, and **Fig.4** shows the control circuit which applies 0 to 12 volts across the primary coil of the pulsing core. Depending on design parameters of primary coil voltage and other factors relevant to core design, the secondary coil of the pulsing core can be set up for a predetermined maximum, such as 2,000 volts.

The cell driver circuit shown in **Fig.5**, allows a gated pulse to be varied in direct relation to voltage amplitude. As noted above, the circuit of **Fig.6** produces a gate pulse frequency. The gate pulse is superimposed on the resonant frequency pulse, to create a duty cycle that determines the number of discrete pulses sent to the primary coil. For example, assuming a resonant pulse of 5 KHz, a 0.5 KHz gating pulse with a 50% duty cycle, will allow 2,500 discrete pulses to be sent to the primary coil, followed by an equal time interval in which no pulses are passed through. The relationship of resonant pulse to the gate pulse is determined by conventional signal addition/subtraction techniques.

The phase lock loop circuit shown in **Fig.7** allows the pulse frequency to be maintained at a predetermined resonant condition sensed by the circuit. Together, the circuits of **Fig.7** and **Fig.8**, determine an output signal to the pulsing core until the peak voltage signal sensed at resonance is achieved.

A resonant condition occurs when the pulse frequency and the voltage input attenuates the covalent bonding forces of the hydrogen and oxygen atoms of the water molecule. When this occurs, current leakage through the water capacitor is minimised. The tendency of voltage to maximise at resonance, increases the force of the electric potential applied to the water molecules, which ultimately disassociate into atoms.

Because resonances of different waters, water volumes and capacitor cells vary, the resonant scanning circuit of **Fig.8** scans frequency from high to low and back to high, until a signal lock is achieved. The ferromagnetic core of the voltage intensifier circuit transformer, suppresses electron surge in an out-of-resonance condition of the fuel cell. In an example, the circuit scans at frequencies from 0 Hz to 10 KHz and back to 0 Hz. In water having contaminants in the range of 1 part per million to 20 parts per million, a 20% variation in resonant frequency is encountered. depending on water flow rate into the fuel cell, the normal variation range is about 8% to 10%. For example, iron in well water affects the status of molecular disassociation. Also, at a resonant condition, harmonic effects occur. In a typical operation of the cell with a representative water capacitor described below, at a frequency of about 5 KHz, with unipolar pulses from 0 to 650 volts, at a sensed resonant condition in the resonant cavity, on average, the conversion into gas occurs at a rate of about 5 US gallons (19 litres) of water per hour. To increase the rate, multiple resonant cavities can be used and/or the surfaces of the water capacitor can be increased, however, the water capacitor cell is preferably small in size. A typical water capacitor may be formed from a 0.5 inch diameter stainless steel rod and a 0.75 inch inside-diameter cylinder which extends over the rod for a length of 3 inches.

The shape and size of the resonant cavity may vary. Larger resonant cavities and higher rates of consumption of water in the conversion process require higher frequencies up to 50 KHz and above. The pulsing rate, to sustain such high rates of conversion, must be increased correspondingly.

From the above description of the preferred embodiment, other variations and modifications of the system disclosed will be evident to those skilled in the art.

CLAIMS

1. A control circuit for a resonant cavity water capacitor cell utilised for the production of a hydrogen- containing fuel gas, including an isolation transformer with a ferromagnetic core, and having one side of a secondary coil connected in series with a high-speed switching diode to one plate of the water capacitor of the resonant cavity, and the other side of the secondary coil connected to the other plate of the water capacitor, to form a closed-loop electronic circuit utilising the dielectric properties of water as part of the electronic circuit, and a primary coil connected to a pulse generator.

2. The circuit of Claim 1. in which the secondary coil includes segments which form a resonant charging choke circuit in series with the water capacitor.

3. The circuit of Claim 1. in which the pulse generator includes an adjustable first frequency generator and a second gated pulse frequency generator which controls the number of pulses produced by the first frequency generator, sent to the primary coil during a period determined by the gate frequency of the second pulse generator.

4. The circuit of Claim 1. further including a means for sensing the occurrence of a resonant condition in the water capacitor of the resonant cavity.

5. The circuit of Claim 4. in which the means for sensing is a pickup coil on the ferromagnetic core of the transformer.

6. The circuit of Claim 4. or Claim 5. in which the sensing means is interconnected to a scanning circuit and a phase-lock-loop circuit, by which the pulsing frequency sent to the primary coil of the transformer is maintained at a sensed frequency corresponding to a resonant condition in the water capacitor.

7. The circuit of Claim 1. including means for adjusting the amplitude of a pulsing cycle sent to the primary coil.

8. The circuit of Claim 6. including further means for maintaining the frequency of the pulsing cycle at a constant frequency regardless of pulse amplitude.

9. the circuit of Claim 3. in which the gated pulse frequency generator is connected to a sensor which monitors the rate of gas production from the cell and controls the number of pulses sent to the cell in a gated frequency, corresponding to the rate of gas production.

10. The circuit of Claim 7. or Claim 8. or Claim 9. further including a gas-pressure sensor in an enclosed water capacitor resonant cavity which also includes a gas outlet, where the gas-pressure sensor is connected to the circuit to determine the rate of gas production with respect to ambient gas pressure in the water capacitor enclosure.

11. The methods and apparatus as substantially described herein.

STEPHEN MEYER

Patent application US 2005/0246059 3rd November 2005 Inventor: Stephen F. Meyer

MLS-HYDROXYL FILLING STATION

This is a patent application from Stephen Meyer, brother of the late Stan Meyer. While this application mentions filling stations, it is clear that the design is aimed at use in vehicles with internal combustion engines. I believe that the impedance-matching interface between the alternator and the cell electrodes is particularly important. The water-splitter cell uses sets of three pipes in a concentric array which results in small gaps between the innermost, middle and outer pipe. Stephen refers to these three electrode pipes as a "wave-guide", so please bear that in mind when reading this patent application. Stephen uses the word "hydroxyl" to refer to the mixture of hydrogen and oxygen gases produced by electrolysis of water. Other people use the word "hydroxy" to describe this mixture, so they should be considered interchangeable.

The operation of this system as described here, calls for the generating power to be removed when the gas pressure in the generating chambers reaches 5 psi. The gas is then pumped into a pressure chamber where the pressure ranges from 40 psi to 80 psi, at which point the compressor is powered down and the excess gas vented to some external storage or using device. It is not until this is completed that the power is applied again to the generating chambers. May I remark that, in my opinion, there is no need to remove the power from at generating chambers at any time when this system is in operation, since all that that does is to lower the generating capacity, unless of course, the production rate is so high that it exceeds the level of demand.

ABSTRACT

The usefullness of this system, it's configuration, design and operation, are the keystone of a new type of automation: the production of hydroxyl gases from renewable sources.

BACKGROUND OF THE INVENTION

Fuel Cell and auto industries have been looking for methods and apparatus that can supply a source of hydrogen and oxygen for its new hybrid industry. This invention is such a device.

SUMMARY OF THE INVENTION

The invention is a computerised, automatic, on-site/mobile hydroxyl gas producing filling station which allows the products being produced to be used, either by the hydrogen fuel cells installed in automobiles, trucks, buses, boats and land-based generating applications, or in any internal combustion engine.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig.1 shows the configuration of the components which go to make up the MLS-hydroxyl Filling Station.

Fig.2 shows the software display which the operator uses to monitor and control the production of hydroxy gases and heat.

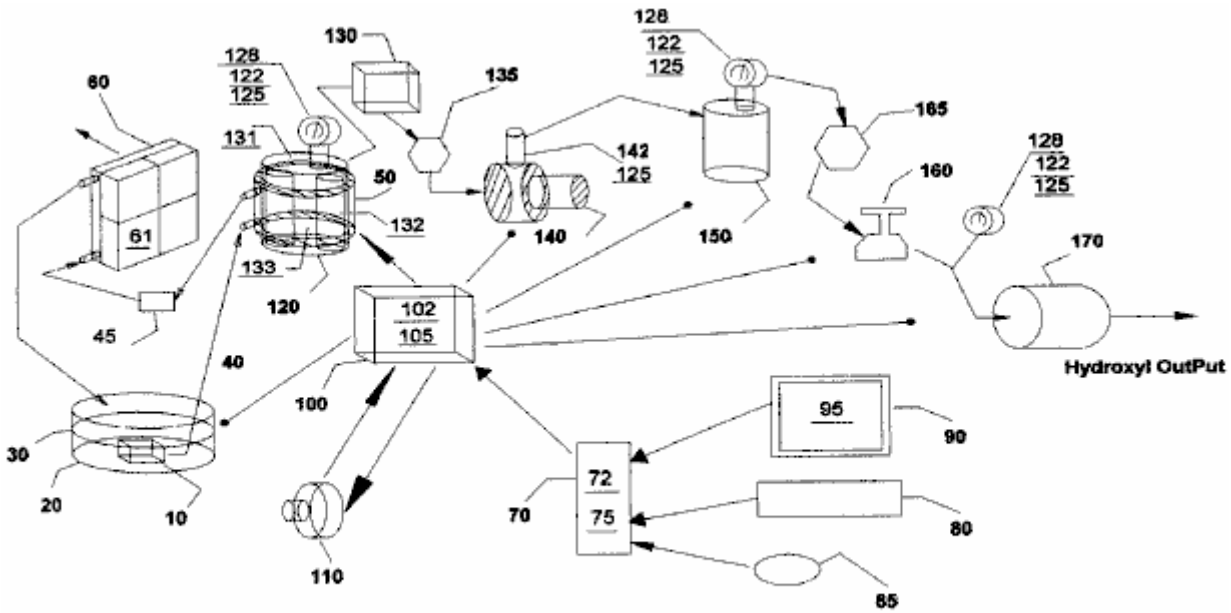
Fig.3 shows the methods, configuration, and apparatus used in the hydroxyl producing cell system **120**.

Fig.4 shows the electronic impedance-matching circuits **102**, connected between the dual three-phase synchronised generators (**110A** and **110B** in **Fig.3**) and each of the electrodes or "waveguide" arrays **132** in cell **120** of **Fig.3**. Note that only generator **A** is depicted in **Fig.4** as being connected to arrays **A**, **B** and **C** using PC cards **1 to 3**. generator **B** is connected to arrays **D**, **E** and **F** using cards **4 to 6**.

Fig.5 Shows the signals emitted by each of the impedance-matching circuits (**102** in **Fig.4** mounted on cards **1 to 6**) which are applied to each of the cylinder arrays (**132** in **Fig.3**) installed in hydroxyl cell **120**. These sets of signals with their offset phase relationship, frequencies and amplitudes, are the driving forces producing the hydroxy gases in cell **120** of **Fig.3**.

Fig.6 shows the high-frequency ringing signal which is produced between points **T1** and **T2** in the impedance-matching circuit **102** in **Fig.4**. It is this ringing which enhances the production of the hydroxyl gas in cell **120** of **Fig.3**.

DETAILED DESCRIPTION OF THE DRAWINGS



MLS- Hydroxyl Filling Station (MLS-HFS)

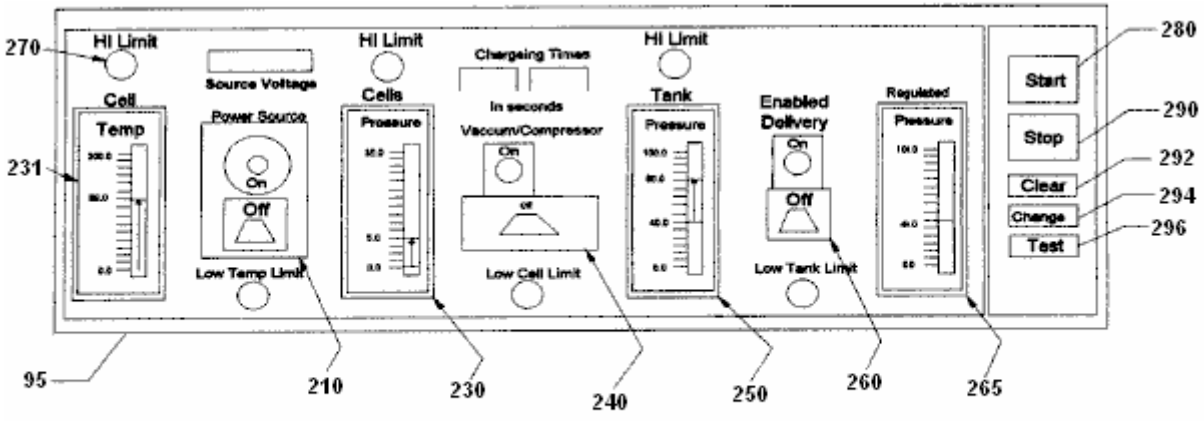
Fig 1

The heat-removing section in **Fig.1** consists of a liquid bath **30** and its container **20**, a liquid circulating pump **10**, conveying-conduits **40**, cooling chamber 50 attached to hydroxyl generating cell **120**, filter **45**, radiator **60** and cooling fans **61** attached to it.

The automatic-control section in **Fig.1** consists of a computer **70**, software program **75**, video monitor **90** and it's graphic operator display **95** (**Fig.2**), pointer **85**, keyboard **80**, interface card **72**, and Input/Output controller **100** with it's driver electronics cards **102** and **105**.

Dual three-phase power sources **110** and impedance-matching circuits **102**, provide the power needed to drive the hydroxyl cell **120**.

The remaining apparatus is used to convey the gases from cells **120**, through liquid trap **130**, through gas flow restriction valve **135**, elevate its gas pressures through compressor **140**, transfer them to storage tank **150**, then deliver the gases through safety cut off **165**, regulators **160** and through flash-back arrestor **170** for external delivery.

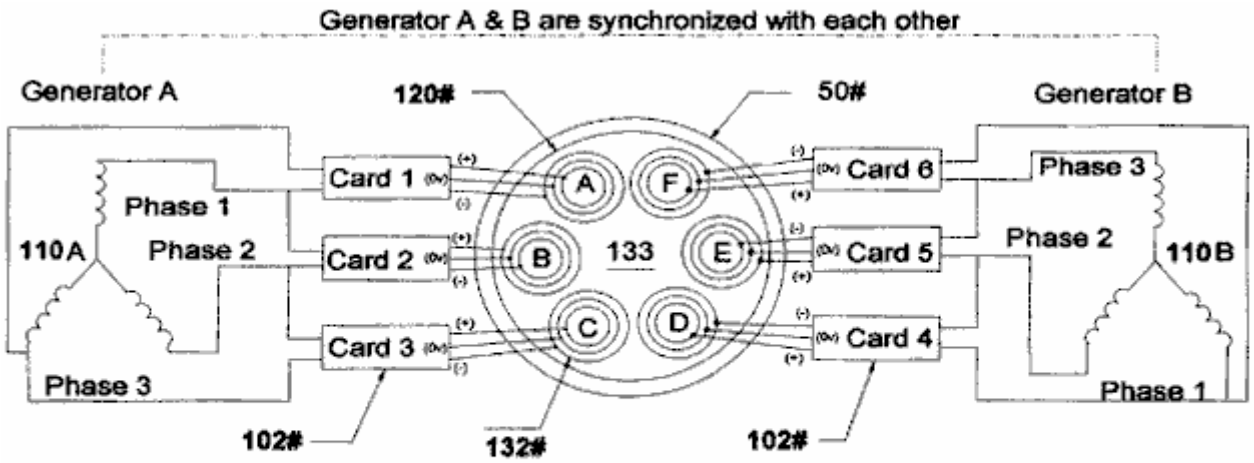


MLS-Hydroxyl Filling Station (MLS-HFS) Graph Display and Operator Control

Fig-2

Fig.2 shows the layout and functions of the operator control display **95** of program **75** in **Fig.1**. It consists of cell temperature indicator **230**, vacuum controller **240**, high-pressure tank indicator **250**, delivery controller **260**, delivery regulated-pressure indicator **265** and related alarm/status indicators **270**. Also, software control buttons

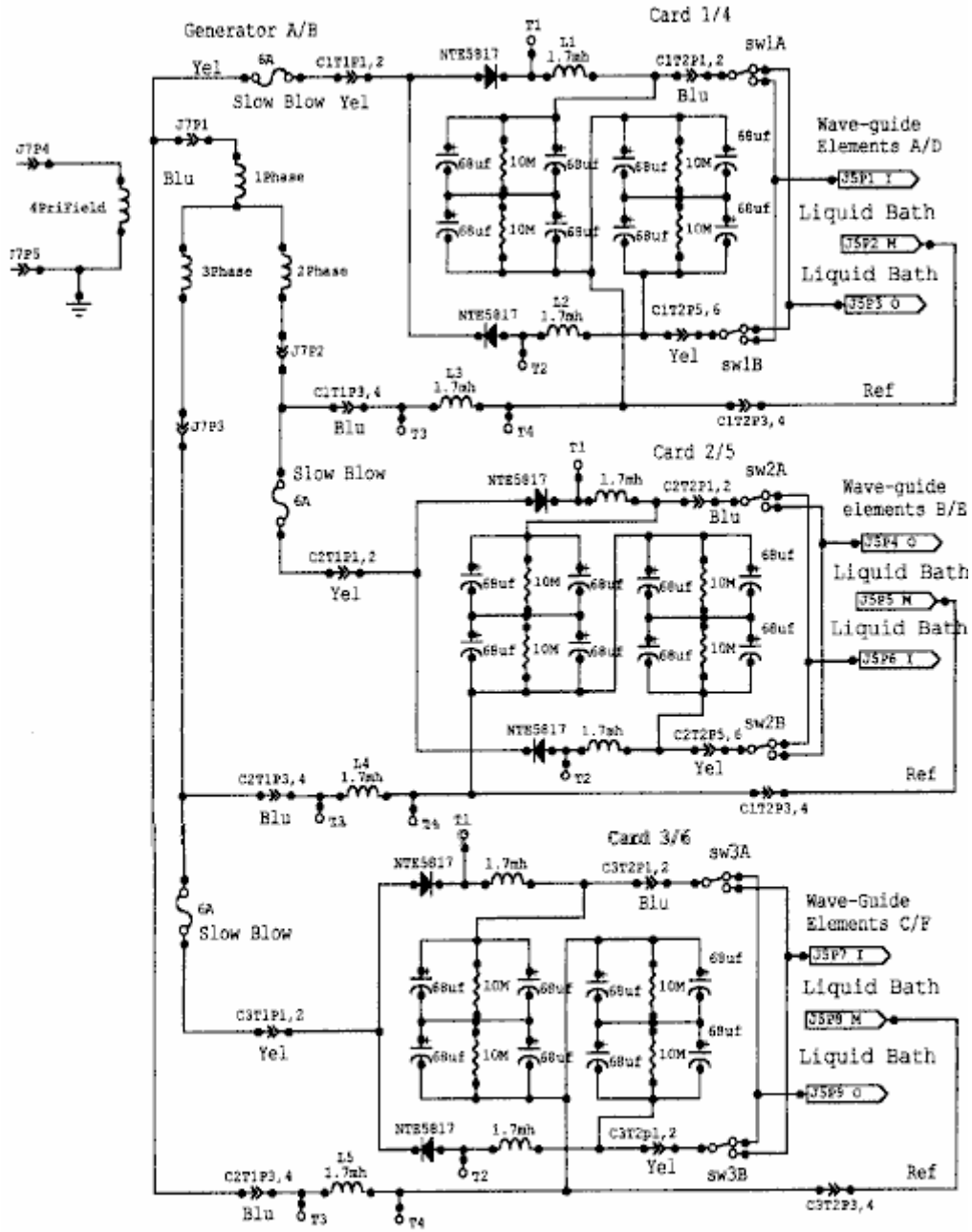
are provided to start **280**, stop **290**, clear data **292**, change setting **294** and the testing of equipment and their sequences **296**.



Configuration of Hydroxyl gas producting appartuses

Fig-3

Fig.3 shows the configuration of our proprietary hydroxyl-producing apparatus **120** consisting of dual three-phase power source **110**, impedance matching electronic circuits **102** and gas converter devices **132** submerged in a bath of water **133** in cell **120**. The drawing also shows the water jacket **50** surrounding the cell **120** that helps lower its temperature and allows more production of the hydroxyl gases at higher voltage signals as shown in **Fig.5**.



Impedance matching circuits 102

Fig-4

Fig.4 shows the electrical circuits **102**, used to drive the gas converting arrays (**132** in **Fig.3**) submerged in a bath of water **133** in cell **120**. **Fig.4** shows three identical circuits connected to each of the three-phase signals from one half of the dual three-phase generator **110A** in **Fig.3**. The circuits **102,** convert the AC signal from each phase of **110** into a modulated signal as depicted by **Fig.5**. These signals are then coupled to the triple array elements **132**, (Inside, Middle and Outside) by alternating the connection between the Inside and Outside elements of the arrays (**132** in **Fig.3**).

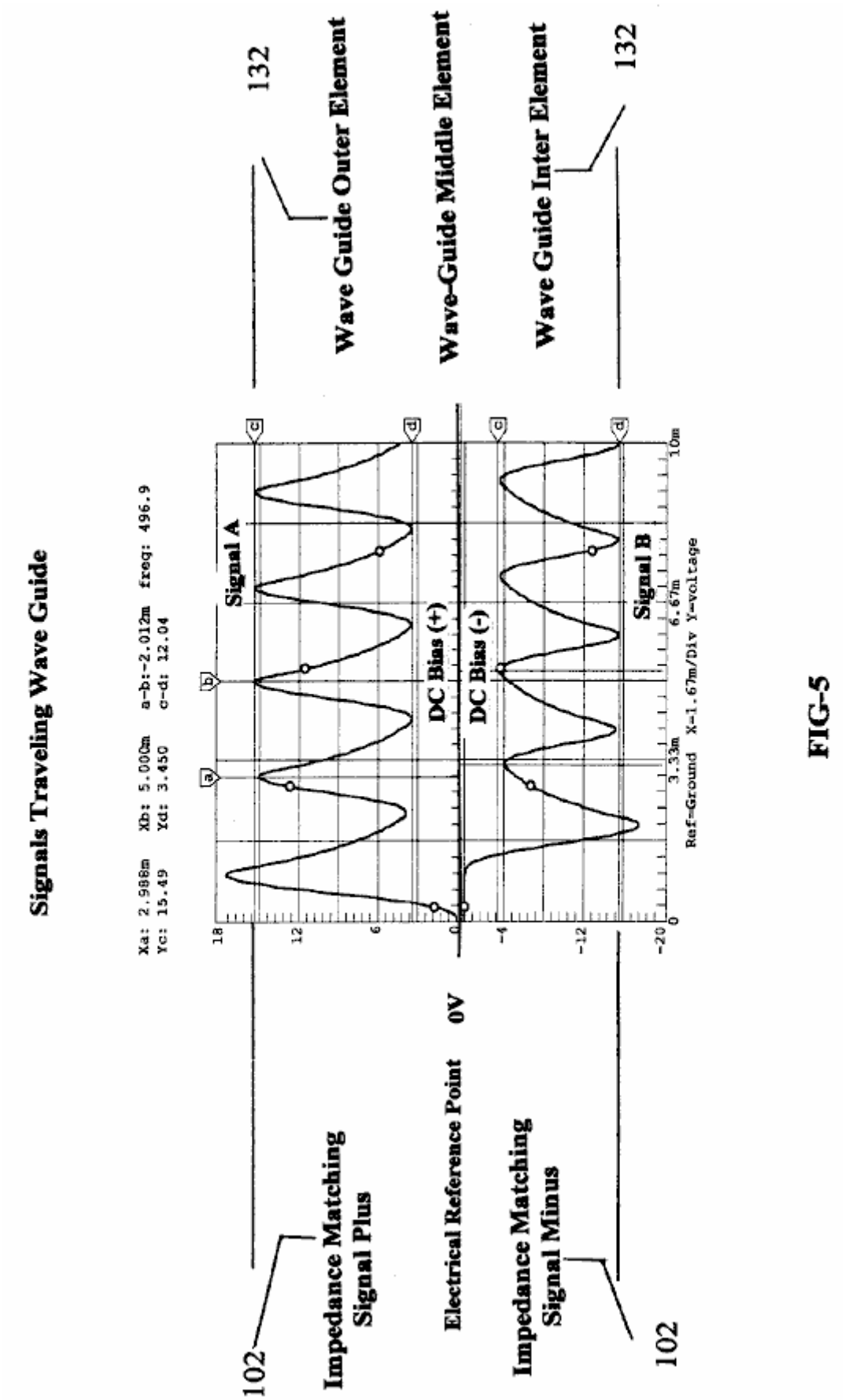


FIG-5

Fig.5 shows the composite signals applied to each of the arrays (**132** in **Fig.3**) submerged in the water bath **133** in cell **120**, and indicates the differential voltages used in the hydroxyl producing process. Note that the Middle wave-guide element is used as the electrical reference point for both the Outside and Inside elements of array **132**. It is this composite signal applied to the surface of the stainless steel elements in array **132** submerged in water bath **133**, heat allows the ions from the elements in array **132** to cross its water surface barriers **133** and contribute to the hy-droxyl production. Note the DC bias voltage +,- on either side of the centre electrical

reference point 0V. It is this bias voltage being modulated by multi-polarity differential signals from **102**, that contributes to the wave-guide action of arrays **132**. Also, the frequency of the waveform shown in **Fig.5** is adjusted to match the electrical wavelength of the arrays **132** of **Fig.3** and the impedance of water bath **133**.

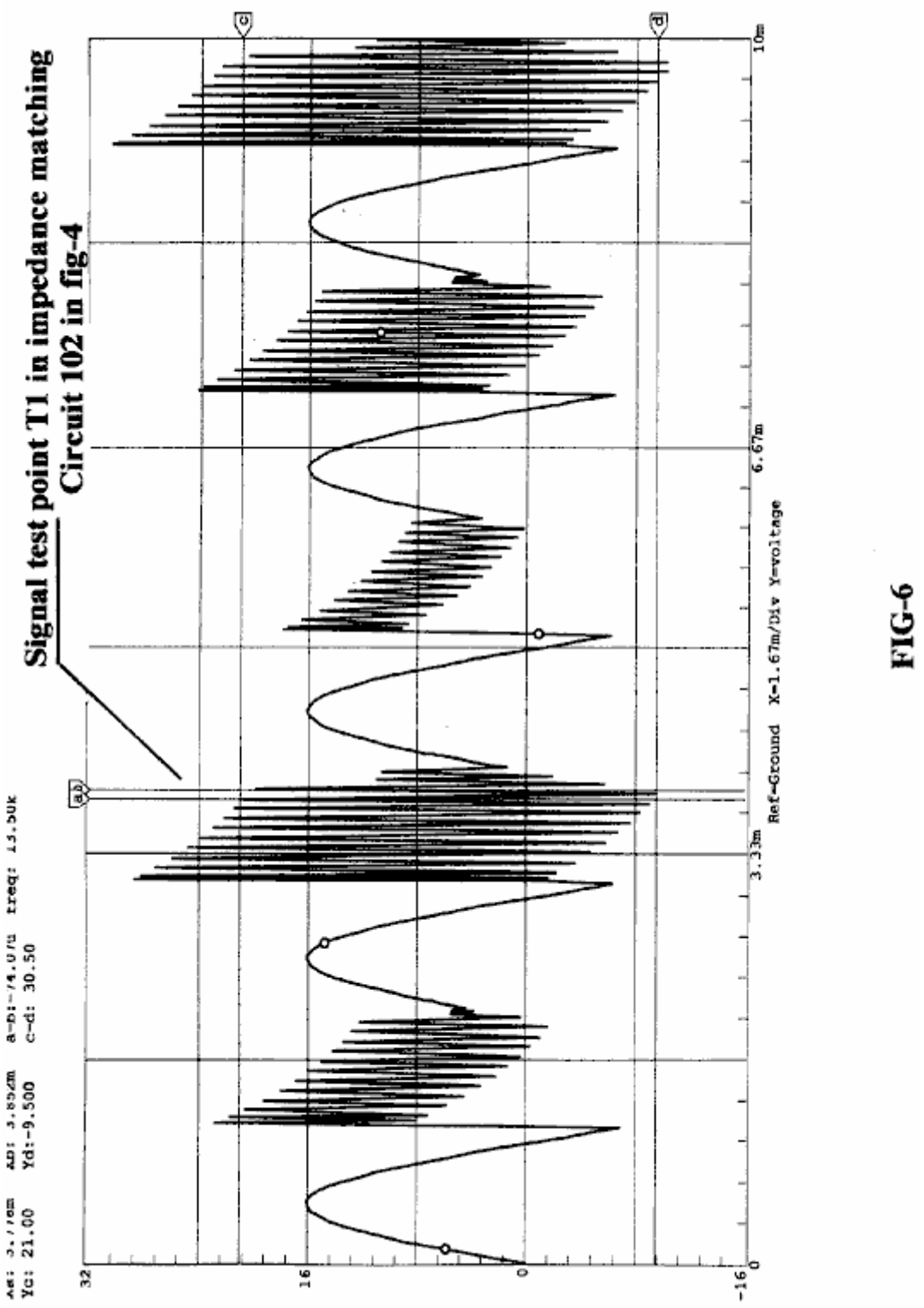


FIG-6

Fig.6 shows the high-frequency ringing signals which contribute to the operation of the hydroxyl production. just as a tuning fork rings when struck by a hammer, so do the wave-guide elements in array **132** immersed in the hydroxyl-generating liquid **133** when struck by the electrical signals shown in **Fig.5** and **Fig.6**, coming from the impedance-matching circuits **102** shown in **Fig.4**.

Brief Description of Sequences

This invention is a computerised Hydroxyl Gas producing filling station "**MLS-HFS**" designed to provide automatic control of its on-site gas production and delivery.

The MLS-HFS shown in **Fig.1**, is a hydroxyl gas and heat generating system which uses a renewable source of liquid supply **30** such as water. It uses a computer control program **75** with display interface **95**, for the monitoring, adjusting and controlling of the electronic and hardware apparatus and process logic. The electronic circuits **102** mounted in driver **100**, control the production of the gases and heating while circuit **105** controls the process and routeing of the hydroxyl gas.

The system consists of a low-pressure hydrolyser cell **120** in **Fig.1**, a liquid trap **130**, an adjustable flow-restriction valve **135**, high-pressure vacuum pump **140**, and check valve **142** installed in **140**. It also contains a high-pressure storage tank **150**, an alarm/low-pressure cut-off valve **165**, gas regulator **160**, flashback arrestor **170**, over-pressure safety release valves **125**, pressure gauges **128**, analogue pressure-sending units **122** installed on cell **120**, and tank **150** at the regulating side of regulator **160**. Also, **125** is installed on Compressor **140** high-pressure output. The computer controller **70**, monitor **90**, keyboard **80**, interface I/O card **72** and software position pointer **85**, are used to control the production process, using electronic driver **100** through it's PC boards **105** and their attached control devices. The power to the cell-driving circuits **102**, installed in driver **100**, is supplied from a dual three-phase isolated power source **110**. The amplitude, signal phases and frequency from this power source is controlled by signal adjustments coming from the computer **70**.

Detailed Description

Sequence of Operation

The system shown in **Fig.1** is monitored and controlled by the software program **75**, computer **70**, monitor **90**, keyboard **80**, pointer **85**, and display interface **95** in **Fig.2**.

The software program has five main functions, namely: to purge the system of ambient air, check and test for any equipment malfunctions, prepare the system for production, monitor and control the current activities of the production process, and the safety shutdown of the system if alarms are detected.

During the initial installation, and again after any repairs, the total system is purged using the vacuum pump **140**, using manual procedures to ensure that all ambient air has been removed from the system. Before the system is put into service, the operator can test the operation of the system by using the graphic display. The main functions of the testing is to ensure that the temperature electronics **131** attached to the hydroxyl cells **120**, transferring compressor **140** and analogue pressure sensors **122** mounted on cells **120**, high-pressure tank **150** and the discharge side of regulator **160** used for control and monitoring, are working properly. the operator can then activate the Run Sequence of the program **75** via the start software button **280** in **Fig.2** on graphic display **95**.

During the initial startup phase of the system, the computer program will configure the system for the purge sequence. this sequence allows the vacuum pump **140** to draw down the hydroxyl cells **120** liquid trap **130** coupled to flow-restriction valve **135**, to remove all ambient air from them. Once the program has done this and detected no leaks in the system, it then prepares the system for gas production by switching the gas flow from cells **120** to high-pressure tank **150** and on to the output flashback protector **170**.

The program starts it's production sequence by turning on the cooling system pump **10** which is submerged in the liquid bath **30**, contained in vessel **20**. The cooling liquid is pumped through the cooling jacket **50** which is attached to the outside of cells **120**, through filter **45** and then through an air-cooled radiator **60**. Fans attached to the radiator are turned on for cooling.

Next, the computer turns on the dual three-phase power source **110**, which supplies operating power to the frequency, phase-shifting, signal amplitude and impedance-matching circuits coupled to the hydroxyl generating cells.

The result of this is just like the operation of a radio transmitter matching it's signal to the air via the antenna impedance. **Fig.3** shows the relationship of this configuration to arrays **132**, water bath **133** and Signals (**Fig.5** and **Fig.6**).

While the power source **110** is operating, the computer **70** is monitoring the pressure **122** and temperature **131** of hydroxyl cells **120**. When the cell pressure reaches a typical level of **5** pounds per square inch, the power source is turned off and compressor **140** is turned on the pump the gas into pressure tank **150**. When the pressure in the hydroxyl cells **120** is drawn down to near zero, the compressor is turned off and the power to the gas generating cells is turned back on again, to repeat the cycle.

The production cycle is repeated until tank **150** reaches a pressure of, typically, 80 psi, at which time the computer enables the output pressure regulator **160** which is typically set to operate at 40 psi, for the delivery of the hydroxyl gas to some external storage system or device. During this operation, the computer program handles all switching and displays the current status and any alerts or warning messages for the operator on the graphical display **95**.

Impedance-Matching Circuit 102:

The impedance-matching circuits **102** in **Fig.4**, convert the sinewave signals coming from the three-phase power source (**110** in **Fig.3**) into multi-polarity differential signals (**Fig.5**) which are applied to the triple wave-guide cluster arrays **132 A, 132B, 132C, 132D, 132E** and **132F** installed in cell **120**.

It is this converted signal, along with the phase relationship of the power source **110** and the triple wave-guide elements in cluster **132** submerged in water bath **133**, which produce the hydroxyl gases. It is important to note that not only is the gas produced between the elements in the array, but also between each array installed in the cell - see the phase relationship of array **A-B-C** shown in **Fig.3**. Also note that the array elements themselves are supplying many of the ions needed for the production of the gases.

Sequence of Hydroxyl Gas Generation:

Once the hydroxyl-generating cell **120** has been purged of ambient air and the production routing completed (**Fig.1**), the dual three-phase power source **110** is activated, supplying frequency, amplitude and phase signals to the impedance-matching circuitry **102**. The converted signals from **102** are then applied to cell array **132** for processing. It is the combination of the impedance-matching circuits signal transformations (as shown in **Fig.5** and **Fig.6**), the cell configuration and materials used in arrays **132**, and the rotational phase relationship between arrays **AD**, **BE** and **CF** and the submersion of these arrays in a bath of water **133**, that allows this system to produce large amounts of hydroxyl gases. The computer program 75 and it's graphic display **95**, is used by the operator to adjust the rate of gas production and set the upper limit to which the low-pressure cell **120** will charge.

After the cell **120** has reached its upper pressure cut-off limit (typically 5 psi), the power source **110** is turned off, enabling the compressor **140** to start its draw-down and transfer of the gases to the high-pressure tank **150**. When the pressure in the cell **120** reaches a low-level limit (near zero psi), **140** stops its charging cycle of **150**. Check valve **142** which is installed in **140**, prevents any back flow of gases to **120** from high-pressure tank **150**. The power source **110** is then turned back on to repeat the cycle. These charging cycles continue until the high-pressure tank **150** reaches it's upper pressure limit (typically 80 psi), at which point the hydroxyl production is stopped. As the gases in the high-pressure tank are being used or transferred to some external storage system, the pressure in **150** is monitored at the output of pressure-regulator **160**, until the low-pressure limit for this tank is reached (typically 40 psi). When this pressure level is reached, the hydroxyl gas production is started again.

During the operation of cell **120**, it's temperature is monitored to ensure that it does not exceed the "out of limits" conditions set by control **231** and monitored via the graphics display **95**. If the temperature exceed the limit set, then the gas production is stopped and the computer program alerts the operator, indicating the problem. The cooling system **30** which uses water jacket **50** surrounding cell **120**, helps to reduce the temperature and allows higher rates of gas production.

After extended running times, the water in cell **120** is replenished from bath **30** and filtered by **45**, to help control the operating impedance of the cell.

CLAIMS

1. The MLS-HFS information in this specification is the embodiment of the claims.

2. The system according to Claim 1 further enhances the production of hydroxyls based on the configuration of the hydroxyl gas-producing apparatuses of **Fig.3**.

3. The system according to Claim 1 further enhances the production of hydroxyls based on the configuration of the impedance-matching circuits of **Fig.4**.

4. The system according to Claim 1 further enhances the production of hydroxyls based on the application of the electrical signals shown in **Fig.5** applied to signal travelling wave-guides **132** submerged in a bath of water **133** installed in cell **120** and configured as depicted in **Fig.3**.

5. The system according to Claim 1 further enhances the production of hydroxyls based on the resonating action of the electrical signals depicted in **Fig.6**.

6. The system according to Claim 1 further enhances the production of hydroxyls based on the software program's ability to control the production of hydroxyl gases; controlling it's process limits, controlling it's storage and controlling it's delivery via operator controller **Fig.2**.

7. The software program **75** according to Claim 6, further enhances the safety of the production of hydroxyls based on the monitoring of high and low limits and either alerting the operator of the conditions and/or stopping the production on device failures via operator controller **Fig.2**.

8. The software according to Claim 6 further enhances the safety of the hydroxyls based on its ability to purge the system of ambient air before starting the production of hydroxyl gases.

Dr HENRY PUHARICH

Dr Andrija Puharich (who later changed his name to Henry Puharich) reportedly drove his motor home for hundreds of thousands of miles around North America in the 1970s using only water as fuel. At a mountain pass in Mexico, he collected snow for water. Here is an article which he wrote:

Cutting The Gordian Knot of the Great Energy Bind

by Andrija Puharich

Introduction

It is hardly necessary to weigh the value of the World Energy bank account for any sophisticated person, these days. It is grim. The oil reserves will dwindle away in a score of years or so, and the coal reserves will be gone in some twelve score years. This is not to say that the outlook is hopeless. There is an abundance of alternative energy sources, but the economics of development and exploitation present an enormous short-term strain on the world political and banking resources.

Visionary scientists tell us that the ideal fuel in the future will be as cheap as water, that it will be non-toxic both in its short-term, and in its long-term, effects, that it will be renewable in that it can be used over and over again, that it will be safe to handle, and present minimal storage and transportation problems and costs. And finally that it will be universally available anywhere on earth. What is this magical fuel, and why is it not being used? The fuel is water. It can be used in its fresh water form. It can be used in its salt water form. It can be used in its brackish form. It can be used in its snow and ice form. When such water is decomposed by electrolytic fission into hydrogen and oxygen gases, it becomes a high energy fuel with three times the energy output which is available from an equivalent weight of high grade gasoline.

Then why is water not being used as a fuel? The answer is simple - it costs too much with existing technology to convert water into hydrogen and oxygen gases. The basic cycle of using water for fuel is described in the following two equations, familiar to every high school student of Chemistry:

H_2O Electrolysis + 249.68 BTU $\Delta G \rightarrow H_2 + (1/2)O_2$ per mole of water.....(1)

(1 mole = 18 gm). This means that it requires 249.688 BTU of energy (from electricity) to break water by electrolysis into the gases hydrogen and oxygen.

$H_2 + (1/2)O_2$ + catalyst $\rightarrow H_2O - \Delta H$ 302.375 BTU per mole of water.....(2)

This means that 302.375 BTU of energy (heat or electricity) will be released when the gases, hydrogen and oxygen, combine. The end product (the exhaust) from this reaction is water. Note that more energy (under ideal conditions) is released from combining the gases than is used to free them from water. It is know that under ideal conditions it is possible to get some 20% more energy out of reaction (2) above, then it takes to produce the gases of reaction (1) above. Therefore, if reaction (1) could be carried out at 100% efficiency, the release of energy from reaction (2) in an optimally efficient engine (such as a low temperature fuel cell), there would be a net energy profit which would make the use of water as a fuel an economically feasible source of energy .

The cost of producing hydrogen is directly related to the cost of producing electricity. Hydrogen as produced today is generally a by-product of off-peak-hour electrical production in either nuclear or hydroelectric plants. The electricity thus produced is the cheapest way of making hydrogen. We can compare the cost of production of electricity and the cost of producing hydrogen. The following table is adapted from Penner whose data source is based on Federal Power Commission, and American Gas Association Figures of 1970 and on a 1973 price evaluation (just before the OPEC oil price escalation.)

Table 1: Relative Prices in Dollars per 106 BTU

Cost Component	Electricity	Electrolytically-Produced H
Production	2.67 (b)	2.95 to 3.23 (b)
Transmission	0.61	0.52 (c)
Distribution	1.61	0.34
Total Cost	$4.89	$3.81 to $4.09

If we compare only the unit cost of production of electricity vs Hydrogen from the above table:

106 BTU H_2 / 106 BTU El = $3.23 / $2.67, or 20.9% higher cost, H_2

It must also be noted that the price of natural gas is much cheaper than either electricity or hydrogen, but because of the price fluctuations due to recent deregulation of gas it is not possible to present a realistic figure. In the opinion of Penner, if the hydrogen production cost component of its total cost could be reduced three fold, it would become a viable alternate energy source. In order to achieve such a three-fold reduction in production costs, several major breakthroughs would have to occur.

1. **Endergonic Reaction** A technological breakthrough which permits 100% conversion efficiency of water by electrolysis fission into the two gases, Hydrogen as fuel and Oxygen as oxidant.

2. **Hydrogen Production in Situ** A technological breakthrough which eliminates the need and cost of hydrogen liquefaction and storage, transmission, and distribution, by producing the fuel in situ, when and where needed.

3. **Exergonic Reaction** A technological breakthrough which yields a 100% efficient energy release from the combination of hydrogen and oxygen into water in an engine that can utilize the heat, steam, or electricity thus produced.

4. **Engine Efficiency** By a combination of the breakthroughs outlined above, 1, 2, and 3 utilized in a highly efficient engine to do work, it is theoretically possible to achieve a 15% to 20% surplus of energy return over energy input.

It is of interest to record that a new invention is now being developed to realise the above outlined goal of cheap, clean renewable and high grade energy. A Thermodynamic Device has been invented which produces hydrogen as fuel, and oxygen as oxidant, from ordinary water or from sea water, eliminating the cost and hazard of liquefaction, storage, transmission, and distribution. The saving of this aspect of the invention alone reduces the total cost of hydrogen by about 25%.

This Thermodynamic Device is based on a new discovery - the efficient electrolytic fission of water into hydrogen gas and oxygen gas by the use of low frequency alternating currents as opposed to the conventional use of direct current, or ultra-high frequency current today. Such gas production from water by electrolytic fission approaches 100% efficiency under laboratory conditions and measurements. No laws of physics are violated in this process.

This Thermodynamic Device has already been tested at ambient pressures and temperatures from sea level to an altitude of 10,000 feet above sea level without any loss of its peak efficiency. The device produces two types of gas bubbles; one type of bubble contains hydrogen gas; the other type contains oxygen gas. The two gases are thereafter easily separable by passive membrane filters to yield pure hydrogen gas, and pure oxygen gas.

The separate gases are now ready to be combined in a chemical fusion with a small activation energy such as that from a catalyst or an electrical spark, and yield energy in the form of heat, or steam, or electricity as needed. When the energy is released by the chemical fusion of hydrogen and oxygen, the exhaust product is clean water. The water exhaust can be released into nature and then renewed in its energy content by natural processes of evaporation, solar irradiation in cloud form, an subsequent precipitation as rain on land or sea, and then collected again as a fuel source. Or, the exhaust water can have its energy content pumped up by artificial processes such as through solar energy acting through photocells. Hence, the exhaust product is both clean and renewable. The fuel hydrogen, and the oxidant oxygen, can be used in any form of heat engine as an energy source if economy is not an important factor. But the practical considerations of maximum efficiency, dictate that a low temperature fuel cell with its direct chemical fusion conversion from gases to electricity offers the greatest economy and efficiency from small power plants of less than 5 kilowatts.

For large power plants, steam and gas turbines are the ideal heat engines for economy and efficiency. With the proper engineering effort, automobiles could be converted rather easily to use water as the main fuel source.

The Thermodynamic Device (“TD”) is made up of three principal components:
Component 1: An electrical function generator which energizes a water cell.
Component 2: The Thermodynamic Device
Component 3: A weak electrolyte.

Component 1: The Electrical Function Generator:

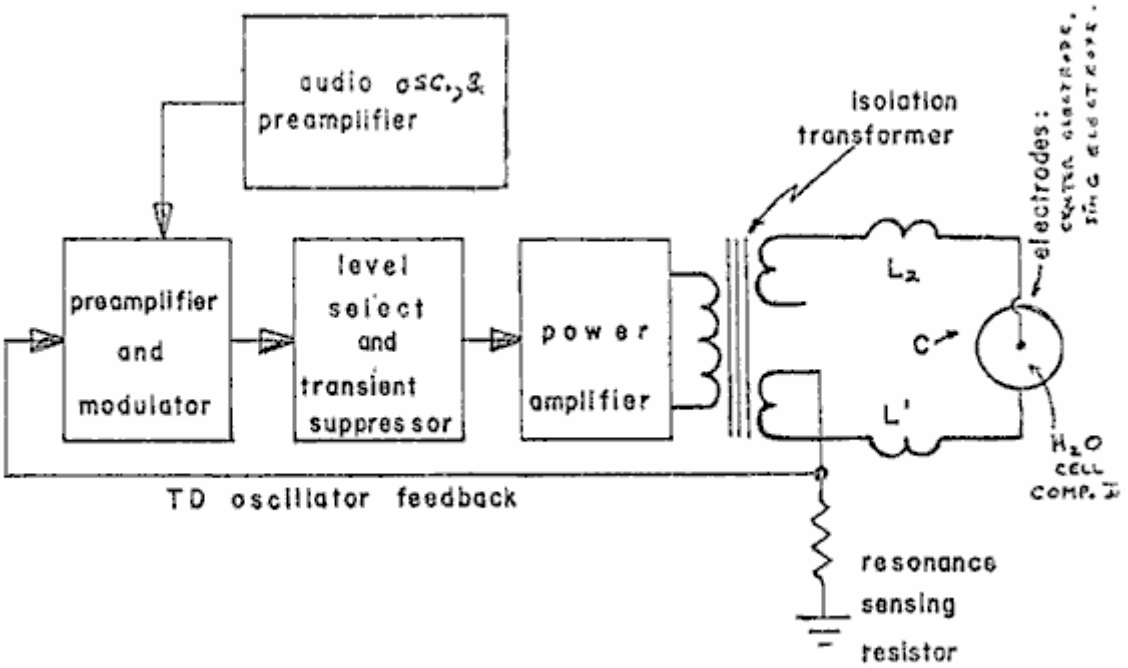


Figure 1: Signal Generator Component Block

This electronic device has a complex alternating current output consisting of an audio frequency (range 20 to 200 Hz) amplitude modulation of a carrier wave (range: 200 to 100,000 Hz). The output is connected by two wires to Component II at the center electrode, and at the ring electrode. See Fig.1. The impedance of this output signal is continuously being matched to the load which is the water solution in Component II.

Component 2: The Thermodynamic Device:

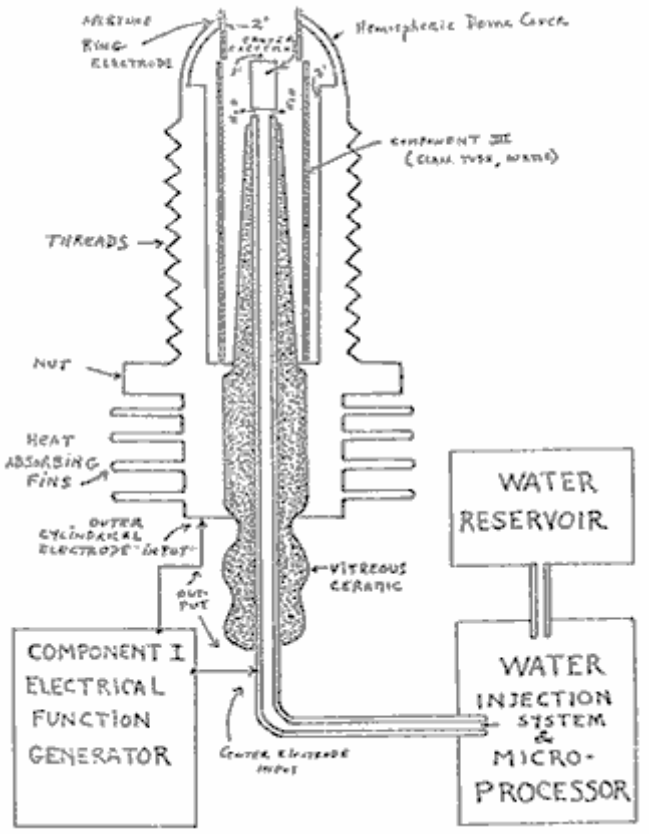


Figure 2: Thermodynamic Device

The TD is fabricated of metals and ceramic in the geometric form of a coaxial cylinder made up of a centered hollow tubular electrode which is surrounded by a larger tubular steel cylinder. These two electrodes comprise the coaxial electrode system energised by Component I. The space between the two electrodes is, properly speaking, Component III which contains the water solution to be electrolysed. The center hollow tubular electrode carries water into the cell, and is further separated from the outer cylindrical electrode by a porous ceramic vitreous material. The space between the two electrodes contains two lengths of tubular Pyrex glass, shown in Figures 2 and 3. The metal electrode surface in contact with the water solution are coated with a nickel alloy.

Component 3: The weak electrolyte water solution:

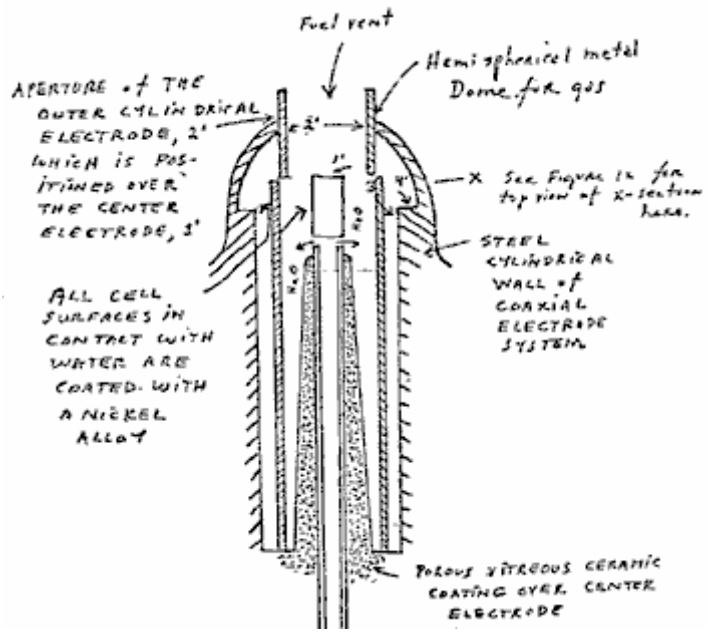


Figure 3: The Water Cell Section of Component 2

This consists of the water solution, the two glass tubes, and the geometry of the containing wall of Component 2. It is the true load for Component 1, and its electrode of Component 2.

The Component 3 water solution is more properly speaking, ideally a 0.1540 M Sodium Chloride solution, and as such, it is a weak electrolyte. In Figure 4 we show the hypothetical tetrahedral structure of water molecule, probably in the form in which the complex electromagnetic waves of Component 1 to see it. The center of mass of this tetrahedral form is the oxygen atom. The geometric arrangement of the p electrons of oxygen probably determine the vectors i (L1) and i (L2) and i (H1) and i (H2) which in turn probably determine the tetrahedral architecture of the water molecule. The p electron configuration of oxygen is shown in Figure 5. Reference to Figure 4, shows that the diagonal of the right side of the cube has at its corner terminations, the positive charge hydrogen (H+) atoms; and that the left side of the cube diagonal has at its corners, the lone pair electrons, (e-). It is to be further noted that this diagonal pair has an orthonormal relationship.

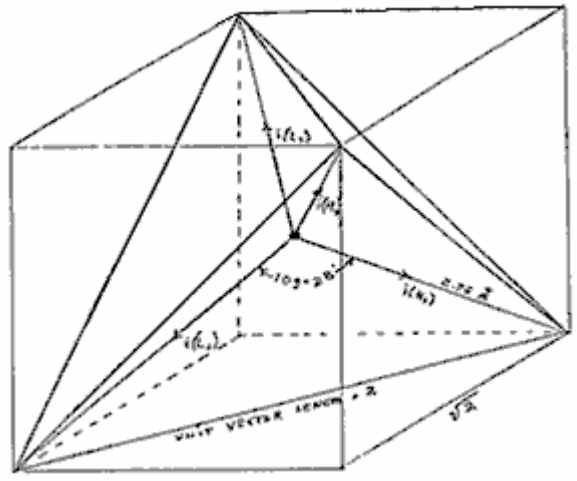


Figure 4: The Water Molecule in Tetrahedral Form:

Hydrogen bonding occurs only along the four vectors pointing to the four vertices of a regular tetrahedron, and in the above drawing we show the four unit vectors along these directions originating from the oxygen atoms at the center. i(H1) and i(H2) are the vectors of the hydrogen bonds formed by the molecule i as a donor molecule. These are assigned to the lone pair electrons. Molecules i are the neighboring oxygen atoms at each vertex of the tetrahedron.

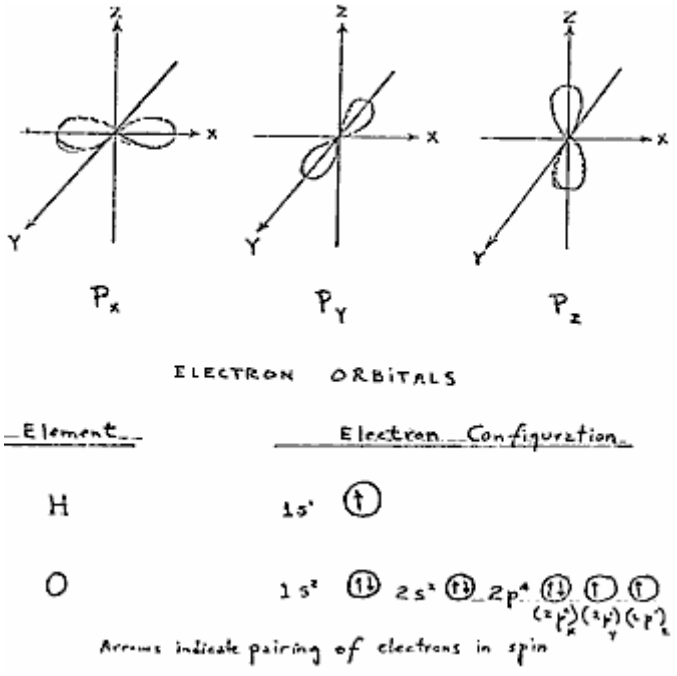


Figure 5: Electron Orbitals

3. Electrothermodynamics

We will now portray the complex electromagnetic wave as the tetrahedral water molecule sees it. The first effect felt by the water molecule is in the protons of the vectors, i (H1) and i (H2). These protons feel the 3-second cycling of the amplitude of the carrier frequency and its associated side bands as generated by Component 1. This sets up a rotation moment of the proton magnetic moment which one can clearly see on the XY plot of an oscilloscope, as an hysteresis loop figure. However, it is noted that this hysteresis loop does not appear in the liquid water sample until all the parameters of the three components have been adjusted to the configuration which is the novel basis of this device. The hysteresis loop gives us a vivid portrayal of the nuclear magnetic relaxation cycle of the proton in water.

The next effect felt by the water molecule is the Component 1 carrier resonant frequency, Fo. At the peak efficiency for electrolysis the value of Fo is 600 Hz +/- 5 Hz.

This resonance however is achieved through control of two other factors. The first is the molal concentration of salt in the water. This is controlled by measuring the conductivity of the water through the built-in current meter of Component 1. There is maintained an idea ratio of current to voltage where I/E = 0.01870 which is an index to the optimum salt concentration of 0.1540 Molal.

The second factor which helps to hold the resonant which helps to hold the resonant frequency at 600 Hz is the gap distance of Y, between the centre electrode, and the ring electrode of Component 2.

This gap distance will vary depending on the size scale of Component 2, but again, the current flow I, is used to set it to the optimal distance when the voltage reads between 2.30 (rms) volts, at resonance Fo, and at molal concentration, 0.1540. The molal concentration of the water is thus seen to represent the electric term of the water molecule and hence its conductivity.

The amplitude modulation of the carrier gives rise to side bands in the power spectrum of the carrier frequency distribution. It is these side bands which give rise to an acoustic vibration of the liquid water, and it is believed, also to the tetrahedral water molecule. The importance of the phonon effect - the acoustic vibration of water in electrolysis - was discovered in a roundabout way. Research work with Component 1 had earlier established that it could be used for the electro-stimulation of hearing in humans. When the output of Component 1 is comprised of flat circular metal plates applied to the head of normal hearing humans, it was found that they could hear pure tones and speech. Simultaneously, acoustic vibration could also be heard by an outside observer with a stethoscope placed near one of the electrodes on the skin. It was observed that the absolute threshold of hearing could be obtained at 0.16 mW (rms), and by calculation that there was an amplitude of displacement of the eardrum of the order of 10^{-11} meter and a corresponding amplitude of the cochlear basilar membrane of 10^{-13}

meter. Corollary to this finding, I was able to achieve the absolute reversible threshold of electrolysis at a power level of 0.16 mW (rms). By carrying out new calculations, I was able to show that the water was being vibrated with a displacement of the order of 1 Angstrom unit (= 10^{-10} meters). This displacement is of the order of the diameter of the hydrogen atom. Thus it is possible that the acoustic phonons generated by audio side bands of the carrier are able to vibrate particle structures within the unit water tetrahedron.

We now turn to the measurement problem with respect to efficiency of electrolysis. There are four means which can be used to measure the reactant product of water electrolysis. For simple volume measurements, one can use a precision nitrometer such as the Pregl type. For both volume and quantitative analysis one can use the gas chromatography with thermal conductivity detector. For a continuous flow analysis of both volume and gas species the mass spectrometer is very useful. For pure thermodynamic measurements the calorimeter is useful. In our measurements, all four methods were examined, and it was found that the mass spectrometer gave the most flexibility and the greatest precision. In the next section we will describe our measurement using the mass spectrometer.

Protocol

4. Methodology for the Evaluation of the Efficiency of Water Decomposition by Means of Alternating Current Electrolysis

Introduction

All systems used today for the electrolysis of water into hydrogen as fuel, and oxygen as oxidant apply direct current to a strong electrolyte solution. These systems range in efficiency from 50% to 71%. The calculation of energy efficiency in electrolysis is defined as follows:

"The energy efficiency is the ratio of the energy released from the electrolysis products formed (when they are subsequently used) to the energy required to effect electrolysis."

The energy released by the exergonic process under standard conditions is

$H_2(g) + (1/2)\ O_2(g) \rightarrow H_2O = 3\ 02.375$ BTU

which is 68.315 Kcal/mol. or, 286,021 Joules/mol, and is numerically equal to the enthalphy charge (ΔH) for the indicated process. On the other hand, the minimum energy (or useful work input) required at constant temperature and pressure for electrolysis equals the Gibbs free energy change (ΔG).

Penner shows that there is a basic relation derivable from the first and second laws of thermodynamics for isothermal changes which shows that

$\Delta G = \Delta H - T\ \Delta S$ ..... (2)

where ΔS represents the entropy change for the chemical reaction and T is the absolute temperature.

The Gibbs free energy change (ΔG) is also related to the voltage (e) required to implement electrolysis by Faraday's equation:

$e = (\Delta G / 23.06\ n)$ volts ..... (3)

where ΔG is in Kcal/mol, and n is the number of electrons (or equivalents) per mole of water electrolysed and has the numerical value 2 in the equation (endergonic process),

$H_2O \rightarrow H_2\ (g) + (1/2)O_2\ (g) + 56.620$ kcal or + 249.68 BTU ..... (4)

Therefore, according to equation (2) at atmospheric pressure, and $300^{O}K$:
$\Delta H = 68.315$ kcal/mol of H_2O, and
$\Delta G = 56.620$ kcal / mol of $H_2O = 236{,}954$ J/mol H_2O for the electrolysis of liquid water.

In view of these thermodynamic parameters for the electrolysis of water into gases, hydrogen and oxygen, we can establish by Eq.(2) numeric values where,

$\Delta G = 236.954$ J/mol H_2O under standard conditions. Thus

$n = \Delta G\ (J/mol) / \Delta Ge\ (J/mol) = <1$ ..... (5)

where ΔGe is the electrical energy input to H_2O (1) in Joules, and ΔG is the Gibbs free energy of H_2O. The conversion between the two quantities is one Watt second (Ws) = one Joule.

Or, in terms of gas volume, as hydrogen, produced and measured,

$n =$ Measured H_2 (cc) / Ideal H_2 (cc) = <1 ..... (6)

In accordance with these general principles we present the methodology followed in evaluating the electrolytic of alternating current on H_2O in producing the gases, hydrogen and oxygen. No attempt has been made to utilize these gases according to the process of Eq.(1). It is to be noted that the process

$H_2\ (g) + (1/2)O_2\ (g) \rightarrow H_2O\ (g)$ ..... (7)

yields only 57.796 kcal /mol. Eq.(7) shows that per mole of gases water formed at 300^OK, the heat released is reduced from the 68.315 kcal/mol at Eq. (1) by the molar heat of evaporation of water at 300^OK (10.5 kcal) and the overall heat release is 57.796 kcal/mol if H_2O (g) is formed at 300^OK.

In the following sections we describe the new method of electrolysis by means of alternating current, and the exact method and means used to measure the endergonic process of Eq.(4) and the governing Eq.(2) and Eq.(5).

5. Thermodynamic Measurement

In order to properly couple Component 2 to a mass spectrometer, one requires a special housing around Component 2 which will capture the gases produced, and permit these to be drawn under low vacuum into the mass spectrometer. Therefore a stainless steel and glass chamber was built to contain Component 2, and provision made to couple it directly through a CO_2 water-trap to the mass spectrometer with the appropriate stainless steel tubing. This chamber is designated as Component 4. Both the mass spectrometer and Component 4 were purged with helium and evacuated for a two hour period before any gas samples were drawn. In this way, contamination was minimized. The definitive measurement were done at Gollob Analytical Services in Berkeley Heights, New Jersey.

We now describe the use of Component 1 and how its energy output to Component 2 is measured. The energy output of Component 1 is an amplitude-modulated alternating current looking into a highly non-linear load, i.e., the water solution. Component 1 is so designed that at peak load it is in resonance across the system (Components 1, 2, and 3) and the vector diagrams show that the capacitive reactance, and the inductance reactance are almost exactly 180^O out of phase with each other, and so the net power output is reactive (the dissipative power is very small). This design ensures minimum power losses across the entire output system. In the experiments to be described, the entire emphasis is placed on achieving the maximum gas yield (credit) in exchange for the minimum applied electrical energy.

The most precise way to measure the applied energy from Component 1 to Component 2 and Component 3, is to measure the power, P, in watts, W. Ideally this should be done with a precision wattmeter, but since we were interested in following the voltage and current separately, it was decided not to use the watt meter. Separate meters were used to continuously monitor the current and the volts.

This is done by precision measurement of the volts across Component 3 as root mean square (rms) volts; and the current flowing in the system as rms amperes. Precisely calibrated instruments were used to take these two measurements. A typical set of experiments using water in the form of 0.9% saline solution 0.1540 molar to obtain high efficiency hydrolysis gave the following results:

rms Current = I = 25mA to 38 mA (0.025 A to 0.038 A.)

rms Volts = E = 4.0 Volts to 2.6 Volts

The resultant ration between current and voltage is dependent on many factors such as the gap distance between the center and ring electrodes, dielectric properties of the water, conductivity properties of the water, equilibrium states, isothermal conditions, materials used, and even the pressure of clathrates. The above current and voltage values reflect the net effect of various combinations of such parameters. When one takes the product of rms current, and rms volts, one has a measure of the power, P in watts.

P = I x E = 25 mA x 4.0 volts =100 mW (0.1 W)

and P = I x E =38 mA x 2.6 volts = 98.8 mW (0.0988 W)

At these power levels (with load), the resonant frequency of the system is 600 Hz (plus or minus 5 Hz) as measured on a precision frequency counter. The waveform was monitored for harmonic content on an oscilloscope, and the nuclear magnetic relaxation cycle was monitored on an XY plotting oscilloscope in order to maintain the proper hysteresis loop figure. All experiments were run so that the power in watts, applied through Components 1, 2, and 3 ranged between 98.8 mW to 100 mW.

Since by the International System of Units 1971 (ST), one Watt-second (Ws) is exactly equal to one Joule (J), our measurements of efficiency used these two yardsticks (1 Ws = 1J) from the debit side of the measurement.

The energy output of the system is, of course, the two gases, Hydrogen (H_2) and Oxygen, $(1/2)O_2$, and this credit side was measured in two laboratories, on two kinds of calibrated instruments, namely gas chromatography machine, and mass spectrometer machine.

The volume of gases H_2 and $(1/2)O_2$ was measured as produced under standard conditions of temperature and pressure in unit time, i.e., in cubic centimeters per minute (cc/min), as well as the possibility contaminating gases, such as air oxygen, nitrogen and argon, carbon monoxide, carbon dioxide, water vapor, etc.

The electrical and gas measurements were reduced to the common denominator of Joules of energy so that the efficiency accounting could all be handled in one currency. We now present the averaged results from many experiments. The standard error between different samples, machines, and locations is at +/- 10%, and we only use the mean for all the following calculations.

2. Thermodynamic Efficiency for the Endergonic Decomposition of Liquid Water (Salininized) to Gases Under Standard Atmosphere (754 to 750 mm. Hg) and Standard Isothermal Conditions @ $25^{O}C = 77^{O}F = 298.16^{O}K$, According to the Following Reaction:

$H_20\ (1) \rightarrow H_2(g) + (1/2)O_2(1) + \Delta G = 56.620$ Kcal /mole ..... (10)

As already described, ΔG is the Gibbs function. We convert Kcal to our common currency of Joules by the formula, One Calorie = 4.1868 Joules

$\Delta G = 56.620$ Kcal x 4.1868 J = 236,954/J/mol of H_2O where 1 mole = 18 gr. ..... (11)

ΔGe = the electrical energy required to yield an equivalent amount of energy from H_2O in the form of gases H_2 and $(1/2)O_2$.

To simplify our calculation we wish to find out how much energy is required to produce the 1.0 cc of H_2O as the gases H_2 and $(1/2)O_2$. There are (under standard conditions) 22,400 cc = V of gas in one mole of H_2O. Therefore

$\Delta G / V =$ 236,954 J / 22,400 cc = 10.5783 J/cc. ..... (12)

We now calculate how much electrical energy is required to liberate 1.0 cc of the H_2O gases (where H_2 = 0.666 parts, and $(1/2)O_2$ = 0.333 parts by volume) from liquid water. Since P = 1 Ws = 1 Joule , and V = 1.0 cc of gas = 10.5783 Joules, then

PV = 1 Js x 10.5783 J = 10.5783 Js, or, = 10.5783 Ws ..... (13)

Since our experiments were run at 100 mW (0.1 W) applied to the water sample in Component II, III, for 30 minutes, we wish to calculate the ideal (100% efficient) gas production at this total applied power level. This is,

0.1 Ws x 60 sec x 30 min = 180,00 Joules (for 30 min.). The total gas production at ideal 100% efficiency is 180 J/10.5783 J/cc = 17.01 cc H_2O (g)

We further wish to calculate how much hydrogen is present in the 17.01 cc H_2O (g).

17.01 cc H_2O (g) x 0.666 H_2 (g) = 11.329 cc H_2 (g) ..... (14)

17.01 cc H_2O (g) x 0.333 $(1/2)O_2$ (g) = 5.681 cc $(1/2)O_2$ (g)

Against this ideal standard of efficiency of expected gas production, we must measure the actual amount of gas produced under: (1) Standard conditions as defined above, and (2) 0.1 Ws power applied over 30 minutes. In our experiments, the mean amount of H_2 and (1/2)O_2 produced, as measured on precision calibrated GC, and MS machines in two different laboratories, where SE is +/- 10%, is,

Measured Mean = 10.80 cc H_2 (g)
Measured Mean = 5.40 cc (1/2) cc (1/2)O_2 (g)
Total Mean = 16.20 cc H_2O (g)

The ratio, n, between the ideal yield, and measured yield,

Measured H_2 (g) / Ideal H_2 (g) = 10.80 cc / 11.33 cc = 91.30%

6. Alternative Method for Calculating Efficiency Based on the Faraday Law of Electrochemistry

This method is based on the number of electrons that must be removed, or added to decompose, or form one mole of, a substance of valence one. In water (H_2O), one mole has the following weight:

H = 1.008 gr /mol
H = 1.008 gr /mol
O = 15.999 gr/mol
Thus, 1 mol H_2O = 18.015 gr/mol

For a univalent substance, one gram/mole contains 6.022 x 10-23 electrons = N = Avogadro's Number. If the substance is divalent, trivalent, etc., N is multiplied by the number of the valence. Water is generally considered to be of valence two.

At standard temperature and pressure ("STP") one mole of a substance contains 22.414 cc, where Standard temperature is 273.15^{O}K = 0^{O}C = T . Standard Pressure (one atmosphere) = 760 mm Hg = P.

One Faraday ("F") is 96,485 Coulombs per mole (univalent).

One Coulomb ("C") is defined as:

1 N / 1 F = 6.122 x 10^{23} Electrons / 96,485 C = one C

The flow of one C/second = one Ampere.
One C x one volt = one Joule second (Js).
One Ampere per second @ one volt = one Watt = one Joule.

In alternating current, when amps (I) and Volts (E) are expressed in root mean squares (rms), their product is Power in watts.

P = IE watts (Watts = Amps x Volts).

With these basic definitions we can now calculate efficiency of electrolysis of water by the method of Faraday's electrochemistry.

The two-electron model of water requires 2 moles of electrons for electrolysis (2 x 6.022 x 10^{23}), or two Faraday quantities (2 x 96,485 = 192,970 Coulombs).

The amount of gas produced will be:

H_2 = 22,414 cc /mol at STP
(1/2)O_2 = 11,207 cc / mol at STP
Gases = 33.621 cc / mol H_2O (g)

The number of coulombs required to produce one cc of gases by electrolysis of water:

193,970 C / 33621 C = 5.739567 C per cc gases.

Then, 5,739 C /cc /sec = 5.739 amp/sec/cc. How many cc of total gases will be produced by 1 A/sec?

0.1742291709 cc.

How many cc of total gases will be produced by 1 A/min ?

10.45375 cc/min

What does this represent as the gases H_2 and O_2 ?

$(1/2)O_2$ = 3.136438721 cc/Amp/min.
H_2 = 6.2728 cc/Amp /min.

We can now develop a Table for values of current used in some of our experiments, and disregarding the voltage as is done conventionally.

1. Calculations for 100 mA per minute:
Total Gases = 1.04537 cc/min
H_2 = 0.6968 cc/min
$(1/2)O_2$ = 0.3484 cc/min
30 min. H_2 = 20.9054 cc/ 30 minutes

2. Calculations for 38 mA per minute:
Total Gases = 0.3972 cc/ 30 minutes
H_2 = 0.2645 cc/min
$(1/2)O_2$ = 0.1323 cc/min
30 min. H_2 = 7.9369 cc/min

3. Calculations for 25mA per minute:
30 min. H_2 = 5.2263 cc/ minute

7. Conclusion

Fig.6 and Fig.7 [not available] show two of the many energy production systems that may be configured to include renewable sources and the present electrolysis technique. Figure 6 shows a proposed photovoltaic powered system using a fuel cell as the primary battery. Assuming optimum operating conditions using 0.25 watt seconds of energy from the photovoltaic array would enable 0.15 watt-seconds to be load.

Figure 7 depicts several renewable sources operating in conjunction with the electrolysis device to provide motive power for an automobile.

****

US Patent 4,394,230 DATE 19th July 1983 INVENTOR: HENRY K. PUHARICH

METHOD AND APPARATUS FOR SPLITTING WATER MOLECULES

This is a re-worded extract from the United States Patent number 4,394,230. It describes how Henry Puharich was able to split water into hydrogen and oxygen gasses by a process which used very little input power.

ABSTRACT

Disclosed herein is a new and improved thermodynamic device to produce hydrogen gas and oxygen gas from ordinary water molecules or from seawater at normal temperatures and pressure. Also disclosed is a new and improved method for electrically treating water molecules to decompose them into hydrogen gas and oxygen gas at efficiency levels ranging between approximately 80-100%. The evolved hydrogen gas may be used as a fuel; and the evolved oxygen gas may be used as an oxidant.

Inventors: Puharich; Henry K. (Rte. 1, Box 97, Delaplane, VA 22025)

BACKGROUND OF THE INVENTION

The scientific community has long realised that water is an enormous natural energy resource, indeed an inexhaustible source, since there are over 300 million cubic miles of water on the earth's surface, all of it a potential source of hydrogen for use as fuel. In fact, more than 100 years ago Jules Verne prophesied that water eventually would be employed as a fuel and that the hydrogen and oxygen which constitute it would furnish an inexhaustible source of heat and light.

Water has been split into its constituent elements of hydrogen and oxygen by electrolytic methods, which have been extremely inefficient, by thermochemical extraction processes called thermochemical water-splitting, which have likewise been inefficient and have also been inordinately expensive, and by other processes including some employing solar energy. In addition, artificial chloroplasts imitating the natural process of photosynthesis have been used to separate hydrogen from water utilising complicated membranes and sophisticated artificial catalysts. However, these artificial chloroplasts have yet to produce hydrogen at an efficient and economical rate.

These and other proposed water splitting techniques are all part of a massive effort by the scientific community to find a plentiful, clean, and inexpensive source of fuel. While none of the methods have yet proved to be commercially feasible, they all share in common the known acceptability of hydrogen gas as a clean fuel, one that can be transmitted easily and economically over long distances and one which when burned forms water.

SUMMARY OF THE PRESENT INVENTION

In classical quantum physical chemistry, the water molecule has two basic bond angles, one angle being 104°, and the other angle being 109°28'. The present invention involves a method by which a water molecule can be energised by electrical means so as to shift the bond angle from the 104° degree. configuration to the 109° degree 28' tetrahedral geometrical configuration.

An electrical function generator (Component 1) is used to produce complex electrical wave form frequencies which are applied to, and match the complex resonant frequencies of the tetrahedral geometrical form of water. It is this complex electrical wave form applied to water which is contained in a special thermodynamic device (Component II) which shatters the water molecule by resonance into its component molecules --- hydrogen and oxygen.

The hydrogen, in gas form, may then be used as fuel; and oxygen, in gas form is used as oxidant. For example, the thermodynamic device of the present invention may be used as a hydrogen fuel source for any existing heat engine --- such as, internal combustion engines of all types, turbines, fuel cell, space heaters, water heaters, heat exchange systems, and other such devices. It can also be used for the desalination of sea water, and other water purification purposes. It can also be applied to the development of new closed cycle heat engines where water goes in as fuel, and water comes out as a clean exhaust.

For a more complete understanding of the present invention and for a greater appreciation of its attendant advantages, reference should be made to the following detailed description taken in conjunction with the accompanying drawings.

DESCRIPTION OF THE DRAWINGS:

Fig.1 is a schematic block diagram illustrating the electrical function generator, Component I, employed in the practice of the present invention:

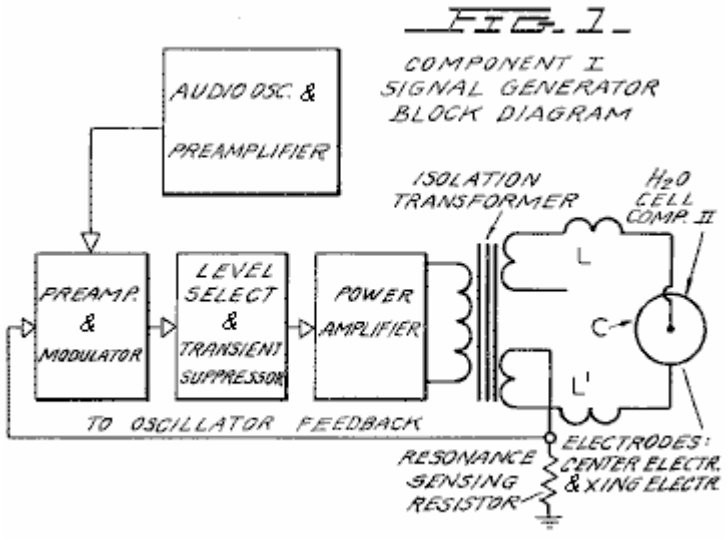


Fig.2 is a schematic illustration of the apparatus of the present invention, including a cross sectional representation of the thermodynamic device, Component II:

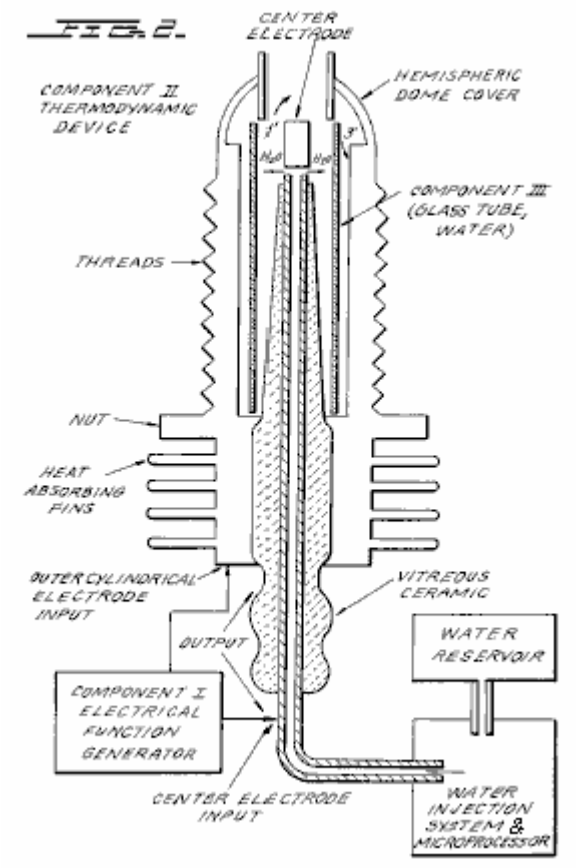


Fig.3 is a cross-sectional view of Component III of the present invention, the water cell section of Component II:

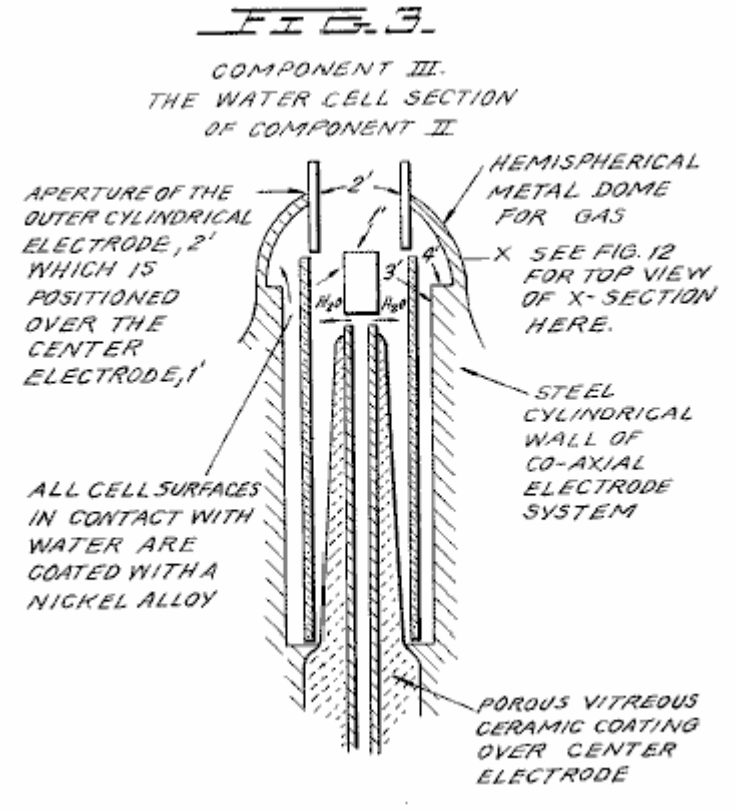


Fig.4 is an illustration of the hydrogen covalent bond:

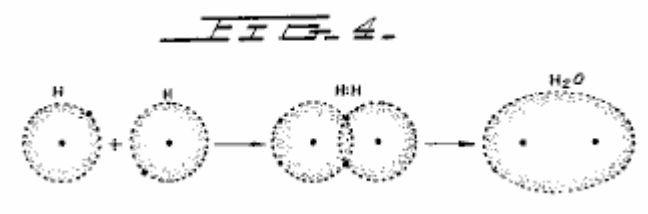


Fig.4A is an illustration of the hydrogen bond angle:

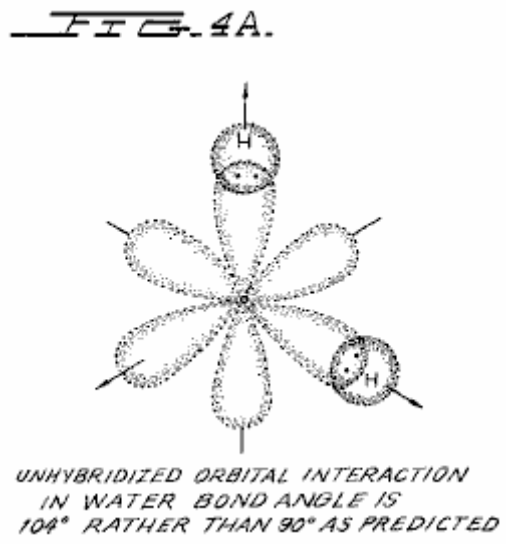


Fig.4B is an illustration of hybridised and un-hybridised orbitals:

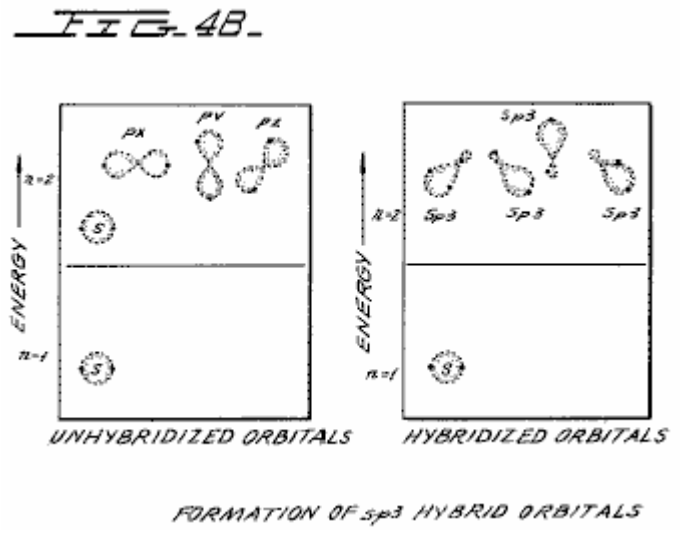


Fig.4C is an illustration of the geometry of methane ammonia and water molecules:

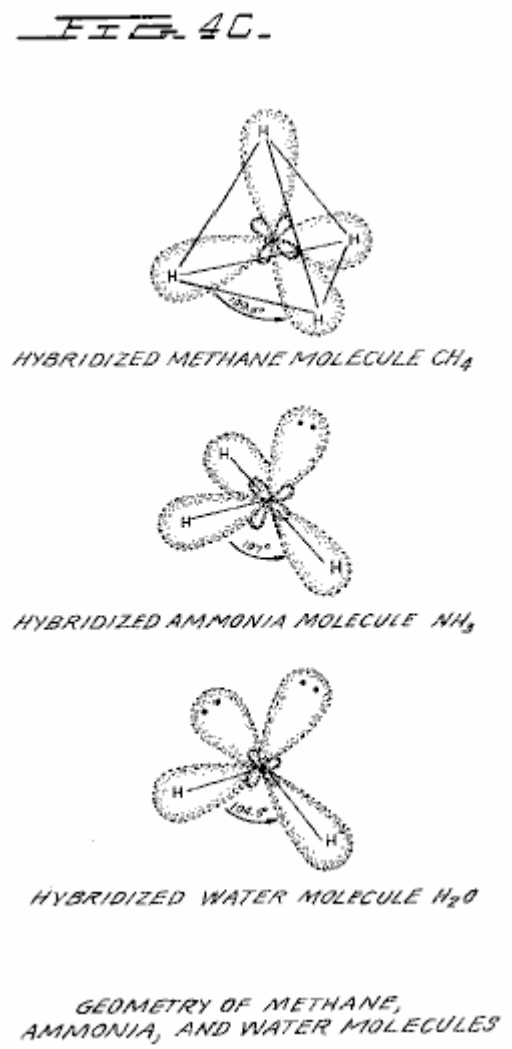


Fig.5 is an illustration of an amplitude modulated carrier wave:

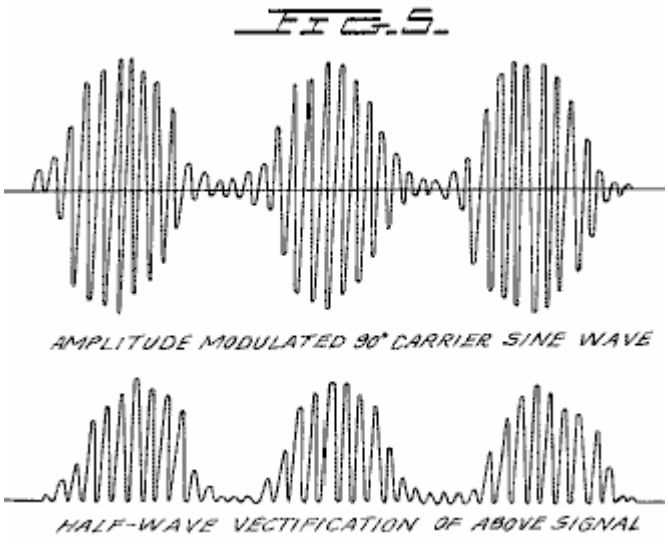


Fig.6 is an illustration of a ripple square wave:

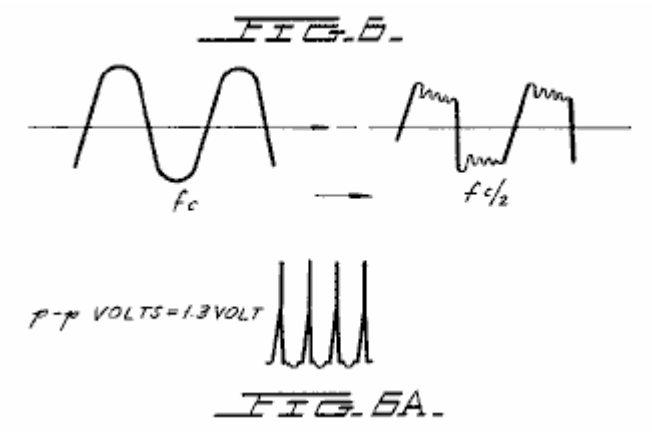


Fig.6A is an illustration of unipolar pulses.

Fig.7 is a diagram showing ion distribution at the negative electrode:

FIG. 7.

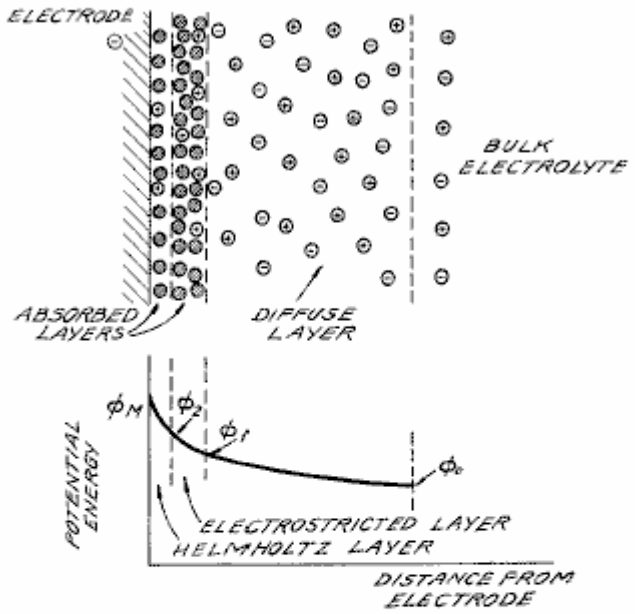


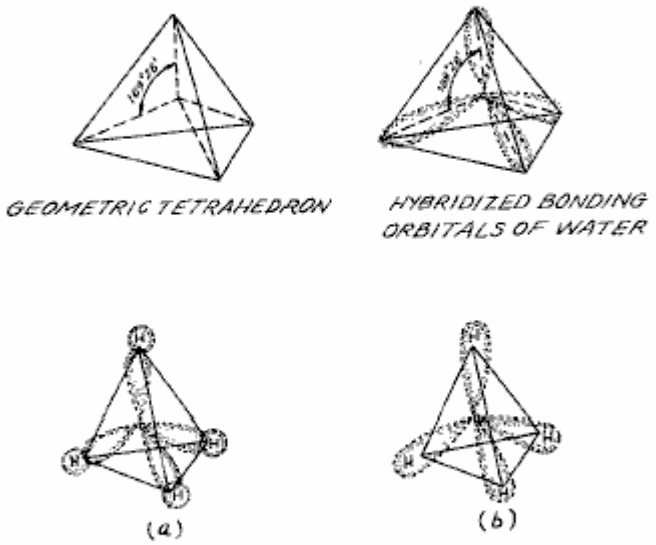
DIAGRAM OF THE DOUBLE LAYER CLOSE TO A NEGATIVE ELECTRODE. THE POTENTIAL ENERGY OF POSITIVE IONS IN THIS REGION WHEN NO CURRENT IS FLOWING IS SHOWN IN THE LOWER DIAGRAM. $\phi_M - \phi_2$ IS THE ELECTRON TRANSFER POTENTIAL: $\phi_2 - \phi_1$ IS RELATED TO THE ACTIVATION OVERPOTENTIAL: AND $\phi_1 - \phi_0$ IS RELATED TO THE DIFFUSION OVERPOTENTIAL.

KEY
- SOLVENT MOLECULE = H_2O
- POSITIVE ION = H^+
- NEGATIVE ION = O^-

Fig.8 is an illustration of tetrahedral bonding orbitals:

FIG.8.

EQUIVALENT TETRAHEDRAL BONDING ORBITALS OF WATER



METHANE OVERLAP OF SPHERICAL 1s ORBITAL OF HYDROGEN WITH sp3 BONDING ORBITALS OF CARBON (a) RESULTS IN EQUIVALENT SIGMA BONDS, THE MOLECULAR ORBITALS OF (b).

Fig.9 is an illustration of water molecules:

FIG.9.

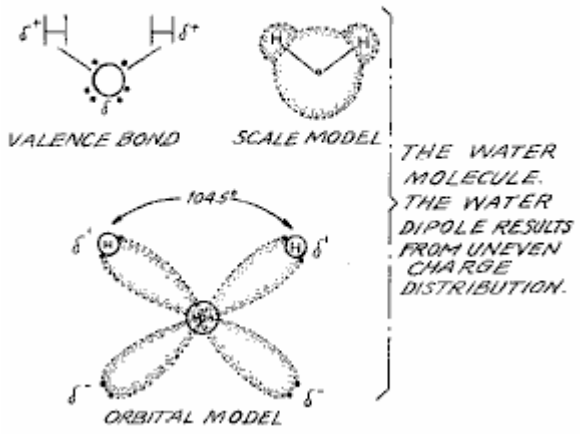


Fig.10 is an illustration of productive and non-productive collisions of hydrogen with iodine:

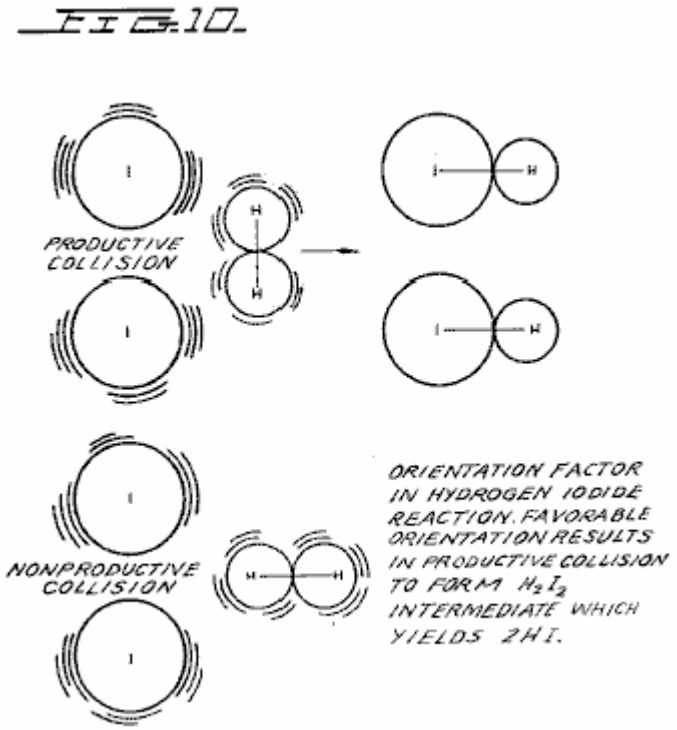


Fig.11 is a wave form found to be the prime characteristic for optimum efficiency:

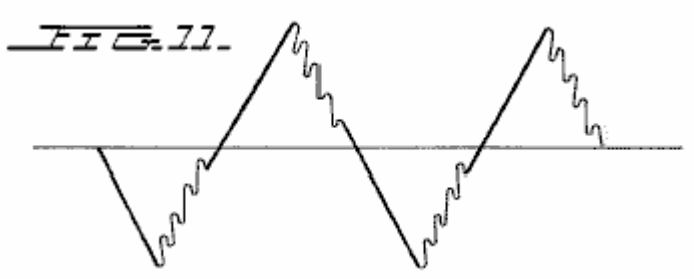


Fig.12 is an illustration of pearl chain formation:

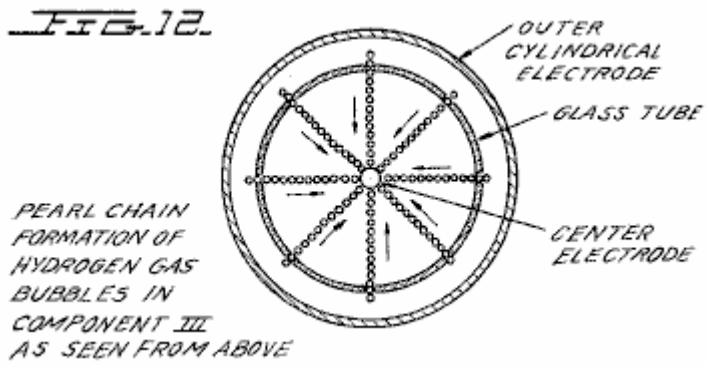
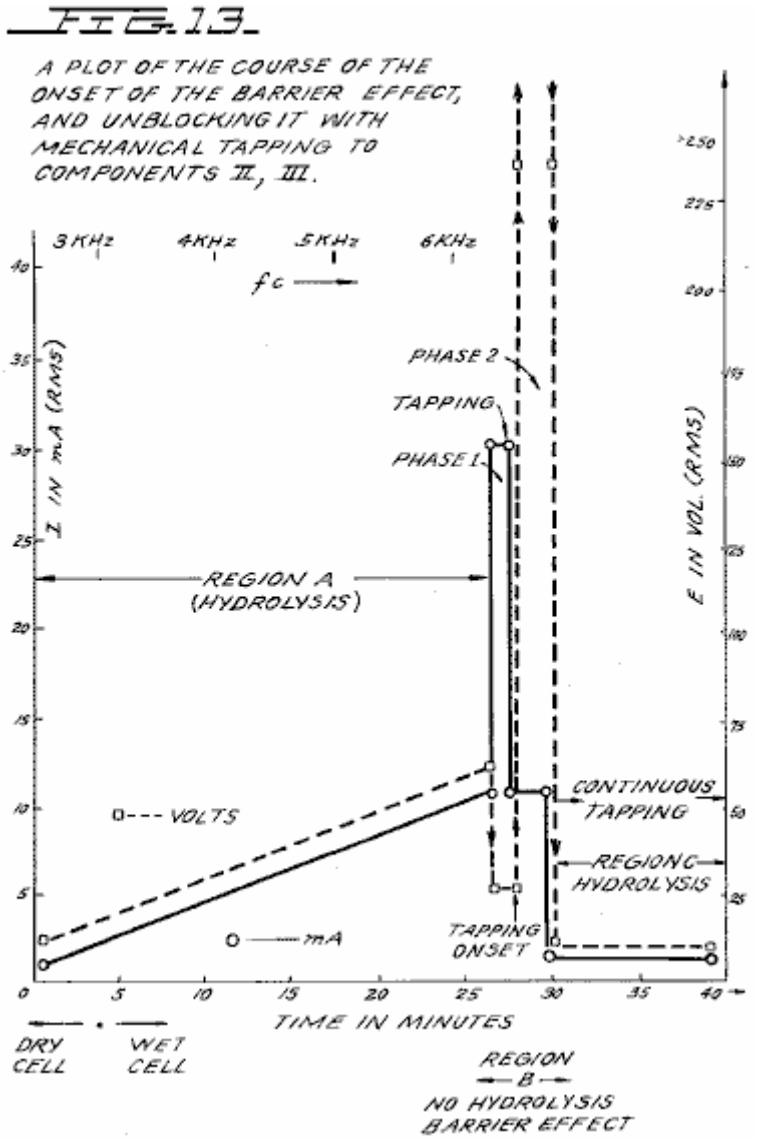


Fig.13 is a plot of the course of the onset of the barrier effect and the unblocking of the barrier effect:



Figs.14A, B, and C are energy diagrams for exergonic reactions:

FIG. 14A.

(a) AN EXERGONIC REACTION. PRODUCTS HAVE A LOWER POTENTIAL ENERGY THAN REACTANTS, THEREFORE, ENERGY IS RELEASED.

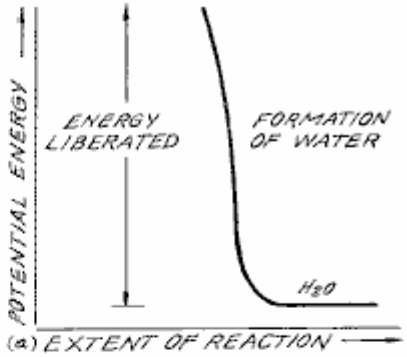


FIG. 14B.

(b) AN ENDERGONIC REACTION PRODUCTS HAVE A HIGHER POTENTIAL ENERGY THAN REACTANTS, CAUSING ENERGY TO BE CONSUMED.

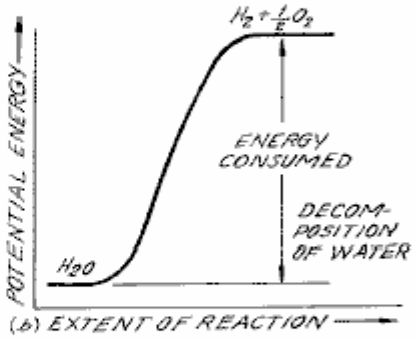
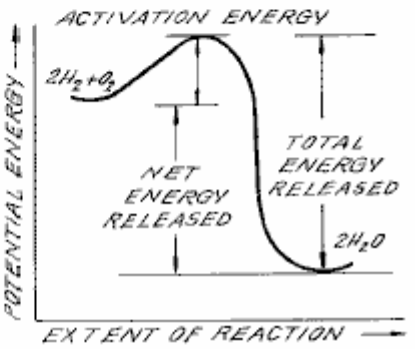


FIG. 14C.

ENERGY DIAGRAM FOR EXERGONIC REACTION. ACTIVATION ENERGY IS BARRIER TO BE OVERCOME FOR REACTION TO PROCEED, ⚡ IS SUPPLIED AS A "SPARK" TO THE GASES TO GET IGNITION.



DETAILED DESCRIPTION OF INVENTION:

Section 1:
Apparatus of Invention;
The apparatus of the invention consists of three components, the Electrical Function Generator, the Thermodynamic Device, and the Water Cell.

Component I: The Electrical Function Generator;
This device has an output consisting of an audio frequency (range 20 to 200 Hz) amplitude modulation of a carrier wave (range 200 Hz to 100,000 Hz). The impedance of this output signal is continuously being matched to the load which is the second component, the thermodynamic device. The electrical function generator represents a novel application of circuitry disclosed in my earlier U.S. Pat. Nos. 3,629,521; 3,563,246; and 3,726,762, which are incorporated by reference herein. See **Fig.1** for the block diagram of Component I.

Component II: The Thermodynamic Device;
The thermodynamic device is fabricated of metals and ceramic in the geometric form of coaxial cylinder made up of a central hollow tubular electrode which is surrounded by a larger tubular steel cylinder, said two electrodes comprising the coaxial electrode system which forms the load of the output of the electrical function generator, Component I. Said central hollow tubular electrode carries water, and is separated from the outer cylindrical electrode by a porous ceramic vitreous material. Between the outer surface of the insulating ceramic vitreous material, and the inner surface of the outer cylindrical electrode exists a space to contain the water to be electrolysed. This water cell space comprises the third component (Component III) of the invention. It contains two lengths of tubular Pyrex glass, shown in **Fig.2** and **Fig.3**. The metal electrode surfaces of the two electrodes which are in contact with the water are coated with a nickel alloy.

The coaxial electrode system is specifically designed in materials and geometry to energise the water molecule to the end that it might be electrolysed. The central electrode is a hollow tube and also serves as a conductor of water to the Component III cell. The central tubular electrode is coated with a nickel alloy, and surrounded with a porous vitreous ceramic and a glass tube with the exception of the tip that faces the second electrode. The outer cylindrical electrode is made of a heat conducting steel alloy with fins on the outside, and coated on the inside with a nickel alloy. The central electrode, and the cylindrical electrode are electrically connected by an arching dome extension of the outer electrode which brings the two electrodes at one point to a critical gap distance which is determined by the known quenching distance for hydrogen. See **Fig.2** for an illustration of Component II.

Component III: The Water Cell;
The water cell is a part of the upper end of Component II, and has been described. An enlarged schematic illustration of the cell is presented in FIG. 3. The Component III consists of the water and glass tubes contained in the geometrical form of the walls of cell in Component II, the thermodynamic device. The elements of a practical device for the practice of the invention will include:
(A) Water reservoir; and salt reservoir; and/or salt

(B) Water injection system with microprocessor or other controls which sense and regulate (in accordance with the parameters set forth here:
a. Carrier frequency
b. Current
c. Voltage
d. RC relaxation time constant of water in the cell
e. Nuclear magnetic relaxation constant of water
f. Temperature of hydrogen combustion
g. Carrier wave form
h. RPM of an internal combustion engine (if used)
i. Ignition control system
j. Temperature of region to be heated;

(C) An electrical ignition system to ignite the evolved hydrogen gas fuel.
The important aspects of Component III are the tubular vitreous material, the geometry of the containing walls of the cell, and the geometrical forms of the water molecules that are contained in the cell. A further important aspect of the invention is the manipulation of the tetrahedral geometry of the water molecule by the novel methods and means which will be more fully described in the succeeding sections of this specification.

The different parts of a molecule are bound together by electrons. One of the electron configurations which can exist is the covalent bond which is achieved by the sharing of electrons. A molecule of hydrogen gas, H_2 is the smallest representative unit of covalent bonding, as can be seen in **Fig.4**. The molecule of hydrogen gas is formed by the overlap and pairing of 1s orbital electrons. A new molecular orbit is formed in which the shared electron pair orbits both nuclei as shown in **Fig.4**. The attraction of the nuclei for the shared electrons holds the atoms together in a covalent bond.

Covalent bonds have direction. The electronic orbitals of an uncombined atom can change shape and direction when that atom becomes part of a molecule. In a molecule in which two or more covalent bonds are present the molecular geometry is dictated by the bond angles about the central atom. The outermost lone pair (non-bonding) electrons profoundly affect the molecular geometry.

The geometry of water illustrates this concept. In the ground state, oxygen has the outer shell configuration:
$$\mathbf{1s^2\ 2s^2\ 2p^2_x\ 2p^1_y\ 2p^1_z}$$
In water the 1s electrons from two hydrogen atoms bond with the $2p_y$ and $2p_z$ electrons of oxygen. Since p orbitals lie at right angles to each other (see **Fig.4A**), a bond angle of 90° might be expected. However, the bond angle is found experimentally to be approximately 104°. Theoretically this is explained by the effect of lone pair electrons on hybridised orbitals.

Combined or hybrid orbitals are formed when the excitement of 2s electrons results in their promotion from the ground state to a state energetically equivalent to the 2p orbitals. The new hybrids are termed sp^3 from the combination of one s and three p orbitals (See **Fig.4B**). Hybrid sp^3 orbitals are directed in space from the centre of a regular tetrahedron toward the four corners. If the orbitals are equivalent the bond angle will be 109°28' (See **Fig.15**) consistent with the geometry of a tetrahedron. In the case of water two of the orbitals are occupied by non-bonding electrons (See **Fig.4C**). There is greater repulsion of these lone pair electrons which orbit only one nucleus, compared to the repulsion of electrons in bonding orbitals which orbit two nuclei. This tends to increase the angle between non-bonding orbitals so that it is greater than 109°, which pushes the bonding orbitals together, reducing the bond angle to 104°. In the case of ammonia, NH_3 where there is only one lone pair, the repulsion is

not so great and the bond angle is 107°. Carbon forms typical tetrahedral forms and components the simplest being the gas methane, CH_4 (See **Fig.4C** and **Fig.8**). The repulsion of lone pair electrons affects charge distribution and contributes to the polarity of a covalent bond. (See **Fig.16**)

As demonstrated in succeeding sections of this patent specification, a significant and novel aspect of this invention is the manipulation, by electronic methods and means, of the energy level of the water molecule, and the transformation of the water molecule into, and out of, the geometrical form of the tetrahedron. This is made possible only by certain subtle dynamic interactions among the Components I, II, and III of the present invention.

Section 2:
Electrodynamics (Pure Water);
The electrodynamics of Components I, II, and III, will be described individually and in interaction during the progress of pure water reaction rate in time. The reactions of saline water will be described in Section 3. It is to be noted that the output of Component I automatically follows the seven stages (hereinafter Stages A-F) of the reaction rate by varying its parameters of resonant carrier frequency, wave form, current voltage and impedance. All the seven states of the reaction herein described are not necessary for the practical operation of the system, but are included in order to explicate the dynamics and novel aspects of the invention. The seven stages are applicable only to the electrolysis of pure water.

Stage A:
Dry Charging of Component II by Component I;
To make the new system operational, the Component I output electrodes are connected to component II, but no water is placed in the cell of Component III. When Component I output is across the load of Component II we observe the following electrical parameters are observed:

Range of current (I) output with (dry) load: 0 to 25 mA (milliamps) rms.
Range of voltage (E) output with (dry) load: 0 to 250 Volts (AC) rms.
There is no distortion of the amplitude modulated (AM), or of the sine wave carrier whose central frequency, f_c' ranges between 59,748 Hz to 66, 221 Hz, with f_c average = 62, 985 Hz

The carrier frequency varies with the power output in that f_c goes down with an increase in amperes (current). The AM wave form is shown in **Fig.5**. It is to be noted here that the electrical function generator, Component I, has an automatic amplitude modulation volume control which cycles the degree of Amplitude Modulation from 0% to 100%, and then from 100% to 0% every 3.0 seconds. This cycle rate of 3.0 seconds corresponds to the nuclear spin relaxation time, tau/sec, of the water in Component III. The meaning of this effect will be discussed in greater detail in a later section.

In summary, the principal effects to be noted during Stage A -dry charging of Component II are as follows:

a. Tests the integrity of Component I circuitry.
b. Tests the integrity of the coaxial electrodes, and the vitreous ceramic materials of Component II and Component III.
c. Electrostatic cleaning of electrode and ceramic surfaces.

Stage B:
Initial operation of Component I, Component II, and with Component III containing pure water. There is no significant electrolysis of water during Stage B. However, in Stage B the sine wave output of Component I is shaped to a rippled square wave by the changing RC constant of the water as it is treated;

There is an `Open Circuit` reversible threshold effect that occurs in Component III due to water polarisation effects that lead to half wave rectification and the appearance of positive unipolar pulses; and
There are electrode polarisation effects in Component II which are a prelude to true electrolysis of water as evidenced by oxygen and hydrogen gas bubble formation.

Appearance of Rippled Square Waves:
Phase 1: At the end of the Stage A dry charging, the output of Component I is lowered to typical values of: I = 1 ma. E = 24V AC. f_c .congruent.66,234 Hz.

Phase 2: Then water is added to the Component III water cell drop by drop until the top of the centre electrode, 1', in **Fig.3** is covered, and when this water just makes contact with the inner surface of the top outer electrode at 2'. As this coupling of the two electrodes by water happens, the following series of events occur:

Phase 3: The f_c drops from 66,234 Hz, to a range from 1272 Hz to 1848 Hz. The current and voltage both drop, and begin to pulse in entrainment with the water nuclear spin relaxation constant, tau =3.0 sec. The presence of

the nuclear spin relaxation oscillation is proven by a characteristic hysteresis loop on the X-Y axes of an oscilloscope.
I = 0 to 0.2 mA surging at .tau. cycle
E = 4.3 to 4.8V AC surging at .tau. cycle
The sine wave carrier converts to a rippled square wave pulse which reflects the RC time constant of water, and it is observed that the square wave contains higher order harmonics. See **Fig.6**:
With the appearance of the rippled square wave, the threshold of hydrolysis may be detected (just barely) as a vapour precipitation on a cover glass slip placed over the Component III cell and viewed under a low power microscope.

The 'Open Circuit' Reversible Threshold Effect:
Phase 4 A secondary effect of the change in the RC constant of water on the wave form shows up as a full half wave rectification of the carrier wave indicating a high level of polarisation of the water molecule in tetrahedral form at the outer electrode.

With the already noted appearance of the rippled square wave, and the signs of faint vapour precipitation which indicate the earliest stage of electrolysis, it is possible to test for the presence of a reversible hydrolysis threshold. This test is carried out by creating an open circuit between Components I and II, i.e., no current flows. This is done by lowering the water level between the two electrodes in the region --- 1' and 2' shown in **Fig.3**; or by interrupting the circuit between Component I and II, while the Component I signal generator is on and oscillating.

Immediately, with the creation of an `open circuit` condition, the following effects occur:
(a) The carrier frequency, f_c, shifts from Phase 4 valve 1272 Hz to 1848 Hz to 6128 Hz.
(b) The current and voltage drop to zero on the meters which record I and E, but the oscilloscope continues to show the presence of the peak-to-peak (p-p) voltage, and the waveform shows a remarkable effect. The rippled square wave has disappeared, and in its place there appear unipolar (positive) pulses as follows in **Fig.6A**.

The unipolar pulse frequency stabilises to ca. 5000 Hz. The unipolar pulses undergo a 0 to 1.3 volt pulsing amplitude modulation with .tau. at 3.0 seconds. Thus, there exists a pure open circuit reversible threshold for water electrolysis in which the water molecules are capacitor charging and discharging at their characteristic low frequency RC time constant of 0.0002 seconds. It is to be noted that pure water has a very high dielectric constant which makes such an effect possible.

The pulsing amplitude modulation of the voltage is determined by the Hydrogen Nuclear Spin Relaxation constant of 3.0 seconds. It is to be noted that the positive pulse spikes are followed by a negative after-potential. These pulse wave forms are identical to the classic nerve action potential spikes found in the nervous system of all of the living species which have a nervous system. The fact that these unipolar pulses were observed arising in water under the conditions of reversible threshold hydrolysis has a profound significance. These findings illuminate and confirm the Warren McCulloch Theory of water "crystal" dynamics as being the foundation of neural dynamics; and the converse theory of Linus Pauling which holds that water clathrate formation is the mechanism of neural anesthesia.

Phase 5: The effects associated with reversible threshold electrolysis are noted only in passing, since they reflect events which are occurring on the electrode surfaces of Component II, the Thermodynamic Device.

A principal effect which occurs in Stage B, Phase 3, in Component II, (the thermodynamic device), is that the two electrodes undergo stages of polarisation. It has been observed in extensive experiments with different kinds of fluids in the cell of Component II , i.e., distilled water, sea water, tap water, Ringers solution, dilute suspensions of animal and human blood cells, etc. that the inner surface of the outer ring electrode at 3' in **Fig.3** (the electrode that is in contact with the fluid) becomes negatively charged. Referring to **Fig.7**, this corresponds to the left hand columnar area marked, "Electrode .crclbar.".

Electrode Polarisation Effects at the Interface Between Components II and III:
Concurrently with the driver pulsing of Component I at the .tau. constant cycle which leads to electrode polarisation effects in Component II, there is an action on Component III which energises and entrains the water molecule to a higher energy level which shifts the bond angle from 104° to the tetrahedral form with angle 109°28' as shown in **Fig.8** and **Fig.15**.

This electronic pumping action is most important, and represents a significant part of the novel method of this invention for several reasons. First, the shift to the tetrahedral form of water increases the structural stability of the water molecule, thereby making it more susceptible to breakage at the correct resonant frequency, or frequencies. Second, increasing the polarisation of the water molecule makes the lone pair electrons, S- connected with the oxygen molecule more electronegative; and the weakly positive hydrogen atoms, S+ more positive. See **Fig.9** and **Fig.22**.

As the outer electrode becomes more electrically negative, the central electrode becomes more electrically positive as will be shown. As the polarity of the water molecule tetrahedron increases, a repulsive force occurs between the two S+ apices of the water tetrahedron and the negatively charged electrode surface within the region of the Helmholtz layer, as shown in **Fig.7**. This effect "orients" the water molecule in the field, and is the well-known "orientation factor" of electrochemistry which serves to catalyse the rate of oxygen dissociation from the water molecule, and thereby causes the reaction rate to proceed at the lowest energy levels. See **Fig.10** for an example of how the orientation factor works. Near the end of Stage B, the conditions are established for the beginning of the next stage, the stage of high efficiency electrolysis of water.

Stage C:
Generation of the complex wave form frequencies from Component I to match the complex wave form resonant frequencies of the energised and highly polarised water molecule in tetrahedral form with angles, 109°28' are carried out in Stage C. In the operation of the invention active bubble electrolysis of water is initiated following Stage B, phase 3 by setting (automatically) the output of Component I to:
I = 1 mA., E = 22V AC-rms, causing the rippled square wave pulses to disappear with the appearance of a rippled sawtooth wave. The basic frequency of the carrier now becomes, f_c = 3980 Hz.

The wave form now automatically shifts to a form found to be the prime characteristic necessary for optimum efficiency in the electrolysis of water and illustrated in **Fig.11**. In the wave form of **Fig.11**, the fundamental carrier frequency, f_c = 3980 Hz., and a harmonic modulation of the carrier is as follows:

1st Order Harmonic Modulation (OHM) = 7960 Hz.
2nd Order Harmonic Modulation (II OHM) = 15,920 Hz.
3rd Order Harmonic Modulation (III OHM) = 31,840 Hz.
4th Order Harmonic Modulation (IV OHM) = 63,690 Hz.

What is believed to be happening in this IV OHM effect is that each of the four apices of the tetrahedron water molecule is resonant to one of the four harmonics observed. It is believed that the combination of negative repulsive forces at the outer electrode with the resonant frequencies just described work together to shatter the water molecule into its component hydrogen and oxygen atoms (as gases). This deduction is based on the following observations of the process through a low power microscope. The hydrogen bubbles were seen to originate at the electrode rim, 4', of **Fig.3**. The bubbles then moved in a very orderly `pearl chain` formation centripetally (like the spokes of a wheel) toward the central electrode, 1' of **Fig.3**, (**Fig.12** shows a top view of this effect).

Thereafter, upon lowering the output of Component I, the threshold for electrolysis of water as evidenced by vapour deposition of water droplets on a glass cover plate over the cell of Component III, is:

I = 1 mA, E = 10V so, Power = 10 mW

with all other conditions and waveforms as described under Stage C, supra. Occasionally, this threshold can be lowered to:

I = 1 ma, E = 2.6V so, Power = 2.6 mW

This Stage C vapour hydrolysis threshold effect cannot be directly observed as taking place in the fluid because no bubbles are formed --- only invisible gas molecules which become visible when they strike a glass plate and combine into water molecules and form droplets which appear as vapour.

Stage D:
Production of hydrogen and oxygen gas at an efficient rate of water electrolysis is slowed in Stage D when a barrier potential is formed, which blocks electrolysis, irrespective of the amount of power applied to Components II and III.

A typical experiment will illustrate the problems of barrier potential formation. Components I, II, and III are set to operate with the following parameters:

I = 1 ma, E = 11.2V so, Power = 11.2 mW (at the start, rising to 100 mW later)

This input to Component III yields, by electrolysis of water, approximately 0.1 cm^3 of hydrogen gas per minute at one atmosphere and 289^{O}K. It is observed that as a function of time the f_c crept up from 2978 Hz to 6474 Hz over 27 minutes. The current and the voltage also rose with time. At the 27th minute a barrier effect blocked the electrolysis of water, and one can best appreciate the cycle of events by reference to **Fig.13**.

Stage E:
The Anatomy of the Barrier Effect:
Region A: Shows active and efficient hydrolysis
Region B: The barrier region effect can be initiated with taps of the finger, or it can spontaneously occur as a function of time.
Phase a: The current rose from 1 mA to 30 mA. The voltage fell from 22 volts to 2.5 V.
Phase b: If component II is tapped mechanically during Phase a supra --- it can be reversed as follows: The current dropped from 30 mA to 10 mA. The voltage shot up from 5 volts to over 250 volts (off scale).

Throughout 'Phase a' and 'Phase b', all hydrolysis has ceased. It was observed under the microscope that the inner surface of the outer electrode was thickly covered with hydrogen gas bubbles. It was reasoned that the hydrogen gas bubbles had become trapped in the electrostricted layer, because the water molecule tetrahedrons had flipped so that the S+ hydrogen apices had entered the Helmholtz layer and were absorbed to the electronegative charge of the electrode. This left the S- lone pair apices facing the electrostricted layer. This process bound the newly forming H^+ ions which blocked the reaction
$H^+ + H^+ + 2e \rightarrow H_2$ (gas)

Stage F:
Region C: It was found that the barrier effect could be unblocked by some relatively simple procedures:
(a) Reversing the output electrodes from Component I to Component II, and/or:
(b) Mechanically tapping the Component III cell at a frequency T/2 = 1.5 seconds per tap.
These effects are shown in FIG. 12 and induce the drop in barrier potential from:

I = 10 mA to 1 ma, E = 250V to 4V so, Power fell from 2.5W to 4 mW

Upon unblocking of the barrier effect, electrolysis of water resumed with renewed bubble formation of hydrogen gas.

The barrier potential problem has been solved for practical application by lowering the high dielectric constant of pure water, by adding salts (NaCl, KOH, etc.) to the pure water thereby increasing its conductivity characteristics. For optimum efficiency the salt concentration need not exceed that of sea water (0.9% salinity) in Section 3, "Thermodynamics of the Invention", it is to be understood that all water solutions described are not "pure" water as in Section B, but refer only to saline water.

Section 3:
The Thermodynamics of the Invention (Saline Water);

Introduction: (water, hereinafter refers to saline water);
The thermodynamic considerations in the normal operations of Components I, II, and III in producing hydrogen as fuel, and oxygen as oxidant during the electrolysis of water, and the combustion of the hydrogen fuel to do work in various heat engines is discussed in this section.

In chemical reactions the participating atoms form new bonds resulting in compounds with different electronic configurations. Chemical reactions which release energy are said to be exergonic and result in products whose chemical bonds have a lower energy content than the reactants. The energy released most frequently appears as heat. Energy, like matter, can neither be created nor destroyed according to the Law of Conservation of Energy. The energy released in a chemical reaction, plus the lower energy state of the products, is equal to the original energy content of the reactants. The burning of hydrogen occurs rather violently to produce water as follows:

$2H_2 + O_2 \rightarrow 2H_2O - \Delta H$ 68.315 Kcal/mol (this is the enthalpy, or heat of combustion at constant pressure) where 18 gms = 1 mol.

The chemical bonds of the water molecules have a lower energy content than the hydrogen and oxygen gases which serve at the reactants. Low energy molecules are characterised by their stability. High energy molecules are inherently unstable. These relations are summarised in the two graphs of **Fig.14**. It is to be noted that **Fig.14B** shows the endergonic reaction aspect of the invention when water is decomposed by electrolysis into hydrogen and oxygen.

Fig.14A shows the reaction when the hydrogen and oxygen gases combine, liberate energy, and re-form into water. Note that there is a difference in the potential energy of the two reactions. **Fig.14C** shows that there are two components to this potential energy. The net energy released, or the energy that yields net work is labelled in the diagram as "Net Energy Released", and is more properly called the free energy change denoted by the Gibbs function, $-.\Delta.G$.

The energy which must be supplied for a reaction to achieve (burning) spontaneity is called the “Activation Energy”. The sum of the two is the total energy released. A first thermodynamic subtlety of the thermodynamic device of the invention is noted in Angus McDougall's Fuel Cells, Energy Alternative Series, The MacMillan Press Ltd., London, 1976, where on page 15 it is stated:

"The Gibbs function is defined in terms of the enthalpy H, and the entropy S of the system:

G = H-T S (where .tau. is the thermodynamic temperature) . A particularly important result is that for an electrochemical cell working reversibly at constant temperature and pressure, the electrical work done is the net work and hence,

. $\Delta G = -w_e$

For this to be a reversible process, it is necessary for the cell to be on `open circuit`, that is, no current flows and the potential difference across the electrodes is the EMF, E. Thus,

. $\Delta G = -zFE$

(where F is the Faraday constant --- the product of the Avogadro Constant + $N_A = 6.022045 \times 10^{23}$ mole^{-1}, and the charge on the electron, $e = 1.602\ 189 \times 10^{-19}$ C --- both in SI units; and z is the number of electrons transported.) when the cell reaction proceeds from left to right."

It is to be noted that the Activation Energy is directly related to the controlling reaction rate process, and thus is related to the Gibbs free energy changes. The other thermodynamic subtlety is described by S. S. Penner in his work: Penner, S. S. and L. Icerman, Energy, Vol, II, Non-Nuclear Energy Technologies. Addison-Wesley Publishing Company, Inc. Revised Edition, 1977. Reading, Mass. where on page 140 it is stated that:

"It should be possible to improve the efficiency achieved in practical electrolysis to about 100% because, under optimal operating conditions, the theoretically-attainable energy conversion by electrolysis is about 120% of the electrical energy input. The physical basis for this last statement will now be considered:

"A useful definition for energy efficiency in electrolysis is the following: the energy efficiency is the ratio of the energy released from the electrolysis products formed (when they are subsequently used) to the energy required to effect electrolysis. The energy released by the process

H_2 (gas) + (1/2)O_2 (gas) $\rightarrow$ H_2O (liquid)

under standard conditions (standard conditions in this example are: (1) atmospheric pressure = 760 mm Hg and (2) temperature = 298.16^{O}K. = 25^{O}C. = 77^{O}F.) is 68.315 Kcal and is numerically equal to the enthalph change (.ΔH) for the indicated process. On the other hand, the minimum energy (or useful work input) required at constant temperature and pressure for electrolysis equals the Gibbs free energy change (.ΔG). There is a basic relation derivable from the first and second laws of thermodynamics for isothermal changes, which shows that:

.$\Delta G = . \Delta H - T. \Delta S$

where .ΔS represents the entropy change for the chemical reaction. The Gibbs free energy change (.ΔG) is also related to the voltage (E) required to implement electrolysis by Faraday's equation, viz.

E = (.ΔG/23.06n) volts

where . ΔG is in Kcal/mol and n is the number of electrons (or equivalents) per mol of water electrolysed and has the numerical value 2.

At atmospheric pressure and 300^{O}K., .ΔH = 68.315 Kcal/mol of H_2O (i) and .ΔG = 56.62 Kcal/mole of H_2O (i) for the electrolysis of liquid water. Hence, the energy efficiency of electrolysis at 300^{O}K. is about 120%.

(When) H_2 (gas) and O_2 (gas) are generated by electrolysis, the electrolysis cell must absorb heat from the surroundings, in order to remain at constant temperature. It is this ability to produce gaseous electrolysis products with heat absorption from the surroundings that is ultimately responsible for energy-conversion efficiencies during electrolysis greater than unity."

Using the criteria of these two authorities, it is possible to make a rough calculation of the efficiency of the present invention.

Section 4:
Thermodynamic Efficiency of the Invention;
Efficiency is deduced on the grounds of scientific accounting principles which are based on accurate measurements of total energy input to a system (debit), and accurate measurements of total energy (or work) obtained out of the system (credit). In principle, this is followed by drawing up a balance sheet of energy debits and credits, and expressing them as an efficiency ration, ε.

$$\eta = \frac{\text{Credit}}{\text{Debit}} = \frac{\text{Energy Out}}{\text{Energy In}} < 1$$

The energy output of Component I is an alternating current passing into a highly non-linear load, i.e., the water solution. This alternating current generator (Component I) is so designed that at peak load it is in resonance (Components I, II, III), and the vector diagrams show that the capacitive reactance, and the inductive reactance are almost exactly 180° out of phase, so that the net power output is reactive, and the dissipative power is very small. This design insures minimum power losses across the entire output system. In the experiments which are now to be described the entire emphasis was placed on achieving the maximum gas yield (credit) in exchange for the minimum applied energy (debit).

The most precise way to measure the applied energy to Components II and III is to measure the Power, P, in Watts, W. This was done by precision measurements of the volts across Component II as root mean square (rms) volts; and the current flowing in the system as rms amperes. Precisely calibrated instruments were used to take these two measurements. A typical set of experiments (using water in the form of 0.9% saline solution = 0.1540 molar concentration) to obtain high efficiency hydrolysis gave the following results:

ms Current = 25 mA to 38 mA (0.025 A to 0.038 A)
rms Volts = 4 Volts to 2.6 Volts

The resultant ratio between current and voltage is dependent on many factors, such as the gap distance between the central and ring electrodes, dielectric properties of the water, conductivity properties of the water, equilibrium states, isothermal conditions, materials used, and even the presence of clathrates. The above current and voltage values reflect the net effect of various combinations of such parameters. The product of rms current, and rms volts is a measure of the power, P in watts:

P = I x E = 25 mA x 4.0 volts = 100 mW (0.1 W)
P = I x E = 38 mA x 2.6 volts = 98.8 mW (0.0988 W)

At these power levels (with load), the resonant frequency of the system is 600 Hz (plus or minus 5 Hz) as measured on a precision frequency counter. The wave form was monitored for harmonic content on an oscilloscope, and the nuclear magnetic relaxation cycle was monitored on an X-Y plotting oscilloscope in order to maintain the proper hysteresis loop figure. All experiments were run so that the power in Watts, applied through Components I, II, and III ranged between 98.8 mW to 100 mW. Since, by the International System of Units --- 1971 (SI), One-Watt-second (Ws) is exactly equal to One Joule (J), the measurements of efficiency used these two yardsticks (1 Ws = 1 J) for the debit side of the measurement.

The energy output of the system is, of course, the two gases, hydrogen (H_2) and oxygen ($1/2O_2$), and this credit side was measured in two laboratories, on two kinds of calibrated instruments, namely, a Gas Chromatography Machine, and, a Mass Spectrometer Machine.

The volume of gases, H_2 and $(1/2)O_2$, was measured as produced under standard conditions of temperature and pressure in unit time, i.e., in ccs per minute (cc/min), as well as the possibly contaminating gases, such as air oxygen, nitrogen and argon; carbon monoxide, carbon dioxide, water vapour, etc.

The electrical, and gas, measurements were reduced to the common denominator of Joules of energy so that the efficiency accounting could all be handled in common units. The averaged results from many experiments follow. The Standard Error between different samples, machines, and locations is plus or minus 10%, and only the mean was used for all the following calculations.

Section 5:
Endergonic Decomposition of Liquid Water;

Thermodynamic efficiency for the endergonic decomposition of saline liquid water into gases under standard atmosphere (754 to 750 m.m. Hg), and standard isothermal conditions @ 25^{O}C. = 77^{O}F. = 298.16^{O}K., according to the following reaction:

$$H_2O(1) \rightarrow H_2\,(g) + (1/2)O_2\,(g) + .\Delta.G\ 56.620\ KCal/mole$$

As already described, $.\Delta.G$ is the Gibbs function (**Fig.14B**). A conversion of Kcal to the common units, Joules, by the formula, One Calorie = 4.1868 Joules was made.

$.\Delta.G = 56.620$ Kcal x 4.1868 J = 236,954 J/mol of H_2O (1) where, 1 mole is 18 gms.

$.\Delta.G$ = the free energy required to yield an equivalent amount of energy from H_2O in the form of the gases, H_2 and $(1/2)O_2$.

To simplify the calculations, the energy required to produce 1.0 cc of H_2O as the gases, H_2 and $(1/2)O_2$ was determined. There are (under standard conditions) 22,400 cc = V, of gas in one mole of H_2O. Therefore:

$$\frac{\Delta G}{V} = \frac{236{,}954\ J}{22{,}400\ cc} = 10.5783\ J/cc$$

The electrical energy required to liberate 1.0 cc of the H_2O gases (where H_2 = 0.666 parts, and $(1/2)O_2$ = 0.333 parts, by volume) from liquid water is then determined. Since P = 1 Ws = 1 Joule, and V=1.0 cc of gas = 10.5783 Joules, then:

PV = 1 x 10.5783 J = 10.5783 Ws

Since the experiments were run at 100 mW (0.1 W) applied to the water sample in Component II, III, for 30 minutes, the ideal (100% efficient) gas production at this total applied power level was calculated.
0.1 Ws x 60 sec x 30 min = 180.00 Joules (for 30 min)
The total gas production at Ideal 100% efficiency is,
180.00 J / 10.5783 J/cc = 17.01 cc H_2O (g)
The amount of hydrogen present in the 17.01 cc H_2O (g) was then calculated.
17.01 cc H_2O (gas) x 0.666 H_2 (g) = 11.329 cc H_2 (g)
17.01 cc H_2O (g) x 0.333 $(1/2)O_2$ (g) = 5.681 cc $(1/2)O_2$ (g)
Against this ideal standard of efficiency of expected gas production, the actual amount of gas produced was measured under: (1) standard conditions as defined above (2) 0.1 Ws power applied over 30 minutes. In the experiments, the mean amount of H_2 and $(1/2)O_2$ produced, as measured on precision calibrated GC, and MS machines in two different laboratories, where the S.E. is +/-10%, was,

_____

Measured Mean = 10.80 cc H_2 (g)
Measured Mean = 5.40 cc $(1/2)O_2$ (g)
Total Mean = 16.20 cc H_2O(g)

_____

The ratio, ε, between the ideal yield, and measured yield is:

$$\eta = \frac{\text{Measured } H_2(g)}{\text{Ideal } H_2(g)} = \frac{10.80 \text{ cc}}{11.33 \text{ cc}} = 91.30\%$$

Section 6:
Energy Release;
The total energy release (as heat, or electricity) from an exergonic reaction of the gases, H_2 and O_2, is given by:

$$H_2(g) + (\tfrac{1}{2})O_2(g) \rightarrow H_2O(P) - \Delta H\ 68.315\ \text{Kcal/mol} = (-\Delta H\ 286{,}021\ \text{Joules})/\text{mol}$$

It is possible (Penner, Op. Cit., p.128) to get a total heat release, or total conversion to electricity in a fuel cell, in the above reaction when the reactants are initially near room temperature (298.16^{O}K.), and the reactant product (H_2O) is finally returned to room temperature. With this authoritative opinion in mind, it is desirable to determine the amount of energy released (ideal) from the exergonic experiment. The total energy of 1.0 cc of H_2O (1), as above is:

$$1.0 \text{ cc } \Delta H = \frac{286{,}021 \text{ J/mol}}{22{,}400 \text{ cc/mol}} = 12.7687 \text{ J/cc } H_2O$$

for H_2 = 12.7687 x 0.666 = 8.509 J/0.66 cc H_2 for O_2 = 12.7687 x 0.333 = 4.259 J/0.33 cc $(1/2)O_2$ The energy produced from the gases produced in the experiments in an exergonic reaction was:

16.20 cc H_2O (g) x 12.7687 J/cc H_2O = 206,8544 J.

The overall energy transaction can be written as:

$$\frac{\text{EXERGONIC}}{\text{ENDERGONIC}} - \eta - \frac{-\Delta H}{+\Delta G} = \frac{206{,}854.4 \text{ J}}{180{,}000 \text{ J}} = 114.92\%$$

In practical bookkeeping terms the balance of debits and credits, n = (-.Δ.H) - (+.Δ.G), so:

n = 206.8544 J - 180.0 = + 26.8544 J (surplus).

Since, in the invention, the gas is produced where and when needed, there is no additional cost accounting for liquefaction, storage, or transportation of the hydrogen fuel, and the oxygen oxidant. Therefore, the practical efficiency, is:

$$\eta P = \frac{26.8544\ J}{180.0000\ J} = 14.919\%\ \text{(as net return on the original energy investment)}$$

In practical applications, the energy output (exergonic) of the Component II System can be parsed between the electrical energy required to power the Component I System, as an isothermal closed loop; while the surplus of approximately 15% can be shunted to an engine (heat, electrical, battery, etc.) that has a work load. Although this energy cost accounting represents an ideal model, it is believed that there is enough return (approximately 15%) on the capital energy investment to yield a net energy profit that can be used to do useful work.

CONCLUSION:

From the foregoing disclosure it will be appreciated that the achievement of efficient water splitting through the application of complex electrical waveforms to energised water molecules, i.e. tetrahedral molecules having bonding angles of 109°28', in the special apparatus described and illustrated, will provide ample and economical production of hydrogen gas and oxygen gas from readily available sources of water. It is to be understood, that the specific forms of the invention disclosed and discussed herein are intended to be representative and by way of illustrative example only, since various changes may be made therein without departing from the clear and specific teachings of the disclosure. Accordingly, reference should be made to the following appended claims in determining the full scope of the method and apparatus of the present invention.

SHIGETA HASEBE

US Patent 4,105,528 8th August 1978 Inventor: Shigeta Hasebe

APPARATUS FOR DECOMPOSITION OF AQUEOUS LIQUID

Please note that this is a re-worded excerpt from this patent. This patent describes an electrolysis system which it is claimed has demonstrated ten times the efficiency that Faraday considered to be the maximum possible.

ABSTRACT

An apparatus for decomposition of liquid, in which spiral negative and positive electrodes are arranged close together but not touching. These two electrodes are supplied with power through external terminals and the electrolyte is caused to flow between the negative and positive electrodes for the electrolysis between two electrodes under the function of the potential magnetic field formed by the coil current which is generated by the electrodes with active movement of an electrolytic ion so that the electrolysis of water takes place smoothly under the spin functions of the atom and electron.

BACKGROUND AND SUMMARY OF THE INVENTION

This invention relates to an apparatus for decomposition of liquid where a flowing electrolyte is subjected to electrolysis for the production of gases.

As is well known, water is composed of hydrogen atoms and oxygen atoms. When water is sufficiently magnetised, each constitutive atom is also weakly magnetised to rotate the elementary particle in a regular direction. This rotation of the elementary particle is generally called "spin". That is, the spin function is caused by an electron, atomic nucleus, atom and even by the molecule. When a negative electrode is immersed in the electrolyte - Sodium Hydroxide (“lye”) solution - with a view to applying a voltage to it in order to cause the elementary particle to react with the electric field, the coupling state of the hydrogen with the oxygen is varied and the electrolysis is facilitated by the spin.

In the present invention, spiral negative and positive electrodes are arranged close together but not touching and these two electrodes are supplied with power through external terminals and the electrolyte is caused to flow between the negative and positive electrodes. Thus, the electrolyte is subjected to the electrolysis between two electrodes while within a magnetic field formed by the coil current which is generated by the electrodes with active movement of an electrolytic ion (Na^+, OH^-) so that the electrolysis of water takes place smoothly under the spin functions of the atom and electron.

It has been confirmed that the rate of the electrolysis of water using this invention is approximately 10 or more times (approximately 20 times when calculated) than that produced by conventional electrolysis.

The design of the electrolytic cell of this invention is such that the electrolyte flowing through the supply ports provided at the lower portion of the electrolytic cell is subjected to the magnetic field produced by a permanent magnet and the electrodes cause it to be further subjected to magnetic and electric fields which cause it to obtain a sufficient spin effect.

It is, therefore, a general object of the invention to provide a novel apparatus for decomposition of liquid in which an electrolyte (NaOH) solution is subjected to magnetic fields to cause electrolysis assisted by the spin of the water molecules which produces a great amount of gas with less consumption of electrical energy.

A principal object of the invention is to provide an apparatus for decomposition of liquid which has a liquid circulating system for the separation of gas and liquid in which positive and negative spiral electrodes are arranged across the flow path of the liquid and the opposite ends of the electrodes being provided with magnetic materials to augment the effect caused by the applied voltage across a liquid passing through a magnetic field caused by the positive and negative spiral electrodes, thereby to promote generation and separation of cat-ions and an-ions with a high efficiency in production of a large quantity of gases.

Other objects and advantages of the present invention will become apparent through the detailed description which follows.

BRIEF DESCRIPTION OF THE DRAWINGS

The invention will be described more in detail in the following with reference to the accompanying drawings, wherein:

Fig.1 is a partially cross-sectional schematic elevation of an apparatus in accordance with the invention;

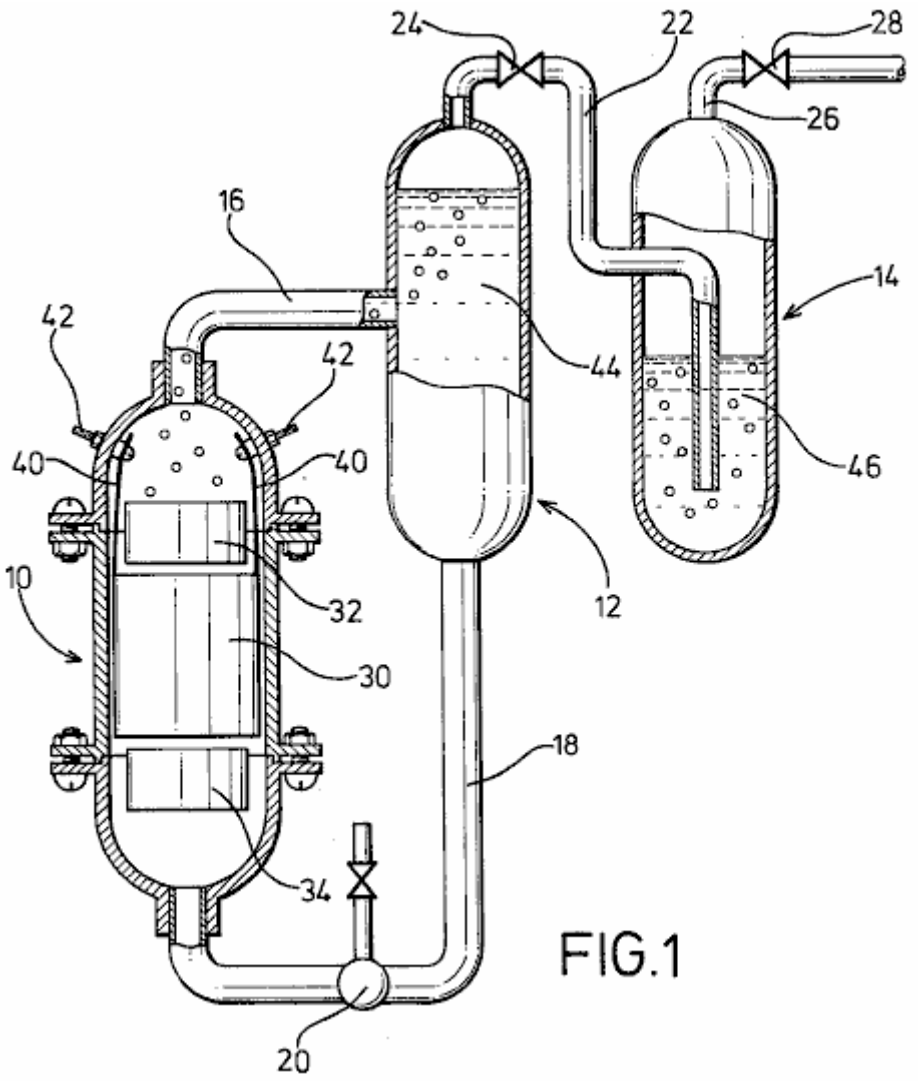


FIG.1

Fig.2 is a perspective view of electrodes arranged in accordance with the invention;

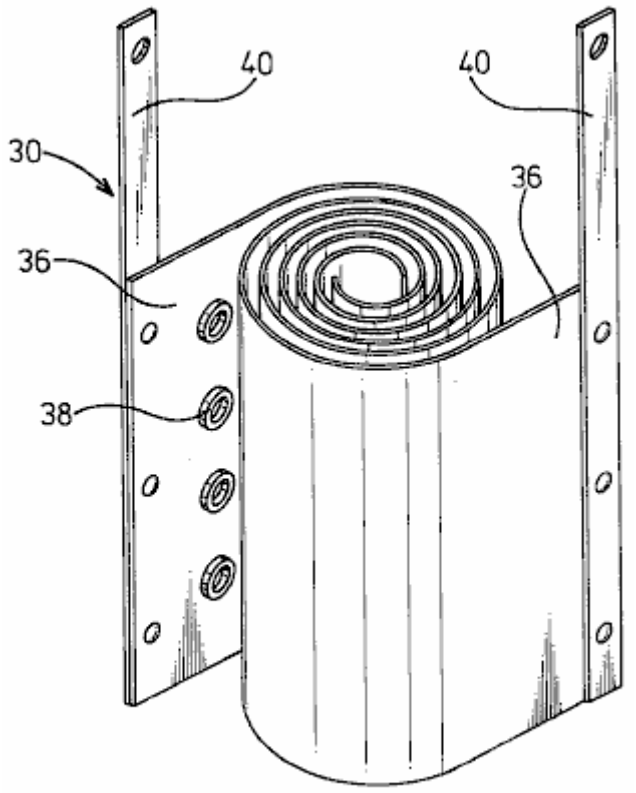


FIG.2

Fig.3 is a plan view of electrodes with magnetic materials.

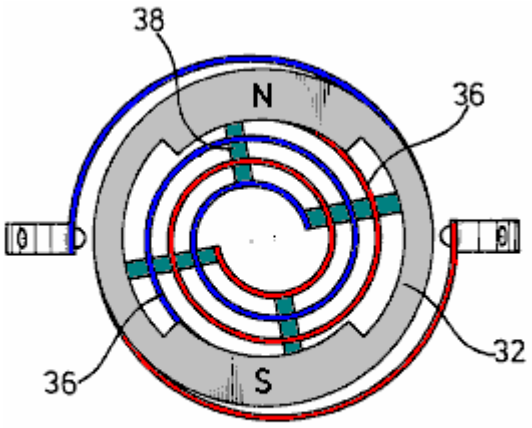


FIG.3

DESCRIPTION OF THE PREFERRED EMBODIMENT

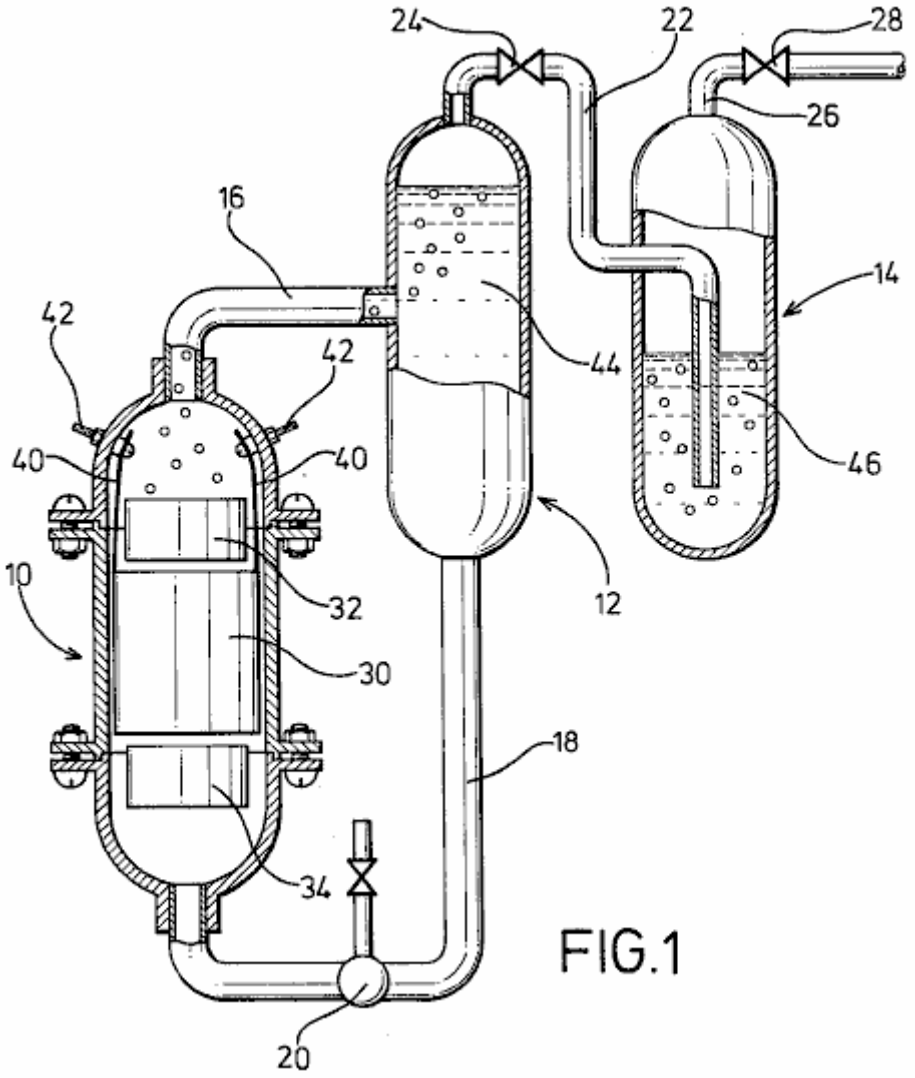


FIG.1

In **Fig.1**, an electrolysis cell **10**, a gas-liquid separation tank **12** and a gas-washing tank **14** are vertically arranged as shown with the electrolytic cell **10** being positioned a little lower than the tanks.

Cell **10** and tanks **12** and **14** are connected together by a delivery pipe **16** which connects the top of the electrolytic cell **10** with the middle of the gas-liquid separation tank **12**. A feed-back pipe **18** containing a pump **20**, is provided to connect the bottom of the gas-liquid separation tank **12,** with the bottom of the electrolytic cell **10**. Also provided is pipe **22**, which runs from the top of the gas-liquid separation tank **12** through a valve **24** to the bottom of the gas-washing tank **14**. A drain pipe **26**, provided with a valve **28**, is taken from the top of the gas-washing tank **14**.

In the electrolytic cell **10**, positive and negative spiral electrodes **30** of diameters suited to the internal diameter of the electrolytic cell **10** are arranged coaxially. At the upper and lower parts of the spiral electrodes **30** are arranged magnet rings **32** and **34** made from ferrite or similar material, positioned so that North and South poles are opposite one another to create a magnetic field which is at right angles to the axis of the electrolytic cell.

Electrodes **30** are composed of two metal strips **36** which are wound into spiral shapes with cylindrical insulating spacers **38** made of rubber or a similar material, placed between them and attached to the surface of the metal strips **36**. From the metal strips **36,** wires **40**, are taken to the positive and negative power supply terminals, via connectors provided in the inner wall of the electrolytic cell.

The electrolytic cell **10** and the gas-liquid separation tank **12** are filled with a electrolyte **44** which is circulated by the pump **20**, while the gas-washing tank **14** is filled with a washing liquid **46** to such a level that gases gushing out of the conduit **22** are thoroughly washed.

The apparatus of the present invention may be well be used for the electrolysis of flowing water for the production of hydrogen gas and oxygen gas at a high efficiency. That is to say, the electrolytic cell **10** and the gas-liquid separation tank **12** are filled with the electrolyte **44** which is caused by pump **20** to flow through a magnetic field in an vortex path in which positive and negative magnetic poles N, S of the magnets **32** and **34** face each other to

produce a transverse field, and through the metal plates **36** of the vortical electrodes **30** to generate an orientation for the electrical migration of cat-ions and an-ions, causing an increased gas separation rate and enhancement of the electrolysis.

In particular, the flowing oxygen gas serves to facilitate an aeration of the electrolyte since it has varying magnetic effects as it passes through the magnetic field. The spiral electrodes **30** of this invention, create a remarkable increase in the rate of electrolysis. This is caused by the continuously decreasing space between the electrodes **30** which causes the flow velocity to increase as the flow progresses along its path. This causes turbulence which instantly removes bubbles of gas from the surface of the electrodes, allowing fresh ions full contact with the metal surfaces, thus raising the efficiency of the cell.

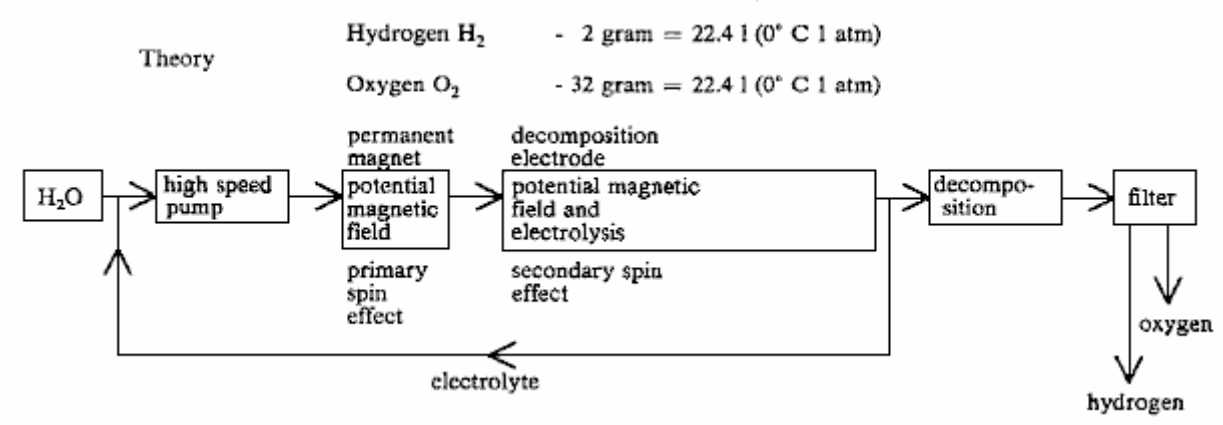
The spiral coiling of the electrodes also enables a very desirable reduction in the size of the cell, while increasing the electrode area and improving its contact with the electrolyte **44**. There is also a relatively short migration distance of ions which also promotes rapid gas production. On the other hand, insulating spacers **38** interposed between the metal strips **36** serves to create the desired turbulence of the electrolyte passing through the cell.

The liquid circulating system for separation of gas and liquid requires no other driving unit except the circulation pump **20** to achieve separation of gas and liquid by utilising differences in water heads between cell **10** and tanks **12** and **14**. In other words, a flow of gas-liquid mixture supplied from electrolytic cell **10** is fed into the gas-liquid separation tank **12** where, due to the difference in buoyancy of gases and liquid, the gas rises and is fed into the gas-washing tank **14** while the liquid moves down and is returned to the electrolytic cell **10**. The washing tank **14** is filled with any convenient washing liquid **46** so that the gases gushing out of conduit **22** are thoroughly washed and fed into the drain pipe **26**. Thus, the apparatus may be constructed at reduced cost and without any complexity.

As described earlier, the magnets **32** and **34** provide positive and negative magnetic poles N, S which are confronted in the annular wall for facilitating an alignment between the cross section of the flow-path of the liquid and the annular portion of the magnets **32** and **34** and a generation of a magnetic field in a direction perpendicular to that of the liquid flow, so that the liquid is forced to flow through the magnetic field.

Experimental data	Value
Room temperature	20^0 Centigrade
Atmospheric pressure	1003 millibars
Electrolyte temperature	25^0 Centigrade
Humidity	43%
Voltage	2.8 Volts
Current	30 Amps
Hydroxy gas production rate	116 cc/sec.
Hydrogen production per Coulomb (1A x 1 sec.)	2.6 cc.
Oxygen production per Coulomb	1.3 cc.

The rate of generation shown by these figures is over 20 times that which could be obtained by standard Faraday electrolysis.



While a preferred embodiment of the invention has been illustrated by way of example in the drawings and particularly described, it will be understood that various modifications may be made in the construction and that the invention is no way limited to the embodiments shown.

STEPHEN CHAMBERS (Xogen Power Inc.)

US Patent 6,126,794 16th July 2002 Inventor: Stephen Chambers

AN APPARATUS FOR PRODUCING ORTHOHYDROGEN AND/OR PARAHYDROGEN

This patent describes an electrolyser system capable of running a small internal combustion engine directly from water alone.

ABSTRACT

An apparatus for producing orthohydrogen and/or parahydrogen. The apparatus includes a container holding water and at least one pair of closely-spaced electrodes arranged within the container and submerged in the water. A first power supply provides a particular first pulsed signal to the electrodes. A coil may also be arranged within the container and submerged in the water if the production of parahydrogen is also required. A second power supply provides a second pulsed signal to the coil through a switch to apply energy to the water. When the second power supply is disconnected from the coil by the switch and only the electrodes receive a pulsed signal, then orthohydrogen can be produced. When the second power supply is connected to the coil and both the electrodes and coil receive pulsed signals, then the first and second pulsed signals can be controlled to produce parahydrogen. The container is self-pressurised and the water within the container requires no chemical catalyst and yet can produce the orthohydrogen and/or parahydrogen efficiently. Heat is not generated, and bubbles do not form on the electrodes.

BACKGROUND OF THE INVENTION

Conventional electrolysis cells are capable of producing hydrogen and oxygen from water. These conventional cells generally include two electrodes arranged within the cell which apply energy to the water to thereby produce hydrogen and oxygen. The two electrodes are conventionally made of two different materials.

However, the hydrogen and oxygen generated in the conventional cells are generally produced in an inefficient manner. That is, a large amount of electrical power has to be applied to the electrodes in order to produce the hydrogen and oxygen. Moreover, a chemical catalyst such as sodium hydroxide or potassium hydroxide must be added to the water to separate hydrogen or oxygen bubbles from the electrodes. Also, the produced gas must often be transported to a pressurised container for storage, because conventional cells produce the gases slowly. Also, conventional cells tend to heat up, creating a variety of problems, including boiling of the water. In addition, conventional cells tend to form gas bubbles on the electrodes which act as electrical insulators and reduce the efficiency of the cell.

Accordingly, it is extremely desirable to produce a large amount of hydrogen and oxygen with only a modest amount of input power. Furthermore, it is desirable to produce the hydrogen and oxygen with "regular" tap water and without any additional chemical catalyst, and to operate the cell without the need for an additional pump to pressurise it. It is also desirable to construct both of the electrodes from the same material. It is also desirable to produce the gases quickly, and without heat, and without bubbles forming on the electrodes.

Orthohydrogen and parahydrogen are two different isomers of hydrogen. Orthohydrogen is that state of hydrogen molecules in which the spins of the two nuclei are parallel. Parahydrogen is that state of hydrogen molecules in which the spins of the two nuclei are antiparallel. The different characteristics of orthohydrogen and parahydrogen lead to different physical properties. For example, orthohydrogen is highly combustible whereas parahydrogen is a slower burning form of hydrogen. Thus, orthohydrogen and parahydrogen can be used for different applications. Conventional electrolytic cells make only orthohydrogen and parahydrogen. Parahydrogen is difficult and expensive to make by conventional means.

Accordingly, it is desirable to produce orthohydrogen and/or parahydrogen cheaply within a cell and to be able to control the amount of either produced by that cell. It is also desirable to direct the produced orthohydrogen or parahydrogen to a coupled machine in order to provide a source of energy for it.

SUMMARY OF THE INVENTION

It is therefore an object of the present invention to provide a cell having electrodes and containing water which produces a large amount of hydrogen and oxygen in a relatively small amount of time, and with a modest amount of input power, and without generating heat.

It is another object of the present invention for the cell to produce bubbles of hydrogen and oxygen which do not bunch around or on the electrodes.

It is also an object of the present invention for the cell to operate properly without a chemical catalyst. Thus, the cell can be run using ordinary tap water. This has the advantage of avoiding the additional costs required for producing the chemical catalyst.

It is another object of the present invention for the cell to be self-pressurising. Thus avoiding the need for an additional pump.

It is another object of the present invention to provide a cell having electrodes made of the same material. This material can, for example, be stainless steel. Thus, the construction of the cell can be simplified and construction costs reduced.

It is another object of the present invention to provide a cell which is capable of producing orthohydrogen, parahydrogen or a mixture thereof and can be set so as to produce any relative amount of orthohydrogen and parahydrogen desired by the user.

It is another object of the invention to couple the gaseous output of the cell to a device, such as an internal combustion engine, so that the device may be powered from the gas supplied to it.

These and other objects, features, and characteristics of the present invention will be more apparent upon consideration of the following detailed description and appended claims with reference to the accompanying drawings, wherein the same reference numbers have been used to indicate corresponding parts in the various figures.

Accordingly, the present invention includes a container for holding water. At least one pair of closely-spaced electrodes are positioned within the container and submerged under the water. A first power supply provides a particular pulsed signal to the electrodes. A coil is also arranged in the container and submerged under the water. A second power supply provides a particular pulsed signal through a switch to the electrodes.

When only the electrodes receive a pulsed signal, then orthohydrogen can be produced. When both the electrodes and coil receive pulsed signals, then parahydrogen or a mixture of parahydrogen and orthohydrogen can be produced. The container is self pressurised and the water within the container requires no chemical catalyst to produce the orthohydrogen and/or parahydrogen efficiently.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig.1 is a side view of a cell for producing orthohydrogen including a pair of electrodes according to a first embodiment of the present invention;

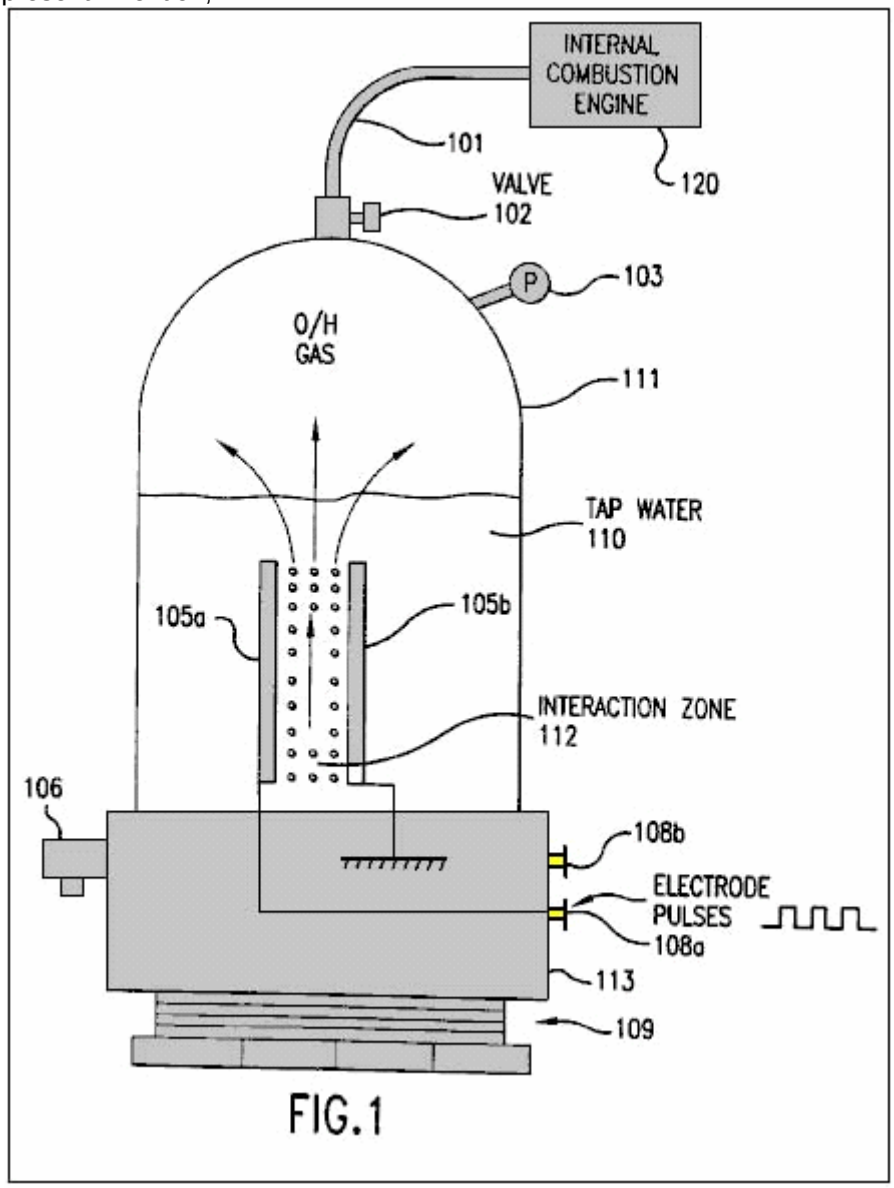


FIG.1

Fig.2 is a side view of a cell for producing orthohydrogen including two pairs of electrodes according to a second embodiment of the present invention;

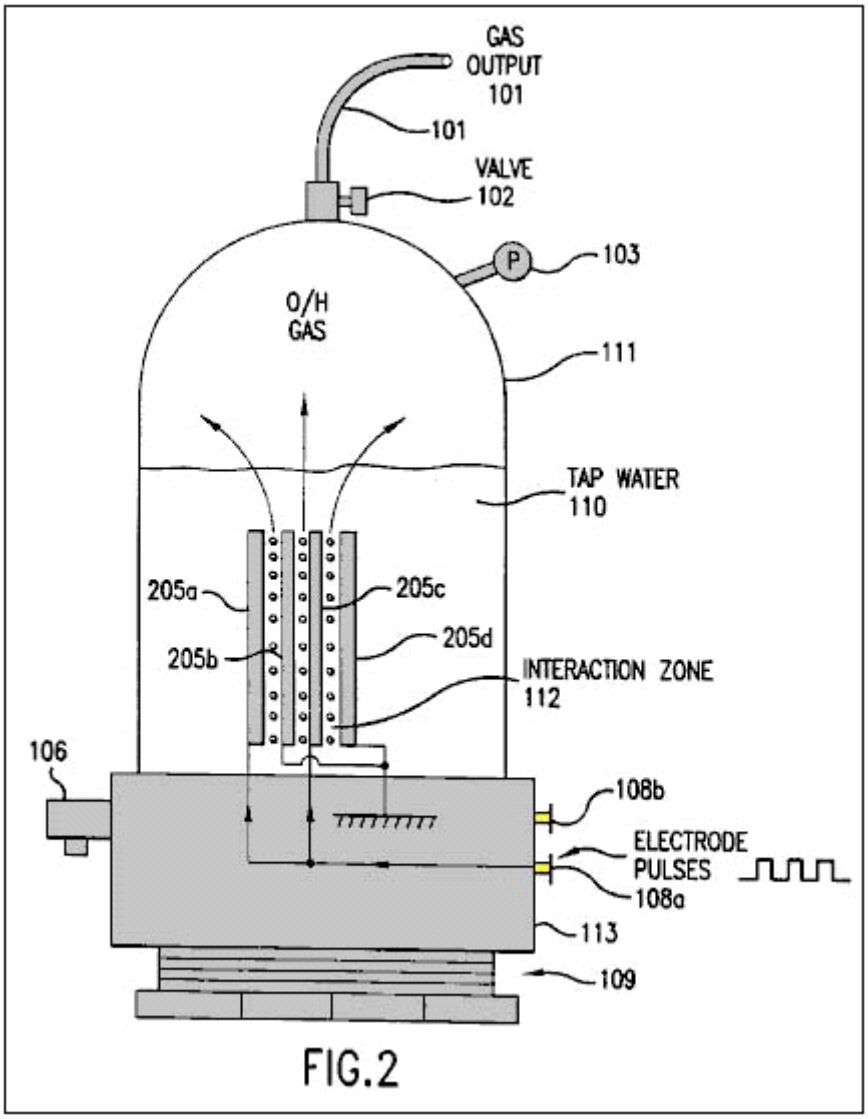


FIG.2

Fig.3 is a side view of a cell for producing orthohydrogen including a pair of cylindrical-shaped electrodes according to a third embodiment of the present invention;

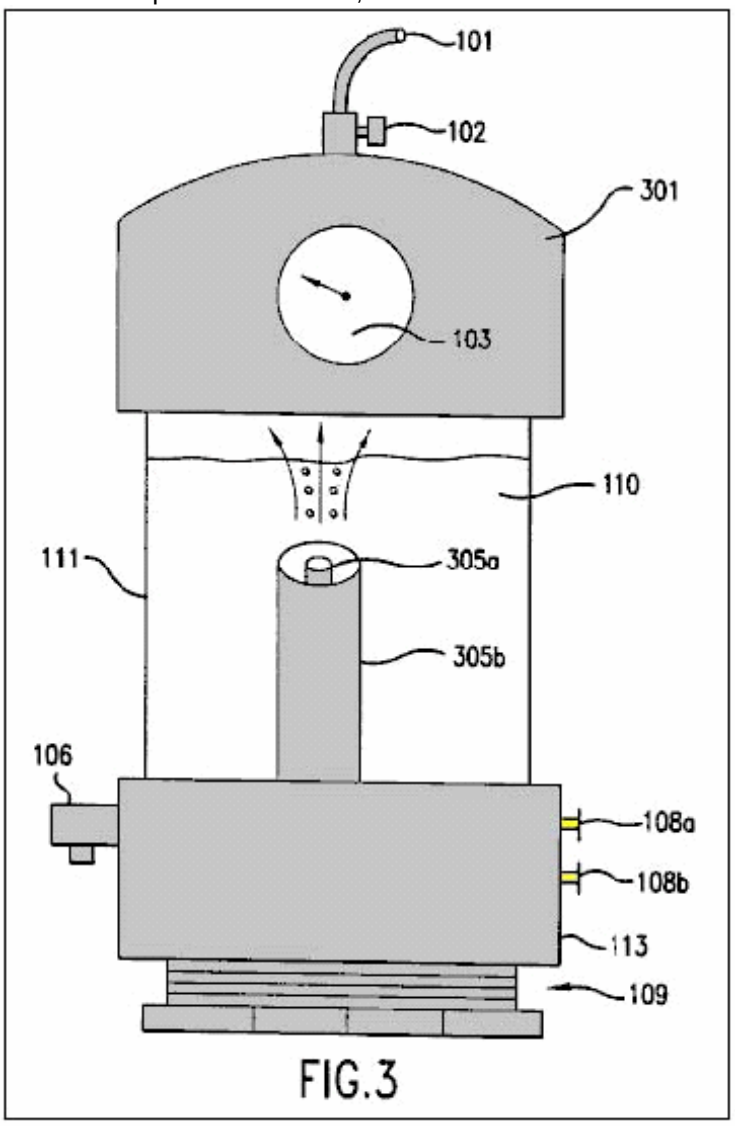


FIG.3

Fig.4a is a diagram illustrating a square wave pulsed signal which can be produced by the circuit of **Fig.5** and applied to the electrodes of **Fig.1** through **Fig.3**;

Fig.4b is a diagram illustrating a saw tooth wave pulsed signal which can be produced by the circuit of **Fig.5** and applied to the electrodes of **Fig.1** through **Fig.3**;

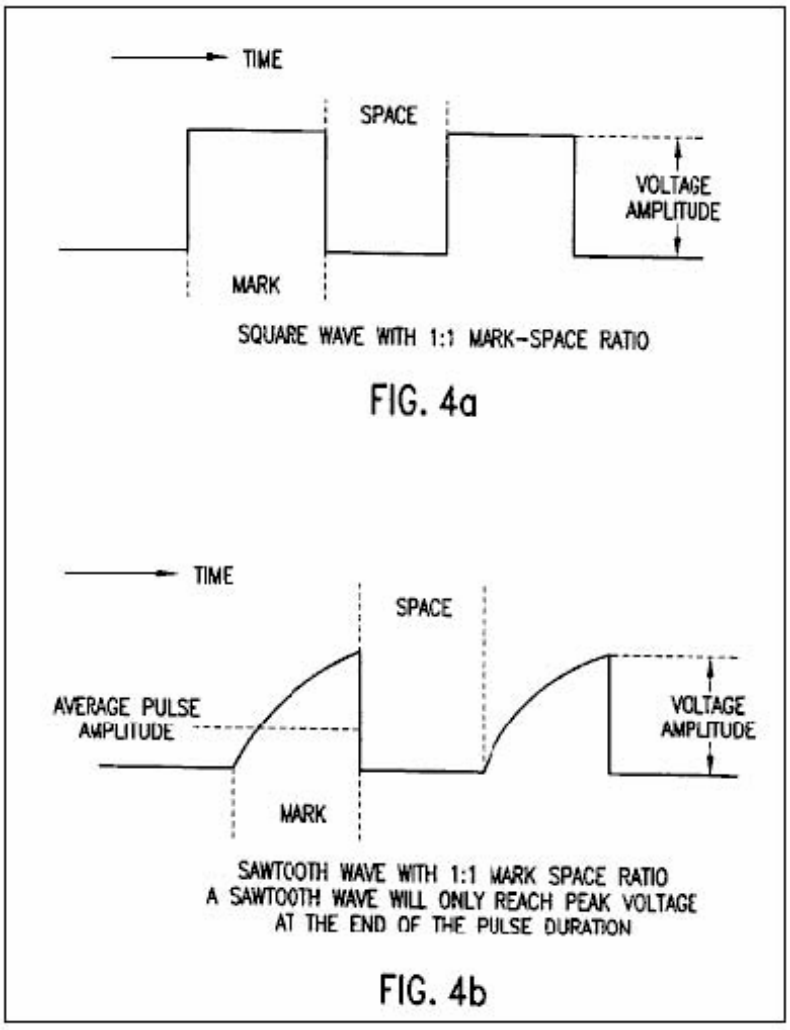


FIG. 4a

FIG. 4b

Fig.4c is a diagram illustrating a triangular wave pulsed signal which can be produced by the circuit of **Fig.5** and applied to the electrodes of **Fig.1** through **Fig.3**;

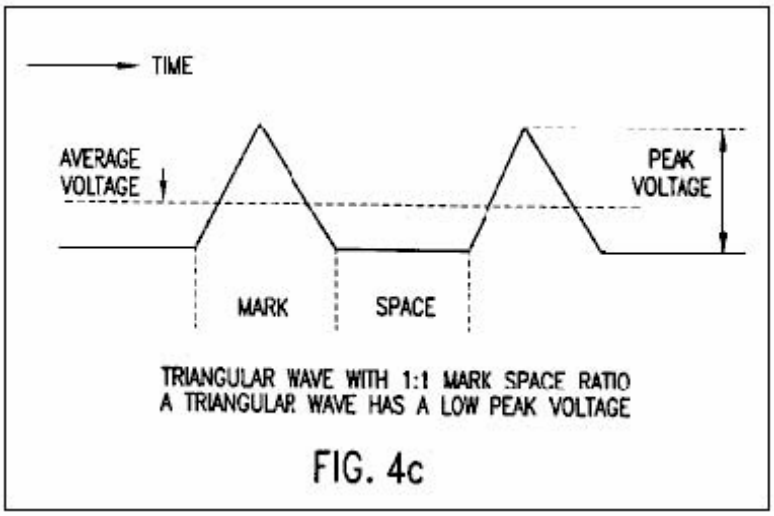


FIG. 4c

Fig.5 is an electronic circuit diagram illustrating a power supply which is connected to the electrodes of **Fig.1** through **Fig.3**;

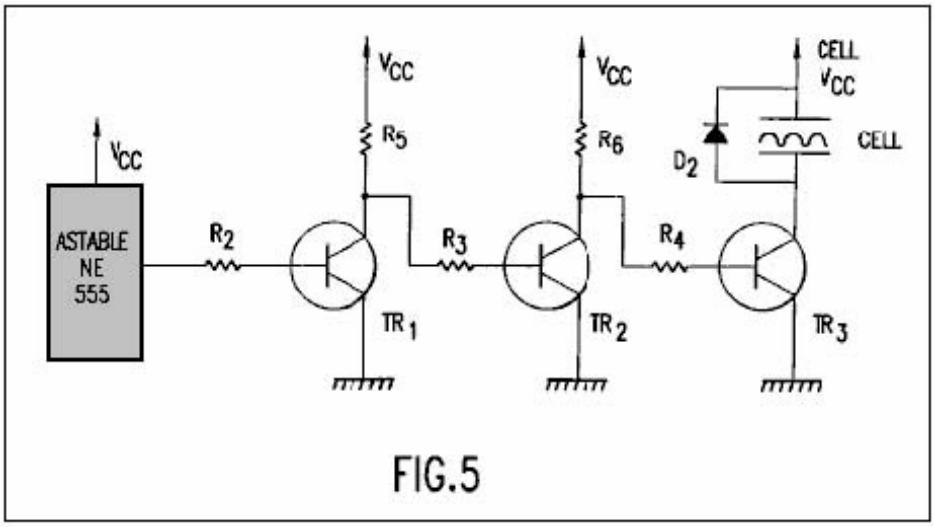


FIG.5

Fig.6 is a side view of a cell for producing at least parahydrogen including a coil and a pair of electrodes according to a fourth embodiment of the present invention;

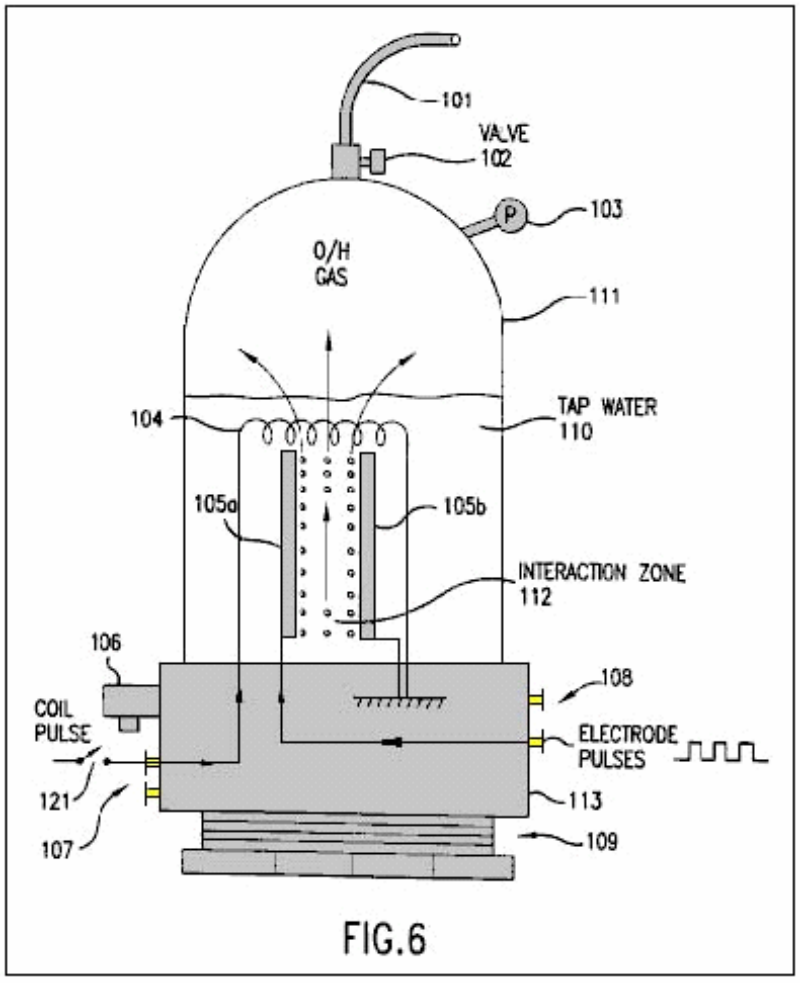


FIG.6

Fig.7 is a side view of a cell for producing at least parahydrogen including a coil and two pairs of electrodes according to a fifth embodiment of the present invention;

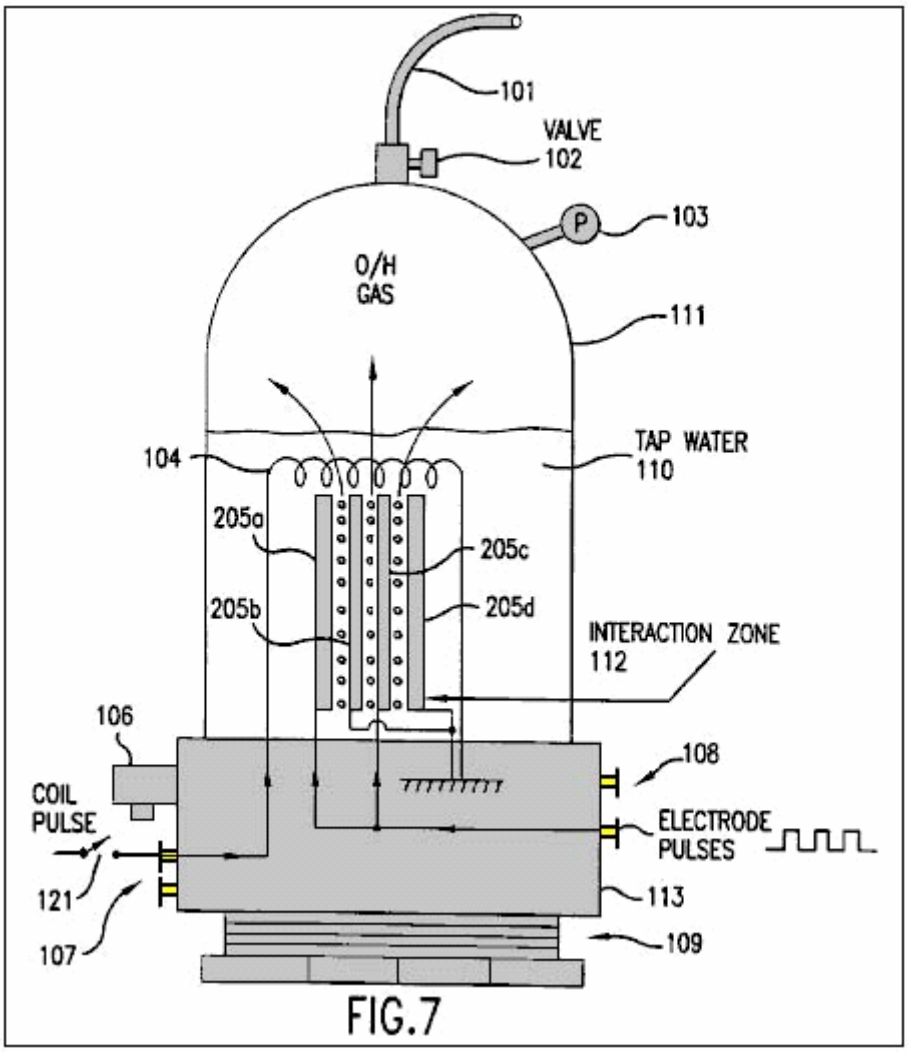


FIG.7

Fig.8 is a side view of a cell for producing at least parahydrogen including a coil and a pair of cylindrical-shaped electrodes according to a sixth embodiment of the present invention; and

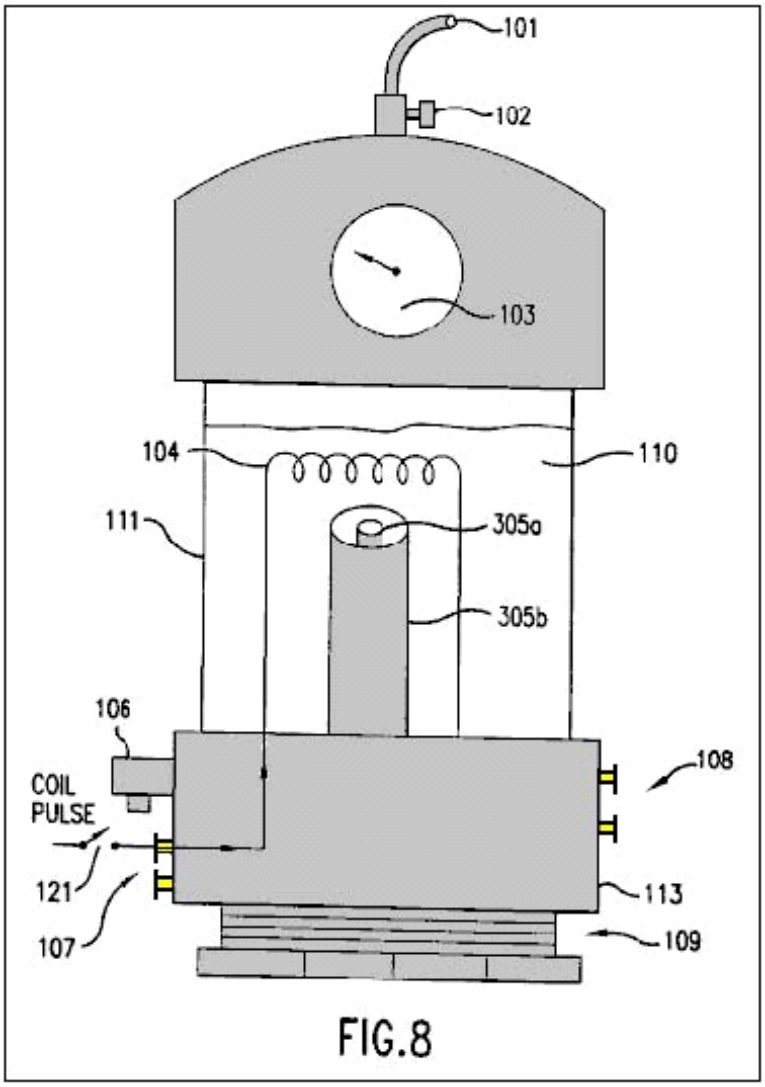


FIG.8

Fig.9 is as electronic circuit diagram illustrating a power supply which is connected to the coil and electrodes of **Fig.6** through **Fig.8**.

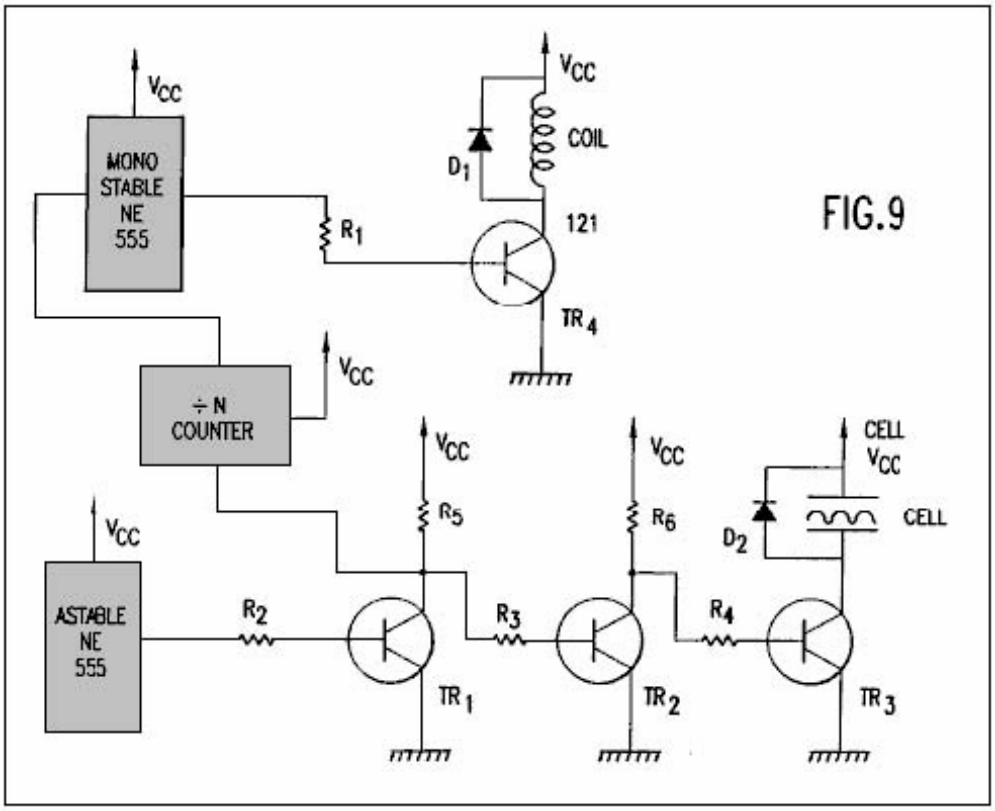


FIG.9

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

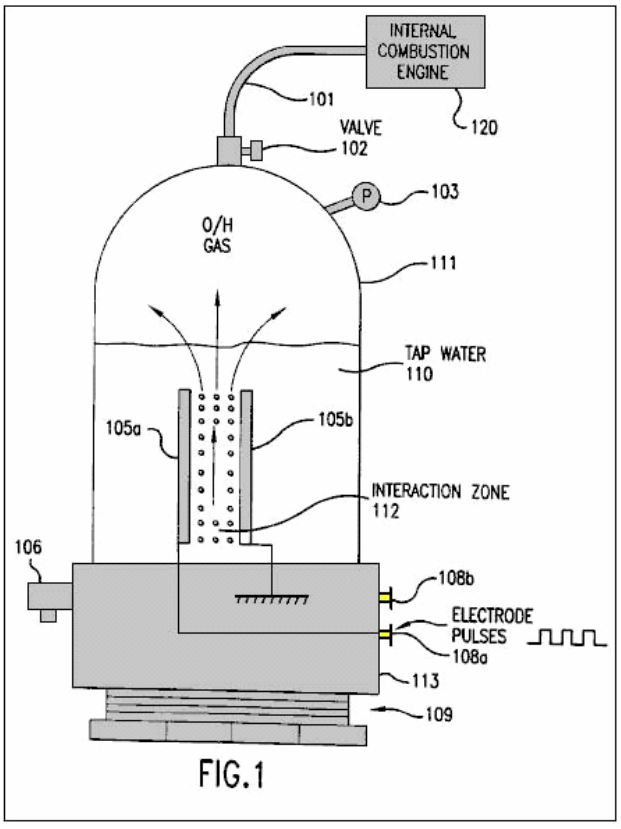


FIG.1

Fig.1 shows a first embodiment of the present invention including a cell for producing hydrogen and oxygen. As will be discussed below in conjunction with **Figs.6-8**, the production of parahydrogen requires an additional coil not shown in **Fig.1**. Thus, the hydrogen produced by the first embodiment of **Fig.1** is orthohydrogen.

The cell includes a closed container **111** which is closed at its bottom portion by threaded plastic base **113** and screw thread base **109**. The container **111** can be made of, for example, Plexiglas and might have a height of 430 mm and a width of 90 mm. The container **111** holds tap water **110**.

The cell also includes a pressure gauge **103** to measure the pressure within the container **111**. An outlet valve **102** is connected to the top of the container **111** to permit any gas within the container to escape into an output tube **101**.

The cell also includes an over-pressure valve **106** connected to a base **113**. The valve **106** provides a safety function by automatically releasing the pressure within the container **111** if the pressure exceeds a predetermined threshold. For example, the valve **106** may be set so that it will open if the pressure in the container exceeds 75 p.s.i. Since the container **111** is built to withstand a pressure of about 200 p.s.i., the cell is provided with a large safety margin.

A pair of electrodes **105a** and **105b** are arranged within the container **111**. These electrodes are submerged under the top level of the water **110** and define an interaction zone **112** between them. The electrodes are preferably made from the same material, such as stainless steel.

In order to produce an optimum amount of hydrogen and oxygen, an equal spacing between the electrodes **105a** and **105b** must be maintained. Moreover, it is preferable to minimise the spacing between the electrodes. However, the electrodes cannot be positioned excessively close together, because arcing between the electrodes would occur. It has been determined that a spacing of 1 mm is the optimum spacing for producing hydrogen and oxygen. Spacing up to 5 mm can work effectively, but spacing above 5 mm has not worked well, except with excessive power.

Hydrogen and oxygen gas may be output through tube **101** to a device **120** which can use those gases, for example an internal combustion engine, such as shown in **Fig.1**. Instead of an internal combustion engine, device **120** may be any device using hydrogen and oxygen, including a reciprocating piston engine, a gas turbine engine, a stove, a heater, a furnace, a distillation unit, a water purification unit, a hydrogen/oxygen jet, or other device using the gases. With an adequately productive example of the present invention, any such device **120** using the output gases can be run continuously without the need for storing dangerous hydrogen and oxygen gases.

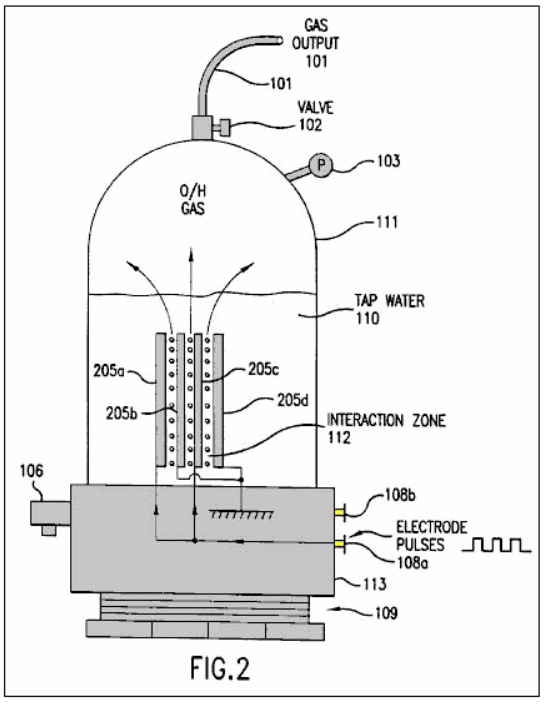


FIG.2

Fig.2 shows a second embodiment of the present invention which includes more than one pair of electrodes **205a-d**. The spacing between the electrodes is less than 5 mm as in the embodiment of **Fig.1**. While **Fig.2** shows only one additional pair of electrodes, it is possible to include many more pairs (e.g., as many as 40 pairs of electrodes) within the cell. The rest of the cell illustrated in **Fig.2** remains the same as that illustrated in **Fig.1**. The multiple electrodes are preferably flat plates closely spaced, parallel to each other.

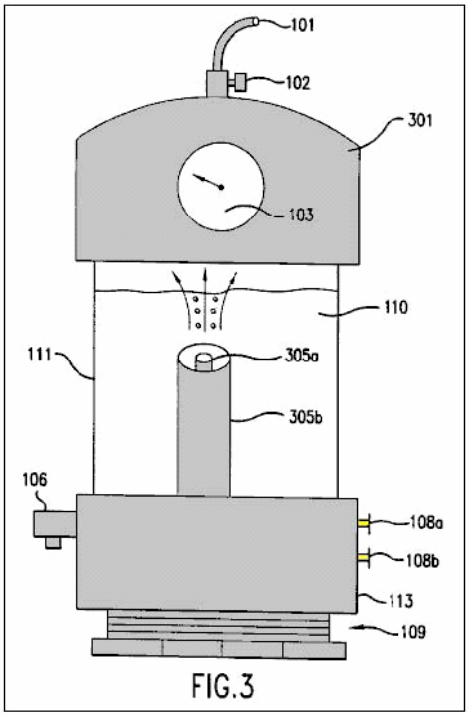


FIG.3

Fig.3 illustrates a cell having a cylindrically shaped electrodes **305a** and **305b**. The outer electrode **305b** surrounds the coaxially aligned inner electrode **305a**. The equal spacing of the electrodes **305a** and **305b** is less than 5 mm and the interactive zone is coaxially arranged between the two electrodes. While **Fig.3** illustrates the top portion of the container **111** being formed by a plastic cap **301**, it will be appreciated by those skilled in the art, that the cap **301** may be used in the embodiments of **Fig.1** and **Fig.2** and the embodiment of **Fig.3** can utilise the same container **111** illustrated in **Figs.1-2**. As suggested by **Fig.3**, the electrodes can be almost any shape such as flat plates, rods, tubes or coaxial cylinders.

The electrodes **105a** and **105b** of **Fig.1** (or electrodes **205a-d** of **Fig.2** or electrodes **305a** and **305b** of **Fig.3**) are respectively connected to power supply terminals **108a** and **108b** so that they can receive a pulsed electrical signal from a power supply. The pulsed signal can be almost any waveform and have a variable current level, voltage level, frequency and mark-space ratio (i.e., a ratio of the duration of a single pulse to the interval between two successive pulses). For example, the power supply providing power to the electrodes can be a mains 110 volts to a 12 volt supply or a car battery.

Fig.4a, **Fig.4b** and **Fig.4c** illustrate a square wave, a saw tooth wave and a triangular wave, respectively which can be applied to the electrodes **105a** and **105b** (or **205a-d** or **305a**, **305b**) in accordance with the present invention. Each of the waveforms illustrated in **Figs.4a-4c** has a 1:1 mark-space ratio. As shown in **Fig.4b**, the saw tooth wave will only reach a peak voltage at the end of the pulse duration. As shown in **Fig.4c**, the triangular wave has a low peak voltage. It has been found that optimal results for producing hydrogen and oxygen in the present invention are obtained using a square wave.

After initiation of the pulsed signal from the power supply, the electrodes **105a** and **105b** continuously and almost instantaneously generate hydrogen and oxygen bubbles from the water **110** in the interaction zone **112**. Moreover, the bubbles can be generated with only minimal heating of the water or any other part of the cell. These bubbles rise through the water and collect in the upper portion of the container **111**.

The generated bubbles are not bunched around or on the electrodes **105a** and **105b** and thus readily float to the surface of the water. Therefore, there is no need to add a chemical catalyst to assist the conduction of the solution or reduce the bubble bunching around or on the electrodes. Thus, only tap water is needed for generation of the hydrogen and oxygen in the present invention.

The gases produced within the container are self-pressurising (i.e., pressure builds in the container by the production of gas, without an air pump). Thus, no additional pump is needed to be coupled to the container **111** and the produced gases do no need to be transported into a pressurised container.

The power supply in the present invention is required to provide a pulsed signal having only 12 volts at 300 mA (3.6 watts). It has been found that an optimal amount of hydrogen and oxygen has been produced when the pulsed signal has mark-space ratio of 10:1 and a frequency of 10-250 KHz. Using these parameters, the prototype cell of the present invention is capable of producing gas at the rate of 1 p.s.i. per minute. Accordingly,

the cell of the present invention is capable of producing hydrogen and oxygen in a highly efficient manner, quickly and with low power requirements.

As noted above, the hydrogen produced by the embodiments of **Figs.1-3** is orthohydrogen. As is well understood by those skilled in the art, orthohydrogen is highly combustible. Therefore, any orthohydrogen produced can be transported from the container **111** through valve **102** and outlet tube **101** to be used by a device such as an internal combustion engine.

The present invention, with sufficient electrodes, can generate hydrogen and oxygen fast enough to feed the gases directly into an internal combustion engine or turbine engine, and run the engine continuously without accumulation and storage of the gases. Hence, this provides for the first time a hydrogen/oxygen driven engine that is safe because it requires no storage of hydrogen or oxygen gas.

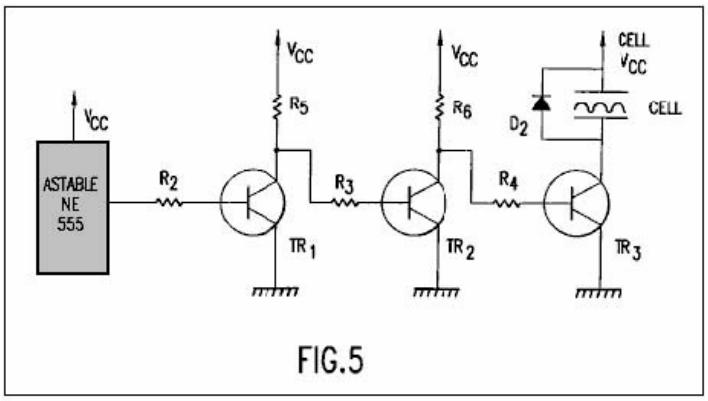


FIG.5

Fig.5 illustrates an exemplary power supply for providing D.C. pulsed signals such as those illustrated in **Figs.4a-4c** to the electrodes illustrated in **Figs.1-3**. As will be readily understood by those skilled in the art, any other power supply which is capable of providing the pulsed signals discussed above can be substituted.
The power supply illustrated in **Fig.5** includes the following parts, components and values:

The astable circuit is connected to the base of transistor **TR1** through resistor **R2**. The collector of transistor **TR1** is connected to voltage supply **Vcc** through resistor **R5** and the base of transistor **TR2** through resistor **R3**. The collector of transistor **TR2** is connected to voltage supply **Vcc** through resistor **R6** and the base of transistor **TR3** through resistor **R4**. The collector of transistor **TR3** is connect to one of the electrodes of the cell and diode **D2**. The emitters of transistors **TR1, TR2** and **TR3** are connected to ground. Resistors **R5** and **R6** serve as collector loads for transistors **TR1** and **TR2**, respectively. The cell serves as the collector load for transistor **TR3**. Resistors **R2, R3** and **R4** ensure that transistors **TR1**, **TR2** and **TR3** are saturated. Diode **D2** protects the rest of the circuit from any induced back emf within the cell.

The astable circuit is used to generate a pulse train at a specific time and with a specific mark-space ratio. This pulse train is provided to the base of transistor **TR1** through resistor **R2**. Transistor **TR1** operates as an invert switch. Thus, when the a stable circuit produces an output pulse, the base voltage of the transistor **TR1** goes high (i.e. close to **Vcc** or logic 1). Hence, the voltage level of the collector of transistor **TR1** goes low (i.e., close to ground or logic 0).

Transistor **TR2** also operates as an inverter. When the collector voltage of transistor **TR1** goes low, the base voltage of transistor **TR2** also goes low and transistor **TR2** turns off. Hence, the collector voltage of transistor **TR2** and the base voltage of Transistor **TR3** go high. Therefore, transistor **TR3** turns on with the same mark-space ratio as the astable circuit. When the transistor **TR3** is on, one electrode of the cell is connected to **Vcc** and the other is connected to ground through transistor **TR3**. Thus, the transistor **TR3** can be turned on (and off) and therefore the transistor **TR3** effectively serves as a power switch for the electrodes of the cell.

Figs.6-8 illustrate additional embodiments of the cell which are similar to the embodiments of **Figs.1-3**, respectively. However, each of embodiments of **Figs.6-8** further includes a coil **104** arranged above the electrodes and power supply terminals **107** connected to the coil **104**. The dimensions of coil **104** can be, for example, 5 x 7 cm and have, for example, 1500 turns. The coil **104** is submerged under the surface of the water **110**.

The embodiments of **Figs.6-8** further include an optional switch **121** which can be switched on or off by the user. When the switch **121** is not closed, then the cell forms basically the same structure as **Figs.1-3** and thus can be operated in the same manner described in **Figs.1-3** to produce orthohydrogen and oxygen. When the switch **121** is closed, the additional coil **104** makes the cell capable of producing oxygen and either (1) parahydrogen or (2) a

mixture of parahydrogen and orthohydrogen.

When the switch **121** is closed (or not included), the coil **104** is connected through terminals **106** and the switch **121** (or directly connected only through terminals **106**) to a power supply so that the coil **104** can a receive a pulsed signal. As will be discussed below, this power supply can be formed by the circuit illustrated in **Fig.9**.

When the coil **104** and the electrodes **105a** and **105b** receive pulses, it is possible to produce bubbles of parahydrogen or a mixture of parahydrogen and orthohydrogen. The bubbles are formed and float to the surface of the water **110** as discussed in **Figs.1-3**. When the coil is pulsed with a higher current, a greater amount of parahydrogen is produced. Moreover, by varying the voltage of the coil **104**, a greater/lesser percentage of orthohydrogen/parahydrogen can be produced. Thus, by controlling the voltage level, current level and frequency (discussed below) provided to the coil **104** (and the parameters such as voltage level, current level, frequency, mark-space ratio and waveform provided to the electrodes **105a** and **105b** as discussed above) the composition of the gas produced by the cell can be controlled. For example, it is possible to produce only oxygen and orthohydrogen by simply disconnecting the coil **104**. It is also possible to produce only oxygen and parahydrogen by providing the appropriate pulsed signals to the coil **104** and the electrodes **105a** and **105b**. All of the benefits and results discussed in connection with the embodiments of **Figs.1-3** are equally derived from the embodiments of **Figs.6-8**. For example, the cells of **Figs.6-8** are self-pressurising, require no-chemical catalyst, do not greatly heat the water **110** or cell, and produce a large amount of hydrogen and oxygen gases from a modest amount of input power, without bubbles on the electrodes.

A considerable amount of time must pass before the next pulse provides current to the coil **104**. Hence, the frequency of the pulsed signal is much lower than that provided to the electrodes **105a** and **105b**. Accordingly, with the type of coil **104** having the dimensions described above, the frequency of pulsed signals can be as high as 30 Hz, but is preferably 17-22 Hz to obtain optimum results.

Parahydrogen is not as highly combustible as orthohydrogen and hence is a slower burning form of hydrogen. Thus, if parahydrogen is produced by the cell, the parahydrogen can be coupled to a suitable device such as a cooker or a furnace to provide a source of power or heat with a slower flame.

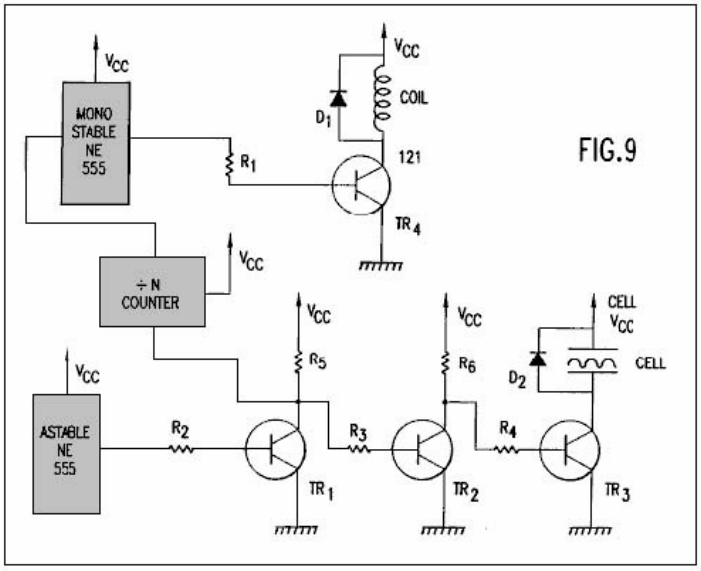


Fig.9 illustrates an exemplary power supply for providing D.C. pulsed signals such as those illustrated in **Figs.4a-4c** to the electrodes illustrated in **Figs.6-8**. Additionally, the power supply can provide another pulsed signal to the coil. As will be readily understood by those skilled in the art, any other power supply which is capable of providing the pulsed signals discussed above to the electrodes of the cell and the coil can be substituted. Alternatively, the pulsed signals provided to the electrodes and the coil can be provided by two separate power supplies.

The portion of the power supply (astable circuit, **R2-R6**, **TR1-TR3**, **D2**) providing a pulsed signal to the electrodes of the cell is identical to that illustrated in **Fig.5**. The power supply illustrated in **Fig.9** further includes the following parts and their respective exemplary values:

The input of the ‘divide-by-N’ counter (hereinafter "the divider") is connected to the collector of transistor **TR1**. The output of the divider is connected to the monostable circuit and the output of the monostable circuit is connected to the base of transistor **TR4** through resistor **R1**. The collector of transistor **TR4** is connected to one end of the coil and a diode **D1**. The other end of the coil and the diode **D1** are connected to the voltage supply

Vcc. Resistor **R1** ensures that **TR4** is fully saturated. Diode **D2** prevents any induced back emf generated within the coil from damaging the rest of the circuit. As illustrated in **Figs.6-8**, a switch **121** can also incorporated into the circuit to allow the user to switch between (1) a cell which produces orthohydrogen and oxygen, and (2) a cell which produces at least parahydrogen and oxygen.

The high/low switching of the collector voltage of transistor **TR1** provides a pulsed signal to the divider. The divider divides this pulsed signal by N (where N is a positive integer) to produce a pulsed output signal. This output signal is used to trigger the monostable circuit. The monostable circuit restores the pulse length so that it has a suitable timing. The output signal from the monostable circuit is connected to the base of transistor **TR4** through resistor **R1** to switch transistor **TR4** on/off. When transistor **TR4** is switched on, the coil is placed between **Vcc** and ground. When the transistor **TR4** is switched off, the coil is disconnected from the rest of the circuit. As discussed in conjunction with **Figs.6-8**, the frequency of pulse signal provided to the coil is switched at a rate preferably between 17-22 Hz; i.e., much lower than the frequency of the pulsed signal provided to the electrodes.

As indicated above, it is not required that the circuit (divider, monostable circuit, **R1, TR4** and **D1**) providing the pulsed signal to the coil be connected to the circuit (astable circuit, **R2-R6**, **TR1-TR3, D2**) providing the pulsed signal to the electrodes. However, connecting the circuits in this manner provides an easy way to initiate the pulsed signal to the coil.

A working prototype of the present invention has been successfully built and operated with the exemplary and optimal parameters indicated above to generate orthohydrogen, parahydrogen and oxygen from water. The output gas from the prototype has been connected by a tube to the manifold inlet of a small one cylinder gasoline engine, with the carburettor removed, and has thus successfully run such engine without any gasoline:

CHARLES GARRETT

US Patent 2,006,676 2nd July 1935 Inventor: Charles H. Garrett

ELECTROLYTIC CARBURETTOR

Please note that this is a re-worded excerpt from this patent. It describes an electrolyser which Charles claimed was able to generate enough gas from hydrolysis of water, to be able to run a car engine without the use of any other fuel. It should be remembered that in Garrett's day, car electrics were all 6-volt systems.

DESCRIPTION

This invention relates to carburettors and it has particular reference to an electrolytic carburettor by means of which water may be broken up into its hydrogen and oxygen constituents and the gases so formed suitably mixed with each other and with air.

Another object of the invention is to provide a means whereby the electrolyte level in the carburettor may be maintained at a more or less constant level regardless of fluctuations in water pressure at the water inlet of the carburettor.

Another object of the invention is to provide a means whereby the relative amount of air mixed with the hydrogen and oxygen may be regulated as desired.

Still another object of the invention is the provision of a means to prevent the loss of hydrogen and oxygen gases during periods in which these gases are not being drawn from the carburettor.

Still another object of the invention is the provision of a means whereby the hydrogen and oxygen resulting from electrolysis may be formed in separate compartments, and a further object of the invention is the provision of a means to periodically reverse the direction of current flow and thereby alternate the evolution of the gases in the separate compartments, to be intermingled at a later time.

With reference to the accompanying drawings: -

Figure 1 is a view in vertical section of one form of carburettor.

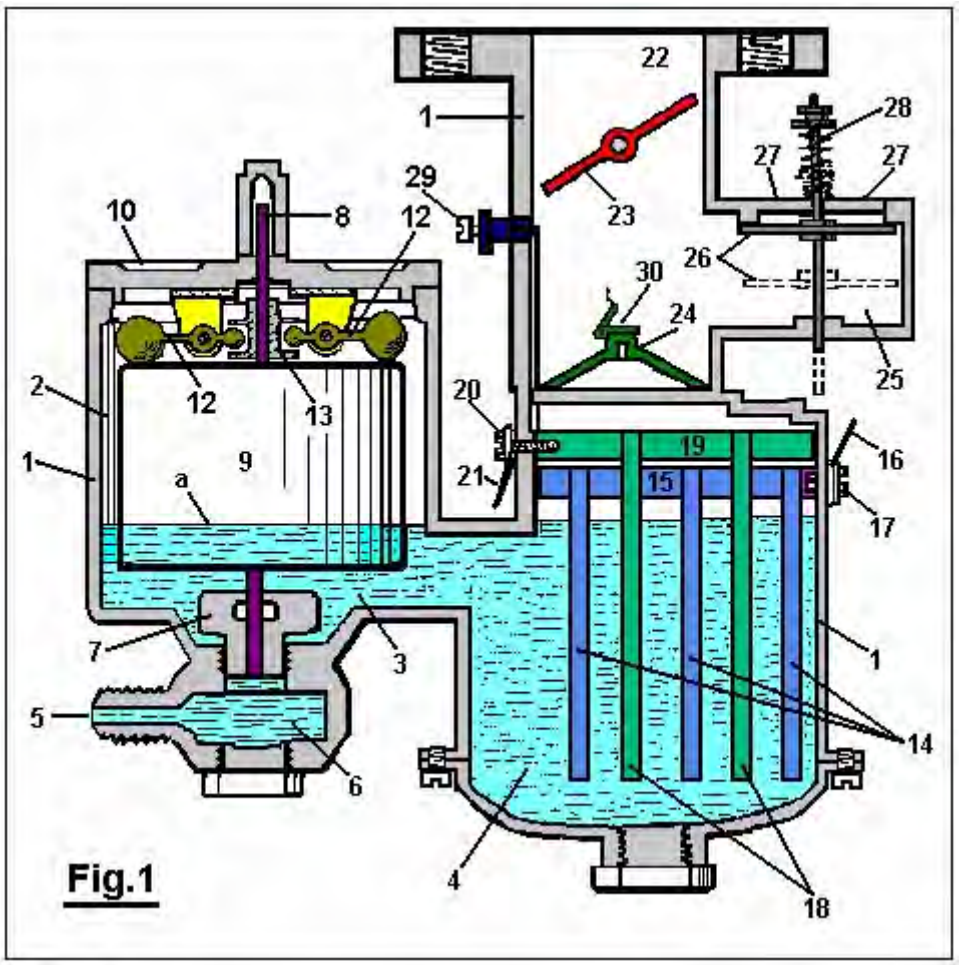


Figure 2 is a modified form.

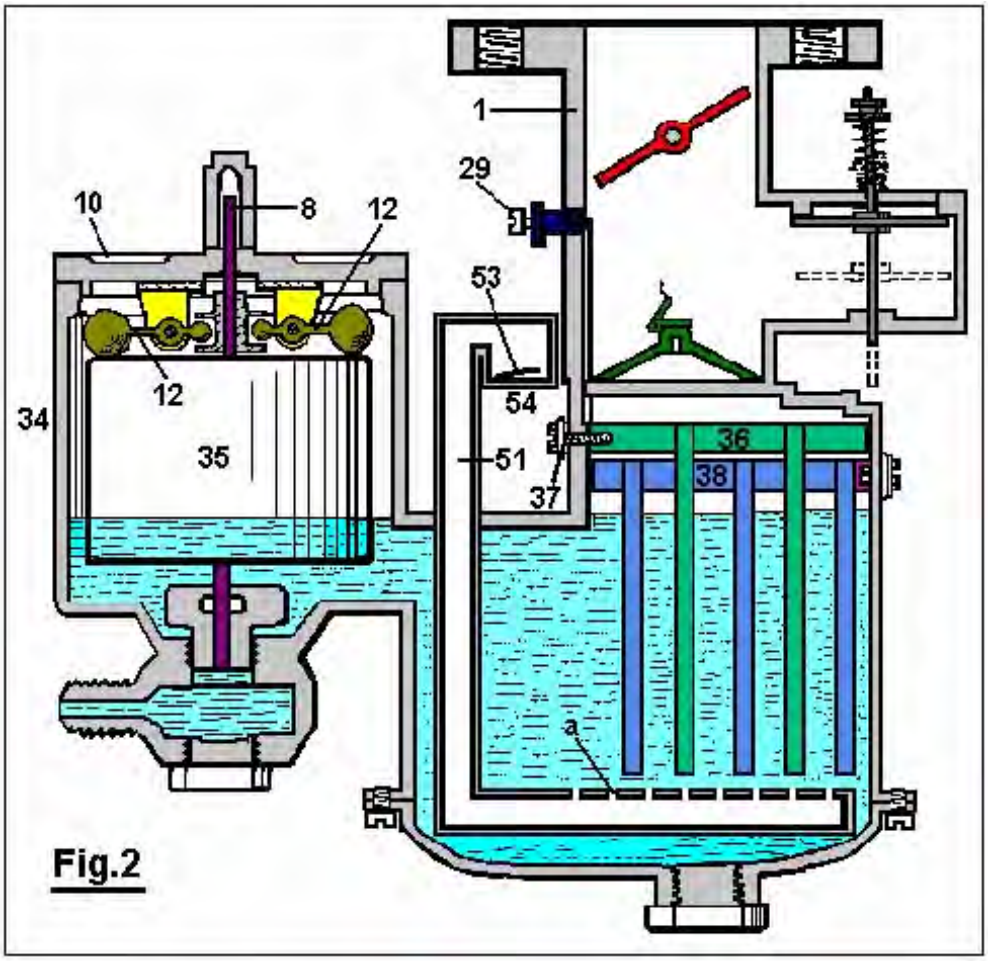


Figure 3 is a diagrammatic view of a pole changer, showing its actuating mechanism, and

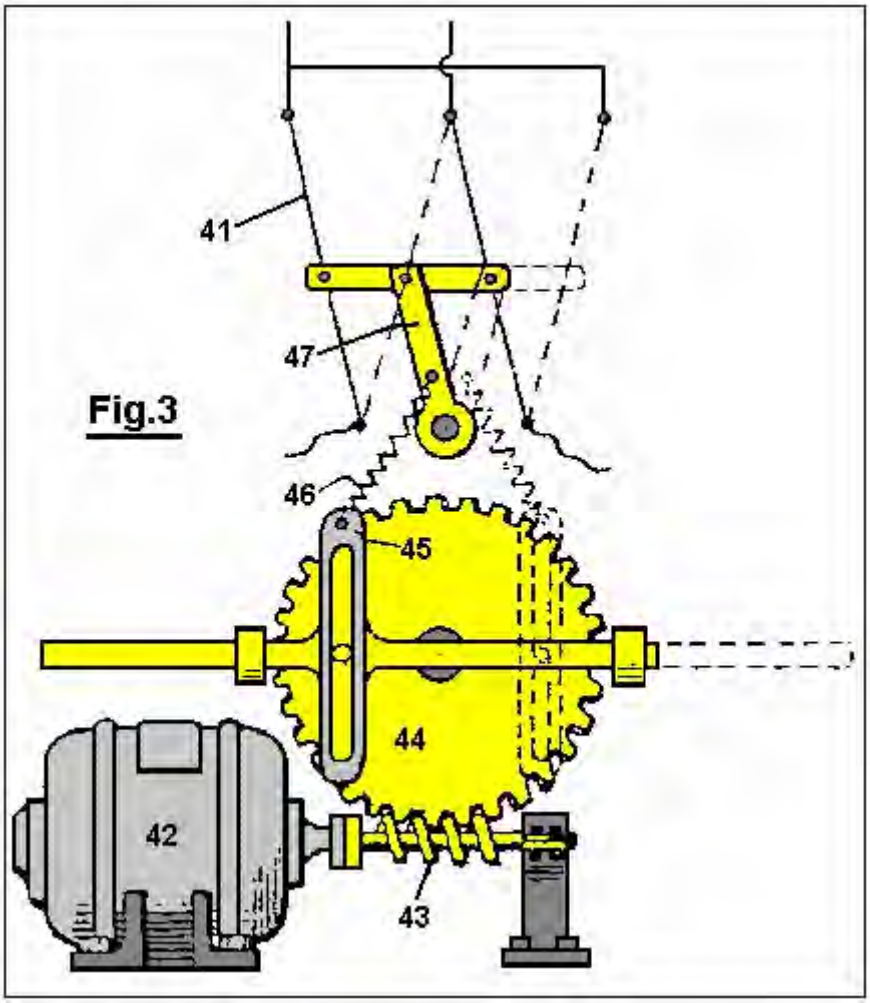


Figure 4 is a wiring diagram for the modified form of carburettor shown in Figure 2.

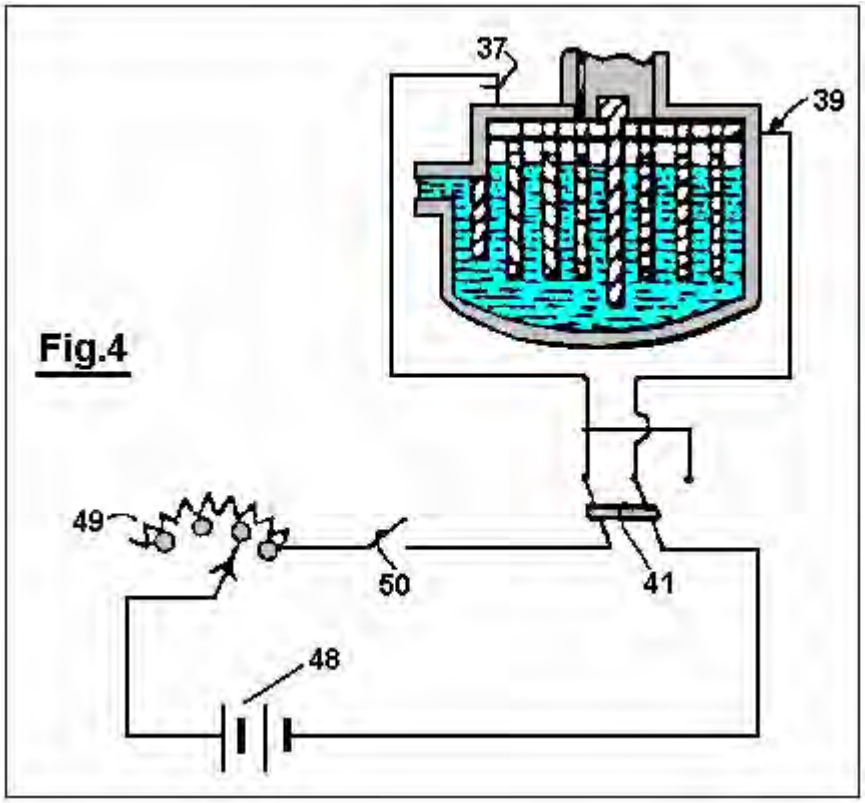


Fig.4

With reference to **Fig.1**: The reference numeral **1** designates the carburettor housing, which is preferably constructed of bakelite or other suitable insulating material. This housing is designed so as to divide the carburettor into a float chamber **2** and gas generating chamber **4**, connected by a fluid passage **3**.

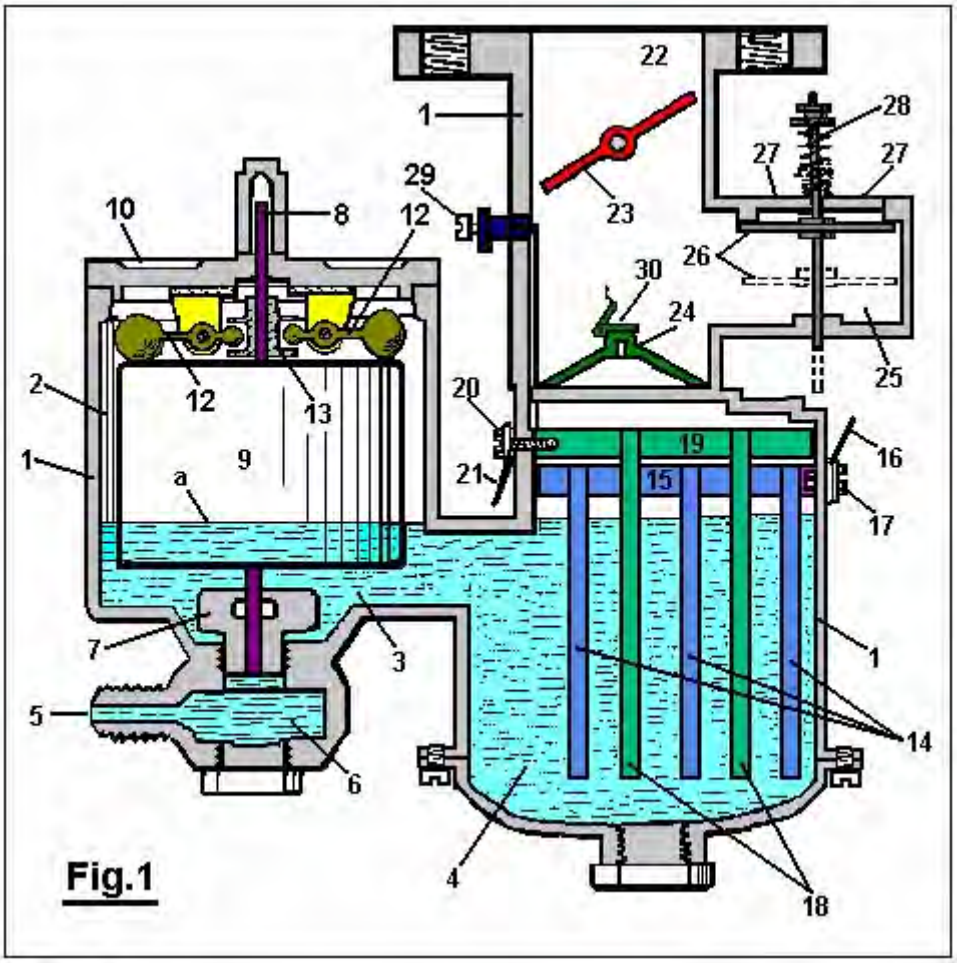


Fig.1

Water under pressure is forced into the carburettor through an opening **5** which communicates with the float chamber **2** through the medium of the sediment chamber **6** and the needle valve orifice **7**, which is closed by a needle valve **8** when the device is not in operation. A float **9** surrounds the needle valve **8** and is free to move vertically relative thereto. Descending from the cover **10** to the float chamber **2** are two ears **11**, located at spaced intervals on opposite sides of the needle valve **8**. The members **12** are pivoted to the ears **11**, as shown. The weighted outer ends of the members **12** rest on top of the float **9**, and their inner ends are received in an annular groove in the collar **13** which is rigidly attached to the needle valve **8**.

Within the gas generating chamber **4**, a series of spaced, descending plates **14** are suspended from a horizontal member **15** to which a wire **16** has electrical contact through the medium of the bolt **17**, which extends inwards through housing **1** and is threaded into the horizontal member **15**.

A second series of plates **18** is located between the plates **14** and attached to the horizontal member **19**, and has electrical contact with the wire **20** through the bolt **21**.

A gas passageway **22**, in which a butterfly valve **23** is located, communicates with the gas generating chamber **4** through an orifice **24**. An air inlet chamber **25** has communication with the gas passageway **22** above the orifice **24**. A check valve **26** which opens downwards, controls the openings **27**, and is held closed and inoperative by means of light spring **28**.

An adjustable auxiliary air valve **29** is provided in the wall of the gas passageway **22**, which air valve is closed by the butterfly valve **23** when the butterfly valve is closed, but communicates with the outside air when the butterfly valve is open.

The operation of the device is as follows :

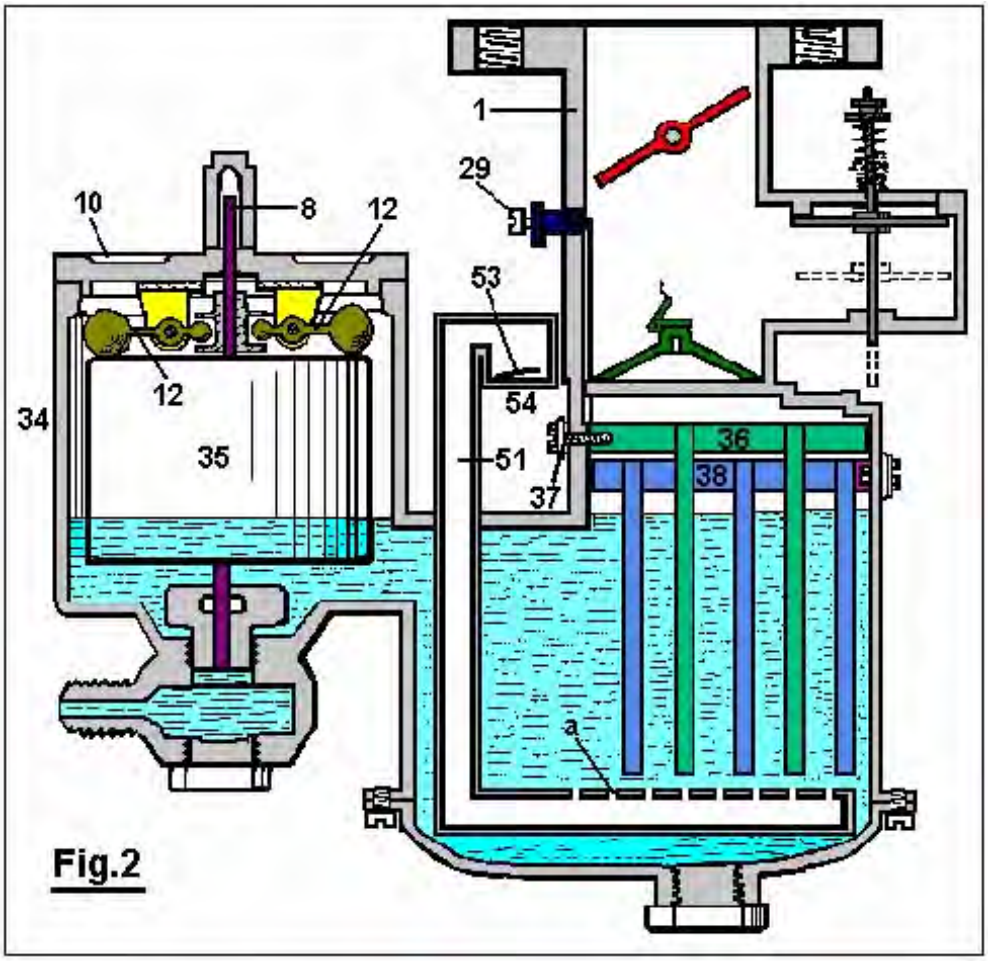
The chambers **2** and **4** are first filled to the level **'a'** with a solution of weak sulphuric acid (or other electrolyte not changed by the passage of current through it), and the opening **5** is connected to a tank of water (not shown).

The wire **16** is next connected to the positive pole of a storage battery or other source of direct current and the wire **20** to the negative pole. Since the solution within the carburettor is a conductor of electricity, current will flow through it and hydrogen will be given off from the negative or cathode plates **18** and oxygen from the positive or anode plates **14**.

The butterfly valve **23** is opened and the gas passageway **22** brought into communication with a partial vacuum. Atmospheric pressure acting on the top of the check valve **26** causes it to be forced downwards as shown in dotted lines. The hydrogen and oxygen liberated from the water at the plates **18** and **14** are drawn upwards through the orifice **24** covered by the check valve **30** where they are mixed with air entering through the openings **27** and through the auxiliary air valve **29**.

When it is desired to reduce the flow of hydrogen and oxygen from the plates **18** and **14**, the current flowing through the device is reduced, and when the current is interrupted the flow ceases. When the butterfly valve **23** is moved to its ‘closed’ position, the check-valve **26** is automatically closed by the spring **28**. Any excess given off during these operations is stored in the space above the fluid where it is ready for subsequent use.

Water is converted into its gaseous constituents by the device herein described, but the dilute sulphuric acid or other suitable electrolyte in the carburettor remains unchanged, since it is not destroyed by electrolysis, and the parts in contact therewith are made of bakelite and lead or other material not attacked by the electrolyte.



The structure shown in **Fig.2** is substantially the same as that shown in **Fig.1** with the exception that the modified structure embraces a larger gas generating chamber which is divided by means of an insulating plate **31** and is further provided with a depending baffle plate **32** which separates the gas generating chamber **33** from the float chamber **34** in which the float **35** operates in the same manner as in **Fig.1**. Moreover, the structure shown in **Fig.2** provides a series of spaced depending plates **36** which are electrically connected to the wire **37**, and a second series of similar plates **38** which are electrically connected to the wire **39** and are kept apart from the plates **36** by the insulating plate **31**.

Gases generated on the surfaces of the plates **36** and **38** pass upward through the orifice **39a** into the gas passageway **40** where they are mixed with air as explained in the description of **Fig.1**.

A pipe **51,** bent as shown in **Fig.2**, passes downwards through the housing of the carburettor and has a series of spaced apertures **'a'** in its horizontal portion beneath the plates **36** and **38**. Check valve **53**, with opens upwards, controls air inlet **54**. When a partial vacuum exists in the chamber **33**, air is drawn in through the opening **54** and then passes upwards through the apertures **'a'**. This air tends to remove any bubbles of gas collecting on the plates **36** and **38** and also tends to cool the electrolyte. The check valve **53** automatically closes when a gas pressure exists within the carburettor and thereby prevents the electrolyte from being forced out of the opening **54**.

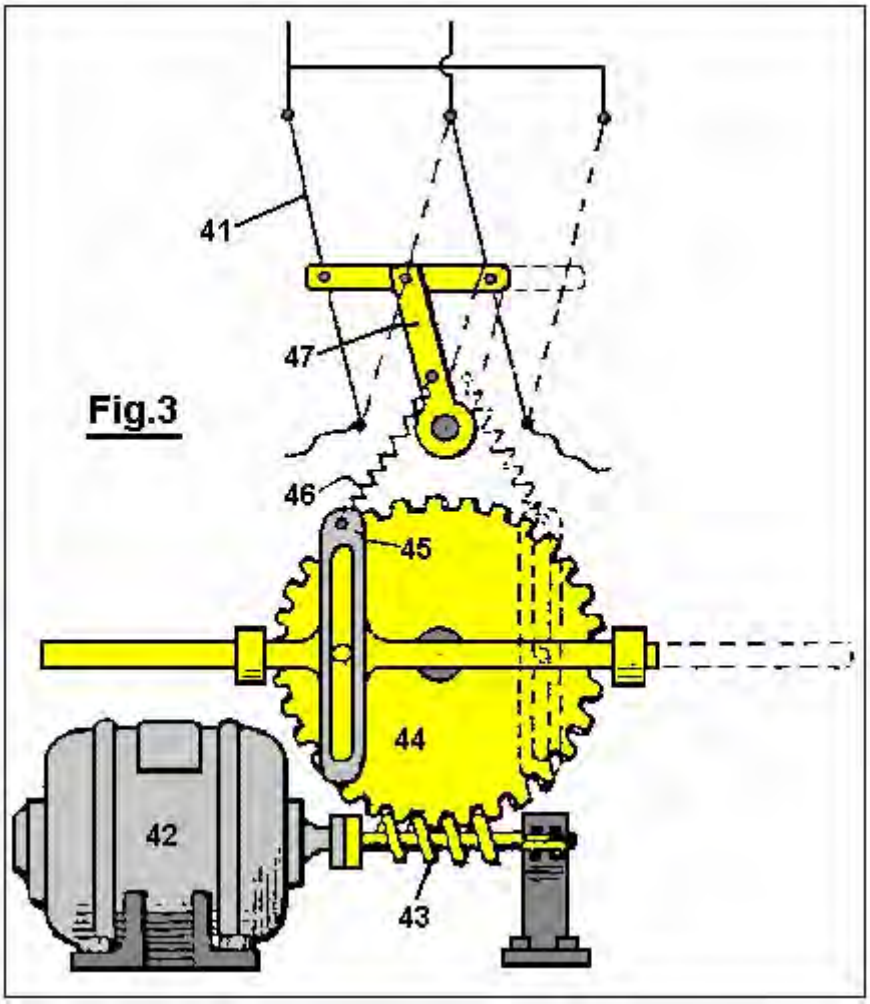


Fig.3

In order to provide for alternate evolution of the gases from the plates **36** and **38**, a pole changer **41**, shown in **Fig.3**, is actuated periodically by the motor **42** which drives the worm **43** and the gear **44** and causes oscillations of the member **45** which is connected by a spring **46** to the arm **47**, thereby causing the pole changer to snap from one position to the other.

In operation, the carburettor shown in **Fig.2** is connected as shown in the wiring diagram of **Fig.4**. A storage battery **48** or other suitable source of direct current is connected to a variable rheostat **49**, switch **50**, pole changer **41** and to the carburettor as shown. Thus the rate of evolution of the gases can be controlled by the setting of the rheostat **49** and the desired alternate evolution of the gases in the compartments of the carburettor is accomplished by means of the periodically operated pole changer **41**.

Manifestly, the construction shown is capable of considerable modification and such modification as is considered within the scope and meaning of the appended claims is also considered within the spirit and intent of the invention.

ARCHIE BLUE

US Patent 4,124,463 7th November 1978 Inventor: Archie H. Blue

ELECTROLYTIC CELL

Please note that this is a re-worded excerpt from this patent. It describes an electrolyser system where air is drawn through the electrolyte to dislodge bubbles from the electrodes.

ABSTRACT

In the electrolytic production of hydrogen and oxygen, air is pumped through the cell while the electrolysis is in progress so as to obtain a mixture of air, hydrogen and oxygen.

BACKGROUND AND BRIEF DESCRIPTION OF THE INVENTION

This invention relates to the production of gases which can be utilised primarily, but not necessarily, as a fuel.

To decompose water electrically, it is necessary to pass direct current between a pair of electrodes which are immersed in a suitable electrolyte. During such electrolysis, it is normal to place some form of gas barrier between the two electrodes, in order to prevent the gases produced forming an explosive mixture. However provided suitable precautions are taken, it has been found that the gases can be allowed to mix and can be fed into a storage tank for subsequent use. Because the gases when mixed form an explosive mixture, it is possible for the mixture to be utilised, for instance, as a fuel for an internal combustion engine. In such circumstances it is desirable that the gases should also be mixed with a certain proportion of air in order to control the explosive force which results when the gases are ignited.

One of the difficulties encountered with electrolysis is that bubbles of gas are liable to remain on the electrodes during the electrolysis thus effectively limiting the area of electrode which is in contact with the electrolyte and preventing optimum current flow between the electrodes. Because it is desirable that the gases evolved during the electrolysis be mixed with air, it is possible for air to be passed through the cell while electrolysis is in progress. The passage of air through the cell can be directed past the electrodes so as to pick up any gas bubbles on the electrodes.

Accordingly, the invention comprises an electrolytic cell with a gas tight casing, several electrodes supported on a central post within the cell, spaced apart and electrically insulated from each other, each alternative electrode being connected to a positive direct current source or a negative direct current source respectively and wherein the central post is in the form of a tube, one end of which is extended out of the cell and connected to a source of air under pressure, with the other end of the central post terminating in an air outlet below the electrodes. The cell also includes a gas outlet to carry the air forced into the cell through the central post and to exhaust the gases produced by electrolysis.

DETAILED DESCRIPTION OF THE INVENTION

Various forms of the invention will now be described with the aid of the accompanying drawings wherein:

Fig.1 is a diagrammatic elevational view partly in section of one form of the invention,

FIG. 1

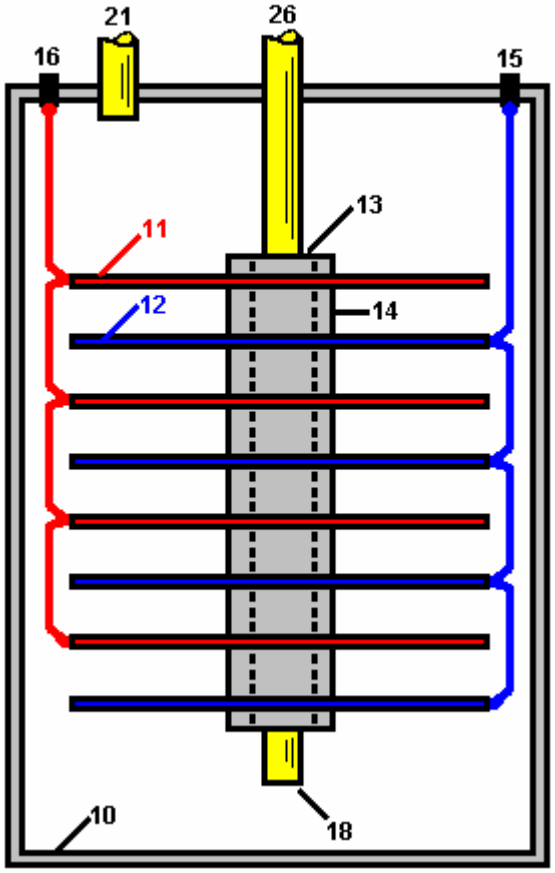


Fig.2 is a diagrammatic elevational view partly in section of a modified form of the invention,

FIG. 2

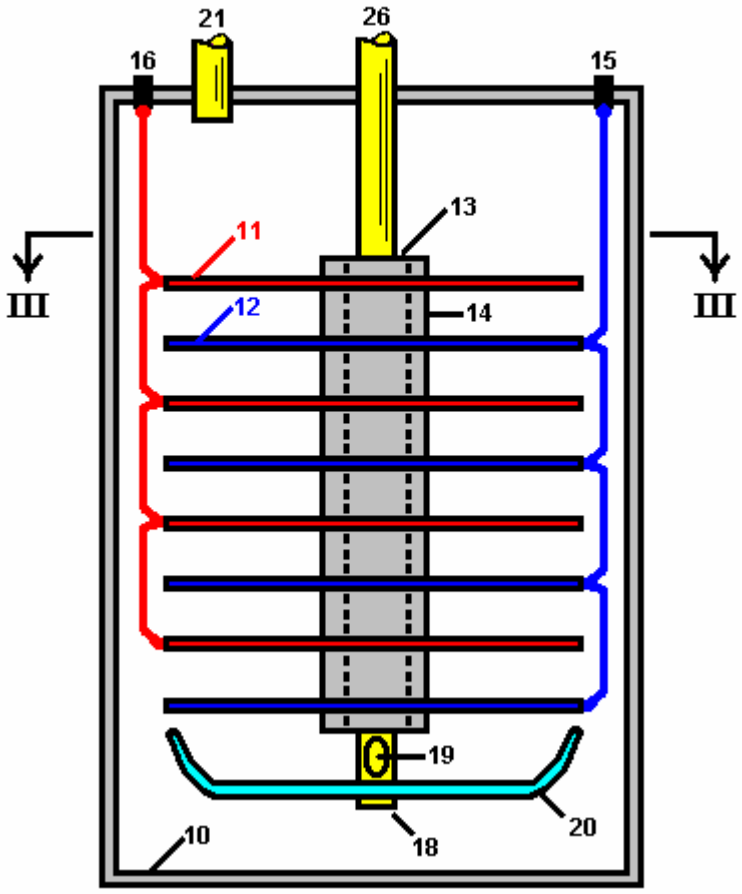
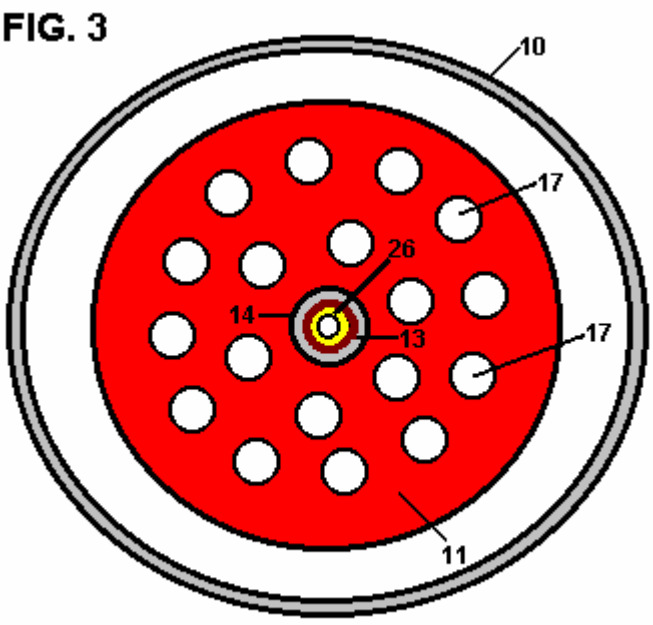


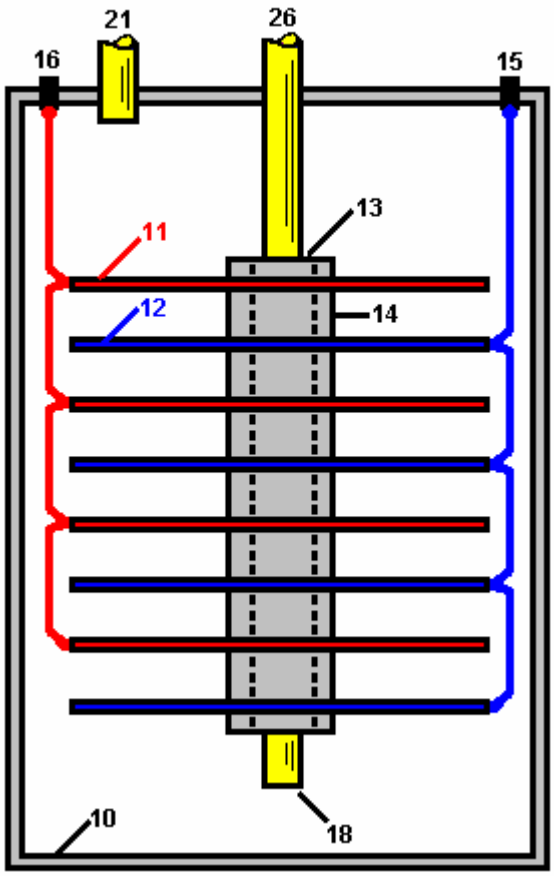
Fig.3 is a section along the line **III--III** of **Fig.2**.

FIG. 3



Section III - III

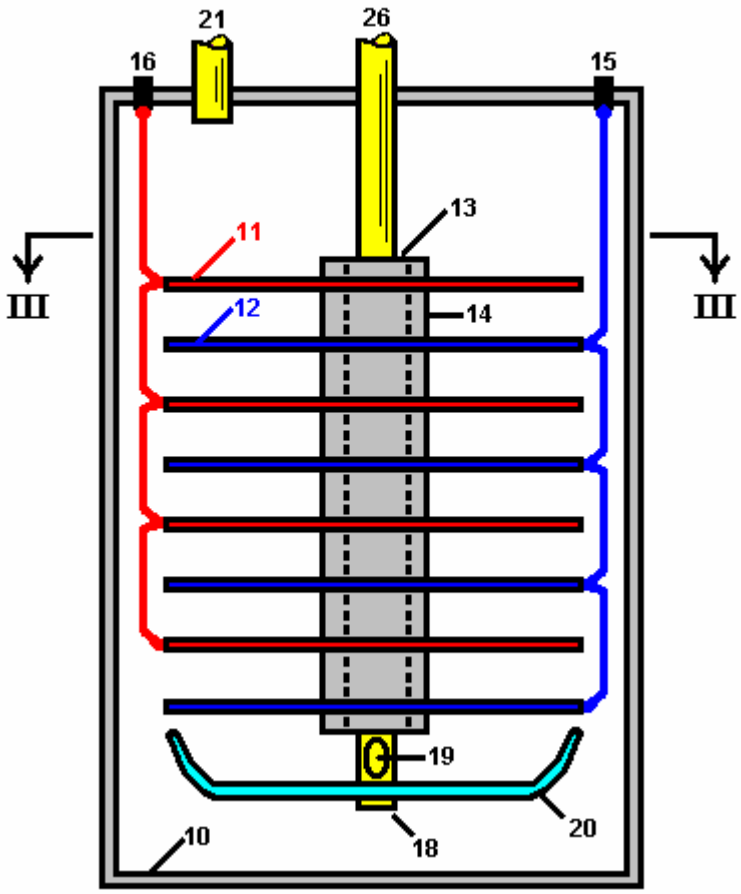
FIG. 1



The cell as shown in **Fig.1** comprises a gas-tight casing **10** which is formed from a material incapable of corrosion, such as plastic. Several cathode plates **11** and several anode plates **12** are supported within the cell on an electrically insulating central post **13**, with the cathode plates and anode plates being spaced apart by means of insulating spacers **14**. The anode plates **12** are all connected in parallel to a positive terminal post **15** while the cathode plates are all connected in parallel to the negative terminal post **16**, these connections being indicated in dotted lines in the drawings. The cathode and anode plates are preferably in the form of discs made from a metal suited to the electrolyte, thus ensuring a satisfactory cell life. These plates may be shaped to conform with the shape of the walls of the cell which may be circular in cross section as indicated or any other desired shape.

The central post **26** is preferably in the form of a tube which extends out of the cell. The lower end of the tube **18** is open so that air can be pumped into the cell through the central post **26** and enter the cell via the lower end **18** where it will pass up through the electrolyte. This keeps the electrolyte in constant motion which assists in the rapid removal of any gas bubbles which may be adhering to the electrode plates.

FIG. 2



In the modification shown in **Fig.2** and **Fig.3**, each electrode plate is provided with holes **17.** The central post **26** is also provided with at least one air hole **19** adjacent to its the lower end. A deflector plate **20** is also supported by the central post **26**, this plate being dish shaped so as to deflect air issuing out of the air hole **19** up through the holes **17** in the electrodes. This further assists in dislodging any bubbles of gas clinging to the electrode plates.

The cell also includes a gas outlet **21** so that the air which enters the cell, together with the gases produced by electrolysis, can be taken out of the cell into a suitable storage tank (not shown in the drawings). If desired, such storage tank can be arranged to accept the gases under pressure and for this purpose the air pumped into the cell will be pumped in under the required pressure. A gas drier (not shown in the drawings) can also be interposed between the gas outlet **21** and the storage tank.

Although the electrolysis will naturally produce considerable heat, nevertheless it can be found advantageous to install a heater in the cell, preferably in the bottom of the cell, to assist and facilitate the warming up of the electrolyte so that the cell reaches its most efficient operating conditions as quickly as possible.

Preferably also, a current-control device should be employed so that the intensity of the electrolytic action can be controlled.

A mechanism may also be provided for the automatic replenishment of water within the cell as the level of the electrolytic drops during use.

While it is recognised that the mixing of hydrogen and oxygen will create a dangerous explosive mixture, nevertheless by carrying out the invention as described above, the risk of explosion is minimised. The gases produced can be utilised, for instance, as a fuel to power an internal combustion engine and for this purpose it is desirable, as already mentioned, to mix a proportion of air with the gases produced during electrolysis, so that when the mixture is ignited within the cylinder or cylinders of the engine, the explosive force so created can be of the desired amount.

While in the foregoing description reference is made to the utilisation of the mixed gases as a fuel, it will of course be understood that the gases can be separated for individual use.

CLAIMS

1. A process for producing, Through the electrolysis of an aqueous liquid, a combustible mixture of hydrogen, oxygen and air. This is achieved in an electrolytic cell having a gas-tight casing, a substantially central tubular post mounted in the casing and having an air inlet at its upper end, and a several electrodes supported on the post and axially spaced along it, alternate electrodes being connected to a first electrical terminal and to a second electrical terminal respectively connected to a respective poles of a current source and being mutually insulated, the post having an air outlet below the electrodes out of which flows air from the air inlet into the cell and over the electrodes; and a source of air under pressure connected to the said air inlet forcing a flow of air through the aqueous liquid contained in the cell; the cell having in its upper region a common outlet exhausting the combustible mixture comprising air forced through the cell, along with hydrogen and oxygen produced by electrolysis in the cell.

2. The process according to claim 1 wherein the electrodes are discs each having a several holes through them.

3. The process according to claim 1 further including a dish-shaped air deflector plate supported on the post below the air outlet.

4. Apparatus for producing by electrolysis of an aqueous liquid, a combustible mixture of hydrogen and oxygen, comprising: an electrolytic cell having a gas-tight casing, a substantially central tubular post mounted in the casing and having an air inlet at its upper end, and a plurality of electrodes supported on the post and axially spaced along it, alternate electrodes being connected to a first electrical terminal and to a second electrical terminal respectively for connection to respective poles of a current source and being mutually insulated, the post having an air outlet below the electrodes for flow of air from the air inlet into the cell and over the electrodes; a dish-shaped air deflector supported on said post below said air outlet; and a source of air under pressure connected to the said air inlet for forcing a flow of air through the aqueous liquid contained in the cell in operation thereof; the cell having in its upper region a common outlet for exhausting the combustible mixture comprising air forced through the cell and hydrogen and oxygen produced by electrolysis of the liquid in the cell.

5. The apparatus according to claim 4 wherein the electrodes are discs each having a several holes through them.

RUGGERO SANTILLI

US Patent 6,183,604 6th February 2001 Inventor: Ruggero M. Santilli

Durable and Efficient Equipment for the Production of a Combustible and Non-Pollutant Gas from Underwater Arcs and Method therefor

Please note that this is a re-worded excerpt from this patent. It shows how electrolysis of water can be carried out on a large scale as a continuous process.

ABSTRACT

A system for producing a clean burning combustible gas comprising an electrically conductive first electrode and an electrically conductive second electrode. A motor coupled to the first electrode is adapted to move the first electrode with respect to the second electrode to continuously move the arc away from the plasma created by the arc. A water-tight container for the electrodes is provided with a quantity of water within the tank sufficient to submerge the electrodes.

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention relates to durable and efficient equipment for the production of a combustible and non-polluting gas from underwater arcs and the method for doing this and more particularly, the invention pertains to producing a combustible gas from the underwater arcing of electrodes which move with respect to each other.

2. Description of the Prior Art

The combustible nature of the gas bubbling to the surface from an underwater welding arc between carbon electrodes was discovered and patented in the last century. Various improved equipment for the production of said combustible gas have been patented during this century. Nevertheless, the technology has not yet reached sufficient maturity for regular industrial and consumer production and sales because of numerous insufficiencies, including excessively short duration of the carbon electrodes which requires prohibitive replacement and service, as well as low efficiency and high content of carbon dioxide responsible for the greenhouse effect. As a result of numerous experiments, this invention deals with new equipment for the production of a combustible gas from underwater arcs between carbon electrodes which resolves the previous problems, and achieves the first known practical equipment for industrial production and sales.

The technology of underwater electric welding via the use of an arc between carbon electrodes to repair ships, was established in the last century. It was then discovered that the gas bubbling to the surface from underwater arcs is combustible. In fact, one of the first U.S. patents on the production of a combustible gas via an underwater electric arc between carbon electrodes dates back to 1898 (U.S. Pat. No. 603,058 by H. Eldridge).

Subsequently, various other patents were obtained in this century on improved equipment for the production of this combustible gas, among which are:

US Pat. No. 5,159,900 (W.A. Dammann abd D. Wallman, 1992); U.S.Pat. No. 5,435,274 (W. H. Richardson, Jr., 1995); U.S. Pat. No. 5,417,817 (W. A. Dammann and D. Wallman, 1995); U.S. Pat. No. 5,692,459 (W. H. Richardson, Jr., 1997); U.S. Pat. No. 5,792,325 (W. H. Richardson, Jr., 1998); and U.S. Pat. No. 5,826,548 (W. H. Richardson, Jr., 1998).

The main process in these inventions is essentially the following. The arc is generally produced by a DC power unit, such as a welder, operating at low voltage (25-35 V) and high current (300 A to 3,000 A) depending on the available Kwh input power. The high value of the current brings the tip of the carbon electrode in the cathode to incandescence, with the consequential disintegration of the carbon crystal, and release of highly ionised carbon atoms to the arc. Jointly, the arc separates the water into highly ionised atoms of Hydrogen and Oxygen. This causes a high temperature plasma in the immediate surrounding of the arc, of about 7,000^{o}F, which is composed of highly ionised H, O and C atoms.

A number of chemical reactions then occur within or near the plasma, such as: the formation of the H_2O_2 molecule; the burning of H and O into H_2O; the burning of C and O into CO; the burning of CO and O into CO_2, and other reactions. Since all these reactions are highly exothermic, they result in the typical, very intense glow of the arc within water, which is bigger than that of the same arc in air. The resulting gases cool down in the water surrounding the discharge, and bubble to the surface, where they are collected with various means. According to numerous measurements conducted at various independent laboratories, the combustible gas produced with the above process essentially consists of 45%-48% H_2, 36%-38% CO, 8%-10% CO_2, and 1%-2% O_2, the remaining gas consisting of parts per million of more complex molecules composed by H, O and C.

This process produces an excellent combustible gas because the combustion exhausts meet all current EPA requirement without any catalytic converter at all, and without the highly harmful carcinogenic pollutants which are contained in the combustion exhausts of gasoline, diesel, natural gas and other fuels of current use.

Despite the indicated excellent combustion characteristics, and despite research and development conducted by inventors for decades, the technology of the combustible gas produced by an underwater arc between carbon electrodes has not reached industrial maturity until now, and no equipment producing said combustible gas for actual practical usages is currently sold to the public in the U.S.A. or abroad, the only equipment currently available for sale being limited to research and testing. The sole equipment currently sold for public use produce different gases, such as Brown's gas which is not suitable for use in internal combustion engines because it implodes, rather than explodes, during combustion.

The main reason for lack of industrial and consumer maturity is the excessively short duration of the carbon electrodes, which requires prohibitive replacement and services. According to extensive, independently supervised, and certified measurements, the electrodes are typically composed of solid carbon rods of about 3/8 inch (9 mm) in diameter and about 1 foot length. Under 14 Kwh power input, said electrodes consume at the rate of about one and one quarter inch (32 mm) length per minute, requiring the halting of the operation, and replacement of the electrodes every ten minutes.

The same tests have shown that, for 100 Kwh power input, said electrodes are generally constituted by solid carbon rod of about 1 inch diameter and of the approximate length of one foot, and are consumed under a continuous underwater arc at the rate of about 3 inch length per minute, thus requiring servicing after 3 to 4 minutes of operation. In either case, current equipment requires servicing after only a few minutes of usage, which is unacceptable on industrial and consumer grounds for evident reasons, including increased risks of accidents for very frequent manual operations in a piece of high current equipment.

An additional insufficiency of existing equipment is the low efficiency in the production of said combustible gas, which efficiency will from now on be referred to as the ratio between the volume of combustible gas produced in cubic feet per hour (cfh) and the real input power per hour (Kwh). For instance extensive measurements have established that pre-existing equipment has an efficiency of 2-3 cfh/Kwh. Yet another insufficiency of existing equipment is the high carbon dioxide content in the gas produced. Carbon dioxide is the gas responsible for the greenhouse effect. In fact, prior to combustion the gas has a CO_2 content of 8%-10% with a corresponding content after combustion of about 15% CO_2, thus causing evident environmental problems.

SUMMARY OF THE INVENTION

In view of the foregoing disadvantages inherent in the known types of traditional equipment for the production of combustible and non-polluting gases now present in the prior art, the present invention provides improved durable and efficient equipment for the production of a combustible and non-polluting gas from underwater arcs and the method of production.

As such, the general purpose of the present invention, which will be described later in greater detail, is to provide new, improved, durable and efficient equipment for the production of a combustible and non-polluting gas from underwater arcs and the method for achieving this, a method which has all the advantages of the prior art and none of the disadvantages.

To attain this, the present invention essentially comprises of a new and improved system for producing a clean burning combustible gas from an electric arc generating plasma under water. First provided is an electrically conductive anode fabricated of tungsten. The anode is solid in a generally cylindrical configuration with a diameter of about one inch and a length of about three inches. Next provided is a generally Z-shaped crank of a electrically conductive material. The crank has a linear output end supporting the anode. The crank also has a linear input end essentially parallel with the output end. A transverse connecting portion is located between the input and output ends.

An electrically conductive cathode is next provided. The cathode is fabricated of carbon. The carbon is in a hollow tubular configuration with an axis. The cathode has a supported end and a free end. The cathode has a length of about 12 inches and an internal diameter of about 11.5 inches and an external diameter of about 12.5 inches. A motor is next provided. The motor has a rotatable drive shaft. The drive shaft has a fixed axis of rotation. The motor is coupled to the input end of the crank and is adapted to rotate the crank to move the output end and anode in a circular path of travel. The circular path of travel has a diameter of about twelve inches with the anode located adjacent to the free end of the cathode. In this manner the anode and the arc are continuously moved around the cathode and away from the plasma created by the arc.

Next provided is an axially shifted support. The support is in a circular configuration to receive the supported end of the cathode and to move the cathode axially toward the anode as the carbon of the cathode is consumed during operation and use. Next provided is a water tight container for the anode, cathode, crank and support. A quantity of water is provided within the tank, sufficient to submerge the anode and the cathode. Next provided is an entrance port in the container. The entrance port functions to feed water and a carbon enriched fluid into the container to supplement the carbon and water lost from the container during operation and use. Next provided is a source of potential. The source of potential couples the anode and the cathode. In this manner an electrical arc is created between the anode and the cathode with a surrounding plasma for the production of gas within the water. The gas will then bubble upwards and collect above the water. Last provided is an exit port for removing the gas which results from the application of current from the source of potential to the anode and the cathode while the anode is rotating and the cathode is shifting axially.

This broad outline indicates the more important features of the invention in order that the detailed description which follows may be better understood and in order that the present contribution to the art may be better appreciated. There are, of course, additional features of the invention that will be described and which will form the subject matter of the claims made.

In this respect, before explaining at least one embodiment of the invention in detail, it is to be understood that the invention is not limited in its application to the details of construction and to the arrangements of the components set forth in the following description or illustrated in the drawings. The invention is capable of other embodiments and of being practised and carried out in various ways. Also, it is to be understood that the phraseology and terminology employed here are for the purpose of descriptions and should not be regarded as limiting the scope of this invention.

It is another object of the present invention to provide new and improved durable and efficient equipment for the production of a combustible and non-polluting gas from underwater arcs and method therefor which may be easily and efficiently manufactured and marketed on a commercial basis.

Lastly, it is an object of the present invention to provide a new and improved system for producing a clean burning combustible gas comprising an electrically conductive first electrode, an electrically conductive second electrode, a motor coupled to the first electrode and adapted to move the first electrode with respect to the second electrode to continuously move the arc away from the plasma created by the arc, and a water-tight container for the electrodes with a quantity of water within the tank sufficient to submerge the electrodes.

These together with other objects of the invention, along with the various novel features which characterise the invention, are pointed out particularly in the claims section of this disclosure. For a better understanding of the invention, its operating advantages and the specific objects attained by its uses, reference should be made to the accompanying drawings and descriptive matter in which there is illustrated preferred embodiments of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

The invention will be better understood and objects other than those set forth above will become apparent when consideration is given to the following detailed description thereof. Such description makes reference to the annexed drawings wherein:

Fig.1 and **Fig.2** are illustrations of prior art equipment for the fabrication of a pollutant-free combustible gas produced by an electric arc under water constructed with prior art techniques.

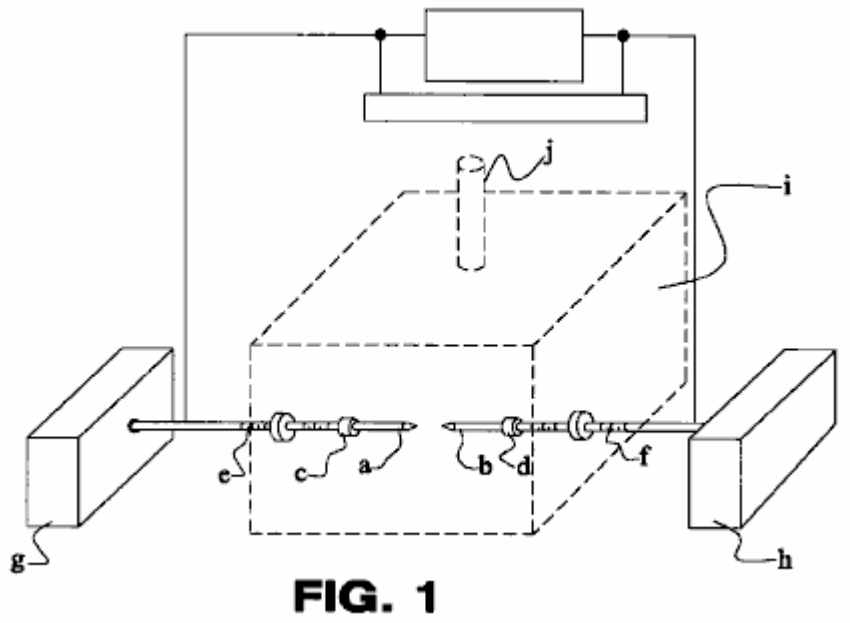


FIG. 1

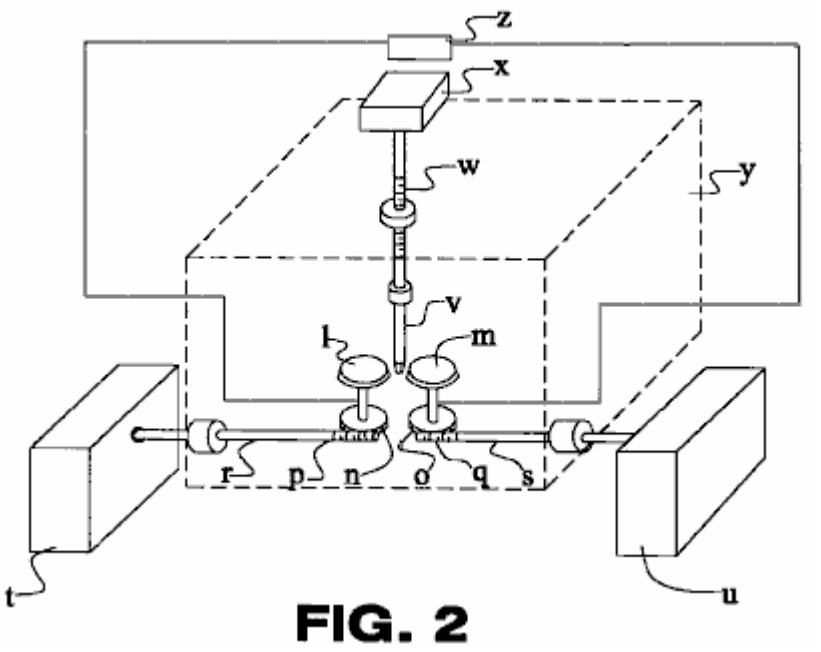


FIG. 2

Fig.3 is a schematic diagram depicting the principles of the present invention.

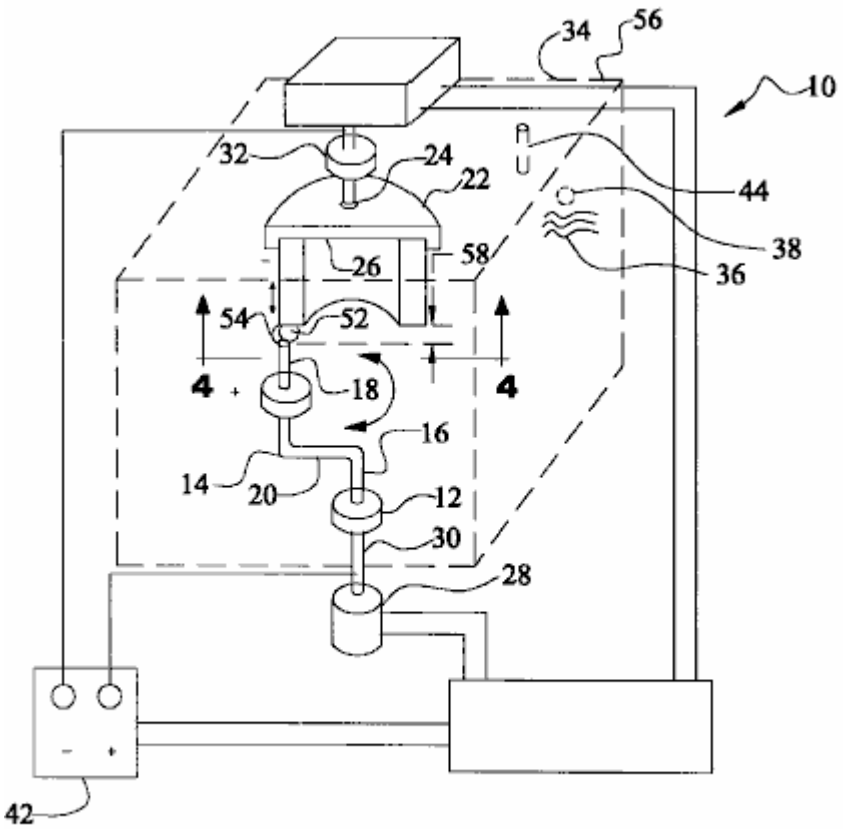


FIG. 3

Fig.4 is a schematic diagram of a partial sectional view taken along line 4--4 of Fig.3, depicting an additional embodiment of the present invention.

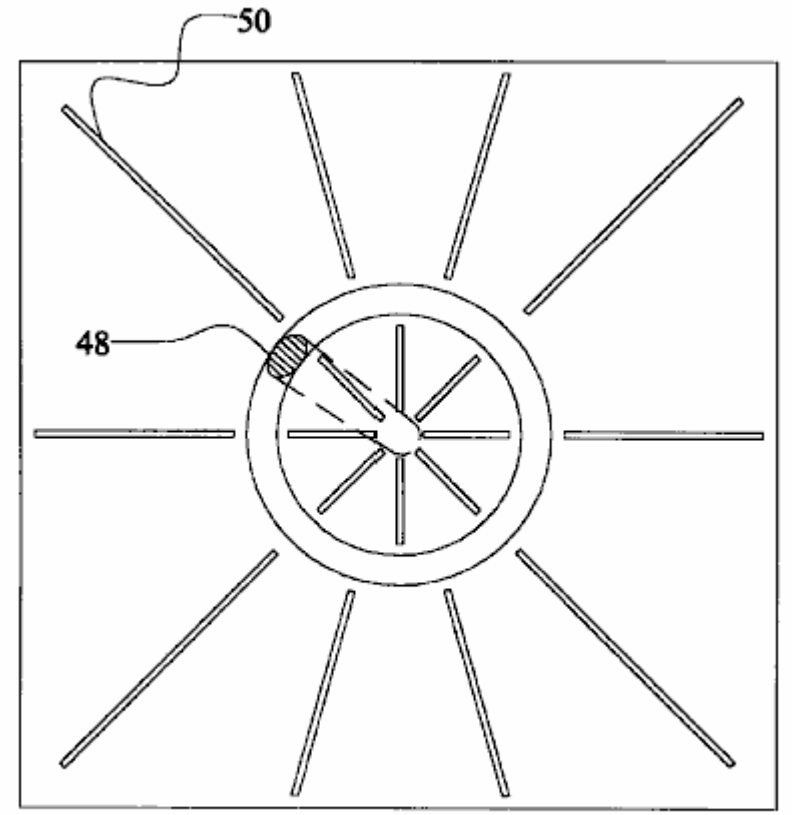


FIG. 4

The same reference numerals refer to the same parts throughout the various Figures.

DESCRIPTION OF THE PREFERRED EMBODIMENT

With reference to **Fig.1**, a typical embodiment of the electrodes of current use for the production of a combustible gas from underwater arcs is that in which one or more pairs of solid carbon rods are immersed within the selected liquid head-on along their cylindrical symmetry axis. The activation of the arc first requires the physical contact of the tips of the two rods, with consequential large surge of electricity due to shorting, followed by a retraction of the electrodes up to the arc gap, which is typically of the order of 1/16 inch (1.5 mm) depending on the input power. The components of such embodiment include:

a, b: carbon electrodes
c, d: holder of a & b
e, f: screws for advancement of a & b
g, h: mechanism for the advancement of a & b
i: reaction chamber
j: exit of combustible gas from chamber

Numerous alternatives to the above typical embodiment have been invented. For instance, in the U.S. Pat. No. 603,058 (H. Eldridge, 1898) one can see a variety of configurations of the electrodes, including rod shaped anodes and disk-shaped cathodes. As a further example also with reference to Fig.1, the embodiment of U.S. Pat. No. 5,159,900 (W. A. Dammann and D. Wallman, 1992) and U.S. Pat. No. 5,417,817 (W. A. Dammann and D. Wallman, 1995), essentially consists of the preceding geometric configuration of the electrodes, complemented by a mechanism for the inversion of polarity between the electrodes, because the cathode experiences the highest consumption under a DC arc, while the anode experiences a much reduced consumption. Even though innovative, this second embodiment also remains manifestly insufficient to achieve the duration of the electrodes needed for industrial maturity, while adding other insufficiencies, such as the interruption of the arc at each time the polarities are inverted, with consequential loss of time and efficiency due to the indicated electrical surges each time the arc is initiated.

As an additional example, and with reference to **Fig.2**, the mechanism of the U.S. Pat. No. 5,792,325 (W. H. Richardson, Jr., 1998), has a different preferred embodiment consisting of one or more pairs of electrodes in

the shape of carbon disks rotating at a distance along their peripheral edges, in between which an electrically neutral carbon rod is inserted. This rod causes the shorting necessary to activate the arc, and then the maintenance of the arc itself. This latter mechanism also does not resolve the main problem considered here. In fact, the neutral carbon rod is consumed at essentially the same rate as that of the preceding embodiments. In addition, the mechanism has the disadvantage of breaking down the single arc between two cylindrical electrodes into two separate arcs, one per each the two couplings of the conducting disk and the neutral rod, with consequential reduction of efficiency due to the drop of voltage and other factors. Numerous means can be envisaged to improve the life of carbon electrodes, such as mechanisms based on barrel-type rapid replacements of the carbon rods. These mechanisms are not preferred here because the arc has to be reactivated every time a rod is replaced, thus requiring the re-establishing of the arc with physical contact, and consequential shortcomings indicated earlier. The components of such embodiment include:

l, m: carbon disk electrodes
n, O: gear rotating l & m
p, q: side gear for rotating n & o
r, s: shaft of gears p & q
t, u: mechanism for rotating shafts r & s
v: electrodes neutral vertical rod
w: advancement of v
x: mechanism for advancement of v
y: reactor chamber
z: electrical power mechanism

This inventor believes that the primary origin of the insufficiency considered here, rests with the carbon rods themselves, which are indeed effective for underwater welding, but are not adequate for the different scope of producing a combustible gas from underwater arcs.

With reference to **Fig.3**, this invention specifically deals with equipment which solves the insufficiency considered here, by achieving the duration of operation desired by the manufacturer, while sustaining a continuous arc without interruptions for the entire desired duration. For the case of large industrial production of this combustible gas with electrical energy input of the order of 100 Kwh, a representative equipment of this invention essentially consists of:

1) One or more arcs produced by a DC current as typically available in commercially sold power units;
2) One or more anodes made of solid rods of about 1 inch in diameter and about 2 inches in length and composed of a high temperature conductor, such as Tungsten or ceramic. Extensive and diversified experiments have established that the consumption of an anode composed of ordinary Tungsten is minimal, and definitely of the order of several weeks of operation.
3) One or more carbon-based cathodes in the configuration of a large hollow rod geometrically defined as a cylinder with the same thickness of the anode, but with a radius and length selected to provide the desired duration. This cathode performs the vital function of becoming incandescent in the immediate vicinity of the arc, thus releasing carbon to the plasma.

More specifically, and with reference to **Fig.3** and **Fig.4**, the present invention essentially comprises a new and improved system **10** for producing a clean burning combustible gas from an electric arc generating plasma under water. First provided is an electrically conductive anode **12** fabricated of tungsten. The anode is solid in a generally cylindrical configuration with a diameter of about one inch and a length of about three inches.

Next provided is a generally Z-shaped crank **14** of a electrically conductive material. The crank has a linear output end **16** supporting the anode. The crank also has a linear input end **18** essentially parallel with the output end. A transverse connecting portion **20** is located between the input and output ends.
An electrically conductive cathode **22** is next provided. The cathode is fabricated of carbon. The carbon is in a hollow tubular configuration with an axis. The cathode has a supported end **24** and a free end **26**. The cathode has a length of about 12 inches and an internal diameter of about 11.5 inches and an external diameter of about 12.5 inches.

A motor **28** is next provided. The motor has a rotatable drive shaft **30**. The drive shaft has a fixed axis of rotation. The motor is coupled to the input end of the crank and is positioned so as to rotate the crank and move the output end and anode in a circular path of travel. The circular path of travel has a diameter of about twelve inches with the anode located adjacent to the free end of the cathode. In this manner the anode and the arc are continuously moved around the cathode and away from the plasma created by the arc.

Next provided is an axially shifted support **32**. The support is in a circular configuration to receive the supported end of the cathode and to move the cathode axially toward the anode as the carbon of the cathode is consumed during operation and use.

A water-tight container **34** for the anode, cathode, crank and support is next provided. A quantity of water **36** is provided within the tank sufficient to submerge the anode and the cathode.

An entrance port **38** is provided in the container. The entrance port functions to feed water and a carbon enriched fluid into the container to supplement the carbon and water lost from the container during operation and use.

Next provided is a source of potential **42**. The source of potential couples the anode and the cathode. In this manner an electrical arc is created between the anode and the cathode with a surrounding plasma for the production of gas within the water. The gas will then bubble upwardly to above the water.

Lastly provided is an exit port **44** for the gas resulting from the application of current from the source of potential to the anode and the cathode while the anode is rotating and the cathode is shifting axially.

Fig.4 is a cross-sectional view taken along line 4--4 of Fig.3, but is directed to an alternate embodiment. In such an embodiment, the anode **48** is wing shaped to cause less turbulence in the water when moving. In addition, various supports **50** are provided for abating turbulence and for providing rigidity.

Again with reference to **Fig.3**, the anode rod is placed head-on on the edge of the cylindrical cathode and is permitted to rotate around the entire periphery of the cylindrical edge via an electric motor or other means. (The inverse case of the rotation of the cathode cylinder on a fixed anode rod or the simultaneous rotation of both, are equally acceptable, although more expensive for engineering production). Extensive tests have established, that under a sufficient rotational speed of the anode rod on the cylindrical cathode of the order of 100 r.p.m. or thereabouts, the consumption of the edge of the cathode tube is uniform, thus permitting the desired continuous underwater arc without the interruptions necessary for the frequent cathode rod replacements in the pre-existing configurations.

For the case of smaller electrical power input the above equipment remains essentially the same, except for the reduction of the diameter of the non-carbon based anode and of the corresponding thickness of the carbon-based cylindrical cathode. For instance, for 14 Kwh power input, the anode diameter and related thickness of the cylindrical cathode can be reduced to about 3/8 inch.

The above new equipment does indeed permit the achievement of the desired duration of the electrodes prior to servicing. As a first illustration for industrial usage, suppose that the manufacturer desires an equipment for the high volume industrial production of said combustible gas from about 100 Kwh energy input with the duration of four hours, thus requiring the servicing twice a day, once for lunch break and the other at the end of the working day, as compared to the servicing only after a few minutes of use for the pre-existing equipment.

This invention readily permits the achievement of this duration with this power input. Recall that carbon rods of about 1 inch in diameter are consumed by the underwater arc from 100 Kwh at the speed of about 3 inches in length per minute. Numerous experiments have established that a cylindrical carbon cathode of 1 inch thickness, approximately one foot radius and approximately two feet in length, permits the achievement of the desired duration of 4 hours of continuous use prior to service. In fact, such a geometry implies that each 1 inch section of the cylindrical cathode is consumed in 6 minutes. Since 4 hours correspond to 240 minutes, the duration of four hours of continuous use requires forty 1 inch sections of the cylindrical cathode. Then, the desired 4 hours duration of said cathode requires the radius R = 40/3.14 or 12.7 inches, as indicated. It is evident that a cylindrical carbon cathode of about two feet in radius and about one foot in length has essentially the same duration as the preceding configuration of one foot radius and two feet in length. As a second example for consumer units with smaller power input than the above, the same duration of 4 hours prior to servicing can be reached with proportionately smaller dimensions of said electrodes which can be easily computed via the above calculations.

It is important to show that the same equipment described above also permits the increase of the efficiency as defined earlier. In-depth studies conducted by this inventor at the particle, atomic and molecular levels, here omitted for brevity, have established that the arc is very efficient in decomposing water molecules into hydrogen and oxygen gases. The low efficiency in the production of a combustible gas under the additional presence of carbon as in pre-existing patents is due to the fact that, when said H and O gases are formed in the plasma surrounding the discharge, most of these gases burn, by returning to form water molecules again. In turn, the loss due to re-creation of water molecules is the evident main reason for the low efficiency of pre-

existing equipment. The very reason for this poor efficiency is the stationary nature of the arc itself within the plasma, because under these conditions the arc triggers the combustion of hydrogen and oxygen originally created from the separation of the water.

The above described new equipment of this invention also improves the efficiency. In fact, the efficiency can be improved by removing the arc from the plasma immediately after its formation. In turn, an effective way for achieving such an objective without extinguishing the arc itself is to keep the liquid and plasma in stationary conditions, and instead, rapidly move the arc away from the plasma. This function is precisely fulfilled by the new equipment of this invention because the arc rotates continuously, therefore exiting the plasma immediately after its formation. Extensive experiments which were conducted, have established that the new equipment of this invention can increase the efficiency from the 2-3 cu. ft. per kWh of current embodiments to 4-6 cu. ft. per kWh.

It is easy to see that the same equipment of this invention also decreases the content of carbon dioxide. In fact, CO_2 is formed by burning CO and O, thus originating from a secondary chemical reaction in the arc plasma following the creation of CO. But the latter reaction is triggered precisely by the stationary arc within the plasma. Therefore, the removal of the arc from the plasma after its formation via the fast rotation of the anode on the cylindrical edge of the cathode while the liquid is stationary implies a decrease of CO_2 content because of the decrease of the ignition of CO and O.

Extensive experimentation has established that a rotation of 100 r.p.m of the anode over the edge of the cylindrical cathode of radius one foot decreases the content of carbon dioxide in the combustible gas at least by half, thus permitting a significant environmental advantage. The decrease of the CO_2 content also implies an increase of the efficiency, alternatively defined as energy content of the gas produced per hour (BTU/hr) divided by the real electric energy absorbed per hour (kWh). In fact, CO_2 is a non-combustible gas, thus having no meaningful BTU content. It is then evident that, since the total carbon content in the gas remains the same, the decrease of the non-combustible CO_2 is replaced in the gas by a corresponding increase of the combustible CO with the same carbon content, thus increasing the energy content of the gas for the same production volume of pre-existing inventions and for the same real power absorbed.

With reference to **Fig.3**, among various possible alternatives, a preferred embodiment of this invention for the high volume industrial production of a combustible gas from underwater arcs with about 100 Kwh real electrical energy essentially comprises:

A) An enclosed reactor chamber **56** of the approximate dimensions 4 feet high, 3 feet wide and 3 feet long fabricated out of steel sheets or other metal of about 1/4 inch thickness, comprising in its interior the electrodes for the creation of the arc and having some means for the exiting of the gas produced in its interior as well as some means for the rapid access or servicing of the internal electrodes;
B) The filling up of said chamber with a liquid generally consisting of water and/or water saturated with carbon rich water soluble substances;
C) One or more anodes consisting of rods of about 1 inch in diameter and about 2 inches in length made of Tungsten or other temperature resistant conductor;
D) One or more cylindrical shaped carbon cathodes with essentially the same thickness as that of the anodes and with radius and length selected for the desired duration;
E) Electromechanical means for the rotation of the anode rod head-wise on the edge of the cylindrical cathode, or the rotation of the edge of the cylindrical cathode on a stationary anode rod, or the simultaneous rotation of both;
F) Automation for the initiation of the arc and its maintenance via the automatic advancement of the carbon cathode, and/or the anode rod and/or both, in such a way to maintain constant the arc gap **58**.
G) Fastenings of the cylindrical carbon cathode so as to permit its rapid replacement; various gauges for the remote monitoring of the power unit, combustible gas, liquid and electrodes; tank for the storage of the gas produced and miscellaneous other items.

An improved version of the above embodiment is conceived to minimise the rotation of the liquid because of drag due to the submerged rotation of the anode, with consequential return to the stationary character of the plasma **54** and the arc, consequential loss of efficiency and increase of CO_2 content for the reasons indicated above.

With reference to **Fig.4**, and among a variety of embodiments, this objective can be achieved by shaping the rotating anode in the form of a wing with minimal possible drag resistance while rotating within said liquid, and by inserting in the interior of the enclosed reactor chamber panels fabricated out of metal or other strong material with the approximate thickness of 1/8 inch, said panels being placed not in contact with yet close to the cathode and the anode in a radially distributed with respect to the cylindrical symmetry axis of the equipment and placed both inside as well as outside said cylindrical cathode. The latter panels perform the

evident function of minimising the rotational motion of said liquid due to drag created by the submerged rotation of the anode.

The remote operation of the equipment is essentially as follows:

1) The equipment is switched on with electric current automatically set at minimum, the anode rod automatically initiating its rotation on the edge of the cylindrical cathode, and the arc being open;
2) The automation decreases the distance between anode and cathode until the arc is initiated, while the amps are released automatically to the desired value per each given Kwh, and the gap distance is automatically kept to the optimal value of the selected liquid and Kwh via mechanical and/or optical and/or electrical sensors;
3) The above equipment produces the combustible gas under pressure inside the metal vessel, which is then transferred to the storage tank via pressure difference or a pump; production of said combustible gas then continues automatically until the complete consumption of said cylindrical carbon cathode.

As to the manner of usage and operation of the present invention, the same should be apparent from the above description. Accordingly, no further discussion relating to the manner of usage and operation will be provided.

With respect to the above description then, it is to be realised that the optimum dimensional relationships for the parts of the invention, to include variations in size, materials, shape, form, function and manner of operation, assembly and use, are deemed readily apparent and obvious to one skilled in the art, and all equivalent relationships to those illustrated in the drawings and described in the specification are intended to be encompassed by the present invention.

Therefore, the foregoing is considered as illustrative only of the principles of the invention. Further, since numerous modifications and changes will readily occur to those skilled in the art, it is not desired to limit the invention to the exact construction and operation shown and described, and accordingly, all suitable modifications and equivalents may be resorted to, falling within the scope of the invention

CHAK CHANG

Patent Application US 2006/060464 23rd March 2006 Inventor: Chak Chang

A METHOD AND APPARATUS FOR GENERATING PLASMA IN A FLUID

This patent application is for a most unusual system which produces a plasma discharge at room temperature and ambient pressure, using voltages as low as 350 volts and currents as low as 50 milliamps and among other things, it is capable of promoting the production of pharmaceuticals, production of nano-particles, the extraction of metals from liquids, low temperature sterilisation of liquid food, use in paper industries to decontaminate the effluent discharge, fragmentation or de-lignifications of cellulose; the removal of odour from discharging liquid in the food industries, and the treatment of fluid effluent. It is also a method of producing hydrogen gas at low cost.

ABSTRACT

A method and apparatus for generating plasma in a fluid. The fluid **3** is placed in a bath **2** having a pair of spaced electrodes **4, 6** forming a cathode and an anode. A stream of bubbles is introduced or generated within the fluid adjacent to the cathode. A potential difference is applied across the cathode and anode such that a glow discharge is formed in the bubble region and a plasma of ionised gas molecules is formed within the bubbles. The plasma may then be used in electrolysis, gas production, effluent treatment or sterilisation, mineral extraction, production of nanoparticles or material enhancement. The method can be carried out at atmospheric pressure and room temperature. The electrodes may carry means to trap the bubbles in close proximity. Partitions may be present between the electrodes.

DESCRIPTION

The invention relates to the provision and utilisation of a plasma formed in a fluid, and in particular to the provision and utility of a plasma formed within bubbles contained in an aqueous medium.

BACKGROUND

Plasma is an electrically conductive gas containing highly reactive particles such as radicals, atoms, plasma electrons, ions and the like. For example plasma may be formed when atoms of a gas are excited to high energy levels whereby the gas atoms lose hold of some of their electrons and become ionised to produce plasma.

Thermal plasma, including plasma arc is known. However plasma arc is associated with high power consumption, the rapid erosion of electrodes when used in electrolysis, the need for catalysts and high-energy loss due to the associated high temperatures.

Clearly therefore, it would be advantageous if a non-thermal plasma could be devised. This would enable the plasma to be used for a number of applications for which plasma is useful without the disadvantages associated with using a high temperature plasma arc.

SUMMARY OF THE INVENTION

According to a first aspect of the present invention, there is provided a method for generating plasma in a fluid, comprising the steps of providing a fluid, introducing and/or generating one or more gas chambers or bubbles within the fluid, whereby the chambers or bubbles are contained by the fluid, and treating the fluid such that a plasma is generated within the chambers or bubbles.

The fluid may be a liquid that is contained within liquid containment means.

The applicant has discovered that a plasma can be generated relatively easily within bubbles within an aqueous medium. This plasma causes dissociation of molecules and/or atoms which can then be treated and/or reacted to obtain beneficial reaction products and/or molecules and/or atoms.

The liquid container may be open to the atmosphere and the process may therefore be carried out at substantially atmospheric pressure. Alternatively the container may be placed inside a sealed reaction chamber, e.g. under

partial vacuum. This reduction in pressure can reduce the energy required to achieve a glow discharge within the bubbles passing over a cathode.

Importantly the process is not required to be carried out in a vacuum.

The plasma may be formed, for example, by applying a potential difference across electrodes which are immersed in the liquid.

Upon passing electricity of sufficient potential between two electrodes, the dielectric barrier associated with the bubble/chamber surface breaks down, with the accompanying formation of a glow discharge and plasma inside the gas bubbles or chambers. This enables plasma formation to be effected at very low voltages, current, temperature and pressure, as compared with known methods of plasma formation.

For example, typical voltages and currents associated with plasma arc are in the region of 5 KV and 200 A respectively, whilst in the present invention, a plasma may be provided with a voltage as low as 350 V and a current as low as 50 mA.

The formation of a glow discharge region adjacent said one electrode is caused by a dielectric breakdown in the bubbles surrounding the electrode. The bubbles have a low electrical conductivity and as a result there is a large voltage drop between the electrodes across this bubble region. This voltage drop accounts for a large portion of the overall voltage drop across the electrodes. The plasma is generated within the bubbles contained within the electrolyte. The liquid electrolyte acts as containment for the plasma within the bubbles.

When plasma discharge occurs, any water vapour inside the bubbles will experience plasma dissociation whereby H+, OH^-, O^-, H, H_3, and other oxidative, reductive and radicals species are formed. The formation of charged plasma species will of course also depend on the chemical composition of the electrolyte.

In the present invention, the voltage needed for plasma generation is much lower than plasma glow discharge generated under gas only conditions. For example experiments have demonstrated that plasma begins to occur at voltages as low as 350 V and the maximum voltage required should not exceed 3,000 V. This requirement is based on a current density of 1 to 3 Amp/cm^2 which can be achieved at the point of discharge whereby the current input ranges from 50 mA to about 900 mA.

Plasma can be created, according to the present invention, in a steady manner with a low voltage and current supply, which leads to an economy in power consumption.

The bubbles may contain precursor materials originating in the fluid, which is preferably a liquid, more preferably being an aqueous electrolyte. This material may have been transferred from the liquid to the bubbles by diffusion or evaporation.

Alternatively the precursor may be introduced directly into the bubbles from outside the system.

The step of generating bubbles within the aqueous medium may be accomplished by one or more of the following: electrolysis, ebullition, ultrasonic cavitations, entrainment, scattering, chemical reaction, dissociation by electrons and ion collisions or local heating or ebullition, hydraulic impingement, ultrasonic waves, laser heating, or electrochemical reaction, electrode heating, releasing of trapped gases in the liquid, and externally introduced gases or a combination of them.

Electrolysis bubbles may be generated by the electrode as a result of the potential differences applied across them, e.g. hydrogen bubbles liberated by the cathode or oxygen bubbles liberated by the anode. Ebullition bubbles may be generated by electrical heating in the region of the electrodes. The bubbles may be generated by direct electrical heating or by heating in proximity to the electrode by a moving wire or grid. Microwave heating and heating using lasers may also be used to generate ebullition bubbles.

Cavitation bubbles may be generated by using an ultrasonic bubble generator or a jet of fluid or a jet of a mixture of gas and liquid injected into the electrolyte in proximity to the electrode. Cavitation bubbles may also be generated by hydrodynamic flow of the electrolyte in proximity to the electrode. Scattering of gas in proximity to the electrode may also be used to generate bubbles.

Bubbles may also be generated by a chemical reaction which evolves gas as a reaction product. Typically such reactions involve thermal decomposition of compounds in the electrolyte or acid based reactions in the electrolyte. Bubbles may also be formed in the electrolyte by adding a frother to it.

Typically the generation of bubbles forms a bubble sheath around one electrode. The bubble sheath may have a thickness of anything from a few nanometres to say, 50 millimetres. Typically the bubble sheath may have a thickness of 1 mm to 5 mm. Further, it should be understood that the bubbles may not be homogeneous throughout the sheath.

Gas or vapour formed external to the container may be pumped or blown into the aqueous medium near the cathode.

Thus the composition of the plasma that is generated within the bubbles may be tailored to suit the application to which the plasma is being put and the bubbles may either be generated within the liquid from components within the liquid or introduced into the liquid from outside the containment means.

The bubbles can assume various sizes and shapes including a sheet form air gap or air pocket covering shrouding the electrodes or spread across the liquid medium in micro bubbles.

Liquid foam may also be considered to be bubbles or gas chambers for the purposes of the present invention. This is a highly concentrated dispersion of gas within a continuous interconnecting thin film of liquid. The gas volume can reach up to 80% of a contained area. Gas generated within or introduced to the reactor externally can also be encapsulated within a foaming agent to enable it to undergo plasma discharge treatment.

Gases trapped inside a thick liquid mist in a confined space are also considered to be gas containing bubbles, which contain the gases, and liquid vapours that provide the condition for generation of non-thermal plasma. The liquid may contribute one or more source materials for dissociation during the plasma discharge.

In practise, gas bubbles evolving near and shrouding an electrode in an electrolysis process create a dielectric barrier which prevents and slows down the flow of current. At the same time the dissolved gas or micro bubbles spread and diffuse in the liquid volume thereby creating a high percentage of void fractions (micro gas bubbles) which in turn increase the electric resistance whereby the voltage across the liquid medium is raised. When the voltage has increased sufficiently, gas trapped inside the bubbles undergoes non-equilibrium plasma transformation. At this point, di-electric breakdown occurs enabling resumption of current flow through the bubbles sheath or air pocket layer.

Any water molecules and atoms lining the gas and liquid interface of a bubble shell will also be subjected to the influence of the plasma to produce H^+ and OH^- and other radical species. Some of these neutralised atoms and molecules will transpose into the gas bubbles as additional gas that increases the size of the bubble. As such the bubbles pick up more liquid vapours before a next succession of plasma discharge. Such a cycle of such repetitive discharge can take place in a fraction of a second to several seconds depending on the make up of the electrode and reactor.

The step of generating bubbles within the aqueous medium may include adding a foaming agent to the aqueous medium such that bubbles are formed within foam. The foam bubbles are confined by an aqueous medium that is electrically conductive. The foam bubbles can vary widely in size down to a fraction of a millimetre.

The step of generating bubbles may include forming an aerosol mist. The gas within the aerosol mist broadly defines bubbles in the sense that there are volumes of gas between liquid droplets. These bubbles in the form of spaces between liquid drops function in a similar way to conventional bubbles within a liquid and a plasma is formed in this gas in the same way as described above.

An advantage of foam and aerosol mist is that it provides for good mixing of gaseous components within the mist and foam. The plasma is generated in the bubbles of the foam and aerosol mist in the same way that they are formed in an aqueous liquid, e.g. by passing electrical current between spaced electrodes within the foam or mist.

The step of forming a glow discharge in the bubble region may be achieved by increasing the potential difference across the electrodes above a certain threshold point.

The formation of a glow discharge and generation of plasma within the bubbles may be assisted by a pulsed or steady power supply, a magnetron field, ultrasonic radiation, a hot filament capable of electron emission, laser radiation, radio radiation or microwave radiation. The energy requirements may also be assisted by a combination of any two or more of the above features. These factors may have the effect of lowering the energy input required to reach the threshold potential difference at which glow discharge is formed.

In conventional electrochemical processes bubbles are regarded as undesirable. As a result concerted efforts are made to avoid the generation of bubbles during the operation of electrochemical cells. By contrast the process of the current invention deliberately fosters the formation of bubbles and utilises bubbles in proximity to the electrode

as an essential feature of the invention. The bubble sheath surrounding the electrode is essential to establishing a plasma region which then gives rise to the plasma deposition on the article.

Thus the plasma is formed within bubbles and the molecules and/or atoms that are ionised are surrounded by liquid which effectively provides a containment structure within which the plasma is contained. The liquid in turn generally opens to the atmosphere.

Plasma glow discharge can be fairly easily accomplished within the cell because the sheath of bubbles has the effect of causing a substantial proportion of the voltage drop to occur across the bubble sheath. It is concentrated in this area rather than a linear drop across the electrode space. This provides the driving force to generate plasma glow discharge and from there deposition of the ionic species.

The electrical charge is preferably applied in pulses, since this enables plasma production at lower voltages.

The fluid is preferably a liquid electrolyte, for example an aqueous medium, whereby in one preferred embodiment, the medium is water.

The electrolyte may comprise a carrier liquid and /or a source or precursor of the material to be ionised by the plasma.

When the liquid is water, charged plasma particles include species such as OH radicals, O^- and H^+, -OH, O_2 and O_3, which will react with the surrounding liquid.

Distilled water is known to be dielectric and non-conductive. It is however when water contains impurities such as dissolved minerals, salts and colloids of particles, whereby water becomes conductive, that ionisation and electrolysis can occur.

The method may further include adding an additive, such as an acidic or alkaline conductivity enhancing agent, to the aqueous medium to enhance this electrical conductivity such as organic salts or inorganic salts, e.g. KCl, $MgCl_2$, NaOH, Na_2CO_3, K_2CO_3, H_2SO_4, HCl.

The method may include adding a surfactant to the aqueous medium for lowering the surface tension of the medium and enhancing the formation of bubbles, e.g. to stabilise bubble formation.

The electrolyte may further include additives in the form of catalysts for increasing the reaction of molecules and/or atoms produced in the plasma, additives for assisting the formation of bubbles, and additives for buffering the pH.

The method may further include cooling the electrolyte to remove excess heat generated by the plasma reaction and regulating the concentration of one or more components within the electrolyte.

The cooling may comprise drawing electrolyte from the bath pumping it through a heat exchanger, and then returning it to the bath.

Plasma creation, according to the present invention can be effected in the absence of extreme conditions, for example plasma according to the present invention may be provide under atmospheric pressure and at room temperature.

During plasma production according to the present invention, a shroud of bubbles preferably builds up and smothers around at least one of the electrodes, whereby electrical charge builds up in the bubble shroud thereby creating a dielectric barrier which impedes current flow, whereby electrical resistance in the fluid medium builds up so that voltage through the medium is raised to a degree such that gas within the bubbles is excited to an energy level at which a plasma is produced.

The method according to the present invention preferably comprises the further step of exposing the plasma to a material, which on contact with the plasma undergoes a chemical and/or physical change.

For example the plasma can be used to cause dissociation of toxic compounds and then break down the compounds and/or cause them to undergo reactions leading to innocuous reaction products.

The plasma produced according to the present invention, which will be referred to as 'under-liquid' plasma has the same physical and chemical properties as plasma produced according to known methods and accordingly also has the utility of such plasma.

The under-liquid plasma according to the present invention can create an active catalytic condition which facilitates gas and liquid interaction. As such, the plasma according to the present invention, may promote any reaction which takes place in a liquid medium, for example chemical reactions, the production of pharmaceuticals, production of nano-particles, the extraction of metals from liquid, low temperature sterilisation of liquid food, use in paper industries to decontaminate the effluent discharge, fragmentation or de-lignifications of cellulose; the removal of odour from discharging liquid in the food industries, and the treatment of fluid effluent. Material may be chemically modified by means comprising one or more of the following: ionisation, reduction, oxidation, association, dissociation, free radical addition/removal, whereby, optionally, following chemical modification, the material is removed.

The invention may be used to tackle existing problems. For example, water that has been used in industrial processes or used in some other way has to be treated to remove harmful components before it is returned to ground water. This is typically achieved by reacting the harmful components with other chemical components introduced to the water to form relatively harmless products. Many undesirable components are treated fairly effectively in this way.

However some harmful components within water are not capable of being treated in this fashion. This poses a problem as these harmful components, e.g. contaminants, need to be removed from the water before it is returned to ground water. One known way of treating some of these components is to use an electric arc process to break down these toxic chemicals. However an electric arc process requires a substantial amount of energy to arc between electrodes within the liquid and is therefore costly. In addition the number of chemicals that are able to be treated in this way is limited. A further limitation of these processes is that they often cause rapid consumption and degradation of electrode material. Applicant believes that this water could be better treated by the method of this invention.

Moreover, the electric arc method of providing plasma, applies a high voltage across closely spaced electrodes causing the break down and ionisation of molecules, and then a surge of electrical current between the electrodes.

Further, many metals or mineral occur naturally in the ground in the form of ores as mineral oxides. The minerals need to be reduced to useful minerals. Typically the reduction is carried out using pyrometallurgical techniques, e.g. such as are used in electric arc furnaces. These treatments are very aggressive and utilise enormous amounts of electrical energy. Clearly it would be advantageous if a simpler more streamlined and more energy efficient method of reducing a mineral oxide to a mineral could be devised. Applicant believes that this could be done by the method of this invention.

Yet further, the generation of electrical energy with fuel cells is seen as an exciting new area of technology. Such fuel cells utilise hydrogen as a fuel. Accordingly a relatively inexpensive source of this hydrogen as a fuel is required. Currently hydrogen is produced by solar cells. However the present invention could be used to provide such a source of hydrogen.

In one form of the current invention, the undesirable compounds may be deposited on an electrode, e.g. the cathode, as a layer or coating. The compound can then be removed from the liquid by simply removing it from the aqueous medium.

In another form, the undesirable component can be reacted with a chemical compound, e.g. within the plasma, to form a solid compound, e.g. a salt in the form of a precipitate, that settles out of the aqueous medium and can then be removed from the aqueous medium.

Typically the undesirable component will be toxic to animals or harmful to the environment. However components that are undesirable in other ways are also included within the scope of the invention.

Applicant envisages that this will be particularly useful for the removal of harmful heavy metals from waste water. It will probably also be useful for the treatment of contaminated gases. Such gases will be introduced to the aqueous medium in such a way that they form part of the bubbles passing over the cathode and then be treated as described above.

Another example is the extraction of a mineral, e.g. a metal, from its metal oxide, the method including: dissolving the mineral oxide in an aqueous medium and then subjecting it to the method described above according to the first aspect of the invention whereby a plasma is generated within bubbles passing over the cathode, and the plasma reduces the mineral oxide to the mineral per se.

The ozone which is formed in the plasma can then be reacted with hydrogen to form an innocuous compound such as water. The reduced mineral which is formed in the plasma, e.g. a metal, may be deposited on the cathode or else may be precipitated out as a solid in the container.

In the case of water, hydrogen and oxygen produced, travel to the anode and cathode and are preferably then removed. As such, the process according to the present invention is an economical, simple and effective way of producing hydrogen.

The hydrogen produced in this fashion may be used as fuel, e.g. in fuel cells for the generation of electricity. Applicant believes that hydrogen can be produced relatively inexpensively in this fashion. Fuel cell technology is currently receiving an increased level of acceptance looking for a cheap source of the supply of hydrogen.

According to another aspect of the present invention, there is provided the use of this 'under-liquid' plasma in one or more of the following: chemical and/or physical treatments of matter, electrolysis, gas production, in particular hydrogen gas production; water, fluid and/or effluent treatment; mineral extraction; sterilisation of drinking water and/or liquid food, production of nano-particles, the enhancement of material chemical and physical properties.

According to a further related aspect of the present invention there is provided an apparatus for providing a plasma comprising; a container in which a plasma is provideable, bubble trapping means, arranged within the container, for trapping gas bubbles at a predetermined location in the container and, plasma creation means, in association with the container, for creating a plasma from the gas within the bubbles.

The plasma creation means preferably comprise electrical discharge means which most preferably comprise a cathode and/or an anode.

The apparatus, in one preferred embodiment being an electrolysis cell, further preferably comprises bubble introduction and/or generating means, for introducing and/or generating bubbles in the container.

Furthermore, the apparatus preferably comprises one or more of the following: enhancing means for enhancing plasma formation and one or more non-conductive partitions arranged between the electrodes, whereby the enhancing means preferably comprise bubble trapping means most preferably associated with the electrodes and wherein the enhancing means may also comprise current concentrating means for concentrating the electrical current at a predetermined position in the container which can take the form of one or more channels arranged through one or more of the electrodes.

The electrodes may take any suitable form, for example the electrodes may be so profiled as to entrap/attract bubbles, in order to help gas bubbles being created or introduced to the discharging electrode to form a dielectric barrier by which the voltage can be raised whereby a suitable current density is provided directly by high input of current or passively created by a current concentrating arrangement, for example, by conducting the current through small holes on the electrodes or by reducing the discharge surface area of the electrodes whereby in the latter case, the electrodes may take the form of pins, wires, rods and the like.

For example, the cathode may be formed by a hollow tube with perforated holes therein, e.g. small perforated holes. The holes allow bubbles introduced into the tube to pass out of the tube into the aqueous medium. Alternatively a cathode may be made of wire mesh or have a roughened surface, e.g. to encourage the attachment of bubbles thereto to slow down the movement of the bubbles.

In one embodiment there are a plurality of cathodes spaced apart from each other and in parallel with each other, and a single rod-like anode, e.g. centrally positioned relative to the cathode.

The other electrode (non discharging) preferably has a larger surface area such than the discharging electrode.

The discharging electrode can either be cathode or anode depending on the application necessity.

In an experimental reactor the separating membrane, non-conductive partition, was nylon cleaning cloth having a tight matrix 0.5 mm thick. This semi-permeable membrane is capable of resisting the passage of oxygen and hydrogen ions through it in the aqueous medium, intermediate the anodes and cathodes thereby to maintain separation of oxygen and hydrogen produced in the plasma.

Most preferably, the apparatus according to the present invention is an electrolytic cell.

A known problem with carrying out electrolysis is that any gas/bubble build up in the electrolytic cell creates a barrier to the flow of current through the electrolyte, thereby impeding electrolysis, which increase in resistance in turn forces the required voltage up. As such, electrolytic cells require a great deal of energy and are often very large in order to effect dispersion of such gas/bubbles. However the present invention actively promotes such bubble build up, in order to effect plasma creation which the inventors have shown is effective in carrying out electrolysis.

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS OF THE INVENTION

A plasma formed in a fluid in accordance with this invention may manifest itself in a variety of forms. It will be convenient to provide a detailed description of embodiments of the invention with reference to the accompanying drawings. The purpose of providing this detailed description is to instruct persons having an interest in the subject matter of the invention how to put the invention into practice. It is to be clearly understood however that the specific nature of this detailed description does not supersede the generality of the preceding statements. In the drawings:

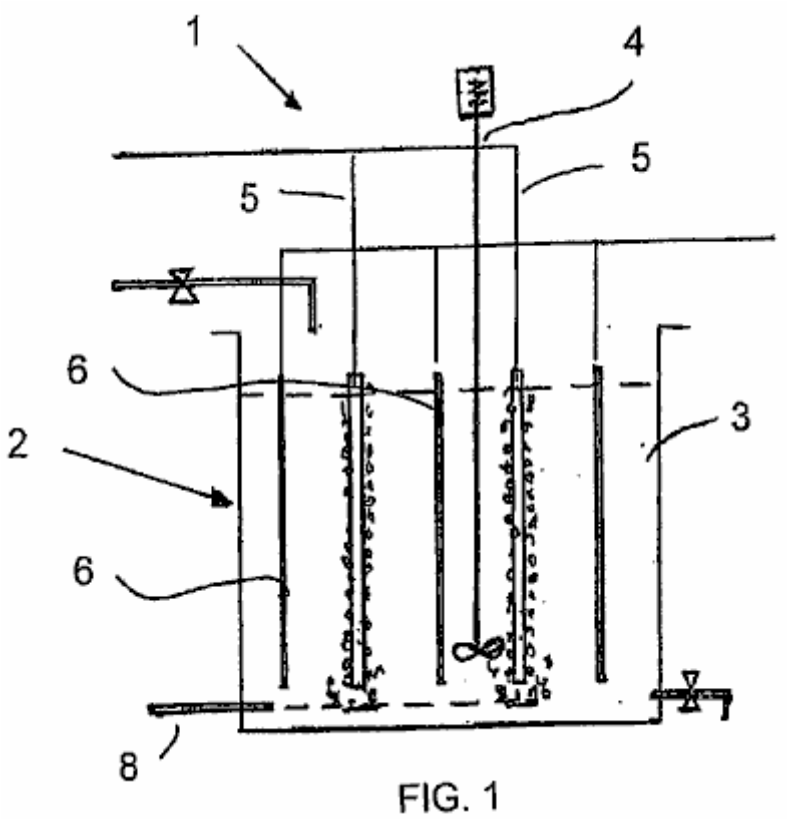


FIG. 1

Fig.1 is a schematic sectional front view of apparatus for carrying out a method in accordance with the invention.

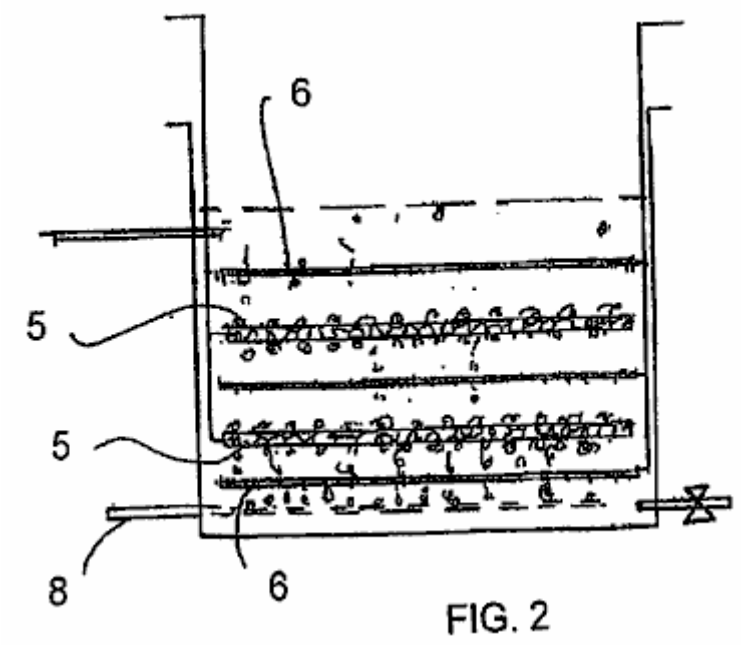


FIG. 2

Fig.2 is a schematic sectional front view of a variation on the apparatus of Fig.1.

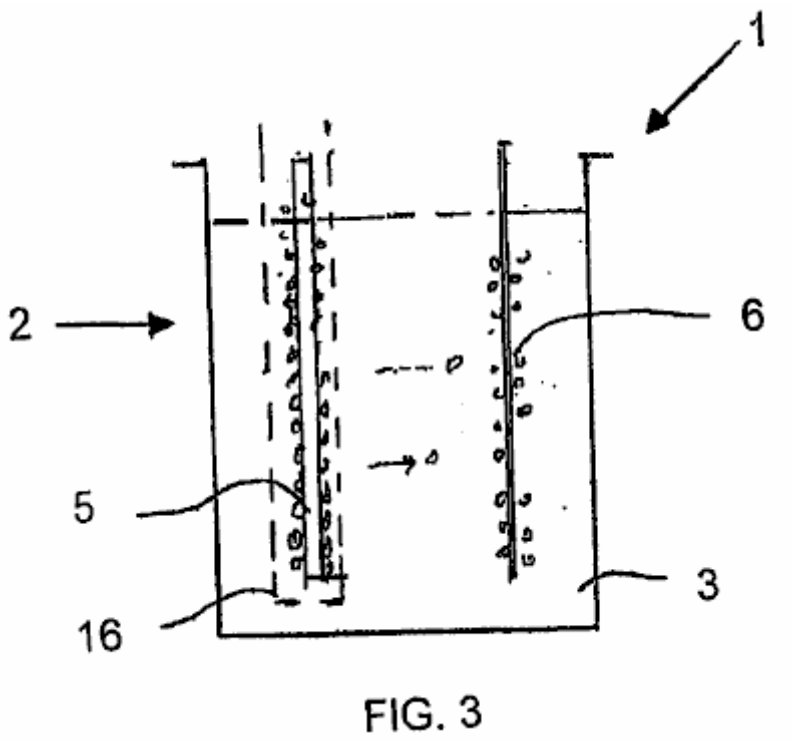


FIG. 3

Fig.3 is a schematic sectional front view of an apparatus in accordance with the invention suitable for producing hydrogen gas.

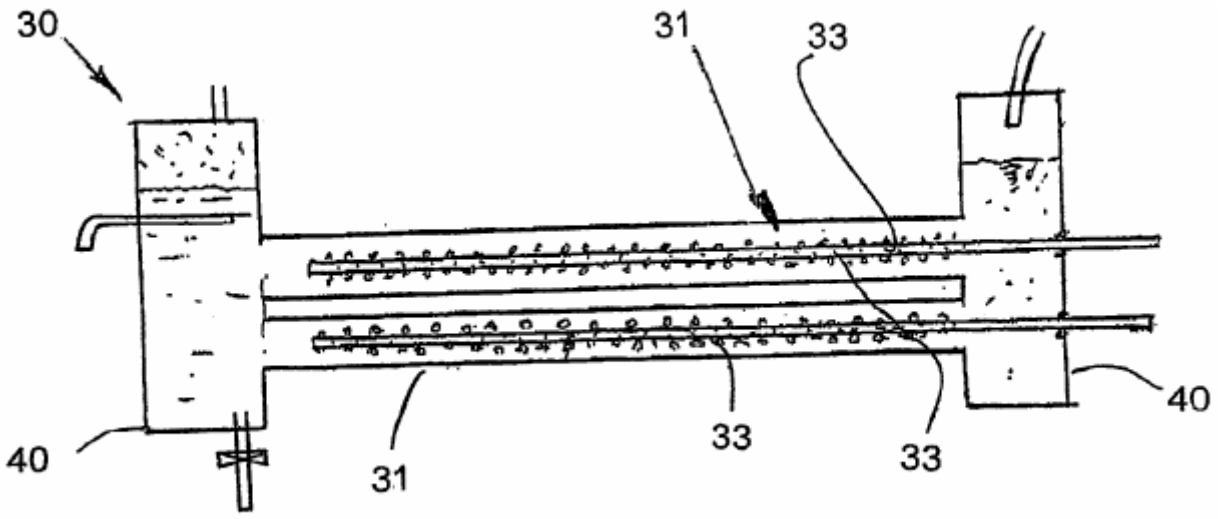


FIG. 4

Fig.4 is a schematic sectional front view of a tubular reactor carrying out a method in accordance with another embodiment of the invention.

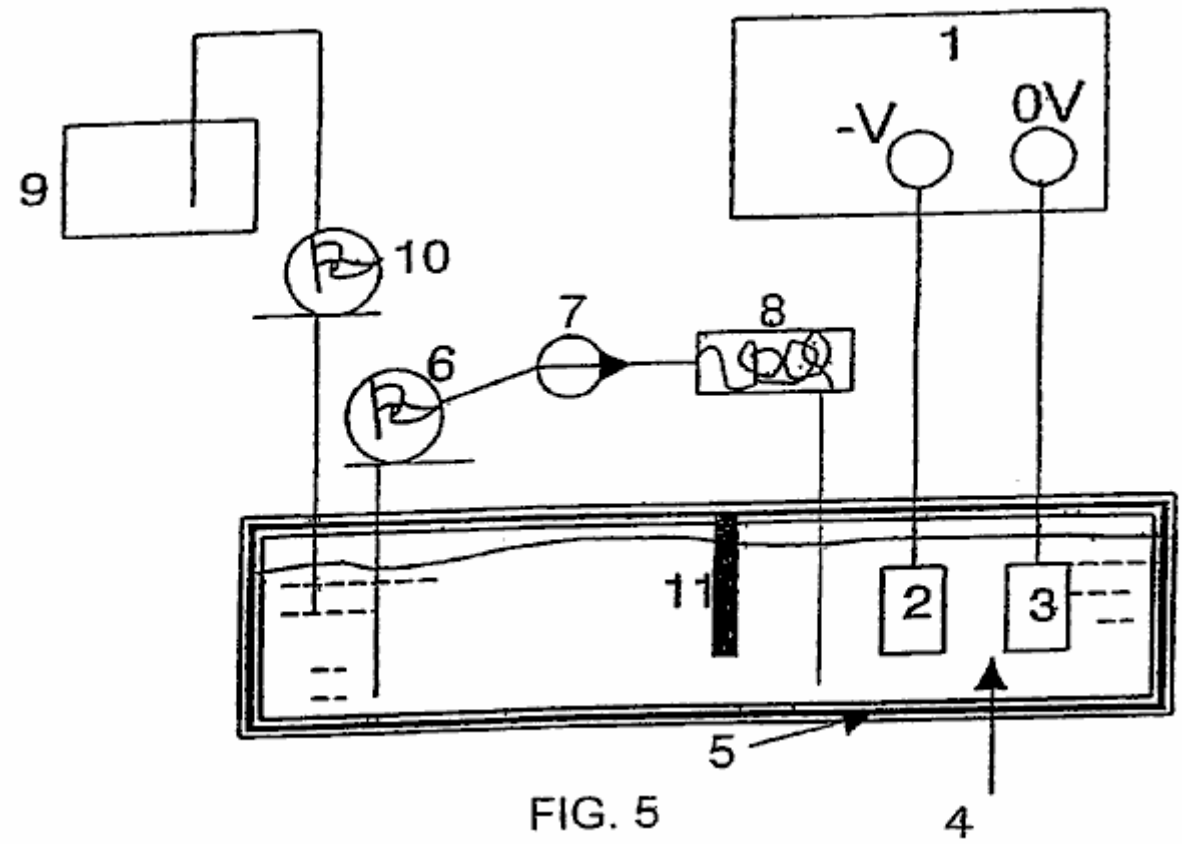


FIG. 5

Fig.5 is a schematic flow sheet of apparatus in the form of a cell for carrying out the invention.

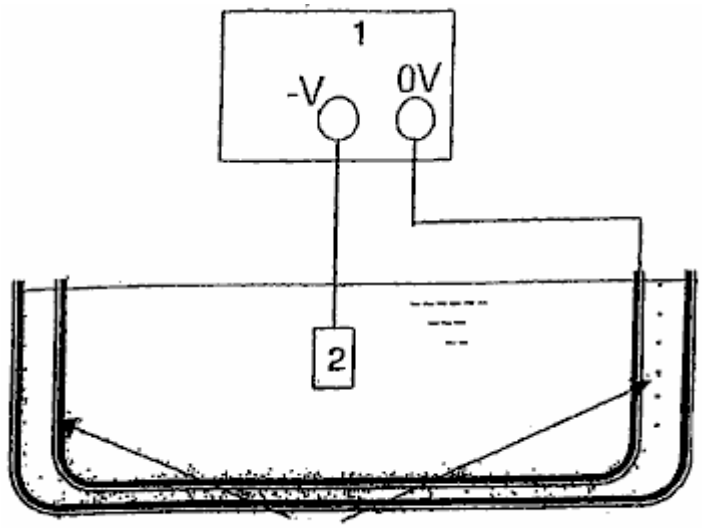


FIG. 6

Fig.6 is a schematic view of a bath for the cell of Fig.5 having an ultrasonic generator for generating bubbles.

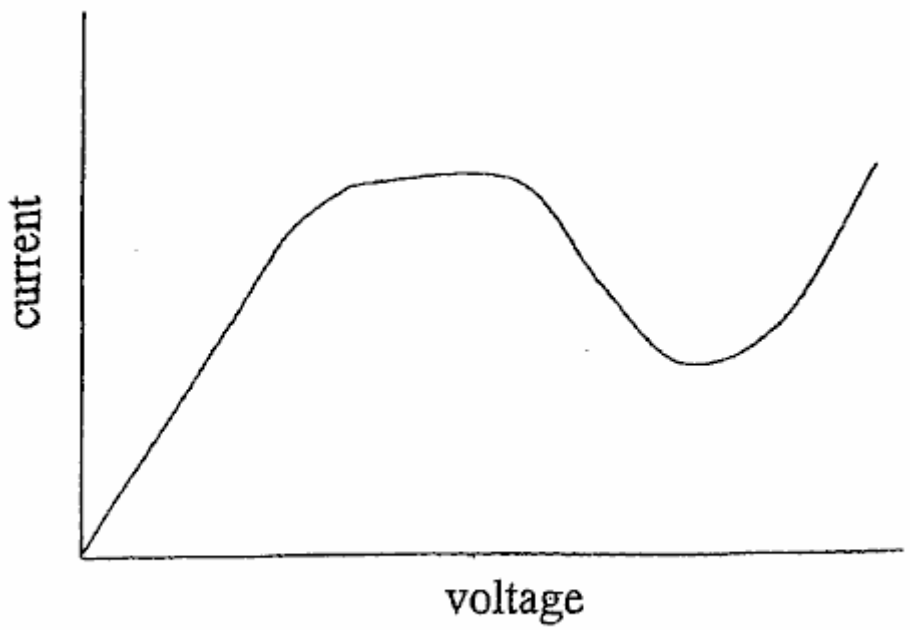


FIG. 7

Fig.7 is a schematic graph of current against voltage in an electrolytic cell.

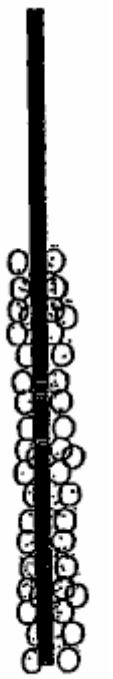


FIG. 8

Fig.8 shows the initial formation of a bubble sheath around the cathode due to the application of voltage across the electrodes.

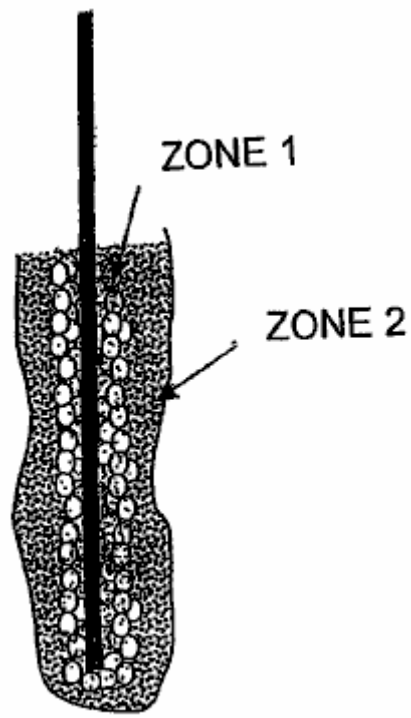


FIG. 9

Fig.9 shows the bubble sheath around the cathode during stable glow discharge within the cell, and

Figs.10-53 refer to further embodiments and experimental results in respect of the present invention.

The present invention relates to the production of non-thermal plasma contained in a liquid by generating corona discharge and or glow plasma discharge inside the bubbles or air pockets present in the liquid.

Upon passing electricity of sufficient potential through the liquid, electric breakdown of the dielectric bubble barrier results in the formation of plasma discharge inside the gas bubbles or pockets present in the liquid. In most cases glow discharge occurs near the electrodes but occasionally glow discharge is also observed away from the electrode.

The bubbles can be produced either by electrolysis, electrochemical reaction, heating of electrodes, releasing of trapped gases in the liquid, ultrasonic cavitations, laser heating, and externally introduced gases.

Bubbles produced by electrolysis of water contain hydrogen gas at the cathode and oxygen gas at the anode. Such bubbles can also contain other chemical vapours originating from the electrolyte or additives.

The liquid serves as an electrolyte which provides conductivity of electricity, the source material from which gases and vapour are produced for plasma dissociation to form, for example, reduction and oxidation, radicals and neutral species. The liquid also provides an active catalytic chemical environment for forming new compounds. It also serves as containment of gases in the form of bubbles or air pockets in which the non-thermal plasma discharge takes place.

In practise, gas bubbles evolving and shrouding the electrodes during electrolysis create a dielectric barrier which inhibits the flow of current.

At the same time the dissolved gas or micro bubbles spread and diffuse in the liquid volume create a high percentage of void fractions (micro gas bubbles) which also increase the electric resistance and so raise the voltage across the liquid medium.

When the voltage between two electrodes reaches a critical level, the gas trapped inside the bubbles undergoes non-equilibrium plasma transformation. This is also known as electric breakdown which enables the resumption of current flow through the bubble sheath or air pocket layer. In the case of water electrolysis, the production of hydrogen will then resume.

During plasma discharge, light emission may be observed in the bubbles in a sporadic or steady manner in short and continuous flashes near the surface of the electrodes and in the liquid medium.

Continuous light spots may also be observed in areas distanced from the electrodes where suspected small air bubbles are trapped and yet remain under the influence of strong electrical field.

The temperature in the electrolyte near the electrodes has been measured to be in the region of $50^{O}C$ to about $90^{O}C$ with an experiment running in water for 30 minutes, which indicates that the plasma is non-thermal plasma.

The temperature variation may be influenced by electrode geometry, electrolyte concentration, level of inception voltage and current density for the glow discharge. The temperature measured directly over the discharging electrode can reach over $200^{O}C$ during reformation of methanol for example.

Configurations of electrodes, size, spacing, dielectric barrier coating, electrolyte temperature, current density, voltage and reactor geometry are factors influencing plasma formation.

A special structure and arrangement to retain gas or gas bubbles close to the electrodes provide favourable circumstances for the ready formation of a steady and cyclical plasma glow discharge with lower voltage and current input.

Electrode configurations can be in following forms: plate to plate, plate to pinned plate, dielectric coated plate to plate or pinned plate or both, wire mesh to plate, wire mesh to wire mesh or to perforated plate, wire or groups of wires in perforated cylinder tube, and tube in tube.

The electrode material may be sponge porous metal electrode, electrode covered with honeycomb non-conductive materials and porous ceramic filter to entrench gas or using non-conductive plate with drilled holes and gas traps that retain gas bubbles and concentrate the current density next to the electrode surface.

In general keeping the bubbles close to the surface of the electrodes can also be achieved by attaching a porous non-conductive nylon foam mattress and/or a honeycomb or porous ceramics slab of suitable thickness, so that the mobility of the bubbles is slowed down and at the same time the conduit for current flow is narrowed by a shading effect of the dielectric materials which in turn raises the current density locally.

For the same reason glass beads, plastic beads and beads of catalytic material i.e. TiO_2, graphite of suitable size can be placed between the electrodes in order to slow down the flow of bubbles.

A non-conductive, heat and corrosion electrode covering material, structured to retain and trap gas bubbles which also concentrates current density through small openings arranged through it whilst providing an adequate exposed electrode surface for electro-chemical and electrolysis reactions, improves the generation of steady and short cyclical reactions under-liquid plasma discharge.

Multiple layers of very fine stainless mesh, sandwiched between two plastic cover plates with small perforated holes, have produced a steady glow plasma. The void space created by the layered wire mesh provides a trap for air bubbles as well as enlarging the contact surface for electrochemical and electrolysis reaction.

In an experiment both vertical or horizontal electrodes were covered and bonded with non-conductive materials (plastic) with patterned perforations to trap gas bubbles while at the same time allowing for electrical contact of the electrodes through the perforations.

The electrode contact surface was enlarged underneath the shielding to increase gas production during electrolysis or heating. Current flow was concentrated through small holes of 1 to 3 mm leading to the trapped gas and bubbles, which underwent plasma transformation. Cyclical and steady plasma was observed with an input DC voltage ranging from 350V to 1900V and current ranging from 50 mA to 800 mA.

A non-conductive diaphragm, which does not restrict the free flow of ions and electrolyte, is placed between two opposite electrodes to prevent crossing of bubbles between two half electrolytic cells avoids re-mixing of the gases which have been separated by electrolysis.

A reactor may be so structured that the electrolyte is able to enter into the reactor through the separating membrane or opening form in the reactor to replenish the loss of electrolyte within the enclosed reactor.

There are other techniques which can be incorporated into the proposed invention for the enhancement of plasma generation such as pulsed power supply, RF power, microwaves, ultrasonic waves, magnetron field, and laser. Some of the above techniques may also be applied in pulsed form.

Ultrasonic cavitations in liquid (sonic-technology) will enhance the plasma formation and the catalytic reactions that benefit a number of under-liquid plasma applications.

The under-liquid plasma requires an input of DC or AC voltage in the range from 350V up to 3000V and current density ranging from 1 Amp to 3 Amp per cm^2 in dealing with a large range of liquid media. The specific voltage and current requirement for a given application depends very much on the chemical and physical properties of electrolytic liquid as well as those factors mentioned above.

The under-liquid plasma method according to the current invention, can operate at atmospheric pressure and ambient temperature. However, an external pressure less than one atmosphere or over one atmosphere with higher temperatures does not deter the generation of plasma in the bubbles. A higher temperature in the liquid also means more active gas molecules within the bubbles, which can benefit plasma formation.

Non-thermal plasma generated in a liquid according to the present invention, has advantages over known types of plasma discharge, for example in gas, under water plasma arc and pulse power electric discharge, these being:

It requires only simple electrolytic cells to be the reactor to perform such discharge. There is little erosion to the electrodes and wider range of electrode materials can be chosen such as stainless steel, graphite, aluminium and good conductive materials which are resistance to chemical erosion. The polarity of the electrode can be reverted if necessary to compensate the lost of electrode materials if so desired.

It works under one atmospheric pressure and ambient temperature. The liquid electrolyte will be primary source of materials for the chemical and physical reaction take part in the process. There are number of ways that bubbles can be produced within the electrolytic cell. Gas can also be introduced to the reactor where plasma catalytic and dissociation is taking place.

It is a low-temperature system as the plasma discharge is non-thermal. Any excessive or undesirable high temperature can be lowered by increasing the circulation rate of the liquid which can lose its temperature through heat exchange. Heat generated can be recovered as secondary energy.

The electrolyte (liquid) will serve as extension of the conducting electrodes in contact with the gases or vapour trapped inside the bubbles. The air gap between two electrodes is reduced to the thickness of the gas bubbles or air pocket which thus enables plasma discharge at a much lower voltage and current compared with other plasma discharge systems. Plasma glow discharge, according to the present invention, can be initiated under conditions of a voltage as low as 350V and the current ranging from 50 mA to 800 mA. Extra energy is not required in splitting the water molecules to transient bubbles as in the other underwater electrical discharge system which requires voltage not less than 5 to 6 KV, and very high current over 200 A in pulsed supply. Plasma discharge will also take place in gas pockets or bubbles away from the electrode as long as the electric field strength is sufficient to cause such discharge.

The electrolyte also serves as a confinement of gas generated within the system, or purposely introduced gas of known properties, instead of ordinary air which may lead to production of unwanted NOx for example. Noble gas such as argon is not necessary to enhance the initiation of glow discharge sometime required in the air discharge system.

The electrolyte also serves as a conductor and passage for the transportation of ionised species and transmission of electrons. The ionised atoms and molecules deriving from the electrolyte will be collected in their respective electrodes in the form of gas or material deposit. These ionised species are either serving as a reduction or oxidation agent in their respective half-cell. Since the gas ions produced during the discharge migrate to their respective poles to be collected individually, hydrogen gas and oxygen gas can be collected separately.

The gas and vapour molecules and atoms inside the bubble which undergo plasma glow discharge are ionised, excited or dissociated to produce the very active species for reduction, oxidation, and the forming of neutral or radical species which in turn react with the chemical elements present in the gas and liquid interface aligning bubbles wall. The large number of bubbles generated near the electrodes and in the nearby liquid, come into contact with a much larger volume of liquid nad so provides effective treatment, breakdown, transformation of chemicals, organic matter or elements which have been targeted.

Liquid is a good medium for transmitting ultrasonic waves. Sonic-excitation is beneficial for the dissociation of materials and extermination of microbes and it aids the breakdown and local melting of colloidal solids during impact which also enhances the plasma oxide reduction process. The generated ultrasonic cavitations may be fully utilised to work in conjunction with the under-liquid plasma discharge. An ultrasonic cavity is micro in size and uniformly distributed in the entire liquid volume. The cavities are a high vacuum which contain liquid vapour and gas, and these favour plasma discharge. The high temperature and pressure reaching $10{,}000^{O}K$ and a

thousand times atmospheric pressure, produced on the collapsing phase of these cavities work is complementary to that of the electro discharge plasma. This enables under-liquid plasma discharge to spread further from the electrodes and be well distributed in the liquid volume which increases its overall effectiveness.

The electrolyte may also be in the form of a mixture, an emulsified liquid, a colloid, or foams encapsulating gas emissions either coming from the liquid or introduced externally. The emulsified liquid of an oil/water mixture and encapsulating gas of hydrocarbon fuel with the ultrasonic irradiation, will facilitate their reformation for hydrogen production.

Fine granular insoluble particles of mineral oxide such as aluminium, titanium, iron, silica etc. can be suspended in the form of colloid with the liquid which is than subjected to reduction with active ionic hydrogen atoms in a highly reactive plasma catalytic environment to become deoxidised and refined. This will be more so, with the assistance of sonic impedance. The Plasma glow discharge has also demonstrated the ability to dissociate soluble ionic metal compounds, whereby subsequently the positively charged metal ions will be segregated near the cathode electrode in the form of precipitation and plasma electroplating deposition.

The electrolyte may be a source of materials for thin-film deposition with the assistance of plasma glow discharge. In addition, nano size particles of certain compounds and elements i.e. metal hydride, oxide, pure metals, semi metals, organic, ceramic etc. can also be produced with the assistance of the under-liquid plasma discharge in conjunction with the ultrasonic cavitations mechanism, to cause breakdown and reformation of certain compounds. The highly catalytic, reactive and dissociation capacity of the glow discharge plasma, reforms and reconstitutes chemical elements and compounds from basic atoms or molecules to form nano particles. These include organic, inorganic, metallic and non-metallic materials such as silica, titanium carbon etc. This is also a very effective way to extract or remove heavy metals from a liquid by oxidising such as Hg to HgO; Cu, Zn, Cr etc. to form hydroxide precipitation and ionic metal solute to be deposited by the plasma electroplating process.

The under-liquid plasma creates a highly catalytic and reactive environment for chemical reactions which would not take place under normal circumstances. The reductive species i.e. H+ and oxidative radicals i.e. O-, O_3, H_2O_2, OH- and other radical species produced in the electrolysis and plasma dissociation derived from the liquid itself. The sonic excitation action which enhances the effectiveness of plasma discharge can only be conducted spontaneously under and within liquid.

The under-liquid plasma technique, coupled with the sonic-excitation and electro-chemical action, creates an environment of localised high temperature up to $10{,}000^{O}K$ and pressure up to thousands of atmospheres which favour the generation of cold-fusion phenomena.

It is a low-energy system. Generally high voltage from 0.35 KV up to 3 KV with low current density rarely required more than 3 Amp/cm^2 will be needed to deal with a vast number of different types of the under-liquid plasma process. If other enhancement method is applied, the high voltage and current requirement will be further reduced.

It is a method for producing hydrogen, oxygen with water or other gases and material deposition with liquid containing chemical solute, other than the conventional exchange of ions. The molecules and atoms are being ionised, excited and subjected to dissociation to form ionised, radicals and neutral species by the influence of plasma discharge. The dissociated species can be produced near either anode or cathode electrodes. The ionised species are then attracted to their respective polarity to be neutralised to produce gas or deposition of materials. The dissociation of atoms or molecules are the result of electron collisions and a wide variety of dissociated species is produced which creates the reactive elements for reduction, oxidation, and highly catalytic environments that facilitate chemical reaction of those relatively stable compounds and elements.

No chemicals are needed as an additive in a decontamination process, of which chemicals, i.e. chlorine and ozone, could become a secondary source of pollution.

EXPERIMENTAL OBSERVATIONS

When sufficient micro bubbles originating from the electrode surface block the current flow, the voltage rises steadily until a point of voltage inception is reached whereby some micro bubbles begin experiencing glow discharge. This precedes an avalanche effect which spreads through other micro bubbles close by.

A massive light is then emitted in a flash with a sound of bursting bubbles. The light is yellow to orange in colour indicating plasma discharge in hydrogen gas at the cathode electrode. Soon after switching on the reactor, temperature in the electrode rises which contributes to the formation of vapour bubbles which in turn creates a large bubble environment full of water vapour whereby the next succession of plasma discharge takes place within a fraction of a second.

The features which enable the trapping of gas, the concentration of current density within a small region, and the continued replenishment of gas, are steady and a self-regulating voltage and current power supply, electrode spacing, electrode configuration and electrolyte concentration, all of which have a bearing on generating desirable steady, and short cycle plasma glow discharges.

The invention has a number of applications including:

Plasma assisted electrolysis for hydrogen generation.
Non-thermal plasma reformation of hydrocarbon and hydrogen rich compounds for the production of hydrogen.
Treatment of polluted and contaminated liquid waste containing chemical and heavy metal pollutants.
Treatment of polluted gas emission and removal of odours.
Sterilisation of drinking water and liquid foods.
Extraction and refinement of mineral from its oxide or oxide ores.
Production of nano particles.
Enhancement of a material's chemical and physical properties by plasma discharge irradiation in under- liquid conditions. This also favours the need of any plasma reaction and treatment under-liquid.

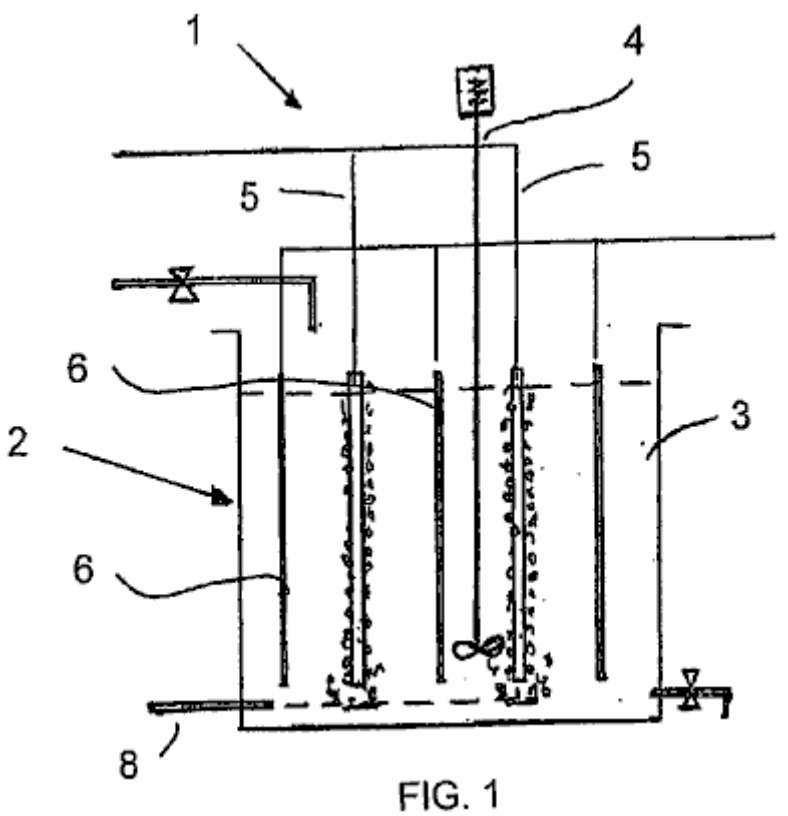


FIG. 1

Fig.1 illustrates a basic apparatus **1** for carrying out the method of the invention, namely, generating a plasma within bubbles formed adjacent to a cathode within an aqueous medium. The apparatus **1** comprises a liquid containment means in the form of an open rectangular tank **2** opening to the atmosphere and containing an aqueous liquid **3**. A stirrer **4** for agitating the aqueous liquids in the tank **2**.

Two spaced cathodes **5** are positioned in the tank **2** alternating with three anodes **6** projecting into the tank **2** and extending generally parallel to the cathodes **5**. A bubble pipe **8** is positioned at the bottom of the tank **2** for introducing bubbles into the aqueous medium in proximity to each of the cathodes **5**.

The application of a suitable potential difference across the anodes and cathodes leads to a glow discharge being formed and a plasma within the bubbles adjacent the cathode. This ionises the atoms and/or molecules within the bubbles and can be used to achieve a number of industrially and commercially useful objectives. For example, it can be used to generate hydrogen gas, one of its uses includes placement in a fuel cell to generate electricity. It can also be used to neutralise harmful compounds within the aqueous medium, e.g. originating in a liquid source or a contaminated gas and treating these harmful compounds. Finally, it can also be used to coat the surface of an article with a particular material.

Each of the cathodes is in the form of a perforated tube. At least one end of the tube is open and typically gas is introduced through such an open end. The side wall of the tube is perforated such that gas issues from the tube into the aqueous medium around the cathode. Alternatively, each of the anodes may be rod-like.

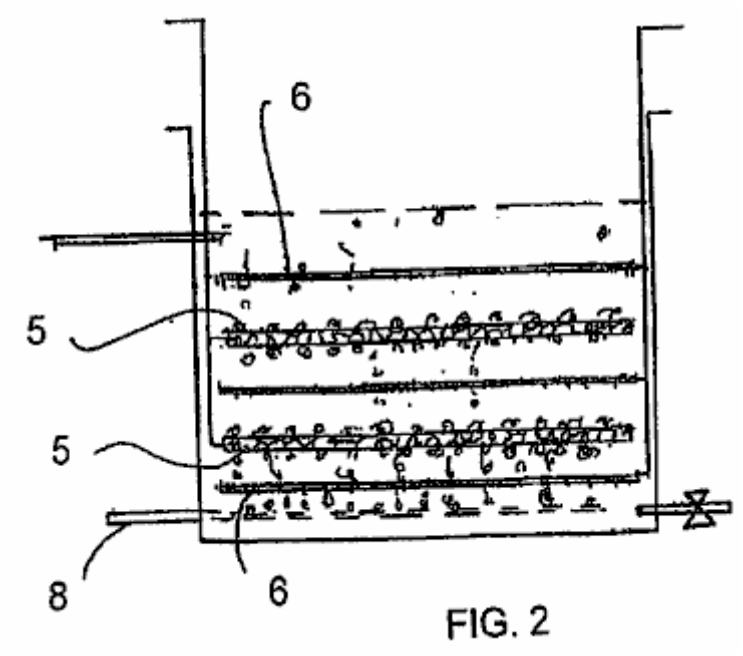


FIG. 2

Fig.2 illustrates a variation on the apparatus of **Fig.1**. This description will be confined to the difference between the **Fig,1** and **Fig.2** apparatuses. In **Fig.2** the electrodes extend horizontally with each cathode positioned between two vertically spaced anodes.

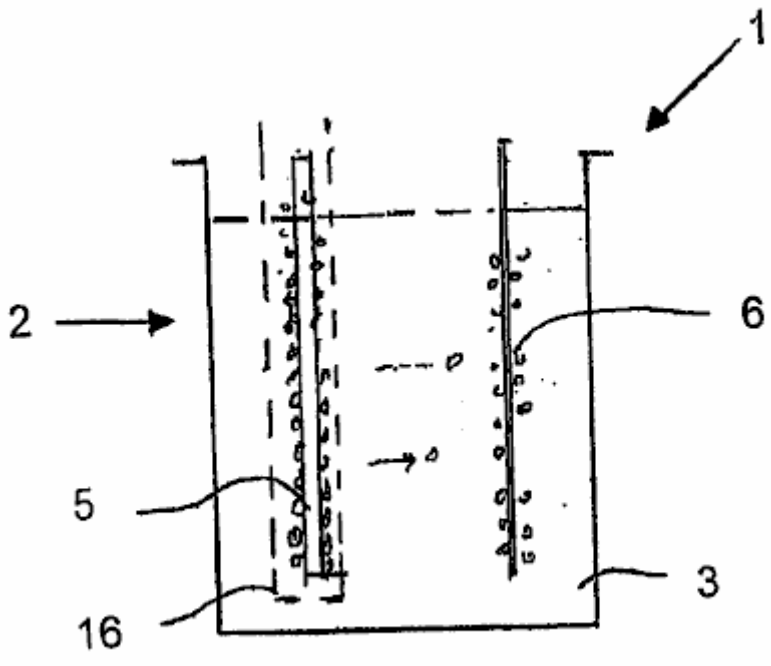


FIG. 3

Fig.3 illustrates an apparatus suitable for the generation of hydrogen. The tank contains an anode and a cathode spaced apart from each other. The electrodes are generally the same as those described above with reference to **Fig.1**. The cathode is surrounded by a semi-permeable membrane. Specifically the membrane is designed to resist the passage of hydrogen and oxygen bubbles through it. Hydrogen gas is formed from the combining the two neutralised hydrogen ions adjacent to the cathode and then is drawn off from the aqueous medium above the cathode and collected for use.

Similarly, oxygen gas is formed adjacent to the anode and this is also drawn off separately and collected for use. An advantage of this method for the formation of hydrogen fuel is that it consumes essentially less energy than other known methods, and as a result, will be a very attractive source of hydrogen for use in fuel cells.

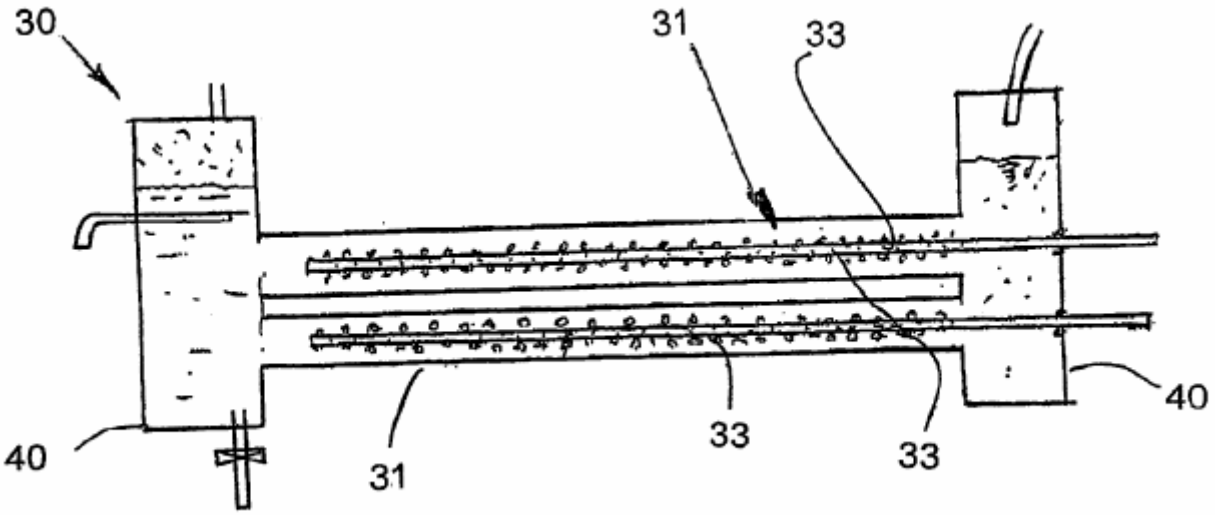


FIG. 4

Fig.4 illustrates a tubular reactor which is quite different to the tank **2** shown in the previous embodiment. The reactor **30** comprises a circular cylindrical body **31** with its longitudinal axis extending horizontally. A pair of electrodes **32**, **33** extend longitudinally through the body, spaced in from the wall of the body **31**. Each cathode **33** is formed by a perforated tube. By contrast, the anode is formed by the body **31**. Thus the single anode **31** extends concentrically around the cathodes **33**, positioned radially inwards from them. A gas, which ultimately forms the bubbles, is pumped into the cathodes, e.g. through their open ends, and then issues through the openings along the length of the cathodes **33**.

Settling tanks are located at each end of the body **31**. The settling tanks **40** permit gas to be separated from the liquid. The gas rises to the top of the tanks **40** from where it can be drawn off. The aqueous liquid can be drawn off through a drain point positioned below this level of aqueous medium in the tank **40**. An aqueous medium can also be introduced into the apparatus, by passing it through an inlet into one of the tanks **40**. Otherwise, the method of generating plasma in bubbles adjacent to the cathodes is very similar to that described above with reference to **Fig.1** to **Fig.3**.

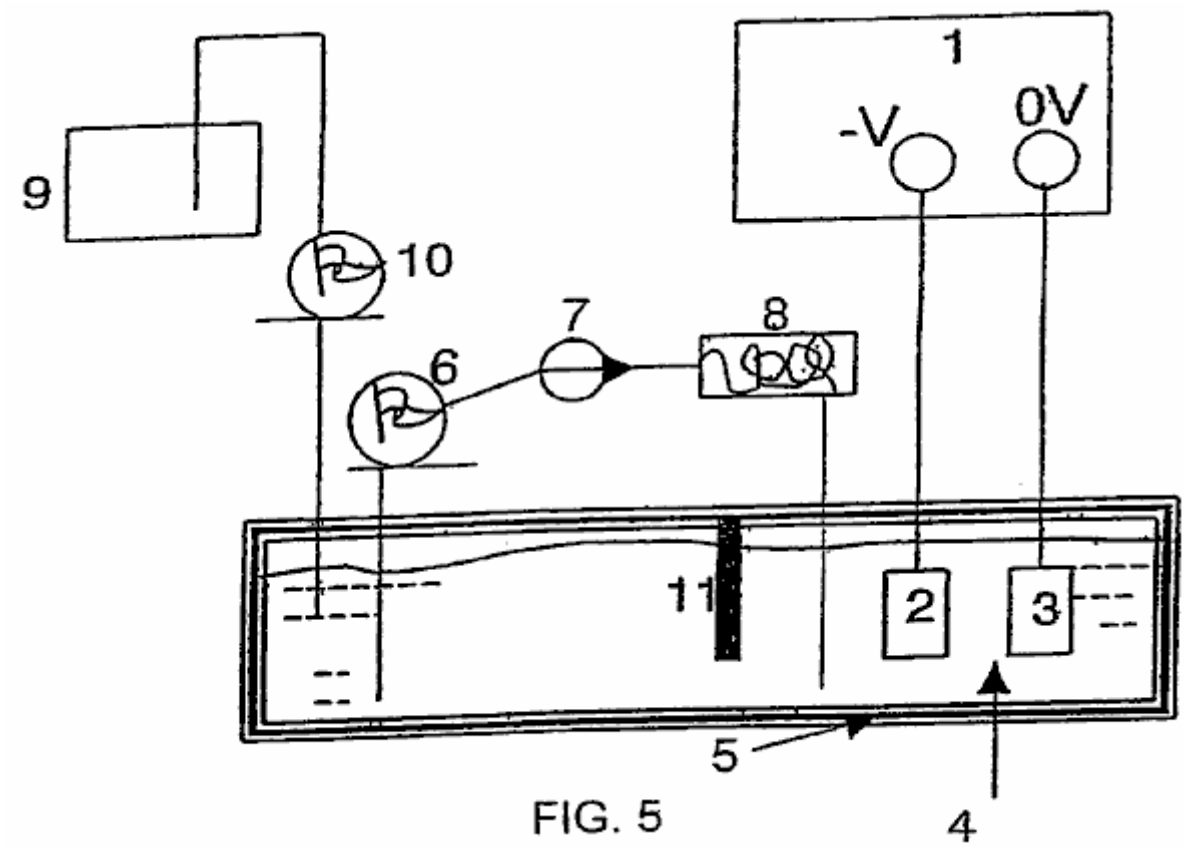


FIG. 5

In **Fig.5**, reference number **1** refers generally to apparatus in the form of a cell and associated components for carrying out a plasma electroplating process (PEP) in accordance with the invention. The cell **1** comprises

broadly, a liquid container in the form of a bath which is filled with an electrolyte which also forms part of the apparatus or cell. A pair of spaced electrodes are positioned in the bath, one being a cathode and the other being an anode. An electrical circuit is formed by electrically connecting up the anode and cathode to a power supply, e.g. a mains power supply. When the bath is being used, a potential difference is applied across the electrodes. A partition divides the bath into an electrode compartment and a circulating compartment. Electrolyte is drawn off the circulating compartment and pumped through a heat exchanger to cool it and then return it to the bath. This helps to keep the temperature of the electrolyte within a suitable range during operation. In addition a make-up tank is positioned adjacent the circulating compartment to replenish the level of electrolyte within the bath as and when required.

The apparatus also includes the means for producing a bubble sheath around the cathode. The bubbles can be generated by gas evolved at the cathode as a result of a cathodic electrochemical reaction. This is one of the ways in which the bubbles were generated in the experiments conducted by the applicant. There are however, alternative ways of generating the bubbles for the bubble sheath. One alternative way, is by boiling the solution (ebullition bubbles). Other ways of producing the bubbles are by cavitation generated by ultrasonic waves or by hydrodynamic flow. Entrainment bubbles can also be produced by a mixture of gas and liquids.

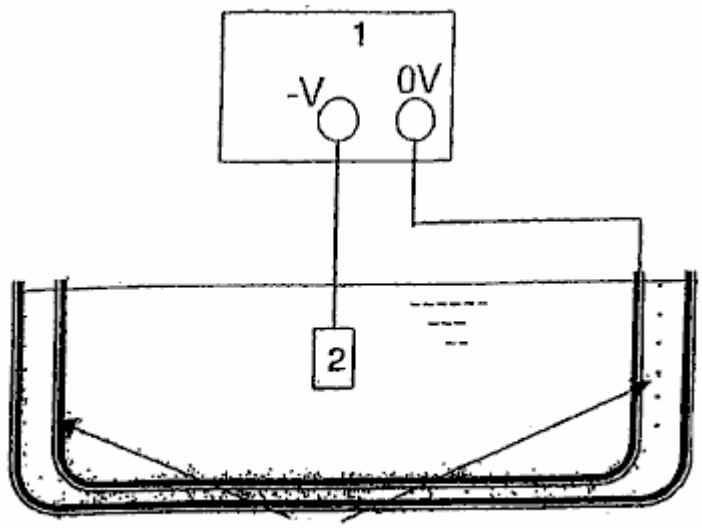


FIG. 6

Fig.6 illustrates an ultrasonic generator surrounding a bath similar to that in **Fig.5**. The generator generates ultrasonic waves which are transmitted into the electrolyte liquid and act to generate bubbles in the electrolyte which then surround the cathode. The cathode, which typically provides the surface for deposition, can be formed of a conductive material, a semi-conductive material or a non-conductive material, coated with a conductive coating. Cathodic materials that have been successfully used in this method are nickel, mild steel, stainless steel, tungsten and aluminium. The cathode can be in the form of either a plate, a mesh, a rod or wire. There may be any number of cathodes and the cathodes can be any shape or size. Any conductive material can be used for the anodes. Graphite, aluminium and stainless steel have all been successfully used to practise this method by the applicant. Generally, aluminium is preferred for the anodes. There may be any number of anodes and the anodes can be any shape.

In use, the bath is filled with an appropriate electrolyte. Broadly speaking, the electrolyte contains a solvent or carrier which provides a liquid environment, within which, electrolysis can occur and which also provides a support for plasma generation in the sense that it provides containment for the plasma generation. The electrolyte also contains a source of the material to be deposited in the form of a precursor. The electrolyte may also include additives for example for enhancing the electrical conductivity of the electrolyte and for assisting in bubble formation and a buffer to maintain a suitable pH in the cell.

In use, the article to be coated is placed in the bath where it typically forms the cathode. In some instances however, it may also form the anode. A voltage or potential difference is then applied across the electrodes and this voltage is set at a level that is higher than the firing point at which the system or cell achieves a stable glow discharge in which glow clusters envelope the cathode surface.

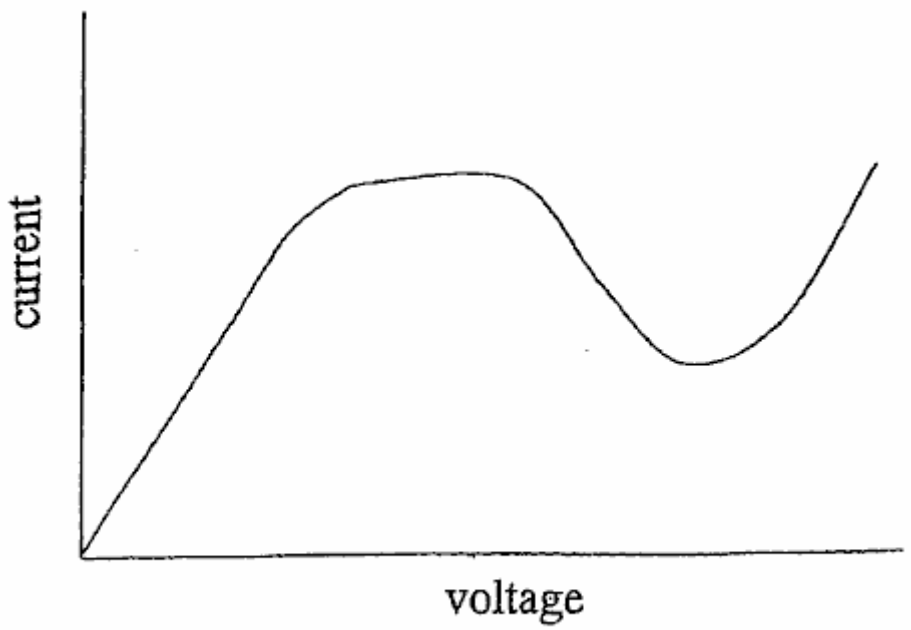


FIG. 7

Fig.7 illustrates a typical current against voltage profile for such a cell as the voltage is progressively increased. Initially there is an ohmic zone where the current increases proportionally with the voltage. After that the curve enters an oscillation zone where the current starts to oscillate. Applicant believes that this condition may be due to the fact that bubbles are evolving out of the solution and partly obscuring the electrodes. The bubbles form plasma, grow and then burst forming a shield shrouding the electrode. These bubbles block the conducting part of the cathode and this might lead to a decrease in apparent current density.

At the cathode, the evolved bubbles include hydrogen generated by the electrolysis of water in the electrolyte and by evaporation of liquid within the electrolyte. The bubbles may also be generated by other means as described above, for example ultrasonic generation. After some time, the number and density of bubbles increases until the entire cathode surface is sheathed in bubbles. At a critical voltage that is constant for a given system, known as the fire point, a glow discharge is formed. Experimental observation shows that this occurs when there is a near continuous bubble sheath around the cathode.

With a wire cathode, a tiny fireball or cluster of fireballs usually appears at the tip of the wire at the fire point. With further increases in voltage a glow discharge is established across the entire cathode. The glow discharge is dynamic and usually shows evidence of glow clusters and/or flashing through the bubble region. The glow discharge is caused by a dielectric breakdown in the bubbles. This is caused mainly by a high electrical field strength. Due to the presence of the bubbles the majority of the voltage drop from the anode to the cathode occurs in the near cathode region occupied by the bubbles. The electric field strength in this region may be of the order of 10,000 to 100,000 V/m. The voltage is set at a setting of 50 to 100 volts higher than the ignition point. This may typically mean a setting of 250 to 1500 volts. A preferred voltage setting would be at the low point of the graph in **Fig.4** within the glow discharge region.

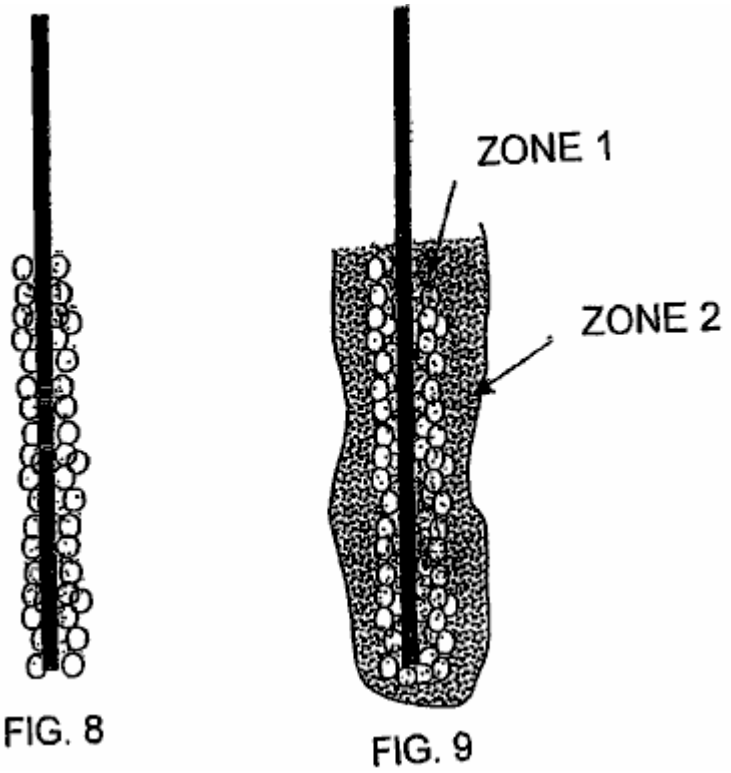


FIG. 8 FIG. 9

The glow discharge causes the generation of a plasma in the bubble. **Fig.8** shows the formation of a bubble sheath around the cathode. **Fig.9** shows the cathode during stable glow discharge. As shown in the drawings, applicant has observed the formation of two distinct zones during stable glow discharge. In zone 1 where the glow discharge clusters are present, there is a plasma envelope that directly shrouds the cathode surface. This envelope is where plasma deposition takes place. The plasma interacts with the cathode surface in a process similar to ion plating and deposition occurs. A film is progressively formed through nucleation and growth on the cathode surface. Zone 2 is a plasma-chemical reaction zone, which forms the interface between the electrolyte and zone 1. This zone envelopes the plasma deposition zone and is often clearly visible as a separate region with a milky appearance.

Dissociation, and possibly also ionisation of the electrolyte components, including the precursor, occur in the outer zone, zone 2. This gives rise to the species that are deposited on the cathode. The species is transferred from the outer zone 2 to the inner zone 1 by the electric field strength, diffusion, and convection. Deposition on the cathode then occurs for as long as these conditions are maintained and the precursor material is available in the electrolyte. After the glow discharge commences the temperature of the electrodes increases in a short space of time. The temperature of the electrolyte must be maintained within acceptable limits for certain type of application. To do this, electrolyte is drawn off from the bath and pumped through a cooling system as shown in **Fig.5**. The cooled electrolyte is then re-introduced into the bath. This cooling is required for both stability and safety reasons. Some of the electrolyte components are flammable. In addition electrolyte is consumed during the deposition reaction. Accordingly, it is necessary to top up the bath with additional electrolyte from time to time. A replenishment tank containing electrolytes is provided to perform this purpose.

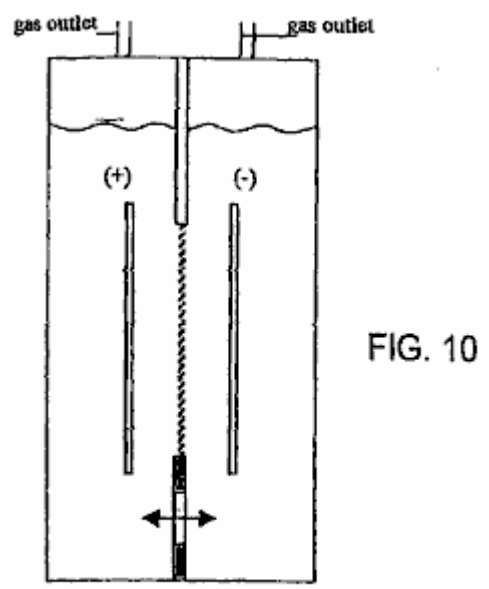


FIG. 10

BASIC TWO ELECTRODE REACTOR
WITH GAS SEPARATOR DIAPHRAGM

As shown in **Fig.10**, the reactor may include a pair of metal electrodes spaced apart and separated by an ion-conducting diaphragm. The electrodes can also be positioned horizontally or vertically.

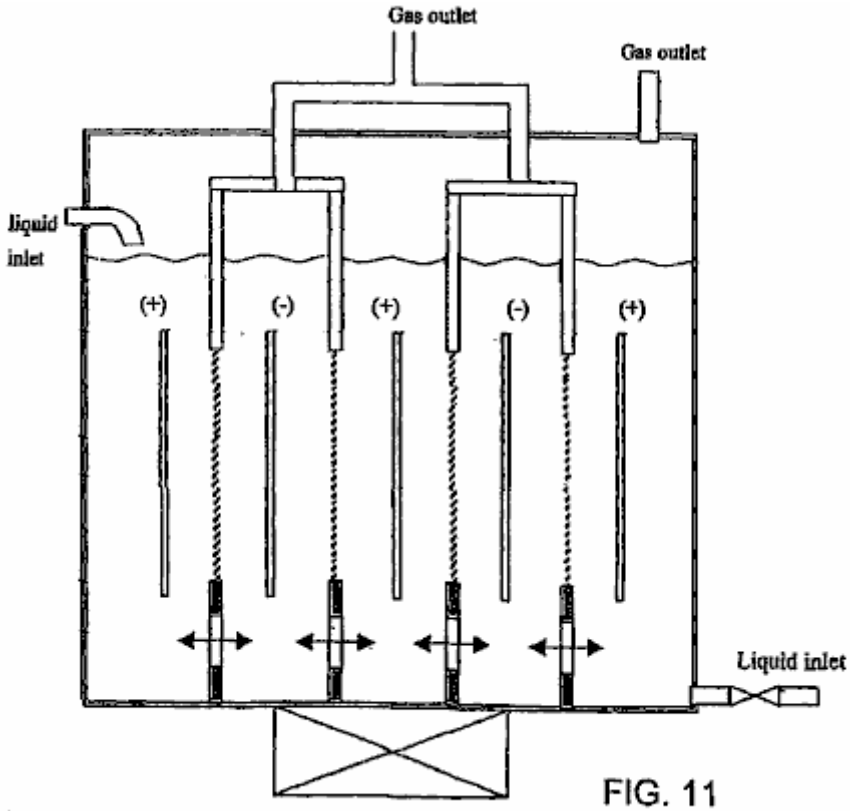


FIG. 11

Multiple Cell Reactors with common bath

As shown in **Fig.11**, the reactor may also include multiple pairs of alternating anodes and cathodes with a diaphragm. The diaphragm can be removed for decontamination and partial oxidation reformation process (**Fig.12**). In the case of reduction process, the hydrogen atoms produced on the side of cathode electrode are kept well separated from mixing back with oxygen by a diaphragm (**Fig.13**). It is possible to increase the throughput capacity of the reactor in treating contaminants with transverse flow through multitudes of alternating electrodes of anode and cathode (**Fig.14**). Wires or rods in tube reactors are suitable to adopt for hydrogen production and reduction process with the metal oxide confined within the narrow space within the cathode half cell and subjecting it to ultrasonic irradiation (**Fig.15** and **Fig.16**).

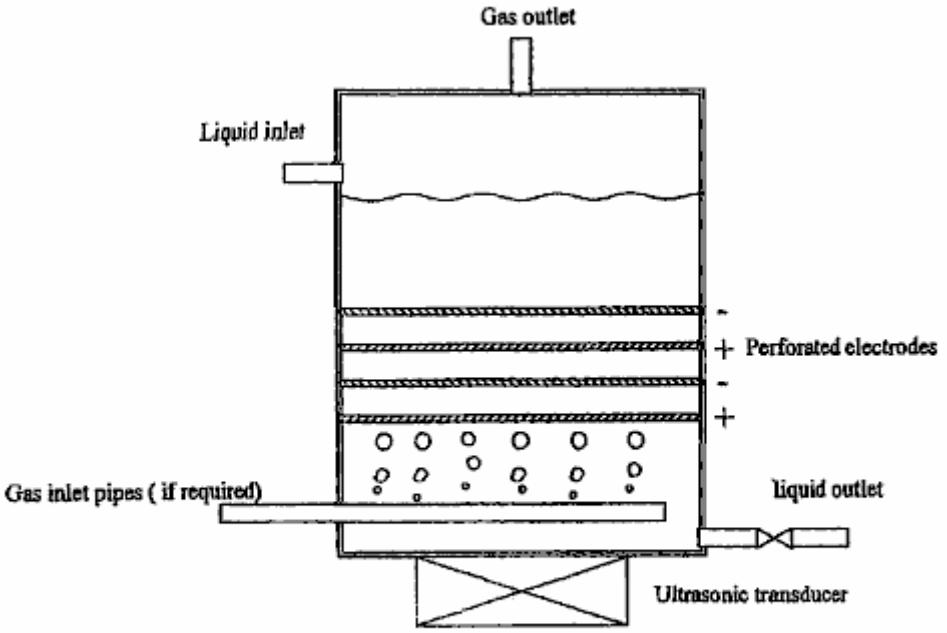


FIG. 12

Tower Reactor with Perforated Electrodes

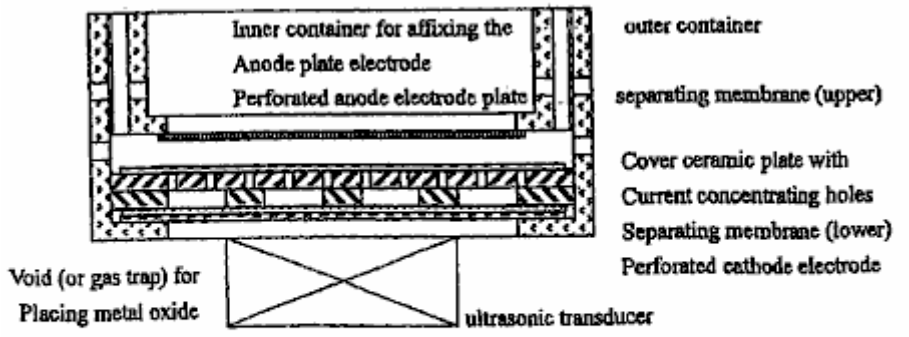
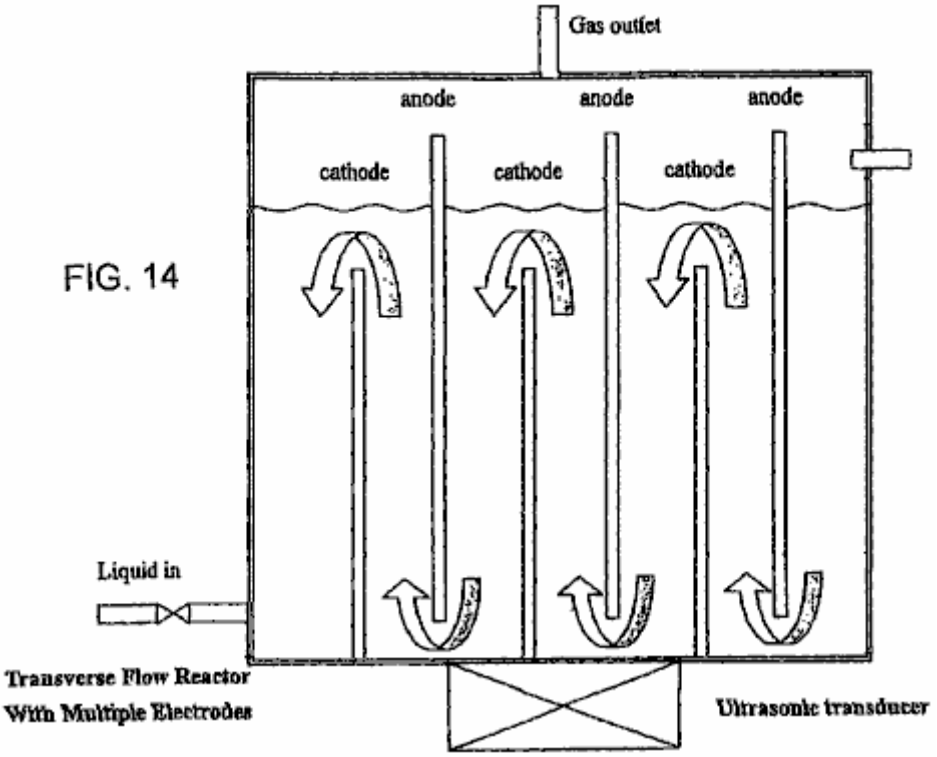


FIG. 13

Reactor for Metal Oxide Reducing Process
(which is to be placed inside an electrolytic bath)

FIG. 14



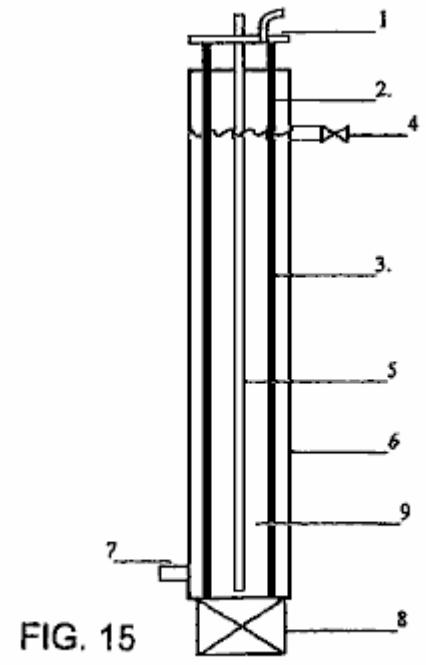


FIG. 15

WIRES OR RODE IN TUBE REACTOR
THE OUTER ELECTRODE SERVING AS BATH

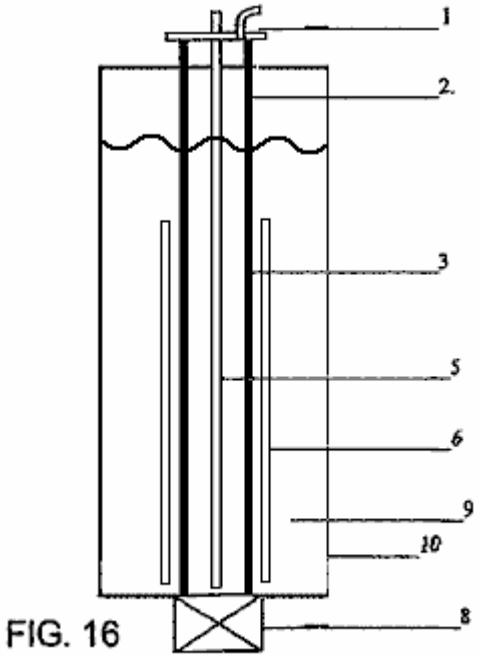


FIG. 16

WIRES OR RODE IN TUBE REACTOR

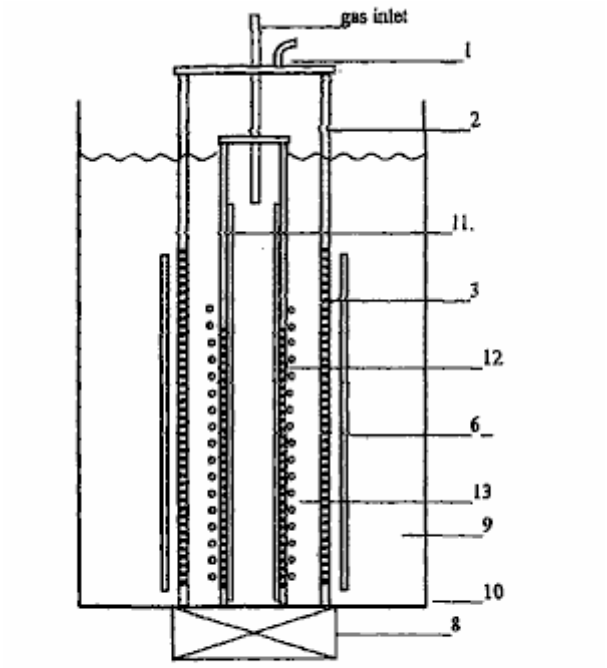


FIG. 17

TUBE IN TUBE REACTOR WITH PERFORATED INNER TUBE COVERED WITH GAS-TRAPPING COVER LAYER

Tube in tube reactor (**Fig.17**) has a tube electrode within the outer tube electrode instead of wire or rod. The inner tube is covered with non-conductive materials of suitable thickness with small diameter holes and gas trap forming in between the inner metal tube which also have small holes formed correspondingly. The gap between the outer electrode and inner electrode is kept close but giving a minimum 3 mm to 5 mm space between the separation diaphragm and the dielectric cover of the inner electrode, to allow free flow of electrolyte and gas. Bubbles of gas will be discharged into the plasma discharging zone with hydrocarbon rich gas i.e. methane, natural gas, H_2S to undergo reformation for the production of hydrogen gas. It can also be adopted for decontamination of polluted gas laden with NOx, SOx and particulates; and reduction process where the metal oxide will flow through the space between the electrodes with the ultrasonic irradiation keeping the fine powder in colloidal and at the same time hydrogen gas or methane gas may also bubble in to provide the extra H_2, H+ and CO to enhance the reduction process.

Configuration of Electrode with Gas Trapping and Current Concentrating Cover Arrangements

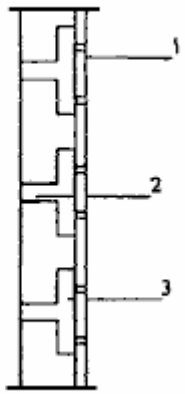


FIG. 18A

Type 1

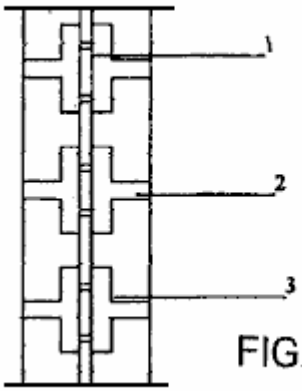


FIG. 18B

Type 2

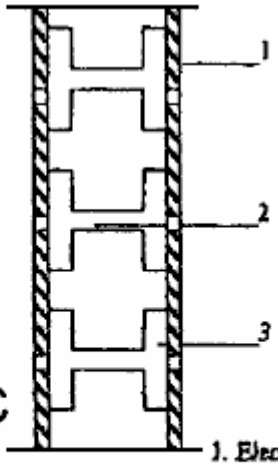


FIG. 18C

Type 3

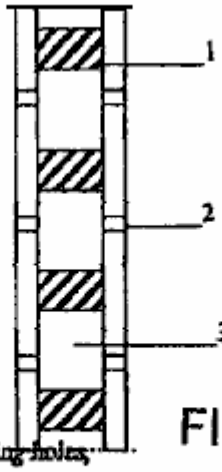


FIG. 18D

Type 4

1. Electrode, 2. Current concentrating holes, 3. Gas traps

Cross Section of Electrodes showing the Gas Trap and Current Concentrating holes

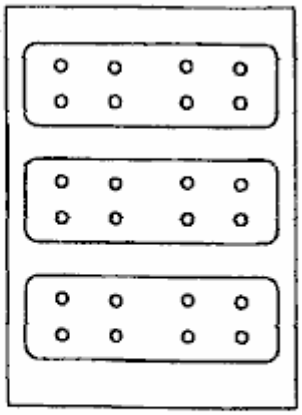


FIG. 18E

Type 5

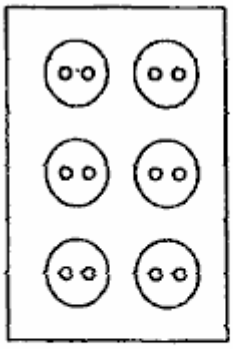


FIG. 18F

Type 6

Plan or Elevation of Gas Trap Cover Plate (with perforated current concentrating holes)

A number of gas trap and bubble retaining arrangements are shown in **Fig.18A** to **Fig.18F**.

The under-liquid plasma discharge, in order to produce various reductive, oxidative, radicals and neutrals species through excitation, ionisation and dissociation of the liquid molecules and atoms, requires high voltage input DC or AC, normally within 3 KV and current density under 3 Amp/cm^2. The electrodes cathode and anode have to be kept as close as possible but not close enough to cause arcing. The electrode surface is preferably flat, even and smooth with no pronounced irregularities. Because of the need of placing diaphragm and complementary gas trapping and retaining construction on the discharging electrode, a minimum distance of 6 mm to 15 mm has been experimented with and shown to produce steady glow plasma under-liquid. With better material choice and engineering capability, there is no reason why the electrode space distant cannot be further reduced. The size, shape and arrangement of the electrodes is not restricted, but the electrodes will usually be somewhat smaller than those required for conventional electrolysis, for the same gas production volume. Both the electrodes, anode and cathode, can be at work at the same time as the plasma discharging electrodes especially if a gas-trapping dielectric cover construction is provided.

Experiments have been conducted to establish the basic criteria to generate steady and rapid cyclical non-thermal plasma glow discharge under-liquid with basic DC high voltage and low current input at atmospheric pressure and ambient temperature leading to the proposal of a phenomenal model of reactor structure and electrode configuration which demonstrate the usefulness of bubbles or gas pocket that creates the under-liquid environment for plasma discharge and it also provides the back ground of further improvement and construction of reactor unite which verify the inventive idea of under-liquid plasma and it subsequent practical applications.

A reactor according to the present invention can basically follow that of a simple water electrolysis cell with one anode electrode separated from the cathode electrode with an ion conducting membrane and yet has the capability to prevent re-mixing of the produced gas on each half-cell. The electrolyte allows moving across the membrane or replenish through the opening in the reactor. In order to increase the proficiency of the reactor the cathode electrode is placed inbetween two anode electrodes and separated from them by a membrane. The hydrogen gas produced is isolated and collected independently. The polarity of the electrode can be reversed with the anode electrode in the middle when oxidative species are needed for the decontamination process. Most importantly, the simple electrode and reactor unit will form the basic module, placed inside a common bath and linked together to form a lage production unit, and these modules can be replaced individually.

Despite the apparent success of the simple perforated plate-to-plate electrode arrangement, it does not preclude other electrode configurations and arrangements such as tube in tube, wire in tube and other flat surface electrodes having different surface structure e.g. wire mesh, expanded metals, pinned plate, sponge porous metal, corrugated plate etc. as long as it is a good electric conductor, corrosion resistant, heat-tolerant material, i.e. stainless steel, aluminium, graphite, platinum etc. The shape and size of the electrode piece is not restricted and sometime it may form the object article which is to undergo plasma surface enhancement treatment.

In practice, a reactor with vertical electrodes, suits plasma-assisted water electrolysis, reformation of hydrocarbon liquid fuel, production of nano materials and decontamination process, while the reactor with horizontal electrodes suits reformation of hydrocarbon gas such as natural gas, methane, hydrogen sulphurs and the like.

This ability to generate steady plasma discharge, can well be adopted for other useful purposes such as thin and thick-film deposition and additional method in the creating of cold fusion.

There have been a series of experiments conducted to generate non-thermal plasma under-liquid by utilising the gas bubbles self generated during electrolysis, electrochemical reaction, heating and releasing of dissolved air or gases in the liquid. Bubbles can also be produce with the influence such as transient bubbles created by shock waves resulted from pulsed power input, ultrasonic cavitations, laser heating and hydraulic impingement. External introduced gas (e.g. air & fuel gas) is found to work well in providing bubbles environment for ready plasma discharge in a steady manner. A number of experiments have also been conducted to test the applicability of under-liquid plasma in the field of hydrogen generation, hydrocarbon fuel reformation, sterilisation and decontamination and reduction of metal oxide. Because of the restriction of the power converter that some result is less than ideal but it all indicate the potential of the under-liquid plasma which is in the first place having the same physical/chemical capability as its counter part operating in gases environment in exciting, ionisation and dissociation, but with some distinctive advantage which has well been described in the foregoing text.

Generation of steady plasma discharge under-liquid has been one of the primary objectives in the research. In general the generation of steady plasma glow discharge are influenced by a number of factors, such as physical and chemical properties of the liquid, its conductivity, temperature, electrode type, electrode spacing, gas retaining or trapping arrangement, current density, voltage input, reactor construction, liquid circulation, influence of ultrasonic irradiation, pulsed power input etc.

There are of course a number of electrode shapes, size and configuration one could choose. In order to find out the how important is the supply of bubbles or gas pocket affects the generation of plasma, a gas retaining or

trapping covering with current concentrating conducting holes over perforated plate electrode is formulated, which has proved effective producing steady glow plasma discharge within the range of 350 V to 2 KV (2,000 V) and current up to 850 mA, but most the time around 100 to 300 mA range. This is considered low in compare with other under-liquid plasma system (i.e. Plasma arc, pulsed high voltage and current electric discharge). Throughout the experiments, a horizontal reactor was used. However an alternative reactor is a vertical reactor.

INTRODUCTION TO THE EXPERIMENTS

Several groups of experiments have been conducted:

1. Preliminary trial experiments
2. Plasma assisted water electrolysis
3. Reformation of methanol
4. Reformation of emulsified diesel
5. Reformation of LPG as hydrocarbon gas (methane is not available in the market)
6. Decontamination or sterilisation of food drink
7. Reduction experiment of TiO_2.

In the preliminary trial experiments a number of electrode types have been adopted and have eventually select the wire to plate configuration and perforated plate to perforated plate or wire mesh as the most suitable under the limiting power supply condition where max. voltage available is 2,000 V and the maximum current is 1,200 mA. In reality, the current input is voluntarily restricted to work below 900 mA for durations not exceeding 30 minutes, to avoid damage to the converter which has happen in a number of occasion which caused stoppage of the experiments for weeks.

To overcome the power supply limitation, and to achieve steady plasma glow discharge, a gas-retaining or trapping cover or layer with current concentration holes has been devised to cover the discharging electrode surface (perforated electrode plate) which is the basic features adopted in the construction of reactor.

In the trial experiments, it has been demonstrated that infrequent visual plasma discharge begins with a voltage of 350 V and steady plasma can be achieved in around 550 V. The initial current input reaches 850 mA and begins to fluctuating in the range of 150 to 650 mA. On many occasions the current fluctuated at 100 mA to 350 mA.

Through these experiments, the mechanism of generating bubbles or gas pocket dielectric barrier which impedes the current flow, leading to an increase of voltage until a threshold voltage is reached which causes the electric breakdown and the formation of plasma inside the bubble, at which point the current immediately returns to its normal level and then another cycle of discharge is established. When the discharge is infrequent it resembles a corona streamer discharge but as the voltage increases, the glow discharge becomes a continuous glow over an extend electrode surface resembling a glow plasma discharge. The colour of the discharge appears as an orange-yellow or red colour in the electrolysis of water and the temperature of the discharging electrode ranges from $50^{O}C$ to about $90^{O}C$ and the temperature of the bath liquid ranges from $40^{O}C$ to $70^{O}C$. No sign of any damage to the electrode or its covering plastic gas trapping plate was observed even after prolong experimentation. When the voltage is allowed to increase beyond the glow plasma region, a plasma arc begin to occurs and becomes an intensive bright blue discharge when voltage is further increased and this causes damage to the metal electrode and plastic covering plate which is easily seen.

On two occasions, hydrogen production was recorded which produced a gas volume with an equivalent energy conversion efficiency up to 56%. Due to damage to the reactor by the plasma arc, that particular experiment cannot be repeated as new model of reactor is designed to achieve low current input and early high voltage response. However with the apparent success of the trial experiment, it shows that a more suitable reactor can be designed specifically for the purpose of hydrogen production by plasma assisted water electrolysis and a higher energy efficiency figure can be achieved with a small reactor.

PLASMA ASSISTED WATER ELECTROLYSIS

Experiments to check the behaviour of plasma discharge at different voltage input levels were carried out. Despite the apparently large volume of bubbles boiling inside the reactor, the total volume of gas produced was unexpectedly low. This may have been caused by the horizontal reactor design adopted throughout the experiments. This may have allowed the hydrogen gas recombine with the hydroxyl ions and convert back into water again. A vertical reactor would be more suited for the plasma assisted water electrolysis where the produced hydrogen gas will rise quickly to the top of the reactor and can be channeled away from the area filled with OH ions.

In this experiments plasma discharge begin to occur at 1,350 V with current fluctuating around 100 mA to 200 mA. At about 1,550 V the reactor produced highest volume of gas. Plasma arc discharge occurs at 1,900 V and is

becoming vigorous when the voltage is increased further. KOH of 0.02% concentration has been used as electrolyte additive throughout the experiment.

The production of gas appears to have a linear relation with time but various substantially with different voltage input. The rate of energy consumption is increasing slowly with time in a constant rate which various with the voltage input and its corresponding energy consumption per unit gas volume produced is having a peak at the first 10 minutes of the experiments and level off with time. The temperature in the electrode rise sharply to from 50^{O}C to 90^{O}C and is maintained more or less at that level throughout the test. The temperature in the bath liquid within the reactor rises slowly from its ambient temperature to around 50^{O}C to 55^{O}C.

EXPERIMENTS WITH METHANOL

Several sets of tests have been conducted with the aim of finding out how different hydrocarbon fuels will be affected by the non-thermal plasma under-liquid system. A methanol / water mixture with methanol concentrations of 5%, 10%, 15%, 20%, 25%, 30% and 40% were tested using the same method and equipment set-up already used for the plasma-assisted water electrolysis. There are three independent tests for each methanol concentration. It has been observed that the gas production is peaked at 25% methanol concentration and the energy consumption per unit gas volume produced is also lower than the others and is nearly at constant rate around 0.0225 Kw.h/L. The voltage input for each test is kept at 1,850 V and the current fluctuating in the range of 100 mA to 200 mA. The temperature measured at the cathode electrode started at 80^{O}C and rose quickly to reach over 200^{O}C at the end of a 30 minute experiment. The temperature recorded in other tests stayed within the range of 60^{O}C to 80^{O}C. The temperature of bath liquid at 25% concentration stayed in the range of 50^{O}C to 60^{O}C, which is typical for each of these tests.

The greatest surprise coming out of the experiments is that the produced gas is composed of two gases. One is hydrogen gas and the other is oxygen gas and no trace of carbon dioxide is found. Repeated examination of the gases produced shows the same result and the hydrogen is having an average value of 51.3% and oxygen 48.7%. This is later found out that the presence of oxygen in the gas is the result of the removal of the separating diaphragm. An acidic electrolyte is preferable in order to increase the hydrogen gas percentage in the output gas mix. This is shown in the latest experiments using sulphuric acid of 0.02% concentration.

A set of experiments with the use of 40 KHz ultrasonic bath having methanol concentration of 10%, 15%, 20% and 25% with the same reactor and equipment arrangement have been conducted to find out the influence of ultrasonic radiation. It has been observed that gas production at 25% is substantially higher than the others and yet the energy consumption per unit gas volume produced is around 0.015 Kw.h/L throughout the 30 minute experiment, which is lower than that without ultrasonic radiation.

The chromatographic analysis of the output gas having an average value of 97.56% hydrogen and 2.4039% of carbon monoxide. Chromatographic analysis of gas produced by reformation of methanol with ultrasonic radiation. Methanol concentration at 25%, and conductive reagent 0.02% sulphuric acid.

TABLE 1

Test	Resident time minutes	Composition V/V %	Gas type
First Test	0.364	98.9937	H_2
	1.047	1.0063	CO
Second Test	0.364	96.7418	H_2
	1.047	3.2582	CO
Third Test	0.354	96.9719	H_2
	1.048	3.0281	CO
Average		97.5691	H_2
		2.4309	CO

EXPERIMENTS WITH LPG

Decomposition of LPG by under-liquid plasma has been conducted (methane or natural gas is preferred but none is available in the market). The LPG is allowed to pass through the horizontal reactor through the perforated anode plate and enter the reactor and trapped at the cathode plate where plasma is taking place at voltage 1980V and current at 100 to 130 mA input. C_3H_8 and C_4H_{10} are the two main components of LPG, it is expected that the volume output having been subjected to plasma dissociation should be larger than the original input volume. This is found to be so that the output gas volume increases by about 50%. The experiment is conducted together with ultrasonic radiation. It is regrettable that the chromatogram is incapable of undertaking analysis of the output gas composition. The next set of experiments should be conducted with methane or natural gas so that more definitive result could be obtained. Rudimentary analysis of the produced gas has shown the presence of H_2, CO_2 and C_3H_6 etc.

REFORMATION OF EMULSIFIED DIESEL AND WATER WITH ULTRASONIC IRRADIATION

Decomposition of emulsified diesel with distilled water has also been carried out. Diesel oil in 25% and 50% by volume has been emulsified by adding 1.25% emulsified agent inside the ultrasonic bath. Since the diesel oil is dielectric, a KOH additive is needed. The emulsified liquid is subjected to plasma discharge at a voltage of 1,850 V and a current fluctuating from 100 mA to 200 mA for a period of 30 minutes. The temperature of the cathode electrode increased from 70^{O}C to about 94^{O}C during the experiment. The gas volume produced was 160 ml with 25% diesel and 1,740 ml with 50% diesel, which is substantially higher and its energy consumption is 0.1213 KWh/L. It is clearly indicated, that gas production is proportional to the diesel contend in the emulsion. Because of the limited power supply capability, the voltage of 1,850 V is merely adequate to produce some plasma discharge but it is far from establishing extensive vigorous plasma with higher current and voltage input, which would produce more gas.

STERILISATION (DECONTAMINATION) OF MULBERRY FRUIT DRINK

The ability of non-thermal plasma to decontaminate noxious chemicals and gases has already established. This experiment is conducted to find out how well the under-liquid plasma may apply in the field of beverage sterilisation with low levels of plasma radiation and keeping the treated liquid within an acceptable temperature.

Two litters of 15% concentrated fruit drink is placed in the bath where a horizontal reactor is submerged. The bacteria count and mold colony count is obtained before the forty minute test. A sample of the fruit drink is extracted at 20 minutes and 40 minutes. The mulberry drink has good natural conductivity so no additive is required. The applied voltage is kept at 1,200 V and the current fluctuates around 200 mA. The temperature at the electrode is maintained at around 62^{O}C and the bath liquid (fruit drink) is kept at around 50^{O}C.

TABLE 2 - The micro-organism count

Time (minutes)	Bacteria count/ml	Mold colony count/ml
0	3,400	37,000
20	1,300	17,000
40	90	10

The favour and colour of the fruit drink had not changed after the test. The bacteria sterilisation is 97.5% and that of mold colony has been sterilised more than 99%. This has given proof that the under-liquid plasma has the same capability as those operated in a gaseous environment.

The time for the treatment could be reduced by providing forced circulation of the liquid and increasing the electrode size. Sterilisation of drinking water imposes no limit on the temperature. Higher voltage input for better plasma glow discharge spreading over larger and multiple electrodes should be able to remove all harmful chemical substance, bacteria, biological matter and microbial matter, thus meeting the municipal requirement for drinking water.

REDUCTION OF METAL OXIDE

One trial experiment to reduce TiO_2 back to Titanium metal has been attempted with little success. It was found that in the X-ray diffraction test, minor traces of titanium nitride and titanium monoxide (TiO) were found. In the experiment, only a minor electrolyte of 0.05% KOH with 25% methanol added to the distilled water was used to increase the production of hydrogen. The applied voltage was fixed at 1,850 V and the current fluctuated in the range of 200 mA to 500 mA. Ultrasonic radiation up to 40 KHz was also provided through an ultrasonic bath. The temperature recorded in the bath liquid rose from $46^{O}C$ to $75^{O}C$ at the end of the 60 minute test. The fine TiO_2 with was suspended with ultrasonic radiation, in the bath liquid in colloidal form, showing as a milky white colour, which gradually became a milky yellow colour towards the end of the experiment. The bath liquid also became viscous.

The X-ray refractive "d" value of TiO_2 were:

Before the experiment: 3.512, 1.892, 2.376 but after the experiment there were two new groups of "d" measurements not seen before the experiment:
a: 2.089, 1.480, 2.400
b: 2.400, 2.329, 2.213
This indicates a new material, positioned between TiO and n-Ti_3N_2-x.

This experiment indicates that a change did happen to the TiO_2, possibly because of the limited voltage and current available as input, which could not provide the intensity of plasma discharge needed to effect the reduction process properly. Higher concentration of either HCl or H_2SO_4 should be use as reagent demonstrated in the following chemical reaction and in the same time serving as electrolyte. The horizontal reactor is not a suitable piece of equipment to undertake such experiment; it is adopted merely for convenience. A wire-in-tube and tube-in-tube reactor would be a suitable candidate, which would keep the metal oxide exposed to plasma discharge throughout the whole of the duration of the experiment. Further, more hydrogen or CO gases produced during the process may be passed back to the reactor to enhance the reaction. (Methane is a suitable gas for this type of reduction process, as both hydrogen and CO gas will be produced to enhance the reaction). The following are the chemical formula, which suggested by transforming TiO_2 to either $TiCl_4$ or $TiOSO_4$ as a soluble ionic compound, will facilitate its reduction with prolong exposure to active atomic hydrogen under the influence of a plasma catalytic environment.

$TiO_2 + 4HCl \rightarrow TiCl_4 + 2H_2O$,
$TiCl_4 + 4H \rightarrow Ti + 4HCl$.
$TiO_2 + H_2SO_4 \rightarrow TiO(SO_4) + H_2O$,
$TiO(SO_4) + 4H \rightarrow Ti + H_2SO_4 + H_2O$
Where $TiCl_4$ is readily produced by an established process from ilmenite.

Similarly, aluminium oxide Al_2O_3 can first be transformed to $AlCl_3$, which is soluble ionic compound, ready to be extracted by electro-deposition enhanced with plasma-reduction and plasma-electroplating process:

$Al_2O_3 + 6HCl \rightarrow 2AlCl_3 + 3H_2O$,

$2AlCl_3 + 6H \rightarrow 2Al + 6HCl$.

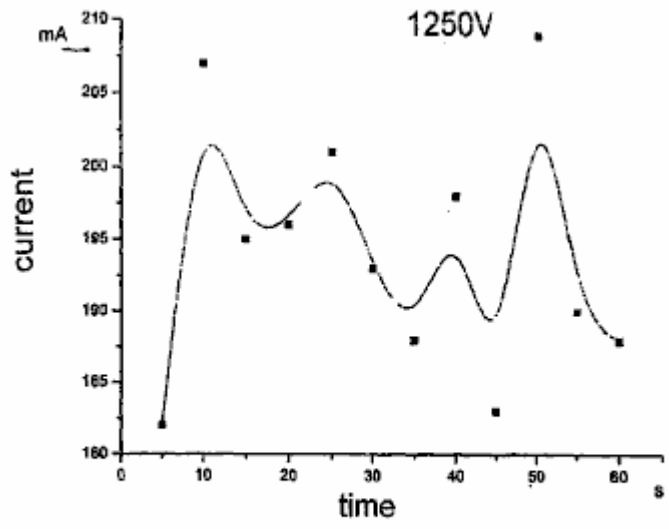
In the case of electrode positive oxide such as Fe_2O_3, it can be reduced in the presence of ionised atomic hydrogen and the presence of carbon monoxide with catalytic reactive plasma irradiation.

Fine metal oxide powder irradiated with ultrasonic waves will maintain in colloidal form allowing it to be exposed to the reduction agent atomic hydrogen and/or carbon monoxide. The process of ultrasonic cavitations and collapse is also known to create extreme localised high temperature up to 10,000^{O}K and thousands of atmospheres of pressure together with the high temperature at the impact point of the fine powder particles which is beneficial to the entire reduction process.

DETAILS OF THE EXPERIMENTS CARRIED OUT

Establishing Generation of Under-Liquid Plasma:

Distilled water is used in the experiments with 0.05% KOH as a conducting reagent. The voltage is controlled at 1,250 V & 1,850 V. The current is raised in steps of 100 mA until it reaches 850 mA. In the beginning the voltage remains low and gradually builds up as more gas bubbles are generated. Once it reaches a certain high level the current drops immediately. The self-regulating current and voltage input of the power unit automatically switches from current input control to voltage input control. At 45 seconds after switching the experiment on, the voltage rose to 470 V and the current dropped below 500 mA. From 3 min. 10 sec to 5 min 20 sec, the voltage rose to a relatively high level while the current kept on fluctuating. After a period of unstable voltage and current movement they become stabilised at 20 min with the characteristic high voltage and low current. At this instant prominent glow is observed at the perforated cover plate (current concentrating holes). The temperature of the cathode electrode has risen and stays steady at around 70^{O}C.



Glow discharge at stable 1250V input

FIG. 25

Fig.25 shows the current fluctuating with stable 1,250 V voltage input and a steady plasma glow discharge. The temperature of the cathode increases rapidly in the early stages and then becomes steady at the 5 min mark, and then rising slowly to it's highest temperature of about 96^{O}C.

OBSERVATION

Generating Under-Liquid Plasma:

In accordance with the experimental results, it is possible to generate non-thermal plasma under-liquid providing that certain conditions are met: a suitable power supply, electrolytic liquid, reactor and other supplementary equipment.

The design of the reactor, with relatively low voltage and limited power rating (restricted current input) requires special construction to trap or retain gas and at the same time to raise the current density at the discharge area.

The gas trap or chamber should be of a suitable size. If the gas trap or chamber is too big, then the trapped gas is too thick which requires a much higher voltage for discharge breakdown and prolongs the time of each cycle of discharge. It becomes difficult to maintain rapid cyclical steady glow discharge. The perforated covering plate, is also an important part of the electrode structure, concentrating the current density. The thickness of the perforated plate and the size of the gas trapping chamber should be carefully controlled so that the electrode spacing gap is not unduly wide as that also influences the voltage requirement. The size and disposition of perforated holes can be determined by trial and error. Wide electrode spacing increase the voltage input requirement and unsuitably close electrode spacing will cause early occurrence of plasma arcing with high current surge and generation of temperatures which will damage the electrodes and their attachments.

The power unit should be of adequate power rating. The electric breakdown is highly dependent on the high voltage supply. If the rating of the power supply unit is inadequate, it could easily be damaged during sudden the high current surge caused at cyclical electric breakdown. There will be no plasma discharge if the power input is inadequate.

The electrolytic liquid should have suitable conductivity, not too low nor too high. Voltage cannot be easily raised between two electrodes the liquid has high conductivity and no plasma discharge will be generated unless there is a high voltage input. The discharging electrode may be fully encapsulated inside a bubble barrier, but high conductivity liquid allows the current to pass through the bubble-liquid interface which in turn, also prevents the voltage rising high enough. If the conductivity of the liquid is too low, then the bubble barrier forms a complete dielectric barrier which requires a much higher inception voltage to cause electric breakdown or discharge and at the same time, the passage of current becomes too low which results in a low current density which also influences the occurrence of discharge. A much higher breakdown voltage (discharging voltage) creates electric arcing in gaseous condition which is no longer considered non-thermal under-liquid plasma discharge.

CONCLUSIONS

1. Gas layer or bubbles form the dielectric barrier that provide the environment for building up the discharge voltage and gaseous space for plasma discharge to take place. High voltage and relatively low current input is characteristic of under-liquid plasma.

2. With the characteristic high voltage and low current requirement, the under-liquid plasma can be generated over a wide range of liquids. The electrolyte liquid can be acidic, alkaline or a solution of salts. Liquids containing conducting impurities or a mixture of organic compounds may also serve as electrolyte such as the case of tape water and fruit drinks.

3. There are a number of factors which would affect the generating of under-liquid plasma such as voltage, current density, configuration of electrodes, area of electrode surface, electrode gap spacing, electrolytic physical and chemical properties, gas retaining and trapping arrangement, provision of plasma enhancement, ultrasonic cavitations, pulsed power supply, ambient temperature and reactor construction. This appears complicated, but the experiments undertaken have demonstrated that all the mentioned factors can be manipulated to achieve generation of stable non-thermal plasma at one atmosphere of pressure.

4. Plasma is the fourth state of matter. It has been widely employed in the field of chemical, electronic, materials and energy industries. Plasma generated under-liquid plasma has its own intrinsic characteristics and advantages, which have already proved to be a useful tool for plasma electroplating or deposition of both metallic and non-metallic materials. It will find its application in the plasma-assisted water electrolysis for hydrogen production; reformation of hydrogen rich compounds or hydrocarbon fuel (gas and liquid); decontamination of both liquid and gas pollution discharges containing persistent harmful chemicals, dissolved heavy metals and organic and biological contaminants; sterilisation of fruit drinks, potable water supply; and reduction of material oxide such as oxide ores, metal oxide as an alternative method metal refinement. It is probable that the proposed under-liquid plasma generation, and this established basic scientific information, would form the basis for further refinements leading to the practical new applications put forward in this patent application.

PLASMA ASSISTED ELECTROLYTES FOR HYDROGEN PRODUCTION

Water electrolysis is still used for the production of pure hydrogen. This hydrogen production is restricted because of it's relatively low energy conversion efficiency. In order to achieve higher energy efficiency, the electric voltage must be kept low to avoid energy loss through heat conversion. There are also claims that the energy efficiency can be improved by better electrode configuration, an increase in the reactive surface area, reduction of the electrode gap and increasing the operatingpressure. The PEM solid electrode system is in its early development and its efficiency remains similar to that of water electrolysis system. In any case the basic principle of water electrolysis has not changed since it was first put to use. Electrolysis as a whole, is considered to be non-competitive with the competing production process of reforming hydrocarbon fuel, but electrolysis has the advantage of being a clean process producing high gas purity and CO_2 is not produced.

The hydrogen bubbles evolving from the electrode surface slow down with time when tiny bubbles gradually built up and smother the electrode surface. These are not easily dislodged and the rate of hydrogen production is reduced further as those tiny bubbles become a barrier to current flow between the two electrodes.

The proposed invention is closely related to the water electrolysis process but the mechanism of separating hydrogen from water molecules is different. Generating non-equilibrium plasma within the bubbles that smother the electrodes will break down the dielectric barrier bubble layer and cause the normal flow of current to be resumed. At the same time, water molecules contained in the bubbles coming into contact with the plasma discharge, will be dissociated to produce extra hydrogen. In addition, the vigorous plasma discharge near the electrode surface will also create an hydrodynamic condition, which will wash away the fine bubbles which block the current flow. The mechanism of producing hydrogen by plasma discharge is different from the conventional electrolysis which splits the ionic water molecules by electro-polarity attraction, while in the plasma discharge the water molecule is broken down as the result of electron collisions. The water molecules under the plasma discharge irradiation would lose one electron due to electron collision to yield $H_2O + e \rightarrow OH + H^+ + e$

The hydrogen produced is of high purity. Ordinary potable water or rainwater with a very low concentration of electrolyte can be used as the main source of material, instead of distilled water, as they contain sufficient impurity to be slightly electro-conductive.

The experiment has demonstrated that hydrogen gas can be produced with plasma glow discharge as a supplementary process to the conventional method. The energy required to produce 1 cubic meter of hydrogen with plasma glow discharge with a very rudimentary reactor has achieved an efficiency of 56% which can be further improved with better engineering, by closing the electrode gap distance, selecting the right concentration of electrolyte, reactor construction and better means of trapping and retaining gas near the discharge electrode.

High temperatures of up to $90^{O}C$ is recorded in the electrolyte, which increases within very short time of the reaction. This may in part due exothermic reaction of recombining H and OH to water. The excessive heat can well be utilised as secondary source of energy. The gas or vapour bubbles by heating assuming greater importance as source materials for plasma dissociation leading to the production of Hydrogen. The high purity oxygen co-produce is also a valuable by-product with many applications.

Since high voltage with moderate current is needed in the plasma process, the production rate per unite area of electrode surface is high, and so only a small reactor is needed for the production of hydrogen, especially when other plasma enhancement methods are employed, such as ultrasonic cavitations, pulsed powers and RF input.

The electrodes could be of any conductive materials such as aluminium, stainless steel, graphite, tungsten, platinum, palladium etc. The size of the electrode for the plasma discharge is much smaller than that required by the conventional electrolysis to produce the same quantity of gas. As a result of this, a smaller reactor is possible.

Sponge porous electrodes will increase the reactive surface area available to produce electrolysis gases. In the experiment, several layers of fine wire mesh were packed tightly together to mimic a sponge porous electrode plate.

Some of the basic electrode configuration is: plate to plate; perforated plate to perforated plate; plate or perforated plate to wire mesh; wire mesh to wire mesh; plate to pinned plate; dielectric coating on one or both electrodes plate or mesh or pinned plate, tube in tube and wire in tube arrangement. It is noted that electrode configuration including any lining or covering materials that help to concentrate the current density and having the ability in retaining gas around the electrode would be adopted which will help to lower the voltage and current requirement to generate steady plasma discharge.

In order to create an environment for steady and short cyclical plasma glow discharge as already mention in the previous text, the electrode configuration should be so structured to retain the bubbles and concentrate the current density and yet keeping the true electrode gap distance to a minimum. This creates a suitable voided

space either in the metal electrode or in the covering materials, capable of retaining gas while at the same time having the mechanism to concentrate the current density to a localised discharge point. This leads to a wide variety of designs and choice of materials to satisfy plasma discharge requirement.

In order to avoid recombination of H^+ and H_2 with OH ions and reverting back to water, the hydrogen atoms after regaining their lost electrons through contacting the cathode should be allowed to escape quickly from the area which abounds with other oxidation species and radicals. This has greatly influenced the productivity of hydrogen gas. If H^+ and OH is allowed to recombined, despite of the apparent bubble boiling in the reactor very little gas can be collected and the temperature in the reactor rises quickly which could well be the exothermic effect of recombination of H^+ and OH.

The hydrogen produced is collected separately from the oxygen. Since the produced hydrogen gas contains a fair amount of water vapour, the hydrogen gas is collected by passing it through a water chiller or other known method, so that the measured gas volume is at room temperature with minimum water vapour content.

The basic plasma assisted electrolysis cell or reactor can be produced in modular form which can be mounted side by side and placed inside a single electrolytic tank with their respective power and output gas collected to form a major production unit. Several reactor types can be employed for the production of hydrogen. Rod or wire in tube reactor, tube in tube reactor, single or multiple cell reactors are also suitable for the plasma assisted water electrolysis. The gas retaining and current concentrating cover will be affixed on the cathode electrode facing the anode electrode. A horizontal reactor whose cathode has a gas-retaining cover can be placed on top of an anode which is separated by a diaphragm and the hydrogen gas will then collect in isolation.

The introduction of ultrasonic cavitations into the electrolytic liquid is easy since the electrolysis bath is also the ultrasonic bath and ultrasonic transducers can be attached to the bath externally. A mixture of sonic frequency should be used to avoid any occurrence of a dead sonic zone. The introduction of sonic excitation through cavitations enhances the production performance of plasma-assisted electrolysis.

Pulsed high-voltage DC supply with single polarity square wave from 5 KHz up to 100 KHz has been found to be beneficial for generating plasma at a much reduced voltage.

The distinct advantage of the under-liquid plasma enables ionised species migrate to the respective half cell and electrodes which will avoid and minimise re-mixing of the produced hydrogen and oxygen causing a reversion to water again and creating a hazardous, explosive condition. The oxygen is considered as a by- product which can be collected for use or it can be channelled to the combustion chamber if hydrogen is used as direct fuel for a combustion engine.

Water is the primary source material for hydrogen production, being economically available and of unlimited supply. It is a completely clean source material that produces no unwanted by-products.

The anode may be gradually losing its materials due to electro transportation, but if so, it will be a very slow process. In practice the polarity of electrodes can be reversed which reverses the materials transportation and deposition. Conductor materials which are inert to electro-chemical corrosion are a good choice to serve as electrodes.

A chemically conductive reagent may be added to water to increase its conductivity and a foaming agent added to enhance generation of bubbles. The electrolyte can be of acidic or alkaline base. The concentration of the electrolyte should be maintained at a steady level for best results. High electrolyte concentration increases liquid conductivity as well as productivity of gas bubbles but it might prevent the rising voltage required for discharge as the current flow between electrode will not be inhibited by the presence of bubbles. However, a very low concentration of electrolyte will favour dielectric breakdown of bubbles, as a lesser current will be carried by the liquid medium inbetween the bubbles. It has been found that either acidic or alkaline electrolyte with 0.02% concentration work extremely well in maintaining steady glow discharge with DC voltage ranging from 350 V to 1,800 V and a current from 100 mA to 800 mA.

Tap water has been used without adding any conducting reagent and it often works unexpected well, most likely due to present of impurity and high pH, in the plasma-assisted electrolysis where steady glow discharge occurs at around 450 V to 900 V and current around 200 mA to 350 mA. The power input requirement varies in accordance to electrode spacing, electrode and reactor configuration, electrolyte concentration and the structure of gas retaining arrangement. Again other plasma assisted method such as pulsed power input and ultrasonic cavitations etc. also help to lower the power input requirement.

The process is in general, conducted at one atmosphere pressure. An increase of pressure will slow down upward movement of the bubbles and raise the temperature of the electrolyte. Some increase in temperature in

the electrolyte is not detrimental to the generation of plasma. Water vapour bubbles provide the source materials and active environment for plasma discharge. In general, electrolyte temperature is well below boiling point as non-thermal plasma produces little heat. The temperature sometime rises quickly in the electrolyte due to occurrence of infrequent plasma arc and exothermic in the recombination of H+ and OH- in quantity.

During the steady glow discharge, vigorous bubbles with yellow/orange/red colour light spots appear all over the plastic perforation. The light spots also appear widely on the electrode surface when the voltage is increased. On examination of the electrode and plastic cover sheet, no burn marks were observed. This proves that the plasma glow is non-thermal after an hour of glow discharge. The temperature in the electrode plate recorded with a thermal couple was around 50^{O}C to about 90^{O}C. The gas produced is composed mainly of hydrogen with some water vapour, which condenses quickly on cooling. The rate of hydrogen production is variable and energy conversion rate also fluctuated throughout the test. This is suspected to cause by the recombination of H and OH, which is affected by the electrode and reactor structure and configuration.

Hydrogen can now be produced with high voltage and low current, which is contrary to the conventional electrolysis system where a small reactor with a high rate of production is becoming possible. This has clearly demonstrated that the mechanism of producing hydrogen with plasma discharge is different from conventional water electrolysis in a number of ways. Steam and gas vapour produced due to heating of the electrodes (cathode) in short space of time are becoming an importance source of materials for plasma dissociation that also influence the productivity of hydrogen.

1.3 Experimental Procedure

1.3.1 A flow diagram for carrying out experiments in relation to this invention is shown in **Fig.28**.

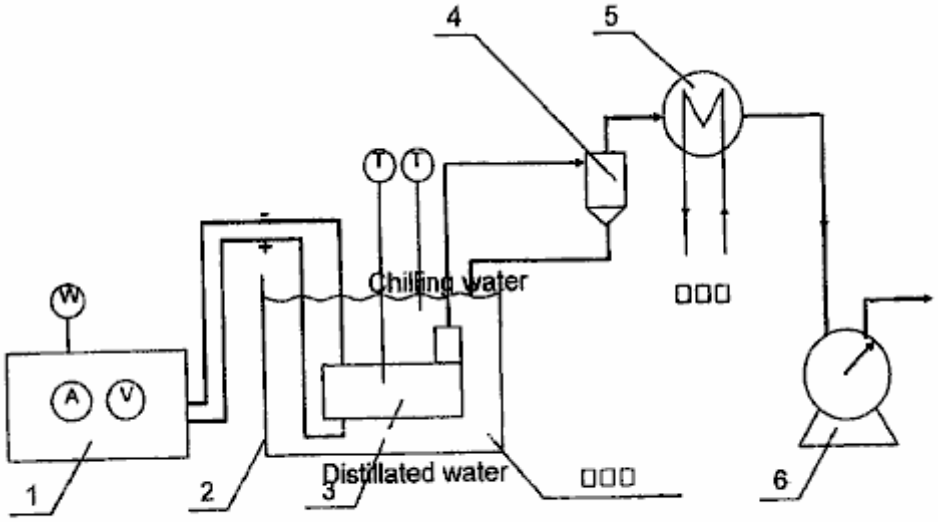


FIG. 28

The apparatus comprises broadly, a DC power source **1**, liquid bath **2**, reactor **3**, gas and liquid separator **4**, water chiller **5**, and gas-volume measuring meter **6**. Gas was produced by electrolysis which was catalysed by the plasma. Hydrogen gas was produced at the cathode and oxygen gas at the anode.

1.3.2 Equipment Function:

DC power source: provides high voltage DC.
Horizontal reactor: generation of non-thermal under-liquid plasma.
Gas and liquid separator: to separate liquid from gas and return as chilled liquid.
Chiller: to condense any liquid vapour admixed in the gas and return to reactor.
Gas-volume measuring meter: to measure the volume of gas flow.

1.4 Method and Operation of the Experiments

(1) The experiment is conducted in according to the occurrence of plasma discharge. Six different levels of voltage are selected to produce under-liquid plasma with same reactor for the generation of hydrogen. They are: 1350 V, 1450 V, 1550 V, 1650 V, 1750 V, and 1850 V. Each experiment lasts 30 minutes and the experiment is repeated three times under the same set of conditions. The data obtained are than averaged out.

1.5 Experimental Observations

Plasma discharge at 1,350 V is observed to have few and limited lighting illumination on the electrode in comparing with those vigorous, steady discharging over a much larger electrode surface at voltage 1,850 V. The corresponding current input is also very much reduced. It has been recorded that the temperature at the cathode electrode rises with time until it reaches about $90^{O}C$ and gradually becomes steady. The colour of the plasma discharge appears to be orange and red and it's colour is greatly different from that of electric arc (plasma arc discharge) which appears to be sharp bright blue in colour.

Applicant also conducted experiments with the same equipment utilising the under-liquid plasma to transform methanol for use in hydrogen production. Applicant found that the plasma was efficacious in producing hydrogen gas from the methanol. CO and CO_2 gases were completely absent from the gas produced. This was unexpected. Without being bound thereby, Applicant believes that CO and CO_2 may have been absorbed by KOH which was added as a conductive agent to the electrolyte. Some oxygen gases were recorded before methanol was added to the electrolyte.

Applicant also conducted experiments with the same equipment utilising the under-liquid plasma to reform hydrocarbons for hydrogen production. Applicant found that the plasma was efficacious in reforming the hydrocarbons and producing amongst other things hydrogen gas.

Applicant also conducted experiments with the same equipment utilising the under-liquid plasma to treat diesel oil. The diesel oil was emulsified in water to disperse it through the body of liquid. After being subjected to plasma conditions near the cathode, a gas was produced that was smoky and resembled an exhaust gas emission that did not easily burn. Applicant established by means of these experiments that diesel oil could be reformed and also dissociated by the in liquid plasma with this equipment.

Reformation of hydrocarbon liquid and gas fuel, and hydrogen rich compounds for hydrogen production:

Water is one of the primary source materials, which serves as carrier, conductor and confinement to the bubbles space where plasma corona and glow discharge would take place when adequate electro-potentials apply across single, or multiple electrodes pairs. The hydrocarbon fuel methane (gas), methanol, diesel, gasoline, kerosene (paraffin), ethane, natural gas, LPG gas, bio-diesel etc. and hydrogen sulphur (H_2S) are also good source material for hydrogen production.

The majority world-wide of hydrogen production conventionally is by high-pressure steam reformation of methane. This requires high pressure and high temperature. The production plant is large and costly to set up. Storage and delivery in association with the production are an added cost for the supply of hydrogen gas. The importance of hydrogen as an alternative environmentally clean fuel is well understood. The upcoming fuel cell technology demands an economic and ready supply of pure hydrogen gas. To produce hydrogen with a small processor to enrich fuels for combustion engines and gas turbines will not only be reducing fuel consumption but it also reduces polluting emissions.

The proposed plasma reformation process can deal with both gaseous fuel and liquid fuel. The gas fuel will be bubbled into the reactor along with an inhibitor to slow down the upward flow of the fuel gas. Since the dissociation of the hydrocarbon fuel will be mainly achieved by plasma dissociation which is similar to the plasma-assisted electrolysis process, but with electrolytic liquid containing hydrogen rich compounds. In the case of liquid fuel, it can either form a mixture with water or be emulsified with water. The percentage of fuel in the mix depends on the type of fuel, its conductivity, boiling point, flammability and electrochemical reaction. The reformation is mainly due to partial oxidation either with the active OH^-, O^-, O_2, O_3 created by the plasma dissociation. At the same time, the hydrogen-rich compound such as CH_4 or CH_3OH will be dissociated directly with electron

collisions. Since carbon dioxide is a major by-product together with some other minor gases coming out from the impurity of the fuel, they will be separated by the conventional absorption method or the membrane separation method.

Transformation of hydrocarbon fuel by corona and glow plasma has been attempted by passing the hydrocarbon gas such as methane, natural gas, LPG and vaporised liquid fuel sometime mixed with water vapours through the plasma reactor. They have all been successful in producing hydrogen-rich gas through corona discharge at atmospheric pressure by subjecting methane, vaporised methanol, diesel fuel mixed with water vapour, by passing it through a plasma gild arc reactor, wire in tube reactor and reactor proposed by MIT plasmatron or other gas phase corona streamer reactor.

The proposed under-liquid plasma reactor has many advantage over the gas-phase plasma reactor as it is able to generate a steady plasma-glow discharge at a very much lower voltage, i.e. from 350 V to (rarely) 1,800 V with current in the range of 100 mA to 800 mA in water. The liquid medium will also permit the application of ultrasonic waves producing an effect which will enhance the generation of glow plasma and thereby increase the overall transformation process. Again, no external air or gas is need be introduced for the reaction. However, the hydrocarbon gas such as methane, natural, LPG or hydrogen sulphurs gas can be introduced to work in conjunction, and complementing the liquid fuel in the reformation process. The fuel gases will enhance plasma-discharge reformation and allow it to take place without having to rely on gas produced by electrolysis.

Those hydrocarbon fuel molecules which come in contact with the plasma-discharge, will be subjected to dissociation and partial oxidation depicted in the following:

$H_2O + e \rightarrow +OH + H^+ + e$ dissociation
$CH_4 + e \rightarrow CH_3 + H^+ + e$ direct plasma dissociation
$CH_4 + H \rightarrow CH_3 + H_2$ reacting with H radicals
$CH_4 + H_2O \rightarrow CO + 3H_2$ partial oxidation
$CO + H_2O \rightarrow CO_2 + H_2$ water shifting
$CH_3OH + H_2O \rightarrow CO_2 + 3H_2$ electrolysis and partial oxidation
$H_2S \rightarrow S + 2H$ without experiencing oxidation
$H_2S + 2H_2O \rightarrow SO_2 + 3H_2$ partial oxidation
$SO_2 + 2H_2O \rightarrow H_2SO_4 + H_2$

Endothermic catalytic conversion of light hydro-carbon (methane to gasoline):

$CnHm + nH_2O \rightarrow nCO + (n + m/2)H_2$

With heavy hydro-carbon:

$CH1{,}4 + 0{,}3H_2O + 0{,}4O_2 \rightarrow 0{,}9CO + 0{,}1CO_2 + H_2$
$C_8H_{18} + H_2O + 9/2O_2 \rightarrow 6CO + 2CO_2 + 10H_2$

The hydrogen gas and carbon dioxide are collected. The CO_2 is separated by establish absorption or the membrane separation method.

The OH radical produced by the plasma dissociation will play an important role in oxidising the CH_4 to produce CO which would further be oxidised to become CO_2. The same applied to methanol CH_3OH and H_2S. The S is being oxidised to form SO_2 and further oxidising to become SO_3 and subsequently reacting with H_2O to produce H_2SO_4. This type of chemical reaction will be possible only with the encouragement of the highly chemical reactive and plasma catalytic environment. Not every CO will become CO_2 and sulphur particles may be observed in the precipitation.

REACTOR

There are number of reactors which can be used for the reformation of hydrogen-rich compounds. Reactors such as the wire in tube, tube in tube; single cell and multiple cell reactors; and the multi-electrodes without diaphragm separation. The tube in tube reactor and tower reactor with horizontal electrodes are suitable for treating both liquid and gas hydrocarbons and both at the same time. The anode and cathode are closely spaced with a gap distance ranging from 6 mm to 12 mm and are covered with dielectric gas-retaining and current-concentrating

construction on one side or both sides of the electrode. One important aspect of the reactor is having the construction, which will accommodate the ultrasonic transducer, which would induce proper sonic cavitations uniformly distributed throughout the reacting volume. The size, shape and arrangement of the electrodes can vary but its size would be restricted by the electric power available. A small reactor electrode plate is quite adequate for good uniform discharge and high productivity. The size of reactor plate use in most of the experiments is in the range of 16 cm^2 to 30 cm^2. It is preferable that the non-discharging electrode has an electrode area larger than the discharging electrode with the dielectric gas-retaining construction. With sufficient power available, both the anode and the cathode electrode can be functioning as plasma discharging electrodes at the same time. This is particularly useful in the partial oxidation process.

In the case of an emulsified oil/water mixture, it is best maintained with ultrasonic excitation which at the same time generates transient micro bubbles which enhance the whole reactive process. Hydrocarbon gas may also introduce to the reactor to form air bubbles or trapped gas pockets for the ready formation of the plasma glow discharge. Since the oily hydrocarbon fuel is highly dielectric this would require a higher concentration of conducting reagent than that required for the plasma-assisted water electrolysis, in order to maintain a suitable level of current density for the discharge to occur.

Reformation of methane gas by the under-liquid non-thermal plasma is by bubbling the gas through the perforated horizontal electrodes of tower a reactor or a tube-in-tube reactor. Since the methane gas is to be oxidised by the plasma dissociated water molecule (OH- + H+) to form carbon monoxide and hydrogen gas ($CH_4 + H_2O \rightarrow CO + 3H_2$). The CO will be further oxidised to form CO_2 with oxygen derived from the plasma dissociated water molecule, releasing two more hydrogen atoms (H_2). The resultant gas is either H_2 or CO_2 with perhaps small amount of CO. The hydrogen gas will be collected with reasonable purity after the CO_2 or CO is removed by absorption or membrane separation. Since the methane gas may not thoroughly reform with one past through the reactor, it is important to regulate the gas flow rate to ensure suitable resident time for the reformation or to have the methane gas recovered by the next round of reformation or to have the gas going through a series of reactors to made sure that the methane gas is fully utilised. The later case may not be energy efficient.

Reformation of methanol for hydrogen production can be achieved in the first place, by ordinary electrolysis or by partial oxidation. When CH_3OH is subjected to plasma discharge irradiation, it will react with the oxidising species and radicals dissociated from the water molecules. Conventional electrolysis will also contribute to the overall production of hydrogen gas. Reformation of methanol/water mixture will achieve better efficiency when plasma discharges is used in conjunction with ultrasonic excitation and cavitation. Several types of reactor can be adopted for the methanol reformation such as a tower reactor with horizontal electrodes, a tube-in-tube reactor, a transverse flow reactor, etc. These types of reactor offer very active oxidising species and hydroxyl radicals needed in the reformation.

Reformation of heavy oil such as diesel by under-liquid plasma discharge will be with emulsified liquid. The best way to maintain a thorough emulsification of diesel fuel and water is by ultrasonic excitation. Micro droplets of diesel will be encapsulated in the water. It is again observed that the conductivity of the emulsified liquid is very low as diesel oil is dielectric and current can only be conducted through the water film inbetween. This has rendered the need of more electrolytes added, especially as the diesel content increases. Bubbles are not easily produced by electrolysis due to its low current flow. It is therefore an advantage to either introduce gas to the reactor from outside or to produce ultrasonic cavitations in the liquid at the same time as the emulsification of the water/oil mixture. The tower reactor, tube-in-tube reactor and the transverse-flow reactor are all suitable for heavy hydrocarbon fuel reformation provided that an adequate ultrasonic transducer is properly located to ensure effective excitation and cavitations distributed throughout the liquid volume. A pulsed power supply will enhance the plasma generation and electrode heating will assist the generation of bubbles at the discharging electrode.

REDUCTION OF METAL AND MINERAL OXIDE PROCESS

Mineral refinment is an expensive and polluting process. To remove oxygen from the oxide, is either by reacting with higher electro-positive elements, which is uneconomic, or by exposing the metal oxide to C, CO, and hydrogen inside a high-temperature furnace such as the case in iron production. The electrolysis of a molten melt of Al_2O_3 or TiO_2 to extract pure metals Al or Ti respectively, consumes a large quantity of electricity, and requires the use of expensive refractory and electrode materials along with polluting emissions, render these two useful metals very expensive and inhibit their common application.

An under-liquid plasma reductive process to reduce oxide of ore or metals is proposed. The plasma discharge irradiation of the metal oxides in a highly catalytic environment, will cause interaction with the active hydrogen atoms produced by the plasma dissociation of water or methane or a methanol/water mix and introduced hydrogen gas together with the assistance of ultrasonic excitation would be sufficient in many instances to dislodge the most stubborn oxide.

It is reported that research is underway to extract Al from Al_2O_3 by electrolysis. Aluminium is electrode wired to cathode from porous Alumina anode electrode. The reduction of TiO_2 and Al_2O_3 by hydrogen plasma discharge is also being actively researched elsewhere with the aim of economically refining these two useful metals. A tube-in-tube reactor, or a wire-in-tube reactor can be used for this reduction process. These two reactors can be easily modified for continuous processing of either the granular form of the mineral or the metal oxide. The metal oxide will be exposed to the influence of highly active hydrogen atoms and subsequently the oxygen in the metal will be removed. This would not be a problem for those electro-positive elements but would present some difficulty for oxides such as Al and Ti.

The oxygen is strongly bonded with the parent metals such as Al_2O_3 and TiO_2 which cannot be reduced easily. This rudimentary horizontal reactor serves to demonstrate that metal oxide can be refined by exposing it in granular form to plasma discharge irradiation, ultrasonic excitation and in a highly reactive environment containing active hydrogen atoms. Additional hydrogen can be derived from the plasma dissociation of methane gas introduced to the reaction chamber where CO and atomic H are produced. Similarly by plasma dissociation of the methane water mixture that active hydrogen and CO_2 are also produced to supplement the reductive atomic hydrogen. Hydrogen gas can also bubble into the reactor and any excess will be collected and passed back to the reactor.

Reduction of Al_2O_3, TiO_2, TiF_3, TiO, $AlCl_3$ will be taking place in the following manner, where:

$TiO_2 + 4H(2H_2) \rightarrow Ti + 2H_2O$
$Al_2O3 + 6H(3H_2) \rightarrow Al + 3H_2O$
$TiF_3 + 3H(3/2H_2) \rightarrow Ti + 3HF$

The alternative is to have:

$TiO_2 + H_2SO_4 \rightarrow TiOSO4 + H_2O$
$TiOSO_4 + 2H \rightarrow TiO + H_2SO_4$
or $TiO + 2H \rightarrow Ti + H_2O$
and
$TiO_2 + 4HCL \rightarrow TiCl4 + 2H2O$
$TiCl_4 + 4H \rightarrow Ti + 4HCl$
where $TiCl_4$ is ionic and is soluble in water

The above reaction is under the influence of a non-thermal plasma so that the oxide of ores or metal is subjected to a highly catalytic environment and comes into contact with the reactive atomic hydrogen whereby the oxygen will be taken out. To enhance the matter further, the whole reaction process is also subjected to sonic excitation. The fine particles in the colloidal suspension of the granular oxide will collide with each other and at the point of impact, the temperature will rise over 1,500^{O}C to 3,000^{O}C and local melting is reported. The high temperature and pressure of a collapsing sonic bubble will work in conjunction with the plasma glow discharge irradiating the oxide particles with atomic hydrogen with localised high temperature due to collision and cavitations implosion which in the end remove the oxygen. The refined metals will be in powder form down to nano size.

The other method of extracting and refining metals from their oxides is to subject the ionic solution of the metal such as $AlCl_3$ to an electrolysis process which is reported to have achieved efficiency of 3 KWh/Kg of Al. The whole process can be further improved with the plasma electroplating technique with the proposed under-liquid glow plasma discharge. The Al will be deposited on the cathode electrode. Part of the chlorine gas will come out from the anode side and will react with the active hydrogen to form Hcl.

The fine granular metal oxide is placed inside a horizontal reactor on top of cathode electrode. A close matrix separator membrane, used to prevent the metal oxide from crossing over, placed above and below the anode electrode is used to separate it from the cathode. The whole reactor is submerged inside an ultrasonic bath. Ultrasonic waves will penetrate the membrane separator to cause the granular metal oxide in colloidal suspension. The oxide will be subjected to the under-liquid plasma glow discharge irradiation and atomic hydrogen reduction. The percentage of metal oxide being reduced after a period of time is evaluated. Metal oxide of TiO_2 will be put to test. A methane/water mixture will be employed as the liquid medium which will produce larger amount of active atomic hydrogen serving as reduction agents.

DECONTAMINATION OF LIQUID

The problem of pollution is a major issue affecting every living being on this planet. A lot of effort has been expended by Governments, universities and private enterprises, seeking a comprehensive process to deal with a vast variety of pollution issues. Polluting gas emissions from industries and motor vehicles produce large quantities of CO_2 causing global warming; NOx, VOC, and particulates causes cancer and smog; SO_2 causes acid rain. Decontamination of the gases discharged from industries is costly to achieve and what is urgently needed is a comprehensive and economical treatment process to reduce the overall trestment cost. Water contamination is another major issue. Contaminated water unfit for human consumption, enters the sea and kills marine life near the shore. Governments worldwide are passing stringent laws setting a pollution standard, which demands the development of efficient and economic ways to control pollutants. The present proposed invention is put forward as a versatile process, which can treat a variety of contaminants either separately or together.

Corona discharge and glow plasma discharge as non-equilibrium plasma has been developed for applications in the decontamination of a wide range of noxious chemical compounds and recalcitrant chlorinated organic compounds such as dichloro-ethane, pentachlorophenol, perchloroethylene, chlorom, carbon tectrachloride, organochlorine presiticides, endocrine disrupter, dioxin etc. It is also capable of sterilising tough microbial, bacteria and biological contaminants present in ground water such as cryptosporidia parvum. Noxious gas emissions such as NOx and SOx can also be neutralised by passing them through the wet reactor, which includes the removal of particulates as well as the pollution emissions. This is mainly due to the ability of plasma to create a very reactive catalytic environment for those normally very stable and inactive compounds to be reduced, oxidised or neutralised by reacting with the OH* radicals, atomic hydrogen H+ and other oxidative species such as O^-, O_2, O_3, H_2O_2 etc. present and is reported to have high efficiency especially in dealing with diluted contaminants.

Microbial bacteria is removed by both oxidations when they come in contact with the oxidative species such as O_3, O_2^-, O^-, H_2O_2, and OH*. At the same time, they are subjected to the electromechanical stretching of the cell wall, which weakens its oxidative resistance, especially when ultrasonic cavitations, implosions and shock waves created by pulse power, are incorporated into the reactive process. Again reports of over 99% sterilisation are not uncommon.

At the present, most of the treatment work is conducted in a gaseous environment, by spraying or vaporising the contaminated liquid over the plasma discharging electrodes, or by producing plasma discharge irradiating over the surface of a liquid which contains the undesirable contaminants, or by passing the polluted gas through a dry reactor sometimes mixed with water vapour or using plasma torch irradiation of the polluted object.

A surface water contact plasma glow discharge system has also been developed as a decontamination process under the name “Plasmate”. Under water plasma by pulsed high voltage electric discharge with high current input to dissociate the water to produce H and OH* radicals to treat bacterial and microbial decontamination has also been reported as being successful.

The proposed under-liquid plasma is a low energy consumption system, which produces steady plasma by utilising the present of bubbles. The voltage required for dealing with a wide range of liquids having variable electrolytic properties, ranges from 350 V to 3,000 V and current intensity ranging from 1 to 2 Amp/cm^2. It produces a highly reactive environment with a supply of oxidative radicals and reductive atomic hydrogen spread over a large volume of liquid, making it highly effective as a decontaminatinf process, and one which is also both economic and easy to operate.

The under-liquid plasma has the advantage of being able to decontaminate several pollutants at the same time and it also has a very active gas and liquid interaction which makes it highly effective as a treatment process. Liquid waste, containing harmful chemical, bacteria, microbial, heavy metals, noxious gas, polluted air and odour can be treated in the same reactor simultaneously.

Recalcitrant organic chlorinated materials in water, which include dichloromethane, pentachlorophenol, chloroform and carbon tetrachloride, will either be oxidised or degraded to CO_2 and chlorine. While the pathogens in drinking water such as cryptosporidia with thick phospholipids wall protecting the trophs is in the first place being stretched and weakened and subsequently broken down by the oxidising species. Some of the oxidative species such as OH radicals, O^-, O_2^-, and O_3 are present in quantity and are more active than chlorine and other mild oxidants. It has the advantage that no chemical is needed as an oxidation agent, which can sometimes result in secondary pollution.

Heavy metals in dilute solution, can be extracted or removed through a simple electrolysis process by turning the metal to hydroxide which could than be removed by filter. Soluble metal ions can also be extracted by deposition on to the cathode electrode, which can be further facilitated by the plasma electroplating process owned by the inventor, and which uses the same under-liquid bubble plasma process.

The treatment of NO, SO_2 and particulates is to pass the polluted gas through the reactor where the particulate will be removed and the NO is either oxidised to become NO_2 or NO_3 by O^-, or O_3. It can also be reduced to N by the active hydrogen. NO_3 will react with water to become nitric acid. NO_2 is not considered to be a noxious gas. SO_2 reacting with O_3 or oxygen radical to form SO_3 can be easily oxidised and then react with water to become H_2SO_4 (sulpheric acid). When the said gas is introduced to the reactor it can be utilised as a gas bubble for plasma discharge especially when this gas bubble is collected or retained near the electrodes.

The effectiveness of non-thermal plasma discharge in treating carcinogenic organic compounds and pollutant gases is well established. Removal or reduction of the amount of heavy metals, arsenic and mercury to an acceptable safe low concentration level from or in water, have been successfully carried out by a simple electrolysis process. The extraction efficiency is further improved by the presence of an under-liquid plasma discharge where some of them will readily react with the OH radicals to become metal hydroxide or to be deposited by the very active plasma electroplating (deposition) method which has been adequately proven as a useful technique.

Further experiments in this area are unnecessary. Adequate information can be drawn upon from much research work which already been carried out. Concentrated effort has already been used to search for a better way of generating steady plasma glow discharge under-liquid by utilising the bubbles which will enable the manufacturing of a simple and economic reactor which requires only low power input and wich will work well in treating a wide scope of contaminants.

Sterilisation of drinking water at municipal scale can be simplified by adopting the under-liquid plasma discharge which will effectively neutralise and degrade carcinogen organic compounds in the water by creating the dissociation and active catalytic environment which encourages the breakdown of the inert chemicals and at the same time subject it to the active reductive and oxidative radicals. The heavy metals dissolved in the water will also be removed or reduced in the same time through the plasma electrolysis and electroplating as described previously. The biological contaminants will be sterilised by the highly oxidative environment existing during the glow discharge. The effectiveness of the combined treatment to produce potable water fit for human consumption is further enhanced by the adoption of ultrasonic cavitation and shock waves with a pulsed power supply.

The entire sterilisation process does not require any added chemicals such as ozone, chlorine or any electrolytic additive. The impurity in the pre-treated liquid will be adequate to serve as conductor for the under-water plasma discharge to take place. Any excessive ozone, which has not been used up in the oxidation process during the plasma discharge, will be easily neutralised by the presence of active hydrogen atoms. Hydroxyl radicals (OH) are one of the most aggressive oxidising agents, which being produced in quantity will do most of the useful work. There will be no chlorine remnant left in the water, as it is unnecessary.

The under-liquid plasma technique will be useful in food industries for low temperature sterilisation and removal of odour. The same method may also find its use in the paper-making industry in fragmentation and de-lignification of the fluidised pulps, treating the highly polluted discharge, and treating fabrics and dyes in the textiles industry.

There are several types of reactors which can be employed in the decontamination process. The separation membrane diaphragm in the wire-in-tube and tube-in-tube reactor is no longer required. Other reactors such as the transverse-flow reactor and the tower reactor can also be adopted.

The reactor can be arrange in such way that the plasma discharge occurs either at the cathode or at the anode provided that a good gas-trapping cover is provided on the electrode. Since much of the decontamination action relies on the presence of strong oxidation agents such as hydroxyl radicals, atomic oxygen, ozone, singlet oxygen and hydroperoxyl radicals, plasma discharge on the side of anode electrode enhanced with the gas retaining cover will cause the formation of said species represented by the following equations:

$H_2O + e \rightarrow +OH + H + e$	dissociation
$H_2O + e \rightarrow + H_2O_+ + 2e$	ionisation
$H_2O_+ + H_2O \rightarrow H_3O_+ + OH$	dissociation
$O_2 + e \rightarrow O_2^* + e$	excitation
$O_2 + e \rightarrow +2O + e$	dissociation
$O_2 + e \rightarrow O^- + O$	dissociation
$O_2 + O \rightarrow O_3$	association
$OH + OH \rightarrow H_2 O_2$	association

Some chemical contaminants can only be broken down by reduction with active atomic hydrogen, which would require plasma discharge at the cathode electrode. In the tower reactor (**Fig.7**) and transverse-flow reactor (**Fig.6**) it is possible to have the gas-retaining cover on one side of electrode facing the side of the opposite electrode with the gas-retaining covers, so that an alternating zone of oxidation and reduction is created in the reactors to deal with a variety of contaminants.

Production of hydrogen by plasma dissociation of water molecules is the result of electron collisions, which is different from the conventional electrolysis, which separates the dipole water molecules by electro-induction. They also have different sets of requirements to dissociate water molecules for the production of hydrogen:

Conventional electrolysis	Plasma glow discharge under water, according to the present invention
1. Low voltage and high current density	High voltage and relatively low current density
2. High concentration of electrolyte (up to 25% KOH)	Low concentration electrolyte (0.01% KOH) low electrolytic requirement
3. Avoid bubble attachment to the electrodes	Bubbles smothering the electrodes is welcome to create a dielectric barrier.
4. Electrode space distance is not restricted.	Electrode space distance has to keep close as far as possible.
5. Water molecules is split by induction	Water molecules are dissociated by electron collision.
6. Large production unit is required for efficiency and productivity	Small production unit favours the decentralisation of production.

The reactors and gas-trapping and retaining structures enclosing the electrode is made of perspex plastic. No sign of burning is observed in the plastic covering plate directly over the discharging electrode and the light emission is an orange/red colour (burning of hydrogen) which is distinctively different from the plasma arc which is bright blue colour when the voltage is brought beyond the glow discharge voltage level. A burn mark will be observed after plasma arc discharge. This proves that the plasma glow discharge with it's orange yellow colour, is non-thermal in nature.

Applicant also conducted experiments with the same equipment utilising the under-liquid plasma to sterilise mulberry juice. Applicant found that the plasma was effective in reducing the bacterial count and the mold colony count in the juice. After 40 minutes the counts of both bacteria and mold had been reduced substantially to less than 100 per ml. This demonstrates that the invention could be used to sterilise potable water, waste water, food, and liquid food and others.

CONCLUSION

A further advantage of the method described above is that plasma can be generated with relative ease within bubbles in the aqueous medium. It does not require excessive amounts of energy and can be done at atmospheric pressure. It certainly does not require a vacuum chamber.

A further advantage of the invention is that it provides a method of treating aqueous waste which contains components that cannot be neutralised or otherwise rendered harmless by the addition of chemicals to the liquid.

It will of course be realised that the above has been given only by way of illustrative example of the invention and that all such modifications and variations thereto as would be apparent to persons skilled in the art are deemed to fall within the broad scope and ambit of the invention as herein set forth.

Figures which are included in the patent application but which are not directly referenced in it:

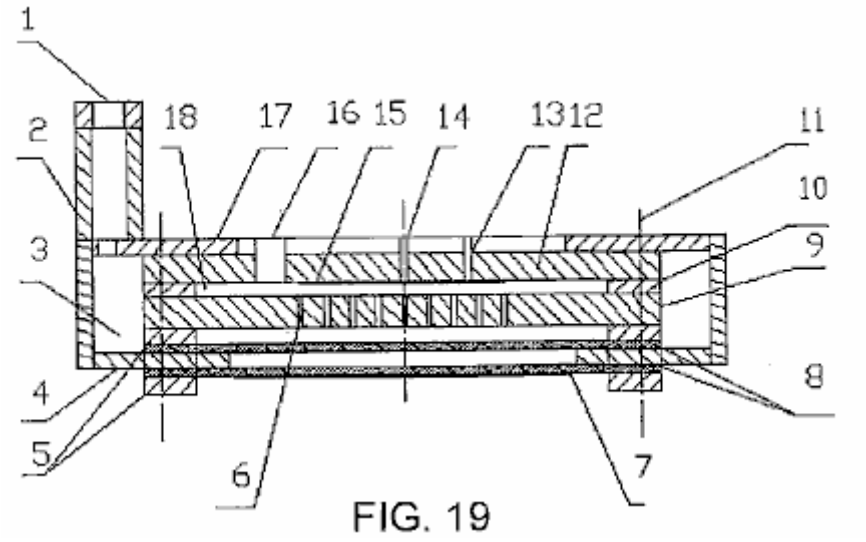


FIG. 19

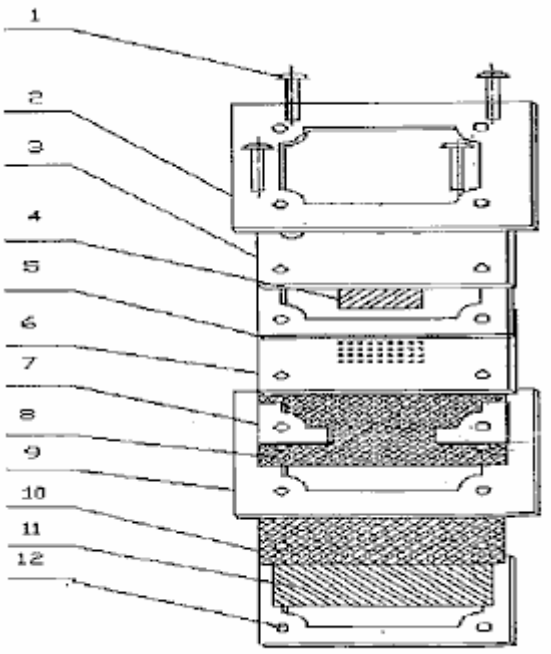


FIG. 20

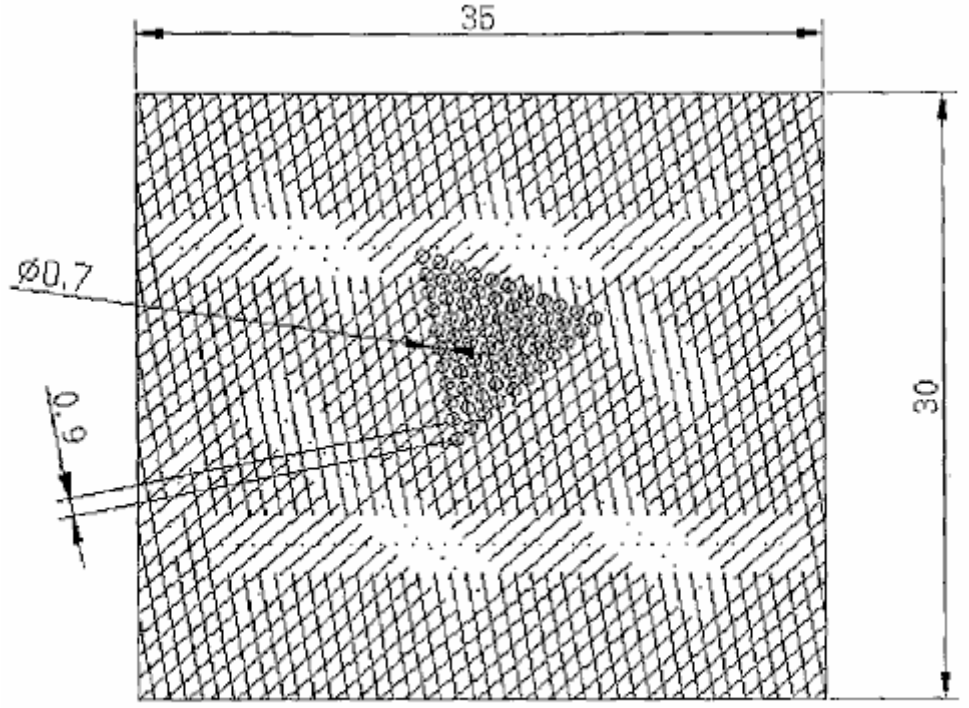


FIG. 21

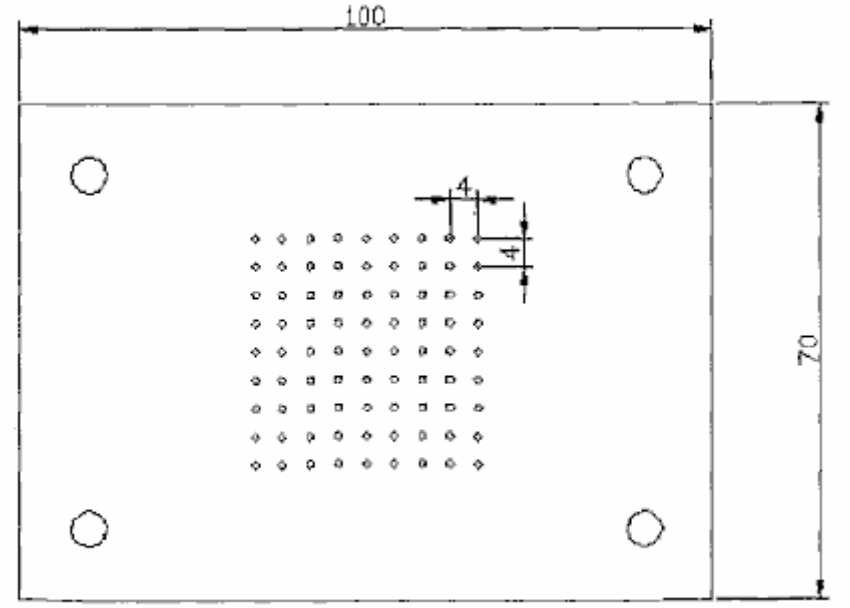


FIG. 22

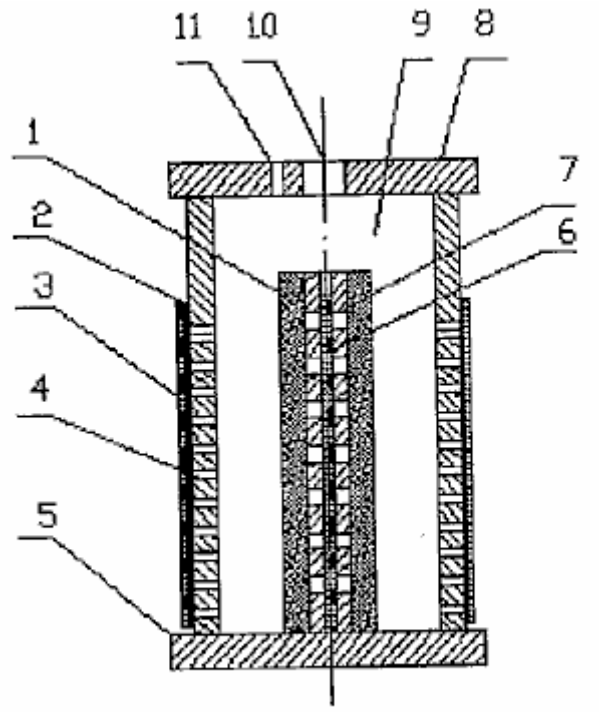


FIG. 23

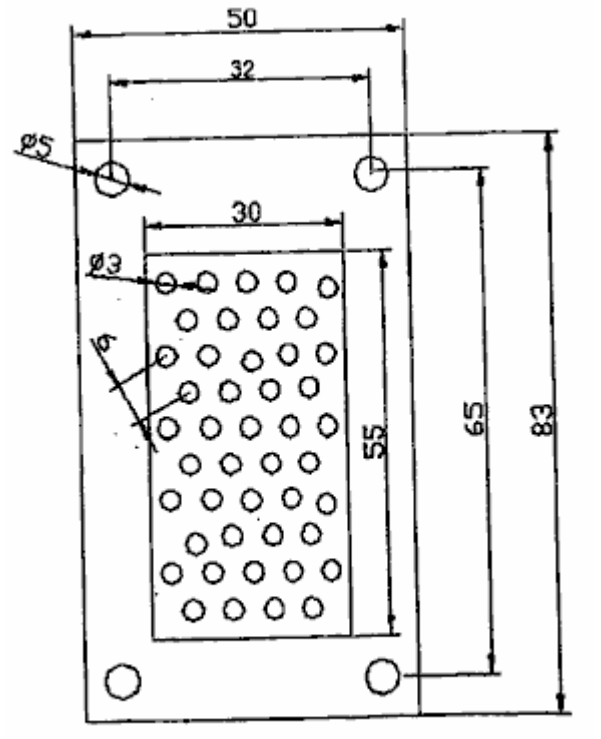
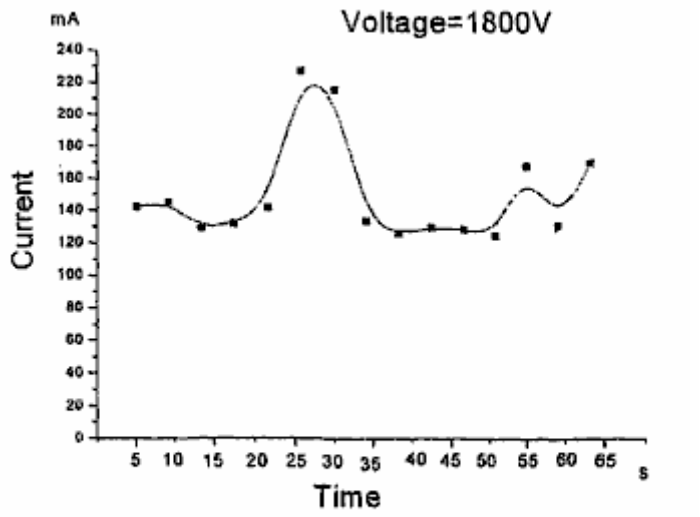
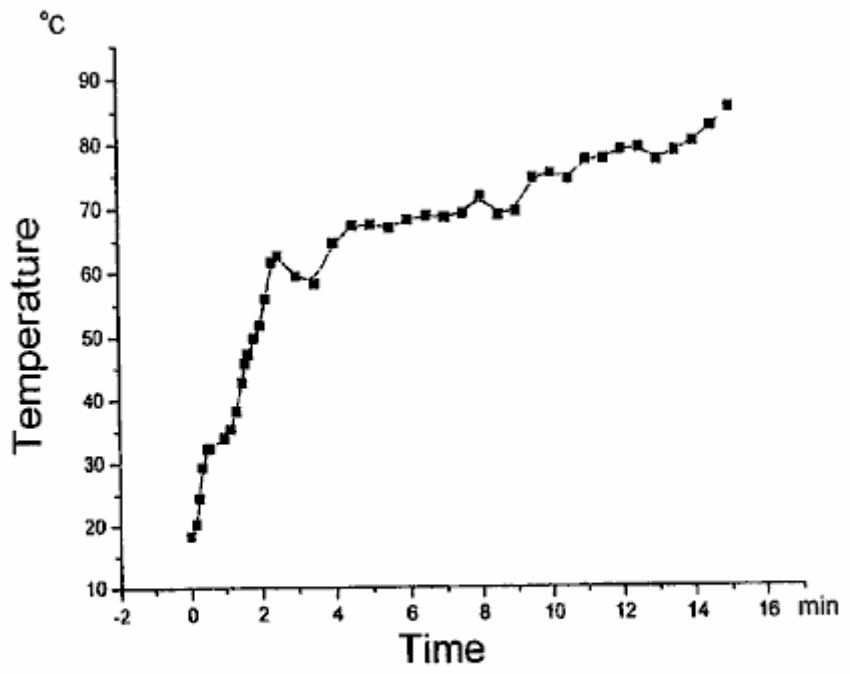


FIG. 24



Glow discharge at stable 1800V voltage input

FIG. 26



Temperature measured in the cathode electrodes

FIG. 27

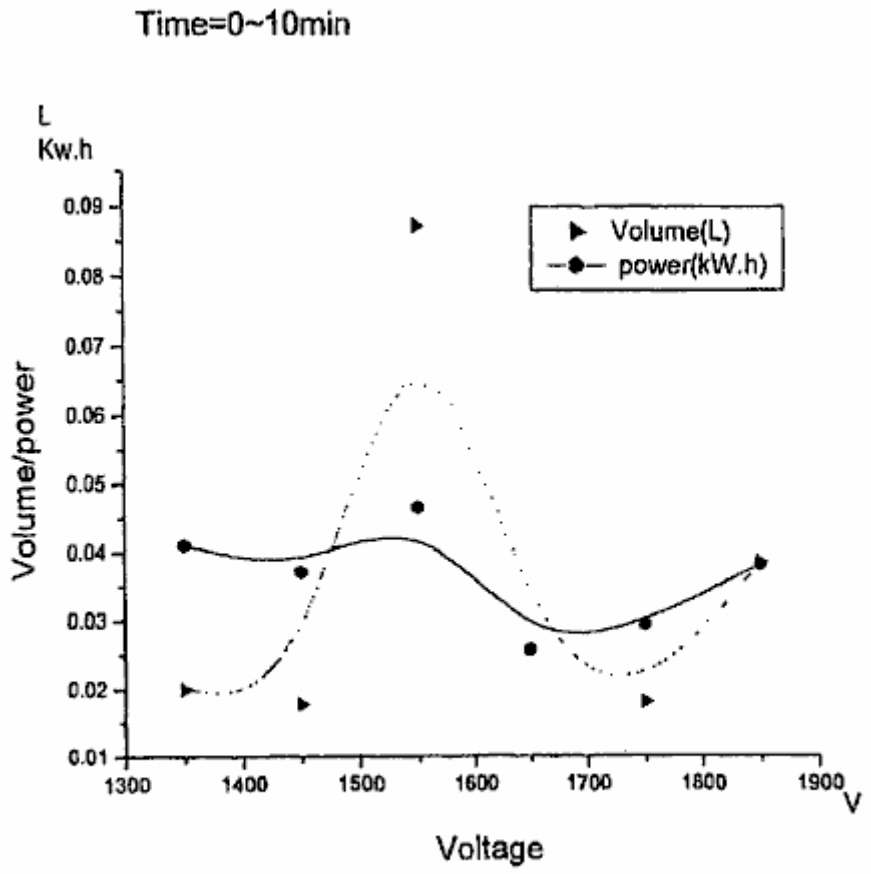

Time=0~10min
L
Kw.h
Volume(L)
power(kW.h)
Volume/power
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
1300
1400
1500
1600
1700
1800
1900
V
Voltage

FIG. 29

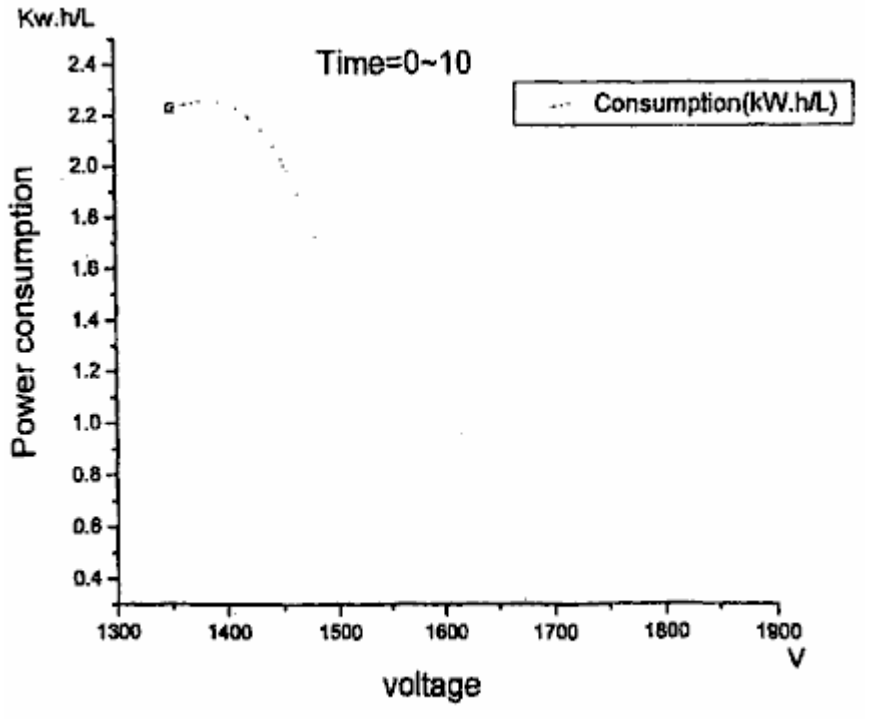

Kw.h/L
Time=0~10
Consumption(kW.h/L)
Power consumption
2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
1300
1400
1500
1600
1700
1800
1900
V
voltage

FIG. 30

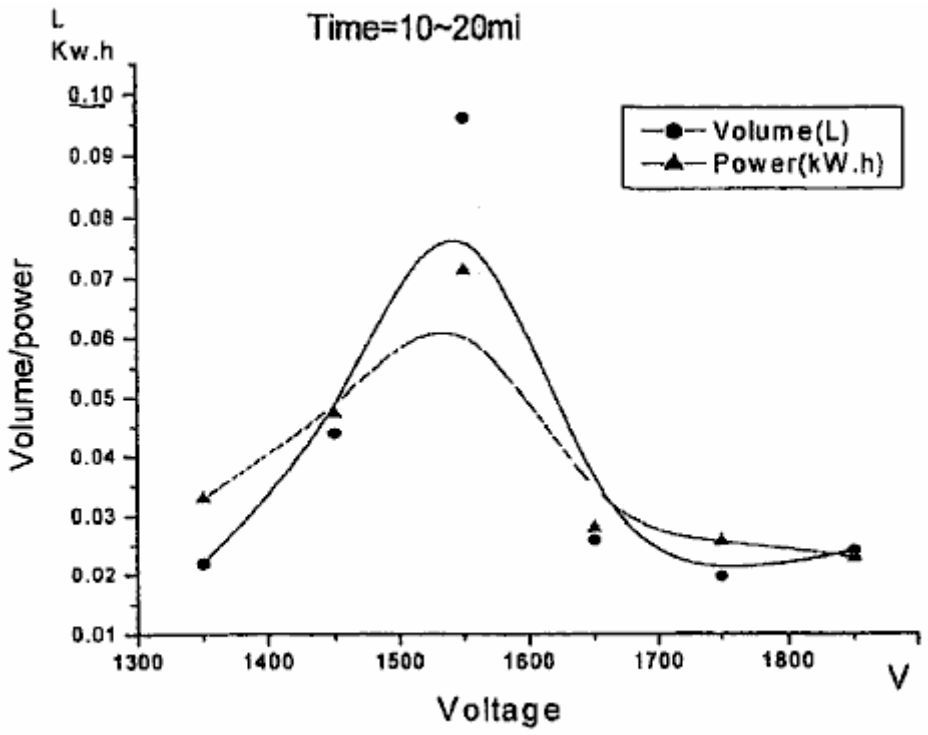

L
Kw.h
Time=10~20mi
Volume(L)
Power(kW.h)
Volume/power
0.10
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
1300
1400
1500
1600
1700
1800
V
Voltage

FIG. 31

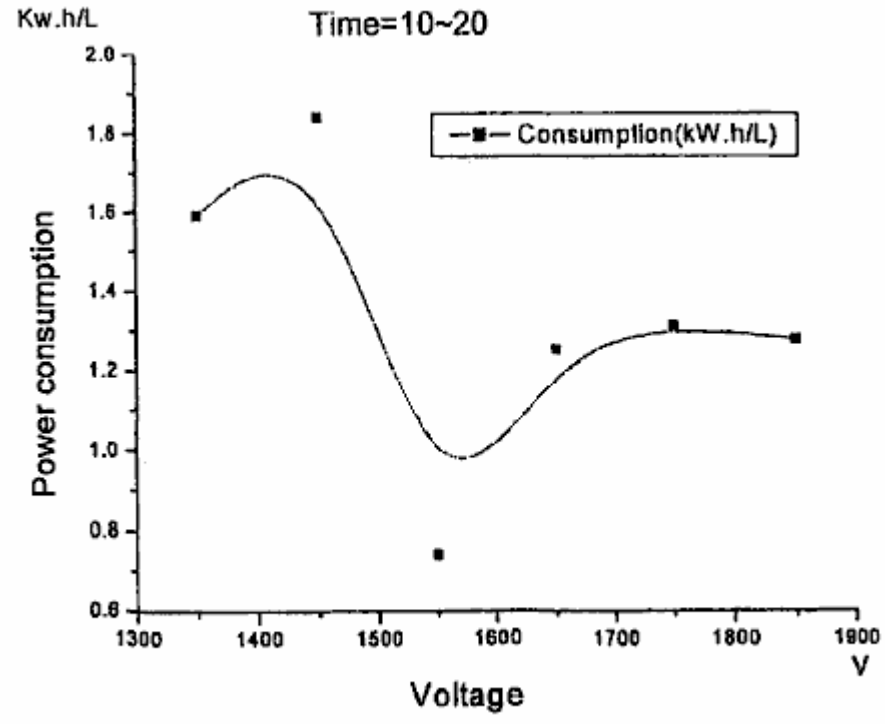

Kw.h/L
Time=10~20
Consumption(kW.h/L)
Power consumption
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
1300
1400
1500
1600
1700
1800
1900
V
Voltage

FIG. 32

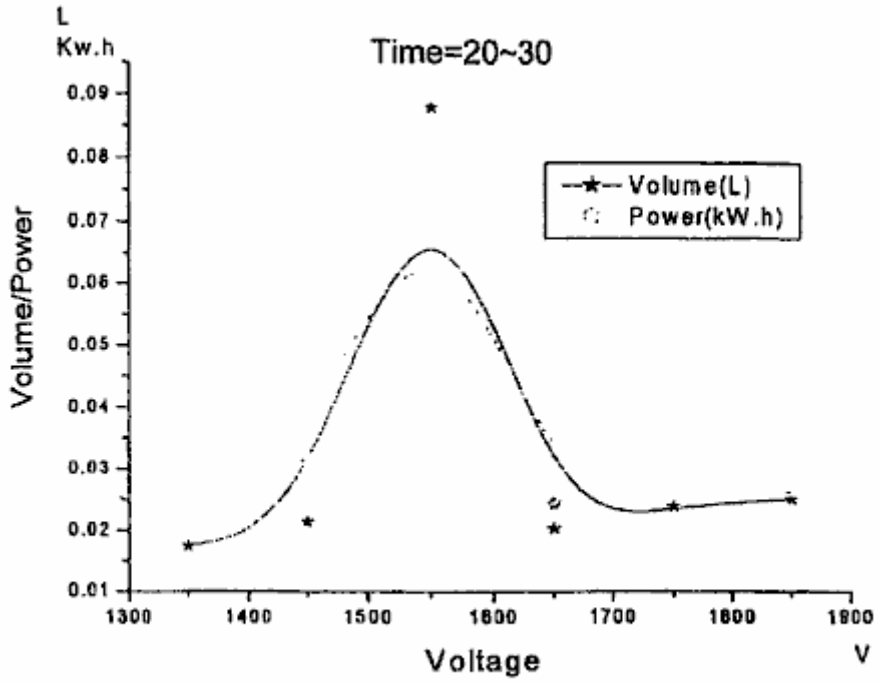

L
Kw.h
Time=20~30
Volume(L)
Power(kW.h)
Volume/Power
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
1300
1400
1500
1600
1700
1800
1900
Voltage
V

FIG. 33

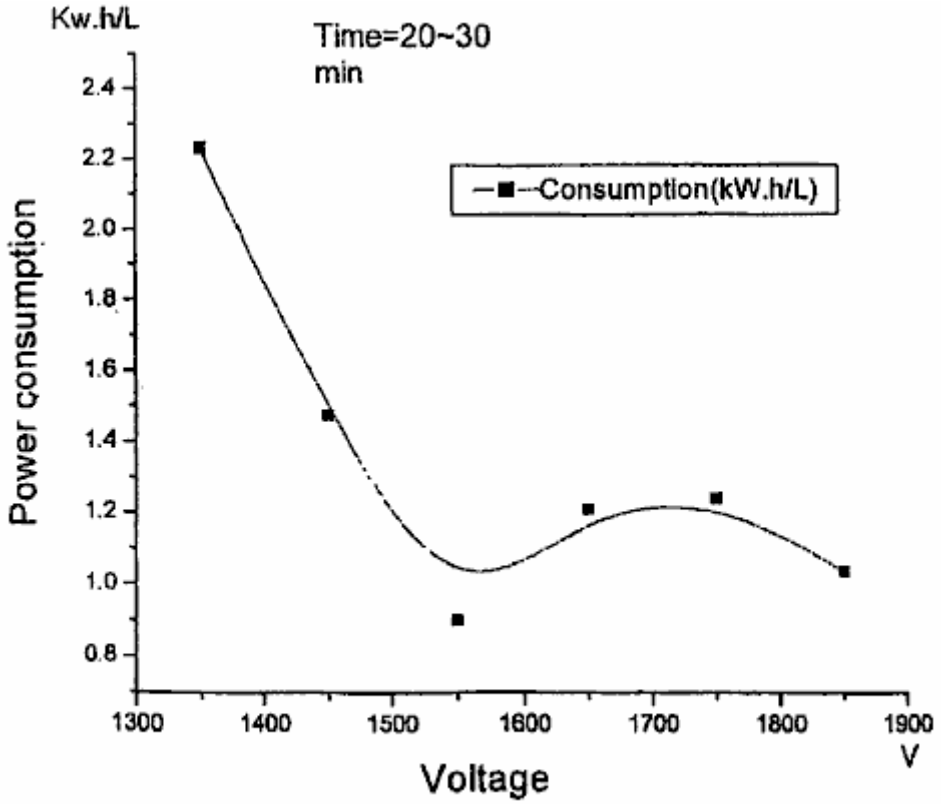

Kw.h/L
Time=20~30
min
Consumption(kW.h/L)
Power consumption
2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
0.8
1300
1400
1500
1600
1700
1800
1900
V
Voltage

FIG. 34

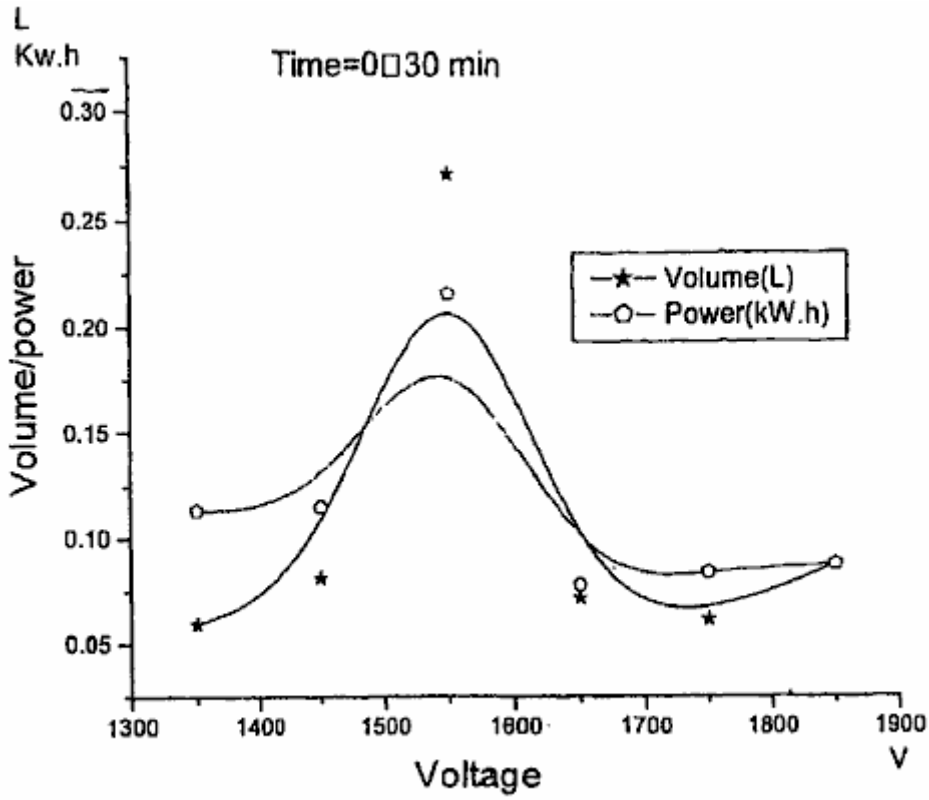

L
Kw.h
Time=0□30 min
0.30
0.25
0.20
0.15
0.10
0.05
Volume/power
Volume(L)
Power(kW.h)
1300
1400
1500
1600
1700
1800
1900
V
Voltage

FIG. 35

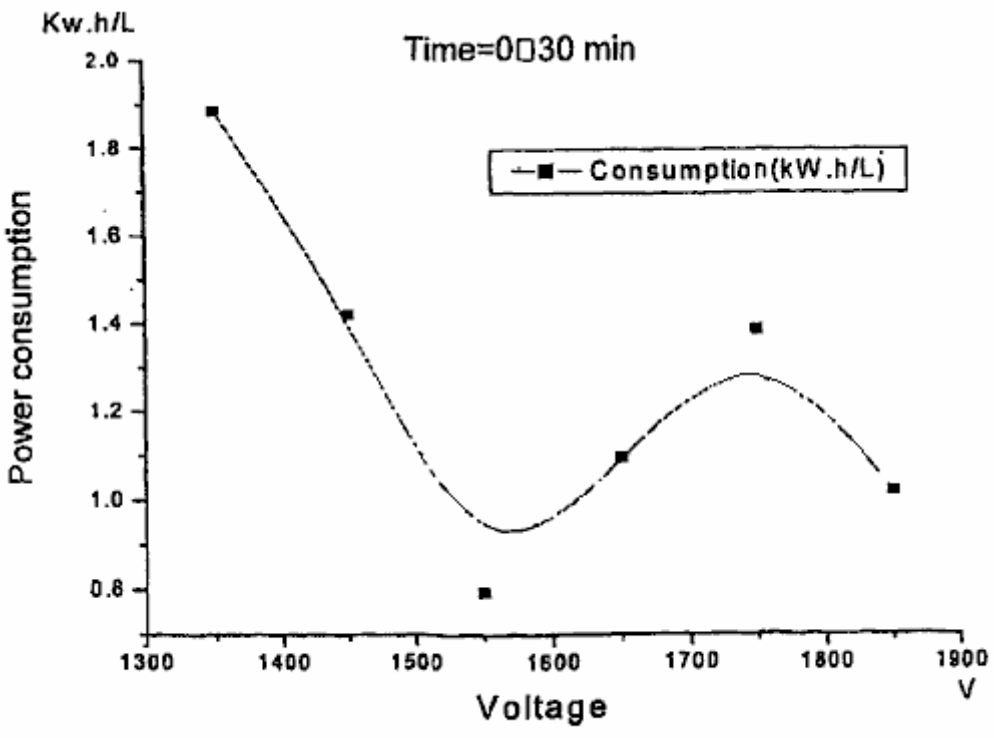

Kw.h/L
Time=0□30 min
2.0
1.8
1.6
1.4
1.2
1.0
0.8
Power consumption
Consumption(kW.h/L)
1300
1400
1500
1600
1700
1800
1900
V
Voltage

FIG. 36

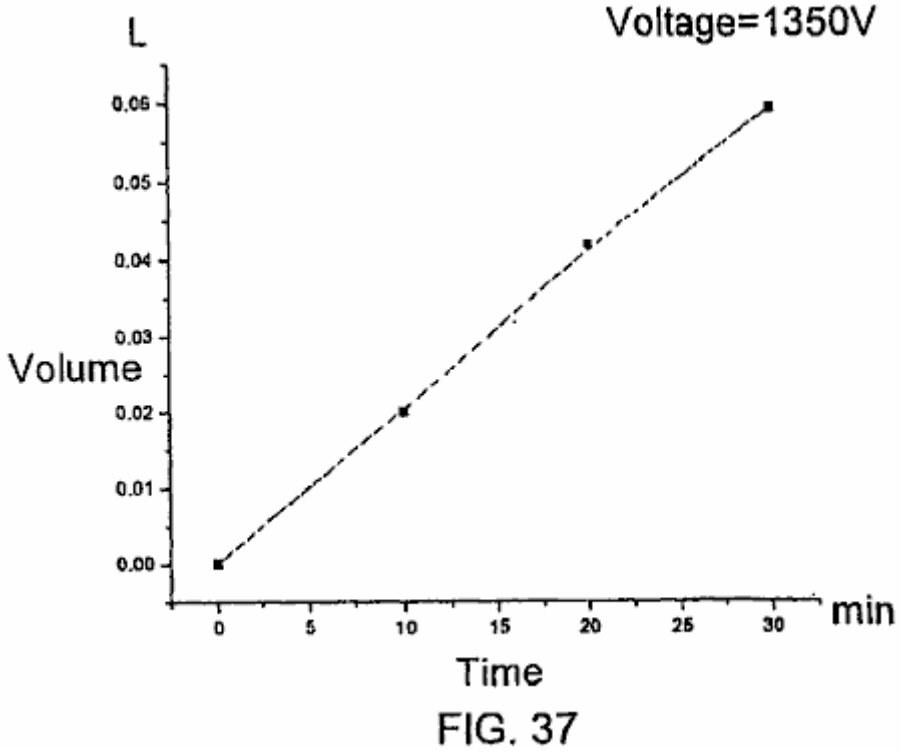

Voltage=1350V
L
0.06
0.05
0.04
0.03
Volume
0.02
0.01
0.00
0
5
10
15
20
25
30
min
Time

FIG. 37

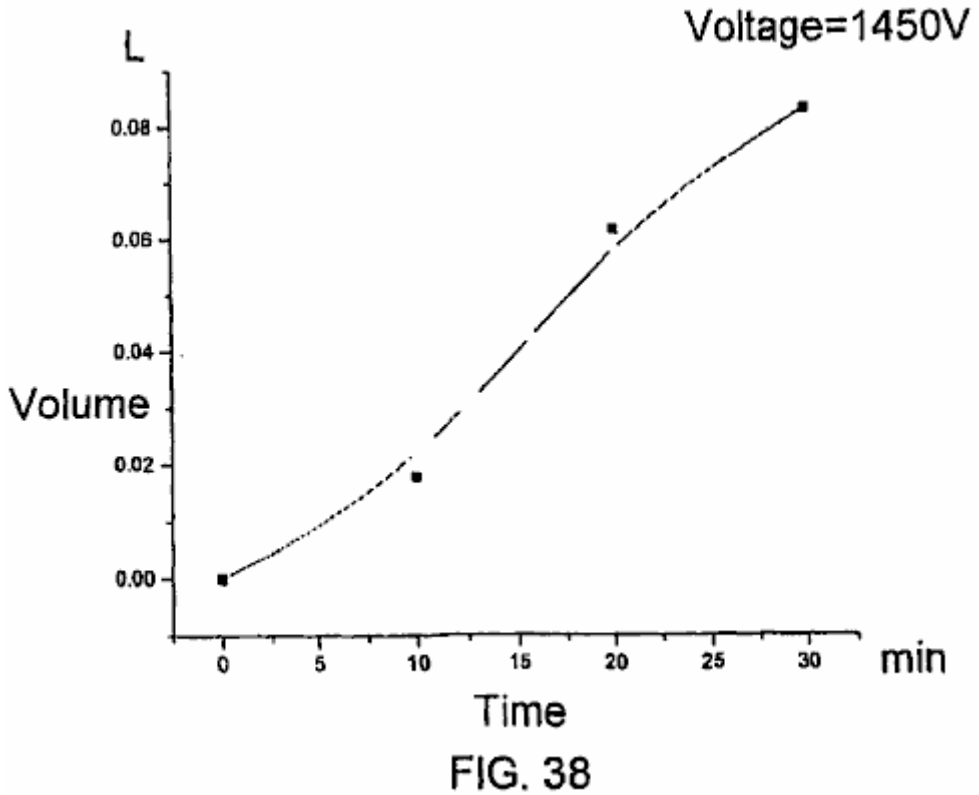

Voltage=1450V
L
0.08
0.06
0.04
Volume
0.02
0.00
0
5
10
15
20
25
30
min
Time

FIG. 38

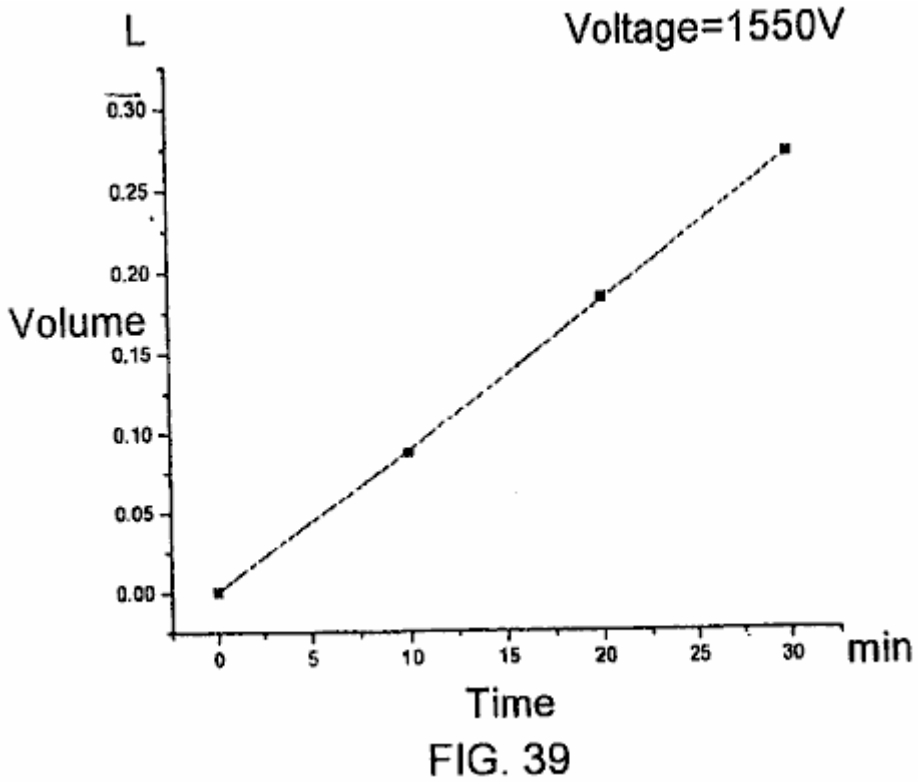

L
Voltage=1550V
0.30
0.25
0.20
Volume
0.15
0.10
0.05
0.00
0
5
10
15
20
25
30
min
Time

FIG. 39

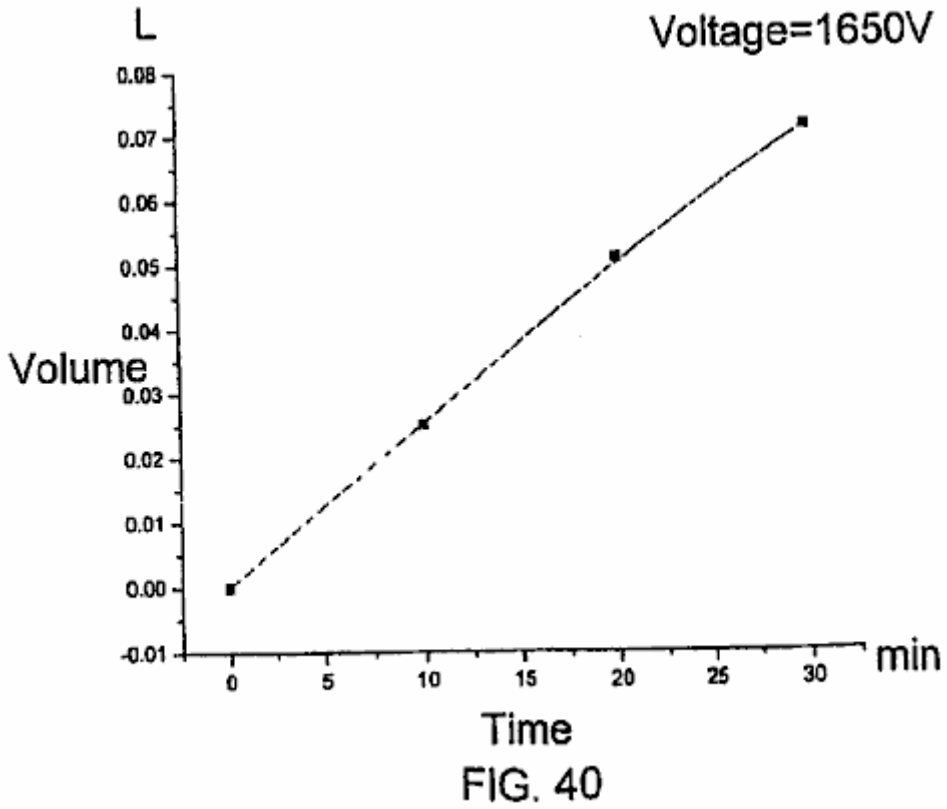

L
Voltage=1650V
0.08
0.07
0.06
0.05
0.04
Volume
0.03
0.02
0.01
0.00
-0.01
0
5
10
15
20
25
30
min
Time

FIG. 40

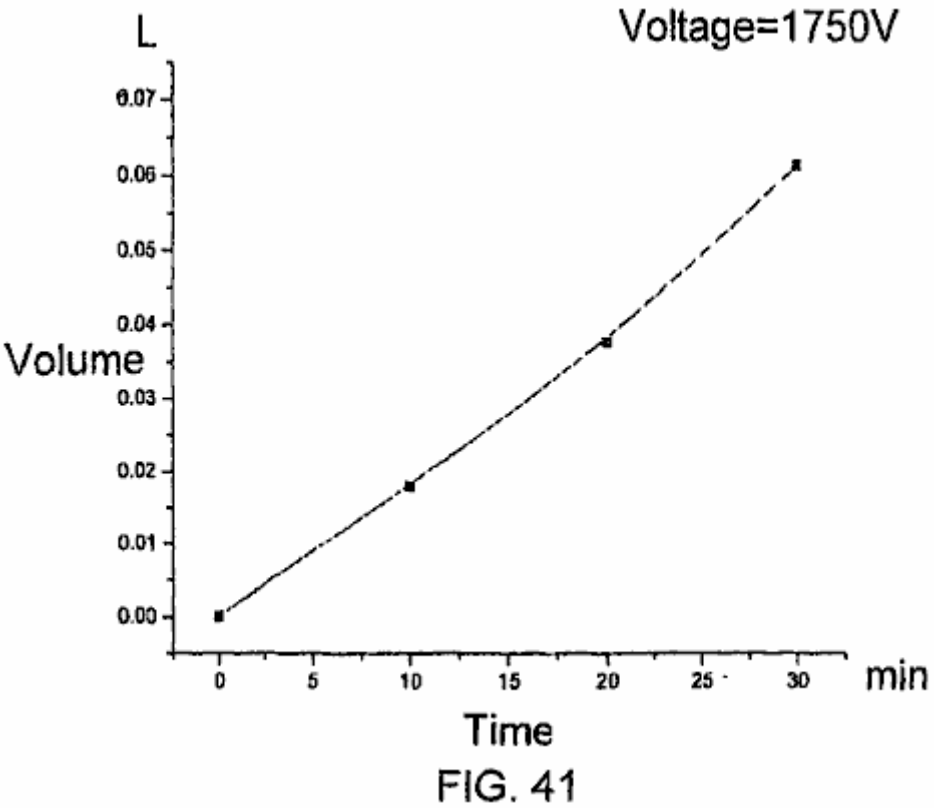

L
Voltage=1750V
Volume
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00
0
5
10
15
20
25
30
min
Time

FIG. 41

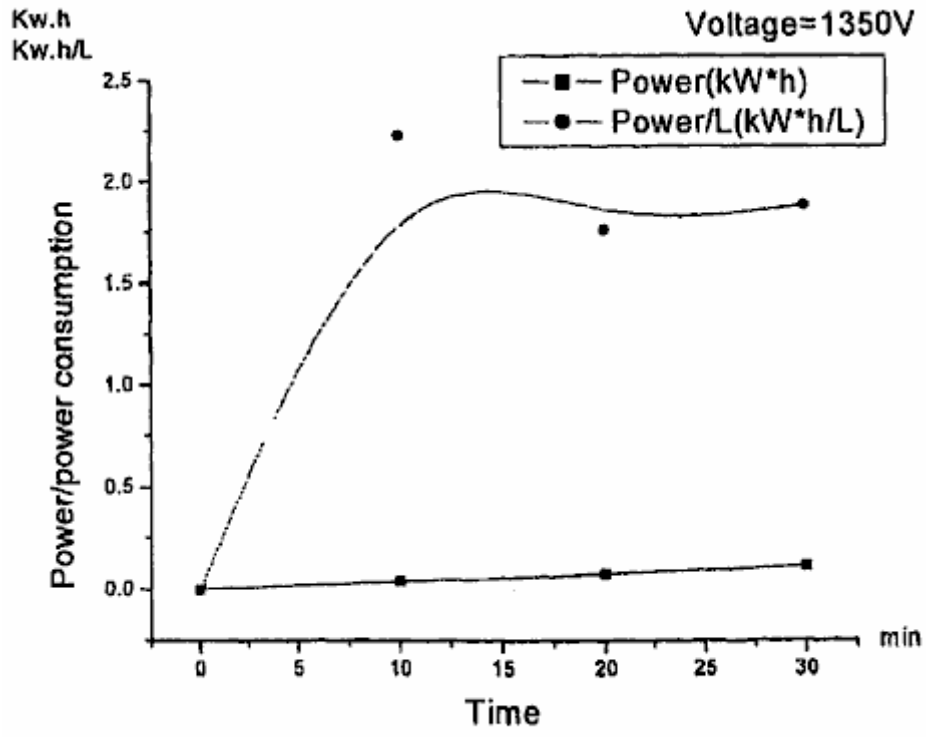

Kw.h
Kw.h/L
Voltage=1350V
Power(kW*h)
Power/L(kW*h/L)
Power/power consumption
2.5
2.0
1.5
1.0
0.5
0.0
0
5
10
15
20
25
30
min
Time

FIG. 42

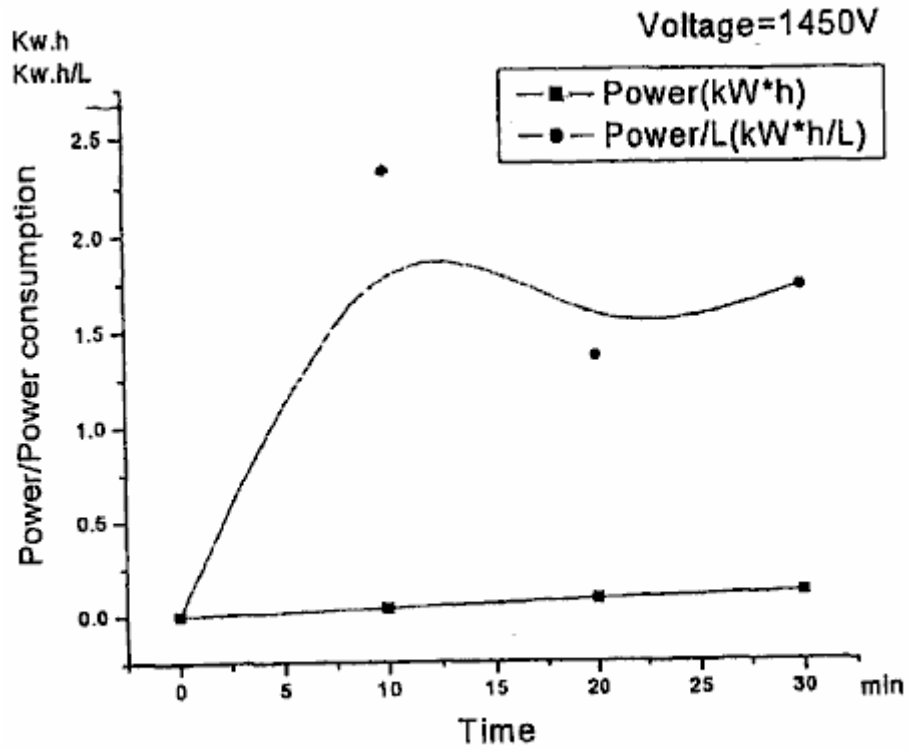

Voltage=1450V
Kw.h
Kw.h/L
Power(kW*h)
Power/L(kW*h/L)
Power/Power consumption
2.5
2.0
1.5
1.0
0.5
0.0
0
5
10
15
20
25
30
min
Time

FIG. 43

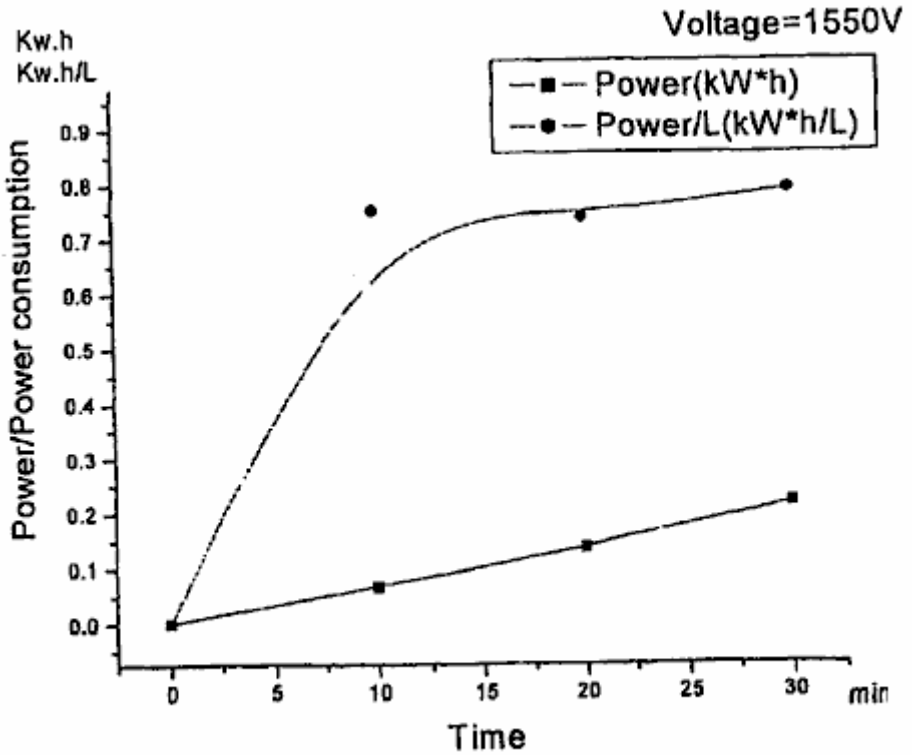

Voltage=1550V
Kw.h
Kw.h/L
Power(kW*h)
Power/L(kW*h/L)
Power/Power consumption
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
5
10
15
20
25
30
min
Time

FIG. 44

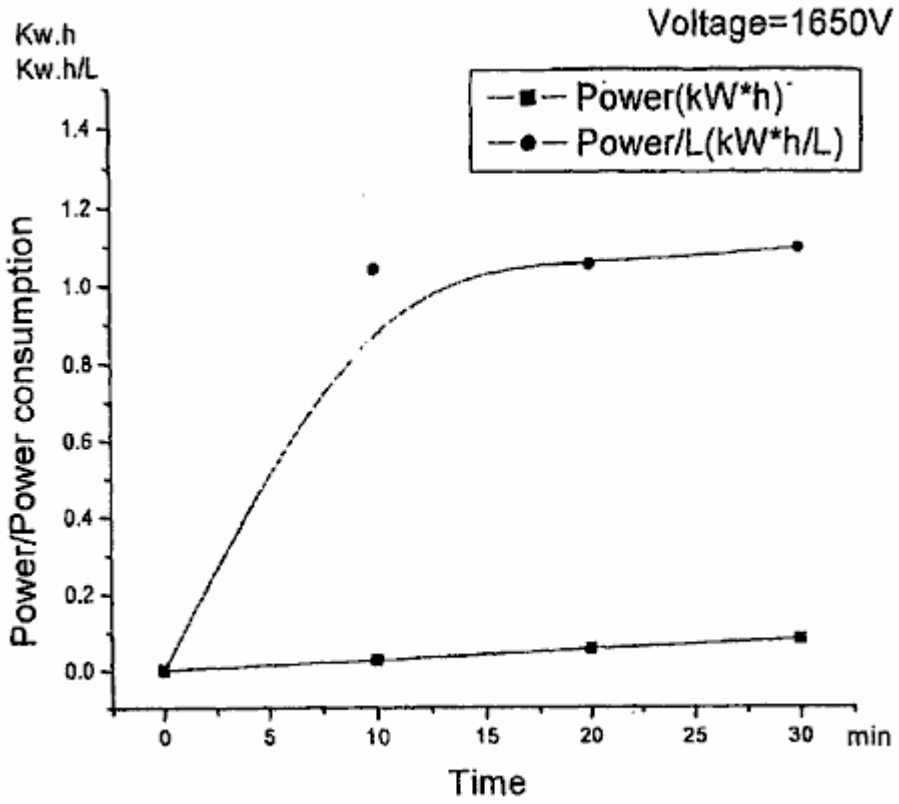

Voltage=1650V
Kw.h
Kw.h/L
Power(kW*h)
Power/L(kW*h/L)
Power/Power consumption
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
30
min
Time

FIG. 45

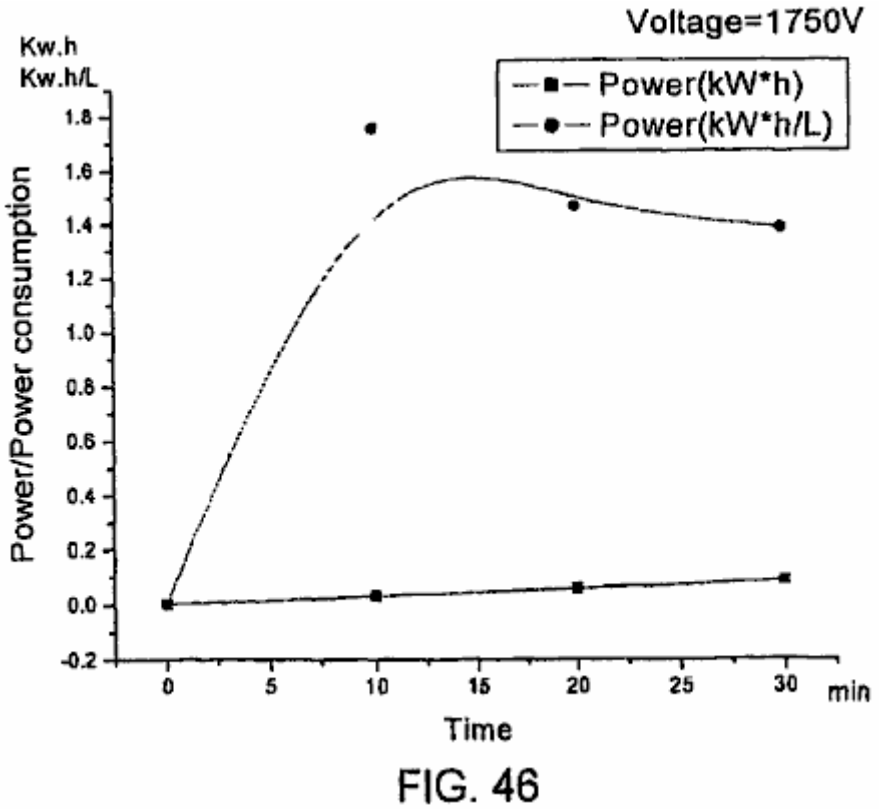

Voltage=1750V
Kw.h
Kw.h/L
Power(kW*h)
Power(kW*h/L)
Power/Power consumption
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
0
5
10
15
20
25
30
min
Time

FIG. 46

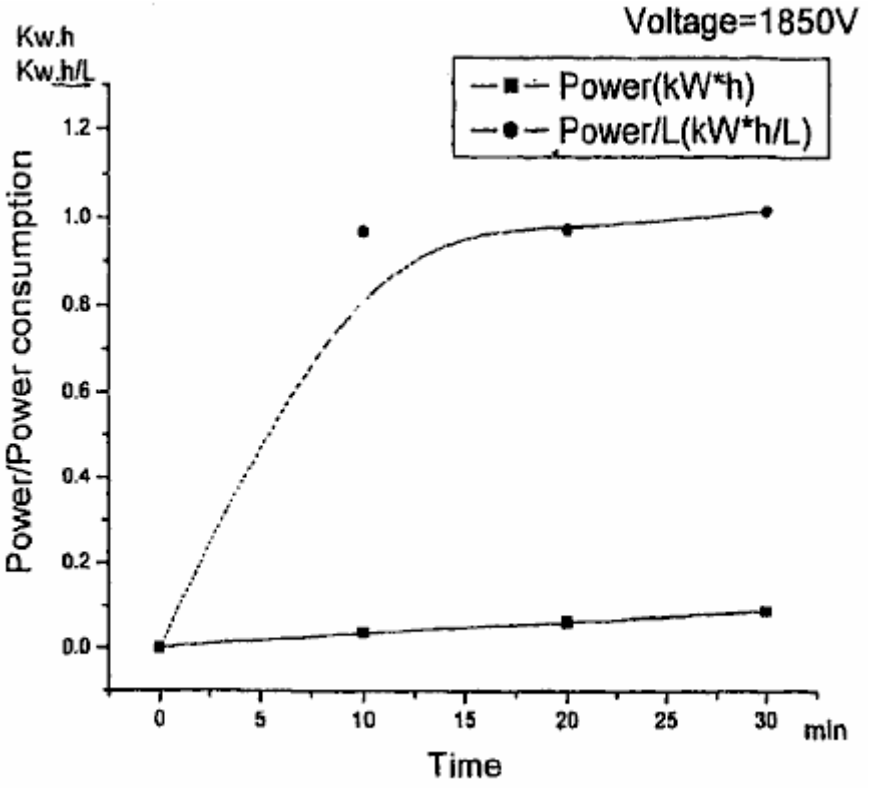

Voltage=1850V
Kw.h
Kw.h/L
Power(kW*h)
Power/L(kW*h/L)
Power/Power consumption
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
30
min
Time

FIG. 47

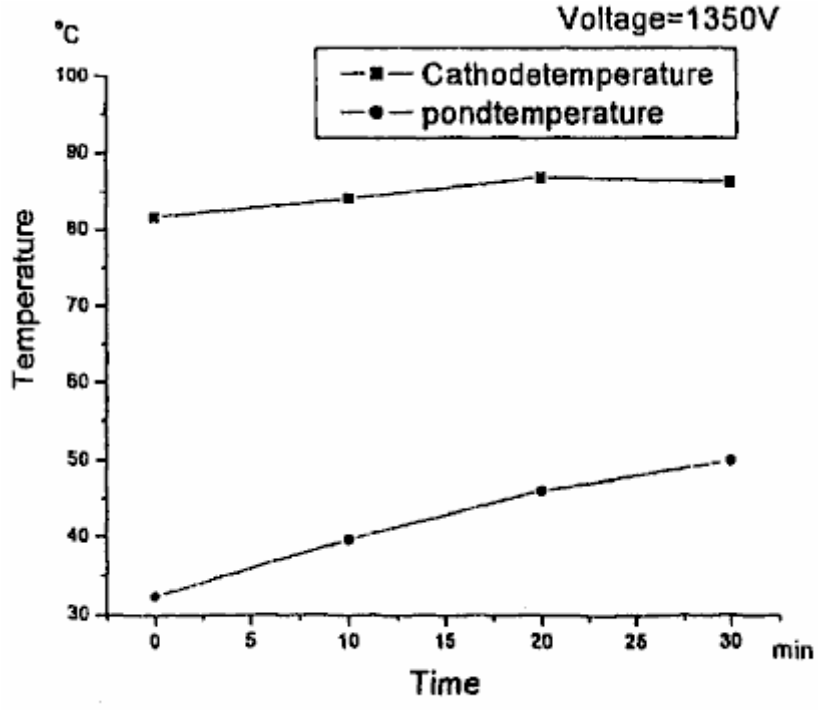

Voltage=1350V
°C
Cathodetemperature
pondtemperature
Temperature
100
90
80
70
60
50
40
30
0
5
10
15
20
25
30
min
Time

FIG. 48

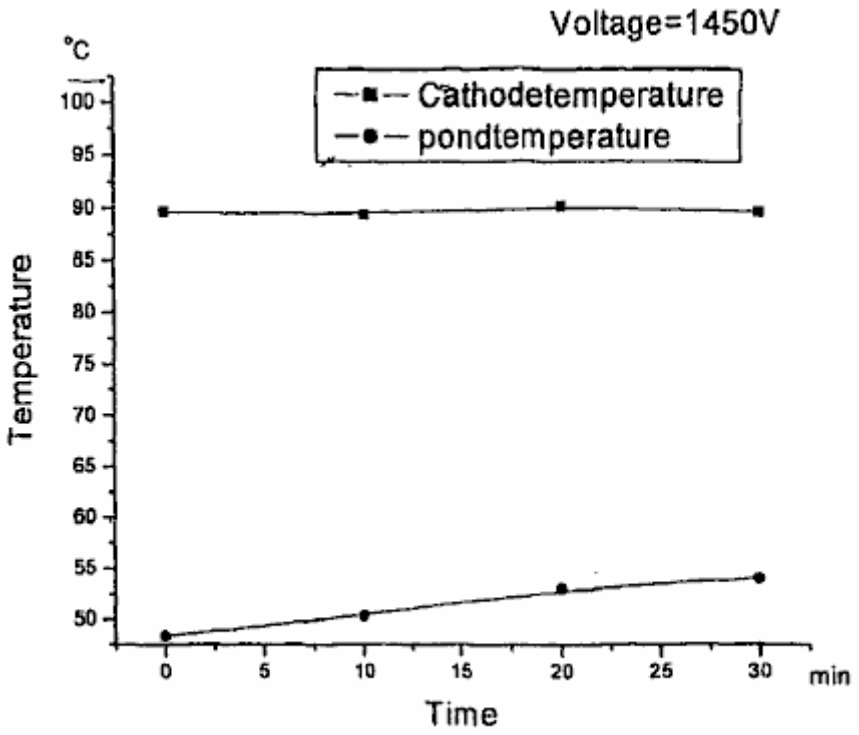

Voltage=1450V
°C
Cathodetemperature
pondtemperature
Temperature
100
95
90
85
80
75
70
65
60
55
50
0
5
10
15
20
25
30
min
Time

FIG. 49

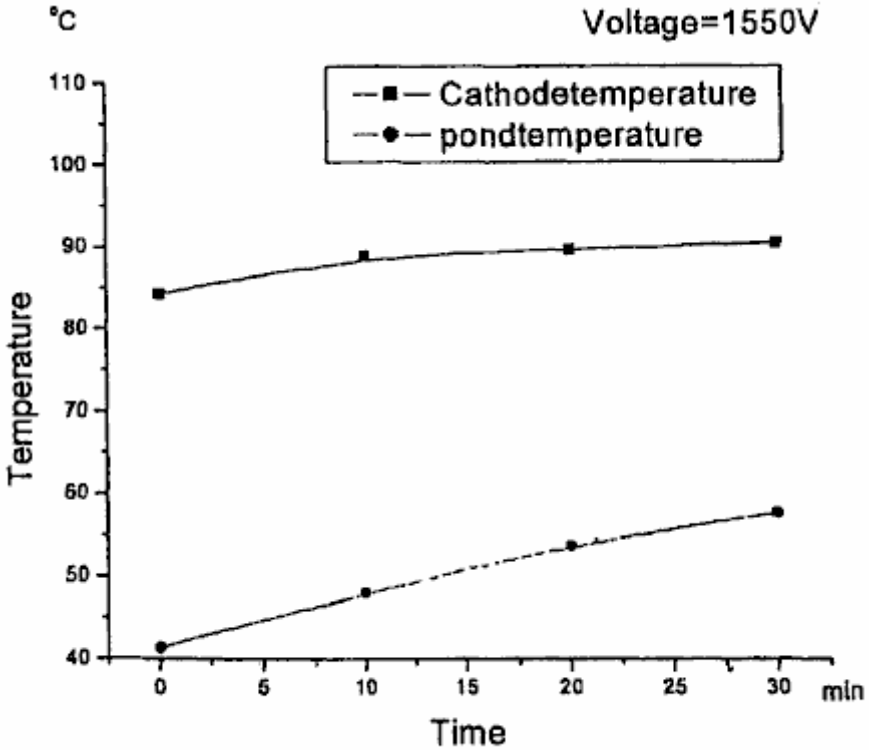

°C
Voltage=1550V
Cathodetemperature
pondtemperature
Temperature
110
100
90
80
70
60
50
40
0
5
10
15
20
25
30
min
Time

FIG. 50

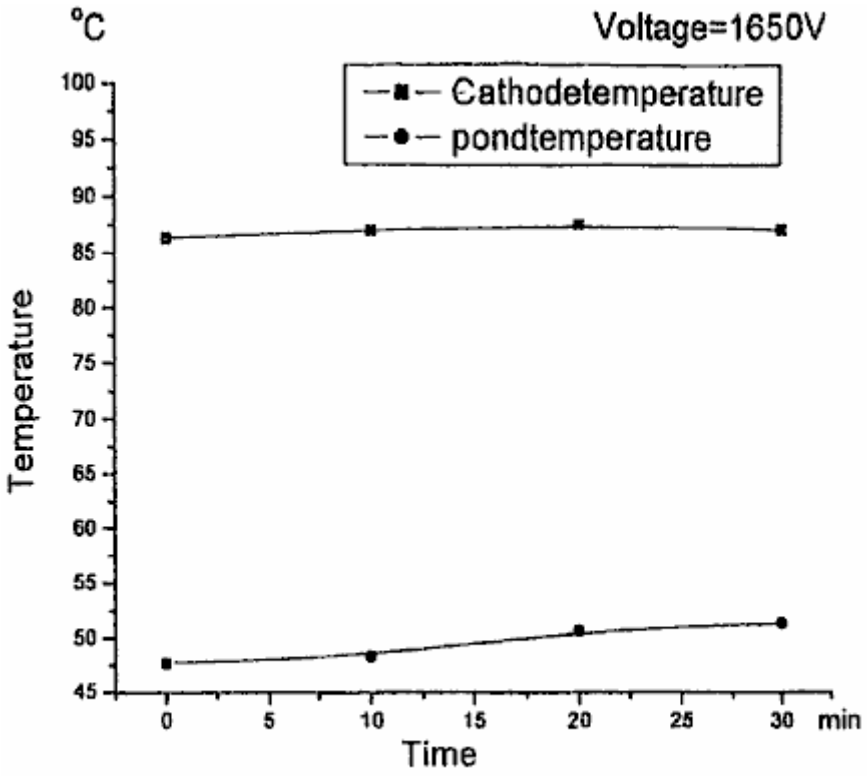

°C
Voltage=1650V
Cathodetemperature
pondtemperature
Temperature
100
95
90
85
80
75
70
65
60
55
50
45
0
5
10
15
20
25
30
min
Time

FIG. 51

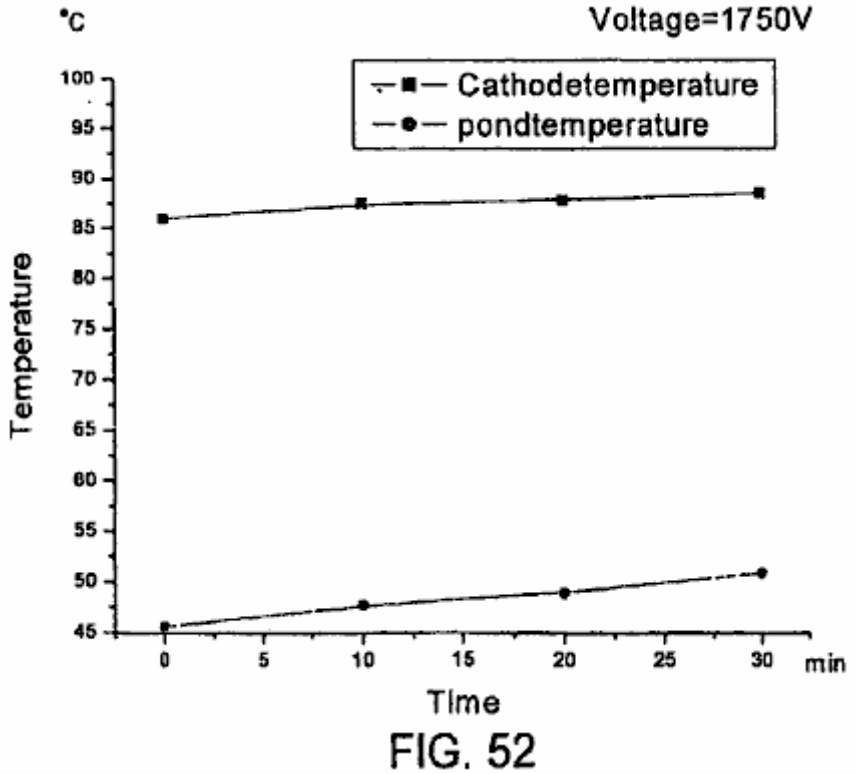
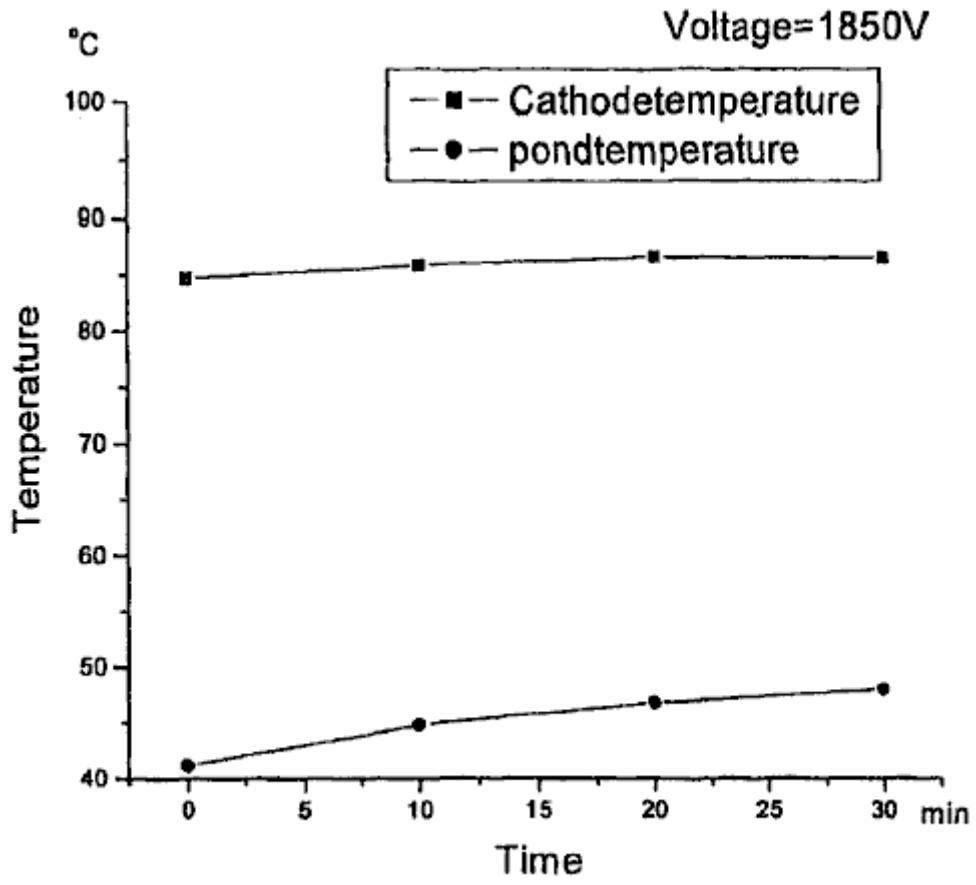

°C
Voltage=1750V
Cathodetemperature
pondtemperature
Temperature
100
95
90
85
80
75
70
65
60
55
50
45
0
5
10
15
20
25
30
min
Time

FIG. 52


Voltage=1850V
°C
Cathodetemperature
pondtemperature
100
90
80
70
60
50
40
Temperature
0
5
10
15
20
25
30
min
Time
FIG. 53

Juan Aguero's Water-Fuel Engine

Patent Application EP0405919 1st February 1991 Inventor: Juan C. Aguero

WATER-PROPELLED INTERNAL-COMBUSTION ENGINE SYSTEM

Please note that this is a re-worded excerpt from this patent application. It describes a method which it is claimed is capable of operating an internal combustion engine from a mixture of steam and hydrogen gas.

ABSTRACT

This is an energy-transforming system for driving, for instance, an internal combustion engine which uses hydrogen gas as its fuel. The gas is obtained by electrolysing water on board and is then injected into the combustion chambers. The electrolysis is carried out in an electrolytic tank **15,** energised with electric current generated by the engine. The hydrogen passes from a reservoir **23,** via collector cylinder **29,** to carburettor device **39**. The hydrogen is then fed into the engine together with dry saturated steam and at least part of the hydrogen may be heated **51** prior to admission. A cooler and more controlled combustion is achieved with the steam and furthermore relatively lesser amounts of hydrogen are required. This is probably caused by the steam acting as a temperature moderator during admission and combustion of the hydrogen and additionally expanding during the expansion stroke.

FIELD OF THE INVENTION

The present invention refers to energy-converter systems, in particular related to an internal combustion engine fuelled by hydrogen gas, i.e. wherein the main propellant admitted to the combustion chambers is hydrogen. More particularly still, the present invention refers to method and means for obtaining hydrogen gas in an efficient and reasonably economical manner, and for supplying the gas to the combustion chambers under conditions for controlled ignition and optimum energy conversion. The present invention also refers to means and method for running an internal-combustion engine system from an available, cheap and non-contaminant hydrogen containing matter such as water as a fuel supply.

In general, the invention may find application in any system employing internal combustion principles, ranging from large installations such as electricity works to relatively smaller automobile systems like locomotives, lorries, motor-cars, ships and motor-boats. In the ensuing description, the invention is generally disclosed for application in the automotive field, however its adaptation and application in other fields may also be considered to be within the purview of the present invention.

BACKGROUND

Dwindling natural resources, dangerous contamination levels, increasing prices and unreliable dependence on other countries are making it increasingly necessary to search an alternative to fossil fuels like oil (hydrocarbons) and oil derivatives as the primary energy source in automobiles. To date, none of the attempted alternatives appears to have proved its worth as a substitute for petrol, either because of inherent drawbacks as to contamination, safety, cost, etc. or because man has not yet been able to find a practical way of applying the alternative energy forms to domestic motor cars.

For instance, electricity is a good alternative in the ecological sense, both chemically and acoustically, however it appears to be the least efficient form of energy known, which together with the high cost of manufacture of electric motors and the severe storage limitations insofar capacity and size have stopped it from coming into the market at least for the time being. The same is generally true even when solar energy is concerned.

Nuclear power is efficient, available and relatively cheap, but extremely perilous. Synthetic fuels may certainly be the answer in the future, however it appears that none practical enough have been developed. Use of gases such as methane or propane, or of alcohol distilled from sugar cane, has also been tried, but for one reason or another its marketing has been limited to small regions. Methanol for instance is a promising synthetic fuel, but it is extremely difficult to ignite in cold weather and has a low energy content (about half that of petrol).

The use of hydrogen gas as a substitute for petrol has been experimented lately. The chemistry investigator Derek P. Gregory is cited as believing that hydrogen is the ideal fuel in not just one sense. Hydrogen combustion produces steam as its only residue, a decisive advantage over contaminating conventional fuels such as petrol and coal. Unfortunately, hydrogen hardly exists on earth in its natural free form but only combined in chemical

compounds, from which it must be extracted using complicated, expensive and often hazardous industrial processes. In addition, if this obstacle were overcome, it would still be necessary to transport and store the hydrogen in service stations and moreover find a safe and practical way of loading and storing it in motor vehicles. Mercedes-Benz for one is experimenting with a vehicle equipped with a special tank for storing hydrogen gas and means for supplying the gas to the injection system, instead of the conventional petrol tank and circuit, without however yet achieving a satisfactory degree of safety and cost-efficiency. The use of dry hydrogen gas as a propellant has heretofore been found to produce a generally uncontrolled ignition, a large temperature excursion upwards which proved too destructive for the chamber walls. The engine life was limited to less than 10,000 km (about 6,000 miles).

DISCLOSURE OF THE INVENTION

The invention is based on the discovery of an energy-converter system to run an internal combustion engine and particularly is based on the discovery of a method and means for reliably, economically, safely and cleanly fuel an internal combustion engine with hydrogen, and obtaining the hydrogen in a usable form to this end from a cheap and plentifully available substance such as water. The hydrogen may be generated in optimum conditions to be fed into the engine.

According to the invention, hydrogen is obtained on board from a readily available hydrogenous source such as ionised water which is subjected to electrolysis, from whence the hydrogen is injected in each cylinder of the engine on the admission stroke. The hydrogen gas is mixed with water vapour (steam at atmospheric temperature) and surrounding air, and when this mixture is ignited within the combustion chamber, the steam (vapour) seems to act as a temperature moderator first and then assist in the expansion stroke. Preferably, the steam is dry saturated steam which, as a moderator, limits the maximum temperature of the combustion, thus helping to preserve the cylinder, valve and piston elements; and in assisting the expansion, the steam expands fast to contribute extra pressure on the piston head, increasing the mechanical output power of the engine. In other words, the inclusion of steam in the hydrogen propellant as suggested by the present invention moderates the negative effects of hydrogen and enhances the positive effects thereof in the combustion cycle.

As a result of this discovery, the amount of hydrogen required to drive the engine is lower than was heretofore expected, hence the electrolysis need not produce more than 10 cc/sec (for example, for a 1,400 cc engine). Thus the amount of electricity required for the electrolysis, a stumbling block in earlier attempts, is lower, so much so, that on-board hydrogen production is now feasible.

The invention includes an apparatus comprising a first system for generating hydrogen and a second system for conditioning and supplying the hydrogen to the admission valves on the cylinder caps. The hydrogen-generating system basically consists of an electrolysis device which receives electrolitically adapted (i.e. at least partially ionised) water or some other suitable hydrogenous substance. An electric power supply is connected to the electrodes of the electrolysis device for generating the hydrogen, and the electricity requirements and the device dimensions are designed for a maximum hydrogen output rate of about 10 cc/sec for a typical automotive application.

The second system comprises means such as a vacuum pump or the like to draw out the hydrogen from the first system, means for supplying the hydrogen gas to the admission valves, means for conditioning the moisture content of the hydrogen, carburettor means or the like for mixing the hydrogen with atmospheric air or some other combustion enabling substance, and means to control and maintain a specified gas pressure valve or range for the hydrogen supplied to the mixing means.

The apparatus was tested and worked surprisingly well. It was discovered that this seemed to be the result of the steam content in the electrolytic hydrogen gas overcoming the pitfalls encountered in the prior art systems which injected relatively dry gas into the cylinder chambers, or at the most with a relatively small proportion of humidity coming from the air itself.

In the preferred embodiment, the electrolysis system is driven with a pulsed DC power signal of up to 80 Amps at between 75 and 100 Volts. The electrolyte is distilled water salted with sodium chloride with a concentration of about 30 grams of salt per litre of water, to 150 grams of salt in 10 litres of water. Other concentrations are possible depending on the kind of engine, fuel and electricity consumption etc. The maximum rate of hydrogen production required for a typical domestic car engine has been estimated at 10 cc/sec. This hydrogen is drawn out by a pump generating a pressure head of around 2 Kg/cm^2 to feed the generated steam-containing hydrogen to a receptacle provided with means for removing the undesired excess of moisture from the gas. The gas is thus mixed with the desired content of steam when it enters the carburettor or mixing device.

In the event that the generated hydrogen does not have enough steam content, dry saturated steam may be added to the hydrogen as it proceeds to the engine. This may done conveniently, before it enters the carburettor

and is mixed with the intake air. Part of the gas may be shunted via a heat-exchanger serpentine connected to the exhaust manifold. This heats some of the gas before it is injected into the base of the carburettor. This heated gas injection operates like a supercharger. The main unheated hydrogen stream is piped directly into the venturi system of the carburettor, where it mixes with air drawn in by the admission stroke vacuum.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig.1 is a schematic layout of the first and second systems and shows the electrolysis device for obtaining hydrogen, and the circuit means for injecting the steam-laden hydrogen into the combustion chambers of a car engine, according to one embodiment of this invention.

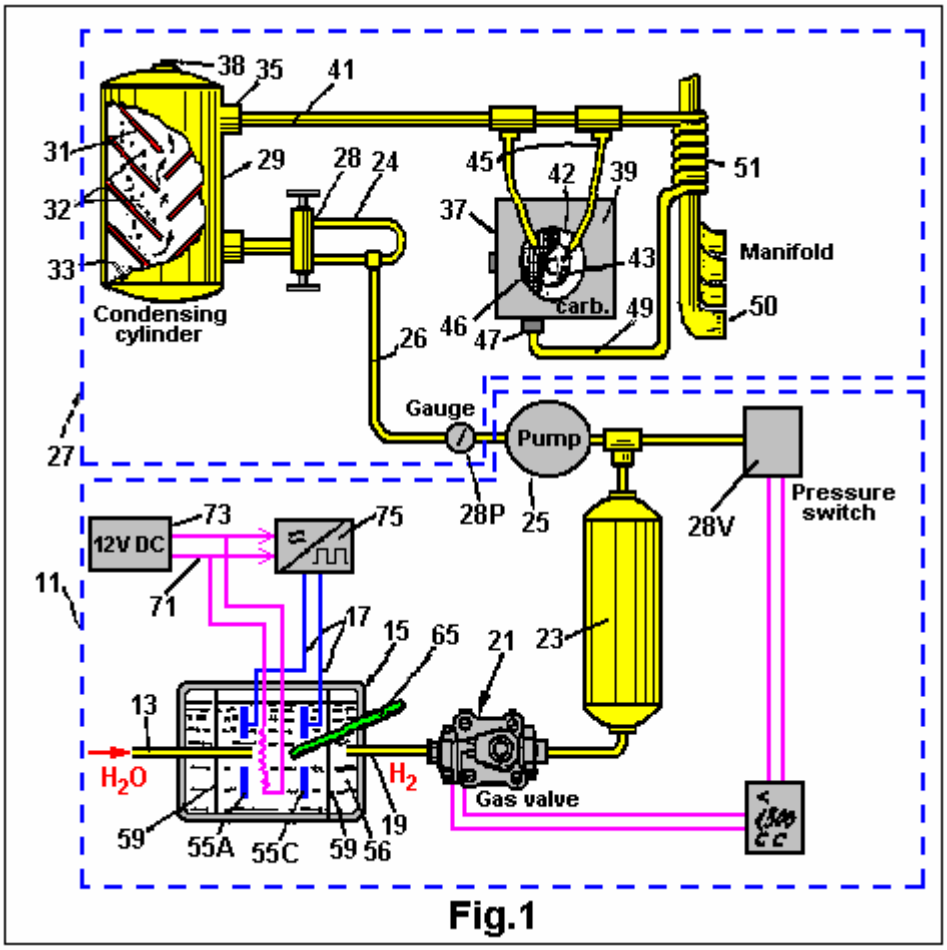


Fig.1

Figure 2 is an elevational view of the electrolysis device of figure 1.

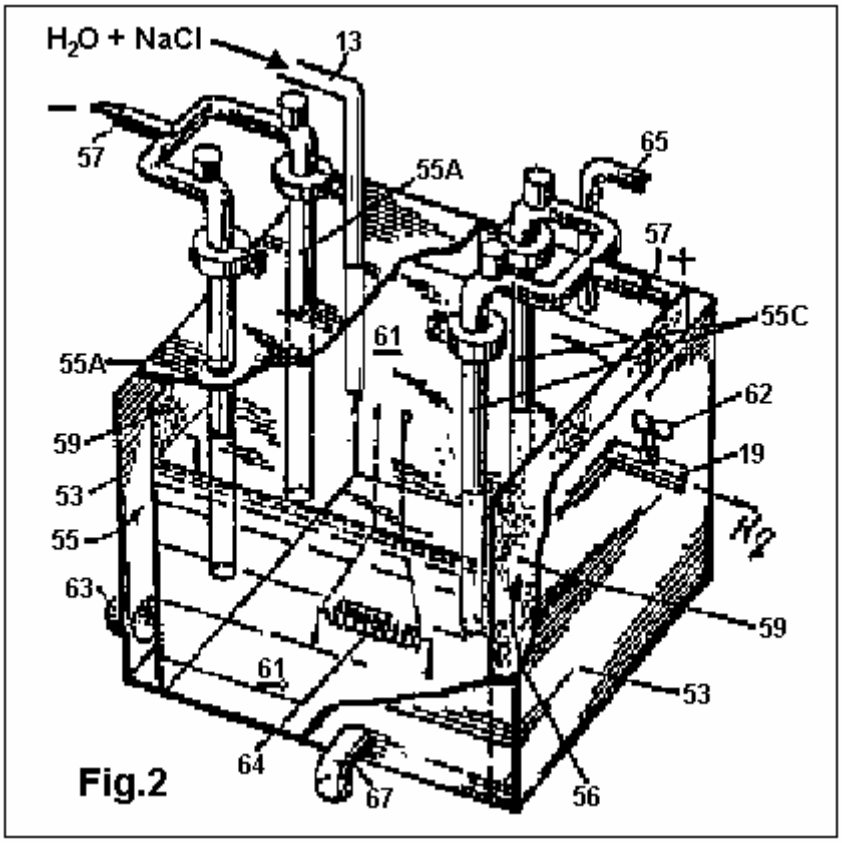


Fig.2

DETAILED ACCOUNT OF AN EMBODIMENT

Fig.1 shows a system **11** for obtaining hydrogen front water piped from a reservoir or tank (not illustrated) to an inlet **13** of an electrolysis cell **15**. The water is salted by adding sodium chloride to ionise it and enable electrolysis when electric power is applied to a pair of terminals **17**. As disclosed in more detail later, the power applied to the terminals **17** is in the form of a DC pulse signal of 65 Amps at 87 Volts, generated via a suitable converter from, in the event that the present system is applied to an automobile, the standard automotive 12 Volt DC level. The device **15** has various outlets, one of which is the hydrogen gas outlet **19** which is connected through a solenoid valve **21** to an accumulator or reservoir cylinder **23**. Other outlets of the electrolysis device **15** are for removing electrolysis effluents such as sodium hydroxide and chlorine gas, to which further reference is made below.

A vacuum pump **25** or similar, extracts gas from the reservoir **23** and channels it through a hydrogen circuit system **27**. Thus the reservoir **23** acts as a pressure buffer of a systems interface between the electrolysis device **15** and the pump **25**. The reservoir **23** may be a 2,000 cc capacity, stainless-steel cylinder with the valve **21** metering the passage of gas through it, so that the reservoir is initially filled with about 1,500 cc of hydrogen at normal pressure and temperature (NPT) conditions. To this end, the cylinder **23** may be provided with a gauge **28V** which controls the state of valve **21** electronically. Valve **21** may be a Jefferson Model SPS solenoid valve, available from OTASI, Santa Rosa 556, Córdoba, Argentina. Vacuum pump **25** is a diaphragm pump with a pulley drive and it is coupled by means of a transmission belt to the engine's crankshaft output. Such a device **25** may be a Bosch model available in Germany. The pulley drive is decoupled by an electromagnetic clutch when the pressure read by a gauge **28P** screwed into the outlet side of pump **25** exceeds 2Kg/sq. cm.

Pump **25** sends hydrogen through tubing **26**, which also includes a by-pass **24** provided for inspection and safety purposes together with a two-way valve **28**, and into a second cylinder **29** which contains means **31** which cause a turbulence or a labyrinthine movement in the gas, in order to condense the heavy mixture, schematically shown as droplets **32**, present in the gas stream. The condensed mixture collects in the form of distilled water **33** at the bottom of cylinder **29**. Near the top of the cylinder, there is an outlet **35** through which hydrogen gas, laden with a good amount of steam, is transported to mixer **37**. Also at the top of collector cylinder **29**, there is a temperature sensor **38** which is connected to an electronic digital thermometer circuit (not shown).

Mixer **37** comprises a carburettor device **39** for mixing hydrogen with air prior to feeding the mixture to the

combustion chambers. The hydrogen is piped through a 3/8" diameter tube **41** from dryer cylinder **29** and then into the venturi section **43** of the carburettor **39** through a pair of 5/16" diameter tubes or hydrogen injecting nozzles **45**. The venturi section **43** is a section of the intake air passage which narrows to increase the air speed at the point where hydrogen is drawn out for mixing. The venturi intake **42** may be covered by a mesh **46**. However, it appears that no air filter is needed for the mixer to operate well. The carburettor device **39** may be a simplified form of a conventional carburettor, since the propellant, i.e. hydrogen gas, is fed directly to the venturi **43**. A butterfly valve, or the like, connected to an accelerator pedal (not illustrated) of the motor-car, controls the air intake rate and therefore the speed of the engine. This mixer device **39** is mounted as is a conventional carburettor, such that its outlet at the bottom communicates with the admission valves in the cylinder caps.

At the bottom part of the carburettor there is a supplementary hydrogen intake **47** connected to another 3/8" diameter pipe **49** which shunts part of the hydrogen through a heater **51**. This heater comprises a serpentine tube **51** of a chromium/cobalt alloy, mounted in close heat-exchange relationship with the body of the exhaust manifold **50** (schematically illustrated) in order to add a portion of heated gas to the fuel mixture before it is drawn into the combustion chambers through the corresponding admission valves on the cylinder caps. This pre-admission heating step, takes the hydrogen mixture to a near critical temperature for detonation. It has been found that this improves performance (e.g. the engine smoothness) at some speed ranges, and it works like a supercharger.

In practice, the engine of the present invention has shown a high efficiency when using three-electrode sparking plugs and an electronic ignition system (not illustrated).

Fig.2 shows the electrolysis cell **15** outlined in **Fig.1** in more detail. It is comprised of a rectangular prism reservoir **53** with a pair of spaced-apart vertical electrodes **55**. The reservoir may measure, for instance, 24 cm long by 20 cm wide and 28 cm high. Both the anode and cathode **55** may each comprise double electrodes of carbon having a spacing between the electrodes **55** of the same polarity of about 10 cm. Alternatively, the anode **55A** may be a ring made of carbon while the cathode **55C** is an iron-mesh cylindrical electrode. Each electrode **55** has a terminal **57** at the top for inputting electric power as mentioned earlier. At each outer side of the electrodes **55** there is a porous membrane **59** made from a sheet of amianto (asbestos) for holding the water solution **61** in whilst at the same time letting the electrolysis products, i.e. hydrogen and oxygen, pass through. Thus, the hydrogen gas passes through the membrane **59** into a gas collector chamber **56** and exits out through pipe **19** to fuel the combustion engine. The hydrogen pipe **19** may have a proportioning valve **62** for regulating the flow of hydrogen. The oxygen on the other hand may be vented out into the atmosphere through an outlet **63**.

There is a heater element **64,** immersed in the salted water **61** fed through a resistor connected to a 12 Volt DC supply. This heats the water to about 85 degrees C (185 degrees F) to enhance the galvanic action of the electrolysis current on the aqueous solution **61**. A thermostat with a solid state silicon thermal sensor may be used to control the water temperature via a threshold comparator driving a relay which controls the current in the heater element **64**.

The electrolysis of the heated salted water solution **61** further produces, as effluents, chlorine gas (Cl_2) and sodium hydroxide (NaOH). The chlorine gas may be vented through an opening **65** at the top of the reservoir **53** or else stored in an appropriate disposal tank (not shown). The sodium hydroxide precipitates and may be removed periodically through tap **67** at the bottom of the electrolysis cell.

It is important to note that the practice of the present invention requires practically no modifications in the engine itself. That is, existing petrol engines may be used with hardly any adjustments. Ignition is initiated at the dead top of the compression stroke or with a 1.5 degree lag at the most, and it has been found convenient to widen the gaps of the admission and exhaust valve pushers and use tri-electrode spark plugs. However it is advisable to use some rust-resistant compound such as plastics for the exhaust pipe and silencer, bearing in mind that the combustion residue is hot steam.

Fig.1 also shows schematically, the electric power supply **71** connected to the terminals **17** of the cube **15**. Electrical current is obtained at 12 volt DC from the car battery/alternator system **73** and processed by an inverter device **75** for generating DC pulses of 65 Amps at 87 Volts. Pulse energisation of the electrolysis appears to maximise the ratio of hydrogen output rate to electric power input.

CLAIMS

1. A method of providing propellant to an internal combustion engine wherein combustion is fuelled on the basis of hydrogen gas admitted into at least one combustion chamber of the engine during the intake stroke, characterised in that the hydrogen is injected into the combustion chamber together with vapour.

2. The method of claim 1, characterised in that the surrounding air enters the combustion chamber, together with the hydrogen and vapour.

3. The method of claim 2, characterised in that the hydrogen gas is obtained from water which is continuously subjected to electrolysis energised by the engine.

4. The method of claim 2 or 3, characterised in that the hydrogen is generated at a rate of not more than 10 cc/sec.

5. The method of any of the preceding claims, characterised in that the engine drives a motor-car.

6. The method of any of preceding claims, characterised in that the vapour is added to the hydrogen prior to entering the combustion chamber.

7. The method of any of claims 1 to 5, characterised in that the vapour is contained in the hydrogen when generated.

8. The method of any of the preceding claims, characterised in that the vapour is dry saturated steam.

9. A method of driving a internal combustion engine with water as its primary source of energy, characterised by the steps of subjecting the water to hydrolysis thereby producing gaseous hydrogen, and
controllably supplying the hydrogen produced by the hydrolysis to the engine combustion chambers during the admission stroke of each cylinder together with a proportion of steam.

10. The method of claim 9, characterised in that the steam is dry saturated steam.

11. The method of any of claims 9 or 10, characterised in that the hydrolysis driven by electric power to produce not more than 10 cc/sec of the hydrogen gas.

12. The method of any of claims 9 to 11, characterised in that the engine drives a motor-car including a water tank as its main propellant supply.

13. The method of any of claims 9 to 12, characterised in that at least part of the hydrogen is heated before injecting it into the chamber.

14. The method of any claims of 9 to 13, characterised in that steam is obtained together with the hydrogen gas from the electrolysis and then subjected to a drying cycle up to a predetermined point of saturation before being passed into the chambers.

15. The method of claim 11, characterised in that the hydrolysis means is supplied with about 5 kW pulsed electrical power.

16.A method of injecting propellant into an hydrogen-driven internal combustion engine cylinder during the admission stroke thereof, characterised in that dry steam is passed into said cylinder during the intake stroke to moderate temperature generation of the hydrogen ignition and enhance expansion after ignition has begun to increase the power of the pistons.

17. A method of obtaining hydrogen capable of being used to fuel an internal combustion engine, characterised by dissociating hydrogen gas from a hydrogenous compound, and admitting the hydrogen gas into each cylinder of said engine together with an amount of dry steam.

18. The method of claim 17, characterised in that the hydrogen gas is admitted to the engine cylinders at a rate of not more than 10 cc/sec.

19. The method of claim 17 or 18, characterised in that the compound is slightly salted water and the steam is saturated steam.

20. A system for obtaining and providing hydrogen propellant to an internal combustion engine including at least one cylinder containing a piston which is subjected to successive combustion cycles and injection means for admitting fuel into the cylinder on the intake or admission stroke of the cycle, characterised by comprising: fuel source means for containing a hydrogenous compound, electrolysis means (**15**) having at least one pair of electrodes (**55**) for receiving electric power and intake means (**13**) connected to the source for supplying the compound to the electrolysis means, a means (**27, 37**) for extracting hydrogen gas from one of the electrodes and supplying it to the cylinder injection means, and control means (**25, 28, 29**) for controlling the supply of hydrogen gas to the cylinder injection means whereby the rate of gas consumption in the engine is not more than 10 cc/sec.

21. The system of claim 20, characterised in that the means supplying hydrogen gas to the cylinder injection means further include means (**37**) for mixing said hydrogen gas with steam.

22. The system of claim 20 or 21, characterised in that the compound is water and the source means includes a water tank, the water including salt to facilitate electrolysis.

23. The system of claim 20, 21 or 22, characterised in that the control means include means (**29**) for removing the excessive moisture from the hydrogen gas extracted from the hydrolysis means.

24. The system of any of claims 20 to 23, characterised in that the electrolysis means is energised by the engine.

25. An internal combustion engine operating on hydrogen and having a water tank as its primary source of combustion fuel, a cylinder block containing at least one cylinder chamber, each chamber, having an associated piston, fuel intake means, ignition means, and exhaust means, and crankshaft means coupled to be driven by the pistons for providing mechanical output power from the engine, and characterised by further comprising: electrolysis means (**15**) connected to the water tank for electrolysing water to obtain hydrogen, electrical means (**17**) connected to supply electric power to at least one pair of electrodes (**55**) of the electrolysis means for carrying out the electrolysis of the water, and hydrogen circuit means (**27**) for extracting the hydrogen gas from the electrolysis means and passing it onto said intake means in a manner enabling controlled ignition and expansion of the fuel in the chamber.

26. The engine of claim 25, characterised in that said hydrogen circuit means passes hydrogen gas to the intake means at a rate of not more than 10 cc/sec.

27. The engine of claim 25 or 26, characterised by further comprising means for adding steam into each chamber before ignition of the hydrogen.

28. The engine of claim 27, characterised in that the steam adder means comprises means (**25**) for extracting steam from the electrolysis means, and means (**29**) for subjecting said steam to a drying process up to a pre-determined point.

29. The engine of any of claims 25 to 28, characterised by further comprising means (**49, 51**) for heating at least part of the hydrogen gas before it is passed into the chambers.

30. The engine of claim 29, characterised in that said heating means is a serpentine (**51**) inserted in a shunt (**49**) of the hydrogen circuit means and mounted in heat-exchange relationship on a manifold exhaust of the engine.

31. The engine of any of claims 25 to 30, characterised in that said electrical means include pulse generator means for supplying electrical pulses to said at least one pair of electrodes.

32. The engine of claim 31, characterised in that said pulse generator means supplies electrical DC pulses of between 50 and 75 Amps at between 60 and 100 Volts.

33. The engine of any of claims 25 to 32, characterised in that said hydrogen circuit means includes drying means (**33**) for removing excess moisture from the hydrogen extracted from the electrolysis means.

34. The engine of any of claims 25 to 33, characterised in that said crankshaft means drives a water-fuelled automobile.

35. The engine of any of claims 25 to 34, characterised in that the electrolysis means is driven by electricity derived from the engine.

The HHO Fuel System of Stephen Horvath

US Patent 3,980,053 14th September 1976 Inventor: Stephen Horvath

FUEL SUPPLY APPARATUS FOR INTERNAL COMBUSTION ENGINES

Please note that this is a re-worded excerpt from this patent. It describes the water-splitting procedure of Stephen Horvath.

ABSTRACT

A fuel supply apparatus generates hydrogen and oxygen by electrolysis of water. There is provided an electrolytic cell which has a circular anode surrounded by a cathode with a porous membrane between them. The anode is fluted and the cathode is slotted to provide anode and cathode areas of substantially equal surface area. A pulsed electrical current is provided between the anode and cathode for the efficient generation of hydrogen and oxygen.

The electrolytic cell is equipped with a float, which detects the level of electrolyte within the cell, and water is added to the cell as needed to replace the water lost through the electrolysis process. The hydrogen and oxygen are collected in chambers which are an integral part of the electrolytic cell, and these two gases are supplied to a mixing chamber where they are mixed in the ratio of two parts hydrogen to one part oxygen. This mixture of hydrogen and oxygen flows to another mixing chamber wherein it is mixed with air from the atmosphere.

The system is disclosed as being installed in an car, and a dual control system, which is actuated by the car throttle, first meters the hydrogen and oxygen mixture into the chamber wherein it is combined with air and then meters the combined mixture into the car engine. The heat of combustion of a pure hydrogen and oxygen mixture is greater than that of a gasoline and air mixture of comparable volume, and air is therefore mixed with the hydrogen and oxygen to produce a composite mixture which has a heat of combustion approximating that of a normal gas-air mixture. This composite mixture of air, hydrogen and oxygen then can be supplied directly to a conventional internal combustion engine without overheating and without creation of a vacuum in the system.

BACKGROUND OF THE INVENTION

This invention relates to internal combustion engines. More particularly it is concerned with a fuel supply apparatus by means of which an internal combustion engine can be run on a fuel comprised of hydrogen and oxygen gases generated on demand by electrolysis of water.

In electrolysis a potential difference is applied between an anode and a cathode in contact with an electrolytic conductor to produce an electric current through the electrolytic conductor. Many molten salts and hydroxides are electrolytic conductors but usually the conductor is a solution of a substance which dissociates in the solution to form ions. The term "electrolyte" will be used herein to refer to a substance which dissociates into ions, at least to some extent, when dissolved in a suitable solvent. The resulting solution will be referred to as an "electrolyte solution".

Faraday's Laws of Electrolysis provide that in any electrolysis process the mass of substance liberated at an anode or cathode is in accordance with the formula

m = z q

where m is the mass of substance liberated in grams, z is the electrochemical equivalent of the substance, and q is the quantity of electricity passed, in coulombs. An important consequence of Faraday's Laws is that the rate of decomposition of an electrolyte is dependent on current and is independent of voltage. For example, in a conventional electrolysis process in which a constant current I amps flows to t seconds, q = It and the mass of material deposited or dissolved will depend on I regardless of voltage, provided that the voltage exceeds the minimum necessary for the electrolysis to proceed. For most electrolytes, the minimum voltage is very low.

There have been previous proposals to run internal combustion engines on a fuel comprised of hydrogen gas. Examples of such proposals are disclosed in U.S. Pat. Nos. 1,275,481, 2,183,674 and 3,471,274 and British specifications Nos., 353,570 and 364,179. It has further been proposed to derive the hydrogen from electrolysis of water, as exemplified by U.S. Pat. No. 1,380,183. However, none of the prior art constructions is capable of producing hydrogen at a rate such that it can be fed directly to internal combustion engines without intermediate

storage. The present invention enables a fuel comprised of hydrogen and oxygen gases to be generated by electrolysis of water at such a rate that it can sustain operation of an internal combustion engine. It achieves this result by use of an improved electrolysis process of the type generally proposed in the parent application hereof.

As disclosed in my aforesaid parent application the prior art also shows electrolytic reactions employing DC or rectified AC which necessarily will have a ripple component; an example of the former being shown for instance in Kilgus U.S. Pat. No. 2,016,442 and an example of the latter being shown in Emich al. U.S. Pat. No. 3,485,742. It will be noted that the Kilgus Patent also discloses the application of a magnetic field to his electrolyte, which field is said to increase the production of gas at the two electrodes.

SUMMARY OF THE INVENTION

The apparatus of the invention applies a pulsating current to an electrolytic solution of an electrolyte in water. Specifically, it enables high pulses of quite high current value and appropriately low voltage to be generated in the electrolyte solution by a direct input supply to produce a yield of electrolysis products such that these products may be fed directly to the internal combustion engine. The pulsating current generated by the apparatus of the present invention is to be distinguished from normal variations which occur in rectification of AC current and as hereinafter employed the term pulsed current will be taken to mean current having a duty cycle of less than 0.5.

It is a specific object of this invention to provide a fuel supply apparatus for an internal combustion engine by which hydrogen and oxygen gases generated by electrolysis of water are mixed together and fed directly to the internal combustion engine.

A still further object of the invention is to provide, for use with an internal combustion engine having inlet means to receive a combustible fuel, fuel supply apparatus comprising:

a vessel to hold an electrolyte solution of electrolyte dissolved in water;

an anode and a cathode to contact the electrolyte solution within the vessel;

electrical supply means to apply between said diode and said cathode pulses of electrical energy to induce a pulsating current in the electrolyte solution thereby to generate by electrolysis hydrogen gas at the cathode and oxygen gas at the anode;

gas collection and delivery means to collect the hydrogen and oxygen gases and to direct them to the engine inlet means; and

water admission means for admission of water to said vessel to make up loss due to electrolysis.

In order that the invention may be more fully explained one particular example of an car internal combustion engine fitted with fuel supply apparatus in accordance with the invention will now be described in detail with reference to the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig.1 is a plan view of part of the car with its engine bay exposed to show the layout of the fuel supply apparatus and the manner in which it is connected to the car engine;

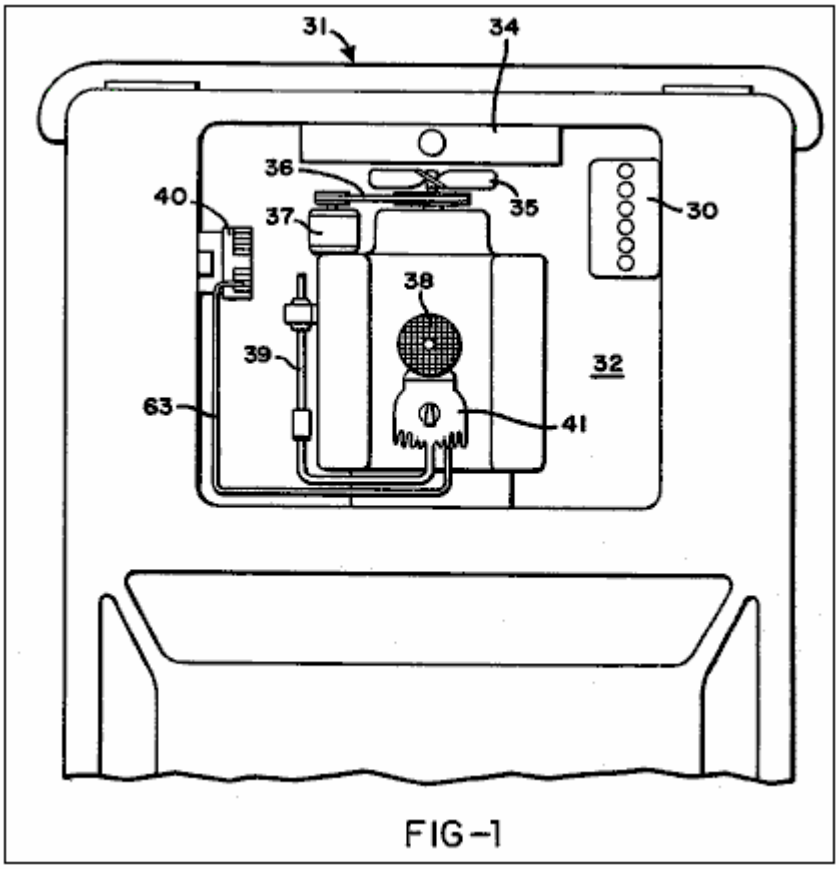


FIG-1

Fig.2 is a circuit diagram of the fuel supply apparatus;

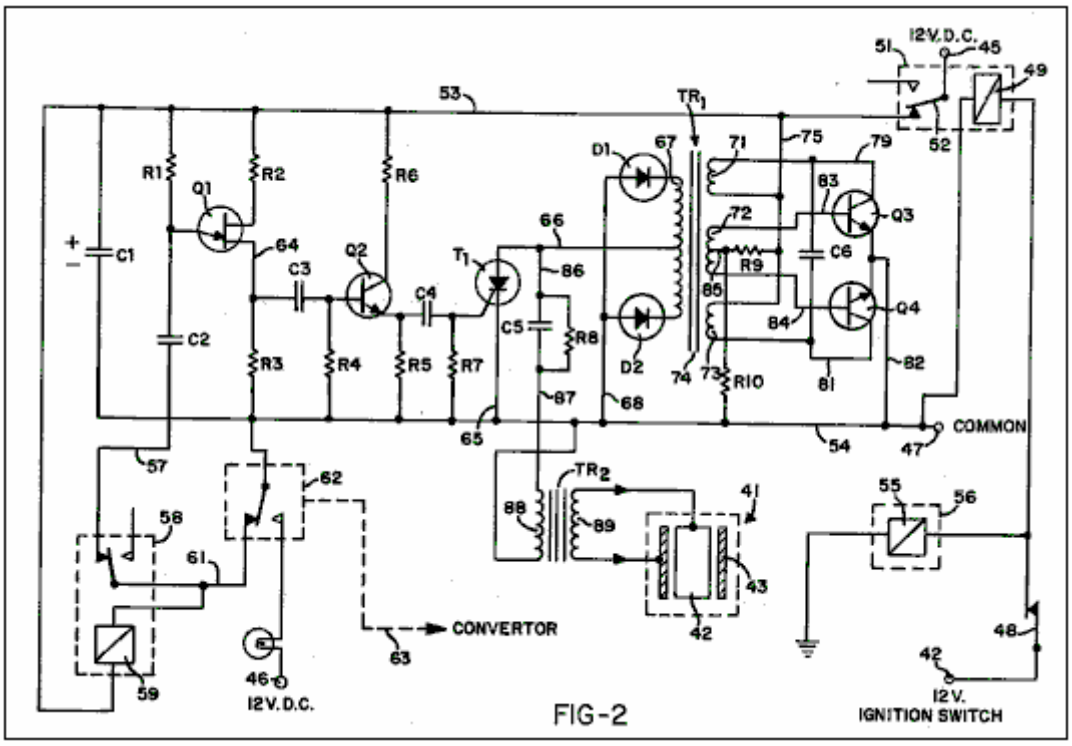


FIG-2

Fig.3 is a plan view of a housing which carries electrical components of the fuel supply apparatus;

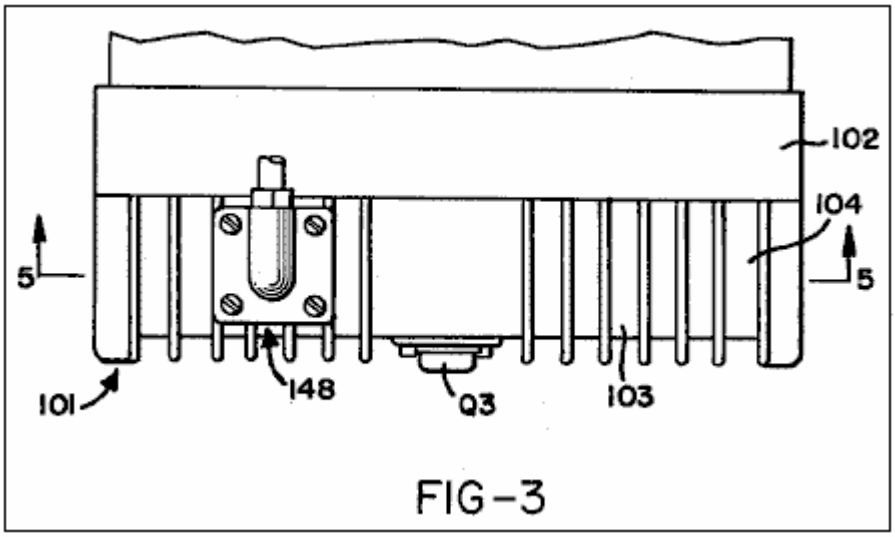


FIG-3

Fig.4 is an elevation view of the housing shown in Fig.3;

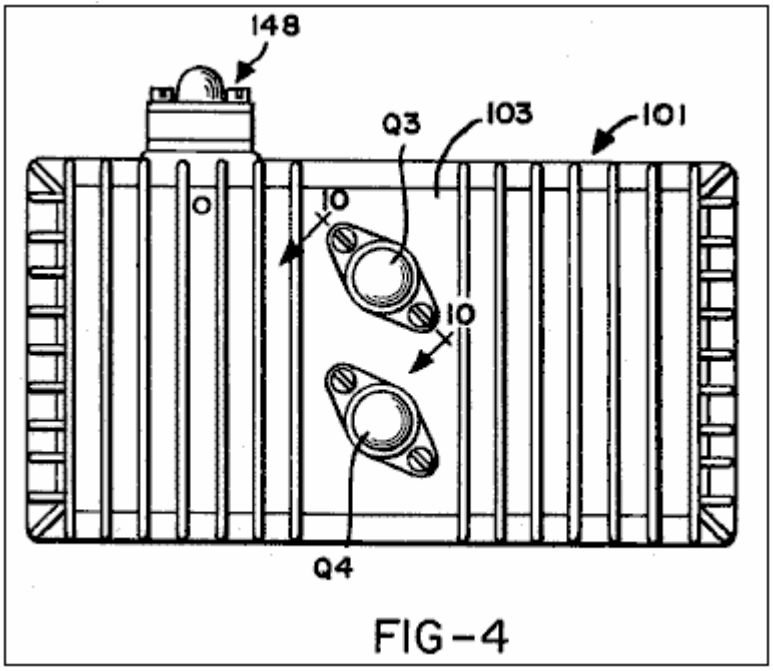


FIG-4

Fig.5 is a cross-section on the line 5--5 in Fig.3;

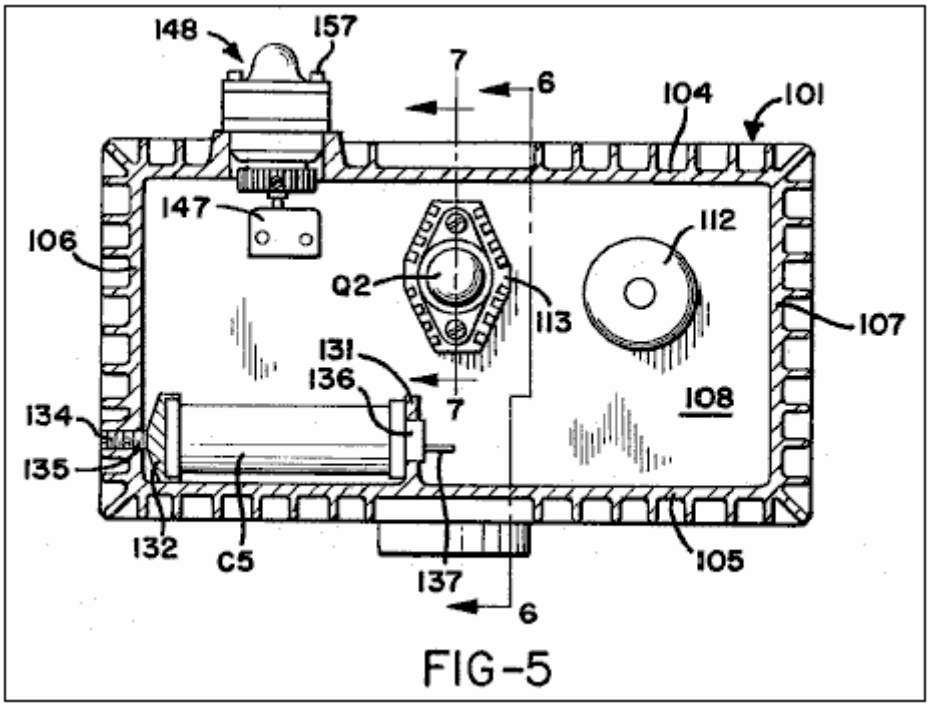


FIG-5

Fig.6 is a cross-section on the line 6--6 in Fig.3;

Fig.7 is a cross-section on the line 7--7 in Fig.5;

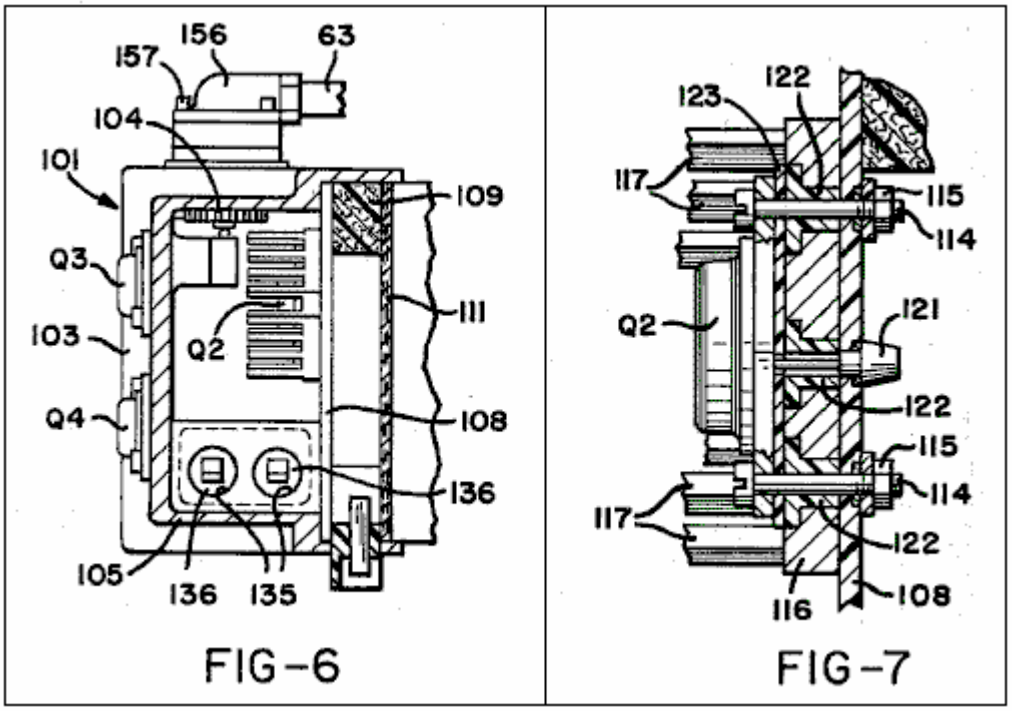


FIG-6 FIG-7

Fig.8 is a perspective view of a diode heat sink included in the components illustrated in Fig.5 and Fig.7;

Fig.9 illustrates a transformer coil assembly included in the electrical components mounted within the housing;

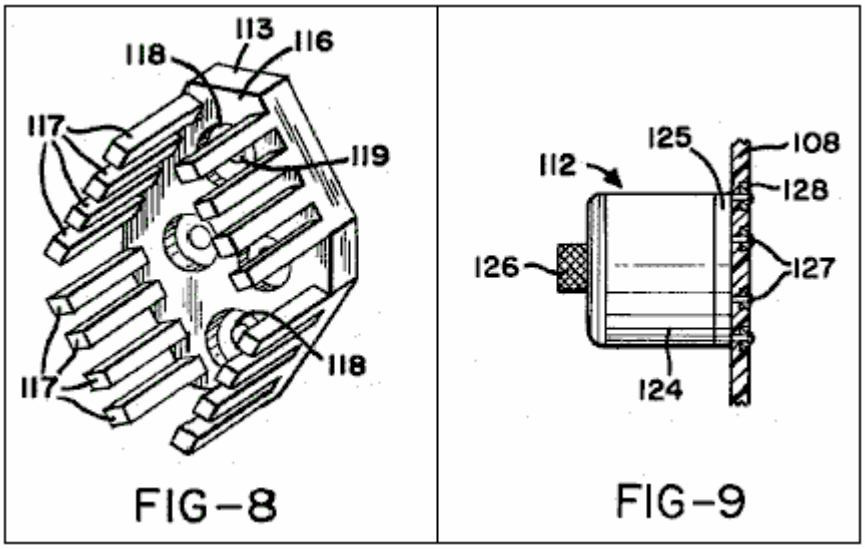


Fig.10 is a cross-section on the line 10--10 in Fig.4;

Fig.11 is a cross-section on the line 11--11 in Fig.5;

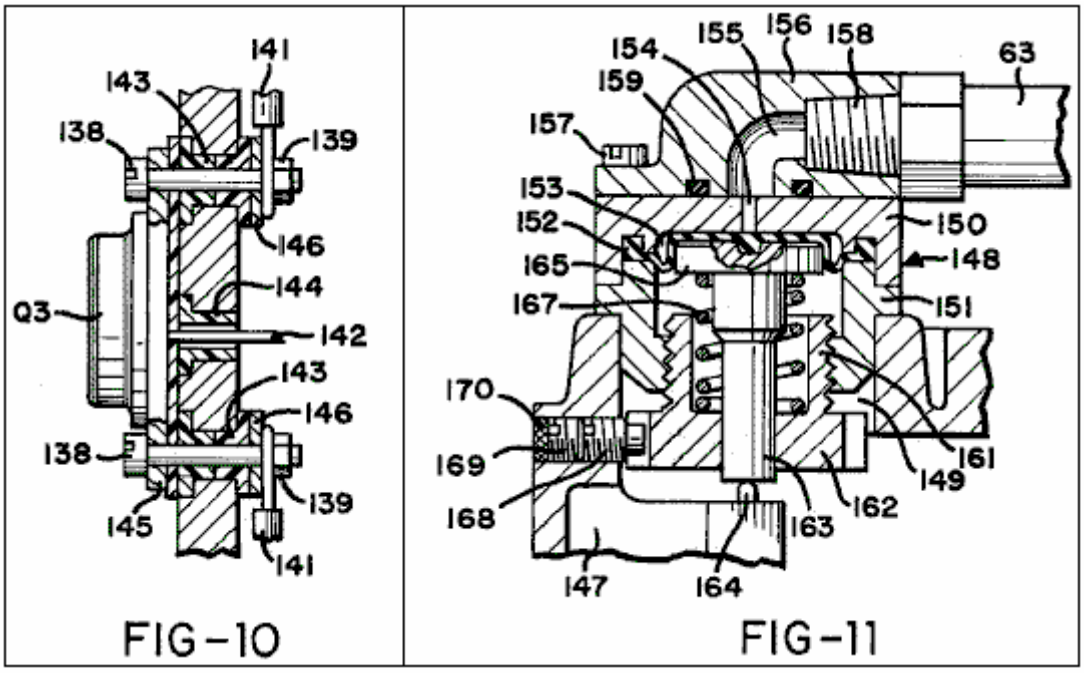


Fig.12 is a cross-section through a terminal block mounted in the floor of the housing;

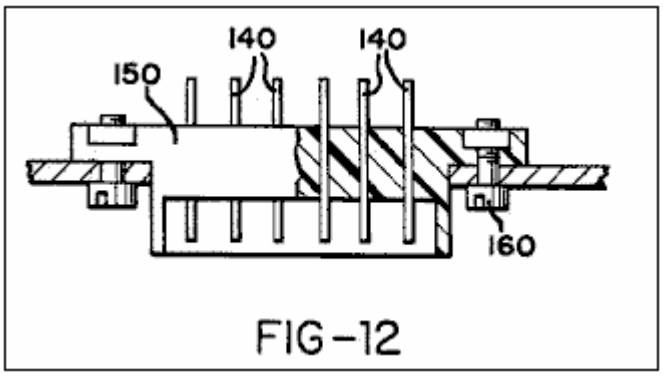


Fig.13 is a plan view of an electrolytic cell incorporated in the fuel supply apparatus;

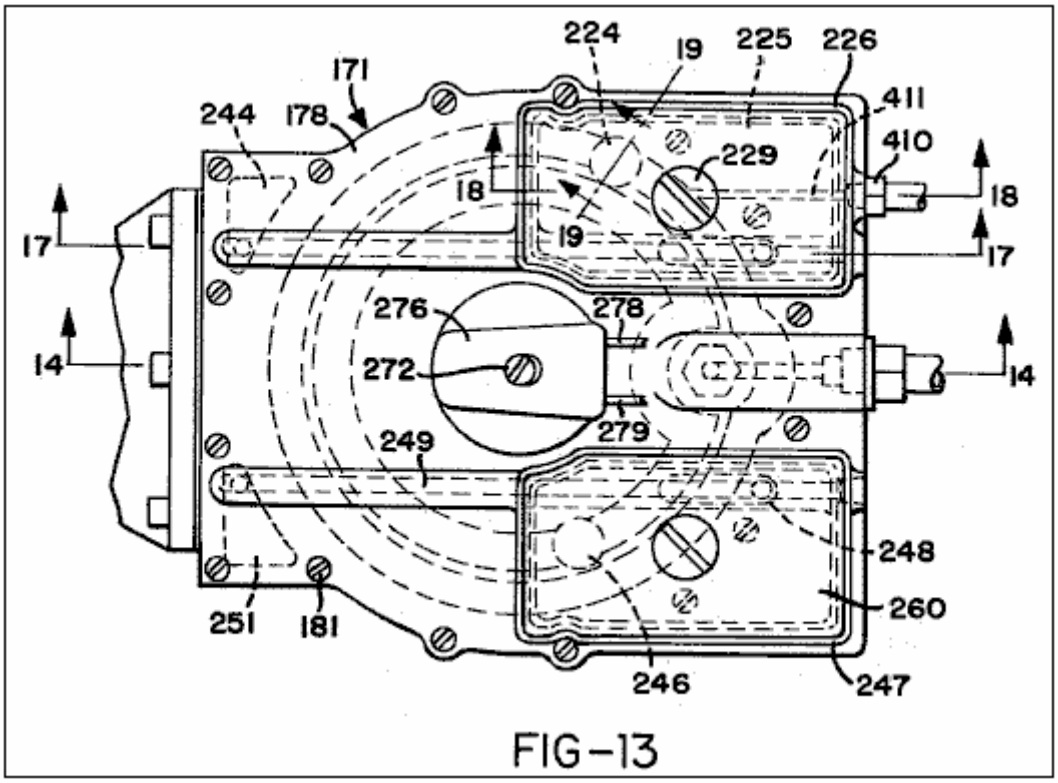


FIG-13

Fig.14 is a cross-section on the line 14--14 in Fig.13;

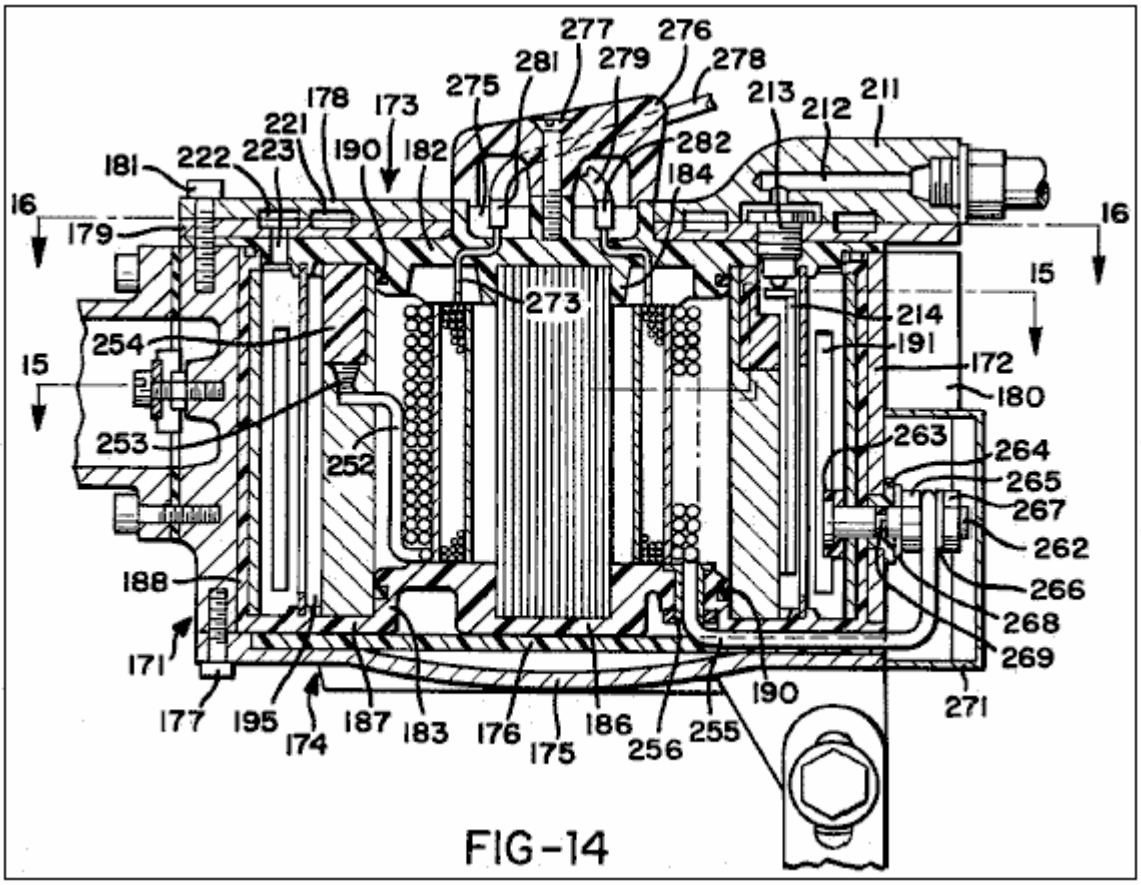


FIG-14

Fig.15 is a cross-section generally on the line 15--15 in Fig.14;

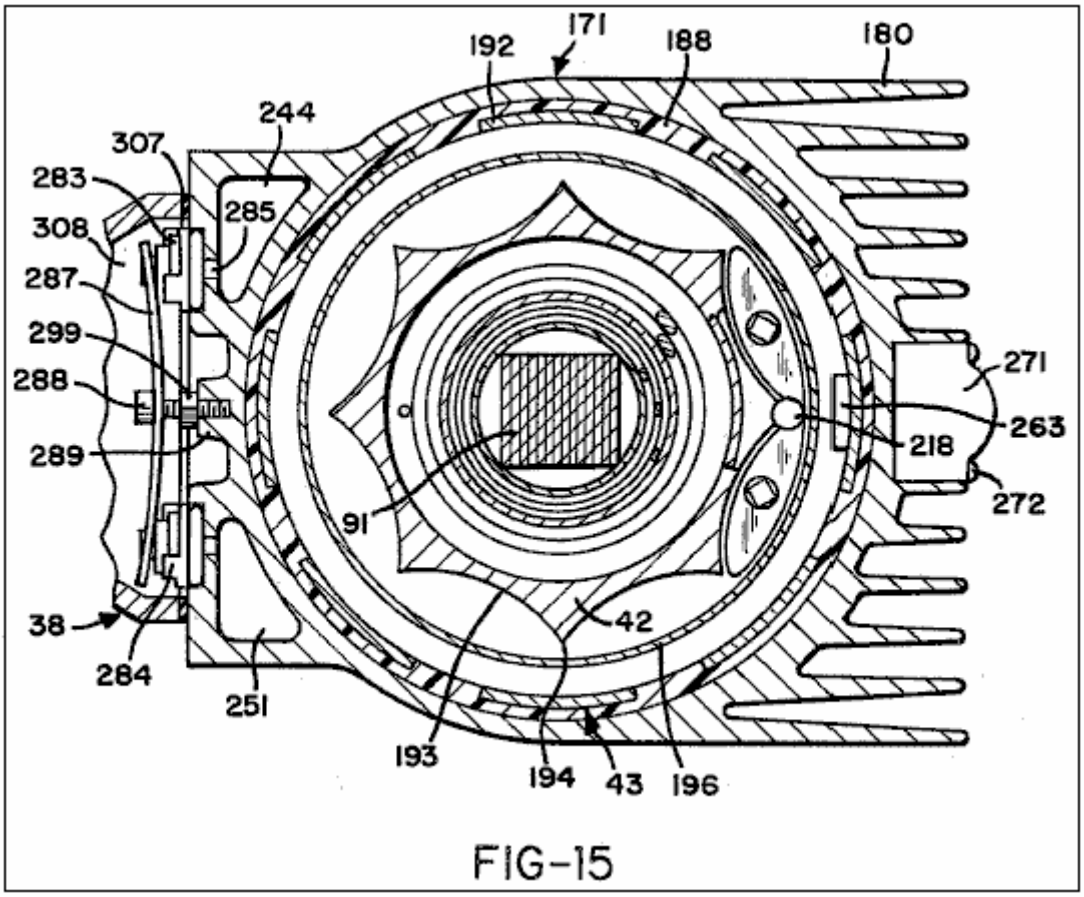


FIG-15

Fig.16 is a cross-section on the line 16--16 in Fig.14;

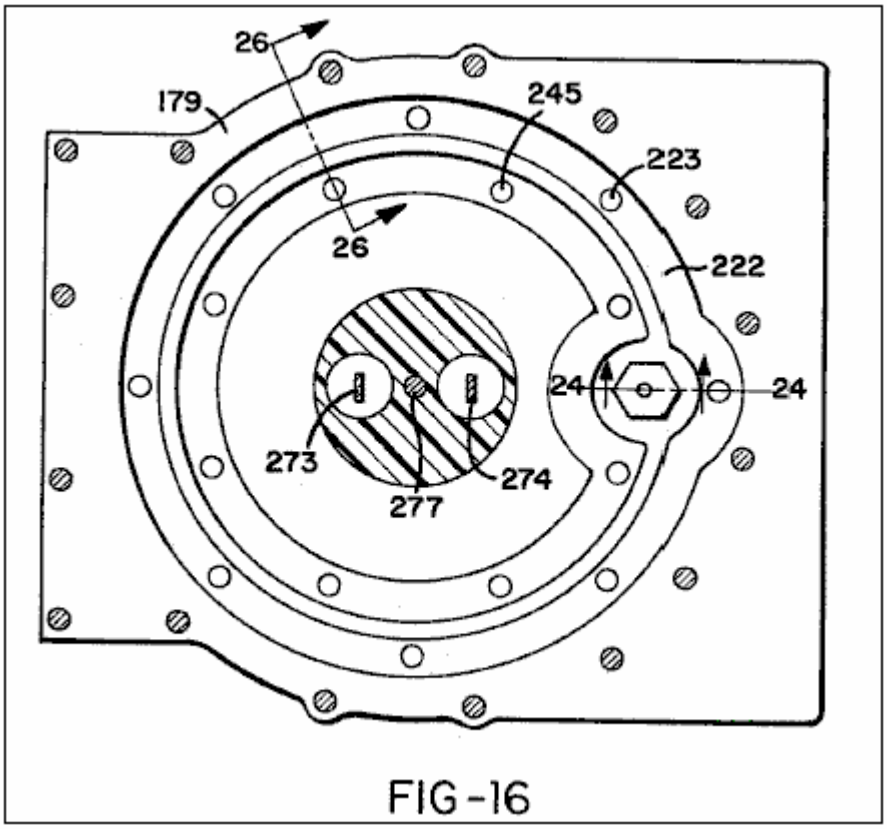


FIG-16

Fig.17 is a cross-section on the line 17--17 in Fig.13;

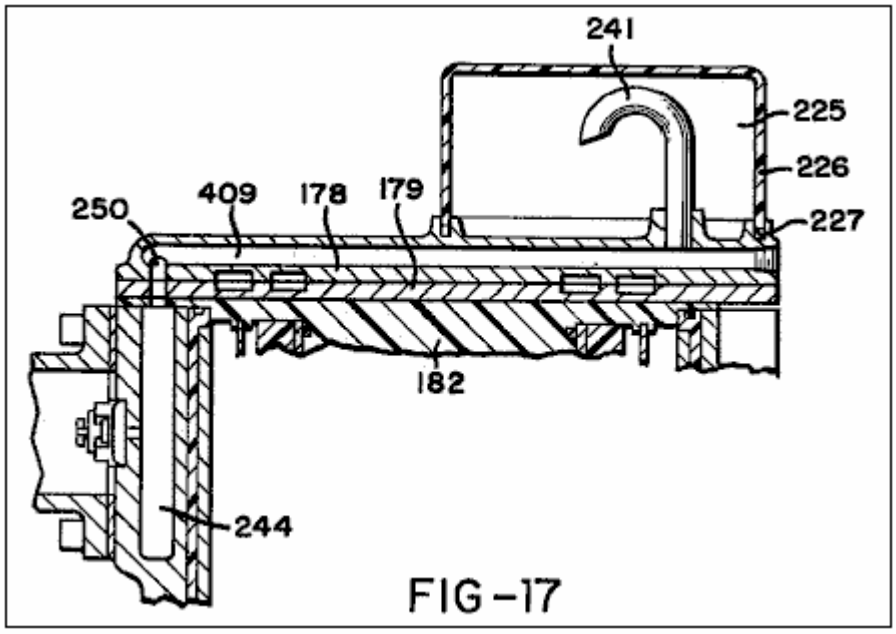


FIG-17

Fig.18 is a cross-section on the line 18--18 of Fig.13;

Fig.19 is a vertical cross-section through a gas valve taken generally on line 19--19 in Fig.13;

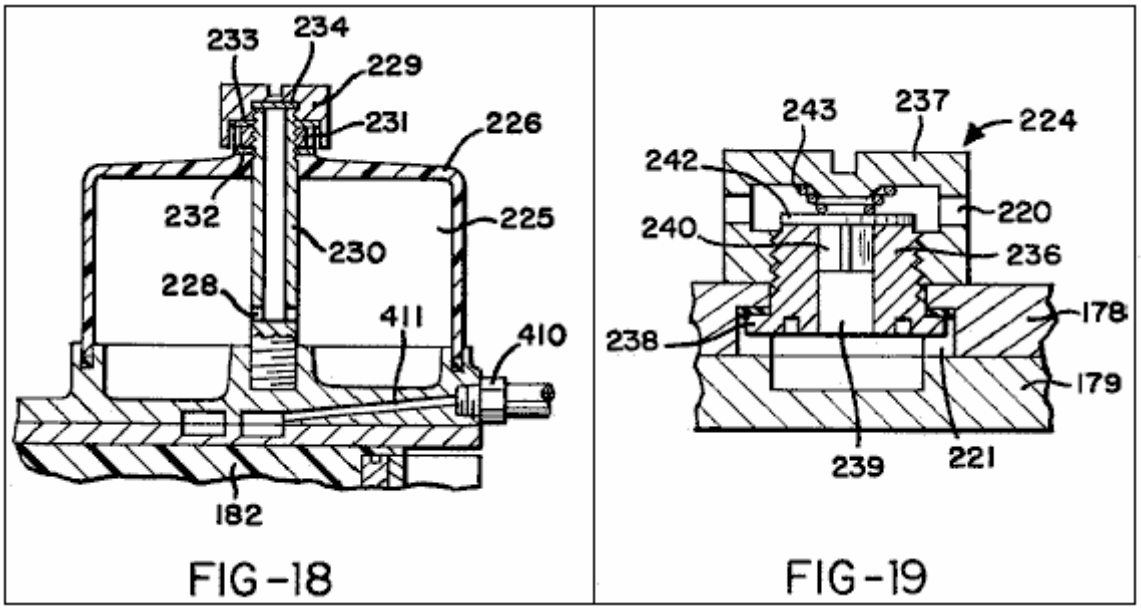


FIG-18 FIG-19

Fig.20 is a perspective view of a membrane assembly disposed in the electrolytic cell;

Fig.21 is a cross-section through part of the membrane assembly;

Fig.22 is a perspective view of a float disposed in the electrolytic cell;

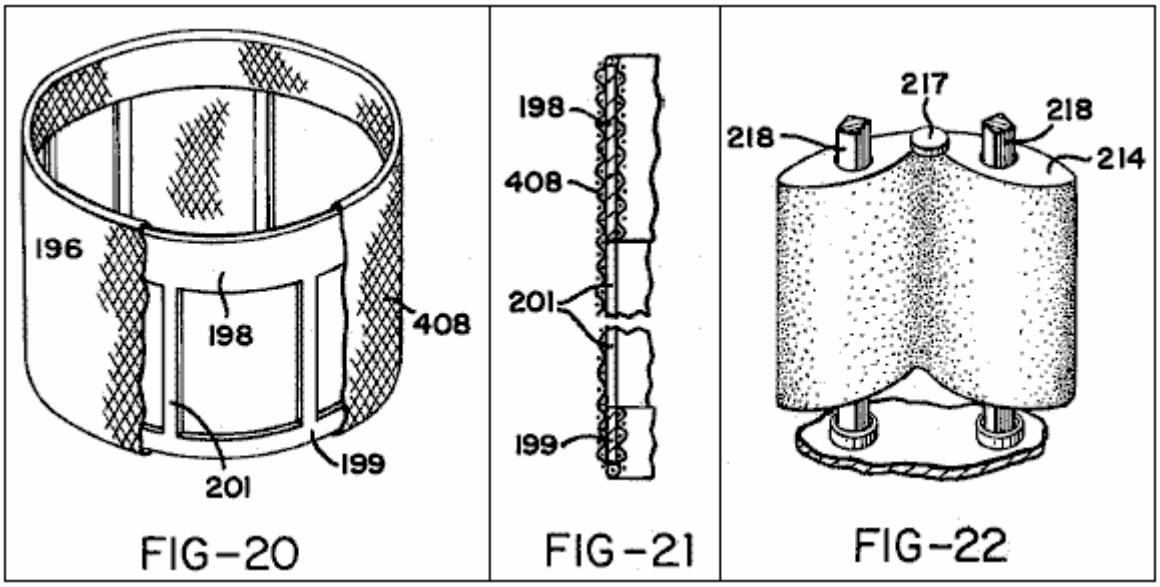


Fig.23 is an enlargement of part of Fig.14;

Fig.24 is an enlarged cross-section on the line 24--24 in Fig.16;

Fig.25 is a perspective view of a water inlet valve member included in the components shown in Fig.24;

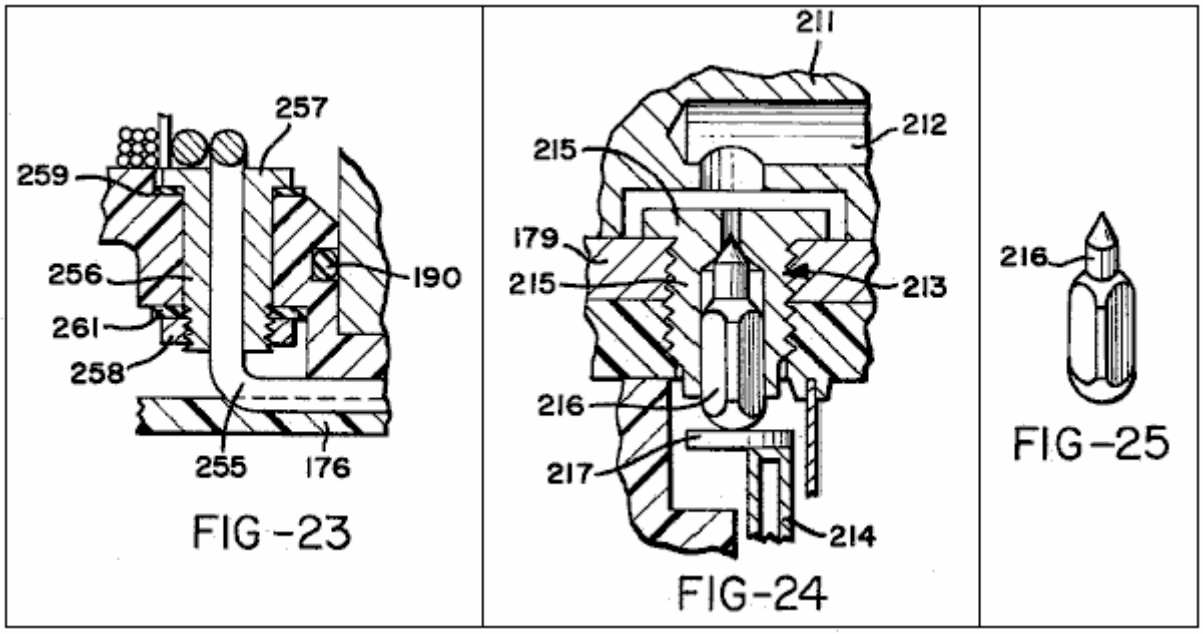


Fig.26 is a cross-section on line 26--26 in Fig.16;

Fig.27 is an exploded and partly broken view of a cathode and cathode collar fitted to the upper end of the cathode;

Fig.28 is an enlarged cross-section showing some of the components of Fig.15;

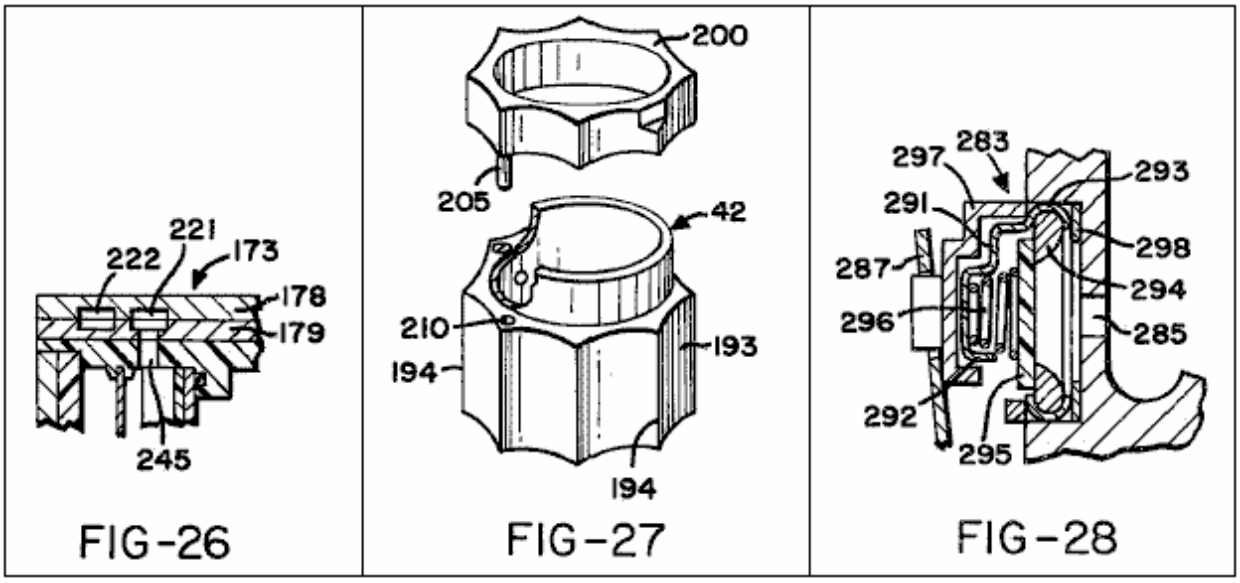


FIG-26 FIG-27 FIG-28

Fig.29 is a perspective view of a valve cover member;

Fig.30 shows a gas mixing and delivery unit of the apparatus generally in side elevation but with an air filter assembly included in the unit shown in section;

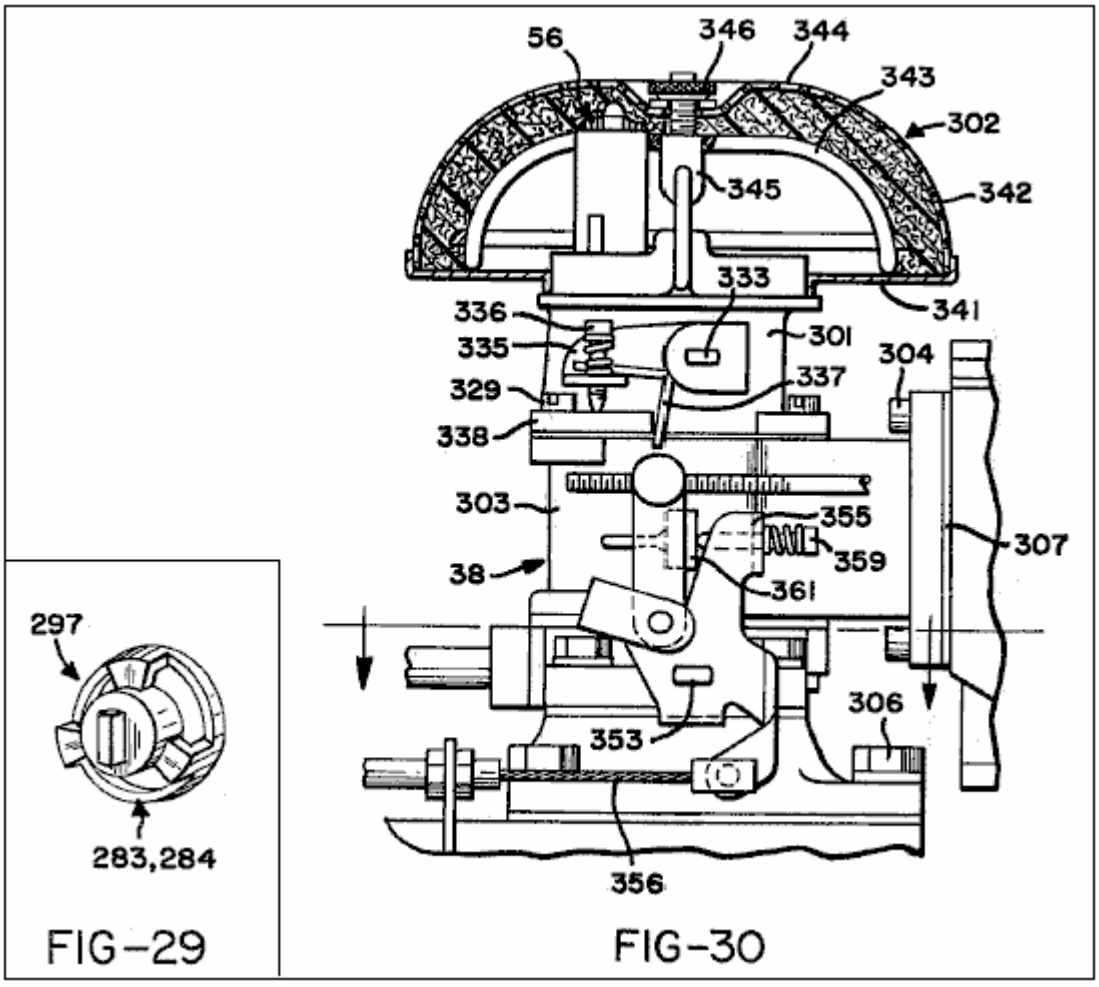


FIG-29 FIG-30

Fig.31 is a vertical cross-section through the gas mixing and delivery unit with the air filter assembly removed;

Fig.32 is a cross-section on the line 32--32 in Fig.31;

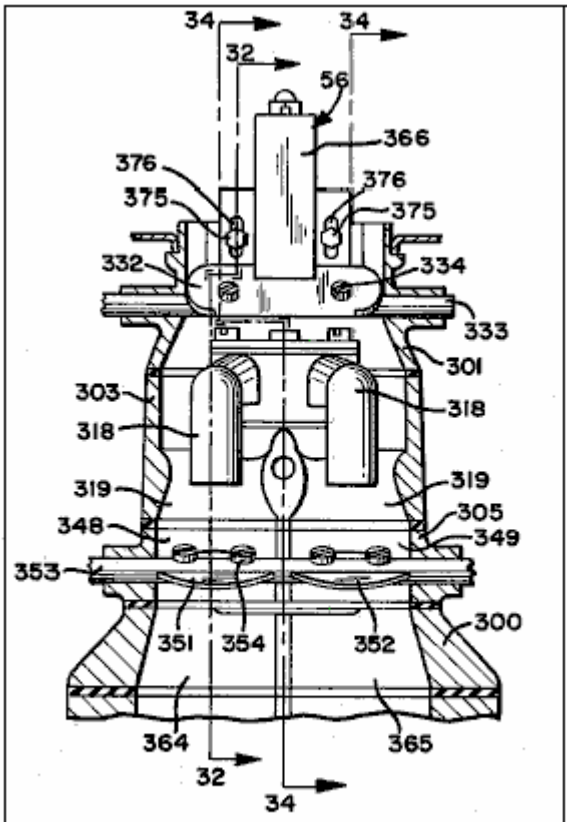


FIG-31

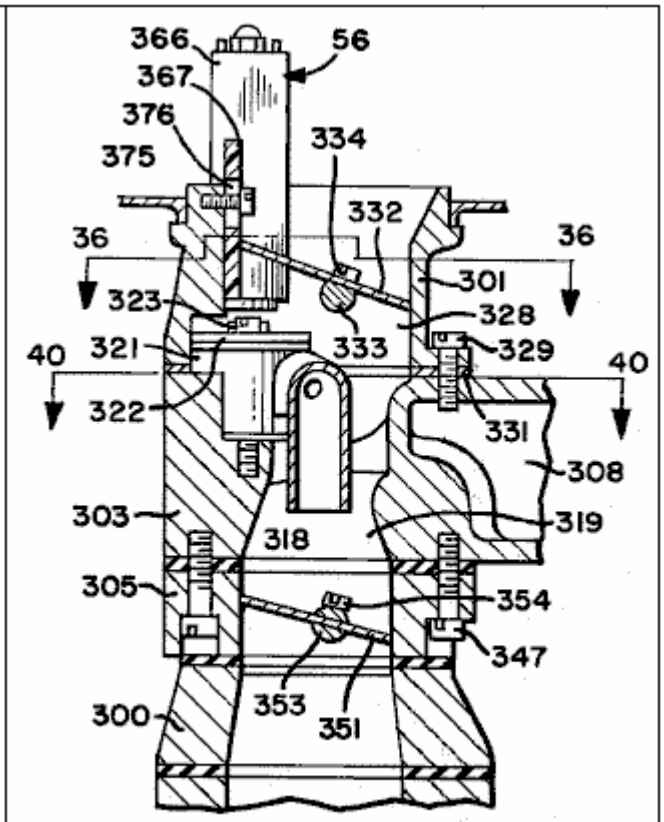


FIG-32

Fig.33 is a perspective view of a valve and jet nozzle assembly incorporated in the gas mixing and delivery unit;

Fig.34 is a cross-section generally on the line 34--34 in Fig.31;

Fig.35 is a cross-section through a solenoid assembly;

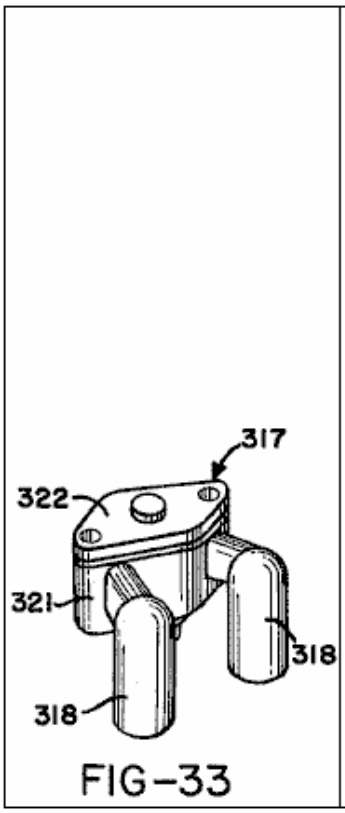


FIG-33

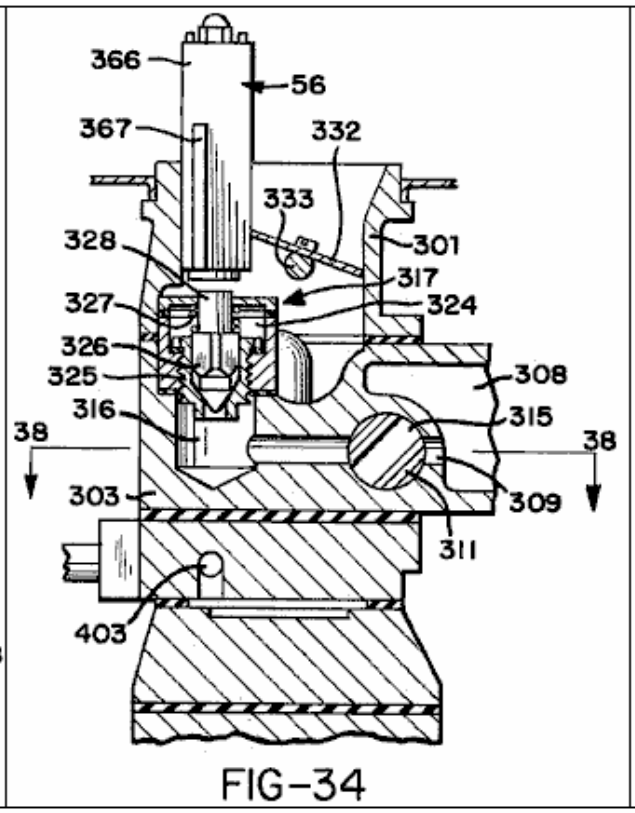


FIG-34

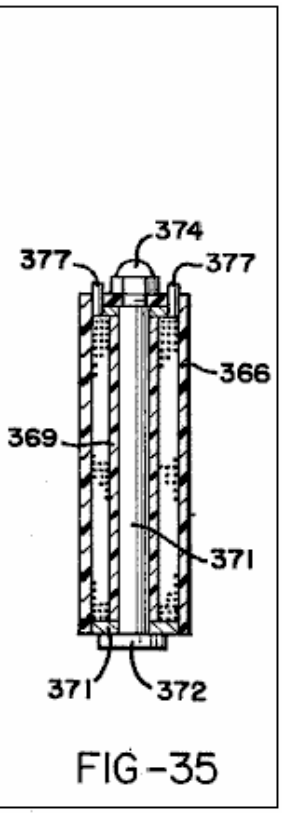


FIG-35

Fig.36 is a cross-section on the line 36--36 in Fig.32;

Fig.37 is a rear elevation of part of the gas mixing and delivery unit;

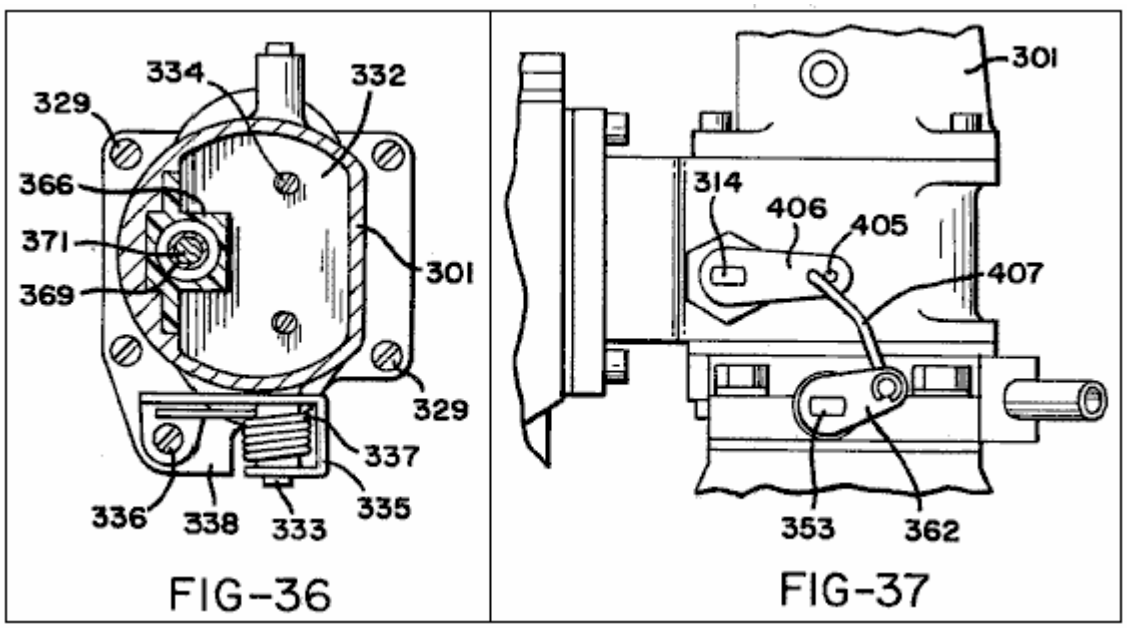


Fig.38 is a cross-section on the line 38--38 in Fig.34;

Fig.39 is a plan view of the lower section of the gas mixing and delivery unit, which is broken away from the upper section along the interface 39--39 of Fig.30;

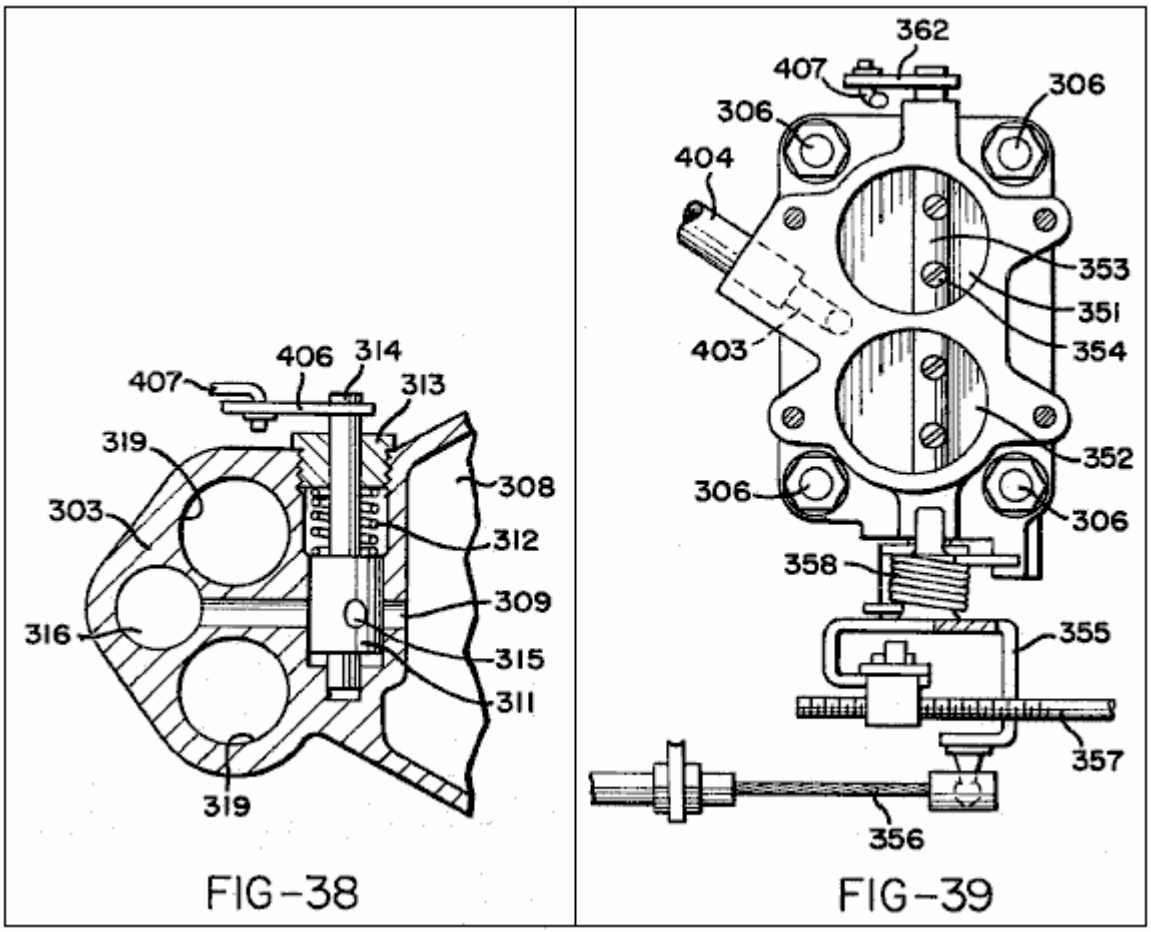


Fig.40 is a cross-section on the line 40--40 in Fig.32; and

Fig.41 is a plan of a lower body part of the gas mixing and delivery unit.

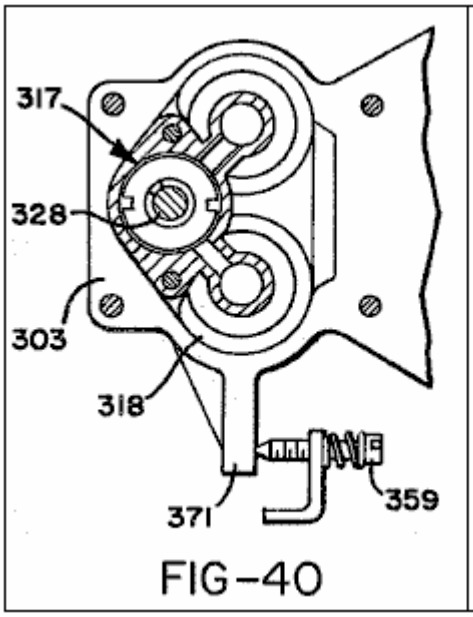


FIG-40

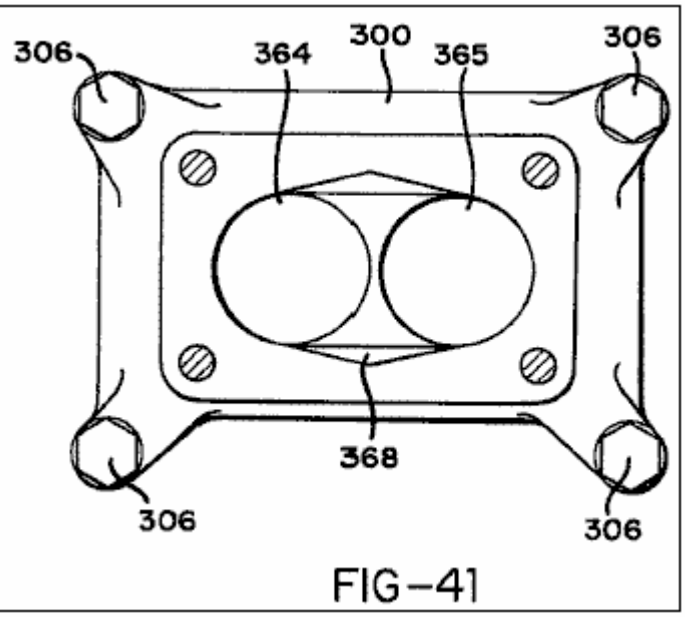


FIG-41

DESCRIPTION OF THE PREFERRED EMBODIMENT

Fig.1 shows an assembly denoted generally as **31** having an engine bay **32** in which an internal combustion engine **33** is mounted behind a radiator **34**. Engine **33** is a conventional engine and, as illustrated, it may have two banks of cylinders in "V" formation. Specifically, it may be a V8 engine. It is generally of conventional construction and **Fig.1** shows the usual cooling fan **34**, fan belt **36** and generator or alternator **37**.

In accordance with the invention the engine does not run on the usual petroleum fuel but is equipped with fuel supply apparatus which supplies it with a mixture of hydrogen and oxygen gases generated as products of a water electrolysis process carried out in the fuel supply apparatus. The major components of the fuel supply apparatus are an electrolytic cell denoted generally as **41** and a gas mixing and delivery unit **38** to mix the hydrogen and oxygen gases generated within the cell **41** and to deliver them to engine **33**. The electrolytic cell **41** receives water through a water delivery line **39** to make up the electrolyte solution within it. It has an anode and a cathode which contact the electrolyte solution, and in operation of the apparatus pulses of electrical energy are applied between the anode and cathode to produce pulses of high current flow through the electrolyte solution. Some of the electrical components necessary to produce the pulses of electrical energy applied between the anode and cathode are carried in a housing **40** mounted on one side of engine bay **32**. The car battery **30** is mounted at the other side of the engine bay.

Before the physical construction of the fuel delivery apparatus is described in detail the general principles of its operation will firstly be described with reference to the electrical circuit diagram of **Fig.2**.

In the illustrated circuit terminals **44, 45, 46** are all connected to the positive terminal of the car battery **30** and terminal **47** is connected to the negative terminal of that battery. Switch **48** is the usual ignition switch of the car and closure of this switch provides current to the coil **49** of a relay **51**. The moving contact **52** of relay **51** receives current at 12 volts from terminal **45**, and when the relay is operated by closure of ignition switch **48** current is supplied through this contact to line **53** so that line **53** may be considered as receiving a positive input and line **54** from terminal **47** may be considered as a common negative for the circuit. Closure of ignition switch **48** also supplies current to one side of the coil **55** of a solenoid **56**. The other side of solenoid coil **55** is earthed by a connection to the car body within the engine bay. As will be explained below solenoid **56** must be energised to open a valve which controls supply of hydrogen and oxygen gases to the engine and the valve closes to cut off that supply as soon as ignition switch **48** is opened.

The function of relay **51** is to connect circuit line **53** directly to the positive terminal of the car battery so that it receives a positive signal directly rather than through the ignition switch and wiring.

The circuit comprises pulse generator circuitry which includes unijunction transistor **Q1** with associated resistors **R1, R2 and R3** and capacitors **C2** and **C3**. This circuitry produces pulses which are used to trigger an NPN silicon power transistor **Q2** which in turn provides via a capacitor **C4** triggering pulses for a thyristor **T1**.

Resistor **R1** and capacitor **C2** are connected in series in a line **57** extending to one of the fixed contacts of a relay **58**. The coil **59** of relay **58** is connected between line **53** and a line **61** which extends from the moving contact of the relay to the common negative line **54** via a normally closed pressure operated switch **62**. The pressure control line **63** of switch **62** is connected in a manner to be described below to a gas collection chamber of electrolytic cell **41** in order to provide a control connection whereby switch **62** is opened when the gas in the collection chamber reaches a certain pressure. However, provided that switch **62** remains closed, relay **58** will operate when ignition switch **48** is closed to provide a connection between lines **57** and **61** thereby to connect capacitor **C2** to the common negative line **54**. The main purpose of relay **58** is to provide a slight delay in this connection between the capacitor **C2** and the common negative line **54** when the circuit is first energised. This will delay the generation of triggering pulses to thyristor **T1** until a required electrical condition has been achieved in the transformer circuitry to be described below. Relay **58** is hermetically sealed and has a balanced armature so that it can operate in any position and can withstand substantial shock or vibration when the car is in use.

When the connection between capacitor **C2** and line **54** is made via relay **58**, unijunction transistor **Q1** will act as an oscillator to provide positive output pulses in line **64** at a pulse rate which is controlled by the ratio of **R1:C1** and at a pulse strength determined by the ratio of **R2:R3**. These pulses will charge the capacitor **C3**. Electrolytic capacitor **C1** is connected directly between the common positive line **53** and the common negative line **54** to filter the circuitry from all static noise.

Resistor **R1** and capacitor **C2** are chosen such that at the input to transistor **Q1** the pulses will be of saw tooth form. This will control the form of the pulses generated in the subsequent circuitry and the saw tooth pulse form is chosen since it is believed that it produces the most satisfactory operation of the pulsing circuitry. It should be stressed, however, that other pulse forms, such as square wave pulses, could be used. Capacitor **C3** discharges

through a resistor **R4** to provide triggering signals for transistor **Q2**. Resistor **R4** is connected to the common negative line **54** to serve as a gate current limiting device for transistor **Q2**.

The triggering signals produced by transistor **Q2** via the network of capacitor **C3** and a resistor **R4** will be in the form of positive pulses of sharply spiked form. The collector of transistor **Q2** is connected to the positive supply line **53** through resistor **R6** while the emitter of that transistor is connected to the common negative line **54** through resistor **R5**. These resistors **R5** and **R6** control the strength of current pulses applied to a capacitor **C4**, which discharges through a resistor **R7** to the common negative line **54**, thereby to apply triggering signals to the gate of thyristor **T1**. The gate of thyristor **T1** receives a negative bias from the common negative line via resistor **R7** which thus serves to prevent triggering of the thyristor by inrush currents.

The triggering pulses applied to the gate of thyristor **T1** will be very sharp spikes occurring at the same frequency as the saw tooth wave form pulses established by unijunction transistor **Q1**. It is preferred that this frequency be of the order of 10,000 pulses per minute and details of specific circuit components which will achieve this result are listed below. Transistor **Q2** serves as an interface between unijunction transistor **Q1** and thyristor **T1**, preventing back flow of emf from the gate of the thyristor which might otherwise interfere with the operation of transistor **Q1**. Because of the high voltages being handled by the thyristor and the high back emf applied to transistor **Q2**, the latter transistor must be mounted on a heat sink.

The cathode of thyristor **T1** is connected via a line **65** to the common negative line **54** and the anode is connected via a line **66** to the centre of the secondary coil **67** of a first stage transformer **TR1**. The two ends of transformer coil **67** are connected via diodes **D1** and **D2** and a line **68** to the common negative line **54** to provide full wave rectification of the transformer output.

First stage transformer **T1** has three primary coils **71, 72, 73** wound together with secondary coil **67** about a core **74**. This transformer may be of conventional half cup construction with a ferrite core. The secondary coil may be wound on to a coil former disposed about the core and primary coils **71** and **73** may be wound in bifilar fashion over the secondary coil. The other primary coil **72** may then be wound over the coils **71, 73**. Primary coils **71** and **73** are connected at one side by a line **75** to the uniform positive potential of circuit line **53** and at their other sides by lines **79, 81** to the collectors of transistors **Q3**, **Q4**. The emitters of transistors **Q3, Q4** are connected permanently via a line **82** to the common negative line **54**. A capacitor **C6** is connected between lines **79, 81** to act as a filter preventing any potential difference between the collectors of transistors **Q3, Q4**.

The two ends of primary coil **72** are connected by lines **83, 84** to the bases of transistors **Q3, Q4**. This coil is centre tapped by a line **85** connected via resistor **R9** to the positive line **53** and via resistor **R10** to the common negative line **54**.

When power is first applied to the circuit transistors **Q3** and **Q4** will be in their non-conducting states and there will be no current in primary coils **71, 73**. However, the positive current in line **53** will provide via resistor **R9** a triggering signal applied to the centre tap of coil **72** and this signal operates to trigger alternate high frequency oscillation of transistors **Q3, Q4** which will result in rapid alternating pulses in primary coils **71, 73**. The triggering signal applied to the centre tap of coil **72** is controlled by the resistor network provided by resistors **R9** and **R10** such that its magnitude is not sufficient to enable it to trigger **Q3** and **Q4** simultaneously but is sufficient to trigger one of those transistors. Therefore only one of the transistors is fired by the initial triggering signal to cause a current to flow through the respective primary coil **71** or **73**. The signal required to hold the transistor in the conducting state is much less than that required to trigger it initially, so that when the transistor becomes conductive some of the signal applied to the centre tap of coil **72** will be diverted to the non-conducting transistor to trigger it. When the second transistor is thus fired to become conductive, current will flow through the other of the primary coils **71, 73**, and since the emitters of the two transistors are directly connected together, the positive output of the second transistor will cause the first-fired transistor to be shut off. When the current drawn by the collector of the second-fired resistor drops, part of the signal on the centre tap of coil **72** is diverted back to the collector of the first transistor which is re-fired. It will be seen that the cycle will then repeat indefinitely so that transistors **Q3, Q4** are alternately fired and shut off in very rapid sequence. Thus current pulses flow in alternate sequence through primary coils **71, 73** at a very high frequency, this frequency being constant and independent of changes in input voltage to the circuit. The rapidly alternating pulses in primary coils **71** and **73**, which will continue for so long as ignition switch **48** remains closed, will generate higher voltage signals at the same frequency in the transformer secondary coil **67**.

A dump capacitor **C5** bridged by a resistor **R8** is connected by a line **86** to the line **66** from the secondary coil of transformer **TR1** and provides the output from that transformer which is fed via line **87** to a second stage transformer **TR2**.

When thyristor **T1** is triggered to become conductive the full charge of dump capacitor **C5** is released to second stage transformer **TR2**. At the same time the first stage of transformer **TR1** ceases to function because of this

momentary short circuit placed across it and consequently thyristor **T1** releases, i.e. becomes non-conductive. This permits charge to be built up again in dump capacitor **C5** for release when the thyristor is next triggered by a signal from transistor **Q2**. Thus during each of the intervals when the thyristor is in its non-conducting state the rapidly alternating pulses in primary coils **71, 73** of transformer **TR1** produced by the continuously oscillating transistors **Q3, Q4** produce, via the transformer coupling, relatively high voltage output pulses which build up a high charge in capacitor **C5**, and this charge is released suddenly when the thyristor is triggered. In a typical apparatus using a 12 volt DC supply battery pulses of the order of 22 amps at 300 volts may be produced in line **87**.

As previously mentioned relay **58** is provided in the circuit to provide a delay in the connection of capacitor **C2** to the common negative line **54**. This delay, although very short, is sufficient to enable transistors **Q3**, **Q4** to start oscillating to cause transformer **TR1** to build up a charge in dumping capacitor **C5** before the first triggering signal is applied to thyristor **T1** to cause discharge of the capacitor.

Transformer **TR2** is a step-down transformer which produces pulses of very high current flow at low voltage. It is built into the anode of electrolytic cell **41** and comprises a primary coil **88** and a secondary coil **89** wound about a core **91**. Secondary coil **89** is formed of heavy wire in order to handle the large current induced in it and its ends are connected directly to the anode **42** and cathode **43** of the electrolytic cell **41** in a manner to be described below.

In a typical apparatus, the output from the first stage transformer **TR1** would be 300 volt pulses of the order of 22 amps at 10,000 pulses per minute and a duty cycle of slightly less than 0.006. This can be achieved from a uniform 12 volt and 40 amps DC supply using the following circuit components:

Components:

R1 2.7 k ohms 1/2 watt 2% resistor
R2 220 ohms 1/2 watt 2% resistor
R3 100 ohms 1/2 watt 2% resistor
R4 22 k ohms 1/2 watt 2% resistor
R5 100 ohms 1/2 watt 2% resistor
R6 220 ohms 1/2 watt 2% resistor
R7 1 k ohms 1/2 watt 2% resistor
R8 10 m ohms 1 watt 5% resistor
R9 100 ohms 5 watt 10% resistor
R10 5.6 ohms 1 watt 5% resistor

C1 2200 mF 16v electrolytic capacitor
C2 2.2 mF 100v 10% capacitor
C3 2.2 mF 100v 10% capacitor
C4 1 mF 100v 10% capacitor
C5 1 mF 1000v ducon paper capacitor 5S10A
C6 0.002 mF 160v capacitor

Q1 2n 2647 PN unijunction transistor
Q2 2N 3055 NPN silicon power transistor
Q3 2n 3055 NPN silicon power transistor
Q4 2n 3055 NPN silicon power transistor
T1 btw 30-800 rm fast turn-off thyristor
D1 a 14 p diode
D2 a 14 p diode

L1 indicator lamp
Sv1 continuously rated solenoid
Rl1 pw5ls hermetically sealed relay
Ps1 p658a-10051 pressure operated micro switch

Tr1 half cup transformer cores 36/22-341
Coil former 4322-021-30390 wound to provide a turns ratio between secondary and primary of 18:1
Secondary coil 67 = 380 turns
Primary coil 71 = 9 turns
Primary coil 73 = 9 turns
Primary coil 72 = 4 turns

The installation of the above circuit components is illustrated in **Fig.3** to **Fig.13**. They are mounted within and on a housing which is denoted generally as **101** and which is fastened to a side wall of the car engine bay **32** via a mounting bracket **102**. Housing **101**, which may be formed as an aluminium casting, has a front wall **103**, top and bottom walls **104**, **105** and side walls **106**, **107**. All of these walls have external cooling fins. The back of housing **101** is closed by a printed circuit board **108** which is held clamped in position by a peripheral frame **109** formed of an insulated plastics material clamped between the circuit board and mounting bracket **102**. An insulating sheet **111** of cork is held between the frame **109** and mounting bracket **102**.

Printed circuit board **108** carries all of the above-listed circuit components except for capacitor **C5** and transistors **Q3** and **Q4**. **Fig.5** illustrates the position in which transistor **Q2** and the coil assembly **112** of transformer **TR1** are mounted on the printed circuit board. Transistor **Q2** must withstand considerable heat generation and it is therefore mounted on a specially designed heat sink **113** clamped to circuit board **108** by clamping screws **114** and nuts **115**. As most clearly illustrated in **Fig.7** and **Fig.8**, heat sink **113** has a flat base plate portion **116** which is generally diamond shaped and a series of rod like cooling fins **117** project to one side of the base plate around its periphery. It has a pair of countersunk holes **118** of the clamping screws and a similar pair of holes **119** to receive the connector pins **121** which connect transistor **Q2** to the printed circuit board. Holes **118, 119** are lined with nylon bushes **122** and a Formica sheet **123** is fitted between the transistor and the heat sink so that the sink is electrically insulated from the transistor.

The coil assembly **112** of transformer **TR1** (See Fig.9) is comprised of a casing **124** which contains transformer coils and the associated core and former and is closed by a plastic closing plate **125**. Plate **125** is held in position by a clamping stud **126** and is fitted with electrical connector pins **127** which are simply pushed through holes in circuit board **108** and are soldered to appropriate copper conductor strips **128** on the outer face of the board.

For clarity the other circuit components mounted on printed circuit board **108** are not illustrated in the drawings. These are standard small size components and the manner in which they may be fitted to the circuit board is entirely conventional.

Capacitor **C5** is mounted within casing **101**. More specifically it is clamped in position between a flange **131** which stands up from the floor **105** of the casing and a clamping pad **132** engaged by a clamping screw **133**, which is mounted in a threaded hole in casing side wall **106** and is set in position by a lock screw **134**. Flange **131** has two holes **135** (See Fig.6) in which the terminal bosses **136** of capacitor **C5** are located. The terminal pins **137** projecting from bosses **136** are connected to the terminal board **108** by wires (not shown) and appropriate connector pins which are extended through holes in the circuit board and soldered to the appropriate conductor strips on the other face of that board.

Transistors **Q3** and **Q4** are mounted on the front wall **103** of casing **101** so that the finned casing serves as an extended heat sink for these two transistors. They are mounted on the casing wall and electrically connected to the printed circuit board in identical fashion and this is illustrated by **Fig.10** which shows the mounting of transistor **Q3**. As shown in that figure the transistor is clamped in position by clamping screws **138** and nuts **139** which also serve to provide electrical connections to the appropriate conductors of the printed circuit board via conductor wires **141**. The third connection from the emitter of the transistor to the common negative conductor of the printed circuit is made by conductor **142**. Screws **130** and conductor **142** extend through three holes in the casing front wall **103** and these holes are lined with electrically insulating nylon bushes **143, 144**. A Formica sheet **145** is sandwiched between casing plate **103** and the transistor which is therefore electrically insulated from the casing. Two washers **146** are placed beneath the ends of conductor wires **141**.

Pressure operated microswitch **52** is mounted on a bracket **147** projecting inwardly from front wall **103** of casing **101** adjacent the top wall **104** of the casing and the pressure sensing unit **148** for this switch is installed in an opening **149** through top wall **104**. As most clearly seen in **Fig.11**, pressure sensing unit **148** is comprised of two generally cylindrical body members **150, 151** between which a flexible diaphragm **152** is clamped to provide a diaphragm chamber **153**. The gas pressure of sensing tube **63** is applied to chamber **153** via a small diameter passage **154** in body member **150** and a larger passage **155** in a cap member **156**. The cap member and body members are fastened together and clamped to the casing top plate **104** by means of clamping screws **157**. Sensing tube **63** is connected to the passage **155** in cap member **156** by a tapered thread connector **158** and the interface between cap member **156** and body member **150** is sealed by an O-ring **159**.

The lower end of body member **151** of pressure sensing unit **148** has an internally screw threaded opening which receives a screw **161** which at its lower end is formed as an externally toothed adjusting wheel **162**. A switch actuating plunger **163** extends through a central bore in adjusting wheel **162** so that it engages at one end flexible diaphragm **152** and at the other end the actuator member **164** of microswitch **62**. The end of plunger **163** which engages the diaphragm has a flange **165** to serve as a pressure pad and a helical compression spring **167** encircles plunger **163** to act between flange **165** and the adjusting wheel **162** to bias the plunger upwardly against the action of the gas pressure acting on diaphragm **152** in chamber **153**. The pressure at which diaphragm **152**

will force plunger **163** down against the action of spring **167** to cause actuation of switch **62** may be varied by rotating screw **161** and the setting of this screw may be held by a setting screw **168** mounted in a threaded hole in the upper part of casing front wall **103** and projecting inwardly to fit between successive teeth of adjusting wheel **162**. After correct setting of screw **161** is achieved set screw **168** will be locked in position by locking screw **169** which is then sealed by a permanent seal **170** to prevent tampering. Microswitch **62** is also electrically connected to the appropriate conductors of the printed circuit board via wires within the housing and connector pins.

Electrical connections are made between the conductors of printed circuit board **108** and the internal wiring of the circuit via a terminal block **150** (**Fig.12**) set in an opening of housing floor **105** by screws **160** and fitted with terminal plates **140**.

The physical construction of electrolytic cell **41** and the second stage transformer **TR2** is illustrated in **Fig.13** to **Fig.29**. The cell comprises an outer casing **171** having a tubular peripheral wall **172** and top and bottom closures **173, 174**. Bottom closure **174** is comprised of a domed cover **175** and an electrically insulated disc **176** which are held to the bottom of peripheral wall **172** by circumferentially spaced clamping studs **177**. Top closure **173** is comprised of a pair of top plates **178, 179** disposed face to face and held by circumferentially spaced clamping studs **181** screwed into tapped holes in the upper end of peripheral wall **172**. The peripheral wall of the casing is provided with cooling fins **180**.

The anode **42** of the cell is of generally tubular formation. It is disposed vertically within the outer casing and is clamped between upper and lower insulators **182, 183**. Upper insulator **182** has a central boss portion **184** and an annular peripheral flange **185** portion the outer rim of which is clamped between upper closure plate **179** and the upper end of peripheral wall **172**. Lower insulator **183** has a central boss portion **186**, an annular flange portion **187** surrounding the boss portion and an outer tubular portion **188** standing up from the outer margin of flange portion **187**. Insulators **182, 183** are moulded from an electrically insulating material which is also alkali resistant. Polytetrafluoroethylene is one suitable material.

When held together by the upper and lower closures, insulators **182, 183** form an enclosure within which anode **42** and the second stage transformer **TR2** are disposed. Anode **42** is of generally tubular formation and it is simply clamped between insulators **182, 183** with its cylindrical inner periphery located on the boss portions **184, 186** of those insulators. It forms a transformer chamber which is closed by the boss portions of the two insulators and which is filled with a suitable transformer oil. O-ring seals **190** are fitted between the central bosses of the insulator plates and the anode to prevent loss of oil from the transformer chamber.

The transformer core **91** is formed as a laminated mild steel bar of square section. It extends vertically between the insulator boss portions **184, 186** and its ends are located within recesses in those boss portions. The primary transformer winding **88** is wound on a first tubular former **401** fitted directly onto core **91** whereas the secondary winding **89** is wound on a second tubular former **402** so as to be spaced outwardly from the primary winding within the oil filled transformer chamber.

The cathode **43** in the form of a longitudinally slotted tube which is embedded in the peripheral wall portion **183**, this being achieved by moulding the insulator around the cathode. The cathode has eight equally spaced longitudinal slots **191** so that it is essentially comprised of eight cathode strips **192** disposed between the slots and connected together at top and bottom only, the slots being filled with the insulating material of insulator **183**.

Both the anode and cathode are made of nickel plated mild steel. The outer periphery of the anode is machined to form eight circumferentially spaced flutes **193** which have arcuate roots meeting at sharp crests or ridges **194** defined between the flutes. The eight anode crests **194** are radially aligned centrally of the cathode strips **192** and the perimeter of the anode measured along its external surface is equal to the combined widths of the cathode strips measured at the internal surfaces of these strips, so that over the major part of their lengths the anode and cathode have equal effective areas. This equalisation of areas generally have not been available in prior art cylindrical anode/cathode arrangements.

As most clearly seen in **Fig.27** the upper end of anode **42** is relieved and fitted with an annular collar **200** the outer periphery of which is shaped to form an extension of the outer peripheral surface of the fluted anode. This collar is formed of an electrically insulated plastics material such as polyvinyl chloride or teflon. A locating pin **205** extends through collar **200** to project upwardly into an opening in upper insulating plate **182** and to extend down into a hole **210** in the cathode. The collar is thus located in correct annular alignment relative to the anode and the anode is correctly aligned relative to the cathode.

The annular space **195** between the anode and cathode serves as the electrolyte solution chamber. Initially this chamber is filled approximately 75% full with an electrolyte solution of 25% potassium hydroxide in distilled water. As the electrolysis reaction progresses hydrogen and oxygen gases collect in the upper part of this chamber and water is admitted to maintain the level of electrolyte solution in the chamber. Insulating collar **200** shields the

cathode in the upper region of the chamber where hydrogen and oxygen gases collect to prevent any possibility of arcing through these gases between the anode and cathode.

Electrolyte chamber **195** is divided by a tubular membrane **196** formed by nylon woven mesh material **408** stretched over a tubular former **197** formed of very thin sheet steel. As most clearly illustrated in **Fig.20** and **Fig.21** former **197** has upper and lower rim portions **198, 199** connected by circumferentially spaced strip portions **201**. The nylon mesh material **408** may be simply folded around the upper and lower insulators **182, 183** so that the former is electrically isolated from all other components of the cell. Material **408** has a mesh size which is so small that the mesh openings will not pass bubbles of greater than 0.004 inch diameter and the material can therefore serve as a barrier against mixing of hydrogen and oxygen generated at the cathode and anode respectively while permitting the electrolytic flow of current between the electrodes. The upper rim portion **198** of the membrane former **197** is deep enough to constitute a solid barrier through the depth of the gas collection chamber above the electrolyte solution level so that there will be no mixing of hydrogen and oxygen within the upper part of the chamber.

Fresh water is admitted into the outer section of chamber **195** via an inlet nozzle **211** formed in upper closure plate **178**. The electrolyte solution passes from the outer to the inner sections of chamber **195** through the mesh membrane **408**.

Nozzle **211** has a flow passage **212** extending to an electrolyte inlet valve **213** controlled by a float **214** in chamber **195**. Valve **213** comprises a bushing **215** mounted within an opening extending down through upper closure plate **179** and the peripheral flange **185** of upper insulator **182** and providing a valve seat which co-operates with valve needle **216**. Needle **216** rests on a pad **217** on the upper end of float **214** so that when the electrolyte solution is at the required level the float lifts the needle hard against the valve seat. The float slides vertically on a pair of square section slide rods **218** extending between the upper and lower insulators **182** and **183**. These rods, which may be formed of polytetrafluoroethylene extend through appropriate holes **107** through the float.

The depth of float **214** is chosen such that the electrolyte solution fills only approximately 75% of the chamber **195**, leaving the upper part of the chamber as a gas space which can accommodate expansion of the generated gas due to heating within the cell.

As electrolysis of the electrolyte solution within chamber **195** proceeds, hydrogen gas is produced at the cathode and oxygen gas is produced at the anode. These gases bubble upwardly into the upper part of chamber **195** where they remain separated in the inner and outer compartments defined by membrane and it should be noted that the electrolyte solution enters that part of the chamber which is filled with oxygen rather than hydrogen so there is no chance of leakage of hydrogen back through the electrolyte inlet nozzle.

The abutting faces of upper closure plates **178, 179** have matching annular grooves forming within the upper closure inner and outer gas collection passages **221, 222**. Outer passage **222** is circular and it communicates with the hydrogen compartment of chamber **195** via eight ports **223** extending down through top closure plate **179** and the peripheral flange of upper insulator **182** adjacent the cathode strips **192**. Hydrogen gas flows upwardly through ports **223** into passage **222** and thence upwardly through a one-way valve **224** (**Fig.19**) into a reservoir **225** provided by a plastic housing **226** bolted to top closure plate **178** via a centre stud **229** and sealed by a gasket **227**. The lower part of housing **114** is charged with water. Stud **229** is hollow and its lower end has a transverse port **228** so that, on removal of a sealing cap **229** from its upper end it can be used as a filter down which to pour water into the reservoir **225**. Cap **229** fits over a nut **231** which provides the clamping action on plastic housing **226** and resilient gaskets **232, 233** and **234** are fitted between the nut and cover, between the cap and the nut and between the cap and the upper end of stud **229**.

One-way valve **224** comprises a bushing **236** which projects down into the annular hydrogen passage **221** and has a valve head member **237** screw fitted to its upper end to provide clamping action on top closure plate **178** between the head member and a flange **238** at the bottom end bushing **236**. Bushing **236** has a central bore **239**, the upper end of which receives the diamond cross-section stem of a valve member **240**, which also comprises a valve plate portion **242** biased against the upper end of the bushing by compression spring **243**. Valve member **240** is lifted against the action of spring **243** by the pressure of hydrogen gas within passage **221** to allow the gas to pass into the interior of valve head **237** and then out through ports **220** in that member into reservoir **225**.

Hydrogen is withdrawn from reservoir **225** via a stainless steel crooked tube **241** which connects with a passage **409**. Passage **409** extends to a port **250** which extends down through the top and bottom closure plates **178, 179** and top insulator **182** into a hydrogen duct **244** extending vertically within the casting of casing **171**. Duct **244** is of triangular cross-section. As will be explained below, the hydrogen passes from this duct into a mixing chamber defined in the gas mixing and delivery unit **38** which is bolted to casing **171**.

Oxygen is withdrawn from chamber **195** via the inner annular passage **221** in the top closure. Passage **221** is not circular but has a scalloped configuration to extend around the water inlet. Oxygen enters it through eight ports **245** extended through top closure plate **179** and the annular flange portion of upper insulator **182**. The oxygen flows upwardly from passage **222** through a one-way valve **246** and into a reservoir **260** provided by a plastic housing **247**. The arrangement is similar to that for withdrawal of hydrogen and will not be described in great detail. Suffice to say that the bottom of the chamber is charged with water and the oxygen is withdrawn through a crooked tube **248**, an outlet passage **249** in top closure plate **178**, and a port which extends down through closure plates **178, 179** and top insulator **182** into a triangular cross-section oxygen duct **251** extending vertically within casing **171** disposed opposite hydrogen duct **244**. The oxygen is also delivered to the gas mixing chamber of the mixing and delivery unit **38**.

The pressure sensing tube **63** for switch **62** is connected via a tapered thread connector **410** and a passage **411** in the top closure plate **178** directly to the annular hydrogen passage **222**. If the pressure within the passage rises above a predetermined level, switch **62** is operated to disconnect capacitor **C2** from the common negative line **54**. This removes the negative signal from capacitor **C2** which is necessary to maintain continuous operation of the pulse generating circuitry for generating the triggering pulses on thyristor **T1** and these triggering pulses therefore cease. The transformer **TR1** continues to remain in operation to charge dumping capacitor **C5** but because thyristor **T1** cannot be triggered dumping capacitor **C5** will simply remain charged until the hydrogen pressure in passage **222**, and therefore in chamber **195** falls below the predetermined level and triggering pulses are applied once more to thyristor **T1**. Pressure actuated switch **62** thus controls the rate of gas production according to the rate at which it is withdrawn. The stiffness of the control springs for gas escape valves **224, 246** must of course be chosen to allow escape of the hydrogen and oxygen in the proportions in which they are produced by electrolysis, i.e. in the ratios 2:1 by volume.

Reservoirs **225, 260** are provided as a safety precaution. If a sudden back-pressure were developed in the delivery pipes this could only shatter the plastic housings **226, 247** and could not be transmitted back into the electrolytic cell. Switch **62** would then operate to stop further generation of gases within the cell.

The electrical connections of secondary transformer coil **89** to the anode and the cathode are shown in **Fig.14**. One end of coil **89** is extended as a wire **252** which extends into a blind hole in the inner face of the anode where it is gripped by a grub screw **253** screwed into a threaded hole extended vertically into the anode underneath collar **200**. A tapered nylon plug **254** is fitted above screw **253** to seal against loss of oil from the interior of the anode. The other end of coil **89** is extended as a wire **255** to pass down through a brass bush **256** in the bottom insulator **183** and then horizontally to leave casing **171** between bottom insulating disc **176** and insulator **183**.

As most clearly shown in **Fig.23**, brass bush **256** has a head flange **257** and is fitted at its lower end with a nut **258** whereby it is firmly clamped in position. Gaskets **259, 261** are disposed beneath head flange **257** and above nut **258** respectively.

At the location where wire **255** is extended horizontally to leave the casing the upper face of disc **176** and the lower face of insulator **183** are grooved to receive and clamp onto the wire. Disc **176** and insulator **183** are also extended radially outwardly at this location to form tabs which extend out beneath casing **171** and ensure proper insulation of the wire through to the outer periphery of the casing.

Outside the casing, wire **255** is connected to a cathode terminal bolt **262**. Terminal bolt **262** has a head which is received in a socket in separate head piece **263** shaped to suit the cylindrically curved inner periphery of the cathode and nickel plated to resist chemical attack by the electrolyte solution. The stem of the terminal bolt extends through openings in the cathode and peripheral wall portion **188** of insulator **183** and air insulating bush fitted in an aligned opening in the casing wall **172**. The head piece **263** of the terminal bolt is drawn against the inner periphery of the cathode by tightening of a clamping nut **265** and the end of wire **255** has an eye which is clamped between nut **265** and a washer **266** by tightening a terminal end nut **267**. A washer **268** is provided between nut **265** and brush **264** and a sealing O-ring **269** is fitted in an annular groove in the bolt stem to engage the inner periphery of the bush in order to prevent escape of electrolyte solution. The terminal connection is covered by a cover plate **271** held in place by fixing screws **272**.

The two ends of the primary transformer coil **88** are connected to strip conductors **273, 274** which extend upwardly through the central portion of upper insulator **183**. The upper ends of conductors **273, 274** project upwardly as pins within a socket **275** formed in the top of upper insulator **183**. The top of socket **275** is closed by a cover **276** which is held by a centre stud **277** and through which wires **278, 279** from the external circuit are extended and connected to conductors **273, 274** by push-on connectors **281, 282**.

The transformer connections shown in **Fig.14** are in accordance with the circuit of **Fig.2**, i.e. the ends of secondary coil **89** are connected directly between the anode and the cathode. Transformer **TR2** is a step-down

transformer and, assuming an input of pulses of 22 amps at 300 volts and a coil ratio between the primary and secondary of 10:1 the output applied between the anode and the cathode will be pulses of 200 amps at a low voltage of the order of 3 volts. The voltage is well in excess of that required for electrolysis to proceed and the very high current achieved produces a high rate of yield of hydrogen and oxygen. The rapid discharge of energy which produces the large current flow will be accompanied by a release of heat. This energy is not entirely lost in that the consequent heating of the electrolyte solution increases the mobility of the ions which tends to increase the rate of electrolysis.

The configuration of the anode and cathode arrangement of electrolytic cell **41** is of significant importance. The fluted external periphery of the anode causes a concentration of current flow which produces a better gas yield over a given electrode area. This particular configuration also causes the surface area of the anode to be extended and permits an arrangement in which the anode and cathode have equal surface areas which is most desirable in order to minimise electrical losses. It is also desirable that the anode and cathode surfaces at which gas is produced be roughened, for example by sand-blasting. This promotes separation of the gas bubbles from the electrode surfaces and avoids the possibility of overvoltages.

The arrangement of the secondary transformer in which the central anode is surrounded by the cathode is also of great importance. The anode, being constructed of a magnetic material, is acted on by the magnetic field of transformer **TR2** to become, during the period of energisation of that transformer, a strong conductor of magnetic flux. This in turn creates a strong magnetic field in the inter-electrode space between the anode and the cathode. It is believed that this magnetic field increases the mobility of the ions in solution thereby improving the efficiency of the cell.

The heat generated by transformer **TR2** is conducted via the anode to the electrolyte solution and increases the mobility of the ions within the electrolyte solution as above mentioned. The cooling fins **180** are provided on casing **171** to assist in dissipation of excess generated heat. The location of the transformer within the anode also enables the connections of the secondary coil **89** to the anode and cathode to be made of short, well protected conductors.

As mentioned above the hydrogen and oxygen gas generated in electrolytic cell **41** and collected in ducts **244, 251** is delivered to a gas mixing chamber of the mixing and delivery unit **38**. More specifically, these gases are delivered from ducts **244, 251** via escape valves **283, 284** (**Fig.15**) which are held in position over discharge ports **285, 286** from the ducts by means of a leaf spring **287**. The outer ends of spring **287** engage the valves **283, 284** and the centre part of the spring is bowed inwardly by a clamping stud **288** screwed into a tapped hole in a boss **289** formed in the cell casing **171**.

Valve **283** is detailed in **Fig.28** and **Fig.29** and valve **284** is of identical construction. Valve **283** includes an inner valve body **291** having a cap portion **292** and an annular end ring portion **293** which holds an annular valve seat **294**. A valve disc **295** is biased against the valve seat by a valve spring **296** reacting against the cap portion **292**. An outer valve cover **297** fits around the inner member **291** and is engaged by spring **287** to force the inner member firmly into a socket in the wall of the cell casing so to cover the hydrogen discharge port **285**. The end ring portion **293** of the inner body member beds on a gasket **298** within the socket.

During normal operation of the apparatus valves **283, 284** act as simple one-way valves by movements of their spring loaded valve plates. However, if an excessive gas pressure should arise within the electrolytic cell these valves will be forced back against the action of holding spring **287** to provide pressure relief. The escaping excess gas then flows to atmosphere via the mixing and delivery unit **38** as described below. The pressure at which valves **283, 284** will lift away to provide pressure relief may be adjusted by appropriate setting of stud **288**, which setting is held by a nut **299**.

The construction of the gas mixing and delivery unit **38** is shown in **Fig.30** and **Fig.40**. It comprises an upper body portion **301** which carries an air filter assembly **302**, an intermediate body portion **303**, which is bolted to the casing of electrolytic cell **41** by six studs **304**, and successive lower body portions **305, 300**, the latter of which is bolted to the inlet manifold of the engine by four studs **306**.

The bolted connection between intermediate body portion **303** and the casing of the electrolytic cell is sealed by a gasket **307**. This connection surrounds valves **283, 284** which deliver hydrogen and oxygen gases directly into a mixing chamber **308** (**Fig.34**) defined by body portion **303**. The gases are allowed to mix together within this chamber and the resulting hydrogen and oxygen mixture passes along small diameter horizontal passageway **309** within body portion **303** which passageway is traversed by a rotary valve member **311**. Valve member **311** is conically tapered and is held within a correspondingly tapered valve housing by a spring **312** (**Fig.38**) reacting against a bush **313** which is screwed into body portion **303** and serves as a mounting for the rotary valve stem **314**. Valve member **311** has a diametral valve port **315** and can be rotated to vary the extent to which this port is

aligned with passageway **309** thereby to vary the effective cross-section for flow through that passageway. As will be explained below, the rotational positions of the valve member is controlled in relation to the engine speed.

Passage **309** extends to the lower end of a larger diameter vertical passageway **316** which extends upwardly to a solenoid freed valve **310** incorporated in a valve and jet assembly denoted generally as **317**.

Assembly **317** comprises a main body **321** (**Fig.32**) closed at the top by a cap **322** when the assembly is clamped to body portion **303** by two clamping studs **323** to form a gas chamber **324** from which gas is to be drawn through jet nozzles **318** into two vertical bores or throats **319** (**Fig.31**) in body portion **303**. The underside of body **321** has a tapped opening into which is fitted an externally screw threaded valve seat **325** of valve **310**. A valve member **326** is biased down against seat **325** by a spring **327** which reacts against cap **322**. Spring **327** encircles a cylindrical stem **328** of valve member **326** which stem projects upwardly through an opening in cap **322** so that it may be acted on by solenoid **56** which is mounted immediately above the valve in upper body portion **301**.

Solenoid **56** is comprised of an outer insulating casing **366** which has two mounting flanges **367**. This casing houses the copper windings constituting coil **55**. These are wound on a plastic bobbin **369** disposed about a central mild steel core **371**. The core has a bottom flange **372** and the bobbin and coils are held clamped in the casing through insulating closure **373** acted on by flange **372** on tightening of a clamping nut **374** which is fitted to the other end of the core.

Upper body portion **301** of unit **38** is tubular but at one side it has an internal face shaped to suit the exterior profile of solenoid casing **366** and mounting flanges **367**. Two mounting screws **375** screw into holes in this face and engage slots **376** in the mounting flanges **367** so that the height of the solenoid above valve **310** can be adjusted. The two terminals **377** are connected into the electrical circuit by wires (not shown) which may be extended into unit **38** via the air filter assembly.

When solenoid **56** is energised its magnetised core attracts valve stem **328** and valve member **326** is lifted until stem **328** abuts the lower flange **372** of the solenoid core. Thus valve **310** is opened when the ignition switch is closed and will close under the influence of spring **327** when the ignition switch is opened. Vertical adjustment of the solenoid position controls the lift of valve member **326** and therefore the maximum fuel flow rate through unit **38**.

Electrolyte cell **41** produces hydrogen in the ratio 2:1 to provide a mixture which is by itself completely combustible. However, as used in connection with existing internal combustion engines the volume of hydrogen and oxygen required for normal operation is less than that of a normal fuel air mixture. Thus a direct application to such an engine of only hydrogen and oxygen in the amount required to meet power demands will result in a vacuum condition within the system. In order to overcome this vacuum condition provision is made to draw make-up air into throats **319** via the air filter assembly **302** and upper body portion **301**.

Upper body portion **301** has a single interior passage **328** through which make-up air is delivered to the dual throats **319**. It is fastened to body portion **303** by clamping studs **329** and a gasket **331** is sandwiched between the two body portions. The amount of make-up air admitted is controlled by an air valve flap **332** disposed across passage **328** and rotatably mounted on a shaft **333** to which it is attached by screws **334**. The valve flap is notched to fit around solenoid casing **366**. Shaft **333** extends through the wall of body portion **301** and outside that wall it is fitted with a bracket **335** which carries an adjustable setting screw **336** and a biasing spring **337**. Spring **337** provides a rotational bias on shaft **333** and during normal running of the engine it simply holds flap **332** in a position determined by engagement of setting screw **336** with a flange **338** of body portion **301**. This position is one in which the flap almost completely closes passage **328** to allow only a small amount of make-up air to enter, this small amount being adjustable by appropriate setting of screw **336**. Screw **336** is fitted with a spring **339** so that it will hold its setting.

Although flaps **332** normally serve only to adjust the amount of make-up air admitted to unit **38**, it also serves as a pressure relief valve if excessive pressures are built up, either due to excessive generation of hydrogen and oxygen gases or due to burning of gases in the inlet manifold of the engine. In either event the gas pressure applied to flaps **332** will cause it to rotate so as to open passage **328** and allow gases to escape back through the air filter. It will be seen in **Fig.32** that flap mounting shaft **333** is offset from the centre of passage **328** such that internal pressure will tend to open the flap and thus exactly the reverse of the air valve in a conventional gasoline carburettor.

Air filter assembly **302** comprises an annular bottom pan **341** which fits snugly onto the top of upper body portion **301** and domed filter element **342** held between an inner frame **343** and an outer steel mesh covering **344**. The assembly is held in position by a wire and eyebolt fitting **345** and clamping nut **346**.

Body portion **305** of unit **38** (**Fig.31**), which is fastened to body portion **303** by clamping studs **347**, carries throttle

valve apparatus to control engine speed. It has two vertical bores **348, 349** serving as continuations of the dual throats which started in body portion **303** and these are fitted with throttle valve flaps **351, 352** fixed to a common throttle valve shaft **353** by fixing screws **354**. Both ends of shaft **353** are extended through the wall of body portion **305** to project outwardly therefrom. One end of this shaft is fitted with a bracket **355** via which it is connected as in a conventional carburettor to a throttle cable **356** and also to an automatic transmission kick-down control linkage **357**. A biasing spring **358** acts on shaft **353** to bias throttle flaps toward closed positions as determined by engagement of a setting screw **359** carried by bracket **355** with a plate **361** projecting from body portion **303**.

The other end of throttle valve shaft **353** carries a lever **362** the outer end of which is connected to a wire link **407** by means of which a control connection is made to the valve stem **314** of valve member **311** via a further lever **406** connected to the outer end of the valve stem. This control connection is such that valve member **311** is at all times positioned to pass a quantity of gas mixture appropriate to the engine speed as determined by the throttle setting. The initial setting of valve member **311** can be adjusted by selection between two connection holes **405** in lever **406** and by bending of link **407**.

Body portion **303** is fastened to the bottom body portion **300** of unit **38** by four clamping studs **306**. The bottom body portion has two holes **364, 365** which form continuations of the dual throats and which diverge in the downward direction so as to direct the hydrogen, oxygen and air mixture delivered through these throats outwardly toward the two banks of cylinder inlets. Since this fuel is dry, a small quantity of oil vapour is added to it via a passage **403** in body portion **305** to provide some upper cylinder lubrication. Passage **403** receives oil vapour through a tube **404** connected to a tapping on the engine tapped cover. It discharges the oil vapour down on to a relieved top face part **368** of body portion **300** between holes **364, 365**. The vapour impinges on the relieved face part and is deflected into the two holes to be drawn with the gases into the engine.

In the illustrated gas mixing and delivery unit **38**, it will be seen that passageway **309**, vertical passageway **316**, chamber **324** and nozzles **318** constitute transfer passage means via which the hydrogen mixture pass to the gas flow duct means comprised of the dual throats via which it passes to the engine. The transfer passage means has a gas metering valve comprised of the valve member **311** and the solenoid operated valve is disposed in the transfer passage means between the metering valve and the gas flow duct means. The gas metering valve is set to give maximum flow rate through the transfer passage means at full throttle setting of throttle flaps **351, 352**. The solenoid operated valve acts as an on/off valve so that when the ignition switch is opened the supply of gas to the engine is positively cut-off thereby preventing any possibility of spontaneous combustion in the cylinders causing the engine to "run on". It also acts to trap gas in the electrolytic cell and within the mixing chamber of the mixing and delivery unit so that gas will be available immediately on restarting the engine.

Dumping capacitor **C5** will determine a ratio of charging time to discharge time which will be largely independent of the pulse rate and the pulse rate determined by the oscillation transistor **Q1** must be chosen so that the discharge time is not so long as to produce overheating of the transformer coils and more particularly the secondary coil 89 of transformer **TR2**. Experiments indicate that overheating problems are encountered at pulse rates below about 5,000 and that the system will behave much like a DC system, with consequently reduced performance at pulse rates greater than about 40,000. A pulse rate of about 10,000 pulses per minute will be nearly optimum. With the saw tooth wave input and sharply spiked output pulses of the preferred oscillator circuit the duty cycle of the pulses produced at a frequency of 10,000 pulses per minute was about 0.006. This pulse form helps to minimise overheating problems in the components of the oscillator circuit at the high pulse rates involved. A duty cycle of up to 0.1, as may result from a square wave input, would be feasible but at a pulse rate of 10,000 pulses per minute some of the components of the oscillator circuit would then be required to withstand unusually high heat inputs. A duty cycle of about 0.005 would be a minimum which could be obtained with the illustrated type of oscillator circuitry.

From the foregoing description it can be seen that the electrolytic cell **41** converts water to hydrogen and oxygen whenever ignition switch **44** is closed to activate solenoid **51**, and this hydrogen and oxygen are mixed in chamber **308**. Closure of the ignition switch also activates solenoid **56** to permit entry of the hydrogen and oxygen mixture into chamber **319**, when it mixes with air admitted into the chamber by air valve flap **332**. As described above, air valve flap **332** may be set to admit air in an amount as required to avoid a vacuum condition in the engine.

In operation the throttle cable **356** causes bracket **355** to pivot about throttle valve shaft **353**, which rotates flap **351** to control the amount of hydrogen-oxygen-air mixture entering the engine. At the same time shaft **353** acts via the linkage shown in **Fig.37** to control the position of shaft **314**, and shaft **314** adjusts the amount of hydrogen-oxygen mixture provided for mixing with the air. As shown in **Fig.30**, bracket **355** may also be linked to a shaft **357**, which is connected to the car transmission. Shaft **357** is a common type of shaft used for down shifting into a passing gear when the throttle has been advanced beyond a predetermined point. Thus there is provided a

compact fuel generation system which is compatible with existing internal combustion engines and which has been designed to fit into a standard passenger car.

While the form of apparatus herein described constitutes a preferred embodiment of the invention, it is to be understood that the invention is not limited to this precise form of apparatus, and that changes may be made therein without departing from the scope of the invention.

CLAIMS

1. For an internal combustion engine having inlet means to receive a combustible fuel, fuel supply apparatus comprising:

a vessel to hold an aqueous electrolyte solution;

an anode and a cathode to contact the electrolyte solution within the vessel;

electrical supply means to apply between said anode and said cathode pulses of electrical energy to induce a pulsating current in the electrolyte solution thereby to generate by electrolysis hydrogen and oxygen gases;

gas collection and delivery means to collect the hydrogen and oxygen gases and to direct them to the engine inlet means; and

water admission means to admit water to said vessel;

said electrical supply means comprising a source of direct current electrical energy of substantially uniform voltage and current and electrical converter means to convert that energy to said pulses, said converter means comprising a transformer means having primary coil means energised by direct current energy from said source and secondary coil means inductively coupled to the primary coil means; a dump capacitor connected to the secondary coil means of the transformer means so as to be charged by electrical output of that coil means; oscillator means to derive electrical pulses from direct current energy of said source; a switching device switchable from a non-conducting state to a conducting state in response to each of the electrical pulses derived by the oscillator means and connected to the secondary coil means of the transformer means and the dump capacitor such that each switching from its non-conducting state to its conducting state causes the dump capacitor to discharge and also short circuits the transformer means to cause the switching means to revert to its non-conducting state; and electrical conversion means to receive the pulse discharges from the dump capacitor and to convert them to said pulses of electrical energy which are applied between the anode and cathode.

2. Fuel supply as claimed in claim 1, wherein the electrical supply means applies said pulses of electrical energy at a frequency of ranging between about 5,000 and 40,000 pulses per minute.

3. Fuel supply apparatus as claimed in claim 2, wherein the electrical supply means applies said pulses of electrical energy at a frequency of about 10,000 pulses per minute.

4. Fuel supply apparatus as claimed in claim 2, wherein the electrical supply means comprises a source of direct current electrical energy of substantially uniform voltage and current and electrical converter means to convert that energy to said pulses.

5. Fuel supply apparatus as claimed in claim 1, wherein the electrical conversion means is a voltage step-down transformer comprising a primary coil to receive the pulse discharge from said dump capacitor and a secondary coil electrically connected between the anode and cathode and inductively coupled to the primary coil.

6. Fuel supply apparatus as claimed in claim 5, wherein said cathode encompasses the anode.

7. Fuel supply apparatus as claimed in claim 1, wherein the cathode encompasses the anode which is hollow and the primary and secondary coils of the second transformer means are disposed within the anode.

8. Fuel supply apparatus as claimed in claim 1, wherein the anode is tubular and its ends are closed to form a chamber which contains the primary and secondary coils of the second transformer means and which is charged with oil.

9. In combination with an internal combustion engine having an inlet for combustible fuel, fuel supply apparatus comprising:

a. an electrolytic cell to hold an electrolytic conductor;

b. a first hollow cylindrical electrode disposed within said cell and provided about its outer surface with a series of circumferentially spaced and longitudinally extending flutes;

c. a second hollow cylindrical electrode surrounding said anode and segmented into a series of electrically connected longitudinally extending strip; said strips being equal in number to the number of said flutes, said strips having a total active surface area approximately equal to the total active surface area of said flutes, and said strips being in radial alignment with the crests of said flutes;

d. current generating means for generating a flow of electrolysing current between said first and second electrodes;

e. gas collection and delivery means to collect hydrogen and oxygen gases from the cell and to direct them to said fuel inlet of the engine; and

f. water admission means to admit water to the cell.

10. The combination claimed in claim 9, wherein said current generating means comprises a transformer situated inside said first electrode.

11. The combination claimed in claim 10, wherein the secondary winding of said transformer is connected whereby said first electrode operates as an anode and said second electrode operates as a cathode.

12. The combination claimed in claim 11, wherein said current generating means further comprising means to generate a pulsed current in the primary winding of said transformer.

13. The combination claimed in claim 9, wherein the roots of said flutes are cylindrically curved.

14. The combination claimed in claim 10, wherein said current generating means comprises a source of direct current; a transformer means having primary coil means energised by direct current energy from said source and secondary coil means inductively coupled to the primary coil means; a dump capacitor connected to the secondary coil means of the transformer means so as to be charged by electrical output of that coil means; oscillator means to derive electrical pulses from direct current energy of said source, a switching device switchable from a non-conducting state to a conducting state in response to each of the electrical pulses derived by the oscillator means and connected to the secondary coil means of the transformer means and the dump capacitor such that each switching from its non-conducting state to its conducting state causes the dump capacitor to discharge and also short circuits the transformer means to cause the switching means to revert to its non-conducting state; and electrical conversion means to receive the pulse discharges from the dump capacitor and to convert them to said pulses of electrical electrical which are applied between said first and second electrodes.

15. The combination claimed in claim 10, wherein the electrical conversion means comprises a voltage step-down transformer having a primary coil to receive the pulse discharge from said dump capacitor and a secondary coil electrically connected between said first and second electrodes.

16. The combination of an internal combustion engine having an inlet to receive a combustible fuel and fuel supply apparatus comprising:

a vessel to hold an aqueous electrolyte solution;

a first hollow cylindrical electrode disposed within said vessel and provided about its outer surface with a series of circumferentially spaced and longitudinally extending flutes;

a second hollow cylindrical electrode surrounding the first electrode and segmented into a series of electrically connected longitudinally extending strips; said strips being equal in number to the number of said flutes and being in radial alignment with the crests of said flutes;

current generating means for generating a pulsating current between said first and second electrodes to produce hydrogen and oxygen gases within the vessel;

gas collection and delivery means to collect the hydrogen and oxygen gases and to direct them to the engine inlet means; and

water admission means to admit water to the vessel.

17. The combination claimed in claim 26, wherein said current generating means comprises a source of direct current; a first transformer means having primary coil means energised by direct current energy from said source and secondary coil means inductively coupled to the primary coil means; a dump capacitor connected to the secondary coil means of the first transformer means so as to be charged by electrical output of that coil means; oscillator means to derive electrical pulses from direct current energy of said source; a switching device switchable from non-conducting state to a conducting state in response to each of the electrical pulses derived by the oscillator means and connected to the secondary coil means of the first transformer means and the dump capacitor such that each switching from its non-conducting state to its conducting state causes the dump capacitor to discharge and also short circuits the first transformer means to cause a second transformer to receive the pulse discharges from the dump capacitor and to transform them to pulses of electrical energy which are applied between said first and second electrodes.

18. The combination claimed in claim 26, wherein the second transformer means has primary coil means energised by the pulse discharges from the dump capacitor and secondary coil means which is inductively coupled to the primary coil means and is connected to the first and second electrodes such that the first electrode operates as an anode and the second electrode operates as a cathode.

The Water Fracture Cell of Christopher Eccles

UK Patent App. 2,324,307 21st October 1998 Inventor: Christopher R. Eccles

FRACTURE CELL APPARATUS

Please note that this is a re-worded extract from the patent and the diagrams have been adapted slightly. It describes a device for splitting water into hydrogen and oxygen gasses via electrolysis using electrodes which are placed on the **outside** of the cell.

ABSTRACT

Fracture cell apparatus including a capacitive fracture cell **20** comprising a container **21** having walls **21a**, and **21b** made of non-electrically conducting material for containing a liquid dielectric **26**, and spaced apart electrodes **22** and **23** positioned outside container **21** with the liquid dielectric **26** between the electrodes, and a mechanism (**8a** and **8b** in **Fig.1** and **Fig.2**) for applying positive and negative voltage pulses to each of the electrods **22** and **23**. In use, whenever one of a positive voltage pulse and a negative voltage pulse is applied to one of the two electrodes, the other of a positive voltage pulse and a negative voltage pulse is applied to the other of the two electrodes, thereby creating an alternating electric field across the liquid dielectric to cause fracture of the liquid dielectric **26**. The apparatus may be used for generating hydrogen gas.

FRACTURE CELL APPARATUS

This invention relates to a fracture cell apparatus and to a method of generating fuel gas from such fracture cell apparatus. In particular, but not exclusively, the invention relates to an apparatus and method for providing fuel gas from water.

Conventionally, the principal methods of splitting a molecular species into its component atomic constituents have been either purely chemical or purely electrolytic:

Purely chemical reactions always involve "third-party" reagents and do not involve the interaction of(l) an applied external electrical influence, and (2) a simple substance. Conventional electrolysis involves the passage of an electric current through a medium (the electrolyte), such current being the product of ion-transits between the electrodes of the cell. When ions are attracted towards either the cathode or the anode of a conventional electrolytic cell, they either receive or donate electrons on contact with the respective electrode. Such electron exchanges constitute the current during electrolysis. It is not possible to effect conventional electrolysis to any useful degree without the passage of this current; it is a feature of the process.

A number of devices have recently been described which purport to effect "fracture" of, particularly, water by means of resonant electrostatic phenomena. In particular one known device and process for producing oxygen and hydrogen from water is disclosed in US-A-4936961. In this known device a so-called fuel cell water "capacitor" is provided in which two concentrically arranged spaced apart "capacitor" plates are positioned in a container of water, the water contacting, and serving as the dielectric between, the "capacitor" plates. The "capacitor" is in effect a charge-dependent resistor which begins to conduct after a small displacement current begins to flow. The ”capacitor" forms part of a resonant charging circuit that includes an inductance in series with the "capacitor". The "capacitor" is subjected to a pulsating, unipolar electric charging voltage which subjects the water molecules within the "capacitor" to a pulsating electric field between the capacitor plates. The "capacitor" remains charged during the application of the pulsating charging voltage causing the covalent electrical bonding of the hydrogen and oxygen atoms within the water molecules to become destabilised, resulting in hydrogen and oxygen atoms being liberated from the molecules as elemental gases.

Such known fracture devices have, hitherto, always featured, as part of their characteristics, the physical contact of a set of electrodes with the water, or other medium to be fractured. The primary method for limiting current flow through the cell is the provision of a high impedance power supply network, and the heavy reliance on the time-domain performance of the ions within the water (or other medium), the applied voltage being effectively "switched off" in each cycle before ion-transit can occur to any significant degree.

In use of such a known system, there is obviously an upper limit to the number of ion-migrations, electron captures, and consequent molecule-to-atom disruptions which can occur during any given momentary application

of an external voltage. In order to perform effectively, such devices require sophisticated current-limiting and very precise switching mechanisms.

A common characteristic of all such known fracture devices described above, which causes them to behave as though they were conventional electrolysis cells at some point in time after the application of the external voltage, is that they have electrodes in actual contact with the water or other medium.

The present invention seeks to provide an alternative method of producing fracture of certain simple molecular species, for example water.

According to one aspect of the present invention there is provided a fracture cell apparatus including a capacitive fracture cell comprising a container having walls made of non-electrically conducting material for containing a liquid dielectric, and spaced apart electrodes positioned outside the container with the liquid dielectric between the electrodes, and a mechanism for applying positive and negative voltage pulses to each of the electrodes so that, whenever one of a positive voltage pulse and a negative voltage pulse is applied to one of the two electrodes, the other voltage pulse is applied to the other electrode, thereby creating an alternating electric field across the liquid dielectric to cause fracture of the liquid dielectric.

In the apparatus of this invention, the electrodes do not contact the liquid dielectric which is to be fractured or disrupted. The liquid to be fractured is the simple dielectric of a capacitor. No purely ohmic element of conductance exists within the fracture cell and, in use, no current flows due to an ion-carrier mechanism within the cell. The required fracture or disruption of the liquid dielectric is effected by the applied electric field whilst only a simple displacement current occurs within the cell.

Preferably the liquid dielectric comprises water, e.g. distilled water, tap water or deuterated water.

Conveniently each electrode comprises a bipolar electrode.

The mechanism for alternately applying positive and negative pulses, provides step voltages alternately to the two electrodes with a short period of time during each charge voltage cycle in which no step voltage is applied to either electrode. Typically, step voltages in excess of 15 kV, typically about 25 kV, on either side of a reference potential, e.g. earth, are applied to the electrodes. In effect, trains of pulses having alternating positive and negative values are applied to the electrodes, the pulses applied to the different electrodes being "phase shifted". In the case where each electrode comprises a bipolar electrode, each bipolar electrode comprising first and second electrode "plates" electrically insulated from each other, a train of positive pulses is arranged to be applied to one electrode plate of each bipolar electrode and a train of negative pulses is arranged to be applied to the other electrode plate of each bipolar electrode. One electrode plate of one bipolar electrode forms a first set with one electrode plate of the other bipolar electrode and the other electrode plate of the one bipolar electrode forms a second set with the other electrode plate of the other bipolar electrode. For each set, a positive pulse is applied to one electrode plate and a negative pulse is applied simultaneously to the other electrode plate. By alternately switching the application of positive and negative pulses from one to the other set of electrode plates, an "alternating" electric field is generated across the dielectric material contained in the container. The pulse trains are synchronised so that there is a short time interval between the removal of pulses from one electrode plate set and the application of pulses to the other electrode plate set.

According to another aspect of the present invention, there is provided a method of generating gas comprising, applying positive and negative voltage pulses alternately to the electrodes (positioned either side of, but not in contact with, a liquid dielectric), the voltage pulses being applied so that, whenever one of a positive voltage pulse and a negative voltage pulse is applied to one of the two electrodes, the other of a positive voltage pulse and a negative voltage pulse is applied to the other of the two electrodes, the applied voltage pulses generating an alternating electric field across the liquid dielectric causing fracture of the liquid dielectric into gaseous media. Preferably, voltages of at least 15 kV, e.g. 25 kV, either side of a reference value, e.g. earth, are applied across the liquid dielectric to generate the alternating electric field.

An embodiment of the invention will now be described by way of example only, with particular reference to the accompanying drawings, in which:

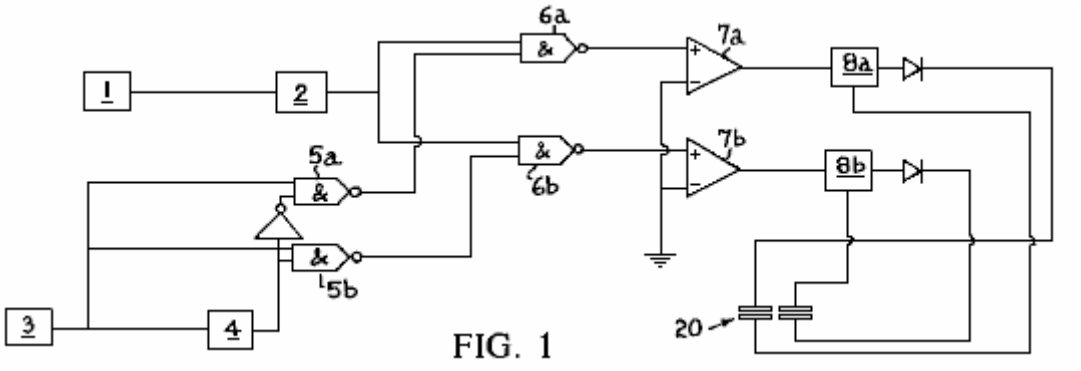


FIG. 1

Fig.1 is a circuit diagram of fracture cell apparatus according to the invention;

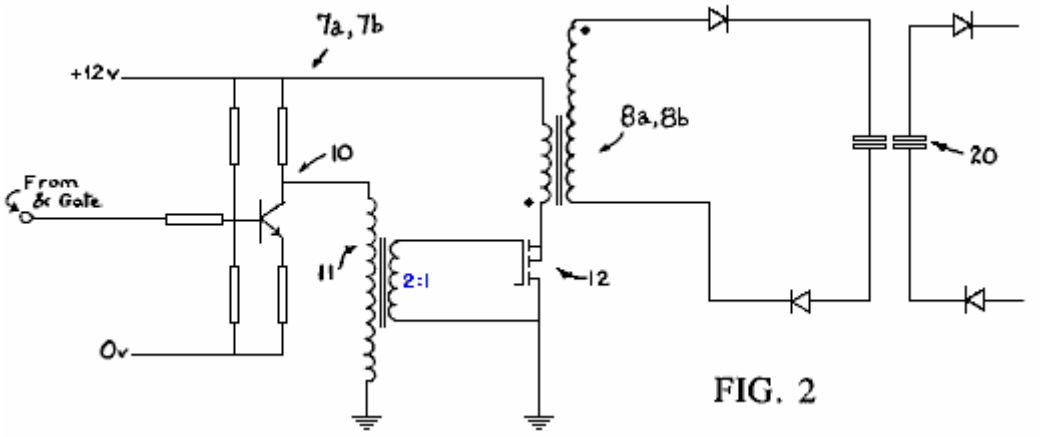


FIG. 2

Fig.2 shows in more detail a part of the circuit diagram of Figure 1;

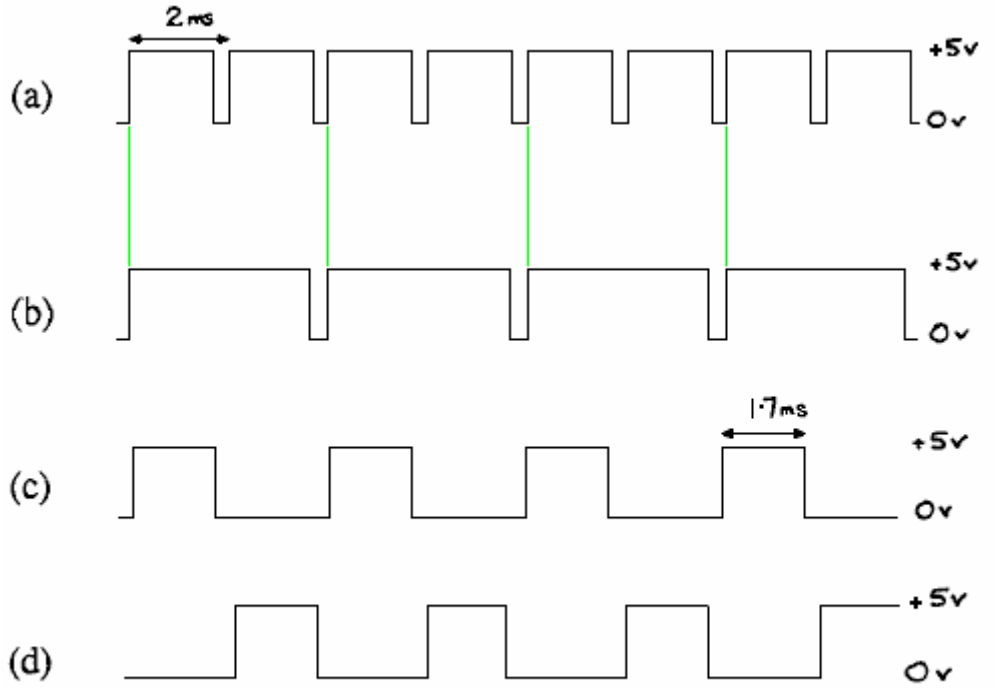


FIG. 3

Fig.3 shows the different waveforms at various parts of the circuit diagram of **Fig.1**;

FIG. 4

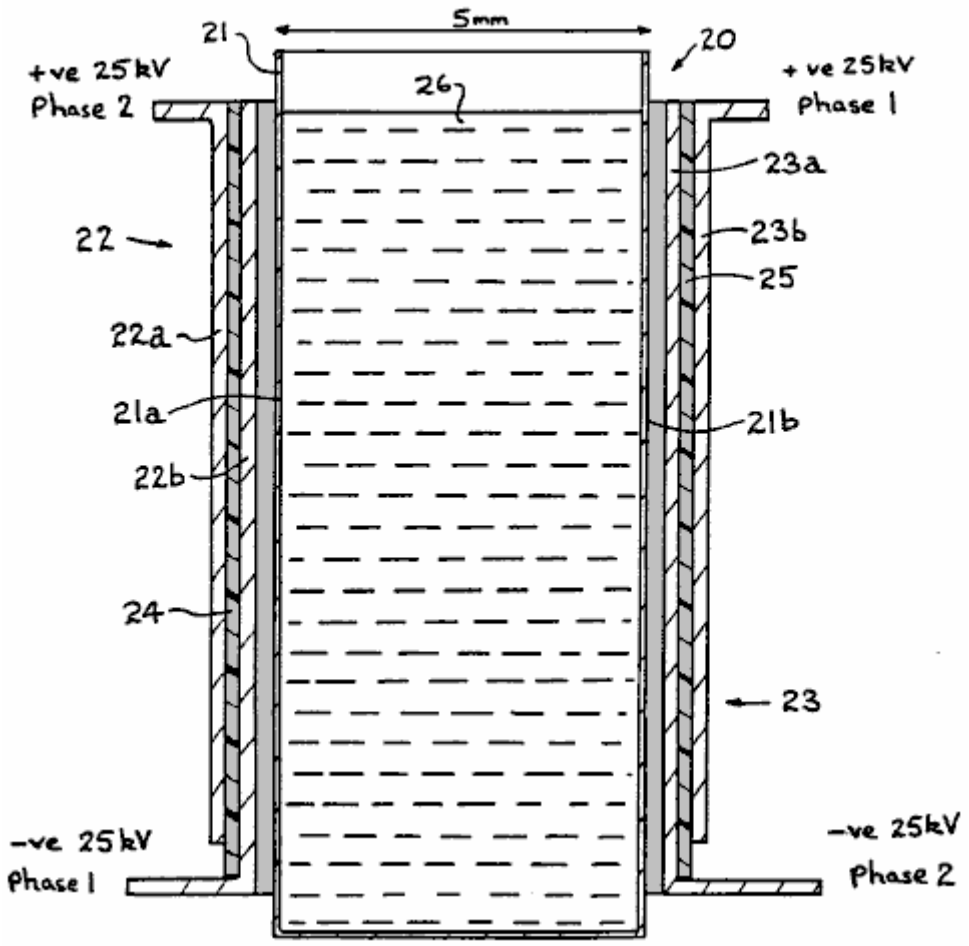


Fig.4 is a schematic diagram of a fracture cell for use in fracture cell apparatus according to the invention,

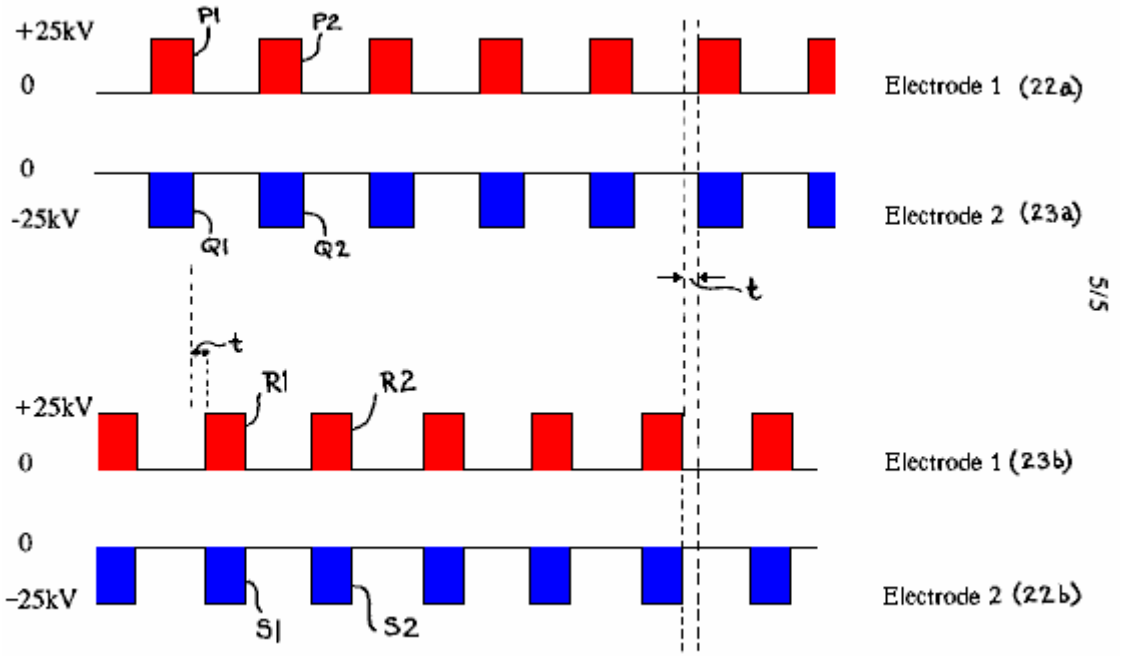


FIG. 5

Fig.5 shows trains of pulses applied to electrodes of the fracture cell apparatus according to the invention.

If a large electric field is applied across a pair of electrode plates positioned either side of a cell containing water, disruption of the water molecules will occur. Such disruption yields hydrogen nuclei and HO- ions. Such a molecular disruption is of little interest in terms of obtaining a usable result from the cell. A proton-rich zone exists for as long as the field exists and quickly re-establishes equilibrium ion-product when the field is removed.

One noticeable side-effect, however, is that the hydroxyl ions (which will migrate to the +ve charged plate) are stripped of electrons as they approach the cell boundary. Any negatively-charged ion will exhibit this behaviour in a strong enough potential well, but the OH ions have a strong tendency to such dissociation. This results, momentarily, in a region of negative-charge close to the positive cell boundary. Thus, on opposite sides of the active cell, there are hydrogen nuclei (free proton zone) and displaced electrons (-ve charge zone), both tending to increase in density closer to the charged plates.

If, at this point, the charge is removed from the plates, there is a tendency for the charge-zones to move, albeit very slowly, towards the centre of the active cell. The ion-transit rates of free electrons and of hydrogen nuclei are, however, some two orders of magnitude greater than either H30+ ions or OH ions.

If the charges are now replaced on the plates, but with opposite polarity, the interesting and potentially useful aspect of the process is revealed. Hydrogen nucleus migration is accelerated in the direction of the new -ve plate and free electron migration takes place towards the new +ve plate. Where there is a sufficient concentration of both species, including the accumulations due to previous polarity changes, monatomic hydrogen is formed with the liberation of some heat energy. Normal molecular association occurs and H2 gas bubbles off from the cell.

Also existing OH radicals are further stripped of hydrogen nuclei and contribute to the process. Active, nascent 0- - ions rapidly lose their electronic space charge to the +ve field and monatomic oxygen forms, forming the diatomic molecule and similarly bubbling off from the cell.

Thus, the continuous application of a strong electric field, changing in polarity every cycle, is sufficient to disrupt water into its constituent gaseous elements, utilising a small fraction of the energy required in conventional electrolysis or chemical energetics, and yielding heat energy of the enthalpy of formation of the diatomic bonds in the hydrogen and oxygen.

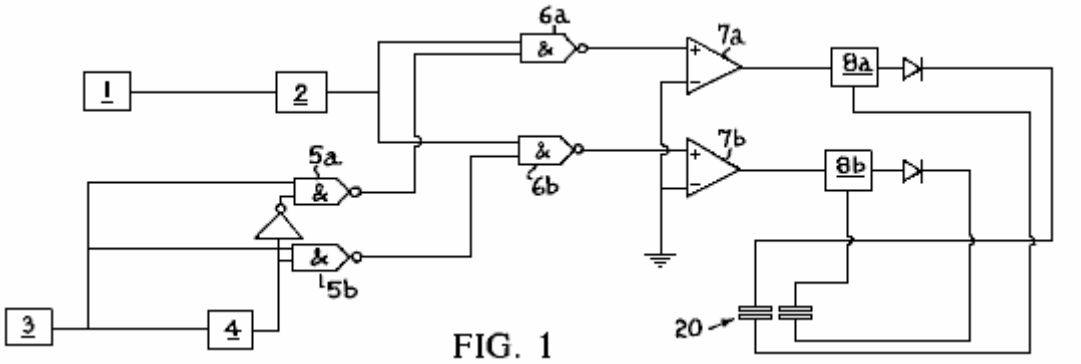


FIG. 1

Apparatus for performing the above process is described below. In particular, electronic circuitry to effect the invention is shown in the simplified block diagram of **Fig.1**. In **Fig.1** a pulse-repetition frequency (PRF) generator **1** comprises an astable multivibrator clock running at a frequency which is preset for any application, but able to be varied across a range of approximately 5-30 kHz. The generator **1** drives, by triggering with the trailing edge of its waveform, a pulse-width (PW) timer **2**.

The output of the timer **2** is a train of regular pulses whose width is determined by the setting of timer **2** and whose repetition frequency is set by the PRF generator **1**.

A gate clock **3** comprises a simple 555-type circuit which produce a waveform (see **Fig.3a**) having a period of 1 to 5 ms, e.g. 2 ms as shown in **Fig.3a**. The duty cycle of this waveform is variable from 50% to around 95%. The waveform is applied to one input of each of a pair of AND gates **5a** and **5b** and also to a binary divide-by-two counter **4**. The output of the counter **4** is shown in **Fig.3b**.

The signal from the divide-by-two counter **4** is applied directly to the AND gate **5b** serving phase-2 driver circuitry **7a** but is inverted before application to the AND gate **5a** serving phase-l driver circuitry **7a**. The output of the AND gate **5a** is therefore ((CLOCK and (NOT (CLOCK)/2)) and the output of the AND gate **5b** is ((CLOCK) and (CLOCK/2)), the waveforms, which are applied to pulse-train gates **6a** and **6b**, being shown in **Fig.3c** and **Fig.3d**.

Trains of 5-30 kHz pulses are applied to drive amplifiers **7a** and **7b** alternately, with a small "off"-period during which no pulses are applied to either amplifier. The duration of each "off" period is dependent upon the original duty cycle of the clock timer **3**. The reason for the small "off" period in the driver waveforms is to prevent local corona arc as the phases change over each cycle.

The drive amplifiers **7a** and **7b** each use a BC182L transistor **10** (see **Fig.2**), small toroidal 2:1 pulse transformer **11** and a BUZll power-MOSFET **12** and apply pulse packets across the primary windings of their respective 25 kV line-output transformers **8a** and **8b** to produce an EHT ac voltage of high frequency at their secondary windings. The secondary windings are 'lifted' from system ground and provide, after simple half-wave rectification, the applied field for application to cell **20** (see **Fig.4**).

Cell **20** comprises a container **21** having walls **21a**, **21b** of electrically insulating material, e.g. a thermoplastics material, such as polymethyl methacrylate, typically spaced about 5 mm apart, and bipolar cell electrodes generally designated **22** and **23** and typically constructed from aluminium foil, positioned outside the walls **21a** and **21b**. Each bipolar cell electrode comprises a pair of electrode plates **22a** and **22b** (or **23a** and **23b**) for each side of the cell **20** separated from each other by an electrically insulating layer **24** (or **25**) , e.g. of polycarbonate plastics material about 0.3 mm thick.

The electrode plates **22a** and **23a** form one set (set A) of electrode plates positioned on opposite sides of container **21** and the electrode plates **22b** and **23b** form another set of electrode plates positioned on opposite sides of the container **21**. An insulating layer **25**, e.g. of polycarbonate material, similar to the insulating layers **24a** or **24b** may be positioned between each bipolar cell electrode **22** (or **23**) and its adjacent container wall **21a**(or **21b**). A liquid electrolyte, preferably water, is placed in the container **21**.

In use, a train of positive pulses is applied to the electrode plates **22a** and **23b** and a train of negative pulses is applied to the electrode plates **23a** and **22b**. The timing of the pulses is shown schematically in **Fig.5**, which illustrates that, for set A (or for set B), whenever a positive pulse is applied to electrode plate **22a** (or **23b**), a negative pulse is also applied to electrode plate **23a** (or **22b**). However the pulses applied to the electrode plate set A are "out of phase" with the pulses applied to the electrode plate set B. In each train of pulses, the duration of each pulse is less than the gap between successive pulses.

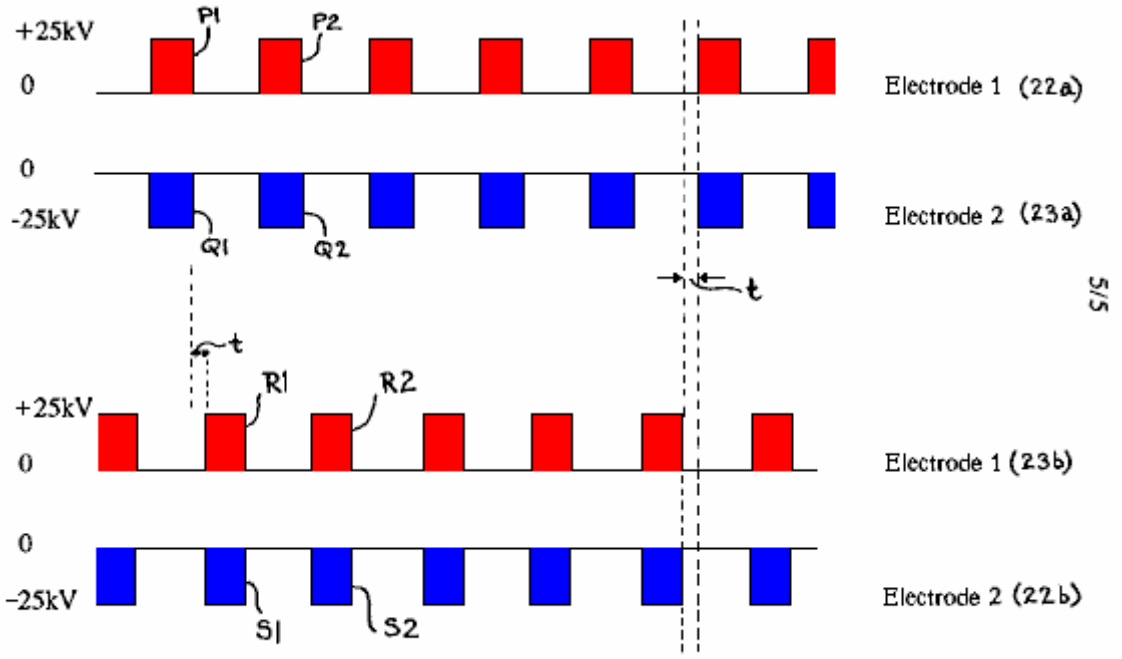


FIG. 5

By arranging for the pulses of electrode plate set B to be applied in the periods when no pulses are applied to the electrode plate set A, the situation arises where pairs of pulses are applied successively to the electrode plates of different sets of electrode plates, there being a short interval of time when no pulses are applied between each successive application of pulses to pairs of electrode plates. In other words, looking at **Fig.5**, pulses **P1** and **Q1** are applied at the same time to the electrode plates **22a** and **23a**. The pulses **P1** and **Q1** are of the same pulse length and, at the end of their duration, there is a short time period **t** before pulses **R1** and **S1** are applied to the electrode plates **23b** and **22b**.

The pulses **R1** and **S1** are of the same pulse length as the pulses **P1** and **Q1** and, at the end of their duration, there is a further time **t** before the next pulses **P2** and **Q2** are applied to the electrode plates **22a** and **23a**. It will be appreciated that whenever a pulse of one sign is applied to one of the electrode plates of a set, a pulse of the opposite sign is applied to the other electrode plate of that set.

Furthermore, by switching from one to the other electrode plate set the polarities applied across the container are repeatedly switched resulting in an "alternating" electric field being created across the "liquid dielectric" water in the container.

The Electrolyser of Spiro Spiros

Patent WO 9528510 26th October 1995 Inventor: Spiro Ross Spiros

IMPROVEMENTS IN ELECTROLYSIS SYSTEMS & THE AVAILABILITY OF OVER-UNITY ENERGY

This patent application shows the details of an electrolyser system which it is claimed, produces greater output than the input power needed to operate it.

ABSTRACT

A looped energy system for the generation of excess energy available to do work is disclosed. The system comprises an electrolysis cell unit **150** receiving a supply of water to liberate separated hydrogen gas **154** and oxygen **156** by electrolysis driven by a DC voltage **152** applied across respective anodes and cathodes of the cell unit **150**. A hydrogen gas receiver **158** receives and stores hydrogen gas liberated by the cell unit **150**, and an oxygen gas receiver **160** receives and stores oxygen gas liberated by the cell unit **150**. A gas expansion device **162** expands the stored gases to recover expansion work, and a gas combustion device **168** mixes and combusts the expanded hydrogen gas and oxygen gas to recover combusted work. A proportion of the sum of the expansion work and the combustion work sustains electrolysis of the cell unit to retain operational gas pressure in the gas receivers **158**, **160** such that the energy system is self-sustaining, and there is excess energy available from the sum of energies.

TECHNICAL FIELD OF THE INVENTION

The present invention relates to the generation of hydrogen gas and oxygen gas from water, either as an admixture or as separated gases, by the process of electrolysis, and relates further to applications for the use of the liberated gas. Embodiments of the invention relate particularly to apparatus for the efficient generation of these gases, and to use of the gases in an internal combustion engine and an implosion pump. The invention also discloses a closed-loop energy generation system where latent molecular energy is liberated as a form of 'free energy' so the system can be self-sustaining.

Reference is made to commonly-owned International patent application No. PCT/AU94/000532, having the International filing date of 6 September 1994.

Background Art

The technique of electrolysing water in the presence of an electrolyte such as sodium hydroxide (NaOH) or potassium hydroxide (KOH) to liberate hydrogen and oxygen gas (H2, 02) is well known. The process involves applying a DC potential difference between two or more anode/cathode electrode pairs and delivering the minimum energy required to break the H-O bonds (i.e. 68.3 kcal per mole @ STP).

The gases are produced in the stoichiometric proportions for O2:H2 of 1:2 liberated respectively from the anode (+) and cathode (-).

Reference can be made to the following texts:
"Modern Electrochemistry, Volume 2, John O'M. Bockris and Amulya K.N. Reddy, Plenum Publishing Corporation",
"Electro-Chemical Science, J. O'M. Bockris and D.M. Drazic, Taylor and Francis Limited" and
"Fuel Cells, Their Electrochemistry, J. O'M. Bockris and S. Srinivasan, McGraw-Hill Book Company".

A discussion of experimental work in relation to electrolysis processes can be obtained from "Hydrogen Energy, Part A, Hydrogen Economy Miami Energy Conference, Miami Beach, Florida, 1974, edited by T. Nejat Veziroglu, Plenum Press". The papers presented by J. O'M. Bockris on pages 371 to 379, by F.C. Jensen and F.H. Schubert on pages 425 to 439 and by John B. Pangborn and John C. Sharer on pages 499 to 508 are of particular relevance.

On a macro-scale, the amount of gas produced depends upon a number of variables, including the type and concentration of the electrolytic solution used, the anode/cathode electrode pair surface area, the electrolytic resistance (equating to ionic conductivity, which is a function of temperature and pressure), achievable current density and anode/cathode potential difference. The total energy delivered must be sufficient to disassociate the

water ions to generate hydrogen and oxygen gases, yet avoid plating (oxidation/reduction) of the metallic or conductive non-metallic materials from which the electrodes are constructed.

DISCLOSURE OF THE INVENTION

The invention discloses a looped-energy system for the generation of excess energy available to do work, the said system comprising of:

An electrolysis cell unit receiving a supply of water for liberating separated hydrogen gas and oxygen gas by electrolysis due to a DC voltage applied across respective anodes and cathodes of the cell;

A hydrogen gas receiver to receive and store the hydrogen gas liberated by the electrolysis cell;

An oxygen gas receiver to receive and store the oxygen gas liberated by the electrolysis cell;

A gas-expansion chamber to allow the expansion of the stored gases to recover expansion work; and

A gas-combustion mechanism for mixing and combusting the expanded hydrogen and oxygen gases to recover combustion work; and wherein a proportion of the sum of the expansion work and the combustion work sustains the electrolysis of the electrolysis cell in order to retain the operational gas pressure in the hydrogen and oxygen gas receivers so that the energy system is self-sustaining and there is excess energy available.

The invention further discloses a method for the generation of excess energy available to do work by the process of electrolysis, said method comprising the steps of: electrolysing water by a DC voltage to liberate separated hydrogen gas and oxygen gas; separately receiving and storing the hydrogen and oxygen gases in a manner to be self-pressuring; separately expanding the stored gas to recover expansion energy; burning the expanded gases to recover combustion energy; and applying a portion of the sum of the expansion work and the combustion work as the DC voltage to retain operational gas pressures and sustain the electrolysis, there being excess energy available to do this.

The invention also discloses an internal combustion engine powered by hydrogen and oxygen comprising of:

At least one cylinder and

At least one reciprocating piston within the cylinder;

A hydrogen gas input port in communication with the cylinder for receiving a supply of pressurised hydrogen;

An oxygen gas input port in communication with the cylinder for receiving a supply of pressurised oxygen; and

An exhaust port in communication with the cylinder and wherein the engine can be operated in a two-stroke manner whereby, at the top of the stroke, hydrogen gas is supplied through the respective inlet port to the cylinder driving the piston downwards, oxygen gas then is supplied through the respective inlet port to the cylinder to drive the cylinder further downwards, after which time self-detonation occurs and the piston moves to the bottom of the stroke and upwards again with the exhaust port opened to force out the water vapour resulting from the detonation.

The invention also discloses an implosion pump comprising of;

A combustion chamber interposed, and in communication with,

An upper reservoir and a lower reservoir separated by a vertical distance across which water is to be pumped, this chamber receiving admixed hydrogen and oxygen at a pressure sufficient to lift a volume of water the distance from there to the top reservoir, the gas in the chamber then being ignited to create a vacuum in the chamber to draw water from the lower reservoir to fill the chamber, whereupon a pumping cycle is established and can be repeated.

The invention also discloses a parallel stacked arrangement of cell plates for a water electrolysis unit, the cell plates alternately forming an anode and cathode of the electrolysis unit, and the arrangement including separate

hydrogen gas and oxygen gas outlet ports respectively linked to the anode cell plates and the cathode cell plates and extending longitudinally along the plate stack. These outlet ports are arranged so as to be insulated from the anode and cathode plates.

DESCRIPTION OF THE DRAWINGS

Figs.1 1a-16 of noted International application no. PCT/AU94/000532 are reproduced to aid description of the present invention, but herein denoted as Figs.la-6:

Fig.1A and **Fig.1B** show an embodiment of a cell plate:

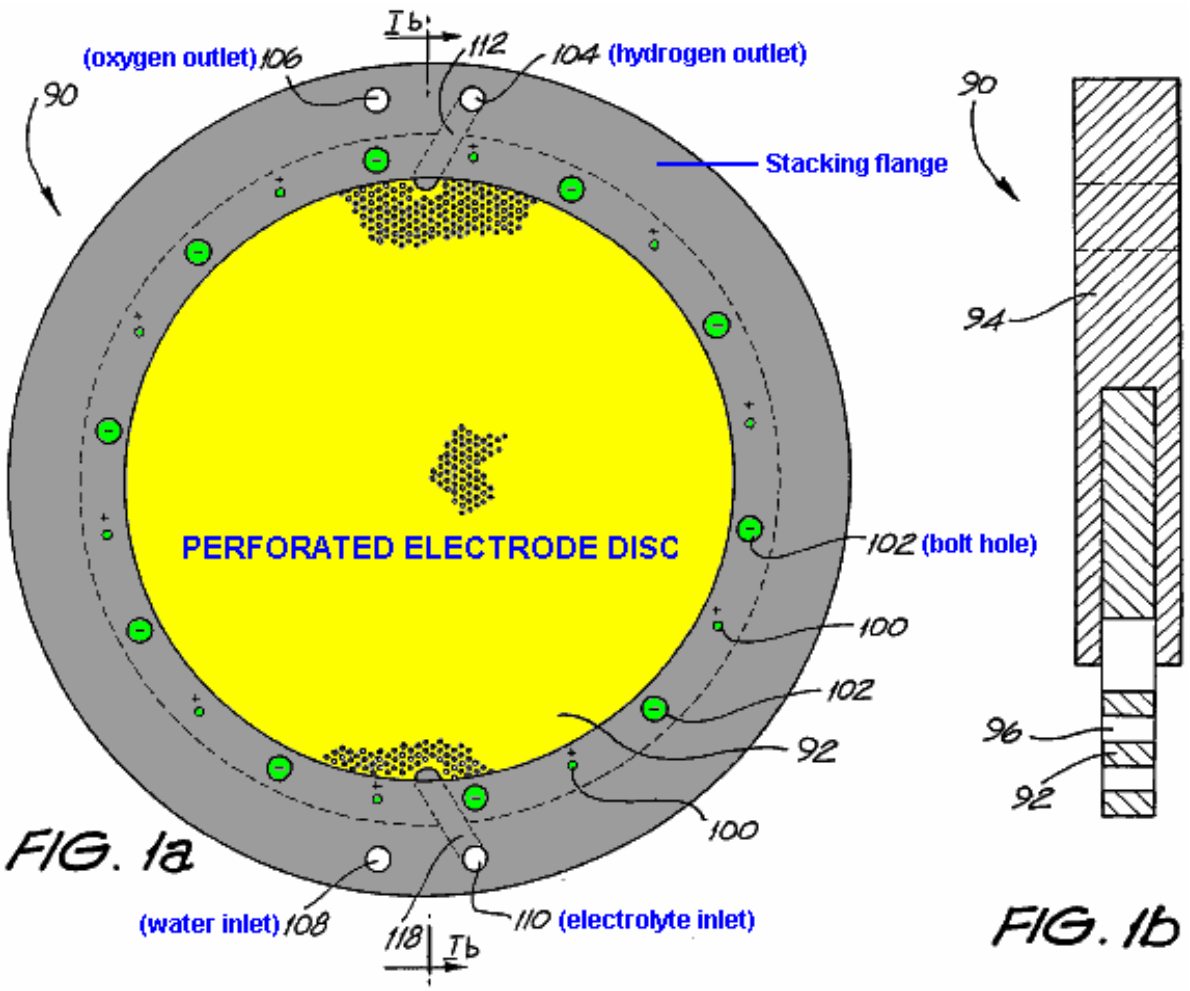


Fig.2A and **Fig.2B** show a complementary cell plate to that of Fig.lA and Fig1B:

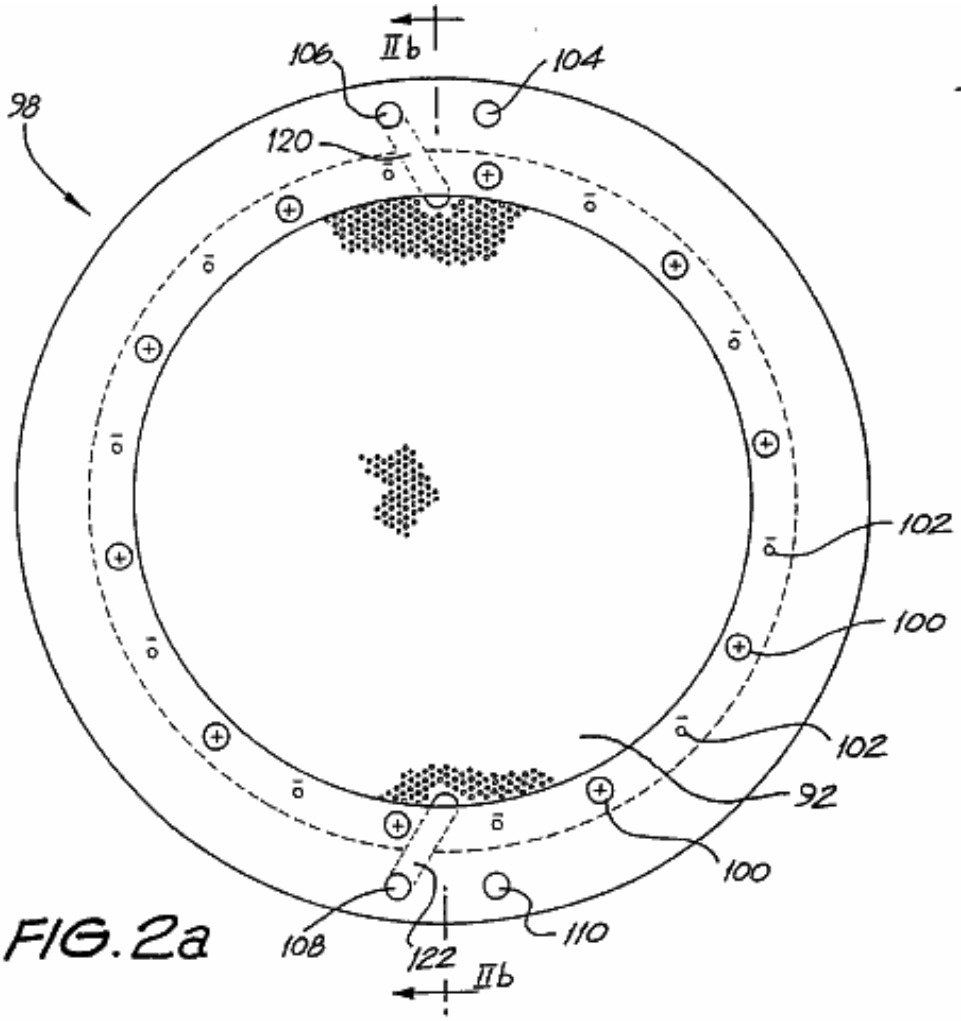


FIG. 2a

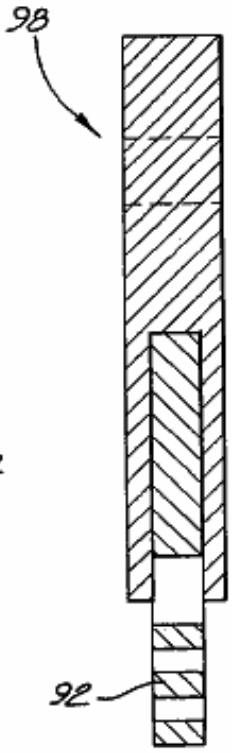


FIG. 2b

Fig.3 shows detail of the perforations and porting of the cell plates of Figs. lA,lB, 2A and 2B:

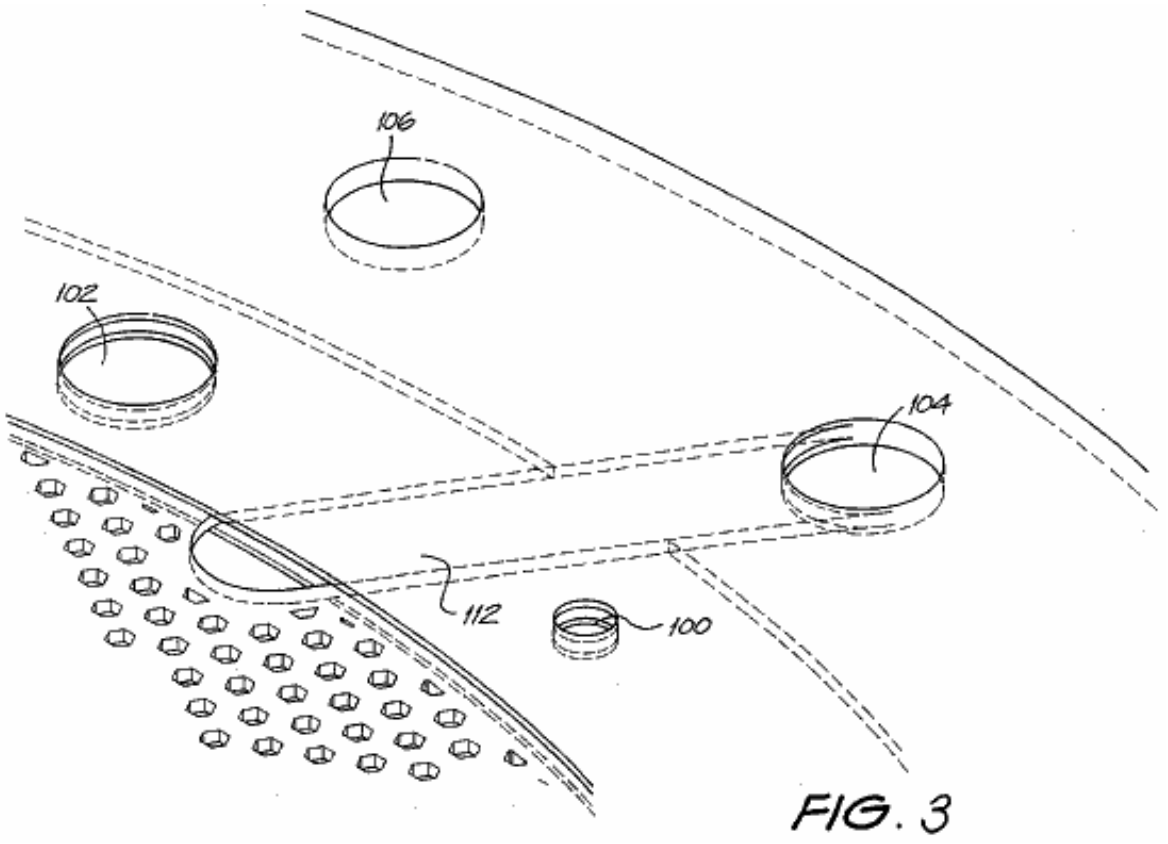


FIG. 3

Fig.4 shows an exploded stacked arrangement of the cell plates of Figs. lA,lB, 2A and 2B:

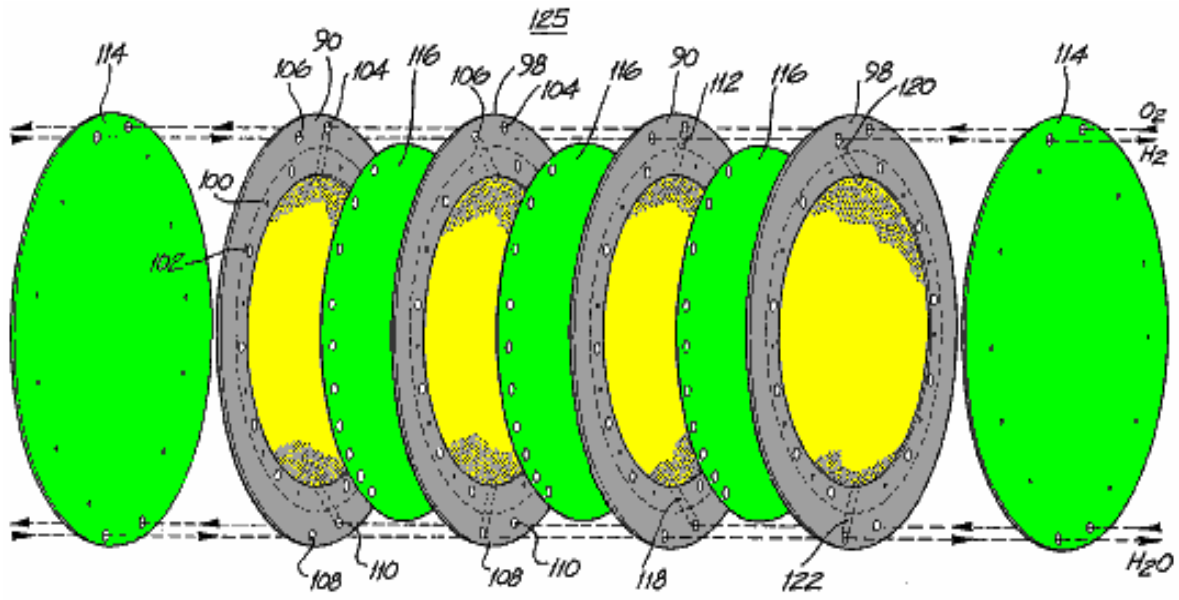


FIG. 4

Fig.5A shows a schematic view of the gas separation system of Fig.4:

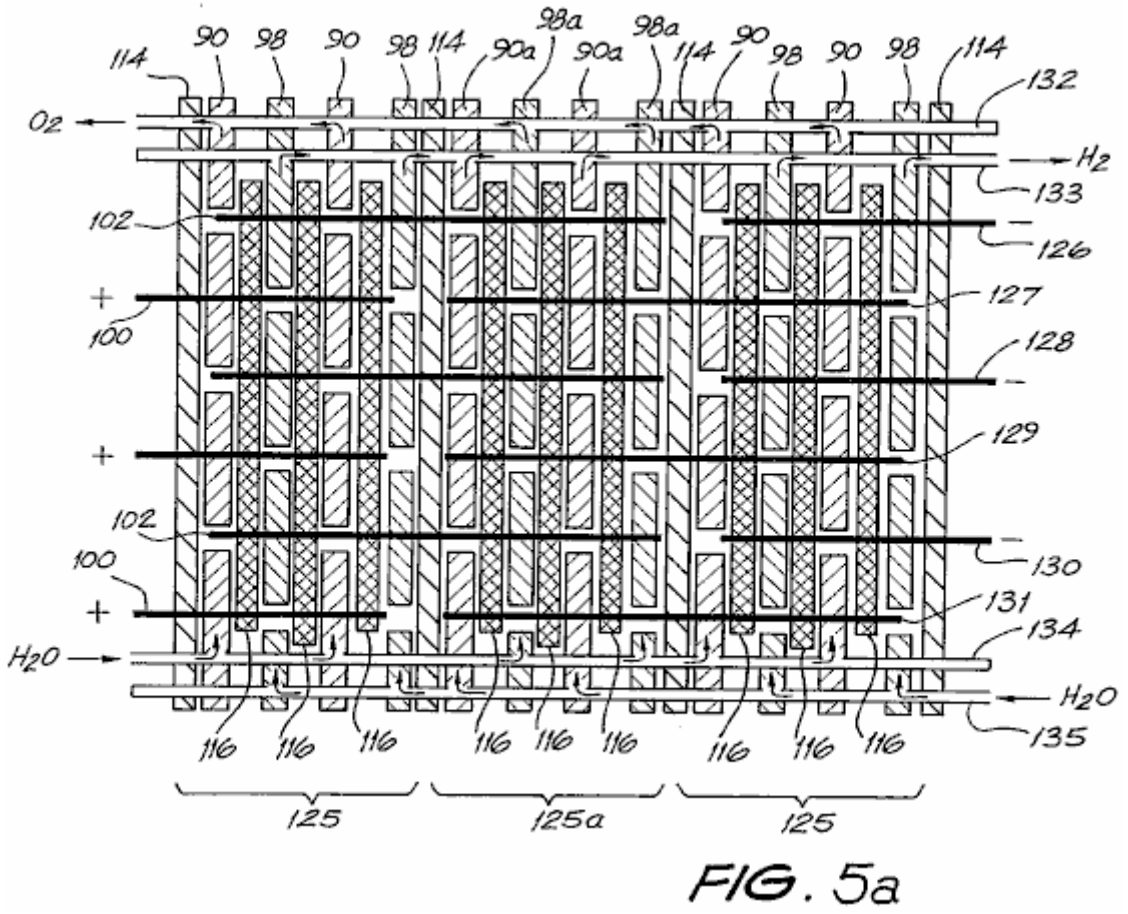


FIG. 5a

Fig.5B shows a stylised representation of Fig.5a:

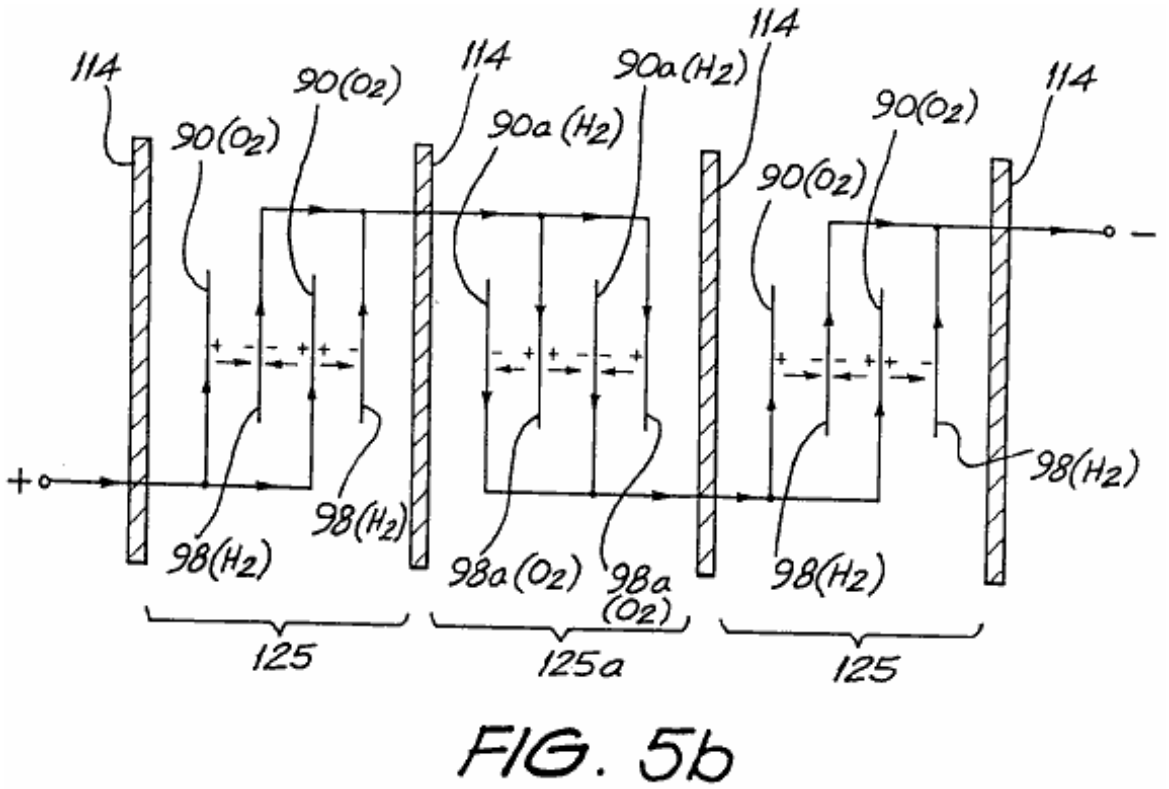


FIG. 5b

Fig.5C shows an electrical equivalent circuit of Fig.5A and

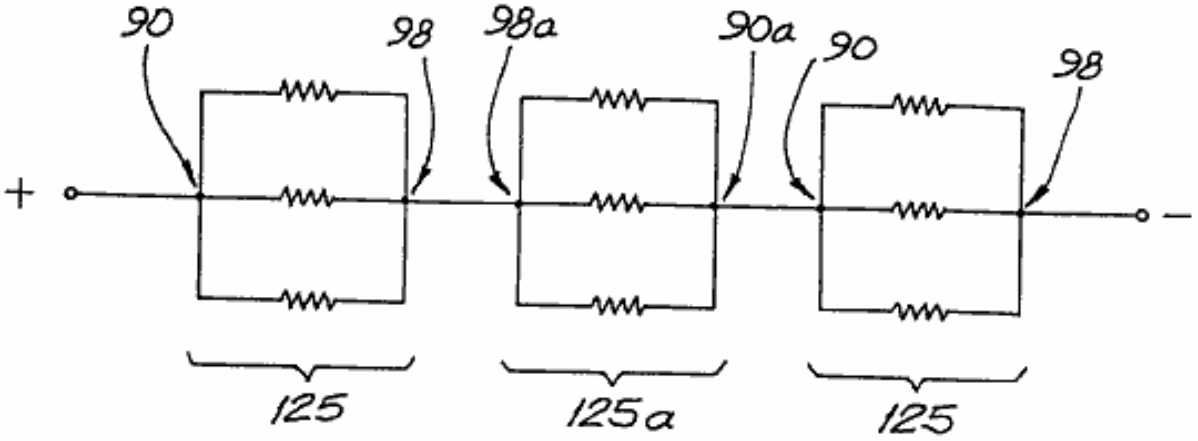


FIG. 5c

Fig.6 shows a gas collection system for use with the cell bank separation system of Figs. 4 and 5a.

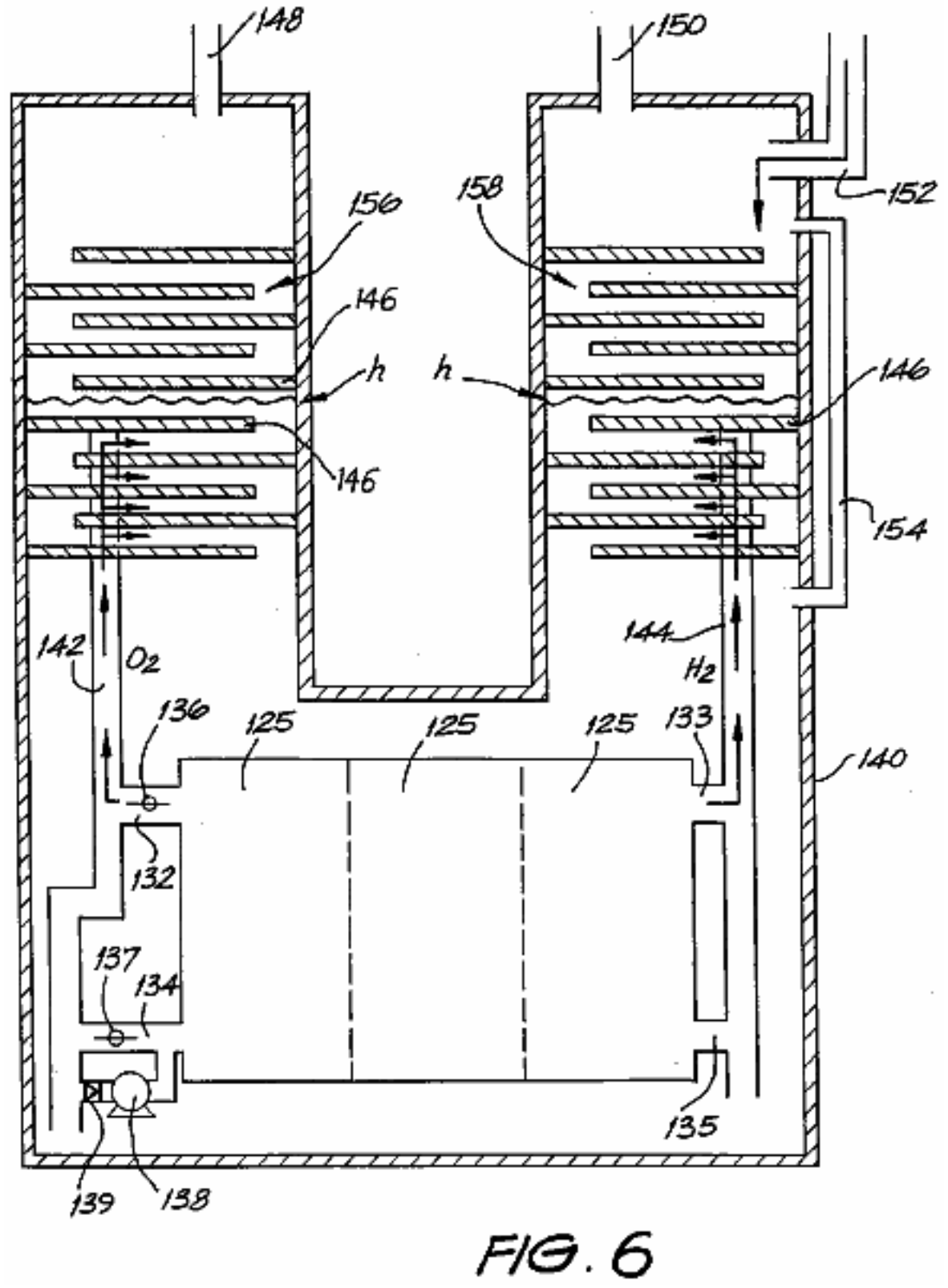


FIG. 6

The remaining drawings are:
Fig.7A and **Fig.7B** are views of a first cell plate:

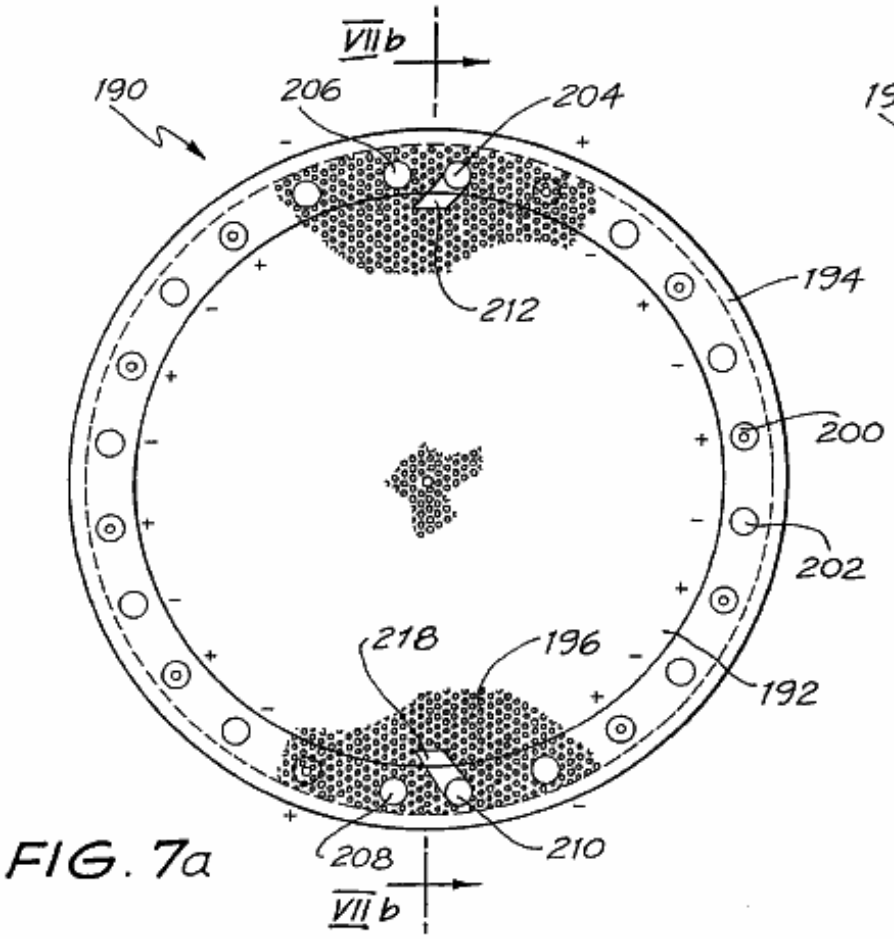


FIG. 7a

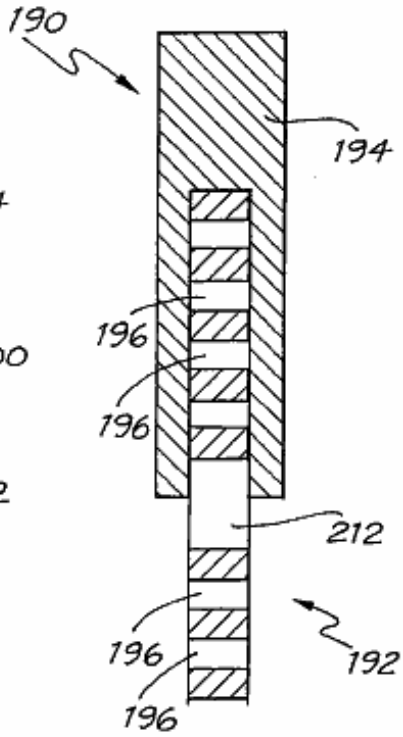


FIG. 7b

Fig.8A and **Fig.8B** are views of a second cell plate:

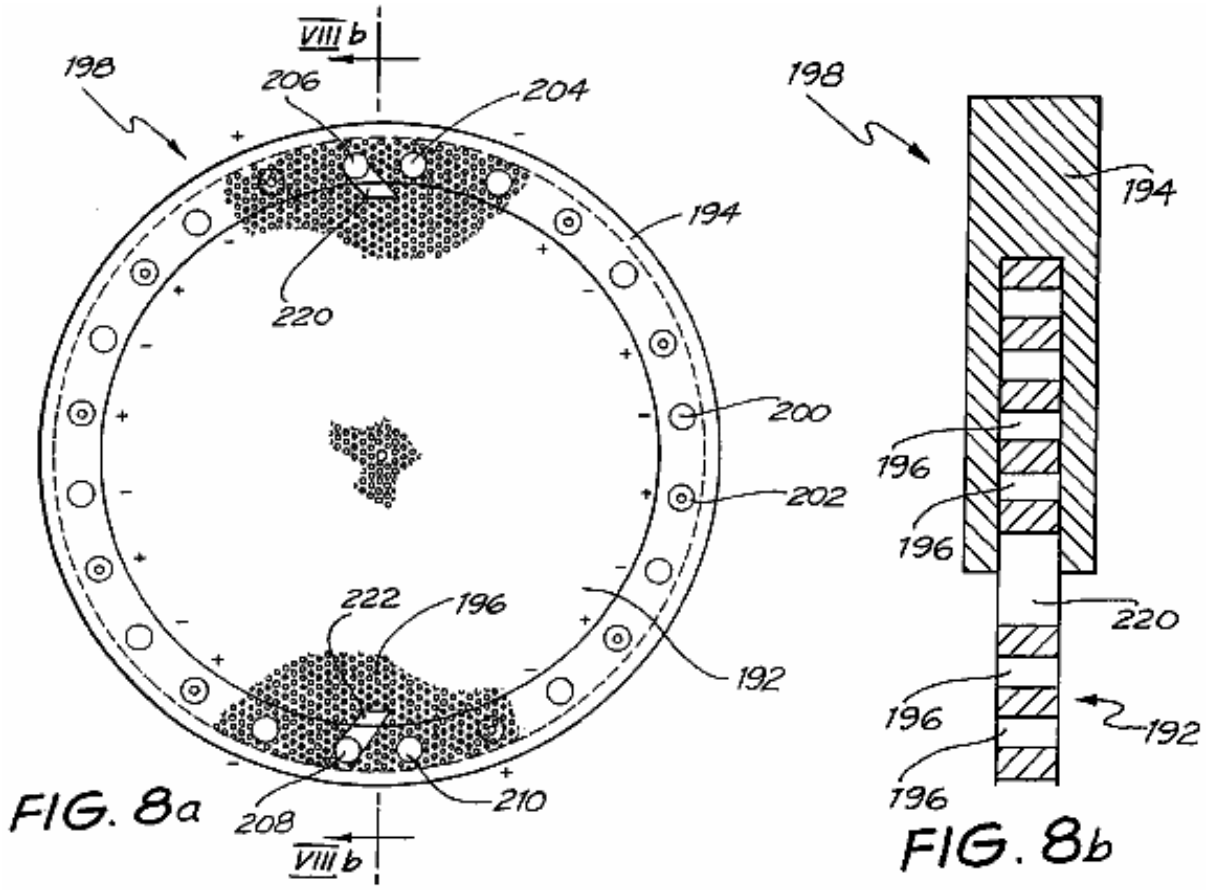


Fig.9 shows detail of the edge margin of the first cell plate:

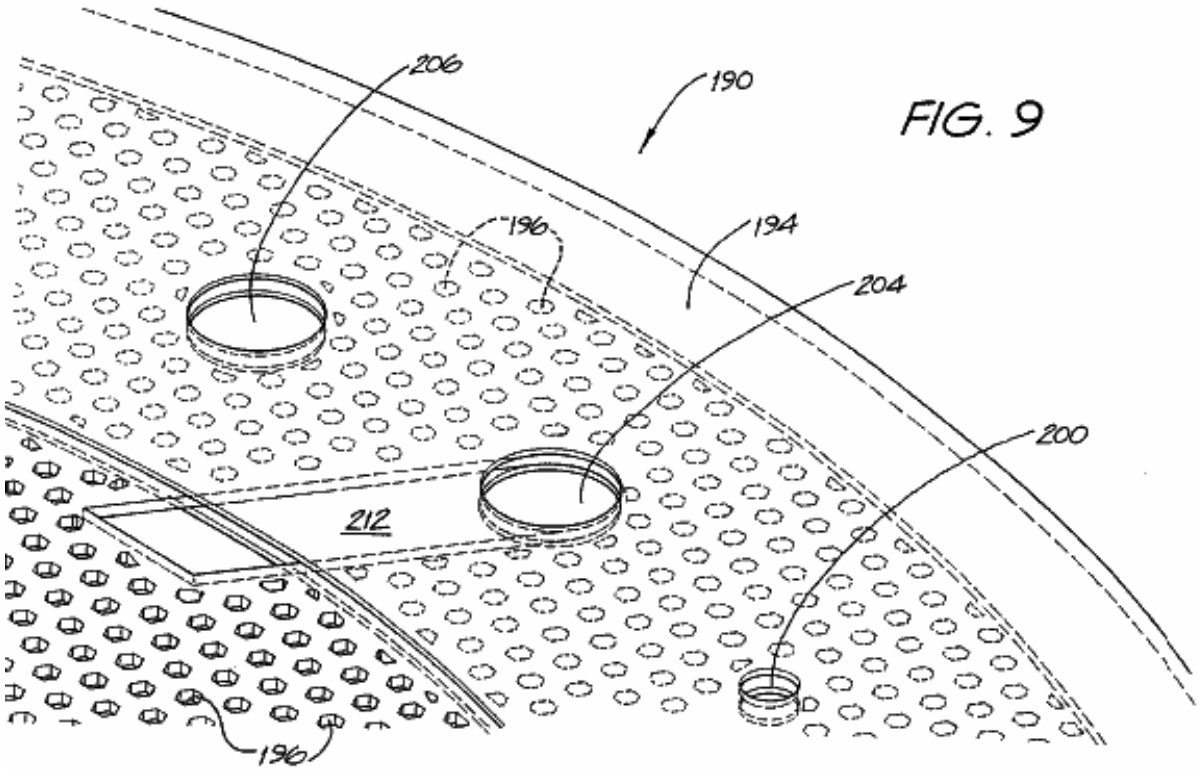


Fig10 shows an exploded stacked arrangement of the cell plates shown in Fig.7A and Fig.8A:

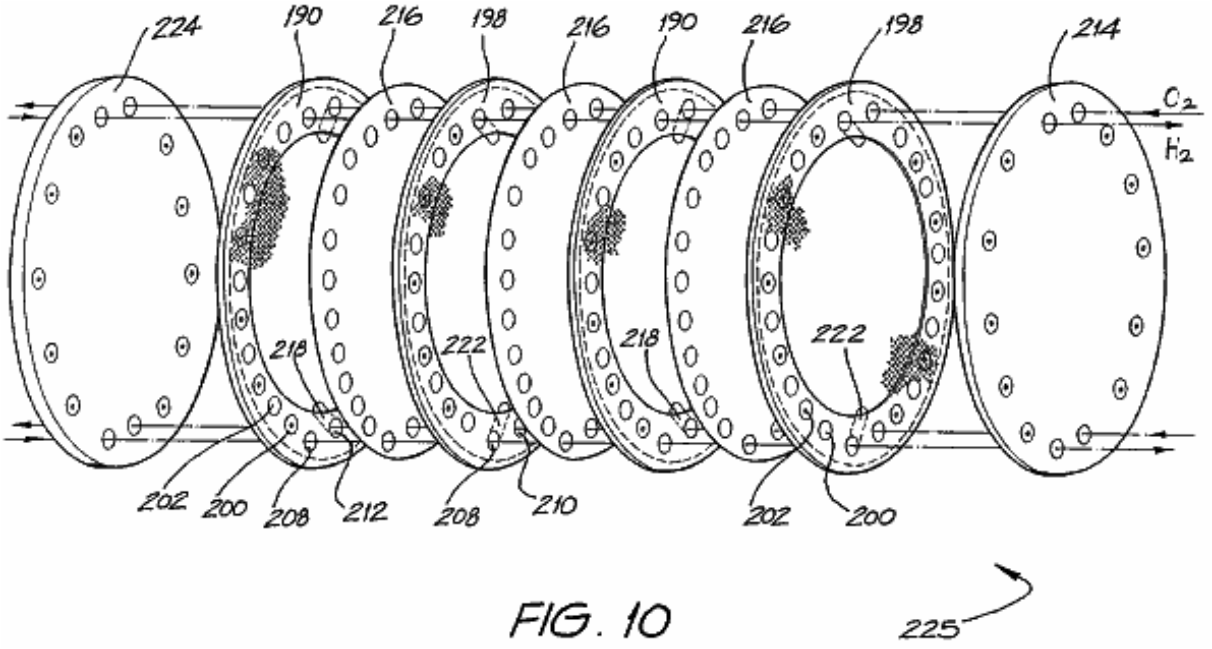


FIG. 10

Fig.11 is a cross-sectional view of three of the stacked cell plates shown in Fig.10 in the vicinity of a gas port:

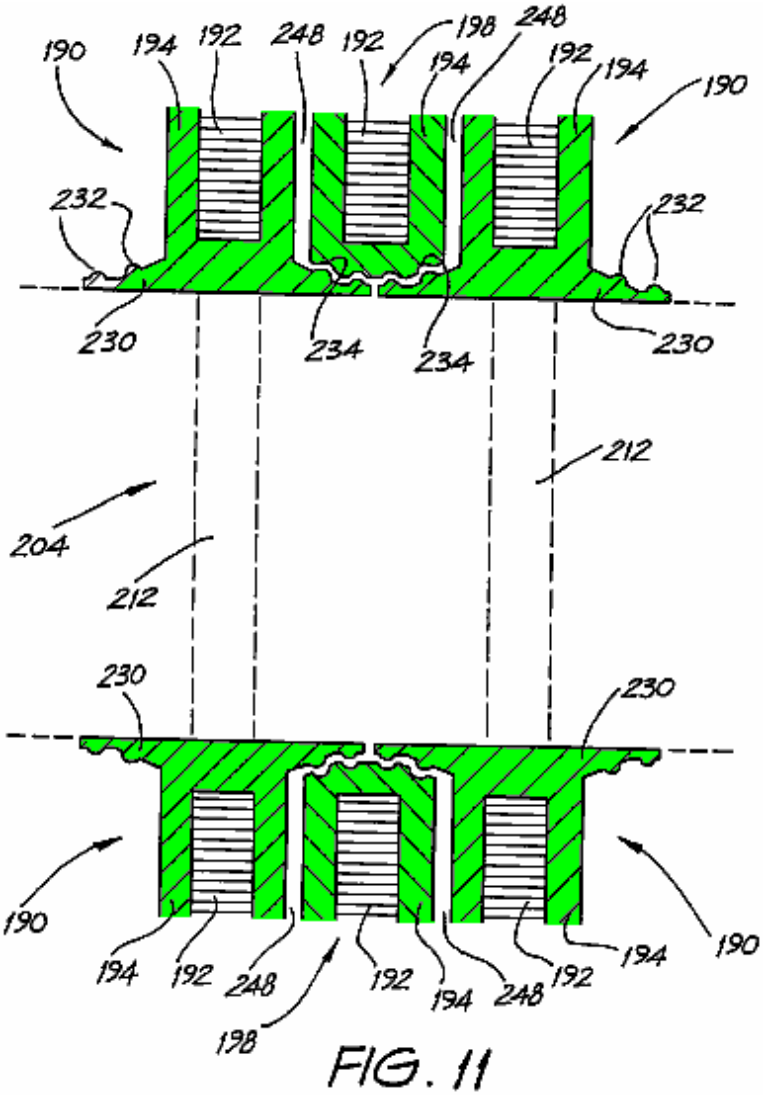


FIG. 11

Fig.12A and **Fig.12B** respectively show detail of the first and second cell plates in the vicinity of a gas port:

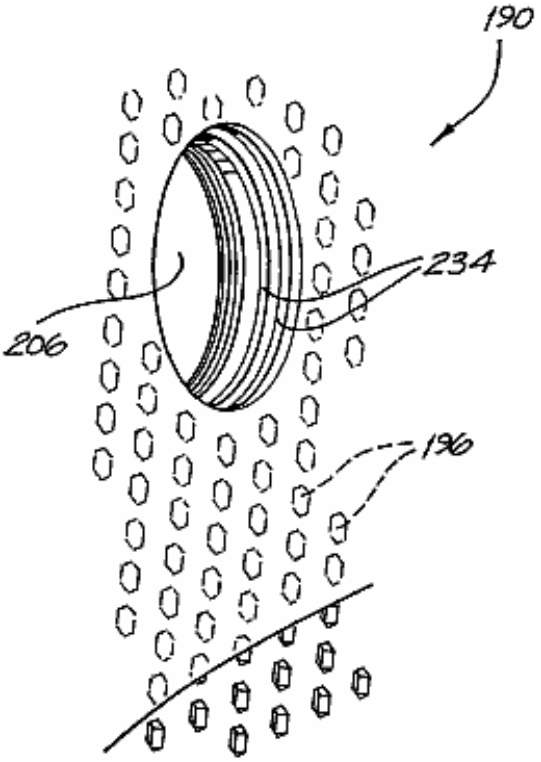


FIG. 12a

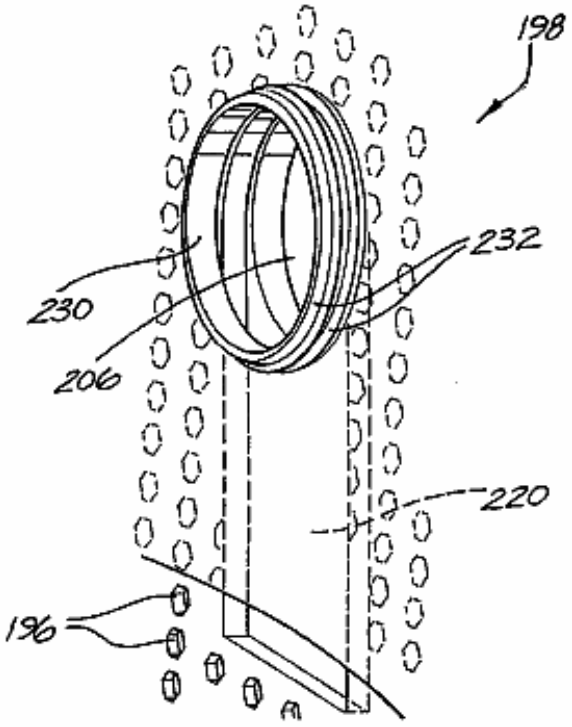


FIG. 12b

Fig.13 is a cross-sectional view of a cell unit of four stacked cell plates in the vicinity of an interconnecting shaft:

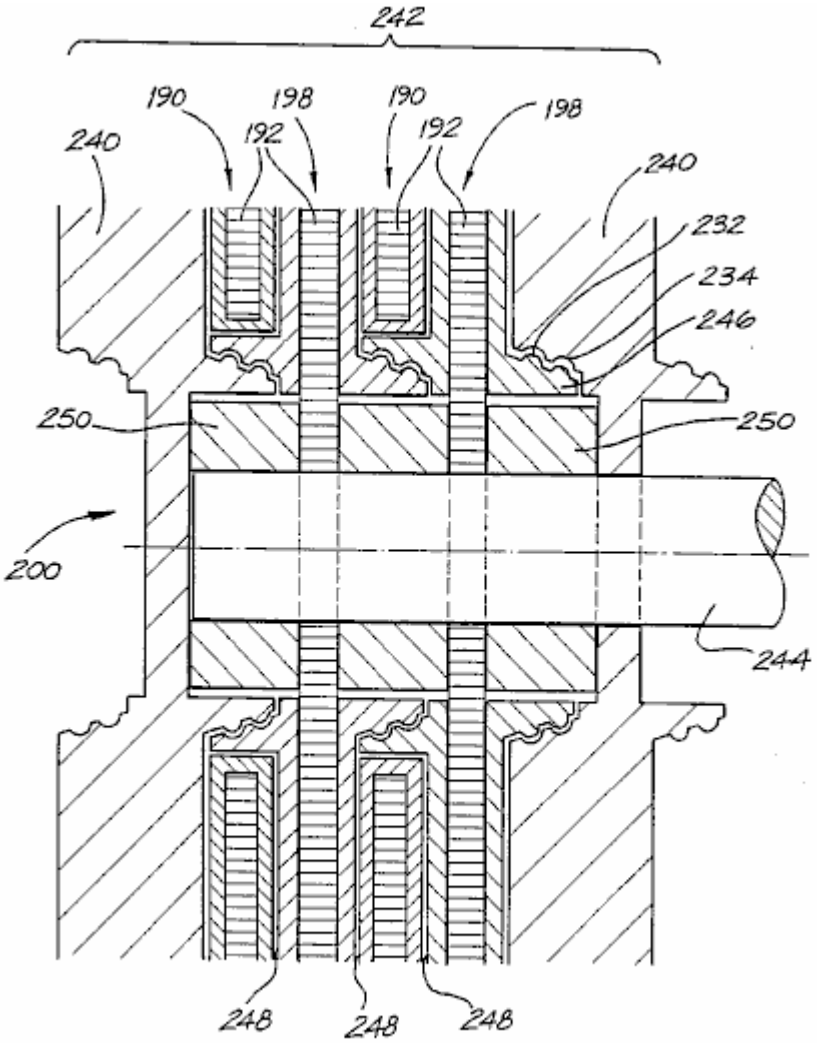


FIG. 13

Fig.14 shows a perspective view of a locking nut used in the arrangement of Fig.13:

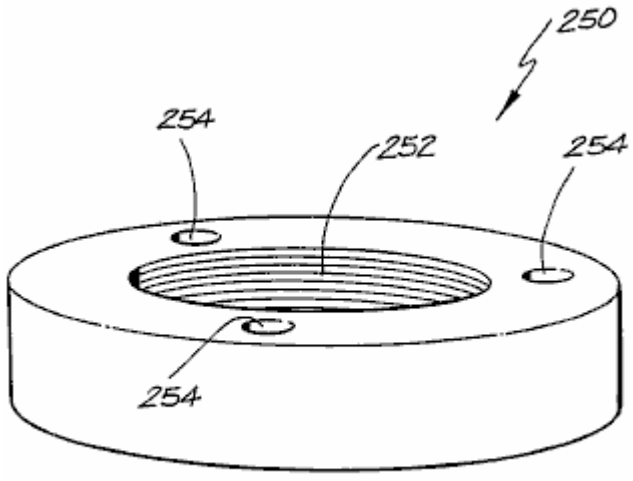


FIG. 14

Fig.15 shows an idealised electrolysis system:

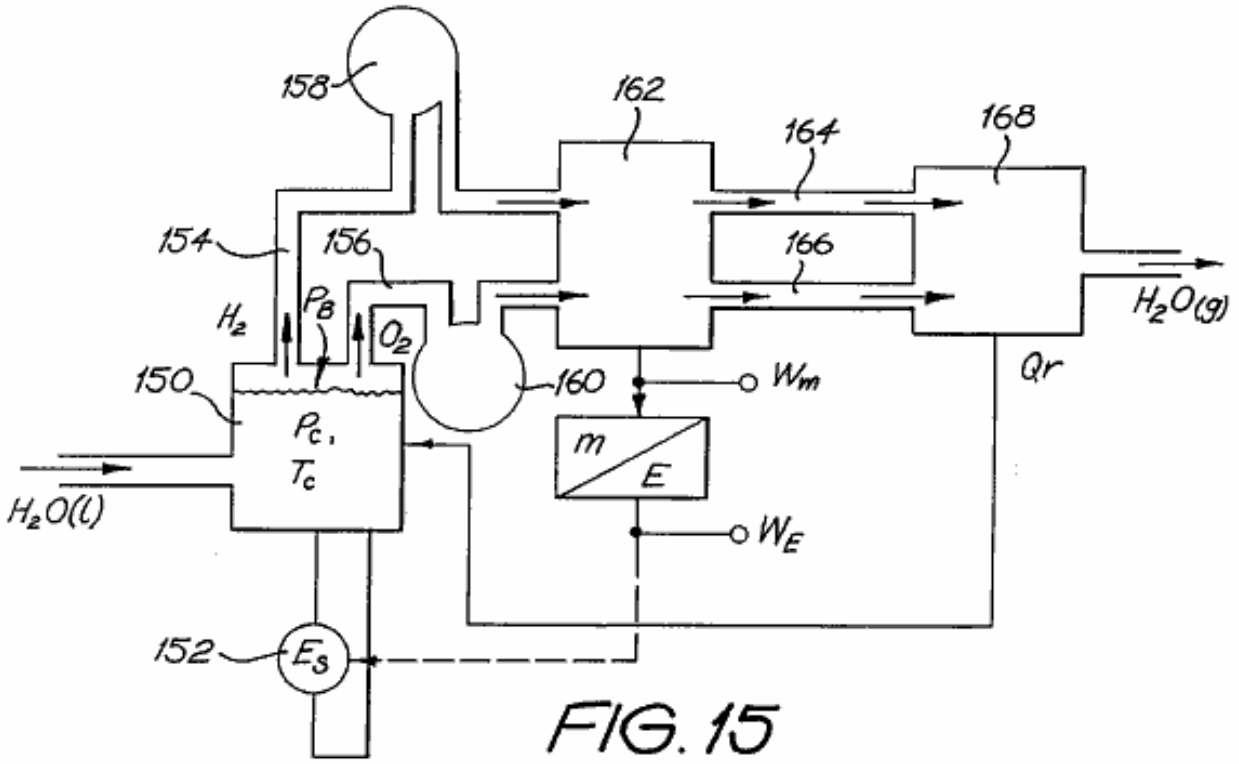


FIG. 15

Figs.16-30 are graphs supporting the system of **Fig.15** and the availability of over-unity energy:

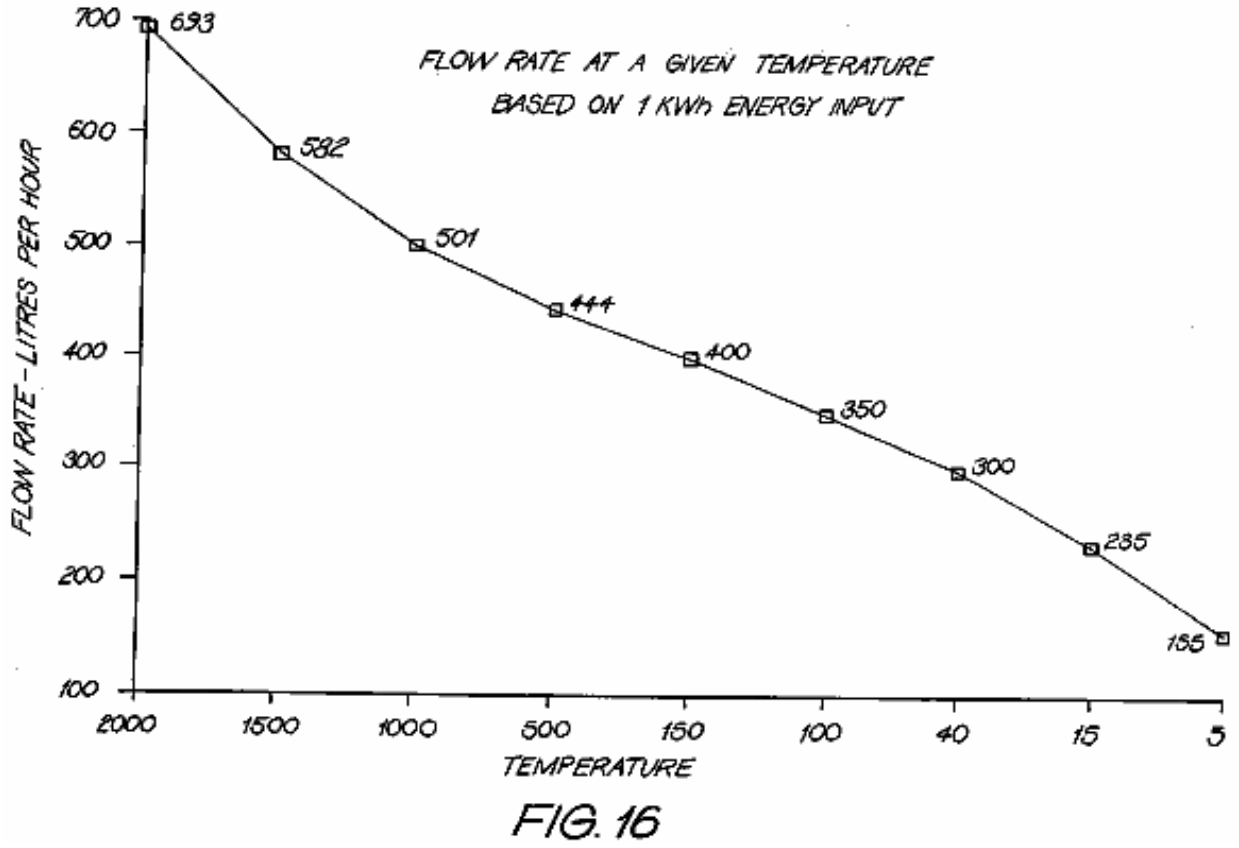


FIG. 16

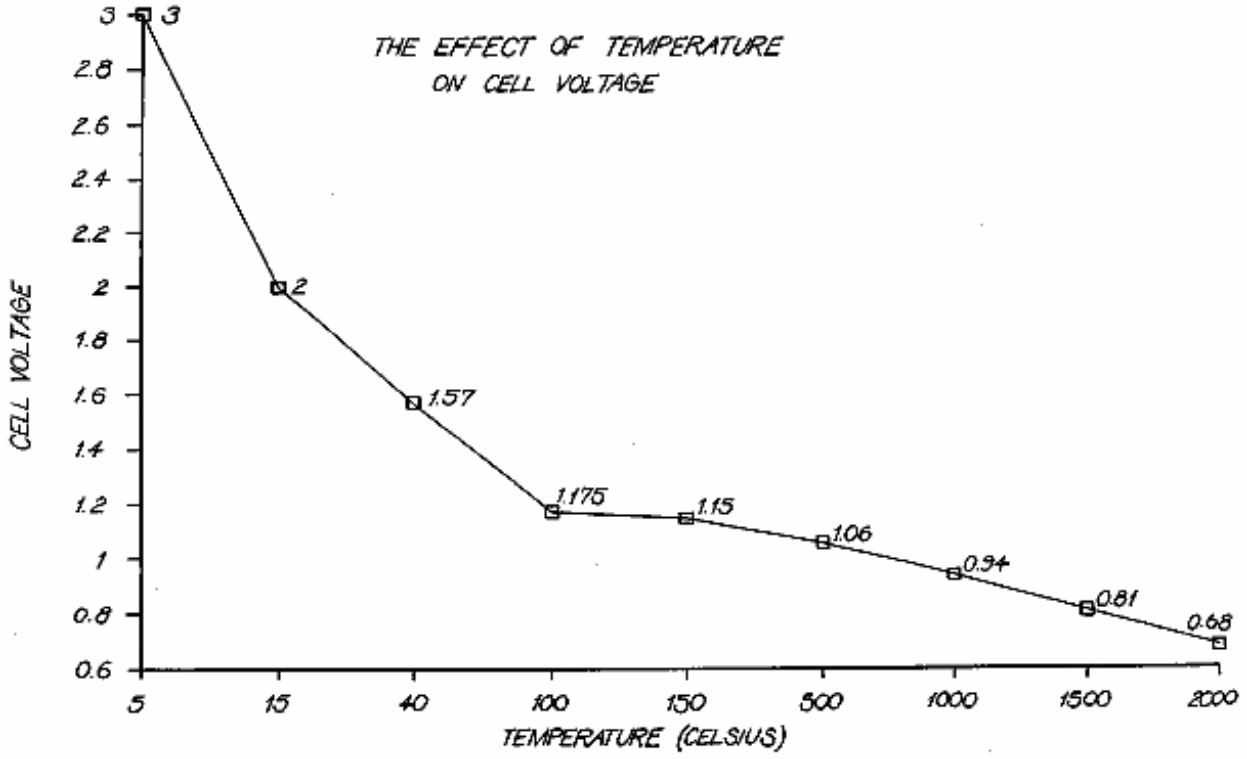

THE EFFECT OF TEMPERATURE
ON CELL VOLTAGE
CELL VOLTAGE
TEMPERATURE (CELSIUS)
3
2.8
2.6
2.4
2.2
2
1.8
1.6
1.4
1.2
1
0.8
0.6
5
15
40
100
150
500
1000
1500
2000
3
2
1.57
1.175
1.15
1.06
0.94
0.81
0.68

FIG. 17

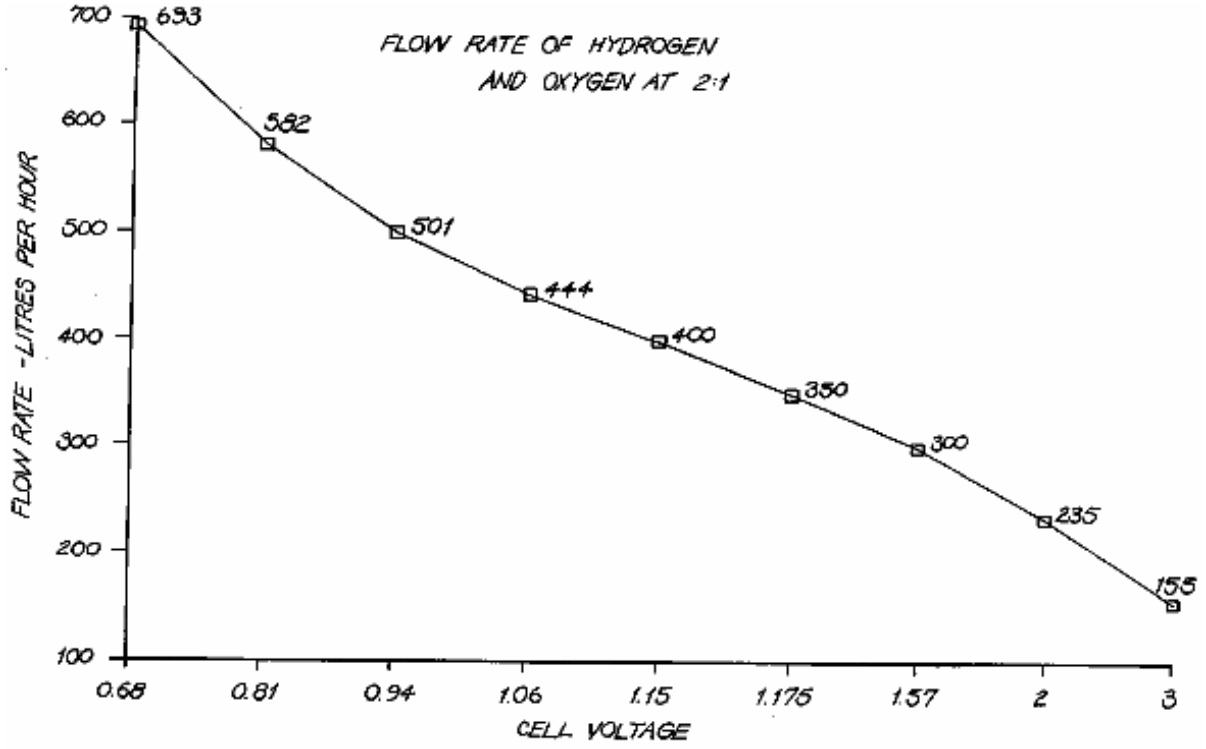

FLOW RATE OF HYDROGEN
AND OXYGEN AT 2:1
FLOW RATE - LITRES PER HOUR
CELL VOLTAGE
700
600
500
400
300
200
100
0.68
0.81
0.94
1.06
1.15
1.175
1.57
2
3
693
582
501
444
400
350
300
235
155

FIG. 18

TEST RUN	AMPS	VOLTS	TEMP C° (INITIAL)	TEMP C° (FINAL)	TIME (SECS.)	WATTS (A×V)	PRESSURE (psi)	FLOW RATE (lph)
1	47.2	38.5	40	-	-	1817.2	1 ATMOSPH.	-
2	54.3	34.4	-	70	222.13	1867.9	1 ATMOSPH.	89
3	65.2	34.4	40	70	26.37	2242.9	100-170	95
4	65.6	29.1	40	70	20.47	1909.0	300-410	97
5	62.9	29.4	40	70	22.93	1849.3	500-610	97
6	64.7	29.4	40	70	24.19	1902.2	700-850	98
7	63.9	29.2	40	70	24.13	1865.9	900-1050	98
8	64.0	29.3	40	70	22.87	1875.2	1100-1250	96
9	64.4	29.3	40	70	21.83	1886.9	1300-1450	98
10	63.7	29.1	40	70	23.34	1853.7	1500-1660	99
11	62.7	29.7	40	70	12.76	1862.2	1700-1890	100
12	61.9	29.9	40	70	11.17	1850.8	1900-1990	-
13	61.7	30.0	40	70	11.19	1851.0	2090-2170	-
14	60.7	30.6	40	70	15.71	1857.4	2290-2400	-
15	66.6	29.9	40	70	-	1991.3	2280-2420	-
16	61.7	30.0	45	70	-	1851.0	2270-2390	-
17	62.5	30.0	57	70	-	1875.0	2350-2380	-
18	62.0	30.1	59	70	-	1866.2	2350-2390	-
19	62.9	29.9	-	-	-	1880.7	2400-2420	-
20	63.0	29.4	-	-	-	1852.2	2430-2450	-

FIG. 19

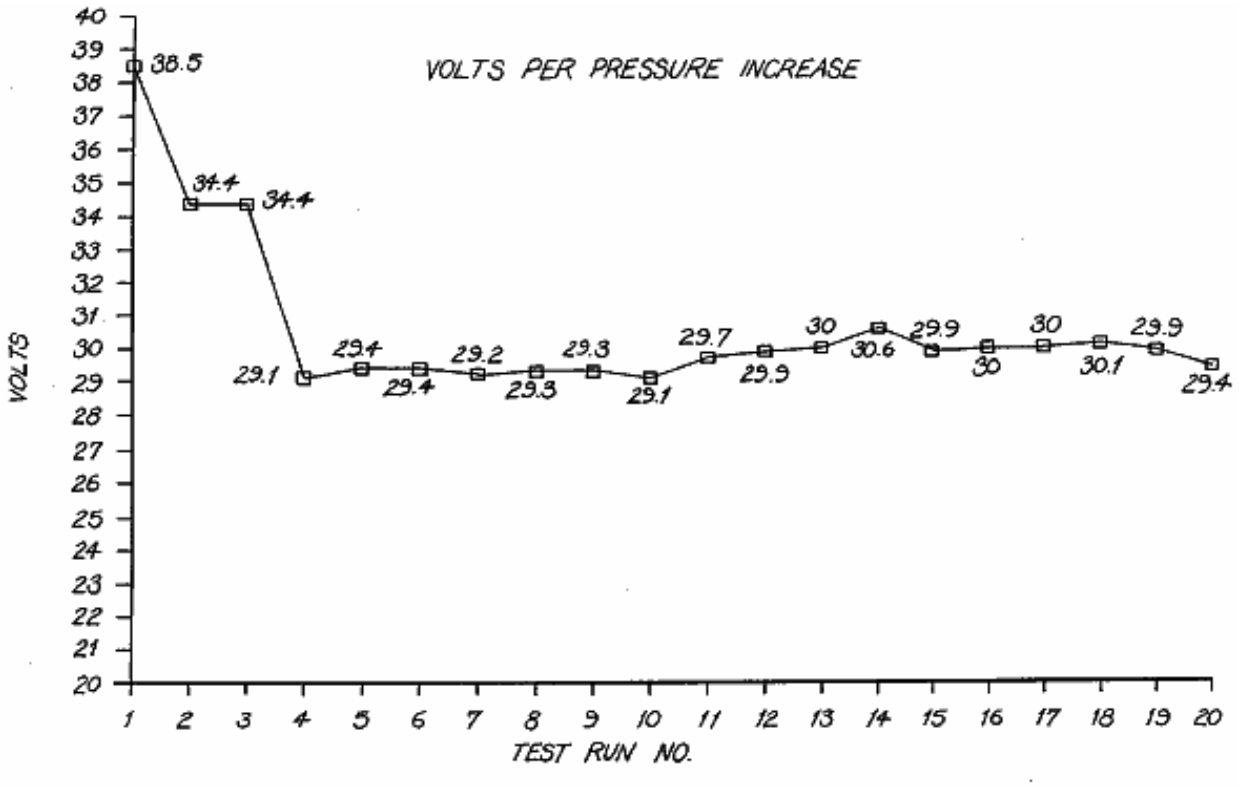


FIG. 20

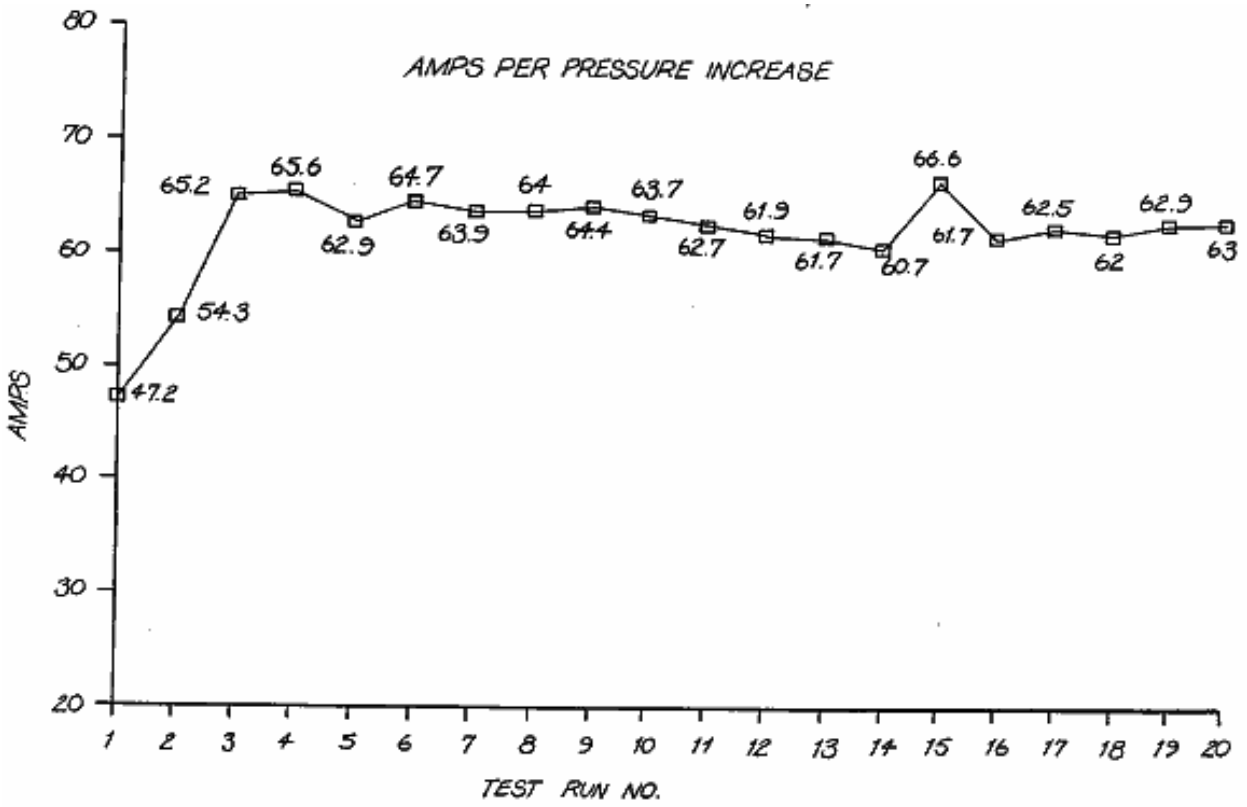

AMPS PER PRESSURE INCREASE
AMPS
TEST RUN NO.
80
70
60
50
40
30
20
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
47.2
54.3
65.2
65.6
62.9
64.7
63.9
64
64.4
63.7
62.7
61.9
61.7
60.7
66.6
61.7
62.5
62
62.9
63

FIG. 21

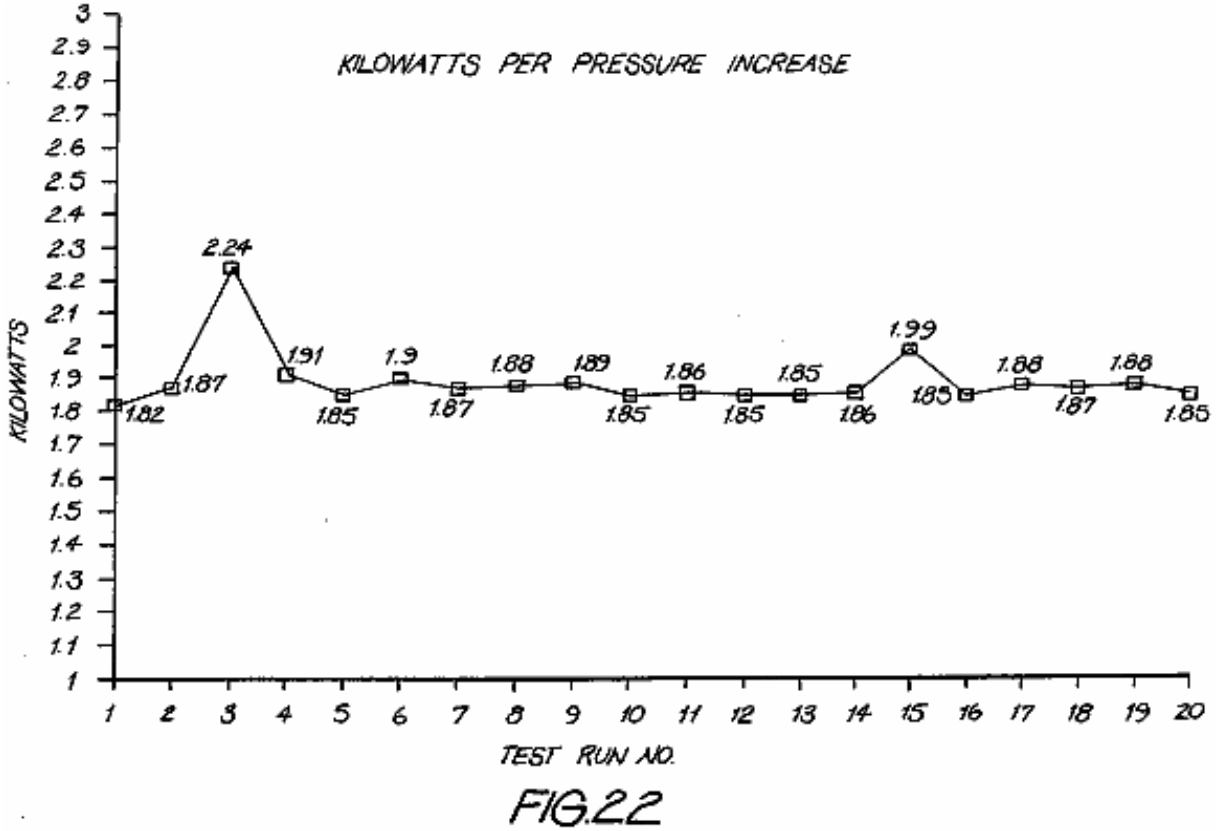

KILOWATTS PER PRESSURE INCREASE
KILOWATTS
TEST RUN NO.
3
2.9
2.8
2.7
2.6
2.5
2.4
2.3
2.2
2.1
2
1.9
1.8
1.7
1.6
1.5
1.4
1.3
1.2
1.1
1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
1.82
1.87
2.24
1.91
1.85
1.9
1.87
1.88
1.89
1.85
1.86
1.85
1.85
1.86
1.99
1.85
1.88
1.87
1.88
1.85

FIG. 22

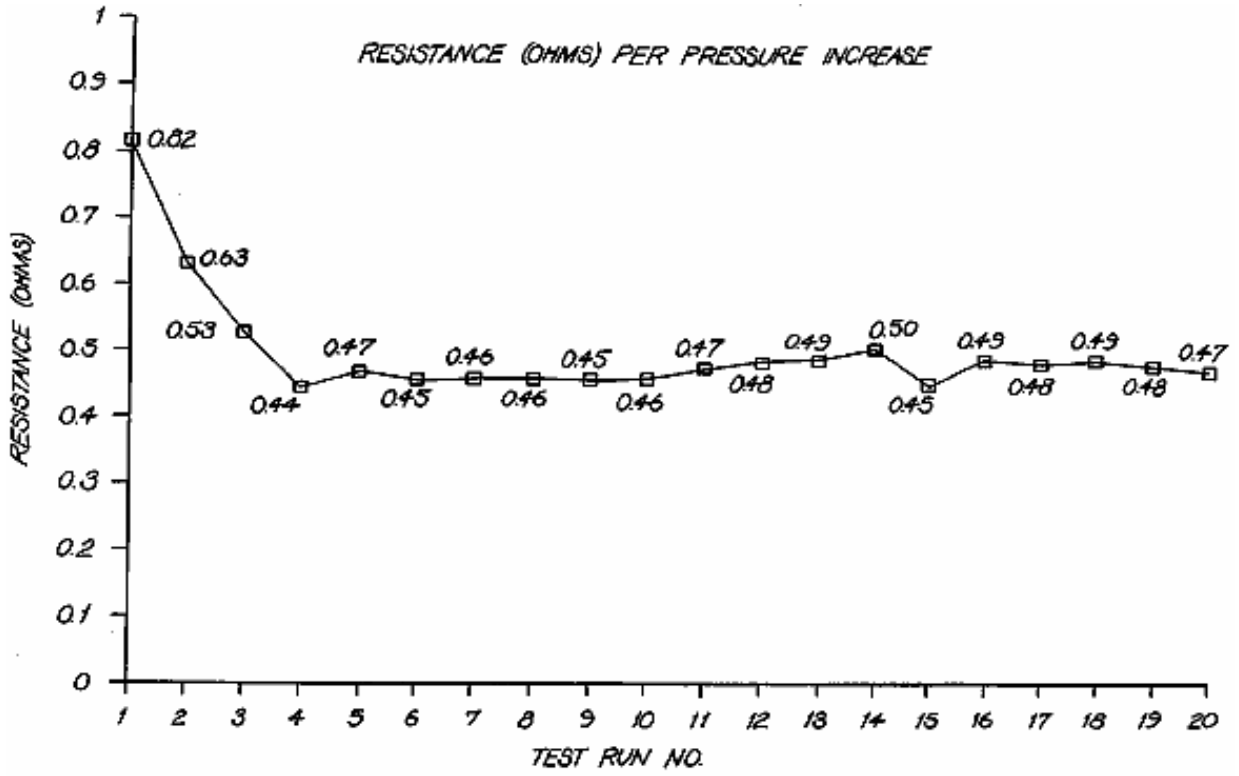

RESISTANCE (OHMS) PER PRESSURE INCREASE
RESISTANCE (OHMS)
TEST RUN NO.
0.82
0.63
0.53
0.44
0.47
0.45
0.46
0.46
0.45
0.46
0.47
0.48
0.49
0.50
0.45
0.49
0.48
0.49
0.48
0.47

FIG.23

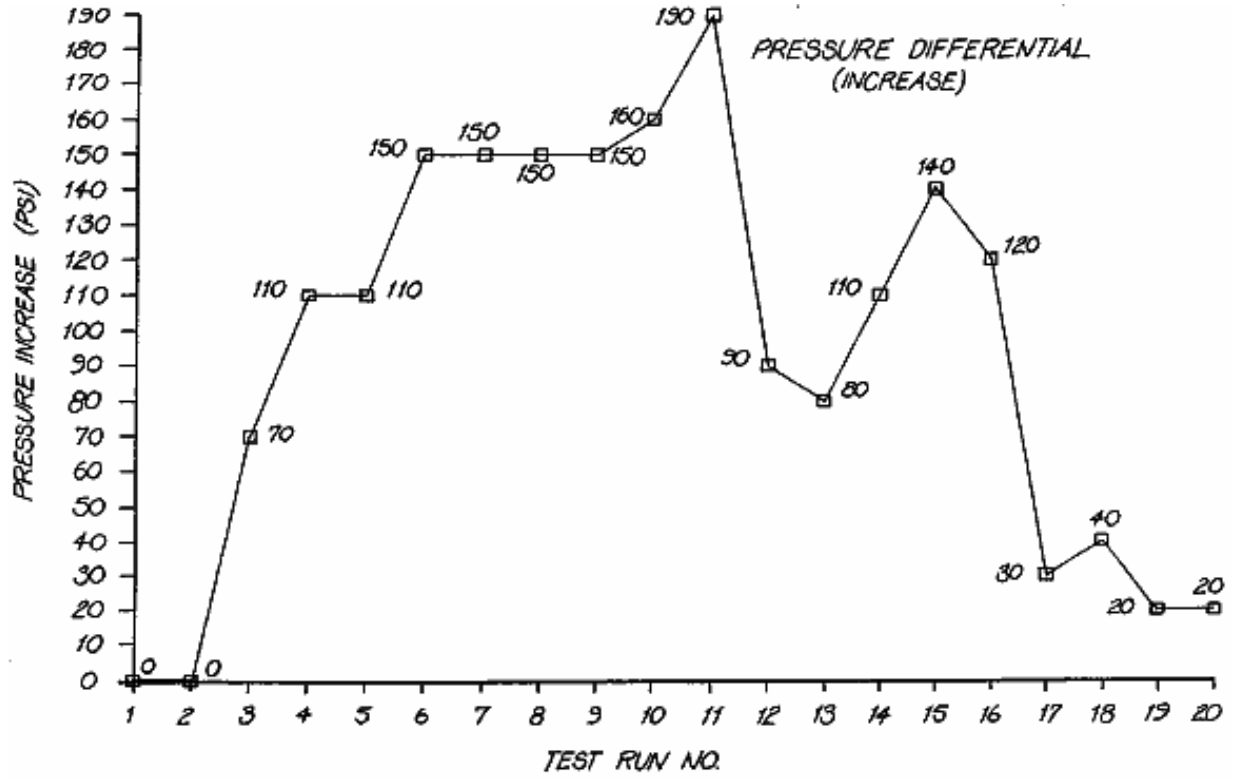

PRESSURE DIFFERENTIAL (INCREASE)
PRESSURE INCREASE (PSI)
TEST RUN NO.
0
0
70
110
110
150
150
150
150
160
190
90
80
110
140
120
30
40
20
20

FIG.24

FLOW RATE ANALYSIS PER PRESSURE INCREASE

RUN	VOLTS	AMPS	TEMP C°	TIME-SECS	VOLUME (LITRES)	LPH	PRESSURE PSI
1	27.5	49.7	70	114.0	2.8	88	14.7
2	34.4	54.3	70	222.13	5.49	89	14.7
3	20.5	51.9	87	190.0	4.7	89	50
4	20	55	80	33.0	1.0	109	170
5	34.4	65.2	70	26.37	0.69	95	200
6	29.1	65.6	70	20.47	0.55	97	410
7	29.4	62.9	70	22.93	0.62	97	610
8	29.4	64.7	70	24.19	0.66	98	850
9	29.2	63.9	70	24.13	0.66	98	1050
10	29.3	64.0	70	22.37	0.61	98	1250
11	29.3	64.4	70	21.83	0.59	98	1450
12	29.1	63.7	70	23.34	0.64	99	1660
13	29.7	62.7	70	12.76	0.35	100	1890

FIG. 25

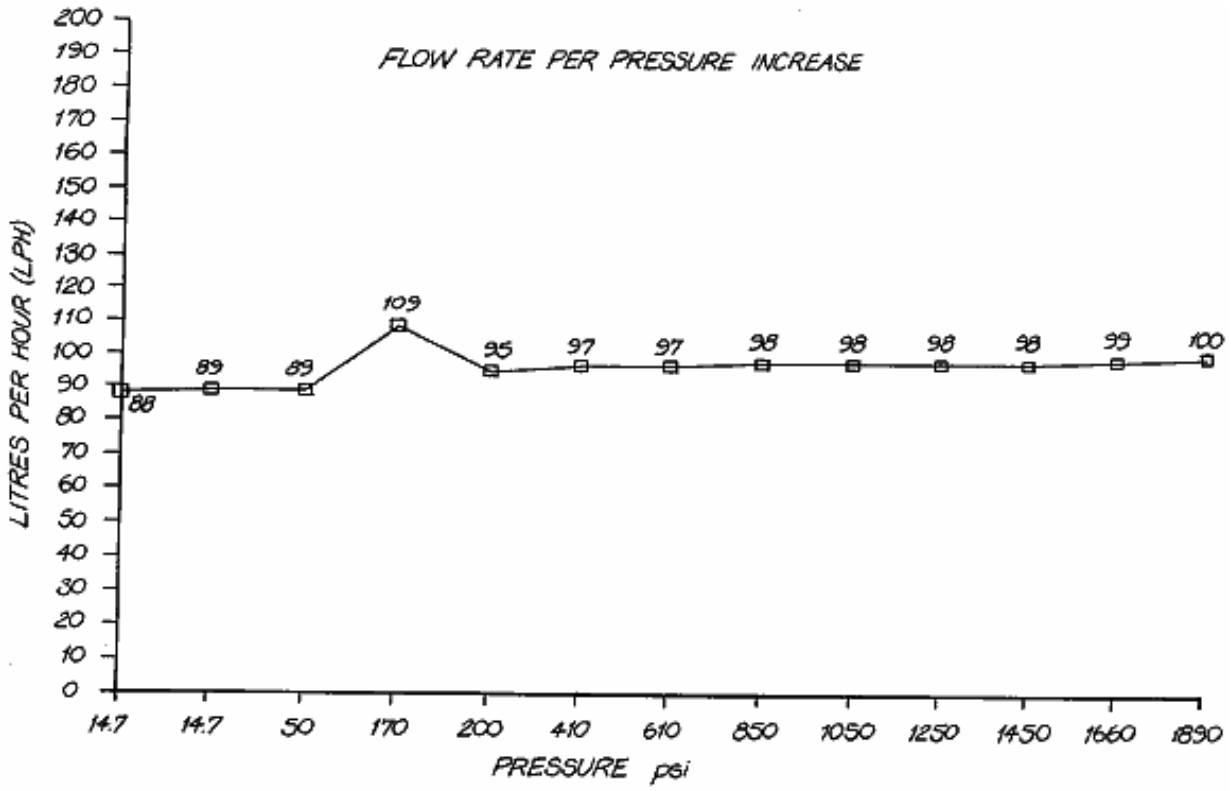


FIG. 26

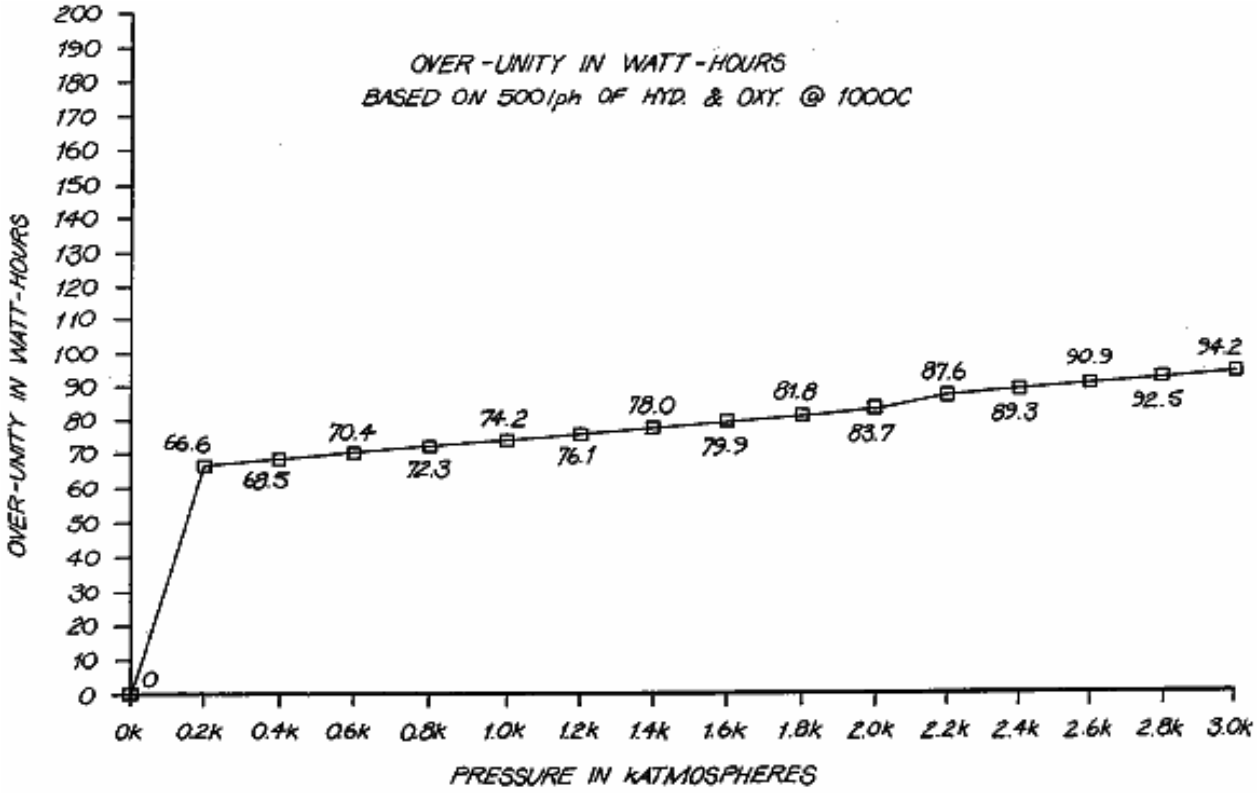

OVER-UNITY IN WATT-HOURS
BASED ON 500 l/ph OF HYD. & OXY. @ 1000C
OVER-UNITY IN WATT-HOURS
PRESSURE IN kATMOSPHERES
0
66.6
68.5
70.4
72.3
74.2
76.1
78.0
79.9
81.8
83.7
87.6
89.3
90.9
92.5
94.2
0k
0.2k
0.4k
0.6k
0.8k
1.0k
1.2k
1.4k
1.6k
1.8k
2.0k
2.2k
2.4k
2.6k
2.8k
3.0k

FIG.27

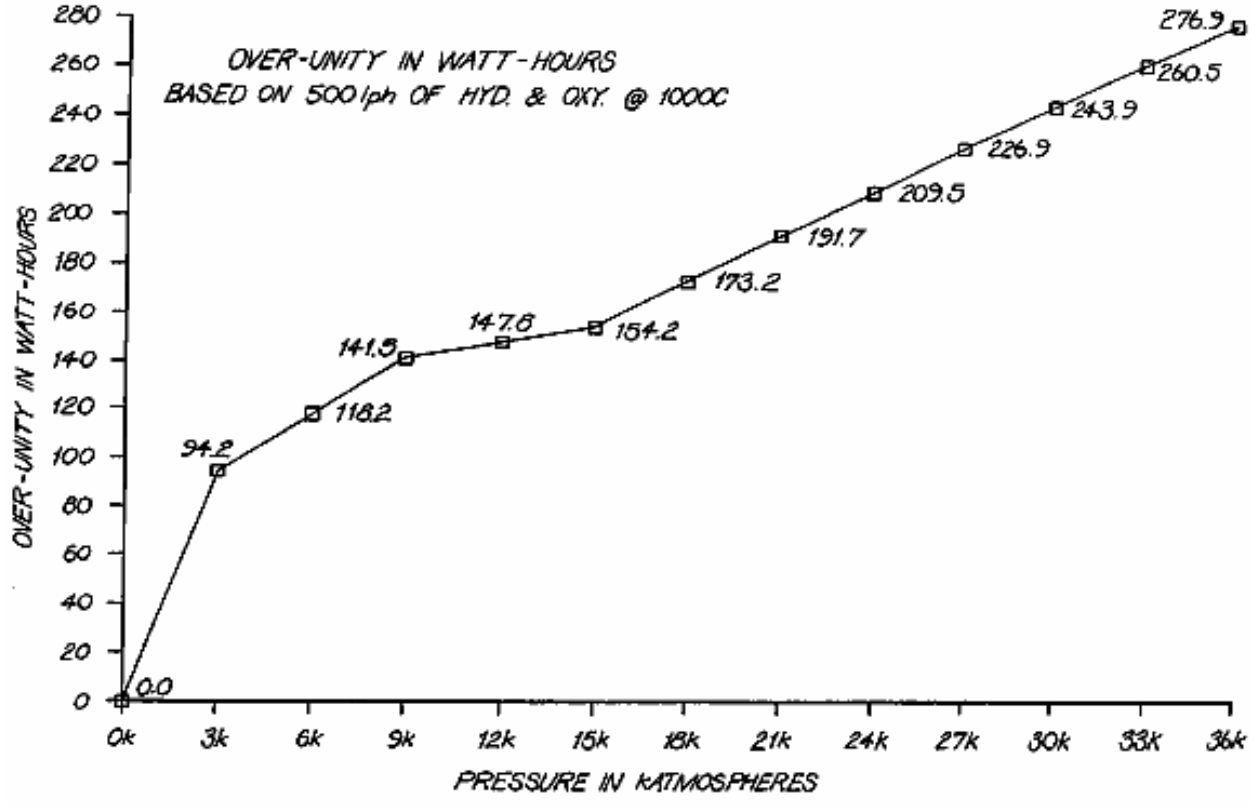

OVER-UNITY IN WATT-HOURS
BASED ON 500 l/ph OF HYD. & OXY. @ 1000C
OVER-UNITY IN WATT-HOURS
PRESSURE IN kATMOSPHERES
0.0
94.2
118.2
141.5
147.6
154.2
173.2
191.7
209.5
226.9
243.9
260.5
276.9
0k
3k
6k
9k
12k
15k
18k
21k
24k
27k
30k
33k
36k

FIG.28

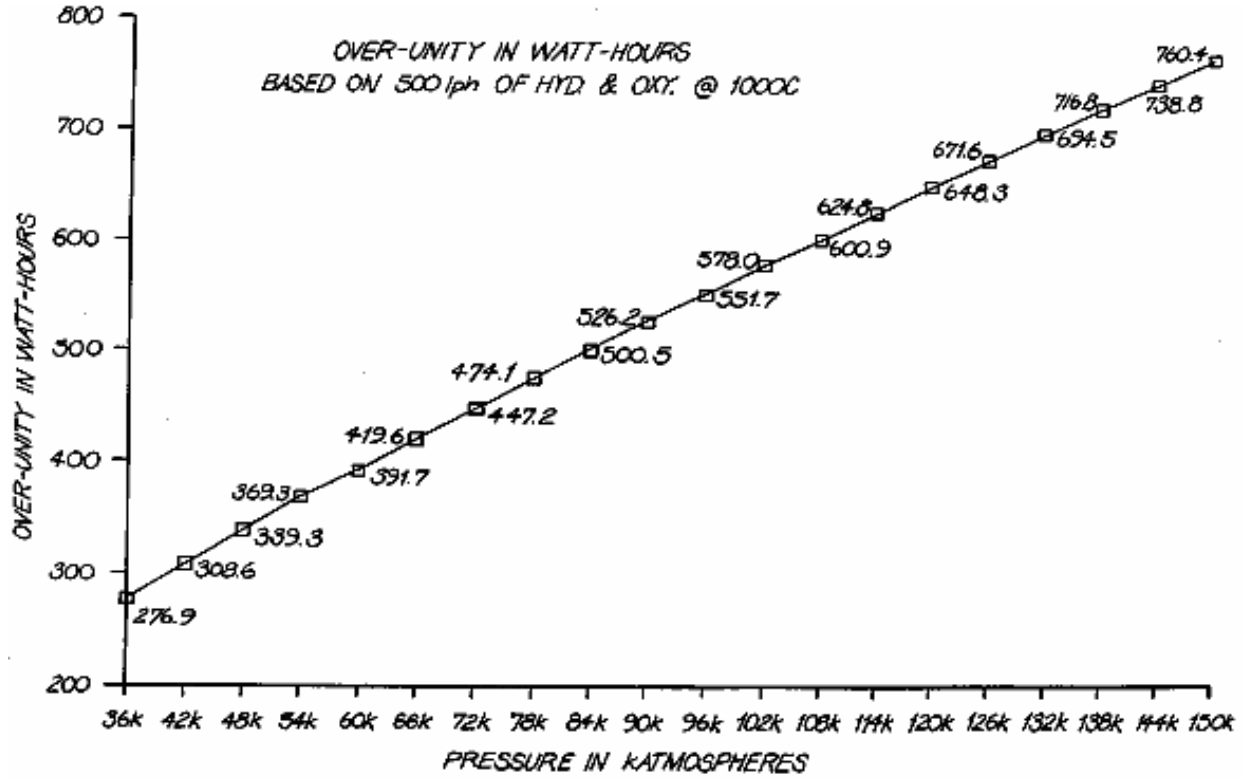


FIG. 29

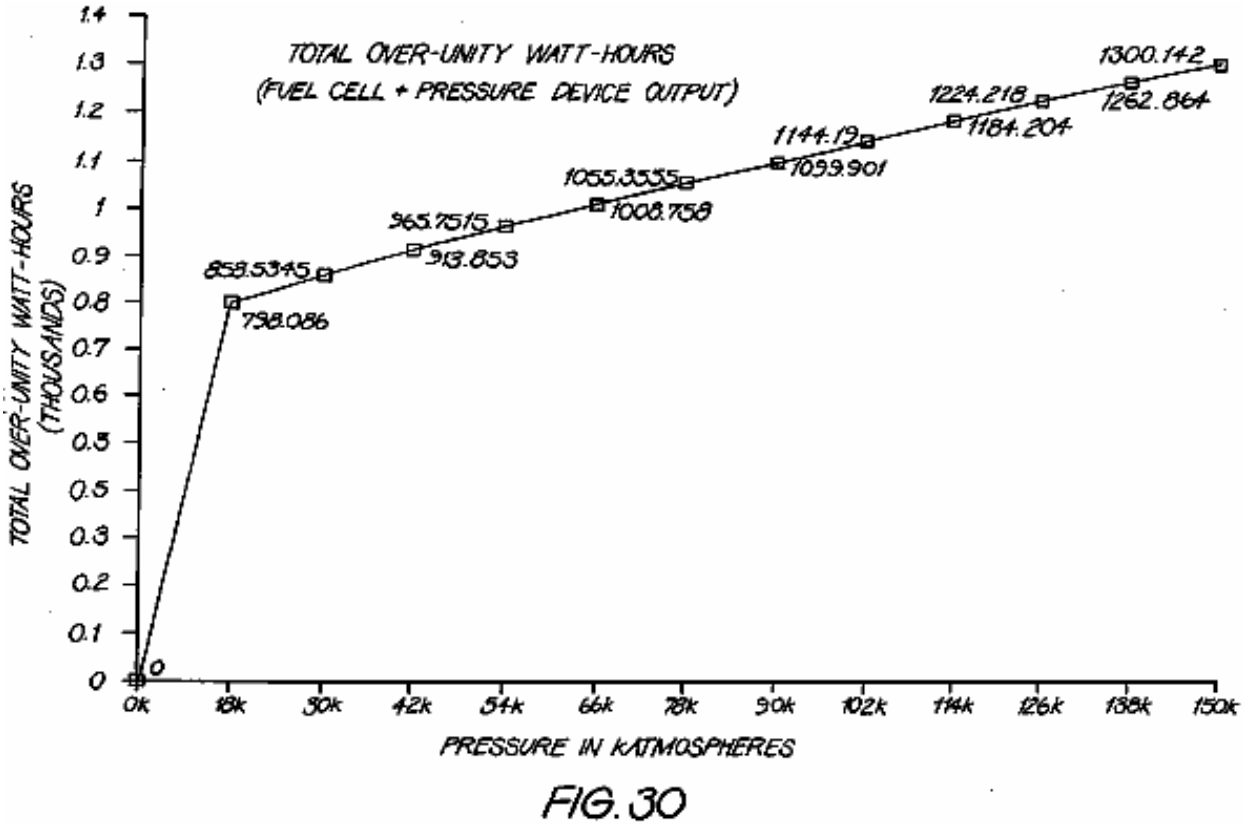


FIG. 30

Figs. 31a to 31e show a hydrogen/oxygen gas-driven internal combustion engine:

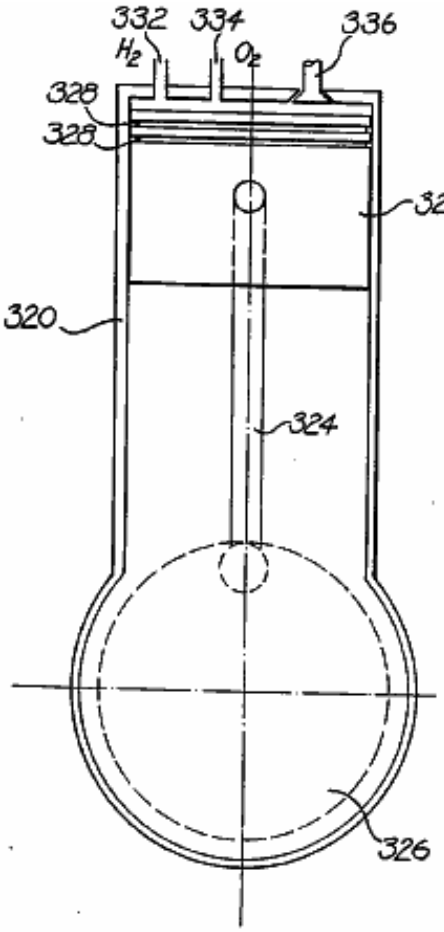


FIG. 31a

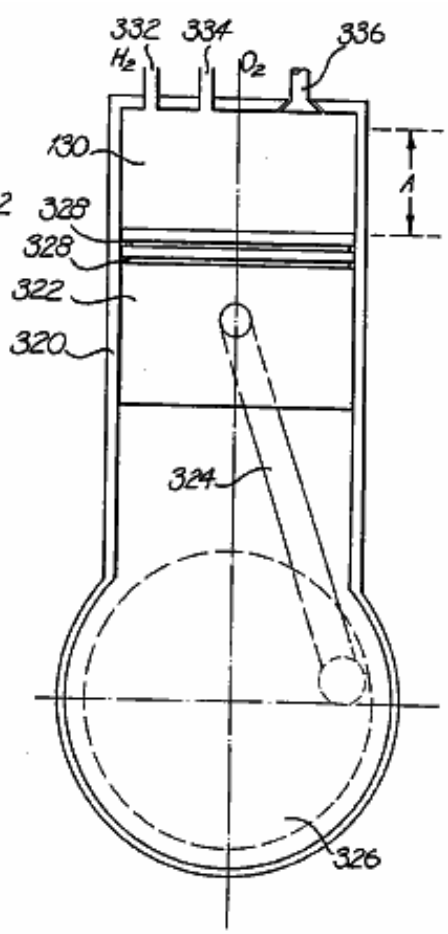


FIG. 31b

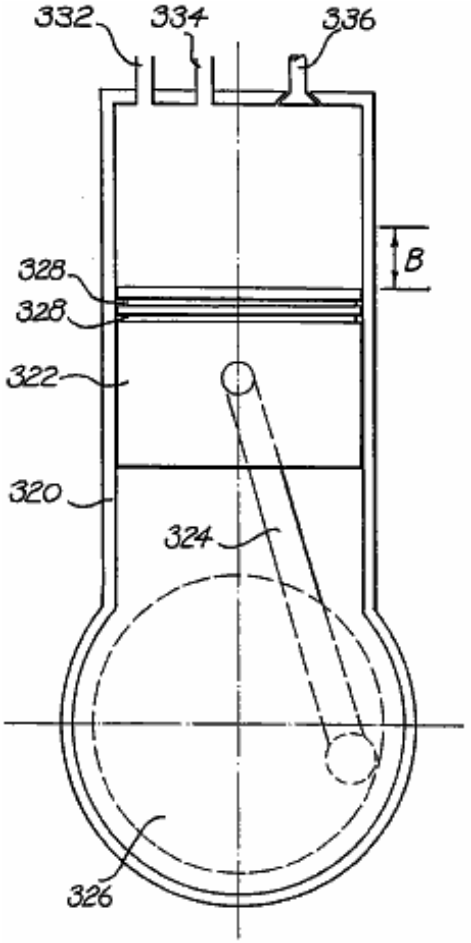


FIG. 31c

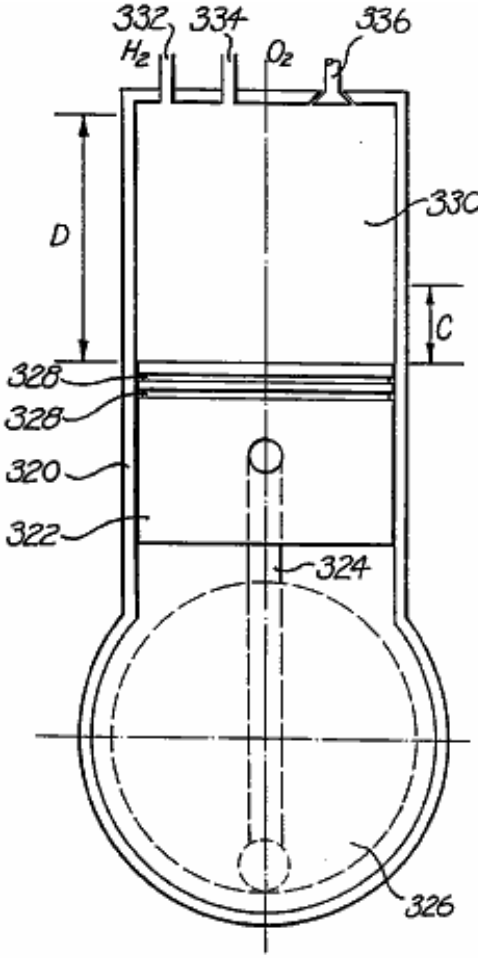


FIG. 31d

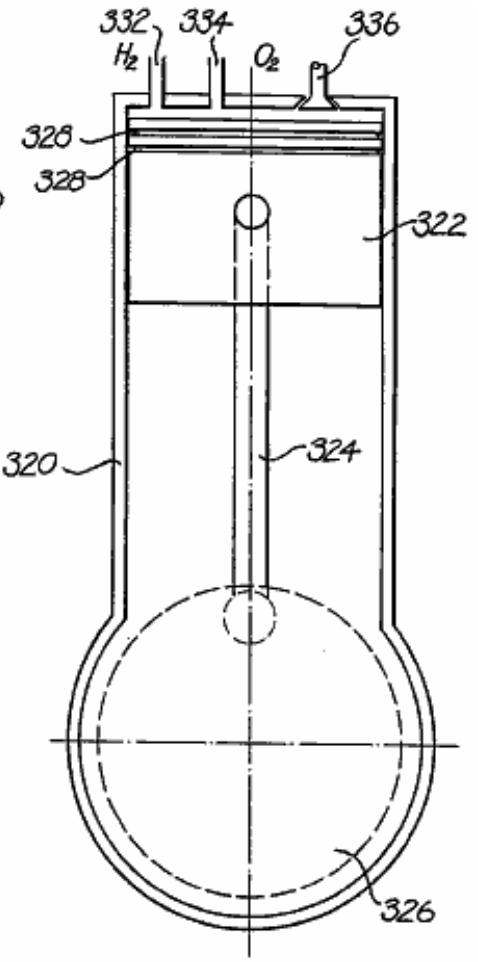


FIG. 31e

Figs. 32a-32c show a gas-driven implosion pump:

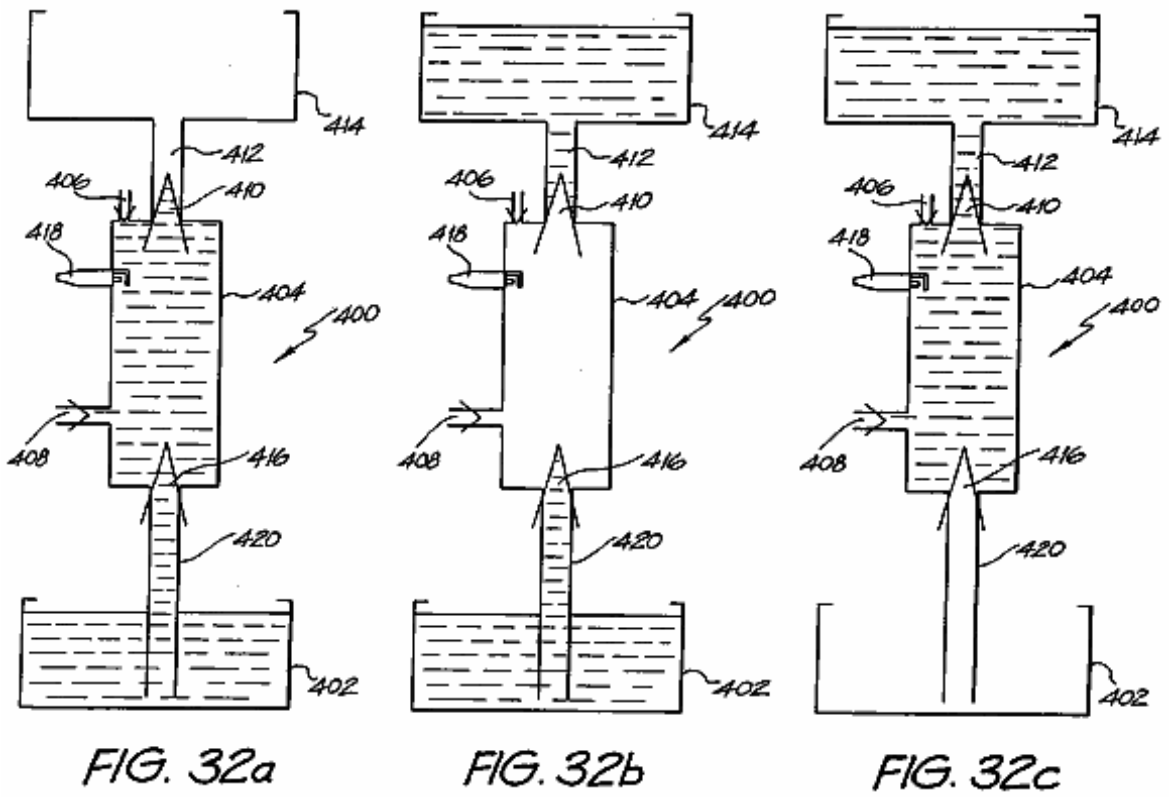


FIG. 32a FIG. 32b FIG. 32c

DETAILED DESCRIPTION AND BEST MODE OF PERFORMANCE

Fig.IA and **Fig.2A** show embodiments of a first and second type of cell plate **90, 98** as an end view. **Fig.1B** and **Fig.2B** are partial cross-sectional views along the respective mid-lines as shown. Common reference numerals have been used where appropriate. The plates **90, 98** can have the function of either an anode (+) or a cathode (-), as will become apparent. Each comprises an electrode disc **92** which is perforated with hexagonally shaped holes **96**. The disc **92** is made from steel or resin-bonded carbon or conductive polymer material. The disc **92** is housed in a circular rim or sleeve **94**. The function of the perforations **96** is to maximise the surface area of the electrode disc **92** and minimise the weight over solid constructions by 45%.

By way of example, for a disc of diameter 280 mm, the thickness of the disc must be 1 mm in order to allow the current density (which ranges from 90 A / 2,650 cm2 - 100 A / 2,940 cm2 of the anode or cathode) to be optimal. If the diameter of the plate is increased, which consequently increases the surface area, it is necessary to increase the thickness of the plate in order to maintain uniformity of conductance for the desired current density.

The hexagonal perforations in a 1 mm disc have a distance of 2 mm between the flats, twice the thickness of the plate in order to maintain the same total surface area prior to perforation, and be 1 mm away from the next adjacent perforation to allow the current density to be optimal. A (flat-to-flat) distance of 1 mm between the hexagonal perforations is required, because a smaller distance will result in thermal losses and a larger distance will add to the overall weight of the plate.

The sleeve **94** is constructed of PVC material and incorporates a number of equally spaced shaft holes **100,102**. The holes are for the passage of interconnecting shafts provided in a stacked arrangement of the plates **90, 98** forming the common conductor for the respective anode and cathode plates. The further two upper holes **104,106** each support a conduit respectively for the out-flow of oxygen and hydrogen gases. The further holes **108,110** at the bottom of the sleeve **94** are provided for the inlet of water and electrolyte to the respective cell plates **90, 98**.

Fig.3 shows an enlarged view of a portion of the cell plate **90** shown in **Fig.IA**. The port hole **104** is connected to the hexagonal perforations **96** within the sleeve **94** by an internal channel **112**. A similar arrangement is in place for the other port hole **106**, and for the water/electrolyte supply holes **108, 110**.

If it is the case that the hydrogen and oxygen gases liberated are to be kept separate (i.e. not to be formed as an admixture), then it is necessary to separate those gases as they are produced. In the prior art this is achieved by use of diaphragms which block the passage of gases and effectively isolate the water/electrolyte on each side of the diaphragm. Ionic transfer thus is facilitated by the conductive nature of the diaphragm material (i.e. a water - diaphragm - water path). This results in an increase in the ionic resistance and hence a reduction in efficiency.

Fig.4 shows an exploded stacked arrangement of four cell plates, being an alternative stacking of two (anode) cell plates **90** and two (cathode) cell plates **98**. The two ends of the stacked arrangement of cell plates delineates a single cell unit **125**.

Interposed between each adjacent cell plate **90, 98** is a PTFE separation **116**. Although not shown in **Fig.4**, the cell unit includes separate hydrogen and oxygen gas conduits that respectively pass through the stacked arrangement of cell plates via the port holes **106, 104** respectively. In a similar way, conduits are provided for the supply of water/electrolyte, respectively passing through the holes **108, 110** at the bottom of the respective plates **90, 98**. Only two pairs of anode/cathode cell plates are shown. The number of such plates can be greatly increased per cell unit **125**.

Also not shown are the interconnecting conductive shafts that electrically interconnect alternative common cell plates. The reason for having a large diameter hole in one cell plate adjacent to a smaller diameter hole in the next cell plate, is so that an interconnecting shaft will pass through the larger diameter hole, and not make an electrical connection (i.e. insulated with PVC tubing) rather only forming an electrical connection between alternate (common) cell plates.

Fig.4 is an exploded view of one cell unit **125** arrangement. When fully constructed, all the elements are stacked in intimate contact. Mechanical fastening is achieved by use of one of two adhesives such as (a) "PUR-FECT LOK" (TM) 34-9002, which is a Urethane Reactive Hot Melt adhesive with a main ingredient of Methylene Bispheny/Dirsocynate (MDI), and (b) "MY-T-BOND" (TM) which is a PVC solvent based adhesive. Both adhesives are Sodium Hydroxide resistant, which is necessary because the electrolyte contains 20% Sodium Hydroxide. In that case the water/electrolyte only resides within the area contained within the cell plate sleeve **94**. Thus the only path for the inlet of water/electrolyte is by bottom channels **118, 122** and the only outlet for the gases is by the top channels **112, 120**. In a system constructed and tested by the inventor, the thickness of the cell plates **90, 98** is 1 mm (2 mm on the rim because of the PVC sleeve **94**), with a diameter of 336 mm. The cell unit **125** is segmented from the next cell by an insulating PVC segmentation disc **114**. A segmentation disc **114** is also placed at the beginning and end of the entire cell bank. If there is to be no separation of the liberated gases, then the PTFE membranes **116** are omitted and sleeve **94** is not required.

The PTFE membrane **116** is fibrous and has 0.2 to 1.0 micron interstices. A suitable type is type Catalogue Code J, supplied by Tokyo Roshi International Inc (Advantec). The water/electrolyte fills the interstices and ionic current flows only via the water - there is no contribution of ionic flow through the PTFE material itself. This leads to a reduction in the resistance to ionic flow. The PTFE material also has a 'bubble point' that is a function of pressure, hence by controlling the relative pressures at either side of the PTFE separation sheets, the gases can be 'forced' through the interstices to form an admixture, or otherwise kept separate. Other advantages of this arrangement include a lesser cost of construction, improved operational efficiency and greater resistance to faults.

Fig.5A is a stylised, and exploded, schematic view of a linear array of three series-connected cell units **125**. For clarity, only six interconnecting shafts **126-131** are shown. The shafts **126-131** pass through the respective shaft holes **102,100** in the various cell plates **90,98** in the stacked arrangement. The polarity attached to each of the exposed end shafts, to which the DC supply is connected also is indicated. The shafts **126-131** do not run the full length of the three cell banks **125**. The representation is similar to the arrangement shown in **Fig.7A** and **Fig.8**. One third the full DC source voltage appears across each anode/cathode cell plate pair **90,98**.

Further, the gas conduits **132,133**, respectively for hydrogen and oxygen, that pass through the port holes **104,106** in the cell plates **90,98** also are shown. In a similar way, water/electrolyte conduits **134,135**, passing through the water port holes **108,110** in the cell plates also are shown.

Fig.5B particularly shows how the relative potential difference in the middle cell bank **125** changes. That is, the plate electrode **90a** now functions as a cathode (i.e. relatively more negative) to generate hydrogen, and the plate electrode **98a** now functions as an anode (i.e. relatively more positive) to generate oxygen. This is the case for every alternate cell unit. The arrowheads shown in **Fig.5B** indicate the electron and ionic current circuit. **Fig.5C** is an electrical equivalent circuit representation of **Fig.5B**, where the resistive elements represent the ionic resistance between adjacent anode/cathode plates. Thus it can be seen that the cell units are connected in series.

Because of the change of function of the cell plates **90a** and **98a**, the complementary gases are liberated at each, hence the respective channels **112** are connected to the opposite gas conduit **132,133**. Practically, this can be achieved by the simple reversal of the cell plates **90,98**.

Fig.6 shows the three cell units **125** of **Fig.5A** connected to a gas collection arrangement. The cell units **125** are located within a tank **140** which is filled with water/electrolyte to the indicated level **h**. The water is consumed as the electrolysis process proceeds, and replenishing supply is provided via the inlet **152**. The water/electrolyte level **h** can be viewed via the sight glass **154**. In normal operation, the different streams of hydrogen and oxygen are produced and passed from the cell units **125** to respective rising columns **142,144**. That is, the pressure of electrolyte on opposed sides of the PTFE membranes **116** is equalised, thus the gases cannot admix.

The columns **142,144** also are filled with the water/electrolyte, and as it is consumed at the electrode plates, replenishing supply of electrolyte is provided by way of circulation through the water/electrolyte conduits **134,135**. The circulation is caused by entrainment by the liberated gases, and by the circulatory inducing nature of the conduits and columns.

The upper extent of the tank **140** forms two scrubbing towers **156,158**, respectively for the collection of oxygen and hydrogen gases. The gases pass up a respective column **142,144**, and out from the columns via openings therein at a point within the interleaved baffles **146**. The point where the gases exit the columns **142,144** is beneath the water level **h**, which serves to settle any turbulent flow and entrained electrolyte. The baffles **146** located above the level **h** scrub the gas of any entrained electrolyte, and the scrubbed gas then exits by respective gas outlet columns **148,150** and so to a gas receiver. The level **h** within the tank **140** can be regulated by any convenient means, including a float switch, again with the replenishing water being supplied by the inlet pipe **152**.

The liberated gases will always separate from the water/electrolyte solution by virtue of the difference in densities. Because of the relative height of the respective set of baffles, and due to the density differential between the gases and the water/electrolyte, it is not possible for the liberated hydrogen and oxygen gases to mix. The presence of the full volume of water within the tank **140** maintains the cell plates in an immersed state, and further serves to absorb the shock of any internal detonations should they occur.

In the event that a gas admixture is required, then firstly the two flow valves **136,137** respectively located in the oxygen gas outlet conduit **132** and water/electrolyte inlet port **134** are closed. This blocks the outlet path for the oxygen gas and forces the inlet water/electrolyte to pass to the inlet conduit **134** via a one-way check valve **139** and pump **138**. The water/electrolyte within the tank **140** is under pressure by virtue of its depth (volume), and the pump **138** operates to increase the pressure of water/electrolyte occurring about the anode cell plates **90,98a** to be at an increased pressure with respect to the water/electrolyte on the other side of the membrane **116**.

This pressure differential is sufficient to cause the oxygen gas to migrate through the membrane, thus admixed oxygen and hydrogen are liberated via the gas output conduit **133** and column **144**. Since there is no return path for the water/electrolyte supplied by the pump **138**, the pressure about the cell plates **90,98a** will increase further, and to a point where the difference is sufficient such that the water/electrolyte also can pass through the membrane **116**. Typically, pressure differential in the range of 1.5 - 10 psi is required to allow passage of gas, and a pressure differential in the range of 10 - 40 psi for water/electrolyte.

While only three cell units **125** are shown, clearly any number, connected in series, can be implemented.

Embodiments of the present invention now will be described. Where applicable, like reference numerals have been used.

Fig.7A and **Fig.7B** show a first type of cell plate **190** respectively as an end view and as an enlarged cross-sectional view along line **VIIb-VIIb**. The cell plate **190** differs from the previous cell plate **90** shown in **Fig.1A** and **Fig.1B** in a number of important aspects. The region of the electrode disc **192** received within the sleeve **194** now is perforated. The function of these perforations is to further reduce the weight of the cell plate **190**. The shaft holes **200,202** again pass through the electrode disc **192**, but so too do the upper holes **204,206** through which the conduits for the out-flow of liberated hydrogen and oxygen gases pass. The bottom holes **208,210**, provided for the inlet of water and electrolyte, now also are located in the region of the sleeve **194** coincident with the perforated edge margin of the electrode disc **192**. The channels **212,218** respectively communicating with the port hole **204** and the supply hole **210** also are shown.

Fig.8A and **Fig.8B** show a second type of cell plate **198** as a companion to the first cell plate **190**, and as the same respective views. The second cell plate **198** is somewhat similar to the cell plate **98** previously shown in **Fig.2A** and **Fig.2B**. The differences between them are the same as the respective differences between the cell

plate shown in **Fig.1A** and **Fig.1B** and the one shown in **Fig.7A** and **Fig.7B**. The arrangement of the respective channels **220,222** with respect to the port **206** and the water supply hole **208** also are shown.

In the fabrication of the cell plates **190,198**, the sleeve **94** is injection moulded from PVC plastics material formed about the edge margin of the electrode disc **192**.

The injection moulding process results in the advantageous forming of interconnecting sprues forming within the perforations **196** in the region of the disc **192** held within the sleeve **194**, thus firmly anchoring the sleeve **194** to the disc **192**.

Fig.9 is a view similar to **Fig.3**, but for the modified porting arrangement and perforations (shown in phantom where covered by the sleeve) of the region of the disc **192** within and immediately outside of the sleeve **194**.

Fig.10 shows a cell unit **225** in the form of an exploded alternating stacking of first and second cell plates **190,198**, much in the same manner as **Fig.4**. Only two pairs of anode/cathode cell plates are shown, however the number of such plates can be greatly increased per cell unit **225**. The membrane **216** preferably is type QR-HE silica fibre with the alternative being PTFE. Both are available from Tokyo Roshi International Inc. (Advantec) of Japan. Type QR-HE is a hydrophobic material having 0.2 to 1.0 micron interstices, and is capable of operation at temperatures up to $1,000^{0}$C. The cell unit **225** can be combined with other such cell units **225** to form an interconnected cell bank in the same manner as shown in **Fig.5A**, **Fig.5B** and **Fig.5C**.

Furthermore, the cell units can be put to use in a gas collection arrangement such as that shown in **Fig.6**. Operation of the gas separation system utilising the new cell plates **190,198** is in the same manner as previously described.

Fig.11 is an enlarged cross-sectional view of three cell plates in the vicinity of the oxygen port **204**. The cell plates comprise two of the first type of plate **190** shown in **Fig.7A** constituting a positive plate, and a single one of the second type of plate **198** shown in **Fig.8A** representing a negative plate. The location of the respective channels **212** for each of the positive cell plates **190** is shown as a dashed representation. The respective sleeves **194** of the three cell plates are formed from moulded PVC plastics as previously described, and in the region that forms the perimeter of the port **204** have a configuration particular to whether a cell plate is positive or negative. In the present case, the positive cell plates **190** have a flanged foot **230** that, in the assembled construction, form the contiguous boundary of the gas port **204**. Each foot **230** has two circumferential ribs **232** which engage corresponding circumferential grooves **234** in the sleeve **194** of the negative plate **198**.

The result of this arrangement is that the exposed metal area of the negative cell plates **198** always are insulated from the flow of oxygen gas liberated from the positive cell plates **190**, thus avoiding the possibility of spontaneous explosion by the mixing of the separated hydrogen and oxygen gases. This arrangement also overcomes the unwanted production of either oxygen gas or hydrogen gas in the gas port.

For the case of the gas port **206** carrying the hydrogen gas, the relative arrangement of the cell plates is reversed such that a flanged footing now is formed on the sleeve **194** of the other type of cell plate **198**. This represents the converse arrangement to that shown in **Fig.11**.

Fig.12A and **Fig.12B** show perspective side views of adjacent cell plates, with **Fig.12A** representing a positive cell plate **190** and **Fig.12B** representing a negative cell plate **198**. The gas port **206** thus formed is to carry hydrogen gas. The mating relationship between the flanged foot **230** and the end margin of the sleeve **194** of the positive cell plate **192** can be seen, particularly the interaction between the ribs **232** and the grooves **234**.

Fig.13 is a cross-sectional view of four cell plates formed into a stacked arrangement delimited by two segmentation plates **240**, together forming a cell unit **242**. Thus there are two positive cell plates **190** and two negative cell plates **198** in alternating arrangement. The cross-section is taken in the vicinity of a shaft hole **202** through which a negative conductive shaft **244** passes. The shaft **244** therefore is in intimate contact with the electrode discs **192** of the negative cell plates **198**. The electrodes discs **192** of the positive cell plates **190** do not extend to contact the shaft **244**. The sleeve **194** of the alternating negative cell plates **198** again have a form of flanged foot **246**, although in this case the complementarily shaped ribs and grooves are formed only on the sleeve of the negative cell plates **198**, and not on the sleeve **194** of the positive cell plates **190**. The segmentation plates **240** serve to delimit the stacked plates forming a single cell unit **242**, with ones of the cell units **242** being stacked in a linear array to form a cell bank such as has been shown in **Fig.5A**.

A threaded shaft nut **250** acts as a spacer between adjacent electrodes connecting with the shaft **244**. **Fig.14** is a perspective view of the shaft nut **250** showing the thread **252** and three recesses **254** for fastening nuts, screws or the like.

In all of **Figs.11 to 13**, the separation membrane material **216** is not shown, but is located in the spaces **248** between adjacent cell plates **190,198**, extending to the margins of the electrode disks **192** in the vicinity of the gas ports **204,206** or the shaft holes **200,202**.

An electrolysis hydrogen and oxygen gas system incorporating a gas separation system, such as has been described above, can therefore be operated to establish respective high pressure stores of gas. That is, the separated hydrogen and oxygen gases liberated by the electrolysis process are stored in separate gas receivers or pressure vessels. The pressure in each will increase with the continuing inflow of gas.

Fig.15 shows an idealised electrolysis system, comprising an electrolysis cell **150** that receives a supply of water to be consumed. The electrolysis process is driven by a DC potential (**Es**) **152**. The potential difference applied to the cell **150** therefore must be sufficient to electrolyse the water into hydrogen and oxygen gas dependent upon, inter alia, the water pressure **PC** and the back pressure of gas **PB** acting on the surface of the water, together with the water temperature **Tc**. The separate liberated hydrogen and oxygen gases, by a priming function, are pressurised to a high value by storage in respective pressure vessels **158,160**, being carried by gas lines **154,156**.

The pressurised store of gases then are passed to an energy conversion device that converts the flow of gas under pressure to mechanical energy (e.g. a pressure drop device **162**). This mechanical energy recovered **WM** is available to be utilised to provide useful work. The mechanical energy **WM** also can be converted into electrical form, again to be available for use.

The resultant exhausted gases are passed via lines **164,166** to a combustion chamber **168**. Here, the gases are combusted to generate heat **QR**, with the waste product being water vapour. The recovered heat **QR** can be recycled to the electrolysis cell to assist in maintaining the advantageous operating temperature of the cell.

The previously described combustion chamber **168** can alternatively be a fuel cell. The type of fuel cell can vary from phosphoric acid fuel cells through to molten carbonate fuel cells and solid oxide cells. A fuel cell generates both heat (**QR**) and electrical energy (**WE**), and thus can supply both heat to the cell **150** or to supplement or replace the DC supply (**Es**) **152**.

Typically, these fuel cells can be of the type LaserCell TM as developed by Dr Roger Billings, the PEM Cell as available from Ballard Power Systems Inc. Canada or the Ceramic Fuel Cell (solid oxide) as developed by Ceramic Fuel Cells Ltd., Melbourne, Australia.

It is, of course, necessary to replenish the pressurised store of gases, thus requiring the continuing consumption of electrical energy. The recovered electrical energy **WE** is in excess of the energy required to drive electrolysis at the elevated temperature and is used to replace the external electrical energy source **152**, thereby completing the energy loop after the system is initially primed and started.

The present inventor has determined that there are some combinations of pressure and temperature where the efficiency of the electrolysis process becomes advantageous in terms of the total energy recovered, either as mechanical energy by virtue of a flow of gas at high pressure or as thermal energy by virtue of combustion (or by means of a fuel cell), with respect to the electrical energy consumed, to the extent of the recovered energy exceeding the energy required to sustain electrolysis at the operational pressure and temperature. This has been substantiated by experimentation. This notion has been termed "over-unity".

"Over-unity" systems can be categorised as broadly falling into three types of physical phenomena:

(i) An electrical device which produces 100 Watts of electrical energy as output after 10 Watts of electrical energy is input thereby providing 90 Watts of overunity (electrical) energy.

(ii) An electro-chemical device such as an electrolysis device where 10 Watts of electrical energy is input and 8 Watts is output being the thermal value of the hydrogen and oxygen gas output. During this process, 2 Watts of electrical energy converted to thermal energy is lost due to specific inefficiencies of the electrolysis system. Pressure - as the over-unity energy - is irrefutably produced during the process of hydrogen and oxygen gas generation during electrolysis. Pressure is a product of the containment of the two separated gases. The Law of Conservation of Energy (as referenced in "Chemistry Experimental Foundations", edited by Parry, R.W.; Steiner, L.E.; Tellefsen, R.L.; Dietz, P.M. Chap. 9, pp. 199-200, Prentice-Hall, New Jersey" and "An Experimental Science", edited by Pimentel, G.C., Chap. 7, pp. 115-117, W.H. & Freeman Co. San Francisco)

is in equilibrium where the 10 watts of input equals the 8 watts thermal energy output plus the 2 watts of losses. However, this Law ends at this point. The present invention utilises the apparent additional energy being the pressure which is a by-product of the electrolysis process to achieve over-unity.

(iii) An electro-chemical device which produces an excess of thermal energy after an input of electrical energy in such devices utilised in "cold fusion" e.g. 10 watts of electrical energy as input and 50 watts of thermal energy as output.

The present invention represents the discovery of means by which the previously mentioned second phenomenon can be embodied to result in "over-unity" and the realisation of 'free' energy. As previously noted, this is the process of liberating latent molecular energy. The following sequence of events describes the basis of the availability of over-unity energy.

In a simple two plate (anode/cathode) electrolysis cell, an applied voltage differential of 1.57 DC Volts draws 0.034 Amps per cm^2 and results in the liberation of hydrogen and oxygen gas from the relevant electrode plate. The electrolyte is kept at a constant temperature of 40^0C, and is open to atmospheric pressure.

The inefficiency of an electrolytic cell is due to its ionic resistance (approximately 20%), and produces a by-product of thermal energy. The resistance reduces, as does the minimum DC voltage required to drive electrolysis, as the temperature increases. The overall energy required to dissociate the bonding electrons from the water molecule also decreases as the temperature increases. In effect, thermal energy acts as a catalyst to reduce the energy requirements in the production of hydrogen and oxygen gases from the water molecule. Improvements in efficiency are obtainable by way of a combination of thermal energy itself and the NaOH electrolyte both acting to reduce the resistance of the ionic flow of current.

Thermal 'cracking' of the water molecule is known to occur at $1{,}500^0$C, whereby the bonding electrons are dissociated and subsequently 'separate' the water molecule into its constituent elements in gaseous form. This thermal cracking then allows the thermal energy to become a consumable. Insulation can be introduced to conserve thermal energy, however there will always be some thermal energy losses.

Accordingly, thermal energy is both a catalyst and a consumable (in the sense that the thermal energy excites bonding electrons to a higher energetic state) in the electrolysis process. A net result from the foregoing process is that hydrogen is being produced from thermal energy because thermal energy reduces the overall energy requirements of the electrolysis system.

Referring to the graph titled "Flow Rate At A Given Temperature" shown in **Fig.16**, it has been calculated that at a temperature of $2{,}000^0$C, 693 litres of hydrogen/oxygen admixed gas (2:1) will be produced. The hydrogen content of this volume is 462 litres. At an energy content of 11 BTUs per litre of hydrogen, this then gives an energy amount of 5,082 BTUs (11 x 462). Using the BTU:kilowatt conversion factor of 3413:1, 5,082 BTUs of the hydrogen gas equate to 1.49 kW. Compare this with l kW to produce the 693 litres of hydrogen/oxygen (including 463 litres of hydrogen). The usage of this apparatus therefore identifies that thermal energy, through the process of electrolysis, is being converted into hydrogen. These inefficiencies, i.e. increased temperature and NaOH electrolyte, reduce with temperature to a point at approximately 1000^0C where the ionic resistance reduces to zero, and the volumetric amount of gases produced per kWh increases.

The lowering of DC voltage necessary to drive electrolysis by way of higher temperatures is demonstrated in the graph in **Fig.17** titled "The Effect of temperature on Cell Voltage".

The data in **Fig.16** and **Fig.17** have two sources. Cell voltages obtained from 0^0C up to and including 100^0C were those obtained by an electrolysis system as described above. Cell voltages obtained from 150^0C up to $2{,}000^0$C are theoretical calculations presented by an acknowledged authority in this field, Prof. J. O'M. Bockris. Specifically, these findings were presented in "Hydrogen Energy, Part A, Hydrogen Economy", Miami Energy Conference, Miami Beach, Florida, 1974, edited by T. Nejat Veziroglu, Plenum Press, pp. 371-379. These calculations appear on page 374.

By inspection of **Fig.17** and **Fig.18** (titled "Flow Rate of Hydrogen and Oxygen at 2:1"), it can be seen that as temperature increases in the cell, the voltage necessary to dissociate the water molecule is reduced, as is the overall energy requirement. This then results in a higher gas flow per kWh.

As constrained by the limitation of the materials within the system, the operationally acceptable temperature of the system is 1000^0C. This temperature level should not, however, be considered as a restriction. This temperature is based on the limitations of the currently commercially available materials. Specifically, this system can utilise material such as compressed Silica Fibre for the sleeve around the electrolysis plate and hydrophobic Silica Fibre

(part no. QR-100HE supplied by Tokyo Roshi International Inc., also known as "Advantec") for the diaphragm (as previously discussed) which separates the electrolysis disc plates. In the process of assembling the cells, the diaphragm material and sleeved electrolysis plates **190,198** are adhered to one another by using high-temperature-resistant silica adhesive (e.g. the "Aremco" product "Ceramabond 618" which has an operational tolerance specification of $1,000^0$C).

For the electrolysis cell described above, with the electrolyte at $1,000^0$C and utilising electrical energy at the rate of 1 kWh, 167 litres of oxygen and 334 litres of hydrogen per hour will be produced.

The silica fibre diaphragm **116** previously discussed separates the oxygen and hydrogen gas streams by the mechanism of density separation, and produce a separate store of oxygen and hydrogen at pressure. Pressure from the produced gases can range from 0 to 150,000 Atmospheres. At higher pressures, density separation may not occur. In this instance, the gas molecules can be magnetically separated from the electrolyte if required.

In reference to the experiments conducted by Messrs Hamann and Linton (S.D. Hamann and M. Linton, Trans. Faraday Soc. 62,2234-2241, specifically, page 2,240), this research has proven that higher pressures can produce the same effect as higher temperatures in that the conductivity increases as temperature and/or pressure increases. At very high pressures, the water molecule dissociates at low temperatures. The reason for this is that the bonding electron is more readily removed when under high pressure. The same phenomenon occurs when the bonding electrons are at a high temperature (e.g. $1,500^0$C) but at low pressures.

As shown in **Fig.15**, hydrogen and oxygen gases are separated into independent gas streams flowing into separate pressure vessels **158,160** capable of withstanding pressures up to 150,000 Atmospheres. Separation of the two gases thereby eliminates the possibility of detonation. It should also be noted that high pressures can facilitate the use of high temperatures within the electrolyte because the higher pressure elevates the boiling point of water.

Experimentation shows that 1 litre of water can yield 1,850 litres of hydrogen/oxygen (in a ratio of 2: 1) gas mix after decomposition, this significant differential(1:1,850) is the source of the pressure. Stripping the bonding electrons from the water molecule, which subsequently converts liquid into a gaseous state, releases energy which can be utilised as pressure when this occurs in a confined space.

A discussion of experimental work in relation to the effects of pressure in electrolysis processes can be obtained from "Hydrogen Energy, Part A, Hydrogen Economy Miami Energy Conference, Miami Beach, Florida, 1974, edited by T. Nejat Veziroglu, Plenum Press". The papers presented by F.C. Jensen and F.H. Schubert on pages 425 to 439 and by John B. Pangborn and John C. Sharer on pages 499 to 508 are of particular relevance.

Attention must be drawn to the above published material; specifically on page 434, third paragraph, where reference is made to "Fig.7 shows the effect of pressure on cell voltage...". Fig. 7 on page 436 ("Effect of Pressure on SFWES Single Cell") indicates that if pressure is increased, then so too does the minimum DC voltage.

These quotes were provided for familiarisation purposes only and not as demonstrable and empirical fact. Experimentation by the inventor factually indicates that increased pressure (up to 2,450 psi) in fact lowers the minimum DC voltage.

This now demonstrable fact, whereby increased pressure actually lowers minimum DC voltage, is further exemplified by the findings of Messrs. Nayar, Ragunathan and Mitra in 1979 which can be referenced in their paper: "Development and operation of a high current density high pressure advanced electrolysis cell".

Nayar, M.G.; Ragunathan, P. and Mitra, S.K. International Journal of Hydrogen Energy (Pergamon Press Ltd.), 1980, Vol. 5, pp. 65-74. Their Table 2 on page 72 expressly highlights this as follows: "At a Current density (ASM) of 7,000 and at a temperature of 80^0C, the table shows identical Cell voltages at both pressures of 7.6 kg/cm^2 and 11.0 kg/cm^2. But at Current densities of 5,000, 6,000, 8,000, 9,000 and 10,000 (at a temperature of 80^0C), the Cell voltages were lower at a pressure of 11.0 kg/cm^2 than at a pressure of 7.6 kg/cm^2. " The present invention thus significantly improves on the apparatus employed by Mr. M.G. Nayar, et al, at least in the areas of cell plate materials, current density and cell configuration.

In the preferred form the electrode discs **192** are perforated mild steel, conductive polymer or perforated resin bonded carbon cell plates. The diameter of the perforated holes **196** is chosen to be twice the thickness of the plate in order to maintain the same total surface area prior to perforation. Nickel was utilised in the noted prior art system. That material has a higher electrical resistance than mild steel or carbon, providing the present invention with a lower voltage capability per cell.

The previously mentioned prior art system quotes a minimum current density (after conversion from ASM to Amps per square cm.) at 0.5 Amps per cm^2. The present invention operates at the ideal current density, established by experimentation, to minimise cell voltage which is 0.034 Amps per cm^2.

When compared with the aforementioned system, an embodiment of the present invention operates more efficiently due to a current density improvement by a factor of 14.7, the utilisation of better conducting cell plate material which additionally lowers cell voltage, a lower cell voltage of 1.49 at 80^0C as opposed to 1.8 volts at 80^0C, and a compact and efficient cell configuration.

In order to further investigate the findings of Messrs. M.G. Nayer, et al, the inventor conducted experiments utilising much higher pressures. For Nayer, et al, the pressures were 7.6 kg/cm^2 to 11.0 kg/cm^2, whereas inventor's pressures were 0 psi to 2,450 psi in an hydrogen/oxygen admixture electrolysis system.

This electrolysis system was run from the secondary coil of a transformer set approximately at maximum 50 Amps and with an open circuit voltage of 60 Volts. In addition, this electrolysis system is designed with reduced surface area in order that it can be housed in an hydraulic container for testing purposes. The reduced surface area subsequently caused the gas production efficiency to drop when compared with previous (i.e. more efficient) prototypes. The gas flow rate was observed to be approximately 90 litres per hour at 70^0C in this system as opposed to 310 litres per hour at 70^0C obtained from previous prototypes. All of the following data and graphs have been taken from the table shown in **Fig.19**.

Referring to **Fig.20** (titled "Volts Per Pressure Increase"), it can be seen that at a pressure of 14.7 psi (i.e. 1 Atmosphere), the voltage measured as 38.5V and at a pressure of 2,450 psi, the voltage measured as 29.4V. This confirms the findings of Nayar et al that increased pressure lowers the system's voltage. Furthermore, these experiments contradict the conclusion drawn by F.C. Jensen and F.H. Schubert ("Hydrogen Energy, Part A, Hydrogen Economy Miami Energy Conference, Miami Beach, Florida, 1974, edited by T. Nejat Veziroglu, Plenum Press", pp 425 to 439, specifically Fig. 7 on page 434) being that "... as the pressure of the water being electrolysed increases, then so too does the minimum DC Voltage". As the inventor's experiments are current and demonstrable, the inventor now presents his findings as the current state of the art and not the previously accepted findings of Schubert and Jensen.

Referring to **Fig.21** (titled "Amps Per Pressure Increase"), it can be seen that at a pressure of 14.7 psi (i.e. 1 Atmosphere being Test Run No. 1), the current was measured as 47.2A and at a pressure of 2,450 psi (Test Run No. 20), the current was measured as 63A.

Referring to **Fig.22** (titled "Kilowatts Per Pressure Increase"), examination of the power from Test Run No. 1 (1.82 kW) through to Test Run No. 20 (1.85 kW) indicates that there was no major increase in energy input required at higher pressures in order to maintain adequate gas flow.

Referring to **Fig.23** (titled "Resistance (Ohms) Per Pressure Increase"), the resistance was calculated from Test Run No. 1 (0.82 ohms) to Test Run No. 20 (0.47 ohms). These data indicate that the losses due to resistance in the electrolysis system at high pressures are negligible.

Currently accepted convention has it that dissolved hydrogen, due to high pressures within the electrolyte, would cause an increase in resistance because hydrogen and oxygen are bad conductors of ionic flow. The net result of which would be that this would decrease the production of gases.

These tests indicate that the ions find their way around the H2 and O2 molecules within the solution and that at higher pressures, density separation will always cause the gases to separate from the water and facilitate the movement of the gases from the electrolysis plates. A very descriptive analogy of this phenomenon is where the ion is about the size of a football and the gas molecules are each about the size of a football field thereby allowing the ion a large manoeuvring area in which to skirt the molecule.

Referring to **Fig.24** (titled "Pressure Differential (Increase)"), it can be seen that the hydrogen/oxygen admixture caused a significant pressure increase on each successive test run from Test Run No. 1 to Test Run No. 11. Test Runs thereafter indicated that the hydrogen/oxygen admixture within the electrolyte solution imploded at the point of conception (being on the surface of the plate).

Referring again to the table of **Fig.19**, it can be noted the time taken from the initial temperature to the final temperature in Test Run No. 12 was approximately half the time taken in Test Run No. 10. The halved elapsed time (from 40^0C to 70^0C) was due to the higher pressure causing the hydrogen/oxygen admixture to detonate which subsequently imploded within the system thereby releasing thermal energy.

Referring to the table shown in **Fig.25** (titled "Flow Rate Analysis Per Pressure Increase"), these findings were brought about from flow rate tests up to 200 psi and data from **Fig.24**. These findings result in the data of **Fig.25** concerning gas flow rate per pressure increase. Referring to **Fig.25**, it can be seen that at a pressure of 14.7 psi (1 Atmosphere) a gas production rate of 88 litres per kWh is being achieved. At 1,890 psi, the system produces 100 litres per kWh. These findings point to the conclusion that higher pressures do not affect the gas production rate of the system, the gas production rate remains constant between pressures of 14.7 psi (1 Atmosphere) and 1,890 psi.

Inferring from all of the foregoing data, increased pressure will not adversely affect cell performance (gas production rate) in separation systems where hydrogen and oxygen gases are produced separately, nor as a combined admixture. Therefore, in an enclosed electrolysis system embodying the invention, the pressure can be allowed to build up to a predetermined level and remain at this level through continuous (on-demand) replenishment. This pressure is the over-unity energy because it has been obtained during the normal course of electrolysis operation without additional energy input. This over-unity energy (i.e. the produced pressure) can be utilised to maintain the requisite electrical energy supply to the electrolysis system as well as provide useful work.

The following formulae and subsequent data do not take into account the apparent efficiencies gained by pressure increase in this electrolysis system such as the gained efficiency factors highlighted by the previously quoted Hamann and Linton research. Accordingly, the over-unity energy should therefore be considered as conservative claims and that such claimed over-unity energy would in fact occur at much lower pressures.

This over-unity energy can be formalised by way of utilising a pressure formula as follows: $E = (P - P_O) V$ which is the energy (E) in Joules per second that can be extracted from a volume (V) which is cubic meters of gas per second at a pressure (P) measured in Pascals and where P_0 is the ambient pressure (i.e. 1 Atmosphere).

In order to formulate total available over-unity energy, we will first use the above formula but will not take into account efficiency losses. The formula is based on a flow rate of 500 litres per kWh at $1{,}000^0$C. When the gases are produced in the electrolysis system, they are allowed to self-compress up to 150,000 Atmospheres which will then produce a volume (V) of 5.07×10^{-8} m^3/sec.

Work [Joules/sec] = $((150\text{-}1) \times 10^8)$ 5.07×10^{-8} m^3/sec = 760.4 Watts

The graphs in **Figs.27-29** (Over-Unity in watt-hours) indicate over-unity energy available excluding efficiency losses. However, in a normal work environment, inefficiencies are encountered as energy is converted from one form to another.

The results of these calculations will indicate the amount of surplus- over-unity energy after the electrolysis system has been supplied with its required 1 kWh to maintain its operation of producing the 500 Iph of hydrogen and oxygen (separately in a ratio of 2:1).

The following calculations utilise the formula stated above, including the efficiency factor. The losses which we will incorporate will be 10% loss due to the energy conversion device (converting pressure to mechanical energy, which is represented by device **162** in **Fig.15**) and 5% loss due to the DC generator $\mathbf{W_e}$ providing a total of 650 watt-hours which results from the pressurised gases.

Returning to the 1 kWh, which is required for electrolysis operation, this 1 kWh is converted (during electrolysis) to hydrogen and oxygen. The 1 kWh of hydrogen and oxygen is fed into a fuel cell. After conversion to electrical energy in the fuel cell, we are left with 585 watt-hours due to a 65 % efficiency factor in the fuel cell (35 % thermal losses are fed back into electrolysis unit **150** via $\mathbf{Q_r}$ in **Fig.15**).

Fig.30 graphically indicates the total over-unity energy available combining a fuel cell with the pressure in this electrolysis system in a range from 0 kAtmospheres to 150 kAtmospheres. The data in **Fig.30** have been compiled utilising the previously quoted formulae where the watt-hours findings are based on incorporating the 1 kWh required to drive the electrolysis system, taking into account all inefficiencies in the idealised electrolysis system (complete the loop) and then adding the output energy from the pressurised electrolysis system with the output of the fuel cell. This graph thereby indicates the energy break-even point (at approximately 66 kAtmospheres) where the idealised electrolysis system becomes self-sustaining.

In order to scale up this system for practical applications, such as power stations that will produce 50 MW of available electrical energy (as an example), the required input energy to the electrolysis system will be 170 MW (which is continually looped).

The stores of high pressure gases can be used with a hydrogen/oxygen internal combustion engine, as shown in **Figs. 31A to 31E**. The stores of high pressure gases can be used with either forms of combustion engines having an expansion stroke, including turbines, rotary, Wankel and orbital engines. One cylinder of an internal combustion engine is represented, however it is usually, but not necessarily always the case, that there will be other cylinders in the engine offset from each other in the timing of their stroke. The cylinder **320** houses a piston head **322** and crank **324**, with the lower end of the crank **324** being connected with a shaft **326**. The piston head **322** has conventional rings **328** sealing the periphery of the piston head **322** to the bore of the cylinder **320**.

A chamber **330**, located above the top of the piston head **322**, receives a supply of regulated separated hydrogen gas and oxygen gas via respective inlet ports **332,334**. There is also an exhaust port **336** venting gas from the chamber **330**.

The engine's operational cycle commences as shown in **Fig.31A**, with the injection of pressurised hydrogen gas, typically at a pressure of 5,000 psi to 30,000 psi, sourced from a reservoir of that gas (not shown). The oxygen gas port **334** is closed at this stage, as is the exhaust port **336**. Therefore, as shown in **Fig.31B**, the pressure of gas forces the piston head **322** downwards, thus driving the shaft **326**. The stroke is shown as distance "**A**".

At this point, the oxygen inlet **334** is opened to a flow of pressurised oxygen, again typically at a pressure of 5,000 psi to 30,000 psi, the volumetric flow rate being one half of the hydrogen already injected, so that the hydrogen and oxygen gas within the chamber **330** are the proportion 2:1.

Conventional expectations when injecting a gas into a confined space (e.g. such as a closed cylinder) are that gases will have a cooling effect on itself and subsequently its immediate environment (e.g. cooling systems/refrigeration). This is not the case with hydrogen. The inverse applies where hydrogen, as it is being injected, heats itself up and subsequently heats up its immediate surroundings. This effect, being the inverse of other gases, adds to the efficiency of the overall energy equation when producing over-unity energy.

As shown in **Fig.31C**, the piston head **322** has moved a further stroke, shown as distance "**B**", at which time there is self-detonation of the hydrogen and oxygen mixture. The hydrogen and oxygen inlets **332,334** are closed at this point, as is the exhaust **336**.

As shown in **Fig.31D**, the piston head is driven further downwards by an additional stroke, shown as distance "**C**", to an overall stroke represented by distance "**D**". The added piston displacement occurs by virtue of the detonation.

As shown in **Fig.31E**, the exhaust port **336** is now opened, and by virtue of the kinetic energy of the shaft **326** (or due to the action of others of the pistons connected with the shaft), the piston head **322** is driven upwards, thus exhausting the waste steam by the exhaust port **336** until such time as the situation of **Fig.31E** is achieved so that the cycle can repeat.

A particular advantage of an internal combustion motor constructed in accordance with the arrangement shown in **Figs.31A to 31E** is that no compression stroke is required, and neither is an ignition system required to ignite the working gases, rather the pressurised gases spontaneously combust when provided in the correction proportion and under conditions of high pressure.

Useful mechanical energy can be extracted from the internal combustion engine, and be utilised to do work. Clearly the supply of pressurised gas must be replenished by the electrolysis process in order to allow the mechanical work to continue to be done. Nevertheless, the inventor believes that it should be possible to power a vehicle with an internal combustion engine of the type described in **Figs.31A to 31E**, with that vehicle having a store of the gases generated by the electrolysis process, and still be possible to undertake regular length journeys with the vehicle carrying a supply of the gases in pressure vessels (somewhat in a similar way to, and the size of, petrol tanks in conventional internal combustion engines).

When applying over-unity energy in the form of pressurised hydrogen and oxygen gases to this internal combustion engine for the purpose of providing acceptable ranging (i.e. distance travelled), pressurised stored gases as mentioned above may be necessary to overcome the problem of mass inertia (e.g. stop-start driving). Inclusion of the stored pressurised gases also facilitates the ranging (i.e. distance travelled) of the vehicle.

Over-unity energy (as claimed in this submission) for an average sized passenger vehicle will be supplied at a continual rate of between 20 kW and 40 kW. In the case of an over-unity energy supplied vehicle, a supply of water (e.g. similar to a petrol tank in function) must be carried in the vehicle.

Clearly electrical energy is consumed in generating the gases. However it is also claimed by the inventor that an over-unity energy system can provide the requisite energy thereby overcoming the problem of the consumption of

fossil fuels either in conventional internal combustion engines or in the generation of the electricity to drive the electrolysis process by coal, oil or natural gas generators.

Experimentation by the inventor shows that if 1,850 litres of hydrogen/oxygen gas mix (in a ratio of 2:1) is detonated, the resultant product is 1 litre of water and 1,850 litres of vacuum if the thermal value of the hydrogen and oxygen gas mix is dissipated. At atmospheric pressure, 1 litre of admixed hydrogen/oxygen (2:1) contains 11 BTUs of thermal energy. Upon detonation, this amount of heat is readily dissipated at a rate measured in microseconds which subsequently causes an implosion (inverse differential of 1,850:1). Tests conducted by the inventor at 3 atmospheres (hydrogen/oxygen gas at a pressure of 50 psi) have proven that complete implosion does not occur. However, even if the implosion container is heated (or becomes heated) to 400C, total implosion will still occur.

This now available function of idiosyncratic implosion can be utilised by a pump taking advantage of this action. Such a pump necessarily requires an electrolysis gas system such as that described above, and particularly shown in **Fig.6**.

Figs. 32A-32C show the use of implosion and its cycles in a pumping device **400**. The pump **400** is initially primed from a water inlet **406**. The water inlet **406** then is closed-off and the hydrogen/oxygen gas inlet **408** is opened.

As shown in **Fig.32B**, the admixed hydrogen/oxygen gas forces the water upward through the one-way check valve **410** and outlet tube **412** into the top reservoir **414**. The one-way check valves **410,416** will not allow the water to drop back into the cylinder **404** or the first reservoir **402**. This force equates to lifting the water over a distance. The gas inlet valve **408** then is closed, and the spark plug **418** detonates the gas mixture which causes an implosion (vacuum). Atmospheric pressure forces the water in reservoir **402** up through tube **420**.

Fig.32C shows the water having been transferred into the pump cylinder **404** by the previous action. The implosion therefore is able to 'lift' the water from the bottom reservoir **402** over a distance which is approximately the length of tube **420**.

The lifting capacity of the implosion pump is therefore approximately the total of the two distances mentioned. This completes the pumping cycle, which can then be repeated after the reservoir **402** has been refilled.

Significant advantages of this pump are that it does not have any diaphragms, impellers nor pistons thereby essentially not having any moving parts (other than solenoids and one-way check valves). As such, the pump is significantly maintenance free when compared to current pump technology.

It is envisaged that this pump with the obvious foregoing positive attributes and advantages in pumping fluids, semi-fluids and gases can replace all currently known general pumps and vacuum pumps with significant benefits to the end-user of this pump.

CLAIMS

1. A looped energy system for the generation of excess energy available to do work, said system comprising:
An electrolysis cell unit receiving a supply of water and for liberating separated hydrogen gas and oxygen gas by electrolysis due to a DC voltage applied across respective anodes and cathodes of said cell unit;
Hydrogen gas receiver means for receiving and storing hydrogen gas liberated by said cell unit;
Oxygen gas receiver means for receiving and storing oxygen gas liberated by said cell unit;
Gas expansion means for expanding said stored gases to recover expansion work; and
Gas combustion means for mixing and combusting said expanded hydrogen gas and oxygen gas to recover combustion work; and in which a proportion of the sum of the expansion work and the combustion work sustains electrolysis of said cell unit to retain operational gas pressure in said hydrogen and oxygen gas receiver means such that the energy system is self-sustaining and there is excess energy available from said sum of energies.

2. A looped energy system for the generation of excess energy available to do work, said system comprising:
An electrolysis cell unit receiving a supply of water and for liberating separated hydrogen gas and oxygen gas by electrolysis due to a DC voltage applied across respective anodes and cathodes of said cell unit;
Hydrogen gas receiver means for receiving and storing hydrogen gas liberated by said cell unit;
Oxygen gas receiver means for receiving and storing oxygen gas liberated by said cell unit;
Gas expansion means for expanding said stored gases to recover expansion work; and
Fuel cell means for recovering electrical work from said expanded hydrogen gas and oxygen gas; and wherein a proportion of the sum of the expansion work and the recovered electrical work sustains electrolysis of said cell

unit to retain operational gas pressure in said hydrogen and oxygen gas receiver means such that the energy system is self-sustaining and there is excess energy available from said sum of energies.

3. An energy system as claimed in Claim 1 or Claim 2 further comprising mechanical-to-electrical energy conversion means coupled to said gas expansion means to convert the expansion work to electrical expansion work to be supplied as said DC voltage to said cell unit.

4. An energy system as claimed in any one of the preceding claims wherein said water in said cell unit is maintained above a predetermined pressure by the effect of back pressure from said gas receiver means and above a predetermined temperature resulting from input heat arising from said combustion work and/or said expansion work.

5. A method for the generation of excess energy available to do work by the process of electrolysis, said method comprising the steps of:
Electrolysing water by a DC voltage to liberate separated hydrogen gas and oxygen gas;
Separately receiving and storing said hydrogen gas and oxygen gas in a manner to be self-pressuring;
Separately expanding said stores of gas to recover expansion work;
Combusting said expanded gases together to recover combustion work; and
Applying a portion of the sum of the expansion work and the combustion work as said DC voltage to retain operational gas pressures and sustain said electrolysing step, there thus being excess energy of said sum available.

6. A method for the generation of excess energy available to do work by the process of electrolysis, said method comprising the steps of:
Electrolysing water by a DC voltage to liberate separated hydrogen gas and oxygen gas;
Separately receiving and storing said hydrogen gas and oxygen gas in a manner to be self-pressuring;
Separately expanding said stores of gas to recover expansion work;
Passing said expanded gases together through a fuel cell to recover electrical work; and
Applying a portion of the sum of the expansion work and the recovered electrical work as said DC voltage to retain operational gas pressures and sustain said electrolysing step, there thus being excess energy of said sum available.

7. An internal combustion engine powered by hydrogen and oxygen comprising:
At least one cylinder and at least one reciprocating piston within the cylinder;
A hydrogen gas input port in communication with the cylinder for receiving a supply of pressurised hydrogen;
An oxygen gas input port in communication with the cylinder for receiving a supply of pressurised oxygen; and
An exhaust port in communication with the cylinder and wherein the engine is operable in a two-stroke manner whereby, at the top of the stroke, hydrogen gas is supplied by the respective inlet port to the cylinder driving the piston downwards, oxygen gas then is supplied by the respective inlet port to the cylinder to drive the cylinder further downwards, after which time self-detonation occurs and the piston moves to the bottom of the stroke and upwardly again with said exhaust port opened to exhaust water vapour resulting from the detonation.

8. An engine as claimed in Claim 7, wherein there are a plurality of said cylinder and an equal plurality of said pistons, said pistons being commonly connected to a shaft and relatively offset in stroke timing to co-operate in driving the shaft.

9. An implosion pump comprising a combustion chamber interposed, and in communication with, an upper reservoir and a lower reservoir separated by a vertical distance across which water is to be pumped, said chamber receiving admixed hydrogen and oxygen at a pressure sufficient to lift a volume of water the distance therefrom to the top reservoir, said gas in the chamber then being combusted to create a vacuum in said chamber to draw water from said lower reservoir to fill said chamber, whereupon a pumping cycle is established and can be repeated.

10. An implosion pump as claimed in Claim 9, further comprising conduit mean connecting a respective reservoir with said chamber and one-way flow valve means located in each conduit means to disallow reverse flow of water from said upper reservoir to said chamber and from said chamber to said lower reservoir.

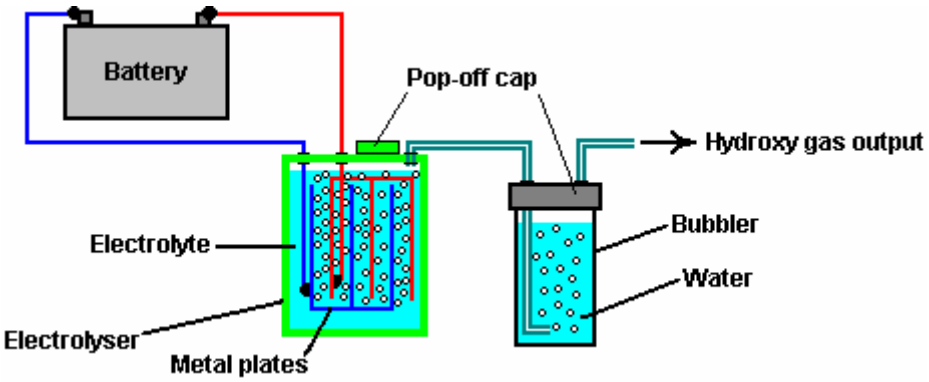
11. A parallel stacked arrangement of cell plates for a water electrolysis unit, the cell plates alternately forming an anode and cathode of said electrolysis unit, and said arrangement including separate hydrogen gas and oxygen gas outlet port means respectively in communication with said anode cell plates and said cathode call plates and extending longitudinally of said stacked plates, said stacked cell plates being configured in the region of said conduits to mate in a complementary manner to form said conduits such that a respective anode cell plate or cathode cell plate is insulated from the hydrogen gas conduit or the oxygen gas conduit.

12. An arrangement of cell plates as claimed in Claim 11, wherein said configuration is in the form of a flanged foot that extends to a flanged foot of the next adjacent like-type of anode or cathode cell plate respectively.

Henry Paine's HHO Fuel Conversion System

This is a very interesting patent which describes a simple system for overcoming the difficult problem of storing the hydrogen/oxygen gas mix produced by electrolysis of water. Normally this "hydroxy" gas mix is too dangerous to be compressed and stored like propane and butane are, but this patent states that hydroxy gas can be converted to a more benign form merely by bubbling it through a hydrocarbon liquid. Henry automatically speaks of turpentine in the patent, which strongly suggests that he used it himself, and consequently, it would probably be a good choice for any tests of the process.

This patent is more than 120 years old and has only recently been brought to the attention of the various "watercar" internet Groups. Consequently, it should be tested carefully before being used. Any tests should be done with extreme caution, taking every precaution against injury or damage should the mixture explode. It should be stressed that hydroxy gas is highly explosive, with a flame front speed far too fast to be contained by conventional commercial flashback arrestors. It is always essential to use a bubbler to contain any accidental ignition of the gas coming out of the electrolyser cell, as shown here:



For the purposes of a test of the claims of this patent, it should be sufficient to fill the bubbler with turpentine rather than water, though if possible, it would be good to have an additional bubbler container for the turpentine, in which case, the bubbler with the water should come between the turpentine and the source of the flame. Any tests should be done in an open space, ignited remotely and the person running the test should be well protected behind a robust object. A disadvantage of hydroxy gas is that it requires a very small orifice in the nozzle used for maintaining a continuous flame and the flame temperature is very high indeed. If this patent is correct, then the modified gas produced by the process should be capable of being used in any conventional gas burner.

US Letters Patent 308,276 18th November 1884 Inventor: Henry M. Paine

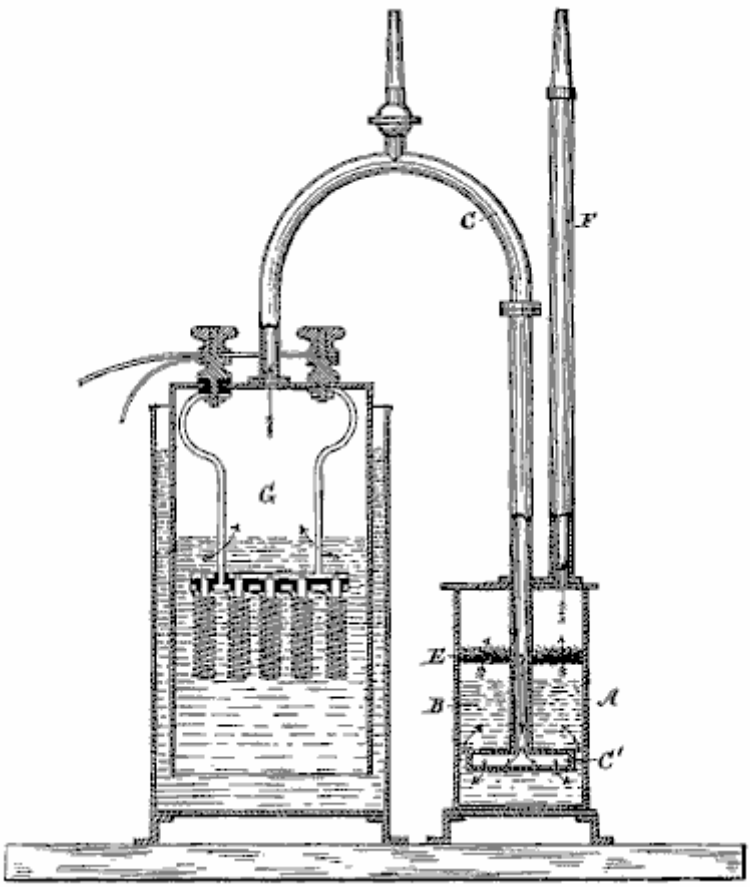
PROCESS OF MANUFACTURING ILLUMINATING GAS

To all whom it may concern:

Be it known that I, Henry M. Paine, a citizen of the United States, residing at Newark, in the county of Essex and State of New Jersey, have invented certain new and useful Improvements in the Process of Manufacturing Illuminating-Gas; and I do hereby declare the following to be a full, clear, and exact description of the invention, such as will enable others skilled in the art to which it appertains, to make and use the same, reference being had to the accompanying drawing, and to letters or figures of reference marked thereon, which form a part of this specification.

The present invention relates to the processes for manufacturing illuminating-gas, as explained and set forth here. Up to now, it has always been found necessary to keep the constituent gases of water separated from each other from the point of production to the point of ignition, as hydrogen and oxygen being present in the proper proportions for a complete reunion, form a highly-explosive mixture. Consequently, the two gases have either been preserved in separate holders and only brought together at the point of ignition, or else the hydrogen alone has been saved and the oxygen to support combustion has been drawn from the open air, and the hydrogen gas thus obtained has been carburetted by itself by passing through a liquid hydrocarbon, which imparts luminosity to the flame.

I have discovered that the mixed gases obtained by the decomposition of water through electrolysis can be used with absolute safety if passed through a volatile hydrocarbon; and my invention consists of the new gas thus obtained, and the process described here for treating the gas mixture whereby it is rendered safe for use and storage under the same conditions as prevail in the use of ordinary coal-gas, and is transformed into a highly-luminiferous gas.



In the accompanying drawing, which shows in sectional elevation, an apparatus adapted to carry out my invention, **G** is a producer for generating the mixed gases, preferably by the decomposition of water by an electric current. **A** is a tank partly filled with turpentine, camphene or other hydrocarbon fluid as indicated by **B**. The two vessels are connected by the pipe **C**, the end of which terminates below the surface of the turpentine, and has a broad mouthpiece **C'**, with numerous small perforations, so that the gas rises through the turpentine in fine streams or bubbles in order that it may be brought intimately in contact with the hydrocarbon.

Above the surface of the turpentine there may be a diaphragm **E**, of wire netting or perforated sheet metal, and above this, a layer of wool or other fibre packed sufficiently tightly to catch all particles of the hydrocarbon fluid which may be mechanically held in suspension, but loose enough to allow free passage of the gases. The pipe **F**, conducts the mixed gases off directly to the burners or to a holder.

I am aware that the hydrocarbons have been used in the manufacturer of water-gas from steam, and, as stated above, hydrogen gas alone has been carburetted; but I am not aware of any attempt being made to treat the explosive mixed gases in this manner.

Experiments have demonstrated that the amount of turpentine or other volatile hydrocarbon taken up by the gases in this process is very small and that the consumption of the hydrocarbon does not appear to bear any fixed ratio to the volume of the mixed gases passed through it. I do not, however, attempt to explain the action of the hydrocarbon on the gases.

What I claim as my invention and desire to secure by Letters Patent, is -

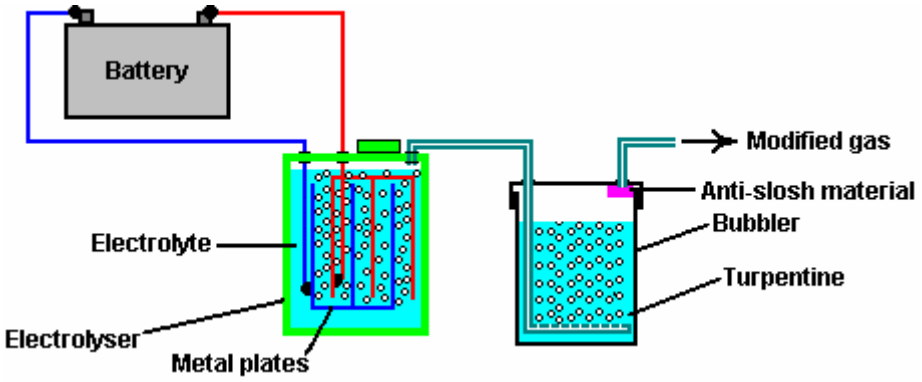
The process described here of manufacturing gas, which consists in decomposing water by electrolysis and conjointly passing the mixed constituent gases of water thus obtained, through a volatile hydrocarbon, substantially as and for the purpose set forth.

In testimony whereof I affix my signature in presence of two witnesses.

HENRY M. PAINE

Witnesses:
Bennet Osborne, Jr.,
W. E. Redding

Henry Paine's apparatus would therefor be:



Boris Volfson's Space Drive

US Patent 6,960,975 Nov.1, 2005 Inventor: Boris Volfson

SPACE VEHICLE PROPELLED BY THE PRESSURE OF INFLATIONARY VACUUM STATE

ABSTRACT

A space vehicle propelled by the pressure of inflationary vacuum state is provided comprising a hollow superconductive shield, an inner shield, a power source, a support structure, upper and lower means for generating an electromagnetic field, and a flux modulation controller. A cooled hollow superconductive shield is energised by an electromagnetic field resulting in the quantised vortices of lattice ions projecting a gravitomagnetic field that forms a space-time curvature anomaly outside the space vehicle. The space-time curvature imbalance, the space-time curvature being the same as gravity, provides for the space vehicle's propulsion. The space vehicle, surrounded by the space-time anomaly, may move at a speed approaching the light-speed characteristic for the modified locale.

US Patent References:

3626605	Dec., 1971	Wallace.
3626606	Dec., 1971	Wallace.
3823570	Jul., 1974	Wallace.
5197279	Mar., 1993	Taylor.
6353311	Mar., 2002	Brainard et al.

Other References:

M.T. French, "To the Stars by Electromagnetic Propulsion", http://www.mtjf.demon.co.uk/antigravp2.htm#cforce.

Evgeny Podkletnov, "Weak Gravitational Shielding Properties of Composite Bulk $YBa_2Cu_33O_{(7-x)}$ Superconductor Below 70K Under E.M. Field", LANL database number cond-mat/9701074, v. 3, 10 pages, Sep. 16, 1997.

N. LI & D.G. Torr, "Effects of a Gravitomagnetic Field on Pure Superconductors", Physical Review, vol. 43, p. 457, 3 pages, Jan. 15, 1991.

Evgeny Podkletnov, Giovanni Modanese "Impulse Gravity Generator Based on Charged $YBa_2Cu_33O_{7-y}$ Superconductor with Composite Crystal Structure", arXiv.org/physics database, #0108005 vol. 2, 32 pages, 8 figures, Aug. 30, 2001.

S. Kopeikin & E. Fomalont, "General Relativistic Model for Experimental Measurement of the Speed of Propagation of Gravity by VLBI", Proceedings of the 6th European VLBI Network Symposium Jun. 25-28, 2002, Bonn, Germany, 4 pages.

Sean M. Carroll, "The Cosmological Constant", http://pancake.uchicago.edu/˜ carroll/encyc/, 6 pages.

Chris Y. Taylor and Giovanni Modanese, "Evaluation of an Impulse Gravity Generator Based Beamed Propulsion Concept", American Institute of Aeronautics and Astronautics, Inc., 2002.

Peter L. Skeggs, "Engineering Analysis of the Podkletnov Gravity Shielding Experiment", Quantum Forum, Nov. 7, 1997, http://www.inetarena'.com/˜ noetic/pls/podlev.html).

BACKGROUND OF THE INVENTION

The existence of a magnetic-like gravitational field has been well established by physicists for general relativity, gravitational theories, and cosmology. The consequences of the effect of electromagnetically-affected gravity could be substantial and have many practical applications, particularly in aviation and space exploration.

There are methods known for converting electromagnetism into a propulsive force that potentially generates a large propulsive thrust. According to these methods, the machine thrust is produced by rotating, reciprocating masses in the following ways: centrifugal thrust, momentum thrust, and impulse thrust. ("To the Stars by Electromagnetic Propulsion", M. T. French, http://www.mtjf.demon.co.uk/antigravp2.htm#cforce).

However, the electromagnetic propulsion in an ambient space, or space that is not artificially modified, is not practical for interstellar travel because of the great distances involved. No interstellar travel is feasible without some form of distortion of space. In turn, no alteration of space is possible without the corresponding deformation of time. Gravitomagnetic alteration of space, resulting in the space-time curvature anomaly that could propel the space vehicle, could be a feasible approach to future space travel.

In the late 1940s, H. B. G. Casimir proved that the vacuum is neither particle nor field-free. It is a source of zero-point-fluctuation (ZPF) of fields such as the vacuum gravitomagnetic field. ZPF fields lead to real, measurable physical consequences such as the Casimir force. The quantised hand-made electromagnetic processes, such as those occurring in superconductors, affect the similarly quantised ZPFs. The most likely reason is the electron-positron creation and annihilation, in part corresponding to the "polarisation effect" sited by Evgeny Podkletnov in explaining the gravitomagnetic effect reportedly observed by him in 1992. ("Weak Gravitational Shielding Properties of Composite Bulk $YBa_2Cu_33O_{(7-x)}$ Superconductor Below 70 K Under E.M. Field", Evgeny Podkletnov, LANL database number cond-mat/9701074, v. 3, 10 pages, 16 Sep. 1997).

The investigation of gravitomagnetism, however, started well before Podkletnov. In the U.S. Pat. No. 3,626,605, Henry Wm. Wallace describes an experimental apparatus for generating and detecting a secondary gravitational field. He also shows how a time-varying gravitomagnetic field can be used to shield the primary background of a gravitoelectric field.

In the U.S. Pat. No. 3,626,606, Henry Wm. Wallace provides a variation of his earlier experiment. A type III-V semiconductor material, of which both components have unpaired nuclear spin, is used as an electronic detector for the gravitomagnetic field. The experiment demonstrates that the material in his gravitomagnetic field circuit has hysterisis and remanence effects analogous to magnetic materials.

In the U.S. Pat. No. 3,823,570, Henry Wm. Wallace provides an additional variation of his experiment. Wallace demonstrates that, by aligning the nuclear spin of materials having an odd number of nucleons, a change in specific heat occurs.

In the U.S. Pat. No. 5,197,279, James R. Taylor discloses Electromagnetic Propulsion Engine where solenoid windings generate an electromagnetic field that, without the conversion into a gravitomagnetic field, generates the thrust necessary for the propulsion.

In the U.S. Pat. No. 6,353,311 B1, John P. Brainard et al. offer a controversial theory of Universal Particle Flux Field, and in order to prove it empirically, provide a shaded motor-type device. This device is also intended for extracting energy from this hypothetical Field.

In the early 1980s, Sidney Coleman and F. de Luca noted that the Einsteinean postulate of a homogeneous Universe, while correct in general, ignores quantised local fluctuation of the pressure of inflationary vacuum state, this fluctuation causing local cosmic calamities. While the mass-less particles propagate through large portions of Universe at light speed, these anomaly bubbles, depending on their low or high relative vacuum density, cause a local increase or decrease of the propagation values for these particles. Scientists disagree about the possibility, and possible ways, to artificially create models of such anomalies.

In the early 1990s, Ning Li and D. G Torr described a method and means for converting an electromagnetic field into a gravitomagnetic field. Li and Torr suggested that, under the proper conditions, the minuscule force fields of superconducting atoms can "couple", compounding in strength to the point where they can produce a repulsion force ("Effects of a Gravitomagnetic Field on Pure Superconductors", N. Li and D. G. Torr, Physical Review, Volume 43, Page 457, 3 pages, 15 Jan. 1991).

A series of experiments, performed in the early 1990s by Podkletnov and R. Nieminen, reportedly resulted in a reduction of the weights of objects placed above a levitating, rotating superconductive disk subjected to high frequency magnetic fields. These results substantially support the expansion of Einstainean physics offered by Li & Torr. Podkletnov and Giovanni Modanese have provided a number of interesting theories as to why the weight reduction effect could have occurred, citing quantum gravitational effects, specifically, a local change in the cosmological constant. The cosmological constant, under ordinary circumstances, is the same everywhere. But, according to Podkletnov and Modanese, above a levitating, rotating superconductive disk exposed to high frequency magnetic fields, it is modified. ("Impulse Gravity Generator Based on Charged $YBa_2Cu_33O_{7-y}$ Superconductor with Composite Crystal Structure", Evgeny Podkletnov, Giovanni Modanese, arXiv.org/physics database, #0108005 volume 2, 32 pages, 8 figures, Aug. 30, 2001).

In the July 2004 paper, Ning Wu hypothesised that exponential decay of the gravitation gauge field, characteristic for the unstable vacuum such as that created by Podkletnov and Nieminen, is at the root of the gravitational shielding effects (Gravitational Shielding Effects in Gauge Theory of Gravity, Ning Wu, arXiv:hep-th/0307225 v 1 23 Jul. 2003, 38 pages incl. 3 figures, July 2004).

In 2002, Edward Fomalont and Sergei Kopeikin measured the speed of propagation of gravity. They confirmed that the speed of propagation of gravity matches the speed of light. ("General Relativistic Model for Experimental Measurement of the Speed of Propagation of Gravity by VLBI", S. Kopeikin and E. Fomalont, Proceedings of the 6th European VLBI Network Symposium Jun. 25-28 2002, Bonn, Germany, 4 pages).

String theory unifies gravity with all other known forces. According to String theory, all interactions are carried by fundamental particles, and all particles are just tiny loops of space itself forming the space-time curvature. Gravity and bent space are the same thing, propagating with the speed of light characteristic of the particular curvature. In light of the Fomalont and Kopeikin discovery, one can conclude that if there is a change in the speed of propagation of gravity within the space-time curvature, then the speed of light within the locality would also be affected.

In general relativity, any form of energy affects the gravitational field, so the vacuum energy density becomes a potentially crucial ingredient. Traditionally, the vacuum is assumed to be the same everywhere in the Universe, so the vacuum energy density is a universal number. The cosmological constant Lambda is proportional to the vacuum pressure:

ρ_Λ: $\Lambda=(8\pi G/3c^2)\rho_\Lambda$

Where:
G is Newton's constant of gravitation and
c is the speed of light
("The Cosmological Constant", Sean M. Carroll, http://pancake.uchicago.edu/~carroll/encyc/, 6 pages). Newer theories, however, permit local vacuum fluctuations where even the "universal" constants are affected:

$\Lambda_1=(8\pi G_1/3c_1^2)\rho_{\Lambda_1}$

Analysing physics laws defining the cosmological constant, a conclusion can be drawn that, if a levitating, rotating superconductive disk subjected to high frequency magnetic fields affects the cosmological constant within a locality, it would also affect the vacuum energy density. According to the general relativity theory, the gravitational attraction is explained as the result of the curvature of space-time being proportional to the cosmological constant. Thus, the change in the gravitational attraction of the vacuum's subatomic particles would cause a local anomaly in the curvature of the Einsteinean space-time.

Time is the fourth dimension. Lorentz and Einstein showed that space and time are intrinsically related. Later in his life, Einstein hypothesised that time fluctuates both locally and universally. Ruggero Santilli, recognised for expanding relativity theory, has developed the isocosmology theory, which allows for variable rates of time. Time is also a force field only detected at speeds above light speed. The energy of this force field grows as its propagation speed declines when approaching light-speed. Not just any light-speed: the light-speed of a locale. If the conditions of the locale were modified, this change would affect the local time rate relative to the rate outside the affected locale, or ambient rate. The electromagnetically-generated gravitomagnetic field could be one such locale modifier.

Analysing the expansion of Einstainean physics offered by Li & Torr, one could conclude that gravity, time, and light speed could be altered by the application of electromagnetic force to a superconductor.

By creating a space-time curvature anomaly associated with lowered pressure of inflationary vacuum state around a space vehicle, with the lowest vacuum pressure density located directly in front of the vehicle, a condition could be created where gravity associated with lowered vacuum pressure density pulls the vehicle forward in modified space-time.

By creating a space-time curvature anomaly associated with elevated pressure of inflationary vacuum state around the space vehicle, with the point of highest vacuum pressure density located directly behind the vehicle, a condition could be created where a repulsion force associated with elevated vacuum pressure density pushes the space vehicle forward in modified space-time. From the above-mentioned cosmological constant equation, re-written as:

$$\rho_\Lambda = \frac{3c^2}{8\pi G}\Lambda$$

it is clear that the increase in the vacuum pressure density could lead to a substantial increase in the light-speed. If the space vehicle is moving in the anomaly where the local light-speed is higher than the light-speed of the ambient vacuum, and if this vehicle approaches this local light-speed, the space vehicle would then possibly exceed the light-speed characteristic for the ambient area.

The levitating and rotating superconductor disk, which Podkletnov used to protect the object of experiment from the attraction produced by the energy of the vacuum, was externally energised by the externally-powered solenoid

coils. Thus, Podkletnov's system is stationary by definition and not suitable for travel in air or space. Even if the superconductive disk is made part of the craft, and if it is energised by the energy available on the craft, the resulting anomaly is one-sided, not enveloping, and not providing the variable speed of light (VSL) environment for the craft.

In a recent (2002) article, Chris Y. Tailor and Modanese propose to employ an impulse gravity generator directing, from an outside location, an anomalous beam toward a spacecraft, this beam acting as a repulsion force field producing propulsion for the spacecraft. ("Evaluation of an Impulse Gravity Generator Based Beamed Propulsion Concept", Chris Y. Taylor and Giovanni Modanese, American Institute of Aeronautics and Astronautics, Inc., 2002, 21 pages, 10 figures). The authors of the article, however, didn't take into account the powerful quantised processes of field dispersion, which would greatly limit the distance of propagation of the repulsive force. At best, the implementation of this concept could assist in acceleration and deceleration at short distances from the impulse gravity generator, and only along a straight line of travel. If the travel goal is a space exploration mission rather than the shuttle-like commute, the proposed system is of little use.

Only a self-sufficient craft, equipped with the internal gravity generator and the internal energy source powering this generator, would have the flexibility needed to explore new frontiers of space. The modification of the space-time curvature all around the spacecraft would allow the spacecraft to approach the light-speed characteristic for the modified locale, this light-speed, when observed from a location in the ambient space, being potentially many times higher than the ambient light-speed. Then, under sufficient local energies, that is, energies available on the spacecraft, very large intergalactic distances could be reduced to conventional planetary distances.

In "The First Men in the Moon" (1903), H. G. Wells anticipates gravitational propulsion methods when he describes gravity repelling "cavorite." Discovered by Professor Cavor, the material acts as a "gravity shield" allowing Cavor's vehicle to reach the Moon. Prof. Cavor built a large spherical gondola surrounded on all sides by cavorite shutters that could be closed or opened. When Prof. Cavor closed all the shutters facing the ground and opened the shutters facing the moon, the gondola took off for the Moon.

Until today, no cavorite has been discovered. However, recent research in the area of superconductivity, nano materials and quantum state of vacuum, including that of Li, Torr, Podkletnov, and Modanese, has resulted in important new information about the interaction between a gravitational field and special states of matter at a quantum level. This new research opens the possibility of using new electromagnetically-energised superconductive materials allowing stable states of energy, the materials useful not only in controlling the local gravitational fields, but also in creating new gravitomagnetic fields.

BACKGROUND OF INVENTION: OBJECTS AND ADVANTAGES

There are four objects of this invention:

The first object is to provide a method for generating a pressure anomaly of inflationary vacuum state that leads to electromagnetic propulsion.

The second object is to provide a space vehicle capable of electromagnetically-generated propulsion. The implementation of these two objects leads to the development of the space vehicle propelled by gravitational imbalance with gravity pulling, and/or antigravity pushing, the space vehicle forward.

The third object is to provide a method for generating a pressure anomaly of inflationary vacuum state, specifically, the local increase in the level of vacuum pressure density associated with the greater curvature of space-time. The speed of light in such an anomaly would be higher than the speed of light in the ambient space.

The fourth object is to provide the space vehicle capable of generating an unequally-distributed external anomaly all around this vehicle, specifically the anomaly with the elevated level of vacuum pressure density. The anomaly is formed in such a way that gravity pulls the space vehicle forward in the modified space-time at a speed possibly approaching the light-speed specific for this modified locale. If the vacuum pressure density of the locale is modified to be substantially higher than that of the ambient vacuum, the speed of the vehicle could conceivably be higher than the ambient light-speed.

SUMMARY OF THE INVENTION

This invention concerns devices self-propelled by the artificially changed properties of the pressure of inflationary vacuum state to speeds possibly approaching the light-speed specific for this modified locale. Furthermore, this invention concerns devices capable of generating the space-time anomaly characterised by the elevated vacuum

pressure density. The devices combining these capabilities may be able to move at speeds substantially higher than the light-speed in the ambient space.

The device of this invention is a space vehicle. The outside shell of the space vehicle is formed by a hollow disk, sphere, or the like hollowed 3-dimensional shape made of a superconductor material, hereinafter a hollow superconductive shield. An inner shield is disposed inside the hollow superconductive shield. The inner shield is provided to protect crew and life-support equipment inside.

A support structure, upper means for generating an electromagnetic field and lower means for generating an electromagnetic field are disposed between the hollow superconductive shield and the inner shield. A flux modulation controller is disposed inside the inner shield to be accessible to the crew.

Electrical energy is generated in a power source disposed inside the hollow superconductive shield. The electrical energy is converted into an electromagnetic field in the upper means for generating an electromagnetic field and the lower means for generating an electromagnetic field.

Electrical motors, also disposed inside the hollow superconductive shield, convert the electrical energy into mechanical energy.

The mechanical energy and the electromagnetic field rotate the hollow superconductive shield, and the upper and the lower means for generating an electromagnetic field, against each other.

The electromagnetic field is converted into a gravitomagnetic field in the hollow superconductive shield.

The gravitomagnetic field, propagated outward, orthogonally to the walls of the hollow superconductive shield, forms a pressure anomaly of inflationary vacuum state in the area of propagation. The pressure anomaly of inflationary vacuum state is comprised of an area of relatively lower vacuum pressure density in front of the space vehicle and an area of relatively higher vacuum pressure density behind the vehicle.

The difference in the vacuum pressure density propels the space vehicle of this invention forward.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig.1 is a cross-sectional view through the front plane taken along the central axis of a space vehicle provided by the method and device of this invention.

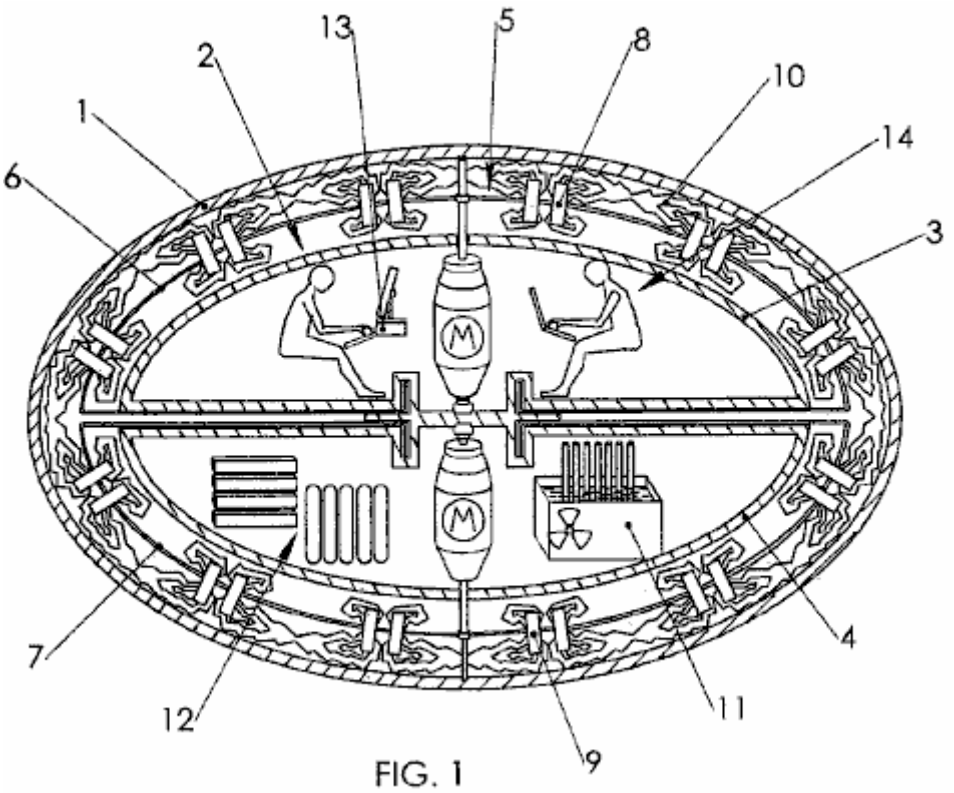


FIG. 1

Fig.2A and **Fig.2B** are diagrams, presented as perspective views, showing some of the physical processes resulting from a dynamic application of an electromagnetic field to a hollow superconductive shield. Only one line of quantised vortices, shown out of scale, is presented for illustration purposes.

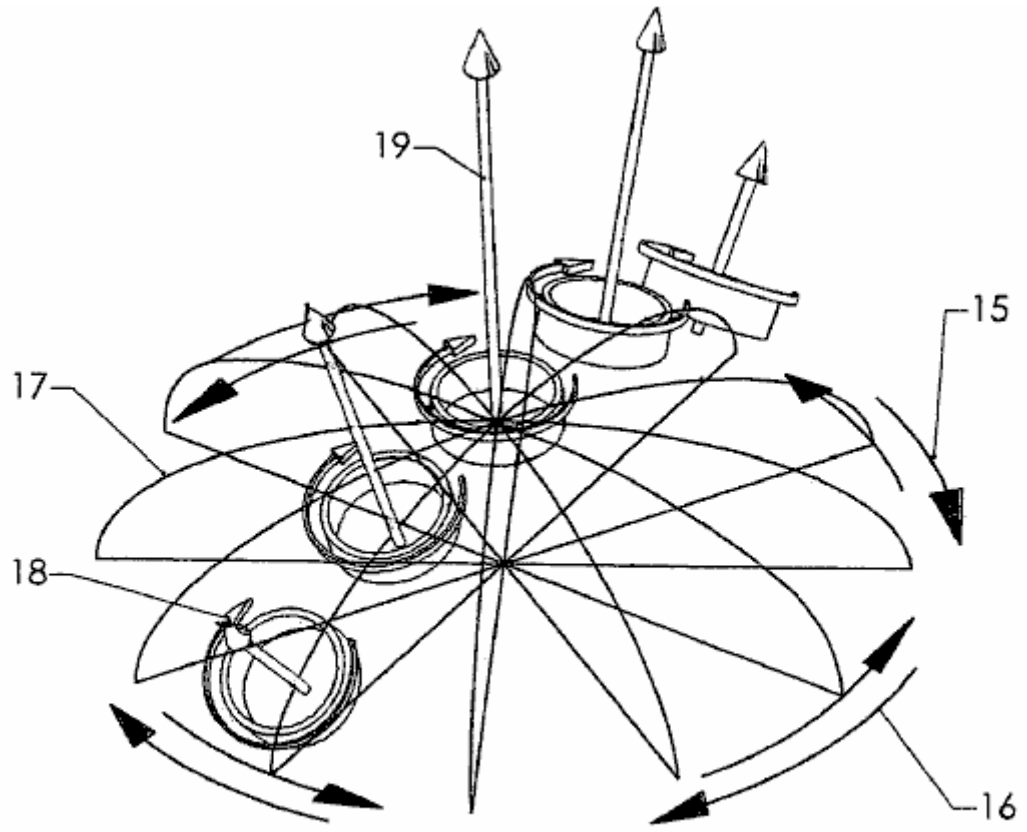


FIG. 2A

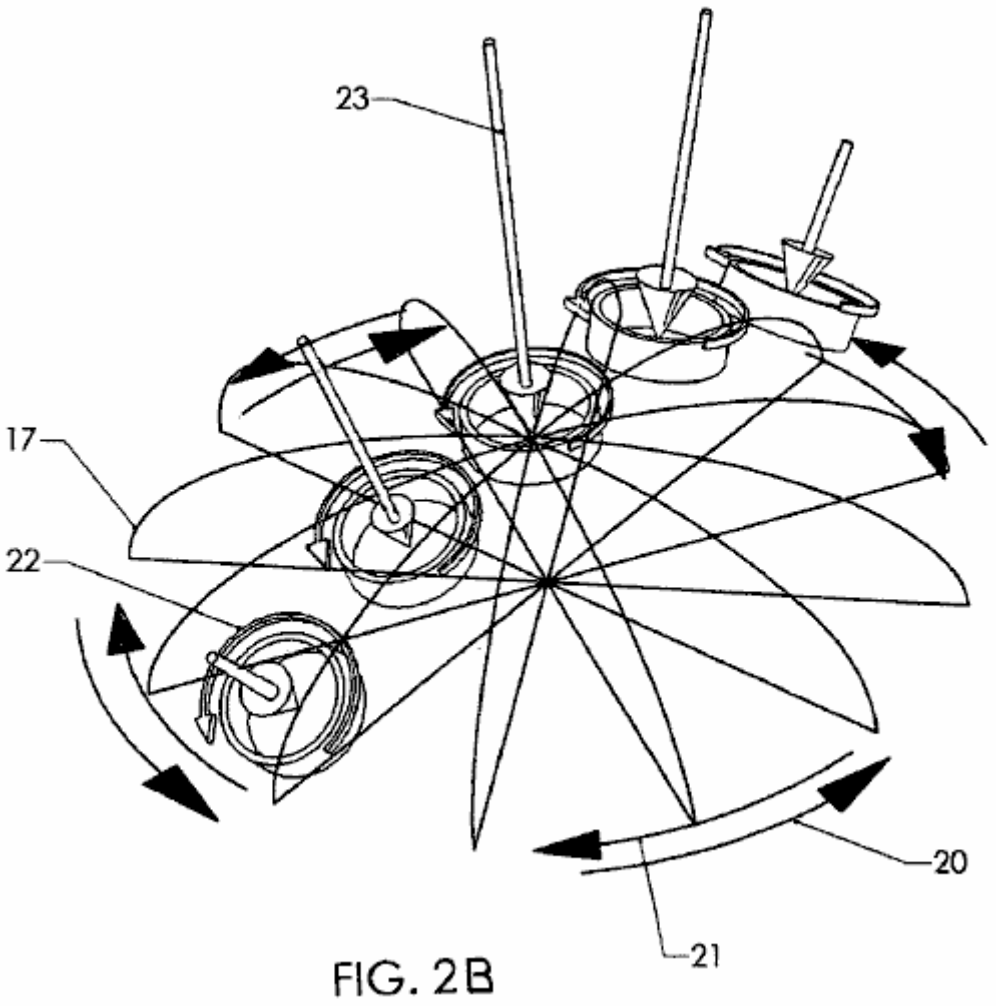


FIG. 2B

Fig.3A and **Fig.3B** are diagrams, presented as perspective views, showing a vacuum pressure density anomaly associated with lowered pressure of inflationary vacuum state and a vacuum pressure density anomaly associated with elevated pressure of inflationary vacuum state, respectively. Both anomalies are shown on the background of Universal curvature of inflationary vacuum state.

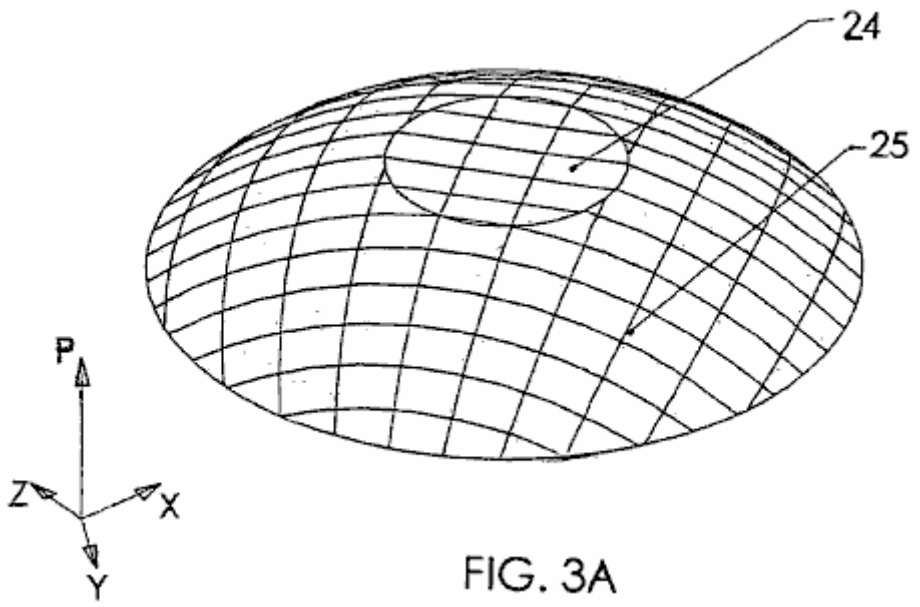


FIG. 3A

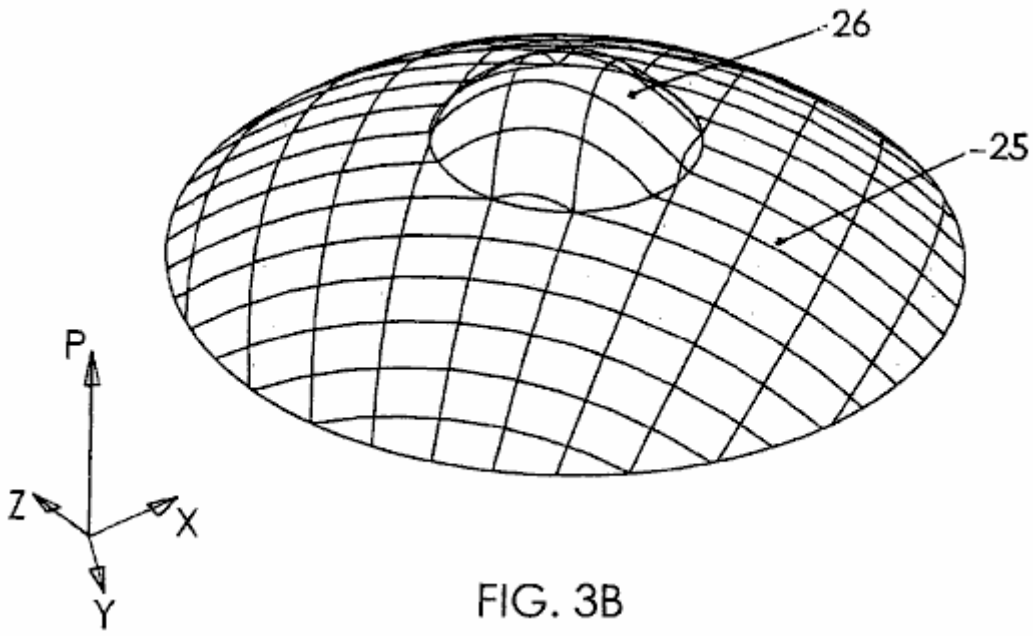


FIG. 3B

Fig.4A and **Fig.4B** are diagrams, presented as perspective views, showing a space-time anomaly associated with lowered pressure of inflationary vacuum state and a space-time anomaly associated with elevated pressure of inflationary vacuum state, respectively. Both anomalies are shown on the background of Universal space-time.

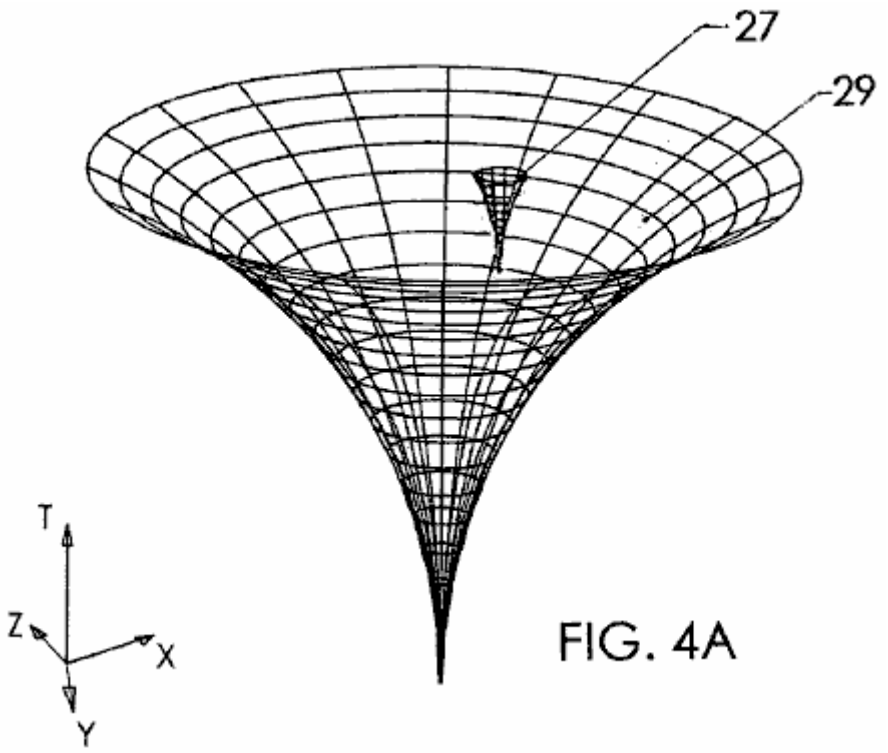


FIG. 4A

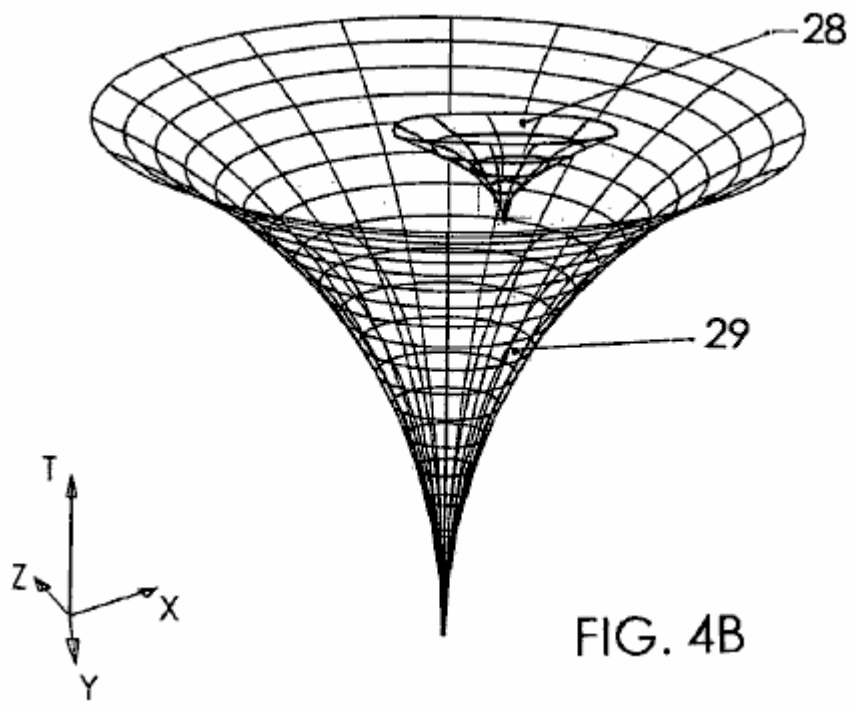
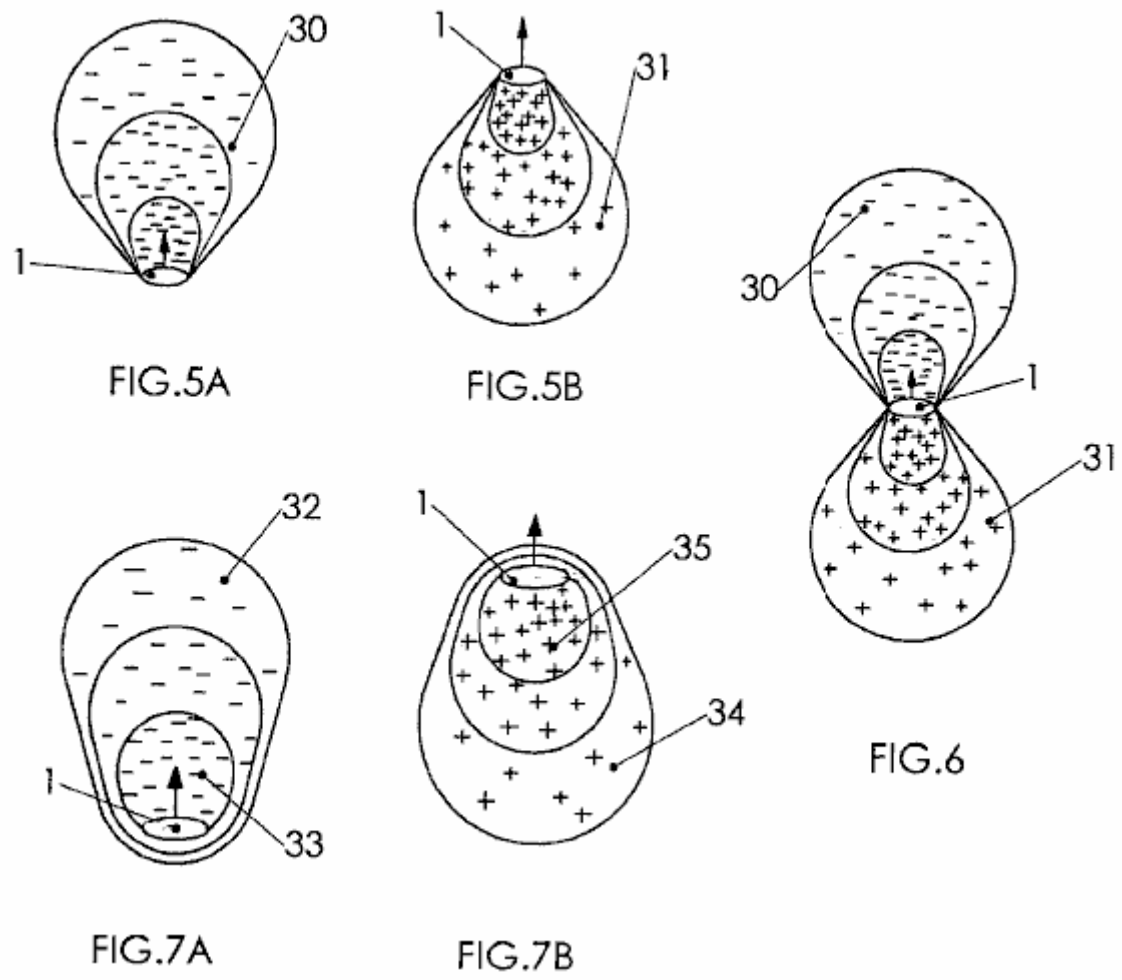


FIG. 4B

Figs.5A, 5B, 6, 7A, & 7B are diagrams of space-time curvature anomalies generated by the space vehicle of the current invention, these anomalies providing for the propulsion of the space vehicle.



DRAWINGS—REFERENCE NUMERALS

#1 hollow superconductive shield
#2 inner shield
#3 upper shell
#4 lower shell
#5 support structure
#6 upper rotating element
#7 lower rotating element
#8 upper means for generating an electromagnetic field
#9 lower means for generating an electromagnetic field
#10 flux lines
#11 power source
#12 life-support equipment
#13 flux modulation controller
#14 crew
#15 clockwise shield motion vector
#16 counter-clockwise EMF motion vector
#17 wire grid
#18 clockwise quantised vortices of lattice ions
#19 outward gravitomagnetic field vector
#20 counter-clockwise shield motion vector
#21 clockwise EMF motion vector
#22 counter-clockwise quantised vortices of lattice ions
#23 inward gravitomagnetic field vector
#24 vacuum pressure density anomaly associated with lowered pressure of inflationary vacuum state
#25 Universal curvature of inflationary vacuum state
#26 vacuum pressure density anomaly associated with elevated pressure of inflationary vacuum state
#27 space-time anomaly associated with lowered pressure of inflationary vacuum state
#28 space-time anomaly associated with elevated pressure of inflationary vacuum state
#29 Universal space-time
#30 substantially droplet-shaped space-time curvature anomaly associated with lowered pressure of inflationary vacuum state
#31 substantially droplet-shaped space-time anomaly associated with elevated pressure of inflationary vacuum state
#32 substantially egg-shaped space-time anomaly associated with lowered pressure of inflationary vacuum state
#33 area of the lowest vacuum pressure density
#34 substantially egg-shaped space-time anomaly associated with elevated pressure of inflationary vacuum state
#35 area of the highest vacuum pressure density

DESCRIPTION OF THE PRESENTLY PREFERRED EMBODIMENT

Fig.1 is a cross-sectional view through the front plane taken along the central axis of a space vehicle provided by the method and device of this invention. A hollow superconductive shield 1 forms a protective outer shell of the space vehicle. The hollow superconductive shield 1 may be shaped as a hollow disk, sphere, or the like 3-dimensional geometrical figure formed by the 2-dimensional rotation of a curve around the central axis.

In the preferred embodiment, the hollow superconductive shield 1 is made of a superconductor such as $YBa_2Cu_33O_{7-y}$, or a like high-temperature superconductor with a composite crystal structure cooled to the temperature of about 40^0K. Those skilled in the art may envision the use of many other low and high temperature superconductors, all within the scope of this invention.

An inner shield **2** is disposed inside the hollow superconductive shield **1**. The inner shield **2** is comprised of an upper shell **3** and a lower shell **4**, the shells **3** and **4** adjoined with each other. Executed from insulation materials such as foamed ceramics, the inner shield **2** protects the environment within the shield from the electromagnetic field and severe temperatures.

A support structure **5** is disposed between the hollow superconductive shield **1** and the inner shield **2**, concentric to the hollow superconductive shield. The support structure **5** is comprised of an upper rotating element **6** and a lower rotating element **7**.

The upper rotating element **6** is pivotably disposed inside the hollow superconductive shield **1** and may envelope the upper shell **3**. The lower rotating element **7** is pivotably disposed inside the hollow superconductive shield **1**

and may envelope the lower shell **4**. Even though the preferred embodiment has two rotating elements, those skilled in the art may envision only one rotating element, or three or more rotation elements, all within the scope of this invention.

Upper means for generating an electromagnetic field **8** are disposed between the hollow superconductive shield **1** and the upper shell **3**. The upper means for generating an electromagnetic field **8** are fixed to the upper rotating element **6** at an electromagnetic field-penetrable distance to the hollow superconductive shield **1**.

Lower means for generating an electromagnetic field **9** are disposed between the hollow superconductive shield **1** and the lower shell **4**. The lower means for generating an electromagnetic field **9** are fixed to the lower rotating element **7** at an electromagnetic field-penetrable distance to the hollow superconductive shield **1**.

The upper means for generating an electromagnetic field **8** and the lower means for generating an electromagnetic field **9** could be solenoid coils or electromagnets. In the process of operation of the space vehicle, the electromagnetic field identified by flux lines **10**, is controllably and variably applied to the hollow superconductive shield **1**.

Electric motors are disposed inside the hollow superconductive shield along its central axis.

A power source **11** is disposed inside the hollow superconductive shield **1** and may be disposed inside the lower shell **4**. The power source **11** is electrically connected with the upper means for generating an electromagnetic field **8**, the lower means for generating an electromagnetic field **9**, and the electric motors. The upper means for generating an electromagnetic field **8**, the lower means for generating an electromagnetic field **9**, and the electric motors provide for the rotation of the upper rotating element **6** and the lower rotating element **7**. The power source **11** may be a nuclear power generator.

Life-support equipment **12** is disposed inside the inner shield **2**, and may be disposed inside the lower shell **4**. The life-support equipment **12** may include oxygen, water, and food.

A flux modulation controller **13** is disposed inside the inner shield **2**, and may be disposed inside the upper shell **3**. The flux modulation controller **13** is in communication with the upper means for generating an electromagnetic field **8**, the lower means for generating an electromagnetic field **9**, the power source **11**, and the electric motors.

The flux modulation controller **8** may be executed as a computer or a microprocessor. The flux modulation controller **8** is provided with a capability of modulating the performance parameters of the upper means for generating an electromagnetic field **8**, the lower means for generating an electromagnetic field **9**, the power source **11**, and the electric motors.

A crew **14** may be located inside the upper shell **3** of the inner shield **2** and may consist of one or more astronauts. The crew has a free access to the life-support equipment **12** and the flux modulation controller **8**. A person skilled in the art, may envision a fully-automated, pilotless craft, which is also within the scope of this invention.

A person skilled in the art, may also envision the embodiment (not shown), also within the scope of this invention, where the hollow superconductive shield is pivotable, and the support structure with the means for generating an electromagnetic field is affixed on the outside of the inner shield.

Fig.2A and **Fig.2B** are diagrams showing the results of the quantised electromagnetic turbulence within the superconductive shell of the hollow superconductive shield provided by the relative rotational motion of the hollow superconductive shield against the upper means for generating an electromagnetic field.

Fig.2A shows the clockwise relative rotational motion of the hollow superconductive shield, this motion identified by a clockwise shield motion vector **15**, and the counter-clockwise relative rotational motion of upper means for generating an electromagnetic field, this motion identified by a counter-clockwise EMF motion vector **16**.

The electromagnetic field, controllably and variably applied by the upper means for generating an electromagnetic field, whose various positions are identified by a wire grid **17**, to the hollow superconductive shield (not shown), causes quantised electromagnetic turbulence within the hollow superconductive shield. This turbulence is represented by a plurality of clockwise quantised vortices of lattice ions **18**. Only one line of the clockwise quantised vortices of lattice ions **18**, (not to scale), is shown for illustration purposes only. Each of the clockwise quantised vortices of lattice ions **18** generates a gravitomagnetic field identified by an outward gravitomagnetic field vector **19** directed orthogonally away from the hollow superconductive shield.

Fig.2B shows the counter-clockwise relative rotational motion of the hollow superconductive shield, this motion identified by a counter-clockwise shield motion vector **20**, and the clockwise relative rotational motion of upper means for generating an electromagnetic field, this motion identified by a clockwise EMF motion vector **21**.

The electromagnetic field, controllably and variably applied by the upper means for generating an electromagnetic field identified by the wire grid **17**, to the hollow superconductive shield (not shown), causes quantised electromagnetic turbulence within the hollow superconductive shield, this turbulence represented by a plurality of counter-clockwise quantised vortices of lattice ions **22**. Only one line of the counter-clockwise quantised vortices of lattice ions **22**, (not to scale), is shown for illustration purposes only. Each of the counter-clockwise quantised vortices of lattice ions **22** generates a gravitomagnetic field identified by an inward gravitomagnetic field vector **23** directed orthogonally toward the hollow superconductive shield.

The electrical requirements for providing the Li-Torr effect are as follows:

Podkletnov has reported using the high frequency current of 105 Hz. He also used 6 solenoid coils @ 850 Gauss each. The reported system's efficiency reached 100% and the total field in the Podkletnov's disk was about 0.5 Tesla. The maximum weight loss reported by Podkletnov was 2.1%.

The preferred embodiment of the device of current invention is capable of housing 2-3 astronauts and therefore is envisioned to be about 5 meters in diameter at the widest point. The preferred space vehicle's acceleration is set at 9.8 m/s/s providing that gravity on board is similar to that on the surface of Earth.

The means for generating an electromagnetic field may be comprised of 124 solenoid coils. At the same 100% efficiency reported by Podkletnov, the total field required providing the acceleration of 9.8 m/s/s is 5,000 Tesla, or about 40 Tesla per coil. Skeggs suggests that on the Podkletnov device, out of 850 Gauss developed on the coil surface, the field affecting the superconductor and causing the gravitomagnetism is only 400 Gauss ("Engineering Analysis of the Podkletnov Gravity Shielding Experiment, Peter L. Skeggs, Quantum Forum, Nov. 7, 1997, http://www.inetarena.com/~noetic/pls/podlev.html, 7 pages). This translates into 47% device efficiency.

In this 47%-efficient space vehicle, the total field required achieving the 9.8 m/s/s acceleration is about 10,600 Tesla, or 85.5 Tesla per each of 124 solenoid coils. It must be noted that at this acceleration rate, it would take nearly a year for the space vehicle to reach the speed of light.

It also must be noted that Skeggs has detected a discrepancy between the Li-Torr estimates and Podkletnov's practical results. If Podkletnov's experimental results are erroneous while the Li-Torr estimates are indeed applicable to the space vehicle of this invention, then the energy requirements for achieving the sought speed would be substantially higher than the above estimate of 10,600 Tesla.

Podkletnov has concluded that, in order for the vacuum pressure density anomaly to take place, the Earth-bound device must be in the condition of Meissner levitation. As are all space bodies, the space vehicle is a subject to the pressure inflationary vacuum state and the gravitational force, which, within the migrating locality of the expanding Universe, in any single linear direction, are substantially in equilibrium. Thus, for the space vehicle, the requirement of Meissner levitation is waved.

The propagation of the gravitomagnetic field identified by the outward gravitomagnetic field vector **19** and the inward gravitomagnetic field vector **23** would cause exotic quantised processes in the vacuum's subatomic particles that include particle polarisation, ZPF field defects, and the matter-energy transformation per $E=mc^2$. The combination of these processes would result in the gravitational anomaly. According to the general relativity theory, gravitational attraction is explained as the result of the curvature of space-time being proportional to the gravitational constant. Thus, the change in the gravitational attraction of the vacuum's subatomic particles would cause a local anomaly in the curvature of the Einsteinean space-time.

Gravity is the same thing as bent space, propagating with the speed of light characteristic for the particular space-time curvature. When bent space is affected, there is a change in the speed of propagation of gravity within the space-time curvature anomaly. The local speed of light, according to Fomalont and Kopeikin always equal to the local speed of propagation of gravity, is also affected within the locality of space-time curvature anomaly.

Creation of space-time curvature anomalies adjacent to, or around, the space vehicle, these anomalies characterised by the local gravity and light-speed change, has been the main object of this invention.

Fig.3A shows a diagram of a vacuum pressure density anomaly associated with lowered pressure of inflationary vacuum state **24** on the background of Universal curvature of inflationary vacuum state **25**. The vacuum pressure density anomaly associated with lowered pressure of inflationary vacuum state **24** is formed by a multitude of the inward gravitomagnetic field vectors. According to the cosmological constant equation,

ρ_Λ: $\Lambda=(8\pi G/3c^2)\rho_\Lambda$

where:
The cosmological constant Lambda, is proportional to the vacuum energy pressure rho-lambda, G is Newton's constant of gravitation, and c is the speed of light, so the curvature of space-time is proportional to the gravitational constant. According to the general relativity theory, the change in the vacuum pressure density is proportional to the change in the space-time curvature anomaly. By replacing rho-lambda with the vacuum pressure density, P times the vacuum energy coefficient kappa, and replacing c with:
delta-distance/delta-time, we derive to the equation:

$$\Lambda=[8\pi G/3(\Delta\text{distance}/\Delta\text{time})^2]P_\kappa$$

and can now construct a vacuum pressure density curvature diagram.

The vacuum pressure density curvature anomaly associated with lowered pressure of inflationary vacuum state **24** is shown here as a flattened surface representing the lowered pressure of the inflationary vacuum state. This anomaly is the result of the exotic quantised processes in the subatomic particles caused by the quantised turbulence occurring in the hollow superconductive shield. The XYZ axes represent three dimensions of space and the P axis represents the vacuum pressure density.

Fig.3B shows a diagram of a vacuum pressure density anomaly associated with elevated pressure of inflationary vacuum state **26** on the background of the Universal curvature of inflationary vacuum state **25**. The vacuum pressure density anomaly associated with elevated pressure of inflationary vacuum state **26** is formed by a multitude of the outward gravitomagnetic field vectors. The anomaly is shown here as a convex surface representing the elevated pressure of inflationary vacuum state. The diagrams of **Fig.3A** and **Fig.3B** are not to scale with the anomaly sizes being exaggerated for clarity.

Fig.4A and **Fig.4B** show diagrams of a space-time anomaly associated with lowered pressure of inflationary vacuum state **27**, and a space-time anomaly associated with elevated pressure of inflationary vacuum state **28**, respectively, each on the background a diagram of Universal space-time **29**.

The quaterised Julia set $Q_{n+1} = Q_n^2 + C_0$ is assumed to be an accurate mathematical representation of the Universal space-time. The generic quaternion Q_0 belongs to the Julia set associated with the quaternion C, and n tends to infinity. If we assume that the quaternion value C_0 is associated with the Universal space-time **29**, C_1 is the value of quaternion C for the space-time anomaly associated with lowered pressure of inflationary vacuum state **27**, and C_2 is the value of quaternion C for the space-time anomaly associated with elevated pressure of inflationary vacuum state **28**, then we can construct two diagrams.

The diagram of **Fig.4A** shows the space-time anomaly associated with lowered pressure of inflationary vacuum state **27** as a quaterised Julia set contained in a 4-dimensional space: $Q_{n+1} = Q_n^2 + C_1$ on the background of the Universal space-time **29** represented by $Q_{n+1} = Q_n^2 + C_0$.

The diagram of **Fig.4B** shows the space-time anomaly associated with elevated pressure of inflationary vacuum state **28** as a quaterised Julia set $Q_{n+1} = Q_n^2 + C_2$, also on the background of the Universal space-time **29** represented by $Q_{n+1} = Q_n^2 + C_0$. On both diagrams, the XYZ axes represent three dimensions of space, and the T axis represents time. The diagrams are not to scale: the anomaly sizes are exaggerated for clarity, and the halves of quaterised Julia sets, conventionally associated with the hypothetical Anti-Universe, are omitted.

Figs. 5A, 5B, 6, 7A, & 7B show simplified diagrams of space-time curvature anomalies generated by the space vehicle of the current invention, these anomalies providing for the propulsion of the space vehicle. In each case, the pressure anomaly of inflationary vacuum state is comprised of an area of relatively lower vacuum pressure density in front of the space vehicle and an area of relatively higher vacuum pressure density behind the space vehicle. Because the lower pressure of inflationary vacuum state is associated with greater gravity and the higher pressure is associated with the higher repulsive force, the space vehicle is urged to move from the area of relatively higher vacuum pressure density toward the area of relatively lower vacuum pressure density.

Fig.5A illustrates the first example of space-time curvature modification. This example shows a substantially droplet-shaped space-time curvature anomaly associated with lowered pressure of inflationary vacuum state **30** adjacent to the hollow superconductive shield **1** of the space vehicle. The anomaly **30** is provided by the propagation of a gravitomagnetic field radiating orthogonally away from the front of the hollow superconductive shield **1**. This gravitomagnetic field may be provided by the relative clockwise motion of the upper means for generating an electromagnetic field, and relative counterclockwise motion of the hollow superconductive field, as observed from above the space vehicle.

In this example, the difference between the space-time curvature within the substantially droplet-shaped space-time anomaly associated with lowered pressure of inflationary vacuum state, and the ambient space-time curvature, the space-time curvature being the same as gravity, results in the gravitational imbalance, with gravity pulling the space vehicle forward.

Fig.5B illustrates the second example of space-time curvature modification. This example shows a substantially droplet-shaped space-time anomaly associated with elevated pressure of inflationary vacuum state **31** adjacent to the hollow superconductive shield **1** of the space vehicle. The anomaly **31** is provided by the propagation of a gravitomagnetic field radiating orthogonally away from the back of the hollow superconductive shield. This gravitomagnetic field may be provided by the relative counter-clockwise motion of the lower means for generating an electromagnetic field, and relative clockwise motion of the hollow superconductive field, as observed from below the space vehicle.

In this example, the difference between the space-time curvature within the substantially droplet-shaped space-time anomaly associated with elevated pressure of inflationary vacuum state, and the ambient space-time curvature, the space-time curvature being the same as gravity, results in the gravitational imbalance, with the repulsion force pushing the space vehicle forward.

Fig.6 illustrates the third example of space-time curvature modification. This example shows the formation of the substantially droplet-shaped space-time anomaly associated with lowered pressure of inflationary vacuum state **30** combined with the substantially droplet-shaped space-time anomaly associated with elevated pressure of inflationary vacuum state **31**. This combination of anomalies may be provided by the relative clockwise motion of the upper means for generating an electromagnetic field and relative clockwise motion of the hollow superconductive field, combined with the relative clockwise motion of the lower means for generating an electromagnetic field, as observed from above the space vehicle.

In this example, the difference between the space-time curvature within the substantially droplet-shaped space-time anomaly associated with lowered pressure of inflationary vacuum state, and the space-time curvature of the substantially droplet-shaped space-time anomaly associated with elevated pressure of inflationary vacuum state, the space-time curvature being the same as gravity, results in the gravitational imbalance, with gravity pulling, and the repulsion force pushing, the space vehicle forward.

Fig.7A illustrates the fourth example of space-time curvature modification. This example shows the formation of a substantially egg-shaped space-time anomaly associated with lowered pressure of inflationary vacuum state **32** around the hollow superconductive shield **1** of the space vehicle. The anomaly **32** is provided by the propagation of gravitomagnetic field of unequally-distributed density, this gravitomagnetic field radiating in all directions orthogonally away from the hollow superconductive shield. The propagation of the unequally-distributed gravitomagnetic field leads to the similarly unequally-distributed space-time curvature anomaly. This unequally-distributed gravitomagnetic field may be provided by the relatively faster clockwise motion of the upper means for generating an electromagnetic field relative to the hollow superconductive field, combined with the relatively slower counter-clockwise motion of the lower means for generating an electromagnetic field, as observed from above the space vehicle.

An area of the lowest vacuum pressure density **33** of the substantially egg-shaped space-time anomaly associated with lowered pressure of inflationary vacuum state **32** is located directly in front of the space vehicle.

In this example, the variation in the space-time curvature within the substantially egg-shaped space-time anomaly associated with lowered pressure of inflationary vacuum state, the space-time curvature being the same as gravity, results in a gravitational imbalance, with gravity pulling the space vehicle forward in modified space-time.

Fig.7B illustrates the fifth example of space-time curvature modification, also with the purpose of providing for a propulsion in modified space-time. This example shows the formation of a substantially egg-shaped space-time anomaly associated with elevated pressure of inflationary vacuum state **34** around the hollow superconductive shield 1 of the space vehicle. The anomaly **34** is provided by the propagation of gravitomagnetic field of unequally-distributed density, this gravitomagnetic field radiating in all directions orthogonally away from the hollow superconductive shield. The propagation of the unequally-distributed gravitomagnetic field leads to the similarly unequally-distributed space-time curvature anomaly. This unequally-distributed gravitomagnetic field may be provided by the relatively slower counter-clockwise motion of the upper means for generating an electromagnetic field relative to the hollow superconductive field, combined with the relatively faster clockwise motion of the lower means for generating an electromagnetic field, as observed from above the space vehicle.

An area of the highest vacuum pressure density **35** of the substantially egg-shaped space-time anomaly associated with elevated pressure of inflationary vacuum state **34** is located directly behind the space vehicle.

In this example, the variation in the space-time curvature within the substantially egg-shaped space-time anomaly associated with elevated pressure of inflationary vacuum state, the space-time curvature being same as gravity, results in a gravitational imbalance, with the repulsion force pushing the space vehicle forward in modified space-time at speeds approaching the light-speed characteristic for this modified area. This light-speed might be much higher than the light-speed in the ambient space.

By creating alternative anomalies and modulating their parameters, the space vehicle's crew would dilate and contract time and space on demand. The space vehicle, emitting a vacuum pressure modifying, controllably-modulated gravitomagnetic field in all directions, would rapidly move in the uneven space-time anomaly it created, pulled forward by gravity or pushed by the repulsion force. The time rate zone of the anomaly is expected to have multiple quantised boundaries rather than a single sudden boundary affecting space and time in the immediate proximity of the vehicle. Speed, rate of time, and direction in space could be shifted on demand and in a rapid manner. The modulated light-speed could make the space vehicle suitable for interstellar travel. Because of the time rate control in the newly created isospace, the accelerations would be gradual and the angles of deviation would be relatively smooth. The gravity shielding would further protect pilots from the ill-effects of gravity during rapid accelerations, directional changes, and sudden stops.

****

If you find the thought of generating a gravitational field, difficult to come to terms with, then consider the work of Henry Wallace who was an engineer at General Electric about 25 years ago, and who developed some incredible inventions relating to the underlying physics of the gravitational field. Few people have heard of him or his work. Wallace discovered that a force field, similar or related to the gravitational field, results from the interaction of relatively moving masses. He built machines which demonstrated that this field could be generated by spinning masses of elemental material having an odd number of nucleons -- i.e. a nucleus having a multiple half-integral value of h-bar, the quantum of angular momentum. Wallace used bismuth or copper material for his rotating bodies and "kinnemassic" field concentrators.

Aside from the immense benefits to humanity which could result from a better understanding of the physical nature of gravity, and other fundamental forces, Wallace's inventions could have enormous practical value in countering gravity or converting gravitational force fields into energy for doing useful work. So, why has no one heard of him? One might think that the discoverer of important knowledge such as this would be heralded as a great scientist and nominated for dynamite prizes. Could it be that his invention does not work? Anyone can get the patents. Study them -- Wallace -- General Electric -- detailed descriptions of operations -- measurements of effects -- drawings and models -- it is authentic. If you are handy with tools, then you can even build it yourself. It does work.

Henry was granted two patents in this field:
US Patent #3626605 -- "Method and Apparatus for Generating a Secondary Gravitational Force Field", Dec 14, 1971 and

US Patent #3626606 -- "Method and Apparatus for Generating a Dynamic Force Field", Dec 14, 1971. He was also granted US Patent #3823570 -- "Heat Pump" (based on technology similar to the above two inventions), July 16, 1973.

These patents can be accessed via http://www.freepatentsonline.com

The First High MPG Carburettor of Charles Pogue

US Patent 642,434 12th November 1932 Inventor: Charles N. Pogue

CARBURETTOR

This patent describes a carburettor design which was able to produce very high mpg figures using the gasoline available in the USA in the 1930s but which is no longer available as the oil industry does not want functional high mpg carburettors to be available to the public.

DESCRIPTION

This invention relates to a device for obtaining an intimate contact between a liquid in a vaporous state and a gas, and particularly to such a device which may serve as a carburettor for internal combustion engines.

Carburettors commonly used for supplying a combustible mixture of air and liquid fuel to internal combustion engines, comprise a bowl in which a supply of the fuel is maintained in the liquid phase and a fuel jet which extends from the liquid fuel into a passage through which air is drawn by the suction of the engine cylinders. On the suction, or intake stroke of the cylinders, air is drawn over and around the fuel jet and a charge of liquid fuel is drawn in, broken up and partially vaporised during its passage to the engine cylinders. However, I have found that in such carburettors, a relatively large amount of the atomised liquid fuel is not vaporised and enters the engine cylinder in the form of microscopic droplets. When such a charge is ignited in the engine cylinder, only that portion of the liquid fuel which has been converted into the vaporous (molecular) state, combines with the air to give an explosive mixture. The remaining portion of the liquid fuel which is drawn into the engine cylinders and remains in the form of small droplets, does not explode and impart power to the engine, but burns with a flame and raises the temperature of the engine above that at which the engine operates most efficiently, i.e. 160^{O} to 180^{O} F.

According to this invention, a carburettor for internal combustion engines is provided in which substantially all of the liquid fuel entering the engine cylinder will be in the vapour phase and consequently, capable of combining with the air to form a mixture which will explode and impart a maximum amount of power to the engine, and which will not burn and unduly raise the temperature of the engine.

A mixture of air and liquid fuel in truly vapour phase in the engine cylinder is obtained by vaporising all, or a large portion of the liquid fuel before it is introduced into the intake manifold of the engine. This is preferably done in a vaporising chamber, and the “dry” vaporous fuel is drawn from the top of this chamber into the intake manifold on the intake or suction stroke of the engine. The term “dry” used here refers to the fuel in the vaporous phase which is at least substantially free from droplets of the fuel in the liquid phase, which on ignition would burn rather than explode.

More particularly, the invention comprises a carburettor embodying a vaporising chamber in the bottom of which, a constant body of liquid fuel is maintained, and in the top of which there is always maintained a supply of “dry” vaporised fuel, ready for admission into the intake manifold of the engine. The supply of vaporised liquid fuel is maintained by drawing air through the supply of liquid fuel in the bottom of the vaporising chamber, and by constantly atomising a portion of the liquid fuel so that it may more readily pass into the vapour phase. This is preferably accomplished by a double-acting suction pump operated from the intake manifold, which forces a mixture of the liquid fuel and air against a plate located within the chamber. To obtain a more complete vaporisation of the liquid fuel, the vaporising chamber and the incoming air are preferably heated by the exhaust gasses from the engine. The carburettor also includes means for initially supplying a mixture of air and vaporised fuel so that starting the engine will not be dependent on the existence of a supply of fuel vapours in the vaporising chamber.

The invention will be further described in connection with the accompanying drawings, but this further disclosure and description is to be taken as an exemplification of the invention and the same is not limited thereby except as is pointed out in the claims.

Fig.1 is an elevational view of a carburettor embodying my invention.

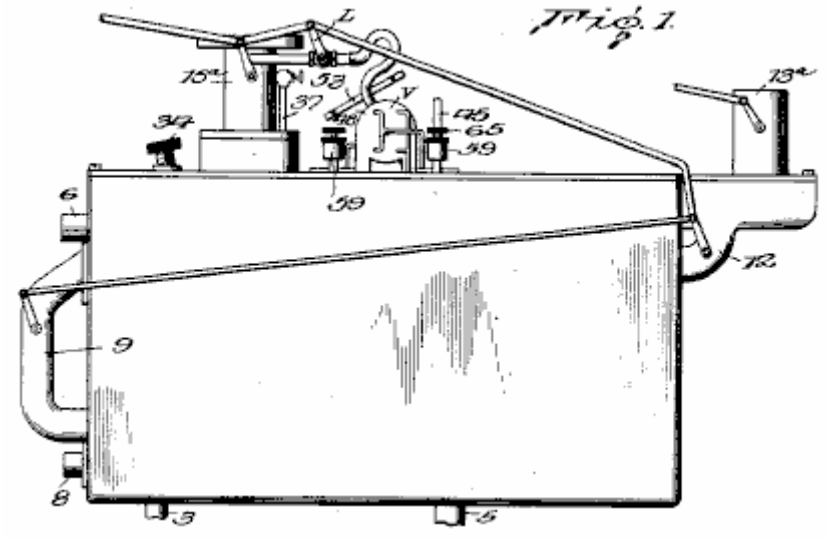


Fig.2 is a vertical cross-sectional view through the centre of Fig.1

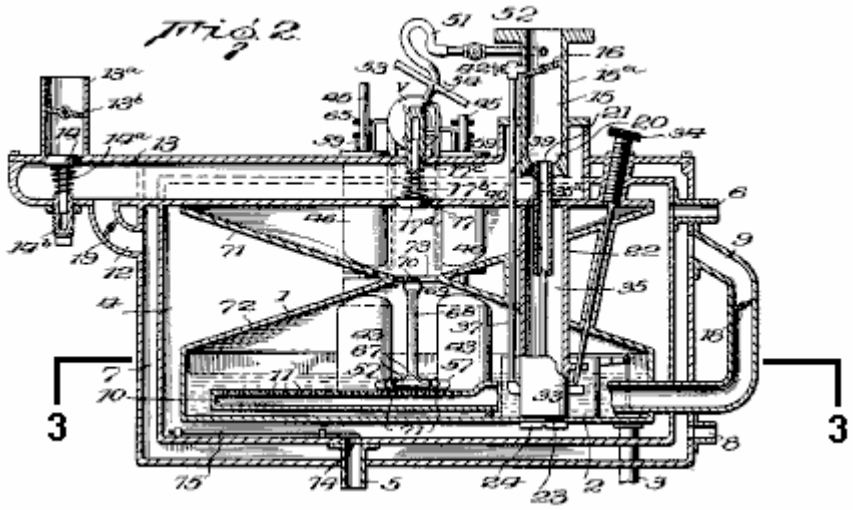
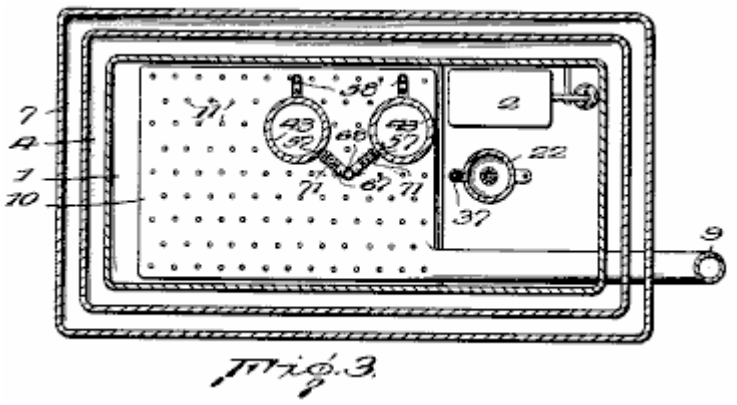


Fig.3 is a horizontal sectional view on line 3--3 of Fig.2.



Фиг. 3.

Fig.4 is an enlarged vertical sectional view through one of the pump cylinders and adjacent parts of the carburettor.

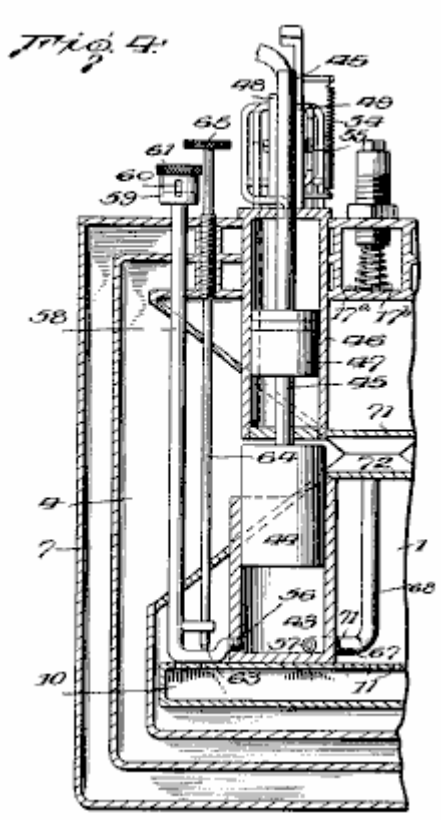


Fig.5 is an enlarged view through the complete double-acting pump and showing the associated distributing valve.

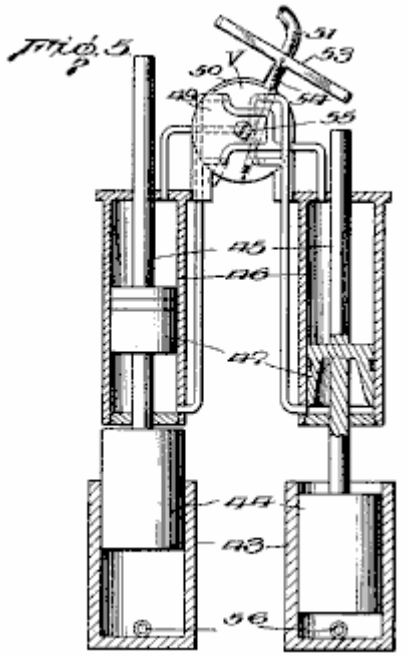


Fig.6 is an enlarged vertical sectional view through the atomising nozzle for supplying a starting charge for the engine.

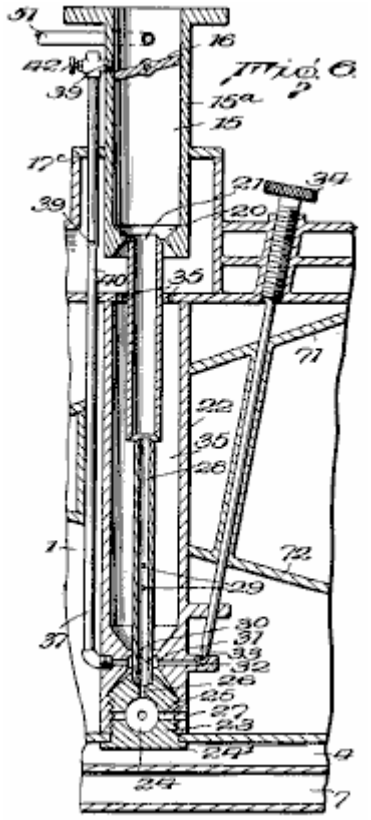


Fig.7 and Fig.8 are detail sectional views of parts 16 and 22 of Fig.6

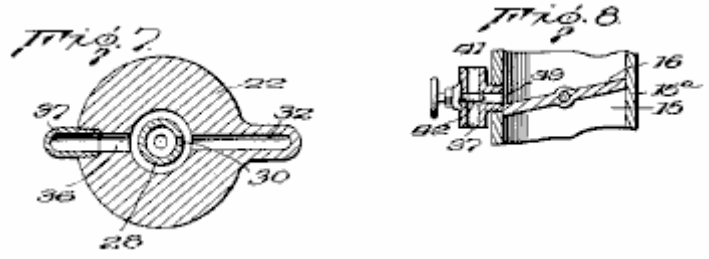
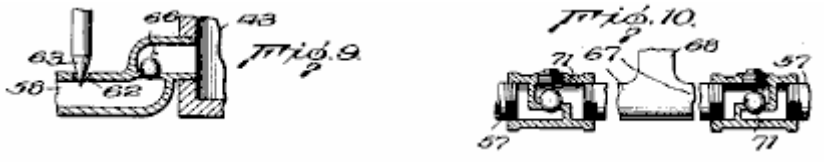


Fig.9 and Fig.10 are detail sectional views showing the inlet and outlet to the cylinders of the atomising pump.



Referring to the drawings, the numeral **1** indicates a combined vaporising chamber and fuel bowl in which liquid fuel is maintained at the level indicated in **Fig.1** by a float-valve **2** controlling the flow of liquid fuel through pipe **3** which leads from the vacuum tank or other liquid fuel reservoir.

The vaporising chamber **1** is surrounded by a chamber **4** through which hot exhaust gasses from the engine, enter through pipe **5** located at the bottom of the chamber. These gasses pass around the vaporising chamber **1** and heat the chamber, which accelerates the vaporisation of the liquid fuel. The gasses then pass out through the upper outlet pipe **6**.

Chamber **4** for the hot exhaust gasses, is in turn surrounded by chamber **7** into which air for vaporising part of the liquid fuel in chamber **1** enters through a lower intake pipe **8**. This air passes upwards through chamber **4** through which the hot exhaust gasses pass, and so the air becomes heated. A portion of the heated air then passes though pipe **9** into an aerator **10**, located in the bottom of the vaporising chamber **1** and submerged in the liquid fuel in it. The aerator **10** is comprised of a relatively flat chamber which extends over a substantial portion of the bottom of the chamber and has a large number of small orifices **11** in its upper wall. The heated air entering the aerator passes through the orifices **11** as small bubbles which then pass upwards through the liquid fuel. These bubbles, together with the heat imparted to the vaporising chamber by the hot exhaust gasses, cause a vaporisation of a portion of the liquid fuel.

Another portion of the air from chamber **7** passes through a connection **12** into passage **13**, through which air is drawn directly from the atmosphere into the intake manifold. Passage **13** is provided with a valve **14** which is normally held closed by spring **14a**, the tension of which may be adjusted by means of the threaded plug **14b**. Passage **13** has an upward extension **13a**, in which is located a choke valve **13b** for assisting in starting the engine. Passage **13** passes through the vaporising chamber **1** and has its inner end communicating with passage **15** via connector **15a** which is secured to the intake manifold of the engine. Passage **15** is provided with the usual butterfly valve **16** which controls the amount of fuel admitted to the engine cylinders, and consequently, regulates the speed of the engine.

The portion of passage **13** which passes through the vaporising chamber has an opening **17** normally closed by valve **17a** which is held against its seat by spring **17b**, the tension of which may be adjusted by a threaded plug **17c**. As air is drawn past valve **14** and through passage **13** on the intake or suction stroke of the engine, valve **17a** will be lifted from its seat and a portion of the dry fuel vapour from the upper portion of the vaporising chamber will be sucked into passage **13** through opening **17** and mingle with the air in it before entering passage **15**.

In order to regulate the amount of air passing from chamber **7** to aerator **10** and into passage **13**, pipe **9** and connection **12** are provided with suitable valves **18** and **19** respectively. Valve **18** in pipe **9** is synchronised with butterfly valve **16** in passage **15**. Valve **19** is adjustable and preferably synchronised with butterfly valve **16** as shown, but this is not essential.

The bottom of passage **15** is made in the form of a venturi **20** and a nozzle **21** for atomised liquid fuel and air is located at or adjacent to the point of greatest restriction. Nozzle **21** is preferably supplied with fuel from the supply of liquid fuel in the bottom of the vaporising chamber, and to that end, a member 22 is secured within the vaporising chamber by a removable threaded plug **23** having a flanged lower end **24**. Plug **22** extends through an opening in the bottom of chamber **1**, and is threaded into the bottom of member **22**. This causes the bottom wall of chamber **1** to be securely clamped between the lower end of member **22** and flange **24**, thus securely retaining member **22** in place.

Plug **23** is provided with a sediment bowl **24** and extending from bowl **24** are several small passages **25** extending laterally, and a central vertical passage **26**. The lateral passages **25** register with corresponding passages **27** located in the lower end of member **22** at a level lower than that at which fuel stands in chamber **1**, whereby liquid fuel is free to pass into bowl **24**.

Vertical passage **26** communicates with a vertical nozzle **28** which terminates within the flaring lower end of nozzle **21**. The external diameter of nozzle **26** is less than the interior diameter of the nozzle **21** so that a space is provided between them for the passage of air or and vapour mixtures. Nozzle **26** is also provided with a series of

inlets **29**, for air or air and vapour mixtures, and a fuel inlet **30**. Fuel inlet **30** communicates with a chamber **31** located in the member **22** and surrounding the nozzle **28.** Chamber **30** is supplied with liquid fuel by means of a passage **32** which is controlled by a needle valve **33**, the stem of which, extends to the outside of the carburettor and is provided with a knurled nut **34** for adjusting purposes.

The upper end of member **22** is made hollow to provide a space **35** surrounding the nozzles **21** and **28**. The lower wall of the passage **13** is provided with a series of openings **35a**, to allow vapours to enter space **35** through them. The vapours may then pass through inlets **29** into the nozzle **28**, and around the upper end of the nozzle **28** into the lower end of nozzle **21**.

Extending from chamber **31** at the side opposite passage **32**, is a passage **36** which communicates with a conduit **37** which extends upwards through passage **13**, and connects through a lateral extension **39**, with passage **15** just above the butterfly valve **16**. The portion of conduit **37** which extends through passage **13** is provided with an orifice **39** through which air or air and fuel vapour may be drawn into the conduit **37** mingle with and atomise the liquid fuel being drawn through the conduit. To further assist in this atomisation of the liquid fuel passing through conduit **37**, the conduit is restricted at **40** just below orifice **39**.

The upper end of conduit **37** is in communication with the atmosphere through opening **41** through which air may be drawn directly into the upper portion of the conduit. The proportion of air to combustible vapours coming through conduit **37** is controlled by needle valve **42**.

As nozzle **21** enters directly into the lower end of passage **15**, suction in the inlet manifold will, in turn, create a suction on nozzle **21** which will cause a mixture of atomised fuel and air to be drawn directly into the intake manifold. This is found to be desirable when starting the engine, particularly in cold weather, when there might not be an adequate supply of vapour in the vaporising chamber , or the mixture of air and vapour passing through passage **13** might be to “lean” to cause a prompt starting of the engine. At such times, closing the choke valve **13b** will cause the maximum suction to be exerted on nozzle **21** and the maximum amount of air and atomised fuel to be drawn directly into the intake manifold. After the engine has been started, only a small portion of the combustible air and vapour mixture necessary for proper operation of the engine is drawn through nozzle **21** as the choke valve will then be open to a greater extent and substantially all of the air and vapour mixture necessary for operation of the engine will be drawn through the lower end **20** of passage **15**, around nozzle **21**.

Conduit **37** extending from fuel chamber **31** to a point above butterfly valve **16** provides an adequate supply of fuel when the engine is idling with vale **16** closed or nearly closed.

The casings forming chambers **1**, **4** and **7**, will be provided with the necessary openings, to subsequently be closed, so that the various parts may be assembled, and subsequently adjusted or repaired.

The intake stroke of the engine creates a suction in the intake manifold, which in turn causes air to be drawn past spring valve **14** into passage **13** and simultaneously a portion of the dry fuel vapour from the top of vaporising chamber **1** is drawn through opening **17** past valve **17a** to mix with the air moving through the passage. This mixture then passes through passage **15** to the intake manifold and engine cylinders.

The drawing of the dry fuel vapour into passage **13** creates a partial vacuum in chamber **1** which causes air to be drawn into chamber **7** around heated chamber **4** from where it passes through connection **12** and valve **19**, into passage **13** and through pipe **9** and valve **18** into aerator **10**, from which it bubbles up through the liquid fuel in the bottom of chamber **1** to vaporise more liquid fuel.

To assist in maintaining a supply of dry fuel vapour in the upper portion of vaporising chamber **1**, the carburettor is provided with means for atomising a portion of the liquid fuel in vaporising chamber **1**. This atomising means preferably is comprised of a double-acting pump which is operated by the suction existing in the intake manifold of the engine.

The double-acting pump is comprised of a pair of cylinders **43** which have their lower ends located in the vaporising chamber **1**, and each of which has a reciprocating pump piston **44** mounted in it. Pistons **44** have rods **45** extending from their upper ends, passing through cylinders **46** and have pistons **47** mounted on them within the cylinders **46**.

Cylinders **46** are connected at each end to a distributing valve **V** which connects the cylinders alternately to the intake manifold so that the suction in the manifold will cause the two pistons **44** to operate as a double-acting suction pump.

The distributing valve **V** is comprised of a pair of discs **48** and **49** between which is located a hollow oscillatable chamber **50** which is constantly subjected to the suction existing in the intake manifold through connection **51**

having a valve **52** in it. Chamber **50** has a pair of upper openings and a pair of lower openings. These openings are so arranged with respect to the conduits leading to the opposite ends of cylinders **46** that the suction of the engine simultaneously forces one piston **47** upwards while forcing the other one downwards.

The oscillatable chamber **50** has a T-shaped extension **53**. The arms of this extension are engaged alternately by the upper ends of the piston rods 45, so as to cause valve **V** to connect cylinders **46** in sequence to the intake manifold.

Spring **54** causes a quick opening and closing of the ports leading to the cylinders **46** so that at no time will the suction of the engine be exerted on both of the pistons **47**. The tension between discs **48** and **49** and the oscillatable chamber **50** may be regulated by screw **55**.

The particular form of the distributing valve **V** is not claimed here so a further description of operation is not necessary. As far as the present invention is concerned, any form of means for imparting movement to pistons **47** may be substituted for the valve **V** and its associated parts.

The cylinders **43** are each provided with inlets and outlets **56** and **57**, each located below the fuel level in chamber **1**. The inlets **56** are connected to horizontally and upwardly extending conduits **58** which pass through the carburettor to the outside. The upper ends of these conduits are enlarged at **59** and are provided with a vertically extending slot **60**. The enlarged ends **59** are threaded on the inside to accept plugs **61**. The position of these plugs with respect to slots **60** determines the amount of air which may pass through the slots **60** and into cylinder **43** on the suction stroke of the pistons **44**.

The upper walls of the horizontal portions of conduits **58** have an opening **62** for the passage of liquid fuel from chamber **1**. The extent to which liquid fuel may pass through these openings is controlled by needle valves **63**, whose stems **64** pass up through and out of the carburettor and terminate in knurled adjusting nuts **65**.

The horizontal portion of each conduit **58** is also provided with a check valve **66** (shown in **Fig.10**) which allows air to be drawn into the cylinders through conduits **58** but prevents liquid fuel from being forced upwards through the conduits on the down stroke of pistons **44**.

Outlets **57** connect with horizontal pipes **67** which merge into a single open-ended pipe **68** which extends upwards. The upper open end of this pipe terminates about half way up the height of the vaporising chamber **1** and is provided with a bail **69** which carries a deflecting plate **70** positioned directly over the open end of pipe **68**.

The horizontal pipes **67** are provided with check valves **71** which permit the mingled air and fuel to be forced from cylinders **43** by the pistons **44**, but which prevent fuel vapour from being drawn from chamber **1** into cylinders **43**.

When operating, pistons **44** on the ‘up’ strokes, draw a charge of air and liquid fuel into cylinders **43**, and on the ‘down’ stroke, discharge the charge in an atomised condition through pipes **67** and **68**, against deflecting plate **70** which further atomises the particles of liquid fuel so that they will readily vaporise. Any portions of the liquid fuel which do not vaporise, drop down into the supply of liquid fuel in the bottom of the vaporising chamber where they are subjected to the vaporising influence of the bubbles of heated air coming from the aerator **10**, and may again pass into the cylinders **43**.

As previously stated, the vaporised fuel for introduction into the intake manifold of the engine, is taken from the upper portion of the vaporising chamber **1**. To ensure that the vapour in this portion of the chamber shall contain no, or substantially no, entrained droplets of liquid fuel, chamber **1** is divided into upper and lower portions by the walls **71** and **72** which converge from all directions to form a central opening **73**. With the vaporising chamber thus divided into upper and lower portions which are connected only by the relatively small opening **73**, any droplets entrained by the bubbles rising from the aerator **10**, will come into contact with the sloping wall **72** and be deflected back into the main body of liquid fuel in the bottom of the chamber. Likewise, the droplets of atomised fuel being forced from the upper end of pipe **68** will, on striking plate **70**, be deflected back into the body of liquid fuel and not pass into the upper portion of the chamber.

In order that the speed of operation of the atomising pump may be governed by the speed at which the engine is running, and further, that the amount of air admitted from chamber **7** to the aerator **10**, and to passage **13** through connection **12**, may be increased as the speed of the engine increases, the valves **18**, **19** and **52** and butterfly valve **16** are all connected by a suitable linkage **L** so that as butterfly valve **16** is opened to increase the speed of the engine, valves **18**, **19** and **52** will also be opened.

As shown in **Fig.2**, the passage of the exhaust gasses from the engine to the heating chamber **4**, located between the vaporising chamber and the air chamber **7**, is controlled by valve **74**. The opening and closing of valve **74** is controlled by a thermostat in accordance with the temperature inside chamber **4**, by means of an adjustable metal

rod 75 having a high coefficient of expansion, whereby the optimum temperature may be maintained in the vaporising chamber, irrespective of the surrounding temperature.

From the foregoing description, it will be understood that the present invention provides a carburettor for supplying to internal combustion engines, a comingled mixture of air and liquid fuel vapour free from microscopic droplets of liquid fuel which would burn rather than explode in the cylinders and that a supply of such dry vaporised fuel is constantly maintained in the carburettor.

The Second High MPG Carburettor of Charles Pogue

US Patent 1,997,497 9th April 1935 Inventor: Charles N. Pogue

CARBURETTOR

This patent describes a carburettor design which was able to produce very high mpg figures using the gasoline available in the USA in the 1930s but which is no longer available as the oil industry does not want functional high mpg carburettors to be available to the public.

DESCRIPTION

This invention relates to a device for obtaining an intimate contact between a liquid in a truly vaporous state and a gas, and particularly to such a device which may serve as a carburettor for internal combustion engines and is an improvement on the form of device shown in my Patent No. 1,938,497, granted on 5th December 1933.

In carburettors commonly used for supplying a combustible mixture of air and liquid fuel to internal combustion engines, a relatively large amount of the atomised liquid fuel is not vaporised and enters the engine cylinder more or less in the form of microscopic droplets. When such a charge is ignited in the engine cylinder, only that portion of the liquid fuel which has been converted into the vaporous, and consequently molecular state, combines with the air to give an explosive mixture. The remaining portion of the liquid fuel which is drawn into the engine cylinders remains in the form of small droplets and does not explode imparting power to the engine, but instead burns with a flame and raises the engine temperature above that at which the engine operates most efficiently, i.e. from 160^{O} F. to 180^{O} F.

In my earlier patent, there is shown and described a form of carburettor in which the liquid fuel is substantially completely vaporised prior to its introduction into the engine cylinders, and in which, means are provided for maintaining a reverse supply of “dry” vapour available for introduction into the engine cylinder. Such a carburettor has been found superior to the standard type of carburettor referred to above, and to give a better engine performance with far less consumption of fuel.

It is an object of the present invention to provide a carburettor in which the liquid fuel is broken up and prepared in advance of and independent of the suction of the engine and in which a reserve supply of dry vapour will be maintained under pressure, ready for introduction into the engine cylinder at all times. It is also an object of the invention to provide a carburettor in which the dry vapour is heated to a sufficient extent prior to being mixed with the main supply of air which carries it into the engine cylinder, to cause it to expand so that it will be relatively lighter and will become more intimately mixed with the air, prior to explosion in the engine cylinders.

I have found that when the reserve supply of dry vapour is heated and expanded prior to being mixed with the air, a greater proportion of the potential energy of the fuel is obtained and the mixture of air and fuel vapour will explode in the engine cylinders without any apparent burning of the fuel which would result in unduly raising the operating temperature of the engine.

More particularly, the present invention comprises a carburettor in which liquid fuel vapour is passed from a main vaporising chamber under at least a slight pressure, into and through a heated chamber where it is caused to expand and in which droplets of liquid fuel are either vaporised or separated from the vapour , so that the fuel finally introduced into the engine cylinders is in the true vapour phase. The chamber in which the liquid fuel vapour is heated and caused to expand, is preferably comprised of a series of passages through which the vapour and exhaust gases from the engine pass in tortuous paths in such a manner that the exhaust gasses are brought into heat interchange relation with the vapour and give up a part of their heat to the vapour, thus causing heating and expansion of the vapour.

The invention will be further described in connection with the accompanying drawings, but this further disclosure and description is to be taken merely as an exemplification of the invention and the invention is not limited to the embodiment so described.

DESCRIPTION OF THE DRAWINGS

Fig.1 is a vertical cross-sectional view through a carburettor embodying my invention.

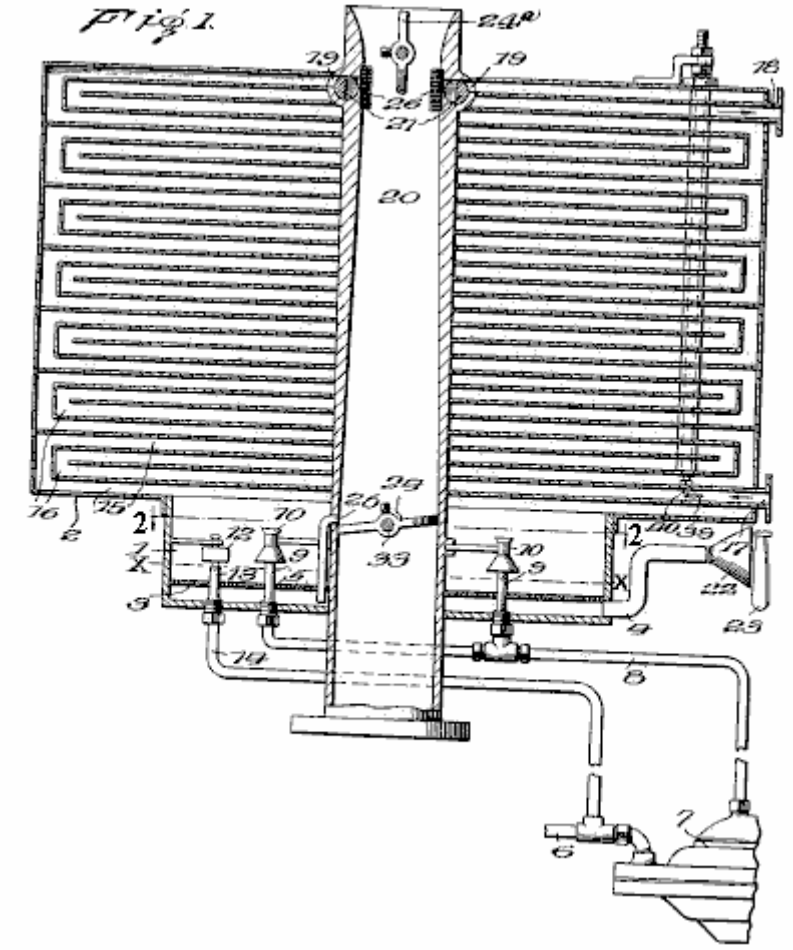


Fig.2 is a horizontal sectional view through the main vaporising or atomising chamber, taken on line 2--2 of Fig.1

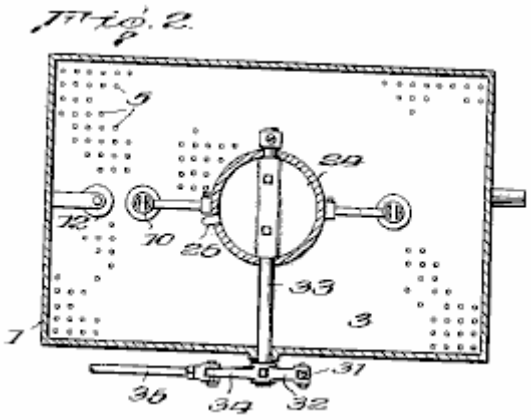


Fig.3 is a side elevation of the carburettor.

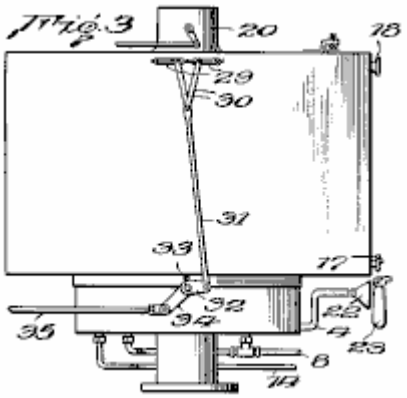


Fig.4 is a detail sectional view of one of the atomising nozzles and its associated parts

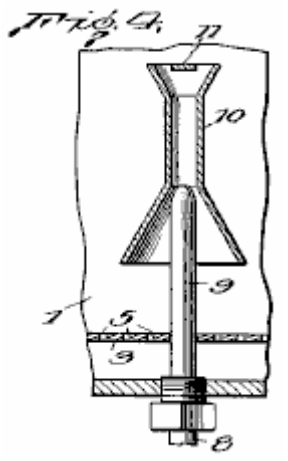


Fig.5 is a detail cross-sectional view showing the means for controlling the passage of gasses from the vapour expanding chamber into the intake manifold of the engine.

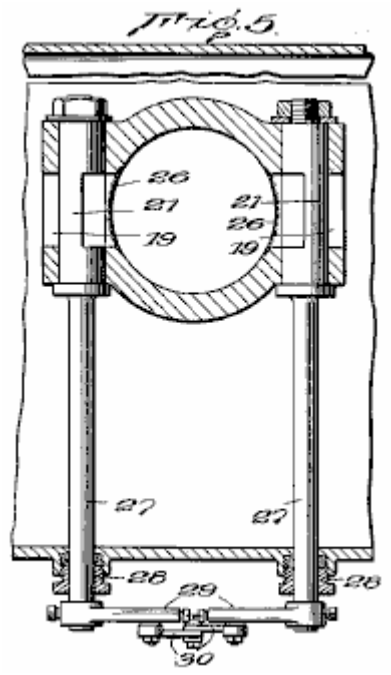


Fig.6 is a perspective view of one of the valves shown in Fig.5

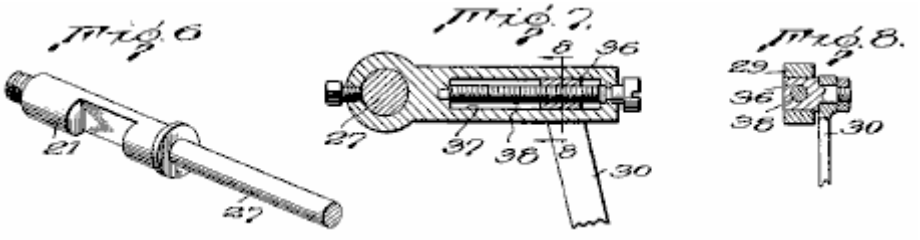


Fig.7 is a cross-sectional view showing means for adjusting the valves shown in Fig.5

Fig.8 is a cross-sectional view on line 8--8 of Fig.7

Referring now to the drawings, the numeral **1** indicates a main vaporising and atomising chamber for the liquid fuel located at the bottom of, and communicating with, a vapour heating and expanding chamber **2**.

The vaporising chamber is provided with a perforated false bottom **3** and is normally filled with liquid fuel to the level **x**. Air enters the space below the false bottom **3** via conduit **4** and passes upwards through perforations **5** in the false bottom and then bubbles up through the liquid fuel, vaporising a portion of it.

To maintain the fuel level **x** in chamber **1**, liquid fuel passes from the usual fuel tank (not shown) through pipe **8** into and through a pair of nozzles **9** which have their outlets located in chamber **1**, just above the level of the liquid fuel in it. The pump **7** may be of any approved form but is preferably of the diaphragm type, as such fuel pumps are now standard equipment on most cars.

The nozzles **9** are externally threaded at their lower ends to facilitate their assembly in chamber **1** and to permit them to be removed readily, should cleaning be necessary.

The upper ends of nozzles **9** are surrounded by venturi tubes **10**, having a baffle **11**, located at their upper ends opposite the outlets of the nozzles. The liquid fuel being forced from the ends of nozzles **9** into the restricted portions of the Venturi tubes, causes a rapid circulation of the air and vapour in the chamber through the tubes **10** and brings the air and vapour into intimate contact with the liquid fuel, with the result that a portion of the liquid fuel is vaporised. The part of the liquid fuel which is not vaporised, strikes the baffles **11** and is further broken up and deflected downwards into the upward-flowing current of air and vapour.

Pump **7** is regulated to supply a greater amount of liquid fuel to the nozzles **9** than will be vaporised. The excess drops into chamber **1** and causes the liquid to be maintained at the indicated level. When the liquid fuel rises above that level, a float valve **12** is lifted, allowing the excess fuel to flow out through overflow pipe **13** into pipe **14** which leads back to pipe **6** on the intake side of pump **7**. Such an arrangement allows a large amount of liquid fuel to be circulated by pump **7** without more fuel being withdrawn from the fuel tank than is actually vaporised and consumed in the engine. As the float valve **12** will set upon the end of the outlet pipe **13** as soon as the liquid level drops below the indicated level, there is no danger of vapour passing into pipe **14** and from there into pump **7** and interfere with its normal operation.

The upper end of the vaporising and atomising chamber **1** is open and vapour formed by air bubbling through the liquid fuel in the bottom of the chamber and that formed as the result of atomisation at nozzles **9**, pass into the heating and expanding chamber **2**. As is clearly shown in **Fig.1**, chamber **2** comprises a series of tortuous passages **15** and **16** leading from the bottom to the top. The fuel vapour passes through passages **15** and the exhaust gasses of the engine pass through passages **16**, a suitable entrance **17** and exit **18** being provided for that purpose.

The vapour passing upwards in a zigzag path through passages **15**, will be brought into heat interchange relation with the hot walls of the passages **16** traversed by the hot exhaust gasses. The total length of the passages **15** and **16** is such that a relatively large reserve supply of the liquid fuel is always maintained in chamber **2**, and by maintaining the vapour in heat interchange relation with the hot exhaust gasses for a substantial period, the vapour will absorb sufficient heat to cause it to expand, with the result that when it is withdrawn from the top of chamber 2, it will be in the true vapour phase, and due to expansion, relatively light.

Any minute droplets of liquid fuel entrained by the vapour in chamber **1** will precipitate out in the lower passages **15** and flow back into chamber **1**, or else be vaporised by the heat absorbed from the exhaust gasses during its passage through chamber **2**.

The upper end of vapour passage **15** communicates with openings **19** adjacent to the upper end of a down-draft air tube **20** leading to the intake manifold of the engine. Valves **21** are interposed in openings **19**, so that the passage of the vapour through them into the air tube may be controlled. Valves **21** are preferably of the rotary plug type and are controlled as described below.

Suitable means are provided for causing the vapour to be maintained in chamber **2,** under a pressure greater than atmospheric, so that when the valves **21** are opened, the vapour will be forced into air tube **20** independent of the engine suction. Such means may comprise an air pump (not shown) for forcing air through pipe **4** into chamber **1** beneath the false bottom **3**, but I prefer merely to provide pipe 4 with a funnel-shaped inlet end **22** and placement just behind the usual engine fan **23**. This causes air to pass through pipe **4** with sufficient force to maintain the desired pressure in chamber **2**, and the air being drawn through the radiator by the fan will be preheated prior to its introduction into chamber **1** and hence will vaporise greater amounts of the liquid fuel. If desired, pipe **4** may be surrounded by an electric or other heater, or exhaust gasses from the engine may be passed around it to further preheat the air passing through it prior to its introduction into the liquid fuel in the bottom of chamber **1**.

Air tube **20** is provided with a butterfly throttle valve **24** and a choke valve **24a**, as is customary with carburettors used for internal combustion engines. The upper end of air tube **20** extends above chamber **2** a distance sufficient to receive an air filter and/or silencer, if desired.

A low-speed or idling jet **25** has its upper end communicating with the passage through air tube **20** adjacent to the throttling valve **24** and its lower end extending into the liquid fuel in the bottom of chamber **1,** for supplying fuel to the engine when the valves are in a position such as to close the passages **19**. However, the passage through idling jet **25** is so small that under normal operations, the suction on it is not sufficient to lift fuel from the bottom of chamber **1**.

To prevent the engine from backfiring into vapour chamber **2**, the ends of the passages **19** are covered with a fine mesh screen **26** which, operating on the principle of the miner's lamp, will prevent the vapour in chamber **2** from exploding in case of a backfire, but which will not interfere substantially with the passage of the vapour from chamber **2** into air tube **20** when valves **21** are open. Air tube **20** is preferably in the form of a venturi with the greatest restriction being at that point where the openings **19** are located, so that when valves **21** are opened, there will be a pulling force on the vapour caused by the increased velocity of the air at the restricted portion of air tube **20** opposite the openings **19**, as well as an expelling force on them due to the pressure in chamber **2**.

As shown in **Fig.3**, the operating mechanism of valves **21** is connected to the operating mechanism for throttle valve **24**, so that they are opened and closed simultaneously with the opening and closing of the throttle valve, ensuring that the amount of vapour supplied to the engine will, at all times, be in proportion to the demands placed upon the engine. To that end, each valve **21** has an extension, or operating stem **27**, protruding through one of the side walls of the vapour-heating and expanding chamber **2**. Packing glands **28** of ordinary construction, surround stems **27** where they pass through the chamber wall, to prevent leakage of vapour at those points.

Operating arms **29** are rigidly secured to the outer ends of stems **27** and extend towards each other. The arms are pivotally and adjustably connected to a pair of links **30** which, at their lower ends are pivotally connected to an operating link **31**, which in turn, is pivotally connected to arm **32** which is rigidly secured on an outer extension **33** of the stem of the throttle valve **24**. Extension **33** also has rigidly connected to it, arm **34** to which is connected operating link **35** leading from the means for accelerating the engine.

The means for adjusting the connection from the upper ends of links **30** to valve stems **27** of valves **21**, so that the amount of vapour delivered from chamber 2 may be regulated to cause the most efficient operation of the particular engine to which the carburettor is attached, comprises angular slides **36**, to which the upper ends of links **30** are fastened, and which cannot rotate but can slide in guideways **37** located in arms **29**. Slides **36** have threaded holes through which screws **38** pass. Screws **38** are rotatably mounted in arms **29**, but are held against longitudinal movement so that when they are rotated, slides **36** will be caused to move along the guideways **37** and change the relative position of links **30** to the valve stems **27**, so that a greater or less movement, and consequently, a greater or less opening of the ports **19** will take place when throttle valve **24** is operated.

For safety, and for most efficient operation of the engine, the vapour in chamber **2** should not be heated or expanded beyond a predetermined amount, and in order to control the extent to which the vapour is heated, and consequently, the extent to which it expands, a valve **39** is located in the exhaust passage **16** adjacent to inlet **17**. Valve **39** is preferably theromstatically controlled, as for example, by an expanding rod thermostat **40**, which extends through chamber **2**. However, any other means may be provided for reducing the amount of hot exhaust gasses entering passage **16** when the temperature of the vapour in the chamber reaches or exceeds the optimum.

The carburettor has been described in detail in connection with a down-draft type of carburettor, but it is to be understood that its usefulness is not to be restricted to that particular type of carburettor, and that the manner in which the mixture of air and vapour is introduced into the engine cylinders is immaterial as far as the advantages of the carburettor are concerned.

The term "dry vapour" is used to define the physical condition of the liquid fuel vapour after removal of liquid droplets or the mist which is frequently entrained in what is ordinarily termed a vapour.

From the foregoing description it will be seen that the present invention provides a carburettor in which the breaking up of the liquid fuel for subsequent use is independent of the suction created by the engine, and that after the liquid fuel is broken up, it is maintained under pressure in a heated space for a length of time sufficient to permit all entrained liquid particles to be separated or vaporised and to permit the dry vapour to expand prior to its introduction into and admixture with the main volume of air passing into the engine cylinders.

The Third High MPG Carburettor of Charles Pogue

US Patent 2,026,798 7th January 1936 Inventor: Charles N. Pogue

CARBURETTOR

This patent describes a carburettor design which was able to produce very high mpg figures using the gasoline available in the USA in the 1930s but which is no longer available as the oil industry does not want functional high mpg carburettors to be available to the public.

DESCRIPTION

This invention relates to carburettors suitable for use with internal combustion engines and is an improvement on the carburettors shown in my Patents Nos. 1,938,497, granted on 5th December 1933 and 1,997,497 granted on 9th April 1935.

In my earlier patents, an intimate contact between such as the fuel used for internal combustion engines, and a gas such as air, is obtained by causing the gas to bubble up through a body of the liquid. The vaporised liquid passes into a vapour chamber which preferably is heated, and any liquid droplets are returned to the body of the liquid, with the result that the fuel introduced into the combustion chambers is free of liquid particles , and in the molecular state so that an intimate mixture with the air is obtained to give an explosive mixture from which nearer the maximum energy contained in the liquid fuel is obtained. Moreover, as there are no liquid particles introduced into the combustion chambers, there will be no burning of the fuel and consequently, the temperature of the engine will not be increased above that at which it operates most efficiently.

In my Patent No. 1,997,497, the air which is to bubble up through the body of the liquid fuel is forced into and through the fuel under pressure and the fuel vapour and air pass into a chamber where they are heated and caused to expand. The introduction of the air under pressure and the expansion of the vaporous mixture ensures a sufficient pressure being maintained in the vapour heating and expanding chamber, to cause at least a portion of it to be expelled from it into the intake manifold as soon as the valve controlling the passage to it is opened.

In accordance with the present invention, improved means are provided for maintaining the vaporous mixture in the vapour-heating chamber under a predetermined pressure, and for regulating such pressure so that it will be at the optimum for the particular conditions under which the engine is to operate. Such means preferably comprises a reciprocating pump operated by a vacuum-actuated motor for forcing the vapour into and through the chamber. The pump is provided with a suitable pressure-regulating valve so that when the pressure in the vapour-heating chamber exceeds the predetermined amount, a portion of the vapour mixture will be by-passed from the outlet side to the inlet side of the pump, and so be recirculated.

The invention will be described further in connection with the accompanying drawings, but such further disclosure and description is to be taken merely as an exemplification of the invention, and the invention is not limited to that embodiment of the invention.

DESCRIPTION OF THE DRAWINGS

Fig.1 is a side elevation of a carburettor embodying the invention.

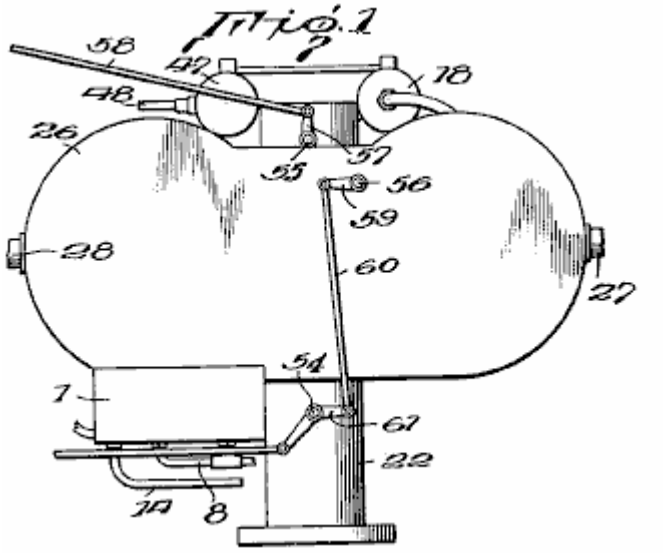


Fig.2 is a plan view of the carburettor

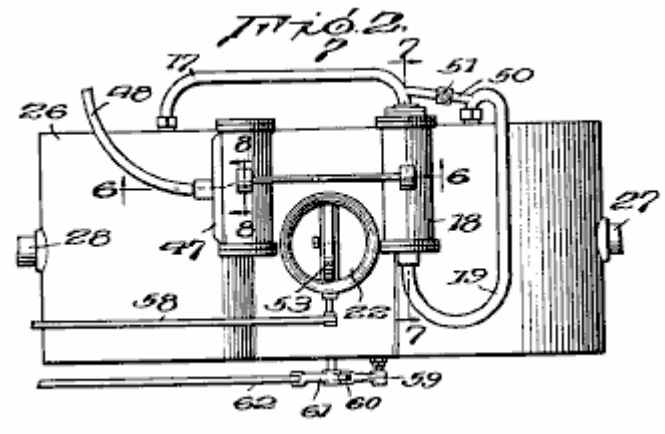


Fig.3 is an enlarged vertical section view.

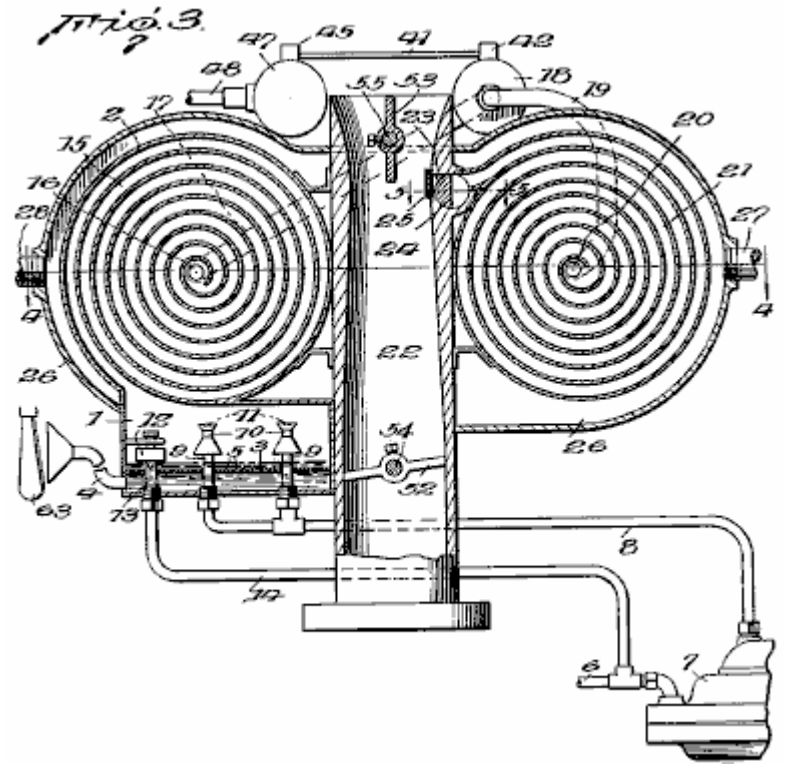


Fig.4 is a transverse sectional view on line 4--4 of Fig.3

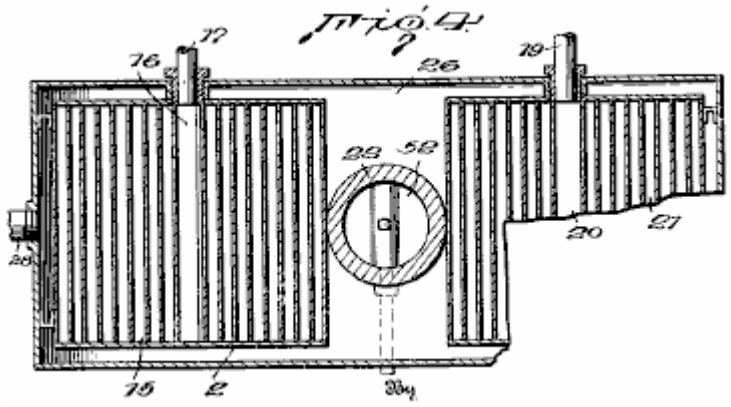


Fig.5 is a detail sectional view on line 5--5 of Fig.3

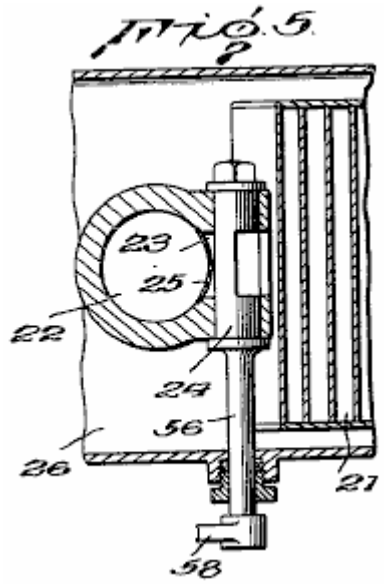


Fig.6 is a transverse sectional view through the pump and actuating motor, taken on line 6--6 of Fig.2

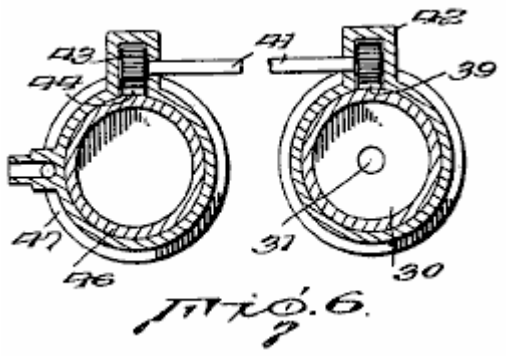


Fig.7 is a longitudinal sectional view through the pump taken on line 7--7 of Fig.2

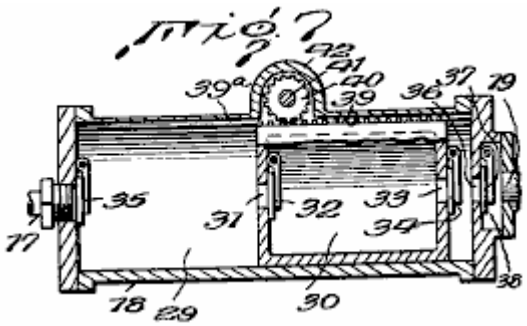
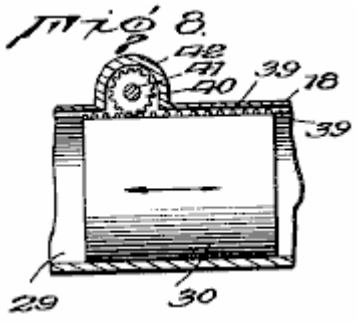


Fig.8 is a longitudinal sectional view through a part of the pump cylinder, showing the piston in elevation.



In the drawings, a vaporising and atomising chamber **1** is located at the bottom of the carburettor and has an outlet at its top for the passage of fuel vapour and air into a primary vapour-heating chamber **2**.

The vaporising chamber **1** is provided with a perforated false bottom **3** and is normally filled with liquid fuel to the level indicated in **Fig.1**. Air is introduced via conduit **4** into the space below the false bottom **3**, and then through the perforations **5** in the false bottom which breaks it into a myriad of fine bubbles, which pass upwards through the liquid fuel above the false bottom.

Liquid fuel for maintaining the level indicated in chamber **1** passes from the usual fuel tank (not shown) through pipe **6**, and is forced by pump **7** through pipe **8** through a pair of nozzles **9** having their outlets located in chamber **1**, just above the level of the liquid fuel in it. Pump **7** may be of any approved form but is preferably of the diaphragm type, as such fuel pumps are now standard equipment on most cars.

The nozzles **9** are externally threaded at their lower ends to facilitate their assembly in chamber **1** and to permit them to be readily removed should cleaning become necessary.

The upper ends of nozzles **9** are surrounded by venturi tubes **10** having baffles **11** located at their upper ends opposite the outlets of the nozzles, as is shown and described in detail in my Patent No. 1,997,497. The liquid fuel being forced from the ends of nozzles **9** into the restricted portions of the venturi tubes, causes a rapid circulation of the air and vapour in the chamber through tubes **10** and brings the air and vapour into intimate contact with the liquid fuel, with the result that a portion of the liquid fuel is vaporised. Unvaporised portions of the liquid fuel strike the baffles **11** and are thereby further broken up and deflected downwards into the upward-flowing current of air and vapour.

Pump **7** is regulated to supply a greater amount of liquid fuel to nozzles **9** than will be vaporised. The excess liquid fuel drops into chamber **1** which causes the liquid there to be maintained at the indicated level. When the liquid fuel rises above that level, float valve **12** opens and the excess fuel flows through overflow pipe **13** into pipe **14** which leads back to pipe **6** on the intake side of pump **7**. Such an arrangement permits a large amount of liquid fuel to be circulated by pump 7 without more fuel being withdrawn from the fuel tank than is actually vaporised and consumed by the engine. As float valve **12** will set upon the end of the outlet pipe **13** as soon as the liquid level drops below the indicated level, there is no danger of vapour passing into pipe **14** and thence into pump **7** to interfere with its normal operation.

The amount of liquid fuel vaporised by nozzles **9** and by the passage of air through the body of liquid, is sufficient to provide a suitably enriched vaporous mixture for introducing into the passage leading to the intake manifold of the engine, through which the main volume of air passes.

Vapour formed by air bubbling through the liquid fuel in the bottom of chamber **1** and that formed by the atomisation at the nozzles **9**, pass from the top of that chamber into the primary heating chamber **2**. As is clearly shown in **Fig.1**, chamber **2** comprises a relatively long spiral passage **15** through which the vaporous mixture gradually passes inwards to a central outlet **16** to which is connected a conduit **17** leading to a reciprocating pump **18** which forces the vaporous mixture under pressure into conduit **19** leading to a central inlet **20** of a secondary heating chamber **21**, which like the primary heating chamber, comprises a relatively long spiral. The vaporous mixture gradually passes outwards through the spiral chamber **21** and enters a downdraft air tube **22**, leading to the intake manifold of the engine, through an outlet **23** controlled by a rotary plug valve **24**.

To prevent the engine from backfiring into vapour chamber **2**, the ends of passage **19** are covered with a fine mesh screen **25**, which, operating on the principle of a miner's lamp, will prevent the vapour in chamber **2** from exploding in case of a backfire, but will not interfere substantially with the passage of the vapour from chamber **21** into air tube **22** when valve **24** is open.

The air tube **22** is preferably in the form of a venturi with the greatest constriction being at that point where outlet **23** is located, so that when valve **24** is opened, there will be a pulling force on the vaporous mixture due to the increased velocity of the air at the restricted portion of the air tube opposite outlet **23**, as well as an expelling force on it due to the pressure maintained in chamber **21** by pump **18**.

Both the primary and secondary spiral heating chambers **15** and **21**, and the central portion of air tube **22** are enclosed by a casing **26** having an inlet **27** and an outlet **28** for a suitable heating medium such as the gasses coming from the exhaust manifold.

Pump **18**, used to force the vaporous mixture from primary heating chamber **2** into and through the secondary chamber **21**, includes a working chamber **29** for hollow piston **30**, provided with an inlet **31** controlled by valve **32**, and an outlet **33** controlled by a valve **34**. The end of the working chamber **29** to which is connected conduit **17**, which conducts the vaporous mixture from primary heating chamber **2**, has an inlet valve **35**, and the opposite end of the working chamber has an outlet **36** controlled by valve **37** positioned in an auxiliary chamber **38**, to which is connected outlet pipe **19** which conducts the vaporous mixture under pressure to the secondary heating chamber **21**. Each of the valves **32**, **34**, **35** and **37** is of the one-way type. They are shown as being gravity-actuated flap valves, but it will be understood that spring-loaded or other types of one-way valves may be used if desired.

One side of piston **30** is formed with a gear rack **39** which is received in a groove **39a** of the wall forming the cylinder of the pump. The gear rack **39** engages with an actuating spur gear **40** carried on one end of shaft **41** and operating in a housing **42** formed on the pump cylinder. The other end of shaft **41** carries a spur gear **43**, which engages and is operated by a gear rack **44** carried on a piston **46** of a double-acting motor **47**. The particular construction of the double-acting motor **47** is not material, and it may be of a vacuum type commonly used for operating windscreen wipers on cars, in which case a flexible hose **48** would be connected with the intake manifold of the engine to provide the necessary vacuum for operating the piston **45**.

Under the influence of the double-acting motor **47**, the piston **30** of the pump has a reciprocatory movement in the working chamber **29**. Movement of the piston towards the left in **Fig.7** tends to compress the vaporous mixture in the working chamber between the end of the piston and the inlet from pipe **17**, and causes valve **35** to be forced tightly against the inlet opening. In a like manner, valves **32** and **34** are forced open and the vaporous mixture in that portion of the working chamber is forced through the inlet **31** in the end of the piston **30**, into the interior of the piston, where it displaces the vaporous mixture there and forces it into the space between the right-hand end of the piston and the right-hand end of the working chamber. The passage of the vaporous mixture into the right-hand end of the working chamber is supplemented by the partial vacuum created there when the piston moves to the left. During such movement of the piston, valve **37** is maintained closed and prevents any sucking back of the vaporous mixture from the secondary heating chamber **21**.

When motor **47** reverses, piston **30** moves to the right and the vaporous mixture in the right-hand end of the working chamber is forced past valve **37** through pipe **19** into the secondary heating chamber **21**. At the same time, a vacuum is created behind piston **30** which results in the left-hand end of the working chamber being filled again with the vaporous mixture from the primary heating chamber **2**.

As the operation of pump **47** varies in accordance with the suction created in the intake manifold, it should be regulated so that the vaporous mixture is pumped into the secondary heating chamber at a rate sufficient to maintain a greater pressure there than is needed. In order that the pressure in the working chamber may at all times be maintained at the optimum, a pipe **50** having an adjustable pressure-regulating valve **51** is connected between the inlet and outlet pipes **17** and **19**. Valve **51** will permit a portion of the vaporous mixture discharged

from the pump to be bypassed to inlet **17** so that a pressure predetermined by the seating of valve **51** will at all times be maintained in the second heating chamber **21**.

Air tube **22** is provided with a butterfly throttle valve **52** and a choke valve **53**, as is usual with carburettors adapted for use with internal combustion engines. Operating stems **54**, **55** and **56** for valves **52**, **53** and **24** respectively, extend through casing **26**. An operating arm **57** is rigidly secured to the outer end of stem **55** and is connected to a rod **58** which extends to the dashboard of the car, or some other place convenient to the driver. The outer end of stem **56** of valve **24** which controls outlet **23** from the secondary heating chamber **21** has one end of an operating arm **59** fixed securely to it. The other end is pivotally connected to link **60** which extends downwards and pivotally connects to one end of a bell crank lever **61**, rigidly attached to the end of stem **54** of throttle valve **52**. The other end of the bell crank lever is connected to an operating rod **62** which, like rod **58**, extends to a place convenient to the driver. Valves **24** and **52** are connected for simultaneous operation so that when the throttle valve **52** is opened to increase the speed of the engine, valve **24** will also be opened to admit a larger amount of the heated vaporous mixture from the secondary heating chamber **21**.

While the suction created by pump **18** ordinarily will create a sufficient vacuum in the primary heating chamber **2** to cause air to be drawn into and upwards through the body of liquid fuel in the bottom of vaporising chamber **1**, in some instances it may be desirable to provide supplemental means for forcing the air into and up through the liquid, and in such cases an auxiliary pump may be provided for that purpose, or the air conduit **4** may be provided with a funnel-shaped intake which is positioned behind the engine fan **63** which is customarily placed behind the engine radiator.

The foregoing description has been given in connection with a downdraft type of carburettor, but it is to be understood that the invention is not limited to use with such type of carburettors and that the manner in which the mixture of air and vapour is introduced into the engine cylinders is immaterial as far as the advantages of the carburettor are concerned.

Before the carburettor is put into use, the pressure-regulating valve **51** in the bypass pipe **50** will be adjusted so that the pressure best suited to the conditions under which the engine is to be operated, will be maintained in the secondary heating chamber **21**. When valve **51** has thus been set and the engine started, pump **18** will create a partial vacuum in the primary heating chamber **2** and cause air to be drawn through conduit **4** to bubble upwards through the liquid fuel in the bottom of the vaporising and atomising chamber **1** with the resulting vaporisation of a part of the liquid fuel. At the same time, pump **7** will be set into operation and liquid fuel will be pumped from the fuel tank through the nozzles **9** which results in an additional amount of the fuel being vaporised. The vapour resulting from such atomisation of the liquid fuel and the passage of air through the body of the liquid, will pass into and through spiral chamber **1** where they will be heated by the products of combustion in the surrounding chamber formed by casing **26**. The fuel vapour and air will gradually pass inwards through outlet **16** and through conduit **17** to pump **18** which will force them into the secondary heating chamber **21** in which they will be maintained at the predetermined pressure by the pressure-regulating valve **51**. The vaporous mixture is further heated in chamber **21** and passes spirally outward to the valve-controlled outlet **23** which opens into air tube **22** which conducts the main volume of air to the intake manifold of the engine.

The heating of the vaporous mixture in the heating chambers **2** and **21**, tends to cause them to expand, but expansion in chamber **21** is prevented due to the pressure regulating valve **51**. However, as soon as the heated vaporous mixture passes valve **24** and is introduced into the air flowing through intake tube **22**, it is free to expand and thereby become relatively light so that a more intimate mixture with the air is obtained prior to the mixture being exploded in the engine cylinders. Thus it will be seen that the present invention not only provides means wherein the vaporous mixture from heating chamber **21** is forced into the air passing through air tube **22** by a positive force, but it is also heated to such an extent that after it leaves chamber **21** it will expand to such an extent as to have a density less than it would if introduced directly from the vaporising and atomising chamber **1** into the air tube **22**.

The majority of the liquid particles entrained by the vaporous mixture leaving chamber **1** will be separated in the first half of the outermost spiral of the primary heating chamber **2** and drained back into the body of liquid fuel in tank **1**. Any liquid particles which are not thus separated, will be carried on with the vaporous mixture and due to the circulation of that mixture and the application of heat, will be vaporised before the vaporous mixture is introduced into the air tube **22** from the secondary heating chamber **21**. Thus only “dry” vapour is introduced into the engine cylinders and any burning in the engine cylinders of liquid particles of the fuel, which would tend to raise the engine temperature above its most efficient level, is avoided.

While the fullest benefits of the invention are obtained by using both a primary and secondary heating chamber, the primary heating chamber may, if desired, be eliminated and the vaporous mixture pumped directly from the vaporising and atomising chamber **1** into the spiral heating chamber **21**.

From the foregoing description it will be seen that the present invention provides an improvement over the carburettor disclosed in my Patent No. 1,997,497, in that it is possible to maintain the vaporous mixture in the heating chamber **21** under a predetermined pressure, and that as soon as the vaporous mixture is introduced into the main supply of air passing to the intake manifold of the engine, it will expand and reach a density at which it will form a more intimate mixture with the air. Furthermore, the introduction of the vaporous mixture into the air stream in the tube **22**, causes a certain amount of turbulence which also tends to give a more intimate mixture of vapour molecules with the air.

The High MPG Carburettor of Ivor Newberry

US Patent 2,218,922 22nd October 1940 Inventor: Ivor B. Newberry

VAPORIZER FOR COMBUSTION ENGINES

This patent describes a carburettor design which was able to produce very high mpg figures using the gasoline available in the USA in the 1930s but which is no longer available as the oil industry does not want functional high mpg carburettors to be available to the public.

DESCRIPTION

This invention relates to fuel vaporising devices for combustion engines and more particularly, is concerned with improvements in devices of the kind where provision is made for using the exhaust gasses of the engines as a heating medium to aid in the vaporisation of the fuel.

One object of the invention is to provide a device which will condition the fuel in such a manner that its potential energy may be fully utilised, thereby ensuring better engine performance and a saving in fuel consumption, and preventing the formation of carbon deposits in the cylinders of the engine and the production of carbon monoxide and other objectionable gasses.

A further object is to provide a device which is so designed that the fuel is delivered to the cylinders of the engine in a highly vaporised, dry and expanded state, this object contemplating a device which is available as an exhaust box in which the vaporisation and expansion of the liquid components is effected at sub-atmospheric pressures and prior to their being mixed with the air component.

A still further object is to provide a device which will condition the components of the fuel in such a manner that they be uniformly and intimately mixed without the use of a carburettor.

A still further object is to provide a device which will enable the use of various inferior and inexpensive grades of fuel.

DESCRIPTION OF THE DRAWINGS

Fig.1 is an elevational view of the device as applied to the engine of a motor vehicle.

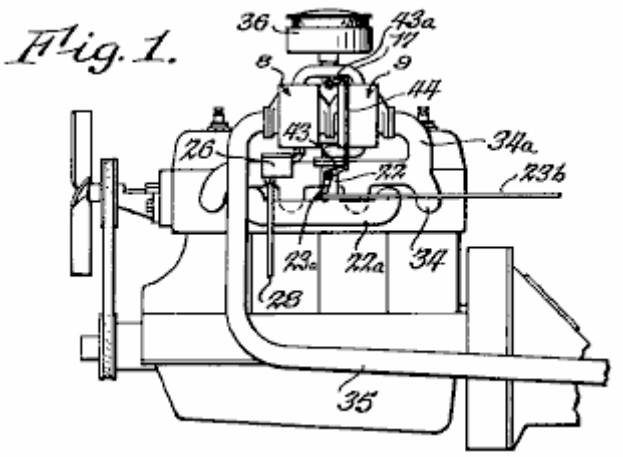


Fig.2 is an enlarged view of the device, partially in elevation and partially in section.

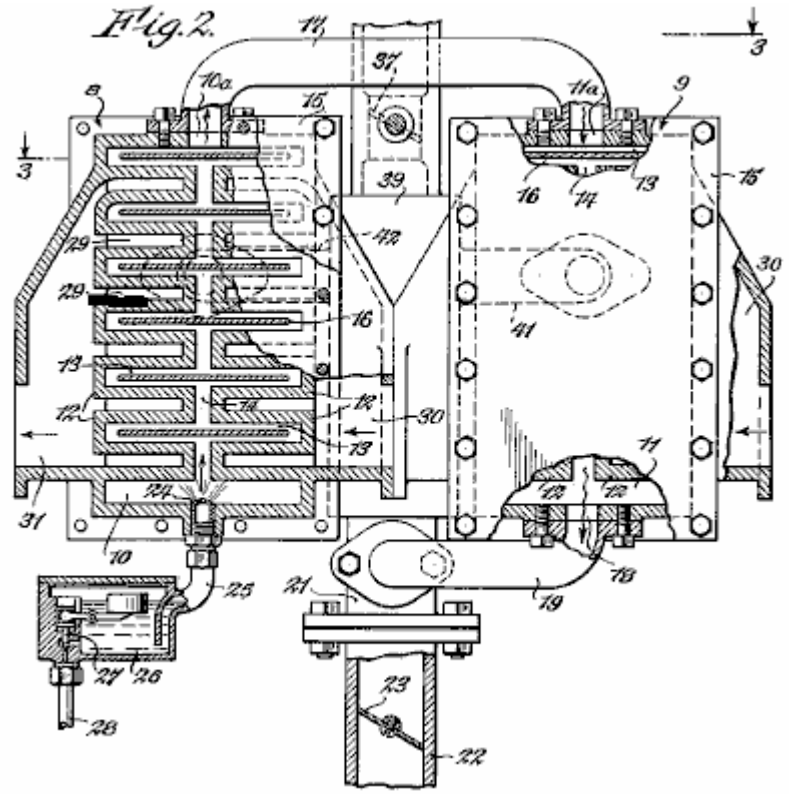


Fig.3 is a section taken along line 3--3 of Fig.2

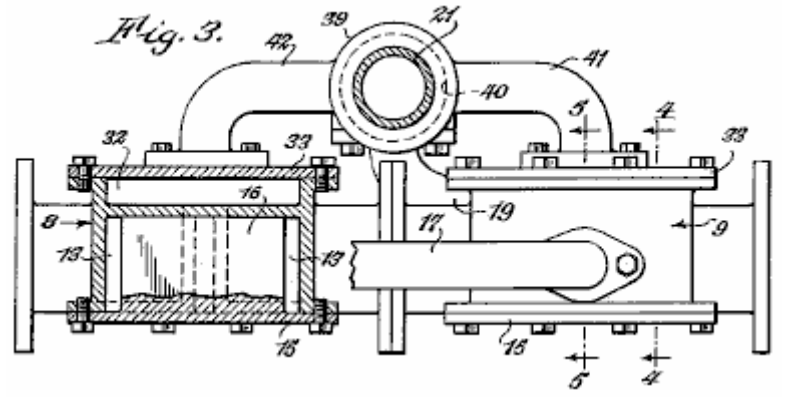


Fig.4 is a section taken along line 4--4 of Fig.3

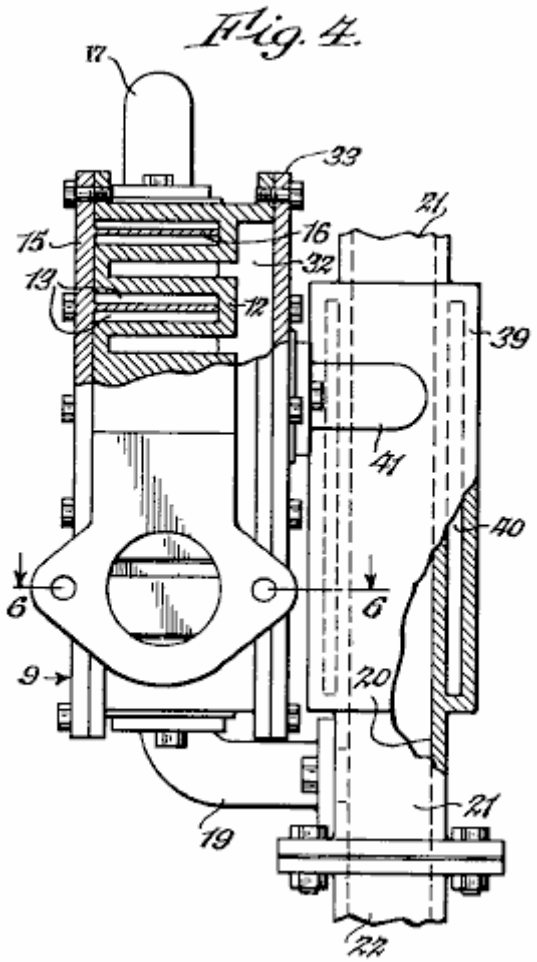


Fig.5 is a fragmentary section taken along line 5--5 of Fig.3

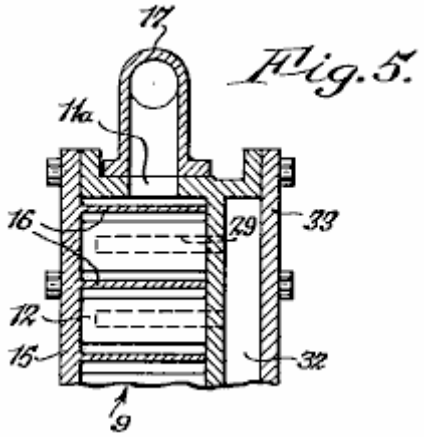
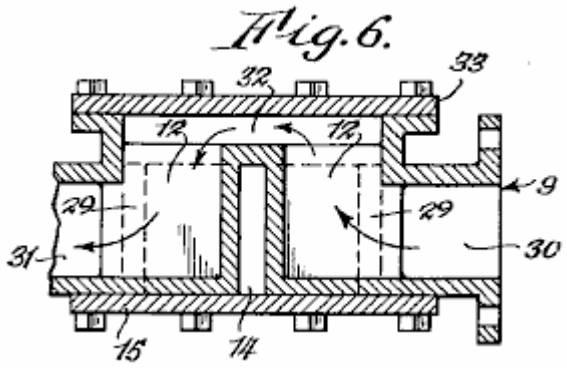


Fig.6 is a section taken along line 6--6 of Fig.4



DESCRIPTION

The device as illustrated, includes similar casings **8** and **9** which are secured together as a unit and which are formed to provide vaporising chambers **10** and **11**, respectively, it being understood that the number of casings may be varied. Two series of ribs **12** are formed in each of the vaporising chambers, the ribs of each series being spaced from one another so as to provide branch passages **13** and being spaced from the ribs of the adjacent series to provide main passages **14** with which the branch passages communicate.

The vaporising chambers are closed by cover plates **15**. The cover plates carry baffles **16** which are supported in the spaces between the ribs **12**. The baffles extend across the main passages **14** and into, but short of the ends of the branch passages **13** to provide tortuous paths. Outlet **10a** of chamber **10** is connected by conduit **17** to inlet **11a** of chamber **11**. Outlet **18** of chamber **11**, is connected by conduit **19** with mixing chamber **20** which is located at the lower end of pipe **21** which in turn is connected to and extension **22** of the intake manifold **22a** of the engine. Extension **22** contains a valve **23** which is connected by a lever **23a** (**Fig.1**) and rod **23b** to a conventional throttle (not shown).

The liquid fuel is introduced into the vaporising chamber **10** through nozzle **24** which is connected by pipe **25** to a reservoir **26** in which the fuel level is maintained by float-controlled valve **27**, the fuel being supplied to the reservoir through pipe **28**.

In accordance with the invention, ribs **12** are hollow, each being formed to provide a cell **29**. The cells in one series of ribs open at one side into an inlet chamber **30**, while the cells of the companion series open at one side into an outlet chamber **31**. The cells of both series of ribs open at their backs into a connecting chamber **32** which is located behind the ribs and which is closed by a cover plate **33**. Casings **8** and **9** are arranged end-to-end so that the outlet chamber of **9** communicates with the inlet chamber of **8**, the gasses from the exhaust manifold **34** being introduced into the inlet chamber of casing **9** through extension **34a**. The exhaust gasses enter the series of cells at the right hand side of the casing, pass through the cells into the connecting chamber at the rear and then enter the inlet chamber of casing **8**. They pass successively through the two series of cells and enter exhaust pipe **35**. The exhaust gasses leave the outlet chamber **31**, and the path along which they travel is clearly shown by the arrows in **Fig.6**. As the gasses pass through casings **8** and **9**, their speed is reduced to such a degree that an exhaust box (muffler) or other silencing device is rendered unnecessary.

It will be apparent that when the engine is operating a normal temperature, the liquid fuel introduced into chamber **10** will be vaporised immediately by contact with the hot walls of ribs **12**. The vapour thus produced is divided into two streams, one of which is caused to enter each of the branch passages at one side of the casing and the other is caused to enter each of the branch passages at the opposite side of the casing. The two streams of vapour merge as they pass around the final baffle and enter conduit **17**, but are again divided and heated in a similar manner as they flow through casing **9**. Each of the vapour streams is constantly in contact with the highly heated walls of ribs **12**. This passage of the vapour through the casings causes the vapour to be heated to such a degree that a dry highly-vaporised gas is produced. In this connection, it will be noted that the vaporising chambers are maintained under a vacuum and that vaporisation is effected in the absence of air. Conversion of the liquid into highly expanded vapour is thus ensured. The flow of the exhaust gasses through casings **8** and **9** is in the opposite direction to the flow of the vapour. The vapour is heated in stages and is introduced into chamber **20** at its highest temperature.

The air which is mixed with the fuel vapour, enters pipe **21** after passing through a conventional filter **36**, the amount of air being regulated by valve **37**. The invention also contemplates the heating of the air prior to its entry into mixing chamber **20**. To this end, a jacket 39 is formed around pipe **21**. The jacket has a chamber **40** which communicates with chamber **32** of casing **9** through inlet pipe **41** and with the corresponding chamber of casing **8**

through outlet pipe **42**. A portion of the exhaust gasses is thus caused to pass through chamber **40** to heat the air as it passes through conduit **21** on its way to the mixing chamber. Valve **37** is connected to valve **23** by arms **43** and **43a** and link **44** so that the volume of air admitted to the mixing chamber is increased proportionately as the volume of vapour is increased. As the fuel vapour and air are both heated to a high temperature and are in a highly expanded state when they enter the mixing chamber, they readily unite to provide a uniform mixture, the use of a carburettor or similar device for this purpose being unnecessary.

From the foregoing it will be apparent that the components of the fuel mixture are separately heated prior to their entry into mixing chamber **20**. As the vapour which is produced is dry (containing no droplets of liquid fuel) and highly expanded, complete combustion is ensured. The potential energy represented by the vapour may thus be fully utilised, thereby ensuring better engine performance and a saving in fuel consumption. At the same time, the formation of carbon deposits in the combustion chambers and the production of carbon monoxide and other objectionable exhaust gasses is prevented. The device has the further advantage that, owing to the high temperature to which the fuel is heated prior to its admission into the combustion chambers, various inferior and inexpensive grades of fuel may be used with satisfactory results.

The High MPG Carburettor of Robert Shelton

US Patent 2,982,528 **2nd May 1940** **Inventor: Robert S. Shelton**

VAPOUR FUEL SYSTEM

This patent describes a carburettor design which was able to produce very high mpg figures using the gasoline available in the USA in the 1930s but which is no longer available as the oil industry does not want functional high mpg carburettors to be available to the public.

DESCRIPTION

This invention relates to improvements in vapour fuel systems which are to be used for internal combustion engines.

An object of this invention is to provide a vapour fuel system which will provide a great saving in fuel since approximately eight times the mileage that is obtained by the conventional combustion engine, is provided by the use of this system.

Another object of the invention is to provide a vapour fuel system which is provided with a reservoir to contain liquid fuel which is heated to provide vapour from which the internal combustion engine will operate.

With the above and other objects and advantages in view, the invention consists of the novel details of construction, arrangement and combination of parts more fully described below, claimed and illustrated in the accompanying drawings.

DESCRIPTION OF THE DRAWINGS

Fig.1 is an elevational view of a vapour fuel system embodying the invention.

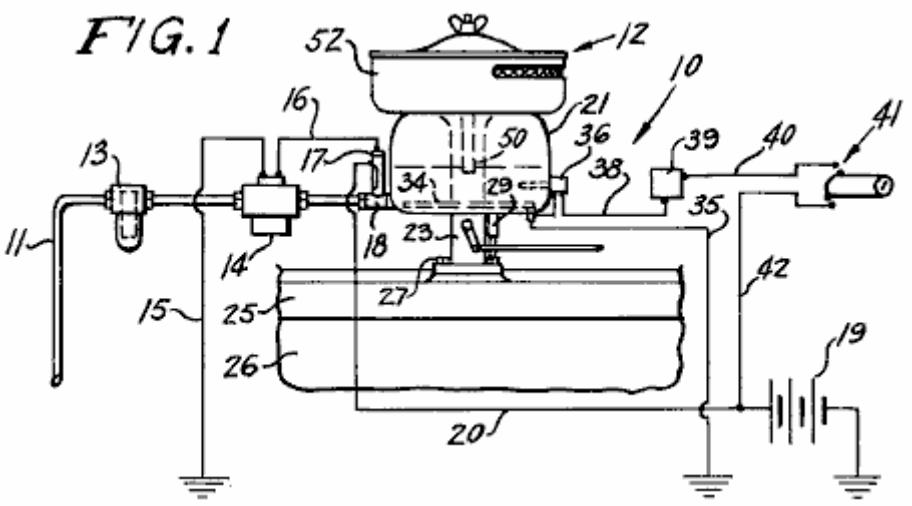


Fig.2 is an enlarged view, partly in section, showing the carburettor forming part of the system shown in Fig.1.

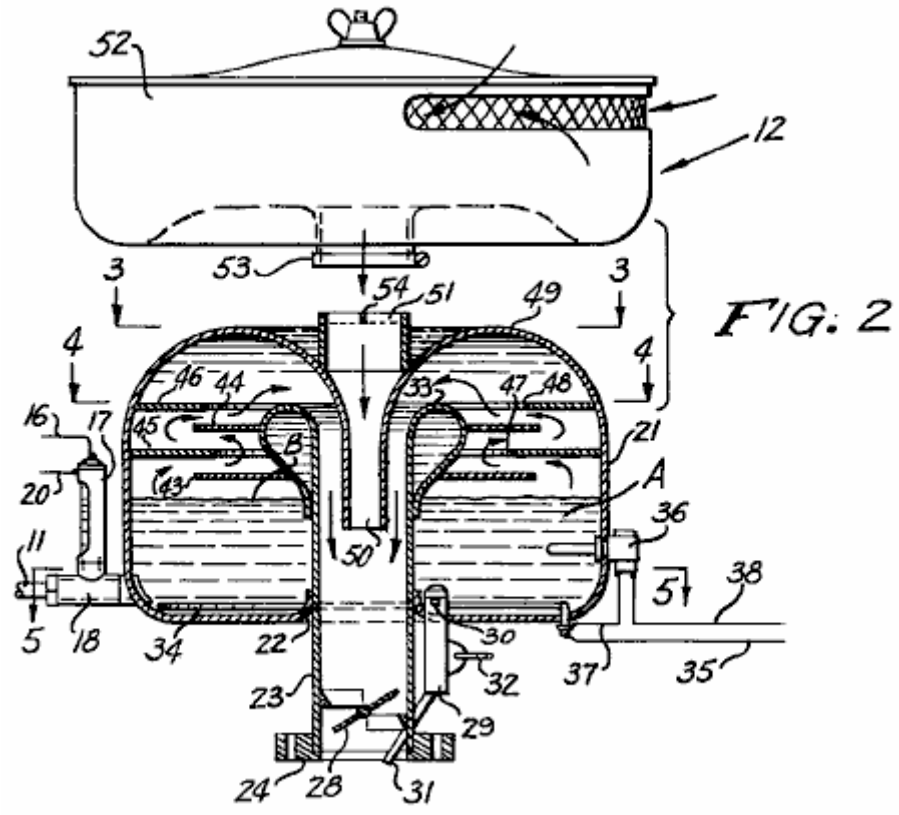


FIG. 2

Fig.3 is a transverse sectional view on line 3--3 of Fig.2

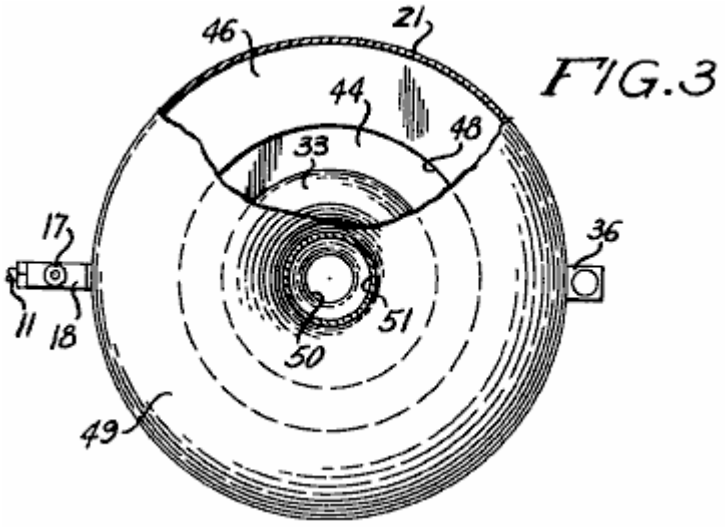


FIG.3

Fig.4 is a transverse sectional view on line 4--4 of Fig.2

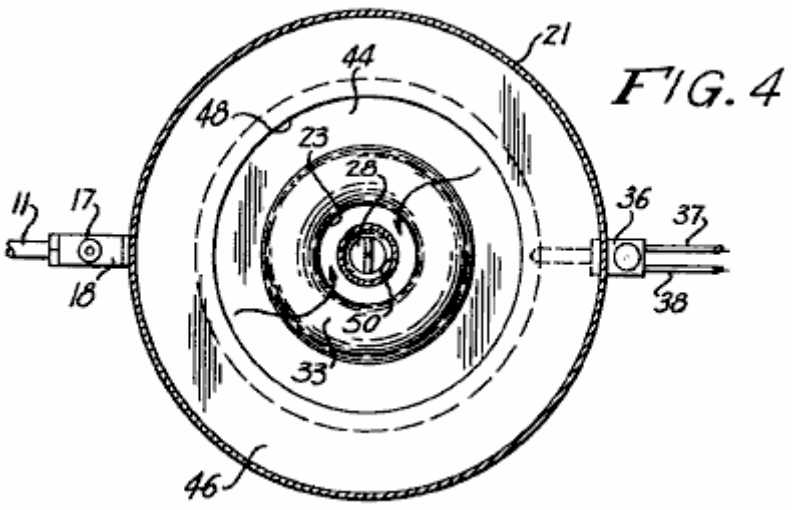
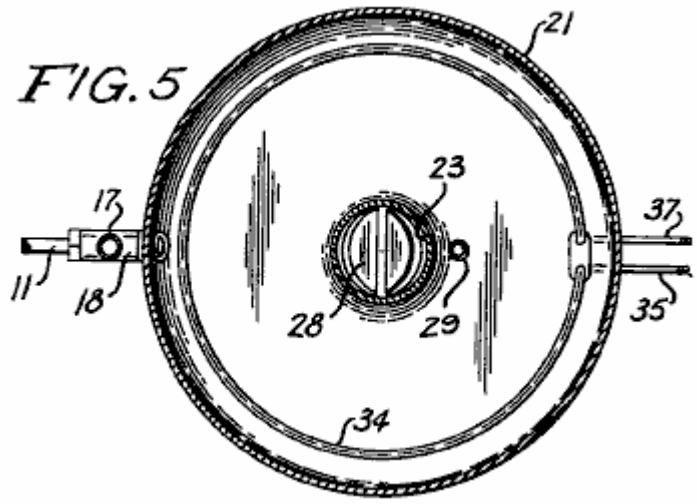


Fig.5 is a transverse sectional view on line 5--5 of Fig.2



The reference numbers used in the drawings always refer to the same item in each of the drawings. The vapour fuel system **10** includes a conduit **11** which is connected to the fuel tank at one end and to a carburettor **12** at the opposite end. In conduit **11** there is a fuel filter **13** and an electric fuel pump **14**. Wire **15** grounds the pump and wire **16** connects the pump to a fuel gauge **18** on which is mounted a switch **17** which is connected to a battery **19** of the engine by wire **20**.

The fuel gauge/switch is of conventional construction and is of the type disclosed in US Patents No. 2,894,093, No. 2,825,895 and No. 2,749,401. The switch is so constructed that a float in the liquid in the gauge, opens a pair of contacts when the liquid rises and this cuts off the electric pump **14**. As the float lowers due to the consumption of the liquid fuel in the body, the float falls, closing the contacts and starting pump **14** which replenishes the liquid fuel in the body.

Carburettor **12** includes a dome-shaped circular bowl or reservoir **21** which is provided with a centrally located flanged opening **22** whereby the reservoir **21** is mounted on a tubular throat **23**. An apratured collar **24** on the lower end of throat **23** is positioned on the intake manifold **25** of an internal combustion engine **26** and fastenings **27** secure the collar to the manifold in a fixed position.

A vapour control butterfly valve **28** is pivotally mounted in the lower end of throat **23** and valve **28** controls the entrance of the vapour into the engine and so controls its speed.

A fuel pump **29**, having an inlet **30**, is mounted in the bottom of the reservoir 21 so that the inlet **30** communicates with the interior of the reservoir. A spurt or feed pipe **31** connected to pump **29** extends into throat **23** so that by means of a linkage **32** which is connected to pump **29** and to a linkage for control valve **28** and the foot throttle of the engine, raw fuel may be forced into throat **23** to start the engine when it is cold.

The upper end of throat **23** is turned over upon itself to provide a bulbous hollow portion **33** within reservoir **21**. An immersion heater **34** is positioned in the bottom of the reservoir and wire **35** grounds the heater. A thermostat **36** is mounted in the wall of the reservoir and extends into it. Wire **37** connects the thermostat to heater **34** and wire **38** connects the thermostat to the thermostat control **39**. Wire **40** connects the control to the ignition switch **41** which in turn is connected to battery **19** via wires **20** and **42**.

A pair of relatively spaced parallel perforated baffle plates **43** and **44**, are connected to the bulbous portion **33** on the upper end of throat **23**, and a second pair of perforated baffle plates **45** and **46** extend inwards from the wall of reservoir **21** parallel to each other and parallel to baffle plates **43** and **44**.

The baffle plates are arranged in staggered relation to each other so that baffle plate **45** is between baffle plates **43** and **44** and baffle plate **46** extends over baffle plate **44**.

Baffle plate **45** has a central opening **47** and baffle plate **46** has a central opening **48** which has a greater diameter than opening **47**. The domed top **49** of reservoir **21**, extends into a tubular air intake **50** which extends downwards into throat **23** and a mounting ring **51** is positioned on the exterior of the domed top, vertically aligned with intake **50**. An air filter **52** is mounted on the mounting ring **51** by a coupling **53** as is the usual procedure, and a spider **54** is mounted in the upper end of mounting ring **51** to break up the air as it enters ring **51** from air filter **52**.

In operation, with carburettor **12** mounted on the internal combustion engine instead of a conventional carburettor, ignition switch **41** is turned on. Current from battery **19** will cause pump **14** to move liquid fuel into reservoir **21** until float switch **18** cuts the pump off when the liquid fuel **A** has reached level **B** in the reservoir. The control **39** is adjusted so that thermostat **36** will operate heater **34** until the liquid fuel has reached a temperature of 105^0 F at which time heater **34** will be cut off. When the liquid fuel has reached the proper temperature, vapour will be available to follow the course indicated by the arrows in **Fig.2**.

The engine is then started and if the foot control is actuated, pump **29** will cause raw liquid fuel to enter the intake manifold **25** until the vapour from the carburettor is drawn into the manifold to cause the engine to operate. As the fuel is consumed, pump **14** will again be operated and heater **34** will be operated by thermostat **36**. Thus, the operation as described will continue as long as the engine is operating and the ignition switch **41** is turned on. Reservoir **21** will hold from 4 to 6 pints (2 to 4 litres) of liquid fuel and since only the vapour from the heated fuel will cause the carburettor **12** to run the engine, the engine will operate for a long time before more fuel is drawn into reservoir **21**.

Baffles **43**, **44**, **45** and **46** are arranged in staggered relation to prevent splashing of the liquid fuel within the carburettor. The level **B** of the fuel in reservoir **21** is maintained constant by switch **18** and with all elements properly sealed, the vapour fuel system **10** will operate the engine efficiently.

Valve **28** controlling the entrance of vapour into intake manifold **25**, controls the speed of the engine in the same manner as the control valve in a conventional carburettor.

There has thus been described a vapour fuel system embodying the invention and it is believed that the structure and operation of it will be apparent to those skilled in the art. It is also to be understood that changes in the minor details of construction, arrangement and combination of parts may be resorted to provided that they fall within the spirit of the invention.

The High MPG Carburettor of Harold Schwartz

US Patent 3,294,381 27th December 1966 Inventor: Harold Schwartz

CARBURETTOR

This patent describes a carburettor design which was able to produce very high mpg figures using the gasoline available in the USA at the time but which is no longer available as the oil industry does not want functional high mpg carburettors to be available to the public.

DESCRIPTION

This invention relates to a carburettor construction. An object of the present invention is to provide a carburettor in which the fuel is treated by the hot exhaust fumes of an engine before being combined with air and being fed into the engine.

Another object of the invention is to provide a carburettor as characterised above, which circulates the fume-laden fuel in a manner to free it of inordinately large globules of fuel, thereby insuring that only finely divided and pre-heated fuel of mist-like consistency is fed to the intake manifold of the engine.

The present carburettor, when used for feeding the six-cylinder engine of a popular car, improved the miles per gallon performance under normal driving conditions using a common grade of fuel, by over 200%. This increased efficiency was achieved from the pre-heating of the fuel and keeping it under low pressure imposed by suction applied to the carburettor for the purpose of maintaining the level of fuel during operation of the engine. This low pressure in the carburettor causes increased vaporisation of the fuel in the carburettor and raises the efficiency of operation.

This invention also has for its objects; to provide a carburettor which is positive in operation, convenient to use, easily installed in its working position, easily removed from the engine, economical to manufacture, of relatively simple design and of general superiority and serviceability.

The invention also comprises novel details of construction and novel combinations and arrangements of parts, which will appear more fully in the course of the following description and which is based on the accompanying drawings. However, the drawings and following description merely describes one embodiment of the present invention, and are only given as an illustration or example.

DESCRIPTION OF THE DRAWINGS

In the drawings, all reference numbers apply to the same parts in each drawing.

Fig.1 is a partly broken plan view of a carburettor constructed in accordance with the present invention, shown with a fuel supply, feeding and return system.

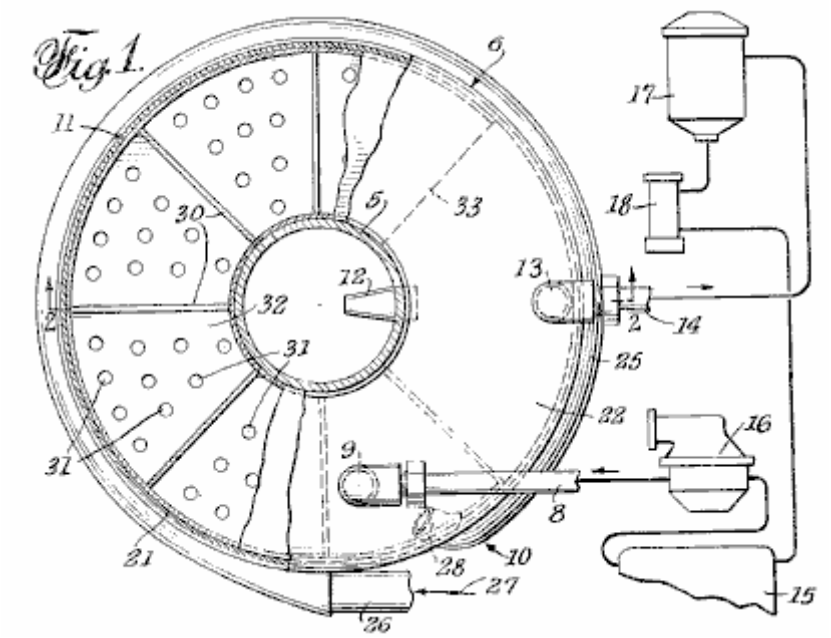


Fig.2 is a vertical sectional view of the carburettor taken on the plane of line 2--2 in Fig.1

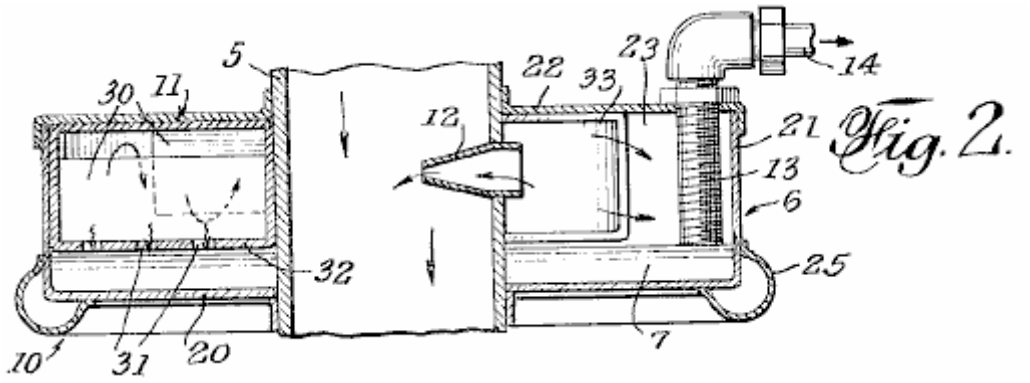
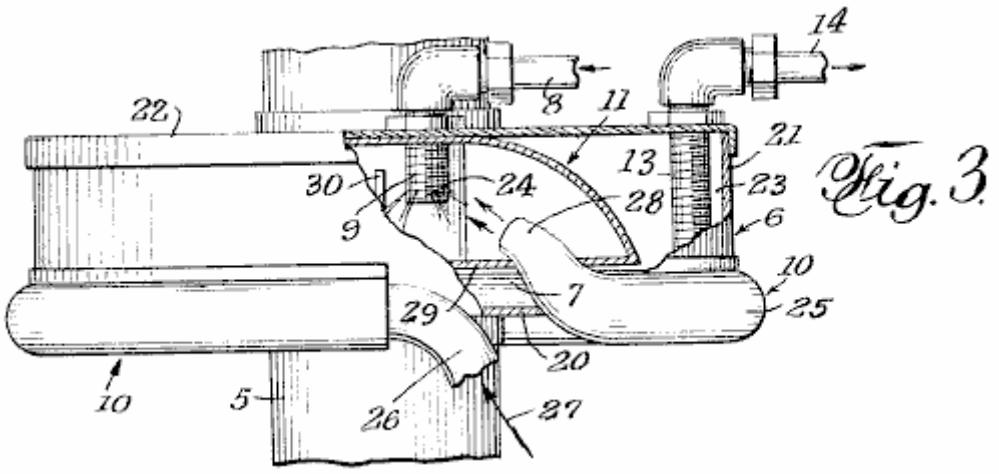


Fig.3 is a partial side elevation and partial sectional view of the carburettor, showing additional structural details



The carburettor is preferably mounted on the usual downdraft air tube **5** which receives a flow of air through the air filter. Tube **5** is provided with a throttle or butterfly valve which controls the flow and incorporates a flow-increasing venturi passage. These common features of the fuel feed to the engine intake manifold are not shown since these features are well known and they are also disclosed in my pending Patent application Serial No.

182,420 now abandoned. The present carburettor embodies improvements over the disclosure of the earlier application.

The present carburettor comprises a housing **6** mounted on air tube **5**, and designed to hold a shallow pool of fuel **7**, a fuel inlet **8** terminating in a spray nozzle **9**, an exhaust gas manifold **10** to conduct heated exhaust gasses for discharge into the spray of fuel coming out of nozzle **9** and for heating the pool of fuel **7** underneath it. Means **11** to scrub the fuel-fumes mixture to eliminate large droplets of fuel from the mixture (the droplets fall into pool **7** underneath), a nozzle tube **12** to receive the scrubbed mixture and to pass the mixture under venturi action into air tube **5** where it is combined with air and made ready for injection into the intake manifold of the engine. Pickup pipe **13** is connected to an outlet **14** for drawing excess fuel from pool **7** during operation of the carburettor.

The system connected to the carburettor is shown in **Fig.1**, and comprises a fuel tank **15**, a generally conventional fuel pump **16** for drawing fuel from the tank and directing it to inlet **8**, a fuel filter **17**, and a pump **18** connected in series between the fuel tank and outlet **14** to place pipe **13** under suction and to draw excess fuel from the carburettor back to tank **15** for re-circulation to inlet **8**.

Carburettor housing **6** may be circular, as shown and quite flat compared to its diameter, so as to have a large flat bottom **20** which, with the cylindrical wall **21**, holds the fuel pool **7**. Cover **22** encloses the top of the housing. The bottom **20** and cover **22** have aligned central openings through which the downdraft tube **5** extends, this pipe forming the interior of the housing, creating an annular inner space **23**.

The fuel inlet **8** is attached to cover **22** by a removable connection. Spray nozzle **9** extends through the cover. While the drawing shows spray-emitting holes **24** arranged to provide a spray around nozzle 7, the nozzle may be formed so that the spray is directional as desired to achieve the most efficient interengagement of the sprayed fuel with the heating gasses supplied by the manifold **10**.

The manifold is shown as a pipe **25** which has and end **26** extending from the conventional heat riser chamber (not shown) of the engine, the arrow **27** indicating exhaust gas flow into pipe **25**. The pipe may encircle the lower portion of the housing **6**, to heat the pool of fuel **7** by transfer of heat through the wall of the housing. The manifold pipe is shown with a discharge end **28** which extends into the housing in an inward and upward direction towards nozzle **9** so that the exhaust gasses flowing in the pipe intermingle with the sprayed fuel and heat it as it leaves the nozzle.

The fuel-scrubbing means **11** is shown as a curved chamber **29** located inside housing **6**, provided with a series of baffle walls **30** which cause the fumes-heated fuel mist to follow a winding path and intercept the heavier droplets of fuel which then run down the faces of the baffle walls, through openings **31** in the bottom wall **32** of scrubbing chamber **29** into the interior space **23** of housing **6** above the level of the fuel pool **7**.

Pickup pipe **13** is also shown as carried by housing cover **22** and may be adjusted so that its lower open end is so spaced from the housing bottom **20** as to regulate the depth of pool **7**, which is preferably below the bottom wall **32** of the scrubbing chamber **29**. Since this pipe is subject to the suction of pump **18** through outlet **14** and filter **17**, the level of pool **7** is maintained by excess fuel being returned to tank **15** by pump **16**.

It will be seen that the surface of pool **7** is subject not only to the venturi action in tube **5**, but also to the suction of pump **18** as it draws excess fuel back to fuel tank **15**. Thus, the surface of the pool is under somewhat less than atmospheric pressure which increases the rate of vaporisation from the pool surface, the resulting vapour combining with the flow from the scrubbing chamber to the downdraft tube **5**..

While this description has illustrated what is now contemplated to be the best mode of carrying out the invention, the construction is, of course, subject to modification without departing from the spirit and scope of the invention. Therefore, it is not desired to restrict the invention to the particular form of construction illustrated and described, but to cover all modifications which may fall within its scope.

The High MPG Carburettor of Oliver Tucker

US Patent 3,653,643 4th April 1972 Inventor: Oliver M. Tucker

CARBURETTOR

This patent describes a carburettor design which was able to produce very high mpg figures using the gasoline available in the USA at the time but which is no longer available as the oil industry does not want functional high mpg carburettors to be available to the public.

ABSTRACT

A carburettor including a housing having a fluid reservoir in the bottom, an air inlet at the top of the housing, a delivery pipe coaxially mounted within the housing and terminating short of the top of the housing, and a porous vaporising filter substantially filling the reservoir. A baffle is concentrically mounted within the housing and extends partially into the vaporising filter in the reservoir to deflect the incoming air through the filter. The level of liquid fuel in the reservoir is kept above the bottom of the baffle, so that air entering the carburettor through the inlet must pass through the liquid fuel and vaporising filter in the reservoir before discharge through the outlet. A secondary air inlet is provided in the top of the housing for controlling the fuel air ratio of the vaporised fuel passing into the delivery pipe.

BACKGROUND OF THE INVENTION

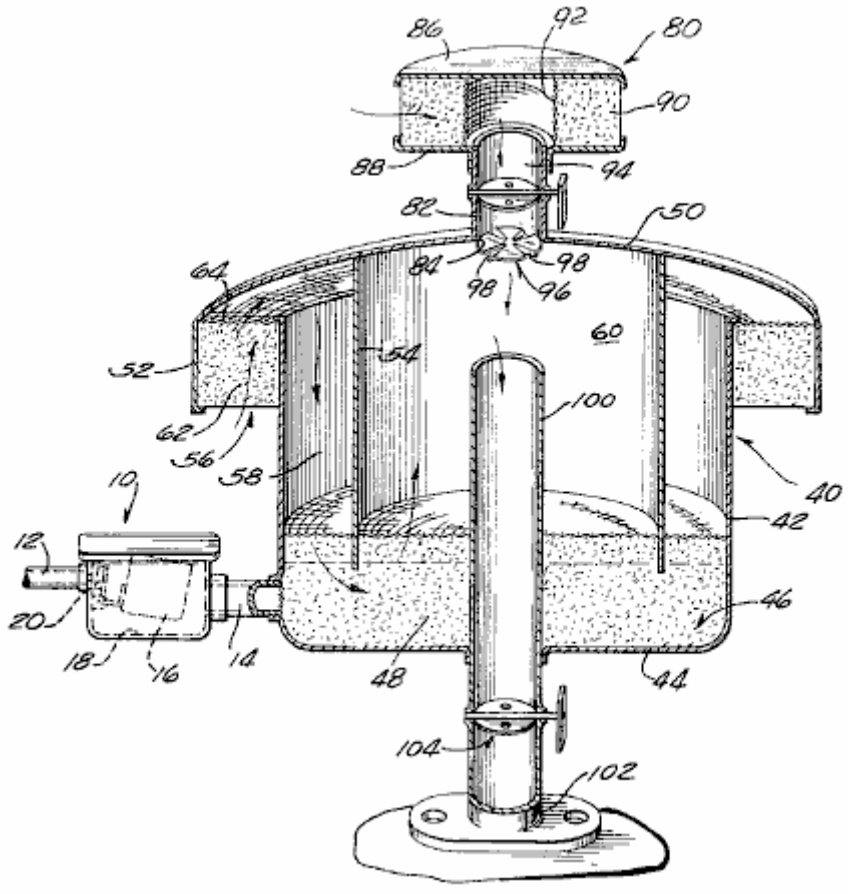
It is generally well known that liquid fuel must be vaporised in order to obtain complete combustion. Incomplete combustion of fuel in internal combustion engines is a major cause of atmospheric pollution. In a typical automotive carburettor, the liquid fuel is atomised and injected into the air stream in a manifold of approximately 3.14 square inches in cross-sectional area. In an eight cylinder 283 cubic inch engine running at approximately 2,400 rpm requires 340,000 cubic inches of air per minute. The air velocity in the intake manifold at this engine speed will be approximately 150 feet per second and it will therefore take approximately 0.07 seconds for a particle of fuel to move from the carburettor to the combustion chamber and the fuel will remain in the combustion chamber for approximately 0.0025 seconds.

It is conceivable that in this short period of time, complete vaporisation of the fuel is not achieved and as a consequence, incomplete combustion occurs, resulting in further air pollution. The liquid fuel particles if not vaporised, can deposit on the cylinder walls and dilute the lubricating oil film there, promoting partial burning of the lubricating oil and adding further to the pollution problem. Destruction of the film of lubricating oil by combustion can also increase mechanical wear of both cylinders and piston rings.

SUMMARY OF THE INVENTION

The carburettor of this invention provides for the complete combustion of liquid fuel in an internal combustion engine, with a corresponding decrease of air pollutant in the exhaust gasses. This is achieved by supplying completely vaporised or dry gas to the combustion chamber. The primary air is initially filtered prior to passing through a vaporising filter which is immersed in liquid fuel drawn from a reservoir in the carburettor. The vaporising filter continuously breaks the primary air up into small bubbles thereby increasing the surface area available for evaporation of the liquid fuel. Secondary air is added to the enriched fuel-air mixture through a secondary air filter prior to admission of the fuel-air mixture into the combustion chambers of the engine. Initial filtration of both the primary and secondary air removes any foreign particles which may be present in the air, and which could cause increased wear within the engine. The carburettor also assures delivery of a clean dry gas to the engine due to the gravity separation of any liquid or dirt particles from the fuel-enriched primary air.

Other objects and advantages will become apparent from the following detailed description when read in conjunction with the accompanying drawing, in which the single figure shows a perspective cross-sectional view of the carburettor of this invention.



DESCRIPTION OF THE INVENTION

The carburettor **40** disclosed here is adapted for use with an internal combustion engine where air is drawn through the carburettor to vaporise the fuel in the carburettor prior to its admission to the engine.

In this regard, the flow of liquid fuel, gas or oil, to the carburettor is controlled by means of a float valve assembly **10** connected to a source of liquid fuel by fuel line **12** and to the carburettor **40** by a connecting tube **14**. The flow of liquid fuel through the float valve assembly **10** is controlled by a float **16**, pivotally mounted within a float chamber **18** and operatively connected to a float valve **20**.

In accordance with the invention, the liquid fuel admitted to the carburettor **40** through tube **14**, is completely evaporated by the primary air for the engine within the carburettor and mixed with secondary air prior to admission into a delivery tube **100** which is connected to the manifold **102** of the engine. More specifically, carburettor **40** includes a cylindrical housing or pan **42**, having a bottom wall **44** which forms a liquid fuel and filter reservoir **46**. A vaporising filter **48** is positioned within reservoir **46** and extends upwards for a distance from the bottom wall **44** of the housing **42**. The vaporising filter **48** is used to continuously break up the primary air into a large number of small bubbles as it passes through the liquid fuel in reservoir **46**. This increases the surface area per volume of air available for evaporation of the liquid fuel, as described in more detail below. This filter **48** is formed of a three-dimensional skeletal material that is washable and is not subject to breakdown under the operating conditions inside the carburettor. A foamed cellular plastic polyurethane filter having approximately 10 to 20 pores per inch has been used successfully in the carburettor.

Housing **42** is closed at the top by a hood or cover **50** which can be secured in place by any appropriate means. The hood has a larger diameter than the diameter of housing **42** and includes a descending flange **52** and a descending baffle **54**. Flange **52** is concentrically arranged and projects outwards beyond the sides of housing **42** to form a primary air inlet **56**. Baffle **54** is concentrically positioned inside housing **42** to create a primary air chamber **58** and a central mixing chamber **60**.

Primary air is drawn into housing **42** through air inlet **56** and is filtered through primary air filter **62** which is removably mounted in the space between flange **52** and the outside of the wall of housing **42** by means of a screen **64**. The primary air filter **62** can be made of the same filtering material as the vaporising filter **48**.

As the primary air enters the primary air chamber **58** it is deflected through the liquid fuel in reservoir **46** by means of the cylindrical baffle **54**. This baffle extends down from hood **50** far enough to penetrate the upper portion of the vaporising filter **48**. The primary air must pass around the bottom of baffle **54** and through both the liquid fuel and the vaporising filter **48** prior to entering the mixing chamber **60**.

The level of the liquid fuel in reservoir **46** is maintained above the bottom edge of baffle **54** by means of the float valve assembly **10**. The operation of the float valve assembly **10** is well known. Float chamber **18** is located at approximately the same level as reservoir **46** and float **16** pivots in response to a drop in the level of the liquid fuel in the float chamber and opens the float valve **20**.

One of the important features of the present invention is the efficiency of evaporation of the liquid fuel by the flow of the large number of bubbles through the reservoir. This is believed to be caused by the continual break up of the bubbles as they pass through the vaporising filter **48**. It is well known that the rate of evaporation caused by a bubble of air passing unmolested through a liquid, is relatively slow due to the surface tension of the bubble. However, if the bubble is continuously broken, the surface tension of the bubble is reduced and a continual evaporating process occurs. This phenomenon is believed to be the cause of the high evaporation rate of the liquid fuel in the carburettor of this invention.

Another feature of the carburettor of this invention is its ability to supply dry gas to the central mixing chamber **60** in housing **42**. Since the flow of primary air in the central mixing chamber **60** is vertically upwards, the force of gravity will prevent any droplets of liquid fuel from rising high enough in the carburettor to enter the delivery tube **100**. The delivery of dry gas to the delivery tube increases the efficiency of combustion and thereby reduces the amount of unburnt gasses or pollutants which are exhausted into the air by the engine.

Means are provided for admitting secondary air into the central mixing chamber **60** to achieve the proper fuel-air ratio required for complete combustion. Such means is in the form of a secondary air filter assembly **80** mounted on an inlet tube **82** provided in opening **84** in hood **50**. The secondary air filter assembly **80** includes an upper plate **86**, a lower plate **88**, and a secondary air filter **90** positioned between plates **86** and **88**. The secondary air filter **90** is prevented from being drawn into inlet tube **82** by means of a cylindrical screen **92** which forms a continuation of tube **82**. The secondary air passes through the outer periphery of the secondary air filter **90**, through screen **92** and into tube **82**. The flow of secondary air through tube **82** is controlled by means of a butterfly valve **94** as is generally understood in the art.

Complete mixing of the dry gas-enriched primary air with the incoming secondary air within housing **42**, is achieved by means of deflector **96** positioned at the end of tube **82**. Deflector **96** includes a number of vanes **98** which are twisted to provide an outwardly-deflected circular air flow into the central mixing chamber **60** and thereby creating an increase in the turbulence of the secondary air as it combines with the fuel-enriched primary air. The deflector prevents cavitation from occurring at the upper end of the outlet tube **100**.

The flow of fuel-air mixture to the engine is controlled by means of a throttle valve **104** provided in the outlet or delivery tube **100**. The operation of the throttle valve **104** and butterfly valve **94** are both controlled in a conventional manner.

THE OPERATION OF THE CARBURETTOR

Primary air is drawn into housing **42** through primary air inlet **56** and passes upwards through primary air filter **62** where substantially all foreign particles are removed from the primary air. The filtered primary air then flows downwards through primary air chamber **58**, under baffle **54**, through fuel filter reservoir **46**, and upwards into central mixing chamber **60**. All of the primary air passes through the vaporising filter **48** provided in reservoir **46**. The vaporising filter **48** continuously breaks the primary air stream into thousands of small bubbles, reducing surface tension and increasing the air surface available for evaporation of the liquid fuel. Since the outer surface of each bubble is being constantly broken up by the vaporising filter **48** and is in constant contact with the liquid fuel as the bubble passes through the vaporising filter **48**, there is a greater opportunity for evaporation of the fuel prior to entering the central mixing chamber **60**. The vertical upward flow of the fuel-enriched primary air in the central mixing chamber, ensures that no liquid fuel droplets will be carried into the delivery tube **100**.

The fuel-enriched primary air is thoroughly mixed with the secondary air entering through tube **82** by means of the deflector system **96** which increases the turbulence of the primary and secondary air within the central mixing chamber and prevents cavitation from occurring in delivery tube **100**. The completely mixed fuel-enriched primary air and the secondary air then pass through delivery tube **100** into the inlet manifold of the engine.

The High MPG Carburettor of Thomas Ogle

US Patent 4,177,779 11th December 1979 Inventor: Thomas H. Ogle

FUEL ECONOMY SYSTEM FOR AN INTERNAL COMBUSTION ENGINE

This patent describes a carburettor design which was able to produce very high mpg figures using the gasoline available in the USA at the time but which is no longer available as the oil industry does not want functional high mpg carburettors to be available to the public.

ABSTRACT

A fuel economy system for an internal combustion engine which, when installed in a motor vehicle, overcomes the need for a conventional carburettor, fuel pump and fuel tank. The system operates by using the engine vacuum to draw fuel vapours from a vapour tank through a vapour conduit to a vapour equaliser which is positioned directly over the intake manifold of the engine. The vapour tank is constructed of heavy duty steel, or the like, to withstand the large vacuum pressure and includes an air inlet valve coupled for control to the accelerator pedal. The vapour equaliser ensures distribution of the correct mixture of air and vapour to the cylinders of the engine for combustion, and also includes its own air inlet valve coupled for control to the accelerator pedal. The system utilises vapour-retarding filters in the vapour conduit, vapour tank and vapour equaliser to deliver the correct vapour/air mixture for proper operation. The vapour tank and fuel contained in it, are heated by running the engine coolant through a conduit within the tank. Due to the extremely lean fuel mixtures used by the present invention, gas mileage in excess of one hundred miles per gallon may be achieved.

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention is related to internal combustion engines and, more particularly, is directed towards a fuel economy system for an internal combustion engine which, when applied to a motor vehicle, overcomes the need for conventional carburettors, fuel pumps and fuel tanks, and enables vastly improved fuel consumption to be achieved.

2. Description of the Prior Art

The prior art evidences many different approaches to the problem of increasing the efficiency of an internal combustion engine. Due to the rising price of fuel, and the popularity of motor vehicles as a mode of transportation, much of the effort in this area is generally directed towards improving fuel consumption for motor vehicles. Along with increased mileage, much work has been done with a view towards reducing pollutant emissions from motor vehicles.

I am aware of the following United States patents which are generally directed towards systems for improving the efficiency and/or reducing the pollutant emissions of internal combustion engines:

Chapin	1,530,882
Crabtree et al	2,312,151
Hietrich et al	3,001,519
Hall	3,191,587
Wentworth	3,221,724
Walker	3,395,681
Holzappfel	3,633,533
Dwyre	3,713,429
Herpin	3,716,040
Gorman, Jr.	3,728,092
Alm et al	3,749,376
Hollis, Jr.	3,752,134
Buckton et al	3,759,234
Kihn	3,817,233
Shih	3,851,633

Burden, Sr.	3,854,463
Woolridge	3,874,353
Mondt	3,888,223
Brown	3,907,946
Lee, Jr.	3,911,881
Rose et al	3,931,801
Reimuller	3,945,352
Harpman	3,968,775
Naylor	4,003,356
Fortino	4,011,847
Leshner et al	4,015,569
Sommerville	4,015,570

The Chapin U.S. Pat. No. 1,530,882 discloses a fuel tank surrounded by a water jacket, the latter of which is included in a circulation system with the radiator of the automobile. The heated water in the circulation system causes the fuel in the fuel tank to readily vaporise. Suction from the inlet manifold causes air to be drawn into the tank to bubble air through the fuel to help form the desired vapour which is then drawn to the manifold for combustion.

The Buckton et al U.S. Pat. No. 3,759,234 advances a fuel system which provides supplementary vapours for an internal combustion engine by means of a canister that contains a bed of charcoal granules. The Wentworth and Hietrich et al U.S. Pat. Nos. 3,221,724 and 3,001,519 also teach vapour recovery systems which utilise filters of charcoal granules or the like.

The Dwyre U.S. Pat. No. 3,713,429 uses, in addition to the normal fuel tank and carburettor, an auxiliary tank having a chamber at the bottom which is designed to receive coolant from the engine cooling system for producing fuel vapours, while the Walker U.S. Pat. No. 3,395,681 discloses a fuel evaporator system which includes a fuel tank intended to replace the normal fuel tank, and which includes a fresh air conduit for drawing air into the tank.

The Fortino U.S. Pat. No. 4,011,847 teaches a fuel supply system wherein the fuel is vaporised primarily by atmospheric air which is released below the level of the fuel, while the Crabtree et al U.S. Pat. No. 2,312,151 teaches a vaporisation system which includes a gas and air inlet port located in a vaporising chamber and which includes a set of baffles for effecting a mixture of the air and vapour within the tank. The Mondt U.S. Pat. No. 3,888,223 also discloses an evaporative control canister for improving cold start operation and emissions, while Sommerville U.S. Pat. No. 4,015,570 teaches a liquid-fuel vaporiser which is intended to replace the conventional fuel pump and carburettor that is designed to mechanically change liquid fuel to a vapour state.

While the foregoing patents evidence a proliferation of attempts to increase the efficiency and/or reduce pollutant emissions from internal combustion engines, no practical system has yet found its way to the marketplace.

OBJECTS AND SUMMARY OF THE INVENTION

It is therefore a primary object of the present invention to provide a new and improved fuel economy system for an internal combustion engine which greatly improves the efficiency of the engine.

Another object of the present invention is to provide a unique fuel economy system for an internal combustion engine which provides a practical, operative and readily realisable means for dramatically increasing the gas mileage of conventional motor vehicles.

A further object of the present invention is to provide an improved fuel economy system for internal combustion engines which also reduces the pollutant emissions.

The foregoing and other objects are attained in accordance with one aspect of the present invention through the provision of a fuel vapour system for an internal combustion engine having an intake manifold, which comprises a tank for containing fuel vapour, a vapour equaliser mounted on and in fluid communication with the intake manifold of the engine, and a vapour conduit which connect the tank to the vapour equaliser for delivering fuel vapour from the former to the latter. The vapour equaliser includes a first valve connected to it for controlling the admission of air to the vapour equaliser, while the tank has a second valve connected to it for controlling the admission of air to the tank. A throttle controls the first and second valves so that the opening of the first valve preceeds and exceeds the opening of the second valve during operation.

In accordance with other aspects of the present invention, a filter is positioned in the vapour conduit to retard the flow of fuel vapour from the tank to the vapour equaliser. In a preferred form, the filter comprises carbon particles and may include a sponge-like collection of, for example, neoprene fibres. In a preferred embodiment, the filter comprises a substantially tubular housing positioned in series in the vapour conduit, the housing containing a central portion comprising a mixture of carbon and neoprene, and end portions comprising carbon, positioned on each side of the central portion.

In accordance with another aspect of the present invention, a second filter is positioned in the vapour equaliser for again retarding the flow of the fuel vapour to the engine intake manifold. The second filter is positioned downstream of the first valve and in a preferred form, includes carbon particles mounted in a pair of recesses formed in a porous support member. The porous support member, which may comprise neoprene, includes a first recessed portion positioned opposite a vapour inlet port in the vapour equaliser to which the vapour conduit is connected, while a second recessed portion is positioned opposite the intake manifold of the engine.

In accordance with still other aspects of the present invention, a third filter is positioned in the tank for controlling the flow of fuel vapour into the vapour conduit in proportion to the degree of vacuum in the tank. The filter more particularly comprises a mechanism for reducing the amount of fuel vapour delivered to the vapour conduit when the engine is idling and when the engine has attained a steady speed. The throttle acts to close the second valve when the engine is idling and when the engine has attained a steady speed, to thereby increase the vacuum pressure in the tank. In a preferred form, the third filter comprises a frame pivotally mounted within the tank and movable between first and second operating positions. The first operating position corresponds to an open condition of the second valve, while the second operating position corresponds to a closed condition of the second valve. The tank includes a vapour outlet port to which one end of the vapour conduit is connected, such that the second operating position of the frame places the third filter in communication with the vapour outlet port.

More particularly, the third filter in a preferred form includes carbon particles sandwiched between two layers of a sponge-like filter material, which may comprise neoprene, and screens for supporting the layered composition within the pivotable frame. A conduit is positioned on the third filter for placing it in direct fluid communication with the vapour outlet port when the frame is in its second operating position.

In accordance with yet other aspects of the present invention, a conduit is connected between the valve cover of the engine and the vapour equaliser for directing the oil blow-by to the vapour equaliser in order to minimise valve clatter. The tank also preferably includes a copper conduit positioned in the bottom of it, which is connected in series with the cooling system of the motor vehicle, for heating the tank and generating more vapour. A beneficial by-product of the circulating system reduces the engine operating temperature to further improve operating efficiency.

BRIEF DESCRIPTION OF THE DRAWINGS

Various objects, features and attendant advantages of the present invention will be more fully appreciated as the same become better understood from the following detailed description of the present invention when considered in connection with the accompanying drawings, in which:

Fig.1 is a perspective view illustrating the various components which together comprise a preferred embodiment of the present invention as installed in a motor vehicle;

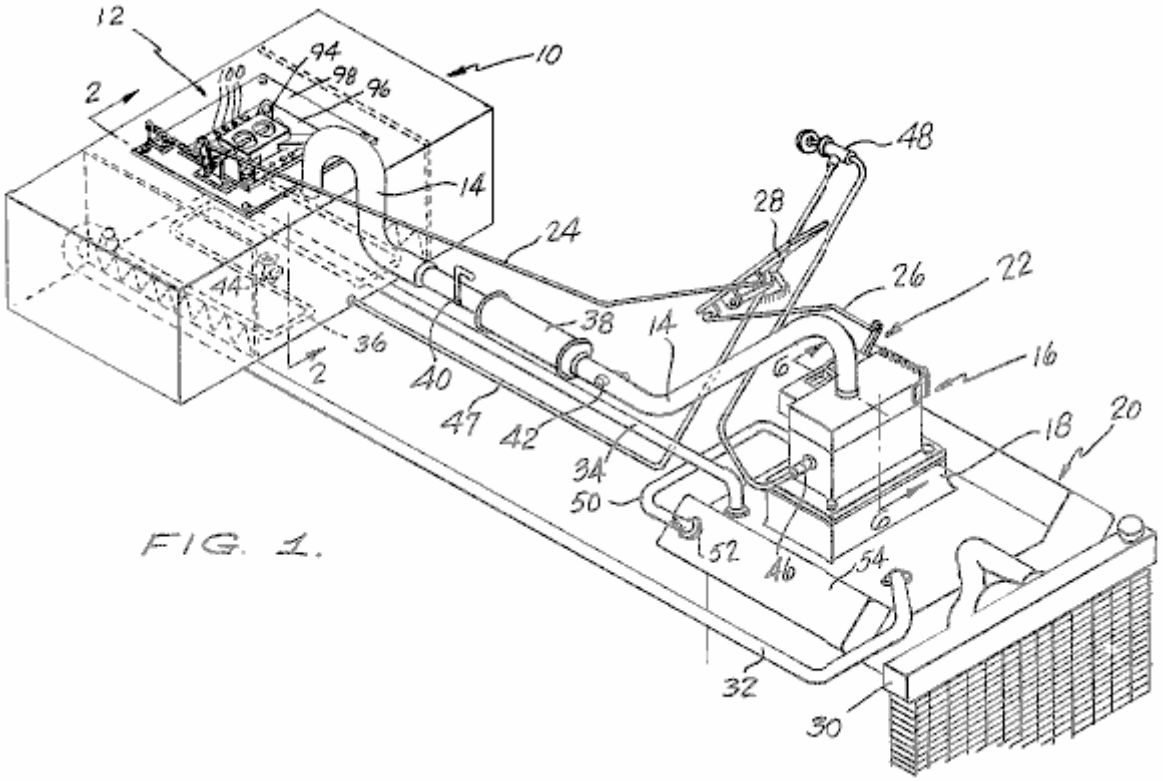


FIG. 1.

Fig.2 is a cross-sectional view of one of the components of the preferred embodiment illustrated in Fig.1 taken along line 2--2

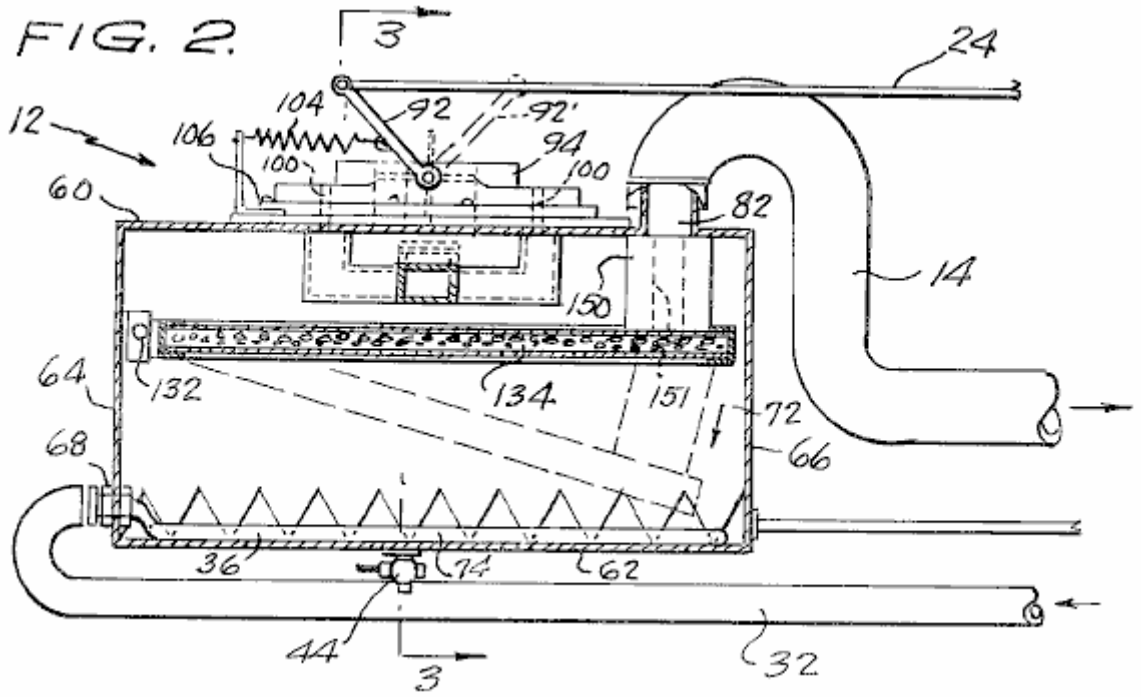


Fig.3 is a sectional view of the vapour tank illustrated in Fig.2 taken along line 3--3

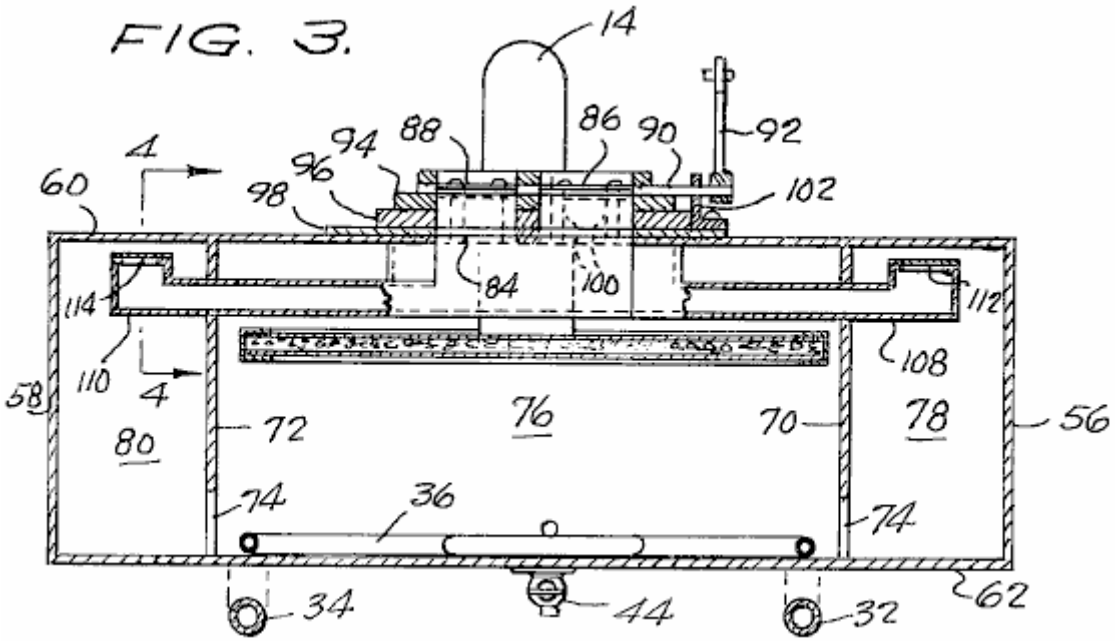


Fig.4 is an enlarged sectional view illustrating in greater detail one component of the vapour tank shown in Fig.3 taken along line 4--4

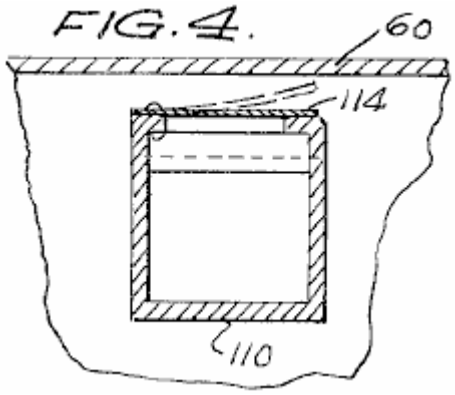


Fig.5 is a perspective, partially sectional view illustrating a filter component of the vapour tank illustrated in Fig.2

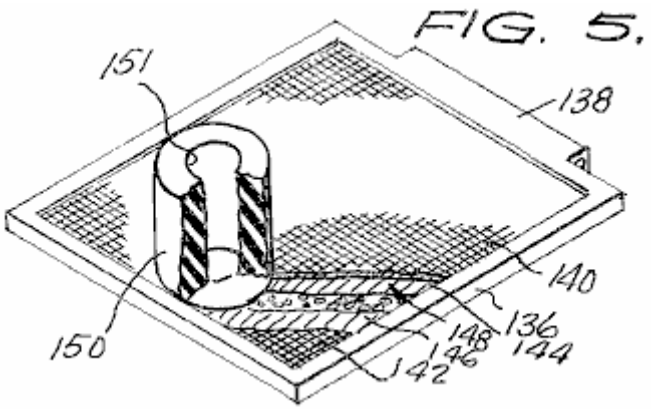


Fig.6 is a cross-sectional view of another component of the preferred embodiment of the present invention illustrated in Fig.1 taken along line 6--6

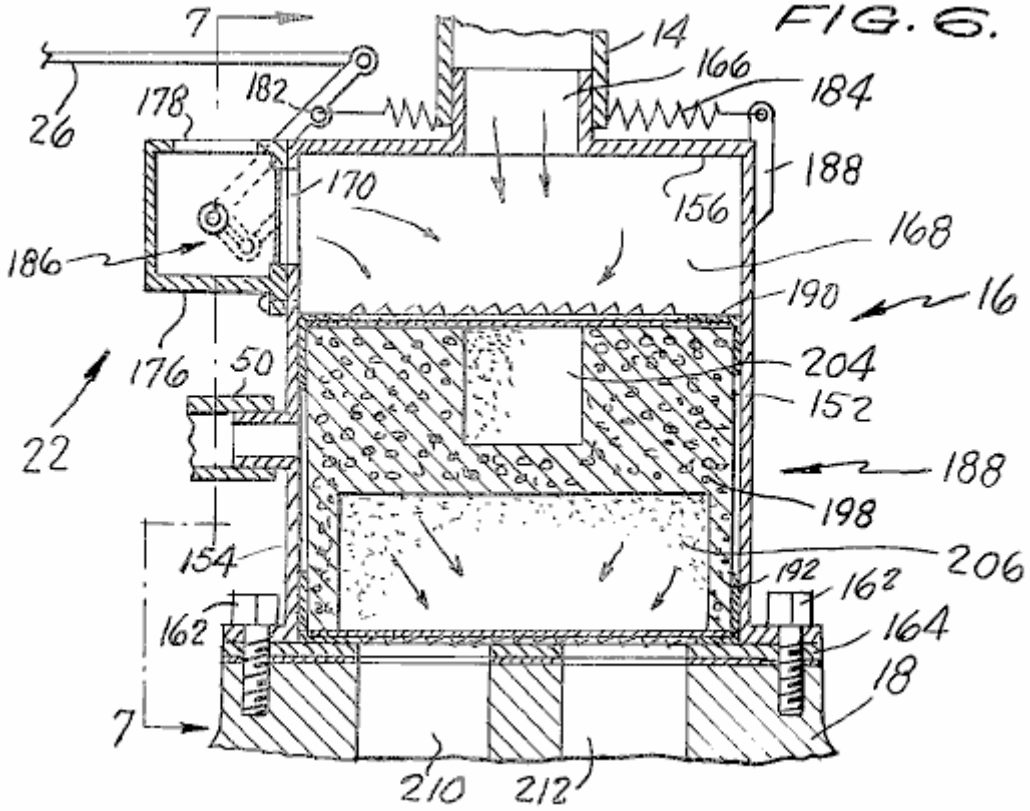


Fig.7 is a partial side, partial sectional view of the vapour equaliser illustrated in Fig.6 taken along line 7--7

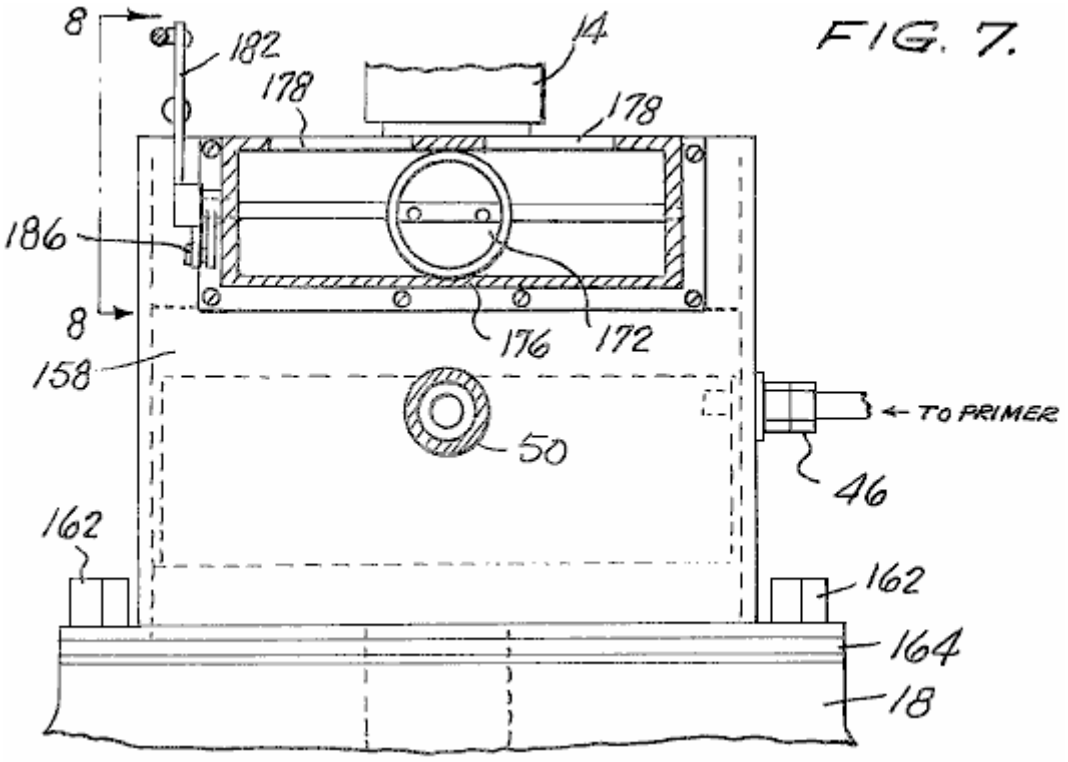


Fig.8 is a side view illustrating the throttle linkage of the vapour equaliser shown in Fig.7 taken along line 8--8

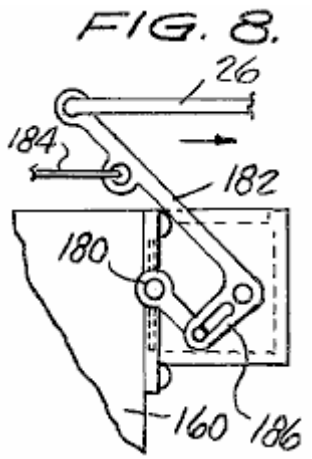


Fig.9 is a longitudinal sectional view of another filter component of the preferred embodiment illustrated in Fig.1

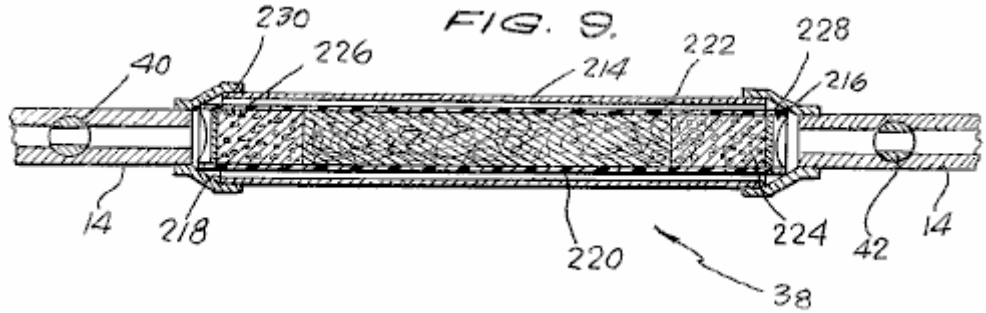


Fig.10 is a view of another component of the present invention

FIG.10.

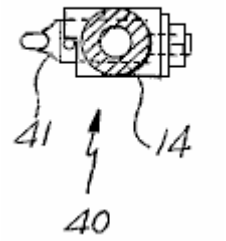
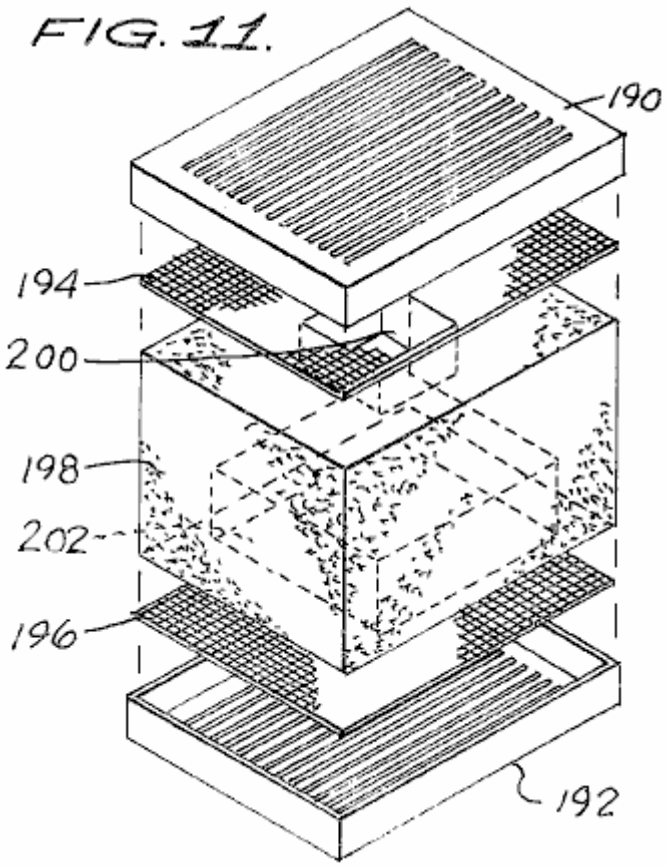


Fig.11 is an exploded, perspective view which illustrates the main components of the filter portion of the vapour equaliser of the present invention.

FIG. 11.



DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

Referring now to the drawings, where parts are numbered the same in each drawing, and more particularly to **Fig.1** which illustrates a preferred embodiment of the present invention as installed in a motor vehicle.

The preferred embodiment includes as its main components a fuel vapour tank **10** in which the fuel vapour is stored and generated for subsequent delivery to the internal combustion engine **20**. On the top of fuel vapour tank **10** is mounted an air inlet control valve **12** whose structure and operation will be described in greater detail below.

The internal combustion engine **20** includes a standard intake manifold **18**. Mounted upon the intake manifold **18** is a vapour equaliser chamber **16**. Connected between the fuel vapour tank **10** and the vapour equaliser chamber **16** is a vapour conduit or hose **14** for conducting the vapours from within tank **10** to the chamber **16**.

Reference numeral **22** indicates generally an air inlet control valve which is mounted on the vapour equaliser chamber **16**. Thus, the system is provided with two separate air inlet control valves **12** and **22** which are respectively coupled via cables **24** and **26** to the throttle control for the motor vehicle which may take the form of a standard accelerator pedal **28**. The air inlet control valves **12** and **22** are synchronised in such a fashion that the opening of the air inlet control valve **22** of the vapour equaliser **16** always precedes and exceeds the opening of the air inlet control valve **12** of the fuel vapour tank **10**, for reasons which will become more clear later.

The cooling system of the vehicle conventionally includes a radiator **30** for storing liquid coolant which is circulated through the engine **20** in the well-known fashion. A pair of hoses **32** and **34** are preferably coupled into the normal heater lines from the engine **20** so as to direct heated liquid coolant from the engine **20** to a warming coil **36**, preferably constructed of copper, which is positioned within vapour tank **10**. I have found that the water circulation system consisting of hoses **32**, **34** and **36** serves three distinct functions. Firstly, it prevents the vapour tank from reaching the cold temperatures to which it would otherwise be subjected as a result of high vacuum pressure and air flow through it. Secondly, the heated coolant serves to enhance vaporisation of the fuel stored within tank **10** by raising its temperature. Thirdly, the liquid coolant, after leaving tank **10** via conduit **34**, has been cooled to the point where engine **20** may then be run at substantially lower operating temperatures to further increase efficiency and prolong the life of the engine.

Included in series with vapour conduit **14** is a filter unit **38** which is designed to retard the flow of fuel vapour from the tank **10** to the vapour equaliser **16**. The precise structure of the filter unit **38** will be described in greater detail below. A thrust adjustment valve **40** is positioned upstream of the filter unit **38** in conduit **14** and acts as a fine adjustment for the idling speed of the vehicle. Positioned on the other side of filter unit **38** in conduit **14** is a safety shut-off valve **42** which comprises a one-way valve. Starting the engine **20** will open the valve **42** to permit the engine vacuum pressure to be transmitted to tank **10**, but, for example, a backfire will close the valve to prevent a possible explosion. The tank **10** may also be provided with a drain **44** positioned at the bottom of the tank.

Positioned on the side of the vapour equaliser chamber **16** is a primer connection **46** which may be controlled by a dash mounted primer control knob **48** connected to tank **10** via conduit **47**. A conduit **50** extends from the oil breather cap opening **52** in a valve cover **54** of the engine **20** to the vapour equaliser **16** to feed the oil blow-by to the engine as a means for eliminating valve clatter. This is believed necessary due to the extreme lean mixture of fuel vapour and air fed to the combustion cylinders of the engine **20** in accordance with the present invention.

Referring now to **Fig.2** and **Fig.3**, the fuel vapour tank **10** of the present invention is illustrated in greater detail in orthogonal sectional views and is seen to include a pair of side walls **56** and **58** which are preferably comprised of heavy duty steel plate (e.g. 1/2" thick) in order to withstand the high vacuum pressures developed inside it. Tank **10** further comprises top wall **60** and bottom wall **62**, and front and rear walls **64** and **66**, respectively.

In the front wall **64** of tank **10** is positioned a coupling **68** for mating the heater hose **32** with the internal copper conduit **36**. Tank **10** is also provided with a pair of vertically oriented planar support plates **70** and **72** which are positioned somewhat inside the side walls **56** and **58** and are substantially parallel to them. Support plates **70** and **72** lend structural integrity to the tank **10** and are also provided with a plurality of openings **74** (**Fig.2**) at the bottom of them to permit fluid communication through it. The bottom of tank **10** is generally filled with from one to five gallons of fuel, and the walls of tank **10** along with plates **70** and **72** define three tank chambers **76**, **78** and **80** which are, by virtue of openings **74**, in fluid communication with one another.

In the top wall **60** of tank **10** is formed an opening **82** for placing one end of vapour conduit **14** in fluid communication with the interior chamber **76** of tank **10**. A second opening **84** is positioned in the top wall **60** of tank **10** over which the air inlet control valve **12** is positioned. The valve assembly **12** comprises a pair of conventional butterfly valves **86** and **88** which are coupled via a control rod **90** to a control arm **92**. Control arm **92** is, in turn, pivoted under the control of a cable **24** and is movable between a solid line position indicated in **Fig.2** by reference numeral **92** and a dotted line position indicated in **Fig.2** by reference numeral **92'**.

Rod **90** and valves **86** and **88** are journaled in a housing **94** having a base plate **96** which is mounted on a cover **98**. As seen in **Fig.1**, the base plate **96** includes several small air intake ports or apertures **100** formed on both sides of the butterfly valves **86** and **88**, which are utilised for a purpose to become more clear later on.

Rod **90** is also journaled in a flange **102** which is mounted to cover **98**, while a return spring **104** for control arm **92** is journaled to cover **98** via flange **106**.

Extending through the baffle and support plates 70 and **72** from the side chambers **78** and **80** of tank **10** to be in fluid communication with apertures **100** are a pair of air conduits **108** and **110** each having a reed valve **112** and **114** positioned at the ends, for controlling air and vapour flow through it. The reed valves **112** and **114** co-operage with the small apertures **100** formed in the base plate **96** to provide the proper amount of air into the tank **10** while the engine is idling and the butterfly valves **86** and **88** are closed.

Mounted to the front wall **64** of tank **10** is a pivot support member **132** for pivotally receiving a filter element which is indicated generally by reference numeral **134** and is illustrated in a perspective, partially cut away view in **Fig.5**. The unique, pivotable filter element **134** comprises a frame member **136** having a pin-receiving stub **138** extending along one side member of it. The actual filter material contained within the frame **136** comprises a layer of carbon particles **148** which is sandwiched between a pair of layers of sponge-like filter material which

may, for example, be made of neoprene. The neoprene layers **144** and **146** and carbon particles **148** are maintained in place by top and bottom screens **140** and **142** which extend within, and are secured by, frame member **136**. ,A thick-walled rubber hose **150** having a central annulus **151** is secured to the top of screen **140** so as to mate with opening **82** of top wall **60** (see **Fig.2**) when the filter assembly **134** is in its solid line operative position illustrated in **Fig.2**. In the latter position, it may be appreciated that the vapour conduit **14** draws vapour fumes directly from the filter element **134**, rather than from the interior portion **76** of tank **10**. In contradistinction, when the filter element **134** is in its alternate operative position, indicated by dotted lines in **Fig.2**, the vapour conduit **14** draws fumes mainly from the interior portions **76**, **78** and **80** of tank **10**.

Fig.4 is an enlarged view of one of the reed valve assemblies **114** which illustrates the manner in which the valve opens and closes in response to the particular vacuum pressure created within the tank **10**. Valves **112** and **114** are designed to admit just enough air to the tank **10** from the apertures **100** at engine idle to prevent the engine from stalling.

Referring now to **Fig.6**, **Fig.7** and **Fig.8**, the vapour equaliser chamber **16** of the present invention is seen to include front and rear walls **152** and **154**, respectively, a top wall **156**, a side wall **158**, and another side wall **160**. The vapour equaliser chamber **16** is secured to the manifold **18** as by a plurality of bolts **162** under which may be positioned a conventional gasket **164**.

In the top wall **156** of the vapour equaliser **16** is formed an opening **166** for communicating the outlet end of vapour conduit **14** with a mixing and equalising chamber **168**. Adjacent to the mixing and equalising chamber **168** in wall **154** is formed another opening **170** which communicates with the outside air via opening **178** formed in the upper portion of housing **176**. The amount of air admitted through openings **178** and **170** is controlled by a conventional butterfly valve **172**. Butterfly valve **172** is rotated by a control rod **180** which, in turn, is coupled to a control arm **182**. Cable **26** is connected to the end of control arm **182** furthest from the centreline and acts against the return bias of spring **184**, the latter of which is journaled to side plate **152** of vapour equaliser **16** via an upstanding flange **188**. Reference numeral **186** indicates generally a butterfly valve operating linkage, as illustrated more clearly in **Fig.8**, and which is of conventional design as may be appreciated by a person skilled in the art.

Positioned below mixing and equalising chamber **168** is a filter unit which is indicated generally by reference numeral **188**. The filter unit **188**, which is illustrated in an exploded view in **Fig.11**, comprises a top plastic fluted cover **190** and a bottom plastic fluted cover **192**. Positioned adjacent to the top and bottom covers **190** and **192** is a pair of screen mesh elements **194** and **196**, respectively. Positioned between the screen mesh elements **194** and **196** is a support member **198** which is preferably formed of a sponge-like filter material, such as, for example, neoprene. The support member **199** has formed on its upper and lower surfaces, a pair of receptacles **200** and **202**, whose diameters are sized similarly to the opening **166** in top plate **156** and the openings formed in the intake manifold **18** which are respectively indicated by reference numerals **210** and **212** in **Fig.6**.

Positioned in receptacles **200** and **202** are carbon particles **204** and **206**, respectively, for vapour retardation and control purposes.

Referring now to **Fig.9**, the filter unit **38** mounted in vapour conduit **14** is illustrated in a longitudinal sectional view and is seen to comprise an outer flexible cylindrical hose **214** which is adapted to connect with hose **14** at both ends by a pair of adapter elements **216** and **218**. Contained within the outer flexible hose **214** is a cylindrical container **220**, preferably of plastic, which houses, in its centre, a mixture of carbon and neoprene filter fibres **222**. At both ends of the mixture **222** are deposited carbon particles **224** and **226**, while the entire filtering unit is held within the container **220** by end screens **228** and **230** which permit passage of vapours through it while holding the carbon particles **224** and **226** in place.

Fig.10 illustrates one form of the thrust adjustment valve **40** which is placed within line **14**. This valve simply controls the amount of fluid which can pass through conduit **14** via a rotating valve member **41**.

In operation, the thrust adjustment valve **40** is initially adjusted to achieve as smooth an idle as possible for the particular motor vehicle in which the system is installed. The emergency shut-off valve **42**, which is closed when the engine is off, generally traps enough vapour between it and the vapour equaliser **16** to start the engine **20**. Initially, the rear intake valves **12** on the tank **10** are fully closed, while the air intake valves **22** on the equaliser **16** are open to admit a charge of air to the vapour equaliser prior to the vapour from the tank, thus forcing the pre-existing vapour in the vapour equaliser into the manifold. The small apertures **100** formed in base plate **96** on tank **10** admit just enough air to actuate the reed valves to permit sufficient vapour and air to be drawn through vapour conduit **14** and equaliser **16** to the engine **20** to provide smooth idling. The front air valves **22** are always set ahead of the rear air valves **12** and the linkages **24** and **26** are coupled to throttle pedal **28** such that the degree of opening of front valves **22** always exceeds the degree of opening of the rear valves **12**.

Upon initial starting of the engine **20**, due to the closed condition of rear valves **12**, a high vacuum pressure is created within tank **10** which causes the filter assembly **134** positioned in tank **10** to rise to its operative position indicated by solid outline in **Fig.2**. In this manner, a relatively small amount of vapour will be drawn directly from filter **134** through vapour conduit **14** to the engine to permit the latter to run on an extremely lean mixture.

Upon initial acceleration, the front air intake valve **22** will open further, while the rear butterfly assembly **12** will begin to open. The latter action will reduce the vacuum pressure within tank **10** whereby the filter assembly **134** will be lowered to its alternate operating position illustrated in dotted outline in **Fig.2**. In this position, the lower end of the filter assembly **134** may actually rest in the liquid fuel contained within the tank **10**. Accordingly, upon acceleration, the filter assembly **134** is moved out of direct fluid communication with the opening **82** such that the vapour conduit **14** then draws fuel vapour and air from the entire tank **10** to provide a richer combustion mixture to the engine, which is necessary during acceleration.

When the motor vehicle attains a steady speed, and the operator eases off the accelerator pedal **28**, the rear butterfly valve assembly **12** closes, but the front air intake **22** remains open to a certain degree. The closing of the rear air intake **12** increases the vacuum pressure within tank **10** to the point where the filter assembly **134** is drawn up to its initial operating position. As illustrated, in this position, the opening **82** is in substantial alignment with the aperture **151** of hose **150** to place the filter unit **134** in direct fluid communication with the vapour conduit **14**, thereby lessening the amount of vapour and air mixture fed to the engine. Any vapour fed through conduit **14** while the filter **134** is at this position is believed to be drawn directly off the filter unit itself.

I have been able to obtain extremely high mpg figures with the system of the present invention installed on a V-8 engine of a conventional 1971 American-made car. In fact, mileage rates in excess of one hundred miles per US gallon have been achieved with the present invention. The present invention eliminates the need for conventional fuel pumps, carburettors, and fuel tanks, thereby more than offsetting whatever the components of the present invention might otherwise add to the cost of a car. The system may be constructed with readily available components and technology, and may be supplied in kit form as well as original equipment.

Obviously, numerous modifications and variations of the present invention are possible in light of the above teachings. For example, although described in connection with the operation of a motor vehicle, the present invention may be universally applied to any four-stroke engine for which its operation depends upon the internal combustion of fossil fuels. Therefore, it is to be understood that within the scope of the appended claims the invention may be practiced otherwise than as specifically described here.

CLAIMS

1. A fuel vapour system for an internal combustion engine having an intake manifold, which comprises:
(a) A tank for containing fuel vapour;
(b) A vapour equaliser mounted on and in fluid communication with the intake manifold of the engine;
(c) A vapour conduit connecting the tank to the vapour equaliser for delivering fuel vapour from the former to the latter;
(d) A vapour equaliser having a valve connected to it for controlling the admission of air to the vapour equaliser;
(e) A tank having a second valve connected to it for controlling the admission of air to the tank;
(f) A throttle for controlling the first and second valves so that the opening of the first valve precedes and exceeds the opening of the second valve.

2. The fuel vapour system as set forth in claim 1, further comprising a filter positioned in the vapour conduit for retarding the flow of fuel vapour from the tank to the vapour equaliser.

3. The fuel vapour system as set forth in claim 2, where the filter comprises carbon particles.

4. The fuel vapour system as set forth in claim 2, where the filter comprises carbon particles and neoprene fibres.

5. The fuel vapour system as set forth in claim 2, where the filter comprises a substantially tubular housing positioned in series in the vapour conduit, the housing containing a central portion comprising a mixture of carbon and neoprene and end portions comprising carbon positioned on each side of the central portion.

6. The fuel vapour system as set forth in claim 1, further comprising a filter positioned in the vapour equaliser, for retarding the flow of the fuel vapour to the engine intake manifold.

7. The fuel vapour system as set forth in claim 6, where the filter is positioned downstream of the first valve.

8. The fuel vapour system as set forth in claim 7, where the filter comprises carbon particles.

9. The fuel vapour system as set forth in claim 8, where the filter further comprises a porous support member having first and second recessed portions for containing the carbon particles, the first recessed portion being positioned opposite a vapour inlet port in the vapour equaliser to which the vapour conduit is connected, the second recessed portion being positioned opposite the intake manifold of the engine.

10. The fuel vapour system as set forth in claim 9, where the porous support member is comprised of neoprene.

11. The fuel vapour system as set forth in claim 1, with a further filter positioned in the tank for controlling the flow of fuel vapour into the vapour conduit in proportion to the degree of vacuum in the tank.

12. The fuel vapour system as set forth in claim 11, where the filter incorporates a method for reducing the amount of fuel vapour delivered to the vapour conduit when the engine is idling and when the engine has attained a steady speed.

13. The fuel vapour system as set forth in claim 12, where the throttle acts to close the second valve when the engine is idling and when the engine has attained a steady speed to thereby increase the vacuum pressure in the tank.

14. The fuel vapour system as set forth in claim 13, where the filter comprises a frame pivotally mounted within the tank and movable between first and second operating positions, the first operating position corresponding to an open condition of the second valve, said second operating position corresponding to a closed condition of the second valve.

15. The fuel vapour system as set forth in claim 14, where the tank includes a vapour outlet port to which one end of the vapour conduit is connected, and where the second operating position of the frame places the filter in direct fluid communication with the vapour outlet port.

16. The fuel vapour system as set forth in claim 15, where the filter includes carbon particles.

17. The fuel vapour system as set forth in claim 16, where the filter includes neoprene filter material.

18. The fuel vapour system as set forth in claim 17, where the filter comprises a layer of carbon particles sandwiched between two layers of neoprene filter material, and a screen for supporting them within the pivotable frame.

19. The fuel vapour system as set forth in claim 18, further comprising a mechanism positioned on the filter for placing the filter in direct fluid communication with the vapour outlet port when the frame is in the second operating position.

20. A fuel vapour system for an internal combustion engine having an intake manifold, which comprises:
(a) A tank for containing fuel vapour;
(b) A vapour equaliser mounted on, and in fluid communication with, the intake manifold of the engine;
(c) A vapour conduit connecting the tank to the vapour equaliser for delivering fuel vapour from the former to the latter;
(d) A vapour equaliser having a first valve connected to it for controlling the admission of air to the vapour equaliser;
(e) A tank having a second valve connected to it for controlling the admission of air to the tank;
(f) A filter positioned in the vapour conduit for retarding the flow of the fuel vapour from the tank to the vapour equaliser means.

21. The fuel vapour system as set forth in claim 20, where the filter comprises a substantially tubular housing positioned in series in the vapour conduit, the housing containing a central portion comprising a mixture of carbon and neoprene and end portions comprising carbon positioned on each side of the central portion.

22. A fuel vapour system for an internal combustion engine having an intake manifold, which comprises:
(a) A tank for containing fuel vapour;
(b) A vapour equaliser mounted on and in fluid communication with the intake manifold of the engine;
(c) A vapour conduit connecting the tank to the vapour equaliser for delivering fuel vapour from the former to the latter;
(d) The vapour equaliser having a first valve connected to it for controlling the admission of air to the vapour equaliser;
(e) The tank having a second valve connected to it for controlling the admission of air to the tank;

(f) A filter positioned in the vapour equaliser for retarding the flow of the fuel vapour to the engine intake manifold.

23. The fuel vapour system as set forth in claim 22, where the filter is positioned downstream of the first valve, the filter comprises carbon particles and a porous support member having first and second recessed portions for containing the carbon particles, the first recessed portion being positioned opposite a vapour inlet port in the vapour equaliser to which the vapour conduit is connected, the second recessed portion being positioned opposite the intake manifold of the engine, and where the porous support member is comprised of neoprene.

The Permanent Magnet Motor of Stephen Kundel

US Patent 7,151,332 19th December 2006 Inventor: Stephen Kundel

MOTOR HAVING RECIPROCATING AND ROTATING PERMANENT MAGNETS

This patent describes a motor powered mainly by permanent magnets. This system uses a rocking frame to position the moving magnets so that they provide a continuous turning force on the output shaft.

ABSTRACT

A motor which has a rotor supported for rotation about an axis, and at least one pair of rotor magnets spaced angularity about the axis and supported on the rotor, at least one reciprocating magnet, and an actuator for moving the reciprocating magnet cyclically toward and away from the pair of rotor magnets, and consequently rotating the rotor magnets relative to the reciprocating magnet.

US Patent References:

0561144	June, 1896	Trudeau
1724446	August, 1929	Worthington
2790095	April, 1957	Peek et al.
3469130	September, 1969	Jines et al.
3703653	November, 1972	Tracy
3811058	May, 1974	Kiniski
3879622	April, 1975	Ecklin
3890548	June, 1975	Gray
3899703	August, 1975	Kinnison
3967146	June, 1976	Howard
3992132	November, 1976	Putt
4011477	March, 1977	Scholin
4151431	April, 1979	Johnson
4179633	December, 1979	Kelly
4196365	April, 1980	Presley
4267647	May, 1981	Anderson et al.
4629921	December, 1986	Gavaletz
4751486	June, 1988	Minato
5402021	March, 1995	Johnson
5594289	January, 1997	Minato
5634390	June, 1997	Takeuchi et al.
5751083	May, 1998	Tamura et al.
5925958	July, 1999	Pirc
6169343	January, 2001	Rich, Sr.
6343419	February, 2002	Litman et al.
6841909	January, 2005	Six
20020167236	November, 2002	Long
20040140722	July, 2004	Long

BACKGROUND OF THE INVENTION

This invention relates to the field of motors. More particularly, it pertains to a motor whose rotor is driven by the mutual attraction and repulsion of permanent magnets located on the rotor and an oscillator.

Various kinds of motors are used to drive a load. For example, hydraulic and pneumatic motors use the flow of pressurised liquid and gas, respectively, to drive a rotor connected to a load. Such motors must be continually supplied with pressurised fluid from a pump driven by energy converted to rotating power by a prime mover, such as an internal combustion engine. The several energy conversion processes, flow losses and pumping losses decrease the operating efficiency of motor systems of this type.

Conventional electric motors employ the force applied to a current carrying conductor placed in a magnetic field. In a d. c. motor the magnetic field is provided either by permanent magnets or by field coils wrapped around clearly defined field poles on a stator. The conductors on which the force is developed are located on a rotor and supplied with electric current. The force induced in the coil is used to apply rotor torque, whose magnitude varies

with the magnitude of the current and strength of the magnetic field. However, flux leakage, air gaps, temperature effects, and the counter-electromotive force reduce the efficiency of the motor.

Permanent dipole magnets have a magnetic north pole, a magnetic south pole, and magnetic fields surrounding each pole. Each magnetic pole attracts a pole of opposite magnetic polarity. Two magnetic poles of the same polarity repel each other. It is desired that a motor be developed such that its rotor is driven by the mutual attraction and repulsion of the poles of permanent magnets.

SUMMARY OF THE INVENTION

A motor according to the present invention includes a rotor supported for rotation about an axis, a first pair of rotor magnets including first and second rotor magnets spaced angularly about the axis and supported on the rotor, a reciprocating magnet, and an actuator for moving the reciprocating magnet cyclically toward and away from the first pair of rotor magnets, and cyclically rotating the first pair of rotor magnets relative to the reciprocating magnet. Preferably the motor includes a second pair of rotor magnets supported on the rotor, spaced axially from the first pair of rotor magnets, the second pair including a third rotor magnet and a fourth rotor magnet spaced angularly about the axis from the third rotor magnet. The reciprocating magnet is located axially between the first and second rotor magnet pairs, and the actuator cyclically moves the reciprocating magnet toward and away from the first and second pairs of rotor magnets.

The magnets are preferably permanent dipole magnets. The poles of the reciprocating magnet are arranged such that they face in opposite lateral directions.

The motor can be started by manually rotating the rotor about its axis. Rotation continues by using the actuator to move the reciprocating magnet toward the first rotor magnet pair and away from the second rotor magnet pair when rotor rotation brings the reference pole of the first rotor magnet closer to the opposite pole of the reciprocating magnet, and the opposite pole of the second rotor magnet closer to the reference pole of the reciprocating magnet. Then the actuator moves the reciprocating magnet toward the second rotor magnet pair and away from the first rotor magnet pair when rotor rotation brings the reference pole of the third rotor magnet closer to the opposite pole of the reciprocating magnet, and the opposite pole of the fourth rotor magnet closer to the reference pole of the reciprocating magnet.

A motor according to this invention requires no power source to energise a field coil because the magnetic fields of the rotor and oscillator are produced by permanent magnets. A nine-volt DC battery has been applied to an actuator switching mechanism to alternate the polarity of a solenoid at the rotor frequency. The solenoid is suspended over a permanent magnet of the actuator mechanism such that rotor rotation and the alternating polarity of a solenoid causes the actuator to oscillate the reciprocating magnet at a frequency and phase relation that is most efficient relative to the rotor rotation.

The motor is lightweight and portable, and requires only a commercially available portable d. c. battery to power an actuator for the oscillator. No motor drive electronics is required. Operation of the motor is practically silent.

Various objects and advantages of this invention will become apparent to those skilled in the art from the following detailed description of the preferred embodiment, when read in light of the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

These and other advantages of the present invention will become apparent to those skilled in the art from the following detailed description of a preferred embodiment when considered in the light of the accompanying drawings in which:

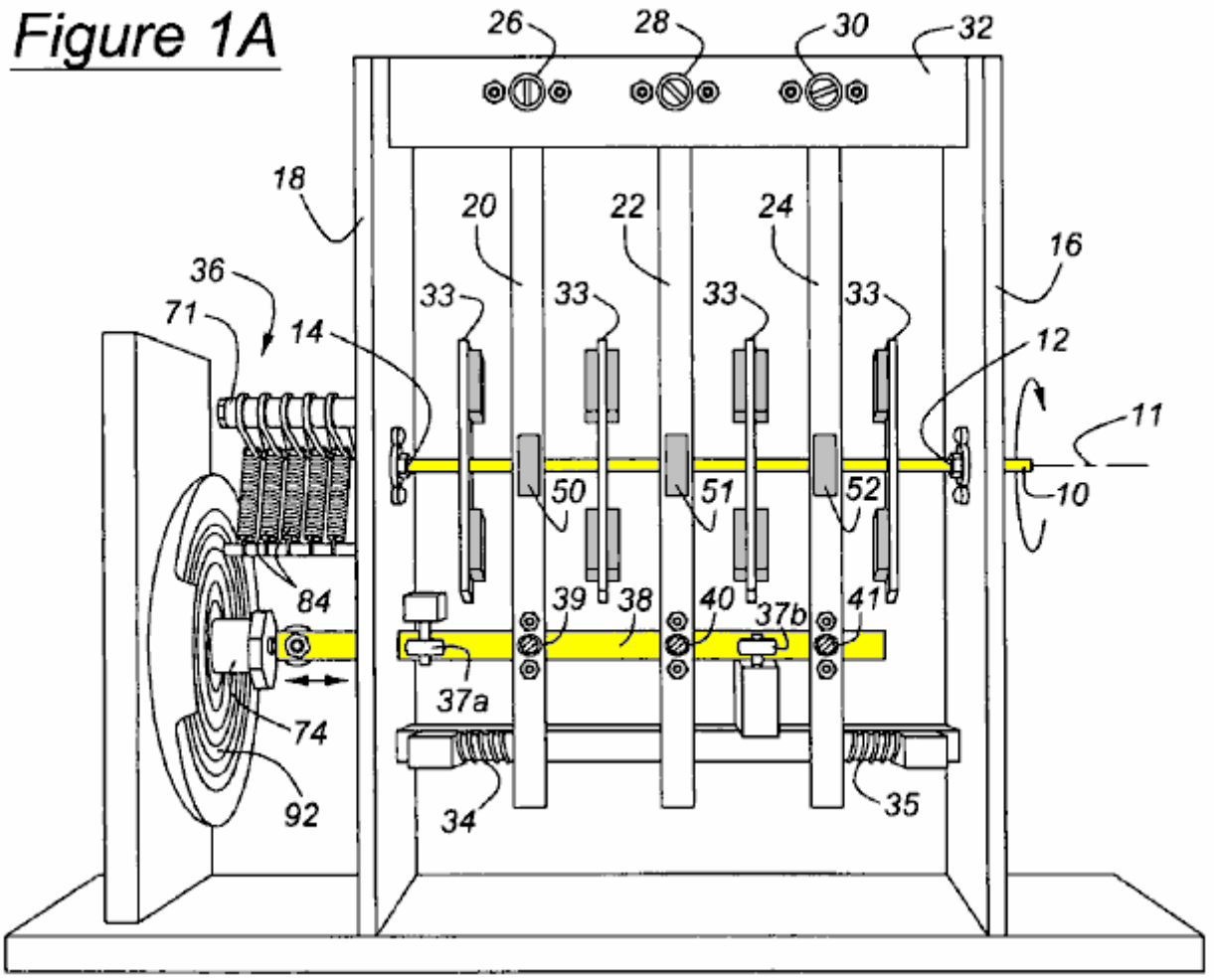


Fig.1A is a side view of a motor according to this invention;

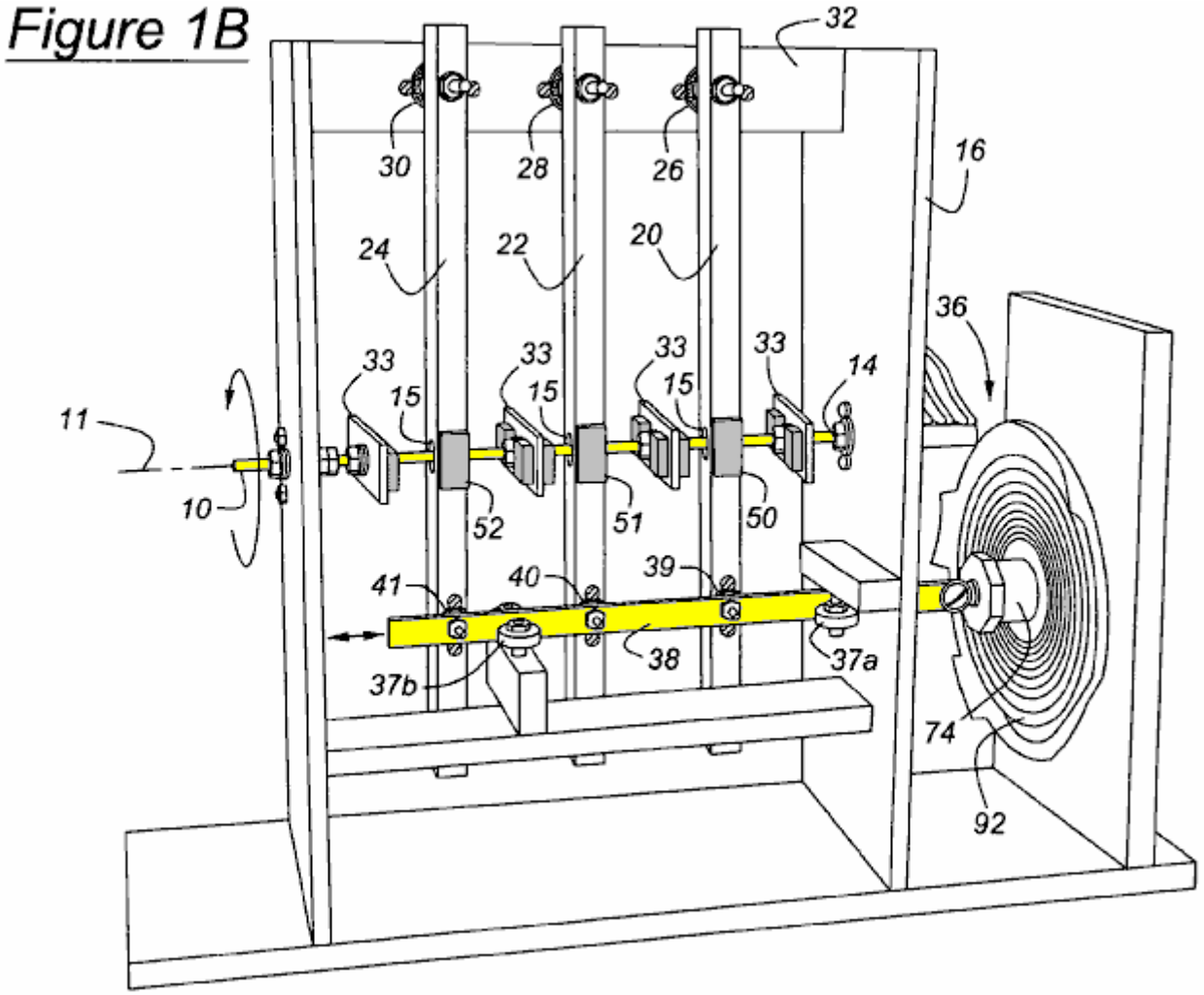


Fig.1B is a perspective view of the motor of **Fig.1A**

Figure 2

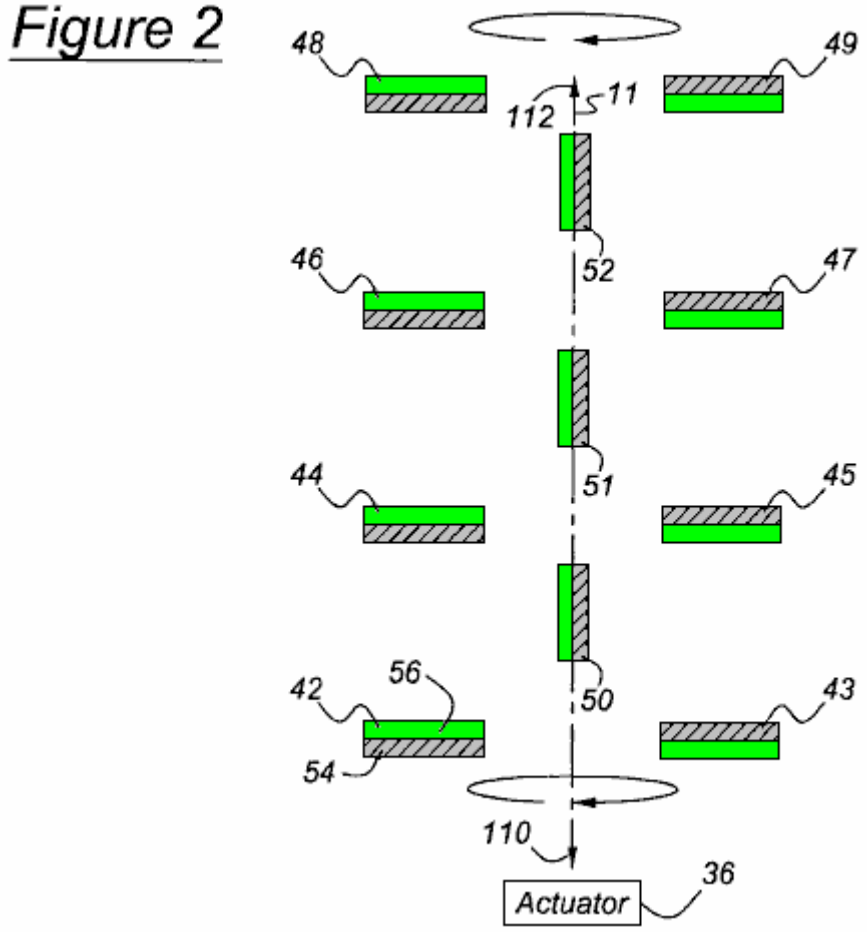


Fig.2 is a top view of the of motor of **Fig.1A** and **Fig.1B** showing the rotor magnets disposed horizontally and the reciprocating magnets located near one end of their range of travel

Figure 3

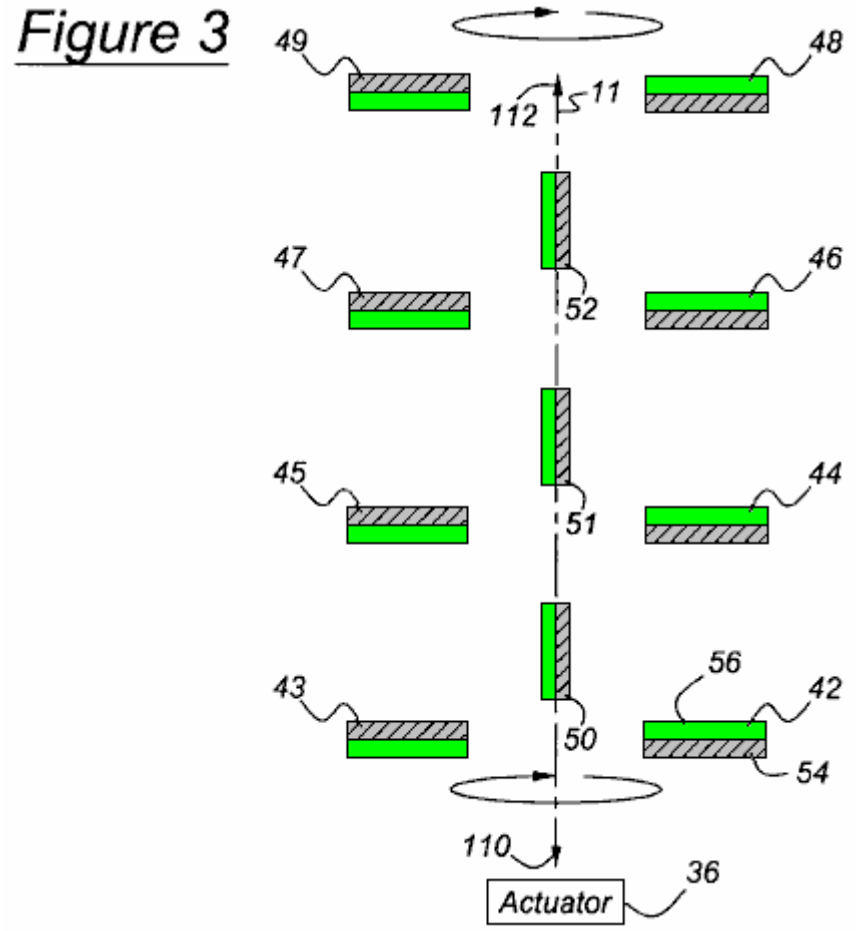


Fig.3 is a top view of the motor of **Fig.2** showing the rotor magnets rotated one-half revolution from the position shown in **Fig.2**, and the reciprocating magnets located near the opposite end of their range of travel

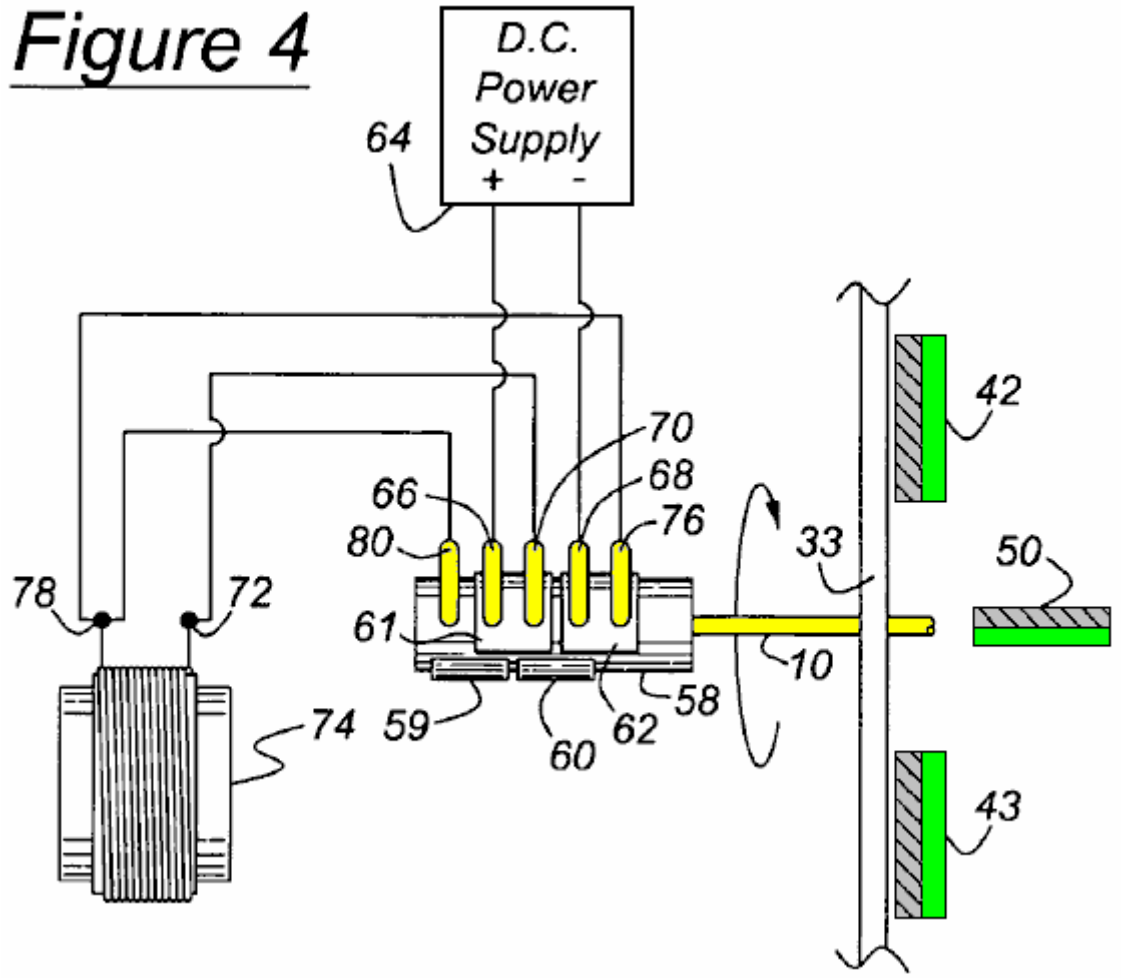


Fig.4 is a schematic diagram of a first state of the actuator switching assembly of the motor of **Fig.1**

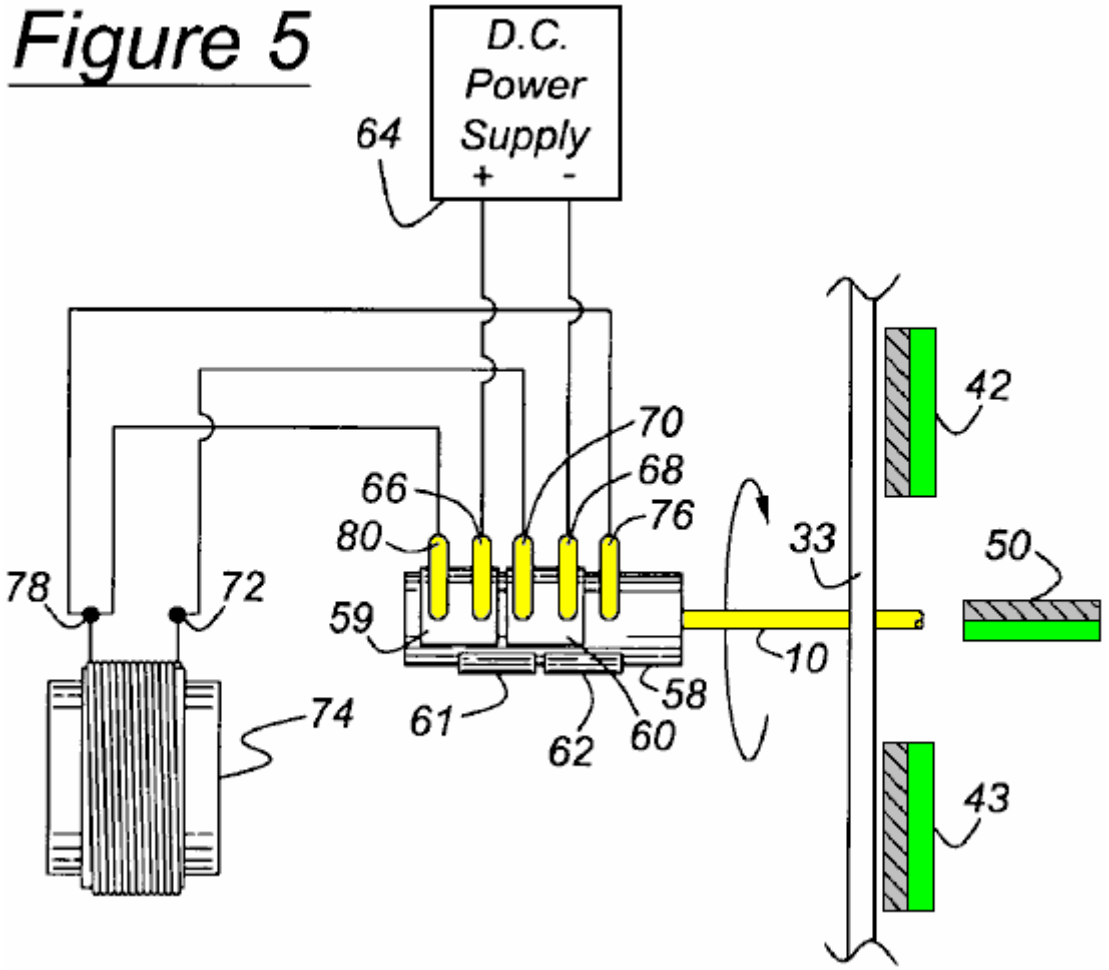


Fig.5 is a schematic diagram of a second state of the actuator switching assembly of the motor of **Fig.1**

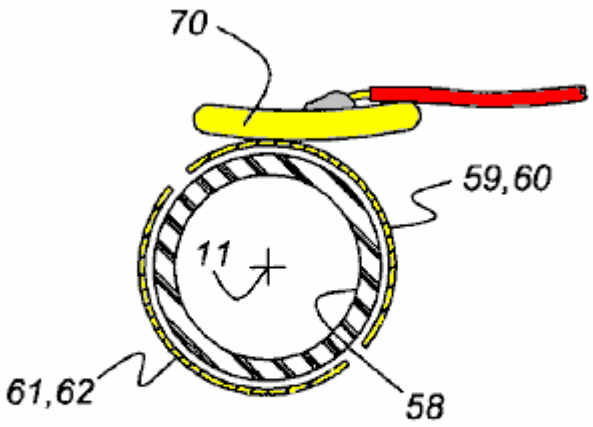


Figure 6

Fig.6 is cross sectional view of a sleeve shaft aligned with the rotor shaft showing a contact finger and bridge contact plates of the switching assembly

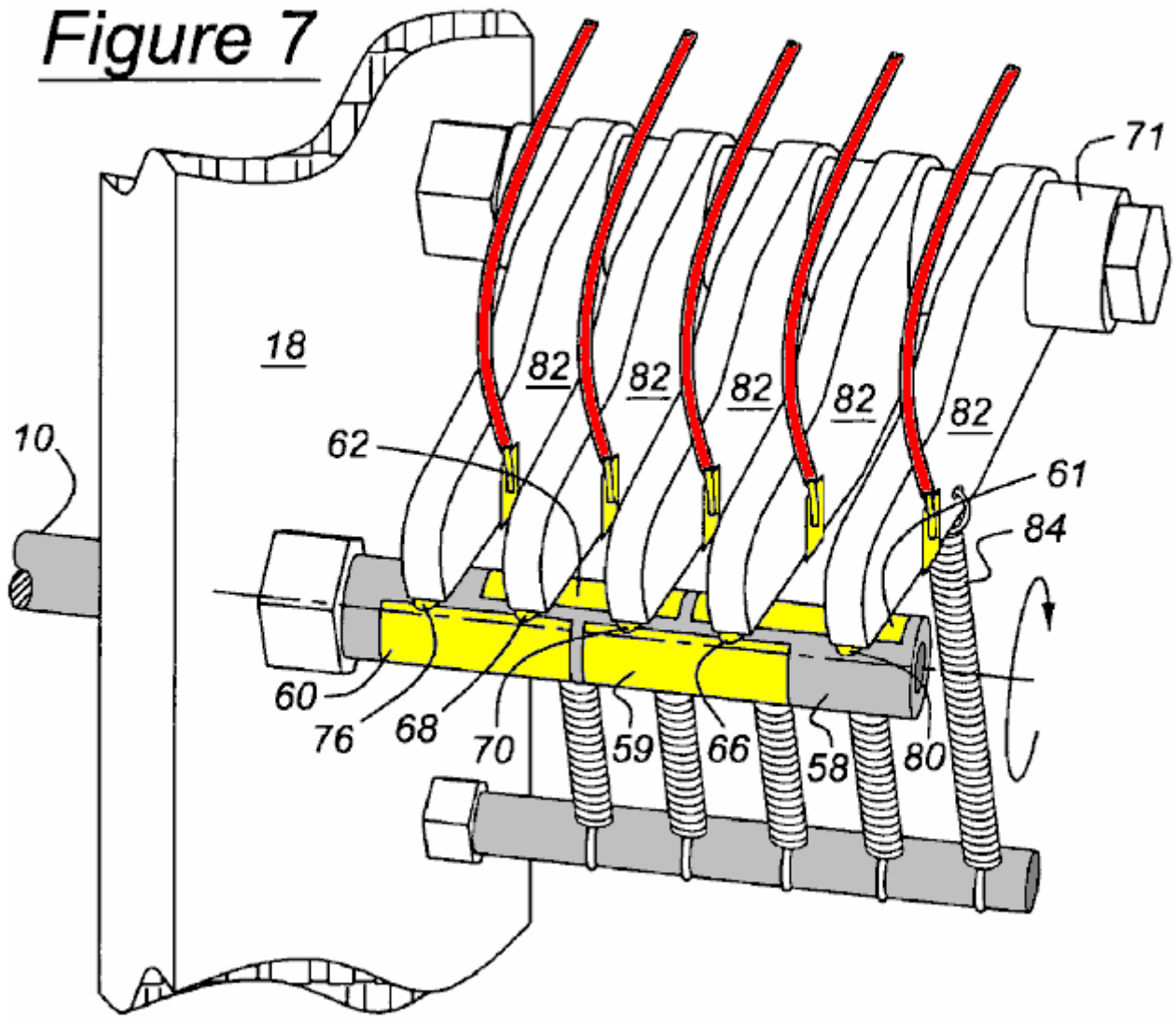


Fig.7 is an isometric view showing the switching contact fingers secured on pivoting arms and seated on the bridge connectors of the switching assembly

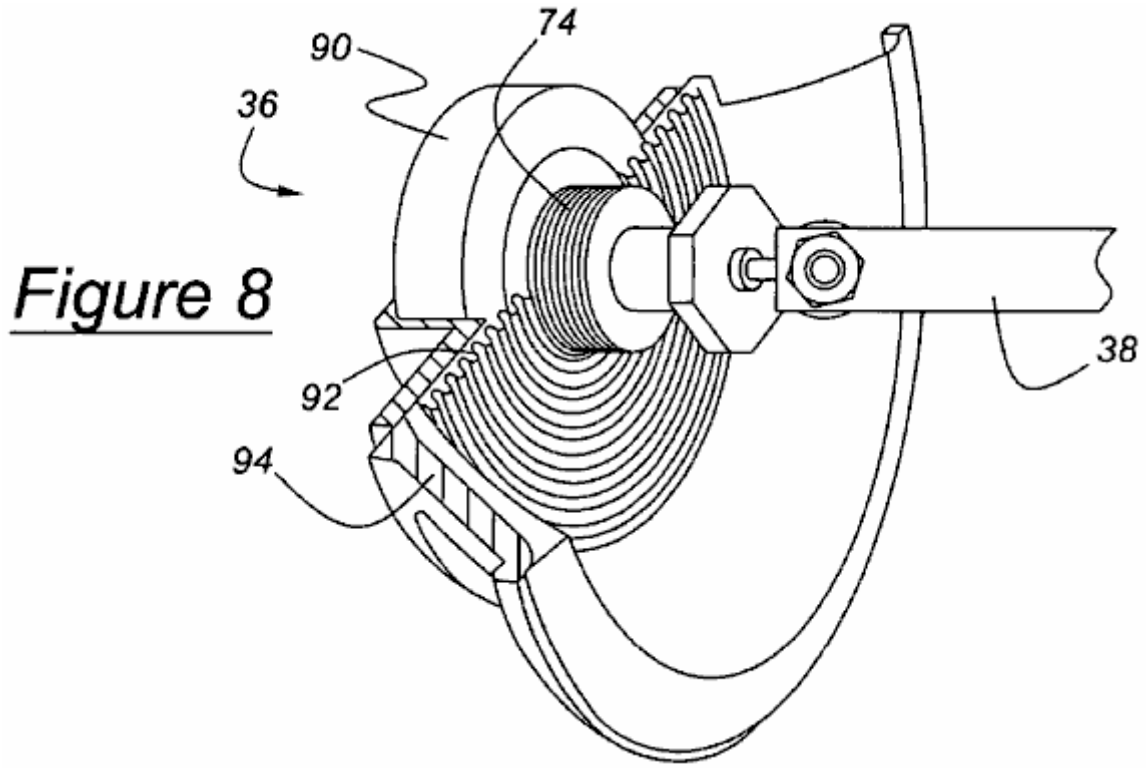


Fig.8 is isometric cross sectional view showing a driver that includes a solenoid and permanent magnet for oscillating the actuator arm in response to rotation of the rotor shaft

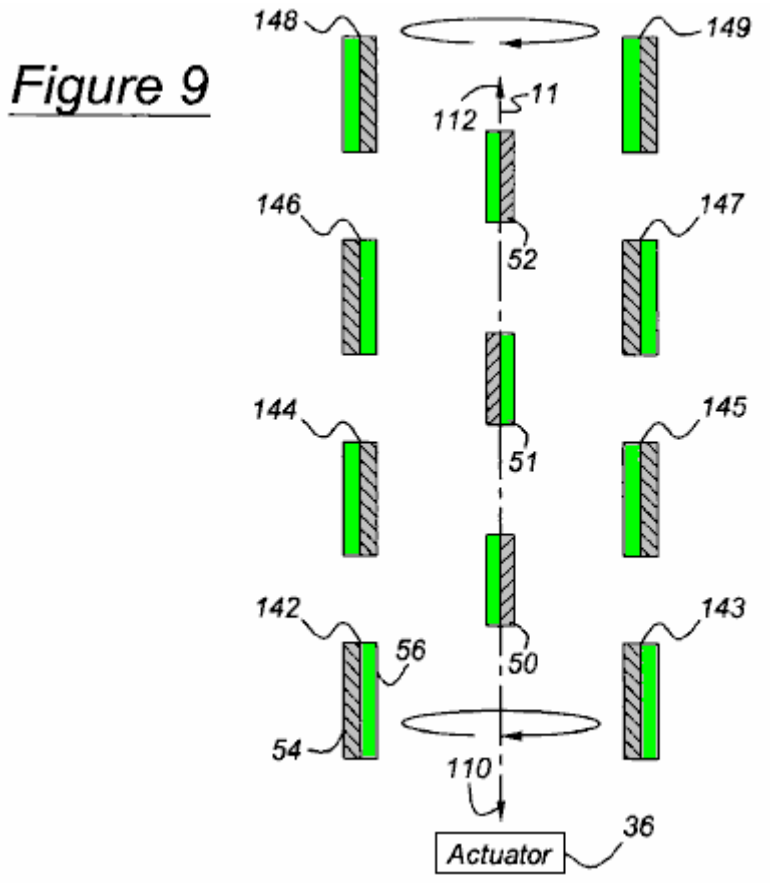


Fig.9 is a top view of an alternate arrangement of the rotor magnets, wherein they are disposed horizontally and rotated ninety degrees from the position shown in **Fig.2**, and the reciprocating magnets are located near an end of their range of displacement

Figure 10

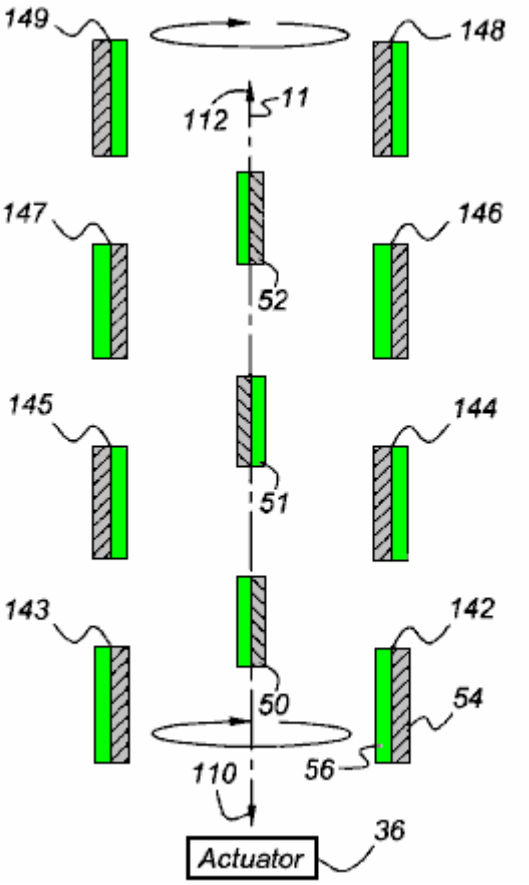


Fig.10 is a top view showing the rotor magnet arrangement of **Fig.9** rotated one-half revolution from the position shown in **Fig.9**, and the reciprocating magnets located near the opposite end of their range of displacement; and

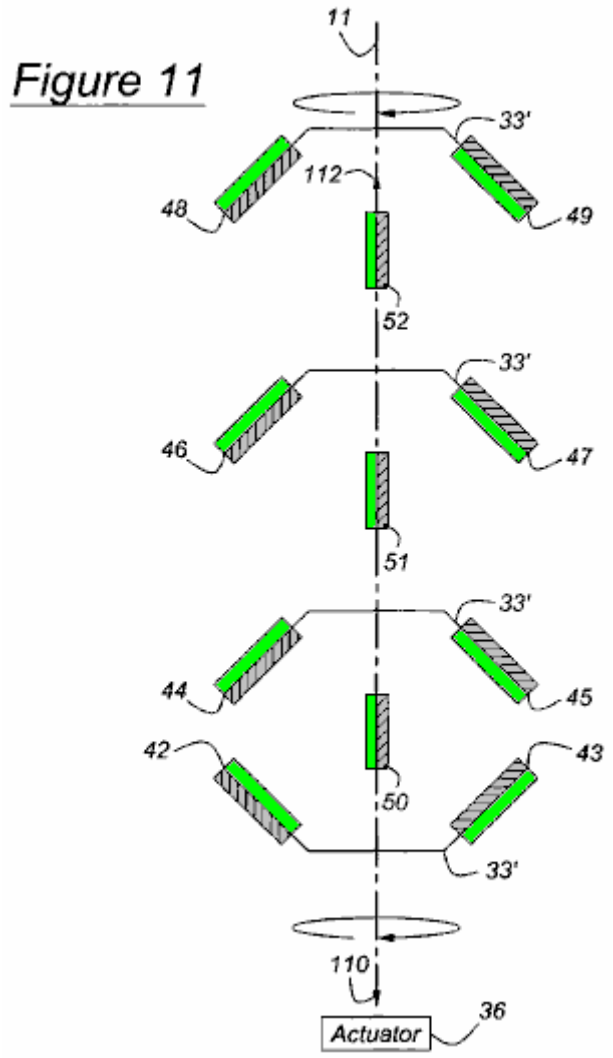


Fig.11 is a top view of the motor showing a third arrangement of the rotor magnets, which are canted with respect to the axis and the reciprocating magnets.

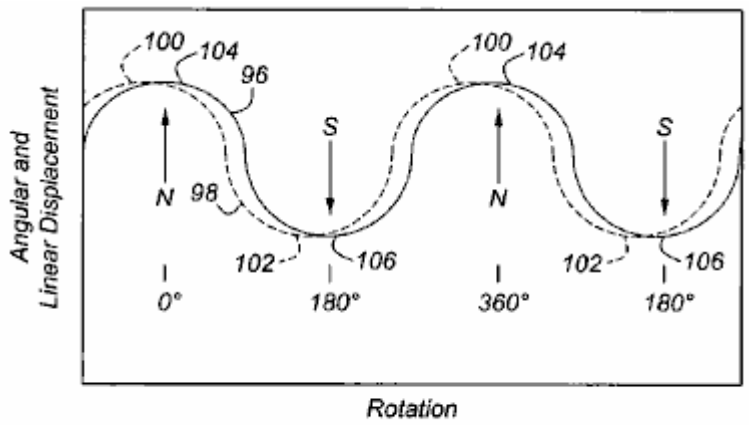


Figure 12

Fig.12 is a graph showing the angular displacement of the rotor shaft **10** and linear displacement of the reciprocating magnets

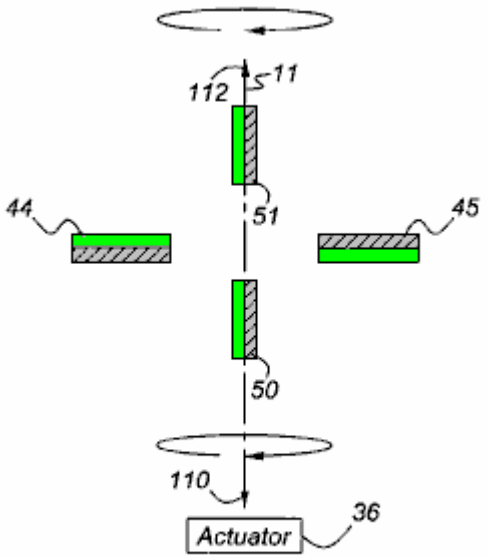


Figure 13

Fig.13 is a top view of a pair of rotor magnets disposed horizontally and reciprocating magnets located near one end of their range of travel

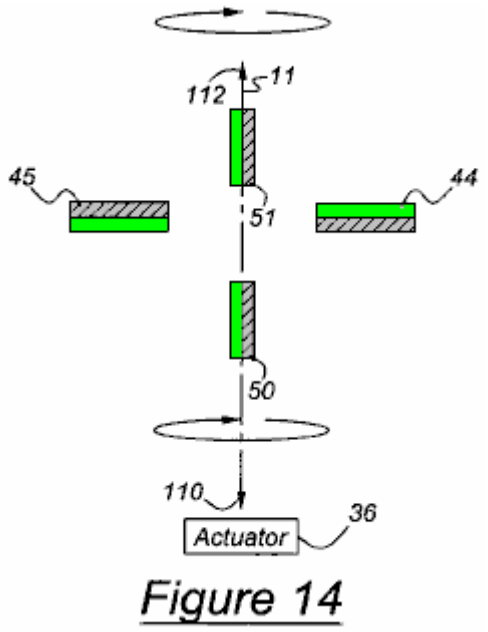


Figure 14

Fig.14 is a top view of the motor of **Fig.13** showing the rotor magnets rotated one-half revolution from the position shown in **Fig.13**, and the reciprocating magnets located near the opposite end of their range of travel; and

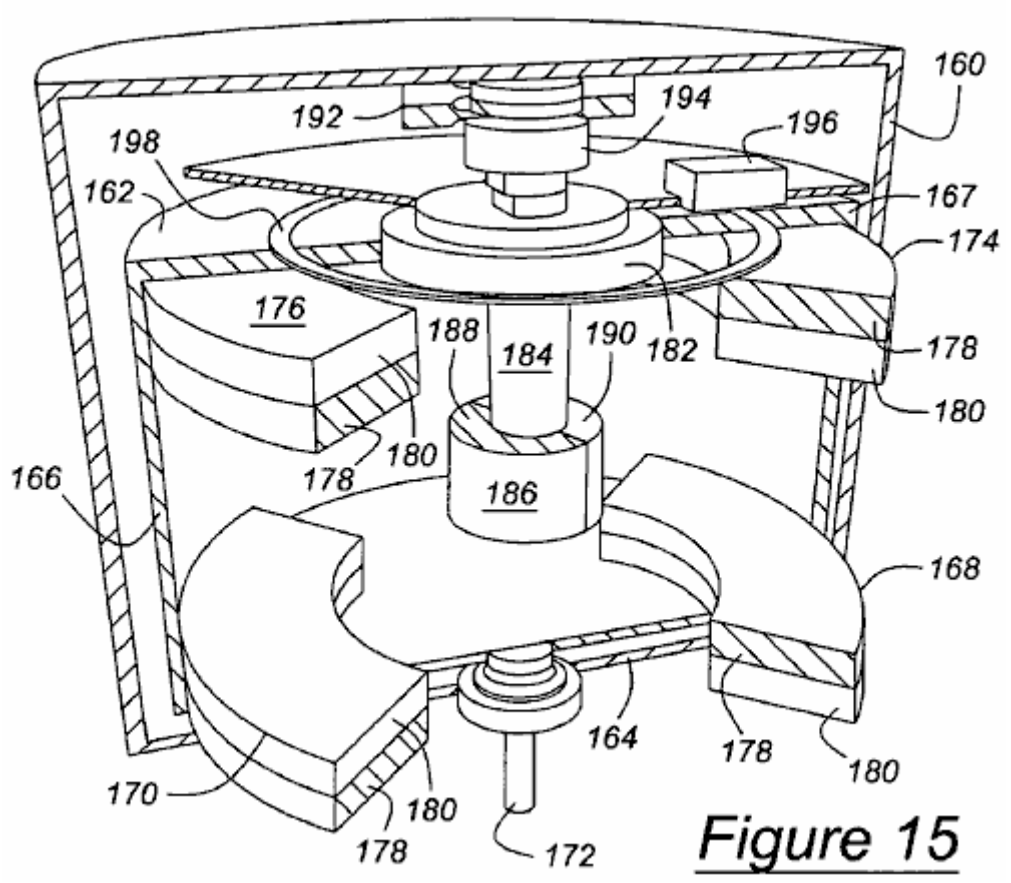


Figure 15

Fig.15 is a perspective cross sectional view of yet another embodiment of the motor according to this invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

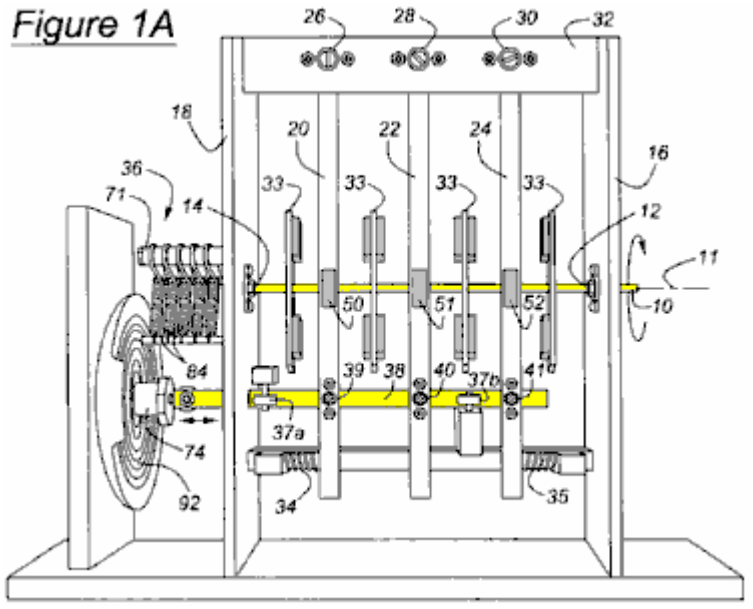


Figure 1A

A motor according to this invention, illustrated in **Fig.1A** and **Fig.1B** includes a rotor shaft **10** supported for rotation about axis **11** on bearings **12** and **14** located on vertical supports **16** and **18** of a frame. An oscillator mechanism includes oscillator arms **20**, **22** and **24** pivotally supported on bearings **26** , **28** and **30** respectively, secured to a horizontal support **32**, which is secured at each axial end to the vertical supports **16** and **18**. The oscillator arms **20**, **22** and **24** are formed with through holes **15** aligned with the axis **11** of rotor shaft **10**, the holes permitting rotation of the rotor shaft and pivoting oscillation of arms without producing interference between the rotor and the arms.

Extending in opposite diametric directions from the rotor axis **11** and secured to the rotor shaft **10** are four plates **33** , axially spaced mutually along the rotor axis, each plate supporting permanent magnets secured to the plate and rotating with the rotor shaft.

Each pivoting oscillator arm **20**, **22** and **24** of the oscillator mechanism support permanent magnets located between the magnets of the rotor shaft. Helical coiled compression return springs **34** and **35** apply oppositely directed forces to oscillator arms **20** and **24** as they pivot about their respective pivotal supports **26** and **30**, respectively. From the point of view of **Fig.1A** and **Fig.1B**, when spring **34** is compressed by displacement of the oscillator arm, the spring applies a force to the right to oscillator arm **20** which tends to return it to its neutral, starting position. When spring **35** is compressed by displacement of arm **24**, the spring applies a force to the left to arm **24** tending to return it to its neutral, starting position.

The oscillator arms **20**, **22** and **24** oscillate about their supported bearings **26**, **28** and **30** , as they move in response to an actuator **36**, which includes an actuator arm **38**, secured through bearings at **39**, **40** and **41** to the oscillator arms **20**, **22** and **24**, respectively. Actuator **36** causes actuator arm **38** to reciprocate linearly leftwards and rightwards from the position shown in **Fig.1A** and **Fig.1B**. The bearings **39**, **40** and **41**, allow the oscillator arms **20**, **22** and **24** to pivot and the strut to translate without mutual interference. Pairs of guide wheels **37a** and **37b** spaced along actuator arm **38**, each include a wheel located on an opposite side of actuator arm **38** from another wheel of the wheel-pair, for guiding linear movement of the strut and maintaining the oscillator arms **20**, **22** and **24** substantially in a vertical plane as they oscillate. Alternatively, the oscillator arms **20**, **22** and **24** may be replaced by a mechanism that allows the magnets on the oscillator arms to reciprocate linearly with actuator arm **38** instead of pivoting above the rotor shaft **10** at **26**, **28** and **30**.

Figure 2

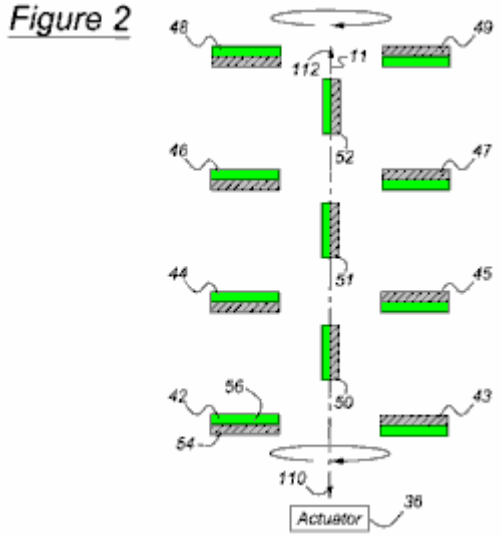


Fig.2 shows a first arrangement of the permanent rotor magnets **42 – 49** that rotate about axis **11** and are secured to the rotor shaft **10**, and the permanent reciprocating magnets **50 – 52** which move along axis **11** and are secured to the oscillating arms **20**, **22** and **24**. Each magnet has a pole of reference polarity and a pole of opposite polarity from that of the reference polarity. For example, rotor magnets **42**, **44**, **46** and **48**, located on one side of axis **11**, each have a north, positive or reference pole **54** facing actuator **36** and a south, negative or opposite pole **56** facing away from the actuator. Similarly, rotation magnets **43**, **45**, **47** and **49**, located diametrically opposite to rotor magnets **42**, **44**, **46** and **48**, each have a south pole facing toward actuator **36** and a north pole facing away from the actuator. The north poles **54** of the reciprocating magnets **50 – 52** face to the right from the point of view seen in **Fig.2** and **Fig.3** and their south poles **56** face towards the left.

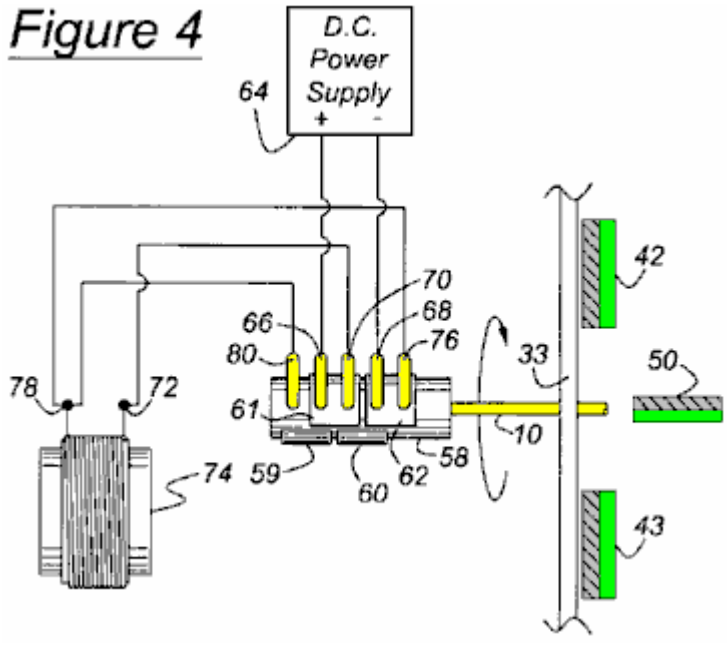


Fig.4 shows a switch assembly located in the region of the left-hand end of rotor shaft **10**. A cylinder, **58**, preferably formed of PVC, is secured to rotor shaft **10**. Cylinder **58** has contact plates **59** and **60**, preferably of brass, located on its outer surface, aligned angularly, and extending approximately 180 degrees about the axis **11**, as shown in **Fig.5**. Cylinder **58** has contact plates **61** and **62**, preferably made of brass, located on its outer surface, aligned angularly, extending approximately 180 degrees about the axis **11**, and offset axially with respect to contact plates **59** and **60**.

A D.C. power supply **64**, has its positive and negative terminals connected electrically through contact fingers **66** and **68**, to contact plates **61** and **62**, respectively. A third contact finger **70**, shown contacting plate **61**, connects terminal **72** of a solenoid **74** electrically to the positive terminal of the power supply **64** through contact finger **66** and contact plate **61**. A fourth contact finger **76**, shown contacting plate **62**, connects terminal **78** of solenoid **74** electrically to the negative terminal of the power supply **64** through contact finger **68** and contact plate **62**. A fifth contact finger **80**, axially aligned with contact plate **59** and offset axially from contact plate **61**, is also connected to terminal **78** of solenoid **74**.

Preferably the D.C. power supply **64** is a nine volt battery, or a D.C. power adaptor, whose input may be a conventional 120 volt, 60 Hz power source. The D.C. power supply and switching mechanism described with reference to **Figs. 4 to 7**, may be replaced by an A.C. power source connected directly across the terminals **72** and **78** of solenoid **74**. As the input current cycles, the polarity of solenoid **74** alternates, the actuator arm **38** moves relative to a toroidal permanent magnet **90** (shown in **Fig.8**), and the reciprocating magnets **50 – 52** reciprocate on the oscillating arms **20**, **22** and **24** which are driven by the actuator arm **38**.

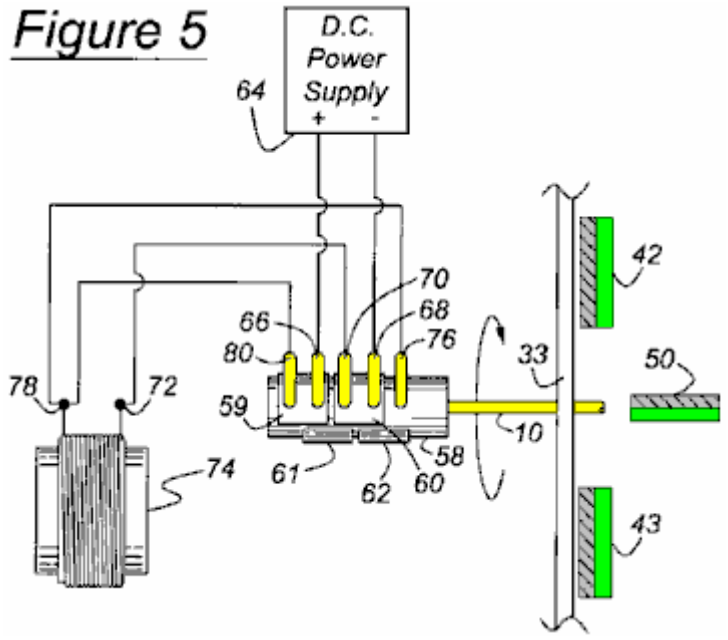


Fig.5 shows the state of the switch assembly when rotor shaft **10** has rotated approximately 180 degrees from the position shown in **Fig.4**. When the switch assembly is in the state shown in **Fig.5**, D.C. power supply **64** has its positive and negative terminals connected electrically by contact fingers **66** and **68** to contact plates **59** and **60**, respectively. Contact finger **70**, shown contacting plate **60**, connects terminal **72** of solenoid **74** electrically to the negative terminal of the power supply **64** through contact finger **68** and contact plate **60**. Contact finger **80**, shown contacting plate **59**, connects terminal **78** of solenoid **74** electrically to the positive terminal through contact finger **66** and contact plate **59**. Contact finger **76**, axially aligned with contact plate **62** and offset axially from contact plate **60**, remains connected to terminal **78** of solenoid **74**. In this way, the polarity of the solenoid **74** changes cyclically as the rotor **10** rotates through each one-half revolution.

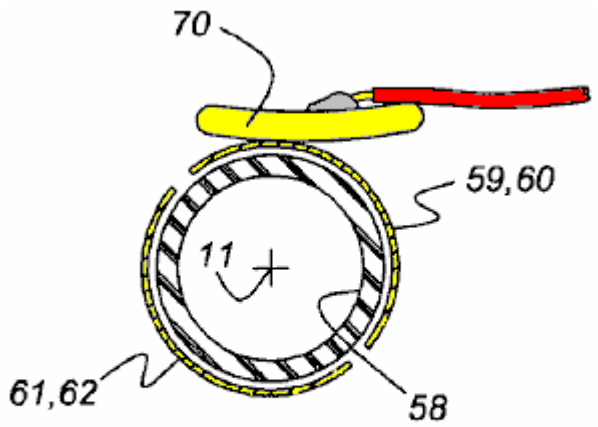
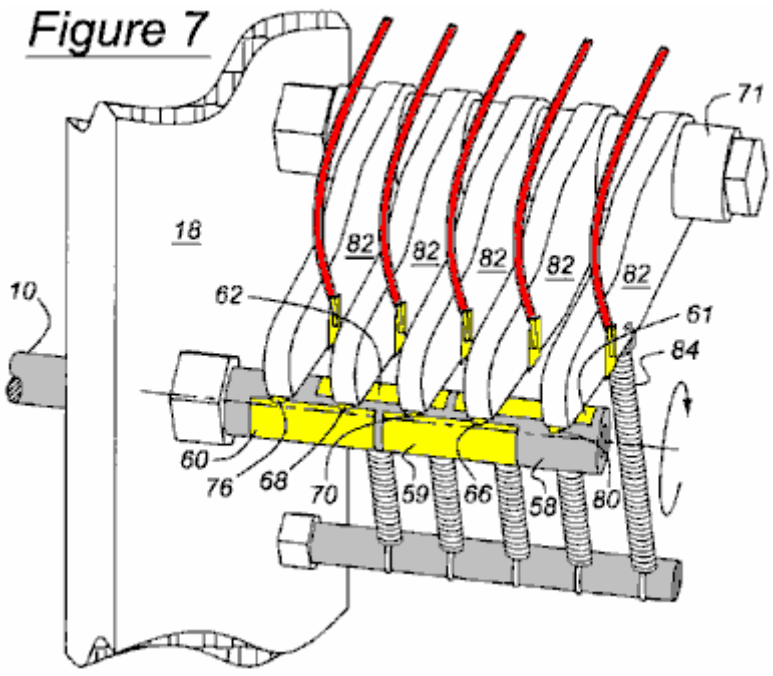


Figure 6

Fig.6 shows in cross-section, the cylinder **58** which is aligned with and driven by the rotor shaft **10**, a contact finger **70**, and the contact plates **59 – 62** of the switching assembly, which rotate with the rotor shaft and cylinder about the axis **11** .

Figure 7



As **Fig.7** illustrates, axially spaced arms **82** are supported on a stub shaft **71**, preferably made of Teflon or another self-lubricating material, to facilitate the pivoting of the arms about the axis of the shaft **71**. Each contact finger **66**, **68**, **70**, **76** and **80** is located at the end of a arm **82**, and tension springs **84**, secured to each arm **82**, urge the contact fingers **66**, **68**, **70**, **76** and **80** continually toward engagement with the contact plates **59 – 62**.

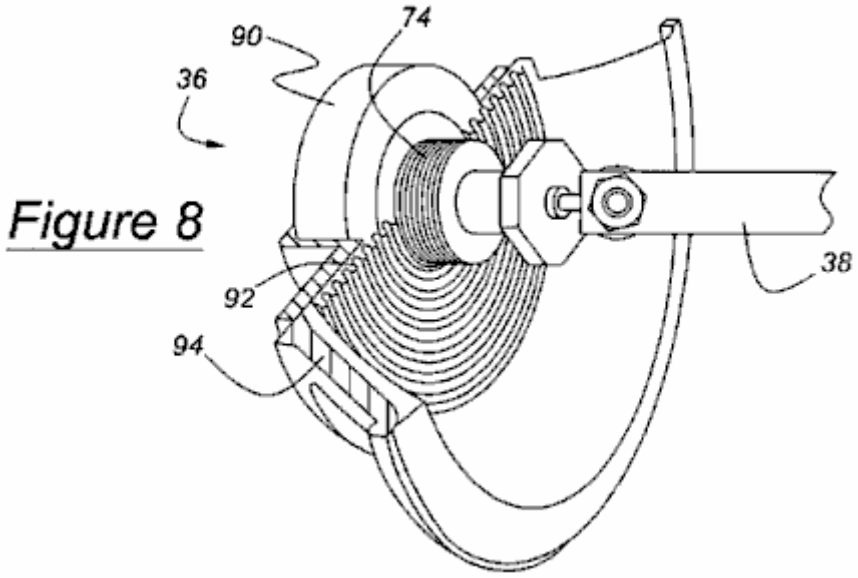


Figure 8

Fig.8 illustrates the actuator **36** for reciprocating the actuator arm **38** in response to rotation of the rotor shaft **10** and the alternating polarity of the solenoid **74**. The actuator **36**, includes the solenoid **74**, the toroidal permanent magnet **90**, an elastic flexible spider **92** for supporting the solenoid above the plane of the magnet, and a basket or frame **94**, to which the spider is secured. The actuator arm **38** is secured to solenoid **74**. The polarity of the solenoid **74** changes as rotor shaft **10** rotates, causing the solenoid and actuator arm **38** to reciprocate due to the alternating polarity of the solenoid relative to that of the toroidal permanent magnet **90**. As the solenoid polarity changes, the actuator arm **38** reciprocates linearly due to the alternating forces of attraction and repulsion of the solenoid **74** relative to the poles of the magnet **90**. The actuator arm **38** is secured to the oscillator arms **20**, **22** and **24** causing them to pivot, and the reciprocating magnets **50 – 52**, secured to the oscillator arms, to reciprocate. Alternatively, the reciprocating magnets **50 – 52** can be secured directly to the arm **38** , so that the magnets **50 – 52** reciprocate without need for an intermediary oscillating component.

It is important to note at this point in the description that, when two magnets approach each other with their poles of like polarity facing each other but slightly offset, there is a tendency for the magnets to rotate to the opposite pole of the other magnet. Therefore, in the preferred embodiment of the instant invention, the angular position at which the switch assembly of the actuator **36** changes between the states of **Fig.4** and **Fig.5** is slightly out of phase with the angular position of the rotor shaft **10** to help sling or propel the actuator arm **38** in the reverse direction at the preferred position of the rotor shaft. The optimum phase offset is approximately 5–8 degrees. This way, advantage is taken of each rotor magnet's tendency to rotate about its own magnetic field when slightly offset from the respective reciprocating magnet, and the repulsive force between like poles of the reciprocating magnets and the rotor magnets is optimised to propel the rotor magnet about the rotor axis **11**, thereby increasing the motor's overall efficiency.

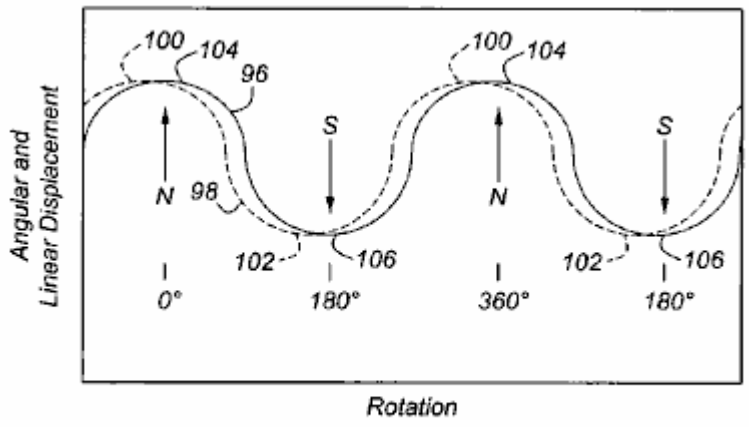


Figure 12

Fig.12 is a graph showing the angular displacement **96** of the rotor shaft **10** and linear displacement **98** of the reciprocating magnets **50 – 52**. Point **100** represents the end of the range of displacement of the reciprocating magnets **50 – 52** shown in FIGS. 2 and 9, and point 102 represents the opposite end of the range of displacement of the reciprocating magnets **50 – 52** shown in FIGS. 3 and 10. Point 104 represents the angular position of the

rotor magnets **42 – 49** when in the horizontal plane shown in FIGS. 2 and 9, and point 106 represents the angular position of the rotor magnets **42 – 49** when rotated one-half rotation to the horizontal plane shown in **Fig.3** and **Fig.10**. Preferably, the reciprocating magnets **50 – 52** and rotor magnets **42 – 49** are out of phase: the reciprocating magnets lead and the rotor magnets lag by several degrees. The reciprocating magnets **50 – 52** reach the respective extremities of their range of travel before rotor rotation moves the rotor magnets **42 – 49** into the horizontal plane.

When the reference poles **54** and opposite poles **56** of the rotor magnets **42 – 49** and reciprocating magnets **50 – 52** are arranged as shown in **Fig.2** and **Fig.3**, the rotor position is stable when the rotor magnets are in a horizontal plane. The rotor position is unstable in any other angular position, and it moves towards horizontal stability from any unstable position, and is least stable when the rotor magnets **42 – 49** are in a vertical plane. The degree of stability of the rotor shaft **10** is a consequence of the mutual attraction and repulsion of the poles of the rotor magnets **42 – 49** and reciprocating magnets **50 – 52** and the relative proximity among the poles. In **Fig.2**, the reciprocating magnets **50 – 52** are located at a first extremity of travel. In **Fig.3**, the reciprocating magnets **50 – 52** have reciprocated to the opposite extremity of travel, and the rotor magnets have rotated one-half revolution from the position shown in **Fig.2**.

When the rotor is stopped, its rotation can be easily started manually by applying torque in either direction. Actuator **36** sustains rotor rotation after it is connecting to its power source. Rotation of rotor shaft **10** about axis **11** is aided by cyclic movement of the reciprocating magnets **50 – 52**, their axial location between the rotor magnet pairs **42 – 43** , **44 – 45** , **46 – 47** and **48 – 49**, the disposition of their poles in relation to the poles of the rotor magnets, and the frequency and phase relationship of their reciprocation relative to rotation of the rotor magnets. Actuator **36** maintains the rotor **10** rotating and actuator arm **38** oscillating at the same frequency, the phase relationship being as described with reference to **Fig.12**.

With the rotor magnets **42** and **49** as shown in **Fig.2**, when viewed from above, the north poles **54** of the rotor magnets on the left-hand side of axis **11** face a first axial direction **110**, i.e., toward the actuator **36**, and the north poles **54** of the rotor magnets on the right-hand side of axis **11** face in the opposite axial direction **112**, away from actuator **36**. When the rotor magnets **42 – 49** are located as in **Fig.2**, the north poles **54** of reciprocating magnets **50 – 52** are adjacent the south poles **56** of rotor magnets **45**, **47** and **49** , and the south poles **56** of reciprocating magnets **50 – 52** are adjacent the north poles **54** of rotor magnets **44**, **46** and **48**.

Figure 2

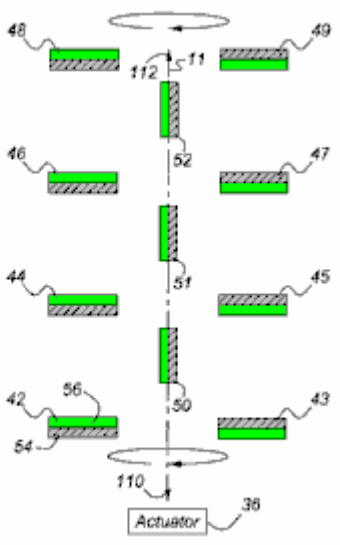
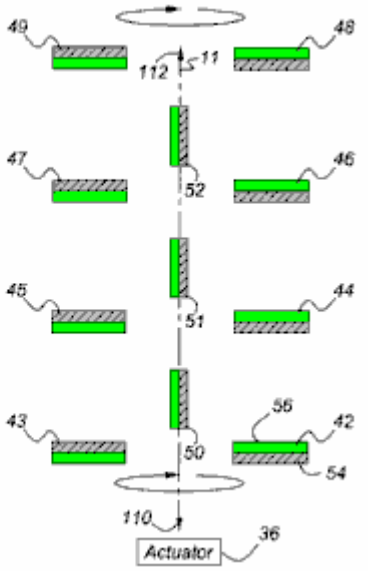


Figure 3



Furthermore, when the rotor shaft 10 rotates to the position shown in **Fig.2**, the reciprocating magnets **50 – 52** are located at, or near, one extremity of their axial travel, so that the north poles **54** of reciprocating magnets **50 – 52** are located close to the south poles **56** of rotor magnets **45**, **47** and **49**, respectively, and relatively more distant from the north poles **54** of rotor magnets **43**, **45** and **47**, respectively. Similarly, the south poles **56** of reciprocating magnets **50 – 52** are located close to the north poles of rotor magnet **44**, **46** and **48**, respectively, and relatively more distant from the south poles of rotor magnets **42**, **44** and **46**, respectively.

With the rotor magnets **42** and **49** rotated into a horizontal plane one-half revolution from the position of **Fig.1B**, when viewed from above as shown in **Fig.3**, the north poles **54** of reciprocating magnets **50 – 52** are located adjacent the south poles of rotor magnets **42**, **44** and **46**, and the south poles **56** of reciprocating magnets **50 – 52** are located adjacent the north poles **54** of rotor magnets **43**, **45** and **47**, respectively. When the rotor **10** shaft is located as shown in **Fig.3**, the reciprocating magnets **50 – 52** are located at or near the opposite extremity of their

axial travel from that of **Fig.2**, such that the north poles **54** of reciprocating magnets **50 – 52** are located close to the south poles **56** of rotor magnet **42**, **44** and **46**, respectively, and relatively more distant from the north poles of rotor magnets **44**, **46** and **48**, respectively. Similarly, when the rotor shaft 10 is located as shown in FIG. 3, the south poles 56 of reciprocating magnets **50 – 52** are located close to the north poles of rotor magnet **43**, **45** and **47**, respectively, and relatively more distant from the south poles of rotor magnets **45**, **47** and **49**, respectively.

In operation, rotation of rotor shaft **10** in either angular direction is started manually or with a starter-actuator (not shown). Actuator **36** causes reciprocating magnets **50 – 52** to oscillate or reciprocate at the same frequency as the rotational frequency of the rotor shaft **10**, i.e. one cycle of reciprocation per cycle of rotation, preferably with the phase relationship illustrated in **Fig.12**. When the reciprocating magnets **50 – 52** are located as shown in **Fig.2**, the rotor shaft **10** will have completed about one-half revolution from the position of **Fig.3** to the position of **Fig.2**.

Rotation of the rotor **10** is aided by mutual attraction between the north poles **54** of the reciprocating magnets **50 – 52** and the south poles **56** of the rotor magnets **43**, **45**, **47** and **49** that are then closest respectively to those north poles of reciprocating magnets **50 – 52**, and mutual attraction between the south poles of reciprocating magnets **50 – 52** and the north poles of the rotor magnets **42**, **44**, **46** and **48** that are then closest respectively to the north poles of the reciprocating magnets.

Assume rotor shaft **10** is rotating counterclockwise when viewed from the actuator **36**, and the rotor magnets **42**, **44**, **46** and **48** are located above rotor magnets **43**, **45**, **47** and **49**. With the rotor shaft **10** positioned so that the reciprocating magnets **50 – 52** are approximately mid-way between the positions shown in **Fig.2** and **Fig.3** and moving toward the position shown in **Fig.2**, as rotation proceeds, the south pole of each reciprocating magnet **50 – 52** applies a downward attraction to the north pole **54** of the closest of the rotor magnets **44**, **46** and **48**, and the north pole **54** of each reciprocating magnet **50 – 52** attracts upwards the south pole **56** of the closest rotor magnet **45**, **47** and **49**. This mutual attraction of the poles causes the rotor to continue rotating counterclockwise to the position of **Fig.2**.

Then the reciprocating magnets **50 – 52** begin to move toward the position shown in **Fig.3**, and rotor inertia overcomes the steadily decreasing force of attraction between the poles as they move mutually apart, permitting the rotor shaft **10** to continue its counterclockwise rotation into the vertical plane where rotor magnets **43**, **45**, **47** and **49** are located above rotor magnets **42**, **44**, **46** and **48**. As rotor shaft **10** rotates past the vertical plane, the reciprocating magnets **50 – 52** continue to move toward the position of **Fig.3**, the south pole **56** of each reciprocating magnet **50 – 52** attracts downward the north pole of the closest rotor magnet **43**, **45** and **47**, and the north pole **54** of each reciprocating magnet **50 – 52** attracts upward the south pole **56** of the closest rotor magnet **42**, **44** and **46**, causing the rotor **10** to rotate counterclockwise to the position of **Fig.3**. Rotor inertia maintains the counterclockwise rotation, the reciprocating magnets **50 – 52** begin to move toward the position shown in **Fig.2**, and the rotor shaft **10** returns to the vertical plane where rotor magnets **43**, **45**, **47** and **49** are located above rotor magnets **42**, **44**, **46** and **48**, thereby completing one full revolution.

Figure 9

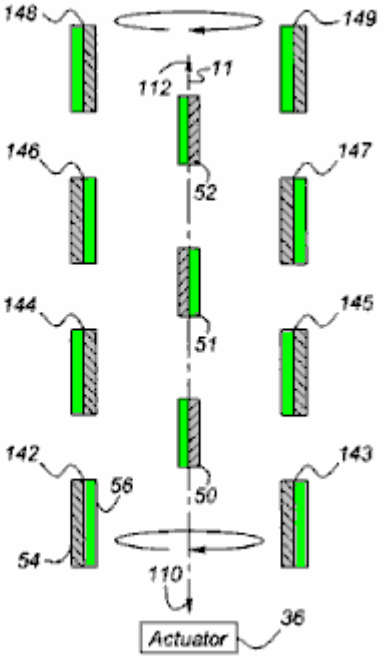


Figure 10

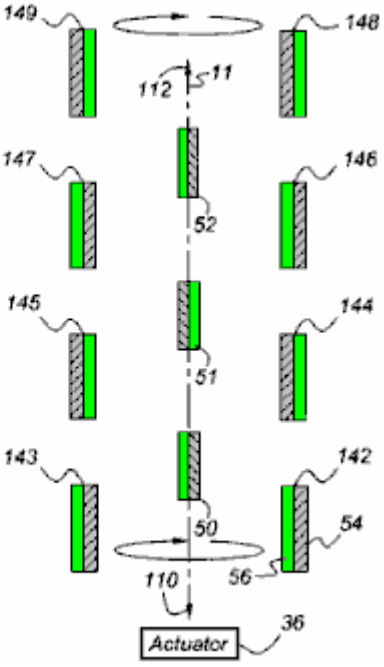


Fig.9 and **Fig.10** show a second arrangement of the motor in which the poles of the rotor magnets **142 – 149** are parallel to, and face the same direction as those of the reciprocating magnets **50 – 52**. Operation of the motor arranged as shown in **Fig.9** and **Fig.10** is identical to the operation described with reference to **Fig.2** and **Fig.3**. In the embodiment of **Fig.9** and **Fig.10**, the poles of the reciprocating magnets **50 – 52** face more directly the poles of the rotor magnets **142 – 149** in the arrangement of **Fig.2** and **Fig.3**. The forces of attraction and repulsion between the poles are greater in the embodiment of **Fig.9** and **Fig.10**, therefore, greater torque is developed. The magnitude of torque is a function of the magnitude of the magnetic forces, and the distance through which those force operate.

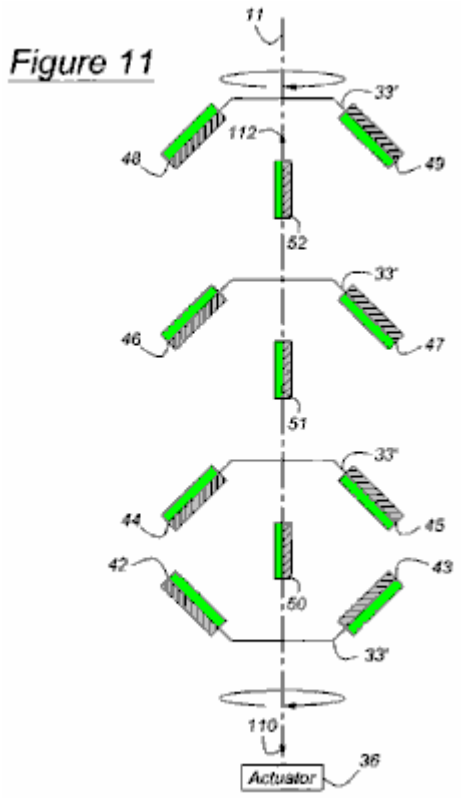


Fig.11 shows a third embodiment of the motor in which the radial outer portion of the rotor plates **33'** are skewed relative to the axis **11** such that the poles of the rotor magnets **42 – 49** are canted relative to the poles of the reciprocating magnets **50 – 52**. Operation of the motor arranged as shown in **Fig.11** is identical to the operation described with reference to **Fig.2** and **Fig.3**.

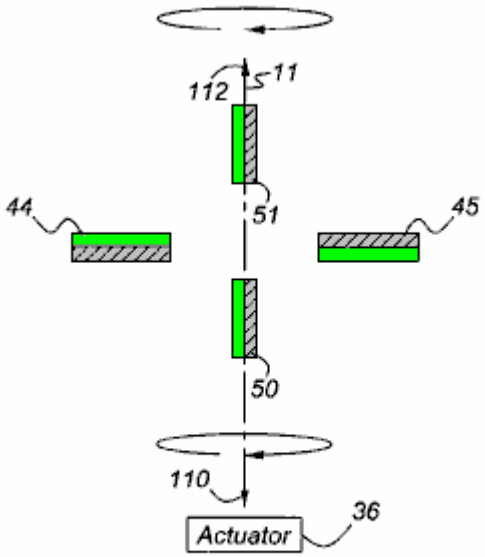


Figure 13

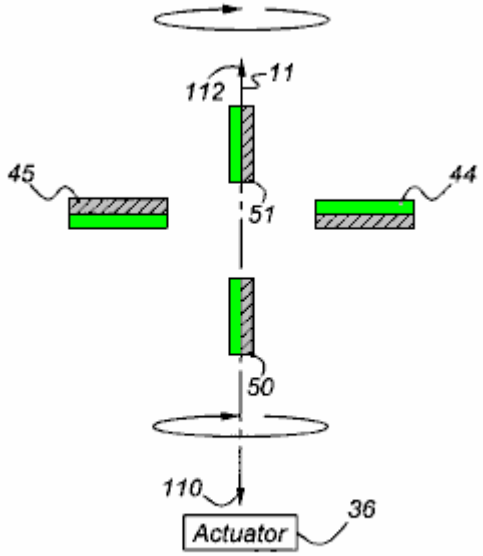


Figure 14

Fig.13 and **Fig.14** show a fourth embodiment of the motor in which each of two reciprocating magnets **50** and **51** is located on an axially opposite side of a rotor magnet pair **44** and **45**. Operation of the motor arranged as shown in **Fig.13** and **Fig.14** is identical to the operation described with reference to **Fig.2** and **Fig.3**.

The direction of the rotational output can be in either angular direction depending on the direction of the starting torque.

The motor can produce reciprocating output on actuator arm **38** instead of the rotational output described above upon disconnecting actuator arm **38** from actuator **36**, and connecting a crank, or a functionally similar device, in the drive path between the actuator and the rotor shaft **10**. The crank converts rotation of the rotor shaft **10** to reciprocation of the actuator **30**. In this case, the rotor shaft **10** is driven rotatably in either direction by the power source, and the output is taken on the reciprocating arm **38**, which remains driveably connected to the oscillating arms **20**, **22** and **24**. The reciprocating magnets **50**, **51** and **52** drive the oscillating arms **20**, **22** and **24**.

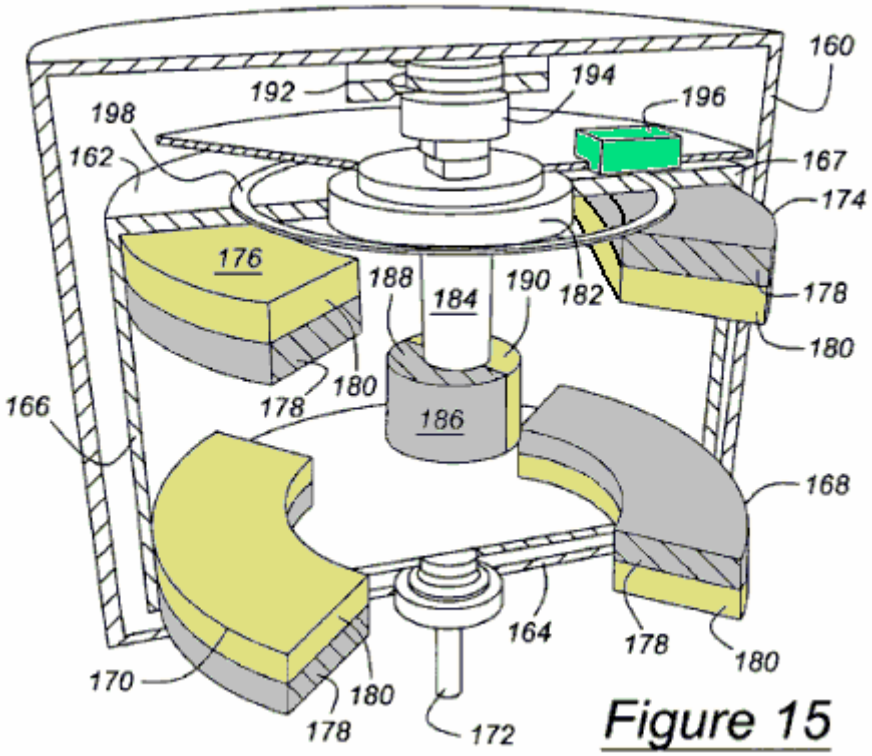


Figure 15

In the perspective cross sectional view shown in **Fig.15**, an outer casing **160** contains a motor according to this invention functioning essentially the same as the embodiment of the more efficient motor shown in **Fig.1A** and **Fig.1B**, but having a commercial appearance. The rotor includes discs **162** and **164** , which are connected by an outer drum **166** of nonmagnetic material. The upper surface **167** of drum **166** forms a magnetic shield surrounding the rotor. Mounted on the lower disc **164** are curved rotor magnets **168** and **170**, which extend angularly about a rotor shaft **172**, which is secured to the rotor. Mounted on the upper disc **162**, are curved rotor magnets **174** and **176**, which extend angularly about the rotor shaft **172**. The reference poles are **178**, and the opposite poles are **180**. A bushing **182** rotates with the rotor.

A reciprocating piston **184**, which moves vertically but does not rotate, supports reciprocating magnet **186**, whose reference pole **188** and opposite pole **190** extend angularly about the axis of piston **184** .

A solenoid magnet **192**, comparable to magnet **90** of the actuator **36** illustrated in **Fig.8**, is located adjacent a solenoid **194**, comparable to solenoid **74** of **Fig.4** and **Fig.5**. The polarity of solenoid **194** alternates as the rotor rotates. Simply stated, as a consequence of the alternating polarity of the solenoid **194**, the reciprocating piston **184** reciprocates which, in turn, continues to advance the rotor more efficiently, using the attraction and repulsion forces between the reciprocating magnets **186** and rotor magnets **168**, 170, **174** and **176** as described above and shown in any of the different embodiments using **Fig.2**, **Fig.3**, **Fig.9**, **Fig.10**, **Fig.11**, **Fig.13** and **Fig.14**. Of course, just as the alternating polarity of the solenoid can put the motor in motion, so can the turning of the rotor, as described above. A photosensor **196** and sensor ring **198** can be used, as an alternative to the mechanical embodiment described in **Fig.4** to **Fig.7**, to determine the angular position of the rotor so as to alternate the polarity of the solenoid **194** with the rotor to correspond with the phase and cycle shown in **Fig.12**.

In accordance with the provisions of the patent statutes, the present invention has been described in what is considered to represent its preferred embodiment. However, it should be noted that the invention can be

constructed otherwise than as specifically illustrated and described without departing from its spirit or scope. It is intended that all such modifications and alterations be included insofar as they come within the scope of the appended claims or the equivalents thereof.

CLAIMS

1. A motor comprising: a rotor supported for rotation about an axis; a first pair of rotor magnets supported on the rotor, including a first rotor magnet and a second rotor magnet spaced angularly about the axis in an opposite radial direction from the first rotor magnet such that the first pair of rotor magnets rotate about the axis along a path having an outermost circumferential perimeter; a first reciprocating magnet supported for movement toward and away from the first and second rotor magnets, the first reciprocating magnet being axially disposed in a first space within a boundary defined by longitudinally extending the outermost circumferential perimeter of the first pair of rotor magnets, and the first reciprocating magnet is a permanent dipole magnet having a reference pole facing laterally from the axis and an opposite pole facing in an opposite lateral direction from the reference pole; and an actuator for moving the first reciprocating magnet cyclically toward and away from the first pair of rotor magnets without passing through a centre of rotation of the first pair of rotor magnets so as to simultaneously create repulsion and attraction forces with the first pair of rotor magnets to cyclically rotate the first pair of rotor magnets relative to the first reciprocating magnet in one rotational direction.

2. The motor of claim 1 further comprising: a second reciprocating magnet axially disposed in a second space within the boundary defined by longitudinally extending the outermost circumferential perimeter of the first pair of rotor magnets at an axial opposite side of the first pair of rotor magnets, and supported for movement toward and away from the first and second rotor magnets without passing through the centre of rotation of the first pair of rotor magnets.

3. The motor of claim 1 further comprising: a second pair of rotor magnets supported on the rotor, spaced axially from the first pair of rotor magnets, the second pair including a third rotor magnet and a fourth rotor magnet spaced angularly about the axis in an opposite radial direction from the third rotor magnet; and wherein the first reciprocating magnet is located in said first space disposed axially between the first and second rotor magnet pairs, and the actuator cyclically moves the first reciprocating magnet toward and away from the first and second pairs of rotor magnets without passing through a centre of rotation of the second pair of rotor magnets.

4. The motor of claim 1 further comprising: a second pair of rotor magnets supported on the rotor, spaced axially from the first pair of rotor magnets, the second pair including a third rotor magnet and a fourth rotor magnet spaced angularly about the axis in an opposite radial direction from the third rotor magnet; a third pair of rotor magnets supported on the rotor, spaced axially from the first and second pairs of rotor magnets, the third pair including a fifth rotor magnet and a sixth rotor magnet spaced angularly about the axis in an opposite radial direction from the fifth rotor magnet; and a second reciprocating magnet disposed in a second space located axially between the second and third rotor magnet pairs and within the boundary defined by longitudinally extending the outermost circumferential perimeter of the first pair of rotor magnets, and the second reciprocating magnet being supported for movement toward and away from the second and third pairs of rotor magnet; and wherein the first reciprocating magnet disposed in the first space is still further located axially between the first and second rotor magnet pairs, and the actuator cyclically moves the first reciprocating magnet toward and away from the first and second pairs of rotor magnets without passing through a centre of rotation of the second pair of rotor magnets, and the second reciprocating magnet toward and away from the second and third pairs of rotor magnets without passing through the centre of rotation of the second pair of rotor magnets and through a centre of rotation of a third pair of rotor magnets.

5. The motor of claim 1 further comprising: an arm supported for pivotal oscillation substantially parallel to the axis, the first reciprocating magnet being supported on the arm adjacent the first and second rotor magnets; and wherein the actuator is driveably connected to the arm.

6. The motor of claim 1 wherein: the first and second rotor magnets are permanent dipole magnets, the first rotor magnet having a reference pole facing axially away from the first reciprocating magnet and an opposite pole facing axially toward the first reciprocating magnet, the second rotor magnet having a reference pole facing axially toward the first reciprocating magnet and an opposite pole facing axially away from the first reciprocating magnet.

7. The motor of claim 1 wherein: the first and second rotor magnets are magnet is a permanent dipole magnets magnet, the first rotor magnet having a reference pole facing axially away from the first reciprocating magnet and an opposite pole facing axially toward the first reciprocating magnet, the second rotor magnet having a reference pole facing axially toward the first reciprocating magnet and an opposite pole facing axially away from the first reciprocating magnet; and the motor further comprising: a second pair of rotor magnets supported on the rotor, spaced axially from the first pair of rotor magnets, the second pair including a third

permanent dipole rotor magnet having a reference pole facing axially toward the first reciprocating magnet and an opposite pole facing away from the first reciprocating magnet, and a fourth permanent dipole rotor magnet spaced angularly about the axis in an opposite radial direction from the third rotor magnet, the fourth permanent dipole rotor magnet having a reference pole facing axially away from the first reciprocating magnet and an opposite pole facing toward the first reciprocating magnet; and wherein the first reciprocating magnet disposed in said first space is still further located axially between the first and second rotor magnet pairs, and the actuator cyclically moves the first reciprocating magnet toward and away from the first and second pairs of rotor magnets without passing through a centre of rotation of the second pair of rotor magnets.

8. The motor of claim 1 wherein: the first and second rotor magnets are permanent dipole magnets, each rotor magnet having a reference pole facing in a first lateral direction relative to the reference pole of the first reciprocating magnet and an opposite pole facing in a second lateral direction opposite the first lateral direction of the respective rotor magnet.

9. The motor of claim 1 wherein: the first and second rotor magnets are permanent dipole magnets, each rotor magnet having a reference pole facing in a first lateral direction relative to the reference pole of the first reciprocating magnet and an opposite pole facing in a second lateral direction opposite the first lateral direction of the respective rotor magnet; and the motor further comprising: a second pair of rotor magnets supported for rotation on the rotor about the axis, the second pair of rotor magnets being spaced axially from the first pair of rotor magnets, the second pair including a third permanent dipole rotor magnet and a fourth permanent dipole rotor magnet, the third and fourth rotor magnets each having a reference pole facing in the second lateral direction and an opposite pole facing in the first lateral direction, and wherein the first reciprocating magnet disposed in the first space is still further located axially between the first and second rotor magnet pairs, and the actuator cyclically moves the first reciprocating magnet toward and away from the first and second pairs of rotor magnets without passing through a centre of rotation of the second pair of rotor magnets.

10. The motor of claim 3 further comprising: a third pair of rotor magnets supported on the rotor, spaced axially from the first and second pairs of rotor magnets, the third pair including a fifth rotor magnet and a sixth rotor magnet spaced angularly about the axis in an opposite radial direction from the fifth rotor magnet; a second reciprocating magnet located in a second space within the boundary defined by longitudinally extending the outermost circumferential perimeter of the first pair of rotor magnets and axially between the second and third rotor magnet pairs, and the second reciprocating magnet being supported for movement toward and away from the second and third pairs of rotor magnet; a first arm supported for pivotal oscillation substantially parallel to the axis, the first reciprocating magnet being supported on the arm adjacent the first and second pairs of rotor magnets; and a second arm supported for pivotal oscillation substantially parallel to the axis, the second reciprocating magnet being supported on the arm adjacent the second and third pairs of rotor magnets; and wherein the actuator is driveably connected to the first and second arms.

11. A motor comprising: a rotor supported for rotation about an axis; a first pair of rotor magnets supported on the rotor, including a first rotor magnet and a second rotor magnet spaced angularly about the axis from the first rotor magnet such that the first pair of rotor magnets rotate about the axis along a circumferential path having an outermost perimeter; a first arm supported for pivotal oscillation along the axis, located adjacent the first and second rotor magnets; a first reciprocating magnet, supported on the first arm for movement toward and away from the first and second rotor magnets, the first reciprocating magnet being disposed axially within a first space within a boundary defined by longitudinally extending the outermost perimeter of the first circumferential path of the first pair of rotor magnets; a second pair of rotor magnets supported on the rotor, spaced axially from the first pair of rotor magnets, the second pair including a third rotor magnet, and a fourth rotor magnet spaced angularly about the axis from the third rotor magnet; a third pair of rotor magnets supported on the rotor, spaced axially from the first and second pairs of rotor magnets, the third pair including a fifth rotor magnet, and a sixth rotor magnet spaced angularly about the axis from the fifth rotor magnet; a second arm supported for pivotal oscillation along the axis between the second and third pairs of rotor magnets; a second reciprocating magnet located axially between the second and third rotor magnet pairs and supported on the second arm for movement toward and away from the second and third pairs of rotor magnet; and an actuator for moving the first reciprocating magnet cyclically toward and away from the first pair of rotor magnets without passing through a centre of rotation of the first pair of rotor magnets so as to simultaneously create repulsion and attraction forces with the first pair of rotor magnets to cyclically rotate the first pair of rotor magnets relative to the first reciprocating magnet in one rotational direction; and wherein the first reciprocating magnet disposed in the first space is still further located axially between the first and second rotor magnet pairs, and the actuator cyclically moves the first arm and first reciprocating magnet toward and away from the first and second pairs of rotor magnets without passing the first reciprocator magnet through a centre of rotation of the second pair of rotor magnets, and moves the second arm and second reciprocating magnet toward and away from the second and third pairs of rotor magnets without passing the second reciprocator magnet through the centre of rotation of the second pair of rotor magnets and through a centre of rotation of the third pair of rotor magnets.

12. The motor of claim 11 wherein the actuator further comprises: a rotor shaft driveably connected to the rotor for rotation therewith; first and second bridge plates, mutually angularly aligned about the axis, extending over a first angular range about the axis; third and fourth bridge plates, offset axially from the first and second bridge plates, mutually angularly aligned about the axis, extending over a second angular range about the axis; an electric power supply including first and second terminals; a first contact connecting the first power supply terminal alternately to the first bridge plate and the third bridge plate as the rotor rotates; a second contact connecting the second power supply terminal alternately to the second bridge plate and the fourth bridge plate as the rotor rotates; a toroidal permanent magnet; a solenoid supported above a pole of the toroidal permanent magnet, including first and second terminals; a third contact connecting the first solenoid terminal alternately to the first and second power supply terminals through the first and fourth bridge plates and first contact as the rotor rotates; a fourth contact alternately connecting and disconnecting the second power supply terminal and the second solenoid terminal as the rotor rotates; and a fifth contact alternately connecting and disconnecting the first power supply terminal and the second solenoid terminal as the rotor rotates.

13. The motor of claim 11 wherein the actuator further comprises: a toroidal permanent magnet; an A.C. power source; and a solenoid supported for displacement adjacent a pole of the toroidal permanent magnet, including first and second terminals electrically connected to the power source.

14. A motor comprising: a rotor supported for rotation about an axis; a first rotor magnet supported for rotation about the axis along a first circumferential path having an outermost perimeter and a centre at the axis, the first rotor magnet having a first permanent reference pole facing laterally toward the axis and a first permanent opposite pole facing in an opposite lateral direction toward the first reference pole; a pair of reciprocating magnets supported for movement toward and away from the rotor magnet, including a first reciprocating magnet and a second reciprocating magnet spaced axially from the first rotor magnet, each reciprocating magnet being at least partially disposed within a first axial space having a boundary defined by longitudinally extending the outermost perimeter of the first circumferential path of the first rotor magnet, wherein the rotor magnet is located axially between the first and second reciprocating magnets; and an actuator for moving the pair of reciprocating magnets cyclically toward and away from the rotor magnet without passing through the centre of the first circumferential path so as to simultaneously create repulsion and attraction forces with the first rotor magnet to cyclically rotate the rotor magnet relative to the pair of reciprocating magnets in one rotational direction.

15. The motor of claim 14 wherein the first and second reciprocating magnets are permanent dipole magnets with each having a reference pole facing laterally from the axis and an opposite pole facing in an opposite lateral direction from its corresponding reference pole.

16. The motor of claim 15 further comprising: a second rotor magnet spaced axially from the first rotor magnet, the second rotor magnet being supported for rotation about the axis along a second circumferential path having an outermost perimeter about the centre, the second rotor magnet including a second permanent reference pole facing laterally toward the axis and a second permanent opposite pole facing in an opposite lateral direction toward the second reference pole; and wherein the second reciprocating magnet is located axially between the first and second rotor magnets and at least partially within a second axial space having a boundary defined by longitudinally extending the outermost perimeter of the second circumferential path of the second rotor magnet, and the actuator cyclically moves the second reciprocating magnet away from and towards the second rotor magnet.

The Magnetic Motor of Charles Flynn

US Patent 5,455,474 **3rd October 1995** **Inventor: Charles Flynn**

MAGNETIC MOTOR CONSTRUCTION

This patent gives details of a permanent magnet motor which uses electromagnet shielding to achieve continuous rotation. The input power is very small with even a 9-volt battery being able to operate the motor. The output power is substantial and operation up to 20,000 rpm is possible. Construction is also very simple and well within the capabilities of the average handyman. It should be realised that the power of this motor comes from the permanent magnets and not from the small battery input used to prevent lock-up of the magnetic fields.

ABSTRACT

The present invention is a motor with permanent magnets positioned so that there is magnetic interaction between them. A coil placed in the space between the permanent magnets is used to control the magnetic interaction. This coil is connected to a source of electric potential and controlled switching so that closing the switch places a voltage across the coil and affects the magnetic interaction between the permanent magnets as to produce rotational movement of the output shaft.

US Patent References:

3096467	Brushless d. c. motor with permanent magnet rotor	July, 1963 Angus et al.	318/138
3569806	Starting Arrangement for Solid-State Motor	March, 1971 Brailsford	318/254
3670189	Gated Permanent Magnet Motor	June, 1972 Monroe	310/181
3796039	Electric Micromotor	March, 1974 Lucien	310/268
3883633	Commutatorless Motor	May, 1975 Kohler	310/152
4151431	Permanent Magnet Motor	April, 1979 Johnson	310/12
4187441	High-power-density Brushless DC Motor	February, 1980 Oney	310/112
4758756	Vernier-type Electrodynamic Machine	July, 1988 Pouillange	310/152
4875110	Rotary-head Apparatus with Motor Magnet	October, 1989 Kazama	310/268
4972112	Brushless DC Motor	November, 1990 Kim	310/181
5179307	Direct Current Brushless Motor	January, 1993 Porter	310/268

Foreign References:

DE210005	July, 1960	310/181
JP0025153	February, 1982	310/181
JP01521078	September, 1982	310/152
JP0002840	January, 1987	310/152

BACKGROUND OF THE INVENTION

The present invention is an improvement over the inventions disclosed in patent applications 07/322,121 and 07/828,703. The devices disclosed in those applications relate to means to produce useful energy using permanent magnets as the driving source. This is also true of the present invention which represents an important improvement over the known constructions and one which is simpler to construct, can be made to be self starting, is easier to adjust, and is less likely to get out of adjustment. The present construction is also relatively easy to control, is relatively stable and produces an amazing amount of output energy considering the source of driving energy that is used. The present construction makes use of permanent magnets as the source of driving energy but shows a novel means of controlling the magnetic interaction between the magnet members in a manner which is relatively rugged, produces a substantial amount of output energy and torque, and in a device capable of being used to generate substantial amounts of energy that is useful for many different purposes.

The present invention resides has a fixed support structure with one or more fixed permanent magnets such as an annular permanent magnet mounted on it with the pole faces of the permanent magnet on opposite faces of the magnet. The device has one or more relatively flat coils positioned around the edge of one of the faces of the magnet, and a shaft extends through the permanent magnet with one or more other permanent magnets attached to it. The spaced permanent magnets and the fixed permanent magnet have their polarities arranged to produce a magnetic interaction between them. The device also includes a circuit for selectively and sequentially energising the coils to control the magnetic interaction between the magnets in such a manner as to produce

rotation between them. Various methods can be used to control the application of energy to the coils including a timer or a control mechanism mounted on the rotating shaft. This design can be made to be self-starting or to be started with some initial help to establish rotation.

OBJECTS OF THE INVENTION

It is a principal object of the present invention to teach the construction and operation of a relatively simple, motor-like device using permanent magnets in an unique manner to generate rotational or other forms of movement.

Another object is to teach the construction and operation of a relatively simple, motor-like device having novel means for coupling and/or decoupling relatively moveable permanent magnets to produce motion.

Another object is to provide novel means for controlling the coupling and decoupling of relatively moveable permanent magnets.

Another object is to make the generation of rotational energy less expensive and more reliable.

Another object is to teach a novel way of generating energy by varying magnetic interaction forces between permanent magnets.

Another object is to provide an inexpensive way of producing energy.

Another object is to provide a substitute source of energy for use in places where conventional motors, generators and engines are used.

These and other objects and advantages of the present invention will become apparent after considering the following detailed specification of preferred embodiments in conjunction with the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

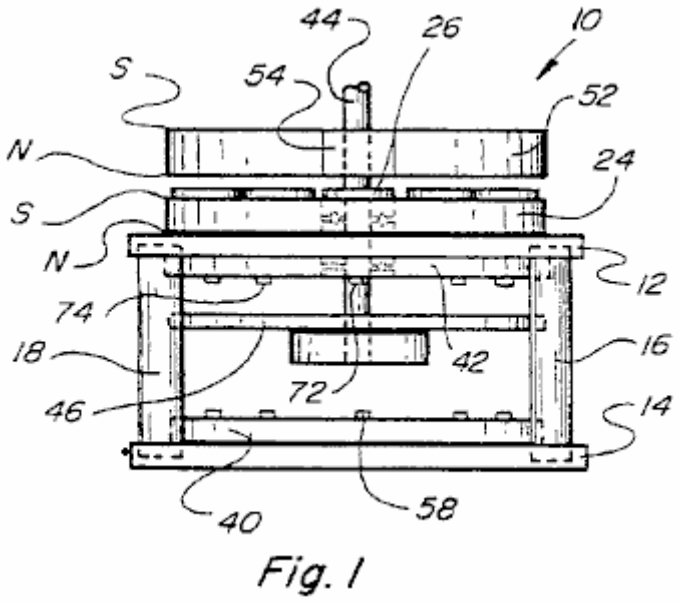


Fig. 1

Fig.1 is a side view of a magnetically powered device constructed according to the present invention.

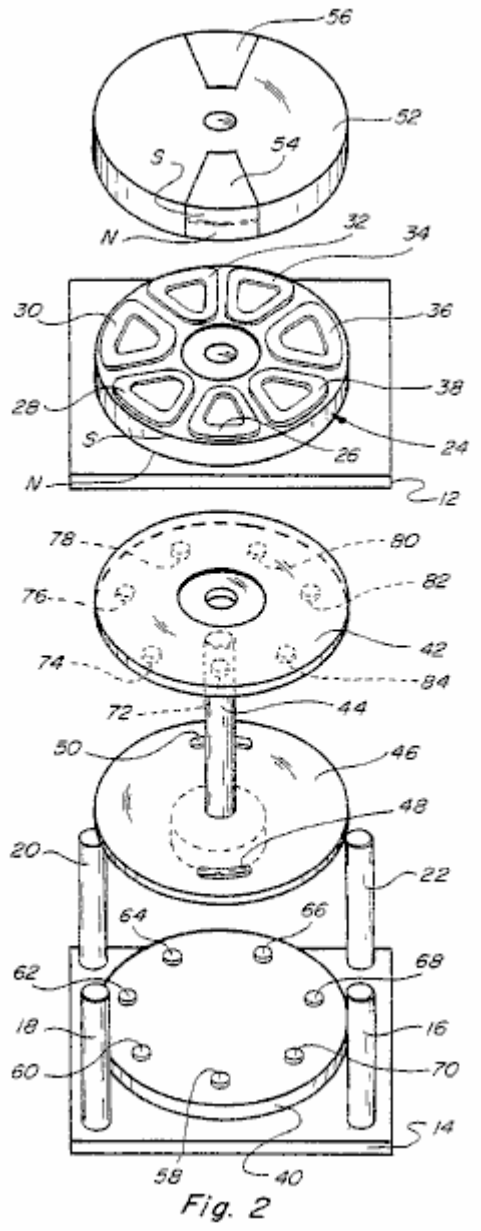


Fig. 2

Fig.2 is an exploded view of the device shown in Fig.1.

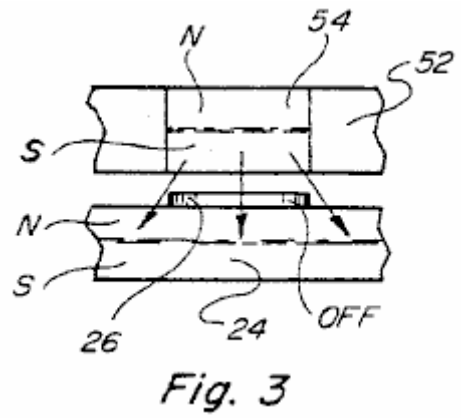


Fig. 3

Fig.3 is a fragmentary side view of one of the movable magnets and the fixed magnet, in one position of the device.

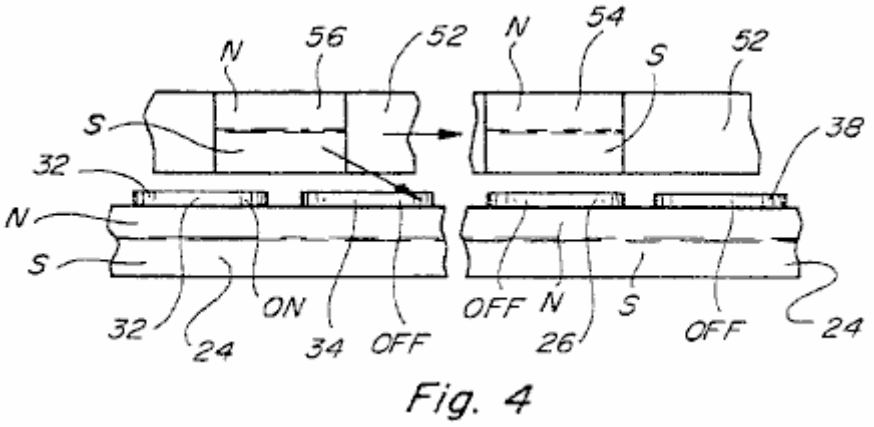


Fig. 4

Fig.4 is a view similar to Fig.3 but showing the relationship between the other movable magnets and the fixed magnet in the same rotational position of the device.

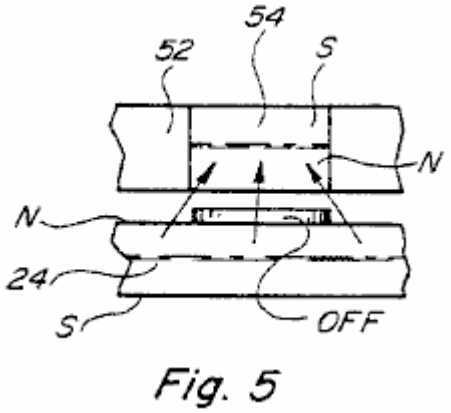


Fig. 5

Fig.5 is a fragmentary view similar to Fig.3 but showing a repulsion interaction between the relatively movable permanent magnets.

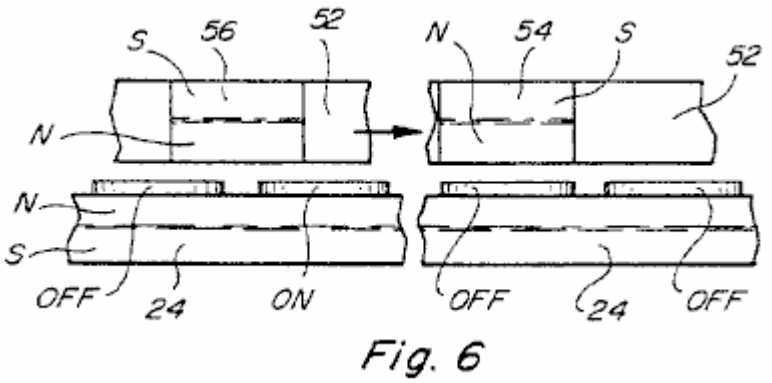


Fig. 6

Fig.6 is a view similar to Fig.4 for the condition shown in Fig.5.

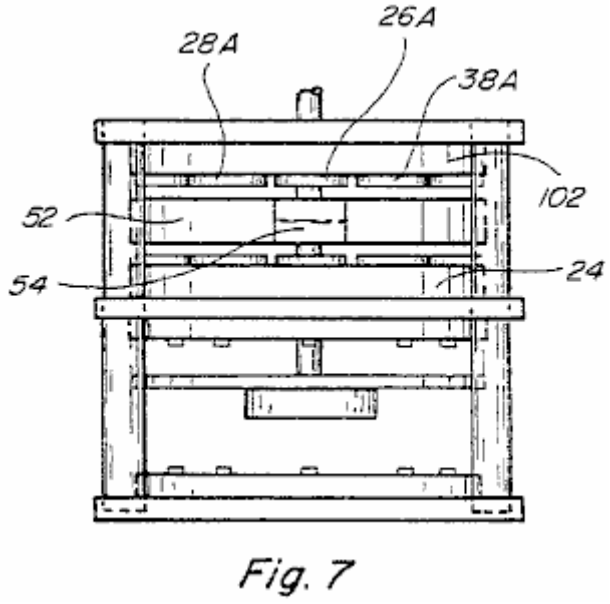


Fig. 7

Fig.7 is a side view showing another embodiment which is capable of producing even greater energy and torque.

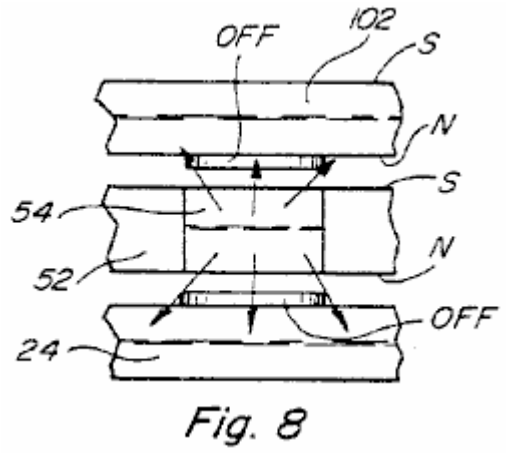


Fig. 8

Fig.8 is a fragmentary elevational view similar to Fig.3 for the device of Fig7.

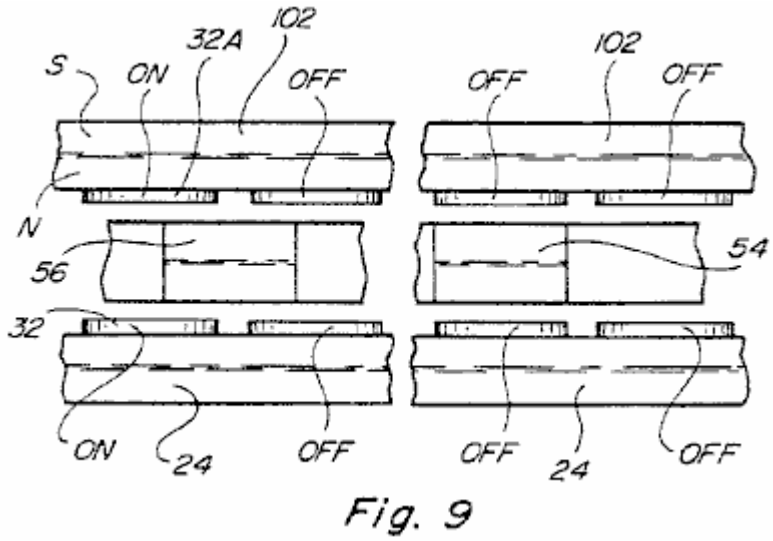


Fig. 9

Fig.9 is a view similar to Fig.4 for the construction shown in Fig.7.

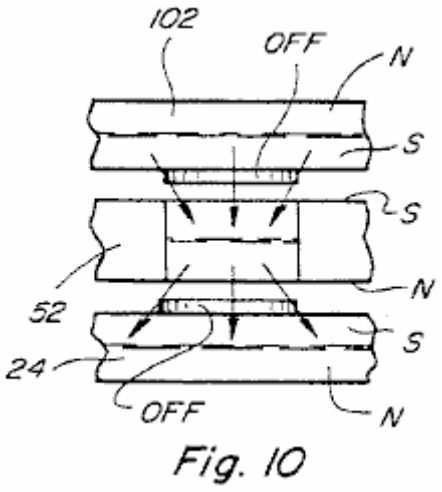


Fig. 10

Fig.10 is a view similar to Fig.3 for the device shown in Fig.7 but with the polarity of one of the fixed permanent magnets reversed.

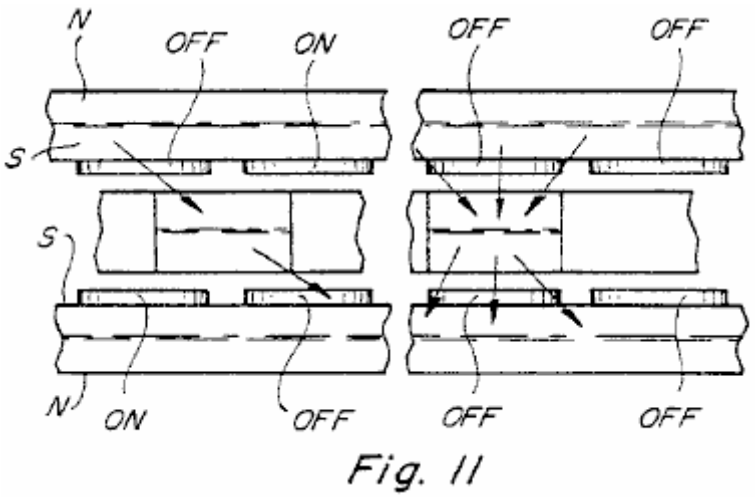


Fig. 11

Fig.11 is a fragmentary view similar to Fig.4 for the device as shown in Fig.7 and Fig.10.

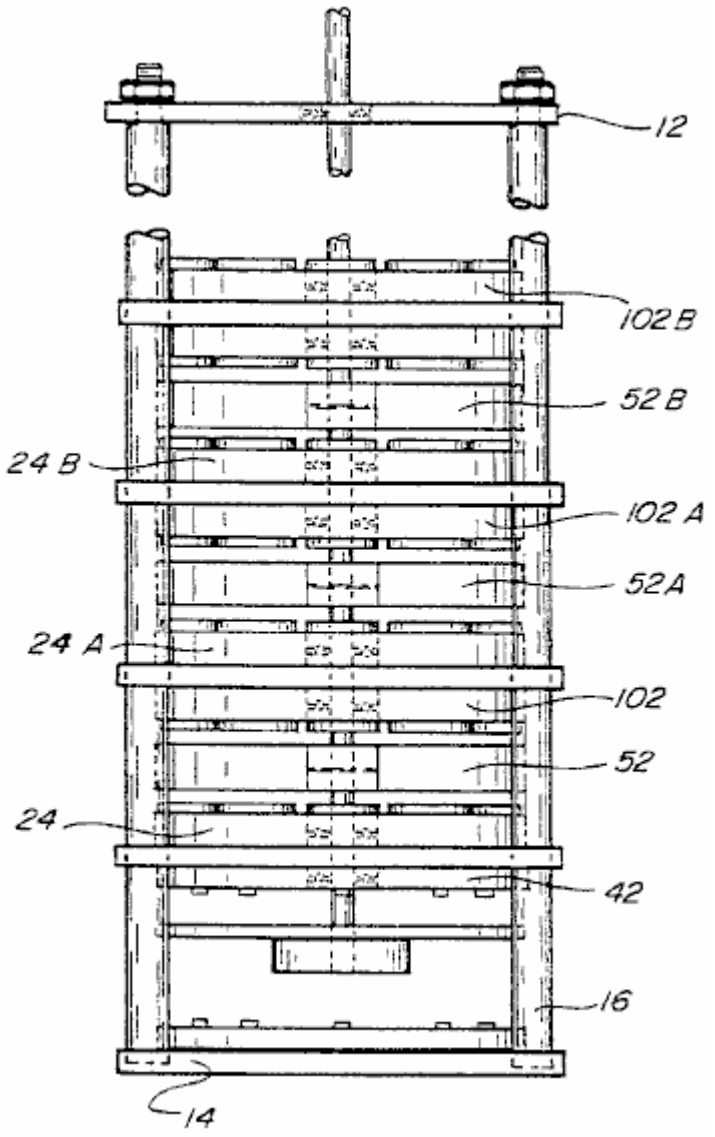


Fig. 12

Fig.12 is a side elevational view of another embodiment of the device.

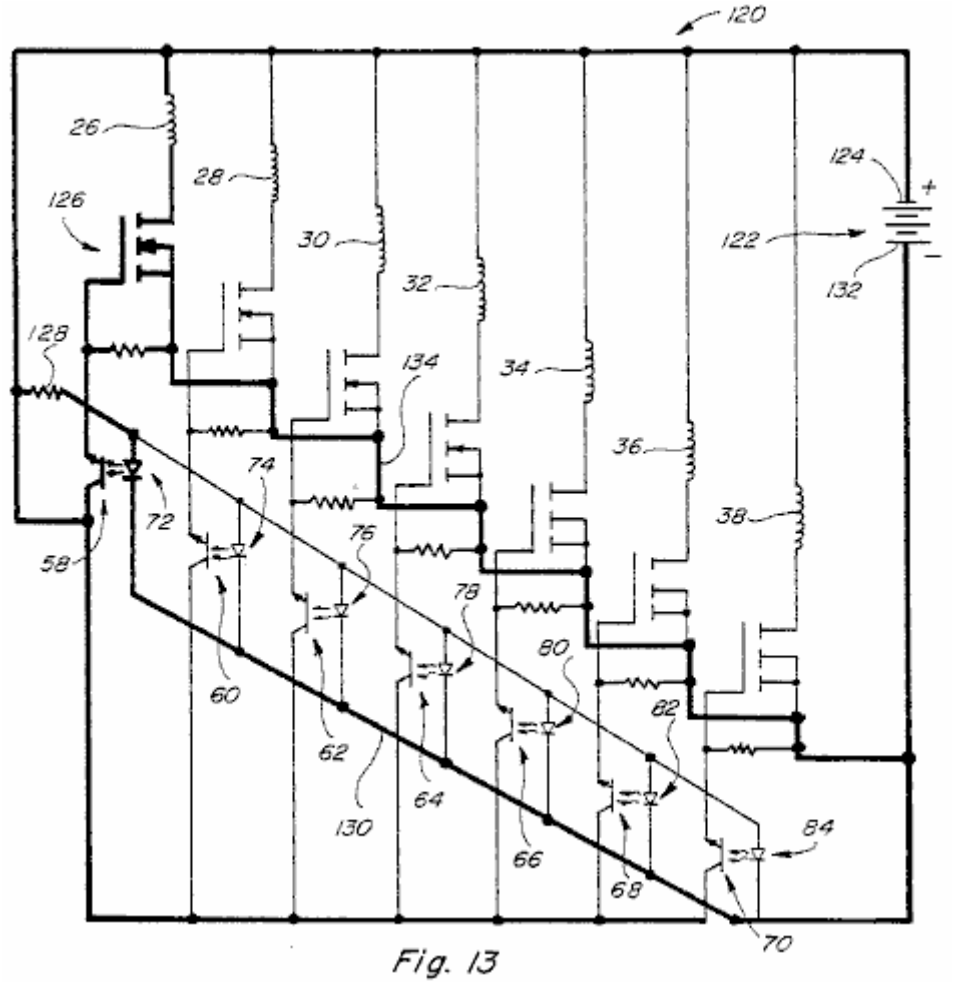


Fig. 13

Fig.13 is a schematic circuit diagram of the circuit for the devices of Figs. 1, 7 and 12.

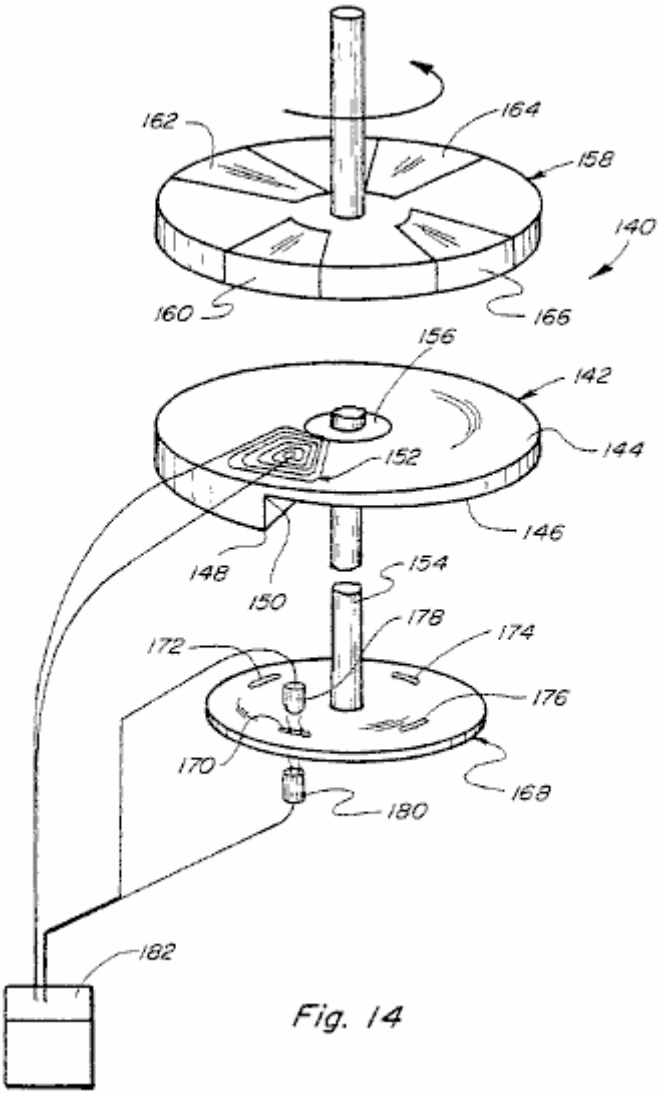


Fig. 14

Fig.14 is a perspective view of another embodiment.

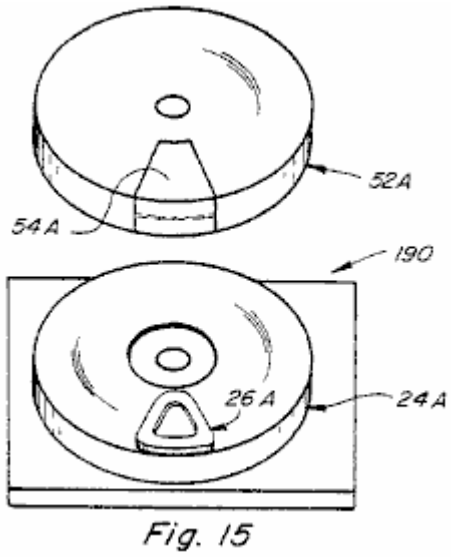


Fig. 15

Fig.15 is a simplified embodiment of the device showing the use of one rotating magnet and one coil positioned in the plane between the rotating and stationary magnets.

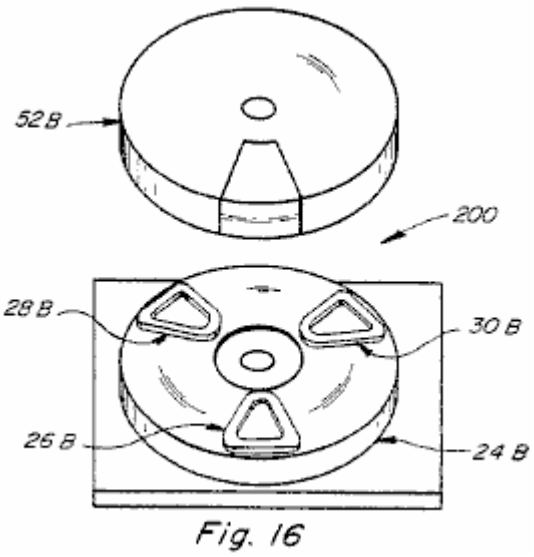


Fig. 16

Fig.16 is a simplified embodiment of the device showing use of one movable magnet and three coils arranged to be in a plane between the rotating and stationary magnets.

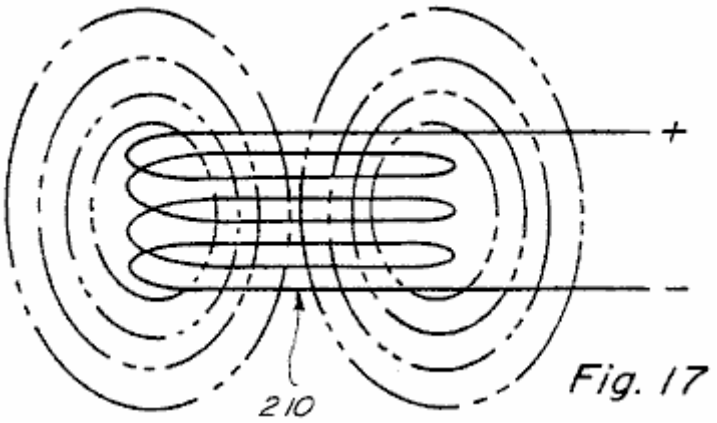


Fig. 17

Fig.17 is a side view of an air coil with a voltage applied across it and showing in dotted outline the field of the coil.

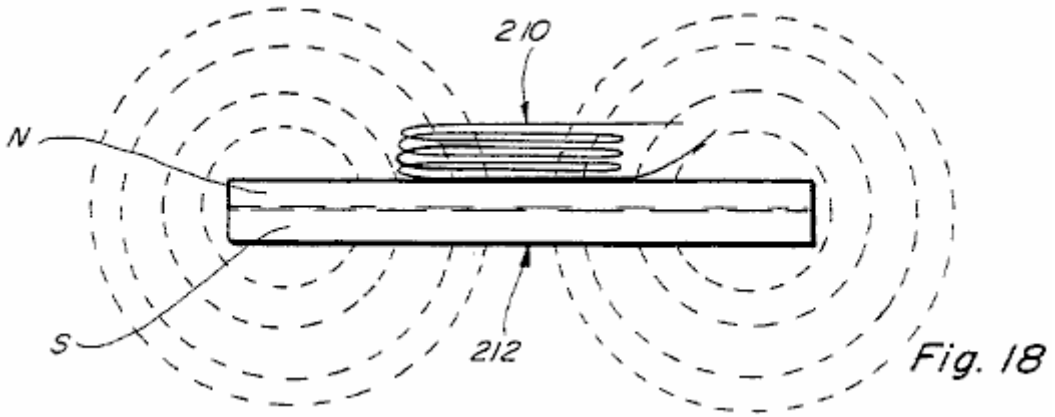


Fig. 18

Fig.18 is a view similar to Fig.17 but showing the air coil positioned adjacent to one side of a permanent magnet showing in dotted outline the magnetic field of the permanent magnet with no electric potential applied across the air coil.

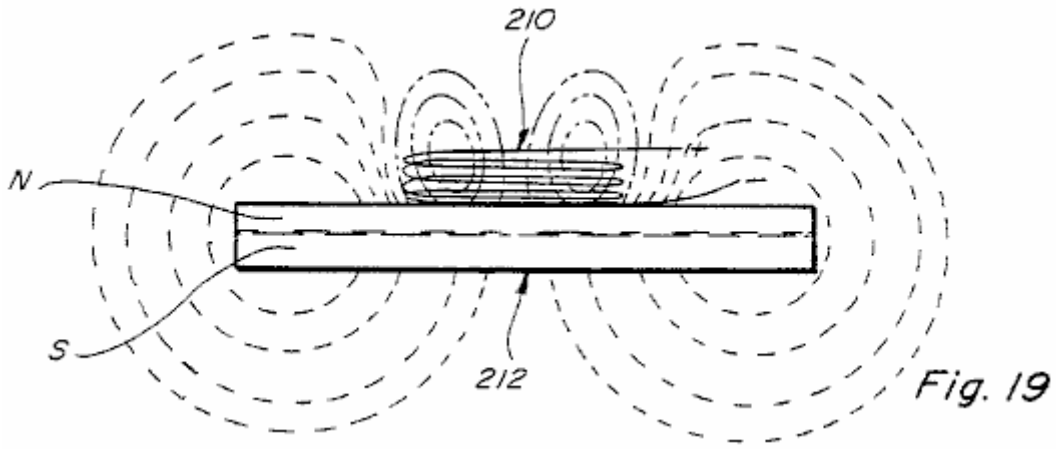


Fig.19 is a side view similar to Fig.18 with an electric potential applied across the air coil, showing in dotted outline the shapes of the electric field of the air coil and the magnetic field of the permanent magnet.

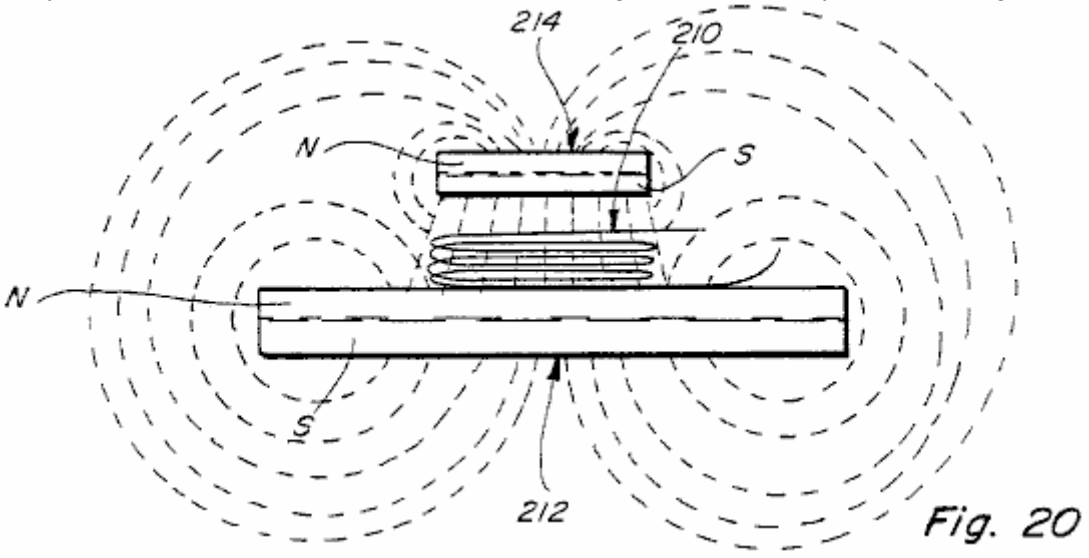


Fig.20 is a side view similar to Fig.19 but showing a second permanent magnet positioned above the first permanent magnet and showing in dotted outline the magnetic fields of the two permanent magnets when no electric potential is connected across the air coil.

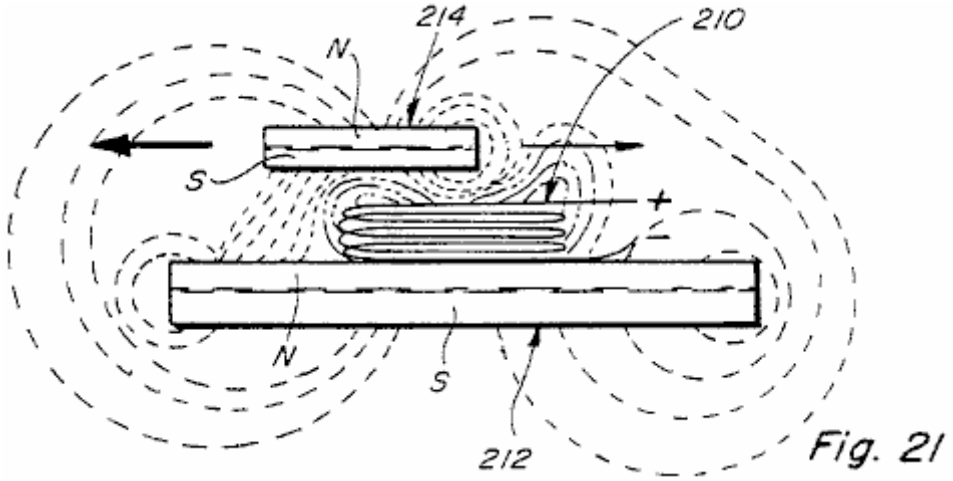


Fig.21 is a view similar to Fig.20 but with the permanent magnets in an different relative position and with a voltage applied across the air coil, said view showing the shapes of the electro-magnetic field of the air coil and the modified shapes of the magnetic fields of the two permanent magnets; and

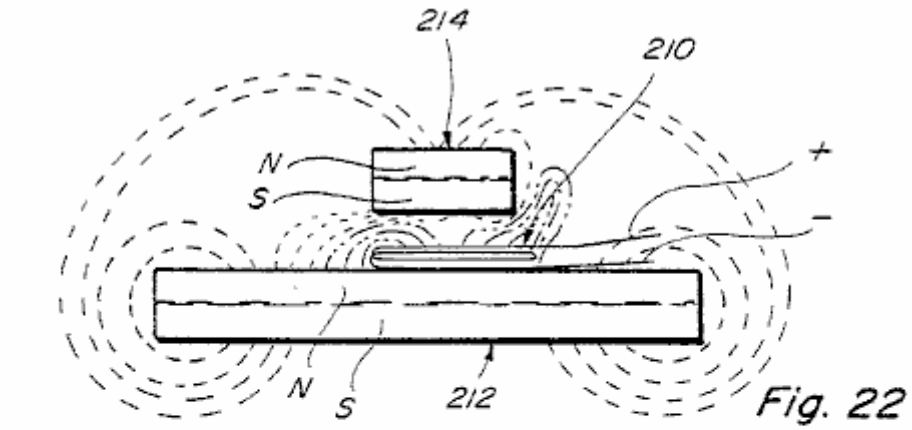


Fig. 22

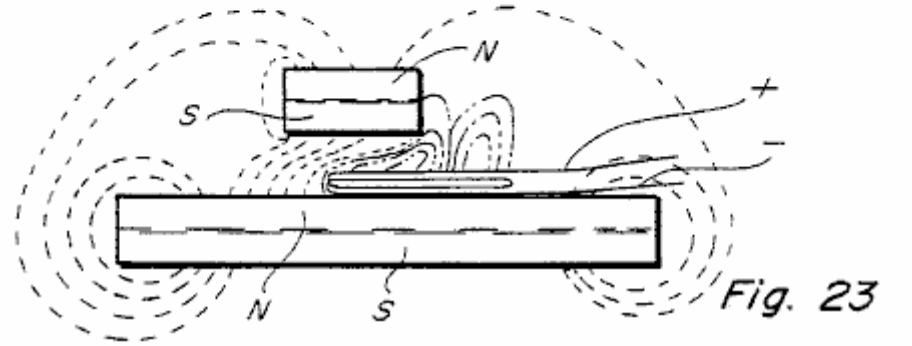


Fig. 23

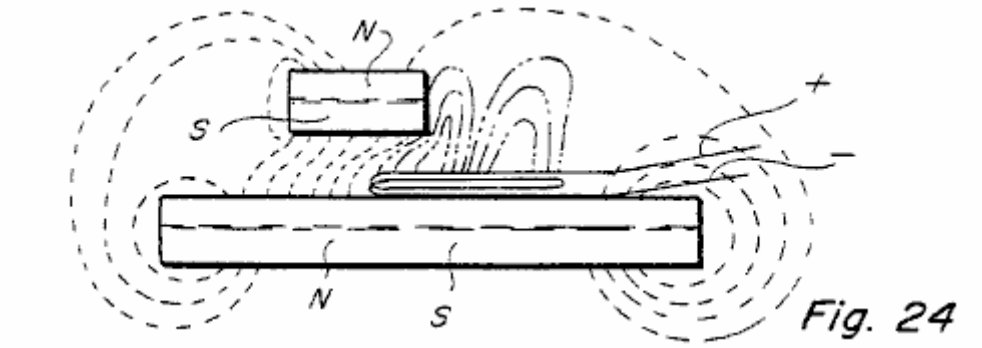


Fig. 24

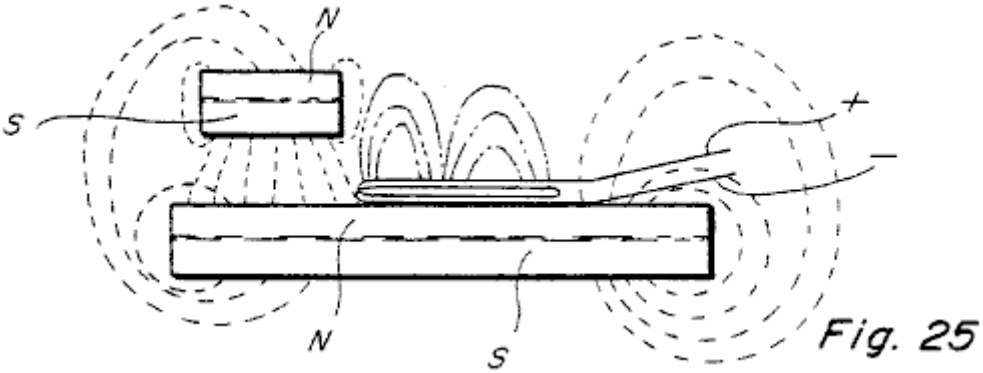


Fig. 25

Fig.22 to **Fig.25** are similar to Fig.21 and show the electro-magnetic field of the air coil and the magnetic fields of the magnets in four different relative positions of the permanent magnets.

DETAILED DESCRIPTION

In the drawings, the number **10** refers to a device constructed according to the present invention. The device **10** includes a stationary base structure including an upper plate **12**, a lower plate **14**, and spaced posts **16-22** connected between them.

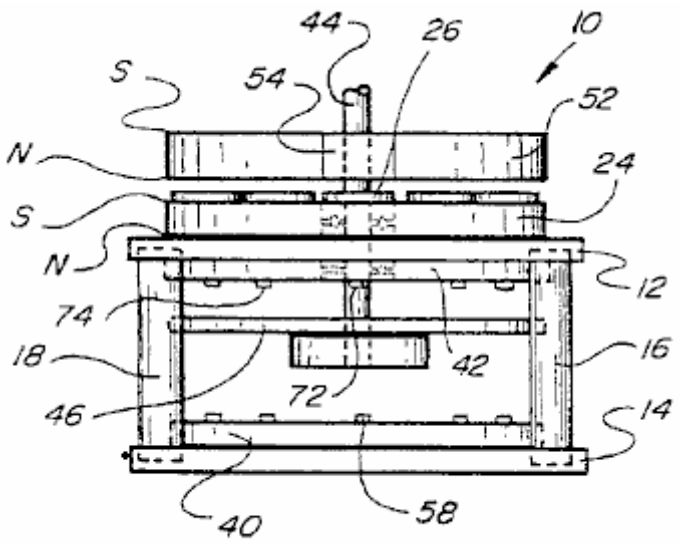


Fig. 1

Mounted on the upper plate **12** is a fixed permanent magnet **24** shown annular in shape which has its North pole adjacent to the upper surface of plate **12** and its South pole facing away from plate **12**.

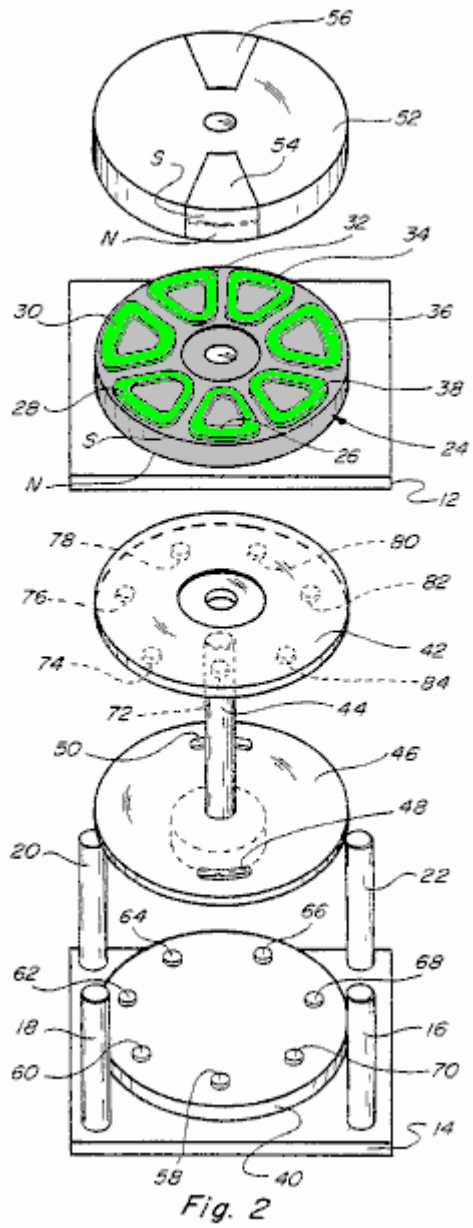


Fig. 2

Referring to **Fig.2**, the permanent magnet **24** is shown having seven coils **26-38** mounted flat on its upper surface. Seven coils are shown, and the coils **26-38** have electrical connections made through plate **12** to other circuit members which will be described later in connection with **Fig.13**. Another member **40** is mounted on the upper surface of the lower plate **14** and a similar member **42** is mounted on the underside of the plate **12**.

A shaft **44**, (shown oriented vertically for convenience) extends through aligned holes in the members **42**, **12** and **24**. The lower end of shaft **44** is connected to disk **46** which has a pair of curved openings **48** and **50** shown diametrically opposite to each other, a little in from the edge of disc **46**. The purpose of these openings **48** and **50** will be explained later on.

Shaft **44** is also connected to another disc **52** which is located on the shaft so as to be positioned adjacent to the coils **26-38**. Disc **52** has a pair of permanent magnets **54** and **56** mounted on or in it positioned diametrically opposite to each other. Magnets **54** and **56** have their north and south poles oriented as shown in **Fig.2**, that is with north poles shown on their lower sides and their south poles on the upper sides. This is done so that there will be mutual magnetic attraction and coupling between the magnets **54** and **56** and the fixed magnet **24**. The polarity of the magnets **54** and **56** and/or of the magnet **24** can also be reversed if desired for some purposes to produce relative magnetic repulsion between them.

Referring again to **Fig.2**, the lower plate **40** is shown having a series of phototransistors **58-70** mounted on its upper surface and spaced out as shown. These phototransistors are positioned under the centres of the coils **26-38** which are mounted on magnet **24**. An equal number of infra red emitters **72-84** are mounted on the under surface of the member **42** aligned with the phototransistors. There are seven infra red emitters **72-84** shown, each of which is in alignment with a respective one of the seven phototransistors **58-70** and with one of the seven coils **26-38**. This arrangement is such that when the shaft **44** and the components attached to it, including discs **46** and **52**, rotate relative to the other members including magnet *24*, the curved openings **48** and **50** pass under the infra red emitters and cause the phototransistors to switch on for a predetermined time interval. This establishes a sequence of energised circuits which powers coils **26-38**, one at a time, which in turn, causes a momentary interruption of the magnetic interaction between one of the permanent magnets **54** and **56** and magnet **24**.

When a coil is mounted on top of a permanent magnet such as permanent magnet **24** and energised it acts to concentrate the flux in a symmetrical magnetic field resulting in a non-symmetrical field when another permanent magnet is above the coil on magnet **24**. This results in uneven or non-uniform forces being produced when the coil is energised and this causes a torque between the two permanent magnets, which tries to move one of the permanent magnets relative to the other.

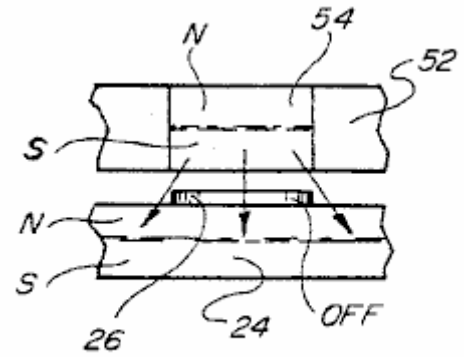


Fig. 3

Fig.3 shows the position when one of the magnets **54** is located immediately above one of the coils, say, coil **26**. In this position there would be magnetic coupling between the magnets **54** and **24** so long as there is no voltage across the coil **26**. However, if a voltage is placed across the coil **26** it will interrupt the magnetic coupling between the magnets **54** and **24** where the coil is located. This means that if there is any torque developed, it will be developed to either side of the coil **26**. Without energising the coil **26** there will be full attraction between the magnets **24** and **54** and no rotational force will be produced.

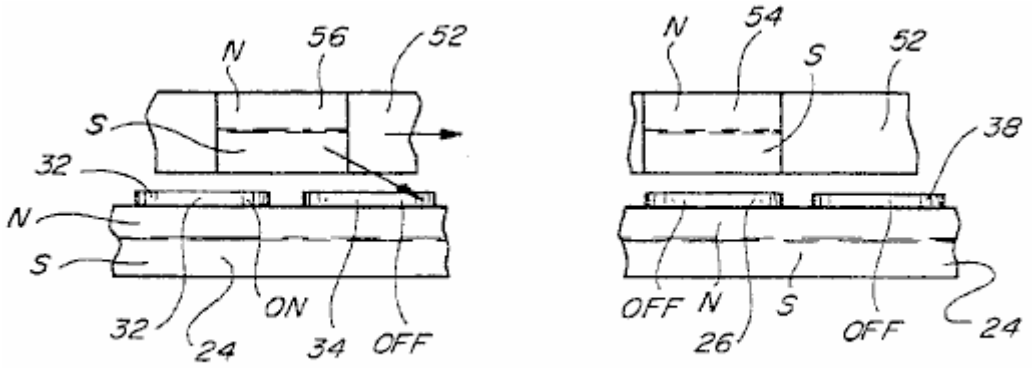


Fig. 4

Referring to **Fig.4** there is shown the relative positions of the movable magnets **54** and **56** for one position of disc **52**. For example, the magnet **54** is shown located immediately above the coil **26** while the magnet **56** is shown straddling portions of the coils **32** and **34**. If, in this position, coil **32** is energised but coils **34** and **26** are not energised, then the magnetic coupling between magnet **56** and magnet **24** will be oriented at an angle shown illustrated by the arrow in **Fig.4**, and this attractive coupling will tend to move disc **52** to the right. Since coil **26** is not powered up, there is full coupling between magnet **54** and magnet **24** but this has no effect since it does not have a directional force. At the same time, coil **38** which is the next coil over which the magnet **54** will move, is also not powered up and so it will have no rotational effect on disc **52**.

As disc **52** continues to rotate, different coils in the group **26-38** will be energised in sequence to continue to produce a rotational magnetic coupling force between disc **52** and magnet **24**. It should be noted, however, that all of the rotational force is produced by interaction between the permanent magnets and none of the rotational force is produced by the coils or by any other means. The coils are merely energised in sequence to control where the magnetic interaction occurs, and this is done in a manner to cause disc **52** to rotate. It should also be understood that one, two, or more than two, permanent magnets such as the permanent magnets **54** and **56** can be mounted on the rotating disc **52**, and the shape and size of the rotating disc **52** can be adjusted accordingly to accommodate the number of permanent magnets mounted in it. Also, disc **52** can be constructed of a non-magnetic material, the only requirement being that sufficient structure be provided to support the permanent magnets during rotation. This means that disc **52** need not necessarily be constructed to be round as shown in the drawing.

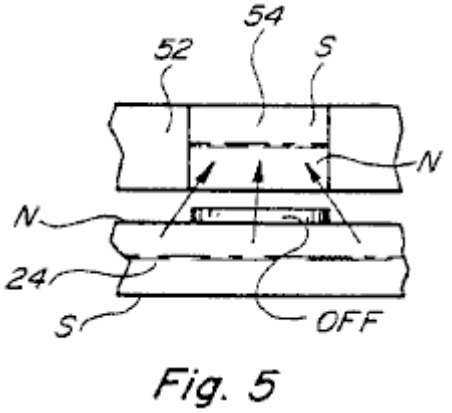


Fig. 5

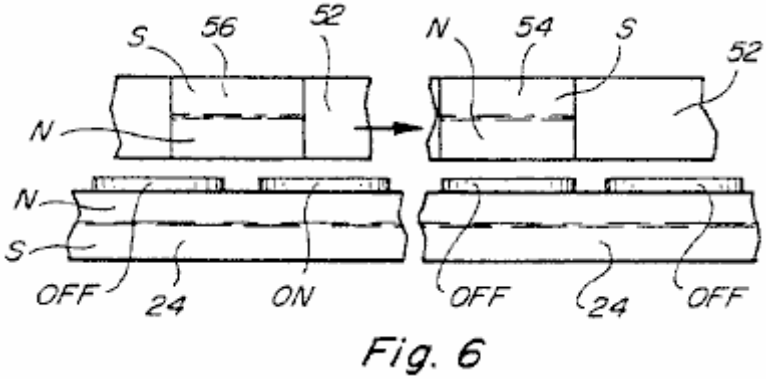


Fig. 6

Fig.5 and **Fig.6** are similar to **Fig.3** and **Fig.4** but show a construction where the permanent magnets **54** and **56** are turned over so that instead of having their north poles facing magnet **24** they have their south poles facing magnet **24** but on the opposite side of the coils such as coils **26-38**. The construction and operation of the modified device illustrated by **Fig.5** and **Fig.6** is similar to that described above except that instead of producing magnetic attraction forces between the magnets **54** and **56** and the magnet **24**, magnetic repulsion forces are produced, and these repulsion forces can likewise be used in a similar manner to produce rotation of the member **52**, whatever its construction.

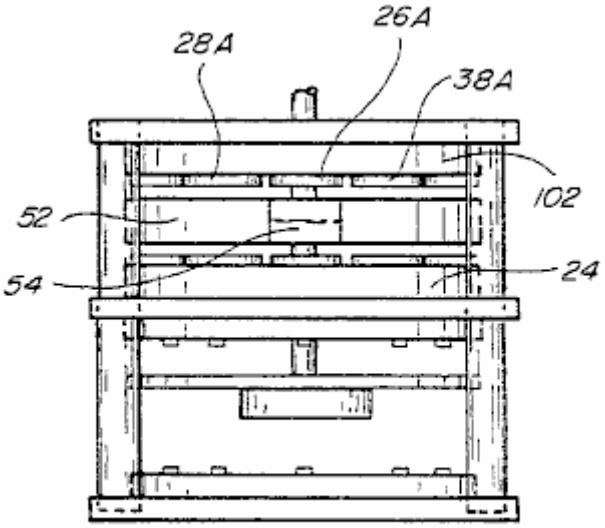


Fig. 7

Fig.7 shows a modified embodiment which includes all of the elements shown in **Fig.1** and **Fig.2** but in addition has a second stationary permanent magnet **102** which is mounted above rotating disc **52** and has its coil members such as coil members **26A-38A** mounted on its underside. Magnet **102** operates with the magnets **54** and **56** similarly to the magnet **24** and can operate in precisely the same manner, that is by producing attraction force between the magnet members or by producing repulsion forces between them, each being used to produce relative rotational movement between the rotor and the stator. It is also contemplated to make the construction shown in **Fig.7** so as to produce attraction forces between the magnets **54** and **56** on one side thereof and cooperating repulsion forces which add to the rotation generating forces produced on the opposite side.

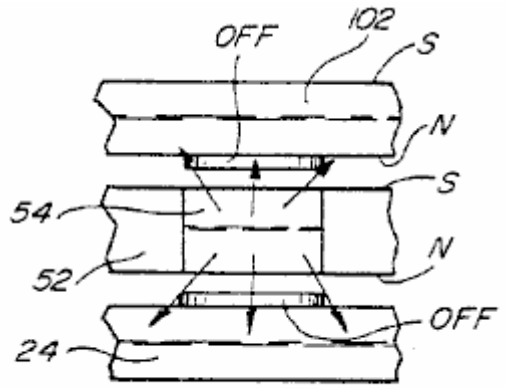


Fig. 8

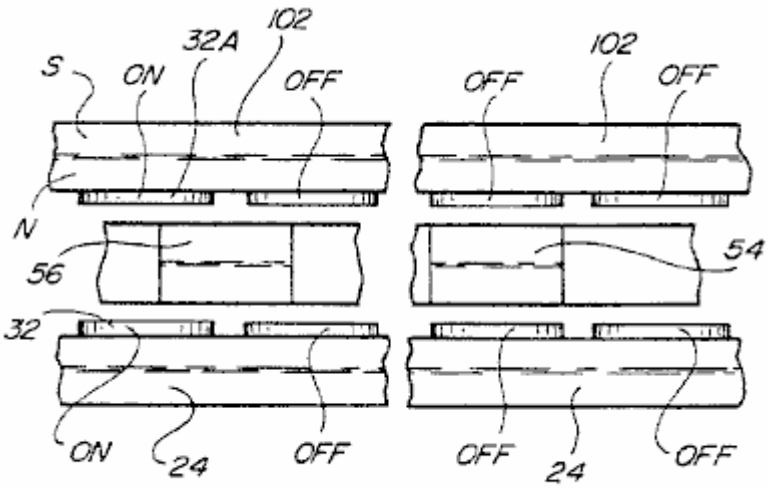


Fig. 9

Fig.8 and **Fig.9** are similar to **Fig.3** and **Fig.4** but show the relationship between the magnets **54** and **56** and the members **24** and **102** located on opposite sides. These figures show one form of interaction between the rotating magnets **54** and **56** and the stationary magnets **24** and **102** located as shown in **Fig.7**. In this construction, the device produces attractive rotating force only.

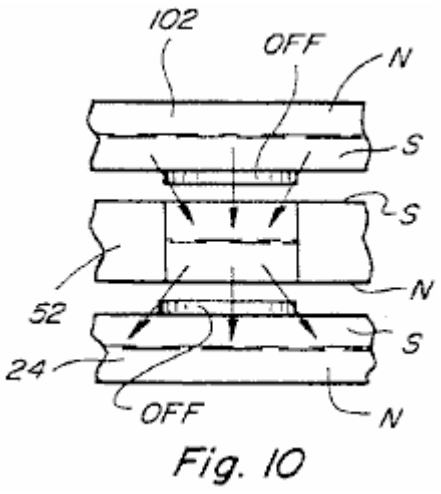


Fig. 10

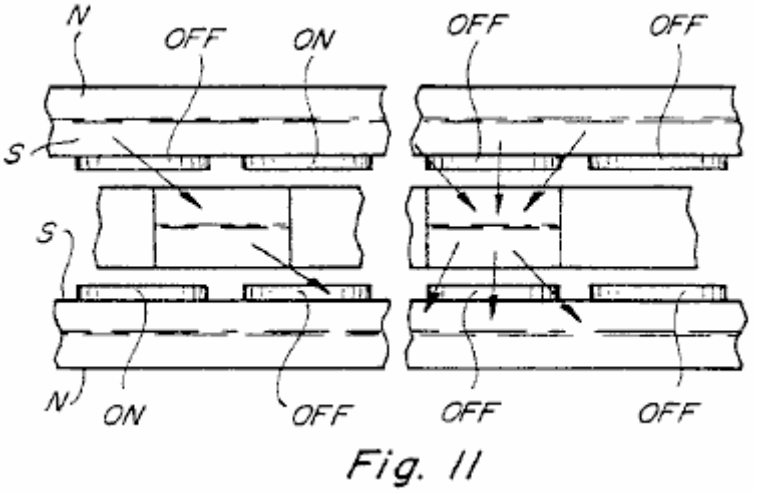


Fig. 11

Fig.10 and **Fig.11** are similar to **Fig.8** and **Fig.9** except that in these figures both attraction and repulsion forces are shown being produced in association with the stationary magnets on opposite sides of the rotating magnets. Note also that the coils being energised on opposite sides of disc **52** are energised in a different arrangement.

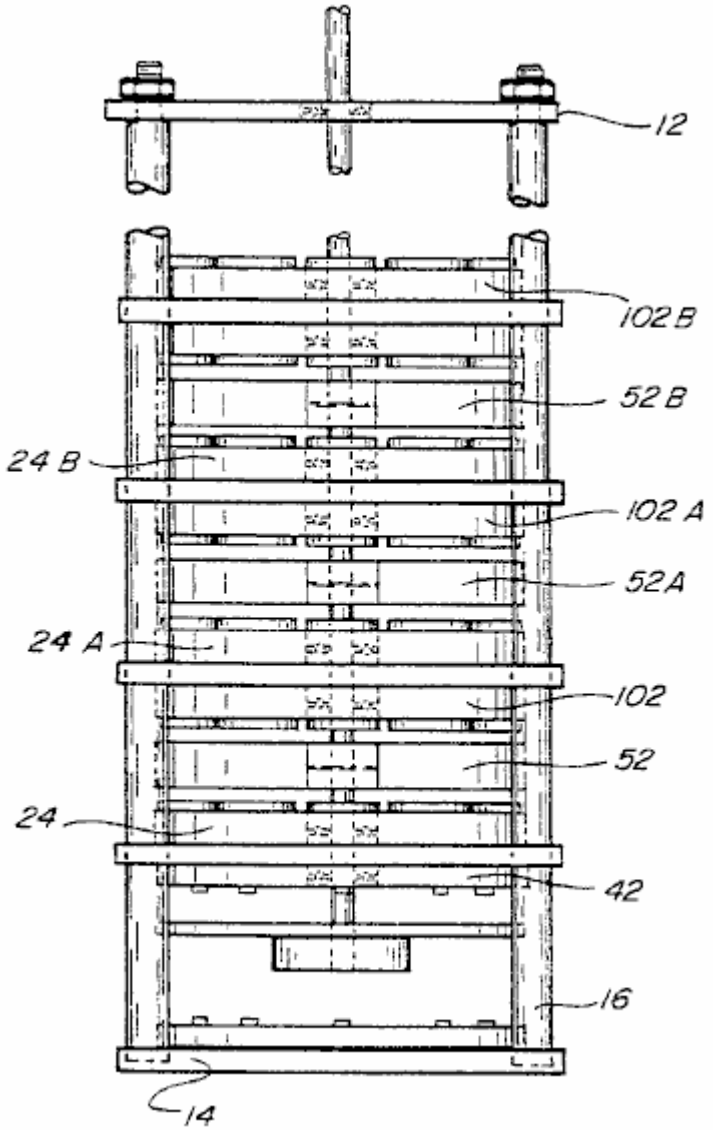


Fig. 12

Fig.12 is a side view similar to **Fig.7** but showing the way in which several stationary and rotating magnetic members such as the discs **24** and **102** can be mounted on the same shaft, in almost any number of repeating groups to increase the amount of torque produced by the device. In **Fig.12**, the same power source and the same circuit arrangement can be used to energise the phototransistors and the infra red emitters. However, depending upon whether attraction or repulsion forces are used to produce the rotation or some combination of

them, will depend upon the order in which the coils associated with the stationary magnetic members are energised.

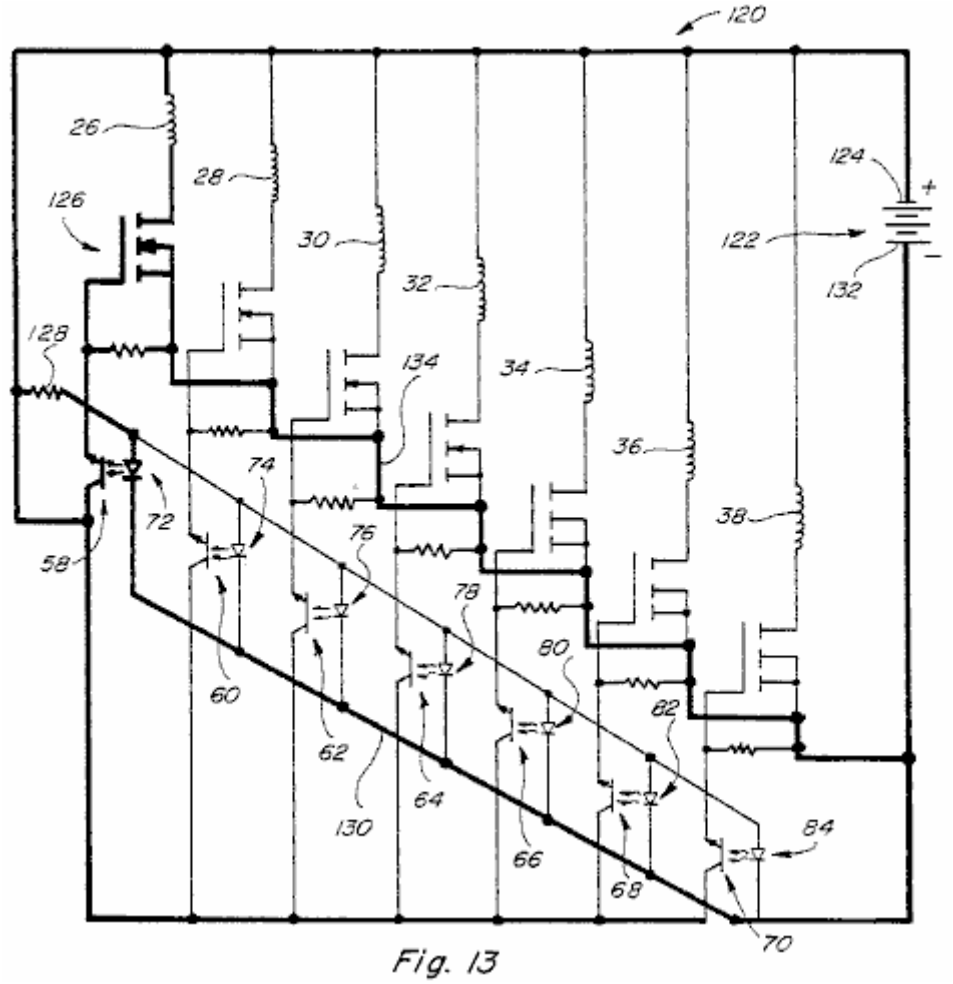


Fig. 13

Fig.13 is a circuit diagram for the device shown in **Fig.1** and **Fig.2**, showing the circuit connections for the coils **26-38** and for the circuit elements associated with them. A similar circuit can be used for the construction shown in **Fig.7** and **Fig.12**. The circuit also includes connections to the various phototransistors and infra red emitters.

In **Fig.13**, the circuit **120** is shown including a power supply **122** which may be a battery power supply, a rectified AC power supply or an AC or pulsed power supply. The positive side **124** of the power supply **122** is shown connected to one side of each of the coils **26-38**, coil **26** and the circuits associated with it being shown in bold outline and including connections to one side of a resistor **128** and to one side of the photo transistors **58-70**. The opposite side of the coil **26** is connected to one terminal of MOSFET **126**. The opposite side of the resistor **128** is connected to one side of the infra red emitter **72**, as well as to the corresponding sides of all of the other infra red emitters **74-84**. The opposite sides of the infra red emitters **72-84** are connected by lead **130** to the negative terminal side **132** of the power supply **122**. With the circuit as shown, the infra red emitters **72-84** are all continuously energised and produce light which can be detected by the respective phototransistors **58-70** when one of the openings **48** or **50** passes between them. When this happens, the respective phototransistor **58** will conduct and in so doing will apply positive voltage on the associated MOSFET **126**, turning the MOSFET on, and causing the voltage of the source **122** to also be applied across the coil **26**. The circuit for this is from the source **122** through the coil **26**, through the MOSFET **126** to and through the lead **134** to the opposite side of the source **122**. When the supply voltage is applied across the coil **26**, it operates to limit or prevent magnetic communication between whichever one of the magnets **54** or **56** happens to be positioned adjacent to the coil **26** which is in the space between that magnet **54** or **56** and the magnet **24**. This circuit is shown in bold in **Fig.13**. By properly timing and controlling the application of voltage to the various coils **26-38** in the manner described, the magnetic coupling between the magnets **54** and **56** and the magnet **24** can be accurately controlled and cause angular magnetic attraction between the magnet **54** (or **56**) and magnet **24**, which angular attraction (or repulsion) is in a direction to cause rotation of the rotating parts of the structure shown in **Figs. 1**, **2**, **7** and **12**. It should be understood that each of the coils **26-38** will be controlled in the same manner, that is, will have a voltage appearing across it at the proper time to control the direction of the magnetic coupling in a manner to produce rotation. The rotating portions will continue to rotate and the speed of rotation can be maintained at any desired speed. Various means can be used to control the speed of rotation such as by controlling the timing of

the DC or other voltage applied to the various coils, such as by using an alternating or pulsed current source instead of a direct current source or by loading the device to limit its rotational speed.

It is especially important to note that the energy required to operate the subject device is minimal since very little electrical energy is drawn when voltage is applied across the various coils when they are energised.

A well known equation used for conventional motor art, is:

Power (in watts) = Speed x Torque / 9.55

Hence,

W = S x T / 9.55

This equation has limited application to the present device because in the present device the torque is believed to be constant while the speed is the variable. The same equation can be rewritten:

T = 9.55 x W / S or S = 9.55 X W / T

These equations, if applicable, mean that as the speed increases, the watts divided by the torque must also increase but by a factor of 9.55. Thus if torque is constant or nearly constant, as speed increases, the power output must increase and at a very rapid rate.

It should be understood that the present device can be made to have any number of stationary and rotating magnets arranged in stacked relationship to increase the power output, (see **Fig.12**) and it is also possible to use any desired number of coils mounted on the various stationary magnets. In the constructions shown in **Figs. 1**, **7**, and **12** seven coils are shown mounted on each of the stationary magnets but more or fewer coils could be used on each of stationary magnet depending upon the power and other requirements of the device. If the number of coils is changed the number of light sources and photo-detectors or transistors will change accordingly. It is also important to note that the timing of the turning on of the various phototransistors is important. The timing should be such as that illustrated in **Fig.4**, for example, when one of the coils such as coil **32** is energised to prevent coupling in one direction between magnet **56** and magnet **24**, the adjacent coil **34** will not be energised. The reasons for this have already been explained.

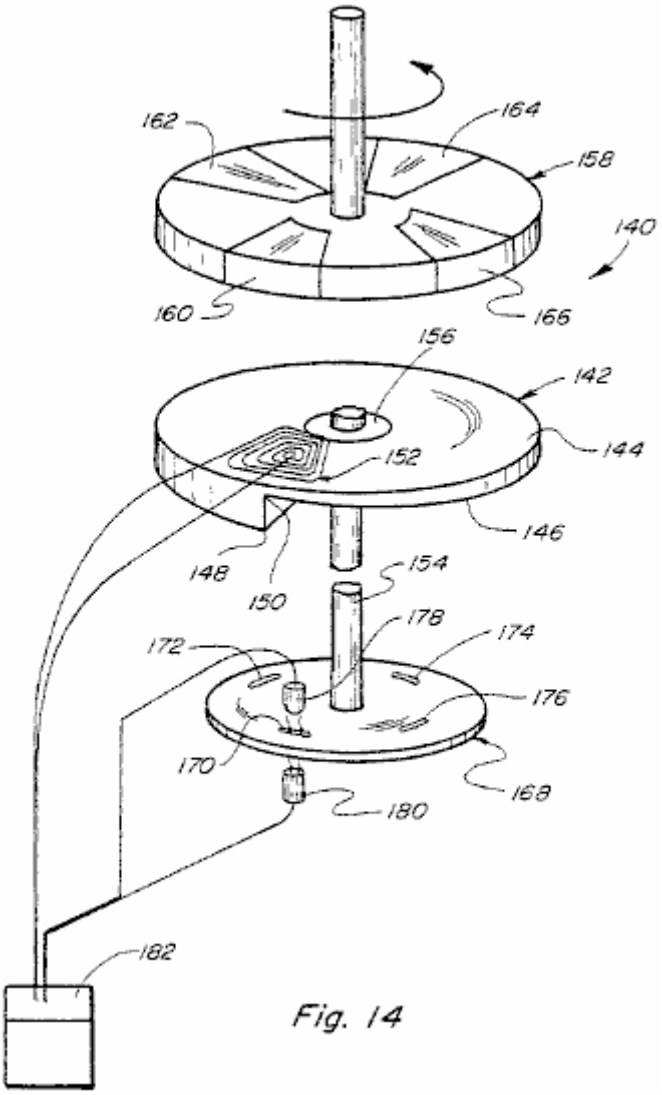


Fig. 14

Fig.14, shows another embodiment **140** of this motor. This includes a stationary permanent magnet **142** which has a flat upper surface **144** and a lower surface **146** that is circumferentially helical so that the member **142** varies in thickness from a location of maximum thickness at **148** to a location of minimum thickness at **150**. The thickness of the member **142** is shown varying uniformly. Near the location of the thickest portion **148** of the permanent magnet **142** and adjacent to the surface **144** is an air coil **152** shown formed by a plurality of windings. A shaft member **154** is journaled by the bearing **156** to allow rotation relative to the stationary permanent magnet **142** and is connected to a rotating disc **158**. The disc includes four spaced permanent magnets **160**, **162**, **164** and **166** mounted on or in it. The permanent magnets **160-166** are positioned to rotate close to the stationary permanent magnet **142** but with the coil **152** positioned between them. Coil **152** is connected into a circuit similar to that shown in **Fig.13** and so the circuit will not be described again.

The principals of operation of the device **140** shown in **Fig.14** are similar to those described above in connection with **Fig.1** and other figures. It is important to note, however, that the permanent magnets **160-166** rotate relative to the permanent magnet **142** because of the increasing coupling between them and the permanent magnet due to the increasing peripheral thickness of the permanent magnet. Thus the member **158** will rotate in a counter-clockwise direction as shown, and each time one of the magnets **160-166** moves into a position adjacent to the thickest portion **148** of the fixed permanent magnet **142** the coil **152** will have voltage applied across it, otherwise there would be a tendency for the member **158** to stop or reduce the rotational force. In order to overcome this the coil **152** is energised each time one of the permanent magnets **160-166** is in the position shown. The rotating disc **158** is connected through the shaft **154** to rotating disc **168** which has four openings **170**, **172**, **174** and **176** corresponding to the locations of the permanent magnets **160-166** so that each time one of the permanent magnets moves to a position adjacent to the thickest portion **148** of the stationary permanent magnet **142** the coil **152** will be energised and this will reduce or eliminate the coupling between the rotating and stationary magnets that would otherwise slow the rotating portions down.

The circuit connected to the coil **152** includes the same basic elements described above in connection with **Fig.13** including varying a photocell **178**, an infra red emitter **180** and a MOSFET **182** connected into a circuit such as

that shown in **Fig.13**. The timing of the energising of the coil **152** is important and should be such that the coil will be energised as the respective permanent magnets **160-166** move to a position in alignment or substantial alignment with the thickened portion **148** of the stationary permanent magnet **142**.

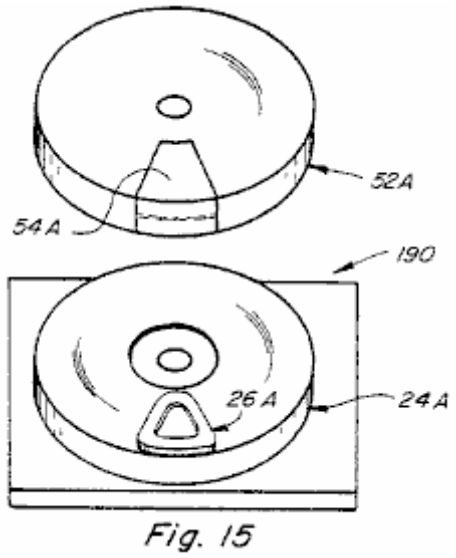


Fig. 15

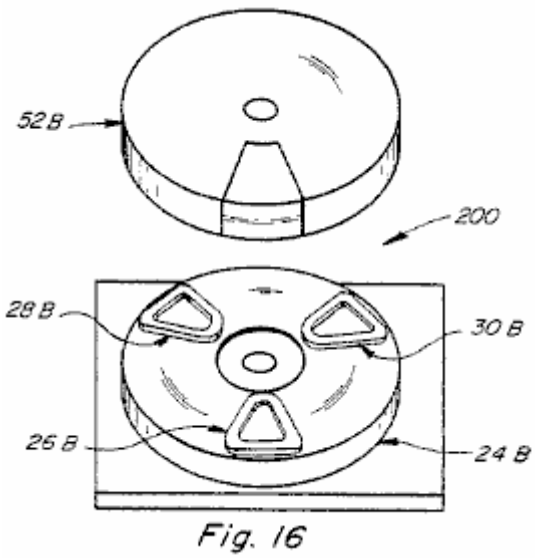


Fig. 16

Fig.15 shows a basic simplified form **190** of the present device which includes a rotary member **52A** having a single permanent magnet portion **54A** mounted on it. The device also has a stationary permanent magnet **24A** with a single air coil **26A** positioned in the space between the members **52A** and **24A** in the manner already described. The construction **190** is not self-starting as are the preferred embodiments such as embodiment **10** but the rotary portions will rotate continuously once the device is started as by manually rotating the rotary portions. The construction **190** will have other portions as described above but the output from the construction will be less than the output produced by the other constructions.

Fig.16 shows another simplified version **200** of the device wherein the member **52B** is similar to the corresponding rotating member **52A** shown in **Fig.15**. However, the fixed structure including the permanent magnet **24B** has three windings **26B**, **28B** and **30B** located at spaced intervals adjacent to the upper surface of it. The construction shown in **Fig.16** will produce more output than the construction shown in **Fig.15** but less than that of the other constructions such as that shown in **Figs. 1**, **2**, **7** and **12**. Obviously, many other variations of the constructions shown in the application are also possible including constructions having more or fewer coils, more or fewer rotating magnetic portions, more or fewer rotating members such as disc **52** and more or fewer stationary members such as magnets **24** and **142**.

Figs.17-25 illustrate some of the underline principles of the present invention.

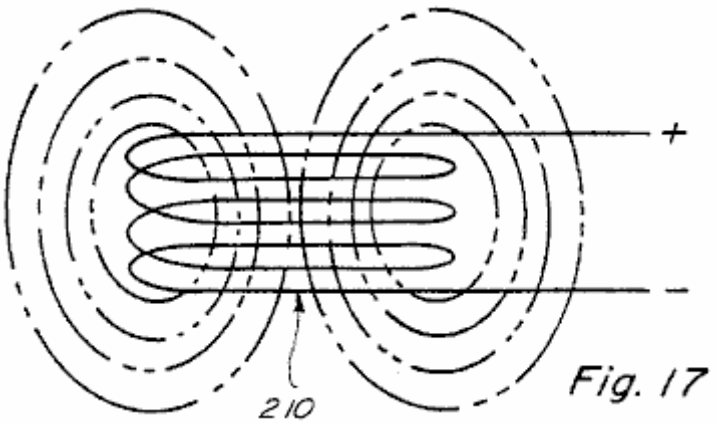


Fig. 17

Fig.17 shows an air coil **210**, positioned in space, with an electric potential applied across it. With the energising voltage applied, the electro-magnetic field of air coil **210** extends substantially equally in the space above and below the coil as shown in dotted outlined.

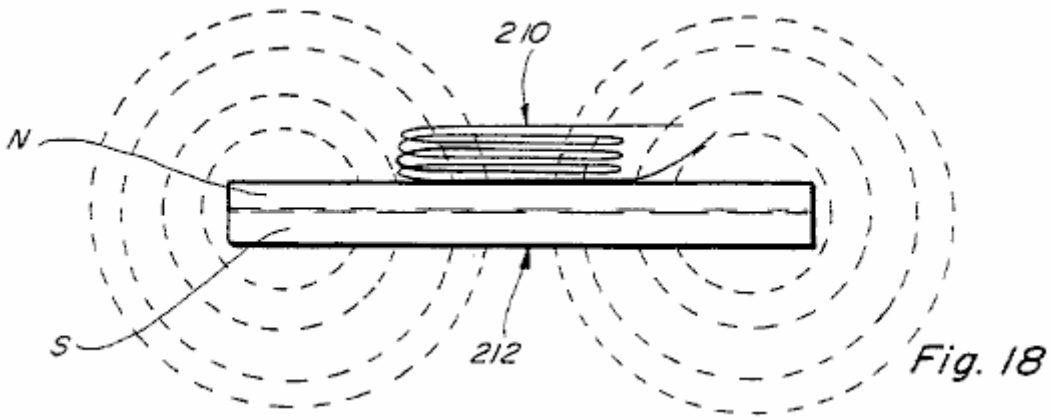


Fig. 18

Fig.18 shows the air coil **210** positioned adjacent to one side (the north side) of permanent magnet **212**. In **Fig.18** no voltage is applied across the air coil **210** and therefore the coil does not produce an electro-magnetic field as in **Fig.17**. Under these circumstances, the air coil **210** has no effect on the magnetic field of the permanent magnet **212** and the field of the permanent magnet is substantially as shown by the dotted outlines in **Fig.18**.

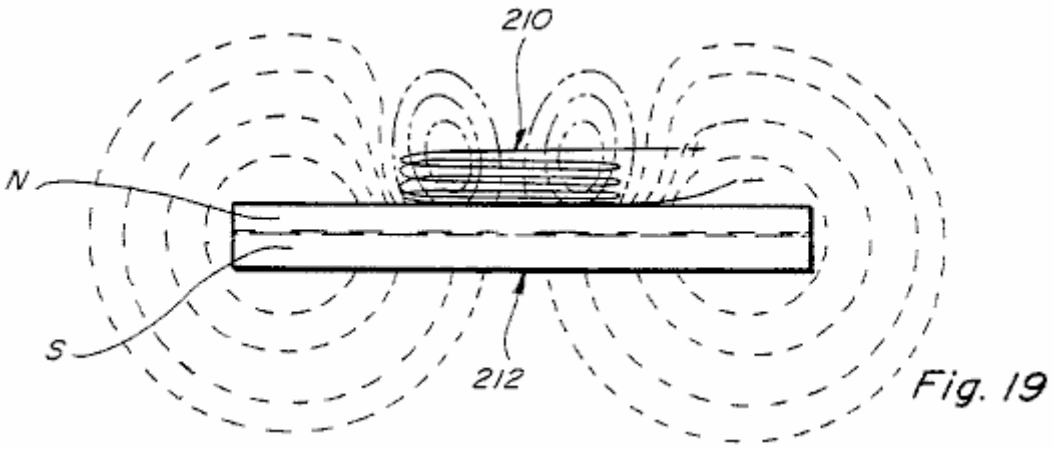


Fig. 19

Fig.19 is similar to **Fig.18** except that in **Fig.19** the air coil **210** has an electric potential applied across it and therefore has an established electro-magnetic field shown again by dotted outline.

The electro-magnetic field of the air coil **210** modifies the magnetic field of the permanent magnet **212** in the manner shown. If coil **210** is placed in contact with, or close to the surface of, the permanent magnet and it is energised so that its polarity is opposite to that of the permanent magnet then the field produced is similar to that shown in **Fig.19**. Note that the field of coil **210** and the field of the permanent magnet **212** directly beneath the air coil **210** are in opposition and therefore act to cancel one another. Coil **210** would be defined to produce a counter-magnetomotive force which acts to cancel the field of the permanent magnet **212** in the region where the air coil **210** exists and the amount of the field in that region of the permanent magnet **212** that is cancelled is the remainder of the difference in magnetomotive force between the region of the permanent magnet **212** and the counter magnetomotive force of the air coil **210**. Note that, since the field of permanent magnet **212** is only altered in the region of the air coil **210**, the geometric magnetic field characteristics of the permanent magnet **212** can be altered selectively based upon the size of the coil **210**, the number of air coils **210** and the amount of counter magnetomotive force being produced by the air coil **210**.

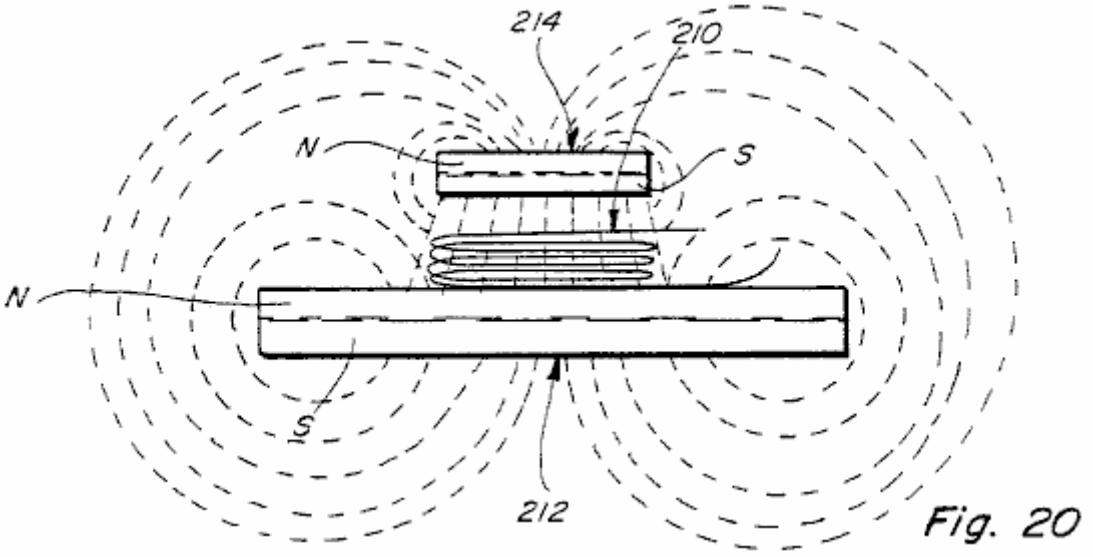


Fig. 20

Fig.20 is similar to **Fig.19** except that a second permanent magnet **214** is positioned at a location spaced above the air coil **210**. In **Fig.20** no voltage is applied across the air coil **210** and therefore the air coil **210** does not have an electro-magnetic field. Thus **Fig.20** shows only the combined affect of the fields of the permanent magnets **212** and **214**. Since the permanent magnets **212** and **214** are positioned so that their respective north and south poles are close together, there will be a strong attractive force between them at the location of the air coil **210**.

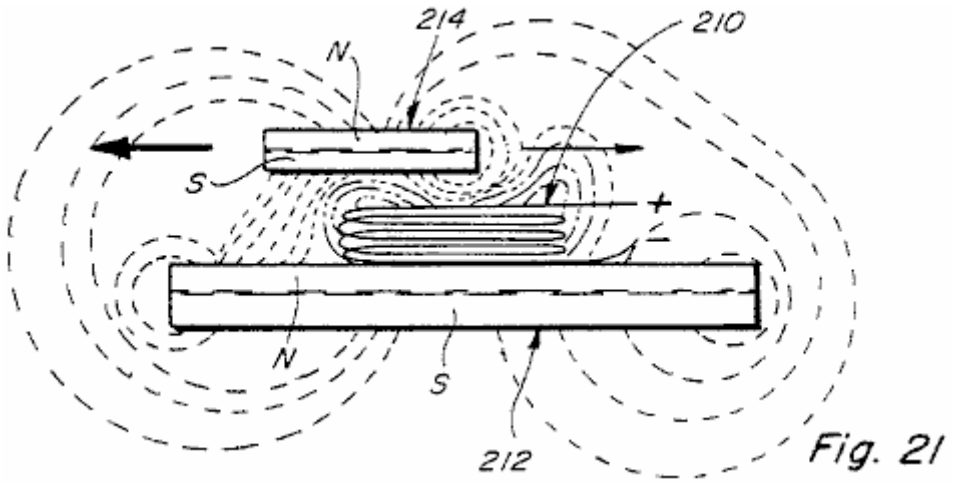


Fig. 21

Fig.21 is a view similar **Fig.20** but with an electric potential applied across the air coil **210** and with the upper permanent magnet **214** displaced to the left relative to its position in **Fig.20**. Note that in **Fig.21** the shape of the electro-magnetic field of the air coil **210** is concentrated and shifted somewhat to the right and upward. This shift of the electro-magnetic field concentrates the magnetic coupling between the magnets **212** and **214** to the left thereby increasing the tendency of the upper permanent magnet **214** to move to the left. A much smaller magnetic coupling occurs between the right end of the permanent magnets **212** and **214** and thus the force tending to move the permanent magnet **214** to the right is much less than the force tending to move it to the left. This is illustrated by the size of the arrows shown in **Fig.21**.

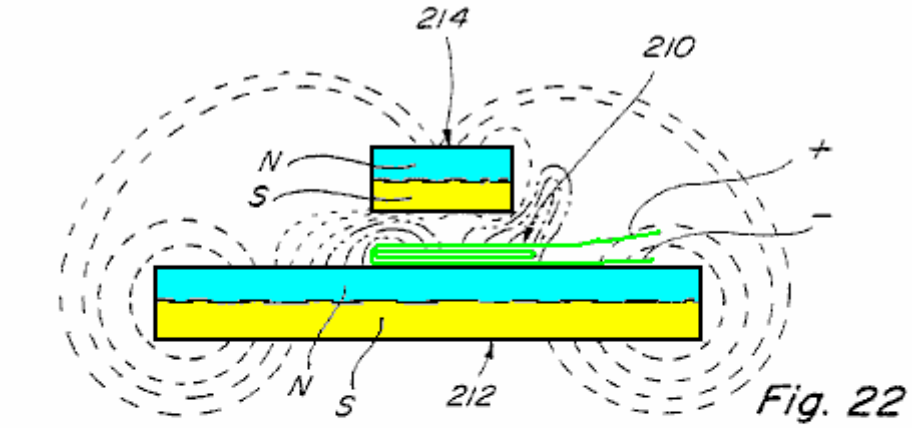


Fig. 22

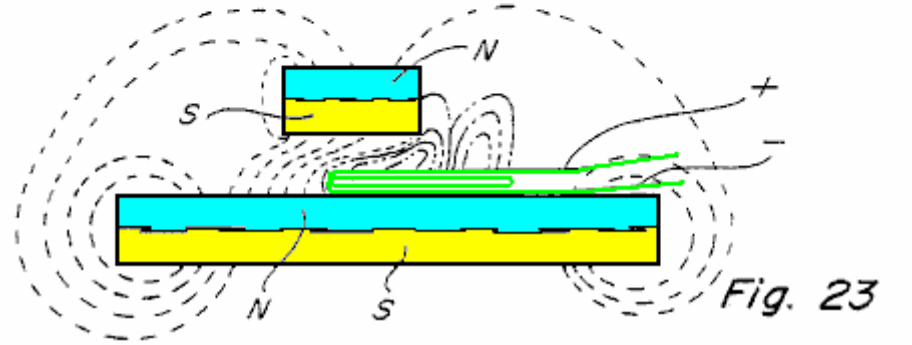


Fig. 23

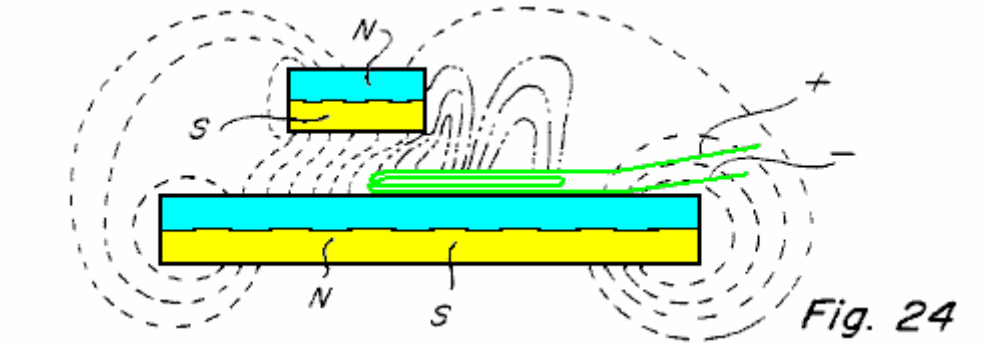


Fig. 24

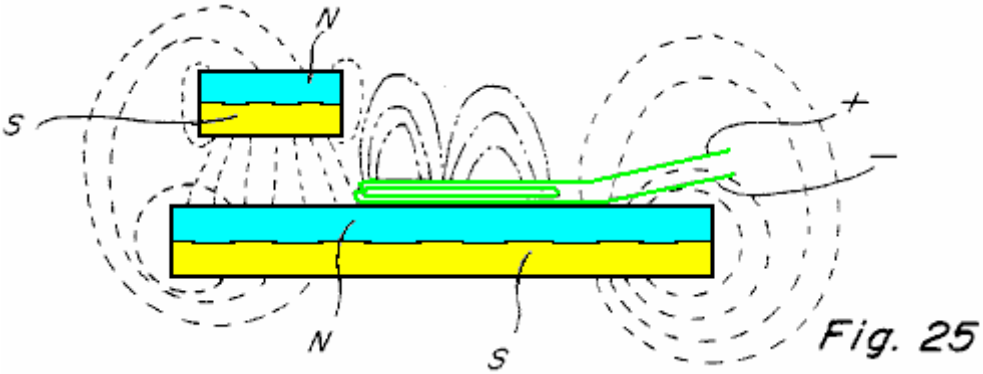


Fig. 25

Figs. 22-25 show four different positions of the upper permanent magnet **214** relative to the lower permanent magnet **212**. In **Fig.22** because of the position of the upper permanent magnet **214** relative to the air coil **210** there is a concentration of the magnetic coupling force tending to move the upper permanent magnet **214** to the left. This force increases in **Fig.23** and **Fig.24** until the upper permanent magnet **214** reaches the position shown in **Fig.25** where all of the magnetic coupling is directed substantially vertically between the permanent magnets **212** and **214** and in this position there is little or no torque as a result of coupling energy between the permanent magnets **212** and **214** tending to move them relative to one another.

The principles illustrated in **Figs. 17-25** are at the heart of the present invention and explain where the energy comes from to produce relative movement between the permanent magnets.

The present device has application for very many different purposes and applications including almost any purpose where a motor or engine drive is required and where the amount of energy available and/or required to produce the driving force may vary little to nil. Applicant has produced devices of the type described herein capable of rotating at very high speed in the order of magnitude of 20,000 RPMs and with substantial torque. Other lesser speeds can also be produced, and the subject device can be made to be self starting as is true of the constructions shown in **Figs. 1**, **2**, **7** and **12**. Because of the low power required to operate the device applicant has been able to operate same using a commercially available battery such as a nine volt battery.

CLAIMS

1. A device to control the magnetic interaction between spaced permanent magnets comprising:

a first permanent magnet having opposite surfaces with north and south poles respectively,

a second permanent magnet spaced from and movable relative to the first permanent magnet and having opposite surfaces with north and south poles respectively, one of which is positioned in close enough proximity to one of the surfaces of the first permanent magnet to produce magnetic interaction between them,

a coil of conductive metal positioned in the space between the first and second permanent magnets,

a source of electrical energy and switch means connected in series therewith across the coil whereby when the switch means are closed the electrical energy from said source is applied across the coil whereby the magnetic interaction between the first and second permanent magnets is changed, and

means to control the opening and closing of the switch means.

2. A device for producing rotational movement and torque comprising:

a member journaled for rotational movement about an axis of rotation, the rotating member having at least a portion adjacent the periphery thereof formed of a permanently magnetized material,

a stationary member formed of permanently magnetized material mounted adjacent to the peripheral portion of the rotating member axially spaced from it whereby a magnetic interaction is produced between the stationary and the rotating members in predetermined positions of the rotating member,

at least one coil positioned extending into the space between the stationary and rotating members,

means including a source of electric potential and switch means connected in series across the coil, and

means to predeterminately control the opening and closing of the switch means during rotation of the rotating member to vary the magnetic interaction in a way to produce rotation of the rotating member.

3. Means to predeterminately vary the magnetic interaction between first and second spaced permanent magnet members comprising a first permanent magnet member having north and south poles, a second permanent magnet member having north and south poles spaced from the first permanent magnet member by a gap between them, a coil positioned extending into the gap between the first and second permanent magnet members, means connecting the coil across a circuit that includes a source of voltage and switch means connected in series therewith so that when the voltage source is connected across the coil it effects the magnetic interaction between the first and second permanent magnet members, and means for mounting the first permanent magnet member for movement relative to the second permanent magnet member and relative to the coil in the gap between them.

4. The device of claim 3 wherein the first and second permanent magnet members are mounted to produce magnetic attraction between them.

5. The device of claim 3 wherein the first and second permanent magnet members are mounted to produce magnetic repulsion between them.

6. The device of claim 3 wherein the means mounting the first permanent magnet member includes means mounting the first permanent magnet member for rotational movement relative to the second permanent magnet member and the switch means includes cooperative optical means having a first portion mounted for

movement with the first permanent magnet member and a second portion associated with the second permanent magnet member.

7. The device of claim 6 wherein the switch means includes a light source and a light sensitive member associated respectively with the first and second permanent magnet members, and control means for them mounted for movement with the first permanent magnet.

8. The device of claim 3 wherein the second permanent magnet member is an annular permanent magnet member having one of its poles on one side of the gap and the other of its poles opposite thereto, means mounting the first permanent magnet member for rotational movement relative to the second permanent magnet member, said first permanent magnet member having one of its poles on one side of the gap, and a plurality of circumferentially spaced coils mounted in the gap between the first and second permanent magnet members.

9. The device of claim 8 wherein the first permanent magnet member includes two circumferentially spaced portions.

10. Means for producing rotational movement comprising:

a support structure having a first permanent magnet mounted thereon, said first permanent magnet having a north pole adjacent one surface and a south pole adjacent to the opposite surface,

means for mounting a second permanent magnet for rotational movement in a plane parallel to the first permanent magnet, the second permanent magnet occupying an curved portion of said mounting means less than the entire circumference of said mounting means and having a north pole adjacent to the opposite surface and positioned so that there is a magnetic interaction between the spaced first and second permanent magnets across a gap between them in at least one position thereof,

at least one air coil positioned in the gap between the first and second permanent magnets,

a source of electric potential and switch means for controlling the application of the electric potential from said source across the air coil, the application of voltage across the air coil effecting the magnetic interaction between the first and second permanent magnet members in certain positions of the second permanent magnet relative to the first permanent magnet and in such a manner as to produce rotational movement of the second permanent magnet.

11. The device for producing rotational movement of claim 10 wherein a third permanent magnet is mounted on the support structure on the opposite side of the second permanent magnet from the first permanent magnet so as to establish a second gap between them and so that there is magnetic interaction between the second and third permanent magnets, and at least one second coil mounted in the gap between the second and third permanent magnets to predeterminately effect the magnetic interaction between them in certain positions of the second permanent magnet relative to the third permanent magnet thereby to contribute to the production of rotational movement of the second permanent magnet member relative to the first and third permanent magnets.

12. The device for producing rotational movement defined in claim 11 wherein the switch means for applying voltage from the source across the coils includes a light source and light sensor one mounted on the support structure and the other on the rotating means to produce a switching action to apply and remove voltage from across the coils in predetermined positions of the second permanent magnet relative to the first and third permanent magnets.

13. Means for producing rotary motion using magnetic energy from permanent magnets comprising:

a fixed permanent magnet having opposite surfaces with north and south poles respectively adjacent thereto,

a shaft having an axis and means journaling the shaft for rotation in a position extending normal to the opposite surfaces of the fixed permanent magnet,

a movable permanent magnet and means mounting the movable permanent magnet on the shaft for rotation therewith, the movable permanent magnet occupying an curved portion of said mounting means less than the entire circumference of said mounting means and having opposite surfaces with associated north and south poles respectively, one pole of said movable permanent magnet being positioned to move in close

enough proximity to one of the opposite surfaces of the fixed permanent magnet to produce magnetic interaction between them,

at least one coil mounted in the space between the fixed permanent magnet and the movable permanent magnet, energising of the coil effecting the magnetic interaction between the fixed and the movable permanent magnets when positioned between them, and

means connecting the coil to a source of energising potential in selected positions of the movable permanent magnet relative to the fixed permanent magnet.

14. The device for producing rotary motion of claim 13 wherein a plurality of coils are mounted in a coplanar relationship in the space between the fixed permanent magnet and the movable permanent magnet, the means connecting the coils to a source of energising potential including means for energising the respective coils in a predetermined sequence.

15. The device for producing rotary motion of claim 13 including a second movable permanent magnet mounted on the means mounting the movable permanent magnet for movement therewith, said second movable permanent magnet being spaced circumferentially from the aforesaid movable permanent magnet.

16. The device for producing rotary motion of claim 13 wherein a second fixed permanent magnet has opposite surfaces with north and south poles respectively adjacent thereto and is mounted on the opposite side of the movable permanent magnet from the aforesaid fixed permanent magnet and at least one coil mounted in the space between the second fixed permanent magnet, and the movable permanent magnet.

17. A device for producing rotary motion defined in claim 13 wherein the means connecting the coil to a source of energising potential includes a fixed light source and a fixed light sensitive member mounted in spaced relationship and means on the mounting means for the movable permanent magnet for predeterminately controlling communication between the light source and the light sensitive member during rotation of the movable permanent magnet.

18. A magnetic motor-like device comprising:

a fixed support structure having a permanent magnet member mounted thereon, said member having opposite side faces with a north magnetic pole adjacent one side face and a south magnetic pole adjacent the opposite side face,

a plurality of coils mounted adjacent to and arranged about one of the opposite side faces,

an orifice through the permanent magnet member at a location intermediate the coils,

a shaft extending through the orifice for rotation about the axis thereof,

a member attached to the shaft for rotation therewith and spaced from the one opposite magnet side faces,

at least one magnet member attached to a segment of said rotating member for rotation therewith, each of said rotating magnetic members having a magnetic pole face positioned in spaced relation to the one opposite pole side face of the fixed permanent magnet member, the plurality of coils being in the space formed by and between the fixed permanent magnet member and the at least one rotating magnet member, and

means to selectively and sequentially energise the coils as the shaft rotates to predeterminately control the magnetic interaction between the at least one magnetic member and that fixed permanent magnet member.

19. The magnetic device of claim 18 wherein there is an odd number of coils mounted in the space between the permanent magnet member and the at least one rotating magnetic member.

20. The magnetic device of claim 18 wherein the at least one magnetic member attached to the rotating member for rotation therewith includes two circumferentially spaced rotating magnet portions.

21. A device for producing rotary motion comprising:

a support structure having a wall member,

a shaft and means journaling the shaft for rotation in the wall member about its axis,

a permanent magnet member mounted on the wall member extending about at least a portion of the shaft, said permanent magnet member having one pole adjacent to the wall member and an opposite pole spaced therefrom,

a member mounted on the shaft having at least two magnetic members oriented to produce magnetic interaction with the permanent magnet member,

a plurality of coils mounted in coplanar relation extending into the space formed by and between the permanent magnet member and the at least two magnetic members and

means to sequentially apply a voltage across the respective coils to vary the magnetic interaction between the permanent magnet member mounted on the wall member and selected ones of the at least two magnetic members.

22. A device for producing rotary motion using magnetic energy from permanent magnets comprising

a fixed permanent magnet having opposite surfaces with north and south poles respectively adjacent thereto,

a shaft and means for journaling the shaft for rotation extending normal to the opposite surfaces of the fixed permanent magnet,

at least two rotatable permanent magnets and means mounting them for rotation with the shaft, the rotatable permanent magnets having opposite surfaces with associated north and south poles respectively, one pole of each rotatable permanent magnet being positioned close enough to one of the opposite surfaces of the fixed permanent magnet to produce magnetic interaction therebetween,

a plurality of spaced coils arranged to be coplanar and positioned in the space formed by and between the fixed permanent magnet and the rotatable permanent magnets, and

means to apply a voltage across respective ones of the coils in a sequence so as to predeterminately affect the interaction between the fixed permanent magnet and the rotatable permanent magnets in a manner to produce rotation of the at least two permanent magnets.

23. A device for producing rotary motion using magnetic energy from permanent magnets comprising:

a fixed annular permanent magnet having a flat surface on one side and an opposite surface of helical shape extending therearound from a location of minimum thickness to a location of maximum thickness approximately adjacent thereto, the annular permanent magnet having one of its poles adjacent to the flat surface and its opposite pole adjacent to the helical opposite surface,

a shaft and means for journaling the shaft for rotation extending substantially normal to the flat surface of the fixed permanent magnet,

a permanent magnet and means mounting it on the shaft for rotation therewith, said permanent magnet having opposite pole faces and being positioned so that there is magnetic interaction between said permanent magnet and the fixed annular permanent magnet,

at least one air coil positioned in the space between the fixed and rotatable permanent magnets, and

means to apply a voltage across the air coil when the rotatable permanent magnet is adjacent to the thickest portion of the fixed permanent magnet to change the magnetic interaction therebetween, said last name means including a source of voltage and switch means in series with the source for controlling the application of voltage across the air coil.

24. The device for producing rotary motion of claim 23 wherein a plurality of rotatable permanent magnets are mounted at circumferentially spaced locations about the shaft for magnetic interaction with the fixed annular permanent magnet, the switch means controlling the application of voltage from the source to the air coil

when one of the rotatable permanent magnets is positioned adjacent to the thickest portion of the fixed annular permanent magnet.

25. The means for producing rotary motion of claim 23 wherein the switch means includes cooperative optical means having a first portion associated with the fixed annular permanent magnet and a second portion associated with the rotatable annular permanent magnet.

The Power Plant of Claude Mead and William Holmes

US Patent 4,229,661 21st October 1980 Inventors: Claude Mead and William Holmes

POWER PLANT FOR CAMPING TRAILER

Note: This patent is not a free-energy patent, but it does provide a suggestion for an integrated and practical system for providing power for people living in a caravan which is frequently off-grid but which occasionally is positioned where electrical mains power is available. It describes a practical system for storing wind energy for high-power electrical power supply, and so is of interest.

ABSTRACT

A power plant for mobile homes, camping trailers, and the like, capable of capturing low-powered wind energy, storing the energy in the form of compressed air, and delivering it on demand in the form of household electrical current. The device comprises a wind turbine which drives an air compressor which feeds a storage tank. When required, the compressed air drives a turbine coupled to an electrical generator. Various pressure regulators are used to control the speed of the generator. The wind turbine is also coupled to an alternator which keeps a bank of batteries charged. A DC motor running on the batteries, is used when necessary, to boost the drive of the air compressor during periods of heavy or long power drain. Provision is made for rapidly recharging the power plant from either a supply of compressed air or from an AC power source.

US Patent References:

2230526	Wind power plant	February, 1941	Claytor	290/44
2539862	Air-driven turbine power plant	January, 1951	Rushing	290/44
3315085	Auxiliary power supply for aircraft	April, 1967	Mileti et al.	290/55
3546474	Electrohydraulic Transmission of Power	December, 1979	DeCourcy et al.	290/1
4150300	Electrical and thermal system for buildings	April, 1979	VanWinkle	290/55

BACKGROUND OF THE INVENTION

The current shortage of fossil fuel and public concern for the quality of the environment have triggered a hurried search for alternate forms of energy. The capture and use of solar energy, and its derivative, wind power, is the object of many new inventions. Due to the inefficiency of the collector device and storage media, use of these forms of energy has been limited to low-power stationery applications. Yet wind power should be adequate for any application requiring very low power or a short, occasional low to medium power supply of energy. These circumstances are encountered, for instance, in a refrigerated railroad car where occasional bursts of power are required to run the refrigerating system in order to maintain a low temperature inside the car. Similar circumstances are found in some mobile housing units such as a camping trailer. There, again, a supply of household current might be necessary for a short time between long periods of travel. In such instances, a system can be devised for accumulating energy generated by a wind turbine powered by the wind or by the air draft created by the motion of the vehicle. It is further desirable that the power system be capable of being replenished from non-polluting energy sources which can be encountered along the travel route.

SUMMARY OF THE INVENTION

It is accordingly an object of the instant invention to provide a novel power plant for mobile homes, and the like, which captures wind energy, stores it in the form of compressed air, and delivers it on demand in the form of household electrical current.

Another object of this invention is to provide a power plant which does not discharge polluting effluents into the atmosphere.

Still another object of the invention is to provide a power plant which can be recharged by capturing the effect of the wind, or the effect of the air stream created by the movement of the vehicle.

A further object of the invention is to provide a power plant which can be recharged from a household current electrical outlet.

It is also an object of this invention to provide a power plant which can be replenished from a source of compressed air such as those found in automotive service stations.

An additional object of the invention is to provide a power plant which is responsive to a very low level of wind energy for a short period of time.

These and other objects are achieved by a power plant which comprises a wind turbine driving an air compressor. The air supply of the compressor is stored in the tank and used on demand to activate a turbine. The turbine, in turn, is coupled to a generator which creates household current. The wind turbine is also coupled to generators which charge a series of electrical batteries. On occasions when the AC power drain requires it, a motor running on the batteries is used to boost the output of the air compressor. Provision is made for driving the compressor from an outside AC power source. The air tank has a separate inlet through which it can be replenished from a source of compressed air.

THE DRAWINGS

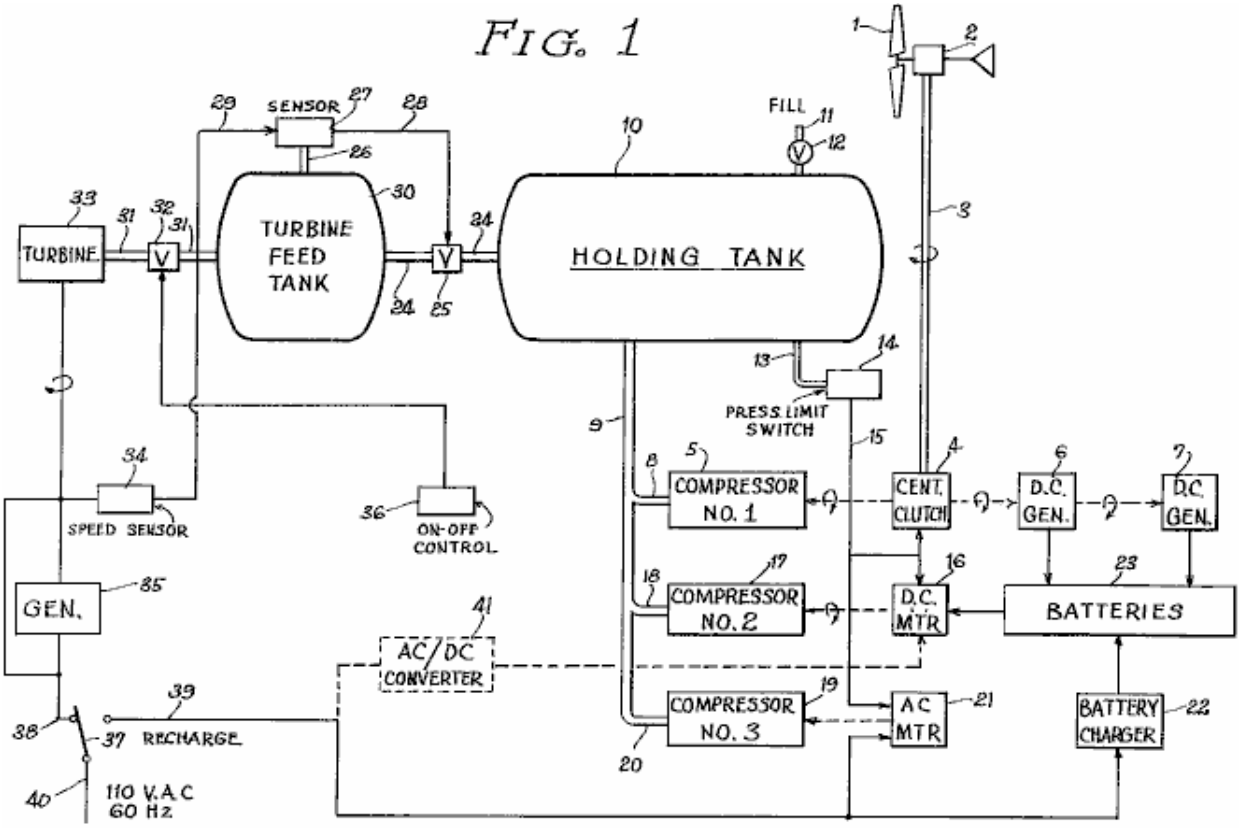


Fig.1 is the general block diagram of the entire power plant;

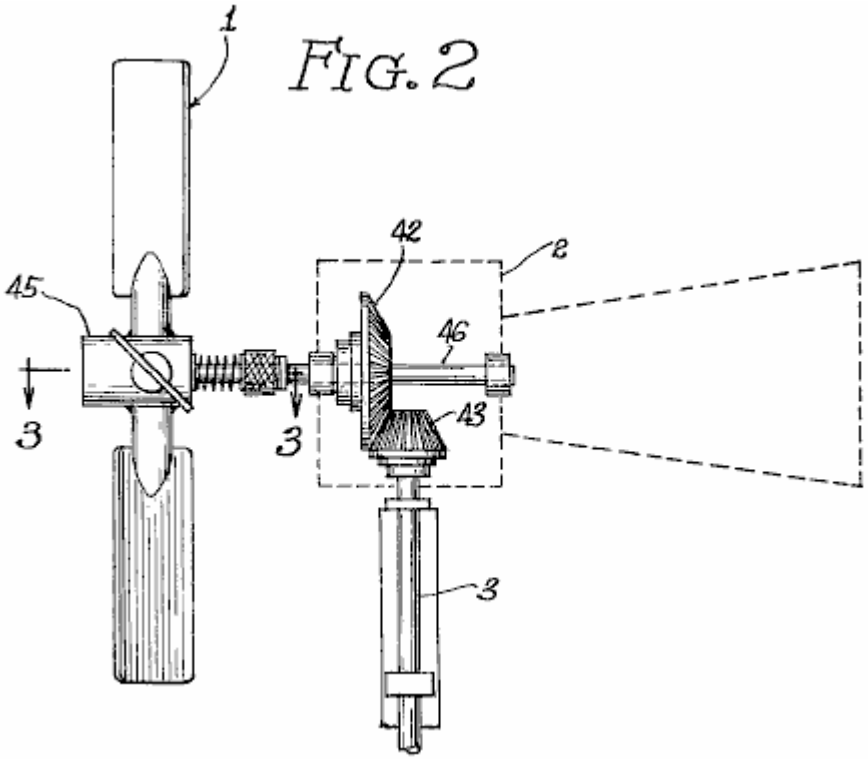


Fig.2 is a front elevation of the wind turbine and of its mechanical coupling to the drive shaft;

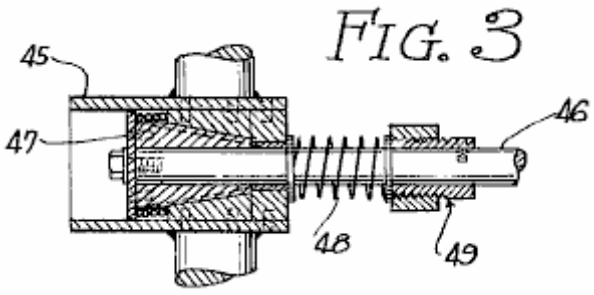


Fig.3 is a cross-sectional view taken along line 3--3 of Fig.2 showing the propeller linkage mechanism in the engaged position;

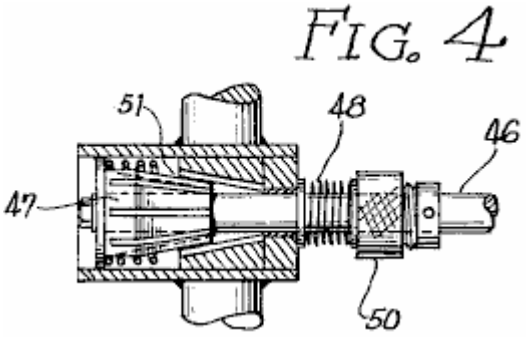
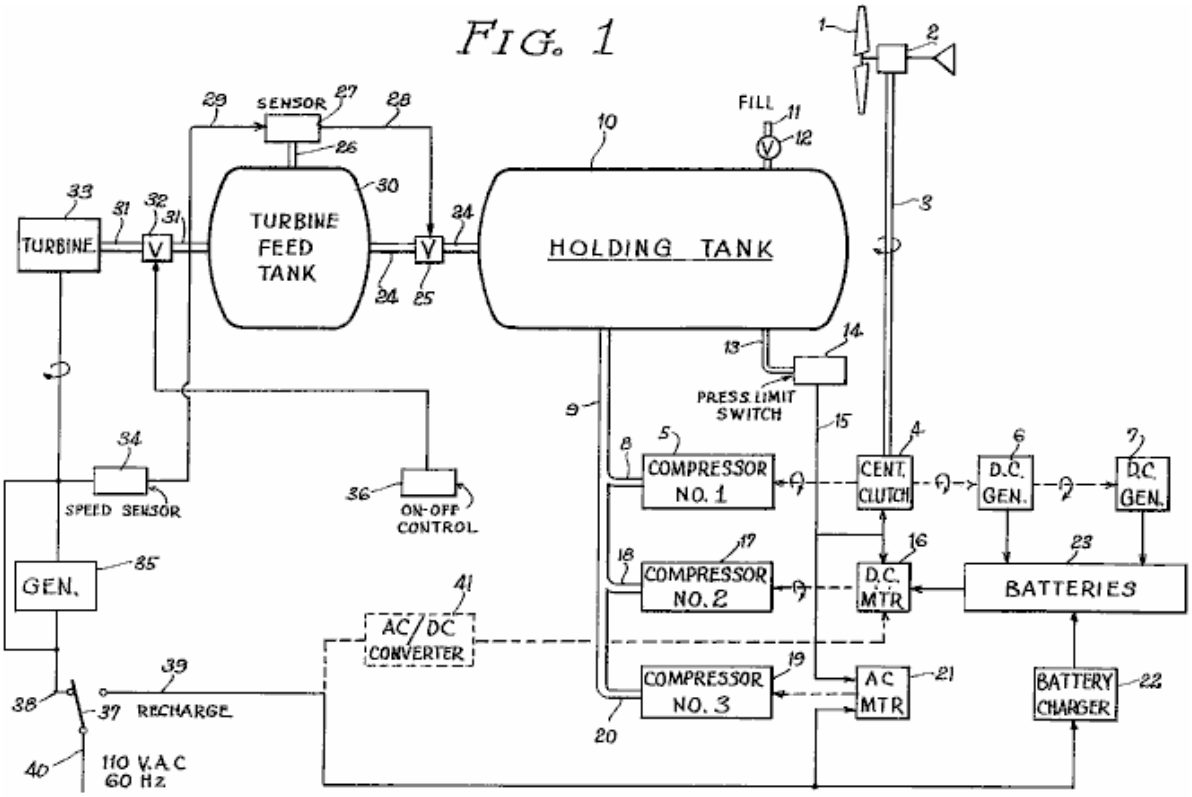


Fig.4 is a view similar to the one illustrated in Fig.3 but showing the propeller linkage mechanism in the disengaged position.

DESCRIPTION OF THE PREFERRED EMBODIMENT OF THE INVENTION



Referring now to **Fig.1**, there is shown a diagramatic representation of the preferred embodiment of the invention. A wind turbine comprising a propeller **1** and an orthogonal coupling assembly **2** drives a shaft **3** connected to a centrifugal clutch **4**. This type of clutch is designed to engage itself when the speed of the drive shaft **3** reaches a certain minimum preset limit. The plate of the clutch is first connected to a compressor **5** and second to two DC generators **6** and **7**. Block **5** represents a adiabatic compressor requiring an input drive of approximately one-fourth horsepower.

The output of the compressors **5** is protected by a check valve and leads into a pipe **8** connected to a tank inlet pipe **9**. The inlet pipe **9** feeds into a holding tank **10** capable of holding sixty gallons of compressed air under a maximum pressure of 200 pounds per square inch. The DC generators **6** and **7** supply a series of electrical batteries **23**. The batteries feed a DC motor **16**. The DC motor is in turn connected to a second compressor **17**. The second compressor **17** is similar to the first compressor **5** and is connected through to pipe **18** to the tank inlet pipe **9**. A third compressor **19** similar to the first and second compressors is also connected to the tank inlet pipe **9** through pipe **20**. The third compressor **19** is powered by an AC motor **21**.

A pressure limit switch assembly **14** senses the pressure in the holding tank through a pipe **13**. A high pressure switch within the assembly **14** is activated when the holding tank reaches the maximum safely allowable pressure. This switch through line **15** causes the disengagement of the clutch **4** and turns off DC motor **16** and AC motor **21**. A second switch within the assembly **14** is activated when the holding pressure falls below a preset limit.

This second switch through line **15** turns on the DC motor **16**. It can now be seen that when the tank pressure is below the lowest limit, both the first and second compressors **15**, **17** will be activated. When the tank pressure goes above the lowest preset limit, only the first compressor **5** will be activated. If the holding tank pressure reaches the maximum tolerable limit all the compressors will be deactivated. The engagement speed of the centrifugal clutch **4** is set to a level corresponding to the minimum power necessary to drive the first compressor **5** and the DC generators **6** and **7**. If the speed of the wind falls below that level, the shaft **3** will be free-running.

The holding tank **10** has a separate inlet **11** protected by a check valve **12**. The holding tank is connected to a turbine feed tank **30** through pipe **24** controlled by valve **25**. The turbine feed tank **30** is connected to the inlet of a turbine **33** through pipe **31** controlled by valve **32**. The turbine **33** is powered by the expansion of the compressed air supplied by the turbine feed tank **30**. The turbine **33** is similar to the compressed air motors used in certain

impactors and drills. The turbine drives an AC generator **35** designed to supply approximately five kilowatts of household current at 60 Hz and 110 volts. The turbine is turned on by means of the valve **32** controlled by an/off switch **36**. The speed of the turbine **33** is determined by the pressure of the air accumulated in the turbine tank **30**. The pressure is monitored by sensor **27** connected to the turbine feed tank **30** by pipe **26**. Sensor **27** contains a set of high and low limits. When the turbine feed tank pressure falls below the low limit, valve **25** is opened through control line **28**. When the pressure in the turbine feed tank **30** reaches the high limit, the valve **25** is closed. The high and low limit of sensors **27** are not fixed but subject to minor variations in response to the speed of the turbine **33**.

The speed of the turbine **33** and of the generator **35** is monitored by speed sensor **34**. The output of the speed sensor **34** is inversely proportional to the speed of the turbine **33**. The speed sensor signal **29** is fed to sensor **27**. If the output frequency of the generator **35** deviates from the required 60 Hz, the high and low limits of the sensor **27** are either increased or decreased. If the speed of the generator is slowed down by an increase in the load current, the high and low limits of sensor **27** are raised in order to raise the pressure in turbine feed tank **30**. The turbine **33** will respond to the pressure change by increasing its rotational speed. The output of the generator **35** is made available for use through lines **38** and **40** controlled by a switch **37**.

The pressure in the holding tank **10** may be boosted from two external sources. First, compressed air may be introduced through inlet **11**. Second, the AC motor **21** may be connected to an external source of electrical energy through lines **39** and **40** controlled by switch **37**. The external electrical source may also be applied to a battery charger **22** which supplies the series of batteries **23**. In an alternate version of the preferred embodiment, it is suggested that an AC/DC converter **41** be used to drive the DC motor **16** from the external electrical supply. In such a case, the AC motor **23** and the third compressor **19** are not necessary.

The power plant just described is primarily designed to be installed on board a camping trailer. This power plant will accumulate wind ("aeolian") energy during the periods when the wind is blowing or the trailer is in motion. The energy is stored in two forms. First, it is stored in the form of compressed air in the holding tank **10**. Second, it is stored in the form of DC current in the series of batteries **23**. Both storage media are ecologically clean. Furthermore, the electrical system can boost the power of the compressed air system during periods of heavy power drain or long use. For added convenience, the system can be refuelled from an external source of electrical energy such as a household outlet or from an external source of compressed air such as those found in service stations for use by vehicle drivers. It should be noted also that this power plant is versatile in that it can be driven not only from the movement of fluids such as air or water, but also from the movement of the vehicle. In the later case, the shaft **3** would be coupled directly to the wheel of the vehicle.

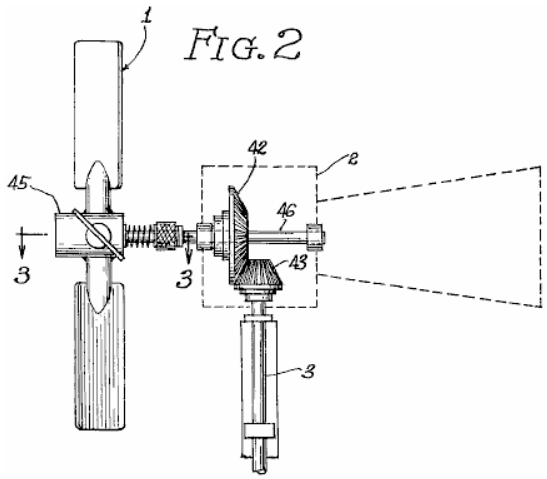


FIG. 2

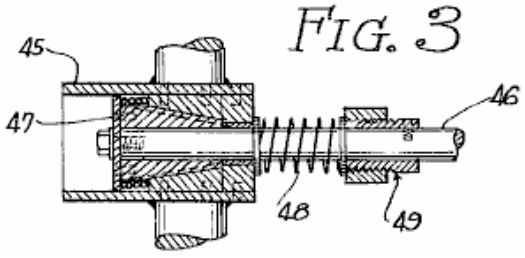


FIG. 3

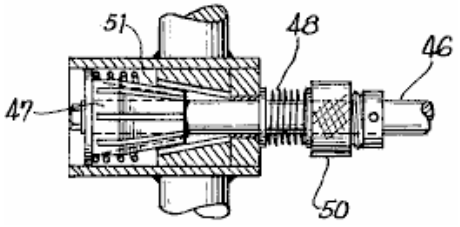


FIG. 4

Referring now to **Figs. 2** through **4**, there is shown the details of the propeller **1** and coupling box **2**. The propeller is noticeable by the fact that it is protected against bursts of wind which could damage the equipment. The hub **45** of propeller **1** is mounted on a shaft **46** by means of a conical spindle **46**. The hub has a central cavity **51** matching the outline of the spindle **47**. The hub **45** is held against the spindle by means of a coil spring **48** resting against an adjustable stop **49**. An excess of pressure of the wind against the propeller **1** will cause the hub **45** to be pulled back against the spring **48**, disengaging it from the spindle **47**. At that point the propeller 1 will rotate freely without driving the shaft **46**. The pressure of the coil spring **48** may be adjusted by turning the ring **50** around the threaded base of the stop **49**.

The various mechanical and electro-mechanical components of the power plant such as the centrifugal clutch, compressors, generators, turbines, valves and pressure-activated switches are well known to those skilled in the art.

The speed sensor **34** may be implemented with an electronic integrator whose output signal **29** amplitude is proportional to the frequency of AC generator **35**. The signal **29** is then used to modulate the sensitivity of sensor switches **27**. This technique is also well known to those skilled in the electro-mechanical arts.

Modifications, other than those suggested, can be made to the embodiment of the invention just described without departing from the spirit of the invention and the scope of the appended claims.

CLAIMS

1. A power plant which comprises:

(a) first rotating means responsive to movement of a fluid;
(b) first fluid compressor driven by the first rotating means;
(c) first means for coupling the first rotating means to the first fluid compressor;
(d) first electrical energy generator driven by the first rotating means;
(e) second means for coupling the first rotating means to the first generator;
(f) means for accumulating electrical energy generated by the first generator;
(g) second rotating means responsive to The accumulated energy;
(h) second fluid compressor driven by the second rotating means;
(i) means for storing compressed fluid;
(j) fluid conduit means for connecting the outputs of the first and second fluid compressors to the means for storing;
(k) means responsive to fluid pressure within the means for storing for controlling the operation of the first and second fluid compressors;
(l) third rotating means responsive to the expansion of compressed fluid;
(m) means for connecting the means for storing to the third rotating means;
(n) second electrical energy generator driven by third rotating means; and
(o) means for coupling the third rotating means to the second electrical energy generator.

2. The power plant claimed in claim 1 wherein the means for controlling the operation of the first and second fluid compressors comprise:

(a) first switch means responsive to high pressure for turning off the second rotating means and for inhibiting the first fluid compressor; and
(b) second switch means responsive to lower pressure for turning on the second rotating means.

3. The power plant claimed in claim 2 wherein the means for storing compressed fluid comprise:

(a) a high pressure tank;
(b) a low pressure tank;
(c) first valve means responsive to fluid pressure in the low pressure tank for regulating the flow of fluid from the high pressure tank to the low pressure tank; and
(d) the means for connecting the means for storing to the third rotating means comprise fluid conduit means and second valve means for controlling the flow of fluid.

4. The power plant claimed in claim 3 wherein The means for storing further comprise means responsive to the rotating speed of the third rotating means for controlling the first valve means.

5. The power plant claimed in claim 4 which further comprises:

(a) fourth rotating means responsive to electrical energy;
(b) third fluid compressor driven by the fourth rotating means;
(c) means for coupling the fourth rotating means to the third fluid compressor;
(d) means for connecting the third fluid compressor to the means for storing; and
(e) means for connecting the fourth rotating means to an external electrical energy source.

6. The power plant claimed in claim 4 wherein The means for accumulating comprise at least one electrical storage battery;

a battery charger connected to The battery; and
means for connecting The battery to an external electrical power source.

7. The power plant claimed in claim 1 wherein The first rotating means comprise: Lp1

(a) a rotating shaft;
(b) a conical spindle at one end of the shaft;
(c) a propeller having in its hub a conical hole engaging The spindle;
(d) means for resiliently holding the propeller engaged around The spindle; and
(e) means for adjusting the pressure of the means for holding against the propeller.

8. The power plant claimed in claim 4 wherein the first means for coupling comprise a centrifugal clutch.

9. The power plant claimed in claim 7 installed into a vehicle.

10. The power plant claimed in claim 9 wherein The high pressure tank comprises a means for connecting The tank to an outside source of compressed air;

A means for accumulating electrical energy comprises at least one electrical storage battery;
A second rotating means comprise a DC motor;
A third rotating means comprise a turbine powered by expansion of compressed air;
A second electrical energy generator comprise a generator of household alternating current; and
A means for distributing the household current to the vehicle electrical appliances.

The Motionless Generator of Richard Willis

This patent application covers a device which is claimed to have a substantially greater output power than the input power required to run it and it has no moving parts.

Patent application WO2009065210 (A1) 28th May 2009 Inventor: Richard Willis

ELECTRICAL GENERATOR

ABSTRACT

An electrical generator comprising an induction coil with a first magnet positioned adjacent to the first end of the induction coil so as to be in the electromagnetic influence of the induction coil when it is energised, and for creating a magnetic field around at least the first end of the induction coil. There is also a second magnet positioned near the second end of the induction coil so as to be in the electromagnetic field of the induction coil when the induction coil is energised, and for creating a magnetic field around at least the second end of the induction coil. A power input circuit powers the induction coil. A timer is placed in the power input circuit in order to create electrical pulses and controlling their timing. A power output circuit receives power from the induction coil.

FIELD OF THE INVENTION

The present invention relates to an electrical power generator, and more particularly to an "over-unity" electrical power generator.

BACKGROUND OF THE INVENTION

Electricity is conventionally generated in a number of ways, including fossil fuel powered electromechanical generators, coal powered electromechanical generators, water-flow powered electromechanical generators, nuclear reactor type generators, and so on. In each case, there are a number of disadvantages associated with these methods, especially inefficiency and also the scarcity of a power source.

Recently, magnetic generators have been developed which produce electrical power from the magnetic field of the Earth. Basically, an input magnetic field is quickly switched on and off, or alternatively more than one input magnetic field is selectively switched on and off, on an alternating basis, to influence a larger magnetic field in an electromagnetic apparatus that is selectively connected to an electrical power output circuit. A resulting electrical power is produced in the power output circuit.

There are even magnetic generator circuits which produce more electrical power than that which is applied to the circuit. While this seems to contradict the laws of physics, it docs not, otherwise, such magnetic generator circuits would not work. These magnetic generator circuits work, on the basic principle that the space-time continuum is very energetic, including energy fields such as the Earth's magnetic field.

It should be understood that electric fields and magnetic fields do not have an independent existence. A purely electromagnetic field in one coordinate system can appear as a mixture of electric and magnetic fields in another coordinate system. In other words, a magnetic field can at least partially turn into an electric field, or vice versa.

It is also well known that a system which is far from equilibrium in it's energy exchange with it's environment can steadily and freely receive environmental energy and dissipate it in external loads. Such a system, can have a Coefficient of Performance ("COP") greater than 1. For a COP greater than 1, an electrical power system must take some, or all of its input energy, from it's active external environment. In other words, the system must be open to receive and convert energy from it's external environment, as opposed to merely converting energy from one form to another.

The US Patent 6,362,718 issued on 26th March 2002 to Patrick et at., discloses an electromagnetic generator without moving parts. This electromagnetic generator includes a permanent magnet mounted within a rectangular

ring-shaped magnetic core having a magnetic path to one side of the permanent magnet and a second magnetic path to the other side of the permanent magnet. A first input coil and a first output coil extend around portions of the first magnetic path, with the first input coil being at least partially positioned between the permanent magnet and the first output coil. A second input coil and a second output coil extend around portions of the second magnetic path, with the second input coil being at least partially positioned between the permanent magnet and the second output coil. The input coils are alternatively pulsed by a switching and control circuit and provide induced current pulses in the output coils. Driving electrical current through each of the input coils reduces a level of flux from the permanent magnet within the magnet path around which the input coil extends.

In an alternative embodiment of the Patrick et al electromagnetic generator, the magnetic core includes circular spaced-apart plates, with posts and permanent magnets extending in an alternating fashion between the plates. An output coil extends around each of these posts. Input coils extending around portions of the plates are pulsed to cause the induction of current within the output coils.

The apparent problems with the electric magnetic generator is disclosed in US Patent 6,362,718 seem to be twofold. First, it is more expensive to produce than necessary as it has four coils. Secondly, while it apparently achieves a Coefficient of Performance of more than 3.0, a much greater Coefficient of Performance is readily achievable. This is believed to be due to the specific physical configuration of the magnetic paths.

It is an object of the present invention to provide an electrical generator having a Coefficient of Performance significantly greater than 1.

SUMMARY OF THE INVENTION

In accordance with one aspect of the present invention there is disclosed a novel electrical generator comprising an induction coil. There is a first magnet positioned beside the first end of the induction coil so as to be in the electro-magnetic field of the induction coil when the induction coil is energised, and for creating a magnetic field around at least the first end of the induction coil. There is also a second magnet positioned near the second end of the induction coil so as to be in the electro-magnetic field of the induction coil when the induction coil is energised, and for creating a magnetic field around at least the second end of the induction coil. A power input circuit provides power to the induction coil. A timing device is placed in the input power circuit in order to create electrical pulses and for controlling the timing of those electrical pulses being passed to the induction coil. A power output circuit receives power from the induction coil.

Other advantages, features and characteristics of the present invention, as well as methods of operation and functions of the related elements of the structure, and the combination of parts and economies of manufacture, will become more apparent upon consideration of the following detailed description and the appended claims with reference to the accompanying drawings which are described here:

BRIEF DESCRIPTION OF THE DRAWINGS

The novel features which are believed to be characteristic of the electrical generator according to the present invention, as to its structure, organisation, use and method of operation, together with it's further objectives and advantages, will be better understood from the following drawings in which a preferred embodiment of the invention will now be illustrated by way of example. It is expressly understood, however, that the drawings are for the purpose of illustration and description only, and are not intended as a definition of the limits of the invention. In the accompanying drawings:

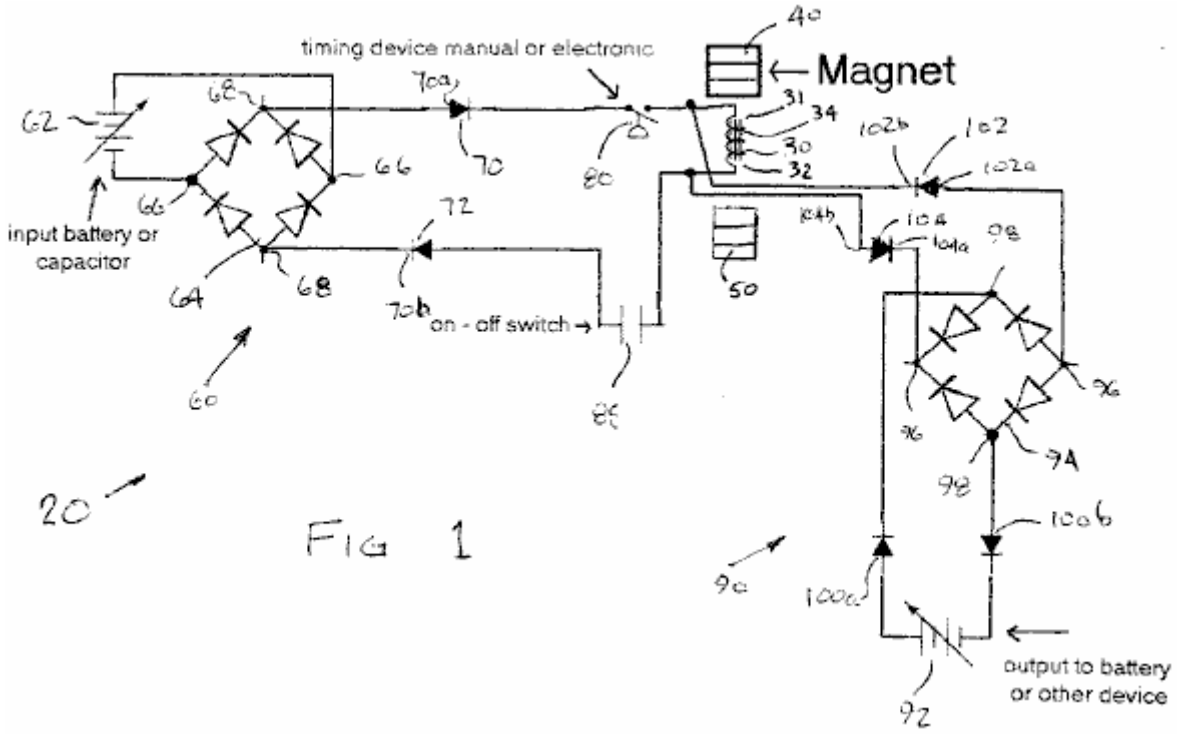


Fig.1 is an electrical schematic of the first preferred embodiment of the electrical generator.

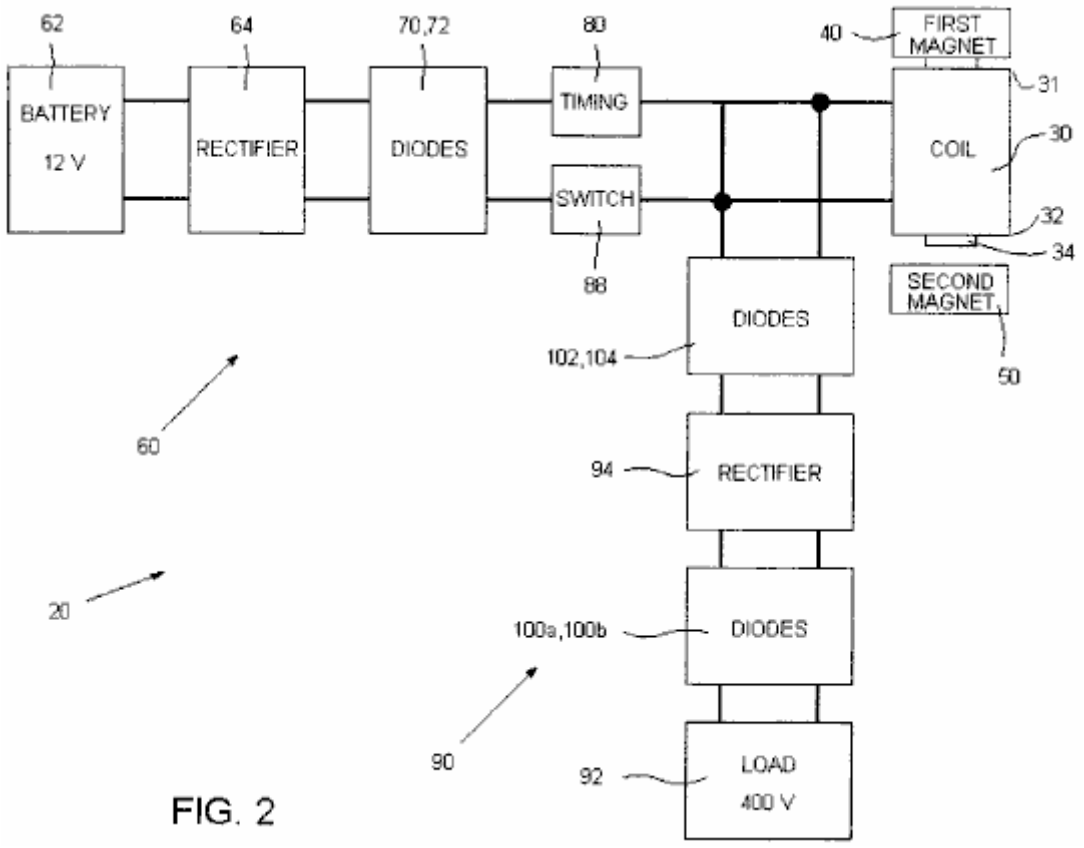


FIG. 2

Fig.2 is a block diagram schematic of the first preferred embodiment of the electrical generator of **Fig.I**.

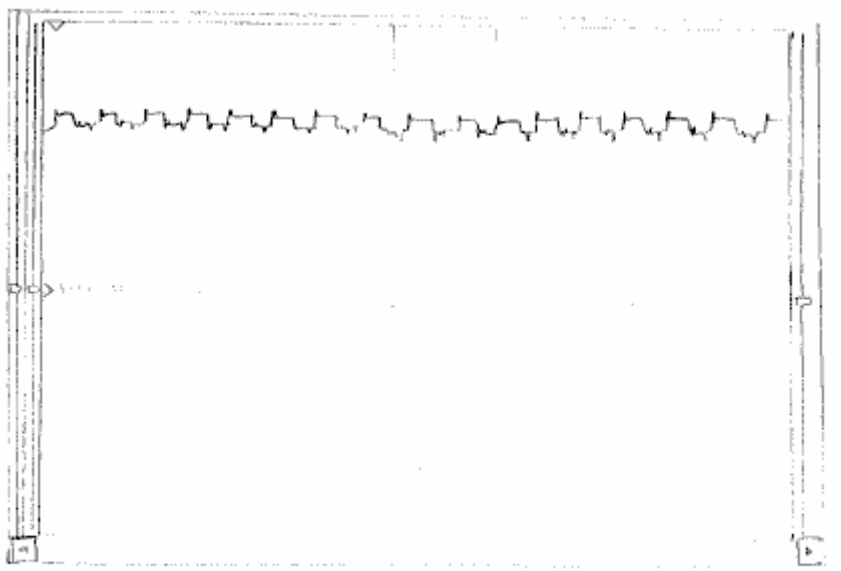


FIG. 3

Fig.3 is an oscilloscope waveform taken at the input power circuit after the timing mechanism.

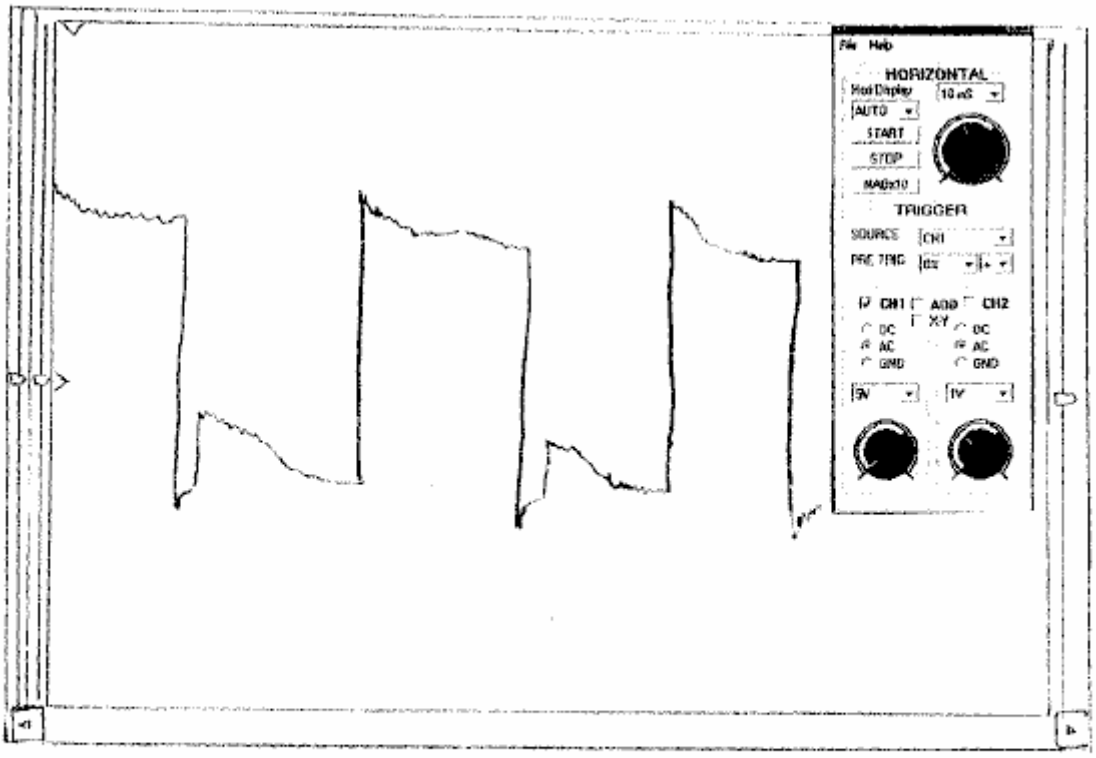


FIG. 4

Fig.4 is an oscilloscope waveform taken at the output power circuit before the first set of diodes immediately after the coil.

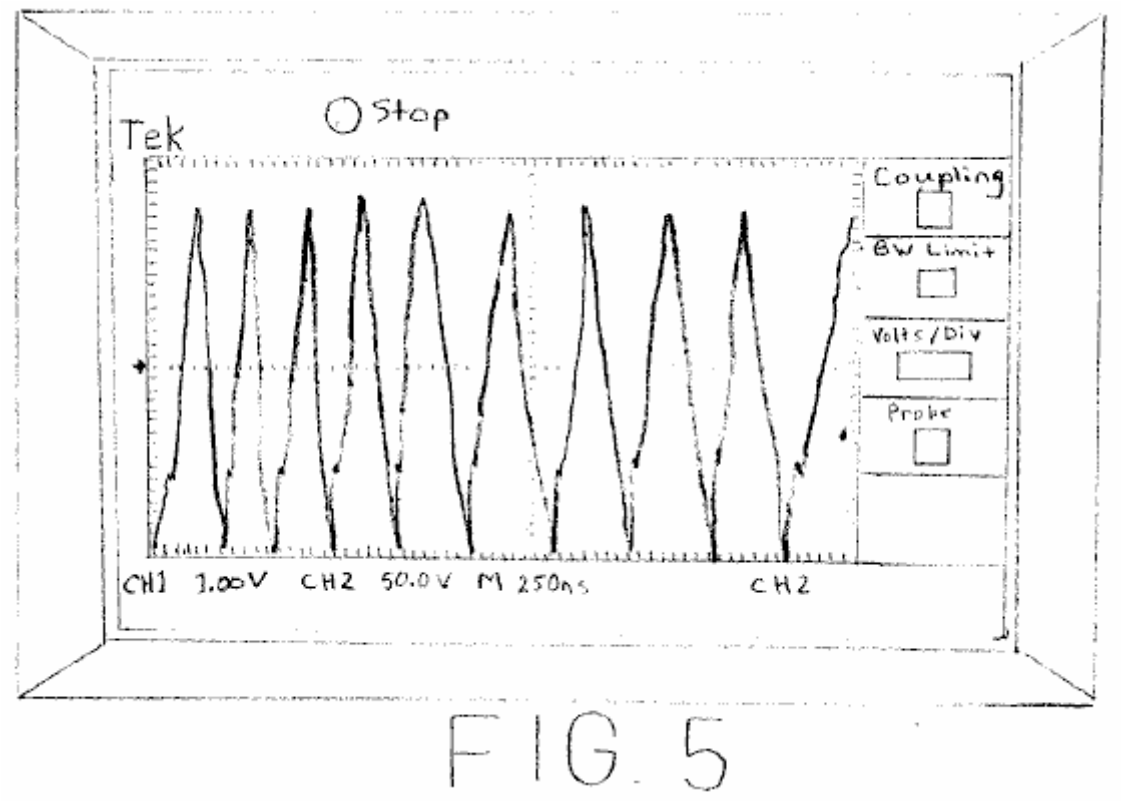


FIG. 5

Fig.5 is an oscilloscope waveform taken at the output power circuit at the load; and,

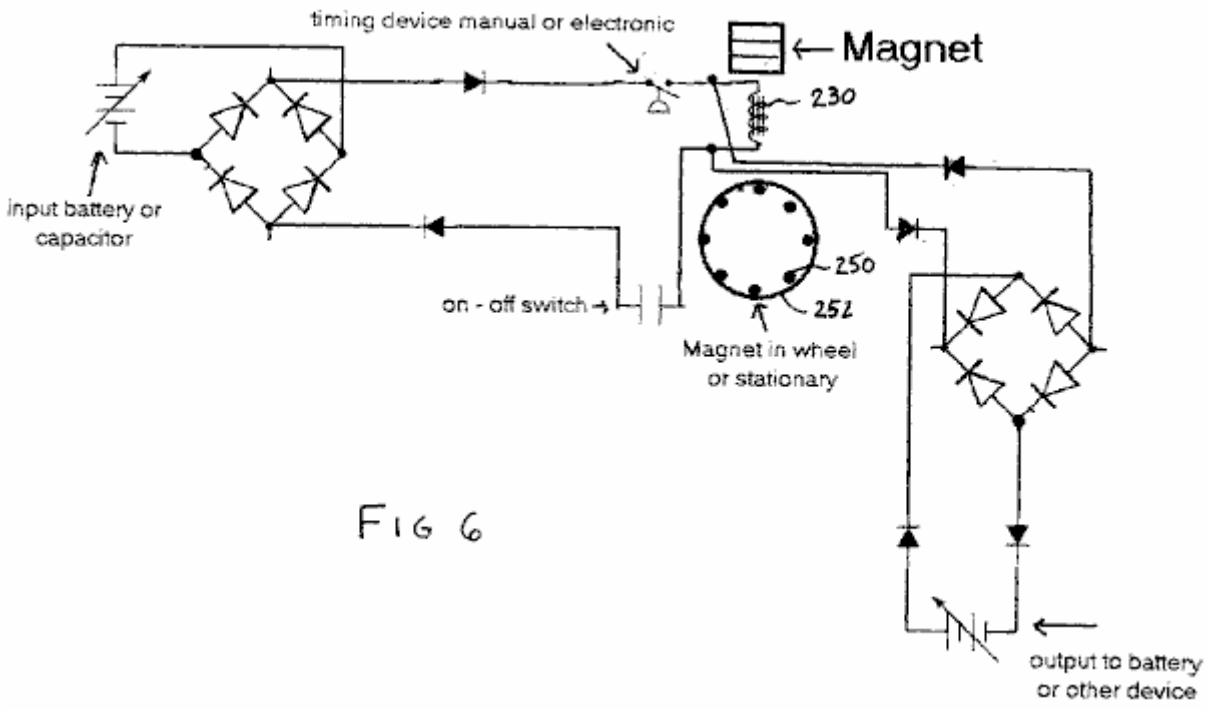


FIG 6

Fig.6 is an electrical schematic of the second preferred embodiment of the electrical generator

.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

Referring to **Fig.1** through **Fig.6** of the drawings, it will be noted that **Fig.1** through **Fig.5** illustrate a first preferred embodiment of the electrical generator of the present invention, and **Fig.6** illustrates a second preferred embodiment of the electrical generator of the present invention.

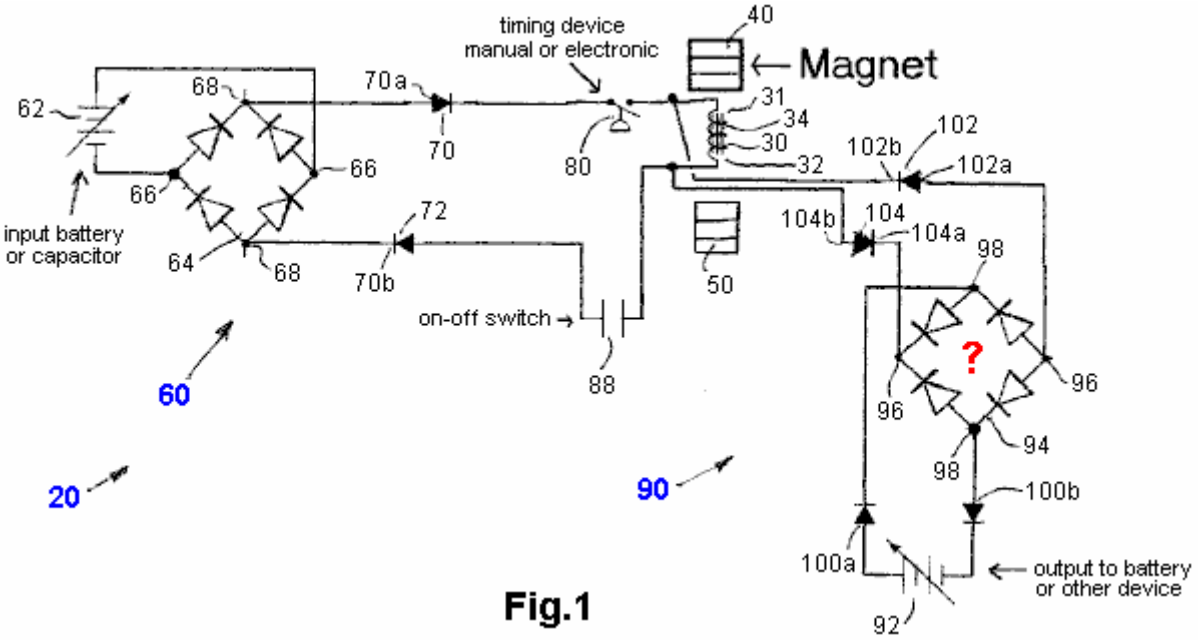


Fig.1

Reference will now be made to **Fig.1** through **Fig.5**, which show a first preferred embodiment of the electrical generator of the present invention, as indicated by general reference numeral **20**. The electrical generator **20** comprises an induction coil **30** having a first end **31** and a second end **32**. The induction coil **30** preferably includes a core **34** which is made from any suitable type of material, such as ferrite, mumetal, permalloy, cobalt, any non-permeable metal material, or any other suitable type of material. The coil **30** is wound with copper wire which can be a single size or multiple sizes depending on the size of the ferrite core **34**.

There is a first magnet **40** positioned adjacent to the induction coil **30**, preferably at the first end **31** so as to be within the electromagnetic field of the induction coil **30** when the induction coil **30** is energised. The first magnet **40** is a permanent magnet which has its North pole facing the first end **31** of the induction coil **30**. In the first preferred embodiment, the first magnet **40** is stationary with respect to the induction coil **30**, and even more preferably is in contact with, or is even secured to, the first end **31** of the induction coil **30**. The size of the coil and the copper wire used to wind the coil also depend on the size of the first magnet **40**. The first magnet **40** is there to create a magnetic field around at least the first end **31** of the first magnet **30**.

There is also a second magnet **50** positioned adjacent to the induction coil **30**, preferably at the second end **32** of the induction coil **30** but at a distance of about 1.0 cm or so from the coil core **34** but within the electromagnetic field of the induction coil **30** when the induction coil **30** is energised. The gap between the second end **32** of the induction coil **30** and the second magnet **50** can be an air gap or can be a vacuum.

The second magnet **50** is a permanent magnet which has it's North pole facing the second end **32** of the induction coil **30**. In the first preferred embodiment, the second magnet **50** is stationary with respect to the induction coil **30**. The size of the coil and the copper wire used to wind it also depends on the size of the second magnet **50**. The second magnet **50** is there in order to create a magnetic field around at least the second end **32** of the induction coil **30**.

As can be seen in **Fig.1**, the first magnet **40** is positioned so it's North pole is facing the first end **31** of the induction coil and its South pole is facing away from the first end **31** of the induction coil **30**. The first end **31** of the induction coil **30** creates a South magnetic field when it is energised. In this manner, the North pole of the first magnet **40** and the South pole of the first end **31** of the induction coil attract each other.

Similarly, but oppositely, the second magnet **50** is positioned so that it's North pole is facing the second end **32** of the induction coil and its South pole is facing away from the second end **32** of the induction coil **30**. The second end **32** of the induction coil **30** creates a North magnetic field when the induction coil **30** is energised. In this manner, the North pole of the second magnet **50** and the North pole of the second end **32** of the induction coil repel each other.

A power input circuit section, as indicated by the general reference numeral **60**, is for providing power to the induction coil and is comprised of a source of electrical power **62**. In the first preferred embodiment, as illustrated, the input source of electrical power **62** comprises a DC power source, specifically a battery **62**, but additionally or alternatively may comprise a capacitor (not shown). The source of electrical power can range from less than 1.0 volt to more than 1,000,000 volts, and can range from less than 1.0 amp to more than 1 million amps. Alternatively, it is contemplated that the input source of electrical power could be an AC power source (not shown).

An input rectifier **64** which is preferably, but not necessarily, a full-wave rectifier **64**, has an input **66** electrically connected to the source of electrical power **62** and also has an output **68**. A first diode **70** is connected at its positive end **70a** to one terminal **68a** of the output **68** of the rectifier **62**. A second diode **72** is connected at its negative end **72a** to the other terminal **68b** of the output **68** of the rectifier **62**.

There is also a timing mechanism **80** in the input power circuit section **60**, which as shown, is electrically connected in series with the first diode **70**. This timing mechanism both creates electrical pulses and controls the timing of those electrical pulses which are fed to the induction coil **30**. The pulses are basically saw-tooth waveforms, as can be seen in **Fig.3**.

In the first preferred embodiment, the timing device **80** is a manual timer in the form of a set of "points" from the ignition system of a vehicle, as they can withstand high voltage and high current levels. Alternatively, it is contemplated that the timing mechanism could be an electronic timing circuit. It is also contemplated that a TGBT unit from a MIG welder could be used as the basis of the timing device **80**. It has been found that a timing device which provides a physical break in its "off" configuration works well as stray currents cannot backtrack through the circuit at that time. The timing mechanism can be of any suitable design so long as it can respond to the placement of the magnets **50** in the rotor **52** in the second preferred embodiment shown in **Fig.6**.

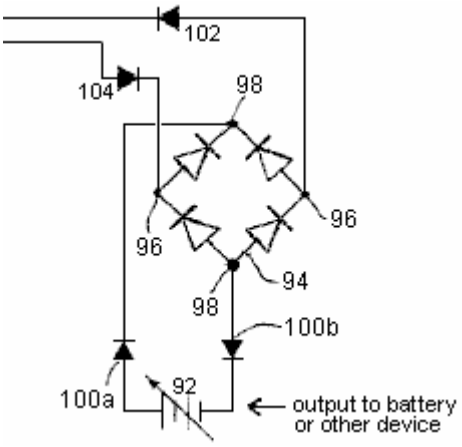
When the device is in use, the magnetic fields created by the first magnet 40 and the second magnet **50** in conjunction with the coil **30**, are each somewhat mushroom shaped, and oscillate back and forth, with respect to their size, in a manner corresponding to the timing of the electrical pulses from the power input circuit **60**, as controlled by the timing mechanism **80**.

The power input circuit **60** has an on/off switch **88** to allow disconnection of the power feed to the induction coil **30**. The on/off switch **88** may alternatively be located in any other suitable place in the power input circuit **60**.

A power output circuit section, indicated by the general reference numeral **90**, is for receiving power from the induction coil and comprises an electrical load **92**, which, in the first preferred embodiment is a battery **92**, but may additionally or alternatively comprise a capacitor (not shown), or any other suitable electrical load device.

The power output circuit portion **90** also has an output rectifier **94** having an input **96** an output **98** electrically connected to the electrical load **92** via a pair of forward biased diodes **100a, 100b** which prevent the electrical load **92** from powering the induction coil **30**. A first diode **102** is electrically connected at its positive end **102a** to one terminal **94a** of the input of the rectifier **94** and is electrically connected at its negative end **102b** to one end of the induction coil **30**. A second diode **104** is connected at its negative end **104a** to the other terminal **94b** of the input of the rectifier **94** and is electrically connected at its positive end **104b** to the other end of the induction coil **30**. The output of the coil, taken before the diodes **102,104** is shown in **Fig.4**.

Note: It is highly likely that there is a clerical error in **Fig.1** because as it is drawn the bridge input is point **98** and not **96** as stated. If this is the case, then the two diode bridges are identical and the output section should be drawn like this:



although it is by no means obvious why diodes 102 and 104 are needed as their function would appear to be provided by the output bridge diodes.

The output to the electrical load **92** of the power output circuit **90** can range from less than 1 volt to more than 1,000,000 volts, and can range from less than one amp to more than 1 million amps. As can be seen in **Fig.5**, the output to the electrical load **92** comprises generally spike-shaped pulses which have both negative and positive components.

As can be readily seen in **Fig.1** and **Fig.2**, the input power circuit **60** is electrically connected in parallel with the induction coil **30** and the output power circuit portion **90** is electrically connected in parallel with the induction coil **30**.

The various diodes and rectifiers in the electrical generator **20** can be of any suitable voltage from about 12 volts to over 1,000,000 volts, and can have slow recovery or fast recovery, as desired. Further, the various diodes and rectifiers may be configured in other suitable formats. There also may be additional capacitors added into the power output circuit adjacent to the electrical load **92** in order to increase the output power before discharge.

It has been found that setting the timing to six hundred pulses per minute (10 Hz) provides a waveform in the power output circuit portion 90 that comprises generally spike-shaped pulses with a period of about 20 nanoseconds. It is believed that the flux of the power pulses that are input into the induction coil **30** is quickly shifting the magnetic field back and forth in the induction coil **30**, which is akin to the flux of the power pulses creating its own echo. The various electromagnetic oscillations in the coil provide a much higher frequency in the power output circuit **90** than in the power input circuit portion **60**.

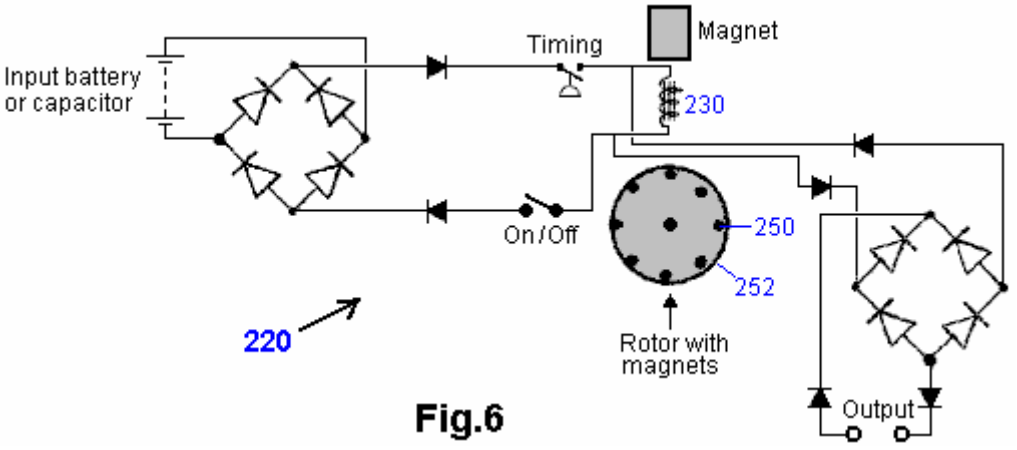


Fig.6

Reference will now be made to **Fig.6**, which shows a second preferred embodiment of the electrical generator of the present invention, as indicated by general reference numeral **220**. The second preferred embodiment electrical generator is similar to the first preferred embodiment electrical generator **20** except that the second magnet comprises several moving magnets **250**, typically eight permanent magnets **250**. These magnets are mounted on a wheel **252**, which is free to rotate. Ideally, these magnets are mounted in an identical way to each other on the rotor disc **252**. If desired, there can be any suitable number of magnets mounted in the rotor. Accordingly, at least one rotor magnet **250** will be within the electromagnetic field of the induction coil **230** when the coil is energised. The rotor magnets can be of any suitable strength and any suitable type of magnet, and they may be mounted on the rotator by any suitable means, such as a suitable adhesive, or moulded into the disc

if the rotor is made of plastic. In practice, the rotor disc is driven round by the magnetic field of the induction coil when it is energised. It is also possible for the first magnet to a rotor magnet in the same manner as described for the second magnet **250**.

As can be understood from the above description and from the accompanying drawings, the present invention provides an electrical generator having a Coefficient of Performance greater than 1.0. and more specifically, an electrical generator which has a Coefficient of Performance significantly greater than 1.0. An electrical generator having a Coefficient of Performance significantly greater than 1.0 is at present, unknown in the prior art.

Other variations of the above principles will be apparent to those who are knowledgeable in the field of the invention, and such variations are considered to be within the scope of the present invention. Further, other modifications and alterations may be used in the design and manufacture of the electrical generator of the present invention without departing from the spirit and scope of the following claims:

CLAIMS

1. An electrical generator comprising:
an induction coil having a first end and a second end;
a first magnet positioned adjacent said first end of said induction coil so as to be in the electromagnetic field of said induction coil when said induction coil is energised, and for creating a magnetic field around at least said first end of said induction coil,
a second magnet positioned adjacent said second end of said induction coil so as to be in the electro-magnetic field of said induction coil when said induction coil is energized, and for creating a magnetic field around at least said second end of said induction coil;
a power input circuit portion for providing power to said induction coil;
a liming means in said power input circuit portion for creating electrical pulses and controlling the timing of said electrical pulses to said induction coil; and,
a power output circuit portion for receiving power from said induction coil.

2. The electrical generator of claim 1 , wherein said first magnet is stationary with respect to said induction coil.

3. The electrical generator of claim 2, wherein said first magnet comprises a permanent magnet.

4. The electrical generator of claim 2, wherein said induction coil includes a core.

5. The electrical generator of claim 4, wherein said first magnet is in contact with said core.

6. The electrical generator of claim 4, wherein said core is made from a material chosen from the group of ferrite, mumetal, permalloy, and cobalt.

7. The electrical generator of claim 4, wherein said core is made from a non-permeable metal material.

8. The electrical generator of claim 3, wherein said second magnet is stationary with respect to said induction coil.

9. The electrical generator of claim 8, wherein said second magnet comprises a permanent magnet.

10. The electrical generator of claim 1, wherein said second magnet comprises at least one movable magnet.

11. The electrical generator of claim 10. wherein said at least one movable magnet is mounted on a rotor.

12. The electrical generator of claim 11 , wherein said at least one movable magnet comprises a plurality of magnets mounted on said rotor.

13. The electrical generator of claim 1, wherein said power input circuit portion comprises a source of electrical power, a input rectifier having an input electrically connected to said source of electrical power and an output, a first diode connected at its positive end to one terminal of said input rectifier, a second diode connected at its negative end to the other terminal of said input rectifier.

14. The electrical generator of claim 13, wherein said timing means is electrically connected in series with said first diode.

15. The electrical generator of claim 14, wherein said power output circuit portion comprising an electrical load, an output rectifier having an output electrically connected to said electrical load via a pair of forward biased diodes and an input, a first diode connected at its negative end to one terminal of said output rectifier, a

second diode connected at its positive end to the other terminal of said output rectifier.

16. The electrical generator of claim 15, wherein said input power circuit portion is electrically connected in parallel with said induction coil and said output power circuit portion is electrically connected in parallel with said induction coil.

17. The electrical generator of claim 1, wherein said input source of electrical power comprises a DC power source.

18. The electrical generator of claim 17, wherein said DC power source comprises a battery.

19. The electrical generator of claim 17, wherein said DC power source comprises a capacitor.

20. The electrical generator of claim 1 , wherein said input source of electrical power comprises an AC power source.

21. The electrical generator of claim 1 where the input rectifier is a Wheatstone bridge rectifier.

22. The electrical generator of claim 1 , wherein said timing means comprises an electronic timing circuit.

23. The electrical generator of claim 1 , wherein said timing means comprises a manual timer.

24. The electrical generator of claim 1, wherein said first magnet comprises a permanent magnet.

25. (Appears to have been omitted from the archived copy)

26. The electrical generator of claim 12, wherein said plurality of movable magnets are each mounted similarly one to another on said rotatable wheel.

27. The electrical generator of claim 1 , wherein said electrical load comprises a battery.

28. The electrical generator of claim 1 , further comprising an on/off switch electrically connected in said power input circuit portion.

The Motionless Generator of Graham Gunderson

Graham Gunderson's Solid-State Electric Generator is shown in US Patent Application 2006/0163971 A1 of 27th July 2006. The details are as follows:

Abstract
A solid-state electrical generator including at least one permanent magnet, magnetically coupled to a ferromagnetic core provided with at least one hole penetrating its volume; the hole(s) and magnet(s) being placed so that the hole(s) intercept flux from the permanent magnet(s) coupled into the ferromagnetic core. A first wire coil is wound around the ferromagnetic core for the purpose of moving the coupled permanent magnet flux within the ferromagnetic core. A second wire is routed through the hole(s) penetrating the volume of the ferromagnetic core, for the purpose of intercepting this moving magnetic flux, thereby inducing an output electromotive force. A changing voltage applied to the first wire coil causes coupled permanent magnet flux to move within the core relative to the hole(s) penetrating the core volume, thus inducing electromotive force along wire(s) passing through the hole(s) in the ferromagnetic core. The mechanical action of an electrical generator is therefore synthesised without the use of moving parts.

Background
This invention relates to a method and device for generating electrical power using solid state means.

It has long been known that moving a magnetic field across a wire will generate an electromotive force (EMF), or voltage, along the wire. When this wire is connected in a closed electrical circuit, an electric current, capable of performing work, is driven through this closed circuit by the induced electromotive force.

It has also long been known that this resulting electric current causes the closed circuit to become encircled with a secondary, induced magnetic field, whose polarity opposes the primary magnetic field which first induced the EMF. This magnetic opposition creates mutual repulsion as a moving magnet approaches such a closed circuit, and a mutual attraction as that moving magnet moves away from the closed circuit. Both these actions tend to slow or cause "drag" on the progress of the moving magnet, causing the electric generator to act as a magnetic brake, whose effect is in direct proportion to the amount of electric current produced.

Historically, gas engines, hydroelectric dams and steam-fed turbines have been used to overcome this magnetic braking action which occurs within mechanical generators. A large amount of mechanical power is required to produce a large amount of electrical power, since the magnetic braking is generally proportional to the amount of electrical power being generated.

There has long been felt the need for a generator which reduces or eliminates the well-known magnetic braking interaction, while nevertheless generating useful electric power. The need for convenient, economical and powerful sources of renewable energy remains urgent. When the magnetic fields within a generator are caused to move and interact by means other than applied mechanical force, electric power can be supplied without the necessity of consuming limited natural resources, thus with far greater economy.

Summary of the Invention
It has long been known that the source of the magnetism within a permanent magnet is a spinning electric current within ferromagnetic atoms of certain elements, persisting indefinitely in accord with well-defined quantum rules. This atomic current encircles every atom, thereby causing each atom to emit a magnetic field, as a miniature electromagnet.

This atomic current does not exist in magnets alone. It also exists in ordinary metallic iron, and in any element or metallic alloy which can be "magnetised", that is, any material which exhibits ferromagnetism. All ferromagnetic atoms and "magnetic metals" contain such quantum atomic electromagnets.

In specific ferromagnetic materials, the orientation axis of each atomic electromagnet is flexible. The orientation of magnetic flux both internal and external to the material, pivots easily. Such materials are referred to as magnetically "soft", due to this magnetic flexibility.

Permanent magnet materials are magnetically "hard". The orientation axis of each is fixed in place within a rigid crystal structure. The total magnetic field produced by these atoms cannot easily move. This constraint aligns the field of ordinary magnets permanently, hence the name "permanent".

The axis of circular current flow in one ferromagnetic atom can direct the axis of magnetism within another ferromagnetic atom, through a process known as "spin exchange". This gives a soft magnetic material, like raw

iron, the useful ability to aim, focus and redirect the magnetic field emitted from a magnetically hard permanent magnet.

In the present invention, a permanent magnet's rigid field is sent into a magnetically flexible "soft" magnetic material. the permanent magnet's apparent location, observed from points within the magnetically soft material, will effectively move, vibrate, and appear to shift position when the magnetisation of the soft magnetic material is modulated by ancillary means (much like the sun, viewed while underwater, appears to move when the water is agitated). By this mechanism, the motion required for generation of electricity can be synthesised within a soft magnetic material, without requiring physical movement or an applied mechanical force.

The present invention synthesises the virtual motion of magnets and their magnetic fields, without the need for mechanical action or moving parts, to produce the electrical generator described here. The present invention describes an electrical generator where magnetic braking known as expressions of Lenz's Law, do not oppose the means by which the magnetic field energy is caused to move. The synthesised magnetic motion is produced without either mechanical or electrical resistance. This synthesised magnetic motion is aided by forces generated in accordance with Lenz's Law, in order to produce acceleration of the synthesised magnetic motion, instead of physical "magnetic braking" common to mechanically-actuated electrical generators. Because of this novel magnetic interaction, the solid-state static generator of the present invention is a robust generator, requiring only a small electric force of operate.

Brief Description of the Drawings

The appended drawings illustrate only typical embodiments of this invention and are therefore not to be considered limiting of its scope, as the invention encompasses other equally effective embodiments.

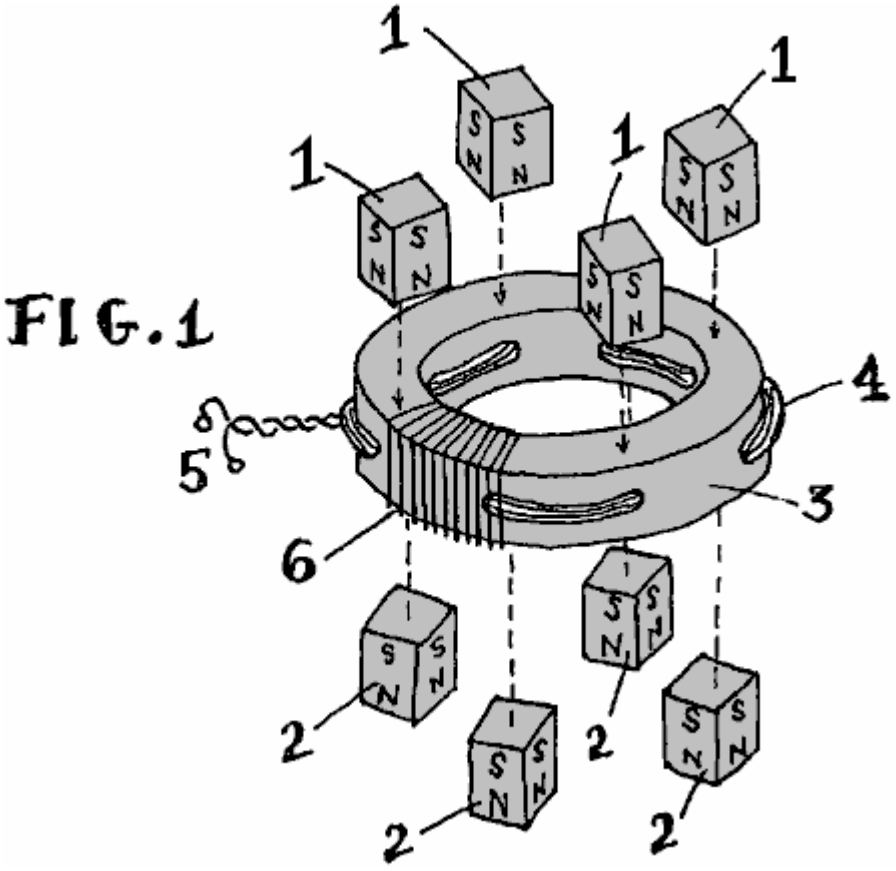


Fig.1 is an exploded view of the generator of this invention.

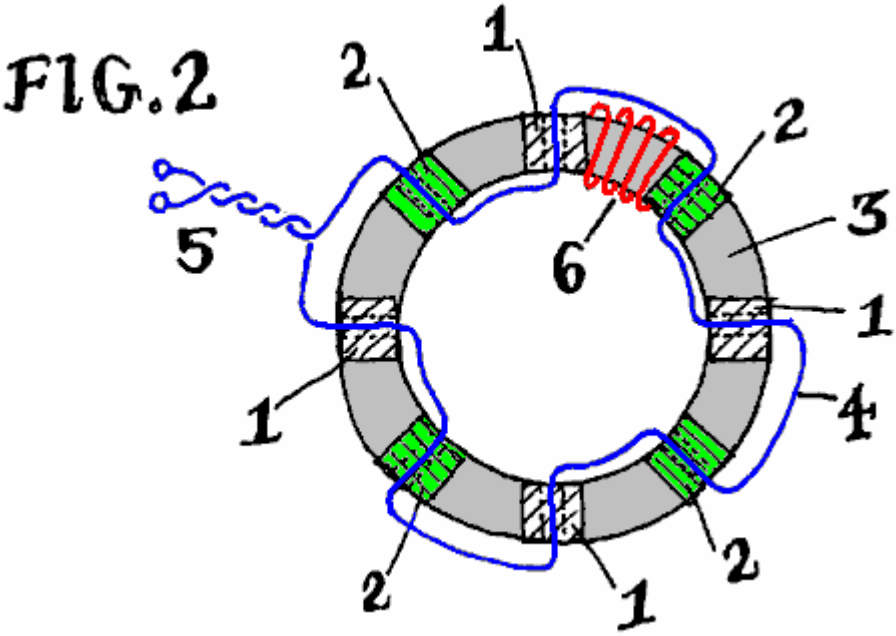


Fig.2 is a cross-sectional elevation of the generator of this invention.

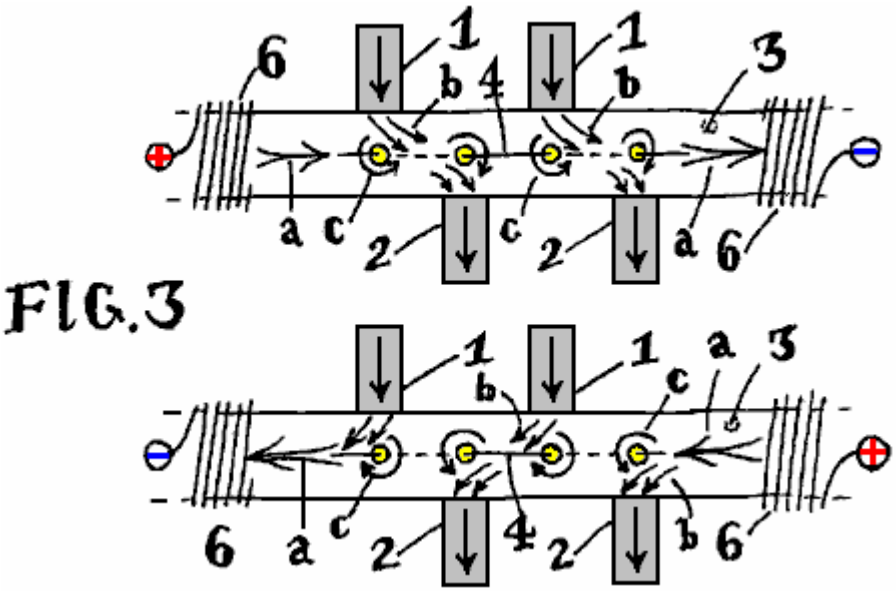


Fig.3 is a schematic diagram of the magnetic action occurring within the generator of **Fig.1** and **Fig.2**.

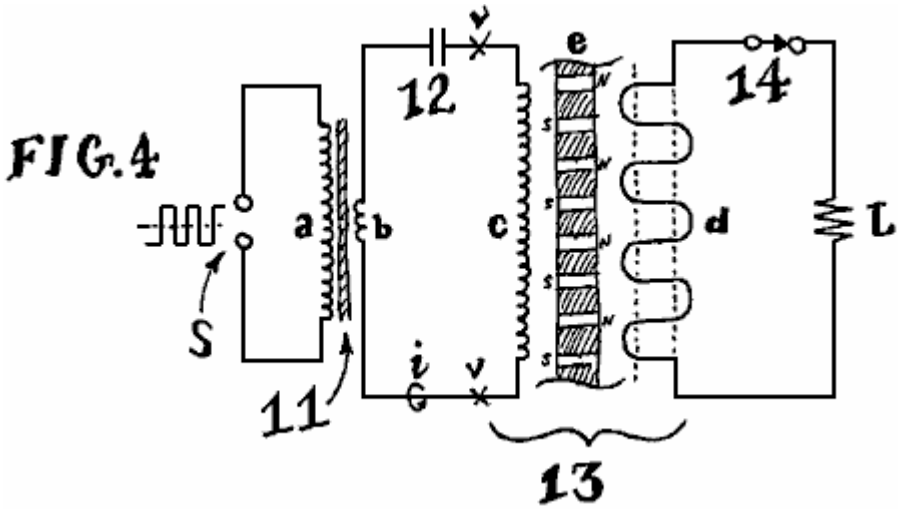
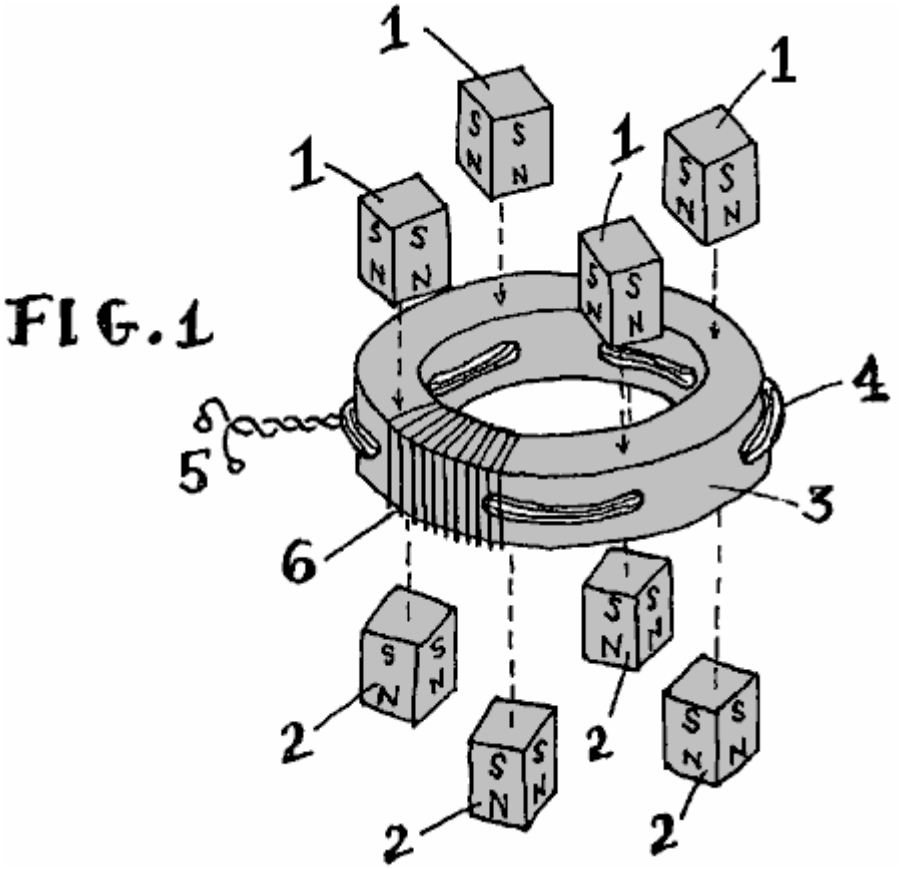


Fig.4 is a circuit diagram, illustrating one method of operating the electrical generator of this invention.

Detailed Description of the Invention

Fig.1 depicts a partially exploded view of an embodiment of an electrical generator of this invention. The part numbers also apply in **Fig.2** and **Fig.3**.



Numeral **1** represents a permanent magnet with it's North pole pointing inward towards the soft ferromagnetic core of the device. Similarly, numeral **2** indicates permanent magnets (preferably of the same size, shape and composition), with their South poles aimed inward towards the opposite side, or opposite surface of the device. The letters "**S**" and "**N**" denote these magnetic poles in the drawings. Other magnetic polarities and configurations may be used with success; the pattern shown merely illustrates one efficient method of adding magnets to the core.

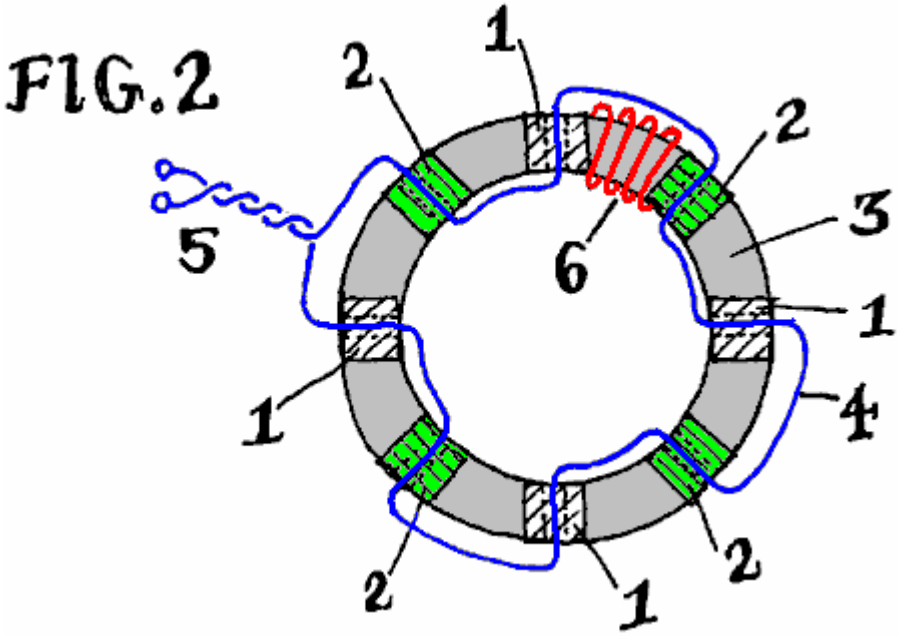
The magnets may be formed of any polarised magnetic material. In order of descending effectiveness, the most desirable permanent magnet materials are Neodymium-Iron-Boron ("NIB"), Samarium Cobalt, AlNiCo alloy, or

"ceramic" Strontium-Barium or Lead-Ferrite. A primary factor determining permanent magnet material composition is the magnetic flux strength of the particular material type. In an embodiment of the invention, these magnets may also be substituted with one or more electromagnets producing the required magnetic flux. In another embodiment of the invention, a superimposed DC current bias can be applied to the output wire to generate the required magnetic flux, replacing or augmenting the permanent magnets.

Numeral **3** indicates the magnetic core. This core is a critical component of the generator. The core determines the output power capacity, the optimum magnet type, the electrical impedance and the operating frequency range. The core may be any shape, composed of any ferromagnetic material, formed by any process (sintering, casting, adhesive bonding, tape-winding, etc.). A wide range of shapes, materials and processes is known in the art of making magnetic cores. Effective common materials include amorphous metal alloys (such as sold under the "Metglas" trademark by Metglas Inc., Conway, S.C.), nanocrystalline alloys, manganese and zinc ferrites as well as ferrites of any suitable element including any combination of magnetically "hard" and "soft" ferrites, powdered metals and ferromagnetic alloys, laminations of cobalt and/or iron and silicon-iron "electrical steel". This invention successfully utilises any ferromagnetic material, while functioning as claimed. In an embodiment of the invention, and for the purpose of illustration, a circular "toroid" core is illustrated. In an embodiment of the invention, the composition may be bonded iron powder, commonly available from many manufacturers.

Regardless of core type, the core is prepared with holes, through which, wires may pass. the holes are drilled or formed to penetrate the core's ferromagnetic volume. The toroidal core **3** shown, includes radial holes pointing towards a common centre. If, for example, stiff wire rods were to be inserted through each of these holes, these rods would meet at the centre point of the core, producing an appearance similar to a wheel with spokes. If a square or rectangular core (not illustrated) is used, then these holes are preferably oriented parallel to the core's flat sides, causing stiff rods passed through the holes to form a square grid pattern, as the rods cross each other in the interior "window" area framed by the core. While in other embodiments of the invention, these holes may take any possible orientation or patterns of orientation, a simple row of radial holes is illustrated as one example.

Numeral **4** depicts a wire, or bundle of wires which pick up and carry the output power of the generator. Typically, this wire is composed of insulated copper, though other materials such as aluminium, iron, dielectric material, polymers and semiconducting materials may be substituted. It may be seen in **Fig.1** and **Fig.2**, that wire **4** passes alternately through neighbouring holes formed in core **3**. The path taken by wire **4** undulates as it passes in opposite direction through each adjacent hole. If an even number of holes is used, the wire will emerge on the same side of the core on which it first entered. Once all the holes are filled, the resulting pair of trailing leads may be twisted together or similarly terminated, forming the output terminals of the generator shown at numeral **5**. Output wire **4**, may also make multiple passes through each hole in the core. Though the winding pattern is not necessarily undulatory, this basic form is shown as an example. Many effective connection styles exist. This illustration shows the most simple.



Numeral **6** in **Fig.1**, **Fig.2** and **Fig.3**, points to a partial illustration of the input winding, or inductive coil used to shift the fields of the permanent magnets, within the core. Typically, this wire coil encircles the core, wrapping around it. For the toroidal core shown, input coil **6** resembles the outer windings of a typical toroidal inductor - a common electrical component. For the sake of clarity, only a few turns of coil **6** are shown in each of **Fig.1**, **Fig.2**

and **Fig.3**. In practice, this coil may cover the entire core, or specific sections of the core, including, or not including the magnets.

Fig.2 shows the same electrical generator of **Fig.1**, looking transparently "down" through it from above, so that the relative positions of the core holes (shown as dotted lines), the path of the output wire **4**, and the position of the magnets (white hatched areas for magnets under the core and green hatched areas for magnets above the core) are made clear. The few representative turns of the input coil **6** are shown in red in **Fig.2**.

The generator illustrated, uses a core with 8 radially drilled holes. The spacing between these holes is equal. As shown, each hole is displaced by 45 degrees from each of it's adjoining holes. The centres of all of the holes lie on a common plane lying half-way down the vertical thickness of the core. Cores of any shape or size may have as few as two or as many as hundreds of holes and a similar number of magnets. Other variations exist, such as generators with multiple rows of holes, zigzag and diagonal patterns, or output wire **4** moulded directly into the core material. In any case, the basic magnetic interaction shown in **Fig.3** occurs for each hole in the core as described below.

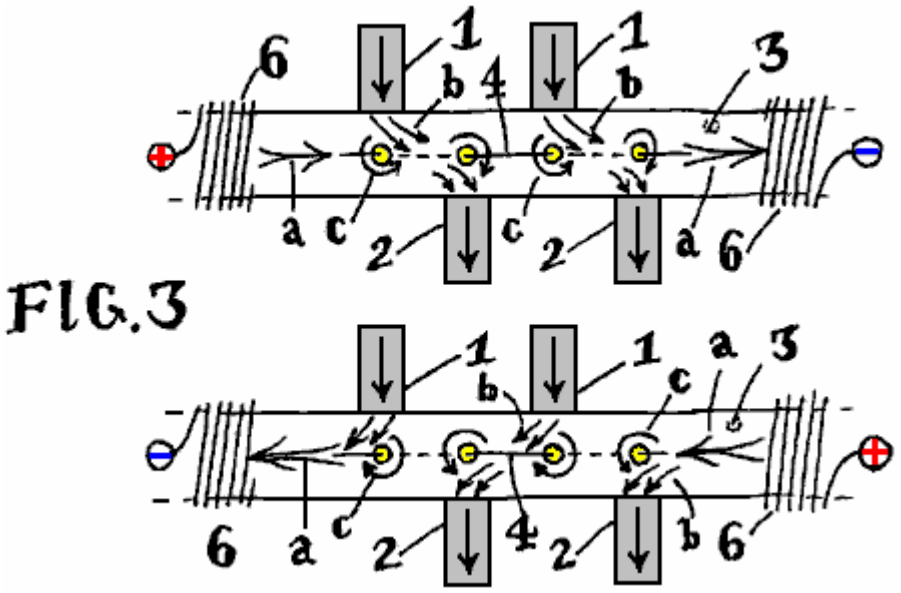


FIG.3

Fig.3 shows the same design, viewed from the side. The curvature of the core is shown flattened on the page for the purpose of illustration. The magnets are represented schematically, protruding from the top and bottom of the core, and including arrows indicating the direction of magnetic flux (the arrow heads point to the magnet's North pole).

In practice, the free, unattached polar ends of the generator's magnets may be left "as-is" in open air, or they may be provided with a common ferromagnetic path linking the unattached North and South poles together as a magnetic "ground". The common return path is typically made of steel, iron or similar material, taking the form of a ferrous enclosure housing the device. It may serve the additional purpose of a protecting chassis. The magnetic return may also be another ferromagnetic core of a similar electric generator stacked on top of the illustrated generator. There can be a stack of generators, sharing common magnets between the generator cores. Any such additions are without direct bearing on the functional principle of the generator itself, and have therefore been omitted from these illustrations.

Two example flux diagrams are shown in **Fig.3**. Each example is shown in a space between schematically depicted partial input coils **6**. A positive or negative polarity marker indicates the direction of input current, applied through the input coil. This applied current produces "modulating" magnetic flux, which is used to synthesise apparent motion of the permanent magnets, and is shown as a double-tailed horizontal arrow (**a**) along the core **3**. Each example shows this double-tailed arrow (**a**) pointing to the right or to the left, depending on the polarity of the applied current.

In either case, vertical flux entering the core (**b,3**) from the external permanent magnets (**1,2**) is swept along within the core, in the direction of the double-tailed arrow (**a**), representing the magnetic flux of the input coil. These curved arrows (**b**) in the space between the magnets and the holes, can be seen to shift or bend (**a** --> **b**), as if they were streams or jets of air subject to a changing wind.

The resulting sweeping motion of the fields of the permanent magnets, causes their flux (**b**) to brush back and forth over the holes and wire **4** which passes through these holes. Just as in a mechanical generator, when the

magnetic flux brushes or “cuts” sideways across a conductor in this way, voltage is induced in the conductor. If an electrical load is connected across the ends of this wire conductor (numeral **5** in **Fig.1** and **Fig.2**), a current flows through the load via this closed circuit, delivering electrical power able to perform work. Input of an alternating current across the input coil **6**, generates an alternating magnetic field (**a**) causing the fields of permanent magnets **1** and **2** to shift (**b**) within the core **3**, inducing electrical power through a load (attached to terminals **5**), as if the fixed magnets (**1**,**2**) themselves were physically moving. However, no mechanical motion is present.

In a mechanical generator, induced current powering an electrical load, returns through output wire **4**, creating a secondary induced magnetic field, exerting forces which substantially oppose the original magnetic field inducing the original EMF. Since load currents induce their own, secondary magnetic fields opposing the original act of induction in this way, the source of the original induction requires additional energy to restore itself and continue generating electricity. In mechanical generators, the energy-inducing motion of the generator’s magnetic fields is being physically actuated, requiring a strong prime mover (such as a steam turbine) to restore the EMF-generating magnetic fields’ motion against the braking effect of the output-induced magnetic fields (the induced field **c** and the inducing field **b**), destructively in mutual opposition, which must ultimately be overcome by physical force, which is commonly produced by the consumption of other energy resources.

The electrical generator of the present invention is not actuated by mechanical force. It makes use of the induced secondary magnetic field in such a way as to not cause opposition, but instead, addition and resulting acceleration of magnetic field motion. Because the present invention is not mechanically actuated, and because the magnetic fields do not act to destroy one another in mutual opposition, the present invention does not require the consumption of natural resources in order to generate electricity.

The present generator’s induced magnetic field, resulting from electrical current flowing through the load and returning through output wire **4**, is that of a closed loop encircling each hole in the core. The induced magnetic fields create magnetic flux in the form of closed loops within the ferromagnetic core. The magnetic field “encircles” each hole in the core which carries output wire **4**. This is similar to the threads of a screw “encircling” the shaft of the screw.

Within this generator, the magnetic field from output wire **4** immediately encircles each hole formed in the core (**c**). Since wire **4** may take an opposing direction through each neighbouring hole, the direction of the resulting magnetic field will likewise be opposite. The direction of arrows (**b**) and (**c**) are, at each hole, opposing, headed in opposite directions, since (**b**) is the inducing flux and (**c**) is the induced flux, each opposing one another while generating electricity.

However, this magnetic opposition is effectively directed against the permanent magnets which are injecting their flux into the core, but not the source of the alternating magnetic input field **6**. In the present solid-state generator, induced output flux (**4**,**c**) is directed to oppose the permanent magnets (**1**,**2**) not the input flux source (**6**, **a**) which is synthesising the virtual motion of those magnets (**1**,**2**) by it’s magnetising action on core **3**.

The present generator employs magnets as the source of motive pressure driving the generator, since they are the entity being opposed or “pushed against” by the opposing reaction induced by output current which is powering a load. Experiments show that high-quality permanent magnets can be magnetically “pushed against” in this way for very long periods of time, before becoming demagnetised or “spent”.

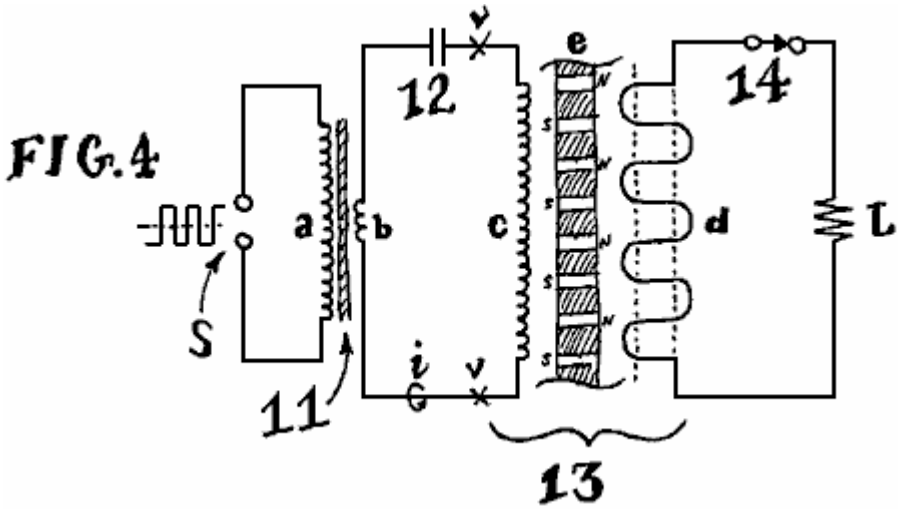
Fig.3 illustrates inducing representative flux arrows (**b**) directed oppositely against induced representative flux (**c**). In materials typically used to form core **3**, fields flowing in mutually opposite directions tend to cancel each other, just as positive and negative numbers of equal magnitude sum to zero.

On the remaining side of each hole, opposite the permanent magnet, no mutual opposition takes place. Induced flux (**c**) caused by the generator load current remains present; however, inducing flux from the permanent magnets (**b**) is not present since no magnet is present, on this side, to provide the necessary flux. This leaves the induced flux (**c**) encircling the hole, as well as input flux (**a**) from the input coils **6**, continuing its path along the core, on either side of each hole.

On the side of each hole in the core where a magnet is present, action (**b**) and reaction (**c**) magnetic flux substantially cancel each other, being directed in opposite directions within the core. On the other side of each hole, where no magnet is present, input flux (**a**) and reaction flux (**c**) share a common direction. Magnetic flux adds together in these zones, where induced magnetic flux (**c**) aids the input flux (**a**). This is the reverse of typical generator action, where induced flux (**c**) is typically opposing the “input” flux originating the induction.

Since the magnetic interaction is a combination of magnetic flux opposition and magnetic flux acceleration, there is no longer an overall magnetic braking or total opposition effect. The braking and opposition is counterbalanced

by a simultaneous magnetic acceleration within the core. Since mechanical motion is absent, the equivalent electrical effect ranges from idling, or absence of opposition, to a strengthening and overall acceleration of the electrical input signal (within coils **6**). proper selection of the permanent magnet (**1**,**2**) material and flux density, core **3** material magnetic characteristics, core hole pattern and spacing, and output medium connection technique, create embodiments where the present generator will display an absence of electrical loading at the input and/or an overall amplification of the input signal. This ultimately causes less input energy to be required in order to work the generator. Therefore, as increasing amounts of energy are withdrawn from the generator as output power performing useful work, decreasing amounts of energy are generally required to operate it. This process continues, working against the permanent magnets (**1**,**2**) until they are demagnetised.



In an embodiment of this invention, **Fig.4** illustrates a typical operating circuit employing the generator of this invention. A square-wave input signal from a transistor switching circuit, is applied at the input terminals (**S**), to the primary (**a**) of a step-down transformer **11**. The secondary winding (**b**) of the input transformer may be a single turn, in series with a capacitor **12** and the generator **13** input coil (**c**), forming a series resonant circuit. The frequency of the applied square wave (S) must either match, or be an integral sub-harmonic of the resonant frequency of this 3-element transformer-capacitor-inductor input circuit.

Generator **13** output winding (**d**) is connected to resistive load **L** through switch **14**. When switch **14** is closed, generated power is dissipated at **L**, which is any resistive load, for example, and incandescent lamp or resistive heater.

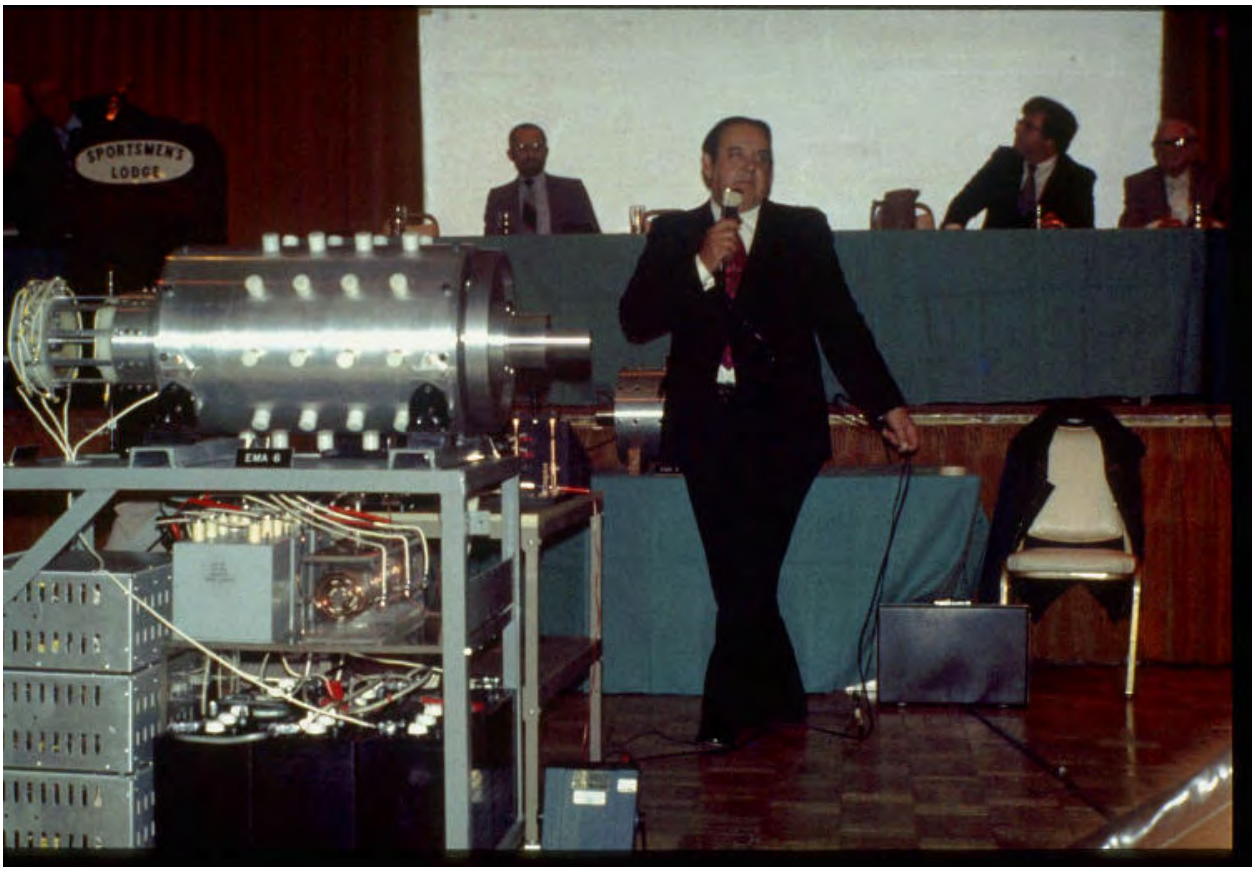
Once input resonance is achieved, and the square-wave frequency applied at **S** is such that the combined reactive impedance of total inductance (**b** + **c**) is equal in magnitude to the opposing reactive impedance of capacitance **12**, the electrical phases of current through, and voltage across, generator **13** input coil (**c**) will flow 90 degrees apart in resonant quadrature. Power drawn from the square-wave input energy source applied to **S** will now be at a minimum.

In this condition, the resonant energy present at the generator input may be measured by connecting a voltage probe across the test points (**v**), situated across the generator input coil, together with a current probe around point (**I**), situated in series with the generator input coil (**c**). The instantaneous vector product of these two measurements indicates the energy circulating at the generator's input, ultimately shifting the permanent magnets' fields in order to create useful induction. This situation persists until the magnets are no longer magnetised.

It will be apparent to those skilled in the art that a square (or other) wave may be applied directly to the generator input terminals (**c**) without the use of other components. While this remains effective, advantageous re-generating effects may not be realised to their fullest extent with such direct excitation. Use of a resonant circuit, particularly with inclusion of a capacitor **12** as suggested, facilitates recirculation of energy within the input circuit, generally producing efficient excitation and a reduction of the required input power as loads are applied.

Mark McKay's Investigation of Edwin Gray's Technology: Part 1

Enter…. The Mallory Connection Mark McKay, PE 3/2/06



E.V. Gray Version 2.0 type Motor EMA6 1977 – Courtesy Dr. Peter Lindemann

Consider the now classic 1977 photo (above) of Mr. E.V. Gray demonstrating his EMA6 motor to investors at the Sportsman Lodge in Burbank, CA. This photo was taken by Tom Valentine, who wrote a series of informative articles about the EV Gray saga. Dr. Peter Lindemann received this original film from Mr. Valentine to support Peter's research for his book "The Free Energy Secrets of Cold Electricity".

In a fruitful attempt to extract additional technical information from this historical photo Dr. Lindemann arranged to have it digitally enhanced. One of the goals of this effort was to decipher the writing on the large gray storage capacitor directly under the motor. It read:

MALLORY
MADE IN U.S.A.
TYPE TVC-606
5.0 MFD 5000 VDC

Mallory is a well known name in the field of electronics. When one thinks of Mallory today they generally think of the premium large blue electrolytic filter capacitors that dominated the high end linear power supply market in the 70's and 80's. At its peak, the P.R. Mallory Company was a power house of US made electrical components. Not only did they make several lines of capacitors but they also made Battery Chargers, Resistors, Rheostats, Rectifiers, Switches, UHF Converters, Noise Filters, Soldering Iron Tips, and Special Television Components. Their 1955 Catalog was 60 pages long.

Mr. P.G. Mallory started out in 1916 with the invention of the Mercury Battery. By 1965 the company developed the well known Duracell Alkaline battery.

The North America Capacitor Company (NACC) is headquartered in Indianapolis, Indiana. Today, NACC continues to manufacture and market Mallory capacitors at its modern manufacturing and warehouse facilities located in Greencastle, Indiana and Glasgow, Kentucky

Mallory Capacitors and Duracell Batteries from Author's Experimental Parts Reserve

Another important Mallory invention, very relative to the EV Gray technology, was the 1920's development of the "Elkonode", better known back then as simply the "vibrator". Today this device is hardly known at all. In its time it served as a vital sub-system in early DC converters. These were used to raise the low voltage levels of storage batteries to the operating levels required by vacuum tubes, which was 200 to 500 VDC. This now forgotten electro-mechanical component was the functional equivalent of two push-pull power transistors in a modern

switch-mode power supply. At the time, when it came to mobile electronics there were two choices. 1) A vibrator based power converter, or 2) A heavy dynamo-motor base converter. For applications under 30 watts the vibrator approach was smaller, lighter, cheaper, and more efficient than the alternative. Therefore, the military had a serious interest this technology, but it was in the mass market demand for small vacuum tube car radios where the real money was made.

The P.G. Mallory Co. almost completely dominated the top end power vibrator market for 40 years and was responsible for almost all of the performance improvements through the 40's and 50's. But, all good things must end. This lucrative product line came to a screeching halt in 1957 with the development of low voltage signal and power transistors. But Mallory still managed to keep a cutting edge in many of its other market areas for several years after that.

So, it is no big surprise when one reads in the 1973 Scagnetti EV Gray article:

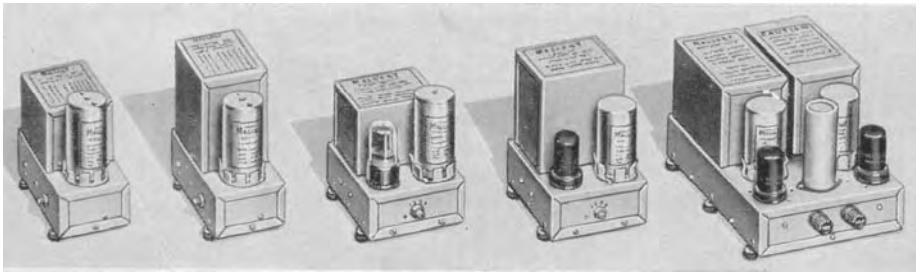
The Engine that Runs Itself

By Jack Scagnetti from `Probe The Unknown' in June 1973.

"Mallory Electric Corporation of Carson City, Nevada, has also made a major contribution toward the design of the electronic pulsing system."

It's all pretty obvious that Mr. Gray had a huge investment in Mallory type components. If his invention did become main stream then the Mallory Co. would have had first shot at a huge new automotive market. Each new vehicle would need between $300 - $600 worth of rugged HV storage capacitors, not to mention an investment of twice that much for vibrator power converters or their equivalent solid state replacements, which Mallory made also.

It is real easy to see how Mr. Gray could have convinced a few executives at Mallory how it would be in their best interests to help him out financially, or at least provide him with a little hardware donation from their Vibrapack division in Irvine CA. Mr. Grays impressive "hands-on" demonstrations were known to be very effective at convincing technical professionals that he was on to something big, providing that he was ever allowed the opportunity to make such presentation to a real decision maker. Most likely some inspired and insightful 3rd level staff person managed to fix him up with a pickup load of surplus vibrator converters that were, or would be, completely obsolete.



Examples of the P.R. Mallory line of "Vibrapacks" (DC Converters) from 1955 Catalog

All models have a 30 Watt power rating except the one on the far right which is rated at 60 Watts

But this story has an important twist in it……..

The Mallory Company that gave Mr. Gray enough money to make mention of it in the above magazine article was not the P. G. Mallory & Company Inc. but the Mallory Electric Company of Carson City, Nevada, designers and manufactures of a multitude of OEM and after-market automotive ignition systems.

A Small Sample of modern Mallory brand name After Market Ignition Products 2006

Mr. Marion Mallory was the rare sort of independent individual who would start a company on Friday the 13th in February of 1925. He was a self-made inventor with a 4th grade education who was not only brilliant at his craft but also had what it takes to manage a business. If he ever met Mr. Gray face to face the two men would have had a lot in common, especially from a “hands-on” creative energy standpoint. Mr. Mallory made his money in a variety of automotive, motor cycle and marine ignition systems. For years he was the main supplier to the Ford Motor Company for ignition distributors and their upgrades. He received about 30 US and 10 international patents for a multitude of significant improvements in ignition technology, both in electrical and mechanical systems. He

was darn good at business, but his personal weakness was high performance auto racing. The market for race car parts is not very big, but the activity it supports is very addictive. Marion sponsored as many as three teams a year in the various classes of professional auto racing. It is also been said that Mr. Mallory looked for and hired like minded creative engineers and technicians. He also despised the union worker mentality that had become so adversarial in the Detroit area between the 50's and 60's.

Mr. Mallory finally got fed up with the stifling and counter-productive demands of the United Auto Workers Union. In a rare act of individualism he decided to make arrangements to move his entire company, lock, stock and, ignition coils to Carson City, NV. At this time Marion was getting along in years and unfortunately never made the move. He died in 1968 at the age of 70. His son 'Boot' Mallory was then handed the reins of this privately held company. 'Boot' terminated all the Union labor and kept 10 of the most productive engineers and technicians who were willing to relocate to the new factory. This facility was opened in 1969. From all accounts the "heir apparent" and only son was very motivated, technically competent, savvy at business, and like his father hopelessly addicted to high performance auto racing.

Given the timing of events it is most likely that Mr. Gray never met Marion Mallory. It is almost certain that the connection to the Mallory Company was entirely between Mr. Gray and 'Boot' Mallory. This was also helped by the fact these two men were about the same age with Mr. Gray being 5 years older.

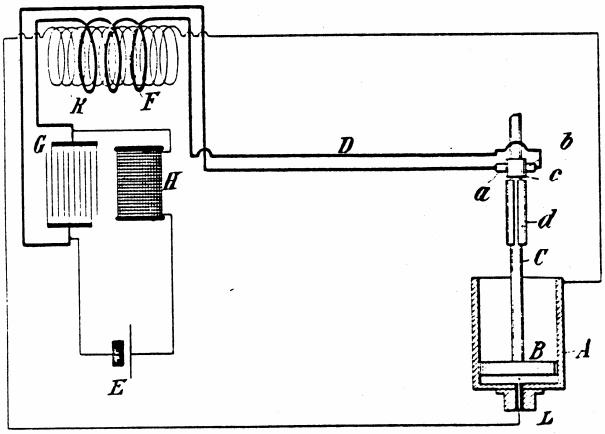
For their entire business careers Marion and 'Boot' Mallory were always on the look out for improved ignition systems, both for good business practice and, of course, a desire to sport the fastest cars at the race track. Their knowledge base and field experience covered all approaches to ignition system design, both in the electrical and mechanical areas. It is interesting to note that they developed and manufactured magneto systems as well as traditional distributor systems. Understand that these two technologies are vastly different to each other.

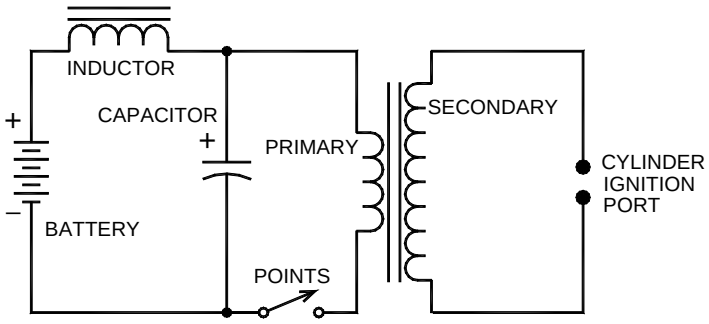
No. 609,250. Patented Aug. 16, 1898.

N. TESLA.

ELECTRICAL IGNITER FOR GAS ENGINES.

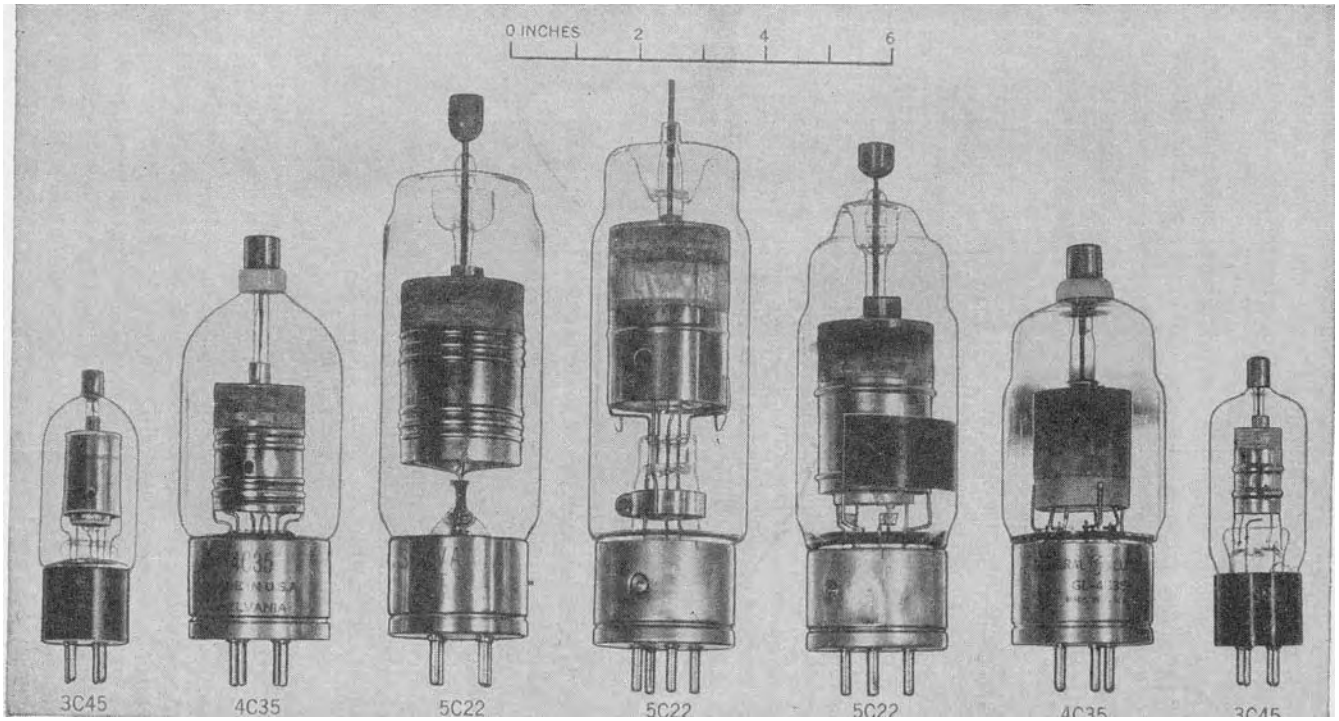
(No Model.)





SCHEMATIC FOR TESLA'S "ELECTRICAL IGNITER FOR GAS-ENGINES"
US PATENT 609,250 AUGUST 1898
FIG. 7 (From The Complete Patents of Nikola Tesla)

In the auto racing circles it has always been known that capacitive discharge ignitions system are far superior to the limitations of the standard Kettering induction system, especially at high RPM. Dr. Tesla patented the first CD ignition system as early as 1898 but it was never produced because of serious design and component limitations. Marion Mallory and his engineers did get a working capacitive-discharge system finally connected to a race car engine in 1948. This first design was built employing a thyratron gas tube and vacuum-tube circuitry. As a result, it was costly, bulky, and unwieldy, not to mention fragile and economical unfeasible. But despite all of its failings the Capacitive Discharge Systems (CD) clearly showed its superior performance in the laboratory and on the track. Had it not been for the random and sudden failure of these alpha-test units (because of vibration) they might have still been used in professional auto racing, regardless of their unit cost.



Glass Hydrogen Thyratrons of the 40's
From "Pulse Generators" Radiation Laboratory MIT 1948

Two new technologies were needed to get CD systems off the ground.

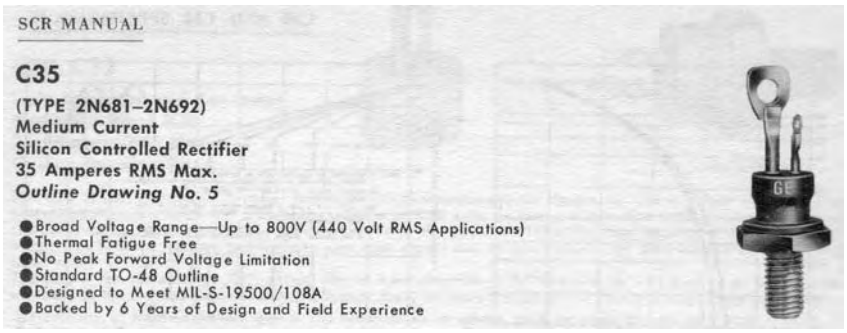
1) Some method to boost the 6 or 12 V DC storage battery voltage to the 400-500 Volt range with an available current of at least 100 mA. (40-50 Watts)

2) A component or technique that would replace the bulky, fragile, and power hungry thyratron that acted as the master timing control switch.

Both solutions came along about the same time. Power transistors became available to the aerospace industry in 1954. These allowed the development of early push-pull switched mode power supplies whose output were way beyond what a mechanical power vibrator could deliver (up to 90 Watts initially). Complete transistor converters were available to the hobbyist in early 1958. So we can assume that prototype power transistors were available to industry in about 1955.

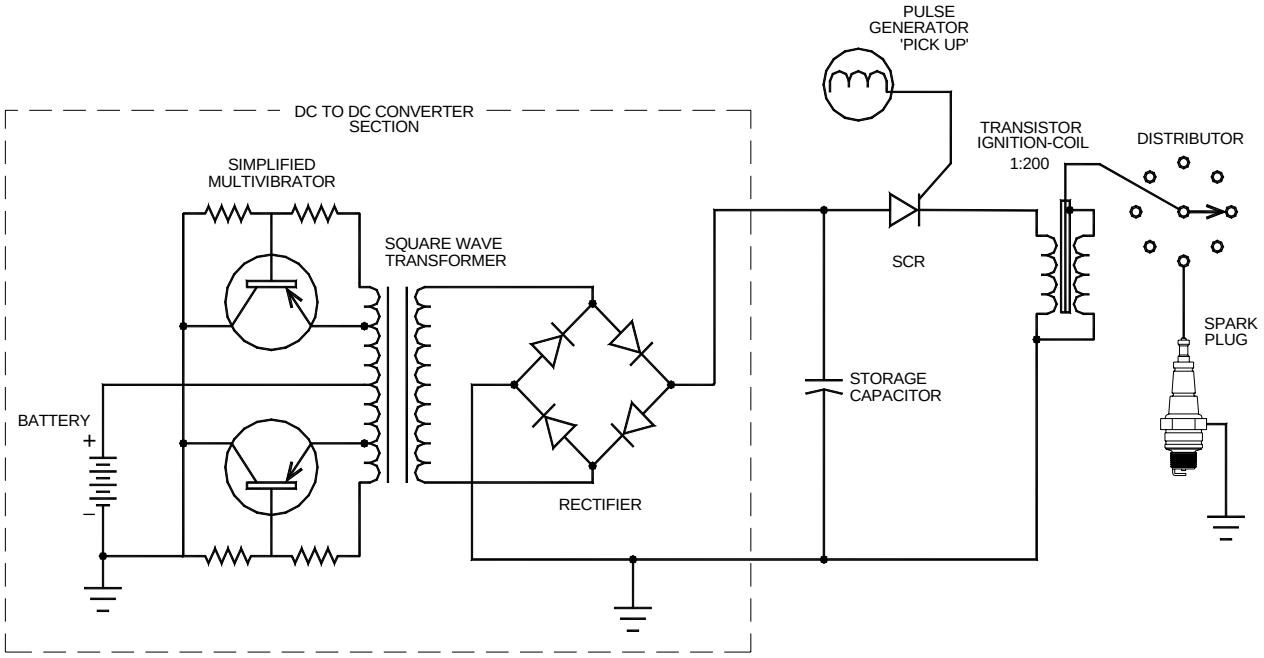
Early advertisement for a 90 Watt (pulsed) Hobbyist 12V to 450V DC Converter
From “QST” magazine January 1958
(Notice size reduction when compared to the 60 Watt Vibrapack)

The second critical breakthrough came with the invention of the Thyristor or Silicon Controlled Rectifier (SCR) by Bell Labs in 1957. General Electric quickly bought the rights for this promising technology and wasted no time in bringing it into production. The manufacture of solid state power rectifiers and transistors was already well underway, so, building an SCR using the existing production equipment was a slam-dunk. According to the GE SCR Handbook 1964 3rd edition, the model C35 had already been in the field since 1958.



Silicon Controlled Rectifier available to Industry and Military in 1958

With these new solid state components at hand Marion & ‘Boot’ Mallory were off and running. Their first beta-test race track CD ignition system was introduced in limited quantities in the fall of 1961. Their first after market production models did not reach distributors until 1964. It took 3 years of detailed development and waiting for the SCR market to settle down before deciding on a final production design. While the basic operating principles of a CD ignition circuit is straight forward getting a long-life circuit that will function well when exposed to the temperature, voltage, and vibration extremes is a different matter. At that time in our country’s industrial heritage new products were not generally rushed, half-baked, to the re-sellers because of some imaginary dead-line imposed by the bean-counters in the marketing department.



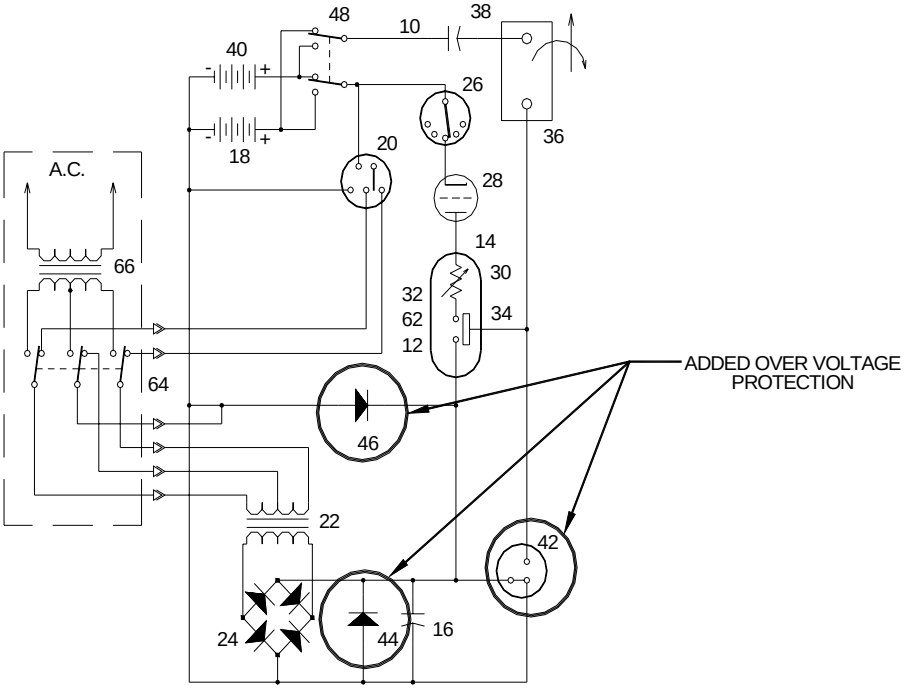
SIMPLIFIED SCHEMATIC OF CAPACITIVE DISCHARGE SYSTEM
CICRA 1975 TO PRESENT
(From Tektronix - Engine Analysis Measurements 1970)

So, in the timeframe of 1960 to 1970 where could Mr. Gray have gone when he needed some rare applied technical expertise on battery operated High Voltage pulse systems? The solution seems almost obvious.

We have no doubt that Mr. Gray and 'Boot' Mallory were on a first name basis. They may have already developed some kind of relationship while the company was still in Detroit, we don't know when they first got together. We do know that Mr. Gray was provided with some significant venture capital along with the fruits of 10 or so years of proprietary field tested solid state CD technology.

It has been pointed out, by knowledgeable sources, that all of the Mallory's after market ignition systems used power transistors for the 6-12V to 450V converter section. So, we wonder, why was Mr. Gray still using obsolete vibrator packs in 1973? 'Boot' would have certainly supplied Mr. Gray with the most modern equipment, along with the SCR and Ignition-Coil components in a small, self contained, custom engineered, and de-bugged package.

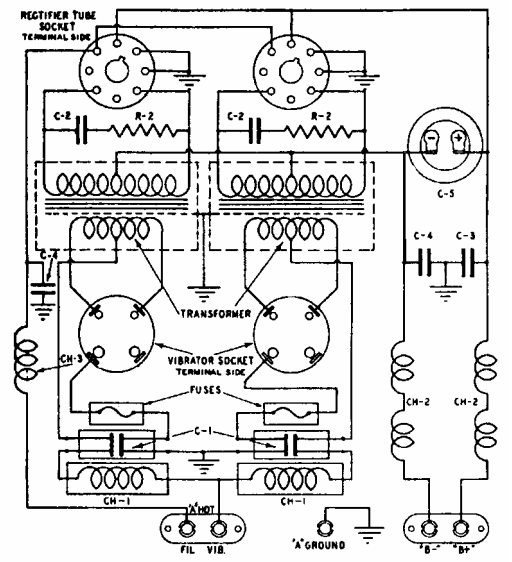
We suspect that 'Boot' did provide these complete transistorized CD systems and that Mr. Gray was eagerly looking forward to the reduced size, increased life time, and improved efficiencies that the new solid state devices promised. Especially after having to constantly fight with vibrators that kept burning out during his trial runs. But, Radiant Energy (RE) generation has its own special challenges to deal with. One major engineering issue is what to do with the Electro Magnetic Pulse (EMP) like effect that happens when a RE circuit reaches a certain power level. If all that excess energy is not properly shunted to the system common (hopefully after doing some serious work) it escapes from the circuit conductors to charge every metal object within 20' or so of the generator. A multitude of blue-white sparks will erupt from every metallic object in a room, due to the induced high voltage. This is certainly an interesting light-show, with the lights turned off, but devastating to any near by transistor or IC that has any amount of wire connected to it. Transistors and IC's that are stored in metalised protective bags or boxes seem to survive.

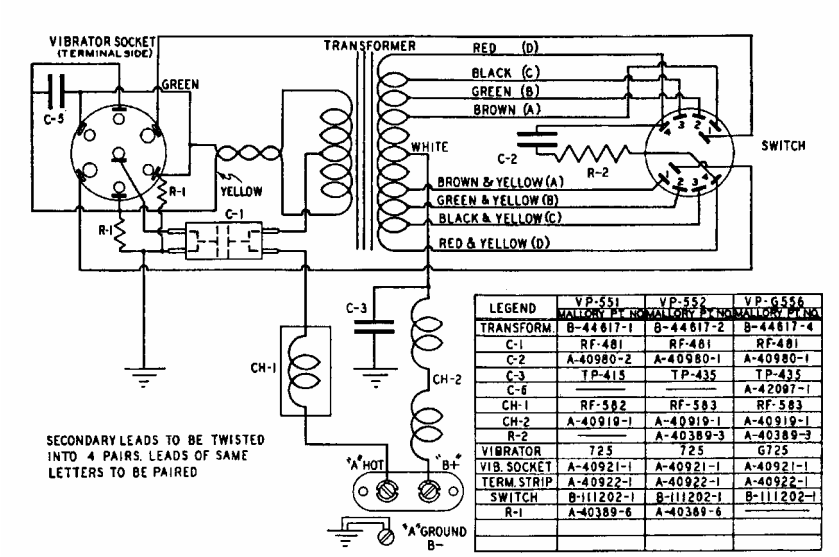


THE GRAY CIRCUIT PER PATENT 4,595,975
JUNE 17, 1986

If this was the case, then we can imagine how disappointed Mr. Gray might have felt when his new transistorized converters started to fail, perhaps even catastrophically. Fortunately, and **we really mean very fortunately,** the SCRs were able to survive the RE onslaught. Had this not been the case the EV Gray technology, because of the constant system failure, would have seriously fallen on its nose by 1965 and never have been able to produce the demonstrated power levels that we would so very much like to recreate. Transistors, fail because they are constructed with super thin base structures that are sensitive to moderate voltage differences. SCRs are constructed with thick silicon layers that are relatively more rugged. However, a poorly designed trigger circuit in an RE application will still destroy a heavy duty SCR, if proper gate transient protection methods are not employed. Because of this first hand experience Mr. Gray went on to install many over-voltage protection devices in his future circuits. This is very apparent in the design of the power supply shown in his Conversion Tube Patent #4,595,975.

It appears that Mr. Gray was forced to go back and use the failure prone obsolete vibrator packs that he started out with. According to the first patent these were used for the primary DC voltage conversion. We suspect that the engineers at Mallory were enlisted to help Mr. Gray marry the vibrator pack to the SCR system. The SCR addition did help solve the failure problem by reducing the arching current across the vibrator contacts. This is not a straight forward interface and it requires some experienced electronic know-how. The challenge is balancing the limited current capacity of the vibrator to the low impedance of the SCR storage capacitor.



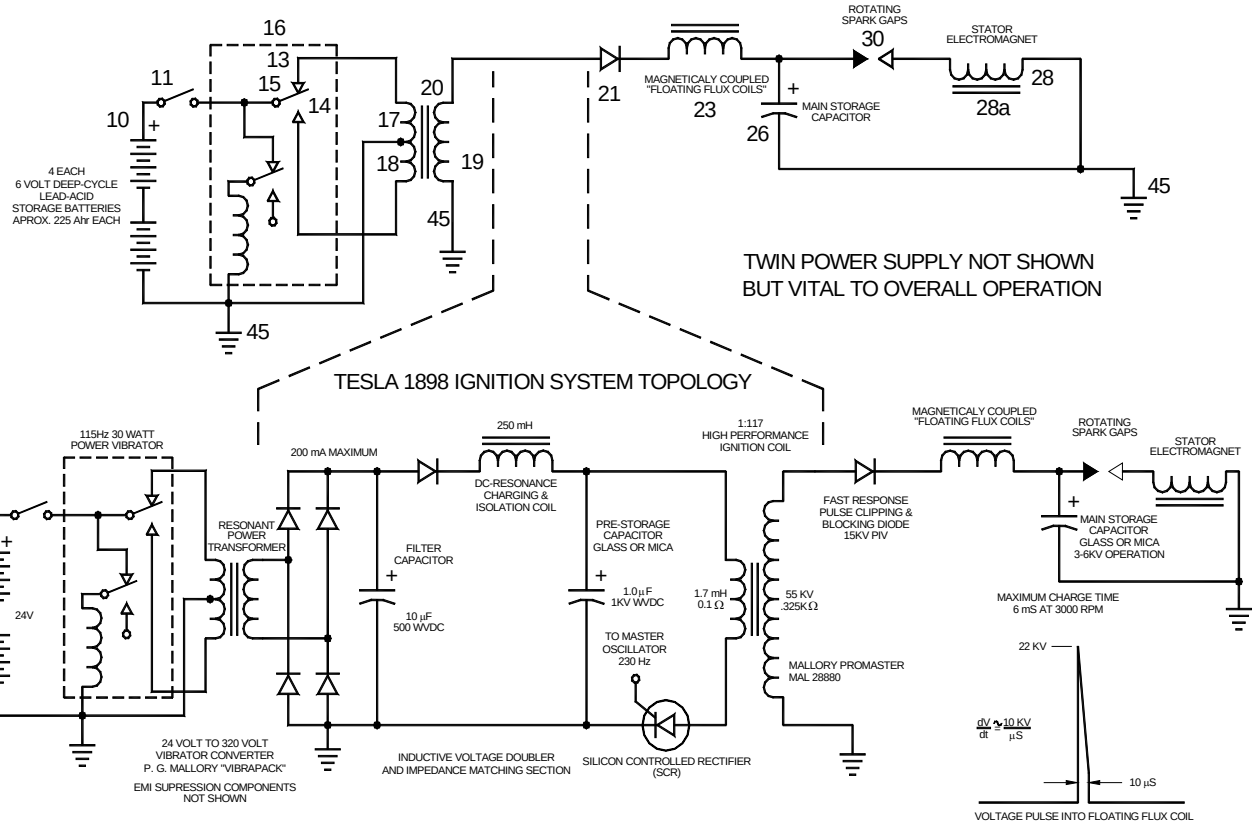


LEGEND	VP-551 MALLORY PT. NO.	VP-552 MALLORY PT. NO.	VP-G556 MALLORY PT. NO.
TRANSFORM.	B-44617-1	B-44617-2	B-44617-4
C-1	RF-481	RF-481	RF-481
C-2	A-40980-2	A-40980-1	A-40980-1
C-3	TP-415	TP-435	TP-435
C-6	———	———	A-42097-1
CH-1	RF-582	RF-583	RF-583
CH-2	A-40919-1	A-40919-1	A-40919-1
R-2	———	A-40389-3	A-40389-3
VIBRATOR	725	725	G725
VIB. SOCKET	A-40921-1	A-40921-1	A-40921-1
TERM. STRIP	A-40922-1	A-40922-1	A-40922-1
SWITCH	B-111202-1	B-111202-1	B-111202-1
R-1	A-40389-6	A-40389-6	———

FIG. 28—SCHEMATIC WIRING DIAGRAM FOR VIBRAPACKS
Nos. VP-551, VP-552, VP-G556

Schematic Wiring Diagrams for two P.R. Mallory Vibrapacks
60 Watt model on the left – 30 Watt model on the right

Other researchers contend that Mr. Gray never intended to use transistors in the first place. This is because one RE theory states that the non-classical process begins in the minute arcs formed during the making and breaking of the vibrator contacts. This technical issue is still open for debate and experimental verification.



PROPOSED NON-DISCLOSED CAPACITIVE DISCHARGE SUB-SYSTEM IN EV GRAY CIRCUIT

However, we all agree that the SCR CD circuit is still a vital sub-system to the EV Gray technology, but it is not the whole story for a complete Over Unity (OU) process. We further believe that Mr. Gray didn't disclose the kernel of his "secret" to 'Boot' or any one else at the Mallory Electric Company. It would appear that 'Boot', because of his unique individualistic upbringing, respected Mr. Gray's right to his own creations. 'Boot' was obviously far sighted enough to see some greater business potential in this venture, not to mention a whole new class of future racing machines. One main reason for this enlightened attitude was that 'Boot' didn't have to contend with a short-sighted governing board of directors whose members were more worried about next quarters stock price than taking risky chances on age changing technologies.

The CD sub-system of the Gray motor was not disclosed in patent #3,890,548. Mr. Gray did mention the use of ignition coils in the patent text, but didn't show them in the schematic diagram. The simplest solution to help protect his "secret" was to just eliminate the CD sub-system from the schematic. Since Mr. Gray was only attempting to disclose a new type of pulse motor in this first patent. The omission of a "minor" power supply "feature" was not going to mean anything to the patent reviewers. But, the devil is in the details, especially when attempting to reconstruct this lost technology 30 years later.

There is a good possibility that Mr. Gray was returning a favor to 'Boot' by not disclosing the proprietary CD circuit designs. They very well could have had a gentlemen's agreement and a joint venture on this issue. 'Boot' didn't need to know Mr. Gray's Free Energy "Secret". His high margin piece of the action was locked in because each new EV Gray motor would need 18 or more complete CD power supplies, including the patented construction details of the Mallory ignition coils. Mr. Gray's success was going to be 'Boot' Mallory's success – BIG TIME. A classic win-win situation. It's no wonder that 'Boot' willingly made out checks to this unknown and un-educated inventor from California. While the P.R. Mallory Company was unknowingly going to reap some benefit from this breakthrough the Mallory Electric Company was going to hit the jackpot.

As a purely speculative observation, it may have been 'Boot' Mallory who clued Mr. Gray in on how to write patents and attempt to protect one's intellectual property form the big business lawyers. What to show and what not to show, what to draw and what not to draw and what to say the rest of the time. With this technology it was going be a feeding frenzy as soon before the first beta-test hit the street and 'Boot' knew it. Mr. Gray probably received a life time of inside information on how to keep secrets, make money, and cover one's assets from a man who had been there and seen how big business really works.

We all know that Mr. Gray suffered a major setback when his research facility was raided in 1974 by the agents of the Los Angles District Attorneys Office for suspected securities fraud. But, by 1977, as shown in the photo above, Mr. Gray had recovered enough to receive his first patent, build, debug, and demonstrate his second generation

motor. What is not generally known, in Free Energy circles, is that Mr. Gray suffered a far greater loss when 'Boot' Mallory was killed in a car wreck in 1978 at the age of 48. He was always known to be somewhat of a lead foot.

Gone was the financial, technical and morel support. As far as we can observe it appears that the EV Gray motor didn't develop significantly much beyond the EMA6 model (above). The surviving Mallory women sold the company to Super Shops of Irvine, California in 1979. Mr. Gray continued to seek a proper level of investment capital so that he could control and manufacture his fuel-less motors in-house. He also improved on his popping-coil demonstration and updated it to a continuous process that hinted at anti-gravity possibilities, very impressive. It has also been rumored that Mr. Gray almost did collect enough money to begin production.

Unfortunately, we also know that ten years later Mr. Gray died under un-resolved circumstances in Sparks, NV in April, 1989. Sparks is just East of Reno, NV which is about 50 miles North of Carson City, NV. Some researchers contend that the main reason why Mr. Gray established one of his multiple laboratories in this town was because of the invaluable technical experience of some of the retired Mallory technicians still living in the area.

We have also been lead to believe that it was 'Boot' Mallory who made the first formal introductions between Mr. Gray and the alternate car inventor Mr. Paul M. Lewis, creator of the "Fascination". You can imagine the possible creative energy that might have flowed between these three unique individuals while they were sitting around the dinner table sharing a host of far-reaching dreams and schemes.

Today, the sold and re-sold fragments of the P.R. Mallory and the Mallory Electric Company have suffered, like so many U.S. businesses, from the now common and insidious blight of globalization. Both organizations are outsourcing their manufacturing operations to China, their engineering departments to India, and their R & D efforts to Canada.

In conclusion all we can say is that this saga is truly a vital lost opportunity for the world, they were so darn close. Had this story been different we most likely wouldn't be bankrupting our country in a vain attempt to secure oil reserves in Iraq. We could have easily had permanent colonies on Mars and not be worrying about the ongoing effects of Green House Gasses. This great country could have re-invested the trillions of our oil dollars into our own economy rather than providing excessively lush life styles for a few privileged Middle Eastern clan leaders.

Note: This document is one in a series produced by Mr. McKay as part of his investigation of the work of Edwin Gray senior and he invites readers to contact him if they have any constructive comments or queries concerning the work of Mr. Gray. Mr McKay's e-mail address is mmckay@tycoint.com

Mark McKay's investigation of Edwin Gray's Technology: Part 2

Taking a closer Look at the Demonstration Equipment
October 24, 2006

This is the classic photo of E.V. Gray's "Popping Coil" Demonstration apparatus. This can be found on Peter Lindemann's web site. This photo was taken by Tom Valentine in 1973. Mr. Gray is the man in the center and Fritz Lens (his new father-in-law) is on the right. The man on the left is unidentified (most likely Richard Hackenburger VP of Engineering).

For years, about all one could say about this photo was that there was a fair amount of equipment involved in these demonstrations. The energy source appears to be a common large automotive 12 volt battery. Identifiable components are the custom made air transformer and the Triplett 630-A multimeter, all the rest of the technical detail is hidden by the black Plexiglas instrument boxes. By itself this photo does not yield much information.
In 2004 a former E.V. Gray investor came forth and presented Peter Lindemann and John Bedini with a period collection of historical snapshots. Five of these photos were of the same apparatus that was shown to Mr. Valentine in the above photo. The location was different, but the equipment and layout appears to be the same. It is assumed that these new investor photos were taken at Mr. Grays shop in Van Nuys, CA. These photos were developed in January and June of 1974 so they could have been taken within a few months of the Valentine 1973 photo. By observing these photos some additional technical information about this novel technology can be extracted.

The Investor Photos:

Investor Photo #013C
Overall View

This is a nice shot of the whole demonstration apparatus from one end of the table showing the supply battery, two popping coils and an end view of the air transformer. Despite the limited focus, this photo shows that the popping coils are connected in parallel since the white leads on the left are both terminated on the negative terminal of the battery. Also connected to the battery is a component that appears to be an analog metering current shunt - a low value high current resistor device. However, there is no meter connected to this component as there would be in a normal application. This suggests that it is being used simply as a low value current limiting resistor. It is doubtful that this component was ever intended to be used in a metering capacity. Its output would have been a very short voltage pulse that could not be recorded or observed on any of the test instrumentation shown in any of these photos.

It is believed that the two black leads on the right of the air transformer are disconnected and hanging straight down to the floor. Compare this situation to the Tom Valentine photo where these heavy black leads are connected to two of the black boxes.

There appears to be four black wires connected to the right side of the electromagnets. The two larger black wires are thought to connect to the wiper of the DPST knife switch. It is not known for sure where the small remaining black wires connect, but most likely to an additional set of electromagnets parked under the air transformer as shown in photo #013B. If so, then there probably was an accompanying demonstration that showed what would happen if additional load was added to the circuit.

Investor Photo #012D

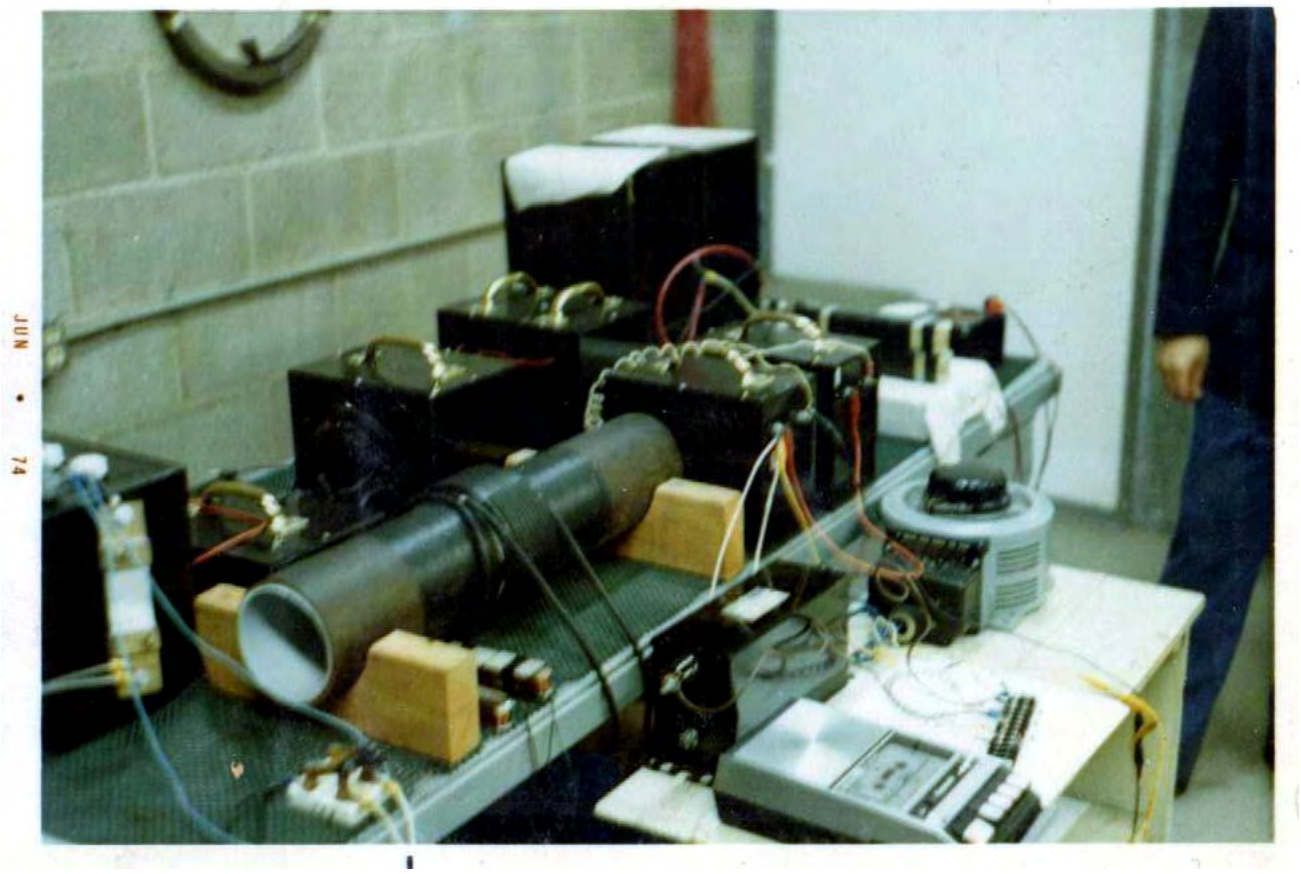
Popping a coil with the second demonstration setup on the “Right”

This photo is taken at the same location some time earlier where the circumstances were slightly different. The small white table and its attending equipment that is shown in the future June 74 photos are not preset. This photo (Jan 74) was developed 6 months before Photo #013C. The equipment on the large table seems to be in the same relative positions. What this photo reveals is that there is a second “Popping Coil” demonstration taking place at the other end (right side) of the table.

It is proposed that this total assembly of “Black Boxes” (a dozen or more subsystems) actually supports two different and independent demonstrations, a “Popping Coil” demo on the left and another similar “Popping Coil” demo on the right. The photos available allow for a better technical analysis of the demonstration equipment on the left side of the table. It is unknown as to what the actual differences between these two demonstrations were, however it is apparent that the coils being popped have obvious size differences. In photo #012D the coil in mid air is about twice the size of the electromagnets shown at the other end of the table in photo #013C. The Tom Valentine photo shows a set of electromagnets (at rest in the lower right hand corner) that are at least four times the size of the coils used for the demonstration that was set up on the left side of the table. However, the launched coil shown above is not the same (being 50% smaller) as the coil shown in the Tom Valentine photograph, even though it is being powered by the same equipment.

It is thought that the demo on the right had something to do with a higher power level or a more advanced method of energy recovery. Most likely, the demo on the left was intended to make the initial technical introduction to the basic idea of a repulsion motor concept, while the demo on the right had some important engineering advancement to display.

Photo #012D is dark but it helps shows that the two white wires from the DPST knife switch for the left demo connect to the two equal size boxes in the middle of the table, one wire per box.



Investor Photo #013B
120VAC Power Source being explored

This June 1974 photo is a nice over view of the “left” demonstration equipment. The major issue here is the additional equipment on the small white table. Here we see some identifiable items, a neon transformer, a 2KW Variac autotransformer, a cassette tape recorder and a barrier type terminal strip. The question is: What is this extra stuff for?

It appears that this setup is a variation from the normal equipment demonstration as seen in the Tom Valentine photo. It seems that the Air Transformer is disconnected from the system and has been replaced by the power provided by the equipment on the white table. Most likely this was an attempt to demonstrate that AC line power could be converted to “Cold Electricity”. It is important to note the variations in this particular circuit layout as it provides some clues as to the function of the various Black Boxes.

First, notice that the two white wires that go to the DPST knife switch have now been connected to one terminal of the black box, while a red jumper connects to the white wires’ previous connection point. Compare this to how these white wires are connected in the Tom Valentine photo.

It is not all together clear how the Neon transformer and Autotransformer are connected but a standard approach would be to have the Variac control the input line voltage to the Neon transformer. This Variac has the ability to increase its output voltage by 25% above its input. If this Neon transformer were a common 15KV 30 mA unit then the RMS output voltage could have been adjusted to a maximum of 18 KV. This is comparable to the output of an auto ignition coil. The peak DC voltage potential would have been about 25KV. However it is unlikely they were operating at this high of voltage for very long because of the size, layout and construction of the temporary conductors.

Since a single pair of conductors (yellow and black jumpers) drop below the top of the white table it is proposed that there is a high voltage diode stack underneath the table on a shelf that is operating in half-wave mode. Had full-wave mode been used then four wires would be seen leaving the top of the table (which is still a possibility).

The utilization of DC pulses is very clear in the Gray motor patent. It has often been wondered why Mr. Gray didn’t use full-wave rectification in his power supply to take advantage of the increased efficiency. Apparently this equipment does not have a taste for straight DC voltage. This concept is reinforced by the use of the half-wave rectification power supply shown in photo #013B. This situation supports the idea that Mr. Gray may have had

capacitors connected in series, without equalization resistors, thus pulsating DC would have been needed to charge them.

Photo #013B shows the best view of the demonstration equipment for the “Right” demonstration. It seems to be composed of five Black boxes, two small ones, two large ones, and one small flat one. If a knife switch was used to launch the popping coil it is not visible in these photos. An air transformer seems to be missing from this equipment collection. However, consider the cylindrical object seen under the large table in photos #012D and #013D. This is about the size of a gallon paint can and has yellow tape on top. Three black wires (and possibly a fourth) can be seen leading to this device. It is proposed that this is the air transformer used for this equipment. It has a larger diameter (8”) than the air transformer that is used for the “Left” demonstration (4”). It is believed that the automotive battery seen at the left end of the large table is the prime source of power for both demonstrations. A Triplett 630-A multimeter can be seen laying down on the far right of the table.

Examine the air transformer in its disconnected configuration. Notice how the two black conductors roll off the coil to the floor. This can only be achieved with two separate layers. The nearest conductor is part of the first layer. From this observation the relative polarity of the air transformer can be determined.

The core of the air transformer appears to be about 4” in diameter, when compared to the 2”x4” support blocks. It appears to be of a dual layer construction like one kind of pipe was slipped over another. The inner pipe resembles gray electrical PVC, but thinner (could be schedule 20 pipe). The outer pipe is a dark brown material that is not a common modern construction material. It is closer to an older fiber-composite material that was used for sewer pipe in the 50’s. Why the need for two nested cores? Is the dielectric breakdown of the core that big of an issue for such a small air transformer? The insulation strength of the (assumed) spark plug wire is near 50KV and should be plenty for the operating voltages expected. In addition there appears to be a hefty layer of electrical black tape between the core and the heavy windings.

It has been proposed that the black tape covers a single layer of #16 AWG magnet wire that forms a winding 3-4 times longer than the observed spark plug wire “primaries”. This feature (if it exists) is considered to be an additional energy recovery subsystem.

Investor Photo #013C
Group Photo Session

This photo is too fuzzy to extract much additional detail, (as compared to photo #013C) however the 35mm camera that is being held by the gentleman on the right is clear enough. Also, note the Flash Cube snapshot camera sitting beside the autotransformer. Cameras are in abundance in this portrait. This suggests that this particular collection of photos (June 74) were the result of a planned event where selected investors were allowed take all the snapshots they wanted. It is believed that this was a rare event. Therefore we can be assured that the equipment displayed at this time had been personally sanitized by Mr. Gray to insure that none of the essentials of his “Secret” would be disclosed.

The well dressed gentleman, on the left, appears to be holding another cassette tape recorder with a black plastic microphone being held in his fingers.

Investor Photo #013D

Count the Turns on the Air Transformer

This is about the best photo available showing the overall layout of both coil popping demonstrations. A lot of the essential details are hidden in this presentation but some of the subsystem interconnections can be determined.

The lower shelf of the white table displays what appears to be a HV “door knob” capacitor that is connected to Yellow and Black jumpers. It is more likely that this is a HV diode.

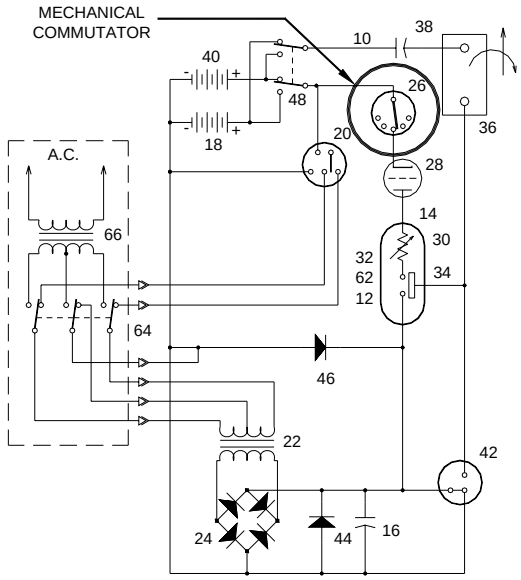
Note: This document is one in a series produced by Mr. McKay as part of his investigation of the work of Edwin Gray senior and he invites readers to contact him if they have any constructive comments or queries concerning the work of Mr. Gray. Mr McKay’s e-mail address is mmckay@tycoint.com

Mark McKay's investigation of Edwin Gray's Technology: Part 3

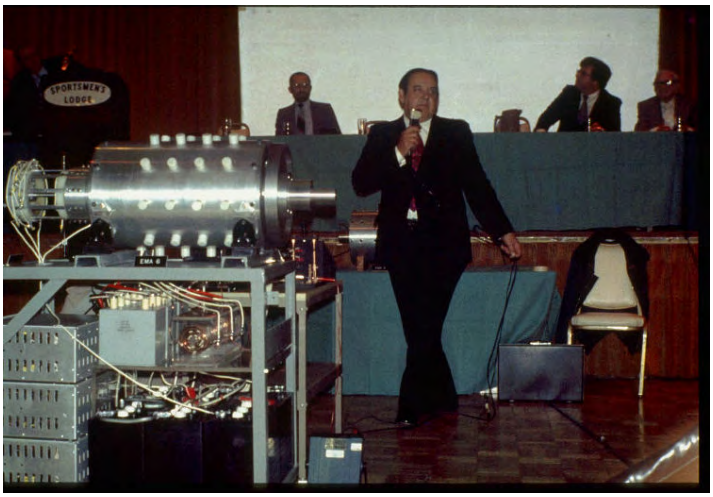
Secrets of the EMA4 and EMA5 Control Commutators (Still Unresolved) Mark McKay, PE

While the technical revelations provided by the disassembly of Mr. Gray's custom electromagnets is important, the observations collected from the EMA4 and EMA5 control commutators are even more interesting (and perplexing).

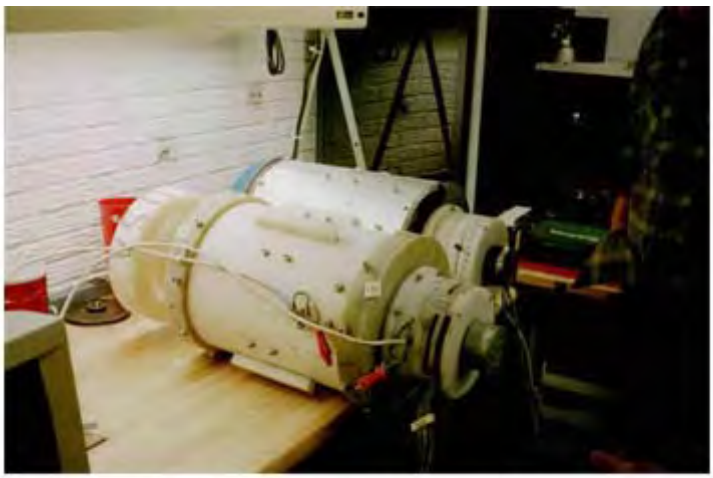
Prior to the recovery of the EMA4 & EMA5 it was thought that the attached white cylindrical device on the back end of the EMA6 was a simple rotary positional timing commutator device. According to patent 4,595,975 a commutator like device was included in the schematic diagram. It appeared to be some kind of mechanical rotary switch that controls timed pulses of power to flow through the anodes of the CSET. So when the patent and the photos are examined together the arrangement seems plausible.



THE GRAY CIRCUIT PER PATENT 4,595,975
JUNE 17, 1986

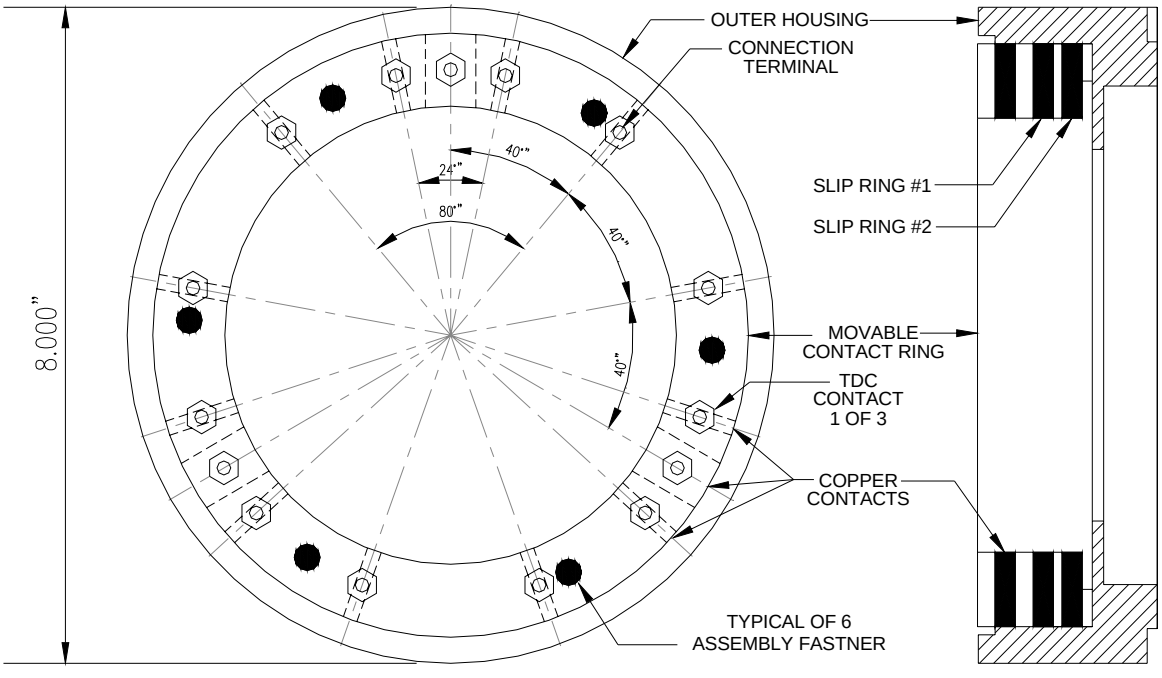

The EMA6 – with Control Commutator on extreme Left Stripped down EMA4 motor on back table

As it turns out the EMA4 and EMA5 motors revealed a much more complex component for researchers to consider. These commutators were constructed in such a way that they contained way more contacts than what would be needed for simple positional feedback. The units that came with each motor were designed to be pretty much the same, however they were wired differently. More control wires were utilized with the EMA5 than with the EMA4. This would be consistent with the fact the EMA4 only had one electromagnet pair to pulse while the EMA5 had three. The EMA5 commutator used 9 of its 15 contacts and was connected with 7 control wires. The EMA4 commutator also used 9 of its contacts but was only connected with 3 control wires.


EMA4 and EMA5 Motors at the time of recovery in 2000
With external Control Commutators mounted on the right

An examination for wear on the commutator contact surfaces, from possible arcing and heating, showed almost no signs of degradation. The conclusion reached from this observation was that whatever energy passed through these devices must have been at a very low level. This being at least two or three orders of magnitude less than what would be needed to pulse all the stator and rotor coils at once. Estimated classical current levels of less than 1 mA at 200 Volts have been proposed as being an upper limit. Mr. Wooten examined these motors from a mechanical point of view, using his professional expertise, and reported that each motor appeared to have logged at least several hundred hours of operation. Yet, you would never conclude that much use by looking at the contact surfaces alone. It is possible that the commutators may have been replaced, prior to being taken out of service, but that is a long shot.

Norman Wooten displaying the Non-Disclosed Complexities of the Timing Commutator from the EMA5 Gray motor at the 2001 KeelyNet Conference[5] – Courtesy Dr. Peter Lindemann



EMA4 CONTROL COMMUTATOR

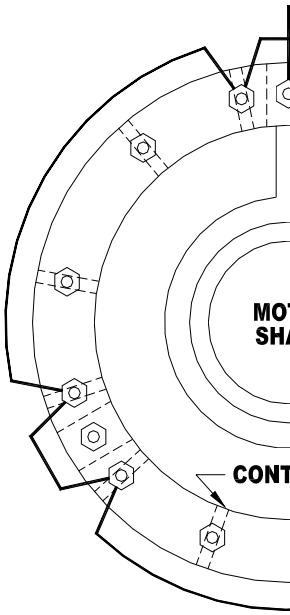
Observing the lack of wear, the new belief is that the commutators were providing both control timing and positional signals to Mr. Gray's energy converter. They were defiantly not directly switching the prime power that went to the stator and rotor coils. Further more, these timing signals were more complex than ever thought. In the recovered motors the commutator section and the motor electromagnets were wired independently.

Observing the lack of wear, the new belief is that the commutators were providing both control timing and positional signals to Mr. Gray's energy converter. They were defiantly not directly switching the prime power that went to the stator and rotor coils. Further more, these timing signals were more complex than ever thought. In the recovered motors the commutator section and the motor electromagnets were wired independently.

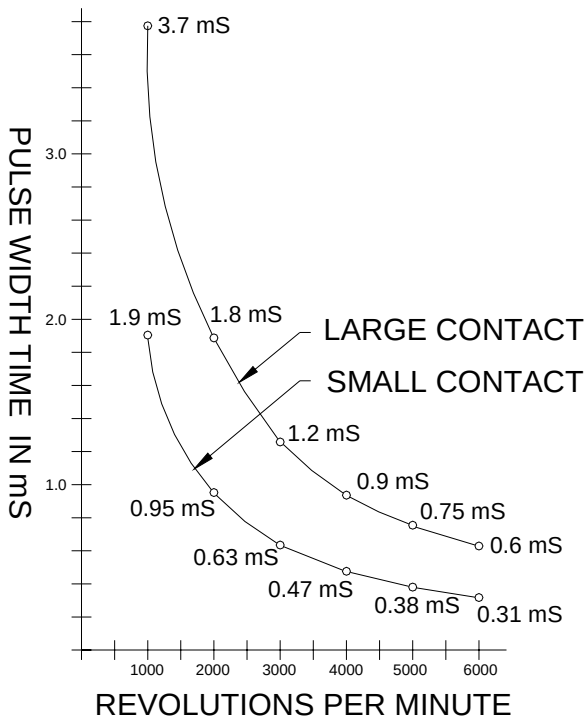
There are 15 contacts and two independent aluminum slip rings in each commutator subassembly. Three of these contacts are rectangular (1/4" x ¾") copper bars that are three times wider than the remaining ¼" diameter copper rod contacts. For both motors there appears to be two general timing patterns that emerge when looking at the angular spacing relationships of these contacts.

1.) The three large rectangular contacts and 6 of the smaller contacts are equally spaced 40° apart from each other around the circumference of the mounting ring. These would provide a continuous evenly spaced train set of short timing pulses, proportional to the speed of the motor, with every third pulse having three times the pulse width of the others. But, this is not what has been wired to go to the energy converter.

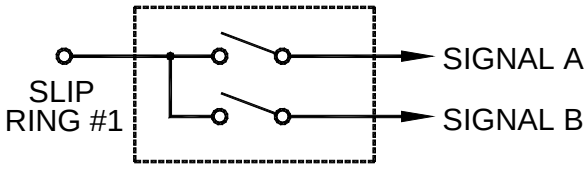
2.) There is also a repeated pattern with three clustered contacts. This group is composed of two small and the one large contact. These seem to be related to the "firing" of the electromagnets when the wiper is about 6° past TDC.



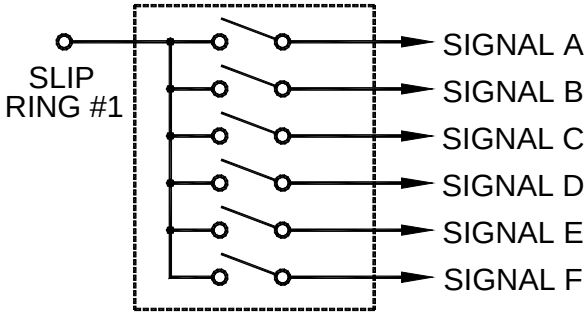
EMA4 COMMNETA



EMA4 COMMENTATOR
EQUILIVANT CIRCUIT

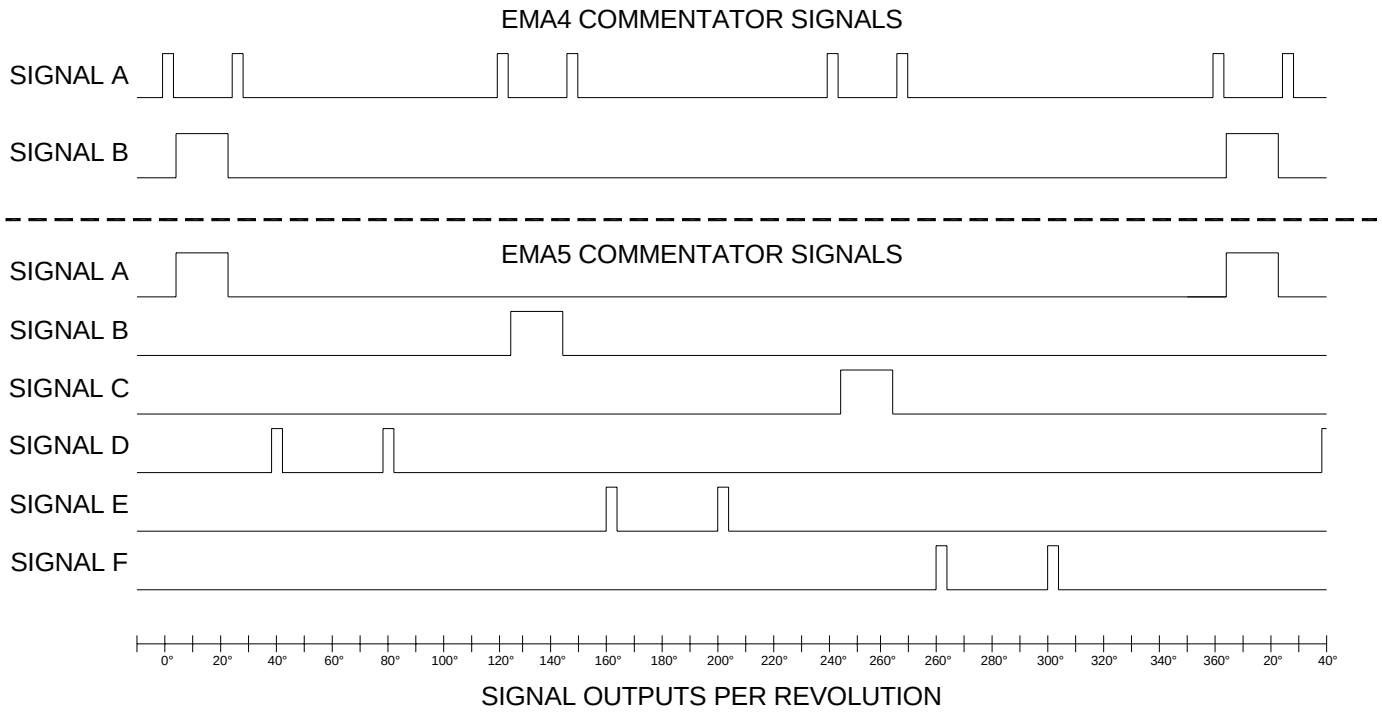


EMA5 COMMENTATOR
EQUILIVANT CIRCUIT



The rotary aluminum shaft wiper houses a spring loaded metallic "brush" that connects each contact to the slip ring in a sequential order. A second aluminum slip ring was installed, but was not utilized

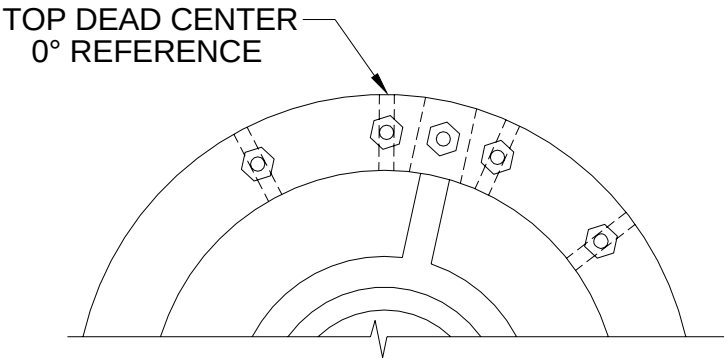
in the EMA4. If the slip ring were considered a circuit common then the timing pattern shown in Diagram 01 would be the result. Again not all of the contacts were used in either motor. This is indeed puzzling. Apparently different circuit configurations were being planned that might have used all these contacts.



Timing Diagram 01 for Control Commutators for the EMA4 and EMA5 EV Gray Motors

Mr. Gray used a construction technique that is not generally seen in rotary equipment. There are three slip ring assemblies used in each of these two motors. One assembly is used in the commutator subassembly and has two slip rings sharing a common wiper. The other two slip ring assemblies are used to conduct pulse power through the rotor electromagnets. One is in front and the other is in the back of the motor. All three of these slip ring assemblies have an uncommon internal design. This is because the wiper and “brush” are rotating around the inside of a stationary slip ring. This is just the opposite to 98% of all other industrial machines in the world that use slip rings. Almost always, the slip rings are attached to the rotating shaft and the contacts or “brushes” are stationary. The obvious advantage of this common approach is that it allows the brushes to be easily replaced when they wear down. Another important advantage is that the “brushes” can easily accommodate some imperfections in the roundness of the slip rings that rub against them. This is because the brushes are mounted in spring loaded holders that allow them to move back and fourth. However, in Mr. Gray’s design, a brush or wiper replacement would require way more disassembly. Also, it doesn’t appear that this design could allow for nearly as much deviation from tolerance as the standard brush and slip ring arrangement can. We just don’t know what the application specific reason was that promoted this kind of solution; it certainly is not obvious from looking at the motors alone. Mr. Wooten contends that he could have designed a much better system to get the power into the rotor as well as several other major mechanical system improvements. So far no one has disputed his claim.

It is interesting to note that the Top Dead Center (TDC), the position where the electromagnets are squarely aligned with each other, takes place when the wiper is on the first small round contact in the cluster of three contacts, rather that the larger rectangular contact. Mr. Gray designated this location as 0°. It has been proposed that a certain amount of angular displacement is needed between opposing electromagnets when operating in the repulsion mode to insure that the generated forces are focused in one direction. Perhaps Mr. Gray determined that the optimum angle, for this size motor, is around 6°. The actual working angular displacement could be adjusted. Perhaps this was just a convenient reference point and had nothing to do with the function of the motor.



According to the jacket information the control conductors leading off from the commutators are rated at 25KV. Yet, their overall diameter is equivalent to common #14 AWG THHN household wire (.12" diameter). This is much smaller than typical electronic high voltage wire that has this kind of voltage rating. This wire was probably an expensive specialty cable in its time.

The small spacing between the wiper and the contacts in the clusters of three suggests that Mr. Gray didn't utilize any classical control voltages that had a differential greater than 200V. If classical electron flow were involved then voltages higher than this would have caused arcing at both the leading and trailing edges of the contacts as the wiper approached and receded from them. Again arcing was not observed. Then what was the purpose of the expensive high voltage cable? One proposal is that all of the control voltages connected to the commentators were elevated to some high value and their differences was less than 200 volts. This means that the whole commutator was "floating" at some high potential above ground. The overall nylon construction of the commentator assembly suggests that it could have easily have supported this kind of high voltage operation (5KV to 20KV). The commutators on the EMA4, EMA5, and EMA6 are all mounted almost independently and external from the motor proper. This construction feature might imply a need for a high degree of isolation between the motor and the commutator. If so, then it is a distinct possibility that the commutator did operate at some high floating voltage.

The purpose of the various timing signals has been discussed within the Free Energy community but so far no general conclusions have been tendered that would explain how they affected the energy converter's circuit operation.

It appears that the energy converter needed at least two data streams, only a portion of which was the simple positional information. The rest of these short contact closures are assumed to be signals that could prepare the energy converter for its next pulse or to, perhaps, facilitate some kind of energy recovery cycle. There are four contacts between each TDC position; therefore there are provisions for as many as four changes of state per each power pulse. Not all of them were used at the time these motors were taken out of service, but they could have been.

Mr. Wooten, in his 2001 video, claims that the commutator compartments were filled with "Luberplate". This is the trade name for premium quality white lithium machine grease. Given that Mr. Gray didn't seem to spare any expense in the construction of this sub assembly, then what Norm could have observed might have been a special High Voltage Teflon/Silicon insulation compound that is used in the X-Ray business. This would have help to extend the voltage differential of Mr. Gray's control signals to maybe 500 volts or so. However smearing insulation grease (or any kind of grease) on moving electrical contacts is a risky business. This is because it is difficult to build a system that will reliably wipe all the grease off the contacts just prior to contact and still provide a consistent low resistance connection.

Both commutators were built so that the contacts are housed in a movable nylon ring. This ring was installed in a larger hollowed out cylinder that acted as a housing so that the whole collection of 15 contacts could be adjusted together in relation to the shaft position. A machine set screw allowed for a wide range of timing angle adjustments (-40° to +40°). At a setting of -16°, according to notes written on the commutator, the pulse motor would run backwards. Probably not at full torque, but this shows that these motors were reversible.

After the recovery of the EMA4 and EMA5 motors the idea that Mr. Gray's energy converters were dirt simple has come to be questioned. The revised thought is that the Mr. Gray's low energy technology may have been simple, but the higher power technology now appears to be more complex.

EMA4 Rear View

EMA4 Front View

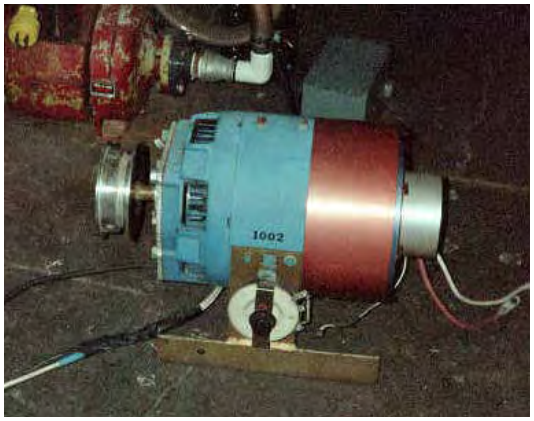
Photos of EMA4 and EMA5 motors are the courtesy of Mr. Norman Wooten via KeelyNet

Note: This document is one in a series produced by Mr. McKay as part of his investigation of the work of Edwin Gray senior and he invites readers to contact him if they have any constructive comments or queries concerning the work of Mr. Gray. Mr McKay's e-mail address is mmckay@tycoint.com

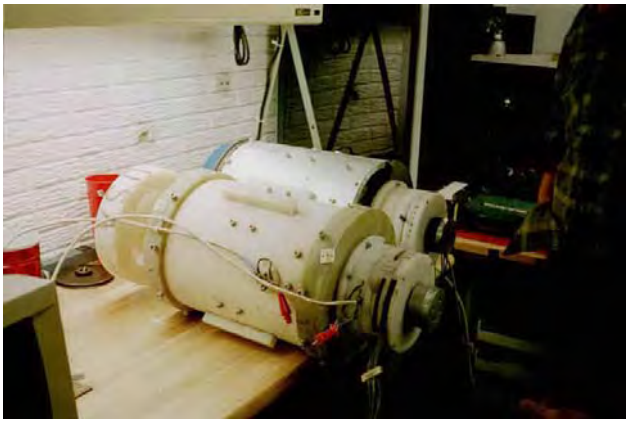
Mark McKay's investigation of Edwin Gray's Technology: Part 4

E. V. Gray Historical Series

Starting with the Start Motor — Mark McKay, PE



. The Start Motor as Found in 2000



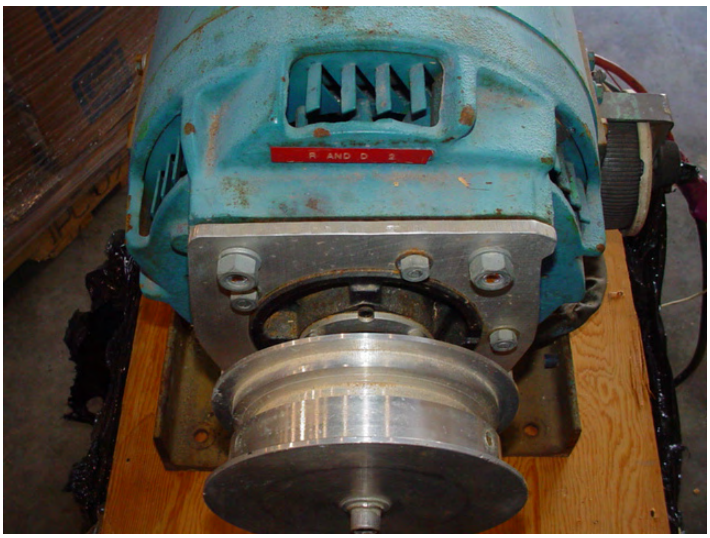
EMA4 and EMA5 Motors as Found in 2000

E. V. Gray once commented to John Bedini that his early free energy experiments were conducted with modified off the shelf industrial motors. It is assumed that when Mr. Gray's finally got adequate funding he went on to build a series of custom made motors that could take better advantage of the unique properties of his non-classical "Cold Electricity". These experimental designs were stamped with the model numbers EMA1 through EMA6. The EMA4-E2 and the EMA6 are his most well know constructions and are always associated with Mr. Gray's work. However, there were other transitional models built.

There may be one recovered example of a pre-EMA series motor that might have served as a functional test bed and very possibly an early investor demonstration model (circa 1963 to 1969).

In 2000 friends of Norm Wooten discovered two original EV Gray motors in a shop somewhere in Texas (most likely Grande Prairie, Texas where Mr. Gray had established a shop in 1986). These were the EMA4 and the EMA5 prototypes. Mr. Wooten acquired these pieces of history from the building land lord. He then took them to his shop where they were carefully disassembled. Later he produced a highly recommended video of his observations for the 2001 Keely conference in Florida. This informative tape is available from Clear-Tech at http://www.free-energy.cc/index.html in DVD and VHS formats. At the time the "Start Motor" was considered insignificant and therefore not looked at very closely.

After considerable mechanical analysis of the EMA4 and EMA5, Mr. Wooten came to the conclusion that this equipment contained no obvious free energy secrets. The vital energy converters that had powered these unique motors were not found. A few years later he decided to sell this collection.



Mr. Allan Francoeur of Penticton, BC, a long time free energy researcher and inventor, bought the entire lot for $5,000 US in 2003. This package included the two prototype evaluation motors (EMA4 and EMA5), one of Mr. Gray's advanced coil popping setups (partial), and an 1940's modified non descript industrial motor. It was assumed, at the time, that this humble looking machine was a high voltage (5KV) generator used by Mr. Gray to charge up his storage capacitors for motor experiments. Later it was proposed that it was a DC motor used to start up Mr. Gray's large experimental motors, thus it finally became known as simply the "Start Motor". The Start Motor could also have been thought to be a dyno-motor. In this capacity it could have acted as a dynamic load to evaluate the performance of Mr. Gray's energy converters.

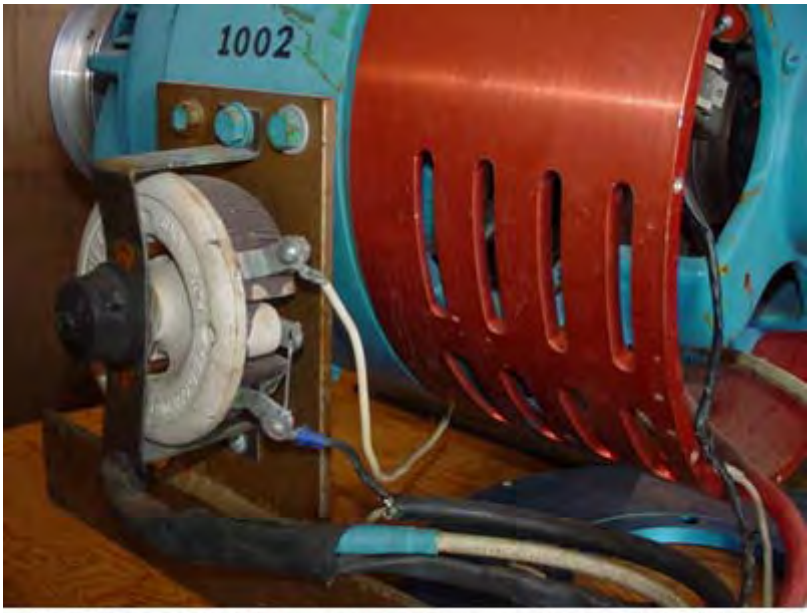
Custom Adapter Flange Added to Front of Motor

For a number of reasons this author contends that this piece of equipment was an actual working EV Gray pulse motor prior to the construction of the custom EMA models

Showmanship Tells All

Mr. Gray spent some serious money to have this simple motor dressed up way beyond any practical bench top need. If he wanted to conceal the details of its internal wiring from the occasional investor visit, then some heavy gauge sheet metal would have been a cost effective solution. Yet, this “Start Motor” was outfitted with a custom built three piece three color (Red, White, and Blue) anodized aluminum cowling set. The large red section was outfitted with a dozen small machined ventilation slots. These three pieces of non-functional eye candy probably cost him 50 times what the motor was worth, but may have been thought important enough, at the time, to help advance his early business development efforts.

As it turns out, the Start Motor is not a motor but a 5 KW DC exciter generator, circa 1940, used to provide field coil power for a larger generator (75KW to 150 KW). The 4-pole salient stator is outfitted with dual field coils that function in a compound wound configuration. It also has an independent set of slip rings that are connected to the armature coils and thus allow for external regulation. It looks odd, when compared to modern generators, because it has a commutator, like a DC motor, plus two additional slip rings like an AC motor. With the advent of solid state power rectifiers the slip rings and commutator bars in small generators have been completely eliminated, so you seldom (if ever) see this kind of construction. Externally mounted exciters have also been eliminated from the larger generator sets as well for much the same reasons. This same design was also called a “Three Wire Generator”. These were used in the 20’s to provide unbalanced three wire DC power for combination motor and lighting loads.



Side Mounted 200 Watt 2 Ohm Rheostat and Attached Cabling

Modification Details

Mr. Gray did a custom retro-fit to the front end of this motor. This modification was intended to be an adapter plate that would allow different flange mounted gear boxes to be attached. He also installed a simple magnetic probe in between two of the stator coils. The Start Motor was also reconfigured to receive its power through a #4 AWG cable (see the discussion about the cable used for the EMA4). There is a 2 Ohm 100 watt rheostat attached to the Start Motor’s side that has one #14 AWG cable going to one slip ring and the other going elsewhere (not connected). The return large red cable (ground?) was connected directly to the generator frame once it got inside the case. Having prime power travel through the frame of a generator or motor is defiantly not a traditional electrical practice. Except for the rewiring of the stator coils, the probe, and the cowling the rest of the motor appears to be “stock”. There were two suppressor capacitors associated with the slip rings that are similar to 50’s automotive distributor condensers. These seemed to be original equipment and had not been replaced. One of the slip ring brushes appears to have been replaced once.

Back End View of the "Start Motor"

The recovery and simple analysis of the Start Motor only reinforces what has already been suspected about Mr. Gray's technology:

1.) There is no obvious over-unity process to be found in this rotary converter. (But that doesn't mean there are none)

2.) This device was designed to have all the stator and rotor coils pulsed at once. This is an operational feature that appears common in Mr. Gray's motor systems.

3.) Applied Voltage considerations: The effective classical voltage potential of the energy that passed through this device certainly did not exceed 600 volts and most likely did not get beyond 300 volts. Had Mr. Gray exceeded these parameters, given the age of these exciter generators windings, he would have risked an insulation failure. The typical classical operation of an exciter generator like this was typically 120 VDC at 50 Amps.

Interesting Thoughts:

Why was Mr. Gray still hanging on to this early prototype demonstration motor (for some 15 years) in the first place? Technically, it would appear that it was a relic from his development past, when compared to the advanced EMA4 and EMA5 evaluation motors. He certainly paid good money to have this equipment shipped from his Van Nuys, CA shop to Texas, so it must have been of some value. The "Start Motor" weighs about 75 lbs. The best speculation to date is that Mr. Gray was probably saving his more important milestone pieces of equipment for a future exhibit in some national technical museum. If this is partially true then the importance of the "Start Motor" should not be over looked.

The schematic for the "Start Motor" below is the author's best attempt, with out disassembling the motor completely, to show the modified internal wiring.

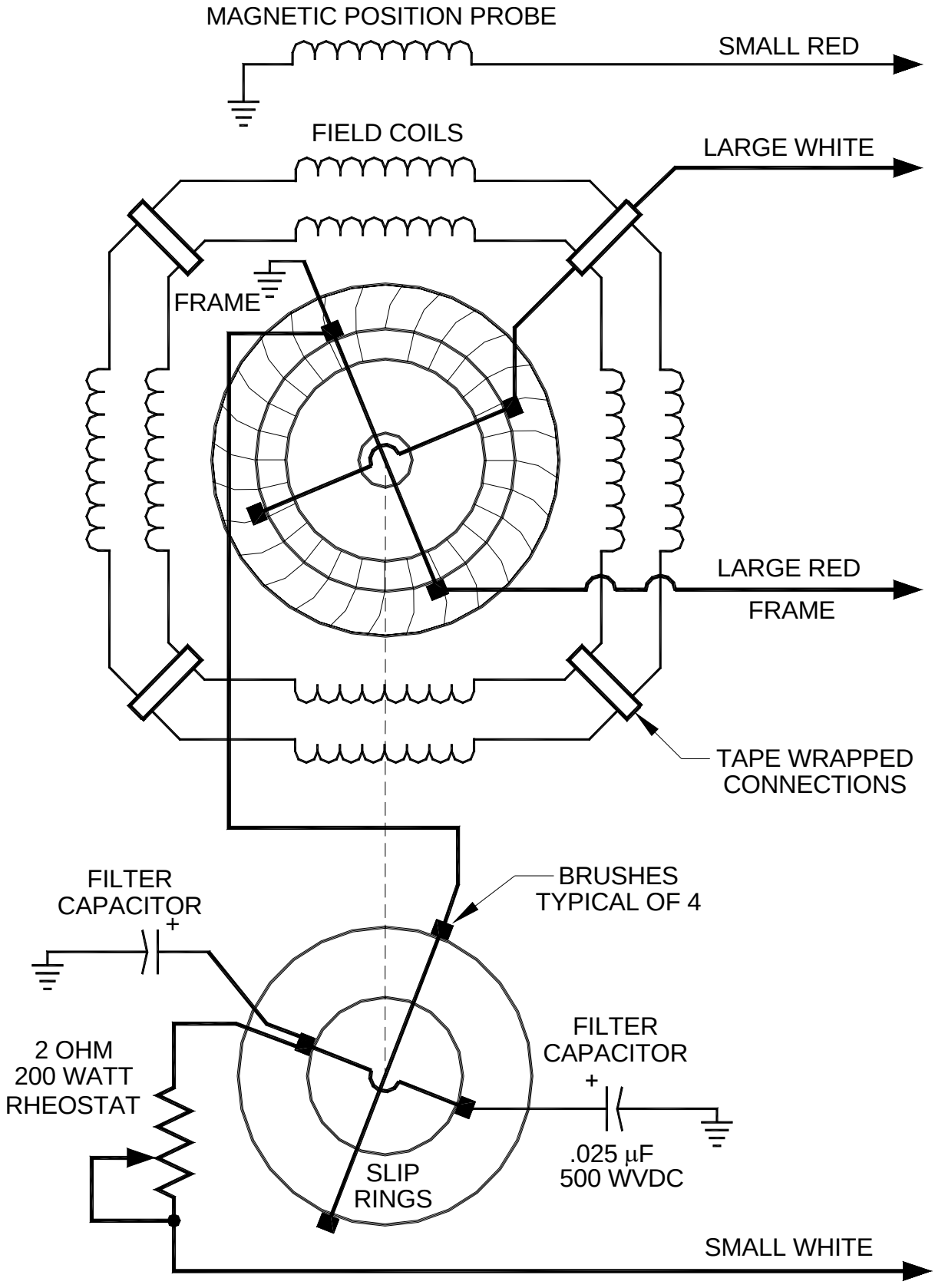
Added Magnetic Probe Next to Stator Winding
Assumed Used for Positional Feedback

Al Francoeur has taken very good care of this earliest surviving example of Mr. Gray's technology. It has been repaired, lubricated, cleaned up and now sports a new paint job. All that is needed is a reproduction EV Gray pulse energy converter to bring the "Start Motor" back to life.

If a breakthrough is ever re-discovered that unlocks the secrets of the methods used to create "Cold Electricity" then this modified exciter motor could well end up as a featured exhibit in the Smithsonian. This could have been what Mr. Gray intended all along.

Backend of the "Start Motor"

View of Compound Stator Coil and Slip Rings



EV GRAY "START MOTOR" SCHEMATIC
(PARTIAL)

Mark McKay's investigation of Edwin Gray's Technology: Part 5

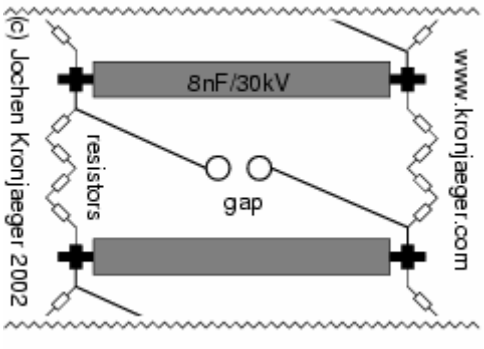
A Compilation of e-mail correspondence from Mr. Tad Johnson and other fellow researches concerning experiments with the “ED Gray” energy conversion device

From: Tad Johnson <h2opowered@c...>
Subject: ERE Produced by Accident **Date:** Thu Feb 13, 2003 2:18 pm

(Tad Johnson) Have a look at the bottom of the page explaining the "problems" Jochen has found when firing this 300KV Marx generator. Looks to be what we are after since he cannot seem to eliminate it through grounding and other means. Also look at the total conduction times (64uS) with rise and fall times substantially lower possibly in the 5-10uS range.

http://www.kronjaeger.com/hv/hv/pro/marx/index.html

“The discharge seems to induce huge voltage transients in ground and/or mains leads. This has resulted in a burnt mains switch and a destroyed ground fault interrupter. Grounding the Marx generator separately and decoupling the charging voltage ground with a resistor helps somewhat. This may turn out to be a major problem, as the Marx generator naturally produces a huge voltage step with a rise-time probably in the microsecond range, and the subsequent discharge produces a similarly steep current pulse which might be kA or more.”



© 2000-2002 Jochen Kronjaeger
hv@Kronjaeger.com
Last modified: 2002-09-08 15:41:04

(Tim Martin) Do you have a plan to allow for easily adjusting the frequency of the impulses? I think it will be important to precisely tune the device so as to discern specific effects.

(Tad Johnson) The frequency is adjustable to a degree through adjustment of the spark gap distance and cap size. The caps I am using are 500pF so frequency should be in the KHz range depending on how much amperage the power supply is charging the stack with. Just got the HV resistors today. All I have left to do is build the CSET and figure out the charging circuit. Hydrogen or magnetically quenched gap on the output might be added later for even higher frequency and more protection against current reversals.

Subject: folder added Hi folks, **Date:** Sat Feb 15, 2003 11:52 am

(Jani V.) I thought you might like to see my version on Ed Gray’s circuit In folder "romisrom" I just created, are some pictures of it, I will add complete schematic with component data as soon as I'm able to draw it...

Tad, I hope from picture "convtube" you will find some hints for your CSET. -Jani-

Subject: CSET design **Date:** Sun Feb 16, 2003 8:28 pm

(Tad Johnson) Thanks for the info. I was going to built it similarly although I was going to use 1.250" acrylic I have already to center the copper pipe. I have some new info on my power supply I will post soon. Looks like the rise time will be ~10nS with a pulse width of 50uS and a fall time of 40uS without a tailbiter circuit or resistive load of about .1Ohm to sharpen the fall time. I may add this later. Frequency should be about 25Khz as is.

Subject: Tesla/Gray device update **Date:** Thu Feb 27, 2003 7:08 pm

(Tad Johnson) My Gray device is now operational although I have foolishly fried a couple of neon sign transformers in the process of trying to loop the collection grid energy back to the power supply without some form of isolation circuitry. It appears I am now at the point that Gary Magratten was when trying to deal with a large pulse of energy and then measure it. Current circuit parameters are:

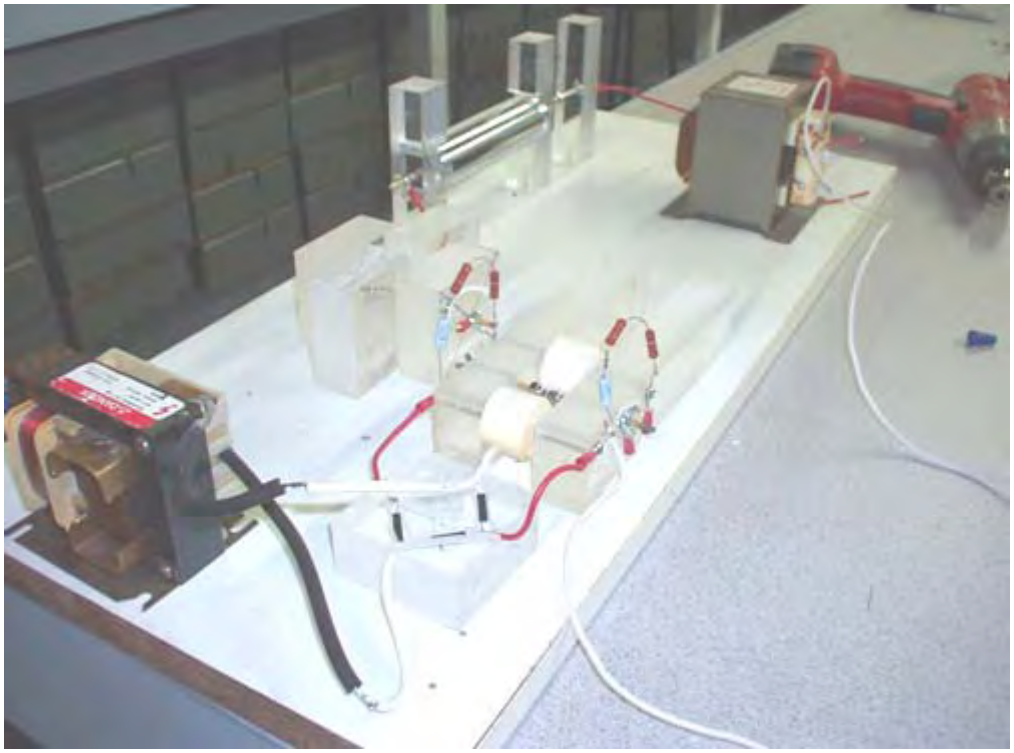
2000VAC @ 19.2Khz @ 20mA into a 12KV/40mA/100nS full wave bridge into a 2 stage marx generator using 400pF/ 30KV ceramic "doorknob" caps into a magnetically quenched spark gap using needle points of brass into the CSET of stainless steel balls on threaded brass rods. Collection grid is 316 stainless 2" diameter tube.

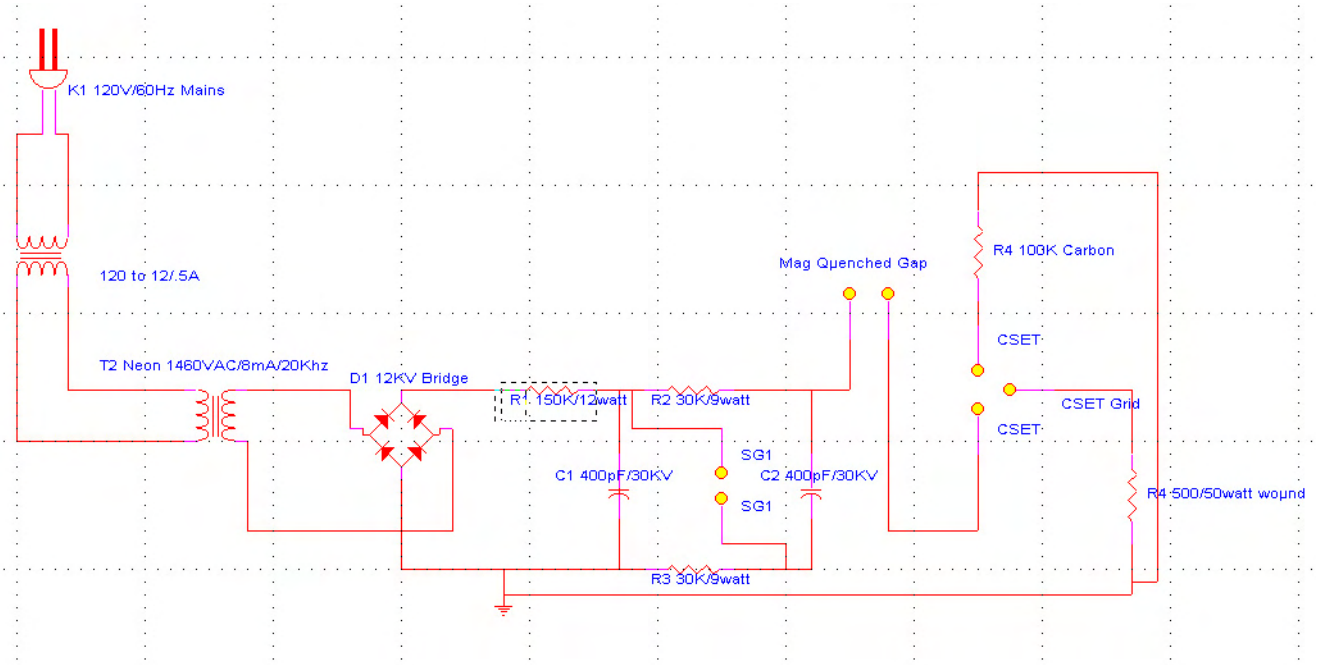
Total output pulse is 54uS wide with ~10nS rise and ~42nS fall.

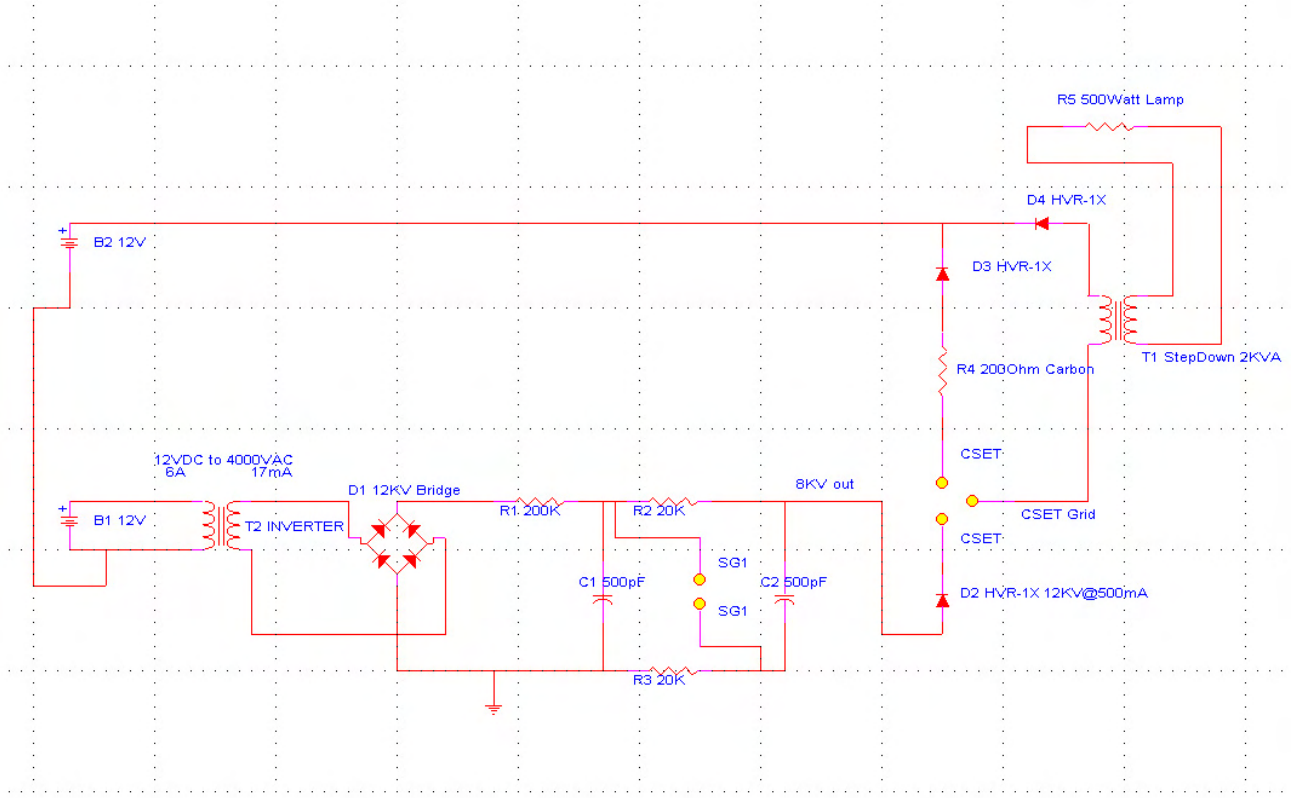
I am thinking of running the output energy in the secondary of a 3KV microwave transformer to power a lower voltage load although I am not sure how the transformer secondary will handle
this input, especially considering the frequency. Another option would be to increase cap size on the marx generator portion of the circuit to lower the frequency to something around 60-120Hz and then use it in a more conventional form.

Pictures and schematics to come soon. Any ideas are much appreciated.

Tad




K1 120V/60Hz Mains
120 to 12/.5A
T2 Neon 1460VAC/8mA/20Khz
D1 12KV Bridge
R1 150K/12watt
R2 30K/9watt
Mag Quenched Gap
R4 100K Carbon
CSET
CSET Grid
CSET
C1 400pF/30KV
SG1
C2 400pF/30KV
SG1
R4 500/50watt wound
R3 30K/9watt



Date: Fri Feb 28, 2003 8:25 pm

(Tim Martin) I have a few questions.

Is it possible to safely measure the voltage and frequency of the CSET output?

(Tad Johnson) Yes, I got the data below by making a 50Megaohm resistor to measure it, although I am reluctant to hook up the 3500 dollar scope to it as of yet. I get more guts to do so after I check the warranty info on it. All data thus far was taken on a true RMS LCR meter.
What is the AC current draw of the neon sign transformer? (Tim Martin)

Should be 1.5 Amp per the specs. But I will check it with my true RMS power-meter(5amp max on the meter).

(Tim Martin) Would it be possible to dump the CSET output into a large lead acid storage battery?

(Tad Johnson) Yes, although I am told it will "cold boil" at that voltage. Seems to be hard on the battery but I don't have much knowledge on it. I would like to step the voltage down before connecting it to the battery to avoid premature failure.

(Tim Martin) Would the neon sign transformer work properly if connected to a small >DC/AC inverter on the 12 volt battery?

(Tad Johnson) Should.

Subject: Gray Circuit Images **Date:** Sat Mar 1, 2003 10:19 pm

(Tad Johnson) New images uploaded showing the Gray circuit running after being tuned. Having issues with long runs because the resistors are not rated for more than 10watt on the Marx generator, they start to get a bit hot. Images show a 120VAC/60HZ/1.5A neon transformer powering it since my two other 12VDC inverters were smoked due to bad judgment. No connection to the CSET grid was present during this test run since I was mostly tuning the Marx stack to the 120V neon supply. Frequency was .5-1Khz on this test.

New power supply got here today so I will try the 12VDC version charging the Marx stack at higher frequencies (20Khz).

Flash on the camera makes it hard to see arc across gaps, but it is there.

Total cost of the entire device is now about $145 American dollars.

Subject: Re: [ElectroRadiantResearch] Re: Gray Circuit Images **Date:** Sun Mar 2, 2003 4:36 pm

(Tim Martin) I noticed in your pictures that you do not have a large high voltage air core as Gray and Magratten used in their circuits. Is this un-necessary?

(Tad Johnson) I am told the air core was a step down to run 120VAC/60HZ lamps and other resistive loads since resistive loads don't care about frequency. I haven't built an air core step down yet, but I might if I can't get a motor built soon.

(Tim Martin) Also, what did you say the clear "Plexiglas" material is? Real Plexiglas(tm) in those dimensions is fairly costly.

(Tad Johnson) Acrylic. Resists about 50KV in that dimension 1-1/8" thick. Very inexpensive. 1.5'X 1.5X square is 20 dollars. I used about half of one.

Subject: Grid Energy **Date:** Sun Mar 2, 2003 11:02 pm

(Tad Johnson) Interesting findings after running the Gray circuit for a couple hours:

ERE does NOT manifest if there is no resistor on the spark gap end of the CSET. Repeat ZERO POWER if no resistor in place. The more resistance, the more the effect appears to manifest.

With 300 Ohm or more of resistance the grid starts to put off a FRIGHTENING amount of power. Enough to smoke a 50watt, 500 ohm resistor in less than 30 seconds. My input was 12 watts
total from the wall. Output from the CSET grid is UNMEASURABLE. Grounding is also becoming an issue since I cannot run the end of the CSET back to ground with a resistor in between. Also,
the energy coming off the grid appears to be harmful even with fast rise and fall times contrary to other information out there.

Anyone have any bright ideas on measuring this high amperage, high voltage energy I would be very happy. We need accurate wattage out at this point. I feel confident already with my input measurements.

Subject: Re: [ElectroRadiantResearch] Re: Grid Energy **Date:** Mon Mar 3, 2003 11:05 am

(Tim Martin) It sounds as though Lindemann was correct in saying that one of the problems Gray had was dealing with the abundance of power.

(Tad Johnson) Yes, but we will see how much power. This is what I am after. If it is possible for a small 12 watt power supply to see a gain of at least twice that, then making the circuit for the application I am interested in will be easy (small motive power, scooter, etc.).

(Tim Martin) Do you think the CSET output is behaving different than "normal" electricity? What I am curious about is your statement regarding additional resistance increasing the effect.

(Tad Johnson) It appears as though there MUST be resistance at the end of the CSET in order for the CSET grid to make power. this appears to be the "bunching up" effect Lindemann was talking about, and that Tesla had experienced. It may be that when this HV pulse hits the resistance is like it hits a brick wall and explodes outward into the grid (path of least resistance).

(Tim Martin) Also, I believe that the frequency will govern whether or not the effect is harmful. Be careful!

(Tad Johnson) I'm being as careful as I can, but I have already had one small incident.

(Tim Martin) Another thing you might try is placing a normal 100 watt incandescent bulb on the output of the CSET without closing the circuit. Single wire power transmission is a related phenomenon.

(Tad Johnson) Yes, this works with a neon bulb, I've already run neon bulbs off the grid energy. they glow beautifully to full brightness.

Subject: Fwd: Re: [alfenergy] Grid Energy **Date:** Sun Mar 2, 2003 11:35 pm

(Willard)I can suggest putting a string of light bulbs together in series as a load. 5 bulbs of 100 watts each for instance.

(Tad Johnson) I will try that although I really need to somehow get an amp meter on it
and the scope. I had to drop the voltage down from 2920 to 1460 just so I could lessen the effect enough to work with the components I am using without it destroying them. Meter overloads when trying to measure grid voltage on the doubled setting from the Marx generator.
I am using a 100Megaohm, 100watt HV probe which should be more than sufficient for these voltages. Very strange.

Subject: Re: [alfenergy] magnetic quenched gap **Date:** Tue Mar 4, 2003 11:35 am

(Peer) The magnetic quenched gap is necessary to prevent continuously arcing. Is this right?

(Tad Johnson) No, it helps quench the arc, and bring the fall times back to something more normal. The waveform as per calculations is ~10nS rise, 50uS wide, with a long fall time, this is how Marx generators work. To bring the fall time back into ~20nS range we need to clip the end of the pulse. You can do this by killing the arc prematurely or you can put a low resistance load on the output of the spark gap (tail-biter circuit), or you can do both. My goal was ~10nS rise, 20uS pulse, ~20nS fall, with a pause of 500uS between pulses.

Subject: Re: [alfenergy] for Tad **Date:** Wed Mar 5, 2003 11:44 am

(Unknown Member) I'm trying to rebuild your circuit in order to better understand the working of the CSET. The original circuit built by Gray himself had a powerful input. Heavy batteries were used to power the circuit. You only use a small current und a much higher resistor at the CSET.

(Tad Johnson) Yes, my idea is to keep the power usage as low as possible but still see
the effect. And I have truly seen it with a 9-12 watt power supply, so it IS there. I am now lighting neon bulbs from the grid energy alone, this should not be possible since it would mean an energy gain of at least 100%, or an additional 9 watts to make a total of 18watts for the entire circuit.

http://www.amazing1.com/voltage.htm

At the bottom of the page you will see the power supply I am currently using (MINIMAX2)

ATTENTION! High Voltage Experimenters
High Voltage Transformers

Low cost thumb sized modules may be battery powered and used for experimental research in: Plasma Guns, Shock Wands, Anti-Gravity, Hovercraft, Tesla Coils, Ion Guns, Force Fields, Electrical Pyrotechnics, Stun Guns, Etc..

MINIMAX5 - 7000 Volt With IOG9 Plans.....**$29.95**
MINIMAX4 - 4000 Volt With IOG9 Plans.....**$19.95**
MINIMAX3 - 3000 Volt With IOG9 Plans.....**$17.95**
MINIMAX2 - 2000 Volt With IOG9 Plans.....**$14.95**
MINIMAX1 - 1000 Volt.....**$9.95**

Bag of five 2 to 3000 volt units-some requiring minor repair, others more.
MINIBAG1 - Includes Basic Schematic.....**$19.95**

(Unknown Member) I try to copy your circuit, using a medium size 6,5kV HeNe-LASER supply.
The output (grid-power) I get, is however tiny small.

(Tad Johnson) That's fine, my supply I use now is only 1460V @ 8mA!! But this voltage is doubled in the Marx generator. The Marx generator is used instead of the large capacitor and vacuum tube switch in the Gray patents. This eliminates the need for expensive and complicated switching techniques since the Marx generator switches on in less than 50nS and off in that
same amount of time unless you are running larger capacitors. 400pF caps @ 1460V @ 8mA gives me 500HZ. But 1900pF in that same supply only gives me about 1-2HZ, but much higher amperage pulse when the gap fires. If more amperage in the power supply (like 20mA) then this rate
would obviously be much higher and much more controllable.

http://home.earthlink.net/~jimlux/hv/marx.htm [Appendix 1]

http://members.tm.net/lapointe/MarxMain.html [Appendix 2]

http://www.kronjaeger.com/hv/hv/src/marx/index.html [Appendix 3]

(Tad Johnson) The capacitors come from:

http://www.alltronics.com/capacito.htm

The 400pF 30KV ones are US $12.50 each. The 6.5KV 1500pF are 99 cents each. The cheaper ones work just as well if not better! If you really want a big power pulse buy the 14uF, 20KV, 2800 joule
cap!

CERAMIC HI-VOLTAGE TRANSMITTING CAP
400pF @ 30KV, TC N4700. Made by TDK.
20P007 $12.50

SANGAMO ENERGY DISCHARGE CAPACITOR

14 uF 20KV 2800 Joule 14" x 8" x 24" --- Mineral oil filled

20P002 $250.00

(Unknown Member) Maybe there is a secret I have not seen yet. My CSET is not a pipe, but a
round cage made by copper wire soldered together. If a measurable radiant energy is made, this one I guess should be noticed by the small CSET grid I have.

(Tad Johnson)You WILL see energy on that grid regardless of it's design. I am using a stainless tube, but any copper, aluminum or anything else should work also. Multiple layers
of different metals (copper inside, aluminum outside should increase power as well).Also, move the CSET spark gap into the tube like Skip said. I should have done this as well, but I was lazy. This should maximize the energy on the grid. Use a couple neon lamps to run off the grid. 220VAC @ 10mA is what my bulbs are, I use two in series and they light up to full brightness off the grid energy alone. One lead to grid, one to ground. They light to half brightness just touching the grid and not grounded. I am trying to figure out what I was doing when I ran the 50watt resistor across the grid output in order to get it as hot as it was getting. This circuit grid output varies greatly depending on how it is tuned so there are many things to test still.

I really want to try a flyback supply soon though.

http://www.electronicsic.com/fly.htm

(Unknown Member) Maybe my quenched spark gap is not working. How is yours built up?

(Tad Johnson) I used a block of plastic on both sides and used a Forstner bit (1/2") to core a hole in the plastic, then I used glue to glue the ceramic magnet into the hole on both pieces of plastic. Then I used a router to make a slot so I could adjust the magnet distance from the gap electrodes. The magnets TWIST the arc and cut it off early, This gives us a faster fall time.

(Unknown Member) Have you enclosed the R4 inside the CSET tube or outside? Is it a high voltage type or a normal one?

(Tad Johnson) Outside and it is a normal 10K, 3 watt resistor, made by Panasonic, ordered from Digikey. The same resistors are used in the Marx stack. I have also tried a HVR-1X, 12KV/550mA diode (THV512T is new part number). This works well also.

http://www.electronicsic.com/diode.htm

POWER DIODES (Use in MICROWAVE OVEN)

BUY **THV512T** **12KV - 550mA** **$3.20 each**

Replacement For : HVR-1X-3 12KV - 550mA
HVR-1X-4 9KV - 550mA

Other diodes I bought were VG3, VG6 and VG12 from

http://www.amazing1.com/parts.htm

VG22	22KV HV Diode For KILOVOLT MAGNIFIERS	**$3.95**
VG4	3KV HV Diode - Used LGU4, IOG3, etc.	**$1.95**

[Apparently out of Stock on the VG3, VG6, and VG12 on 5/4/03]

Subject: Gray Circuit Modifications **Date:** Wed Mar 5, 2003 11:18 pm

(Tad Johnson) I finished my circuit modifications as per suggestions. I tripled the capacitance in the Marx bank, installed the CSET gap in the center of the collection grid and added a 25nF cap on the output of the CSET grid in line with the load. The lamps glow at least as twice as bright as they did before. But what is really exciting to me was that I was going to work on the Marx gap so I went to short the cap bank. At the instant I shorted this bank of caps I felt the "wave of energy" which actually pushed my shirt in the direction of the blast.

Has anyone else seen this when discharging a cap bank and being of close proximity? Very strange anomaly. Makes me believe that Tesla must have been working with much higher voltage and much higher capacity than this circuit in order to feel this wave constantly at each gap firing. This is obviously what we are looking to reproduce.

Subject: Re: [alfenergy] Magnetic Quenched Gap **Date:** Thu Mar 6, 2003 9:16 am

(Alan Francoeur) I have tested the function of a magnetic quenched gap. I used a Marx generator to create short HV pulses. The spark gap was simple two ends of a copper wire facing each other with a distance of about 2 mm. I used a vice and put a strong Neodymium magnet at each side of the vise jaw. The gap between the two magnets was about 17 mm. (The magnets were attracting each other) the arrangement was so that you could easily remove the vice with magnets without changing the spark gap.

Without magnets an arc occurred many times after a spark and the frequency of the spark was changing all times and there was a small interval without a spark, partially. From that view I can conclude the spark gap without magnet is not so well functioning because of the lower spark frequency and the occurring arcs.

(Tad Johnson) Yes, I have found this myself as well. This is why I like the magnetic gap so much.

(Alan Francoeur) With the magnets, the spark's frequency was higher, and there was no standing arc at all. Each time an arc liked to occur the arc got blown out like a candle in the wind.

When I was connecting a small (8 Watt) neon-bulb between the vice ,which was made of steel and somehow served as grid, and ground the neon-light lit weekly and the ark frequency changed a bit also the ark noise changed! And this although there is no galvanic contact between the Marx generator and the neon-bulb.

(Tad Johnson) I don't understand why frequency changes when you connect a load to the grid, but I have seen this as well.

(Alan Francoeur) But I also measured the current flowing back to ground after the mentioned spark gap. This was done by a 50 Ohm resistor a HV-probe and an oscilloscope.

(Tad Johnson) I am making a new HV probe, 1GOhm will be the size. A bit high, but I have many problems with the 100MOhm one I now use.

(Alan Francoeur) Without magnets: the time duration of the spark could be hardly measured but seemed to be >500 ns.

With magnets: the time duration of the spark was definitely shorter and the picture on the scope was more clear. The time duration was 100 us to 200 ns.

(Tad Johnson) Great! This is what we are after.

(Alan Francoeur) In both cases, you see a positive high voltage pulse that exceeds the capacity of the screen of the scope. Then a small negative pulse, like the half of a sine wave, follows. After that there are fast oscillations. Maybe this picture does not show the true current flow, because of parasitic capacities of the used resistor.

(Tad Johnson) The ringing is what has been messing my frequency counter up I think. I might not be getting the correct frequency of pulses measured. Inductors can be used in place of the resistors to reduce loss, although the output will obviously be different and need to be rectified or sharpened up.

(Alan Francoeur) Another investigation was, that using no magnet, a multi-discharge could occur (many tiny discharges). With magnet there was always one discharge. Maybe you have the same experience.

(Tad Johnson) Yes, exactly. This is why Tesla also used these magnets around the gap. He was trying for a smaller and tighter discharge of energy.

(Alan Francoeur) Tad, have you tried to put magnets inside the gray tube? Therefore you would not need to have a separate spark gap and maybe more power inside the Gray tube.

(Tad Johnson) I have not tried this yet, but I can try it soon.

Subject: Progress **Date:** Thu Mar 13, 2003 10:42 pm

(Tad Johnson) No progress on the Gray circuit this week as I have been working on getting a lathe to make parts and do better quality work so I have not been financially able to buy the HV resistor for measurement nor the Thyratron, or spark tubes.

I pulled my Hydrogen combustion enhancement device out of the shop since fuel prices are getting ridiculous. Car already gets 33mpg, but 38-40 would be better.

I will put pictures of it when I get it running again.

I will be working on the Gray circuit again within a week or two though. Stay tuned,

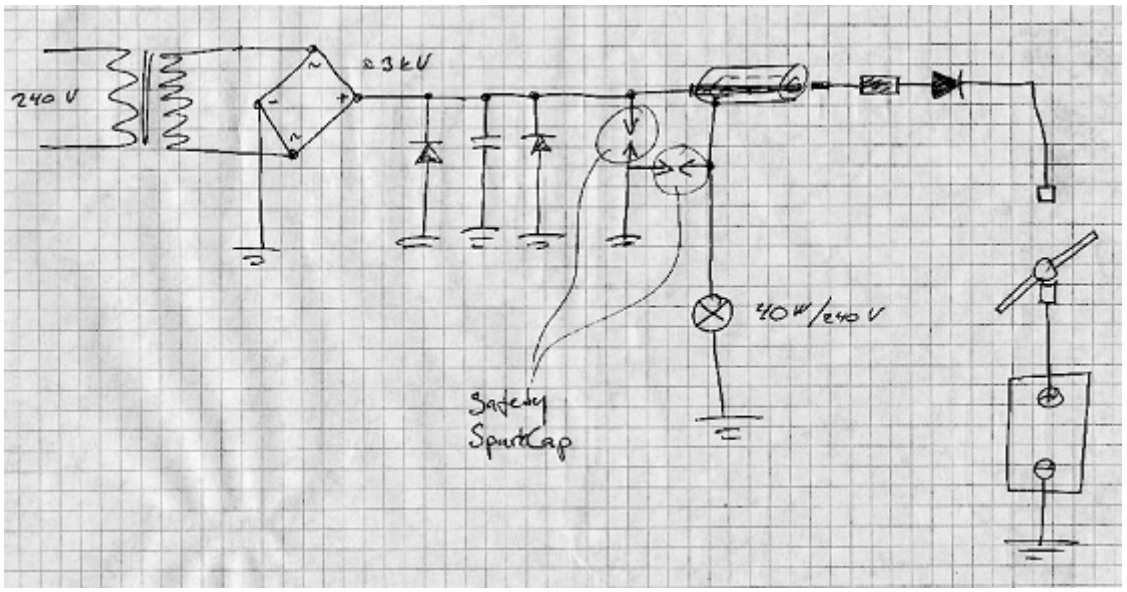
Subject: Re: [ElectroRadiantResearch] Success ??? **Date:** Fri Mar 21, 2003 9:17 pm

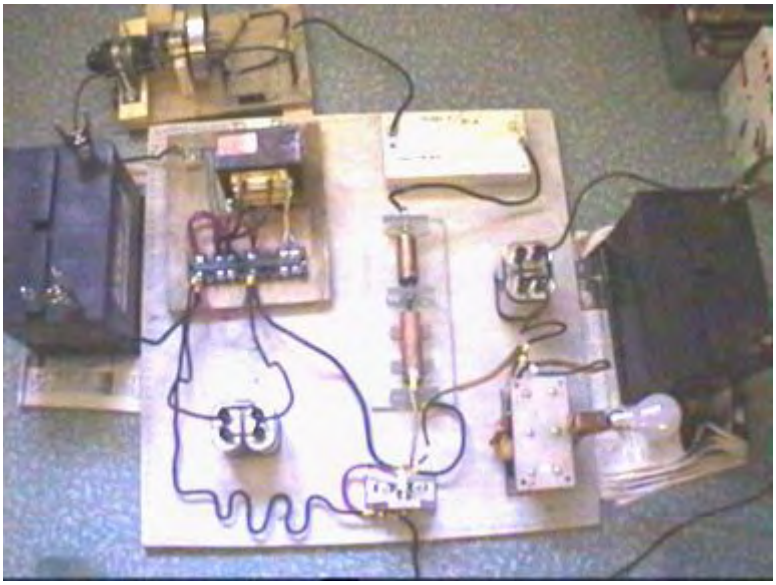
(Jani V.) Last weekend I finally got a chance to test my Ed Gray machine and I think the Electro-Radiant-Event manifested once. When I ran the test, 40 W light bulb flashed before the whole bunch of charge, which was collected to the grids,

discharge though the safety spark gap (schematic Test1a, look my folder romisrom). I tried to duplicate the Radiant-Event but it didn't manifest again. I think the interrupter-rotating rod burned somehow because it's resistance raised near two meg-ohms!!! I also have to make the carbon resistor different because it is not very stable, resistance range between 50 - 500 ohms depending temperature. I've also added in the spark-gap a strong NIB magnet to cut arc more faster. I think this magnetically quenched spark is very important to produce ERE. Anyway, test must be done again to make sure that it was ERE that manifest neither some other discharge.....unfortunately my testing is very slow because I live in another place due to my work and my test equipment are another place. So, it may take awhile.

(Tad Johnson) Congratulations!, sounds like a successful test run. You should get constant power off the grid once the circuit is tuned and stabilized. 300 Ohms on the end of the CSET seem to be perfect in my last test run.

Keep up the good work, no matter how slow it goes, it's worth it to humanity.





Subject: Progress **Date:** Sun Mar 30, 2003 5:21 pm

Hi folks,

I have not felt like doing much on the Gray device for a couple weeks since I have seen a relationship of mine fall apart after 8 years of being with this woman.

I am excited to see progress being made by Jani and Peer on their circuits and will hopefully find some "drive" to work on my system again soon.

Best wishes,

Tad

Note: This document is one in a series produced by Mr. McKay as part of his investigation of the work of Edwin Gray senior and he invites readers to contact him if they have any constructive comments or queries concerning the work of Mr. Gray. Mr McKay's e-mail address is mmckay@tycoint.com

Mark McKay's investigation of Edwin Gray's Technology: Part 6

Conversation between Mark Gray and Mark McKay on 5/19/07

Mark Gray is E.V. Gray's 6th child born in1958 in southern California. For the past several years he has been a parts-room manager for a school district repair shop which maintains over 200 buses. He is a single parent who currently lives with his three young adult children. (Two daughters and one son).

Mark Gray was employed by his father, E.V. Gray, for the majority of the time between 1979 and early 1988. In this time period, he served in the capacity of a general assistant. He traveled and worked at seven different locations, including a two week long trip to Israel.

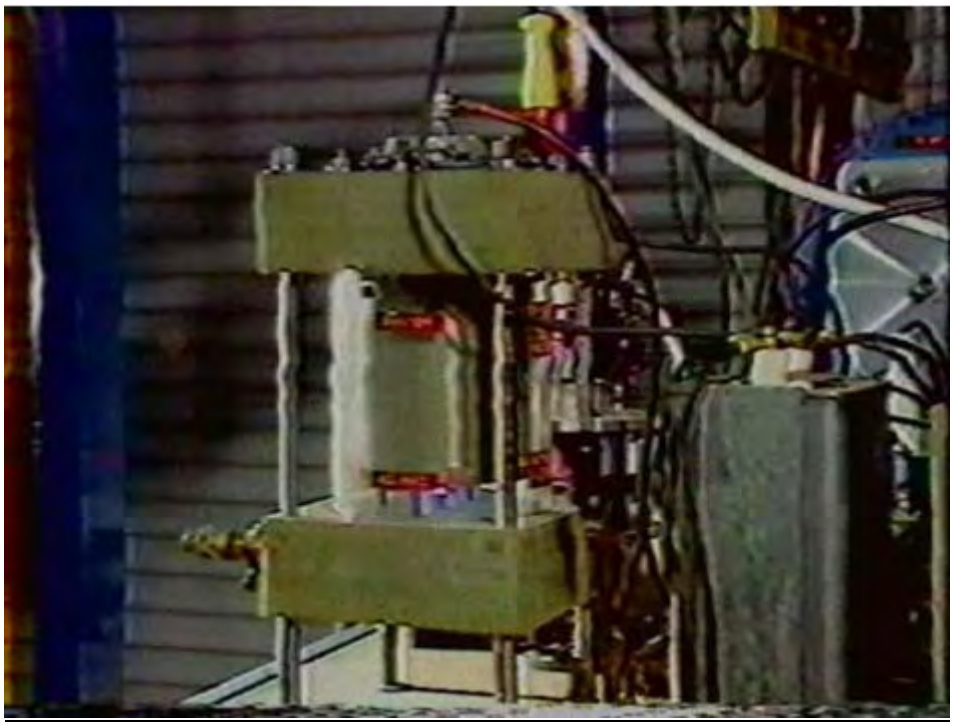
Under his father's direction he assisted in the building of the majority of the "Trigger Carts" (The converter systems under the pulse motors) that are displayed in the 1896 ZTEX promotion video. He also assisted in securing parts from custom vendors, video taped the technology, assisted with various demonstrations, drove the company truck, and wrote licensing agreements. These are just a few of the multitude of tasks he did during his tenure of service.

Mark parted on good terms from his father in early 1988 when funding ran out due to differences between E.V. Gray and certain investors, over the control and future of the technology. These differences were heightened when an alleged government contact, interested in a possible R&D program on the switching/triggering aspect of the technology, came into the picture late 1987 – early 1988.

While Mark had a tremendous exposure to his father's later technology (1979-1988), his detailed understanding of the underlying functioning principles is almost gone. He did what he was told to do and was compensated appropriately for his services, but never got deeply involved with the workings of the technology. For the past twenty years Mark has been completely divorced from his father's technology and has forgotten almost everything he knew about it. He regrets not having paid more attention and not having taken a real interest in the "nuts and bolts" of the processes.

Mark was most willing to share these anecdotal technical Tid-Bits that might have a bearing on rediscovering this lost technology.

The Mark I (Converter Switching Element Tube)



Tthe cylindrical glass enclosure is a Colman gas lantern cover

● COMMENTARY: This really limits the magnitude of the internal pressure of what ever gas may have been present. The size of the end caps could support pressures up to 6000 psi. With such a thin glass envelop anything over 3 psi would be difficult. "He didn't want to pay the high price for a machined enclosure"

● all electrical connections were made from the top

COMMENTARY: I only see two electrical connections at the top of this device (the black center conductor and the white conductor with the large yellow single pin connector. Therefore the "Grid" is not connected to anything, unless it is connected to one of the electrodes.

● the gap was adjustable

● the internal gas was presumed to be Nitrogen from a welding supply house

COMMENTARY: Mr. E.V. Gray was very familiar with welding gasses. "He didn't get involved with anything that exotic" (Referring to S6F)

● Purpose of the Grids: "Possibly to cover up something he didn't want people to see?"

COMMENTARY: Like an additional series component, perhaps an HV RF coil?

● Was there an electrical connection to the "Grids"? "I don't recall"

● "the electrodes were made of Tungsten or Titanium. Which ever material Russia is famous for." [Titanium]

Ignitrons installed on the "Red Motor Cart"

<u>The Mark II "Silver Cylinder" (Ignitron)</u>

● This was an off the shelf commercial device that was a metal cylinder about 2" in diameter and 6" long.

● The terminal insulators were glass

● It was a two terminal device only, with wires connected to the top and the bottom.

● The round flanges were custom made end pieces to secure additional finned aluminum heat sinks that were attached around the periphery.

● The band in the center was a radiator clamp to hold it all together. Sometimes two clamps were used.

● These units did occasionally wear out or fail. New units were stocked on the shelf

● These devices contained Mercury and therefore retired units were treated with respect in storage.

● When these units arced inside you could see a blue flash through the terminal glass.

COMMENTARY: It appears these devices are Class A Ignitrons. They are the right size, right form factor and contain Mercury. However an Ignitron is a three, or more, terminal device. It operates much like a very high current thyratron. If there were no control connections for the igniter, then one use might have been a fixed-distance spark gap and just overvoltaged until it fired. One advantage of this approach would be a clean Mercury surface after each pulse. The pulse rate observed in the 1986 video is on the order of 2 Hz.

It is unclear wither these ignitrons were a replacement for the CSET or components in addition to the CSET. So far, the best explanation supports the idea that the ignitrons replaced the function of the rotating spark gaps that were in the commutator section of E.V. Gray's early motor designs. The 1986 Promotion video will show that E.V. Gray used several of these devices for his motors (up to six per cart). E.V. Gray probably developed a new system where the complexity of the old front end rotary spark gap array was no longer needed, thus greatly reducing the fabrication costs per motor.

Magnet wire for the Popping coils:

●All the wire for the construction of the projectile coils was standard copper magnet wire

●One company was contracted to machine the aluminum or plastic coils forms (Normally Nylon). Another company was hired to wind the coils. "We attempted to wind a few of our own coils. But not many"

Wire used in special places:

"That wire there was the expensive silicone filled wire that had to be used at that connection" pointing to the photo of the battery charger converter and the wires coming off the storage capacitor.

COMMENTARY: In the Cannady Interview it was noted how "Cold Electricity' would destroy the insulation on conductors. Apparently E.V. Gray did find a tentative solution to this problem by using special wire in the locations where it was required.

A Trip to the Capacitor Vendor

Mark Gray recounted an experience he had when he was instructed to return some defective capacitors to a custom supplier in Southern California.

The internal connection between the external capacitor terminal and the internal plates had opened up because the wire gauge was too small, thus causing it to fail. To explore this complaint first hand, the vendor opened up one defective unit with a can opener. Since the connection had been separated at this point there was still a substantial charge still left in the unit. There was an unexpected accidental discharged that caused a loud bang. Apparently the vendor quickly made repair modifications to all of the returned capacitors at no charge. Mark reports that the plates were gray with layers of a white material in between them. The entire unit was filled with a thick clear gel. Mark Gray claims he recalls values of 500 mF at 5 KV.

COMMENTARY: This type of construction implies a low inductance plate capacitor rather that the higher inductance rolled designs. The residual stored charge implies a low loss construction. I don't know about the dielectric, it could have been a standard poly material. Another authority claims E.V. Gray used Mica. I don't know what color mica is when installed in a large capacitor. "Cold electricity" is also known for its loud discharges.

The "Trigger Cart"

Mark Gray claims that the heart and soul of the E.V. Gray technology is the “Trigger Cart”. This is the power supply that was the source of the anomalous energy for all of the projectile demonstrations. What is interesting about this system, is that it operates from 220 V AC, counter to all of E.V. Gray’s previous motors and circuits.

COMMENTARY: Some researchers have proposed that the E.V. Gray technology required the use of wet cell lead-acid batteries for the generation of “Cold Electricity”. Apparently this is not the case with the existence of this cart. However, the overall OU qualities of this technology may be impaired with the use of utility power. But at the time, E.V. Gray was seeking military customers who could benefit from the propulsion features of this equipment.

Trigger Cart Operation: "Slowly crank up the Auto-transformer until the tubes started to fire, then watch the volt meter. When it got to 5,000 volts I would quickly turn down the Auto-transformer and fire the projectile."

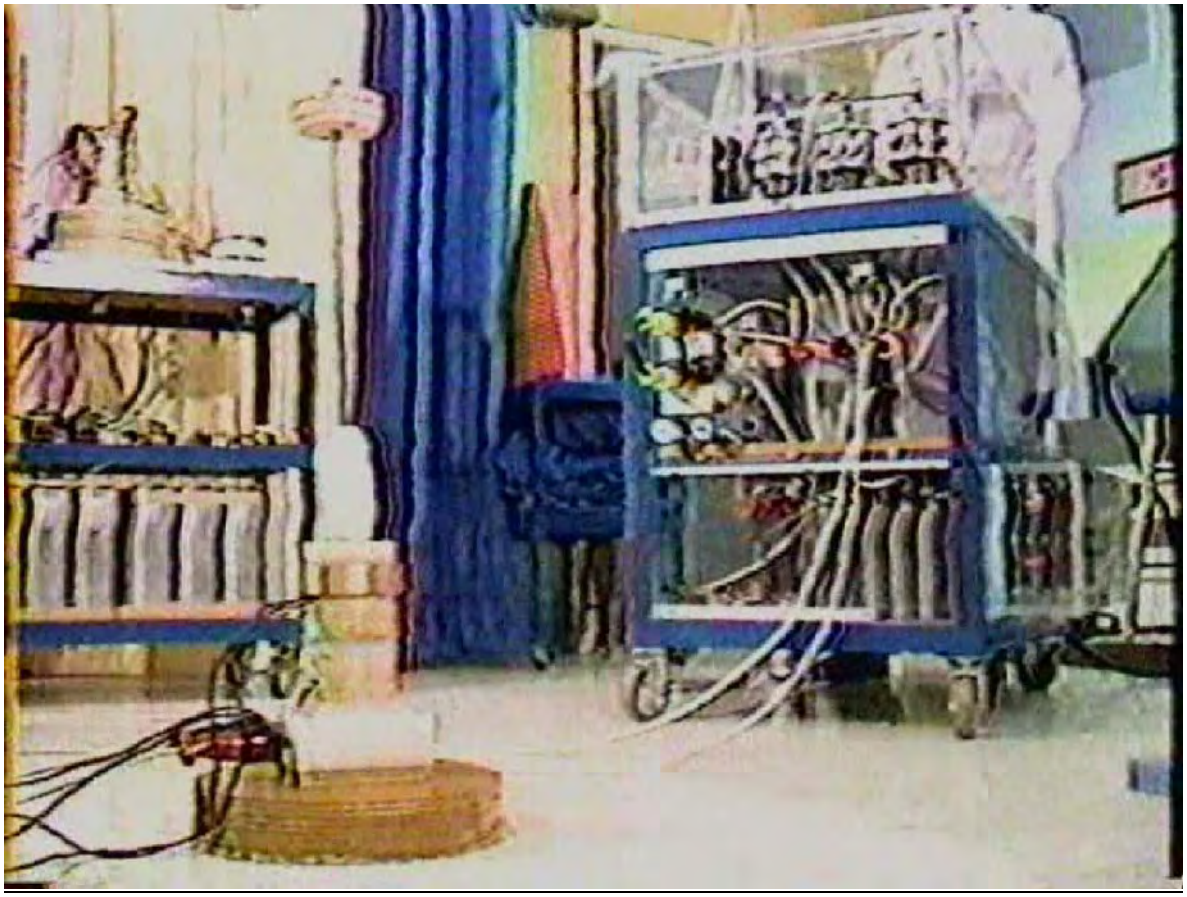
COMMENTARY: In the background sound of the demonstration video we hear about 20 pops before the projectile is ready for launch. It seems E.V. Gray was discharging one capacitor into another capacitor. Once this charging operation was complete he would discharge the collected anomalous energy through his opposing coils to launch a projectile. I don't know what he used for a discharge switch.

If Mark Gray was reading an analog voltage meter then we can be pretty sure that the anomalous “Cold electricity”, when stored in a capacitor, can be observed as a positive classical voltage. This is very consistent with Tom Bearden’s description of “Negative Mass Energy” - if the two phenomena are at all related. Earlier photos show E.V. Gray using an analog Triplett 630-A multimeter to measure the voltage of “Black Boxes” that are assumed to be storage capacitors in his early “Popping Coil” demonstrations (1973).

If the Pops we hear (20 or so per launch) are from the four Ignitrons on top of the cart, then it is reasonable to assume that the source DC supply voltage was in excess of 5 KV. If the Ignitrons were connected so that they would self-trigger by connecting the igniter to the anode, then there would be a sudden break-over pulse every time the voltage difference between the anode and cathode reached about 1500 V DC. This would imply that the source supply voltage was at least no lower than 8 KV.

Since there was a concerted effort to turn down the auto-transformer after reaching 5 KV, I would guess that E.V. Gray was charging his custom capacitors right to their design limits.

Auxiliary Capacitors:



COMMENTARY: In this photo, note the “Projectile Cart” on the left. Six different types of projectile are launched from this demonstration platform. The bottom of this cart contains a pretty substantial capacitor bank array. You can see only 70% of the cart. This would imply that there are about 9 large capacitors in the first rank. If two rows are employed, then a total of 18 capacitors are needed. I suppose this sort of stored energy was needed to support the “Hover” demonstrations or the large 71 lb launch.

Mark Gray claims that this cart was in E.V. Gray’s possession at the time of his death. He plans to enquire among family members as to where this piece of equipment went.

COMMENTARY: It is my contention that if this cart was saved from the one way trip to the surplus re-seller, then who ever got it couldn’t make it operational. According to Mark Gray, his father spent his last days disassembling this equipment. This system would be high on the list of things to do first.

“Split the Positive?”

When asked if his father ever told him about the fundamental energy conversion process Mark Gray recalled one experience where his father told him “The energy starts from the positive terminal [of the storage capacitor/dipole] then part of it goes back to the supply battery and part of it goes to the load

COMMENTARY: This type of topology is shown in patent 4,595,975, but the actual technical meaning is anybody’s guess.

The “Wireless Projectile”

Mark Gray claims that some potential investors would ask “What good is this system if you have to have wires connected to projectile? That is not going to work”. So he developed this demonstration apparatus to show that the projectiles really didn’t need wires. Actually, they are needed for only a short distance, beyond which the magnitude of the repulsive forces drops off quickly. The above setup provided a sliding contact that is in the little black & white tower on the left of the larger black cylinder. This arrangement allows for about 6-8” of travel before electrical contact is broken. By that time, the travelling mass has received most of the shock impulse it is going to get. The black repulsing coils are composed of copper magnet wire that is about 2” deep. The outside is covered with black vinyl electricians tape. Mark also said that it was hard to reconnect the sliding contact because of rotation after a shot. Apparently it took a broom stick and a ladder to rest the demo.

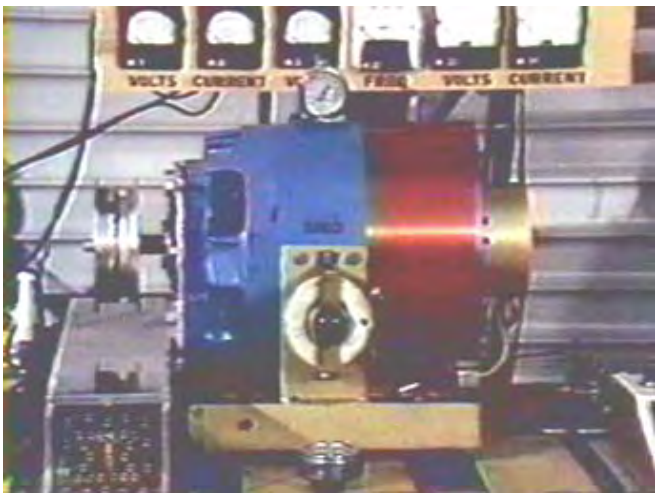
COMMENTARY: The measurable voltage of the energy that propelled the small black cylinder on top with the (white plastic saucer on the bottom) was said to be 5KV. Now look at the length of the arc trail [about 12”] of the little contact tower (at the left) after lift-off. Consider what kind of voltage was being generated at this point.

<u>The State of the Storage Batteries prior to a test or demonstration for a Motor Cart</u>

“When a motor cart was prepared for a test (or demonstration) both sets of batteries were fully charged”

COMMENTARY: So much for the idea of having to start with a dead battery. This theory comes from the idea that the lead-sulfite was the medium that might have converted a pulse of classical electricity into “Cold Electricity”

<u>Another Cold Electricity Demo using the “Start Motor”</u>



The white round dial instrument sitting on top of the “Start Motor” on the Multi-demonstration Cart is a thermometer. The other round dial instrument lying down on the table just below the round rheostat is a mechanical RPM indicator. [Biddle Meter]

The Importance of the Spark Gap

E.V. Gray told Mark Gary that the spark gap was very important.

COMMENTARY: A lot of other researchers think so too.

The Purple Motor

A Family Group Photo

Motor Names:

While the older E.V. Gray motors were numbered, the newer versions in the 80's were named according to a color. There was the Red Motor, The Blue Motor, The Purple Motor, The White Motor and the Black Motor. Each one was intended to demonstrate some particular aspect of this technology or head off any common questions that had continually arisen over the years.

Stump the Expert Time:

Once, a professional researcher, from MIT, was allowed to examine the equipment while development was taking place in Canyon Country, CA, (Possibly for some investor review). He had flight arrangements to leave the following Monday and had the whole weekend plus a day for his investigation. Apparently there were no restrictions placed on what he could look at. This man was alleged to be one of the co-inventers who developed the first anti-shark repellants. He examined and observed for at least one whole day and then made a comment to the effect, "If I can't figure this out, then all of my academic training is worthless". He worked all through the weekend and left the following Monday with no tentative classical explanation.

COMMENTARY: It would sure be nice to see if this individual would grant a phone interview. I'm sure he didn't talk a whole lot about his experience when he returned to Boston. I wonder if he would now?

Other Questions Asked through e-mail:

To your knowledge did your father (or his assistants) own or use any of these common electronics shop instruments?

Oscilloscope
Radio Frequency (RF) Generator

General Signal Generator
Pulse Generator
Transistor Tester
Q-Meter
Grid Dip Meter
Frequency Meter
Digital counter
Capacitor Tester
Battery Tester
Spectrum Analyzer
DC Power Supply

Of course any information about a general description, perhaps a Make and Model number (ha,ha), and an idea as to what the instrument was used for. When it was used and by whom.

Response 1) There were some meters involved, but I do not remember what meters might have been used or for they would have been used for.

2) The "kernel" of the technology appears to reside on the circuit trigger boards and the specific wiring to the off board components. From the photos we know that large power transistors were used. It is pretty obvious that other board components were used as well.

Do you happen to know what kinds of major components were on these boards? *We can assume that there were a number of supporting resistors and small capacitors*

Silicon controlled Rectifier (SCR)
Control Relays
Large Power Resistors
Transformers
Inductors or Chokes
Radio Frequency Coils
Vacuum Tubes
Diodes
Rectifiers
Power MOSFETS
Varisters
Potentiometers - Variable Resistors
Others
Model number of Power Transistors?

Of course a general description, approximate count, and any idea as to their function would be helpful.

Response 2) The most knowledgeable on the circuit boards may be Nelson 'Rocky' Shlaff (or Schlaff) from the Los Angeles area. I do remember that the circuit boards were developed in Canyon Country and for awhile the services of an electronics consultant was acquired to help development some of this circuitry. I do not remember the name of the consultant.

3) We know that you did a majority of the work on this equipment.

Was there any specific part of these "Carts" that your father reserved for himself to work on exclusively?

Response 3) Actually, my father did not protect any specific area of any of the technology that I can remember. Many people had cast their eyes on and all over the technology that was built. Nelson Schlaff and myself did most the assembly of the technology. There were others from time to time that were involved with the technology built.

4) Concerning the "Trigger Cart". You said that during its operation you would charge a certain capacitor to 5,000 volts before launching a projectile. You also said the voltage input was 220V AC.
Here are some general questions about the over all construction of the cart.

What Size Breaker was needed to power the "Trigger Cart" 30 Amp, 40 Amp, 50 Amp, higher?

Was a transformer use to raise the voltage from 220V AC to a higher voltage?

If 5,000 volts was the final measurable output voltage, then was there a higher voltage used somewhere else in the circuit that you know of?

Were Inductors or "Chokes" included on this Cart?

Did you ever have to make repairs on the "Trigger Cart", if so what was replaced and how often?

There are 4 "Ignitrons" on the Trigger Cart. Were all of these used at all times, or did different demonstrations use a different number of these devices?

Response 4) The only thing I remember about the voltage was charging the capacitors to 5,000v ?? for a one-time discharge (propulsion of a magnet), however, the hovering of magnets was achieved by a constant firing of the tubes.

5) Concerning the origins and nature of the transistor circuit boards used for the "converters".

Were these circuits made in house or contracted out? Did you make them? Did the design change over the years? If these boards failed who repaired them? Were replacements kept on hand?

Response 5) I do not recall much, if any was needed, maintenance on the circuit boards, nor do I recall having any made up as spares. I believe that all R & D and constructions of the technology happened in-house.

Mark McKay's investigation of Edwin Gray's Technology: Part **7**

Edwin Vincent Gray (1925-1989)

Edwin Gray was born in Washington, DC in 1925. He was one of 14 children. At age eleven, he became interested in the emerging field of electronics, when he watched some of the first demonstrations of primitive radar being tested across the Potomac River. He left home at 15 and joined the Army, but was quickly discharged for being under age. At 18 he joined the Navy and served three years of combat duty in the Pacific. He narrowly escaped death when a bomb exploded on his ship's deck during an attack. He received an honorable medical discharge after spending some time in a navel hospital with head injuries.

After World War 2, he married his first wife, Geraldine, and started a family in Maryland. He worked as an auto-body and fender repair man. In 1956 he moved his family to Venice, California. A few months later he moved to Santa Monica where he began his first business named "Broadway Collision". A couple of years later, he opened a second shop in West Los Angeles. Both locations failed early in 1960 due to an economic downturn. He relocated to Prescott Arizona, and then to Littleton, Colorado in 1961. From 1962 until 1964, he worked in Las Vegas, Nevada, always in the auto-body repair business.

By 1965, Gray relocated to southern California again, and established a partnership with George Watson. Watson was a master car painter with an established clientele of Hollywood celebrities. A new location was established in Van Nuys, California on Calvert Street called "The Body Shop". It was a one-stop, high-end custom auto-body & painting shop. This business prospered well for the next three years until a conflict of romantic interests ended his first marriage (with seven children) in early 1968. A divorce followed in 1969.

(In 1971, Gray married Renate Lenz, the daughter of Fritz Lenz. They had three children. This relationship lasted 7 years. Gray married three more times after that.)

Towards the end of 1969, Gray terminated his auto-body business, never to practice it again. He sold 2/3^{rds} of the Van Nuys building to his nephew and re-outfitted the remaining portion to build and promote his next business enterprise. Somehow, Ed Gray had made a sudden and dramatic shift from the auto-body business to an independent inventor with an extraordinary technology, with hardly any previous background in electronics.

Members of his family are still baffled by the quick transition. Some say their father was occasionally struck with flashes of profound inspiration. Other researchers say that Gray must have been working secretly on the motors for years, but family members dispute this. Gray himself told one of his partners that he received this information from a Russian immigrant named Dr. Popov, who had gotten it from Nikola Tesla. But again, family members claim no knowledge of these supposed events. While there are similarities between Gray's technology from 1970 and Tesla's "Method of Conversion" technology from 1893, there is no known lineage to trace the connection between these two processes. No one ever saw Gray studying the work of Tesla, or running any preliminary experiments. No one who is still alive, who was associated with these events, knows where the technology came from or how it developed.

In 1971, Gray formed a limited partnership named EVGRAY Enterprises, Ltd. By 1972, Gray had gathered enough investment and development expertise to build a 10 HP prototype motor. This unit was submitted to Crosby Research Laboratories for evaluation at Cal-Tech. Crosby Research Institute was owned by Bing Crosby and run by his brother, Larry Crosby. This motor demonstrated an output of 10 HP (7460 watts of mechanical energy) for the extremely low electrical input of 26.8 watts. This is an apparent energy gain of 278 times the input! This left the Cal-Tech scientists very uncomfortable. The report states the motor operated at "over 99% efficiency", but the rest of the data is a little confusing.

On the strength of this report, Bing Crosby came on board as a major investor. So did 'Boot' Mallory, of the Mallory Electric Company, who made the high voltage ignition coils used in Gray's circuits. By early 1973, EVGRAY Enterprises, Inc. had completed a 100 HP prototype motor called the EMA4-E2. Fifteen private investors were now involved. Ed Gray also received a "Certificate of Merit" from Ronald Reagan, then Governor of California, during this period.

By the summer of 1973, Gray was doing demonstrations of his technology and receiving some very positive press. Later that year, Gray teamed up with automobile designer Paul M. Lewis, to build the first fuel-less, electric car in America. But trouble was brewing when a disgruntled ex-employee made a series of unfounded complaints to the local authorities.

On July 22, 1974, the Los Angeles District Attorney's Office raided the office and shop of EVGRAY Enterprises, and confiscated all of their business records and working prototypes. For 8 months, the DA tried to get Gray's

stockholders to file charges against him, but none would. Since he only had 15 investors, many of the SEC regulations did not apply. By March 1976, Gray pleaded guilty to two minor SEC violations, was fined, and the case closed. After this investigation ended, the DA's office never returned any of his working prototypes.

In spite of these troubles, a number of good things were happening. His first U.S. Patent, on the motor design, issued in June of 1975, and by February 1976, Gray was nominated for "Inventor of the Year" by the Los Angeles Patent Attorney's Association, for "discovering and proving a new form of electric power". Despite this support, Gray kept a much lower profile after this time.

But there were also other set-backs. Paul Lewis pulled out of his deal with Gray in 1975 when Gray couldn't deliver a production motor for Lewis's Fascination car. Gray made a last ditch effort to secure the needed capital to get his motor into production by calling a press conference in 1976 and demonstrating his nearly complete, second generation 100 HP motor, the EMA-6. Unfortunately, this event didn't secure any additional funds for the company. Shortly thereafter, Bing Crosby died in 1977, followed by 'Boot' Mallory in 1978. This left Gray without his two strongest supporters.

In 1979 Gray reorganized himself into ZETEX, Inc. and EVGRAY Enterprises, Inc. ceased to exist. In the process of this corporate restructuring, all of his earlier stockholders lost all of their money. Gray then moved his development operations to Kalona, Iowa where new investors were supporting his research. This working relationship also failed when these new partners attempted a hostile take over. In a sudden midnight flight, in the middle of winter, Gray loaded up the technology with all his belongings and headed to San Diego, CA where stayed for 18 months.

In 1982, he relocated his operations to Canyon Country, California where he hired three assistants to help build several large demonstration carts. After a year of work, Gray got suspicious of the loyalty of his employees. He abruptly fired all of them when they reported for work one morning. He then moved to a second location in Canyon Country and continued with the construction until early 1984. Later that year, he moved his operation back to Las Vegas where he stayed till the spring of 1985. In the summer of that year, he moved to the almost abandoned town of Council, ID (population of 816), where his oldest son 'Eddie' had settled down.

In Council, Gray finished up the construction of five different motor prototypes and several other kinds of demonstration equipment. He then began to produce promotional videos and invited local TV stations to report on his work. Gray then sought out the services of a Wild Cat oil exploration lawyer and found Mr. Joe Gordon of Texas doing work in Montana. The two men formed a partnership under Mr. Gordon's established business Western States Oil. They also established a branch holding company in the Cayman Islands from which to sell stock in the new venture. Gray decided to move again, this time to Grand Prairie, Texas to improve his exposure to international investors.

On the strength of his videos alone, the Cayman Island operation was selling stock and raising capital quickly. Interested investors from Israel convinced Gray to spend two weeks in the Holy Land where a series of emotional group negotiations took place. An agreement was never reached. They conceded that the technology held a lot of promise, but it was not mature enough to be immediately employed on the battlefield. In addition Gray insisted on maintaining a controlling interest in what ever deal was cut. For whatever reasons, Gray came back with a much different attitude.

Meanwhile the agents who had been selling his stock in the Cayman Islands decided to give themselves large commissions, plus whatever other funds they had control of, and quickly move to Israel themselves. Apparently, they had also oversold the original stock issue by about three times.

Feeling swindled himself, Gray made a final, desperate attempt to get proper recognition for his achievements. He actually wrote letters to every member of Congress, Senators and Representatives, as well as to the President, Vice President, and every member of the Cabinet, offering the US Government his technology for Reagan's "Star Wars" program. Remarkably, in response to this letter writing campaign, Gray did not receive a single reply or even an acknowledgment!

In 1987, a person named Reznor Orr presented himself, claiming to be a "Government Contact". Mr. Orr first made straightforward offers to buy all of Gray's technology outright for a modest price. These initial proposals did not meet with Gray's approval, and he turned them all down. At about this time, Gray's income stream from the Cayman Islands stopped. Mr. Orr's next offers were much less friendly, and mixed with certain veiled threats. When Mr. Orr left town, "to let Mr. Gray think about it", Gray realized he had a serious problem. Out of money and under threat, he quickly held a massive liquidation sale, including personal belongings and family furniture he had had for years. Only the equipment and materials he could stuff into his Ford F-700 box van were spared. Gray drove to Portland, Oregon and hid out for six months.

Some time during 1987 - 1988, Gray became ill with a serious case of pneumonia and was hospitalized. He had been a heavy smoker all his life. He never fully recovered from this illness and required Oxygen from this point on. His reduced lung capacity made it much more difficult to continue his work.

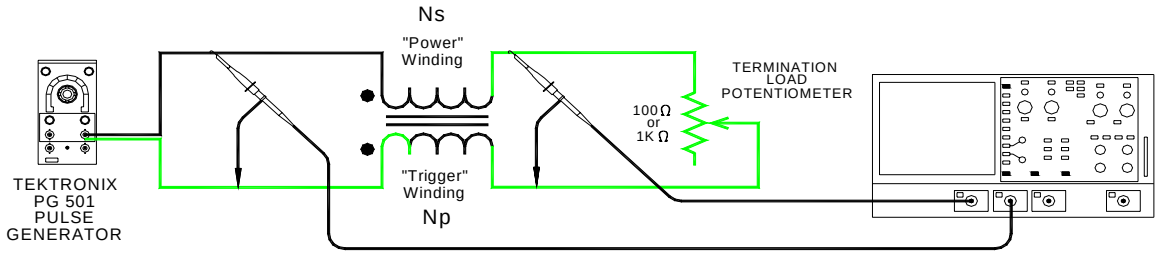
From Portland he moved to Sparks, Nevada. Gray rented a combination living quarters and shop space in a light industrial area. He unloaded his truck and began to disassemble all of his demonstration carts. He was living with Dorothy McKellips at the time who claims that Gray still did experiments during the day but in the evening all the components were once again taken apart and mixed with other parts. Early, one morning in April of 1989, about 2:00 am, somebody suddenly started banging hard on one of the shop windows. Gray, in his compromised health condition, got out his gun and went down stairs to frighten off the intruder with a warning shot. The gun failed to fire. A few minutes later, Dorothy found Ed on the floor. It is presumed that the resulting stress caused Gray to suffer a fatal heart attack, although the exact cause of death was never determined. He was 64. The identity of the late night visitor is not known.

Gray's oldest son "Eddie" flew to Sparks, Nevada to identify his father's body. Later, he spent several months attempting to help a Kansas group recover the technology. But, Dorothy would not release any of Gray's equipment until she had received a large payment for herself. The Kansas group then got a court order to take possession of the technology. But the document was poorly worded and did not define exactly what "technology" really meant. The order did state that they had rights to all of the motors. Dorothy caught this fact and gave them just the bare motors, keeping all the power converters and other things in her possession. Dorothy then decided to have the last laugh before this looming legal battle could escalate much further. She had all the remaining equipment, videos, parts, drawings, and laboratory notes hauled away and dumped in the local land fill. Apparently none of the remaining systems that the Kansas group had on hand were complete enough to reconstruct. Meanwhile, the remaining millions of dollars of investor capital in the Cayman Islands bank account were tainted by the fraud of the over-sale of the stock. Ultimately, these funds were either confiscated by the local government in fines or simply swallowed by the bank, since no one could withdraw the funds without being arrested.

[This account of the life and times of Edwin V. Gray was compiled by Mark McKay, of Spokane, Washington, after numerous interviews with a number of Ed Gray's surviving children. This account is an attempt to piece together the most accurate retelling of Ed Gray's story ever made available to the public. Many of the details in this account are in direct contradiction of earlier accounts as reported in the newspaper clippings from the 1970's. These earlier accounts should now be considered to be in error.]

Mark McKay's investigation of Edwin Gray's Technology: Part 8

Evaluating Common FE Coupled Inductor Systems in Terms of Delay Line Parameters



DETERMINING DELAY TIME T_d & CHARACTERISTIC IMPEDANCE Z_o

Coupled Inductors are a central component in a number of established Free Energy technologies. They have been used by Robert Prentice, Marvin Cole (E.V. Gray), Eric Dollard, John Bedini, Stan Meyer, and possibly Lester Hendershot. This is in addition to the vast array of coupled inductors that Dr. Tesla employed in his decades of research. Generally, modern independent researchers approach these devices from the standpoint of classical transformer theory and tend to view their operation in this way. I propose that, in many cases, these devices were intended to be used as Transmission Lines or Delay lines to take advantage of the unique features available with this topology. This is especially important when the characteristics of a high energy sparks are being engineered to achieve fast rise and fall times (<10 nS).

Volumes of detailed technical books are devoted to this complex subject. Specific applications are numerous because so many power and information signals are carried by transmission lines of one sort or another. However, in the realm of Free Energy the function of a Delay line appears to be relatively straight forward. Its common purpose is to act as a special kind of DC charged capacitor that will quickly deliver a fixed amount of disruptive energy to a spark gap. In applications that don't involve a spark, like the John Bedini motor, it is used (among other purposes) for sharp transition pulse formation using the same principles of operation.

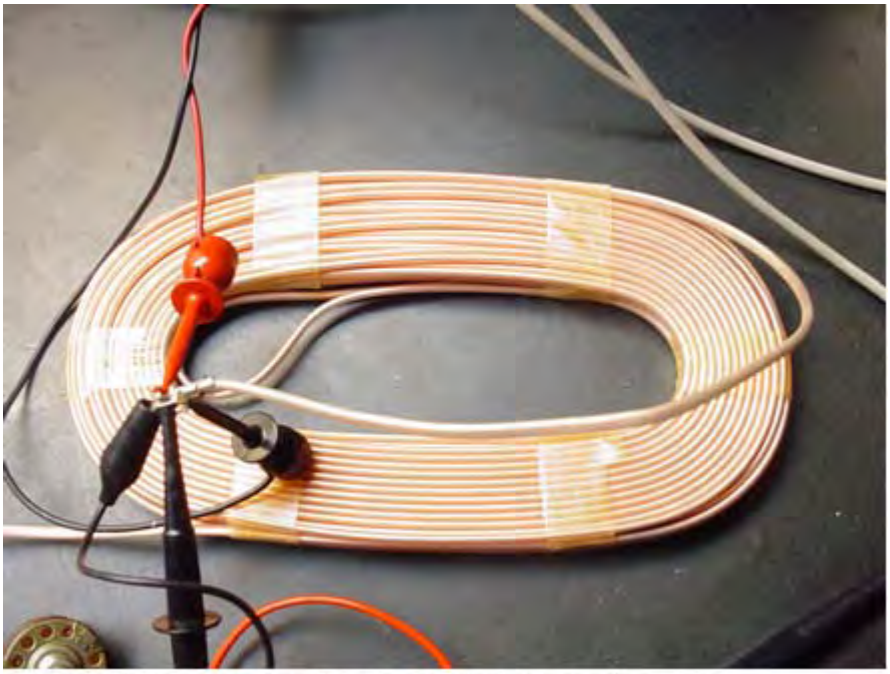
There are two measurable parameters of a Delay line which are the foundation of most engineering analysis that will involve these devices.

1) The effective voltage time delay from one end to the other, abbreviated as T_d measured in seconds

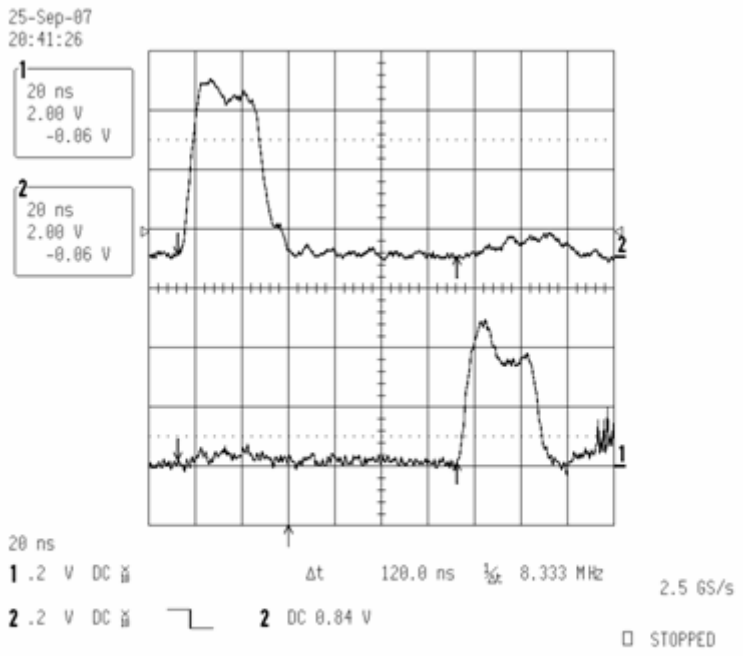
2) The characteristic impedance Z_o measured in Ohms

Both of these values can be easily measured with standard electronics equipment. This paper will utilize a LeCroy 9361 dual channel 300 MHz Oscilloscope with two standard 10:1 10 Meg probes and a Tektronix PG 501 pulse generator. A Fluke 87 VOM will be used to determine the resistance of potentiometer settings.

A good place to start this subject is to observe how a commercial Delay line functions. In this example an old 465 Tektronix oscilloscope twin-lead vertical input Delay line is evaluated. To best see its operation, the PG 501 was set to the narrowest pulse it could produce (25 nS) and applied directly to the Delay line input. A 100 Ohm potentiometer was set to 50 Ohms and connected to the Delay line output. The second oscilloscope probe was connected in shunt with the termination potentiometer.



Vertical Delay line for Tektronix 465 Oscilloscope



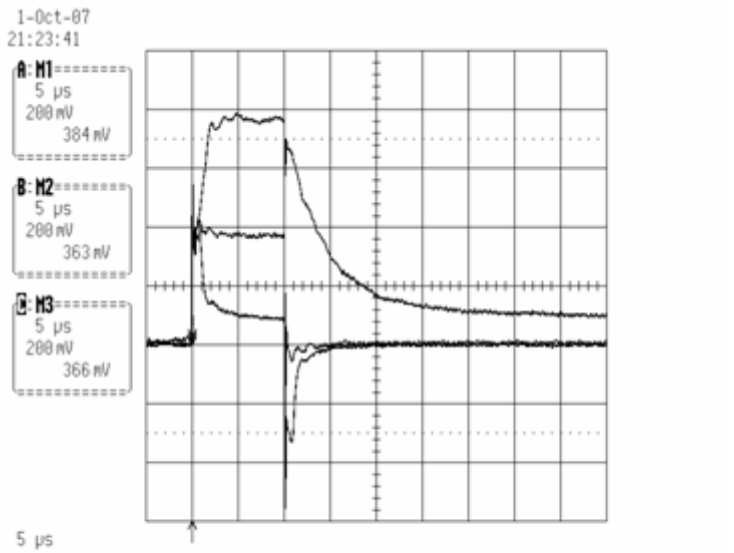
Resulting Trace using Two Probes

The two-channel trace from the oscilloscope (above) clearly shows the input pulse (Upper trace on Channel 2) and the output pulse (Lower trace Channel 1) delayed by 120 nS. While this straightforward approach will easily determine the delay time in a very low loss instrument Delay line, establishing delay times in homemade coupled inductors requires a different approach. If this present method were applied to most real-world coupled inductors, the output pulse will become so attenuated that it will be barely visible. The degradation of the input pulse increases as the coil under test becomes larger.

As it turns out, the energy in a 25 nS pulse is just too feeble to be observed in any homemade coupled inductor. This is because the parasitic capacitance filters out all of the high frequency components. Short pulses are just swallowed up in the unavoidable losses inherent in hand-wound inductors. However, another simple method, using the same equipment, can be employed to overcome these limitations. If the test input pulse is widened to some convenient length (to increase the applied energy) then the reflected pulse wave forms can be viewed. The actual delay time will be ½ of the observed time between the leading edge of the applied pulse and the change in response that is caused by the termination resistance.

A good example would be to make measurements on a typical Bedini SG motor coil. The coil being measured is a bifilar design using #19 AWG magnet wire for the “Power Winding” and #24 AWG magnet wire for the “Trigger Winding” with 420 turns wound on a Radio Shack wire spool. The soft iron welding rods used for the core were removed.

Typical John Bedini SG Bifilar Motor Coil



Dynamic Pulse Response

The first step is to establish the value of a load resistance R_L that will closely match the effective Z_o of the coupled inductor under test. This is done by applying a suitable pulse to the input of the Delay line (in this example we are using a 10 uS pulse) and then storing three traces:

a) Upper Trace: Delay Line is open at the output end

b) Middle Trace: Delay Line is terminated to a potentiometer adjusted to match Z_o Adjusted for “maximum squareness”

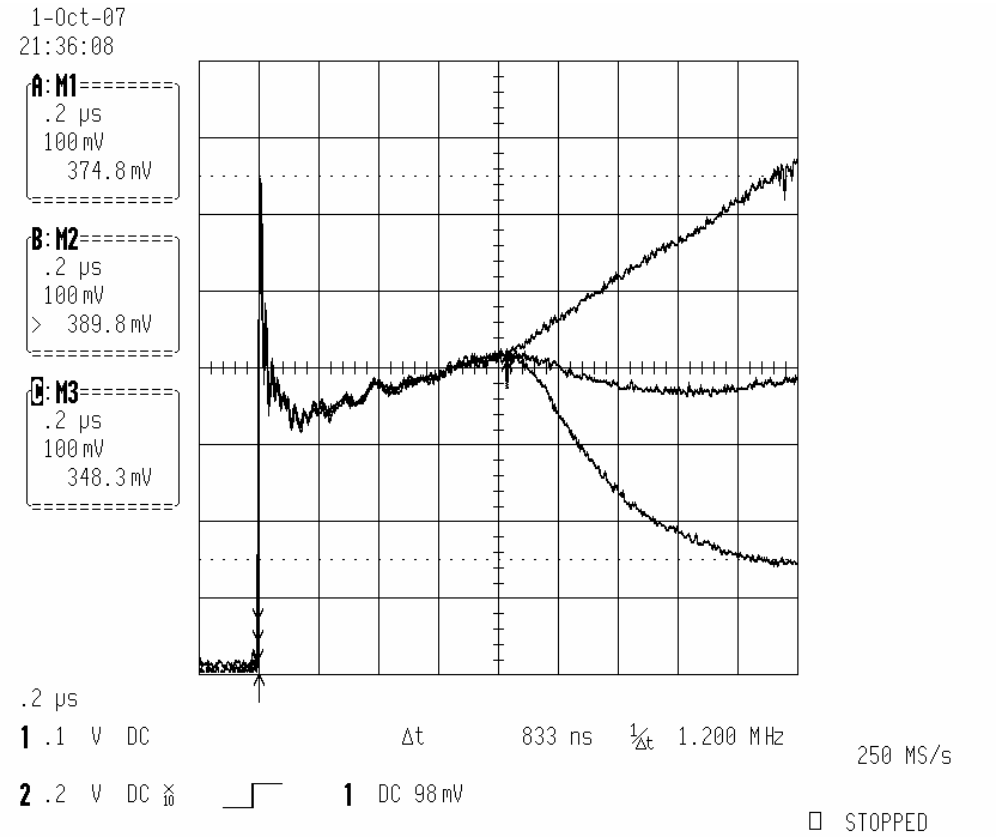
c) Lower Trace: Delay Line is shorted at its output end

What “maximum squareness” means is a matter of personal taste since there is always ringing and overshoots to have to deal with. However, when the potentiometer is close to the optimum value, small variations will make a big difference in the observed shape.

When the potentiometer is “dialed in”, it is then removed from the test bed and its resistance value measured with a VOM. In this example the result was 40.6 ohms.

If the iron welding rods are inserted into the core, no observable change is noticed in this series of measurements.

The next step is to expand our time base on the above pulse and store another three traces, following the same procedures as above.

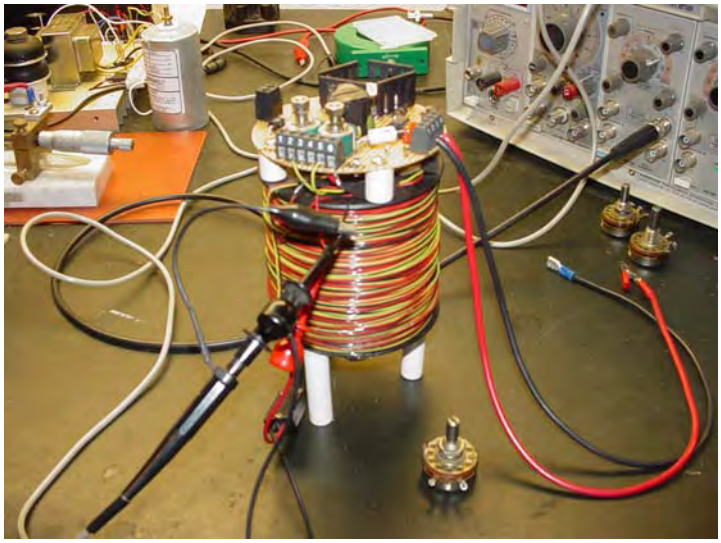


Leading edge of a pulse applied to a Bedini SG coupled inductor under three load conditions

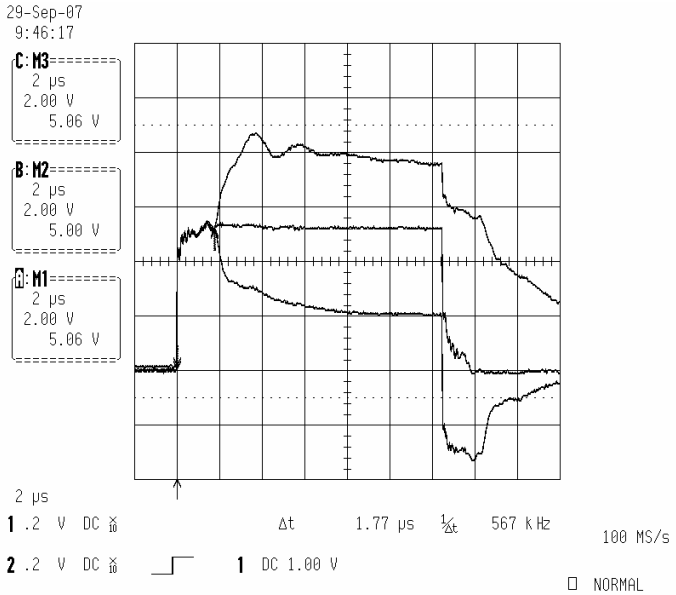
Here, the time base has been expanded by a factor of 10X to view the leading edge of the applied pulse at 200 nS/div. The upper trace is the open condition. The middle trace is done with matched Z_o loading and the lower trace is the shorted condition. All three of these waveforms converge at one point. This point establishes how long it takes the applied pulse leading edge to travel to the end of the coupled inductor and return. The kind of load it finds attached at the end, then determines how it will respond from there on.

Measuring the time between the leading edge and this intersection, then dividing by 2 we arrive at the one way Delay Time for the coupled inductor under test. For this Bedini Coil we measure a T_d of 415.5 nS.

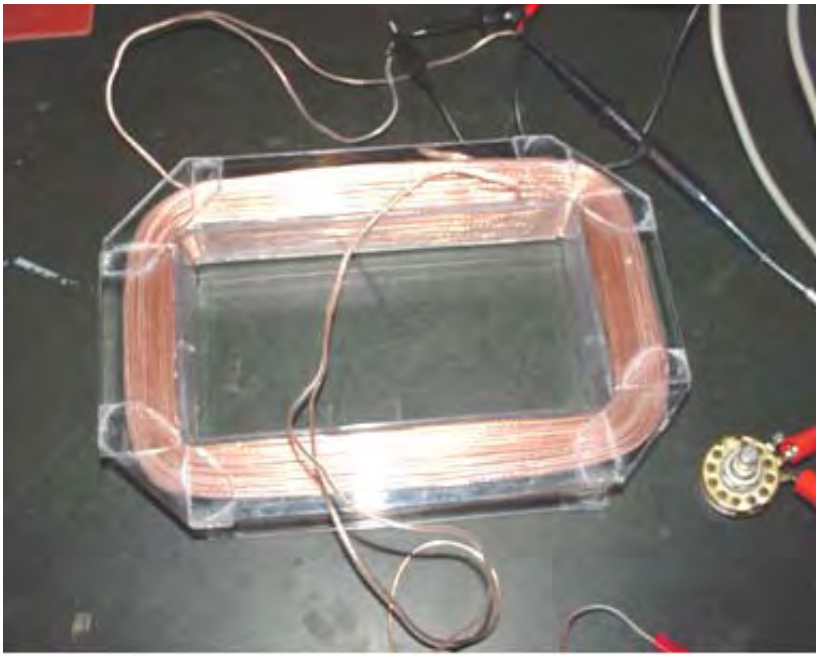
With this procedure we can go on to evaluate other kinds of FE coupled inductor systems:



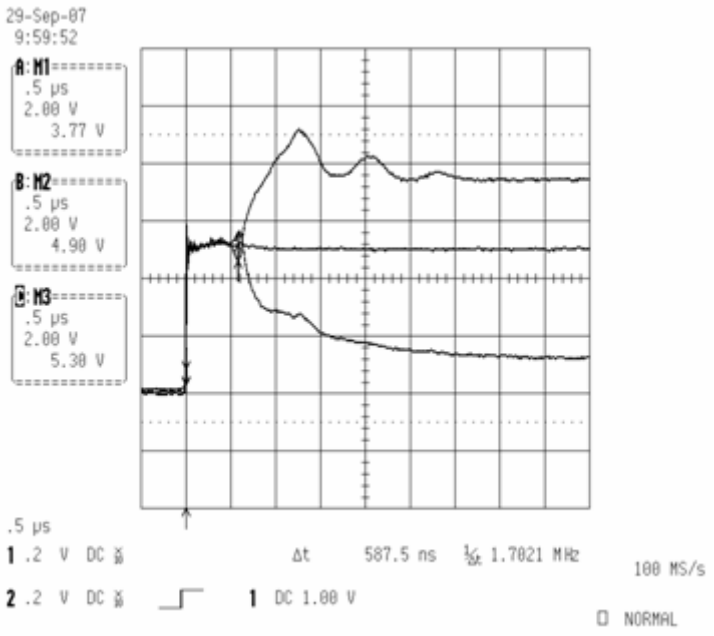
The Trifilar Lindemann Coil – 1000 Turns



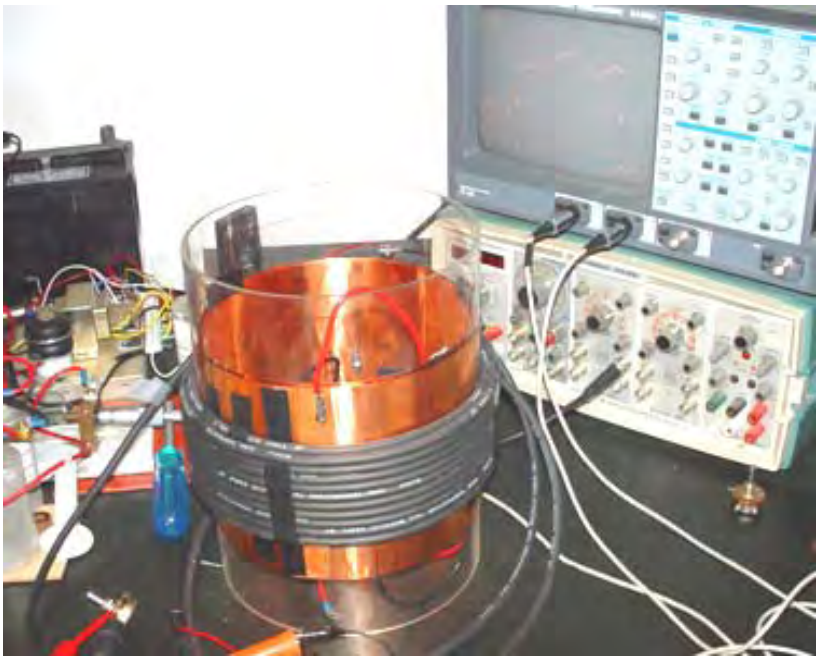
Z_o = 108 Ohms T_d of 885 nS.



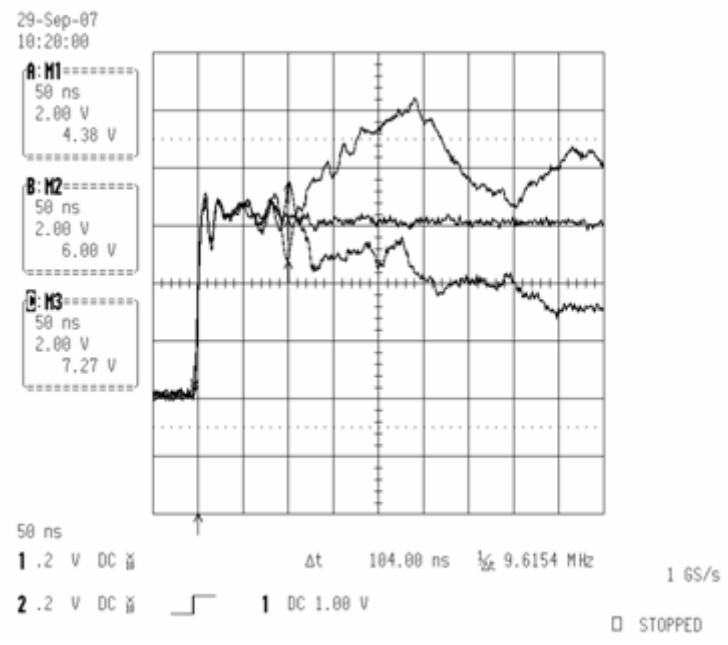
The Mike Motor Coil – 100' #22 Speaker Wire



Z_o = 112 Ohms T_d of 293 nS.



50 KV 8” Prototype Cole FFF



Z_o = 180 Ohms T_d of 52 nS.

Mike Brady's "Perendev" Magnet Motor

Patent Application WO 2006/045333 A1 4th May 2006 Inventor Mike Brady

PERMANENT MAGNET MACHINE

ABSTRACT

The invention provides a magnetic repellent motor which comprises: a shaft (**26**) which can rotate around it's longitudinal axis, a first set (**16**) of magnets (**14**) arranged around the shaft (**26**) in a rotor (**10**) for rotation with the shaft, and a second set (**42**) of magnets (**40**) arranged in a stator (**32**) surrounding the rotor. The second set of magnets interacts with the first set of magnets, and the magnets of both sets are at least partially screened so as to concentrate their magnetic field strength in the direction of the gap between the rotor (**10**) and the stator (**32**).

BACKGROUND

This invention relates to a magnetic repellent motor, or drive mechanism. Such a mechanism may be useful for driving an electrical generator, a vehicle, a ship, an aircraft, or the like.

Conventional power sources rely on fossil fuels or secondary power sources such as nuclear power, or electricity derived by whatever means, for its source of driving power. All of these sources of power suffer from disadvantages such as being the cause of pollution, requiring transportation or transmission over long distances to the point of use, and being costly to purchase. Thus, there is a need for a power source which is substantially pollution-free in operation, requiring substantially no external power, and which is simple to maintain.

SUMMARY

This invention provides a magnetic repellent motor which comprises: a shaft which can rotate about its longitudinal axis, a first set of magnets which are arranged around the shaft and which rotate with the shaft, and a second set of magnets arranged in a stator surrounding the rotor, where the second set of magnets reacts with the first set of magnets, both sets being partially screen magnetically in order to direct their magnetic field into a gap between the two sets of magnets. Thus, the interaction of at least some of the magnets of the first and second sets urge the shaft to rotate.

The interaction may be the net force of like magnetic poles repelling each other thereby urging the magnets away from each other, however, since only the rotor magnets can be moved by this urging force, the shaft is urged to rotate into a position where the repelling force is less.

The rotor may be substantially disc-shaped and the first set of magnets may be located in a peripheral region of the rotor which rotates with the shaft. The stator may be in the form of a pair of arms aligned with the rotor. These stator arms can be moved relative to each other and away from the rotor, in order to allow the gap between the rotor and the stator to be set selectively. The gap may be set manually, for example, by a hand wheel, or automatically, for example by a system of weights which move centrifugally and so form a rotational speed control which acts automatically, i.e. the smaller the gap, the greater the repulsion forces between the magnets of the rotor and stator.

Both the rotor and the stator may have more than one set of magnets. The magnets may be placed in sockets which extend towards the circumference of the rotor. These sockets may be substantially cylindrical and arranged in a plane which is perpendicular to the longitudinal axis of the rotor shaft. These sockets may also be arranged at an acute angle relative to the tangent to the circumference of the rotor disc where the mouth of the cylindrical socket is located. Similarly, the stator magnet sockets may be angled relative to the inner circumference of the stator. These angles may be between 18 degrees and 40 degrees, but preferably between 30 degrees and 35 degrees.

These sockets may have a socket lining consisting at least partially of a magnetic screening material. The socket lining may line the entire extent of the sockets so that only the opening to the exterior remains unlined. In another embodiment of the invention, the magnetic screen lining may cover a substantial percentage of the whole of the socket lining, e.g. 50% of the socket lining.

The magnets may be Nd-Fe-B of dimensions which fit snugly inside the linings of the sockets. These magnets may be cylindrical in shape and have a 37 mm diameter, a 75 mm length and a magnetic strength of 360,000 gauss. The socket lining, magnetic shield and magnet may all have a hole through them to receive a securing pin, preferably positioned so that it is parallel to the longitudinal axis of the shaft.

The number of sockets in the rotor and the corresponding stator may differ so that there is not a one-to-one relationship between the sockets in the rotor and the sockets in the corresponding stator. Similarly, the number of magnets in any additional rotor/stator sets may differ from the first rotor/stator sets in order that the two sets are out of register at any given time. Some sockets may be left empty in either the rotor or the corresponding stator, or both. The motor may have one or more rotor/stator pairs of this type arranged in a stack. It is preferable for the magnets of adjacent rotors to be out of register, i.e. staggered or offset relative to each other.

DESCRIPTION OF THE DRAWINGS

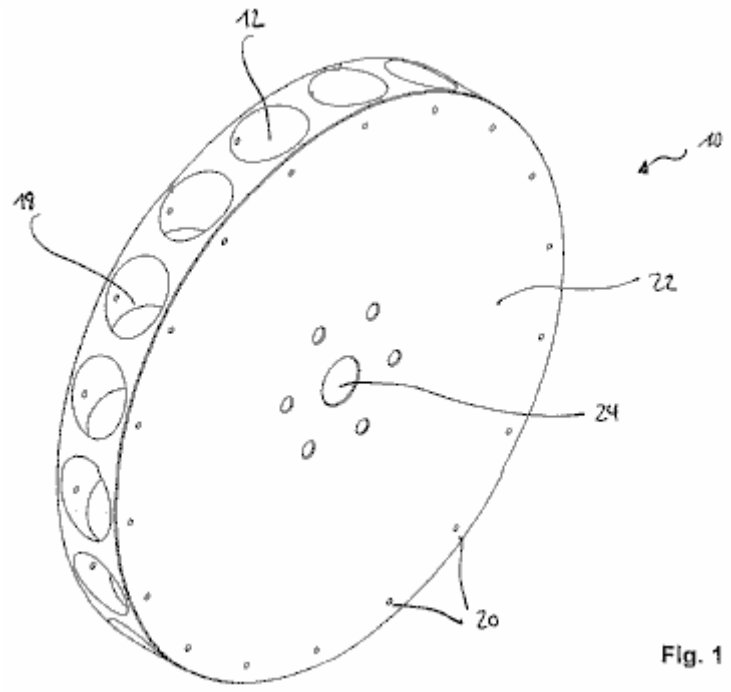


Fig. 1

Fig.1 is a perspective view which shows one rotor disc.

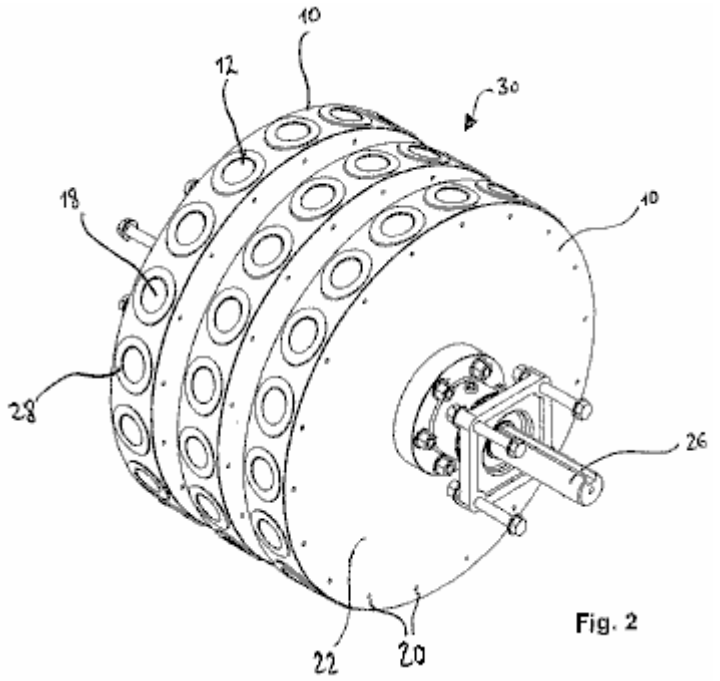


Fig. 2

Fig.2 is a perspective view showing a stack of the Fig.1 rotors in an assembled arrangement.

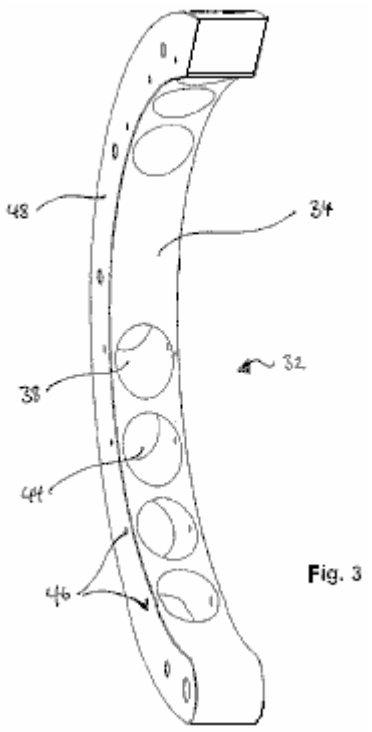


Fig. 3

Fig.3 is a perspective view showing a left arm of a stator.

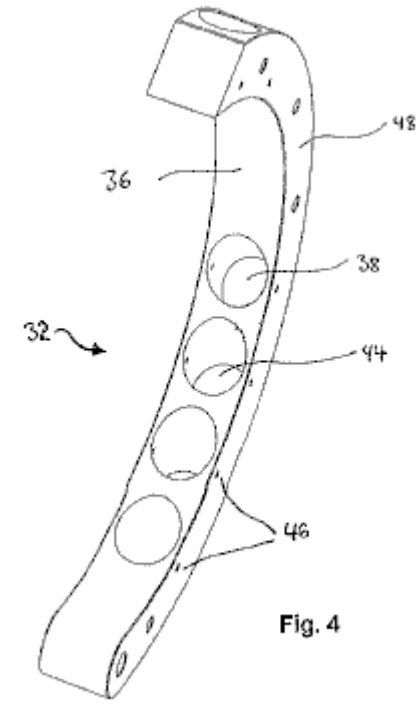


Fig. 4

Fig.4 is a perspective view showing a right arm of a stator

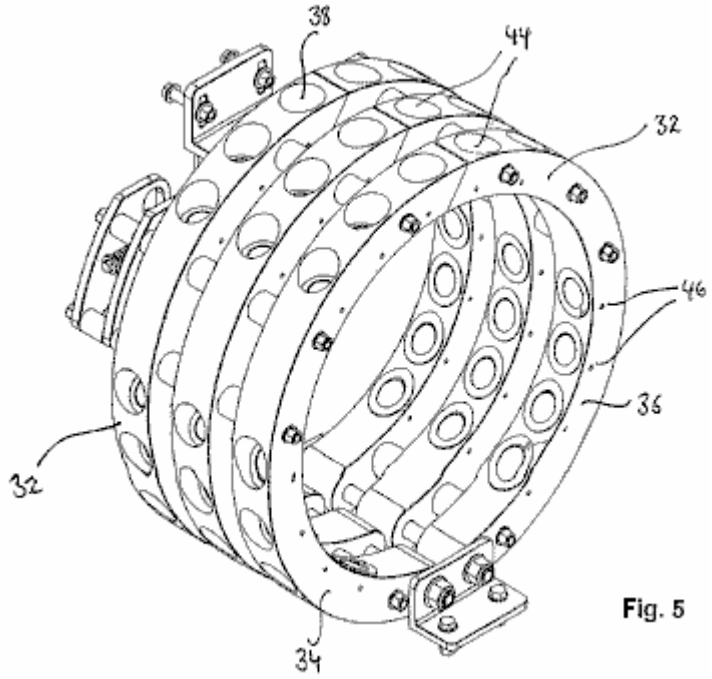


Fig. 5

Fig.5 is a perspective view showing a stack of the stators or Fig.3 and Fig.4 in an assembled arrangement.

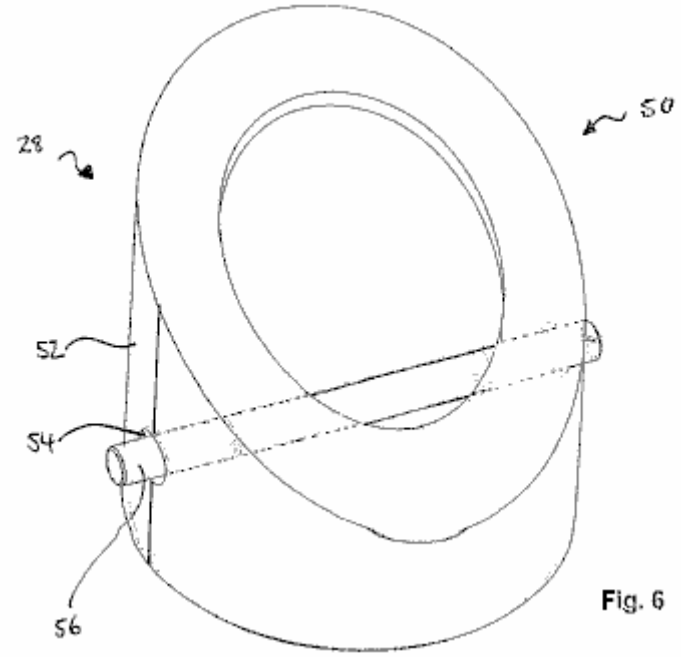


Fig. 6

Fig.6 is a perspective view showing a socket lining of a stator or a rotor.

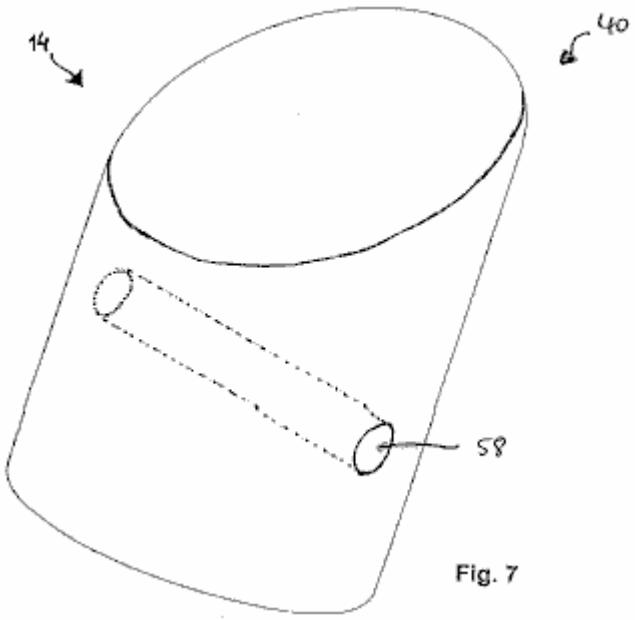


Fig. 7

Fig.7 is a perspective view showing one of the magnets.

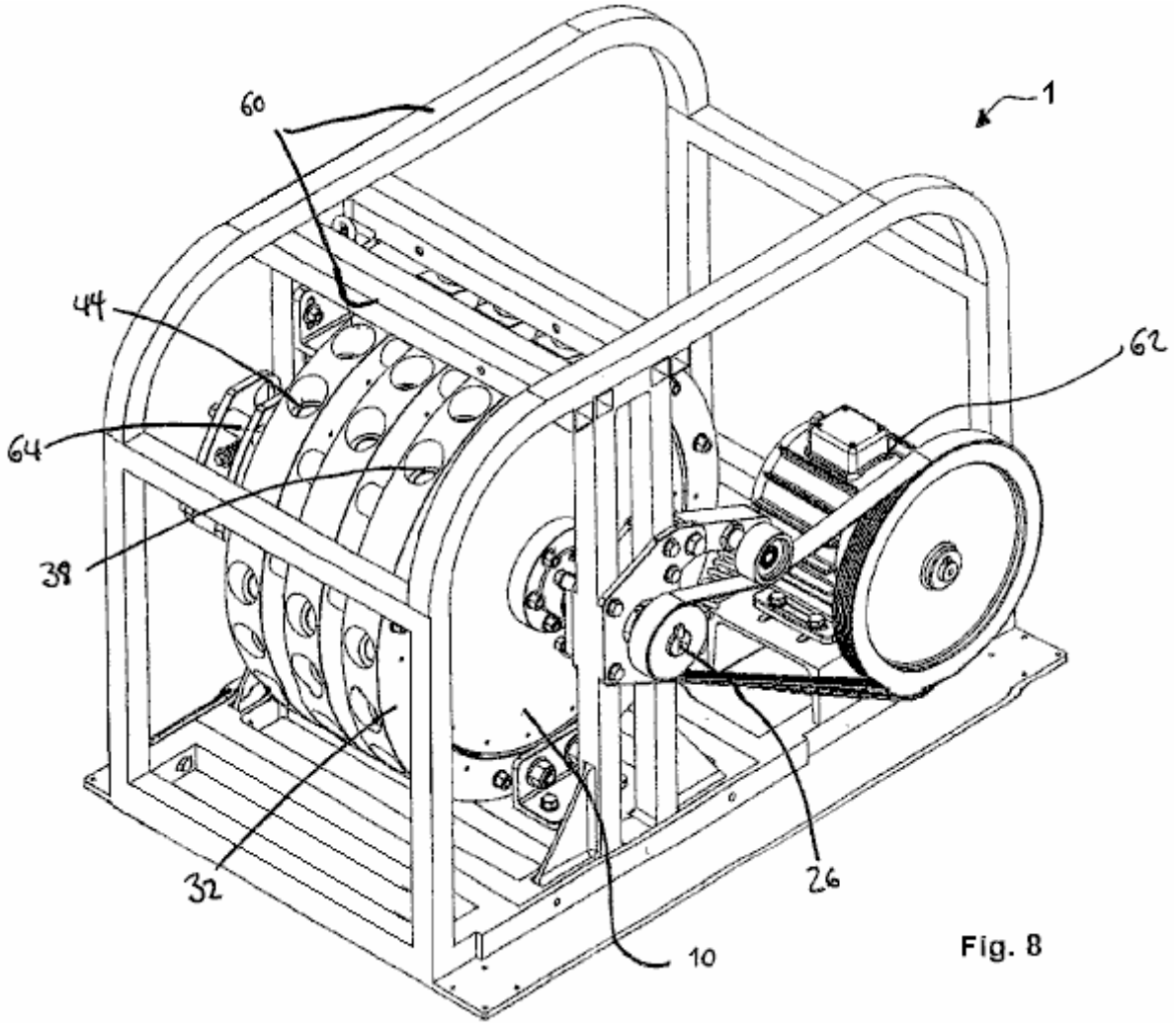


Fig. 8

Fig.8 is a perspective view showing one embodiment of the magnetic repellent motor coupled to an electrical generator.

DESCRIPTION OF PREFERRED EMBODIMENTS

Referring to **Fig.1**, a substantially disc-shaped rotor **10**, is made from a non-magnetic material. The rotor **10** has a plurality of magnet receiving zones **12**, provided in it for receiving magnets **28** (shown in later figures)

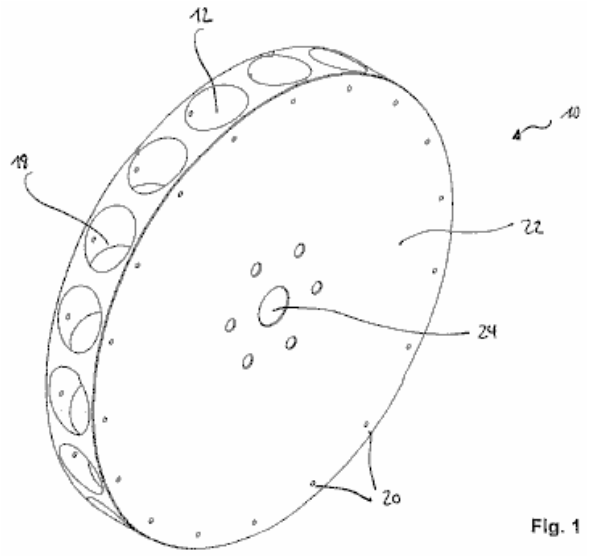


Fig. 1

of a first set **16** of magnets. The receiving zones **12** are in the form of circumferentially extending, spaced apart, and substantially cylindrical sockets **18** which are located in a plane which is perpendicular to the rotational axis **10** of the rotor and in a peripheral region of the disc.

In the region of the sockets **18**, the rotor **10** also has through holes **20** in it's side surfaces **22**, extending parallel to the rotational axis of the rotor. The rotor **10**, also has a centre hole **24**, to receive shaft **28** which is shown in later figures. The sockets **18**, are preferably angled at an acute angle relative to the tangent to the circumference of the rotor disc **10**, at the mouth opening of the sockets **18**. Ideally, this angle is between 18 and 40 degrees, and preferably between 30 and 35 degrees. In one particularly preferred embodiment, the angle is 34 degrees.

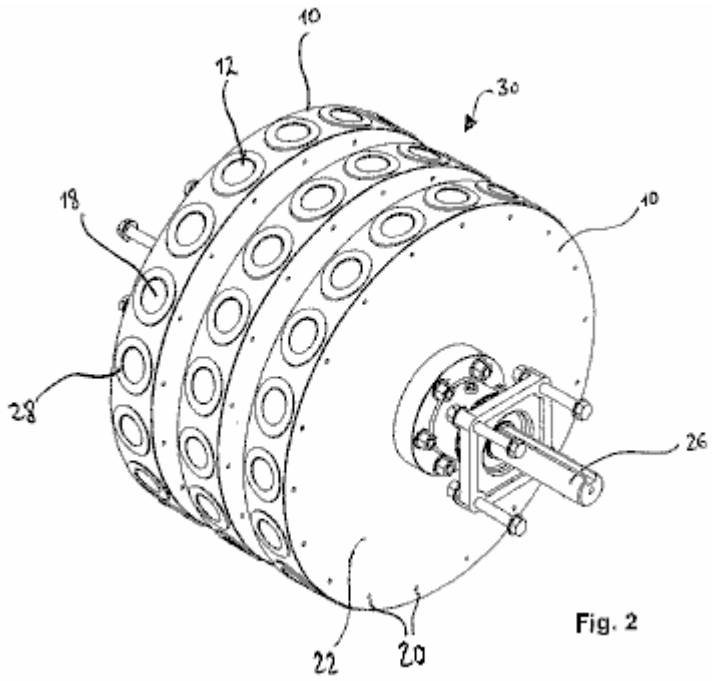


Fig. 2

As shown in **Fig.2**, the sockets **18**, receive (or incorporate) a socket lining **28** (shown in more detail in later figures) which is at least partially made of a magnetic screening material, whether metallic or non-metallic, for example, graphite. The socket lining **28**, covers the entire extent of the sockets **18**, so that only the opening to the exterior remains uncovered.

In the rotor assembly **30** of **Fig.2**, three rotors discs **10**, have been stacked in a row on the shaft **26**. The connection between the rotor discs **10** and shaft **26**, as well as between the rotor discs themselves, can be established via linking means which are widely known. In general, the motor may have any number of rotor discs **10**, and corresponding stators **32**, since the effect of using several rotor discs **10** in parallel, is cumulative. However, it may be useful for smooth operation of the motor **1**, to arrange the rotor discs **10** so that the magnets of adjacent rotor discs are staggered, or offset relative to each other.

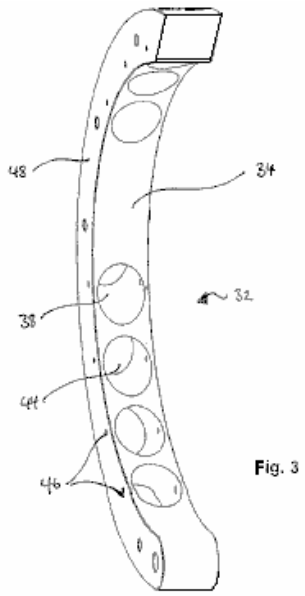


Fig. 3

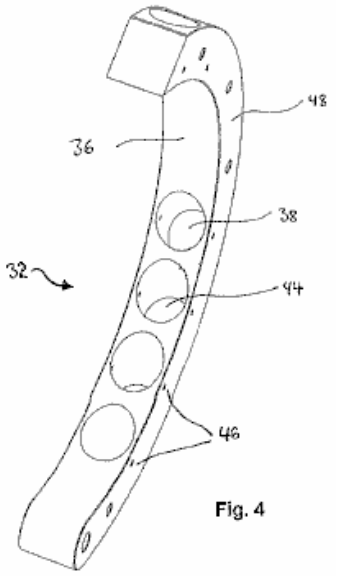


Fig. 4

Referring to **Fig.3** and **Fig.4**, a stator **32** is shown. This stator is made of a non-magnetic material. The left arm **34**, and the right arm **36**, combine to form the stator **32**. Each of the arms, **34** and **36**, has a substantially semi-circular shape and is sized so as to enclose the corresponding rotor disc **10** in the radial direction, while still leaving a gap between the stator **32** and the rotor disc **10**. The arms **34** and **36** of one stator **32**, can be moved relative to each other and their corresponding rotor disc **10**, so that the gap between the arms and the rotor disc can be set at different values.

The stator **32** has several magnet receiving zones **38**, ready to accept the magnets **40**, (which are shown in a later figure) of the magnet set **42**. These receiving zones are again in the form of circumferentially extending, substantially cylindrical sockets **44** which are positioned in a plane which is perpendicular to the longitudinal axis of shaft **26**. In the region of the sockets **44**, the stator **32** has through holes 46 arranged in it's side surfaces **48**, these holes extending parallel to the longitudinal axis of the shaft **26**.

These sockets **44** are again angled at an acute angle relative to a tangent to the inner circumference of the stator **32** at the mouth opening of the sockets **44**. This angle is preferably between 18 and 40 degrees and more preferably, between 30 and 35 degrees. The angle of the sockets **18** and **44**, and the relative positioning between them, has to be adjusted to allow for a good performance of the motor.

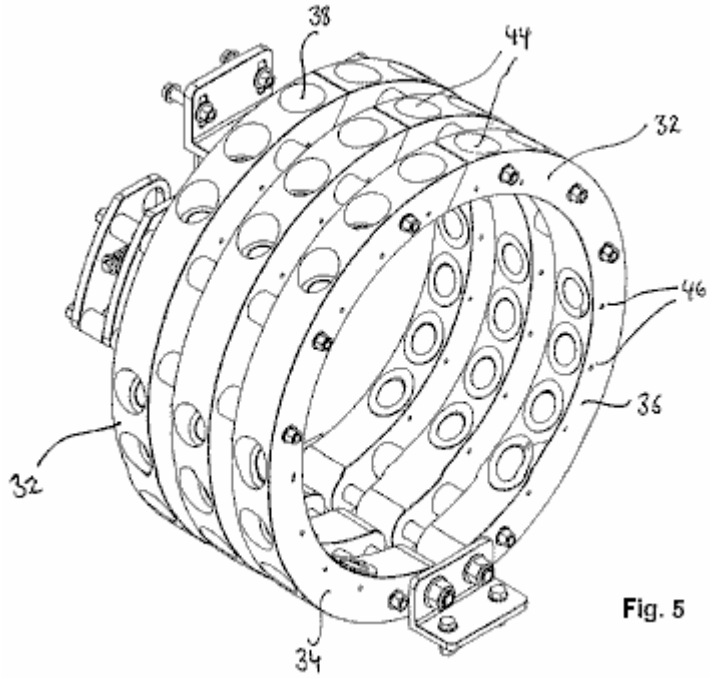


Fig. 5

Fig.5 shows a stator assembly consisting of three stators designed to fit the rotor assembly of **Fig.2**. As described with reference to the sockets **18** of **Fig.2**, the sockets **44** receive (or incorporate) a socket lining **50** (shown in more detail in later figures), which is at least partially made of a magnetic screening material. The socket lining **50**, covers the entire extent of the sockets **44** so that only the opening to the exterior remains uncovered.

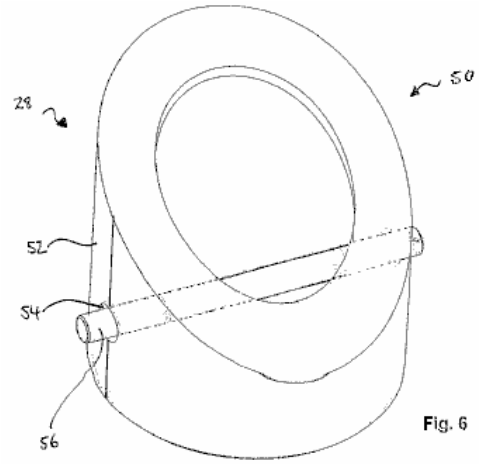


Fig. 6

Referring to **Fig.6**, a socket lining **28, 50** of the rotor disc **10**, or the stator **32**, is shown in more detail. The socket lining **28, 50** is formed to fit into the sockets **18, 44** and may be made completely of a material which has magnetic screening properties. In one preferred embodiment, the socket lining **28, 50** is made of diamagnetic graphite and is partially surrounded by an additional shield **52** of a material having strong magnetic screening properties, e.g. stainless steel. In the embodiment shown in **Fig.6**, the shield **52** surrounds about 50% of the socket lining surface.

Thus, by at least partially covering the sockets **18, 44** with a magnetic screening material, the magnetic field of the inserted magnets **14, 40** is, so to say, focussed axially with the socket **18, 44**, rather than dissipated about the magnets.

Further, holes **54** through the socket linings **28, 50** are provided and these correspond to the through-holes **20** and **46** in the rotor disc **10** and the stator **32**, respectively. Thus, a retaining pin **56** may be inserted after magnet **14, 40** has been put in socket **18, 44** to make a detachable fixing for magnet **14, 40** to the socket lining **28, 50** and the socket **18, 44** so as to prevent expulsion of the magnetic sources during operation.

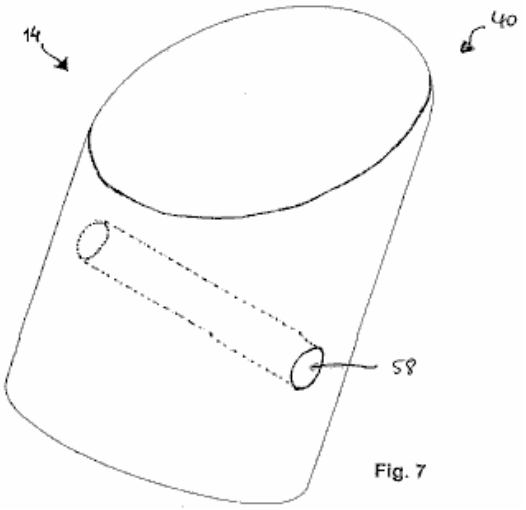


Fig. 7

Fig.7 shows a typical magnetic source **14,40** used in this motor design. The magnetic sources **18, 40** may be natural magnets, induced magnets or electromagnets. The magnetic source for example, is a Nd-fe-B magnet which has the necessary dimensions needed to fit neatly into socket **18, 44** and socket lining **28, 50**, respectively. In one preferred embodiment, the magnetic source **18, 44** is a substantially cylindrically shaped magnet with a diameter of 37 mm, a length of 75 mm and provides 360,000 gauss. However, the magnetic source **18, 44** may be shaped differently to cylindrical and may have different characteristics. In any case, the magnetic source **18, 44** must have a through-hole **58** to receive the retaining pin 56.

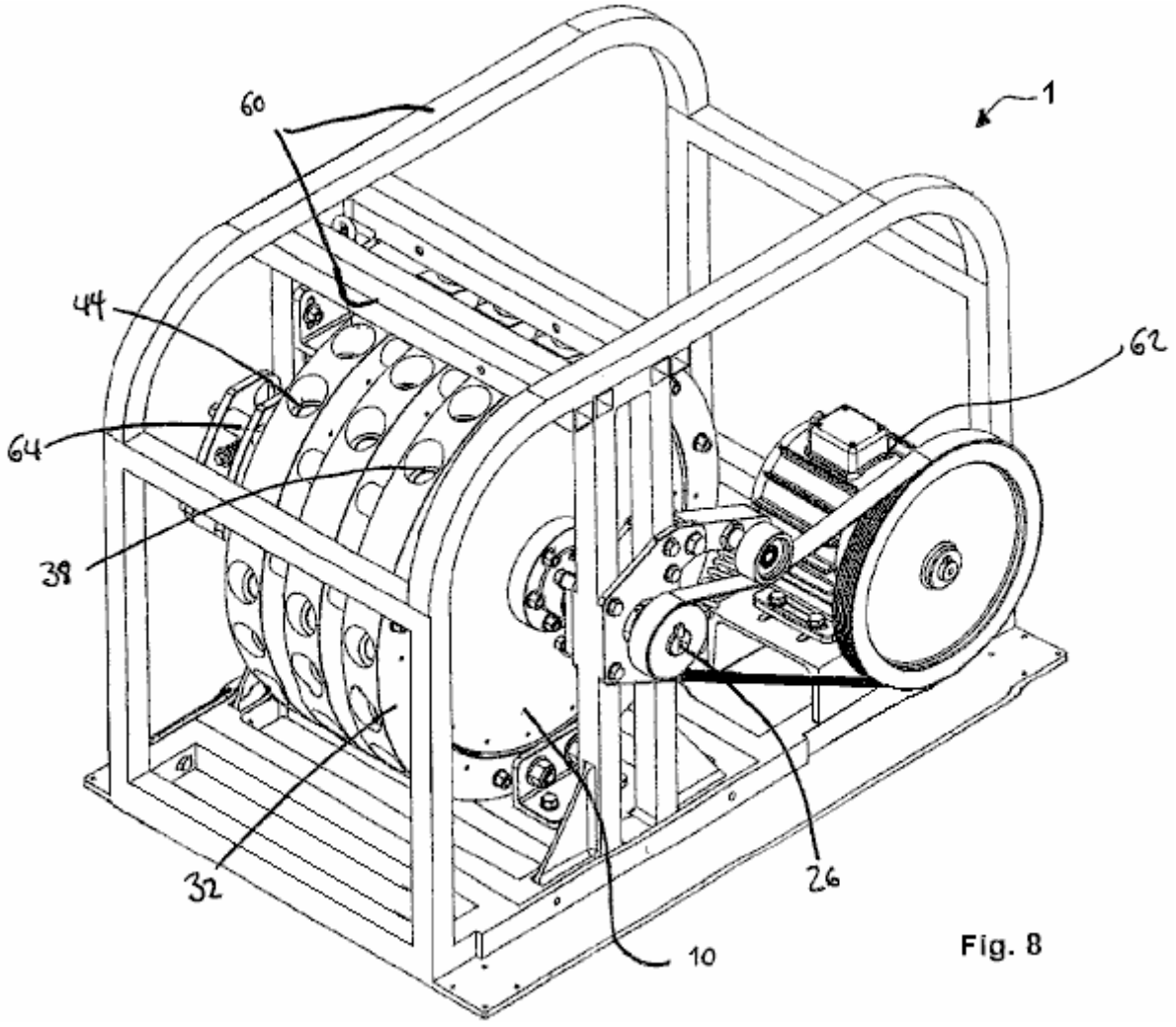


Fig. 8

The magnet motor shown in **Fig.8** is mounted on frame **60** and is coupled to an electrical generator **62**. In this specific embodiment, the motor has three rotor discs **10** of the type already described. These are mounted on a single rotating shaft **26** and are driven by three stators **32**, as already described, causing shaft **26** to rotate about it's longitudinal axis. Shaft **26** may be connected to a gearbox in order to gain a mechanical advantage. The stator arms can be moved by a stepper motor **64**.

The number of sockets in the rotor discs **10** and their corresponding stators **32** may differ so that there is not a one-to-one relationship between the sockets **18** in the rotor disc **10** and sockets **44** in the corresponding stator **32**. Similarly, the number of magnetic sources in the stator **32** and the rotor disc **10** may differ so that a proportion of the magnetic sources **14, 40** are out of register at any given time. Some sockets may be empty, i.e. without a magnetic source, in either the rotor disc **10** or the stator **32**, or both.

The sockets **18** of the rotor discs **10** can be staggered, i.e. offset relative to the sockets of adjacent rotors, or they can line up in register. Thus, the magnet motor may be time-tuned by the relative positioning of the magnetic sources **14** of adjacent rotor discs **10**.

Thus, the interaction of at least some of the magnetic sources **14, 40** of the first and second set **16, 42** urges the shaft **26** to rotate. Once the shaft begins to rotate, the plurality of simultaneous interactions causes shaft **26** to continue rotating.

As mentioned before, the motor can have any number rotor discs **10** and corresponding stator sets **32**. Although the precise adjustment of the motor elements is important, one may imagine other embodiments covered by this invention.

The Magnet Motor of Donald A. Kelly

Patent US 4,179,633 **18th December 1979** **Inventor: Donald A. Kelly**

MAGNETIC DISC DRIVE

ABSTRACT

This permanent magnet disc drive consists of two basic magnetic components, one large driven flat disc containing a uniform series of identical magnet segments, and a second magnetic driving means comprising multiple oscillating magnetic pairs of opposite identical magnet segments. The magnetic mechanism simulates the action of a clock escapement mechanism in that the oscillating magnet pairs uniformly oscillate between the disc magnet segments to induce continuous disc rotation. All of the multiple oscillating magnet pairs are oscillated by a motor, or motors, which provide an eccentric movement through a suitable gear reduction unit. The small DC motors are powered by multiple arrays of silicon solar photovoltaic cells at some convenient rooftop location.

US Patent References:

4,082,969	Magnetic torque converter	April, 1978	Kelly	310/103
4,100,441	Magnetic transmission	July, 1978	Landery	310/103

BACKGROUND OF THE INVENTION

At the present time the magnetic disc drive has reached the stage of development where the oscillating magnet pairs will rotate the magnetic segmented disc when the oscillations is done manually. The disc rotation is smooth and continuous when the manual oscillation is uniform and continuous, and the disc speed may be increased as the oscillation rate is increased.

Since the adequate functioning of the magnetic/mechanical-conversion concept has now been proven with a working prototype, a practical and economical self and/or external oscillation means for the oscillating magnetic pairs must now be developed. The magnetic disc drive was originally designed to be self-actuated by means of a multi-lobe cam and push rod arrangement, but this approach has not been proven successful to date.

A disadvantage for the self-actuated type of magnetic disc drive is that the disc is locked-in with a low, fixed speed output which is dependant on the natural magnetic field interaction between the involved interacting magnet segments.

A mid-diameter direct displacement multi-lobe cam was used for the first prototype, but this did not work because of the high rotational resistance imposed by the high cam lobe angles. A peripheral, direct displacement multi-lobe cam was also tried but this was not successful because of the moderate and sufficient cam lobe resistance to push rod displacement.

Other cam lobe configurations are being planned and developed to make sure that no possible trade-off to self-actuated mechanical oscillation is overlooked. Another possible approach to self-actuation for the magnetic disc drive is by the application of a twin level magnetic commutator which is directly connected to the disc drive shaft. The magnetic commutator segments alternately attract corresponding radial magnets on pull-rods which are pivoted on each of the oscillation plates of the magnetic pairs.

While auto-actuation of the magnetic disc units may be desirable for some self-contained power applications, the low, fixed speed output is not considered attractive and promising for a wide range of household power applications. Because of the inflexibility of speed output of the auto-actuated type of unit the, the development of a variable speed, externally oscillated type of disc unit is required to meet the growing demand for alternate and auxiliary power means for many applications.

The matching of a large magnetic disc drive and small solar powered DC electric motors is a nearly ideal arrangement since a single or series of small precision DC motors can be readily powered by modest arrays of silicon photovoltaic cells located at some convenient rooftop location. Small high-efficiency, ball bearing DC motors are available which, when connected to suitable gear reduction drives, can revolve a simple eccentric mechanism with sufficient power and variable speed, to cause oscillation of a series of four to six magnetic oscillating pairs of stator magnets.

This series of magnetic oscillating pairs will all be connected together with straight linkage to transmit the reciprocating motion from the driving oscillating shaft to the other oscillating shafts of the series. This is a more desirable multiple driving arrangement rather than separate small DC motors since synchronism is automatically assured, rather than more complex and less reliable electrical synchronization requirements. Because there is no locked-in synchronism for this type of external oscillation means, the multiple magnetic oscillation pairs must be of the minimum interference type, in that they must not become jammed into the disc magnet segments. Although the proper functioning of the magnetic disc unit requires that the oscillating magnet pairs must enter the disc's magnet segment interference circle, deflection means must be added to all of the oscillation plates to insure that the continuously revolving disc will readily by-pass all of the oscillating magnet pairs.

The large magnetic disc unit will consist of a basic non-magnetic circular disc, on which multiple high energy permanent magnet segments are equally spaced around the rim of the disc. The drive shaft of the disc rotates on precision ball-bearings and may be chosen to revolve in either a horizontal or a vertical plane. The disc is the driven component of the magnetic drive assembly, and it can be connected to the load or an electrical generator.

The multiple oscillating magnet pairs are the driving component of the disc drive unit and consist of flat, non-magnetic oscillation plates, on which identical high-energy permanent magnets are secured at each end of these oscillating plates. The magnet segments are placed with opposite poles exposed at the sides, relative to each other so that a north-south pole couple reacts on the disc's magnet segments. The driven disc's direction of rotation depends on the polarity of the disc's magnets in relation to the oscillating magnetic pairs.

The oscillating magnetic pairs will make a full back and forth oscillation between two adjacent local disc magnet segments so that an alternate "pull and push" effect is induced on the magnetic segmented disc. The basic synchronism between the disc's magnet segments and the multiple oscillating magnet pairs closely simulates the action of a watch or clock escapement mechanism in respect to the natural "cogging" action between the functioning components.

This general magnetic disc drive arrangement insures smooth and continuous rotation for the driven disc with an optimum of magnetic energy interchange between the oscillation stations and the magnetic disc because of near pole face to pole face exposure. It is now believed that this present type of magnetic disc drive is approaching a theoretical maximum of conversion performance possible, especially when compared with other types of magnetic/mechanical arrangements such as magnetic worm and worm discs, spur couples, mitre couples, and all types of inferior, linear magnetic devices.

The attractiveness of the basic magnetic disc and oscillating pairs is that a nearly ideal leverage factor is introduced in magnetic/mechanical conversion arrangements. Simply stated, considerably less energy is needed to oscillate the oscillating pairs than is produced from the near pole face to pole face magnetic interaction between the functioning magnetic components.

The alternating and uniform "pull and push" force imposed by the oscillating magnet pairs on the disc magnet segments produces no direct back or counter force reaction on the driving oscillating magnet segments which is the master key for a useful and practical magnetic/mechanical conversion drive. The back or counter-reacting force on the oscillating magnet pairs is taken directly by the fixed pivots of the oscillation plates, with a minimum of load penalty imposed on the drive of the oscillating magnet pairs.

All other types of rotary magnetic/mechanical conversion devices, with the possible exception of the worm and worm disc type, produce an undesirable back reaction force on the driving component and resulting ineffective performance. The magnetic worm and worm disc units have not proven to be sufficiently worthwhile for commercial applications because of the very high permanent magnetic energy necessary and due to the low speed output of these mechanisms.

When configuration comparisons are made of all types of possible magnetic/mechanical conversion devices it will be noted that the combination of a magnetic disc driven by multiple oscillating magnet pairs will stand out as a practical and useful permanent magnetic conversion arrangement. The incentive for the development of this magnetic disc drive was the direct outgrowth of overall disappointing performance of solar energy conversion efforts and the frustrations encountered with component costs, conversion efficiency and a lack of suitable energy storage means. While solar energy is being widely hailed for its future potential as a viable alternate energy source, relatively few engineers speak out about relatively poor overall cost/effectiveness due to days-on-end of overcast skies during the winter months when the energy is most needed, especially in northern latitudes.

Because of the less-than-adequate solar energy conversion outlook for the vast majority of American homeowners, other alternate, small scale, decentralised, energy sources must be explored and developed on a crash program basis. If this is not done within the next several decades we must accept the alternative of a greatly reduced standard of living because of the alarming rise in the rate of energy costs.

This magnetic disc drive represents a practical solution in applying permanent magnetism in the development and commercialism of a decentralised, silent, fuel-free, household-sized alternate power system. While the power output from an individual magnetic disc unit may be small, the power output is constant and does not generally depend on the intensity of an external energy source, as do present solar energy systems.

SUMMARY OF THE INVENTION

The magnetic disc drive unit is comprised of a large driving disc made of non-magnetic metal on which several permanent magnets are equally spaced around the rim. The disc drive shaft rotates on trunnion supported ball bearings and may revolve in nearly any conventional position, and may be constructed with any practical large diameter.

The identical oscillating magnet pairs are the driving component of the disc drive and consist of flat, non-magnetic plates on which, pairs of identical permanent magnets are secured at both sides of the oscillation plates. These magnet pairs have opposite pole faces facing each other. The disc's direction of rotation is determined by the polarity of all the disc's magnets relative to the polarity of the oscillating magnet pairs.

The oscillating pair of magnets make a full back and forth oscillation while each rotor disc magnet passes by. This produces a pull on the disc magnet as it approaches the oscillator magnet and then when the oscillator moves that magnet away, a push force is applied to the magnet on the rotating disc by the second magnet of the oscillating pair of magnets. The synchronisation of the disc and the oscillating magnet pairs must be maintained for continuous and smooth rotation of the disc. This movement is similar to the action of a clock escapement-mechanism.

The method of moving the oscillating pairs of magnets is one or more solar-powered DC motors. These motors drive push rods which are in contact with ball bearings mounted on the oscillation plates. Since the eccentrics must move at relatively slow speeds, suitable gear reduction units must be used between the motors and the rocker arms.

In order to maintain proper synchronisation of all of the oscillating components, straight links are used to connect all of the driven oscillation shafts to the driving oscillation shaft. Four or five oscillation stations can be driven from one driver oscillation shaft so that a disc drive with a large number of oscillation stations will require several D.C. motors to drive all of the other oscillation shafts.

It is important that the multiple, identical oscillation plates and their magnet pairs be slightly shorter in width than the space between two adjacent disc magnet segments, so that an optimum pull and push force is induced on the local disc magnet segments. One side of the oscillating magnet couple "pulls" on the disc's permanent magnet and then the other oscillator magnet "pushes" the disc's permanent magnet onwards as it has been moved into place by the oscillation.

All of the oscillating magnet pairs oscillate on stationary rods, or shafts, and all of the eccentrics and DC motor drives remain fixed on a base plate. The other ends of the oscillating rods or shafts must be supported by some form of bracket to keep the oscillation plates parallel to the disc magnet segments. Each eccentric which moves a ball bearing attached to arms on the oscillation plates must make one full 360 degree revolution within the angular displacement arc between two adjacent rotor disc magnet segments. Two small pivot brackets are attached to the extreme, non-magnetic ends of the oscillation plates to allow these plates to oscillate freely with a minimum of friction.

The basic rotational relationship between the magnetic oscillating pairs, and the magnetic segmented disc, will have a bearing on the gear reduction ratio required for the gear drive unit coupled to the small DC motors. Fairly rapid oscillation is necessary to maintain a reasonably acceptable disc speed which will be required for most power applications. The size of the eccentrics which oscillate the oscillating magnet pairs will be determined by the full oscillating arc needed and the mechanical advantage required by the oscillation plate in order to cause the optimum rotation of the magnetic disc drive unit.

Proper magnetic disc drive functioning requires the pulling magnets of the oscillating magnet pairs to enter the disc's interference circle within the mutual magnetic field zone between the two local interacting magnets on the disc's rim. Since the disc will revolve continuously, the withdrawing phase of the "pulling" magnets brings the "pushing" magnets of the couple into the disc's interference circle within the mutual magnetic field zone, for effective interaction with the adjacent disc magnet segment.

All of the magnet segments on the oscillation plates which form the magnetic couples must be in line with the corresponding disc magnet segments in order to maintain an optimum interaction between them.

Because there is no natural, lock-in synchronism for this type of magnetic disc drive, the multiple magnetic oscillating magnet pairs must be of the minimum interference type, which consists of adding plastic deflectors to the oscillation plates to prevent the pulling magnets of the couple from jamming into the disc magnet segments. Since the oscillating magnet pairs must never jam into the disc and stop its rotation, the plastic deflectors will allow the oscillation plates and magnet pairs to be deflected away from all of the disc magnet segments.

The permanent magnets selected for both components of the disc drive must be uniformly identical and have the highest possible energy product or magnetic induction plus coercivity. Both of these magnetic properties will play a significant role in determining the true value of the magnetic disc drive unit. At the present time the rare-earth/cobalt permanent magnets offer the highest possible magnetic properties for this application, but their cost is very high and currently not considered cost effective for the magnetic disc drive. Since costs will also play a major role in the competitive value of the disc drive, the magnets selected must show the highest possible cost/effectiveness ratio, along with long operating life.

Rectangular ceramic permanent magnets with large flat pole faces are preferred for the disc drive prototypes, and there is no theoretical limit to the size of both interacting components. A practical limit to the actual size of the components is imposed by weight and material cost restrictions plus available space, but nearly any practical number and size of uniformly identical magnets may be used to make up the magnetic disc drive.

It will be advantageous to build up each disc magnet station into clusters of up to about twelve to twenty four individual magnets which are arranged in lengths of four or five units and double or triple widths depending on the disc diameter. A large diameter disc unit is always desirable since the torque output for the disc unit depends on the tangential magnetic force produced by all of the oscillating magnet couple stations multiplied by the disc radius.

The large diameter disc speed will be relatively slow, in the 20 to 30 r.p.m. range, so that the disc output speed must be stepped up to a useful 750 to 1200 r.p.m. speed range, by a belt drive arrangement. The magnetic disc drive output is best adapted to run an electrical generator or alternator to produce electrical power for various household purposes.

An advantage to using silicon photovoltaic solar cells on an exposed rooftop location as a power source, is that they are capable of providing a partial E.M.F. under non-sunlight/overcast sky conditions. With full sunlight exposure the electrical energy produced will run the magnetic disc drive at its maximum possible speed, with reduced sunlight levels producing a corresponding proportionate reduction in the disc output speed.

A workable option exists for using a greater number of silicon photocells than would be normally necessary for full sunlight operation. The number of cells selected would be capable of running the magnetic disc drive at full speed under overcast sky conditions, with any excess full sunlight current bypassed to storage batteries. This option is a desirable arrangement since the disc will be assured of full electrical input power each day, with battery power available to make up the loss from any dark daytime sky conditions.

The principal object of the invention is to provide the highest torque output for the large driven disc from the lowest possible torque input for the multiple oscillating magnet pairs, as a useful power step-up means for electrical generating applications.

Another object of the invention is to provide a step-up power source which can be produced at competitive costs, requires no combustible fuel and is non-polluting while running silently and continuously.

It is a further object of the invention to provide a natural energy source which has an extremely long operating life, with a maximum of operating effectiveness, component resistance to degradation, with a minimum of parts replacement and maintenance.

The various features of the invention with its basic design geometry will be more apparent from the following description and drawings which illustrate the preferred embodiment. It should be understood that variations may be made in the specific components, without departing from the spirit and scope of the invention as described and illustrated.

Referring to the Drawings:

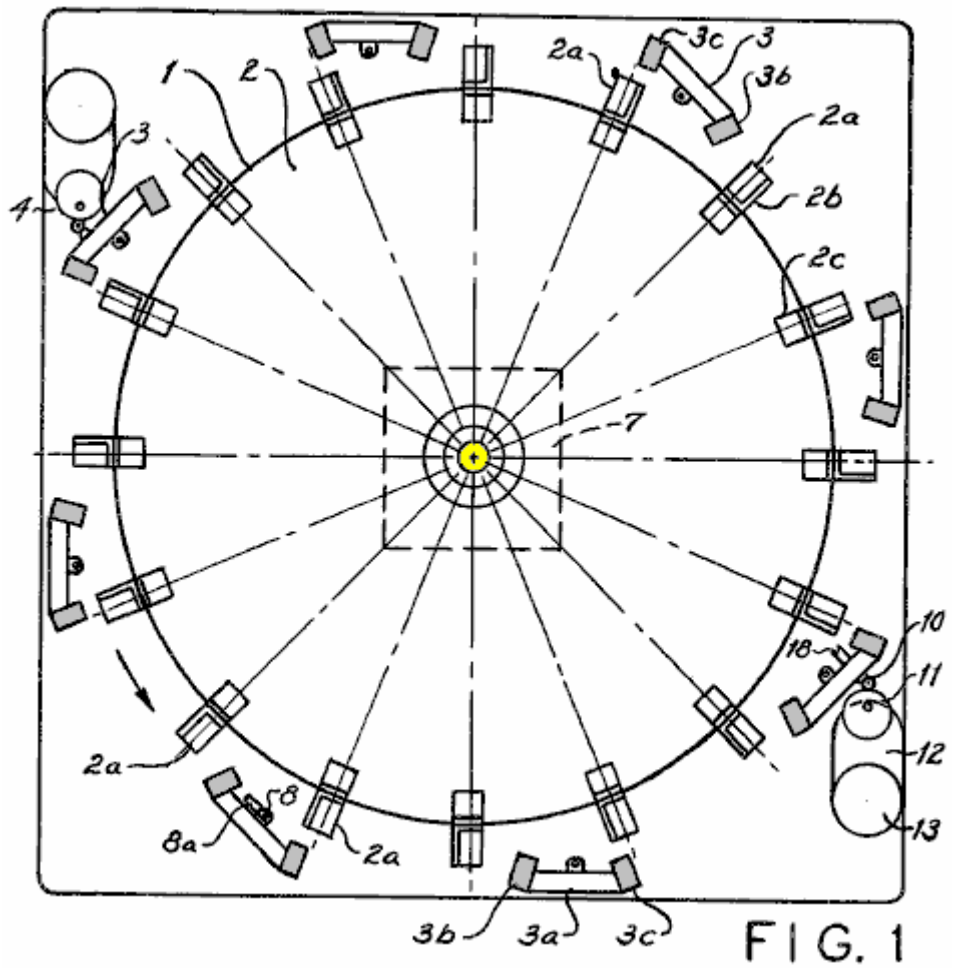


FIG. 1

Fig.1 is a top, external view of the magnetic disc drive.

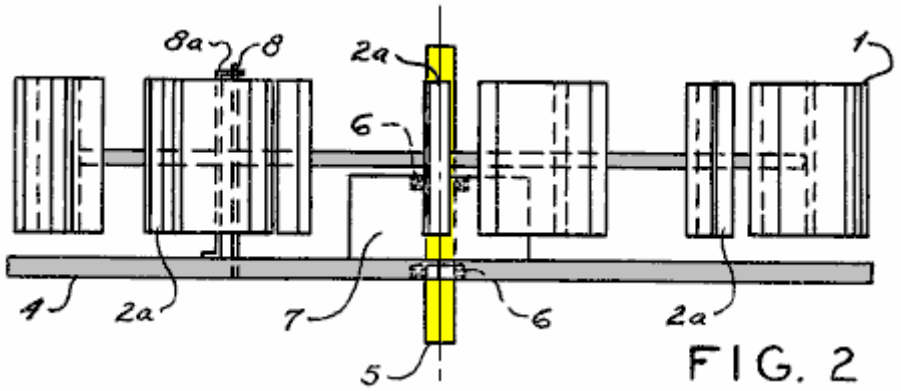


FIG. 2

Fig.2 is an external side view of the magnetic disc drive.

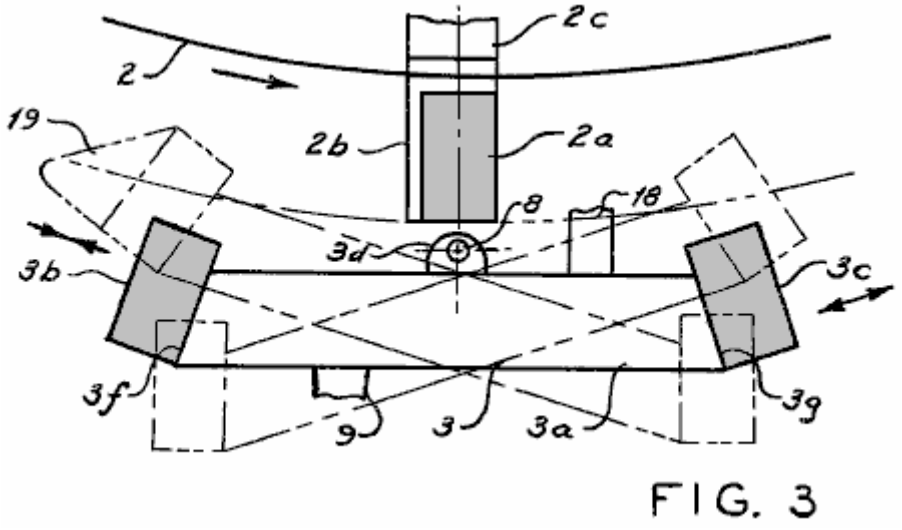


FIG. 3

Fig.3 is an enlarged top view of one oscillating magnet couple.

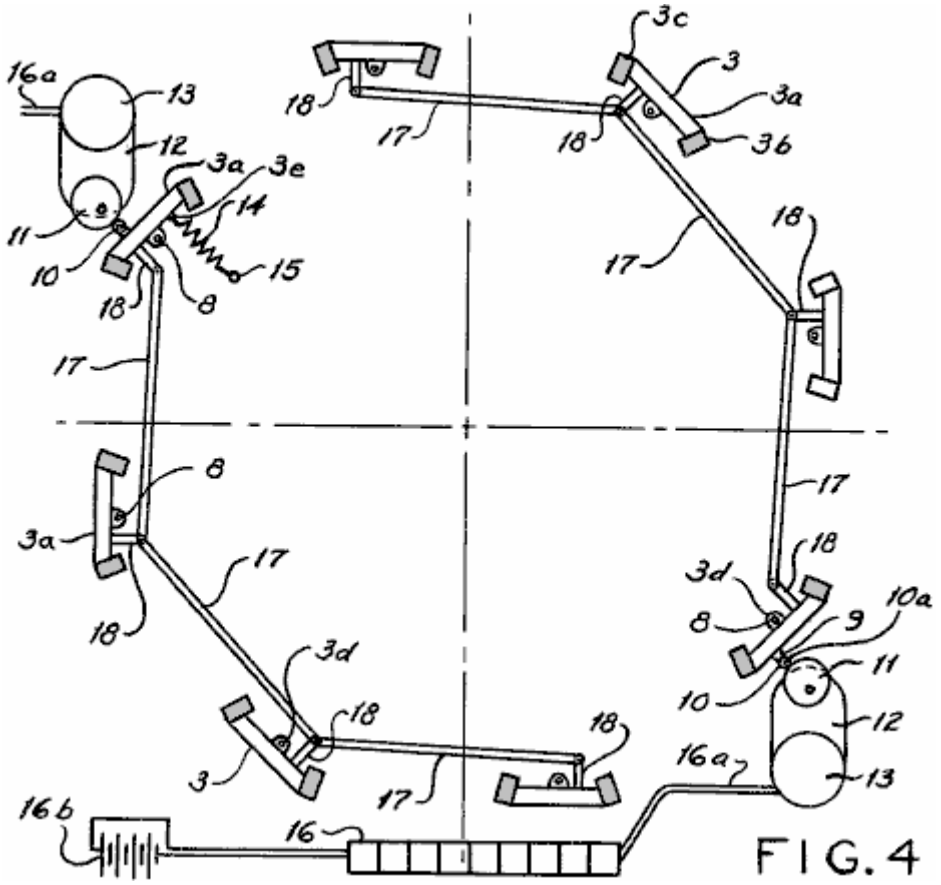


FIG. 4

Fig.4 is a top, break-away view of several oscillating magnet pairs connected together with linkage.

DESCRIPTION OF THE PREFERRED EMBODIMENT

The invention **1**, is comprised of two basic components: a large driven disc **2**, and multiple oscillating magnet pairs **3**, which are closely interrelated and mounted on a common base plate **4**.

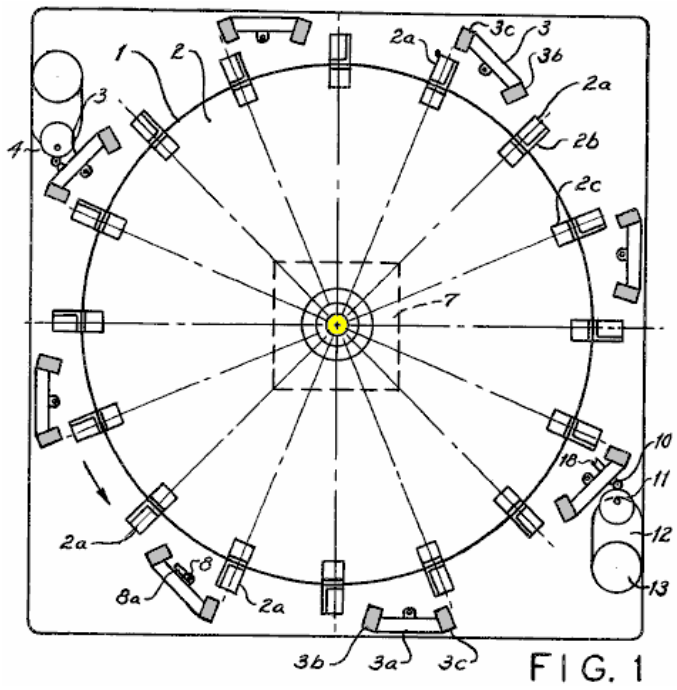


FIG. 1

Multiple, identical permanent magnets **2a**, are equally spaced around the periphery of the large driven disc **2**, by means of support angles **2b**, and angle brackets **2c**, which are secured to the disc **2**, with standard hardware.

A drive shaft **5**, is fastened to the disc **2**, by means of a hub **2d**, and supported by two ball bearings **6**. One of the ball bearings **6**, is fitted into a bore within the base plate **4**, while the other ball bearing **6**, is fitted into a box-base **7**, which is fastened to the base plate **4**, with standard hardware.

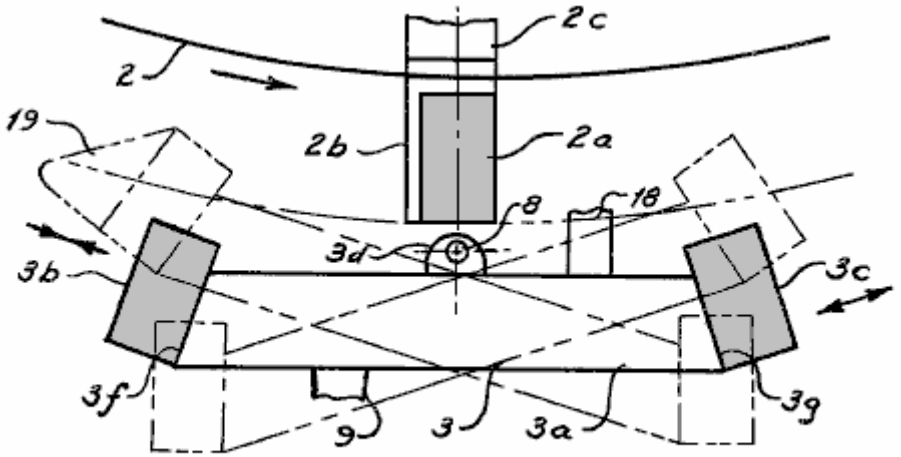


FIG. 3

The multiple oscillating magnet pairs **3**, are a flat, non-magnetic plate **3a**, with opposite pole magnet segments **3b** and **3c**, respectively, attached to the side of the flat oscillation plate **3a**. Two pivot brackets **3d**, are attached to the top and bottom of the flat plate **3a**, which pivot the oscillation plate **3a**, on the pivot rod **8**. One end of the pivot rod **8**, is fitted into the base plate **4**, and the opposite end is supported by an elongated Z-shaped bracket **8a**.

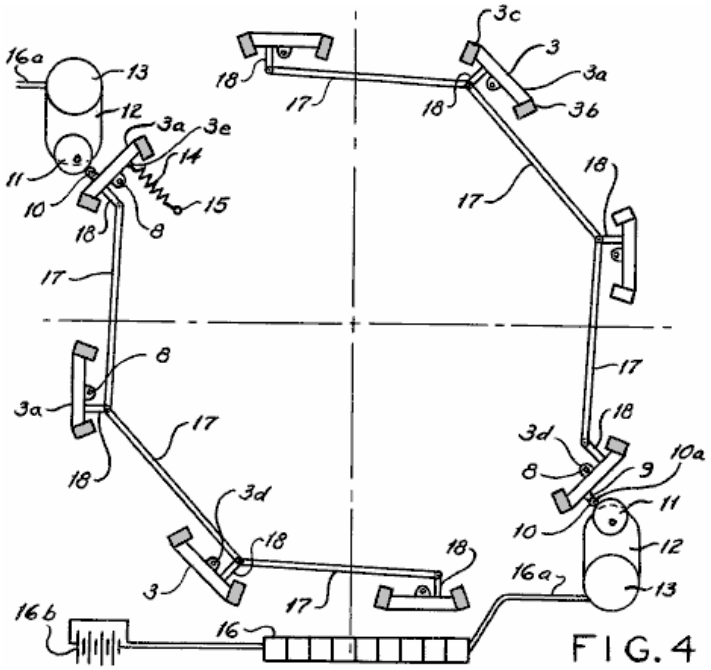


FIG. 4

An arm **9**, is fastened to a flat face of the flat plate **3a**, which supports the pin **10a**, which carries the ball bearing **10**, as it rolls on the eccentric disc **11**. The off-centre disc **11**, is fastened to the slow speed shaft of the gear reduction unit **12**, which is driven by the small DC motor **13**. A return tension spring **14**, is connected to the oscillation plate **3a**, by eyelet **3e**. The opposite end of the return tension spring **14**, is retained by the post **15**, which is pressed into the base plate **4**. Motors **13**, are powered by multiple arrays of silicon photovoltaic solar cells **16**. Electrical leads **16a**, conduct solar converted electricity to the motors **13**, with any excess current stored in the batteries **16b**.

The motor driven oscillation stations become the master stations for this invention **1**, from which three to five slave oscillation stations are driven. The reciprocating motion is transmitted by straight links **17**, which are pinned to the link arms **18**, which in turn are secured to the flat plates **3a**.

All of the slave oscillation stations must be precisely adjusted to exactly the same angular position as the master driving oscillation station so that all stations are synchronised to allow proper functioning of the rotating disc **2**.

For very large discs **2**, with many disc magnets, several master oscillation stations, with a fixed number of slave oscillation stations will be required. All of the master oscillation driving-stations will have to be electrically synchronised to maintain overall synchronisation, with all of the eccentrics **11**, set at the same angle at start-up of the disc.

Either end of the drive shaft **5**, may be connected with a speed step-up belt drive arrangement, which is not shown here.

Plastic deflectors **19**, are added to either side of the oscillation plates **3a**, adjacent to the opposite magnets segments **3b**, and **3c**, their exact position depending on the direction of rotation of disc **2**. These act as an anti-jamming device for the magnets.

Magnetic field bias angles **3f** and **3g** (**Fig.3**), are required for the sides of plates **3a**, in order to assure an optimum "pull-push" sequence on the large drive disc **2**, as the magnetic oscillation pairs **3**, are actuated. The bias angle **3f**, is matched to the magnet segment **3b**, while bias angle **3g** is matched to magnet segment **3c**.

None of the load components which are external to the device, such as an electric generator or alternator, are shown as a part of this invention, since a variety of load devices and arrangements are possible for the magnetic disc drive.

Bob Neal's Compressed Air Engine

US Patent 2,030,759 **11th Feb. 1936** **Inventor: Bob Neal**

COMPRESSOR UNIT

This invention relates to the construction of a compressor, and more particularly to a combined fluid-operated engine and compressor.

The primary object of the invention, is the provision of a compressor of this character, wherein there is arranged an automatically counterbalanced crankshaft and fluid equalisers within a storage tank, which makes it possible for the engine to operate on constant reserve tank pressure, so as to actuate additional equipment, the pistons for the engine also being automatically balanced and suspended when the engine is operating.

Another object of the invention is the provision of an engine which is operated by air under pressure, the air being supplied by compressors which are in a bank with the engine construction.

A further object of this invention is the provision of an engine of this type of novel construction as the engine and the compressors are operated from the same crankshaft, which is of the automatically balanced type, so that high efficiency is attained.

A still further object of the invention is the provision of an engine of this character which is comparatively simple in construction, thoroughly reliable and efficient in its operation, strong, durable, and inexpensive to manufacture.

With these and other objects in view, the invention consists in the features of construction, combination and arrangement of parts as will be described more fully here, illustrated in the accompanying drawings which disclose the preferred embodiment of the invention, and pointed out in the appended Claim.

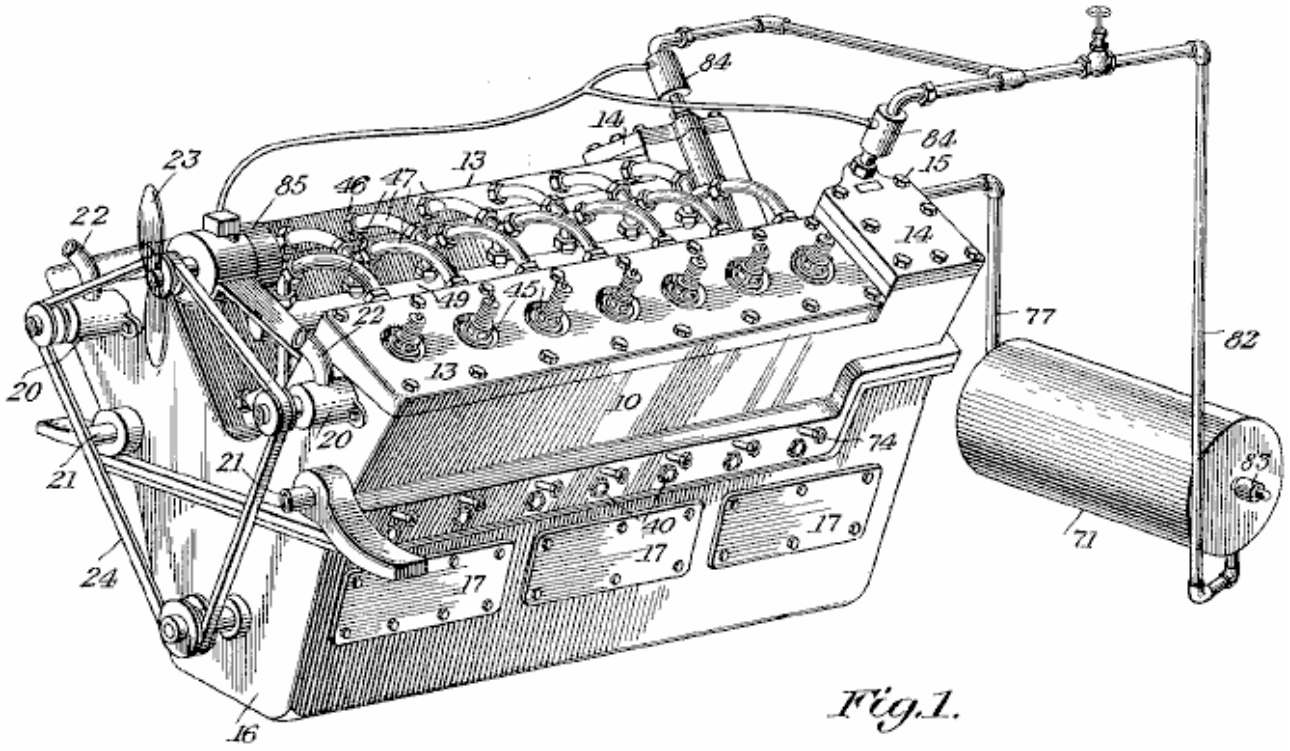


Fig.1.

In the drawings, **Fig.1** is a perspective view of the engine constructed in accordance with the invention.

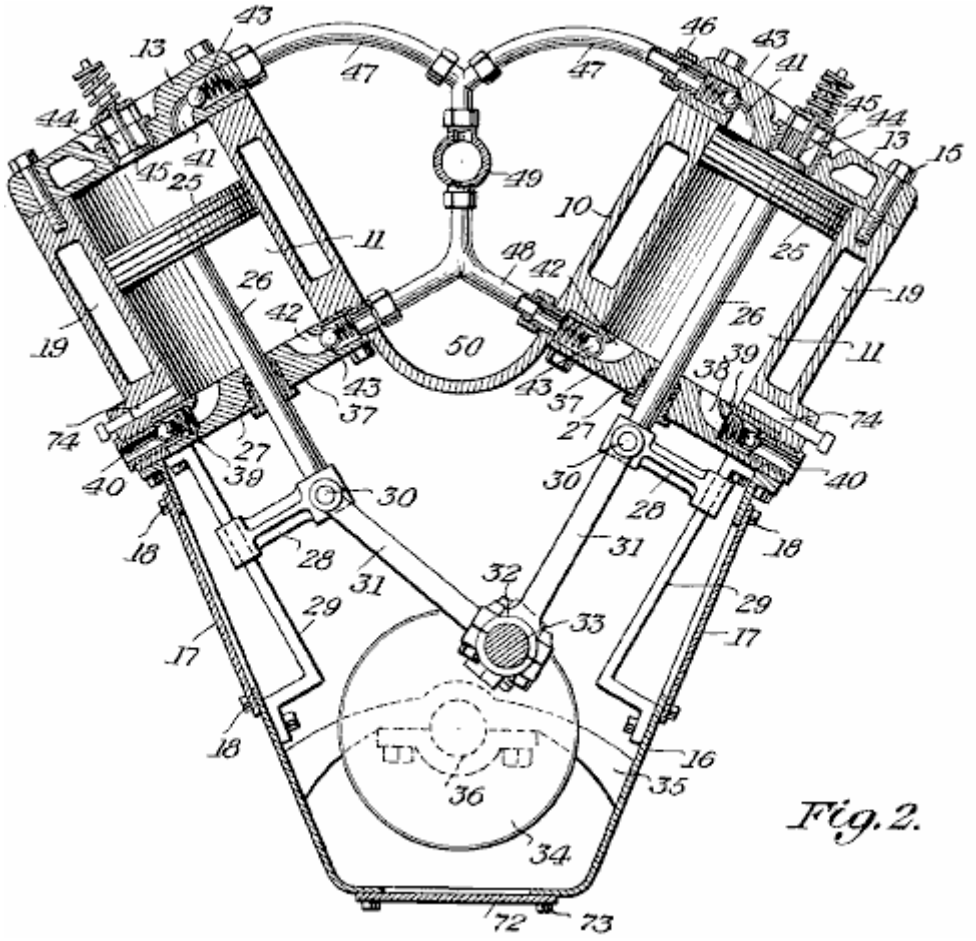


Fig. 2.

Fig.2 is a vertical transverse cross-section view through the compressor part of the engine.

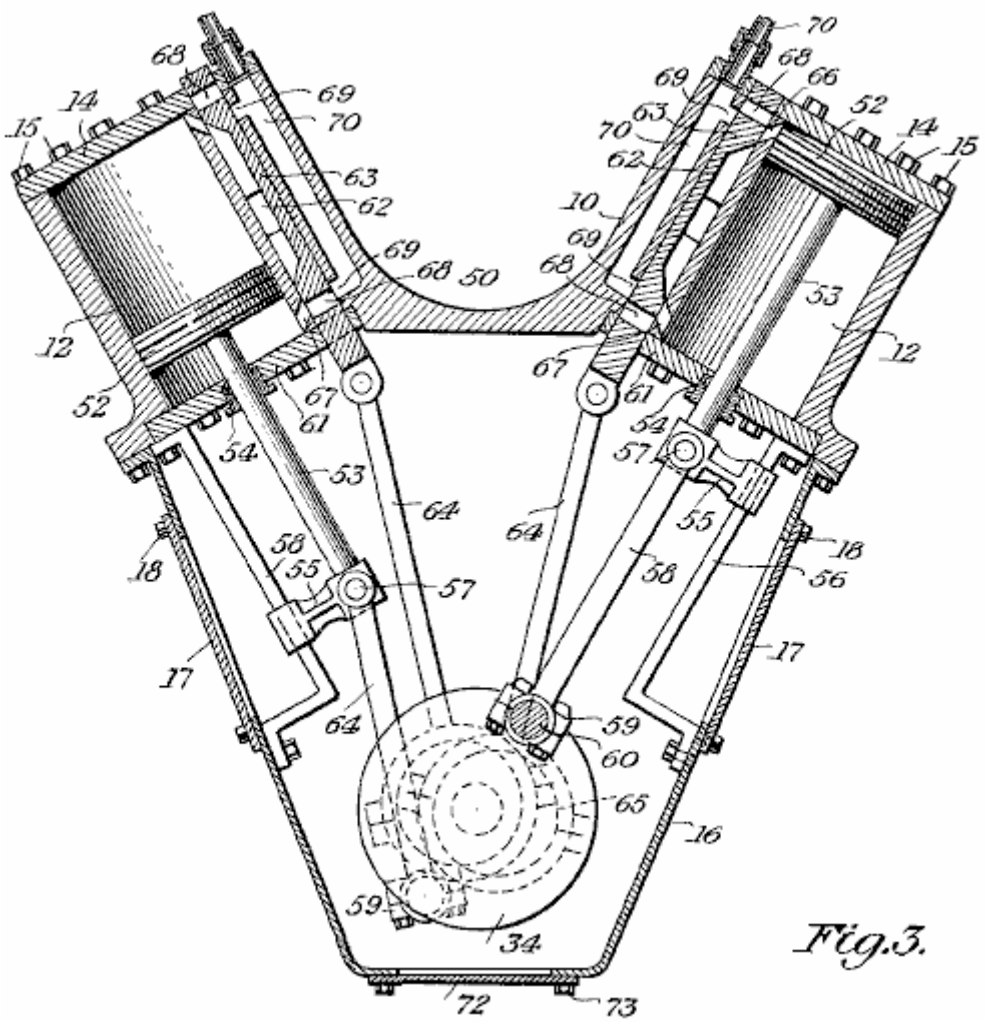


Fig.3 is a vertical cross-sectional view through the power part of the engine.

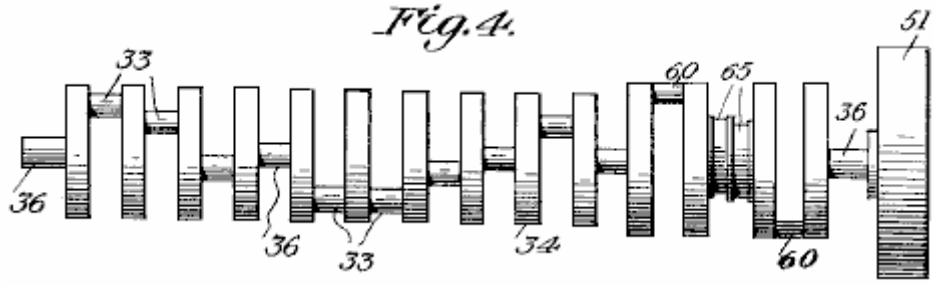


Fig.4 is a detail elevation of the crankshaft of the engine.

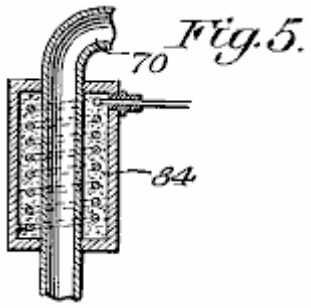


Fig.5 is an enlarged cross-sectional view through one of the electric heaters for the engine.

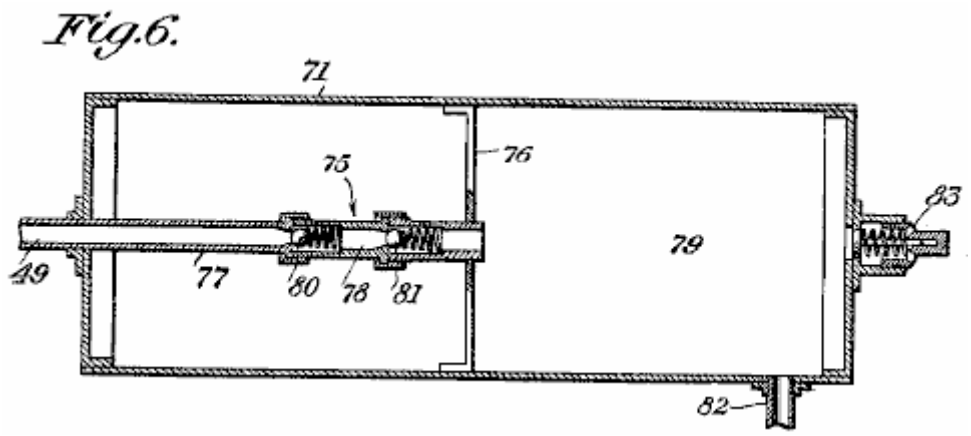
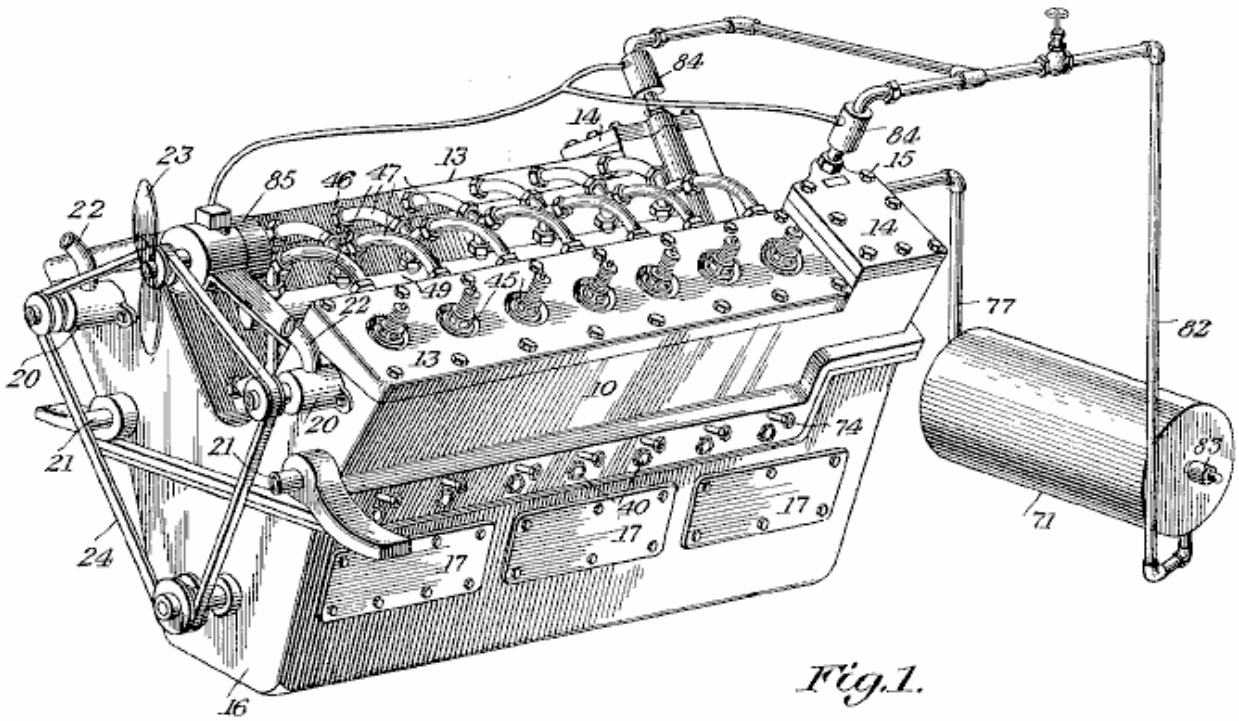


Fig.6 is a vertical, longitudinal, cross-sectional view through the air storage tank, including the equaliser.

The same reference numbers are used for each individual part in every view in every drawing.



Referring to the drawings in detail, the engine in its entirety, composes a cylinder block **10** having inside it, the series of compressor cylinders **11** and the power cylinders **12**. The block **10** is of the V-type and the upper ends of the cylinders are closed off by the removable heads **13** and **14** which are held in place by conventional head bolts **15**. Beneath block **10** is the crank case **16**, which has detachable plates **17** at opposite sides, held in place by fasteners **18**, and seated so as to be leak proof. The block **10** is chambered to provide a water jacket **19** surrounding the cylinders, while at the forward end of the block are water pumps **20**, circulating water through the inlet pipe **21** which leads into the jacket and the water exits from the jacket through the outlet pipe **22**. Beside the pumps **20**, is a fan 23 which is operated from the same belt **24** which drives the pumps.

Working inside the cylinders **11**,are the reciprocating pistons **25**, their rods **26** sliding through packing glands **27** and fixed to crossheads **28** which slide on their mounting guides **29** which are secured to the walls of the crank case **16**. These crossheads **28** are fitted with wrist pins **30**, forming a pivoting connection with the connecting rods **31**, which are connected to their cranks **33** by their bearings **32**. The cranks **33** form part of a counter balanced crankshaft **34**, which is mounted in supports **35** attached to the crank case **16**, the shaft being provided with the required bearings **36**.

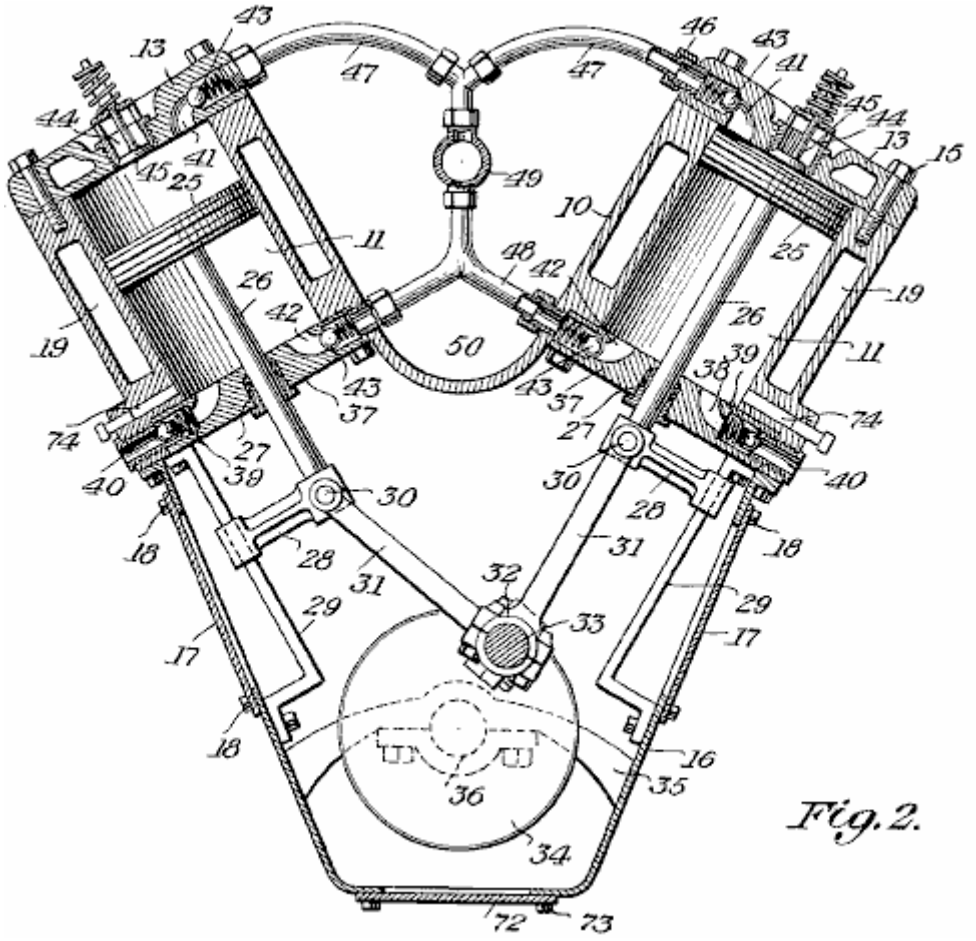
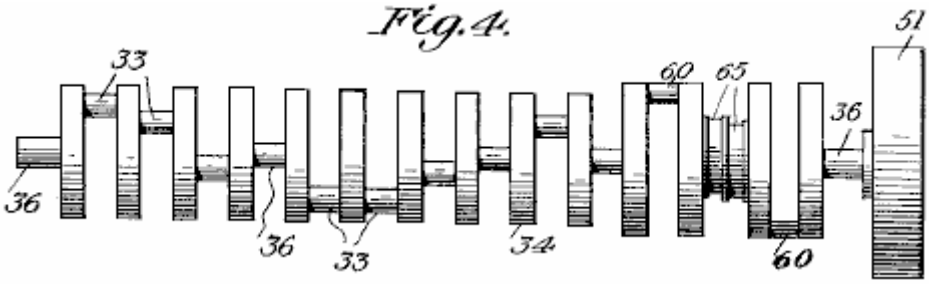


Fig. 2.

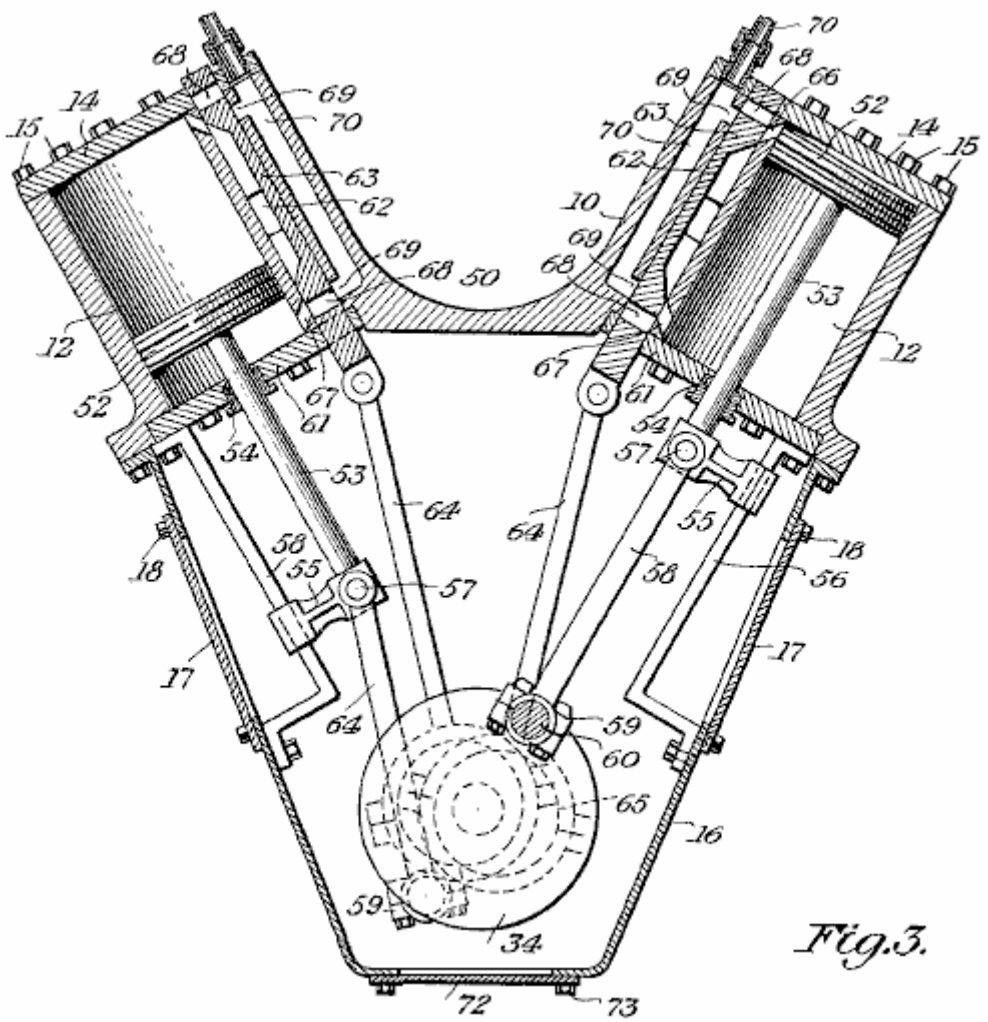
The inner ends of the cylinders **11** are fitted with inner end heads **37**, which are provided with air intake ports **38** fitted with spring ball inlet checks **39**, the air entering through passages **40** which open outside the block **10**. Glands **27** are mounted in the heads **37**.

The heads **13** and **37** are provided with the compressed air outlets **41** and **42**, which are fitted with spring ball checks **43**. The heads **13** are also provided with the central air inlets **44**, which are fitted with spring checks **45**. Couplings **46** attach the air outlets **41** and **42** to their outlet feed pipes **47** and **48**. These pipes lead to a main conduit **49** which is located in the centre channel **50** of the block **10**.

Fig. 4.



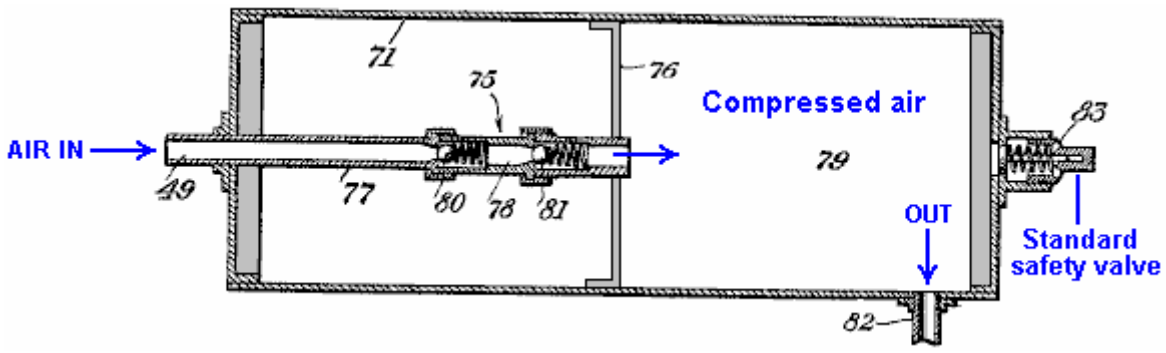
At the rear end of the block **10**, mounted on shaft **36**, there is a conventional flywheel **51**.



Working inside the cylinders **12** are the pistons **52**, with their piston rods **53** sliding through packing glands **54** and fixed in crossheads **55** which slide along their mounting guides **56**, mounted on the inner walls of the crank case **16**. The crossheads **55** have wrist pins **57** which provide a pivoting joint for the connecting rods **58** which are connected by their bearings **59** to their cranks **60** of the crank shaft **34**, the inner ends of the cylinders **12** being closed by the inner heads **61** and their associated glands **54**.

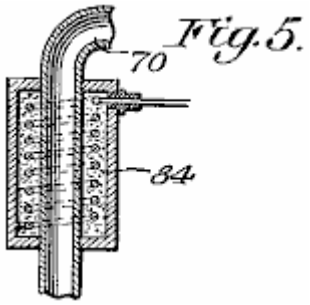
On the cylinders **12** are slide valve chests **62** in which are the slide valves **63**, these being operated by throw rods **64** actuated by cams **65** and the valves controlling the admission and exhaust of air into and out of the cylinders **12**, through the ports **66** and **67**, and these valves **63** are provided with ports **68** for the delivery of air under pressure from the inlet passages **69** common to a pipe **70** coming from a compressed air storage tank **71**.

The bottom of the crank case **16** is fitted with a removable plate **72** which is secured in place by fasteners **73**, and when this plate is removed, it provides access to the crank shaft **34** and the bearings for the engine, as well as other parts inside the crank case.



Leading into the cylinders **11** are the passages **74** of a lubricating system (not shown). The compressed air storage tank **71** has inside it a double-check discharge nozzle **75**, supported by member **76**. Leading to this equaliser is an air inlet pipe **77** which connects through its valved section **78** to the compressed air reservoir **79**.

In the equaliser **75**, are the spaced spring ball checks **80** and **81**, one being for the inlet side and the other for the outlet side of the equaliser. This pipe **77** is connected with the main conduit **49**, while a pipe **82** connects to pipe **70**. The tank is also fitted with an automatic relief valve **83** and this valve can be of any approved type.



Placed around the pipes **70** which connect to the air passages **69** (**Fig.3**) are electric heating units **84** to heat the pressurised air to above freezing temperature when delivered from tank **71** to the cylinders **12**. Supported on the block **10** is an electric generator **85** which is driven from the shaft **34** (**Fig.2**) through a belt **24** (**Fig.1**) and this generator is included in an electric circuit which also has the heaters **84** so that these will operate from current supplied by the generator.

The compressed air storage tank **71** with the equaliser is constructed so that it is possible to pump air into it while it contains an air pressure of 200 pounds per square inch while the compressors are only pumping against 15 pounds per square inch of (atmospheric) pressure. An outside air pressure source can be coupled with the tank to augment that pressure derived from the cylinders **11** of the engine.

CLAIMS

What is claimed is:
In a structure of the kind described, a V-shaped cylinder block provided with upwardly divergent cylinders, end heads fitted to said cylinders at opposite ends thereof, each head having valved inlets and outlets, a main outlet lead between the cylinders of the block for a storage tank and having lateral branches to the outlets at the inner sides of said heads, one inlet being located at the centre of each head at the outer ends of said cylinders while the remaining inlets are at the outer sides of the heads at the inner ends of said cylinders, a substantially V-shaped crank case fitted to the block beneath the cylinders, a counterbalanced crank shaft journaled in the crank case, pistons operating in the cylinders and having rods extended into the crank case, crosshead guides fitted to the interior sides of said case, crossheads connecting the rods with the guides and sliding on them and connecting rods operated by the crank shaft and pivoted at the crossheads in order to allow reciprocation of the pistons.

Leroy K. Rogers' Compressed Air Engine

Patent US 4,292,804 6th October 1980 Inventor: Leroy K. Rogers

METHOD AND APPARATUS FOR OPERATING AN ENGINE ON COMPRESSED GAS

ABSTRACT

The present invention relates to a method and apparatus for operating an engine having a cylinder containing a reciprocating piston driven by a compressed gas. The apparatus comprises a source of compressed gas connected to a distributor which conveys the compressed gas to the cylinder. A valve is provided to admit compressed gas to the cylinder when the piston is in an approximately Top Dead Centre position.

In one embodiment of the present invention, the timing of the opening of the valve is advanced so that the compressed gas is admitted to the cylinder progressively further before the Top Dead Centre position of the piston as the speed of the engine increases.

In a further embodiment of the present invention, a valve actuator is provided which increases the length of time over which the valve remains open to admit compressed gas to the cylinder as the speed of the engine increases.

A still further embodiment of the present invention relates to an apparatus for adapting a conventional internal combustion engine for operation on compressed gas.

US Patent References:

3,881,399	May., 1975	Sagi et al.	91/187.
3,885,387	May., 1975	Simington	60/407.
4,018,050	Apr., 1977	Murphy	60/412.

DESCRIPTION

BACKGROUND AND SUMMARY OF THE PRESENT INVENTION

The present invention is a method and apparatus for operating an engine using a compressed gas as the motive fluid. More particularly, the present invention relates to a apparatus for adapting a pre-existing internal combustion engine for operation on a compressed gas.

Air pollution is one of the most serious problems facing the world today. One of the major contributors to air pollution is the ordinary internal combustion engine which is used in most motor vehicles today. Various devices, including many items required by legislation, have been proposed in an attempt to limit the pollutants which an internal combustion engine exhausts to the air. However, most of these devices have met with limited success and are often both prohibitively expensive and complex. A clean alternative to the internal combustion engine is needed to power vehicles and other machinery.

A compressed gas, preferably air, would provide an ideal motive fluid for an engine, since it would eliminate the usual pollutants exhausted from an internal combustion engine. An apparatus for converting an internal combustion engine for operation on compressed air is disclosed in U.S. Pat. No. 3,885,387 issued May 27, 1975 to Simington. The Simington patent discloses an apparatus including a source of compressed air and a rotating valve actuator which opens and closes a plurality of mechanical poppet valves. The valves deliver compressed air in timed sequence to the cylinders of an engine through adapters located in the spark plug holes. However, the output speed of an engine of this type is limited by the speed of the mechanical valves and the fact that the length of time over which each of the valves remains open cannot be varied as the speed of the engine increases.

Another apparatus for converting an internal combustion engine for operation on steam or compressed air is disclosed in U.S. Pat. No. 4,102,130 issued July 25, 1978 to Stricklin. The Stricklin patent discloses a device

which changes the valve timing of a conventional four stroke engine such that the intake and exhaust valves open once for every revolution of the engine instead of once every other revolution of the engine. A reversing valve is provided which delivers live steam or compressed air to the intake valves and is subsequently reversed to allow the exhaust valves to deliver the expanded steam or air to the atmosphere. A reversing valve of this type however does not provide a reliable apparatus for varying the amount of motive fluid injected into the cylinders when it is desired to increase the speed of the engine. Further, a device of the type disclosed in the Stricklin patent requires the use of multiple reversing valves if the cylinders in a multi-cylinder engine were to be fired sequentially.

Therefore, it is an object of the present invention to provide a reliable method and apparatus for operating an engine or converting an engine for operation with a compressed gas.

A further object of the present invention is to provide a method and apparatus which is effective to deliver a constantly increasing amount of compressed gas to an engine as the speed of the engine increases.

A still further object of the present invention is to provide a method and apparatus which will operate an engine using compressed gas at a speed sufficient to drive a conventional automobile at highway speeds.

It is still a further object of the present invention to provide a method and apparatus which is readily adaptable to a standard internal combustion engine, to convert the internal combustion engine for operation with a compressed gas.

Another object of the invention is to provide a method and apparatus which utilises cool expanded gas, exhausted from a compressed gas engine, to operate an air-conditioning unit and/or an oil-cooler.

These and other objects are realised by the method and apparatus of the present invention for operating an engine having at least one cylinder containing a reciprocating piston and using compressed gas as the motive fluid. The apparatus includes a source of compressed gas, a distributor connected it for conveying the compressed gas to the cylinder or cylinders. A valve is provided for admitting the compressed gas to the cylinder when the piston is in an approximately Top Dead Centre position within the cylinder. An exhaust is provided for exhausting the expanded gas from the cylinder as the piston returns to approximately the Top Dead Centre position.

In a preferred embodiment of the present invention, a device is provided for varying the duration of each engine cycle over which the valve remains open to admit compressed gas to the cylinder, dependent upon the speed of the engine. In a further preferred embodiment of the present invention, an apparatus for advancing the timing of the opening of the valve is arranged to admit the compressed gas to the cylinder progressively further and further before the Top Dead Centre position of the piston, as the speed of the engine increases.

Further features of the present invention include a valve for controlling the amount of compressed gas admitted to the distributor. Also, a portion of the gas which has been expanded in the cylinder and exhausted through the exhaust valve, is delivered to a compressor to be compressed again and returned to the source of compressed gas. A gear train can be engaged to drive the compressor selectively at different operating speeds, depending upon the pressure maintained at the source of compressed air and/or the speed of the engine. Still further, a second portion of the exhaust gas is used to cool a lubricating fluid for the engine or to operate an air-conditioning unit.

In a preferred embodiment of the present invention, the valve for admitting compressed gas to the cylinder is operated electrically. The device for varying the duration of each engine cycle, over which the intake valve remains open, as the speed of the engine increases, comprises a rotating element whose effective length increases as the speed of the engine increases, causing a first contact on the rotating element to be electrically connected to a second contact on the rotating element, for a longer period of each engine cycle. The second contact operates the valve causing it to remain in an open position for a longer period of each engine cycle, as the speed of the engine increases.

Still further features of the present invention include an adaptor plate for supporting the distributor above the intake manifold of a conventional internal combustion engine after a carburettor has been removed to allow air to enter the cylinders of the engine through the intake manifold and conventional intake valves. Another adaptor plate is arranged over an exhaust passageway of the internal combustion engine to reduce the cross-sectional area of the exhaust passageway.

BRIEF DESCRIPTION OF THE DRAWINGS

Preferred embodiments of a method and apparatus for operating an engine according to the present invention will be described with reference to the accompanying drawings in which components have the same reference numbers in each drawing.

Fig.1 is a schematic representation of an apparatus according to the present invention arranged on an engine:

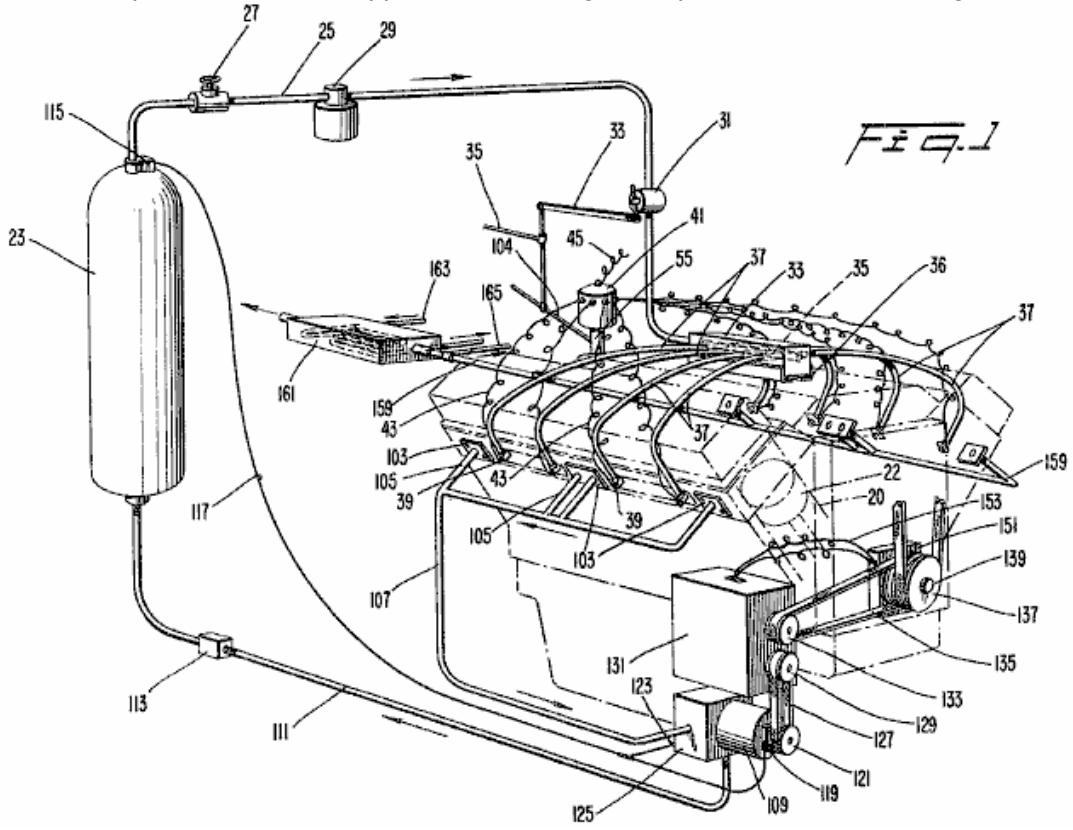


Fig.2 is a side view of one embodiment of a valve actuator according to the present invention.

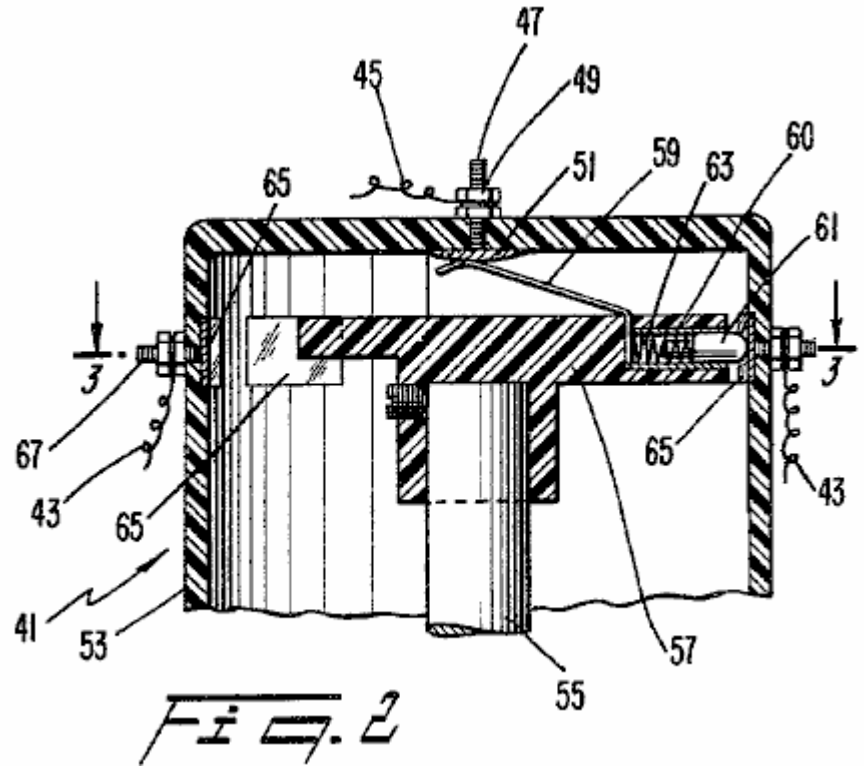


Fig.3 is a cross-sectional view taken along the line **3--3** in **Fig.2**.

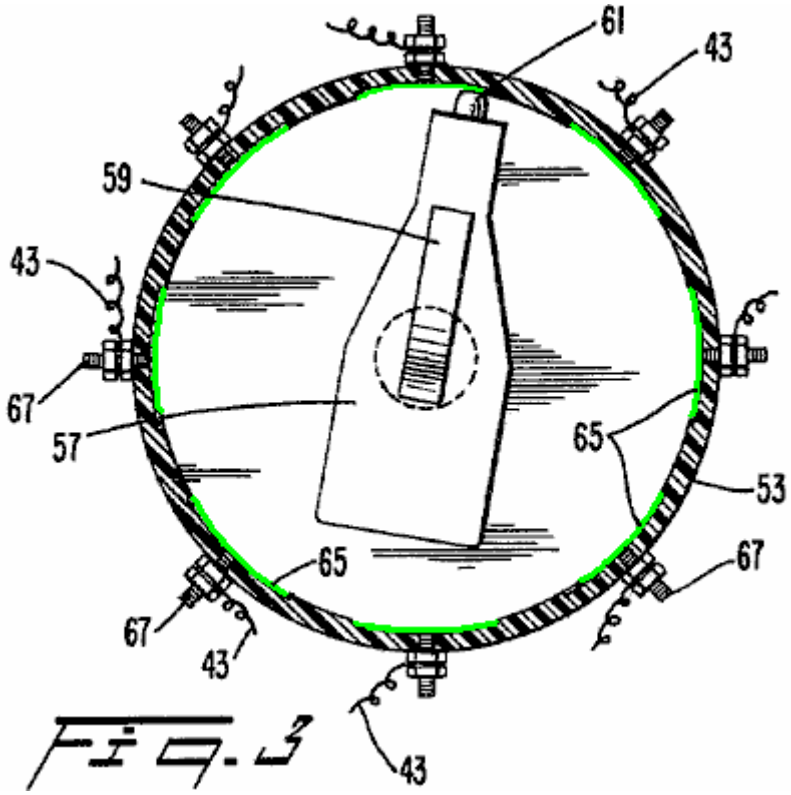


Fig. 3

Fig.4 is a cross-sectional view of a second embodiment of a valve actuator according to the present invention.

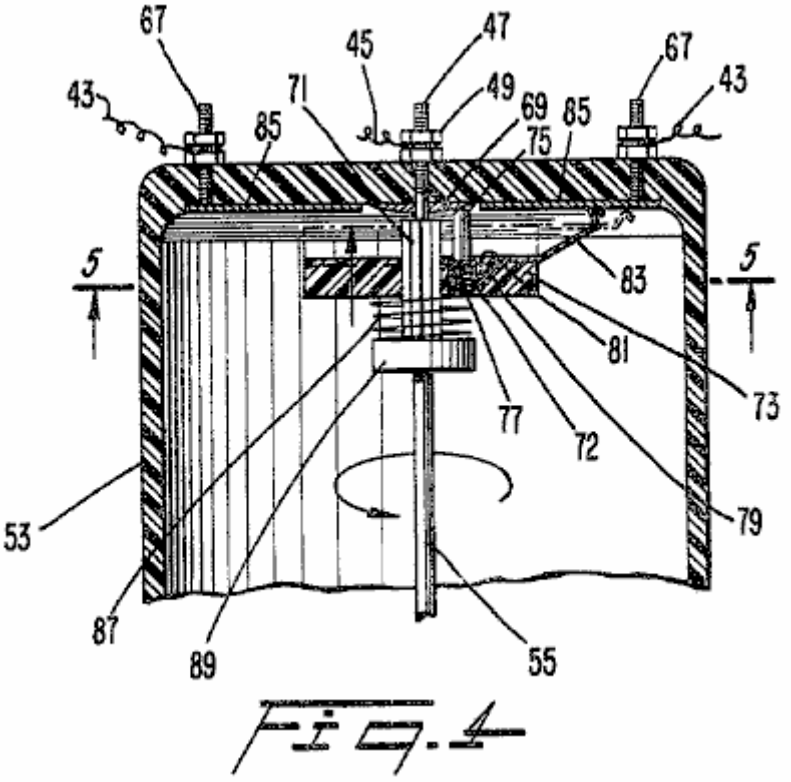


Fig. 4

Fig.5 is a view taken along the line **5--5** in **Fig.4**.

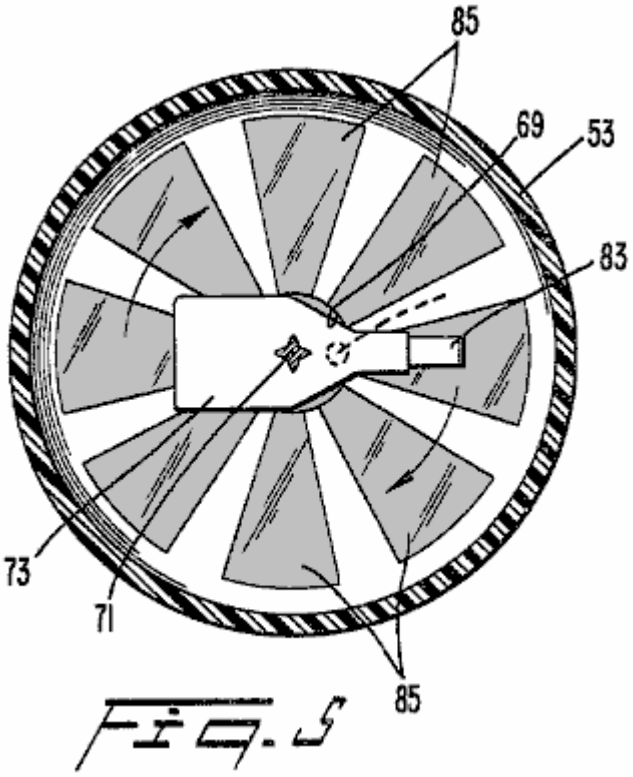


Fig. 5

Fig.6 is a cross-sectional view of a third embodiment of a valve actuator according to the present invention;

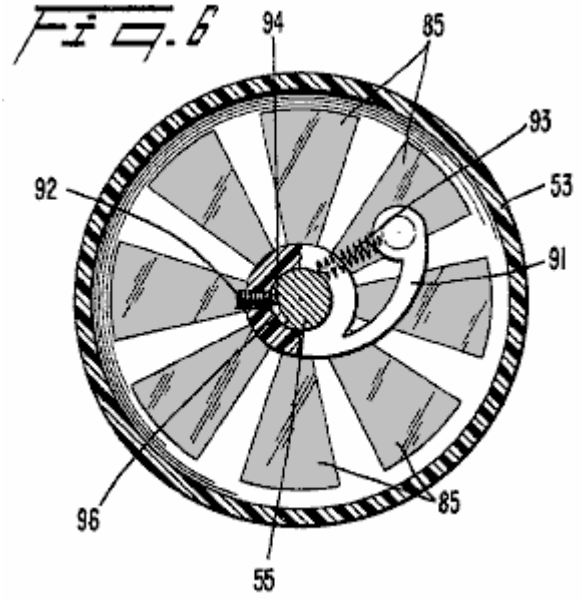


Fig.7 is a view taken along the line **7--7** in **Fig.6**.

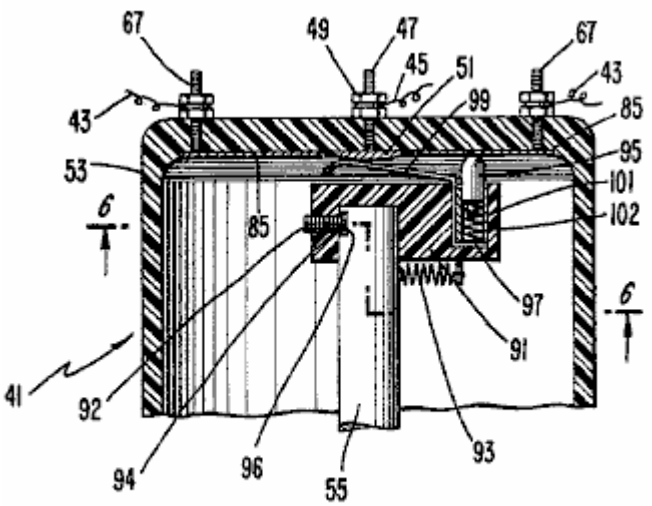
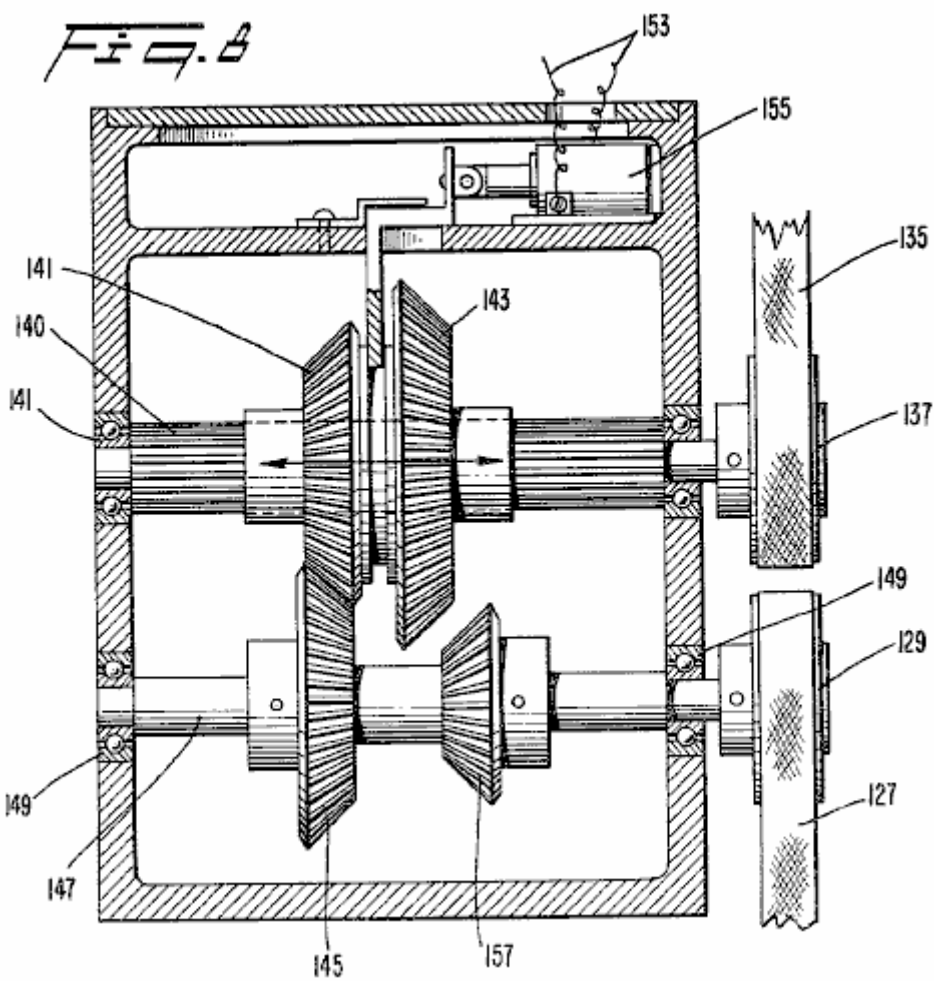

67
49
47
67
45
51
43
43
99
85
95
53
101
6
85
102
97
6
91
41
92
94
96
55
93

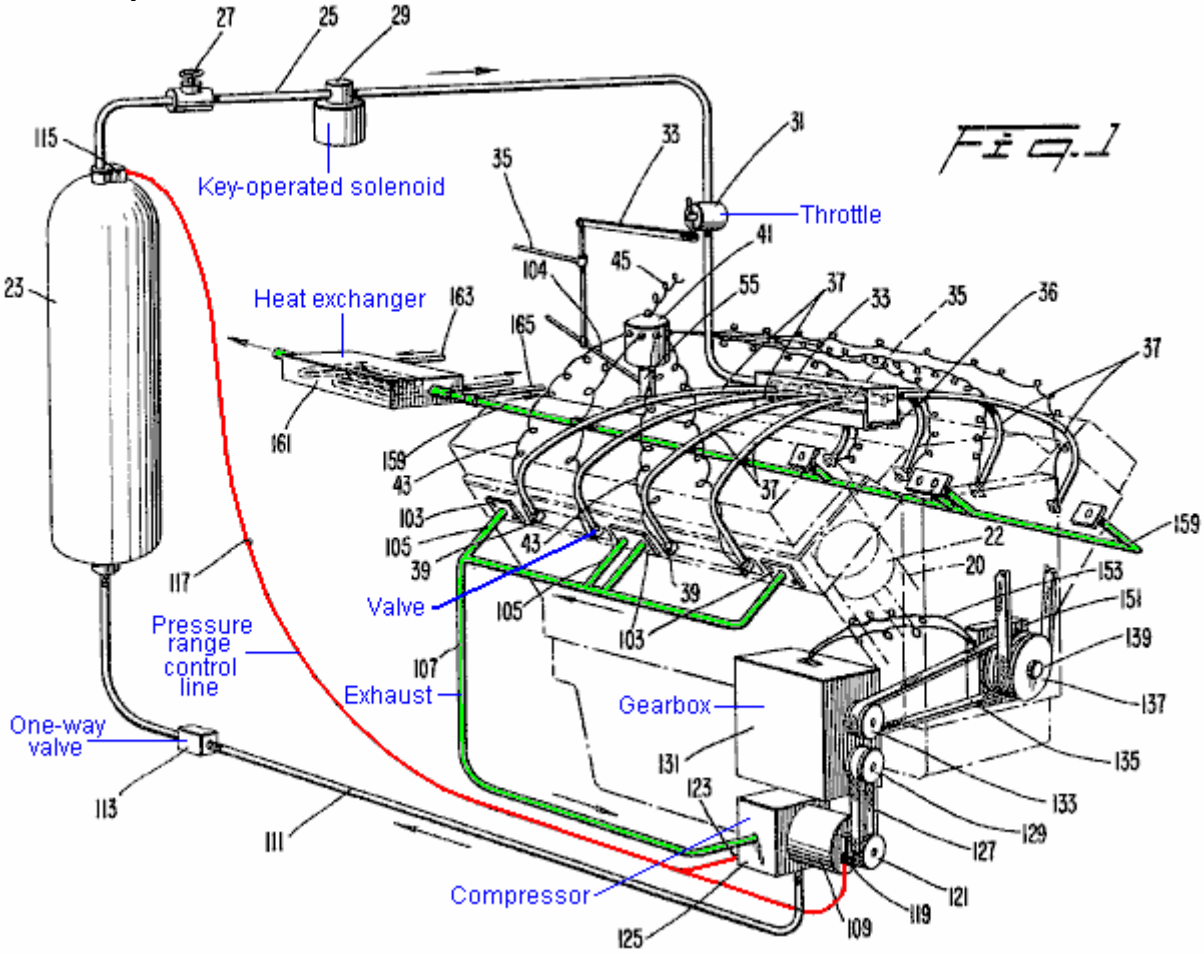
Fig. 7

Fig.8 is a cross-sectional view of a gearing unit to drive a compressor according to the present invention.



DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

With reference to **Fig.1**, an engine block **21** (shown in phantom) having two banks of cylinders with each bank including cylinders **20** having pistons **22** which reciprocate in them in a conventional manner (only one of which is shown in phantom). While the illustrated engine is a V-8 engine, it will be apparent that the present invention is applicable to an engine having any number of pistons and cylinders with the V-8 engine being utilised for illustration purposes only. A compressed gas tank **23** is provided to store a compressed gas at high pressure. It may also be desirable to include a small electric or gas compressor to provide compressed gas to supplement the compressed gas held in the tank **23**. In a preferred embodiment, the compressed gas is air which can be obtained from any suitable source.



A line **25** transports the gas withdrawn from the tank **23** when a conventional shut-off valve **27** is open. In addition, a solenoid valve **29** preferably operated by a suitable key-operated engine switch (not shown) is also placed in the line **25**. In normal operation, the valve **27** is maintained open at all times with the solenoid valve **29** operating as a selective shut off valve to start and stop the engine **21**.

A suitable regulating valve **31** is arranged downstream of the solenoid valve **29** and is connected by a linkage **33** to a throttle linkage **35** which is operator-actuated by any suitable apparatus such as a foot pedal (not shown). The line **25** enters an end of a distributor **33** and is connected to an end of a pipe **35** which is closed at the other end. A plurality of holes, which are equal to the number of cylinders in the engine **21**, are provided on either side of the pipe **35** along the length of the pipe **35**.

When the present invention is used to adapt a conventional internal combustion engine for operation on compressed gas, an adaptor plate **36** is provided to support the distributor **33** in spaced relation from the usual intake opening in the intake manifold of the engine after a conventional carburettor has been removed. In this way, air is permitted to enter the internal combustion engine through the usual passageways and to be admitted to the cylinders through suitable intake valves (not shown). The adaptor plate **36** is attached to the engine block **21** and the distributor **33** by any suitable apparatus, e.g., bolts.

Each of the holes in the pipe **35** is connected in fluid-tight manner to a single line **37**. Each line **37** carries the compressed gas to a single cylinder **20**. In a preferred embodiment, each of the lines **37** is 1/2 inch high pressure plastic tubing attached through suitable connectors to the distributor **33** and the pipe **35**. Each of the lines **37** is connected to a valve **39** which is secured in an opening provided near the top of each of the cylinders **20**. In the case of a conversion of a standard internal combustion engine, the valves **39** can be conveniently screwed into a

tapped hole in the cylinder **20** typically provided for a spark plug of the internal combustion engine. In a preferred embodiment, the valves **39** are solenoid actuated valves in order to provide a fast and reliable opening and closing of the valves **39**.

Each of the valves **39** is energised by a valve actuator **41** through one of a plurality of wires **43**. The valve actuator **41** is driven by a shaft of the engine similar to the drive for a conventional distributor of an internal combustion engine. That is, a shaft **55** of the valve actuator **41** is driven in synchronism with the engine **21** at one half the speed of the engine **21**.

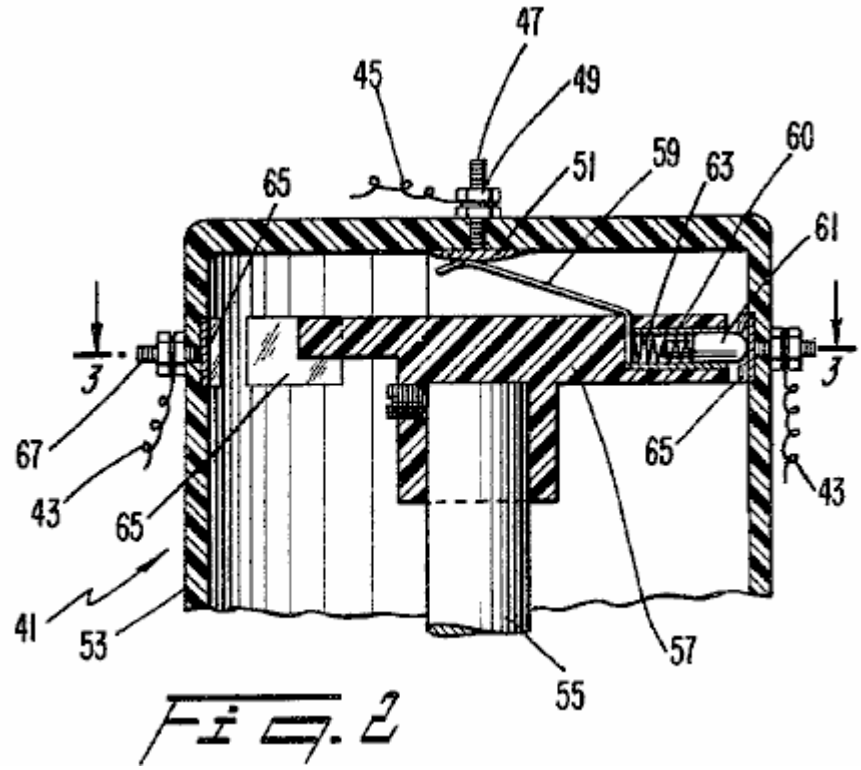


Fig. 2

A first embodiment of the valve actuator **41** (**Fig.2** and **Fig.3**), receives electrical power through a wire **45** which is energised in a suitable manner by a battery, and a coil if necessary (not shown) as is conventional in an internal combustion engine. The wire **45** is attached to a central post **47** by a nut **49**. The post **47** is connected to a conducting plate **51** arranged in a housing **53** for the valve actuator **41**. Within the housing **53**, the shaft **55** has an insulating element **57** secured to an end of the shaft **55** and rotates with it when the shaft **55** is driven by the engine **21**. A first end of a flexible contact **59** is continuously biased against the conducting plate **51** to receive electricity from the battery or other suitable source. The other end of the contact **59** is connected to a conducting sleeve **60** which is in constant contact with a spring biased contact **61** which is arranged within the sleeve **60**. The contact **61** is pressed by a spring **63** which pushes contact **61** towards a side wall of the housing **53**.

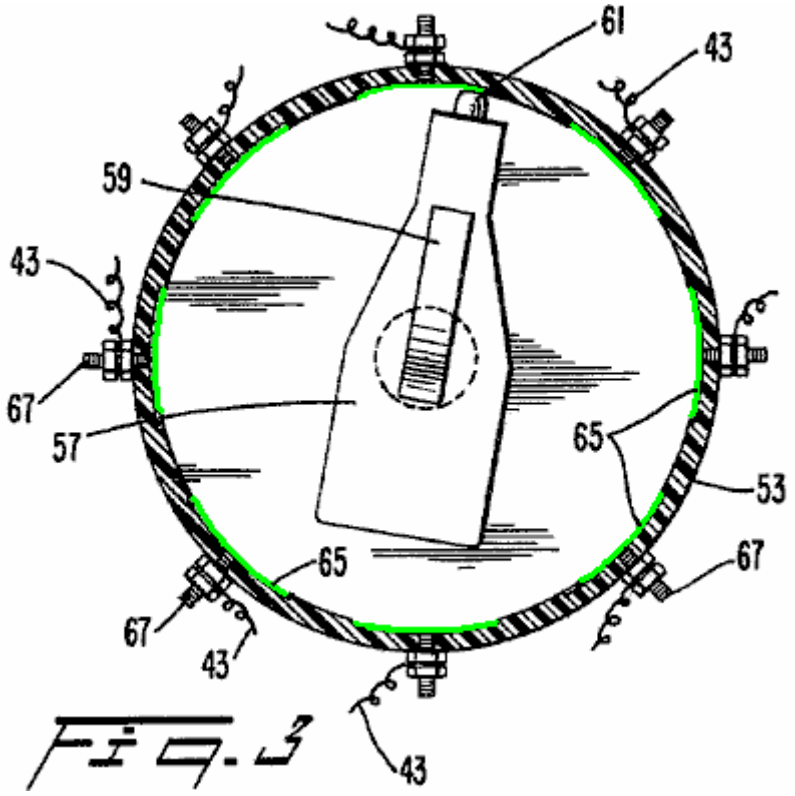


Fig. 3

With reference to **Fig.3**, a plurality of contacts **65** are spaced from one another and are arranged around the periphery of the housing **53** at the same level as the spring biased contact **61**. Each contact **65** is electrically connected to a post **67** which extends outside of the housing **53**. The number of contacts **65** is equal to the number of cylinders in the engine **21**. One of the wires **43**, which actuate the valves **39**, is secured to each of the posts **67**.

In operation, as the shaft **55** rotates in synchronism with the engine **21**, the insulating element **57** rotates and electricity is ultimately delivered to successive pairs of the contacts **65** and wires **43** through the spring loaded contact **61** and the flexible contact **59**. In this way, each of the electrical valves **39** is activated and opened in the proper timed sequence to admit compressed gas to each of the cylinders **20** to drive the pistons **22** on a downward stroke.

The embodiment illustrated in **Fig.2** and **Fig.3** is effective in causing each of the valves **39** to remain open for a long enough period of time to admit sufficient compressed gas to each of the cylinders **20** of the engine **21** to drive the engine **21**. The length of each of the contacts **65** around the periphery of the housing **53** is sufficient to permit the speed of the engine to be increased when desired by the operator by moving the throttle linkage **35** which actuates the linkage **33** to further open the regulating valve **31** to admit more compressed gas from the tank **23** to the distributor **33**. However, it has been found that the amount of air admitted by the valves **39** when using the first embodiment of the valve actuator **41** (**Fig.2** and **Fig.3**) is substantially more than required to operate the engine **21** at an idling speed. Therefore, it may be desirable to provide a valve actuator **41** which is capable of varying the duration of each engine cycle over which the solenoid valves **39** are actuated, i.e., remain open to admit compressed gas, as the speed of the engine **21** is varied.

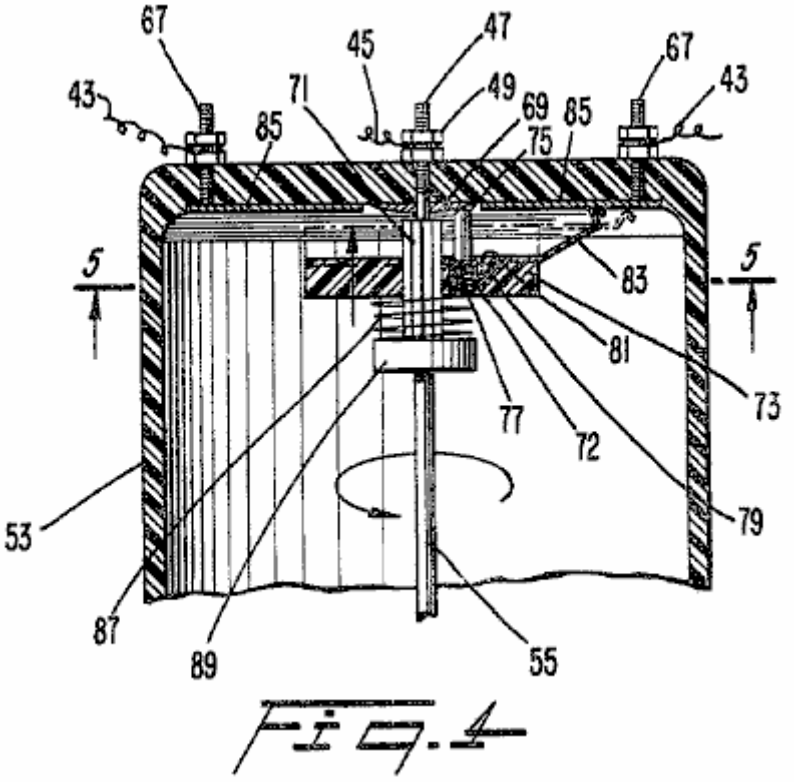


Fig. 4

A second embodiment of a valve actuator **41** which is capable of varying the duration of each engine cycle over which each of the valves **39** remains open to admit compressed gas to the cylinders **20** dependent upon the speed of the engine **21** will be described with reference to **Fig.4** and **Fig.5** wherein members corresponding to those of **Fig.2** and **Fig.3** bear like reference numbers. The wire **45** from the electricity source is attached to the post **47** by the nut **49**. The post **47** has a annular contact ring **69** electrically connected to an end of the post **47** and arranged within the housing **53**. The shaft **55** rotates at one half the speed of the engine as in the embodiment of **Fig.2** and **Fig.3**.

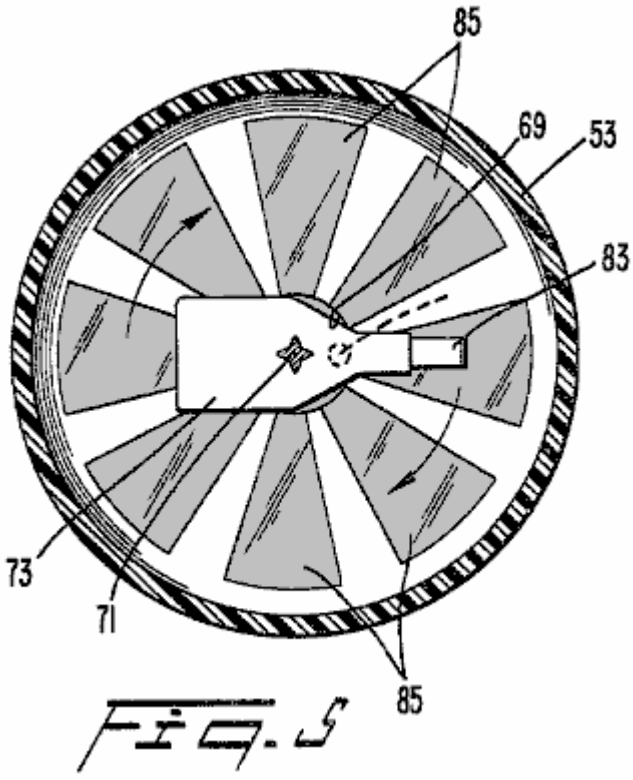


Fig. 5

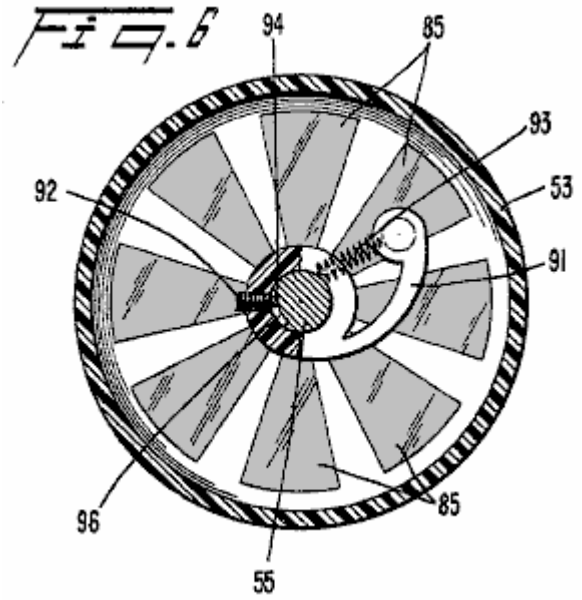
At an upper end of the shaft **55**, a splined section **71** receives a sliding insulating member **73**. The splined section **71** of the shaft **55** holds the insulating member **73** securely as it rotates with shaft **55** but permits the insulating member **73** to slide axially along the length of the splined section **71**. Near the shaft **55**, a conductive sleeve **72** is arranged in a bore **81** in an upper surface of the insulating element **73** generally parallel to the splined section **71**.

A contact **75**, biased towards the annular contact ring **69** by a spring **77**, is arranged within the conductive sleeve **72** and in contact with it. The conductive sleeve **72** also contacts a conductor **79** at a base of the bore **81**.

The conductor **79** extends to the upper surface of the insulating element **73** near an outer periphery of the insulating element **73** where the conductor **79** is electrically connected to a flexible contact **83**. The flexible contact **83** connects, one after the other, with a series of radial contacts **85** which are positioned on an upper inside surface of the housing **53**. A weak spring **87** arranged around the splined section **71** engages a stop member **89** secured on the shaft **55** and the insulating element **73** to slightly bias the insulating element **73** towards the upper inside surface of the housing **53** to ensure contact between the flexible contact **83** and the upper inside surface of the housing **53**. As best seen in **Fig.5**, the radial contacts **85** on the upper inside surface of the housing **53** are arranged generally in the form of radial spokes extending from the centre of the housing **53** with the number of contacts being equal to the number of cylinders **20** in the engine **21**. The number of degrees covered by each of the radial contacts **85** gradually increases as the distance from the centre of the upper inside surface of the housing **53** increases.

In operation of the device of **Fig.4** and **Fig.5**, as the shaft **55** rotates, electricity flows along a path through the wire **45** down through post **47** to the annular contact member **69** which is in constant contact with the spring biased contact **75**. The electrical current passes through the conductive sleeve **72** to the conductor **79** and then to the flexible contact **83**. As the flexible contact **83** rotates along with the insulating member **73** and the shaft **55**, the tip of the flexible contact **83** successively engages each of the radial contacts **85** on the upper inside of the housing **53**. As the speed of the shaft **55** increases, the insulating member **73** and the flexible contact **83** attached to it, move upwards along the splined section **71** of the shaft **55** due to the radial component of the splines in the direction of rotation under the influence of centrifugal force. As the insulating member **73** moves upwards, the flexible contact **83** is bent so that the tip of the contact **83** extends further outwards radially from the centre of the housing **53** (as seen in phantom lines in **Fig.4**). In other words, the effective length of the flexible contact **83** increases as the speed of the engine **21** increases.

As the flexible contact **83** is bent and the tip of the contact **83** moves outwards, the tip remains in contact with each of the radial contacts **85** for a longer period of each engine cycle due to the increased angular width of the radial contacts with increasing distance from the centre of the housing **53**. In this way, the length of time over which each of the valves **39** remains open is increased as the speed of the engine is increased. Thus, a larger quantity of compressed gas or air is injected into the cylinders as the speed increases. Conversely, as the speed decreases and the insulating member **73** moves downwards along the splined section **71**, a minimum quantity of air is injected into the cylinder due to the shorter length of the individual radial contact **85** which is in contact with the flexible contact **83**. In this way, the amount of compressed gas that is used during idling of the engine **21** is at a minimum whereas the amount of compressed gas which is required to increase the speed of the engine **21** to a level suitable to drive a vehicle on a highway is readily available.



Shown in **Fig.6** and **Fig.7**, is a third embodiment of a valve actuator **41** according to the present invention. This embodiment includes a curved insulating element **91** having it's first end able to pivot, being secured by any suitable device such as screw **92** to the shaft **55** for co-rotation with the shaft **55**. The screw **92** is screwed into a tapped hole in the insulating element **91** so that a tab **94** at an end of the screw **92** engages a groove **96** provided in the shaft **55**. In this way, the insulating element **91** rotates positively with the shaft **55**. However, as the shaft **55**

rotates faster, the other end **98** of the insulating element **91** is permitted to pivot outwards under the influence of centrifugal force because of the groove **96** provided in the shaft **55**. A spring **93**, connected between the second end **98** of the element **91** and the shaft **55** urges the second end of the element **91** towards the centre of the housing **53**.

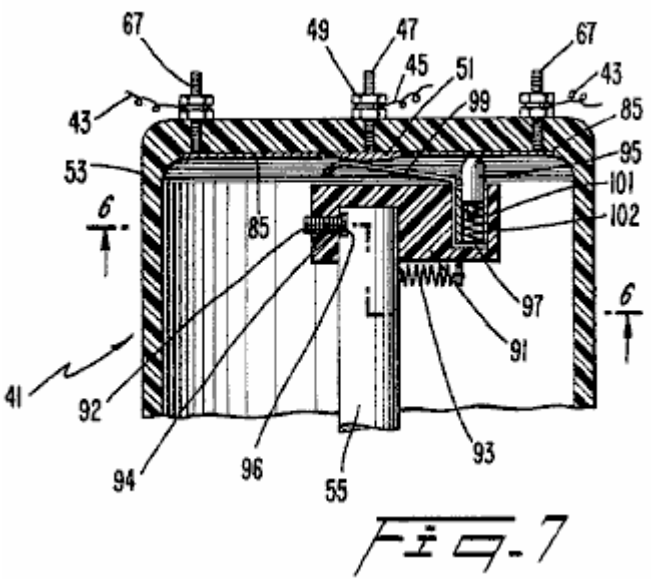


Fig. 7

A contact **99** similar to the contact **59** (Fig.2) is arranged so that one end of the contact piece **99** is in constant contact with the conducting plate **51** located centrally within the housing **53**. The other end of the contact **99** engages a conductive sleeve **101** arranged in bore **102**. A contact element **95** is arranged in the conductive sleeve **101** in constant contact with the sleeve **101**. The bore **102** is arranged generally parallel to the shaft **55** near the second end of the curved insulating element **91**. The contact **95** is biased by a spring **97** towards the upper inside surface of the housing **53** for selective contact with each of the plurality of radial contacts **85** which increase in arc length towards the outer peripheral surface of the housing **53** (**Fig.6**).

When the device shown in **Fig.6** and **Fig.7** is operating, as the shaft **55** rotates the curved insulating element **91** rotates with the shaft **55** and the second end **98** of the insulating element **91** tends to pivot about the shaft **55** due to centrifugal force. Thus, as the effective length of the contact **95** increases, i.e., as the curved insulating element **91** pivots further outwards, the number of degrees of rotation over which the contact **95** is in contact with each of the radial contacts **85** on the upper inside surface of the housing **53** increases thereby allowing each of the valves **39** to remain open for a longer period of each engine cycle, which in turn, allows more compressed gas enter the respective cylinder **20** to further increase the speed of the engine **21**.

With reference to **Fig.1**, a mechanical advance linkage **104** which is connected to the throttle linkage **35**, advances the initiation of the opening of each valve **39** such that compressed gas is injected into the respective cylinder further before the piston **22** in the respective cylinder **20** reaches a Top Dead Centre position as the speed of the engine is increased by moving the throttle linkage **35**. The advance linkage **104** is similar to a conventional standard mechanical advance employed on an internal combustion engine. In other words, the linkage **104** varies the relationship between the angular positions of a point on the shaft **55** and a point on the housing **53** containing the contacts. Alternatively, a conventional vacuum advance could also be employed. By advancing the timing of the opening of the valves **39**, the speed of the engine can more easily be increased.

The operation of the engine cycle according to the present invention will now be described. The compressed gas injected into each cylinder of the engine **21** drives the respective piston **22** downwards to rotate a conventional crankshaft (not shown). The movement of the piston downwards causes the compressed gas to expand rapidly and cool. As the piston **22** begins to move upwards in the cylinder **20** a suitable exhaust valve (not shown), arranged to close an exhaust passageway, is opened by any suitable apparatus. The expanded gas is then expelled through the exhaust passageway. As the piston **22** begins to move downwards again, a suitable intake valve opens to admit ambient air to the cylinder. The intake valve closes and the ambient air is compressed on the subsequent upward movement of the piston until the piston reaches approximately the Top Dead Centre position at which time the compressed gas is again injected into the cylinder **20** to drive the piston **22** downwards and the cycle begins again.

In the case of adapting a conventional internal combustion engine for operation on compressed gas, a plurality of plates **103** are arranged, preferably over an end of the exhaust passageways, in order to reduce the outlet size of the exhaust passageways of the conventional internal combustion engine. In the illustrated embodiment, a single

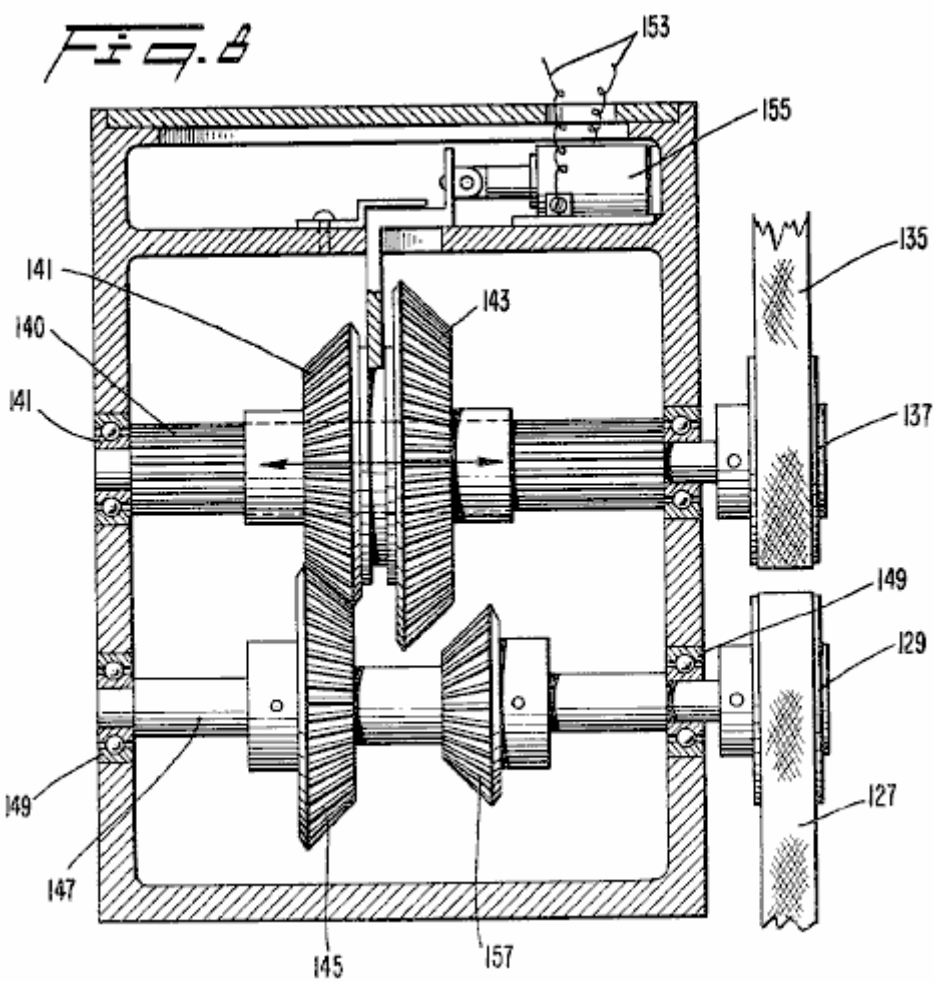
plate having an opening in the centre is bolted to the outside exhaust passageway on each bank of the V-8 engine, while another single plate having two openings in it, is arranged with one opening over each of the interior exhaust passageways on each bank of the V-8 engine. A line **105** is suitably attached to each of the adaptor plates to carry the exhaust to an appropriate location. In a preferred embodiment, the exhaust lines **105** are made from 1.5" plastic tubing.

In a preferred embodiment, the exhaust lines **105** of one bank of the V-8 engine are collected in a line **107** and fed to an inlet of a compressor **109**. The pressure of the exhaust gas emanating from the engine **21** according to the present invention is approximately 25 p.s.i. In this way, the compressor **109** does not have to pull the exhaust into the compressor since the gas exhausted from the engine **21** is at a positive pressure. The positive pressure of the incoming fluid increases the efficiency and reduces wear on the compressor **109**. The exhaust gas is compressed in the compressor **109** and returned through a line **111** and a check valve **113** to the compressed gas storage tank **23**. The check valve **113** prevents the flow of compressed gas stored in the tank **23** back towards the compressor **109**.

A suitable pressure sensor **115** is arranged at an upper end of the tank **23** and sends a signal along a line **117** when the pressure exceeds a predetermined level and when the pressure drops below a predetermined level. The line **117** controls an electrically activated clutch **119** positioned at the front end of the compressor **109**. The clutch **119** is operated to engage and disengage the compressor **109** from a drive pulley **121**. Also, the signal carried by the line **117** activates a suitable valve **123** arranged on compressor housing **125** to exhaust the air entering the compressor housing **125** from the line **107** when the clutch **119** has disengaged the compressor **109** from the drive pulley **121**.

In a preferred embodiment, when the pressure is the tank **23** reaches approximately 600 p.s.i., the clutch **119** is disengaged and the compressor **109** is deactivated and the valve **123** is opened to exhaust the expanded gas delivered to the compressor **109** from the line **107** to the atmosphere. When the pressure within the tank **23** drops below approximately 500 p.s.i., the sensor **115** sends a signal to engage the clutch **119** and close the valve **123**, thereby operating the compressor **109** for supplying the tank **23** with compressed gas.

The pulley **121** which drives the compressor **109** through the clutch **119** is driven by a belt **127** which is driven by a pulley **129** which operates through a gear box **131**. With reference to **Fig.1** and **Fig.8**, a second pulley **133** on the gear box is driven by a belt **135** from a pulley **137** arranged on a drive shaft **139** of the engine **21**. The pulley **137** drives a splined shaft **140** which has a first gear **141** and a second larger gear **143** placed on it, which rotates with the splined shaft **140**. The splined shaft **140** permits axial movement of the gears **141** and **143** along the shaft **140**.



In normal operation (as seen in **Fig.8**), the first gear **141** engages a third gear **145** arranged on a shaft **147** which drives the pulley **129**. The shafts **140** and **147** are arranged in suitable bearings **149** positioned at each end of it. When the speed of the engine **21** drops below a predetermined level, a suitable sensor **151** responsive to the speed of the drive shaft **139** of the engine **21** generates a signal which is transmitted through a line **153** to a solenoid actuator **155** arranged within the gear box **131**. The solenoid actuator **155** moves the first and second gears **141**, **143** axially along the splined shaft **140** to the right as seen in **Fig.8** so that the second, larger gear **143** engages a fourth smaller gear **157** which is arranged on the shaft **147**. The ratio of the second gear **143** to the fourth gear **157** is preferably approximately 3 to 1.

In this way, when the speed of the engine **21** drops below the predetermined level as sensed by the sensor **151** (which predetermined level is insufficient to drive the compressor **109** at a speed sufficient to generate the 500-600 pounds of pressure which is preferably in the tank **23**), the solenoid actuator **155** is energised to slide the gears **143**, **141** axially along the splined shaft **140** so that the second, larger gear **143** engages the fourth, smaller gear **157** to drive the pulley **129** and hence the compressor **109** at a higher rate, to generate the desired pressure. When the speed of the engine increases above the predetermined level, which, in a preferred embodiment is approximately 1500 rpm, the solenoid actuator **155** is deactivated by the sensor **151** thereby moving the gears **143** and **141** to the left as seen in **Fig.8** so that the first gear **141**, engages again with the third gear **145** to effectuate a 1 to 1 ratio between the output shaft **139** of the engine **21** and the pulley **129**.

The other bank of the V-8 engine has its exhaust ports arranged with adapter plates **103** similar to those on the first bank. However, the exhaust from this bank of the engine **21** is not collected and circulated through the compressor **109**. In a preferred embodiment, a portion of the exhaust is collected in a line **159** and fed to an enlarged chamber **161**. A second fluid is fed through a line **163** into the chamber **161** to be cooled by the cool exhaust emanating from the engine **21** in the line **159**. The second fluid in the line **163** may be either transmission fluid contained in a transmission associated with the engine **21** or a portion of the oil used to lubricate the engine **21**. A second portion of the exhaust from the second bank of the V-8 engine is removed from the line **159** in a line **165** and used as a working fluid in an air conditioning system or for any other suitable use.

It should be noted that the particular arrangement utilised for collecting and distributing the gas exhausted from the engine **21** would be determined by the use for which the engine is employed. In other words, it may be

advantageous to rearrange the exhaust tubing such that a larger or smaller percentage of the exhaust is routed through the compressor **109**. It should also be noted that since the exhaust lines **105** are plastic tubing, a rearrangement of the lines for a different purpose is both simple and inexpensive.

In operation of the engine of the present invention, the engine **21** is started by energising the solenoid valve **29** and any suitable starting device (not shown), e.g., a conventional electric starter as used on an internal combustion engine. Compressed gas from the full tank **23** flows through the line **25** and a variable amount of the compressed gas is admitted to the distributor **33** by controlling the regulator valve **31** through the linkage **33** and the operator actuated throttle linkage **35**. The compressed gas is distributed to each of the lines **37** which lead to the individual cylinders **20**. The compressed gas is admitted to each of the cylinders **20** in timed relationship to the position of the pistons within the cylinders by opening the valves **39** with the valve actuator **41**.

When it is desired to increase the speed of the engine, the operator moves the throttle linkage **35** which simultaneously admits a larger quantity of compressed gas to the distributor **33** from the tank **23** by further opening the regulator valve **31**. The timing of the valve actuator **41** is also advanced through the linkage **104**. Still further, as the speed of the engine **21** increases, the effective length of the rotating contact **83** (**Fig.4**) or **95** (**Fig.6**) increases thereby electrically contacting a wider portion of one of the stationary radial contacts **85** to cause each of the valves **39** to remain open for a longer period of each engine cycle to admit a larger quantity of compressed gas to each of the cylinders **20**.

As can be seen, the combination of the regulating valve **31**, the mechanical advance **104**, and the valve actuator **41**, combine to produce a compressed gas engine which is quickly and efficiently adaptable to various operating speeds. However, all three of the controls need not be employed simultaneously. For example, the mechanical advance **104** could be utilised without the benefit of one of the varying valve actuators **41** but the high speed operation of the engine may not be as efficient. By increasing the duration of each engine cycle over which each of the valves **39** remains open to admit compressed gas to each of the cylinders **20** as the speed increases, conservation of compressed gas during low speed operation and efficient high speed operation are both possible.

After the compressed gas admitted to the cylinder **20** has forced the piston **22** downwards within the cylinder to drive the shaft **139** of the engine, the piston **22** moves upwards within the cylinder **20** and forces the expanded gas out through a suitable exhaust valve (not shown) through the adapter plate **103** (if employed) and into the exhaust line **105**. The cool exhaust can then be collected in any suitable arrangement to be compressed and returned to the tank **23** or used for any desired purpose including use as a working fluid in an air conditioning system or as a coolant for oil.

When using the apparatus and method of the present invention to adapt a ordinary internal combustion engine for operation with compressed gas it can be seen that considerable savings in weight are achieved. For example, the ordinary cooling system including a radiator, fan, hoses, etc. can be eliminated since the compressed gas is cooled as it expands in the cylinder. In addition, there are no explosions within the cylinder to generate heat. Further reductions in weight are obtained by employing plastic tubing for the lines which carry the compressed gas between the distributor and the cylinders and for the exhaust lines. Once again, heavy tubing is not required since there is little or no heat generated by the engine of the present invention. In addition, the noise generated by an engine according to the present invention is considerably less than that generated by an ordinary internal combustion engine since there are no explosions taking place within the cylinders.

The principles of preferred embodiments of the present invention have been described in the foregoing specification. However, the invention which is intended to be protected is not to be construed as limited to the particular embodiments disclosed. The embodiments are to be regarded as illustrative rather than restrictive. Variations and changes may be made by others without departing from the spirit of the invention. Accordingly, it is expressly intended that all such variations and changes which fall within the spirit and the scope of the present invention as defined in the appended claims be embraced thereby.

Eber Van Valkinburg's Compressed Fluid Engine

Patent US 3,744,252 10th July 1973 Inventor: Eber Van Valkinburg

CLOSED MOTIVE POWER SYSTEM UTILISING COMPRESSED FLUIDS

ABSTRACT

Stored energy in a compressed elastic fluid is utilised in a controlled manner to pressurise an inelastic fluid and to maintain such pressurisation. The pressurised inelastic fluid is throttled to the impeller of a prime mover. Only a portion of the output energy from the prime mover is utilised to circulate the inelastic fluid so as to maintain a nearly constant volumetric balance in the system.

DESCRIPTION

The objective of the invention is to provide a closed-loop power system which utilises the expansive energy of a compressed elastic fluid, such as air, to pressurise and maintain pressurised throughout the operational cycle of the system a second non-elastic and non-compressible fluid, such as oil. The pressurised non-elastic fluid is released in a controlled manner by a throttle to the rotary impeller of a turbine or the like, having an output shaft. This shaft is coupled to a pump for the non-elastic fluid which automatically maintains the necessary circulation needed for the operation of the prime mover, and maintains a near volumetric balance in the system between the two fluids which are separated by self-adjusting free piston devices. The pump for the non-elastic fluid includes an automatic by-pass for the non-elastic fluid which eliminates the possibility of starving the pump which depends on the discharge of the non-elastic fluid at low pressure from the exhaust of the turbine. Other features and advantages of the invention will become apparent during the course of the following detailed description.

BRIEF DESCRIPTION OF DRAWING FIGURES

Fig.1 is a partly schematic cross-sectional view of a closed motive power system embodying the invention.

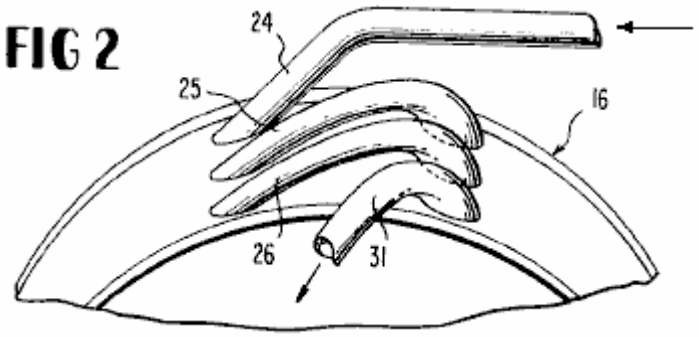


Fig.2 is a fragmentary perspective view of a rotary prime mover utilised in the system.

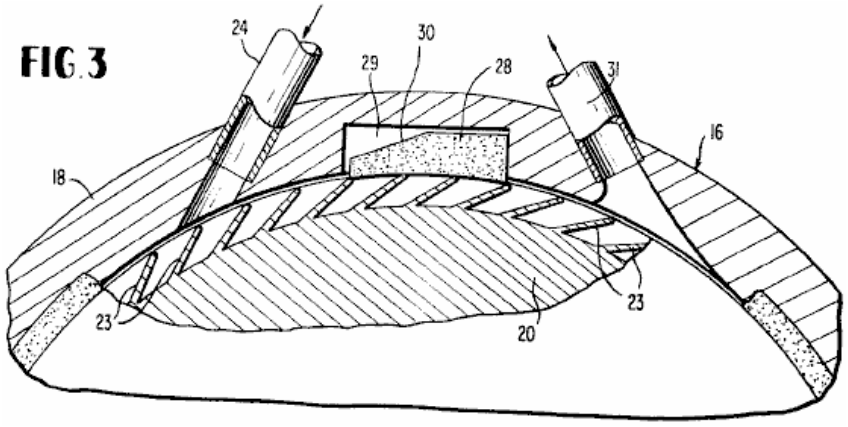


Fig.3 is an enlarged fragmentary vertical section through the prime mover taken at right angles to its rotational axis.

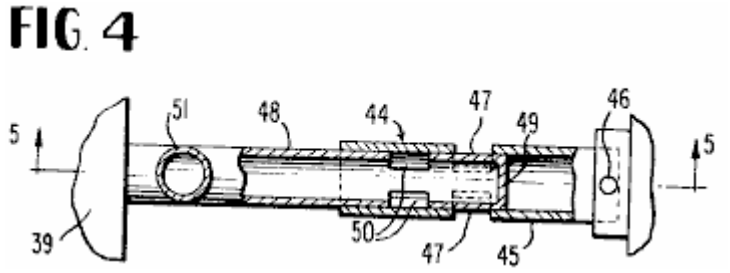


Fig.4 is an enlarged fragmentary vertical section taken on line **4--4** of **Fig.1**.

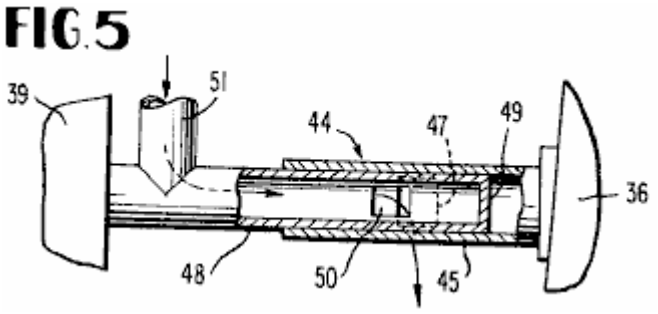
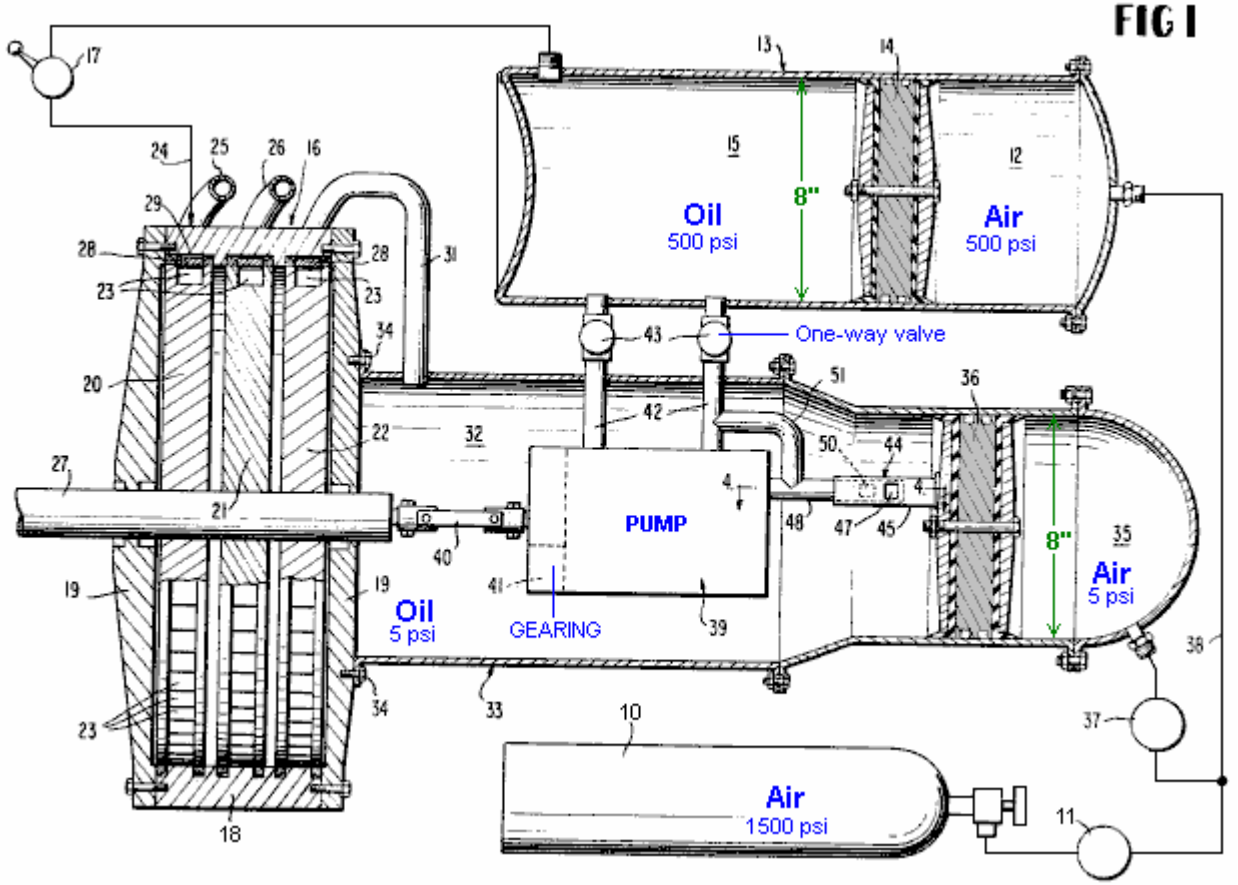


Fig.5 is a similar section taken on line **5--5** of **Fig.4**.

DETAILED DESCRIPTION

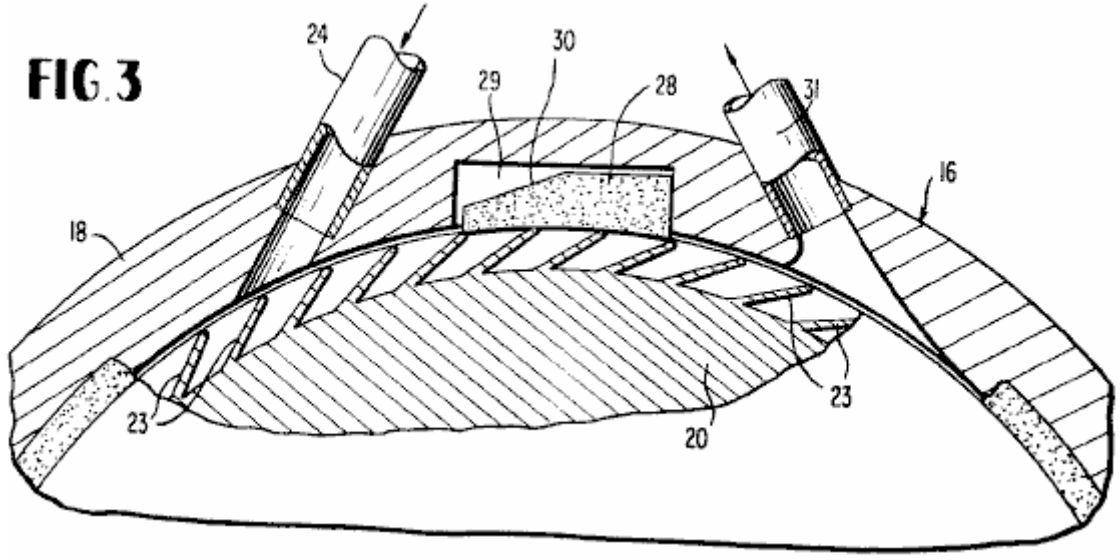


Referring to the drawings in detail, in which the same numbers refer to the same parts in each drawing, the numeral **10** designates a supply bottle or tank for a compressed elastic fluid, such as air. Preferably, the air in the bottle **10** is compressed to approximately 1,500 p.s.i. The compressed air from the bottle **10** is delivered through a suitable pressure regulating valve **11** to the chamber **12** of a high pressure tank **13** on one side of a free piston **14** in the bore of such tank. The free piston **14** separates the chamber **12** for compressed air from a second chamber **15** for an inelastic fluid, such as oil, on the opposite side of the free piston. The free piston **14** can move axially within the bore of the cylindrical tank **13** and is constantly self-adjusting there to maintain a proper volumetric balance between the two separated fluids of the system. The free piston has the ability to maintain the two fluids, air and oil, completely separated during the operation of the system.

The regulator valve **11** delivers compressed air to the chamber **12** under a pressure of approximately 500 p.s.i. The working inelastic fluid, oil, which fills the chamber **15** of high pressure tank **13** is maintained under 500 p.s.i. pressure by the expansive force of the elastic compressed air in the chamber **12** on the free piston **14**. The oil in the chamber **15** is delivered to a prime mover **16**, such as an oil turbine, through a suitable supply regulating or throttle valve **17** which controls the volume of pressurised oil delivered to the prime mover.

The turbine **16** embodies a stator consisting of a casing ring **18** and end cover plates **19** joined to it in a fluid- tight manner. It further embodies a single or plural stage impeller or rotor having bladed wheels **20**, **21** and **22** in the illustrated embodiment. The peripheral blades **23** of these turbine wheels receive the motive fluid from the pressurised chamber **15** through serially connected nozzles **24**, **25** and **26**, connected generally tangentially through the stator ring **18**, as shown in **Fig.3**. The first nozzle **24** shown schematically in **Fig.1** is connected directly with the outlet of the throttle valve **17**. The successive nozzles **25** and **26** deliver the pressurised working fluid serially to the blades **23** of the turbine wheels **21** and **22**, all of the turbine wheels being suitably coupled to a central axial output or working shaft **27** of the turbine **16**.

FIG. 3



Back-pressure sealing blocks **28**, made of fibre, are contained within recesses **29** of casing ring **18** to prevent co-mingling of the working fluid and exhaust at each stage of the turbine. A back-pressure sealing block **28** is actually only required in the third stage between inlet **26** and exhaust **31**, because of the pressure distribution, but such a block can be included in each stage as shown in **Fig.1**. The top surface, including a sloping face portion **30** on each block **28**, reacts with the pressurised fluid to keep the fibre block sealed against the adjacent, bladed turbine wheel; and the longer the slope on the block to increase it's top surface area, the greater will be the sealing pressure pushing it against the periphery of the wheel.

Leading from the final stage of the turbine **16** is a low-pressure working fluid exhaust nozzle **31** which delivers the working fluid, oil, into an oil supply chamber or reservoir **32** of a low pressure tank **33** which may be bolted to the adjacent end cover plate **19** of the turbine, as indicated at **34**. The oil entering the reservoir chamber **32** from the exhaust stage of the turbine is at a pressure of about 3-5 p.s.i. In a second chamber **35** of the low pressure tank **33** separated from the chamber **32** by an automatically moving or self-adjusting free piston **36**, compressed air at a balancing pressure of from 3-5 p.s.i. is maintained by a second pressure regulating valve **37**. The pressure regulating valve **37** is connected with the compressed air supply line **38** which extends from the regulating valve **11** to the high pressure chamber **12** for compressed air.

Within the chamber **32** is a gear pump **39** or the like having its input shaft connected by a coupling **40** with the turbine shaft **27**. Suitable reduction gearing **41** for the pump may be provided internally, as shown, or in any other conventional manner, to gear down the rotational speed derived from the turbine shaft. The pump **39** is supplied with the oil in the filled chamber **32** delivered by the exhaust nozzle or conduit **31** from the turbine. The pump, as illustrated, has twin outlet or delivery conduits **42** each having a back-pressure check valve **43** connected therein and each delivering a like volume of pressurised oil back to the high pressure chamber **15** at a pressure of about 500 p.s.i. The pump **39** also has twin fluid inlets. The pump employed is preferably of the type known on the market as "Hydreco Tandem Gear Pump," Model No. 151515, L12BL, or equivalent. In some models, other types of pumps could be employed including pumps having a single inlet and outlet. The illustrated pump will operate clockwise or counter-clockwise and will deliver 14.1 g.p.m. at 1,800 r.p.m. and 1,500 p.s.i. Therefore, in the present application of the pump **39**, it will be operating at considerably less than capacity and will be under no undue stress.

FIG. 4

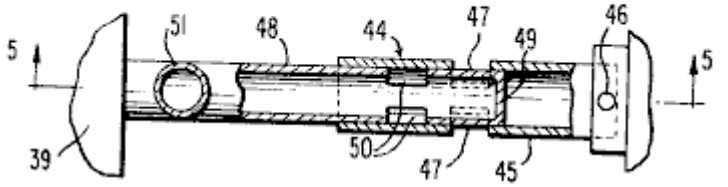
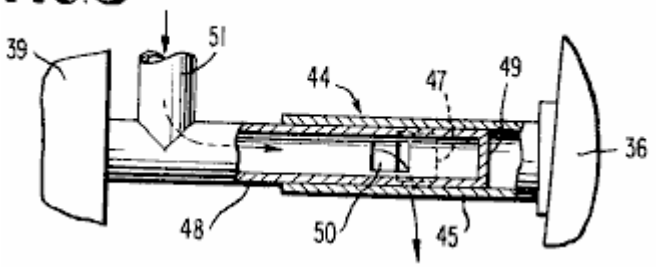


FIG.5



Since the pump depends for its supply of fluid on the delivery of oil at low pressure from the turbine **16** into the chamber **32**, an automatically operating by-pass sleeve valve device **44** for oil is provided as indicated in **Fig.1**, **Fig.4** and **Fig.5**. This device comprises an exterior sleeve or tube **45** having one end directly rigidly secured as at **46** to the movable free piston **36**. This sleeve **45** is provided with slots **47** intermediate its ends. A co-acting interior sleeve **48** engages telescopically and slidably within the sleeve **45** and has a closed end wall **49** and ports or slots **50** intermediate its ends, as shown. The sleeve **48** communicates with one of the delivery conduits **42** by way of an elbow **51**, and the sleeve **48** is also connected with the adjacent end of the pump **39**, as shown.

As long as the chamber **32** is filled with low pressure oil sufficient to balance the low air pressure in the chamber **35** on the opposite side of free piston **36**, such piston will be positioned as shown in **Fig.1** and **Fig.4** so that the slots **47** and **50** of the two sleeves **45** and **48** are out of registration and therefore no flow path exists through them. Under such circumstances, the oil from the chamber **32** will enter the pump and will be delivered by the two conduits **42** at the required pressure to the chamber **15**. Should the supply of oil from the turbine **16** to the chamber **32** diminish so that pump **39** might not be adequately supplied, then the resulting drop in pressure in the chamber **32** will cause the free piston **36** to move to the left in **Fig.1** and bring the slots **47** into registration or partial registration with the slots **50**, as depicted in **Fig.5**. This will instantly establish a by-pass for oil from one conduit **42** back through the elbow **51** and tubes **48** and **45** and their registering slots to the oil chamber **32** to maintain this chamber filled and properly pressurised at all times. The by-pass arrangement is completely automatic and responds to a diminished supply of oil from the turbine into the chamber **32**, so long as the required compressed air pressure of 3-5 p.s.i. is maintained in the chamber **35**.

Briefly, in summary, the system operates as follows. The pressurised inelastic and non-compressible fluid, oil, from the chamber **15** is throttled into the turbine **16** by utilising the throttle valve **17** in a control station. The resulting rotation of the shaft **27** produces the required mechanical energy or work to power a given instrumentality, such as a propeller. A relatively small component of this work energy is utilised through the coupling **40** to drive the pump **39** which maintains the necessary volumetric flow of oil from the turbine back into the high pressure chamber **15**, with the automatic by-pass **44** coming into operation whenever needed.

The ultimate source of energy for the closed power system is the compressed elastic fluid, air, in the tank or bottle **10** which through the regulating valves **11** and **37** maintains a constant air pressure in the required degree in each of the chambers **12** and **35**. As described, the air pressure in the high pressure chamber **12** will be approximately 500 p.s.i. and in the low pressure chamber **35** will be approximately 3-5 p.s.i.

It may be observed in **Fig.1** that the tank **33** is enlarged relative to the tank **13** to compensate for the space occupied by the pump and associated components. The usable volumes of the two tanks are approximately equal.

In an operative embodiment of the invention, the two free pistons **14** and **36** and the tank bores receiving them are 8 inches in diameter. The approximate diameters of the bladed turbine wheels are 18 inches. The pump **39** is approximately 10 inches long and 5 inches in diameter. The tank **13** is about 21 inches long between its crowned end walls. The tank **33** is 10 inches in diameter adjacent to the pump **39**.

The terms and expressions which have been employed herein are used as terms of description and not of limitation, and there is no intention, in the use of such terms and expressions, of excluding any equivalents of the features shown and described or portions thereof but it is recognised that various modifications are possible within the scope of the invention claimed.

Josef Papp's Inert Gas Engine

US Patent 4,428,193 **31st January 1984** **Inventor: Josef Papp**

INERT GAS FUEL, FUEL PREPARATION APPARATUS AND SYSTEM FOR EXTRACTING USEFUL WORK FROM THE FUEL

ABSTRACT

An inert gas fuel consisting essentially of a precise, homogeneous mixture of helium, neon, argon, krypton and xenon. Apparatus for preparing the fuel includes a mixing chamber, tubing to allow movement of each inert gas into and through the various stages of the apparatus, a plurality of electric coils for producing magnetic fields, an ion gauge, ionises, cathode ray tubes, filters, a polarise and a high frequency generator. An engine for extracting useful work from the fuel has at least two closed cylinders for fuel, each cylinder being defined by a head and a piston. A plurality of electrodes extend into each chamber, some containing low level radioactive material. The head has a generally concave depression facing a generally semi-toroidal depression in the surface of the piston. The piston is axially movable with respect to the head from a first position to a second position and back, which linear motion is converted to rotary motion by a crankshaft. The engine's electrical system includes coils and condensers which circle each cylinder, an electric generator, and circuitry for controlling the flow of current within the system.

BACKGROUND OF THE INVENTION

This invention relates to closed reciprocating engines, i.e., ones which do not require an air supply and do not emit exhaust gases, and more particularly to such engines which use inert gases as fuel. It also concerns such inert gas fuels and apparatus for preparing same.

Currently available internal combustion engines suffer from several disadvantages. They are inefficient in their utilisation of the energy present in their fuels. The fuel itself is generally a petroleum derivative with an ever-increasing price and sometimes limited availability. The burning of such fuel normally results in pollutants which are emitted into the atmosphere. These engines require oxygen and, therefore, are particularly unsuitable in environments, such as underwater or outer space, in which gaseous oxygen is relatively unavailable. Present internal combustion engines are, furthermore, relatively complex with a great number of moving parts. Larger units, such as fossil-fuel electric power plants, escape some of the disadvantages of the present internal combustion engine, but not, inter alia, those of pollution, price of fuel and availability of fuel.

Several alternative energy sources have been proposed, such as the sun (through direct solar power devices), nuclear fission and nuclear fusion. Due to the lack of public acceptance, cost, other pollutants, technical problems, and/or lack of development, these sources have not wholly solved the problem.
Moreover, the preparation of fuel for nuclear fission and nuclear fusion reactors has heretofore been a complicated process requiring expensive apparatus.

SUMMARY OF THE INVENTION

Among the several objects of the present invention may be noted the provision of an engine which is efficient; the provision of an engine which does not require frequent refuelling; the provision of an engine which develops no pollutants in operation; the provision of an engine which is particularly suited for use in environments devoid of free oxygen; the provision of an engine which requires no oxygen in operation; the provision of an engine having a relatively small number of moving parts; the provision of an engine of a relatively simple construction; the provision of an engine which can be used in light and heavy-duty applications; the provision of an engine which is relatively inexpensive to make and operate; the provision of a fuel which uses widely available components; the provision of a fuel which is relatively inexpensive; the provision of a fuel which is not a petroleum derivative; the provision of relatively simple and inexpensive apparatus for preparing inert gases for use as a fuel; the provision of such apparatus which mixes inert gases in precise, predetermined ratios; and the provision of such apparatus which eliminates contaminants from the inert gas mixture. Other objects and features will be in part apparent and in part pointed out hereinafter.

Briefly, in one aspect the engine of the present invention includes a head having a generally concave depression in it, the head defining one end of a chamber, a piston having a generally semi-toroidal depression in its upper surface, the piston defining the other end of the chamber, and a plurality of electrodes extending into the chamber for exciting and igniting the working fluid. The piston can move along its axis towards and away from the head, causing the volume of the chamber to alter, depending on the position of the piston relative to the head.

In another aspect, the engine of the present invention includes a head which defines one end of the chamber, a piston which defines the other end of the chamber, a plurality of magnetic coils wound around the chamber for generating magnetic fields inside the chamber, and at least four electrodes extending into the chamber for exciting and igniting the working fluid. The magnetic coils are generally coaxial with the chamber. The electrodes are generally equidistantly spaced from the axis of the chamber and are each normally positioned 90 degrees from the adjacent electrodes. Lines between opposed pairs of electrodes intersect generally on the axis of the chamber to define a focal point.

In a further aspect, the engine of the present invention includes a head which defines one end of a chamber, a piston which defines the other end of the chamber, at least two electric coils wound around the chamber for generating magnetic fields inside the chamber, and a plurality of electrodes extending into the chamber for exciting and igniting the working fluid. The electric coils are generally coaxial with the chamber. And the working fluid includes a mixture of inert gases.

The apparatus of the present invention for preparing a mixture of inert gases for use as a fuel includes a chamber, electric coils for generating predetermined magnetic fields inside the chamber, tubing adapted to be connected to sources of preselected inert gases for flow of the gases from the sources to the chamber, and ionisers for ionising the gases.

The fuel of the present invention includes a mixture of inert gases including approximately 36% helium, approximately 26% neon, approximately 17% argon, approximately 13% krypton, and approximately 8% xenon by volume.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig.1 is a side elevation of an engine of this invention:
Fig.2 is a rear elevation of an engine of this invention:

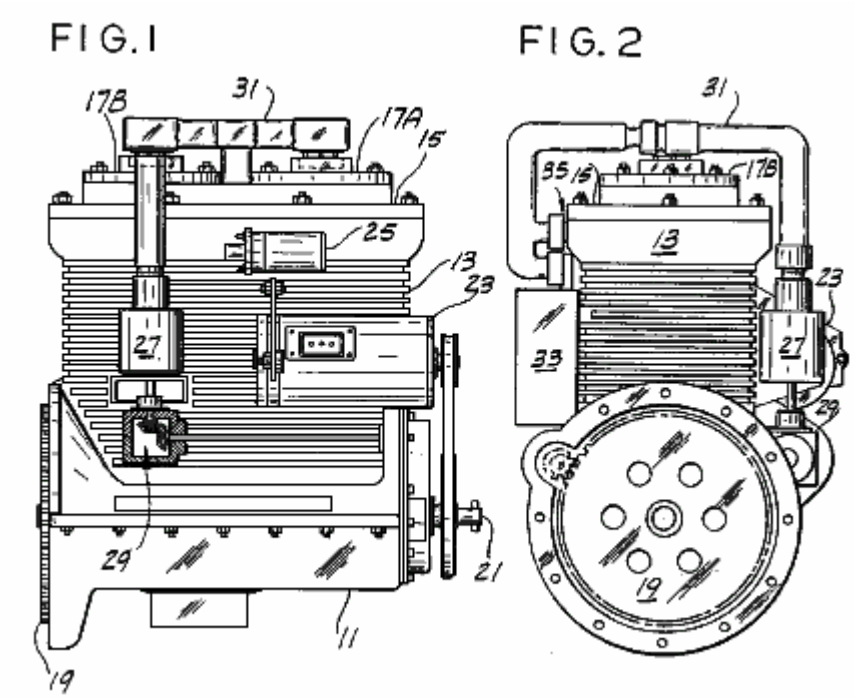


Fig.3 is a top view of an engine of this invention:

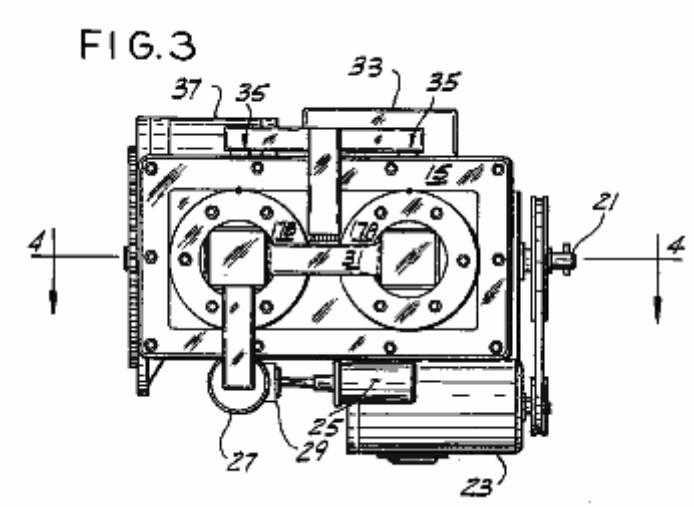


Fig.4 is a cross-sectional view generally along line **4--4** of **Fig.3** of an engine of this invention:

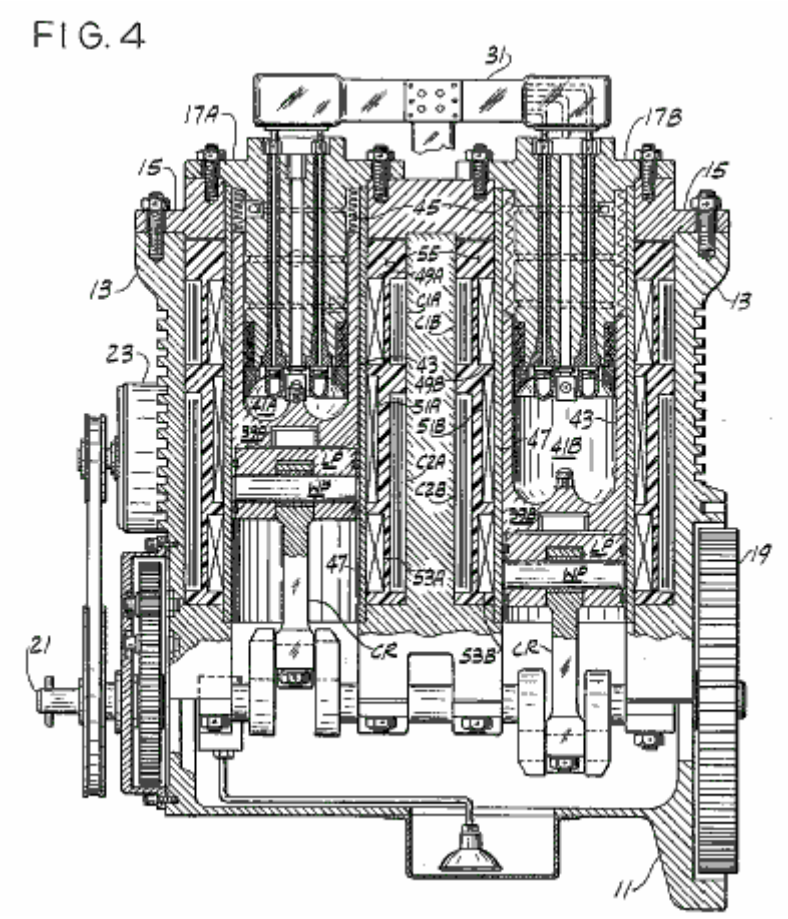


Fig.5 is a cross-sectional view of a cylinder of an engine of this invention:

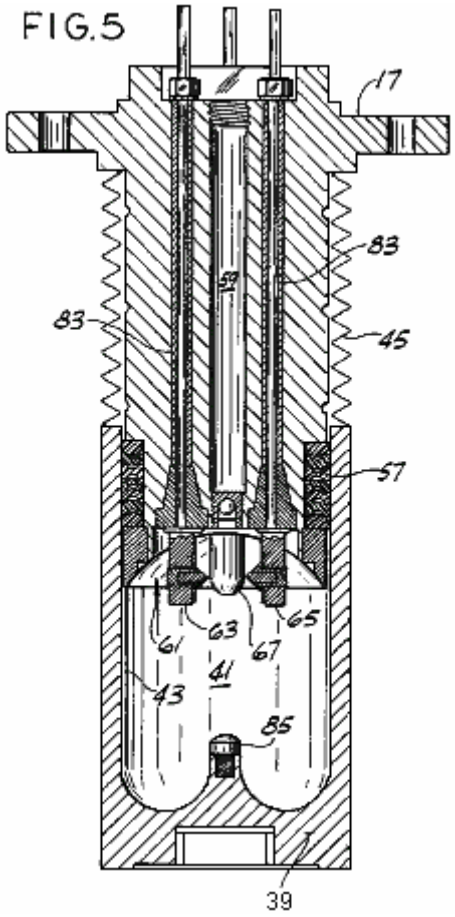


Fig.6 is a plan of the base of a cylinder head of an engine of this invention:

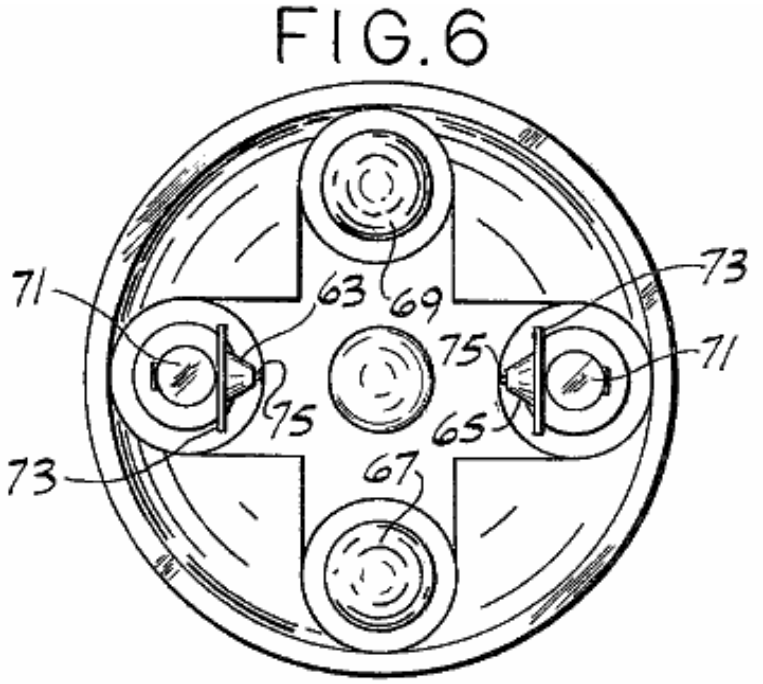


Fig.7 is an elevation of an electrode rod of an engine of this invention:

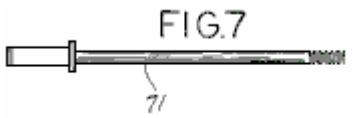


Fig.8 is an elevation, with parts broken away, of one type of electrode used in an engine of this invention:

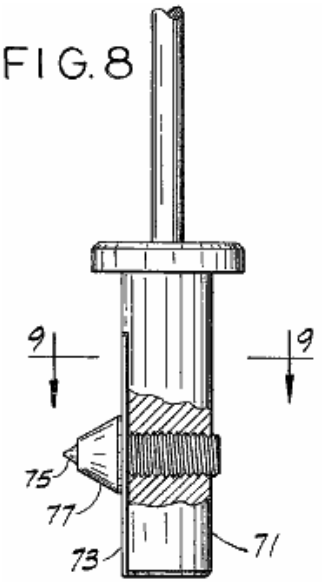


Fig.9 is a view taken generally along line **9--9** of **Fig.8**:

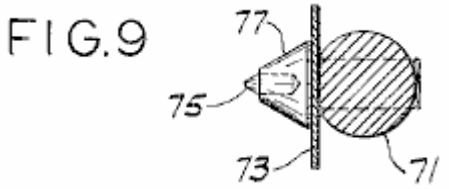


Fig.10 is a cross-sectional view of a second type of electrode used in an engine of this invention:

Fig.11 is a cross-sectional view similar to **Fig.5** showing the piston in its uppermost position:

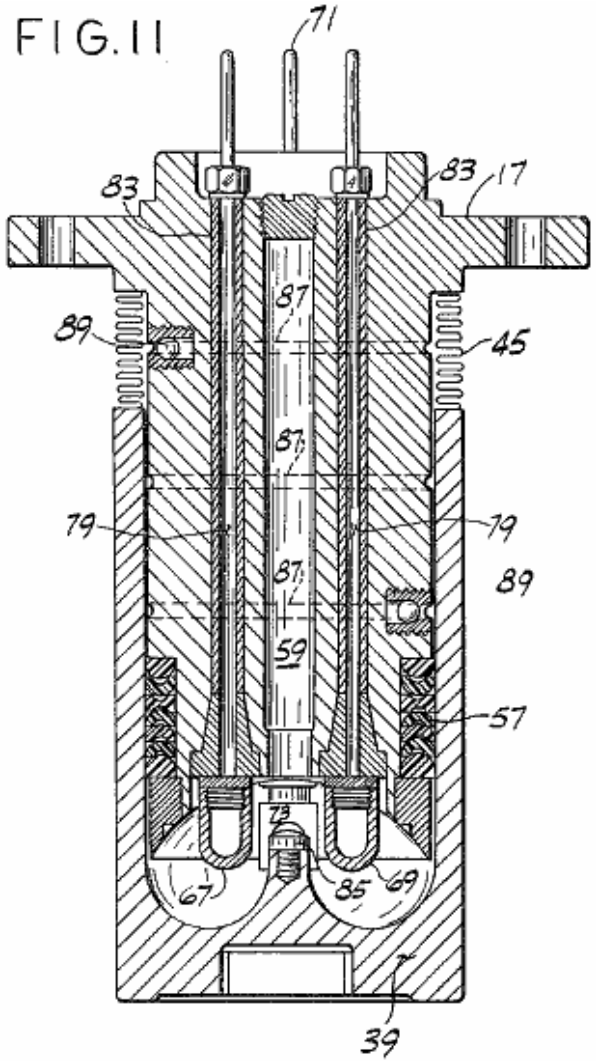


Fig.12 is a cross-sectional view similar to **Fig.5** showing an alternative cylinder used in an engine of this invention:

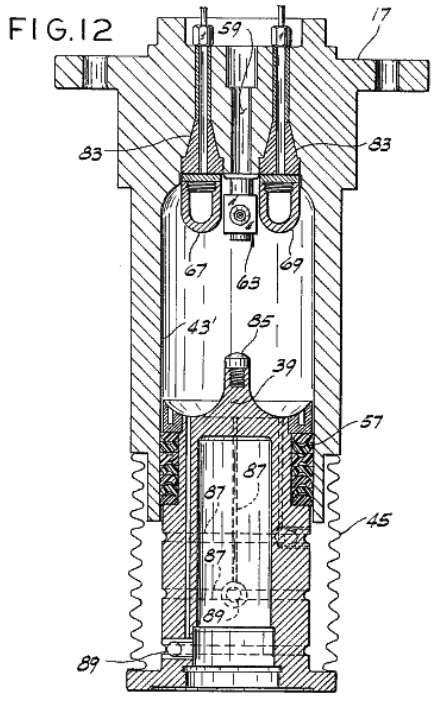


Fig.12A is a cross-sectional view similar to **Fig.5** and **Fig.12**, but on a reduced scale and with parts broken away, showing an additional embodiment of a cylinder head used in an engine of this invention:

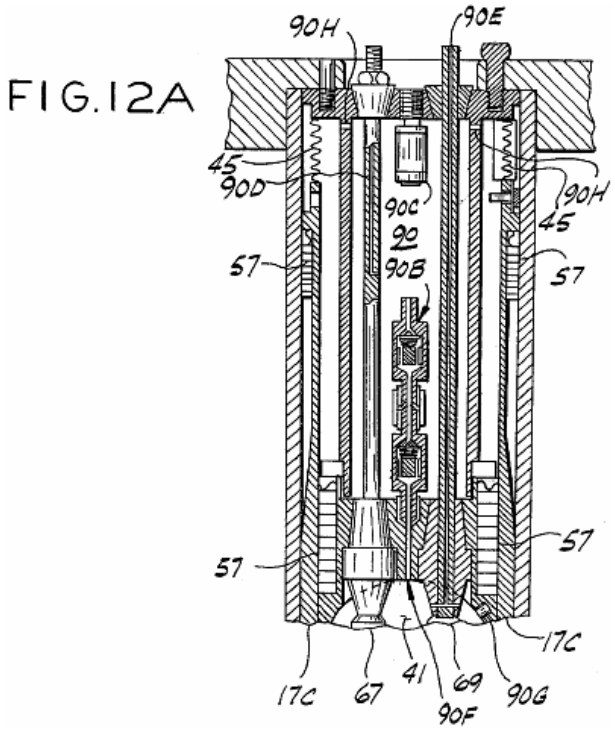


Fig.13A and **Fig.13B** are schematic diagrams of the electrical circuitry for an engine of this invention:

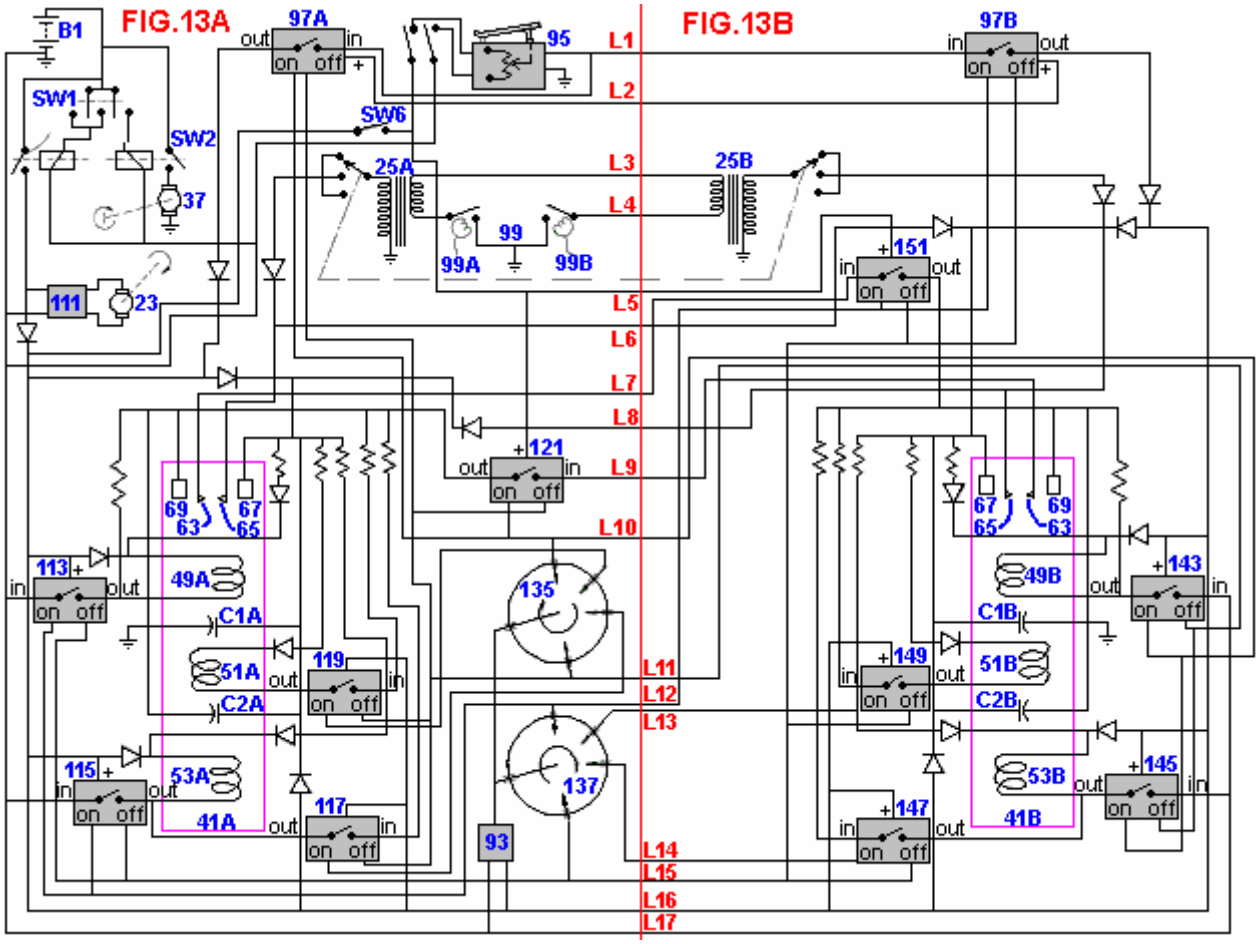


Fig.14 is a schematic diagram of an alternative high-voltage ignition system for an engine of this invention:

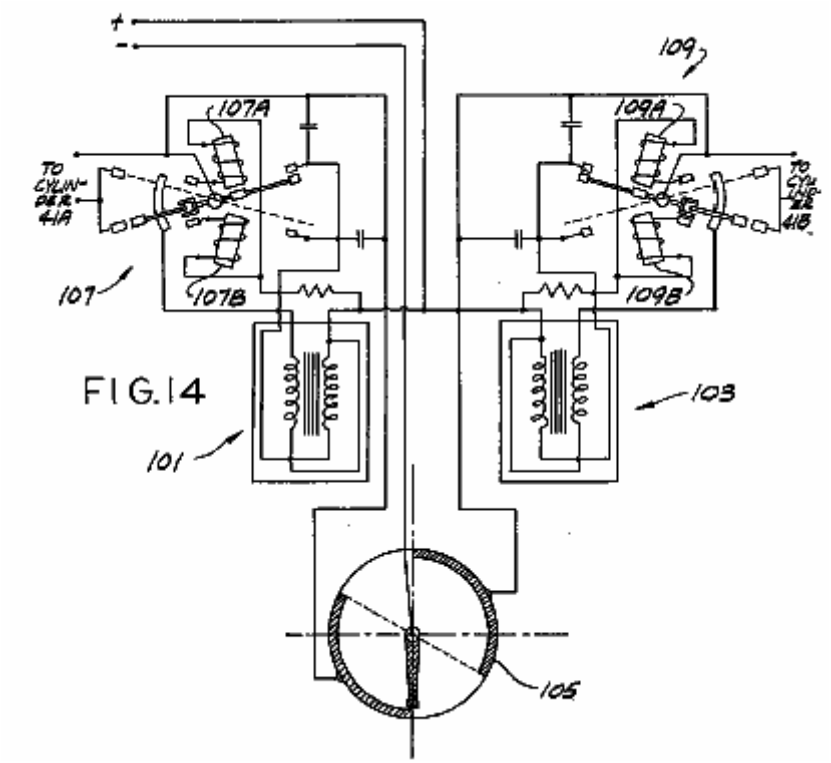


FIG.14

Fig.15 is a schematic diagram of an electronic switching unit for an engine of this invention:
Fig.16 is a schematic diagram of a regulator/electronic switching unit for an engine of this invention:

FIG.15

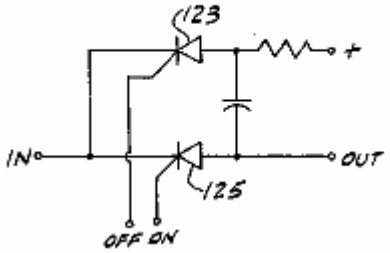
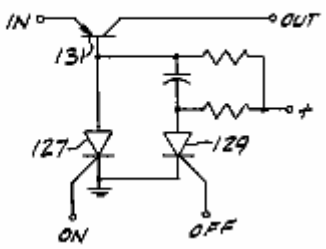
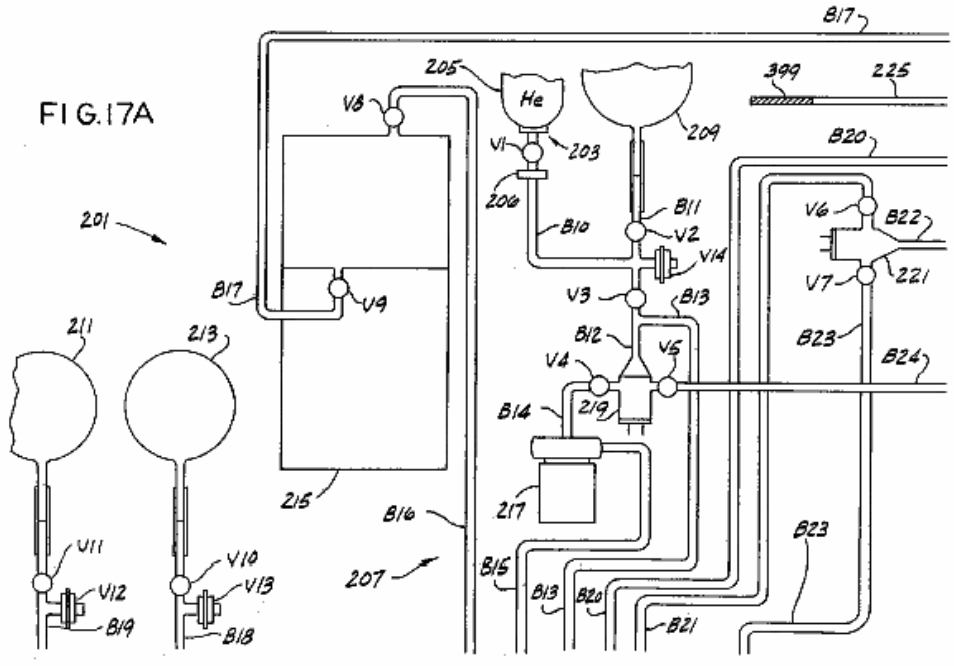
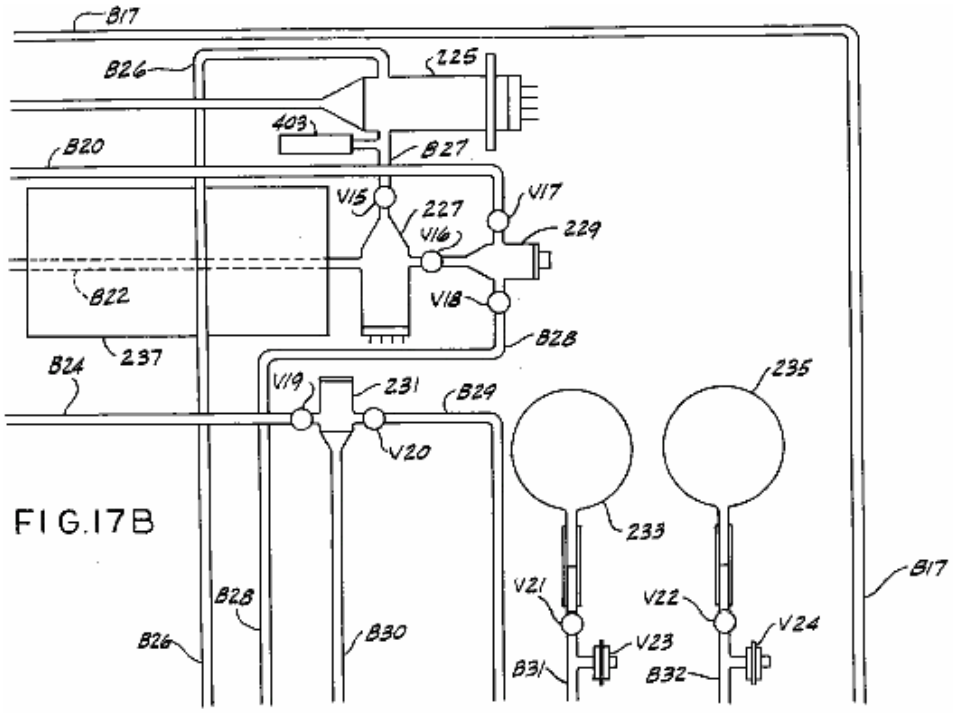


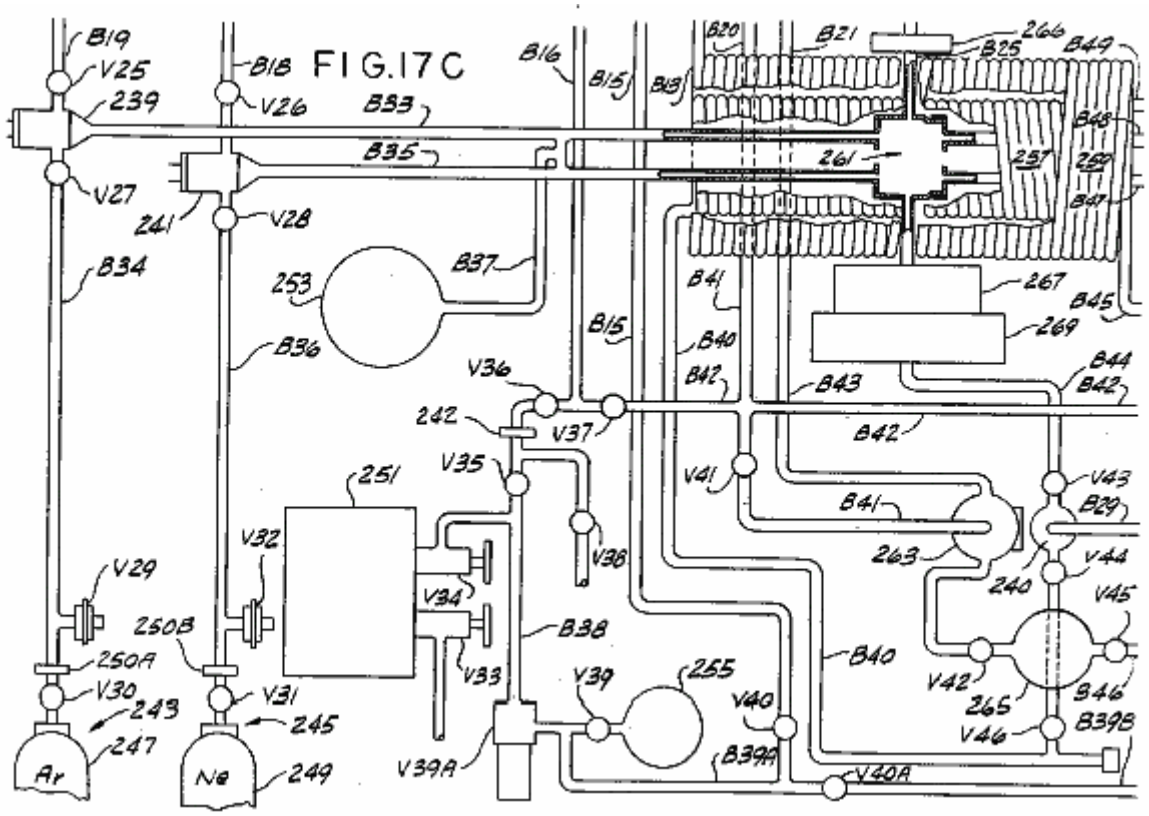
FIG.16



Figs.17A-17D are schematic diagrams of a fuel mixer of the present invention:







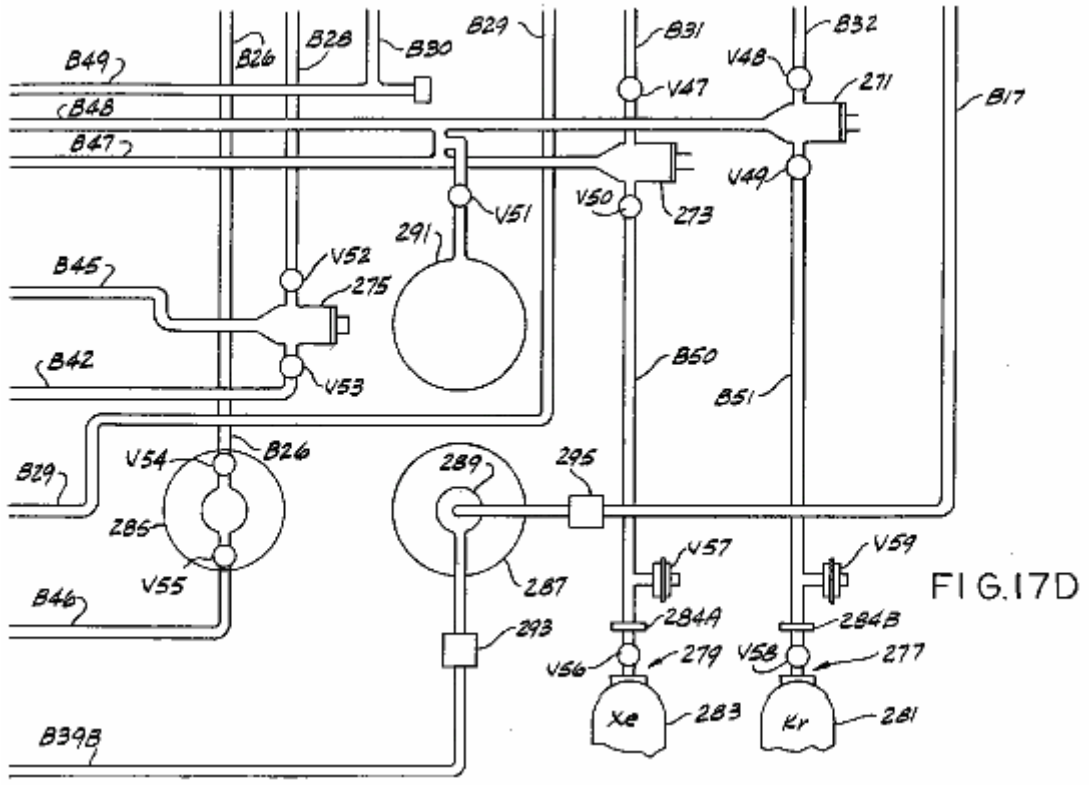
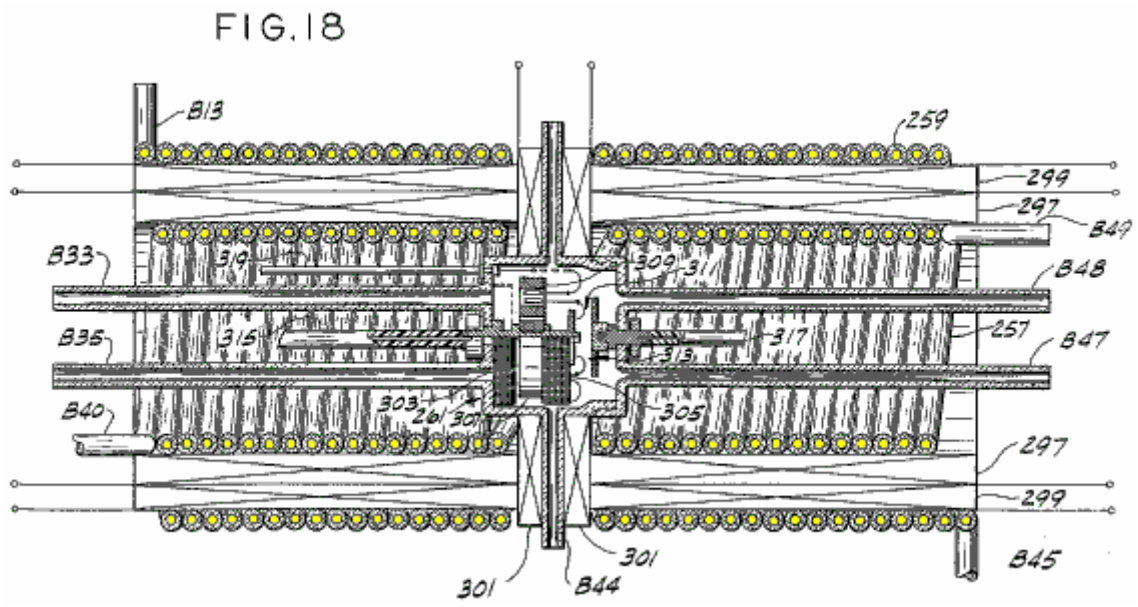


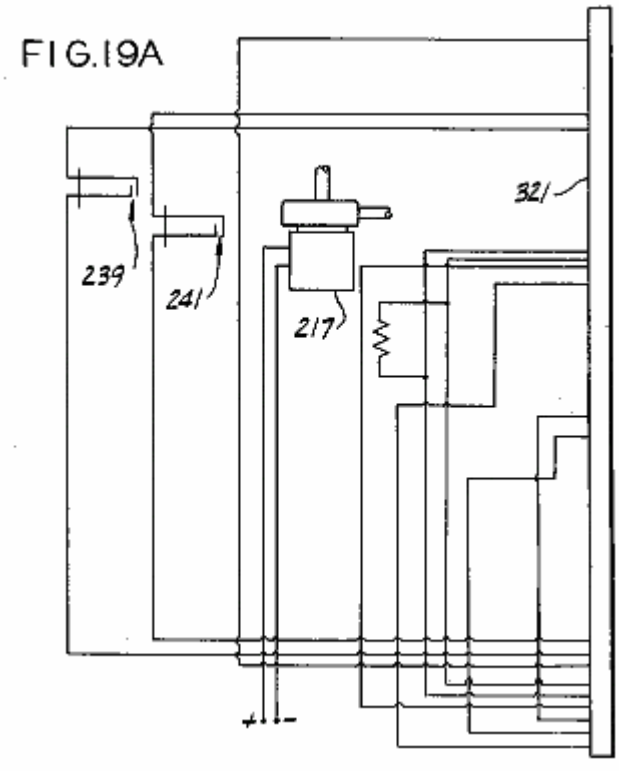
FIG.17C

FIG.17D

Fig.18 is a schematic diagram of the mixing chamber portion of the fuel mixer shown in **Figs.17A-17D**:



Figs.19A-19E are schematic diagrams of a portion of the electrical circuitry of the fuel mixer shown in **Figs.17A-17D**:



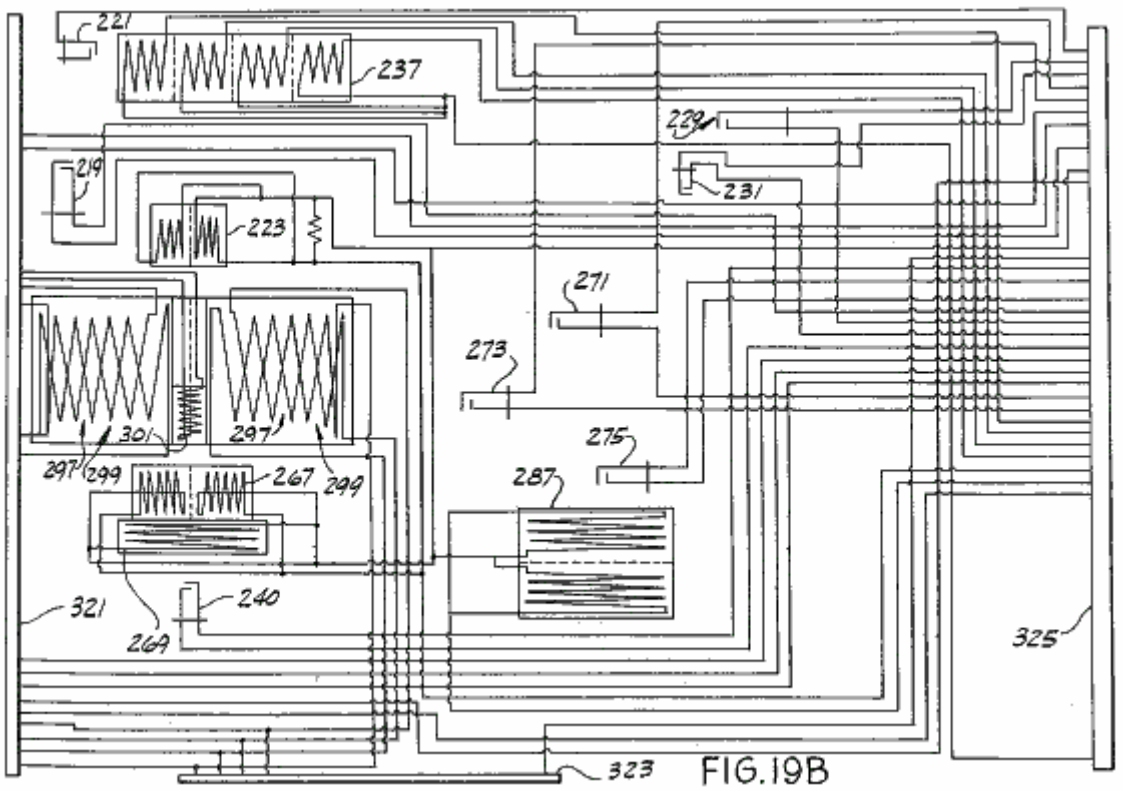


FIG.19B

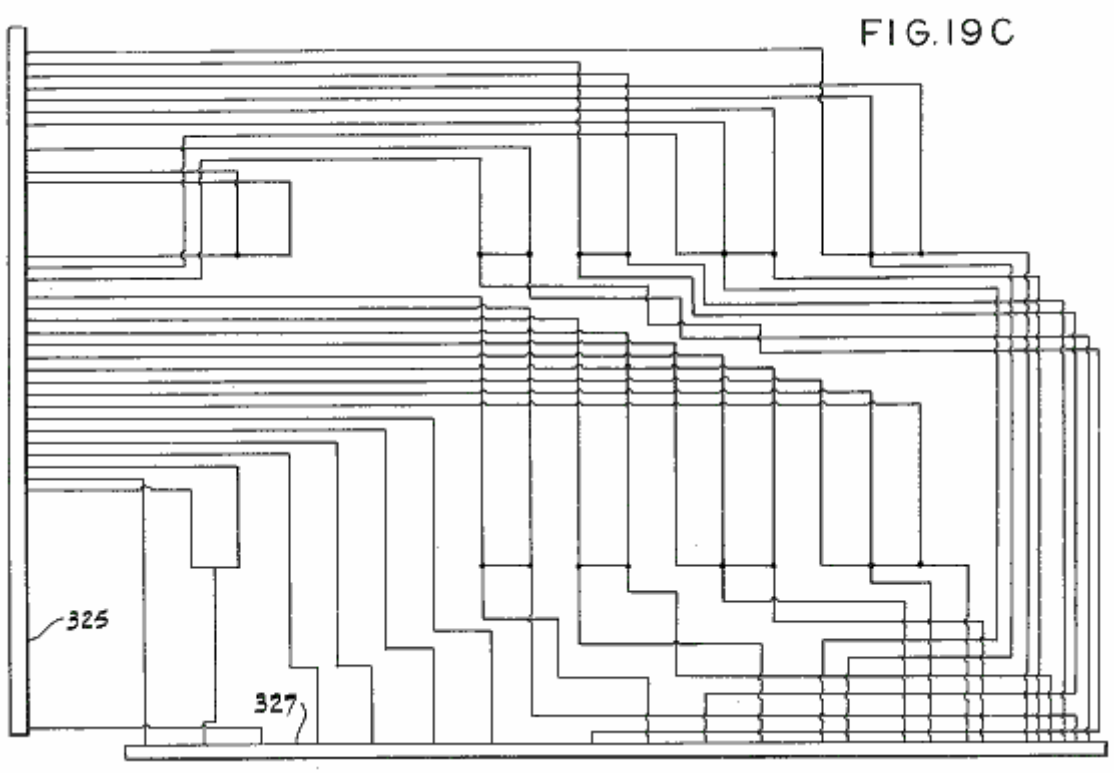


FIG.19C

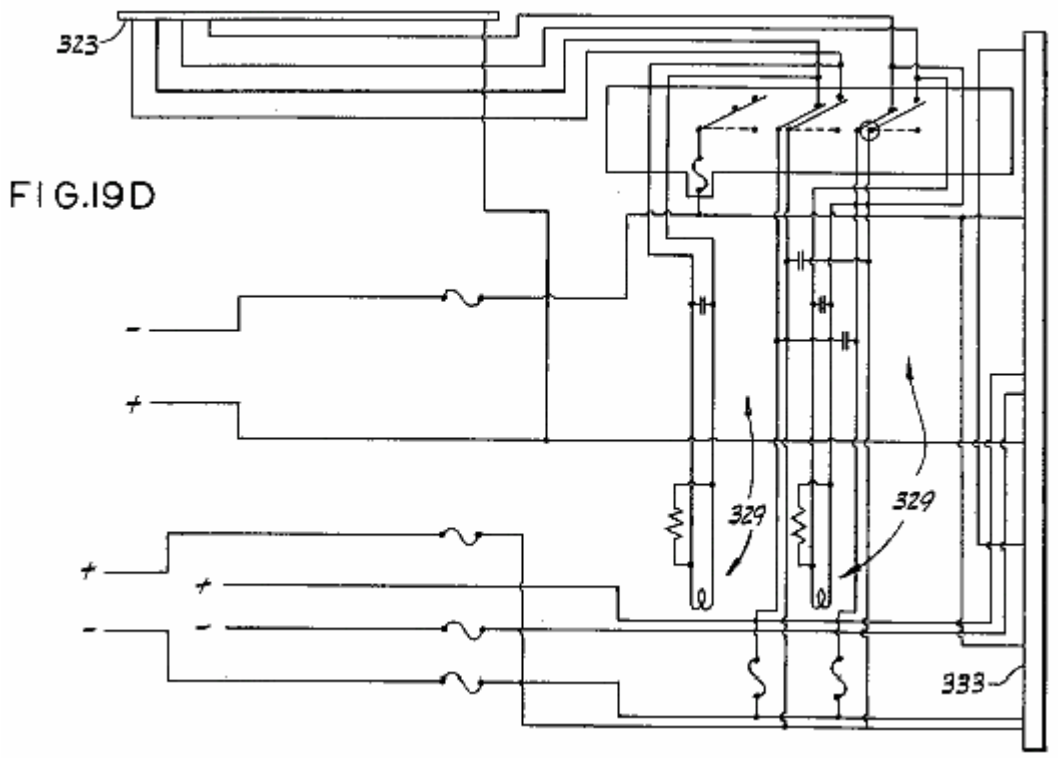


FIG.19D

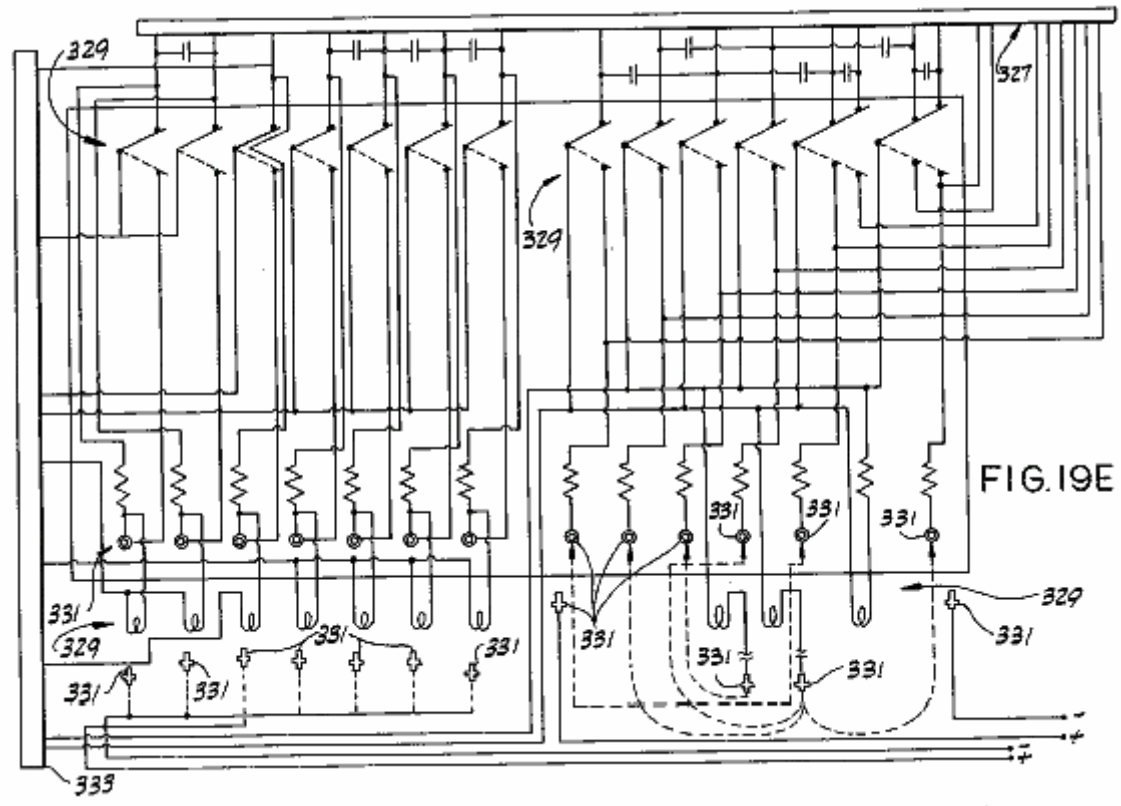


FIG.19E

Figs.20A-20F are schematic diagrams of the rest of the electrical circuitry of the fuel mixer shown in **Figs.17A-17D**:

FIG.20A

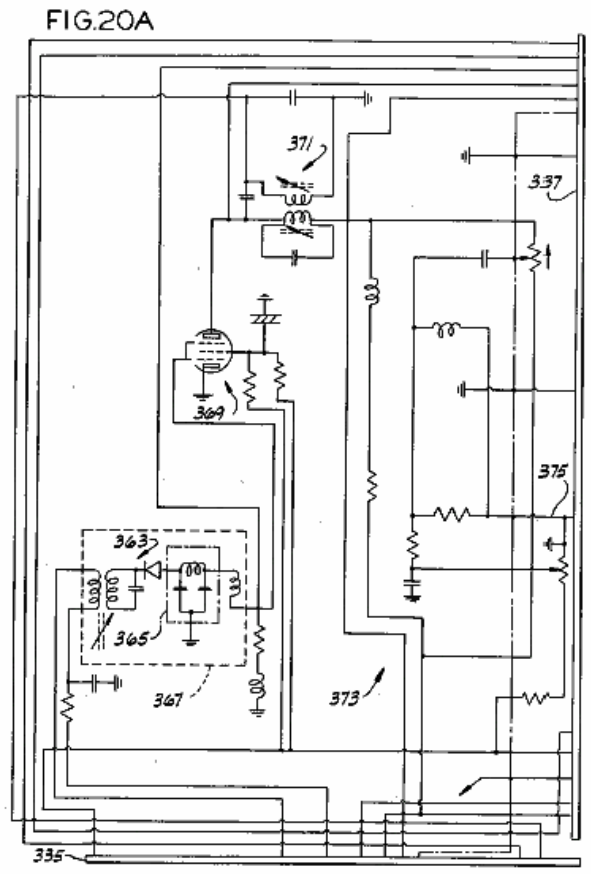


FIG.20B

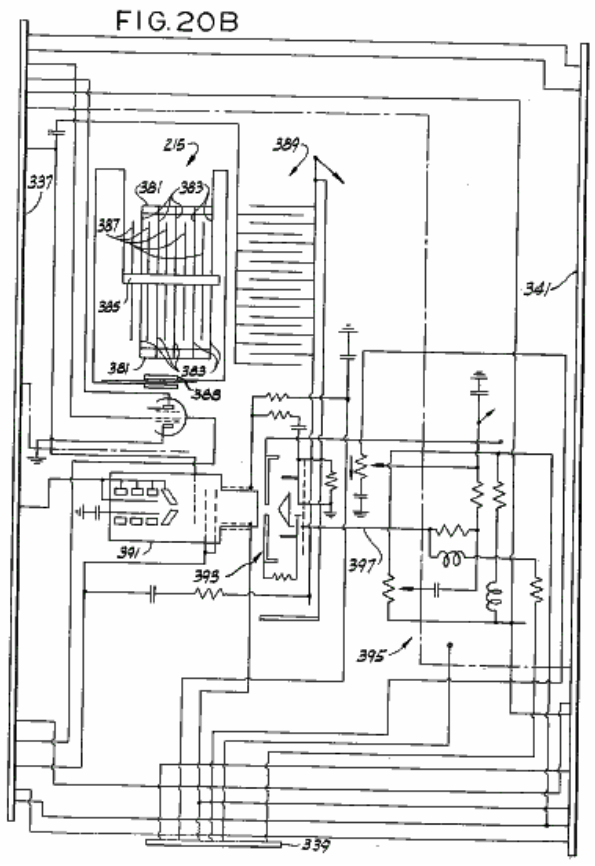
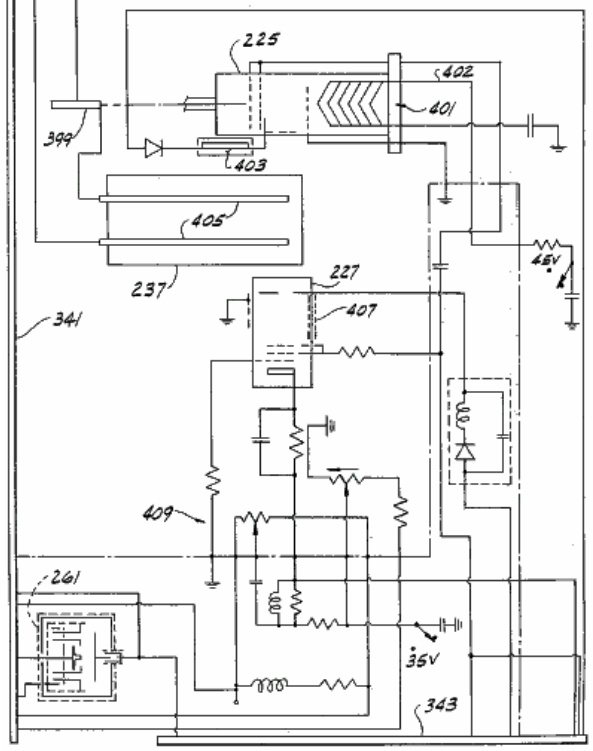


FIG. 20C



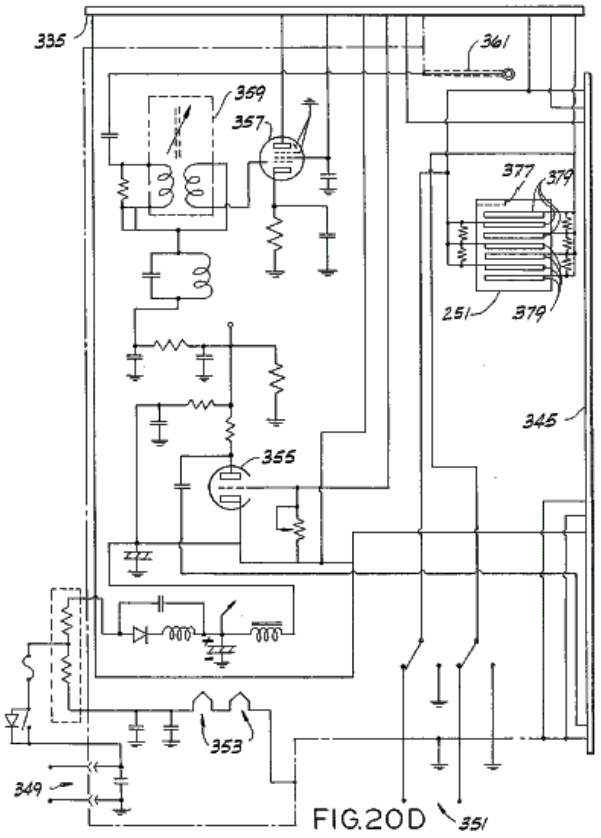
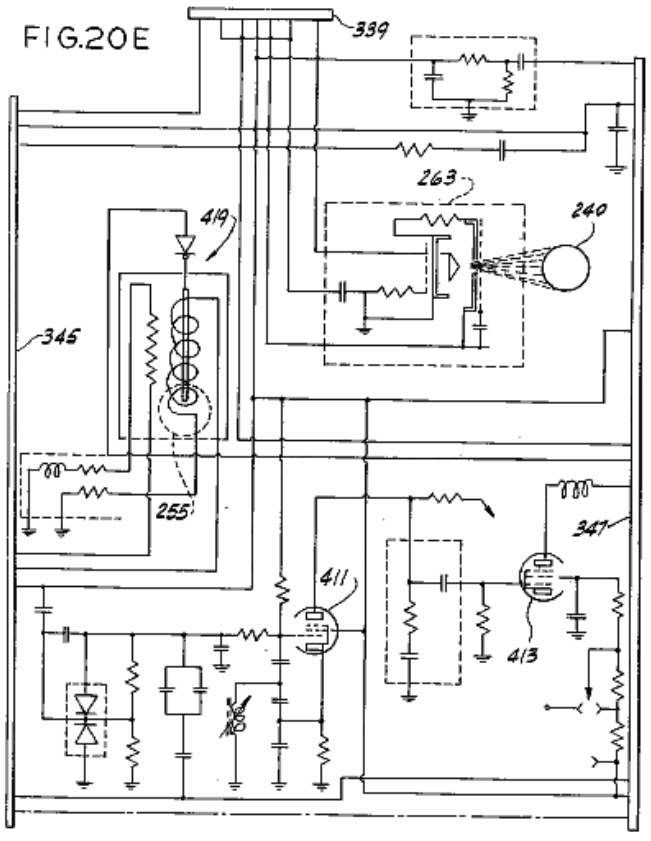


FIG. 20D



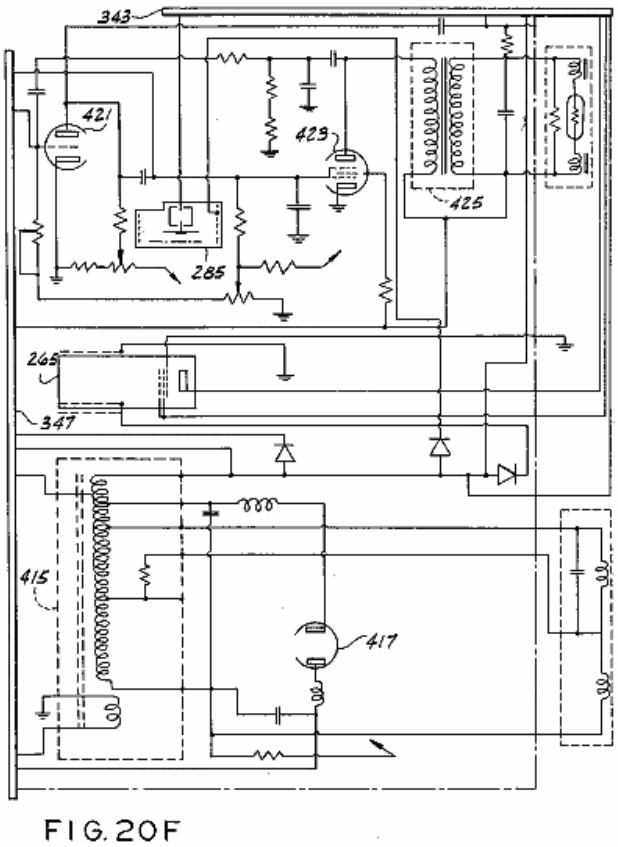
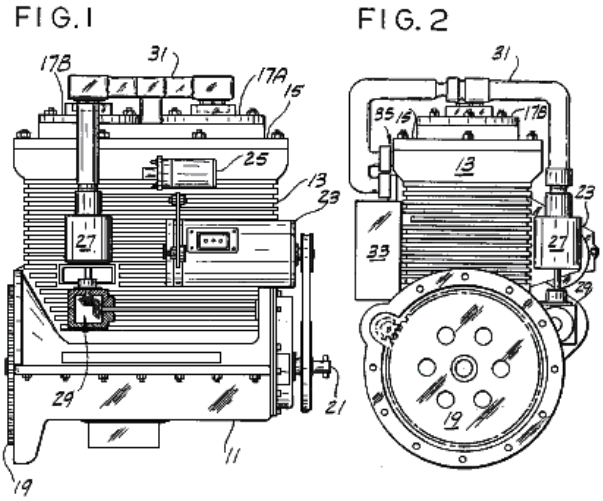


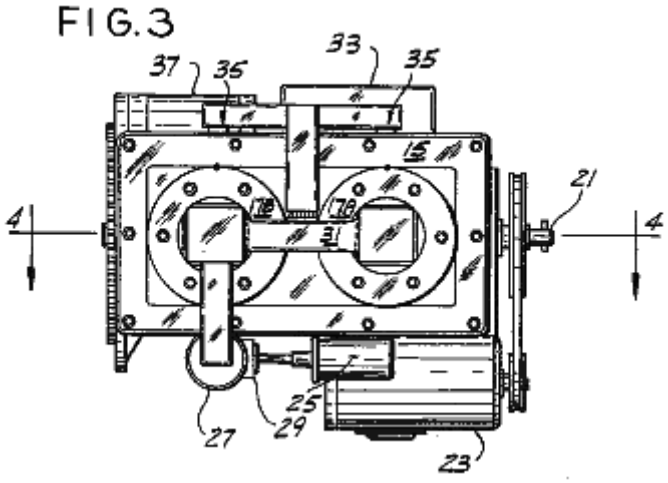
FIG. 20F

Note: Corresponding reference characters indicate corresponding parts throughout all of the views of the drawings.

DESCRIPTION OF A PREFERRED EMBODIMENT

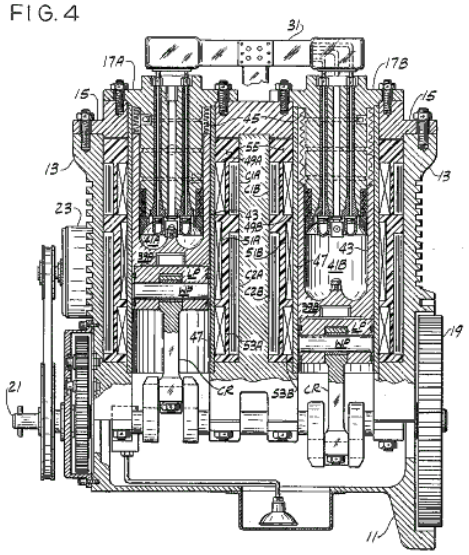


Referring to the drawings, there is shown in **Fig.1** a two-cylinder engine **11** comprising a block **13** preferably of a nonmagnetic material such as aluminium, a nonmagnetic head **15**, and a pair of cylinder heads **17A** and **17B** of a magnetisable material such as 0.1-0.3% carbon steel. Also shown in **Fig.1** is a flywheel **19** attached to a crankshaft **21**, a generator **23**, a high-voltage coil **25**, a distributor **27** attached by a gear arrangement shown in part at **29** to the crankshaft, and an electrical cable **31** which is connected to the distributor and to both cylinders. Cable **31** (see **Fig.2**) is also electrically connected to a switching unit **33** which preferably comprises a plurality of silicon controlled rectifiers (SCRs) or transistors. Also shown in **Fig.2** is a second electrical connection of the cable to the cylinders, which connection is indicated generally at **35**. Turning to **Fig.3**, there is shown a starter motor **37** as well as a clearer view of the connections **35** to each cylinder.



A cross section of the engine is shown in **Fig.4**. The cylinder heads have associated with them, pistons marked **39A** and **39B**, respectively, the heads and pistons define opposite ends of a pair of chambers or cylinders **41A** and **41B** respectively. The pistons are made of a magnetisable material. Although only two chambers are shown, the engine can include any number. It is preferred, however, for reasons set forth below, that there be an even number of cylinders. Pistons **39A** and **39B** move axially with respect to their corresponding heads from a first position (the position of piston **39A** in **Fig.4**) to a second position (the position of piston **39B**) and back, each piston being suitably connected to crankshaft **21**. As shown in **Fig.4**, this suitable connection can include a connecting rod **CR**, a wrist pin **WP**, and a lower piston portion or power piston **LP**. The connecting rods and/or power pistons must be of non-magnetisable material. When a split piston is used, pistons **39A** and **39B** are suitably connected to lower piston portions **LP** by bolting, spring-loaded press fitting, or the like. Pistons **39A** and **39B** are attached 180 degrees apart from each other with respect to the crankshaft so that when one piston is at top dead centre (TDC) the other will be at bottom dead centre (BDC) and vice versa. Additional pairs of cylinders may be added as desired but the pistons of each pair should be attached to the crankshaft 180 degrees from each other. Of course, the relative position of each piston with respect to its respective head determines the volume of its chamber.

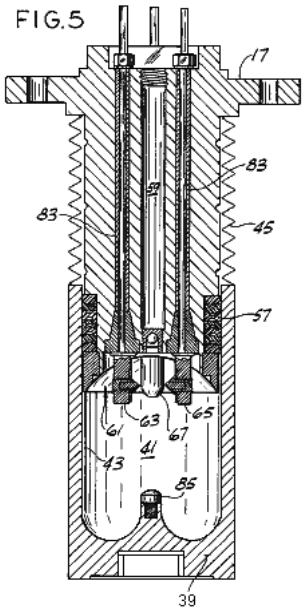
FIG. 4



Integral with the piston bodies are walls **43** which form the walls of the chambers. Preferably, a set of air-tight bellows **45**, of similar construction to that sold under the designation ME 197-0009-001 by the Belfab Company of Daytona Beach, Fla., are suitably secured between walls **43** and cylinder heads **17A** and **17B** respectively to form an airtight seal between each piston and its cylinder head. While walls **43** and piston **39** can be made of one magnetisable piece, a preferable and more efficient construction has walls **43** separate from piston **39** and made of a non-magnetisable material. The length of time that a given engine will run is a function of the efficacy of its sealing system. Means, such as bellows **45**, for hermetically sealing the cylinders will optimise said length of time. Such a hermetic seal should be secured between walls **43** and cylinder heads **17** to form an airtight seal between them. This seal could be the airtight bellows system shown or some other sealing system such as an oil sealing system.

Cylinder bodies **47** (see **Fig.4**), made of nonmagnetic material such as stainless steel, extend from the point of attachment of each bellows to its cylinder head to the base of the corresponding pistons, forming sleeves for each piston in which each piston moves. Three sets of electric coils **49A, 49B, 51A, 51B**, and **53A, 53B**, are wound around sleeves **47**, and hence around chambers **41A** and **41B**, respectively, for generating magnetic fields in the chambers, those coils being generally coaxial with their respective chambers. Each of these coils has an inductance of approximately 100 mH. It is preferred that 14-19 gauge wire be used to wind these coils and that the coils be coated with a suitable coating, such as #9615 hardener from Furane Plastics, Inc., of Los Angeles, California, or the coating sold by the Epoxylite Corp. of South El Monte, California under the trade designation Epoxylite 8683. Each chamber is also surrounded by a pair of capacitors, **C1A**, **C1B** and **C2A, C2B** wound around it, capacitors **C1A**, **C1B** having a capacitance of approximately 1.3 microfarads and capacitors **C2A, C2B** having a capacitance of approximately 2.2 microfarads. The coils and capacitors are potted in hardened epoxy of fibreglass material **55**. The epoxy resin and hardener sold under the designations EPI Bond 121 and #9615 hardener by Furane Plastics, supra, are satisfactory, but other epoxy material which will remain stable at temperatures up to 200 degrees F would probably also be acceptable. It is preferred that a small amount of graphite such as that sold under the trade designation Asbury 225 by Asbury Graphite, Inc. of Rodeo, Calif., be included in the epoxy potting to prevent nuclear particles formed in the chamber from escaping from the apparatus. Ten to 15% graphite to epoxy by weight is more than enough.

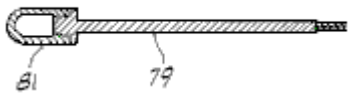
FIG.5



A typical cylinder is shown in section in **Fig.5**, showing the piston in its fully extended position with respect to the head and showing many details on a somewhat larger scale than that of **Fig.4**. A set of seals **57**, made of a material such as that sold under the trade designation Teflon by the DuPont Company of Delaware, is positioned between the cylinder head and wall **43** to prevent escape of the working fluid from chamber **41**. A filler tube **59** with a ball valve at its lower end is used in filling the chamber with the working fluid but is closed during operation of the engine.

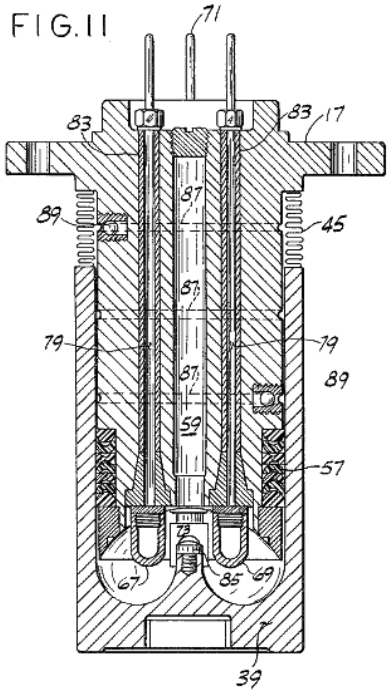
The cylinder head has a generally concave depression therein, indicated at **61**, which defines the top end of the chamber. A plurality of electrodes for exciting and igniting the working fluid extend through the cylinder head into the chamber. Two of those electrodes, shown in section in **Fig.5** and labelled **63** and **65**, have tungsten points **75**, while the other two, labelled **67** and **69** (see **Fig.6** for electrode **69**) are containers called, respectively, the anode and the cathode. The electrodes are generally equidistantly spaced from the axes of their chambers and are generally coplanar to each other, their mutual plane being perpendicular to the axes of their chambers. Each electrode is positioned 90 degrees from adjacent electrodes in this embodiment and are generally positioned so that a line from the anode to the cathode and a line between the other two electrodes intersect at a focal point generally on the axis of the chamber. The radial distance of each electrode from the focal point is fixed for a reason discussed below. The general construction of electrodes **63** and **65** is shown in **Fig.6** to **Fig.9**. These electrodes include a conductive rod **71** (see **Fig.7**) preferably of brass or copper; a conductive, generally rectangular plate **73** (see **Fig.6**, **Fig.8** and **Fig.9**); and tungsten point **75** mounted in a conductive base **77** generally at right angles to the plate (see **Fig.8** and **Fig.9**).

FIG.10



The construction of the anode and cathode is shown in **Fig.10**. Each includes a conductive rod **79** and a container **81**. The cathode container is substantially pure aluminium. If desired, aluminium alloys with, e.g., less than 5% copper, 1% manganese and 2% magnesium may be used. In one embodiment, the cathode container contains approximately four grams of thorium-232 and is filled with argon. In this same embodiment the anode container is copper or brass and contains approximately two grams of rubidium-37 and approximately three grams of phosphorus-15 hermetically sealed in mineral oil. In a second embodiment, the cathode is still aluminium, but it contains at least two grams of rubidium-37 in addition to the approximately four grams of thorium-232 in either argon or mineral oil. In this second embodiment, the anode is also aluminium and contains at least 4 grams of phosphorus-15 and at least 2 grams of thorium-232 in argon or mineral oil. Alternatively, mesothorium may be used for the thorium, strontium-38 may be used for the rubidium, and sulphur-16 may be used for the phosphorus. Rods **71** and **79** extend through cylinder head **17** to the exterior where electrical connections are made to the electrodes. Each rod is surrounded by one of four insulating sleeves **83**, the lower portion of each of which being flared outwards to seat firmly in the cylinder head.

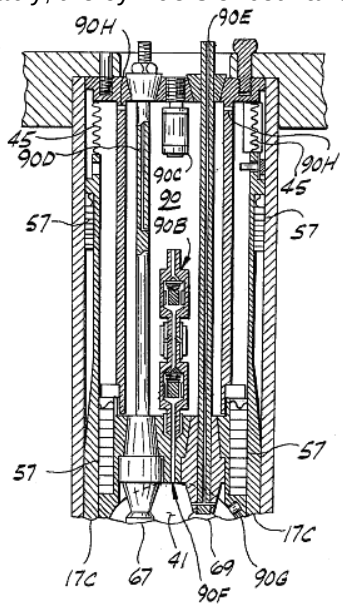
FIG.11



The piston has a generally semi-toroidal depression in its upper surface (see **Fig.4**, **Fig.5** and **Fig.11**) and carries a conductive discharge point **85** of copper, brass or bronze generally along the axis of the chamber. When the piston is generally extended, the discharge point is a substantial distance from the electrodes. But when the piston is in its upper position (see **Fig.11**), the discharge point is positioned generally between all four electrodes and close to them, there being gaps between the electrodes and the discharge point. When the piston is in this upper position, the electrodes extend somewhat into the semi-toroidal depression in the piston's upper surface and the chamber is generally toroidal in shape. The volume of the chamber shown in **Fig.11** can be from approximately 6.0 cubic inches (100 cc) or larger. Given the present state of the art, 1500 cubic inches (25,000 cc) appears to be the upper limit. A plurality of ports **87** and one-way valves **89** return working fluid which escapes from the chamber back into it, so long as a sealing system such as bellows **45** is used.

An alternative cylinder head/piston arrangement is shown in **Fig.12**. The main difference between this arrangement and that of **Fig.5** is that the chamber walls, here labelled **43'** are integrally formed with the head. As a result seals **57** are carried by the piston rather than by the head, the attachment of bellows **45** is somewhat different, and the fluid-returning valves and ports are part of the piston rather than of the head. Otherwise these arrangements are substantially the same. Preferably, the cylinders of both arrangements are hermetically sealed.

FIG.12A



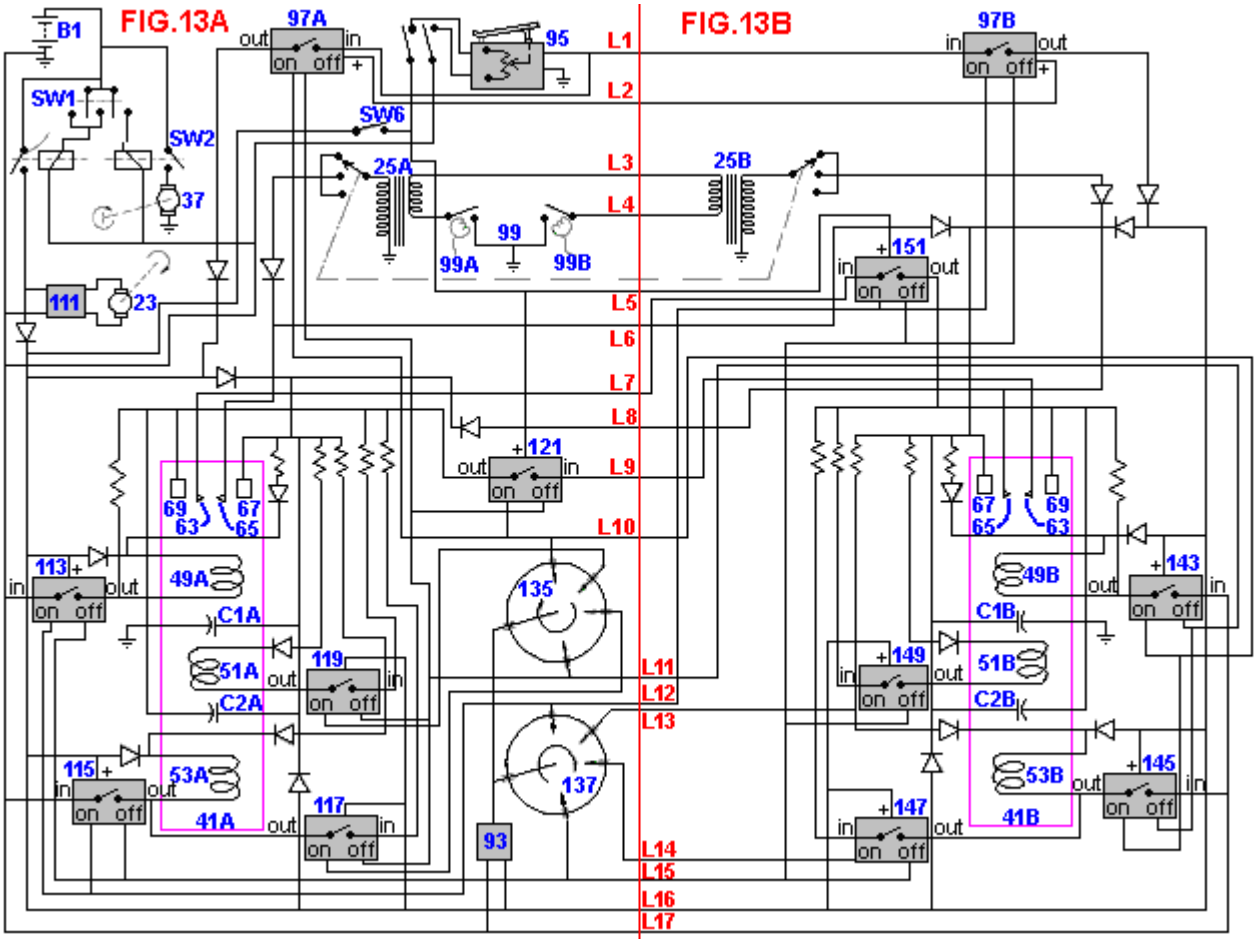
An additional embodiment of a cylinder head/piston arrangement used in the present invention is shown in **Fig.12A**. In this arrangement, a tapered sleeve **17C** mates between cylinder head **17** and piston **39**, a plurality of seals **57** are provided, and electrodes **67** and **69** have a somewhat different shape. Also, in this embodiment, a chamber **90** is provided in cylinder head **17** for storing additional working fluid, i.e., the purpose of chamber **90** is to extend the operating time between refuelling by circulating the working fluid, viz. the mixture of inert gases

described, between cylinder **41** and chamber **90** as needed so that the reactions in cylinder 41 are not adversely affected. To accomplish this, this embodiment further includes a two-way circulation valve **90B**, a relief valve **90C**, and duct or passageway **90D** for evacuating and filling chamber **90**, a duct or passageway **90E** for evacuating and filling cylinder **41**, a passageway **90F** between chamber **90** and cylinder **41** in which two-way valve **90B** is disposed, a sensor **90G** and a plurality of small pressure relief holes **90H**. Relief holes **90H** serve to relieve the pressure on bellows **45** as the piston moves from BDC to TDC.

In larger engines holes **90H** should be replaced with one way valves. Two-way valve **90B** is either controlled by sensor **90G** or is manually operated, as desired, to allow the circulation of gases between chamber **90** and cylinder **41**. The sensor itself detects a condition requiring the opening or closing of valve **90B** and signals that condition to the valve. For example, sensor **90G** can measure pressure in cylinder **41** while the piston is at top dead centre. A predetermined cylinder pressure can cause a spring to compress, causing the valve to open or close as appropriate. A subsequent change in the cylinder pressure would then cause another change in the valve. Another sensor (not shown) could measure the physical location of the piston by a physical trip switch or an electric eye, or it could measure angular distance from top dead centre on the distributor or the crankshaft. The sensor must keep the gas pressure in chamber **90** at one atmosphere, plus or minus 5%, and at top dead centre, cylinder **41** should also be at that pressure. If gas is lost from the system, it is more important to maintain the proper pressure in cylinder **41**. Alternatively, a small passage between cylinder **41** and chamber **90** could function in a passive manner to satisfactorily accomplish the same result. From the above, it can be seen that this embodiment utilises the hollowed out centre of the cylinder head for storing additional working fluid, which fluid is circulated between chamber **90** and cylinder **41** through a valve system comprising valve **90B** and sensor **90G** with the moving piston causing the gases to circulate.

The electrical circuitry for engine **11** includes (see **Fig.13A**) a 24 V battery **B1**, an ignition switch **SW1**, a starter switch **SW2**, starter motor **37**, a main circuit switch **SW4**, a step-down transformer **93** (e.g., a 24 V to 3.5 V transformer), a switch **SW6** for supplying power to ignition coil **25** (shown in **Fig.13A** and **Fig.13B** as two separate ignition coils **25A** and **25B**), and various decoupling diodes.

The circuitry of **Fig.13A** also includes a high frequency voltage source or oscillator **95** for supplying rapidly varying voltage through two electronic current regulators **97A, 97B** (see **Fig.13B** for regulator **97B**) to the anode and cathode electrodes of each cylinder, and a high-voltage distributor **99** for distributing 40,000 volt pulses to the cylinders. Distributor **99** has two wipers **99A** and **99B** and supplies three pulses to each cylinder per cycle. Wipers **99A** and **99B** are 180 degrees out of phase with each other and each operates to supply pulses to its respective cylinder from TDC to 120 degrees thereafter. More pulses are desirable and therefore a better distributor arrangement (shown in **Fig.14**) may be used. The arrangement shown in **Fig.14** includes two ignition coils **101**, **103**, a simple distributor **105** and a pair of magnetic ignition circuits **107** and **109**, described below. Of course many other ignition systems could also be developed. For example, a single circuit might be used in place of circuits **107**, **109**, additional induction coils might be added to the ignition coils to assist in starting or a resistor could be added to the ignition coils to ensure a constant 40,000 volt output regardless of engine rpm. Also, a solid-state distributor could be used instead of the mechanical distributor labelled **99**.



Referring back to **Fig.13A**, for engines of more than 1000 hp a high frequency source **95** could be used to control engine RPM. The output frequency is controlled by a foot pedal similar to an accelerator pedal in a conventional vehicle. The output frequency varies through a range of from approximately 2.057 MHz to approximately 27.120 MHz with an output current of approximately 8.4 amps. The speed of engine **11** is controlled by the output frequency of source **95**. The high frequency current, as described below, is directed to each cylinder in turn by circuitry described below. For engines producing from 300 to 1000 hp (not shown), a high frequency source having a constant output of 27.120 MHz with a constant current of 3.4 amps which is continually supplied to all cylinders could be used. In this case an autotransformer, such as that sold under the trade designation Variac by the General Radio Company, controlled by a foot pedal varies the voltage to each cylinder from 5 to 24 volts DC at 4.5 amps, using power from the batteries or the alternator. The DC current from the Variac is switched from cylinder to cylinder by two small electronic switching units which in turn are controlled by larger electronic switching units. For the smallest engines (not shown), a high frequency generator could supply a constant output of 27.120 MHz with a constant current of 4.2 amps to the cylinders during starting only. Speed control would be achieved by a Variac as described above which controls the DC voltage supplied to the cylinders in turn within a range of from 5 to 24 volts at a current of 5.2 amps. In this case, once the engine is running, the full voltage needed to ignite the (smaller) quantity of gases is obtained from the electrodes in the other cylinder of the pair.

The circuitry of **Fig.13A** also includes the generator, a voltage regulator and relay **111**, five electronic switching units **113, 115, 117, 119** and **121**, electrodes **63** and **65** associated with chamber **41A** (hereinafter chamber **41A** is sometimes referred to as the "**A**" cylinder and chamber **41B** is sometimes referred to as the "**B**" cylinder), anode **67**, cathode **69**, magnetic coils **49A, 51A** and **53A**, capacitors **C1A** and **C2A**, and various decoupling diodes. The electronic switching units can take a variety of forms. For example, one simple form (see **Fig.15**) includes a pair of SCRs **123** and **125**. The switching unit is connected at terminal IN to the corresponding line on the input side and at terminal OUT to the corresponding line on the output side. When a voltage of 3.5 volts is supplied from the battery through a distributor, for example, to the ON terminal, SCR **125** conducts, thereby completing a circuit through the switching unit. Conversely, when 3.5 volts is applied to the OFF terminal, SCR **123** conducts and the circuit is broken. Likewise, the circuit for regulators **97A** and **97B** (see **Fig.16**) includes two SCRs **127** and **129** and a PNP transistor **131**. In this circuit when SCR **127** is gated on, it forces transistor **131** into conduction, thereby completing the circuit through the regulator. When SCR **129** is gated on, the circuit through transistor **131** is broken. A number of other configurations may be used in place of those of **Fig.15** and **Fig.16** and not all would use SCRs. For example, one triode could be used to replace two main SCRs, or transistors could be used instead of SCRs.

A pair of low-voltage distributors **135** and **137** are also shown in **Fig.13A**. Distributors **135** and **137** provide gating pulses for the electronic switching units of **Fig.13A** and **Fig.13B**. Of course, solid-state distributors could also replace mechanical distributors **135** and **137**.

In addition, the engine circuitry includes (see **Fig.13B**) five electronic switching units **143, 145, 147, 149** and **151** corresponding to units **113, 115, 117, 119** and **121** of **Fig.13A**, electrodes **63** and **65** of the "**B**" cylinder, anode **67**, cathode **69**, electric coils **49B, 51B** and **53B**, capacitors **C1B** and **C2B**, and various decoupling diodes. The circuitry of **Fig.13B** is generally the same as the corresponding portions of **Fig.13A**, so the description of one for the most part applies to both. Of course, if more than two cylinders are used, each pair of cylinders would have associated with them, circuitry such as that shown in **Fig.13A** and **Fig.13B**. The circuitry of **Fig.13A** is connected to that of **Fig.13B** by the lines **L1-L17**.

The working fluid and the fuel for the engine are one and the same and consist of a mixture of inert gases, which mixture consists essentially of helium, neon, argon, krypton and xenon. It is preferred that the mixture contain 35.6% helium, 26.3% neon, 16.9% argon, 12.7% krypton, and 8.5% xenon by volume, it having been calculated that this particular mixture gives the maximum operation time without refuelling. Generally, the initial mixture may contain, by volume, approximately 36% helium, approximately 26% neon, approximately 17% argon, approximately 13% krypton, and approximately 8% xenon. This mixture results from a calculation that equalises the total charge for each of the gases used after compensating for the fact that one inert gas, viz. radon, is not used. The foregoing is confirmed by a spectroscopic flashing, described below, that occurs during the mixing process. If one of the gases in the mixture has less than the prescribed percentage, it will become over-excited. Similarly, if one of the gases has more than the prescribed percentage, that gas will be under-excited. These percentages do not vary with the size of the cylinder.

Operation of the engine is as follows: At room temperature, each cylinder is filled with a one atmosphere charge of the fuel mixture of approximately 6 cubic inches (100 cc) /cylinder (in the case of the smallest engine) by means of filler tube **59**. The filler tubes are then plugged and the cylinders are installed in the engine as shown in **Fig.4**, one piston being in the fully extended position and the other being in the fully retracted position. To start the engine, the ignition and starter switches are closed, as is switch **SW6**. This causes the starter motor to crank the engine, which in turn causes the wiper arms of the distributors to rotate. The starting process begins, for example, when the pistons are in the positions shown in **Fig.4**. Ignition coil **25** and distributor **99** (see **Fig.13A**) generate a 40,000 volt pulse which is supplied to electrode **65** of chamber **41A**. Therefore, a momentary high potential exists between electrodes **63** and **65** and the plates on each. The discharge point on piston **39A** is adjacent these electrodes at this time and sparks occur between one or more of the electrodes and the discharge point to partially excite, e.g. ionise, the gaseous fuel mixture.

The gaseous fuel mixture in cylinder **41A** is further excited by magnetic fields set up in the chamber by coil **49A**. This coil is connected to the output side of electronic switching unit **121** and, through switching unit **113**, to the battery and the generator. At this time, i.e., between approximately 5 degrees before TDC and TDC, distributor **135** is supplying a gating signal to unit **121**. Any current present on the input side of unit **121**, therefore, passes through unit **121** to energise coil **49A**. Moreover, high frequency current from oscillator **95** is supplied via regulator **97A** to coil **49A**. This current passes through regulator and relay **97A** because the gating signal supplied from distributor **135** to unit **121** is also supplied to relay **97A**. The current from switching unit **121** and from oscillator **95** also is supplied to the anode and the cathode. It is calculated that this causes radioactive rays (x-rays) to flow between the anode and the cathode, thereby further exciting the gaseous mixture.

As the starter motor continues cranking, piston **39A** begins moving downward, piston **39B** begins moving upward, and the wiper arms of the distributors rotate. (Needless to say, a solid-state distributor would not rotate. The distributor could utilise photo cells, either light or reflected light, rather than contact points). After 45 degrees of rotation, distributor **135** supplies a gating pulse to electronic switching unit **119**, thereby completing a circuit through unit **119**. The input to unit **119** is connected to the same lines that supply current to coil **49A**. The completion of the circuit through unit **119**, therefore, causes coil **51A** to be energised in the same manner as coil **49A**. After an additional 45 degrees of rotation, distributor **135** gates on electronic switching unit **117** which completes a circuit to the same lines. The output terminal of unit **117** is connected to coil **53A**, and so this coil is energised when unit **117** is gated on. All three coils of the "**A**" cylinder remain energised and, therefore, generating magnetic fields in chamber **41A** until piston **39A** reaches BDC.

As piston **39A** moves from TDC to BDC, two additional 40,000 volt pulses (for a total of three) are supplied from distributor **99** to the "**A**" cylinder. These pulses are spaced approximately 60 degrees apart. If more pulses are desired, the apparatus shown in **Fig.14** may be used. In that case, the solenoids indicated generally at **107A, 107B** and **109A, 109B** are energised to create a number of rapid, high-voltage pulses which are supplied as indicated in **Fig.14** to the cylinders, distributor **105** operating to supply pulses to only one of the pair of cylinders at a time.

As piston **39A** reaches BDC, distributor **135** sends a pulse to the OFF terminals of electronic switching units **121, 117** and **119**, respectively, causing all three coils **49A, 51A** and **53A** to be de-energised. At about the same time, i.e., between approximately 5 degrees before TDC and TDC for piston **39B**, distributor **137** supplies a gating pulse to the ON terminals of electronic switching units **113** and **115**. The power inputs to units **113** and **115** come from the generator through regulator **111** and from the battery, and the outputs are directly connected to coils **49A** and **53A**. Therefore, when units **113** and **115** are gated on, coils **49A** and **53A** are reenergised. But in this part of the cycle, the coils are energised with the opposite polarity, causing a reversal in the magnetic field in chamber **41A**. Note that coil **51A** is not energised at all during this portion of the cycle. Capacitors **C1A** and **C2A** are also charged during the BDC to TDC portion of the cycle. (During the TDC to BDC portion of the cycle, these capacitors are charged and/or discharged by the same currents as are supplied to the anode and cathode since they are directly connected to them).

As piston **39A** moves upwards, electrodes **63** and **65** serve as pick-up points in order to conduct some of the current out of chamber **41A**, this current being generated by the excited gases in the chamber. This current is transferred via line **L7** to electronic switching unit **151**. The same gating pulse which gated on units **113** and **115** was also supplied from distributor **137** via line **L12** to gate on switching unit **151**, so the current from the electrodes of chamber **41A** passes through unit **151** to the anode, cathode and capacitors of chamber **41B**, as well as through switching units **147** and **149** to coils **49B, 51B** and **53B**. Thus it can be seen that electricity generated in one cylinder during a portion of the cycle is transferred to the other cylinder to assist in the excitation of the gaseous mixture in the latter. Note that this electricity is regulated to maintain a constant in-engine current. It should be noted, that twenty four volts from the generator is always present on electrodes **63** and **65** during operation to provide for pre-excitement of the gases.

From the above it can be seen that distributors **135** and **137** in conjunction with electronic switching units **113, 115, 117, 119, 121, 143, 145, 147, 149** and **151** constitute the means for individually energising coils **49A**, **49B, 51A, 51B, 53A** and **53B**. More particularly, they constitute the means to energise all the coils of a given cylinder from the other cylinder when the first cylinder's piston is moving from TDC to BDC and operate to energise only two (i.e., less than all) of the coils from the alternator when that piston is moving from BDC to TDC. Additionally, these components constitute the means for energising the coils with a given polarity when the piston of that cylinder is moving from TDC to BDC and for energising the first and third coils with the opposite polarity when that piston is moving from BDC to TDC.

As can also be seen, switching units **121** and **151** together with distributors **135** and **137** constitute the means for closing a circuit for flow of current from chamber **41A** to chamber **41B** during the BDC to TDC portion of the cycle of chamber **41A** and for closing a circuit for flow of current from chamber **41B** to chamber **41A** during the TDC to BDC portion of the cycle of chamber **41A**. Oscillator **95** constitutes the means for supplying a time varying electrical voltage to the electrodes of each cylinder, and oscillator **95**, distributors **135** and **137**, and regulators **97A** and **97B** together constitute the means for supplying the time varying voltage during a predetermined portion of the cycle of each piston. Moreover, distributor **99** together with ignition coils **25A** and **25B** constitute the means for supplying high-voltage pulses to the cylinders at predetermined times during the cycle of each piston.

The cycle of piston **39B** is exactly the same as that of piston **39A** except for the 180 degree phase difference. For each cylinder, it is calculated that the excitation as described above causes the gases to separate into layers, the lowest atomic weight gas in the mixture, namely helium, being disposed generally in the centre of each chamber, neon forming the next layer, and so on until we reach xenon which is in physical contact with the chamber walls. The input current (power) to do this is the calculated potential of the gas mixture. Since helium is located in the centre of the chamber, the focal point of the electrode discharges and the discharges between the anode and cathode is in the helium layer when the piston is near TDC. As the piston moves slightly below TDC, the electrons from electrodes **63** and **65** will no longer strike the tip of the piston, but rather will intersect in the centre of the cylinder (this is called "focal point electron and particle collision") as will the alpha, beta and gamma rays from the anode and cathode. Of course, the helium is in this exact spot and is heavily ionised at that time. Thus the electrodes together with the source of electrical power connected thereto constitute the means for ionising the inert gas.

It is calculated that as a result of all the aforementioned interactions, an ignition discharge occurs in which the helium splits into hydrogen in a volume not larger than 2 or 3 x 10^{-6} cubic millimetres at a temperature of approximately 100,000,000 degrees F. Of course this temperature is confined to a very small space and the layering of the gases insulates the cylinder walls from it. Such heat excites the adjacent helium so that a plasma occurs. Consequently, there is a minute fusion reaction in the helium consisting of the energy conversion of a single helium atom, which releases sufficient energy to drive the piston in that chamber toward BDC with a force similar in magnitude to that generated in a cylinder of a conventional internal combustion engine. Electrodes **63** and **65** extend into the argon layer while each piston is in its BDC to TDC stroke so as to pick up some of the

current flowing in that layer. It may take a cycle or two for the gases in the cylinders to become sufficiently excited for ignition to occur.

Once ignition does occur, the electrical operation of the engine continues as before, without the operation of the starter motor. Distributor **99** supplies three pulses per cycle (or more if the magnetic ignition system of **Fig.14** is used) to each cylinder; and distributors **135** and **137** continue to supply "on" and "off" gating pulses to the electronic switching units. The rpm of the engine is, as explained above, governed by the frequency of the current from oscillator **95** (or in the case of smaller horsepower units, by the DC voltage supplied to the cylinders from the Variac).

Because of the minute amount of fuel consumed in each cycle, it is calculated that a cylinder can run at 1200 rpm approximately 1000 hours, if not more, on a single charge of gas. Note that even at 1200 rpm, there will be intense heat occurring only 0.002% of the time. This means that input power need be applied only sporadically. This power can be supplied to a cylinder from the other cylinder of its pair by means of electronic switching units which, in the case of SCRs, are themselves triggered by low voltage (e.g. 3.5 V) current. Thus, since electrical power generated in one cylinder is used to excite the gases in the other cylinder of a pair, it is practical that the cylinders be paired as discussed above. Capacitors are, of course, used to store such energy for use during the proper portion of the cycle of each cylinder.

From the above, it should be appreciated that the engine of this invention has several advantages over presently proposed fusion reactors, such as smaller size, lower energy requirements, etc. But what are the bases of these advantages? For one, presently proposed fusion reactors use hydrogen and its isotopes as a fuel instead of inert gases. Presumably this is because hydrogen requires less excitement power. While this is true, the input power that is required in order to make hydrogen reactors operate makes the excitation power almost insignificant. For example, to keep a hydrogen reactor from short circuiting, the hydrogen gas has to be separated from the reactor walls while it is in the plasma state. This separation is accomplished by the maintenance of a near vacuum in the reactor and by the concentration of the gas in the centre of the reactor (typically a toroid) by a continuous, intense magnetic field. Accordingly, separation requires a large amount of input energy.

In the present invention, on the other hand, the greater excitation energy of the fuel is more than compensated for by the fact that the input energy for operation can be minimised by manipulation of the unique characteristics of the inert gases. First, helium is the inert gas used for fusion in the present invention. The helium is primarily isolated from the walls of the container by the layering of the other inert gases, which layering is caused by the different excitation potential (because of the different atomic weights) of the different inert gases, said excitation being caused by the action of the electrodes, anode and cathode in a magnetic field. This excitation causes the gases each to be excited in inverse proportion to their atomic numbers, the lighter gases being excited correspondingly more. Helium, therefore, forms the central core with the other four gases forming layers, in order, around the helium. The helium is secondarily isolated from the walls of the container by a modest vacuum (in comparison to the vacuum in hydrogen reactors) which is caused partially by the "choking" effect of the coils and partially by the enlargement of the combustion chamber as the piston moves from TDC to BDC. (Unexcited, the gases are at one atmosphere at TDC). Second, argon, the middle gas of the five, is a good electrical conductor and becomes an excellent conductor when (as explained below) it is polarised during the mixing process. By placing the electrodes such that they are in the argon layer, electrical energy can be tapped from one cylinder for use in the other. During a piston's movement from BDC to TDC, the gases are caused to circulate in the cylinder by the change in the polarity of the coils, which occurs at BDC.

During such circulation, the gases remain layered, causing the argon atoms to be relatively close to each other, thereby optimising the conductivity of the argon. This conductivity optimisation is further enhanced by a mild choking effect that is due to the magnetic fields. The circulation of the highly conductive argon results in a continuous cutting of the magnetic lines of force so that the current flows through the electrodes. This production of electricity is similar to the rotating copper wire cutting the magnetic lines of force in a conventional generator except that the rotating copper wire is replaced by the rotating, highly conductive argon. The amount of electricity that can be produced in this manner is a function of how many magnetic field lines are available to be cut. If one of the coils, or all three of the coils or two adjacent coils were energised, there would be only one field with electricity produced at each end. By energising the top and the bottom coil, two separate fields are produced, with electricity produced at four points.

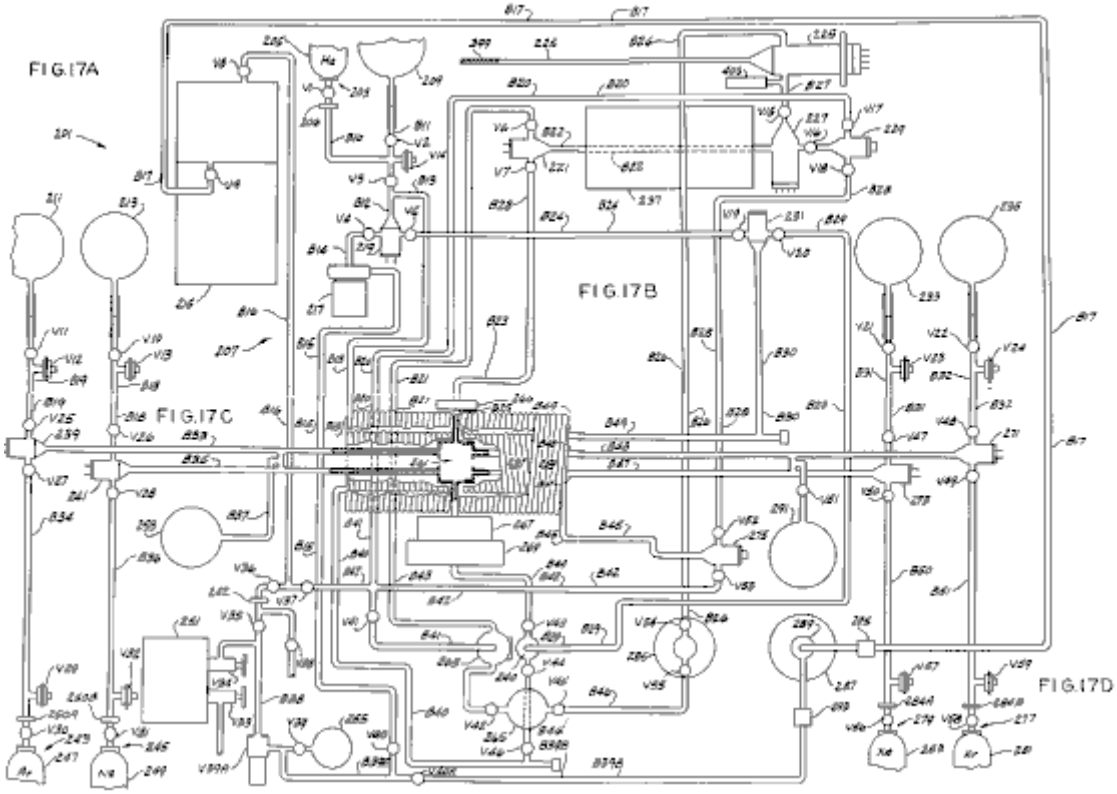
A five coil system, if there were sufficient space, would produce three fields with the top, bottom and middle coils energised. Six points for electricity production would result. The number of coils that can be installed on a given cylinder is a function of space limitations. The recombination of gas atoms during the BDC to TDC phase causes the radiation of electrical energy which also provides a minor portion of the electricity that the electrode picks up. Additional non-grounded electrodes in each cylinder would result in more electricity being tapped off. It should be noted that during the BDC to TDC phase, the anode and the cathode are also in the argon layer and, like the electrodes, they pick up electricity, which charges the capacitors around the cylinder. Third, inert gases remain a

mixture and do not combine because of the completeness of the electron shells. They are therefore well suited to a cycle whereby they are continually organised and reorganised. Fourth, as the helium atoms are consumed, the other gases have the capacity to absorb the charge of the consumed gas so that the total charge of the mixture remains the same.

The second basis of these advantages of the present engine over proposed fusion reactors concerns the fact that hydrogen reactors develop heat which generates steam to turn turbines in order to generate electrical power. This requires tremendous input energy on a continuous basis. The present invention operates on a closed cycle, utilising pistons and a crankshaft which does not require a continuous plasma but rather an infrequent, short duration (10^{-6} second) plasma that therefore requires much less input energy. In the present invention, a plasma lasting longer than 10^{-6} second is not necessary because sufficient pressure is generated in that time to turn the engine. A plasma of longer duration could damage the engine if the heat were sufficiently intense to be transmitted through the inert gas layers to the cylinder walls. A similar heat build-up in the engine can occur if the repetition rate is increased. Such an increase can be used to increase the horsepower per engine size but at the cost of adding a cooling system, using more expensive engine components, and increasing fuel consumption. Note that even though layers of inert gases insulate the cylinder walls, there might be some slight increase in the temperature of the gas layers after a number of cycles, i.e., after a number of ignitions.

Whereas hydrogen fusion reactors cannot directly produce power by driving a piston (because of the required vacuum), the present invention uses the layered inert gases to transmit the power from the plasma to each gas in turn until the power is applied to a piston, which can easily be translated into rotary motion. The layered gases also cushion the piston from the full force of the ignition. Moreover, the fields inside the cylinder undergoing expansion cause the gases to shrink, thereby taking up some of the pressure generated by the explosion and preventing rupturing of the cylinder walls.

Turning now to **Fig.17A** to **Fig.17D**, there is shown apparatus **201** for preparing the fuel mixture for engine **11**. For convenience apparatus **201** is called a mixer although it should be understood that the apparatus not only mixes the gases which form the fuel but also performs many other vital functions as well. The five constituent inert gases are introduced in precise, predetermined proportions. The mixer extracts, filters and neutralises the non-inert gases and other contaminants which may be found in the gas mixture. It also increases the potential capacity of gas atoms, discharges the krypton and xenon gases, polarises the argon gases, ionises the gases in a manner such that the ionisation is maintained until the gas has been utilised and otherwise prepares them for use as a fuel in engine **11**. In particular, the mixer makes the gases easier to excite during operation of the engine. Mixing does not mean an atomic or molecular combination or unification of gases because inert gases cannot chemically combine, in general, due to the completeness of the outer shell of electrons. During mixing, the various gases form a homogeneous mixture. The mixing of the five inert gases in apparatus **201** is somewhat analogous to preparing a five part liquid chemical mixture by titration. In such a mixture, the proportions of the different chemicals are accurately determined by visually observing the end point of each reaction during titration. In apparatus **201**, a visible, spectroscopic flash of light accompanies the desired end point of the introduction of each new gas as it reaches its proper, precalculated proportion. (Each gas has its own distinctive, characteristic, spectroscopic display). The ends points are theoretically calculated and are determined by pre-set voltages on each of a group of ionising heads in the apparatus, as described below.

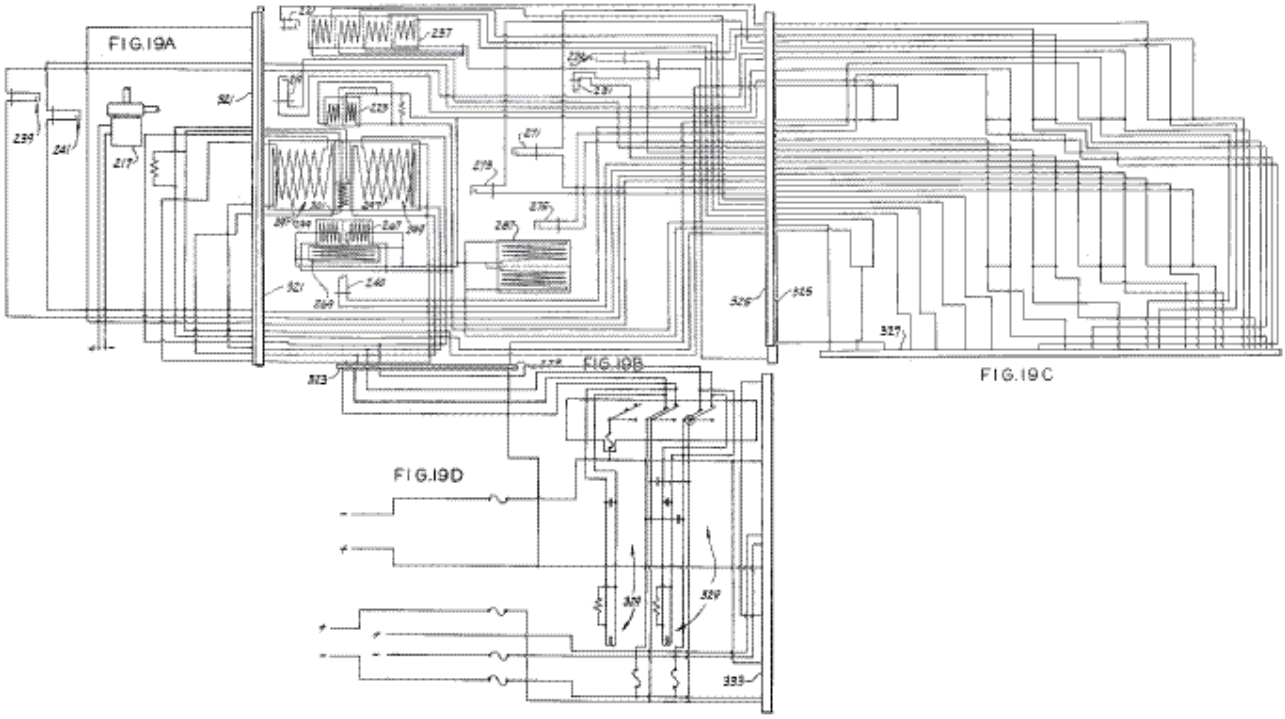


Mixer **201** includes (see **Fig.17A**) an intake port, indicated generally at **203**, which during operation is connected to a source **205** of helium gas, a gauge **206**, glass tubing **207** comprising a plurality of branches **B10-B25** for flow of the gases through the mixer, a plurality of valves **V1-V11** in the branches, which valves may be opened or closed as necessary, three gas reservoirs **209, 211** and **213** for storing small quantities of helium, argon and neon gas respectively, an ionising and filtering unit **215** for filtering undesired non-inert gases and contaminants out of the fuel mixture, for regulating the gas atom electron charge and to absorb the free flowing electrons, a gas flow circulation pump **217**, two ionising heads **219** and **221**, and three quality control and exhaust valves **V12-V14**. The mixer also comprises (see **Fig.17B**) a high frequency discharge tube **225**, a non-directed cathode ray tube **227**, two more ionising heads **229** and **231**, two additional gas reservoirs **233** and **235** for storing small quantities of xenon and krypton, a quadruple magnetic coil **237**, a group of valves **V15-V24**, valves **V23** and **V24** being quality control and exhaust valves, and a plurality of additional glass tubing branches **B26-B32**.

Turning to **Fig.17C**, mixer **201** also includes additional ionising heads **239, 240** and **241**, additional valves **V25-V46**, **V39A** and **V40A**, valves **V29** and **V32** being quality control and exhaust valves and valve **V39A** being a check valve, a vacuum and pressure gauge **242** between valves **V35** and **V36**, tubing branches **B34-B49** (branch **B39** consisting of two parts **B39A** and **B39B**), a pair of intake ports **243** and **245** which during operation are connected to sources **247** and **249** of argon and neon gas respectively, gauges **250A** and **250B**, a spark chamber **251**, a hydrogen and oxygen retention chamber **253** containing No. 650 steel dust in a silk filter, an ion gauge **255** (which can be an RG 75K type Ion Gauge from Glass Instruments, Inc. of Pasadena, Calif.) for removing excess inert gases from the mixture, inner and outer coils of glass tubing **257** and **259** surrounding a mixing chamber **261**, a focused x-ray tube **263** for subjecting the mixture flowing through it to 15-20 millirem alpha radiation and 120-125 millirem beta radiation, a directed cathode ray tube **265**, two twin parallel magnetic coils **266** and **267**, and a focusing magnetic coil **269**. It is important that coils **266** and **267** be immediately adjacent mixing chamber **261**. And (see **Fig.17D**) the mixer also comprises three more ionising heads **271**, **273** and **275**, two entry ports **277** and **279** which during operation are connected to sources **281** and **283** of krypton and xenon respectively, gauges **284A** and **284B**, a high frequency discharge tube **285**, a twin parallel magnetic coil **287** surrounding a polariser **289** for polarising the argon, said polarise containing fine steel particles which are polarised by coils **287** and which in turn polarise argon, a second hydrogen retention chamber **291**, a pair of tubing branches **B50** and **B51**, two filters **293** and **295** and a plurality of valves **V47-V59**, valves **V57** and **V59** being quality control and exhaust valves.

Inner and outer glass tubing coils **257** and **259** and mixing chamber **261** are shown in cross section in **Fig.18**. Intermediate glass coils **257** and **259** are two magnetic coils **297** and **299** having an inductance of approximately 130 mH. A yoke coil **301** is positioned in a semi-circle around mixing chamber **261**. Inside mixing chamber **261** are located a pair of screens **303** and **305**, insulators **307** and **309**, and a pair of spark gaps indicated generally at **311** and **313**. A high frequency amplitude modulated source provides 120 V AC, 60 Hz, 8.4 amp, 560 watt, 27,120 to 40,000 MHz plus or minus 160 KHz current via heavily insulated wires **315** and **317** to the chamber.

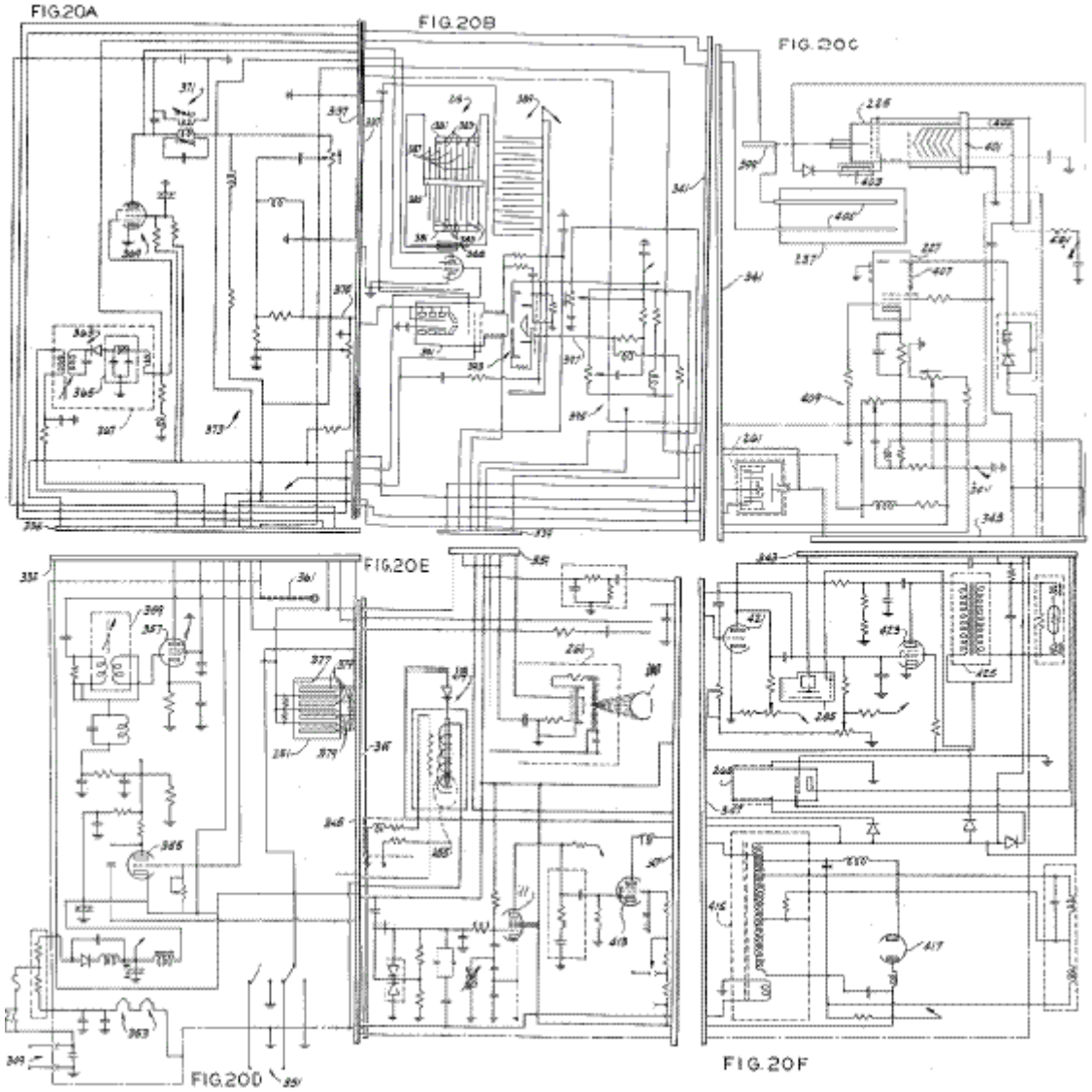
These wires are about twelve gauge, like those used as spark plug wires on internal combustion engines. Additionally 95 volt Direct Current is supplied via a smaller (e.g. sixteen to eighteen gauge) insulated wire **319**. As described below, the gases to be mixed and prepared flow through chamber **261** and are suitably treated therein by the action of the various fields present in the chamber.



The magnetic coils, ionisation heads, and pump **217**, along with the required electrical interconnections, are schematically shown in **Fig.19A** to **Fig.19E**. More particularly, heads **239** and **241** are shown in **Fig.19A**, as is pump **217**. Each ionising head has two electrodes with a gap between them to cause ionisation of gases flowing through the head, the electrodes being connected to a source of electrical power. Pump **217** is directly connected to a source of power (either AC or DC as required by the particular pump being used). The connections between the circuitry on **Fig.19A** and that on **Fig.19B** are shown as a plug **321**, it being understood that this plug represents a suitable one-to-one connection between the lines of **Fig.19A** and those of **Fig.19B**.

The remaining ionising heads and all the magnetic coils are shown in **Fig.19B**. For clarity, the coils are shown in an unconventional form. Quadruple coil **237** (shown at the top of **Fig.19B**) has one side of each winding connected in common but the other sides are connected to different lines. Coil **223** consists of two windings in parallel. Coils **297** and **299**, the ones around the mixing chamber, are shown overlapping, it being understood that coil **297** is actually interior of coil **299**. Yoke coil **301**, as shown, extends half-way from the bottom to the top of coils **297** and **299**. Twin parallel magnetic coils **267** are connected in parallel with each other, both sides of focusing coil **269** being connected to one node of coils **267**. Likewise coils **287** are connected in parallel. The connections between the lines of **Fig.19B** and those of **Fig.19C** and **Fig.19D** are shown as plugs **323** and **325**, although other suitable one-to-one connections could certainly be made. **Fig.19C** shows the interconnecting lines between **Fig.19B** and **Fig.19E**. A plug **327** or other suitable one-to-one connections connects the lines of **Fig.19C** and **Fig.19E**.

A plurality of power sources, like the above-mentioned Variacs, of suitable voltages and currents as well as a plurality of relays **329**, and plugs **331** are shown on **Fig.19D** and **Fig.19E**. The connections between these two Figures is shown as a plug **333**. It should be appreciated that the Variacs can be adjusted by the operator as necessary to supply the desired voltages to the aforementioned coils and ionising heads. It should also be realised that the desired relays can be closed or opened as needed by connecting or disconnecting the two parts of the corresponding plug **331**. That is, by use of plugs **331**, the operator can control the energising of the ionising heads and magnetic coils as desired. Plugs **331** are also an aid in checking to ensure that each component is in operating condition just prior to its use. Of course, the manipulation of the power sources and the relays need not be performed manually; it could be automated.



The remaining circuitry for the mixer is shown on **Fig.20A** to **Fig.20F**. For convenience, plugs **335**, **337**, **339, 341, 343, 345** and **347** are shown as connecting the circuitry shown in the various Figures, although other suitable one-to-one connections may be used. The chassis of the apparatus is shown on these Figures in phantom and is grounded. The power supply for the apparatus is shown in part on **Fig.20A** and **Fig.20D** and includes an input **349** (see **Fig.20D**) which is connected to 120 volt, 60 Hz power during operation and an input **351** which is connected to the aforementioned high frequency generator or some other suitable source of approximately 27,120 MHz current. The power supply includes a pair of tuners **353**, numerous RLC circuits, a triode **355**, a pentode **357** with a ZnS screen, a variable transformer **359**, an input control **361**, a second variable transformer **363** (see **Fig.20A**) which together with a filter **365** forms a 2.0 volts (peak-to-peak) power supply **367**, a pentode **369**, a variable transformer **371**, and a resistor network indicated generally at **373**. Exemplary voltages in the power supply during operation are as follows: The anode of triode **355** is at 145 V, the control grid at 135 V and the cathode at -25 V. The voltage at the top of the right-hand winding of transformer **359** is -5 V. The anode of pentode **357** is at 143 V, the top grid is grounded (as is the ZnS screen), the bottom grid is connected to transformer **359**, and the control electrode is at 143 V. The input to supply **367** is 143 volts AC while its output, as stated above, is 2 V (peak-to-peak). The anode of pentode **369** is at 60 V, the grids at -1.5 V, the control electrode at 130 V, and the cathode is substantially at ground. The output of resistor network **373**, labelled **375**, is at 45 V.

Also shown on **Fig.20D** is spark chamber **251**. Spark chamber **251** includes a small amount of thorium, indicated at **377**, and a plurality of parallel brass plates **379**. When the gases in the mixer reach the proper ionisation, the alpha particles emitted by the thorium shown up as flashes of light in the spark chamber.

Turning now to **Fig.20B**, ionising and filtering unit **215** includes a pair of conductive supports **381** for a plurality of conductors **383**, said supports and conductors being connected to a voltage source, an insulating support **385** for additional conductors **387**, and a ZnS screen **388** which emits light when impurities are removed from the gaseous fuel mixture. Unit **215** also includes a second set of interleaved conductors indicated generally at **389**, a

cold-cathode tube **391**, and an x-ray tube indicated generally at **393**. Also shown on **Fig.20B** is an RLC network **395** which has an output on a line **397** which is at 35 V, this voltage being supplied to the x-ray tube.

High frequency discharge tube **255** (see **Fig.20C**) has a conductive electrode **399** at one end to which high frequency current is applied to excite the gases in the mixer, and an electrode/heater arrangement **401** at the other, a voltage of 45 V being applied to an input **402** of the tube. It is desirable that a small quantity of mercury, indicated at **403**, be included in tube **225** to promote discharge of the helium gas. Magnetic coils **237** have disposed therein a pair of generally parallel conductors **405** to which a high frequency signal is applied. When gas flows through coils **237** and between parallel conductors **405**, therefore, it is subjected to the combination of a DC magnetic field from the coil and high frequency waves from the conductors, which conductors act as transmitting antennas. The resulting high frequency magnetic field causes the atoms to become unstable, which allows the engine to change a given atom's quantum level with much less input power than would normally be required. The volume of each gas atom will also be smaller. Also shown on **Fig.20C** is non-directed cathode ray tube **227**. The grids of tube **227** are at 145 V, the control electrode is at ground, while the anode is at 35 V to 80 V (peak-to-peak). The purpose of non-directed cathode ray tube **227** is to add photons to the gas mixture. To generate these photons, tube **227** has a two layer ZnS coating indicated generally at **407**. Chamber **261**, described above, is also shown schematically on **Fig.20C**, along with an RLC network **409**.

The power supply for the mixer (see the lower halves of **Fig.20E** and **Fig.20F**) also includes two pentodes **411** and **413**, a transformer **415**, and a diode tube **417**. The control electrode of pentode **411** is at 5 V to 40 V (peak-to-peak), the grids are at 145 V, the anode is at 100 V, and the cathode is at 8 V to 30 V (peak-to-peak). The control electrode of pentode **413** is at 115 V, while its grids and cathode are at -33 V. The anode of tube **413** is connected to transformer **415**. Also shown on **Fig.20E** are a relay **419** associated with ion gauge **255**, and focused x-ray tube **263** associated with ionisation head **240**. The upper input to tube **263** is at 45 V to 80 V (peak-to-peak).

Turning to **Fig.20F**, there is shown tubes **265** and **285**. Directed cathode ray tube **265** is a pentode connected like tube **227**. High frequency discharge tube **285** includes a phosphor screen and is connected to a high frequency source. Also shown on **Fig.20F** is a triode **421** with its anode at 30 V, its cathode at ground, and its control grid at -60 V; a pentode **423** with its anode at 135 V to 1000 V peak to peak, its cathode at ground, its control electrode at 143 V, its grids at 20 V; and a transformer **425**. It should be understood that various arrangements of electrical components other than those described above could be designed to perform the same functions.

The operation of the mixer is best understood with reference to **Fig.17A** to **Fig.17D** and is as follows: Before and during operation, the mixer, and particularly chamber **261** is kept hermetically sealed and evacuated. To begin the mixing process, helium is admitted into the mixer via intake port **203**. Then a vacuum is again drawn, by a vacuum pump (not shown) connected to valve **V38**, to flush the chamber. This flushing is repeated several times to completely cleanse the tubing branches of the mixer. The mixer is now ready. The ionisation heads next to mixing chamber **261** are connected to a voltage corresponding to approximately 36% of the calculated total ionising voltage, DC current is allowed to flow through magnetic coils **297** and **299** around chamber **261**, and high frequency current is allowed to pass through the mixing chamber. Helium is then slowly admitted, via port **203**, into the mixer. From port **203**, the helium passes through ionisation head **219** into glass tubing coil **259**. This glass coil, being outside magnetic coils **297** and **299**, is in the diverging portion of a magnetic field. The helium slowly flowing through glass coil **259** is gently excited. From coil **259**, the helium flows through branch **B45** to ionisation head **275** and from there, via branch **B28**, to ionisation head **229** (see **Fig.17B**). From head **229**, the gas flows through non-directed cathode ray tube **227** to high-frequency discharger **225**. The high frequency discharger **225**, with heating element, discharges, separates or completely neutralises the charge of any radioactive and/or cosmic particles that are in the helium atom in addition to the protons, neutrons and electrons.

The gas exits discharger **225** via branch **B26** and passes to high-frequency discharger **285**. The high frequency discharger **285**, without heating element, disturbs the frequency of oscillation which binds the gas atoms together. This prepares the helium atoms so that the electrons can more easily be split from the nucleus during the excitation and ignition process in the engine. Discharger **285** includes a phosphorus screen or deposit (similar to the coating on a cathode ray tube) which makes discharges in the tube visible. From discharger **285**, the helium passes through directed cathode ray tube **265** and focused x-ray tube **263**. Directed cathode ray tube **265** produces cathode rays which oscillate back and forth longitudinally underneath and along the gas carrying tube. After that, the helium passes successively through branch **B21**, ionisation head **221**, branch **B23**, twin parallel magnetic coil **266**, and branch **B25** into mixing chamber **261**. Helium flows slowly into and through apparatus **201**. The helium atoms become ionised as a result of excitation by magnetic force, high frequency vibrations and charge acquired from the ionisation heads. When sufficient helium has entered the apparatus, the ionisation energy (which is approximately 36% of the total) is totally absorbed. A spectroscopic flash of light in the mixing

chamber signals that the precise, proper quantity of helium has been allowed to enter. The entry of helium is then immediately halted by the closing of valve **V3**.

The next step in preparing the fuel is to add neon to the mixture. The potential on the relevant ionisation heads, particularly head **241** (see **Fig.17C**), is raised by the addition of approximately 26% which results in a total of approximately 62% of the total calculated potential and valve **V31** is opened, thereby allowing neon to slowly enter the mixer via port **245**. This gas passes through branch **B36**, ionisation head **241**, and branch **B35** directly into the mixing chamber. Since the previously admitted helium is fully charged, the neon absorbs all of the increased ionisation potential. As soon as the neon acquires the additional charge, a spectroscopic flash of light occurs and the operator closes valve **V31**.

In the same manner, the potential on the ionisation heads is increased by the addition of approximately 17% for a total of approximately 79% of the total calculated potential and then valve **V30** is opened to admit argon into the mixer via port **243**. This gas passes through branch **B34**, ionisation head **239**, and branch **B33** into mixing chamber **261**. Again, when the proper amount of argon has been admitted, it emits a spectroscopic flash of light and the operator closes valve **V30**. Next, the potential on the ionisation heads is increased by the addition of approximately 13% to result in a total of approximately 92% of the total calculated potential and valve **V58** (see **Fig.17D**) is opened to admit krypton into the system. The krypton gas passes through branch **B51**, ionisation head **271** and branch **B48** into chamber **261**. Upon the emission of a spectroscopic flash of light by the gas, the operator closes valve **V58**. Finally, the potential on the ionisation heads is increased by the addition of approximately 8% which brings the ionisation potential to the full 100% of the calculated ionisation voltage and valve **V56** is opened to admit xenon into the mixer via port **279**. This gas passes through branch **B50**, ionisation head **273** and branch **B47** to the mixing chamber. When the proper amount of gas has been admitted, a spectroscopic flash of light occurs signalling the operator to close valve **V56**. Note that there are two filter/absorber units, labelled **253** and **291**. Unit **253** is connected to the neon and argon inlet branches **B33** and **B35** while unit **291** is connected to the krypton and xenon inlet branches **B47** and **B48**. These two units absorb hydrogen residue and immobilise the water vapour created when the pump circulates the gases and generates vacuum states.

After all the gases are admitted in the desired proportions, all the valves are closed. (The mixture in the mixing chamber and in the adjacent tubing is at one atmosphere pressure at this time). Once this is done, the interval valves of the system are all opened (but the inlet and outlet valves remain closed) to allow the mixture to circulate throughout the tubing as follows: branch **B44**, magnetic coils **267** and **269**, ionisation head **240**, branch **B29**, ionisation head **231**, branch **B24**, ionisation head **219**, pump **217**, branches **B15** and **B39A**, ionisation gauge **255**, branches **B38** and **B42**, ionisation head **275**, branch **B28**, ionisation head **229**, non-directed cathode ray tube **227**, quadruple magnetic coil **272**, ionisation head **221**, branch **B23**, twin parallel magnetic coil **266**, branch **B25** and mixing chamber **261**. When this circuit is initially opened, the pressure of the mixture drops 40-50% because some of the tubing had previously been under vacuum. Pump **217** is then started to cause the gases to be slowly and evenly mixed.

Because of dead space in the tubing and the reaction time of the operator, it may occur that the proportions of the gases are not exactly those set forth above. This is remedied during the circulation step. As the gas flows through ionisation gauge **255**, excess gas is removed from the mixture so that the correct proportions are obtained. To do this the grid of gauge **255** is subjected to 100% ionisation energy and is heated to approximately 165 degrees F. This temperature of 165 degrees F is related to xenon's boiling point of -165 degrees F in magnitude but is opposite in sign. Xenon is the heaviest of the five inert gases in the mixture. As the gas mixture flows through ionisation gauge **255**, the gas atoms that are in excess of their prescribed percentages are burned out of the mixture and their charge is acquired by the remaining gas atoms from the grid of the ionisation gauge. Because the gases are under a partial vacuum, the ionisation gauge is able to adjust the gas percentages very precisely. (Note: The steps described in the last two paragraphs are repeated if the finished gases are rejected in the final quality control step described below).

The next step involves purifying the mixture so that only the five inert gases remain, absorbing any free electrons and regulating the electrical charge in the mixture. To do this, the circuit consisting of the following components is opened: Branch **B44**, magnetic coil **267**, magnetic coil **269**, ionisation head **240**, branch **B29**, ionisation head **231**, branch **B24**, ionisation head **219**, pump **217**, branches **B15** and **B39**, magnetic coil **287** (see **Fig.17D**) polariser **289**, branch **B17**, ionising and filtering unit **215**, branches **B16**, **B42**, and **B41**, x-ray tube **263**, branch **B21**, ionisation head **221**, branch **B23**, magnetic coil **266**, branch **B25**, and mixing chamber **261**. The gases should complete this circuit at least three times.

The last step required to prepare the mixture for bottling is polarisation of the argon. The circuit required to do this consists of the following components: mixing chamber **261**, branch **B44**, magnetic coil **267**, magnetic coil **269**, ionisation head **240**, cathode ray tube **265**, branch **B40**, tubing coil **257**, branches **B49** and **B30**, ionisation head **231**, branch **B24**, ionisation head **219**, pump **217**, branches **B15** and **B39**, twin parallel magnetic coil **287** (see

Fig.17D), polariser **289**, branch **B17**, ionising and filtering unit **215**, branches **B16**, **B42** and **B20**, ionisation head **229**, cathode ray tube **227**, magnetic coil **237**, ionisation head **221**, branch **B23** and magnetic coil **266**. This too is repeated at least three times. The key to the polarisation of argon is polariser **289** and twin parallel magnetic coil **287** that encircles it. Polariser **289** is a glass bottle which is filled with finely powdered soft iron which can be easily magnetised. The filled bottle is, in effect, the iron core of the coils. The iron particles align themselves with the magnetic lines of force, which lines radiate from the centre toward the north and south poles. The ionised gas mixture is forced through the magnetised iron powder by means of pump pressure and vacuum, thereby polarising the argon gas. Filters **293** and **295** are disposed as shown in order to filter metallic particles out of the gas.

The mixture is now double-checked by means of spark chamber **251** at atmospheric pressure since the fusion reaction in the engine is started at one atmosphere. Because the gases in mixing apparatus **201** are at a partial vacuum, sufficient gases must be pumped into spark chamber **251** to attain atmospheric pressure. To do this valves **V33**, **V36** and **V40A** are closed and circulating pump **217** pumps the gases in the mixing apparatus via branches **B15** and **B39A**, through check valve **V39A** into spark chamber **251** until the vacuum and pressure gauge **242** indicates that the gases within spark chamber **251** are at atmospheric pressure. Valve **V34** is then closed. The spark chamber is similar to a cloud chamber. Six or more high capacity brass capacitor plates are spaced 1/8" to 1/4" apart in the chamber. A small plastic container holds the thorium 232. One side of the chamber is equipped with a thick glass window through which sparks in the chamber may be observed. A potential is placed on the brass plates in the chamber and the current flowing between the plates is measured. If this current exactly corresponds to the ionisation current, the mixture is acceptable. A difference of greater than 5% is not acceptable. A lesser difference can be corrected by recirculating the gas in the mixer and particularly through ionisation gauge **255** as previously described in the circulation step. A second test is then given the gases that pass the first test. A calculated high frequency current is gradually imposed on the spark chamber capacitor plates. This excitation causes neutrons to be emitted from the thorium 232 which, if the mixture is satisfactory, can be easily seen as a thin thread of light in the chamber. If the mixture is not satisfactory, light discharges cannot be seen and the high frequency circuit will short out and turn off before the desired frequency is reached.

To bottle the mixture, valve **V33** is opened and valves **V36** and **V40** are closed. During bottling polariser **289**, twin parallel magnetic coil **287**, ionisation unit **215** and ion gauge **255** are electrically energised (all electrical circuits are previously de-energised) to improve the stability of the mixture. The prepared gases are withdrawn from the mixing apparatus via branches **B24** and **B16**, ionisation unit **215**, branch **B17**, filters **293** and **295**, polariser **289**, twin parallel magnetic coil **287**, branch **B39**, ion gauge **255**, check valve **V39A**, branch B38 and spark chamber 251. If desired, after bottling the mixer may be exhausted by opening valves **V12, V13, V14, V23, V24, V29, V32, V57** and **V59**. Of course, one can also automate the fuel preparation process to be continuous so that it would never be necessary to exhaust the gas.

In operation of mixing apparatus **201**, certain operational factors must be considered. For one, no electrical devices can be on without the pump being in operation because an electrical device that is on can damage adjacent gas that is not circulating. For another, it should be noted that directed cathode ray tube **265**, non-directed cathode ray tube **227** and focused x-ray tube **263** serve different functions at different points in the mixing process. In one mode, they provide hot cathode radiation, which can occur only in a vacuum. When gases are flowing through these devices, they provide a cold cathode discharge. For example, during argon polarisation and the circulation step, focused x-ray tube **263** is under vacuum and affects the gases flowing through ionisation head **240** by way of hot cathode radiation. During the introduction of the different gases into mixing apparatus **201** and during the recirculation step, the gases are flowing through focused x-ray tube **263**, which affects the gases by way of a cold cathode discharge.

It is preferred that each switchable electrical component in mixing apparatus **201** be wired into a separate circuit despite the fact that one of the poles of each could be commonly wired. In a common ground circuit if one device is turned on, all of the other units may also turn on because the gases in the device are conductive. In addition, if one unit on a common circuit were energised with high frequency current, the others would also be affected. In the same vein, the high frequency current cannot be used when the cathode ray tubes, the x-ray tubes or the dischargers are heated and under vacuum because the heater filaments will burn out.

Finally, the current source, the variable rectifiers and the electrical measuring instruments must be located more than ten feet from mixing apparatus **201** because the high frequency current is harmful to the rectifiers, causing them to burn out or short out.

It is hoped that a brief summary of the concepts used by the inventor in developing the above invention will be helpful to the reader, it being understood that this summary is in no way intended to limit the claims which follow or to affect their validity. The first concept is that of using an inert gas mixture at approximately one atmosphere at TDC (at ignition) as a fuel in a thermonuclear energy production process. The second concept is the layering of

the various inert gases, which layering is designed to confine the input energy in the innermost layers during pre-excitement and ignition, to provide thermal insulation for the container walls during and after ignition, to transmit power resulting from the ignition through the layers in turn to the piston, to absorb the pressure generated during ignition to protect the cylinder walls, and to provide an orderly, predictable positioning of the argon layer during the BDC to TDC portion of the engine cycle. The third concept of this invention involves utilising electric current produced in one cylinder of a pair to perform functions in the other cylinder of that pair. This concept includes the sub-concepts of generating electric current by atomic recombination and of electric generation in place resulting from the rotation of layered inert gases within each cylinder because of the changed polarity of the encircling coils at BDC, from judicious placement of coils which produce magnetic field lines which are cut by a near perfect conductor (polarised argon), and from movement of said near perfect conductor through the magnetic field.

The fourth and fifth concepts of this invention are the transformation of rapid, intense, but short duration thermonuclear reactions into pressure that is transmitted from inert gas to inert gas until it creates linear kinetic energy at the piston, which energy is converted into rotary kinetic energy by a crankshaft, and the use of a shaft-driven generator to provide power to spaced field coils during the BDC to TDC portion of the cycle of each cylinder.

The sixth concept concerns adequate pre-excitement of the inert gas fuel and more particularly involves the sub-concepts of pre-exciting the fuel in the mixing process, of manipulation of the currents in the coils surrounding each cylinder, of discharging the capacitors surrounding each cylinder at predetermined times in the cycles, of causing a stream of electrical particles to flow between electrodes and a conductive discharge point on the piston, of emitting alpha, beta and gamma rays from an anode and a cathode containing low level radioactive material to the piston's discharge point, of accelerating the alpha, beta and gamma rays by the application of a high-voltage field, and of situating capacitor plates 90 degrees from the anode and cathode to slow and reflect neutrons generated during ignition. The seventh concept involves the provision of a minute, pellet-type fission ignition, the heat from which causes a minute fusion as the result of the ignition chamber shape and arrangement, as a result of the collision of the alpha, beta and gamma rays and the electrical particles at a focal point in conjunction with the discharge of the capacitors that surround the cylinder through the electrodes, and as a result of increasing the magnetic field in the direction of the movement of each piston.

Robert Britt's Inert Gas Engine

US Patent 3,977,191 **31st August 1976** **Inventor: Robert G. Britt**

ATOMIC EXPANSION REFLEX OPTICS POWER SOURCE (AEROPS) ENGINE

ABSTRACT

An engine is provided which will greatly reduce atmospheric pollution and noise by providing a sealed system engine power source which has no exhaust nor intake ports. The engine includes a spherical hollow pressure chamber which is provided with a reflecting mirror surface. A noble gas mixture within the chamber is energised by electrodes and work is derived from the expansion of the gas mixture against a piston.

SUMMARY OF THE INVENTION

An atomic expansion reflex optics power source (AEROPS) engine, having a central crankshaft surrounded by a crankcase. The crankcase has a number of cylinders and a number of pistons located within the cylinders. The pistons are connected to the crankshaft by a number of connecting rods. As the crankshaft turns, the pistons move in a reciprocating motion within the cylinders. An assembly consisting of a number of hollow spherical pressure chambers, having a number of electrodes and hollow tubes, with air-cooling fins, is mounted on the top of each cylinder. The necessary gaskets are provided as needed to seal the complete engine assemblies from atmospheric pressure. A means is provided to charge the hollow spherical pressure chamber assembly and the engine crankcase with noble gas mixtures through a series of valves and tubes. A source of medium-voltage pulses is applied to two of the electrodes extending into each of the hollow spherical pressure chambers.

When a source of high-voltage pulses is applied from an electrical rotary distributor switch to other electrodes extending into each of the hollow spherical pressure chambers in a continuous firing order, electrical discharges take place periodically in the various hollow spherical pressure chambers. When the electrical discharges take place, high energy photons are released on many different electromagnetic frequencies. The photons strike the atoms of the various mixed gases, e.g., xenon, krypton, helium and mercury, at different electromagnetic frequencies to which each is selectively sensitive, and the atoms become excited. The first photons emitted are reflected back into the mass of excited atoms by a reflecting mirror surface on the inside wall of any particular hollow spherical pressure chamber, and this triggers more photons to be released by these atoms. They are reflected likewise and strike other atoms into excitation and photon energy release. The electrons orbiting around the protons of each excited atom in any hollow spherical pressure chamber increase in speed and expand outward from centre via centrifugal force causing the atoms to enlarge in size. Consequently, a pressure wave is developed, the gases expand and the pressure of the gas increases.

As the gases expand, the increased pressure is applied to the top of the pistons in the various cylinders fired selectively by the electrical distributor. The force periodically applied to the pistons is transmitted to the connecting rods which turn the crankshaft to produce rotary power. Throttle control valves and connecting tubes form a bypass between opposing hollow spherical pressure chambers of each engine section thereby providing a means of controlling engine speed and power. The means whereby the excited atoms are returned to normal minimum energy ground-state and minimum pressure level, is provided by disrupting the electrical discharge between the medium-voltage electrodes, by cooling the atoms as they pass through a heat transfer assembly, and by the increase in the volume area above the pistons at the bottom of their power stroke. The AEROPS engine as described above provides a sealed unit power source which has no atmospheric air intake nor exhaust emission. The AEROPS engine is therefore pollution free.

BRIEF OBJECTIVE OF THE INVENTION

This invention relates to the development of an atomic expansion reflex optics power source (AEROPS) engine, having the advantages of greater safety, economy and efficiency over those disclosed in the prior art. The principal object of this invention is to provide a new engine power technology which will greatly reduce atmospheric pollution and noise, by providing a sealed system engine power source which has no exhaust nor intake ports.

Engine power is provided by expanding the atoms of various noble gas mixtures. The pressure of the gases increases periodically to drive the pistons and crankshaft in the engine to produce safe rotary power. The objects and other advantages of this invention will become better understood to those skilled in the art when viewed in light of the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig.1 is an elevational view of the hollow spherical pressure chamber assembly, including sources of gas mixtures and electrical supply:

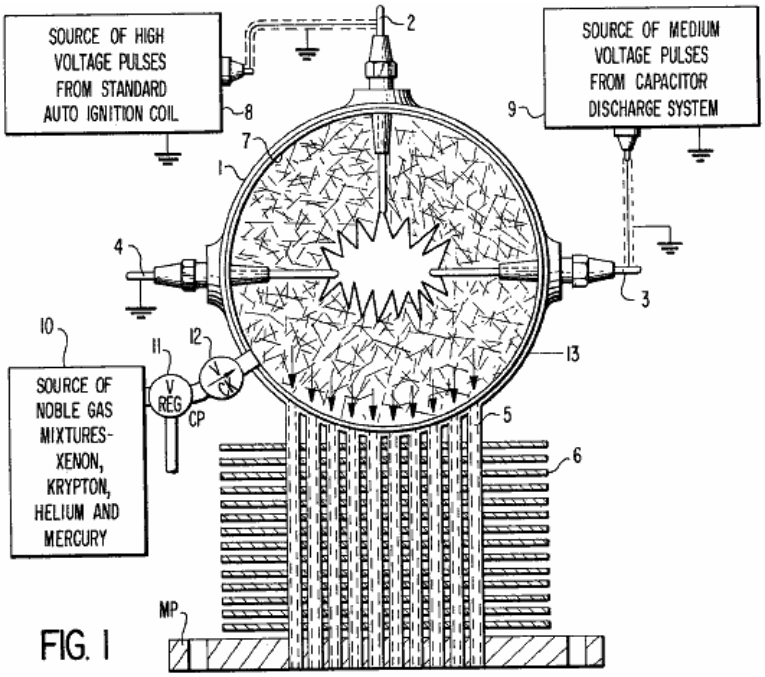


FIG. 1

Fig.2 is an elevational view of the primary engine power stroke:

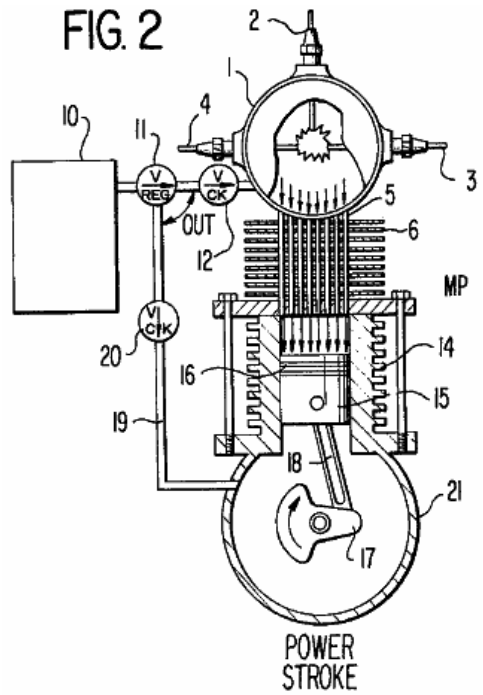


Fig.3 is an elevational view of the primary engine compression stroke:

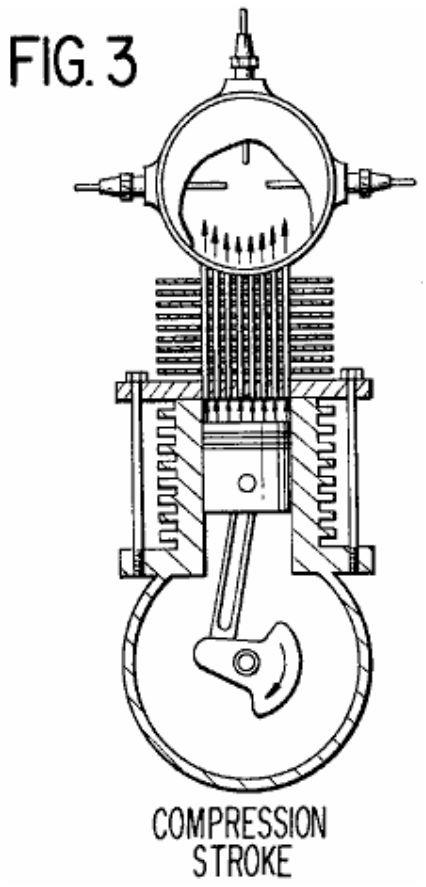


Fig.4 is a rear elevational view of a six cylinder AEROPS engine:

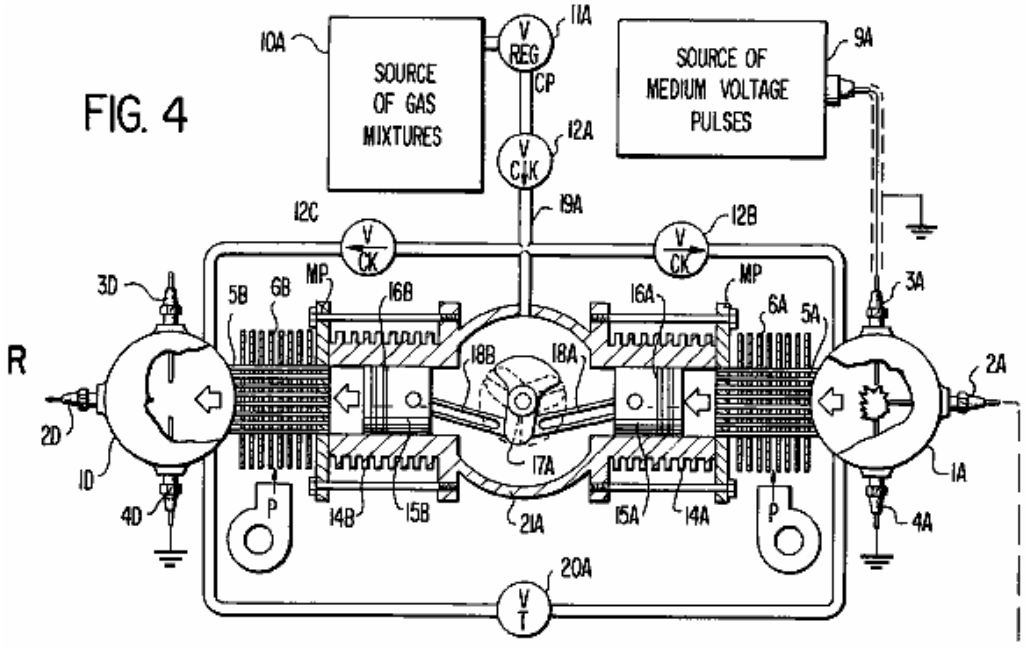


Fig.5 is a top view of the six cylinder AEROPS engine:

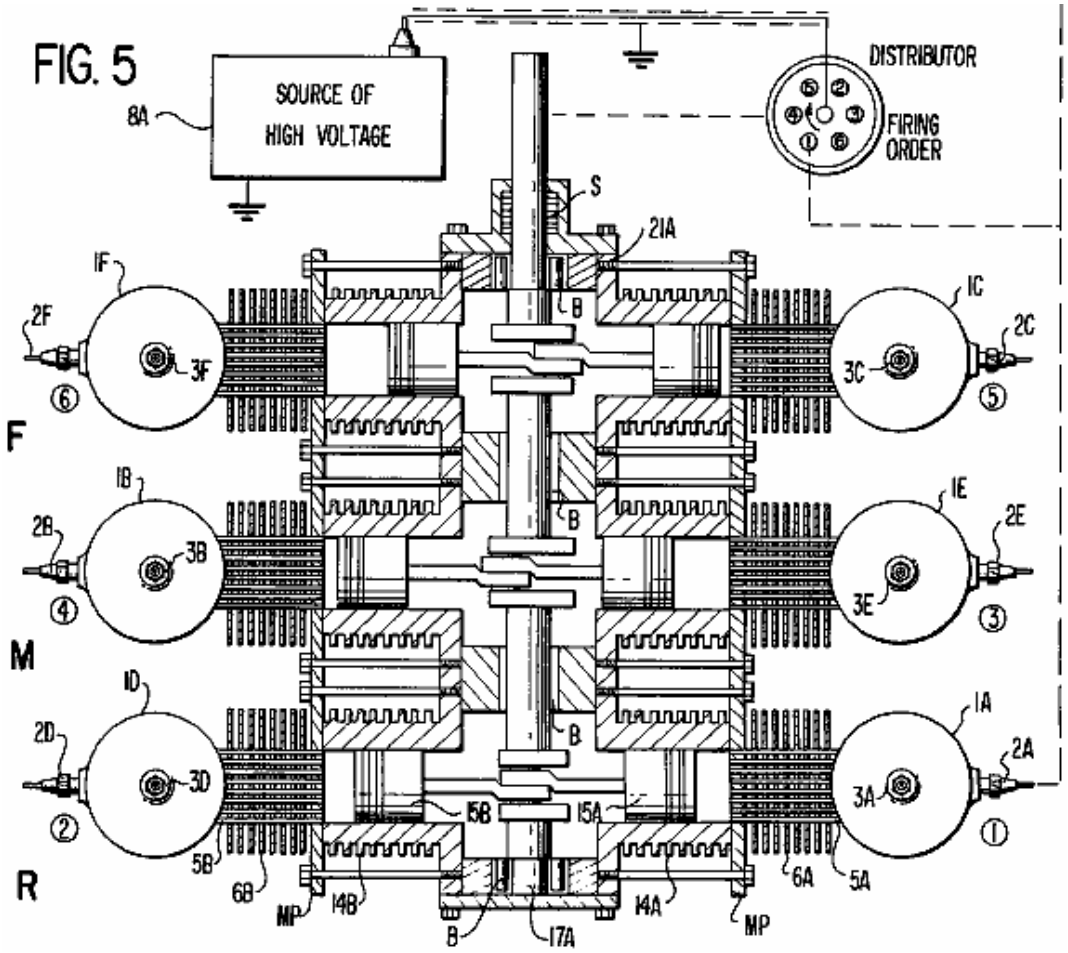


Fig.6 is an electrical schematic of the source of medium-voltage:

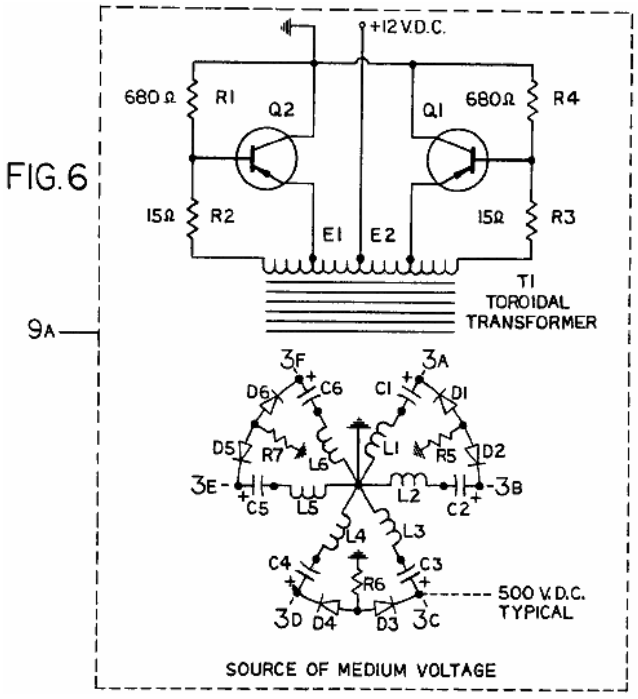
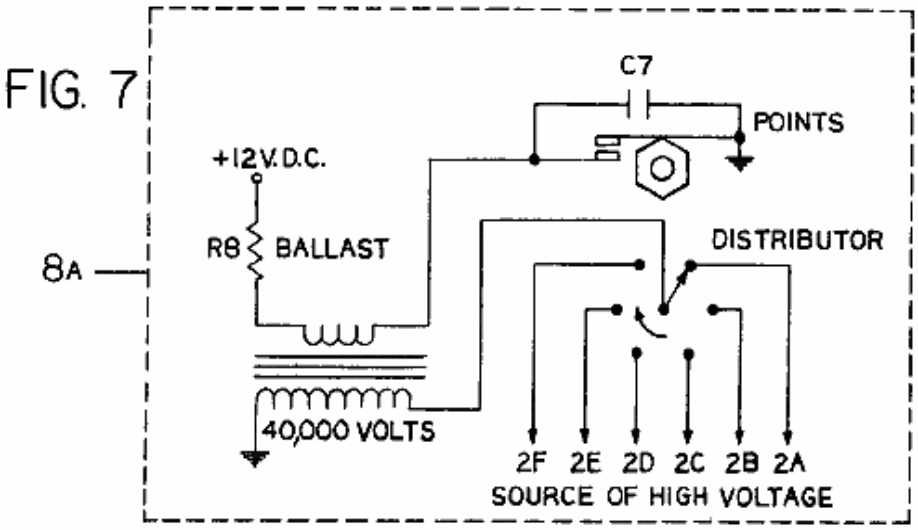


Fig.7 is an electrical schematic of the source of high-voltage:



DETAILED DESCRIPTION

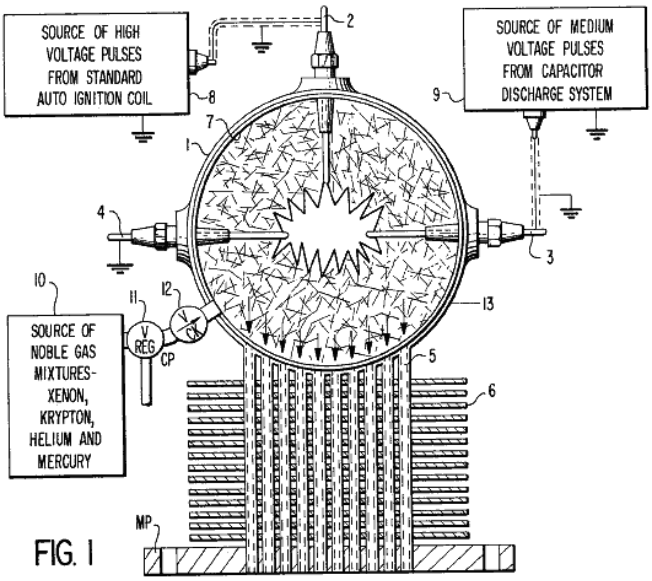
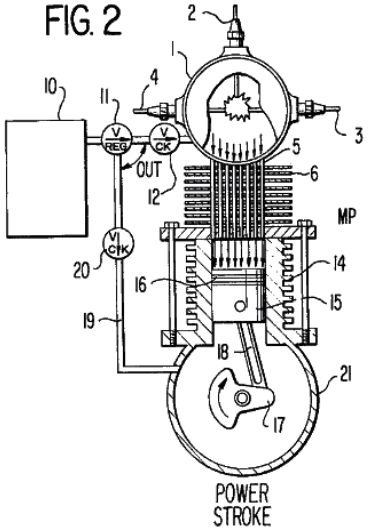


FIG. 1

Referring to **Fig.1** of the drawings, the AEROPS engine comprises a hollow spherical pressure chamber **1** having an insulated high-voltage electrode **2** mounted on the top, an insulated medium-voltage electrode **3** mounted on the right, and an insulated common ground electrode **4** mounted on the left, as shown in this particular view. Electrodes **2, 3** and **4** extend through the wall of the hollow spherical pressure chamber **1** and each electrode forms a pressure seal. A plurality of hollow tubes **5** arranged in a cylindrical pattern extend through the wall of the hollow spherical pressure chamber **1**, and each hollow tube is welded to the pressure chamber to form a pressure seal. The opposite ends of hollow tubes **5** extend through the mounting plate **MP** and are welded likewise to form a pressure seal. A plurality of heat transfer fins **6** are welded at intervals along the length of said hollow tubes **5**. A bright reflecting mirror surface **7** is provided on the inner wall of the hollow spherical pressure chamber **1**. A source of high-voltage **8** is periodically connected to the insulated high-voltage electrodes **2** and **4**. A source of medium-voltage **9** from a discharge capacitor is connected to the insulated medium-voltage electrodes **3** and **4**. A source of noble gas mixtures **10**, e.g., xenon, krypton, helium and mercury is applied under pressure into the hollow spherical pressure chamber **1** through pressure regulator valve **11** and check valve **12**.

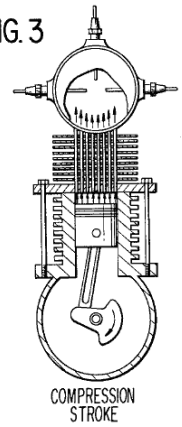
FIG. 2



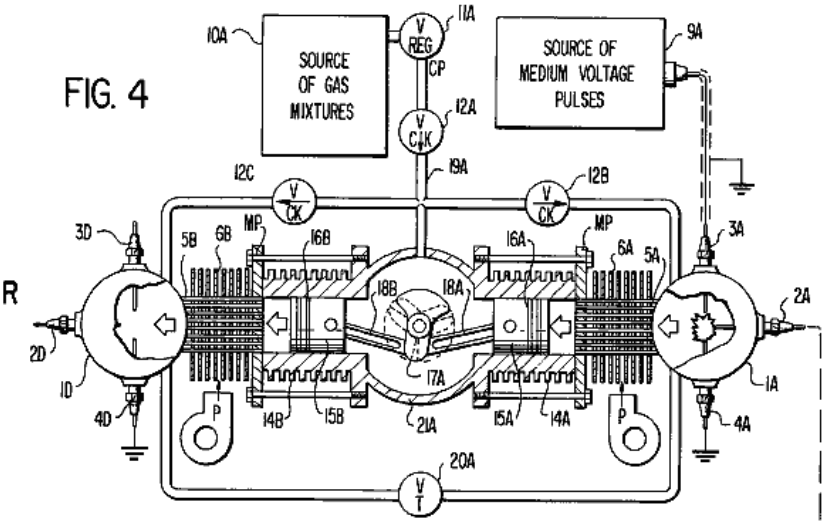
Referring now to **Fig.2** of the drawings, the complete assembly **13** shown in **Fig.1** is mounted on the top of the cylinder **14** via mounting plate **MP**. The necessary gaskets or other means are provided to seal the engine and prevent loss of gases into the atmosphere. The piston **15** located within cylinder **14** has several rings **16** which seal against the inner wall of the cylinder. The piston **15** is connected to the crankshaft **17** by connecting rod **18**. The source of noble gas mixtures **10** is applied under pressure into the crankcase **21** through pressure regulator valve **11**, check valve **12** and capillary tube **19**. The piston **15** is now balanced between equal gas pressures. Assuming that the engine is running and the piston **15** is just passing Top-Dead-Centre (TDC), a source of medium-voltage from a capacitor discharge system **9** (**Fig.6**, a single typical capacitor section) is applied to electrodes **3** and **4**. A source of high-voltage pulses from a standard ignition coil **8** (such as shown in **Fig.7**) is applied to electrodes **2** and **4** and the gases within the hollow spherical pressure chamber **1** are ionised and made electrically conductive. An electrical discharge takes place between electrodes **3** and **4** through the gases in the hollow spherical pressure chamber **1**.

The electrical discharge releases high energy photons on many different electromagnetic frequencies. The photons strike the atoms of the various gases, e.g., xenon, krypton, helium and mercury at different electromagnetic frequencies to which each atom is selectively sensitive and the atoms of each gas become excited. The first photons emitted are reflected back into the mass of excited atoms by the reflecting mirror surface **7**. This triggers more photons to be released by these atoms, and they are reflected likewise from the mirror surface 7 and strike other atoms into excitation and more photons are released as the chain reaction progresses. The electrons orbiting around the protons of each excited atom increase in speed and expand outward in a new orbital pattern due to an increase in centrifugal force. Consequently, a pressure wave is developed in the gases as the atoms expand and the overall pressure of the gases within the hollow spherical pressure chamber **1** increases. As the gases expand they pass through the hollow tubes **5** and apply pressure on the top of piston **15**. The pressure pushes the piston **15** and the force and motion of the piston is transmitted through the connecting rod **18** to the crankshaft **17** rotating it in a clockwise direction. At this point of operation, the power stroke is completed and the capacitor in the medium-voltage capacitor discharge system **9** is discharged. The excited atoms return to normal ground state and the gases return to normal pressure level. The capacitor in the medium-voltage capacitor discharge system **9** is recharged during the time period between (TDC) power strokes.

FIG. 3

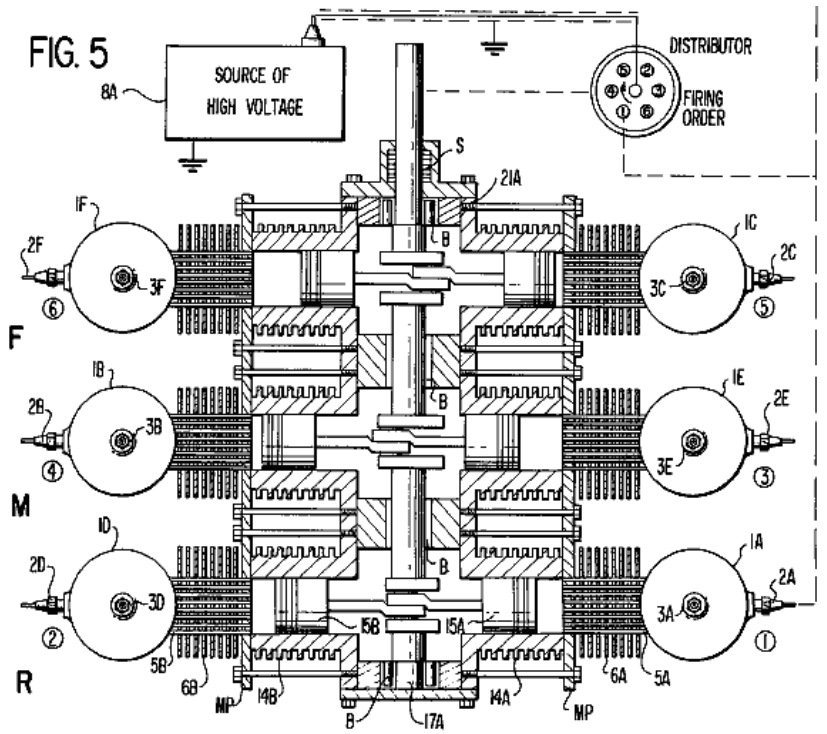


Referring now to **Fig.3** of the drawings, the compression stroke of the engine is shown. In this engine cycle the gases above the piston are forced back into the hollow spherical pressure chamber through the tubes of the heat transfer assembly. The gases are cooled as the heat is conducted into the fins of the heat transfer assembly and carried away by an air blast passing through the fins. An example is shown in **Fig.4**, the centrifugal air pump **P** providing an air blast upon like fins.



Some of the basic elements of the invention as set forth in **Fig.1**, **Fig.2**, and **Fig.3** are now shown in **Fig.4** and **Fig.5** which show complete details of a six-cylinder horizontally-opposed AEROPS engine.

Referring now to **Fig.4** and **Fig.5** of the drawings. **Fig.4** is a view of the rear section of the engine showing the crankshaft, centre axis and two of the horizontally-opposed cylinders. In as much as the rear **R**, middle **M** and front **F** sections of the engine possess identical features, only the rear **R** engine section will be elaborated upon in detail in order to prevent repetition and in the interest of simplification. The crankshaft **17A** consists of three cranks spaced 120 degrees apart in a 360 degree circle as shown. Both connecting rods **18A** and **18B** are connected to the same crank. Their opposite ends connect to pistons **15A** and **15B**, located in cylinders **14A** and **14B** respectively. Each piston has pressure sealing rings **16A** and **16B**. The hollow spherical pressure chamber assemblies consisting of **1A** and **1D** are mounted on cylinders **14A** and **14B** via mounting plates **MP**. The necessary gaskets are provided as needed to seal the complete engine assemblies from atmospheric pressure.



The source of gas mixtures **10A** is applied under pressure to pressure regulator valve **11A** and flows through check valve **12A**, through check valve **12B** to the hollow spherical pressure chamber **1A**, and through check valve

12C to the hollow spherical pressure chamber **1D**. The gas flow network consisting of capillary tubes below point **19A** represents the flow of gases to the rear section **R** of the engine. The middle section **M** and the front section **F** both have gas flow networks identical to that consisting of capillary tubes below point **19A**, while the gas flow network above is common to all engine sections. Throttle valve **20A** and the connecting tubing form a variable bypass between hollow spherical pressure chambers **1A** and **1D** to control engine speed and power. Engine sections **R**, **M** and **F** each have this bypass throttle network. The three throttle valves have their control shafts ganged together. A source of medium-voltage pulses **9A** is connected to medium-voltage electrodes **3A** and **3D**. In one particular embodiment the medium-voltage is 500 volts. A source of high-voltage pulses **8A** is connected to electrode **2A** through the distributor as shown. Electrode **4A** is connected to common ground. Centrifugal air pumps **P** force air through heat transfer fins **6A** and **6B** to cool the gases flowing in the tubes **5A** and **5B**.

Fig.5 is a top view of the AEROPS engine showing the six cylinders and crankshaft arrangement consisting of the rear **R**, middle **M** and front **F** sections. The crankshaft **17A** is mounted on bearings **B**, and a multiple shaft seal **S** is provided as well as the necessary seals at other points to prevent loss of gases into the atmosphere. The hollow spherical pressure chambers **1A, 1B, 1C, 1D, 1E** and **1F** are shown in detail with high-voltage electrodes **2A, 2B, 2C, 2D, 2E, 2F** and medium-voltage electrodes **3A, 3B, 3C, 3E** and **3F**. The common ground electrodes **4A, 4B, 4C, 4D, 4E, 4F** are not shown in **Fig.5** but are typical of the common ground electrodes **4A** and **4D** shown in **Fig.4**. It should be noted that the cranks on crankshaft **17A** are so arranged to provide directly opposing cylinders rather than a conventional staggered cylinder design.

FIG. 6

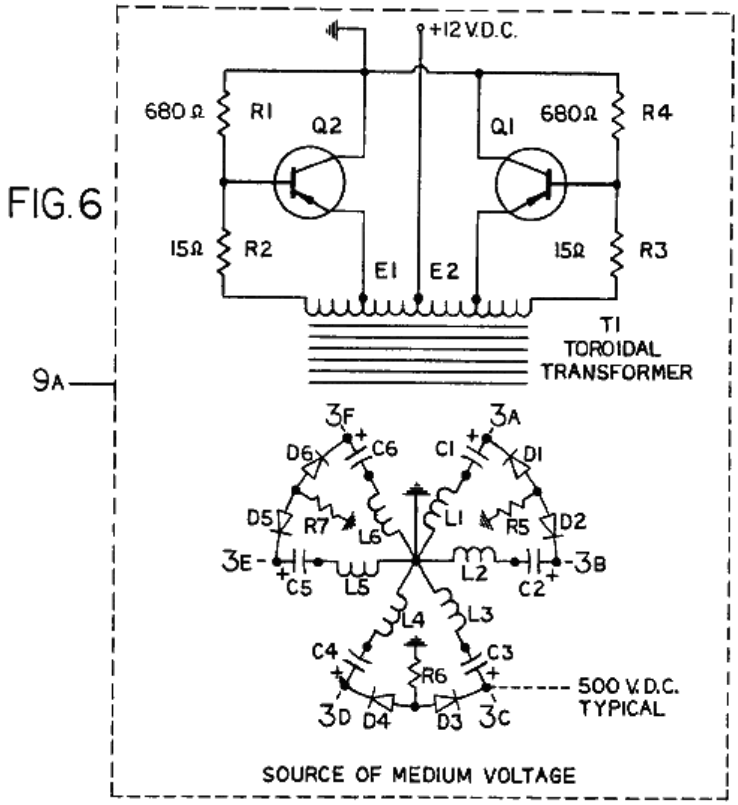
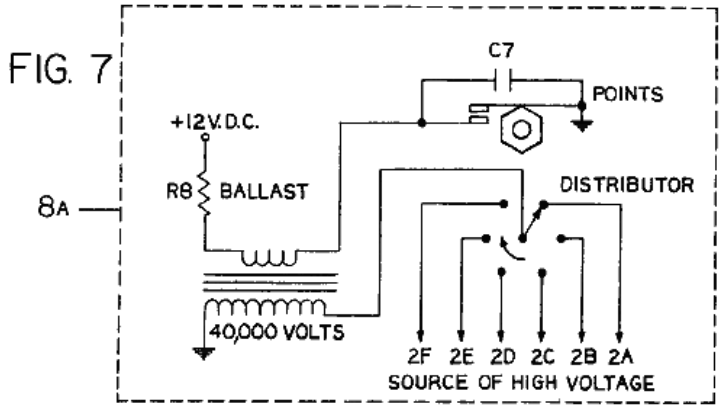


Fig.6 is an electrical schematic of the source of medium-voltage **9A**. The complete operation of the converter is explained as follows: The battery voltage 12 VDC is applied to transformer **T1**, which causes currents to pass through resistors **R1, R2, R3** and **R4**. Since it is not possible for these two paths to be exactly equal in resistance, one-half of the primary winding of **T1** will have a somewhat higher current flow. Assuming that the current through the upper half of the primary winding is slightly higher than the current through the lower half, the voltages developed in the two feedback windings (the ends connected to **R3** and **R2**) tend to turn transistor **Q2** on and transistor **Q1** off. The increased conduction of **Q2** causes additional current to flow through the lower half of the transformer primary winding. The increase in current induces voltages in the feedback windings which further drives **Q2** into conduction and **Q1** into cut-off, simultaneously transferring energy to the secondary of **T1**. When the current through the lower half of the primary winding of **T1** reaches a point where it can no longer increase due to the resistance of the primary circuit and saturation of the transformer core, the signal applied to the transistor from the feedback winding drops to zero, thereby turning **Q2** off. The current in this portion of the primary winding drops immediately, causing a collapse of the field about the windings of **T1**. This collapse in field flux, cutting across all of the windings in the transformer, develops voltages in the transformer windings that are opposite in polarity to the voltages developed by the original field. This new voltage now drives **Q2** into cut-off

and drives **Q1** into conduction. The collapsing field simultaneously delivers power to the secondary windings **L1, L2, L3, L4, L5** and **L6**. The output voltage of each winding is connected through resistors **R5, R6** and **R7** and diode rectifiers **D1, D2, D3, D4, D5** and **D6**, respectively, whereby capacitors **C1, C2, C3, C4, C5** and **C6** are charged with a medium-voltage potential of the polarity shown. The output voltage is made available at points **3A, 3B, 3C, 3D, 3E** and **3F** which are connected to the respective medium-voltage electrodes on the engine shown in **Fig.4** and **Fig.5**.



Referring now to **Fig.7** of the drawings, a conventional "Kettering" ignition system provides a source of high-voltage pulses **8A** of approximately 40,000 volts to a distributor, which provides selective voltage output at **2A, 2B, 2C, 2D, 2E** and **2F**, which are connected to the respective high-voltage electrodes on the engine shown in **Fig.4** and **Fig.5**. The distributor is driven by the engine crankshaft **17A** (**Fig.5**) at a one to one mechanical gear ratio.

Referring again to **Fig.4** and **Fig.5** of the drawings, the operation of the engine is as follows: Assuming that a source of noble gas mixtures, e.g., xenon, krypton, helium and mercury is applied under pressure to the hollow spherical pressure chambers **1A, 1B, 1C, 1D, 1E** and **1F** and internally to the crankcase **21A** through pressure regulator valve **11A** and check valves **12A**, **12B** and **12C**; and the source of medium-voltage **9A** is applied to electrodes **3A, 3B, 3C, 3D, 3E** and **3F**; and a source of high-voltage pulse **8A** is applied to electrode **2A** through the timing distributor, the gas mixtures in the hollow spherical pressure chamber **1A** is ionised and an electrical discharge occurs immediately between electrodes **3A** and **4A**.

High-energy photons are released on many different electromagnetic frequencies. The photons strike the atoms of the various gases, e.g., xenon, krypton, helium and mercury at different electromagnetic frequencies to which each is particularly sensitive and the atoms of each gas become excited. The first photons emitted are reflected back into the mass of excited atoms by the internal reflecting mirror surface on the inside wall of the hollow spherical pressure chamber **1A**. This triggers more photons to be released by these atoms and they are reflected likewise from the mirror surface and strike other atoms into excitation and more photons are released as the chain reaction progresses. The electrons orbiting around the protons of each excited atom in the hollow spherical pressure chamber **1A** increase in speed and expand outward in a new orbital pattern due to an increase in centrifugal force. Consequently, a pressure wave is developed in the gases as the atoms expand and the overall pressure of the gases within the hollow spherical pressure chamber **1A** increases.

As the gases expand they pass through the hollow tubes **5A** applying pressure on the top of piston **15A**. The pressure applied to piston **15A** is transmitted through connecting rod **18A** to the crankshaft **17A** rotating it in a clockwise direction. As the crankshaft **17A** rotates it pushes piston **15B** via connecting rod **18B** in the direction of a compression stroke, forcing the gases on the top of the piston through hollow tubes **5B** into the hollow spherical pressure chamber **1D**. As the gases pass through the hollow tubes **5A** and **5B** the heat contained in the gases is conducted into the heat transfer fins **6A** and **6B**, where it is dissipated by a blast of air passing through said fins from the centrifugal air pumps **P**. At this point of operation the power stroke of piston **15A** is completed and the capacitor in the medium-voltage capacitor discharge system **9A** is discharged. The excited atoms return to normal ground state and the gases return to normal pressure level. The capacitor in the medium-voltage capacitor discharge system **9A** is recharged during the time period between the power strokes of piston **15A**.

The above power stroke cycle occurs exactly the same in the remaining cylinders as the high-voltage firing order progresses in respect to the position of the distributor switch. In as much as the AEROPS engine delivers six power strokes per single crankshaft revolution, the crankshaft drives the distributor rotor at a one to one shaft ratio. The complete high-voltage firing order is 1, 4, 5, 2, 3, 6, whereas, the high-voltage is applied to electrodes **2A, 2B, 2C, 2D, 2E** and **2F** respectively. A means of controlling engine speed and power is provided by a plurality of throttle control valves and connecting tubes which form a bypass between opposing hollow spherical pressure chambers of each engine section.

The AEROPS engine as described above provides a sealed unit power source which has no atmospheric air intake nor exhaust emission and is therefore pollution free.

Floyd Sweet

Recently, some additional information on Floyd Sweet's device, has been released publicly by an associate of Floyd's who goes just by his first name of "Maurice" and who, having reached the age of seventy has decided that it is time to release this additional information.

Maurice says: After observing the comments made over the past year regarding the Sweet-VTA Energy Device, I decided to "come out of the woodwork" and explain what basically is NOT known regarding Floyd Sweet ("Sparky") and his energy device.

Keep in mind that I am 70 years old, quite computer illiterate, my background Being mainly Political Science (Graduate Degree); consulting with State Legislatures; Mental Health (former Executive Director of five clinics); and, acquiring Venture Capital for High Tech. Equipment (such as medical equipment) and various Projects. My story is very unusual and strange, but, nevertheless TRUE! At my age I have no one to impress with what I am about to tell you. My only interest is to correct error where possible and to make certain information known!

Remember, that I have never had any education in electronics. This was a real advantage for me because I did not have any electrical principles which I had to UN-LEARN in anything that Floyd told us. Unfortunately, one of my brothers who trained for 35 years in electronics was "blown away" when Floyd told him that "he needed to reverse the concepts which he was taught about the action of an electron and treat it like it was positive". Therefore, for Sparky's modelling, electrons were flowing and acting in the opposite direction to what was normally modelled by a trained physicist. See what I mean? The Dean of the School of Science of MIT that verified that Sparky had an MSEE degree and came third in his class of more than two hundred.

Hopefully sincere researchers will be able to obtain some useful information in what I attempt to explain in the future that will help them to duplicate what Floyd had. In this respect, one day after Floyd had repeatedly asked me: "What is this device Maurice?" and I repeatedly gave him the wrong answer, saying that it was an energy device, I finally realised that what was important to him was that he considered the device to be a TIME MACHINE - his emphasis was NOT on the energy. He told me never to forget that the most important thing was that the device was a "Time Machine".

Maurice draws attention to the fact that Floyd Sweet graduated as an M.S.E.E. from the Massachusetts Institute of Technology in 1969 and his thesis "Dynamics of Magnetic Domains" is considered by the M.I.T. scientific community to be unparalleled in magnetic concepts. He received the coveted Dean's Award for his scientific research and his academic level in Electrical Engineering achievement ranks third in the history of the M.I.T. School of Science. He has an extraordinary talent in the area of Engineering Mathematics not to mention his concept of electromagnetic and related electrical phenomena and understanding of abstract intangibles needed to predict the unforeseen.

Maurice says: In about 1988 John, who my two brothers and I were involved with in the High Tech field realised that my brother, who was a Doctor (Doctor brother), was interested in negative energy devices for the treatment of the physical body (similar to Rife/Tesla Frequency Machines). John had formerly been employed at NASA with Floyd Sweet. John lived in California close to Floyd (Sherman Oaks).

My doctor brother and I were introduced to Floyd by John and we waited patiently for the time when we could see the VTA device. We saw it on the table at his house during various visits but it was not operating. Floyd was like many inventors who played games with you. Each time we would drive 13 hours to see him thinking we could see the device operating, but he would have some excuse for not turning it on, or he would just ignore the purpose of our visit.

On one visit, I looked over at Floyd and he was “showing off” his Barium Ferrite bar magnet. The magnet was approximately 1/2” thick, 7” long and 3” wide. He had a small piece of metal that was standing on the top of the magnet at a 45 degree angle. As I recall, he claimed that the 45 degree angle was needed in the treatment of the magnet so that it could capture Scalar waves. The magnets were mainly functioning as a “gate” for the Scalar waves. Additionally, if you placed a piece of thin “flexible” (ribbon type) metal flat on the top of the magnet, the middle of the “ribbon metal” would be “sucked down” flat at the middle of the magnet and both ends of the “ribbon metal” would be bowed-up at each end of the magnet. Also, I came to understand from another inventor that we introduced later on to Floyd that the “figure eight” design (flux flow?) on the top of the magnet played an important part in the functioning of the magnet - I don’t really know about the concept and can’t relay any additional information.

On another visit, Floyd demonstrated the flowing flux of the magnet. He had a TV monitor and he would place the magnet by the screen and you could see all the beautiful colours of the flux as it moved across the monitor screen. My electronics brother told me that Floyd had told him that he had a way of treating the magnet by calibrating the Scalar wave angle coming in using the TV monitor. A side note is that Floyd delighted in telling people, when they asked how he treated his magnets, they should get the magnets real hot first. This apparently "screwed up" the magnetism and he enjoyed doing this for some weird reason!

Finally, after 12 trips across the California Desert, Floyd agreed to show us the Device in operation. In his defence, Floyd did claim that on some earlier planned demonstrations that his magnets had been "pulverized" by artificial earthquakes coming up through Mexico. He designed some type of buffer in the Device that eliminated the problem, but, it was an on-going problem for quite a period of time. This reminds me now that I must digress because I need to tell you about the Government (or who?) involvement with us.

When we first started to visit Floyd, our phones were all "tapped" - I do not know by whom. My electronics brother worked full-time with the Air National Guard and his specialty was electronic Security, Crypto, etc. tied in with SAC bases in our area and the surrounding States. Additionally, he had set-up the "clean room" for the President of the United States when he visited our State. I mention this because even my electronics brother was doubtful in the beginning that we were all being monitored. On one occasion, my doctor brother had his complete prior telephone conversation played back to him when he answered the phone (twenty minutes later) - I think it was probably some type of "screw-up" by whoever was monitoring our phones. My biggest complaint was the consistent early morning 3am call and then a "hang-up" when you answered - for what reason I don't know other than for harassment purposes.

I give you the above information so that you can understand the seriousness of what we were involved with.

Floyd's Energy Device was mainly three things:

(1) It was a healing device - negative electricity - negative time. In theory, you could re-set the template in your DNA with this energy source and therefore cleanse the body of all impurities that your ancestors had acquired over time. Additionally, you could kill current disease (virus/bacteria) in the body by using the right frequencies, and this did not disturb any other body cells. This is why Floyd needed my doctor brother to help him arrive at the proper medical protocol for using his technology. Additionally, if you note in the Payroll Expenses attachment of this e-mail, a one-line item of expenditure is for AIDS-related materials in which Floyd and my doctor brother had a real interest. My doctor brother had an agreement with Floyd to build three medical interferometers which would all have a noble gas plasma inside them. I actually witnessed one of these devices in operation. At the end of the (approximately 20 inch long) tube-like structure you could feel a pulsing being emitted at the end of the tube on to whichever part of the body you wanted treated. My doctor brother had ordered two Interferometers from Floyd which were about 4 feet long.

(2) The VTA energy device is probably the world's worst weapon. Floyd claimed that like Nicola Tesla, you could cause "artificial earthquakes" - besides destroying buildings. As I understood from people in the intelligence world, which we de-briefed after we saw the device operate, three countries have what is called the "Tesla Cannon"; Russia, America and I never found out who the third country was. As mentioned earlier, this energy source is what disabled Floyd's VTA equipment over many months until he got his "buffer" built into his device. Further, this is why the Federal Government had such an interest in what we were doing with Floyd during the time we spent with him.

(3) The device was an Energy source for the home (could change negative energy to positive energy). It was also an energy source for the car and many other purposes. The cost of building one of these energy devices was only about US $200.00 - incredible!

Description of the VTA device:

On the day that we finally got to see the device operating, my doctor brother and I had finally convinced my electronics brother to accompany us to Sherman Oaks, California to see the demonstration. My doctor brother and I had made ALL the preliminary trips to see Floyd minus our electronics brother because he was literally a "doubting Thomas", being heavily involved in the electronics field and full of Maxwell's Theories of electronics, etc. Yes, you could say that he was a traditional electronics person. But, for this reason, we needed my electronics brother to be our DEBUNKER in case the device was not what it was portrayed to be. We had one other witness "Gary", an associate of mine who was to bring in

the venture capital funding if the device proved to be as good as claimed.

The day when we witnessed the VTA device operating is a day which I shall never forget. To actually see a device working, which cost only $200 dollars to make and which could create all the clean energy you would ever need, was "awesome". I know I have been "altered" ever since knowing that such a device existed. Now for a brief description of the Device:

These are not exact measurements but only approximations. The device was on what I believe to be "Plexi glass" (acrylic). Nothing was hidden. You could see everything, top and bottom through the plastic. The Plexi glass structure was approximately 18" square. We were allowed to pick-up the device and carry it around Floyd's living room so you could see that there were no other electrical connections to it.

On top of the Plexi glass case there were three toroidal coils wound with thin windings of varnished copper wire. There were two barium ferrite bar magnets (approx. 7"x 3"x 1/2"). Present was a volt meter which displayed 120v when the device was turned on. Also, there was an ampere meter which measured the electrical currents flowing when Floyd switched different things on-and-off during the demonstration. The items used for load demonstration included the burner part of the stove, a hair dryer, a fan, and five one-hundred watt globe lights. The fascinating thing to me about the light demo was that the lights had a glow like the overhead lights in your kitchen - a very soft, COOL appearance. Not the look of a traditional bright light bulb such as you have in your lamp on a traditional night stand.

I forgot to mention that the device was started by attaching a 9-volt battery which, I understand, started the magnetic flux in motion. Floyd would then connect the "pigtail" on the device and it would become just one circular energy unit.

As Floyd put more load on the device, the ambient temperature around the device (coils) would start to get lower. Additionally, depending on how much load you added, the device would start to lose some of its weight and you then had levitation beginning to take place. I should note at this point that on one meeting with Floyd, his wife Rose, used some expletives when telling how one day, Floyd kept adding more-and-more load to the device and he almost "brought down" the Apartment Complex he lived in at Sherman Oaks. He turned off the equipment, went out on his patio and pretended that it was a California Earthquake! His neighbours never did know what he had in his apartment. In this respect, I never did find out what the big piece of equipment was in his bedroom. It literally stretched from the ceiling to the floor. It was so heavy that the floor was bowed-in and sunken and that "big sucker" had a growling noise when it was on - I never did find out what it was. It was big like some kind of transformer.

The Rest of The Story:

You are probably wondering what the article on Ron Brandt is about. It's a long story, but after I moved Ron and his laboratory all the way from the mouth of Zion's National Park to "someplace" Oregon to hide him out - he was using "Tachyon Beams" with his medical equipment and after only a couple of minutes the "Black Helicopters" would show up - soooo at my doctor brother's request I moved Ron to Oregon. At the time I thought Ron was a "real flake" because when I helped him forward his mail from a small town in Southern Utah, he asked me how to spell the word "electric" so he could put in the full address of "Brandt Electric". Further, Ron said he was only here on this Earth until 2012 - It was now 1987-88 - and then he had to leave to go to another planet! I now wanted to shoot my doctor brother who got me into this whole moving-Ron thing! My doctor brother told me that Ron had to move fast because Ron had told him that an earthquake was coming in the next few days - Right!

Well, guess what happened a few days later? The largest earthquake in many years in that particular location took place and it even wiped out the hot springs at the Resorts along the Virgin River which runs through Zion's National Park and through the small town of Virgin where Ron lived. I since found out that Ron had invented earthquake equipment along with Philo T. Farnsworth's (Inventor of Television) grandson and six months ahead, they had actually predicted the previous great earthquake in California and their prediction was off by only six minutes! The Government is insisting that they want the equipment, so that is one of the reasons for everyone "hiding out".

Now, why am I giving you all this preliminary information regarding Ron Brandt? Well it seems that Ron has a Magnet Motor which weighs only 75 pounds and which can generate power equivalent to that of a 300 horsepower internal combustion engine. Also, the motor can be a retro-fit in any existing car without the need to design a whole new car. This is the connection I will explain later regarding Ron who could not even spell "electric" and Floyd who was placed 3rd in all the inventions to ever come out of MIT - All I can say is "WOW"!

EVENTS SURROUNDING FLOYD'S DEATH:

I will now leave it up to you to decide whether or not Floyd died of natural causes or was "taken out" by some person, group, or some Government.

In the summer of 1994, my doctor brother suddenly "passed out" at one of our Venture Capital meetings and was rushed to the hospital. After an MRI of his head, it was discovered that he had a brain tumour and it was of the worst kind (very fast growing). This seemed impossible as my doctor brother had always monitored his body daily as he did an occasional experiment on himself with certain medicines. By 11th November 1994, my doctor brother had died. He told us prior to death that "they" (whoever "they" were) had succeeded in placing the fastest growing cancer tumour into his brain - How? - I have no idea! I never did find out. What is important to the free-energy field was that my doctor brother was in daily contact with Floyd and his Associates regarding the energy devices. I was not that important and basically only accompanied my doctor brother to meetings and kind of "got lost in the woodwork". Intellectually, I really was not a threat to anyone. I was only there at meetings to help acquire venture capital.

On the very day that my doctor brother died, my electronics brother and I were at the home of John, (Floyd's Associate from NASA) who for some strange reason had followed my brothers and I to our home city where we lived, bought a home and took up residence there. We did not complain as he was our go-between with Floyd. But the move still seemed strange to me. The reason my electronics brother and I were with John is that John had arranged a conference call with Floyd and us, to see if there was a possibility for Floyd to make some type of energy device which could power the magnet motor that Ron Brandt had. My brothers and I had all the contractual rights to Ron's Magnet Motor which could be used in any car. I thought to myself that now I can really find out how "real" Ron (who could not even spell "electric") was when I matched him up with Floyd from MIT. I could not believe what I heard as Floyd and Ron conversed at the highest electronic levels - "who the 'hell' is Ron?" I thought. Floyd agreed that he would have no problem doing the prototype for Ron's Magnet Motor to power the car.

Floyd mainly worked with my electronics brother on this project as Floyd needed old vacuum tubes which my electronics brother had to acquire for the device and my electronics brother was a real "bench" person which Floyd seemed to favour over academic Electrical Engineers.

During the Spring of 1995, while Floyd was working on our energy device for the car, John (from NASA) and Floyd were elated that there was supposed to be an announcement from the White House regarding Floyd's VTA Energy Device. It seems that Floyd was a past friend of Senator John Glen (the former NASA astronaut) and he had given Glen one of the energy devices. Unfortunately, Glen gave the device to the Department of Energy, who, according to Floyd, passed the device on to General Motors. Floyd was furious and as I understood Floyd was then going to sue GM for two hundred million dollars. As far as I know Floyd never got the device back. I will always remember the extreme disappointment on the faces of Floyd and John when they realised that the trip to Washington DC for the announcement, was not going to take place.

In July 1995, Floyd let us know that the Energy Device was finished and we were to take possession of it. Floyd now lived in Desert Palms, California and that is where we would pick it up. After much thought, we decided we better not board a plane with the device as we were not sure of any magnetic effects on the instruments of the plane in having it transported - it was new technology which still had many questions to be answered. Instead, we decided to drive our car to Desert Palms and bring the device back ourselves.

Floyd called us the day before we were to leave and asked us if he could keep the device for a couple of extra days. He said he had "someone" coming (I thought he said China) and wanted to show them the device. We said ok, we would plan to pick it up when he was done.

A day later, at about 7:00 am Pacific time, there was a frantic call from Floyd's wife Violet (Floyd's wife Rose had died and he had re-married) to my electronics brother's house. My electronics brother was not at home and my sister-in-law, his wife, took the call from Violet. Violet was very traumatised when she told my sister-in-law that Floyd was dead. There was a lot of shouting going on in the background. The people who were there claimed they were from the FBI and that Floyd's equipment belonged to them. Rose was extremely confused with the death of Floyd and people she had never seen before taking all the equipment out of her house to waiting vans. She asked my sister-in-law what to do and my sister-in-law had NO idea as she was not aware of what my brothers and I had going on!

Violet also said that about 5:00 pm the previous night, two men whom she had never seen before, showed up to see Floyd. Floyd was with them for a period of time and then they left. At about 8:00 pm, Floyd was having a cup of coffee when he fell out of the chair on to the floor. She called for an ambulance and when they arrived they would not let her ride with them. Violet was 75 years old and didn't drive. About twenty minutes later the ambulance called back to Violet and told her they didn't think Floyd was going to "make it"!! As I understand it, Floyd's body was cremated. How soon afterwards, I don't know. The end result for my brothers and I is that ALL of our energy equipment that Floyd made for us was taken - By Whom??

Who were the two men who met with Floyd a few hours before his death? Was anything put in Floyd's coffee by these men? Violet said she had never seen them before and they seemed strange!
Why could Violet not go with her husband in the ambulance? I have seen it happen many times when family is allowed, especially where age is concerned!

How did the FBI (if that is who they were) know that Floyd was dead and show up in the very early morning (about 6:00 am) just hours after he died late at night?

YOU BE THE JUDGE - ALL I KNOW IS THAT ALL OF OUR ENERGY DEVICES (MEDICAL AND CAR-MAGNET MOTOR) ARE GONE!!! WHERE ARE THEY AND WHO ARE THE ONES RESPONSIBLE FOR TAKING THEM ??

Here are some of the known facts about Floyd's energy device:

The invention is a unified-field device and so combines both electromagnetic and gravitational effects in the same unit. For a tiny power input of just 0.31 milliwatt, the unit produces over 500 watts of output power, which is an energy gain of more than 1,500,000. The prototype, has no moving parts, is about 6" x 6" x 4" in size and taps an inexhaustible source of energy. To date, up to one kilowatt of power has been produced in actual tests which required only tiny input power to make the device operate.

Our normal day-to-day energy is "positive energy". The energy produced by Floyd's device is "negative energy" but in spite of this, it powers ordinary equipment, producing light and heat as normal. A device like this has to have a major impact on the world as we know it, because:

1. **It can be easily built. The components are quite ordinary and the cost of the materials in the demonstration prototype was only a few hundred US dollars and it was constructed in just a few hours, using simple tools and equipment.**

2. **The test results are so impressive that there can be no question of errors of measurement when the energy gain is of the order of 1,500,000 times.**

3. **It demonstrates with laboratory precision that the 'law' of Conservation of Energy does not appear to apply during the operation of this device, which is something which most scientists have difficulty in accepting.**

The device has very high performance. When a 1-milliwatt 60Hz sine wave is fed into it, the out put powers 500 watts of standard mains-voltage light bulbs, producing both heat and light. The device has a positive-feedback loop so it's gain is depends directly on the output load and the input power remains unchanged. So to increase the output power, all that is necessary is to connect extra light bulbs or equipment across the output.

When a motor was connected in addition to the light bulbs, the motor ran perfectly well under load and the light bulbs remained as bright as ever. Because it is a "cold electricity" device, the wires feeding the load can be very much smaller in diameter than would be normal for the load and these wires run cold at all times. When the power hits the resistance of the filaments of the light bulbs, it converts into conventional "hot electricity" and the filaments perform in exactly the same way as they do when powered by "hot electricity".

In 1988, Floyd produced a paper which he considered to be very important. The following text is an attempt to reproduce the content his highly mathematical style of presentation. If you are not into complicated mathematical presentations, then just move on past and don't worry about the following technical material, or alternatively, take a quick skim through it and don't bother with the maths. Floyd says:

What is thought of as "empty space" actually contains almost everything in the universe. It is home to all kinds of invisible energy fields and is seething with all kinds of very real forces.

Every kind of matter produces an energy field and these energy fields interact with each other in many complicated ways, producing all sorts of additional effects. These energy fields are the "stuff" of space, or as it is sometimes described, "the virtual vacuum". Space is packed full of all sorts of things but because it does not contain air, we tend to think that there is nothing at all in it. Most people think that "vacuum" means "without air" but when scientists speak of space as "the vacuum" they do not mean that at all, and they use the word "vacuum" to describe to describe (loosely speaking) the place which is between the stars and planets of the universe, and Floyd refers to that vast place as "the vacuum", so please don't think that it has anything to do with air, as it definitely doesn't.

Floyd says: We all think that we know what light is, but the reality is that a particle of light is nothing more than a large interference in the electromagnetic field. Unless it interacts with matter or with another field, any electromagnetic field with not be changed in any way by the vacuum. Electromagnetic fields are a fundamental part of the structure of the vacuum itself. The whole universe is permeated by a constant magnetic field. That field is made up of countless numbers of North and South pole magnets in a completely random scatter.

Einstein has pointed out that $E = mC^2$ which is one way of saying that energy and matter are interchangeable (or are two different faces of the same thing). The energy everywhere in the universe is so great that new particles of matter pop into existence and drop back into their energy form many trillions of times per second. Actually, they exist for such a very short time that calling them "particles" is not really appropriate, so perhaps "virtual particles" might be a better description.

However, if we generate a moving magnetic field, it alters the random nature of this energy in the tiny part of the vacuum where we happen to be, and the vacuum energy becomes much less random and allows a very large amount of vacuum energy to be drawn into our equipment and do what we think of as "useful work" - producing heat and light, powering motors and vehicles, etc. This was proved in laboratory experiments during the week of 19th June 1988 and it is the underlying operating principle of my "Phase-Conjugated Vacuum Triode" device.

The energy produced by this device is "negative energy" which is the reverse of the energy with which we are familiar. The spark caused by a short-circuit in a negative energy system is excessively bright and cold and it produces a barely audible hiss with no explosive force. Melting of wires does not occur and this type of negative current passes through the human body with only the feeling of a chill.

Wires which carry a lot of negative energy remain cool at all times and so tiny wires can feed equipment with hundreds of watts of power. This has been demonstrated in the laboratory and the source of energy is unlimited as it is the virtual vacuum of space itself.

The Nature of Space:

Space itself is the ability to accommodate energy. Consider for a moment, the following illustration:

> **A signal (energy) is transmitted from point "A" to point "B" which are separated by a finite distance. Consider three periods of time:**
>
> **1. The signal is launched from point A.**
> **2. The signal resides in the space between point A and point B.**
> **3. The signal arrives at point B.**
>
> **If 3. occurs simultaneously with 1. we say that the signal has travelled at infinite velocity. If that were the case, then the signal never resided in the intervening space and therefore there must be no space between point A and point B and so both points A and B must be at the same location. For real space to exist between the two points, it is necessary that a signal moving between them has to get "lost" to both points, that is, out of touch with both points for a finite period of time.**

Now, we know that for real space to exist between two points, a signal passing between them has to move at a finite speed between them and if it can't do that, then there can't be any space between them. If space can't accommodate a signal passing between two points, then it has no function and no reality. We are left then with the only real space, the home of the real and virtual vacuum - space which supports a finite, non-zero signal velocity.

A similar argument applies to the impedance of space. A medium can only accommodate positive energy if the medium resists it to a reasonable degree. Neither an infinitely strong spring nor an infinitely weak spring can absorb energy by being compressed. Neither an infinitely large mass nor an infinitely small mass can absorb or accommodate energy imparted by a collision and the same holds true for space.

Energy cannot enter a space of zero impedance any more than a force can bear on a mass of zero magnitude. Similarly, energy could not enter space which has an infinite impedance. It follows therefore, that real space must have:

1. Finite propagation velocity and
2. Finite impedance.

Another way of looking at this is instead of considering the actual speed of propagation of a signal through space, to consider the length of time "t" which it takes the signal to pass through that part of space. We can think of a section of space as being, say, 1 nanosecond wide if it takes a signal 1 nanosecond to traverse it. That is, the energy or signal entering that part of space, leaves it again 1 nanosecond later. Signal propagation speed in the space in which we live is at the speed of light.

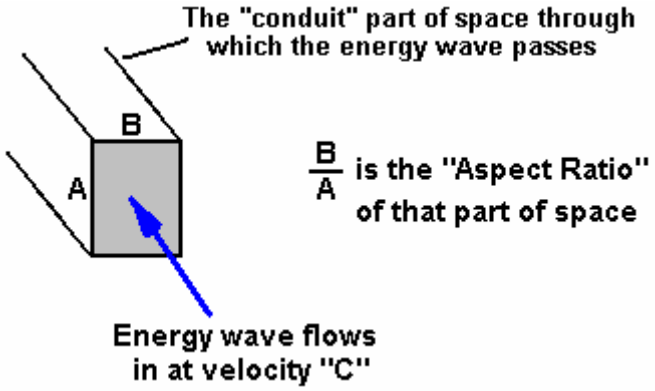
General Description of Energy Transfer:
Consider energy flowing straight and level down a transmission line. The energy does not "know" the width of the channel through which it is passing. If the energy flow reaches a point where the conductivity of the channel lowers but the size and shape of the channel remain the same, then not as much energy can flow and some gets reflected back along the channel. The energy current will not "know" if (a) the conductivity has changed or (b) the geometry has changed. The energy current can change direction very easily and so as far as it is concerned, the change caused by (a) is equivalent to the change caused by (b).

The channel through which the energy flows has width and height and the width divided by the height is called the "aspect ratio" of the channel. Energy current has an aspect ratio and if that aspect ratio is forced to change, then some of the flowing energy will reflect so as to keep the *overall* aspect ratio unchanged.

The aspect ratio of energy current is much like the aspect ratio of space itself. While the aspect ratio of space itself can change, it's fundamental velocity of "C" the speed of light in space can't really change. That speed is just our way of visualising time delay when energy resides in a region of space. Uniform space has only two parameters:

(1) Aspect ratio and
(2) Time delay

Aspect ratio defines the shape (but not the magnitude) of any energy flow which enters a given region of space. Velocity or length define the time during which that energy can be accommodated in a region of space.



Does an energy flow travel unimpeded through an interface, or does a large part of it get reflected? Space has quiet zones through which energy glides virtually unreflected. It also has noisy zones where the energy current becomes incoherent, bounces around and splits apart. These noisy zones in space either have either rapidly changing geometry or rapidly changing impedance.

Electromagnetic Energy:

The rate of flow of energy through a surface can be calculated using "E" the Electric field, and "H" the Magnetic field intensity. The energy flow through space is E x H per unit area (of it's "conduit's" cross-sectional area) and the energy density is E x H / C where C is the speed of light in space.

If there happen to be two signals of exactly the same strength, passing through each other in opposite directions in such a way that their "H" fields cancel out, then if each has a strength of E/2 and H/2, the

energy density will be E x H / 2C and it will have the appearance of a steady E-field. In the same way, if the E fields cancel out, the result will appear to be a steady "H" field.

Modern physics is based on the faulty assumption that electromagnetics contains two kinds of energy: electric and magnetic. This leads to the Baroque view of physical reality. Under that view, energy seems to be associated with the square of the field intensity, rather than a more reasonable view that it is directly to the field intensity. It is worth remembering that neither Einstein nor most modern physicists were, or are, familiar with the concept of "energy current" described here. However, their work still survives by ignoring the energy current concept, scalar electromagnetics, the works of Tom Bearden, kaluza-Klein and others who dispute Heaviside's interpretations of Maxwell's equations.

The Fallacy of Displacement Current:

Conventional electromagnetic theory proposes that when an electric current flows down a wire into a capacitor, it spreads out across the plate, producing an electric charge on the plate which in turn, leads to an electric field between the plates of the capacitor. The valuable concept of continuity is then retained by postulating a displacement current "after Maxwell". This current is a manipulation of the electric field "E" between the plates of the capacitor, the field having the characteristics of electric current, thus completing the flow of electricity in the circuit. This approach allows Kirchoff's laws and other valuable concepts to be retained even though superficially, it appears that at the capacitor there is a break in the continuous flow of electric current.

The flaw in this model appears when we notice that we notice that the current entered the capacitor at only one point on the capacitor plate. We are then left with the major difficulty of explaining how the electric charge flowing down the wire suddenly distributes itself uniformly across the entire capacitor plate at a velocity in excess of the speed of light. This paradoxical situation is created by a flaw in the basic model. Work in high-speed logic carried out by Ivor Catt has shown that the model of lumped capacitance is faulty and displacement current is an artefact of the faulty model. Since any capacitor behaves in a similar way to a transmission line, it is no more necessary to postulate a displacement current for the capacitor than it is necessary to do so for a transmission line. The removal of "displacement current" from electromagnetic theory has been based on arguments which are independent of the classic dispute over whether the electric current causes the electromagnetic field or vice versa.

The Motional E-Field:

Of all of the known fields; electric, magnetic, gravitational and motional E-field, the only ones incapable of being shielded against are the induced motional E-field and the gravitational field. The nature of the motionally-induced electric field is quite unique. In order to understand it more fully, we must start by discarding a few misleading ideas. When magnetic flux is moved perpendicularly across a conductor, an electromotive force ("e.m.f.") is electromagnetically induced "within" the conductor. "Within" is a phrase which comes from the common idea of comparing the flow of electric current within a wire to the flow of water in a pipe. This is a most misleading comparison. The true phenomenon taking place has little been thought of as involving the production of a spatially- distributed electric field. We can see that the model's origins are likely to have arising from the operation called "flux cutting" which is a most misleading term. A better term "time-varying flux modulation" does not imply any separation of lines of flux. Truly, lines of flux always form closed loops and are expressed mathematically as line integrals.

It is a fallacy to use the term "cutting" which implies time-varying separation which does not in fact ever occur. A motionally-induced E-field is actually created within the space occupied by the moving magnetic flux described above. The field is there whether or not a conductor is present in the space. In terms of a definition, we can say that when magnetic flux of vector intensity B-bar is moved across a region of space with vector velocity V-bar, an electromagnetically induced electric field vector B x V appears in the space at right angles to both B-bar and V-bar. Therefore:

$$E = \text{B-bar} \times \text{V-bar} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

It is this field which is related to gravity and which is virtually unshieldable. This field may be called the Motional E-field. According to Tom Bearden, "It seems that the charged particles in the atom act like tiny magnets and their motion in the space surrounding the atom would create this motional E-field". The fields created by both the positive and negative charges would cancel to some degree, but due to the high orbital velocity of the negative electron relative to that of the positive proton, the induced field of the electron would dominate the resulting field. The field produced as a result of these charges would vary in proportion to the inverse square of the distance as gravity does. The field produced by the translational motion of the charges would vary inversely as the cube of distance. This concept totally unites the

electromagnetic and gravitational field theories and accounts for the strong and weak force within the atom.

Field Super-Position and the Vacuum Triode:

Electromagnetic induction with no measurable magnetic field is not new. It is well known that in the space surrounding a properly wound toroidal coil, there is no magnetic field. This is due to the superposition of the fields. However, when alternating current is surging through a transformer, an electric field surrounds it. When we apply the principle of super-position to the vacuum triode, it becomes more obvious how the device is operating.

The principle of super-position states that "in order to calculate the resultant intensity of superimposed fields, each field must be dealt with individually as though the others were not present" The resultant is produced by the vector addition of each of the fields considered on its own. Consider for a moment, the construction of the triode which includes two bi-filar coils located within the fields of two conditioned magnets. When the current in one half of the conductors in the coils (that is, just one strand of the twin windings in each coil) is increasing, both the current and the magnetic field follow the right-hand rule. The resulting motional E-field would be vertical to both and directed inwards. At the same time, the current in the other strand of each winding is decreasing and both the current and the magnetic field also follow the right-hand rule. The resulting motional E-field is again vertical to both, and directed inwards. So, the resultant combined field intensity is double the intensity produced by either one of the conductors considered on its own. Expressed mathematically, this is:

$$E = (B \times V) + (-B \times -V) \text{ or}$$

$$E = 2\,(B \times V) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2)$$

Where: E is the electric field intensity
B is the magnetic field intensity and
V is the electron drift velocity

(B x V), the first term in the equation, represents the flow of the magnetic field when the electrons are moving in one direction, while (-B x -V), the second term in the equation, defines the flow of the magnetic field when the electrons are moving in the other direction. This indicates that field intensity is directly proportional to the square of the current required by the load placed on the device. This is due to it's proportional relationship with the virtual value of the magnetic field which theory states is proportional to the current. Electrometer readings were always close to parabolic, indicating that the source was of infinite capacity. It was further determined through experiment, that the magnetic field does not change with temperature. Also, there is no reason yet identified, which would lead one to believe that electron drift velocity changes. It has been found remarkable that the vacuum triode runs approximately $20^{O}F$ below ambient.

Induced Electromotive Force - Positive Energy:

When an e.m.f. ("electromotive force") is applied to a closed metallic circuit, current flows. The e.m.f. along a closed path "C" in space is defined as the work per unit charge (that is, W / Q) done by the electromagnetic fields on a small test charge moved along path C. Since work is the line integral of Force ("F"), the work per unit charge is the line integral of force per unit charge (in Newtons per Coulomb) we have:

$$\text{e.m.f.} = \int_C F / Q \times \text{dtdl volts} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (3)$$

The scalar product "(F/Q) x dtdl" is the product of (F/Q) x Cosθ x dl where θ denotes the angle between the vectors F/Q and dl.

The electric force per unit charge is the electric field intensity ("E") in volts per metre. The magnetic force per unit charge is V x B where "V" denotes the velocity of the test charge in metres per second and "B" denotes the magnetic flux density in webers per metre squared. In terms of the smaller angle θ between V and B, the cross product of V and B is a vector having the magnitude VBSinθ. The direction of vector V x B is at right angles to the plane which contains vectors V and B in accordance with the right-hand rule (that is, V x B is in the direction of the thumb while the fingers curl through the angle θ from V towards B). Since the total force per unit charge is E + VB, the total e.m.f. in terms of the fields is:

$$\text{e.m.f.} = \int_C (E + V \times B)\, dt dl \quad \ldots\ldots\ldots\ldots\ldots\ldots (4)$$

It appears from equation (4) that the e.m.f. depends on the forward velocity with which the test charge moves along the path C. This, however, is not the case. If V and dl in equation (4) have the same direction, then their associated scalar product is zero. So, only the component of V which is not aligned with dl (that is, with $\theta = 0$), can contribute to the e.m.f. This component has value only if the differential path length dl has a sideways motion. So, V in equation (4), represents the sideways motion of dl, if there is any. The fields E and B in equation (4) could well be represented as functions of time as well as functions of the space co-ordinates. In addition, the velocity V of each differential path length dl, may vary with time. However, equation (4) correctly expresses the e.m.f. or voltage drop along path C as a function of time. That component of the e.m.f. consisting of the line integral V x B is the motional E-field since it has value only when path C is ,moving through a magnetic field, traversing lines of magnetic flux. For stationary paths, there is no motional E-field and the voltage drop is simply the integral of the electric field "E". Devices which separate charges, generate e.m.f.s and a familiar example of this is a battery which utilises chemical forces to separate charge. Other examples include the heating of a thermocouple, exposure of a photovoltaic cell to incident light or the rubbing together of different material to produce electrostatic charge separation. Electric fields are also produced by time-varying magnetic fields. This principle is already exploited extensively in the production of electrical power by the utility companies.

The line integral of electric field intensity "E" around any closed path "C" equals $-d\varphi/dt$ where φ represents the magnetic flux over any surface "S" having the closed path "C" as it's contour. The positive side of the surface S and the direction of the line integral around contour C, are related by the right-hand rule (the curled fingers are oriented so as to point around the loop in the direction of integration and the extended thumb points out the positive side of the surface S). The magnetic flux φ is the surface integral of magnetic flux density "B" as shown here:

$$\varphi = \iint_S B \times ds \quad \text{webers} \ldots\ldots\ldots\ldots\ldots (5)$$

In Equation (5), the vector differential surface "ds" has an area of ds and in direction, it is perpendicular to the plane of ds, projecting out of the positive side of that surface. The partial time derivative of φ is defined as:

$$\frac{\partial \varphi}{\partial t} = \iint_S \frac{\partial b}{\partial t} \times ds \quad \text{volts} \ldots\ldots\ldots\ldots\ldots\ldots (6)$$

This is referred to as the magnetic current through surface S. For a moving surface S, the limits of the surface integral in equation (6) are functions of time, but the equation still applies. It is important to clarify at this point, that when we evaluate the value of $d\varphi/dt$ over a surface which is moving in proximity to magnetic field activity, we treat the surface as though it were stationary for the instant under consideration. The partial time derivative of φ, is the time rate of change of flux through surface S, due only to the changing magnetic field density B. Any increase of φ due to the motion of the surface in the B-field, is not included in that calculation.

Continuing this discussion leads us to note that an electric field must be present in any region containing a time-varying magnetic field. This is shown by the following equation:

$$\oint_C E \times dl = \frac{-\partial \varphi}{\partial t} \quad \ldots\ldots\ldots\ldots\ldots\ldots (7)$$

In this equation, φ is the magnetic flux in webers out of the positive side of any surface having path C as its contour. Combining equations (7) and (4), we are able to calculate the e.m.f. about a closed path C as shown here:

$$\text{e.m.f.} = \oint_C E \times dl + (V \times B)\, dl \quad \ldots\ldots\ldots\ldots\ldots\ldots (8)$$

or in another form:

$$\text{e.m.f.} = \frac{-\partial \varphi}{\partial t} + \oint_C (V \times B)\, dl \quad \ldots\ldots\ldots\ldots\ldots\ldots (9)$$

So, the e.m.f. around a closed path consists in general of two components. The component dφ/dt is the variational e.m.f. and the second component is the motional E-field. In equation (9), (V x B)dl can, by means of a vector identity, be replaced with B x (V x dl)A. V is the sideways velocity of d: the vector V x dl has magnitude Vdl and a direction normal to the surface ds swept out by the moving length dl in time dt. Letting Bn denote the component of B normal to this area, we can see that the quantity -B x (V x dl) becomes -BnVdl and equation 9 can be re-written as:

$$\text{e.m.f.} = \frac{-\partial \varphi}{\partial t} + \oint_C Bn\ V\ dl \quad \ldots\ldots\ldots\ldots\ldots\ldots (10)$$

Clearly, the integral of BnV around the closed contour C with sideways velocity of magnitude V for each length dl traversed, is simply the time rate of change of the magnetic flux through the surface bounded by C. This change is directly due to the passage of path C through lines of magnetic flux. Hence, the complete expression for e.m.f. in equation (10) is the time rate of change of the magnetic flux over any surface S, bounded by the closed path C, due to the changing magnetic field and the movement of the path through the magnetic field. Equation (10) may be written:

$$\text{e.m.f.} = -d\varphi / dt \quad \ldots\ldots\ldots\ldots\ldots\ldots (11)$$

Note: The distinction between equations (7) and (11) is that equation (7) contains only the variational e.m.f. while equation (11) is the sum of the variational and motional e.m.f. values. In equation (7), the partial time derivative of magnetic flux φ is the rate of flux change due only to the time-varying magnetic field, while equation (11) includes the total time derivative of the rate of flux change due to the time-varying magnetic field and path C's passage through the magnetic field. If the closed path C is not passing through lines of magnetic flux, then equation (7) and equation (11) are equivalent.

It is also important to point out that dφ/dt in equation (11) does not necessarily mean the total time rate of change of the flux φ over the surface S. For example, the flux over surface S is bounded by the closed contour C of the left portion of the electric circuit shown in Fig.1.

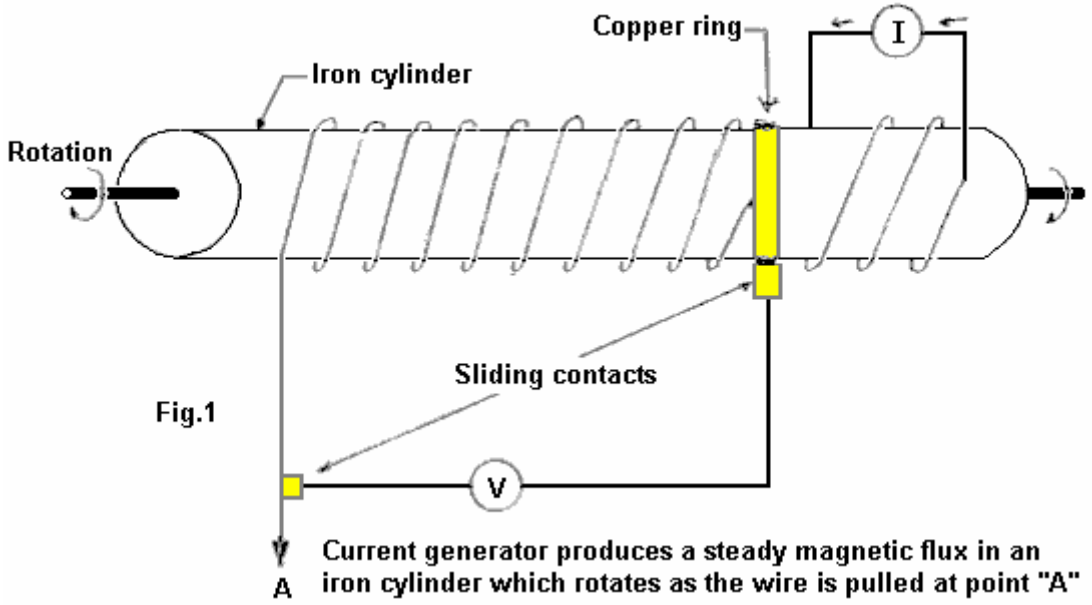


Fig.1

The flux is changing as the coil is unwound by the rotation of the cylinder, as illustrated. However, since B is static, there is no variational e.m.f. and since the conductors are not modulating lines of flux, there is

no motional e.m.f. Thus, dφ/dt in equation (11) is zero, even though the flux is changing with time. Note that dφ/dt was defined as representing the right hand part of the expression in equation (10) and dφ/dt must not be interpreted more broadly than that.

In the application of the present equations, it is required that all flux densities and movements are referred to a single, specified co-ordinate system. In particular, the velocities will all be with respect to this system alone and not interpreted as relative velocities between conductors or moving lines of flux. The co-ordinate system is selected arbitrarily and the magnitudes of variational and motional fields depend upon the selection.

Example 1:

A fundamental electric generator is shown in Figure 2:

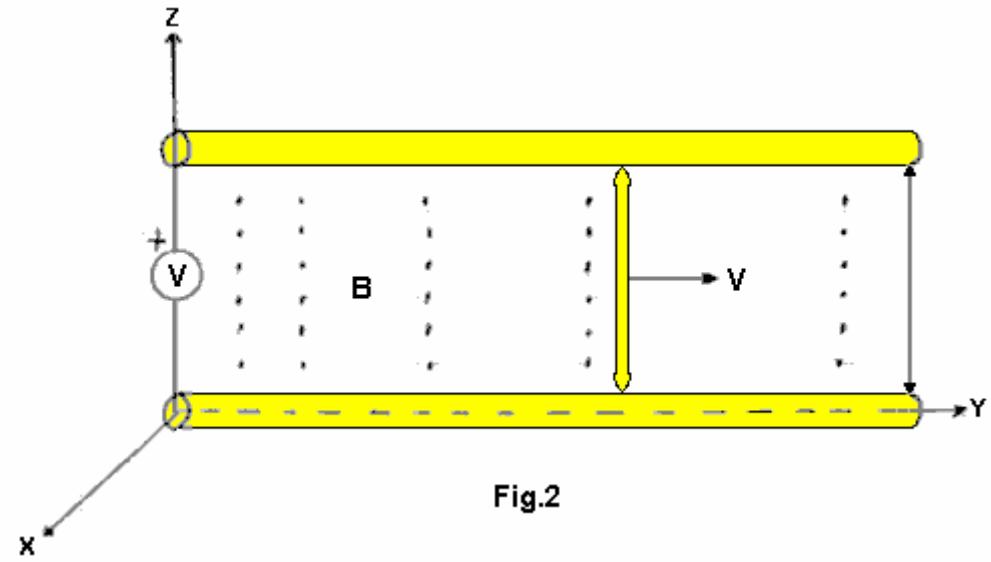


Fig.2

The parallel, stationary conductors, separated by distance "l", have a stationary voltmeter connected across them. The circuit is completed by a moving conductor connected to the parallel conductors by means of two sliding contacts. This conductor is connected at y = 0 at time t = 0, and it moves to the right at a constant velocity V = Vay. The applied flux B is represented by dots on Fig.2 and has a magnitude of $B = B_O$ CosBy Coswt ax. The unit vectors in the direction of the co-ordinate axes are ax, ay and az respectively.

Solution: Let S denote the plane rectangular surface bounded by the closed electric circuit, with a positive side selected as the side facing you. The counter-clockwise e.m.f. around the circuit is dφ/dt with φ signifying the magnetic flux out of the positive side of S (As ds = 1 dy ax). The scalar product B x ds is B_O l CosBy, Coswt dy; integrating from y = 0 to y = y gives:

$$\varphi = B_0 l \sin By_1 \cos wt \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (12)$$

With y_1 denoting the instantaneous y position of the moving wire. The counter-clockwise e.m.f. is found by replacing y with vt and evaluating dφ/dt. The result is:

$$\text{e.m.f.} = wB_0 l / B\sin Bvt \sin wt - B_0 l v \cos Bvt \cos wt \quad \ldots\ldots\ldots\ldots\ldots\ldots (13)$$

The variational (transformer) component is determined with the aid of equation (12) and is $wB_Ol/B\sin By$ sinwt where y = vt. This is the first component on the right hand side of equation (13).
Note: y_1 was treated as a constant when evaluating the partial time derivative of φ.

The motional E-field is the line integral of V x B along the path of the moving conductor. As V x B is $-B_O$ $vcosBy_1$ coswt ax and As dl is dz ax, evaluation of the integral $-B_O$ $vcosBy_1$ coswt dz from Z = 0 to Z = 1 results in a motional E-field of $-B_O$ $lvcosBv_1$ coswt. This component results from modulation of the lines of flux by the moving conductor. If the voltmeter draws no current, there can be no electromagnetic force on the free electrons of the wire. Therefore, the e.m.f. along the path of the metal conductors including the moving conductor, is zero.

Example 2:

Suppose the conductor with the sliding taps is stationary (V = 0) and it is located at $y = y_1$. Also, suppose that the magnetic field B is produced by a system of moving conductors which are not shown in Fig.2 and those conductors are travelling with a constant velocity V = Vay. At time t = 0, the magnetic field B is B_O sinBy ax. Determine the voltage across the voltmeter.

Solution: There is no motional E-field because the conductors in Fig.2 are at rest (stationary) with respect to our selected co-ordinate system. However, the magnetic field at points fixed with respect to the co-ordinate system is changing with time and as a result, there is a variational e.m.f. Since the B-field at time t = 0 is B_O sinBy ax and has a velocity of V = Vay, it can be calculated that the B-field as a function of time is BOsin[B(y-vt)] ax. This is verified by noting that an observer located at time t = 0 who is travelling at the constant velocity (V = Vay) of the moving current, would have a y co-ordinate of y = y + Vt and an accordingly different expression for B. He would observe a constant field where the magnetic current density is:

$$\frac{\partial B}{\partial t} = -BvB_o \cos B(y - Vt)\, ax$$

The counter-clockwise e.m.f. can be arrived at by taking the negative of an integral of the above expression for the rectangular surface bounded by the electric circuit with the positive side facing you, with the limits of zero and y. The resulting e.m.f. equals:

$$B_o lv[\sin B(y_1 - vt) + \sin Bvt]$$

which is the voltage across the meter.

Induced Motional Field - Negative Energy:

Conventional theory says that electric fields and magnetic fields are different things. Consider for a moment, a charge with an electric field around it. If the charge is moved, then a magnetic field develops and the moving charge constitutes a current. If an observer were to move along with the charge, then he would see no relative motion, no current and no magnetic field. A stationary observer would see motion, current and a magnetic field. It would appear that a magnetic field is an electric field observed from a motional reference frame. Similarly, if we take a mass with a gravity field around it, and we move the mass and create a mass current, a new field is also created. It is a different kind of gravity field with no source and no sink. It is called the "Protational field" and is also known as the "Lense-Thirring Effect". This field and it's governing principles will form the basis for future anti-gravitational devices (see figures 1 to 4).

Fig.3

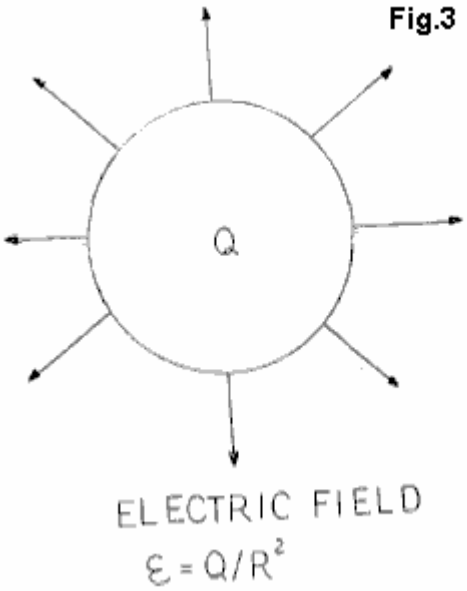
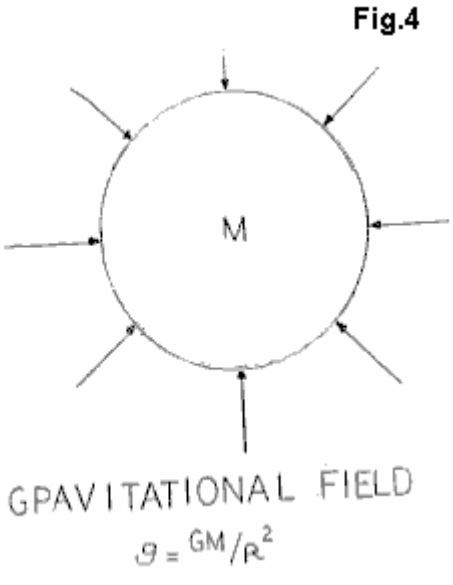


Fig.4



Within the confined are of the Vacuum Triode box, the space-time continuum is reversed by the fields which are produced in the presence of excited coherent space flux quanta. These quanta have been attracted form, and ultimately extracted from the virtual vacuum, the infinitely non-exhaustible Diac Sea. For a more detailed mathematical format see Tom Bearden's paper "The Phase Conjugate Vacuum Triode" (23rd April 1987). Much of the theory which likely applies to the vacuum triode has been developed in the field of phase-conjugate optics.

With regards to over-unity phenomena, it is important to note that so long as positive energy is present in a positively-flowing time regime, then unity and over-unity power gains are not possible. The summation of the losses due to resistance, impedance, friction, magnetic hysteresis, eddy currents and windage losses of rotating machinery will always reduce overall efficiency below unity for a closed system. The laws of conservation of energy always apply to all systems. However, the induced motional E-field changes the system upon which those laws need to be applied. Since the vacuum triode operates in more than four dimensions and provides a link between the multi-dimensional reality of the quantum state and the Dirac Sea, we are now dealing with an open-ended system and not the "closed system" within which all conservation and thermodynamic laws were developed.

To achieve unity, the summation of all magnetic and ohmic losses must equal zero. To achieve this state, negative energy and negative time need to be created. When this is achieved, all ohmic resistance becomes zero and all energy then flows along the outside of conductors in the form of a special space field. Negative energy is fully capable of lighting incandescent lights, running motors and performing all of the functions of positive energy tested to date. When run in parallel with positive energy however, cancellation (annihilation) of opposing power types occurs. This has been fully tested in the laboratory.

Once unity has been achieved and the gate to the Dirac sea opened, over-unity is affected by loading the open gate more and more, which opens it further to the point where direct communication / interaction with the nucleus of the atom itself is achieved. Output of the vacuum triode is not proportional to the excitation input as the output produced by the device is directly proportional to the load which is placed on it. That load is the only dependent variable for device output. The triode's output voltage and frequency always remains constant due to the conditioning of the motional E-field in the permanent magnets and the small regulated excitation signal which is provided through a small oscillator. Regulation remains constant and the triode output looks into an in-phase condition ($\cos\theta$=1 Kvar=1) under all load characteristics.

the vacuum triode is a solid-state device consisting of conditioned permanent magnets capable of producing a motional field. This field opens the gate to the Dirac Sea from where negative energy flows into the triode's receiving coils. The coils are wound with very small-diameter wire but in spite of that, they are capable of producing more than 5 kilowatts of useful power. This in itself, is a clear indicator that the type of electrical energy collected by the device is not conventional electrical energy. The wire sizes used in the construction of the device would not be capable of carrying such large currents without excessive heat gain, however, the triode's coils actually run cooler when loaded at 5 kilowatts.

The fundamental magnets have been broken free of the binding forces which constrained them to be steady-state single-pole uniform magnetic flux devices. They are now able to simply support mass, as demonstrated with the transformer steel illustration. They can now easily be made to adopt a dynamic motional field by applying a tiny amount of excitation. Specifically, 1 milliamp at 10 volts (10 milliwatts) of excitation at 60 Hz enables the coils of the triode to receive from the Dirac Sea, more than 5,000 watts of usable negative energy. It has not yet been determined how much more energy can be safely removed.

Meguer Kalfaian's Energy Generator

There is a patent application which has some very interesting ideas and claims. It has been around for a long time but it has not been noticed until recently. Personally, I get the impression that it is more a concept rather than a solidly based prototype-proven device, but that is only my impression and you need to make up your own mind on the matter.

Patent Application GB 2130431A **31st May 1984** **Inventor: Meguer Kalfaian**

Method and means for producing perpetual motion with high power

ABSTRACT

The perpetual static energies, as provided by the electron (self spin) and the permanent magnet (push and pull) are combined to form a dynamic function. Electrons emitted from a heated coil **F** are trapped permanently within the central magnetic field of a cylindrical magnet **M5**. A second magnet **M6**, in opposite polarity to the poles of the electrons causes polar tilt, and precession. This precession radiates a powerful electromagnetic field to a coil **L** placed between the cylindrical magnet and a vacuum chamber **C** - wound in a direction perpendicular to the polar axes of the electrons. Alternatively, the electromagnetic radiation is emitted as coherent light. The original source of electrons is shut off after entrapment.

SPECIFICATION

Method and means for producing perpetual motion with high power. This invention relates to methods and means for producing perpetual motion. An object of the invention is, therefore, to produce useful perpetual motion for utility purposes.

BRIEF EMBODIMENT OF THE INVENTION

The electron has acquired self spin from the very beginning of its birth during the time of creation of matter, and represents a perpetual energy. But self spin alone, without polar motion is not functional, and therefore, useful energy cannot be derived from it. Similarly, the permanent magnet represents a source of perpetual energy, but since its poles are stationary, useful energy cannot be derived from it.

However, the characteristics of these two types of static energies differ one from the other, and therefore the two types of energies can be combined in such a manner that, the combined output can be converted into perpetual polar motion.

In one exemplary mode, a cylindrical vacuum chamber having a filament and a cathode inside, is enclosed within the central magnetic field of a cylindrical permanent magnet, the magnetisation of which can be in a direction either along the longitudinal axis, or from the centre to the circumferential outer surface of the cylinder. When current is passed through the filament, the electrons emitted from the cathode are compressed into a beam at the centre of the cylindrical chamber by the magnetic field of the cylindrical magnet. Thus, when the current through the filament is shut off, the electrons in the beam remain permanently trapped inside the magnetic field.

In such an arrangement, the poles of the electrons are aligned uniformly. When a second permanent magnet is held against the beam in repelling polarity, the poles of the electrons are pushed and tilted from their normal longitudinal polar axes. In such tilted orientations, the electrons now start wobbling (precessing) in gyroscopic motions, just like a spinning top when it is tilted to one side. The frequency of this wobbling (precessional resonance) depends upon the field strengths of the two magnets, similar to the resonance of the violin string relative to its tensional stretch. The polar movements of the electrons radiate an electromagnetic field, which can be collected by a coil and then converted into any desired type of energy. Because of the uniformly aligned electrons, the output field is coherent, and the output power is high.

Observed examples upon which the invention is based:

The apparatus can best be described by examples of a spinning top in wobbling motion. Thus, referring to the illustration of **Fig.1**, assume that the spinning top **T** is made of magnetic material, as indicated by their pole signs (**S** and **N**). Even though the top is magnetic, the spin motion does not radiate any type of field, which can be received and converted into a useful type of energy. This is due to the known fact that, radiation is created only when the poles of the magnet are in motion, and in this case, the poles are stationary.

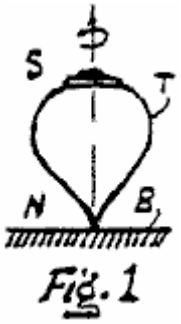


Fig. 1

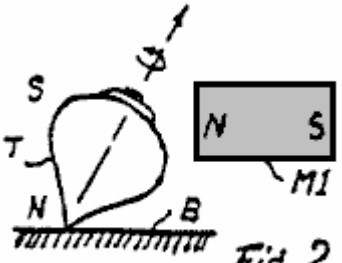


Fig. 2

When a magnet **M1** is held from a direction perpendicular to the longitudinal polar axis of the top, as shown in **Fig.2**, the polar axis of the top will be tilted as shown, and keep on spinning in that tilted direction. When the magnet **M1** is removed, however, the top will try to regain its original vertical posture, but in doing so, it will wobble in gyroscopic motion, such as shown in **Fig.3**. The faster the top spins, then the faster the wobbling motion will be.

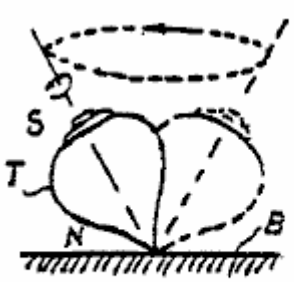


Fig. 3

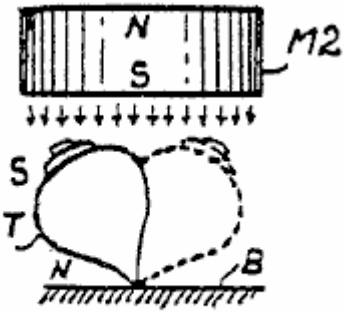


Fig. 4

The reason that the top tilts angularly, but does not wobble when the magnet **M1** is held from horizontal direction, is that, the one-sided pull prevents the top from moving away from the magnetic field for free circular wobble. Instead of holding the magnet **M1** from the side of the top, we may also hold the magnet from a direction above the top, as shown in **Fig.4**. In this case, however, the polar signs between the magnet and the top are oriented in like signs, so that instead of pulling action, there is pushing action between the magnet and the top - causing angular tilt of the top, such as shown in **Fig.4**. The pushing action of the magnetic field from above the top is now equalised within a circular area, so that the top finds freedom to wobble in gyroscopic rotation.

The important point in the above given explanation is that, the top tries to gain its original vertical position, but it is prevented from doing so by the steady downward push from the static magnetic field of magnet **M2**. So, as long as the top is spinning, it will wobble in a steady state. Since there is now, polar motion in the wobbling motion of the top, this wobbling motion can easily be converted into useful energy. To make this conversion into perpetual energy, however, the top must be spinning perpetually. Nature has already provided a perpetually spinning magnetic top, which is called, "the electron" - guaranteed to spin forever, at a rate of 1.5×10^{23} (one hundred fifty thousand billion billion revolutions per second).

BRIEF DESCRIPTION OF THE DRAWINGS

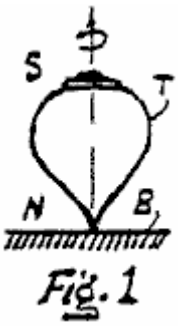


Fig. 1

Fig.1 illustrates a magnetic spinning top, used to describe the basic principles of the invention.

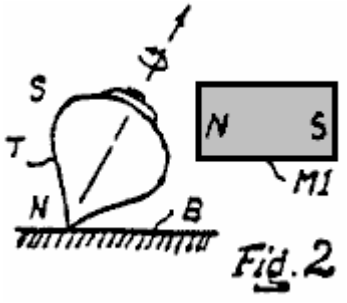


Fig. 2

Fig.2 illustrates a controlled top for describing the basic principles of the invention.

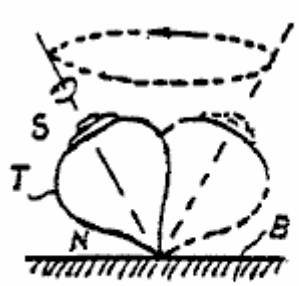


Fig. 3

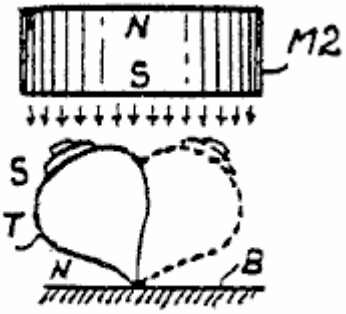


Fig. 4

Fig.3 and **Fig.4** illustrate spinning tops in wobbling states for describing the basic principles of the invention.

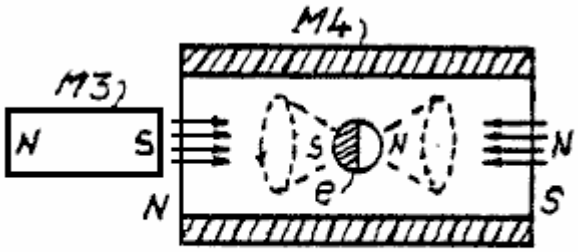


Fig. 5

Fig.5 shows how an electron can be driven into a wobbling state under the control of permanent magnets.

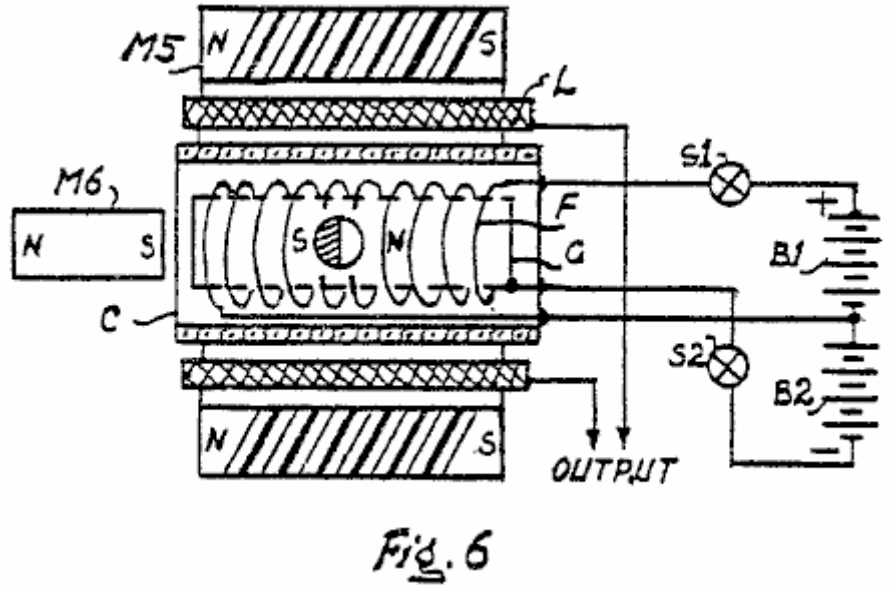


Fig. 6

Fig.6 is a practical arrangement for obtaining perpetual motion.

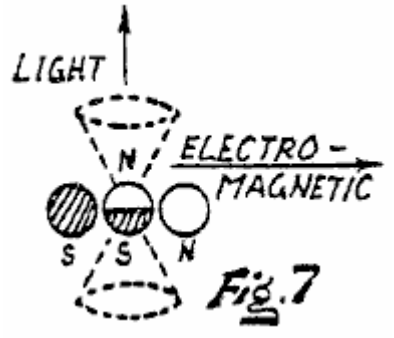


Fig. 7

Fig.7 shows a natural atomic arrangement for obtaining precessional resonance.

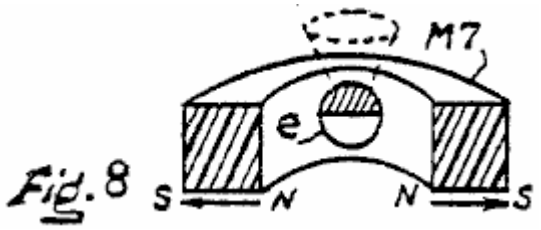


Fig. 8

Fig.8 shows a different type of electron trapping permanent magnet, to that used in **Fig.6**.

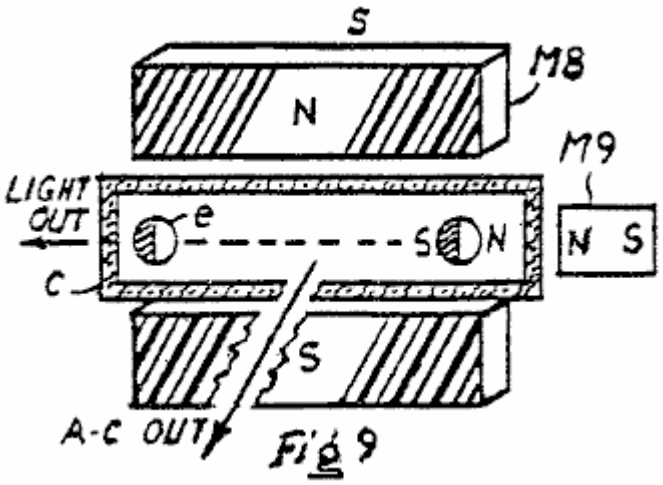


Fig 9

Fig.9 is a modification of **Fig.6**; and

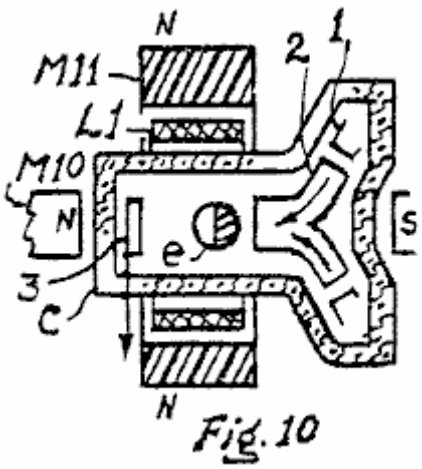


Fig. 10

Fig.10 is a modification of the electron trapping magnets, used in **Fig.6**.

BEST MODE OF CARRYING OUT THE INVENTION

Referring to the exemplary illustration of **Fig.4**, the spinning top **T** is pivoted to the base B by gravity.

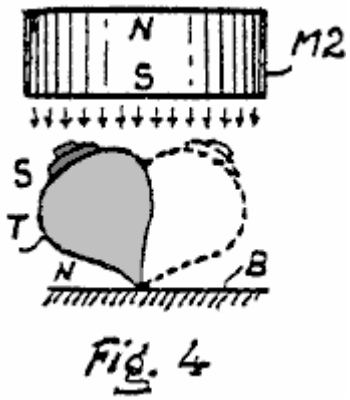


Fig. 4

In the case of the electron, however, it must be held tightly between some magnetic forces. So, referring to the illustration of **Fig.5**, assume that an electron **e** is placed in the centre of a cylindrical magnet **M4**. The direction of

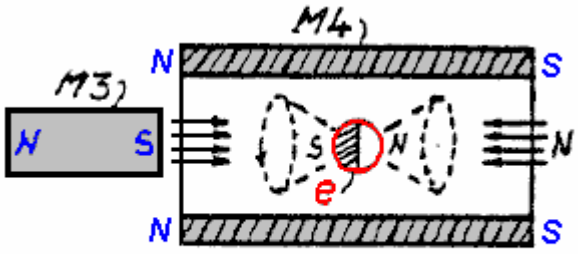


Fig. 5

magnetisation of the magnet **M4**, and the polar orientation of the electron **e** are marked in the drawing. In this case, when a permanent magnet **M3** is placed at the open end of the cylindrical magnet **M4**, the electron **e** will precess, in a manner, as described by way of the spinning top. The difficulty in this arrangement is that, electrons cannot be separated in open air, and a vacuum chamber is required, as in the following:

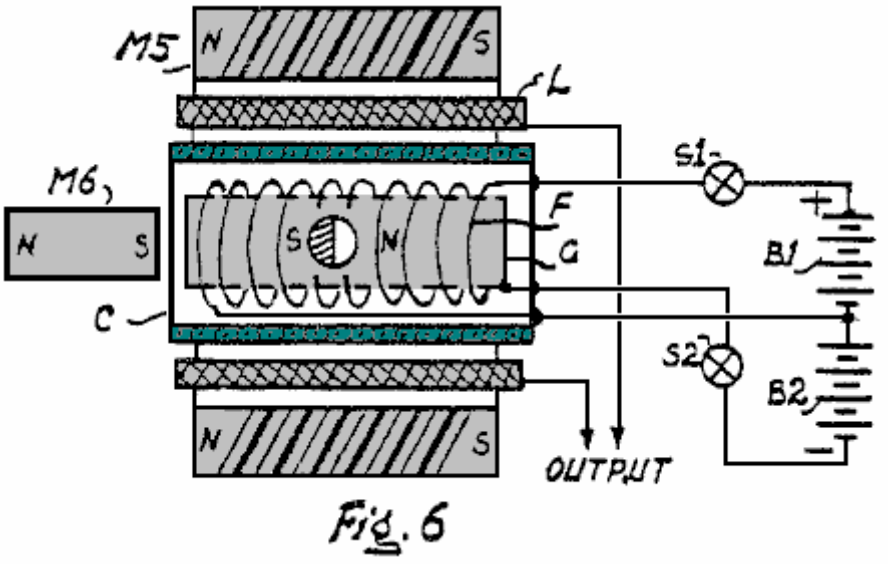


Fig. 6

Fig.6 shows a vacuum chamber **C**, which contains a cylindrically wound filament **F**, connected to the battery **B1** by way of the switch **S1**. Thus, when the switch **S1** is turned ON, the filament **F** is lighted, and it releases electrons. External to the vacuum chamber **C** is mounted a cylindrical permanent magnet **M5**, which compresses the emitted electrons into a beam at the centre of the chamber.

When the beam is formed, the switch is turned OFF, so that the beam of electrons is permanently trapped at the centre of the chamber.

The permanent trapping of the electrons in the chamber **C** represents a permanent storage of static energy. Thus, when a permanent magnet **M6** is placed to tilt the polar orientations of the uniformly poled electrons in the beam, they start precessing perpetually at a resonant frequency, as determined by the field strengths of the magnets **M5** and **M6**.

The precessing electrons in the beam will radiate quadrature phased electromagnetic field in a direction perpendicular to the polar axes of the electrons.

Thus, a coil **L** may be placed between the magnet **M5** and the vacuum chamber **C**, to receive the radiated field from the beam. The output may then be utilised in different modes for practical purposes, for example, rectified for DC power use.

The electron beam-forming cylindrical magnet **M5**, which may also be called a focusing magnet, is shown to be bipolar along the longitudinal axis. The direction of magnetisation, however, may be from the central opening to

the outer periphery of the magnet, as shown by the magnet **M7**, in **Fig.8** but the precessing magnet **M6** will be needed in either case.

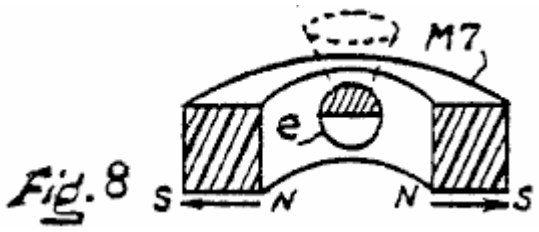


Fig. 8

In the arrangement of **Fig.6**, I have included a current control grid **G**. While it is not essential for operation of the arrangement shown, it may be connected to a high negative potential **B2** by the switch **S2** just before switching the **S1** in OFF position, so that during the cooling period of the filament, there will occur no escape of any electrons from the beam to the cathode. Also, the grid **G** may be switched ON during the heating period of the cathode, so that electrons are not forcibly released from the cathode during the heating period, and thereby causing no damage to the cathode, or filament.

Biological precessional resonance

Electron precessional resonance occurs in living tissue matter, as observed in laboratory tests. This is called ESR (Electron Spin Resonance) or PMR (Paramagnetic Resonance). In tissue matter, however, the precessing electron is entrapped between two electrons, as shown in **Fig.7**, and the polar orientations are indicated by the polar signs and shadings, for clarity of drawing.

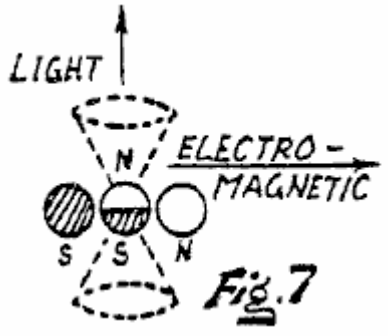


Fig. 7

Simulation

The arrangement of **Fig.7** may be simulated artificially in a manner as shown in **Fig.9**, wherein, the electron trapping magnet is a pair of parallel spaced magnets **M8**. In actual practice, however, the structure of this pair of magnets **M8** can be modified. For example, a second pair of magnets **M8** may be disposed between the two pairs, so that the directions of the transverse fields between the two pairs cross mutually perpendicular at the central longitudinal axis of the vacuum chamber. The inner field radiating surfaces of these two pairs of magnets may be shaped circular, and the two pairs may be assembled, either by physical contact to each other, or separated from each other.

Modifications

Referring to the arrangements of **Fig.6**, **Fig.9** and **Fig.10**, when the electron is in precessional gyroscopic motion, the radiated field in a direction parallel to the polar axis of the electron, is a single phased corkscrew waveform, which when precessed at light frequency, the radiation produces the effect of light.

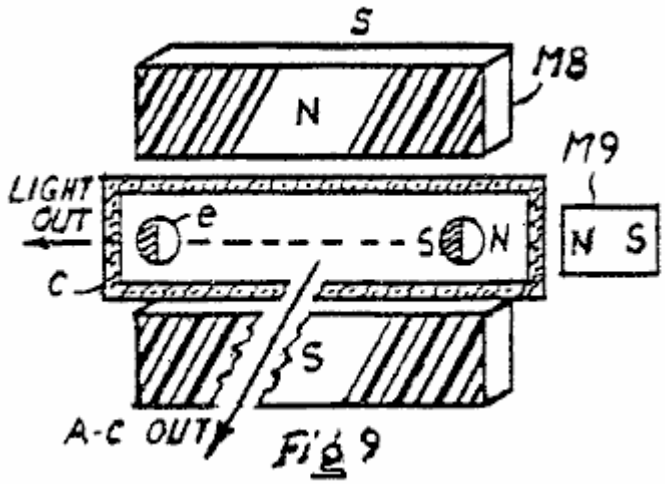


Fig 9

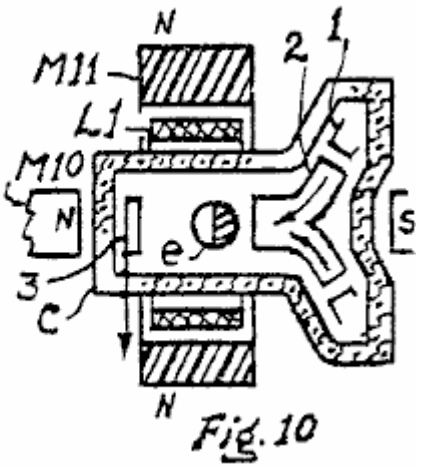


Fig. 10

Whereas, the field in a direction perpendicular to the axis of the electron produces a quadrature phased electromagnetic radiation. Thus, instead of utilising the output of electron precession for energy purposes, it may be utilised for field radiation of either light or electromagnetic waves, such as indicated by the arrows in **Fig.9**. In this case, the output will be coherent field radiation.

In reference to the arrangement of **Fig.6**, the electron emission is shown to occur within the central magnetic field of the focusing magnet **M5**. It may be practically desired, however, that these electrons are injected into the central field of the cylindrical magnet from a gun assembly, as shown in an exemplary arrangement of **Fig.10**. In this case, the vacuum chamber **C** is flanged at the right hand side, for mounting an electron emitting cathode **1** (the filament not being shown), and a curved electron-accelerating gun **2**. The central part of this flange is recessed for convenience of mounting an electron-tilting magnet (as shown), as close as possible to the electron beam. In operation, when current is passed through the filament, and a positive voltage is applied (not shown) to the gun **2**, the emitted electrons from the cathode are accelerated and injected into the central field of the magnet **11**. Assuming that the open end of the gun **2** overlaps slightly the open end of the cylindrical central field of the magnet **M1**, and the positive accelerating voltage applied to the gun **2** is very low, the accelerated electrons will enter the central field of the magnet **M1**, and travel to the other end of the field. Due to the low speed acceleration of the electrons, however, they cannot spill out of the field, and become permanently entrapped therein.

In regard to the direction in which the coil **L1** is positioned, its winding should be in a direction perpendicular to the longitudinal axis of the beam to which the polar axes of the electrons are aligned uniformly in parallel. In one practical mode, the coil **L1** may be wound in the shape of a surface winding around a tubular form fitted over the cylindrical vacuum chamber.

In regard to the operability of the apparatus as disclosed herein, the illustration in **Fig.7** shows that the field output in a direction parallel to the polar axis of the electron is singular phased, and it produces the effect of light when the precessional frequency is at a light frequency. Whereas, the output in a direction perpendicular to the polar axis of the electron is quadrature phased, which is manifested in practiced electromagnetic field transmission.

In regard to experimental references, an article entitled "Magnetic Resonance at high Pressure" in the "Scientific American" by George B. Benedek, page 105 illustrates a precessing nucleus, and indicates the direction of the electromagnetic field radiation by the precessing nucleus. The same technique is also used in the medical apparatus "Nuclear magnetic resonance" now used in numerous hospitals for imaging ailing tissues (see "High Technology" Nov. Dec. 1982. Refer also to the technique of detecting Electron Spin Resonance, in which electrons (called "free radicals") are precessed by the application of external magnetic field to the tissue matter. In all of these practices, the electromagnetic field detecting coils are directed perpendicular to the polar axes of the precessing electrons or the nuclei.

In regard to the production of light by a precessing electron, in a direction parallel to the polar axis of the precessing electron, see an experimental reference entitled "Free electrons make powerful new laser" published in "high Technology" February 1983 page 69.

In regard to the aspect of producing and storing the electrons in a vacuum chamber, it is a known fact by practice that the electrons are entrapped within the central field of a cylindrical permanent magnet, and they will remain entrapped as long as the magnet remains in position.

With regard to the performance of obtaining precessional resonance of the electron, the simple example of a

wobbling top is sufficient, as proof of operability.

Having described the preferred embodiments of the invention, and in view of the suggestions of numerous possibilities of modifications, adaptations, adjustments and substitutions of parts, it should be obvious to the skilled in related arts that other possibilities are within the spirit and scope of the present invention.

CLAIMS

1. The method of effecting perpetual retaining and precession of electrons, for obtaining perpetual field radiation from the polar motions of said precessing electrons, comprising the steps of:

producing electrons;
compressing said produced electrons into a perpetually retainable state; and
precessing said compressed electrons for effecting perpetual field radiation by the polar motions of said precessing electrons.

2. The method of producing perpetual field radiation for conversion into perpetual energy, the method comprising the steps of:

producing electrons;
imposing a first perpetually occurring electron controlling force from a first direction upon said produced electrons into a perpetually retainable state; and
imposing a second perpetually occurring electron controlling force from a second direction upon said retained electrons, for inducing precessional motions to the electrons, and thereby obtaining said perpetual field radiation for conversion into perpetual energy.

3. The method of generating perpetual simultaneous single phased and quadrature phased coherent field radiations, comprising the steps of:

producing electrons;
imposing a first perpetually occurring electron controlling force from a first direction upon said produced electrons into a uniformly polarised perpetually retainable compressed state; and
imposing a second perpetually occurring electron controlling force from a second direction upon said compressed electrons, for effecting precessional motions of the electrons, thereby causing a quadrature phased coherent field radiation in a direction perpendicular to the uniformly polarised polar axes of said electrons, and a simultaneous single phased coherent field in a direction parallel to the polar axes of said electrons.

4. The method of producing perpetual dynamic motions for conversion into energy, comprising the steps of:

trapping and compressing a concentrated quantity of electrons within a first electron controlling field in a vacuum space, whereby forming a tightly confined permanent concentration of statistically spinning electrons, both of their polar axes and polar orientations being uniformly aligned; and
tilting the polar axes of said trapped electrons by a second permanent electron controlling field, for inducing precessional gyrations to the electrons in the form of perpetual dynamic motions, which are adaptively convertible into energy.

5. Apparatus for producing perpetual dynamic motions, which comprises:

a vacuum chamber having an electron-emitting means; an auxiliary means for causing emission of electrons from said electron-emitting means;
a first permanent magnet disposed externally of said chamber for trapping and compressing a quantity of said emitted electrons within its magnetic field, with uniform alignments of the polar axes and polar orientations of said electrons;
means for stopping said auxiliary means from further causing emission of electrons from said electron emitting means, whereby forming a tightly confined concentration of statistically spinning electrons permanently entrapped within said first permanent magnet; and
a second permanent magnet, the field projection of which is oriented to tilt the polar a axes of said trapped electrons, for causing precessional gyrations of the electrons, as representation of said dynamic motions.

6. Apparatus comprising:
a vacuum chamber having an electron emitting means;

an auxiliary means for causing emission of electrons from said electron emitting means;
a first permanent magnet disposed externally of said chamber for permanently trapping and compressing a quantity of said emitted electrons within its magnetic field, with uniform alignments of the polar axes and polar orientations of said electrons; and
a second permanent magnet so oriented with respect to said entrapped electrons that, the field projection from the second magnet causes precessional gyrations of the uniformly aligned entrapped electrons.

7. The apparatus as set forth in claim 6, wherein said first permanent magnet is cylindrical magnet surrounding said chamber, and the magnetisation of said first magnet is in a direction along the longitudinal axis of the cylinder.

8. The apparatus as set forth in claim 6, wherein said first permanent magnet is cylindrical magnet surrounding said chamber, and the magnetisation of said first magnet is in a direction from the central hollow space to the outer surface of said cylinder.

9. The apparatus as set forth in claim 6, wherein the polar sign of field projection from said second magnet to said entrapped electrons is in repelling polar sign.

10. The apparatus as set forth in claim 6, wherein is included a field responsive coil mounted between said first magnet and said vacuum chamber, for receiving the field radiation that is effected by the motions of said gyrating electrons.

11. The apparatus as set forth in claim 6, wherein is included a field responsive coil mounted between said first magnet and said vacuum chamber, the turns of winding of said coil being in a direction perpendicular to the polar axes of said compressed electrons.

12. Apparatus for producing perpetual motion, the apparatus being substantially as hereinbefore described with reference to, and as illustrated by, the accompanying drawings.

Theodore Annis & Patrick Eberly

US Patent App. 20090096219 16th April 2009 Inventors: Theodore Annis & Patrick Eberly

ENERGY GENERATION APPARATUS AND METHODS BASED UPON MAGNETIC FLUX SWITCHING

ABSTRACT

Methods and apparatus generate electricity through the operation of a circuit based on a single magnetic flux path. A magnetisable member provides the flux path. One or more electrically conductive coils are wound around the member, and a reluctance or flux-switching apparatus is used to control the flux. When operated, the switching apparatus causes a reversal of the polarity (direction) of the magnetic flux of the permanent magnet through the member, thereby inducing alternating electrical current in each coil. The flux-switching apparatus may be motionless or rotational. In the motionless embodiments, two or four reluctance switches are operated so that the magnetic flux from one or more stationary permanent magnet(s) is reversed through the magnetisable member. In alternative embodiments, the flux-switching apparatus comprises a body composed of high-permeability and low-permeability materials, such that when the body is rotated, the flux from the magnet is sequentially reversed through the magnetisable member.

FIELD OF THE INVENTION

The present invention relates to methods and apparatus wherein the magnetic flux from one or more permanent magnets is reversed repeatedly in polarity (direction) through a single flux path around which there is wound a conducting coil or coils for the purpose of inducing electricity in the coils.

BACKGROUND OF THE INVENTION

The electromechanical and electromagnetic methods involved in motional electric generators and alternators are well known. Alternators and generators often employ permanent magnets and usually have a rotor and a stator and a coil or coils in which an EMF (electromotive force) is induced. The physics involved for producing electricity is described by the generator equation $V = \int(vxB).dl$.

Permanent magnets made of materials that have a high coercively, a high magnetic flux density a high magnetic motive force (mmf), and no significant deterioration of magnetic strength over time are now common. Examples include ceramic ferrite magnets (Fe2O3); samarium cobalt (SmCO5); combinations of iron, neodymium, and boron; and others.

Magnetic paths for transformers are often constructed of laminated ferrous materials; inductors often employ ferrite materials, which are used for higher frequency operation for both devices. High performance magnetic materials for use as the magnetic paths within a magnetic circuit are now available and are well suited for the (rapid) switching of magnetic flux with a minimum of eddy currents. An example is the FINEMET® nanocrystalline core material made by Hitachi of Japan.

According to Moskowitz, "Permanent Magnet Design and Application Handbook" 1995, page 52, magnetic flux may be thought of as flux lines which always leave and enter the surfaces of ferromagnetic materials at right angles, which never can make true right-angle turns, which travel only in straight or curved paths, which follow the shortest distance, and which follow the path of lowest reluctance.

A "reluctance switch" is a device that can significantly increase or decrease (typically increase) the reluctance (resistance to magnetic motive force) of a magnetic path in a direct and rapid manner and subsequently restore it to its original (typically lower) value in a direct and rapid manner. A reluctance switch typically has analog characteristics. By way of contrast, an off/on electric switch typically has a digital characteristic, as there is no electricity "bleed-through." With the current state of the art, reluctance switches have magnetic flux bleed-through. Reluctance switches may be implemented mechanically, such as to cause keeper movement to create an air gap, or electrically by several means, or by other means. One electrical means is that of using control coils wound around the flux paths.

Another electrical means is the placement within the flux path of certain classes of materials that change (typically increase) their reluctance upon the application of electricity. Another electrical means is to saturate a region of the switch material so that the reluctance increases to that of air by inserting conducting electrical wires into the material as described by Konrad and Brudny in "An Improved Method for Virtual Air Gap Length Computation," in IEEE Transactions on Magnetics, Vol. 41, No. 10, October 2005.

The patent literature describes a number of constructs that have been devised to vary the amounts of magnetic flux in alternate flux paths by disproportionately dividing the flux from a stationary permanent magnet or magnets between or among alternate flux paths repeatedly for the purpose of generating electricity. The increase of flux in one magnetic path and the corresponding decrease in the other path(s) provide the basis for inducing electricity when coils are wound around the paths. The physics involved for producing electricity by these constructs is described by the transformer equation $V = -\int dB/dt.ds$. A variety of reluctance switching means have been employed to cause the flux to be increased/decreased through a particular alternate path with a corresponding decrease/increase in the other path and to do so repeatedly.

A means of switching flux along alternate paths between the opposite poles of a permanent magnet have included the flux transfer principle described by R. J. Radus, Engineers' Digest, July, 1963.

A result of providing alternate flux paths of generally similar geometry and permeability is that, under particular conditions, the alternate path first selected or the path selected for the majority of the flux will remain a "preferred path" in that it will retain more flux and the other path, despite the paths having equal reluctance. (There is not an automatic equalization of the flux among similar paths.)

Moskowitz, "Permanent Magnet Design and Application Handbook" 1995, page 87 discusses this effect with regard to the industrial use of permanent magnets to lift and release iron and steel by turning the permanent magnet on and (almost) off via reluctance switching that consists of the electric pulsing of coils wound around the magnetic flux paths (the reluctance switches).

Experimental results with four iron rectangular bars (relative permeability=1000) placed together in a square with a bar permanent magnet (flux density measured at one pole=5000 Gauss) between two of the opposing bars roughly in a centre position showed that removal and replacement of the one of the end bars that is parallel to the bar magnet will result in about 80% of the flux remaining in the bar that remained in contact. The results further showed that the preferred path must experience an increase of reluctance about IOx of that of the available alternate path before its disproportionate flux condition will yield and transfer to the alternate path.

Flynn U.S. Pat. No. 6,246,561; Patrick, et al. U.S. Pat. No. 6,362,718; and Pedersen U.S. Pat. No. 6,946,938 all disclose a method and apparatus for switching (dividing) the quantity of magnetic flux from a stationary permanent magnet or magnets between and among alternate paths for the purpose of generating electricity (and/or motive force). They provide for the increase of magnetic flux in one path with a corresponding decrease in the other path(s). There are always at least two paths.

SUMMARY OF THE INVENTION

The present invention relates to methods and apparatus for the production of electricity through the operation of a circuit based upon a single magnetic flux path. A magnetisable member provides the flux path. One or more electrically conductive coils are wound around the member, and a reluctance or flux switching apparatus is used to control the flux. When operated, the switching apparatus causes a reversal of the polarity (direction) of the magnetic flux of the permanent magnet through the member, thereby inducing alternating electrical current in each coil.

According to the invention, the flux switching apparatus may be motionless or rotational. In the motionless embodiments, four reluctance switches are operated by a control unit that causes a first pair of switches to open (increasing reluctance), while another pair of switches close (decreasing reluctance). The initial pair is then closed as the other pair is opened, and so on. This 2x2 opening and closing cycle repeats and, as it does, the magnetic flux from the stationary permanent magnet(s) is reversed in polarity through the magnetisable member, causing electricity to be generated in the conducting coils. An alternative motionless embodiment uses two reluctance switches and two gaps of air or other materials.

In alternative embodiments, the flux switching apparatus comprises a body composed of high-permeability and low-permeability materials, such that when the body is rotated, the flux from the magnet is sequentially reversed through the magnetisable member. In the preferred embodiment the body is cylindrical having a central axis, and the body rotates about the axis. The cylinder is composed of a high-permeability material except for section of low-permeability material that divided the cylinder into two half cylinders. At least one electrically conductive coil is wound around the magnetisable member, such that when the body rotates an electrical current is induced in the coil. The body may be rotated by mechanical, electromechanical or other forces.

A method of generating electrical current, comprises the steps of providing a magnetisable member with an electrically conductive coil wound therearound, and sequentially reversing the flux from a permanent magnet through the member, thereby inducing electrical current in the coil.

BRIEF DESCRIPTION OF THE DRAWINGS

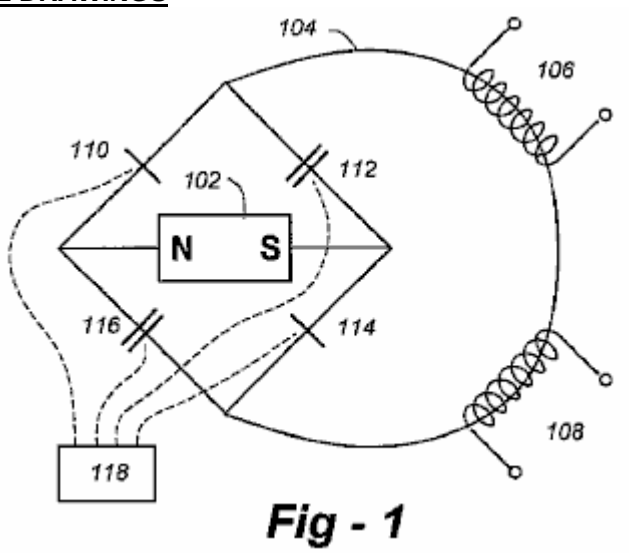


Fig - 1

Fig.1 is a schematic diagram of a magnetic circuit according to the invention.

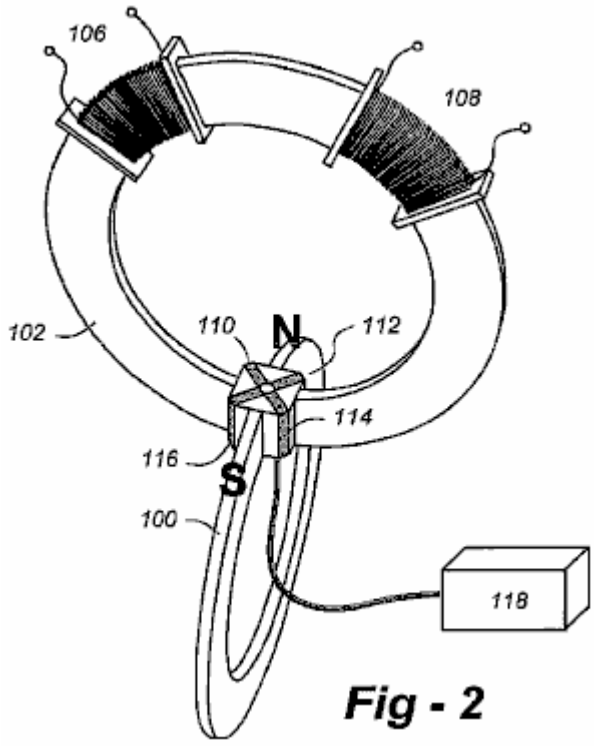


Fig - 2

Fig.2 is a perspective view of an embodiment of the invention based upon motionless magnetic flux switches.

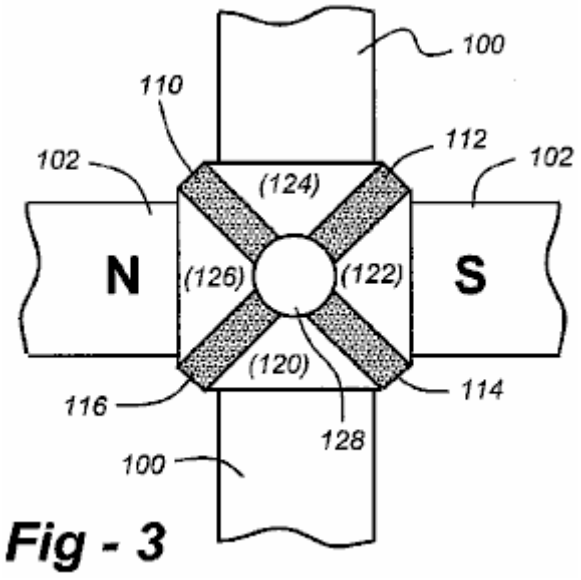


Fig - 3

Fig.3 is a detail drawing of a motionless flux switch according to the invention.

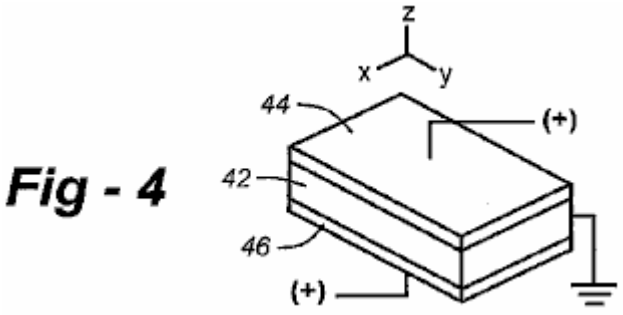


Fig - 4

Fig.4 is a detail drawing of a reluctance switch according to the invention.

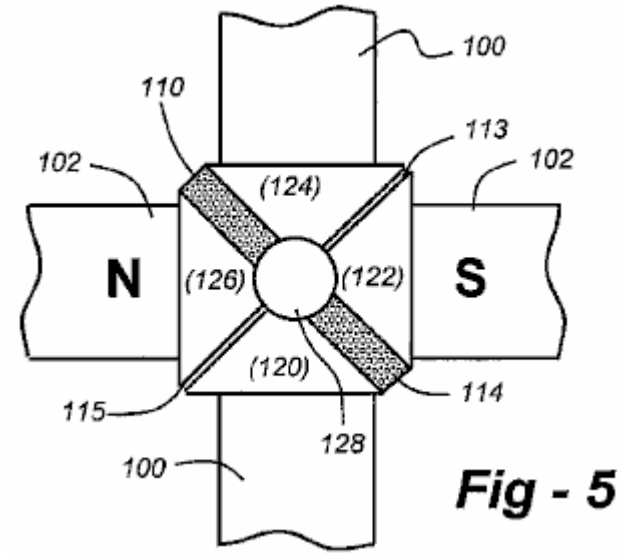


Fig - 5

Fig.5 is a detail drawing of an alternative motionless flux switch according to the invention which utilizes gaps of air or other materials.

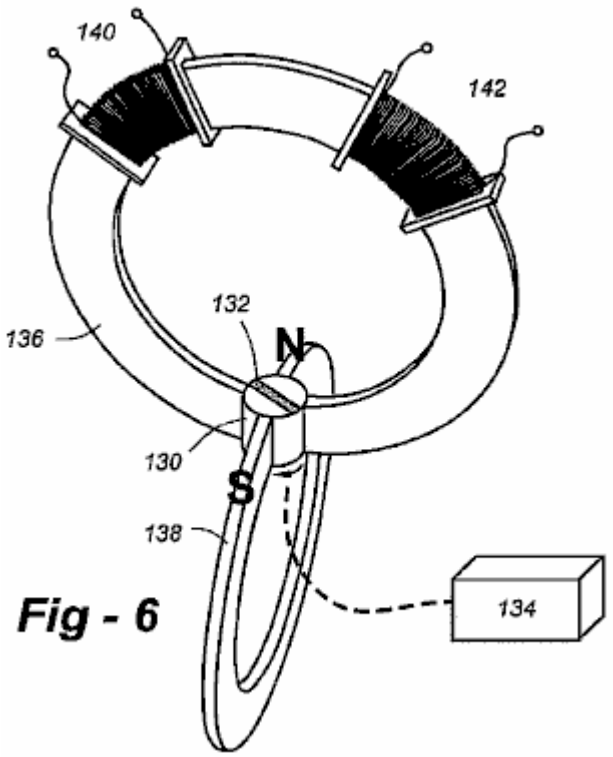


Fig - 6

Fig.6 is a schematic diagram of a system using a rotary flux switch according to the invention.

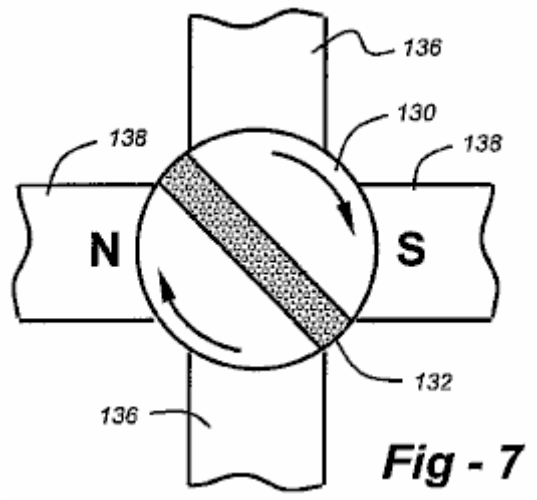


Fig - 7

Fig.7 is a detail drawing of a rotary flux switch according to the invention.

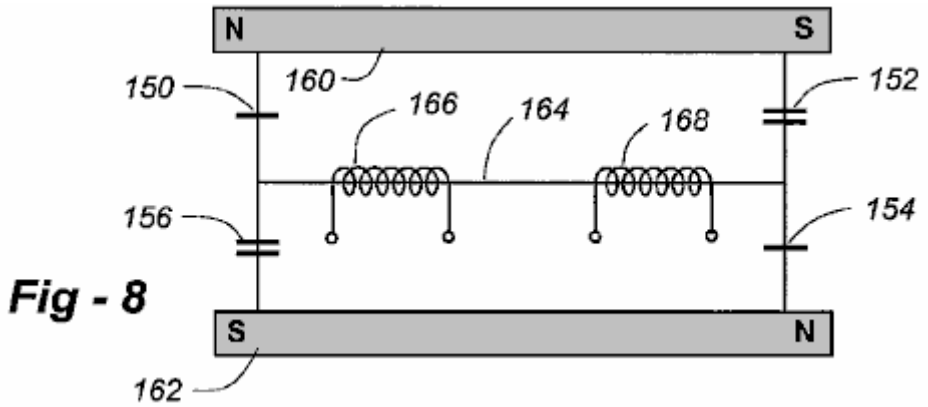


Fig - 8

Fig.8 is a schematic diagram of a circuit according to the invention utilizing two permanent magnets and a single flux path.

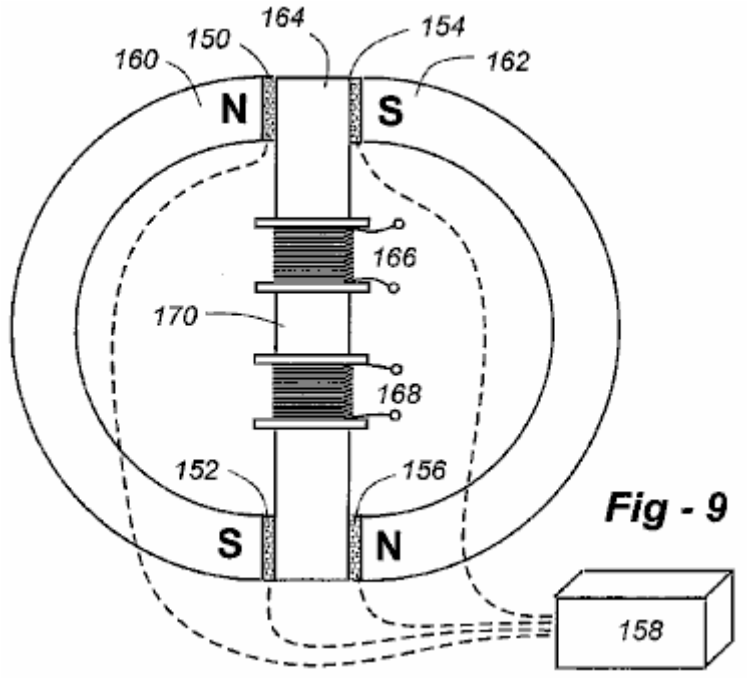


Fig - 9

Fig.9 shows one possible physical embodiment of the apparatus with the components of FIG. 8, including a reluctance switch control unit.

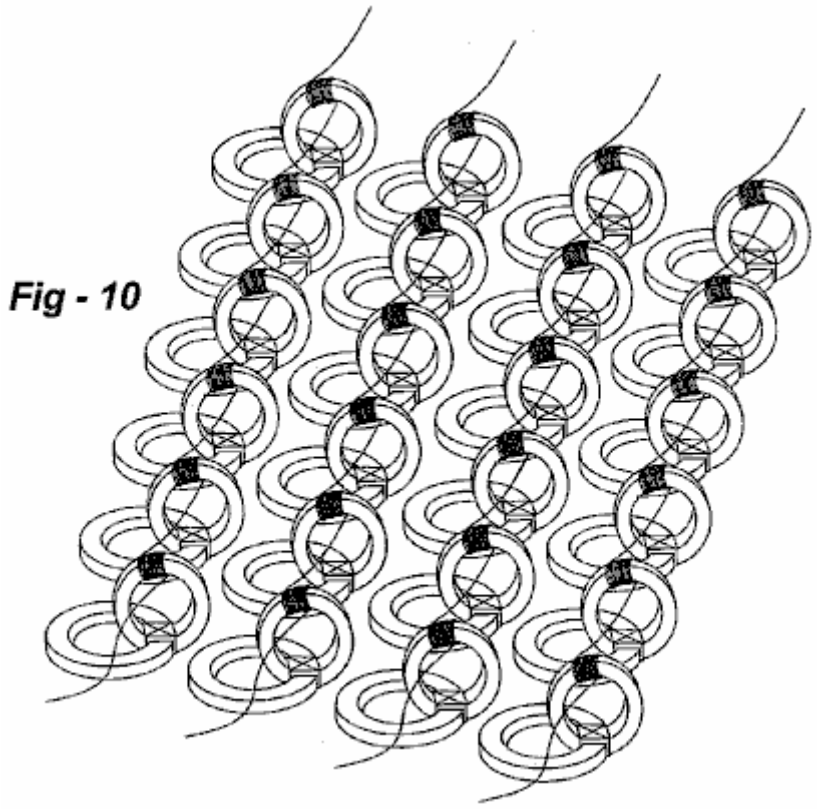


Fig - 10

Fig.10 shows and array of interconnected electrical generators according to the invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

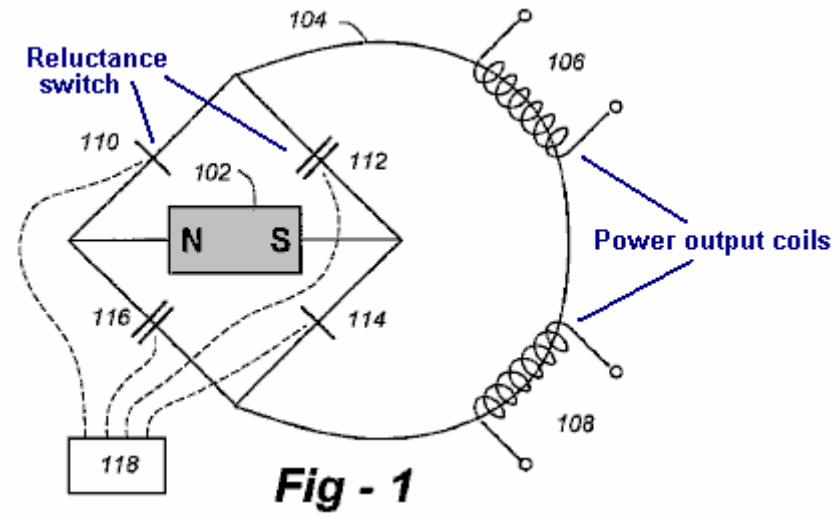


Fig - 1

Fig.1 is a schematic diagram of a magnetic circuit according to the invention utilizing a motionless flux switch. The circuit includes the following components: a permanent magnet 102, single flux path 104, conducting coils 106, 108, and four reluctance switches 110, 112, 114, 116. Under the control of unit 118, reluctance switches 110, 114 open (increasing reluctance), while switches 112, 116 close (decreasing reluctance). Reluctance switches 110, 114 then close, while switches 112, 116 open, and so on. This 2x2 opening and closing cycle repeats and, as it does, the magnetic flux from stationary permanent magnet 102 is reversed in polarity through single flux path 104, causing electricity to be generated in conducting coils 106, 108.

An efficient shape of permanent magnet 102 is a "C" in which the poles are in close proximity to one another and engage with the flux switch. The single flux is carried by a magnetisable member 100, also in a "C" shape with ends that are in close proximity to one another and also engage with the flux switch. In this, and in other embodiments, the 2x2 switching cycle is carried out simultaneously. As such, control circuit 118 is preferably implemented with a crystal-controlled clock feeding digital counters, flip-flops, gate packages, or the like, to adjust rise time, fall time, ringing and other parasitic effects. The output stage of the control circuit may use FET (Field-Effect Transistor switches) to route analog or digital waveforms to the reluctance switches as required.

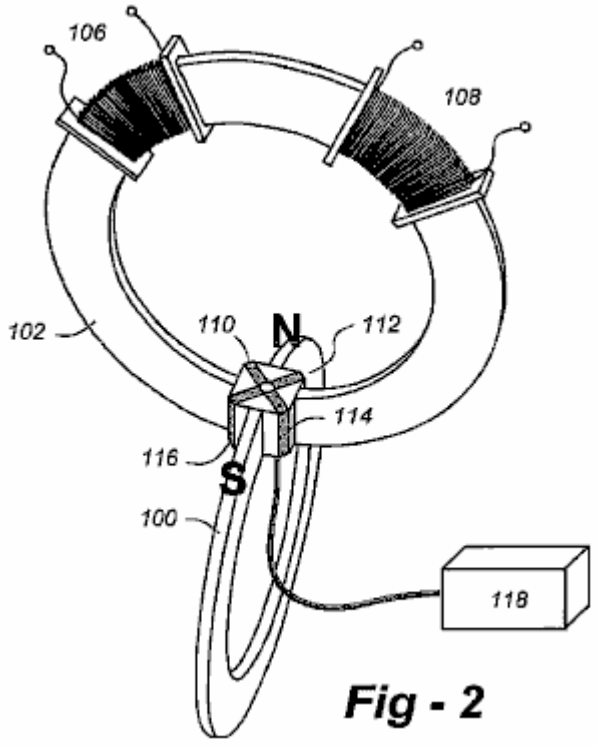


Fig - 2

Fig.2 is a perspective of one possible physical embodiment of the apparatus using the components of Fig.1, showing their relative positions to one another. Reluctance switches 110, 112, 114, 116 may be implemented differently, as described below, but will usually occupy the same relative position within the apparatus.

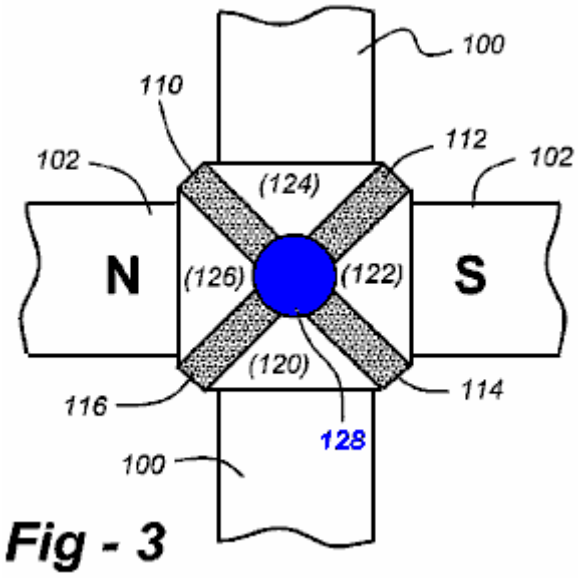


Fig - 3

Fig.3 is a detail drawing of the motionless flux switch. Connecting segments 120, 122, 124, 126 must be made of a high-permeability ferromagnetic material. The central volume 128 may be a through-hole, providing an air gap, or it may be filled with glass, ceramic or other low-permeability material. A super-conductor or other structure exhibiting the Meissner effect may alternatively be used.

In the embodiment depicted in Fig.2 and Fig.3, reluctance switches 110, 112, 114, 116 are implemented with a solid-state structure facilitating motionless operation. The currently preferred motionless reluctance switch is described by Toshiyuki Ueno & Toshiro Higuchi, in the paper "Investigation on Dynamic Properties of Magnetic Flux Control Device composed of Lamination of Magnetostrictive Material Piezoelectric Material," The University of Tokyo 2004, the entirety of which is incorporated herein by reference. As shown in Fig.4, this switch is made of a laminate of a GMM (Giant Magnetostrictive Material 42), a TbDyFe alloy, bonded on both sides by a PZT (Piezoelectric) material 44, 46 to which electricity is applied. The application of electricity to the PZT creates strain on the GMM, which causes its reluctance to increase.

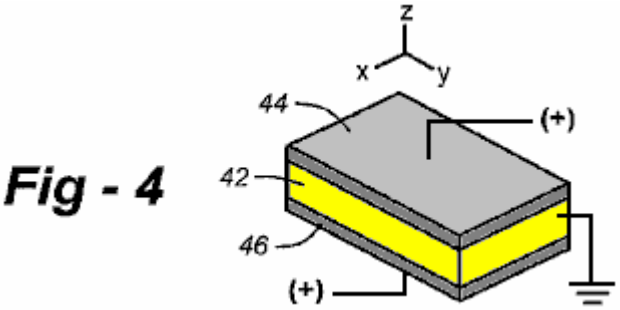


Fig - 4

Other arrangements are applicable, including those disclosed in pending U.S. Patent Application Serial no. 2006/0012453, the entire content of which is incorporated herein by reference. These switches disclosed in this reference are based upon the magnetoelectric (ME) effects of liquid crystal materials in the form of magnetorestrictive and piezoelectric effects. The properties of ME materials are described, for example, in Ryu et al, "Magnetoelectric Effect in Composites of Magnetorestrictive and Piezoelectric Materials," Journal of Electroceramics, Vol. 8, 107-119

Filipov et al, "Magnetoelectric Effects at Piezoresonance in Ferromagnetic-Ferroelectric Layered Composites," Abstract, American Physical Society Meeting (March 2003) and Chang et al., "Magneto-band of Stacked Nanographite Ribbons," Abstract, American Physical Society Meeting (March 2003). The entire content of each of these papers are also incorporated herein.

Further alternatives include materials that may sequentially heated and allowed to cool (or cooled and allowed to warm up or actively heated and cooled) above and below the Currie temperature, thereby modulating reluctance. Gadolinium is a candidate since its Currie point is near room temperature. High-temperature superconductors are other candidates, with the material being cooled in an insulated chamber at a temperature substantially at or near the Currie point. Microwave or other energy sources may be used in conjunction with the control unit to effectuate this switching. Depending upon how rigidly the switches are contained, further expansion-limiting `yokes' may or may not be necessary around the block best seen in Fig.4.

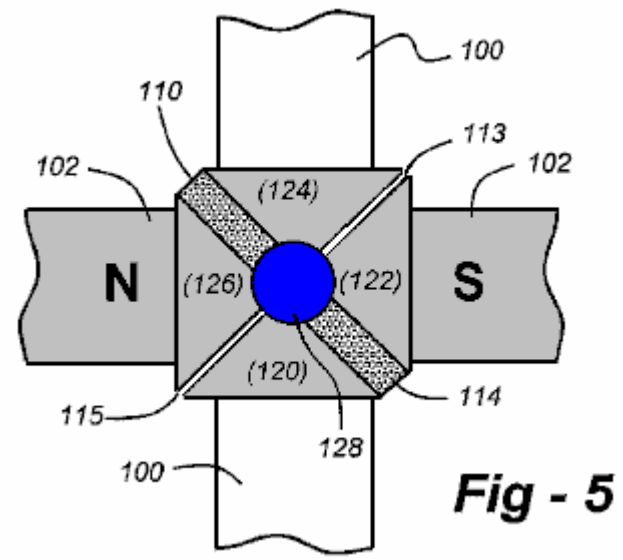


Fig - 5

Fig.5 is a detail drawing of an alternative motionless flux switch according to the invention which utilizes gaps of air or other materials. This embodiment uses two electrically operated reluctance switches 110, 114, and two gaps 113, 115, such that when the switches are activated in a prescribed manner, the flux from the magnet 102 is blocked along the switch segments containing the switches and forced through the gap-containing segments, thereby reversing the flux through the magnetisable member 100. Upon activation of the two reluctance switches 110, 114, the flux, seeking a path of significantly lower reluctance, flips back to the original path containing the (non deactivated) reluctance switches, thereby reversing the flux through the member 100. Note that the flux switches may also be electromagnetic to saturate local regions of the switch such that reluctance increases to that of air (or gap material), creating a virtual gap as described by Konrad and Brudny in the Background of the Invention.

More particularly, flux switching apparatus according to this embodiment uses a permanent magnet having a north pole `N' and a south pole `S' in opposing relation across a gap defining a volume. A magnetisable member with ends `A' and 'B' is supported in opposing relation across a gap sharing the volume, and a flux switch comprises a stationary block in the volume having four sides, 1-4, with two opposing sides interfaced to N and S, respectively and with the other two opposing sides being interfaced to A and B, respectively. The block is composed of a magnetisable material segmented by two electrically operated magnetic flux switches and two gaps filled with air or other material(s). A control unit in electrical communication with the flux switches is operative to:

a) passively allow a default flux path through sides 1-2 and 3-4, then
b) actively establish a flux path through sides 2-3 and 1-4, and
c) repeat a) and b) on a sequential basis.

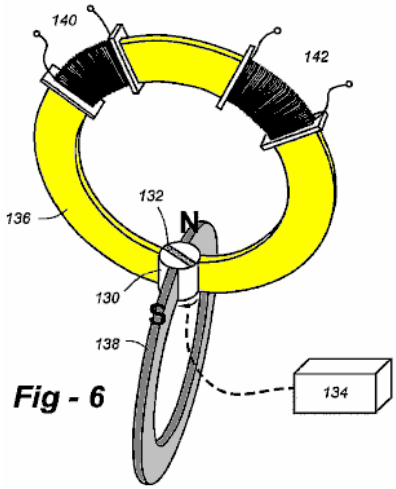


Fig - 6

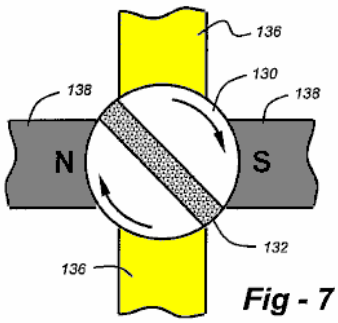


Fig - 7

As an alternative to a motionless flux switch, a rotary flux switch may be used to implement the 2x2 alternating sequence. Referring to Fig.6 and Fig.7, cylinder 130 with flux gap 132 is rotated by a motive means 134. This causes the halves of cylinder 130 to provide two concurrent and separate magnetic flux bridges (i.e., a "closed" reluctance switch condition), in which a given end of magnetisable member 136 is paired up with one of the poles

of stationary permanent magnet 138. Simultaneously, the other end of single flux path carrier 136 is paired up with the opposite pole of stationary permanent magnet 138.

Fig.7 is a detail view of the cylinder. Each 90° rotation of the cylinder causes the first flux bridges to be broken (an "open" reluctance switches condition) and a second set of flux bridges to be created in which the given end of member 136 is then bridged with the opposite pole of stationary permanent magnet 138. A full rotation of cylinder 130 causes four such reversals. Each flux reversal within single flux path 2 causes an electric current to be induced in conducting coil(s) 140, 142. In this embodiment, it is important to keep a precise, consistent spacing between each of the "halves" of (rotating) cylinder 130 in relation to the poles of permanent magnet 138 and the ends of flux path carrier 136 as the magnetic flux bridges are provided by the cylinder 130 as it rotates.

Rotating cylinder 130 is made of high magnetic permeability material which is divided completely by the flux gap 132. A preferred material is a nanocrystalline material such as FINEMET® made by Hitachi. The flux gap 132 may be air, glass, ceramic, or any material exhibiting low magnetic permeability. A superconductor or other structure exhibiting the Meissner effect may alternatively be used.

An efficient shape of magnetisable member 136 is a "C" in which its opposing ends are curved with a same radius as cylinder 130 and are in the closest possible proximity with rotating cylinder 130. Permanent magnet 138 is also preferably C-shaped in which the opposing poles are curved with a same radius as cylinder 130 and are in the closest possible proximity with rotating cylinder 130. Manufacturing and assembly considerations may dictate other shapes.

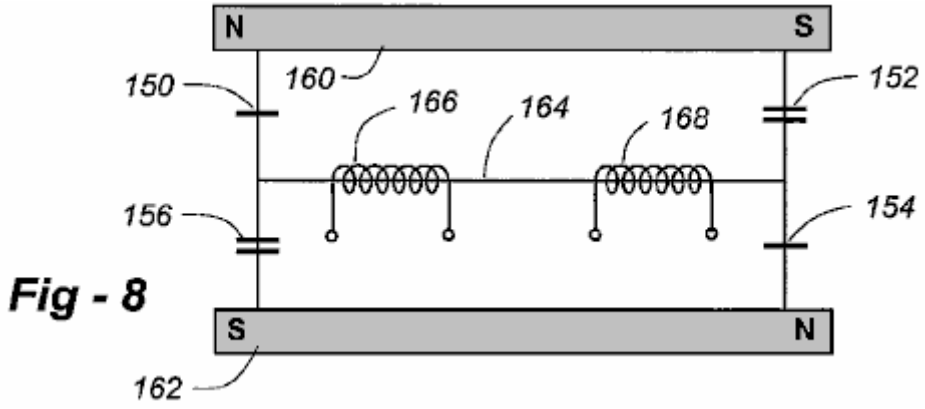


Fig - 8

While the embodiments described thus far utilize a single permanent magnet, other embodiments are possible according to the invention utilizing a plurality of permanent magnets while nonetheless generating a single flux path. Fig.8 depicts a circuit utilizing two permanent magnets and a single flux path. Fig.9 shows one possible physical embodiment of the apparatus based upon the components of Fig.8, including a reluctance switch control unit 158.

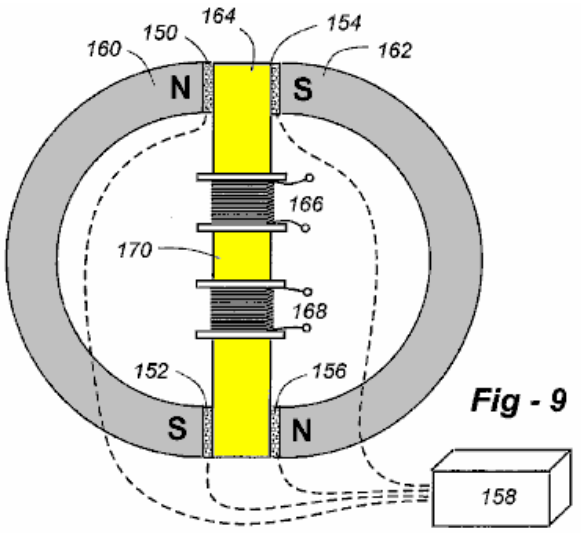


Fig - 9

Under the control of unit 158, reluctance switches 150, 152 open (increasing reluctance), while switches 154, 156 close (decreasing reluctance). Reluctance switches 150, 152 then close, while switches 154, 156 open, and so on. This 2x2 opening and closing cycle repeats and, as it does, the magnetic flux from stationary permanent magnets 160, 162 is reversed in polarity through the magnetisable member, causing electricity to be generated in conducting coils 166, 168.

In the preferred implementation of this embodiment, the magnets are arranged with their N and S poles reversed. The magnetisable member is disposed between the two magnets, and there are four flux switches, SW1-SW4, two between each end of the member and the poles of each magnet. The reluctance switches are implemented with the structures described above with reference to Figs. 1 to 3.

For added particularity, assume the first magnet has north and south poles, N1 and S1, the second magnet has north and south poles, N2 and S2 and the member has two ends A and B. Assuming SW1 is situated between N1 and A, SW2 is between A and S2, SW3 is between N2 and B, and SW4 is between B and S1, the control circuitry operative to activate SW1 and SW4, then activate SW2 and SW3, and repeat this process on a sequential basis. As with the other embodiments described herein, for reasons of efficiency, the switching is carried out simultaneously.

In all of the embodiments described herein the material used for the permanent magnet(s) may be either a magnetic assembly or a single magnetized unit. Preferred materials are ceramic ferrite magnets (Fe_20_3), samarium cobalt ($SmCO_5$), or combinations of iron, neodymium, and boron. The single flux path is carried by a material having a high magnetic permeability and constructed to minimize eddy currents. Such material may be a laminated iron or steel assembly or ferrite core such as used in transformers. A preferred material is a nanocrystalline material such as FINEMET®. The conducting coil or coils are wound around the material carrying the single flux path as many turns as required to meet the voltage, current or power objectives. Ordinary, standard, insulated, copper magnet wire (motor wire) is sufficient and acceptable. Superconducting materials may also be used. At least some of the electricity induced in the conducting coils may be fed back into the switch control unit. In this mode of operation, starting pulses of electricity may be provided from a chemical or solar battery, as required.

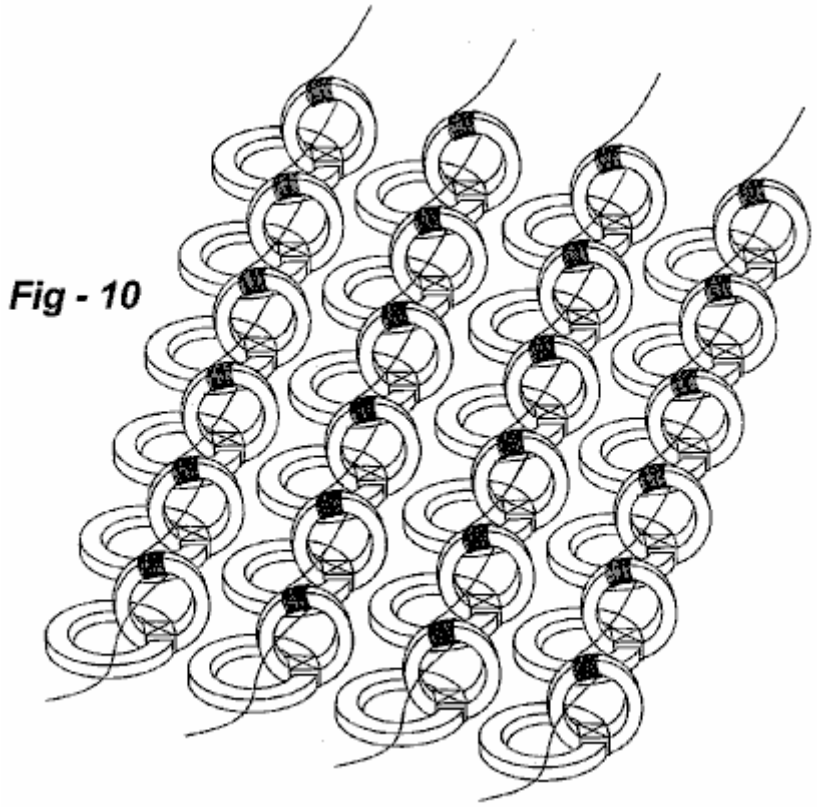


Fig - 10

Although in the embodiments of Fig.2 and Fig.6 the magnet and flux-carrying materials are flat elements lying in orthogonal planes with flux-carrying material lying outside the volume described by the magnet, the flux path may be disposed `within' the magnet volume or configured at an angle. The physical scale of the elements may also be varied to take advantage of manufacturing techniques or other advantages. Fig.10, for example, shows an array of magnetic circuits, each having one or more coils that may be in series, parallel, or series-parallel combinations, depending upon voltage or current requirements. In each case the magnets may be placed or fabricated using techniques common to the microelectronics industry. If mechanical flux switches are used they may be fabricated using MEMs-type techniques. If motionless switches are used, the materials may be placed and/or deposited. The paths are preferably wound in advance then picked and placed into position as shown. The embodiment shown in Fig.9 is also amenable to miniaturization and replication.

CLAIMS

1. An energy generator, comprising: at least one permanent magnet generating flux; a magnetisable member; an electrical conductor wound around the member; and a plurality of magnetic flux switches operative to sequentially reverse the flux from the magnet through the member, thereby inducing electricity in the electrical conductor.

2. The energy generator of claim 1, comprising: first and second loops of magnetisable material; the first loop having four segments in order A, 1, B, 2; the second loop having four segments in order C, 3, D, 4; the magnetisable member coupling segments 2 and 4; the permanent magnet coupling segments 1 and 3, such that the flux from the magnet flows through segments A, B, C, D and the magnetisable member; four magnetic flux switches, each controlling the flux through a respective one of the segments A, B, C, D; and a controller operative to activate switches A-D and B-C in an alternating sequence, thereby reversing the flux through the segment and inducing electricity in the electrical conductor.

3. The energy generator of claim 2, wherein the loops and magnetisable member are composed of a nanocrystalline material exhibiting a substantially square BH intrinsic curve.

4. The energy generator of claim 2, wherein each magnetic flux switch is operative to add flux to the segment it controls, thereby magnetically saturating that segment when activated.

5. The energy generator of claim 2, wherein: each segment has an aperture formed therethrough; and each magnetic flux switch is implemented as a coil of wire wound through one of the apertures.

6. The energy generator of claim 2, wherein the controller is at least initially operative to activate the switches with electrical current spikes.

7. The energy generator of claim 2, wherein the first and second loops are toroids.

8. The energy generator of claim 2, wherein the first and second loops are spaced apart from one another, with A opposing C, 1 opposing 3, B opposing D and 2 opposing 4.

9. The energy generator of claim 2, wherein the first and second loops intersect to form the magnetisable member.

10. The energy generator of claim 2, wherein the flux flowing through each segment A, B, C, D is substantially half of that flowing through the magnetisable member prior to switch activation.

The Energy Conversion Device of William McDavid junior

US Patent 6,800,955 5th Oct. 2004 Inventor: William McDavid jnr.

Fluid-powered energy conversion device

Note: The wording of this patent has been altered to make it easier to understand. An unaltered copy can be downloaded from www.freepatentsonline.com. In this patent, William relates sections of his design according to the direction of flow through the housing and so he calls the first section the "downstream" chamber and the following chamber as the "upstream" chamber. Although water could be used, this patent essentially describes a high-efficiency wind-powered generator. For dimensions: one inch = 25.4 mm.

Abstract:

A fluid-powered energy conversion device which converts energy in a moving fluid into mechanical energy. A rigid cylindrical frame of toroidal baffles forms an "upstream" annular or ring-shaped chamber and a "downstream" annular chamber, each of the chambers having open sides to allow the entry of the fluid. The toroidal baffles create an upstream **drive** vortex in an upstream central vortex chamber, and a downstream **extraction** vortex rotating in the opposite direction in a downstream central vortex chamber. A set of hinged louvers surround the vortex chambers and these allow the fluid to enter each chamber only in the direction of vortex rotation, and prevent the fluid from exiting through the sides of the device. The driving vortex passes through, and rotates, a turbine positioned in a central aperture between the two chambers. The turbine blades are rotated by the rotational momentum of the driving fluid vortex, plus the lift generated by each turbine blade, plus the additional momentum imparted by the vortex reversal.

US Patent References:

McDavid, Jr. – US 6,710,469
McDavid, Jr. – US 6,518,680
Walters – US 5,664,418

Description:

BACKGROUND OF THE INVENTION

1. Technical Field of the Invention

The present invention relates generally to electrical generation and energy conversion devices, and more particularly to a fluid-powered energy conversion device which converts the energy of wind or flowing water into mechanical or electrical energy.

2. Description of Related Art

The use of wind or flowing water to provide power for various uses dates back many centuries. In modern times, wind and water have been used to generate electricity. Hydroelectric generating plants have been used to generate large quantities of electrical energy for widespread distribution. However, this requires major permanent environmental changes to the areas where dams are built and reservoirs rise. Wind-powered devices, in general, have been used to perform mechanical work, or to generate electricity, only on a limited scale. With the ever increasing demand for additional, or alternative energy sources, all possible sources are being given more scrutiny. This is particularly true for sources which are non-polluting and inexhaustible. Free-flowing hydro-electric and wind-powered systems provide such sources, and the capturing of increased energy from wind and water has received much consideration.

However, commercial hydro-electric and wind-powered electrical generation devices which are currently in use have several disadvantages. Wind-powered devices, in particular, are expensive, inefficient, dangerous, noisy, and unpleasant to be around. To capture a large volume of wind, existing wind-powered devices are very large. As a result, they cannot be distributed throughout population centres, but must be installed some distance away. Then, like dams with hydro-electric generators, the electrical energy they generate must be transmitted, at considerable cost and with considerable energy loss, to the population centres where the energy is needed.

It would be desirable to distribute smaller water-powered and wind-powered units throughout the population centres. For example, it would be desirable to have a wind-powered unit for each building structure, thus distributing the generating capacity over the entire area, and making the energy supply less vulnerable to local

events such as storms or earthquakes. Such distributed generation would also solve the most common and valid objection to wind power, namely, that the wind does not blow all the time. In a large geographical area, however, wind is almost always blowing somewhere. Therefore, with wind-powered generators which are distributed throughout the area, power could be generated in the areas where the wind is blowing, and then transmitted to the rest of the power grid. However, with existing technology, smaller units suitable for distributing throughout a population area are not efficient enough to provide a sufficient amount of energy to power a structure such as a house or office building. In addition, such units are visually obtrusive and noisy, making them unsuitable for use in residential or other highly populated settings.

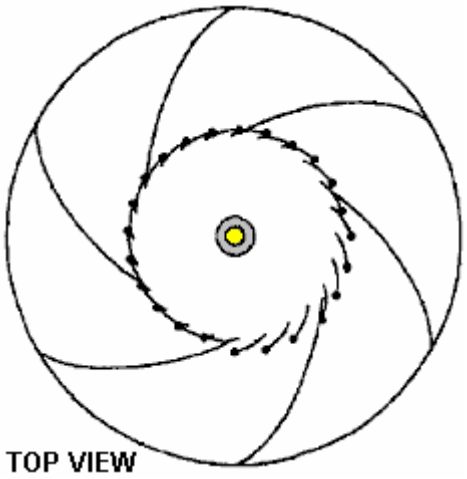
Existing wind-powered electrical generation devices commonly utilise a propeller mounted on the horizontal shaft of a generator which, in turn, is mounted at the top of a tower. This is an inefficient design because energy is extracted from the wind by reducing the wind velocity as it passes through the propeller. This creates a pocket of slow-moving air cantered behind the propeller, which the ambient wind blows around. Therefore, only the outer portion of the propeller blades use the wind efficiently.

To counter this effect, modern windmill designs utilise extremely long propeller blades. The use of such massive blades, however, has its own disadvantages. Firstly, the propellers are known to kill or injure thousands of large birds each year. Secondly, the massive blades can be dangerous if the device fails structurally and the propeller breaks loose. In this case, the propeller can fly a considerable distance and cause serious damage or injury to anything or anyone in its path. Thirdly, the propeller design contains an inherent gravitational imbalance. The rising blades on one side of the propeller's hub are opposing gravity, while the descending blades on the other side of the hub are falling with gravity. This imbalance creates a great deal of vibration and stress on the device. Consequently, the device must be structurally enhanced, at great expense, to withstand the vibration and stress, and thus avoid frequent maintenance and/or replacement.

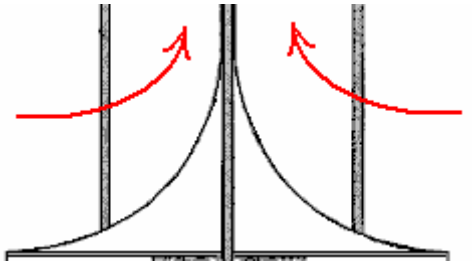
It would therefore be advantageous to have a fluid-powered energy conversion device which overcomes the shortcomings of existing devices. Such a device could utilise wind energy or the energy of flowing water to provide mechanical energy or electrical energy. The present invention provides such a device.

SUMMARY OF THE INVENTION

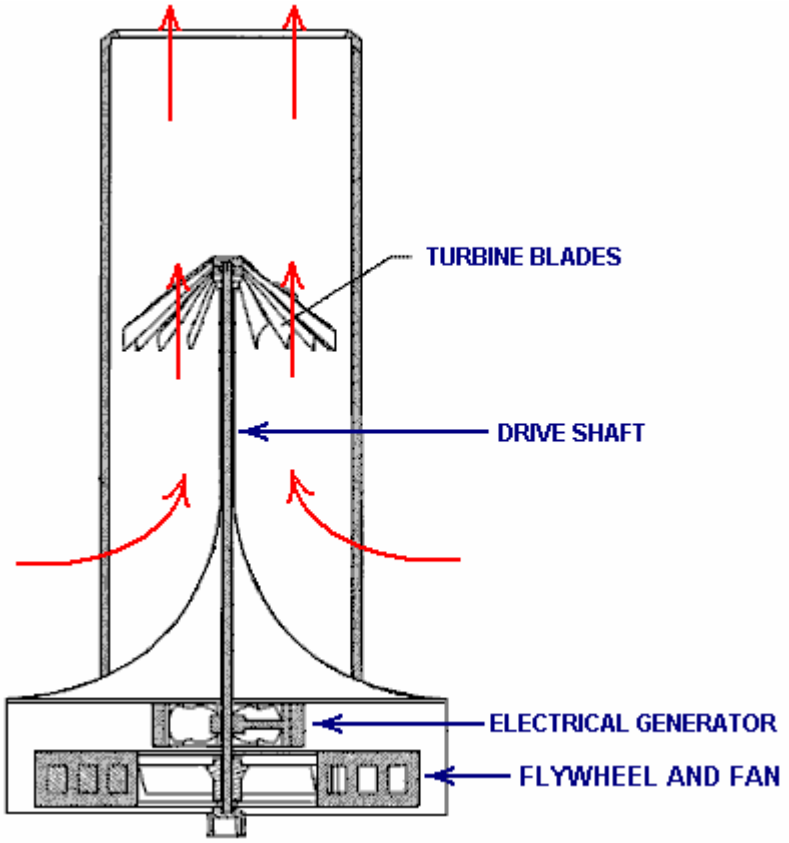
One aspect, the present invention is a fluid-powered energy-conversion device for converting energy in a moving fluid into mechanical energy. The device includes a rigid cylindrical frame which has an “upstream” annular (ring-shaped) chamber and a “downstream” annular chamber. Each of the chambers has sides which are open to allow entry of the moving fluid. A first set of baffles are mounted longitudinally in the upstream chamber, and these create a driving vortex which rotates in a first direction when the moving fluid enters the upstream chamber through the upstream chamber's open sides. A set of hinged louvers are positioned in the openings between these baffles, creating a central vortex chamber centered on the longitudinal axis of the device.



This first set of louvers permits entry of the moving fluid into the upstream central vortex chamber only when the fluid is rotating in the first direction. They also prevent the fluid from exiting from the upstream central vortex chamber through the sides of the device. The device also includes a floor of the upstream annular chamber which slopes upwards towards the downstream chamber as the floor approaches the central longitudinal axis of the device.



This sloping floor causes the drive vortex to flow “downstream” (upwards for air) through the upstream central vortex chamber and pass through a central aperture located between the upstream annular chamber and the downstream annular chamber. A longitudinal drive shaft is mounted centrally in the central aperture, and a turbine is mounted on the drive shaft in the central aperture. The turbine is rotated by the drive vortex as the drive vortex passes through the central aperture.



The device may also include a second set of baffles longitudinally mounted in the “downstream” (upper for air) chamber which operate to create an extraction vortex which rotates in the opposite direction when the moving fluid enters the downstream chamber through the downstream chamber's open sides. Additionally, a second set of hinged louvers may be positioned in the openings between the second set of baffles, encircling a downstream central vortex chamber. The second set of louvers permit entry of the moving fluid into the downstream central vortex chamber only when the fluid is rotating in the direction opposite to the direction of flow in the “upstream” camber. These louvers also prevent the fluid from exiting the downstream central vortex chamber through the sides of the device. In this manner, the turbine is rotated by the drive vortex as the drive vortex passes through the turbine and reverses direction to match the direction of the extraction vortex.

For high-wind conditions or when powered by water flow, the driving vortex and extraction vortex may rotate in the same direction. The first set of hinged louvers form the upstream central vortex chamber, and the second set of hinged louvers form the downstream central vortex chamber. The first set of louvers permit entry of the wind or water into the upstream central vortex chamber only when the fluid is rotating in the first direction.

BRIEF DESCRIPTION OF THE DRAWINGS

The invention will be better understood and its numerous objects and advantages will become more apparent to those skilled in the art by reference to the following drawings, in conjunction with the accompanying specification, in which:

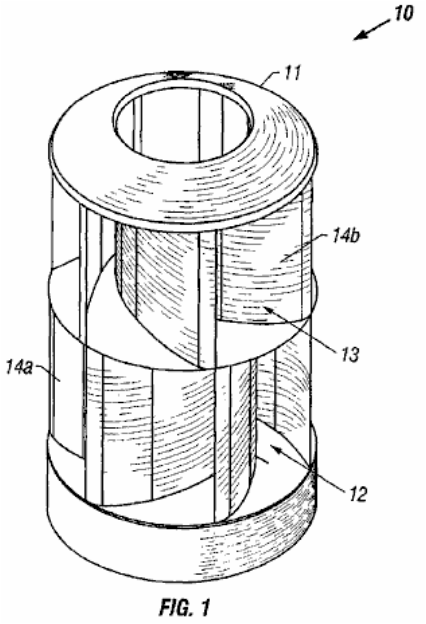


FIG. 1

FIG.1 is a perspective view of a first embodiment of the present invention that converts wind energy to mechanical or electrical energy;

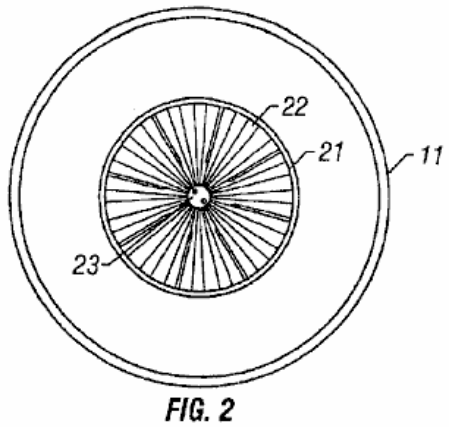


FIG. 2

FIG.2 is a top plan view of the embodiment of Fig.1

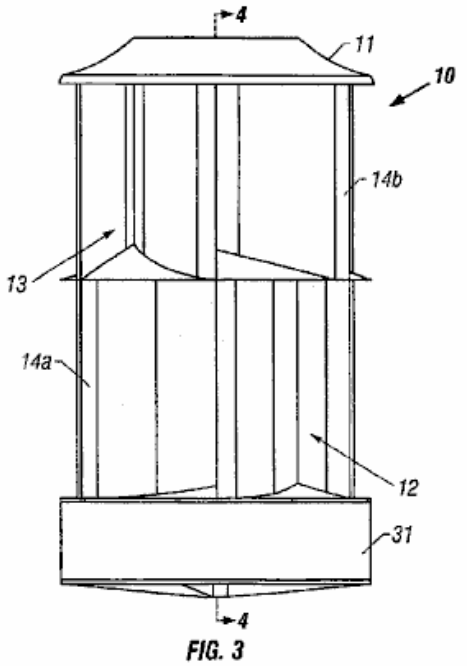


FIG. 3

FIG.3 is a side elevational view of the embodiment of Fig.1

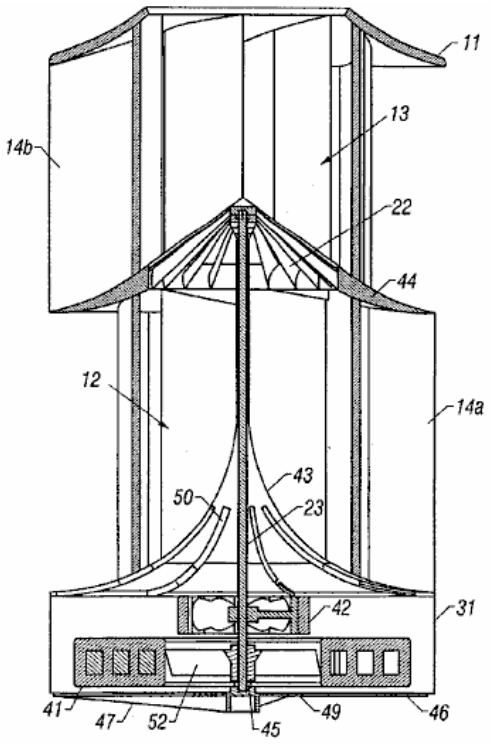


FIG. 4

FIG.4 is a cross-sectional view of the embodiment of Fig.1 taken along line **4** — **4** of Fig.3 with an electrical generator installed to produce electrical energy;

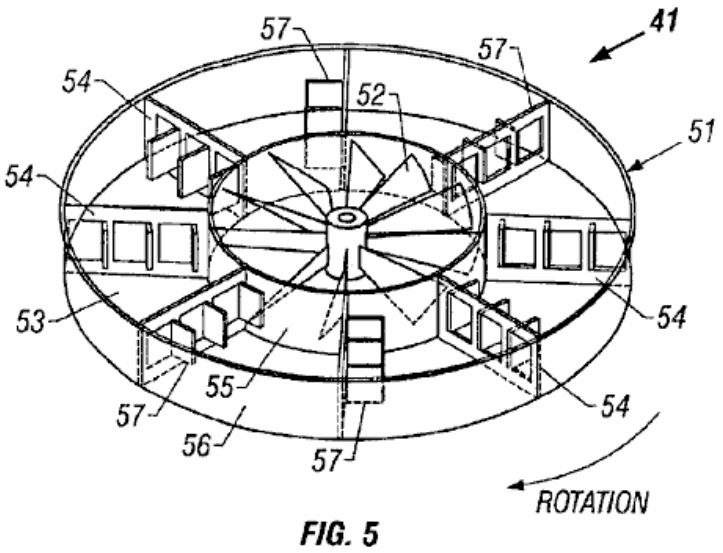


FIG. 5

FIG.5 is a perspective view of a fluid-filled flywheel suitable for use with the present invention;

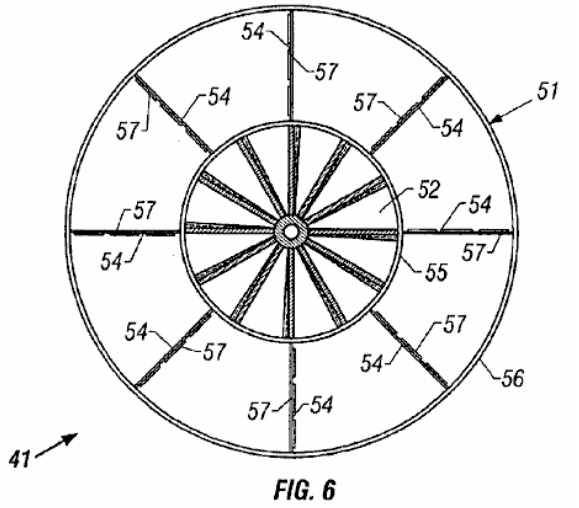


FIG. 6

FIG.6 is a top plan view of the fluid-filled flywheel of Fig.5

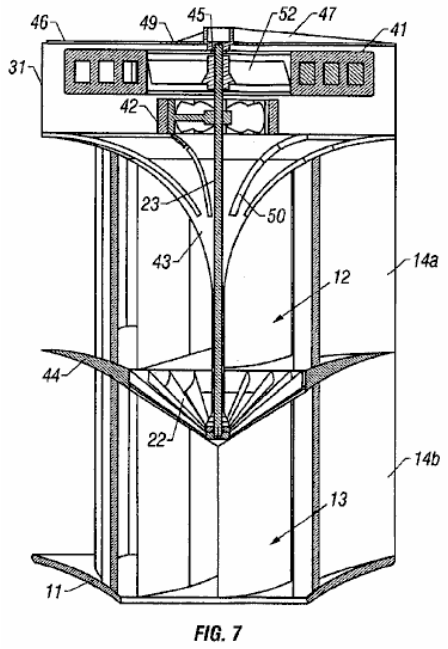


FIG. 7

FIG.7 is a cross-sectional view of an embodiment of the present invention that converts the energy of flowing **water** to electrical energy;

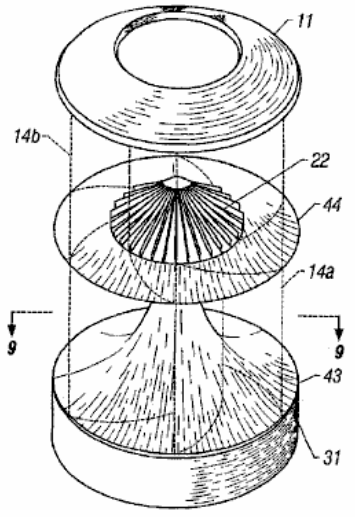


FIG. 8

FIG.8 is a perspective view of the embodiment of Fig.1 with the longitudinal baffles drawn in phantom so that the annular central divider (mid-deck) and turbine can be seen

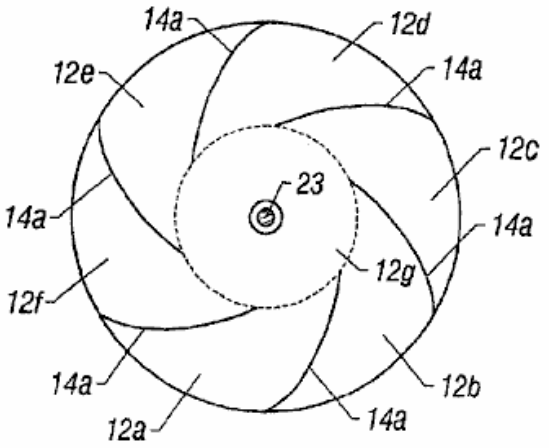


FIG. 9

FIG.9 is a horizontal cross-sectional view of the embodiment of Fig.1 taken along line **9** — **9** of Fig.8

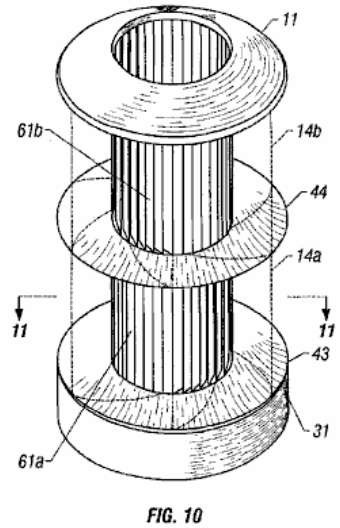


FIG. 10

FIG.10 is a perspective view of a second embodiment of the present invention that converts wind energy to mechanical or electrical energy, with the longitudinal baffles drawn in phantom so that a set of hinged longitudinal louvers can be seen; and

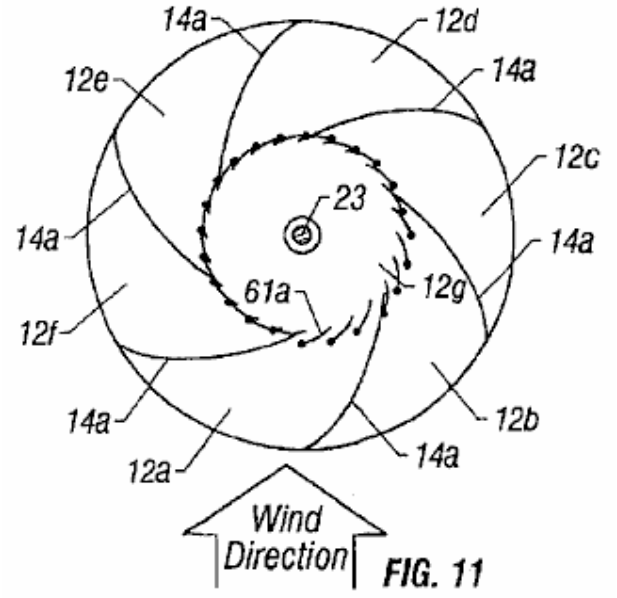


FIG. 11

FIG.11 is a horizontal cross-sectional view of the embodiment of Fig.10 taken along line **11** — **11** .

In the drawings, like or similar elements are designated with identical reference numerals throughout the various views, and the various elements shown are not necessarily drawn to scale.

DETAILED DESCRIPTION OF EMBODIMENTS

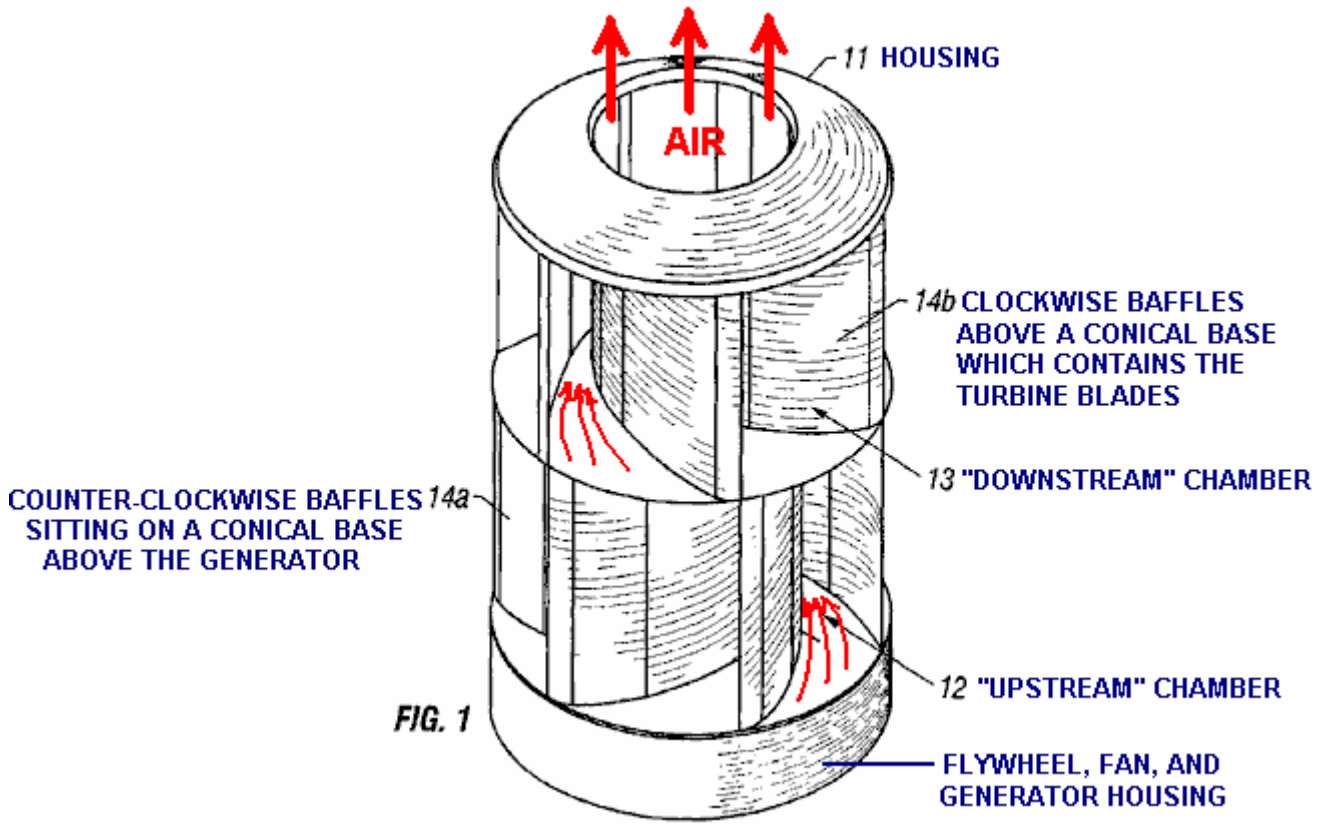


FIG. 1

Fig.1 is a perspective view of an embodiment of the present invention which converts wind energy to mechanical or electrical energy. The energy conversion device **10** includes a stationary cowling **11** surrounding an upstream (lower) ring-shaped or doughnut-shaped chamber **12** and a downstream (upper) ring-shaped chamber **13**. The cowling may be constructed of any suitable rigid material such as wood, plastic, metal, or similar. The cowling may be constructed from a transparent material, making the device visually unobtrusive. In the preferred embodiment of the present invention, the cowling is cylindrical and is constructed of a high-grade, ultraviolet-protected plastic.

The cowling **11** includes a set of longitudinal baffles which are curved and arranged in a toroidal pattern. Upstream baffles **14a** are mounted in the upstream annular chamber **12**, and downstream baffles **14b** are mounted in the downstream annular chamber **13**. In the preferred embodiment of the present invention, approximately six toroidal longitudinal baffles are mounted in each chamber. The baffles function to guide the wind into each chamber. The narrowing cross-sectional area between the baffles causes the air to accelerate as it moves toward the centre of the device, creating two high-velocity vortices (an upstream drive vortex and a downstream extraction vortex). Although the invention is described here primarily as a vertically-oriented cylinder, it should be understood that the device may be installed in other positions, such as a horizontal orientation, which results in the device having an upstream annular chamber and a downstream annular chamber which are at the same height. Alternatively, as noted below in connection with **Fig.7**, the device may be inverted when used in water since water vortices move more readily downwards rather than upwards.

In the embodiment illustrated in **Fig.1**, in which low-speed wind is the input energy source, the upstream baffles **14a** and the downstream baffles **14b** are curved in opposite directions. The baffles therefore create two high-velocity vortices which rotate in opposite directions. As described below in connection with **Fig.4** , the direction of the vortex flow is reversed in a turbine located between the upstream annular chamber **12** and the downstream annular chamber **13**, thereby adding additional rotational power to the turbine. In the hydro-electrical embodiment in which flowing water is the input energy source, and in high-speed wind conditions such as when the device is mounted on a vehicle, the upstream baffles and the downstream baffles may be curved in the same direction. In those particular embodiments, therefore, the baffles create two high-velocity vortices which rotate in the same direction. The device may be converted from a low-wind device to a high-wind device by removing the counter-

rotational downstream annular chamber **13** and replacing it with a downstream annular chamber which creates a vortex rotating in the same direction as the drive vortex.

In the preferred embodiment of the present invention, plastic mesh (not shown) may surround the entry and exit openings of the cowling **11** to prevent birds, animals, or debris from entering the device **10**. In addition, should the device fail structurally, any broken parts are contained by the mesh instead of flying out into the vicinity and causing damage or injury.

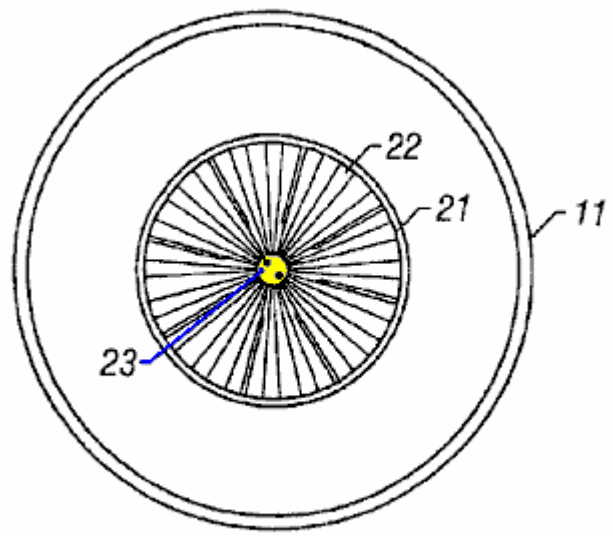


FIG. 2

Fig.2 is a top plan view of the embodiment of **Fig.1**. The top of the cowling **11** includes a central aperture **21** through which the air in the extraction vortex exits the device. In the preferred embodiment, the extraction vortex exits the device rotating in a counter-cyclonic direction (clockwise in the Northern Hemisphere) so that it dissipates rather than creating potentially damaging whirlwinds. The turbine **22** is visible through the aperture. The turbine rotates around a central drive shaft **23** .

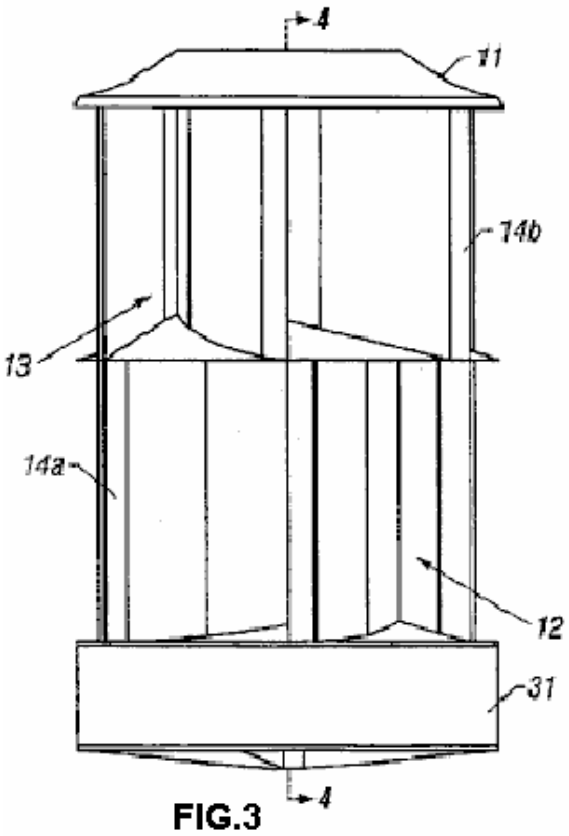


FIG.3

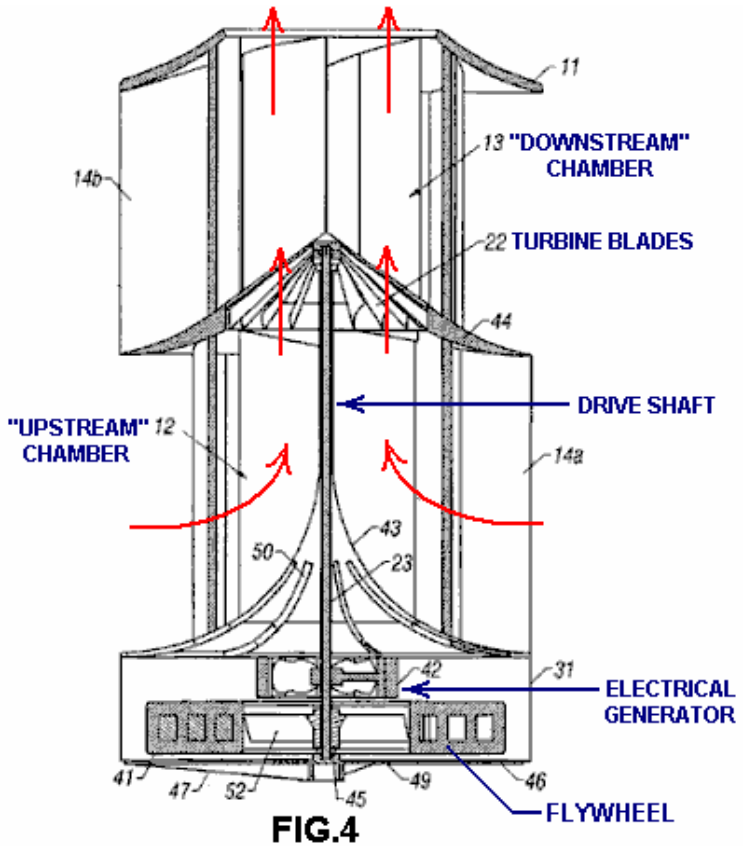


FIG.4

Fig.3 is a side-elevational view of the embodiment of **Fig.1** illustrating the profile of the cowling **11**, the upstream annular chamber **12**, the downstream annular chamber **13**, and the baffles **14a** and **14b**. The cowling may be

mounted on a base **31** and the base **31** may also be used to enclose additional mechanical assemblies such as a flywheel and/or an electrical generator.

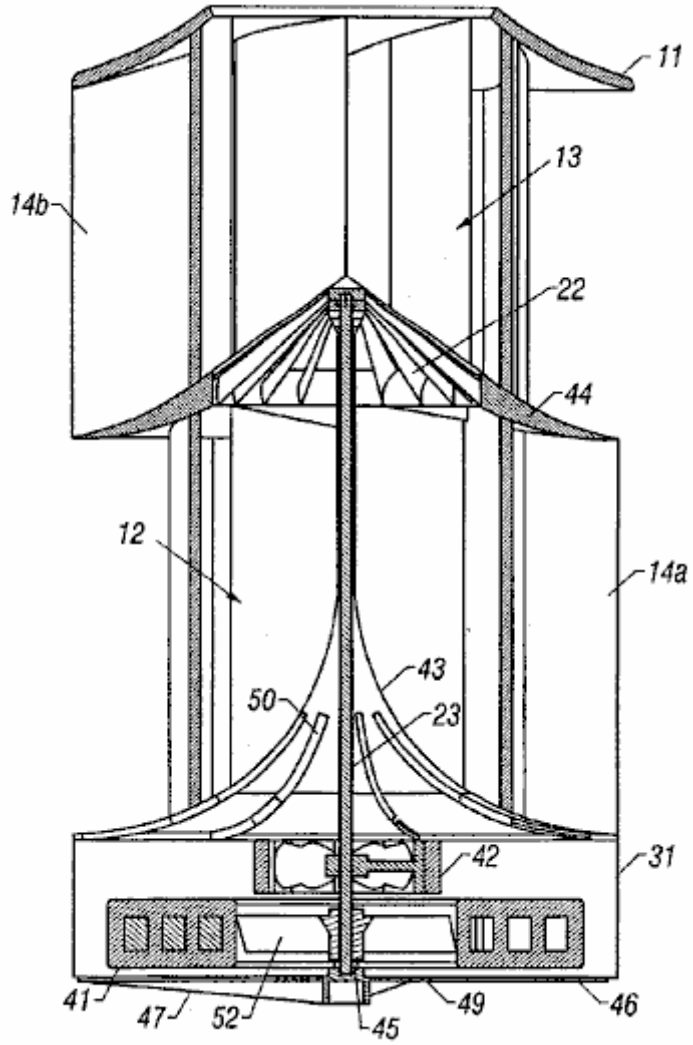


FIG. 4

Fig.4 is a cross-sectional view of the embodiment of **Fig.1** taken along line **4** — **4** of **Fig.3** with a flywheel **41** installed in the base **31** along with an electrical generator **42** to produce electricity. Ambient wind flows simultaneously into the upstream annular chamber **12** through upstream baffles **14a**, and into the downstream annular chamber **13** through the downstream baffles **14b** through the sides of the cowling **11**. The baffles guide the ambient wind towards the centre of the device **10**. A sloping parabolic floor (deck) **43** of the upstream annular chamber **12** causes the wind to flow downstream into the centrally mounted turbine **22** that rotates on the central drive shaft **23**. The device **10** produces power by guiding ambient wind flows into two high-velocity vortices arranged upstream and downstream of the turbine which converts the wind flows to mechanical energy by turning the drive shaft **23**. High-RPM and high-torque are produced by the turbine due to three primary factors:

(1) each blade of the turbine is shaped like a scoop which captures the rotational momentum of the drive vortex;

(2) each blade of the turbine has a cross-sectional shape of an airfoil that generates lift in the direction of rotation of the turbine; and

(3) in low wind conditions, the reversal of the direction of the vortex rotation adds additional force to the turbine in the direction of rotation.

The large flywheel **41** may be attached to the rotating turbine drive shaft **23**. In one embodiment, the flywheel may be a permanent magnet, surrounded by copper windings. The flywheel may serve both as an internal energy storage device due to its angular momentum, and as a dynamo for the generator **42** mounted under the deck **43** of the upstream annular chamber **12**. A solid-state electronic regulator (not shown) may be utilised to control the electrical current load. The regulator maintains a zero load until a preset rotational velocity (RPM) is reached. The load is then increased in order to generate electricity while maintaining the RPM of the turbine at a preselected level.

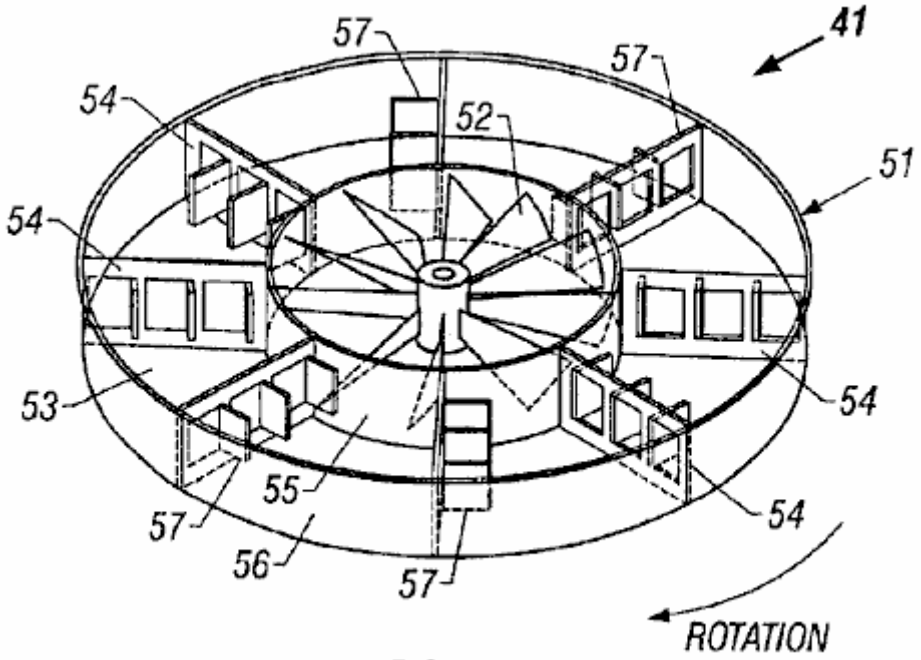


FIG. 5

In **Fig.5** is shown a perspective view of another embodiment of the flywheel **41**. In this embodiment, the flywheel (shown in phantom) includes a hollow disk-shaped shell **51** which is filled with a fluid such as water. The design shown also includes a cooling fan **52** in the hub of the flywheel which rotates with the drive shaft **23** and the flywheel to produce a flow of cooling air that is used to cool the adjacent generator **42** (**Fig.4** and **Fig.7**). The placement of the fan in the hub of the flywheel creates an annular chamber **53** whicht holds the fluid. Within the chamber, there is a set of radial bulkheads **54** extending from the interior wall **55** to the exterior wall **56** of the chamber. Each of the radial bulkheads includes hinged gates or hatches **57**. In the example version shown here, each radial bulkhead has three hinged gates.

During acceleration of the flywheel **41** , these gates **57** open in the opposite direction of rotation. This allows the fluid to flow through the radial bulkheads **54**, reducing start-up inertia. The fluid then slowly comes up to speed due to friction with the interior and exterior walls **55** and **56** of the annular chamber, and due to the motion of the radial bulkheads through the fluid. During deceleration of the flywheel, the gates close because of the forward momentum of the fluid. This creates solid radial bulkheads and causes the flywheel to perform as a solid flywheel. The angular momentum of the flywheel then helps to maintain the angular velocity of the drive shaft **23** when the input power of the wind drops off.

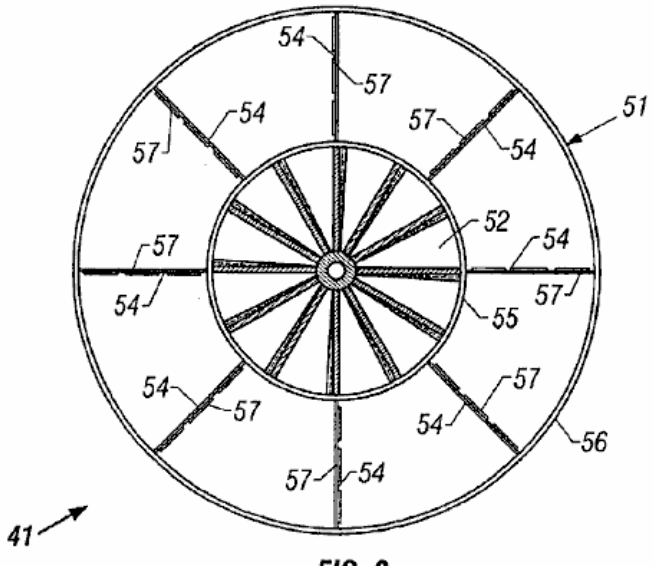


FIG. 6

Fig.6 is a top plan view of the fluid-filled flywheel **41** of **Fig.5** , showing the blades of the cooling fan **52** in the hub of the flywheel, the annular chamber **53**, the radial bulkheads **54**, and the gates **57** in the closed (decelerating) position.

Thus, the fluid-filled flywheel **41** is particularly well suited for use with this energy conversion device **10** of the present invention. The fluid-filled flywheel allows rapid spin-up of the drive shaft **23** by reducing the start-up inertia, but resists deceleration like a solid flywheel. These features can significantly boost the efficiency of a wind-powered or water-powered device that operates with varying input power levels. By simply inverting the flywheel, the fluid-filled flywheel can be used with systems that spin either clockwise or counter-clockwise. As an additional feature, shipping weight is greatly reduced because the fluid can be added at the point of use.

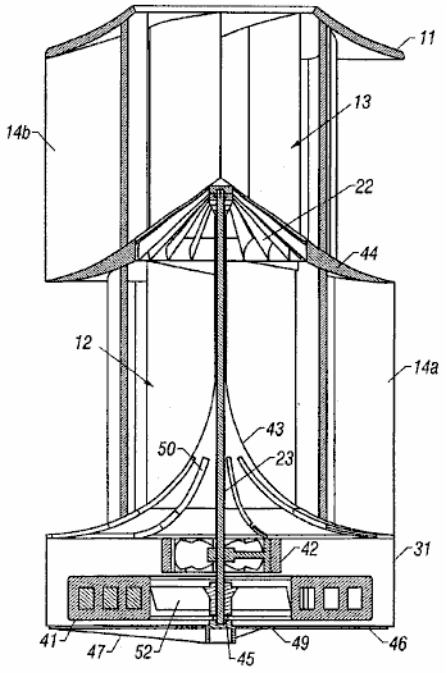


FIG. 4

Referring again to **Fig.4** , an annular central divider (mid-deck) **44** divides the upstream annular chamber **12** from the downstream annular chamber **13**. The top of the mid-deck slopes away from the turbine, causing the ambient wind entering the downstream annular chamber to flow away from the turbine. This creates an area of reduced air pressure on the downstream side of the turbine **22** that increases the flow of air from the upstream annular chamber **12** through the turbine. Each blade of the turbine **22** is a curved airfoil which receives rotational impetus from the rotation of the drive vortex, the reversal of the vortex direction, and aerodynamic lift that is generated by the airfoil in the direction of rotation of the turbine.

In the preferred embodiment of the present invention, the turbine **22** and flywheel **41** may be made of metal. Further, all metal parts may be coated with, for example, plastic, chrome, or paint to prevent corrosion. As discussed above, the flywheel may be a permanent magnet or may be a fluid-filled flywheel. All bearings such as bearing **45** may be magnetic-repulsion-levitation bearings so that there is no physical contact between the moving and stationary elements of the device. The base **31** may be mounted on a support plate **46** and/or a support brace **47**, depending on the structure on which the device is mounted and the orientation of the device.

The central drive shaft **23** may also drive the cooling fan **52** that draws cooling air through vents **49** in the support plate and directs the air through the generator **42**. The heated air may exit through louvers **50** in the parabolic deck **43** of the upstream annular chamber **12** where it then mixes with the driving airflow in the upstream annular chamber to defrost the interior of the device and the turbine **22**.

The device **10** may vary in its dimensions, depending upon the specific application for which it is utilised. For example, the dimensions of a wind-powered device that is mounted on the roof of a house may be between 40 inches and 48 inches in diameter, and between 60 inches and 78 inches in height. In this configuration, the turbine **22** has a diameter approximately one-half the diameter of the exterior of the cowling **11** (i.e. approximately 20 to 24 inches in diameter). Larger versions may be utilised for larger buildings such as factories or office buildings with increased economies of scale. For example, an office building may use a device that is 20 feet in diameter and 20 feet tall with a turbine that is 10 feet in diameter. A vehicle-mounted device (for example, for a passenger car), designed for high-wind conditions, may be about 24 inches in diameter and 6 inches in height. The generator and flywheel, if any, may be mounted inside the contour of the vehicle, or on a luggage rack. A

small hydro-electric version of the device that is placed in a running stream or river may have similar dimensions to the vehicle-mounted device. In addition, since the outflow of the hydro-electric version is directed downward, a deflector may be utilised in shallow bodies of water to prevent erosion of the stream bed.

It should be noted that when the present invention is oriented vertically, the turbine **22**, the generator **42**, and the flywheel **41** rotate around a vertical axis. Therefore, the supporting structures are not subject to the vibration and stress produced by gravity effects in prior art devices in which propellers rotate around a horizontal axis. Moreover, exceptional wind-conversion efficiency is realized from the present invention as it diverts and accelerates the ambient wind flow into vortices that have several times the velocity of the ambient wind flow when they reach the turbine. Additionally, the acceleration of the air flow into the upstream and downstream annular chambers creates a low pressure area that pulls air into the device from an effective cross-sectional area that is greater than the physical cross-sectional area of the device. As a result, the present invention provides a new and improved wind-power conversion device which is quieter, safer, more efficient, and more cost effective than existing devices.

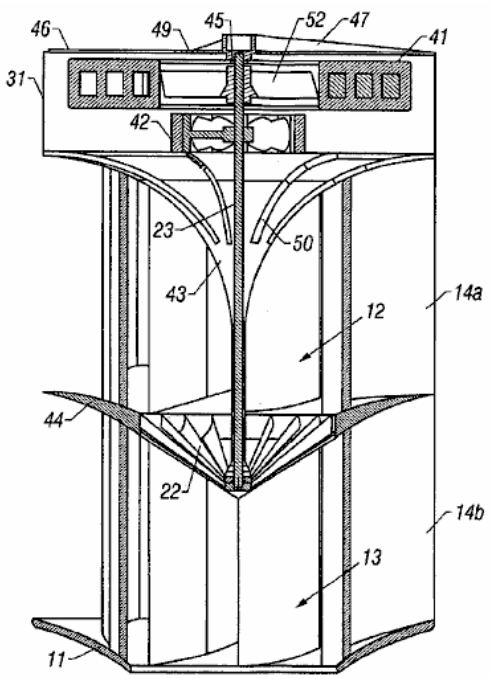


FIG. 7

Referring now to **Fig.7**, there is shown a cross-sectional view of a version of the present invention which converts the energy of flowing water to electrical energy (i.e. a hydro-electrical device). There are three key differences between the hydro-electrical embodiment from the low-wind-powered embodiment of **Figs. 1 to 4**. Firstly, the upstream baffles **14a** and the downstream baffles **14b** curve in the same direction. The baffles therefore create two high-velocity vortices which rotate in the same direction. This is a more efficient design when the fluid flowing through the device is an incompressible fluid such as water. Secondly, the device operates more efficiently when it is inverted and mounted vertically since water vortices move downward due to the force of gravity. The third difference is the ratio of the height of the device to the diameter of the device. As noted above, the hydro-electric embodiment of the device may have a height that is shorter when compared to its diameter, and may have a height that is equal to or less than its diameter.

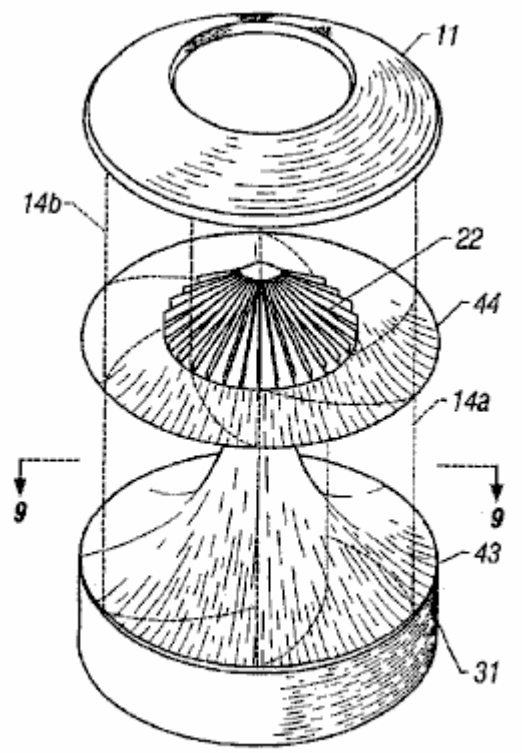


FIG. 8

Fig.8 is a perspective view of the embodiment of **Fig.1** with the toroidal longitudinal baffles **14a** and **14b** drawn in phantom so that the annular central divider (mid-deck) **44** and turbine **22** can be seen.

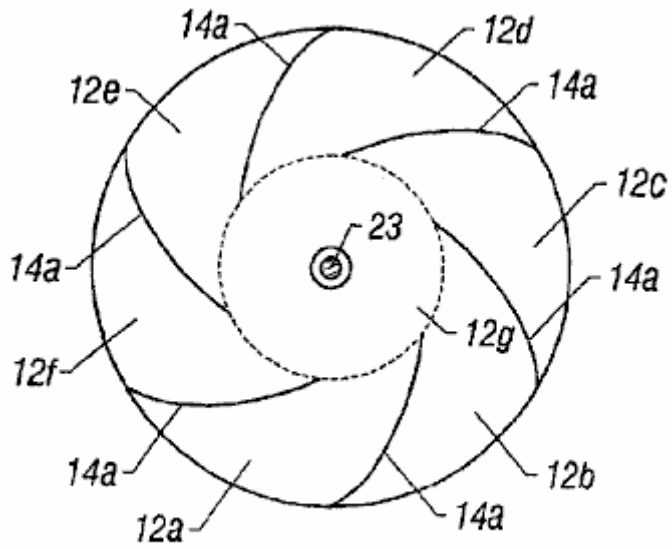


FIG. 9

Fig.9 is a horizontal cross-sectional view of the embodiment of **Fig.1** taken along line **9** — **9** of **Fig.8**. In this view, it can be seen that the upstream annular chamber **12** is divided into a set of smaller chambers **12a** through **12f** by the toroidal longitudinal baffles **14a**. The interior ends of the longitudinal baffles define a central vortex chamber **12g** (illustrated by a dashed circle) in which the upstream vortex is formed, and from which the upstream vortex enters the turbine **22**. The central vortex chamber **12g** has a diameter approximately equal to the diameter of the turbine.

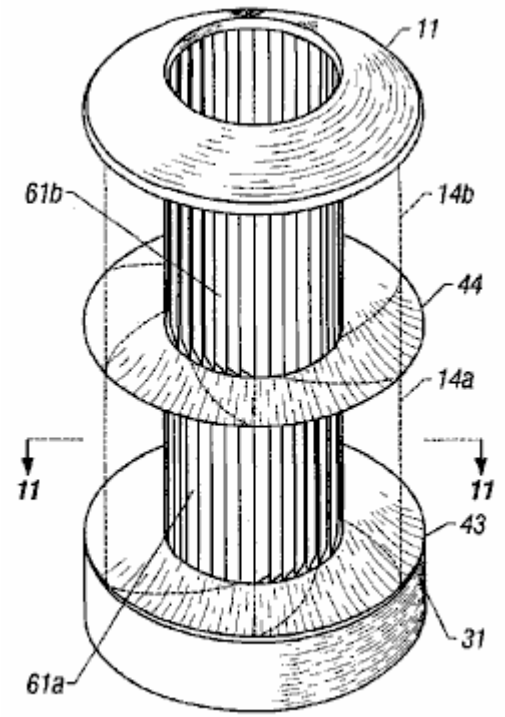


FIG. 10

Fig.10 is a perspective view of a second embodiment of the present invention that converts wind energy to mechanical or electrical energy, with the longitudinal baffles **14a** and **14b** drawn in phantom so that a set of hinged longitudinal louvers **61a** and **61b** can be seen. The hinged louvers are mounted in the openings between the longitudinal baffles. The louvers may be mounted in a circular configuration anywhere from the outside edge of the longitudinal baffles to the inside edge of the baffles. In the version shown, the louvers are longitudinally mounted at the inside edge of the baffles, around the perimeter of the central vortex chamber **12g**. Each of the louvers is hinged on one side (i.e., the windward side as wind enters through the baffles) so that the louver may be opened toward the central vortex chamber by the force of the incoming wind. The width of each louver is slightly greater than the distance between louvers so that each louver slightly overlaps the hinged edge of the next louver. This prevents the louvers from opening outward.

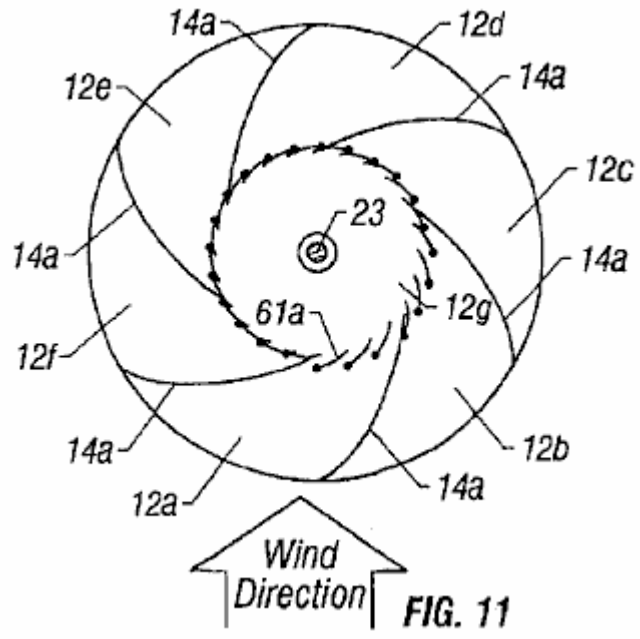


FIG. 11

In **Fig.11** there is shown a horizontal cross-sectional view of the embodiment of **Fig.10** taken along line **11** — **11**. During operation, wind blowing in the direction shown from the outside of the energy conversion device is funnelled by the toroidal longitudinal baffles **14a** into upstream chambers **12a** and **12b**. The baffles block the wind from entering the other chambers **12c** through **12f**. The wind flows through chambers **12a** and **12b**, and enters

the central vortex chamber **12g** by opening the hinged longitudinal louvers **61a** which are mounted between the baffles in the openings defining chambers **12a** and **12b**. The remaining louvers remain closed, preventing the wind from exiting through the sides of the device. Thus, the wind-activated louvers are, in effect, one-way valves allowing the wind to flow into the central vortex chamber through the sides of the device, but only allowing the wind to exit through the top of the chamber, and through the turbine **22**.

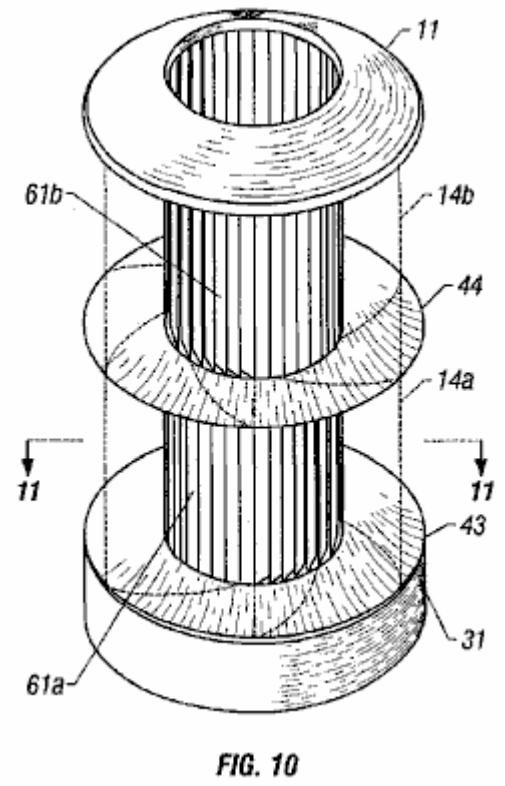


FIG. 10

Referring again to **Fig.10**, it can be seen that the longitudinal louvers **61a** mounted in the upstream chamber **12** are hinged on the opposite side from the louvers **61b** mounted in the downstream chamber **13**. This is because the vortex in the downstream chamber rotates in the opposite direction from the vortex in the upstream chamber, and the downstream toroidal baffles **14b** funnel the wind into the louvers **61b** in the opposite direction. Like the louvers **61a** in the upstream chamber **12**, the louvers **61b** in the downstream chamber **13** act as one-way valves allowing the wind to flow into the central vortex chamber through the sides of the device, but only allowing the wind to exit through the top of the chamber, and out of the device. This configuration helps to maintain the strength of both the upstream and the downstream vortices during operation of the device.

It should be recognized that some degree of improved energy-conversion performance may be obtained in a configuration in which there are toroidal baffles **14a** and hinged louvers **61a** only in the upstream annular chamber **12** because this ensures that all of the wind or other fluid entering the sides of the upstream chamber flows through the turbine. The addition of toroidal baffles **14b** in the downstream annular chamber **13** provides additional improved performance, particularly when the direction of rotation of the downstream vortex is opposite the direction of the upstream vortex. Optimum energy-conversion performance is provided by a device having oppositely configured toroidal baffles **14a** and **14b**, and oppositely hinged louvers **61a** and **61b**, for both the upstream annular chamber **12** and the downstream annular chamber **13**.

It is to be understood that even though numerous characteristics and advantages of the present invention have been set forth in the foregoing description, the disclosure is illustrative only, and changes may be made in detail, especially in matters of size, shape, and arrangement of parts within the principles of the invention to the full extent indicated by the broad general meaning of the terms in which the appended claims are expressed.

Claims:

What is claimed is:

1. A fluid-powered energy conversion device for converting energy in a moving fluid into mechanical energy, said device comprising: a rigid cylindrical frame having an upstream annular chamber and a downstream annular chamber, each of said chambers having sides that are open to allow entry of the moving fluid; a first set of baffles

longitudinally mounted in the upstream chamber that operate to create in the upstream chamber, an upstream drive vortex rotating in a first direction when the moving fluid enters the upstream chamber through the upstream chamber's open sides and through openings between the baffles; a first set of hinged louvers positioned in the openings between the first set of baffles and encircling an upstream central vortex chamber cantered around a central longitudinal axis of the device, said first set of louvers being operable to permit entry of the moving fluid into the upstream central vortex chamber only when the fluid is rotating in the first direction, and to prevent the fluid from exiting the upstream central vortex chamber through the sides of the device; a floor of the upstream annular chamber that slopes toward the downstream chamber as the floor approaches the central longitudinal axis of the device, said floor causing the drive vortex to flow downstream through the upstream central vortex chamber and pass through a central aperture located between the upstream annular chamber and the downstream annular chamber; a longitudinal drive shaft centrally mounted in the central aperture; and a turbine mounted on the drive shaft in the central aperture, said turbine being rotated by the drive vortex as the drive vortex passes through the central aperture.

2. The fluid-powered energy conversion device of claim 1 further comprising a second set of baffles longitudinally mounted in the downstream chamber that operate to create in the downstream chamber, a downstream extraction vortex rotating in a direction opposite to the first direction when the moving fluid enters the downstream chamber through the downstream chamber's open sides and through openings between the baffles, whereby the turbine is rotated by the drive vortex as the drive vortex passes through the turbine and reverses direction to match the direction of the extraction vortex.

3. The fluid-powered energy conversion device of claim 2 further comprising an annular central divider between the upstream chamber and the downstream chamber, said divider having a downstream surface that slopes downstream as it approaches the central longitudinal axis of the device, said downstream surface causing the extraction vortex to flow downstream, thereby creating an area of reduced fluid pressure downstream of the turbine.

4. The fluid-powered energy conversion device of claim 2 further comprising a second set of hinged louvers positioned in the openings between the second set of baffles and encircling a downstream central vortex chamber cantered around the central longitudinal axis of the device, said second set of louvers being operable to permit entry of the moving fluid into the downstream central vortex chamber only when the fluid is rotating in the direction opposite to the first direction, and to prevent the fluid from exiting the downstream central vortex chamber through the sides of the device.

5. The fluid-powered energy conversion device of claim 4 wherein said first set of baffles are curved to form a toroidal pattern in the first direction, and said second set of baffles are curved to form a toroidal pattern in the direction opposite to the first direction.

6. The fluid-powered energy conversion device of claim 5 wherein said turbine comprises a set of rotating blades, each of said blades having a cross-sectional shape of a curved airfoil that generates a lift force, said lift force being directed in the direction of rotation of the turbine.

7. The fluid-powered energy conversion device of claim 1 further comprising a flywheel mounted on the drive shaft, said flywheel having sufficient mass to operate as an internal energy storage device due to its angular momentum.

8. The fluid-powered energy conversion device of claim 7 wherein said flywheel is a permanent magnet.

9. The fluid-powered energy conversion device of claim 7 wherein said flywheel is a fluid-filled flywheel that rotates with the drive shaft in a direction of rotation, said fluid-filled flywheel comprising: a hollow disk-shaped shell filled with fluid; and a set of radial bulkheads that separate the interior of the shell into separate sections, each of said bulkheads having at least one gate pivotally mounted thereon to open in a direction opposite to the direction of rotation, said gate covering an aperture in the bulkhead when the gate is pivoted to a closed position, and said gate opening the aperture when the gate is pivoted to an open position; whereby the gates are opened by the fluid when the flywheel accelerates in the direction of rotation, thus allowing the fluid to flow through the apertures in the bulkheads and reduce start-up inertia of the flywheel, and whereby the gates are closed by the fluid when the flywheel decelerates, thus preventing the fluid from flowing through the apertures, and causing the flywheel to maintain angular momentum like a solid flywheel.

10. The fluid-powered energy conversion device of claim 9 wherein the hollow disk-shaped shell includes: an annular compartment filled with the fluid; and a cooling fan mounted in a central hub section of the shell.

11. The fluid-powered energy conversion device of claim 1 further comprising an electrical generator mounted on the drive shaft, said generator converting mechanical energy from the rotation of the shaft into electrical energy.

12. The fluid-powered energy conversion device of claim 11 further comprising a cooling fan mounted on the drive shaft, said cooling fan directing cooling air through the generator.

13. A wind-powered energy conversion device for converting wind energy into mechanical energy, said device comprising: a rigid cylindrical frame having an upstream annular chamber, a downstream annular chamber, and an annular central divider between the upstream chamber and the downstream chamber, each of said chambers having sides that are open to allow entry of ambient wind, and said annular central divider having a central aperture therein and having a downstream surface that slopes downstream as it approaches a central longitudinal axis of the device; a first set of baffles longitudinally mounted in the upstream chamber and curved to form a toroidal pattern that operates to create in an upstream central vortex chamber cantered around a central longitudinal axis of the device, an upstream drive vortex rotating in a first direction when the ambient wind enters the upstream chamber through the upstream chamber's open sides and through openings between the baffles; a first set of hinged louvers positioned in the openings between the first set of baffles and encircling the upstream central vortex chamber, said first set of louvers being operable to permit entry of the wind into the upstream central vortex chamber only when the wind is rotating in the first direction, and to prevent the wind from exiting the upstream central vortex chamber through the sides of the device; a second set of baffles longitudinally mounted in the downstream chamber and curved to form a toroidal pattern operable to create in a downstream central vortex chamber cantered around the central longitudinal axis of the device, a downstream extraction vortex rotating in a direction opposite to the first direction when the ambient wind enters the downstream chamber through the downstream chamber's open sides and through openings between the baffles; a second set of hinged louvers positioned in the openings between the second set of baffles and encircling the downstream central vortex chamber, said second set of louvers being operable to permit entry of the wind into the downstream central vortex chamber only when the wind is rotating in the direction opposite to the first direction, and to prevent the wind from exiting the downstream central vortex chamber through the sides of the device; a floor of the upstream annular chamber that slopes downstream as the floor approaches a central longitudinal axis of the device, said floor causing the drive vortex to flow downstream and pass through the central aperture in the annular central divider; a longitudinal drive shaft centrally mounted in the central aperture; and a turbine mounted on the drive shaft in the central aperture, said turbine comprising a set of rotating blades, each of said blades having a cross-sectional shape of a curved airfoil that generates a lift force, said lift force being directed in the direction of rotation of the turbine, said turbine being rotated by the drive vortex as the drive vortex passes through the turbine and reverses direction to match the direction of the extraction vortex.

14. The wind-powered energy conversion device of claim 13 further comprising a flywheel mounted on the drive shaft, said flywheel having sufficient mass to operate as an internal energy storage device due to its angular momentum.

15. The wind-powered energy conversion device of claim 13 further comprising an electrical generator mounted on the drive shaft, said generator converting mechanical energy from the rotation of the shaft into electrical energy.

16. The wind-powered energy conversion device of claim 13 wherein the extraction vortex rotates in a counter-cyclonic direction so that the extraction vortex dissipates after it exits the downstream chamber.

17. A wind-powered energy conversion device for converting high-speed wind energy into mechanical energy, said device comprising: a rigid cylindrical frame having an upstream annular chamber and a downstream annular chamber, each of said chambers having sides that are open to allow entry of the high-speed wind; a first set of baffles longitudinally mounted in the upstream chamber that create in an upstream central vortex chamber, an upstream drive vortex rotating in a first direction when the high-speed wind enters the upstream chamber through the upstream chamber's open sides and through openings between the baffles; a first set of hinged louvers positioned in the openings between the first set of baffles and encircling the upstream central vortex chamber, said first set of louvers being operable to permit entry of the wind into the upstream central vortex chamber only when the wind is rotating in the first direction, and to prevent the wind from exiting the upstream central vortex chamber through the sides of the device; a second set of baffles longitudinally mounted in the downstream chamber that create in a downstream central vortex chamber, a downstream extraction vortex rotating in the first direction when the high-speed wind enters the downstream chamber through the downstream chamber's open sides and through openings between the baffles; a second set of hinged louvers positioned in the openings between the second set of baffles and encircling the downstream central vortex chamber, said second set of louvers being operable to permit entry of the wind into the downstream central vortex chamber only when the wind is rotating in the first direction, and to prevent the wind from exiting the downstream central vortex chamber through the sides of the device; a floor of the upstream annular chamber that slopes downstream as the floor approaches a central longitudinal axis of the device, said floor causing the drive vortex to flow downstream and pass through a central aperture located between the upstream annular chamber and the downstream annular chamber; a longitudinal drive shaft centrally mounted in the central aperture; and a turbine mounted on the drive

shaft in the central aperture, said turbine being rotated by the drive vortex as the drive vortex passes through the turbine.

18. The wind-powered energy conversion device of claim 17 further comprising a fluid-filled flywheel mounted on the drive shaft that rotates with the drive shaft in a direction of rotation, said fluid-filled flywheel comprising: a hollow disk-shaped shell filled with fluid; and a set of radial bulkheads that separate the interior of the shell into separate sections, each of said bulkheads having at least one gate pivotally mounted thereon to open in a direction opposite to the direction of rotation, said gate covering an aperture in the bulkhead when the gate is pivoted to a closed position, and said gate opening the aperture when the gate is pivoted to an open position; whereby the gates are opened by the fluid when the flywheel accelerates in the direction of rotation, thus allowing the fluid to flow through the apertures in the bulkheads and reduce start-up inertia of the flywheel, and whereby the gates are closed by the fluid when the flywheel decelerates, thus preventing the fluid from flowing through the apertures, and causing the flywheel to maintain angular momentum like a solid flywheel.

19. A water-powered energy conversion device for converting energy in a moving stream of water into mechanical energy, said device comprising: a rigid cylindrical frame having an upstream annular chamber and a downstream annular chamber, each of said chambers having sides that are open to allow entry of the stream of water; a first set of baffles longitudinally mounted in the upstream chamber that create in an upstream central vortex chamber, an upstream drive vortex rotating in a first direction when the stream of water enters the upstream chamber through the upstream chamber's open sides and through openings between the baffles; a first set of hinged louvers positioned in the openings between the first set of baffles and encircling the upstream central vortex chamber, said first set of louvers being operable to permit entry of the water into the upstream central vortex chamber only when the water is rotating in the first direction, and to prevent the water from exiting the upstream central vortex chamber through the sides of the device; a second set of baffles longitudinally mounted in the downstream chamber that create in a downstream central vortex chamber, a downstream extraction vortex rotating in the first direction when the stream of water enters the downstream chamber through the downstream chamber's open sides and through openings between the baffles; a second set of hinged louvers positioned in the openings between the second set of baffles and encircling the downstream central vortex chamber, said second set of louvers being operable to permit entry of the water into the downstream central vortex chamber only when the water is rotating in the first direction, and to prevent the water from exiting the downstream central vortex chamber through the sides of the device; a floor of the upstream annular chamber that slopes downstream as the floor approaches a central longitudinal axis of the device, said floor causing the drive vortex to flow downstream and pass through a central aperture located between the upstream annular chamber and the downstream annular chamber; a longitudinal drive shaft centrally mounted in the central aperture; and a turbine mounted on the drive shaft in the central aperture, said turbine being rotated by the drive vortex as the drive vortex passes through the turbine.

20. The water-powered energy conversion device of claim 19 further comprising a fluid-filled flywheel mounted on the drive shaft that rotates with the drive shaft in a direction of rotation, said fluid-filled flywheel comprising: a hollow disk-shaped shell filled with fluid; and a set of radial bulkheads that separate the interior of the shell into separate sections, each of said bulkheads having at least one gate pivotally mounted thereon to open in a direction opposite to the direction of rotation, said gate covering an aperture in the bulkhead when the gate is pivoted to a closed position, and said gate opening the aperture when the gate is pivoted to an open position; whereby the gates are opened by the fluid when the flywheel accelerates in the direction of rotation, thus allowing the fluid to flow through the apertures in the bulkheads and reduce start-up inertia of the flywheel, and whereby the gates are closed by the fluid when the flywheel decelerates, thus preventing the fluid from flowing through the apertures, and causing the flywheel to maintain angular momentum like a solid flywheel.

21. A fluid-powered energy conversion device for converting energy in a moving fluid into mechanical energy, said device comprising: a rigid cylindrical frame having an upstream annular chamber and a downstream annular chamber cantered around a longitudinal axis, each of said chambers having sides that are open to allow entry of the moving fluid in a direction approximately perpendicular to the longitudinal axis, said upstream and downstream chambers being separated by an annular divider having a central aperture therein; a longitudinal drive shaft centrally mounted along the longitudinal axis and passing through the central aperture; a turbine mounted on the drive shaft in the central aperture; means for creating in the upstream chamber an upstream drive vortex rotating in a first direction when the moving fluid enters the upstream chamber through the upstream chamber's open sides: means for creating in the downstream chamber, a downstream extraction vortex rotating in a second direction opposite to the first direction when the moving fluid enters the downstream chamber through the downstream chamber's open sides; end means for causing the drive vortex to flow downstream and pass through the turbine, said turbine being rotated by the drive vortex as the drive vortex passes through the turbine and reverses direction to match the direction of the extraction vortex.

22. The fluid-powered energy conversion device of claim 21 wherein the means for creating an upstream drive vortex in the upstream chamber includes a first set of longitudinally mounted baffles having openings between

them through which the moving fluid enters the upstream chamber, said first set of baffles being curved to form a toroidal pattern in the first direction.

23. The fluid-powered energy conversion device of claim 22 wherein the means for creating an upstream drive vortex in the upstream chamber includes a first set of hinged louvers positioned in the openings between the first set of baffles and encircling the upstream chamber, said first set of louvers being operable to permit entry of the moving fluid into the upstream chamber only when the fluid is rotating in the first direction, and to prevent the fluid from exiting the upstream chamber through the sides of the device.

24. The fluid-powered energy conversion device of claim 23 wherein the means for creating a downstream extraction vortex in the downstream chamber includes a second set of longitudinally mounted baffles having openings between them through which the moving fluid enters the upstream chamber, said second set of baffles being curved to form a toroidal pattern in the second direction.

25. The fluid-powered energy conversion device of claim 24 wherein the means for creating a downstream extraction vortex in the downstream chamber includes a second set of hinged louvers positioned in the openings between the second set of baffles and encircling the downstream chamber, said second set of louvers being operable to permit entry of the moving fluid into the downstream chamber only when the fluid is rotating in the second direction, and to prevent the fluid from exiting the downstream chamber through the sides of the device.

26. The fluid-powered energy conversion device of claim 25 wherein the means for causing the drive vortex to flow downstream includes means for creating a pressure differential in which the fluid pressure in the downstream chamber is less than the fluid pressure in the upstream chamber.

27. The fluid-powered energy conversion device of claim 26 wherein the means for creating a pressure differential includes a downstream surface of the annular divider that slopes downstream as it approaches the central longitudinal axis of the device, said downstream surface causing the extraction vortex to flow downstream, thereby creating an area of reduced fluid pressure downstream of the turbine.

28. The fluid-powered energy conversion device of claim 27 wherein the means for causing the drive vortex to flow downstream includes a floor of the upstream annular chamber that slopes toward the downstream chamber as the floor approaches the central longitudinal axis of the device, said floor causing the drive vortex to flow downstream and pass through the turbine.

29. The fluid-powered energy conversion device of claim 21 further comprising a fluid-filled flywheel that rotates with the drive shaft in a direction of rotation, said fluid-filled flywheel comprising: a hollow disk-shaped shell filled with fluid: and a set of radial bulkheads that separate the interior of the shell into separate sections, each of said bulkheads having at least one gate pivotally mounted thereon to open in a direction opposite to the direction of rotation, said gate covering an aperture in the bulkhead when the gate is pivoted to a closed position, and said gate opening the aperture when the gate is pivoted to an open position; whereby the gates are opened by the fluid when the flywheel accelerates in the direction of rotation, thus allowing the fluid to flow through the apertures in the bulkheads and reduce start-up inertia of the flywheel, and whereby the gates are closed by the fluid when the flywheel decelerates, thus preventing the fluid from flowing through the apertures, and causing the flywheel to maintain angular momentum like a solid flywheel.

Nilson Barbosa and Cleriston Leal

WO Patent 2013/104043 18th July 2013 Inventors: Nilson Barbosa and Cleriston Leal

ELECTRIC ENERGY GENERATION SYSTEM WITH FEEDBACK

Note: These three patents are in Portuguese and what is shown here is a low-quality attempt at translation into English using a translation program. The originals can be downloaded free from: http://worldwide.espacenet.com/singleLineSearch?locale=en_EP.

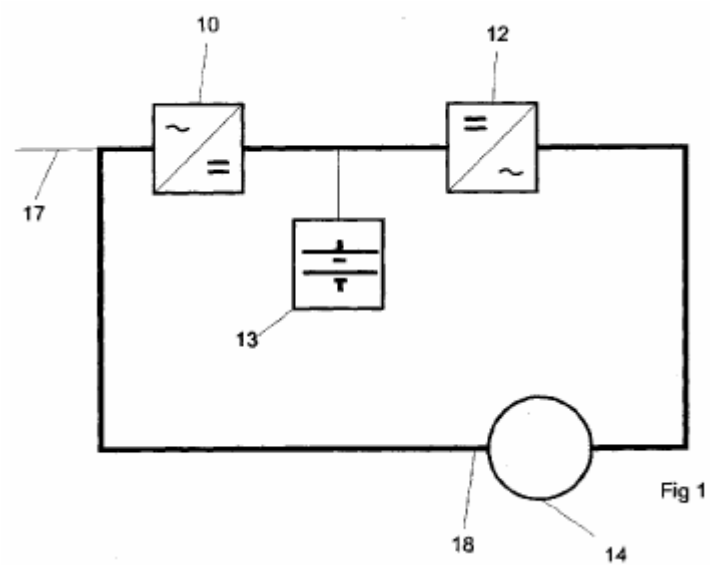


Fig 1

Abstract:
The present invention relates to electric energy generation equipment comprising a basic circuit formed by a rectifier (**10**), for example, an AC/DC converter connected in series to an inverter (**12**), for example, a DC/AC converter, and a bank of batteries (**13**) connected in series between the rectifier (**10**) and the inverter (**12**). An electron-capturing element (**14**), which can be either a free space electron-capturing element or, alternatively, an earth electron-capturing element, is connected in series to the basic circuit formed by the rectifier (**10**), the inverter (**12**) and the battery assembly (**13**). The bank of batteries (**13**) powers the basic circuit because it is connected to the system. Consequently, the inverter (**12**) converts direct current into alternating current and supplies this current to the electron-capturing element (**14**). After receiving the electric current from the inverter (**12**), the electron-capturing element (**14**) starts capturing electrons from the alternating current and powering the rectifier (**10**), which converts the alternating current into a direct current in order to recharge the bank of batteries (**13**) and power the inverter (**12**) which powers the electron-capturing element, closing the feedback loop, and also providing electric energy for consumption by external loads.

WIPO Patent Application WO/2013/104043 Filing Date: 01/11/2013
Application Number: BR2013/000016 Publication Date: 07/18/2013

Assignee: EVOLUÇÕES ENERGIA LTDA (Rua Santa Tereza 1427-B Centro - Imperatriz -MA, CEP -470 - Maranhão, 65900, BR)

SELF-POWERED ELECTRICITY GENERATOR.

Technical field
The present invention relates to a device for generating electricity, in particular self-powered equipment for generating electricity.

Description of the Related Art
There are many methods for generating electricity using electromagnetism, but all of these are electromechanical devices using magnets and have limited generating capacity and an ecological impact which makes them unsuited to large scale projects.

Objectives of the Invention

The aim of this invention is the sustainable generation of electricity, using a generator which is able to produce large amounts of electricity from an extremely low input current, which initially is supplied by a bank of batteries, but subsequently is supplied by the output from the generator which is also able to power external loads.

The above objective, and other objectives, are achieved by the present invention through the use of a typical Uninterruptible Power Supply circuit comprising of an AC/DC rectifier feeding a battery bank which powers a DC/AC inverter, which is connected to a device to trap electrons from space (as described in Brazilian patent application No. BR1020120008378 of 13th January 2012) or alternatively, a device which extracts electrons from the Earth (as described in Brazilian patent application No. BR1020120008386 of 13th January 2012), which then passes the extracted electrons to the AC/DC rectifier, charging the battery bank, thus closing the loop as well as providing electricity to power external loads.

The self-powered system for generating electricity from the present invention can be fixed or mobile. It is fixed when using electron capture from the earth due to the ground connection, or mobile when using electron capture from space.

The self-powered electricity generating system of this invention may be configured in several different ways, each using the same inventive concept but using different arrangements of components. Different versions include single-phase, two-phase or three-phase versions, producing outputs of any power and voltage.

Brief Description of the Drawings

The present invention will now be described with the aid of drawings, but this patent is not limited to the versions and details shown in these drawings, although they show additional details and advantages of the present invention.

The drawings:

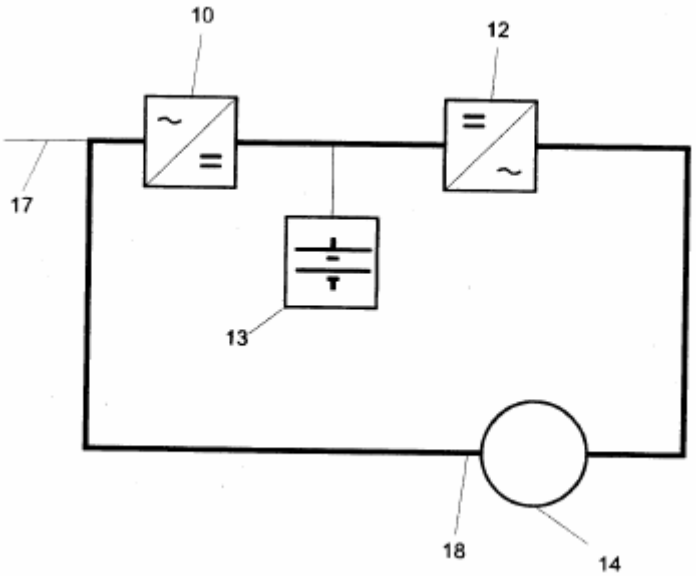


Figure 1 - shows a basic circuit system for self-powered electricity generation of the present invention

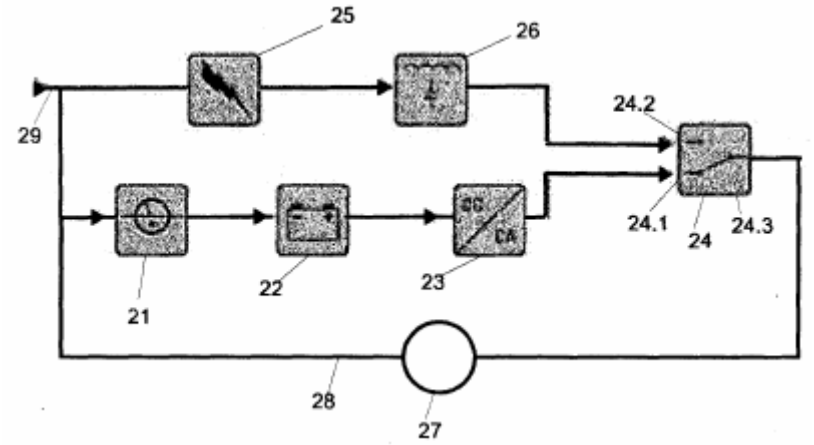


Figure 2 - shows a first embodiment of the constructive system for self-powered electricity generation of the present invention;

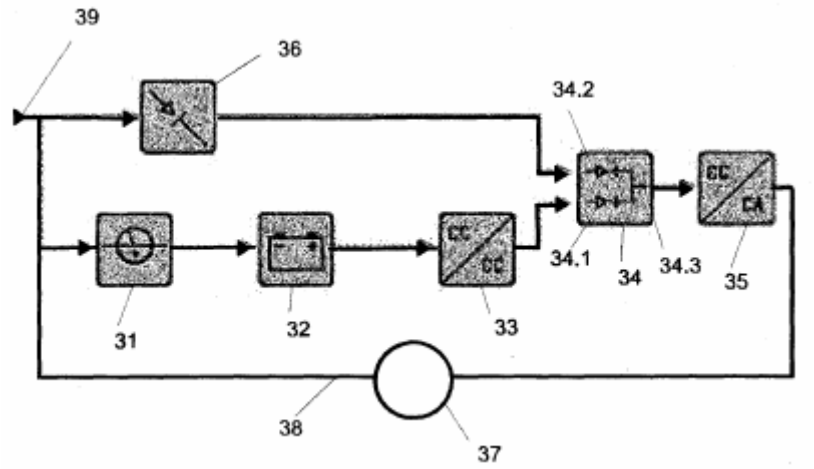


Figure 3 - shows a second embodiment of the self-powered system for generating electricity of the present invention;

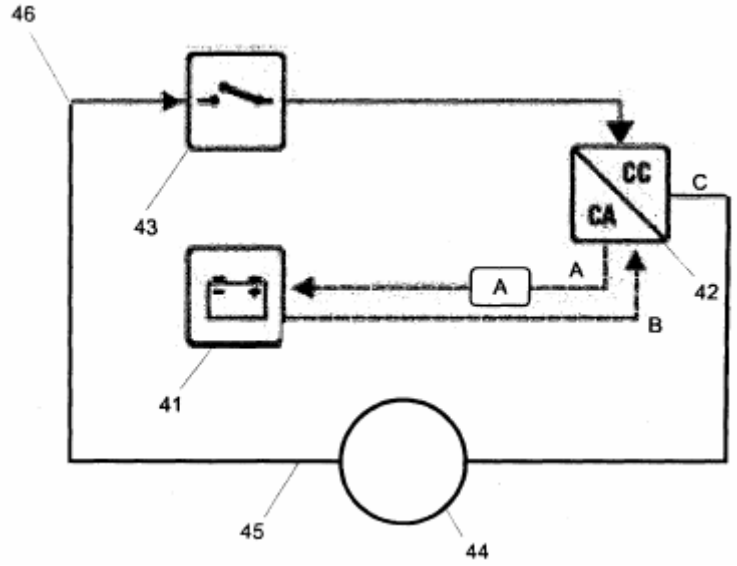


Figure 4 - shows a third embodiment of the self-powered system for generating electricity of the present invention;

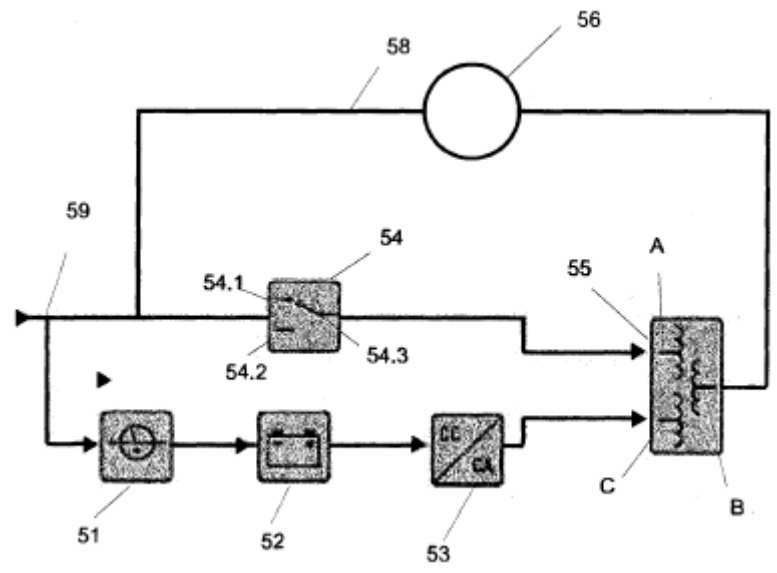


Figure 5 - shows a fourth embodiment of the self-powered system for generating electricity of the present invention;

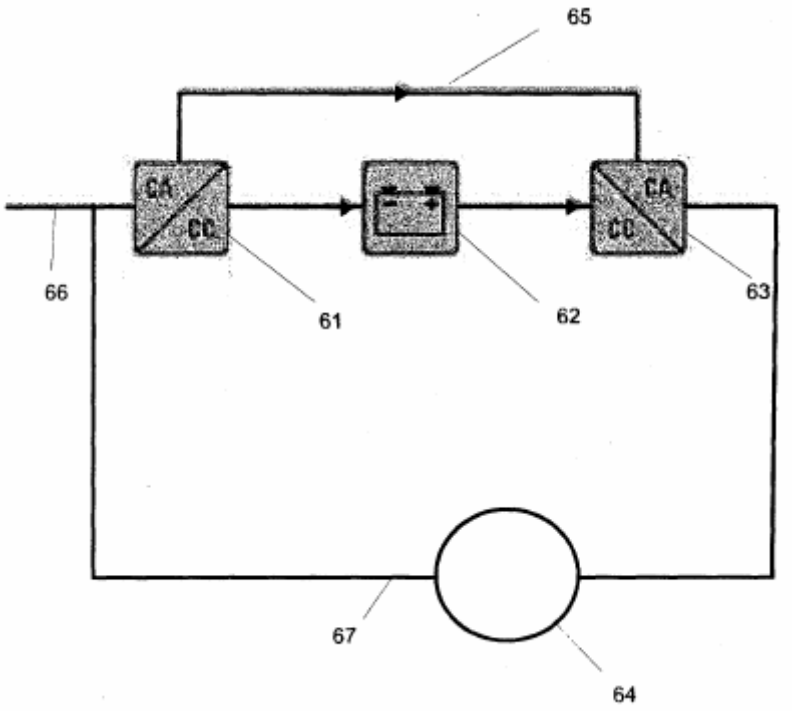


Figure 6 - shows a fifth embodiment of the self-powered system for generating electricity of the present invention;

Detailed description of the Invention:
There are different ways of closing the self-feeding cycle depending on the circuit configuration chosen. Some of these arrangements are shown in **Figures 2** to **6**, wherein the main circuitry continues to oscillate, continuously generating instant electricity.

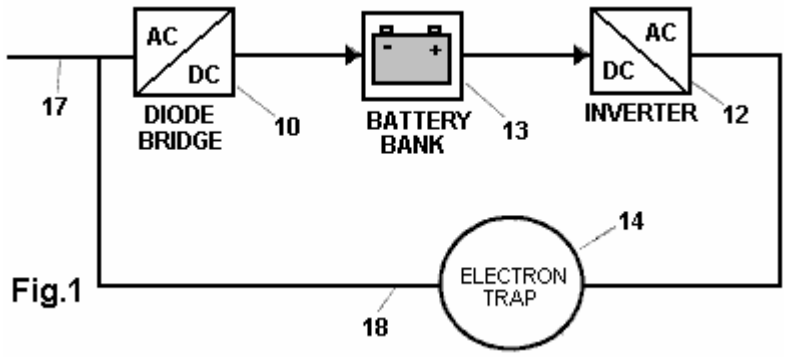


Fig.1

As shown in **Fig.1**, the self-powered system for generating electricity comprises a basic circuit consisting of a rectifier (AC/DC converter) **10** which is connected in series to an inverter (DC/AC) **12**. A bank of batteries **13** is connected between the rectifier **10** and the inverter **12**. The output from the DC/AC inverter **12**, connects to an electron-trap **14** which can extract electrons from space (as described in Brazilian patent application No. BR1020120008378 of 13th January 2012) or alternatively, extracts electrons from the Earth (as described in Brazilian patent application No. BR1020120008386 of 13th January 2012).

When connected, the battery bank **13** provides power to the DC/AC inverter **12** which converts the direct current into alternating current and provides current to the electron-trap **14**. The output of the electron trap **14** is passed through wire **18**, to the AC/DC bridge rectifier **10**, which keeps the battery bank charged as well as powering the DC/AC inverter **12**. Additional power is passed to external equipment through wire **17**.

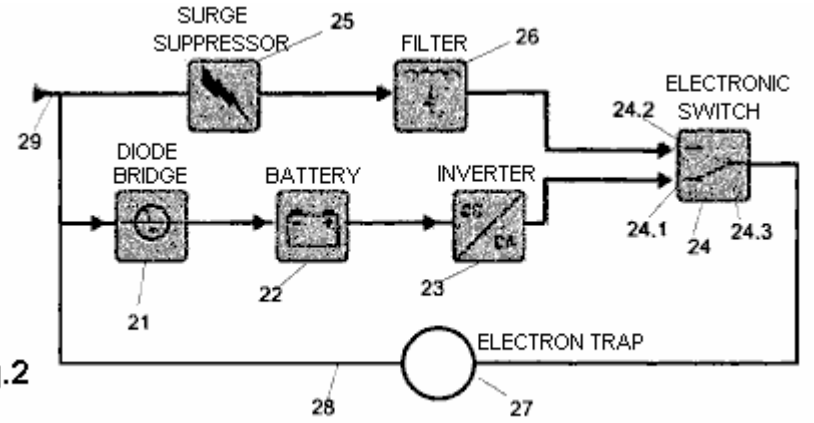


Fig.2

Fig.2, shows another embodiment of the system of this self-powered electric power generation equipment. It comprises a typical Uninterruptible Power Supply circuit of a battery charger (AC/DC converter) **21** connected to a drive device (a DC/AC inverter) **23** and between them, a battery bank **22** forming the basic circuit. Additional devices are an electron-trap **27** which may collect free electrons from space (as defined in Brazilian patent application No. BR1020120008378 of 13th January 2012) or, alternatively, collects electrons from the Earth (as described in Brazilian patent application No. BR1020120008386 of 13th January 2012). The 3-phase electronic switch **24** normally connects **24.1** to **24.3** connecting the electron trap **27** to inverter **23**. Connected in parallel is the surge suppressor **25**, which, when activated, via filter **26**, causes switch **24** to disconnect the **24.3** to **24.1** link and instead, connect **24.3** to **24.2**.

An alternative arrangement for use in emergency situations, is to use the system no longer self-powered. For this, the system is comprised of a power input from an external power source, directly to the interconnection point **29** to provide power to surge suppressor **25**, which provides power to feed the power output point **28** in order to power external loads. When the electron-trap **27** is turned off, the electronic transfer switch **24** reverts to its default position which connects point **24.1** to point **24.3** causing the circuit to function, once again, in its self-feeding mode. As soon as the electron sensor **27** provides sufficient power to the over-voltage sensor **25**, it operates the transfer switch **24** through filter **26**, ending the self-feeding phase and supplying energy directly to the power output point **28**, in order to feed external loads.

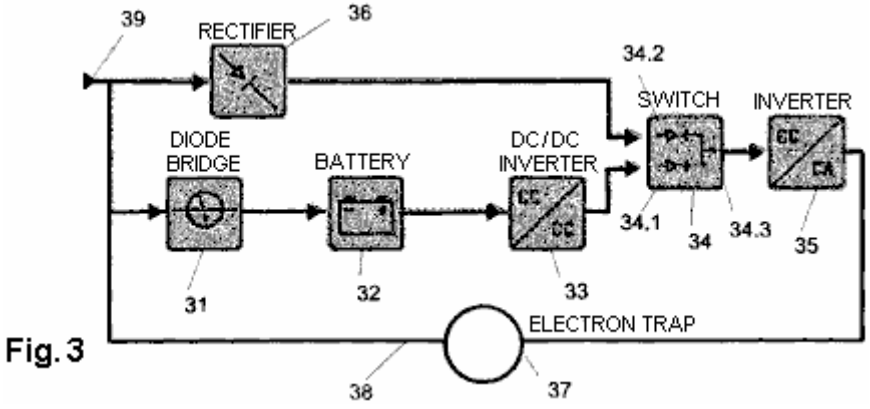


Fig. 3

Fig.3 shows another embodiment of the self-powered system for generating electricity, comprising a device which includes the basic circuit of a typical Uninterruptible Power Supply, consisting of a battery charger (AC/DC converter) **31** connected to a drive device (inverter DC/AC) **35** and attached to them, a battery bank **32**. This basic circuit together with other devices is connected to an electron-trap **37** for collecting free electrons from surrounding space or, alternatively, an Earth-connected electron trap **37**. We have then, a bank of batteries **32** connected to the DC/DC converter **33**, which is connected to the phase transfer switch **34 / 34.1** which is connected to point **34.3**, which connects to the inverter **35**, and so, the electron-trap **37**.

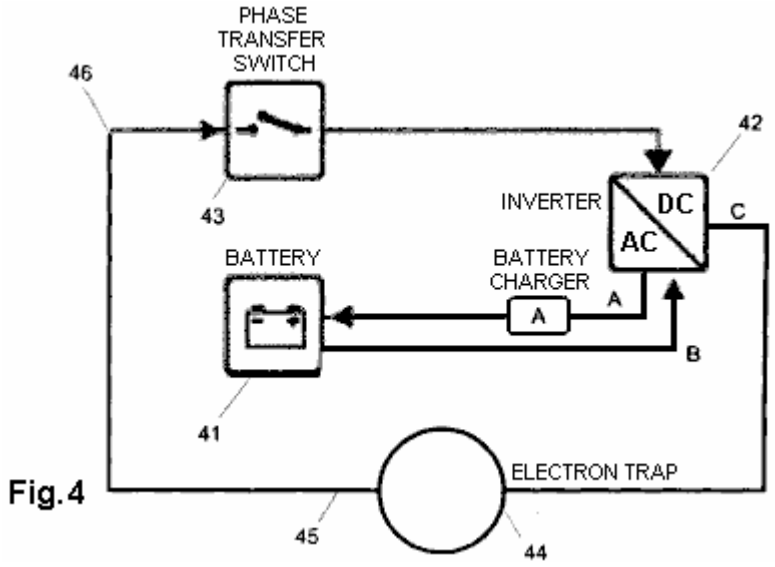


Fig. 4

Fig.4 shows another embodiment of the system for self-powered electricity generation which is comprised of a basic circuit of a typical uninterruptible power supply, consisting of a battery charger (AC/DC converter) **A** connected to an inverter (DC/AC) **42** and attached to them, battery bank **41**, and this basic circuit together with other devices are connected to a free space electron-capture device **44** or an earth-connection electron-trap **44**. Comprising thus, a battery charger **A** connected to a battery bank **41**, which is connected in series with inverter **42** at point **B** which is in series with point **C** of inverter **42** which is in series with the electron sensor **44**, which is in series with the phase transfer switch **43** via the three-phase load output connection point **45**. The phase transfer switch **43** is in series with the inverter **42**, which is connected in series the (AC/DC converter) battery charger **A** feeding the battery bank **41**.

An alternative construction for use in emergency situations, in which the system ceases to be self-powered, the system may include power input from an external power source, via the interconnection point **46**, thus providing electricity output **45**, to power external loads. The battery bank **41** provides power to the inverter **42** which converts the direct current into alternating current and feeds the electron trap **44**. The phase transfer switch closes when the batteries need recharging.

Sensor **44** captures electrons, producing alternating current, which feeds the phase transfer switch **43** with alternating current input power. The phase transfer switch **43** feeds the inverter **42** which charges the batteries, closing the self-powering loop which provides power at the output **45**, feeding both the power input and any external loads.

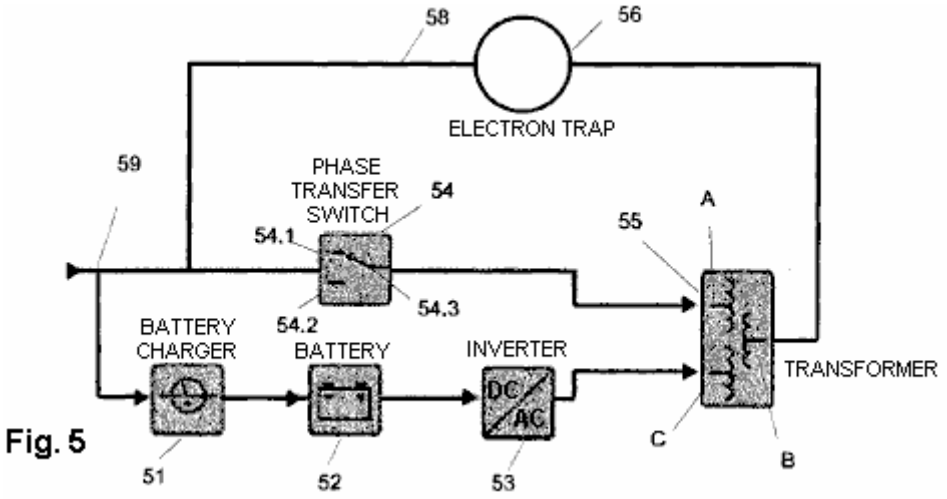


Fig. 5

Fig.5 shows another embodiment of the system for self-powered electric power generation equipment comprising a circuit which includes a typical uninterruptible power supply comprising a battery charger (AC/DC converter) **51** connected to a DC/AC inverter **53** and attached to them, a battery bank **52**. This basic circuit together with other devices are connected to a space free-electron capture device **56** (as defined in Brazilian patent application No. BR1020120008378 of 13/1/12) or, alternatively, an earthed free-electron collector **56** (as defined in Brazilian patent application No. BR1020120008386 of 13/1/12). This then comprises a battery charger **51** which is connected in series with a battery bank **52**, which is connected in series with the inverter **53**, which is connected in series with the transformer **55** at its point **C**, which is in series with its point **B** which is in series with the electron collector **56**, which is in series with the battery charger **51** which is connected to the load exit point **58**, which is also the circuit entry point **59**, which is in series with the phase transfer switch **54** section **54.1**, which is connected to terminal **54.3**, which is in series with point **A** of the transformer **55** which exits at point **B**. Points **A** and **54.3** as well as the parallel points **54.1** and **54.2**, are all parallel to the battery charger **51**, the battery bank **52**, the inverter **53** and to point **C** of the transformer **55**.

An alternative construction for use in emergency situations, in which the system ceases to be self-powered, the system may include an external power input point **59**, allowing phase transfer switch **54** to provide power output **58**, to feed external loads. Battery bank **52** provides power to the inverter **53**, which converts the direct current into alternating current, feeding point **C** of the transformer, which comes out at points **B** and **A** of the transformer **55**. Point **B** of the transformer feeds the electron-trap **56** producing alternating current which feeds the battery charger **51**, recharging the battery bank **52**.

The battery charger **51** is connected in parallel with the transfer switch **54** via connection points **54.1** and **54.3**, feeding point **A** of the transformer, which comes out at point **B**. Point **A** of the transformer and the switch transfer points **54.3** and **54.1** are in parallel to the battery charger **51**, the battery **52**, the inverter **53** and point **C** of the transformer **55**.

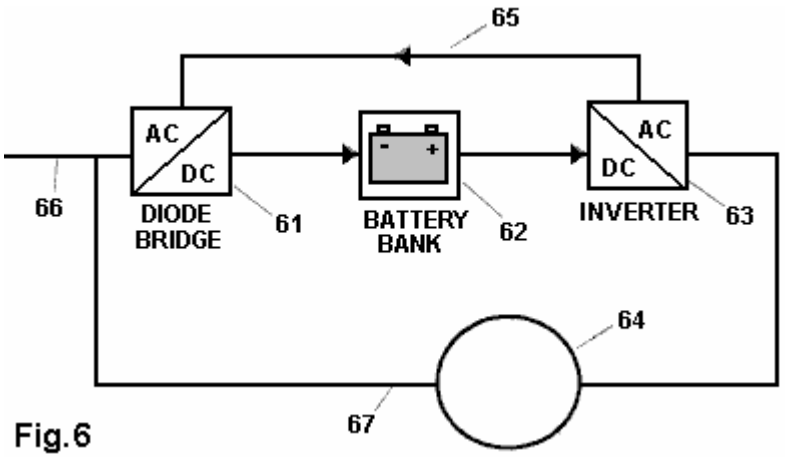


Fig.6

Fig.6 shows another embodiment where a rectifier **61** is connected to an inverter **63** and a battery bank **62**, and to a space free-electron trap **64** or alternatively, an earth electron trap **64** comprising thus, a delta (AC/DC) converter **61**, which is connected in series to a battery bank **62**, which is connected in series with the (DC/AC)

inverter **63**, which is in series with the electron collector **64** which is connected in series with the delta converter (AC/DC) **61** whose AC part is in series with the alternating AC current inverter **63** via a connecting wire **65** which is in parallel with the DC part of the delta converter **61** with the battery bank **62** and the DC part of inverter **63**. An alternative construction for use in emergency situations, in which the system ceases to be self-powered, the system may comprise a power input from an external power source, via the interconnection point **66** connected to the delta converter **61**, the output **67** supplying power, to the external loads.

Battery bank **62** provides power to the inverter **63**, which converts the direct current into alternating current, powering the free-electron collector **64**. The captured electrons from collector **64** form an alternating current which feeds the delta converter **61** via an output power load wire **67**.

The alternating part of the three-phase delta converter **61** is fed with alternating current from inverter **63** via connecting wire **65**, which is connected in parallel to the continuous DC delta converter **61**, which feeds the battery bank **62** and with the continuous portion the inverter **63**, closing the cycle of self-feeding and supplying power at the output **67**, which is the output power point.

Having described examples of preferred embodiments, it should be understood that the scope of the present invention encompasses other possible forms of construction, using the electron collectors connected to a basic circuit of a typical uninterruptible power supply of energy, known as a UPS, comprising a rectifier device (an AC/DC converter) **10**, connected to one inverter (DC/AC converter) **12**, and attached between them, an energy storage device (typically, a battery bank).

A very important part of the above patent is the device described as a “collector of free-electrons”, either from the earth or from space. We have to go to the patent applications mentioned above to find the details of these designs:

Application Number: BR2013/000015, Publication Date: 07/18/2013, Filing Date: 01/11/2013

Assignee: EVOLUÇÖES ENERGIA LTDA (Rua Santa Tereza 1427-B Centro - Imperatriz, MA- CEP -470 - Maranhäo, 65900, BR)

ELECTROMAGNETIC ELECTRON TRAP FOR ELECTRIC POWER GENERATION

Technical Field

The present invention refers to electromagnetic equipment for electric power generation or alternatively for thermal power generation. More specifically equipment capable of producing abundant electricity and thermal energy from a tiny amount of input electrical energy

Description of the Related Art

According to Lenz's law, any induced current has a direction such that the magnetic field it generates opposes the change in magnetic flux which produced it. Mathematically, Lenz's Law is expressed by the negative sign (-) that appears in the formula of Faraday's Law, as follows.

The magnitude of the induced emf (ε) in a conducting loop is equal to the rate of change of magnetic flux (Φ_B) with time:

$$\varepsilon = -\frac{d\Phi_B}{dt} \qquad \text{Equation 1}$$

As an example of application of Faraday's Law, we can calculate the electromotive force induced in a rectangular loop that moves in or out, with constant speed, a region of uniform magnetic field. The magnetic field flux through the surface limited by the loop is given by:

$$\phi = xLB \qquad \text{Equation 2}$$

and its variation in time:

$$\frac{\Delta\phi}{\Delta t} = \left(\frac{\Delta x}{\Delta t}\right)LB = vLB \qquad \text{Equation 3}$$

So:

$$\varepsilon = vLB \qquad \text{Equation 4}$$

and if the coil has a resistance (R) and the induced current:

$$i = \frac{\varepsilon}{R} = \frac{vLB}{R} \qquad \text{Equation 5}$$

A conductor traversed by an electric current immersed in a magnetic field undergoes the action of a force given by:

$$F = IL \times B \qquad \text{Equation 6}$$

Thus, the effect of the current induced in the loop appears as forces Ff, and F - FM. The first two cancel each other out and the third is cancelled by an external force P_{EXT} needed to maintain the constant speed loop.

As the force FM must oppose the force F_{EXT}, current (i) induced in the loop by varying the magnetic flux must have the meaning indicated in **Fig.3**. This fact is a particular example of Lenz's Law.

Considering the experimental activities discussed with Faraday's law, when a magnet approaches a coil, the induced current in the coil has a direction as shown in **Fig.1**. This generates a magnetic field whose north pole is facing the north pole of the magnet, that is, the field generated by the induced current opposes the motion of the magnet.

When the magnet is moved away from the coil, the current induced in the coil has a direction opposite to that shown in **Fig.1**, thereby generating a magnetic field whose south pole is facing the north pole of the magnet. The two poles attract each other, that is, the field generated by the induced current opposes the movement of the

magnet away from the coil. This behaviour is present in all current power generators, and known as ‘engine brake’ is highly undesirable as it increases the resistance and so, the energy loss.

When two electromagnetic coils are placed facing each other, as shown in **Fig.2**, there is no current in either of them. At the instant of power-up of one of the coils, the current in the coil, generates an induced current in the second coil. When powered up, the current in the coil goes from zero to its maximum value, and then remains constant.

Thus, when the current is changing, the magnetic field generated by it, (whose north pole faces the second coil) is also changing and so the magnetic flux of this field through the second coil is also changing. Then there is a current induced in the second coil whose sense is such that the magnetic field it generates tends to decrease the flow mentioned above, that is, its north pole confronts the north pole of the first field coil.

When the power switch is opened, the current in the first coil drops from its maximum value to zero, and correspondingly its magnetic field decreases. The flux of the magnetic field in the second coil also decreases, and the induced current now flows in the opposite direction. This current flow direction produces an enhancing magnetic field, that is, it has a south pole facing the north pole of the field of the first coil.

Thus, there is a realisation of the principle of conservation of energy, expressed by Lenz's law, wherein any induced current has an effect which opposes the cause that produced it. Assuming that the induced current acts to favour the variation of the magnetic flux that produced the magnetic field of the coil, it would have a south pole facing the north pole of the approaching magnet, causing the magnet to be attracted towards the coil.

If the magnet were then released, it would experience an acceleration toward the coil, increasing the intensity of the induced current and thus create an enhanced magnetic field. This field, in turn, would attract the magnet with increasing force, and so on, with a continuing increase in the kinetic energy of the magnet.

If energy were to be withdrawn from the magnet-coil system at the same rate at which the kinetic energy of the magnet increases, then there would be an endless supply of energy. So it would be a perpetually operating motor, which would violate the principle of conservation of energy. Therefore, it can be concluded that current generators feature a large energy loss during the generation of electricity.

Objectives of the Invention

An objective of the present invention is to contribute to the generation of sustainable energy, proposing an electromagnetic machine capable of producing abundant electricity from an extremely low input of electrical energy.

The above objective and other objectives are achieved by the present invention by a device comprised of at least one electromagnetic field-generating device (without a core or with at least one core) powered by an electrical power source (without a core or with at least one core) having their coils, or sets of coils, wound on at least one common conductive member in a closed circuit which itself has a polarised voltage which is connected to at least one conductive interconnection element which is connected to a grounding grid, these interconnections creating a new technical effect, namely, the appearance of an electric current which keeps circulating in a closed conductive loop, and which can therefore be used to power external loads.

The device which is the object of the present invention operates as follows: the electromagnetic field generating device, powered by a power source, produces an electromagnetic field which induces an electric current in a closed conductive circuit, creating an interaction between the magnetic poles of the equipment and the magnetic poles of the earth - through both electromagnetic attraction and repulsion. An endless supply of electrons is drawn from the earth into the conductive closed loop, which is connected to the ground through a conductive interconnected grid. Attracted electrons add to the current already flowing in the conductive closed loop, making power available for driving high-power loads, although the device itself is only supplied with a small amount of power. Thus, advantageously, the device which is the object of the present invention, acts as a trap for electrons from the earth and this allows the generation of electricity.

Advantageously, the present electromagnetic equipment generates either electricity or thermal energy, providing access to this new source of energy is through an electromagnetic field. The interconnections of the components of the electron-trap of the present invention, cause an advantageous new technical effect, namely, the appearance of an electric current which keeps circling in the conductive closed circuit, with or without voltage being applied and even without a load being connected to the loop - provided that the electron-trap is connected.

The proposed sensor can also be used to generate thermal power, depending on the form in which you want to use the effect of the flow of electrical current produced in this electromagnetic equipment.

For the generation of thermal energy in amounts proportional to the power of the electron-trap, through the movement of electrons in the conductive closed loop itself, the resistance should be increased by increasing the number of turns around the cores in the conductive element of the closed circuit, and in that instance, the coils of the electromagnetic field generating device, will then be made of heat-insulated electrical circuit components, bearing in mind the required temperature which is to be produced. The thermal energy generated by the electron-trap can be used in any application from domestic to industrial applications.

This technology can also be used for various technical purposes in electric machines. By "electrical machines", it should be understood to include: static electrical machines, transformers, ballasts, rotating electrical machines, synchronous machines, dual power supply machines, current rectifiers in synchronous cascade, external pole machines, synchronous current machines alternating current machines and/or direct current machines, electronic equipment and electrical resistances. The capture of electrons can provide single-phase, two-phase or three-phase supplies, operating at low, medium or high voltage.

The capture of electrons by induction, does not impact on the environment. The fact is that we use as the capturing force, only a negligible amount of electricity relative to the current captured by the sensor. The relationship between power input and the quantity of electricity generated by the electron-trap is at least 1 to 100, that is, for each 1 watt provided to the sensor, there is at least 100 watts of power available for external loads. This relationship, however, is not limited, as it depends on the mounting of the electron-trap and the objectives of the circuit, and so, the generated power can be greater than 100 times the input power.

Another advantage of the earthed electron- trap proposed in the present invention is that the electron-trap can transport electrons from point "A" to point "B" without a voltage drop across the closed-loop conductive element - if it is biased with a voltage - regardless of the distance between the points depending on the strength and quantity of the electromagnetic field generating devices. It is also possible to transport electrons when the conductive element in a closed circuit is itself not polarised. Thus, the electric current is transported without voltage, just by the magnetic field formed between the device and the generator of the electromagnetic field.

Brief description of the Drawings

The present invention will now be described with the aid of drawings, but the design is not limited to the implementations shown in these drawings, although they show other details and advantages of the present invention.

The figures show:

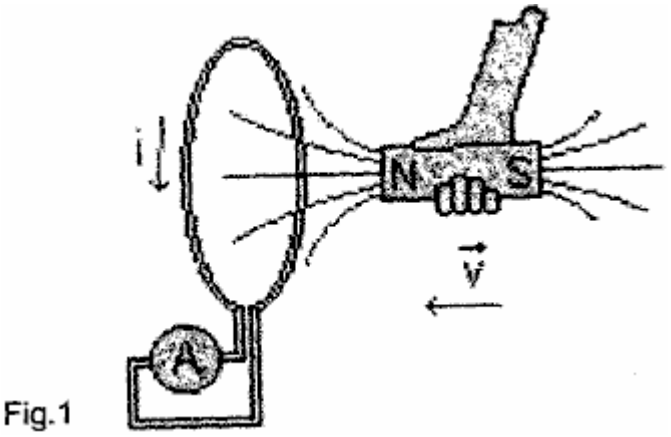


Fig.1

Fig.1 - illustrates Faraday's law.

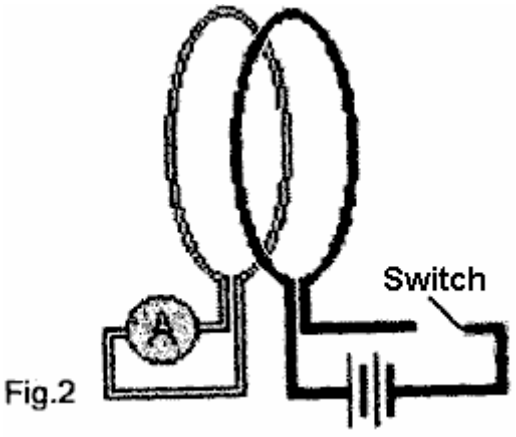


Fig.2 – is a representation of Faraday's law.

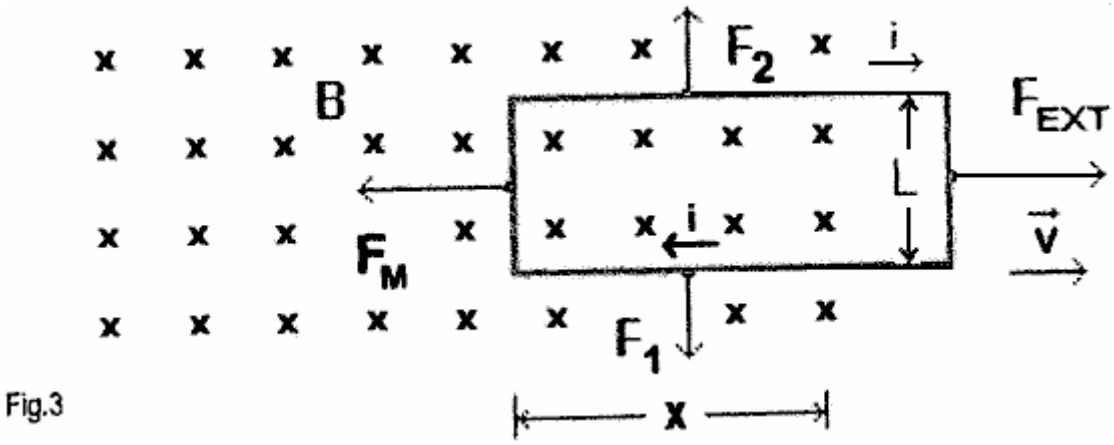


Fig. 3 – is a representation of Faraday's law.

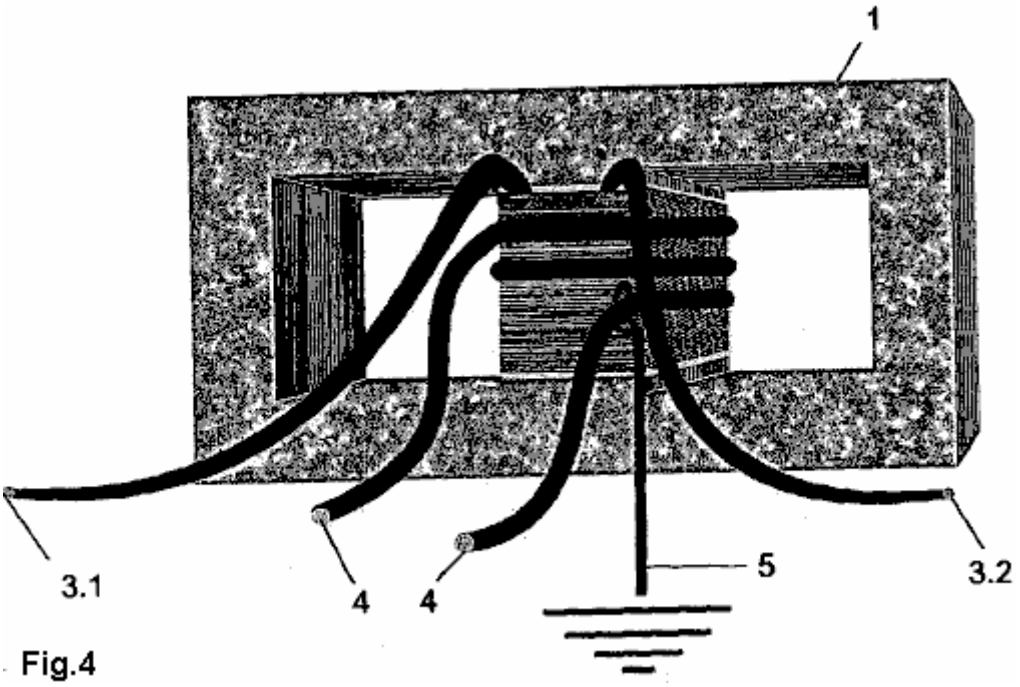


Fig. 4 - is a perspective view of an electron-trap with a single phase coil.

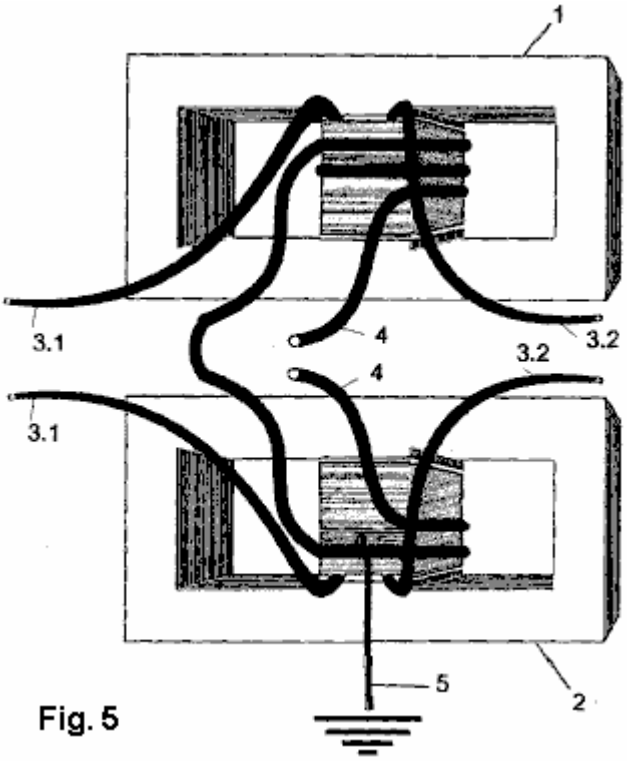


Fig. 5

Fig.5 – is a perspective view of a single-phase electron trap with two coils.

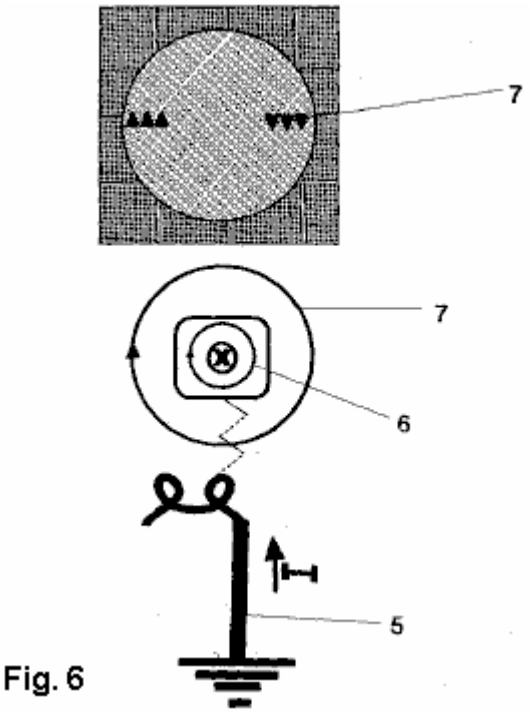


Fig. 6

Fig.6 - is a representation of the effect of electromagnetic flux in the coils around the cores of the electron trap.

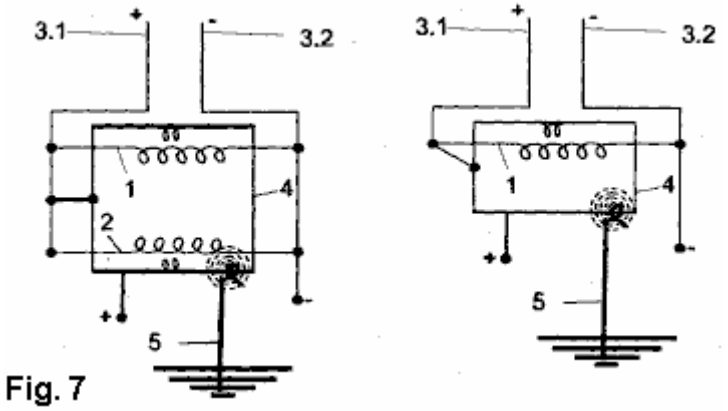


Fig.7 - is a representation of an electrical circuit with two coils of the link/coil conductor polarised.

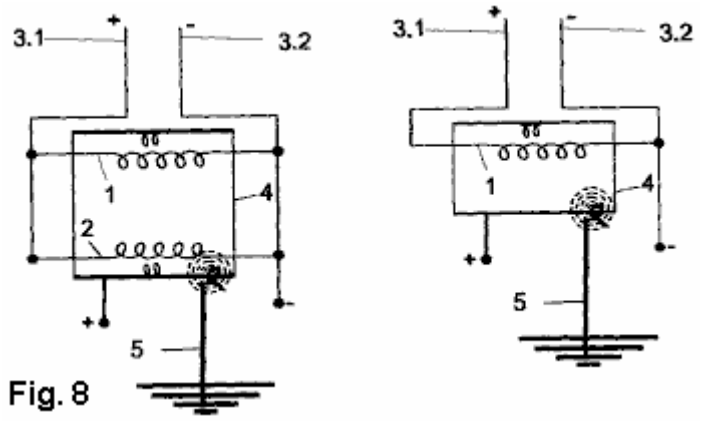


Fig.8 - is a representation of an electrical circuit with two coils of the link/coil conductor not polarised.

Detailed Description of the Drawings

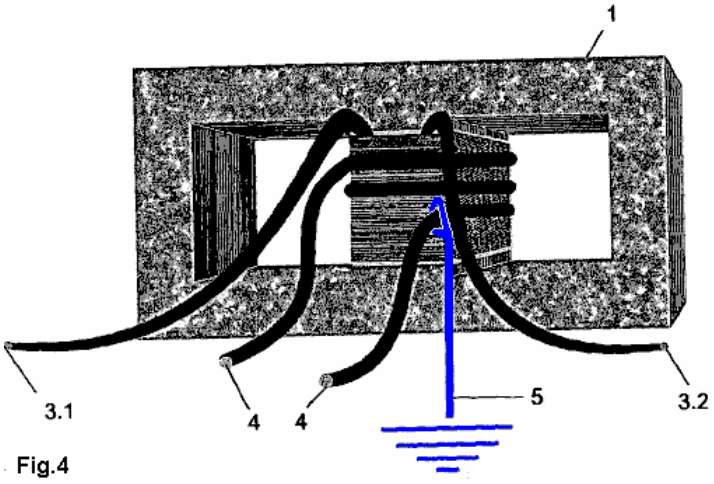


Fig.4 shows one of several types of electron-trap proposed by the present invention, where the electron-trap is single-phase and consists of at least one electromagnetic field-generating device with at least one set of coils, in this case it happens to be an electromagnetic type coil with one common magnetic core, but it could alternatively have any number of windings of any kind and shape. However, the electron-trap proposed by the present invention can be constructed with a different type of electromagnetic field generating device, such as an electromagnetic inductor or magnet of any type or shape, or any combination of them, and in unlimited numbers for each phase of the electron trap.

When winding these coils, for example, coil **4-4**, each coil must have at least one complete turn, preferably two turns if the objective is to generate electricity, and preferably four turns if the objective is provide thermal energy. The number of turns in the coils wound around the common core, is directly related to the amount of current to be generated.

At least one conductive interconnection element, in this case the driving member **5** - which can be copper or any other suitable conductive, material whether insulated or not insulated, connects or loop-links wire **4** to the ground grid. The connection between the conductor **5** and wire **4** is by electromagnetic induction. Winding **4** is also the power supply for the loads which are to be powered by the captured electrons.

Also in **Fig.4**, the power wires **3.1** and **3.2** (live phase and neutral) have an input from an external power coil **1** which can be energised from any external source of electricity such as a power grid. The trapped electrons can be configured to supply DC or AC current. Thus, if the coil **1** power source is alternating electrical current - AC, then the electron-trap provides alternating electrical current. If the power source is continuous electrical current - DC, then the electron-trap provides continuous electrical current - DC. The electrical supply provided by the trapped electrons can be single-phase, two-phase or three-phase, and at low, medium or high voltage.

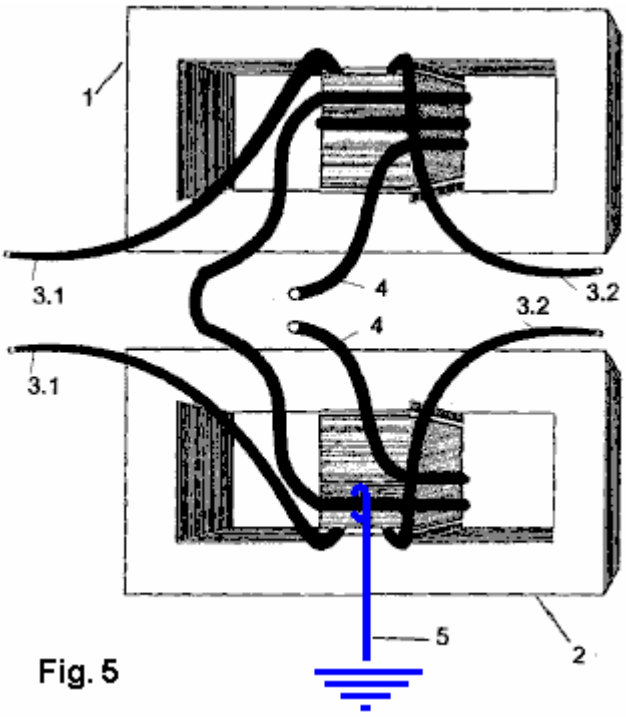


Fig. 5

Fig.5 shows an electron-trap with two single-core phase coils: **1** and **2**, although these coils may be of any type and shape. However, the electron-trap proposed by the present invention can be constructed with other types of electromagnetic field generating device, with at least one electromagnetic inductor or electromagnet which can be of any type and shape, with any combination of them, and in unlimited quantities in each phase of the electron-trap.

The coils on frames **1** and **2** may have other shapes, but they must each have at least one complete turn, particularly in coil **4**. The number of turns in this winding are directly related to the amount of current to be generated. This coil also makes the interconnection between the coils **1** and **2** forming the link between their two cores.

At least one conductive interconnection element, in this case the driving member **5** - which can be copper or any other suitable conductive, material whether insulated or not insulated, connects or loop-links wire **4** to the ground grid. The connection between the conductor **5** and wire **4** is by electromagnetic induction.

In electron-traps which have numerous sets of coils **1** and **2**, the ends of all of the power-supply conductors **3.1** can all be connected to each other, and all of the **3.2** conductor ends may be connected together. Thus, all of the coils **1** and **2** can be fed exactly the same voltage. The power to energise coils **1** and **2** can be provided from any external source of supply of electricity such as a power grid.

In electron-traps which have numerous coils **1** and **2**, a single coil winding **4** connects the cores of all of the coils **1** and **2**.

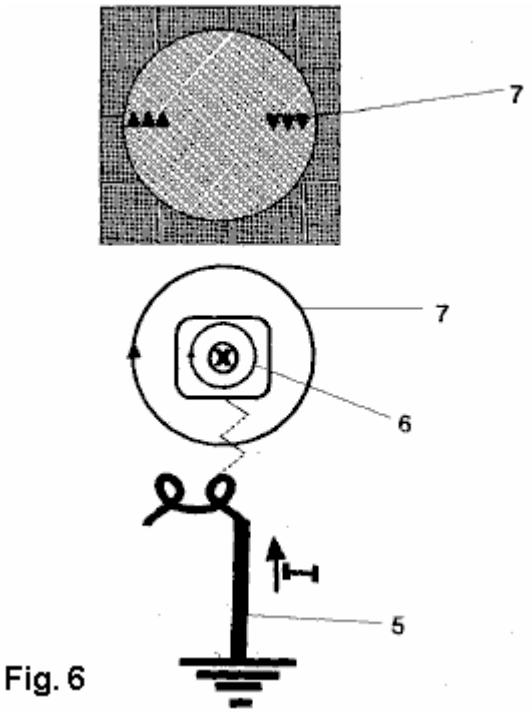


Fig. 6

The diagram shown in **Fig.6**, illustrates the magnetic induction **6** around the core "**X**" of the coil **1**. This induction causes electrical current flow in the conductor coil link **7/4**, attracting electrons from the earth, through the conductive member **5**, to the magnetic field of the electron-trap, where those electrons are added to the current generated by induction in the link coil **4** conductor loop circulating between north and south magnetic poles.

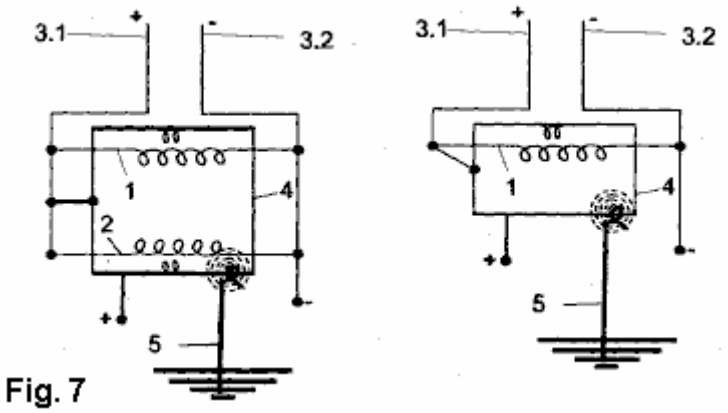


Fig. 7

Fig.7 shows how the connections should be made in one version of the electrical circuit of the electron-trap proposed in this invention. The diagram shows the electrical circuit of an electron-trap where the link/coil driver **4** is polarised with a voltage. This is one form of construction for an electron-trap which has two coils **1** and **2**, where a link/coil loop conductor **4** is biased with a voltage, that is, there is a link connecting the coil conductors **4** of a power supply **3.1** or **3.2**, whatever the stage.

In this way, earth electron-traps, by adopting this circuit, that is, with the link/conductor loop **4** and polarised voltage on coils **1** and **2**, besides being used as a power source for external loads, can also be used for thermal power generation.

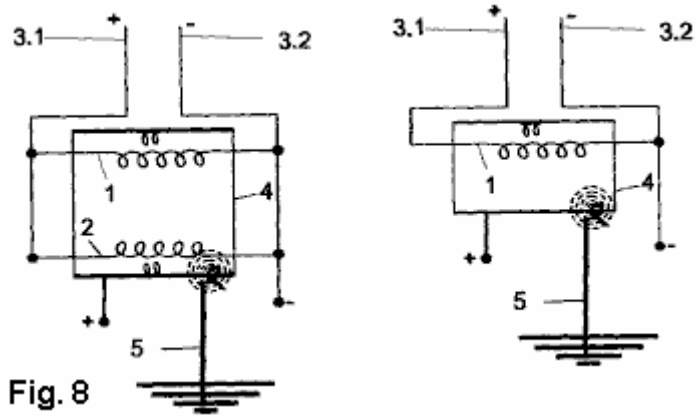


Fig. 8

Fig.8 shows how connections should be made in another electric circuit electron-trap proposed in this invention. the circuit illustrates a circuit of an electron-trap with a non-polarised link / coil driver **4**. This is one form of construction of the electron-trap where a link / coil conductor **4** of the spiral conductor coils **1** and **2** is not polarised, that is, there is no such link connecting conductor / conductor coil conductors **4** of a coil **3.1** or **3.2**.

Thus, earth electron-traps adopting this circuit, that is, with the link coil not polarised, the current flows without there being voltage in the link/coil conductor **4** joining the first and second coils by electromagnetic induction. They can also be used for generating thermal energy.

The structure of the circuit - in the open or closed coils **1** and **2**, and always in the closed link / loop lead **4** - makes it possible to generate current by induction and electron capture by electromagnetism on the link conductor **4** - where current is generated and stays in motion with or without voltage, as the coils **1** and **2** are being fed. Thus, the present invention provides a new concept for electrical energy generation, since it is obtained from an electric current circling without consumption and even without an output load being attached to it.

Additionally, because the induced electrical current flows regardless of the voltage present, it can be used as a current stabiliser for electrical networks whether they be single-phase, two-phase or three-phase, with low, medium or high voltage.

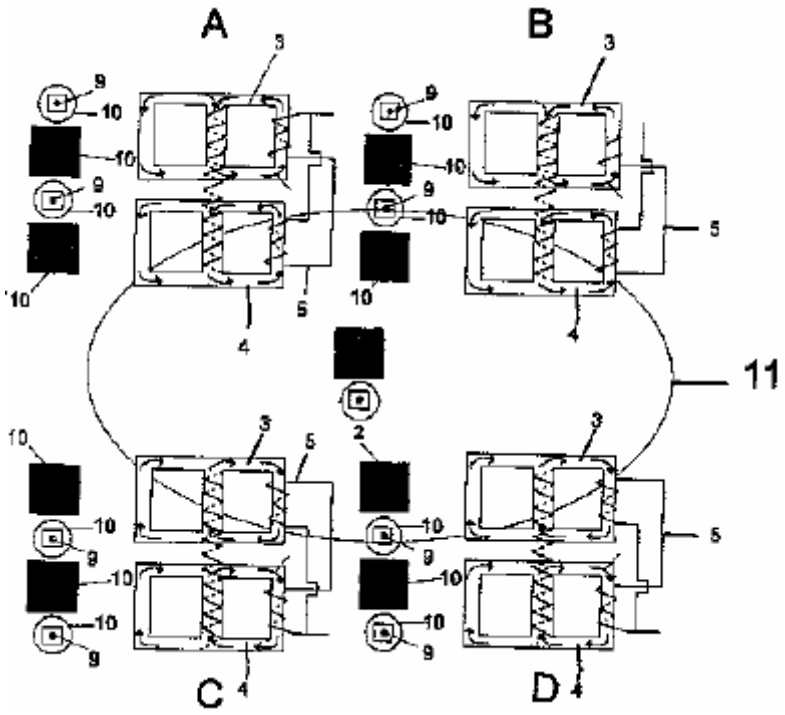
ELECTROMAGNETIC DEVICE FOR CAPTURING FREE SPACE ELECTRONS TO GENERATE ELECTRICITY

Application Number: BR2013/000014

Publication Date: 07/18/2013

Filing Date: 01/11/2013

Assignee: EVOLUÇÕES ENERGIA LTDA (Rua Santa Tereza 1427-B Centro - Imperatriz -, MA - CEP -470 - Maranhão, 65900, BR)



Abstract:

The invention relates to a device that comprises at least three sets (**A, B, C, D**) of at least one device for generating an electromagnetic field (**3**) and (**4**), powered by an electricity source (without a core or with at least one core) the cores thereof or any extension thereof, preferably the windings or sets of windings thereof, being surrounded by at least a single conductive element forming a polarised and energised closed-circuit (**5**), the sets of electromagnetic-field generating devices (**3**) and (**4**) being linked together by their opposing poles to encourage the interaction of their electromagnetic fields, which ideally, are located between two hollow metal hemispheres (**1**) so as to concentrate and enhance the electromagnetic fields, these interconnections causing, as a novel technical effect, the emergence of an electrical current that circulates, with or without voltage, in the conductive element forming a closed-circuit (**5**) - even if no load is connected.

Description:

"ELECTROMAGNETIC EQUIPMENT FOR FREE ELECTRON-CAPTURE FROM SPACE, FOR ELECTRICITY GENERATION".

Technical Field

The present invention relates to electromagnetic equipment for electrical power generation and/or thermal power generation. More specifically, equipment capable of producing abundant electricity and thermal energy from a tiny input of electrical energy.

Description of the Related Art

According to Lenz's law, any induced current has a direction such that the magnetic field it generates opposes the change in magnetic flux that produced it. Mathematically, Lenz's Law is expressed by the negative sign (-) that appears in the formula of Faraday's Law, as follows.

The magnitude of the induced emf (ε) in a conducting loop is equal to the rate of change of magnetic flux (Φ_B) with time:

$$\varepsilon = -\frac{d\Phi_B}{dt} \qquad \text{Equation 1}$$

As an example of application of Faraday's Law, we can calculate the electromotive force induced in a rectangular loop that moves in or out, with constant speed, a region of uniform magnetic field. The magnetic field flux through the surface limited by the loop is given by:

$$\phi = xLB \qquad \text{Equation 2}$$

and its variation in time:

$$\frac{\Delta\phi}{\Delta t} = \left(\frac{\Delta x}{\Delta t}\right)LB = vLB \qquad \text{Equation 3}$$

So:

$$\varepsilon = vLB \qquad \text{Equation 4}$$

and if the coil has a resistance (R) and the induced current:

$$i = \frac{\varepsilon}{R} = \frac{vLB}{R} \qquad \text{Equation 5}$$

A conductor traversed by an electric current immersed in a magnetic field undergoes the action of a force given by:

$$F = I\vec{L} \times B \qquad \text{Equation 6}$$

Thus, the effect of the current induced in the loop appears as forces Ff, and F - FM. The first two cancel each other out and the third is cancelled by an external force P_{EXT} needed to maintain the constant speed loop.

As the force FM must oppose the force F_{EXT}, current (i) induced in the loop by varying the magnetic flux flux must have the meaning indicated in **Fig.1**. This fact is a particular example of Lenz's Law.

Considering the experimental activities discussed with Faraday's law, when a magnet approaches a coil, the induced current in the coil has a direction as shown in **Fig.2**. This generates a magnetic field whose north pole is facing the north pole of the magnet, that is, the field generated by the induced current opposes the motion of the magnet.

When the magnet is moved away from the coil, the current induced in the coil has a direction opposite to that shown in **Fig.2**, thereby generating a magnetic field whose south pole is facing the north pole of the magnet. The two poles attract each other, that is, the field generated by the induced current opposes the movement of the magnet away from the coil. This behaviour is present in all current power generators, and known as ‘engine brake’ is highly undesirable as it increases the resistance and so, the energy loss.

When two electromagnetic coils are placed facing each other, there is no current in either of them. At the instant of power up one of the coils, the current in the coil, generates an induced current in the second coil. When powered up, the current in the coil goes from zero to its maximum value, and then remains constant.

Thus, when the current is changing, the magnetic field generated by it, (whose north pole faces the second coil) is also changing and so the magnetic flux of this field through the second coil is also changing. Then there is a current induced in the second coil whose sense is such that the magnetic field it generates tends to decrease the flow mentioned above, that is, its north pole confronts the north pole of the first field coil.

When the power switch is opened, the current in the first coil drops from its maximum value to zero, and correspondingly its magnetic field decreases. The flux of the magnetic field in the second coil also decreases,

and the induced current now flows in the opposite direction. This current flow direction produces an enhancing magnetic field, that is, it has a south pole facing the north pole of the field of the first coil.

Thus, there is a realisation of the principle of conservation of energy, expressed by Lenz's law, wherein any induced current has an effect which opposes the cause that produced it. Assuming that the induced current acts to favour the variation of the magnetic flux that produced the magnetic field of the coil, it would have a south pole facing the north pole of the approaching magnet, causing the magnet to be attracted towards the coil.

If the magnet were then released, it would experience an acceleration toward the coil, increasing the intensity of the induced current and thus create an enhanced magnetic field. This field, in turn, would attract the magnet with increasing force, and so on, with a continuing increase in the kinetic energy of the magnet.

If energy were to be withdrawn from the magnet-coil system at the same rate at which the kinetic energy of the magnet increases, then there would be an endless supply of energy. So it would be a perpetually operating motor, which would violate the principle of conservation of energy. Therefore, it can be concluded that current generators feature a large energy loss during the generation of electricity.

Objectives of the Invention

The present invention aims to contribute to the generation of sustainable energy, proposing electromagnetic equipment capable of producing abundant electricity from an extremely low input of electrical energy.

The above objective and other objectives are achieved in the present invention by a device comprising at least three sets of at least one electromagnetic field generating device (without a core or with at least one core) powered by an electrical power source, having their cores or any extension of them, with their coils or sets of coils, wound on at least one common conductive member in a closed circuit which is polarised by a voltage source, and these sets of electromagnetic field generating devices are arranged with their poles in confrontation, to promote the interaction of electromagnetic fields, and, preferably, positioned between two hollow metallic hemispheres, in order to focus and enhance their electromagnetic fields - these interactions cause a new technical effect - the emergence of an electric current which keeps flowing in a closed loop, with or without voltage being applied to that closed loop, current which is capable of powering external loads - even if no load is attached to it.

The device which is the object of the present invention operates as follows: Sets of electromagnetic field generating devices to be powered by an electrical power source, produce an electromagnetic field which induces an electric current in a closed conductive circuit, creating an interaction between the magnetic poles, and through repeated electromagnetic attraction and repulsion, provides an endless supply of electrons to the conductive closed loop itself.

The electrons attracted by this technique, augment the current flowing in the closed conductive loop, which provides the current to power external loads of high power, in spite of the fact that the device itself is supplied with only a small level of power. Thus, advantageously, the device which is disclosed in the present invention forms a trap for electrons from space, resulting in the generation of electricity. The interconnections of the components of the electron-trap cause, a new technical effect, namely, the appearance of an electric current which keeps circling in a closed circuit, even without any voltage being applied to the closed circuit, and even without a load being connected to it. The present electromagnetic equipment generates electricity or thermal energy, providing access to this new source of energy through the use of an electromagnetic field.

The proposed sensor can also be used for the generation of thermal energy depending on the form of circuit which is to be used, resulting from the flow of electric current produced by this electromagnetic equipment.

This field generates a flow of electric current induced by electromagnetic coils, which appears in the linking interconnecting devices generating electromagnetic fields with electromagnets, inductors or magnets. This chain operates in a manner favourable to the variation of the magnetic flux produced by the magnetic field in the electron-trap. Thus, it creates a north pole and a south pole, providing an endless supply of electric current without resistance between the links which interconnect the devices which are generating the electromagnetic fields. So, induced electric current is generated with or without voltage in the interconnection links of electromagnetic field-generating devices, depending on the connection method of the electrical circuit of the electron-trap.

The free-electrons collected by the space electron-trap can form alternating current (AC) or direct current (DC). The ratio of input power to output power is 1 to 100, that is, the generated power can be 100 times greater than the input power when there is at least one link / coil driver between the coils and the inductors or electromagnets. This relationship, however, is not limited to a factor of 100, as it depends on the shape of the electron-trap and its objective.

Another advantage of the free space electron-trap of the present invention is that, with thermal insulation of the components in the electric circuit, it is possible to produce thermal energy at low, medium or high temperature, through the movement of the electrons in the conductors, coils and/or electromagnets. The temperature generated is linked directly to the number of turns in the coils.

Thermal power generation performed by the sensor can be used for boiling and/or evaporation of liquids to be used in other types of energy generation, for example, replacing the use of coal and natural gas.

Another advantage of the proposed electron-trap of the present invention is that the electron-trap can transport electrons from one point "A" to a point "B", without a voltage drop in the link - if it is polarised - regardless of the distance between the points, depending on the strength and quantity of the electromagnetic field-generating devices. It is also possible to transport the electrons when the link devices generating the electromagnetic field are not polarised. In this way, the electric current is conveyed without voltage but only by the magnetic field formed between the coils. This methodology can be used in various fields.

Because of its simple construction, the electron-trap is a simple device which is compact, and performs low-cost power generation which can be used in all types of machinery, equipment and devices of all kinds, and many areas of application which require electricity in order to operate. The electron-trap can have single-phase, two-phase or three-phase output, and can generate electric current at low, medium or high voltage.

Brief description of the Drawings

The present invention will now be described with the aid of drawings, but the design is not limited to the implementations shown in these drawings, although they show other details and advantages of the present invention.

The figures show:

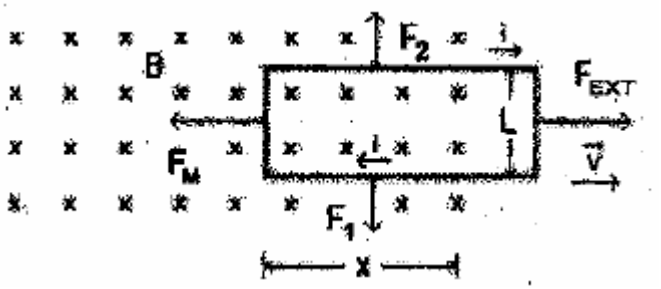


Fig.1 - illustrates Faraday's law.

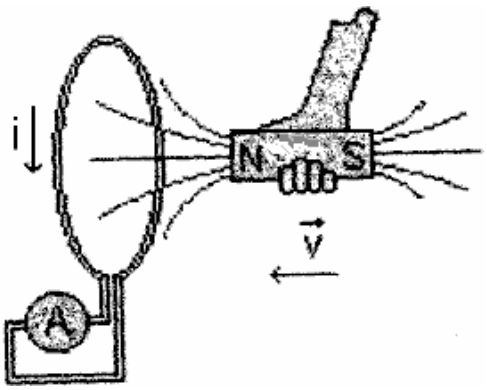


Fig.2 illustrates Faraday's law where a magnet approaches a coil of just one turn..

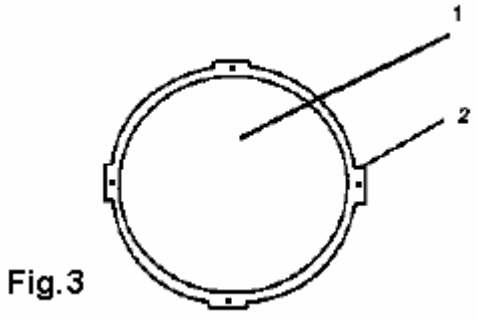


Fig.3

Fig.3 is a view of one metallic hemisphere seen from above.

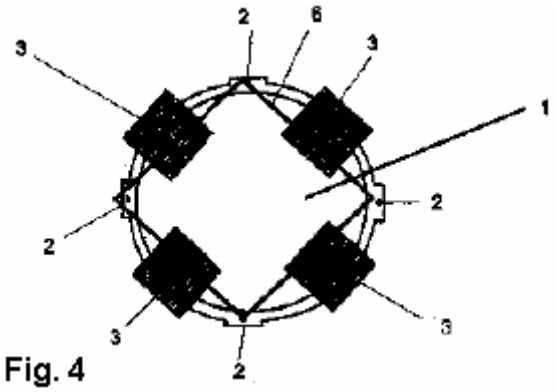


Fig. 4

Fig.4 is a bottom view of the hemisphere with the coils in place.

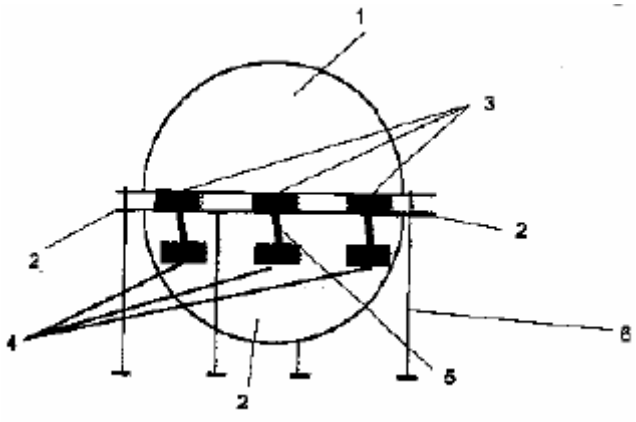


Fig.5 is a side view of the free-space electron-trap.

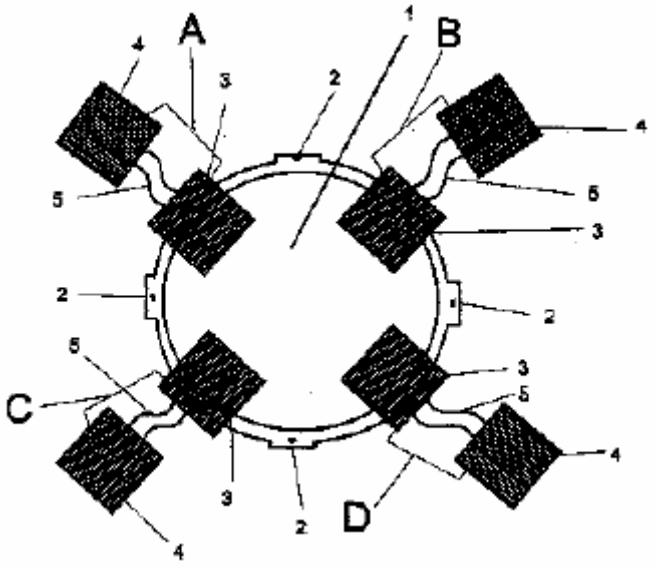


Fig.6 is an underside view of the space electron-trap, with its coils and electromagnets

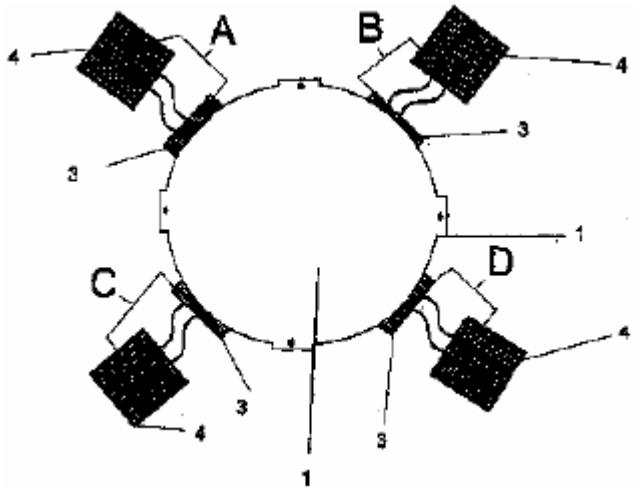


Fig.7 a view from above of the space electron-trap with its coils and electromagnets.

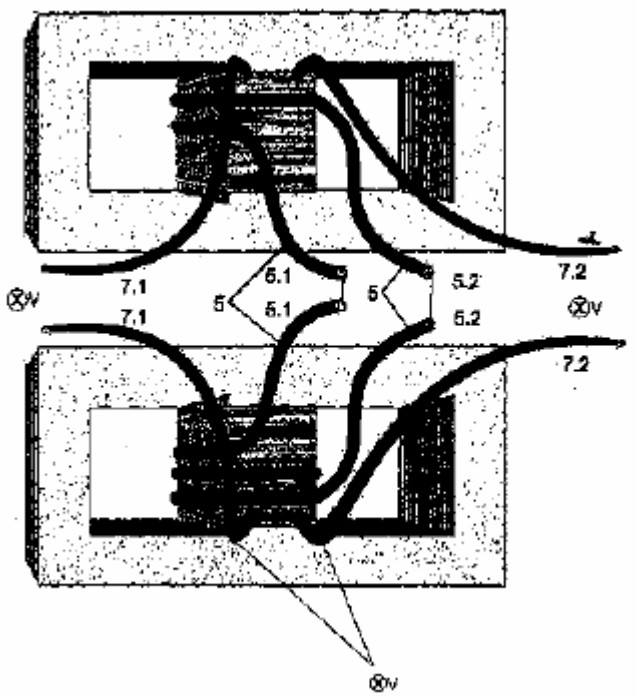


Fig.8 is a perspective view of an electron-trap with its coils.

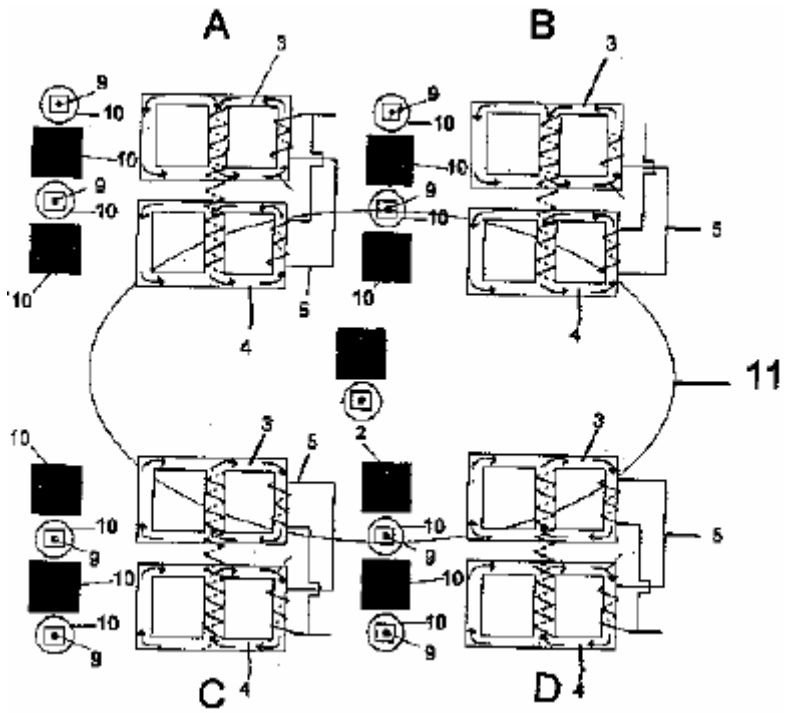


Fig.9 shows the circuit diagram of the device, indicating the effect of electromagnetic field.

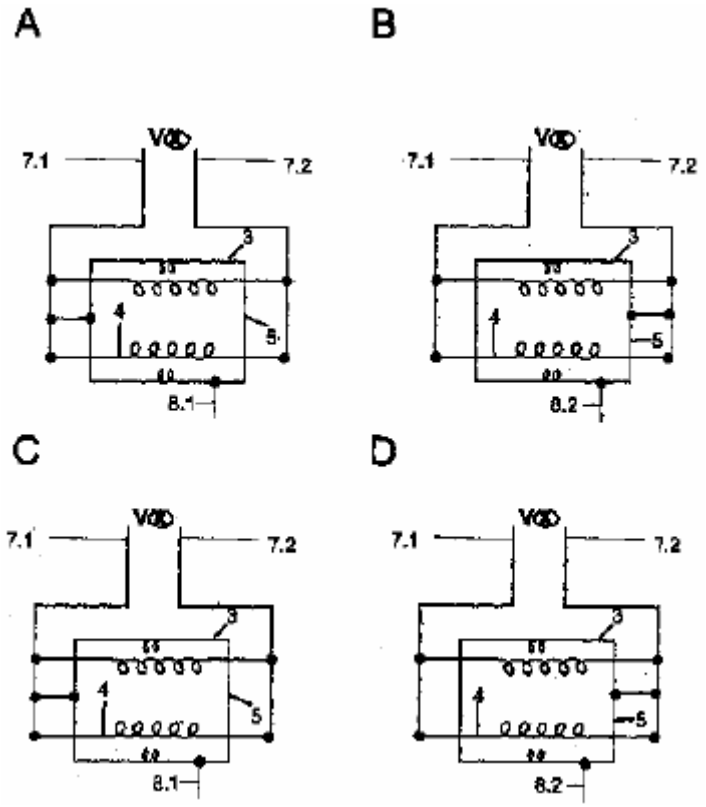


Fig.10 - shows the circuit diagram of the connection of the inductor coils in sets (**A, B, C** and **D**).

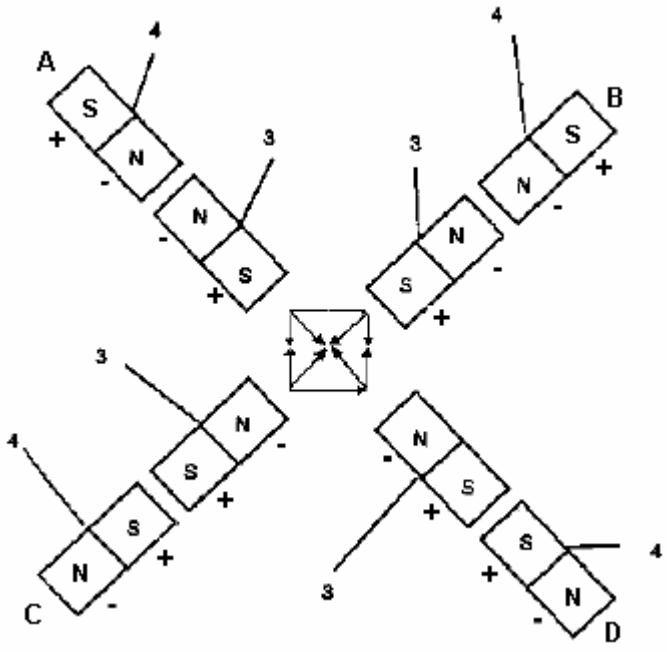


Fig.11 - is an electromagnetic diagram representation of north and south poles of the sets of coils (**A, B, C** and **D**).

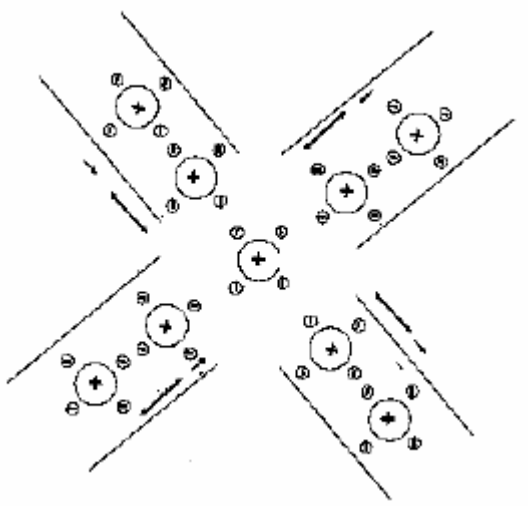


Fig.12 is a representation of the electrons being attracted and repelled by the device.

Detailed Description of the Drawings

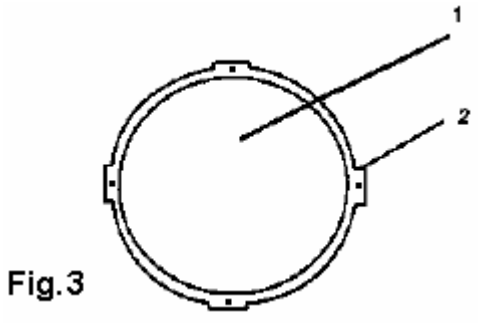


Fig.3 is a top view of one of the two hollow metallic hemispheres **1** which is part of the electron trap of free space proposed in this invention. Hemisphere **1** is preferably made from, but not limited to, aluminium, and it has mounting tabs **2**.

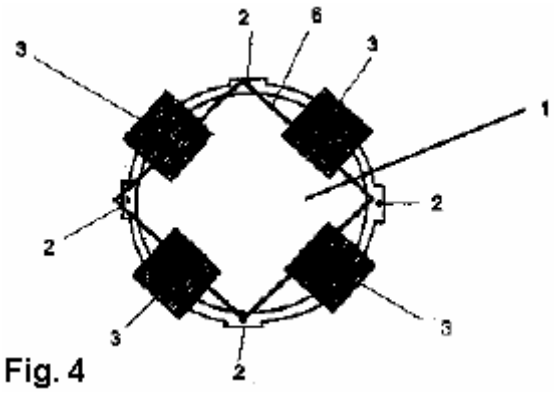


Fig.4 is a bottom view of metallic hemisphere **1**. It has four electromagnetic field generating devices **3**, positioned around the hemisphere and fixed to support **6** which is attached to hemisphere **1** by mounting tabs **2**.

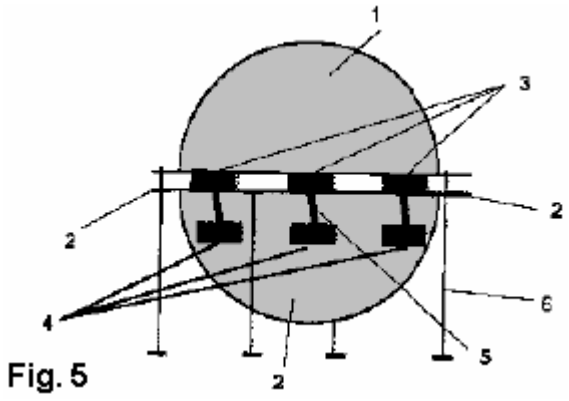


Fig. 5

Fig.5 is a side view of the free space electron-trap. It shows the two metallic hemispheres **1** and **2** (which form an imperfect sphere), and three of the coils **3** which are attached to the mounting tabs **2** and three inductors **4** which form the closed circuit itself, and which are attached by conductors **5**, and support member **6** on which are mounted coils **3** and their components.

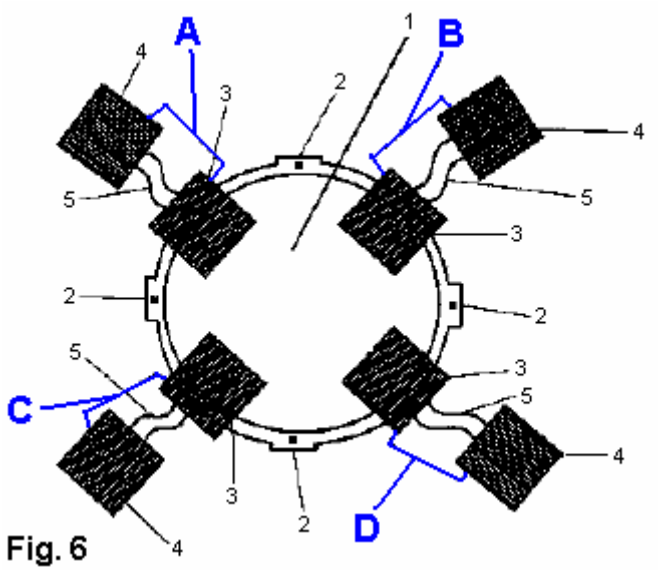


Fig. 6

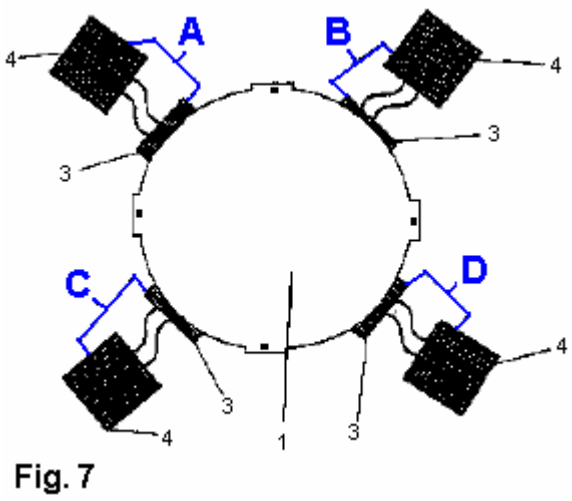


Fig. 7

Fig.6 and **Fig.7** show the top and bottom views of the metallic hemisphere **1** which accommodates four coils **3** attached to the holder **6** (not shown) which is secured to the hemisphere **1** by its mounting tabs **2**. **Fig.6** also shows the inductors or electromagnets **4** their corresponding coils **3** and their interconnecting conductors **5**. Each coil **3** and its linked inductor **4** forms a set. In **Figures 6** and **7** there are four such sets, marked **A, B, C** and **D**. The coils **3**, connected by their links **5**, each have at least one turn, and if the objective is to generate electricity, then preferably two turns, and if the objective is thermal energy, then four turns. The coils **3** may have various different shapes. The number of turns in the coil **3** are directly related to the amount of current to be generated, and the connecting links **5** may be either a single conductor or more than one conductor, the cross-sectional area of conductor **5** being selected to carry the current which is to be generated.

In sets **A, B, C** and **D**, the link conductors **5** have at least one turn around coils **3**. This winding is connected to the respective electromagnets **4** of each set (**A, B, C** and **D**) as shown in **Figures 6** and **7**. Please note that the inductors and electromagnets **4** can be any type of inductor, and other types of coil may be used.

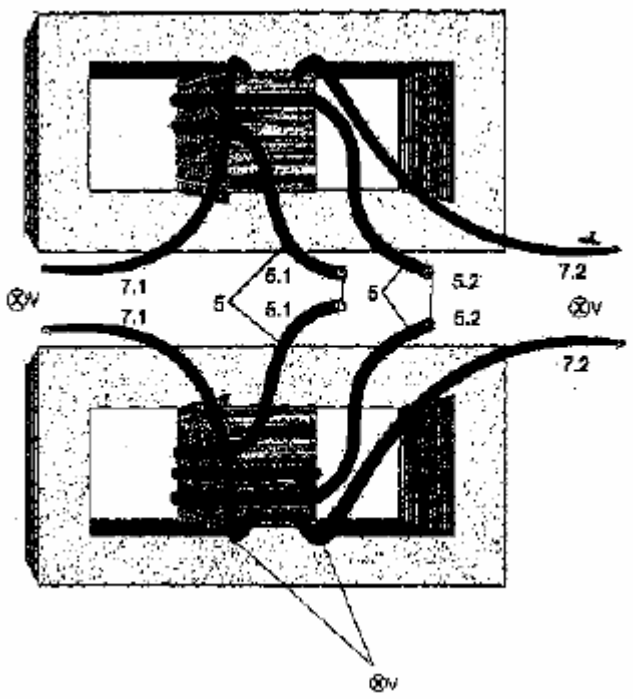


Fig.8 shows the inter-connecting coils **5** for each of the five sets **A, B, C** or **D** linking between coils **3** and **4** in each set. As shown in **Fig.6** and **Fig.7**, the link **5** makes the connection between coils **3** and **4**. This means that the wires marked **5.1** are all connected together, and the wires marked **5.2** are all connected together. Doing this, establishes the interconnection links **5** shown in the drawings. The power supply wires marked **7.1** are connected together as are the wires marked **7.2**. The wires marked **7.1** are connected to the live phase of the external power supply, while the other ends marked **7.2** are connected to the neutral of the external power supply.

In the space free-electron trap of the present invention, the coils **3** can be either single-phase, two-phase or three-phase. Also, the coils **3** may be powered by any voltage (V). The power coil **3** can be energised by any source of electrical energy such as a power grid. The electron-trap can be configured to produce alternating current or direct current. So, if the external power supply is alternating electrical current - AC, then the electron-trap provides an alternating electrical current output. If the power supply is DC, then the electron-trap provides an output of continuous electrical current - DC. The electron-trap can be configured for single-phase, two-phase or three-phase operation, with low, medium or high voltage outputs.

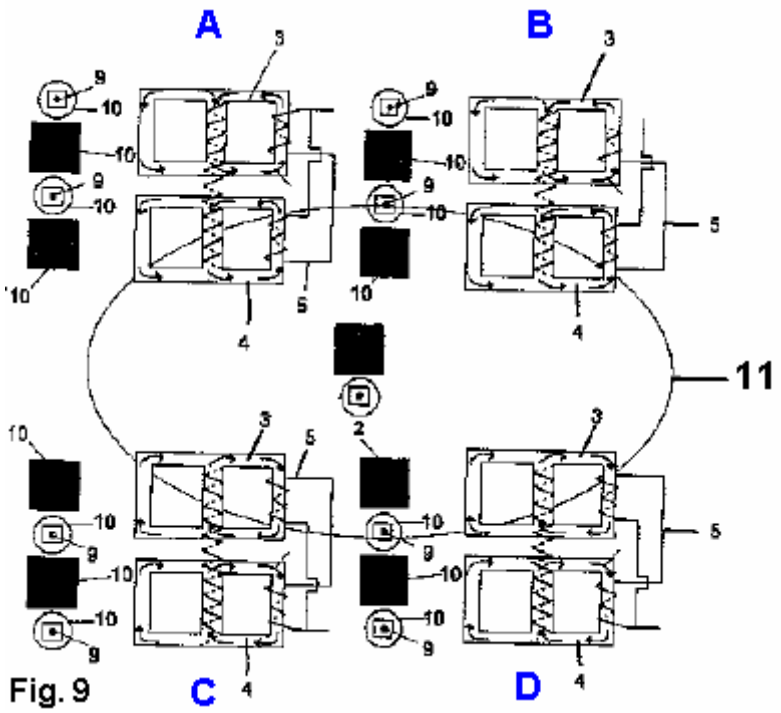


Fig.9 shows an electron-trap circuit diagram with four sets **A, B, C** and **D** of inductor coils **3** and **4**. Induction is produced around core **9** of the three sets of coils **A, B, C** and **D**. The effect of the interaction of the

electromagnetic fields **11** is shown. The induction via core **9**, causes the circulation of electric current in the links **5**, attracting the free electrons through the electromagnetic field of the trap. Then, the electrons join with the current generated by induction on link **5**, circulating between the magnetic poles north-south and south-north.

By way of example, the coils **3** are shown wound on a single phase column type core, but these can also be of any kind or shape. The electron-trap proposed by the present invention can be constructed with another type of electromagnetic field generating device which has at least one electromagnetic coil or magnet or electromagnetic inductor which can be of any kind or shape, or any combination of those, and with any number in each phase of the electron-trap.

The electron capture occurs through an electromagnetic field which is formed with the connection of coils **3** with the electromagnets or inductors **4** through the links **5** between the eight components.

This closure produces the displacement of the electrons in the coil **3** set (**A**) (for simplicity, referred to as coil **3A**), these electrons are attracted by the protons of coil **3D**, and are repelled by the electrons of the electromagnetic field of the coil **3D** itself. These coil **3D** electrons are attracted by the protons of the coil **3B**, and are repelled by the electrons of the electromagnetic field of coil **3B**. These electrons of coil **3B** are attracted by the protons of coil **3C**, and are repelled by the electrons of the electromagnetic field of the coil **3C** itself. Similarly, the **3C** coil electrons are attracted by protons of the **3A** coil, and are repelled by the electrons of the electromagnetic field of the coil **3A** itself. These coil **3A** electrons are attracted by the protons of the **3D** coil, and are repelled by the electrons of the electromagnetic field of the **3D** coil itself. Analogously, the coil **3D** electrons are attracted by the protons of the coil **3B**, and are repelled by the electrons of the electromagnetic field of the coil **3B** itself. These **3B** coil electrons are attracted by the protons of coil **3C**, and are repelled by the electrons of the coil itself induced **3C**, and then the coil **3C** electrons are attracted by protons of coil **3A**, and are repelled by the electrons of the electromagnetic field of the coil **3A** itself. That cycle continues as the sets of coils **A, B, C** and **D** are being fed by a voltage. These endless attractions and repulsions generate an electric current in the link coil **5**.

In the electron-trap, the voltage is stable. Regardless of the amount of current generated-which can be very high, the voltage will be the same in the electric circuit of the sensor, because the current moves through the attraction and repulsion of the electrons, regardless of voltage.

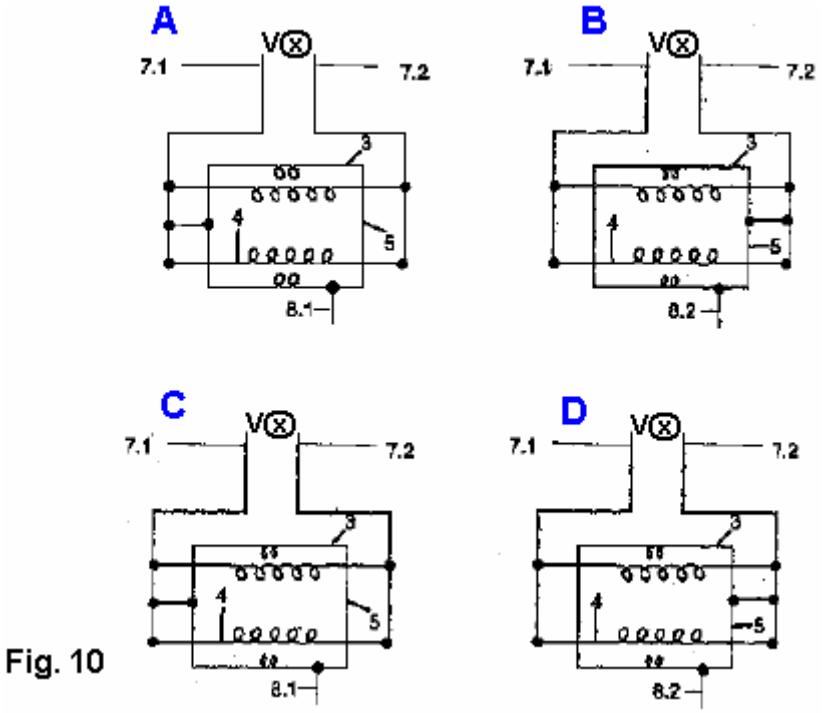


Fig. 10

Fig.10 illustrates a circuit diagram of the electrical connection between the coils **3** and **4** in sets **A, B, C** and **D**. It can be seen that the sets **A, B, C** and **D** are enclosed between the coils **3** and their associated inductors or electromagnets **4**. The supply conductors **7.1** and **7.2**, of sets **A, B, C** and **D** must be interconnected. When feeding power to the coils **3** and **4** the phase should be connected to **7.1** and the neutral to **7.2**.

The sets **A, B, C** and **D** after being fed with electric current, generate voltage through the attraction and repulsion of the electrons in the linking coil **5**, where there is at least one output load **8.1**, which should be connected joining sets **A** and **C**, and at least one load output **8.2**, which should be connected joining sets **B** and **D**. The output points **8.1** and **8.2** are the respective phases and neutral of power points **7.1** and **7.2**.

In this way, a singl-phase electron-trap is created by two pairs of sets of coils/inductors **3** and **4**.

The **3/4** electromagnet coil set can be replaced by a **3/3** coil set, without any disadvantage to the electron-trap.

Sets **A, B, C** and **D**, are inserted into a hollow metal hemisphere **1** preferably constructed from - but not limited to - aluminium. The hemisphere **1**, whose function is to concentrate and maximise their electromagnetic fields, simulating an electron cloud, has a fixed support **6** connected to attachment tabs **2**, and to which the coils **3** are fixed.

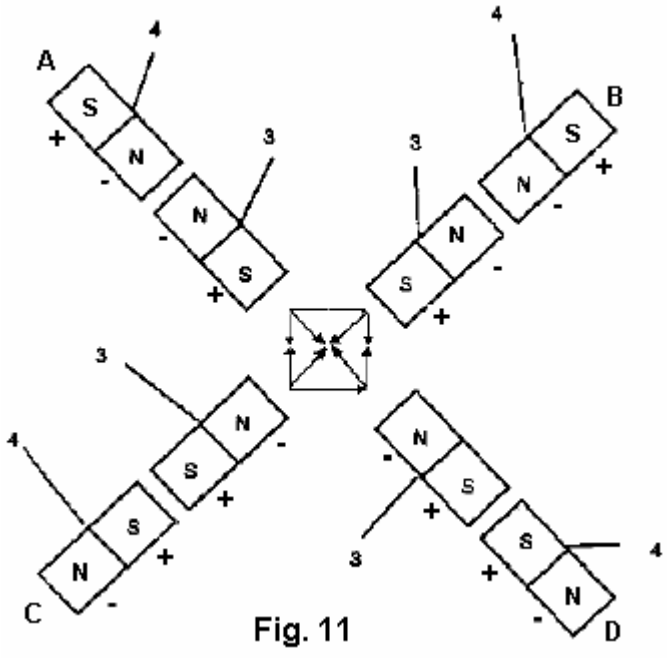


Fig. 11

Fig.11 is a diagram of the electromagnetic north and south poles of the inductor coils **3** and **4** of sets **A, B, C** and **D** of the electron-trap. The electromagnetic behaviour described for **Fig.9** is again demonstrated by the formation of the magnet assembly to the North Pole and South Pole being attracted and repelled by the lines of force of the magnet from the point "**A**" to point "**D**", point "**A**" to point "**B**", the point "**B**" to point "**C**", point "**C**" to point "**A**", and so on, as long as there is an electromagnetic field. The electromagnetic field of the space electron-trap provides that induced current in a direction similar to the variations of the magnetic flux that produced it. So, the magnetic field creates a north pole and a south pole in each of the sets **A, B, C** and **D**, as shown in **Fig.11**.

By feeding the coils **3** of the electron-trap with a desired voltage a magnetic field is generated in coils **3**, between the four sets **A, B, C** and **D**, which form a flow of electrons. This flow of electrons augments the electron flow which is circulating in the closed-loop link-coil **5**, thus implementing free electron capture from space. The electromagnetic field of the coil **3A** runs north to south, the electromagnetic field of the coil **3B** runs north to south, the electromagnetic field of the coil **3C** flows from south to north, and the electromagnetic field of the coil **3D** flows from south to north, as shown in **Fig.11**. It should be noted that the sets **A, B, C** and **D** can be formed by any combination of coil, magnet and electromagnet.

The south to north electromagnetic field induces current flow in the coil **3A**. The north to south electromagnetic field induces current flow in the coil **3B**. The north to south electromagnetic field induces current flow in the coil **3C** and The north to south electromagnetic field induces current flow in the coil **3D**. The induced current flow can have any power and it can be single-phase, two-phase or three-phase current.

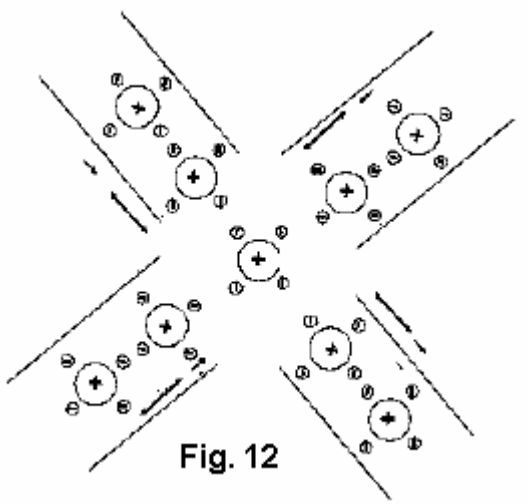


Fig. 12

Fig.12 shows the electrons being attracted and repelled by the induction coils **3** and **4**. Being repelled and attracted by electromagnetic induction, the electric current flows without resistance.

The electron-trap produces electromagnetic waves which can be used for various purposes, including signal transmission at any frequency and for any purpose. The capture is caused by these electromagnetic waves. The same physical effect can be achieved by the combination of the capture devices of other technologies, including electromechanical, electric, electronic, electromagnetic, or through the combination of a magnet or any other magnetised materials.

The space free electron-trap of the present invention is a renewable source of electrical power production and a new way of generating energy through the capture effect, generating flows of electrons, generating ordered movement of electrons - electric current - as shown in **Figures 9, 11,** and **12**. Electrons can move without any voltage difference in the continuous loop **5**. Alternatively, the loop may be biased with any chosen voltage.

The relevant Barbosa and Leal patents in Portuguese can be downloaded here:

http://www.free-energy-info.tuks.nl/Barbosa1.pdf

http://www.free-energy-info.tuks.nl/Barbosa2.pdf

http://www.free-energy-info.tuks.nl/Barbosa3.pdf

Scientific Papers

The following links connect to various scientific papers and documents of interest. As web-based resources are very prone to change and disappear, if you download any of these to read, I suggest that you store them on your local drive in case they become unavailable at a later date. If, for any reason, the www.free-energy-info.co.uk web site is not available, then you can try www.free-energy-info.com or www.free-energy-devices.com which are mirror sites.

http://www.free-energy-info.tuks.nl/P1.pdf 4 Mb Pulsed DC electrolysis

http://www.free-energy-info.tuks.nl/P2.pdf 360 Kb Water arc explosions

http://www.free-energy-info.tuks.nl/P3.pdf 388 Kb Plasma electrolysis

http://www.free-energy-info.tuks.nl/P4.pdf 321 Kb Cold water fog explosions

http://www.free-energy-info.tuks.nl/P5.pdf 151 Kb Engine operation with hydrogen added to the fuel

http://www.free-energy-info.tuks.nl/P6.pdf 63 Kb Bubbles and steam electricity

http://www.free-energy-info.tuks.nl/P7.pdf 600 Kb Stan Meyer's Water Fuel Cell

http://www.free-energy-info.tuks.nl/P8.pdf 3.5 Mb Stan Meyer's Water Fuel Cell

http://www.free-energy-info.tuks.nl/P9.pdf 303 Kb Water as Fuel by Stan Meyer

http://www.free-energy-info.tuks.nl/P10.pdf 68 Kb Solar Hydrogen Production

http://www.free-energy-info.tuks.nl/P11.pdf 328 Kb Passive Cooling

http://www.free-energy-info.tuks.nl/P13.pdf 347 Kb A Solar Ice-maker

http://www.free-energy-info.tuks.nl/P14.pdf 711 Kb Smart-Skin Shielding Technology

http://www.free-energy-info.tuks.nl/P15.pdf 215 Kb Physics for Engineers

http://www.free-energy-info.tuks.nl/P16.pdf 2.5 Mb Fuel Cell Handbook

http://www.free-energy-info.tuks.nl/P17.pdf 62 Kb Prof. Kanarev's Low-current Electrolysis

http://www.free-energy-info.tuks.nl/P18.pdf 8 Mb Principles of Ultra Relativity by Shinichi Seike

http://www.free-energy-info.tuks.nl/P21.pdf 754 Kb The Theory of Anti-gravity

http://www.free-energy-info.tuks.nl/P22.pdf 13.3 Mb Physics Without Einstein by Dr Harold Aspden

http://www.free-energy-info.tuks.nl/P23.pdf 6.9 Mb Modern Aether Science by Dr Harold Aspden

http://www.free-energy-info.tuks.nl/P24.pdf 10 Mb Unified Physics by Dr Harold Aspden

http://www.free-energy-info.tuks.nl/P25.pdf 1.5 Mb Permanent Magnet Motor Design

http://www.free-energy-info.tuks.nl/P26.pdf 402 Kb The Sea of Energy by Thomas Henry Moray

http://www.free-energy-info.tuks.nl/P31.pdf 14.5 Mb Tesla's Science of Energy by Dr Thomas Valone

http://www.free-energy-info.tuks.nl/P32.pdf 605 Kb High Voltage, High Frequency Alternating Currents by Nikola Tesla

http://www.free-energy-info.tuks.nl/P33.pdf 632 Kb "Prodigal Genius" biography of Nikola Tesla

http://www.free-energy-info.tuks.nl/P34.pdf 488 Kb Damped Waves by Nikola Tesla

http://www.free-energy-info.tuks.nl/P41.pdf 3.2 Mb Electrogravitics Systems by Dr Thomas Valone

http://www.free-energy-info.tuks.nl/P42.pdf 2.5 Mb Zero-Point Energy Extraction by Dr Thomas Valone

http://www.free-energy-info.tuks.nl/P63.pdf 181 Kb Scalar Waves by Konstantin Meyl

http://www.free-energy-info.tuks.nl/P65.pdf 592 Kb EsotericScience.com Presentation

http://www.free-energy-info.tuks.nl/P66.pdf 450 Kb The Genesis Project

http://www.free-energy-info.tuks.nl/Lodge.pdf 27 Mb. Modern Views of Electricity by Sir Oliver Lodge

http://www.free-energy-info.tuks.nl/Lodge2.pdf 9 Mb. The Ether of Space by Sir Oliver Lodge

http://www.free-energy-info.tuks.nl/Lodge3.pdf 22 Mb. The Nature and Properties of Negative Electricity by Sir Oliver Lodge

http://www.free-energy-info.tuks.nl/TeslaBook.pdf 36 Mb. The Inventions, Researches and Writings of Nikola Tesla by Thomas Commerford Martin

http://www.free-energy-info.tuks.nl/DonKelly.pdf 13 Mb. The 'Manual of Free Energy Devices and Systems' by Donald A. Kelly

http://www.free-energy-info.tuks.nl/Blitz.pdf 30 Mb. Ultrasonics by Jack Blitz

http://www.free-energy-info.tuks.nl/Magnetos.pdf 9.5 Mb. Magnetos Simply Explained by F. H. Hutton

http://www.free-energy-info.tuks.nl/TeslaPatents/TeslaPatents.pdf 100 Mb. All of Tesla's published patents

http://www.free-energy-info.tuks.nl/Magnetic_Motor.pdf 511 Kb Magnetic Motors by Lawrence Tseung

http://www.free-energy-info.tuks.nl/Maxwell.pdf 2.2 Mb The Electromagnetic Field by James Clerk Maxwell

http://www.free-energy-info.tuks.nl/Davson.pdf 46.7 Mb Davson's book on Karl Schappeller's work

http://www.free-energy-info.tuks.nl/Newman.pdf 97 Mb The Energy Machine of Joseph Newman

http://www.free-energy-info.tuks.nl/Combine.pdf 2.1 Mb Radiant Energy and Over-Unity by Dan Combine

http://www.free-energy-info.tuks.nl/Rodin.pdf 3 Mb The Rodin Solution Project by Marko Rodin

http://www.free-energy-info.tuks.nl/SEG.pdf 594 Kb Magnetic Gravity Research by Roschin and Godin

http://www.free-energy-info.tuks.nl/Stan_Meyer_Full_Data.pdf 3.8 Mb Technical Brief by Stan Meyer

http://www.free-energy-info.tuks.nl/LTseung.pdf 3.2 Mb Over-unity Devices by Lawrence Tseung

http://www.free-energy-info.tuks.nl/Smith.pdf 2.0 Mb. Resonant Energy Methods by Donald Lee Smith

http://www.free-energy-info.tuks.nl/Cottrell.pdf 2.8 Mb. How Gravity Works by Maurice Cottrell

http://www.free-energy-info.tuks.nl/Davidson.pdf 4.5 Mb. Shape Power by Dan Davidson

Videos

http://www.free-energy-info.tuks.nl/Meyer.wmv 4.6 Mb
http://www.free-energy-info.tuks.nl/Newman.avi 53 Mb
http://www.free-energy-info.tuks.nl/pyramid.avi 25 Mb
http://www.free-energy-info.tuks.nl/stage1.wmv 1.9 Mb
http://www.free-energy-info.tuks.nl/stage2.wmv 1.9 Mb
http://www.free-energy-info.tuks.nl/stage3.wmv 3.7 Mb
http://www.free-energy-info.tuks.nl/WFCrep2.wmv 1 Mb
http://www.free-energy-info.tuks.nl/stan.wmv 4.5 Mb
http://www.free-energy-info.tuks.nl/WFCrep.wmv 5.3 Mb
http://www.free-energy-info.tuks.nl/Bills.avi 17 Mb

http://www.youtube.com/watch?v=-6MDHF39XmU&NR

THE IMPULSE GENERATOR

Dr. Evgeny Podkletnov on the Impulse Gravity-Generator

By Tim Ventura & Dr. Evgeny Podkletnov, April 10th, 2006

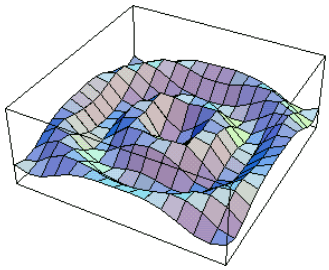
In 2001, Dr. Evgeny Podkletnov began publishing a series of scientific papers detailing the experimental results of what he called an "impulse gravity generator". The device reportedly produced hundreds of pounds of gravitational force in a non-diverging beam, capable of "punching holes through concrete and warping metal like hitting it with a sledgehammer". Podkletnov further added that this beam produced no recoil on the superconducting emitter itself, and that a radiation had been produced behind the device creating a molecular juxtaposition between plastics, metals, and living tissues similar to that described in the Hutchison Effect. This document is the result of 3 interviews between 2004 and 2006 that attempt to document and further clarify his remarkable experimental claims...

AAG: Thanks again for taking the time to speak with us about your research; it's truly been pioneering in the field of superconductive antigravity. A lot of the papers that I read online – and a lot of the references – date directly back to your research from the early 1990's, and more recently your experiments with Dr. Giovanni Modanese. Is there anything that you can tell us about what you've been up to lately?

Podkletnov: My research is based on the idea that it is possible to change or to modify the local gravitational field, and to consequently change the gravitational properties of objects within that field to make them lighter or heavier. This all can be done by creating special conditions in what might be described as the polarization of the ubiquitous vacuum, or by modifying the curvature of Einstein's space-time. There are several ways to do this – we can use high-voltages, large magnetic fields, and extremely high-speed rotation of various objects. We can also take advantage of topological effects in superconductive materials. Altogether, there are a number of possibilities, including potential combinations of some or all the parameters that I've just described.

Dr. Evgeny Podkletnov: Specialist in Type-II superconductive gravitomagnetic systems.

My latest experiment is a device called the "gravity impulse generator", which utilizes a Marx-Generator discharge through a superconducting emitter in a high-magnetic field to create a wave in time-space with properties very close to gravity-waves. The similarities are apparent enough we're almost positive that it actually is a form of gravity. These impulses can be directionally projected in any direction in space, and they exert a force on any object in the path of propagation. We haven't quite uncovered a detailed mechanism to explain how this force is generated, but we nonetheless have the technological possibility for industrial and commercial applications of these results, and this is the most important.



HFGW: A High Frequency Gravitational Wave model.

AAG: You'd mentioned that this effect generated hundreds of pounds of force in a very short period of time, and I'd wondered if

this effect has been able to punch holes through lightweight substances, or does it produce more of just a motion on them?

Podkletnov: The force of the impulse depends entirely on the structure of the superconducting emitter and the voltage that we apply to it. Given the materials & voltages we currently have available, we can obtain large impulses capable of deforming metal plates with a thickness of a couple of inches, and we've also been able to demonstrate punching holes in thick concrete walls. Obviously these effects are not limited to just lightweight substances – a concrete wall is something very solid. The impulse deforms metal in the way that a hydraulic press might do it, but the pulse-duration is very short in time, so we've been discussing a system utilizing several Marx-Generators to give a series of impulses that will definitely improve the overall effect. We've experimented with using the impulse-generator to treat a variety of materials, and we've also made another important find: the beam can hit a target over very large distances with a minimum of divergence and what appears to be zero loss in energy, even after passing through other objects in the beam-path.

By using extreme parameters we can create an even larger impulse forces. We discovered in the last year and a half that if we go to 5-million volts we can generate hundreds of pounds of force. At that time, we were also trying to measure the speed of propagation for these impulses. The results were extremely interesting, and hard for even us to believe, because to some extent it contradicts many aspects of modern physics. Nonetheless, we have experimental proof, and we're going to continue developing all of these experiments.

Marx Generator: An example 2 megavolt model constructed by Information Unlimited.

AAG: Oh, so these are all generated by Marx-Generator discharges?

Podkletnov: It is possible to use high-voltage discharges up to 5 million volts and specially prepared superconductors. These impulses are very short in time, we're talking about 1 millionth of a second and shorter, and they can be very powerful, and can be used for rocket propulsion in space. At the same time, they can be used for the correction of satellite orbits from the earth with minimum expense, making this a very promising technology.

AAG: In terms of putting holes through concrete, it makes sense that the holes would remain, but with metals, after the metal is deformed, do they snap back after the beam is gone, or do they remain deformed?

Podkletnov: No, they just remain deformed. It's just like a punch – very short in time, so it's close to an explosive action or something like that.

AAG: One of the things that I was wondering about is whether you've been able to do efficiency calculations for the force-beam.

Podkletnov: Well, Dr. Giovanni Modanese made some preliminary measurements that gave us the force in joules, but we did not try to make predictions – we wanted to simply see the results of how different objects reacted to the action of this impulse, so we didn't make any precise calculations.

AAG: The type-II YBCO emitter measured 47mm in diameter, if I remember correctly. Does changing the size or the shape change the beam output – maybe making it stronger or refocusing it?

Podkletnov: Well, in terms of the size of the superconducting emitter, there are some limitations. The diameter of the superconductor shouldn't be smaller than 4 inches, because of the Schwartzchild radius. Now to answer your question, in terms of the shape of the emitter, the superconductor can in fact have different shapes, and the projected impulse will maintain the emitter's cross-sectional shape, so is important.

Impulse Gravity Generator (initial setup)

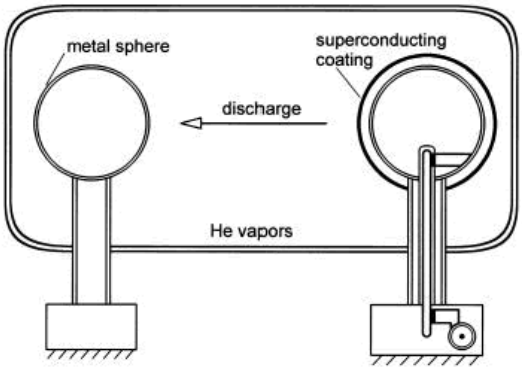


Fig. 1

AAG: The 4-inch minimum size for the superconductor seems to limit the public's ability to perform experimental replications. The largest commercial available superconductor that I've seen is a 1" superconductor – so it sounds like the emitters require custom fabrication?

Podkletnov: Well the size is very important – we didn't get any good results with smaller superconductors, and it's also more difficult to create the correct flat-glow discharge with smaller superconductors.

Impulse Gravity Generator (improved variant)

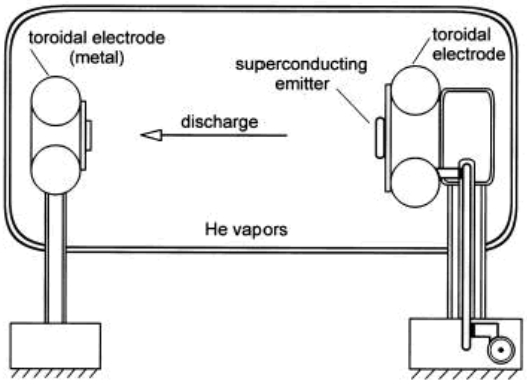


Fig. 2

Emitter Configuration: Schematic of two superconducting emitter apparatus designs.

AAG: Have you noticed any changes in the molecular structure, internal materials deformation, or maybe simple mechanical compression in the targets that you'd sent the beam through? You'd mentioned metals deforming and holes through concrete...

Podkletnov: We didn't see any compression effects or any change of the molecular structure – just a large scale deformation of the target material from the beam's force.

AAG: I should ask whether the beam loses energy as it penetrates materials. Does it naturally decrease or diverge with distance?

Podkletnov: That's an interesting question...and to our great surprise the beam practically does not lose energy when it meets the materials. It can pass through a brick wall or concrete, or metal-plates – very thick ones – or plastics also, and it seems that it doesn't lose energy at all. This is the consistent, long-term evidence from a lot of test discharges we performed at an installation I've been working at for about 4 years.

These results seem a bit strange, but we don't believe we're breaking any natural laws...simply that the system we're working in is not a closed one, and therefore the second law of thermodynamics is not applicable in the traditional sense.

In terms of action at a distance – and the dependence of distance on the beam energy – we don't have much experimental data, but what we do have is a first measurement at a distance of 1.2 kilometers without any loss in energy. Our latest experiment was conducted over a distance of 5 kilometers, and the beam penetrated through several houses made of concrete. We did not measure any loss of energy, but after closely evaluating some of the calculations that we've made, we should get some decrease in beam-energy at distances greater than 100 kilometers. This is research that awaits us in the future.

AAG: You'd mentioned 5 kilometers – did you notice any change in the focus of the beam. Did it widen or perhaps get smaller as it travels?

Podkletnov: If the main solenoid which is wound around the chamber is made in a good way, then we have a very good discharge and it effectively maintains a non-divergent cross-sectional pattern of the emitter it was projected from. However, at a distance of 5 kilometers, the beam begins to lose focus– it gets a bit wider than it was, indicating minor deviations in the shape of the impulse as it propagates.

AAG: I collected a few questions online, and one person wrote me, "Is it possible to generate more work along the path of the beam than energy put into the beam?" I think they were asking about potential "over-unity" applications.

Podkletnov: It's been surprising to find that the energy that we put into the discharge is much less than the energy that the impulse seems to generate, but it doesn't mean that it's an over-unity device, simply that we're creating a set of special space-time conditions through the interactions of the electromagnetic pulse-discharge with the bose-einstein condensate (superconductor). By manipulating these

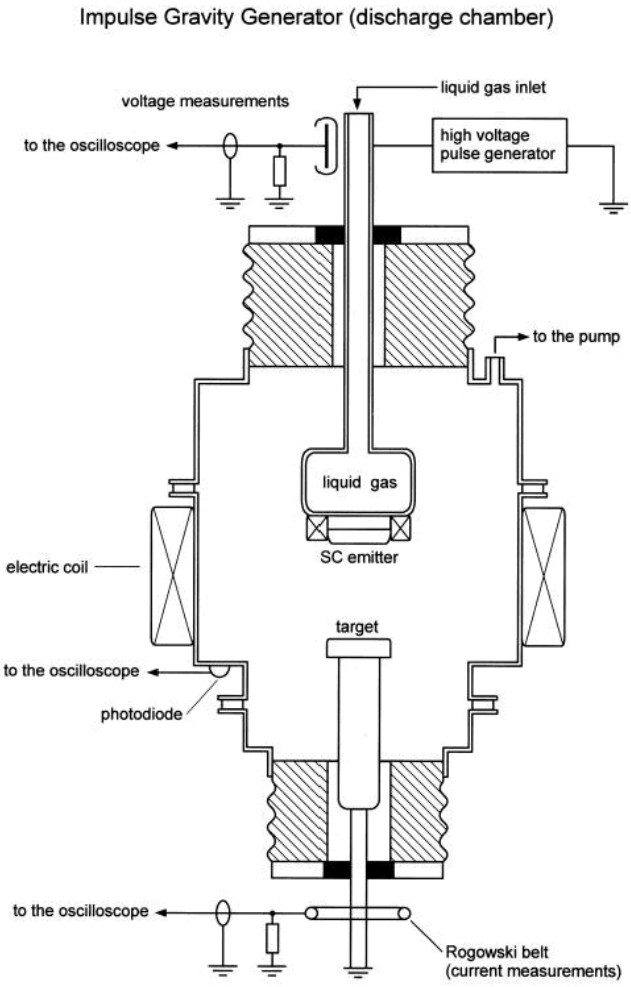


Fig. 3

Discharge Chamber: A complete layout-schematic of the chamber showing experimental components.

parameters we allow the interaction of electromagnetic fields with the sub-atomic field environment. We may call it zero-point energy or maybe even aether, but whatever it is, when normal matter interacts with it this special energy is obtained, and we can harness this energy. Our technology may be like a key that opens the door to the energy of the subatomic realm. This is at least our fervent hope...

AAG: Does the impulse-generator seem to defy conventional Relativity Theory? I'm wondering if you might have any specific examples or observations of things that Relativity Theory simply can't explain...

Podkletnov: I would never say that that our experiments defy conventional Relativity Theory because they don't. You know, people also suggested that about my earlier experimentation with the rotating disks, but in that case the rotations around the disk's own axis are absolute rotation. This is absolute movement, not relative, which is why relativity theory is not applicable to our rotating disks – it's a totally different thing. However, I don't think the force-beam experiment defies Relativity Theory either.

AAG: Does the inertia change in proportion to changes in mass? The reason that I wrote that down was really in relation to Fran De Aquino's research. He believes that inertial & gravitational mass are not equal, they just appear equal to us in our frame of reference.

Podkletnov: Well, based on our understanding and our experience, we didn't differentiate between gravitational and inertial mass, and we believe that they are equal. Experimentally, the pulse-width of the impulses is too short to notice any difference, and we never conducted any special experiments just to distinguish gravitational and inertial mass.

AAG: In Relativity Theory, gravitation and time are integral components of the "fabric of time and space" –and modifying one inherently changes the other. Given your description of the impulse-beam as being gravitational in nature, have you noticed any experimental time effects?

Podkletnov: It's difficult to say, because in a practical sense it's not truly a gravity-beam – it's a gravity-impulse with a very short time-duration, and because of this limitation -- because it has such a short pulse-width -- we didn't notice any time effects and didn't even try to measure them, but it's possible they are present. These effects were definitely present when we were experimenting with the rotating disks – definitely – and we had some experience with this, but with the gravity impulse-generator we simply didn't conduct any measurements.

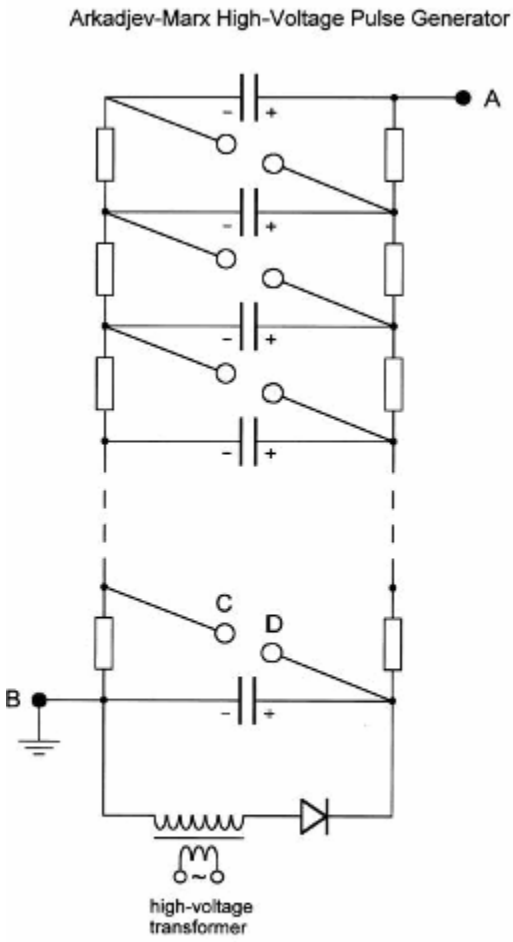


Fig. 4

Marx Circuitry: An overview of an idealized circuit for the Marx-Generator.

AAG: Now you've mentioned working on some scientific publications, and I'd like to find out more about what topics you plan on publishing about in the future, and what your schedule for publication might be?

Podkletnov: Well, of course there is a growing interest in the concept of manipulating gravity in a number of countries. For instance, we got a very interesting offer from China-- they have a very special project at the University of Beijing – and there's also been some interest from the private sector in the United States as well.

As soon as we get funding we'll try to publish on a variety of experiments in a more detailed manner, but at present we're extremely interested in attempting to measure the interaction of the impulse-beam with visible light, and we published some preliminary results about this in the journal of low-temperature physics in August of 2004. The details are all in my article with co-author Dr. Giovanni Modanese, and we're continue this research now, again with a focus on attempting to measure the propagation speed of the impulse.

We're very cautious about what we write, because we don't want to frighten the scientific community, and also we want to be absolutely sure that the results are checked and rechecked several tens of times-- but it seems that based on what we have now, and we've already been working for a year and a half, the speed of the impulse is much higher than the speed of light.

With the parameters that we use now – using our current emitter designs and a voltage of 3 and 5 million volts, the speed is about 63 to 64c – which means that the propagation speed of the impulse is close to 64 times the speed of light. Of course, we would like to measure all these parameters using different measurement systems and different approaches. At present we use two atomic clocks, and we think that our measurements were precise enough, but we would welcome the advice of the international community. It would help to have additional input on how to measure the speed of the impulse in a very specific way. As soon as we get a good confirmation of these results, we will try to publish all of this information.

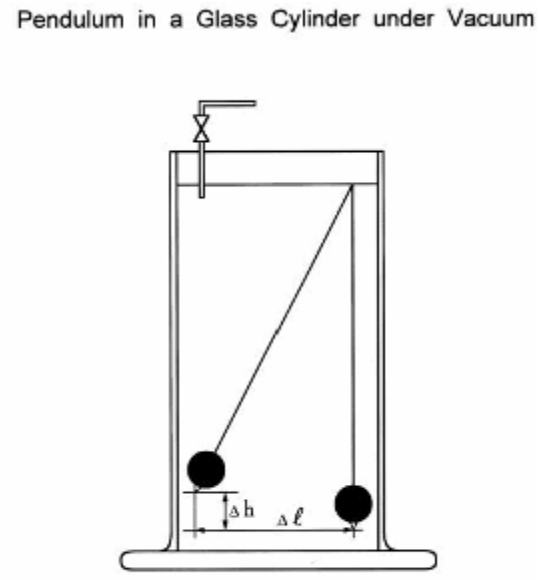


Fig. 5

Pendulum: Podkletnov's initial force-beam experimental measurement apparatus.

AAG: I was wondering about funding: you've mentioned that there was interest from the private sector in the United States and from China – do you know if there are any replication efforts underway by other groups to duplicate your research?

Podkletnov: I know that there was a lot of interest from Boeing, but I don't know the details. I also know that the Department of Defense in the United States is very interested in this technology, and that's why they invited Dr. Ning Li to lead the scientific laboratory funded by AMCOM, but at present I don't know of any official replication of my gravity experiments. This is because my research is rather hard to categorize, it's costly, and the official attitude of the

"politically correct" scientific establishment towards gravity-modification is negative, which creates a lot of difficulties.

However, I don't hide anything, and when people contact me directly or by email I usually try to give all the advice that I have and to share my experience with them. The problem is too complicated for one country or for one lab to succeed; gravity should be studied all over the world using the best forces and the brains of difference scientists. That's the key to success.

Dr. Modanese: Podkletnov's research partner, physicist, and force-beam co-author.

AAG: I think we've already agreed that one of the chief obstacles to a successful replication is the 4-inch superconductor – do you know if those are manufactured and sold anywhere, or is that a process that everyone has to go through to build their own?

Podkletnov: Frankly speaking, in the case of extremely effective emitters that's a part of my professional know-how, but if you're only talking about emitters that allow you to generate small effects, then it's not a problem. I believe that American Superconductor can help to easily make emitters of this kind, and also there is a nice firm called "Superconductive Components" in Columbus, Ohio – they're more or less familiar with my technology, and I think that they are up to the task of building emitter components.

AAG: So essentially, by using a smaller superconductor you have a smaller effect, but that can be tested by using more sensitive equipment.

Podkletnov: The diameter of the disk should still not be less than 4-inches, but I'm talking about the physical structure of the ceramic itself. The structure for efficient emitters is very difficult to build, and requires a lot of experience to build correctly, so even if I provided a detailed description, it would be difficult to construct without my help. Normal emitters which allow you to push a thick book away from a table are possible to construct, as they aren't quite so complicated.

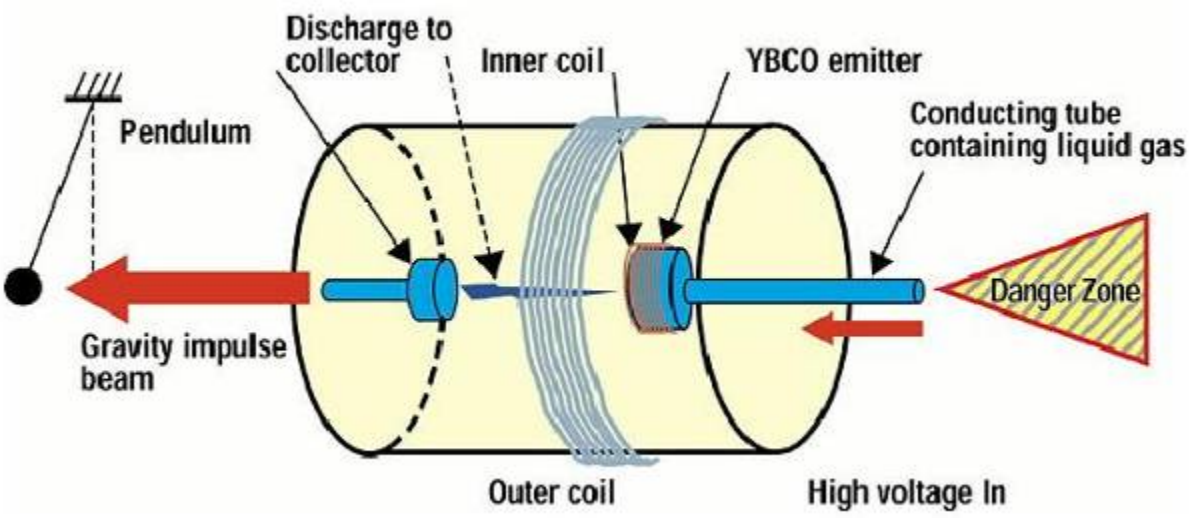
Dr. Ning Li: A colleague of Podkletnov's and specialist in gravitomagnetic coupling effects.

AAG: I had a question from another person who was asking about video, but they were also asking about a flat-glow discharge.

Podkletnov: Well, in order to make a video for the flat-glow discharge, we should use a high-speed video camera that we don't presently have, so we typically have to rely on our keen sight, but it's a bit difficult. We're planning to film it, but even with a normal discharge – a spark, as in a van de graaf generator, or in a flat-glow discharge, which repeats the configuration of the emitter, it's possible to see it with your own eyes -- you don't need a camera for that.

AAG: Now in terms of experimental setup, because I'm getting ready to buy parts...it looks like, from reading through your experimental setup, that essentially you're discharging a high-voltage onto a superconductor – that, and the superconductor itself is inside a magnetic field. It looks in some ways deceptively simplistic, but I understand that there is a definite challenge involved with it. Have you found any real tricks to testing with this?

Podkletnov: Well, this technology is really nothing special, but I work with a team of experienced people who are experts in high-voltage discharges, experts in ceramic materials & superconductors, and experts in magnetic fields, and only by combining the knowledge different fields and people from different countries can we reach something positive. This international collaboration becomes inevitable, and becomes the key to success.



Emitter Experimental Setup: An overview of the emitters subjected to a high-voltage discharge.

AAG: Dr. Bob Baker is working on putting together a 2006 High-Frequency Gravity-Wave Conference as a follow-up to the earlier one in 2003. Has he contacted you about this, and would you be open to attending?

Podkletnov: No, he hasn't sent me anything. You know, I have a job here and I don't want to risk it, especially without firm opportunities for funding. I'm hoping to find investors wise enough to see how much money this research can generate. Unfortunately, up to now they don't understanding it – if I release this technology, if they put money in it – then we will definitely succeed in it, and in 2 or 3 years they will be richer than Bill Gates, but this understanding comes very slowly. We've been hoping for commercial funding, because the impulse generator is basically a gravity-gun, and I'm not really interested in providing a weapon of strategic importance to the United States. If we're talking about new propulsion systems, though, I'm eager to share my knowledge. Of course the gravity-beam can also used for propulsion, but the rotating superconductor can be used only for creating propulsion – a vertical column to help lift a craft off the surface of the earth.

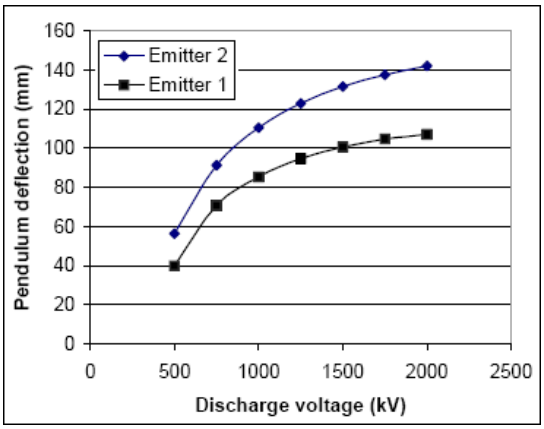
AAG: The force-beam that you'd mentioned coming out of the back of the impulse generator had a negative effect on living organisms, right?

Podkletnov: Yes, it produced a strange kind of non-focused radiation that was very harmful to biological tissue -- fortunately for us; the impulses were very short in time. It was very harmful to living organisms; we also observed a melting or molecular-juxtaposition effect of biological tissues with plastic materials, and even with metals to some extent.

AAG: So this was a very different force than the impulse coming out of the front of the generator?

Podkletnov: Yes, it was an entirely different force, and it was not focused like the gravity-beam.

AAG: In "The Hunt for Zero Point", Nick Cook described an experiment in the 1940's called the "Nazi-Bell" that reportedly exhibited side-effects like the crystallization of cell-membranes in plants and the breakdown of living tissue into a greasy substance after repeated exposure. Does this sound similar in any way to what you saw emanating from the back of the impulse-generator?



Deflection: Deflection of pendulum during testing as a function of discharge voltage.

Podkletnov: Well, I can't speak for what they might have seen – during the 40's I hadn't even been born yet! However, it might be that the side-effects are similar to ours. There are several effects that are similar to this – one is the Hutchison Effect, where he reported molecular distortions & materials juxtapositions, which is similar to the anecdotal claims of "sailors melting into decks" in the Philadelphia Experiment legend.

AAG: Oh, so you're actually talking about a radically different set of associated effects...did you witness anything like atomic transmutations or Low-Energy Nuclear Reaction (LENR) effects occurring, or just the molecular-level anomalies?

Podkletnov: Well, the radiation that emanates from the rear of the impulse-generator might somehow create conditions suitable to LENR effects, but this wasn't an area of focus in our research, so I really don't know. Our primary focus was on studying the gravity-beam itself.

AAG: In terms of the force-beam, you've said that it was hundreds of pounds of force, and I realize that it was very difficult to measure. You've said that it was also enhanced with 5-million volt discharges. Can you tell me a bit more about this?

Podkletnov: Yes, it output force depends on the voltage and also on the effect of how sharp the the impulse is. If the rise-time of the impulse is long in duration, then we have a lower-power impulse, and if it is very sharp – I mean the voltage rises very fast – then the effect is very large and is able to bend metal plates.

AAG: Ok, did you test on anything besides metal or concrete, to try some other substances?

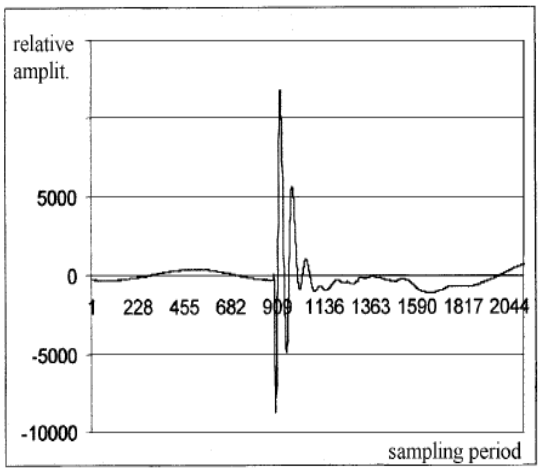
Podkletnov: It's not a drilling-machine, sorry. We're interested mainly in the use of this technology for propulsion in space, and frankly speaking it really is what you'd call "propellantless propulsion".

AAG: Let's focus on that for a second: when you noticed the impulse force on objects in the beam path, was there an equal & opposite force on the superconductor itself?

Podkletnov: No, there wasn't. There was no reaction force at all.

AAG: So if you put the superconductor inside of a spacecraft then, I guess that it wouldn't produce a conventional forward propulsion?

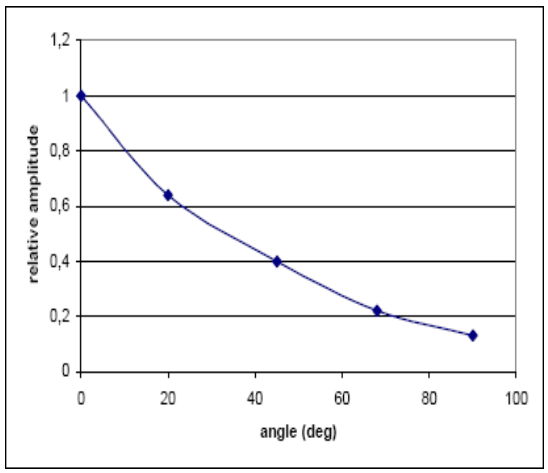
Podkletnov: Yes, it doesn't act according to Newtons third law – that every force has an equal an opposite reaction. But simply the superconducting emitter, whether it is a rotating disk or the discharge-system, is able to create a gravitational wave in what you might call the aether or the space of subatomic particles. Or you call it a gravity-wave. So this wave propagates through space, through all the objects, sometimes it can interact with normal matter, but to my surprise it does not lose its energy at very high distances – it remains collimated as it propagates.



Amplitude: Relative pulse-amplitude from impulse-discharge experimental testing.

AAG: Did you notice any other strange effects in the general area? Anything in addition to the impulse itself, the effect on objects, or the strange radiation coming out behind?

Podkletnov: No, we didn't, but frankly speaking it is not very easy to measure all these additional effects. You know, when you're dealing with millions of volts, it's better to keep a bit of safe distance. We use a faraday cage and special rubber-metal coatings to shield the radiation, because otherwise the high magnetic field strength from the discharge will erase computer hard drives and damage nearby test-equipment.



Deflection: Angle of pendulum deflection correlated with adjusted pulse amplitude.

AAG: When you analyzed the metal sample that was bent with the gravity-beam, did it appear to be only bent forward in the direction of beam propogation? Dr. Bob Baker postulates that gravity-waves may be quadripolar in nature, and I'm wondering if these waves might leave a detectable signature in the materials they deform to help validate his theory...

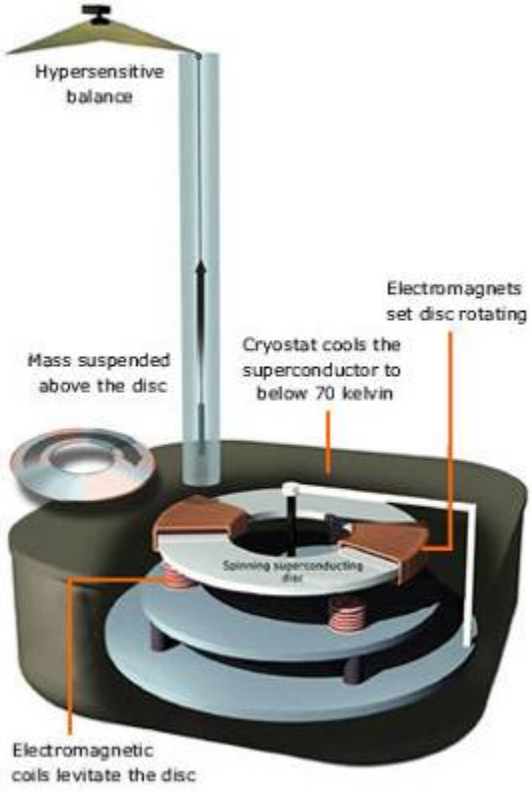
Podkletnov: In our experiments we were able to observe only the forward action.

AAG: Do you think that there might be any torsion physics involved with the impulse-generator?

Podkletnov: I think that torsion or axial fields may be present in our experiments, but it's more suitable to suggest a role for torsion physics in our original rotating disk experiments - the mechanism behind the impulse-generator is a bit different.

AAG: Speaking of which, in terms of the failed NASA replication for the rotating-disk experiment, I'd heard from a NASA insider that they only spun it up to about 200rpm. Does that mean that they didn't fully complete your experiment?

Podkletnov: I cringe every time I hear that NASA failed to replicate my experiment, because no, they didn't fail. They made their own disks, and they were big enough: about 12 inches in diameter. They published some initial test information indicating that they had definitely noticed some unusual effects. Then I got involved in participating to helping them to replicate my experiments, and they practically had everything ready when they ran out of money. So at the last stage they were not able to rotate the superconductor in the magnetic field, and shortly after that the department of defense came in and grabbed all the experiments. All of this research was transferred to Dr. Ning Li – so now NASA has nothing, and we have nothing either.



Rotating Disk: A cutaway of the Type-II disk apparatus from the first experiment.

AAG: So is Dr. Ning Li still in charge of the project then?

Podkletnov: I think so, but the point is that NASA never completed the experiment, and thus nobody can claim that they failed to see results, because they never completed it.

AAG: Have you ever been approached by the Department of Defense for any kind of funding?

Podkletnov: No, they would not do it. I am a citizen of Russia so they wouldn't even try.

AAG: Well, the Chinese, though, have a gravity-lab: have they expressed any interest in your research?

Podkletnov: They expressed real interest and they've invited me to China several times, but they are like kindergarteners: they want me to come just and give them completed technologies. They have a rather good theoretical school but they haven't achieved much in practice, and they seem rather limited in funding, so I am not eager to travel to China and work there at my own expense.

AAG: Have you ever been able to gravity-research lab in China? Bob Baker visited a couple of years ago, and was impressed by their efforts.

Podkletnov: No, but I'm perfectly aware of the projects they have and their progress in development, because I have very good relationships and discussions with my colleagues from different countries, although I haven't seen it with my own eyes.

AAG: One of the responses that we had online was the question, "with nearly a decade of experience, why haven't we been able to see videos or photos of your experiments up until now?"

Podkletnov: Well first of all, when we began those experiments, in Tampere in the early 90's, it just wasn't common practice to make videos or photos of the equipment or of the experiment. I know that it's typical in the United States, but here in Europe it's different...and the same thing goes for Russia, especially with the last experiment in the Moscow Chemical Research Center, because the whole center is a very secure facility and some of the research laboratories are closed to the general public. In our case, nothing was secret, but it was a dangerous environment because we use high-voltages, so it's a closed lab. We have special signs on the walls of the laboratory which do not allow us to make any photos. It's an accepted, established policy at this scientific research center, and I didn't want to break the rules.

AAG: Well, in light of the publicity that you've had recently, have you given thought to doing any photos or video in the near future?

Podkletnov: I've discussed this possibility with the administration and they think it might be possible, but at present we don't have any photos or video to share with you.

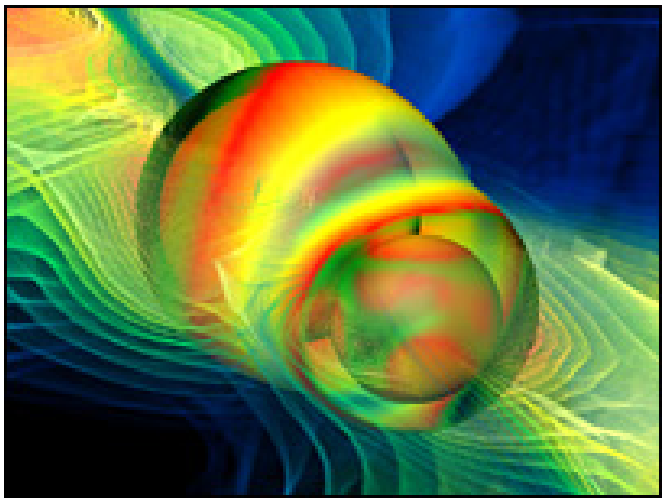
AAG: OK – well, I look forward to hopefully seeing some in the future if it turns out that it's possible. To move along, have you been able to obtain funding from government or private interests? Has there been a fair amount of interest that's come forward and tried to provide funding for your experiments?

Project Delta-G: The now-defunct NASA group tasked with examing superconductive gravity-coupling.

Podkletnov: There has been some interest from the private sector, because by policy the government does all of the research into gravity and experimental gravity. This type of research is not popular, so we don't get much money from the government. Of course, we use the right installations at the technological center. There is some interest in general all over the world – like the United States and Great Britain, but we didn't get any funding from the government. We get rather small funding from the private sector, but our plans are really amazing and we need considerable fundin g. So this topic of gravity is a bit unusual, and we need some exotic materials...we need special installations, we need cryogenic systems, we need the help of top people who are qualified in their own areas, and these all cost a lot.

AAG: Just to wrap things up, because we're almost out of time, again it sounds like you're getting amazing results – you're able to put holes through concrete and bend metals with this at 5 million volts – these are remarkable results. Is there anything you'd like to say in close?

Podkletnov: I don't think that these results are remarkable in general. This subject called experimental gravity research has a very big potential, and if we compare how complicated this problem is to the problem of, let's say, nuclear explosives, I think that gravity research is much more complicated. But even if we speak about nuclear power, there was a period in the United States when everybody was interested, and the military and people wanted to make some research, and the government was interested. Then people came and said, "could you please make a small explosion, and then we will give you money for a big one?"...but it's impossible to make a small nuclear explosion. The same thing refers to gravity – it's an enormous problem, and we can't get much if we don't have an organized approach as it was in the nuclear program for example in the United States. Only combining the knowledge of different fields, and of different physicists, chemists, and materials scientists & theoretical physicists – only through them working together can we make a breakthrough in this field, because it's very, very serious research.



Gravity-Wave: The core of a gravity-wave event as predicted by Dr. Robert Baker's HFGW research.

Dr. Evgeny is a research scientist living in Finland, with an impressive body of experimental research relating to EM/gravitational coupling in superconductive materials. His published abstracts about the impulse generator experiment are online at: Investigation of HV discharges through large ceramic superconducting electrodes, Superconductor Impulse Gravity Generator

Is a Graviton Detectable?

Poincare Prize Lecture
International Congress of Mathematical Physics
Aalborg, Denmark, August 6, 2012

Freeman Dyson, Institute for Advanced Study, Princeton, New Jersey

1. Introduction

It is generally agreed that a gravitational field exists, satifying Einstein's equations of general relativity, and that gravitational waves traveling at the speed of light also exist. The observed orbital shrinkage of the double pulsar [Weisberg and Taylor, 2005] provides direct evidence that the pulsar is emitting gravitational waves at the rate predicted by the theory. The LIGO experiment now in operation is designed to detect kilohertz gravitational waves from astronomical sources. LIGO has not yet detected a signal, but nobody doubts that gravitational waves are in principle detectable.

This talk is concerned with a different question, whether it is in principle possible to detect individual gravitons, or in other words, whether it is possible to detect the quantization of the gravitational field. The words "in principle" are ambiguous. The meaning of "in principle" depends on the rules of the game that we are playing. If we assert that detection of a graviton is in principle impossible, this may have three meanings. Meaning (a): We can prove a theorem asserting that detection of a graviton would contradict the laws of physics. Meaning (b): We have examined a class of possible graviton detectors and demonstrated that they cannot work. Meaning (c): We have examined a class of graviton detectors and demonstrated that they cannot work in the environment provided by the real universe. We do not claim to have answered the question of "in principle" detectability according to meaning (a). In Section 3 we look at detectors with the LIGO design, detecting gravitational waves by measuring their effects on the geometry of space-time, and conclude that they cannot detect gravitons according to meaning (b). In Sections 4 and 5 we look at a different class of detectors, observing the interactions of gravitons with individual atoms, and conclude that they cannot detect gravitons according to meaning (c). In Sections 6 and 7 we look at a third

class of detectors, observing the coherent transitions between graviton and photon states induced by an extended classical magnetic field, and do not reach any definite conclusion.

This paper is a report of work in progress, not a finished product. It raises the question of the observability of gravitons but does not answer it. There is much work still to do.

2. The Bohr-Rosenfeld Argument

Before looking in detail at graviton detectors, I want to discuss a general theoretical question. In 1933 a famous paper by Niels Bohr and Leon Rosenfeld, [Bohr and Rosenfeld, 1933], was published in the proceedings of the Danish Academy of Sciences with the title, "On the Question of the Measurability of the Electromagnetic Field Strengths". An English translation by Bryce de Witt, dated 1960, is in the library at the Institute for Advanced Study in Princeton, bound in an elegant hard cover. This paper was a historic display of Bohr's way of thinking, expounded in long and convoluted German sentences. Rosenfeld was almost driven crazy, writing and rewriting fourteen drafts before Bohr was finally satisfied with it. The paper demonstrates, by a careful and detailed study of imaginary experiments, that the electric and magnetic fields must be quantum fields with the commutation relations dictated by the theory of quantum electrodynamics. The field-strengths are assumed to be measured by observing the motion of massive objects carrying charges and currents with which the fields interact. The massive objects are subject to the rules of ordinary quantum mechanics which set limits to the accuracy of simultaneous measurement of positions and velocities of the objects. Bohr and Rosenfeld show that the quantum-mechanical limitation of measurement of the motion of the masses implies precisely the limitation of measurement of the field-strengths imposed by quantum electrodynamics. In other words, it is mathematically inconsistent to have a classical electromagnetic field interacting with a quantum-mechanical measuring apparatus.

A typical result of the Bohr-Rosenfeld analysis is their equation (58),

$$\Delta E_x(P)\Delta E_x(Q) \sim \hbar|A(P,Q) - A(Q,P)|. \tag{1}$$

Here the left side is the product of the uncertainties of measurement of two averages of the x-component of the electric field, averaged over two space-time regions P and Q. On the right side, $A(P,Q)$ is the double average over regions P and Q of the retarded electric field produced in Q by a unit dipole charge in P. They deduce (1) from the standard Heisenberg uncertainty relation obeyed by the measuring apparatus. The result (1) is precisely the uncertainty relation implied by the commutation rules of quantum electrodynamics. Similar results are found for other components of the electric and magnetic fields.

The question that I am asking is whether the argument of Bohr and Rosenfeld applies also to the gravitational field. If the same argument applies, then the gravitational field must be a quantum field and its quantum nature is in principle observable. However, a close inspection of the Bohr-Rosenfeld argument reveals a crucial feature of their measurement apparatus that makes it inapplicable to gravitational fields. In the last paragraph of Section 3 of the Bohr-Rosenfeld paper, they write: "In order to disturb the electromagnetic field to be measured as little as possible during the presence of the test body system, we shall imagine placed beside each electric or magnetic component particle another exactly oppositely charged neutralizing particle". The neutralizing particles have the following function. Suppose we have a mass carrying a charge or current J whose movement is observed in order to measure the local electric or magnetic field. The movement of the charge or current J produces an additional electromagnetic field that interferes with the field that we are trying to measure. So we must compensate the additional field by adding a second mass, carrying the charge or current $-J$ and occupying the same volume as the first mass. The second mass is constrained by a system of mechanical linkages and springs to follow the movement of the first mass and cancels the fields generated by the first mass. This cancellation is an essential part of the Bohr-Rosenfeld strategy. It is then immediately obvious that the strategy fails for measurement of the gravitational field. The test-objects for measuring the gravitational field are masses rather than charges, and there exist no negative masses that could compensate the fields produced by positive masses.

The conclusion of this argument is that the Bohr-Rosenfeld analysis does not apply to the gravitational field. This does not mean that the gravitational field cannot be quantized.

It means only that the quantization of the gravitational field is not a logical consequence of the quantum behavior of the measuring apparatus. The fact that the electromagnetic field must be quantized does not imply that the gravitational field must be quantized.

3. Can LIGO Detect a Graviton?

In the LIGO experiment, if it is successful, we shall detect a classical gravitational wave, not an individual quantum of gravity. A classical wave may be considered to be a coherent superposition of a large number of gravitons. LIGO is supposed to detect a wave with a strain amplitude f of the order of 10^{-21}. According to [Landau and Lifshitz, 1975], page 370, the energy density of this wave is

$$E = (c^2/32\pi G)\omega^2 f^2, \tag{2}$$

where G is Newton's constant of gravitation and ω is the angular frequency. For a wave with angular frequency 1 Kilohertz and amplitude 10^{-21}, Eq. (2) gives an energy density of roughly 10^{-10} ergs per cubic centimeter. A single graviton of a given angular frequency ω cannot be confined within a region with linear dimension smaller than the reduced wavelength (c/ω). Therefore the energy density of a single graviton of this frequency is at most equal to the energy of the graviton divided by the cube of its reduced wave-length, namely

$$E_s = (\hbar\omega^4/c^3). \tag{3}$$

For an angular frequency of 1 Kilohertz, the single graviton energy density is at most 3.10^{-47} ergs per cubic centimeter. So any gravitational wave detectable by LIGO must contain at least 3.10^{37} gravitons. This wave would be barely detectable by the existing LIGO. For a LIGO apparatus to detect a single graviton, its sensitivity would have to be improved by a factor of the order of 3.10^{37}. Even this vast improvement of sensitivity would probably not be sufficient, because the detection of weak signals is usually limited not only by the sensitivity of the apparatus but also by the presence of background noise. But to see whether detection of single gravitons is possible in principle, we disregard the problem of background radiation and analyze the structure and operation of a super-sensitive LIGO detector.

For a rough estimate of the sensitivity of a LIGO apparatus required to detect a single graviton, we equate (2) with (3). This gives the strain f to be detected by the apparatus,

$$f = (32\pi)^{1/2}(L_p\omega/c), \tag{4}$$

where L_p is the Planck length

$$L_p = (G\hbar/c^3)^{1/2} = 1.4 \times 10^{-33} \text{ cm}. \tag{5}$$

The strain is derived from a measurement of the variation of distance between two masses separated by a distance D. The variation of the measured distance is equal to fD, so long as D does not exceed the reduced wave-length (c/ω) of the graviton. For optimum detectability we take D equal to (c/ω). Then the variation of distance is by (4)

$$\delta = (32\pi)^{1/2}L_p. \tag{6}$$

Up to a factor of order unity, the required precision of measurement of the separation between the two masses is equal to the Planck length, and is independent of the frequency of the graviton.

Is it possible in principle for a LIGO apparatus to measure distances between macroscopic objects to Planck-length accuracy? The following simple arguments give a negative answer to this question. First consider the case in which the objects are floating freely in space. The Heisenberg uncertainty relation between position and momentum of freely floating objects gives the lower bound

$$M\delta^2 \geq \hbar T, \tag{7}$$

for the variation of distance δ, where M is the mass of each object and T is the duration of the measurement. Now T must be greater than the time (D/c) required to communicate between the two masses. If δ is equal to the Planck length, (5) and (7) imply

$$D \leq (GM/c^2). \tag{8}$$

So the separation between the two objects is less than the Schwarzschild radius of each of them, the negative gravitational potential pulling them together is greater than Mc^2, and they are bound to collapse into a black hole before the measurement can be completed.

We next consider the situation that arises when the two masses are clamped in position by a rigid structure. In this case the precision of measurement of the distance between the two objects is limited by quantum fluctuations of the rigid structure. We use a simple dimensional argument to estimate the magnitude of the fluctuations. Let s be the velocity of sound in the structure, let D be the separation between the objects, and let M be the mass of the structure. There will be at least one mode of sound-vibration of the structure which gives a displacement affecting the measurement of D. The mean-square quantum fluctuation amplitude of the displacement in this mode will then be, up to a factor of order unity, at least as large as the zero-point fluctuation,

$$\delta^2 \geq (\hbar D/Ms). \tag{9}$$

The duration of the measurement must be of the order of (D/c), the time it takes the graviton to travel through the apparatus. This duration is shorter than the period (D/s) of the sound-vibration, since s cannot exceed c. Therefore the uncertainty of the measurement is at least equal to the instantaneous vibration-amplitude δ. If the uncertainty is as small as the Planck length (5), then (9) implies

$$(GM/c^2) \geq (c/s)D > D. \tag{10}$$

Again we see that the separation between the two masses is smaller than the Schwarzchild radius of the apparatus, so that the negative gravitational potential of the two masses is greater than Mc^2 and the apparatus will collapse into a black hole. It appears that Nature conspires to forbid any measurement of distance with error smaller than the Planck length. And this prohibition implies that detection of single gravitons with an apparatus resembling LIGO is impossible.

It is clear from Eq. (3) that we have a better chance of detecting a single graviton if we raise the frequency into the optical range and use a different kind of detector. When

the frequency is of the order of 10^{15} Hertz or higher, a single graviton can kick an electron out of an atom, and the electron can be detected by standard methods of atomic or particle physics. We are then dealing with the gravito-electric effect, the gravitational analog of the photo-electric effect which Einstein used in 1905, [Einstein, 1905], to infer the existence of quanta of the electromagnetic field, the quanta which were later called photons. The possibility of detecting individual gravitons in this way depends on two quantities, (a) the cross-section for interaction of a graviton with an atom, and (b) the intensity of possible natural or artificial sources of high-frequency gravitons. Most of this talk will be concerned with estimating these two quantities.

4. Graviton Detectors

The simplest kind of graviton detector is an electron in an atom, which we may approximate by considering the electron to be bound in a fixed non-relativistic potential $V(r)$. We choose coordinate axes so that the z-axis is the direction of motion of a graviton. There are then two orthogonal modes of linear polarization for the graviton, one with the wave-amplitude proportional to xy, and the other with the amplitude proportional to $(x^2 - y^2)$. We choose the x and y-axes so that they make angles of 45 degrees to the plane of polarization of the graviton. Then the matrix element for the electron to absorb the graviton and move from its ground state a to another state b is proportional to the mass-quadrupole component,

$$D_{ab} = m \int \psi_b^* xy \psi_a d\tau, \tag{11}$$

where m is the electron mass. Eq. (11) is the quadrupole approximation, which is valid so long as the wave-length of the graviton is large compared with the size of the atom. The total cross-section for absorption of the graviton by the electron is

$$\sigma(\omega) = (4\pi^2 G\omega^3/c^3) \sum_b |D_{ab}|^2 \delta(E_b - E_a - h\omega), \tag{12}$$

where E_a and E_b are the energies of the initial and final states. It is convenient to consider a logarithmic average of the cross-section over all frequencies ω,

$$S_a = \int \sigma(\omega) d\omega/\omega. \tag{13}$$

Integration of (12) gives the sum-rule

$$S_a = 4\pi^2 L_p^2 Q, \tag{14}$$

where the Planck length L_p is given by (4), and

$$Q = \int |(x\partial/\partial y + y\partial/\partial x)\psi_a|^2 d\tau \tag{15}$$

is a numerical factor of order unity. It is remarkable that the average cross-section (14) is independent of the electron mass and of the nuclear charge. The same formula (14) holds for the absorption of a graviton by a neutron or proton bound in a nuclear potential.

For simplicity we assume that the electron is in an s-state with a wave-function $f(r)$ which is a function of distance r from the nucleus. Then (15) becomes

$$Q = (\int r^4 [f'(r)]^2 dr)/(3 \int r^2 [f(r)]^2 dr). \tag{15}$$

The inequality

$$\int r^4 [f' + (3/2r)f]^2 dr > 0 \tag{16}$$

implies that for any $f(r)$

$$Q > 3/4. \tag{17}$$

On the other hand, if

$$f(r) = r^{-n} \exp(-r/R), \tag{18}$$

then

$$Q = 1 - (n/6). \tag{19}$$

From (17) and (19) it appears that for any tightly-bound s-state Q will be close to unity. The cross-section for absorption of a graviton by any kind of particle will be of the same magnitude

$$4\pi^2 L_p^2 = 4\pi^2 G\hbar/c^3 = 8 \times 10^{-65} \text{ cm}^2, \tag{20}$$

spread over a range of graviton energies extending from the binding-energy of the particle to a few times the binding-energy. For any macroscopic detector composed of ordinary matter, the absorption cross-section will be of the order of 10^{-41} square centimeters per gram.

5. Thermal Graviton Generators

We have a splendid natural generator of thermal gravitons with energies in the kilovolt range, producing far more gravitons than any artificial source is likely to generate. It is called the sun. Stephen Weinberg long ago calculated [Weinberg, 1965] the graviton luminosity of the sun, caused by gravitational bremsstrahlung in collisions of electrons and ions in the sun's core. A later calculation [Gould, 1985] corrected a mistake in Weinberg's paper but does not substantially change the result. For an electron-ion collision with energy E, the differential cross-section $p(\omega)$ for producing a graviton of energy $\hbar\omega$ is divergent at low energies, so that the total cross-section has no meaning. The physically meaningful quantity is the integral of the differential cross-section multiplied by the energy of the graviton,

$$\int p(\omega)\hbar\omega d\omega = (320/9)Z^2\alpha^2 L_p^2 E, \tag{21}$$

where α is the electromagnetic fine-structure constant and Z is the charge of the ion. Including a similar contribution from electron-electron collisions, (21) gives a total graviton luminosity of the sun

$$L_g = 79 \text{ Megawatts}, \tag{22}$$

or about 10^{24} gravitons per second with energy in the kilovolt range. This gives a flux at the earth of

$$F_g = 4 \times 10^{-4} \text{ gravitons per cm}^2 \text{ per second.} \tag{23}$$

If we imagine the whole mass of the earth to be used as a graviton detector, with the cross-section (20) per electron and the flux (23), the counting-rate is 2.4×10^{-17} per second. If the experiment continues for the life-time of the sun, which is 5 billion years, the expected total number of gravitons detected will be 4. The experiment barely succeeds, but in principle it can detect gravitons.

According to [Gould, 1985], there exist in the universe sources of thermal gravitons which are stronger than the sun, namely hot white dwarfs at the beginning of their lives, and hot neutron stars. Gould estimates the graviton luminosities of a typical white dwarf and a typical neutron star to be respectively 10^4 and 10^{10} times solar. Their luminosities are roughly proportional to their central densities. But the life-times during which the stars remain hot are shorter than the life-time of the sun, being of the order of tens of millions of years for the white dwarf and tens of thousands of years for the neutron star. The life-time output of gravitons will therefore be respectively 100 and 10^5 times solar. To stretch the theoretical possibilities of detection to the limit, we may suppose the detector to have mass equal to the sun and to be orbiting around the source of gravitons at a distance of 0.01 astronomical unit with an orbital period of 8 hours. Then the expected number of gravitons detected will be of the order of 10^{13} for the white dwarf and 10^{16} for the neutron star. The detection rate is roughly one per minute for the white dwarf and 3×10^4 per second for the neutron star. The conclusion of this calculation is that graviton detection is in principle possible, if we disregard the problem of discriminating the graviton signal from background noise.

The most important source of background noise is probably the neutrinos emitted by the sun or the white dwarf or the neutron star as the case may be. These neutrinos can mimic graviton absorption events by ejecting electrons from atoms as a result of neutrino-electron scattering. The neutrinos have higher energy than the gravitons, but only a small fraction of

the neutrino energy may be transferred to the electron. From the sun, about 10^{14} neutrinos are emitted for each graviton, and the cross-section for neutrino-electron scattering is about 10^{20} times the cross-section for graviton absorption, [see Fukugita and Yanagida, 2003]. Therefore there will be about 10^{34} neutrino background events for each graviton absorption event.

For white-dwarfs and neutron-stars the ratio of background to signal is even larger, since neutrino production and scattering cross-sections increase with temperature more rapidly than graviton production and absorption cross-sections. Without performing detailed calculations, we can assert that for all thermal sources of gravitons the ratio of neutrino background to graviton signal will be of the order of 10^{34} or greater. In all cases, the total number of detected graviton events is vastly smaller than the square-root of the number of background events. The graviton signal will be swamped by the statistical scatter of the background noise.

Before jumping to conclusions about the detectability of gravitons, we must explore possible ways in which the neutrino background events might be excluded. The first possible way is to surround the detector with a shield thick enough to stop neutrinos but let gravitons pass. If the shield is made of matter of ordinary density, its thickness must be of the order 10^{10} kilometers, and its mass is so large that it will collapse into a black hole. The second possible way is to surround the graviton detector with neutrino detectors in anti-coincidence, to catch the outgoing neutrino after each scattering event. This way fails for the same reason as the shield. The neutrino detectors would need to be at least as massive as the shield. The third possible way is to build a shield or a set of anti-coincidence detectors out of some mythical material with super-high density. The known laws of physics give us no clue as to how this might be done. We conclude that, if we are using known materials and known physical processes, detection of thermal gravitons appears to be impossible.

6. Non-thermal Gravitons

It is possible to imagine various ways in which energetic objects such as pulsars may emit non-thermal gravitons of high energy. One such way is a process first identified by

[Gertsenshtein, 1961], the coherent mixing of photon and graviton states in the presence of an extended classical magnetic field. The graviton emission from various celestial objects resulting from the Gertsenshtein process was calculated by [Papini and Valluri, 1989]. Some interestingly high graviton luminosities were predicted.

The Gertsenshtein process results from the interaction energy

$$(8\pi G/c^4)h_{ij}T_{ij}, \tag{24}$$

between the gravitational field h_{ij} and the energy-momentum tensor T_{ij} of the electromagnetic field. This interaction expresses the fact that electromagnetic fields have weight, just like other forms of energy. Now suppose that h_{ij} is the field of a graviton traveling in the z direction and

$$T_{ij} = (1/4\pi)(B_i + b_i)(B_j + b_j), \tag{25}$$

is the energy-momentum of the photon magnetic field b_i superimposed on a fixed classical magnetic field B_i. Then the interaction (24) contains the term

$$I = (4G/c^4)h_{xy}B_x b_y, \tag{26}$$

bilinear in the graviton and photon fields. The effect of this bilinear term is to mix the photon and graviton fields, so that a particle that is created as a photon may be transformed into a graviton and vice versa. There is an oscillation between graviton and photon states, just like the oscillation between neutrino states that causes neutrinos to change their flavors while traveling between the sun and the earth. If a photon travels a distance D though a uniform transverse magnetic field B, it will emerge as a graviton with probability

$$P = \sin^2(G^{1/2}BD/2c^2) = \sin^2(B/L), \tag{27}$$

with the mixing-length

$$L = (2c^2/G^{1/2}B) \tag{28}$$

independent of wave-length. In all practical situations, D will be small compared with L, so that

$$P = (GB^2D^2/4c^4). \tag{29}$$

The quadratic dependence of P on D makes this process interesting as a possible astrophysical source of gravitons. The numerical value of L according to (28) is roughly

$$L = (10^{25}/B), \tag{30}$$

when L is measured in centimeters and B in Gauss.

We may also consider the Gertsenshtein process as the basis of a graviton detector consisting of a hollow pipe of length D filled with a transverse magnetic field B. The tube must be accurately pointed at a putative source of gravitons in the sky. At the far end of the tube is a shield to block incident photons, and at the near end is a detector of photons resulting from the conversion of gravitons on their way through the tube. If D is one astronomical unit (10^{13} cm), then (29) gives

$$P = 10^{-24}B^2. \tag{31}$$

The field B must be very strong to obtain a reasonable rate of conversion of gravitons to photons. A detector with the same design has been used in a real experiment to detect axions that might be created by thermal processes in the core of the sun [Zioutas et al., 2005]. The axion field is supposed to interact with the electromagnetic field with an interaction energy similar to (26), but with a much larger coupling constant. The experimenters at CERN in Switzerland are using a surplus magnet from the Large Hadron Collider project as an axion-detector, pointing it at the sun and looking for kilovolt photons resulting from conversion of axions into photons. The length of the magnet is 9 meters and the magnetic field is 9×10^4 Gauss. They have not yet detected any axions.

The Gertsenshtein process does not require the classical magnetic field to be uniform. For a non-uniform field, the conversion of photons to gravitons still occurs with probability

given by (27), if we replace the product BD by the integral of the transverse component of B along the trajectory of the photons. Likewise, the conversion will not be disturbed by a background gravitational field, even when the field is strong enough to curve the photon trajectory, because the gravitational field acts in the same way on photons and gravitons. In a curved space-time, the photons and the gravitons follow the same geodesic paths, and the photon and graviton waves remain coherent.

7. Non-linear Electrodynamics

However, there is an important disturbing factor which was neglected in previous discussions of the Gertsenshtein process. The disturbing factor is the non-linearity of the electromagnetic field caused by quantum fluctuations of electron-positron pairs in the vacuum, [Euler and Heisenberg, 1936; Wentzel, 1943]. The fourth-order term in the electromagnetic field energy density is [Wentzel, 1943, page 190],

$$(\alpha/360\pi^2 H_c^2)[(E^2 - H^2)^2 + 7(E.H)^2], \tag{32}$$

where α is the fine-structure constant and

$$H_c = (m^2c^3/e\hbar) = 5.10^{13}\text{Gauss} \tag{33}$$

is the critical magnetic field at which electron-positron pair fluctuations become noticeable.

When the field in (32) is divided into classical and photon components as in (25), there is a term quadratic in both the classical and photon fields,

$$(\alpha/360\pi^2 H_c^2)(4(B.b)^2 + 7(B.e)^2), \tag{34}$$

where b and e are the magnetic and electric fields of the photon. From (34) it follows that the photon velocity v is not equal to c but is reduced by a fraction

$$g = 1 - (v/c) = (k\alpha B^2/360\pi^2 H_c^2). \tag{35}$$

The coefficient k is equal to 4 or 7 for a photon polarized with its magnetic field or its electric field parallel to B. We consider the case $k = 4$, since that case is more favorable to the Gertsenshtein process. Since the graviton field is not affected by the non-linear electromagnetic interaction (32), the graviton velocity is precisely c, and the photon and graviton waves will lose coherence after traveling for a distance

$$L_c = (c/g\omega) = (90\pi^2 c H_c^2/\alpha B^2 \omega) = (10^{43}/B^2\omega). \tag{36}$$

If the propagation distance D is larger than L_c, the Gertsenshtein process fails and the formula (29) for the photon-graviton conversion probability is incorrect. A necessary condition for the Gertsenshtein process to operate is

$$DB^2\omega \leq 10^{43}. \tag{37}$$

Furthermore, even when the Gertsenshtein process is operating, the probability of photon-graviton conversion according to (29) and (37) is

$$P \leq (10^{36}/B^2\omega^2). \tag{38}$$

We are interested in detecting astrophysical sources of gravitons with energies up to 100 kilovolts, which means frequencies up to 10^{20}. With $\omega = 10^{20}$, (37) and (38) become

$$D \leq (10^{23}/B^2), \quad P \leq (10^{-4}/B^2). \tag{39}$$

We consider two situations in which (39) has important consequences. First, with typical values for the magnetic field and linear dimension of a pulsar, $B = 10^{12}$ and $D = 10^6$, (39) shows that the Gertsenshtein process fails by a wide margin. The calculations of the graviton luminosity of pulsars in [Papini and Valluri, 1989] assume that the Gertsenshtein process is producing high-energy gravitons. These calculations, and the high luminosities that they predict, are therefore incorrect. Second, in the hollow pipe graviton detector which we considered earlier, (39) shows that the Gertsenshtein process can operate with a modest

field, $B = 10^5$ Gauss, and a pipe length $D = 10^{13}$ cm, but the probability of detection of each graviton traveling through the pipe is only 10^{-14}. If the field is made stronger, the length of the pipe must be shorter according to (39), and the probability of detecting a graviton becomes even smaller.

8. Conclusions

Many papers have been published, for example [Eppley and Hannah, 1977; Page and Geilker, 1981], claiming to demonstrate that the gravitational field must be quantized. What these papers demonstrate is that a particular theory with a classical gravitational field interacting with quantum-mechanical matter is inconsistent. Page and Geilker assume that the classical gravitational field is generated by the expectation value of the energy-momentum tensor of the matter in whichever quantum state the matter happens to be. They performed an ingenious experiment to verify that this assumption in fact gives the wrong answer for a measurement of the gravitational field in a real situation.

In this paper I am not advocating any particular theory of a classical gravitational field existing in an otherwise quantum-mechanical world. I am raising two separate questions. I am asking whether either one of two theoretical hypotheses may be experimentally testable. One hypothesis is that gravity is a quantum field and gravitons exist. A second hypothesis is that the gravitational field is a statistical concept like entropy or temperature, only defined for gravitational effects of matter in bulk and not for effects of individual elementary particles. If the second hypothesis is true, then the gravitational field is not a local field like the electromagnetic field. The second hypothesis implies that the gravitational field at a point in space-time does not exist, either as a classical or as a quantum field.

Now I assert that both of the two hypotheses may or may not be experimentally testable. Analysis of the properties of graviton-detectors, following the methods of this paper, might be able to throw light on both hypotheses. Three outcomes are logically possible. If a graviton detector is possible and succeeds in detecting gravitons, then the first hypothesis is true and the second is false. If a graviton detector is possible and fails to detect gravitons, then the first hypothesis is false and the second is open. If a graviton detector is in principle

impossible, then both hypotheses remain open. Even if their existence is not experimentally testable, gravitons may still exist.

The conclusion of our analysis is that we are still a long way from settling the question whether gravitons exist. But the question whether gravitons are in principle detectable is also interesting and may be easier to decide.

References

N. Bohr and L. Rosenfeld, 1933. Kgl. Danske Vidensk. Selsk., Mat.-Fys. Med. 12, 8. The English translation by Bryce DeWitt, "On the Question of the Measurability of the Electromagnetic Field Strengths", preserves the Germanic style of the original.

A. Einstein, 1905. "Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt", Ann. der Phys. 17, 132-148.

K. Eppley and E. Hannah, 1977. "On the Logical Necessity of Quantizing the Gravitational Field", Found. Phys. 7, 51-68.

H. Euler and W. Heisenberg, 1936. "Folgerungen aus der Diracschen Theorie der Positrons", Zeits. für Phys., 98, 714-732.

M. Fukugita and T. Yanagida, 2003. "Physics of Neutrinos and Applications to Astrophysics", [Berlin, Springer Verlag].

M. E. Gertsenshtein, 1961. "Wave Resonance of Light and Gravitational Waves", JETP, 41, 113-114, English translation in Soviet Physics JETP, 14, 84-85 (1962).

R. J. Gould, 1985. "The Graviton Luminosity of the Sun and other Stars", Astrophys. Journal, 288, 789-794.

L. D. Landau and E. M. Lifshitz, 1975. "The Classical Theory of Fields, Course of Theoretical Physics, volume 2", fourth revised English edition, translated by Morton Hamermesh, [Oxford, Elsevier].

D. N. Page and C. D. Geilker, 1981. "Indirect Evidence for Quantum Gravity", Phys. Rev. Letters, 47, 979-982.

G. Papini and S.-R. Valluri, 1989. "Photoproduction of high-frequency gravitational radiation by galactic and extragalactic sources", Astronomy and Astrophysics, 208, 345-350, and earlier papers there cited.

S. Weinberg, 1965. "Infrared Photons and Gravitons", Phys. Rev. B140, 516-524.

J. M. Weisberg and J. H. Taylor, 2005. "The Relativistic Binary Pulsar B 1913+16: Thirty years of observations and analysis", in F. Rasio and L. H. Stairs, ed., Astronomical Society of the Pacific Conference Series 328 [Aspen, Colorado, July 2005].

G. Wentzel, 1943. "Einführung in die Quantentheorie der Wellenfelder", [Wien, Franz Deuticke], reprinted by Edwards Brothers, Ann Arbor, 1946.

K. Zioutas et al., 2005. "First results from the CERN Axion Solar Telescope (CAST)", preprint hep-ex/0411033v2. For this reference I am indebted to E. Witten.

A Model for a Gravity Wave at the Quantum Level

Karl V. Thompson

May 7, 2012

Abstract

A classical model that meets the requirements of the Stochastic interpretation of Quantum Mechanics as laid down by D. Bohm and J.P. Vigier is proposed for further study. It appears to explain the wave/particle duality of both photons and electrons. It suggests that a steady state solution is applicable to DeBroglie waves. It yields an explanation for both the nature and source mechanism of the gravitational wave, both on the quantum and astronomical scales. A computation for the gravitational wave of the electron at rest in the earth's gravitational field is given. It suggests a direct link between gravity and electrodynamics. Both elastic and inelastic collisions between photons and electrons are discussed as well as the mechanism of the influence of a magnetic field on an electron. This model does not conflict unduly with the existing work by Einstein, Maxwell, or DeBroglie, nor with any experimental evidence the author is aware of.

A. THE MODEL

The generalized model to be investigated has been suggested and discussed many times in the past by DeBroglie, Dirac, Bohm, Vigier, and others. It is a binary model consisting of a wave that acts as an oscillator and propagates a second wave thru an aether. Unfortunately a good understanding of the oscillator has kept the model from reaching maturity. This following development of a model of this sort is sorely needed for examination by the physics community because it opens up new possibilities for investigation and leads directly to a model showing that a single wave may act as the source of both the gravitational and electromagnetic fields simultaneously, as well as furnish a detailed explanation of wave/particle duality for the electromagnetic wave and the electron.

It differs from previous efforts in four ways. First, the oscillator is proposed to be a wave that always travels with a constant wave momentum[1] under the influence of gravity while the generated wave is proposed to be a wave that travels with a varying wave momentum. The oscillator is only affected by gravity at the time of emission[1]. Secondly, there is no need to propose an aether as a separate entity. The generated waves, collectively from all the oscillating systems throughout the universe, form the needed aether. Thirdly, a set of boundary conditions have been worked out that satisfy existing experimental evidence and make new predictions. Fourth, the model applies equally well to the electron as to the electromagnetic wave.

B. OUTLINE OF DEVELOPMENT

First, the author wants to make it clear that this entire article is only a model and it only attempts to justify qualitatively the effects of the gravitational effect derived from this model. Such justifications and proofs, outside of the gravitational effect, will be left to others to verify or discredit.

Assumptions will be made early in the development and all of these assumptions will seemingly become justified by the results of the predictions of the model being in agreement quantitatively with all experimental observations, and with Einstein's General Theory, and in qualitative agreement with Maxwell's Equations. Despite a difference in form, the model will agree quantitatively with DeBroglie's conditions for matter waves. The energy considerations of the model will differ radically with those of the standard model of the electromagnetic wave.

Classified in terms of wave momentum, only two types of waves can exist. They are waves that travel with wave momentum constant, and waves that travel with their wave momentum not constant. If everything is believed to have a wave nature, then following this reasoning everything could have at its base a model no more complicated than a binary wave system if you ignore geometry.

In setting up the initial model we will seek to divest ourselves of the problem of describing observer dominated events and initially develop, in Euclidian space, the physical properties of a modified electromagnetic wave, having a mass, that always travels through changes in gravitational potential with a wave momentum that is constant. We want to see just what such a wave's characteristics will be.

It should be noted that in section J. we will address the objection, held by so many, to assignment of a mass greater than $hf/C^2$to the electromagnetic wave.

Einstein (1911)[2] had already reasoned, that the electromagnetic wave must travel, under the gravitational influence in Euclidian space, with a constant wave momentum and that this would cause light to bend as it traveled around the sun. Subsequently, observations [3] by Campbell and Trumpler in 1923 verified this prediction in a qualitative manner based on this reasoning. Quantitatively, their measurement was twice the arc predicted by the 1911 article, however the General Theory , in 1916, corrected the 1911 prediction and agreed with Campbell and Trumpler. Therefore it seems that the assigning of a constant wave momentum to an electromagnetic wave is neither new nor unreasonable. Light does exhibit a wave particle duality, and we normally associate a mass with a particle. The assignment of a mass to the electromagnetic wave allows investigation of the gen-

eralized case.

We will develop the relativistic equations describing such a wave and consider them against experiment. There will be a discrepancy in energy at emission when considering the measured shift in wavelength between photons produced in different gravitational intensities. We will work out the boundary conditions for our wave to satisfy the experimental results, and in so doing will conclude that there had to be a three way energy transfer at emission, and that part of the energy appeared to be transferred from the gravitational field to the modified electromagnetic wave.

It will next be postulated that the modified electromagnetic wave must be more complicated. We will then conclude that our model is incomplete and will then propose a new model that consists of two components. We will postulate that a photon wave, acting as an oscillator, coupled to an associated matter wave exists. We will proceed to develop the boundary conditions for this model. We will then postulate that this model can exist in two geometries, one for the electron and one for the electromagnetic wave. Using the results previously developed, we next derive a steady state solution for matter waves in the case of the electron. This model strongly suggests that gravitational waves are components of DeBroglie waves[4] in this steady state.

We will numerically compute the properties for an electron in this state and break the matter wave into two components thereby uncovering how conservation of energy can work at its most basic level. We will then compute and find that this model does indeed match DeBroglie's calculations on a quantitative basis. This model conforms exactly to the precepts of the General Theory and confirms some of Einstein's opinions on the base nature of matter and the existence of an aether as a possibility. We will find that there is a remarkable similarity of this DeBroglie wave's geometry to the characteristics of the electric and magnetic fields and this will be discussed.

We should note here that such a model has been suggested before. A good review of efforts along these lines is given in a discussion by L.V.Chebotarev [5], "The Bohm-Vigier Approach in Quantum Mechanics". Indeed, we already know that a binary system exists at the base level from experiment[6] showing that an electron in motion has a wave associated with it.

Ç. DEFINITIONS and TERMINOLOGY

Before beginning the development of the model we need to make definitions to make the reading of the development easier.

1. The first definition is that, since this entire presentation is basically about gravitational effects, the concept of a force opposing, or in the direction of motion, is really a statement of the gravitational intensity decreasing or increasing respectively. Force is taken to be a function of the gravitational intensity Φ, i. e., $F = f(\Phi)$. Force increases when Φ increases. Early in the development the concept of force is replaced with change in gravitational intensity. Actually later in the development we will show that force is a function of both energy and geometry and that Newton"s "Law of Universal Gravitation" does not hold between systems of different DeBroglie wave geometry.

2. The second definition is that when we refer to a photon we are referring to the standard model of the photon without any electromagnetic components and modified by the assigning of a mass to it.

3. Any time we refer to a photon without mass we will call it "The standard model".

4. Anytime we refer to the standard model with its electromagnetic attributes we will use the term "electromagnetic wave".

5. The fifth definition is that when we we refer to a wave of "constant wave momentum" we are referring to the model described above in the second definition.

6. The sixth definition is that when we refer to a wave of "non-constant wave momentum" we are speaking of a DeBroglie wave.

7.When we speak of the induction of energy between the DeBroglie wave of our model and other DeBroglie waves we are referring to constructive/destructive wave interference that alters the energy distribution within our model's DeBroglie wave.

8. The last definition is that all results are obtained using the boundary condition that when Φ approaches zero that $hfo = moC_0^2$. This allows one to eliminate observer dominated results by referencing all discussion to this boundary condition.

D. THE PHOTON IN A GRAVITATIONAL FIELD

We will initially be interested in deriving all of the characteristics of a electromagnetic wave with mass under the influence of gravity, however we will only be interested initially in the energy of this wave. Since, under standard theory no energy is attributed to the electromagnetic characteristics of the wave we will ignore them. We will call the wave itself a photon and its wave energy will be defined by Planck's hypothesis

$$E_\lambda = nhf \tag{1}$$

, and its wave characteristics will obey the wave equation,

$$C_0 = f\lambda \tag{2}$$

. Since we are assigning an arbitrary mass to the wave its mass energy must obey Einstein's equation,

$$E = mC_0{}^2 \tag{3}$$

. We are initially interested in seeing what effect gravity will have on this photon system and we will assume a boundary condition that will be a function of the gravitational intensity, Φ, of the form

$$nhf_0 = m_0C_0{}^2 \tag{4}$$

where f_0 and m_0 are the values of the frequency and the mass when $\Phi = 0$. We will set n=1 for the remainder of this article.

Because we are building a model and we want this model to obey general relativity we will take the liberty of introducing a negative sign in equation (10) to insure that the frequency decreases as the gravitational intensity Φ increases[7][1].We will also introduce a negative sign to Newton's Law of Gravitation to assure that Φ represents the gravitational intensity, not the gravitational potential energy.

Consider first a quanta of energy, E, at some great distance from a large mass, say the Sun. Let it have a mass, m, defined according to Equation (3), $E = mC_0^2$ Let it be free from any influence except the gravitational field of the Sun. It is "at rest". Now let it fall solely under the gravitational influence on its mass to some distance r from the Sun.

Since for a mass,

$$F = C_0^2\frac{dm}{dr} , \tag{5}$$

and

$$F = -\frac{GMm}{r^2} , \tag{6}$$

we can then say

$$C_0^2\frac{dm}{dr} = -\frac{GMm}{r^2} ,$$

and for a quanta of energy of mass m falling into a gravitational field from infinity we would say

$$\int_{m_0}^{m}\frac{dm}{m} = -G\frac{M}{C_0^2}\int_{\infty}^{r}\frac{dr}{r^2} , \tag{7}$$

where m varies from m_0 at infinity to m at some distance r from the center of the source of the gravitational intensity. Defining the gravitational intensity Φ, at r as:

$$\Phi = \frac{GM}{r} \; and \; d\Phi = -\frac{GM}{r^2}\,dr ,$$

and changing variables, we have

$$\int_{m_0}^{m_\Phi}\frac{dm}{m} = \frac{1}{C_0^2}\int_0^{\Phi} d\Phi , \tag{8}$$

and integrating with respect to Φ we get

$$\ln\frac{m_\Phi}{m_0} = \frac{\Phi}{C_0^2} ,$$

or

$$m_\Phi \;=\; m_0 e^{\frac{\Phi}{C_0^2}} \;. \tag{9}$$

Now assume that the quanta of energy is in the form of a wave that obeys equation (2), the wave equation. We will assume that this is a photon. The frequency of the photon is linked to the mass of the photon by the boundary condition of equation (4)which describes a stable condition at $\Phi \;=\; 0$.

We then have,

$$-h\frac{df}{dr} \;=\; -\,G\frac{Mm}{r^2} \;. \tag{10}$$

But since we saw that

$$m_\Phi \;=\; m_0 e^{\frac{\Phi}{C_0^2}} \; or \; m_\Phi \;=\; m_0 e^{\frac{GM}{rC_0^2}} \;,$$

using the substitutions and noting that Φ_0 was taken to be zero,

$$\Phi \;=\; \frac{GM}{r} \; and \; d\Phi \;=\; -\,\frac{GM}{r^2}dr$$

we can obtain

$$-h\int_{f_0}^{f} df \;=\; m_0 C_0^2 \int_{\Phi_0}^{\Phi} \frac{1}{C_0^2}\, e^{\frac{\Phi}{C_0^2}}\, d\Phi \;. \tag{11}$$

This can be shown to yield when , Φ_0 equals zero,

$$f_\Phi \;=\; f_0 e^{\frac{-\Phi}{C_0^2}} \;, \tag{12}$$

for the frequency in terms of the photon's initial frequency f_0, and the potential difference, Φ. Our photon obeys the wave equation $C \;=\; \lambda f$ at emission, and wavelength was assumed to be thereafter a constant for our model. Summarizing, the general equations describing the internal attributes of a photon's wave in a potential Φ are:

$$m_\Phi \;=\; m_0 e^{\frac{\Phi}{C_0^2}} \;. \tag{9}$$

$$f_\Phi \;=\; f_0 e^{\frac{-\Phi}{C_0^2}} \;, \tag{12}$$

The approximate forms are:

$$m_\Phi \;=\; m_0\,[1+\frac{\Phi)}{C_0^2}] \tag{9a}$$

and,

$$f_\Phi \;=\; f_0\,[1-\frac{\Phi)}{C_0^2}] \tag{12a}$$

where C_0 is the speed of the photon's internal wave at emission and is defined by equation (19).

As a matter of convenience, we will define m_0 and f_0 by

$$hf_0 \;=\; m_0 C_0^2 \tag{13}$$

where m_0,f_0 , and C_0^2 are the values obtained when a photon is emitted in a zero gravitational intensity ($\Phi \;=\; 0$), our boundary condition .

These same results were predicted in 1916 in Einstein's "The Foundation of the General Theory of Relativity"[7].

Calculation of the ratio of the periods of like photons on the Sun and the Earth gives,

$$\frac{Period\;(sun)}{Period\;(earth)} \;=\; 1.0000021 \;,$$

which is in very good agreement with the predictions of the General Theory of Relativity according to Lieber[9].

The photon's wave momentum was assumed always constant and , if so, would be given by

$$p_\lambda \;=\; h\frac{f_\Phi}{C_0} \;=\; \frac{h}{\lambda} \;. \tag{14}$$

If this is so then we deduce that the frequency[1] of the wave can not change as the wave travels thru changing gravitational intensities since C_0 and λ are constants. This implies, that for our model, the wave energy does not change as the photon(the oscillator) travels thru varying gravitational intensities. Therefore, it must be concluded that equation (12) applies only at emission. However we note that there are no such limits placed on the mass of our model. It will obey equation(9) as Φ changes.

At this point in this paper no discussion or assumptions are made about what happens in elastic

or inelastic collisions or interactions with an electromagnetic field. This is really outside of the scope of this article and will be covered later in the summation.

It can be shown that the difference, at emission, between the wave energy, calculated using (12) with (1), and the mass energy, calculated using (9) with (3), is twice the difference between the rest mass energy and the wave energy. The mass energy represents the total energy our photon model has according to $E = mC_0{}^2$ .This result will be shown later in this discussion to have a significant physical meaning and this difference is probably the most significant step toward the results of the rest of this article.

A correspondence between the model and General Theory seems to be emerging. The photon's frequency seems to be equivalent to the rate at which a relativistic unit clock precesses, and the increase in the photon's mass in Euclidian space seems to be equivalent to a contraction of a relativistic unit measuring rod[8] lying along the radius between the gravitational source and the photon as described by the General Theory. This can be confirmed by comparing Einstein's g_{11} and g_{44} tensors of the second rank with the case of the photon in a gravitational field.

Note that f_Φ is a spacial variable related only to the gravitational intensity, while m_Φ is also a spacial variable but is related both to the gravitational intensity and to the total energy of the photon model through both (3) and (13). This means that we can express f_Φ and m_Φ in terms of (12) and of (9) taking $\Phi_0 = 0$. We used this property of the mass in setting up the differential of (11).

E. THE EMISSION OF A PHOTON

We are now in a position to discuss the photon attributes in more detail. Equation (12) tells us in what manner our photon model's frequency, at emission, adjusts to it's potential environment. It decreases as Φ increases. Equation (9) shows that the mass, $\frac{totalenergy}{C_0{}^2}$ adjusts in a similar manner at emission ,however, for our model, there were no limits placed on the mass so it is assumed that, not only at emission, the mass also changes, according to equation (9), as it moves thru varying changes in Φ.

We have already concluded that for equation (12) the assumption of constant wave momentum precludes any changes in frequency as the photon moves from one Φ to a different Φ .

There were results first reported by Evershed and Adams[10] that showed that the wavelengths of spectral lines arriving on Earth from Sirius, Procyon, and Arcturus were longer than the wavelengths of spectral lines originating on the Earth. Obviously, something happens at the birth of a photon that is a function of Φ but is not adequately described by our development of equations (9) and (12) to date. Our understanding of the photon at the boundary condition of birth is incomplete.

We will assume that the wavelength is a function of the potential environment, Φ, at birth, and that this wavelength is given by

$$\lambda_{\Phi'} = \frac{C_0}{f_{\Phi'}} \ , \ \Phi' \text{ being the potential of birth.} \quad (15)$$

If an amount of energy E is transferred to the photon at birth, it must be defined by Einstein's mass energy relationship (3),

$$E = mC_0^2 \ ,$$

which gives an expression for the boundary condition of birth of the form

$$\lambda_{\Phi'} = \frac{1}{f_{\Phi'}} (\frac{E}{m})^{\frac{1}{2}} \ , \quad (16)$$

where the ratio,$(\frac{E}{m})$, must be a constant independent of the birth potential if the results of the Evershed and Adams experiments are to be satisfied.

Consider what happens if the ratio $(\frac{E}{m})$ is adjusted to the potential of the birth environment.

If, using the wave and total energies as a function of Φ , then:

$$\lambda_{\Phi'} = \frac{1}{f_{\Phi'}} (\frac{hf_{\Phi'}}{m_{\Phi'}})^{\frac{1}{2}} \ , \ f_{\Phi'} = f_0 e^{-\frac{\Phi'}{C_0^2}} \ ,$$

and

$$m_{\Phi'} \;=\; m_0 e^{\frac{\Phi'}{C_0^2}} \;,$$

then

$$\lambda_{\Phi'} \;=\; \frac{1}{f}_0 (\frac{E_0}{m_0})^{\frac{1}{2}} \quad . \tag{17}$$

We find that the wavelength is a constant for all like processes and environments, i.e., the wavelength of a sodium D photon emitted on the sun would be the same as the wavelength of a sodium D photon emitted on the earth. This would cause our model to violate experimental observations[10].

Now look at what happens when the ratio $(\frac{E}{m})$ is a constant. If we define this ratio as $(\frac{E_0}{m_0})$, where m_0 is the mass of a photon that is in a force free environment such that $E_0 \;=\; m_0 c_0^2 \;=\; hf_0$, then

$$\lambda_{\Phi'} \;=\; \frac{1}{f}_{\Phi'} (\frac{E_0}{m_0})^{\frac{1}{2}} \;,$$

and we have

$$\lambda_{\Phi'} \;=\; \frac{e^{\frac{\Phi'}{C_0^2}}}{f_0} (\frac{E_0}{m_0})^{\frac{1}{2}} \;. \tag{18}$$

Then if Φ_1 is greater than Φ_2, $f_{\Phi 1}$ will be less than $f_{\Phi 2}$, and thus the wavelength of a photon born on the sun will be longer than that of a photon emitted on earth, satisfying experimental evidence[10] and General Relativity. If we substitute, into the wave equation, (2), the wavelength λ', and $f_{\Phi'}$ and solve this simultaneously with (19) under the definition of (13), eliminating E_0 we obtain the result that $C' = C_0$.

All photons are born with an initial wave velocity equal to C_0. C_0 is linked to the energy given to the photon at birth according to the relation

$$E_0 \;=\; m_0 C_0^2 \tag{19}$$

, where m_0 is the mass that photon would have in a force free environment at rest, i.e. when Φ approaches zero.

The circumstances of a photon's emission define the frequency(12) and the mass(9) as functions of Φ. The wavelength, defined by the wave equation is $\lambda \;=\; \frac{C_0}{f'_\Phi}$.

Immediately we see from the boundary conditions at emission, that the photon development requires that the creation of a new photon involves a three way energy transfer. The emitting system gives up an amount of energy E_0, the photon wave itself gets energy

$$E_\lambda \;=\; hf_{\Phi'} \;, \tag{20}$$

and an amount of energy dE must be transferred to some other entity, where

$$dE \;=\; E_0 \;-\; hf_{\Phi'} \;=\; h(f_0 \;-\; f_{\Phi'}) \;. \tag{21}$$

Now being perfectly objective about this missing energy, there is an alternative explanation. It could be argued that the energy is transferred back into the emitting system itself, i. e. , the atom. Otherwise it seems that the gravitational field must be the recipient of this energy. It is the third party in this event.

We should note for future discussion that the emission of our photon is the result of an electromagnetic event, i.e. a change in the status of an electrical system. The discussion of the necessity for a three way energy interchange,(21), gives a first indication that our model is incomplete. We note that we have been discussing a wave that travels always with its wave momentum constant. We note also that if we accept the premise that all energy is in the form of waves, that there is only one option available to us to describe the form of the missing energy. It has had to be transferred to a wave that travels with its wave momentum not constant. These waves are generally called DeBroglie waves. Classified in terms of wave momentum only two types of energy can exist.

This implies that the gravitational field is composed of DeBroglie waves and this is a profound conclusion that cannot be reached under the standard model of the photon in such a direct manner if at all. This implication means that the difference in energy between equation (9) and equation (12) using the birth potential, Φ' , gives the exact energy of the DeBroglie wave at emission.

This three way transfer, at emission, can be shown to satisfy conservation of energy over the photon's life cycle. The Evershed and Adams experiments cannot be satisfied for a photon by any other conclusions. This is the first indication that, for a photon to have a mass, a binary model is required. At this point we will assume, following DeBroglie[11], D. Bohm, and J.P Vigier[12][13][14][15] that the mechanism for producing the DeBroglie wave is that the our photon model (wave of constant wave momentum)acts as an oscillator creating a disturbance in the gravitational aether and generates a DeBroglie wave in the aether creating a wave (of non constant wave momentum) having a wave energy of

$$hf_0 \ - \ hf_{\Phi'}$$

, and that, at emission, a similar equal amount of energy is induced from the gravitational aether into the generated DeBroglie wave, thus accounting for the discrepancy in energy between equations (9) and 12) of the amount, $\Delta E \ = \ m_0 C_0^2 \ [e^{\frac{2\Phi'}{C_0^2}} \ - \ 1]$.

Now, an interpretation of the scope of influence of equations (9) and (12) is possible. It is immediately apparent that we are forced to consider a binary model. We will discuss this in the context of energy considerations as this binary model climbs out of the gravitational well of the sun and travels to earth.

If the photon wave (of constant wave momentum), the oscillator, of the our binary model must travel unchanged from sun to earth it is easily seen that it is the Debroglie wave which must lose energy in the amount of the change in potential energy between the sun and the earth, the same amount as said to be lost from the standard photon model due to a decrease in its frequency.

Furthermore, as the system travels, the energy lost from the debroglie wave can be considered lost by the mechanism of a reduction in the energy induced from the aether. This is easily deduced because the oscillator will continue to generate the same amount of energy into the debroglie wave but the level of energy in the aether will decrease as the system moves out of the well so the interaction (induction) will decrease proportionately to the to the decrease in gravitational intensity.

This reasoning of this section furnishes the basis for the discussions in the remainder of this paper.

F. GENERAL PROPERTIES OF WAVES WHEN P IS NOT CONSTANT

In this section we want to digress from our discussion of the photon to discuss a wave with mass that travels with non-constant momentum and increasing energy when a force is applied in the direction of motion. The wave equation for such a wave would be given by

$$v \ = \ \lambda \ \times \ f \ , \tag{22}$$

instead of (2),

$$c \ = \ \lambda \ \times \ f \ .$$

For such a system, if its wave energy is to increase, the force differential describing its frequency would be given by

$$F \ = \ nh\frac{df}{dx} \ , \tag{23}$$

and the force differential describing its mass would be given by,

$$F \ = \ C_0^2 \frac{dm}{dx} \ ,$$

and Newton's hypothesis would apply.

$$F \ = \ \frac{d(mv)}{dt} \ , \tag{24}$$

i.e., force is equivalent to the derivative of the momentum of a system with respect to time.

We solve for mass as a function of velocity. Taking the force to be in the direction of motion we have

$$C_0^2 \frac{dm}{dx} \ = \ m\frac{dv}{dt} \ + \ v\frac{dm}{dt} \ ,$$

which yields

$$C_0^2 dm \ = \ mvdv \ + \ v^2 dm \ ,$$

or

$$\frac{dm}{m} \ = \ \frac{vdv}{(C_0^2 - v^2)} \ .$$

Letting $u = C_0^2 - v^2$ and $du = -2vdv$, and integrating and substituting for u and du, we have

$$\ln\ m\ |_{m_0}^{m} \ = \ -\frac{1}{2}\ln(C_0^2 - v^2)\ |_{v_1}^{v}\ ,$$

which evaluates to

$$\ln(\frac{m}{m_0}) \ = \ -\frac{1}{2}\ln\ \frac{(C_0^2 - v^2)}{(C_0^2 - v_1^2)}\ ,$$

which then gives

$$m \ = \ m_0[1 \ + \ \frac{1}{C_0^2}(v_1^2 - v^2)]^{-\frac{1}{2}}\ . \qquad (25)$$

Similarly, we can solve equations (23) and (24) simultaneously by eliminating force and we will obtain an expression for frequency as a function of velocity. Taking force to be in the direction of motion we have

$$h\frac{df}{dx} \ = \ m\frac{dv}{dt} \ + \ v\frac{dm}{dt}\ ,$$

or

$$hdf \ = \ mvdv \ + \ v^2dm\ . \qquad (26)$$

using $m = \frac{hf}{C_0^2}$ and $dm = \frac{hdf}{C_0^2}$ we can obtain

$$C_0^2df \ = \ fvdv \ + \ v^2df\ . \qquad (27)$$

Integrating f from f_1 to f, and v from v_1 to v, we get;

$$\frac{f}{f_1} \ = \ [\ (\frac{C_0^2 - v^2}{C_0^2 - v_1^2})\ ]^{-\frac{1}{2}}\ ,$$

or

$$f \ = \ f_1[1 \ + \ \frac{1}{C_0^2}(v_1^2 - v^2)]^{-\frac{1}{2}}\ . \qquad (28)$$

Now (25), when $v_1 = 0$, is the well-known relationship for relativistic mass containing the Lorentz factor. We now have relationships relating the mass and frequency to the motion for waves that travel with momentum not constant.

G. A BINARY SYSTEM

Now we essentially proposed a modification to our model in section E. It was proposed that our photon model was only part of a more complicated system. This extended photon system, we will refer to a "binary system" because it is composed of two interrelated waves of different characteristics. We will propose that exactly the same binary system can have dual geometries. Both geometries are composed of a photon and an associated DeBroglie wave. Both geometries will share exactly the same energy characteristics except for the effect of geometry on the energy distribution in space.

Both of the geometries will consist of a photon(acting as an oscillator) that generates an attached DeBroglie wave. The geometry of the oscillator can exist in either a linear(open geometry) or in a circular(closed geometry) state. The electron will have closed geometry and the photon will have open geometry. Closed geometry will imply that the photon(the oscillator)of our binary model is in a closed circular orbit. Open geometry will imply that the photon (the oscillator)of our binary model travels in a linear manner. This will be discussed in more detail in a later section.

As mentioned earlier in this article, this proposal is neither new nor radical. The Stochastic, Causal, interpretation of quantum mechanics as put forth by D. Bohm and J.P. Vigier and resulting in the Bohm-Vigier model[5][13][15] describes a similar model. This model consisted of two waves. A linear pilot wave, P wave, that described the particle's drift motion and a non linear soliton wave, S wave, that followed the flow of the P wave provided that the phases of the two waves coincided. The S wave that represented the core of the particle acted as an oscillator.

The proposed dual states of geometry herein described were not considered in the earlier work referenced above. Unfortunately, the state of experimental evidence supporting the General Theory was not as complete as it now is.

This proposal is prompted by the missing energy of equation (21) in section E. , the difference in the total energy and the wave energy as dictated by (9) and (12) respectively, and by the realization that if all energy resides in waves, then classified in terms of wave momentum, a binary system could exist.

Because the mass of (9) indicated total energy in

excess of that lost in (21), it seemed reasonable to assume that the gravitational field, the only other party to these equations, must be contributing energy to our photon system at emission and that because our photon model is proposed as a wave that always travels with its wave momentum constant under the gravitational influence, that the energy of the gravitational field must exist in the form of a different sort of wave. It must be composed of waves that travel with the wave momentum not constant, i. e. , DeBroglie waves. We will proceed to show that this is possible in terms of energies and will calculate the exact gravitational energy and wavelength of the DeBroglie wave of an electron at rest in the laboratory on the earth's surface using the energy for the wave predicted by our model. It will be shown that this predicted wavelength, using the energy for the wave predicted by our model, is exactly that which would be predicted for this model by DeBroglie. Also the established expression for the escape velocity will be derived directly from the model. These two calculations will establish that the model is correct.

H. MATTER WAVES and DEBROGLIE WAVES

We have experimentally observed waves associated with particles, first proposed by Louis DeBroglie[4], and first confirmed by Davisson and Germer[6], that travel with momentum not constant. These waves are called either matter or DeBroglie waves and are seemingly part of the electron system, as well as part of the systems of other particles such as the proton and even the neutron, etc. We will , however confine our efforts to the electron.

DeBroglie waves[4] obey the following conditions:

$$W = \frac{C_0^2}{v} , \tag{29}$$

$$E = h\gamma , \tag{30}$$

$$P = \frac{h}{\lambda} , \tag{31}$$

and the wave equation (2), written as

$$W = \lambda\gamma , \tag{32}$$

where W is the speed of the DeBroglie wave, and v is the group velocity of the wave. Its wavelength is λ , and its frequency is γ. The momentum, P, is equal to the product of its total mass times its group velocity. We see that when v approaches zero, its momentum and energy approach zero, but its wavelength and wave speed approach infinity. Consequently, its frequency approaches zero. We also see that the wavelength of the DeBroglie wave is a variable.

Now it can be shown using the equations of section F., "General Properties of Waves When P is Not Constant", and the DeBroglie conditions, equations (29),(30),(31) and (32) that the DeBroglie wave already exists, even at the boundary condition that the group velocity $v = 0$ in the laboratory. The initial energy in the DeBroglie wave when the system is at rest in the Earth's gravitational intensity, Φ depends on the rest mass of the entire system. This indicates that the system has a DeBroglie wave component even when at rest in the laboratory. It is the energy and wavelength of this DeBroglie component that we will be interested in because it's energy will be the energy of a gravitational wave for the electron in the Earth's Φ.

It should be noted that the definition of rest mass used in this presentation may vary from the conventional definition of rest mass. In section C., equation (4), we effectively defined the rest mass to be that of a zero gravitational environment such that $hf_0 = m_0 C_0^2$. The conventional definition of rest mass , for our model of the electron is, I believe, given by (9),

$$m = m_0 e^{\frac{\Phi}{C_0^2}} . \tag{9}$$

where m is the conventional rest mass in the laboratory and Φ is the gravitational intensity of the laboratory. This is a point that should be carefully examined when using published rest masses in calculations. Equation (25), to be complete should read as,

$$m = m_0 e^{\frac{\Phi}{C_0^2}} [1 + \frac{1}{C_0^2}(v_1^2 - v^2)]^{-\frac{1}{2}} . \tag{25a}$$

when considering the total mass of the electron.

The energy in the matter wave, in the steady state, can be expressed as the difference between the total mass energy of the electron calculated using (9), and the energy in the oscillator component of our model for the electron calculated using (12)

$$E_{db} \;=\; E_{total} \;-\; E_{\lambda}$$

and this yields, for the steady state energy of the matter wave,

$$E_{db} \;=\; m_0 C_0^2 \; [e^{\frac{2\Phi}{C_0^2}} \;-\; 1] \; . \qquad (33)$$

in terms of the gravitational rest mass, m_0.

We have two equations, (9) and (25) for the mass of our model for the electron. Equation (9) gives the mass as a function of Φ and equation (25) gives us the mass as a function of the group velocity, v, of the matter wave. Both are referenced to the boundary condition that when $\Phi \;=\; 0$, $E_{db} \;=\; 0$. You might consider that v represents the velocity of the laboratory when $\Phi \;=\; [\Phi_{earth}]$.

$$m_0 e^{\frac{\Phi}{C_0^2}} \;=\; m_0 [1 \;-\; (\frac{v_g^2}{C_0^2})]^{-\frac{1}{2}} \; . \qquad (34)$$

Now this can be shown to reduce to,

$$v_g \;=\; C_0 \; [1 \;-\; e^{\frac{-2\Phi}{C_0^2}}]^{\frac{1}{2}} \; . \qquad (35)$$

and further reducion will lead to,

$$v_g \;=\; (2\;\Phi)^{\frac{1}{2}} \; . \qquad (35a)$$

The reader will notice that v_g of (35a) is exactly that of the quantity conventionally called the "escape" velocity. However (35) is the relativistic form and agrees with the Newtonian expression (35a) at the lower gravitational intensities while at extremely large intensities it increases asymptotically to a value of C. This in effect predicts that the speed of a particle can never reach the speed of light under the influence of gtavity alone. This is , of course, in complete agreememt with experimental physics.

This yields , to a first order approximation, a group velocity of $v_g \;=\; 11.182 \times (10)^5$ cm/sec when $[\Phi_{earth}]$ is, in cgs units, estimated at $62.496 \;\times\; (10)^{10}$.

What we are primarily interested in at this point is a description of this DeBroglie wave, for the model of closed geometry, i.e., a free electron, in the steady state, i. e. at rest in the frame of the laboratory such that the group velocity relative to the laboratory is zero. The reason, as mentioned previously, that this is of such importance is because it should yield, quantitatively, a description of a gravitational wave.

So far in this article we have been primarily interested in the development of the model. It is time to point out a dichonomy that arises between the macro scale and the quantum scale. We will do this by showing that two electrons can have the same energy in the debroglie wave at the quantum scale but differring velocities on the macro scale. This will serve to illustrate more clearly how the model seperates out macro effects based on quantum characteristics.

Case 1 following is a calculation for an electron , either created in the laboratory and brought to escape velocity, (35a), in the laboratory or captured after falling into the laboratory from empty space without encountering any accelerations except those provided by the Earth's gravitational field. We will use Debroglie's equation (31), $P \;=\; \frac{h}{\lambda}$, to calculate it's debroglie wavelength.

Case 2 following is a calculation for an electron simply created in the laboratory , brought stationary between two capacitor plates, and then released to fall a short distance to the measuring apparatus for a measurement of it's wavelength. One half of the total energy of our debroglie wave given by equation (33) will be used in the calculation of P using equation (36) so that the escape velocity is eliminated from the calculation. We will find that the calculated wavelength using (31), is the same wavelength calculated in case 1. The numerical calculations follow next.

Case 1

Using the published mass, $9.1066 \;\times\; (10)^{-28}$ gm for m_{Φ}, we can compute the momentum

$P = m_\Phi \times v_g$ to be $1018.3 \times (10)^{-24}$ gm cm/sec. By DeBroglie's condition (31), we compute λ to be equal to be on the order of $650 \times (10)^{-8}$ cm or 650 angstroms.

Case 2

Now as a cross check on the above computations we will compute λ using the computed energy of the matter wave as given by (33). We will consider that 1/2 this energy is equivalent to the kinetic energy,T, of our model of the electron. It can be shown that (33) reduces to $E_{db} = 2m_0\ \Phi$, and this is approximately $1.14 \times (10)^{-16}$ ergs. Using

$$P = (2\ m_\Phi\ T)^{\frac{1}{2}}\ , \tag{36}$$

we calculate P to be $1018.3 \times (10)^{-24}$. Using (31) we obtain for λ a wavelength of $650 \times (10)^{-8}$ cm or 650 angstroms. These two estimates of the electron's steady state λ are obviously the same.

The author would like to state that an analysis of the basic model of a binary system in section E. in the discussion of the open geometry binary system climbing out of a gravitational well suggests that only 1/2 of E_{db} as given in (33) contributes to the changes in the motion. The oscillator, staying unchanged after emission would therefore continue to generate a constant energy into the wave. This result of 650 angstroms should be slightly too small because for convenience we used the approximation $m_0 = m_\Phi$, the published value for the mass of the electron.

This is exactly the wavelength obtained using the computation $P = m_\Phi \times v_g$ and the v_g from (35). That computation did not require the use the energy of motion, instead using the gravitational group velocity of the electron that a free electron in the laboratory would have if it had been created in the Φ of the laboratory with no kinetic energy other than that equal to the escape velocity. The difference in computation of λ is resolved, and additionally we see that, indeed, only 1/2 of the energy of (33) is related to motion. This knowledge will be useful in extending the model past this presentation to date.

There is a principle of locality that applies here but we are ignoring it presently for the sake of clarity of presentation.

Furthermore, the reader will note that this division of the DeBroglie wave into two components allows an elegantly simple scheme for the conservation of gravitational energy in a fixed Φ . The oscillator is continually generating an essentially spherical wave of energy $m_0 C_0^2\ [e^{\frac{\Phi}{C_0^2}} - 1]$, into the gravitational aether and at the same time this wave is inducing from the aether an equal amount of energy, while in the Φ of emission, essentially along a wavefront perpendicular to a vector corresponding to the radius between gravitational source and the closed geometry system, into the DeBroglie wave component thus creating a relativistic mass and conserving energy.

We raise the question here as to whether we may have succeeded in defining a gravity wave for the electron on the quantum scale. One might object to this conclusion as being misleading because there seems to be an alternative interpretation to these results. The apparently induced energy may not be stored in the DeBroglie wave, but may be merely the energy due to motion, rather than the energy causing motion. This is a very fine example of the difficulty one experiences in following this model. This induced energy is indeed, the energy of motion, the source of the kinetic energy. However motion is only the effect. The cause is the anisotropic energy distribution of the debroglie wave as discussed in section K, Motion. At this point the results of the closed geometry model seem to fit experimental evidence.

The reader will note that although we earlier proposed an exact description, in the form of closed geometry, for the oscillator, that it matters not if the oscillator is a point oscillator or a closed geometry oscillator. All that really matters is that the oscillator and resultant DeBroglie wave obey the boundary equations so far presented.

We can also express the energy of the total De-

Broglie wave, both that generated by the oscillator and also the component induced from the gravitational field, in terms of the rest frequency of the EM wave as,

$$E_{db} \;=\; hf_0\,[e^{\frac{2\Phi}{C_0^2}} \;-\; 1]\;. \qquad (37)$$

Now let us consider our model to be an electron with rest mass m_0. Let it be at rest so that its DeBroglie component holds only the gravitational tie to the electromagnetic wave.

So the wave energy of the DeBroglie wave, P not constant, using (28) and (30), is given by;

$$E_\lambda \;=\; h\gamma_0[1 \;+\; \frac{1}{C_0^2}(v_0^2 - v^2)]^{-\frac{1}{2}}\;, \qquad (38)$$

and if we set this equal to the total energy of the electromagnetic wave component less the wave energy of the EM component of the electron we have, using (33) and (38),

$$m_0C_0^2\;[e^{\frac{2\Phi}{C_0^2}} \;-\; 1]$$

$$=\; h\gamma_0[1 \;+\; \frac{1}{C_0^2}(v_0^2 - v^2)]^{-\frac{1}{2}}\;, \qquad (39)$$

which gives

$$\gamma_0 \;=\; \frac{m_0C_0^2}{h}\;[\;\frac{e^{-\frac{2\Phi}{C_0^{(2)}}} - 1}{[1 \;+\; \frac{1}{C_0^2}(v_0^2 - v^2)]^{-\frac{1}{2}}}\;]\;, \qquad (40)$$

which finally yields, when the electron is in its steady state in the laboratory, at rest and $v \;=\; v_0$,

$$\gamma_0 \;=\; \frac{m_0C_0^2}{h}\;[e^{\frac{-2\Phi}{C_0^2}} \;-\; 1]\;, \qquad (41)$$

This shows us again that for any v_0 of equation (28), the initial energy in the DeBroglie wave is equal to twice the energy lost by the electromagnetic system. It is a very small amount of energy contributed both by the gravitational environment and the oscillator wave. It should be noted that equations (37) and (41) assume that the electron was created in the Φ of the laboratory. If it were were created, born, in any other Φ then the potential energy difference between the two Φs must betaken into account by adding or subtracting as appropriate this potential energy difference from the energies of (37) and (41).

This is understandable. DeBroglie assumed a model referenced to his laboratory. The binary model is a different model. It is always referenced to a gravitational intensity where Φ approaches zero. At the risk of becoming tedious let us restate this in a different manner. The V_0 of DeBroglie's equations are equivalent, when the General Theory is considered, to the classical Escape Velocity, $v_g \;=\; (2\;\Phi)^{\frac{1}{2}}$. The task is to bring these models together so that we do not contradict experimental evidence. It appears that DeBroglie's model may be a special case of our binary model where everything is referenced to the boundary conditions at $\Phi \;=\; 0$. It should be possible to extend DeBroglie's conditions to encompass the stationary state relative to the boundary conditions of this article. There is no error implied about DeBroglie's conditions. We are looking at the problem from a different perspective. This difference in velocity is so small compared to C that it would have no significance on normal observations. However, when we are looking for such small quantities as the momentum and energy of the electron's gravity wave then this becomes a very important quantity. This is the only place in this article where we have to discuss information from an observer dominated position.

For consistency we should express the Debroglie conditions to the boundary conditions at $\Phi \;=\; 0$. To do this we must modify the DeBroglie expressions (30) and (32), and so the modified DeBroglie conditions are:

$$W \;=\; \frac{C_0^2}{v'}\;, \qquad (42)$$

$$E \;=\; h\gamma_0\;[1 \;+\; \frac{1}{C_0^2}(v_0^2 - v^2)]^{-\frac{1}{2}}\;, \qquad (43)$$

$$P \;=\; \frac{h}{\lambda}\;, \qquad (44)$$

$$\text{and } W \;=\; \lambda\gamma_0\;[1 \;+\; \frac{1}{C_0^2}(v_0^2 - v^2)]^{-\frac{1}{2}}\;, \qquad (45)$$

where γ_0 is given by (40) evaluated at some time t such that $v = v_0$.

Under these modified conditions DeBroglie waves have a steady state solution consistent with the closed binary system. Measurement of the wavelength of the matter waves of the electron should, in practice, yield the same results as are found for DeBroglie waves. Normally the initial energy of the waves is too small to be observed at P equal to 0.

To this extent at least, you can consider that an photon(the oscillator) wave with its associated DeBroglie wave could constitute one of the basic systems of our universe.This would be true whether the photon has either open or closed geometry.

When we make a measurement of a given binary system, such as the measurement of the mass or momentum of an electron in an accelerator, or in the cases of an elastic or inelastic collision, we are measuring its total system as dictated by equations (25a) and (28), not just the energy in it's DeBroglie wave.

We see now, that for a binary system, it seems that it is the oscillator wave that is associated with a particle. The DeBroglie wave contains the energy associated with the interactions between electromagnetic systems(i.e., the forces) and with gravity.

As a matter of clarification, let us note that the reference frame, in which V equaled V_0, was chosen carefully and was assumed to be at rest in the gravitational field. We have referred to this situation frequently as the "steady state". It is apparent that for the electron to stay at rest in the laboratory we would have to apply some artificial force and that since the electron contains internally the energy of acceleration, when the restraining force is terminated the electron will continue to fall into the gravitational field.

Any initial energy in the DeBroglie component shown by (41) resides in the gravitational link to that field. It was further assumed that there was no residual relativistic energy residing in the DeBroglie component due to forces applied to the electron that brought it to rest in our frame of reference. This , of course , is a fictional convenience.

When the group velocity in the laboratory is zero, if the electron is released to move freely in the gravitational field it will still have a gravitational group velocity, v_Φ such that,

$$v_\Phi = v' - v \,, \tag{46}$$

This means that the group velocity, v , of (29) must be replaced in (42) by v'. Then v_Φ can be defined by solving (42) and (45) simultaneously for the v' of equation(42), taking $v = v_0$.

The development to date, proposes an actual mechanical basis and a description thereof, of Einstein's aether[16] in the form of the DeBroglie waves associated with oscillators(photons).

I. CLOSED GEOMETRY AND THE PHOTON

Earlier we only considered the photon as a one dimensional entity. We have set forth the proposition, in the previous section, that a photon, a wave of constant wave momentum, can exist in two states. We have called the state of the photon free geometry and that of the electron closed geometry.

Consider what you would have if the photon, a wave of constant wave momentum, existed in a closed geometry state such as in a closed circular orbit. From the viewpoint of an observer on the binary system, the closed binary system would still retain the wave properties it had before the event, but to an outside observer the binary system would no longer have the property of spacial motion unless the observer could use a measuring device several magnitudes smaller than the closed binary system's circumference. In any energy transference experiment the outside observer would still observe for an element dx of the turn, the same results as for a free binary model, and conclude that the closed binary system had a mass. You would effectively have a particle with mass, not a wave with mass.

The really interesting consequence of this closed geometry is the change, to an outside observer, in the attributes of the DeBroglie component. The photon, the oscillator, would now effectively be infinite in length, and thus extremely stable. It could not lose energy. It would effectively be an infinite power

line. Furthermore, it would have a dipole, and to an observer magnitudes greater in size than its orbital radius, would appear to have a Debroglie wave emanating from a point source.

This Debroglie wave would appear as a spherical wave, with the elements of spin, caused by the spin of the oscillator, a wavefront pressure, and an amplitude decreasing with R^2.

The wavefronts of the DeBroglie wave would correspond exactly to the geometries of an electron's magnetic lines of force, and the equal energy pressure vectors on the wavefront would correspond exactly to the lines of force of the electric field, everywhere perpendicular to the magnetic lines of force. Additionally the amplitudes, when at a distance, fall off at a rate corresponding to $1/R^2$. In the case of the gravitational interaction the spin of our closed geometry DeBroglie wave is not canceled out by the gravitational field. It will contribute to the interaction with the gravitational field, a low frequency wave of spin = 0, where all of the other random spins of all the other closed geometry systems have canceled each other out.

It appears possible that this single DeBroglie wave interacting with other systems could account for both the gravitational influence and the electromagnetic influence.

J. MECHANISMS

At this point we should elaborate on several points not sufficiently discussed in the previous sections of this article.

Forces

First let us explain why a photon is said by many to have no mass except possibly one that would be defined at all times by the relationship $m = hf/C^2$.

We have discussed, in the previous section, the spherical nature of the closed geometry binary system's DeBroglie wave and the similar geometries of its DeBroglie wave attributes to those of the electron. Now for the photon, which has open (linear) geometry and its attached DeBroglie wave which will have cylindrical wavefronts with the pressure of the wavefront everywhere perpendicular to the cylindrical geometry of each wavefront one can immediately recognize that this is the geometry normally associated with an electromagnetic wave.

Now it is suggested that Newton's Law of Universal Gravitational Force is not only "proportional to the product of the masses of the particles and inversely proportional to the square of the distance between them" but is also a function of the geometry of the DeBroglie waves attached to the particles. The empirical constant "G" was derived using systems of closed geometry. "G" is not valid for expressing the force between a system of closed geometry and a system of open geometry, as defined herein. This would explain why the electromagnetic wave does not exhibit a force proportional to its mass of $m_\Phi = m_0 e^{\frac{\Phi}{C_0^2}}$ similar to that of a system of closed geometry when under the influence of gravity.

It is suggested that the laws of electrodynamics governing the interactions of the electromagnetic fields for a current flowing through a single turn of wire and the same current flowing along a linear line may be used demonstrate this and to possibly derive a constant, " G' " that will illustrate the gravitational force exerted on one open geometry particle by the other, a particle of closed geometry. It is further suggested that the interaction between a single turn of current carrying wire will be greater and is proportional in some manner to the gravitational forces between two particles, closed geometry, of matter. However It is realized that a proper verification of this model's field characteristics matching experiment, both for the gravitational and electric fields and for variations in geometry can only be done by extensive modeling of energy distributions within the model's own DeBroglie wave for various situations. This is outside the scope of this article.

Pair Production

It is felt that pair production[17], a well established experiment shows without any doubt that at least one mechanism exists for conversion of an open geometry

system to one of closed geometry. This experiment shows that a gamma ray of energy 1.02 Mev or more can, in the vicinity of a nucleus, be converted into an electron/positron pair each of 0.51 Mev energy. Any energy in excess of 1.02 Mev held originally by the gamma ray is converted into kinetic energy shared equally by the electron and the positron.

There is also evidence[18], for inelastic light-by-light scattering involving only real electromagnetic waves, of an electron-positron pair resulting from the collision between a high-energy electromagnetic wave and several laser photons. This however is not felt to influence the discussion of the model presented in this paper.

Elastic and Non Elastic Collisions

Although it is outside the scope of this article it is felt that some discussion of the mechanisms of elastic and non elastic collisions between open geometry photons (oscillators) and closed geometry oscillators (electrons) is warranted. The electron is known to be extremely stable and this is because electromagnetic waves in power lines and antennas do not leak energy except at their ends. The electron, a closed geometry system, has no ends to its' oscillator so it is extremely stable. The open geometry of the photon's oscillator allows it to alter its energy state in elastic and non elastic (Compton effect)[19][20] encounters with particles having oscillators of closed geometry. The wave energy, hf, and wave momentum, $\frac{h}{\lambda}$ of the open geometry oscillator can be changed, probably transferring energy and momentum to the electron's DeBroglie wave. A transference of energy and momentum between their DeBroglie waves may occur. This is a process that others, perhaps, can examine and explain more thoroughly in terms of our model.

This effect is also seen to occur in the process of pair production, but the exact mechanism is not so clear. This case does not fit into the definition of an elastic collision. Furthermore a reverse effect is seen in pair annihilation. This case does not really fit the definition of a non elastic collision. In both of these cases the oscillators change form and the exact mechanisms are not definitely known. It does seem , however, to fall into the category of an "electromagnetic event" because the energy in the oscillator of an electromagnetic wave appears to be affected. If that is so the same reasoning stated concerning the emission of an electromagnetic wave should apply. The author will not speculate further on these events.

When a non elastic collision occurs the open geometry (photon) system appears to transfer all of the energy of both its oscillator and DeBroglie systems, its total energy, into the DeBroglie system of the closed geometry (electron) system resulting in an increase in kinetic energy and increased momentum of the electron.

The laws of conservation of total energy and total momentum, of course, are honored.

Aether

The question of the existence of an Aether has plagued physics for over a hundred years. If there is any substance to this model then the question is answered. The model provides for every quanta of energy of closed geometry to have an attached DeBroglie wave and except for the electromagnetic wave that these DeBroglie waves should exhibit the property of a spin. If this is so the gravitational field consists simply of a sea of these DeBroglie waves with the spins well canceled out. This is easy to see as positive and negative charges are distributed fairly evenly throughout the universe. Also the amplitudes of these waves cannot help but be fairly low frequency at any reasonable distance from a source and would of course, exhibit a decrease in energy at any given point as r increases according to $\frac{1}{R^2}$.

Unity

In the last sentence of the previous section we stated "It appears possible that this single DeBroglie wave interacting with other systems could account for both the gravitational influence and the electromagnetic influence" As also previously stated the geometries of the electromagnetic fields seem to correspond exactly to the geometries of the DeBroglie wave portion of our binary model. If so, then why are they so different in field strengths and the geometries of their energy vectors?

The answer to this has to be that the binary model of the electron has spin.

Our binary model's DeBroglie wave can interact with the gravitational field which has no spin. Our very same model with spin can interact with other similar binary models that do have spin and the results are totally different geometrically and in magnitude. The disturbance caused in the energy distribution of the Debroglie wave in the case of gravity is totally different from the disturbance in the energy distribution of the very same Debroglie wave in the case of interacting with another DeBroglie wave that has spin. Whether the DeBroglie wave of our closed geometry model acts like a charge or like a mass to the outside observer depends on the characteristics of the system that it is interacting with. It can actually interact with both a charge and a gravitational field at the same time!

We will not comment further on the case of a open geometry (electromagnetic wave) system interacting with a closed geometry system with spin (the electron).

We will have to leave it to others to map the energy distributions and prove or disprove this model.

K. MOTION

We ask the question, how in a purely classical sense, could this induced energy cause the motion of our model of the electron through the gravitational aether of Euclidian space? The author suggests that the mechanism is the attempt at reassertion of equilibrium of the DeBroglie wave from a anisotropic energy distribution thereby causing a distorted geometry back to essentially a spherical geometry. As the wave attempts to return to equilibrium its center of gravity will shift. As long as it is subjected to an induced energy change in a given region of space it will continue to shift its center of gravity and would be accelerated. If the inducing force is terminated, say the accelerator is turned off, the DeBroglie wave will reassert equilibrium and cease to be accelerated. It will then continue to move with a constant velocity. This is the same mechanism as was described by D. Bohm [12] [13] for his pilot wave in terms of the P wave which was in an anisotropic energy state attempting to return to an isotropic energy state.

L. NONLOCALITY of REST MASS

This model demonstrates that there is strong evidence for a prediction of nonlocality dependent on the gravitational intensity at the quantum level.

The quantity that we refer to as rest mass appears to be a function of the Φ of creation.

The most obvious example is apparent when considering the emission of electromagnetic waves in different gravitational intensities. The photon component(the wave of constant wave momentum) for a given spectral emmision , say the sodium D line, of the electromagnetic wave varies with Φ . The author sees no experimental evidence that the total emission energy varies so this would imply that the energy in the matter wave component might also be non local. This would have the effect of predicting that the charge of electrons created in different gravitational environments would be different. The difficulty of confirming such an effect would be in obtaining unaltered electrons from known sources of sufficiently different Φ .

CONCLUSIONS AND COMMENTS

At this time we would like to acknowledge that the section on the emission of a binary photon, an electromagnetic wave, was a matter of testing the various possibilities afforded by equations (9) and (12)against experiment until we found the one that satisfied all experimental requirements and preexisting theory of Einstein, DeBroglie and Maxwell. It is hoped that it is understood that the initial object of this article was merely to explore the characteristics of an electromagnetic wave with mass and subsequently we found that there were multiple forks in this path so we followed experiment keeping faith with the General Theory.

It should be noted that C. O. Alley's article, "Proper Time Experiments in Gravitational Fields with Atomic Clocks, Aircraft , and Laser Light Pulses"[1], figures, 43, 45, and 47 along with his closing discussion appear to demonstrate clearly the need for a binary model of the electromagnetic wave. In the article's conclusion, he discusses the fact that regarding the question of a change in wave energy, E

= hf, by an electromagnetic wave propagating thru a changing gravitational potential, his experiments show that this does not happen and that this was further confirmed by experience with test GPS equipment in orbit. This further strengthens the case for our binary model. As an electromagnetic wave climbs out of a gravitational well it must, to conserve energy, suffer a loss of energy. This energy must come from its' DeBroglie wave. This makes a binary model a necessity.

Equation (35), a known gravitational quantity, also lends a strong credibility to the concept of a binary model based on a photon, having a constant wave momentum and a mass, being possible. The results leading to (35)encompass the entire derivation of this paper, including the assumption that the energy imbalance between (9) and (12) is due to the gravitational influence.The expression of this imbalance given by equation (33) led to the prediction of a wavelength of approximately 650 angstroms for an electron's gravity wave when at rest in the Earth's Φ .

It should be possible to devise, at low temperatures, an experiment to look for a fourier component, a strong spectral line near 650 angstroms, in Debroglie waves associated with the free electron by the diffraction method. The discussion of Section H. , Matter Waves suggests that, perhaps at low temperatures in a weak electric field, some statistical experiment similar to the Millikan oil drop experiment[21] might be employed to produce electrons with low enough energies for diffraction results to be meaningful. The object of the experiment is to simply bring an electron to rest, and then to drop it and let it move solely under the influence of gravity through the diffraction grating.

This presentation was basically confined to gravitational forces only for simplicity and brevity of presentation.

This presentation has been confined to the quantum level but the implications of this model could have far reaching significance on the astrophysical scale and has been shown to apply qualitatively to further development of a unified theory. It should aid experimentalists searching for a new avenue for the detection of gravity waves on the macro level.

It should be noted that the references given were kept to a minimum because the author could find no later experimental or theoretical material that would discredit or significantly enhance this model, the principles used being of such a basic nature. There are many speculations on various points of the model but the author does not feel inclined to argue them, since none found seem to form a complete picture. Reiterating, this presentation is of a physical model, not of a mathematical theoretical one.

Ending, we see that our model, in either geometry, is indeed an electromagnetic wave. What this article has accomplished is to break that electromagnetic wave into its individual components and show how each component fits into the world around us. The author has no illusions to the effect that this article is a completed work.

References

[1] Alley, C.,Proper Time Experiments in Gravitational Fields with Atomic Clocks ,Aircraft, and Laser Light Pulses, Quantum Optics, Experimental Gravity, and Measurement Theory, eds. Pierre Meystre and Martin O. Sculley, Proceedings at Conference, Bad Windsheim 1981,1983, (Plenum Press New York, ISBN 0-306-41354-x, pg 363 - 427) (Doubleday, Garden City, N.Y., 1985)

[2] Einstein, A., On the Influence of Gravitation on the Propagation of Light, Jeffery, trans. (Dover, New York, 1952)

[3] Campbell, W. W. and Trumpler, R. J., Astronomical Society of the Pacific, 35, 158 (1923)

[4] DeBroglie L., Wave Mechanics, H. Flint, trans. (Methuen, London, 1930)

[5] Chebotarev,L.V.,Jean-Pierre Vigier and the Stochastic Interpretation of Quantum mechanics, Intro, Apeiron, Montreal (2000)

[6] Davisson, C. and Germer, L. H., Physical Review. 30, 705 (1927)

[7] Pound and Snyder confirmed the slowing down of time in a high gravitational field from a different standpoint; Pound, R. V. and Snider, R. L., Physics Review. 140, 788 (1965) ; Pound, R. V. and Rebka, G. A., Physics Review Letters. 4, 337 (1960)

[8] Einstein, A., The Foundation of the General Theory of Relativity, section 22, in 2.

[9] Lieber,L.R., The Einstein Theory of Relativity.(Holt, Rienhart and Winston, New York, 1945)

[10] Adams, W. S. , Astrophysical Journal. 33, 64 (1911) ; W.S. Adams, Astronomy. 11, 384 (1925) ; J. Evershed, Monthly Notices of the Royal Astronomical Society. 95, 503 (1935)

[11] DeBroglie, Louis, Waves and Quanta, Comptes rendus, Vol. 177, 1923, pp.507-510)

[12] Bohm, D.,Physical Review, 85-2, 180, (1952)

[13] Bohm, D.and Vigier,J.P., Physical Review, 96-1, 208, (1954)

[14] Bohm, D.,Physical Review, 85-2, 166, (1952)

[15] Bohm, D.,Physical Review, 89-2, 458, (1953)

[16] Einstein, A. , Aether and the Theory of Relativity, Lecture at the University of Leyden, 5 May 1920

[17] Eisberg, R. M. , Fundamentals of Modern Physics, 516, John Wiley and Sons, New York, (1960)

[18] D. L. Burke, G. Horton, J.D. Spencer, D. Walz, et al., Physical Review Letters, Vol. 79, No. 9, 1626, (1997)

[19] Compton, A. H., Physics Review, 22, 409 (1923)

[20] Eisenberg, J., Fundamentals of Modern Physics, 85, John Wiley and Sons (1961)

[21] Sears F. W. and Zemansky, M. W. , University Physics, 442, Addison Wesley (1955)

SYMBOLS

Co denotes a universal constant
C denotes the speed of an electromagnetic wave.
E denotes energy
f denotes the frequency of the oscillator.
F denotes a force applied to the quantum of energy (in direction of motion).
G denotes the universal gravitational constant.
h denotes Planck's constant.
K denotes a constant.
m denotes the total mass of the model.
M denotes the mass of a gravitational source.
n is a constant, usually set to 1.
p denotes momentum.
r denotes the distance from the center of a gravitational source.
t denotes time.
v denotes group velocity in general.
W denotes the speed of a DeBroglie wave.
x denotes distance in general.
λ denotes the wavelength of a wave.
Φ denotes the gravitational intensity.
γ denotes the frequency of a DeBroglie wave.

More Articles
Millennium Relativity home page

Rep. Prog. Phys. **63** (2000) 1317–1427. Printed in the UK PII: S0034-4885(00)07909-4

Detection of gravitational waves

L Ju, D G Blair and C Zhao
Department of Physics, The University of Western Australia, Nedlands, WA 6907, Australia

Received 4 January 2000

Abstract

Gravitational wave detectors have been under development since the pioneering work of Weber in the 1960s. The long and painstaking research effort has yielded enormous improvements in detector sensitivity. Astronomical observations of binary pulsar systems have confirmed the existence of gravitational radiation. Direct detection is inevitable once planned detectors reach sensitivity goals.

This review begins by introducing the concept of gravitational waves, and discusses their significance. Section 2 discusses sources of gravitational waves, giving estimates of signal characteristics and signal strengths. Section 3 presents an overview of gravitational wave detection and the critical issues of data processing.

In the fourth section the physics of resonant-mass gravitational wave detectors is discussed in some detail, covering all areas from antenna materials to transducers and the quantum limits to measurement. This section reviews the major operating antennas in the existing worldwide array but also discusses the prospects for achieving substantial increases in sensitivity in the future.

The fifth section presents the concepts and designs for laser interferometer gravitational wave detectors. Large-scale devices will be in operation in the first decade of the twenty-first century and should eventually be certain of detecting a known class of gravitational wave source. At their predicted sensitivity, space interferometers will be able to detect numerous known galactic sources of gravitational waves and also will be able to detect black hole mergers that are thought to have occurred as primordial galaxies merged and grew in the early universe.

(Some figures in this article are in colour only in the electronic version; see `www.iop.org`)

0034-4885/00/091317+111$90.00 © 2000 IOP Publishing Ltd

Contents

	Page
1. Introduction to gravitational waves	1320
1.1. Listening to the universe	1320
1.2. Gravitational waves in stiff-elastic spacetime	1321
1.3. Gravitational waves in general relativity	1326
2. Sources of gravitational waves	1328
2.1. Introduction	1328
2.2. Classification of sources	1329
2.3. Supernovae	1331
2.4. Rough guide to signal amplitudes	1333
2.5. Neutron star coalescence sources	1334
2.6. Low-frequency sources	1335
2.7. Gravitational waves from binary systems	1336
2.8. Stochastic background from the era of early star formation	1337
2.9. Cosmological gravitational waves from the big bang	1339
3. Detection of gravitational waves	1341
3.1. An overview of detector technology	1341
3.2. Space laser interferometer gravitational wave detectors	1344
3.3. The world array of resonant-mass detectors	1346
3.4. Laser interferometer detectors	1349
3.5. Issues of data processing and signal detection	1350
4. Resonant-mass detectors	1355
4.1. Introduction	1355
4.2. Intrinsic noise in resonant-mass antennas	1356
4.3. The signal-to-noise ratio for burst detection	1359
4.4. Transducers	1363
4.5. Antenna materials	1366
4.6. Antenna suspension and isolation systems	1368
4.7. Present status of resonant-mass detectors	1370
4.8. Performance of resonant bars	1371
4.9. Multiple antenna correlation	1374
4.10. Future prospects	1377
4.11. Vibration isolation and suspension developments	1381
4.12. Conclusion	1382
5. Interferometer detectors	1382
5.1. History	1382
5.2. Configurations	1383
5.3. Recycling	1388
5.4. Vibration isolation	1394
5.5. Thermal noise	1402
5.6. Control systems	1406

5.7. Laser stabilization 1409
5.8. Optics 1412
5.9. High-power lasers 1414
6. Conclusion 1416
Acknowledgments 1416
References 1416

1. Introduction to gravitational waves

1.1. Listening to the universe

Our sense of the universe is provided predominantly by electromagnetic waves. During the twentieth century the opening of the electromagnetic spectrum has successively brought dramatic revelations. For instance, optical astronomy gave us the Hubble law expansion of the universe. Radio astronomy gave us the cosmic background radiation, the giant radio jets emerging from black holes in galactic nuclei and neutron stars in the form of radio pulsars. X-ray astronomy gave us interacting neutron stars and black holes. Infrared astronomy gave us evidence for a massive black hole in the nucleus of our own Galaxy.

Gravitational waves offer us a new sense with which to understand our universe. If electromagnetic astronomy gives us eyes with which we can see the universe, then gravitational wave astronomy offers us ears with which to hear it. We are presently deaf to the myriad gravitational wave sounds of the universe. Imagine you are in a forest: you see a steep hillside, massive trees and small shrubs, bright flowers and colourful birds flitting between the trees. But there is much more to a forest: the sound of the wind in the treetops, the occasional crash of a falling branch, the thump thump of a fleeing kangaroo, the pulse of cicadas, the whistles of parrots and honking of bell frogs. The sense of hearing dramatically enriches our experience.

The gravitational wave universe is likely to be rich with 'sounds' across a frequency range from less than one cycle per month (below one microhertz) up to tens of kilohertz. Frequencies in the audio frequency band will be detectable using Earth-based detectors. But lower frequencies will require observatories in space. Gravitational waves are produced whenever there is non-spherical acceleration of mass–energy distributions. The lowest frequencies will consist of extremely red-shifted signals from the very early universe, as well as the slow interactions of very massive black holes, and a weak background from binary star systems. Signal frequencies often scale inversely as the mass of the relevant systems. Black holes below 100 solar masses, and neutron stars will produce gravitational waves in the audio frequency range: nearly monochromatic whistles from millisecond pulsars, short bursts from their formation, and chirrups from the coalescence of binary pairs.

During the twentieth century, at each opening of a new window in the electromagnetic spectrum, the universe surprised us with unexpected phenomena. Our imagination and ability to predict is limited. The sources we predict today are probably just a fraction of what we will hear when our detectors reach sufficient sensitivity.

Gravitational waves are waves of tidal force. They are vibrations of spacetime which propagate through space at the speed of light. They are registered as tiny vibrations of carefully isolated masses. Their detection is primarily an experimental science, consisting of the development of the necessary ultra-sensitive measurement techniques. While the gravitational waves can be considered as classical waves, the measurement systems must be treated quantum mechanically since the expected signals generally approach the limits set by the uncertainty principle.

The binary pulsar PSR 1913+16 has played a key role in the unfolding story of gravitational waves. This system has proved Einstein's theory of general relativity to high precision, including the quadrupole formula which states that the total emitted gravitational wave power from any system is proportional to the square of the third time derivative of the system's quadrupole moment. The pulsar loses energy exactly as predicted by this formula [1]. Hulse and Taylor, who discovered the system more than 20 years ago [2], were rewarded by a Nobel prize in 1993, by which time careful monitoring had shown gravitational wave energy loss from the system in agreement with theory to better than 1%.

1.2. Gravitational waves in stiff-elastic spacetime

In Newtonian physics spacetime is an infinitely rigid conceptual grid. Gravitational waves cannot exist in this theory. They would have infinite velocity and infinite energy density because in Newtonian gravitation the metrical stiffness of space is infinite. Conversely general relativity introduces a finite coupling coefficient between curvature of spacetime, described by the Einstein curvature tensor, and the stress energy tensor which describes the mass–energy which gives rise to the curvature. This coupling is expressed by the Einstein equation

$$\boldsymbol{T} = \frac{c^4}{8\pi G}\boldsymbol{G}. \tag{1.1}$$

Here $\boldsymbol{T}$ is the stress energy tensor and $\boldsymbol{G}$ is the Einstein curvature tensor, c is the speed of light and G is Newton's gravitational constant. The coupling coefficient $c^4/(8\pi G)$ is an enormous number, of order 10^{43}. This expresses the extremely high stiffness of space which is the reason that the Newtonian law of gravitation is an excellent approximation in normal circumstances, and why gravitational waves have a small amplitude, even when their energy density is very high.

The existence of gravitational waves is intuitively obvious as soon as one recognizes that spacetime is an elastic medium. The basic properties of gravity waves can be easily deduced from our knowledge of Newtonian gravity, combined with knowledge that spacetime curvature is a consequence of mass distributions.

First, consider how gravitational waves might be generated. Electromagnetic waves are generated when charges accelerate. Because a negative charge accelerating to the left is equivalent to a positive charge accelerating to the right, it is impossible to create a time-varying electric monopole. The process of varying the charge on one electrode always creates a time-varying dipole moment. Hence it follows that electromagnetic waves are generated by time-varying dipole moments. In contrast to electromagnetism, gravity has only one charge: there is no such thing as negative mass! Hence it is not possible to create an oscillating mass dipole. Action equals reaction. That is, momentum is conserved and the acceleration of one mass to the left creates an equal and opposite reaction to the right. For two equal masses, their spacing can change but the centre of mass is never altered. This means that there is a time-varying quadrupole moment, but there is no variation in monopole or dipole moment.

To be certain of the quadrupole nature of gravitational waves, think of a system which collapses under its own gravity. First think of a spherically symmetrical array of masses that collapse gravitationally towards a point. At a distance there is no difference between the gravitational field of a point mass and that of the same mass distributed in a uniform spherical distribution. (This is a consequence of the inverse square law, and is also true for electric fields.) Hence the process of gravitational collapse of a spherical distribution creates no variation in the external gravitational field, and hence no gravitational waves. Clearly gravitational waves must be created by non-spherical motions of masses. Consider a ring of eight test masses, such as the one illustrated in figure 1.

The simplest non-spherical motion is one in which the edge masses move inwards and the top and bottom masses move apart as shown in figure 1(*a*). Such a quadrupole motion does vary the external field and does create gravitational waves. For a small amount of vertical stretching, and an equal horizontal shrinking, it is obvious that the diagonally placed masses have no radial motion. There is clearly a second polarization rotated 45° from the first in which the diagonal masses move radially, and the top, bottom and edge masses have no radial motion. Unlike electromagnetic waves, gravitational wave polarizations are just 45° apart.

Gravitational wave detection can be easily understood from the symmetry between sources and detectors—time reversal invariance. A gravitational wave will distort a ring of test masses

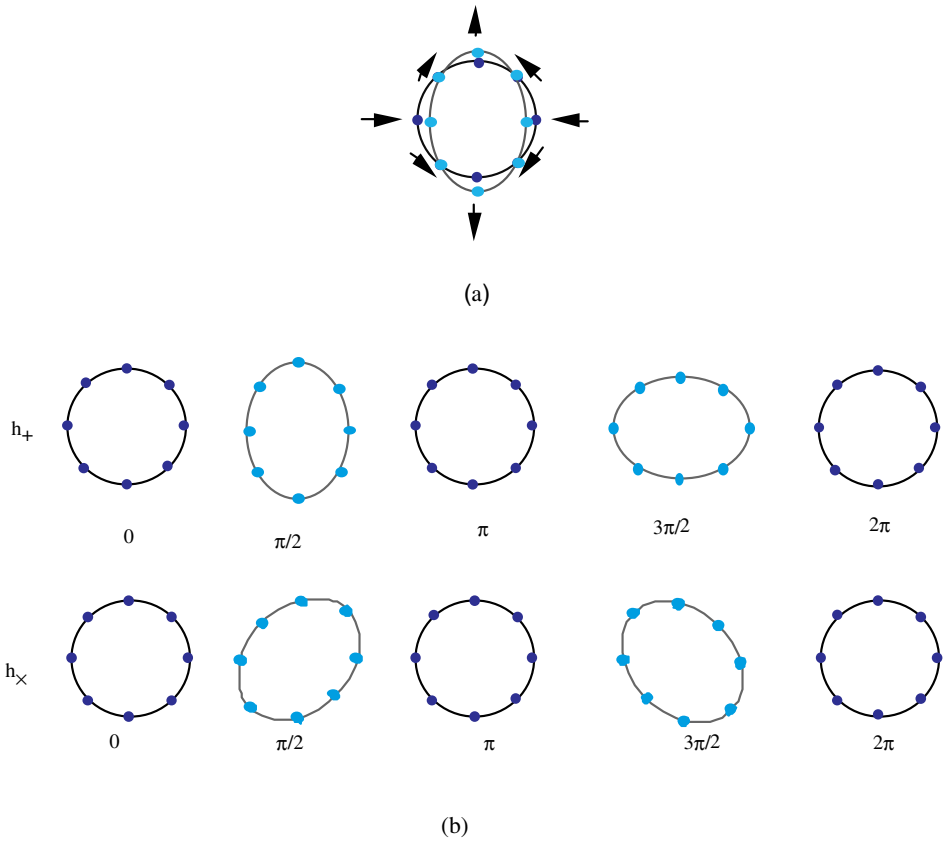


Figure 1. (*a*) The lowest order non-spherical deformation of a ring: the diagonal masses *are not moved*. (*b*) The deformation of a ring of test particles in one cycle of a gravitational wave field.

in exactly the same way that the distortion of a ring of test masses creates gravitational waves. The non-spherical deformation pattern we just observed is exactly like the tidal deformation of the Earth created by the gravity gradient due to the Moon. A gravitational wave is indeed a wave of time-varying gravity gradient. The amplitude of a gravitational wave is measured by the relative change in spacing between masses. That is, the wave amplitude, usually denoted h, is given by $\Delta L/L$, where L is the equilibrium spacing and ΔL is the change of spacing of two test masses. Whereas electromagnetic luminosity depends on the square of the second time derivative of the electric dipole moment, the gravitational wave luminosity is proportional to the square of the *third time* derivative of the mass quadrupole moment. The extra derivative arises because gravitational wave generation is associated with the differential acceleration of masses.

The above deformation patterns also apply to solid or fluid bodies. The rigidity of normal matter is so low compared with that of spacetime that the stiffness of the matter is utterly negligible. Considering the deformations of figure 1(*a*) applying to a solid sphere, such as the Earth, it also follows that the 45° points must involve circumferential motions since the deformation shown acts to transfer matter from the 'equator' to the poles in the same way that the lunar tides act on the Earth.

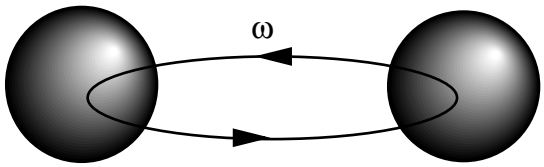


Figure 2. A rotating dumbbell or a binary star system, viewed edge-on, has a maximal variation of quadrupole moment.

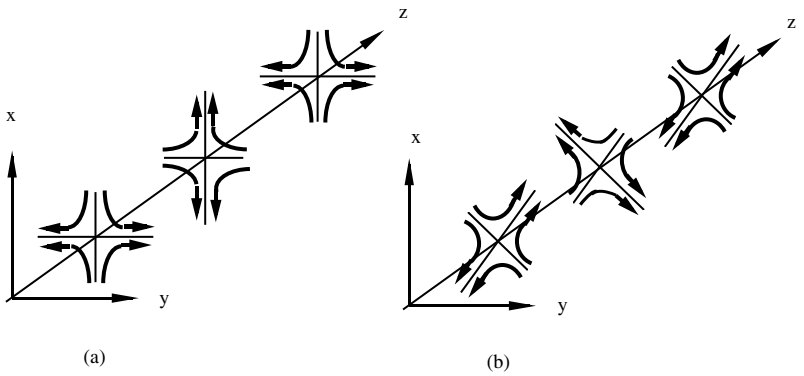


Figure 3. Gravitational wave field force lines. (*a*) '+' polarization; (*b*) '×' polarization.

The gravitational wave has an effective force field determined by the displacement vectors of the test masses. The force field is discussed further below, and is shown in figure 3. The force field indicates that detectors can be designed to couple to gravity waves in several different ways. They may detect straight linear strains, orthogonal strains, or circumferential strains.

The weak coupling of gravitational waves to matter is a consequence of the enormous elastic stiffness of spacetime. If the elastic stiffness of spacetime were infinite (Newtonian physics) the coupling would be zero. In general relativity the generation of gravitational waves is given quantitatively by combining the third time derivative of the quadrupole moment described above, with the appropriate coupling constant. The latter can only depend on the constants G and c (for classical waves) and by dimensional analysis this constant must have the form G/c^5. The luminosity of a source is given by

$$L_G \sim \frac{G}{c^5}\left(\frac{\mathrm{d}^3 D}{\mathrm{d}t^3}\right)^2. \tag{1.2}$$

Except for a numerical factor, this is the Einstein quadrupole formula [3]. There are two useful formulae one can derive from equation (1.2). The first is the formula for a hypothetical terrestrial source or binary star system. The second is for an interacting black hole system. The terrestrial source might be a pair of oscillating masses joined with a spring. Ideally, the spacing of the masses should change from zero to L. This is achieved in the edge on view of a rotating dumbbell or binary star system in a circular orbit as shown in figure 2. Viewed edge-on the masses appear to move in and out periodically twice per rotation cycle. The quadrupole moment for two masses distance x apart is Mx^2. If the motion is sinusoidal at an angular frequency of ω, the square of the third time derivative is $\sim M^2L^4\omega^6$. Thus the gravitational

wave luminosity of such a system is

$$L_c \sim \frac{G}{c^5} M^2 L^4 \omega^6. \tag{1.3}$$

This equation applied to any natural or artificial source in our solar system gives a depressingly small luminosity, due to the extraordinarily small value of G/c^5. However, the situation is different in an astrophysical context.

Suppose that the system is a similar binary system, except that it consists of a pair of gravitationally bound masses, of size such that their escape velocity approaches c and their radius is near to the Schwartzchild radius: that is, a pair of black holes. In this case we can express L in units of the Schwartzchild radius, using $r_s = 2GM/c^2$ and replace $L\omega$ with velocity expressed in light speed units. Then it follows immediately that

$$L_G \sim \frac{c^5}{G} \left(\frac{v}{c}\right)^6 \left(\frac{r_s}{r}\right)^2. \tag{1.4}$$

The remarkable difference between equation (1.3) and equation (1.4) expresses the difference between the physics of normal matter and black holes. Equation (1.3) is scaled by the tiny factor G/c^5, while equation (1.4) is scaled by its enormous reciprocal. Normal matter in our solar system creates negligible curvature of spacetime. A black hole creates an extreme distortion of spacetime. Hence normal matter sources are intrinsically extremely weak, while very large amplitude waves are created in events such as the coalescence of a pair of black holes (for which we would expect $v \sim c$ when $r_s \sim r$). The factor c^5/G is roughly the total electromagnetic luminosity of the universe. This is the upper limit to the gravitational wave luminosity of black hole systems. In reality, equation (1.4) does not take into account the gravitational redshift effects and other spacetime curvature effects which act to reduce the maximum luminosity. However, to order of magnitude, equation (1.4) indicates the extreme luminosity of gravitational waves that can be expected in short bursts when gravitationally collapsed systems with strong gravity, such as black holes (escape velocity $= c$) and neutron stars (escape velocity $\sim 0.1c$), are involved.

As we saw above, a gravitational wave is a wave of gravity gradient which causes relative displacements, or strains between test masses. The detection of gravitational waves requires the detection of small strain amplitudes. We should now consider the typical size of such strain amplitudes. One can very crudely estimate this by scaling the amplitude of the gravitational wave relative to the extreme amplitude at the point of coalescence of two masses to form a black hole. At the point of black hole formation spacetime curvature is very large. For example, the deflection of light for a light beam passing near to the event horizon can approach a complete orbit of a black hole. At the point of generation the dynamic curvature of space that will become the outgoing gravitational wave is unlikely to be able to exceed the static curvature represented by the maximal deflections of light past a black hole. The strain $\Delta L/L$ represented by such deflections can be estimated from the difference in light travel time for the deflected path around the black hole (say half an orbit) and the direct path between the same points in the absence of the black hole. For a half-orbit (in Euclidean geometry) the circular path is $\pi/2$ longer than the direct path, so roughly $\Delta L \sim L$, and the maximum possible strain amplitude is $\sim$unity. But by the inverse square law, the amplitude of the wave reduces as $1/r$. (The energy density which is proportional to the square of the amplitude reduces as $1/r^2$.) So for such a black hole source we can give the strain amplitude at distance r as simply $h \sim r_s/r$. For more realistic sources only a fraction of the total energy will participate in quadrupole motion. Thus it is more reasonable to include an efficiency factor ε which characterizes the fraction of the total system rest mass which can convert to gravitational waves. In this case we

can write

$$h \sim \varepsilon^{1/2} \frac{r_s}{r}. \tag{1.5}$$

Since the Schwartzchild radius of a solar mass is a few kilometres, the maximum strain amplitude that can be expected from any stellar source is numerically equal to the reciprocal of its distance in kilometres. Because r_s is linearly proportional to the mass, gravitational wave amplitudes from very high mass sources, such as colliding 10^9 solar mass black holes in galactic nuclei, will be of correspondingly larger amplitude. Putting in numbers, equation (1.5) gives $h \sim 10^{-16}$ for 10 solar masses and 100% efficiency at the galactic centre, and $h \sim 10^{-13}$ for 3 billion solar masses at 3 Gpc (towards the edge of the visible universe).

Clearly these maximal numbers are very small. It might seem that the supermassive black hole sources might be much more detectable than the stellar mass source. The strain amplitude in this case corresponds to the detection of a motion equal to the size of an atomic nucleus on a one-metre baseline, or one metre between here and Neptune. In fact the detection of such small strains on Earth is probably impossible, because the frequency of the waves from supermassive black hole sources must always be very low. The peak frequency, or its reciprocal, the burst duration, can be estimated from the time the binary black hole system takes to complete its final orbit before coalescence. Its value is about 10 kHz for one solar mass, reducing inversely as the mass. Thus, the peak frequency will be about 1 kHz for the above galactic centre source, and 3×10^{-6} Hz for the distant massive black holes. The latter frequency will be reduced towards 10^{-6} Hz by cosmological redshifts. At such low frequencies environmental effects, and particularly gravity gradients associated with tides and weather variations in the surrounding environment, create perturbations which greatly exceed the desired signal. The only known means around this obstacle is by using drag-free satellite technology to create very stable free-floating masses in space, and laser interferometry between them. In this case detection does look relatively straightforward, though expensive, since it requires several widely spaced spacecraft. For frequencies above 1 Hz, terrestrial detection appears to be possible, limited only by fundamental quantum measurement limits.

For even lower frequencies than 10^{-6} Hz it is possible that radio pulsars can replace man-made spacecraft in detection systems. The pulsar ideally provides a perfect monochromatic timing signal. The radio beams from the pulsar are traversed by incoming gravitational waves. If several pulsars are observed in the same part of the sky, gravitational wave signals would appear as correlated arrival time variations of pulses from pulsars in different directions. In this case it is more convenient to consider the gravitational wave acting not on the pulsar itself, but on spacetime geometry near to the Earth through which the pulsar signal propagates.

Today detectors are in long-term operation which exceed the 10^{-16} sensitivity indicated above by more than two orders of magnitude. Advanced detectors of two types are under development which should achieve another three orders of magnitude in amplitude sensitivity.

For 50 years after Einstein predicted gravitational waves [3] physicists considered them to be of academic interest only. It was not until after the pioneering work of Joseph Weber [4], and his reported discoveries [5, 6] that a growing number of physicists around the world started to develop different types of antennas to search for gravitational waves. Since Weber's first reports, which were never confirmed, the improvement in detectors has been quite remarkable. Relating them to optical telescopes, the improvement achieved so far is equivalent to the step from a 3 cm diameter optical telescope to a 3 m diameter instrument. In the next decade it is hoped that the improvement will be equivalent to a step up in size from 3 m to 3 km. At this sensitivity gravitational wave detection is practically certain, and the field of gravitational astronomy will be able to slowly map and explore the new spectrum, and the objects that it reveals.

In the following sections we will discuss gravitational wave sources in more detail before going on to discuss detectors, first in the form of a general overview, and then with specific emphasis on existing ground-based detectors and their future prospects.

1.3. Gravitational waves in general relativity

Here we give a brief summary of the mathematical basis for gravitational waves. For further details on the theory of gravitational waves, readers are referred to [7–12].

The geometry of spacetime can be expressed by the metric tensor $g_{\alpha\beta}$ which connects the spacetime coordinate $\mathrm{d}x^{\alpha}(\alpha, \beta = 1, 2, 3, 4)$ to the spacetime interval $\mathrm{d}s$ by way of the relation

$$\mathrm{d}s^2 = g_{\alpha\beta}\,\mathrm{d}x^{\alpha}\,\mathrm{d}x^{\beta}. \tag{1.6}$$

In reality, gravitational waves in the vicinity of the Earth will always be very weak. The background curvature can be ignored and the background metric can be approximated as that of the Minkowski flat metric η. An approximation of the gravitational wave field can then be expressed in the form [7]

$$g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta}, \tag{1.7}$$

where $\eta_{\alpha\beta}$ is the metric of the flat background, and $|h_{\alpha\beta}| \ll 1$ is the perturbation on this background. If there is no stress-energy source term in Einstein's field equation, i.e. $T = 0$ in equation (1.1), we are left with the weak field vacuum approximation to the Einstein equations. To obtain an explicit statement of the metric perturbations h it is necessary to make a gauge choice. The most useful gauge is the transverse traceless gauge in which the coordinates are defined by the geodesics of freely falling test bodies. In this gauge, and in the weak field limit discussed above, the equations of general relativity become a system of linear equations, specifically a system of wave equations [7]

$$\left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right) h_{\alpha\beta} = 0. \tag{1.8}$$

Equation (1.8) is a three-dimensional wave equation, telling us that gravitational waves travel at the speed of light c. The gravitational wave curvature tensor $\boldsymbol{h}$ can be considered as the gravitational wave field. The wave field is transverse and traceless, and for waves travelling in the z-direction may be expressed as follows:

$$h_{\alpha\beta} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & h_{xx} & h_{xy} & 0 \\ 0 & h_{yx} & h_{yy} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{1.9}$$

There is no z-component due to the transverse nature of the waves, and to be traceless h satisfies

$$h_{xx} = -h_{yy}. \tag{1.10}$$

Because the Riemann tensor is symmetric, h also satisfies

$$h_{xy} = h_{yx}. \tag{1.11}$$

The symmetry of h means that there are just two possible independent polarization states which are usually denoted h_{+} and $h_{\times}$. In the case of sinusoidal gravitational waves we can express these polarizations as

$$h_{+} = h_{xx} = \mathrm{Re}\,[A_{+}\mathrm{e}^{-\mathrm{i}\omega(t-z/c)}], \tag{1.12}$$

$$h_{\times} = h_{xy} = \mathrm{Re}\,[A_{\times}\mathrm{e}^{-\mathrm{i}\omega(t-z/c)}]. \tag{1.13}$$

Here A_{+} and $A_{\times}$ are the strain amplitudes of each polarization.

We have already seen that a gravitational wave field moves masses in the same way that an electromagnetic wave sets charged particles in motion. Each wave field exerts tidal forces on objects through which it passes. The corresponding lines of force have a quadrupole pattern as shown in figure 3. Figure 3(*a*) shows the force lines of the '+' polarization and (*b*) shows the '×' polarization which is rotated 45° with respect to the '+' state. These time-varying tidal forces can deform an elastic body or change the distance between mass points in free space. A ring of particles placed perpendicular to the wave propagation direction will be distorted as we already saw in figure 1.

Einstein's famous quadrupole formula describes the gravitational wave amplitude from a source. Einstein derived his formula in a slow-motion weak field approximation, but Thorne [8] emphasizes that the result is accurate as long as the reduced wavelength exceeds source size. This condition applies to all but the most compact sources such as forming or coalescing black holes. The latter are potentially the strongest and most detectable sources. It is unfortunate that these are just the ones where the nonlinearity of general relativity, and in particular the gravitational redshift of the outgoing gravitational waves due to the gravitational energy of the spacetime curvature itself, makes the gravitational radiation amplitude extremely difficult to estimate.

The quadrupole formula states that the gravitational wave amplitude h at a distance R from a source is proportional to the second time derivative of the transverse traceless projection of the quadrupole moment evaluated at the retarded time $t - r/c$. That is

$$h_{jk} = \frac{2}{r}\frac{G}{c^4}\frac{\partial^2}{\partial t^2}[D_{jk}(t - R/c)]^{TT}, \qquad (j, k = 1, 2, 3) \tag{1.14}$$

where $[D_{jk}(t - R/c)]^{TT}$ is the transverse traceless projection of the quadrupole moment evaluated at retarded time $(t - R/c)$. The transverse traceless requirement relates to the transverse nature of gravitational waves, and the lack of wave generation from spherically symmetrical motions. For weak fields, for which gravitational self-energy is small (see Damour [13]) $\boldsymbol{D}$ is given by the second moment of the source mass density ρ:

$$D_{jk} = \int \rho(t)[x^j x^k - \tfrac{1}{3}x^2\delta^{jk}]\,\mathrm{d}^3x. \tag{1.15}$$

In this equation the term with the Kronecker delta ensures that D is trace-free.

The total gravitational wave power is proportional to the square of the third time derivative of the mass quadrupole moment [14]. In general, the total energy radiation rate L_G is given by the sum of the squares of all the projections of the quadrupole moment

$$L_G = \frac{1}{5}\frac{G}{c^5}\sum_{jk}\left|\frac{\mathrm{d}^3 D_{jk}}{\mathrm{d}t^3}\right|^2. \tag{1.16}$$

The very small universal constant

$$\frac{G}{5c^5} = 5.49 \times 10^{-54}\ \mathrm{s\,J^{-1}} \tag{1.17}$$

sets the characteristic gravitational radiation power output.

The presence of the factor (G/c^5) in equation (1.16) indicates that unless the $\dddot{D}_{jk}$ involves energy of astronomical proportions, the gravitational wave power will be extremely small. It is easy to show that it is impossible to generate detectable gravitational waves on the laboratory scale, even at extreme limits of known technology. We can only hope to observe gravitational waves emitted by astrophysical sources.

It is useful, however, to consider a laboratory source simply as an application of equations (1.15) and (1.16). Suppose the source consists of a pair of masses distance L apart,

and joined by a spring to allow sinusoidal oscillation of their spacing at angular frequency ω. From equation (1.15), $D = ML^2$, and if $L = L_0 + a \sin \omega t$, it follows that

$$D = ML_0^2 + 2ML_0 a \sin \omega t + Ma^2 \sin^2 \omega t. \tag{1.18}$$

From equation (1.18), taking the third time derivative, it follows that this source will produce gravitational waves at the frequencies of ω and 2ω. If the amplitude a is small compared with L_0, the 2ω term is small and the gravitational wave luminosity is given by

$$L_G = \frac{G}{5c^5} 4M^2 L_0^2 a^2 \omega^6. \tag{1.19}$$

For any practical harmonic oscillator on Earth L_G is infinitesimal. However, such mass quadrupole oscillators have been created as sources of near-field dynamic gravity gradients (not waves) for the purpose of calibrating gravitational wave detectors. Such systems has been successfully used for low-frequency detectors tuned to the Crab pulsar [15] and also by the Rome group to calibrate their resonant-bar detector and measure the inverse square law of gravitation [16].

2. Sources of gravitational waves

2.1. Introduction

Astrophysics provides us with a variety of candidate systems which should be observable in the spectrum of gravitational waves. However, it is important to remember that our powers of prediction of new phenomena are poor, so any list of sources is almost certain to be incomplete.

Amongst stellar mass systems we expect detectable gravitational radiation from the formation of black holes and neutron stars, and the coalescence of binary neutron stars and final collapse of such binaries to form a black hole. We would expect not only discrete sources, but also continuous stochastic backgrounds created from large numbers of discrete sources. In our Galaxy the very large populations of binary stars create a stochastic background in the 10^{-2} to 10 Hz range. In the universe as a whole all of the above neutron star and black hole formation events are likely to merge to form a continuous background in the audio frequency part of the spectrum. This particular background provides an exciting opportunity to observe the earliest epochs of Galaxy formation, and the birth and growth of the supermassive black holes that appear to reside in the nuclei of many galaxies and quasars. We may also be able to observe gravitational waves from the big bang, amplified during the inflationary era, and possible signatures of cosmological phase transitions and topological defects such as cosmic strings. These very earliest sources in the universe would constitute a probe of physics at energy scales far beyond those accessible in particle accelerators and hence represent the best opportunity we have to obtain experimental data from the era of inflation.

Back in our own Galaxy we would also expect to find many quasi-monochromatic sources of gravitational waves such as individual binary star systems, including binary neutron stars as they evolve towards coalescence, and various rotating neutron star systems such as millisecond and x-ray pulsars.

Figure 16 shows the gravitational wave spectrum across nine decades. The spectrum conveniently divides into a terrestrial detection band, above 1 Hz (generally within the audio frequency band), mainly associated with stellar mass compact objects, and a space detection band, from 10^{-6} to 10^{-1} Hz, where sources include both binary star systems in our Galaxy, and cosmological sources associated with massive black hole interactions.

2.2. Classification of sources

All the above sources can be naturally divided into three distinct classes, according to the methods of data processing and signal extraction. The first class consists of catastrophic *burst sources* such as the final coalescence of compact binary star systems, or the formation of neutron stars and black holes in supernova events. The binary coalescence events can consist of binary neutron stars, binary black holes, or neutron-star–black-hole binaries. The burst signal consists of a very short single event, consisting of one or very few cycles, and hence is characterized by a broad bandwidth, roughly determined by the reciprocal of the event duration.

The second class consists of *narrow-band sources*. These include the rotation of single nonaxisymmetric stars, particularly pulsars and accreting neutron stars, as well as binary star systems far from coalescence. All such systems are quasi-periodic because gravitational wave energy loss must cause period evolution, and in general they are also periodically Doppler shifted by binary motion and Earth's orbital motion. Such sources are generally weaker than the burst sources, but in principle they are always amenable to long-term integration to extract signals from the noise. This requires accurate knowledge of the frequency modulations to maintain a coherent integration. Assuming a white noise background and perfect knowledge of the frequency evolution, the signal-to-noise ratio increases as $N^{1/2}$ where N is the number of cycles. For narrow-band sources it may be possible to integrate for 10^8 s, compared with less than 10 ms for a burst source. Thus, at 100 Hz, N can be 10^{10} allowing a 10^5-fold improvement in signal-to-noise ratio.

The third class of sources are the *stochastic backgrounds* produced from the integrated effects of many weak periodic sources in our Galaxy, or from a large population of burst sources at very large distances, as well as the above-mentioned cosmological processes in the early universe. Stochastic backgrounds are difficult to detect in a single detector because they are practically indistinguishable from instrument noise. If the source was not isotropically distributed (such as a population of binary stars towards the centre of our Galaxy), it might be detectable from the variation of observed instrument noise as the detector orientation varied on the rotating Earth. However, a much better way of detecting stochastic backgrounds is by cross-correlating two nearby detectors. In this case the correlated stochastic signal will integrate up in relation to the uncorrelated instrument noise (assuming both detectors to be truly independent). In this case the signal-to-noise ratio increases as $N^{1/4}$, where N is the effective number of cycles, determined by the observation frequency. This technique allows a 300-fold improvement in signal-to-noise ratio in 10^8 s of integration, (compared again with a 10 ms burst source).

Binary neutron star systems can produce gravitational waves in all the three classes. First, a large population of binary neutron stars in our Galaxy, with orbital periods in the range from days to minutes, can produce a stochastic background of individually unresolvable sources in our Galaxy in the frequency band $\sim 10^{-2}$–10^{-5} Hz. Nearby individual systems which are far from binary coalescence could produce detectable nearly monochromatic waves at any frequency up to 0.1 Hz. In addition to the binary orbit, the individual rotation of the stars themselves (if they are nonaxisymmetric), will also give rise to quasi-periodic gravitational waves. For example, the spin-down of a millisecond pulsar can be entirely due to gravitational wave emission if just $10^{-7} M_\odot$ is located in a nonaxisymmetric configuration on the star [17]. As the binary evolves and radiates away gravitational potential energy, it will gradually spiral inwards. As a result, the frequency of the gravitational wave signal will *increase* with time. At the same time, any periodic waves from the rotation of the individual stars will cause loss of rotational kinetic energy, so that this frequency will *decrease* with time. Eventually the stars will coalesce, resulting in a short intense burst of gravitational waves.

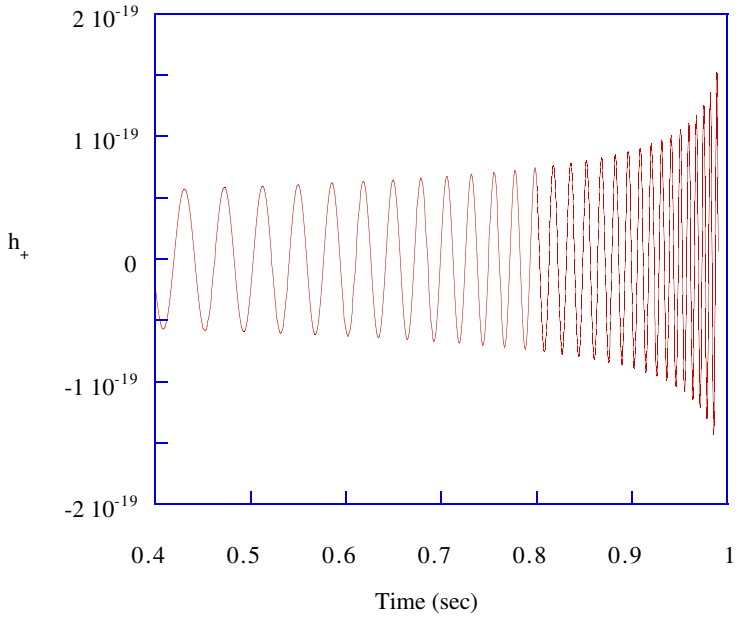


Figure 4. Predicted gravitational waveform from the inspiral of $10M_{\odot}$ black hole binaries [18]. To show the individual cycles near coalescence, the orbital frequency in this graph has been artificially reduced.

Figure 4 is a predicted gravitational waveform produced by the inspiral of a binary made of two black holes [18]. During the final minutes of a coalescing binary the waveform is highly distinctive.

The time evolution of the frequency of two $1.4M_{\odot}$ neutron stars in a binary system is shown in figure 5. Over a period of about 1000 s the frequency rises from about 10 Hz to 1 kHz as the neutron stars spiral together. This part of the merger begins when the stars are within about 1000 km of each other. The orbital velocity is $\sim 0.1c$. Signal detection can make use of exactly the same principles used to extract narrow-band signals due to the fact that the time evolution of the signal frequency and phase is predictable. Matched filtering, based on the existence of a family of accurately predictable waveforms, can allow integration over all of the observed signal cycles. A terrestrial detector may be able to observe more than 1000 gravitational wave cycles from a neutron star binary. The total number of cycles observable increases strongly as the lower cut of frequency is reduced. This provides a strong incentive for creating detectors at the lowest possible frequency. For 1000 observable cycles, the signal-to-noise ratio is improved by the square root of this number, or about 30.

Exactly the same concepts may be applied to supermassive black hole binaries. The signal frequency decreases inversely with the black hole mass. Thus a pair of 10^8 solar mass black holes would produce a chirp of gravitational waves rising from one cycle per year to a cycle per day over a period of 10^{11} s! This is much too long to observe the entire event, so in reality one could only expect to observe rather few cycles. However, since there appear to be a large population of quasars and galaxies containing massive black holes, as well as a large population of interacting galaxies, such events may not be uncommon, and may give rise to numerous strongly detectable sources at very low frequencies.

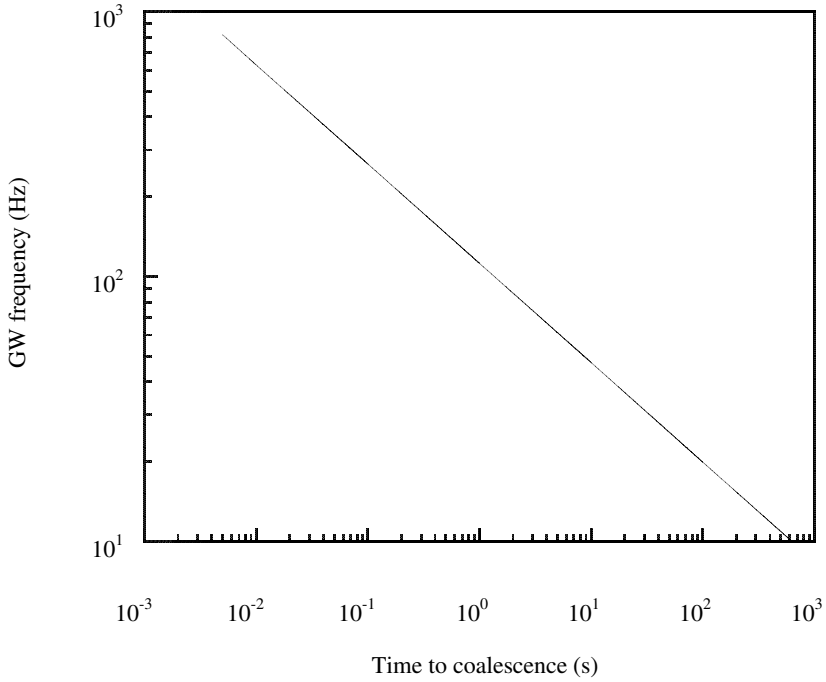


Figure 5. The frequency evolution for coalescence of a binary system consisting of a pair of $1.4M_\odot$ neutron stars.

2.3. Supernovae

Supernovae have long been considered a primary source of gravitational wave bursts. Unfortunately, astrophysics gives us few clues to their efficiency in producing gravitational waves during core collapse. Indeed, the true nature of the various supernova classes is still uncertain. In particular, it is uncertain whether a type I(a) supernova occurs through detonation or collapse in a solitary or binary white dwarf system. Some supernovae, such as the Crab supernova of the year 1054AD, do create neutron stars, but others such as supernova 1987A have failed to yield an identifiable neutron star. It is also unknown whether supernovae can directly create black holes.

The possibility of strong gravitational radiation emission only occurs if the event consists of gravitational collapse to a neutron star or black hole. Even in this case the efficiency of gravitational radiation emission is contentious. (The efficiency ε is defined as the fraction of the rest mass of the system concerned converted to gravitational waves.) Estimates of gravitational wave emission have been based on two quite independent approaches. The first follows the gravitational collapse of a system in two or three dimensions, considering as much of the physics as possible—magnetohydrodynamics, neutrino physics, and general relativity—in an attempt to deduce the time dependence of the quadrupole moment and hence the conversion efficiency to gravitational waves.

The second approach assumes that the collapse has occurred and follows the time evolution of a newly formed hot and rapidly rotating neutron star. Any phenomenon that creates non-axial symmetry will convert rotational kinetic energy into gravitational waves. In addition, convective motions, vibrational modes of oscillation of the star, and nonaxisymmetric emission of neutrinos can convert into gravitational wave emission. To date, many core collapse

calculations have predicted rather low efficiency—say 10^{-6} to 10^{-10} [19, 20], while the post-collapse evolution calculations have predicted far higher efficiencies, of order 10^{-3} to 10^{-4}. We will discuss some of these results below. However, it must be emphasized that all the models used so far are deficient due to uncertainty regarding the equation of state and the viscosity and the difficulty in constructing a full 3D numerical general relativistic hydrodynamical code which must also include magnetic and neutrino phenomena [21]. Due to the enormous difficulties involved it seems most unlikely that theory alone will be able to answer the primary observational questions on the efficiency and the waveforms generated in supernovae. However, almost all models show an unsurprising common feature: the efficiency of gravitational wave production depends on the angular momentum of the progenitor star.

Lai and Shapiro [22] have considered the time evolution of a new-born rapidly rotating neutron star. They have shown that the new-born star is driven by gravitational radiation into a non-asymmetric configuration due to a bar-mode instability. A unique gravitational wave signature ensues: the wave frequency sweeps rapidly downward from a few hundred hertz towards zero, while the wave amplitude increases rapidly from zero at the onset of the instability to a maximum at a few hundred hertz, and then reduces steadily as the frequency falls. Additional gravitational wave signals can also arise in rapidly rotating neutron stars. The rotating stars are modelled as nonaxisymmetric ellipsoids. A secularly unstable Maclaurin spheroid [23] will evolve away from the axisymmetric configuration due to gravitational radiation, and proceed ultimately toward a Dedekind ellipsoid [23].

According to Lai and Shapiro the characteristic amplitude of a gravitational wave during the evolution from a Maclaurin spheroid to a Dedekind ellipsoid is given (within 20% accuracy) by

$$h_c \sim 1.8 \times 10^{22} \left(\frac{10\ \mathrm{Mpc}}{R}\right) \left(\frac{M}{10M_\odot}\right)^{3/4} \left(\frac{r_0}{10\ \mathrm{km}}\right) f^{1/2} \tag{2.1}$$

where M and r_0 are the mass and radius of the star, respectively. Here h_c refers to the effective amplitude which takes into account the number of cycles that the signal is within the detector bandwidth (see below).

At high frequency, gravitational radiation can be expected from the evolution of a Jacobi-like ellipsoid [23] toward a Maclaurin spheroid. This Maclaurin spheroid can evolve further to a Dedekind ellipsoid. The characteristic amplitude during the Jacobi-like evolution can be fitted to the form

$$h \sim 2.7 \times 10^{-20} \left(\frac{10\ \mathrm{Mpc}}{R}\right) \left(\frac{M}{1.4M_\odot}\right)^{3/4} \left(\frac{r_0}{10\ \mathrm{km}}\right) f^{-1/5}. \tag{2.2}$$

Houser *et al* [24] have modelled the gravitational radiation from a bar-mode instability in rapidly rotating neutron stars. Their calculation using Newtonian gravity and without consideration of further collapse to a black hole, nor other hydrodynamic instabilities, gives a gravitational radiation conversion efficiency of $\varepsilon \sim 0.1\%$.

The above examples seem to indicate that supernovae which produce rapidly rotating neutron stars may be reasonably efficient sources of gravitational radiation. The nature of the production process is likely to be through shape instabilities such as those discussed, but it is unlikely that predictions of waveform are accurate. Large amounts of angular momentum may be radiated away in gravitational waves but if the duration and frequency evolution are unknown this presents an additional complication when it comes to trying to dig a signal out of the detector noise. The fraction of supernovae for which high gravitational wave emission occurs is unknown. In the following discussion where we need to use a numerical value, we shall adopt an efficiency of 0.1%. However, the uncertainty of this number must be recalled.

To start with, it is useful to note that a supernova of 0.1% efficiency produces a characteristic strain amplitude of $\sim 10^{-18}$ at 10 kpc (within our Galaxy), and 10^{-21} at 10 Mpc (halfway to the Virgo Cluster of galaxies). The chance of detecting gravitational wave bursts obviously depends strongly on the rate of the burst events. Due to the isolation of the Milky Way Galaxy, and to the large distances required to substantially increase the size of the target population, the amplitude distribution of bursts is extremely non-uniform. Strong events from our Galaxy are almost certainly rare, and to increase the event rate substantially one needs to be able to detect events in the Virgo Cluster. Thus, to have a chance of detecting several events per year, the sensitivity must be able to detect characteristic amplitudes of less than 10^{-21}.

2.4. Rough guide to signal amplitudes

It is useful to have some formulae with which to make rough estimates of signal amplitudes. For a continuous gravitational wave of frequency f_g, the strain amplitude h is related to the power density w through the relation [25, 26]

$$w \approx \frac{\pi c^3}{4G} f_g^2 \langle h^2 \rangle = 3.18 \times 10^{35} f_g^2 \langle h^2 \rangle \ \mathrm{Wm}^2. \tag{2.3}$$

where $h^2 = h_+^2 + h_\times^2$. Because of the large numerical constant in equation (2.3) the amplitude h is extremely small even for a fairly large power density. For a gravitational wave with strain amplitude of $h \sim 10^{-21}$ (typical of possible signals from the Virgo Cluster) at a frequency of 1 kHz, the flux would be 0.3 W m^{-2}, which is about 10^{20} times bigger than typical radio astrophysical energy fluxes. The strain amplitude can be written as

$$h = 4\pi R^2 w \sim \left(\frac{G}{\pi^2 c^3}\right)^{1/2} \frac{L^{1/2}}{f_g R} \sim 1.7 \times 10^{22} \left(\frac{1\ \mathrm{kHz}}{f_g}\right) \left(\frac{10\ \mathrm{Mpc}}{R}\right) \left(\frac{L}{10^{46}\ \mathrm{W}}\right)^{1/2}. \tag{2.4}$$

As we saw above, the modelling of gravitational wave forms in gravitational collapse is extremely uncertain. However, for a gravitational wave burst event, the characteristic time scale of the event τ_g, and the total gravitational energy released E_g, provide a reasonable basis for estimating source parameters. The energy radiation rate L is related to τ_g and E_g by

$$L \sim E_g/\tau_g. \tag{2.5}$$

Burst sources naturally have a broadband spectral distribution. The characteristic frequency of a burst of duration τ_g is roughly

$$f_g = \frac{1}{2\pi\tau_g}. \tag{2.6}$$

This frequency roughly defines the peak frequency in the spectrum. For a roughly Gaussian burst, the width of the spectrum Δf is of the same order of f_g. The strain amplitude can then be written as [8]

$$h \sim \left(\frac{G}{\pi^2 c^3}\right)^{1/2} \frac{(E_g/\tau_g)^{1/2}}{f_g R} \sim 5.8 \times 10^{-20} \left(\frac{E_g}{M_\odot c^2}\right)^{1/2} \left(\frac{1\ \mathrm{kHz}}{f_g}\right)^{1/2} \left(\frac{10\ \mathrm{Mpc}}{R}\right). \tag{2.7}$$

If a gravitational collapse forms a black hole, we can be more specific in estimating the event duration. Defining a characteristic time τ to be the time for the gravitational wave to travel across the region of strong gravitation d_s, which is assumed to be about twice the gravitational radius $2GM/c^2$, and assuming τ_g is approximately the same as the characteristic time τ, we have

$$\tau \sim \frac{1}{2\pi f_g} \sim \frac{d_s}{c} \sim \frac{4GM}{c^3}. \tag{2.8}$$

For a system of several solar masses, this corresponds to a frequency of a few kHz. Putting equation (2.8) into equation (2.7) and using $E_g = Mc^2$, the strain amplitude of a burst event then becomes [8]

$$h \sim \frac{1}{2\pi}\frac{c}{f_g R}\varepsilon^{1/2} \sim 5 \times 10^{-21}\left(\frac{1\ \text{kHz}}{f_g}\right)\left(\frac{10\ \text{Mpc}}{R}\right)\left(\frac{\varepsilon}{10^{-3}}\right)^{1/2}. \tag{2.9}$$

2.5. Neutron star coalescence sources

The modelling of gravitational wave emission from neutron star coalescence has been studied extensively. For most of their evolution, the neutron stars can be considered as point masses, and much of their evolution is well described by the quadrupole formula equations (1.15) and (1.16). The waveform as shown in figure 4 is quite distinctive and amenable to the method of matched filtering for signal detection. A numerical template is used, defined by the set of possible waveforms. When this is cross-correlated with the data and correctly matched in phase, it will produce a large positive correlation. The signal-to-noise ratio is substantially enhanced by this means. The apparent signal enhancement achievable is expressed in terms of the characteristic amplitude h_c. The characteristic amplitude represents the effective amplitude detected after optimal filtering of the waveform. Roughly, h_c includes an enhancement of the signal by the square root of the number of cycles within the spectral band of interest and is roughly a factor of 30 for a neutron star coalescence detected by a laser interferometer detector, although this increases strongly if the waveform is detectable at much lower frequencies where the frequency evolution is slow. For example, the number of observable cycles increases almost 50 times if the detector is able to observe down to 10 Hz instead of 100 Hz. This means that we can only roughly estimate the size of the detectable signal, as it depends on the detailed frequency response of the detector.

Thorne [18] gives the characteristic strain amplitude of the waves from inspiralling binaries as

$$h \sim 0.237\frac{\mu^{1/2}M^{1/3}}{r_0 f_c^{1/6}} = 4.1 \times 10^{-22}\left(\frac{\mu}{M_\odot}\right)^{1/2}\left(\frac{M}{M_\odot}\right)^{1/3}\left(\frac{100\ \text{Mpc}}{R}\right)\left(\frac{100\ \text{Hz}}{f_c}\right)^{1/6}. \tag{2.10}$$

Here M and μ are the total and reduced masses: $M = M_1 + M_2$, $\mu = M_1 M_2/M$, and f_c is roughly the frequency of maximum detector sensitivity.

Lai and Shapiro [22, 27] have modelled neutron star coalescence taking into account the dissipative hydrodynamics of the systems. They showed that a hydrodynamical instability arises through tidal interactions, which significantly accelerates the coalescence at small separations. This leads to a reduction in the coalescence time, and an increase in h compared with a non-viscous system, as shown in figure 6.

The rate of coalescence events is such that the chance of an event in our Galaxy is negligible. Based on the statistics of observed single and binary neutron star systems and on the supernovae rate in external galaxies [28, 29], it is estimated that the merger rate of binary neutron stars in our Galaxy is between 10^{-6} yr^{-1} and 10^{-4} yr^{-1}. For galaxies at $R \leqslant 200$ Mpc a lower limit of this rate is roughly 1–3 yr^{-1}. However, Tutukov *et al* [30] and Yamaoka *et al* [31] have shown that the above merger rate, calculated by means of statistics of observed binaries, is probably wrong because of the short lifetimes of most new-born neutron star binaries. From models based on stellar evolution, they estimate a neutron star binary merger rate up to 100 yr^{-1} in galaxies out to 200 Mpc (assuming a Hubble period of $T_{\text{Hubble}} = 1.5 \times 10^{10}$ yr, which corresponds to a Hubble constant of $H_0 \sim 66$ km s^{-1} Mpc^{-1}).

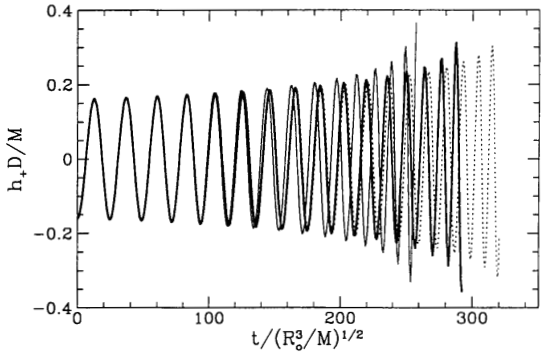


Figure 6. Waveform from neutron star binary coalescence [22]. The thick solid curve corresponds to zero viscosity. The thin solid curve assumes a mass-averaged shear kinetic viscosity $v = 0.5(M/r_0)^{1/2}$, where M and r_0 are the mass and the radius of the star respectively. The dotted curve is the case for two point masses (i.e. as figure 4).

The neutron star coalescence signals amplitude and waveform can be predicted with reasonable confidence, so that such sources are certain to be detected when sufficient sensitivity is achieved. The burst sources are also very promising for advanced gravitational wave detectors [32–34], but suffer from uncertainty in the value of ε. From equation (2.9), the gravitational wave strain amplitude for burst events at 200 Mpc is comparable to that from binary coalescence if $\varepsilon \sim 0.01$. In this case, the event rate will be much greater.

From the estimation of the stellar population with distance, the merger rate would be $\sim$0.1 yr^{-1} for a distance of $R \sim 20$ Mpc, the typical distance to the Virgo Cluster. The burst rate by contrast could be 30 yr^{-1} [35].

2.6. Low-frequency sources

As discussed above, very intense low-frequency gravitational wave sources can be expected from gravitational waves associated with the merger of massive black holes. Rees [36, 37] has argued that massive black holes are inevitable in the cores of young galaxies. There is very strong evidence that such black holes exist in many objects, with masses ranging from 10^6 to 10^9 solar masses. Galactic mergers are likely to give rise to such black hole mergers so one estimate of the rate of powerful gravity wave events can be obtained by estimating the rate of galactic mergers. For large galaxies with central black holes, Haehnelt [38] has estimated this rate at about one per century. This does not include the far more frequent mergers of smaller galaxies for which central black holes have not been confirmed. Vecchio [39] has shown that for the black hole merger rate to reach one per year practically all galaxies out to $z = 1$ would have to contribute black holes to feed the merger process. The latter is not such a strong constraint, however, since the horizon for detecting black hole mergers could be far beyond $z = 1$. For example, if one considers mergers to $z = 3$, only a few per cent of galaxies are required to have a central black hole to achieve one event per year.

Potentially detectable low-frequency gravitational waves can also be created by low-mass objects orbiting massive black holes. The low-mass objects could be smaller black holes or neutron stars, white dwarfs or even main sequence stars. Such sources could exist in the nucleus of our own Galaxy, and could in principle be detectable well beyond the Virgo Cluster (which multiplies up the number of potential sources by several thousand).

The basic physics behind gravity wave emission from sources of this type relates to the question of whether the gravity gradient from the central black hole is sufficient to tidally disrupt the incoming object. Clearly, main sequence stars will be most easily disrupted. Only if the central black hole is capable of swallowing whole the incoming object, will the gravitational radiation be strong. Otherwise the tidally disrupted star will form an accretion disc and slowly

accrete the material with negligible gravitational wave emission. To avoid tidal disruption, the Schwartzchild radius of the black hole must be large compared with the radius of the infalling object. Roughly, the central black hole must be 10^7–10^8 solar masses for main sequence stars, 10^4 solar masses for white dwarfs, and 10 solar masses for neutron stars.

Sigurdsson [40] has estimated the rate of capture of stellar mass black holes by massive black holes in galactic nuclei. The gravitational potential of the central black hole creates a cusp-like stellar density profile. It is difficult to estimate the space density of objects in the central cusp. It depends on star formation in the central high-density region of galaxies. However, the population in the cusp will never achieve dynamic equilibrium because stars approaching too close to the central black hole will be lost into the hole. Sigurdsson estimates the rate of black hole capture by a central object as 10^{-8} per year, meaning that a realistic detectable rate (one per year) requires observations to a range of the order of 3 Gpc. If the high densities in galactic nuclei favour higher mass star formation the event rate could increase by an order of magnitude.

2.7. Gravitational waves from binary systems

Short-period binary systems can create interesting amplitudes of gravitational waves in the 10^{-1}–10^{-5} Hz range. Such binaries exist in several classes. One of the most definite classes consist of the W Ursa Majoris binaries (WUMas), which are contact binary stars, with orbital periods of hours. They are generally low-mass systems. About one star in 150 with mass $>0.6M_\odot$ is a contact binary [41]. Lower mass binaries are difficult to detect: there could be an equal population of such systems with even shorter orbital periods.

A second important class of short-period binary stars are the cataclysmic variables, consisting of an interacting main sequence and white dwarf binary. Cataclysmic variables have orbital periods in the range 1000 s to one day: the shortest-period systems are probably white-dwarf–white-dwarf systems.

Neutron star binaries occur in various forms, from the NS–NS binaries such as the Hulse–Taylor pulsar PSR 1913 + 16, to rather more common systems in which neutron stars have white dwarf or main sequence star companions. The latter often occur as interacting binaries—low-mass or high-mass x-ray binaries, in which x-ray emission occurs due to mass transfer on to the neutron star.

Verbunt [42] has summarized the density and strain amplitude expected from the short-period binary star systems. Table 1 below is based on his review. The table shows the rough number density and mass parameters, and the distances of typical sources. Most produce gravitational waves $\sim 10^{-3}$–10^{-4} Hz, at an amplitude $\sim 10^{-20}$–10^{-22}. There is clearly an abundance of sources in the categories of nearby sources ($<$100 pc), such that the total population creates a stochastic background of gravitational wave noise.

Several x-ray binary systems have been shown to contain neutron stars spinning in the range 250–350 Hz. It is suggested that their spin rate is determined by the balance between the mass accretion which provides a source of energy and angular momentum, and gravitational wave emission which is the dominant energy sink. Emission could occur by a variety of symmetry breaking instabilities. A specific suggestion is the so called r-mode instability, in which rotational fluid flow patterns are induced in the neutron star [43].

Stellar evolution studies suggest that globular clusters are a breeding ground for close-spaced binary black holes [44]. It has been recently suggested that these binaries get ejected from globular clusters by three-body interactions, creating a halo population of binaries which will coalesce in less than a Hubble time. Such binaries might coalesce without electromagnetic signature, and could have a sufficiently large population that they could be detected at a

Table 1. Binary sources of low frequency gravitational waves. The table shows the rough galactic density of each source, the distance within which sources are expected, the typical mass of each binary component, and the typical frequency and strain amplitude from the nearest sources.

Type	Density/number	d (pc)	$M/M_\odot$	$m/M_\odot$	$\log f$ (s^{-1})	$\log h$
WUMa (0.3–0.6$M_\odot$)	2×10^{-4} pc^{-3}	15	0.6	0.3	−4.0	−20.4
WUMa (0.1–0.3$M_\odot$)	2×10^{-4} pc^{-3}	15	0.3	0.1	−3.7	−20.7
Cataclysmic variables	10^{-5} pc^{-3}	45	0.3	0.6	−3.7	−20.7
Double degenerates (AM CVn)		100	0.04	0.6	−2.7	−21.2
Low-mass x-ray binaries ($P_b < 2 \times 10^{-4}$ s)	30	1 000	0.4	1.4	−3.8	−21.7
Low-mass binary pulsars (PSR 2051-08)		1 300	0.03	1.4	−3.8	−21.7
High-mass x-ray binaries (Cyg X-3)	1	10 000	4.0	1.4	−3.9	−21.9
NS–NS binary pulsars (1913 + 16)	$<10^{-5}$	500	1.4	1.4	−3.7	−20.8
Binaries in globular clusters (4U1820-30)		8 100	0.06	1.4	−2.5	−22.3

reasonable rate at $h \sim 10^{-21}$.

Gamma ray bursts which emit energy comparable to a solar mass at cosmological distances could be due to the formation of black holes, or to neutron star binary coalescence, or to the coalescence of neutron-star–black-hole binaries. Except in the case of spherically symmetrical black hole formation, these mechanisms should all include strong gravitational wave bursts. Searches have failed to find correlations between bursts and existing detectors, but this is not unexpected as sensitivity is still not high enough.

Neutron star black hole coalescences, whether or not they are associated with gamma bursts, can allow neutron star structure and microphysics to be probed because the break point in the coalescence waveform is set by the tidal disruption of the neutron star. This depends strongly on the neutron star radius and equation of state.

2.8. Stochastic background from the era of early star formation

We now consider the effects of supernovae and neutron star births at cosmological distances. In this case, we are extrapolating from a radius of 10 Mpc to a radius of several Gpc. For example, we consider supernovae from galaxies at redshift $z = 2$. Such galaxies are much older than massive stellar lifetimes, and the rate of supernovae in such systems is generally thought to be 10–100 times greater than supernovae in contemporary galaxies [45]. This is supported by observations by Cowie *et al* [46] which indicate a fourfold enhancement in faint blue galaxies at $z \geqslant 1$. At greater distances, millimetre wavelength studies of the Hubble deep field region show the presence very-high-luminosity objects consistent with dust-enshrouded galaxies at $z \sim 2$–4. Observations indicate that star formation is occurring at rates ∼50 times that in the present epoch [47]. The significance of these increased star formation rates, and hence increased supernova rates, is that it leads to the possibility that the gravitational waves from supernovae create a nearly continuous stochastic background.

First consider a simple case. Suppose that all supernovae have a gravitational wave burst duration τ, and a mean rate of occurrence within some horizon distance (say 3 Gpc), of R bursts per second. Then the mean duty cycle D of supernova bursts is given by $D = R\tau$. If $\tau \sim 10^{-3}$ s, then D reaches unity when R reaches 3×10^{10} y^{-1}, or 10^3 s^{-1}. When D reaches unity the supernova bursts create an effectively continuous stochastic background. To make an accurate determination of the supernova stochastic background, one needs to take into account both evolutionary effects and cosmological effects. If every population I (second-generation) star was the result of a single prior supernova, then there would need to have been $\sim 10^{21}$ supernovae to create the observed population of second-generation stars. This corresponds

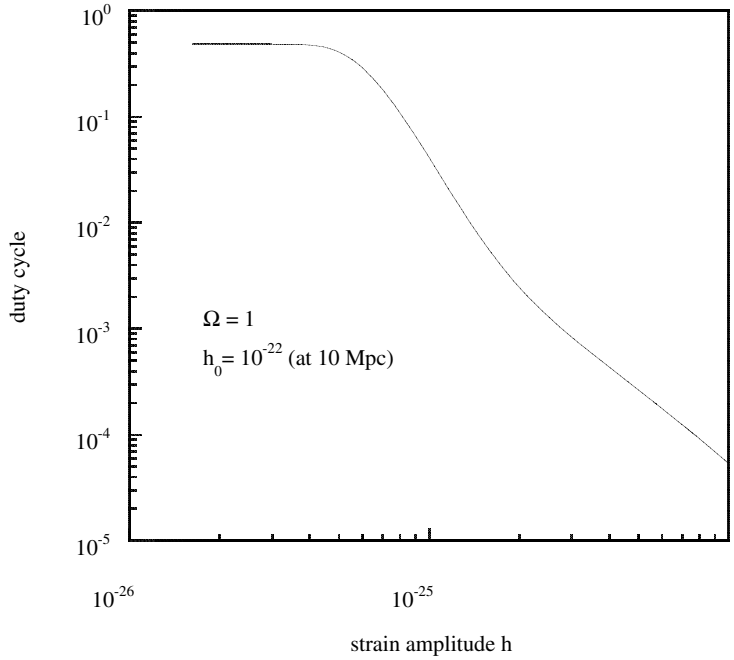


Figure 7. Duty cycle versus strain amplitude for supernova-generated gravitational radiation (with an initial Gaussian burst of star formation at $z = 2$) [49].

to an average rate Excedrin 1000 supernovae s^{-1}. Redshifts both stretch the pulses and the mean time between bursts. The luminosity distance in non-Euclidean geometry changes the observed amplitude of each burst.

A preliminary analysis in flat spacetime with assumed star formation rates showed that the event rate could be as high as 10^4 s^{-1} [48]. Burman *et al* [49] refined the predictions of [48] using various predictions for star formation rates to determine the duty cycle of short bursts of gravitational waves from supernovae within the observable universe for various cosmological models. Ferrari *et al* have separately considered supernovae [50] and gravitational waves from young neuron stars [51]. They obtained event rates $\sim$20 s^{-1}.

Figure 7 gives a typical result [49]. This result uses a Gaussian burst of star formation about $z = 2$. It shows the burst amplitude versus duty cycle for supernovae assumed to have an amplitude $h_0 = 10^{-22}$ at 10 Mpc. The background can only be considered a true stochastic background as D tends to unity: for this model this occurs at $h \sim 10^{-25}$. Most of this signal is due to events occurring during the initial burst of star formation. The amplitude of this stochastic background will be characteristic of the more distant sources at $z = 2$.

Like all stochastic backgrounds, the supernova background can in principle be detected by cross-correlation of signals between nearby detectors (less than half a wavelength separation: for $\tau \sim 10^{-3}$ s, they should be less than 100 km apart) [52]. As discussed above, the signal-to-noise ratio is increased as the $\frac{1}{4}$ power of the effective measurement time defined by the cross-correlation integration time. For 10^7 s, integration at 1 kHz, this represents a 300-fold improvement. Thus, the combined effects of all supernovae is to create a signal which can be detected at a signal-to-noise ratio comparable to that of an individual supernova at 20–30 Mpc. Thus, a detector capable of detection extragalactic supernovae can, with cross-correlation, detect a stochastic background produced at 30 times greater distance.

There are other aspects of the supernova stochastic background which are worth mentioning. Its spectrum represents the average spectrum over all supernovae, but it will be reddened according to the contribution of high-redshift supernovae. The duty cycle is clearly amplitude dependent. Nearer sources will create less-frequent, larger-amplitude bursts. At low duty cycle the background will be like popcorn noise, while for $D > 1$, it will approximate Gaussian noise. There is an important difference in this regard. The presence of a popcorn noise component [53] means that unlike true white noise, the individual short bursts create broadband intensity correlations which might allow more powerful digging into the noise. For example, the broadband correlations might allow the background to be detected as spectral intensity correlations within a single detector. This might be combined with cross-correlation between two detectors to dig still deeper into the noise. Further work in this area is badly needed.

The energy density of the supernovae background Ω_{SN}, expressed as a fraction of closure density, is given by

$$\Omega_{SN} = \Omega f_s f_{SN} \bar{\varepsilon} \tag{2.11}$$

where Ω is the usual fraction of closure density for the universe as a whole, f_S is the fraction of this matter which forms into stars in a Hubble time, f_{SN} is the mass fraction which takes part in supernova events in a Hubble time, and $\bar{\varepsilon}$ the mean gravity wave conversion efficiency for supernovae. It is possible that Ω_{SN} could be in the range 10^{-6}–10^{-8}. However, if $\bar{\varepsilon}$ is low and the duty cycle is low, Ω_{SN} could be $\ll 10^{-10}$.

If the majority of the gravitational waves are generated in relatively long duration spin-downs of neutron stars the spectrum will be dominated by a continuous stochastic component, but if it is emitted in short supernova bursts, the popcorn component will dominates. If the energy density of the early star formation stochastic background is $\sim 10^{-8}$, then it should be eventually detectable by pairs of advanced detectors [54].

To show the signals on the same scale, which relates to detectability, stochastic sources are assumed to have been integrated up for times $\sim 10^8$ s, binary coalescences have been integrated over the coalescence frequency range, while the burst sources signal strengths are the only ones representing the instantaneous signal amplitude.

Finally, to summarize our discussion of sources, we present a graph (figure 8) containing estimates of various events. The comparison is approximate, as it compares various sources detected by various techniques. Supernovae signals would appear as bursts requiring no special signal processing. However, estimates must allow for a wide range of efficiencies and source distances. Black hole formation is similarly uncertain.

2.9. Cosmological gravitational waves from the big bang

Various sources of gravitational waves from the early universe have been hypothesized. These may be thought of as the gravitational wave analogue of the microwave cosmic background radiation. The cosmic microwave background originated at the epoch of last scattering, at a redshift $z \sim 10^3$ when neutral gas first formed in the universe. Thus, the microwave background probes the universe when it was $\sim 10^5$ years old. A similar background of neutrinos should also exist, a relic from their epoch of last scattering, about 0.1 s after the big bang, at a redshift $z \sim 10^{10}$. Due to the weak coupling of gravitational waves with matter, their epoch of release would have been much earlier still, at around the Planck time $\sim 10^{-43}$ s, or $z \sim 10^{30}$.

Thus primordial gravitational waves offer the tantalizing possibility of probing the universe very near to the moment of creation. Unfortunately, we do not have accurate predictions about their amplitude. It has been suggested that the background could have been parametrically

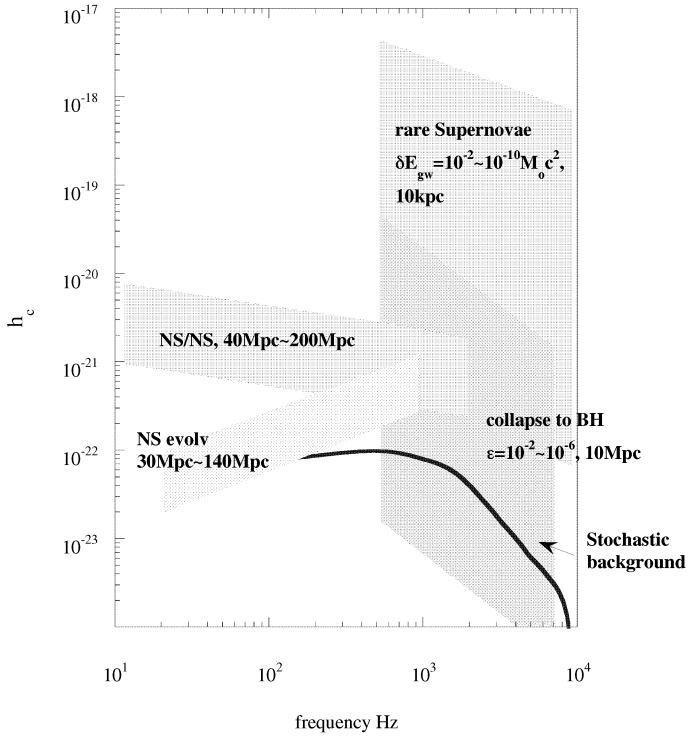


Figure 8. Spectrum of gravitational wave sources [18, 22]. In this figure, the abbreviations are: BH, collapse to black hole; NS/NS, neutron star coalescence; NS evol, secular evolution of a nonaxisymmetric neutron star.

amplified during a period of inflation, or that phase transitions in the early universe (for which there is no experimental evidence) could have created an enhanced background. If there was no process to enhance the background amplitude, then we need consider simply a thermal background that was in equilibrium at the extremely high energies of the Planck era. The background will then have been redshifted like any other radiation. Today this radiation would be in the microwave regime and have an amplitude $h \sim 10^{-35}$, which is beyond the possibility of detection.

If the universe contained an initial inhomogeneity of amplitude h_g [55], then today it would have an amplitude at frequency f given roughly by

$$h \sim 10^{-20} h_g / f. \tag{2.12}$$

We have little idea of the initial amplitude, except for limits set by the cosmic microwave background which implies that inhomogeneities traced by matter had an amplitude $\sim 10^{-5}$. This would imply that the cosmological background amplitude could be 10^{-28} at 1 kHz, (which is beyond terrestrial experiments) and 10^{-21} at 10^{-4} Hz, (which is experimentally accessible by space laser interferometers).

A constraint on the cosmological background is set by considering the energy density, and relating it to cosmological models. Thus cosmological backgrounds are often parametrized in terms of the closure density fraction Ω_g. If the spectrum contains equal energy in each decade, it has a slope of -1 on a log h–log f plot. For example, for the universe to be closed by gravitational waves, $\Omega_g = 1$, the amplitude h_g would be 10^{-13} at 10^{-5} Hz, falling to 10^{-21} at 1 kHz. For $\Omega_g = 10^{-6}$, h_g falls by three orders of magnitude to 10^{-16} at 10^{-5} Hz and 10^{-24} at 1 kHz.

It would appear most unlikely that the universe be closed by gravitational waves, and it would be surprising if gravitational waves contributed significantly to the missing mass. The binary pulsar, in addition to confirming the emission of gravitational waves, also allows limits

to be set on the stochastic background of gravitational waves, and on the time variation of the gravitational constant. Taylor and co-workers [56] have used pulsar timing of the binary pulsar PSR 1913 + 16 to set limits on cosmological backgrounds. The method only works at very low frequencies where both pulsars and atomic frequency standards have sufficient frequency precision that measurement of the arrival time variations of pulsar signals gives useful sensitivity to gravitational waves. Taylor's timing measurements over several years set a limit to $\Omega_g \sim 4 \times 10^{-2}$ in the frequency range 10^{-9} to 10^{-12} Hz [56].

3. Detection of gravitational waves

3.1. An overview of detector technology

The development of gravitational wave detectors was pioneered by Joseph Weber in the early 1960s [4]. He used a massive aluminium bar as the antenna. Following his work, researchers all over the world have been working hard to build different types of gravitational wave detectors. The detection of gravitational waves is based on the following idea (as discussed in section 1.2). A gravitational wave can be considered as a time-dependent strain in space having two linear polarization states (h_+ and $h_\times$). When the wave passes test masses in space it will cause displacements of the test masses, as shown in figure 1. A measurement of the relative displacements of the test masses is a measure of the wave. The gravitational wave does work on electromagnetic field, such as a capacitance or a laser light field. Because the displacements are very small, the momentum of the gravitational wave is in general limited by the uncertainty principle. The *quantum limit* presents a significant but not insurmountable barrier for future detectors.

A gravitational wave detector can be constructed from a pair of masses which can move 'freely' with respect to each other. They can be suspended as pendulums, so that in the horizontal direction they can be treated as nearly free masses above the pendulum resonant frequency. A pair of masses connected by a spring (figure 9(*a*)) can also be used to form a *resonant* gravitational wave detector. Such a mechanical resonator will be driven by a gravitational wave as long as it has spectral components at the resonant frequency of the mass–spring system. If the detector is a high-Q resonator, it will continue to oscillate long after the gravitational wave has passed. That is, the resonator *remembers* the effect of the gravitational wave. A measure of the oscillation of the resonator will give information about the passing gravitational wave.

The resonator need not necessarily be two masses connected by a spring. A lump of metal such as a cylindrical bar is well suited to the purpose (figure 9(*b*)). The difference between the bar detector and two point masses with a spring in between is that the bar detector has a set of higher-order resonant modes. However, for the lowest resonant frequency, the bar can be modelled just like two masses connected by a spring, with an effective mass equal to half the mass of the idealized detector. A multi-spring mass resonator (figure 9(*c*)) can be constructed to detect not only the amplitude of the gravitational waves but also the direction of the waves. This leads to the idea of a spherical antenna [34, 57–59], as shown in figure 9(*d*).

All resonant-mass detectors use cryogenic techniques to reduce the thermal noise and to enable the use of high-sensitivity superconducting transducers. A high Q-factor ensures that the thermal noise (which even after locking has a very large amplitude compared with the signal) is manifested as a highly sinusoidal, and hence predictable, waveform. A gravitational wave induces a small change in the amplitude and phase of this waveform. Typically, the memory effect is used to resolve the signal component averaged over perhaps a few hundred cycles of the antenna. The time over which the signal can be resolved depends most strongly on

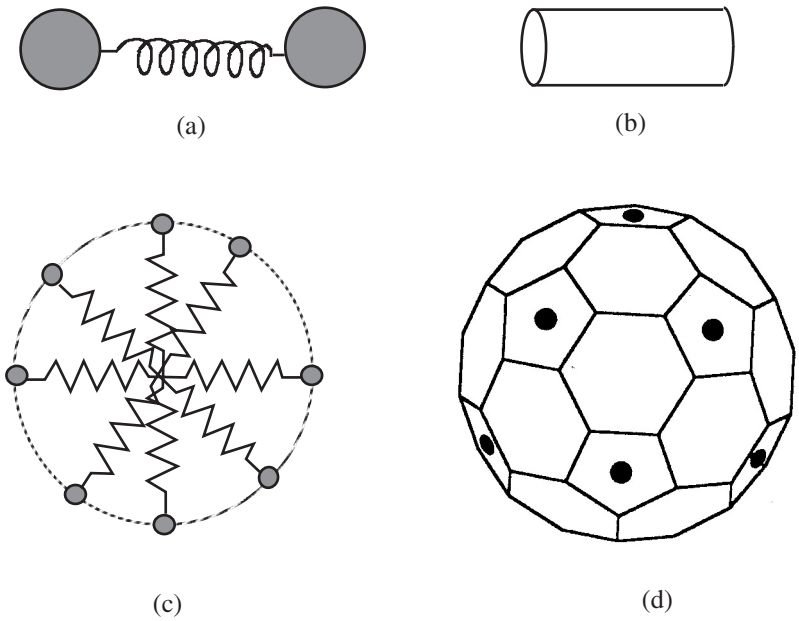


Figure 9. Model of resonant-mass gravitational wave detectors. (*a*) Two masses joined by a spring. (*b*) Bar detector. (*c*) Multi-spring mass detector. (*d*) Spherical detector, in this case formulated as a truncated icosahedron, with six transducers located as indicated.

the sensitivity and electromechanical coupling of the transducer which is used on the antenna. If the transducer is sufficiently strongly coupled, the signal could be resolved in perhaps a single cycle, and then the measured bandwidth would be comparable to the antenna frequency. Practical bar antennas to date have only demonstrated bandwidth of about 10 Hz. However, improved transducers should soon allow the bandwidth to reach $\sim$100 Hz and eventually near-quantum-limited sensitivity [60–63].

As indicated in figure 1, a passing gravitational wave will make a ring of particles stretch and shrink alternately in orthogonal directions. An interferometer configuration which can detect the relative position change between two orthogonal masses is a natural detector configuration [64–66], as shown in figure 10. When a gravitational wave passes, the lengths of the two arms of the interferometer change in antiphase. This results in a change in the interference intensity at the output. This change in light intensity is a measure of the gravitational wave. Since the test masses simply follow the passing wave pattern, these devices can be expected to accurately trace the waveform. The advantage of this type of system is the nonresonant feature and the fact that the scale of an instrument is set, not by the velocity of sound (which limits a resonant bar to a few metres in length if the detector is to detect radiation at $\sim$1 kHz), but by the velocity of light. Thus, a laser interferometer can detect gravitational waves over a wide band of frequencies and can in principle be scaled up to 150 km arm length, (for the same 1 kHz upper frequency) for which the absolute displacement $\Delta L = hL$ is much larger. Because 150 km arm length is impossible on the surface of the Earth (due to the curvature of the Earth), practical instruments must be scaled down to several kilometres.

In reality, the problem of noise always limits the bandwidth of laser interferometers. In their widest bandwidth configuration laser interferometers are limited by seismic and thermal noise at low frequency and photon shot noise at high frequency, as we discuss in detail in

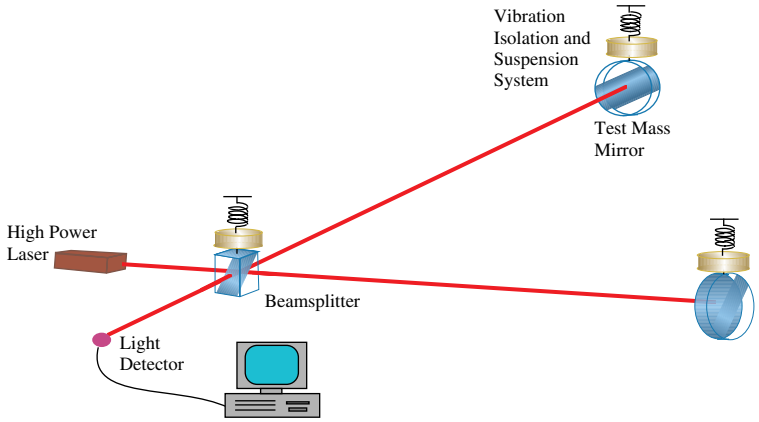


Figure 10. Schematic diagram of a Michelson interferometer for use as a gravitational wave detector.

section 5. This leaves a high-sensitivity bandwidth in the 100–500 Hz range. Optical tuning can in principle be used to restrict the bandwidth considerably, allowing optimum sensitivity to be achieved in narrower bands of several tens of Hz anywhere in the 10 Hz to several kHz range.

There is not a similar size constraint if a laser interferometer detector is placed in space. Laser interferometers can be conceived, of scale size millions of kilometres, in Earth or solar orbit. In this case practical considerations make such devices best suited to search for low-frequency sources in the range of 10^{-1} to 10^{-5} Hz. Free-floating spacecraft carrying test masses shielded from the solar wind, and low-power lasers should be able to achieve very high sensitivity, limited eventually by the confusion limit of gravitational wave 'noise' from the large number of binary star sources in the Galaxy (see section 3.2 for further discussion).

At even lower frequencies, signals from pulsars have been used as gravitational wave detectors. Pulsars, and especially millisecond pulsars, represent precise-frequency sources, close to the limit of the best man-made clocks. While we are accustomed to thinking of gravitational waves changing the relative spacing of test masses, this picture can be confusing when thinking of a pulsar as one of the test masses perhaps 1000 light years away. It is easier to consider the pulsar beam passing through the curved spacetime due to the passage of very-low-frequency gravitational waves in the vicinity of the Earth. (Both pictures are equivalent however.) The result is that the gravitational wave causes changes in the arrival time of the pulsar signal. Because the pulsar signal is weak, and because atomic clocks give best precision over long periods of time, the optimum precision of this method of detection occurs for frequencies $\sim 10^{-7}$ to 10^{-8} Hz [56].

Interferometric detection has the advantage that it gives intrinsic immunity to laser frequency noise. Indeed, a laser interferometer can in principle use white light. However, any single-beam detector is sensitive to frequency fluctuations in the source. The sensitivity limit is set directly by the frequency stability of the radiation source: $h \sim \Delta f/f$. Since frequencies can only be compared against a standard, the limit of a single-beam detector such as a pulsar signal is either the stability of the pulsar, or the stability of the frequency reference: today $\Delta f/f \sim 10^{-16}$.

The pulsar timing technique can to some extent avoid this single-beam clock stability problem by using several pulsars in different directions. A gravitational wave creates correlated fluctuations depending on the pulsar direction so that in principle it should be possible to dig at least an order of magnitude below the clock stability limit. Unfortunately, solar wind refractive

effects and interstellar scintillation are very strong in the MHz–GHz frequency range of strong pulsar signals. This provides an additional noise source that could mimic a gravitational wave.

There is no need in principle to rely on pulsars to provide the timing source for single or multibeam detection of gravitational waves. Interplanetary spacecraft generally transpond a ground-station-generated precision clock frequency back to the ground station. Comparison of the return signal frequency and the transmitted frequency predominantly records the spacecraft velocity, seen as a Doppler shift. Gravitational waves will appear as a perturbation in the Doppler tracking signal. Because the signal strength is stronger than for pulsars, and the path lengths are shorter, Doppler tracking is sensitive in the 10^{-4} to 10^{-5} Hz band. Solar wind refraction is again a very serious limit, but in this case it could be overcome if tracking took place at various frequencies, and if the frequencies used were increased into the millimetre wavelength range. Doppler tracking experiments [67, 68] have taken place successfully as low-cost add-ons to space missions, but the technique is unlikely to compete with space laser interferometers. Future interplanetary spacecraft equipped with nanosecond pulsed lasers avoid all local refractive errors from solar plasma, but solar wind buffeting causes additional noise which must be overcome by drag-free spacecraft technology.

3.2. Space laser interferometer gravitational wave detectors

A joint NASA–ESA space mission has been proposed, to place into solar orbit a set of three laser transponding spacecraft which would form a space laser interferometer for gravitational wave detection. The idea is to create a nested set of interferometers in a triangular configuration, and to use active transponding rather than passive interferometry as used in terrestrial detectors. The space environment allows the path length to be increased to 5 million kilometres. Consider first one vertex of the triangle. A pair of stabilized CW laser beams are transmitted through telescopes in 60° Y-shaped arms of the spacecraft, to two identical target spacecraft 5 million kilometres away. Each receives a very weak signal, but one sufficient to allow an on-board laser to be coherently phase locked to the incoming beam with a slight offset frequency. This laser is then directed back to the originating spacecraft. Thus the distant spacecraft acts like an active mirror, returning the incoming signal to its origin.

LISA (laser interferometer space antenna) as proposed uses three spacecraft in a special solar orbit. All six laser beams create three independent Michelson interferometers. The incoming beam at each arm has its phase compared (by beating) with some of the outgoing signal. That effectively measures the changes in the length of one arm. The same measurement is done in the second arm by the second laser beam. The phase differences are compared, to create a monitor of the arm length differences which could indicate the passage of a gravitational wave.

The above concept can give excellent sensitivity, as shown in the sensitivity curve for LISA in figure 11. This is only achieved if many noise sources are greatly reduced. The first is the buffeting by the solar wind. To overcome this noise the spacecraft are centred around a free-floating test mass, which is shielded from the wind. The spacecraft are controlled by miniature ion drives—field emission electric propulsion thrusters—to maintain the spacecraft location centred on the test mass. The forces required are $\sim 10^{-6}$ N.

A second potential source of noise is the gravity gradient due to the thermal expansion and contraction of the spacecraft structure. This can be overcome with careful design and thermal shielding.

To be able to use low-power thrusters the spacecraft must be placed in an orbit in which their relative positions have very high intrinsic stability. An orbit that achieves this requirement [69] is a heliocentric orbit of about 1 AU, 20° behind the Earth. The orbits of the spacecraft have a small eccentricity of $e = d/(D\sqrt{3})$, and a small inclination $i = d/D$, where d is the

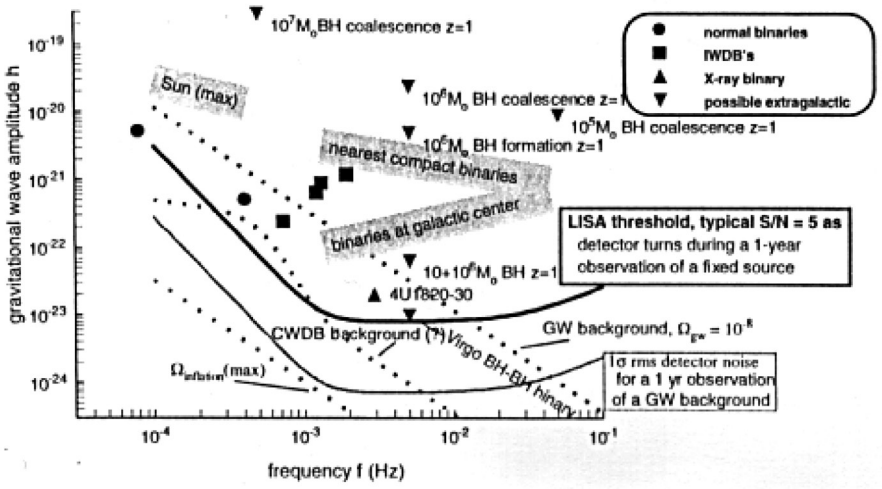


Figure 11. The sensitivity curve of LISA, along with its prime gravitational wave sources [322]. This is the sensitivity achieved after one year of integration.

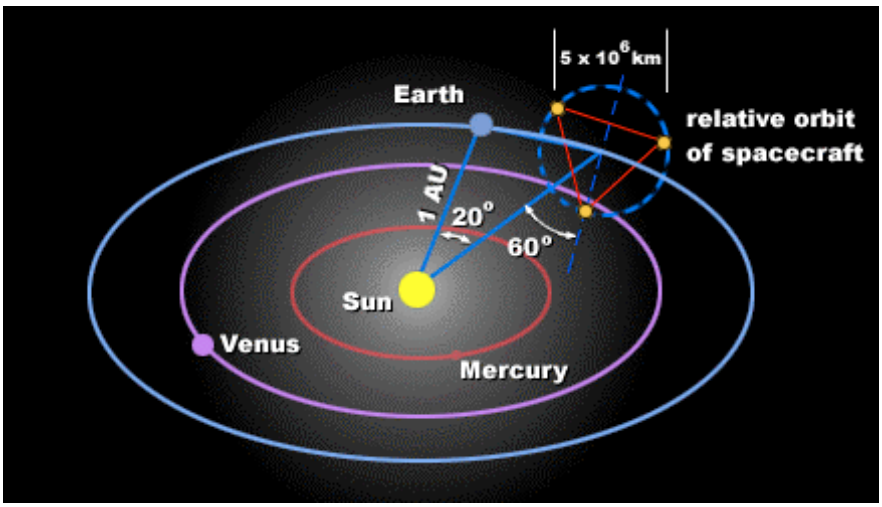


Figure 12. The positions of the proposed LISA spacecraft in heliocentric orbit [323].

triangle arm length and $D = 1$ AU is the semimajor axis of the orbit. Remarkably, three pairs of spacecraft in these orbits, with careful specification of their nodes, appears to maintain a nearly 'rigid' configuration which rotates slowly while maintaining the triangle in a plane which is inclined at 60° to the ecliptic. The positions of the spacecraft are shown in figure 12.

An important question is the orbital stability. If the relative spacing of the satellites changes too rapidly the fringe rate becomes high and the noise contribution from the local oscillator (which is used to apply frequency offsets and measures the fringe rate) becomes relatively larger. Thus it is essential that the fringe rate be low. Ideally the difference in arm length changes between two arms should be reduced to mm s^{-1}, less than the nominal metres per second arm length changes predicted for the proposed orbits. Stabilization schemes for the

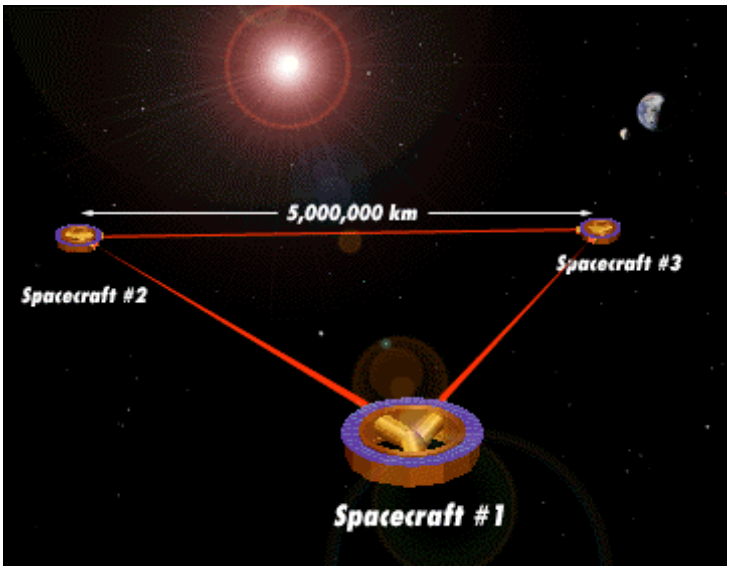


Figure 13. Artist's impression of space interferometer (LISA).

local oscillator through modulating the light in one arm or through the use of improved local oscillators should allow these problems to be overcome. Spacecraft manoeuvres are required occasionally to compensate for accumulated orbital drift.

Figure 13 shows an artist's impression of this remarkably ambitious conception. A large team of scientists is pursuing this project, which should fly in the second decade of the twenty-first century. More details are available from the LISA web page (http://lisa.ipl.nasa.gov/).

3.3. The world array of resonant-mass detectors

The improvement of resonant-bar detectors since they were first reported by Weber has been enormous. Their amplitude sensitivity has been increased several hundred times, corresponding to an improvement in flux sensitivity of about five orders of magnitude. Current detectors are sensitive to narrow frequency bands near 700 Hz and 900 Hz, but improvements will increase the bandwidth of each to >50 Hz. Figure 14 shows one of the present resonant-bar detectors.

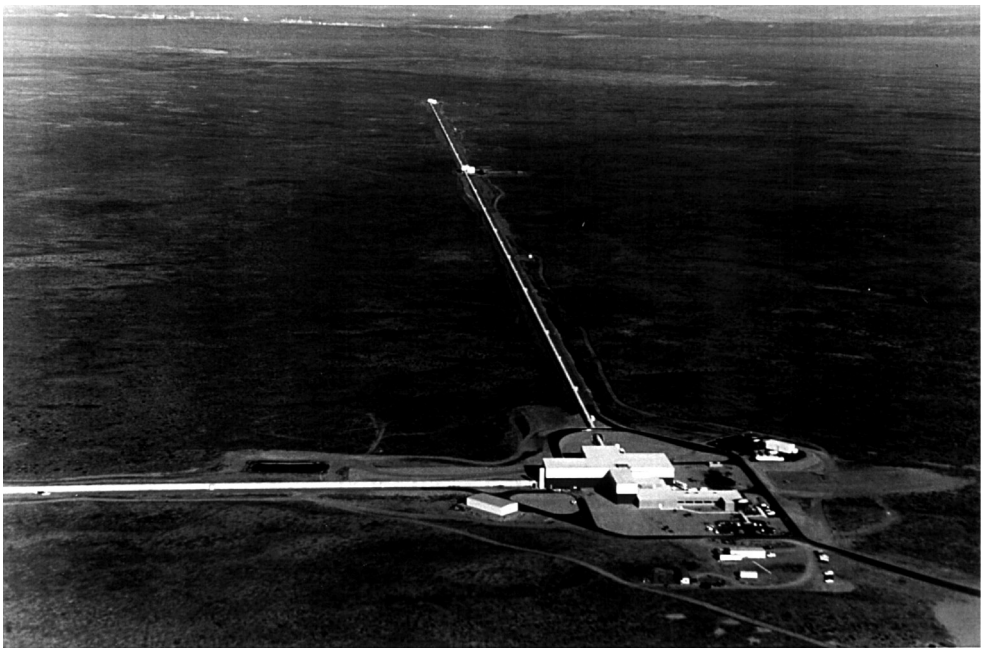
An array of five resonant-mass gravitational wave detectors, coordinated under the International Gravitational Events Collaboration is in operation [70, 71, 110, 114, 156]. The detectors are located at Baton Rouge (Louisiana), CERN (Geneva), Legnaro (near Padova), Frascati (near Rome) and Perth (Australia). The data are available at a web address (http://axln01.lnl.infn.it/igec/). Since 1993 between two and four antennas have been in coincident operation searching for bursts at a strain sensitivity better than 10^{-18} (see figure 16). This is sufficient to detect strong galactic gravitational wave bursts, but insufficient for detection beyond our Galaxy. Over the past two decades the limits to the strength and rate of gravitational wave burst events impinging on the Earth has been reduced substantially, but these limits are still below astrophysical predictions. Thus, so far, it can be stated that the rate of gravitational wave bursts is not two orders of magnitude greater than expected from conservative astrophysical predictions, or else that their strength is not at the high end of predicted signal strengths ($\varepsilon > 10^{-2}$, see figure 29 for quantitative results).

Figure 14. Photograph of the antenna NAUTILUS at Frascati, showing the bar and its cryogenic shields.

The resonant bars have been used for rather deep pulsar searches in certain directions ($h \sim 10^{-23}$) [71] (see figure 30) and, by using cross-correlation, have been used to set impressive limits on the stochastic background of gravitational waves ($h \sim 10^{-22}$) [84].

Some of the resonant-bar detectors are being improved with better transducers and amplifiers, allowing their bandwidth to be increased towards 100 Hz. This improves the burst sensitivity and time resolution which in the immediate future should allow an order-of-magnitude improvement in burst sensitivity. Efforts are underway to create quantum-limited transducer systems (e.g. [72]) which should eventually allow a further order-of-magnitude improvement in amplitude sensitivity.

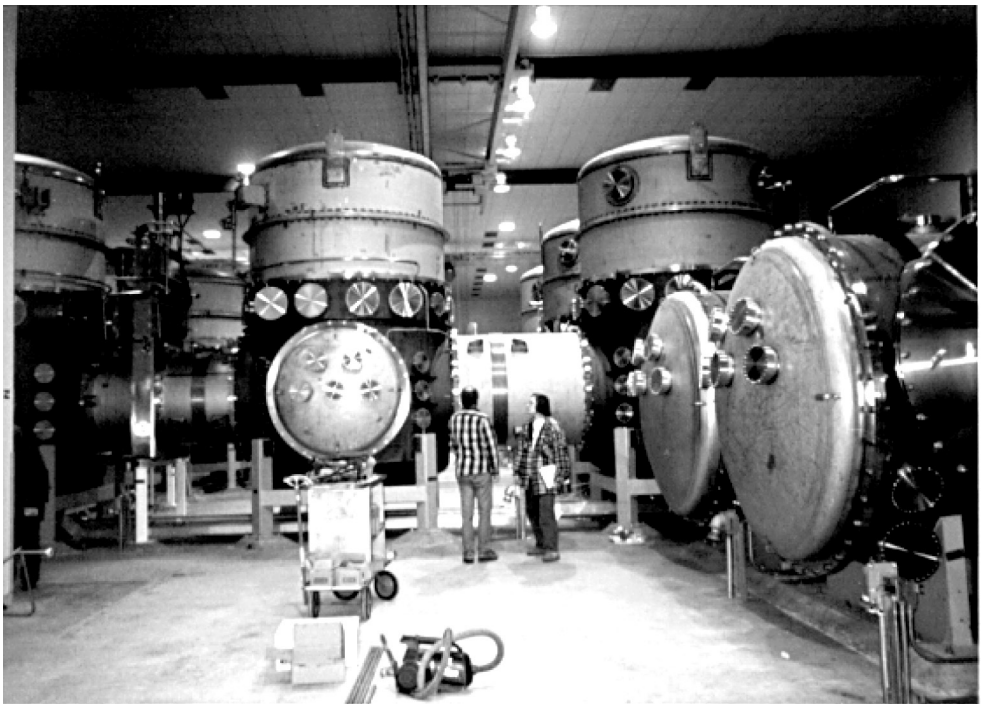
To match the ultimate sensitivity of long-baseline laser interferometer detectors (see below), it is necessary to increase detector mass from a few tonnes to a few hundred tonnes. Such massive spherical detectors have been proposed and development work is underway on small prototypes in Frascati, Leiden, and Sao Paulo [73–75]. As already noted they have the advantage of omnidirectional sensitivity. They use the proven cryogenic techniques of the existing resonant-mass detectors, but to scale up to hundred tonnes represents an exciting major engineering challenge (see section 4.10 for more details).



(*a*)

(*b*)

Figure 15. The LIGO detector at Hanford. (*a*) An aerial view, (*b*) the vacuum pipe tunnel and (*c*) the main cornerstation.



(*c*)

Figure 15. (Continued)

3.4. Laser interferometer detectors

Three long-baseline laser interferometer gravitational wave detectors have been or are being constructed. The US LIGO Laboratory consists of two 4 km arm length detectors, at Hanford, Washington State, and Livingstone, Louisiana (http://www.ligo.caltech.edu). The Italian/French VIRGO project is completing a 3 km arm length instrument at Casina, near Pisa (http://www.virgo.infn.it). Smaller detectors are under construction at Hannover (the German/British GEO project, 600 m arm length, http://www.geo600.uni-hannover.de), Tokyo (TAMA 300 m arm length, http://tamago.mtk.ano.ac.jp) and Perth (80 m arm length, http://www.gravity.pd.uwa.edu.au). Figure 15 shows some photographs of the LIGO detector in Hanford: (*a*) an aerial view, (*b*) the vacuum pipe tunnel and (*c*) the main cornerstation. The long-baseline laser interferometer detectors are initially expected to achieve sensitivity as shown in figure 16.

Laser interferometers are complex instruments limited by a range of noise sources: internal thermal noise in the mirror test masses, seismic noise, radiation pressure noise, laser frequency noise, control system noise, residual gas refraction noise etc. All noise sources must be reduced as far as possible to allow the detectors to achieve high sensitivity. The first decade of the twenty-first century will see steady improvement of the detectors. Late in that decade it is likely that major improvements to the detectors will be possible using improved lasers, improved test masses and improved vibration isolation. Figure 16 shows the expected improvements.

Ultimately, both laser interferometers and resonant masses can be improved by using various quantum measurement techniques. Laser interferometers can in principle also be improved by use of cryogenic methods, and by increasing their arm length.

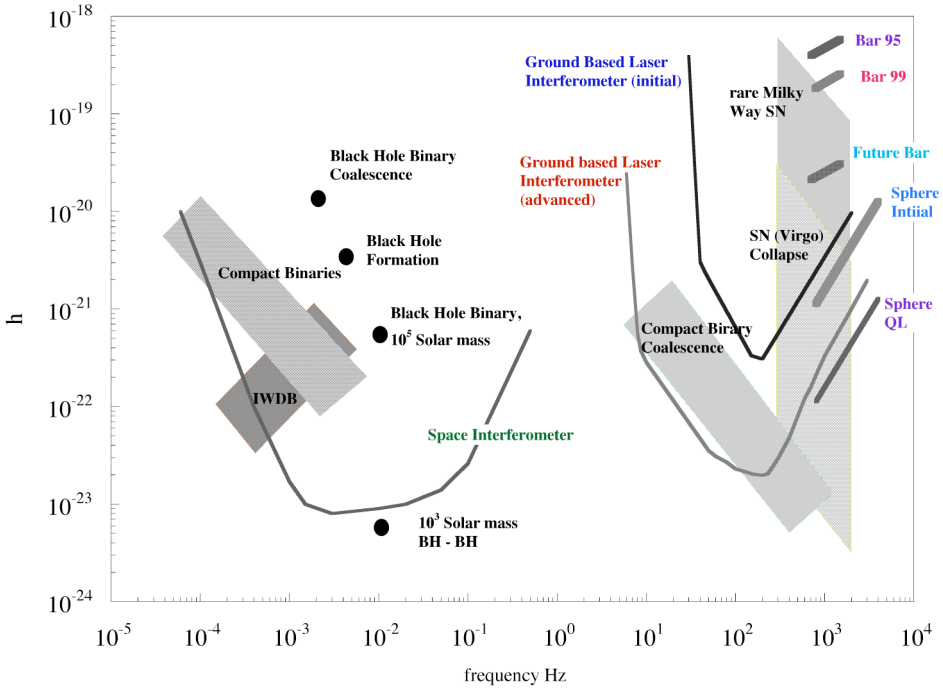


Figure 16. Rough comparison of various detector technologies and some sources. Note that the sensitivity depends on the type of signal searched for: the resonant-mass sensitivities quoted refer to burst sources (e.g. the narrow-band sensitivity of bars has already been demonstrated at 10^{-23}). At low frequencies interacting white dwarf binaries (IWDB) and neutron star binaries are known sources. At high frequency the only known source is neutron star binary coalescence: all the others have unknown signal strength.

All gravitational wave detectors provide data dominated by noise. All face difficult data processing challenges. In the following we shall briefly review some of the main issues in the context of the various classes of gravitational wave sources, before going on to examine resonant-mass and laser interferometer detectors in more detail.

3.5. Issues of data processing and signal detection

There are many important and still unresolved issues of data processing which it is necessary to solve to enable detection of signals to the levels anticipated from a simple analysis. These issues affect burst detection, stochastic background detection and narrow-band detection. Since the general analysis principles are independent of the type of detector used, we will discuss some of them in generality.

The raw data from a detector must first be filtered. In section 2.2 we discussed the use of *matched* filters or optimal filters to extract particular signals. An *optimal filter* is one which maximizes the signal-to-noise power ratio for a particular signal waveform. The term matched filter arose because in the special case where the noise is white, the optimal filter is one with an impulse response which *matches* the shape of the input signal. In these ideal circumstances the matched filter is simply a template of the waveform one desires to detect.

Think of the data from a detector over a particular period of time, due to some input signal. In a perfect noise-free detector it would create an output signal $s(t)$ which is uniquely

determined by incoming signal and the detector impulse response. After sampling, a real detector output consists of a two-dimensional array of output values o_i and time values t_i: $\{o_i, t_i\}$. But $o_i = s_i + n_i$ is a sum of the input signal s_i and the noise n_i. The template is a noise-free array representing $s(t)$, but it has an arbitrary phase, so can be represented $\{s_j, t_j\}$ where the difference between i and j represents a time difference. When the template is aligned with the signal $i = j$ and the sum of the products $o_i s_i (= s_i^2 + s_i n_i)$ over all values of i will be maximal. At times when s_i is zero there is no noise contribution. The template which accurately matches the ideal signal response of the detector is therefore the optimal filter function.

In practice one could apply a large calibration signal and measure the output signal to create such an optimal template.

Often it is more convenient to consider the problem in frequency space. In frequency space the phases of the noise frequency components are random and uncorrelated. If you multiply the Fourier transform of the output data $O(f)$ with the complex conjugate of the Fourier transform of the signal waveform $s^*(f)$, the signal component will be positive definite while the noise phases remain random. Summing over frequency space, the signal will add up but the noise frequency components will tend to cancel. Thus, in frequency space the matched filter transfer function is simply the complex conjugate of the Fourier transform of the input signal $s^*(f)$.

In almost all real situations the noise is not white. However, coloured noise can be whitened by passing it through a filter with a transfer function equal to $S_n^{-1/2}$ where S_n is the noise power spectral density of the noise. Thus, the optimal filter transfer function is simply $s^*(f)/S_n^{1/2}$. In the time domain the same correction for coloured noise is made using the noise autocorrelation function whose Fourier transform is S_n.

All optimal filters require a sum over frequency or time. If the signal is transient, then the sum will be zero after the signal has passed and the detector has stopped responding. For a resonant bar this will be several ring-down times of the bar (after which all memory of the signal is lost). In the same time *before* the signal arrives the bar loses memory of its instantaneous mechanical state. Thus it is only necessary to integrate over a modest time interval before and after the signal arrives. The Louisiana State University (LSU) group have implemented a time domain optimal filter by applying a large 'signal' pulse using a calibrator, and measuring the detector response, as well as the detector noise in the absence of a signal. Other bar groups have used frequency domain filters and obtained similar results [76].

There are many ways to implement optimal filters and the best choice often involves minimization of computation requirements for the particular search. For some systems optimal filters can be implemented in quasi-real time whereas for others the need for prior data requires the search to be conducted off-line.

A particular issue is that of noise stationarity. In practice, stationarity of the noise is not a good approximation. This means that the noise spectrum at the output varies with time, due, for example, to environmental effects such as variable microseismic noise from changes in the ocean wave conditions [77]. To overcome slow changes in the detector noise distribution it is possible to always use recent noise data for the creation of the optimum filter. Several groups use noise from the previous few hours to continually adapt the filter [78].

Another problem relates to the fact that in gravitational wave detectors it is impossible to turn off the signal, and possible signal-like noise events (such as cosmic ray events). Thus, in principle it is impossible to measure the noise spectrum in the absence of signal. In practice today this is not a critical issue (signals are rare and very small), but for future detectors it may be important to ensure that the filter does not suppress signal by confusing it with noise. Heng has shown [79] that periodic transient bursts are indeed suppressed by an optimum filter if the

bursts are present during the measurement of the noise distribution.

For CW signals, a Fourier transform of the entire data set represents an optimal filter. By this method the signal appears as a narrow spectral line. If the Earth's orbital and rotational Doppler corrections are included this method can be used to search for unknown CW signals (such as those due to isolated spinning neutron stars) in a particular direction (because the Doppler correction is direction dependent, see figure 30).

Traditionally the lock-in amplifier or phase-sensitive detector (PSD) has been used to create a matched filter operating in the time domain for CW or near-CW signals. For modern systems this type of filter can be implemented by software (hence not in real time) which allows it to be used for multiple searches through the same piece of data. It also allows CW searches with arbitrarily long integration times. (The traditional analogue lock-in amplifier is limited by the charge storage time of a capacitor.)

The lock-in amplifier form of the matched filter multiplies the data by +1 during the positive going signal cycle, and -1 during negative parts of the cycle. All results are summed. The sum represents the phase coherent integral of the absolute value of the signal over the observation time: hence the alternative name for this method: coherent integration. The noise components average towards zero, and a small signal component previously buried in the noise will emerge.

Assuming that the phase of the incoming signal is unknown, it is necessary to repeat the matched filter at 90° phase shift to determine the magnitude of the orthogonal components of the signal. (For example, if the signal was a pure sine function, the matched filter would give a zero output if the +1 and -1 multipliers were defined by a cosine function.) Clearly, if there were slowly varying phase errors due to errors in the timing of data acquisition, errors in direction for Doppler corrections, or errors in the prediction of the waveform, the accumulated result from a matched filter could average to zero.

For a CW source of fixed but unknown frequency it is usually simpler to replace matched filtering with the fast Fourier transform. However, if the frequency is modulated in a known pattern (such as occurs when you search for a signal from a known binary pulsar) [80] the coherent integration is computationally simpler.

Attempts at gravitational wave burst detection have normally followed techniques first introduced by Weber. By this method the data is first filtered, as discussed above, and then thresholded to obtain a list of candidate events. Some events will be due to spurious effects (see below). These can be vetoed if the appropriate monitor channels are used. Then the event lists for widely spaced pairs of detectors are correlated to search for coincidences. If the time axes for the event lists are randomly displaced one expects all correlation due to possible gravitational wave bursts to be absent. Thus it is possible to compare true-time coincidences with time-shifted coincidences to determine whether there is an excess of 'zero time delay coincidences'. This method is powerful since it allows the probability of accidental coincidence to be experimentally determined by simply running the random time shift analysis a sufficient number of times to obtain an accurate estimation of the probability of the observed peak. However, it is only relevant when the number of accidental coincidences is large.

Consider, for example, a six-month coincidence run which might yield 5000 'candidate events' in each detector, most of which are assumed to be noise or interference. The coincidence analysis might yield 30 true-time coincidences. For such analysis with existing resonant bars the coincidence window is usually more than 0.1 s, so that propagation delays across the Earth ($\sim$40 ms) can be neglected. Then 1000 random time delays are applied to the data of one detector and of these 10 might show 30 or more coincidences. If this was the case, then the zero time delay peak would have a probability that it was accidental of 10/1000 or 1%. This probability could be resolved with more accuracy if more random time delays were used. In

this type of experiment the time delays which are chosen must either be small compared with total duration of the experiment, or they should be modulo the experiment duration so that coincidences are not lost at the ends of the record.

A serious pitfall can occur in this type of analysis if the threshold for candidate event selection, or any other vetoing scheme, is variable during the coincidence experiment. To have such a variable accessible to the persons analysing the data can allow conscious or unconscious selection which can completely invalidate the statistical significance analysis. It can be very difficult to estimate the true significance after such selection has taken place. Since the days of Weber, researchers have been in general very careful to pre-set all thresholds to try to avoid such pitfalls.

For the existing array of five resonant-mass detectors it has been suggested that the entire body of data from all detectors should be used rather than candidate events. As discussed in the next section, the accidental coincidence rate is extremely low for five detectors. The thresholding method means that all phase information is lost. However, it need not be lost if the data from separate detectors are recombined along the lines used by VLBI radioastronomy. If this was done the detectors could represent a single telescope with angular resolution for incoming bursts set by the ratio of the time resolution (in principle $\sim$100 μs) and the near-Earth-diameter baseline (40 ms). However, the noise increases exponentially as one digs to lower and lower energy so the overall amplitude sensitivity of the array would be limited to $\sim$3 times.

Most data from terrestrial gravitational wave detectors can be idealized as the sum of a pair of thermal distributions. The first is the intrinsic Boltzmann distribution of the detector noise that one would measure using a spectrum analyser. This is usually due to some well understood noise sources such as thermal noise, electronics noise or shot noise. The second distribution is described as excess noise. Excess noise arises from rare and poorly understood sources. In the detector NIOBE at the University of Western Australia (UWA) some of the excess noise was correlated with electromagnetic pulses and seismic noise [81]. However, the majority was not identified, but might be due to strain relief events. Surprisingly, the excess noise distribution is rather similar for widely differing types of resonant-mass gravity wave detector. Figure 17 shows the idealized form of these distributions. Both may be expressed as

$$N = N_0 \mathrm{e}^{-E_e/kT}. \tag{3.1}$$

For a typical resonant-bar detector the intrinsic distribution parameters would be: $N_0 \sim 10^5$ per day, and $E_e \sim$ few millikelvin (Kelvin is simply a convenient energy scale). The excess distribution typically would have $N_0 \sim 10$ per day, and $E_e \sim$ few 100 mK: that is, the excess noise distribution has an event rate 10^4 times lower than the thermal distribution, and an effective energy 100 times larger. The presence of these two distributions allow improved data analysis, as discussed in section 4.8.

Periodic signals can be detected by coherent integration or Fourier transform methods. Because, as always, signals are near to the limit of detection sensitivity, long integration is needed. However, in long integrations signals will be smeared out by the Earth Doppler motion unless the source direction is known. Equally, if the source is a member of a binary system such as a binary millisecond pulsar it is necessary to know the ephemeris of the system itself as well as the precise source direction to prevent Doppler smearing of the signal. If coherent integration can be achieved, then the signal-to-noise ratio improves as $\tau^{1/2}$, which can allow very deep searches over a 1–3 year period.

The nearest millisecond pulsar PSR 0437-4715 is a typical such potential source. In this case the pulsar source is a binary system, which is extremely well defined by pulsar timing [82]. The observed radio pulse timing gives very accurate information which can be used to gate the

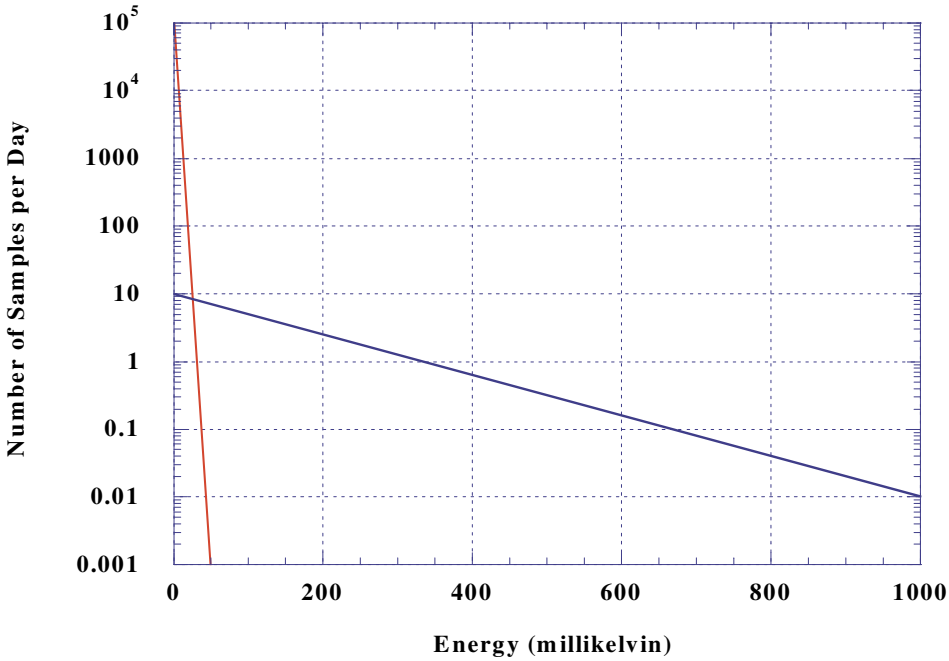


Figure 17. Detector noise is often characterized by a pair of Boltzmann distributions. The steep curve is the typical antenna thermal noise distribution for burst detection in a resonant bar. The flatter curve is the excess noise distribution due to rare disturbances.

phase-coherent integration of the detector output [80]. Such methods were pioneered by the Tokyo group [83] led by Hirakawa searching for gravitational radiation from the Crab pulsar.

It is likely that many such narrow-band sources will not be detected in advance by radio astronomy. Pulsar beaming ensures that the majority of pulsars are not visible in radio, and normal pulsars older than 10^8 years appear to cease to emit radio beams. There could easily be 1000 rotating neutron stars within the range of the nearest observed pulsars ($\sim$100 pc) meaning that a few old neutron stars could exist within 10 pc of the solar system.

Unfortunately, our ability to search for such sources is very poor because of the difficulty of searching all directions in the sky. The most obvious search strategy involves applying a separate ephemeris correction for each of typically 10^{10} source directions, and re-integrating the data 10^{10} times. This already is a daunting computational exercise, but it becomes 10^{10} times harder if the search space has to involve all possible binary orbits of the pulsar. Many investigators are searching for efficient algorithms to solve this problem, based on alternative filtering techniques or hierarchical searches [84].

As discuss in sections 2.2 and 2.8, stochastic signals can be detected by cross-correlation of nearby gravitational wave detectors. For optimum sensitivity the detectors must be located within about one reduced wavelength ($\lambda/2$) of each other. If the spacing requirement is satisfied, the signal-to-noise ratio increases as the $\frac{1}{4}$-power of the number of cycles observed [85]:

$$\frac{S}{N} = \left(\frac{S_{gw}^2}{S_1 S_2} B\tau\right)^{1/4}. \tag{3.2}$$

Here S_{gw} is the gravitational wave background power spectral density, S_1 and S_2 are the spectral

noise density of two detectors, B and τ are the bandwidth and duration of the observation respectively. Thus, a stochastic signal of characteristic frequency $\sim$100 Hz can be cross-correlated to gain a factor of $10^{2.5}$ in signal-to-noise ratio (compared with the observation of a single cycle) after 10^8 s integration. Schutz [85] has shown that this method can be particularly effectively applied to signals in detectors of quite different types: specifically resonant-mass detectors and laser interferometers, which conveniently are planned to be located in sufficiently close proximity to each other to satisfy the above spacing requirement.

The simplest cross-correlation experiment merely gives an output consisting of a single number (and a measure of its statistical significance). This can be quite misleading since there are many ways that cross-correlation can give a false positive result. The major difficulty in cross-correlation signal processing is to ensure that no correlated technical noise components exist in the signals. Remembering that the signal detected is generally going to be at least 100 times lower than the spectral noise floor accessible using normal fast Fourier transform diagnostic instrumentation, correlated features can exist which could never be detected in short-term monitoring. The correlated features could be weak spectral lines such as those created by electrical power harmonics, (which are phase coherent over the electrical grid) or occasional transients such as those due to lightning flashes. To prove that a positive cross-correlation signal was associated with gravitational waves it would be necessary to show that the correlation was preserved across the accessible frequency band, and that it was not due to intermittent transients due to electrical interference. The individual output spectra of the detectors would need to be resolved for weak spectral features to the same depth as the correlation detection threshold. Much effort at developing algorithms and solving some of these practical problems is underway [53, 86].

Giazotto [87] has shown that the stochastic background signal from the combined old radio pulsars in our Galaxy should be detectable in a single detector, due to the fact that they are non-isotropically distributed relative to the solar system. The central concentration of pulsars means that there should be a strong sideband modulation of the stochastic background intensity as the detector sweeps the sky during Earth rotation. This signal should appear in the technically very demanding 1–10 Hz frequency band.

The next two sections present resonant-mass detectors (section 4), and then laser interferometer detectors (section 5), with emphasis on techniques and the solution to various technological challenges.

4. Resonant-mass detectors

4.1. Introduction

Resonant-mass detectors are designed to measure acoustic signals induced in a large mass due to its coupling to a gravitational wave. Resonant detectors were first developed by Weber during the 1960s [4]. They consisted of large vibration-isolated aluminium cylinders instrumented with piezoelectric crystals glued on the surface near to the centre. A low-noise amplifier and lock-in amplifier allowed detection of the energy of the fundamental longitudinal resonance of the bar. A gravitational wave applies a time-varying quadrupole deformation and does mechanical work on the bar. The absorption cross section of the bar to gravitational waves depends on its mass and sound velocity. The cross section is highest at the fundamental resonant frequency. The latter is linked to its length and sound velocity, since its length must be half an acoustic wavelength at the fundamental longitudinal resonance. Weber chose aluminium because of its high sound velocity and availability in large pieces, and because it has quite low acoustic losses. Following Weber, many new resonant-mass detectors using similar techniques, but with

variations and improvements, were developed in the early 1970s. Following null results these were abandoned, but advanced resonant-mass detectors using cryogenics and superconductivity continued to be developed. Fairbank *et al* [88] and Hamilton *et al* [89] first proposed such cryogenic detectors, and proposed cooling to millikelvin temperatures to minimize thermal noise. Today two detectors in Italy are in operation at temperatures below 100 mK.

In all resonant-mass detectors the large amplitude of thermal vibration considerably exceeds the gravitational wave amplitude expected from astrophysical sources. Without methods to suppress this noise the principle of detection by resonant masses would be impossible. Weber's key contribution was the realization that in a high-Q antenna—one with a low acoustic loss—the effective noise energy is reduced by a factor $\sim\tau_i/\tau_a$, where τ_i is the effective measurement integration time, and τ_a is the antenna ring-down time. The advantage of using a low acoustic loss antenna follows directly from the fluctuation–dissipation theorem [90]: the greater the dissipation the greater the fluctuations or noise level imposed by the thermal reservoir. A high-Q antenna approaches an ideal harmonic oscillator, whose motion is exactly predictable at a time in the future from the observed amplitude, frequency and phase at an earlier time. High levels of predictability means that very small deviations from sinusoidal behaviour can be resolved given a sufficiently sensitive transducer for monitoring its motion.

4.2. Intrinsic noise in resonant-mass antennas

To understand the operation of a resonant-mass gravitational wave detector it is convenient to start with an old-fashioned approach first introduced by Gibbons and Hawking [91]. This approach is intuitively obvious but is not consistent with the optimal filter theory discussed in section 3. The instantaneous state of the antenna can be described by the pair of symmetrical harmonic oscillator coordinates X_1 and X_2 given by

$$\begin{aligned} X_1 &= A\cos\phi \\ X_2 &= A\sin\phi, \end{aligned} \tag{4.1}$$

where A is the antenna amplitude and ϕ is the phase. Experimentally, X_1 and X_2 can be easily measured using two lock-in amplifiers or PSDs in a configuration shown schematically in figure 18. They may be analogue or digital or software devices. The state of the antenna can be represented by a point P_1 in the (X_1, X_2) plane. The amplitude $A = |P| = (X_1^2 + X_2^2)^{1/2}$ and phase $\phi = \tan^{-1} X_2/X_1$. This is illustrated in figure 19. A gravitational wave causes the antenna to move from P_1 to P_2. The direction of this motion depends on the relative phase of the gravitational wave and the resonant mass. To extract a signal the measuring system should monitor $\Delta|P_1 - P_2| = (\Delta X_1^2 - \Delta X_2^2)^{1/2}$. The quantity Δ^2 is described as the energy innovation and its magnitude, properly calibrated and expressed in the units of Kelvin (1.38×10^{-23} J) describes the effective temperature of the antenna.

In a noiseless antenna the motions of the vector P would only be due to gravitational waves, but in practice P is driven by thermal fluctuations in the bar. Thermal fluctuations cause the state vector P to execute a random walk in the X_1X_2 plane. A high-Q mode is weakly coupled to the thermal reservoir which is made up of all the higher modes of the system. The antenna loses energy slowly into the reservoir, and equally it is only weakly excited by the reservoir. The relaxation time $\tau_a = 2Q/\omega_a$ thus characterizes both the rate of decay after a high-energy excitation and the rate of amplitude change when the mode is in thermal equilibrium with the reservoir.

Clearly, if τ_a is large and the rate of fluctuation is low, the antenna becomes more deterministic on time scales that are short compared with τ_a. The mean energy is still kT,

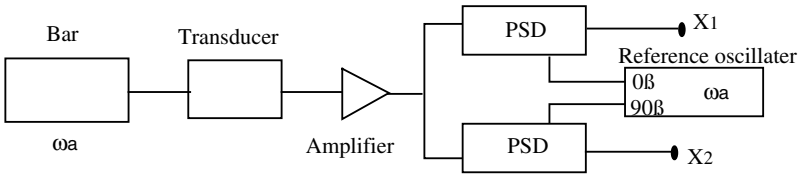


Figure 18. Antenna readout systems for obtaining harmonic oscillator coordinates X_1 and X_2. The down-conversion with the PSD was conventionally done with analogue electronics but today can be achieved in software using fast digital sampling [78].

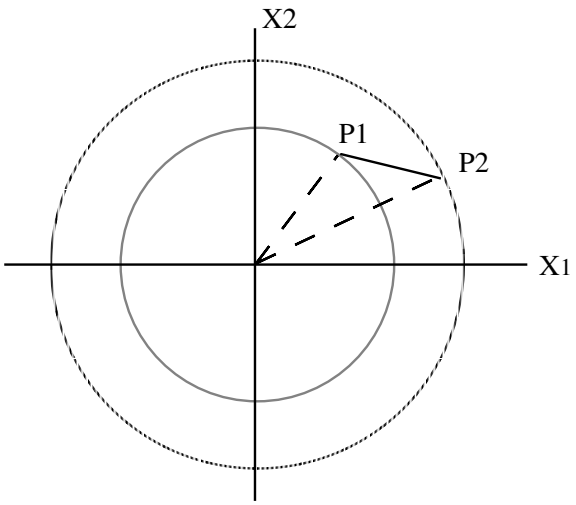


Figure 19. X_1–X_2 representation of the state of the antenna.

but the expected change in energy in a sampling time τ_i is $kT(\tau_i/\tau_a)$. The temperature $T(\tau_i/\tau_a)$ is the effective temperature or noise temperature of the resonator, and quite clearly can be made less than the actual temperature. Indeed, even when transducer readout noise is included, a noise temperature of less than 2×10^{-4} of the thermodynamic temperature has been demonstrated in the detector NIOBE. To use temperature to describe the antenna noise implies that the distribution of Δ^2 has a Boltzmann distribution. This is an excellent approximation (see figures 17 and 28). The slope of the distribution as well as its mean value gives the antenna noise temperature.

The above analysis describes a predictive filter for the detection of short bursts. In this case our prediction is that the amplitude and phase of the detector will remain unchanged over the integration time. Today all operating detectors use optimum filters for the detection of short bursts. The optimum filter improves on the simple predictor discussed above because it implicitly takes into account the dynamics of the system for times of the order of the antenna ring-down time over which the motion is correlated. They are implemented as discussed in section 3 and enable the antenna noise temperature to be improved by a small factor. The most popular optimum filter or matched filter is the Weiner–Kolmogoroff filter which is designed to detect short delta function bursts [92]. Today, such filters are routinely used since they were shown by Pizzella *et al* [93] to give substantially improved performance.

Every antenna must use a sensitive transducer to read out the motion. The transducer, like all electronic devices, has a residual broadband noise floor. This noise floor is equivalent to the Johnson or Nyquist noise of a resistor, given by $V_n^2 = 4kTRB$, where R is the effective resistance of the transducer and B is the measurement bandwidth. However, the bandwidth is roughly the reciprocal of the measurement integration time τ_i. Thus it follows that the noise power is $4kT/\tau_i$. Comparing this with the Brownian motion noise, we see that one noise source varies as τ_i while the other varies as $1/\tau_i$. Thus there is clearly an optimum bandwidth set by the value of τ_i that minimizes the total noise.

Gibbons and Hawking [91] introduced a parameter β to characterize the coupling between the bar and the transducer. They defined β as the proportion of the elastic energy of the detector that can be extracted electrically through the transducer in one cycle. A bar–transducer system with low β (weak coupling) requires more time for the signal energy to appear in the transducer. The longer the energy transfer takes, the more time there is for fluctuations in the antenna to dominate the noise. This point can be clarified by two alternative viewpoints. One is a thermodynamical model. The antenna is considered as a thermal bath at temperature $T_{\text{eff}} = T_a\tau_i/\tau_a$, coupled to a transducer with noise temperature T_t which itself is coupled to an amplifier of noise temperature T_n.

A gravitational wave causes slight 'heating' of the fundamental mode and energy flows through the coupling β. As long as $\beta > 0$ the transducer will eventually come into equilibrium with the bar, but for a rapid response β has to be large. The thermodynamic approach emphasizes that the coupling is not unidirectional: thermal fluctuations in the amplifier or the transducer act back on the antenna producing *back-reaction noise*. Indeed, it is clear that the transducer is a source of thermal fluctuations comparable to those originating within the bar. Voltage noise in the transducer will convert into force noise acting on the bar. Like the forces which act on the bar from the thermal reservoir, the back reaction will produce an additional noise contribution which will diminish as τ_i reduces to zero, as does the Brownian noise.

The second viewpoint is that the antenna–transducer system is effectively a transmission line designed to couple energy into the transducer. One can think in terms of phonons in the bar which may be absorbed by the transducer, with the emission of a photon into the amplifier, or they may be reflected back into the bar. Then β determines the *impedance match* between the output impedance of the bar, Z_{out}, and the transducer's mechanical input impedance Z_{ll}. The ratio Z_{ll}/Z_{out} is simply the coupling coefficient β. See below for more discussion of this point.

Once we begin to think in terms of quanta we are led to ask: what happens if the induced strain in the antenna is equivalent to less than one quantum $h\omega_a$? The profound significance of the quantum mechanical limit to macroscopic measurements was realized independently by several groups, particularly by Braginsky [94] and Giffard [95]. Giffard used the much earlier result of Heffner [96] who showed that, by the uncertainty principle, a linear amplifier has a fundamental limit to its sensitivity, given approximately by $\hbar\omega_a$. Similarly, Giffard showed that a transducer used with a linear amplifier (an amplifier which preserves phase and changes the amplitude by a multiplicative factor) has a maximum sensitivity corresponding to a gravitational wave which produces an equivalent of two quanta in the bar. The term *equivalent* is used because the actual energy absorbed by the antenna depends on its instantaneous amplitude. For linear systems the signal-to-noise ratio is independent of the amplitude, and corresponds exactly to the signal produced in an ideal stationary antenna at absolute zero. Giffard's result meant that the maximum achievable sensitivity of an antenna would be limited to about the single phonon level corresponding to a strain sensitivity $\sim 10^{-21}$. This is described as the standard quantum limit.

Meanwhile, at least as early as 1974, Braginsky and Vorontsov proposed that in principle it might be possible to devise *quantum non-demolition devices* which could read out the state of a system without disturbing it. Braginsky *et al* [97], Caves *et al* [98], Unruh [99] and others went on to identify methods whereby gravitational waves of amplitude less than that required to induce one quantum can in principle be detected using *quantum non-demolition* or back-action-evading techniques. The possibility of mechanical measurements down to and below the quantum limit in sapphire bars has been investigated in detail by Tobar *et al* [100]. The most promising technique is through the use of amplitude-modulated parametric transducers. This represents a small elaboration of parametric transducers of the type we will discuss in section 4.4. The pump signal is amplitude modulated at the signal frequency to create an intrinsically phase-sensitive measuring system. However, classical noise sources need to be beaten down close to the quantum limit before such techniques can successfully pass the quantum limit on real antennas.

4.3. The signal-to-noise ratio for burst detection

Resonant-mass detectors may be used to detect all of the signal classes discussed in section 4.2. However, most effort has concentrated on the detection of bursts. In general, when a gravitational wave in the right frequency range arrives, it excites all normal modes of the bar that have a high quadrupole moment. A transducer attached to the bar will pick up the signal, which must be discerned in the presence of transducer noise and a large Brownian motion background, as discussed in the previous section.

The efficiency of the detector is determined by the fraction of the wave energy absorbed and converted into acoustic energy inside the bar. Clearly, it is important that the bar absorb as much as possible of the energy of the passing gravitational wave. This can be quantified in terms of the antenna cross section as discussed below. Denoting the incident spectral energy density of gravitational waves as $w(f)$, the energy deposited in the bar is given by

$$\varepsilon_g = \sigma w(f). \tag{4.2}$$

The term σ in the above equation is the so-called cross section of a bar, which is the ratio of the absorbed energy to the incoming energy, and thus a measure of the sensitivity of the bar. The cross section is actually a function of frequency $\sigma(f)$ because the detector will absorb energy more readily around the resonant frequency f_o of the antenna. The total energy deposited in the bar is then

$$E = \int \sigma(f) w(f)\, \mathrm{d}f. \tag{4.3}$$

Since for a high-Q system, $\sigma(f)$ is sharply peaked around the resonant frequency f_0, only a narrow portion of the gravitational wave signal around the resonant frequency of the bar can be picked up by the detector. In this case we may write

$$\sigma(f) w(f)\, \mathrm{d}f = w(f_0) \int \sigma(f)\, \mathrm{d}f. \tag{4.4}$$

The cross section, first elaborated by Weber, can be expressed in several forms [7]. The cross section of the bar, integrated over the frequency band, can be expressed as [101]

$$\int \sigma(f)\, \mathrm{d}f = \frac{8GM}{\pi c} \left(\frac{v_s}{c}\right)^2 \ \mathrm{m^2\ Hz}, \tag{4.5}$$

where v_s is the sound velocity in the bar, and M is the mass of the bar. Clearly, to obtain high sensitivity it is desirable to build detectors as massive as possible and from a material having a sound velocity as high as possible. Usually, the size of the bar is chosen such that

the fundamental mode is at about 1 kHz—about the expected frequency for the collapse of a massive stellar core to a black hole. Aluminium has been used for most resonant bars. Niobium is used for the bar detector at the University of Western Australia. Other materials could give substantial advantages, as discussed in section 4.5 below.

A gravitational wave carries an energy flux S(J m^{-2} s^{-1}) given by

$$S = \frac{c^3}{16\pi G}\langle h_+^2 + h_x^2\rangle, \tag{4.6}$$

where h_+ and $h_\times$ denote the dimensionless strain amplitudes of the two possible polarizations of the wave. Since the shape of expected gravitational wave pulses from gravitational collapse events is not accurately known, we cannot accurately determine the expected excitation of an antenna even knowing the total pulse energy. We need to know both the spectral distribution of the pulse energy, and the relationship between h and its time derivative. The details of the expected pulses depend not only on the dynamics of the gravitational collapse, but also on the mass of the collapsing object, both of which are uncertain.

If we assume only knowledge of the pulse duration τ_g (expected to be $\sim 10^{-3}$ s), and that it is predominantly a single cycle, it is sufficient to assume that $\mathrm{d}h/\mathrm{d}t \sim 2h/\tau_g$. Then equation (4.6) can be rewritten

$$S = \frac{c^3}{16\pi G}\frac{4h^2}{\tau_g^2}. \tag{4.7}$$

The total pulse energy E is then given by $S \cdot \tau_g$:

$$E = \frac{c^3}{16\pi G}\frac{4h^2}{t_g^2}. \tag{4.8}$$

If we assume that the spectral distribution of the pulse energy $F(\omega)$ is uniform over a bandwidth $\Delta\omega \sim 1/\tau_g$, it follows that

$$F(\omega) \sim E/\Delta\omega, = E\tau_g \sim c^3 h/4\pi G \text{ J m}^{-2} \text{ Hz}^{-1}. \tag{4.9}$$

Numerically $F(\omega) \sim 20 \times 10^{34}\ h^2$.

The assumption used in obtaining the result must be emphasized: the result is simply an order-of-magnitude estimate of the expected signal spectral densities. Moreover, variations in the pulse durations could make any chosen antenna frequency only suitable for a small proportion of actual events.

The energy deposited in an initially stationary antenna of mass M by a signal pulse $F(\omega)$ follows directly from equation (4.5) combined with geometrical terms:

$$U_s \sim F(\omega)\sin^2\theta\cos^2 2\phi\frac{8}{\pi}\frac{G}{c}\left(\frac{v_s}{c}\right)^2 M \tag{4.10}$$

where θ and ϕ are coordinates describing the orientation of the bar relative to the incoming wave (as given in figure 20).

For a short burst of gravitational waves the bandwidth of the pulse is roughly the inverse of the pulse duration which is roughly equal to the peak frequency. Under these circumstances the strain amplitude $\delta 1/1$ induced in the bar is roughly equal to the incoming wave amplitude h; there is no resonant excitation.

The incoming gravitational wave will only be detectable if the signal U_s is greater than the noise in the antenna U_n. From an engineering point of view it is useful to characterize the noise U_n; we generalize the transducer to a two-port device described by a 2×2 impedance matrix Z_{ij}. The transducer accepts force and velocity inputs F and v, giving current and voltage outputs I and V:

$$\begin{pmatrix} F \\ V \end{pmatrix} = \begin{pmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{pmatrix}\begin{pmatrix} v \\ l \end{pmatrix}. \tag{4.11}$$

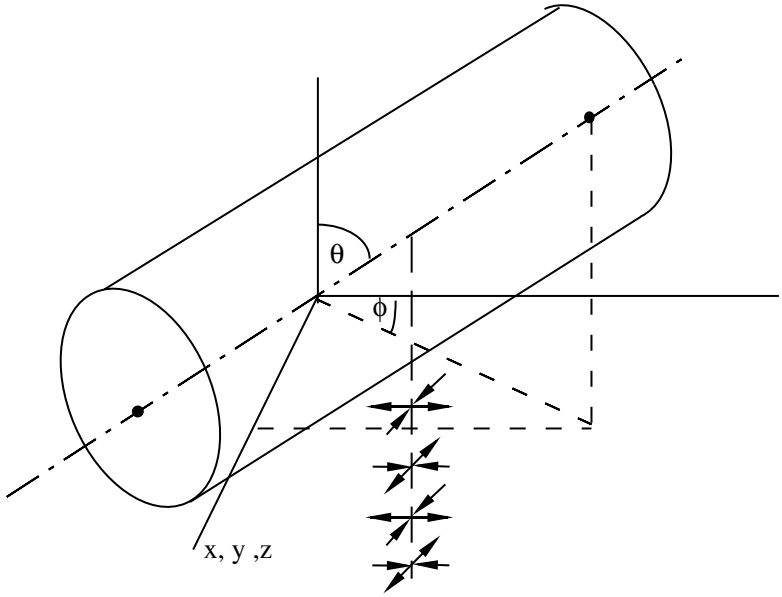


Figure 20. Coordinate system for resonant antenna.

The transducer has input impedance Z_{ll}, measured in kg s^{-1}, and output impedance Z_{22}, measured in ohms. The forward transductance Z_{21}, measured in V m^{-1} s, determines the transducer sensitivity whereas the reverse transductance Z_{12}, measured in kg A^{-1}, determines the back-acting force on the antenna due to currents in the output circuit: see [100, 102–104] for more details. Quantum mechanics tells us that it is impossible to make $Z_{12} = 0$: it is impossible to create a perfect one-way valve.

All the noise sources in the transducer and amplifier can be expressed as equivalent spectral densities of current and voltage noise at the input of the amplifier, denoted $S_i(\omega)$ and $S_e(\omega)$, respectively, as illustrated in figure 21. The current noise S_i is the source of back-action noise in the antenna, whereas S_e describes the series noise contribution. In terms of these quantities the total system noise is given by

$$U_n = 2kT_a\frac{\tau_i}{\tau_a} + \frac{|Z_{12}|^2}{2M}S_i(\omega)\tau_i + \frac{2M}{|Z_{21}|^2}\frac{S_e(\omega)}{\tau_i}. \tag{4.12}$$

The first term in equation (4.12) is the Brownian motion or thermal noise in the antenna discussed above. The second term describes the energy fluctuations arising from the current noise acting back through the reverse transductance, and giving the back-action noise, also proportional to integration time. The third term is the series electronics noise, which for given S_e is reduced as Z_{21} increases. As we saw above this term is proportional to the bandwidth τ_I^{-1}. Only the first term in equation (4.12) can ever be reduced below the quantum limit.

The problem of detecting gravitational waves with resonant-bar antennas to a large extent consists of minimizing equation (4.12). The technical means of achieving this requires some or all of the following:

(a) Reduce the antenna temperature T_a.
(b) Use a transducer with high Z_{21} and low Z_{12}.
(c) Use amplifiers with S_e and S_i as low as possible.

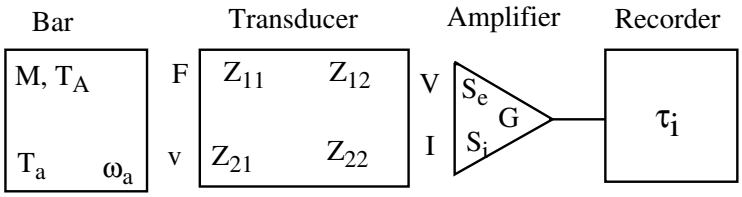


Figure 21. The various quantities used to characterize a gravitational wave antenna system.

(d) Reduce the acoustic loss in the antenna as well as the acoustic and electrical losses in the transducer to obtain high-Q or large relaxation time.
(e) Obtain a reasonable impedance match between antenna and transducer to enable τ_i to be small.

The last requirement can only be achieved by good impedance matching between the mechanical input impedance of the transducer and the mechanical output impedance of the antenna, which we shall discuss further below.

It is convenient to scale the noise in the system relative to the *standard quantum limit* of one equivalent quantum induced in the bar. To do this we rewrite the noise equation (4.12) in terms of *noise number A* (a quantity first used by Weber to characterize noise in masers):

$$A = U_n/\hbar\omega_a = A_T + A_B + A_S. \tag{4.13}$$

Here A_T, A_B and A_S are the equivalent numbers of noise quanta due to thermal noise, back-reaction noise and series noise in the measurement system. The experimentalists need to achieve a total system noise number A approaching unity. For a 1 kHz resonant bar this corresponds to a noise temperature of $\sim$0.1 μK. To achieve this it is necessary not only to have a low-noise transducer, but also to use a low-acoustic-loss antenna material, and to suspend and isolate the antenna so as not to increase the acoustic loss, nor to couple in excess noise from the environment. Because the noise number contributions are additive there is no particular advantage in reducing one of them far below the others.

It is useful to express the thermal and transducer noise contributions directly in terms of the gravitational wave strain equivalent noise. The *Brownian motion noise* h_B is given by [105]

$$h_B \sim \left(\frac{k_B T \tau_i}{M\omega_a l^2 Q}\right)^{1/2}. \tag{4.14}$$

For a bar detector, $\omega_a l = \pi v_s$, so the Brownian noise is given by

$$h_B \sim \left(\frac{k_B T \tau_i \omega_a}{\pi^2 M v_s^2 Q}\right)^{1/2}. \tag{4.15}$$

Expressing the above equation numerically, we have

$$h_B = 10^{-21}\left[\left(\frac{f_0}{1\ \text{kHz}}\right)\left(\frac{10^{10}\ \text{J}}{M v_s^2}\right)\left(\frac{10^9}{Q}\right)\left(\frac{T}{0.1\ \text{K}}\right)\left(\frac{\tau_i}{10^{-2}\ \text{s}}\right)\right]^{1/2}. \tag{4.16}$$

This means that a resonant-mass detector with the hypothetical parameters implied in the above equation will have a Brownian motion noise strain of about 10^{-21}. However, some of the parameters given above are difficult to achieve in practice. For example, the energy term is satisfied by about 10 tonnes of sapphire or 100 tonnes of bronze or niobium. In the case of the UWA niobium bar with mass of 1.5×10^3 kg, resonant frequency of 700 Hz, sound velocity of 3.4×10^3 m s^{-1}, temperature of 4 K, Q-factor of 3×10^7 and measurement integration time 1 s, the Brownian noise strain amplitude is 2.2×10^{-19}. It can be seen that the bar must

have a very high acoustic quality factor or much shorter integration time to reduce the effect of the thermal noise. With today's multimode impedance matching techniques, τ_i can indeed be reduced to 10^{-2} s.

The sensor noise lower limit (series noise and back-reaction noise $A_B + A_S$) is set by the standard quantum limit [105]

$$h_{\mathrm{SQL}} \sim \left(\frac{2\hbar\omega_a}{\pi^2 M v_s^2}\right)^{1/2} \sim 1.1 \times 10^{-21} \left(\frac{f_m}{1\ \mathrm{kHz}}\right)^{1/2} \left(\frac{10^3\ \mathrm{kg}}{M}\right)^{1/2} \left(\frac{10^4\ \mathrm{m\ s^{-1}}}{v_s^2}\right). \tag{4.17}$$

This sets a strain amplitude limit $\sim 10^{-22}$ for a 100 tonne resonant-mass detector such as the proposed spherical detectors.

4.4. Transducers

Transducers for resonant-mass gravitational wave antennas fall into two categories: *passive* transducers and parametric transducers. Passive transducers have no external power source, and their power gain is less than unity. They must always be used with a high-gain, low-noise amplifier at the frequency of the antenna. Parametric transducers, on the other hand, have an external power source (a pump oscillator at frequency ω_p) which is modulated by the antenna motion. They have intrinsic power gain associated with the transfer from the antenna frequency ω_a to the higher frequency ω_p. A laser interferometer is a parametric transducer operating at an optical pump frequency. Parametric transducers for resonant-mass readouts may be optical, microwave or radiofrequency devices.

Most parametric transducers use a high-frequency resonator combined with a low-noise high-frequency amplifier. Passive transducers use an inductive or capacitive readout, coupled to a SQUID amplifier. Figure 22 illustrates their basic structure. The passive transducer illustrated uses a superconducting inductor whose inductance (if constructed in a planar fashion) is proportional to the gap spacing between the coil and the superconducting ground plane on the antenna. Relative motion modulates persistent current trapped in the inductor (since the magnetic flux LI must be conserved). The changing current is detected by a SQUID magnetometer. The parametric transducer illustrated uses a capacitor in a resonant circuit. The capacitance is modulated by the gap spacing between the capacitor and the antenna. The change in capacitance due to motion modulates the resonant frequency of the circuit, creating modulation sidebands on the output signal. Both types of transducer may use capacitive or inductive sensing.

Fundamentally, there is little difference between passive and active transducers. Active transducers use a transduction process that is combined with power amplification but additional amplification of the high-frequency signal is usually necessary. Passive transducers have a complete separation between the transduction process and the amplification process. However, the amplifier itself (such as a SQUID) makes intrinsic use of a parametric up-conversion process. Thus the difference between passive and active transducers is simply in the choice of whether the parametric up-conversion occurs during or after transduction. In the case of an optical pump frequency, amplification is unnecessary: the entire power gain is realized through the up-conversion of the signal frequency to the optical frequency.

One important difference between passive and parametric transducers is in the transducer impedance mismatch ratio or coupling factor β. For the parametric transducer

$$\beta_{\mathrm{para}} \frac{1}{2} \frac{C V_p^2 Q_e}{m\omega_a^2 x^2}. \tag{4.18}$$

In the limit $Q_e > \omega_p/\omega_a$, the electrical Q-factor of the transducer resonator Q_e is replaced by the frequency ratio ω_p/ω_a.

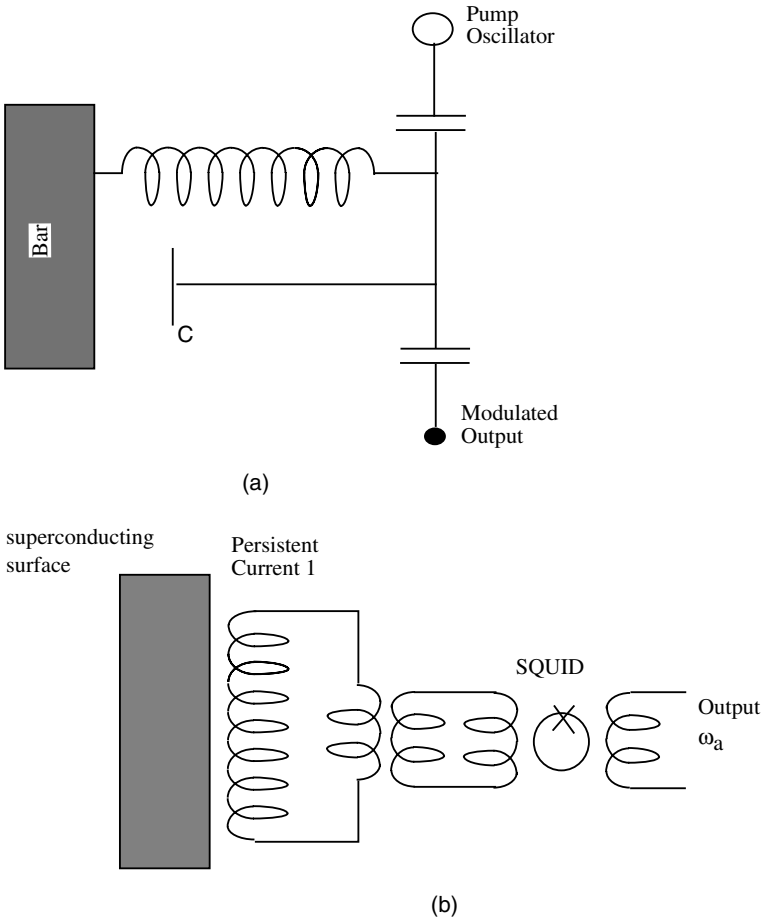


Figure 22. Inductance or capacitance is modulated by a gap spacing. (*a*) Active or parametric transducer use a low-loss resonant circuit pumped by an external oscillator. (*b*) Passive transducers use an inductive or capacitive readout, coupled to a SQUID amplifier. All the circuits are made from superconducting and very-low-loss components.

For the passive inductive transducer

$$\beta_{\text{pass}} = \frac{1}{2}\frac{LI^2}{m\omega_a^2 x^2}. \tag{4.19}$$

Note that the passive transducer coupling factor is not enhanced by a Q-factor term. For a capacitive passive transducer the inductive stored energy $\frac{1}{2}LI^2$ is replaced by the capacitive stored energy $\frac{1}{2}CV^2$. The parametric transducer effectively samples the incoming signal Q_e times per cycle up to a maximum value of ω_p/ω_a, and therefore increases its coupling by the same factor.

Parametric coupling is *reactive* as a result of the position-dependent mechanical forces which act across the electrical resonator. The mechanical forces vary strongly over the transducer *position bandwidth*, defined as the halfwidth of the electrical resonator, measured in terms of displacement. Typically the position bandwidth is ∼pm (10^{-12} m). The stored electrical energy exerts forces across the capacitance which vary strongly over the position

bandwidth. Hence the effective spring constant can be very large, and due to its reactive nature can create problems in maintaining stability.

The passive transducer does not have the same coupling advantage. However, the advantage is to some extent illusory because by moving the coupling structure to a high frequency one reduces its size, so that the absolute value of the L or C is significantly reduced.

The practical problems of the two types of transducer are quite different. Passive transducers are limited by a poorly understood problem of AC losses in their superconducting circuits, and the performance of available SQUID amplifiers. Parametric transducers, on the other hand, are limited by phase noise in the pump oscillator, tuning difficulties, noise degradation in amplifiers at high microwave power, and sometimes by the effects of low-frequency seismic noise. All transducers, both active and passive, are limited in noise performance by the noise of the amplifier with which they are used. In terms of noise number equation (4.13), at 1 kHz a noise number of 1 corresponds to a noise temperature of about 100 nK, whereas at 10GHz the same performance corresponds to $T_n \sim 1$K. In principle, both types of transducer can reach close to the quantum limit. In practice, none have reached this level to date, although SQUID amplifiers have been developed close to 1 μK, and amplifiers for microwave parametric transducers have long been available with $T_n \sim 10$K. The microwave parametric transducer on NIOBE at UWA has achieved about 1 mK noise temperature, while the SQUID transducer on NAUTILUS has achieved similar noise performance. It is not obvious which type of transducer will ultimately be the most successful.

Johnson and Bocko [106–108] and Tobar *et al* [100] have presented designs for quantum-limited microwave transducers, while Richard and the Legnaro group have presented design for optical transducers [109, 110]. See [94, 98, 111, 112] for further discussion of these issues.

A sensitive transducer and a low-loss resonant mass are not sufficient to create a sensitive gravitational wave antenna. There is a major problem at the interface: mechanical impedance matching. The impedance mismatch ratio (introduced as the coupling factor β) arises because the mechanical output impedance of the bar is very high, characterized by the elastic stiffness of the bar itself, whereas the mechanical impedance of the electric or magnetic field which couples this motion into the transducer is not large. The solution is to create an acoustic transformer at the end of the bar. Such a transformer is analogous to acoustic horns used in loudspeakers, or to the mechanisms in the human ear that couple the motion of the air into the fluid of our cochlea.

All successful impedance matching schemes have consisted of low-mass secondary acoustic resonators tuned to the antenna frequency. This creates a two-mode resonator with a pair of normal modes. The acoustic energy beats between the high-mass resonator and the low-mass resonator, while the transducer is coupled to the low mass. The scheme can be generalized to multimode transformers, consisting of nested sets of resonators reducing in mass by a geometric progression.

Three secondary resonator configurations have been used successfully on antennas: diaphragms (first developed by Paik at Stanford [113]), mushrooms (developed by the Rome group [114]) and bending flaps (developed at UWA [115]). Pang and Richard [109], Hamilton *et al* [116] and Tobar [63] have proposed and tested 3–5 mode systems but these have yet to be implemented. Four of the systems are illustrated in figure 23. The antenna NIOBE uses a 400 g bending flap (figure 23(a)). The bending flap is a convenient form of secondary resonator, which has open geometry suitable for an attachment of a microwave re-entry cavity parametric transducer readout. The microwave readout system consists of a carrier suppression interferometer, and a microwave amplifier followed by a demodulation stage. This achieves reasonably high coupling to the microwave transducer, as demonstrated by the fact that the elastic forces provided by the transducer are sufficient to detune the mechanical oscillator by

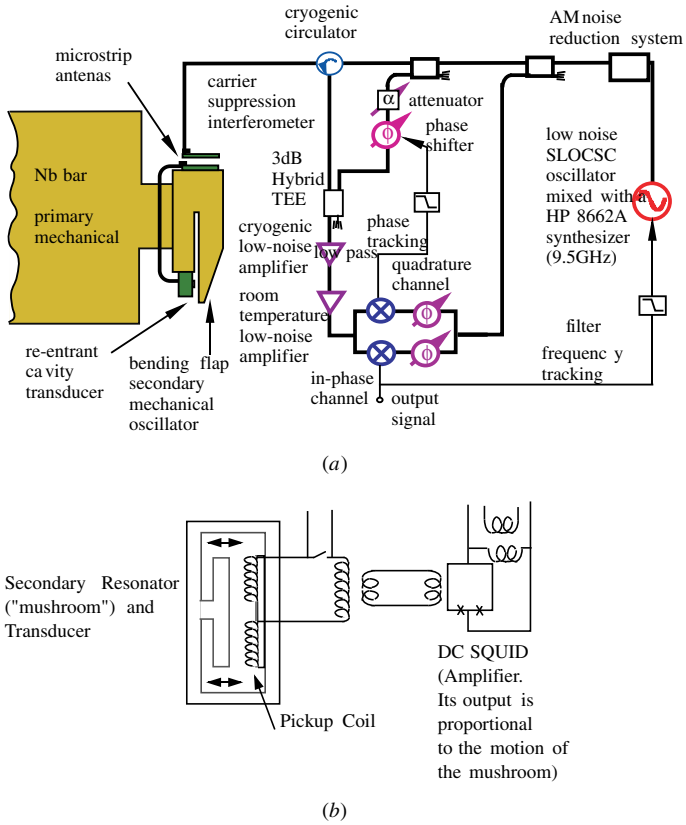


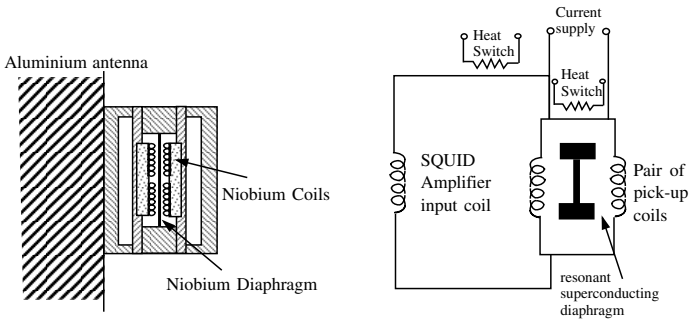
Figure 23. Bar antenna transducer readout systems. (*a*) Bending flap [115]; (*b*) mushroom [114]; (*c*) diaphragm [113]; (*d*) multimode transducer [116].

several Hz. The following three configurations, (*b*) the mushroom, (*c*) the diaphragm and (*d*) a multimode transducer, use similar SQUID amplifier readout circuits. The moving mass modulates the inductance of a flat coil, which by flux conservation leads to a modulated current through a SQUID amplifier. The Rome group has used capacitive readout for a mushroom, in which case currents are induced by charge conservation in capacitor. The multimode transducer uses a massive 'diaphragm' coupled to a small tertiary mass plate supported by small niobium cantilever springs.

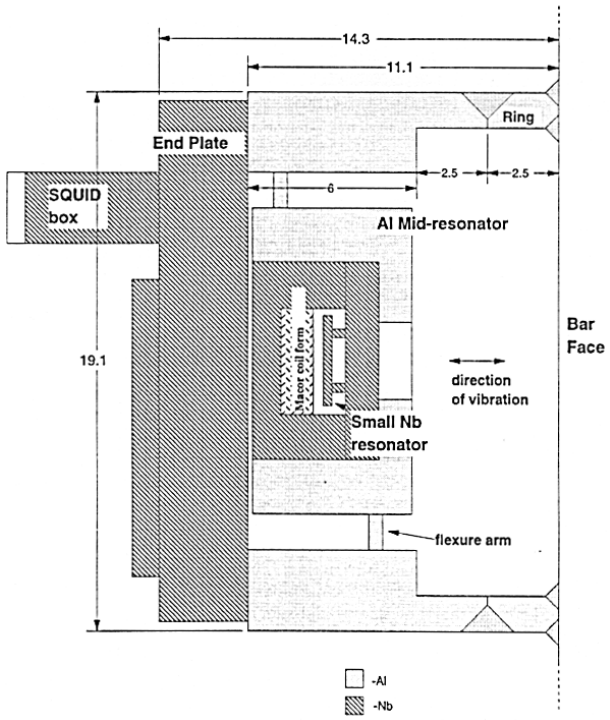
4.5. Antenna materials

An ideal resonant bar would consist of a piece of nuclear matter, with high density and a velocity of sound comparable to the velocity of light! Since this is not available except in neutron stars, we must find a form of molecular matter which, to maximize coupling to gravitational waves, combines high velocity of sound v_s, and high density ρ. To reduce the thermal noise we require a low acoustic loss Q^{-1}.

For an antenna limited by thermal noise the best antenna material (at a practical frequency) will have the largest value of $Q\rho v_s^3$. This quantity is proportional to the ratio of energy absorbed ($\sim\rho v_s^3$) and the thermal noise in the antenna ($\sim Q^{-1}$). Of the three controlling parameters, only the Q-factor can be modified significantly in a particular material, depending on its preparation



(*c*)



(*d*)

Figure 23. (Continued)

and suspension.

Table 2 lists the values of ρ, v_s and ρv_s^3 for various materials, along with the maximum achieved Q-value to date, and the signal-to-noise ratio figure of merit, $Q\rho v_s^3$. The table shows that nearly one order of magnitude improvement is obtained (at a given frequency) in ρv_s^3 by changing from aluminium or niobium to sapphire, and when the Q-factor is included the very low losses in sapphire make it about 500 times superior to Nb or Al (at a given operating temperature). Silicon is more than 100 times better than Nb and Al. Unfortunately, at present silicon and sapphire are not available in sufficiently large masses for these apparent advantages to be useful. Note that a lower Q-factor can always be compensated for by sufficient

Table 2. Comparison between antenna materials.

Material	ρ(g cm^{-3})	v_s(km s^{-1})	Q	ρv_s^3 (10^{13} kg s^{-3})	$Q\rho v_s^3$ (10^{20} kg s^{-3})
Aluminium 6061	2.7	5.1	5×10^6	36	18
Aluminium 5056	2.7	5.1	7×10^7	36	250
Niobium	8.57	3.4	2.3×10^8	34	800
Silicon	2.33	8.5	2×10^9	140	2.8×10^4
Sapphire	3.98	9.4	3×10^9	330	10^5
Lead	11.36	1.1		1.5	
Tungsten	18.8	4.3		150	
Copper(94)/Al(6)	8.0	4.6	2×10^7	77.8	155

cooling, so that fundamentally only the ρv_s^3 term need be considered. Copper-based alloys have been selected as preferred materials for very-high-mass spherical antennas, chiefly because superconducting materials (for which the thermal conductivity becomes very low) are very difficult to cool to ultralow temperatures.

For comparison, table 2 also shows lead and tungsten. Lead is very poor, because of its low sound velocity, whereas tungsten is comparable to silicon. If massive high-Q tungsten masses could be obtained, they would have the significant advantage that the cryogenic system necessary to house the antenna would be smaller (and cheaper and simpler) than that needed for lower-density materials.

4.6. Antenna suspension and isolation systems

Typical *seismic noise* has the spectrum of $x_s = \alpha f^{-2}$ m Hz$^{-\frac{1}{2}}$ in the frequency range 1 Hz to a few kilohertz, where f is the frequency and α is a constant. Measurements by gravitational wave research groups at various sites have shown that the constant varies between 10^{-6} and 10^{-9} (e.g. [117–119]). A vibration isolation system is needed to isolate seismic noise to well below the signal level at the antenna resonant frequency. A variety of metallic suspensions have been developed for resonant gravitational wave detectors (e.g. [120–122]). All are designed to have the normal-mode resonant frequencies of the isolator well below the antenna frequency, and the internal modes of the isolator elements above the frequency range of interest. In general, the normal modes define a set of low-frequency resonances. Internal modes of the mass and spring elements are generally at high audio frequencies. Such isolators therefore have good isolation above a low-frequency corner, and also below the high-frequency internal resonances. For resonant-bar antennas, the isolation band required is from a few hundred hertz to a few kilohertz. In principle, it is fairly easy to realize a mechanical isolator which will attenuate the seismic motion at 1 kHz ($\sim 10^{-12}$ m Hz) by 10^{10} (e.g. [121, 122]). However, practical problems such as nonlinear up-conversion can degrade the performance [123].

Figures 24 and 25 illustrate two of the resonant-bar antennas constructed to date. For antennas operated at 4 K the bar is supported by a low-loss multistage suspension in an experimental chamber with which the antenna has no direct contact. A room temperature vibration isolation stage suspends the cryogenic suspension stage. In the case of NIOBE, vibration can only act on the antenna by traversing the entire 18 stages of vibration isolation, or through transmission through the residual gas in the experimental vacuum. There are no wires connecting to the antenna (which can transmit vibration) because the transducer is

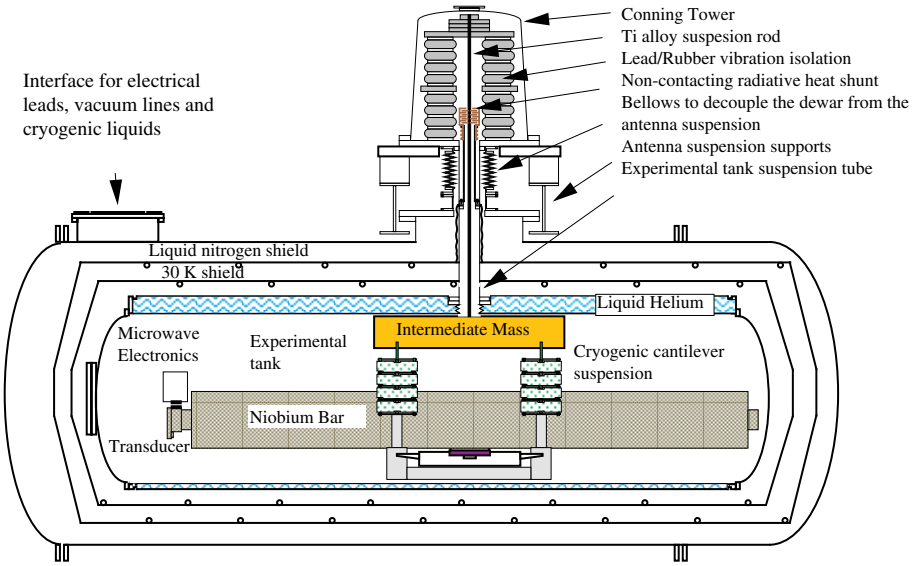


Figure 24. Cross section of the NIOBE antenna. The cryostat is 5 m long and 3 m high.

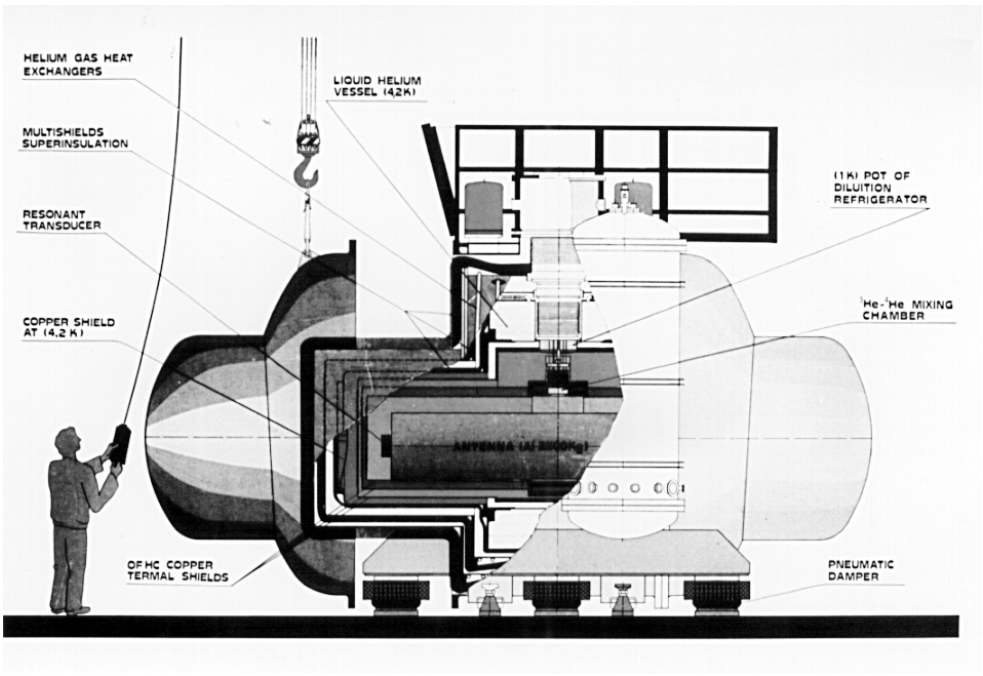


Figure 25. View of the ultracryogenic antenna AURIGA showing cryogenic shields and the Al bar.

interrogated entirely by radiative contact. In the case of antennas using passive transducers superconducting cables can act as transmission lines to conduct vibration, so that great care must be taken to isolate cables using *Taber isolators* [124], consisting of additional mass spring stages (using thin wires as springs) to which the cables are bonded.

The *cryogenic suspension and isolation stage* of an antenna is one of the most critical components. It must isolate against the noise which bypasses room temperature isolation, as well as the thermal noise of the room temperature stage. In addition, it must suspend the antenna using low-loss elements so as to maintain a high antenna quality factor. It is essential that the cryogenic stage, at least, does not have any resonant modes near the antenna frequency. Several different systems have been used with reasonable success. Traditionally, cable suspensions have been used: a cable slung around the belly of the resonant bar. This method has the risk of introducing vibration by the violin string mechanism: slip–stick frictional transitions associated with the motion of the belly cable at the point of tangential contact with the bar. This process arises because of the seismically driven low-frequency pendulum motion of the bar. Such boundary condition effects can be avoided if the antenna is somehow bonded or clamped to the bar at the tangent point. The now defunct Stanford group and the LSU group used welded rods instead of wires to avoid this problem. The third method consists of a cantilever suspension. High Q-factor curved cantilever springs such as the Catherine wheel used on NIOBE support the bar from below. This can have well defined contact points to the antenna to minimize nonlinear processes, and has given the highest Q-factor ever observed in a metal [125]. A fourth method, first suggested by Coccia [126], is the use of a nodal point suspension. In the case of a bar or sphere, the nodal point for the fundamental mode is located at the centre of mass. This means boring a hole to the centre and attaching a rod or cable. While this has many advantages in reducing sensitivity to external noise, it has not been used in a full-scale antenna.

Antennas cooled to ultracryogenic temperatures (below 100 mK) have a particular problem to contend with. Helium exchange gas can no longer be used at such temperatures (the vapour pressure is too low) so the antenna must be cooled by conduction. This means that there must be direct cryogenic conduction paths to the antenna from the dilution refrigerator. Pure and nonsuperconducting metals must be used (such as OFHC copper). Yet the thermal conduction must exist without significant vibration conduction, especially since the dilution refrigerator is likely to be a substantial source of vibration.

Ultracryogenic detectors used to date have exhibited excess sensitivity to local vibration, due presumably to the inadequate performance of the thermal conduction/vibration isolation system. However, recently both NAUTILUS and AURIGA have yielded improved performance, down to a noise temperature of 1 mK.

4.7. Present status of resonant-mass detectors

At the time of writing five resonant-mass gravitational wave detectors are in operation. These consist of three liquid helium temperature detectors, ALLEGRO, EXPLORER and NIOBE at Louisiana State University, CERN (operated by the University of Rome Group) and the University of Western Australia, and two ultracryogenic detectors, NAUTILUS and AURIGA at INFN Frascati and INFN Legnaro, Italy. The latter detectors have been successfully cooled to below 100 mK. Table 3 below summarizes the basic parameters of these antennas. The noise performances quoted are typical/best rms noise levels for the detection of broadband bursts.

As well as undertaking long periods of operation, the antennas are all in the process of upgrade, either through installation of improved transducers, or through installation of improved vibration isolation. Due to the large size of the cryogenic systems, resonant antennas have a rather long cycle time (several months) of cooling and warm-up. Combined with the probability of malfunction in experimental apparatus (a particular problem during the 1980s) the rate of progress has been slower than predicted.

Table 3. The resonant mass detectors which belong to the international Gravitational Events Collaboration.

Antenna	Location	Material	Temperature	Noise Temperature (mK)	Frequency (Hz)	Strain sensitivity
ALLEGRO	Baton Rouge	Al	4 K	6	900	7×10^{-19}
EXPLORER	CERN	Al	2 K	6	900	7×10^{-19}
NIOBE	Perth	Nb	5 K	1	700	5×10^{-19}
NAUTILUS	Frascati	Al	100 mK	4	900	6×10^{-19}
AURIGA	Legnaro	Al	100 mK	1	900	3×10^{-19}

4.8. Performance of resonant bars

The long-term operation of cryogenic resonant bars has been invaluable in characterizing their instantaneous performance, and evaluating various possible filtering techniques for extracting events from a long data record in the presence of excess noise. With regard to instantaneous performance it has been important to verify that the noise performance is consistent with the noise parameters of the bar, transducer and readout system. In nearly all cases extremely good agreement has been obtained, such as illustrated in figure 26. In the case of parametric transducers their performance can be characterized not only by noise spectra, but also by their variable interaction with the antenna, as discussed in section 4.5.

From the experimentally observed noise spectral density, such as figure 26, one can determine the sensitivity of the antenna to various signals such as stochastic background, CW signals and bursts. Figure 27 shows the calibrated burst sensitivity of NIOBE during a 24 h period. The data are presented as mean noise temperature (bottom curve) and the largest noise temperature observed in 100 s. This allows the antenna performance to be quickly assessed, including the presence of excess noise. Figure 28 presents the same data in the form of Brownian motion noise histograms. From the single antenna data there are clearly few events above 10 mK, corresponding to $h \sim 10^{-15}$ (all of these could normally be eliminated by coincidence analysis.

The resonant-bar detector network has recently been able to set new upper limits to the strength and event rate of gravitational wave burst signals from coincidence analysis of three- and four-antenna data. Figure 29 shows this result in comparison with previous upper limits [127].

We saw in section 4.2 that the noise energy for burst detection is reduced by the factor τ_i/τ_a, where τ_i is the signal integration time and τ_a is the antenna ring-down time. The reduction of the noise with τ_a is a manifestation of the fluctuation dissipation theorem. In the case of CW signal detection the noise energy reduces proportional to $(\tau_i \tau_a)^{-1}$. This means again that the best detector is one with very high Q-factor, and that very long integration times improve the amplitude sensitivity as $\tau_i^{1/2}$. Figure 30 presents an FFT analysis of one month's data from the ALLEGRO detector in the search for pulsar signals form the globular cluster 47 Tucanae. In this case the deep integration, over a narrow frequency band, has set limits for CW pulsar signals $\sim 10^{-23}$. The analysis has been repeated for various directions in the two low-noise bands of ALLEGRO.

For the detection of stochastic backgrounds one simply has to multiply the output of two nearby detectors and integrate the result. However, as discussed in section 3, the detectors need to be sufficiently close that the incoming waves are correlated. Their space should be within about $\lambda/3$ where λ is the gravitational wave wavelength c/f. Cross correlation between the detectors NAUTILUS and EXPLORER (as mentioned in section 3) have yielded a limit to

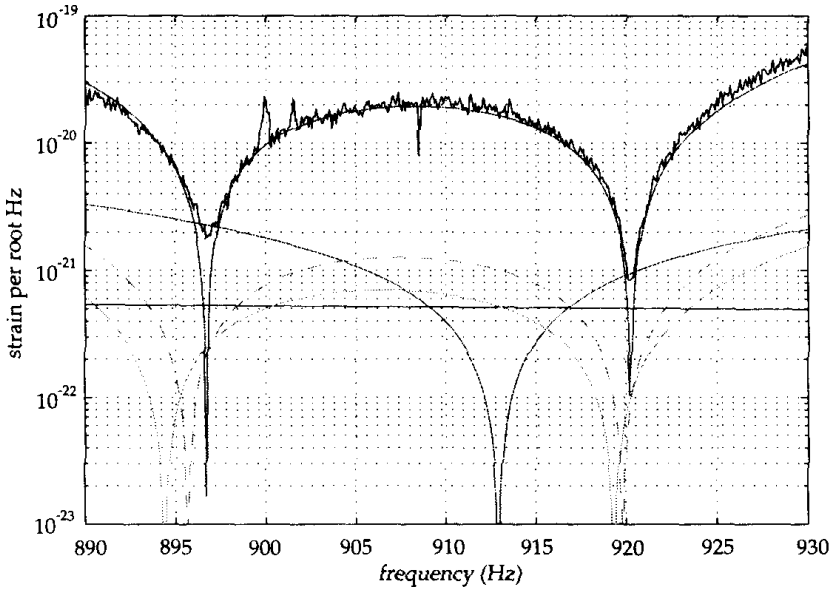


Figure 26. The measured strain noise of ALLEGRO, shown as the irregular trace. The various noise contributions estimated from the noise model are shown as smooth curves. The noise is dominated by the SQUID's wide-band and the transducer's narrow-band noise [71].

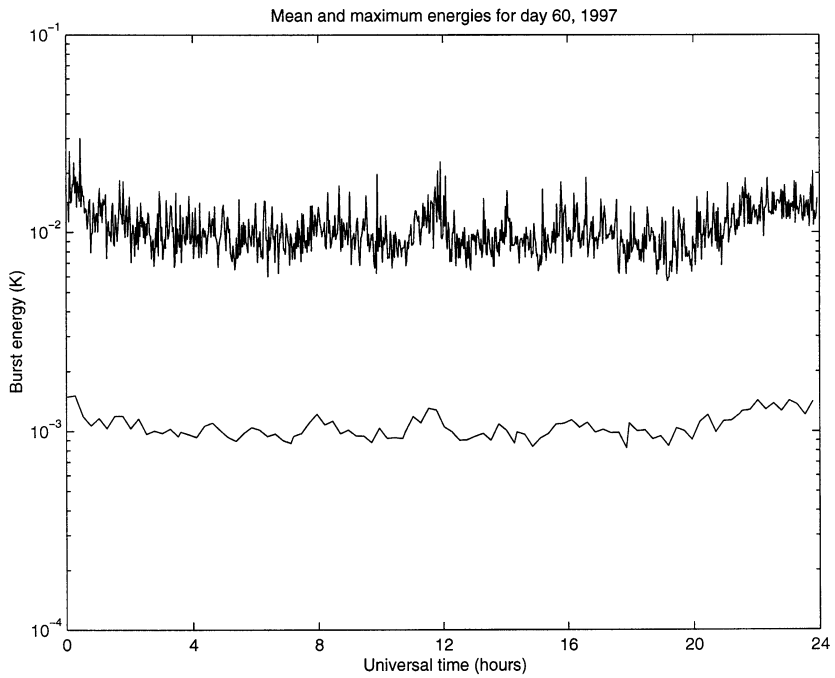


Figure 27. The calibrated burst sensitivity of NIOBE during a 24 h period. The data are presented as mean noise temperature (bottom curve) and the largest noise temperature observed in 100 s [78].

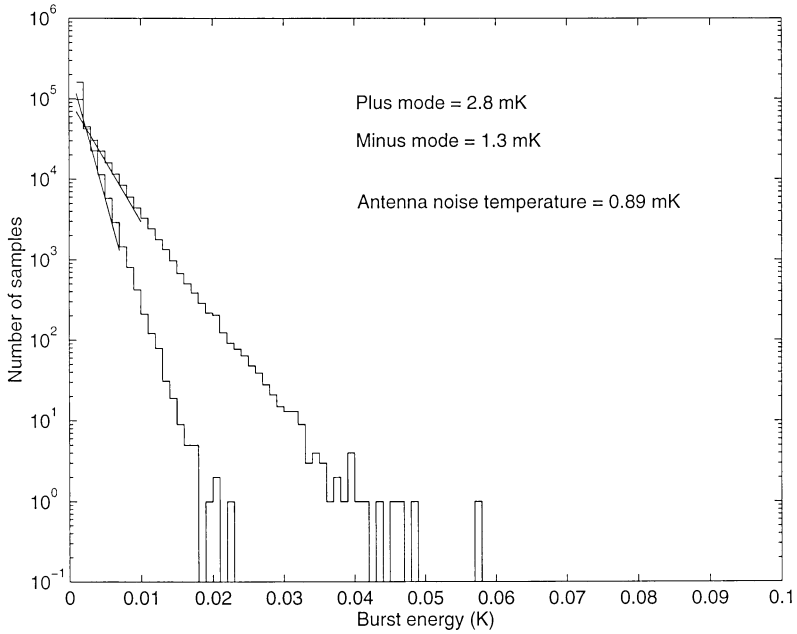


Figure 28. Histogram of the noise of NIOBE after optimal filtering for burst signals, showing the expected Boltzmann distribution. When the weighted mean of both normal modes is evaluated the overall system noise temperature is 0.89 mK.

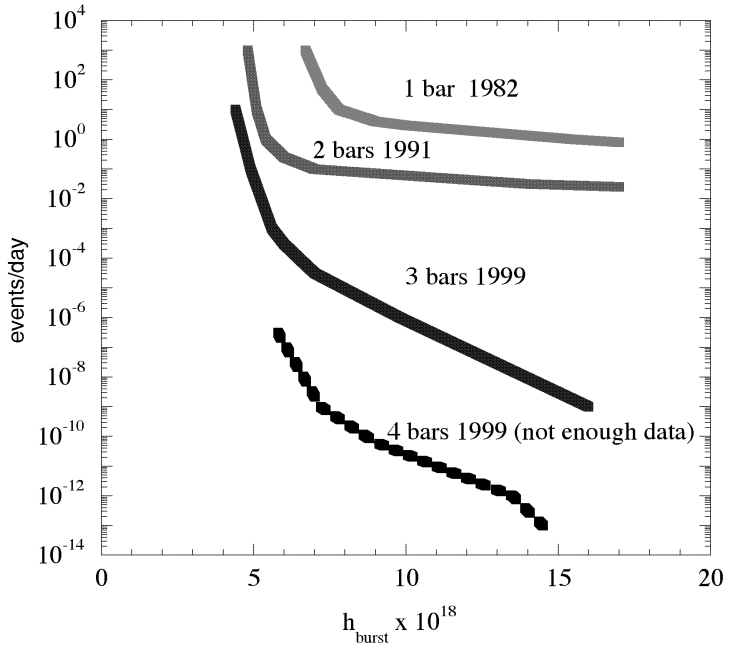


Figure 29. The rate of burst events versus strain amplitude set by one-two-three- and four-antenna experiments. The 4-antenna result is based on a short period of data and hence has weaker significance.

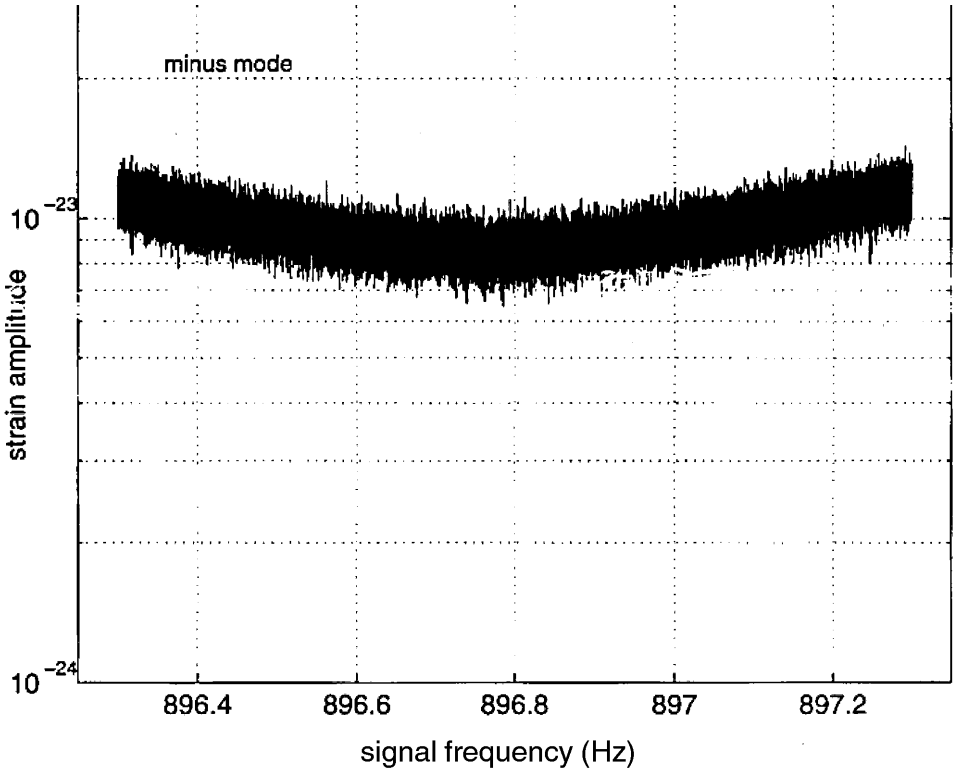


Figure 30. Search for narrow-band continuous waves from possible pulsars in 47 Tucanae [71]. Over about 1 Hz range no CW signal is visible above a strain amplitude $\sim 10^{-23}$.

stochastic gravitational waves of $\sim 10^{-22}$ at 900 Hz. The results achieved were limited because the detectors were too far apart. Vitale *et al* [128] have shown that NAUTILUS and AURIGA can be expected to achieve improvement of more than two orders of magnitude over this figure. However, they should be located within about 100 km of each other to achieve best sensitivity. This would bring the sensitivity to within the range of possible stochastic signals from the era of early star formation.

4.9. Multiple antenna correlation

In the future we can expect detailed correlation experiments to take place between a worldwide array consisting of both resonant-mass and laser interferometer gravitational wave detectors. The analysis would take into account the relative amplitude of the signals observed by detectors with different orientations relative to an incoming signal, and the phase delay due to the propagation time of signals through the Earth. Such a combined analysis would allow source direction and polarization to be accurately determined. Today we are still far from achieving this goal. Here we will discuss the analysis performed to date on data from a far less optimal array.

Besides seismic noise excitation, all cryogenic antennas have shown evidence of excess noise of indeterminate origin [81], as mentioned in section 3. Low noise performance may be achieved for considerable periods of time; but interspersed are periods of excess noise which

is not identifiably correlated with known noise sources. This can degrade performance by many orders of magnitude. While there are opportunities (as discussed above) for nonlinear up-conversion driven by low-frequency pendular modes, and also for thermal-stress-driven excitation as regions of the cryostat vary in temperature, no firm correlation is generally apparent. Thus the seismic, acoustic, electromagnetic pulse and cosmic ray shower detectors that are generally used to discriminate possible noise signals are not sufficient to eliminate excess noise, and much careful work still needs to be done. It has been shown that at a noise temperature of $\sim$1 mK about 2 cosmic ray events will occur per day, and this increases to 103 events per day at the 1 μK noise sensitivity [129]. The Rome group has convincingly demonstrated the presence of cosmic ray excitation events in their data [130].

Multiple antenna coincidence correlation can minimize the effects of excess noise, as demonstrated by many coincidence experiments (e.g. [131]), which sets new lower limits on the gravitational wave flux. When four or more antennas are operated in coincidence the rate of accidental coincidences becomes extremely small, as we discuss further below.

Candidate gravitational wave events consist of either unknown environmental perturbations, occasional rare Gaussian high-energy excursions, and possibly real gravitational wave signals. These may be idealized as an independent set of background events, occurring at a constant rate R per unit time. There is evidence that the background events are not entirely independent but to some extent are clustered. This can occur if a local vibration source acts over a period of time. In spite of this, the data are generally well described by a Poisson distribution.

The probability of a background event in one antenna during the antenna resolving time τ_r is given by

$$P_1 = R\tau_r. \tag{4.20}$$

Now if there are N independent antennas, the probability of accidental coincident excitation of all N antennas in a coincidence window τ_c that must always satisfy $\tau_c \geqslant \tau_r$, is given by the product

$$P_N = \tau_c^N \prod_{i=1,N} R_i. \tag{4.21}$$

If we simplify by assuming that all antennas experience the same background at the rate R, equation (4.21) becomes

$$P_N = R^N \tau_c^N. \tag{4.22}$$

From this it follows that the mean number of accidental coincidences during a coincidence experiment of duration $t_{\rm tot}$ is given by

$$N_{ac} = R^N \tau_c^{N-1} \cdot t_{\rm tot}. \tag{4.23}$$

Table 4 summarizes the mean number of accidental coincidences for experiments with one or more detectors, assuming various coincidence windows. The coincidence window depends on several factors. It cannot be smaller than the poorest clock precision of the detectors in the experiment. Second, it depends on the timing resolution of the optimal filters in the antenna readout. The latter depends on the signal size, but for the typical thresholds used to extract candidate events it is about 0.1 of the optimum integration time. At the improved antenna sensitivity expected in the near future, cosmic ray events could produce one event per 100 s (10^3 events per day). To detect rare events such as gravitational waves from supernovae in our Galaxy, we should be looking for $N_{ac} < 1$ in any coincidence experiment. The probability that a single coincidence is accidental is given roughly by the value of N_{ac} (for N_{ac} less than 1).

Table 4 emphasizes that multiple antenna operation is essential to reduce the background coincidence rate. Two-way coincidence experiments to date have been meaningful because

Table 4. Mean number of accidental coincidences for $R = 10^{-2}$ s^{-1} in a four-month (10^7 s) coincidence experiment.

Number of antennas	Coincidence window (s)			
	1	0.1	0.01	0.001
1	10^5	10^5	10^5	10^5
2	10^3	10^2	10	1
3	10	0.1	10^{-3}	10^{-5}
4	0.1	10^{-4}	10^{-7}	10^{-10}
5	10^{-3}	10^{-7}	10^{-11}	10^{-15}

$R < 3 \times 10^{-4}$ has usually been achieved. At the increased rate of candidate events expected due to cosmic rays as sensitivity improves towards $h \sim 10^{-20}$, it will be difficult to obtain significant data unless at least three detectors are used, or else if separate vetos can eliminate the cosmic rays themselves. The very low probability of accidental coincidences achievable with four- and five-antenna arrays apparent in table 4 demonstrates that with high time resolution it is possible to eliminate a substantially larger background than the assumed value of R. For example, even if R was increased to 1 s^{-1}, the probability of a five-way accidental coincidence using the smallest assumed window is 10^{-5} in a four-month experiment. Note that if the antennas have different resolving times, the τ_c used in equation (4.23) must be greater than the longest resolving time.

The issue of multiple antenna coincidences is not as simple as indicated above due to the fact that the antennas have varying orientations on the surface of the globe. The angular dependence of the signal S observed in a single resonant-bar antenna is given by

$$S(\theta, \phi, \varepsilon) = (0.5(1 - \varepsilon) + \varepsilon \cos^2 2\phi) \sin^4 \theta, \tag{4.24}$$

where θ is the angle of the incoming plane wave relative to the cylinder axis of the antenna, and ϕ is the polarization angle of the wave measured relative to the plane of the antenna and the source. The polarization fraction ε measures the fraction of linear polarization of the wave. For $\varepsilon = 0$ the wave is circularly polarized, whereas for $\varepsilon = 1$ the wave is 100% linearly polarized, with polarization angle ϕ.

To assess the probability of multiple antenna coincidences we must investigate the antenna pattern of a set of antennas on the globe. Since all antennas are horizontal, their orientation with respect to the sky is largely determined by their locations on the Earth, and considered as a whole this leads to a complex antenna pattern when the responses S as given by equation (4.24) are combined into a synthetic multiple antenna pattern. Antenna patterns for one and two resonant-bar antennas have been analysed by Frasca [132] and Nitti [133]; Blair and Frasca *et al* have analysed multiple antenna arrays [134], while Schutz and Tinto [135] have analysed antenna patterns for pairs of laser interferometers.

An analysis of antenna patterns for the geographical locations of four antennas [134] shows that four antennas are sufficient to obtain near-100% sky coverage for two-way coincidences. That is, if we are content with only two antennas being suitably aligned for a random source, we can observe practically 100% of the sky. On the other hand, if we are to demand four-antenna coincidences then we require the antenna orientations to be adjusted such that they optimally search the same part of the sky, and the sky coverage is reduced to about 50%. With eight operating antennas one can achieve near-100% sky coverage and a minimum of four antennas in coincidence for any one event. Note, however, that since spherical antennas have 100% sky coverage, such an eight-antenna array could be replaced by three or four spherical antennas, thus avoiding the sky coverage problem.

We now return to the practical problems of the existing array of bars. We have seen that the presence of local sources of background noise lead to the minimum detection requirement that a two-antenna zero time delay coincidence be observed. With unknown source direction we can apply no constraint on the signal based on the relative size of the signal observed in each antenna. However, it is possible to utilize the fact that the rotation of the Earth causes the antenna sensitivity to be strongly modulated for specific source directions. For example, it may be reasonable to consider that the nearest sources (such as coalescing black holes) will be concentrated near the galactic centre. Under this assumption one can reject data when the antenna sensitivity in this direction (defined by equation (4.24)) is below some predetermined value. Due to the $\sin^4\theta$ factor in equation (4.20) this provides a rather steep angular cut-off and in practice one can eliminate $\sim$50% of data by applying such a *direction filter*. This automatically reduces the number of accidental coincidences (since these occur randomly in sidereal time) and increases the statistical significance of any zero time delay excess). Such a procedure is statistically dangerous, however. The reason for this is that there will always be some direction in which an apparent excess is statistically significant. The direction and the cut-off threshold must always be preset before the analysis.

A second means of improving the statistical significance of the data rests on the fact that the energy distribution for true signals (such as gravitational waves or calibration pulses) is different from the distribution of excess noise events. This is easy to verify experimentally. Astone *et al* [136] have shown that the energy resolution achievable in a typical resonant bar with a threshold for candidate events about 10 times the mean noise energy is uncertain within a range typically 0.3–3 (for signal energy $= 1$). This uncertainty has been verified experimentally, indicating that the use of relative amplitudes to determine source direction can only be used effectively at large signal-to-noise ratio [137]. While this energy spread is large, the energy spread of accidental coincidences is always much larger, typically spanning up to two orders of magnitude. This follows immediately from the high effective temperature of the excess noise distribution (see figure 17). As a result one can apply an energy ratio filter to coincidence data, thereby eliminating from consideration any coincidences for which the energy ratio exceeds a predetermined value. (This depends on the relative orientation of the detectors, and on their excess noise distribution.) Appropriate criteria can be set for such filters based on the measured distribution of non-true-time coincidences. In practice this scheme allows about 50% of the total coincidences to be excluded, again allowing substantial improvement in the statistical significance.

In an unpublished NIOBE–EXPLORER coincidence experiment a total of about 25 zero time delay coincidences was reduced to a total of seven coincidences by the application of both an energy ratio filter and a galactic centre direction filter. This figure was substantially above the background, but timing uncertainties and the fact that these filters were developed during the data analysis period ruled out attribution of statistical significance to the result. If it had been a blind experiment and there had not been timing uncertainties the probability of the zero time delay peak occurring by chance would have been improved from $\sim$0.01 to $\sim 3 \times 10^{-4}$. Future experiments can utilize one or both of the above filters to improve statistics. It is a technique that is likely to be applicable in the search for burst sources in laser interferometer detectors.

4.10. Future prospects

4.10.1. Spherical gravitational wave detectors. Spherical gravitational wave detectors have been analysed extensively by Coccia and co-workers [138–143], Johnson and Merkowitz [34, 144], Zhou and Michelson [145] and Stevenson *et al* [146–149]. A spherical detector consists

of a large approximately spherical mass instrumented with five or more transducers to read out the orthogonal quadrupole modes and the gravitational wave insensitive monopole mode. Such a detector has a high cross section for gravitational waves and omnidirectional sensitivity. A spherical detector can be suspended close to its centre of mass to achieve quite good decoupling of the suspension from the antenna normal modes. This cannot be perfect however, because of another problem: the matching of transducers to the large normal mode masses. All operating resonant-bar transducers use a secondary resonator to match the antenna normal mode mass (about one tonne) to the transducer (see section 4.4). Typically, the secondary resonator (such as the UWA bending flap, figure 23) has a mass of several hundred grams. This is roughly the maximum mass that can be coupled to existing transducers with electromechanical coupling coefficient in the range $1–10^{-3}$.

In the case of a 100 tonne sphere aiming for high sensitivity to bursts in the kHz range, the impedance matching problem is more complex. The maximum effective bandwidth of a two-mode system is set by the beat frequency between the two normal modes. This is given by $\Delta f = f_0(M_1/M_2)^{1/2}$. For a 1000:1 mass ratio this represents a bandwidth of 3%. Since bandwidth translates directly into burst sensitivity, it is necessary to reduce this ratio, which can best be done with a multistage impedance matching network. A reasonable choice is a factor of about 30, which gives a fractional bandwidth of $(\frac{1}{30})^{1/2}$, or about 18%. However, to match to a transducer using such mass ratios between a 50 tonne normal mode mass and a 50 g transducer coupling mass requires the use of four stages, with the second-stage resonator exceeding 1 tonne, a third-stage resonator of tens of kilos, coupled to the final low-mass transducer stage. The sphere becomes a rather hairy sphere, possessing not five fundamental quadrupole modes that couple to gravitational waves, but 20. Care must be taken that there are no adverse couplings between the normal modes, since perfect orthogonality will be difficult to achieve in such a complex structure. Negligible cross-coupling can easily be achieved by deliberately choosing coupling mass parameters such that all modes are offset from each other by a few Hz.

A scheme for creating some of the additional resonators required on a sphere is illustrated in figure 31. It makes use of bending flaps, which have the advantage that they have low surface-to-mass ratio (to minimize surface losses), simple geometry for fabrication, and minimum wasted volume. The first intermediate mass is realized by simple machining of the sphere. The second could be welded or cast *in situ*, or it could also be cut from the first intermediate mass by suitable machining (the latter would save more space but has been omitted in the figure for clarity). Just two of the necessary five resonator structures are shown here. This design concept ensures that the antenna and most of the impedance matching system can be fabricated with complete mechanical integrity, thus avoiding unmodelled losses from bolts or glue joints. Here it is worth pointing out that the presence or absence of cylindrical symmetry is not an issue in antenna design, as demonstrated in the NIOBE bending flap at UWA. For large mass ratios the symmetry breaking creates a very small torsional reaction at the suspension point, which could only modify losses if the suspension itself had very large differences in acoustic losses for different degrees of freedom. In reality, suspension systems must be designed to isolate all linear and rotational degrees of freedom since intrinsic cross-coupling always occurs, determined by the Poisson ratio which has a value $\sim$0.3 in all pure homogeneous materials.

The final transducer stage on a spherical antenna can follow the noncontacting microwave parametric transducer concept to avoid the need for cable isolation and save space in the ultracryogenic volume. The transducer could be a sapphire transducer [150–152] or it could be a superconducting re-entrant cavity, as shown in figure 23. For ultralow-temperature applications it is necessary that the transducer have low power dissipation, generally below 1–10 μW. This requires the transducer losses to be lower than niobium transducers demonstrated

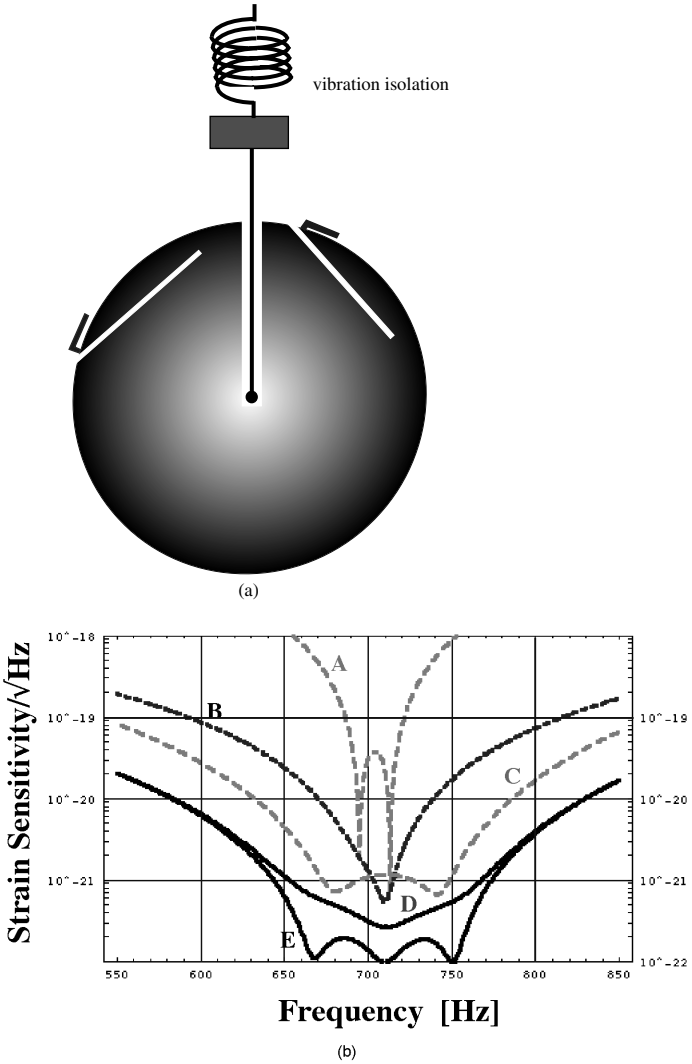


Figure 31. (*a*) A 100 tonne spherical antenna with two of the five cuts required to create ∼1 tonne secondary resonators to readout the orthogonal normal modes. To each secondary resonator is bonded a bending flap ∼30 kg to create the third normal mode of the impedance matching network. The fourth resonator (too small to show here) would consist of an additional bending flap of mass ∼50 g. A parametric transducer could read out such an antenna using radiative coupling thus avoiding vibration coupling through wires. (*b*) [153] Noise performance of a similar spherical detector. Curves A–E show the increasing bandwidth as the number of secondary resonators increased from 1 to 5 respectively.

so far. However, this can easily be achieved by using a larger gap spacing, and a frequency below 10 GHz. In the case of sapphire transducers, sideband pumping allows the dissipated power to be reduced below 1 μW. Figure 31(*b*) [153] shows that broadband noise performance ∼250 Hz can be achieved with a four-mode transducer on a large sphere. The bandwidth is significantly degraded as the number of modes is reduced.

At the transducer it becomes necessary to make a materials transition from the Al or Cu–

Al antenna material, to niobium or sapphire. No definitive solutions for such transitions have been defined, but shrink fitting (by thermal differential contraction), glue bonding or brazing are likely to allow suitably low-loss assembly. It will be important to develop bonding systems in small-scale tests before construction of large-scale spheres.

Spherical detectors cooled to ultralow temperatures can offer difficulties in cooling due to their small surface-to-mass ratio. Frossati has determined that forced convective cooling is essential if the cooling time is to be reasonable. If the material is superconducting, thermal conductivity freeze-out makes this a much more severe problem. For this reason Cu–Al and Be–Cu have been proposed for the antenna material.

4.10.2. Arrays of small high-frequency detectors. A part of the gravitational wave spectrum that has had insufficient attention is the high-frequency band between 2 and 20 kHz. In this band one expects gravitational waves from stellar mass black hole formation and their normal modes. At high frequencies, laser interferometers lose sensitivity due to shot noise, while resonant-mass detectors lose sensitivity due to the smaller size required for such high resonant frequency. Frasca and Papa [154] has proposed a solution consisting of phase coherent arrays of short stumpy antennas, designed so that that all five lower quadrupolar modes (the longitudinal mode, the two discoidal modes and the two 'pantograph' modes) have comparable quadrupole moment and sensitivity. Using five transducers, such stumpies can be read out similarly to the five quadrupole modes of a sphere, and thus have omnidirectional sensitivity. In principle, the total energy sensitivity increases linearly with the total detector mass, so that an array can be enlarged arbitrarily by simply adding additional identical elements.

Whereas a large sphere achieves high sensitivity in a single device, but at the expense of a rather low resonant frequency, the array of stumpies is claimed to achieve the same sensitivity at an arbitrarily high frequency, chosen by astrophysical considerations. It can have cost advantages too, through replication of identical elements.

Such an array obviously requires the individual elements to be sensitive to the same part of the spectrum. However, with adequate impedance matching the stumpies should have bandwidth $\sim$15% similar to that achievable in spheres and bars. This means that there is minimal difficulty in tuning individual detectors.

However, the practical limitations to achieving high sensitivity, such as the limits imposed by noise, have yet to be demonstrated. Further research in this area would be very valuable, including an in-principle demonstration of the concept using calibration signals. The array of stumpies provides an opportunity to achieve great increases in high-frequency sensitivity in the future, but will probably not be implemented until positive results are obtained with existing detectors, or astrophysics provides a very strong justification for improved high-frequency sensitivity.

4.10.3. Transducer developments and prospects. The implementation of improved transducers and impedance matching structures is essential to increasing the bandwidth and burst sensitivity of resonant-mass detectors. Successful transducers have been developed based on capacitative or inductive modulated superconducting circuits with RF or DC SQUID amplifiers, as well as microwave parametric transducers based on superconducting cavities and cryogenic GaAsFET amplifiers. Both systems have comparable sensitivity and a variety of advantages and disadvantages.

The sensitivity of a SQUID-based transducer depends on the SQUID noise performance, and on the mechanical and electrical quality factors of the superconducting elements. A persistent problem has been AC electrical losses in the niobium superconducting coils used for

sensing and coupling transformers in passive transducers. It has sometimes been difficult to obtain sufficient trapped persistent current in inductive sensing circuits, and sufficient electric field across capacitative sense elements. AC losses introduce thermal noise, while low current or electric field reduces the forward transductance, reducing the electromechanical coupling and increasing the antenna noise temperature.

A critical factor for SQUID transducers is the SQUID noise. Recently there has been progress in this area. Two groups have created SQUIDs with noise measured to be below about $30\hbar$, one to two orders of magnitude better than commercial devices. However, recent results [155] indicate that noise is degraded in a transducer environment due probably to trapped flux.

Parametric transducer performance also depends on the mechanical and electrical quality factor of the transducer structure. This is less of a problem however, since the simpler mechanical structures used have shown Q-factors $\sim 10^7$, and electrical Q-factors $\sim 10^5$–10^6, both of which are high enough to be negligible noise sources at current sensitivity. Microwave amplifiers have for many years been shown in radio astronomy to have near-quantum-limited noise, but in the only successful such implementation the amplifier noise contribution is in excess of this [156]. Critical to achieving excellent amplifier noise is very low signal levels, requiring excellent carrier suppression, since otherwise pump power reflected from the cavity greatly exceeds the signal sidebands.

Another critical problem is pump oscillator phase noise. The NIOBE transducer required an ultralow phase noise microwave oscillator to be especially developed.

To achieve sensitivity near to the quantum limit with a parametric transducer will require a higher Q-factor electrical resonator and a lower noise pump oscillator. Fortunately, oscillators 30 dB better than that used on NIOBE have now been developed, and a sapphire transducer with mechanical and electrical Q-factor exceeding 10^8 should allow this technology to advance to within a factor $\sim$30 of the quantum limit [100, 152, 157].

4.11. Vibration isolation and suspension developments

The vibration isolation and suspension system for resonant-mass detectors has always consisted of both room temperature and cryogenic isolation stages such as those discussed above. In the case of ultracryogenic detectors there has been greater emphasis on cryogenic isolation but careful isolation of the cryostat structure itself has been necessary to prevent local disturbances. It is tempting but incorrect to consider the vibration isolation problem solved, since improvements of more than 1000-fold in energy sensitivity are projected, and already detectors show signs of inadequate isolation. In the case of NIOBE there are signs of variable noise temperature by about a factor of two which is observed as degrading mode temperature without change to the wideband noise. Sometimes diurnal variations associated with human activity are observed. This can be understood as arising from vibrational short circuits, perhaps due to a small piece of superinsulation or whiskers of solid air crossing the narrow spaces in the radiative cooler in the suspension tube, or residual conduction through the low pressure (10^{-5} torr) gas in the experimental volume.

EXPLORER is operated at 2 K, below the helium superfluid transition to eliminate the vibrational effects of boiling liquid helium, which otherwise causes degraded performance. NAUTILUS operates well only when local activity is low, again indicating inadequate vibration isolation. Thus it appears that all currently operating detectors are operating close to a vibration isolation limited noise floor, at which improved antennas and transducers will not provide the anticipated noise advantages.

In principle, isolation can be easily improved by further isolation stages, such as the

cantilever spring stages used at UWA. However, short circuits can easily occur, especially through unmodelled resonances in components such as masses, springs or wires, which can cause catastrophic degradation of performance. Not only great care, but excellent design and deep understanding of complex mechanical structures is essential. One point in favour of advanced detectors such as spheres is the advantage of near-nodal-point suspension achieved using centre of mass suspension [126], which can improve isolation by 40–60 dB depending on the proximity of the suspension point to the node. As already emphasized, this can never be perfect due to the finite size of suspension elements and the multimode nature of detectors.

4.12. Conclusion

Resonant bars have been brought to a high level of development. There is now an excellent understanding of the technology and clear proposals for substantial future advances. In particular, it is likely that large spherical detectors will be found to be the best solution for obtaining high sensitivity in the 1–10 kHz range. Meanwhile, an extremely vigorous research effort in laser interferometer detectors is underway, as discussed in the next section.

5. Interferometer detectors

5.1. History

The Michelson interferometer has long been known as an extremely sensitive instrument to measure length changes. The idea of using a laser interferometer as a gravitational wave detector was suggested as early as the 1960s [65, 158] and experimentally investigated in the 1970s [66, 159, 160]. But the first experimental attempt, giving high sensitivity to the displacement of test masses was due to Forward [161]. Forward used a retro-reflector to reflect a beam to a beamsplitter and used active control for locking the interferometer to a fringe. He obtained a spectral strain sensitivity of 2×10^{-16} Hz$^{-\frac{1}{2}}$. The state of the art in sensitive interferometers at the time of writing is represented by three prototype instruments of 10 to 40 m in arm length. These are at Garching [162], Glasgow [163] and CalTech [164]. The 30 m delay line interferometer at Garching has achieved test mass differential position sensitivity of 2.5×10^{-18} m Hz$^{-\frac{1}{2}}$ dominated by shot noise between 1 and 6 kHz. The 10 m Fabry–Perot interferometer at Glasgow has reached $\sim 7 \times 10^{-19}$ m Hz$^{-\frac{1}{2}}$ from 500 Hz to 3 kHz and is close to being limited by shot noise. The 40 m Fabry–Perot interferometer at CalTech has achieved its best displacement sensitivity of 3×10^{-19} m Hz$^{-\frac{1}{2}}$ near 450 Hz. The broadband noise background (neglecting the narrow peaks which can be removed by appropriate filtering of the data) between 300 and 1000 Hz gives a rms differential displacement of less than 10^{-17} m, corresponding to an rms gravitational strain noise level of 2×10^{-19} which is comparable to the sensitivity of current resonant-bar detectors.

In pursuit of increasing the interferometer sensitivity, several optical schemes have been invented. *Power recycling*, proposed by Drever [165], reuses the otherwise unused laser power from the interferometer bright fringe. The light is reflected back towards the beamsplitter, thus increasing the total light power entering the interferometer. *Dual recycling* proposed by Meers [166] allows both laser power and signal power to be recycled, thus increasing sensitivity at the expense of signal bandwidth. Various other schemes such as synchronous *recycling* [165], *detuned resonant recycling* [167], and *resonant sideband extraction* [168, 169] provide various advantages. In general, recycling schemes allow high power built-up in the interferometer arms (which may or may not contain optical cavities) to increase the sensitivity while controlling the signal sideband storage time to maintain detection bandwidth.

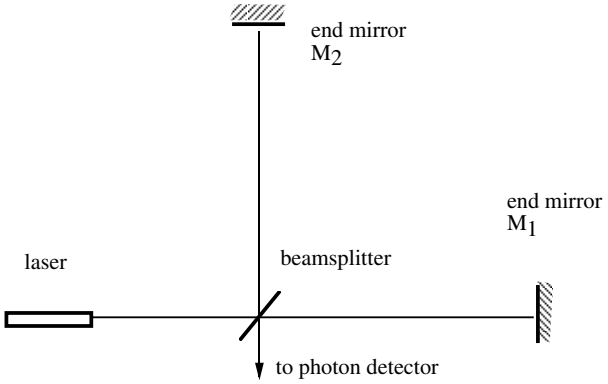


Figure 32. Simple Michelson interferometer.

Power recycling has been demonstrated on both tabletop [170–172] and suspended interferometers [173, 174] with success. Power recycling factors of 300 and 450 have been achieved [173, 174]. The broadband and tuned signal recycling has been demonstrated on tabletop by ANU group [175]. Dual recycling has been tested experimentally by both Glasgow [176] and Max Plank groups [177].

Long-baseline detectors are under construction at: Hanford, Washington and Livinston, Louisiana, USA (American LIGO project) [32], Pisa, Italy (3 km Italy/France VIRGO project) [178], Hannover, Germany (600 m British/Germany GEO project) [179] and Tokyo, Japan (300 m TAMA project) [180].

During the past few years the development of diode-pumped Nd:YAG lasers and ultralow-loss optical coatings has offered greatly improved performance in laser interferometer devices. Nd:YAG lasers are intrinsically stable and efficient. Low-loss coatings (for mirrors working in the infrared frequencies) offer very high recycling gain. Development of polishing techniques for sapphire means that improved material less susceptible to thermal lensing is becoming available. New ideas in suspension and isolation means that it appears likely that future interferometers need only be limited by fundamental limits, while there remains substantial room for future advances.

5.2. Configurations

5.2.1. Simple Michelson. A simple Michelson interferometer detector is shown schematically in figure 32. The interferometer consists of three 'free masses'—one beamsplitter, and two test masses at right angles to form the end mirrors. These masses are vibration isolated and suspended so that at frequencies well above resonance they can move freely as inertial test masses in the direction of the optical path of the interferometer in the frequency range of interest. When a gravitational wave passes it creates relative displacements of the test masses. The relative motion of the end mirrors is read out as intensity variations in the interferometer output, giving information about the incoming gravitational wave.

For simplicity, consider the case of an incident gravitational wave perpendicular to the plane of the interferometer with a polarization direction parallel to the interferometer arms. The passing wave will make one arm of the interferometer shorter and the other longer in half of the wave period, and reverse the contraction–elongation process in the other half-period. The relative change of optical length of the two arms $\Delta L = L_2 - L_1$ can be described as a

phase shift,

$$\Delta F = 2\pi \Delta L/\lambda. \tag{5.1}$$

This results in a change in the interference pattern at the output of the beamsplitter. The relative difference in optical path ΔL is proportional to the arm length $\Delta L = hL$. Generally, an interferometer is sensitive to a linear combination of the two polarization fields, and h in the above equation is

$$h = F_+ h_+ + F_\times h_\times \tag{5.2}$$

where F_+ and $F_\times$ are coefficients depending on the direction to the source and the orientation of the interferometer.

Because the gravitational signal is extremely small, it is very difficult to monitor the small time-varying changes in the interference pattern due to the passing gravitational wave. In practice, the phase difference arising from the optical arm length variations is obtained by the so-called 'nulling method'. The idea is to always keep the light returning from the two arms 180° out of phase so that its output is a dark 'fringe'. When the end mirrors are moved by the passing gravitational waves, the error signals applied to end mirrors to maintain the dark fringe contain the information of the gravitational wave signals. In this way the effect of power fluctuations in the laser beam can be minimized, and the shot noise level can be reduced.

Since the detection of gravitation waves with an interferometer is achieved by measuring the relative optical path (phase) change between the two arms, and since this path difference is proportional to the optical path L, it is clear that the size of the signal can be increased by lengthening the optical path of each arm. However, there is an optimum length $L_{\rm opt}$. At such a length the storage time of the light within the interferometer arms is equal to half the period of the gravitational wave. For arms longer than $L_{\rm opt}$, the gravitational wave signal will change sign during the light travelling time in the arm and the effect will partially cancel out. For example, for a gravitational wave signal of frequency $f_g \sim 1$ kHz, the optimum arm length is $L_{\rm opt} = c/(2f_g) = 150$ km. Such a long optical path length can be easily realized in a space-based interferometer (see section 3.2). In an Earth-based interferometer this is made possible by using multi-pass techniques. One such technique is the multi-pass Michelson interferometer in which an optical delay line [66] is inserted in each of the interferometer arms. The other is a Fabry–Perot Michelson interferometer [181] in which a second mirror is inserted in each arm to form a Fabry–Perot cavity.

5.2.2. Delay line Michelson interferometer. A delay line Michelson interferometer is shown schematically in figure 33. The two beams coming out of the beamsplitter are reflected many times between the beamsplitter and the end mirrors before they are recombined. For example, a 3 km long interferometer can have an optical length of 150 km by having 50 bounces. Apart from the restriction that the optical path length be shorter than $L_{\rm opt}$, the useful number of reflections is in practice limited by the reflection losses at the mirror.

A practical difficulty of the delay line Michelson interferometer configuration is the scattered light problem. The delay line uses a large number of beams zigzagging back and forth between mirrors. It is easy for light to scatter from the mirrors or from the side of vacuum pipes into the main beam at the photodetector. This scattered light could have a large phase difference $\alpha = 2\pi f \delta l/c$ with respect to the main beam, where f is the laser light frequency and δl is the path difference from the main beam. When the scattered light interferes with the main beam, it results in a phase shift β of the recombined light given by $\beta \sim a \sin\alpha$, where a is the fraction of the scattered light. This change of phase will be sensed by the output photodetector when combined with the beam from the other arm. It can be seen that any fluctuation of

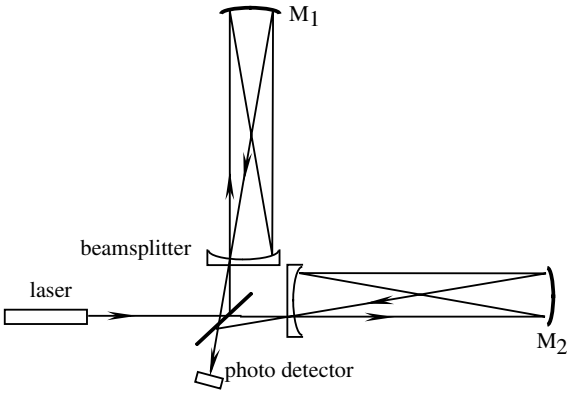


Figure 33. Schematic diagram of a delay line Michelson interferometer.

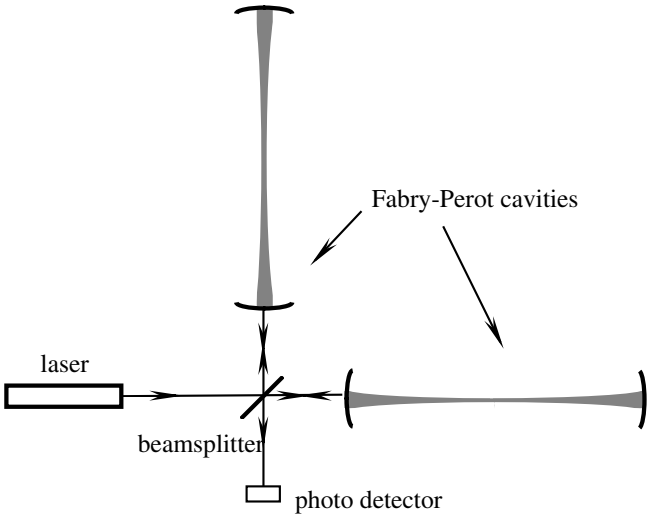


Figure 34. Schematic diagram of a Fabry–Perot cavity interferometer.

laser frequency or vibration of the vacuum pipes (which results in a change of δl) will cause a phase fluctuation of each beam, and thus a fluctuation in the final signal. This effect can be reduced by stabilizing the laser frequency. Modulation of the laser light [182] can improve the performance to some extent. This is done by changing the laser frequency f such that over a certain measurement time the scattered light phase difference α changes from 0–$2n\pi$. The average effect of the scattered light is then zero. This technique is limited by the dynamic range of the modulation. Another disadvantage of a delay line Michelson interferometer is that it needs large mirrors. However, excellent sensitivities of the order $h \sim 10^{-19}/\mathrm{Hz}^{1/2}$ have been obtained using a delay line Michelson interferometer of arm length 30 m [183].

5.2.3. Fabry–Perot cavity interferometer. The Fabry–Perot cavity interferometer gravitational wave detector was first introduced by Drever *et al* [181]. The idea is to add two additional mirrors near the beamsplitter, as shown in figure 34.

The near mirror and the end mirror in each arm form an optical cavity. Light travelling

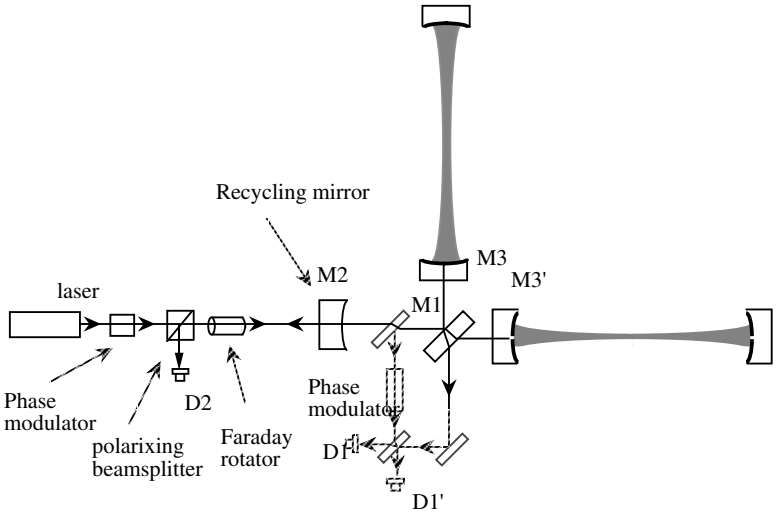


Figure 35. Basic arrangement of a Fabry–Perot interferometer system [324].

in one arm is reflected between the same pair of spots on the two mirrors, forming a sharp resonance. This Fabry–Perot cavity is very sensitive to the changes in cavity length, and to the change in frequency of the light. Because no laser has achieved the required stabilization level to directly detect the change in cavity length induced by gravitational waves, two cavities are needed to detect differential changes. The light from a single source passes through a beamsplitter and illuminates the cavities in the two arms of the interferometer. By looking at the differential phase change of the two arms, the effect of frequency fluctuations of the light source can be eliminated, and the signal is a measure of the passing of a gravitational wave. The size of the suppression of the laser frequency fluctuation noise depends on the balance of the two cavities. Typically, with a highly stabilized laser, the cavities must be balanced to one part in 10^3 or better.

Another approach is to lock the laser in wavelength to one of the cavities, and then to lock the second cavity to the laser wavelength. The locking signal of the second cavity then gives the relative cavity length change with respect to the first cavity, due to the passing gravitational wave. A practical arrangement of a Fabry–Perot cavity interferometer with such a readout system is shown in figure 35.

The Fabry–Perot cavity interferometer has the advantage of having smaller mirror size, and thus smaller vacuum pipe size, than that in a delay line Michelson interferometer. Also, the scattered light problem can be reduced greatly in a Fabry–Perot cavity arrangement because the scattered light is made to travel the same path as the main beam in the cavity.

5.2.4. Sagnac interferometer. A Sagnac interferometer is schematically shown in figure 36 in which light beams travelling in opposite directions experience common optical paths. With its common-path nature, the Sagnac interferometer is insensitive to a range of noise sources that affect other interferometers. Noise from low-frequency mirror displacements, laser frequency and intensity fluctuations, laser beam pointing fluctuations, thermally induced birefringence, and reflectivity asymmetry in the arms are all suppressed. This means that the Sagnac configuration can have a simplified control system and reduced optical tolerance requirements. Low-temporal-coherence illumination can be used in a common-path interferometer to reduce

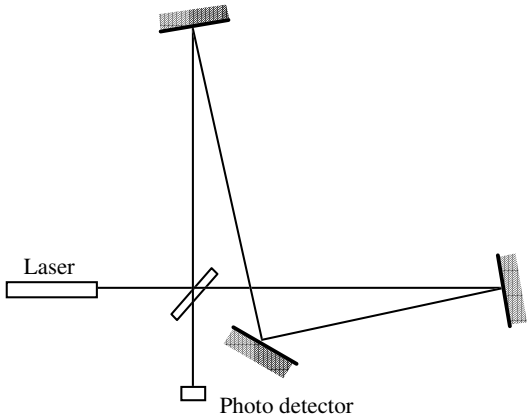


Figure 36. Schematic diagram of a Sagnac interferometer.

the noise caused by parasitic paths introduced by scattered light. In 1986, Weiss proposed an open-area Sagnac interferometer for gravitational wave detection. Recently, the zero-area Sagnac interferometer shown in figure 36 has been analysed and experimentally investigated as a topology for an advanced gravitational wave detector [184,185]. For 4-km 20-bounce (storage time of 0.53 ms) LIGO-scale interferometers illuminated with a 1064 nm Nd:YAG laser, the Sagnac interferometer has its first peak response at 690 Hz with 3 dB bandwidth from 220 to 1250 Hz [184], as shown in figure 37. Proper setting of the storage time allows the peak response frequency to be tuned to the gravitational wave band of interest. A shot-noise-limited phase sensitivity of 9×10^{-10} rad $\mathrm{Hz}^{-1/2}$ has been achieved on a tabletop Sagnac interferometer [185]. It has been demonstrated that precision phase measurement can be performed with a laser bearing a substantial amount of frequency and amplitude noise [185]. The Sagnac interferometer with resonant sideband extraction has been demonstrated on the tabletop [186].

The comparison between the Sagnac and Michelson interferometers is detailed by Mizuno [187] and colleagues. Their conclusion is that, once cavities are used either in the arms or for power- or signal-recycling, the advantages of the Sagnac interferometer over the Michelson interferometer disappear. Without power recycling, $\gg$1 kW laser with quite high stability is required to achieve the desired sensitivity for Sagnac interferometer.

5.2.5. All-reflective interferometer. Extremely high light power incident on the beamsplitter is needed to reduced the photon shot noise which is a major limiting factor in the high-frequency regime ($\sim$1 kHz) in laser interferometers. However, high light power poses problems for transmissive optics such as beamsplitters and the input/output mirrors of Fabry–Perot cavities, because there is always some optical absorption causing heating. The optical components are heated both by absorption in the substrate and in the reflective coatings. Heat from both regions leads to thermal lensing and birefringence in the substrate, and also to distortion of the optical surfaces. The Sagnac interferometer might alleviate these problems since the counterpropagating laser beams in principle share the same optical path.

A more certain method of reducing adverse thermal effects is to eliminate transmissive optical components completely. The losses in optical coatings are generally less than absorption losses, so by using reflective diffractive components instead of lenses and beamsplitters, it should in principle be possible to make interferometers much more tolerant of high optical powers. The idea of an entirely diffractive reflective interferometer has

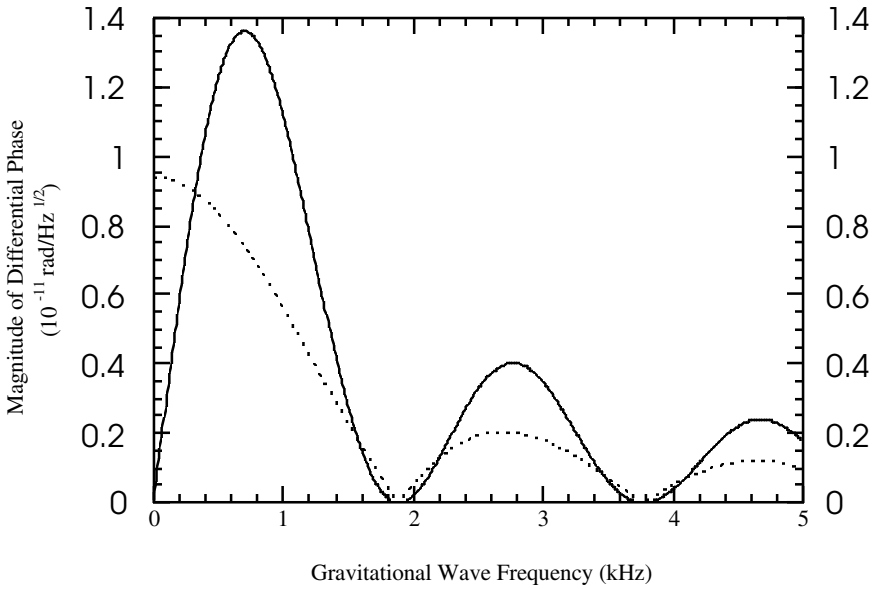


Figure 37. Theoretically calculated differential phase generated by a gravitational wave with strain $h_g \sim 10^{-23}/\sqrt{\text{Hz}}$, in a 20-bounce delay line, 4 km arm length Sagnac interferometer illuminated by a 1064 nm laser. The Michelson interferometer response (dotted curve) is shown for comparison [184].

been discussed by Drever [188]. The concept would allow test masses to be created using nontransmissive materials such as silicon or niobium, where low thermal noise did not have to be combined with excellent optical properties. One possible interferometer configuration using a diffractive beamsplitter is shown in figure 38. Veitch *et al* [189] have reported tests of a holographic beamsplitter used to remove optical aberrations and guarantee high fringe visibility and complete destructive interference at the interferometer output even in the presence of aberrations. An experimental demonstration of various grating beamsplitter tabletop interferometers (Michelson, Sagnac and Fabry–Perot) has been performed at Stanford [190]. For a practical diffractive interferometer it will be necessary to perfect reflective diffractive elements. If optical cavities are to be used, the mirror coatings will have to include a diffractive coupling beam. For example, a mirror may be required that has 99.999% reflection, and 0.001% coupling into a beam that leaves the mirror at a suitable angle. This problem sets a challenge for optics in the next decade, which can lead to major improvements in gravitational wave detection in the future.

5.3. Recycling

5.3.1. Power recycling. The sensitivity of an interferometer is ultimately limited by shot noise due to photon quantum statistics. The standard quantum limit for an interferometer can be obtained from the balance of two competing quantum noise sources as described by Caves [191, 192], Braginsky *et al* [112] and others. The first is the photon-counting error due to $N^{1/2}$ fluctuations in the number of output photons from the interferometer. The second is the radiation-pressure error. This arises from the perturbations on the end mirrors produced by fluctuating radiation-pressure forces which also scale as $N^{1/2}$. As the input laser power P increases, the relative photon-counting error decreases as $N^{1/2}/N$, while the radiation-pressure

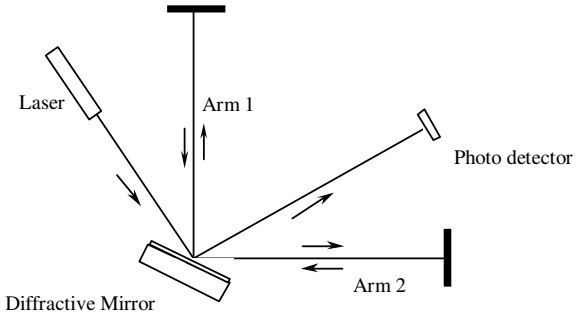


Figure 38. Schematic diagram of a Michelson interferometer based on the use of reflective diffraction grating as a beamsplitter.

error increases as $N^{1/2}$. Minimizing the total error with respect to P yields a minimum error of order of the standard quantum limit and an optimal input power for a simple Michelson interferometer,

$$P_{\text{opt}} = \frac{\lambda mc}{8\pi\tau^2\sqrt{\eta}}, \tag{5.3}$$

at which the minimum error can be achieved. Here, m is the mass of an end mirror, λ is the wavelength of the light, τ is the measurement duration, and c is the light velocity.

With a reasonable set of values for interferometer parameters, $m \sim 10^2$ kg, $\tau \sim 10^{-3}$ s, $\lambda \sim 1\ \mu$m, the optimum laser power P_{opt} is approximately 6×10^7 W—a power far higher than the power of present CW lasers. The low available input power means that the interferometer for use as gravitational wave detectors will be limited not by the standard quantum limit, but rather by photon-counting statistics (shot noise) which scales inversely to the square root of incident power at beamsplitter.

As mentioned above, the technique of power recycling [165, 166] can be used to increase the power incident at the beamsplitter and improve the sensitivity of interferometer detectors. The basic idea is that because the interferometer detector operates at a dark fringe output, almost all of the light (reduced only by losses in mirrors and beamsplitters) is reflected back towards the laser, and can therefore be used again as long as it is phase coherent with the input laser beam. This technique is realized by inserting a recycling mirror M_2 in between the laser and the beamsplitter as shown in figure 35. The position of the recycling mirror (or the laser frequency) is then carefully adjusted so that the recycling mirror combined with the two main cavities and the beamsplitter form a large resonant optical cavity containing the interferometer. By doing so, an effective laser power perhaps 1000-fold larger than the original laser may be built up inside this cavity, thus reducing the shot noise.

5.3.2. Resonant recycling. The resonant recycling technique [165] uses a mirror arrangement such that after each half gravitational wave period, the light in the two arms exchange arms instead of recombining at the output photodetector. In this way, the light of each beam always experiences phase shift in the same direction. The phase shift builds up during the total storage time over many gravitational wave cycles. In the end, a large phase difference between the two arms can be detected at the output of the interferometer. This technique can be used in both delay line Michelson interferometers and in Fabry–Perot interferometers. Figure 39 shows resonant recycling arrangements for both delay line interferometer and Fabry–Perot cavity interferometer.

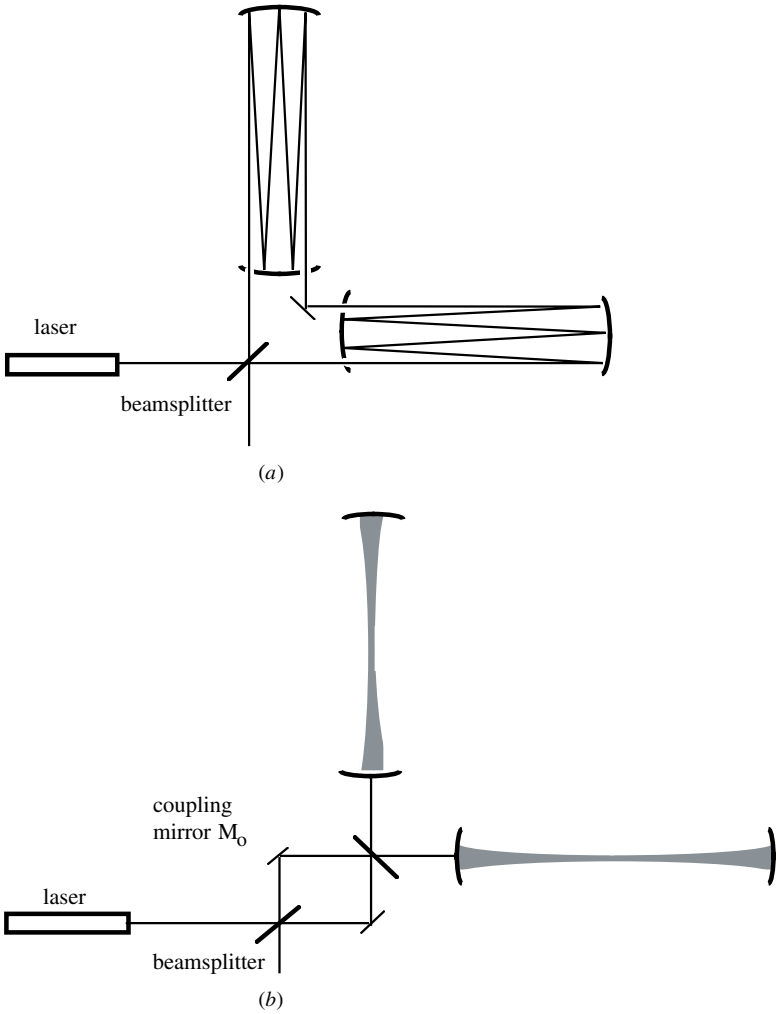


Figure 39. (*a*) Resonant recycling arrangement for a delay line interferometer. (*b*) Resonant recycling arrangement for a Fabry–Perot cavity interferometer.

In the case of resonant recycling in a Fabry–Perot cavity interferometer, it is considered that the two cavities are coupled through the high-reflectivity mirror M_0. This coupled system then has two modes of oscillation. The interferometer is tuned so that one of the modes of the coupled optical cavity system matches the frequency of the laser light and the other matches the frequency of the optical sideband produced by the motion of the end mirrors due to an incident gravitational wave. Both the laser light and the sideband signal produced by gravitational wave are enhanced at the output. In principle, resonant recycling gives better sensitivity for detecting periodic signals of known frequency. The total storage time is limited by the losses of the mirrors.

5.3.3. Dual recycling. A simple arrangement for dual recycling [193] is shown in figure 40. In addition to the power recycling mirror M_3, a new signal recycling mirror M_4 is placed at the output of the interferometer. Generally, the light at laser frequency (carrier) cannot

reach M_4 when the interferometer is locked on a dark fringe. However, when the gravitational wave arrives, it modulates the interferometer arm length difference and generate sidebands at frequencies offset from the laser frequency by the gravitational wave frequency. The sidebands leak from the interferometer output towards M4 as gravitational wave signal. So M_4 has no effect on the carrier but reflects the signal sidebands back to the interferometer. If we adjust the position of the mirror M_4 to allow the reflected signal sidebands and the signal in the interferometer arms to be in phase, the signal power (sensitivity) will be increased. The relative position of the recycling mirror M_4 determines the tuning frequency of the dual recycling. The bandwidth is determined by the reflectivity of the signal recycling mirror and the loss of the two arms. The dual recycling with tuning frequency at zero is called broadband dual recycling. The relative signal-to-noise ratio of a 3 km, 16-reflection delay line interferometer with different degrees of tuned dual recycling is shown in figure 41. The central frequencies are at 200 and 1000 Hz respectively. Even a relatively short arm interferometer can obtain impressive sensitivity. For example, figure 42 shows the predicted shot-noise-limited strain sensitivity of a 400 m, four-pass delay line [194]. The input power is 5 W. The tunned frequency is at 200 Hz. Curves (a) and(c) are based on application of recycling factors (450) already achieved in the laboratory [174], curves (b) and (d) employ somewhat higher factors.

5.3.4. Simple dual recycling instruments. The advantage of the Fabry–Perot Michelson interferometer is that high laser power and low mechanical noise requirements are predominantly restricted to optical cavities. Thus one needs only four very-high-performance components (the main Fabry–Perot mirrors) while the beamsplitter and other components are far less critical. When this scheme is extended to dual recycling as shown in figure 43. The interferometer consists of nested cavities, a pair of Fabry–Perot cavities within the overall power recycling and signal recycling cavities. This requires rather complex control systems.

A much simpler arrangement would be a four-pass dual recycling Michelson interferometer (GEO project) as illustrated in figure 43 [195]. Such a system utilizes only two cavities. Now, to regain the sensitivity of the Fabry–Perot Michelson, the interferometer must use high levels of power recycling, plus moderate signal recycling. Because of the thermal lensing problem [196], this arrangement requires excellent mechanical and optical performance of the beamsplitter.

It appears that very-low-loss silica beamsplitters may allow recycling factors up to 10^4 [197]. An alternative would be to use sapphire beamsplitters, which have intrinsic advantages associated with their high rigidity, and high thermal conductivity. However, the disadvantage of sapphire is its optical birefringence which requires control of the orientation of the crystal relative to the input and output beams.

5.3.5. Resonant sideband extraction. Resonant sideband extraction [168, 169] is a similar configuration to dual recycling for laser-interferometric gravitational wave detectors with Fabry–Perot cavities in the arms. This scheme reduces the thermal load on the beamsplitter and the coupling mirrors of the cavities and allows one to adapt the frequency response of the detector to a variety of requirements.

To obtain a good sensitivity in interferometric gravitational wave detectors one requires high light power in the arms of the interferometer to increase the photon shot-noise-limited signal-to-noise ratio. This can be done by increasing the finesses of the arm cavities. But since high finesse cavities have narrow bandwidth, (i.e. long optical storage time) this sets a limit to the detector bandwidth. The same power build-up can equally be obtained in principle by using power recycling to compensate for the limitation to the power enhancement in the

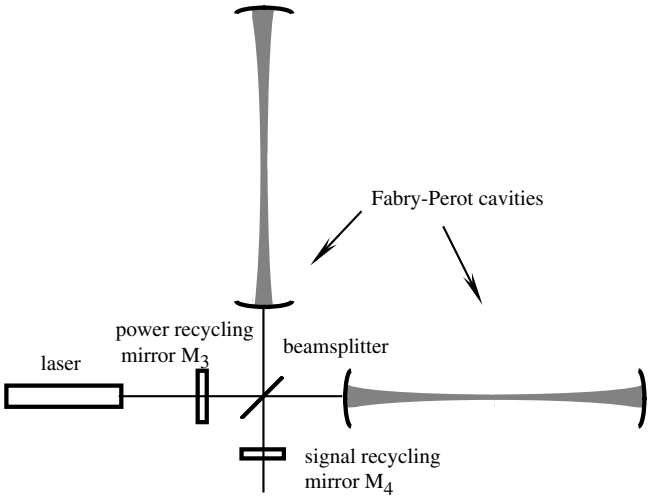


Figure 40. Dual recycling arrangement of an interferometer.

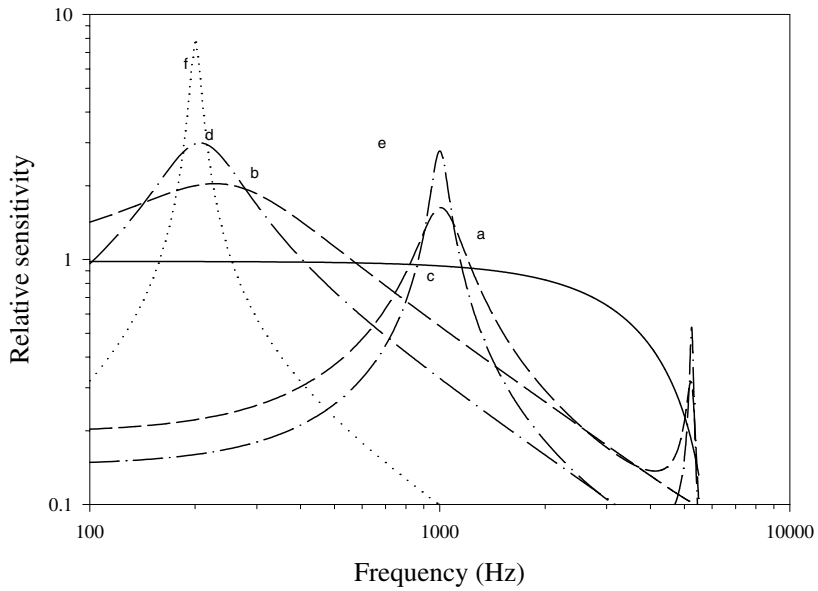


Figure 41. Relative signal-to-noise ratio of a 3 km 16-reflection delay line with different degrees of tuned dual recycling. The tuning frequencies are at 200 Hz and 1 kHz respectively. The vertical axis is in arbitrary units. The solid curve is for power recycling only; the dashed curve is for signal recycling mirror reflectivity, $R_s = 75\%$; the dash-dotted curve for $R_s = 90\%$; the dotted curve for $R_s = 99\%$ [325].

arm cavities. However, in practice the power recycling gain achievable is likely to be limited by imperfect contrast as well as losses in the beamsplitter and the coupling mirrors of the arm cavities. Furthermore, as already discussed, high power may induce thermal lensing and birefringence in the beamsplitter.

Resonant sideband extraction allows this dilemma to be avoided in an interferometer with

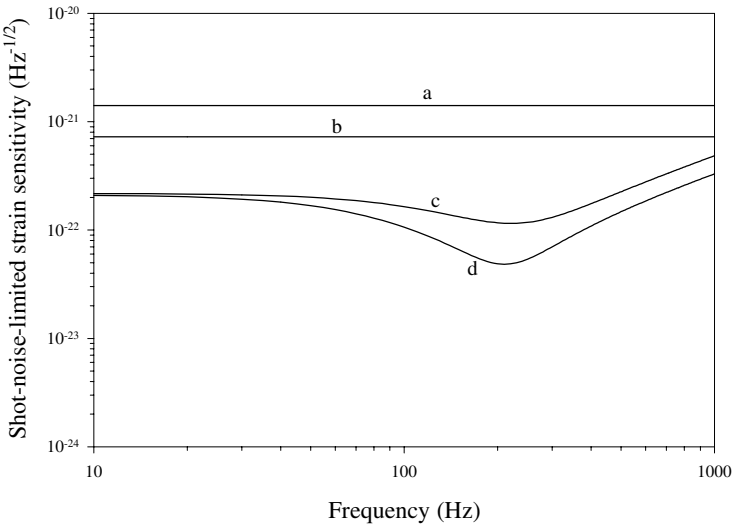


Figure 42. Frequency response of a 400 m interferometer. Input power is 5 W. Curve (*a*) power recycling factor of 450, no signal recycling; curve (*b*) power recycling factor of 1700; curve (*c*) dual recycling with power and signal recycling factors of 450; curve (*d*) power recycling factor of 1700, signal recycling factor of 800 [194].

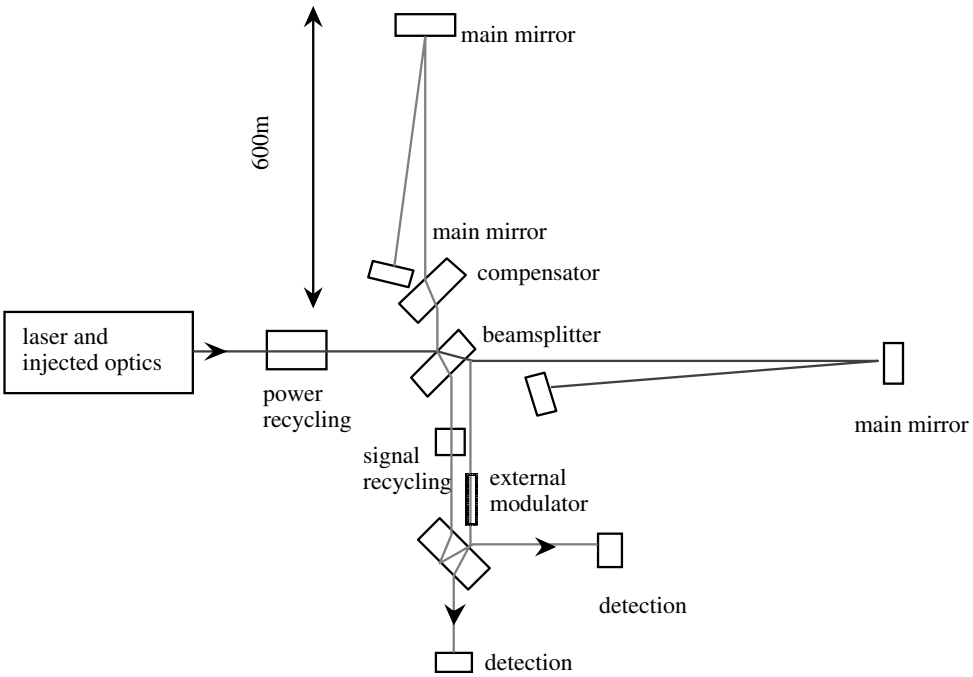


Figure 43. Schematic diagram of a four-pass delay line interferometer with both power and signal recycling (i.e. dual recycling). The angles and lengths are not to scale. The test masses of the interferometer are formed by the main mirrors and the beamsplitter [195].

arm cavities. The scheme resembles signal recycling, but uses a *signal extraction mirror* between the beamsplitter and the photo-detector. The purpose of this mirror is to *decrease* the storage time of the signal sidebands and therefore increase the detector bandwidth. This is achieved because the signal extraction mirror and arm cavities form a three-mirror-coupled cavity. Because the interferometer is locked to dark fringe the signal extraction mirror has no effect on the carrier. The signal extraction cavity formed by the signal extraction mirror and the coupling mirrors of the arm cavities forms a compound mirror which has frequency-dependent transmittance and reflectivity. Tuning the signal extraction cavity allows the transmittance for the signal frequencies of interest to be higher than that of arm cavities alone. For these frequencies the storage time in the three-mirror cavity is shorter than that in the unmodified arm cavity. In this case, the reduction of the storage time results in an increased detection bandwidth and unchanged high-finesse arm cavities for the carrier. In principle, the power enhancement in the arm cavities could be so great that no power recycling would be required. Yet the power passing through the beamsplitter and the coupling mirrors of the arm cavities could be low enough to have little thermal load. Figure 44 shows the frequency response of 3 km arm length interferometer with resonant sideband extraction configuration at various conditions [168].

5.4. Vibration isolation

At Earth-based sites for gravitational wave detectors, the ground is continuously vibrating with a rms amplitude of $x_s \approx \alpha f^{-2}$ m Hz$^{-\frac{1}{2}}$, where $\alpha \approx 10^{-6}$–10^{-9}. This is far greater than the signals we want to measure. Thus it is of great importance that terrestrial gravitational wave detectors, both resonant-bar and laser interferometers, are isolated from the seismic noise background. High-performance mechanical vibration isolation systems are required for this purpose. An ideal vibration isolator would not only cut out all significant seismic vibration in the pass band of the gravitational wave detector, but also cut out the seismic noise at much lower frequencies, so that the suspended test masses were effectively stationary with respect to the laser light field. If the total rms motion were much less than an optical fringe width, the servo control requirements would be minimized: components would be as stable as if rigidly attached to an optical table, or placed in interplanetary space. Operation of an interferometer would then be very simple. We show below that total rms motion from all frequencies above 0.2 Hz can in principle be reduced to about 1 nm.

This ideal level of isolation has not yet been achieved but there are various approaches that when combined, should approach the ideal performance discussed above. We subdivide the approaches to this problem under the headings passive isolation, active isolation and pre-isolation.

5.4.1. Passive mechanical isolation. Passive mechanical vibration isolators are mass–spring low-pass filters. For a multistage isolator, each stage of the isolator with resonant frequency f_0 will attenuate vibration by a factor $(f_0/f)^2$ at frequencies $f \gg f_0$. The total attenuation of the multistage isolator at frequencies above the *corner frequency* (the highest normal mode frequency of the isolator) is $(f_1 f_2 \ldots f_N/f^N)^2$, where N is the number of stages.

Figure 45 shows the typical behaviour of such an isolator. The figure shows a typical transfer function of a five-stage isolator with $f_0 = 2$ Hz. Below the corner frequency the normal modes amplify the seismic noise, while above it the isolation improves as the tenth power of the frequency ratio.

Traditionally, vibration isolators for gravitational wave detectors were based on industrial vibration isolators, using systems such as alternating layers of lead (or steel) and rubber [5].

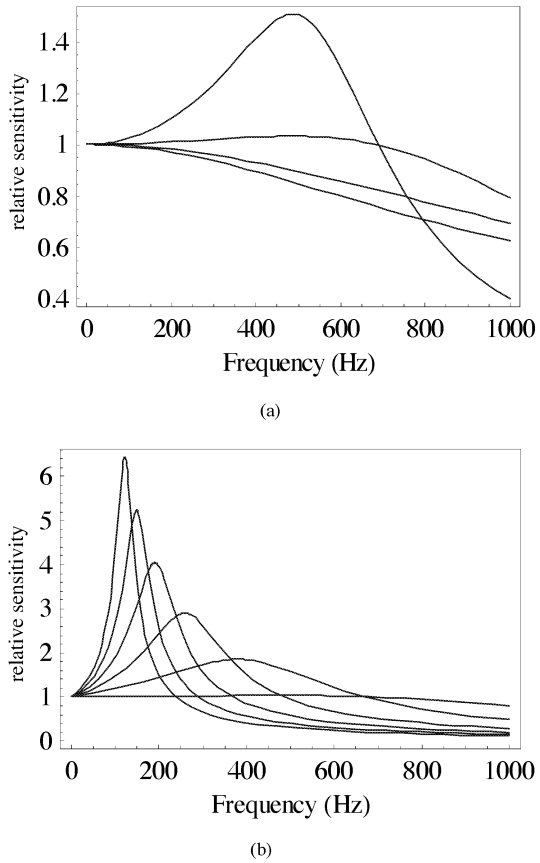


Figure 44. Typical frequency response of a 3 km arm length interferometer with resonant sideband extraction configuration. (*a*) Dependence of the frequency response on the length of the sideband extraction cavity (SEC) when the carrier is resonant in SEC. The curves, in order of increasing response at 400 Hz, are for SEC length of 1 m, 30 m, 100 m, 300 m. (*b*) Dependence on the tuning of SEC to carrier. The broadest response corresponds to the carrier resonant in SEC, and the others are detuned from it (carrier resonant condition). The length of SEC is 100 m, and detuning is by steps $2\pi/1000$. In both figures, the vertical units are arbitrary [168].

Isolators of this type have been used for room temperature vibration isolation in both resonant-mass gravitational wave detectors and prototype laser interferometer detectors. The disadvantage of these isolators is that they are not suitable for use in high vacuum (unless the rubber is outgassed or packaged in metal bellows); moreover, they are not suitable for uses at low temperatures where elastomer materials harden. In addition, this type of isolator generally shows large temperature coefficients and drift, due to the properties of the rubber, as well as having a relatively high corner frequency, limited by the compressive yield of the elastic elements.

Ideally, a multistage pendulum could be used to provide sufficient horizontal isolation for a laser interferometer gravitational wave antenna. Several groups have used double or triple pendulum suspensions combined with rubber and steel isolation stacks [198–201]. However, the curvature of the Earth creates intrinsic cross-coupling between the horizontal and vertical directions, due to the fact that the laser beam in an interferometer can only be perpendicular to the local vertical at one location. In a large-scale laser interferometer detector, the magnitude

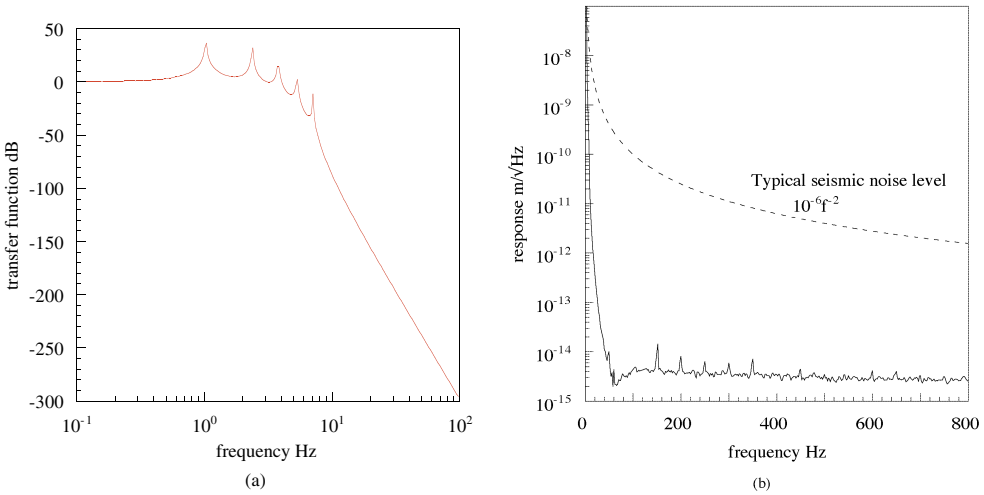


Figure 45. (*a*) Theoretical transfer function of a five-stage vibration isolator. Each stage has a natural frequency of 2 Hz. (*b*) Experimental upper limit of performance of a five-stage isolator [326], with a typical seismic curve of $10^{-6} f^{-2}$. At high frequencies the isolation reaches the noise floor of the sensing transducer of 3×10^{-15} m Hz$^{-1/2}$. Below 58 H the mechanical resonant frequency of the transducer, the sensitivity of the transducer degraded. Practically all the data above 50 Hz are transducer noise.

of the cross-coupling in ideal circumstances is $\sim 10^{-3}$. In real isolators the cross-coupling is likely to be degraded by mechanical imperfections. Thus it is important that the vertical vibration isolation should also be very high. The isolators must also be strong enough to support a total weight $\sim 10^2$–10^3 kg.

In interferometer antennas, the corner frequency needs to be pushed as low as possible to create the broadest possible bandwidth for observations. There is a great advantage in operating an interferometer antenna at the lowest possible frequencies. Not only does it extend the range of sources accessible to the detector, but for specific sources such as binary neutron star coalescence events, it increases the number of cycles of the coalescence that can be observed, thus allowing the signal-to-noise ratio to be increased. Giazotto pioneered the development of low-frequency vibration isolation. A very-large-scale multistage low-frequency superattenuator based on a gas spring was developed at Pisa [202, 203]. It had a high load-bearing capacity and a corner frequency of 2–3 Hz. However, gas springs have strong temperature coefficients, and so such isolators have problems of thermal stability and complexity. At the University of Western Australia, tapered metal cantilever spring vibration isolators [118], were developed, which showed excellent performance but somewhat higher corner frequency. The Pisa group replaced gas springs with similar tapered cantilevers and reduced the mechanical frequencies by use of magnetic antisprings, created by using magnetic repulsion between like poles of permanent magnets. This allowed low-frequency behaviour similar to the gas spring system to be achieved [204].

5.4.2. Active isolation. Active isolation techniques have been investigated extensively in gravitational wave research [205–210]. The basic idea is simple, as shown in figure 46: the relative displacement between the test mass to be isolated and the suspension point (the error signal) is sensed and fed back to an actuator to servo the suspension platform so that the motion of suspension point follows the motion of the test mass. In this way, the motion of the

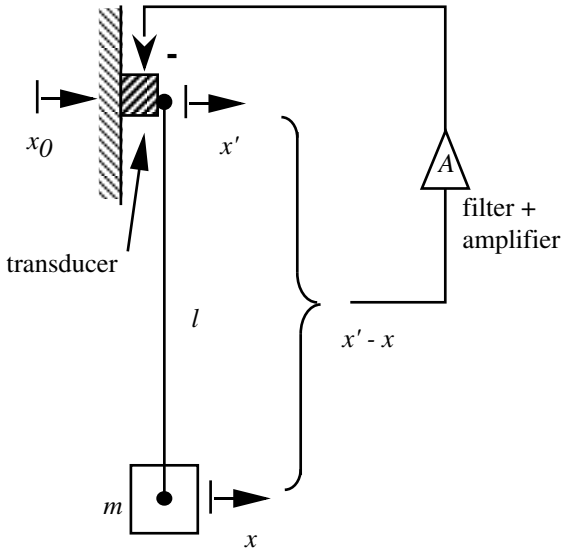


Figure 46. Schematic diagram of an active pendulum [205].

suspension platform is reduced and hence the motion of the test mass. If we look at the simple case of figure 46, the equation of motion of the mass is (for simplicity we ignore the damping term)

$$\ddot{x} = \omega_0^2(x' - x_0), \qquad \omega_0^2 = \frac{g}{l}. \tag{5.4}$$

Assuming the error signal is fed back to the suspension point so that $x' = x_0 - A(x' - x)$, then the transfer function of the system is

$$\frac{x}{x_0} = \frac{\omega_0^2}{-\omega_2(1+A) + \omega_0^2} = \frac{\omega_0^2/(1+A)}{-\omega^2 + \omega_0^2/(1+A)}. \tag{5.5}$$

It can be seen that this system behaves like a pendulum with an equivalent resonant frequency of $\omega_0/\sqrt{(1+A)}$. At frequency $f \gg f_o$, the attenuation is $[f_o/f/(1+A)]^2$ instead of $(f_o/f)^2$. This means that the isolation of mass m is improved by a factor of $(1+A)$.

The basic arrangement for vertical active isolation is shown in figure 47. The final transfer function is slightly complex. However, from the block diagram (figure 48), it can be seen that

$$x = G'x', \qquad G' < 1, \tag{5.6}$$

$$x' = G[xo - H(x' - x)] \qquad G < 1, \tag{5.7}$$

where GG' is the passive transfer function of the system and H is the loop gain of the feedback. The closed loop transfer is then

$$\frac{x}{x_0} = \frac{GG}{1 + H(G' - G'G)}. \tag{5.8}$$

The performance of this active system is improved by a factor of $[1 + H(G' - G'G)]$. However, it is impractical to assume very large gain H to obtain high-performance isolation. There are several limitations. One is the presence of internal resonances of the isolation structure. At each resonance there will be a phase shift added to the loop transfer function which can make the servo unstable. To ensure stability, the gain of the servo should be within the limit that

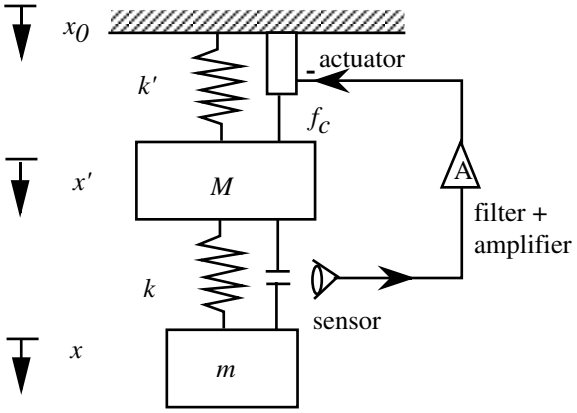


Figure 47. Schematic diagram of the active isolation model [206].

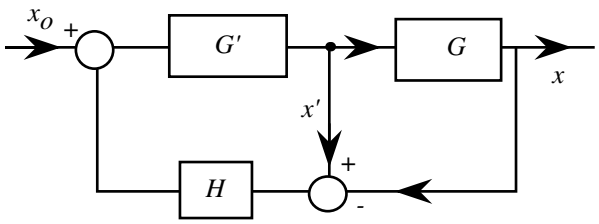


Figure 48. Block diagram of the active isolator [209].

all the internal resonant peaks are below unity gain. Another limitation is the noise in the sensor. This noise is treated the same by the servo system as the error signal between test mass and suspension platform. This noise will be amplified and fedback to the platform to create a displacement. At low frequency, where the attenuation of the test mass from the platform is not high, sensor noise plays an important role since the test mass tends to follow the movement of the platform.

5.4.3. Residual motion and the need for ultralow-frequency pre-isolation. A suspended test mass in a laser interferometer is required to be very precisely located to within $\leqslant 10^{-6}$ of an optical wavelength. Residual motion of the test mass in interferometers operated to date is typically ∼few microns. This has caused two problems. First, it makes it difficult to acquire lock, because very large forces must be applied to decelerate the test mass and locate it within operating range. Second, it makes it impossible to act directly on the test mass (by say a magnetic actuator) because the electronic noise of the actuator circuit is never less than 10^9 times smaller than the maximum signal which can be applied. (This is the dynamic range of a low-noise amplifier.) If the maximum signal is able to correct 10^{-6} m of motion, then the noise level will create noise motion just 10^9 times smaller, i.e. 10^{-15} m. This is unacceptably large.

There are two possible solutions to the problem. One is to create a more complex servo system and, in particular, to apply control forces to the stage above the test mass. All detectors constructed to date use this approach. The second possibility is to greatly reduce the residual motion.

Residual motion, and in particular residual accelerations, can be greatly reduced if a

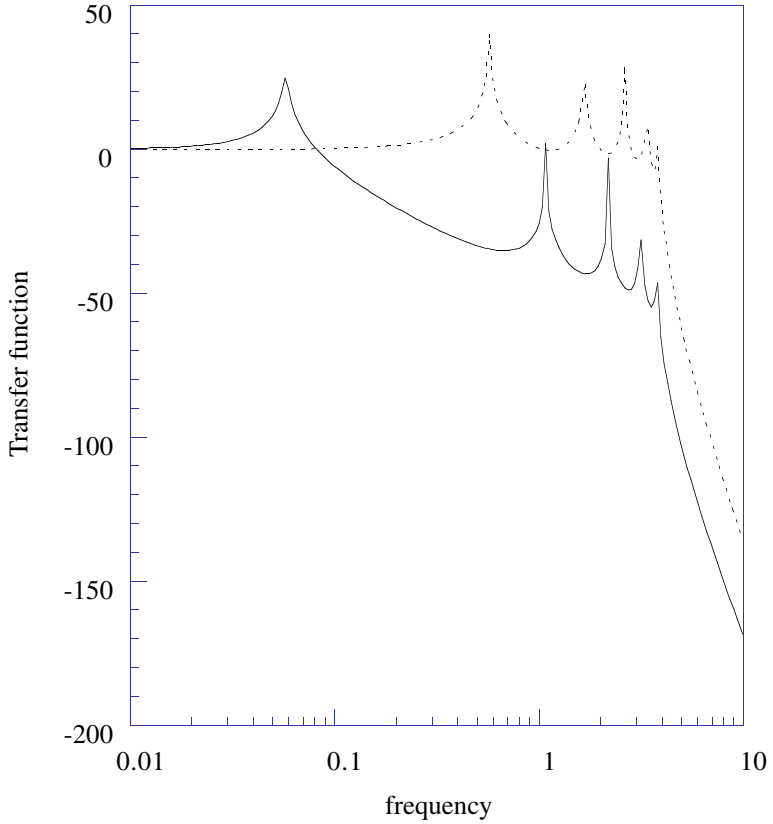


Figure 49. Comparison of transfer functions of a multistage isolator with and without an ultralow-frequency stage. Dotted curve: five-stage 2 Hz isolator, solid curve: four-stage 2 Hz isolator plus a 0.1 Hz ultralow-frequency pre-isolator.

conventional isolator is suspended by a pre-isolator with a much lower resonant frequency. This is illustrated in figure 49, where the addition of a single low-frequency stage reduces the amplitude of the normal mode peaks by about 50 dB. To realize this advantage requires the development of ultralow-frequency (ULF) mechanical suspension stages. As mentioned before, the isolation performance of an isolator with resonant frequency of f_o above the corner frequency is $(f_o/f)^2$. Thus the lower the resonant frequency f_o, the better isolation performance and the lower the corner frequency. This does not guarantee that with one or several ULF stages one can obtain very high isolation performance at high frequency. The problem is that internal mechanical resonances usually occur at frequencies that are typically about 10^2–10^3 times the fundamental resonant frequency of the mechanical resonant structures. Thus, for example, a stage with 10^{-1} Hz resonant frequency is likely to have an internal resonance from 100 Hz upwards and these will corrupt its isolation performance.

As a result, ULF stages are best used as pre-isolation stages in conjunction with low-frequency isolators. For an N-stage isolator, there are N normal mode peaks with amplitude up to 100 times the seismic background, depending on the Q-factor of the isolation elements. It is predominantly these normal modes which make it difficult to control and lock the interferometer. As discussed in section 5.4.1, different methods have been used to damp the normal modes, such as magnetic eddy current damping [200,211] and vibration absorber [212].

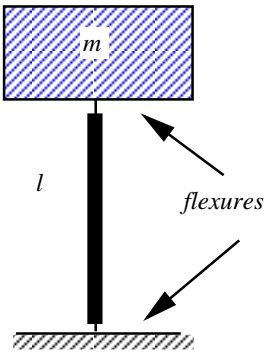


Figure 50. Simple model of an inverted pendulum.

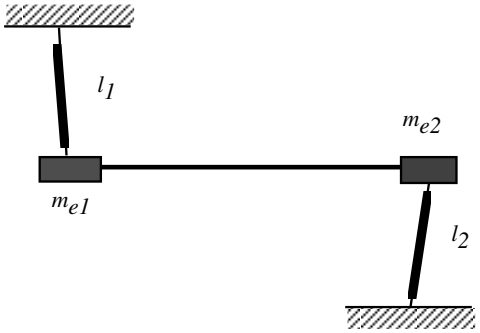


Figure 51. Simplified diagram of a folded pendulum.

A ULF-stage pre-isolator efficiently and passively reduces the normal mode amplitude as shown in figure 49. All the normal modes are on the cut-off slope of the ULF stage and are greatly reduced, leaving one very-low-frequency mode which is easy to control. This greatly simplifies the control system of the interferometer.

There are two main approaches to achieving very low frequency. One is to use a negative spring or antispring. A simple version is the inverted pendulum [213,214] as shown in figure 50. The resonant frequency of an inverted pendulum is given by

$$\omega_0 = \sqrt{\frac{k}{m} - \frac{g}{l}}, \tag{5.9}$$

where k is the spring constant of the flexure elements, m is the mass and l is the length of the inverted pendulum. Gravity provides a negative spring constant. The device has a low frequency when $g/l \sim k/m$, and it becomes unstable for $g/l > k/m$. Inverted pendulums for gravitational wave detectors were first investigated in UWA [213]. Because a mechanical Hooke's law spring is being used to counter a gravitational spring, and because the temperature coefficient of length differs from the temperature coefficient of Young's modulus, such antispring devices generally have relatively large temperature sensitivity. Thus servo control is needed to maintain a stable operating position. A full-scale inverted pendulum pre-isolation stage about 6 m high has been built for VIRGO's supperattanuator [215], and a 1 m stage has been developed at UWA [216].

For horizontal isolation, more elegant devices minimize the contribution of mechanical springs. One such device is the folded pendulum [217–219]. It combines a positive and negative pendulum as shown in figure 51. The resonant frequency of the folded pendulum is given as

$$\omega = \sqrt{\frac{1}{M_e}\left(\frac{m_{e1}g}{l_1} - \frac{m_{e2}g}{l_2}\right) + \gamma}, \tag{5.10}$$

where γ is a small additional term which takes into account elastic contributions from the flexures, $M_e = m_{e1} + m_{e2}$ is the equivalent mass of the pendulum and l_1, l_2 are the lengths of the positive and negative arm respectively. This device has low temperature sensitivity and has achieved a resonant frequency of 15 mHz [220]. With a resonant frequency of 17 mHz, this device gives isolation of more than 90 dB at 10 Hz. Above 15 Hz, the attenuation is degraded due to internal resonances of the isolator structure, as discussed above.

Another type of ULF stage is the X-pendulum [221, 222]. It uses two cross-wire linkage arrangements that mimic the motion of a very long pendulum. A resonant frequency of

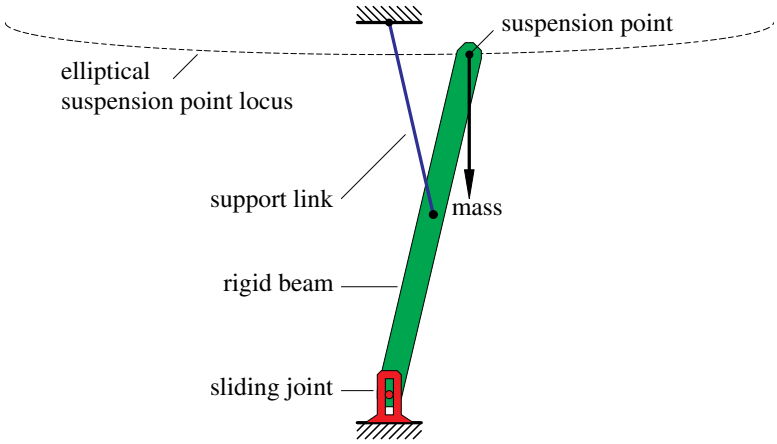


Figure 52. Scott–Russel linkage.

50 mHz has been demonstrated. With a resonant frequency of 0.21 Hz, the X-pendulum has demonstrated 30 dB isolation at 3 Hz. Above 3 Hz, the internal resonances dominate.

Both the folded pendulum and X-pendulum are most easily implemented as one-dimensional isolators. Although they can be cascaded to form a two-dimensional horizontal isolator, there are significant construction problems. Winterflood and Blair [223] have used the Scott–Russel linkage, shown schematically in figure 51. This device mimics the motion of a very long conical pendulum, achieving two-dimensional horizontal isolation in one single stage. A full-size prototype pre-isolator [224] has demonstrated a resonant frequency of 7 mHz and with a resonant frequency of 17 mHz, it achieves an isolation exceeding 75 dB at 0.5 Hz.

It is interesting to note that only one of the four horizontal stages discussed above is dependent on a carefully designed spring. This is the inverted pendulum, where a gravitational antispring balances on angular mechanical spring. In the other cases, mechanical springs are eliminated, except in so far as being an intrinsic, but small, component of a flexure suspension. All the linkage-based devices have the advantage that the temperature coefficients and nonlinearity of springs are minimized.

As mentioned above, cross-coupling requires a high level of isolation for both horizontal and vertical isolation. Thus there is little point in building an isolator with excellent horizontal pre-isolation unless it also has good vertical pre-isolation. Since the vertical isolation must always counteract a large gravitational force, it is almost impossible to avoid the use of large mechanical springs for vertical load bearing. However, several practical means of counteracting the spring constant have been demonstrated. The first is the magnetic antispring, demonstrated at Pisa [204, 225]. Pairs of magnets in a repulsive arrangement create a potential hill in the middle line of the magnet pairs. The negative spring constant represented by the potential hill partially cancels the positive spring constant, substantially reducing the total stiffness of the spring. Unfortunately, most magnets have high temperature coefficients, so that the magnetic antispring must be very carefully temperature controlled.

An alternation which avoids magnets is the geometric antispring. This was first demonstrated by LaCoste in his seismometer design [226], in which a zero length coil spring suspends a horizontal arm by acting on it at an angle (figure 53). The torsion crank linkage developed by Winterflood and Blair [224] uses a different geometrical antispring concept, as illustrated in figure 54. This design makes use of the nonlinearity produce by the torsion arm connected to a suspension link. It is arranged so that the effective spring constant

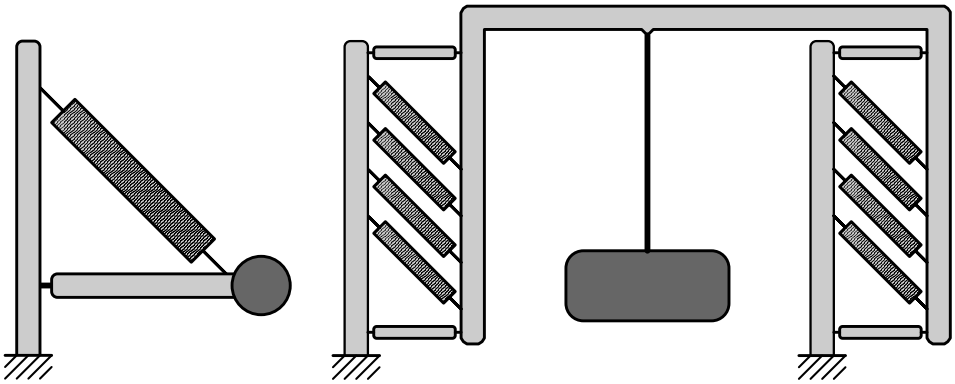


Figure 53. (*a*) LaCoste linkage. (*b*) Vertical preisolation using LaCoste linkage.

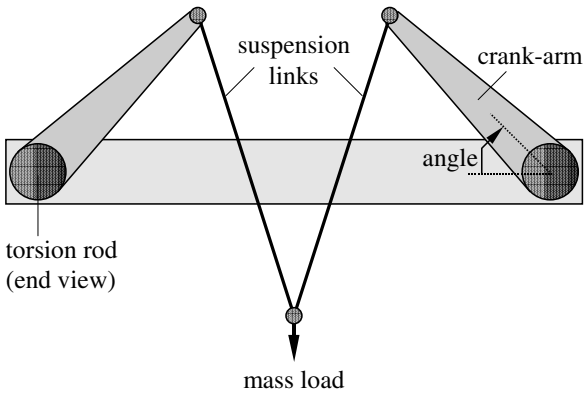


Figure 54. The torsion crank linkage arrangement [224].

$k = \partial_F/\partial_y (F = \text{force}, y = \text{vertical displacement})$ is almost zero and constant in a certain range. In a simple model, the torsion crank achieved a resonant frequency of 50 mHz. This vertical ULF stage can be combined with a Scott–Russel stage to create a three-dimensional pre-isolator as shown in figure 55.

De Salvo has demonstrated a geometric antispring based around a pair of cantilever blade springs [227]. The torsion rods of figure 54 are replaced by cantilever blades, and using appropriated angled suspension wires the same geometry effect causes nulling of the spring constant for a certain deflection angle.

Figure 56 shows the residual motion predicted for a passive isolator being developed at UWA [216]. This system utilizes passive eddy current damping, as well as seismic tilt control. The predicted performance is 10^{-9} m rms above 0.2 Hz. If such an isolator can be realized, interferometer operation will be greatly simplified.

5.5. Thermal noise

Once the seismic noise cut-off is lowered sufficiently through the use of high-performance vibration isolators, thermal noise will become the critical source of noise. From the fluctuation–

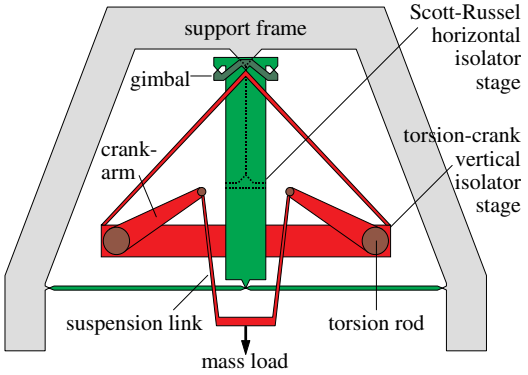


Figure 55. Three-dimensional pre-isolation stage, consisting of a torsion crank vertical stage and a Scott–Russel horizontal stage.

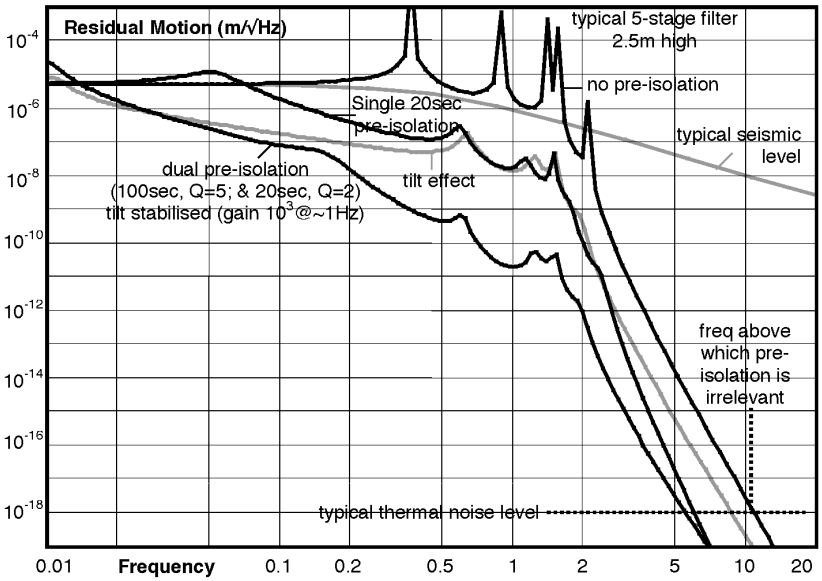


Figure 56. Successive reduction of residual motion. The bottom curve is the predicted overall isolation performance with pre-isolation stage, eddy current damping and tilt control.

dissipation theorem [90] the general power spectrum of the minimal fluctuation force is

$$F_{\mathrm{th}}^2 = 4k_{\mathrm{B}}TR(\omega), \tag{5.11}$$

where k_{B} is the Boltzmann constant and $R(\omega)$ is the real part of the impedance of the system. Using $Z = F/v$, the above equation can be expressed as

$$x_{\mathrm{th}}^2 = \frac{4k_{\mathrm{B}}T\sigma(\omega)}{\omega}. \tag{5.12}$$

$\sigma(\omega)$ is the real part of the admittance $Y(\omega) = 1/Z(\omega)$. For a simple harmonic motion system with spring constant k, mass m and a damping r, the thermal noise displacement power spectrum is [228]

$$x_{\mathrm{th}}^2 = \frac{4k_{\mathrm{B}}Tr}{(k - m\omega^2)^2 + r^2\omega^2}. \tag{5.13}$$

If a damping mechanism of structural damping is assumed, which can be described with a complex modulus of elasticity [229], $k = k_0[1+\mathrm{i}\varphi]$, $\varphi =$ constant (i.e. $r = k\varphi$), equation (5.18) becomes

$$x_{\mathrm{th}}^2 = \frac{4k_{\mathrm{B}}Tk\varphi}{\omega[(k - m\omega^2)^2 + k^2\varphi^2]} = \frac{4k_{\mathrm{B}}T\omega_0\varphi}{\omega m[(1 - \frac{\omega^2}{\omega_0^2})^2 + \varphi^2]}, \tag{5.14}$$

where $\omega_0 = \sqrt{\frac{k}{m}}$ is the angular resonant frequency.

In a laser interferometer gravitational wave detector, thermal noise is mainly divided into two classes—the suspension thermal noise and the internal thermal noise of the test mass.

5.5.1. Suspension thermal noise and Q-factor of the pendulum. Suspension thermal noise is mainly the Nyquist noise of the test mass pendulum of the interferometer. The Q-factor of a pendulum suspension can in principle be very high because the energy storage is predominantly in the effectively loss-less gravitational field. However, some elastic energy must always be stored in the flexure which supports the pendulum. The Q-factor of a pendulum is limited by the losses in this element. The Brownian motion noise amplitude of a simple pendulum at angular frequency ω, for any normal mode at frequency $\omega_p \ll \omega$ is given by

$$x_p^2(\omega) = \frac{4k_{\mathrm{B}}T\omega_{p^2}}{Q_p m\omega^5}. \tag{5.15}$$

The Q-factor of the pendulum is given by $Q_p = \gamma Q_0$, where $Q_0 = 1/\varphi$ is the intrinsic Q-factor of the flexure material, and γ is the enhancement factor, which depends on the geometry and material of the pendulum flexure [228, 230].

Applying equation (5.20) to a simple pendulum, it has been shown that the thermal noise in a simple pendulum scales as [228]

$$x_{\mathrm{th}}^2 \propto m^{1/2}. \tag{5.16}$$

It can be seen that the suspension thermal noise can be reduced by increasing the mass of the pendulum and the quality factor of the test mass pendulum stage. To obtain numerical estimates, consider a typical interferometer with parameters of $L = 3$ km, $T = 300$ K, $m = 30$ kg, $f_p = \omega_p/2\pi = 1$ Hz, and $Q_p = 10^9$. We then have

$$h = \frac{x}{L} \sim 10^{-21}\left(\frac{10\ \mathrm{Hz}}{f}\right)^{5/2} \Big/ \sqrt{\mathrm{Hz}}. \tag{5.17}$$

This shows that extremely-low-loss pendulums are essential. For example, to obtain the sensitivity goals of $10^{-23}/\sqrt{\mathrm{Hz}}$ at 10 Hz, a pendulum Q-factor of 10^{10} is required.

At current sensitivity levels, it is barely possible to measure directly the thermal noise floor for high Q-factor test masses over the frequency range of interest for gravitational wave detectors. The thermal noise has to be inferred from Q-factor measurement of some resonance at other frequencies. However, the Q-factor and frequency relation is model dependent, and the calculated thermal noise floor differs with different damping mechanism assumptions. Frequency-independent Q-factors over a large frequency range have been approximately confirmed in some materials [214, 231–234]. The mechanism of structural damping is now widely accepted. However, it still needs more investigation, particularly to confirm the phenomenon in low-loss single-crystal materials.

Various types of suspension have been studied to obtain a high pendulum Q-factor. Most researchers have assumed that wires are necessary. Suspensions with a thin wire double-loop simple pendulum configuration are widely used [199, 200, 235–237]. A pendulum Q-factor of

$Q \sim 10^8$ with wire suspension is expected [238], assuming reasonable Q-factors for the wires. Pendulum Q-factors exceed 10^7 had been observed with fused silica fibre suspension [239]. The highest pendulum Q-factor reported so far is $>3 \times 10^8$ by Braginsky *et al* [240] with a monolithic silica fibre and a 30 g mass.

The simple pendulum suspension system inevitably will have a set of middle-frequency violin string modes from the suspension wire. Those modes will contaminate the window of gravitational wave detection. The problem of the thermal noise of pendulums suspended by wires has been intensively studied [232–234, 241–243]. Studies have shown that high pendulum Q-factor is related to a high violin string mode Q-factor. So each violin string mode is confined within a very narrow frequency band.

Theoretical analysis has shown that significant improvements can be achieved if the pendulum is replaced by a compound pendulum supported by a thin membrane which acts as a hinge [230]. A pendulum with hinge suspension has been shown to achieve $Q \sim 10^7$ [244]. Using known high-Q materials such as niobium which has a Q-factor $\sim 10^5$ in thin membrane, and assuming that Q is independent of frequency, it can be shown that Q-factors exceeding 10^7 [245] can be expected. Also, since the test mass pendulum is suspended by a very short membrane, the violin string mode will be high enough to be neglected. The membrane flexure can be made much thinner than the wire flexure and thus has lower thermoelastic effect.

5.5.2. Test mass internal resonance thermal noise. The test mass will have many internal resonances. The total thermal noise thermal noise spectral density can be expressed as [228]

$$x_{\mathrm{th}}^2 = \frac{4k_{\mathrm{B}}T}{\omega} \sum_i \frac{\omega_i^2 \varphi_i}{m_i[(\omega_i^2 - \omega^2)^2 + \omega_i^4 \varphi_i^2]}. \tag{5.18}$$

Here i is the index of the ith mode, m_i is the effective of ith mode and $\varphi_i = 1/Q_i$, where Q_i is the quality factor of the ith mode. Typical internal resonances of a test mass are at about several kilohertz. The thermal noise spectral density far below the internal resonant frequency $\omega \ll \omega_i$ (assuming loss factor $\varphi_i = \varphi =$ constant) is given by

$$x_{\mathrm{th}}^2 \frac{4k_{\mathrm{B}}T\varphi}{\omega} \sum_i \frac{1}{m_i \omega_i^2}. \tag{5.19}$$

Although it can be seen from the above equation that the lowest internal resonance contributes most to the thermal noise, detailed theoretical studies [246, 247] has shown that higher-order modes cannot be neglected, especially when the laser beam is not perfectly aligned.

From equation (5.19) it can be seen that to reduce internal thermal noise it is required that the test masses have high internal resonances and very high Q-factors $(1/\varphi)$. Since ω_i is proportional to the velocity of sound, high sound velocity materials are required. It is worthwhile pointing out that the dimensions of the test mass should not be too large so as to keep the frequency of internal resonances as high as possible. This is contradictory to the requirement of using a big test mass to reduce the pendulum thermal noise. There is a compromise in choosing the size and mass of the test mass. High Q-factor materials (such as quartz and silicon) have been investigated in prototype laser interferometer detectors [248]. At present, fused silica test masses are widely used in prototype interferometer detectors. The Q-factor of silica test masses were observed to be of the order of 10^6 [249, 250] with a higher value of $>10^7$ in some resonant modes recently observed by Beilby and Startin [251]. The highest Q-factor of silica was reported by Braginsky of 6×10^7 [252].

Using another high Q-factor material, sapphire, as test mass has been proposed [230]. The excellent thermal and mechanical properties of sapphire makes it a promising material for use in test masses. Compared with silica material, the high thermal conductivity means the thermal

lensing problem can be minimized, while the very high Young's modulus means the internal resonance will be high. The highest reported sapphire Q-factor is 4×10^8 by Braginsky [112] and recent measurements reported a reproducible result of $2–3 \times 10^8$ Q-factors [239, 253]. Theoretical analysis [254] suggests that the thermal noise can improve by a factor of >10 by using sapphire test mass instead of silica test masses.

In practice, the suspension joint to the test mass plays an important role in obtaining both high-Q pendulum motion and high internal Q-resonance. Monolithic suspension systems have been suggested and investigated. These include bonding the silica fibre to the silica test mass using silicate bonding method [255], and bonding niobium flexure to sapphire test mass using active alloy bonding [256] and possibly silicate bonding [257].

5.6. Control systems

As mentioned above, the test masses in the interferometer must be suspended to isolate against seismic and environmental vibration. To achieve operation of an interferometer a feedback system must be used control the test mass position to high precision. Firstly, a local control system is necessary to suppress large-scale motions and align the mirrors to the point where the best interference contrast may be achieved. Secondly, to allow maximum sensitivity, a global control system is required to control the interferometer arm lengths to a relative motion of $\sim 10^{-12}$ m rms [183]. This is achieved by locking the interferometer to a dark fringe.

5.6.1. Local controls. Various damping methods to suppress low-frequency normal mode in vibration isolators have been investigated. Passive damping such as magnetic eddy current damping has been investigated [200, 211]. The problem with this type of damping is that the damping is usually relative to a support structure and can introduce noise into the isolators both through seismic noise and resonant peaks of the support structure. This degrades the high-frequency performance of the isolator.

An alternative method of magnetic damping uses the narrow-band resonant absorber. In this case, individual modes of an isolator can be damped using a tuned resonator which is itself magnetically damped.

The low-frequency normal modes can also be attenuated by active damping [162, 199, 258–261]. A correction force corresponding to the motion of the suspended masses (measured by an inertial or non-inertial sensor) is applied to the appropriate part of the suspension system. A non-inertial sensor such as shadow sensor [162] is widely used on prototype interferometers [162, 163]. It consists of a small vane mounted on the sensed surface, a light-emitting diode (LED) and an opposing photodiode (PD) mounted on the reference surface. The vane is free to move between the LED and PD and develops a signal proportional to the displacement of the vane by partially interrupting the light. The correction forces are applied by a small permanent magnet mounted integrally with the vane, and a coil mounted on the reference surface. The reference surface is usually the support frame, which is attached to the ground. Then the problem arises that the seismic noise can be injected into the servo loop both in sensing and in force actuation. If motion sensing is done with respect to the frame, it is impossible to avoid seismic noise in the sensing. This sensing noise has components in the signal band which must be prevented from driving the test mass and appearing in the signal output. This is readily achieved by electronic filtering of the signal [162, 260]. The noise injection through the coil can be overcome by carefully positioning the coil so that the magnetic field gradient is maximized at the magnet. This results in the best decoupling of the forces applied on the controlled masses [162]. Another way to overcome frame vibration is to use two coils to linearize the field and create a magnetic force which is independent of the position and motion

of the frame [260]. If the reference surface is a reaction mass [262] which is well isolated from the seismic noise, the noise injection will not be such a problem. However, such a sensor will not damp the common-mode motion of the test and reference masses. An inertial sensor such as a mass-loaded piezoelectric accelerometer [263] or a wideband accelerometer [264, 265] can in principle avoid the problem of sensing frame vibrations. Inertial sensing avoids any coupling with seismic noise. Its ultimate sensitivity is set by electronics noise, thermal noise and thermal drifts of the accelerometer and the feedback actuators.

Another kind of non-inertial sensor is the capacitance transducer [266]. A convenient implementation for use on dielectric test mass surfaces consists of two interleaving combs of parallel conductors etched onto a circuit board. The dielectric constant of the sensed surface contributes to the capacitance between the combs. A voltage applied to the capacitor will exert a force on the test mass, while motion of the test mass modulates the capacitance. Thus this system can provide both sensing and feedback forces. There is no need for lossy magnets or vanes to be attached to the test mass, so the test mass acoustic losses need not be degraded [249].

5.6.2. Global controls. It is necessary to extract error signals to control the interferometer lengths or mirror positions to lock the output at the dark fringe and the recycling cavity or arm cavities on resonant with the laser frequency. In order to avoid the noise due to the laser power fluctuations and $1/f$ electronics noise at low frequencies, the measurement has to be shifted to the quieter MHz domain using modulation–demodulation techniques. Various modulation configurations have been proposed and extensively studied. Schemes known as external [176, 267–269] and frontal [270–274] modulation have been particularly investigated because their modulators are outside the interferometer. Thus, as opposed to internal modulation, these schemes avoid introducing losses or wavefront distortion within the interferometer [275]. Although quantum-noise-limited sensitivity has been achieved with internal modulation at low laser power level [276–278], all the proposed large-scale laser interferometer detectors will not use this configuration because it introduces losses and wavefront distortion.

In the external modulation configuration (figure 36), a reference beam from the antireflective coated face of the beamsplitter is extracted, phase modulated and mixed with the main interference as in a Mach–Zehnder interferometer. This configuration needs additional optical components that, in order to avoid noise, must be suspended. The length of the reference Mach–Zehnder arm must be controlled to maximize the phase sensitivity. The modulation index m in the external arm is an independent variable and can be set to maximize $J_1(m)$ when m reaches its optimum value of 1.84 radians. A larger modulation index leads to more power being transferred to higher-order sidebands, which are normally not utilized by the demodulation process. Furthermore, in power-recycling and arm-cavity configurations, another modulation is needed to extract error signals to control the recycling mirror and arm cavities. Because the arm lengths of the interferometer can be equal and because the reference beam extracted from the back face of the beamsplitter has travelled almost the same path length as the main beam, this configuration is insensitive to the laser frequency noise.

In the frontal modulation configuration (figure 57) the interferometer has two arms with a small difference in length. The laser beam is phase modulated before entering the interferometer. The laser field at the interferometer input may be expressed as a superposition of three monochromatic plane waves if the modulation index is not too high: the carrier with the original laser frequency and two sidebands with the frequency shift of the modulation frequency f_m. In the simple Michelson, when the dark fringe condition is fulfilled for the carrier, the sidebands transmitted to the interferometer are maximum when arm length difference is a quarter of the modulation wavelength, $c/4f_m$.

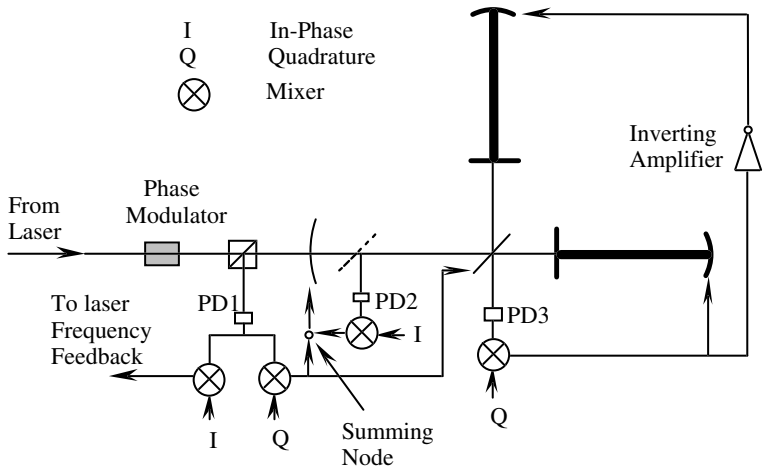


Figure 57. Global control configuration with frontal modulation [272].

If a recycling mirror is added, the modulation frequency and arm length difference must fulfil two conditions. First, the sidebands must resonant in the recycling cavity, otherwise the effective modulation index will be very small. Second, the signal must have maximum phase sensitivity. If R_r and R_m are respectively the intensity reflectivity of the recycling mirror and the interferometer, the modulation frequency f_m and the arm length difference Δl have to satisfy the following condition [273]:

$$\cos\left(2\pi f_m \frac{\Delta l}{c}\right) = \sqrt{R_r R_m}. \tag{5.20}$$

This condition also ensures the optimum enhancement of the modulation index inside the recycling cavity.

Even with cavities in two arms, one modulator is enough to get all the error signals to control each mirror. This configuration is simple and rather easy to realize. But because of the nonsymmetric configuration, the interferometer is sensitive to laser frequency and beam jitter noise. Frontal modulation was first demonstrated by successfully locking a tabletop prototype of a power-recycled Michelson interferometer with Fabry–Perot cavities in the arms [272].

5.6.3. Analogue and digital controls. Implementing the control discussed above can be achieved by either analogue or digital control systems. Analogue control is a well developed technology. Conventional PID servo control techniques can be used. Most prototype interferometers use analogue control for local damping, alignment and fringe locking [279]. The results have been successful, with sensitivity of 3×10^{-19} m Hz$^{-1/2}$ achieved [164] and generally not limited by servo system noise. However, for large-scale interferometers like LIGO and VIRGO, there are a large number of degrees of freedom (more than 200) [280] which need to be controlled and many of them are coupled each other. It is also planned that many error signals will be monitored and archived, to allow cross-correlation with the signal from the interferometer output PD. Many automated features are also required. All of the above points incline towards the use of digital control systems. Since Barone *et al* [281] introduced the idea of digital control into the automatic alignment of a Michelson interferometer, they have demonstrated theoretically and experimentally that all the specification on the noise requirements, the dynamic range and the control bandwidth can be satisfied using all-

digital control systems [282]. A fully digitally controlled interferometer prototype has been successfully operated in Naples, Italy for development and test of some VIRGO subsystems.

Heflin and Kawashima [283, 284] implemented both analogue and digital alignment control for the TENKO-100 DL interferometer. The results indicate that digital feedback control for application onto interferometer systems with target search frequency $\sim$1 kHz or less is a feasible alternative to analogue feedback systems. It has been shown to reproduce many of the same desirable features as the analogue systems.

Improved vibration isolation, especially that with very low residual motion, is likely to allow the undoubted complexity of the control problems discussed above to be substantially simplified.

5.7. Laser stabilization

Laser interferometers are designed to be sensitive to the optical phase difference of two arms, and should not be sensitive to common-mode fluctuations of the input light. But in practice, because of asymmetry between the two arms, fluctuations in the input light will couple into the output signal. In the frontal modulation scheme some asymmetry is unavoidable. In addition, differences in the optical components will cause, for example, intensity fluctuations to give rise to a differential radiation pressure force between both arms. Thus, laser intensity and frequency fluctuations must be strongly suppressed.

5.7.1. Laser pre-stabilization. The laser frequency noise Δf can couple into interferometer phase fluctuation $\Delta\varphi$ via arm length difference ΔL [162]:

$$\Delta\varphi \cong 2\pi\,\Delta f\,\Delta L/c. \tag{5.21}$$

The interference of the scattered light with the main beam can also couple the frequency noise into the output signal [182]. To ensure that frequency noise is sufficiently low that it does not compromise the sensitivity in the signal frequency regime, an active stabilization system is necessary to reduce the laser frequency noise.

An effective method of laser frequency stabilization was proposed by Drever and Hall *et al* [285], modelled on a technique used in microwave systems which was first proposed by Pound [286]. In optics the technique is called Pound–Drever–Hall (PDH) modulation. Light incident on a cavity is frequency modulated. The cavity creates intensity modulation whose phase depends on the relative frequency between the laser and the cavity resonance. The reflected light can be thought of as containing two beams: light simply reflected from the cavity input mirror and the light which has entered the cavity, resonated in the cavity and leaked back via the input mirror. The cavity leakage has a strong phase shift with respect to the directly reflected light from the input mirror, depending on the detuning from the cavity resonance. The interference between these two beams allows the detection of the phase difference, and hence the frequency difference, of the laser frequency compared with the cavity resonant frequency. The light from the laser is usually modulated at a RF frequency to shift the measurement to the quiet high-frequency domain. This helps overcome technical noise such as low-frequency electronics noise.

The schematic diagram in figure 58 shows a typical laser frequency stabilization scheme. A small fraction of laser light is phase modulated at a radio frequency by a Pockel cell (PC) and injected into a reference cavity. The reflected light is detected by a PD and mixed with the RF reference signal. If the laser frequency is tuned to one of the cavity resonances the reflected light has two balanced sidebands with opposite phase and a carrier. The mixer output is zero. If the laser frequency fluctuates the two sidebands will be unbalanced and the mixer will give

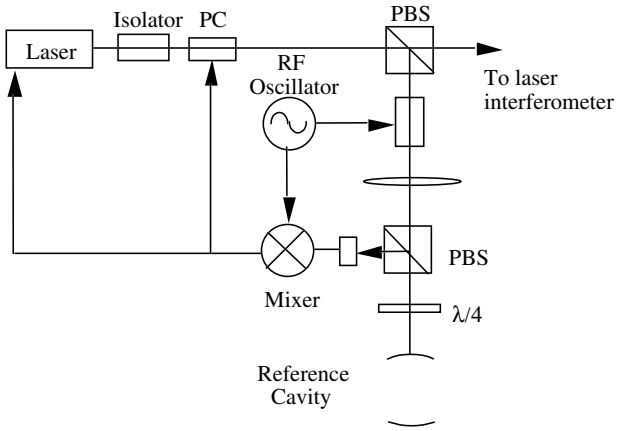


Figure 58. Frequency stabilization.

a signal proportional to the laser frequency difference with respect to the cavity resonance frequency. This signal can be fed back to a piezomirror (for low-frequency correction) and an extra-cavity Pockel cell (for high-frequency correction) to lock the laser frequency to the cavity.

The performance achievable in PDH stabilization is theoretically set by the quantum limit: the balance of shot noise and radiation pressure fluctuations which deform the reference cavity. However, in practice the laser shot noise limit is always dominant. This sets a limit to the frequency noise spectral density Ssn, and if the cavity has no losses this is given by [287]

$$S_{sn}(\mathrm{Hz}/\sqrt{\mathrm{Hz}}) = \frac{\Delta\nu}{J_0(\beta)}\sqrt{\frac{h\nu}{8\eta P_i}}. \tag{5.22}$$

Here $\Delta\nu$ is the cavity linewidth, P_i is the power incident on the cavity, ν is the laser frequency, η is the quantum efficiency of the PD, β is the modulation index, and $J_0(\beta)$ is the zeroth Bessel function.

Every gravitational wave detection group is involved in the development of laser stabilization. A frequency noise on the order of $\sim$10 mHz $\mathrm{Hz}^{-\frac{1}{2}}$ at 1 kHz has been achieved with a diode pumped Nd:YAG laser actively stabilized to a rigid reference cavity [288–291].

Another noise source in PDH locking is the mechanical and thermal noise of the reference cavity. The cavity resonant-frequency fluctuation $\Delta\nu$ is directly linked to the cavity length fluctuations ΔL. That is,

$$\frac{\Delta\nu}{\nu} = \frac{\Delta L}{L}. \tag{5.23}$$

If the reference cavity length is very long the noise contribution from a given cavity length fluctuation is much less than that of a short cavity. For this reason, the laser frequency is often stabilized to one arm cavity [163, 289] of an interferometer for further stabilization of the light frequency, and a frequency noise on the order of $\sim$10 μ Hz $\mathrm{Hz}^{-\frac{1}{2}}$ at 1 kHz has been achieved. In recent designs special-purpose *mode cleaner* cavities have been used [292] (see below) as a reference cavity. The Max Planck group has stabilized to an interferometer's power recycling cavity [173].

5.7.2. Mode cleaners. It is inevitable that the laser beam has geometric fluctuations (beam jitter) because of the vibration of the laser cavity. Vibration of the injection optics may also

introduce the beam jitter. The beam jitter noise will couple into the interferometer output signal if the two arms are not perfectly symmetric. For example, if the beamsplitter is misaligned from the optimal symmetric orientation by a small angle θ the lateral movement Δx of incident beam may produce an effective differential displacement signal Δs given by [293]

$$\Delta s = 4\theta \Delta x. \tag{5.24}$$

A simple way to clean the beam jitter noise of a low-power laser beam is to pass it through a single mode fibre [162]. In a single-mode fibre geometric fluctuations are transformed into fluctuations of the power coupling into the fibre. The problem in that because the core of a single-mode fibre has a small radius (of the order of a few wavelengths) the intensity inside the core is very high and this generates nonlinear effects including stimulated Brillouin scattering [294, 295]. This sets an upper limit on the maximum power that can be transported in a single-mode fibre. In addition, fibres themselves can be subject to vibration which can introduce additional beam jitter.

Another method to suppress beam jitter is to use a long optical cavity in transmission, called a mode cleaner. The geometric fluctuations of the laser beam are suppressed because they are not resonant within the cavity. The geometric beam fluctuations can be described in terms of higher transverse modes of the resonant cavity. In rectangular coordinates Hermite–Gaussian functions can be used to describe the eigenmodes of a cavity to a good approximation. Since the Hermite–Gaussian modes form a complete set, an incident beam with small translational movement Δx, angular fluctuation $\Delta\theta$, beam waist size mismatch $\Delta\omega_0$, and beam waist position mismatch Δb can be expanded in terms of cavity eigenmodes [292] as follows:

$$E_{\rm in} = E_{00}\left(\frac{\Delta x}{\omega_0} + j\frac{k\omega_0}{2}\Delta\theta\right)E_{10} + \frac{1}{\sqrt{2}}\left(\frac{\Delta\omega_0}{\omega_0} + j\frac{\Delta b}{k\omega^2}\right)(E_{02} + E_{20}). \tag{5.25}$$

Here E_{ij} are the amplitudes of the fundamental, first and second eigenmodes, while $E_{\rm in}$ is the amplitude of the incident beam and ω_0 is the beam waist.

For a cavity consisting of two mirrors with radii of curvature R_{c1}, R_{c2}, intensity transmission T_1, T_2, and reflectivity R_1, R_2, separated by a distance L, the fractional transmission of the incident light amplitude through this cavity is [296]

$$T_{mn} = \frac{\sqrt{T_1T_2}}{1-\sqrt{R_1R_2}}\frac{1}{\sqrt{1+(\frac{2\sqrt{R_1R_2}}{1-R_1R_2}\sin((m+n)\Psi))^2}}, \tag{5.26}$$

where $\Psi = \cos^{-1}(\sqrt{(1-L/R_{c1}(1-L/R_{c2})}$.

For the fundamental mode, $T_{00} = \frac{\sqrt{T_1T_2}}{1-\sqrt{R_1R_2}}$. The amplitude attenuation of the higher-order modes compared with the fundamental mode is given by

$$\sqrt{1+\left(\frac{2\sqrt{R_1R_2}}{1-R_1R_2}\sin((m+n)\Psi)\right)^2} \approx \frac{2F}{\pi}\sin((m+n)\Psi), \tag{5.27}$$

where $F = \frac{\sqrt{R_1R_2}}{1-R_1R_2}$ is the finesse of the cavity. Thus a high-finesse cavity mode cleaner can strongly suppress the higher-order modes.

However, the power inside a high-finesse mode cleaner cavity will also be much higher than the incident light power. Thus, thermal damage to the mirror coating becomes a key issue in defining the cavity configuration. If P is the power we want to transmit, and ω is the beam radius on the mirrors, then the power density on mirrors is given by $I = \frac{2FP}{\pi^2\omega^2}$. To avoid optical damage to the coatings of the mirrors, the spot area should be kept above $\pi\omega_{\rm min}^2 = \frac{2FP}{\pi I_{\rm max}}$, where $I_{\rm max}$ is the power density limit of the coatings.

The power transmission of an optical cavity for incident light with frequency offset f from the resonant frequency is given by the well known Lorentzian response: $T_f = \frac{T_0}{1+(2f/\Delta\nu)^2}$, where $\Delta\nu = c/2LF$ is the cavity linewidth. The cavity acts as a second-order low-pass filter. With a large length L and high finesse F the cut-off frequency becomes low. For example, for $L =$ 100 m and $F = 1000$ the cut-off frequency is 1.5 kHz. Since all the laser fluctuations (amplitude and frequency) can be understood in terms of the generation of sidebands, a mode cleaner also acts to suppress such fluctuations at offset frequencies higher than the cut-off frequency.

5.8. Optics

As we have already seen, a fundamental limitation of the sensitivity of interferometric gravitational wave detectors is shot noise, or photon-counting errors. In order to reduce shot noise, high light power must circulate in the interferometer. Because of the limitation of laser power currently available from stabilized CW lasers, power recycling appears to be an indispensable technique for large-scale instruments.

The maximum power recycling factor is determined by the total losses of the interferometer when the transmission of the recycling mirror is properly chosen. The losses of the interferometer may result from energy loss due to absorption or scattering of the mirrors, or by imperfect recombination of the two beams on the beamsplitter due to misalignments and wavefront distortions. Power losses can arise from imperfect mirror surfaces, coating inhomogeneities, diffraction losses through limited apertures, beamsplitter and mirror substrate wavefront distortion and depolarization. The minimization of these losses presents a formidable challenge.

5.8.1. Surface quality. The relevant specifications of the surface figure for optical components has been considered extensively by all of the groups building large-scale laser interferometers [196, 297–299]: see, for example, the discussion by Winkler and coworkers [196], which is based on a requirement for a power build-up in a power-recycled DL interferometer. The surface deformations can be characterized by their amplitude s and spatial wavelength λ_s. Surface deformation with λ_s smaller than the beam diameter (micro-roughness) cause the scattering loss. The relative power loss by scattering due to the micro-roughness is given by

$$\frac{\Delta P}{P} = \left(4\pi \frac{s_{\rm rms}}{\lambda_2}\right)^2, \tag{5.28}$$

in which $s_{\rm rms}$ is the rms value of the micro-roughness amplitude. The tolerable micro-roughness in a DL interferometer with power recycling gain of 100 and 34 reflections is $s_{\rm rms} < \lambda/730$.

Surface deformations with λ_s in the order of the beam diameter contribute to the beam wavefront distortion and the deterioration of the dark fringe of a perfect interference. The tolerable deformation in this scale for the same interferometer is $s_{\rm rms} < \lambda/230$. For surface deformation with λ_s larger than the beam diameter (aberration) the demands are slightly less.

Using the computer-mode-based code of Vinet and Hello [297], LIGO group derived the requirements [299] of the optical components of the 4 km Fabry–Perot Michelson interferometer. The tolerable rms amplitude of the surface micro-roughness is 0.4 nm. The surface figure error (spatial wavelength larger than beam diameter) should be less than 0.8 nm.

Techniques of superpolishing are now well established and mirrors exceeding the tight specifications required have been developed [299–301]. LIGO optics is being polished by General Optics (GO) and Commonwealth Scientific and Industrial Research Organization (CSIRO). Metrology indicated that the polished surfaces with rms deviation (after removeing focus and astigmatism) <1 nm over 20 cm were produced. Surface roughness

measurements showed a micro-roughness of ∼3 Å for CSIRO substrates and ∼0.9 Å for LIGO substrates [299].

The mirror coating is also critical for scattering level and wavefront preservation in the interferometer. VIGO group has reached ∼1 ppm scattering level on 80 mm diameter coated mirror surface. The peak-to-valley coated surface deformation is ∼14 nm on 70 mm diameter [302].

5.8.2. Absorption. Optical absorption may take place at dielectric coatings, or inside the substrate material of transmitting optical components, such as the beamsplitter and the coupling mirror of Fabry–Perot cavities. Apart from losses introduced by the absorption itself, the absorbed light power will heat the local area, and consequently deform the optical components causing wavefront distortion. This arises due to the limited thermal conductivity of the substrate. The local heating forms a temperature gradient inside the substrate and consequently introduces radius of curvature changes on surfaces. Winkler *et al* [196] has analysed the problem and shown that the change δs of the sagitta by the absorption of the reflected beam is given by

$$\delta s = \frac{\alpha}{4\pi\kappa} P_a, \tag{5.29}$$

in which P_a is the absorbed light power, α is the thermal expansion coefficient, κ is the heat conductivity of the substrate material. The crucial quantity for the magnitude of the effect is the ratio α/κ of thermal expansion to heat conductivity. The α/κ of fused silica and sapphire are 33 and 28 respectively [196].

When the beam is transmitted through a material with a temperature gradient, because of a temperature dependence of the refraction index, the refraction indices of the beam axis and the outer parts of the beam are different and thermal lensing may result. The path difference δl between beam axis and outer parts of the beam, introduced by thermal lensing, is approximately [196]

$$\delta l \approx \frac{\beta}{4\pi\kappa} P_a, \tag{5.30}$$

in which $\beta = \delta n/\delta T$ is the temperature dependence of the refraction index and P_a is the power absorption there. It is clear that one wants to keep the ratio β/κ small. The β/κ of fused silica and sapphire are 1000 and 60 [196]. As regards thermal lensing, sapphire is much better than fused silica as a substrate material.

5.8.3. Depolarization. The temperature gradient inside the substrate will also produce a stress distribution. The stresses generate strains in the substrate, which in turn produce refractive index variations (or birefringence) via the photoelastic effect. Stresses may also be introduced during the manufacturing process. The magnitude of the birefringence may be defined by the phase difference δ introduced between orthogonal polarizations. In general, the polarization of the input beam will not be parallel to one of the principal axes of the birefringent component, especially since the birefringence may vary locally.

Power loss occurs due to the depolarization of the original input beam polarization. The measured birefringence of a very homogeneous 10 cm thick Corning 7940 grade 0A fused silica plate is $\delta = 1.2^\circ \pm 0.2^\circ$ in the central area [303] which corresponds to a maximum loss of 10^{-4} per pass. The reported birefringence in sapphire is comparable, about 0.1° cm^{-1} [304]. The lowest birefringences in coatings obtainable today are between 2 to 10 μrad per reflection—but so far only for mirrors and beams with a size of a few cm and mm, respectively [197].

Hello and Vinet [305, 306] have analysed the thermal effects in massive mirrors heated by high-power laser beams. The complete numerical simulation of thermal effects in a GW interferometer [307] indicate that the first-generation detectors using silica components may require substrate absorption coefficients of $\sim$2 ppm cm^{-1} $\pm$ 15%, and coating absorption coefficient of $\sim$2 ppm $\pm$ 10%.These values are within present capabilities for mirror and coating technologies. The reported lowest absorption is a level of less than 1 ppm cm^{-1} for light of 1 μm wavelength in a fused silica sample [308] and a level of 3.1 ppm cm^{-1} in a sapphire sample [309]. The present coating absorption coefficient is $\sim$1 ppm [302]. Recycling factors of 300 [173] and 450 [174] have been achieved with power of $\sim$100 W built up inside the recycling cavity. This corresponds to $>$1 kW cm^{-2} power on the beamsplitter and mirrors. Expected limits due to thermal effects have still not been reached.

Winkler *et al* [196,197] and Strain *et al* [308] with their co-workers at MPI have evaluated the effects of thermal deformation, thermal lensing and thermally induced birefringence on a recycling interferometer. Their model assumes that Gaussian-profile laser beams heat the optical substrates through either uniform bulk absorption or uniform absorption in the coatings, and that the optics has an aperture much larger than the beam diameter. They concluded that the power limit set by thermal lensing is a problem only for advanced interferometers operating at higher power than those presently under development. They find that inherent and thermally induced birefringence will not be the dominant loss mechanism [197] assuming the lowest values for absorption and inherent birefringence reported. They propose that resonant sideband extraction is the best way of reducing the effects of thermal lensing to reach sensitivity appropriate to a 'second-generation' detector [308].

5.9. High-power lasers

The choice of the laser wavelength is an important effect on the design of a long-baseline interferometer. At short wavelengths, the beam diffracts less and thus the diameter of the interferometer mirrors and the vacuum tube can be reduced; thereby reducing the cost of the interferometer. Short wavelengths would also allow a better shot-noise-limited strain sensitivity which is $\propto\lambda^{1/2}$ [189]. Other important factors in the choice of laser wavelength are the losses (absorption and scattering) in the mirrors and beamsplitter, and the availability of suitably quiet and powerful lasers.

Contrasting with prototype interferometers which have used argon-ion lasers, all the long-baseline interferometers will use diode-laser-pumped Nd:YAG lasers ($\lambda = 1064$ nm). The use of Nd:YAG lasers is driven by their much better efficiency and generally quieter characteristics with regard to practically all types of laser noise [293]. Diode-laser-pumped miniature ring lasers can fulfil the requirements of a interferometric gravitational wave detector concerning amplitude and frequency stability [310]. But direct use of these devices in an interferometer for gravitational wave detector is not possible, because the output power of these system is limited to values below 2 W CW in a single axial mode [311].

One possible technique to increase the output power is injection locking which coherently couples a low-power master and a high-power slave oscillator resulting in a high-power output with the frequency characteristics of the master [312]. As shown in figure 59, this is accomplished by injecting the output power from the master laser into the slave laser's resonator. The PDH reflection locking is used to lock the frequency of the slave laser to the master's by adjusting the position of the slave's mirrors according to the error signal [311].

Application of the injection locking technique to Nd:YAG lasers has been extensively investigated [311,313–317]. At Laser Zentrum Hannover (LZH), a maximum single-frequency output power of 20 W has been generated by injection locking to a monolithic ring laser. The

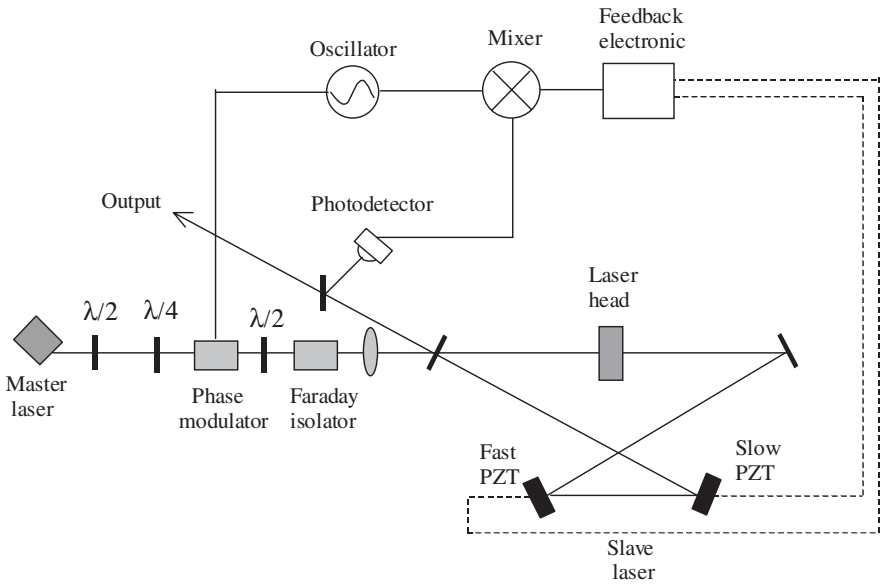


Figure 59. Schematic for injection locking of two lasers [311].

amplitude and frequency stability is investigated. The amplitude noise reaches the shot noise level beyond a few MHz. The relaxation oscillation of the miniature ring laser coupling at ∼700 kHz and the disturbance in the excitation of the slave laser caused by the diode laser power supplies at 80 kHz are two noise sources which could be further stabilized using active amplitude stabilization [318]. Using fibre-coupled diode lasers as pump sources, 62 W CW TEM_{00} mode output has been achieved [319].

At Stanford, an output of 5.5 W single-frequency, 'nearly diffraction-limited', TEM_{00} power was produced by using 50.4 W of pump power. The frequency noise of the unstabilized master laser of ∼20 Hz $Hz^{-\frac{1}{2}}$ at 1 kHz was reproduced at the output of the slave. The relative intensity noise at the output of the slave is $1.7 \times 10^{-6}/\sqrt{\text{Hz}}$ which is 10 times higher that that at the output of the master laser [314]. A 40 W CW, TEM_{00} diode-laser-pumped, Nd:YAG miniature-slab laser has been built and demonstrated with 212 W pumping power [320]. A 10 W laser-diode-pumped Nd:YAG master-oscillator power amplifier is spatially and temporally filtered by a fixed Fabry–Perot cavity, which produced a 7.6 W TEM_{00} beam with 1% higher-order transverse mode content and reduced the relative power fluctuations at 10 MHz to $2.8 \times 10^{-9}/\sqrt{\text{Hz}}$.

At VIRGO, a 10 W laser-diode-pumped Nd:YAG laser has been developed by using injection locking a high-power slave laser to a low power master laster [302]. The master laser is a 700 mW laser-diode-pumped miniature ring Nd:YAG laser operating at single-frequency. The slave laser is Nd:YAG laser transverse pumped on one side by 10 fibre-coupled diodes. TEM_{00} operation has been achieved with a slight contribution of TEM_{01} using a diaphragm inside the X-shaped ring cavity. Using a spatial filter, the TEM_{00} component can be extracted to give a 9 W output power for an effective pumped power of 60 W. The 10 W laser will be frequency pre-stabilized to a reference cavity, and be actively power stabilized.

At TAMA, with 22.3 W pump input from two fibre-coupled laser diodes, combined with 700 mW of power injected by a single-frequency master laser, the injection-locked slave laser emitted 10 W of linearly polarized TEM_{00} beam. The measured relative intensity

noise and the frequency noise are $2 \times 10^{-5}/\sqrt{\text{Hz}}$ and 50 Hz $\text{Hz}^{-\frac{1}{2}}$ respectively. When the frequency of the injection-locked laser is stabilized to an external high-finesse reference cavity, a minimum frequency noise of 40 mHz $\text{Hz}^{-\frac{1}{2}}$ was measured from the locking loop error signal at 1 kHz [313].

At Adelaide, an efficient, medium power, diode-pumped Nd:YAG slab, stable resonator, ring laser based on a new diode-pumping geometry [321] was developed. Using 18 W of absorbed laser diode power (20 W diode output power), 5.8 W TEM_{00} output beam has been produced [189].

6. Conclusion

Gravitational wave researchers have expected to detect gravitational waves 'within the next decade' for the last three decades. Detectors have been dramatically improved and a steadily increasing band of physicists has been able to devote more and more resources to the problem. In the process they have uncovered new physics and new technology. Gravitational wave detectors are the most sensitive devices ever invented.

Like the Great South Land which was rumoured for centuries before it was discovered, the spectrum of gravitational waves is a rumoured continent, first to be detected, and then to be explored. It seems not unreasonable that the exploration will begin within the next decade, but whatever happens the search will continue to motivate physicists and drive a continuing process of innovation.

Acknowledgments

The authors wish to acknowledge the Australian Research Council, which has supported gravitational wave research in Australia for many years, and our colleagues in Perth and around the world with whom we have had rich and rewarding interactions. The field is too large to be able to reference all important papers, and we apologize to those whose work we have failed to reference.

References

[1] Weisberg J M and Taylor J H 1984 Observations of post-Newtonian timing effects in the binary pulsar PSR 1913 + 16 *Phys. Rev. Lett.* **52** 1348
[2] Hulse R A and Taylor J H 1975 Discovery of a pulsar in a binary system *Astrophys. J.* **195** L51
[3] Einstein A 1916 *Preuss. Akad. Wiss. Berlin, Sitzungsberichte der Physikalisch-mathematischen Klasse* p 668
[4] Weber J 1960 Detection and generation of gravitational waves *Phys. Rev.* **117** 306
[5] Weber J 1969 Evidence for discovery of gravitational radiation *Phys. Rev. Lett.* **22** 1320
[6] Weber J 1970 Anisotropy and polarization in the gravitational-radiation experiments *Phys. Rev. Lett.* **25** 180
[7] Misner C W, Thorne K S and Wheeler J A 1973 *Gravitation* (San Francisco: Freeman)
[8] Thorne K S 1987 Gravitational radiation *300 Years of Gravitation* ed S Hawking and W Israel (Cambridge: Cambridge University Press)
[9] Saulson P R 1994 *Fundamentals of Interferometric Gravitational Wave Detectors* (Singapore: World Scientific)
[10] Weinberg S 1972 *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity* (New York: Wiley)
[11] Giazotto A 1989 Interferometer detection of gravitational waves *Phys. Rep.* **182** 365
[12] Damour T 1986 An introduction to the theory of gravitational radiation *Gravitation in Astrophysics* ed B Carter and J B Hartle (New York: Plenum)
[13] Damour T 1987 *Gravitation in Astrophysics* ed B Carter and J Hartle (New York: Plenum)
[14] Einstein A 1918 *Preuss. Akad. Wiss. Berlin, Sitzungsberichte der Physikalisch-mathematischen Klasse* p 154
[15] Suzuki T, Tsubono K, Kuroda K and Hirakawa H 1981 Calibration of gravitational radiation antenna by dynamic Newton field *Japan. J. Appl. Phys.* **20** L498

[16] Astone P *et al* 1991 Evaluation and preliminary measurement of the interaction of a dynamical gravitational near field with a cryogenic gravitational wave antenna *Z. Phys.* C **50** 21
[17] Dhurandhar S V, Blair D G and Costa M E 1996 Possibility of detection gravitational waves from millisecond pulsars by resonant bar antennas *Astron. Astrophys.* **311** 1043
[18] Thorne K S 1992 Sources of gravitational waves and prospects for their detection *Recent Advances in General Relativity* ed A Janis and J Porter (Boston: Birkhauser)
[19] Nakamura T 1995 *Proc. 7th Marcel Grossmann Meeting on General Relativity* ed R T Jantzen and G M Keiser (Singapore: World Scientific)
[20] Marck J-A 1995 *Proc. 7th Marcel Grossmann Meeting on General Relativity* ed R T Jantzen and G M Keiser (Singapore: World Scientific)
[21] Laguna P 1997 *Gravitational Waves: Sources and Detectors* ed I Ciufolini and F Fidecaro (Singapore: World Scientific) p 61
[22] Lai D and Shapiro S L 1995 Hydrodynamics of coalescing binary neutron stars: ellipsoidal treatment *Astrophys. J.* **443** 705
[23] Chandrasekhar S 1969 *Ellipsoidal Figures of Equilibrium* (New Haven, CT: Yale University Press)
[24] Houser J L, Centrella J M and Smith S C 1994 Gravitational radiation from nonaxisymmetric instability in a rotating star *Phys. Rev. Lett.* **72** 1314
[25] Hawking S W 1971 Gravitational radiation from colliding black holes *Phys. Rev. Lett.* **26** 1344
[26] Press W H and Thorne K S 1972 *Annu. Rev. Astron. Astrophys.* **10** 338
[27] Lai D and Shapiro S L 1995 Gravitational radiation from rapidly rotating nascent neutron stars *Astrophys. J.* **442** 259
[28] Phinney E S 1991 The rate of neutron star binary mergers in the universe: minimal predictions for gravity wave detectors *Astrophys. J.* **380** L17
[29] Narayan R, Piran T and Shemi A 1991 Neutron star and black hole binaries in the galaxy *Astrophys. J.* **379** L17
[30] Tutukov A V and Yungelson L R 1993 The merger rate of neutron star and black hole binaries *Mon. Not. R. Astron. Soc.* **260** 675
[31] Yamaoka H, Shigeyama T and Nomoto Z 1993 Formation of double neutron star systems and asymmetric supernova explosions *Astron. Astrophys.* **267** 433
[32] Abramovici A *et al* 1992 LIGO: The laser interferometer gravitational-wave observatory *Science* **256** 325
[33] Bradaschia C *et al* 1992 VIRGO: Very wide band interferometric gravitational wave antenna *Nucl. Phys.* B **28A** 54
[34] Johnson W W and Merkowitz S M 1993 Truncated icosahedral gravitational wave antenna *Phys. Rev. Lett.* **70** 2367
[35] Blair D G 1983 Resonant bar detectors for gravitational waves *Gravitational Radiation* ed N Deruelle and T Piran (Amsterdam: North-Holland)
[36] Rees M J 1993 *Proc. Natl Acad. Sci. USA* **90** 4840
[37] Rees M J 1997 Gravitational waves from galactic centres *Class. Quantum Grav.* **14** 1411
[38] Haehnelt M G 1994 Low-frequency gravitational waves from supermassive black holes *Mon. Not. R. Astron. Soc.* **269** 199
[39] Vecchio A 1997 Galaxy mergers and implications for massive black hole binary coalescence *Class. Quantum Grav.* **14** 1431
[40] Sigurdsson S 1997 Estimating the detectable rate of capture of stellar mass black holes by massive central black holes in normal galaxies *Class. Quantum Grav.* **14** 1425
[41] Rucinski S M 1994 M/sub V/=M/sub V/(log P, log T/sub e/) calibrations for W Ursae Majoris systems *Publ. Astron. Soc. Pac.* **106** 462
[42] Verbunt F 1997 Waiting for LISA: binaries with orbital periods less than $\sim 10^4$ s *Class. Quantum Grav.* **14** 1417
[43] Bildsten L 1998 Gravitational radiation and rotation of accreting neutron stars *Astrophys. J. Lett.* **501** L89
[44] Zwart S F P and McMillan S L W 1999 Black hole mergers in the universe *Astrophys. J. Lett. preprint* 9910061
[45] Ostriker J P and Cowie L L 1981 Galaxy formation in an intergalactic medium dominated by explosions *Astrophys. J. Lett.* **243** L127
[46] Cowie L L, Hu E M and Songaila A 1995 Detection of massive forming galaxies at redshifts $z > 1$ *Nature* **377** 603
[47] Hughes D H, Serjeant S, Dunlop J *et al* 1998 High-redshift star formation in the Hubble deep field revealed by a submillimetre-wavelength survey *Nature* **394** 241
[48] Blair D G and Ju L 1996 A cosmological background of gravitational waves produced by supernovae in the early universe *Mon. Not. R. Astron. Soc.* **283** 648

[49] Burman R R, Woodings S J and Blair D G 1998 The cosmic background of supernova-generated gravitational radiation *Publ. Astron. Soc. Australia*
[50] Ferrari V, Matarrese S and Schneider R 1999 Gravitational wave background from a cosmological population of core-collapse supernovae *Mon. Not. R. Astron. Soc.* **303** 247
[51] Ferrari V, Matarrese S and Schneider R 1999 Stochastic background of gravitational waves generated by a cosmological population of young, rapidly rotating neutron stars *Mon. Not. R. Astron. Soc.* **303** 258
[52] Flanagan E E 1993 Sensitivity of the laser interferometer gravitational wave observatory to a stochastic background, and its dependence on the detector orientations *Phys. Rev.* D **48** 2389
[53] Allen B 1996 *Proc. Les Houches Astronomical Sources of Gravitational Waves* ed J-A Marck and J-P Lasoto (Cambridge: Cambridge University Press)
[54] Blair D G 1999 Stochastic background detection with LLO and ALLEGRO *LIGO Report* U990003-00-0
[55] Sazhin M V 1988 *Experimental Gravitational Physics* ed P F Michelson (Singapore: World Scientific) p 179
[56] Taylor J H and Weisberg J M 1989 Further experimental tests of relativistic gravity using the binary pulsar PSR 1913+16 *Astrophys. J.* **345** 434
[57] Weber J 1967 Gravitational radiation *Phys. Rev. Lett.* **18** 498
[58] Forward R L 1971 Multidirectional, multipolarization antenna for scalar and tensor gravitational radiation *Gen. Rel. Grav.* **2** 149
[59] Wagoner R V and Paik H J 1977 *Gravitasione Sperimentale* (Rome: Academia Nazionale dei Lincei)
[60] Richard J-P 1982 A multimode antenna for the detection of short pulses of gravitational radiation *Proc. 2nd Marcel Grossmann Meeting* ed R Ruffini (Amsterdam: North-Holland)
[61] Richard J-P 1984 Wide-band bar detectors of gravitational radiation *Phys. Rev. Lett.* **52** 165
[62] Solomonson N, Johnson W W and Hamilton W O 1992 Comparative performance of two-, three- and four-mode gravitational radiation detectors *Phys. Rev.* D **46** 2299
[63] Tobar M E 1995 Characterizing multi-mode resonant-mass gravitational wave detectors *J. Phys. D: Appl. Phys.* **28** 1729
[64] Gertsenshtein M E and Pustovoit V I 1962 On the detection of low frequency gravitational waves *Sov. Phys.–JETP* **16** 433
[65] Weber J 1960s Weber proposed interferometer detector idea but never published it; see Thorne K S 1978
[66] Weiss R 1972 *MIT Res. Lab. Electron. Q. Report* vol 105, p 54
[67] Iess L and Armstrong J W 1997 *Gravitational Waves: Sources and Detectors* ed I Ciufolini and F Fidecaro (Singapore: World Scientific) p 323
[68] Bertotti B and Giampieri G 1998 Solar coronal plasma in Doppler measurements *Sol. Phys.* **178** 85
[69] Folkner W M, Hechler F, Vincent M A and Bender P L 1997 LISA orbit and stability *Class. Quantum Grav.* **14** 1405
[70] Astone P *et al* 1997 The gravitational wave detector Nautilus operation at $T = 0.1$ K *Astropart. Phys.* **7** 231
[71] Geng Z K, Hamilton W O, Johnson W W, Mauceli E, Merkowitz S, Morse A and Solomonson N 1995 Operation of the Allegro detector at LSU *Proc. 1st Eduardo Amaldi Conf. on Gravitational Wave Experiments* ed E Coccia, G Pizzella and G Ronga (Singapore: World Scientific)
[72] Locke C R, Tobar M E, Ivanov E N and Blair D G 1998 Parametric interaction of the electric and acoustic fields in a sapphire monocrystal transducer with a microwave readout *J. Appl. Phys.* **84** 6523
[73] Coccia E *et al* 1997 Experimental study of spherical resonators at very low temperatures *Proc. Omnidirectional Gravitational Radiation Observatory* ed W G Velloso Jr, O D Aguiar and N S Magalhaes (Singapore: World Scientific)
[74] Frossati G *et al* 1997 Very low temperature measurements of quality factors of copper alloys for resonant gravitational wave antennae *Proc. Omnidirectional Gravitational Radiation Observatory* ed W G Velloso Jr, O D Aguiar and N S Magalhaes (Singapore: World Scientific)
[75] Velloso W F Jr *et al* 1997 On the use of the finite elements method to design the structures of mechanical isolation to resonant mass antennas *Proc. Omnidirectional Gravitational Radiation Observatory* ed W G Velloso Jr, O D Aguiar and N S Magalhaes (Singapore: World Scientific)
[76] Astone P, Bonifazi P, Pallottino G V and Pizzella G 1994 Wiener filters for gravitational-wave antennas: characteristics and application *Nuovo Cimento* C **17** 713
[77] Liu J F, Blair D G and Ju L 1998 Near shore ocean wave measurement using a very low frequency folded pendulum *Meas. Sci. Technol.* **9** 1772
[78] Astone P, Buttigilione C, Frasca S, Palottin G V and Pizzella G 1997 The fast matched filter for gravitational-wave analysis *Nuovo Cimento* C **20** 9
[79] Heng I S 1999 Long term performance of the cryogenic resonant-mass gravitational wave antenna, NIOBE *PhD Thesis* University of Western Australia
[80] Mohanty S D, Heng I S, Blair D G, Dhurandhar S V, Tobar M and Ivanov E 1998 A data analysis approach for

detecting gravitational waves from PSR 0437-4715 *Mon. Not. R. Astron. Soc.* **301** 469
[81] Heng I S, Blair D G, Ivanov E N and Tobar M E 1996 Report on two years of the UWA gravitational radiation antenna *Phys. Lett.* A **218** 190
[82] Johnston S 1996 *Pulsars: Problems and Progress (ASP Conf. Ser. 105)* ed S Johnston, M A Walker and M Bailes p 501
[83] Suzuki T 1993 Low resonant frequency resonant detectors *Int. Workshop of New Technology for Gravitational Astronomy (Perth, 1993)*
[84] Astone P 1998 On the detection of monochromatic and stochastic gravitational waves with resonant gravitational wave detectors *Proc. 2nd Eduardo Amaldi Conf. on Gravitational Waves* ed E Coccia, G Pizzella and G Veneziano (Singapore: World Scientific)
[85] Schutz B F 1995 Sources of gravitational radiation for detectors of the 21st century *Proc. 1st Eduardo Amaldi Conf. on Gravitational Wave Experiments* ed E Coccia, G Pizzella and G Ronga (Singapore: World Scientific)
[86] Astone P *et al* 1997 Contributions from the AIE to the GWDAW2 workshop, Orsay *Albert Einstein Institut Preprint* AEI-057
[87] Giazotto A 1996 *Int. Conf. on Gravitational Waves: Source and Detectors (Pisa, Italy)*
[88] Opfer J E, Boughn S P, Fairbank W M, McAshan M S, Paik H J and Taber R C 1974 *Proc. Conf. Low Temperature Physics-LT13* vol 4, ed K D Timmerhaus, W J O Sullivan and E F Hammel (New York: Plenum) p 59
[89] Hamilton W O, Pipes P B and Nayar P S 1974 *Proc. Conf. Low Temperature Physics-LT13* vol 4, ed K D Timmerhaus, W J O Sullivan and E F Hammel (New York: Plenum) p 555
[90] Callen H B and Green R F 1952 On a theorem of irreversitable thermodynamics *Phys. Rev.* **86** 702
[91] Gibbons G W and Hawking S W 1971 Theory of the detection of short bursts of gravitational radiation *Phys. Rev.* D **4** 2191
[92] Papoulis A 1977 *Signal Analysis* (New York: McGraw-Hill)
[93] Pallotino G V and Pizzella G 1991 *The Detection of Gravitational Waves* ed D G Blair (Cambridge: Cambridge University Press)
[94] Braginsky V B 1970 *Physical experiments with test bodies NASA Tech. Trans. F672NTIS* (Springfield: NASA)
[95] Giffard R P 1976 Ultimate sensitivity limit of a resonant gravitational wave antenna using a linear motion detector *Phys. Rev.* D **14** 2478
[96] Heffner H 1962 *Proc. IRE* **50** 1604
[97] Braginsky V B, Vorontsov Y I and Kalili F Y 1978 *JETP Lett.* **27** 276
[98] Caves C M, Thorne K S, Drever R W P, Sandberg V D and Zimmerman M 1980 On the measurement of a weak classical force coupled to a quantum-mechanical oscillator I. Issues of principle *Rev. Mod. Phys.* **52** 341
[99] Unruh W G 1978 Analysis of quantum-nondemolition measurement *Phys. Rev.* D **18** 1764
[100] Tobar M E, Ivanov E N, Oi D K L, Cuthbertson D B and Blair D G 1997 Sapphire test-masses for measuring the standard quantum limit and achieving quantum non-demolition *J. Appl. Phys.* B **64** 153
[101] Paik H J and Wagoner R V 1976 Calculation of the absorption cross section for a cylindrical gravitational-wave antenna *Phys. Rev.* D **13** 2694
[102] Blair D G 1980 *Gravitational Radiation, Collapsed Objects and Exact Solutions* vol 314, ed D Edward (Berlin: Springer) p 299
[103] Veitch P J 1991 *The Detection of Gravitational Waves* ed D G Blair (Cambridge: Cambridge University Press)
[104] Cuthbertson B D, Tobar M E, Ivanov E N and Blair D G 1996 Parametric back action effects in a high-Q cryogenic sapphire transducer *Rev. Sci. Instrum.* **67** 2435
[105] Douglass X H and Braginsky V B 1979 *General Relativity, An Einstein Centenary Survey* ed S W Hawking and W Israel (Cambridge: Cambridge University Press)
[106] Johnson W W and Bocko M 1981 Approaching the quantum 'limit' for force detection *Phys. Rev. Lett.* **47** 1184
[107] Bocko M and Johnson W W 1982 Surpassing the amplifier limit for force detection *Phys. Rev. Lett.* **48** 1371
[108] Bocko M and Johnson W W 1984 Phase-sensitive parametric motion transducer *Phys. Rev.* A **30** 2135
[109] Pang Y and Richard J P 1997 Development of an optical transducer *Proc. 1st Int. Workshop for an Omnidirectional Gravitational Radiation Observatory* ed W F Velloso Jr, O D Aguiar and N S Magalhaes (Singapore: World Scientific)
[110] Conti L, Prodi G A, Vitale S, Marin F, De Rosa M, Cerdonio, Taffarello L and Zendri J P 1999 An optical transduction chain for the AURIGA detector *Proc. 3rd Eduardo Amaldi Conf. on Gravitational Waves* ed S Meshkov (New York: AIP) at press
[111] Braginsky V B 1983 *Gravitational Radiation* ed N Deruelle and T Piran (Amsterdam: North-Holland)
[112] Braginsky V B, Mitrofanov V P and Panov V I 1985 *System with Small Dissipation* (Chicago, IL: University of Chicago Press)

[113] Paik H J 1976 Superconducting tunable-diaphragm transducer for sensitive acceleration measurements *J. Appl. Phys.* **47** 1168
[114] Astone P *et al* 1993 Long-term operation of the Rome explorer cryogenic gravitational wave detector *Phys. Rev.* D **47** 362
[115] Linthorne N P, Veitch P J, Blair D G, Turner P J, Tobar M E and Mann A G 1990 Niobium gravitational radiation antenna with superconducting parametric transducer *Physica* B **165** 9–10
[116] Allen Z *et al* 1999 Development of a multi-mode transducer using a double SQUID preamplifier for the ALLEGRO detector *Proc. 3rd Eduardo Amaldi Conf. on Gravitational Waves* ed S Meshkov (New York: AIP) at press
[117] Giazotto A 1987 *Proc. 7th Conf. on General Relativity and Gravitational Physics* ed U Bruzzo, R Cianci and E Massa (Singapore: World Scientific)
[118] Ju L, Blair D G, Peng H and van Kann F 1992 High dynamic range measurements of an all metal isolator using a sapphire transducer *Meas. Sci. Technol.* **3** 463
[119] Tsubono K 1992 Private communication
[120] Richard J-P 1976 Sensor and suspensions for a low-temperature gravitational wave antenna *Rev. Sci. Instrum.* **47** 423
[121] Aldcroft T L, Michelson P and Taber R C 1992 *Proc. 6th Marcel Grossmann Meeting* ed H Sato and T Nakamura (Singapore: World Scientific)
[122] Blair D G, Tuner P J, Berinson D, Ju L, van Kann F J, Linthorne N P, Mann A G, Peng H and Tobar M E 1992 *Proc. 6th Marcel Grossmann Meeting* ed H Sato and T Nakamura (Singapore: World Scientific)
[123] Blair D G 1982 *Proc. 2nd Marcel Grossmann Meeting* ed R Ruffine (Amsterdam: North-Holland)
[124] Michelson P F and Taber R C 1984 Can a resonant-mass gravitational-wave detector have wideband sensitivity? *Phys. Rev.* D **29** 2149
[125] Veitch P J 1986 *PhD Thesis* University of Western Australia
[126] Coccia E 1984 Nodal point supported gravitational wave antennas *Rev. Sci. Instrum.* **55** 1980
[127] Allen Z A *et al* 2000 First search of gravitational wave bursts with a network of detectors *Phys. Rev. Lett.* submitted
[128] Vitale S, Cerdonio M, Coccia E and Ortolan A 1997 Gravitational-wave stochastic background detection with resonant-mass detectors *Phys. Rev.* D **55** 1741
[129] Astone P *et al* 1997 *Proc. 1st Int. Workshop for an Omnidirectional Gravitational Radiation Observatory* ed W F Velloso Jr, O D Aguiar and N S Magalhaes (Singapore: World Scientific)
[130] Modestino G *et al* 1999 Detection of cosmic rays by NAUTILUS *Proc. 3rd Eduardo Amaldi Conf. on Gravitational Waves* ed S Meshkov (New York: AIP) at press
[131] Hamilton W O 1987 *Int. Symp. on Experimental Gravitational Physics* ed P F Michelson (Singapore: World Scientific)
[132] Frasca S 1980 Analogical device for a rough localization of gravitational-wave sources *Nuovo Cimento* C **3** 237
[133] Nitti G 1980 Choice of optimal orientation of one or two cylindrical gravitational antennas *Nuovo Cimento* C **3** 420
[134] Blair D G, Frasca S and Pizzella G 1988 Antenna pattern for four gravitational wave antennas *Nuovo Cimento* **11** 185
[135] Schutz B F and Tinto M 1987 Antenna patterns of interferometric detectors of gravitational waves I. Linearly polarized waves *Mon. Not. R. Astron. Soc.* **224** 131
[136] Astone P, Pallottino G V and Pizzella G 1998 On the efficiency of the coincidence search in gravitational wave experiments *Gen. Rel. Grav.* **30** 105
[137] Heng I S, Bonifazi P, Blair D G, Tobar M E and Ivanov E N 1999 Background noise reduction in gravitational wave detectors through use of an amplitude ratio filter *Gen. Rel. Grav.* to be published
[138] Coccia E, Fafone V and Frossati G 1995 On the design of ultralow temperature spherical gravitational wave detectors *Proc. 1st Eduardo Amaldi Conf. on Gravitational Wave Experiments* ed E Coccia, G Pizzella and G Ronga (Singapore: World Scientific)
[139] Coccia E, Lobo J A and Ortega J A 1995 Proposed gravitational wave observatory based on solid elastic spheres *Phys. Rev.* D **52** 3735
[140] Coccia E and Fafone V 1996 Coalescing binaries and spherical gravitational wave detectors *Phys. Lett.* A **213** 16
[141] Coccia E, Fafone V, Frossati G, ter Haar E and Meisel M W 1996 Eigenfrequencies and quality factors of vibration of aluminium alloy spherical resonators *Phys. Lett.* A **219** 263
[142] Bianchi N, Coccia E, Colacino C N, Fafone V and Fucito F 1996 Testing theories of gravity with a spherical gravitational wave detector *Class. Quantum Grav.* **13** 2865

[143] Vitale S, Cerdonio M, Coccia E and Ortolan A 1997 Gravitational-wave stochastic background detection with resonant-mass detectors *Phys. Rev.* D **55** 1741
[144] Merkowitz S M and Johnson W W 1995 Spherical gravitational wave antennas and the truncated icosahedral arrangement *Phys. Rev.* D **51** 2546
[145] Zhou C Z and Michelson P F 1994 Spherical resonant-mass gravitational wave detectors *Phys. Rev.* D **51** 2517
[146] Stevenson T R 1995 Signal to noise analysis for a spherical gravitational wave antenna instrumented with multiple transducers *Proc. 1st Eduardo Amaldi Conf. on Gravitational Wave Experiments* ed E Coccia, G Pizzella and G Ronga (Singapore: World Scientific)
[147] Stevenson T R 1997 Limits on the sensitivity of spherical gravitational wave detectors and on the accuracy of reconstructed signals *Phys. Rev.* D **56** 564
[148] Harry G M, Stevenson T R and Paik H J 1996 Detecability of gravitational wave events by spherical resonant-mass antennas *Phys. Rev.* D 54 2409
[149] Lobo J A 1995 What can we learn about gravitational wave physics with an elastic spherical antenna? *Phys. Rev.* D **52** 591
[150] Blair D G, Peng H and Ivanov E N 1992 Sapphire dielectric resonator transducers *J. Phys. D: Appl. Phys.* **25** 1110
[151] Tobar M E, Ivanov E N, Blair D G and Cuthbertson B D 1997 *Proc. 1st Int. Workshop for an Omnidirectional Gravitational Radiation Observatory* ed W F Velloso Jr, O D Aguiar and N S Magalhaes (Singapore: World Scientific)
[152] Cuthbertson B D, Tobar M E, Ivanov E N and Blair D G 1998 Sensitivity and optimization of a high-Q sapphire dielectric motion sensing transducer *IEEE Trans. UFFC* **45** 1303
[153] Tobar M E, Ivanov E N and Blair D G 1999 Parametric transducers for the advanced cryogenic resonant-mass gravitational wave detectors *Gen. Rel. Grav.* to be published
[154] Frasca S and Papa M 1995 Local array of high frequency antennas *Proc. 1st Eduardo Amaldi Conf. on Gravitational Wave Experiments* ed E Coccia, G Pizzella and G Ronga (Singapore: World Scientific)
[155] Harry G M, Jin I, Paik H J, Stevenson T R and Wellstood F C 2000 Two stage superconducting quantum interference device amplifier in a high-Q gravitational wave transducer *Appl. Phys. Lett.* **76** 1446
[156] Tobar M, Ivanov E and Blair D G 1999 NIOBE: improved noise temperature performance and background noise suppression *Proc. 3rd Eduardo Amaldi Conf. on Gravitational Waves* ed S Meshkov (New York: AIP) at press
[157] Tobar M E, Ivanov E N and Blair D G 1996 Future prospects for the University of Western Australia's cryogenic resonant-mass gravitational wave detector *Czech. J. Phys.* **46** 2909
[158] Gertsenshtein M E and Pustovoit V I 1962 On the detection of low frequency gravitational waves *Sov. Phys.–JETP* **16** 433
[159] Moss G E, Miller L R and Forward R L 1971 Photon-noise-limited laser transducer for gravitational antenna *Appl. Opt.* **10** 2495
[160] Forward R L and Mosse G E 1972 Wideband interferometric gravitational radiation antenna *Bull. Am. Phys. Soc.* **17** 1183
[161] Forward R L 1978 Wide-band laser interferometer gravitational-radiation experiment *Phys. Rev.* D **17** 379
[162] Shoemaker D, Schilling R, Schnupp L, Maischberger K and Rüdiger A 1988 Noise behaviour of the Garching 30-meter prototype gravitational-wave detector *Phys. Rev.* D **38** 423
[163] Robertson D I, Morrison E, Hough J, Killbourn S, Meers B J, Newton G P, Robertson N A, Strain K A and Ward H 1995 The Glasgow 10 m prototype laser interferometric gravitational wave detector *Rev. Sci. Instrum.* **66** 4447
[164] Abramovici A *et al* 1996 Improved sensitivity in a gravitational wave interferometer and implications for LIGO *Phys. Lett.* A **218** 157
[165] Drever R W P 1983 *Gravitational Radiation* ed N Deruelle and T Piran (Amsterdam: North-Holland)
[166] B J Meers 1988 Recycling in laser-interferometric gravitational-wave detectors *Phys. Rev.* D **38** 2317
[167] Vinet J-Y, Meers B J, Man C N and Brillet A 1988 Optimization of long-baseline optical interferometers for gravitational-wave detection *Phys. Rev.* D **38** 433
[168] Mizuno J, Strain K A, Nelson P G, Chen J M, Schilling R, Rüdiger A, Winkler W and Danzmann K 1993 Resonant sideband extraction: a new configuration for interferometric gravitational wave detectors *Phys. Lett.* A **175** 273
[169] Heinzel G, Mizuno J, Schilling R, Winkler W, Rüdiger A and Danzmann K 1996 An experimental demonstration of resonant sideband extraction for laser-interferometric gravitational wave detectors *Phys. Lett.* A **217** 305
[170] Man C N, Shoemaker D, Manh Pham Tu and Davey D 1990 External modulation technique for sensitive interferometric detection of displacement *Phys. Lett.* A **148** 8
[171] Fritschel P, Shoemaker D and Weiss R 1992 Demonstration of light recycling in a Michelson interferometer

with Fabry–Perot cavities *Appl. Opt.* **31** 1412
[172] Regehr M W, Raab F J and Whitcomb S E 1995 Demonstration of a power-recycled Michelson interferometer with Fabry–Perot arms by frontal modulation *Opt. Lett.* **20** 1507
[173] Schnier D, Mizuno J, Heinzel G, Lück H, Rüdiger A, Schilling R, Schremple M, Winkler W and Danzmann K 1997 Power recycling in the Garching 30 m prototype interferometer for gravitational-wave detection *Phys. Lett.* A **225** 210
[174] Fritschel P, Gonzalez G, Lantz B and Zuker M 1998 High power phase measurements limited by quantum noise and application to detection of gravitational waves *Phys. Rev. Lett.* **80** 318
[175] Gray M B, Stevenson A J, Bachor H-A and McClelland D E 1998 Broadband and tuned signal recycling with a simple Michelson interferometer *Appl. Opt.* **37** 5886
[176] Strain K A and Meers B J 1991 Experimental demonstration of dual recycling for interferometric gravitational-wave detectors *Phys. Rev. Lett.* **66** 1391
[177] Heinzel G, Stain K A, Mizuno J, Skeldon K D, Willke B, Winkler W, Schilling R, Rüdiger A and Danzmann K 1998 Experimental demonstration of a suspended dual recycling interferometer for gravitational wave detection *Phys. Rev. Lett.* **81** 5493
[178] Bradaschia C *et al* 1990 The VIRGO project: a wide band antenna for gravitational wave detection *Nucl. Instrum. Methods* A **289** 518
[179] Danzmann K *et al* 1995 GEO 600: A 600 m laser interferometric gravitational wave antenna *Proc. 1st Eduardo Amaldi Conf. on Gravitational Wave Experiments* ed E Coccia, G Pizzella and G Tonga (Singapore: World Scientific)
[180] Tsubono K 1995 300-m laser interferometer gravitational wave detector (TAMA 300) *Proc. 1st Eduardo Amaldi Conf. on Gravitational Wave Experiments* ed E Coccia, G Pizzela and G Tonga (Singapore: World Scientific)
[181] Drever R W P, Ford G M, Hough J, Kerr I, Munley A J, Pugh J R, Robertson N A and Ward H 1983 *Proc. 9th Int. Conf. on General Relativity and Gravitation* ed E Shmutzer (Cambridge: Cambridge University Press)
[182] Shilling R, Schnupp L, Winkler W, Billing H, Maischerger K and Rüdiger A 1981 A method to bolt out scattered light effects and its application to a gravitational wave detector *J. Phys. E: Sci. Instrum.* **14** 65
[183] Winkler W 1991 A Michelson interferometer using delay line *The Detection of Gravitational Waves* ed D G Blair (Cambridge: Cambridge University Press)
[184] Sun K X, Fejer M M, Gustafson E and Byer R L 1996 Sagnac interferometer for gravitational-wave detection *Phys. Rev. Lett.* **76** 3053
[185] Sun K X, Gustafson E K, Fejer M M and Byer R L 1997 Polarization-based balanced heterodyne detection method in a Sagnac interferometer for precision phase measurement *Opt. Lett.* **22** 1359
[186] Shaddock D A, Gray M B and McClelland D E 1998 Experimental demonstration of resonant sideband extraction with a Sagnac interferometer *Appl. Opt.* **37** 7995
[187] Mizuno J, Rüdiger A, Schilling R, Winkler W and Danzmann K 1997 Frequency response of Michelson- and Sagnac-based interferometers *Opt. Commun.* **138** 383
[188] Drever R W P 1995 Concepts for extending the ultimate sensitivity of interferometric gravitational wave detectors using non-transmissive optics with diffractive or holographic coupling *Proc. 7th Marcel Grossman Meeting on General Relativity* ed M Keiser and R T Jantzen (Singapore: World Scientific)
[189] Veitch P J, Munch J, Hamilton M W, Ottaway D, Greentree A and Tikhomirov A 1995 High power laser and novel optics for laser interferometric gravitational wave detectors *Aust. J. Phys.* **48** 999
[190] Sun K X and Byer R L 1998 All-reflective Michelson, Sagnac, and Fabry–Perot interferometers based on grating beam splitters *Opt. Lett.* **23** 567
[191] Caves C M 1980 Quantum-mechanical radiation-pressure fluctuations in an interferometer *Phys. Rev. Lett.* **45** 75
[192] Caves C M 1981 Quantum-mechanical noise in an interferometer *Phys. Rev.* D **23** 1693
[193] Meers B J 1988 Recycling in laser interferometric gravitational-wave detectors *Phys. Rev.* D **38** 2317
[194] McClelland D E 1995 An overview of recycling in laser interferometric gravitational wave detectors *Aust. J. Phys.* **48** 953
[195] GEO 600 1994 Proposal for a 600 m laser-interferometric gravitational wave antenna *Max-Plank-Institut Für Quantenoptik Report* MPQ 190
[196] Winkler W, Danzmann K, Rüdiger A and Schilling R 1991 Optical problems in interferometric gravitational wave antennas *Proc. 6th Marcel Grossmann Meeting* ed H Sato and T Nakamura (Singapore: World Scientific)
[197] Winkler W, Rüdiger A, Schilling R, Strain K A and Danzmann K 1994 Birefringence-induced losses in interferometers *Opt. Commun.* **112** 245
[198] Plissi M V *et al* 1998 Aspect of the suspension system for GEO 600 *Rev. Sci. Instrum.* **69** 3055
[199] Stephens M, Saulson P and Kovalik J 1991 A double pendulum vibration isolation system for a laser

interferometric gravitational wave antenna *Rev. Sci. Instrum.* **62** 924
[200] Tsubono K, Arays A, Kawabe K, Moriwaki S and Mio N 1993 Triple-pendulum vibration isolation system for a laser interferometer *Rev. Sci. Instrum.* **64** 2237
[201] Veitch P J, Robertson N A, Cantley C A and Hough J 1992 Active control of a balanced two-stage pendulum vibration isolation system and its application to laser interferometric gravity wave detectors *Rev. Sci. Instrum.* **64** 1330
[202] Del Fabbro R, Di Virgilio A, Giazotto A, Kautzky H, Montelatici V and Passuello D 1988 Performance of a gas spring harmonic oscillator *Rev. Sci. Instrum.* **59** 293
[203] Del Fabbro R, Di Virgilio A, Giazotto A, Kautzky H, Montelatici V and Passuello D 1988 First results from the Pisa seismic noise super-attenuator for low frequency gravitational wave detection *Phys. Lett.* A **132** 237
[204] Braccini S *et al* 1996 Seismic vibration mechanical filters for the gravitational waves detector VIRGO *Rev. Sci. Instrum.* **67** 2899
[205] Robertson N A, Drever R W P, Kerr I and Hough J 1982 Passive and active seismic isolation for gravitational radiation detectors and other instruments *J. Phys. E: Sci. Instrum.* **15** 1101
[206] Saulson P 1984 Vibration isolation for broadband gravitational wave antennas *Rev. Sci. Instrum.* **55** 1315
[207] Campani E, Giazotto A and Passuello D 1986 Performance of an analogue phase follower *Rev. Sci. Instrum.* **57** 79
[208] Giazotto A, Passuello D and Stefani A 1986 One-mile equivalent length interferometric pendulum for seismic noise reduction *Rev. Sci. Instrum.* **57** 1145
[209] Nelson P G 1991 An active vibration isolation system for inertial reference and precision measurement *Rev. Sci. Instrum.* **62** 2069
[210] Stebbins R, Newell D, Richman S, Bender P, Faller J and Mason J 1994 Low-frequency active vibration isolation system *Proc. SPIE* **2264** 27
[211] McLoughlin F A 1990 Integrated structural damping and control system design for high-order flexible systems *PhD Thesis* Stanford University
[212] Beccaria M *et al* 1997 Extending the VIRGO gravitational wave detection band down to a few Hz: metal blade springs and magnetic antisprings *Nucl. Instrum. Methods* A **394** 397–408
[213] Pinoli M, Blair D G and Ju L 1993 Test on a low-frequency inverted pendulum system *Meas. Sci. Technol.* **4** 995
[214] Saulson P R, Stebbins R T, Dumont F D and Mock S E 1994 The inverted pendulum as a probe of anelasticity *Rev. Sci. Instrum.* **65** 182
[215] Losurdo G *et al* 1999 An inverted pendulum preisolator stage for the VIRGO suspension system *Rev. Sci. Instrum.* **70** 2507
[216] Winterflood J, Zhou Z B and Blair D G 1999 Reducing low-frequency residual motion in vibration isolation to the nanometre level *Proc. 3rd Eduardo Amaldi Conf. on Gravitational Waves* ed S Meshkov (New York: AIP) at press
[217] van Kann F 1993 Private communications
[218] Blair D G, Liu J, Moghaddam E F and Ju L 1994 Performance of an ultra low-frequency folded pendulum *Phys. Lett.* A **193** 223
[219] Liu L, Winterflood J and Blair D G 1995 Transfer function of an ultralow frequency vibration isolation system *Rev. Sci. Instrum.* **66** 3216
[220] Liu J, Ju L and Blair D G 1997 Vibration isolation performance of an ultra-low frequency folded pendulum *Phys. Lett.* A **228** 243
[221] Barton M A and Kuroda K 1994 Ultra-low frequency oscillator using a pendulum with crossed suspension wires *Rev. Sci. Instrum.* **65** 3775
[222] Kanda N, Barton M A and Kuroda K 1994 Transfer function of a crossed wire pendulum isolation system *Rev. Sci. Instrum.* **65** 3780
[223] Winterflood J and Blair D G 1996 A long-period conical pendulum for vibration isolation *Phys. Lett.* A **222** 141
[224] Winterflood J and Blair D G 1998 Ultra-low frequency pre-isolation in three dimensions *Proc. 2nd Eduardo Amaldi Conf. on Gravitational Waves* ed E Coccia, G Pizzella and G Veneziano (Singapore: World Scientific)
[225] Braccini S *et al* 1993 An improvement in the VIRGO super attenuator for interferometric detection of gravitational wave: the sue of a magnetic antispring *Rev. Sci. Instrum.* **64** 310
[226] LaCoste L J B 1934 *Phys.* **5** 178
[227] Bertolini A, Cella G, DeSalvo R and Sannibale V 1999 Seismic noise filters, vertical resonance frequency reduction with geometrical anti-springs: a feasibility study *Nucl. Instrum. Methods* A **435** 475
[228] Saulson P 1990 Thermal noise in mechanical experiments *Phys. Rev.* D **42** 2437
[229] Kimball A L and Lovell D E 1927 *Phys. Rev.* **30** 948

[230] Ju L, Blair D G and Notcutt M 1993 Ultrahigh *Q* pendulum suspensions for gravitational wave detectors *Rev. Sci. Instrum.* **64** 1899
[231] Quinn T J, Speake C C and Brown L M 1992 Materials problems in the construction of long-period pendulums *Phil. Mag.* A **65** 261
[232] Kovalik J and Saulson P R 1993 Mechanical loss in fibres for low noise pendulums *Rev. Sci. Instrum.* **64** 2942
[233] Gillespie A and Raab F 1993 Thermal noise in the test mass suspensions of a laser interferometer gravitational-wave detector prototype *Phys. Lett.* A **178** 357
[234] Gillespie A and Raab F 1994 Suspensions losses in the pendula of laser interferometer gravitational-wave detectors *Phys. Lett.* A **190** 213
[235] Zuker M 1992 The LIGO 40 m prototype laser interferometer gravitational wave detector *Proc. 6th Marcel Grossmann Meeting* ed H Sato and T Nakamura (Singapore: World Scientific)
[236] Logan J E, Hough J and Robertson N A 1993 Aspect of the thermal motion of a mass suspended as a pendulum by wires *Phys. Lett.* A **183** 145
[237] Bernardi A, Majorana E, Puppo P, Rapagnani P, Ricci F and Testi E 1999 Suspension last stages for the mirrors of the VIRGO interferometric gravitational wave antenna *Rev. Sci. Instrum.* **70** 3463
[238] Hough J 1989 Prospect for gravitational wave detection with laser interferometer detectors *Proc. 5th Marcel Grossmann Meeting* ed Blair D G and Buckingham M J (Singapore: World Scientific)
[239] Rowan S *et al* 1999 Mechanical loss factors of material and suspension systems for advanced gravitational wave detectors *Proc. 3rd Eduardo Amaldi Conf. on Gravitational Waves* ed S Meshkov (New York: AIP) at press
[240] Braginsky V B, Mitrofanov V P and Vyatchanin S P 1994 Isolation of test masses in the advanced laser interfeormetric gravitational wave antennae *Rev. Sci. Instrum.* **65** 3771
[241] Braginsky V B, Mitrofanov V P and Tokmakov K V 1994 On the thermal noise from the violin modes of the test mass suspension in the gravitational wave antennae *Phys. Lett.* A **186** 18
[242] Logan J E, Hough J and Robertson N A 1992 An investigation of limitations to quality factor measurements of suspended masses due to resonances in the suspension wires *Phys. Lett.* A **170** 352
[243] González G I and Saulson P R 1994 Brownian motion of a mass suspended by an anelastic wire *J. Acoust. Soc. Am.* **96** 207
[244] Martin W 1987 Experiments and techniques for the detection of gravitational and pulsed electromagnetic radiation from astrophysical sources *PhD Thesis* University of Glasgow
[245] Ju L, Blair D G, Taniwaki M and Andrew R 1999 The quality factor of niobium flexure pendulums *Phys. Lett.* A **254** 239
[246] Bondu F and Vinet J-Y 1995 Mirror thermal noise in interferometric gravitational-wave detectors *Phys. Lett.* A **198** 74
[247] Gillespie A and Raab F 1995 Thermally excited vibration isolations of the mirrors of laser interferometer gravitational-wave detectors *Phys. Rev.* D **52** 577
[248] Logan J E, Robertson N A, Hough J and Veitch P J 1992 An investigation of some mechanical properties of materials for test masses in laser gravitational wave detectors *Proc. 6th Marcel Grossmann Meeting* ed H Sato and T Nakamura (Singapore: World Scientific)
[249] Traeger S, Willke B and Danzmann K 1997 Monolithically suspended fused silica substrates with very high mechanical *Q Phys. Lett.* A **225** 39
[250] Taniwaki M, Ju L, Blair D G and Tobar M E 1998 Design and verification of low acoustic loss suspension system for measuring the *Q*-factor of a gravitational wave detector test mass *Phys. Lett.* A **246** 273
[251] Startin W J, Beilby M A and Saulson P R 1998 Mechanical quality factors of fused silica resonators *Rev. Sci. Instrum.* **69** 3681
[252] Braginsky V B, Mitorfanov V P and Okhrimenko O A 1992 Oscillators for free mass gravitational antennas *JETP Lett.* **55** 432
[253] Locke C 1999 University of Western Australia Private communication
[254] Ju L, Notcutt M, Blair D G, Bondu R and Zhao C N 1996 Sapphire beamsplitters and test masses for advanced laser interferometer gravitational wave detectors *Phys. Lett.* A **218** 197
[255] Rowan S, Twyford S M, Hough J, Gwo D-H and Route R 1998 Mechanical losses associated with the technique of hydroxide-catalysis bonding of fused silica *Phys. Lett.* A **246** 471
[256] Ju L, Taniwaki M, Blair D G, Benabid F and Notcutt M 1998 Using sapphire test masses and niobium flexure suspension in laser interferometer gravitational wave detectors *Proc. 2nd Eduardo Amaldi Conf. on Gravitational Waves* ed E Coccia, G Pizzella and G Veneziano (Singapore: World Scientific)
[257] Gwo D H 1998 Stanford University Private communication
[258] Bradaschia C, Del Fabbro R, Di Virgilio A, Giazotto A, Hautzky H, montelatici V and Passuello D 1989 First results on the electronic cooling of the Pisa seismic noise super-attenuator for gravitational wave detection

Phys. Lett. A **137** 329
[259] Blair D G, Ju L and Peng H 1993 Vibration isolation for gravitational wave detection *Class. Quantum Grav.* **10** 2407
[260] Winterflood J, Blair D G, Schilling R and Notcutt M 1995 Position control system for suspended masses in laser interferometer gravitational wave detectors *Rev. Sci. Instrum.* **66** 174
[261] Ke W, Blair D G, Notcutt M and Ju L 1995 Normal mode suppression in all metal cantilever vibration isolators *Phys. Lett.* A **197** 275
[262] Veitch P J, Robterson N A, Cantley C A and Hough J 1993 Active control of a balanced two-stage pendulum vibration isolation system and its application to laser interferometric gravity wave detectors *Rev. Sci. Instrum.* **64** 1330
[263] Veitch P J, Hough J, Morrison E and Robertson D I 1996 Can piezoelectric accelerometers be used to actively damp the mechanical suspensions in laser interferometric gravitational wave detectors? *Rev. Sci. Instrum.* **67** 633
[264] Braccini S *et al* 1995 Low noise wideband accelerometer using an inductive displacement sensor *Rev. Sci. Instrum.* **66** 2672
[265] Luiten A N, Beccaria M and Zhang Z 1997 Ground tilt spectrum measured with a new high sensitivity rotational accelerometer *Rev. Sci. Instrum.* **68** 1889
[266] Linsay P S and Shoemaker D H 1982 Low-noise RF capacitance bridge transducer *Rev. Sci. Instrum.* **53** 1014
[267] Drever R W P 1982 Interferometric detectors for gravitational radiation *Gravitational Radiation, Proc. Les Houches Summer Institute* ed T Piran and N Deruelle
[268] Man C N, Shoemaker D, Manh Pham Tu and Dewey D 1990 External modulation technique for sensitive interferometric detection of displacements *Phys. Lett.* A **148** 8
[269] Gray M B, Stevenson A J, Harb C C, Bachor H-A and McClelland D E 1996 External phase-modulation interferometry *Appl. Opt.* **35** 1623
[270] Schnupp L 1988 Contribution to the European collaboration meeting on interferometric detection of gravitational waves (Sorrento) Unpublished
[271] Takahashi R, Mizuno J, Miyoki S and Kawashima N 1994 Control of a 10 m delay line laser interferometer using the pre-modulation method *Phys. Lett.* A **187** 157
[272] Regehr M W, Raab F J and Whitcomb S E 1995 Demonstration of a power-recycled Michelson interferometer with Fabry–Perot arms by frontal modulation *Opt. Lett.* **20** 1507
[273] Flaminio R and Heitmann H 1996 Longitudinal control of an interferometer for the detection of gravitational waves *Phys. Lett.* A **214** 112
[274] Ando M, Kawabe K and Tsubono K 1997 Signal separation technique for a power-recycled interferometric gravitational wave detector *Phys. Lett.* A **237** 13
[275] Billing H, Maischberger K, Rüdiger A, Schilling R, Schnupp L and Winkler W 1979 An argon laser interferometer for the detection of gravitational radiation *J. Phys. E: Sci. Instrum.* **12** 1043
[276] Shoemaker D, Fritschel P, Giaime J, Christensen N and Weiss R 1991 Prototype Michelson interferometer with Fabry–Perot cavities *Appl. Opt.* **30** 3133
[277] Stevenson A J, Gray M B, Bachor H A and McClelland D E 1993 Quantum noise limited interferometer phase measurements *Appl. Opt.* **32** 3481
[278] Stevenson A J, Gray M B, Harb C C, Bachor H A and McClelland D E 1995 Interferometer with internal and external phase modulation *Aust. J. Phys.* **48** 971
[279] Lyons T T 1997 An optically recombined laser interferometer for gravitational wave detection *PhD Thesis* California Institute of Technology
[280] Barone F, Di Fiore L, Milano L and Russo G 1994 A digital approach to automatic control of a long-baseline interferometric antenna for gravitational wave detection *Meas. Sci. Technol.* **5** 1187
[281] Barone F, Di Fiore L, Milano L, Russo G and Solimeno S 1992 Automatic alignment of a Michelson interferometer *IEEE Trans. Nucl. Sci.* **39** 232
[282] Barone F, Calloni E, Di Fiore L, Grado A, Milano L and Russo G 1996 Digitally controlled interferometer prototype for gravitational wave detection *Rev. Sci. Instrum.* **67** 4353
[283] Heflin E G and Kawashima N 1995 *The Institute of Space and Astronautical Science, Japan ISAS Research Note* ISAS RN 567
[284] Heflin E G and Kawashima N 1995 *The Institute of Space and Astronautical Science, Japan ISAS Research Note* ISAS RN 563
[285] Drever R W P, Hall T L, Kowalski F V, Hough J, Ford G M, Munley A G and Ward H 1983 Laser phase and frequency stabilization using an optical resonator *J. Appl. Phys.* B **31** 97
[286] Pound R V 1946 *Rev. Sci. Instrum.* **17** 490
[287] Day T, Gustafson E K and Byer R L 1992 Sub-hertz relative frequency stabilization of two-diode laser-pumped

Nd:YAG lasers locked to a Fabry–Perot interferometer *IEEE Trans. Quantum Electron.* **28** 1106
[288] Shoemaker D, Brillet A, Man C N and Cregut O 1989 Frequency-stabilized laser-diode-pumped Nd:YAG laser *Opt. Lett.* **14** 609
[289] Kawamura S, Abramovici A and Zuker M E 1997 Improved multistage wideband laser frequency stabilization *Rev. Sci. Instrum.* **68** 223
[290] Bondu F, Fritschel P, Man C N and Brillet A 1996 Ultrahigh-spectral-purity laser for the VIRGO project *Opt. Lett.* **21** 582
[291] Nakagawa K, Katsuda T and Ohtsu M 1995 Short-term frequency stabilization of diode-laser pumped Nd:YAG lasers using double pendulum suspended cavities *J. Appl. Phys.* B **66** 489
[292] Araya A, Mio N, Tsubono K, Suehiro K, Telada S, Ohashi M and Fujimoto M 1997 Optical mode cleaner with suspended mirrors *Appl. Opt.* **36** 1446
[293] Skeldon K D, Strain K A, Grant A I and Hough J 1996 Test of an 18-m-long suspended mode cleaner cavity *Rev. Sci. Instrum.* **67** 2443
[294] Cotter D 1983 Stimulated Brillouin scattering in monomode optical fibre *J. Opt. Commun.* **4** 10
[295] Shiraki K, Ohashi M and Tateda M 1995 Suppression of stimulated Brillouin scattering in a fibre by changing the core radius *Electron. Lett.* **31** 668
[296] Rudiger A, Shilling R, Schnupp L, Winkler V, Billing H and Maischberger K 1998 A mode selector to suppress fluctuations in laser beam geometry *Opt. Acta* **28** 641
[297] Vinet J Y, Hello P H, Man C N and Brillet A 1992 A high accuracy method for the simulation of non-ideal optical cavity *J. Physique* I **2** 1287
[298] McClelland D E, Savage C M, Trigdell A J and Mavaddat R 1993 Tolerance of dual recycling laser interferometric gravitational wave detectors to mirror tilt and curvature errors *Phys. Rev.* D **48** 5474
[299] Fritschel P 1997 The LIGO project: progress and plans *Proc. 2nd Eduardo Amaldi Conf. on Gravitational Waves* ed E Coccia, G Pizzella and G Veneziano (Singapore: World Scientific)
[300] Walsh C, Davis G, Farrant D, Leistner A, Lesha F, Oreb B, Pavlovic E, Seckold J and Stuart W 1998 Fabrication and measurement of optics for gravitational wave detectors *13th National Congress of the Australian Institute of Physics (Fremantle, Western Australia, Sept. 1998)*
[301] Blair D G *et al* 1997 Development of low-loss sapphire mirrors *Appl. Opt.* **36** 337
[302] Mackowski J M, Pinard L, Dognin L, Ganau P, Lagrange B, Michel C and Morgue M 1998 VIRGO mirrors: wavefront control *Report* Univ. Lyon I, 5
[303] Logan J E, Robertson N A and Hough J 1994 Measurements of birefringence in a suspended sample of fused silica *Opt. Commun.* **107** 342
[304] Benabid F, Notcutt M, Ju L and Blair D G 1998 Birefringence measurements of sapphire test masses for laser interferometer gravitational wave detector *Phys. Lett.* A **237** 337
[305] Hello P and Vinet J Y 1990 Analytical models of thermal aberrations in massive mirrors heated by high power laser beams *J. Physique* **51** 1267
[306] Hello P and Vinet J Y 1990 Analytical models of transient thermoelastic deformations of mirrors heated by high power cw laser beams *J. Physique* **51** 2243
[307] Hello P and Vinet J Y 1993 Simulation of thermal effects in interferometric gravitational-wave detectors *Phys. Lett.* A **178** 351
[308] Strain K A, Danzmann K, Mizuno J, Nelson P G, Rüdiger A, Schilling R and Winkler W 1994 Thermal lensing in recycling interferometric gravitational wave detectors *Phys. Lett.* A **194** 124
[309] Blair D, Cleva F and Man C N 1997 Optical absorption measurements in monocrystalline sapphire at 1 μm *Opt. Mater.* **8** 233
[310] Kane T J and Byer R L 1985 Monolithic, unidirectional single-mode Nd:YAG ring laser *Opt. Lett.* **10** 65
[311] Freitag I, Golla D, Knoke S, Schöne W, Zellmer H, Tünnermann A and Welling H 1995 Amplitude and frequency stability of a diode pumped Nd:YAG laser operating at a single-frequency continuous-wave output power of 20 W *Opt. Lett.* **20** 462
[312] Siegman A E 1986 *Lasers* (Mill Valley, CA: University Science Books)
[313] Yang S T, Imai Y, Oka M, Eguchi N and Kubota S 1996 Frequency-stabilized, 10-W continuous-wave, laser-diode, end-pumped, injection-locked, Nd:YAG laser *Opt. Lett.* **21** 1676
[314] Farinas A D, Gustafson E K and Byer R L 1994 Design and characterization of a 5.5-W, CW, injection-locked fibre-coupled, laser-diode-pumped Nd:YAG miniature-slab laser *Opt. Lett.* **19** 114
[315] Harb C C, Ralph T C, Huntington E H, Freitag I, McClelland D E and Bachor H A 1996 Intensity noise properties of injection locked lasers *Phys. Rev.* A **54** 4370
[316] Harb C C, Ralph T C, Huntington E H, Bachor H A and McClelland D E 1997 Intensity noise dependence of Nd:YAG lasers on their diode-laser pump sources *J. Opt. Soc. Am.* B **14** 2936
[317] Huntington E H, Harb C C, Ralph T C, McClelland D E and Bachor H A 1998 Feedback control of the intensity

noise of injection locked lasers *Opt. Commun.* **145** 359
[318] Harb C C, Gray M B, Bachor H-A, Schilling R, Rottengatter P, Freitag I and Welling H 1994 Suppression of the intensity noise in a diode-pumped neodymium:YAG nonplanar ring laser *IEEE J. Quantum Electron.* **25** 767
[319] Golla D, Bode M, Knoke S, Schone W and Tunnermann A 1996 62-W CW TEM/sub 00/ Nd:YAG laser side-pumped by fibre-coupled diode lasers *Opt. Lett.* **21** 210
[320] Shine R J Jr, Alfrey A J and Byer R L 1995 40-W CW, TEM00-mode, diode-laser-pumped, Nd:YAG miniature-slab laser *Opt. Lett.* **20** 459
[321] Richards J and McInnes A 1995 Versatile, efficient, diode-pumped miniature slab laser *Opt. Lett.* **20** 371
[322] Hough J 1997 LISA—new development in design *Proc. 2nd Edoardo Amaldi Conf. on Gravitational Waves* ed E Coccia, G Pizzella and G Veneziano (Singapore: World Scientific)
[323] LISA web page: http://lisa.jpl/nasa/gov
[324] Drever R W P 1991 Fabry–Perot cavity gravity-wave detectors *The Detection of Gravitational Waves* ed D G Blair (Cambridge: Cambridge University Press)
[325] Meers B J 1989 The frequency response of interferometric gravitational wave detectors *Phys. Lett.* A **142** 465
[326] Ju L and Blair D G 1994 Low resonant frequency cantilever spring vibration isolator for gravitational wave detectors *Rev. Sci. Instrum.* **65** 3482

Gravity Waves

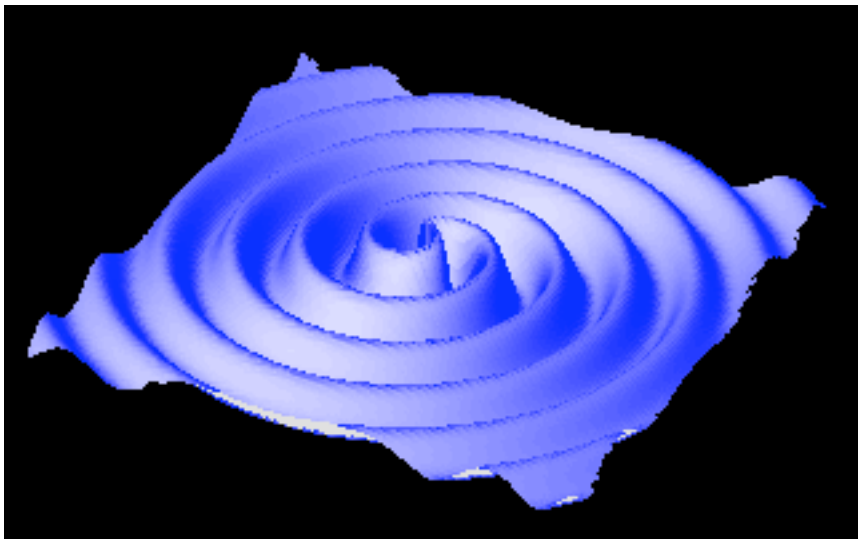
Another new prediction of Einstein's theory of gravity (general relativity) is *gravity waves*, "ripples of spacetime" that travel at the speed of light. Notice that these are ripples *of* spacetime, not ripples *in* spacetime like electromagnetic waves (also called light). Gravity waves are also called *gravitational radiation.* There is no such phenomenon in Newton's theory of gravity.

(Side note: There is a much more ordinary phenomenon also called "gravity waves" in the Earth's oceans and atmosphere: this is simply the fact that if a fluid rises then it will fall again due to gravity. An example is surface waves on the ocean. So when you find an article about gravity waves, check which type they are talking about. From now on, when I say "gravity wave" I mean waves of spacetime as predicted by general relativity)

Gravity waves are generated by the motion of, or changes in, masses in spacetime. For example, if the Sun were to suddenly disappear, the Earth would continue to move in its orbit along the straightest path in the curved spacetime generated by the Sun for eight more minutes. This is because it takes light 8 minutes to get from the Sun to the Earth, so it will take 8 minutes for the change in the spacetime curvature to get from the previous position of the Sun to the Earth. In other words the Earth's orbit would change at the same time that we saw the Sun disappear on Earth, 8 minutes after it actually happened. Don't worry, we're very sure the Sun will not actually disappear.

Actually, the special theory of relativity demands that gravitational changes move no faster than the speed of light, otherwise you could use the gravitational effect of a mass to transmit signals faster than light. This would contradict the understanding of spacetime we have from special relativity.

The detection of gravity waves is an important test of Einstein's general relativity. These waves are very difficult to detect because their effects are very weak for gravity waves generated by most things that happen in the Universe. For example, a rotating spherical mass (like the sun) does not generate any gravity waves at all. Planets orbiting a star generate gravity waves that are far too weak to detect. Normal stars orbiting each other generate gravity waves that stronger, but still too weak to detect from a great distance. Here's a picture of the gravity wave emitted from two stars orbiting each other.



A representation of the gravity waves emitted by two stars orbiting each other.

Only very large masses interacting very close to each other will generate gravity waves that we can hope to detect here on Earth. The strongest gravity waves would be generated by the collision of two black holes. Such collisions are rare, however. More likely is a large mass falling into a black hole such as the one in the center of our Galaxy. This is also rare, but would happen more often than a black hole collision.

Therefore gravity waves are difficult to detect. As I'll describe below, there are several experiments underway to detect them. The best evidence to date, however, is indirect, found by examining the orbit of two very massive but close stars around each other. I'll also describe this below.

What Does a Gravity Wave Look Like?

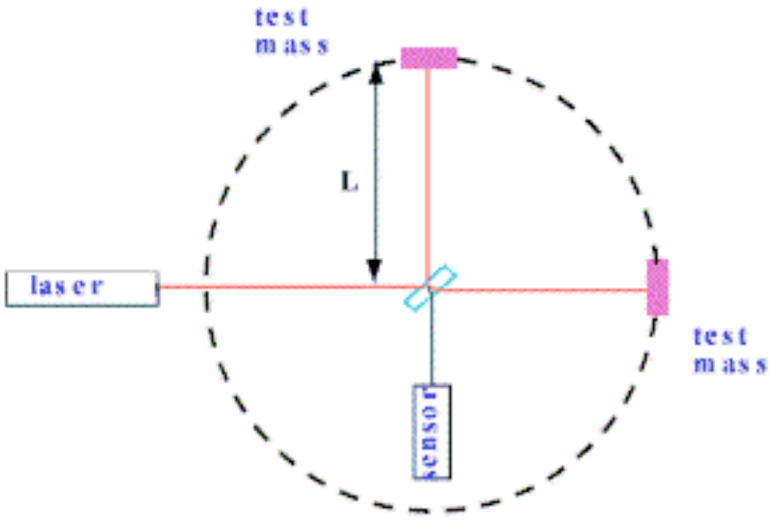
Gravity waves are ripples *of* spacetime, not waves of some material in spacetime. Gravity waves causes space to warp as time passes. Because it is space itself that is warping, everything in that space will warp with it. Therefore you won't see a gravity wave as it passes.

Luckily for physicists, for many gravity waves the warping in one direction will be different from the warping in another direction. We can therefore detect a gravity wave by observing lengths in different directions. If the length changes one way in one direction and a different way in another direction there is a good chance that a gravity wave has passed by.

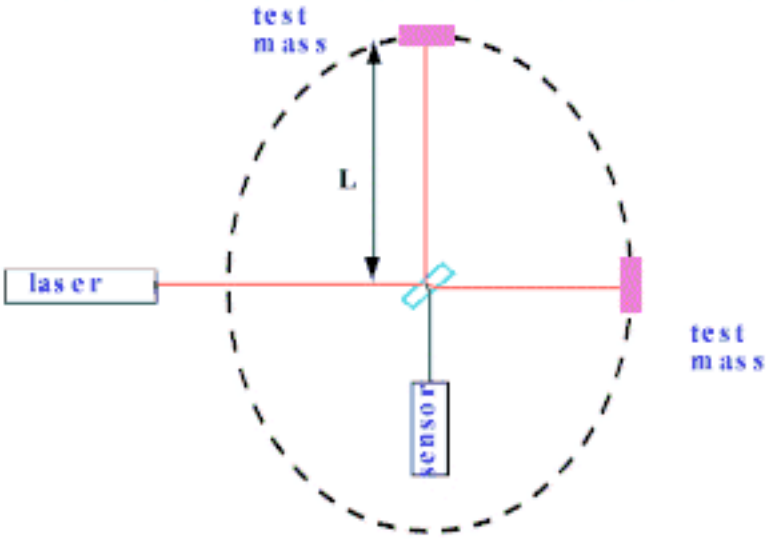
The first attempt to detect gravity waves in this way was by Joseph Webber in the late 60's. He placed very sensitive strain gauges on a large metal bar. If a gravity wave goes by, then the different stretching of the bar in various directions would appear as varying strains. Webber never detected gravity waves, because the bar was not large enough and was subject to too much noise due to changes in temperature and so on.

Joseph Webber and his gravity wave detector.

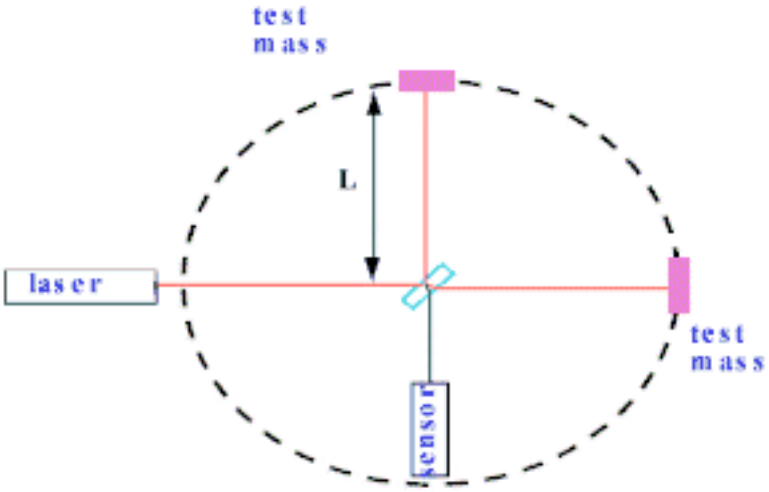
A number of experiments are now coming into operation that try to measure differences in length in various directions by using the interference of light: a laser beam is split in two, sending the two beams along perpendicular paths, then reflecting them back so they can be recombined. The *interference* of the recombined beams will show a length difference as variation in the brightness of the recombined beams. Here's some pictures showing such an experiment in operation:



An interference-based gravity wave detector. The line is the path of the laser beam. This is the picture when no gravity wave is passing.



Here is the situation when a gravity wave is passing: first the beams are stretched in one direction…

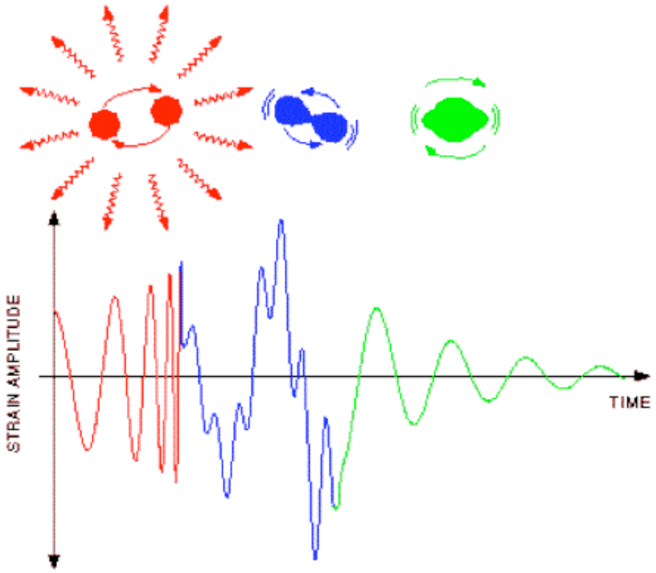


then they are stretched in the other. The sensor will change in brightness in a characteristic way.

Two large interference-based experiments are now operating in the US, one in Livingston, Louisiana and the other in Hanford, Washington. Both are part of the Laser Interferometer Gravity-Wave Observatory (LIGO), run by Cal Tech and MIT. These devices are quite large, with the arm of each laser path (L in the above diagrams) being 2.5 miles. The reason there are two of them far apart is to make sure any signals detected at one is not due to some local noise. A real gravitational wave would be detected at both experiments. LIGO has been taking data since April of this year, with increasing sensitivity over time. Hopefully we'll see an announcement of gravity wave detection soon, though LIGO is just barely at the required sensitivity to detect gravity waves from very large sources.

The LIGO facility in Hanford, Washington, showing one 2.5 mile arm, with the other arm going out of the picture to the left.

Here's a picture of the signal the LIGO experiment may detect from two massive stars coalescing after orbiting each other very closely:

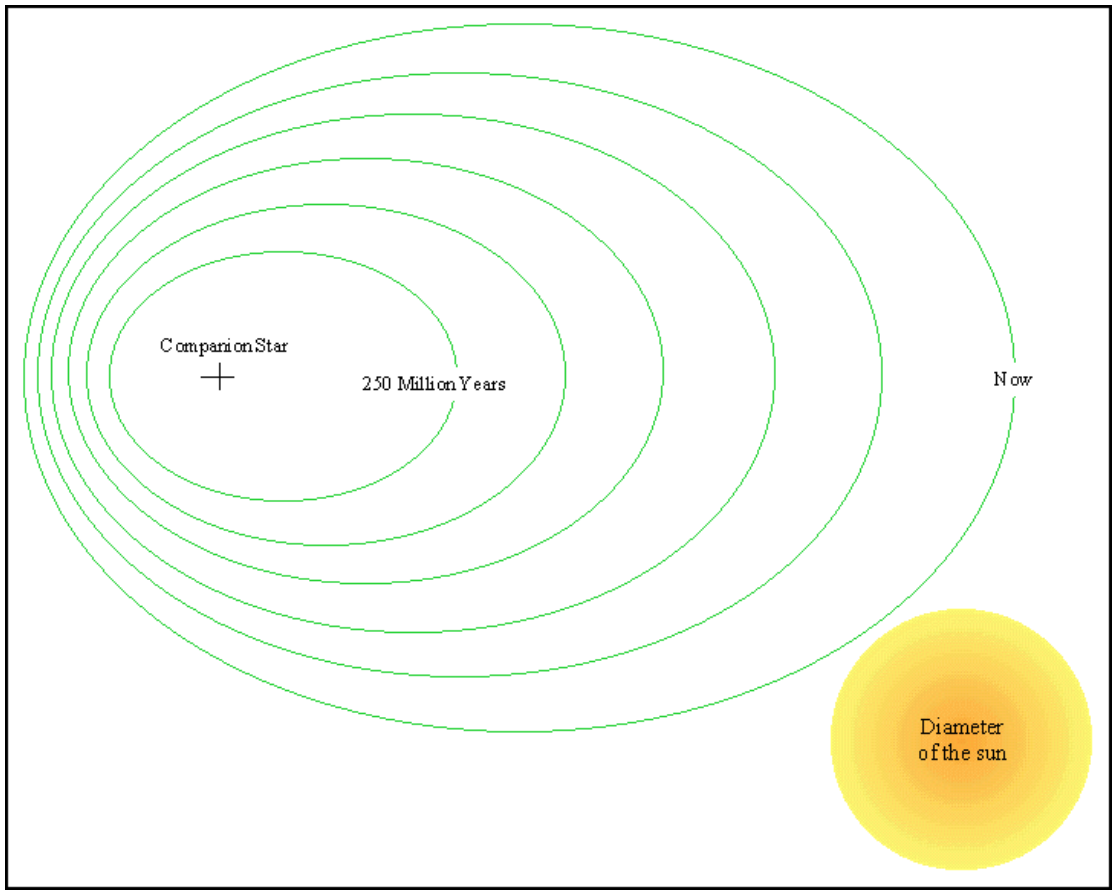


The top of the figure shows two very massive stars coalescing after being in a close orbit. The bottom shows the gravity wave signal that LIGO would detect from this event.

Several other interferometer-based experiments are being built in France, Germany, Japan and one is proposed in Australia. The most ambitious and sensitive proposed interferometer-based experiment is the Laser Interferometer Space Antenna (LISA), a fleet of three spacecraft that would form an equilateral triangle in space, measuring the changes in the length of the triangle with laser beams between the satellites. The sides of the triangle would be about 3 million miles.

The Indirect Detection of Gravity Waves

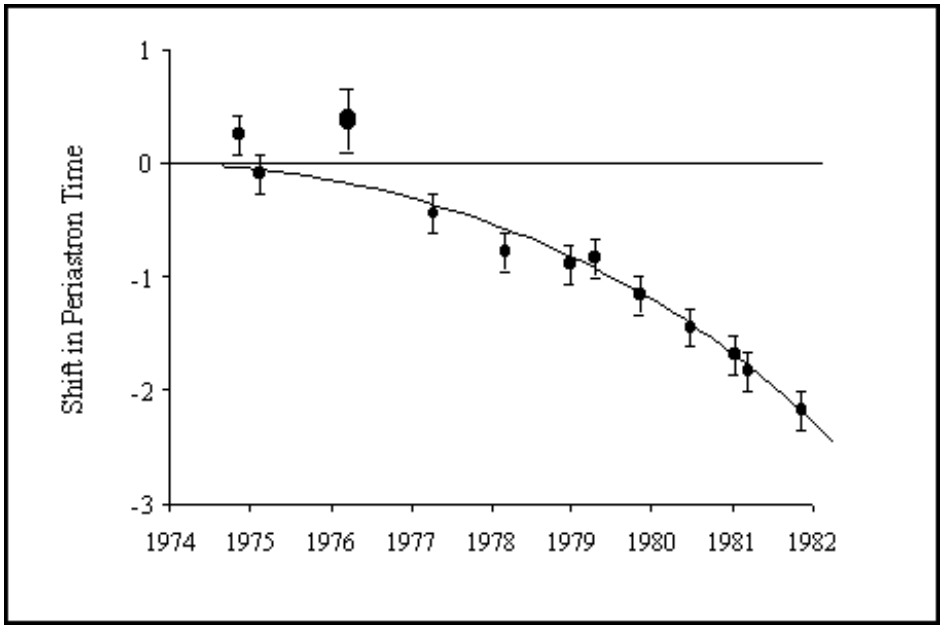
The best evidence for gravity waves to date is indirect, based on observations of the binary star system PSR1913+16. One of the stars in this system is a pulsar, which is a massive star in the last stages of life that has collapsed to a very small size, on the order of 10 miles in diameter. Both stars in PSR1913+16 are about 1.4 times the mass of the sun, and are orbiting very close together. Their orbits are very elliptic, with a closest approach of only 776,000 miles (a little more than the radius of the sun) and a furthest distance of 2 million miles (for comparison, the average distance between the Earth and the sun is about 93 million miles). The orbital period is only 7.75 *hours*.



The orbits of PSR1913+16 (from the point of view of one of the stars) compared with the size of the sun. The ellipse marked “now” is the current orbit. The other ellipses show the future orbits as the stars lose energy by emitting gravity waves, which cause them to slow down and fall into smaller orbits.

Because each star in PSR1913+16 is so massive and they get so close together, their gravitational field is very strong (in other words the spacetime between the pulsars is highly curved). Therefore they emit stronger gravity waves than most objects in the sky. These gravity waves are still too weak to detect even with LIGO, but they carry energy away from PSR1913+16. This energy loss causes the stars in PSR1913+16 to slow down, which causes a change in their orbit that we can detect from Earth. From these observations we can reconstruct the energy loss, and it exactly matches the energy of the gravity waves predicted by general relativity!

If you want a little more detail, here it is. The energy loss is measured by observing the time of closest approach between the stars (which is observed via the very fast pulses of the pulsar). If there were no energy loss then the time between closest approaches would be constant. If energy were being lost due to gravity waves, then the time between closest approaches would become shorter as the stars moved closer together. The closer the stars are together, the larger the gravity waves so the greater the energy loss, which moves the stars closer together still. Detailed calculations from general relativity predict that the time between closest approaches would decrease at an ever increasing rate. When the observations of the closest approach times are plotted over several years, they dramatically match the prediction of general relativity.



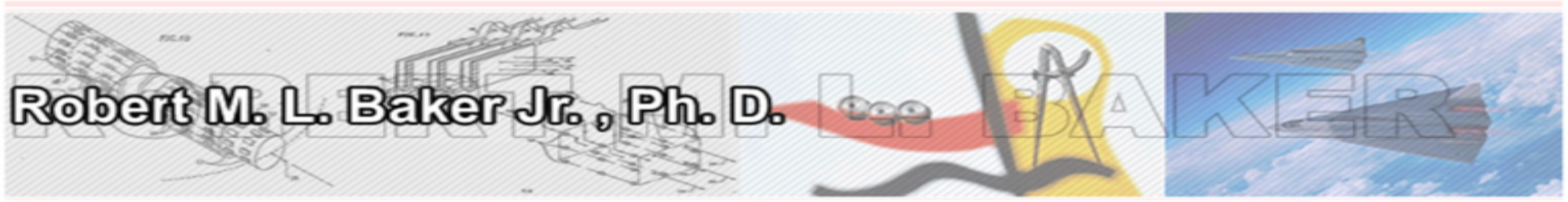
The change in time of closest approach of the stars observed over several years (dots with error bars) compared to the prediction of general relativity (curved line). The horizontal straight line (at 0) is the prediction if there were no gravity waves.

Though indirect, this observation is a dramatic vindication of general relativity.

GRAVWAVE ® LLC.

Applications of High-Frequency Gravitational Waves to Aerospace Systems

AIAA ASAT Conference, May 1, 2010



| PHYSICS | ENGINEERING | SCIENCE | ASTRODYNAMICS | MATHAMATICS | ASTRONAUTICS | RESEARCH | HIGH-FREQUENCY GRAVITATIONAL WAVES |

PREFACE

The predictions in this presentation of the benefits of high-frequency gravitational wave based Aerospace applications are theoretical at this time. Evidence of their value is contingent upon laboratory experiments in their generation and detection. Nonetheless, given their potential vital aerospace, strategic military and economic importance, I believe that these possible applications are important motivations for research and development.— Robert M L Baker, Jr.

EXECUTIVE SUMMARY

GRAVWAVE ® LLC.

- For several years the Peoples Republic of China has funded HFGW research programs involving dozens of their scientists and well-known Russian scientists have been involved in HFGW research for over four decades. HFGW detectors have been built by the British, Italians and the Japanese
- Technology developed by GravWave LLC and other institutions overseas can lead to devices, some already constructed overseas, that can generate and detect HFGWs in the laboratory.
- Because of their unique characteristics, HFGWs could be utilized for uninterruptible, very low-probability-of-intercept (LPI) communications among satellites.
- Other potential very theoretical aerospace applications are propulsion, including “moving” space objects and missiles in flight, frustrating anti-missile and anti-satellite systems, surveillance through buildings and the Earth itself, and remote initiation of nuclear events.
- The important potential aerospace applications of this enabling technology are motivations for research and development and such an R&D program in the United States is recommended for immediate initiation.

HFGW COMMUNICATION

GRAVWAVE ® LLC.

Operational Capability:

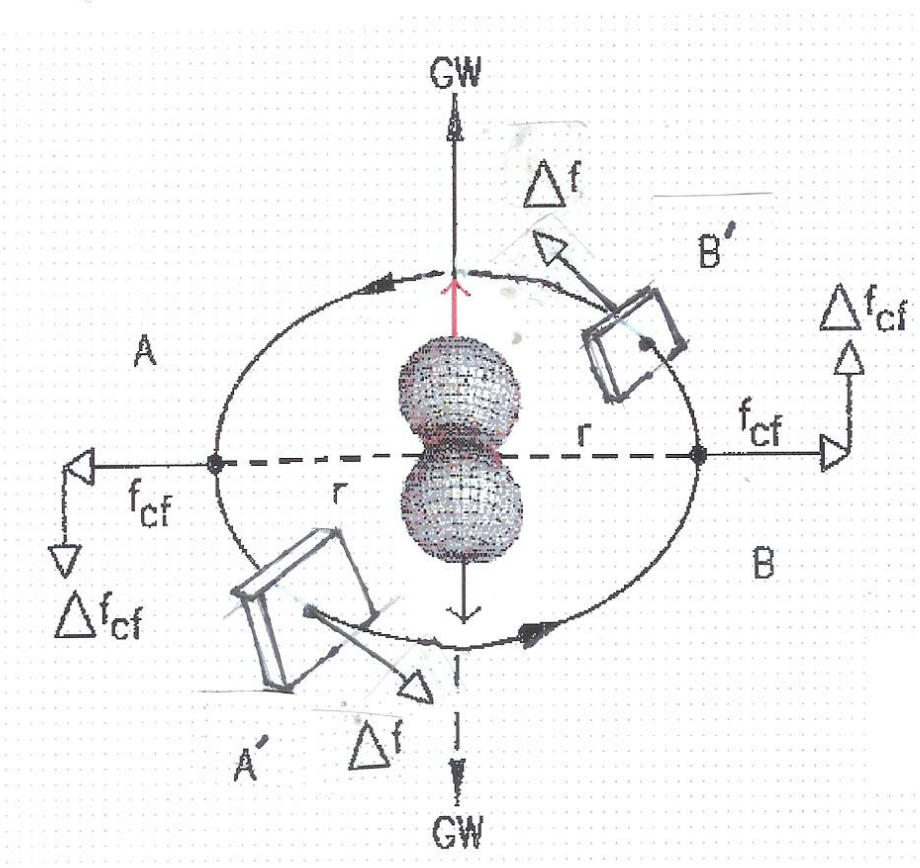
DATE

Russia 2017
USA 2019

Broadband Communications
Worldwide through the Earth

[Operational capability predictions are based on very rough estimates by the author from conversations and impressions gained during four international HFGW Workshops (MITRE2003, Austin2007, MSFC 2009 and Johns Hopkins APL 2010) and trips to China in 2004, 2006 and 2008 and to Europe (2002 and 2009) and the Middle East in 2009]

Radiation pattern calculated by Landau & Lifshitz, Section 110 Page 356.



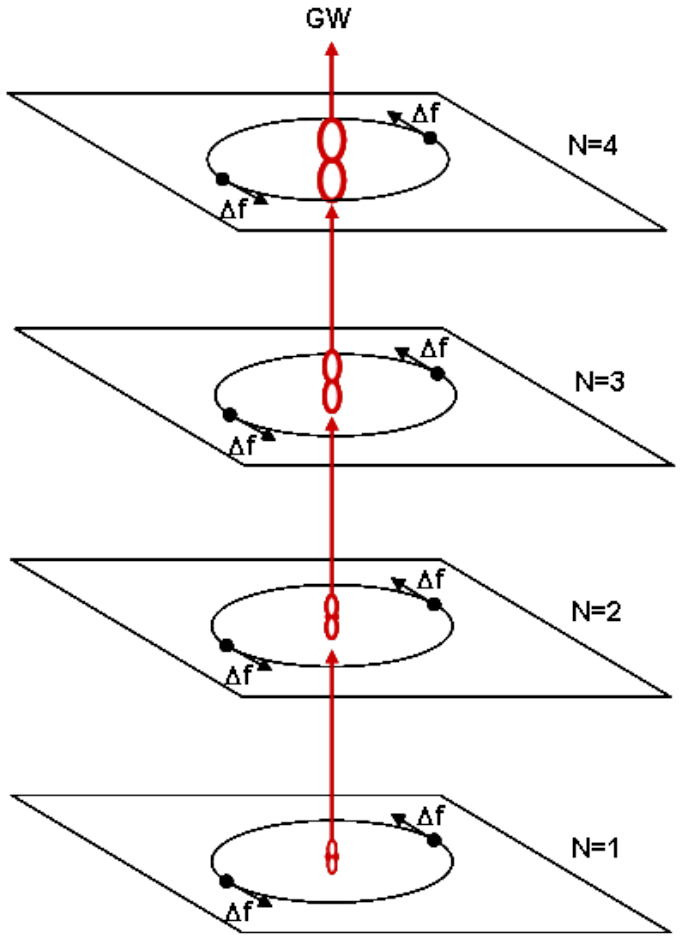
KEY EQUATIONS

- $P = 1.76 \times 10^{-52} \, (2r \, \Delta f / \Delta t)^2$ W,

- $S(1.0) = F(1.0) = N^2 F(1.0)^{N=1}$
  $= N^2 \, (0.336) \, P_i$ W/m^2 and

- $A = 1.28 \times 10^{-18} \sqrt{S} / \, v_{GW}$ m/m,

GW Flux Growth Analogous to Stack of *N* Orbital Planes



Magnetron FBAR (Piezoelectric Crystal) HFGW Generator.

GRAVWAVE ® LLC.

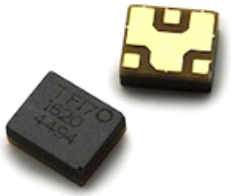
HFGW Generator

Using Magnetron-FBAR (Piezoelectric Crystals)

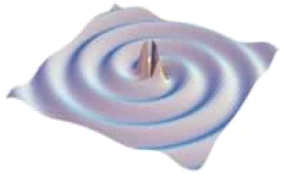
Similar to Romero and Dehnen (1981)

Magnetrons
(1000s)

Microwave
radiation
(2.45 GHz)



Film Bulk Acoustic
Resonator (FBAR)
piezoelectric crystals
(millions)

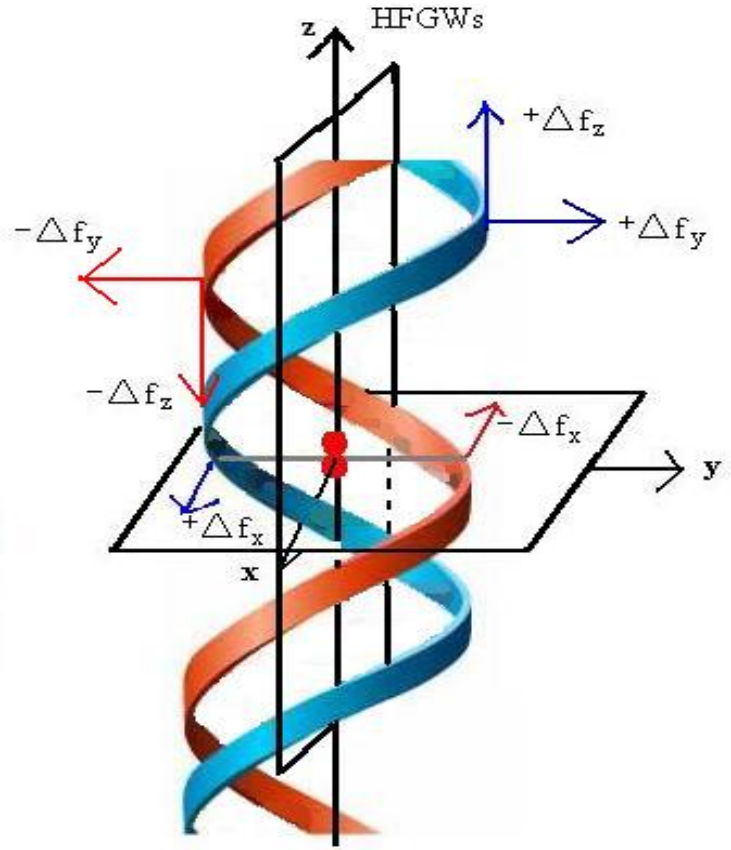


HFGWs

(4.9 GHz)

Double-Helix HFGW Generator FBAR Array (Patent Pending).

GRAVWAVE ® LLC.



Birmingham University HFGW Detector.

GRAVWAVE ® LLC.

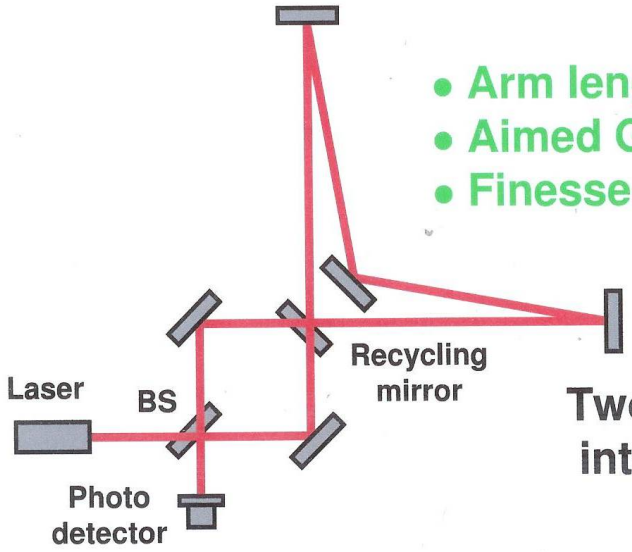
INFN Genoa HFGW Detector.

GRAVWAVE ® LLC.

The National Astronomical Observatory of Japan 100MHz Detector. Cruise (2008)

UNIVERSITY OF BIRMINGHAM

Development of 100MHz GW detectors at National Astronomical Observatory of Japan

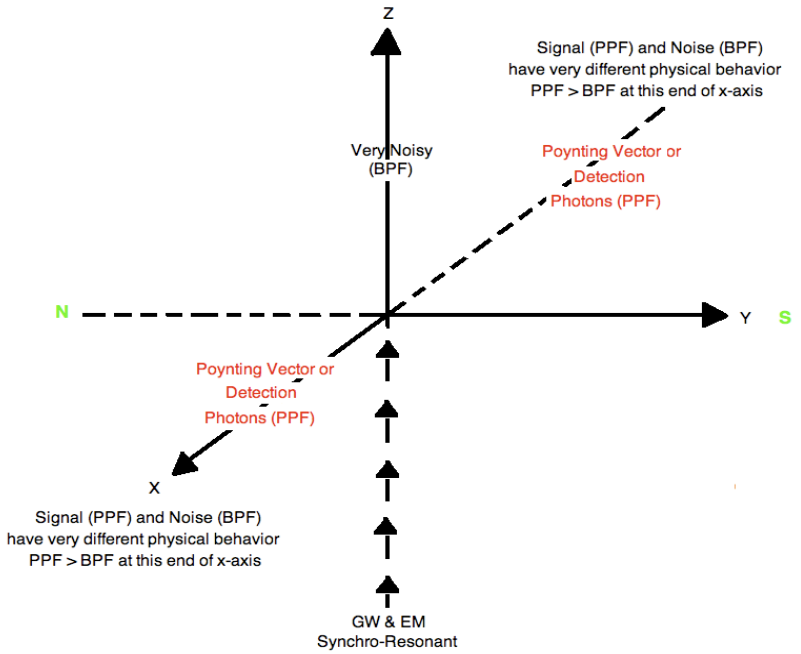


- Arm length: 75 cm
- Aimed GW frequency: 100 MHz
- Finesse: 100

Two synchronous recycling interferometers were built!

Synchronous recycling Interferometer (Concept: Drever 1983)

Li-effect detection photons directed to locations at the ends of the x-axis that are less affected by noise



The result of the intersection of the parallel and superimposed EM and GW beams, according to the Li effect, is *new EM photons moving off in a direction (both ways on the x-axis) perpendicular to the directions* of the *beams (GB and HFRGWs) on the z-axis and of the magnetic field (on the y-axis)*

Li Effect vs. classical *(inverse)* Gertsenshtein-Effect.

GRAVWAVE ® LLC.

- First published in 1992. Subsequently the “Li Effect” was validated by eight journal articles, independently peer reviewed by scientists well versed in general relativity, (Li, Tang and Zhao, 1992; Li and Tang, 1997; Li, Tang, Luo, 2000; Li, Tang and Shi, 2003; Li and Yang, 2004; Li and Baker, 2007; Li, et al., 2008; Li, et al., 2009) including capstone papers, Li, et al.(2008) and Li et al. (2009).

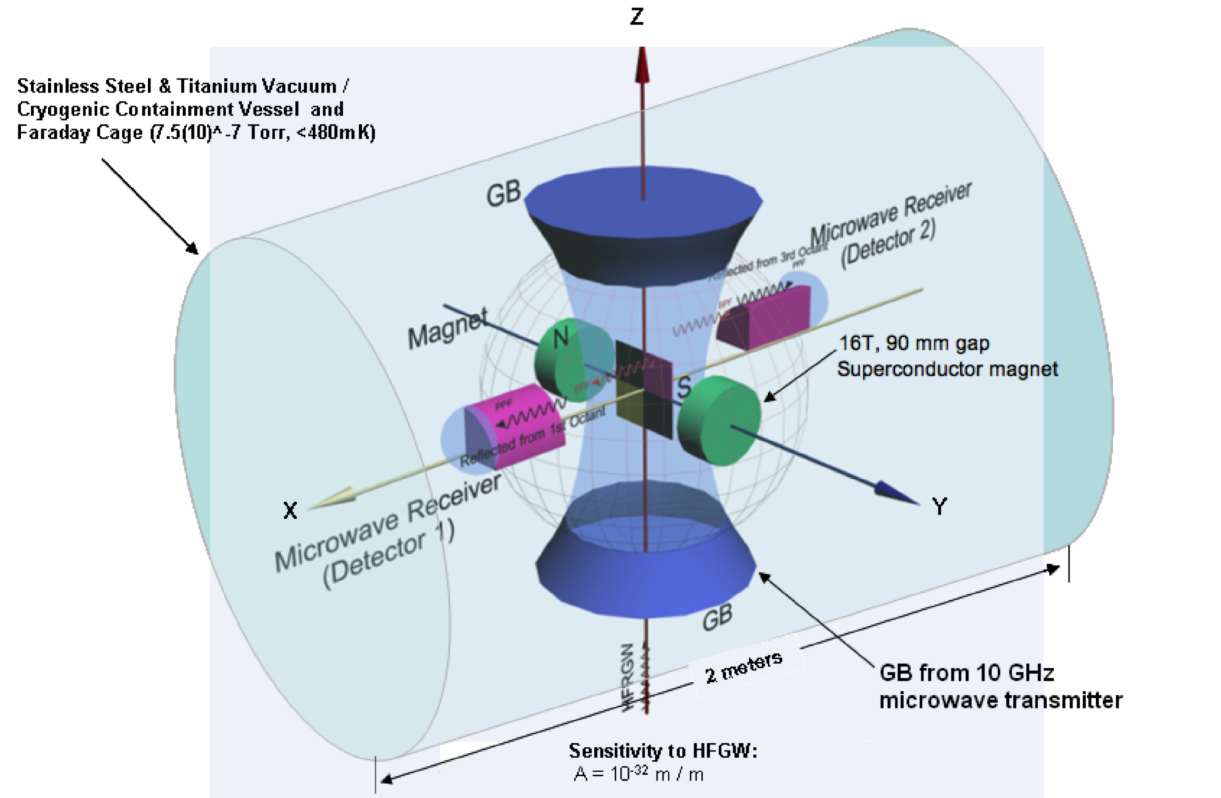
- The Li-Effect is *very different* from the classical *(inverse)* Gertsenshtein-Effect. the Li-effect produces a *first-order* perturbative photon flux (PPF), proportional to the amplitude of the gravitational wave (GW) A not A^2 as is the case of the Gertsenshtein-Effect. Thus the Gertsenshtein- Effect is 10^{-30} weaker than the Li-Effect for the detection of relic HFGWs. **The JASON report (Eardley, 2008) confuses the two effects and erroneously suggests that the Li-Baker HFGW Detector utilizes the inverse Gertsenshtein effect.**

Rayleigh scattering mean free path of the EM beam in a Li-Baker HFGW Detector

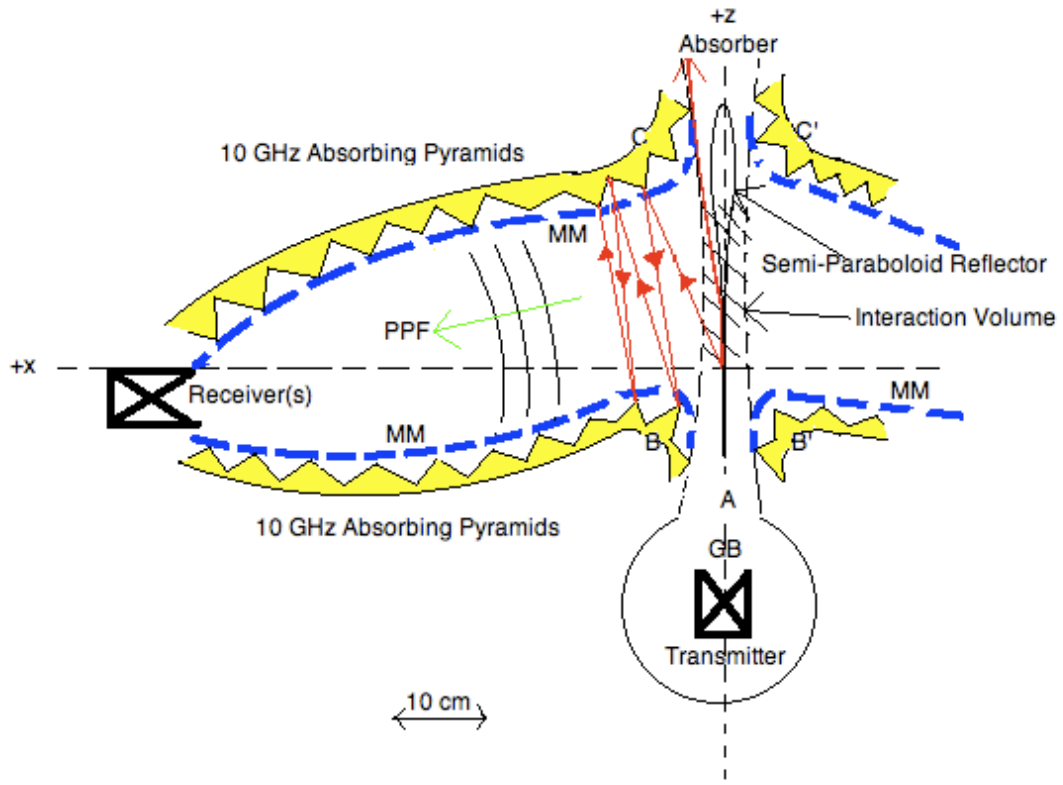
- $l_{ray} \approx 1/(n\sigma_{ray}) = 1/([N_m P//T][\sigma_{ray}(H_2)] = 1/([1.51\text{x}10^{13}][1.047\text{x}10^{-46}]) = 6\text{x}10^{32}$ cm
- $I = I_0 e^{-z/lray}$,
- I_0 is the initial intensity of the GB = $1.51\text{x}10^{26}\text{s}^{-1}$ photons per second
- $I_{(10)} - I_{(40)} = I_0 (e^{-10/lray} - e^{-40/lray}) \approx (1.51\text{x}10^{26})(-1 + 10/6\text{x}10^{32} + 1 - 40/6\text{X}10^{32}) = 3\text{x}10^{-7}$ photons per second scattered in the 30 cm long GB interaction volume, **which is negligible**.

Schematic of Li-Baker HFGW Detector (Peoples Republic of China Patent Number 0510055882.2)

GRAVWAVE ® LLC.



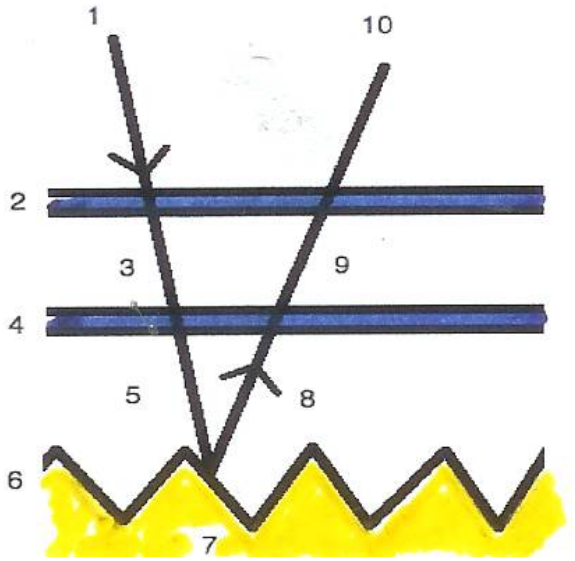
Side-view Schematic of the Li-Baker HFGW Detector Exhibiting Microwave Absorbent Walls in the Anechoic Chamber .



Schematic of the multilayer metamaterial or *MM* absorbers and pyramid absorber/reflector (-45 dB per layer). Patent Pending

GRAVWAVE ® LLC.

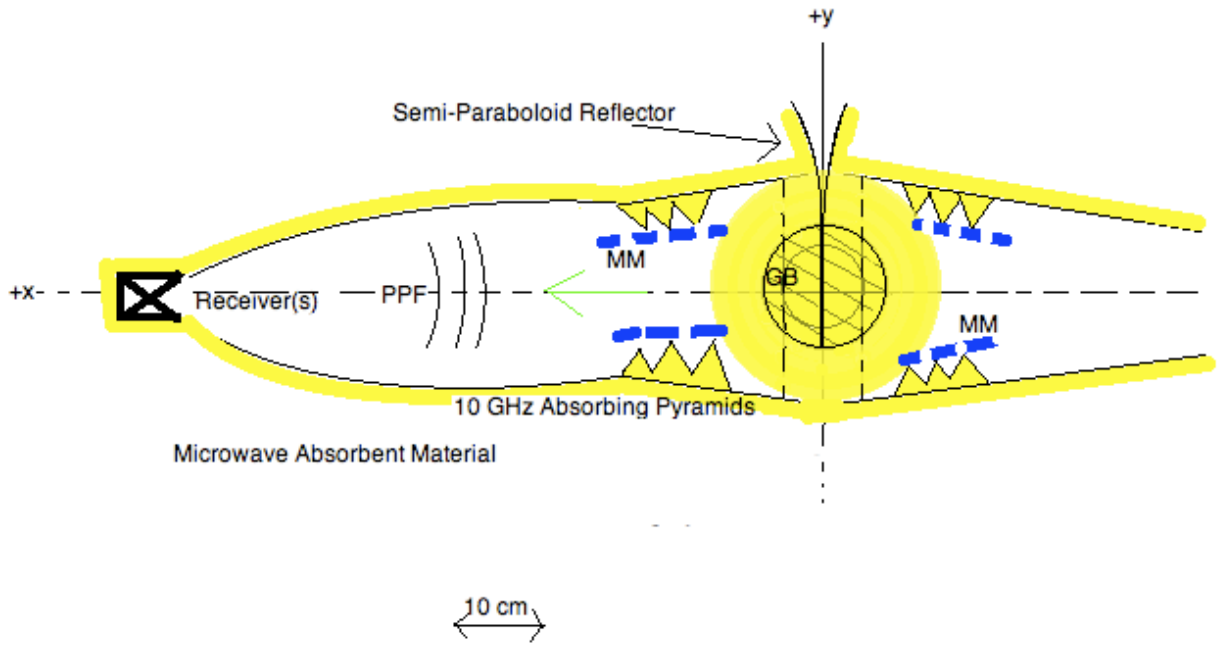
1 Incident
2 1st metamaterial (MM) layer
3 transmitted
4 typical MM layer
6 conventional
microwave absorber
8 reflected
10 remaining



The incident ray can have almost any inclination: Service (2010)

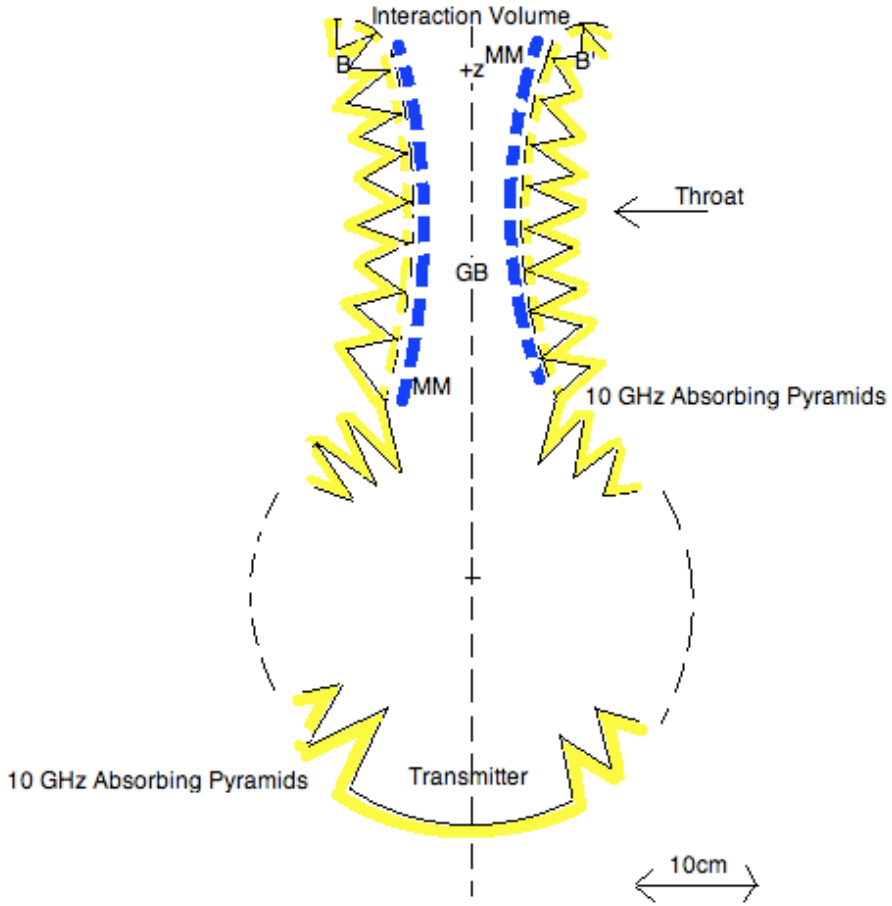
Plan-view Schematic of the Li-Baker HFGW Detector Exhibiting Microwave Absorbent Walls in the Anechoic Chamber

GRAVWAVE ® LLC.



Gaussian-Beam Transmitter Compartment (Patents Pending).

GRAVWAVE ® LLC.



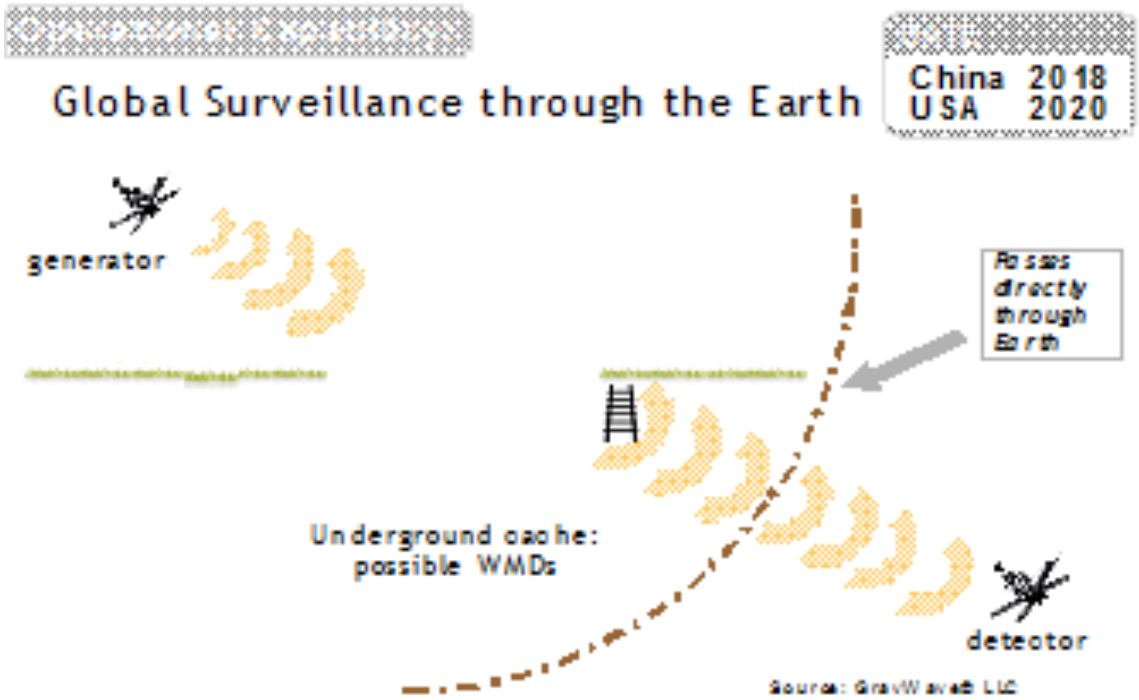
Required Detector-Wall Absorption

- Power db =10 $\log_{10}$ (power out/power in) = 10$\log_{10}$ (6.626x10-27/1000) = -290 db

- Absorption “mat” consisting of four MM layers, each layer a quarter wavelength from the next (in order to cancel any possible surface reflection) and provide a - 45 db -45 db - 45 db = -135 db absorption (Landy, et al., 2008 indicate **no reflection -- only transmission**). Behind these MM layers would be a sheet of 10 GHz microwave pyramid absorbers providing a -40 db absorption before reflection back into the four MM layers. Thus the total absorption would be -135 db -40 db –135db = -310 db.

HFGW Surveillance

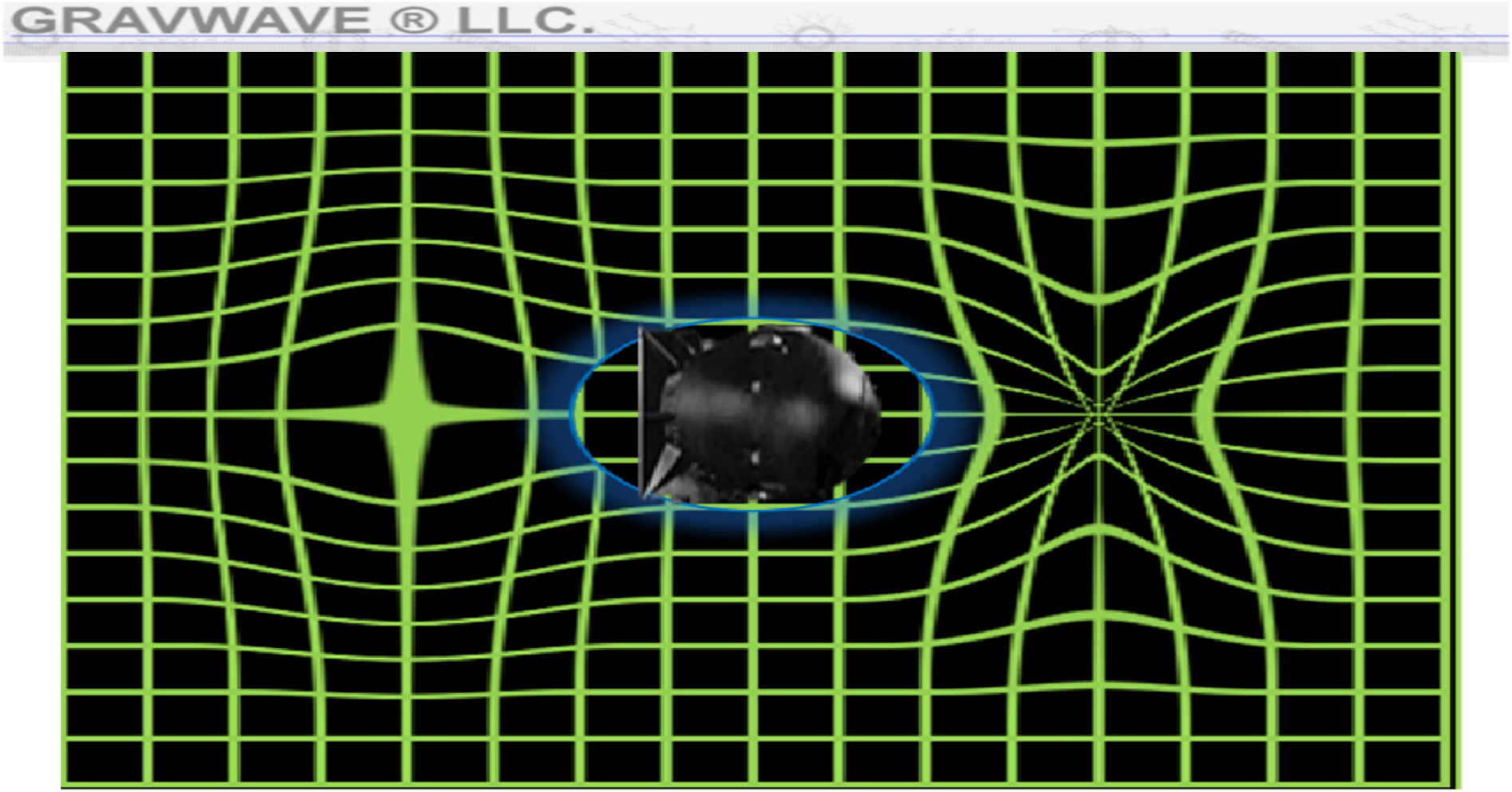
GRAVWAVE ® LLC.



Theoretical Surveillance Application

- Misner, Thorne and Wheeler (1973): “In the real universe there are spacetime curvatures due not only to the energy of gravitational waves, but also more importantly to the material [objects and structures] content of the universe ... its wavelength changes [based on gravitational red shift] and [the gravitational wave] backscatters off the curvature to some extent. If the wave is a pulse, then the backscatter will (change) its shape and *polarization*....” It is extremely difficult to theoretically establish the actual magnitude of the changes, especially at very high frequencies (10^9 Hz and higher) and to quantify them prior to the proof-of-concept HFGW generation/ detection laboratory experiments.

REMOTE DISPLACEMENT OF MASSES



Section 108, page 349 of Landau and Lifshitz (1975)

GRAVWAVE ® LLC.

- "Since it has definite energy, the gravitational wave is itself is the source of some additional gravitational field [static g-field]. Like the energy producing it, this field is a second-order effect in the h_{ik}. ***But in the case of high-frequency gravitational waves the effect is significantly strengthened*****:** the fact that the pseudotensor t_{ik} is quadratic in the derivatives of the h_{ik} introduces the large factor λ^{-2}. In such a case we may say that the wave itself produces the background field [static g-field] on which it propagates. This [static g] field is conveniently treated by carrying out the averaging described above over regions of four-space with dimensions large compared to λ. Such an averaging, smoothes out the short-wave 'ripple' and leaves the slowly varying background metric [static g-field]." (Brackets and italics added for clarity and emphasis.)

Extremely Theoretical Calculations

- Landau and Lifshitz (1975) offer no elaboration of the physics and mathematics behind their assertion; but their textbook reference certainly is based upon some of their analyses or could be derived. In any event, the concept is clear.

- Using the standard astrodynamics equations found, for example, on pages 90 and 91 of Herrick (1971), a computer program yields from a 26.8 to a 2.7 mile perturbative g-field change in missile entry location for 6,200 mile ICBM trajectories (*with 50 to 100 mile length, 0.1 to 0.01 g-field perturbations*). For short-range 1,400 mile trajectories, it yields from a 2.0 to a 0.41 mile perturbative g-field change in missile entry location (*with 25 to 50 mile length, 0.1 to 0.01 g-field perturbations*). Such modest changes would not greatly reduce the damage caused by terrorist ICBM nuclear strikes, but would frustrate anti-missile systems (by perturbing their trajectories) or defend against, for example, surgical strikes against submerged submarine assets.

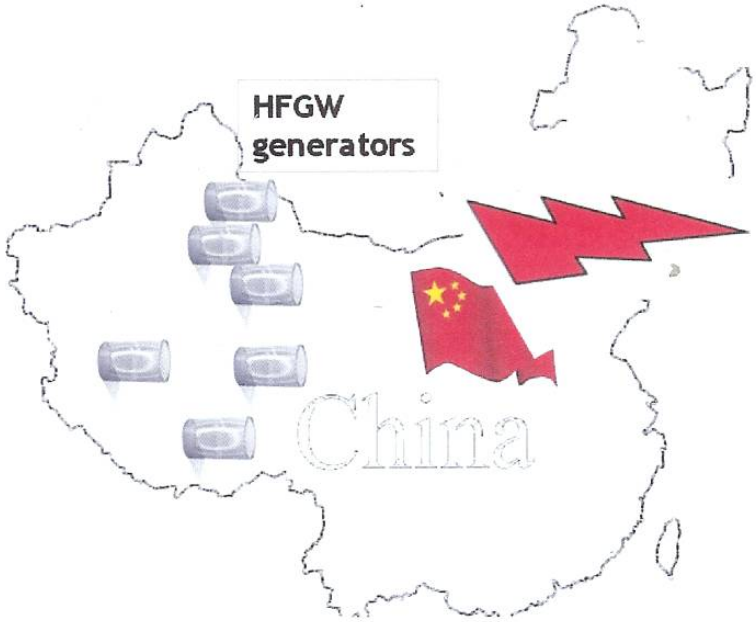
Remote HFGW-Induced Nuclear Fusion

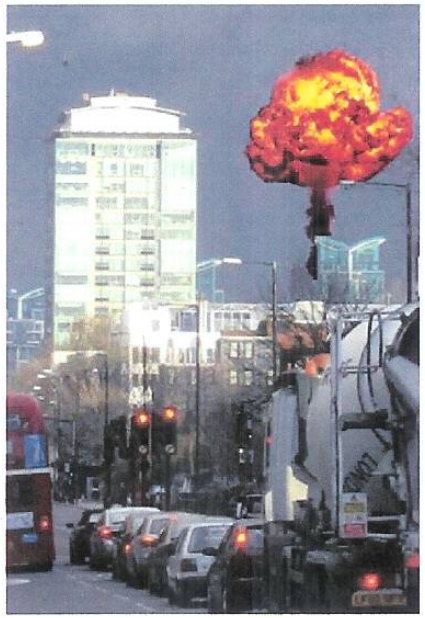
GRAVWAVE ® LLC.

Operational Capability:

Remote Nuclear Fusion

DATE
China 2029
USA 2031



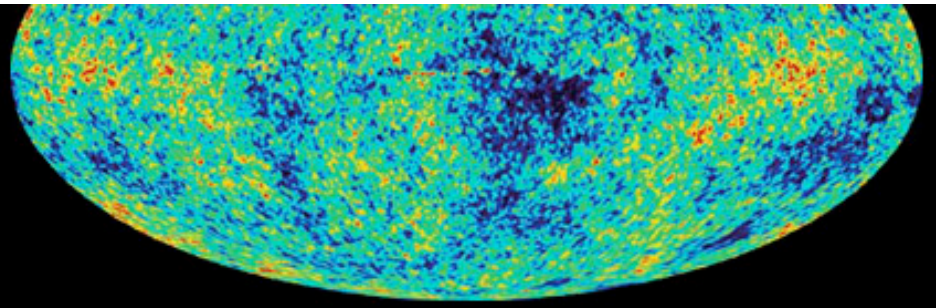


Conclusions

GRAVWAVE ® LLC.

- High-frequency gravitational wave (HFGW) generators have been proposed theoretically by the Russians, Germans, Italians and Chinese. HFGW detectors are a reality and three have been actually constructed outside the United States by the British, Italians and Japanese.
- A theoretically more sensitive detector than these, the Li-Baker, utilizing metamaterial and off-the-shelf microwave absorbers to eliminate noise, together with a theoretical, multi-FBAR HFGW generator in a double-helix configuration that are discussed, could be utilized for transglobal, low-probability of intercept (PPI) communications. The multi-elements of the transmitter (HFGW generator) are off-the-shelf piezoelectric film-bulk acoustic resonators or FBARs energized by off-the-shelf' modified Magnetrons. In theory a large number of these FBAR elements could lead to HFGW generator-detector communications for a laboratory proof-of-concept experiment.
- Pending the recommended proof-of-concept HFGW experiment other HFGW applications could be of value. These theoretical applications, yet to be quantified, but discussed herein, include surveillance and remote displacement of masses such as missiles and anti-missiles and the remote generation of nuclear fission.

CMB and High Energy Physics
from Early Universe to clusters
of galaxies and Large Scale Structure
La Palma, July 16-24, 2012



Primordial Gravitational waves and the polarization of the CMB

José Alberto Rubiño Martín

(IAC, Tenerife)

Outline

Lecture 1.

- Theory of CMB polarization.
  - E and B modes.
  - Primordial Gravitational waves.
- Observational status: CMB polarization measurements. E-mode detections and B-mode constraints.

Lecture 2.

- Foregrounds of the CMB.
- The future (CMB polarization experiments).
  - The Planck satellite.
  - QUIJOTE CMB experiment.
  - Core satellite.

Lecture 1. Bibliography

Books:

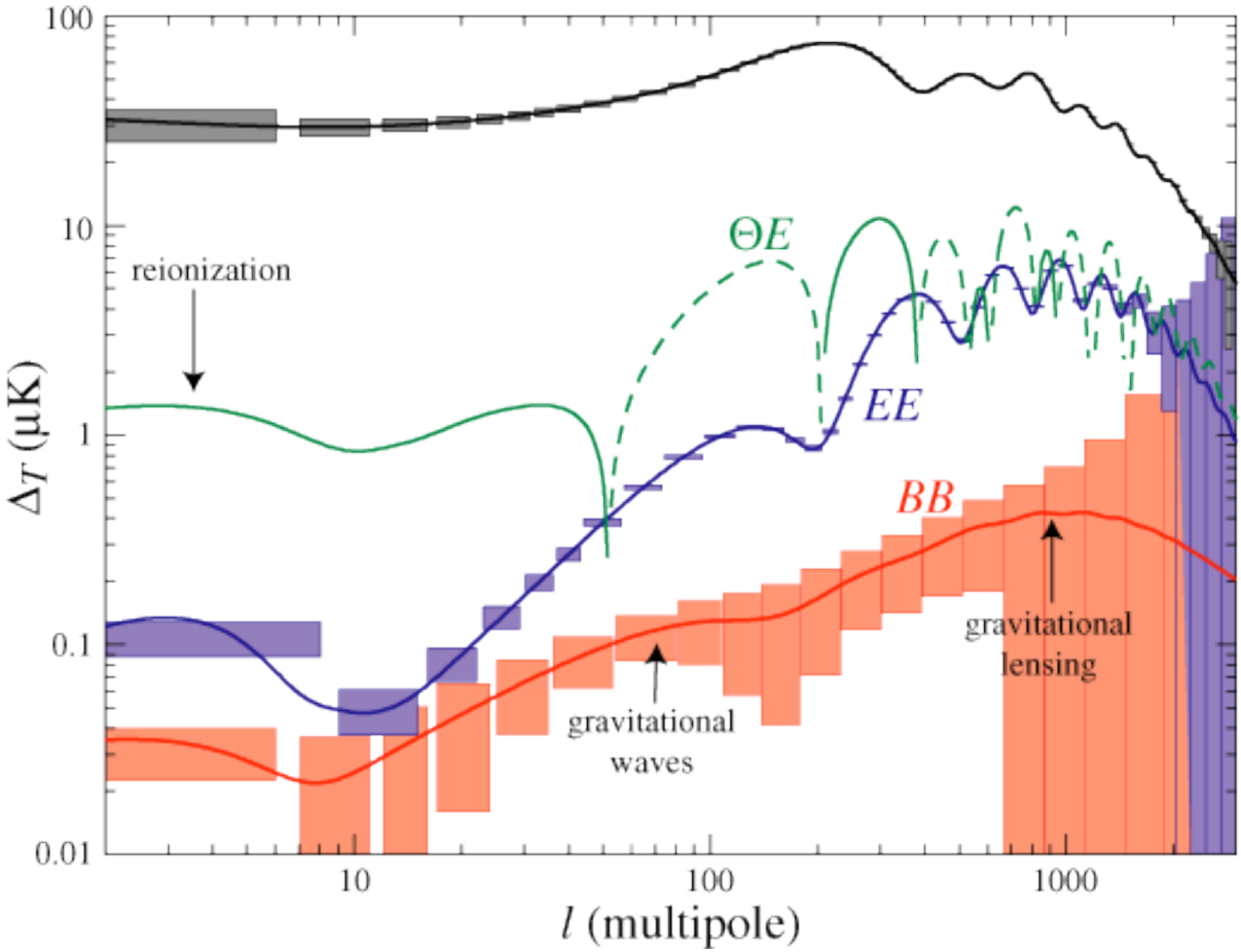
- ***Cosmological Inflation and Large Scale Structure.*** Liddle & Lyth (CUP 2000).
- ***Cosmology.*** S Weinberg (Oxford Univ. Press 2008)***.***
- ***The Cosmic Microwave Bakcground.*** Durrer (CUP 2008).
- ***The primordial density perturbation.*** Lyth & Liddle (CUP 2009)
- **The Cosmic Microwave Background: from quantum fluctuations to the present Universe.** Eds. Rubiño-Martin, Rebolo, Mediavilla (Cambridge Univ. Press 2010).

Papers (Polarization of the CMB):

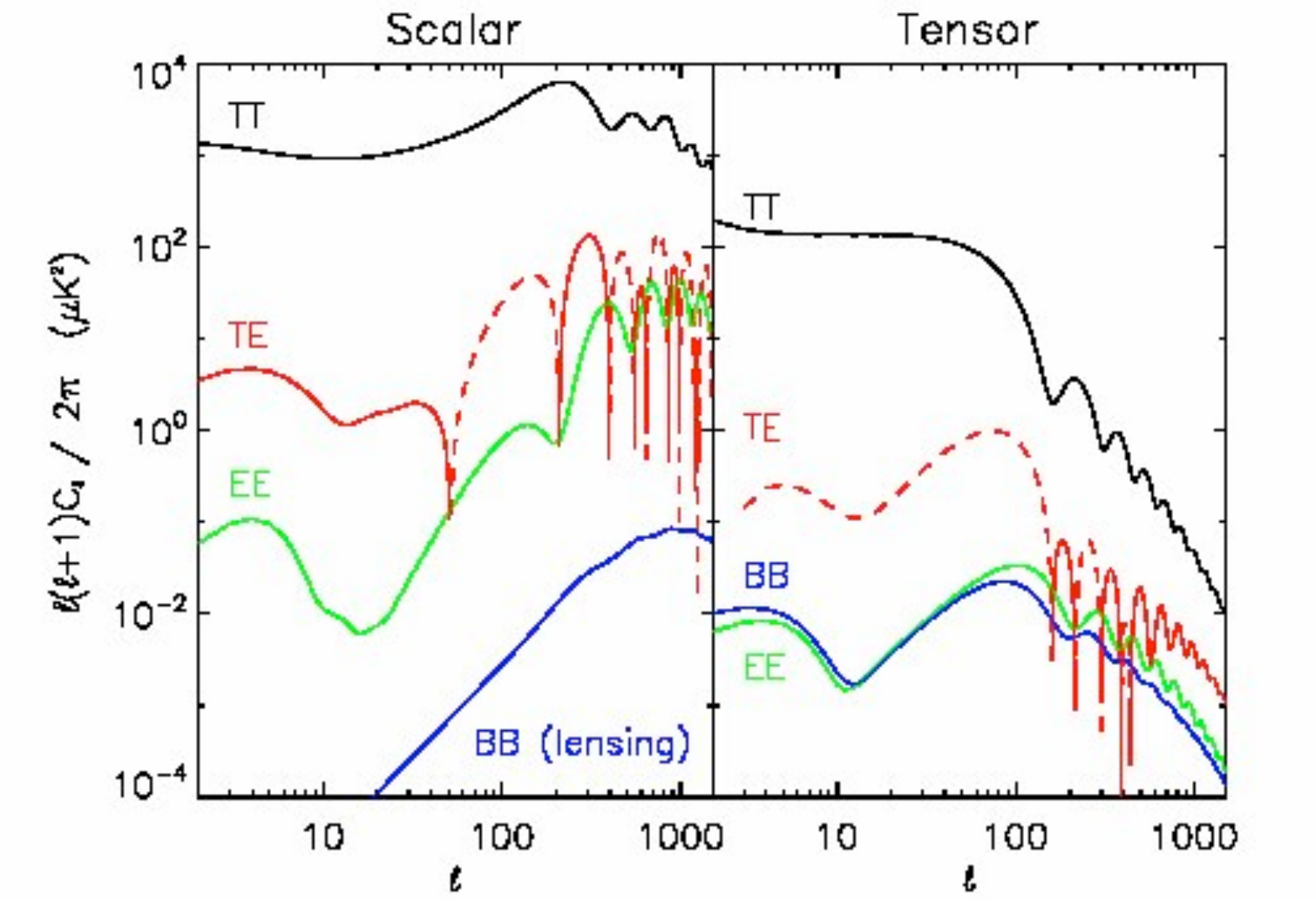
- Ma & Bertschinger (1995), ApJ, 455, 7.
- Hu & Sugiyama (1995), Phys. Rev. D 51, 2599
- Seljak & Zaldarriaga (1996), ApJ, 469, 437.
- Zaldarriaga & Seljak (1997), Phys. Rev. D. 55, 1830.
- Kamionkowski, Kosowsky & Stebbins (1997), Phys. Rev. D 55, 7368.

Remember A. Lasenby´s talk

Physics of the CMB anisotropies: the angular power spectrum



Power spectra - Theory



Stokes parameters

❖The polarization state of a electromagnetic wave propagating in a direction **n** can described by the intensity matrix

$$P_{ij} = \left\langle E_i(\hat{n}) E_j^*(\hat{n}) \right\rangle$$

where **E** is the electric field vector and the brackets denote time averaging.

❖ P is a hermitian 2x2 matrix and thus can be decomposed into the Pauli basis

$$P = I(\hat{n})\sigma_0 + Q(\hat{n})\sigma_3 + U(\hat{n})\sigma_1 + V(\hat{n})\sigma_2$$

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

❖ In terms of the electric field, the Stokes parameters are defined as:

$$I = |E_1|^2 + |E_2|^2 \qquad Q = |E_1|^2 - |E_2|^2$$

$$U = (E_1^* E_2 + E_2^* E_1) = 2\mathrm{Re}(E_1^* E_2) \qquad V = 2\mathrm{Im}(E_1^* E_2)$$

Stokes parameters

❖ Linear polarization is described by Stokes Q and U parameters.

❖ Q measures the different of intensities in the to axes x and y, while U measures the difference of intensities in a coordinate system at 45º.

❖ Not all Stokes parameters are rotationally invariant. Under a rotation of ψ degrees of the coordinate system, we have

$$I' = I \qquad V' = V$$

$$Q' = Q\cos(2\psi) - U\sin(2\psi) \qquad U' = U\cos(2\psi) + Q\sin(2\psi)$$

or in a more compact form

$$Q' \pm iU' = e^{\pm 2i\psi}(Q \pm iU)$$

❖Hence, (Q±iU) transforms like a spin-2 variable under rotations.

Spin weighted spherical harmonics

- Spin-s spherical harmonics.
- Under rotations, they transform as: $${}_sY_{\ell m} \rightarrow e^{\pm si\psi}\, {}_sY_{\ell m}(\hat{n})$$
- Orthogonality and completeness

$$\int d\hat{\mathbf{n}}\, {}_sY^*_{\ell m}(\hat{\mathbf{n}})\, {}_sY_{\ell m}(\hat{\mathbf{n}}) = \delta_{\ell\ell'}\delta_{mm'}$$

$$\sum_{\ell m} {}_sY^*_{\ell m}(\hat{\mathbf{n}})\, {}_sY_{\ell m}(\hat{\mathbf{n}}') = \delta(\phi-\phi')\delta(\cos\theta-\cos\theta')$$

- Relation to Wigner rotation matrices:

$${}_sY_{\ell m}(\theta,\phi) = (-1)^m \sqrt{\frac{2\ell+1}{4\pi}} e^{-is\psi} D^{\ell}_{-m,s}(\phi,\theta,-\psi)$$

Statistical representation

❖ All-sky decomposition:

$$(Q \pm iU)(\hat{n}) = \sum_{\ell=2}^{\infty} \sum_{m=-\ell}^{+\ell} a_{\ell m}^{\pm 2} \; {}_{\pm 2}Y_{\ell m}(\hat{n}) = \sum_{\ell=2}^{\infty} \sum_{m=-\ell}^{+\ell} (a_{E,\ell m} \pm i a_{B,\ell m}) \; {}_{\pm 2}Y_{\ell m}(\hat{n})$$

❖ Here, $a_{lm}^{\pm 2}$ is a decomposition into positive and negative helicity. The helicity basis

$$e^{\pm} = \frac{1}{\sqrt{2}}(e_\theta \pm i e_\phi)$$

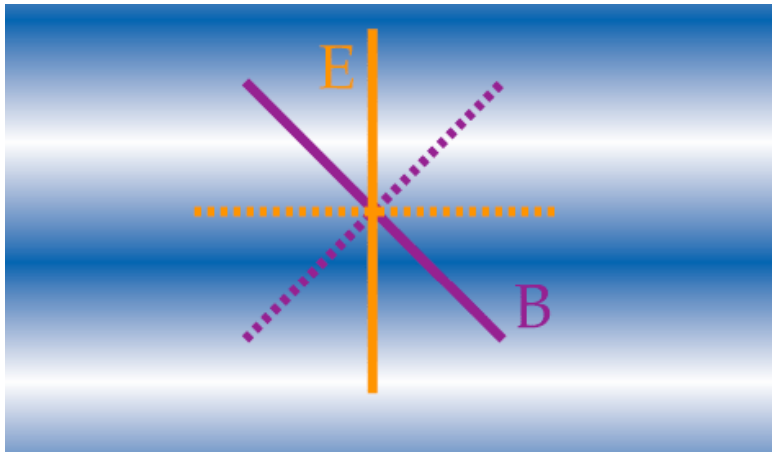
❖ In the last equality we have defined E- and B-modes:

$$a_{E,\ell m} = \frac{1}{2}(a_{\ell m}^{+2} + a_{\ell m}^{-2})$$

$$a_{B,\ell m} = \frac{-i}{2}(a_{\ell m}^{+2} - a_{\ell m}^{-2})$$

❖ Under parity transformations (**n→-n**), the E-modes remain invariant, while B-modes change sign.

E and B modes



A plane wave moving from top to bottom. The direction of the polarization vector defines if they are E or B modes.

- **Full-sky** polarization maps can be decomposed into two components usually called **E-modes** (analog of the gradient component) and **B-modes** (analog of the curl component) (see Kamionkowski et al. 1997; Seljak & Zaldarriaga 1997).
- These modes are independent on the coordinate system, and are related to the Q and U Stokes parameters by a non-local transformation.

(A pure E-mode turns into pure B-mode if we turn all polarization vectors by 45º).

Angular power spectra

$$C_{Tl}=\frac{1}{2l+1}\sum_m \langle a^*_{T,lm}a_{T,lm}\rangle,$$

$$C_{El}=\frac{1}{2l+1}\sum_m \langle a^*_{E,lm}a_{E,lm}\rangle,$$

$$C_{Bl}=\frac{1}{2l+1}\sum_m \langle a^*_{B,lm}a_{B,lm}\rangle,$$

$$C_{Cl}=\frac{1}{2l+1}\sum_m \langle a^*_{T,lm}a_{E,lm}\rangle,$$

in terms of which,

$$\langle a^*_{T,l'm'}a_{T,lm}\rangle = C_{Tl}\delta_{l'l}\delta_{m'm},$$

$$\langle a^*_{E,l'm'}a_{E,lm}\rangle = C_{El}\delta_{l'l}\delta_{m'm},$$

$$\langle a^*_{B,l'm'}a_{B,lm}\rangle = C_{Bl}\delta_{l'l}\delta_{m'm},$$

$$\langle a^*_{T,l'm'}a_{E,lm}\rangle = C_{Cl}\delta_{l'l}\delta_{m'm},$$

$$\langle a^*_{B,l'm'}a_{E,lm}\rangle = \langle a^*_{B,l'm'}a_{T,lm}\rangle = 0.$$

The polarization of the CMB anisotropies

- **Four parity-independent power spectra can be formed:**

$$C_{TT} = \frac{1}{2l+1} \sum_m \left\langle a^*_{T,lm} a_{T,lm} \right\rangle \qquad C_{BB} = \frac{1}{2l+1} \sum_m \left\langle a^*_{B,lm} a_{B,lm} \right\rangle$$

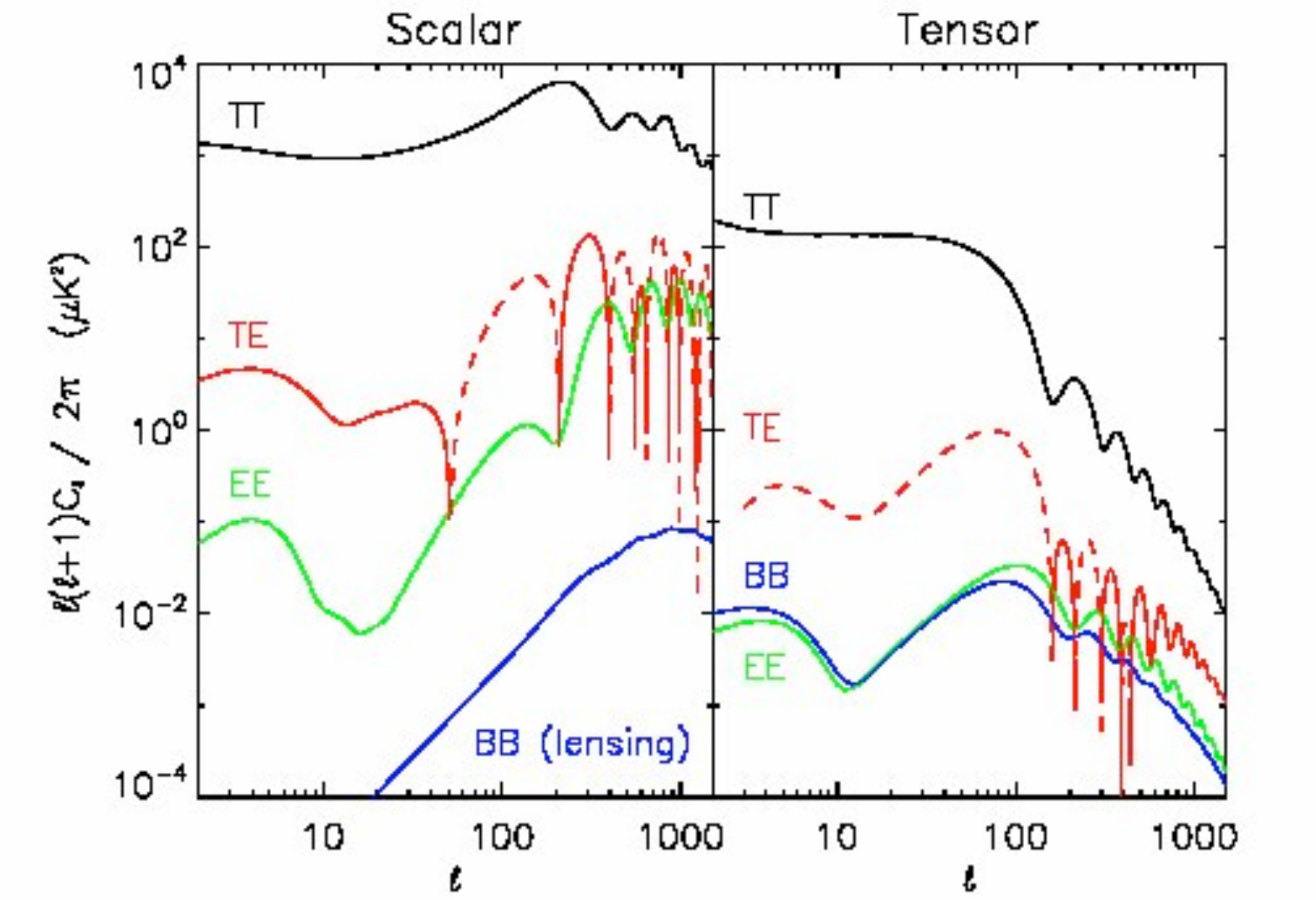
$$C_{EE} = \frac{1}{2l+1} \sum_m \left\langle a^*_{E,lm} a_{E,lm} \right\rangle \qquad C_{TE} = \frac{1}{2l+1} \sum_m \left\langle a^*_{T,lm} a_{E,lm} \right\rangle$$

- **Physics of generation of the Polarization.** Different sources of anisotropies generate different types of modes:

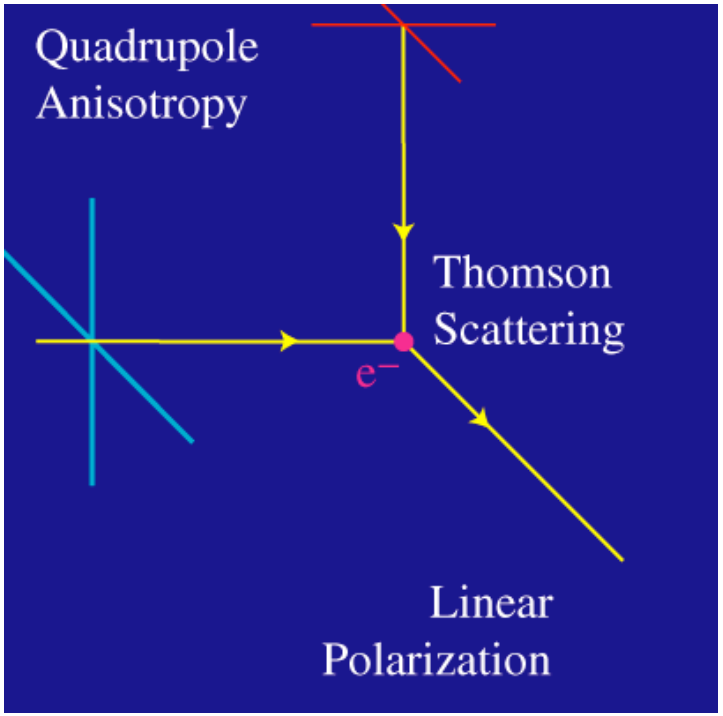
	Scalar (density perturbations)	Tensor (gravitational waves)
E-modes	Yes	Yes
B-modes	No	Yes

- **B-modes probe the existence of primordial gravitational waves.**

Power spectra - Theory



Thomson scattering



❖Differential cross-section

$$\frac{d\sigma}{d\Omega} = \frac{3}{8\pi}\left|\hat{\varepsilon}' \cdot \hat{\varepsilon}\right|^2 \sigma_T$$

❖A net polarization is generated during recombination if there is a quadrupole anisotropy in the radiation field.

❖ Defining the y-axes of the incoming and outgoing coordinate systems to be in the scattering plane, the Stokes parameters of the outgoing beam, defined with respect to the x-axis, we have:

$$I = \frac{3\sigma_T}{8\pi R^2} I'(1+\cos^2(\beta)) \qquad Q = \frac{3\sigma_T}{8\pi R^2} I' \sin^2(\beta) \qquad U = V = 0$$

❖ The net polarization produced by the scattering of an incoming, unpolarized ratiation field of intensity I′(θ,ϕ) is determined by integrating over all incoming directions.

$$I(\hat{z}) = \frac{3\sigma_T}{8\pi R^2} \int d\Omega\, I'(\theta,\phi)(1+\cos^2(\theta))$$

$$Q(\hat{z}) - iU(\hat{z}) = \frac{3\sigma_T}{8\pi R^2} \int d\Omega\, \sin^2(\theta) e^{i2\phi} I'(\theta,\phi)$$

Only a_{2m} components contribute.

Polarization: scalar perturbations

❖Breakdown of tight-coupling leads to a quadrupole anisotropy.

❖ For scalar perturbations, the polarization signal arises from the gradient of the peculiar velocity of the photon fluid (e.g. Zaldarriaga & Harari 1995).

❖ The basic argument is:

❖ Consider a scattering ocurring at x_0. The mean free path is λ_T. Photons coming from a direction **n** roughly come from

$$x = x_0 + \lambda_T \hat{n}$$

❖ The photon-baryon fluid at that point was moving with velocity

$$v(x) = v(x_0) + \lambda_T \hat{n}_i \partial_i v(x_0)$$

❖ Due to Doppler effect, the temperature by the scatterer at x_0 is

$$\delta T(x_0, \hat{n}) = \hat{n} \cdot [v(x) - v(x_0)] \approx \lambda_T \hat{n}_i \hat{n}_j \partial_i v_j(x_0)$$

Polarization: scalar perturbations

❖Breakdown of tight-coupling leads to a quadrupole anisotropy.

❖ For scalar perturbations, the polarization signal arises from the gradient of the peculiar velocity of the photon fluid (e.g. Zaldarriaga & Harari 1995).

❖ Gradient of the velocity is along the direction of the wavevector, so the polarization is purely E-mode:

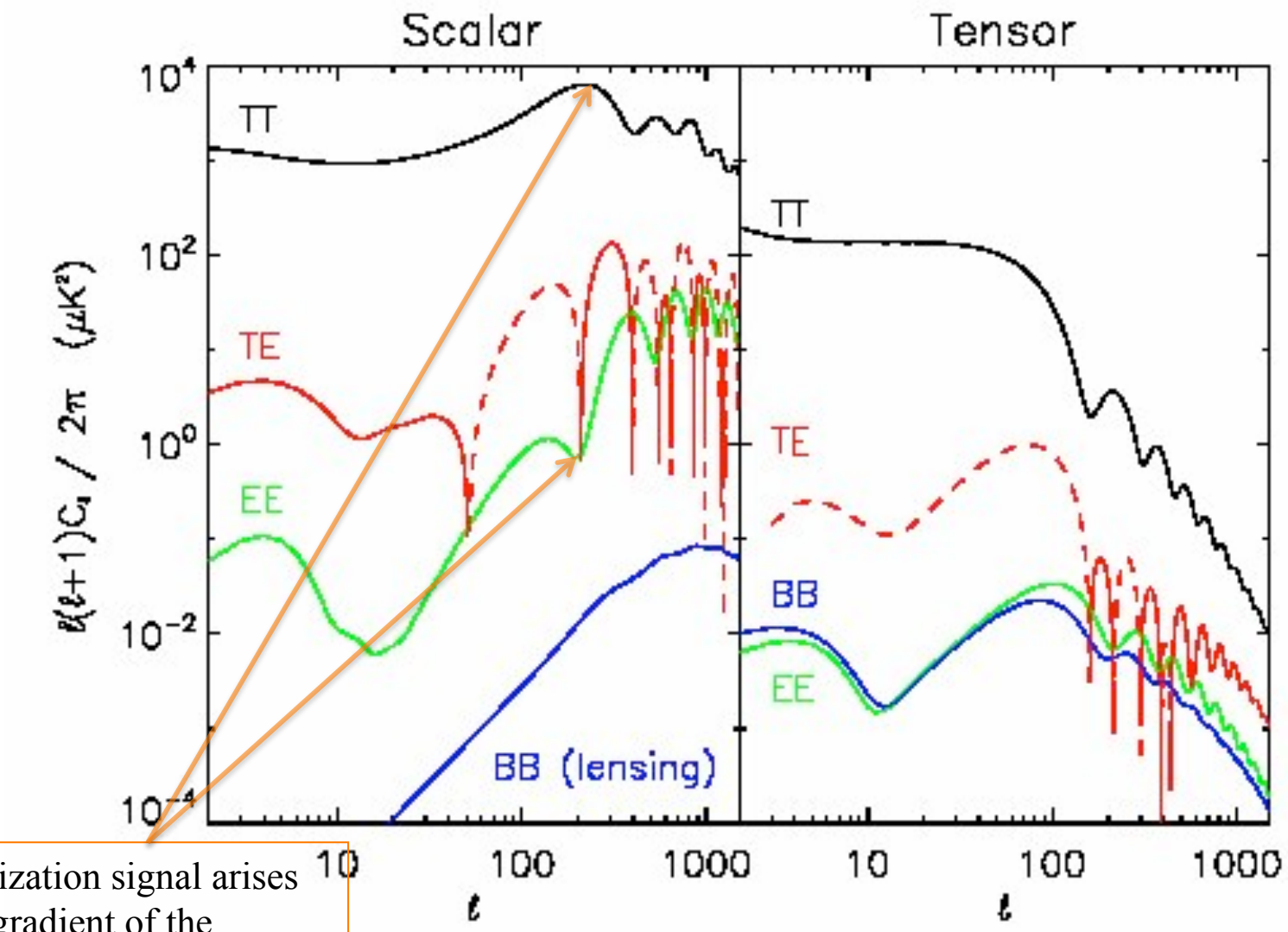
$$\Delta_E \approx -0.17(1-\mu^2)\Delta\eta_{dec} k\, v_\gamma(\eta_{dec})$$

❖ Velocity is 90º out of phase with respect to temperature – turning points of oscillator are zero points of velocity:

$$\Delta_T \propto \cos(kr_s) \qquad v_\gamma \propto \sin(kr_s)$$

❖ Polarization peaks are at troughs of temperature peaks.

Power spectra - Theory



The polarization signal arises from the gradient of the peculiar velocity of the photon fluid => TT and EE peaks are out of phase.

TE Polarization and acoustic peaks

❖Cross-correlation of temperature and polarization

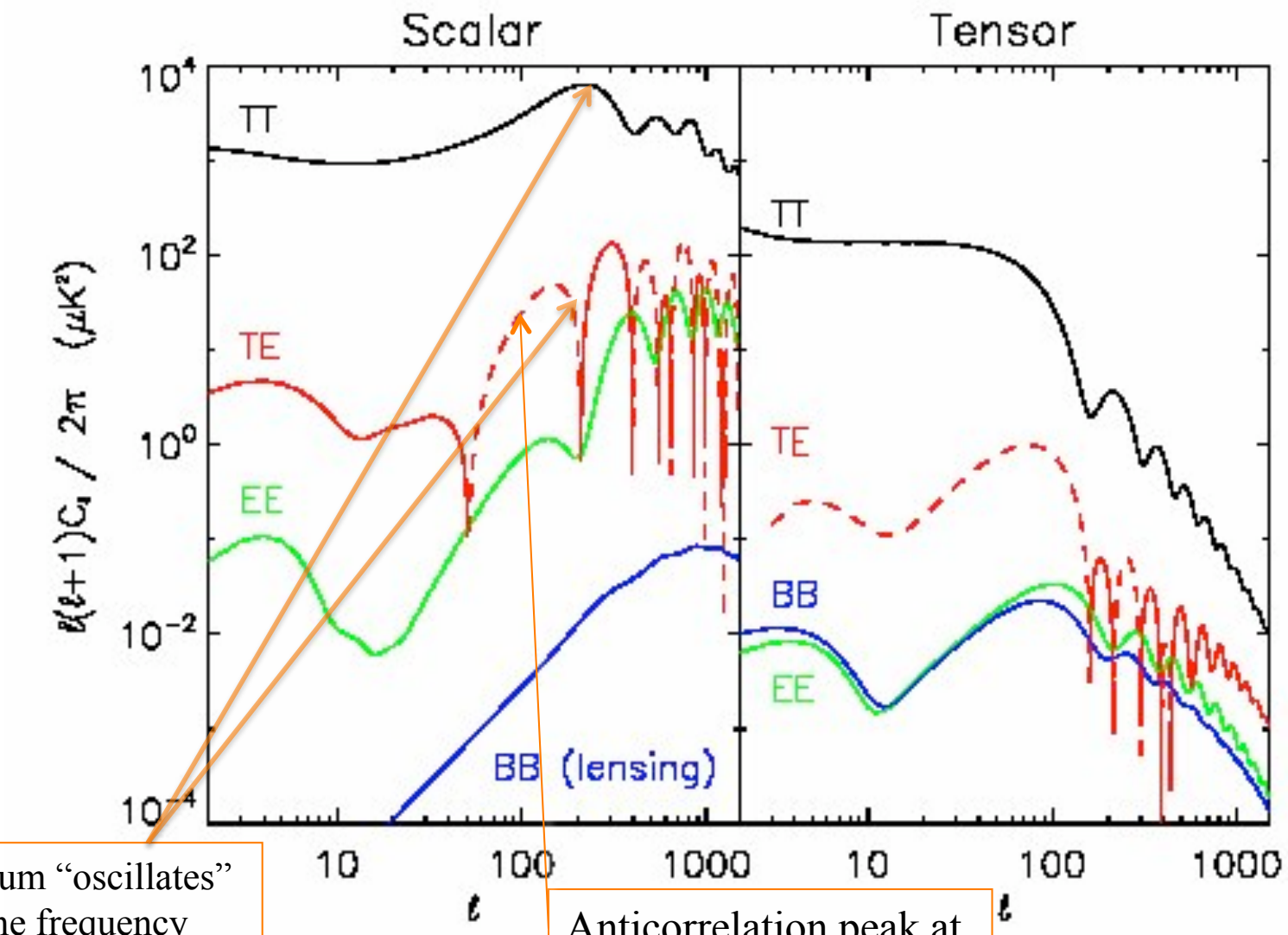
$$\Delta_T \Delta_E \propto \cos(kr_s)\sin(kr_s) \propto \sin(2kr_s)$$

❖ TE spectrum “oscillates” at twice the frequency

❖ TE correlation is radial or tangential around hot spots (see later, WMAP result).

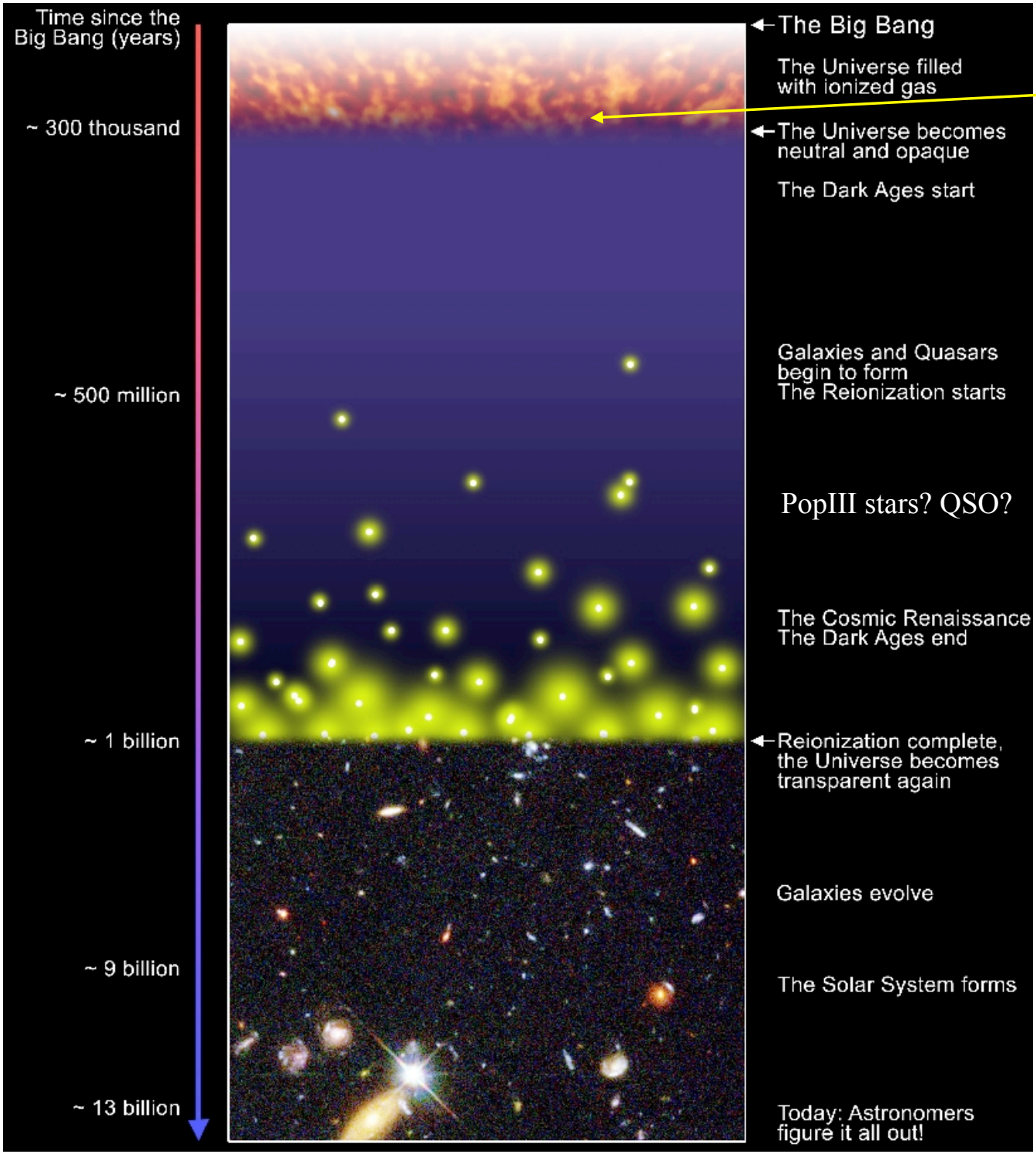
❖ Large scales: anticorrelation peak around l=150, a distinctive signature of primordial adiabatic fluctuations (Peiris et al. 2003).

Power spectra - Theory



TE spectrum "oscillates" at twice the frequency

Anticorrelation peak at l=150 is a signature of superhorizon adiabatic fluctuations



Recombination (z ~ 1000)
(Microwave background)

Dark Ages

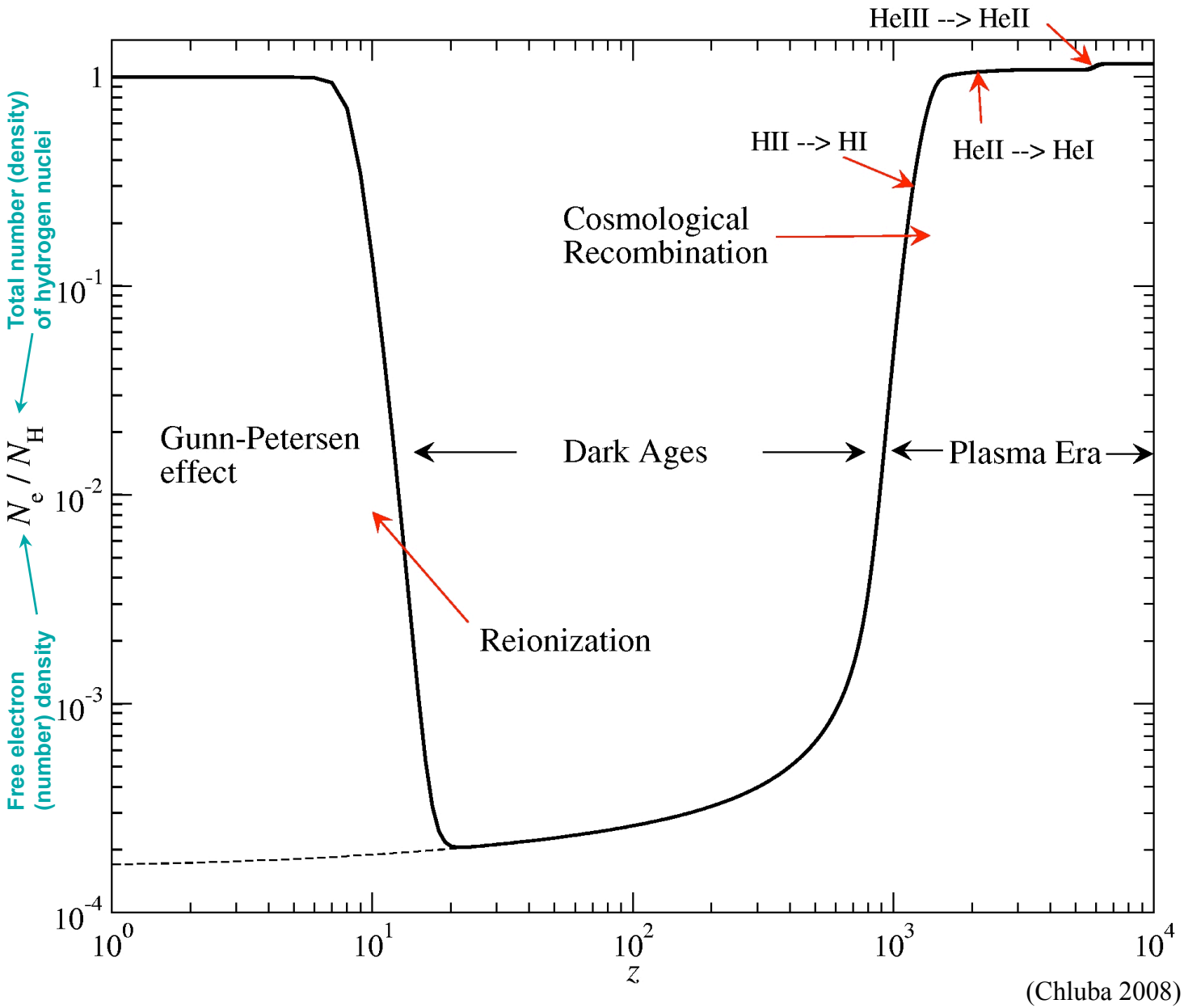
(Neutral hydrogen. Accesible with 21cm line surveys as SKA or LOFAR)

z~6
(GP trough observed in QSO)

Galaxies form and evolve

(Universe accesible with optical/IR telescopes)

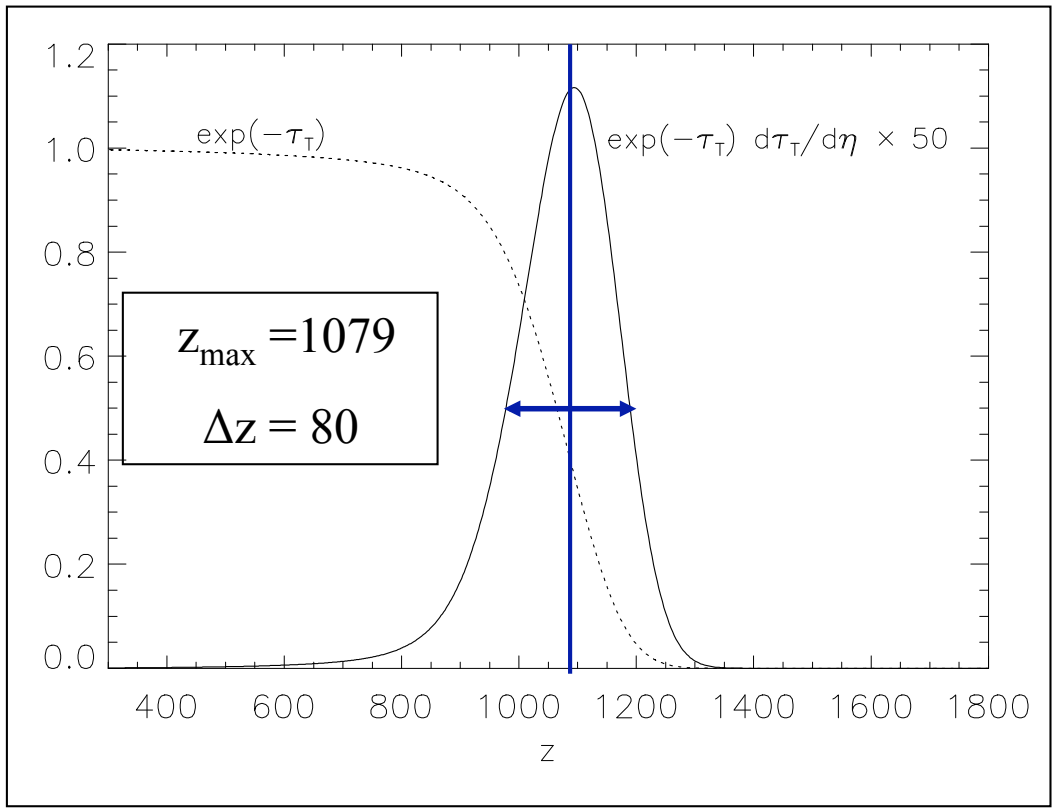
How did the Universe became neutral? Sketch of the Ionization History



(Chluba 2008)

The Thomsom visibility function

Prior to the recombination epoch, the photons and the electrons are tighly coupled due to Thomson scattering. When the number density of electrons decreases, the photons are released and the CMB is formed.



- The relevant quantity is the Thomson optical depth.

$$\tau = \int \sigma_T\, n_e\, dl$$

- The **visibility function** gives the probability that a photon observed now was last scattered at redshift z:

$$V = \exp(-\tau)\, d\tau/d\eta$$

Reionization

❖Optical depth to Thomson scattering to reionization

$$\tau(z) = \int d\eta n_e \sigma_T a = \int d\ln a \frac{n_e \sigma_T}{H(a)} \propto (\Omega_b h^2)(\Omega_m h^2)^{-1/2}(1+z)^{3/2}$$

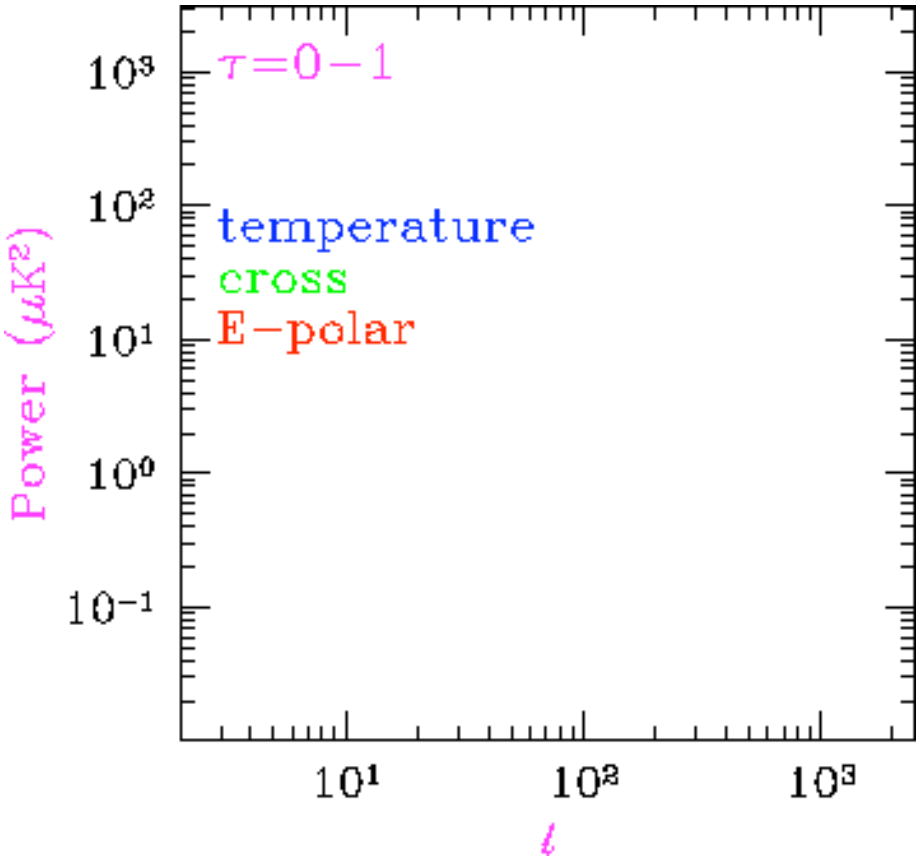
$$= \left(\frac{\Omega_b h^2}{0.02}\right)\left(\frac{\Omega_m h^2}{0.15}\right)^{-1/2}\left(\frac{1+z}{61}\right)^{3/2}$$

❖ CMB re-scatters off re-ionized gas. Generation of new anisotropies at large scales (Doppler) and absorption at small scales:

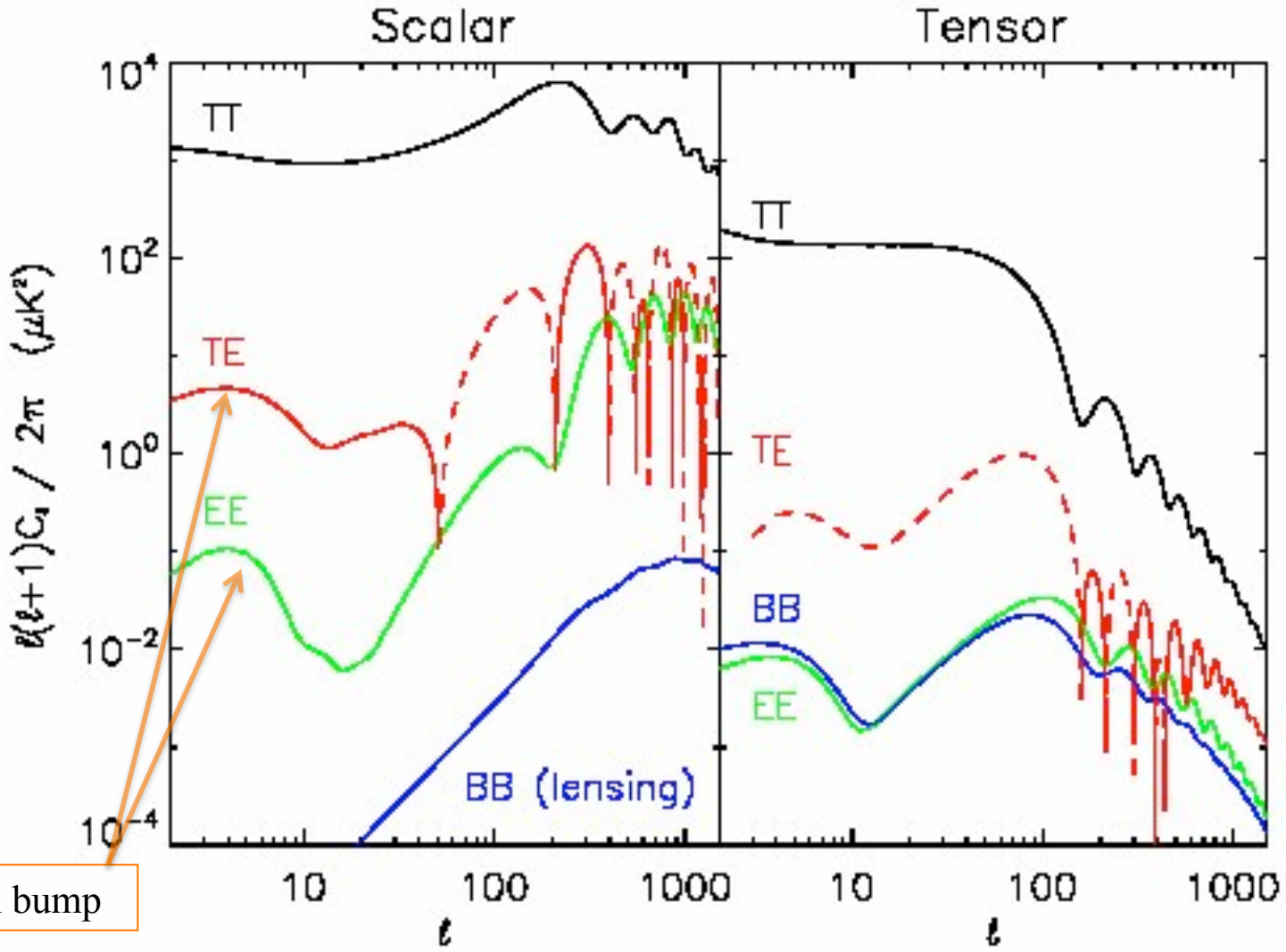
$$\Delta_T \to \Delta_T e^{-\tau} \qquad \Delta_E \to \Delta_E \tau$$

$$C_\ell^{TE} \propto \tau e^{-\tau} \qquad C_\ell^{EE} \propto \tau^2$$

❖ Effect peaks at horizon scale at recombination (l≈2-3). If the optical depth is very large, primordial anisotropies are erased.



Power spectra - Theory



Reionization bump

Gravitational lensing

❖ Large scale density fluctuations in the Universe induce random deflections in the direction of CMB photons as they propagate from last scattering surface to us. In the small scale limit, we have

$$Q(\boldsymbol{\theta}) = \tilde{Q}(\boldsymbol{\theta} + \delta\boldsymbol{\theta}) = (2\pi)^{-2} \int d^2\boldsymbol{l} e^{i\boldsymbol{l}\cdot(\boldsymbol{\theta}+\delta\boldsymbol{\theta})} \times [E(\boldsymbol{l})\cos(2\phi_l) - B(\boldsymbol{l})\sin(2\phi_l)],$$

$$U(\boldsymbol{\theta}) = \tilde{U}(\boldsymbol{\theta} + \delta\boldsymbol{\theta}) = (2\pi)^{-2} \int d^2\boldsymbol{l} e^{i\boldsymbol{l}\cdot(\boldsymbol{\theta}+\delta\boldsymbol{\theta})} \times [E(\boldsymbol{l})\sin(2\phi_l) + B(\boldsymbol{l})\cos(2\phi_l)].$$

(Zaldarriaga & Seljak 1998)

Gravitational lensing

❖ Zaldarriaga & Seljak (1998), PRD 58, 23003.

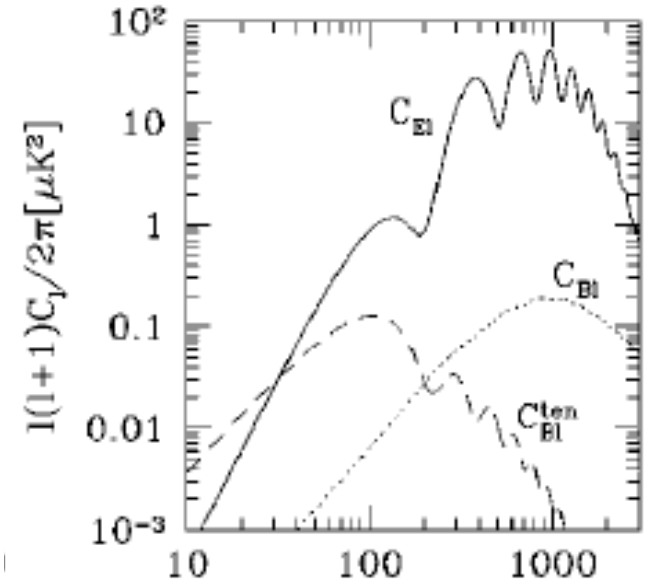
$$C_{Tl} = C_{\tilde{T}l} + \mathcal{W}^{l'}_{1l} C_{\tilde{T}l'} ,$$

$$C_{El} = C_{\tilde{E}l} + \frac{1}{2}[\mathcal{W}^{l'}_{1l} + \mathcal{W}^{l'}_{2l}] C_{\tilde{E}l'} + \frac{1}{2}[\mathcal{W}^{l'}_{1l} - \mathcal{W}^{l'}_{2l}] C_{\tilde{B}l'} ,$$

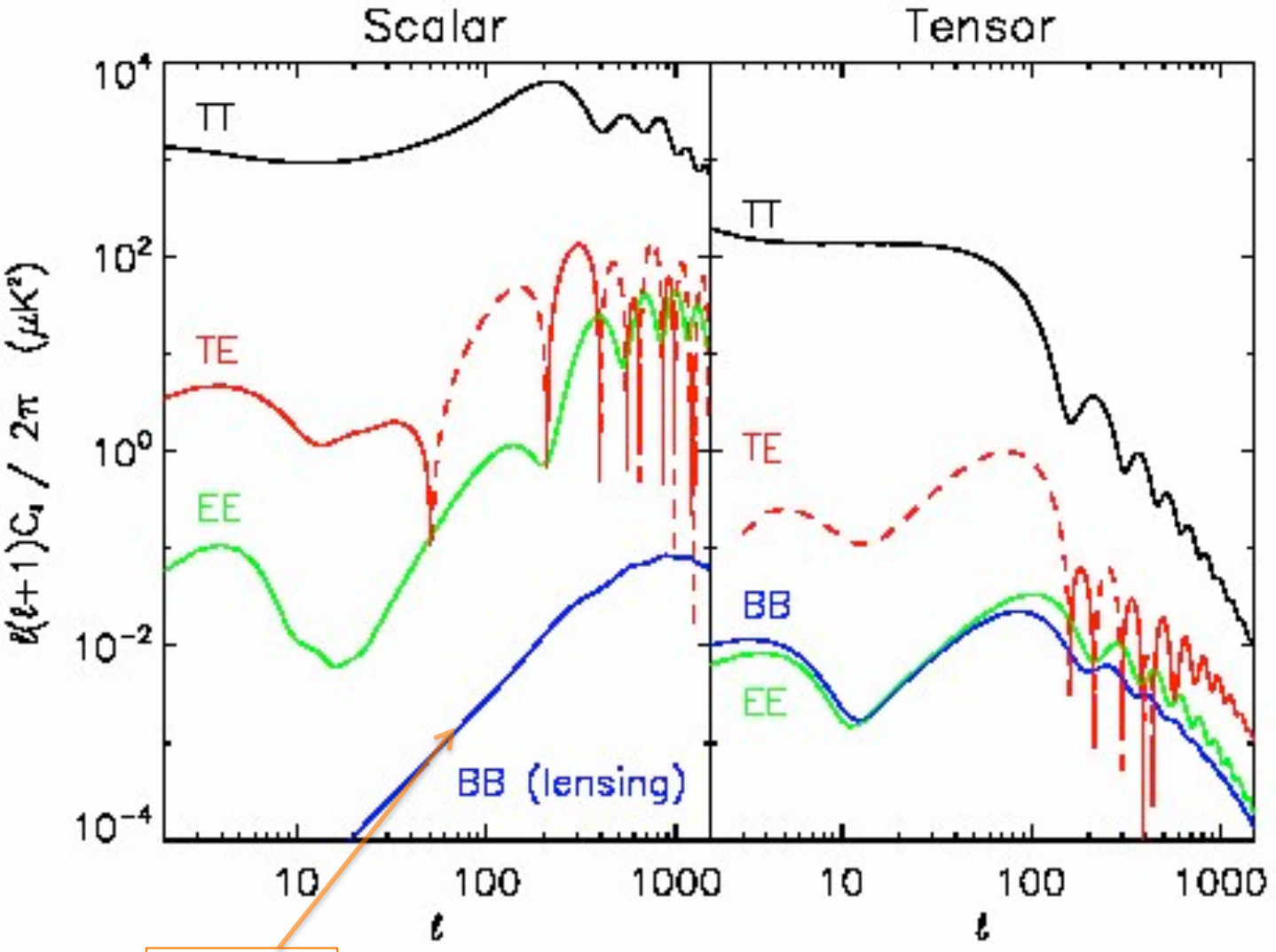
$$C_{Bl} = C_{\tilde{B}l} + \frac{1}{2}[\mathcal{W}^{l'}_{1l} - \mathcal{W}^{l'}_{2l}] C_{\tilde{E}l'} - \frac{1}{2}[\mathcal{W}^{l'}_{1l} + \mathcal{W}^{l'}_{2l}] C_{\tilde{B}l'} ,$$

$$C_{Cl} = C_{\tilde{C}l} + \mathcal{W}^{l'}_{3l} C_{\tilde{C}l'} ,$$

❖ Lensing mixes E and B polarization modes. Even for pure scalar fluctuations, B-modes are generated at small scales.



Power spectra - Theory



Polarization and tensor modes: Gravitational waves

❖ Gravitational waves are a natural consequence of inflationary models (Grishchuk 1974; Rubakov et al. 1982; Starobinsky 1982, 1983; Abbott & Wise 1984).

❖ GW created as vacuum fluctuations (exactly as density perturbations).

❖ GW correspond to spatial metric perturbations. From A. Lasenby´s talk:

- Tensor metric perturbations $ds^2 = a^2[d\eta^2 - (\delta_{ij} + 2h_{ij})dx^i dx^j]$ where h_{ij} is traceless and transverse
- If substitute this into Einstein equations, and use modes of the form $\exp(ik_i x^i)$, then find

$$\ddot{h}_{ij} + 2aH\dot{h}_{ij} + k^2 h_{ij} = 8\pi G\Sigma^T_{ij}$$

  where Σ^T_{ij} is the transverse and traceless part of the anisotropic stress
- Express h_{ij} in terms of the two independent gravitational wave polarization components, h_+ and $h_\times$

Polarization and tensor modes: Gravitational waves

❖From A. Lasenby´s lecture:

- Two independent solutions:

$$h \propto \frac{1}{\eta^3}(\cos(k\eta)k\eta - \sin(k\eta)) \quad \text{and} \quad h \propto \frac{1}{\eta^3}(\cos(k\eta)k\eta + \sin(k\eta))$$

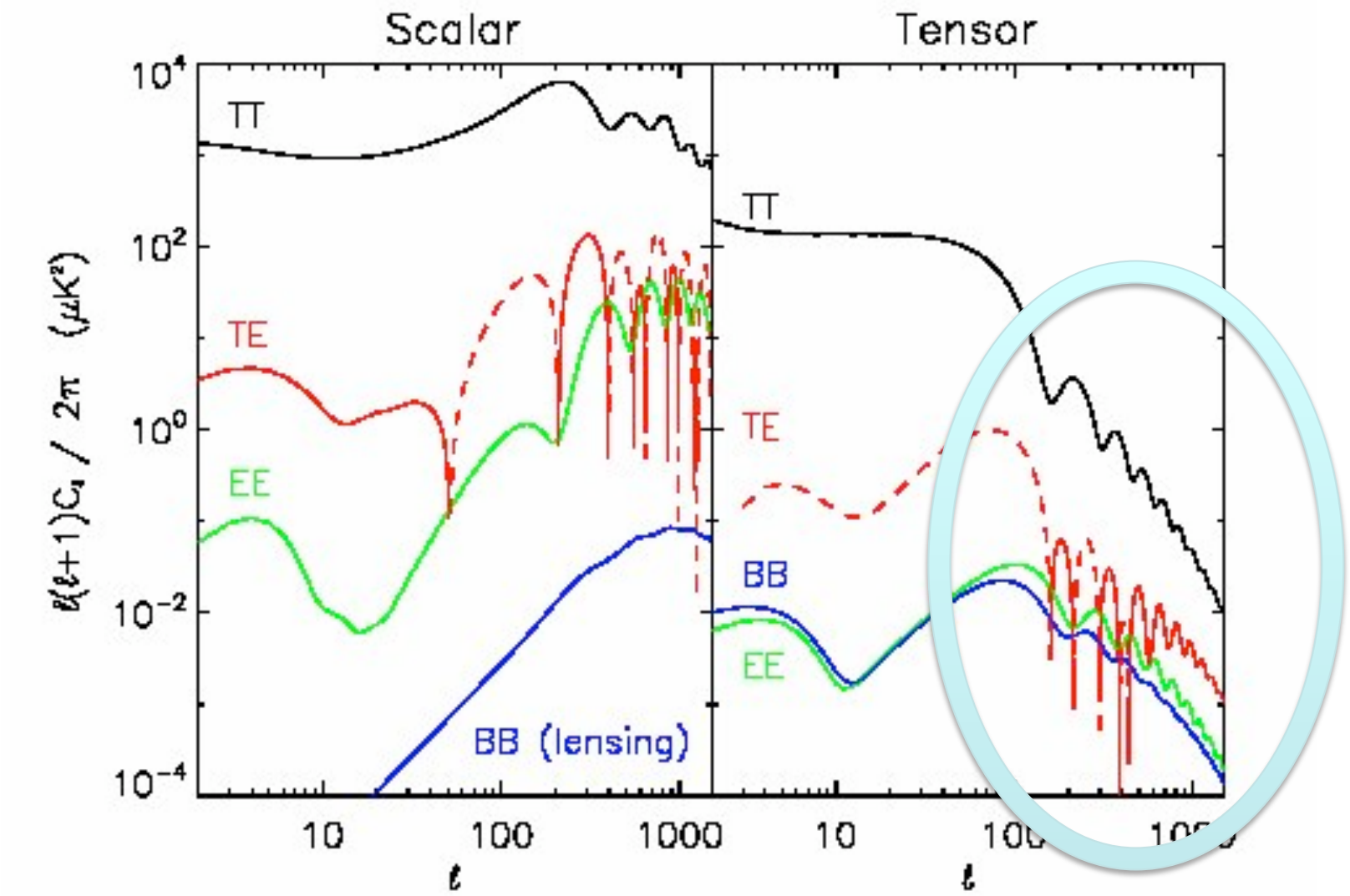
- Find following behavior as one goes back in time: only one of these is non-singular, and here can write

$$h_{+,\times}(\eta) = -\frac{3}{k^3\eta^3}(\cos(k\eta)k\eta - \sin(k\eta))h_{+,\times}(0) = 3\sqrt{\frac{\pi}{2}}\frac{J_{3/2}(k\eta)}{(k\eta)^{3/2}}h_{+,\times}(0)$$

- So this is solution for scales entering the horizon well after matter domination
- Note that overall η dependence for large η is $1/\eta^2 \propto 1/t^{2/3} \propto 1+z$, so the gravitational perturbations redshift away inside the horizon — quite unlike the scalar perturbations
- Therefore only important at degree scales and above

❖ GW oscillate and decay at horizon crossing.

Power spectra - Theory



Effects only on large scales because gravity waves damp inside horizon.

Polarization and tensor modes: Gravitational waves

❖Gravitational waves produce a quadrupolar distortion in the temperature of the CMB.

❖ B-mode polarization is produced, because the symmetry of a plane wave is broken by the transverse nature of gravity wave polarization.

$$C_{El}^{(T)} = (4\pi)^2 \int k^2\, dk\, P_h(k) \left| \int d\tau\, g(\tau)\Psi(k,\tau)\left[-j_l(x) + j_l''(x) + \frac{2j_l(x)}{x^2} + \frac{4j_l'(x)}{x} \right] \right|^2,$$

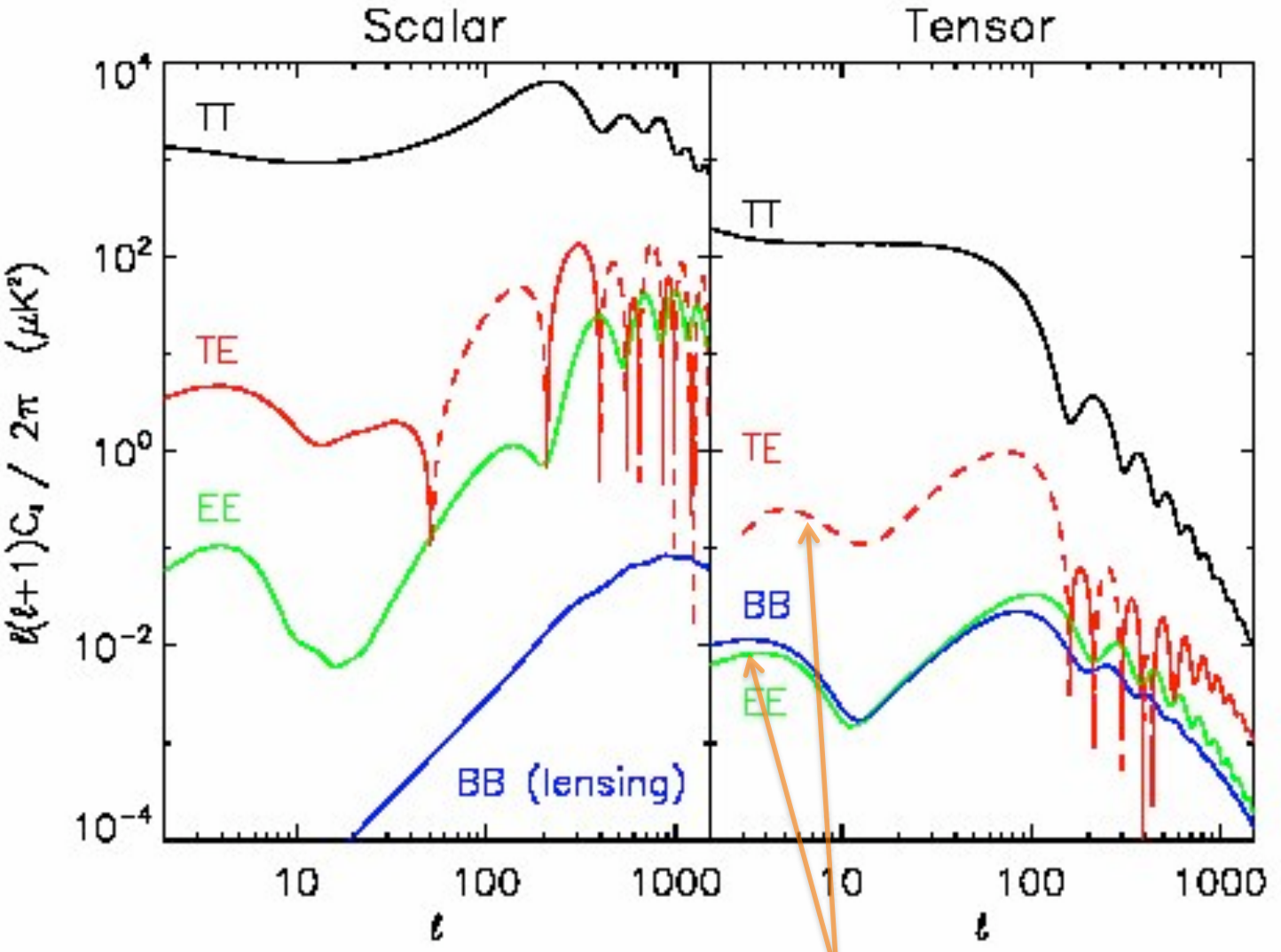
$$C_{Bl}^{(T)} = (4\pi)^2 \int k^2\, dk\, P_h(k) \left| \int d\tau\, g(\tau)\Psi(k,\tau)\left[2j_l'(x) + \frac{4j_l}{x} \right] \right|^2,$$

(Seljak & Zaldarriaga 1997)

❖ E and B modes have similar amplitude.

❖Again, polarization is only generated at last scattering surface (or reionization).

Power spectra - Theory



What would a detection of GW tell us?

- It would provide strong evidence that inflation happened!
- The amplitude of the power spectrum is a (model-independent) measurement of the energy scale of inflation.

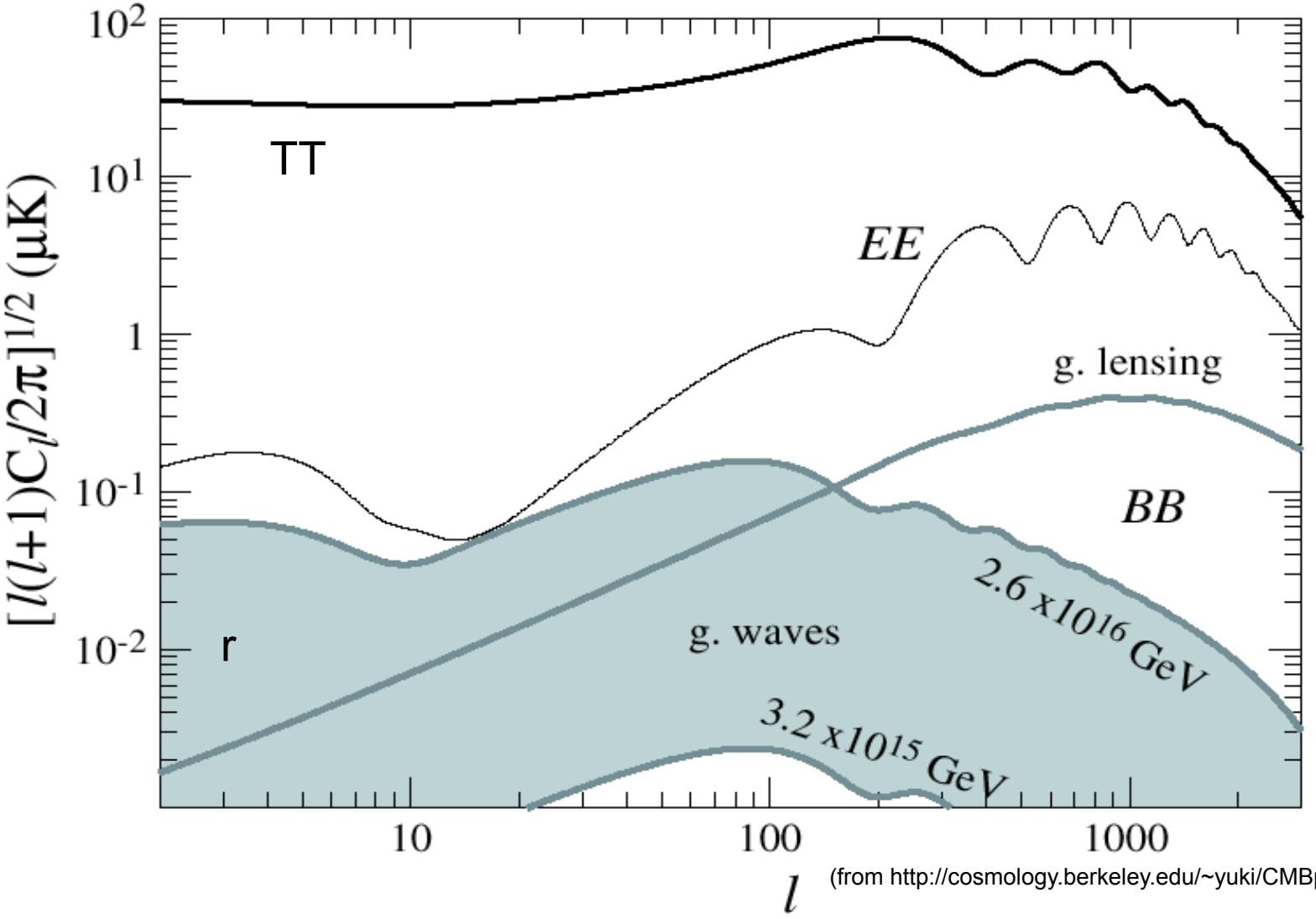
$$P_{tensor} = \frac{8}{m_{Pl}^2}\left(\frac{H}{2\pi}\right)^2 \propto E_{\inf}^4$$

- Defining the tensor-to-scalar ratio (r) at a certain scale k_0 (typically 0.001 Mpc^{-1}), we have

$$r \equiv \frac{P_{tensor}(k_0)}{P_{scalar}(k_0)} = 0.008\left(\frac{E_{\inf}}{10^{16}GeV}\right)^4$$

- Values of r of the order of 0.01 or larger would imply that inflation occurred at the GUT scale.
- These scales are 12 orders of magnitude larger than those achievable at LHC!

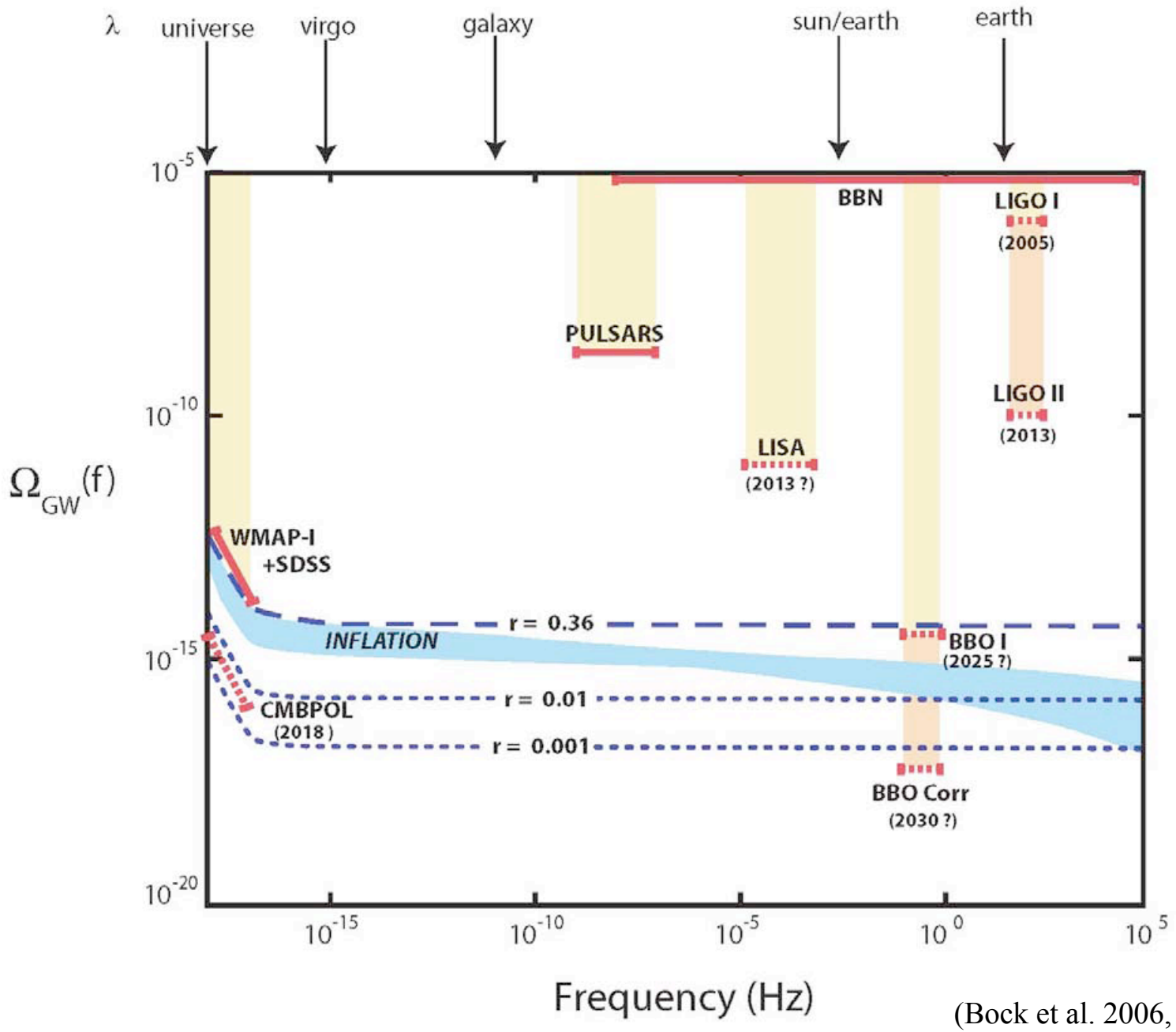
Primordial gravitational waves and B-modes



(from http://cosmology.berkeley.edu/~yuki/CMBpol/CMBpol.htm)

- r=0.1 corresponds to an energy scale of inflation around 2×10^{16} GeV.

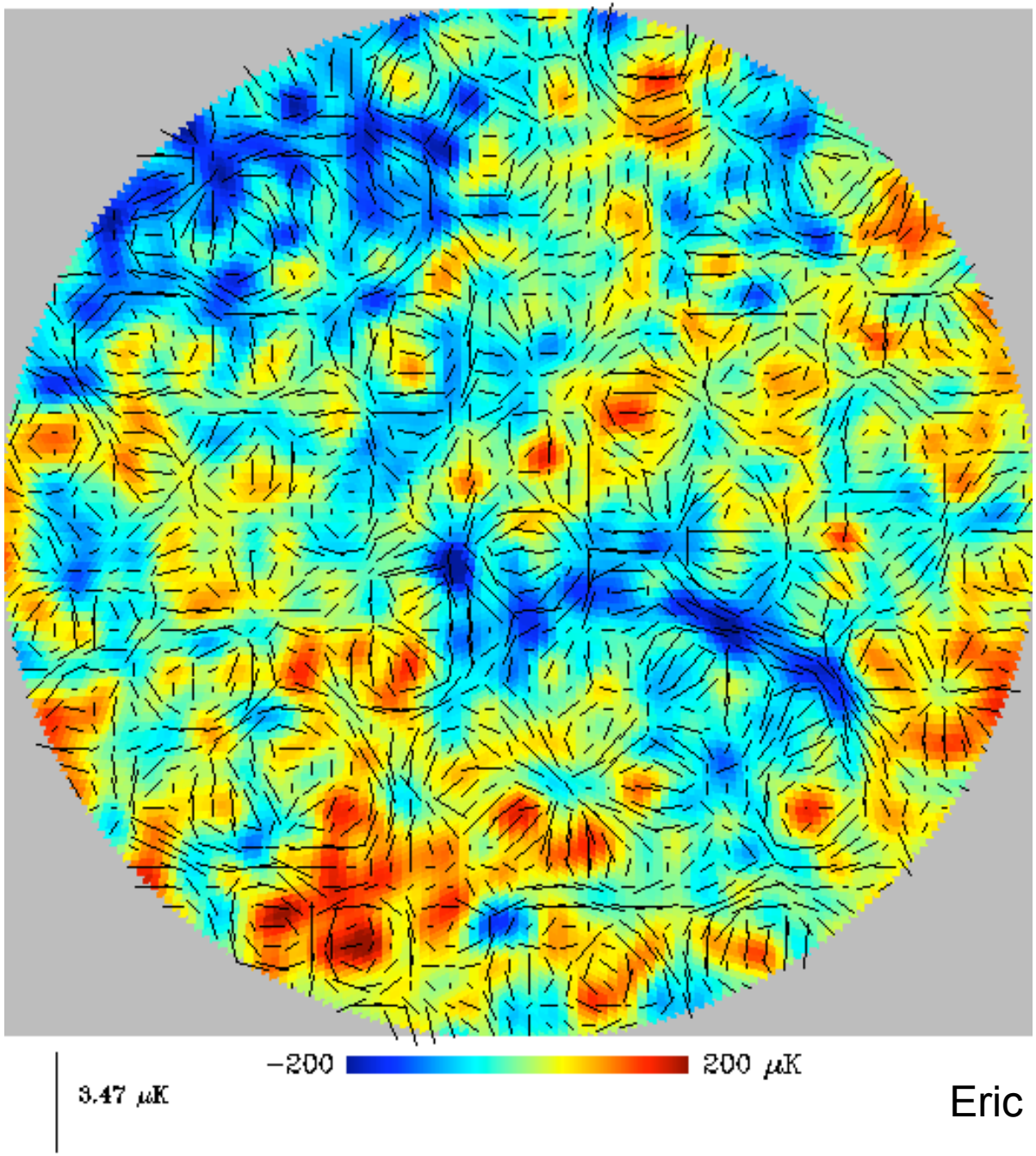
Direct measurements of primordial Gravitational Waves



(Bock et al. 2006, arXiv:0604101)

No gravitational waves (r = 0)

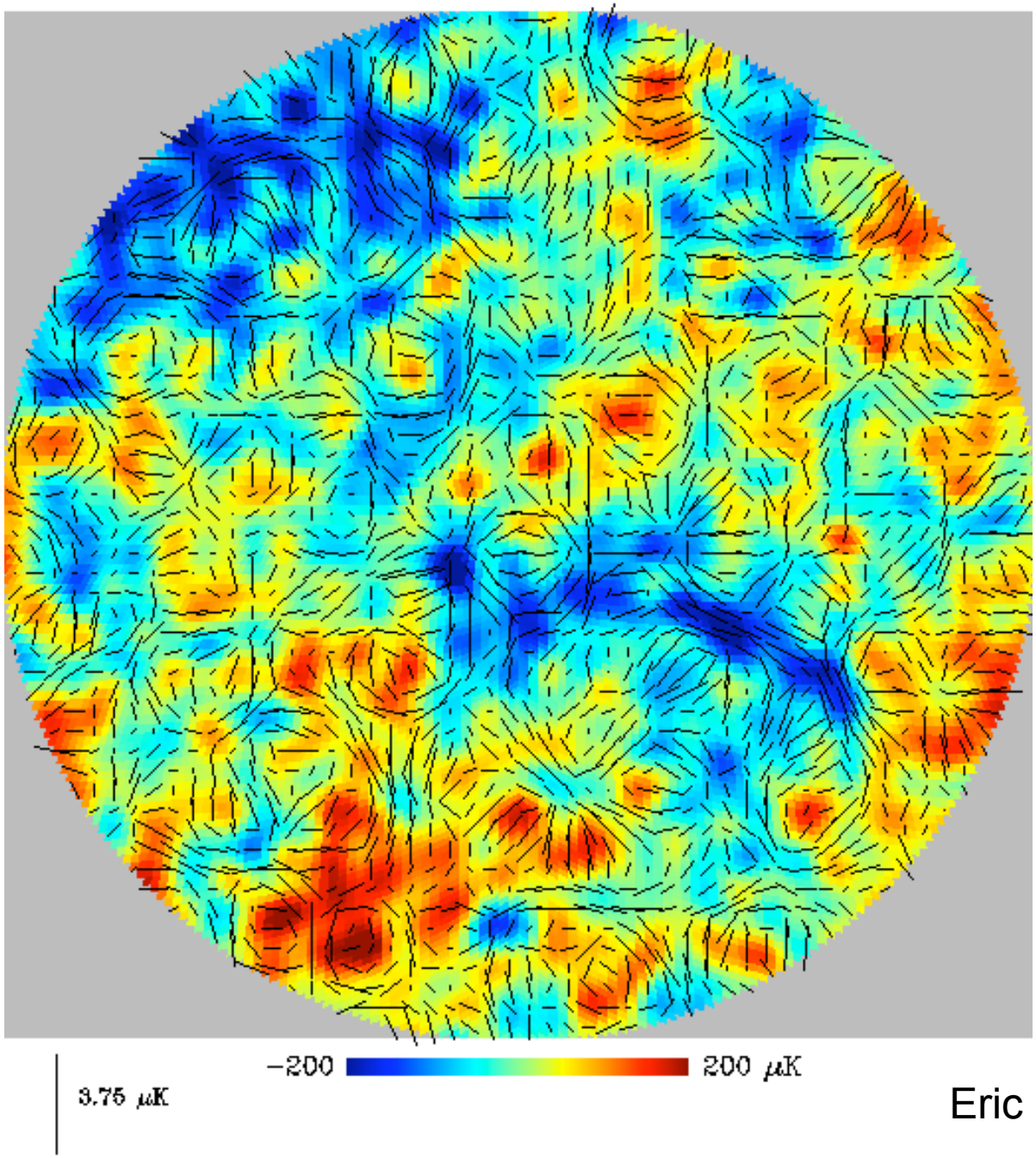
30 degrees



Eric Hivon

Gravitational waves (r = 0.3)

30 degrees



Eric Hivon

Raising/lowering operators

An important property of spin-s functions is that there exists a spin raising (lowering) operator $\eth$ ($\bar{\eth}$) with the property of raising (lowering) the spin weight of a function, $(\eth_s f)' = e^{-i(s+1)\psi}\eth_s f$, $(\bar{\eth}_s f)' = e^{-i(s-1)\psi}\bar{\eth}_s f$. Their explicit expression is given by

$$\eth_s f(\theta,\phi) = -\sin^s(\theta)\left[\frac{\partial}{\partial\theta} + i\csc(\theta)\frac{\partial}{\partial\phi}\right]\sin^{-s}(\theta)\,{}_s f(\theta,\phi),$$

$$\bar{\eth}_s f(\theta,\phi) = -\sin^{-s}(\theta)\left[\frac{\partial}{\partial\theta} - i\csc(\theta)\frac{\partial}{\partial\phi}\right]\sin^{s}(\theta)\,{}_s f(\theta,\phi). \tag{A2}$$

Raising/lowering operators

❖ They can be used to define spin zero quantities:

$$\bar{\eth}^{\,2}(Q+iU)(\hat{\boldsymbol{n}})=\sum_{lm}\left[\frac{(l+2)!}{(l-2)!}\right]^{1/2}a_{2,lm}Y_{lm}(\hat{\boldsymbol{n}}),$$

$$\eth^{2}(Q-iU)(\hat{\boldsymbol{n}})=\sum_{lm}\left[\frac{(l+2)!}{(l-2)!}\right]^{1/2}a_{-2,lm}Y_{lm}(\hat{\boldsymbol{n}}).$$

❖ For real-space computations, it is useful to define:

$$\begin{aligned}\chi_E &\equiv [\bar{\eth}\,\bar{\eth}(Q+iU)+\eth\eth(Q-iU)]/2\\ &=-\sum_{lm}[(l+2)!/(l-2)!]^{1/2}a_{E,lm}Y_{lm},\\ \chi_B &\equiv i[\bar{\eth}\,\bar{\eth}(Q+iU)-\eth\eth(Q-iU)]/2\\ &=\sum_{lm}[(l+2)!/(l-2)!]^{1/2}a_{B,lm}Y_{lm}.\end{aligned}$$

E/B mixing

❖In finite patches of sky, the separation between E and B modes can not be done perfectly.

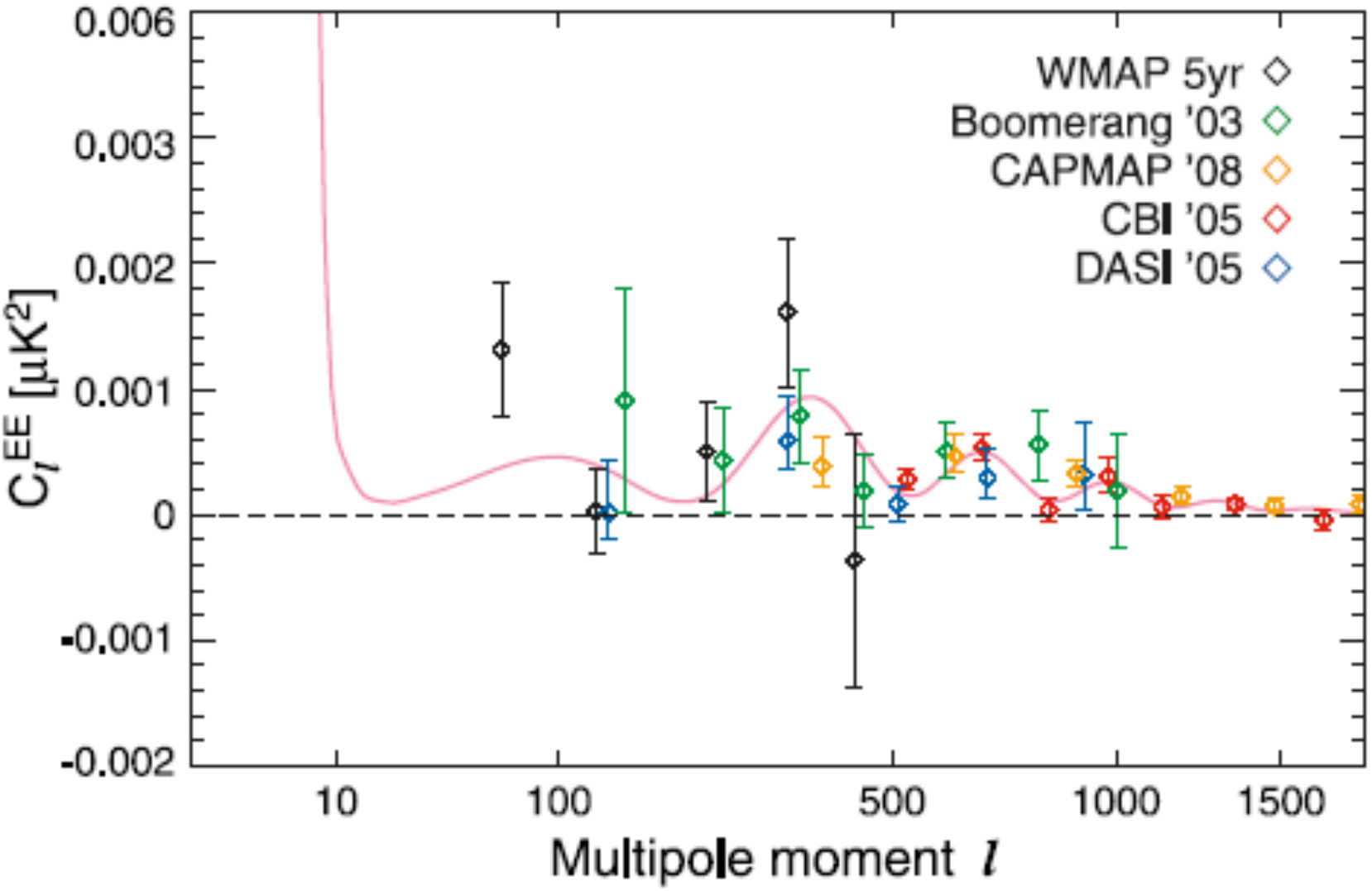
❖ Aliasing effects from E modes into B modes are very important, as E-mode spectrum is much larger in amplitude.

❖Specific methods in Fourier space or real space to minimize this problem.

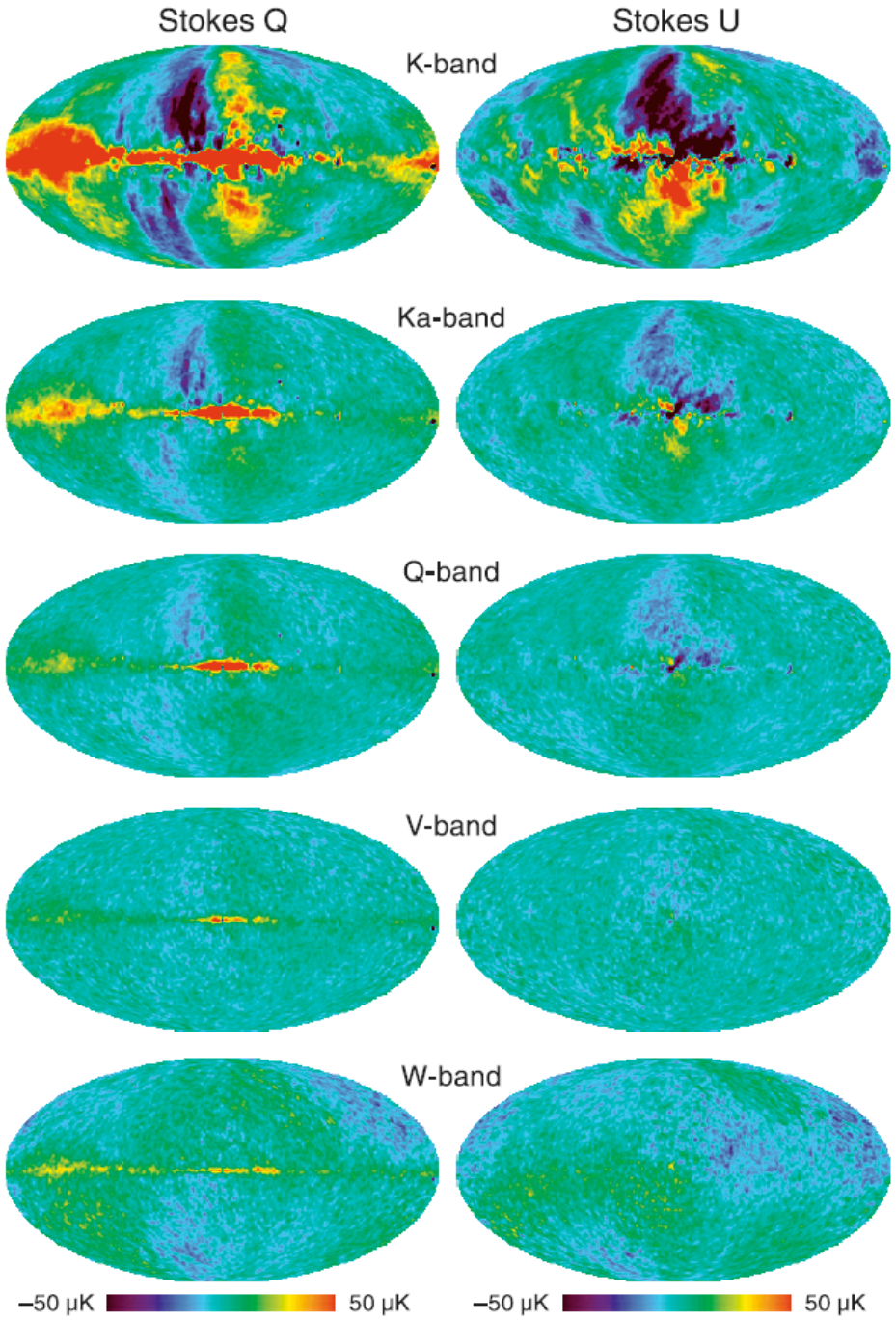
Observational status: polarization of the CMB

First Observations of CMB polarization

E-mode detections: DASI (Kovac et al. 2002, Nature), WMAP, CAPMAP, CBI, Boomerang.

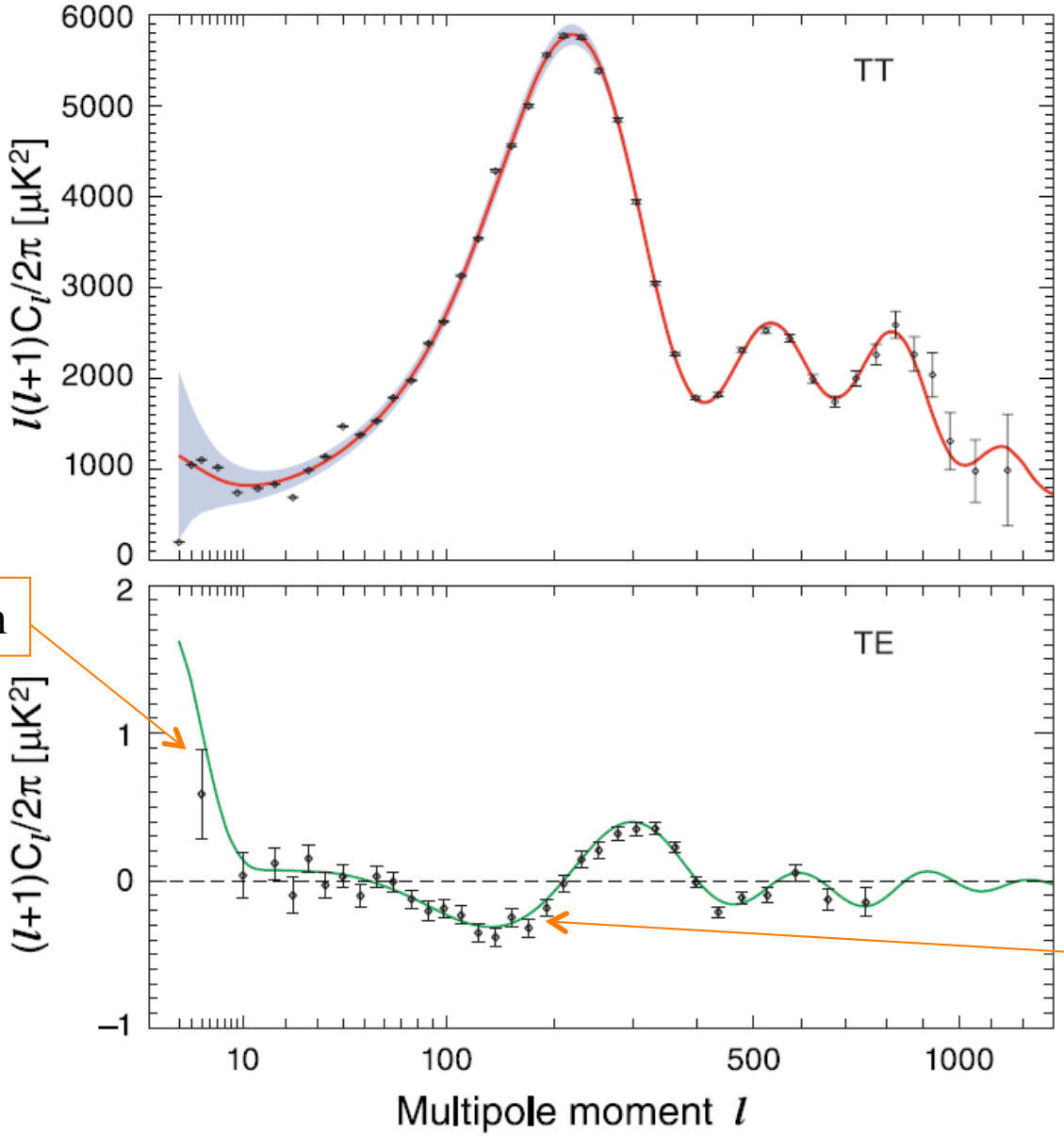


WMAP 7yr results



Jarosik et al. 2010

WMAP7 power spectrum



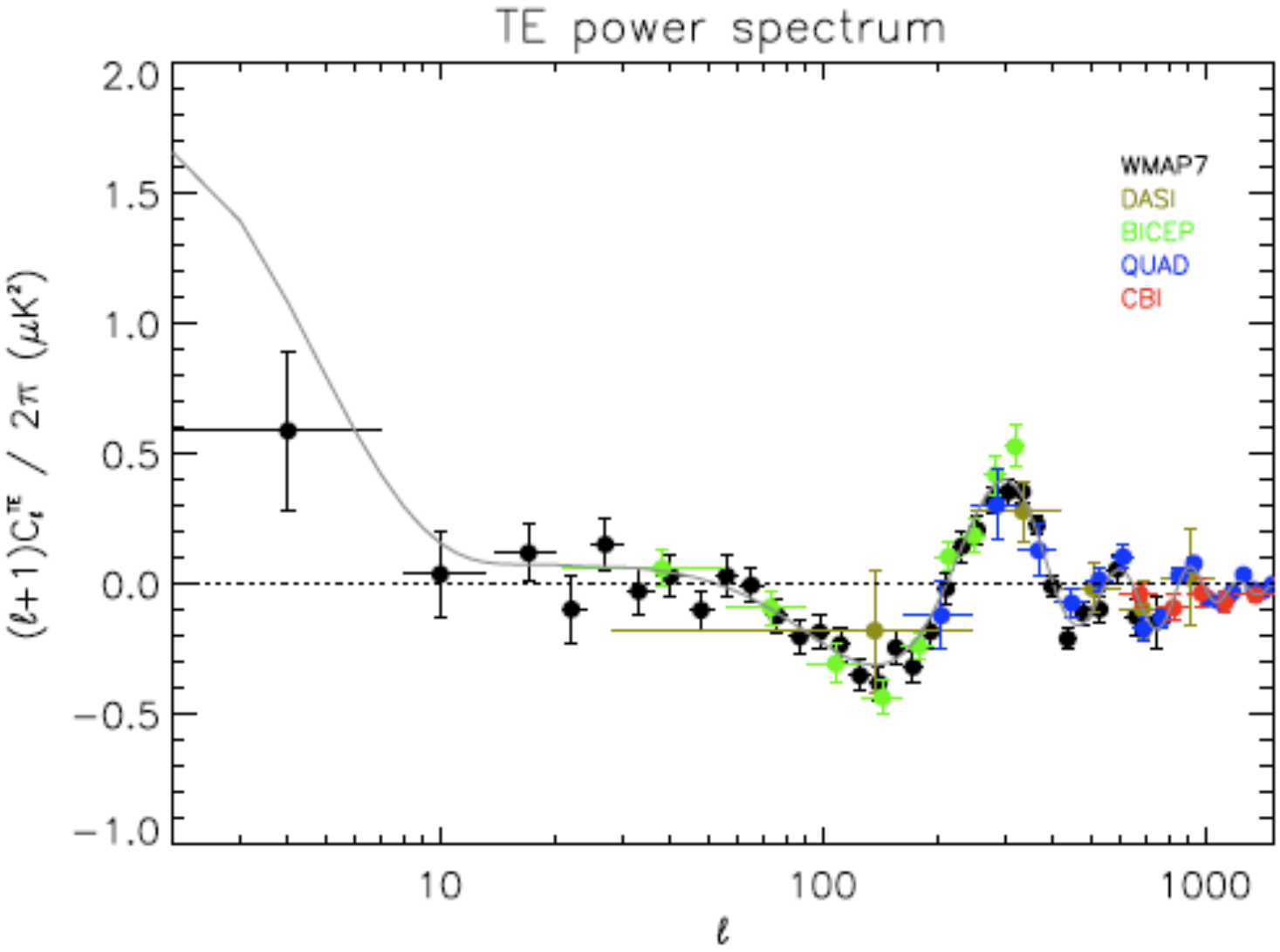
(Larson et al. 2011)

Reionization

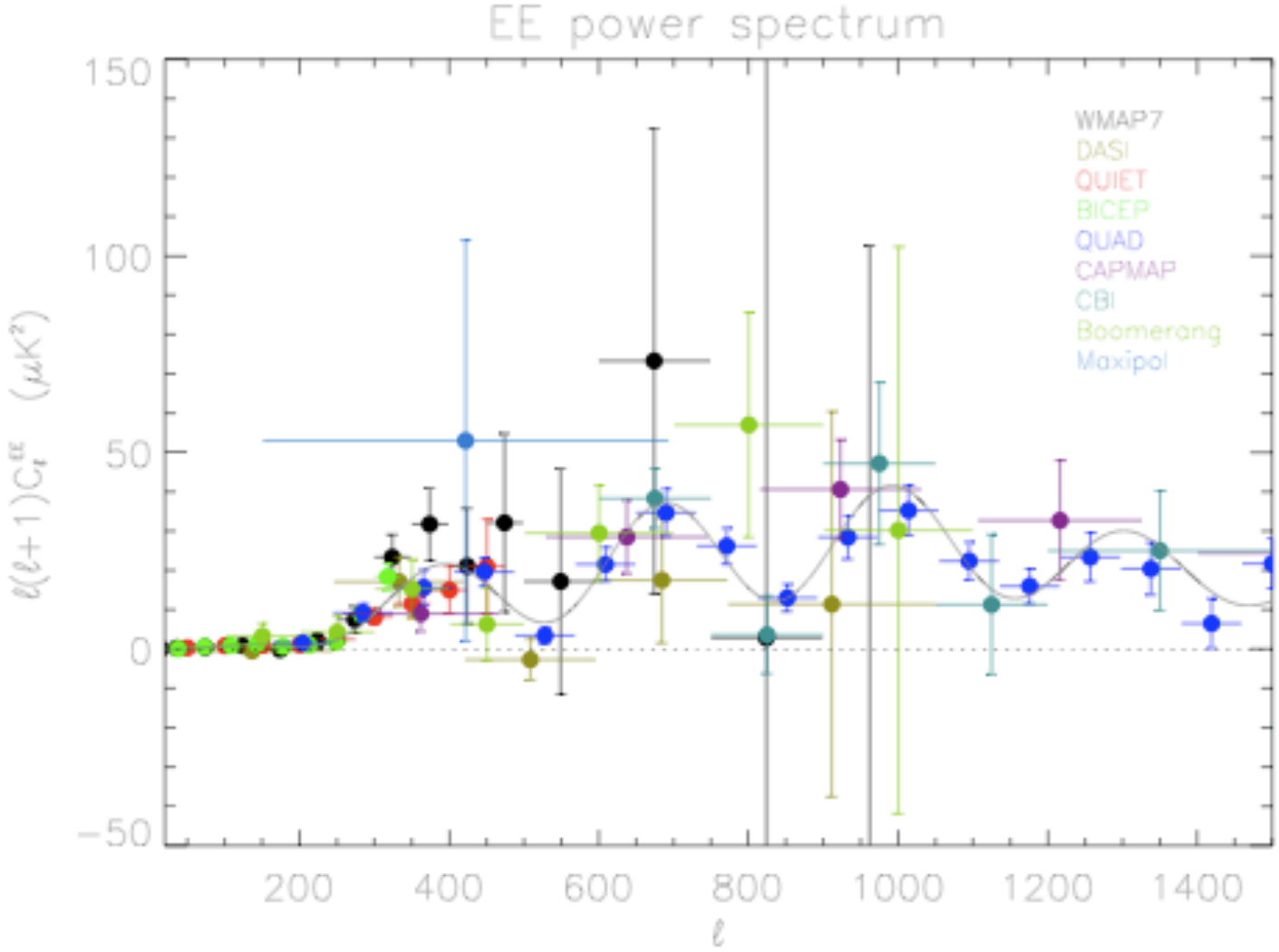
Green line is the ΛCDM prediction!

Anticorrelation peak at l=150 is a signature of superhorizon adiabatic fluctuations

CMB polarization. TE spectrum



CMB polarization. EE spectrum



Real space correlations

$$\langle T_i T_j \rangle \equiv \sum_l \left(\frac{2l+1}{4\pi} \right) P_l(z) C_l^T$$

$$\langle T_i Q_j \rangle \equiv -\sum_l \left(\frac{2l+1}{4\pi} \right) F_l^{10}(z) C_l^{TE},$$

$$\langle T_i U_j \rangle \equiv -\sum_l \left(\frac{2l+1}{4\pi} \right) F_l^{10}(z) C_l^{BT},$$

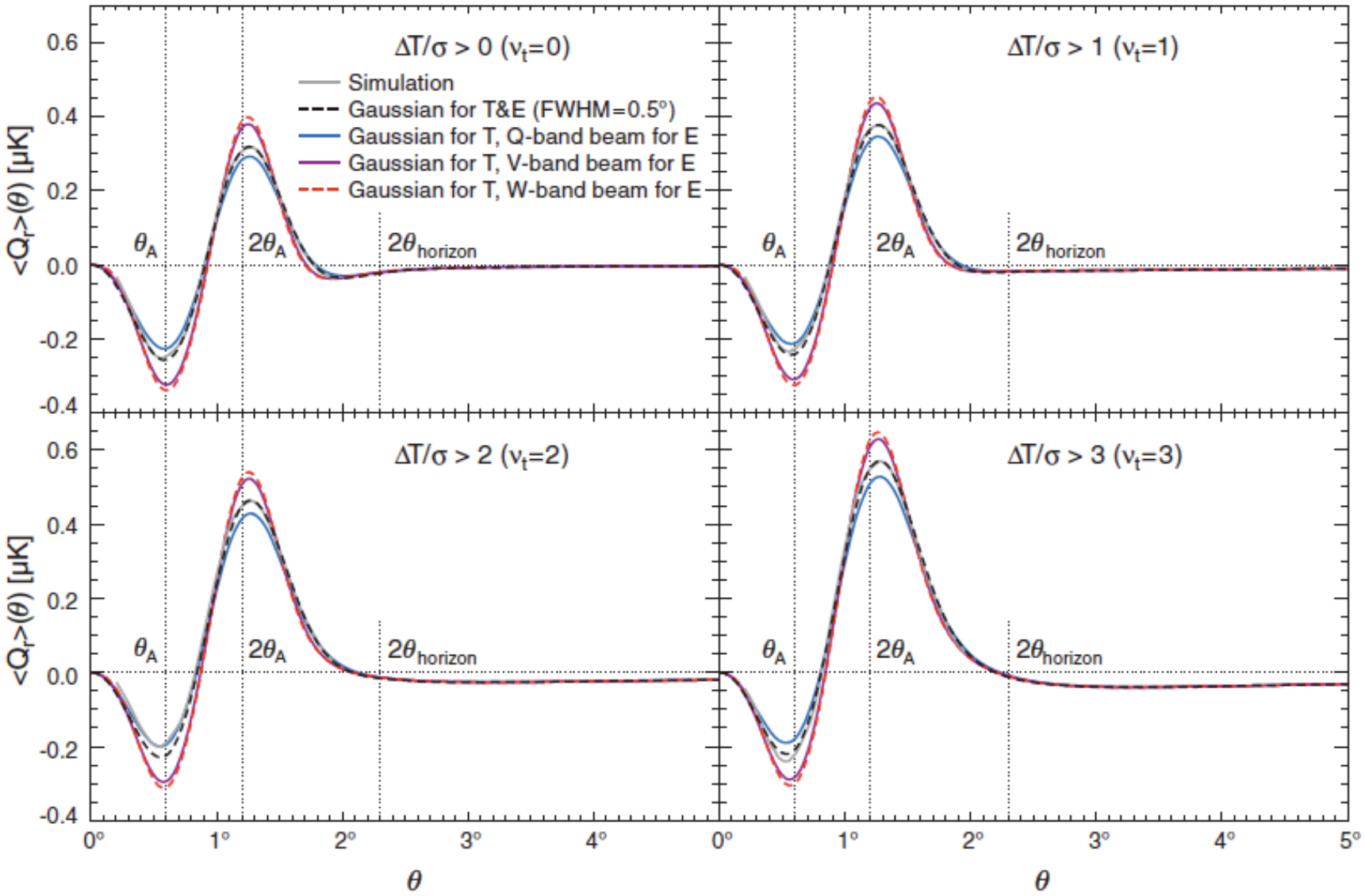
$$\langle Q_i Q_j \rangle \equiv \sum_l \left(\frac{2l+1}{4\pi} \right) [F_l^{12}(z) C_l^E - F_l^{22}(z) C_l^B],$$

$$\langle U_i U_j \rangle \equiv \sum_l \left(\frac{2l+1}{4\pi} \right) [F_l^{12}(z) C_l^B - F_l^{22}(z) C_l^E],$$

$$\langle Q_i U_j \rangle \equiv \sum_l \left(\frac{2l+1}{4\pi} \right) [F_l^{12}(z) + F_l^{22}(z)] C_l^{EB},$$

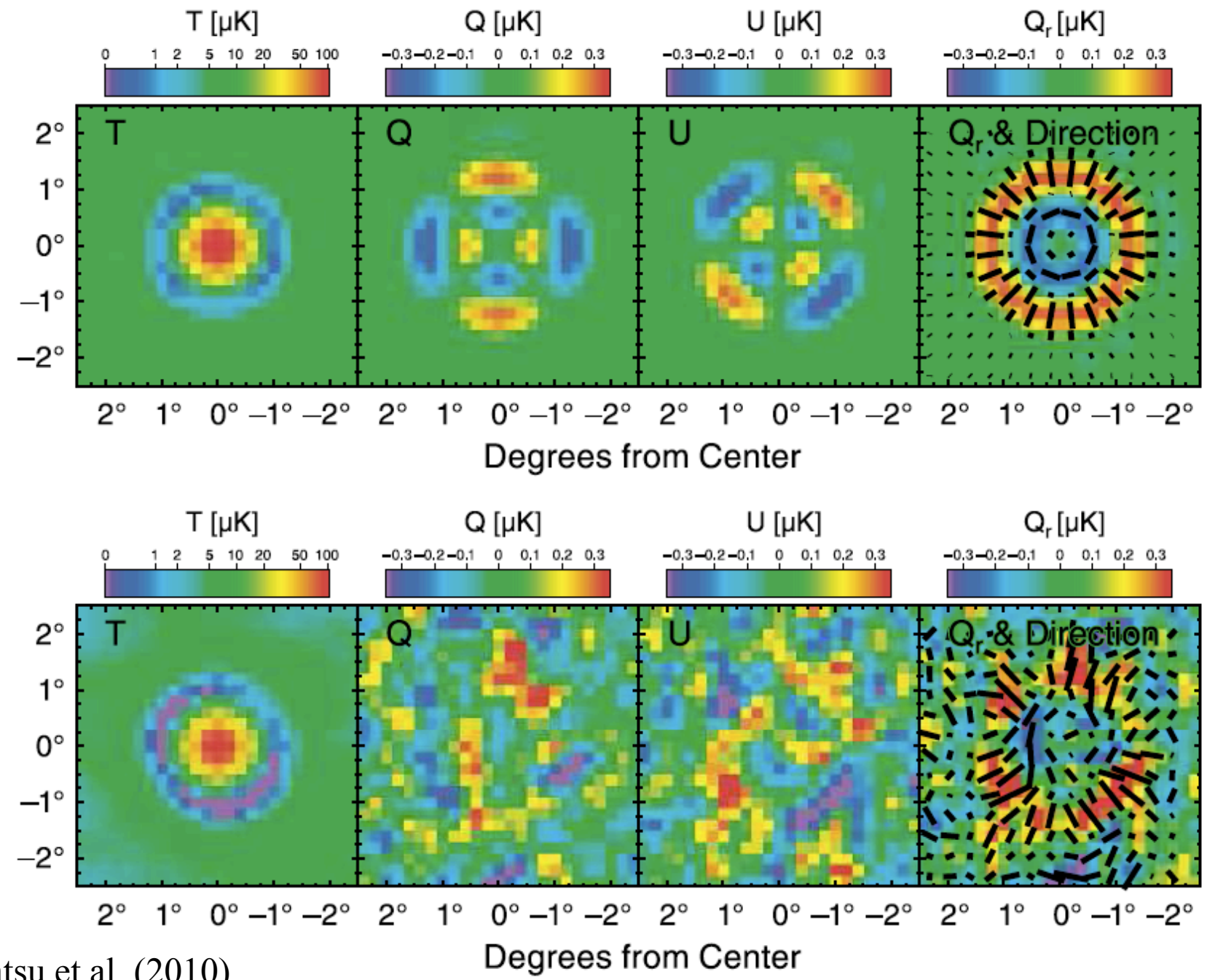
(See e.g. Tegmark et al. 2006)

WMAP7: real space correlations



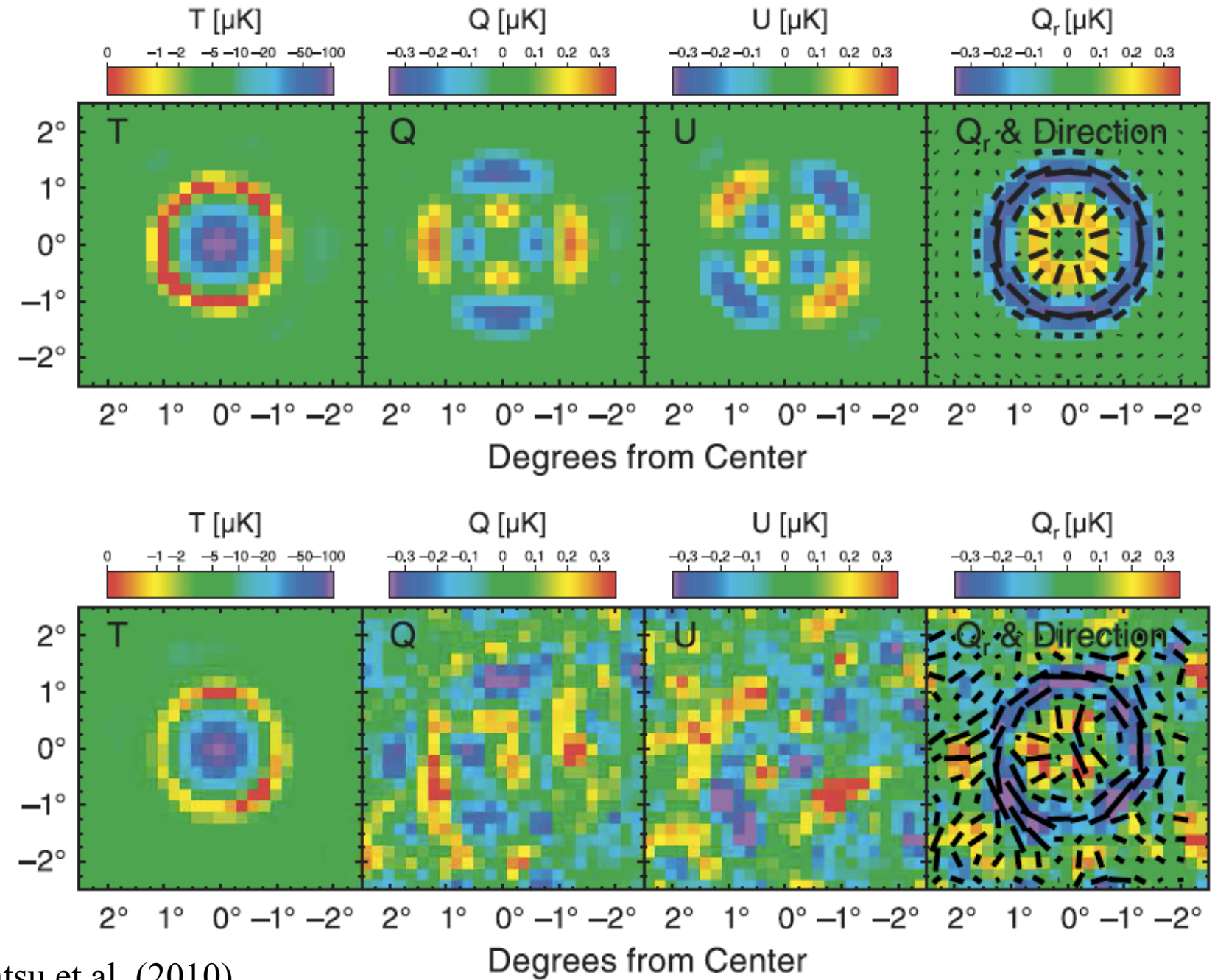
Komatsu et al. (2010)

WMAP7 hot spots



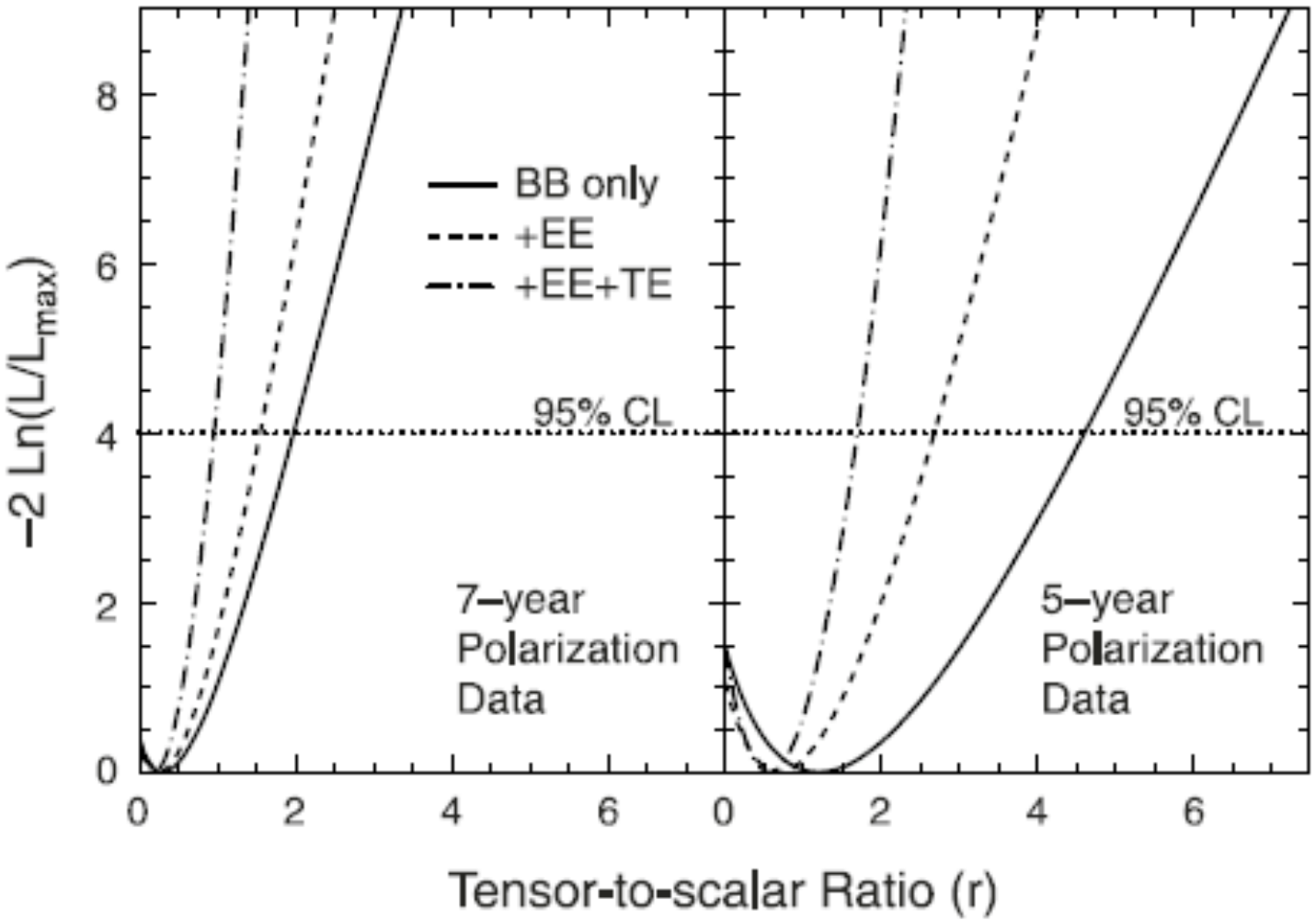
Komatsu et al. (2010)

WMAP7 cold spots



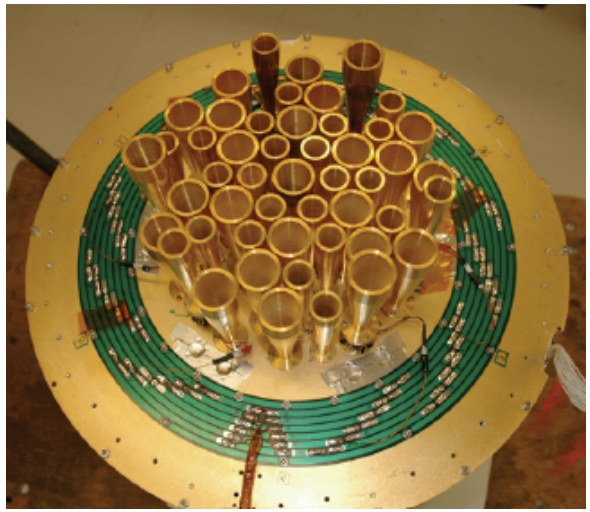
Komatsu et al. (2010)

WMAP7 constraints on r



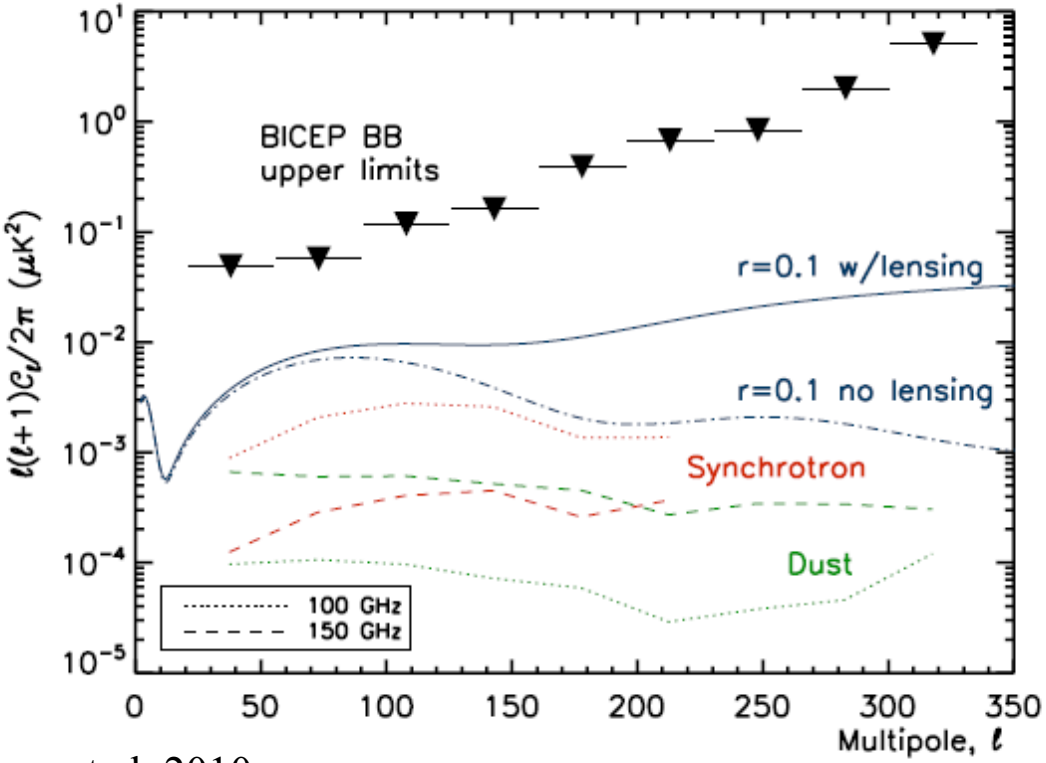
Komatsu et al. (2010)

BICEP results



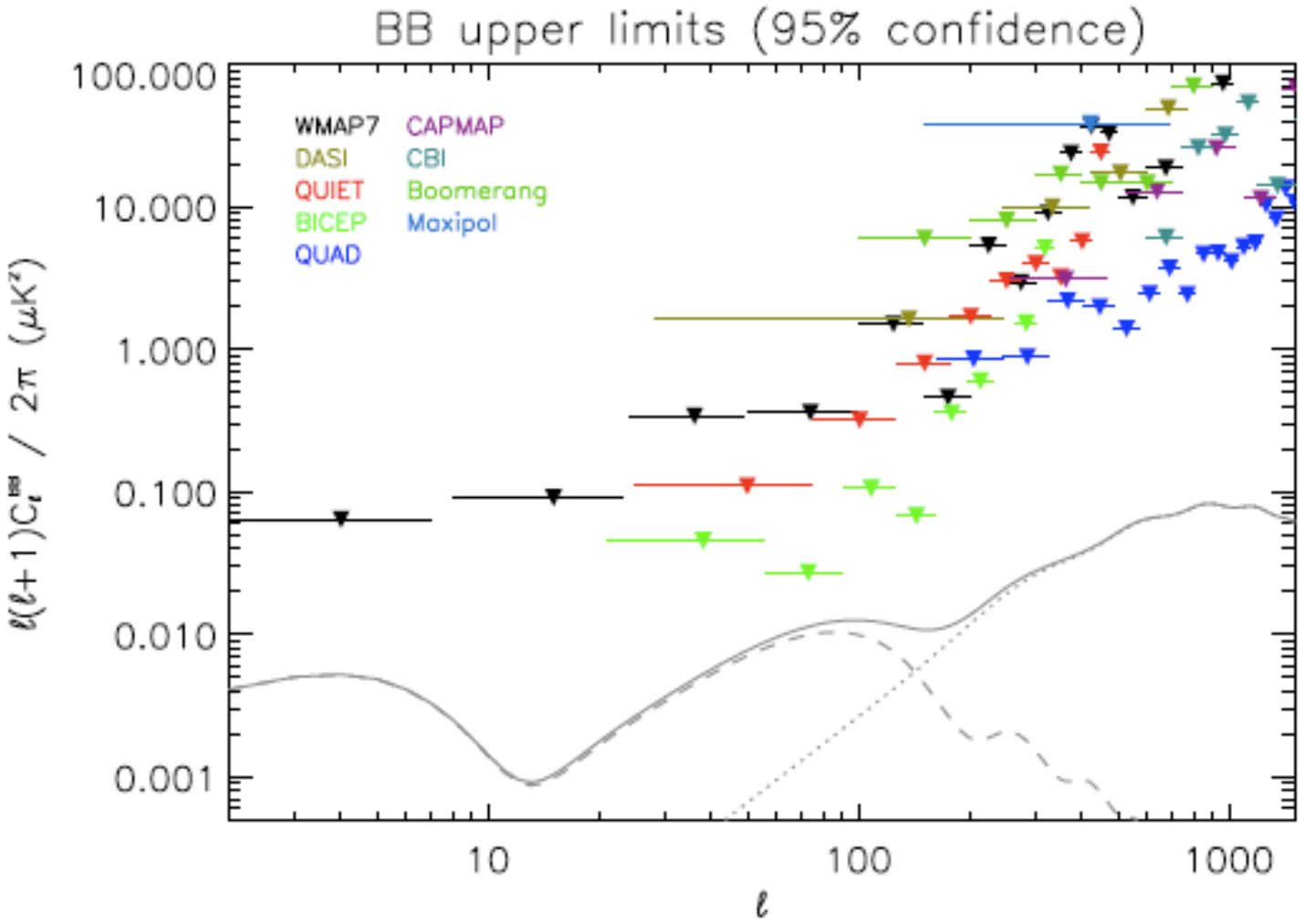
• BICEP (Background Imaging of Cosmic Extragalactic Polarization). Caltech, Princeton, JPL, Berkeley and others collaboration.
• Uses 100GHz and 150GHz polarization sensitive bolometers. Operates from South Pole.

• Detection of E-mode signal.
• Upper limit on r, based on BB spectrum only:
r<0.72 at 95% confidence (Chiang et al. 2010).

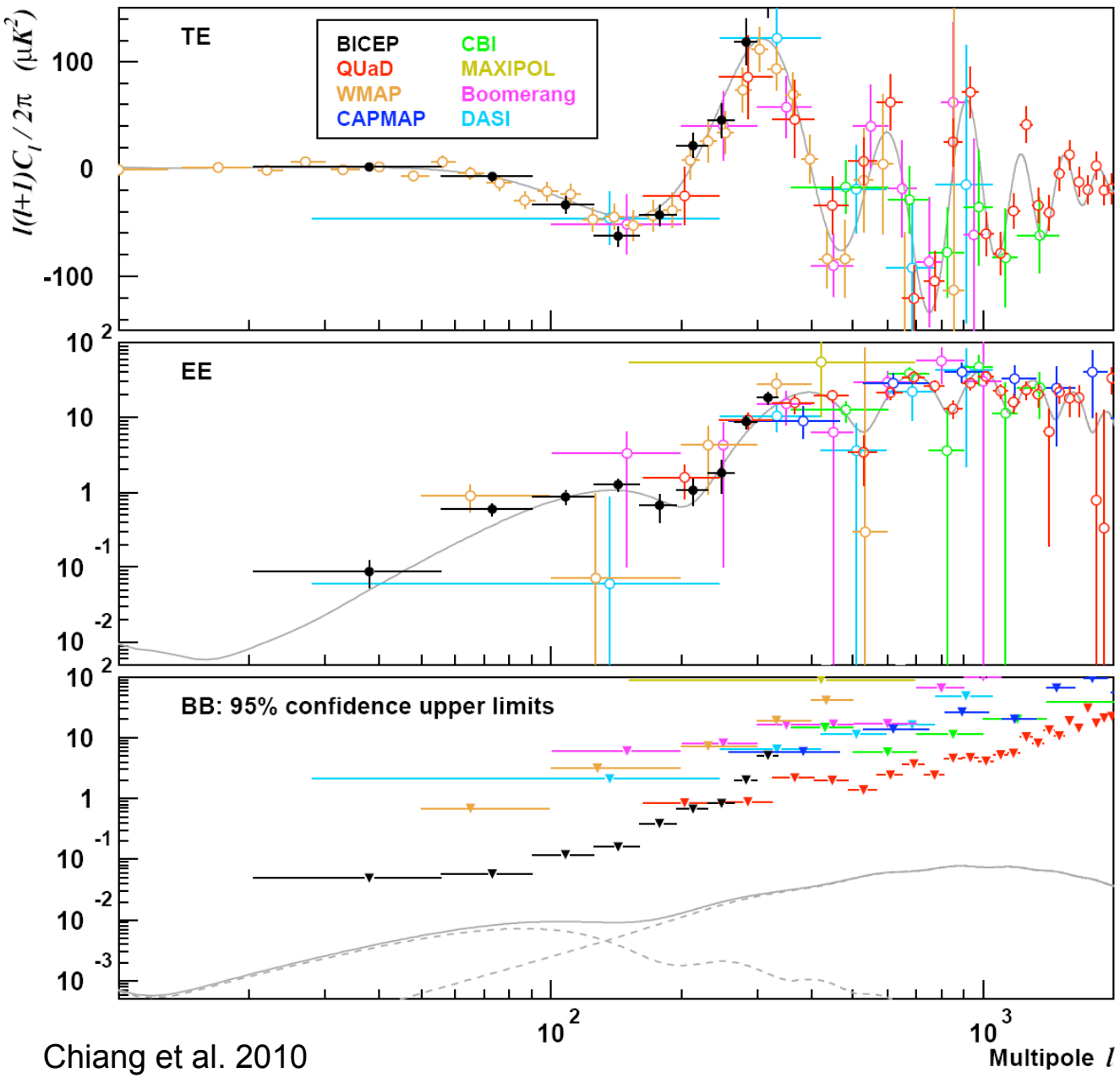


Chiang et al. 2010

CMB polarization. BB spectrum



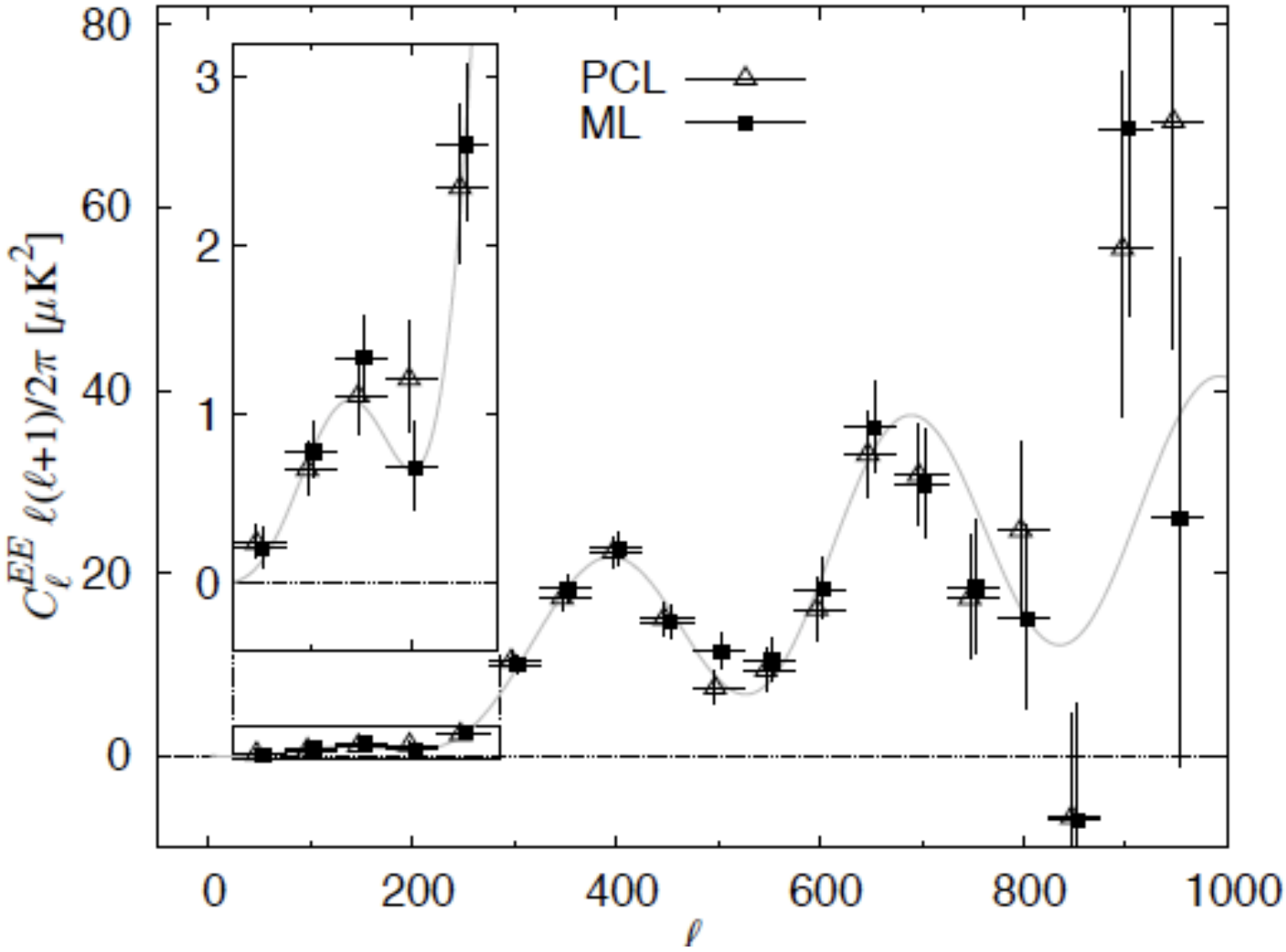
CMB polarization: observational status



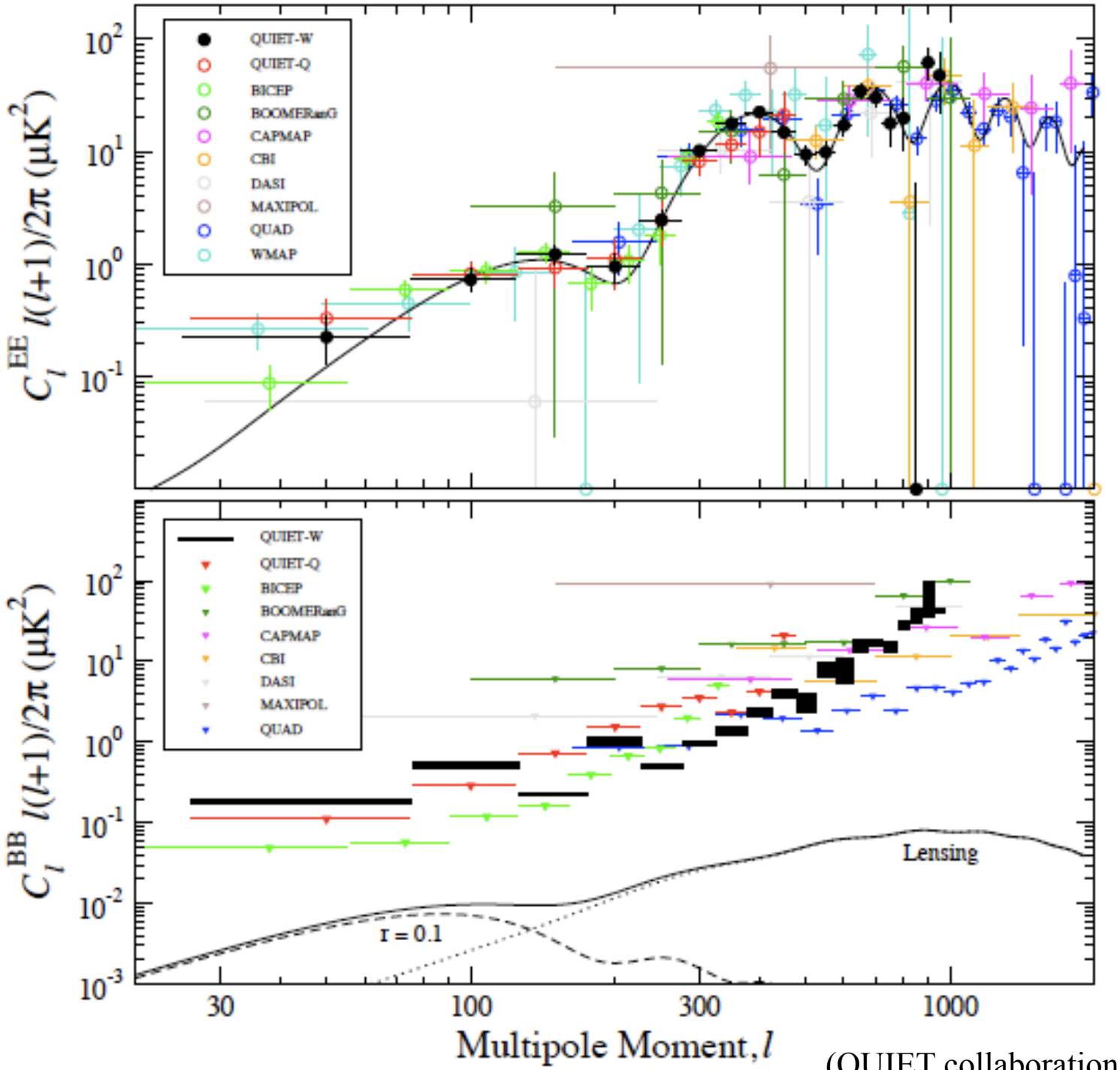
Chiang et al. 2010

- Several E-mode detections: DASI, CBI, CAPMAP, Boomerang, WMAP, QUAD, BICEP, QUIET, etc.
- WMAP7 gives r<0.93 at 95% using TE/EE/BB, and r<2.1 at 95% with BB alone.
- WMAP7+BAO+SN gives r<0.2 (Komatsu et al. 2010).
- BICEP: r<0.72 at 95% with BB only (Chiang et al. 2010).
- QUIET: $r=0.35^{+1.06}_{-0.87}$ with BB only (Bischoff et al. 2010)

QUIET results at 95 GHz



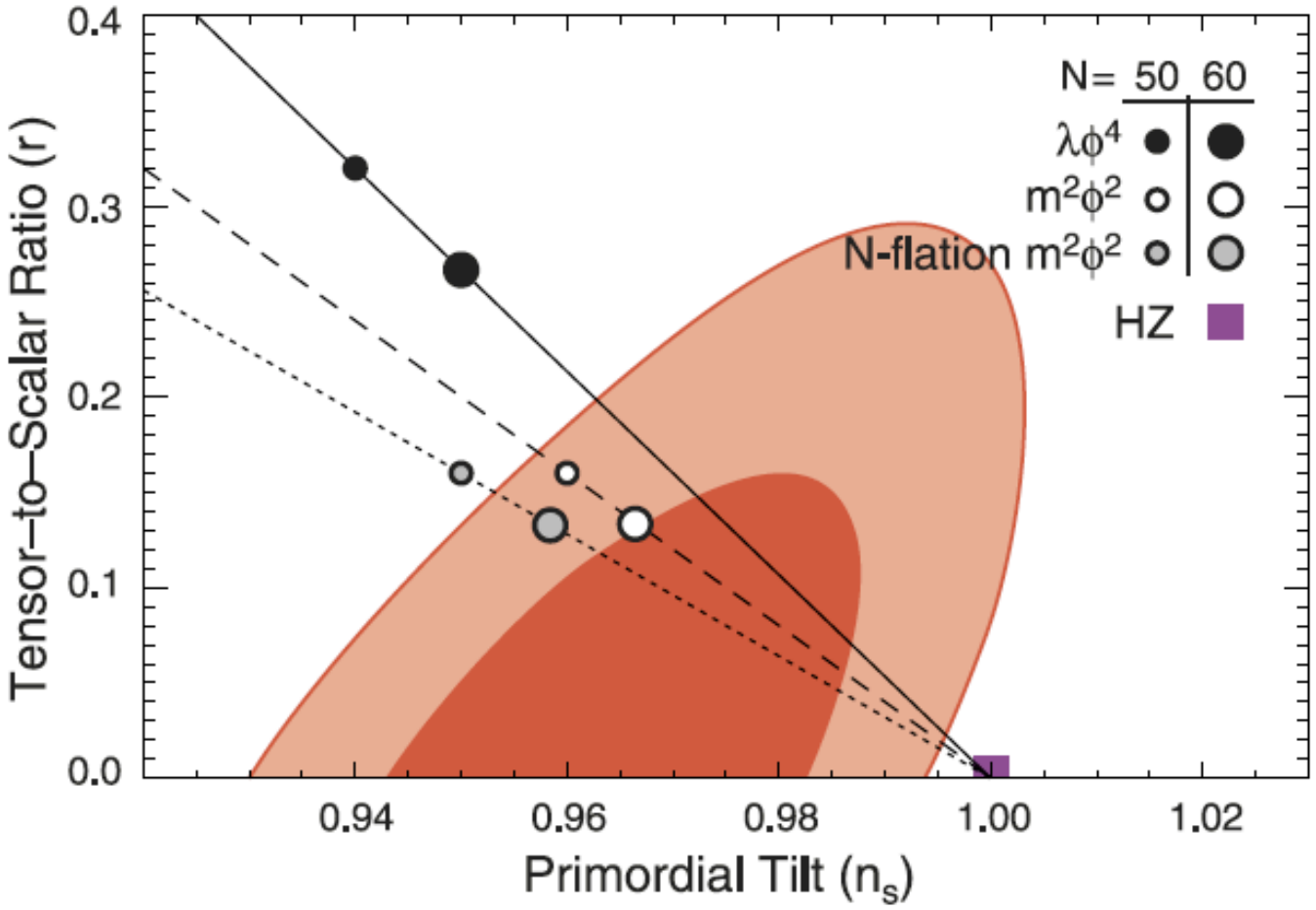
(QUIET collaboration: arXiv1207.5034)



r < 2.7 at
95% C.L.

(QUIET collaboration: arXiv1207.5034)

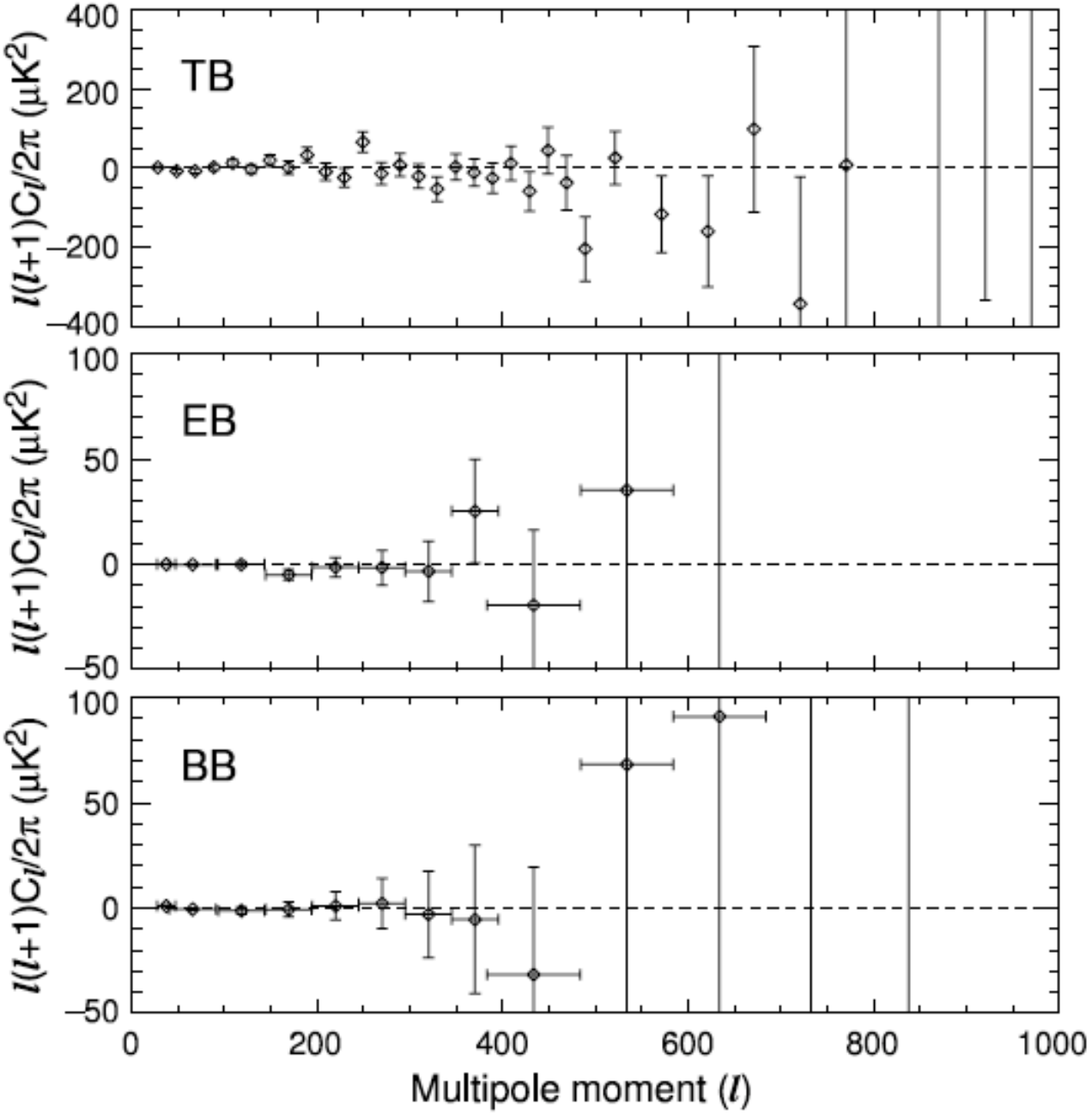
Probing inflation



• Komatsu et al. (2010): r<0.24 (at 95% C.L.) using WMAP7+BAO+H_0

• For chaotic inflationary models with $V=\lambda\phi^p$ and N e-folds of inflation, the predicted tensor-to-scalar ratio is $r=4p/N$, with $n_S=1-(p+2)/2N$. Data excludes p>3 for N=60 at 95% C.L.

TB, EB and BB spectra: hints for new physics?



(WMAP3: Kogut et al. 2007)

TB, EB and BB spectra: hints for new physics?

Testing CPT Symmetry

if terms like Chern-Simons have to be added to the standard electrodynamic Lagrangian, than Lorentz, P and CPT symmetries will be violated. This will induce a rotation of the polarization direction of each photon as it propagates from the LSS to us. This effect is called "Cosmic Birefringence"

$$C_\ell^{\mathrm{TE,obs}} = C_\ell^{\mathrm{TE}} \cos(2\Delta\alpha),$$

$$C_\ell^{\mathrm{TB,obs}} = C_\ell^{\mathrm{TE}} \sin(2\Delta\alpha),$$

$$C_\ell^{\mathrm{EE,obs}} = C_\ell^{\mathrm{EE}} \cos^2(2\Delta\alpha),$$

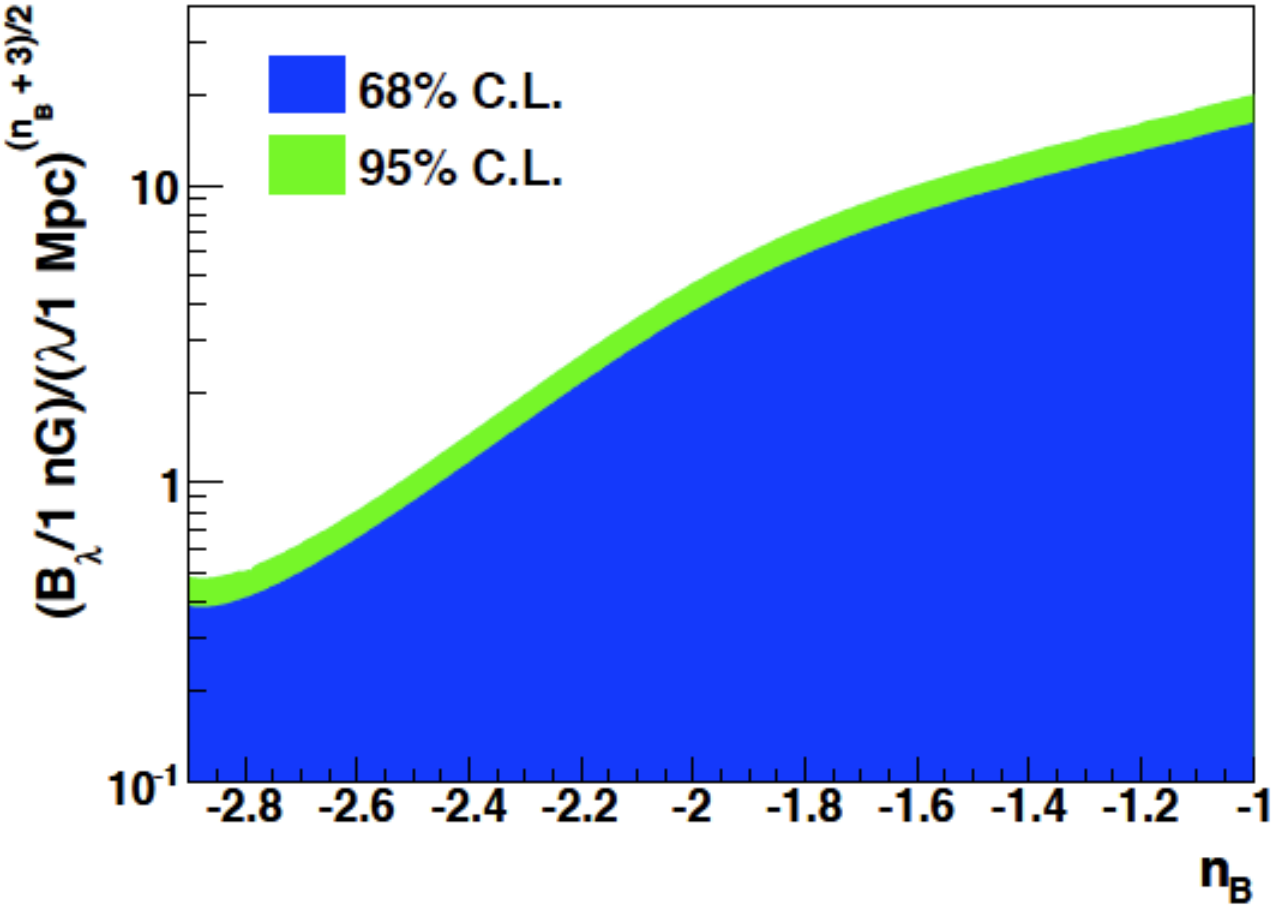
$$C_\ell^{\mathrm{BB,obs}} = C_\ell^{\mathrm{EE}} \sin^2(2\Delta\alpha),$$

$$C_\ell^{\mathrm{EB,obs}} = \tfrac{1}{2}(C_\ell^{\mathrm{EE}}) \sin(4\Delta\alpha).$$

alpha = Birefringence angle

From these equations it is possible to build estimators for alpha (as long as it is zero these violating terms are excluded)

Constraining Primordial magnetic fields



- Faraday rotation by a primordial magnetic field produce B-mode polarization.
- Kahniashvili et al. (2008). At scales of 1 Mpc and n_B=-2.9, constraints are better than 1nG (extrapolated to present).

Observability of B-modes

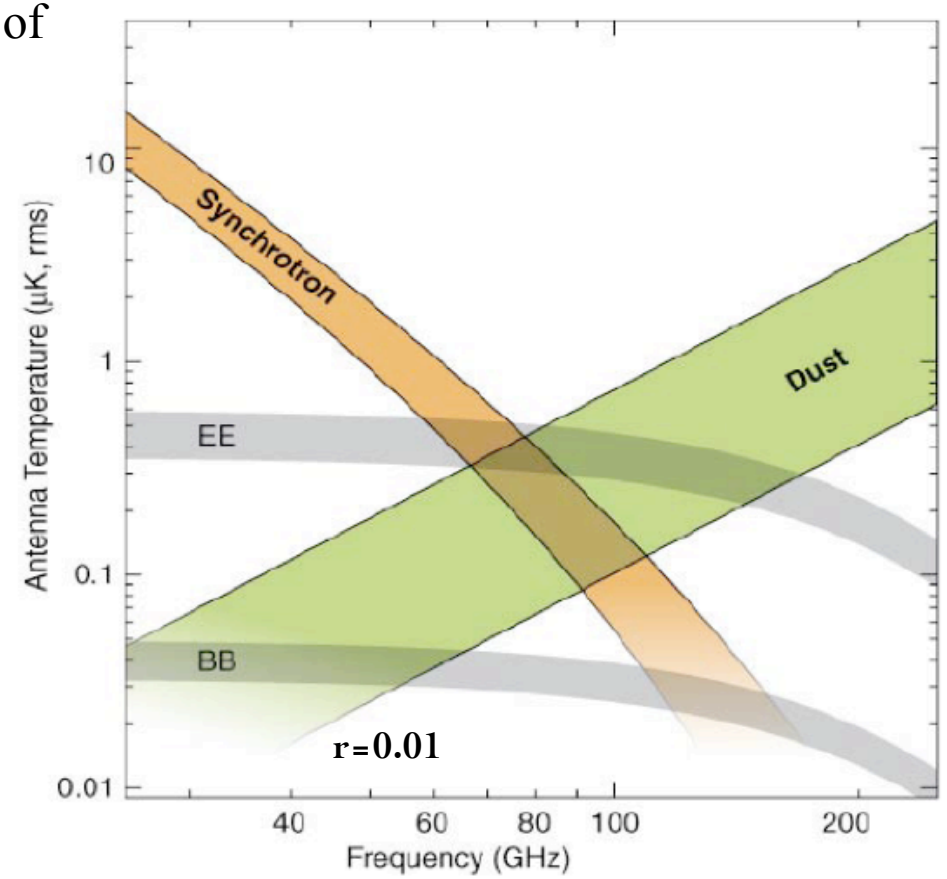
❖ Critical issues:

• **Signals are extremely small** ⇒ large number of receivers with large bandwidths are required.

• Accurate **control of systematics** (cross-pol, spillover,...) is mandatory.

• **Foregrounds**. B-mode signal is subdominant over Galactic foregrounds

- Free-free, low-freq, not polarized
- Synchrotron, low-freq, pol ~10%
- Thermal dust, high-freq, pol ~10%
- Anomalous emission, 20-60 GHz, pol ~3%?
- Point sources, low-freq, pol ~5%



GRAVITY WAVE GENERATION BY VORTICAL FLOWS IN A ROTATING FRAME

Rupert Ford

Churchill College
Cambridge

A dissertation submitted for
the degree of Doctor of Philosophy
at the University of Cambridge

October, 1993

To my grandparents

Margaret and Ted

Preface

This thesis is the result of my own work, and includes nothing which is the outcome of work done in collaboration. Except where stated otherwise, the research reported in this thesis is original. No part of it has been submitted for a degree or any similar qualification at any other university.

This thesis would not have been possible without the enthusiasm, suggestions and encouragement of my supervisor, Professor Michael E. McIntyre. His enduring love of fluid dynamics will remain an inspiration to me for the rest of my life. All members of the atmospheric dynamics group have, at some time, made a valued contribution to my academic life in DAMTP. I would particularly like to thank Peter Haynes for discussions about instability theory, David Dritschel for discussions about vortex dynamics and conformal mappings, and Saravanan for suggestions about numerical modelling. Stephen Cowley generously allowed me use of his computational resources to enable me to complete the simulations in chapter 4 quickly. I would also like to thank the people who have given so generously of their time over the last few months in reading the manuscript and suggesting improvements to it. They are Ben Edgington, Stefan Llewellyn-Smith, Gudrun Magnusdottir, John Scinocca and David Tan — my deepest thanks to you all. I would also like to thank all other members of DAMTP, research staff, computer officers and secretaries, who have helped me in any way.

Throughout my thesis work I have profited from discussions with many scientists at other institutions. Particular thanks go to Glenn Flierl, Joe Keller, Bernard Legras, Lorenzo Polvani, Rich Rotunno, Ted Shepherd and Joe Tribbia. During the summer of 1992 I was a fellow at the Woods Hole Summer School in Geophysical Fluid Dynamics. I thank the organisers and participants for an enjoyable and rewarding summer.

For financial and travel support I thank the Natural Environment Research Council, and for travel support I thank additionally Churchill College and the European Geophysical Society. Finally, I would like to thank my parents for getting me this far, and for their unending love.

Contents

1 Introduction **1**

1.1 Balance and Imbalance 2

1.2 Potential vorticity inversion 6

1.3 Slow manifold hypothesis 14

1.4 Approach of this thesis 18

2 Asymptotic analysis at low Froude number **21**

2.1 Introduction 22

2.2 Gravity wave generation as a singular perturbation problem 28

2.2.1 Perturbation expansion for the vortical flow 30

2.2.2 Dynamics of the wave zone resulting from the vortical flow 39

2.2.3 The effect of wave radiation on the vortical flow 42

2.3 Implications for balance and potential vorticity inversion 44

2.4 The effect of gravity wave radiation upon a rotating elliptical patch of uniform potential vorticity – an explicit example of generalized adjustment 48

2.4.1 Evolution of the vortex boundary 49

2.4.2 The solution at leading order in F 51

2.4.3 The solution to $O(F^2)$ 53

2.4.4 V-states for weakly divergent flow 57

2.4.5 The effect of gravity wave radiation 61

2.5 Discussion 66

2.6 Appendix A 68

2.6.1 Region I 70

2.6.2 Region II 72

2.6.3 Region III 72

2.7 Appendix B 73

3 Instability of vortices and jets **75**

3.1 Introduction 76

3.2 Instability of a circular vortex with discontinuous potential vorticity 81

3.2.1 The basic state 81

3.2.2 The disturbance equations 82

3.2.3 The limit $F \ll 1$ 83

3.2.4 Numerical eigenvalue calculations 89

3.2.5 The limit of large mode number 118

3.2.6 The Sozou modes 129

3.3 Instability of a circular vortex with smooth potential vorticity 133

3.3.1 Disturbance equations 133

3.4 Relevance to the existence or non–existence of a slow manifold 143

3.5 Instability of a parallel jet 148

3.5.1 The basic state 149

3.5.2 Ripa's theorem 149

3.5.3 The disturbance equations 151

3.5.4 Numerical technique 152

3.5.5 Results of the eigenvalue calculations . . . 153
3.5.6 The limit of large wavenumber . . . 155
3.5.7 The limit $q \to \infty$ . . . 157
3.5.8 Eigenvalue calculations for large q . . . 164
3.5.9 Remarks . . . 166
3.6 Discussion . . . 166

4 Nonlinear simulation of gravity wave generation 169

4.1 Introduction . . . 170
4.2 A "Lighthill" theory of gravity wave generation for a periodic parallel flow . . . 177
4.3 Pseudoenergy and pseudomomentum for the shallow water equations . . . 180
4.4 The numerical model . . . 183
4.5 Overview of numerical simulations . . . 186
4.6 Jets with moderate Rossby number and variable Froude number . . . 191
4.7 Cyclonic strips at large Rossby number . . . 210
4.8 Anticyclonic strips with non-negative potential vorticity . . . 223
4.9 Anticyclonic strips with negative potential vorticity . . . 234
4.10 Discussion . . . 248

5 Discussion 253

5.1 Review of aims and objectives of this thesis . . . 254
5.2 Suggestions for further research . . . 256

6 References 258

Chapter 1

Introduction

1.1 Balance and Imbalance

> *"The striking errors in the 'forecast', which has been obtained by computing forms, may be traced back to the large apparent convergence of wind. It may be asked whether this spurious convergence arises from the errors of observations with balloons, or from the finite horizontal differences being to large, or thirdly from the process by which winds at points, arranged in a rectangular pattern, are interpolated between the observing stations ..."*

Thus Richardson (1922) first gave prominence to the idea that inaccuracies in the data used to initialize his pioneering numerical forecast could have given rise to the failure of that forecast. The failure consisted of large spurious variations in the computed pressure field over a 6 hour forecast interval. In fact, Richardson had forecast a pressure variation in one location of 145mb in 6 hours, whereas the actual change in pressure over that period had been of the order of 1mb.

The phenomenon that Richardson had come upon is well known today both to theoreticians and to practitioners of numerical weather prediction - that large scale atmospheric flow away from the equator is, to a good first approximation, in a state of geostrophic balance, in which the Coriolis and pressure gradient terms constitute the principal balance in the horizontal momentum equation.

It now seems almost certain (Platzman 1967, 1968) that Richardson developed some appreciation of this fact in the following twenty years, and that the difficulty had even been foreseen by Margules (1904). Eventually, however, it was when Charney (1951) applied his recently-developed "quasi-geostrophic" theory (Charney 1947, 1948, 1949) to numerical forecasting that forecasts were able to predict an evolution of fields at least qualitatively in line with observations. Even more remarkable in Charney's (1948) paper is his far-sighted remark that the conservation of potential temperature and potential vorticity on fluid particles may alone be sufficient to determine the nature of large scale atmospheric motions. In his words

> *"The motion of large-scale atmospheric disturbances is governed by the laws of conservation of of potential temperature and potential vorticity, and by the conditions that the horizontal velocity be quasi-geostrophic and the pressure be quasi-*

hydrostatic."

In present day terminology, adopted in this thesis, the quasi-geostrophic and quasi-hydrostatic conditions are "balance" conditions, which enable one to determine the velocity and temperature fields knowing only the mass under each isentropic surface, and the distribution of potential vorticity on each isentropic surface. The resulting system of equations is referred to as a balanced model.

Some appreciation of the importance of potential vorticity had already been developed in the ten years preceeding Charney's work. Rossby (1936), Rossby *et al.* (1939) and Ertel (1942) had been independently developing some fundamental understanding of the dynamical implications of the conservation of potential vorticity on fluid particles, although they did not postulate that it could be "inverted" to give other dynamical information, as Charney (1948) had noted in his "quasi-geostrophic" approximation. The first explicit application of potential vorticity inversion was made by Kleinschmidt (1950a,b;1951,1955,1957), who used an exact form of potential vorticity inversion for axisymmetric structures in his studies of upper level cyclones.

The potential vorticity Q is defined by

$$Q = \rho^{-1}(2\boldsymbol{\Omega} + \nabla \times \boldsymbol{u}).\nabla\theta, \tag{1.1}$$

where ρ is the air density, Ω is the Earth's angular rotation vector, $\boldsymbol{u}$ is the fluid velocity in coordinate axes fixed on the Earth, and θ is the potential temperature (or any function of it). In the absence of friction or diabatic processes, Q is a materially conserved quantity:

$$\frac{DQ}{Dt} = 0. \tag{1.2}$$

Although the hydrostatic approximation is valid for large scale motion throughout the atmosphere, the quasi-geostrophic approximation is seldom sufficiently accurate to render it of practical use in numerical weather prediction. On the other hand, the absence of any balancing procedure is known to give rise to large errors of the type encountered by Richardson. Consequently, balanced equations have been developed for atmospheric modelling which are more accurate than the quasi-geostrophic equations, while still retaining the robustness of a balanced system of equations, being insensitive to errors in the initial data. Indeed, although the "quasi-

geostrophic" approximation remains the best known example of a balanced system, many others have been proposed, each with varying degrees of analytical tractibility and practical accuracy.

At this point, it is essential to be clear as to exactly what is required of a balanced model. Following Hoskins *et al.* (1985), we may define a balanced model by prescribing a method for determining the velocity field everywhere given only the mass under each isentrope (θ-surface), the potential vorticity distribution on each isentrope, and the θ distribution on the boundary of the domain. The proposed "inversion" procedure should not depend on knowledge of the PV distribution at any time other than the current time. The integration forward in time proceeds in principle by advecting the potential vorticity according to (1.2).

Before proceeding, in §1.2, to a discussion of how one might set about inverting the potential vorticity to obtain the other dynamical fields, it is instructive to consider the features of the flow which no balanced model, however accurate, can capture.

To fix ideas, let us consider the shallow water equations on an f-plane:

$$\frac{D\boldsymbol{u}}{Dt} + f\boldsymbol{k} \times \boldsymbol{u} + g\nabla h = 0 \tag{1.3}$$

$$\frac{Dh}{Dt} + h\nabla.\boldsymbol{u} = 0, \tag{1.4}$$

Firstly, we consider the dispersion relation for the linearized system corresponding to (1.3,1.4):

$$\omega(\omega^2 - f^2 - c_0^2(k^2 + l^2)) = 0, \tag{1.5}$$

where ω is the wave frequency and $c_0 = (gH)^{1/2}$ is the gravity wave phase speed. There are three roots. One root, $\omega = 0$, corresponds to "slow" waves, which are the analogue of Rossby waves in this system. Their frequency would be modified in the presence of a background planetary vorticity gradient. The other two roots correspond to propagating waves only if $\omega > f$. These correspond to "fast" surface inertio-gravity waves, which owe their existence to the restoring effect of gravity on the free surface, and whose frequency is modified by background rotation.

In a balanced model, however, there is only one time derivative, resulting from the advection of potential vorticity. The resulting dispersion relation will therefore have only one root, which must, by symmetry, correspond to some appropriate generalization of the slow Rossby wave. A balanced model is therefore a model which filters from the equations of motion gravity waves, which are generally of much greater phase speed than typical fluid velocities.

Moreover, since a balanced model contains two fewer time derivatives that the full system, the inversion procedures required by a balanced model to obtain the velocity and height fields from the potential vorticity can in general be defined by imposing two constraints on the full shallow water system. The selection of the two constraints must be made carefully, however, if the model is still to possess analogues of potential vorticity, which should be conserved on fluid particles, and analogues of total mass, energy, and enstrophy, which should be conserved globally. Indeed, several popular balanced models, such as the Balance Equations (Lorenz, 1960) in the form in which they are usually integrated, do not possess a full set of conservation laws. However, from a theoretical viewpoint, it nonetheless seems clear that the unifying concept in any discussion of balanced dynamics must be that of advection of potential vorticity on fluid particles, together with the specification of sufficient constraints that the potential vorticity can be inverted to give all the other dynamical information necessary to proceed with the integration.

A second thought experiment which one can conduct, and which also illustrates that there cannot be freely propagating gravity waves in a balanced model, is to consider the classical "Rossby adjustment" problem (see, e.g., Gill, 1982). In this thought experiment we consider a rotating shallow water fluid consisting of two semi-infinite regions of uniform potential vorticity of different values, which intersect along a straight line of infinite length. If we suppose that initially there is no motion then, by considering the potential vorticity, there must be a discontinuity in surface elevation along the line of intersection. Now suppose that ultimately the fluid evolves to remove the surface discontinuity. By symmetry, the final state must also be one of two semi-infinite regions of uniform potential vorticity intersecting along a straight line of infinite length, but the energy associated with this final state is always less than the energy associated with the initial state.

Now, since the potential vorticity is uniform everywhere in the fluid away from the interface, its conservation on fluid particles reduces the number of independent time derivatives away from the interface from three to two in the fluid away from the interface. The resulting waves in these regions are therefore gravity waves, which can propagate away from the interface to infinity, and account for the loss of energy. To a balanced model, however, which depends only upon knowledge of the potential vorticity field, both initial and final states look exactly the same, and would possess the same velocity and height fields. There is no Rossby adjustment, and no

gravity wave radiation.

1.2 Potential vorticity inversion

Over the years, many balanced models have been proposed for use in different circumstances. In this section, we start by deriving two very simple balanced models via scaling arguments, and proceed to discuss different ways in which balanced models can be constructed.

To begin, let us return briefly to the shallow water equations, this time on a β-plane:

$$\frac{D\boldsymbol{u}}{Dt} + f\boldsymbol{k} \times \boldsymbol{u} + g\nabla h = 0 \tag{1.6}$$

$$\frac{Dh}{Dt} + h\nabla.\boldsymbol{u} = 0, \tag{1.7}$$

in which the Coriolis parameter f represents $f_0 + \beta y$, where f_0 and β are constants. Let us further assume that $\beta y \ll f_0$ (in a sense to be defined), so that the Coriolis parameter does not change much over the range of latitudes which we are considering (these assumptions are made here to facilitate a pedagogical development of a "potential vorticity view" of quasi-geostrophy, and will be removed in subsequent discussions of balance). The potential vorticity is defined as

$$Q = \frac{f + \zeta}{H + h}, \tag{1.8}$$

where H is the mean layer depth and h represents departures of the total layer depth from the mean; and ζ is the relative vorticity. Let us now assume, as is implicit in the scalings of Charney (1948), that Q has only small departures from its background value f_0/H. This requires that the Rossby number $Ro = U/fL \ll 1$, so that

$$Q \sim \frac{f_0}{H} + \frac{\zeta + \beta y}{H} - \frac{f_0 h}{H^2}. \tag{1.9}$$

Then the advective timescales D/Dt must be long compared with the inertial timescale f_0^{-1}, so the shallow water momentum equation (1.6) gives

$$f_0\boldsymbol{k} \times \boldsymbol{u} + g\nabla h = 0 \tag{1.10}$$

at leading order. Taking the curl of this equation implies $\nabla.\boldsymbol{u} = 0$. Therefore we can introduce a streamfunction ψ such that $\boldsymbol{u} = \boldsymbol{k} \times \nabla\psi$, and from (1.10) we have $h = (f_0/g)\psi$. The

approximation (1.9) to the potential vorticity becomes

$$Q \sim \frac{1}{H}\left(f_0 + \beta y + \nabla^2\psi - \frac{f_0^2}{gH}\psi\right). \tag{1.11}$$

Scaling arguments show that f_0 is an order of Ro^{-1} larger than either of the terms involving ψ in (1.11), which are themselves of the same order of magnitude, and hence the effects of local potential vorticity anomalies and the background planetary vorticity gradient enter at the same order in the quasi-geostrophic scaling if $\beta y/f_0 = O(Ro)$.

It follows that, knowing Q, one can in principle invert (1.11) with evanescence or no normal flow boundary conditions to give ψ, and hence $\boldsymbol{u}$ and h.

An alternative, but equally valid, balance condition is to relax the assumption that potential vorticity perturbations are small compared with the background value f_0/H, but to continue to assume, as is implicit in the quasi-geostrophic scaling, that the Froude number $F = U/(gH)^{1/2} \ll 1$. The implication of this assumption (see chapter 2 for details of the scaling analysis) is that $h/H = O(F^2)$. This means that height perturbations are small, the flow is approximately non-divergent, and the potential vorticity perturbations are dominated by the relative vorticity ζ and variations in the planetary vorticity βy, which is not now required to be small compared with f_0. Once again we may introduce a streamfunction ψ, but now the potential vorticity becomes

$$Q \sim \frac{1}{H}\left(f_0 + \beta y + \nabla^2\psi\right). \tag{1.12}$$

The result is the Euler equations for a two-dimensional incompressible perfect fluid. Again, decay of velocity to infinity or no normal flow boundary conditions can be used to invert (1.12) to give velocity and height fields.

The quasi-geostrophic and Euler equations have been widely used for both analytical and numerical work. Analytically, and from the perspective of the "contour dynamics" algorithm (Dritschel, 1989, 1993; Pullin 1992; Dritschel & Saravanan 1993) they are attractive because a linear operator (Laplacian or Helmholtzian) must be inverted to obtain the other dynamical fields from the potential vorticity. Along with the Planetary Geostrophic or "thermocline" equations (Robinson & Stommel, 1959), in which the effects of inertia are neglected entirely, they are the only balanced models to have been proposed in which only linear operators need be inverted. From the perspective of numerical methods other than contour dynamics, their

advantage also derives from their simplicity. Indeed, few other balanced models can be integrated as quickly as the full shallow water equations, provided the latter are integrated using a large implicit time-step (Allen, personal communication).

The disadvantage of these simple balanced models lies in that they are only leading order approximations, and it is quite possible that one could proceed to higher order in Rossby or Froude number before nonlinear coupling between vortical motions and gravity waves led inevitably to the appearance of freely propagating gravity waves in the solution, and made any further improvements fundamentally impossible.

Indeed, the quasi-geostrophic approximation places two additional constraints upon the flow: the Coriolis parameter can vary only by an amount of order Rossby number throughout the entire domain of interest, and the surface elevation must have a gradient of the order of the Rossby number. These restrictions can be lifted by using the semi-geostrophic approximation (Salmon, 1985), which is also a low Rossby number approximations, formally valid to the same order in Rossby number as the quasi-geostrophic approximation. An interesting derivation of the semi-geostrophic system with variable Coriolis parameter and order-one surface gradients was given by Salmon (1988a), in which he showed that the semi-geostrophic equations can be derived from the shallow water system by making an approximation to the symplectic tensor in a constrained Hamiltonian, whereas an equivalent derivation of the quasi-geostrophic equations required the arbitrary introduction of a reference state and a metric tensor. The Euler equations were also derived via the constrained Hamiltonian approach, making approximations in the symplectic tensor. This method ensures that any balanced equations which are derived possess a full set of conserved quantities.

The essential point is that, although in principle the quasi-geostrophic system had solved the basic problem of unbalanced fields which Richardson had encountered, by the mid-1950s it was becoming clear that something more accurate was needed for practical purposes. A significant development in balanced models occurred when Charney (1955) proposed a balanced system with a higher order of approximation than the quasi-geostrophic system.

For his new system, he assumed that the flow was non-divergent ($\nabla.\boldsymbol{u} = 0$). Taking the divergence of the momentum equation (1.6), and applying $\nabla.\boldsymbol{u} = 0$, one obtains a second

constraint:

$$\nabla.(f\nabla\psi) + 2\left(\frac{\partial^2\psi}{\partial x^2}\frac{\partial^2\psi}{\partial y^2} - \left(\frac{\partial^2\psi}{\partial x\partial y}\right)^2\right) = \nabla^2 h. \tag{1.13}$$

The two constraints, (1.13) and $\nabla.\boldsymbol{u} = 0$, thus constitute a balanced model.

Condition (1.13) was also one of the balance constraints used by Lorenz (1960) in his derivation of what have now become known as the Balance Equations. However, although he used only the rotational component of the velocity in his balance constraint (1.13), some divergence remains in his system. The divergence is obtained via a diagnostic relation, and is used in the prognostic equation for the vorticity. The system constitutes a balanced model, since there is one prognostic equation and two constraints, although vorticity, rather than potential vorticity, has typically been used as the prognostic variable of integration in numerical work (Norton *et al.*, 1986).

McWilliams (1992) presents a scaling analysis to show that (1.13) is formally valid when $F^2 \ll 1$ and $F^2/Ro \ll 1$. At high Rossby number, the condition is simply one of small divergence, and at leading order (1.13) reduces to the barotropic vorticity equation. At small Rossby number one recovers the quasigeostrophic scaling, and (1.13) reduces to quasigeostrophic balance.

The Balance Equations have been the focus of several numerical studies over the last decade. Norton *et al.* (1986) show how the balance equations may be implemented in a numerical scheme, and McWilliams *et al.* (1986) present a numerical study of oceanic vortices on a β plane, in which there is practically no discernable difference between the potential vorticity fields of the Balance Equation integrations and the shallow water integrations. In their cases, Rossby numbers are of order one, but they consider vortices which are either barotropic or have at most two baroclinic modes. The structures are thus rather deep, and their Froude numbers are consequently rather small, so explaining the good agreement between Balance Equation and shallow water integrations.

An extensive numerical and analytical investigation of the properties of a large number of different balanced models has been presented by Allen *et al.* (1990a,b) and Barth *et. al* (1990). In addition to considering existing balanced models, they also present some models of their own. In interpreting their results, however, it should be noted that their models impose only one, not two, constraints, on the dynamics, and two fields must be integrated in the subsequent

evolution (albeit that one is integrated using an implicit time step). Their models should not therefore be regarded as balanced models in the sense described above, although their dynamics are certainly constrained in some way. For the purposes of this discussion, I shall refer to these as "semi-balanced" models.

The first paper (Allen *et al.*, 1990a), presents the models they considered, and investigates properties of their stationary solutions and coastal Kelvin waves. Following Gent & McWilliams (1983a), they use boundary conditions in such a form that coastally trapped Kelvin waves are retained in all balanced models except the quasi-geostrophic model. In the subsequent analysis of the performance of the different models, presented in the second and third papers, they take as a measure of model accuracy the root mean square departure of the surface height from that of the corresponding shallow water integration.

The second paper (Barth *et al.*, 1990) investigates the evolution when the models are initialized with sinusoidal flow over a Gaussian topography in a doubly periodic geometry. The main conclusion is that the classical Balance Equations perform best - almost as well as the shallow water equations. One of the "semi-balanced" models constructed by Allen *et al.* (1990a) performs about equally well though, for reasons described above, should probably not be regarded as a balanced model. The quasigeostrophic model invariably performs the worst, all the more so when the topography is of order one height.

The third paper (Allen *et al.*, 1990b) is similar in spirit to the second, but the geometry is now a confined channel, which makes desirable the inclusion of coastal waves in as many balanced models as will support them. To satisfy the channel wall boundary conditions, the initial flow is uniform, rather than sinusoidal. Similar results on the relative accuracy of the various balanced models are found here when compared with those presented in the second paper, although in this case one of the "semi-balanced" systems performs better than the Balance Equations, and even better than the shallow water equations with a large implicit time step.

Spall & McWilliams (1992) addressed the question of whether there might be significantly more accurate balance schemes than the Balance Equations by comparing integrations of the Balance Equations and of a higher order balance scheme with integrations of the shallow water equations initialized with the same (balanced) fields in a doubly-periodic domain. They define the "unbalanced" amplitude for any field as the difference between the balanced model inte-

gration and the shallow water integration with the same balanced initial fields. In their study, they find that, at very low Rossby numbers, the higher order balance scheme initially gives rise to less energy in the unbalanced motions than the Balance Equations, but in time both orders of balance lead to the same (very small) degree of energy in the unbalanced flow. At higher Rossby number, agreement between the two different orders of balance is even closer for earlier times, and remains good for the whole integration, but both model integrations depart significantly further from the corresponding shallow water integrations than was the case at low Rossby number.

They present two different scaling analysis to investigate the nature of the "unbalanced" flow, which predict how its amplitude should scale with Froude and Rossby numbers. At low Rossby numbers, the unbalanced flow is consistent with a single scale analysis, which assumes that it has length and time scales of the same order as the balanced flow. The so-called unbalanced flow therefore has characteristics very similar to the balanced flow, and it seems plausible that it might be possible to obtain it as a higher order correction to balance in a more accurate balanced model. At high Rossby numbers, however, the amplitude of the unbalanced flow is consistent with a multiple scale analysis. The multiple scale analysis assumes that the unbalanced flow is dominated by free gravity waves, which appear to inherit the length scale of the balanced motions, but possess significantly higher frequencies. They conclude that at low Rossby numbers further corrections may be possible to their balance schemes, since the dynamics which is not captured by the balanced models seems nonetheless similar in character to the balanced dynamics. On the other hand, at high Rossby numbers the principal component of the unbalanced flow is similar for both balanced models, but is essentially different from the balanced dynamics. The principal unbalanced flow in this case corresponds to freely propagating gravity waves. Such motions are, for reasons described in §1.1, essentially unbalanced, and can almost certainly not be slaved to the potential vorticity evolution, as is required for a balanced model.

Charney balance (condition (1.13) with $\nabla.\boldsymbol{u} = 0$) was the basis for a study of a series of balanced models by Norton (1988), and McIntyre & Norton (1993). Their hierarchy of balanced models was constructed by successively eliminating higher order Eulerian or Lagrangian time derivatives of the divergence field. Their "first order" balance sets the divergence field and its first Eulerian time derivative to zero, and is therefore equivalent to Charney's (1955) balance

scheme. Their "second order" balance sets the first and second time derivatives of the divergence field to zero, their "third order" the second and third time derivatives to zero, and so on. With two constraints now supplied, they use the advection of potential vorticity as their temporal evolution equation, and at every time-step the velocity fields required to advect the potential vorticity are obtained through instantaneous nonlinear elliptic inversion operators acting on the potential vorticity field.

They present integrations of a forced polar vortex in a single layer of shallow water on a hemisphere, with balanced initial conditions. They compare their integrations using successive order of balanced models with integrations of the full shallow water equations, and find that successive models up to the third order of their balanced models achieve improvements over lower orders in the hierarchy, but that their fourth order model seldom achieved noticeable improvements over their third order one. Boundary conditions imposed at the equator are equivalent to imposing antisymmetry. The simulations do not therefore have physical boundaries at the equator, and any equatorial Kelvin waves which develop must be regarded as part of the unbalanced flow (for a discussion of equatorial Kelvin waves in balanced models, see Gent & McWilliams, 1983b).

Comparing the divergence field from the shallow water and balanced integrations, they find that, aside from small differences in the magnitude of features which are common to both integrations, the most significant difference between balanced and shallow water simulations is a long equatorial Kelvin wave, trapped at the hemispherical "boundary", but apparently propagating freely around the equator, with a group velocity consistent with the dispersion relation for equatorial Kelvin waves, and not apparently slaved to the vortical motions. This Kelvin wave is not present in any of their balanced integrations, but is present in almost all the corresponding shallow water integrations. This fact should not be surprising in view of the discussion of the inadmissibility of inertio-gravity waves in balanced models at the end of §1.1. What might be regarded as surprising, however, is that the wave has exceptionally low amplitude (at most about 10% of the maximum divergence in the flow), even when the Rossby numbers are of order one throughout a large region of the domain, and the Froude number reaches a maximum value of 0.7. It seems that the coupling between vortical motions and free gravity waves tends to be rather weak, even when there are no obvious small parameters in the flow, and hence, for systems without a large initial amount of free gravity wave activity, free

gravity waves develop in the flow only to very low amplitude. On the other hand, the fact that the fourth order inversion scheme does not perform significantly better than the third, and the ubiquitous emergence of a well-defined Kelvin wave in the shallow water simulations, suggest, but do not prove, that there is some ultimate limitation to potential vorticity inversion and balanced dynamics. We shall return to this point in §1.3, and again in chapter 3.

A somewhat different framework of balance was also considered by Norton (1988), following on the ideas of Non Linear Normal Mode Initialization, developed independently by Machenhauer (1977) and Baer & Tribbia (1977). The technique consists of first decomposing the equations into linear normal modes. The full equations are then cast in the framework of the normal mode decomposition, and a slaving relation derived to bound the gravity modes to the Rossby modes, on the assumption that the natural frequencies of the gravity modes are large compared with typical inverse time scales for the flow. A succession of balance inversion operators can be derived by this procedure. The zeroth order normal mode inversion consists in setting the amplitudes of the gravity wave mode amplitudes to zero. The first order inversion sets the time derivative of the gravity modes to zero. The second order inversion, obtained by Tribbia (1984), effectively sets the second time derivative of the gravity modes to zero. Like McIntyre & Norton's (1993) primitive equation inversions, this procedure can proceed indefinitely although, also like McIntyre & Norton's procedure, it will quickly become impractical as its complexity increases.

It has occasionally been suggested that the generation of gravity waves by vortical motions could be thought of as a generalization of the Rossby adjustment process discussed in §1.1, and that construction of an accurate balanced model might require the adjustment to be parametrized in some way. Although there are some difficulties with the idea, which will become clear as the thesis develops, two interesting proposals have developed from it.

In a recent paper, Vallis (1992) suggested that the production of gravity waves by vortical motions might be parametrized as a form of generalized Rossby adjustment. Following this idea, he proposed a balanced model in which the total energy of the system is minimized, subject to the condition that the potential vorticity field remain unchanged. Two constraints can be derived from the associated Euler-Lagrange equations, which can be used to define a balance system. Curiously, however, the method does not reconstruct steady axisymmetric solutions of the shallow water equations, which is perhaps related to the fact that it does not respect the

Casimir invariants of the shallow water system.

Another proposal (T. G. Shepherd, personal communication), is to divide the system into fast and slow components, very much like the normal mode decomposition. In the case of a canonical Hamiltonian system, in which fast variables have only fast conjugate variables, and slow variables have only slow conjugate variables, it is clear that, regarding the slow variables as fixed, the fast variables on their own constitute a reduced Hamiltonian system. To parametrize the adjustment process one might think of extremizing the "fast" Hamiltonian with respect to the fast variables, holding the slow variables fixed, since they are not supposed to change appreciably while the adjustment process is taking place. This procedure provides a slaving relation for the fast variables to the slow variables.

In truncated models of the atmosphere, there are typically twice as many gravity modes as Rossby modes. The reason for this is the presence of a Casimir density - the potential vorticity - which makes the full system non-canonical and, in the case of the shallow water system, reduces its order from 4 to 3. The canonical description involves either particle positions and momenta (two of each), or three Clebsh potentials and height (Salmon, 1988b). Truncated models of the atmosphere tend to inherit the non-canonical structure of their infinite dimensional counterpart (Bokhove, personal communication), and it is no longer clear whether the fast variables constitute a Hamiltonian system by themselves, since the existence of a Hamiltonian for the fast variables, which could be extremized to parametrize an adjustment process, is not assured if the full system is non-canonical. If Darboux' theorem is invoked to make the system canonical, it is not clear that the canonical variables will separate into canonically conjugate fast variables and canonically conjugate slow variables.

1.3 Slow manifold hypothesis

In assessing different schemes for potential vorticity inversion, with different properties with respect to accuracy, analytical tractibility and attempts to parametrize gravity wave generation in a physically motivated way, one is constantly left to wonder whether there might not be some more fundamental property of the shallow water equations involved – whether they might not possess a true "slow manifold", on which the entire flow is determined exactly by the potential vorticity field, and which, once on, the solution will remain on for all time.

The origination of the "slow manifold" hypothesis originated in a paper of Leith (1980), in which he postulates the existence of a manifold $\mathcal{M}$ with a phase space one third the dimension of the phase space of a full normal mode decomposition of the shallow water equations, for which, in his words

> *"any state vector in $\mathcal{M}$ would continue to move slowly in it, while any state vector not in $\mathcal{M}$ would oscillate about it".*

The implication of the existence of a slow manifold for any finite dimensional system is that the gravity wave modes can be slaved exactly to the Rossby wave modes. Similarly, for an infinite dimensional system, there would exist elliptic inversion operators acting on the instantaneous potential vorticity field such that, if the velocity and height fields associated with a given potential vorticity field were consistent with the inversion operators at an initial instant in time, then they would remain consistent with the given inversion operators for all subsequent times under evolution of the full (shallow water) system. A slow manifold can then be regarded as an exact balance, to which all known balanced models are approximations.

Questions regarding the existence or non-existence of a slow manifold have almost always been posed within the context of low order models. There are obvious conceptual simplifications in doing so. In particular, the fields are typically broken down into $3 \times N$ normal modes of the system. One third of the modes are "Rossby" modes ($\mathcal{R}_i$), and two thirds are gravity modes ($\mathcal{G}_i$). The evolution of the $\mathcal{R}_i$ and the $\mathcal{G}_i$ is given by

$$\frac{d\mathcal{R}_i}{dt} + \Omega_i \mathcal{R}_i = \mathcal{N}_{\mathcal{R}_i}(\mathcal{R}, \mathcal{G}) \tag{1.14}$$

$$\frac{d\mathcal{G}_i}{dt} + \omega_i \mathcal{G}_i = \mathcal{N}_{\mathcal{G}_i}(\mathcal{R}, \mathcal{G}) \tag{1.15}$$

where the $\mathcal{N}_{\mathcal{R}_i}$ and $\mathcal{N}_{\mathcal{G}_i}$ are nonlinear functions required to account for the nonlinear interaction of Rossby modes $\mathcal{R}_i$ and gravity modes $\mathcal{G}_i$, and no summation over repeated indices is intended.

The existence of a slow manifold for this system is equivalent to the existence of functions $\mathcal{F}_i$ for $i = 1, \ldots 2N$, such that, given suitable initial conditions, the gravity wave modes can be slaved exactly to the Rossby wave modes by the functional relationship

$$\mathcal{G}_i = \mathcal{F}_i(\mathcal{R}_1, \ldots, \mathcal{R}_N). \tag{1.16}$$

On the slow manifold, we would have

$$\frac{d\mathcal{G}_i}{dt} \equiv \frac{d\mathcal{F}_i}{d\mathcal{R}_j}\frac{d\mathcal{R}_j}{dt}, \tag{1.17}$$

where summation is understood for $j = 1, \ldots, N$, and hence for a slow manifold to exist there would equivalently have to be exist a solution to the nonlinear slaving relation

$$\sum_{j=1}^{N} \frac{d\mathcal{F}_i}{d\mathcal{R}_j}\left(\mathcal{N}^{\mathcal{R}}{}_j(\mathcal{R},\mathcal{F}) - \Omega_j\mathcal{R}_j\right) + \omega_i\mathcal{F}_i = \mathcal{N}^{\mathcal{G}}{}_i(\mathcal{R},\mathcal{F}) \tag{1.18}$$

for $i = 1, \ldots, 2N$.

The suitable boundary conditions are that the amplitudes of the gravity wave modes $\mathcal{G}_i$ should tend to zero as the amplitudes of the Rossby wave modes $\mathcal{R}_j$ tend to zero. The leading order of the solution should then be equivalent to a quasi-geostrophic normal mode expansion. If this is not assured, then any invariant manifold found could not really be thought of as a slow manifold, since it would be dominated by gravity wave activity in the low Rossby number limit.

A further subtlety is that the the slow manifold, if it exists, might not exist over the whole range of $\mathcal{R}$–space, but rather might only exist in some region $\mathbf{M}$ of Rossby mode phase space. The important point is that, if this is so, any trajectory initially in $\mathbf{M}$ must remain in $\mathbf{M}$. It is not sufficient, for example, to be able to define a slow manifold in the neighbourhood of the origin if trajectories can then depart from that neighbourhood, perhaps to re-enter it not on the slow manifold.

Lorenz (1980) considered a 9 component model based on the shallow water equations on the f-plane. He found that for sufficiently small Rossby numbers the system appeared evolve towards a three-dimensional attractor, which corresponded approximately to quasi-geostrophic balance. Subsequently, he showed (Lorenz, 1986) that an unforced 5-component system, with three Rossby wave components and two gravity wave components, possessed a slow manifold. The proof, however, relied on an additional symmetry of the equations, in part an artifact of the small number of gravity wave components in the system, and it was not clear that the nine-component model used in Lorenz (1980), or a continuous fluid model, would possess sufficient symmetries to allow similar construction of a slow manifold..

Experiments with 9 component systems by Errico (1982), Warn & Ménard (1986), and Vautard & Legras (1986) have suggested that gravity wave generation is an almost universal

property of models with the appropriate ratio of gravity wave modes to Rossby wave modes, and the result of Lorenz (1986) should be regarded as somewhat atypical of low order models.

There are various other ways in which simple models can be made to have slow manifolds. Tribbia (personal communication) has investigated a model in which the Rossby wave modes force the gravity wave modes, but the gravity wave modes do not react back upon the Rossby wave modes. It follows at once that the Rossby modes can be integrated on their own, and the gravity modes can then be integrated *a posteriori*. Imposing the condition that the gravity mode amplitudes tend to zero as the Rossby mode amplitudes tend to zero determines the gravity modes uniquely, and the model thus possesses a slow manifold, on which the gravity modes can be diagnosed from the Rossby modes, and are not required for the temporal integration of the model. One might in some sense regard the model of a description of how gravity waves are generated by vortical motions, but not including the effect describing how the vortical motions are affected by gravity wave generation. This philosophy has strong parallels with the Lighthill theory of aerodynamic sound generation (Lighthill, 1952), which we shall discuss in chapter 2 but is almost certainly fundamentally different from a realistic model of Rossby wave – gravity wave interaction.

When forcing and dissipation are introduced into simple models, additional difficulties may arise. Lorenz & Krishnamurthy (1987) demonstrated numerically that the 5 component model of Lorenz (1986) did not possess a slow manifold when forcing was introduced and was sufficiently strong. Some doubt has recently been cast on their result by Jacobs (1991), who shows for the same model how to obtain the slow manifold in the neighbourhood of fixed points of the system. His proof, however, relies on constructing analytic solutions by Taylor series, whose radius of convergence vanishes when the forcing and dissipation are turned off, and may be quite small otherwise. The Rossby numbers required for the slow manifold to exist thus depend on the amplitudes of forcing and dissipation. Moreover, phase space trajectories are guaranteed to remain in the region where the slow manifold exists only when the dissipation dominates over forcing and nonlinear evolution, in which case the entire system evolves towards a fixed point. The problem in all other cases is exactly that which was discussed earlier – the trajectories are not assured to remain in the region in which the slow manifold is defined, and may leave the region, perhaps to subsequently re-enter it not on the slow manifold. The study of Lorenz & Krishnamurthy (*op. cit.*) was not confined to the limited regime where all trajectories approach

the fixed point, so explaining the apparent discrepancy between the two papers. This has been discussed in the literature by Lorenz (1992).

1.4 Approach of this thesis

The aims of this thesis are three-fold:

- To investigate the generation of gravity waves in the shallow water system, and the dependence of their amplitudes and character on the Froude and Rossby numbers of the basic flow:
- To quantify the effect of the generation of gravity waves in the shallow water system on the potential vorticity distributions which generate them;
- To establish whether, without additional regard to flow geometry, there can or cannot exist a slow manifold for the shallow water equations in any region of Froude and Rossby number space.

The approach adopted to these questions in this thesis is as follows:

In chapter 2, we analyse the nature of interaction between vortical motions and gravity waves in the limit of low Froude number F and order-one Rossby number Ro. We begin by developing a theory of gravity wave generation by vortical motions. It is straightforward to show that there is a formal equivalence between two-dimensional compressible gas dynamics and the non-rotating shallow water system, in the case where the adiabatic gas exponent γ in the former is taken equal to two. The theory is thus an extension of the existing theory of aerodynamic sound generation, first proposed by Lighthill (1952), in which a compact vortical source generates acoustic waves having the same frequency as the vortical motions, but a wavelength which is greater than the vortical length scale by the order of F^{-1}.

The analysis proceeds via a singular perturbation expansion in F, generalizing the analysis of Crow (1970) to include the effect of background rotation upon both the high order Froude number corrections to the vortical source motion and gravity waves it generates. Within this framework we can also investigate the effect which the radiated waves have on the vortical motions, which is inaccessible via Lighthill's (1952) analysis. The chapter ends with a specific

example of the effect of gravity wave radiation on a rotating ellipse of uniform potential vorticity, which is a consistent leading order solution to the low Froude number dynamics (Kida, 1981; Lamb; 1932), generalizing a study by Zeitlin (1988, 1991).

Chapter 3 is motivated by the observation that when a small disturbance is made to the boundary of an axisymmetric vortex of uniform potential vorticity in the shallow water equations, the low Froude number analysis of chapter 2 predicts that, provided the frequency of the disturbance exceeds the inertial frequency in magnitude, the disturbance at a rate of order F^{2m}, where m is the mode number of the disturbance (Broadbent & Moore 1979; Kop'ev & Leont'ev, 1983, 1985, 1986).

We note that the instability may be regarded as a mixed Rossby-gravity instability in the sense of Sakai (1989), and investigate its growth rates over a range of F—Ro space. In the spirit of Knessel & Keller (1992), we develop a WKBJ analysis to show that the instability persists for all Froude and Rossby numbers of the basic vortex, and can be generalized in a limited fashion to vortices with a continuous and monotonic potential vorticity profile. The implications for the existence of a slow manifold for the shallow water system are discussed.

The analysis for the axisymmetric vortex is compared with a similar analysis for parallel flow, with a single discontinuity of potential vorticity. Ripa's (1983) theorem states that the parallel flow will be stable of the ratio of the potential vorticities in the two layers is less than 4:1. In the final part of chapter 3, I show that the stability boundary provided by Ripa's theorem is indeed the stability boundary of the flow, and that at potential vorticity ratios exceeding 4:1 the flow is unstable to short wave disturbances.

In chapter 4 we return to the Lighthill theory of gravity wave generation, to assess its predictive power in cases where the Froude number is not small. Webster (1970) has investigated the validity of the F^7 power scaling law experimentally for a non-rotating jet in shallow water, and found that it applied for Froude number up to approximately $F = 1$. In this thesis I present a numerical investigation of barotropically unstable jets which generate gravity waves as a result of coupling between vortex motions and gravity waves as the instability saturates and the potential vorticity rolls up.4:1

It is found that, when background rotation is included, the vortical dynamics will tend to become less unsteady as the Froude number is increased. This effect, combined with the fact

that gravity waves can not propagate with frequencies below the inertial frequency, acts to reduce the amplitude of gravity wave radiation at larger Froude numbers, unless the potential vorticity in the strip is zero or of opposite sign to the background potential vorticity.

In the final chapter, I summarise the results obtained in the thesis, discuss possible connections between this work and observed phenomena in atmospheres and oceans, and make some suggestions for further research.

Chapter 2

Asymptotic analysis at low Froude number

2.1 Introduction

It is now more than forty years since Lighthill (1952) proposed his successful theory for predicting the intensity acoustic emissions from turbulent jets. The starting point for any theory of aerodynamic sound generation must be the compressible fluid equations:

$$\frac{\partial}{\partial t}\rho u_i + \frac{\partial}{\partial x_j}\rho u_i u_j + \frac{\partial p}{\partial x_i} = 0 \tag{2.1}$$

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_i}\rho u_i = 0 \tag{2.2}$$

$$\left(\frac{p}{p_0}\right) = \left(\frac{\rho}{\rho_0}\right)^{\gamma}, \tag{2.3}$$

where γ is the adiabatic gas exponent. For a diatomic gas, one takes $\gamma = 1.4$, but for our purposes it is important to note that if the fluid is assumed to be two-dimensional, the choice $\gamma = 2$ yields the (non-rotating) shallow water equations, with density ρ in place of the shallow water height. Gravity waves in the shallow water system are the counterpart of acoustic waves in the compressible gas equations, and it follows that any analysis of sound wave generation by vortical motions in two-dimensional compressible fluids should be equally valid in analysing gravity wave generation by vortical motions in the shallow water equations.

The crucial step of Lighthill's (*op. cit.*) so-called "acoustic analogy" theory is to combine (2.1) – (2.3) to yield

$$\left(\frac{\partial^2}{\partial t^2} - c_0^2\nabla^2\right)\rho' = \frac{\partial^2}{\partial x_i \partial x_j}T_{ij}, \tag{2.4}$$

where $c_0^2 = \gamma p_0/\rho_0$, $\rho = \rho_0 + \rho'$, ρ_0 is constant, and

$$T_{ij} = \rho u_i u_j + (p - c_0^2\rho)\delta_{ij}. \tag{2.5}$$

The left hand side of (2.4) is then simply the acoustic wave operator acting on the density field, and the right hand side is to be regarded as the effective acoustic source.

So far we have proceeded by formal manipulation of (2.1 - 2.3). The predictive power of (2.4), however, comes only as a result of making certain assumptions about the nature of the source tensor T_{ij}.

Lighthill (*op. cit.*) proceeded by assuming that the eddy Mach number is small, so that each eddy is acoustically compact. Scaling arguments for small Mach numbers imply that $T_{ij} \approx \rho_0 u_i u_j$. In deriving scaling laws for the wave field due to a low Mach number eddy, it is

convenient to follow Crighton (1975) and work in the frequency domain $(\boldsymbol{x},\omega)$, rather than the time domain $(\boldsymbol{x},t)$.

The solution to (2.4) may then be written down as

$$\rho'(\boldsymbol{x},\omega) = -\frac{1}{4\pi c_0^2}\int \frac{\partial^2}{\partial x_i' \partial x_j'} T_{ij}(\boldsymbol{x}',\omega)\frac{\exp(i\omega|\boldsymbol{x}-\boldsymbol{x}'|/c_0)}{|\boldsymbol{x}-\boldsymbol{x}'|}d^3\boldsymbol{x}' \tag{2.6}$$

in three space dimensions, and

$$\rho'(\boldsymbol{x},\omega) = \frac{1}{16\pi i c_0^2}\int \frac{\partial^2}{\partial x_i' \partial x_j'} T_{ij}(\boldsymbol{x}',\omega)\mathrm{H}_0(\omega|\boldsymbol{x}-\boldsymbol{x}'|/c_0)d^2\boldsymbol{x}' \tag{2.7}$$

in two space dimensions, where H_0 is the Hankel function of the first kind of order zero. If the source T_{ij} is assumed acoustically compact, we may expand the integrals (2.6) and (2.7) to obtain far field expressions for the radiated waves as

$$\rho' \sim \frac{1}{4\pi c_0^2}\frac{\omega^2}{c_0^2}\frac{x_i x_j}{|\boldsymbol{x}|^2}\frac{1}{|\boldsymbol{x}|}\exp(i\omega|\boldsymbol{x}|/c_0)\int T_{ij}(\boldsymbol{x}',\omega)d^3\boldsymbol{x}' \tag{2.8}$$

in three space dimensions, and

$$\rho' \sim \frac{i}{16\pi c_0^2}\frac{\omega^2}{c_0^2}\frac{x_i x_j}{|\boldsymbol{x}|^2}\frac{1}{(\omega|\boldsymbol{x}|/c_0)^{1/2}}\int T_{ij}(\boldsymbol{x}',\omega)d^2\boldsymbol{x}' \tag{2.9}$$

in two. The asymptotic expressions (2.8) and (2.9) enable us to obtain the dependence of the radiated far wave field upon the parameters of the source flow. In three dimensions, we have

$$\rho' \sim \frac{1}{c_0^2}\frac{\omega^2}{c_0^2}\frac{1}{|\boldsymbol{x}|}\rho_0 u^2 l^3 = \rho_0 \frac{l}{|\boldsymbol{x}|}\left(\frac{u}{c_0}\right)^2\left(\frac{\omega l}{c_0}\right)^2 \tag{2.10}$$

and in two dimensions

$$\rho' \sim \frac{1}{c_0^2}\frac{\omega^{3/2}}{c_0^{3/2}}\frac{1}{|\boldsymbol{x}|^{1/2}}\rho_0 u^2 l^2 = \rho_0\left(\frac{l}{|\boldsymbol{x}|}\right)^{1/2}\left(\frac{u}{c_0}\right)^2\left(\frac{\omega l}{c_0}\right)^{3/2}. \tag{2.11}$$

Typically, one assumes that the frequency ω is set by the velocity and length scales of the eddies, so that $\omega \sim u/l$. These expressions then lead directly to the most startling results of the theory – that at low Mach number the acoustic intensity $I \equiv (c_0^3/\rho_0)\rho'^2$ of the radiated waves at fixed $l/|\boldsymbol{x}|$ obeys the scaling laws

$$I \sim \begin{cases} M^8 & 3 \text{ dimensions} \\ M^7 & 2 \text{ dimensions} \end{cases}, \tag{2.12}$$

where M is Mach number of the turbulence in the eddy, holding c_0 and ρ_0 constant. The high exponent in the power scaling law derives from three facts: power is a quadratic quantity in

wave amplitude; the power radiated will depend on the intensity of the turbulence generating it; and the effective source term for the acoustic radiation corresponds to a quadrupole source distribution in the compact limit, cancelling out the potentially stronger effects of monopoles or dipoles. The implication is that acoustic disturbances generated by vortical motions will be exceptionally weak, at least in the limit of low Mach number, and can be expected to have a negligible impact on the flows which generate them.

One shortcoming of the Lighthill theory is that there is no way to quantify the effect of acoustic wave radiation on the vortical flows which generate the acoustic waves. A significant advance in the general theory of aerodynamic sound generation was made when several workers (Obermeier, 1967; Lauvstad, 1968; Crow, 1970) independently developed analyses of the phenomenon using the method of matched asymptotic expansions, taking the Mach number M as the small asymptotic parameter. Two asymptotic regions are established: a source region of length scale l, in which the flow is nearly incompressible, and a wave region of length scale lM^{-1}, in which acoustic waves are linear at the leading two orders in Mach number at which they appear.

One generally assumes that the vorticity in the eddy is zero outside a region of finite extent on the l length scale. In the vortical flow region, the flow is incompressible at leading order in Mach number. Therefore, if the vorticity field is specified, the velocity and pressure fields at leading order in Mach number can be obtained directly from it. If the vorticity is zero except in some bounded domain, the velocity fields will decay as r^{-3} in three dimensions, or r^{-1} in two dimensions. Higher order Mach number corrections to the velocity and pressure in the vortical region are obtained by perturbation expansion. The expansion is singular, with acoustic waves on a length scale of order lM^{-1}, which is long compared with the scale of the vortical region. At high order in Mach number, details of the matching conditions are required to uniquely determine the velocity and height fields in the vortical regions.

The principal acoustic waves obtained by the matched asymptotic analysis correspond exactly to those predicted by the original Lighthill theory as described above, provided T_{ij} is formed from the incompressible velocity field associated with the vorticity distribution for the eddy. The analysis of Crow is distinguished among these studies in that he carries out the Mach number expansion in the source region to a sufficiently high order in M to establish that corrections due to the compressibility of the vortical region do not introduce monopoles or

dipoles, which could in principle give rise to waves of the same amplitude as the leading order quadrupole.

In addition to a general analysis of the radiated sound field, the method of matched asymptotic expansions has also been used to obtain expressions for the sound radiated from flows whose leading order incompressible solution may be expressed analytically. Two-dimensional flows analysed in this way include point vortex flows (Klyatskin 1966; Stuber 1970; Rahman 1971; Crighton 1972a; Gryanik 1983), vortex sheets (Crighton 1972b; Crighton & Leppington, 1974) and a limited class of flows with distributed vorticity (Zeitlin 1988, 1991).

Application of the theory to simple three-dimensional flows is less common since, unlike a large class of two-dimensional point vortex solutions, there are comparatively few three dimensional flows with confined vorticity whose solution can be expressed in closed analytical form. The only reported calculations of sound generation by three dimensional ideal flows in free space are for interacting vortex rings (Klyatskin 1966; Kambe & Minota 1981), and for the Hasimoto soliton (Kimura 1989). Kambe and Minota have subsequently obtained good agreement between experimental measurements of the acoustic sound field generated by low Mach number colliding vortex rings in the laboratory and the corresponding theoretical predictions (Kambe & Minota 1983; Minota & Kambe 1986; Kambe, 1986).

Despite considerable interest in the sound field generated by two and three dimensional vortical motions, comparatively little attention has been paid to the effect of sound radiation upon the flow which generates it. Exceptions are the papers of Klyatskin (1966) and Gryanik (1983) for simple point vortex flows, and Zeitlin (1988, 1991) for two dimensional flows with distributed vorticity.

In the point vortex paper of Gryanik (1983), the flows consist of rotating configurations of point vortices. The point vortices are of equal strength, and placed at the vertices of a regular polygon. The acoustic wave field is calculated using the leading order of matched asymptotic expansions, and the radiated power computed from it. The interaction energy of the point vortex configurations depends on a single parameter: the distance from the centre of the polygon to the vortices, which is assumed to adjust slowly to account for the energy lost due to the acoustic wave radiation. However, a difficulty with point vortex calculations for essentially compressible fluid dynamics problems is that the point vortex model collapses at order M in the expansion as an evacuated region forms at a finite $O(M)$ distance from the

vortex core (see Barsony-Nagy *et al.*, 1987; Moore & Pullin, 1987). The resulting complexity of the asymptotic expansions makes it practically impossible to continue the analysis to sufficiently high order in M to account for the effect of acoustic radiation within a matched asymptotic framework, and we feel it is therefore desirable to investigate the effect of radiation on some distributed vortex flows, such as those analysed by Zeitlin (1988, 1991).

However, one can raise concerns about the analysis of Zeitlin (*op. cit.*). In his study, he took the initial vorticity distribution to be a member of the class of steadily rotating two dimensional vortical flows with vorticity of finite extent investigated by Abrashkin & Yakubovich (1984). The flows belong to distinct one-parameter families, each family being characterised by its circulation and degree of rotational symmetry. Again, using the leading order of matched asymptotics, Zeitlin obtains the wave field, and hence rate of energy radiation. He then assumes that the flows evolve to take account of this energy loss by adjusting the available free parameter, conserving circulation but reducing the energy of the vortical flow.

The important point is that, although distributed vorticity makes it quite possible to construct a singular perturbation expansion in M which remains regular over the vortical scale, and requires rescaling only to account for acoustic wave radiation, there seems to be no reason to assume, as Zeitlin does, that the effect of the acoustic radiation is to cause the vortex to evolve slowly between different states of the one parameter family. Williams (1992) has undertaken a weakly nonlinear analysis of the case with 180° rotational symmetry, in which the steadily rotating states are ellipses with small aspect ratio. His analysis showed that the vortex could split at a time of $O(F^{-4})$, which seems inconsistent with the supposition of Zeitlin that the ellipses would simply elongate, conserving their area, but remaining elliptical.

It is therefore of interest to perform a full analysis of this flow using matched asymptotic expansions, expanding the vortical and acoustic scale flows in the small parameter M until the matching condition between them requires a term to be introduced to the vortical scale flow which accounts for the loss of energy to the acoustic waves at some (high) order in M. The behaviour would then be described entirely in terms of advection of potential vorticity in the vortical region. The flow evolution should be consistent with energy arguments in cases where the latter apply, but would also be able to show when energy arguments would not apply.

Norton (personal communication) has shown how to recast the f-plane shallow water equa-

tions

$$\frac{D\boldsymbol{u}}{Dt} + f\boldsymbol{k} \times \boldsymbol{u} + g\nabla h = 0 \tag{2.13}$$

$$\frac{Dh}{Dt} + h\nabla.\boldsymbol{u} = 0 \tag{2.14}$$

into a similar form to (2.4), viz:

$$\left(\frac{\partial^2}{\partial t^2} + f^2 - c_0^2\nabla^2\right)\frac{\partial h}{\partial t} = \frac{\partial^2}{\partial x_i \partial x_j}T_{ij}, \tag{2.15}$$

where now

$$T_{ij} = \frac{\partial}{\partial t}(hu_iu_j) + f\varepsilon_{ik}hu_ju_k + \frac{1}{2}g\frac{\partial}{\partial t}(h-h_0)^2\delta_{ij}. \tag{2.16}$$

Now, (2.15) will have propagating wave solutions if and only if their frequency exceeds the magnitude of the inertial frequency $|f|$. If T_{ij} is formed as the result of some fully nonlinear evolution, it would seem unlikely that all frequencies in all of its components will fall below the inertial frequency f, however small the Rossby number of the flow may be in some averaged sense. Therefore, it seems highly likely that gravity wave radiation is the inevitable consequence of fully nonlinear flow evolution.

Since balanced models do not admit freely propagating gravity waves, one might suppose that the present analysis shows that it is unlikely that one could construct a balanced model of the vortical flow at all orders in the Froude number expansion. This conclusion, however, should not be drawn too quickly, and we are certainly in no position to draw it as a result of the analysis presented so far.

To summarize, the essential feature of the matched asymptotic approach when used in aeroacoustics is that the Mach number is low, so that the source scale flow is nearly incompressible, and can be described by its vorticity alone. The sound field is obtained by an asymptotic matching process, which does not affect the source flow at leading order.

In the shallow water equations, if a low Froude number approximation is made, and T_{ij}, given by (2.16), is constructed using the incompressible velocity field, a similar matched asymptotic analysis may be used to predict the leading order gravity wave radiation due to a localized vortical flow knowing only the evolution of the potential vorticity field. With the gravity wave field now obtained, we might tentatively suggest that the vortical flow could be corrected at subsequent orders in Froude number to take account of the both the effects of divergence in

the vortical region and of gravity wave radiation. If so, then the entire perturbation expansion in Froude number for all fields could be constructed knowing only the potential vorticity. This would therefore constitute a balanced model which actually takes into account all the gravity wave radiation that is necessarily generated by the Lighthill mechanism, and would remain valid as the vortical flow is adjusted to take account of the gravity wave radiation. This would seem contrary to the intuitive notion that balanced models should not be able to take propagating gravity waves into account.

It is clear that the only way to address the possibility of a balanced model incorporating gravity wave radiation is to perform a matched asymptotic analysis of the entire problem. That analysis forms the basis for the rest of this chapter, which is organised as follows:

In §2.2 I present a matched asymptotic analysis of a compact vortical source in the f-plane shallow water equations at low Froude number. I assume throughout that the potential vorticity is confined to a finite compact region for all time. The applicability of the method of matched asymptotic expansions to gravity wave generation in the two-dimensional shallow water equations in a rotating frame is established. The analysis is carried sufficiently far to obtain the principal effect of the wave radiation on the source region.

In §2.3 the effect of the radiation on the source region is discussed within the context of balance and potential vorticity inversion.

In §2.4 the effect of gravity wave radiation on an ellipse of uniform potential vorticity is examined. The degree to which the solution agrees with, and differs from, the analysis of Zeitlin (1988, 1991) is clarified.

In §2.5, I offer some concluding remarks about the nature of the effect of gravity wave radiation on vortical sources, with particular reference to the "generalized adjustment" ideas of Vallis (1992) and Shepherd (personal communication), which were discussed briefly in chapter 1.

2.2 Gravity wave generation as a singular perturbation problem

To establish the nature of gravity wave generation by vortical motions at low Froude number, we are considering single layer shallow water dynamics on an f-plane. In the absence of motion,

we assume the layer has uniform depth H_0, and hence uniform potential vorticity $Q_0 = f/H_0$.

Suppose now that the potential vorticity Q differs from Q_0 only in some vortical region of finite extent, which is of dimension L. We must now turn to deriving a scaling of the shallow water equations valid in the limit of order–one Rossby number and small Froude number. From the definition of potential vorticity, it follows that the relative vorticity ζ scales as $\zeta \sim Q'H_0$, where $Q' \sim Q - Q_0$, and hence that the typical velocities U in the region of non-uniform potential vorticity scale according to $U \sim Q'H_0L$. The Rossby number Ro and Froude number F are defined by $F \sim U/(gH_0)^{1/2}$, $Ro = U/fL \sim Q'/Q$. We shall assume for the remainder of this chapter that $Q'/Q_0 = O(1)$, so that $Ro = O(1)$, and we shall also assume that $F \ll 1$.

Now we let the total layer depth be H. By considering the balance of terms in the momentum equation, it follows from the assumption of order-one Rossby number that the advective and Coriolis terms are of the same order. Consistency requires that all terms in the momentum equation are of the same order. This implies that $H - H_0 = O(F^2)$, and hence that the continuity equation reduces to the incompressibility constraint at leading order.

With velocity $\boldsymbol{u}$ scaled on U, length scaled on L, time t scaled on U/L, and writing $H - H_0 = F^2H_0h$, the shallow water equations take the nondimensional form:

$$\frac{D\boldsymbol{u}}{Dt} + f\boldsymbol{k} \times \boldsymbol{u} + \nabla h = 0 \tag{2.17}$$

$$F^2\left(\frac{Dh}{Dt} + h\nabla.\boldsymbol{u}\right) + \nabla.\boldsymbol{u} = 0. \tag{2.18}$$

The nondimensionalised potential vorticity takes the form

$$Q = \frac{f + \zeta}{1 + F^2h}. \tag{2.19}$$

The leading order dynamics are thus simply two-dimensional incompressible vortex dynamics. The presence of background rotation plays no part in the leading order evolution of the potential vorticity field, although the associated height field h will be modified due to background rotation.

With the nondimensional form of the equations now established for the vortical region which "drives" the flow, we may proceed to develop the singular perturbation expansion for the entire

flow which will enable us to determine the radiated wave field and the nature of its effect on the vortical flow.

2.2.1 Perturbation expansion for the vortical flow

We begin by developing the perturbation expansion for the vortical flow region. We write $\boldsymbol{u} = \nabla\phi + \boldsymbol{k} \times \nabla\psi$, and expand ϕ, ψ and h in perturbation expansions as

$$\psi = \psi_0 + F^2\psi_2 + \cdots; \quad \phi = F^2\phi_2 + \cdots; \quad h = h_0 + F^2 h_2 + \cdots. \tag{2.20}$$

The expansion for ϕ starts at $O(F^2)$ because from (2.18) we see that at leading order the flow is non-divergent.

We consider a flow in which initially the potential vorticity Q is specified as a function of position.

As the flow evolves, the potential vorticity is advected by all orders of the velocity field $\boldsymbol{u}$:

$$\frac{\partial Q}{\partial t} + (\boldsymbol{u}_0 + F^2\boldsymbol{u}_2 + \cdots).\nabla Q = 0. \tag{2.21}$$

As it stands, this equation cannot be correct, since there are advective terms at all orders in F, but the time derivative term $\partial Q/\partial t$ occurs only at the leading order in F.

For times of order unity, the difficulty may be easily resolved by expanding Q as an asymptotic series in F:

$$Q = Q_0 + F^2 Q_2 + \cdots. \tag{2.22}$$

However, over longer times, it is possible that this formulation might give rise to secular growth in Q_2 – for example, if there is a component of $\boldsymbol{u}_2$ which acts to slow down the rotation of vortices, or to move two vortices apart. To resolve these difficulties, we must introduce a long time scale $T = F^2 t$, and the potential vorticity evolution equation at $O(F^2)$ becomes

$$\frac{\partial Q_2}{\partial t} + \frac{\partial Q_0}{\partial T} + \boldsymbol{u}_0.\nabla Q_2 + \boldsymbol{u}_2.\nabla Q_0 = 0. \tag{2.23}$$

In general, there is no straightforward way to distinguish between the advective terms in (2.23) which give rise to slow secular behaviour in the order-unity potential vorticity field from those which give rise to $O(F^2)$ fluctuations of the potential vorticity on the fast ($O(1)$) timescale. Fortunately, this distinction is not important in analysing gravity wave generation and its effect

on the vortical flow, although it is important when specific flows are to be described, as in §2.4 below.

We shall now assume that $Q = f$ everywhere except in some finite domain, which we refer to as the vortex. Then, since $\nabla.\boldsymbol{u}_0 = 0$, we may write $\boldsymbol{u}_0 = \boldsymbol{k} \times \nabla\psi_0$, and obtain

$$\nabla^2\psi_0 = Q_0 - f. \tag{2.24}$$

$Q_0 - f$ represents the leading order relative vorticity $\boldsymbol{k}.(\nabla \times \boldsymbol{u})$, which for convenience we shall refer to as ζ_0, so that ζ_0 is zero outside some finite region on the vortical length scale. In general, the Laplacian can only be inverted up to non-singular harmonic terms. By Liouville's theorem, any nonsingular harmonic terms must be unbounded at infinity. Assuming, that ψ_0 does not contain any non-singular growing terms, we can obtain ψ_0 from (2.24) in terms of a Green function integral

$$\psi_0 = \frac{1}{2\pi}\int \zeta_0(\boldsymbol{x}')\ln|\boldsymbol{x}-\boldsymbol{x}'|d^2\boldsymbol{x}'. \tag{2.25}$$

For the present analysis, we are primarily concerned with obtaining expressions for the vortical flow fields $\psi_0, h_0, \psi_2, \phi_2, \ldots$ in sufficient generality that we may evaluate them in the limit as $|\boldsymbol{x}| \to \infty$, in preparation for matching onto the wave region. First, we expand $\ln|\boldsymbol{x}-\boldsymbol{x}'|$ for $|\boldsymbol{x}| > |\boldsymbol{x}'|$, using the Taylor expansion for the logarithm:

$$\ln|\boldsymbol{x}-\boldsymbol{x}'| = \ln|\boldsymbol{x}| + x_i'\frac{\partial}{\partial x_i}\ln|\boldsymbol{x}| + \frac{1}{2}x_i'x_j'\frac{\partial^2}{\partial x_i\partial x_j}\ln|\boldsymbol{x}| + \cdots, \tag{2.26}$$

and substitute this for $\ln|\boldsymbol{x}-\boldsymbol{x}'|$ in (2.25). Since $Q - f$ is of compact support, all the resulting moment integrals converge, and it is simple to show that the remainder is of smaller order than the last moment integral taken. Taking terms to $O(r^{-2})$, we have

$$\begin{aligned}
\psi_0 &= \frac{1}{2\pi}\ln|\boldsymbol{x}|\int\zeta_0(\boldsymbol{x}')d^2\boldsymbol{x}' && (2.27)\\
&+ \frac{1}{2\pi}\frac{x_i}{|\boldsymbol{x}|^2}\int x_i'\zeta_0(\boldsymbol{x}')d^2\boldsymbol{x}' && (2.28)\\
&+ \frac{1}{2\pi}\left(\frac{\delta_{ij}}{|\boldsymbol{x}|^2} - \frac{2x_ix_j}{|\boldsymbol{x}|^4}\right)\int\frac{1}{2}x_i'x_j'\zeta_0(\boldsymbol{x}')d^2\boldsymbol{x}' && (2.29)\\
&+ O(r^{-3}).
\end{aligned}$$

Now, we may note that the integral in (2.27) is an $O(1)$ approximation to the circulation $\int\zeta d^2\boldsymbol{x}'$, and that in (2.28) is an $O(1)$ approximation to the Kelvin impulse $\int\zeta\boldsymbol{x}'d^2\boldsymbol{x}'$. These

are both constant on the $O(1)$ timescale t, and so the fluctuating part of the leading order streamfunction ψ_0 is $O(r^{-2})$ as $r \to \infty$. Since $\int(x^2+y^2)\zeta$ is also independent of time on the $O(1)$ timescale, the time-dependent part of (2.29) may be represented as

$$\frac{A(t)}{r^2}\cos 2\theta + \frac{B(t)}{r^2}\sin 2\theta. \tag{2.30}$$

This represents a quadrupolar far field for ψ_0, and will clearly match to a radiating wave field of quadrupole form.

It is straightforward to see from dimensional grounds, and it will be verified in §2.2.2 below, that if the vortical source has a length scale of order unity, then the gravity waves generated will have a long wavelength of order F^{-1}. It follows that the leading order temporally fluctuating quadrupole in ψ_0, given by (2.30), will be of order F^2 in the wave region, where the radial variable is $R \equiv Fr$.

On the other hand, it is clear that $O(F^2)$ fields will arise in the vortical region as a direct result of the local effects of divergence. If any of these fields should have a temporally fluctuating monopolar far field, they would give rise to monopolar gravity wave radiation at $O(F^2)$, which would be formally of the same order as the quadrupolar radiation obtained from the incompressible source scale dynamics. We shall therefore proceed to obtain the far field expressions for ψ_2 and ϕ_2, to determine whether the leading order quadrupole is the only $O(F^2)$ wave field.

To obtain expressions for ψ_2 and ϕ_2, we must first obtain an expression for h_0. If we attempt to use the divergence of the momentum equation (2.17):

$$\nabla.(\boldsymbol{u}_0.\nabla\boldsymbol{u}_0) + f\nabla.(\boldsymbol{k}\times\boldsymbol{u}_0) + \nabla^2 h_0 = 0 \tag{2.31}$$

directly to obtain h_0, we obtain the integral expression

$$h_0 = f\psi_0 - \frac{1}{2\pi}\int \frac{\partial^2}{\partial x_i' \partial x_j'}\left(u_{0i}(\boldsymbol{x}')u_{0j}(\boldsymbol{x}')\right)\ln|\boldsymbol{x}-\boldsymbol{x}'|d^2\boldsymbol{x}'. \tag{2.32}$$

Although $\partial^2(u_{0i}u_{0j})/\partial x_i \partial x_j = O(r^{-4})$ as $r \to \infty$, and therefore the integral in (2.32) converges, it is not straightforward to extract the asymptotic form of h_0 for large $|\boldsymbol{x}|$ from (2.32). To see this, substitute the expansion for $\ln|\boldsymbol{x}-\boldsymbol{x}'|$ into (2.32). The $\ln r$ and r^{-1} terms vanish after integration by parts, and we are left with a divergent integral:

$$h_0 \sim f\psi_0 - \frac{1}{2\pi}\left(\frac{\partial^2}{\partial x_i \partial x_j}\ln|\boldsymbol{x}|\right)\int (u_{0i}(\boldsymbol{x}')u_{0j}(\boldsymbol{x}'))\, d^2\boldsymbol{x}'. \tag{2.33}$$

The integrand is now of order r^{-2} for large r, and therefore the moment integrals for the far field of h_0 fail to converge at the first order at which they are non-zero. We are therefore unable to predict the far-field behaviour of h_0 directly from (2.31), and some other method must be found to obtain the far-field expression for h_0.

The resolution to the difficulty is to introduce the Bernoulli variable $B = h + \frac{1}{2}\boldsymbol{u}^2$ in (2.17), which was used extensively by Howe (1975) for his study of the aerodynamic sound generation problem, to obtain Crocco's equation:

$$\frac{\partial \boldsymbol{u}}{\partial t} + (f+\zeta)(\boldsymbol{k}\times\boldsymbol{u}) + \nabla B = 0. \tag{2.34}$$

As with ψ, ϕ and h, we expand B in an asymptotic series in F:

$$B = B_0 + F^2 B_2 + \cdots. \tag{2.35}$$

From (2.34) one then obtains at order unity

$$\nabla^2 B_0 = -\nabla.(\zeta_0 \boldsymbol{k}\times\boldsymbol{u}) + f\nabla^2\psi_0. \tag{2.36}$$

Inverting the Laplacian, we obtain[1]

$$B_0 = f\psi_0 + \frac{1}{2\pi}\int \nabla.(\zeta_0\nabla\psi_0)\ln|\boldsymbol{x}-\boldsymbol{x}'|d^2\boldsymbol{x}'. \tag{2.37}$$

The important point is that ζ_0 is zero outside some finite domain. Therefore we may obtain the far field expansion for B_0 by expanding the logarithm in the form (2.26) and substituting it into the integral in (2.37). Moreover, since the leading order term in the far field expansion of (2.37) has the coefficient

$$\int \zeta_0 \nabla\psi_0 d^2\boldsymbol{x}' = 0, \tag{2.38}$$

[1] In (2.37) and, expect where otherwise stated, in all subsequent integrals in §2.2.1, all terms in integrands are functions of $\boldsymbol{x}'$, and all gradients, represented by ∇, when acting on integrands, are gradients with respect to $\boldsymbol{x}'$.

it follows from conservation of Kelvin's impulse on timescale t (see, e.g. Batchelor, 1967) that the far field expansion of B_0 to $O(r^{-2})$ takes the form

$$B_0 \sim f\psi_0 - \frac{1}{2\pi}\left(\frac{\delta_{ij}}{|\boldsymbol{x}|^2} - \frac{2x_i x_j}{|\boldsymbol{x}|^4}\right)\int \zeta_0(\boldsymbol{x}')x_i'\frac{\partial\psi_0}{\partial x_j'}d^2\boldsymbol{x}' + O(r^{-3}) \tag{2.39}$$

in the limit of large r. Consequently, the leading order Bernoulli function B_0 has a quadrupolar far field. To obtain the far field for h_0, we then use $h_0 = B_0 - \frac{1}{2}\boldsymbol{u}_0^2$. From the far field expression for ψ_0 (2.27-2.29),we can see that $u_0 = O(r^{-1})$ in the far field, and the only r^{-1}-like term in $\boldsymbol{u}_0$ is the circulation, which is independent of time. Therefore, at $O(r^{-2})$, B_0 and h_0 differ in the far field by a time-independent term only, and the principal time fluctuating part of h_0 is therefore also quadrupolar in the far field.

This resolves our first difficulty, which arose due to the nonconvergence of the integral in (2.33). We must now proceed to resolve whether there are competing monopoles introduced by compressibility at $O(F^2)$. If the time fluctuating parts of the far fields of ψ_2 and ϕ_2 at $O(1)$ have only $e^{2i\theta}$ dependence, then they will correspond to higher order matching for the quadrupole in the wave region. If, on the other hand, they contain components which are independent of θ, these would correspond to an $O(F^2)$ monopole. The principal wave field could not then be deduced directly from the incompressible velocity field, and we could not truly refer to the radiation as "quadrupole".

To resolve this, we carry out the perturbation expansion to the next order in F in the source region. Writing $\boldsymbol{u}_2 = \nabla\varphi_2 + \boldsymbol{k}\times\nabla\psi_2$, we have

$$\nabla^2\phi_2 = -\frac{\partial h_0}{\partial t} - \nabla.(\boldsymbol{u}_0 h_0) \tag{2.40}$$

$$\nabla^2\psi_2 = (f+\zeta_0)h_0 + Q_2. \tag{2.41}$$

From (2.37), and using the identity

$$\boldsymbol{u}_0^2 \equiv |\nabla\psi|^2 = \frac{1}{2}\nabla^2\psi_0^2 - \psi_0\nabla^2\psi_0, \tag{2.42}$$

we may write h_0 as

$$h_0 = \left(f + \frac{1}{2}\zeta_0\right)\psi_0 - \frac{1}{4}\nabla^2\psi_0^2 + \frac{1}{2\pi}\int \nabla.(\zeta_0\nabla\psi_0)\ln|\boldsymbol{x}-\boldsymbol{x}'|d^2\boldsymbol{x}'. \tag{2.43}$$

In addition, it is helpful to note the identity:

$$\nabla.(\psi_0 \boldsymbol{u}_0) = 0, \tag{2.44}$$

and hence to use the expression

$$\nabla.(h_0 \boldsymbol{u}_0) = \nabla.((h - f\psi_0)\boldsymbol{u}_0) \tag{2.45}$$

in (2.40). The point here is that $h_0 = O(\ln r)$ as $r \to \infty$, whereas, from (2.39), it is straightforward to show that $h_0 - f\psi_0 = O(r^{-2})$. This turns out to be essential in determining the far field expression for ϕ_2.

With h_0 now in a suitable form for integration, the solution to (2.40) may be written as

$$\begin{aligned}
\phi_2 = \quad & - \frac{f}{8\pi}\frac{\partial}{\partial t}\int \zeta_0 |\boldsymbol{x} - \boldsymbol{x}'|^2 (\ln|\boldsymbol{x} - \boldsymbol{x}'| - 1) d^2\boldsymbol{x}' && (2.46)\\
& - \frac{1}{4\pi}\frac{\partial}{\partial t}\int \psi_0 \zeta_0 \ln|\boldsymbol{x} - \boldsymbol{x}'| d^2\boldsymbol{x}' && (2.47)\\
& + \frac{1}{2}\psi_0 \frac{\partial \psi_0}{\partial t} && (2.48)\\
& - \frac{1}{8\pi}\frac{\partial}{\partial t}\int \nabla.(\zeta_0 \nabla\psi_0)|\boldsymbol{x} - \boldsymbol{x}'|^2 (\ln|\boldsymbol{x} - \boldsymbol{x}'| - 1) d^2\boldsymbol{x}' && (2.49)\\
& - \frac{1}{2\pi}\int \nabla.((h_0 - f\psi_0)\boldsymbol{u}_0) \ln|\boldsymbol{x} - \boldsymbol{x}'| d^2\boldsymbol{x}'. && (2.50)
\end{aligned}$$

The first four terms (2.46 - 2.49) all come from inversion of the Laplacian on $-\partial h_0/\partial t$, using (2.43) to first obtain an expression for $-\partial h_0/\partial t$. The last term (2.50) comes from inversion of the Laplacian on $-\nabla.(h_0\boldsymbol{u}_0)$. However, $h_0\boldsymbol{u}_0 = O((\ln r)/r)$, and it is not immediately clear that the Laplacian can be inverted via an integral expression such as (2.50). Using expression (2.45) makes it clear that the integrand in (2.50) is of order $O(r^{-4})$ as $r \to \infty$, and hence the integral converges.

The solution to (2.41) for ψ_2 can also be written down as an integral expression:

$$\begin{aligned}
\psi_2 = \quad & - \frac{1}{4} f \psi_0^2 && (2.51)\\
& + \frac{1}{2\pi}\int Q_2 + \zeta_0 \left(h_0 + \frac{1}{2} f\psi_0\right) \ln|\boldsymbol{x} - \boldsymbol{x}'| d^2\boldsymbol{x}' && (2.52)\\
& + \frac{1}{8\pi} f^2 \int \zeta_0 |\boldsymbol{x} - \boldsymbol{x}'|^2 (\ln|\boldsymbol{x} - \boldsymbol{x}'| - 1) d^2\boldsymbol{x}' && (2.53)\\
& + \frac{1}{8\pi} f \int \nabla.(\zeta_0 \nabla\psi_0)|\boldsymbol{x} - \boldsymbol{x}'|^2 (\ln|\boldsymbol{x} - \boldsymbol{x}'| - 1) d^2\boldsymbol{x}'. && (2.54)
\end{aligned}$$

We must now obtain the far field expansions of these expressions for ϕ_2 and ψ_2.

Consider first the far field form of ϕ_2 (2.46-2.50). We shall consider the far field form of each of expressions (2.46-2.50) in turn.

Starting with (2.46), we have

$$\begin{aligned}\frac{\partial}{\partial t}\int \zeta_0|\boldsymbol{x}-\boldsymbol{x}'|^2(\ln|\boldsymbol{x}-\boldsymbol{x}'|-1)d^2\boldsymbol{x}' &= |\boldsymbol{x}|^2(\ln|\boldsymbol{x}|-1)\frac{d}{dt}\int \zeta_0 d^2\boldsymbol{x}' && (2.55)\\ &+ \frac{\partial}{\partial x_i}(|\boldsymbol{x}|^2(\ln|\boldsymbol{x}|-1))\frac{d}{dt}\int x_i'\zeta_0 d^2\boldsymbol{x}' && (2.56)\\ &+ \frac{1}{2}\frac{\partial^2}{\partial x_i\partial x_j}(|\boldsymbol{x}|^2(\ln|\boldsymbol{x}|-1))\frac{d}{dt}\int x_i'x_j'\zeta_0 d^2\boldsymbol{x}' && (2.57)\\ &+ O(r^{-1}\ln r). && (2.58)\end{aligned}$$

Now, since

$$\frac{d}{dt}\int \zeta_0 d^2\boldsymbol{x}' = 0 \tag{2.59}$$

by conservation of circulation, and

$$\frac{d}{dt}\int x_i'\zeta_0 d^2\boldsymbol{x}' = 0 \tag{2.60}$$

by conservation of Kelvin's impulse, it follows that

$$\begin{aligned}\frac{\partial}{\partial t}\int \zeta_0|\boldsymbol{x}-\boldsymbol{x}'|^2(\ln|\boldsymbol{x}-\boldsymbol{x}'|-1)d^2\boldsymbol{x}' &= \frac{1}{2}\frac{\partial^2}{\partial x_i\partial x_j}(|\boldsymbol{x}|^2(\ln|\boldsymbol{x}|-1))\frac{\partial}{\partial t}\int x_i'x_j'\zeta_0 d^2\boldsymbol{x}' && (2.61)\\ &+ O(r^{-1}\ln r). && (2.62)\end{aligned}$$

We can now show that (2.46) does not contribute a monopolar source term. We start by noting that

$$\frac{d}{dt}\int (x'^2+y'^2)\zeta_0 d^2\boldsymbol{x}' = 0. \tag{2.63}$$

It then follows that the tensor integral in (2.61) is traceless The $O(1)$ far field component of (2.61) therefore corresponds to the higher order in Froude number matching conditions for the quadrupolar wave field, but contributes no further monopolar wave field of its own.

Continuing to analyse (2.47), it follows from conservation of energy in the source region on the timescale of the vortex dynamics that

$$\frac{d}{dt}\int \zeta_0\psi_0 d^2\boldsymbol{x}' = 0. \tag{2.64}$$

Expanding the logarithm and substituting, it follows that the far field of (2.47) is $O(r^{-1})$.

Since $\psi_0 \sim \ln r$, and $\partial\psi_0/\partial t \sim r^{-2}$, it follows that (2.48) is $O(r^{-2}\ln r)$ in the far field, so we need not consider it further.

For (2.49), it follows again from conservation of Kelvin's impulse that

$$\begin{aligned}\frac{\partial}{\partial t}\int \nabla.(\zeta_0\nabla\psi_0)|\boldsymbol{x}-\boldsymbol{x}'|^2(\ln|\boldsymbol{x}-\boldsymbol{x}'|-1)d^2\boldsymbol{x}' = \\ \frac{1}{2}\frac{\partial^2}{\partial x_i\partial x_j}(|\boldsymbol{x}|^2(\ln|\boldsymbol{x}|-1))\frac{\partial}{\partial t}\int x_i'\zeta_0\frac{\partial\psi_0}{\partial x_j'}d^2\boldsymbol{x}' \\ + \quad O(r^{-1}\ln r)\end{aligned} \tag{2.65}$$

Like (2.46), this also has a fluctuating far field at $O(1)$, but also like (2.46), we can show that the trace of the tensor integral is zero, and therefore this too corresponds only to higher order matching for the quadrupolar waves, and does not introduce a monopole either.

To see this, note that the trace of (2.65) is proportional to $\int(xv_0 - yu_0)$. One can see that this is zero from conservation of $\int(x^2+y^2)\zeta_0$, and therefore we can conclude that (2.49) has only $e^{2i\theta}$ behaviour at $O(1)$ in the far field. Consequently, (2.49) does not contribute a monopole in the far field.

Finally we consider (2.50). If we proceed by expanding the logarithm and integrating by parts, as before, the first term we obtain is

$$\int \nabla.((h_0 - f\psi_0)\boldsymbol{u}_0)\ln|\boldsymbol{x}-\boldsymbol{x}'|d^2\boldsymbol{x}' = \frac{\boldsymbol{x}}{|\boldsymbol{x}|^2}.\int(h_0 - f\psi_0)\boldsymbol{u}_0 d^2\boldsymbol{x}'. \tag{2.66}$$

However, we should note that, unlike the other integrals for which far field expressions have been obtained, the integrand here is not of compact support, and is of order $O(r^{-3})$ as $r \to \infty$. At the next order in the expansion of the logarithm, a factor x' is introduced into the integrand, making the integral non-convergent. It follows that the order of the remainder in this case cannot be determined from subsequent terms in the asymptotic expansion of the logarithm, and extensive analysis is required to show that the order of the remainder is $(r^{-2}\ln r)$, thus confirming that (2.50) is of order $O(r^{-1})$ for large r, and therefore does not contribute a monopole to the wave field. The analysis to show that the order of the remainder in (2.66) is $r^{-2}\ln r$ is presented in Appendix A.

We turn now to examining the far field of ψ_2.

Two of the four integrals in the representation of ψ_2 have counterparts in the representation of ϕ_2. They are (2.53), which is the counterpart of (2.46), and (2.54), which is the counterpart of (2.49). The only difference between (2.46) and (2.53) is a factor of $f^{-1}\partial/\partial t$. This is also the difference between (2.49) and (2.54). It follows from the analyses presented for (2.46) and (2.49) that neither (2.53) nor (2.54) can have time dependent monopoles in their far fields, although both terms do contain components of the time dependent quadrupole.

To show that (2.51) does not have a time dependent monopolar far field, we recall that although $\psi_0 = O(\ln r)$ for large r, $\psi_{0t} = O(r^{-2})$. Therefore, the time dependent part far field of (2.51) is of order $O(r^{-2}\ln r)$, and hence does not constitute a time dependent monopole.

To show that (2.52) does not have a time dependent monopolar far field, we first recall that $\int \zeta_0\psi_0$ is independent of time, and therefore we need only consider the term

$$\frac{1}{2\pi}\ln r \int Q_2 + \zeta_0 h_0. \tag{2.67}$$

However, it is simply the $O(F^2)$ contribution to the circulation. Since conservation of circulation is true of a general shallow water system, and does not rely on any small Froude number limit, it should be independent of time at all orders. To show that (2.67) is independent of time, we note that

$$\frac{\partial \zeta_0}{\partial t} = -\nabla.(\boldsymbol{u}_0\zeta_0); \quad \frac{\partial h_0}{\partial t} = -\nabla.(\boldsymbol{u}_0 h_0 + \boldsymbol{u}_2). \tag{2.68}$$

Hence

$$\frac{\partial}{\partial t}(\zeta_0 h_0) = -\nabla.(\zeta_0 h_0 \boldsymbol{u}_0) - \zeta_0 \nabla.\boldsymbol{u}_2. \tag{2.69}$$

Moreover, if we also recall that

$$\frac{\partial Q_2}{\partial t} + \frac{\partial \zeta_0}{\partial T} = -\boldsymbol{u}_0.\nabla Q_2 - \boldsymbol{u}_2.\nabla Q_0, \tag{2.70}$$

we obtain

$$\frac{d}{dt}(Q_2 + \zeta_0 h_0) = -\nabla.(\boldsymbol{u}_0 Q_2 + h_0\zeta_0\boldsymbol{u}_0 + \zeta_0\boldsymbol{u}_2) - \frac{\partial \zeta_0}{\partial T}. \tag{2.71}$$

It follows that

$$\frac{d}{dt}\int Q_2 + \zeta_0 h_0 = -\frac{d}{dT}\int \zeta_0 - \int (\boldsymbol{u}_0 Q_2 + \zeta_0\boldsymbol{u}_2 + \zeta_0 h_0\boldsymbol{u}_0).d\boldsymbol{s}. \tag{2.72}$$

Now since ζ_0 vanishes ouside some finite region, the boundary integral on the right hand side of (2.72) vanishes. Then, since $\int \zeta_0$ is independent of t, it follows that (2.52) does not have a time dependent far field at $O(1)$ in r for large r.

To obtain an integral representation for h_2, we expand (2.34) at order F^2 to obtain

$$\nabla^2\left(B_2 + \frac{\partial \phi_2}{\partial t} - f\psi_2 - \frac{1}{2}f^2\psi_0^2\right) = \nabla.(h_0\zeta_0\nabla\psi_0 + (h_0 - f\psi_0)\nabla\psi_0 - \zeta_0(\boldsymbol{k}\times\boldsymbol{u}_2)). \qquad (2.73)$$

Now ϕ_2 and ψ_2 have already been shown to contribute only quadrupoles. It is straightforward to show that the right hand side contributes a dipole only (in an analysis similar to that given for (2.50) in appendix A). Therefore B_2, and hence h_2, are also quadrupole at leading order in the far field. This is reassuring, since gravity waves involve perturbations in both velocity and height fields, and it should not have been possible to have a monopole in h_2 without a corresponding monopole in at least one of ψ_2 or ϕ_2.

2.2.2 Dynamics of the wave zone resulting from the vortical flow

Considering the orders of the unsteady terms in the limit $r \to \infty$ in the vortical flow, we have terms of orders $1 \times r^{-2}$ and $F^2 \times 1$, reminding us that the expansion becomes disordered when $r \sim F^{-1}$. To obtain the dynamics of the wave zone, we must rescale the equations using the long length scale F^{-1} corresponding to the wavelength of gravity waves. We therefore introduce the wave region spatial variable $\boldsymbol{X}$, defined such that $\boldsymbol{X} = F\boldsymbol{x}$. Rescaling the equations, we obtain

$$\frac{\partial}{\partial t}(\nabla\phi + \boldsymbol{k}\times\nabla\psi) + (f+\zeta)\boldsymbol{k}\times(\nabla\phi + \boldsymbol{k}\times\nabla\psi) + \nabla h + \frac{1}{2}F^2\nabla(\nabla\phi + \boldsymbol{k}\times\nabla\psi)^2 = 0 \qquad (2.74)$$

$$\frac{\partial h}{\partial t} + \nabla^2\phi + F^2\nabla.((\nabla\phi + \boldsymbol{k}\times\nabla\psi)h) = 0, \qquad (2.75)$$

where now ∇ represents $\partial/\partial\boldsymbol{X}$. These equations admit propagating gravity waves as solutions. In the wave region it is the nonlinear terms, rather than the divergence terms, which are of small order in the limit of small Froude number.

We now proceed to the details of matching the asymptotic expansions in the source region and the wave zone together. We start by expanding ψ, ϕ and h in asymptotic series in F:

$$\psi = \psi_0 + F\psi_1 + F^2\psi_2 + \cdots \tag{2.76}$$

$$\phi = \phi_0 + F\phi_1 + F^2\phi_2 + \cdots \tag{2.77}$$

$$h = h_0 + Fh_1 + F^2h_2 + \cdots. \tag{2.78}$$

Firstly, since, for the vortical region, ψ_0 and h_0 are independent of t at orders 1 and r^{-1} in their far field limits, it is sufficient in the expansion in the wave zone to take ψ_0, ψ_1, h_0 and h_1 to be independent of t, and to take $\phi_0 = \phi_1 = 0$. Thus, in the wave zone,

$$\{h_0, h_1\} = f\{\psi_0, \psi_1\} \tag{2.79}$$

$$(\nabla^2 - f^2)\{\psi_0, \psi_1\} = 0, \tag{2.80}$$

and it follows from applying decaying boundary conditions at infinity that

$$\psi_0 = C_0 \mathrm{K}_0(|f|R) \tag{2.81}$$

$$\psi_1 = C_1 \mathrm{K}_1(|f|R). \tag{2.82}$$

The values of the constants C_0 and C_1 are determined by matching conditions onto the source scale flow in the limit $R \to 0$.

The expansion of $\mathrm{K}_0(|f|R)$ for small R implies that constant corrections to h will be required at $O(1)$ and $O(\ln F)$, which will in turn affect the velocity field at $O(F^2)$ and $O(F^2 \ln F)$. Since these are time-independent terms, will shall not discuss them further here. We shall return to them in §2.4, where the complete solution to $O(F^2)$ is obtained for the model problem of the Kirchoff ellipse (Lamb, 1932).

At the next two orders ($O(F^2)$ and $O(F^3)$) in the wave zone, it is convenient to separate the fields into two parts: one part which is independent of t, and one part which is oscillatory in t. Taking, for example, ψ_2, we use $\hat{\psi}_2$ to represent the part which is independent of t, and $\tilde{\psi}_2$ to represent the oscillatory part. For the t-independent parts, we obtain from (2.74) at $O(F^2)$:

$$(\nabla^2 - f^2)\hat{\psi}_2 = -\frac{1}{2} f |\nabla \psi_0|^2 + \frac{1}{2} f^3 \psi_0^2. \tag{2.83}$$

A similar equation can be obtained for $\hat{\psi}_3$. Again, the boundary conditions on these equations are decaying at infinity, and matching conditions onto the source flow as $|\boldsymbol{X}| \to 0$.

For the oscillatory parts, we then have

$$\left(\frac{\partial^2}{\partial t^2} + f^2 - \nabla^2\right)\left\{\tilde{\phi}_2, \tilde{\psi}_2, \tilde{h}_2, \tilde{\phi}_3, \tilde{\psi}_3, \tilde{h}_3\right\} = 0, \tag{2.84}$$

with

$$\frac{\partial}{\partial t}\{\tilde{\psi}_2, \tilde{\psi}_3\} + f\{\tilde{\phi}_2, \tilde{\phi}_3\} = 0. \tag{2.85}$$

Here, the boundary conditions which we shall impose are a radiation condition at infinity, and matching onto the source flow as $|\boldsymbol{X}| \to 0$. When matching $\tilde{\phi}_2$ and $\tilde{\phi}_2$ onto the vortical region, we do not match ϕ and ψ individually, because the decomposition of the velocity field into streamfunction ψ and velocity potential ϕ is non-unique. In the vortical region, the leading order far field quadrupole comes about directly as a result of incompressible flow, and in the previous section we chose to represent that flow using only a streamfunction. Therefore, its far field does not satisfy the condition (2.85). However, since the vorticity is confined to a finite area of the vortical region, the flow is both incompressible and irrotational in the limit $r \to \infty$, and hence we are free to represent it far from the vortex using either a streamfunction or a velocity potential, or any combination of the two, including a combination which satisfies the constraint (2.85).

It is convenient to work in frequency space for the details of wave matching, in which the general solutions of (2.84) are the Hankel functions

$$\mathrm{H}_m^{(\cdot)}\left(\sqrt{\omega^2 - f^2}R\right)e^{im\theta}, \tag{2.86}$$

where m is the order, determined by the required θ dependence of the match on to the source flow; $(\cdot) = 1, 2$ is the type, determined by the radiation condition; and $R = |\boldsymbol{X}|$ as before. The order $m = 0$ corresponds to a monopole wave, $m = 1$ a dipole, $m = 2$ a quadrupole, and so on. Normally the Hankel functions $H_m^{(1,2)}$ are defined to be $J_m \pm iY_m$, where J and Y are Bessel and associated Bessel functions respectively (Abramowitz & Stegun, 1965). For convenience in the present analysis, we shall multiply H by $\pm i$, so that in our case $H_m^{(1,2)} = -Y_m \pm iJ_m$. Throughout, we shall represent the Fourier transform of a function $f(t)$ by $\tilde{f}(\omega)$, where

$$\tilde{f}(\omega) = \int_{-\infty}^{\infty} f(t)e^{i\omega t}dt. \tag{2.87}$$

Now, the time-dependent far field form of ψ_0 in the vortical region, given in (2.30), may be taken as

$$\psi_0 \sim \frac{4\alpha(t)}{\pi r^2}e^{2i\theta}. \tag{2.88}$$

Therefore, the outer limit of the inner expansion for ψ takes the form $r^{-2}e^{2i\theta}$. To satisfy the matching conditions for the wave fields at $O(F^2)$, we will require them to have 2θ angular dependence. Hence we choose mode $m = 2$ and write the general solution for the $O(F^2)$ wave fields in frequency space:

$$\tilde{\phi_2} = A(\omega)(\omega^2 - f^2)\mathrm{H}_2^{(\cdot)}\left(\sqrt{\omega^2 - f^2}R\right)e^{2i\theta} \tag{2.89}$$

$$\tilde{\psi_2} = B(\omega)(\omega^2 - f^2)\mathrm{H}_2^{(\cdot)}\left(\sqrt{\omega^2 - f^2}R\right)e^{2i\theta} \tag{2.90}$$

$$\tilde{h_2} = C(\omega)(\omega^2 - f^2)\mathrm{H}_2^{(\cdot)}\left(\sqrt{\omega^2 - f^2}R\right)e^{2i\theta}. \tag{2.91}$$

One matches the velocity $\nabla\phi + \boldsymbol{k} \times \nabla\psi$ and the height field h to get

$$A = \frac{f\tilde{\alpha}}{\omega + f}; \quad B = \frac{i\omega\tilde{\alpha}}{\omega + f}; \quad C = (\omega - f)\tilde{\alpha}, \tag{2.92}$$

where $\tilde{\alpha}(\omega)$ is the Fourier transform of $\alpha(t)$, defined by (2.87).

Since we have shown that there are no $O(F^2)$ monopoles in the source region, this is the complete representation of the leading order (i.e. $O(F^2)$) wave field. Thus we have confirmed that the principal radiation is quadrupole. In general, the $O(F^3)$ wave field will be composed of a dipole and an octupole.

2.2.3 The effect of wave radiation on the vortical flow

We have now determined the principal wave radiation in terms of integrals over the vortex. These integrals can be regarded as known functions of time. At any time, they can be determined by knowing only the potential vorticity distribution at that time.

We turn now to considering the "back-reaction" problem. That is, we calculate the effect of the radiation on the flow which is generating it. This consists in matching the wave solution back onto the source flow. To do this, we must consider the expansion of the Hankel function in the limit $R \to 0$:

$$\pi(\omega^2 - f^2)\mathrm{H}_2^{(\cdot)}(\sqrt{\omega^2 - f^2}R) =$$

$$\frac{4}{R^2} \tag{2.93}$$

$$+ \quad (\omega^2 - f^2) \tag{2.94}$$

$$- \quad \frac{1}{4}(\omega^2 - f^2)^2 R^2 \ln\left(\frac{1}{2}\sqrt{\omega^2 - f^2}R\right) \tag{2.95}$$

$$+ \quad \frac{1}{8}(\omega^2 - f^2)^2(\psi(1) + \psi(3))R^2 \tag{2.96}$$

$$\pm \quad i\pi\frac{1}{8}(\omega^2 - f^2)^2 R^2 \tag{2.97}$$

$$+ \quad O(R^4 \ln R),$$

where $\psi(\cdot)$ is the logarithmic derivative of the Γ function (Abramowitz & Stegun, 1965. p.258).

Expanding this in the source scale variables we get

$$F^2\pi(\omega^2 - f^2)\mathrm{H}_2^{(\cdot)}(\sqrt{\omega^2 - f^2}R) \quad =$$

$$\frac{4}{r^2} \tag{2.98}$$

$$+ \quad F^2(\omega^2 - f^2) \tag{2.99}$$

$$- \quad F^4 \ln F \frac{1}{4}(\omega^2 - f^2)^2 r^2 \tag{2.100}$$

$$+ \quad F^4\frac{1}{4}(\omega^2 - f^2)^2 r^2\left(\ln 2 - \ln(\sqrt{(\omega^2 - f^2)}r)\right) \tag{2.101}$$

$$+ \quad F^4\frac{1}{8}(\omega^2 - f^2)^2\left(\frac{3}{2} - 2\gamma\right)r^2 \tag{2.102}$$

$$\pm \quad i\pi\frac{F^4}{8}(\omega^2 - f^2)^2 r^2 \tag{2.103}$$

$$+ \quad O(F^4 \ln F),$$

where $\gamma \approx 0.5772$ is Euler's constant.

At this point, any attempt to separate the matching procedure into terms which are "due to the radiation" and "other terms" might seem somewhat arbitrary. For a start, there is the term (2.100) of order $F^4 \ln F$, which could certainly not have been predicted from considering the expansion for the source flow alone. However, it has none of the features that we would expect of a term due to the radiation. In particular, it is independent of the sign of $\pm$ in (2.103) taken to satisfy the radiation condition! The only term which actually depends on application of the radiation condition is (2.103). However, one can see that, when we take the convolution with $\tilde{\alpha}(\omega)$ to invert the Fourier transform, we recover an integral which depends on both past and future values of $\alpha(t)$, which appears to violate the radiation condition. Thus it seems misleading to regard term (2.103) alone as representing the effect of radiation on the source. Instead, we

amalgamate the $\ln(\omega^2 - f^2)$ term in (2.101) with (2.103) and the γ term from (2.102) to give

$$- F^4(\omega^2 - f^2)^2 r^2 \left(\frac{1}{8} \left(\ln(\omega^2 - f^2) r + 2\gamma \right) \mp \frac{i\pi}{8} \right) \tilde{\sigma}(\omega), \tag{2.104}$$

where $\tilde{\sigma}$ is the Fourier transform of a source function. For the present analysis, $\tilde{\sigma}$ is $A(\omega)$, $B(\omega)$ or $C(\omega)$, where A, B and C are given by (2.92), depending on whether we are considering the matching conditions for ϕ, ψ or h, respectively. When (2.104) is converted back into the time domain, we recover an integral in the matching condition which depends only on the past history of the dynamics:

$$r^2 \int_0^\infty \ln \tau \left[\cos(f\tau) g(t - \tau) \right]_\tau d\tau, \tag{2.105}$$

where $g(t)$ is the inverse Fourier transform of $\tilde{g}(\omega) \equiv (\omega^2 - f^2)^2 \tilde{\sigma}(\omega)/16$ (the details are given in appendix B). A term (2.105) $\times e^{2i\theta}$ is thus required in ϕ, ψ and h, in the vortical region, at order $O(F^4)$. The exact form of g depends, of course, on whether we are performing the match for ϕ, ψ or h. The important point is that, since $\tilde{\sigma}$ is related to $\tilde{\alpha}$ through (2.92), with $\sigma = A, B$, or C, $g(t)$ may be expressed as a sum of instantaneous time derivatives of $\alpha(t)$ defined by (2.88), and is therefore a known function of the dynamics of the vortex region. The choice of $\mp$ in (2.104), required to obtain an integral (2.105) over the past history only, and not involving the future evolution, can be shown to be equivalent to requiring an outgoing wave form for $H_2^{(\cdot)}$ as $R \to \infty$.

We further remark that $O(F^3)$ dipoles in the wave zone also give rise to non-local time integrals at $O(F^4)$ in the source region. In this case the flow induced in the source region is a uniform velocity which has no effect on the energy of the source at leading order, although it must be retained if the $O(F^6)$ and $O(F^6 \ln F)$ flows in the source are to be computed.

2.3 Implications for balance and potential vorticity inversion

The foregoing study was motivated by a desire to understand the fundamental limitations of the concept of balanced dynamics. We shall now discuss this concept in the context of the isolated low Froude number vortex which we have studied here.

At $O(1)$, the dynamics are just two-dimensional incompressible vortex dynamics. This clearly does not depend on any knowledge of the past history of the flow in the vortex region, or of the gravity wave field, and so the dynamics are balanced in the sense that the instantaneous

potential vorticity field is sufficient to determine the velocity and height fields, and hence to integrate the equations of motion by advecting the potential vorticity field.

At $O(F^2 \ln F)$, there is a non-uniform correction to the vorticity of the vortical region as a result of the matching conditions. Although it arises as a result of the matching conditions between the vortical region and the wave region, it is independent of time. It can therefore be obtained from instantaneous fields, and should be regarded as a balanced correction to the vortical dynamics.

At $O(F^2)$, the corrections to the dynamics are expressible entirely in terms of inversions of the Laplacian on the $O(1)$ fields and further time-independent terms introduced by the asymptotic matching conditions. At this order, still no knowledge of the gravity wave field or of the past history of the vortex dynamics is required. This means that we can think of corrections at this order as balanced corrections to the dynamics, which represent the effects of local divergence, but not of gravity wave generation.

However, at $O(F^4)$, terms arise which come about as a result of matching to the wave zone, which are non-local in time, and which depend on the gravity wave radiation condition. We should note, however, that the integral (2.105) is not strictly an integral over the past history of an actual flow. The source term $g(t-\tau)$ in the integral is to be regarded as a term $g(t-\tau, T_2, T_4, \ldots)$ in which only the shortest time variable t is allowed to vary, and all other time variables are kept fixed. Formally we regard the integral as bounded because, implicit in our scaling analysis, we have assumed that terms associated with two-dimensional vortex dynamics remain bounded as the integration proceeds for arbitrarily long times. In reality this may be true only for a very small number of vorticity distributions, such as the solutions presented by Abrashkin and Yakubovich (1984).

Interestingly, then, this integral can be computed, in principle, by knowing only the instantaneous potential vorticity, and then integrating the two-dimensional incompressible vortex dynamics equations backwards in time. One might be tempted to think that we could therefore regard this correction as known in terms of the potential vorticity evolution, and therefore part of the balanced dynamics. However, it is then not clear how far back into the past they should be integrated. If we assume that the disturbance is turned on at some time t_0 then, for consistency, we would have to integrate back to time t_0. We would then expect (2.105) to be sensitive to the value of t_0 chosen. In the standard situation we are given *no* information about

the past history of the flow: we are given a potential vorticity distribution and nothing else. The choice of t_0 then becomes entirely arbitrary, and we cannot satisfactorily say that we have evaluated (2.105).

If we were to try to integrate (2.105) back to $\tau = -\infty$, then there is no reason to expect the integral to converge, even if it remains bounded. One might be slightly worried by this, since it is the result of inverting an apparently well-behaved Fourier transform. Formally one should say that the integral (2.105) is defined by its Cauchy Principal Value, so that $\int_0^\infty \ln\tau \cos\Omega\tau d\tau$ is given by the corresponding Fourier Transform (2.104), with $\omega \equiv \Omega$ and $\tilde{\sigma} \equiv 1$.

Another interesting feature is the logarithmic kernel, which means that contributions to (2.105) are significant for times arbitrarily far into the past. This may seem surprising because one would expect activity in the source region not to continue to affect the source region at the same strength long after the waves have propagated away. If g is truly of compact support, then we can integrate (2.105) once by parts to obtain

$$\int_{t-a}^{t-b} \tau^{-1} \cos(f\tau) g(t-\tau) d\tau, \tag{2.106}$$

where a and b are the limits of the range over which g is non-zero. It follows that if g is bounded then the importance of its contribution to the back reaction decays as t^{-1} for large t. However, we cannot dispense with past history in this manner unless we know that g is zero for all times before some initial time t_0. This is because although the individual impact of finite time sources decays as t^{-1}, there are in principle an infinite number of such sources.

Finally, we should note that the "radiation term" (2.104) is non-vanishing even when all the frequencies in $\tilde{\sigma}(\omega)$ lie below the inertial frequency f in magnitude. This corresponds to the case where all the gravity waves are evanescent. Therefore, they do not transport energy to infinity, and we would not wish to refer to this term as the effect of radiation in this context. However, we can see that in the case where $|\omega| < |f|$, we can obtain an expression for (2.104) which requires only instantaneous time derivates of $a(t)$ at the current time to be known. If $|\omega| < f$ we may write

$$-F^4\pi(\omega^2 - f^2)^2 r^2 \left(\frac{1}{8\pi}\left(\ln(f^2-\omega^2)r + i\arg(\omega^2 - f^2) + 2\gamma\right) \mp i\frac{1}{8}\right)\tilde{\sigma}(\omega). \tag{2.107}$$

Then, selecting the sign $\mp$ appropriately, according to whether we take $\arg(-1)$ as $\pm\pi$, we

are left with the so-called radiation term expressed as

$$\frac{1}{8}F^4(f^2-\omega^2)^2r^2\left(\ln(f^2-\omega^2)r+2\gamma\right)\tilde{\sigma}(\omega). \tag{2.108}$$

The selection of the $\mp$ here is equivalent to an evanescence boundary condition in the wave far field.

The logarithmic term $\ln(f^2-\omega^2)$ can now be expanded as

$$\ln(f^2-\omega^2)=\ln f^2+\ln(1-\omega^2/f^2)=2\ln f-\omega^2/f^2+\omega^4/2f^4-\omega^6/3f^6+\cdots, \tag{2.109}$$

i.e. $2\ln f$ plus a convergent Taylor series in which only integer powers of ω appear. The recovery of only integral powers of ω implies that we need only find the inverse Fourier transform of $\tilde{\sigma}(\omega)$ multiplied by integral powers of ω, and hence only instantaneous time derivatives of $\sigma(t)$ are required: no integral over the past history of evolution, such as (2.105), need be evaluated.

Technically, any finite truncation of this expansion defines a balanced model for the single compact vortex. One could perhaps argue that, in a source whose frequencies lie largely, but not entirely, below the inertial frequency, taking a few terms in the expansion of $\ln(f^2-\omega^2)$ might lead to a general improvement to the $O(F^4)$ dynamics, even though the series would ultimately diverge. This sort of behaviour is qualitatively familiar from the studies of Norton (1988) and McIntyre & Norton (1993), where a hierarchy of balanced models produces successively better approximations to shallow water dynamics at the first few orders, but ultimately fails to converge.

The notions discussed above have some bearing on the concept of "superbalance", as proposed by J. J. Tribbia (personal communication). In a "superbalance", one is supposed to diagnose all the past history of gravity wave radiation consistent with the given potential vorticity distribution. "Superbalance" might therefore be regarded as a gravity wave minimization, and as such deserves some consideration.

Tribbia has proposed the following scheme:

- Choose height and velocity fields consistent with the current PV distribution
- Integrate backwards in time (in principle to $t=-\infty$)
- Compute a time-integrated measure of total gravity wave activity

- Iterate the height and velocity velocity fields at $t = 0$ (subject to the prescribed PV field) to minimise the integrated gravity wave activity

In a sense, we might regard the flow given by the low Froude number matched asymptotic analysis as precisely the flow which minimises the total gravity wave activity, since we include only that gravity wave field which is essential for an asymptotic match onto the vortex dynamics. If so, then the foregoing analysis gives some idea of the characteristics which supposed superinversion operators are likely to possess. In particular, it suggests (a) that they are unlikely to converge unless some "initial instant" is assumed for the model, before which there is assumed to be no gravity wave activity, and (b) that the superinversion operator will in general be sensitive (at $O(F^4)$) to the choice of initial instant.

2.4 The effect of gravity wave radiation upon a rotating elliptical patch of uniform potential vorticity – an explicit example of generalized adjustment

In a pioneering paper, Broadbent & Moore (1979) investigated the stability of a Rankine vortex to two-dimensional perturbations. A Rankine vortex is an axisymmetric vortical column, with uniform vorticity within the column and zero vorticity without. In the low Mach number limit, therefore, the Rankine vortex corresponds to a vortex of uniform potential vorticity. They showed numerically that disturbances with axial mode numbers of two, three, four and six were unstable over a wide range of Mach numbers, and obtained an expression for the growth rate of the mode two instability at low Mach number by a matched asymptotic analysis.

Subsequently, Kop'ev & Leont'ev (1983) argued that the expression for the growth rate at low Mach number which was obtained by Broadbent & Moore (1979) could also be derived from energetic arguments, which made it possible to obtain growth rates for higher axial modes without significantly greater effort. Their analysis depends on assuming that the boundary perturbations are of small amplitude, of the form $e^{im\theta}$.

Recently, Zeitlin (1988, 1991) has suggested that the energetic arguments of Kop'ev & Leont'ev (1983) could be applied to flows with nonlinear departures from axisymmetry. He used these arguments to predict the way in which a class of exact solutions to the two-dimensional Euler equations found by Abrashkin & Yakubovich (1984) would respond to acoustic wave

radiation.

The Abrashkin & Yakubovich flows are categorized according to their degree of rotational symmetry, and the simplest class, with two-fold rotational symmetry, corresponds to the rotating ellipse of uniform vorticity found by Lamb (1932).

The analysis of Zeitlin (1988, 1991) assumed that the Abrashkin & Yakubovich flows adjusted to take account of gravity wave radiation simply by evolving to another such flow with the same degree of rotational symmetry and the same circulation but with less energy. For a rotating ellipse, this corresponds to increasing the aspect ratio of the ellipse while keeping its area constant. Most significantly, it implicitly assumes that, to a first approximation at least, the ellipse remains elliptical as it adjusts to the effect of gravity wave radiation.

Zeitlin's analysis might be regarded as unsatisfactory, however, since nothing is done to check that ellipses will simply elongate without significant change of form. The purpose of this section is to investigate the evolution of the rotating ellipse by a matched asymptotic analysis, thus enabling us to obtain the significant characteristics of the flow at every order in Froude number up to the order at which the vortical flow loses energy in response to gravity wave radiation. The present analysis is therefore able to investigate the validity of Zeitlin's assumption that the ellipse remains elliptical.

2.4.1 Evolution of the vortex boundary

A compact description of the rotating elliptical vortex requires construction of an "elliptical" coordinate system in order to describe the flow outside the vortex by means of elementary functions.

It is convenient to work in the complex plane, in which $z = x + iy$, where x and y are cartesian coordinates in physical space. One can then define elliptical coordinates ξ and η, via a new complex variable $w = \xi + i\eta$, such that $z = c\cosh w$. In w-space, the azimuthal coordinate η runs from 0 to 2π, and the ellipse in z-space lies within the region $0 < \xi < 1$ in w-space. For an ellipse with semi-major axis a and semi-minor axis b, Lamb (1932) showed that c is given by

$$c = (a^2 + b^2)^{1/2}. \tag{2.110}$$

An alternative description, which appears to be more readily generalizable in the present

context, is to work in $\ln w$-space, to which we assign the complex variable ζ. In the following analysis, we use a conformal mapping from z-space to ζ-space such that the ellipse in z-space is mapped onto the unit disc $|\zeta| < 1$ in ζ-space.

Recently, Legras & Zeitlin (1992) have used this formalism to develop the elliptical moment model of Legras & Dritschel (1991), Dritschel & Legras (1991). They begin by postulating a conformal mapping $f(\zeta)$ from the exterior of the unit disk in the ζ-plane to the exterior of a simply connected area in the z-plane, given by

$$f(\zeta) = \Gamma\zeta + \frac{\nu}{\zeta} + \sum_{i=3}^{\infty} \frac{\mu_i}{\zeta^i}. \tag{2.111}$$

Without loss of generality, Γ is taken to be real throughout. They were interested in the interaction between a number of different ellipses in two dimensional incompressible flow, in which the shape of each ellipse was represented in the form (2.111). The essence of their approach was that the evolution of the ellipses could be described by obtaining evolution equations for ν and the μ_i for each ellipse. They developed their so-called elliptical moment model by assuming that the μ_i were small, and so neglecting nonlinear products of the μ_i in the derivation of the evolution equations.

To obtain the evolution equations for the coefficients of the conformal map, it is convenient to consider the Lagrangian time derivative of z.

We start from the statement

$$\dot{z} \equiv \frac{Dz}{Dt} = u + iv \tag{2.112}$$

and convert this via (2.111) into its equivalent statement in the ζ–plane

$$\frac{Df}{Dt} \equiv \frac{\partial f}{\partial t} + \frac{\partial f}{\partial \zeta}\dot{\zeta} = u + iv. \tag{2.113}$$

Rearranging (2.113), we obtain

$$\frac{\dot{\zeta}}{\zeta} + \frac{1}{\zeta\frac{\partial f}{\partial \zeta}}\frac{\partial f}{\partial t} = \frac{u + iv}{\zeta\frac{\partial f}{\partial \zeta}}. \tag{2.114}$$

Now, since the elliptical vortex boundary is a material line in z-space, so it must also be a material line in ζ–space. Recalling that in ζ–space the vortex boundary is the unit circle, particle velocities on $|\zeta| = 1$ must be tangent to the unit circle, and therefore on $|\zeta| = 1$, we

will have

$$\mathrm{Re}\left[\frac{\dot{\zeta}}{\zeta}\right] = 0, \tag{2.115}$$

and hence

$$\mathrm{Re}\left[\frac{1}{\zeta f'}\frac{\partial f}{\partial t}\right] = \mathrm{Re}\left[\frac{u+iv}{\zeta f'}\right]. \tag{2.116}$$

Legras & Zeitlin (1992) now invoke a classical theorem of complex analysis, which states that a complex function which is analytic in $|z| < 1$ is known, up to a purely imaginary constant, if its real part is known in the unit circle $|z| = 1$. Hence, the right hand side of (2.116) can be used to define a function $h(\zeta)$, which is analytic in $1/\zeta$, and for which

$$\mathrm{Re}\, h(\zeta)|_{|\zeta|=1} = \mathrm{Re}\left[\frac{u+iv}{\zeta f'}\right]_{|\zeta|=1}. \tag{2.117}$$

The evolution of the conformal mapping function $f(\zeta)$ can then be obtained from

$$\frac{\partial f}{\partial t} = \zeta\frac{\partial f}{\partial \zeta}h(\zeta). \tag{2.118}$$

The temporal evolution of the vortex is now described entirely by the temporal evolution of the conformal map $f(\zeta)$, using (2.118). If $h(\zeta)$ is expressed in the form of a Laurent series in ζ^{-1}, setting the (arbitrary) imaginary part of the constant term in the expansion to zero is equivalent to imposing that Γ remains real throughout the evolution of the vortex.

2.4.2 The solution at leading order in F

We shall begin by obtaining Kirchoff's famous solution for the rotation rate of the elliptical vortex (see Lamb 1932) using the method outlined in the previous section. Although there are many ways to derive the solution, the presentation given here is intended to describe how the method presented above will work when we proceed to obtain evolution equations for the coefficients of the conformal map at higher orders in F. Without loss of generality, we shall assume that the potential vorticity within the ellipse exceeds its value outside the ellipse by unity.

To obtain the leading order streamfunction ψ_0, we must solve

$$\nabla^2\psi_0 = \begin{cases} 1 & \text{inside ellipse} \\ 0 & \text{outside ellipse} \end{cases}. \tag{2.119}$$

Using the complex variables z inside the ellipse and ζ outside it, (2.119) becomes

$$4\frac{d^2\psi_0^{(i)}}{dzd\bar{z}} = 1, \qquad |\zeta| < 1 \tag{2.120}$$

$$4\frac{d^2\psi_0^{(e)}}{d\zeta d\bar{\zeta}} = 0, \qquad |\zeta| > 1, \tag{2.121}$$

where $\psi_0^{(i)}$ and $\psi_0^{(e)}$ are the expressions for the leading order streamfunction for the interior and exterior of the ellipse respectively. Integrating these equations, and imposing continuity of ψ and its normal derivative on the boundary of the ellipse, we obtain

$$\psi_0^{(i)} = \frac{1}{4}z\bar{z} - \frac{1}{8\Gamma}\left(\bar{\nu}z^2 + \nu\bar{z}^2\right) - \frac{1}{4}\left(\Gamma^2 - |\nu|^2\right) \tag{2.122}$$

$$\psi_0^{(e)} = \frac{1}{4}\left(\Gamma^2 - |\nu|^2\right)\ln\zeta\bar{\zeta} + \frac{1}{8\Gamma}\left(\Gamma^2 - |\nu|^2\right)\left(\nu\zeta^{-2} + \bar{\nu}\bar{\zeta}^{-2}\right) \tag{2.123}$$

as the leading order solution for ψ.

It is straightforward to show that the velocity, on the boundary of the ellipse, is then

$$\frac{1}{2\Gamma}i(\Gamma^2 - |\nu|^2)\zeta. \tag{2.124}$$

To obtain the evolution equations for Γ and ν, we note that

$$\mathrm{Re}\left[\frac{u+iv}{\zeta f'}\right] = \mathrm{Re}\left[\frac{i}{2\Gamma}\frac{\Gamma^2 - |\nu|^2}{\Gamma - \nu/\zeta^2}\right] = \mathrm{Re}\left[\frac{i}{2\Gamma^2}\frac{\nu(\Gamma^2 - |\nu|^2)}{\Gamma\zeta^2 - \nu}\right], \tag{2.125}$$

and hence, by (2.118),

$$\frac{\partial f}{\partial t} = \frac{1}{2}i\nu\left(1 - \frac{|\nu|^2}{\Gamma^2}\right)\zeta^{-1}, \tag{2.126}$$

which implies

$$\frac{d\Gamma}{dt} = 0; \quad \frac{d\nu}{dt} = \frac{1}{2}i\nu\left(1 - \frac{|\nu|^2}{\Gamma^2}\right), \tag{2.127}$$

with solution

$$\Gamma = \text{constant}; \quad |\nu| = \text{constant}; \quad \nu = |\nu|e^{i\Omega t}; \quad \Omega = \frac{1}{2}\left(1 - \frac{|\nu|^2}{\Gamma^2}\right). \tag{2.128}$$

There are no higher order matching terms which would require the introduction of any of the μ_i terms in (2.111), thus confirming that at leading order the ellipse remains elliptical.

The rotation rate of the vortex is therefore $(1 - |\nu|^2/\Gamma^2)/4$, the factor of four rather than two coming from the fact that the conformal mapping description of the vortex boundary is unable to distinguish between 0^o and 180^o.

2.4.3 The solution to $O(F^2)$

To proceed to higher order in F, we must first compute the leading order height field h_0. It is convenient to work with the Bernoulli function B, for which

$$\nabla^2 B_0 = \begin{cases} 1+f & \text{inside ellipse} \\ 0 & \text{outside ellipse} \end{cases} . \tag{2.129}$$

In the interior of the ellipse, h_0 can be readily obtained from B_0 via the relation $h_0 = B_0 - \frac{1}{2}\boldsymbol{u}_0^2$. Outside the ellipse, a possible difficulty arises, since the calculation of $\boldsymbol{u}_0$ from ψ_0 will introduce a Jacobian factor:

$$h_0^{(e)} = B_0^{(e)} - 2\left(\frac{d\psi_0^{(e)}}{d\zeta}\frac{d\psi_0^{(e)}}{d\bar{\zeta}}\right) \times \left(\frac{df}{d\zeta}\frac{d\bar{f}}{d\bar{\zeta}}\right)^{-1} . \tag{2.130}$$

At first, the appearance of the Jacobian factor in (2.130) would seem to imply that, outside the ellipse, the expression for h_0 will contain all even inverse powers of ζ. and the simplicity of the analytical expression of the rotating elliptical vortex solution would be lost. Fortunately, however, (2.130) can be simplified. After imposing continuity of h_0 and its normal derivative across the elliptical boundary, we obtain

$$\begin{aligned} h_0^{(i)} &= \frac{1}{16}\left(1+2f-|\nu|^2\right)\left(2z\bar{z} - \frac{\bar{\nu}}{\Gamma}z^2 - \frac{\nu}{\Gamma}\bar{z}^2\right) \\ &- \frac{1}{4}\left(\Gamma^2(1+f) - |\nu|^2(2+f)\right) + \text{constant} \qquad (2.131) \\ h_0^{(e)} &= \frac{1}{4}f\left(\Gamma^2 - |\nu|^2\right)\ln\zeta\bar{\zeta} + \frac{1}{16\Gamma^3}\left(\Gamma^4(1+2f) - 2|\nu|^2\Gamma^2(1+f) + |\nu|^4\right)\left(\nu\zeta^{-2} + \bar{\nu}\bar{\zeta}^{-2}\right) \\ &- \frac{1}{8\Gamma^2}\left(\Gamma^2 - |\nu|^2\right)^2(\zeta\bar{\zeta})^{-1} + \text{constant}. \qquad (2.132) \end{aligned}$$

The Jacobian factor has been cancelled exactly from the denominator by the derivates of ψ. It follows that, as in the case of ψ_0, only quadratic terms in z or ζ^{-1} and their conjugates need be retained in the solution.

The arbitrary constant in (2.131) and (2.132) must be determined through the matching conditions to the outer flow.

We first note that in the limit $r \to \infty$, (2.123) implies

$$\psi_0^{(e)} \sim \frac{1}{2}\left(\Gamma^2 - |\nu|^2\right) \ln r. \tag{2.133}$$

From (2.81), this implies an outer scale flow streamfunction

$$\psi_0 = -\frac{1}{2}\left(\Gamma^2 - |\nu|^2\right) K_0(|f|Fr), \tag{2.134}$$

and a height field

$$h_0 = -f\frac{1}{2}\left(\Gamma^2 - |\nu|^2\right) K_0(|f|Fr). \tag{2.135}$$

This in turn induces a uniform correction to h in the vortical flow:

$$h_{0l} = f\frac{1}{2}\left(1 - |\nu|^2\right)(\ln(|f|F\Gamma) + \gamma), \tag{2.136}$$

which implies an $O(F^2)$ streamfunction in the vortical region:

$$\nabla^2\psi_{2l} = \begin{cases} (1+f)h_{0l} & \text{inside ellipse} \\ fh_{0l} & \text{outside ellipse} \end{cases}, \tag{2.137}$$

which has components at $O(F^2)$ and $O(F^2 \ln F)$.

Converting into complex coordinates and solving this equation, we obtain the streamfunction inside the elliptical vortex:

$$\psi_{2l}^{(i)} = h_{0l}\left(\frac{1}{4}z\bar{z}(1+f) - \frac{1}{8\Gamma}(\bar{\nu}z^2 + \nu\bar{z}^2) - \frac{1}{4}(\Gamma^2 - |\nu|^2)\right), \tag{2.138}$$

and the streamfunction outside the ellipse:

$$\begin{aligned} \psi_{2l}^{(e)} &= \frac{1}{8}h_{0l}\Big(2f\Gamma^2\zeta\bar{\zeta} + 2f\Gamma(\bar{\nu}\zeta\bar{\zeta}^{-1} + \nu\bar{\zeta}\zeta^{-1}) \\ &+ (\Gamma^2 - |\nu|^2)\left(2\ln\zeta\bar{\zeta} + \frac{\nu}{\Gamma}\zeta^{-2} + \frac{\bar{\nu}}{\Gamma}\bar{\zeta}^{-2}\right) + 2f|\nu|^2\zeta^{-1}\bar{\zeta}^{-1}\Big). \end{aligned} \tag{2.139}$$

Repeating the procedure to obtain the leading order solution described in §4.2, we obtain an evolution equation for ν at $O(F^2 \ln F)$:

$$\frac{d\nu}{dT_l} = \frac{1}{2}ih_{0l}\nu(2f + 1 - (|\nu|/\Gamma)^2). \tag{2.140}$$

There is no correction to the higher μ_i coefficients at $(F^2 \ln F)$.

Equation (2.140) implies a correction to the rotation rate ω at $O(F^2 \ln F)$, such that

$$\omega = \frac{1}{4}\left(\Gamma^2 - |\nu|^2 + F^2 \ln F\frac{1}{2}f(\Gamma^2 - |\nu|^2)(2f + 1 - (|\nu|/\Gamma)^2) + O(F^2)\right). \tag{2.141}$$

Recalling that $\ln F < 0$, this corresponds to a reduction in the rotation rate if and only if $f(2f + 1 - (|\nu|/\Gamma)^2) > 0$. The implication is that, as $|f|$ becomes large, the vortex always slows down compared with its incompressible rotation rate. Its rotation rate increases only if $-(1 - (|\nu|/\Gamma)^2)/2 < f < 0$.

We are now in a position to write down the equations for the second order velocity potential and streamfunction ϕ_2 and ψ_2 respectively.

$$4\frac{d^2\phi_2^{(i)}}{dzd\bar{z}} = -\frac{1}{32\Gamma^5}i(\Gamma^2 - |\nu|^2)(\Gamma^2(1+2f) - |\nu|^2)(\bar{\nu}z^2 - \nu\bar{z}^2) \tag{2.142}$$

$$\begin{aligned} 4\frac{d^2\phi_2^{(e)}}{d\zeta d\bar{\zeta}} = &- \frac{1}{32\Gamma^4}i(\Gamma^2 - |\nu|^2)^2(\Gamma^2(1+2f) - |\nu|^2)(\nu^2\zeta^{-4} - \bar{\nu}^2\bar{\zeta}^{-4}) \\ &- \frac{1}{32\Gamma^3}i(\Gamma^2 - |\nu|^2)^2(\Gamma^2(1-2f) - |\nu|^2)(\nu\zeta^{-2} - \bar{\nu}\bar{\zeta}^{-2}) \\ &- \frac{1}{16\Gamma^5}(\Gamma^2 - |\nu|^2)^2(\Gamma^4 - |\nu|^4 - f\Gamma^2|\nu|^2)(\nu\zeta^{-4}\bar{\zeta}^{-2} - \bar{\nu}\bar{\zeta}^{-4}\zeta^{-2}) \\ &+ \frac{1}{8\Gamma^3}i(\Gamma^2 - |\nu|^2)^3(\nu\zeta^{-3}\bar{\zeta}^{-1} - \bar{\nu}\bar{\zeta}^{-3}\zeta^{-1}) \end{aligned} \tag{2.143}$$

$$4\frac{d^2\psi_2^{(i)}}{dzd\bar{z}} = (1+f)\left[\frac{1}{16}\left(1 + 2f - |\nu|^2\right)\left(2z\bar{z} - \frac{\bar{\nu}}{\Gamma}z^2 - \frac{\nu}{\Gamma}\bar{z}^2\right) - \frac{1}{4}\left(\Gamma^2(1+f) - |\nu|^2(2+f)\right)\right] \tag{2.144}$$

$$\begin{aligned} 4\frac{d^2\psi_2^{(e)}}{d\zeta d\bar{\zeta}} = f(\Gamma - \nu\zeta^{-2})(\Gamma - \bar{\nu}\bar{\zeta}^{-2}) \Big[&\tfrac{1}{4}f\left(\Gamma^2 - |\nu|^2\right)\ln\zeta\bar{\zeta} \\ &+ \tfrac{1}{16\Gamma^3}\left(\Gamma^4(1+2f) - 2|\nu|^2\Gamma^2(1+f) + |\nu|^4\right)(\nu\zeta^{-2} + \bar{\nu}\bar{\zeta}^{-2}) \\ &\tfrac{1}{8\Gamma^2}\left(\Gamma^2 - |\nu|^2\right)^2(\zeta\bar{\zeta})^{-1}\Big] \end{aligned} \tag{2.145}$$

The result is that ϕ_2 and ψ_2 will require only constant, logarithmic, quadratic and quartic terms in their expressions, and the full solution of the problem to $O(F^2)$ can be written down in a finite expression involving only products of algebraic and logarithmic functions. The evolution of the boundary of the ellipse will require μ_3 to be non-zero at order $O(F^2)$, but all the other μ_i may be set to zero without approximation.

A considerable amount of manipulation is required to obtain expressions for ϕ_2 and ψ_2, and a program was written to undertake the integration and matching using the REDUCE symbolic manipulation language.

After obtaining the $O(F^2)$ velocity fields and substituting into (2.111 - 2.118), one obtains:

$$\begin{aligned}
\frac{d\mu_3}{dt} &= -\frac{i}{2}\bar{\mu_3}\frac{\nu^4}{\Gamma^4} - 2i\mu_3|\frac{\nu|^2}{\Gamma^2} + \frac{3i}{2}\mu_3 \\
&+ \frac{i\Gamma\nu^2}{96}\left(1-(|\nu|/\Gamma)^2\right)\ \Big[f^2\left(6-2\left(|\nu|/\Gamma\right)^2\right) \\
&\qquad +f\left(1-(|\nu|/\Gamma)^2\right)\left(7+3\left(|\nu|/\Gamma\right)^2\right) \\
&\qquad +2\left(1-(|\nu|/\Gamma)^2\right)^2\left(1+(|\nu|/\Gamma)^2\right)\Big] \qquad (2.146)\\
\frac{d\mu_1}{dt}+\frac{d\nu}{dT} &= \frac{i\nu|\nu|^2}{2\Gamma^3}(\mu_{-1}+\bar{\mu_{-1}}) \\
&+ \frac{i}{2}(\mu_1 - 2(|\nu|/\Gamma)^2\mu_1 - (\nu/\Gamma)^2\bar{\mu_1}) \\
&+ \frac{i}{4}(1-(|\nu|/\Gamma)^2)(4(\nu/\Gamma)^3\bar{\mu_3}+2(\bar{\nu}/\Gamma)\mu_3+(\bar{\nu}|\nu|^2/\Gamma^3)\mu_3) \\
&+ \frac{i\nu}{96}(1-(|\nu|/\Gamma)^2)\ (f^2(4+18(|\nu|/\Gamma)^2-2(|\nu|/\Gamma)^4) \\
&\qquad -f(1-(|\nu|/\Gamma)^2)(29-2(|\nu|/\Gamma)^2-3(|\nu|/\Gamma)^4) \\
&\qquad -(1-(|\nu|/\Gamma)^2)^2(11+(|\nu|/\Gamma)^2-2(|\nu|/\Gamma)^4)) \\
&+ \frac{i}{4}f\nu(\Gamma^2-|\nu|^2)(2f+1-(|\nu|/\Gamma)^2)(\ln(|f|\Gamma)+\gamma) \qquad (2.147)\\
\frac{d\mu_{-1}}{dt} &= \frac{1}{4}((\nu^2|\nu|^2/\Gamma^4)\bar{\mu_3}+(\bar{\nu}/\Gamma)^2\mu_3-(\nu/\Gamma)^2\bar{\mu_3}-(\bar{\nu}^2|\nu|^2/\Gamma^4)\mu_3). \qquad (2.148)
\end{aligned}$$

The solution for ν is known on the shortest timescale t, and this system of equations describes oscillations of μ_3, μ_1 and μ_{-1}, where μ_1 and μ_{-1} are used to represent $O(F^2)$ corrections of ν and Γ respectively. The oscillations are forced in the μ_3 equation, with a forcing frequency of 2Ω – i.e. four times the rotational frequency of the vortex. Forcing in the μ_1 equation is secular, and is absorbed into the $d\nu/dT$ component, leading to an $O(F^2)$ modification to the rotation rate of the vortex.

It is convenient to investigate solutions for μ_3 with the frequency of the forcing. These solutions will give weakly perturbed elliptical vortex shapes which rotate without change of form. Vortical patches which rotate without change of form have been termed V-states by Deem & Zabusky (1978), the simplest example of a V-state being the rotating ellipse of uniform

vorticity in the two-dimensional incompressible Euler equations.

Once the V-states of our system are found, the forcing is effectively removed from the system of equations describing the $O(F^2)$ perturbations to the elliptical vortex shape. They are then the same equations which describe free perturbations to an elliptical vortex boundary in incompressible flow. The perturbations will therefore be oscillatory provided the vortex in stable to incompressible mode-4 perturbations, and exponential otherwise.

The stability of an incompressible elliptical vortex was investigated by Love (1893). He found that the ellipse was always stable if it had an aspect ratio $\lambda = (\Gamma + |\nu|)/(\Gamma - |\nu|)$ of less than 3. At an aspect ratio $\lambda = 3$, the ellipse became unstable to mode 3 perturbations (i.e. $\mu_2 \neq 0$). In the present study, we impose 180^o rotational symmetry, and therefore we do not encounter mode 3 perturbations. As the aspect ratio is increased further, so successively more modes become unstable. The first one of interest to us here is mode 4, which becomes unstable when the aspect ratio $\lambda = (\sqrt{2} + 2\sqrt{\sqrt{2}-1})(1 + 1/\sqrt{2}) \approx 4.61$. In the variables Γ and $|\nu|$ to be used here, this corresponds to $|\nu|/\Gamma = \sqrt{\sqrt{2}-1}$.

2.4.4 V-states for weakly divergent flow

In finding the V-state with 180^o rotational symmetry for incompressible flow, the conformal map (2.111) from the unit disc in ζ-space to the shape of the rotating vortex patch in z-space was postulated. The evolution equations for the coefficients showed that $\mu_i = 0$ for all i was a consistent solution. A single evolution equation remained for ν, and it followed that $|\nu|$ was a constant, corresponding to vortex patch rotating at a constant rate.

With the effects of divergence introduced, we have seen that μ_3 may not be assumed to be zero for all time, although all the other μ_i for $i > 3$ may be taken to be zero.

We start by searching for a solution for μ_3 of the form

$$\mu_3(t) = Me^{2i\Omega t}, \tag{2.149}$$

where M is a constant (which will depend on $|\nu|$), and $\Omega = (\Gamma^2 - |\nu|^2)/2$. This solution should be regarded as the particular integral of (2.148). Substituting (2.149) into (2.148) and factorizing,

we obtain

$$M = -\frac{\Gamma|\nu|^2(1-(|\nu|/\Gamma)^2)}{48(1-2(|\nu|/\Gamma)^2-(|\nu|/\Gamma)^4)}\ (f^2(6-2(|\nu|/\Gamma)^2)+f(1-(|\nu|/\Gamma)^2)(7+3(|\nu|/\Gamma)^2) \\ +2(1-(|\nu|/\Gamma)^2)^2(1+(|\nu|/\Gamma)^2))\,. \tag{2.150}$$

Thus, provided that the denominator $1-2(|\nu|/\Gamma)^2-(|\nu|/\Gamma)^4 \neq 0$, there is a unique solution for M in terms of $|\mu|$. This is equivalent to the condition $(|\nu|/\Gamma)^2 \neq \sqrt{2}-1$. One can readily verify that μ_1 and μ_{-1} are may be set to zero with this solution. The rotation rate of the vortex is changed by $O(F^2)$.

It is also straightforward to show that the solution describes a vortex which is rotating without change of form. The conformal mapping function $f(\zeta)$ is now

$$f(\zeta) = \Gamma\zeta + |\nu|e^{i\Omega t}\zeta^{-1} + F^2 M e^{2i\Omega t}\zeta^{-3}. \tag{2.151}$$

Moving into a frame of reference rotating with the vortex, we define the complex coordinate in the rotating frame Z by $\zeta = Ze^{i\Omega t/2}$. Then

$$f(\zeta) = e^{i\Omega t/2}\left(\Gamma Z + |\nu|Z^{-1} + F^2 M Z^{-3}\right). \tag{2.152}$$

In the rotating Z–frame, therefore, the conformal mapping is rotating at a constant rotation rate $\Omega/2$. Apart from this uniform background rotation, the coefficients of the mapping are constants, and hence the vortex rotates without change of shape.

From (2.150), we can see that M can be positive or negative, depending upon $|\nu|/\Gamma$ and f. In figures 2.1–2.4 rotating V-state shapes are shown for various $|\nu|$ and various M, both positive and negative.

The ellipses are normalised to have a semi-major axis of unit length. A clear distinction is to be drawn between positive M and negative M. When M is positive, the ellipses seem to become more distorted towards square shapes, and for larger values of M appear to develop cusps. When M is negative, however, the ends of the ellipses tend to round out, and the V-states become more peanut-shaped. Now, it also follows from (2.150) that for moderate to large $|f|$, we will have $M < 0$, and therefore the $M < 0$ solutions should be qualitatively similar to the V-states of the single layer quasi-geostrophic equations. A study of these V-states was carried

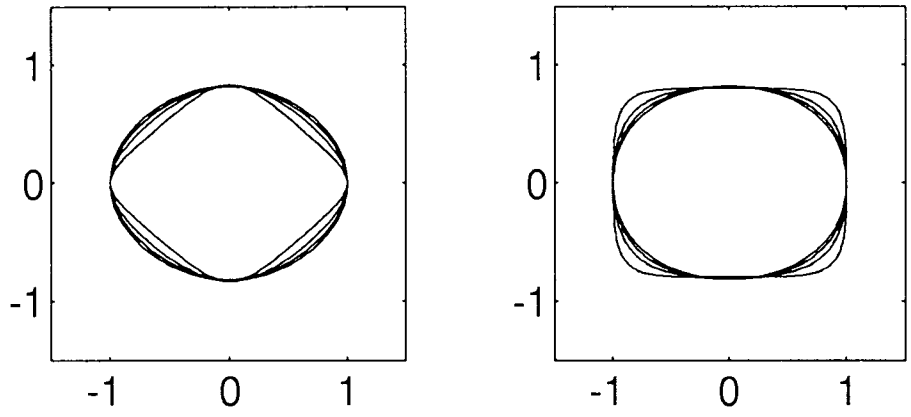


Figure 2.1: Shapes of rotating weakly compressible steadily rotating V-states for positive μ_3 (left) and negative μ_3 (right), with small aspect ratio ($|\nu|/\Gamma = 0.1$)

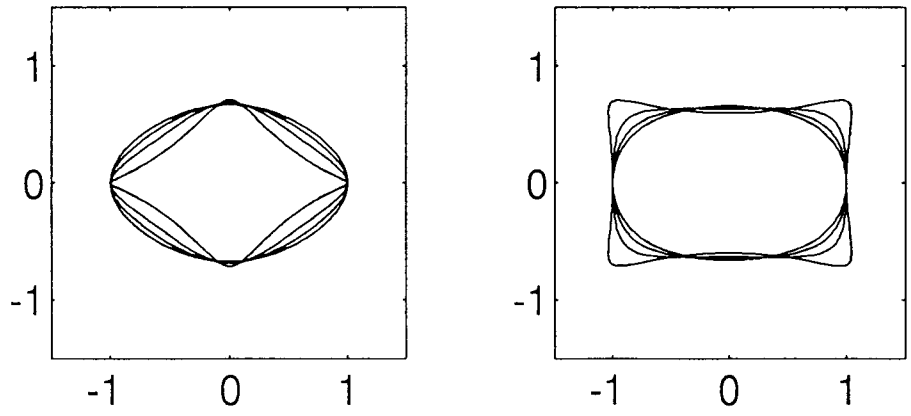


Figure 2.2: As figure 2.1, but with $|\nu|/\Gamma = 0.2$

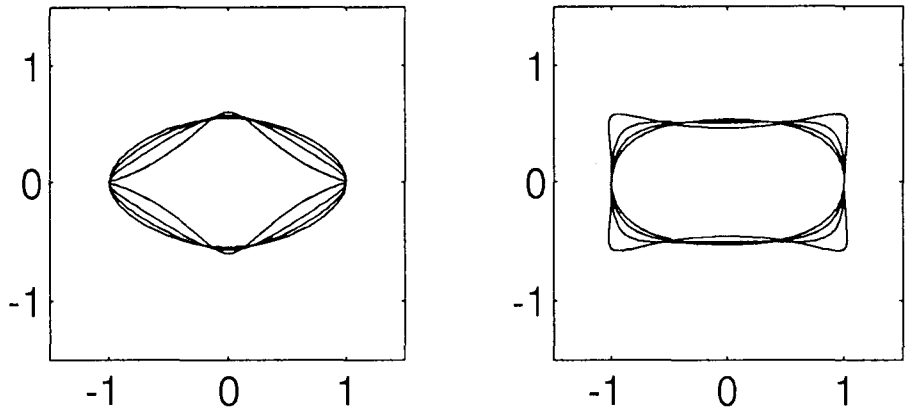


Figure 2.3: As figure 2.1, but with $|\nu|/\Gamma = 0.3$

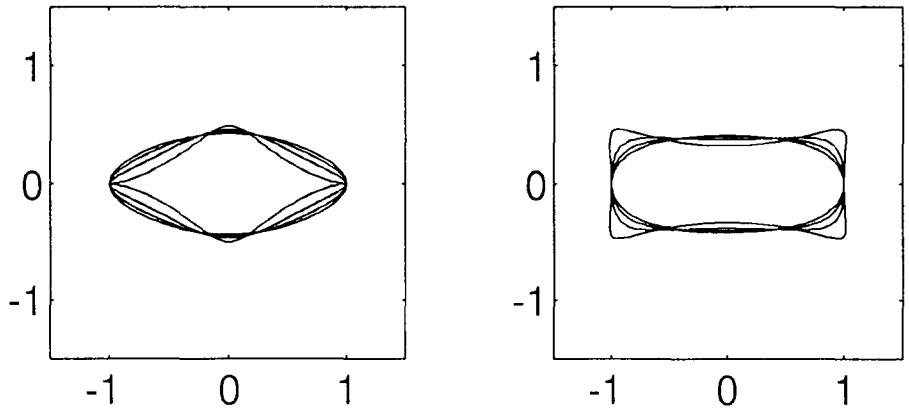


Figure 2.4: As figure 2.1, but with $|\nu|/\Gamma = 0.4$

out by Polvani *et al.* (1989), who found that the V-states of the single layer quasigeostrophic equations were more peanut-shaped than their elliptical counterparts for the incompressible Euler equations. It seems, therefore, that the asymptotic expansion presented here is capable of capturing the qualitative of V-states of the quasigeostrophic system.

In addition to the $O(F^2)$ change of shape of the V-states, the rotation rate is adjusted by an amount of $O(F^2)$, due to local divergence effects. However, except when the background rotation is absent, the dominant correction to the rotation rate comes about from matching conditions with the wave scale flow. The correction introduced from the matching conditions when background rotation is present is $O(F^2 \ln F)$ (equation 2.141).

2.4.5 The effect of gravity wave radiation

We turn now to the effect of gravity wave radiation on the rotating nearly-elliptical vortex. As we have seen, the effect of wave radiation does not enter the problem at $O(F^2)$, but does enter the problem at $O(F^4)$, in the form of a matching condition in the far field, imposing a uniform straining flow.

If one attempts to perform a complete asymptotic analysis of the fields to $O(F^4)$, it is not possible to express them in closed analytical form, unlike the $O(1)$, $O(F^2)$ and $O(F^2 \ln F)$ fields. The problem arises because, although the Bernoulli function to second order in F^2 outside the ellipse can be expressed in a finite sum of powers of ζ and $\bar{\zeta}$, multiplied, by logarithms of ζ and $\bar{\zeta}$, obtaining the height field from the Bernoulli function involves dividing by the Jacobian of the transformation. At $O(F^2)$, unlike the leading order analysis, it does not seem possible to factor out the Jacobian from the corresponding numerator. Thus although the second order height field may be expressed in closed (albeit cumbersome) form, integration with respect to ζ and $\bar{\zeta}$ will not lead to a closed form expression for the velocity potential at $O(F^4)$.

It follows that at $O(F^4)$, the representation of any V-state by a conformal mapping of form (2.111)) will require $\mu_i \neq 0$ for all odd i. However, it is clear that even if it were possible to obtain equations for the evolution of the μ_i, the contributions from the fields obtained in the way described above would give rise only to adjustments to the rotation rate of the structure at $O(F^4)$, and some further $O(F^4)$ corrections to the amplitudes of ν and the μ_i.

On the other hand, matching of the leading order solution to the gravity wave region will

give rise to a radiating wave field, one result of which is to introduce in the wave field a Hankel function of the appropriate type, corresponding to a radiating boundary condition. Matching conditions then imply that a large scale straining flow of form $r^2 \sin 2\theta$ will be introduced at $O(F^4)$ in the region of the vortex, in a sense which will cause the ellipse to elongate and lose energy, consistent with the associated radiating wave field. Since we know that ellipses remain elliptical under the influence of a large scale straining flow (Kida, 1981), it follows that we may simply regard ν as a slowly varying parameter.

To see this in detail, we must first obtain the wave-like part of the far field. First, therefore, we must obtain the time fluctuating part of the far field of $\psi_0^{(i)}$ in the limit $r \to \infty$. In doing so, however, we must recall that the ζ-coordinate is itself evolving in time, and hence the $\ln \zeta\bar{\zeta}$ term will contribute to the quadrupolar far field in (2.123).

From (2.123), the time-fluctuating part of $\psi_0^{(e)}$ is

$$-\frac{\Gamma}{8r^2}|\nu|(\Gamma^2-|\nu|^2)\left(e^{-i(2m\theta-\Omega t)}+e^{i(2m\theta-\Omega t)}\right). \tag{2.153}$$

Matching onto the wave field, this implies a leading order ($O(F^2)$) wave field, which consists of wavelike perturbations for streamfunction ψ and velocity potential ϕ, related via (2.85). In this case, we therefore have

$$-i\Omega\psi + f\phi = 0 \tag{2.154}$$

$$\phi + i\psi \sim \frac{i\Gamma}{8r^2}|\nu|(\Gamma^2-|\nu|^2)\left(e^{-i(2m\theta-\Omega t)}+e^{i(2m\theta-\Omega t)}\right) \quad \text{as } r \to 0, \tag{2.155}$$

which implies $O(F^2)$ wave fields:

$$\psi_2 = \frac{i\pi\Gamma}{16}\frac{f}{\Omega+f}|\nu|(\Gamma^2-|\nu|^2)(\Omega^2-f^2)\mathrm{H}_2^{(1)}((\Omega^2-f^2)^{1/2}R)e^{i(2m\theta-\Omega t)} \tag{2.156}$$

$$\phi_2 = -\frac{\pi\Gamma}{16}\frac{\Omega}{\Omega+f}|\nu|(\Gamma^2-|\nu|^2)(\Omega^2-f^2)\mathrm{H}_2^{(1)}((\Omega^2-f^2)^{1/2}R)e^{i(2m\theta-\Omega t)}. \tag{2.157}$$

The imaginary part of this corresponds to a streamfunction at $O(F^4)$ in the vortical region:

$$\psi_4 = -\frac{\pi\Gamma}{128}\frac{\Omega-f}{\Omega+f}(\Gamma^2-|\nu|^2)(\Omega^2-f^2)^2(\bar{\nu}z^2-\nu\bar{z}^2). \tag{2.158}$$

Adding this streamfunction to the velocity field at $O(F^4)$ and using (2.116), we obtain

$$\frac{d\Gamma^2}{dt} = F^4\frac{\pi}{32}\Gamma^2|\nu|^2(\Gamma^2-|\nu|^2)(\Omega^2-f^2)(\Omega-f)^2 \tag{2.159}$$

$$\frac{d|\nu|^2}{dt} = F^4\frac{\pi}{32}\Gamma^2|\nu|^2(\Gamma^2-|\nu|^2)(\Omega^2-f^2)(\Omega-f)^2. \tag{2.160}$$

Conservation of the circulation, $(\Gamma^2 - |\nu|^2)/4$, was not used explicitly in deriving these equations, and serves as an independent check on the analysis.

If we introduce the constant C, defined such that $C^2 = \Gamma^2 - |\nu|^2$, we can obtain an evolution equation for Γ^2:

$$\frac{d\Gamma^2}{dt} = F^4 \frac{\pi}{32\Gamma^6} C^2(\Gamma^2 - C^2)\left(\frac{1}{4}C^4 - \Gamma^4 f^2\right)\left(\frac{1}{2}C^2 - f\Gamma^2\right)^2. \tag{2.161}$$

Let us consider first the non-rotating case, which was considered by Zeitlin (1991). In this case, (2.161) reduces to:

$$\frac{d\Gamma^2}{dt} = \frac{\pi F^4}{512\Gamma^6} C^{10}(\Gamma^2 - C^2). \tag{2.162}$$

In this case, we may obtain the solution:

$$\Gamma^6/3 + C^2\Gamma^4/2 + C^4\Gamma^2 + C^6 \ln(\Gamma^2 - C^2) = \frac{\pi F^4 C^{10} t}{512} + \text{constant}. \tag{2.163}$$

Since $\Gamma^2 > C^2$, we know that Γ is a monotonically increasing function of time, and hence as $t \to \infty$ we will have

$$\Gamma \sim \left(\frac{3\pi}{512}\right)^{1/6} C^{5/3}(F^4 t)^{1/6} + O(1). \tag{2.164}$$

Since Γ increases without bound in this case, it follows from conservation of $\Gamma^2 - |\nu|^2$ that, for large t, $|\nu| \sim \Gamma - O(\Gamma^{-1})$. The ellipse has thus become infinitely long and this, with semi-major axis $\Gamma + |\nu| \to \infty$, and semi-minor axis $\Gamma - |\nu| \to 0$ as $t \to \infty$. The result is that the vortex evolves through a succession of V-states, with rapid oscillations about them on the $O(1)$ timescale.

However, we already know that if $(|\nu|/\Gamma)^2 = \sqrt{2} - 1$, the perturbation expansion breaks down because the ellipse departs from elliptical form as μ_3 in the conformal mapping (2.111) becomes unbounded. In the case $f = 0$, we also know that $\mu_3 < 0$, and hence we can make some qualitative statement about the nature of distortion of the ellipse.

In figure 2.5, the evolution of an ellipse with $f = 0$ is shown approaching the critical aspect ratio $(|\nu|/\Gamma)^2 = \sqrt{2} - 1$. The perturbation expansion suggests that, as the critical aspect ratio is reached, the ellipse becomes progressively more "peanut"-shaped. Eventually, the expansion procedure breaks down, and the mapping $f(\zeta)$ ceases to be single-valued. By this time, the lobes of the "peanut" have grown significantly, and there is a tendency to pinch in the middle, creating two separate vortices.

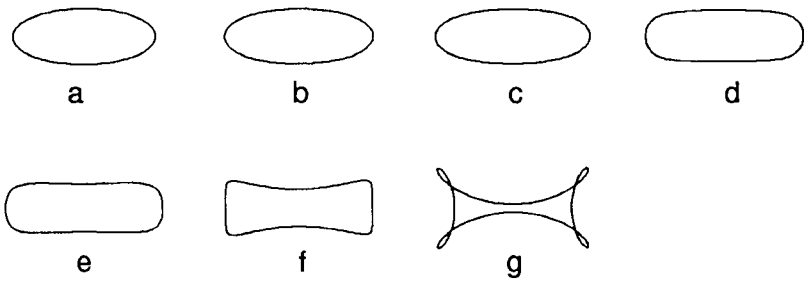


Figure 2.5: V-states of the rotating ellipse in the non-rotating frame ($f = 0$). Fig (a) represents the aspect ratio at which mode 3 instability would first occur, if 180^o symmetry were relaxed. Figures (a) to (g) show the V-states as the ellipse extends due to wave radiation. The splitting ratio is $(|\nu|/\Gamma)^2 = 0.4142$.

Some support for this scenario is provided by the analytical work of Williams (1992), and the numerical work of Chan *et al.* (1993). Williams performs a weakly nonlinear version of the present analysis, equivalent to assuming that $|\nu|/\Gamma = O(F^2)$. He shows that the weakly nonlinear development of the elliptical perturbation also leads to a "pinch" at a time of order $\ln F/F^4$. The numerical work of Chan *et al.* confirms this, but their Froude number is of order one. It seems that this analysis is in agreement with the existing literature, though not with Zeitlin's (1991) remarks, where he supposed that a secondary barotropic instability mechanism would have to be invoked to explain the subsequent breakup of the ellipse into two vortices. That is not the conclusion of the present study, nor is it the conclusion of Williams. The present study provides an interesting counterpart to the study of Williams in that it shows that $O(F^2)$ terms are sufficient to explain the principal dynamics, and one does not need to take into account terms at high order in F to obtain the nonlinear breakdown of the ellipse, even though they are required for the weakly nonlinear analysis.

Now, we notice that, as the aspect ratio of the ellipse increases, the frequency of the radiated waves $\Omega = (1-(|\nu|^2/\Gamma^2)/2 = C^2/(2\Gamma^2)$ decreases. Provided $\Omega > |f|$, waves will radiate and the ellipse will continue to elongate. However, if $\Omega < |f|$, then waves will not radiate, Γ^2 and $|\nu|^2$ will be constant, and the ellipse will not elongate with time. The values of the constants $|f|$ and C will determine whether the ellipse reaches the splitting aspect ratio, at which it breaks in two. At the splitting aspect ratio, the frequency of radiated waves is $\Omega_c = 1 - 1/\sqrt{2}$. Therefore,

if $|f| > 1 - 1/\sqrt{2}$, the ellipse will gradually approach the aspect ratio, at which it no longer excites radiating gravity waves. If, on the other hand, $|f| < 1 - 1/\sqrt{2}$, we expect the ellipse to split in two as it approaches the critical aspect ratio. It will then be necessary to determine the sign of μ_3 to describe the nature of the breakup of the ellipse.

To determine the sign of μ_3 at the splitting aspect ratio, we set $(|\nu|/\Gamma)^2 = \sqrt{2} - 1$ in (2.150) and determine the roots of $M = 0$ as a function of f. This leads to the equation

$$(4 - \sqrt{2})f^2 + (\sqrt{2} - 1)f + 6\sqrt{2} - 8 = 0, \tag{2.165}$$

which has solutions

$$f_1 = -(1 - 1/\sqrt{2}) \approx -0.29; \qquad f_2 = -2(3\sqrt{2} - 2)/7 \approx -0.64. \tag{2.166}$$

If f lies between f_1 and f_2, then $M > 0$, otherwise $M < 0$.

The magnitude of f_1, the smaller of the two roots of (2.166), is the same as the magnitude of the critical background rotation rate, above which the inhibiting effect of background rotation on the radiation will prevent ellipses from reaching their splitting aspect ratio. It follows that if $|f| < 1 - 1/\sqrt{2}$, we will always have $M < 0$ at the critical aspect ratio $(|\nu|/\Gamma)^2 = \sqrt{2} - 1$. Therefore, the nature of the breakup of an elliptical vortex by elongation due to gravity wave radiation can be described entirely by considering a succession of V-states in which $\mu_3 \to -\infty$ while $(|\nu|/\Gamma)^2 = \sqrt{2} - 1$ — the same as for the case without background rotation.

Finally, in cases where the background rotation is sufficiently strong to inhibit the breakup of the vortex, it is interesting to note the differences in the approach to the equilibrium aspect ratio of the ellipse between positive and negative f.

Returning to (2.161), we can write

$$\int \left(\frac{\alpha}{\Gamma^2 - C^2} + \frac{\beta}{f\Gamma^2 + C^2/2} + \frac{\gamma}{f\Gamma^2 - C^2/2} + \frac{\delta}{(f\Gamma^2 - C^2/2)^2} + \frac{\epsilon}{(f\Gamma^2 - C^2/2)^3} \right) d\Gamma^2 = \int \frac{F^4 \pi C^2}{32} dt \tag{2.167}$$

for constants $\alpha, \beta, \gamma, \delta, \epsilon$. For $f > 0$, at large times we will have

$$\int \frac{\epsilon d\Gamma^2}{(f\Gamma^2 - C^2/2)^3} \sim \int \frac{F^4 \pi C^2}{32} dt, \tag{2.168}$$

where $\epsilon = -C^2/(64f^2(1 - 2f))$. Hence, for large times, we will have

$$\Gamma^2 \sim \frac{C^2}{2f} - \frac{1}{2\sqrt{\pi} f (1 - 2f)^{1/2}} (F^4 t)^{-1/2}. \tag{2.169}$$

The cyclonic ellipse approaches its final aspect ratio algebraically in time for large times.

On the other hand, if $f < 0$ then, for large t, we have

$$\int \frac{\beta d\Gamma^2}{f\Gamma^2 + C^2/2} \sim \int \frac{F^4 \pi C^2}{32} dt, \tag{2.170}$$

where $\beta = 1/(64f^2C^2(1+2f))$. From this it follows that

$$\Gamma^2 - \frac{C^2}{2|f|} \propto \exp(-2\pi C^4 |f|^3 (1-2|f|) F^4 t), \tag{2.171}$$

so that, when $-1/2 < f < 0$, the final approach of the anticyclonic ellipse to its steady aspect ratio is exponential in time, and hence rather more rapid than the corresponding cyclonic case.

The particular case $f = 1 - 1/\sqrt{2}$ requires special attention. In this case, the radiating ellipse equilibrates at the critical aspect ratio. However, μ_3 is unbounded at the critical aspect ratio (unlike the corresponding anticyclonic case $f = -(1 - 1/\sqrt{2})$, in which case it remains bounded as it approaches its critical aspect ratio). Since only other radiation terms can affect the growth of Γ, it is probably necessary to consider higher order radiation fields to resolve this singularity, perhaps with f differing from $1 - 1/\sqrt{2}$ by $O(F^2)$. However, it seems that the principal features of the response of the ellipse to gravity wave radiation have been captured by the present analysis, and no further analysis has been attempted to clarify this remaining limiting case.

2.5 Discussion

The material presented in this chapter has focussed on the asymptotic limit of small Froude number. This is the limit in which the Lighthill analysis of aerodynamic sound generation is formally valid.

We have seen that it is possible to perform a matched asymptotic analysis and recover a quadrupolar far field, in a manner similar to Crow (1970) for the three-dimensional case. That result alone is non-trivial: if the incompressible velocity field is used in computing the Lighthill source term for a single two-dimensional eddy, the resulting integrals for the quadrupole moments are divergent.

The matched asymptotic analysis enabled us to analyse the general form of the back-reaction — i.e. the effect of the radiating gravity waves on the vortical flow. We showed that it was not

in general possible to determine the back reaction, which occurs at $O(F^4)$ in the source flow, in terms of the instantaneous potential vorticity field, even if one is permitted to integrate the equations of motion backwards in time. This is because history integrals are in general sensitive to the initial instant which is selected in the integration over the past times. However, if no radiating gravity waves were present, it was possible to obtain an expression for the $O(F^4)$ flow evaluating only instantaneous time derivatives of the potential vorticity field, and without reference to any initial instant. It follows that, at low Froude number, it is not in general possible to construct a balanced model which incorporates gravity wave emission without specifying a time in the past before which there are supposed to be no gravity waves.

An example of a flow subject to gravity wave radiation is the rotating ellipse of uniform potential vorticity. This flow was analysed to elucidate some more features of the effect of gravity wave radiation in a low Froude number limit. It was found that the ellipse tended to elongate in response to gravity wave radiation until it reached a certain aspect ratio, at which point it appeared to split into two vortices, and beyond which point the asymptotic analysis is not valid. The effect of the elongation is to reduce the radiation frequency of the quadrupolar gravity waves. Therefore, if the background rotation is sufficiently rapid, the ellipse might not reach its splitting aspect ratio, whereas for weaker background rotation the splitting aspect ratio is always attained on times such that $t = O(F^{-4})$.

If the splitting aspect ratio is not achieved, a remarkable degree of asymmetry in f is observed in the way in which the final aspect ratio of the ellipse is approached – exponentially in F^4t for anticyclones, but algebraically in F^4t for cyclones.

From time to time it is suggested that the nonlinear shallow water equations might exhibit a phenomenon of generalized adjustment, in which the vortical flow is continually adjusting itself to emit as few gravity waves as possible. Although the general analysis suggests that this is never truly possible for flows with a full spectrum of frequencies. the generalized adjustment hypothesis is supported by the model problem of the rotating ellipse. The ellipse elongates, and if the background rotation is sufficiently strong, will ultimately stop emitting gravity waves. If the background rotation is not sufficiently strong, it will split into two vortices, which will presumably move apart until they no longer radiate gravity waves. At that point, if they are sufficiently far apart, elliptical perturbation might once again grow on the boundaries of each separately, and the process will repeat itself. By this mechanism, the overall intensity of

radiation will decrease as more and more smaller vortices are responsible for the radiation. To see this, first note that the scaling law (2.9) can be modified to take account of the effect of the Coriolis force:

$$h' \sim \begin{cases} h_0 \left(\frac{l}{|\boldsymbol{x}|}\right)^{1/2} \left(\frac{u}{c_0}\right)^2 \left(\frac{(\omega^2 - f^2)^{1/2} l}{c_0}\right)^{3/2} & \omega^2 > f^2 \\ 0 & \omega^2 < f^2 \end{cases} . \tag{2.172}$$

Now, ω is set by the vorticity, and $u \sim \omega l$. Hence, keeping $l/|\boldsymbol{x}|$ fixed, we have

$$h' \sim l^{7/2}. \tag{2.173}$$

Each time the vortex splits, two vortices are formed, each with $l \longrightarrow l/\sqrt{2}$. Ultimately, we suppose that there are n vortices, each of dimension $l/\sqrt{n}$. Then

$$h' \sim n \times \left(\frac{l}{n^{1/2}}\right)^{7/2} \sim n^{-3/4}. \tag{2.174}$$

Alternatively, keeping $|\boldsymbol{x}|$ fixes, $h' \sim n^{-1}$. Consequently, as $n \to \infty$, the overall radiated gravity wave amplitude decreases, supporting the concept of generalized adjustment, in which the flow evolves so as to decrease the amplitude of the gravity waves it radiates.

There is, however, one key point about the generalized adjustment hypothesis. In the usual Rossby adjustment problem, the adjustment happens on the timescale of the inertial period, which is assumed short compared with the timescale of motions. Rossby adjustment is a fast process, as far as the balanced dynamics are concerned. For the type of adjustment due to gravity wave radiation experienced by the rotating ellipse, the opposite is true: the adjustment occurs on a very long timescale compared with the balanced dynamics.

2.6 Appendix A

The aim of this appendix is to show that the far field of (2.50) takes the form of the expected dipole when the logarithmic kernel is expanded for $|\boldsymbol{x}'| \ll |\boldsymbol{x}|$, and the error is an order of magnitude smaller, except for a logarithmic factor. We note that Crow (1970) remarks that his integrals will fail to converge after a certain number of moment expansions, but he does not rigorously address their asymptotic form. Ting & Miksis (1990) have presented a more careful analysis of the corresponding asymptotic forms in the three-dimensional problem, and

the analysis presented here is in the spirit of their paper, although the details are somewhat different.

In this appendix, we are considering the far field form of

$$\int_{\mathbb{R}^2} \nabla.((h_0 - f\psi_0)\boldsymbol{u}_0) \ln|\boldsymbol{x} - \boldsymbol{x}'| d^2\boldsymbol{x}'. \tag{2.175}$$

We note that, integrating (2.175) once by parts, this is equivalent to

$$\int_{\mathbb{R}^2} (h_0 - f\psi_0)\boldsymbol{u}_0 . \frac{\boldsymbol{x} - \boldsymbol{x}'}{|\boldsymbol{x} - \boldsymbol{x}'|^2} d^2\boldsymbol{x}'. \tag{2.176}$$

We start by considering the form of the far field for $\psi_0(\boldsymbol{x})$. In (2.25), we suppose that ζ_0 is a C^∞ function of two variables. We assume further that ζ_0 is of compact support, so $\exists \rho > 0$ such that $\zeta_0(\boldsymbol{x}) = 0 \ \forall |\boldsymbol{x}| > \rho$. Therefore $|\zeta_0|$ is bounded with $|\zeta_0| < \mathcal{Q} \ \forall \boldsymbol{x}$. It follows that, $\forall |\boldsymbol{x}| > 2\rho$,

$$\left|\frac{\partial \psi}{\partial x_i}\right| = \frac{1}{2\pi}\left|\int_{\mathbb{R}^2} \zeta_0(\boldsymbol{x}') \frac{x_i - x_i'}{|\boldsymbol{x} - \boldsymbol{x}'|^2} d^2\boldsymbol{x}'\right| \tag{2.177}$$

$$\leq \frac{\mathcal{Q}}{2\pi} \int_{|\boldsymbol{x}'|<\rho} \frac{1}{|\boldsymbol{x}| - \rho} d^2\boldsymbol{x}' \tag{2.178}$$

$$\leq \frac{\mathcal{Q}\rho^2}{|\boldsymbol{x}|}, \tag{2.179}$$

where we obtain (2.178) from (2.177) by using Schwartz' formula in the form $|\boldsymbol{x} - \boldsymbol{x}'| > |\boldsymbol{x}| - |\boldsymbol{x}'|$. Similarly

$$\left|\frac{\partial^2 \psi_0}{\partial x_i \partial x_j}\right| = \frac{1}{2\pi}\left|\int_{\mathbb{R}^2} \zeta_0(\boldsymbol{x}') \left(\frac{\delta_{ij}}{|\boldsymbol{x} - \boldsymbol{x}'|^2} - \frac{2(x_i - x_i')(x_j - x_j')}{|\boldsymbol{x} - \boldsymbol{x}'|^4}\right) d^2\boldsymbol{x}'\right| \tag{2.180}$$

$$\leq \frac{\mathcal{Q}}{2\pi} \int_{|\boldsymbol{x}'|<\rho} \left(\frac{1}{(|\boldsymbol{x}| - \rho)^2} + \frac{2}{(|\boldsymbol{x}| - \rho)^2}\right) d^2\boldsymbol{x}' \tag{2.181}$$

$$\leq \frac{6\mathcal{Q}\rho^2}{|\boldsymbol{x}|^2}. \tag{2.182}$$

Now let $\chi = B_0 - f\psi_0$. Since ζ_0 is a C^∞ function of compact support, so is $\zeta_0 \nabla\psi_0$, which is bounded with bound $\mathcal{U}$ (say). Then

$$|\chi(\boldsymbol{x})| = \frac{1}{2\pi}\left|\int_{\mathbb{R}^2} \zeta_0 \frac{\partial \psi_0}{\partial x_j'} \left(\frac{x_j - x_j'}{|\boldsymbol{x} - \boldsymbol{x}'|^2} - \frac{x_j}{|\boldsymbol{x}|^2}\right) d^2\boldsymbol{x}'\right| \tag{2.183}$$

$$= \frac{1}{2\pi}\left|\int_{\mathbb{R}^2} \zeta_0 \frac{\partial \psi_0}{\partial x_j}\left(\frac{2x_j \boldsymbol{x}.\boldsymbol{x}' - x_j'|\boldsymbol{x}|^2 - x_j|\boldsymbol{x}'|^2}{|\boldsymbol{x}|^2|\boldsymbol{x}-\boldsymbol{x}'|^2}\right) d^2\boldsymbol{x}'\right| \tag{2.184}$$

$$\leq \frac{\mathcal{U}}{2\pi}\int_{|\boldsymbol{x}'|<\rho} \frac{3\rho|\boldsymbol{x}|^2 + \rho^2|\boldsymbol{x}|}{|\boldsymbol{x}|^2(|\boldsymbol{x}|-\rho)^2} d^2\boldsymbol{x}' \tag{2.185}$$

$$\leq \frac{7\mathcal{U}\rho^3}{|\boldsymbol{x}|^2}, \tag{2.186}$$

where here we obtain (2.185) from (2.184) by expanding the numerator as a sum of the modulii of its components, and expand the denominator in the same way as in the step from (2.177) to (2.178). Similarly,

$$\left|\frac{\partial \chi}{\partial x_i}\right| = \frac{1}{2\pi}\left|\int_{\mathbb{R}^2} \zeta_0 \frac{\partial \psi_0}{\partial x_j'}\left(\frac{\delta_{ij}}{|\boldsymbol{x}-\boldsymbol{x}'|^2} - \frac{2(x_i-x_i')(x_j-x_j')}{|\boldsymbol{x}-\boldsymbol{x}'|^4} - \frac{\delta_{ij}}{|\boldsymbol{x}|^2} + \frac{2x x_i x_j}{|\boldsymbol{x}|^4}\right) d^2\boldsymbol{x}'\right| \tag{2.187}$$

$$\leq \frac{\mathcal{U}}{2\pi}\int_{|\boldsymbol{x}'|<\rho}\left(\frac{2|\boldsymbol{x}|\rho+\rho^2}{|\boldsymbol{x}|^2(|\boldsymbol{x}|-\rho)^2} + \frac{12|\boldsymbol{x}|^2\rho + 6|\boldsymbol{x}|^4\rho^2 + 2|\boldsymbol{x}|^2\rho^4}{|\boldsymbol{x}|^4(|\boldsymbol{x}|-\rho)^4}\right) d^2\boldsymbol{x}' \tag{2.188}$$

$$\leq \frac{126\mathcal{U}\rho^3}{|\boldsymbol{x}|^3}. \tag{2.189}$$

Now let $\boldsymbol{g} = (h_0 - f\psi_0)\boldsymbol{u}$. From (2.179, 2.182, 2.186, 2.189) we have:

Lemma 1 *$\exists A, B > 0$ such that $|g_i| < A|\boldsymbol{x}|^{-3}$ and $|\partial g_i/\partial x_j| < B|\boldsymbol{x}|^{-4}$ $\forall |\boldsymbol{x}| > 2\rho$ and for $i = 1,2$; $j = 1,2$.*

To proceed with our analysis of (2.176), we split the domain of integration $\mathbb{R}^2$ into three parts:

I $|\boldsymbol{x}'| < |\boldsymbol{x}| - 1$

II $|\boldsymbol{x}| - 1 < |\boldsymbol{x}'| < |\boldsymbol{x}| + 1$

III $|\boldsymbol{x}'| > |\boldsymbol{x}| + 1$

2.6.1 Region I

In region (I), we consider

$$\int_{|\boldsymbol{x}'|<|\boldsymbol{x}|-1} \boldsymbol{g}(\boldsymbol{x}')\cdot\left(\frac{\boldsymbol{x}-\boldsymbol{x}'}{|\boldsymbol{x}-\boldsymbol{x}'|^2} - \frac{\boldsymbol{x}}{|\boldsymbol{x}|^2}\right) d^2\boldsymbol{x}' \tag{2.190}$$

in (2.176). We split region (I) into three subregions:

Ia $|\boldsymbol{x}'| < 2\rho$

Ib $2\rho < |\boldsymbol{x}'| < |\boldsymbol{x}|/2$

Ic $|\boldsymbol{x}|/2 < |\boldsymbol{x}'| < |\boldsymbol{x}| - 1$

In region (Ia) we have

$$\left| \int_{Ia} g_j(\boldsymbol{x}') \frac{2x_j \boldsymbol{x}.\boldsymbol{x}' - x_j' |\boldsymbol{x}|^2 - x_j |\boldsymbol{x}'|^2}{|\boldsymbol{x}|^2 |\boldsymbol{x} - \boldsymbol{x}'|^2} \right| \tag{2.191}$$

$$\leq \mathcal{G} \int_{|\boldsymbol{x}'|<2\rho} \frac{6\sigma|\boldsymbol{x}|^2 + 4\sigma^2|\boldsymbol{x}|}{|\boldsymbol{x}|^2(|\boldsymbol{x}| - \sigma)^2} d^2\boldsymbol{x}' \tag{2.192}$$

$$\leq 8\pi\mathcal{G}\rho^3 \frac{3|\boldsymbol{x}| + 2\rho}{|\boldsymbol{x}|(|\boldsymbol{x}| - 2\rho)^2} \tag{2.193}$$

$$\sim O\left(\frac{1}{|\boldsymbol{x}|^2}\right), \tag{2.194}$$

where, in (2.192), $\sigma = |\boldsymbol{x}'|$.

In region (Ib) we have

$$\left| \int_{Ib} g_j(\boldsymbol{x}') \frac{2x_j \boldsymbol{x}.\boldsymbol{x}' - x_j' |\boldsymbol{x}|^2 - x_j |\boldsymbol{x}'|^2}{|\boldsymbol{x}|^2 |\boldsymbol{x} - \boldsymbol{x}'|^2} \right| \tag{2.195}$$

$$\leq 4\pi \int_{2\rho}^{\frac{1}{2}|\boldsymbol{x}|} \frac{A}{\sigma^3} \frac{3\sigma|\boldsymbol{x}|^2 + \sigma^2|\boldsymbol{x}|}{|\boldsymbol{x}|^2(|\boldsymbol{x}| - \sigma)^2} \sigma d\sigma \tag{2.196}$$

$$= \frac{4\pi A}{|\boldsymbol{x}|^2} \left[8 - \frac{4|\boldsymbol{x}|}{|\boldsymbol{x}| - 2\rho} + 3\ln(|\boldsymbol{x}| - 2\rho) - 3\ln 2\rho \right] \tag{2.197}$$

$$\sim O\left(\frac{\ln|\boldsymbol{x}|}{|\boldsymbol{x}|^2}\right). \tag{2.198}$$

In region (Ic) we have

$$\left| \int_{Ic} g_j(\boldsymbol{x}') \frac{2x_j \boldsymbol{x}.\boldsymbol{x}' - x_j' |\boldsymbol{x}|^2 - x_j |\boldsymbol{x}'|^2}{|\boldsymbol{x}|^2 |\boldsymbol{x} - \boldsymbol{x}'|^2} \right| \tag{2.199}$$

$$\leq \int_{\frac{1}{2}|\boldsymbol{x}|}^{|\boldsymbol{x}|-1} \frac{3}{|\boldsymbol{x}| - \sigma} \frac{A}{\sigma^3} \sigma d\sigma \tag{2.200}$$

$$= \frac{3A}{|\boldsymbol{x}|^2} \left[\ln(|\boldsymbol{x}| - 1) + 2 - \frac{|\boldsymbol{x}|}{|\boldsymbol{x}| - 1} \right] \tag{2.201}$$

$$\sim O\left(\frac{\ln|\boldsymbol{x}|}{|\boldsymbol{x}|^2}\right). \tag{2.202}$$

2.6.2 Region II

In region II we must consider the subregions

IIa $|\boldsymbol{x}-\boldsymbol{x}'|<1$

IIb $|\boldsymbol{x}-\boldsymbol{x}'|>1$

First we note that, from lemma 1 and Taylor's theorem with differential form of the remainder, we have that, $\forall|\boldsymbol{x}|,|\boldsymbol{x}'|>2\rho$, $\exists C$ independent of $|\boldsymbol{x}|$ such that

$$|\boldsymbol{g}(\boldsymbol{x})-\boldsymbol{g}(\boldsymbol{x}')|<\frac{C}{|\boldsymbol{x}|^4}|\boldsymbol{x}-\boldsymbol{x}'|. \tag{2.203}$$

Thus, we write (IIa) as

$$\int\limits_{|\boldsymbol{x}-\boldsymbol{x}'|<1}\frac{(\boldsymbol{g}(\boldsymbol{x}')-\boldsymbol{g}(\boldsymbol{x})).(\boldsymbol{x}-\boldsymbol{x}')}{|\boldsymbol{x}-\boldsymbol{x}'|^2}d^2\boldsymbol{x}'+\boldsymbol{g}(\boldsymbol{x})\cdot\int\limits_{|\boldsymbol{x}-\boldsymbol{x}'|<1}\frac{\boldsymbol{x}-\boldsymbol{x}'}{|\boldsymbol{x}-\boldsymbol{x}'|^2}d^2\boldsymbol{x}'. \tag{2.204}$$

Now,

$$\left|\int\limits_{|\boldsymbol{x}-\boldsymbol{x}'|<1}\frac{(\boldsymbol{g}(\boldsymbol{x}')-\boldsymbol{g}(\boldsymbol{x})).(\boldsymbol{x}-\boldsymbol{x}')}{|\boldsymbol{x}-\boldsymbol{x}'|^2}d^2\boldsymbol{x}'\right|<\frac{C}{|\boldsymbol{x}|^4}\int_0^{2\pi}\int_0^1 r\,dr\,d\theta=\pi C/|\boldsymbol{x}|^4, \tag{2.205}$$

whereas

$$\boldsymbol{g}(\boldsymbol{x}).\int\limits_{|\boldsymbol{x}-\boldsymbol{x}'|<1}\frac{\boldsymbol{x}-\boldsymbol{x}'}{|\boldsymbol{x}-\boldsymbol{x}'|^2}d^2\boldsymbol{x}'=|\boldsymbol{g}(\boldsymbol{x})|\int_0^{2\pi}\int_0^1\frac{1}{r}r\,dr\,d\theta=0. \tag{2.206}$$

In (IIb) we have

$$\left|\int_{IIb}\frac{\boldsymbol{g}(\boldsymbol{x}').(\boldsymbol{x}-\boldsymbol{x}')}{|\boldsymbol{x}-\boldsymbol{x}'|^2}d^2\boldsymbol{x}'\right|<8\pi A/|\boldsymbol{x}|^2, \tag{2.207}$$

since $|\boldsymbol{x}-\boldsymbol{x}'|^{-1}<1$ in IIb.

2.6.3 Region III

In III, we consider

$$\left|\int\limits_{|\boldsymbol{x}'|>|\boldsymbol{x}|+1}\boldsymbol{g}(\boldsymbol{x}')\cdot\frac{\boldsymbol{x}-\boldsymbol{x}'}{|\boldsymbol{x}-\boldsymbol{x}'|^2}d^2\boldsymbol{x}'\right| \tag{2.208}$$

$$<\ 2A\int\limits_{|\boldsymbol{x}'|>|\boldsymbol{x}|+1}|\boldsymbol{x}'|^{-3}\left(\frac{1}{|\boldsymbol{x}'|-|\boldsymbol{x}|}\right)d^2\boldsymbol{x}' \tag{2.209}$$

$$= 4\pi A \int_{|\boldsymbol{x}+1|}^{\infty} \frac{1}{r^2}\frac{1}{r-|\boldsymbol{x}|}dr \tag{2.210}$$

$$= 4\pi A \left(\frac{1}{|\boldsymbol{x}|^2}\ln(|\boldsymbol{x}|+1) - \frac{1}{|\boldsymbol{x}|(|\boldsymbol{x}|+1)}\right) \tag{2.211}$$

$$\sim O\left(\frac{\ln|\boldsymbol{x}|}{|\boldsymbol{x}|^2}\right). \tag{2.212}$$

Collecting (2.194, 2.198, 2.202, 2.205, 2.206, 2.207, 2.212) together we arrive at

$$\int_{\mathbb{R}^2} \nabla.((h_0 - f\psi_0)\boldsymbol{u}_0)\ln|\boldsymbol{x}-\boldsymbol{x}'|d^2\boldsymbol{x}' \tag{2.213}$$

$$\sim \frac{\boldsymbol{x}}{|\boldsymbol{x}|^2}.\int_{\mathbb{R}^2}(h_0 - f\psi_0)\boldsymbol{u}_0(\boldsymbol{x}')d^2\boldsymbol{x}' + O\left(\frac{\ln|\boldsymbol{x}|}{|\boldsymbol{x}|^2}\right) \qquad \square \tag{2.214}$$

2.7 Appendix B

We define the Fourier Transform $\tilde{f}(\omega)$ if the function $f(t)$ by

$$\tilde{f}(\omega) = \int_{-\infty}^{\infty} f(t)e^{i\omega t}dt, \tag{2.215}$$

as in equation (2.87). From Lighthill (1958), the Fourier Transform of $\ln|x|H(x)$ is

$$\ln|t|\tilde{H}(t) = -\frac{1}{i\omega}\left(\frac{i\pi}{2}\mathrm{sgn}\omega + \ln|\omega|\right) + \frac{\gamma}{i\omega}. \tag{2.216}$$

Using shifting formulae, it follows that the Fourier Transform of $e^{ift}\ln|t|H(t)$ is

$$e^{ift}\ln|t|\tilde{H}(t) = -\frac{1}{i(\omega-f)}\left(\frac{i\pi}{2}\mathrm{sgn}(\omega-f) + \ln|\omega-f|\right) + \frac{\gamma}{i(\omega-f)}, \tag{2.217}$$

and hence the Fourier transform of the integral

$$\int_0^{\infty} \ln\tau e^{if\tau}\left(g'(t-\tau) - ifg(t-\tau)\right) \tag{2.218}$$

is

$$-\tilde{g}(\omega)\left(\frac{i\pi}{2}\mathrm{sgn}(\omega-f) + \ln|\omega-f| + \gamma\right). \tag{2.219}$$

and hence the Fourier transform of

$$2\int_0^{\infty} \ln\tau\frac{d}{d\tau}\left[\cos f\tau g(t-\tau)\right] \tag{2.220}$$

is

$$-\tilde{g}(\omega)\left[\frac{i\pi}{2}\left(\mathrm{sgn}(\omega-f) + \mathrm{sgn}(\omega+f)\right) + \ln|\omega^2-f^2| + 2\gamma\right]. \tag{2.221}$$

Now,

$$\ln|\omega^2 - f^2| + \mathrm{sgn}(\omega - f) + \mathrm{sgn}(\omega + f) = \ln(\omega^2 - f^2) \mp i\pi, \tag{2.222}$$

where the sign of $\mp$ is (2.222) is set by the choice of analysic branch of the logathm function. Therefore, with the appropriate choice of the branch of the logaritm,

$$-(\omega^2 - f^2)^2\tilde{\sigma}(\omega)\left(\frac{1}{8}\left(\ln(\omega^2 - f^2) + 2\gamma\right) \mp \frac{i\pi}{8}\right) \tag{2.223}$$

has an inverse Fourier transform

$$\int_0^\infty \ln\tau \frac{d}{d\tau}(\cos f\tau g(t-\tau)), \tag{2.224}$$

where

$$\tilde{g}(\omega) = \frac{1}{16}(\omega^2 - f^2)^2\tilde{\sigma}(\omega), \tag{2.225}$$

and where (2.224) is an integral over the entire past history of the dynamics. If the other analytic branch of the logarithm is taken in (2.222), the result is that the integral (2.224) becomes an integral over the entire future evolution of the dynamics. The radiation condition is equivalent to imposing the choice of branch of logarithm which gives rise to (2.224).

Chapter 3

Instability of vortices and jets

3.1 Introduction

In chapter 2, some general properties of gravity wave generation by vortical motions were derived in the asymptotic limit of low Froude number F. It was shown that, in the vortical region, the first corrections to the incompressible vortical dynamics which are required to take account of non-zero divergence occur at $O(F^2)$ and $O(F^2 \ln F)$. On the basis of this elementary scale analysis, one might therefore expect gravity wave activity to be observed in the vortical flow at $O(F^2)$. However, the $O(F^2)$ and $O(F^2 \ln F)$ corrections can be determined entirely in terms of the instantaneous vortical flow, and thus represent balanced corrections to the leading order dynamics. The effect of gravity wave radiation is felt in the vortical region at $O(F^4)$, two orders higher in the expansion parameter than might have been expected from elementary scale analysis. The particular flow of the Kirchoff ellipse of uniform potential vorticity was analysed. The ellipse experienced elongation as a result of gravity wave radiation. The elongation either ceased at the aspect ratio at which the radiation from the ellipse was reduced to the inertial frequency, or continued until the ellipse reached a "splitting" aspect ratio of 4.6:1, at which the ellipse split in two. The elongation of the ellipse represents the effect of gravity wave radiation in the vortical flow which generates it, and the elongation rate of the ellipse was therefore $O(F^4)$. If we suppose now that the eccentricity of the ellipse is small, we may regard this elongation as the growth of a mode of instability on an axisymmetric vortex, with an eigenfunction of the form $f(r)e^{i(2\theta-\omega t)}$, where θ is the azimuthal angle, and ω is the eigenfrequency. Here, ω is real at $O(1)$ and $O(F^2)$, but contains an imaginary part at $O(F^4)$. The eigenmode consists of a vortical or Rossby wave on the boundary of the vortex, coupled to a gravity wave far from the vortex. The growth rate of the instability is a measure of the degree of interaction between vortical motions and gravity waves in the system.

The concept of instabilities due to coupling between different waves has been the subject of a paper by Sakai (1989). In the most general form of the theory, he advocates decomposing the system of eigenvector equations to be studied into distinct wave vectors, which are neutral non-interacting modes of some reference system, but which interact in the system of interest due to the presence of a mean flow, second layer or other destabilizing phenomenon. Once the decomposition into a system of wave functions is accomplished, the resulting matrix eigenvalue analysis for the system of interest is lengthy but routine. The structure of the unstable eigenmodes in terms of the modes of the reference system elucidates which wave – wave inter-

actions are responsible for the instability. However, if the eigenmodes can be analysed in some asymptotic limit, as here in the case of low Froude number or high mode number, it is possible to elucidate the nature of the wave mode interactions which are responsible for the instability without recourse to lengthy matrix eigenvalue analysis. This asymptotic approach has been used by Knessel & Keller (1992) in their analysis of a short wavelength instability found by Griffiths *et al.* (1982) which is due to gravity wave – gravity wave interaction.

Within this wave mode instability framework, conventional barotropic instability and baroclinic instability are to be regarded as due to Rossby wave – Rossby wave interactions. They both arise due to opposite signs of the potential vorticity gradient in the flow, either at different horizontal locations (in the case of barotropic instability) or on different vertical levels (in the case of baroclinic instability).

Gravity wave – gravity wave instabilities are also fairly common, and have been found in single layer shear flows with uniform potential vorticity (Satomura 1981; Kubokawa 1986; Hayashi & Young 1987). Kelvin – Helmholtz instability is also a manifestation of gravity wave – gravity wave instability. The instability exists due to gravity waves which develop on either side of a density jump in stratified shear flow, just as Rossby waves develop on either side of a strip of potential vorticity in barotropic instability.

Rossby wave – gravity wave instability, however, is a comparatively unfamiliar phenomenon. Sakai (1989) has clearly demonstrated a Rossby wave – gravity wave instability in a two-layer parallel shear flow at moderate Rossby number. He has also suggested the several previously discovered instabilities are candidates for a description in terms of Rossby wave –gravity wave interaction. In many of the studies he discusses, an interface between two fluids of different density intersects a free surface, and the "gravity wave" part of the Rossby wave – gravity wave instability is supposed to be the wave at the free surface intersection.

Within the context of barotropic shallow water vortex dynamics, which is the topic of this thesis, the model problem consists of a semi-infinite layer of light fluid, of finite depth, on top of an infinite layer of heavy fluid of infinite depth (see figure 3.1). Paldor (1983) has shown that, if the potential vorticity in the upper layer is uniform everywhere where it has non-zero depth, the eigenmodes of the system are always neutral. Killworth & Stern (1982) have shown that an instability will always be present when the potential vorticity in the upper layer increases away from the interface. I believe that this wave is more properly a generalization of a Rossby

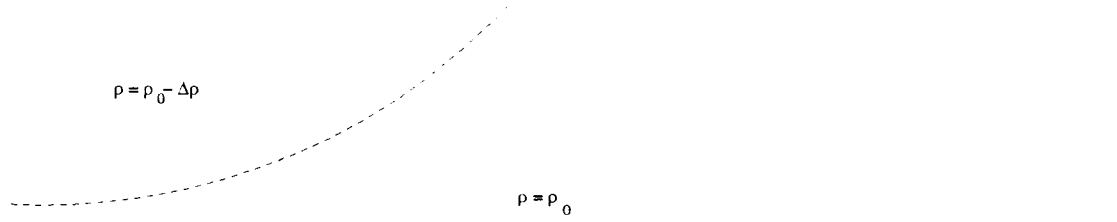


Figure 3.1: Schematic diagram depicting a semi-infinite single layer of fluid.

wave in the upper layer, with finite potential vorticity on one side of the interface, and infinite potential vorticity (zero depth) on the other side. Consider first the case where the potential vorticity is finite on both sides of the interface. In such circumstances, a wave on the boundary between the two regions would clearly be regarded as a Rossby wave. Then, since at all finite values of potential vorticity we would regard the wave on the interface between the two regions of different potential vorticity as a Rossby wave, it seems perverse to suggest, as Sakai (1989) does, that we should regard it as a gravity wave in the singular case of infinite potential vorticity on one side of the interface. The condition for instability found by Killworth & Stern (1982), that the potential vorticity increase away from the interface, is simply a condition which ensures that the interface is a minimum of potential vorticity, and that there is therefore a change in the potential vorticity gradient in the upper layer at the interface.

The shallow water vortex instability, discussed partially in chapter 2 in terms of the elongation of the elliptical vortex, is a clear case of Rossby wave – gravity wave instability in a single layer model. It was discovered first in the non-rotating two dimensional ideal gas equations by Broadbent & Moore (1979). They consider a vortex of radius a, with uniform vorticity 2Ω for $r < a$, and zero vorticity for $r > a$. Such a vortex is commonly referred to as a "Rankine vortex". They define a Mach number M based on the vortex boundary velocity, so that $M = \Omega a/c_0(a)$, where c_0 is the sound speed $c_0^2(r) = \gamma p_0(r)/\rho_0(r)$. They consider eigenfunctions of the form $f(r)\epsilon^{i(m\theta-\omega t)}$, where m represents the azimuthal mode number of the disturbance. They show numerically that, as M is increased away from zero, the growth rates of instabilities are non-zero for mode numbers $m = 2, 3, 4, 6, 8$. For each m, they investigate only one eigenmode, which reduces to the neutral mode found by Kelvin (1880) in the limit $M \to 0$, with eigenfrequency $\omega = \Omega - 1$. At small M the growth rates scale as M^{2m} (Kop'ev & Leont'ev, 1983), meaning that large mode numbers are less unstable than small ones for small M. Indeed, it can be clearly seen from figure 3 of Broadbent & Moore (1979) that, for moderate Mach number M, growth

rates of the instability remain smaller for larger mode numbers than smaller ones.

For $M < 2.236$, modes $m = 2, 3, 4, 6, 8$ are found to be unstable for all M. The maximum growth rates obtained by Broadbent & Moore for a given M were always found at mode $m = 2$. The overall maximum growth rate occurred at $M \approx 1.5$, with $\mathrm{Im}(\omega) \approx 0.012$. The growth rates remain small over the whole range of values of $M < 2.236$, suggesting that the strength of Rossby wave – gravity wave coupling might be small even at quite large Froude numbers. When $M = 2.236$ Broadbent & Moore are obliged to terminate their numerical eigenvalue calculations, since then the density vanishes at the centre of the vortex.

Sozou (1987) showed that the mode investigated by Broadbent & Moore (1979) is not the only unstable mode in the compressible Rankine vortex, and that there are an infinite number of modes for which $\omega = \Omega - O(M^2)$. These modes are singular in the limit $M = 0$, but non-singular for all non-zero M. They typically have growth rates at least an order of magnitude smaller than the corresponding Broadbent – Moore mode (Sozou 1987, table 1).

A general stability theorem for axisymmetric shallow water vortices was given by Ripa (1987). It was derived by variational methods, and therefore provides a sufficient condition for stability. The condition is that the axisymmetric flow with azimuthal velocity $V(r)$ and height field $H(r)$ is stable if

$$(V - \lambda r)^2 < H(r) \tag{3.1}$$

$$(V - \lambda r)\frac{dQ}{dr} \geq 0 \tag{3.2}$$

$$\text{for some } \lambda, \tag{3.3}$$

where $Q = (f + dV/dr + V/r)/H(r)$ is the potential vorticity in the basic state. In the case of an isolated vortex in unbounded shallow water, (3.1) can only be satisfied if $\lambda = 0$. Then, if the vortex is cyclonic, with the potential vorticity inside the vortex exceeding the value outside, we would expect the azimuthal velocity $V(r)$ at the boundary of the vortex to be anticlockwise, i.e. positive. Consequently, VdQ/dr will be zero everywhere except at the boundary of the vortex, where it will be negative. Similarly, if the vortex is anticyclonic, with $dQ/dr > 0$, we expect $V(r) < 0$ at the vortex boundary. In either case, $V(r)dQ/dr < 0$ at the vortex boundary, and (3.2) is not satisfied. Therefore it is not possible to obtain any stability results for an axisymmetric vortex with a monotonic potential vorticity profile by variational methods, no matter how small the Froude and Rossby numbers might be.

A general stability theorem for parallel flow stability in shallow water was given by Ripa (1983). In this case, the sufficient condition for stability of the flow with jet velocity $U(y)$ and layer depth $H(y)$ is

$$(\lambda - U(y))\frac{dQ}{dy} \geq 0 \tag{3.4}$$

$$(\lambda - U(y))^2 < H(y) \tag{3.5}$$

$$\text{for some } \lambda. \tag{3.6}$$

Now, if the potential vorticity gradient is concentrated in a single jump, at $y = 0$, say, then $U(y)dQ/dy < 0$ at $y = 0$, and we may take $\lambda = U(0)$ to satisfy (3.4). The remaining condition (3.5) then states that the flow must be everywhere subsonic with respect to the velocity of the potential vorticity jump $(U(0) - U(y))^2 < H(y)$. Unlike the case of axisymmetric flow, it is now possible to show that the flow is stable provided the ratio of potential vorticities across the jump is less than 4.

One of the aims of this thesis is to investigate the strength of Rossby wave – gravity interactions in the shallow water system. The growth rates of Rossby wave – gravity wave instabilities in axisymmetric and parallel flows are therefore of fundamental interest, and the remainder of the chapter is organised as follows:

In §3.2, the instability of an axisymmetric vortex with a single discontinuity in potential vorticity is analysed. The results at finite Froude number will differ from those of Broadbent & Moore (1979) at finite Mach number, in that in this study the potential vorticity, rather than vorticity, is taken to be constant within the vortex. This prevents the evacuation of the basic state found by Broadbent & Moore. As the background rotation rate f is increased, the low mode number disturbances cease to be unstable. A WKBJ analysis is therefore performed, and it is found that large mode number disturbances to an axisymmetric vortex are always unstable, no matter how small the Froude and Rossby numbers might be.

In §3.3 we investigate the effect of smoothing the potential vorticity profile within the context of the WKBJ analysis. It is found that the instability persists, provided the potential vorticity gradient is confined to within a region of size $o(m^{-1})$, so that Rossby wave critical layers at $\omega = mV/r$ induce phase shifts which are no more than exponentially weak in m. The Sozou modes are shown to exist due to the non-uniform potential vorticity within the Rankine vortex at non-zero Mach number, which therefore supports a spectrum of Rossby waves within

the vortex in addition to the single Rossby wave on the vortex boundary.

In §3.4, we discuss the results of the analysis in §3.2 in relation to the question of existence or non-existence of a slow manifold for the shallow water equations. Since the WKBJ analysis has shown that a Rossby wave – gravity wave instability exists at arbitrarily small Rossby number, the result is that a true slow manifold, in the sense of no identifiable gravity wave – like structure at small non-zero Rossby number, cannot exist, because the large mode number eigenmodes will always possess a gravity wave - like tail. Mathematically, one can show that a slow manifold, if it exists at all, cannot then be unique, precisely because of the presence of the instability at arbitrarily small Rossby number.

In §3.5 the instability of a parallel flow with a single discontinuity of potential vorticity is investigated. The potential vorticity in the region $y < 0$ is taken to be unity, while the potential vorticity in the region $y > 0$ takes a value $q > 1$. It is found via a WKBJ analysis that the stability bound found by Ripa (1983) is an exact stability boundary, with unstable modes for all $q > 4$. The limit $q \to \infty$ is singular, and if q is rescaled such that $q = O(k^{4/3})$, the growth rate of the instability tends to zero, as $q/k^{4/3} \to \infty$, consistent with the analysis of Paldor (1983), and supporting the contention that the interfacial wave should be viewed as a Rossby wave, rather than a gravity wave.

In §3.6, some final conclusions to this chapter are offered.

3.2 Instability of a circular vortex with discontinuous potential vorticity

3.2.1 The basic state

In this section, we shall restrict our attention to basic states with a single radial discontinuity of potential vorticity. Following the same scaling of the shallow water equations as in chapter 2, but not now assuming $F \ll 1$, the basic state equations are

$$\frac{d\overline{h}}{dr} = f\overline{v} + \frac{\overline{v}^2}{r} \tag{3.7}$$

$$\frac{d\overline{v}}{dr} + \frac{\overline{v}}{r} = Q(1 + F^2\overline{h}) - f. \tag{3.8}$$

These must be solved subject to regularity conditions at $r = 0$, and decay conditions as $r \to \infty$. The nonlinearity in (3.7) means that we must use numerical means to solve (3.7,3.8), except in

the limit $F \ll 1$, where they can be solved by matched asymptotic expansion.

The free parameters in the problem are f and F. The nondimensionalizations of chapter 2 allow us to take the vortex to be of unit radius, the potential vorticity jump across the boundary of the vortex to be of unit strength, and the layer to be of unit mean depth. Then $Q = f$ for $r > 1$, and $Q = f + 1$ for $r < 1$.

3.2.2 The disturbance equations

Disturbance equations are derived for a single azimuthal wavenumber m and frequency ω, so that the disturbance variables are expressible in the form $g(r)e^{im\theta - i\omega t}$. They are

$$i\sigma v_r - (f + 2\overline{v}/r)v_\theta + \frac{dh}{dr} = 0 \tag{3.9}$$

$$i\sigma v_\theta + Q(1 + F^2\overline{h})v_r + \frac{1}{r}imh = 0 \tag{3.10}$$

$$F^2\left(i\sigma h + (f\overline{v} + \overline{v}^2/r)v_r\right) + (1 + F^2\overline{h})\left(\frac{dv_r}{dr} + \frac{v_r}{r} + \frac{1}{r}imv_\theta\right) = 0, \tag{3.11}$$

where $\sigma = -\omega + m\overline{v}/r$.

This set of equations must be solved subject to regularity conditions at $r = 0$, and a radiation/evanescence condition as $r \to \infty$.

The equations (3.9 - 3.11) contain only two radial derivatives, and therefore represent a second order differential equation for the eigenfunctions. If one chooses v_r and h as variables, v_θ must be obtained from (3.10). The possibility of a critical layer at $\sigma = 0$ makes it impossible to ensure that this will be a successful method. Instead, we choose to eliminate h. After some algebra, one obtains a pair of first order ordinary differential equations for the eigenmodes:

$$\frac{dv_r}{dr} = A(r)v_r + B(r)v_\theta \tag{3.12}$$

$$\frac{dv_\theta}{dr} = C(r)v_r + D(r)v_\theta, \tag{3.13}$$

where

$$A(r) = \frac{F^2\sigma Q r}{m} - \frac{1}{r}\frac{d}{dr}\ln H \tag{3.14}$$

$$B(r) = \frac{iF^2\sigma^2 r}{Hm} - \frac{im}{r} \tag{3.15}$$

$$C(r) = \frac{iF^2 HQ^2 r}{m} + \frac{im}{r} + \frac{iH}{\sigma}\frac{dQ}{dr} \tag{3.16}$$

$$D(r) = -\frac{F^2 Q\sigma r}{m} - \frac{1}{r}, \tag{3.17}$$

and $H = 1 + F^2\overline{h}$.

The advantage of this formulation over the standard formulation (3.9 - 3.11) is that the critical layer singularity $\sigma = 0$ occurs only where there are radial gradients of the basic state potential vorticity $Q(r)$. If the potential vorticity gradients are confined to a single discontinuity in potential vorticity at the vortex boundary, the eigenfunction equations are non-singular, and at the vortex boundary we impose continuity of h and v_r. Dividing (3.10) by v_r, we thereby obtain a continuity equation

$$\left[\sigma \frac{v_\theta}{v_r}\right] = \left[iQ(1 + F^2\overline{h})\right], \tag{3.18}$$

where $[\alpha]$ represents the jump in a quantity α across the vortex boundary.

3.2.3 The limit $F \ll 1$

Before proceeding to numerical calculation of the eigenvalues of the system (3.12 - 3.17), we consider the limit $F \ll 1$.

In chapter 2 we showed that matching conditions between the vortical region and the gravity wave region led to a wave field in which the principal wave component was a quadrupole. This was forced by unsteady terms of the form $r^{-2}e^{2i\theta}$ present in the streamfunction in the limit of $r \to \infty$. In the case of a mode 2 perturbation to an axisymmetric vortex, these terms appear due to the unsteadiness in velocity and height fields induced by the perturbation of the vortex boundary. The effect of the radiation appears at $O(F^4)$, the effect of which is to destabilize the vortex.

From the viewpoint of an instability, we therefore have growth of a mode 2 perturbation to an axisymmetric vortex, with a growth rate of order F^4. Higher order perturbations to the vortex do not produce quadrupole far fields, since they have m-fold rotational symmetry, and their streamfunction far fields take the form $r^{-m}e^{im\theta}$. Rescaling the radial coordinate to $R = Fr$, appropriate for the wave region, we have wave fields in streamfunction, velocity potential and height of form $H_m(\sqrt{\omega^2 - f^2}R)e^{i(m\theta - \omega t)}$ at order F^m, and therefore the first out of phase contribution in the matching conditions for the inner problem occurs at order $O(F^{2m})$. Hence in general at low Froude number the vortex has mode m instabilities with growth rates $O(F^{2m})$.

However, we must also consider the real part of the frequency, which must lie above the

inertial frequency $|f|$ if waves are to propagate and the out of phase terms are to appear in the matching conditions. The analysis for the circular vortex at $F = 0$ was first performed by Kelvin (1880), who found that

$$\omega = \frac{1}{2}(m-1). \tag{3.19}$$

It follows that we shall require $m > 1+2|f|$ for instability at low F. The fastest growing $m = 2$ modes are inhibited by the presence of background rotation, and for small to moderate Rossby numbers the first unstable mode could occur at quite a large value of m, and hence quite a small growth rate of order F^{2m}.

The basic state

We begin our analysis by establishing the form of the basic state from considering the basic state equations (3.7-3.8) in the limit $F \ll 1$. From the analysis of chapter 2, we require two asymptotic regions – $r = O(1)$ and $r = O(F^{-1})$.

We begin with the region $r = O(1)$. We expand

$$\overline{v} = \overline{v}_0 + F^2\overline{v}_2 + \cdots \tag{3.20}$$

$$\overline{h} = \overline{h}_0 + F^2\overline{h}_2 + \cdots. \tag{3.21}$$

Here we allow for the possibility that the matching conditions might introduce $\ln F$ terms into the expansions (3.20,3.21). Substituting (3.20,3.21) into (3.8) gives, at leading order

$$\frac{d\overline{v}_0}{dr} + \frac{\overline{v}_0}{r} = \begin{cases} 1 & \text{for } r < 1 \\ 0 & \text{for } r > 1 \end{cases}. \tag{3.22}$$

Imposing regularity ar $r = 0$, and r^{-1} decay ar $r \to \infty$, we have

$$\overline{v}_0 = \begin{cases} \frac{1}{2}r & \text{for } r < 1 \\ \frac{1}{2}r^{-1} & \text{for } r > 1 \end{cases}. \tag{3.23}$$

The equation for $\overline{h}_0$ is

$$\frac{d\overline{h}_0}{dr} = f\overline{v}_0 + \frac{\overline{v}_0^2}{r}. \tag{3.24}$$

Substituting (3.23) into (3.24) we find

$$\overline{h}_0 = \begin{cases} \frac{1}{8}(2f+1)r^2 + C_1 & \text{for } r < 1 \\ \frac{1}{2}f\ln r - \frac{1}{8}r^{-2} + C_2 & \text{for } r > 1 \end{cases}, \tag{3.25}$$

where C_1 and C_2 are constants, to be determined by continuity of $\overline{h}_0$ at $r = 1$, and matching to the outer solution.

Proceeding now to the outer region, we define a long range variable $R = Fr$. Since $\overline{v}_0 \sim r^{-1}$ as $r \to \infty$, we rescale $\overline{v}_0$ by F, so

$$\overline{h}_0 = H(R) + O(F^2) \tag{3.26}$$

$$\overline{v}_0 = FV(R) + O(F^3). \tag{3.27}$$

Then substituting (3.26, 3.27) into (3.7,3.8) we obtain

$$\frac{dV}{dR} + \frac{V}{R} = fH \tag{3.28}$$

$$\frac{dH}{dR} = fV. \tag{3.29}$$

Combining (3.28) & (3.29) into a single equation for $V(R)$, we have

$$\frac{d^2V}{dR^2} + \frac{1}{R}\frac{dV}{dR} - \left(f^2 + \frac{1}{R^2}\right)V = 0. \tag{3.30}$$

Imposing decaying boundary conditions as $R \to \infty$, (3.30) has solution

$$V(R) = A\mathrm{K}_1(|f|R), \tag{3.31}$$

where K_1 is the modified Bessel function of order one, and A is a constant. Matching to the velocity field in the vortical region (3.23) gives

$$A = \frac{1}{2}|f|. \tag{3.32}$$

It follows that

$$H(R) = -\frac{1}{2}f\mathrm{K}_0(|f|R). \tag{3.33}$$

For small R, we have

$$H(r) = \frac{1}{2}f\left(\ln\left(\frac{1}{2}|f|R\right) + \gamma\right) + O(R^2 \ln R). \tag{3.34}$$

and matching to (3.25) gives

$$C_2 = \frac{1}{2}f\left(\ln\left(\frac{1}{2}|f|F\right) + \gamma\right). \tag{3.35}$$

Note the appearance of the $\ln F$ term. This means that $\overline{h}_0$ starts at $O(\ln F)$. Then finally

$$C_1 = \frac{1}{2}f\left(\ln\left(\frac{1}{2}|f|F\right) + \gamma\right) - \frac{1}{4}(f+1). \tag{3.36}$$

The analysis presented is valid in the limit $f \to 0$, in which limit

$$V(R) = \frac{1}{2R}; \quad H(R) = O(F^2); \quad C_1 = -\frac{1}{4}; \quad C_2 = 0. \tag{3.37}$$

Having obtained the leading order solutions for $\overline{v}_0$ and $\overline{h}_0$ in both asymptotic regions, we could now go on to develop perturbation expansions to obtain equations for $\overline{v}_2, \overline{h}_2 \ldots$. However, as we shall see, such corrections are not required to compute the growth rate of the instability, and serve only to correct the real part of the eigenfrequency. Indeed, the expression for $\overline{h}_0$ is not really essential to the $F \ll 1$ analysis, although it is required to initialize the numerical eigenvalue calculations in the $F \ll 1$ limit.

The disturbance equations

As for the analysis of the basic state, we expand v_r and v_θ in asymptotic series with F^2 as the expansion parameter. The asymptotic expansions are matched between two regions, with range variables r and $R \equiv Fr$ as before (see also chapter 2). Additionally, the eigenvalue ω must be expanded in powers of F^2, with logarithmic terms as necessary.

Substituting (3.14 - 3.17) into (3.12,3.13) in the limit $F \ll 1$ gives equations for v_r and v_θ

$$\frac{dv_r}{dr} = -\frac{1}{r}v_r - \frac{im}{r}v_\theta \tag{3.38}$$

$$\frac{dv_\theta}{dr} = \frac{im}{r}v_r - \frac{1}{r}v_\theta. \tag{3.39}$$

The solutions are

$$v_r = r^{m-1}; \qquad v_\theta = ir^{m-1} \tag{3.40}$$

or

$$v_r = r^{-m-1}; \qquad v_\theta = -ir^{-m-1}. \tag{3.41}$$

Regularity as $r \to 0$ implies we take an eigenfunction of form (3.40) in $r < 1$. Decay as $r \to \infty$ implies we take an eigenfunction of form (3.41) in $r > 1$. Then, substituting into the continuity equation (3.18) gives $\omega = m\overline{v}(1) - 1/2$, which gives Kelvin's result (3.19) on substitution of the basic state (3.23), for which $\overline{v}(1) = 1/2$, valid in the limit $F = 0$.

Rather than proceeding to consider higher orders in F in the inner expansion, we will

consider the leading order outer expansion. There the disturbance equations are

$$\frac{dv_r}{dR} = -\left(\frac{\omega f R}{m} + \frac{1}{R}\right) v_r + \left(\frac{i\omega^2 R}{m} - \frac{im}{R}\right) v_\theta \tag{3.42}$$

$$\frac{dv_\theta}{dR} = \left(\frac{if^2 R}{m} + \frac{im}{R}\right) v_r + \left(\frac{f\omega R}{m} - \frac{1}{R}\right) v_\theta. \tag{3.43}$$

They have solution

$$v_r = A\left[i\omega\lambda \mathrm{H}_{m-1}(\lambda R) - i\frac{m}{R}(\omega + f)\mathrm{H}_m(\lambda R)\right] \tag{3.44}$$

$$v_\theta = A\left[f\lambda \mathrm{H}_{m-1}(\lambda R) - \frac{m}{R}(\omega + f)\mathrm{H}_m(\lambda R)\right], \tag{3.45}$$

where $\lambda^2 = (\omega^2 - f^2)$, H_m is a Hankel function of degree m, of the first or second kind, and A is an arbitrary constant. The kind of the Hankel function is chosen to give radiation conditions at infinity. With $e^{-i\omega t}$ time dependence, one chooses Hankel functions of the first kind for radiation conditions.

Now the outer limit of the inner expansions for v_r and v_θ have r^{-m-1} behaviour. Therefore in the outer region we start our expansions for v_r and v_θ at $O(F^{m+1})$. This is consistent with waves in ψ, ϕ and h at $O(F^m)$. If a formal expansion is carried out for the inner region, we see from (3.14 - 3.17) that it proceeds by first obtaining the fields at a given order, and then obtaining ω at that order from the matching conditions. To obtain v_r and v_θ to order F^{2n} for some integer n, one need only know ω to order $F^{2(n-1)}$. The continuity condition (3.18) then gives ω at order F^{2n}.

Since the eigenvalue equations are linear in v_r and v_θ, the amplitude of the leading order solution given by (3.40) for $r < 1$ and (3.41) for $r > 1$ is not specified uniquely, and at higher orders in the perturbation expansion it is possible to add to the solution at that order a component of the leading order eigenfunction. To specify the eigenfunction uniquely, we impose the condition that the coefficients of r^{m-1} in the expansions of v_r and v_θ in the limit $r \to 0$ are zero at all orders in F except the leading order. This then fixes the amplitude and phase of the eigenfunction at all orders in F. Then, without reference to the outer expansion, one can see that the velocity components v_r and v_θ remain $\pi/2$ out of phase with each other at all subsequent orders in F in the inner expansion, and therefore ω remains real until the matching conditions with the wave region must be taken into account.

Matching (3.41) onto (3.44,3.45) to determine the amplitude of the gravity waves in the outer expansion gives

$$A = -\frac{\pi\lambda^m}{2^m m!(\omega + f)}. \tag{3.46}$$

Then matching the out-of-phase J_m terms in the Hankel functions (3.44, 3.45) back onto the inner solution at order $O(F^{2m})$ gives a contribution to the inner solution of the form

$$v_r = -iBr^{m-1}; \qquad v_\theta = Br^{m-1}, \tag{3.47}$$

where B is determined from A and the properties of the Hankel function in the limit $R \to 0$. It is given by

$$B = \frac{\pi m}{(m!)^2}\left(\frac{\omega^2 - f^2}{4}\right)^m \left(\frac{\omega - f}{\omega + f}\right). \tag{3.48}$$

As we discussed above, to make the expansion procedure well defined, the expansions for v_r and v_θ in $r < 1$ are to be $o(r^{m-1})$ for $r \ll 1$, and hence we must add to (3.47) functions of form (3.41) to ensure continuity of v_r at $r = 1$. All other contributions to v_r and v_θ at all orders up to and including F^{2m} contribute only to the real part of ω, and not to the growth of the instability.

To calculate the growth rate of the instability, it is therefore sufficient to consider the velocity fields

$$v_r = \begin{cases} r^{m-1} & \text{for } r < 1 \\ r^{-m-1} - iBF^{2m}(r^{m-1} - r^{-m-1}) & \text{for } r > 1 \end{cases} \tag{3.49}$$

$$v_\theta = \begin{cases} ir^{m-1} & \text{for } r < 1 \\ -ir^{-m-1} + BF^{2m}(r^{m-1} + r^{-m-1}) & \text{for } r > 1 \end{cases}. \tag{3.50}$$

Returning to the jump condition (3.18), we obtain

$$\mathrm{Im}(\omega) = \frac{\pi m F^{2m}}{2(m!)^2}\left(\frac{\omega^2 - f^2}{4}\right)^m \left(\frac{\omega - f}{\omega + f}\right) + O(F^{2m+2}). \tag{3.51}$$

3.2.4 Numerical eigenvalue calculations

Numerical technique

We start by considering techniques for solving numerically the basic equations (3.7, 3.8) for the basic state variables $\overline{v}$ and $\overline{h}$. The numerical difficulties arise in that both $r = 0$ and $r = \infty$ are singular points of the system, and we can only apply boundary conditions there by making use of some expansion of the solutions for small r and large r respectively. Given initial guesses for $\overline{h}(0)$ and $\overline{h}(\infty)$, the basic state equations can then be integrated to $r = 1$. and iteration on $\overline{h}(0)$ and $\overline{h}(\infty)$ used to make $\overline{h}$ and $\overline{v}$ continuous across the vortex boundary $r = 1$.

We consider first the singular point $r = \infty$. In this limit, we will impose conditions that $\overline{v}$ and $\overline{h}$ tend to zero as $r \to \infty$. By manipulation of the equations (3.7, 3.8), one obtains

$$\frac{d^2\overline{h}}{dr^2} + \frac{1}{r}\frac{d\overline{h}}{dr} - F^2 f^2 \overline{h} = \frac{2\overline{v}}{r}\frac{d\overline{v}}{dr} \tag{3.52}$$

$$\frac{d^2\overline{v}}{dr^2} + \frac{1}{r}\frac{d\overline{v}}{dr} - \left(F^2 f^2 + \frac{1}{r^2}\right)\overline{v} = F^2 f \frac{\overline{v}^2}{r}. \tag{3.53}$$

Imposing evanescence conditions as $r \to \infty$, we note that the nonlinear right hand sides of (3.52) and (3.53) will be of small order compared with $\overline{v}$ and $\overline{h}$ in the limit $r \to \infty$. A convenient form of solution for $r \to \infty$ is therefore

$$h = \mathrm{K}_0(\alpha r) H(r) \tag{3.54}$$

$$v = \mathrm{K}_1(\alpha r) V(r), \tag{3.55}$$

where $H(r)$ and $V(r)$ tend to some finite non-zero limits H_∞ and V_∞ respectively as $r \to \infty$, and $\alpha^{-1} = (F|f|)^{-1}$ is the Rossby deformation radius for this problem. Differential equations for H and V are

$$\frac{dV}{dr} = \frac{\mathrm{K}_0(\alpha r)}{\mathrm{K}_1(\alpha r)}(\alpha V + F^2 f H) \tag{3.56}$$

$$\frac{dH}{dr} = \frac{\mathrm{K}_1(\alpha r)}{\mathrm{K}_0(\alpha r)}\left(fV + \alpha H + \frac{\mathrm{K}_1(\alpha r)}{r} V^2\right), \tag{3.57}$$

from which it follows that

$$V_\infty = -\frac{\alpha}{f} H_\infty. \tag{3.58}$$

Then, given an initial estimate for H_∞, and taking some $R \gg \alpha^{-1}$, we take $H = H_\infty$ at $r = R$, and (3.58) as the relationship between H and V at $r = R$. We can then integrate (3.56, 3.57) backwards to the boundary of the vortex.

The sole exception is in the case $f = 0$, where the form of the solution for $\overline{v}$ is known outside the vortex boundary:

$$\overline{v} = \frac{A}{r}; \quad \overline{h} = -\frac{A^2}{2r^2}, \tag{3.59}$$

where A is an arbitrary constant, and hence $\overline{v}$ and $\overline{h}$ are known in terms of the constant A rather than the constant H_∞.

In the vicinity of $r = 0$, $\overline{h} \sim O(1)$ and $\overline{v} \sim O(r)$ as $r \to 0$. In this case, simple power series expansions in r of the form

$$\overline{h}(r) = h_0 + h_2 r^2 + h_4 r^4 + h_6 r^6 + \cdots \tag{3.60}$$

$$\overline{v}(r) = v_1 R + v_3 r^3 + v_5 r^5 + v_7 r^7 + \cdots \tag{3.61}$$

are sufficient to give $\overline{h}(r)$ and $\overline{v}(r)$ at some small, non-zero r, given a guess h_0 for $h(0)$. We can then integrate (3.7, 3.8) forwards in r to the boundary of the vortex.

At the boundary of the vortex, we impose continuity of $\overline{h}$ and $\overline{v}$. A Newton method was used to iterate on h_0 and H_∞ (and h_0 and A in the case of $f = 0$) until the discontinuities in $\overline{h}$ and $\overline{v}$ across the boundary fell below some prescribed tolerance. To employ a Newton method, the differential equations (3.7, 3.8, 3.56, 3.57) and the starting series (3.58, 3.60, 3.61) were differentiated with respect to the parameters h_0 and H_∞, so that the variation of $\overline{v}$ and $\overline{h}$ with respect to h_0 and H_∞ on either side of the vortex boundary could be obtained (in the case $f = 0$, it is trivial to obtain expressions for the derivatives of the boundary velocity and height with respect to A). For $F \ll 1$, good initial guesses for h_0 and H_∞ or A were obtained from a matched asymptotic analysis, with $\overline{h}$ given by C_1 due to (3.25) and (3.36), and $H_\infty = -f/2$ due to (3.33), or $A = 1/2$.

We now turn to the disturbance equations.

The analysis of the eigenvalue equations is simplified by realising that the amplitude of the eigenfunctions is irrelevant, and that only the ratio of v_r/v_θ on either side of the vortex boundary is required to compute the eigenvalue ω.

As for the basic state, both $r = 0$ and $r = \infty$ are singular points of the equations, and

some use must be made of the asymptotic form of the solution in these limits to apply the appropriate boundary conditions, which are boundedness at $r = 0$ and evanescence as $r \to \infty$.

We start by considering the limit $r \to 0$. Here, we may take $v_r = r^{m-1} + k_1 r^{m+1} + \cdots$. Regularity at $r = 0$ then implies that $v_\theta = ir^{m-1} + k_2 r^{m+1} + \cdots$. The constants k_1 and k_2 can then be obtained from the eigenvalue equations, and

$$v_r/v_\theta = -i + K_2 r^2 + O(r^4). \tag{3.62}$$

The constant K_2 depends on the eigenvalue ω. Hence, given an initial guess for ω, the ratio of v_r to v_θ is known in the limit $r \to 0$ from the series (3.62). The ratio is used to initialize the r-integration of the eigenvalue equations with a starting value of $r = 0.01$. In all eigenvalue calculation, the starting series (3.62) was terminated at $O(r^2)$. Hence, the ratio v_r/v_θ is known just inside the vortex boundary.

At infinity, since the background flow is exponentially small there, we have

$$h \sim H_m^{(1)}(\lambda r) \tag{3.63}$$

with exponentially small corrections, where

$$\lambda^2 = (\omega^2 - f^2)F^2, \tag{3.64}$$

and the type of the Hankel function is chosen to satisfy the radiation condition. Rearranging (3.9, 3.10, 3.11) we have a relationship between v_r and v_θ through

$$(f^2 - \omega^2)v_r + i\omega\frac{dh}{dr} + \frac{1}{r}imfh = 0 \tag{3.65}$$

$$(f^2 - \omega^2)v_\theta - f\frac{dh}{dr} - \frac{1}{r}m\omega h = 0. \tag{3.66}$$

In the case where $\omega^2 \sim f^2$ we can still find a non-singular expression for the ratio v_r/v_θ by expanding the Hankel function representation of h. Again, the amplitude of h is not important – it is sufficient to know the ratio of the velocities v_r/v_θ.

Note that there is no need to integrate the disturbance equations in from the radiation far field of the Hankel function. This is fortunate, since, if $\omega^2 \sim f^2$, this would require a very large range of integration indeed. The equations (3.65, 3.66) assume negligible basic state velocity and height fields, but not a far field form of the wave field. Therefore, it is sufficient to integrate in from a radius where the basic state is sufficiently small. Since the basic state decay

is independent of the eigenvalue, the required position of the integration for the eigenfunction equations in the region $r > 1$, is fixed for given F and $|f|$, and decreases as F and $|f|$ increase.

Given an initial guess for ω, we must iterate on ω to satisfy the continuity equation (3.18). Here it is not so convenient to use a Newton method, due to the non-analytic behaviour of the solution at $\omega = \pm f$. However, iteration from two initial guesses using linear interpolation to predict the location of the root was found to perform reasonably well, typically converging in 5 or 6 iterations.

At low Froude number, we supply an initial guess for ω by considering the Kelvin modes on a circular incompressible vortex, with frequency given by (3.19). This can be used as a first guess ω_0 for the eigenfrequency, and a second guess can be supplied which is not too far away ($\omega_0 - 0.05$ say).

At low Froude number, growth rates are typically very low, and this places a strain on the numerical accuracy required. In all cases presented here, NAG routines were used for integrating the ODEs for both the basic state problem and the eigenvalue problem. In the exterior of the vortex, the differential equations are stiff, and a backward differencing method was used. A forward differencing method was used for the interior of the vortex. Tolerances of 10^{-15} were specified in all numerical integrations. In the disturbance problem, the computational values of the eigenfunctions were kept of order unity throughout the ranges of integration by applying the scaling function $(1+(Fr)^{2m})^{-1}r^{-1/2}e^{-\lambda_r r}$ to the variables in the exterior of the vortex (where λ_r is the real part of λ, and λ is given by (3.64)), and the function r^{m-1} to the interior variables. The very low tolerance specification, together with the rescaling to make the computational variables of order unity throughout the domains of integration, means that we can have some confidence in the imaginary parts of the eigenvalues down to around 10^{-14}.

Results of the eigenvalue calculations

The growth rate of the instability of the axisymmetric vortex depends on three parameters: the Froude number F; the Rossby number f^{-1}; and the azimuthal mode number of the instability m.

We begin by investigating the growth rates of the instability in the case of infinite Rossby number ($f = 0$). This limit is most similar to the study of Broadbent & Moore (1979), except

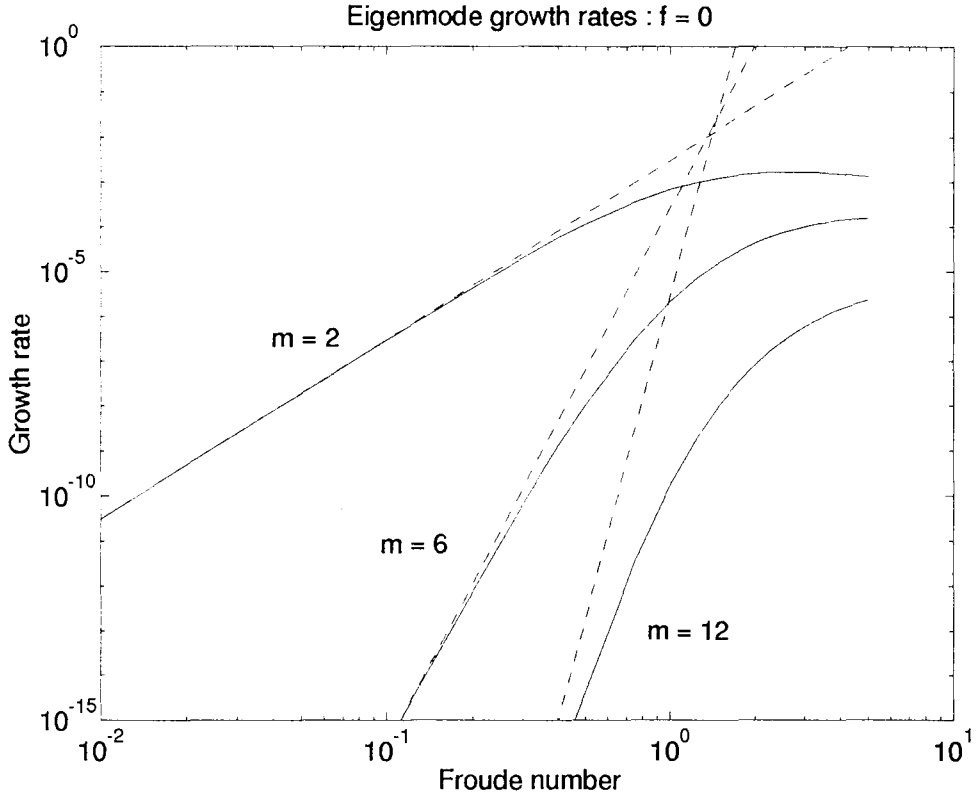


Figure 3.2: Axisymmetric vortex with $f = 0$. Growth rates of unstable modes (solid lines), and their low Froude number limits (dashed lines).

here we take the vortex to have uniform potential vorticity, rather than uniform vorticity as they did.

Figure 3.2 shows the growth rates of the instability for Froude numbers from 0 to 5 and for mode numbers $m = 2, 6, 12$. Solid lines are growth rates obtained from the numerical eigenvalue calculation, and dotted lines are the growth rates given by the low Froude number analysis (equation 3.51). The graph is plotted on a log-log scale, so that at low Froude number the slope of the lines is $2m$. Mode 2 is found to be the most unstable, and grows more slowly with Froude number than the higher mode numbers. The numerical eigenvalue calculations are apparently reproducing the eigenvalues obtained from low Froude number limit with growth rates down to 10^{-14}, and we can be confident that numerical algorithm used to integrate (3.12) & (3.13) is performing adequately.

The real part of the eigenfrequency divided by m is shown in figure 3.3. It is always bounded between 0 and $m \times V(1)$, apparently approaching $m \times V(1)$ at large m.

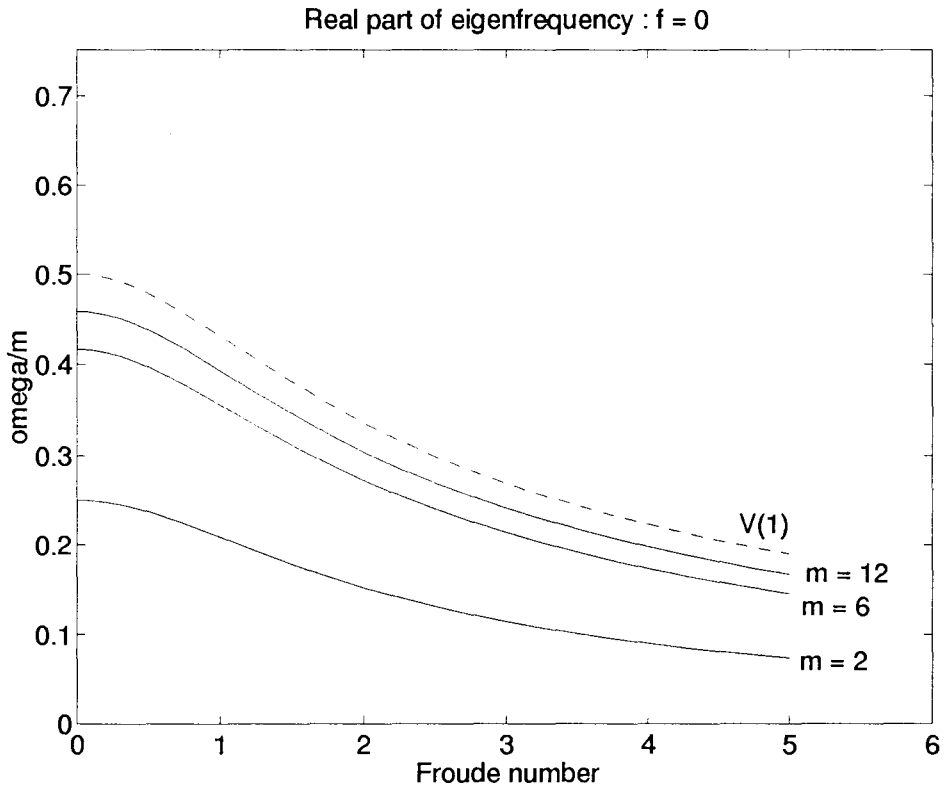


Figure 3.3: Axisymmetric vortex with $f = 0$. Real parts of eigenfrequencies / m (solid lines), with $V(1)$ (dashed line) for comparison.

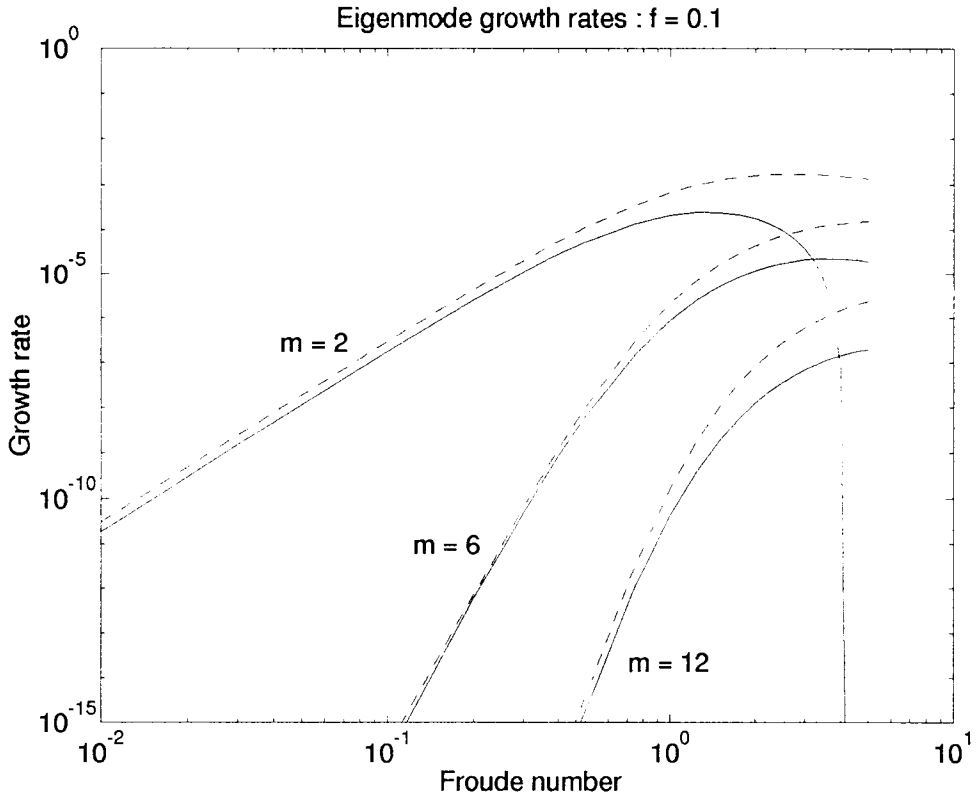


Figure 3.4: Axisymmetric vortex with $f = 0.1$. Growth rates of unstable modes (solid lines), together with growth rates of modes with $f = 0$ (dashed lines) for comparison.

We now turn to investigating the instability at non-zero f. In particular, it is important to note that for f between 0 and -1, the vortices are anticyclones with potential vorticity of opposite sign to the background potential vorticity. For $f > 0$ the vortices are always cyclones, and for $f < -1$, they are anticyclones with potential vorticity of the same sign as the background. If $f \neq 0$, we expect the effect of background rotation to inhibit the instability if the real part of the eigenfrequency is below $|f|$. We start by investigating the effect on the instability of weak background rotation, in the first case taking $|f| = 0.1$.

In figure 3.4, growth rates are shown for $f = 0.1$, and in figure 3.5 they are shown for $f = -0.1$. For reference, the dashed lines in figures 3.4 & 3.5 show the corresponding growth rates of the instability for $f = 0$. In the case $f = 0.1$, the vortices should be regarded as very intense cyclones, with potential vorticity equal to 11 times the background value, whereas in the case $f = -0.1$ they are very intense anticyclones, with potential vorticity equal to -9 times the background value. The corresponding real parts of the eigenfrequencies divided by m are

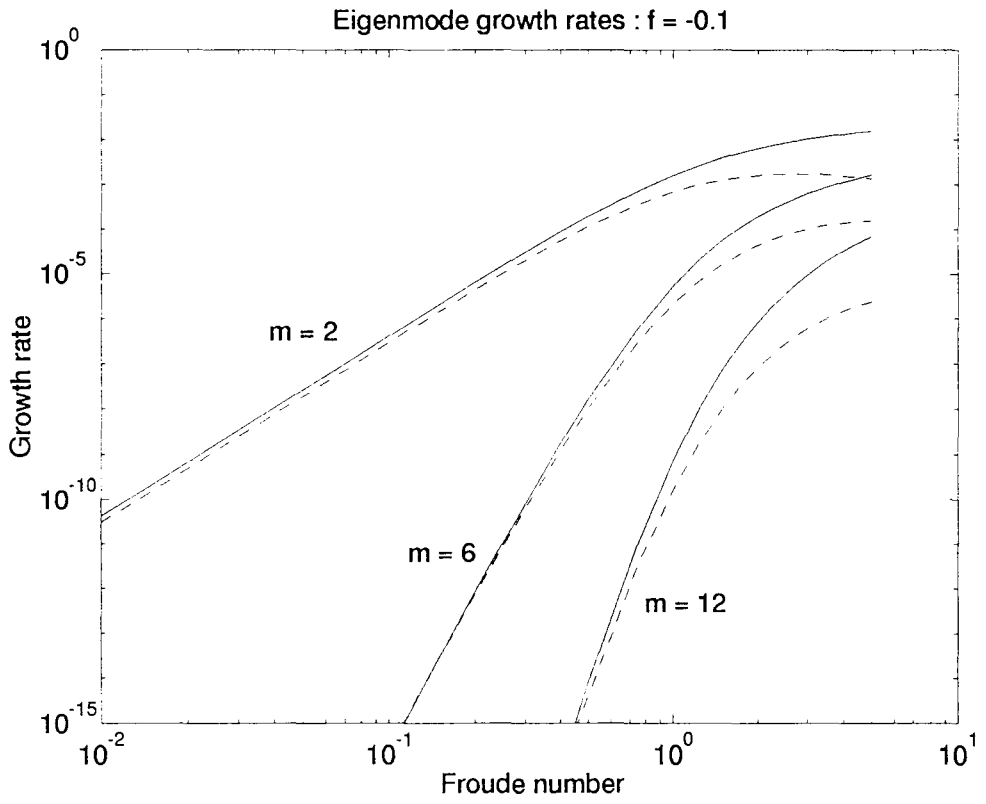


Figure 3.5: Axisymmetric vortex with $f = -0.1$. Growth rates of unstable modes (solid lines), together with growth rates of modes with $f = 0$ (dashed lines) for comparison.

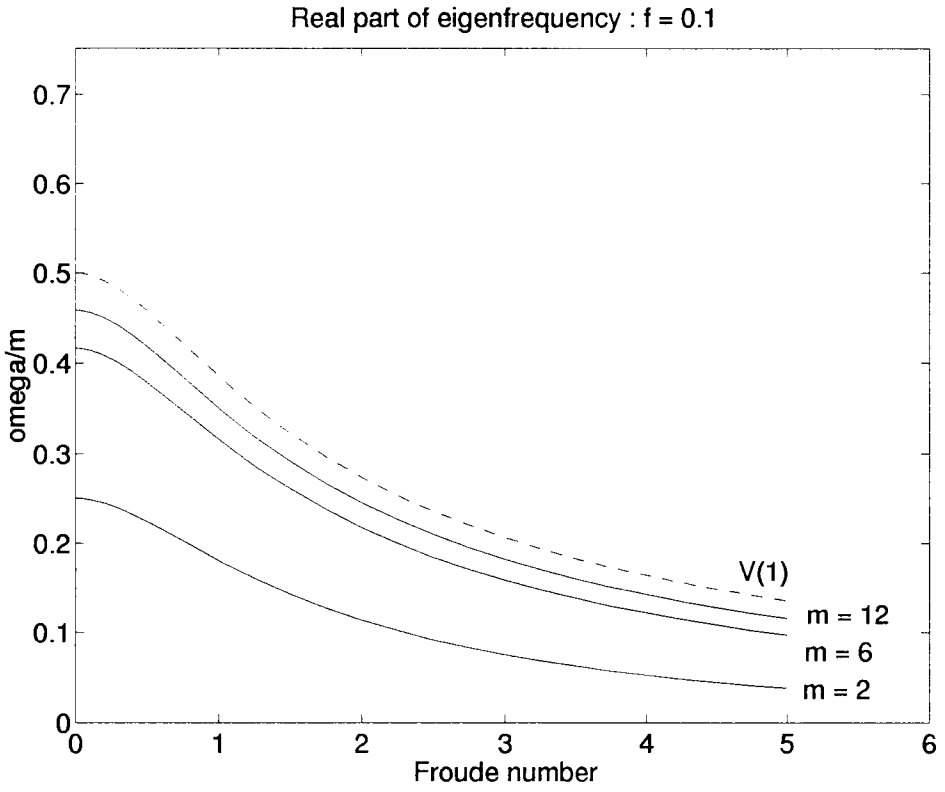


Figure 3.6: Axisymmetric vortex with $f = 0.1$. Real parts of eigenfrequencies / m (solid lines), with $V(1)$ (dashed line) for comparison.

shown in figures 3.6 & 3.7.

Notice that $V(1)$ is quite different for $f > 0$ and $f < 0$ as F is varied, but that the real part of the eigenfrequency is still bounded between 0 and $m \times V(1)$, approaching $m \times V(1)$ at large m.

Figures 3.4 & 3.5 show that a small amount of background rotation has a significant effect on the growth rate of the instability in both the cyclonic and anticyclonic cases, although in opposite senses. In figure 3.4 we see that the background rotation inhibits the instability of the cyclone. One would expect it to do so, but to completely cut off the instability the real part of the eigenfrequency must be reduced from 0.5 at $F = 0$ to 0.1 at $F \approx 4$. It is perhaps surprising that the velocity at the boundary of the vortex, and hence the real part of the eigenfrequency, decreases so markedly with increasing Froude number. By contrast, in the anticyclonic case $f = -0.1$, the growth rates of the instability are markedly increased over their corresponding $f = 0$ values, and the real part of the eigenfrequency does not decrease nearly so rapidly with

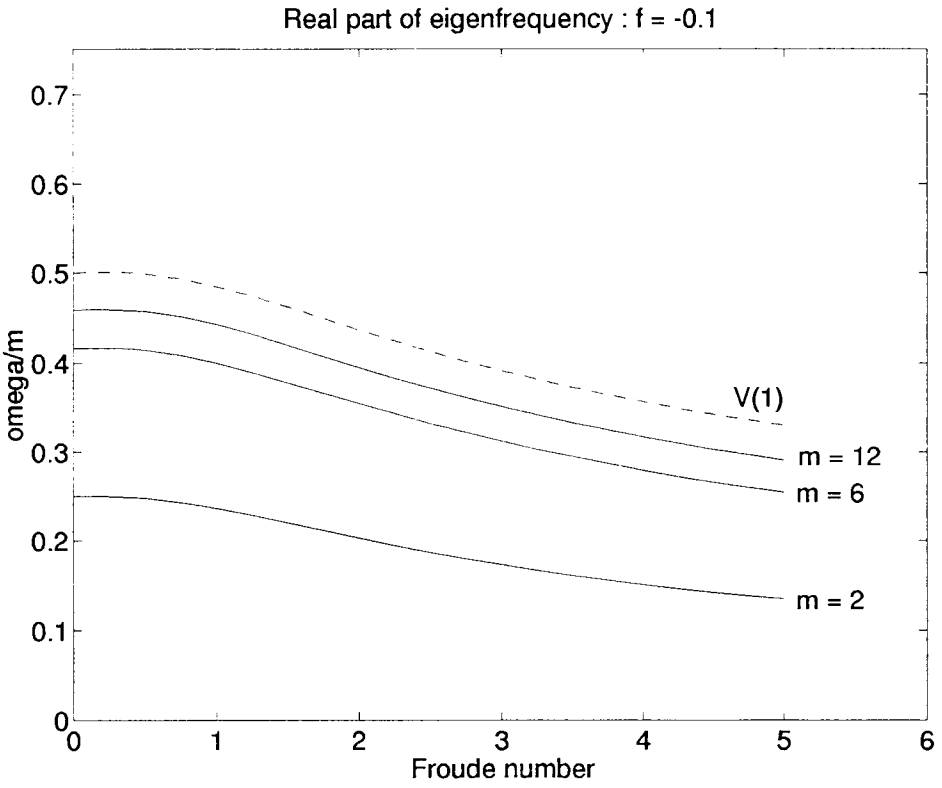


Figure 3.7: Axisymmetric vortex with $f = -0.1$. Real parts of eigenfrequencies / m (solid lines), with $V(1)$ (dashed line) for comparison.

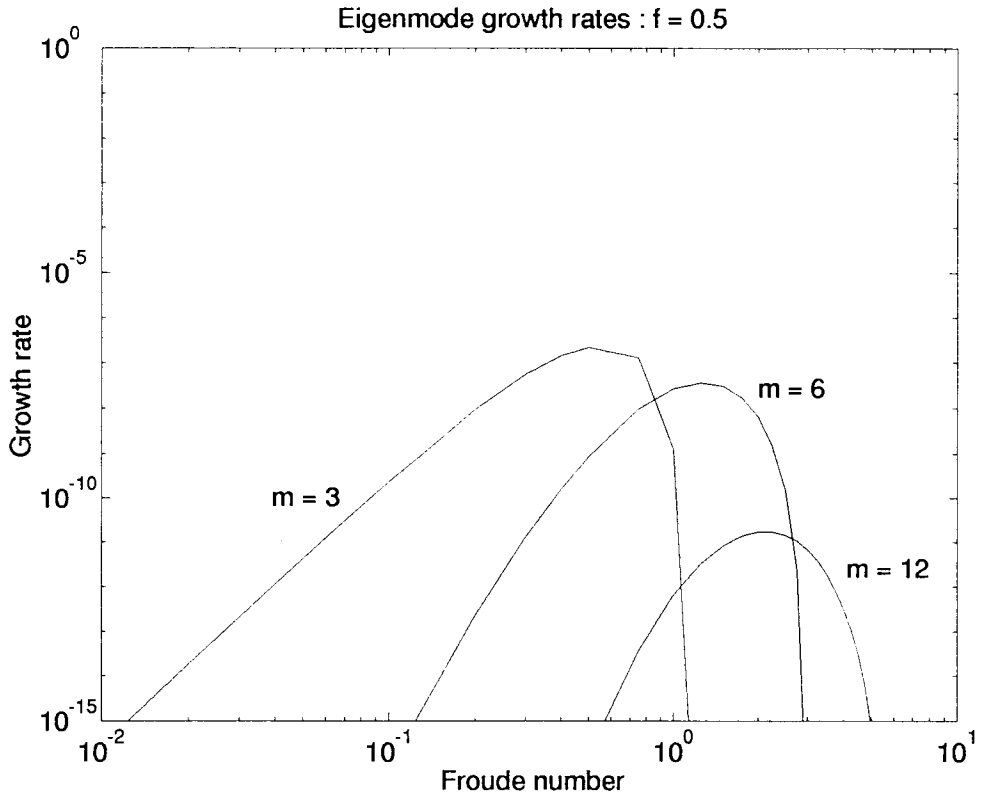


Figure 3.8: Axisymmetric vortex with $f = 0.5$. Growth rates of unstable modes.

increasing Froude number.

When $|f| = 0.5$, we reach the situation where the mode $m = 2$ instability is not unstable at leading order in Froude number, since then we have $\omega = |f|$ at leading order in F. In figure 3.8 growth rates are shown for $f = 0.5$. Mode 2 is stable over all Froude number shown on the graph, and growth rates for modes 3,6 and 12 are shown. As F is increased, so more and more modes become stabilized, as the vortex boundary velocity decreases and the real part of the eigenfrequency falls below f. This is shown in figure 3.9. The instability of the vortex at $f = 0.5$ is summarized in figure 3.10. In figure 3.10a we see that the maximum growth rate, over all Froude numbers, occurs at a Froude number of about 0.9. As F is increased further, the mode number of the fastest growing mode increases, apparently almost linearly with F (see figure 3.10b), and the growth rate of the fastest growing mode decreases, being barely greater than 10^{-15} at $F = 5.0$.

In the anticyclonic case $f = -0.5$, shown in figure 3.11, the mode 2 instability persists for

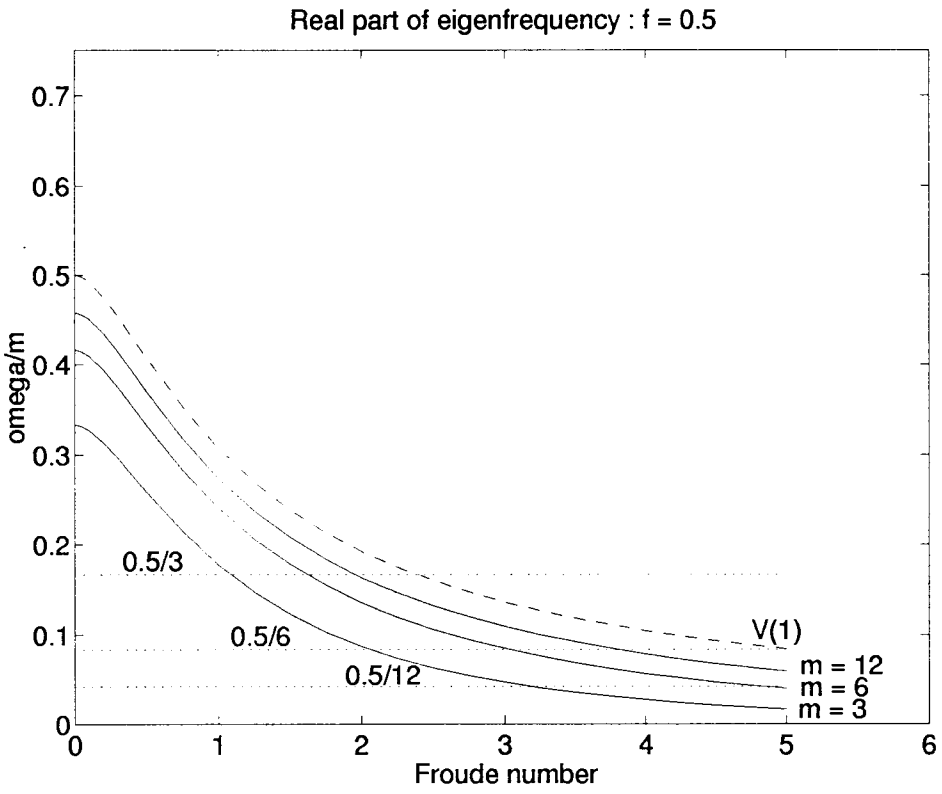


Figure 3.9: Axisymmetric vortex with $f = 0.5$. Real parts of eigenfrequencies / m (solid lines), with $V(1)$ (dashed line) for comparison. Dotted lines are the critical values of ω/m, below which the instability is inhibited for each given m. Modes $m = 3$ and $m = 6$ are inhibited at some $F < 5$. Mode 12 is not inhibited for $F < 5$, but the growth rate of mode 12 is less than 10^{-15} at $F = 5$ (see figure 3.8).

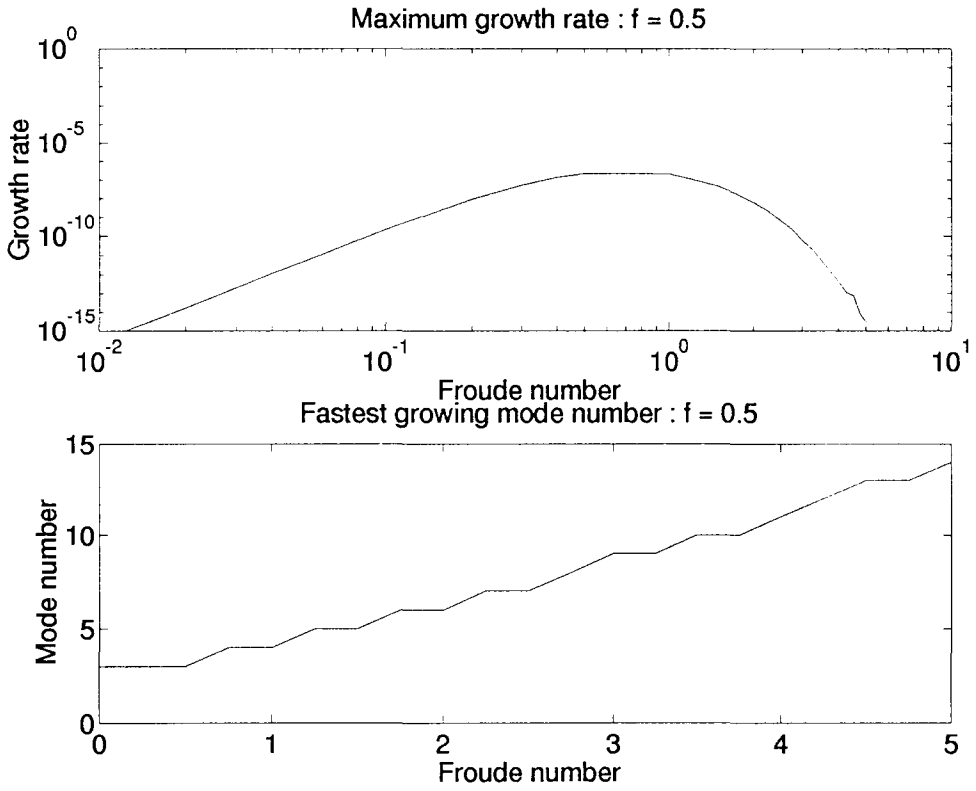


Figure 3.10: Axisymmetric vortex with $f = 0.5$. (a): Growth rate of fastest growing mode, and (b): Mode number of fastest growing mode

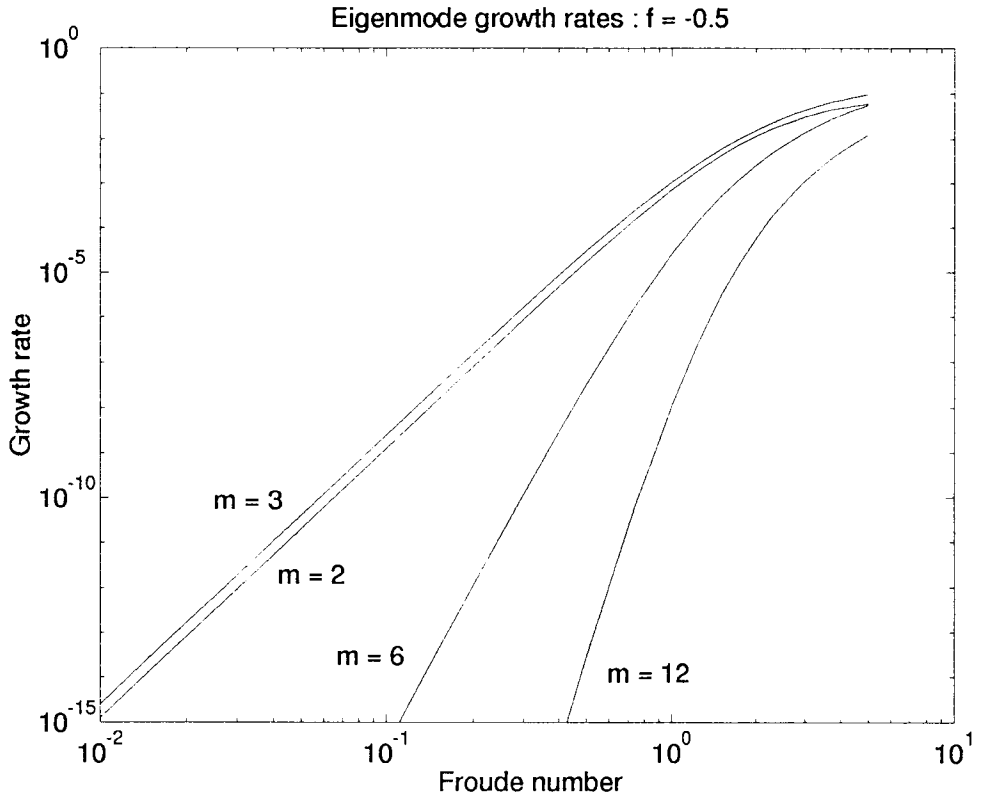


Figure 3.11: Axisymmetric vortex with $f = -0.5$. Growth rates of unstable modes.

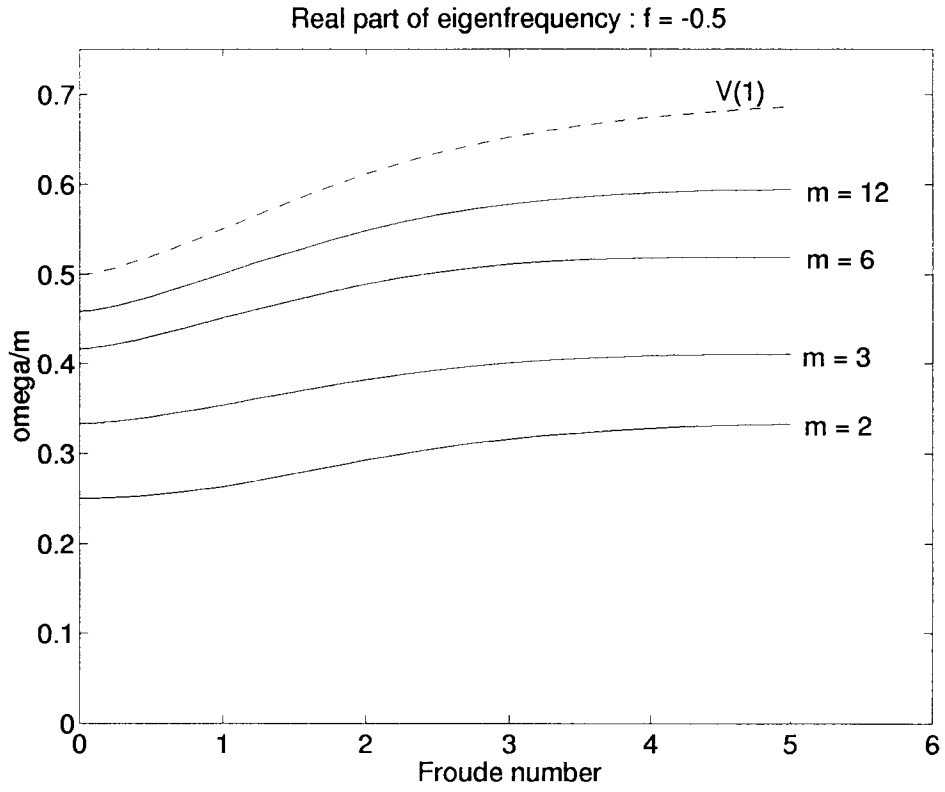


Figure 3.12: Axisymmetric vortex with $f = -0.5$. Real parts of eigenfrequencies / m (solid lines), with $V(1)$ (dashed line) for comparison.

all non-zero F, but its behaviour at small F is changed, increasing with the 6th power of the Froude number, rather than the 4th. In this case, the real part of the eigenfrequency tends to increase as F is increased, consistent with an increase of $V(1)$ (figure 3.12).

When $|f| = 1.0$, mode 3 ceases to be unstable at small Froude number. Moreover, for $|f| > 1.0$ the anticyclones have potential vorticity $(f + 1)$ of the same sign as the background potential vorticity f, and the case $f = -1$ corresponds to a vortex of zero potential vorticity. Figure 3.13 shows the growth rates of the instability for $f = 1$. The maximum growth rate never exceeds 10^{-10}, and the growth rate of mode $m = 12$ never exceeds 10^{-15}. The real part of the eigenfrequency behaves much as before, with $V(1)$ now decreasing very rapidly as F is increased (figure 3.14).

In figure 3.15, the growth rates of the instability are shown for $f = -1.0$. In contrast to the case $f = 1.0$, the instabilities persist over a wide range of F. The vortex boundary velocity is always 0.5, and the real parts of the eigenfrequencies do not decrease markedly as

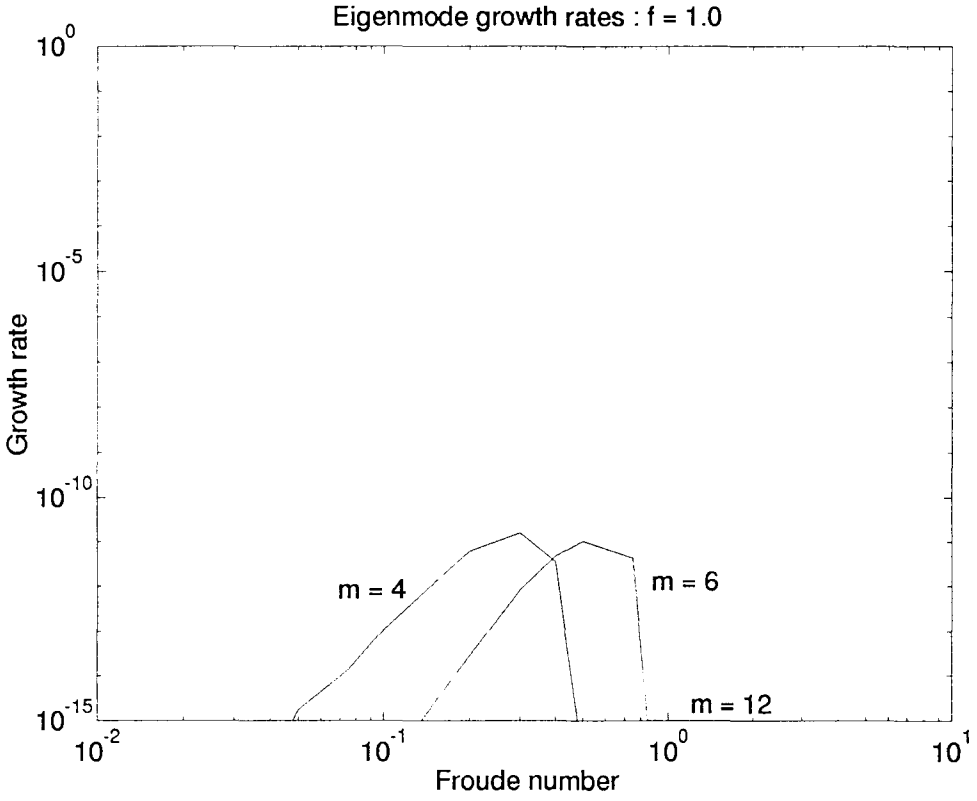


Figure 3.13: Axisymmetric vortex with $f = 1.0$. Growth rates of unstable modes.

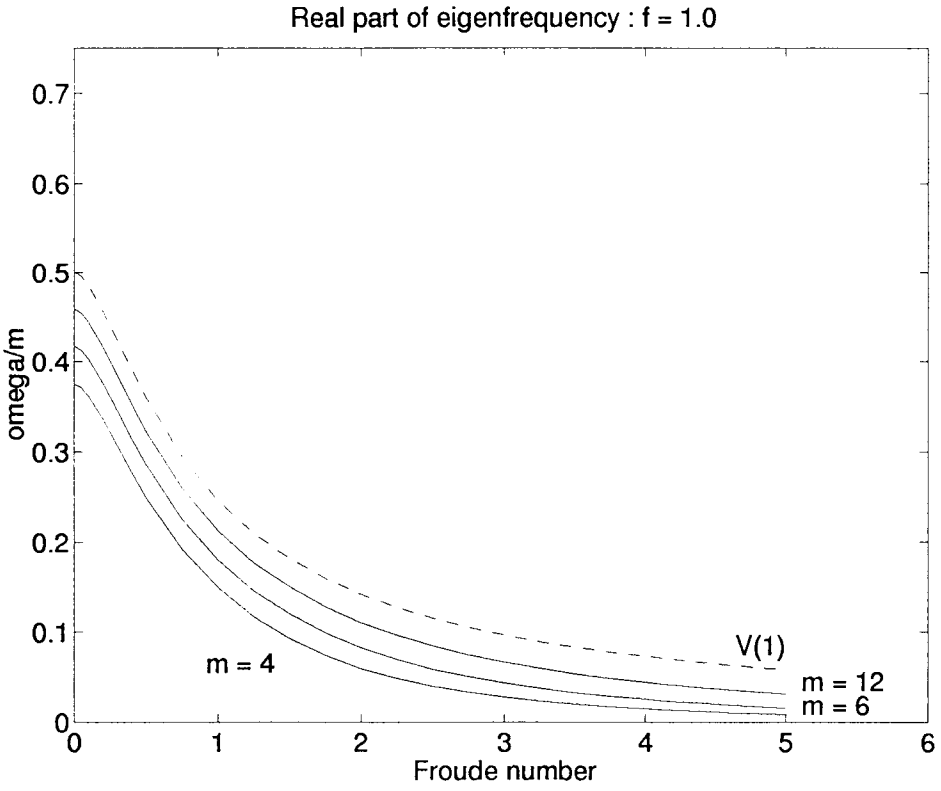


Figure 3.14: Axisymmetric vortex with $f = 1.0$. Real parts of eigenfrequencies / m (solid lines), with $V(1)$ (dashed line) for comparison.

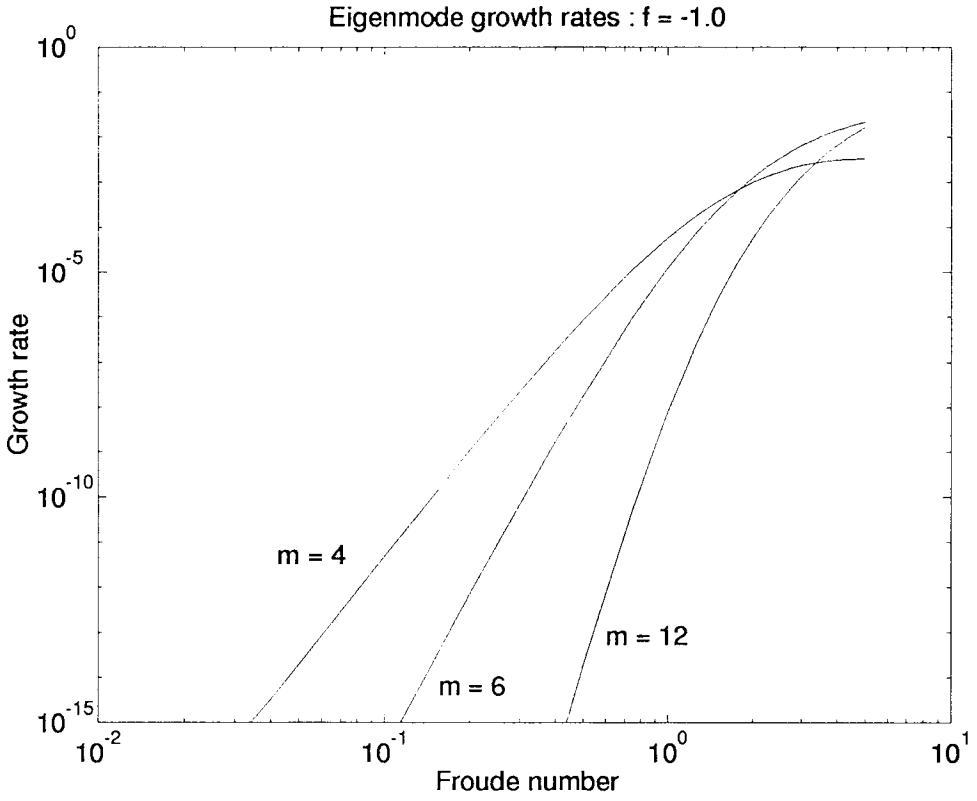


Figure 3.15: Axisymmetric vortex with $f = -1.0$. Growth rates of unstable modes.

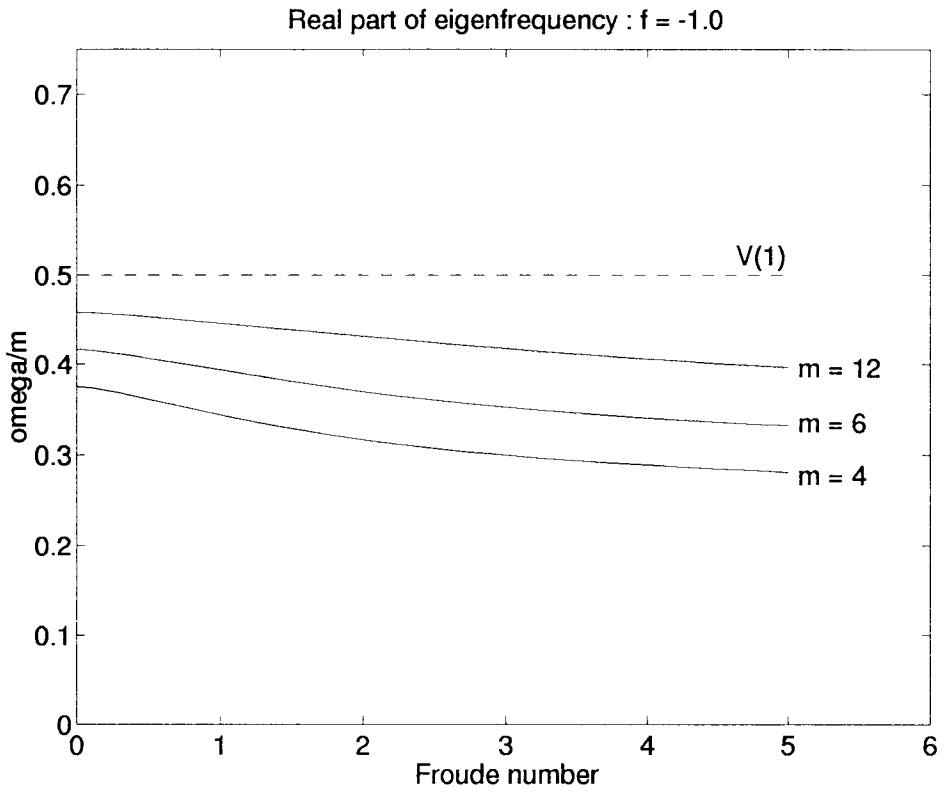


Figure 3.16: Axisymmetric vortex with $f = -1.0$. Real parts of eigenfrequencies / m (solid lines), with $V(1)$ (dashed line) for comparison.

F is increased (see figure 3.16). Mode 3, which is marginally stable at $F = 0$, for $f = \pm 1.0$, is never unstable for $|f| = 1.0$, either for positive or negative f. This is quite different from the case $f = -0.5$, where the marginally stable mode, $m = 2$, became unstable in the anticyclonic case, but remained stable in the cyclonic case.

For $f \sim 1$ it appears that the cyclones have instabilities with only very weak growth rates, if at all, and it seems likely that increasing f still further will continue to reduce the growth rates, at the same time requiring larger and larger m for the instability to exist at all. However, for $f \sim -1$, the anticyclones still have comparatively large growth rates, and in figures 3.17 & 3.18 we see the the effect of increasing $|f|$ above 1 in the anticyclonic case.

In figure 3.17 the growth rates of mode 4 instabilities are shown for $f = -1.1, -1.2, -1.3$. We can see that as $|f|$ is increased beyond 1.0, the effect of background rotation very rapidly reduces instability at mode number 4, the maximum growth rate at $f = -1.3$ being four order of magnitude smaller than the maximum growth rate at $f = -1.1$. Figure 3.18 shows growth

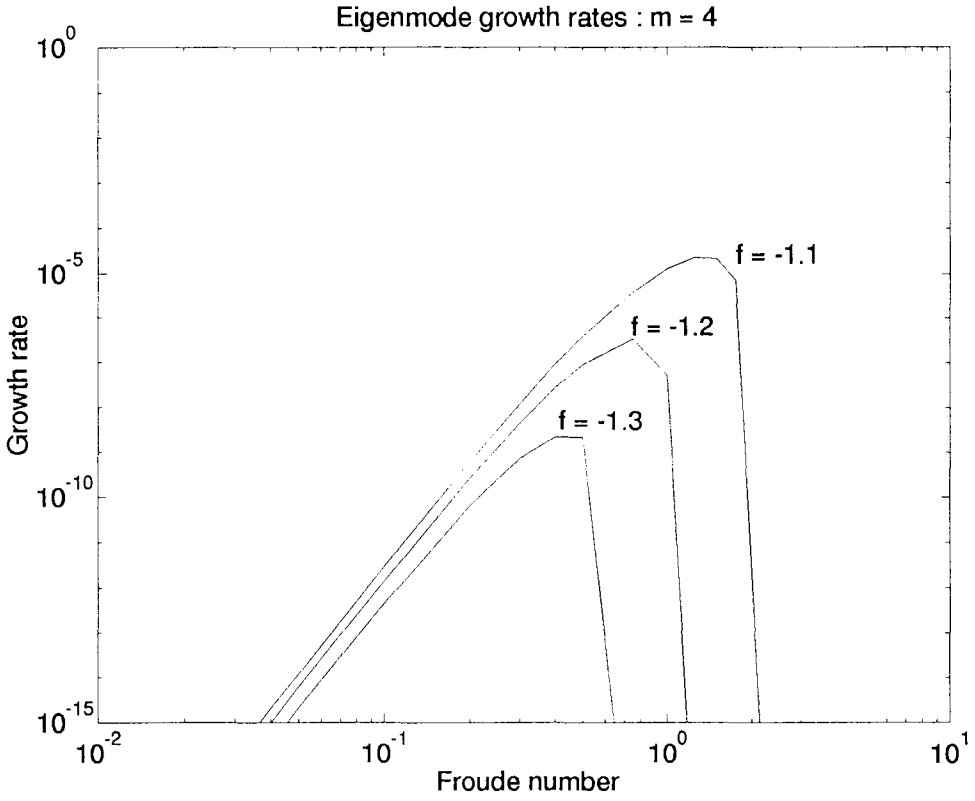


Figure 3.17: Growth rates of mode number 4 disturbances to vortices with $f = -1.1, -1.2, -1.3$.

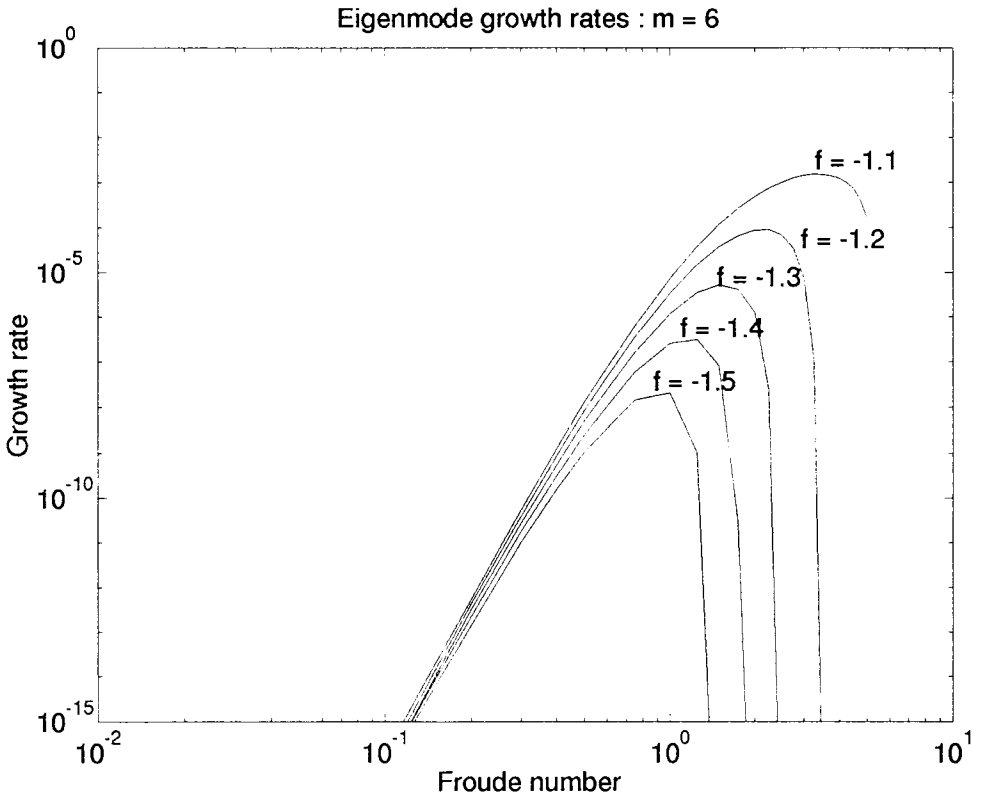


Figure 3.18: Growth rates of mode number 6 disturbances to vortices with $f = -1.1, -1.2, -1.3, -1.4, -1.5$.

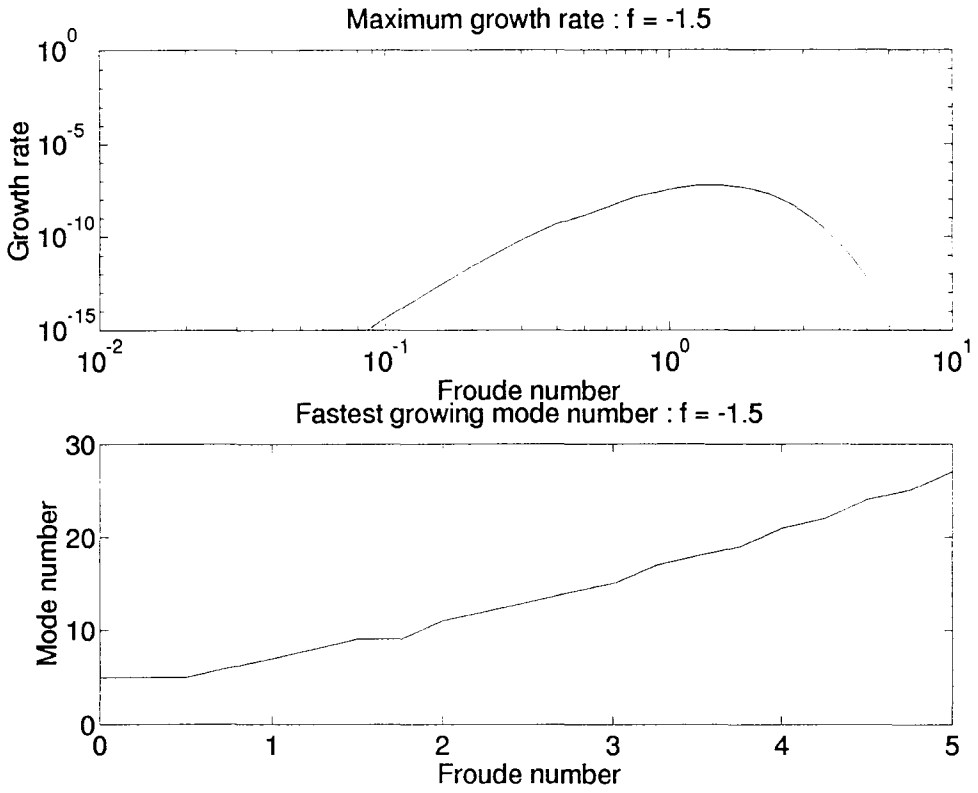


Figure 3.19: Axisymmetric vortex with $f = -1.5$. (a): Growth rate of fastest growing mode, and (b): Mode number of fastest growing mode.

rates of mode 6 instabilities for $f = -1.1, -1.2, -1.3, -1.4, -1.5$. For this higher mode number the instabilities persist for larger $|f|$ than in the case $m = 4$, but are still significantly inhibited as $|f|$ is increased. It seems that the instabilities of the anticyclones with $f < -1$ are very similar to instabilities of the cyclones $f > 1$, with the instability at moderate to large Froude numbers persisting only at very high mode numbers.

The instability of the vortex at $f = -1.5$ is summarized in figure 3.19, which is to be compared with figure 3.10 for the vortex with $f = 0.5$. The overall maximum growth rate now occurs at a Froude number of about 1.1, with the fastest growth rate reducing to about 10^{-12} at $F = 5.0$. The mode number of the fastest growing mode increases almost linearly with m, reaching $m = 27$ at $F = 5.0$. The picture is broadly similar to the case $f = 0.5$, shown in figure 3.10, although the mode numbers in this case are larger, and the growth rates slightly greater at larger values of F. It is in contrast to what is found for $0 > f \geq -1$, where the growth rates apparently increased without bound as F was increased.

We turn now to an explanation of the difference in behaviour of the instability between cyclonic and anticyclonic vortices. At low Froude number, a difference between positive and negative f exists in the expression (3.51) for the growth rate of the instability in the $F \ll 1$ limit. The difference is a factor of $(\omega - f)/(\omega + f)$, which is smaller if ω and f have the same sign than it is if they are of opposite sign. Since ω is always positive, this factor means that anticyclones are always more unstable to Rossby wave – gravity wave instability than cyclones. The difference lies in the nature of the coupling between the vortical region and the wave region. This difference is small, however, and amounts only to a constant factor. It cannot explain why low wave number disturbances to cyclonic vortices are stabilized as the Froude number is increased, whereas low wave number disturbances to anticyclonic vortices are destabilized.

The key to understanding the difference lies in the nature of the basic states. For small F, the azimuthal velocity at the vortex boundary is always 0.5. Since the wave on the boundary of the vortex is a Rossby wave, we expect it to propagate in a pseudo-westward direction – i.e. with a real frequency between 0 and $m \times V(1)$ in the case of a vortex of unit radius. Therefore the azimuthal velocity at the boundary of the vortex sets a bound for the real part of the eigenfrequency.

Of course, one reason why eigenmodes with comparatively high frequencies will be more unstable is that they are less influenced by the inertial cut-off $\omega = |f|$. While this is one explanation, it is not the only one, and is probably not the most significant. Regardless of the inertial cut-off frequency, we would still expect higher frequency eigenmodes to be more unstable than lower frequency modes, because at higher frequencies the interaction between vortical and gravitational motions becomes stronger.

To see this, first recall that we expect the growth rate of the instability to depend on the strength of coupling between the Rossby wave, on the vortex boundary, and the gravity waves, found distant from the vortex boundary. The character of the eigenmodes changes from being Rossby wave - like to gravity wave - like at the "sonic radius", at which the intrinsic angular phase speed of disturbances, $\omega r/m - V(r)$, is equal to the gravity wave phase speed $F^{-1}H^{1/2}(r)$. At a given F, it follows that, the larger the value of ω, the smaller r will have to be to achieve a given angular phase speed, and hence the closer the sonic radius will be to the boundary of the vortex. We should expect this to lead to an enhanced coupling between the Rossby

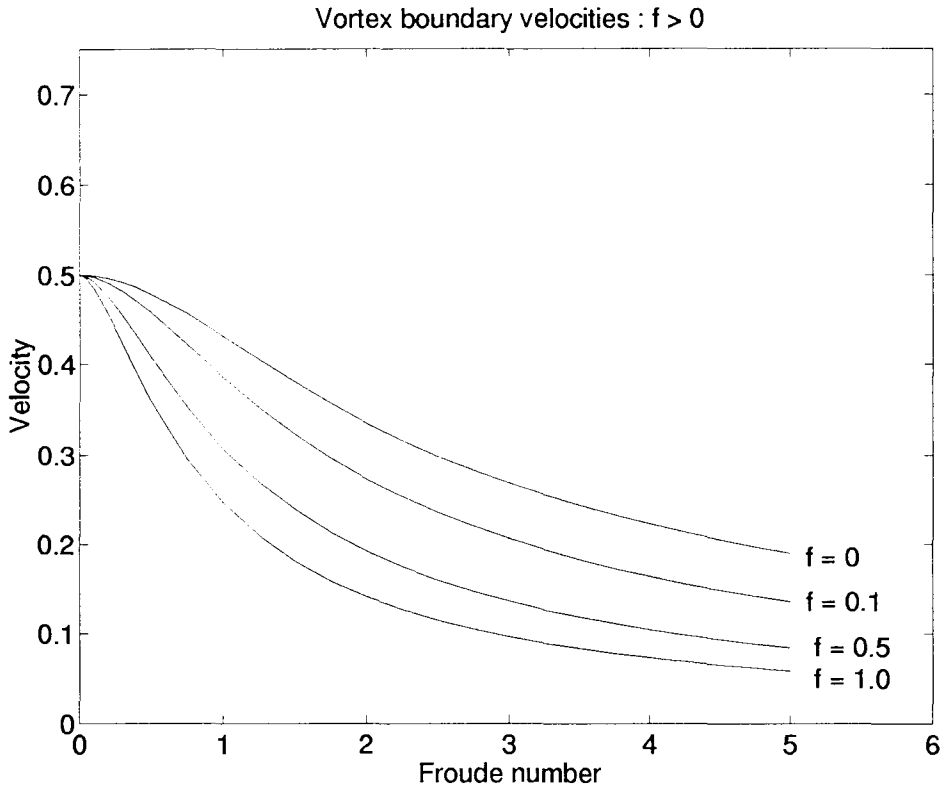


Figure 3.20: Azimuthal velocity at boundary of vortices, $f > 0$.

wave on the vortex boundary and the gravity wave beyond the sonic radius in cases where ω is comparatively large.

We shall return to this point in the next section, with reference to the WKBJ analysis. The analysis enables us to distinguish between the effect of the inertial cut-off, which primarily affects the growth rates of low mode number disturbances, and the effect of variation in the location of the sonic radius, which affects the growth rates at all mode numbers. For the present, however, we shall simply examine the large degree of variation in the vortex boundary velocities which is experienced as F and f are varied, and defer a quantitative discussion to the end of the next section.

In figure 3.20, the azimuthal vortex boundary velocity is shown as a function of Froude number for cyclones with various values of f. In all cases the vortex boundary velocity decreases very rapidly as the Froude number is increased, from a velocity of 0.5 at $F = 0$ to less than 0.2 at $F = 5$. Therefore, we expect the real part of the eigenfrequency to decrease as F is

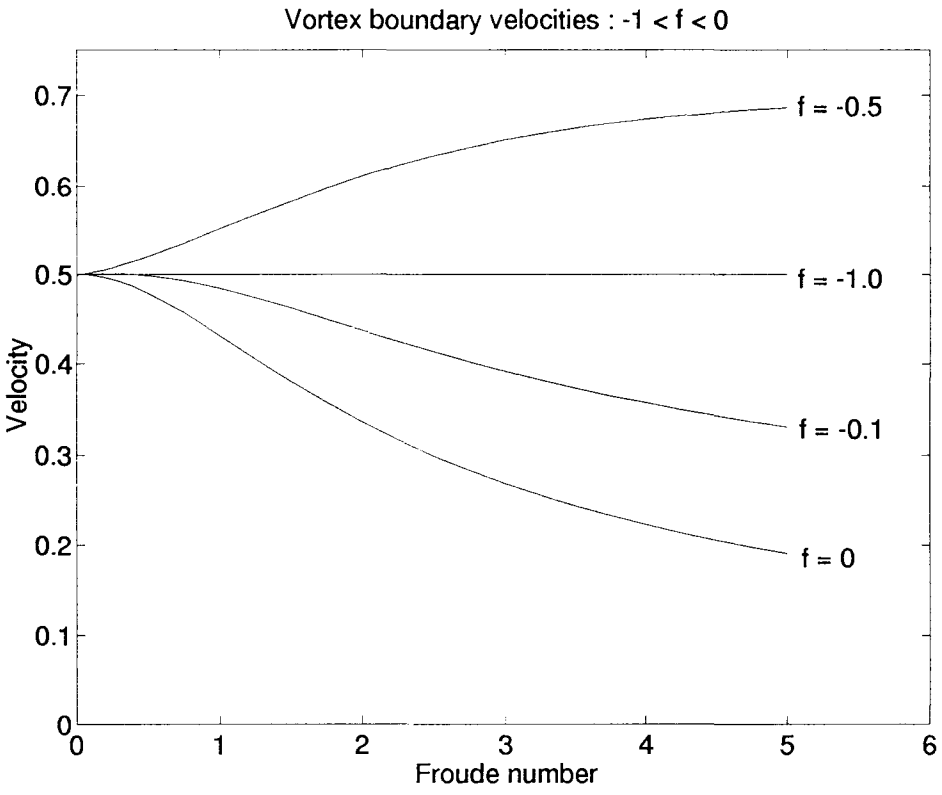


Figure 3.21: Azimuthal velocity at boundary of vortices, $-1 < f < 0$.

increased.

Vortex boundary velocities for anticyclones with $|f| < 1$ are shown in figure 3.21. For small $|f|$, the vortex boundary velocity decreases, just as in the cyclonic case, but much less rapidly. For example, with $f = -1/10$, the velocity decreases from 0.5 to 0.33 as F increases from 0 to 5, whereas with $f = 1/10$ it decreases from 0.5 to 0.14. The most striking difference comes for larger f, where the vortex boundary velocity increases as F is increased. The most extreme example is the case $f = -1/2$, where the velocity increases from 0.5 at $F = 0$ to 0.68 at $F = 5$. When $f = -1$, the vortex boundary velocity is exactly 0.5 for all F.

For $f < -1$, the vortex boundary velocities always decrease as F is increased. This is shown in figure 3.22. The growth rates of these instabilities tend to behave more like those of cyclones than the anticyclones with $-1 < f < 0$.

The results of the eigenvalue calculations may be summarized as follows. In the absence of background rotation, the eigenmode with $m = 2$ is always unstable in the range $0 < F < 5$,

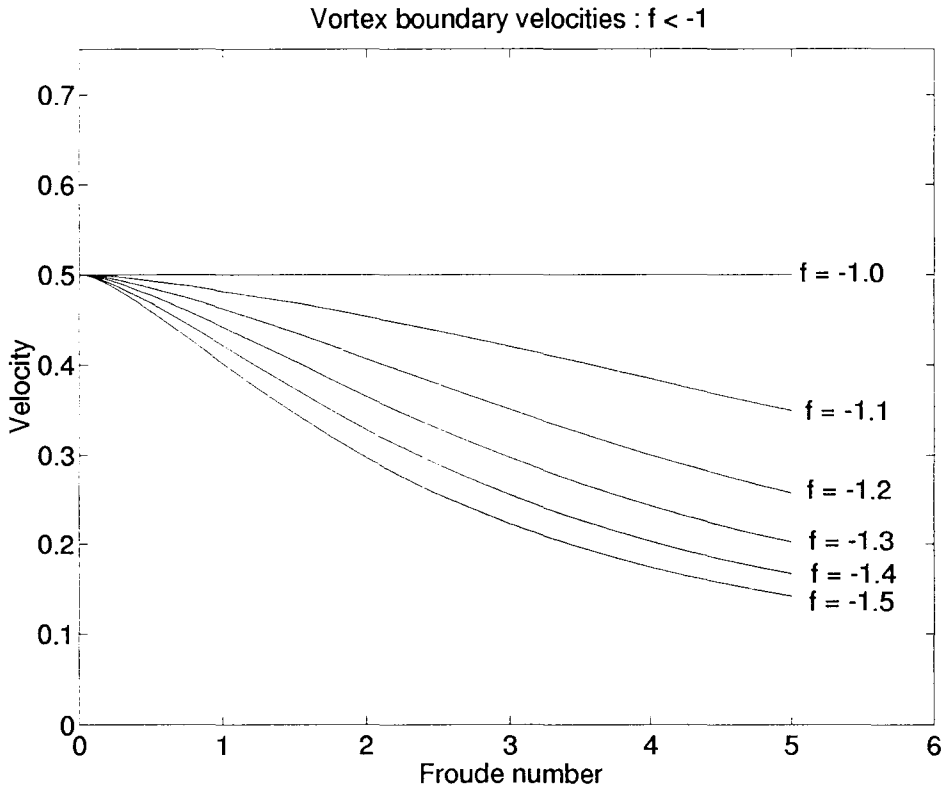


Figure 3.22: Azimuthal velocity at boundary of vortices, $f < -1$.

and is probably unstable for all F. With background rotation present, low mode number eigenmodes may be unstable at small Froude numbers, but become stabilized as the Froude number is increased. This is because the real part of the eigenfrequency decreases as the Froude number is increased, and eventually falls below $|f|$, at which point the instability mechanism no longer exists.

The maximum growth rates (taken over all m) are shown as a function of F and f in figures 3.23a & 3.24a. The growth rate and fastest growing mode number surfaces are shown from two different orientations for ease of visualization. With the exception of the vortices for which $-1 < f < 0$, the growth rate of the instability generally peaks at a finite value of the Froude number, and with a fairly low growth rate. However, when $-1 < f < 0$, it seems that the growth rate of the instability may increase without bound as F is increased. The corresponding fastest growing mode number is shown in figures 3.23b & 3.24b. For $-1 < f < 0$, the fastest growing mode number is typically small (either 2 or 3), whereas for values of f outside this range, the fastest growing mode number can be quite large. In the most extreme case shown, when $F = 5$ and $f = -1.5$, the fastest growing mode number is 27.

Thus, the cyclones all tend to exhibit the same dependence on F and f. As f and F are increased, successively higher mode numbers must be taken to retain the instability, whereas the anticyclones fall into two categories. When $-1 < f < 0$, the vortex has potential vorticity of opposite sign to the background. In these cases, there is no marked tendency to stabilize as F is increased. In the particular case $f = -1/2$ the instability becomes markedly stronger as F increases, with maximum growth rate for $m = 3$. When $f < -1$, however, the dependence of the instability on $|f|$ and F is similar to the cyclonic cases, and again large mode numbers are required to retain the instability. Strong instability by the Rossby wave - gravity wave instability mechanism seems possible only when $-1 < f < 0$ and the mode number m is not too large. These are the cases where the sonic radius is closest to the boundary of the vortex.

From a practical point of view, we have almost certainly identified those cases in which this instability is sufficiently vigorous that it might be observed in nature. However, we are interested not only in cases of practical significance, but also in whether coupling between vortical motions and gravity waves exists at all Rossby numbers, or whether these is a value of $|f|$ above which the instability ceases to exist, at least for some range of F, and in which cases the vortical motions could then be separated from the gravity waves in the problem.

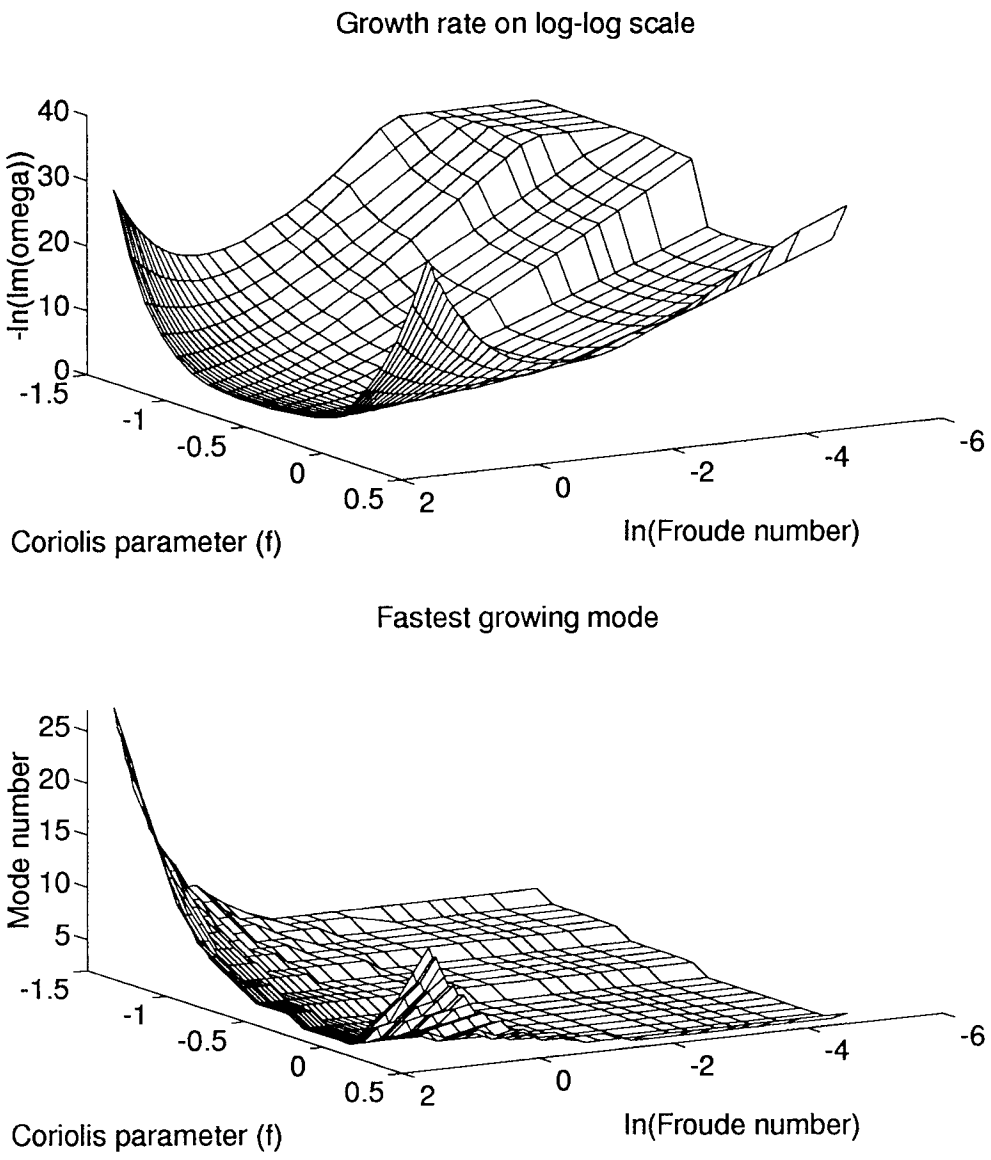


Figure 3.23: Summary of axisymmetric vortex instability calculations, viewed from orientation giving most complete view of growth rate surface.

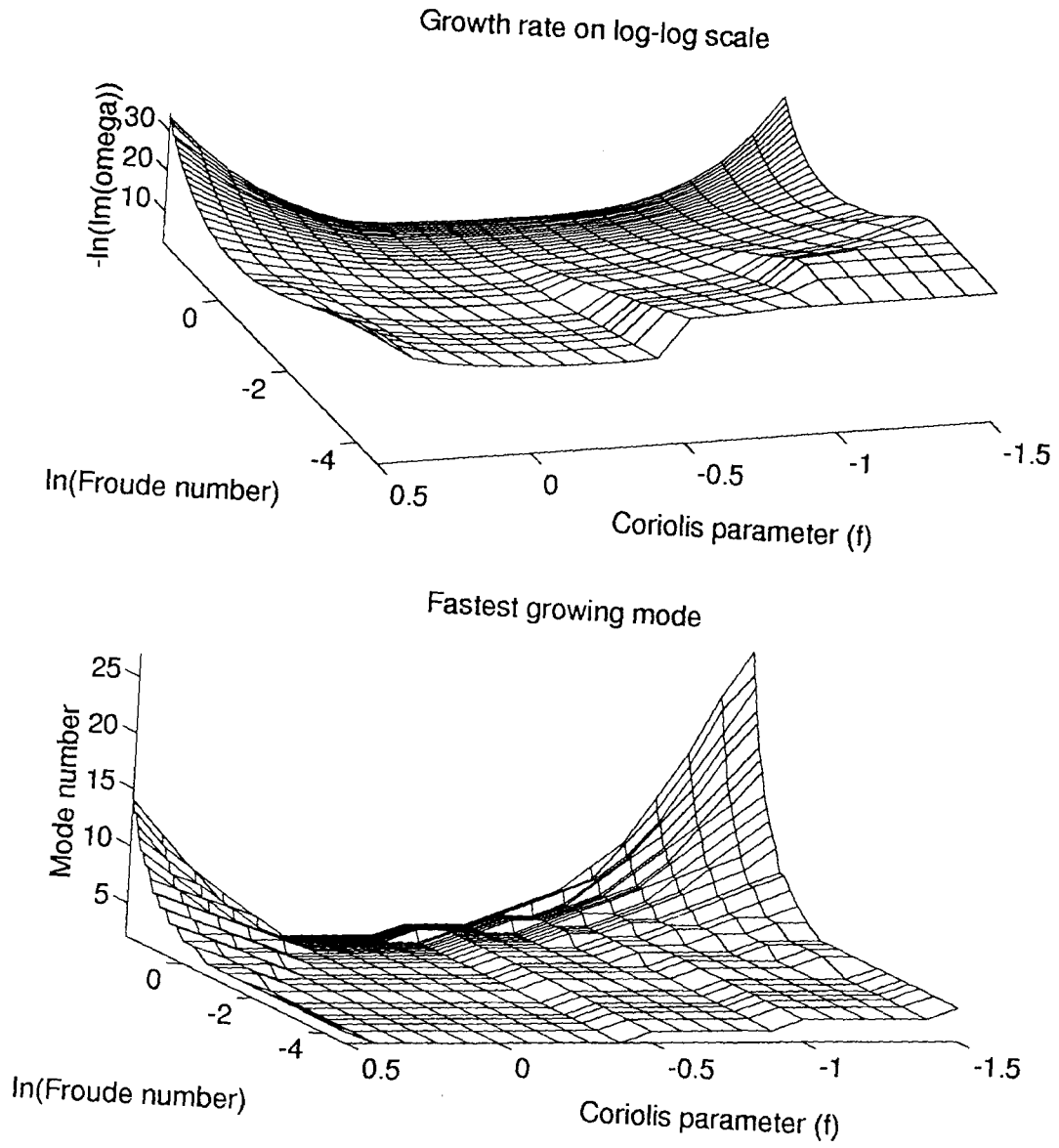


Figure 3.24: As figure 3.23, but viewed from orientation giving most complete view of fastest growing mode number surface. Note that (*a*) is viewed from a greater elevation than (*b*).

The next section therefore examines this instability for arbitrary f and F in the limit $m \gg 1$. The WKBJ analysis has the added advantage that it assumes *a priori* that $\omega \gg f$. Therefore any distinction between cyclones and anticyclones which can be accounted for by the WKBJ analysis must be due to the general nature of the basic state, and the strength of Rossby wave – gravity wave interactions on it, and any distinction not accounted for must be due to the specific effects of the inertial cut-off frequency.

3.2.5 The limit of large mode number

We start by writing (3.12,3.13) as a single second order ordinary differential equation for v_θ:

$$\frac{d^2 v_\theta}{dr^2} + F(r)\frac{dv_\theta}{dr} + G(r)v_\theta = 0, \tag{3.67}$$

where

$$F = -\frac{C'}{C} - A - D; \qquad G = \frac{C'D}{C} + AD - BC - D', \tag{3.68}$$

where here A, B, C and D are given by (3.14-3.17). Equation (3.67) must be solved subject to certain boundary conditions. These are regularity at $r = 0$, and an evanescence/radiation condition as $r \to \infty$. The latter of these conditions requires some care. Any unstable mode must be exponentially weak at infinity, and therefore an evanescence condition should suffice. However, we shall for convenience state the radiation condition, bearing in mind that when the eigenfrequency becomes slightly complex, perhaps at some high order in the expansion parameter m, the solution of (3.67) acquires some weak exponential decay.

Following standard WKBJ analysis, we write

$$v_\theta = A(r)e^{im\psi(r)}, \tag{3.69}$$

where now $A(r)$ is an amplitude function and $\psi(r)$ is a phase function, made rapidly varying through the factor of m in the exponent. Substituting (3.69) into (3.67) gives

$$A'' + 2im\psi'A' + im\psi''A - m^2\psi'^2 A + F(r)(A' + im\psi' A) + G(r)A = 0. \tag{3.70}$$

In general ψ and A must be expanded in inverse powers of the WKBJ expansion parameter m. The decomposition of v_θ into $A(r)$ and $\psi(r)$ is not unique. A small ($O(m^{-1})$, say), correction to v_θ through ψ could just as easily be made by an $O(m^{-1})$ correction to A. However, the decomposition can be made unique, and a great deal more convenient to work with, if we insist

that $A(r)$ and $\psi(r)$ naturally take the roles of amplitude and phase in that, if $\{A(r), \psi(r)\}$ is a solution of (3.70), then so is $\{A(r), -\psi(r)\}$.

Equation (3.70) can then be separated into two equations:

$$A'' - m^2\psi'^2 A + F(r)A' + G(r)A = 0 \tag{3.71}$$

$$2\psi' A' + \psi'' A + \psi' A F(r) = 0, \tag{3.72}$$

which latter equation can be integrated at once to

$$2\ln|A| + \ln|\psi'| + \int F = 0. \tag{3.73}$$

We now start our asymptotic analysis of (3.71). Since we know from (3.19) that ω is of order m, at least in the limit $F \ll 1$, it is convenient to work with an intrinsic angular frequency c defined by $c \equiv \omega/m$. Bearing Kelvin's expression $\omega = (m-1)/2$ in mind, and in the spirit of asymptotic expansions, we must expand c in the form

$$c = c_0 + m^{-1}c_1 + \cdots. \tag{3.74}$$

In fact, it will turn out that c has not only an algebraic expansion in m^{-1}, but also requires terms exponentially small in m to describe the instability.

From (3.14-3.17) we know that, expanding in m, $F(r)$ is $O(1)$, while $G(r)$ takes the form

$$G(r) = m^2\left[\frac{(-c + \overline{v}/r)^2 F^2}{1 + F^2\overline{h}} - \frac{1}{r^2}\right] + O(1), \tag{3.75}$$

which can, by virtue of the expansion (3.74) have terms of order $O(m^2)$ and $O(m)$. From (3.71), at $O(m^2)$, the equation for ψ_0 is

$$\psi_0'^2 = \frac{(-c_0 + \overline{v}/r)^2 F^2}{1 + F^2\overline{h}} - \frac{1}{r^2}. \tag{3.76}$$

The expression for A_0 can be obtained from (3.73). Notice that, since $F(r)$ is regular for all $r \neq 0$, there is a singularity in the expression for A_0 at a radius $r = r_c$ where $\psi_0'(r_c) = 0$. We can see at once that, for $c_0 \neq 0$, there is at least one such radius, since $\psi_0'^2 \sim -r^{-2} < 0$ as $r \to 0$, but $\psi_0'^2 \sim c_0^2 F^2 > 0$ as $r \to \infty$.

Now, $v_\theta = A(r)e^{im(\psi_0 + m^{-1}\psi_1 + \cdots)}$, and since we will require an expression for both the amplitude and phase of v_θ, we must obtain expressions for both ψ_0 and ψ_1. The equation for

ψ_1 is obtained from (3.71) at $O(m)$, and is

$$\psi_0'\psi_1' + \frac{F^2 c_1(-c_0 + \overline{v}/r)}{1 + F^2\overline{h}} = 0. \tag{3.77}$$

Now, the nature of the solutions for v_θ differs significantly on either side of $r = r_c$. On one side, and in particular in the limit $r \to \infty$, the solutions are oscillatory (for real c_0), whereas on the other side, and in particular in the limit $r \to 0$, they are growing and decaying exponentials. We refer to the radius $r = r_c$ as the "turning radius", and in the following analysis we shall assume that there is precisely one turning radius for the vortex. There is always precisely one turning radius for $F \ll 1$, whatever the magnitude of the Coriolis parameter f, and hence the assumption includes the basic state in the "quasigeostrophic" limit, in which one might expect interactions between vortices and gravity waves to be weakest. All vortices investigated numerically in the previous section were found to have precisely one turning radius.

The expansion (3.69) for v_θ is not valid in the neighbourhood of $r = r_c$, since ψ_0' is zero there, and so by (3.73) A would be unbounded. We must therefore rescale the equations in the neighbourhood of $r = r_c$ and derive equations for the perturbation expansion in this inner asymptotic region. In the vicinity of $r = r_c$, we have

$$G(r) = m^2 a_1 (r - r_c) + \cdots, \tag{3.78}$$

where a_1 is a positive constant. Rescaling the radial coordinate as $x = m^{2/3}(r - r_c)$ the leading order equation for v_θ from (3.67) becomes

$$\frac{d^2 v_\theta}{dx^2} + a_1 x v_\theta = 0. \tag{3.79}$$

Equation (3.79) has solutions $v_\theta = A_i(-a_1^{1/3}x), B_i(-a_1^{1/3}x)$, where $A_i(z)$ and $B_i(z)$ are Airy functions (Abramowitz & Stegun, 1965).

Four asymptotic regions have now been established, and the method of matched asymptotic expansions will now be used to relate the solutions in the different regions. The asymptotic regions are

1 : $r > r_c$

2 : $r - r_c = O(m^{-2/3})$

3 : $1 < r < r_c$

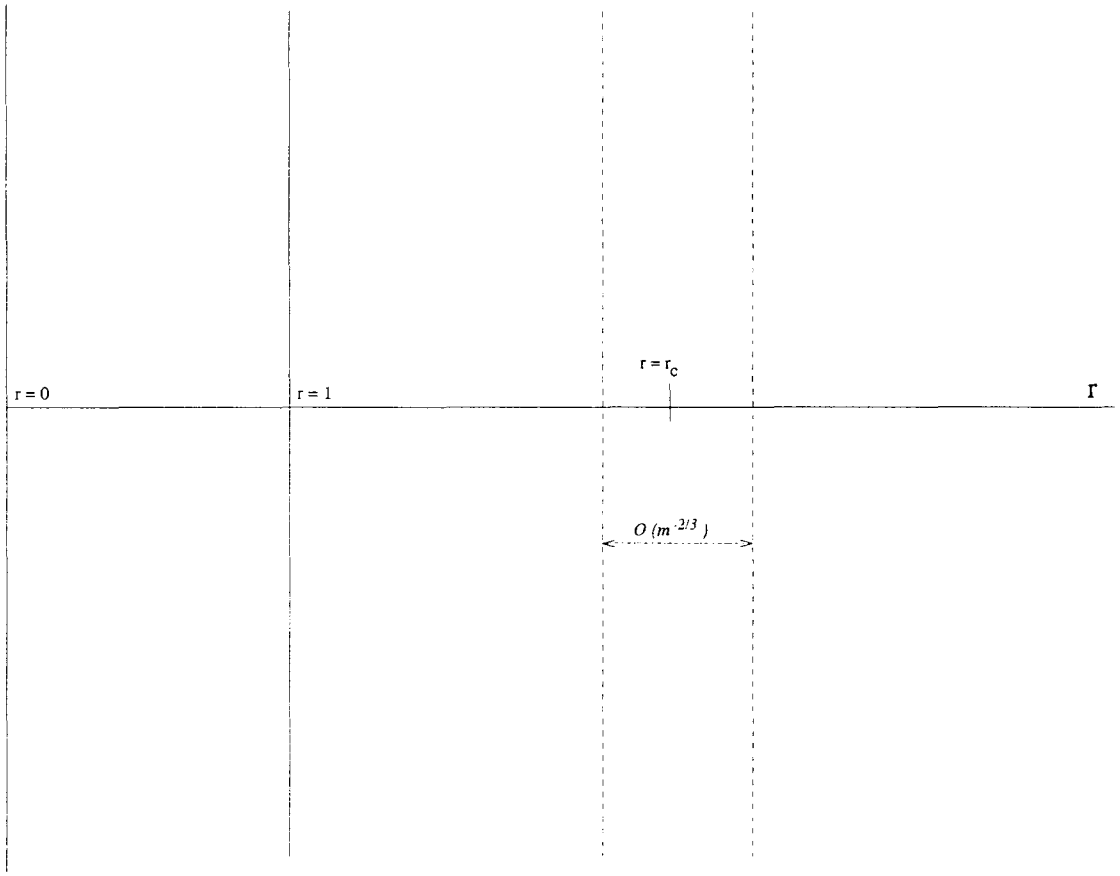


Figure 3.25: Asymptotic regions of shallow water vortex instability analysis.

4 : $0 < r < 1$

The situation is depicted in figure 3.25.

Region 1: $r > r_c$

As $r \to \infty$, we can place conditions on the branch of square root to take in the solution of (3.76) by imposing a radiation boundary condition, so that $\psi_0' > 0$ as $r \to \infty$. This immediately implies, through the matching condition in the limit $x \to \infty$, the ratio of A_i to B_i to take in the solution of (3.79). The expansions of $A_i(z)$ and $B_i(z)$ in the limit $z \to \infty$ then gives, through the matching conditions, the ratio of the growing to decaying exponential solutions to (3.67) for $1 < r < r_c$.

Region 2: $r - r_c = O(m^{-2/3})$

For $z \to -\infty$, the Airy functions take the asymptotic forms

$$A_i(-z) \sim \pi^{-1/2} z^{-1/4} \sin(2/3 z^{3/2} + \pi/4) \tag{3.80}$$

$$B_i(-z) \sim \pi^{-1/2} z^{-1/4} \cos(2/3 z^{3/2} + \pi/4). \tag{3.81}$$

Therefore, if we write $v_\theta = \alpha A_i(-a_1^{1/3} x) + \beta B_i(-a_1^{1/3} x)$, matching to a radiating field as $x \to \infty$ implies $\alpha/\beta = i$.

In the limit $x \to -\infty$ we have

$$A_i(z) \sim \frac{1}{2} \pi^{-1/2} z^{-1/4} e^{-2/3 z^{3/2}} \tag{3.82}$$

$$B_i(z) \sim \pi^{-1/2} z^{-1/4} e^{2/3 z^{3/2}}. \tag{3.83}$$

Region 3: $1 < r < r_c$

Now if we let

$$\Psi_0(r) = -\int_r^{r_c} \left(\frac{1}{r^2} - \frac{(-c_0 + \overline{v}/r)^2 F^2}{1 + F^2 \overline{h}} \right)^{1/2} dr \tag{3.84}$$

and

$$\Psi_1(r) = \int_r^{r_c} \frac{c_1(-c_0 + \overline{v}/r)}{\Psi_0'(1 + F^2 \overline{h})} dr \tag{3.85}$$

then, for $1 < r < r_c$, we have

$$v_\theta = A(r) \left(a e^{m\Psi_0 + \Psi_1} + b e^{-m\Psi_0 - \Psi_1} \right), \tag{3.86}$$

where the ratio of b to a is to be determined by the matching conditions to the Airy function region. Now the $B_i(-a_1^{1/3} x)$ term is, by (3.83), exponentially growing away from $r = r_c$ - i.e. it is exponentially decaying with radius in $1 < r < r_c$. It therefore matches to the term $be^{-m\Psi_0 - \Psi_1}$ in the solution in the region $1 < r < r_c$. Correspondingly, the $A_i(-a_1^{1/3} x)$ term matches to the solution which is exponentially growing with radius – that is, the term $ae^{m\Psi_0 + \Psi_1}$. Hence by considering (3.82-3.83) in the limit $z \to \infty$, the ratio of exponentially growing to exponentially decaying terms in the region $1 < r < r_c$ satisfies the relation $a/b = i/2$. Up to an arbitrary amplitude, therefore, this completes the solution in the range $r > 1$.

Region 4: $0 < r < 1$

In the range $r < 1$, the crucial boundary condition is regularity at $r = 0$. This implies only solutions of the form

$$v_\theta = de^{m\Psi_0 + \Psi_1} \tag{3.87}$$

are possible, where Ψ_0 is given by (3.84), and Ψ_1 by (3.85).

Jump conditions at $r = 1$

The final task, which determines the eigenvalue c at successive orders, is to apply continuity of v_r and h at $r = 1$, which is expressed in the continuity condition (3.18).

To find v_r, we may use to sufficient accuracy

$$\frac{dv_\theta}{dr} \sim \frac{im}{r} v_r. \tag{3.88}$$

For $r < 1$, this leads to

$$\left.\frac{v_\theta}{v_r}\right|_{r=1^-} = i + O(m^{-1}). \tag{3.89}$$

On the other hand, for $r > 1$, we obtain

$$\left.\frac{v_\theta}{v_r} = i\frac{e^{-m\Psi(1)} + i/2e^{m\Psi(1)}}{-e^{-m\Psi(1)} + i/2e^{m\Psi(1)}}\right|_{r=1+}, \tag{3.90}$$

where here

$$\Psi(1) = \Psi_0(1) + m^{-1}\Psi_1(1). \tag{3.91}$$

Recalling that $\Psi < 0$, this takes the asymptotic form for large m:

$$\frac{v_\theta}{v_r} \sim \frac{i}{r\Psi_0'}\left(1 + ie^{2m\Psi(1)}\right). \tag{3.92}$$

There are algebraic corrections to these terms, but because of our judicious decomposition of A and ψ, the expression (3.92) expresses the principal exponential terms. In particular, then, it has the correct expression for the leading order (exponentially small) imaginary part of v_θ/v_r.

Returning to the continuity condition (3.18), at order $O(m)$, we simply have

$$c_0 = \overline{v}(1). \tag{3.93}$$

At the next order, we obtain

$$c_1 = -\frac{1}{2}H_1\left(1 - \frac{1}{2}ie^{2m\Psi(1)}\right), \tag{3.94}$$

where $H_1 = 1 + F^2\overline{h}(1)$.

There are two things to note about (3.94). Firstly, the correction to the real phase speed of the disturbance is negative, and is in accordance with Kelvin's expression (3.19) in the limit $F \ll 1$. This is consistent with the claim that the disturbance on the boundary of the vortex has locally the character of a Rossby wave, with a pseudo-westward phase speed. The second thing is that the imaginary correction to c is always positive, corresponding to a temporally growing disturbance of form e^{-imct} at any value of F or f. Since the asymptotic limit in this case is of a particular type of disturbance, rather than a particular type of basic state, as it was in the $F \ll 1$ analysis of chapter 2, we conclude that all basic states which satisfy the assumption of a single turning point are unstable, at least to large mode number disturbances.

The low Froude number limit

We can obtain an independent check on the expression (3.94) in the limit of low Froude number by expanding it for $F \ll 1$ and comparing the result with the expression (3.51) expanded in the limit of large m.

We start by expanding the expression (3.84) for $\Psi_0(1)$ in the limit $F \ll 1$. Firstly, we recall that $\overline{v}$ is of order $O(F)$ everywhere except where r is of order $O(1)$. Therefore, to leading order in F, and recalling that $c_0 = 1/2$ at leading order in F, we have that $r_c = 2F^{-1}$ at leading order in F.

Now the integrand in (3.84) takes the asymptotic forms

$$\left(\frac{1}{r^2} - \frac{(-c_0 + \overline{v}/r)^2F^2}{1 + F^2\overline{h}}\right)^{1/2} \sim \begin{cases} \frac{1}{r} + O(F) & \text{for } r = O(1) \\ F\left(\frac{1}{F^2r^2} - \frac{1}{4}\right)^{1/2} + O(F^2) & \text{for } r = O(F^{-1}) \end{cases}. \tag{3.95}$$

The technique is to split up the range of integration into two sub domains: $[1, F^{-1/2}]$ and $[F^{-1/2}, 2F^{-1}]$. In the latter of these two domains we use a rescaled variable $R \equiv Fr$. The leading order expansion of $\Psi_0(1)$ is then

$$\Psi_0(1) \sim \int_2^{F^{1/2}} \left(\frac{1}{R^2} - \frac{1}{4}\right)^{1/2} dR + \int_{F^{-1/2}}^1 \frac{1}{r}dr. \tag{3.96}$$

Integrating, substituting, and expanding for small F, we obtain

$$\Psi_0(1) \sim 1 + \ln F - \ln 4 + O(F). \tag{3.97}$$

The asymptotic evaluation of (3.85) for $\Psi_1(1)$ is simpler, because the integrand has a regular expansion in F. The leading contribution comes from the range $r = O(F^{-1})$, and $\Psi_1(1)$ takes the form

$$\Psi_1(1) \sim \frac{1}{4}\int_2^0 \left(\frac{1}{R^2} - \frac{1}{4}\right)^{-1/2} = -1. \tag{3.98}$$

It follows from (3.94) that the growth rate of the instability is

$$\frac{1}{4}e^{2m\Psi(1)} = \frac{1}{4}\left(\frac{F}{4}\right)^{2m} e^{2(m-1)}. \tag{3.99}$$

To compare the expression (3.99) with the predictions of the analysis for $F \ll 1$, we expand $\mathrm{Im}(\omega)$ for large m, where $\mathrm{Im}(\omega)$ is given by (3.51). For large m we have

$$\mathrm{Im}(\omega) \sim \frac{\pi F^{2m}}{2m((m-1)!)^2}\frac{1}{4^{2m}}(m-1)^{2m}. \tag{3.100}$$

Using Stirling's formula for $(m-1)!$:

$$(m-1)! = (2\pi)^{1/2}(m-1)^{m-1/2}e^{-(m-1)}, \tag{3.101}$$

we have

$$\mathrm{Im}(\omega) \sim \frac{1}{4}\frac{e^{2(m-1)}}{4^{2m}}F^{2m}, \tag{3.102}$$

and hence the low Froude number and large mode number analyses agree in the limit in which they are both valid.

Comparison with numerical eigenvalue calculations

In addition to the analysis in the low Froude number limit, we can also test the validity of the general expression (3.94) given for the growth rate of the instability in the WKBJ limit against the numerical eigenvalue calculations.

Growth rates of the instability against mode number are shown in figures 3.26 & 3.27.

The basic states chosen all have Froude number $F = 2.5$ in figure 3.26, and $F = 5$ in figure 3.27, with varying values of f indicated on the graphs. Overall there seems to be good

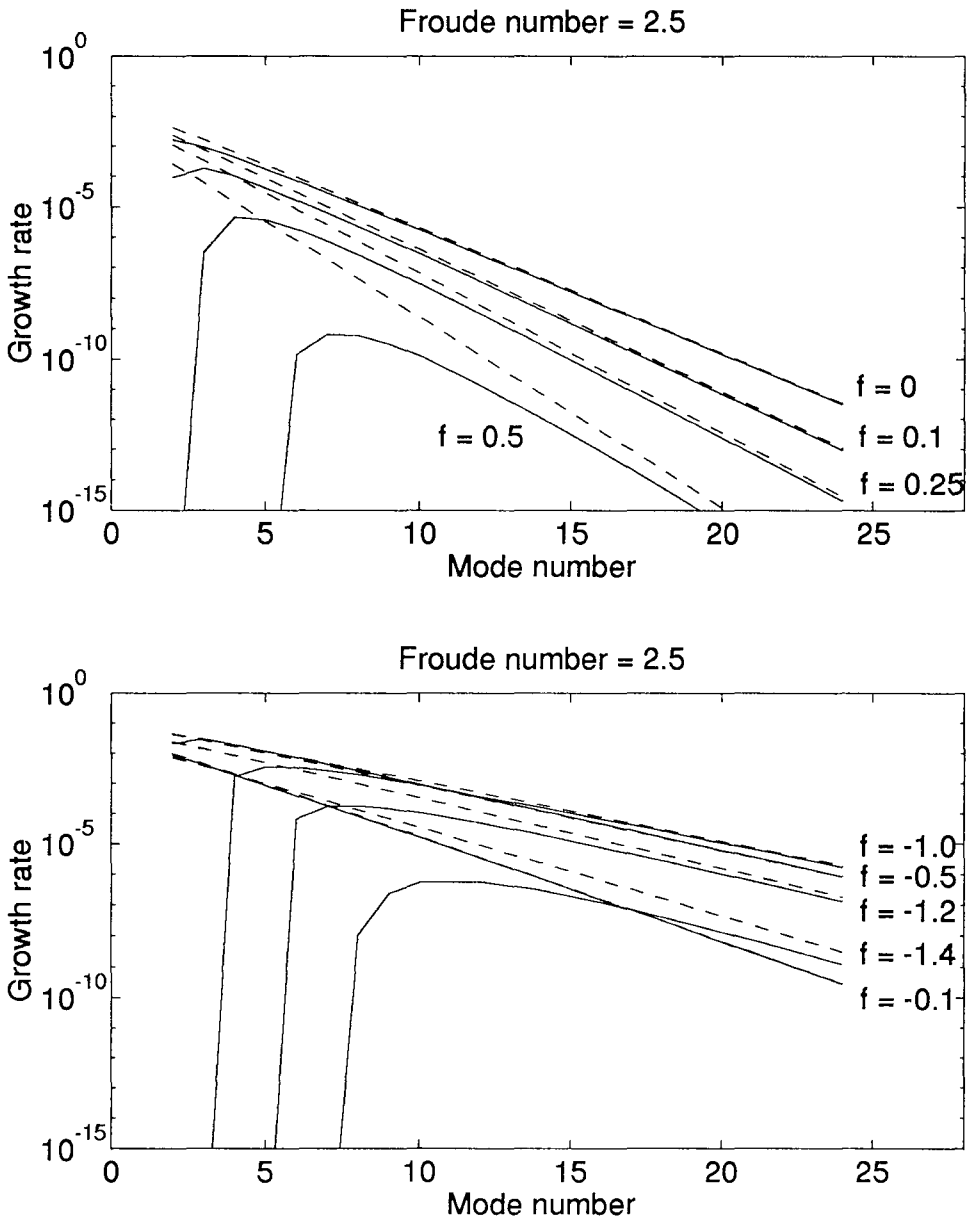


Figure 3.26: Comparison of eigenmode growth rates (solid lines) with WKBJ predictions (dashed lines), $F = 2.5$, for $m = 0$ to 25, at various values of f.

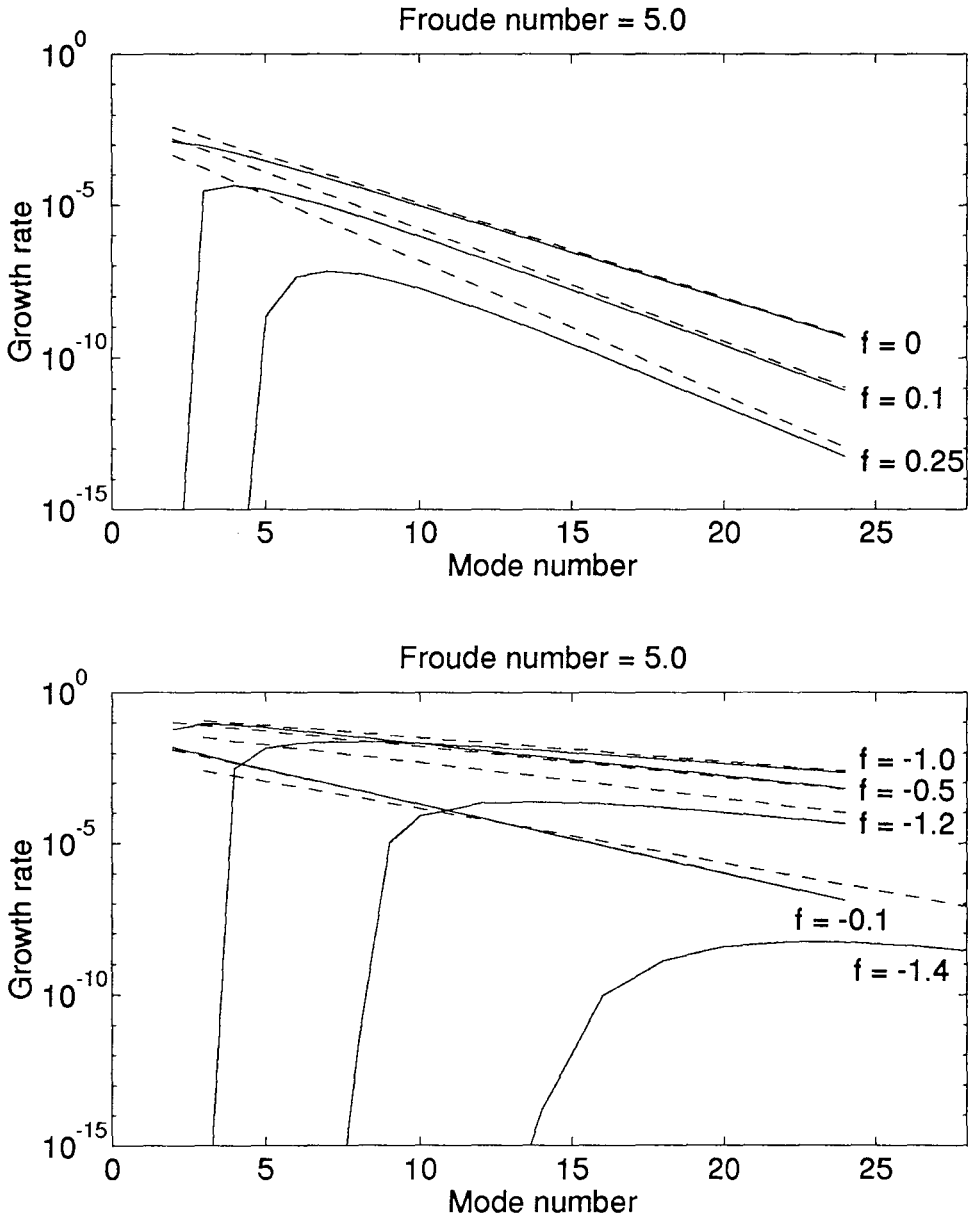


Figure 3.27: Comparison of eigenmode growth rates (solid lines) with WKBJ predictions (dashed lines), $F = 5.0$, for $m = 0$ to 25.

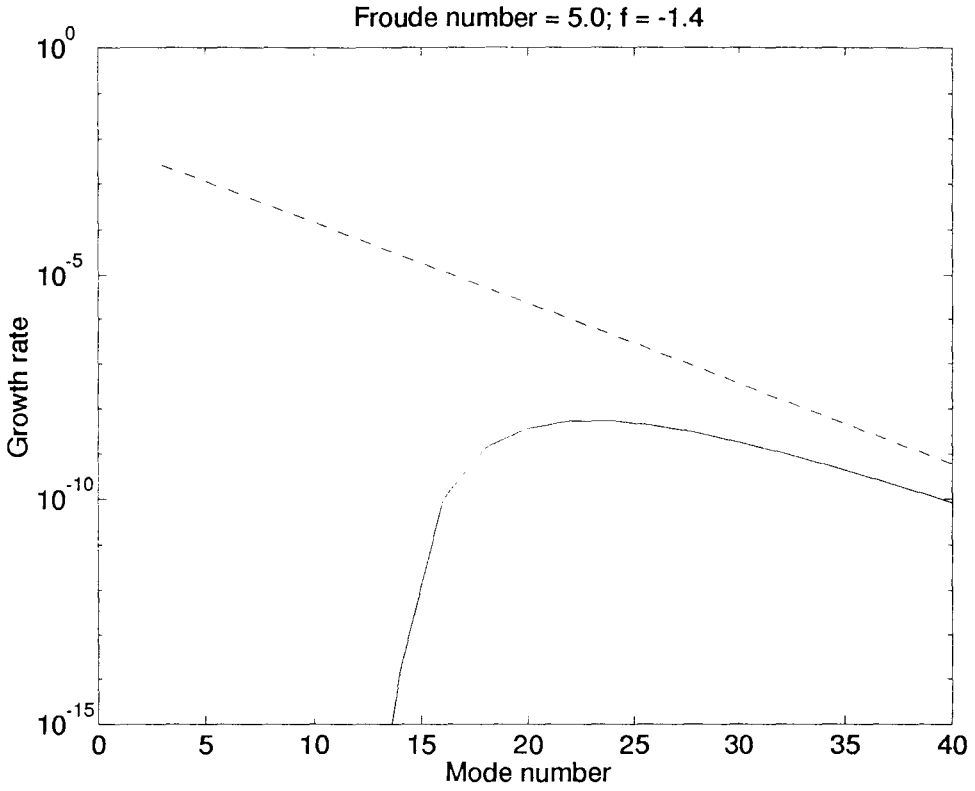


Figure 3.28: Comparison of eigenmode growth rate (solid line) with WKBJ prediction (dashed line) for $F = 5.0$ and $f = -1.4$, extended to $m = 40$.

agreement between the WKBJ analysis (dashed lines) and the numerical eigenvalue code (solid lines), even for quite moderate values of m. The case $F = 5.$ $f = -1.4$ is extended to $m = 40$ in figure 3.28 to demonstrate agreement in this case.

A significant distinction is to be drawn between $f > 0$ and $f < 0$. For given F, the magnitudes of $\Psi_0(1)$ and $\Psi_1(1)$ are always greater when $f > 0$ than when $f < 0$. It follows immediately that the instability will be weaker for $f > 0$ in the limit of large m. This is significant, in that it is dependent entirely upon the nature of the basic state, and independent of the effects of inertial cut-off, which are assumed not to apply, since the WKBJ analysis assumes that f is of order unity, whereas ω is of order $m \gg 1$. Where the numerical eigenvalues are in good qualitative agreement with the WKBJ prediction, any differences between them for different values of f must be directly attributable to the nature of the basic state alone, and independent of the effects of the inertial cut-off.

For small m and small $|f|$, the eigenmodes are unstable, and their growth rates are broadly

in line with the values predicted by the WKBJ analysis. However, as $|f|$ is increased, the eigenmodes become stable at small m. Since this effect is not predicted by the WKBJ analysis, it must be due to the fact that the eigenfrequencies are comparable with the magnitudes of the inertial frequency, and hence the inertial cut-off is significant.

As a heuristic way of thinking about the behaviour of the growth rates, I propose that in the region where they are seen to be increasing with increasing m, it is because the real part of the eigenvalue is increasing. The growth rate is therefore dominated by effects of the inertial cut-off. Examination of figures 3.26 & 3.27 shows that, once the instability has set in, this range of m tends to be quite small. However, in the range where the growth rate of the instability decreases with m, I propose that the WKBJ analysis is broadly valid in this limit, and the dependence of the growth rate of the instability on f is determined largely by the nature of induced basic state, and not directly by the effect of the inertial cut-off. Even in the range where the inertial cut-off is dominating the lower eigenmodes, the nature of the basic state is clearly important, in that it sets the real part of the eigenfrequency, and hence determines whether the eigenmodes are unstable or not.

In conclusion, the axisymmetric vortex is always unstable, but as F and $|f|$ are increased, larger and larger values of m must be taken to obtain the instability. Almost all the asymmetry between cyclones and anticyclones can be explained in terms of the basic state. At large m, the nature of the basic state determines the strength of Rossby wave – gravity wave interactions, and at small m it determines whether and by how much the real part of the eigenfrequency will exceed $|f|$, and hence how significantly the instability is affected by the inertial cut-off.

3.2.6 The Sozou modes

The analysis of the previous sections is distinguished by the fact that only one eigenmode has been found, whereas Sozou (1987) found an infinite set of modes in the Rankine vortex at non-zero Mach number.

The purpose of this section is to explain the difference between Sozou's results, both analytical and numerical, and the present analysis. It turns out that the difference lies in the fact that the Rankine vortex has uniform vorticity and non-uniform potential vorticity for all $M \neq 0$, whereas in the analysis of the previous section, we have assumed the potential vorticity to be piecewise constant.

We start by obtaining the potential vorticity of the Rankine vortex. The Rankine vortex has uniform vorticity, which we shall set to unity. Then

$$\overline{v}(r) = \begin{cases} \frac{1}{2}r & r < 1 \\ \frac{1}{2r} & r > 1 \end{cases} . \tag{3.103}$$

Using $d\overline{h}/dr = \overline{v}^2/r$, and applying continuity of $\overline{h}$ at $r = 1$, we have

$$\overline{h}(r) = \begin{cases} -\frac{1}{4} + \frac{1}{8}r^2 & r < 1 \\ -\frac{1}{8}r^{-2} & r > 1 \end{cases} . \tag{3.104}$$

Hence the potential vorticity and its gradient in the radial direction are given respectively by

$$Q = \frac{1}{1 - \frac{1}{8}F^2(2 - r^2)}; \quad \frac{dQ}{dr} = -\frac{F^2 r}{\left[2 - \frac{1}{4}F^2(2 - r^2)\right]^2}. \tag{3.105}$$

From this we see that $dQ/dr < 0$ for all $r < 1$, and the potential vorticity is zero for $r > 1$. Hence the potential vorticity gradient is monotonic, and the vortex is not unstable to a barotropic Rossby wave – Rossby wave type of instability. However, the non-zero potential vorticity gradient means that, in addition to the Rossby wave on the boundary of the vortex, it is capable of supporting an infinite set of Rossby wave modes within the vortex.

We now consider the limit $m \gg 1$. Returning to the eigenvalue equations (3.12 - 3.17), if additional modes are to be introduced in this large m limit as a result of the non-zero gradient of potential vorticity, it can only be because $(H/\sigma)dQ/dr \sim O(m)$ over a range of order unity within the vortex. Only then is the balance of terms in the equation for ψ_0 changed at leading order in m.

In fact, the uniform vorticity of the Rankine vortex is significant, as well as its non-uniform potential vorticity. Within the vortex, $\sigma = -\omega + m$, i.e. σ is constant within the vortex, and there are no critical layers there (unless $\sigma = 0$ throughout the vortex). In this case we may take $\sigma = m + m^{-1}c_2$, where c_2 is constant, so that σ is of order m^{-1} throughout the vortex. It turns out that the equation for v_θ is singular in the general cases to be described below, and it is more convenient to work with an equation for v_r. Writing $v_r(r) = A(r)e^{im\psi(r)}$, the equation for ψ_0 then becomes

$$\psi_0'^2 = -\left(\frac{1}{r^2} + \frac{F^2}{4c_2 H(r)}\right), \tag{3.106}$$

where $H(r) = 1 + F^2\overline{h}(r)$. In this case, it is convenient to note that $c_1 = 0$, and hence $\psi_1 = 0$.

Now, we consider the continuity condition (3.18). At order unity, it is

$$-\frac{c_2}{m}\left[\frac{v_\theta}{v_r}\right] = [iH(1)Q], \tag{3.107}$$

and therefore we require v_θ/v_r to be of order m at $r = 1^-$. Since $v_\theta \sim (ir/m)dv_r/dr$, this can only be achieved if $v_r = O(m^{-1})$ at $r = 1$.

We must now consider two possible cases: $c_2 > 0$ and $c_2 < 0$. If $c_2 > 0$, we can write down the solution for v_θ over the range $0 < r < 1$ as

$$v_r = A(r)\exp\left[m\int_0^r \left(\frac{1}{r^2} + \frac{F^2}{4c_2H(r)}\right)^{1/2} dr\right]. \tag{3.108}$$

In this case, v_r is always of order 1, and hence it will not be possible to satisfy the continuity conditions at $r = 1$.

Therefore, we seek solutions with $c_2 < 0$. For convenience, we introduce a new constant κ, and let $c_2 = -4F^2\kappa^{-2}$. Then

$$\psi_0'^2 = \left(\frac{\kappa^2}{H(r)} - \frac{1}{r^2}\right). \tag{3.109}$$

There is now a turning point at r_c, given by $H(r_c)/r_c^2 = \kappa^2$ and, provided κ is sufficiently large, this turning point will occur at some $r_c < 1$. It is simple to show from the form $H(r)$ that the turning point is unique for a given κ. The solution in $0 < r < 1$ is now divided into three asymptotic regions

$r < r_c$ Exponentially growing solution

$r - r_c = O(m^{-2/3})$ Airy functions

$r_c < r < 1$ Oscillatory functions

For $0 < r < r_c$, we will have

$$v_r = A(r)\exp\left[m\left(\frac{1}{r^2} - \frac{\kappa^2}{H(r)}\right)^{1/2} dr\right]. \tag{3.110}$$

In the region $r - r_c = O(m^{-2/3})$, the solution will be represented in terms of Airy functions. What is important for the eigenvalue equation is the solution in $r > r_c$. Now, if the solution

in $r < r_c$ consists of only an exponential term which is growing in r, then the solution in the neighbourhood of $r = r_c$ will be represented by $A_i(z)$, where z is proportional to $r^{2/3}$. Then in $r > r_c$, matching to the Airy function gives a solution for v_r:

$$v_r = A(r)\cos\left[m\int_{r_c}^{r}\left(\frac{\kappa^2}{H(r)} - \frac{1}{r^2}\right)^{1/2} dr - \frac{\pi}{4}\right]. \tag{3.111}$$

If we now wish to have $v_r = 0$ at leading order in m then,

$$m\int_{r_c}^{r}\left(\frac{\kappa^2}{H(r)} - \frac{1}{r^2}\right)^{1/2} dr = \left(n - \frac{1}{4}\right)\pi; \quad n = 1, 2, 3, \ldots. \tag{3.112}$$

This equation is the leading order eigenvalue equation for κ at arbitrary Froude number F. To make further progress, and to compare this result with that obtained by Sozou (1987), we shall make the assumption $F \ll 1$, but $\kappa = O(1)$. The only approximation to make is that $H(r) = 1$ in (3.112). Then we obtain a dispersion relation for κ

$$\sqrt{\kappa^2 - 1} - \arccos\kappa^{-1} = m^{-1}\left(n - \frac{1}{4}\right)\pi. \tag{3.113}$$

Sozou works with a frequency divided by the vorticity in the vortex $f = m - 2m\omega^2/s^2$, where ω is the vorticity in the vortex. After rescaling Sozou's (1987) equations, one can show that his variable s should correspond to our variable $m\kappa$. Sozou claims that roots of his dispersion relation satisfy $J_m(s) = 0$. It is more convenient to work with κ than s, so we solve $J_m(m\kappa) = 0$ for large m and κ of order unity.

From Abramowitz & Stegun (1965),

$$J_m(m\sec\beta) \sim \sqrt{2/(\pi m\tan\beta)}\cos(m\tan\beta - m\beta - \pi/4). \tag{3.114}$$

Identifying $\sec\beta \equiv \kappa$ gives $\tan\beta \equiv \sqrt{\kappa^2 - 1}$, and hence Sozou's condition is

$$m\sqrt{\kappa^2 - 1} - m\arccos\kappa^{-1} - \pi/4 = (n - 1/2)\pi, \tag{3.115}$$

which is the same as equation (3.113).

The important point is that the additional roots found by Sozou are only possible at large m because of the uniform vorticity and non-uniform potential vorticity in the Rankine vortex

at non-zero Mach number. The large mode number analysis reveals that there is a turning point within the vortex in these cases, implying that the eigenmodes have a wave-like structure for $r_c < r < 1$. The turning point is possible only because $\sigma = O(m^{-1})$ and $dQ/dr \neq 0$ throughout the vortex region, neither of which can be true for a vortex with uniform potential vorticity, except in the singular limit $F = 0$. The conclusion is that they are associated with Rossby waves propagating on the potential vorticity gradient within the Rankine vortex. The fact that $c_2 < 0$ once again reminds us that they have a retrograde or pseudo-westward phase progression, consistent with their being Rossby-wave like.

An expression for the growth rate of the instability can also be obtained, by considering the leading order real terms in (3.18). Outside the vortex, the equations for the disturbances are the same as before, except that now c_1 in (3.74) is zero, and hence $\Psi_1 = 0$. Outside the vortex, therefore, the first real terms in v_θ/v_r are $\epsilon^{2m\Psi_0(1)}$, where Ψ_0 is given by (3.84).

Now, at $r = 1^+$, $v_\theta/v_r = -i$, whereas at $r = 1^-$, $v_\theta/v_r = -imH_1/c_2 = imH_1\kappa^2/(4F^2)$. Therefore, the real part of (3.18) is

$$-c_i m^2 H_0/c_2 + \frac{c_2}{m}\epsilon^{2m\Psi_0(1)} = 0, \tag{3.116}$$

and hence we may write down an expression for the principal imaginary part of ω

$$\omega_i = c_2^2/(m^2 H_0)e^{2m\Psi_0(1)}. \tag{3.117}$$

3.3 Instability of a circular vortex with smooth potential vorticity

We turn now to showing that, under certain assumptions, the vortex instability found in the previous section can be generalized to vortices with a smooth (analytic) potential vorticity profile, provided the potential vorticity gradient is sufficiently steep. I show how to scale the potential vorticity gradient so that the instability threshold is found at a finite value of the gradient.

3.3.1 Disturbance equations

To begin, we consider how the behaviour of differential equations (3.9-3.11) are modified when there is a non-zero potential vorticity gradient in the basic state.

The equations for the eigenfunctions were stated in the previous section as

$$\frac{dv_r}{dr} = A(r)v_r + B(r)v_\theta \tag{3.118}$$

$$\frac{dv_\theta}{dr} = C(r)v_r + D(r)v_\theta, \tag{3.119}$$

where

$$A(r) = \frac{F^2\sigma Q r}{m} - \frac{d}{dr}\ln H\frac{1}{r} \tag{3.120}$$

$$B(r) = \frac{iF^2\sigma^2 r}{Hm} - \frac{im}{r} \tag{3.121}$$

$$C(r) = \frac{iF^2HQ^2r}{m} + \frac{im}{r} + \frac{iH}{\sigma}\frac{dQ}{dr} \tag{3.122}$$

$$D(r) = -\frac{F^2Q\sigma r}{m} - \frac{1}{r}, \tag{3.123}$$

and $H = 1 + F^2\overline{h}$. We now consider the order of terms in the case where $dQ/dr \neq 0$.

In general, σ is of order $O(m)$, and the principal behaviour of the eigenmodes will not be affected by the term $(H/\sigma)dQ/dr$ in $C(r)$, not present in the case of piecewise uniform potential vorticity. However, we know from the eigenmodes found above that $\sigma > 0$ at the vortex boundary, but tends to a negative constant value as $r \to \infty$. Therefore, there must exist at least one zero of σ in $r > 1$ and, if $\overline{v}$ approaches $r = 1$ faster than linearly in r, there will exist another zero of σ in $r < 1$. In particular, if f is large, then the basic state flow could be taken to be approximated by the quasigeostrophic equations, in which case the velocity field takes the form $I_0(F|f|r)$ for $r < 1$, and $K_0(F|f|r)$ for $r > 1$.

To derive a distinguished asymptotic scaling for the problem, we consider a limit in which the potential vorticity gradient at the boundary of the vortex, although finite, is very sharp. In this limit, we expect to recover the results of the previous section, in which the eigenmode was exponentially growing in r for $r < 1$, and exponentially decaying in r for $r > 1$. To achieve this transition in the form of the eigenmode there must be an asymptotic region in the neighbourhood of $r = 1$ in which the potential vorticity gradient enters the eigenfunction equations at the leading order in m. Across this region the nature of the solution will change from exponentially growing to exponentially decaying. There is only one way in which this can be achieved, and that is by bringing $(H/\sigma)dQ/dr$ into leading order balance in (3.119).

Now, to change the balance in (3.119), we shall require $(H/\sigma)dQ/dr$ to be of order m, in order to balance m/r. Furthermore, in the limit of a very sharp potential vorticity jump

(abbreviated to PV jump), we know that $\sigma = H(1)/2$ at the boundary of the vortex. Over a range of order δ, σ is of order $1 + m\delta$, and if Q varies by order unity across this region, its gradient is of order $1/\delta$.

It remains to show that the required scaling is $\delta \sim m^{-1}$. If we take $\delta \ll m^{-1}$, then $(H/\sigma)dQ/dr \sim 1/\delta$, which is required to balance m/r. This contradicts the assumption $\delta \ll m^{-1}$, and hence we require $\delta \geq O(m^{-1})$. On the other hand, if $\delta \gg m^{-1}$, then $(H/\sigma)dQ/dr \sim 1/(m\delta^2)$. Requiring this to balance m/r leads to another contradiction, and hence we arrive at the scaling $\delta = O(m^{-1})$.

The situation now being considered is a potential vorticity jump of order one over a distance of order m^{-1}, with a gradient of order m. This is internally consistent – if Q' is of order m over a range of order m^{-1} then this will lead to a net PV jump of order one over this range. The amplitude of the eigenfunction $e^{\pm im\Psi}$ varies by order unity over this range, and hence $\delta \sim m^{-1}$ represents the distinguished scaling for the problem, in which the potential vorticity changes by order unity over the length scale of amplitude variations of order unity in the eigenfunctions.

Now recall that the vortex is of unit radius, so the region of sharp PV gradients is centred on $r = 1$. Introducing a new range variable $z = m(r - 1)$, and assuming $\sigma \sim O(1)$, we obtain an equation for v_r at leading order in m:

$$\frac{d^2 v_r}{dz^2} - \left(1 + \frac{H}{\sigma}\frac{dQ}{dz}\right) v_r = 0. \tag{3.124}$$

To make further progress, we may assume initially that the PV jump occupies a small region in z (of order ϵ). In fact, it turns out that this is sufficient to understand the problem in full. We shall assume as before that $Q(r = 1^-) = Q(r = 1^+) + 1$. To fix ideas, we shall assume that the PV across the vortex boundary takes a hyperbolic tangent profile. Then

$$\frac{dQ}{dz} = -\frac{1}{2\epsilon}\mathrm{sech}^2(z/\epsilon), \tag{3.125}$$

so that

$$Q = -\frac{1}{2}\tanh(z/\epsilon) + b, \tag{3.126}$$

where b is an arbitrary constant, required for matching to the outer flow. The equations for the velocity and height fields of the basic state in the asymptotic region $z = O(1)$ are

$$\frac{m}{F^2}\frac{dH}{dz} = f\overline{v} + \frac{\overline{v}^2}{r} \tag{3.127}$$

$$m\frac{d\overline{v}}{dz} = QH - f - \overline{v}(1 + z/m)^{-1}. \tag{3.128}$$

From (3.128) it follows that $\overline{v}(z) = V_0(\text{constant}) + O(m^{-1})$, and from (3.127) it follows that $H(z) = H_0(\text{constant}) + O(m^{-1})$. If we expand $\overline{v} = V_0 + m^{-1}V_1 + \cdots$, then from (3.128) we obtain an equation for V_1:

$$\frac{dV_1}{dz} = QH_0 - f - V_0, \tag{3.129}$$

with solution

$$V_1 = -\frac{1}{2}\epsilon H_0 \ln\cosh(z/\epsilon) + (bH_0 - f - V_0)z. \tag{3.130}$$

Without loss of generality, we may take the associated constant of integration to be zero in this case by absorbing it into V_0. The variable $\sigma = m(-c + V/r)$ now takes the form

$$\sigma = -c_1 - \frac{1}{2}\epsilon H_0 \ln\cosh(z/\epsilon) + (bH_0 - f - 2V_0)z + O(m^{-1}). \tag{3.131}$$

The requirement $\sigma = O(1)$ is met by setting $c_0 = V_0$. This is consistent with equation (3.94) the limit of a single discontinuity in potential vorticity.

We now have sufficient details of the basic state in the region $z = O(1)$ to enable us to analyse the nature of the unstable eigenmode.

We start with the inner region, $z = O(\epsilon)$, across which the PV varies by an amount of order $O(1)$. Introducing a further variable $\zeta = z/\epsilon$, and expanding $v_r = v_r^{(0)} + \epsilon v_r^{(1)} + \cdots$ gives

$$v_r^{(0)} = \text{constant} = 1 \quad \text{say}. \tag{3.132}$$

Then

$$\frac{d^2 v_r^{(1)}}{d\zeta^2} = \frac{H_0}{2c_1}\text{sech}^2\zeta. \tag{3.133}$$

Integrating for $v_r^{(1)}$ gives

$$v_r^{(1)} = \frac{H_0}{2c_1}\ln\cosh\zeta + k_1\zeta + k_2, \tag{3.134}$$

where k_1, k_2 and c_1 are found by matching to v_r in the region $z = O(1)$.

Now, for $z = O(1)$, σ takes the form

$$\sigma \sim -c_1 - \frac{H_0}{2}(|z| - \alpha z), \tag{3.135}$$

where $\alpha = 2(bH_0 - f - 2V_0)/H_0$. If $|\alpha| < 1$, then there are two roots of the equation $\sigma = 0$. As discussed above, this will be the generic case for low Rossby number flow, in which σ has a

zero on either side of the PV jump. Since the WKBJ ($m \gg 1$) analysis was motivated by the difficulty of finding unstable modes numerically in the low Rossby number limit, I shall assume $|\alpha| < 1$ for the remainder of the analysis of this section. The equation governing v_r is then

$$\frac{d^2 v_r}{dz^2} - \left(1 - \frac{1}{2\epsilon\sigma}\mathrm{sech}^2(z/\epsilon)\right) v_r = 0. \tag{3.136}$$

It follows that, to leading order of approximation,

$$\frac{d^2 v_r}{dz^2} - v_r = -\frac{4}{\epsilon}(-2c_1/H_0 - |z| + \alpha z)^{-1} e^{-z^2/\epsilon^2} v_r. \tag{3.137}$$

We retain the exponentially small term on the right hand side of (3.137) due to its possible singularities at $\sigma = 0$. We denote these positions by $z = z_-$ and $z = z_+$.

Analysis of these points shows that they are regular singular points, with indicial exponents 0 and 1. Moreover, writing the solution with the logarithmic term in the form

$$v_r = 1 + \delta(z - z_0)\log(z - z_0) + \cdots, \tag{3.138}$$

(where here z_0 stands for z_- or z_+) we can easily see that δ will be exponentially small in ϵ. It follows that the critical layers $\sigma = 0$ are insignificant to the eigenfunctions at all algebraic orders in ϵ, and the eigenfunctions and their derivatives will be continuous across $\sigma = 0$ at all algebraic orders in ϵ. Hence the leading order solutions are $v_r = e^{\pm z}$.

Now, we know from the analysis for the case of a sharp PV interface that we will require v_r to be exponentially decaying at leading order away from $r = 1$, and hence we take $v_r = e^{-|z|}$ as the leading order solutions in the regions $|z| = O(1)$. It follows from matching to the inner region $z = O(\epsilon)$ that $v_r^{(1)} \sim -|\zeta|$ as $|\zeta| \to \infty$, and this matching to (3.134) fixes $k_1 = k_2 = 0$, and

$$c_1 = -\frac{1}{2}H_0. \tag{3.139}$$

This is consistent with the expression for c_1 given by (3.94) in the analysis for the discontinuous potential vorticity. Hence,

$$\sigma = \frac{H_0}{2}(1 - |z| + \alpha z), \tag{3.140}$$

and the zeros of σ occur at two locations z_+ and z_-, given by

$$z_+ = \frac{1}{1-\alpha} > 0; \qquad z_- = -\frac{1}{1+\alpha} < 0. \tag{3.141}$$

We now turn to analysing the effect of the exponentially small δ in the neighbourhood of the critical layers. We start by analysing v_r in the neighbourhood of $z = z_-$. Writing $y = z - z_-$, and substituting (3.139) into (3.137), we have

$$\sigma = \frac{H_0}{2|z_-|} y \tag{3.142}$$

$$\frac{d^2 v_r}{dy^2} - \left(1 - \frac{4|z_-|e^{-z_-^2/\epsilon^2}}{\epsilon y}\right) v_r = 0. \tag{3.143}$$

Substitution of (3.138) with $z_0 \equiv z_-$ into (3.143) gives

$$\delta = -4|z_-|\epsilon^{-1} e^{-z_-^2/\epsilon^2}. \tag{3.144}$$

Now, if we are constructing an unstable mode, it follows that $\mathrm{Im}\sigma < 0$, and that therefore there is not really a singularity at $\sigma = 0$ at all. If σ has a small negative imaginary part as we pass through $y = 0$, it follows from (3.142) that y must also have a small negative imaginary part ($-i\mu$, say). Since $\ln(z - z_-)$ has branch points at the zeros of σ, the effect of this small negative imaginary part of σ is to displace the branch point of the logarithm from $y = 0$ to $y = i\mu$ for some small positive μ. As y is increased through zero along the real axis, the argument of y must therefore increase by π to be consistent with the branch cut of the logarithm. The situation is depicted in figure 3.29. Since $\ln y = \ln|y| + i\mathrm{Arg}y$, and since v_r takes the value $e^{-|z_-|}$ to leading order at $y = 0$, it follows that the leading order contribution to the phase shift of the solution as we pass through $y = 0$ is given by $i\pi\delta e^{-|z_-|} y + \cdots$. The solution satisfies (3.137), and therefore for $y > 0$, the leading order contribution to the imaginary part of the solution introduced by the branch cut is

$$v_r^i = -4\pi i|z_-|e^{-|z_-|}\epsilon^{-1} e^{-|z_-|^2/\epsilon^2} \sinh y. \tag{3.145}$$

In addition to an imaginary part of the solution due to the presence of critical layers in the flow, an imaginary contribution to v_r can also be introduced by having a small imaginary part of c, which we shall denote as ic_i. Near $z = 0$, in the region where $\zeta = O(1)$, $v_r^{(i)}$ is expanded in powers of ϵ:

$$v_r^i = v_0^i + \epsilon v_1^i + \cdots. \tag{3.146}$$

Then v_0^i is constant, and v_1^i satisfies

$$\frac{d^2 v_1^i}{d\zeta^2} = \frac{H_0 v_0^i}{2c_1}\mathrm{sech}^2\zeta - \frac{iH_0 c_i}{2c_1^2}\mathrm{sech}^2\zeta. \tag{3.147}$$

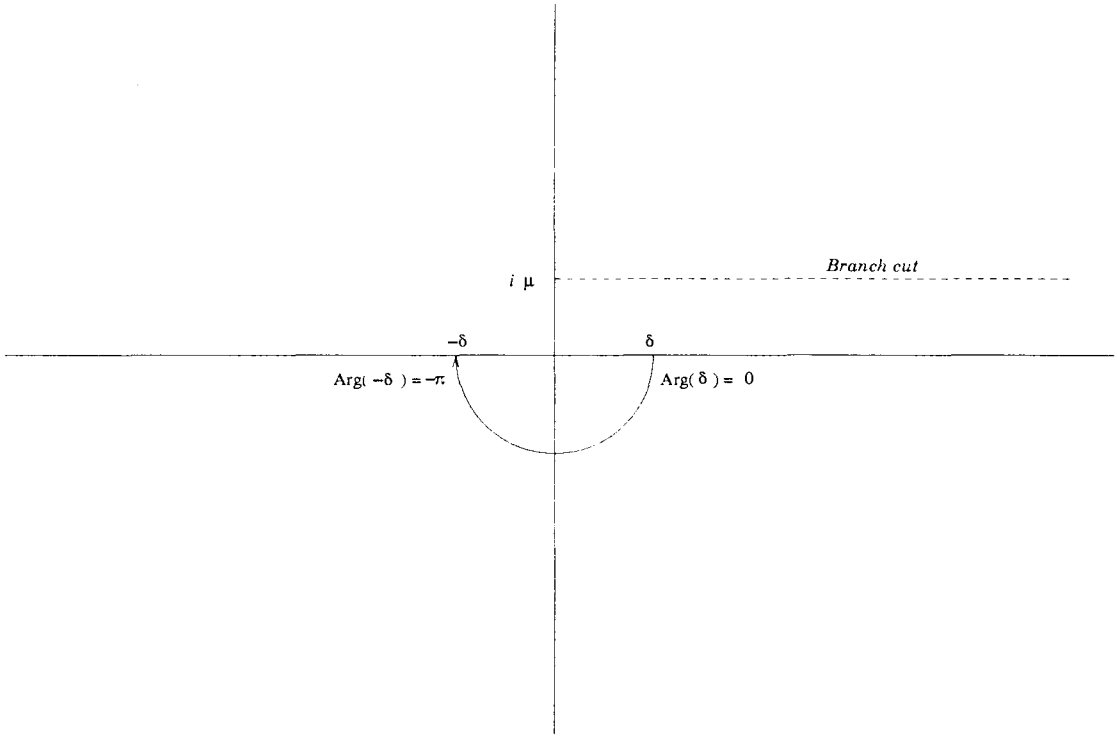


Figure 3.29: Schematic showing how argument changes in the neighbourhood of a branch point displaced slightly from the real axis.

This equation has solution

$$v_1^i = -\left(v_0^i + \frac{2ic_i}{H_0}\right) \ln\cosh\zeta + k_1\zeta + k_2. \tag{3.148}$$

As $\zeta \to -\infty$,

$$v_r^i \sim -4\pi i|z_-|e^{-|z_-|}\epsilon^{-1}e^{-|z_-|^2/\epsilon^2}(\sinh|z_-| + \epsilon\zeta\cosh|z_-| + O(\epsilon^2)). \tag{3.149}$$

Hence

$$v_0^i = -4\pi i|z_-|\epsilon^{-|z_-|}\sinh|z_-|\epsilon^{-1}e^{-|z_-|^2/\epsilon^2} \tag{3.150}$$

$$v_0^i + \frac{2ic_i}{H_0} + k_1 = -4\pi i|z_-|\epsilon^{-|z_-|}\cosh|z_-|\epsilon^{-1}e^{-|z_-|^2/\epsilon^2}. \tag{3.151}$$

Then, as $\zeta \to \infty$, we will have

$$v_r^i \sim v_0^i + \epsilon\left(-v_0^i - \frac{2ic_i}{H_0} + k_1\right)\zeta. \tag{3.152}$$

The solution in the region $z > 0$ is now written as

$$v_r^i = Ae^z + Be^{-z}. \tag{3.153}$$

Matching implies

$$A + B = v_0^i; \quad A - B = -v_0^i - \frac{2ic_i}{H_0} + k_1. \tag{3.154}$$

It is the exponentially growing part of this solution which will match to the exponentially growing (and imaginary) part of the solution in $r > 1$, introduced in the region $1 < r < r_c$ by the matching procedure across the sonic radius $r = r_c$. Of particular interest is therefore the value of A, which is given by

$$A = \frac{k_1}{2} - \frac{ic_i}{H_0} = -\frac{2ic_i}{H_0} - 2\pi i|z_-|\epsilon^{-2|z_-|}. \tag{3.155}$$

Finally, we must attend to the matching conditions across the second critical layer at $z = z_+$. In the same manner as the critical layer at $z = z_-$, this introduces a further small imaginary part to v_r as the leading order solution crosses the critical layer. The existing imaginary part is not changed at leading order. At this second critical layer we use the variable $y = z - z_+$. Then

$$\sigma = -\frac{H_0}{2|z_+|}y \tag{3.156}$$

$$\frac{d^2v_r}{dy^2} - \left(1 + \frac{4|z_+|e^{-|z_+|^2/\epsilon^2}}{\epsilon y}\right)v_r = 0. \tag{3.157}$$

Now recall that the singularity is to be removed by introducing a small negative imaginary part to σ. In the case $z \sim z_+$, we see from (3.156) that y therefore has a small positive imaginary part as we pass through $y = 0$ along the real line. This is the opposite sign to the case $z = z_-$, in which y was required to have a small negative imaginary part. In the present case the branch

point of the logarithm is therefore displaced to $y = -i\mu$, and the argument of y decreases by π as y is increased through $y = 0$. Hence to the imaginary part of the solution we add

$$-4\pi i|z_+|e^{-|z_+|}\epsilon^{-1}e^{-|z_+|^2/\epsilon^2}\sinh y. \tag{3.158}$$

The exponentially growing part of this solution is

$$-2\pi i|z_+|e^{-2|z_+|}\epsilon^{-1}e^{-|z_+|^2/\epsilon^2}e^z. \tag{3.159}$$

Hence the sum of the exponentially growing imaginary terms in the region $z > z_+$ will be

$$\left[-2\pi i\epsilon^{-1}\left(|z_+|\epsilon^{-2|z_+|}e^{-z_+^2/\epsilon^2}+|z_-|e^{-2|z_-|}e^{-z_-^2/\epsilon^2}\right)-\frac{2ic_i}{H_0}\right]e^z. \tag{3.160}$$

Exponentially growing real terms can be removed by a small change to the real part of c.

We now turn to matching this to the region $1 < r < r_c$. In this region, the expansion of the leading order term for $r \to 1^+$ gives an exponential function which is decaying with increasing r. However, as discussed in the analysis for the vortex with discontinuous PV, the radiation condition introduces an exponentially growing term, which is exponentially small (in m), and out of phase with the leading order exponentially decaying term. The ratio of growing to decaying terms in v_r as $r \to 1^+$ is $-(i/2)e^{2m|\Psi(1)|}$, where $\Psi(1)$ is given by (3.84, 3.85, 3.91). Since we took v_r to be one at leading order in this region, matching conditions in the limit $z \to \infty$ for the imaginary component of v_r from (3.160) imply

$$\frac{2\pi}{\epsilon}\left(|z_+|e^{-2|z_+|}\epsilon^{-z_+^2/\epsilon^2}+|z_-|e^{-2|z_-|}e^{-z_-^2/\epsilon^2}\right)+\frac{2ic_i}{H_0}=\frac{1}{2}e^{2m\Psi(1)}. \tag{3.161}$$

Now let $Z = \max(|z_+|, |z_-|)$. Then if

$$\epsilon = Z/(2|\Psi(1)|)^{1/2}m^{-1/2}, \tag{3.162}$$

with small corrections, then (3.161) is solved with $c_i = 0$, and the necessary imaginary part of v_r in the limit $z \to \infty$ exists without introducing an imaginary part to c. If ϵ is smaller than, but of the same order as, the value given by (3.162), then c_i will be of order $e^{2m|\Psi(1)|}$.

If $\epsilon \ll m^{-1/2}$, then the principal phase shift across the PV gradient is due to the imaginary part of c, and

$$c_i = \frac{1}{4}e^{2m\Psi(1)}. \tag{3.163}$$

Again, this is consistent with equation (3.94), derived in the limit of discontinuous potential vorticity.

Finally, if $\epsilon > Z/(2|\Psi(1)|)^{1/2}m^{-1/2}$, the equation (3.161) can only be satisfied with $c_i < 0$. This is a contradiction, since the analytic continuation used at the critical layers assumed $c_i > 0$. Consequently the eigenmode structure of the problem is lost if $\epsilon > Z/(2|\Psi(1)|)^{1/2}m^{-1/2}$.

It is interesting to see how some of these conclusions could have been reached without the detailed analysis presented above. In particular, the equation (3.143) is precisely the equation which would be obtained without any effects of divergence or gravity waves. It is Rayleigh's equation for a monotonic PV profile. The difference is that, without a turning point and gravity waves in the far field, the boundary conditions for (3.143) would normally be that the eigenfunctions should be exponentially decaying for $|z| \to \infty$ at all orders, and in particular there would be no exponentially growing imaginary term. That would force $c_i < 0$, which is a contradiction, since c_i was assumed to be greater than zero.

Now, since it is possible to prove Rayleigh's theorem without recourse to eigenmodes, we should know that any attempt to construct solutions to (3.143), which are exponentially decaying as $|z| \to \infty$, by introducing an infinitessimally small imaginary part to c must ultimately contradict itself, and this is exactly what has happened. Reversing the signs of the imaginary terms in v_r throughout the analysis leads to the same contradiction for any postulated temporally decaying modes with $c_i < 0$. There are no analytic solutions to (3.143) satisfying exponentially decaying boundary conditions, and Rayleigh's theorem holds.

It follows that, if the structure of the instability is to be retained, the Rossby wave critical layers at $\sigma = 0$ must not introduce an imaginary part in the eigenfunction greater than that required for matching to the imaginary part of the radiating solution in $r > 1$. Since the amplitude of the imaginary part introduced by critical layers will depend linearly on the strength of the PV gradient at the critical layers, and since the radiating solution has only an exponentially small (in m) imaginary component as $r \to 1^+$, it follows that the PV gradient must be exponentially weak in m at the locations of the critical layers. Consequently, a rapidly decaying profile for the PV gradient, such as $\mathrm{sech}^2\zeta$, will require only a small enhancement of the peak PV gradient (in this case given by $\epsilon^{-1} = O(m^{1/2})$), whereas a broader function, such as $1/(1+\zeta^2)$ for example, would require PV gradients exponentially large (in m) at $\zeta = 0$ in order that they will decay to exponentially small amplitudes for z of order 1 where the critical

layers occur. The order of the peak PV gradient required to retain the instability is thus very dependent on the profile of the potential vorticity chosen. The common feature is the order of the PV gradient at the critical layers of the flow, which must be exponentially small in m.

3.4 Relevance to the existence or non–existence of a slow manifold

Having shown that an axisymmetric shallow water vortex is unstable for all Froude and Rossby numbers, one may ask how this result relates to the question of existence or non-existence of a slow manifold for the f-plane shallow water equations.

As a preliminary, let us recall two general properteis of the inviscid shallow water equations.

Firstly, suppose we have a solution of the shallow water equations $\{(u_r^{(1)}, u_\theta^{(1)}), h^{(1)}\}$:

$$u_r^{(1)} = \mathcal{U}_r(r,\theta,t;f); \quad u_\theta^{(1)} = \mathcal{U}_\theta(r,\theta,t;f); \quad h^{(1)} = \mathcal{H}(r,\theta,t;f), \tag{3.164}$$

where (u_r, u_θ) are the radial and azimuthal components of the velocity field, and h the height field. Then another solution of the shallow water equations, $\{(u_r^{(2)}, u_\theta^{(2)}, h^{(2)}\}$, is given by

$$u_r^{(2)} = -\mathcal{U}_r(r,\theta,-t;-f); \quad u_\theta^{(2)} = -\mathcal{U}_\theta(r,\theta,-t;-f); \quad h^{(2)} = \mathcal{H}(r,\theta,-t;-f). \tag{3.165}$$

This is a general transformation property of the shallow water equations, which states that if all particle velocities are reversed, including the sense of the background rotation, then the flow evolution is reversed. The effect on the potential vorticity is to transform $Q \to -Q$. This we shall refer to as the particle symmetry of the equations.

Secondly, suppose again that we have a solution of the shallow water equations given by (3.164). Then another solution of the shallow water equations, $\{u_r^{(3)}, u_\theta^{(3)}, h^{(3)}\}$, is given by

$$u_r^{(3)} = \mathcal{U}_r(r,-\theta,t;-f); \quad u_\theta^{(3)} = -\mathcal{U}_\theta(r,-\theta,t;-f); \quad h^{(3)} = \mathcal{H}(r,-\theta,t;-f). \tag{3.166}$$

This is also a general transformation property of the shallow water equations, which states that the motion is unchanged if the system is viewed in a mirror, with the radial coordinate unchanged, but with the sence of the azimuthal coordinate reversed. As in the case of (3.165), the effect on the potential vorticity is to transform $Q \to -Q$. However, unlike (3.165), time is not reversed under this transformation.

Now, in order to discuss the existence or non-existence of a slow manifold, we must say what properties we shall assume a slow manifold to possess. Such a set of properties, consistent with the notion of a slow manifold as it is usually understood, is the following:

(a) The dimension of the slow manifold is one-third the dimension of the phase space of the shallow water equations.

(b) The slow manifold is tangent to the quasigeostrophic manifold at zero Rossby number, so that flow on the slow manifold is identical to quasigeostrophic flow at the leading order in Rossby number.

(c) Steady axisymmetric vortices and steady parallel flows are on the slow manifold.

(d) Viewing the system in a mirror ($f \to -f, u_r \to u_r, u_\theta \to -u_\theta, h \to h$) keeps the system on the same slow manifold as that which it was on when viewed in the original frame.

(e) Reversing the absolute direction of motion of all particles ($f \to -f, u_r \to -u_r, u_\theta \to -u_\theta, h \to h$) keeps the system on a slow manifold.

(f) At any instant in time, the entire velocity field and the entire height field on the slow manifold can be determined uniquely from the potential vorticity distribution $Q(x, y)$ at that instant in time.

Notice that *(c)*, *(d)*, *(e)* and *(f)* all follow from *(a)* and *(b)* in the small Rossby number limit. In particular, for quasigeostrophic flow, the potential vorticity field can be shown to determine all other dynamical fields uniquely. Notice also that *(f)* is a strong form of *(a)*. Furthermore, notice the distinction between property *(d)*, which states that, under the specified transformation, the system must remain on the same slow manifold, and property *(d)*, which states that under a different transformation, the system must simply remain on a slow manifold. It seems reasonable to make this distinction, in that the transformation under consideration in *(d)* can be effected entirely by regarding the system under reflection in a mirror, and its evolution should therefore be identical to the original flow, whereas the transformation under consideration in *(e)* involves a change in the physical state of the system, and is therefore not directly connected to the original flow.

Since, on any slow manifold having property *(f)*, Q and f are sufficient to determine the entire velocity field and the entire height field, it is natural to impose one further condition:

(g) Under the transformation $Q \rightarrow -Q; f \rightarrow -f$, the implied flow on th... $\{\boldsymbol{u}_s, h_s\}$, will remain on the slow manifold, and be transformed acc... $-\boldsymbol{u}_s; h_s \rightarrow h_s$.

This last property was termed "sign consistency" by McIntyre & Norton (19... property of all standard balanced models, such as the quasigeostrophic and ... equations. Property *(g)*, where it applies, is a strong form of property *(e)*, an... to giving the same force to property *(e)* as that which was given to proper... that under the relevant transformation, the slow manifold is not simply mapp... manifold, but it is mapped onto the same slow manifold. If $\mathcal{M}$ were assum... then property *(g)* and property *(e)* would be equivalent.

It will now be shown that, for the shallow water equations at least, a slow ma... properties *(a)* – *(g)* cannot exist. This is done by first supposing that ther... manifold, $\mathcal{M}$, say, and then deriving a contradiction.

Firstly, recall the vortex instability result. In an unbounded domain, it wa... vortex was temporally unstable for all Froude and Rossby numbers. Moreov... unique temporally growing mode for each azimuthal mode number m, provide... large, and the amplitude of the mode is assumed to be finite at infinity. It ... unique temporally decaying mode for each such m, assuming that the amplitu... finite at infinity. We shall fix attention on a single such azimuthal mode numbe... the growing and decaying eigenmodes for that mode number by $\phi_\uparrow(r)$ and ϕ... Here, the functions $\phi_\uparrow(r)$ and $\phi_\downarrow(r)$ are taken to represent the three-vector... case. In the limit $Ro \rightarrow 0$, $\phi_\uparrow$ and $\phi_\downarrow$ coincide, since the growth rate of the... small order in Rossby number.

Now, let us consider the nature of the evolution of the vortex with a pert... under the assumption of linearized perturbations. The boundary of the vortex ... tion on it, of small amplitude, and of fixed mode number m. On the slow manif... from property *(f)* and linearization that, at any instant in time, the amplitu... the vortex boundary undulation are sufficient to determine the perturbation ve... fields everywhere. Since we are considering linearized perturbations, we can repr...

location of the vortex boundary as a function $r(\theta, t)$, in the form

$$r(\theta, t) = \mathrm{Re}\left\{1 + a(t)e^{im\theta}\right\}, \tag{3.167}$$

for some complex function $a(t)$. Moreover, the assumption of linearized perturbations implies that the amplitude of the perturbations of the velocity and height fields depends linearly on the amplitude of the vortex boundary undulations, so we may represent the perturbation fields φ in the form

$$\varphi(r, \theta, t) = \mathrm{Re}\left\{a(t)\Phi(r)e^{im\theta}\right\}, \tag{3.168}$$

where here, as before, φ represents the three-vector (u, v, h). By substitution of (3.168) into the linearized shallow water equations for disturbances to an axisymmetric vortex, it can be shown using elementary techniques from the theory of solution of partial differential equations by separation of variables that

$$\frac{da}{dt} = \lambda a, \tag{3.169}$$

for some possibly complex constant λ, and therefore

$$\varphi(r, \theta, t) = \mathrm{Re}\left\{A\Phi(r)e^{i(m\theta - \lambda t)}\right\} \tag{3.170}$$

for some complex constant A.

Now, we know from the shallow water vortex instability analysis that, in the unbounded domain, only two bounded solutions of form (3.170) exist, and that they are the temporally growing eigenmode, $\phi_\uparrow$, and the temporally decaying eigenmode, $\phi_\downarrow$. Therefore, in the limit of linearized perturbations of fixed mode number to the boundary of an axisymmetric vortex, the flow on the slow manifold $\mathcal{M}$ consists of either the temporally growing mode alone, or the temporally decaying mode alone.

Suppose that the flow on the slow manifold consists of the temporally growing mode alone. Suppose also, without loss of generality, that $f > 0$. Then, under the transformation $Q \to -Q; f \to -f$, it follows, from property *(g)* and the time symmetry of the shallow water equations (3.164 & 3.165), that the temporally growing mode is mapped onto the temporally decaying mode. However, by property *(d)* and the reflectional symmetry of the shallow water equations (3.164 & 3.166), the transformation $Q \to -Q; f \to -f$ must map the temporally growing mode onto the temporally growing mode. This implies that the flow on the slow manifold for

$\{-Q, -f\}$ consists simultaneously of the temporally growing mode alone and the temporally decaying mode alone, which is a contradiction.

Now let us suppose we relax the condition *(g)*, the condition of sign consistency, and ask whether there can exist a slow manifold satisfying only conditions *(a)* – *(f)*. The velocity field on such a slow manifold, which we shall denote by $\boldsymbol{u}_s$, can then be written as the sum of two components, $\boldsymbol{u}_s = \boldsymbol{v} + \boldsymbol{w}$, where, under the transformation $Q \to -Q$; $f \to -f$, the velocity fields $\boldsymbol{v}$ and $\boldsymbol{w}$ transform according to the transformation laws $\boldsymbol{v} \to -\boldsymbol{v}$ and $\boldsymbol{w} \to \boldsymbol{w}$. One can regard $\boldsymbol{v}$ as the sign-consistent part of the slow manifold flow, and $\boldsymbol{w}$ as the sign-inconsistent part of the flow. The sign-inconsistent part, $\boldsymbol{w}$, is in general non-zero, and in particular it is non-zero, at arbitrarily small Froude and Rossby numbers, in the neighbourhood of an axisymmetric vortex. On the other hand, if $\boldsymbol{v} + \boldsymbol{w}$ is the velocity field on a slow manifold for a given $\{Q(x, y), f\}$ then, from the general transformation properties of the shallow water equations (3.164, 3.165), $-\boldsymbol{v} - \boldsymbol{w}$ constitutes a velocity field for $\{-Q(x, y), -f\}$, which must also be on a slow manifold. Then, from the transformation properties assumed for $\boldsymbol{v}$ and $\boldsymbol{w}$, $\boldsymbol{v} - \boldsymbol{w}$ constitutes the flow on a slow manifold for a potential vorticity distribution and Coriolis parameter $\{Q(x, y), f\}$. We have thus now arrived at a situation in which two flows, $\boldsymbol{v} + \boldsymbol{w}$ and $\boldsymbol{v} - \boldsymbol{w}$, are both flows on a slow manifold corresponding to the same potential vorticity distribution $Q(x, y)$ and Coriolis parameter f, and that $\boldsymbol{w}$ is in general non-zero, even at arbitrarily small Rossby numbers. It follows at once that, if condition *(g)*, the condition of sign consistency, is not imposed, then a slow manifold, if one exists, is not uniquely defined by the potential vorticity distribution, $Q(x, y)$, even in the limit of arbitrarily small Rossby number.

We might now attempt to make the definition of the slow manifold unique by insisting that its gravity wave field satisfies a causality condition, and hence, in the case of the shallow water instability problem, that the temporally growing mode is taken for both $f > 0$ and $f < 0$. This is feasible in the case of the vortex instability problem, because we can require that there are no gravity waves in the flow at $t = -\infty$. In this case, the limit $t \to -\infty$ is well defined, and the flow approaches an axisymmetric balanced vortex in that limit. For a general shallow water flow, however, the limit $t \to -\infty$ is not necessarily well defined, and imposing a causality condition will require the specification of an initial instant t_0 (say), at which time the vortex flow was "turned on", and before which time no gravity waves had been radiated. The choice of initial instant is clearly not unique, and yet in general we must expect the details of the

gravity waves generated, and their effect on the vortical flow, to depend on that initial instant. This point was made explicitly in analysis of the effect of gravity wave generation on a localized vortical disturbance at small Froude number, discussed in chapter 2. Each flow corresponding to a different choice of t_0 can be thought of as a "slow manifold", but with t_0 required to select one of the infinite number of possible branches.

There can now be little doubt now that the original concept of the unique slow manifold, as defined by Leith (1980), and clarified by Lorenz (1986), cannot apply to the shallow water equations. The shallow water vortex instability analysis showed that, at sufficiently large mode number, an axisymmetric vortex is unstable at arbitrarily small Rossby number, from which it follows that a slow manifold cannot be sign consistent, and therefore cannot be unique. Furthermore, the eigenfunctions have a gravity wave-like tail, albeit an exponentially weak one. One might argue from a physical viewpoint that the clear appearance of gravity-wave like structures at arbitrarily small Rossby number is sufficient to show that no so-called slow manifold could be truly slow, even if an invariant manifold specified only by potential vorticity distributions did exist, since it would inevitably contain "fast" gravity wave-like structures. From a mathematical viewpoint we have seen that the appearance of these gravity wave-like structures at arbitrarily small Rossby number is indeed inextricably linked to our inability to define a sign-consistent slow manifold, and hence our inability to define a unique slow manifold.

3.5 Instability of a parallel jet

In this section, we consider the linear stability of a parallel jet. The potential vorticity in the jet is taken to be piecewise uniform, with a single discontinuity at $y = 0$. The discontinuity is sometimes referred to as a front.

We choose to work with a nondimensionalization that is different from the one used in the study of the axisymmetric vortex instability. Instead of nondimensionalizing such that the jump in potential vorticity across the interface is unity, we shall set the potential vorticity in the region $y < 0$ to unity. There is no loss of generality here, since if we wish to have a layer of finite depth, with velocity decaying away from $y = 0$, the potential vorticity will have to be positive in both $y < 0$ and $y > 0$. The timescale must then be set by the Coriolis parameter f, and not the potential vorticity jump. We take $f = 1$, and so the layer depth is unity in the

limit $y \to -\infty$. The Rossby deformation radius is effectively $Q^{-1/2}$, where Q is the potential vorticity. A result of these nondimensionalizations is therefore to set the deformation radius to unity in $y < 0$.

3.5.1 The basic state

With the nondimensionalizations introduced above, the equations for the basic state velocity $U(y)$ and height $H(y)$ are

$$U = -\frac{dH}{dy} \tag{3.171}$$

$$1 - \frac{dU}{dy} = QH. \tag{3.172}$$

It follows that

$$\frac{d^2H}{dy^2} - QH = -1. \tag{3.173}$$

Now we let Q take the form

$$Q(y) = \begin{cases} q & \text{for } y > 0 \\ 1 & \text{for } y < 0 \end{cases}, \tag{3.174}$$

then the equations for the basic state can be solved in terms of elementary functions. Imposing continuity of H and U at $y = 0$ gives

$$H(y) = \begin{cases} \frac{1}{q} + \frac{\sqrt{q}-1}{q} e^{-\sqrt{q}y} & y > 0 \\ 1 - \frac{\sqrt{q}-1}{\sqrt{q}} e^{y} & y < 0 \end{cases}; \qquad U(y) = \begin{cases} \frac{\sqrt{q}-1}{\sqrt{q}} e^{-\sqrt{q}y} & y > 0 \\ \frac{\sqrt{q}-1}{\sqrt{q}} e^{y} & y < 0 \end{cases}. \tag{3.175}$$

3.5.2 Ripa's theorem

Before proceeding to the linear instability calculations for the basic state given by (3.175), we shall discuss Ripa's theorems for the stability of parallel and axisymmetric shallow water flows, and show linear stability for the basic state (3.175) in the case $q < 4$.

The basic technique is to consider first and second variations to some linear combination of the conserved quantities of energy and momentum, plus a Casimir. The first variation of the sum is required to be zero, which sets the Casimir, and the second variation is required to be of definite sign. The requirement of sign definiteness gives conditions on the basic flow which are sufficient for linear stability.

For the shallow water equations, the energy is given by

$$E = \frac{1}{2}\int\int (h\boldsymbol{u}^2 + h^2)dxdy, \tag{3.176}$$

while the x-momentum is given by

$$M = \int\int h(u - fy)dxdy. \tag{3.177}$$

General Casimirs are given by

$$C = \int\int hF(Q)dxdy, \tag{3.178}$$

where $F(Q)$ is an arbitrary function of the potential vorticity Q.

It follows that the combination

$$\mathcal{H} \equiv E + \alpha M + C \tag{3.179}$$

is a conserved quantity. Taking the first variation of $\mathcal{H}$ gives

$$\delta\mathcal{H} = \int\int \left[h\left(u + \alpha + \frac{\partial}{\partial y}C'(Q)\right)(\delta u) - \frac{\partial}{\partial x}C'(Q)(\delta v) \right. \tag{3.180}$$

$$\left. + \left(\frac{1}{2}u^2 + h + \alpha(u - fy) + C(Q) - QC'(Q)\right)(\delta h)\right]. \tag{3.181}$$

For stability we require $\delta\mathcal{H} = 0$. This imposes conditions on the Casimir C, such that

$$C''(Q) = -\frac{h(u+\alpha)}{Q'(y)}. \tag{3.182}$$

For linear stability, we now require that the second variation $\delta^2\mathcal{H}$ should be of definite sign. The second variation of $\mathcal{H}$ is given by

$$\delta^2\mathcal{H} = \int\int \left[h(\delta v)^2 + (\delta h + (u+\alpha)\delta u)^2 + \left(h - (u+\alpha)^2\right)(\delta u)^2 - \frac{h^2(u+\alpha)}{Q'(y)}(\delta Q)^2 \right]. \tag{3.183}$$

If we impose the condition $\delta^2\mathcal{H} > 0$, we obtain

$$(u+\alpha)Q'(y) < 0 \quad \text{and} \quad (u+\alpha)^2 < h. \tag{3.184}$$

If we assume that $Q'(y) > 0$ for all y, then the maximum value for α which we may take to obtain stability by the first of these conditions is the peak jet velocity $-u_{max}$. This choice gives us the greatest stability information from the second of the conditions, so we have

$$Q'(y) > 0 \quad \text{and} \quad (u_{max} - u(y))^2 < h(y). \tag{3.185}$$

The condition is just the standard condition of monotonic PV required to ensure that the flow is not unstable to barotropic instability, with the further condition that the flow must be everywhere subsonic with respect to the peak jet velocity, since $h(y)^{1/2}$ is just the nondimensionalized gravity wave phase speed. This condition is analogous to Arnol'd's first theorem.

From (3.175), we see that these conditions are satisfied when $q < 4$.

In passing we note that, by considering the subclass of variations in which $\delta q = 0$ and $\delta u = 0$, the remaining terms are always positive definite, and hence it is not possible to impose the condition $\delta^2 \mathcal{H} < 0$. Consequently there is no analogue of Arnol'd's second theorem for the shallow water equations [1].

3.5.3 The disturbance equations

The disturbance equations for the parallel flow are very similar, except for geometrical factors, to those for axisymmetric flow.

Now we assume that perturbations are of the form $f(y)e^{i(kx-\omega t)}$. The wavenumber k is assumed real, and again the growth rate of any instability is represented by an imaginary part to ω. The disturbance equations then become

$$\frac{du}{dy} = A(y)u + B(y)v \tag{3.186}$$

$$\frac{dv}{dy} = C(y)u + D(y)v, \tag{3.187}$$

where Q is the (piecewise uniform) potential vorticity of the background flow, and

$$A(y) = \frac{Q(-\omega + kU)}{k} \tag{3.188}$$

$$B(y) = i\left(k + \frac{Q^2 H}{k}\right) \tag{3.189}$$

$$C(y) = i\left(\frac{(-\omega + kU)^2}{Hk} - k\right) \tag{3.190}$$

$$D(y) = \left(\frac{U}{H} - \frac{(-\omega + kU)Q}{k}\right) \tag{3.191}$$

Continuity of v and h at $y = 0$ reduce to

$$(-\omega + kU(0))\left[\frac{u}{v}\right] + iH(0)[Q] = 0. \tag{3.192}$$

[1] This observation arose in discussions with Prof. T. G. Shepherd and Dr. T. E. Dowling

3.5.4 Numerical technique

The equations (3.186-3.187) represent an eigenvalue problem for ω. As for the axisymmetric problem, we use a shooting method based on an initial guess for ω. In the far field the disturbances take the form $e^{i(kx+ly-\omega t)}$, where l is given by $\omega^2 - 1 = Q^{-1}(k^2 + l^2)$. Note that l, which takes different values as $y \to \pm\infty$, always has non-zero imaginary part l_i. The disturbance variables are therefore rescaled by $e^{l_i y}$, so that the computed variables remain of order one throughout the domain $-\infty < y < \infty$. Evanescence conditions at $y = \pm\infty$ are used to impose the ratio of u/v to be taken at some large value of $|y|$, as in the axisymmetric vortex calculations. The equations are then integrated in to $y = 0$ using a stiff integrator for ordinary differential equations from the NAG library, and the continuity condition (3.192) at $y = 0$ is imposed.

Initialization for small q

Initial values for the eigenfrequencies supplied to the shooting method were taken, in the case of a small jump in potential vorticity, using a quasi-geostrophic analysis. The magnitude of the potential vorticity jump could then be increased, taking the eigenfrequency for one value of the potential vorticity jump as the initial guess for the eigenfrequency at the next.

The equation for the basic state streamfunction ψ:

$$\left(\frac{d^2}{dy^2} - 1\right)\psi = \begin{cases} q & \text{for } y > 0 \\ 1 & \text{for } y < 0 \end{cases} . \tag{3.193}$$

Continuity of ψ and ψ_y at $y = 0$ lead to

$$\psi = \begin{cases} \frac{1}{2}(q-1)e^{-y} - q & y > 0 \\ -\frac{1}{2}(q-1)e^{y} - 1 & y < 0 \end{cases} . \tag{3.194}$$

If we now introduce a disturbance streamfunction $\tilde{\psi}(y)e^{i(kx-\omega t)}$, the equation for $\tilde{\psi}$ is

$$\left(\frac{d^2}{dy^2} - k^2 - 1\right)\tilde{\psi} = 0. \tag{3.195}$$

At this point it is convenient to introduce

$$\gamma \equiv \left(1 + k^2\right)^{1/2}, \tag{3.196}$$

so that

$$\tilde{\psi} = \begin{cases} \tilde{a}e^{-\gamma y} & \text{for } y > 0 \\ \tilde{b}e^{+\gamma y} & \text{for } y < 0 \end{cases} . \tag{3.197}$$

We must also introduce a disturbance to the position of the interface between the two uniform potential vorticity regions, such that

$$y = \tilde{\eta}e^{i(kx-\omega t)}. \tag{3.198}$$

Imposing continuity of ψ and ψ_y leads to a kinematic condition on $\tilde{\eta}$, which can then be solved for ω:

$$\omega = \frac{1}{2}\left(1 - \frac{1}{\gamma}\right)(q-1)k. \tag{3.199}$$

Note that this eigenfrequency is real, regardless of the magnitude of $q-1$. It is valid in the low Rossby number limit $q - 1 \ll 1$, and is not valid for $q - 1$ of order one. It therefore gives no information in the regime $q > 4$, in which we might expect instability to occur.

3.5.5 Results of the eigenvalue calculations

In the present study, the growth rate of any instability depends on two parameters: the potential vorticity q in the region $y > 0$; and the wavenumber k of the instability.

In figure 3.30, the growth rate of the instability is shown as a function of wavenumber for $q = 6, 10, 20$ and 30. One can see that, at low wavenumbers, the flow is stable, whereas at higher wavenumbers the instability occurs. This is to be expected when we consider the nature of the instability mechanism. Far from the jet, the disturbances satisfy the linear gravity wave dispersion relation

$$\omega^2 - f^2 = c^2(k^2 + l^2). \tag{3.200}$$

Now, the instability exists due to coupling between the vortical wave on the potential vorticity jump, and a gravity wave at infinity. This implies that, in one of $y < 0$ or $y > 0$, we will need l to be real at leading order in the wavenumber k. If we suppose that $\omega \sim Uk$, then the condition $l^2 > 0$ is $k^2(U^2 - c^2) - f^2 > 0$. Thus, the larger the value of k, the less the disturbances are influenced by the Coriolis parameter, and the remaining condition is that the jet speed must exceed the gravity wave phase speed at infinity – exactly Ripa's condition. Hence, we should expect to see Rossby wave – gravity wave instabilities at large wavenumber.

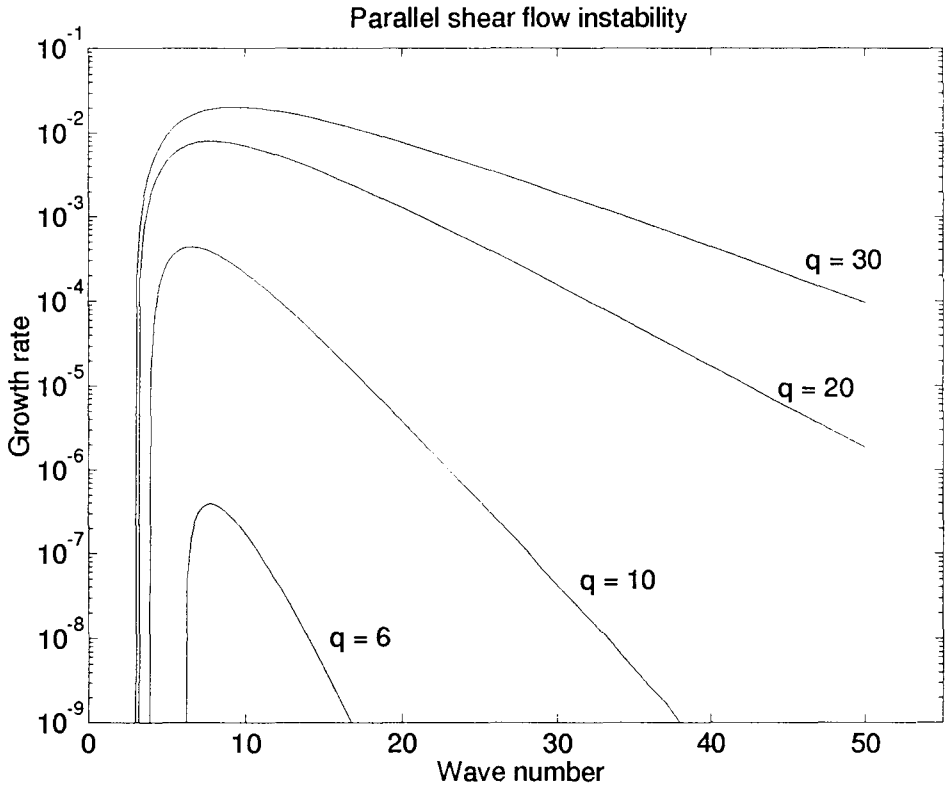


Figure 3.30: Growth rates of linear eigenmodes on a parallel shear flow with a single discontinuity in potential vorticity. Potential vorticity ratios of $6, 10, 20, 30$ are show, with growth rates of eigenmode wavenumbers $k = 0$ to 50.

We shall therefore proceed to analyse the instability in the high wavenumber limit, as we did for the axisymmetric vortex at large mode number.

3.5.6 The limit of large wavenumber

The results of the eigenvalue calculations have shown that there is a tendency towards instability at high wavenumbers. Motivated by the success of the WKBJ analysis for the axisymmetric vortex instability, we shall now perform a WKBJ analysis for the parallel flows discussed in this section. In particular, we shall show that the range of stability found by Ripa's theorem is complete - all flows that do not satisfy Ripa's theorem are unstable.

The details of the parallel flow and axisymmetric vortex analyses are very similar, and only a summary will be given here. The principal difference is that we now impose an evanescence/radiation condition as $y \to \pm\infty$, and this will lead to zero, one or two turning points in the flow.

As in the analysis for the axisymmetric flow, we choose to work with a single second order ordinary differential equation with a single variable. In the interests of consistency with the axisymmetric analysis, we choose to work with the along-stream variable u.

Again, one works with a frequency $c = \omega/k$. After rearranging the disturbance equations (3.186, 3.187) one obtains

$$\frac{d^2u}{dy^2} + F(y)\frac{du}{dy} + G(y)u = 0. \tag{3.201}$$

The function $F(y)$ is everywhere of order unity, and the function $G(y)$ is given by

$$G(y) = \left[\frac{(-c + U(y))^2}{H(y)} - 1\right] k^2 + O(1). \tag{3.202}$$

Again, it is straightforward to see that the jump condition at $y = 0$ can only be satisfied if $c_0 = U(0)$. This gives a condition for a turning point to occur:

$$(u(0) - U(y))^2/H(y) = 1, \tag{3.203}$$

which is precisely the same as the stability boundary predicted by Ripa's theorem.

Now $U(0) = 1 - 1/\sqrt{q}$, and $U \to 0$ as $y \to \pm\infty$. If we consider first $y > 0$, then there will be a turning point if and only if

$$(U(0) - U(y))^2 > H(y). \tag{3.204}$$

Now $H(y) > 1/q$, and $U(0) - U(y) < U(0)$, and therefore a necessary condition for a turning point to exist is

$$U(0)^2 > 1/q; \quad \text{i.e. } q > 4. \tag{3.205}$$

Taking $y = \infty$ in (3.203) ensures that (3.205) is also a sufficient condition.

If we now consider $y < 0$, then $H(y) < H(0)$. There

$$U(0) - U(y) = \frac{\sqrt{q} - 1}{\sqrt{q}}(1 - e^y) < 1, \tag{3.206}$$

and therefore $H(y) - (U(0) - U(y)) = 1/\sqrt{q}$, and hence

$$1 > H(y) > U(0) - U(y) > ((U(0) - U(y))^2, \tag{3.207}$$

and so there is no turning point in $y < 0$.

The situation is now very similar to that of the axisymmetric vortex, in which there is a single turning point and a single potential vorticity discontinuity.

The turning point at $y = y_c$ means that the domain is split into four regions, and the analysis is given below. The four regions and the character of the solutions in them are:

$y < 0$ Exponential decay of solution away from $y = 0$.

$0 < y < y_c$ Exponential decay of real part of solution away from $y = 0$, with exponentially growing imaginary part, such that real and imaginary parts are of the same order of magnitude at $y = y_c$.

$y - y_c = O(k^{-2/3})$ Combination of Airy functions, appropriate for matching onto radiating wave solution in $y > y_c$.

$y > y_c$ Oscillatory functions, satisfying radiating boundary conditions as $t \to \infty$. Small imaginary part of eigenfrequency ensures exponential decay of eigenfunction at very large range.

Continuity conditions at $y = 0$ are then used to obtain the eigenvalue. If we let the eigenvalue ω be written as $\omega = k(c_0 + k^{-1}c_1 + \cdots)$, then

$$c_0 = 1 - 1/\sqrt{q}; \quad c_1 = -\frac{q-1}{2\sqrt{q}}. \tag{3.208}$$

Then if we define

$$\Psi_0 = -\int_0^{y_c} \left(1 - \frac{(\sqrt{q}-1)^2(1-e^{-y\sqrt{q}})}{1+(\sqrt{q}-1)e^{-y\sqrt{q}}}\right)^{1/2} dy \tag{3.209}$$

and

$$\Psi_1 = -\frac{q-1}{2\sqrt{q}} \int_0^{y_c} \frac{U(0)-U(y)}{\Psi_0' H(y)} dy \tag{3.210}$$

then the growth rate c_i is given by

$$c_i = \frac{q-1}{4\sqrt{q}} e^{2(k\Psi_0+\Psi_1)}. \tag{3.211}$$

Comparison with numerical eigenvalue calculations

In figure 3.31, t he growth rate curves shown in figure 3.30 are reproduced, together with their predictions from the WKBJ analysis, valid in the limit $k \gg 1$. On average, there seems to be reasonable agreement at large k, though at moderate k the agreement is not generally as good as it was for the WKBJ analysis of the axisymmetric vortex instability at moderate m.

3.5.7 The limit $q \to \infty$

As discussed in the section 1, the instability of a potential vorticity jump has already been analysed in the limit $q - \infty$ by Paldor (1983), who showed that all eigenmodes are stable in that limit. However, one can show that, in the limit $q \to \infty$,

$$\Psi_0 \sim -\frac{\pi}{4} q^{-3/4}; \quad \Psi_1 \sim -\frac{1}{2}, \tag{3.212}$$

and hence

$$c_i \sim \frac{1}{4} q^{1/2} e^{-\left(\frac{1}{2}k\pi q^{-3/4}+1\right)}, \tag{3.213}$$

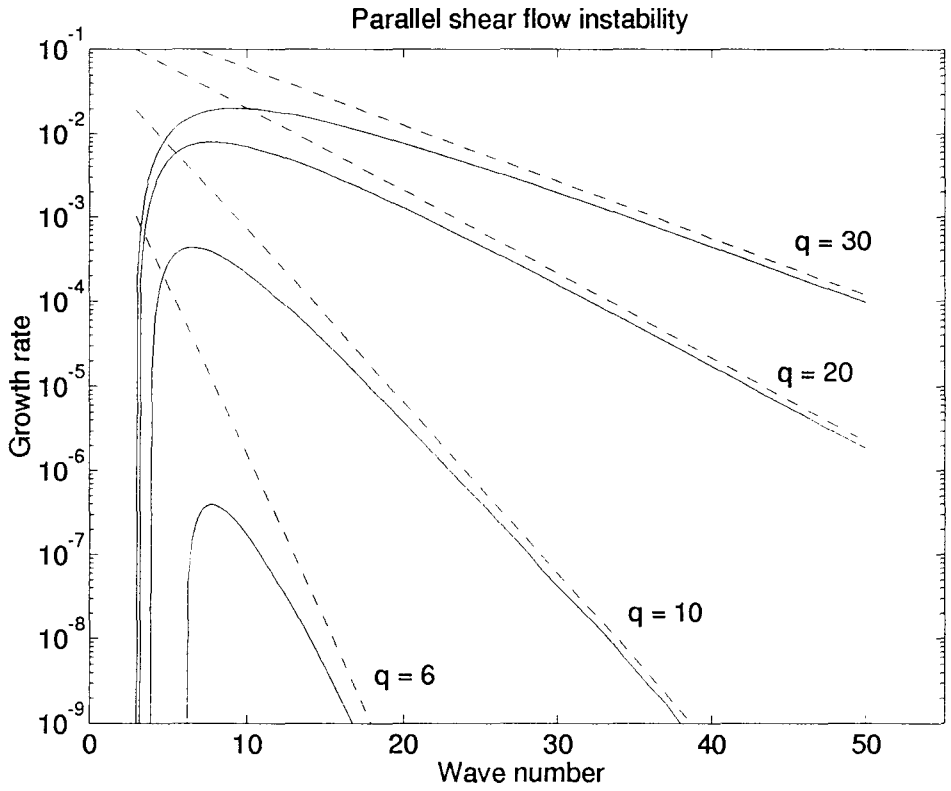


Figure 3.31: Solid lines as figure 3.30, and dashed lines are corresponding growth rate predictions obtained by the WKBJ analysis.

which becomes unbounded in the limit $q \to \infty$. The purpose of this section is to show that the limit $q \to \infty$ is a singular limit. The point is that, at large fixed k, as q is increased, we will approach a limit in which the potential vorticity interface falls within the turning point region. The region $0 < y < y_c$ effectively disappears, and we will be left with only three regions. The appropriate scaling in this limit is to take $q = O(k^{4/3})$.

The scaling $q = O(k^{4/3})$

The aim here is to appreciate the origins of the scaling $q = O(k^{4/3})$. We are seeking to eliminate the inner region $0 < y < y_c$, and therefore we impose that the region including the turning point $y = y_c$ should also include the potential vorticity interface $y = 0$. In the limit $q \to \infty$, the turning point and the interface coincide.

To derive the necessary scaling, we first note that

$$U(0) - U(y) = \left(1 - \frac{1}{\sqrt{q}}\right)\left(1 - \epsilon^{-\sqrt{q}y}\right). \tag{3.214}$$

Notice that $1/\sqrt{q}$ is the natural length scale in the region $y > 0$, and in this analysis we seek to place the turning point at some $y \ll 1/\sqrt{q}$. Therefore,

$$U(0) - U(y) \sim -y\sqrt{q}. \tag{3.215}$$

At the turning point $y = y_c$, we will have $(U(0) - U(y_c))^2 = H(y_c)$. Now $H(y) \sim 1/\sqrt{q}$ for small $\sqrt{q}y$, and therefore the turning point occurs at

$$y^2 q \sim 1/\sqrt{q}; \quad \text{i.e.} \quad y \sim q^{-3/4}. \tag{3.216}$$

On the other hand, the natural length scale for disturbances is k^{-1}, and hence the turning point will lie in the same asymptotic region as the interface if

$$\frac{1}{k} \sim y \sim q^{-3/4} \quad \text{i.e.} \quad q \sim k^{4/3}. \tag{3.217}$$

This scaling could have been predicted with reference to (3.213). The asymptotic analysis of the previous section assumed $k\Psi_0 \gg \Psi_1$, and it can clearly be seen that this assumption breaks down when $q = O(k^{4/3})$. Throughout the remainder of this chapter, we shall therefore replace q with $k^{4/3}\mathsf{q}$, where q is to be taken to be of order unity.

Equations in the region $y = O(k^{-1})$

With the scaling for the potential vorticity jump now fixed, we are able to write down an equation governing disturbances in the region $y = O(k^{-1})$. We introduce an inner length coordinate $Y = ky$, for which the interface is to be found at $Y = 0$, and the turning point at $Y = \mathsf{q}^{-3/4}$.

We express the eigenfrequency ω as

$$\omega = k(c_0 + k^{-1/3}\mathsf{q}^{1/2}c_1 + \cdots). \tag{3.218}$$

The leading order is satisfied by taking $c_0 = U(0)$. The factor of $\mathsf{q}^{1/2}$ is introduced to scale c_1, which is convenient for analysing the limits $\mathsf{q} \to 0$ and $\mathsf{q} \to \infty$.

Then, after some routine algebra, we obtain an equation for u at leading order in k:

$$\frac{d^2u}{dY^2} + (\mathsf{q}^{3/2}(Y + c_1)^2 - 1)u = 0. \tag{3.219}$$

The solution to (3.219) may be represented in terms of parabolic cylinder functions (see Abramowitz & Stegun, 1965, chapter 19).

In the limit $Y \to \infty$, we shall seek to match the solution of (3.219) onto a radiating wave form. We therefore take the solution

$$u(Y) = E(a, 2^{1/2}\mathsf{q}^{3/8}(Y + c_1)), \tag{3.220}$$

where

$$a = \frac{1}{2}\mathsf{q}^{-3/4}. \tag{3.221}$$

The function $E(a, x)$ is related to the real parabolic cylinder function $W(a, x)$ via

$$E(a, x) = \lambda^{-1/2}W(a, x) + i\lambda^{1/2}W(a, -x), \tag{3.222}$$

where

$$\lambda = \sqrt{1 + e^{2\pi a}} - e^{\pi a}; \quad \lambda^{-1} = \sqrt{1 + e^{2\pi a}} + e^{\pi a}. \tag{3.223}$$

In the limit $Y \to \infty$, $u(Y)$ then takes the form

$$u(Y) \sim \sqrt{2/Y}\exp\{i(\mu^2(Y + c_1)^2/4 - a\ln\mu Y + \arg(\Gamma(ia + 1/2)) + \pi/4)\}, \tag{3.224}$$

where

$$\mu = \sqrt{2}\mathsf{q}^{3/8}. \tag{3.225}$$

If c_1 has an imaginary part, one can see from (3.224) that $u(Y)$ is exponentially decaying in the limit $Y \to \infty$.

The continuity equation at $Y = 0$ becomes

$$c_1 \left(\frac{1}{(1 - c_1^2 \mathsf{q}^{3/2})^{1/2}} - \frac{(1 + \mathsf{q}^{3/2}) E(a, \sqrt{2}\mathsf{q}^{3/8} c_1)}{\sqrt{2}\mathsf{q}^{3/8} E'(a, \sqrt{2}\mathsf{q}^{3/8} c_1) + \mathsf{q}^{3/2} c_1 E(a, \sqrt{2}\mathsf{q}^{3/8} c_1)} \right) = -1, \tag{3.226}$$

where the prime denotes differentiation with respect to the second argument. The second term in the bracket is the ratio of u/v at $Y = 0^+$. Recalling that $du/dy = A(y)u + B(y)v$, the complicated form of the denominator comes from the fact that, in this scaling, du/dy, $A(y)u$ and $B(y)v$ are all of the same order of magnitude, whereas in the limit of large k and order one potential vorticity jumps, $A(y)u$ may be neglected. The first term is the ratio of u/v at $Y = 0^-$. Notice that, unlike the previous analysis, it is not unity at leading order. Although there is no turning point in $Y < 0$, the expression

$$\Psi' = \left(1 - \frac{(c - U(y))^2}{H(y)} \right)^{1/2} \tag{3.227}$$

becomes

$$\Psi' = (1 - c_1^2 \mathsf{q}^{3/2}) \tag{3.228}$$

This explains the modification in the first term in (3.226).

The equation (3.226) is a nonlinear equation for the eigenvalue c_1, and it is in general difficult to prove results about the existence or uniqueness of its solutions. For the present analysis, we shall be concerned only with the development of the Rossby wave – gravity wave mode, already discussed in this section, in the limit as the potential vorticity jump is made very large.

The limit $\mathsf{q} \to 0$

We consider first the limit $\mathsf{q} \to 0$. In this limit, we introduce a new variable $z = \mathsf{q}^{3/4} Y$. In this variable, (3.219) becomes

$$\mathsf{q}^{3/2} \frac{d^2 u}{dz^2} + \left(z^2 - 1 + 2c_1 \mathsf{q}^{3/4} z + c_1^2 \mathsf{q}^{3/2} \right) = 0. \tag{3.229}$$

Moreover, c_1 is also expanded in powers of $\mathsf{q}^{3/4}$:

$$c_1 = c_1^0 + \mathsf{q}^{3/4} c_1^1 + \cdots. \tag{3.230}$$

For the present purposes, only c_1^0 will be required to obtain the leading order expression for the growth rate of the instability.

The solution to (3.229) is then written in the form

$$u(z) = A(z) e^{\mathsf{q}^{-3/4}\psi_0(z) + \psi_1(z)} + O(\mathsf{q}^{3/4}). \tag{3.231}$$

Substituting (3.231) into (3.229), one obtains

$$\frac{d\psi_0}{dz} = (1 - z^2)^{1/2} \tag{3.232}$$

$$\frac{d\psi_1}{dz} = -\frac{c_1^0 z}{(1 - z^2)^{1/2}}. \tag{3.233}$$

The analysis from this point is very similar to the WKBJ analyses in the limit of large k. The turning point in the WKBJ analysis occurs at $z = 1$, and so we set $\psi_0(1) = 0$ and $\psi_1(1) = 0$. Then the continuity equation (3.192) becomes

$$c_1^0 \left[1 + 1 + i e^{2(\mathsf{q}^{-3/4}\psi_0(0) + \psi_1(0))} \right] = -1. \tag{3.234}$$

From this it follows that $c_1^0 = -1/2$, and hence

$$\psi_0(0) = -\frac{\pi}{4}; \quad \psi_1(0) = -\frac{1}{2}. \tag{3.235}$$

The growth rate of the instability is then given by

$$\mathrm{Im}(c) \sim \frac{1}{4} e^{-(\mathsf{q}^{-3/4}\pi/2 + 1)}. \tag{3.236}$$

This expression is in agreement with the expression (3.213) for the growth rate of the instability obtained by assuming $q = O(1)$, and then taking the limit $q \to \infty$.

The limit $\mathsf{q} \to \infty$

In the limit $\mathsf{q} \to \infty$, the equation (3.226) is rather difficult to analyse. The principal difficulty comes from the fact that there may be roots with $c_1 = O(\mathsf{q}^{-3/8})$, for which the second argument

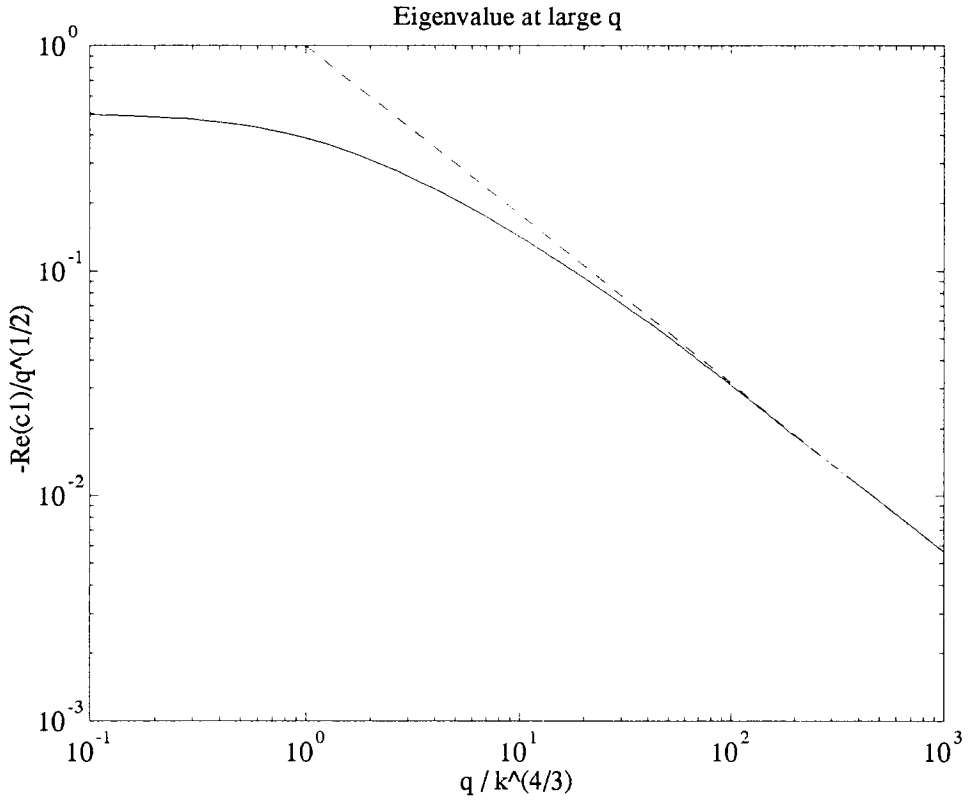


Figure 3.32: Real part of c_1, as a function of q (solid line), with $\mathrm{Re}c_1 = -\mathsf{q}^{-3/4}$ (dashed line).

of the parabolic cylinder function is of order unity. However, we are concerned in this thesis with the character of the Rossby – gravity wave instability as q is varied, and we shall therefore concentrate on the eigenmode which reduces to the above in the limit $\mathsf{q} \to 0$. In the limit $\mathsf{q} \to \infty$, it will be seen that the appropriate root of (3.226) comes by taking

$$c_1 = \mathsf{q}^{-3/4} c_1^1 + \mathsf{q}^{-3/2} c_1^2 + \cdots. \tag{3.237}$$

There is a unique root under this scaling. It shall further turn out that c_1^1 is real, whereas c_1^2 is complex, and the growth rate of the eigenmode, which is $k^{2/3}\mathsf{q}^{1/2}c_1$, thus tends to zero in the limit $\mathsf{q} \to \infty$.

To see this first examine figure 3.32, which shows the real part of c_1, computed numerically by solution of (3.219). At large q, on the log-log scale, one can see that $\mathrm{Re}c \to -\mathsf{q}^{-3/4}$ as $\mathsf{q} \to \infty$. The point about this is that then $1 - c_1^2\mathsf{q}^{3/2}$ is of small order, and hence the balance of terms in (3.226) may change. Expanding

$$c_1 = -\mathsf{q}^{-3/4} + C\mathsf{q}^{-3/2} \tag{3.238}$$

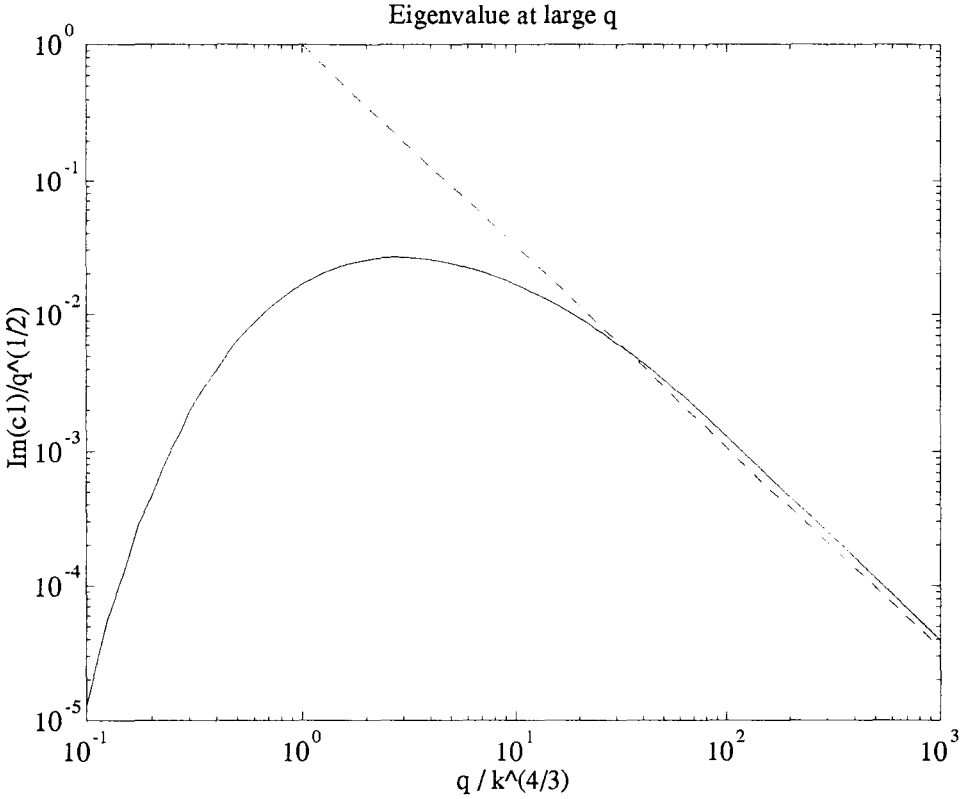


Figure 3.33: Imaginary part of c_1, as a function of q (solid line), with $\mathrm{Im}c_1 = \frac{1}{8}(\Gamma(1/4)/\Gamma(3/4))^2\mathsf{q}^{-3/2}$ (dashed line)

we obtain

$$\frac{1}{(2C)^{1/2}} + \sqrt{2}\frac{E'(0,0)}{E(0,0)} = 0 \tag{3.239}$$

for which the solution is

$$C = \frac{i}{8}\left(\frac{\Gamma(1/4)}{\Gamma(3/4)}\right)^2 \approx 1.09i. \tag{3.240}$$

The imaginary part of c_1 is shown in figure 3.33, with the value given by (3.240) for comparison. In this case, quite large values of q seem to be required for the validity of (3.240).

Thus, c_1 is imaginary at order $\mathsf{q}^{-3/2}$, and hence $\mathrm{Im}(\omega) = O(\mathsf{q}^{-1})$. Therefore, this mode is stable in the limit $\mathsf{q} \to \infty$, consistent with the analysis of Paldor (1983).

3.5.8 Eigenvalue calculations for large q

In figure 3.34, the growth rates of the instabilities are shown for potential vorticity jumps of 10^2, 10^3 and 10^6. The ordinate is $\mathsf{q}^{-3/4}$, or $k/q^{3/4}$ in terms of the original problem. On the

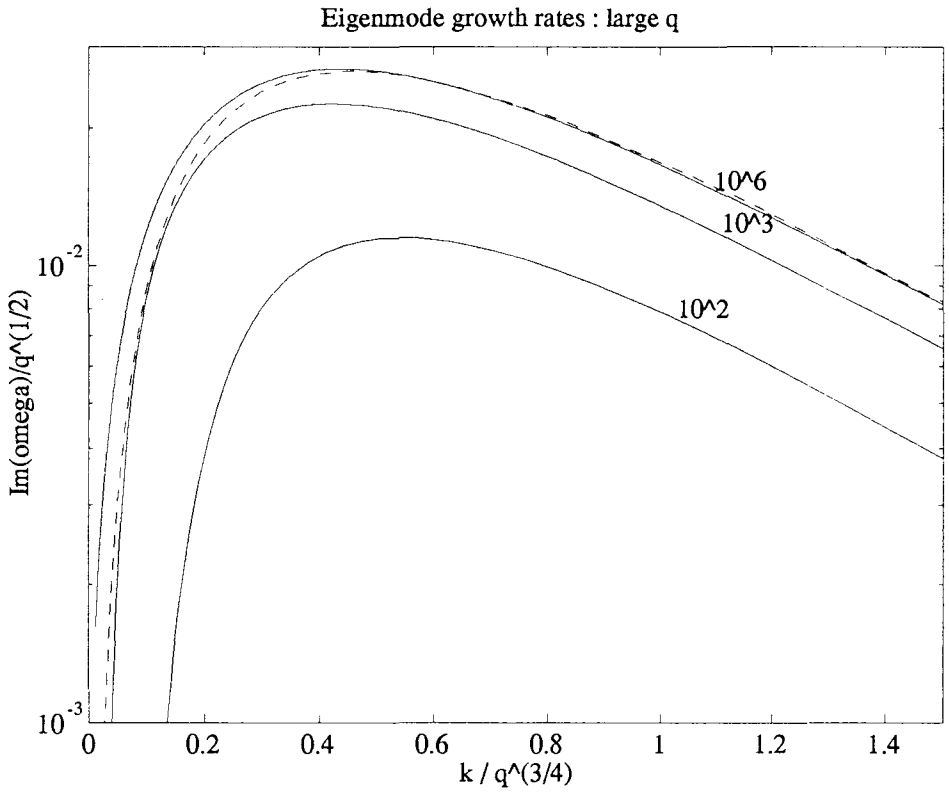


Figure 3.34: Solid lines are growth rates of linear eigenmodes for parallel flows with single potential vorticity jumps of 10^2, 10^3 and 10^6. Dashed line is the growth rate predicted in the limit $q \gg 1$, $k \gg 1$, $\mathsf{q} \equiv q/k^{4/3} = O(1)$

abscissa is $\mathrm{Im}(\omega)/\sqrt{q}$. In the limit of large q, with q of order unity, all the curves should fall on the dashed curve, which is obtained from the eigenvalue equation (3.226).

Comparison of the computed with the predicted growth rates shows that quite large values of q are required to obtain good agreement between computed growth rates and the large q limit. With $q = 10^6$, there is very good agreement for $k/q^{3/4}$ between 0.5 and 1.0. Outside this range, agreement is less good. It would improve with a larger value of q. However, the differential equations that must be integrated to obtain the eigenvalue become very stiff, and there is some doubt as to whether the numerical eigenvalue finder is sufficiently accurate to proceed to values of q in excess of about 10^6. However, calculation with $q = 2 \times 10^7$ (not shown) did appear to show some improvement over $q = 10^6$ for $k/q^{3/4}$ down to around 0.2, apparently confirming that we can ultimately expect to see agreement between the computed and theoretical curves for all but the smallest values of $k/q^{3/4}$, for which the analysis, which assumes $k \gg 1$, can never be valid.

3.5.9 Remarks

The instability of a parallel shear flow with a single discontinuity of potential vorticity has been investigated numerically, and analytically in the limit of short wavelength instabilities. The instability at short wavelength is present whenever the stability criteria of Ripa (1983) fail.

In the limit of a large potential vorticity jump, the instability becomes very strong, with a growth rate of order $k^{2/3}$ if the potential vorticity jump is of order $k^{4/3}$. However, as the jump is made large compared with $k^{4/3}$, the growth rate decays algebraically, and the limiting case of a surfacing front, investigated by Paldor (1983), is stable, as Paldor found.

3.6 Discussion

This chapter has furthered the study of vortex – gravity wave interactions in the shallow water equations by investigating instabilities which owe their existence to a coupling between Rossby waves and gravity waves.

First we investigated the stability of an axisymmetric vortex with a single radial discontinuity in potential vorticity. Numerical results showed that the presence of background rotation had a significant effect on the growth rates of any instabilities. This appeared to be due mainly

to differences in the character of the basic states. The sonic radius tended to be closer to the vortex boundary for values of the Coriolis parameter between 0 and -1. Asymptotic analysis for large azimuthal mode number showed that the vortex was always unstable, provided some rather weak assumptions were made about the basic state – assumptions which were always found to hold in practice. The growth rates are predicted to be exponentially small in the mode number for large mode number. There was good agreement between analysis and numerical eigenvalue calculations for quite moderate values of the mode number. The analysis was generalized to a limited class of flows with smooth potential vorticity profiles, and the connections between the current work and the study of Sozou (1987) were clarified.

The implications for the nature of a slow manifold for the shallow water equations were then discussed. Although there is no absolute proof that an invariant manifold of dimension 1/3 the dimension of the full phase space cannot be found, two things are clear. One is that any such invariant manifold is not unique. Another is that it is not truly slow, in that gravity wave - like features are bound to be present on it, albeit exponentially weakly at small Rossby number.

Finally, the stability of a parallel flow with a single discontinuity of potential vorticity was analysed. The numerical eigenvalue problem and large wavenumber limits are similar to those for the axisymmetric vortex. The problem was nondimensionalized somewhat differently from the axisymmetric vortex problem, and it allows a limit in which the potential vorticity on one side of the discontinuity becomes unbounded, corresponding to a vanishing layer depth. This limit has already been shown to be stable, and it turns out that a further asymptotic regime, in which the potential vorticity q in the shallow layer scales as $k^{4/3}$, where k is the wavenumber of the disturbance, is required to capture this limiting behaviour. Thus, although the maximum growth rate of the instability becomes unbounded as $q \to \infty$, it occurs at a wavenumber with a fixed value of $k/q^{3/4}$, with the growth rate tending to zero as $k/q^{3/4}$ tends to infinity.

In conclusion, a number of Rossby wave – gravity wave instabilities have been investigated in this chapter. They are generally short wave instabilities, and are therefore ideally suited to analysis by the WKBJ technique. In all cases investigated, the growth rates of the instabilities tended to be rather small, with the possible exception of an axisymmetric vortex at fairly large Froude number, with potential vorticity of opposite sign to the background. Nonetheless, it was shown that the instability persists, albeit exponentially weakly, for arbitrarily small

Rossby number, and this has significant implications for the slow manifold hypothesis. It seems increasingly unlikely that a slow manifold exists for the shallow water equations. In this chapter, it has been shown that if one does exist it is not unique, and must contain structures on it which are definitely gravity wave - like. It is unlikely that the concept of a truly slow manifold is therefore a useful one.

Chapter 4

Nonlinear simulation of gravity wave generation

4.1 Introduction

Since the development of the theory of aerodynamic sound by Lighthill (1952), several attempts have been made to verify it, and to investigate the range of Mach numbers over which it is valid (see, for example, Lighthill, 1954; Lush, 1971; Bridges & Hussain, 1992). In this chapter, we shall be concerned with two questions:

1 Over what range of Froude and Rossby numbers does the Lighthill theory of aerodynamic sound generation, appropriately modified to incorporate the effects of background rotation, give a quantitatively accurate description of the gravity wave field radiated by vortical motions in the shallow water equations?

2 How does the gravity wave radiation depend on the Froude and Rossby numbers as the Froude number is increased beyond the regime in which the Lighthill theory is valid, and, specifically, what conditions must exist in the vortical flow to enable gravity waves with order-one fluxes to be generated?

A comparison of the theory with early experimental data was made by Lighthill (1954). In general, there was reasonable agreement between theory and experimental data. In accordance with equation (2.10), the power of the acoustic waves radiated from the jet scaled approximately as the eighth power of the jet speed, providing some support for the validity of the theory when applied to actual turbulent jet flows.

Subsequent experiments by Lush (1971) confirmed the U^8 scaling law for the radiated power, where U is the speed of the jet, in somewhat more convincing detail than the earlier experiments discussed by Lighthill (1954), but also observed that when the acoustic waves passed through the jet for several wavelengths, they do not obey Lighthill's scaling law. This raises doubts about the claim, implicit in Lighthill's original formulation of his theory, that eddies in compressible jets will tend to be incoherent over a typical acoustic wavelength, and will not interfere with the radiated acoustic waves. Much of the literature on the aeroacoustics of jets over the last 20 years has been concerned with this point (see, for example, Moore, 1977; Mankbadi & Liu, 1984).

Recently, Bridges & Hussain (1992) have returned to the issue of the detailed verification of the Lighthill theory in a jet undergoing shear instability, in which the entire sound generation

region is acoustically compact with respect to all but the shortest wavelengths of the radiated acoustic wave field. They present a detailed analysis of the jet flow, identifying vorticity as the most robust measurable quantity from which to reconstruct velocities in the jet, under the assumption that the Mach number in the jet is sufficiently low that the velocity field can be approximated by the incompressible velocity field due to a given vorticity distribution, and hence to effective quadrupole source term for the radiated sound field. Comparisons between measurements of the acoustic field and its reconstruction from the Lighthill theory of aerodynamic sound generation are exceptionally good, thus verifying that the compact source assumption can be applied successfully to actual jet flows. If the vortical source of the radiated sound field is well modelled by a compact source approximation, one can show that for axisymmetric vortical flows, an extinction angle occurs at 54^o to the axis of symmetry. In the experiments of Bridges & Hussain, the jet axis serves as the axis of symmetry, and the extinction angle is well captured in the experimental data. However, agreement between theory and observations is not perfect, since the theory also predicts a second extinction angle which is not observed in their experimental data. Bridges & Hussain suggest that noise associated with the nozzle from which the jet issues is the most likely cause of the discrepancy.

A set of experiments even more idealized than those of Bridges & Hussain, also designed to test the validity of the Lighthill aeroacoustic theory, has been carried out by Kambe & Minota (1983), and Minota & Kambe (1986), who investigate experimentally the sound generated by a pair of colliding vortex rings. These experiments allow an especially detailed investigation of the validity of the theory of sound generation, since the flows may be made localized and compact for sufficiently small Mach number, and there is somewhat more detailed knowledge of the velocity field than in the experiments of Bridges & Hussain (1992). Moreover, since the sound generation is most intense when the vortex rings are interacting, and the point of interaction is well separated from the position of the generation of the vortex rings, there is no difficulty equivalent to the "nozzle noise" problems encountered by Bridges & Hussain (1992). The radiated sound field is compared with that predicted by the analytical work of Kambe & Minota (1981). They find that there is good agreement between experiments and theory, provided viscosity is taken into account. This is clear from their plots of directivity, which are initially consistent with a compact quadrupolar source, but develop a strong isotropic component from an effective monopole, which is due to viscosity, as the experiments proceed

in time.

Although there has been much experimental work with three dimensional jet flows, there appears to have been comparatively little experimental study of gravity wave generation in two dimensional flows. The only study to have been reported in the literature was undertaken by Webster (1970), who verified that a jet in shallow water excites surface waves whose intensity scales with the seventh power of the jet speed for Froude numbers up to approximately 1.0, consistent with (2.11). The Froude number must also be large enough that the jet noise exceeds the noise generated by interaction between the fluid and the nozzle from which the jet is issued. This confirms that the Lighthill theory appears to work well for quite large Froude numbers in two dimensional flows, just as it works well for quite large Mach numbers in three dimensional flows – a fact which could not have been assumed *a priori*, since there is no reason to expect the domain of validity of a small Froude number approximation in two dimensions to be the same as the domain of validity of a small Mach number approximation in three dimensions

Recently, Lele & Ho (1993) have presented a numerical study of a compressible mixing layer in a two dimensional numerical model. The mixing layer is unstable, and the waves which grow on the edges of the mixing layer eventually develop into a train of coherent vortices. The simulations are performed in a domain which is periodic in the the direction of the initial flow, and has absorbing regions placed near the boundaries which are parallel to the initial flow. Since it can be shown that the Lighthill theory is formally valid in the limit of low Mach number, it might be asked whether the source term to be taken in the Lighthill theory should come directly from the velocity and density fields in the numerical simulation, or whether it should be constructed from the vorticity field, under the assumption of incompressible flow, as Bridges & Hussain did in their experimental study. Lele & Ho found that the agreement between the simulated radiated wave field and its reconstruction using the Lighthill theory was significantly better when actual velocity and density fields were used to compute the Lighthill source term than when the incompressible fields were used, and that the Lighthill reconstruction could work quite well at Mach numbers of up to about 0.6, at which they found errors in the acoustic pressure field predicted by the Lighthill theory of only about 10%. For Mach numbers up to 0.2, the errors were only about 1–2%. It seems that the the low Mach number approximation is violated in the source region at quite small Mach numbers, whereas the assumption that the source region is small compared with the scale of the acoustic waves generated, although

formally a low Mach number approximation itself, is valid over a much wider range of Mach numbers.

The basic conclusion from the studies discussed above is that, leaving aside subtleties associated with extensive jets and directivity of the radiated sound, all experiments to date have demonstrated good agreement with the Lighthill theory up to quite large values of the Mach number, with the intensity of the 'sound' scaling with the eighth power of the Mach number in three dimensions, and the seventh power in two dimensions. This implies that the sound radiated, although weak at low Mach numbers, will rise very sharply in intensity as the Mach number is increased. Moreover, when a complete knowledge of the source flow is available, as in the numerical experiments of Lele & Ho (1993), the acoustic sound field can be reconstructed with excellent accuracy for Mach numbers up to 0.6 from the Lighthill source term. They also found that the momentum flux in the acoustic wave region scaled as the sixth power of the Mach number, consistent with scaling arguments which can be derived for one dimensional wave radiation from a vortical source (see Ffowcs-Williams, 1969).

Although the Lighthill theory of aerodynamic sound generation is therefore apparently well established for simple flows and turbulent jet flows over a range of Mach numbers, there remains no numerical or experimental study of the corresponding gravity wave generation problem, for cases where the rotation of the reference frame should be taken into account.

Recently, McIntyre & Norton (1993) presented simulations of a forced polar vortex in a hemispherical shallow water model. The forcing is applied over a long time scale, so it does not generate gravity waves directly, but acts to disrupt the polar vortex, thereby generating gravity waves as a result of the subsequent vortical motions. They compared the shallow water simulations with simulations using balanced models of varying accuracy, and found that the amplitude of gravity waves generated appeared to be very low, even when the Froude number in the flow reached 0.7. Without further investigation, their results might appear to contradict the established body of literature on aerodynamic sound generation and the validity of the Lighthill (1952) theory, which predicts that the amplitude of gravity waves generated by vortical motions will increase rapidly with increasing Froude number.

The complexity of the flows simulated by McIntyre & Norton (1993) makes it difficult to analyse nature of the gravity wave generation process in detail for their simulations. In this chapter, we shall investigate only simple flows, in which the vortical field is comparatively well

understood. We shall use the shallow water equations on an f-plane. This way, large scale Rossby waves are avoided, and the potential vorticity will be non-uniform only in some finite region. A modified version of the Lighthill theory of aerodynamic sound generation can then be used to investigate the validity of the compact source approximation as the Froude number is increased.

The nature of gravity waves generated by vortical motions, and the domain of validity of the modified Lighthill theory, are not only of theoretical interest. Frequently, the passage of a jet streak at the level of the tropopause can give rise to a significant amount of gravity wave activity. Usually, the gravity waves generated will propagate vertically upwards into the stratosphere. Occasionally the tropospheric conditions can act as a wave guide, and a significant amount of the waves generated by the passage of the jet streak remain trapped in the troposphere.

Such an event was observed and analysed by Koch & Dorian (1988) using a mesoscale observing network. In addition to presenting an analysis of the gravity wave field, they investigated possible source mechanisms. After ruling out convection as a possible source, they concluded that the gravity waves were generated either as a result of stratified-shear-flow instability, or through nonlinear interaction with the unsteady vortical motions associated with the jet streak itself. They were unable to demonstrate conclusively which of these two possible mechanisms was responsible for the observed waves. However, if we can gain some more insight into how to apply and interpret the modified Lighthill theory, and have some knowledge of its domain of validity as the Froude number is increased, we may lay the foundations for the development of a test to determine whether the amplitude and form of the observed waves is consistent with their being generated by coupling with the vortex dynamics. Gravity waves of long wavelength and period, apparently generated by vortical motions, have also been observed in simulations of baroclinic lifecycles by D. J. O'Sullivan (personal communication).

Although we may bear in mind these applications as a motivation for the present study, I have chosen not to concentrate on any particular flow which might be observed in the atmosphere, but rather to concentrate on flows which can be analysed and interpreted readily. I have therefore required three things of the idealized experiments to be discussed in this chapter:

1 Any gravity waves present in the simulation must be unequivocally due to vortical motions, and should not be due to any forcing or initialization procedure.

2 The simulations should depend on the minimum number of adjustable parameters, which should be related to the Froude number and the Rossby number.

3 Whatever mechanism is used to generate the vortical flow should apply in the limit of low Froude number, so that in this limit the modified Lighthill theory of gravity wave generation may be tested.

Of these requirements, it is undoubtedly the first requirement which places the most severe restrictions on the nature of the initial flow whose subsequent evolution we may study. Suppose we wish to initialize our simulation with an arbitrary potential vorticity field. A small amount of gravity wave activity is inevitable, due to the non-existence of a true slow manifold for the shallow water equations, established in the previous chapter. How much gravity wave activity there is depends inevitably on which balanced model we choose. In any case, even if the very minimum of gravity waves are introduced by the initialization procedure, we will be in the position of having gravity waves present from the initial instant in our simulation, but no knowledge of the flow which is supposed to have generated them. This is a totally unsuitable situation for testing the validity of a modified Lighthill theory.

The only flows from which we can unambiguously eliminate gravity waves are the steady solutions of the shallow water equations – parallel flow and axisymmetric flow. Since these are steady flows, they will never develop gravity waves. However, if they are unstable, then an arbitrarily small disturbance may develop into a vigorous vortical motion from which gravity waves may be radiated. From the viewpoint of numerical simulation, it is simpler to work with initially parallel flow than initially axisymmetric flow, and that is the sole focus of the study presented here.

The most familiar form of parallel flow instability is the so-called "barotropic" shear flow instability, described, for example, in Hoskins *et al.* (1985). This instability can occur when there is a reversal in the sign of the potential vorticity gradient across the flow. This means that there are potential vorticity gradients of both signs in the flow. Physically, the direction of the potential vorticity gradient sets the direction of Rossby wave propagation, and therefore we can have counter-propagating Rossby waves in the flow, propagating on potential vorticity gradients of opposite sign. For some critical values of wavelength, these two counter-propagating waves can lock onto each other, and thereby cause each other to grow. For this reason, the instability

is sometimes referred to as "Rossby wave – Rossby wave" instability.

An extensive review of shear flow instability, including barotropic shear flow instability, was given by Ho & Huerre (1984). As the instability develops, the potential vorticity rolls up into a train of vortices, spatially periodic in the streamwise direction. Many previous studies have addressed the subsequent merging of adjacent vortices as subharmonics of the fundamental wavelength of the instability develop (see Ho & Huerre (*op. cit.*) and refs. therein). In this study, we are concerned with the degree to which the vortex motions, following the roll up of the primary instability, will excite gravity waves which radiate away from the shear layer. To enable a thorough investigation of the Froude and Rossby number parameter space, I have chosen to concentrate resources on simulating just one wavelength of the primary instability. It follows that subharmonics do not develop, and so the extent to which the vortex merging process excites gravity waves has not been investigated.

An alternative mechanism for parallel flow instability in the shallow water equations was presented in chapter 3, and can be regarded as a "Rossby wave – gravity wave" instability. However, there are three reasons for preferring to use barotropic instability for generating the vortical motions which will excite gravity waves in our experiments.

1 Typical growth rates of Rossby wave – gravity wave instability are very small, so the instability will take a long time to develop, and might saturate at low amplitude.

2 The Rossby number and Froude number are not independent parameters for an initial flow with a single discontinuity of or sharp jump in potential vorticity, so it is not possible to vary them independently in simulations initialized in this way

3 The Rossby wave – gravity wave instability is not present at arbitrarily small Froude number, so rendering a comparison with the modified Lighthill theory impossible in the small Froude number limit in which the theory is formally valid.

For these reasons, all the numerical simulations presented in this chapter are simulations of barotropically unstable parallel flows. Initially, the potential vorticity is everywhere uniform, equal to Q_0, say, except in some strip of finite width. The velocity tends to zero far from the strip, and the height tends to a uniform value. Inside the strip, the potential vorticity takes a different uniform value, Q_1. The potential vorticity jump is smoothed over at least 5 grid

points in the numerical model. The ratio of the difference between the strip and background potential vorticities, $Q_1 - Q_0$, to the background value, Q_0, defines an effective Rossby number $Ro = Q_1/Q_0 - 1$ for the flow. As the strip width is increased, so the velocities in the strip increase, and hence the Froude number is governed by the width of the strip.

The rest of this chapter is organized as follows.

In §4.2 the Lighthill analysis for the shallow water equations is developed, as appropriate for periodic jets.

In §4.3 the pseudomomentum and pseudoenergy are introduced as possible diagnostic quantities for quantifying the intensity of gravity wave radiation, and it is shown that the pseudoenergy flux is the natural wave quantity to choose in this study.

In §4.4 the numerical model to be used for the nonlinear simulations is described, and an overview of the simulations performed with the model is presented in §4.5.

In §4.6, the gravity wave radiation from a train of vortices generated by the roll up of a cyclonic strip of potential vorticity is discussed. The strip has potential vorticity equal to six times the background value.

In §4.7, the effect of increasing the potential vorticity in the strip is investigated.

In §4.8, anticyclonic strips are investigated, with potential vorticity in the strips equal to 0.1 and 0.0 times the background value.

In §4.9, the effects of negative potential vorticity is investigated, with potential vorticities in the strip of -0.1, -1.0 and -19.0 times the background value.

Some conclusions are offered in §4.10.

4.2 A "Lighthill" theory of gravity wave generation for a periodic parallel flow

In this section the analysis necessary to investigate the quantitative accuracy of the Lighthill theory, applied to periodic flows in a rotating frame, is developed. In further sections the analysis will be used to investigate the degree to which the theory remains useful as the Froude number is increased and the compact source assumption becomes less valid.

Returning to equation (2.4), we recall that manipulation of the shallow water equations gives

$$\left(\frac{\partial^2}{\partial t^2} + f^2 - c_0^2\nabla^2\right)\frac{\partial h}{\partial t} = \frac{\partial^2}{\partial x_i \partial x_j}T_{ij}, \tag{4.1}$$

where

$$T_{ij} = \frac{\partial}{\partial t}(hu_iu_j) + f\varepsilon_{ik}hu_ju_k + \frac{1}{2}g\frac{\partial}{\partial t}(h-h_0)^2\delta_{ij}, \tag{4.2}$$

h_0 is the layer depth far from the region of vortical motion, and $c_0 = (gh_0)^{1/2}$ is the gravity wave phase speed far from the region of vortical motion. These equations have formal solution

$$\frac{\partial h}{\partial t} = \frac{1}{4\pi c_0}\frac{\partial^2}{\partial x_i \partial x_j}\int_{t-t'>|\boldsymbol{x}-\boldsymbol{x}'|/c_0} \frac{T_{ij}(\boldsymbol{x}',t')\cos(f((t-t')^2 - |\boldsymbol{x}-\boldsymbol{x}'|^2/c_0^2)^{1/2})}{((t-t')^2 - |\boldsymbol{x}-\boldsymbol{x}'|^2/c_0^2)^{1/2}} d^2\boldsymbol{x}'dt'. \tag{4.3}$$

The key point in Lighthill's theory is now to assume that the source term T_{ij} is only non-zero over a small enough region that it may be approximated by a quadrupole point source, so that we may write the solution as

$$\frac{\partial h}{\partial t} = \frac{1}{4\pi c_0}\frac{\partial^2}{\partial x_i \partial x_j}\int_{t-t'>|\boldsymbol{x}|/c_0} \frac{S_{ij}(t')\cos(f((t-t')^2 - |\boldsymbol{x}|^2/c_0^2)^{1/2})}{((t-t')^2 - |\boldsymbol{x}|/c_0^2)^{1/2}} dt', \tag{4.4}$$

where

$$S_{ij}(t') = \int_{\mathbb{R}^2} T_{ij}(\boldsymbol{x}',t')d^2\boldsymbol{x}'. \tag{4.5}$$

In the classical problem of aerodynamic sound generation, this is the compact source approximation – that the length-scale of the source is small compared with the length-scale of the waves. Here we are imposing an additional condition – that the length-scale of the source is also small compared with a Rossby deformation radius.

If we now assume that we know the way in which the source term scales as the flow parameters are varied, then we may estimate how the amplitude of the wave field depends on parameters such as the Froude number and the Rossby number. Lighthill (1954) proposed that the theory could be applied to an extensive body of turbulence by assuming that each turbulent eddy was itself compact, and not correlated with other eddies.

It is clearly not possible to apply this analysis directly to a radiating vortex train, since the eddies in the train are highly correlated with each other. We must therefore develop a version of the theory specifically for the case where the source, and therefore the waves, are periodic in the x direction.

In comparing the gravity wave radiation with that which is predicted by the Lighthill theory, we shall concentrate on that component of the wave field which is independent of the streamwise coordinate x. In the asymptotic limit of small Froude number, it can be shown by the method of matched asymptotic expansions that all other components of the radiated wave field are exponentially small compared with the x–independent component, and so this seems the natural component on which to focus the analysis.

Now, let the x–average of a quantity a be denoted by $\overline{a}^x$. To derive the effective source for the x–independent component of the wave field, we must take the x–average of equation (4.1). The x derivatives, in both the wave operator and the quadrupole term, then vanish, and we are left with a one-dimensional Klein-Gordon equation, with a one-dimensional quadrupole source term:

$$\left(\frac{\partial^2}{\partial t^2} + f^2 - \frac{1}{c_0^2}\frac{\partial^2}{\partial y^2}\right)\overline{\frac{\partial h}{\partial t}}^x = \frac{\partial^2}{\partial y^2}\overline{T_{22}}^x. \tag{4.6}$$

To proceed, we need only assume that the source term $\overline{T_{22}}^x$ is compact with respect to the cross stream direction y. Making that assumption, we obtain

$$\overline{\frac{\partial h}{\partial t}}^x(y,t) = \frac{1}{2c_0}\frac{d^2}{dy^2}\int_{-\infty}^{t-|y|/c_0} S(t')\mathrm{J}_0(\sqrt{((t-t')^2 - y^2/c_0^2)})dt', \tag{4.7}$$

where

$$S(t') = \int_{-\infty}^{\infty} \overline{T_{22}}^x(y',t')dy', \tag{4.8}$$

and J_0 is the regular Bessel function of order zero (see Morse & Feshbach, 1953). The factor $\frac{1}{2}$ arises from the fact that waves propagate in both positive and negative y directions away from the source at $y = 0$.

To reconstruct the wave field from a numerical simulation using equation (4.7), we now need only know the function $S(t)$, a single function of time, which is obtained for each time t during the simulation by evaluating the integral (4.8) over the entire flow.

We may readily proceed from here to show how $\partial h/\partial t$ in the wave region scales with properties of the vortical flow in the vortex train. Taking the Fourier transform of (4.6) with respect to t, we have, making the compact source approximation,

$$i\omega\tilde{h} \sim \frac{\sqrt{\omega^2 - f^2}}{2c_0}e^{i\sqrt{\omega^2-f^2}y/c_0}\int \tilde{\overline{T_{22}}}^x(y',\omega)dy', \tag{4.9}$$

where, as usual, ω is the transform variable. Now, scale analysis of (4.2) gives a scaling for T_{22}: $\widetilde{\overline{T_{22}}^x}(y',\omega) \sim \omega u^2 \lambda \tilde{h_0}$, where λ is the length scale of the vortical region, and hence, for $\omega^2 > f^2$,

$$h' \sim h_0 \left(\frac{u}{c_0}\right)^2 \left(\frac{\left(\sqrt{\omega^2 - f^2}\right)\lambda}{c_0}\right). \tag{4.10}$$

4.3 Pseudoenergy and pseudomomentum for the shallow water equations

Since we are interested in the amplitude of gravity wave radiation over a wide range of Froude and Rossby numbers, it is desirable to take the flux of an appropriate wave quantity, such as pseudoenergy or pseudomomentum, as a measure of the strength of gravity wave radiation by the vortical flow.

In general, a wave activity density A is a field, second order in disturbance amplitude in the limit of small disturbances, for which there exists an associated wave flux $\boldsymbol{F}$, such that

$$\frac{\partial A}{\partial t} + \nabla.\boldsymbol{F} = 0. \tag{4.11}$$

Definition of wave activities and their corresponding fluxes must be made with respect to some reference state. In this chapter, we shall take the initial parallel flow as the reference state for the definition of wave activities and their fluxes. Throughout this section, $\boldsymbol{u}_0$, h_0 and Q_0 represent the velocity, height and potential vorticity of the basic state respectively, and $\boldsymbol{u}'$ and h' represent departures of velocity and height from the basic state values. The basic state possesses symmetries with respect to x-translation and time translation, and therefore it is possible to define both pseudoenergy and x–pseudomomentum fluxes. For the present study, we are interested in developing a wave quantity which can be used to characterise the strength of gravity wave radiation. If we are to take a single quantity, such as a wave flux strength, to characterise gravity wave radiation, then that quantity should satisfy a monotonic radiation property – that is, all radiating waves should give rise to fluxes of the same sign, at least in the limit of linear waves. It turns out that this "monotonic radiation property" is not satisfied by the pseudomomentum, but is satisfied by the pseudoenergy.

The technique for obtaining wave activities and fluxes is now quite well established (McIntyre & Shepherd, 1987; Haynes, 1988), and the details of the calculations of the pseudoenergy and

pseudomomentum for the shallow water equations will not be reproduced here.

The pseudomomentum density for the shallow water equations is given by

$$A_m = h'u' + h\int_{y_0}^{y} (Q - Q_0(y'))\, h_0(y')dy' \tag{4.12}$$

and its flux is given by

$$\boldsymbol{F}_m = \boldsymbol{u}A_m + \left(\frac{1}{2}h_0\left(u'^2 - v'^2\right) + \frac{1}{2}gh'^2, u'v'h_0\right), \tag{4.13}$$

so that

$$\frac{\partial A_m}{\partial t} + \nabla.\boldsymbol{F}_m = 0. \tag{4.14}$$

The pseudoenergy density is

$$A_e = \frac{1}{2}h\left(u'^2 + v'^2\right) + u_0u'h' + \frac{1}{2}gh'^2 + h\int_{y_0}^{y} (Q - Q_0(y'))\, h_0(y')u_0(y')dy' \tag{4.15}$$

and its flux is

$$\boldsymbol{F}_e = \boldsymbol{u}A_e + h_0\boldsymbol{u}_0.\boldsymbol{u}'\boldsymbol{u}' - \frac{1}{2}h_0\boldsymbol{u}'^2\boldsymbol{u}_0 + \frac{1}{2}gh'^2\boldsymbol{u} + gh_0h'\boldsymbol{u}'. \tag{4.16}$$

In (4.12) and (4.15), y_0 is the initial y location of the fluid in the reference state which is at location (x, y) at time t at which the wave activity or flux is to be evaluated. The integrals in (4.12) and (4.15) represent "Lagrangian information" about the motion of fluid particles. Notice, however, that they will be non-zero only if the potential vorticity in the basic state is non-uniform between y and y_0, implying that Rossby waves exist at the location y. In this chapter, we shall be concerned with gravity wave fluxes in a region of uniform potential vorticity. In computing the fluxes far from the vortical region, therefore, the contribution to (4.12) and (4.15) from the potential vorticity integrals will be identically zero.

To investigate the properties of the pseudoenergy and pseudomomentum fluxes in the far field, we consider the linearized shallow water momentum equations about a state of no motion. We assume that disturbances are wavelike, of the form $e^{i(kx+ly-\omega t)}$. The equations are

$$-i\omega u' - fv' + ikgh' = 0 \tag{4.17}$$

$$-i\omega v' + fu' + ilgh' = 0. \tag{4.18}$$

Considering the quadratic terms in the pseudoenergy flux in the limit where $u_0 = 0$, we see that the y-component of the pseudoenergy flux is dominated by the term $h_0h'v'$. To evaluate

$h'v'$ we eliminate u from (4.17-4.18). This gives

$$gh' = \frac{\omega^2 - f^2}{\omega^2 l^2 + f^2 k^2}(\omega l + ikf)v'. \tag{4.19}$$

Averaging over wave periods gives

$$g\overline{h'v'} = \frac{\omega^2 - f^2}{\omega^2 l^2 + f^2 k^2}\frac{\omega l}{2}|v'|^2. \tag{4.20}$$

For the pseudomomentum, we see that in the far field the y-component of the pseudomomentum flux is dominated by $h_0 u'v'$. So, to analyse the pseudomomentum flux in the far field, we eliminate h' from (4.17-4.18). This leads to

$$u' = \frac{(\omega^2 - f^2)kl + i\omega f(k^2 + l^2)}{\omega^2 l^2 + f^2 k^2}v'. \tag{4.21}$$

Averaging over wave periods here gives

$$\overline{u'v'} = \frac{\omega^2 - f^2}{\omega^2 l^2 + f^2 k^2}\frac{kl}{2}|v'|^2. \tag{4.22}$$

To check for radiation properties, we consider the dispersion relation for gravity waves

$$\omega^2 - f^2 = c_0^2(k^2 + l^2), \tag{4.23}$$

where $c_0 = \sqrt{gh_0}$.

We wish to impose a radiation condition with respect to the propagation in the y direction, so that

$$\frac{\partial \omega}{\partial l} \equiv c_0^2 \frac{l}{\omega}\begin{cases} > 0 & \text{for } y > 0 \\ < 0 & \text{for } y < 0 \end{cases}. \tag{4.24}$$

It follows that l and ω should be of the same sign as $y \to +\infty$, but should be of opposite sign as $y \to -\infty$.

Now we want to select the wave flux such that $\boldsymbol{F}.\boldsymbol{n}$ is the same sign on the upper and lower y–boundaries, recalling that the normal is in the positive y direction for $y > 0$, and is in the negative y direction for $y < 0$. This forces us to choose the pseudoenergy flux, since $h_0\overline{h'v'}$ takes the same sign as ωl provided $\omega^2 > f^2$, so that the waves are radiating. If $\omega^2 < f^2$ then l is imaginary, the waves are evanescent, and there is no wave flux.

It follows that, with the present sign convention for pseudoenergy, there should be a net flux of pseudoenergy in time away from the jet due to gravity wave radiation, whereas there is no corresponding monotonic radiation property for the flux of pseudomomentum.

To see how the intensity of the pseudoenergy flux scales with the Froude number, we note that

$$F_e = h_0 g h' v' = \frac{g^2 h_0 \omega l}{\omega^2 - f^2} h'^2. \tag{4.25}$$

From (4.10), this implies

$$F_e \sim h_0 c_0^2 \left(\frac{u}{c_0}\right)^4 \lambda^2 \omega l. \tag{4.26}$$

From the dispersion relation (4.23), with $k = 0$, we therefore obtain

$$F_e \sim \begin{cases} h_0 c_0 \left(\frac{u}{c_0}\right)^4 \lambda^2 \omega \sqrt{\omega^2 - f^2} & \omega^2 > f^2 \\ 0 & \omega^2 < f^2 \end{cases} . \tag{4.27}$$

Thus if, at low Froude numbers, we assume that the velocity u scales as $u \sim \omega\lambda$, and the Froude number $F \sim \omega\lambda/c_0$, it follws that

$$F_e \sim \begin{cases} h_0 c_0^3 F^6 \sqrt{1 - f^2/\omega^2} & \omega^2 > f^2 \\ 0 & \omega^2 < f^2 \end{cases} . \tag{4.28}$$

Thus, if ω, c_0 and h_0 are held fixed, and the Froude number is increased by increasing λ, the scale of the vortical motions, the intensity of the pseudoenergy flux will increase as the sixth power of the Froude number, in agreement with the analysis of one-dimensional sound radiation by vortical motions presented by Ffowcs-Williams (1969). However, if the Froude number is increased by increasing the magnitude of the vorticity in the vortical region, and therefore increasing $|\omega|$, no such simple scaling law applies. In general, for a given Froude number, the intensity of wave fluxes will be greatest when $\omega^2 \gg f^2$, and the f^2/ω^2 term in (4.28) becomes insignificant.

4.4 The numerical model

The numerical model used in this chapter integrates the nonlinear f-plane shallow water equations in a channel geometry. The channel is periodic in the x-direction, and has solid boundaries at $y = \pm Y$, for some (usually large) Y. The shallow water equations are integrated in nondimensional form, viz:

$$\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u}.\nabla\boldsymbol{u} + \boldsymbol{k} \times \boldsymbol{u} + \nabla h = 0 \tag{4.29}$$

$$\frac{\partial h}{\partial t} + \nabla.(h\boldsymbol{u}) = 0, \tag{4.30}$$

where $\boldsymbol{u}$ is the two-dimensional velocity field, h is the layer depth, and $\boldsymbol{k}$ is the unit vector normal to the plane. The nondimensionalization sets the length scales and time scales for the flow, such that the unit timescale is the inertial period f^{-1}, and the unit length scale is the Rossby deformation radius where the fluid is at rest $\sqrt{gh_0}/f$. In (4.29) and (4.30), h has been rescaled so that g may be set to unity. The potential vorticity away from the region of vortical motions will therefore be unity. In subsequent sections, all results of model integrations will be presented using this nondimensionalization of length and time scales.

Several numerical schemes for integrating the shallow water equations have been presented in the literature, of which the best known is that of Arakawa & Lamb (1981). A possible advantage of the Arakawa & Lamb scheme is that it can be shown to conserve energy and enstrophy exactly. The proof relies on an assumption that a regularly-spaced grid is used to discretize the equations. In this chapter, however, we are interested in flows which will generally consist of a highly nonlinear but highly localized region of vortical motion, from which low amplitude gravity waves are radiated, with a wavelength much longer than the scale of the vortical region. In the interests of computational efficiency, the model used in this chapter will therefore allow variable resolution in both the streamwise and cross-stream directions. There is therefore no particular reason to favour a scheme such as that of Arakawa & Lamb, and the model used here was developed so as to allow the fields with variable resolution to be represented in a convenient way, but not to possess any particular conservation properties.

The geometry is periodic in x, but of finite extent in y. Therefore, the equations were discretized using finite differences in the y direction, with fields at each value of y being represented as a sum of Fourier modes in x. The y-discretization is performed on a staggered grid, with u, v on physical grid levels, and h at intermediate levels. The model is a pseudo-spectral model, with streamwise derivatives being computed in spectral space, and nonlinear products computed in physical space.

The variable cross stream resolution was implemented using a coordinate transformation. The numerical model is written in terms of a cross-stream variable η. The grid positions are equally spaced in the model coordinate η, and the physical coordinate transformation is implemented by specifying a relationship between the model coordinate η and the physical

coordinate y. Cross-stream differencing is performed on the η grid, where centred differencing can be used. By a simple application of the chain rule, the centred differences are then multiplied by $d\eta/dy$ to give the value of the derivative with respect to y. In the vortical region the value of $d\eta/dy$ is unity, whereas in the wave region its value is generally less than unity. In the region of variable resolution, a self-similar transformation is employed, in which $d\eta/dy = e^{-a(\eta-\eta_1)}$, where η_1 is the outer limit of the region of high resolution, and a is a constant which sets the size of the region of variable cross-stream resolution.

The variable streamwise resolution is implemented in a very simple way. The streamwise resolution is specified in the vortical region by specifying a maximum number of Fourier components which are to be used to represent the flow there. A full set of Fourier harmonics is retained throughout the vortical region, and in the region of varying cross-stream resolution. Beginning at the start of the region of low resolution, the number of Fourier components used to represent the solution is then reduced by a factor of two at every fifth cross-stream grid point, until a specified minimum number of streamwise components are left. This number of components is then used to represent the solution in the region of low resolution. In all simulations presented in this chapter, 64 Fourier coefficients in the x direction were used to resolve the vortical region.

Several shallow water models have been presented which use vorticity, divergence and height as prognostic variables (e.g. Bourke, 1972). They have the advantage of a flux formulation with only quadratic nonlinearity, ensuring that quadratic quantities are conserved (but not energy, which is cubic in the shallow water system) when integrated using the leapfrog timestepping scheme. However, these models require inversion of the Laplacian operator at each time step, and the velocity field at one location depends on the vorticity, divergence and height at all locations in the model. This seems an undesirable property for a model when a large domain, with regions of greatly differing character, is to be integrated. The present model uses the two components of velocity and height as the prognostic variables. Therefore, the model does not require inversion of elliptical operators to obtain the velocity and height fields and, although it is not a conservative scheme, there is no reason to expect it to be less accurate locally than a model which uses a flux formulation.

Some diffusion is required to remove fine scales, especially in the vortical region. This was implemented by applying a small amount of hyperdiffusion to the Fourier components in each

of the fields u, v and h. The hyperdiffusion is proportional to k^6, where k is the wave number. Since the hyperdiffusion acts only in the x-direction, its anisotropic nature may make it appear to be a particularly unphysical type of hyperdiffusion. However, the rotational nature of vortical flow means that it should be sufficient to apply a diffusion in the x-direction only. Any fine scale structure which builds up in the cross-stream direction in the vortical flow will rotate into a streamwise orientation, where it will be removed by the hyperdiffusion. This type of hyperdiffusion was used in a model with streamwise spectral representation and cross stream grid-point representation in a numerical study of nonlinear critical layer evolution by Haynes (1989), and appears to control fine scale structure adequately for the present purposes. One advantage is that it is quite simple to implement, and in particular may be implemented in an implicit diffusion scheme using only division, and not tridiagonal matrix inversion, as would be required for an isotropic diffusion with a grid point discretization in the cross stream direction.

At the lateral boundaries, sponge layers may be placed over a variable number of grid levels. Rayleigh friction is applied there, which relaxes the flow back to the initial state. The value of the Rayleigh friction is small at the point where it is first encountered by waves propagating towards the boundary, and increases linearly towards the boundary. In the model, the number of grid levels over which the Rayleigh friction is applied, and its amplitude, are arbitrary. However, for all simulations presented here, it was found that applying Rayleigh friction over 30 grid levels, with a maximum value of 1.0 at the boundary, corresponding to an e-folding decay time of 1.0 for disturbances at the boundary, generally prevented any significant reflection from the boundary, and these values were used throughout.

The equations are integrated forward in time using an explicit leapfrog time step, but with the hyperdiffusion and Rayleigh friction being performed implicitly. The basic leapfrog method is unstable, however, and a time filter must be used to prevent the growth of a rapidly oscillating computational mode. In the present simulations, a Robert-Asselin time filter (Robert, 1966; Asselin, 1972) was used with a value of 0.02.

4.5 Overview of numerical simulations

Using the numerical model described in §4.4, gravity wave generation by vortical motions in the shallow water equations is now studied by numerical simulation of the nonlinear evolution

of a barotropically unstable strip of potential vorticity. The potential vorticity in the strip is Q_1, and the background value of the potential vorticity is unity. The Rossby number Ro can then be defined as $Ro = Q_1 - 1$. The only other significant parameter in the initial flow is the width of the strip, Δ, which determines the Froude number of the flow.

In the following sections, simulations with a range of values of Q_1 and Δ are discussed. The simulations are labelled by letters **A** – **H**, with each letter corresponding to a different value of Q_1. Simulations with the same value of Q_1 are distinguished by Roman numerals (e.g. **A**i–vi), with the Roman numeral increasing with increasing initial strip width. Table 4.1a shows the values of Q_1 and Δ investigated, with their corresponding labels.

In all cases, the initial conditions for the numerical simulation consist of a strip of potential vorticity Q_1 centred on $y = 0$. The model domain in the cross-stream direction extends from $y = -Y$ to $y = +Y$, for some large Y, to accomodate several wavelengths of the radiating gravity waves within the computational domain. Only one wavelength of the fastest growing unstable linear eigenmode is simulated in all cases. The wavenumber of the fastest growing eigenmode is shown in table 4.1b, and a matrix method is used to obtain the eigenmode. In all cases except simulations **C**i–iv and **H**i–iv, a small amplitude disturbance of the form of the fastest growing eigenmode is added to the initial parallel flow at $t = 0$. If a large amplitude of the fastest growing eigenmode is added, then the flow will adjust, radiating gravity waves. The amplitude of the eigenmode added was therefore controlled such that any gravity waves generated by this initial adjustment had an amplitude of no more than 1–2% of the gravity waves subsequently generated by the vortical motions, when viewed in the $\partial h/\partial t$ field. In cases **C**i–iv and **H**i–iv, the strips are very narrow compared with a deformation radius. Consequently, the eigenmode decays very slowly with distance away from the strip, when compared with the width of the strip. This presents very large memory requirements for the matrix method used to find the eigenmode, and in these cases the flow was disturbed by the addition of random noise. Again, the amplitude of the random noise added was chosen to be sufficiently small that any gravity wave radiation associated with it was of small amplitude compared with the gravity waves subsequently generated by the vortical motions.

In the nonlinear evolution of the flow, the strip of potential vorticity rolls up into a train of vortices, which then nutate, radiating gravity waves as they do so. In general, it was found that the maximum value of the Froude number, found in the vortical region, can increase by

(a)

Simulation	*Ro*	Initial strip width i	ii	iii	iv	v	vi
A	5	0.018	0.035	0.07	0.14	0.21	0.28
B	8	0.018	0.035	0.07	0.21		
C	20	0.011	0.021	0.041	0.081		
D	-0.9	0.42	0.84	1.68	3.36	6.00	9.00
E	-1.0	0.42	0.84	1.26	2.52	6.00	
F	-1.1	0.42	0.84	1.26	1.68	2.52	3.04
G	-2	0.07	0.21	0.42	0.63		
H	-20	0.007	0.013	0.021	0.031		

(b)

Simulation	*Ro*	Wavenumber of fastest growing mode i	ii	iii	iv	v	vi
A	5	28.0	14.0	7.0	4.2	3.05	2.3
B	8	28.0	13.75	6.6	2.6		
C	20	50.0	30.0	14.0	5.6		
D	-0.9	2.0	1.15	0.75	0.475	0.245	0.165
E	-1.0	2.0	1.15	0.9	0.55	0.255	
F	-1.1	2.0	1.15	0.9	0.75	0.55	0.47
G	-2	7.2	3.5	1.85	1.2		
H	-20	65.0	44.0	30.0	20.0		

(c)

Simulation	*Ro*	Froude number i	ii	iii	iv	v	vi
A	5	0.11	0.20	0.36	0.50	0.59	0.66
B	8	0.17	0.32	0.53	0.84		
C	20	0.23	0.37	0.66	1.06		
D	-0.9	0.25	0.44	0.64	0.88	1.05	1.08
E	-1.0	0.30	0.51	0.64	0.99	1.78	
F	-1.1	0.33	0.58	0.75	0.93	1.36	1.92
G	-2	0.17	0.38	0.73	1.31		
H	-20	0.20	0.31	0.49	0.75		

Table 4.1: Initial strip widths, fastest growing wave numbers and Froude numbers for all simulations.

as much as 50% during the initial development of the instability, and then tends to fluctuate as the vortices nutate. In subsequent sections, we shall consider how the amplitude of the flux of the radiated gravity waves scales with the Froude number, with reference to the predictions made in §4.2, and in particular equation (4.28). Therefore, as a measure of the Froude number of the flow, we shall require a Froude number associated with the gravity wave generation phase of the flow, rather than the initial roll-up phase. For this reason, the maximum value of the Froude number over all times was taken to characterize the Froude number of the flow, and for each simulation this value is shown in table 4.1c.

With the values of Q_1 and Δ selected, the remaining parameter which must be adjusted is the hyperviscosity. Since it has no physical meaning, it was selected by experimentation with the simulation **A**iii, with the aim of finding the minimum value which prevented an unacceptable build-up of noise on small scales in the potential vorticity field. On the other hand, it was important to ensure that the vortices did not become axisymmetric too quickly as a result of excessive hyperdiffusion, and thereby lose their ability to radiate gravity waves. It was found that 64 Fourier coefficients were required to represent the fields at each cross stream grid level if the required level of hyperdiffusion was not to be so great as to cause the vortices to become axisymmetric after one or two nutations. With the value of the hyperviscosity ν selected, it was then adjusted between simulations so as keep $\nu k_{\max}^6 \times (\delta Q)$ constant, where $k_{\max}$ is the maximum wavenumber in the simulation, and (δQ) is the magnitude of the potential vorticity difference between the vortices and the surroundings, meaning that in all simulations it takes the same amount of time to damp out noise in the highest wave numbers on typical vortex dynamical timescales, set by the magnitude of the potential vorticity variations. For $k = 1.0$ and $(\delta Q) = 1.0$, the value of ν used is $\nu = 6.4 \times 10^{-9}$.

To accommodate the wide range of strip widths (see table 4.1a), the grid-point spacing in the cross-stream direction had to be changed between simulations. In all cases, the region of uniform potential vorticity in the strip was distributed over at least seven grid intervals, and the potential vorticity was adjusted to the background value over five grid intervals. Details of the cross-stream resolution for all simulations are given in table 4.2. For the streamwise resolution, the 64 Fourier coefficients which were used in the vortical region were frequently more than sufficient to resolve the flow in the wave region. Therefore, in addition to varying the cross-stream resolution, the streamwise resolution was also varied, as described in §4.4. The

Simulation		$(\Delta y)^i$	Ny^i	$(\Delta y)^e$	Ny^t	Nx^e	Δt
A	i	0.0025	150	0.0625	1000	8	0.000625
	ii	0.005	150	0.125	600	8	0.00125
	iii	0.01	350	0.2	1000	8	0.0025
	iv	0.01	350	0.2	1000	8	0.0025
	v	0.01	350	0.2	1000	8	0.0025
	vi	0.01	350	0.2	1000	8	0.0025
B	i	0.0025	150	0.0625	1000	8	0.000625
	ii	0.005	150	0.125	600	8	0.00125
	iii	0.01	350	0.2	1000	8	0.0025
	iv	0.01	350	0.2	1000	8	0.0025
C	i	0.001	350	0.02	1000	8	0.00025
	ii	0.001	350	0.02	1000	8	0.00025
	iii	0.001	350	0.02	1000	8	0.00025
	iv	0.004	350	0.04	1000	16	0.001
D	i	0.02	350	0.2	1000	8	0.005
	ii	0.04	350	0.2	1000	8	0.01
	iii	0.04	350	0.2	1000	16	0.01
	iv	0.04	500	0.2	1000	16	0.005
	v	0.2	600	0.2	600	64	0.025
	vi	0.2	600	0.2	600	64	0.025
E	i	0.02	350	0.2	1000	8	0.005
	ii	0.04	350	0.2	1000	8	0.01
	iii	0.04	350	0.2	1000	16	0.01
	iv	0.04	425	0.2	1000	16	0.005
	v	0.2	600	0.2	600	64	0.0125
F	i	0.02	350	0.2	1000	8	0.005
	ii	0.04	350	0.2	1000	8	0.01
	iii	0.04	350	0.2	1000	16	0.01
	iv	0.04	350	0.2	1000	16	0.005
	v	0.04	2000	0.04	2000	64	0.005
	vi	0.04	1000	0.04	1000	64	0.0025
G	i	0.01	350	0.2	1000	8	0.0025
	ii	0.01	350	0.2	1000	8	0.0025
	iii	0.02	350	0.2	1000	8	0.005
	iv	0.03	2000	0.03	2000	64	0.005
H	i	0.001	350	0.025	1000	8	0.00025
	ii	0.001	350	0.02	1000	8	0.00025
	iii	0.001	350	0.02	1000	8	0.00025
	iv	0.001	350	0.02	1000	16	0.000125

Table 4.2: Details of numerical resolution used in all experiments. $(\Delta y)^i$ is the cross-stream grid spacing in the vortical region; Ny^i is the number of cross-stream gridpoints in the vortical region; $(\Delta y)^e$ is the cross-stream grid spacing in the wave region; Ny^t is the total number of cross-stream gridpoints; Nx^e is the number of Fourier coefficients in the streamwise direction used to resolve the wave region; Δt is the timestep.

number of Fourier coefficients used in the wave region are given in table 4.2.

4.6 Jets with moderate Rossby number and variable Froude number

In the first set of numerical experiments, to be described in this section, the Rossby number is fixed, and the effect of varying the Froude number is investigated. In all cases to be described in this section, therefore, the model is initialized with a strip in which the potential vorticity is equal to six times the background value. This can be regarded as fixing the Rossby number at 5. The Froude number then is varied by varying the width of the strip.

Six simulations were performed with the potential vorticity in the strip equal to six times the background value. Initial strip widths of 0.018, 0.035, 0.07, 0.14, 0.21, and 0.28 were used. The corresponding Froude numbers found in the subsequent evolution varied from 0.11 to 0.67, increasing monotonically as the initial strip width increased.

Figure 4.1 shows the potential vorticity field during the initial development and saturation of the instability in the simulation **A**i. All potential vorticity plots are shown with an aspect ratio of unity, and consequently they show only a small portion of the domain in the cross-stream direction. While the model domain accomodates only one wavelength of the primary instability in the streamwise direction, in figure 4.1, and in all subsequent figures shown, two periods of the model domain are displayed.

In figure 4.1, one can see that, as the instability develops, waves develop on the edges of the potential vorticity strip, which can be seen clearly in figures 4.1b and 4.1c. Nonlinear saturation of the instability is shown in figure 4.1d, where the strip has rolled up into a train of coherent vortices, connected by thin filaments.

Figure 4.2 shows the potential vorticity field at four subsequent equally spaced time intervals until the end of the simulation. Although a small amount of hyperdiffusion is applied in the x–direction to prevent an accumulation of noise on the grid scale of the numerical model, it seems that the inviscid character of the vortical flow has been well simulated in this experiment. The vortices appear to nutate between almost axisymmetric and substantially elliptical shapes, whereas if the hyperviscosity is increased significantly they become axisymmetric after only two or three rotations. It is possible that the presence of background shear in the experiment is

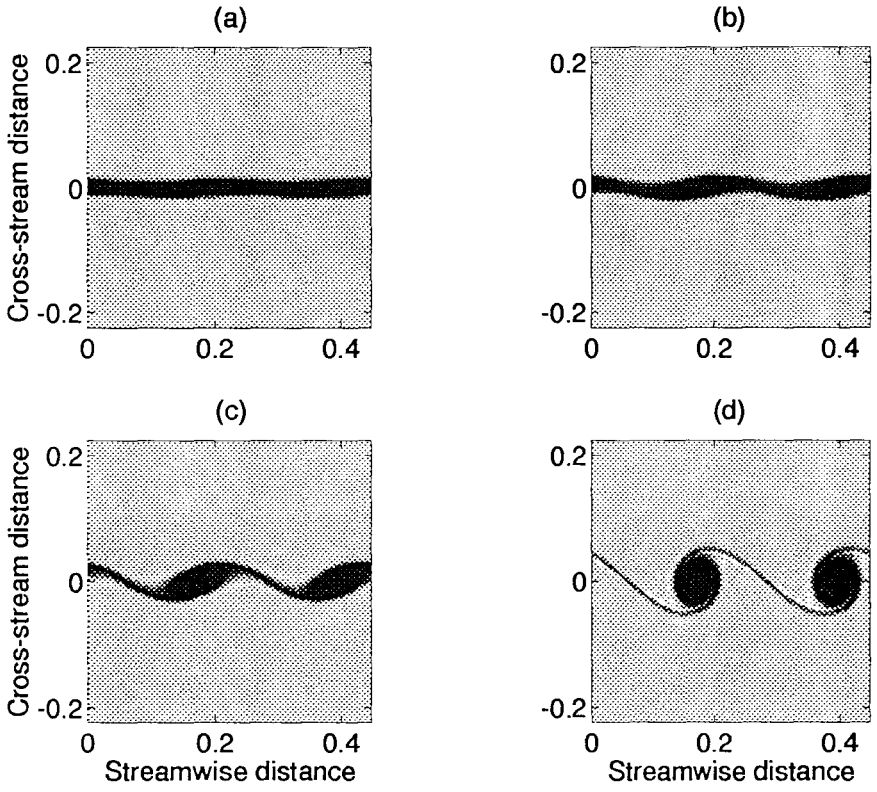


Figure 4.1: Early stages of potential vorticity evolution, simulation Ai. Figures a,b,c,d correspond to times 1.56, 2.81, 4.06 and 5.31 respectively. Two periods of the model domain are shown.

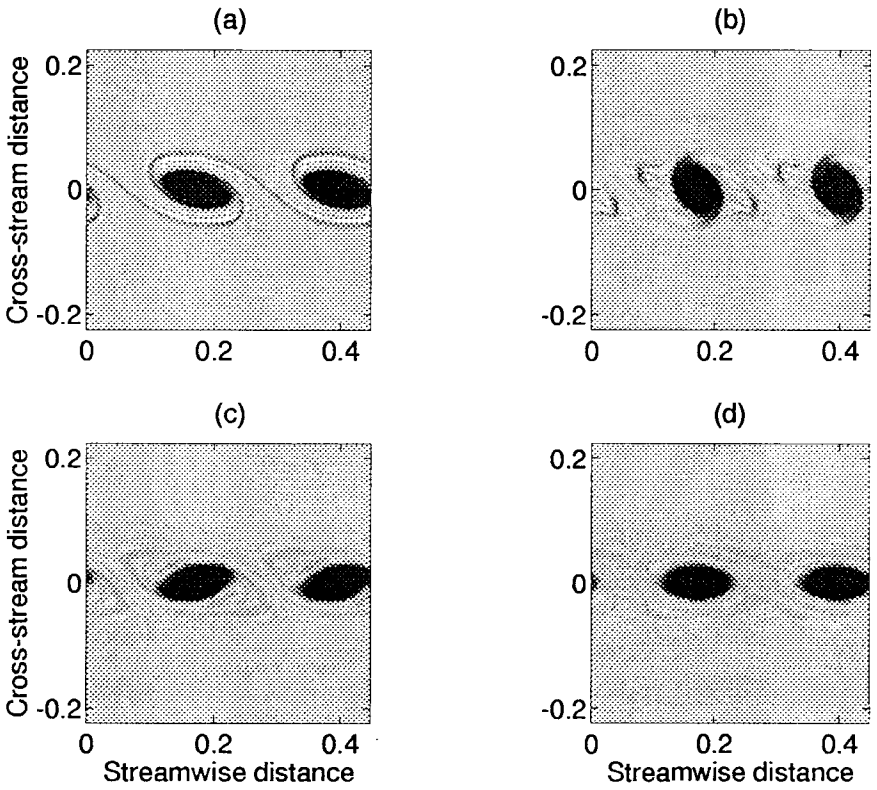


Figure 4.2: Nonlinear evolution of potential vorticity field, simulation Ai. Figures a,b,c,d correspond to times 6.56, 12.2, 17.8 and 23.4 respectively. Two periods of the model domain are shown.

responsible for this nutating behaviour, as proposed by Kida (1981). The filamentary structure appears to have been almost completely smeared out after one rotation of the vortices.

Linear gravity waves have zero potential vorticity perturbation, and so they are not observed in the potential vorticity plots shown in figures 4.1 and 4.2. However, we recall from (4.6) that, in the absence of nonlinearity, the Eulerian time derivative of the height field, $\partial h/\partial t$, satisfies the linear wave equation for gravity waves, i.e.

$$\left(\frac{\partial^2}{\partial t^2} + f^2 - c_0^2\nabla^2\right)\frac{\partial h}{\partial t} = \text{nonlinear terms.} \tag{4.31}$$

It is therefore natural to use $\partial h/\partial t$ to investigate gravity waves radiated by the shear instability. For each of the times shown in figure 4.2, the corresponding $\partial h/\partial t$ field is shown in figure 4.3. Only the region $y > 0$ is shown, since all simulations are symmetric about $y = 0$. The entire computational domain for $y > 0$ is shown in all plots of $\partial h/\partial t$. In figure 4.3, to show the entire computational domain requires that the cross-stream direction be compressed compared with the streamwise direction. As in the case of the potential vorticity field, two periods of the model domain in the streamwise direction are shown in all plots of $\partial h/\partial t$. The colourmap is chosen to saturate at the peak amplitudes in the wave region, even though typical values of $\partial h/\partial t$ in the vortical region may be more than 10 times greater than those found in the wave region. No quantitative information should be inferred from the the greyscale plots of $\partial h/\partial t$ shown in this chapter. We shall address the quantitative aspects of the gravity wave radiation when we compare the waves generated with those predicted by the Lighthill theory, and discuss the way in which the amplitude of the radiating wave fluxes depend on the Froude and Rossby numbers of the vortical flow.

In the greyscale figure 4.3, however, two features of the radiated gravity wave field are nonetheless particularly striking. Firstly, the wavelength of the gravity waves is much longer than the cross-stream scale of the vortical motions. Secondly, the radiated gravity wave field is almost independent of the streamwise coordinate x. This means that we should expect the reconstruction of the gravity wave field from the Lighthill source term to be quite good for this simulation.

Figure 4.4 shows the source function $S(t)$ for simulation **Ai**, computed from (4.8). During the growth phase of the instability, $S(t)$ remains almost zero, and only becomes significant during the subsequent rotation of the vortices in the train.

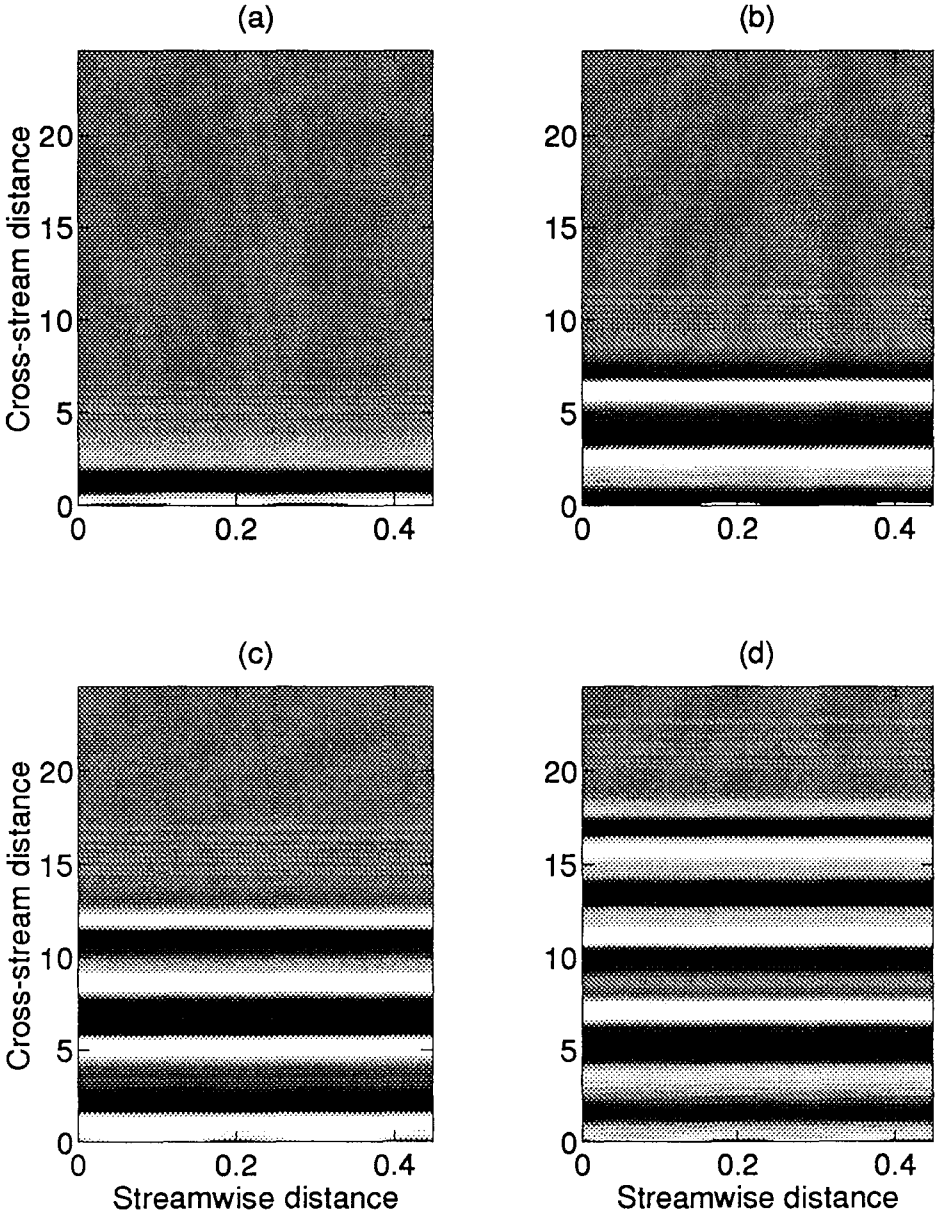


Figure 4.3: Development of gravity wave radiation, simulation Ai. Figures a,b,c,d correspond to times 6.56, 12.2, 17.8 and 23.4 respectively. Two periods of the model domain are shown.

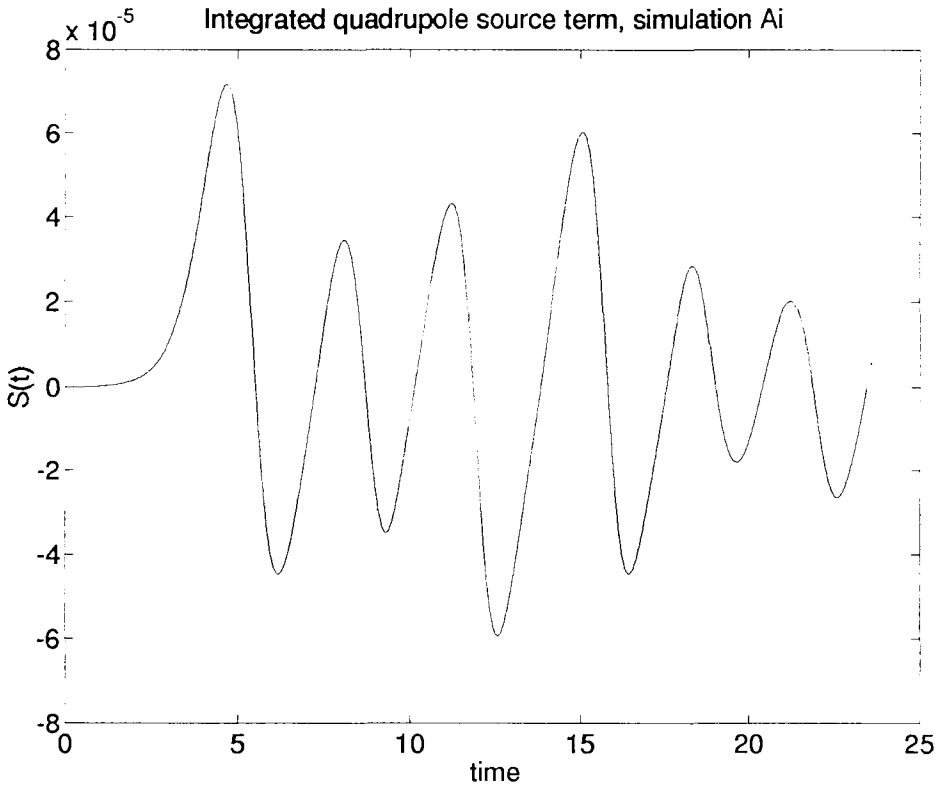


Figure 4.4: Integrated quadrupole source term, $S(t)$, for simulation **Ai**

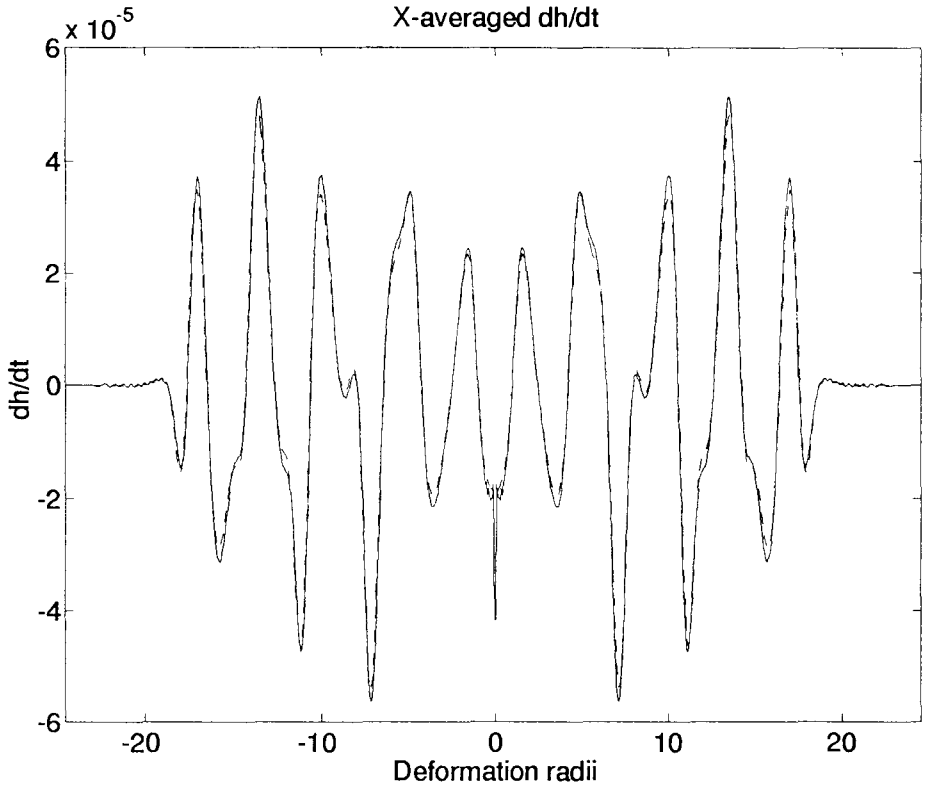


Figure 4.5: Comparison of x–averaged $\partial h/\partial t$ field at termination of simulation **Ai**: solid line is from nonlinear simulation, dashed line is reconstruction from Lighthill convolution integral

In figure 4.5, the x–averaged $\partial h/\partial t$ field is shown at the termination of simulation **Ai**. The solid line in figure 4.5 is the x–average of the height field obtained directly from the nonlinear simulation. The dashed line in figure 4.5 is the result of using the source term $S(t)$, shown in figure 4.4, in (4.7), and is therefore the wave field predicted by this modified form of the Lighthill theory, using the actual velocity and height fields from the simulation to evaluate the source term. For this simulation, the agreement between the full simulation and the Lighthill reconstruction seems very good. This is reassuring, partly because the Froude number is quite small, and the compact source approximation should be reasonably accurate for this flow; and partly because it means that the numerical model is performing sufficiently well to capture the Lighthill mechanism, and that the resolution employed is adequate to reconstruct the Lighthill source term for the convolution integral (4.7).

Figures 4.6 and 4.7 show the potential vorticity and $\partial h/\partial t$ respectively from simulation **Aii** at four equally spaced time intervals during the simulation, and figures 4.8 and 4.9 show the

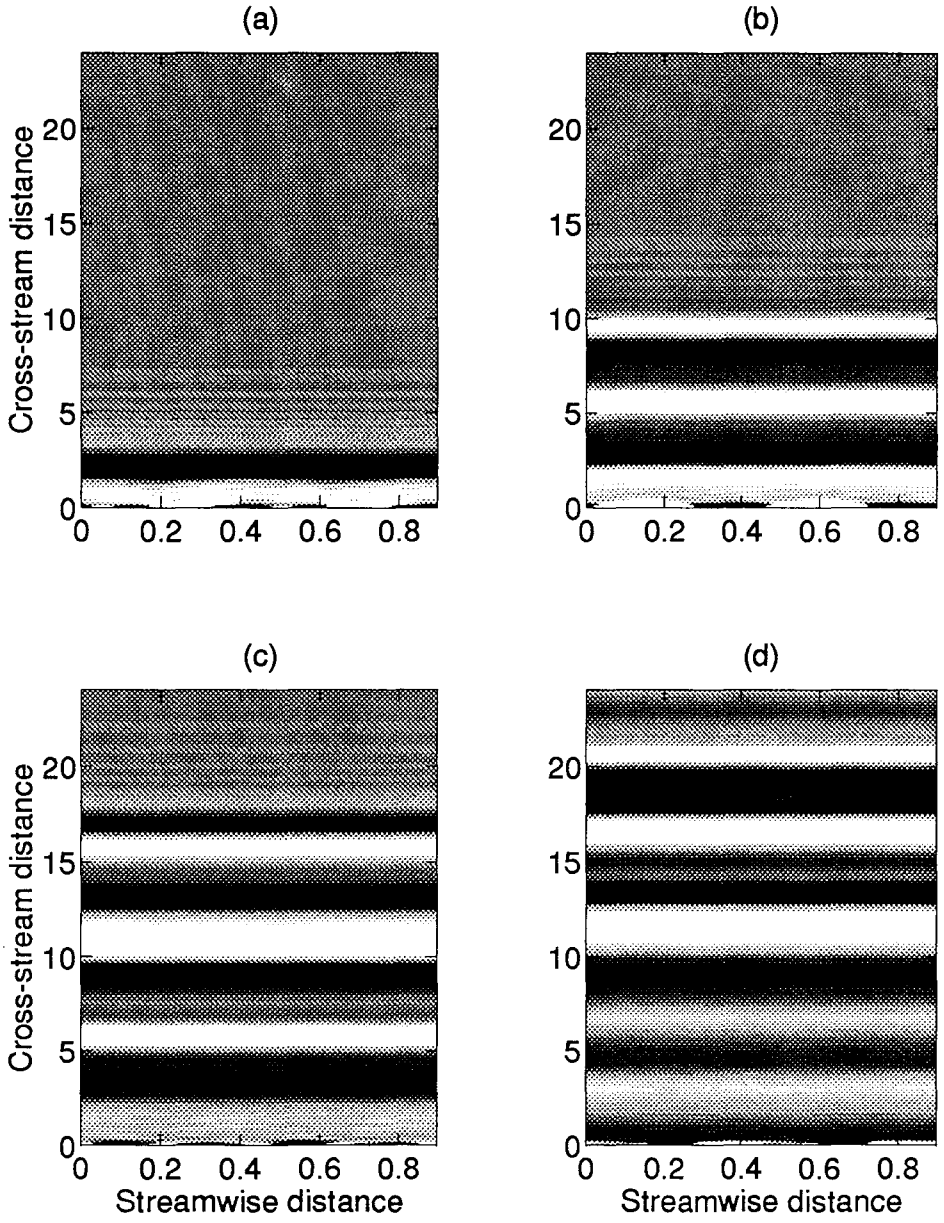


Figure 4.7: Development of gravity wave radiation, simulation Aii. Figures a,b,c,d correspond to times 6.9, 15.0, 23.1 and 31.2 respectively. Two periods of the model domain are shown.

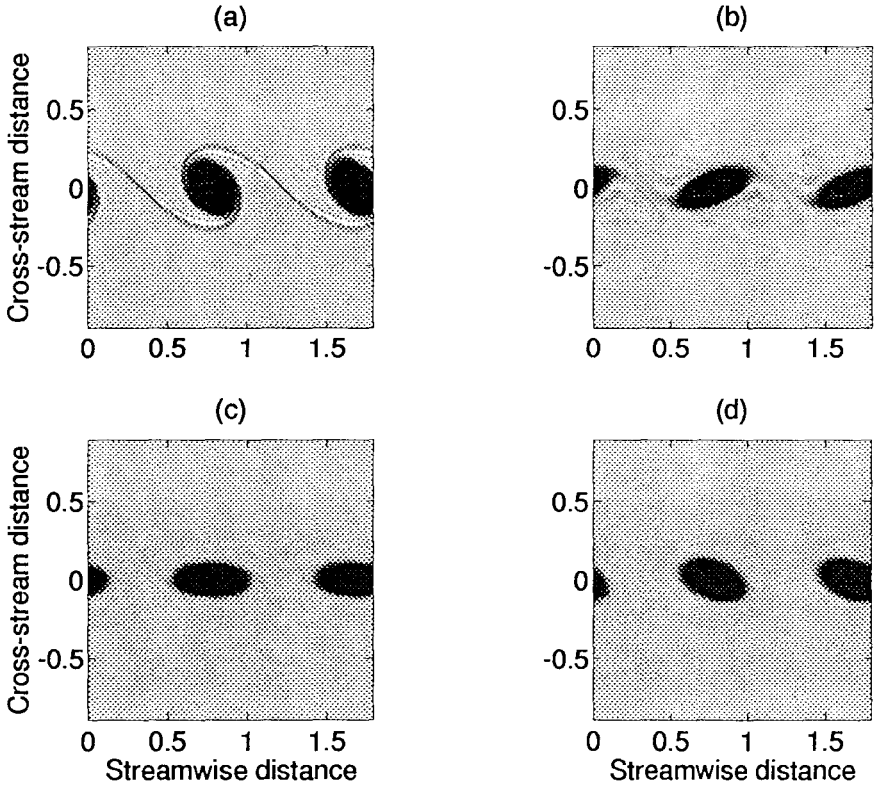


Figure 4.8: Nonlinear evolution of potential vorticity field, simulation Aiii. Figures a,b,c,d correspond to times 8.13, 26.3, 44.4 and 62.5 respectively. Two periods of the model domain are shown.

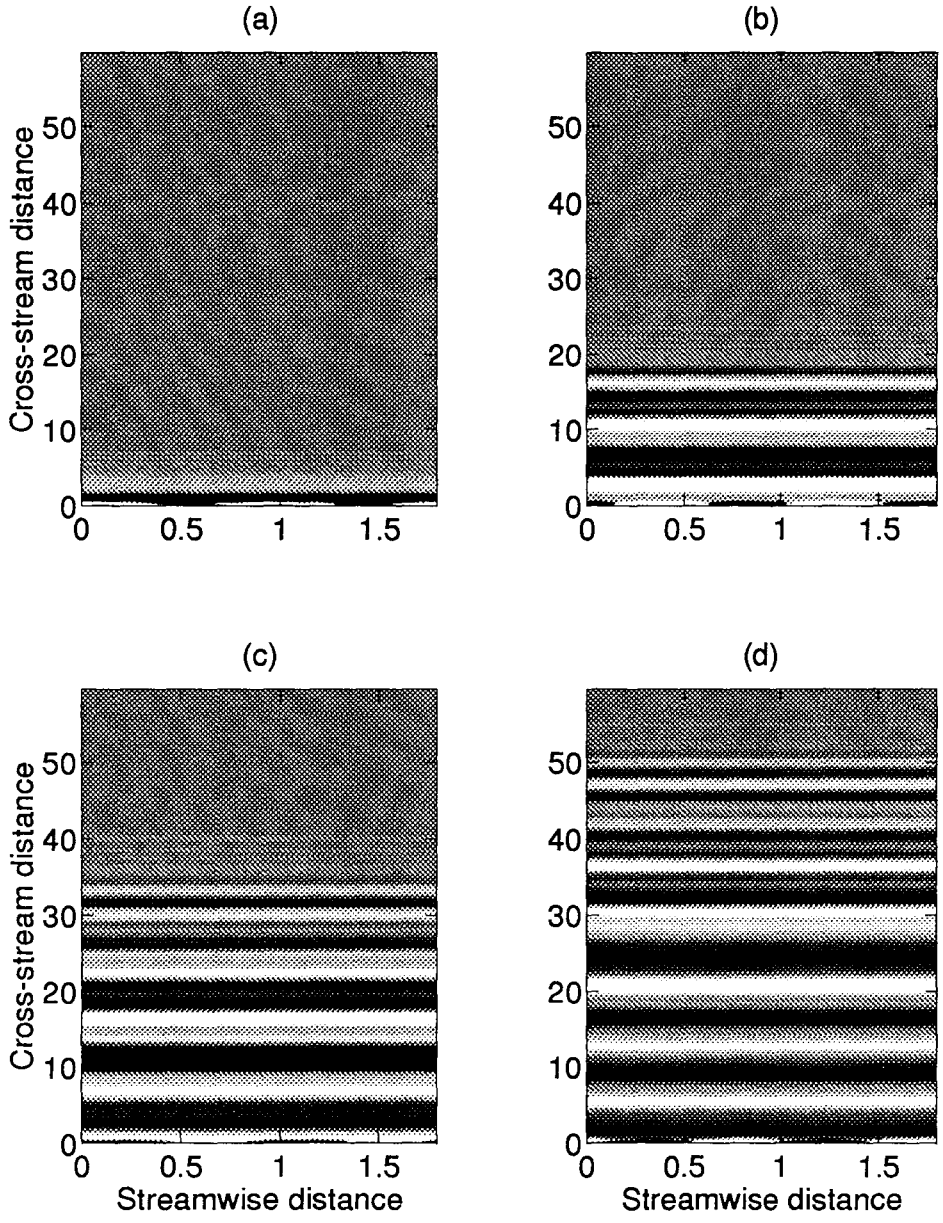


Figure 4.9: Development of gravity wave radiation, simulation Aiii. Figures a,b,c,d correspond to times 8.13, 26.3, 44.4 and 62.5 respectively. Two periods of the model domain are shown.

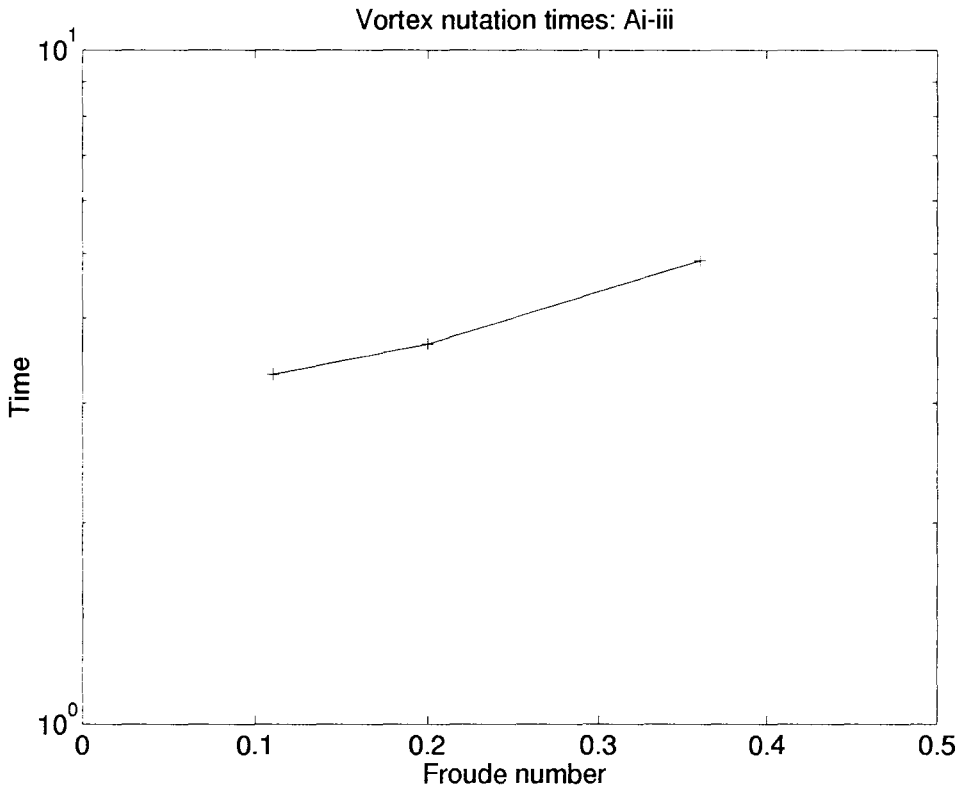


Figure 4.10: Vortex nutation times: simulations Ai-iii

same fields respectively for simulation Aiii.

On visual inspection, there appears to be little difference in the vortical aspects of the dynamics between each of the three simulations Ai–iii. In each case, the strip of high potential vorticity rolls up into a periodic train of vortices, which subsequently nutate. The only perceptible difference is in the nutation time, shown in figure 4.10, which increases from 3.44 inertial periods (simulation Ai) to 4.88 inertial periods (simulation Aiii). The vortex nutation times are a significant feature of the flow, in that they are a broad measure of the degree of unsteadiness of the vortical flow compared with the unit inertial period.

Considering now the gravity wave aspects of the flows Aii and Aiii, we see from figures (4.7) and (4.9) that the form of the radiated gravity waves remains similar in these two cases to those observed in simulation Ai. In figure 4.11, the x-averaged $\partial h/\partial t$ field at the end point of each of the numerical simulation is represented by the solid lines, and its reconstruction using the compact source approximations is shown by the dotted lines. The agreement seems to be

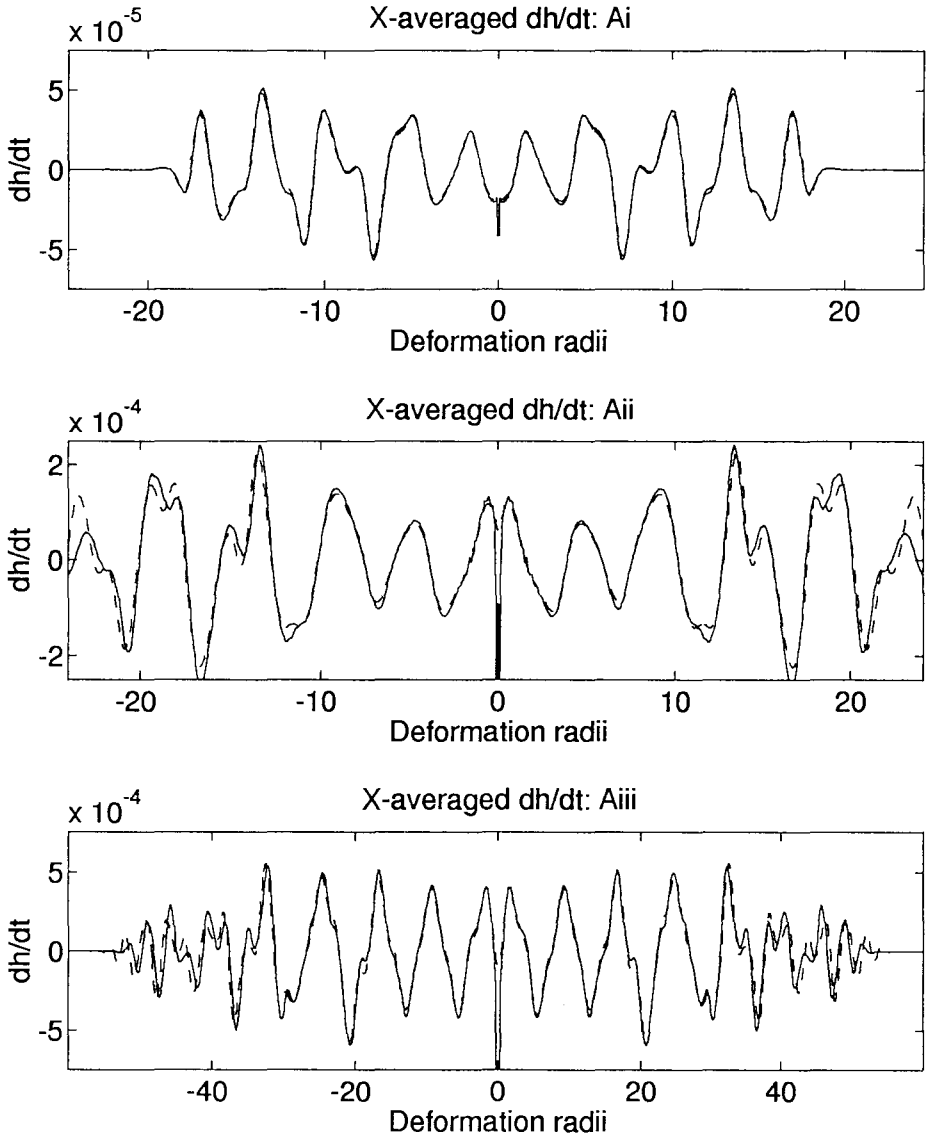


Figure 4.11: The x–averaged $\partial h/\partial t$ field at the termination of simulations Ai-iii. Solid lines are from the nonlinear simulation, and dashed lines are from the Lighthill convolution integral

quite good in each case.

From figure 4.11, we can also see that the amplitude of the radiated gravity waves increases by a factor of about 12 between **A**i and **A**iii. The maximum Froude number found in the evolution increases from 0.11 (**A**i) to 0.36 (**A**iii) (see table 4.1c). However, we should recall from (4.10) that, as the Froude number is increased, we expect the amplitude of the radiated gravity waves to increase as the third power of the Froude number. Thus, on the basis of the low Froude number asymptotic analysis, we should expect the gravity wave amplitude to change by a factor of 35. Evidently, the increase in gravity wave amplitude as a function of Froude number over this range is not as large as one might expect on the basis of the Lighthill theory. Two assumptions were made to predict that the amplitude of the radiated waves would scale with the third power of the Froude number – the point source assumption, and the scaling assumption $u \sim \omega l$. The fact that the wave amplitudes do not increase as rapidly as the theory predicts is thus not necessarily an indication that the point source approximation will fail for these simulations.

Three further simulations (**A**iv – **A**vi) were then performed, with strip widths of 0.14, 0.21, and 0.28. The vortical aspects of the flow remained qualitatively unchanged as the width of the strip, and hence the Froude number, was increased. In each case, the strip initially rolled up into a periodic train of vortices, which were allowed to rotate several times before the simulation was terminated. The most significant quantitative difference in the vortical aspects of the flow between these three cases is in the nutation rate of the vortices in the train, shown in figure 4.12, which increase super-exponentially with increasing Froude number.

Figure 4.13 shows the $\partial h/\partial t$ field at the termination of the simulations **A**iii–**A**vi. The main point is that, as the Froude number is increased, the radiated gravity waves develop more x–dependent structure. The gravity wave field is expected to be x–independent at low Froude numbers, and it is reasonable to expect more x–dependence in the field as the Froude number is increased, as the lengthscale separation between the vortices and the gravity waves diminishes.

Figure 4.14 shows the x–averaged $\partial h/\partial t$ field at the termination of simulations **A**iii-vi, and its reconstruction from the Lighthill source term. The most striking feature of the sequence of figures is that, in contrast to the rapid increase in wave amplitude with Froude number predicted by Lighthill's scaling argument, the radiated wave amplitudes actually decrease. This clearly differs from the scaling arguments presented by numerous authors on aeroacoustics, which were

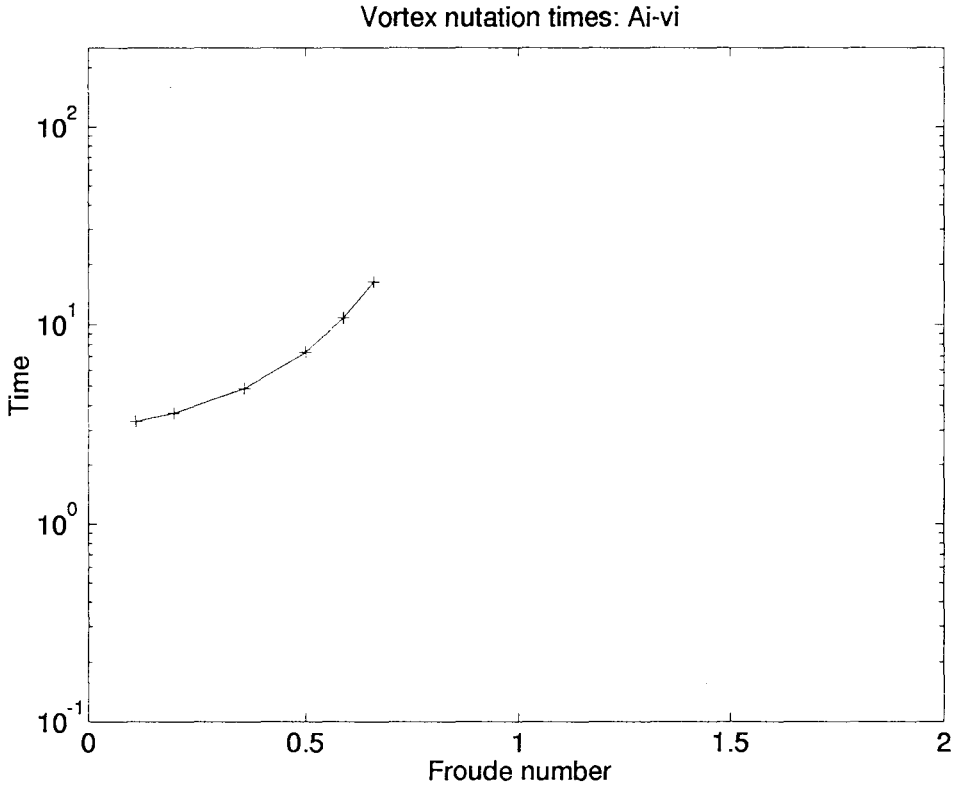


Figure 4.12: Vortex nutation times: simulations Ai-vi

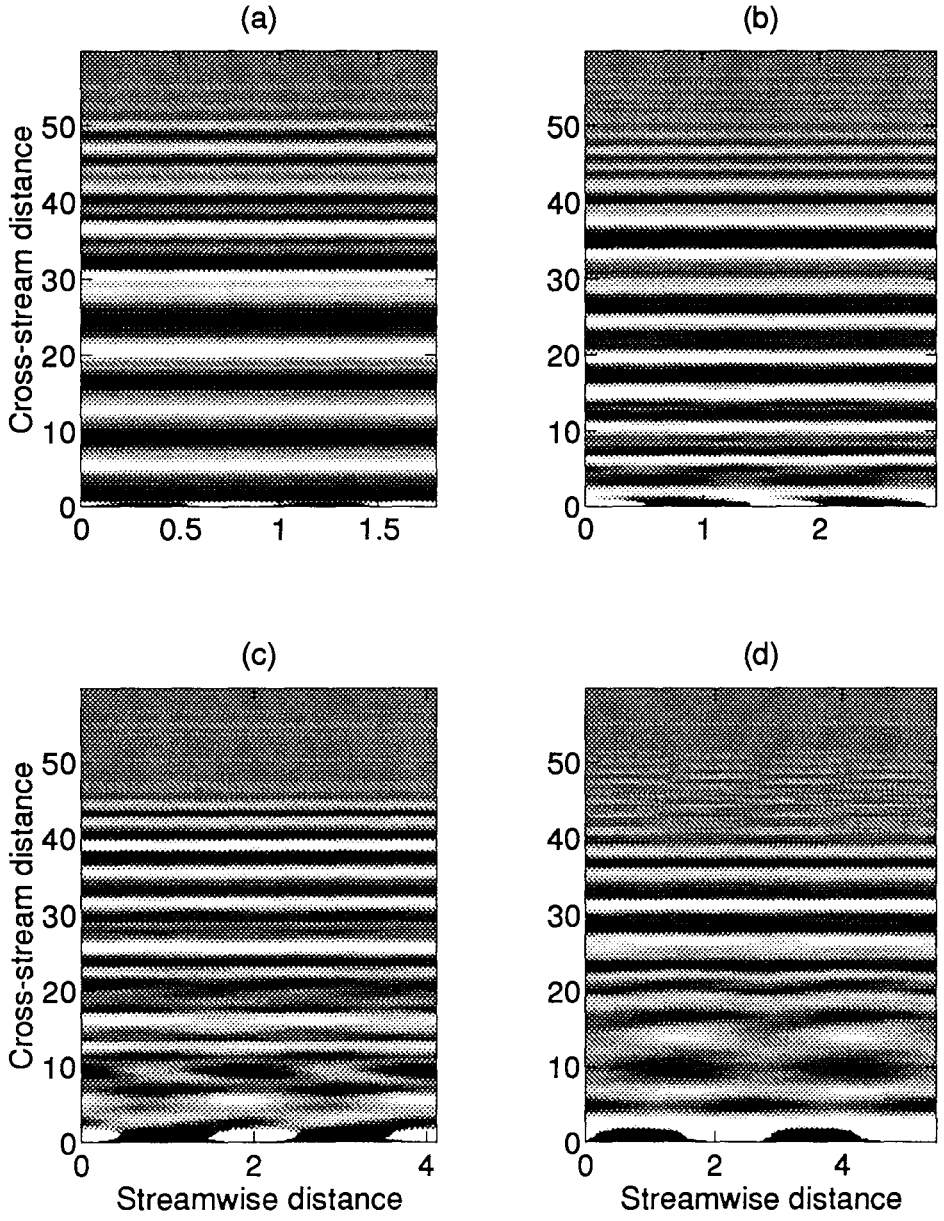


Figure 4.13: The $\partial h/\partial t$ field at the temination of simulations Aiii-vi (time 62.5). Two periods of the model domain are shown.

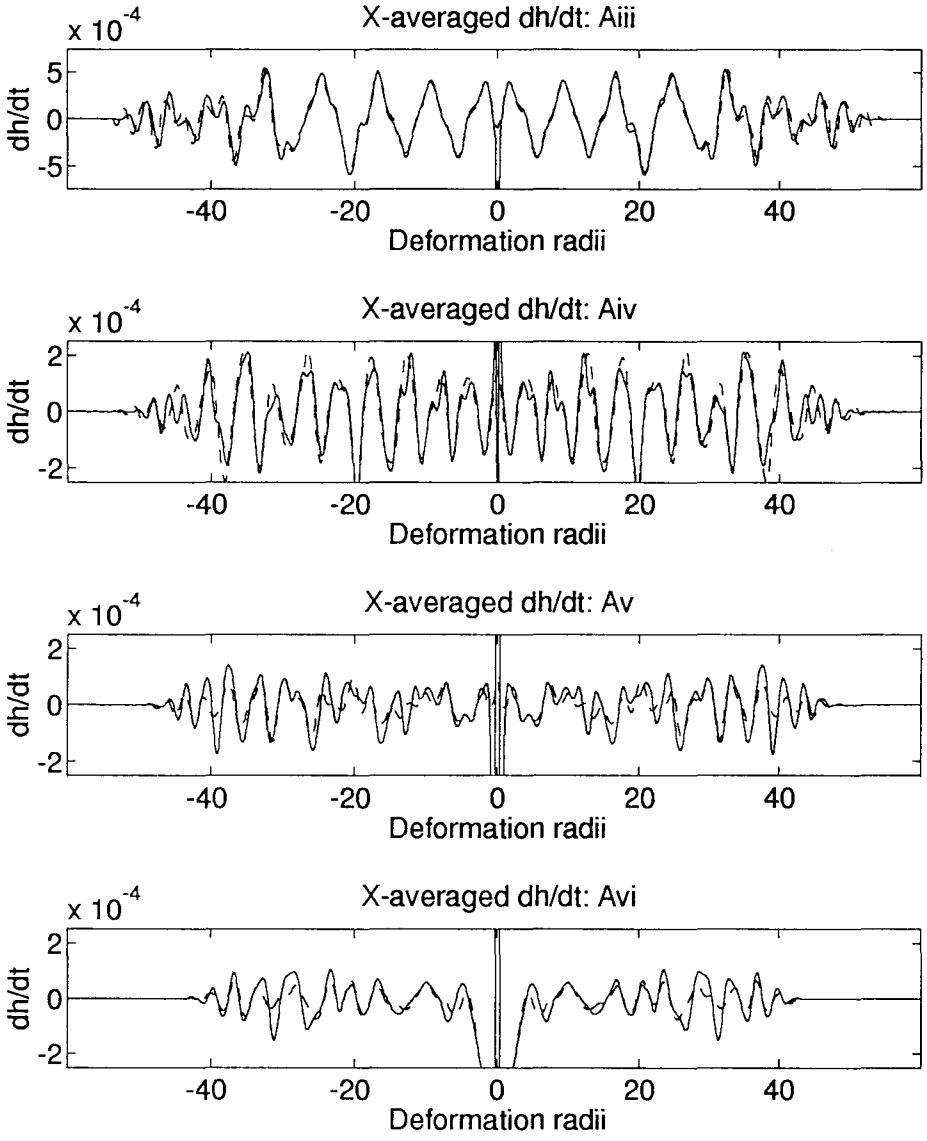


Figure 4.14: The x–averaged $\partial h/\partial t$ field at the end of simulations **A**iii-vi

discussed at the beginning of this section, and from the mass of experimental data that confirm the Lighthill scaling in practice.

Since the only new effect in the present study is the inclusion of background rotation, it follows that background rotation must be responsible for decreasing the wave amplitude as the Froude number is increased. Moreover, since the reconstruction of the wave field from the Lighthill source term effectively captures the decrease in wave amplitude, it must be explicable within the framework of the Lighthill theory.

There are two possibilities. One is that the quadrupole source strength decreases, rather than increases, as the Froude number is increased. However, we should bear in mind that in order for the quadrupole source to excite propagating waves, it must have a significant component of its frequency above the inertial frequency. Therefore, it is possible that, although the magnitude of the source terms might actually increase as the Froude number is increased, their frequency component above the inertial frequency might decrease.

Figure 4.15 shows the integrated source term for each of the simulations **A**i–vi. In general, the amplitude of the source term increases with increasing Froude number, but with frequency decreasing very rapidly with increasing Froude number. It seems, therefore, that the effect of the inertial cut-off, inhibiting gravity wave radiation at frequencies below the inertial frequency, is now dominant over the increased magnitude of the source term, leading to reduced wave amplitudes as the Froude number is further increased.

We may conclude that, although the potential vorticity in the strip was chosen to be six times the background value, both the vortical flow and the wave generation process are significantly affected by the presence of background rotation. For very small Froude numbers (between 0.11 and 0.35), the amplitude of the radiated gravity waves increases with Froude number, although not quite as rapidly as F^3, as found in the asymptotic limit $F \ll 1$. When the Froude number exceeds 0.35, however, the radiated gravity wave amplitudes are found to decrease with increasing Froude number, even though the typical amplitudes of $S(t)$ are generally increasing. This behaviour has never been found in numerical or laboratory experiments in a non-rotating frame. The presence of background rotation thus severely inhibits gravity wave radiation at moderate Rossby numbers, and there seems little doubt that the present study has investigated a case in which the effect of background rotation is quite strong, despite a notional Rossby number of 5. Therefore, in the next section, we investigate the effect of increasing the potential

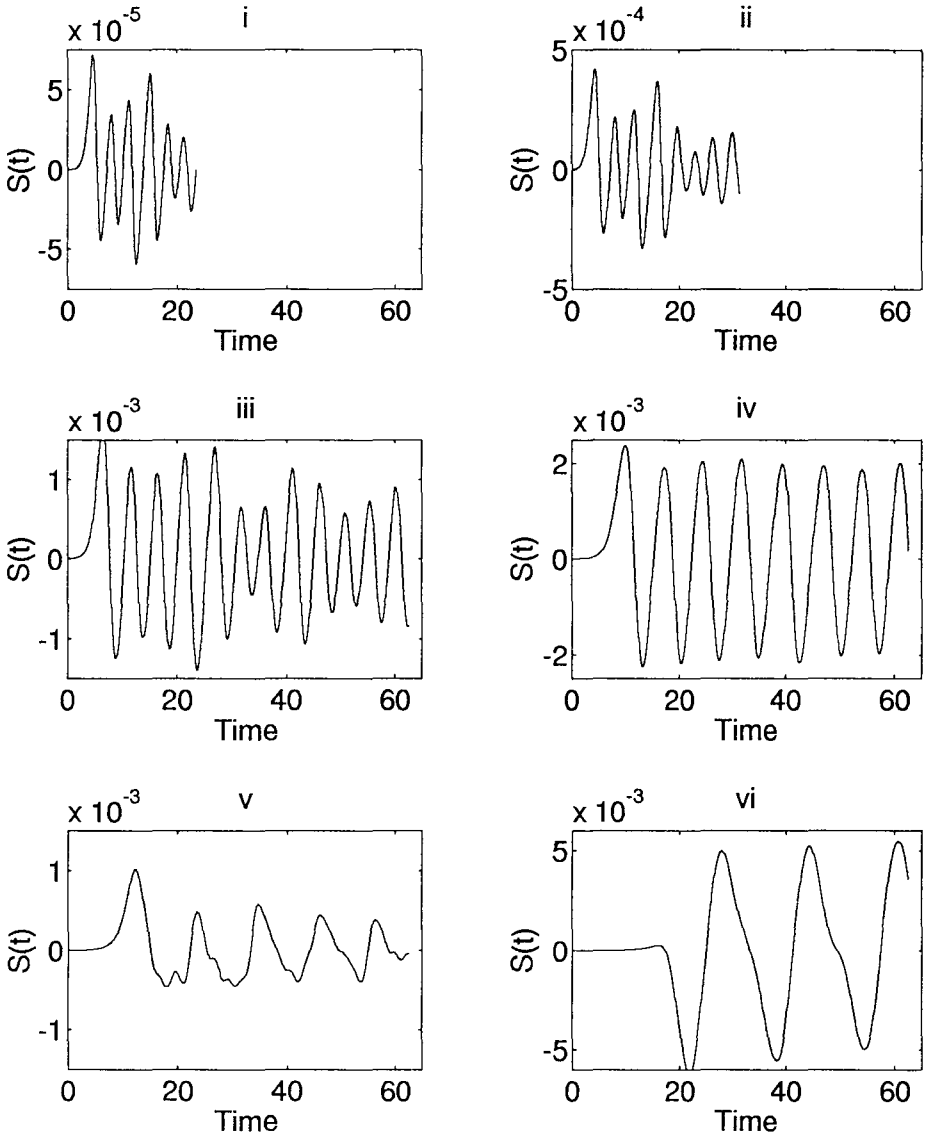


Figure 4.15: The source term $S(t)$ for simulation **A**i-vi. Note the differing scales on the $S(t)$ axes.

vorticity in the strip. The strip remains cyclonic, but the effect of background rotation should decrease as the potential vorticity in the strip increases.

4.7 Cyclonic strips at large Rossby number

In this section the results of simulations with strip potential vorticities of 9 and 21 are presented. There are four simulations with strip potential vorticity of 9, and four with strip potential vorticity of 21. The simulations with strip potential vorticity of 9 are labelled **B**i–iv, and those with strip potential vorticity of 21 are labelled **C**i–iv. In neither set of simulations should we expect to observe any qualitative difference from the simulations with potential vorticity of 6 in the strip. Quantitatively, we should expect to see more rapid vortical motions, and consequently a larger pseudoenergy flux in the radiated wave field.

In simulations **B**i–iv, the initial strip widths are 0.018, 0.035, 0.07 and 0.21, resulting in Froude numbers of 0.17, 0.32, 0.53 and 0.84 respectively. In each case, the strip rolled up into a periodic train of vortices, which then nutated, radiating gravity waves. The vortical flow appeared very similar to that observed in experiments **A**i–vi, and potential vorticity plots are not shown here. The nutation periods for these vortices are shorter than for the corresponding simulations in sequence **A**, but still appear to increase super-exponentially with Froude number (figure 4.16), in common with sequence **A**.

In figure 4.17, the $\partial h/\partial t$ field is shown for the end of each simulation **B**i–iv. The wave field follows broadly the same pattern as in the lower Rossby number simulations **A**i-vi. At small Froude numbers, the wave field is almost independent of x, and as the Froude number is increased, amplitudes of the x–dependent modes increase.

In figure 4.18, the x–independent wave field is shown at the end of each simulation (solid line), and compared with its reconstruction from the Lighthill theory (dashed line). Overall, there is good agreement between the Lighthill theory and the numerical simulations, as there was in the lower Rossby number simulations **A**i–vi. As before, the agreement is best when the Froude numbers are lowest. The radiated wave amplitudes are seen to increase with increasing Froude number for simulations **B**i–iii, but then decrease between **B**iii and **B**iv. The corresponding pseudoenergy fluxes are shown in figure 4.19, and the maximum pseudoenergy flux in the radiated wave field for simulations **A** and **B** is shown in figure 4.20. Notice in figure 4.19 that

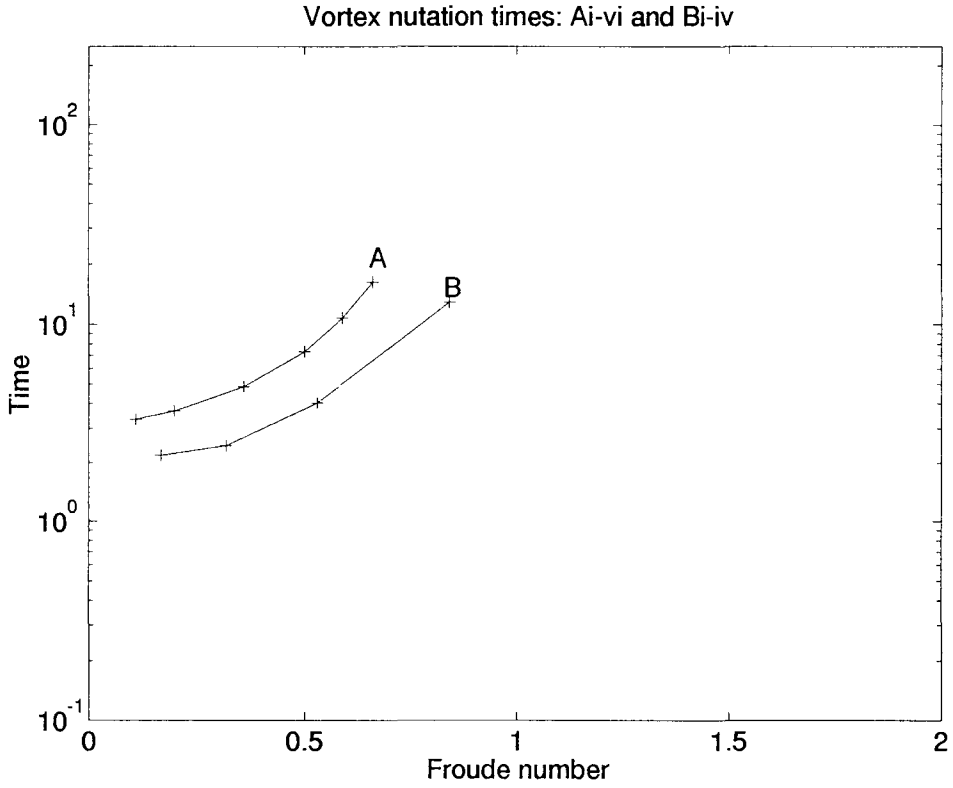


Figure 4.16: Vortex nutation times: simulations Bi-iv

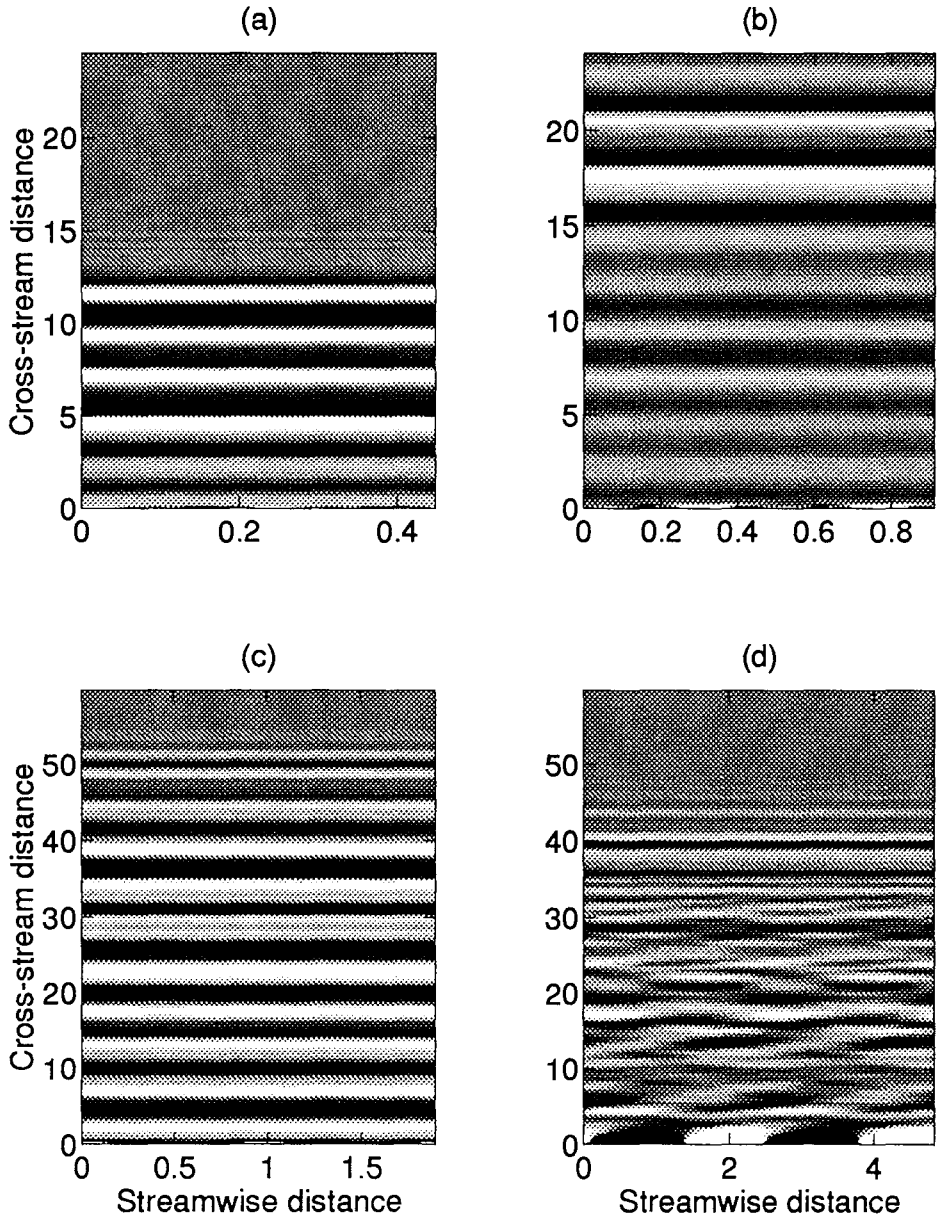


Figure 4.17: The field $\partial h/\partial t$ at the termination of simulations Gi-iv. Two periods of the model domain are shown.

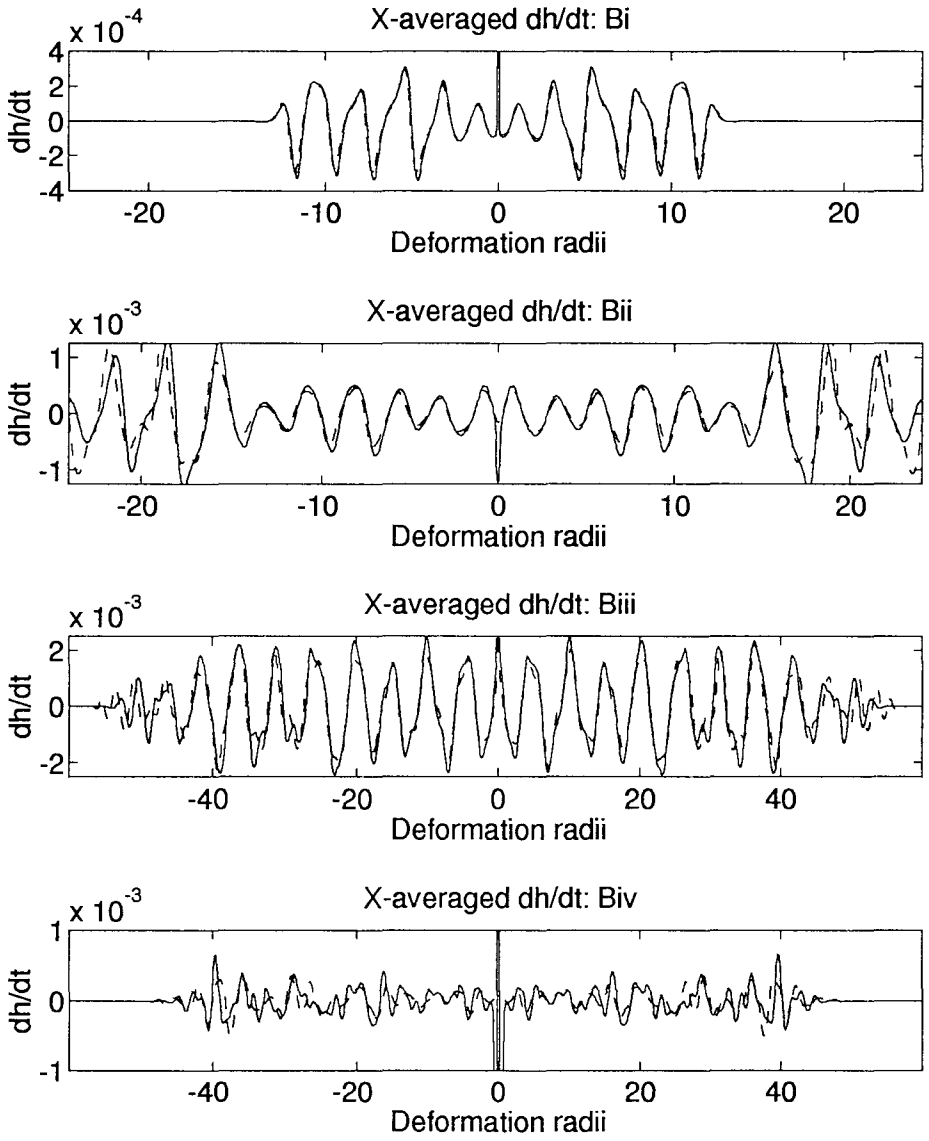


Figure 4.18: The x–averaged $\partial h/\partial t$ field at the termination of simulations **B**i-iv. Solid lines are from the nonlinear simulation, and dashed lines are from the Lighthill convolution integral

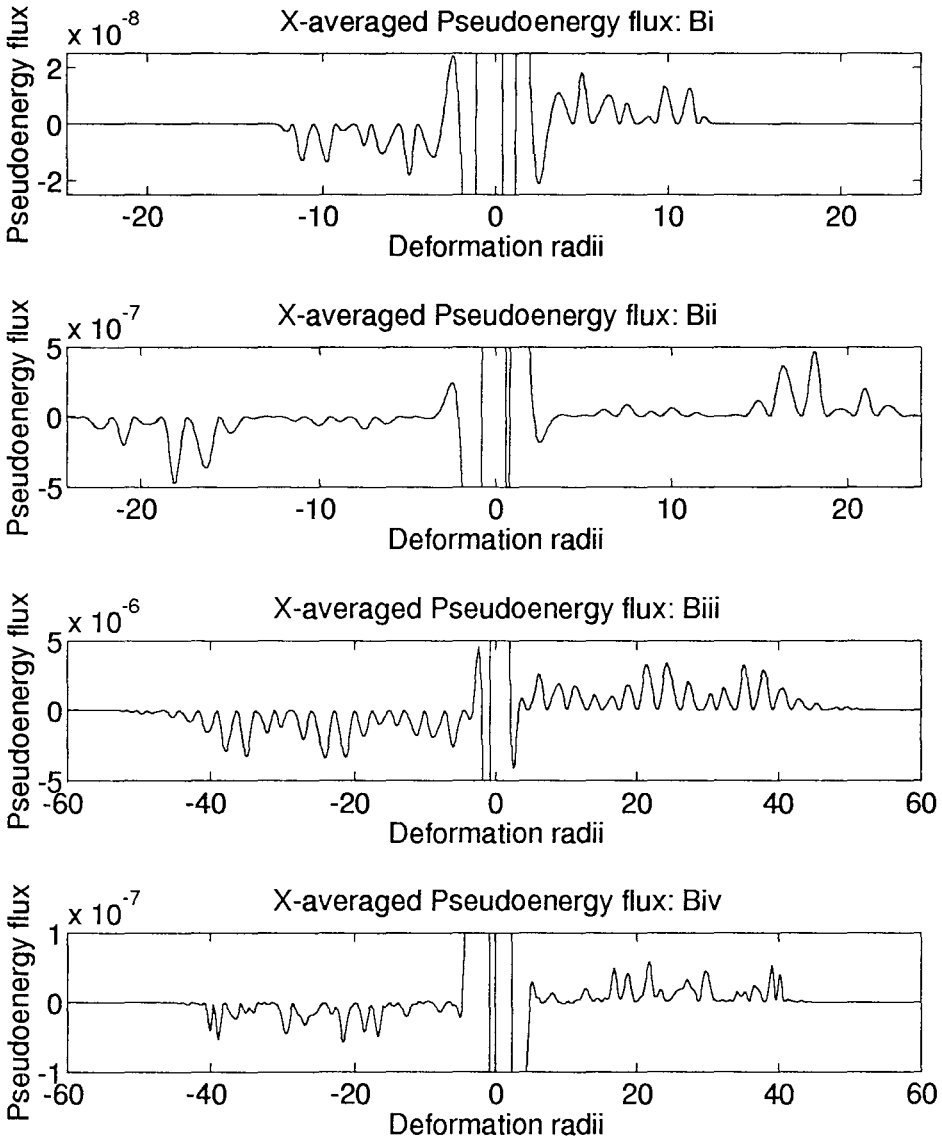


Figure 4.19: The x–averaged pseudoenergy flux at termination of simulations Bi-iv

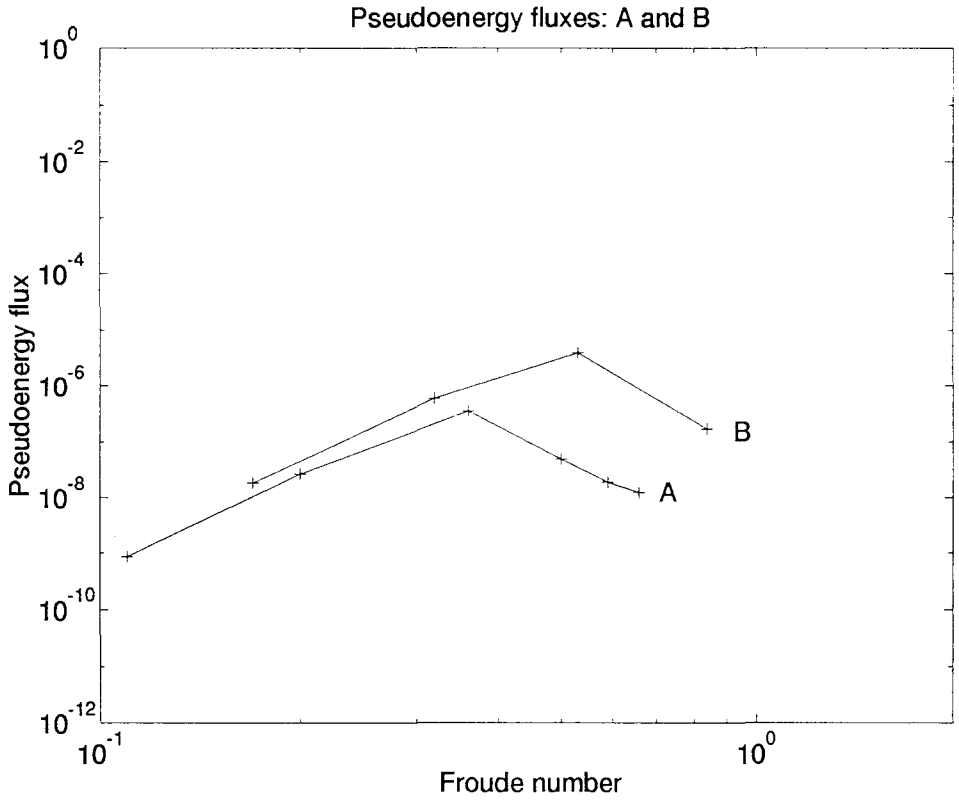


Figure 4.20: The peak pseudoenergy flux due to gravity wave radiation throughout simulations **A** and **B**

the pseudoenergy fluxes are negative for $y < 0$, and positive for $y > 0$, confirming that the flux of pseudoenergy away from the vortical flow at $y = 0$ is single-signed, as demonstrated by the analysis of §4.3.

The Froude number at which the maximum pseudoenergy flux occurs is greater in case **B** than in case **A**, but the strip width at which the transition from increasing to decreasing pseudoenergy fluxes appears to be about the same, at around 0.1.

Overall, increasing the potential vorticity in the strip from 6 to 9 has the effect of increasing the amplitude of radiated gravity waves. However, in both cases, there exists a Froude number (or equivalently a strip width) which is in some sense optimal for gravity wave radiation. The amplitude of the x–averaged flux of radiated gravity waves increases with increasing Froude number up to this critical Froude number, but then decreases as the Froude number is increased further. This property appears to be specific to flows in a rotating frame, and is due to the reduced unsteadiness of the vortical flow at larger Froude numbers, manifested in the rapidly increasing nutation times at larger Froude numbers, combined with the inhibiting effect of the inertial cut-off frequency on gravity wave radiation. It has not been observed in non-rotating aeroacoustic experiments or simulations.

Four further experiments were then performed with a strip potential vorticity of 21. The experiments are labelled **C**i–iv. These correspond to initial strip widths of 0.011, 0.021, 0.041 and 0.081 respectively. Once again, in each case the strip rolls up into a periodic train of vortices, which then nutate several times before the simulation is terminated. The dependence of the nutation time upon the Froude number is shown in figure 4.21, and is consistent with the behaviour found in experiments **A** and **B** – i.e., it appears to be increasing super-exponentially with Froude number.

The $\partial h/\partial t$ field at the end of each simulation is shown in figure 4.22, and again has similar properties to those for simulations **A** and **B**, although at the largest Froude number simulated smaller scale features seem to appear in the $\partial h/\partial t$ field. The comparison of the x–averaged $\partial h/\partial t$ field with its reconstruction by the Lighthill theory is shown in figure 4.23, and again good agreement is generally obtained. Indeed, it seems that the agreement improves as the Rossby number is increased, when compared with simulations **A**i-vi and **B**i-iv. Even in simulation **C**iv, in which the Froude number exceeds unity in the vortical region, the general form of the radiated wave field is captured by the Lighthill theory, although the fine details are not captured

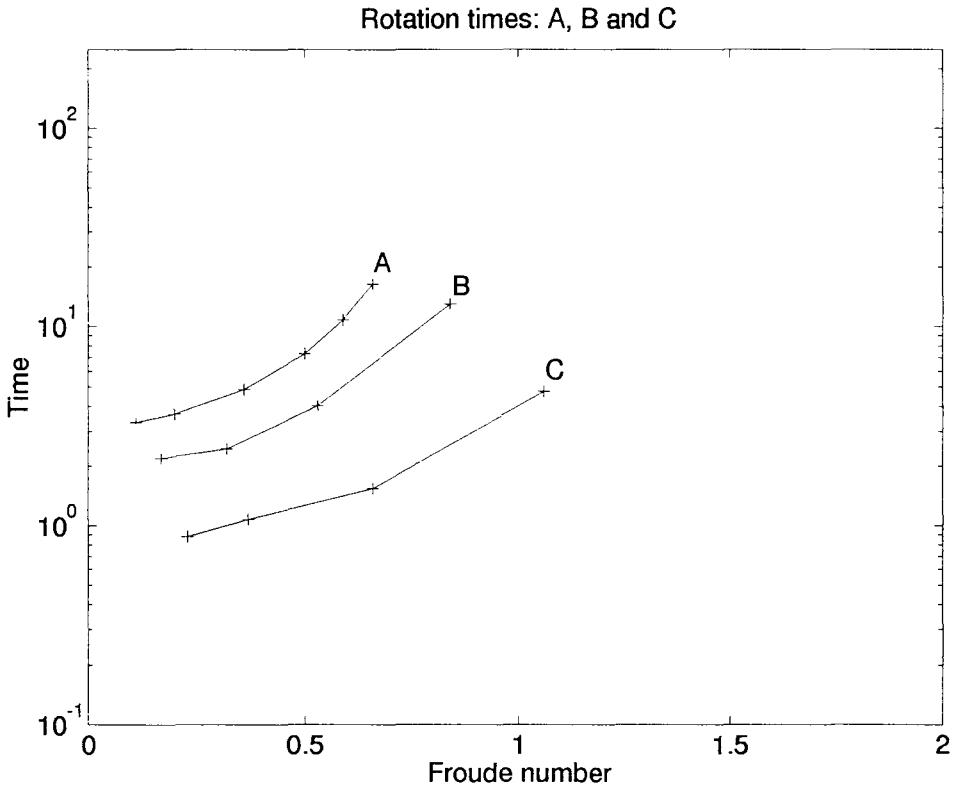


Figure 4.21: Vortex nutation times for simulations Ci-iv

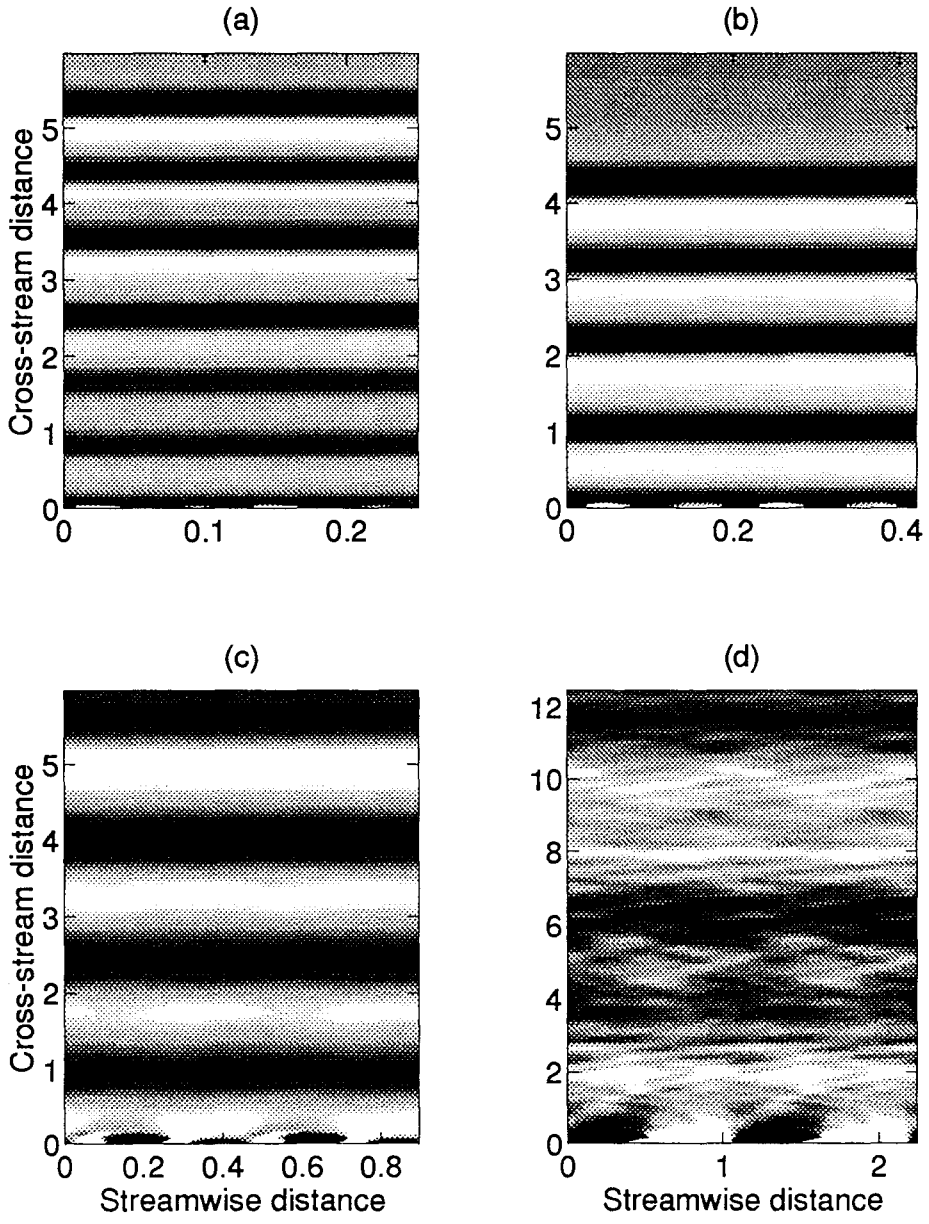


Figure 4.22: The field $\partial h/\partial t$ at the termination of the simulations Ci-iv. Two periods of the model domain are shown.

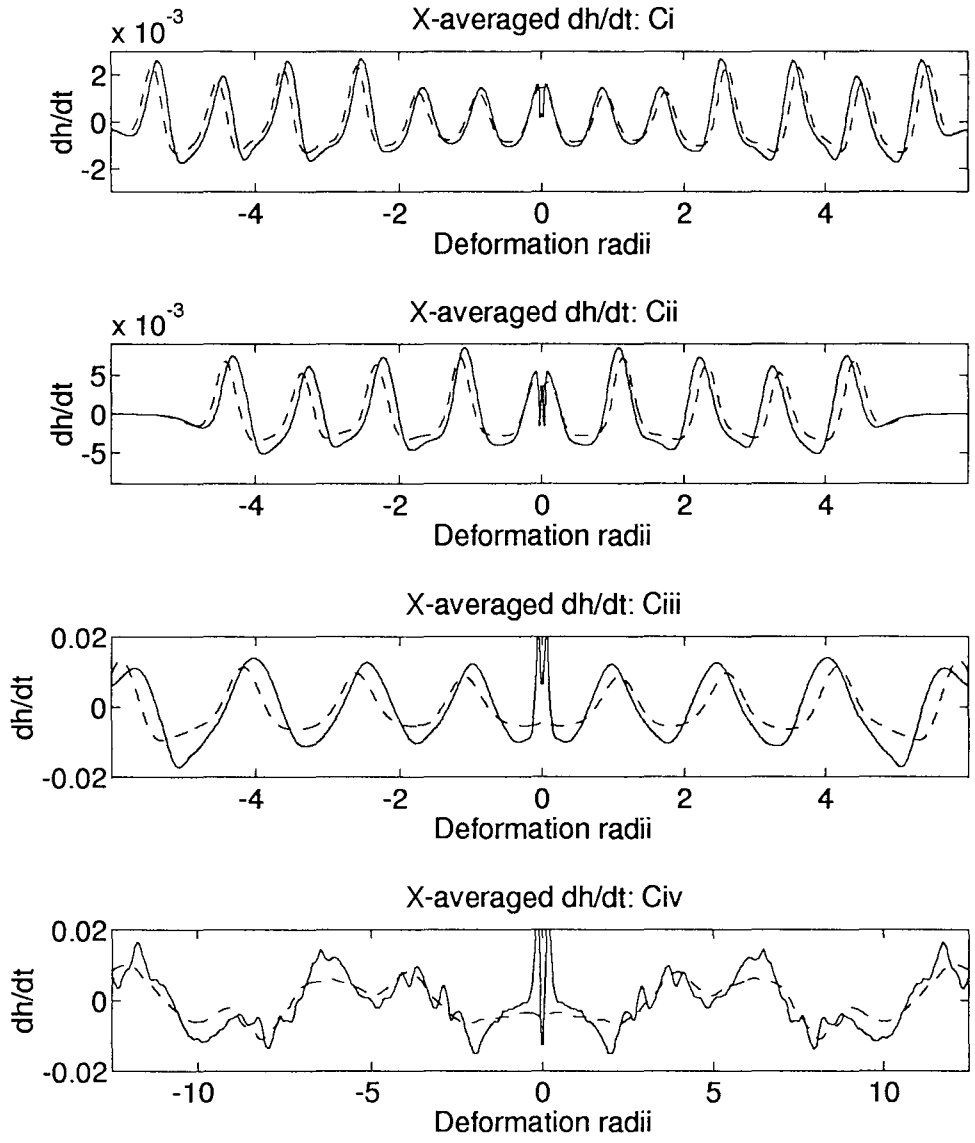


Figure 4.23: The x-averaged $\partial h/\partial t$ field at the termination of simulations Ci-iv (solid lines), and its reconstruction from the Lighthill convolution integral (dashed lines)

by the Lighthill theory.

The pseudoenergy flux at the end of each simulation is shown in figure 4.24. In this case, the pseudoenergy flux appears to be increasing as the Froude number is increased, unlike simulations **A** and **B**, where it decreased as the Froude number was increased beyond a certain critical value. It is interesting to note that in simulation **C**iv the x–averaged pseudoenergy flux appears to be a smoother measure of wave radiation than the x–averaged $\partial h/\partial t$ field, shown in figure 4.23.

In figure 4.25 the maximum x–average pseudoenergy flux over all time in the wave region for each simulation is plotted against Froude number, for the sequences of simulations **A**i–vi, **B**i–iv and **C**i–iv. The dashed line is a line of gradient 6 on the log-log scale, which would be predicted by the compact source asymptotics (equation 4.28). In all cases the pseudoenergy flux increases as the sixth power of the Froude number at small Froude number but, as the Froude number is increased, the pseudoenergy flux increases less rapidly with increasing Froude number. It seems plausible from figure 4.25 that the very large Rossby number simulations **C** will exhibit an optimal Froude number for gravity wave radiation, in the same manner as simulations **A** and **B**, but that it exceeds the values of Froude number investigated by the nonlinear simulations.

From this section, we may conclude that several of the features exhibited by the strips with a potential vorticity of 6 are robust features of cyclonic strip roll-up, subsequent nutation and gravity wave radiation. In particular, the nutation time increases super-exponentially with increasing Froude number. In general the pseudoenergy flux associated with the radiating gravity waves increases with the sixth power of the Froude number at small Froude number, but ultimately decreases as the Froude number is made very large, due to the increased nutation times, and consequently reduced fraction of the effective gravity wave source spectrum lying above the inertial frequency.

It is important to realize at this point that thus far only cyclonic strips have been investigated in this study. All cyclonic strips appear to exhibit similar behaviour, in respect of their vortex nutation times and pseudoenergy flux dependences upon the Froude number. In the remainder of this chapter, we shall investigate gravity wave generation by anticyclonic vortex trains, and compare it with that found for the cyclonic cases. The anticyclonic parameter space is divided into three parts: strips with positive potential vorticity, strips with negative potential vorticity, and the single case of a strip with a potential vorticity of zero. In the next section, §4.8, we

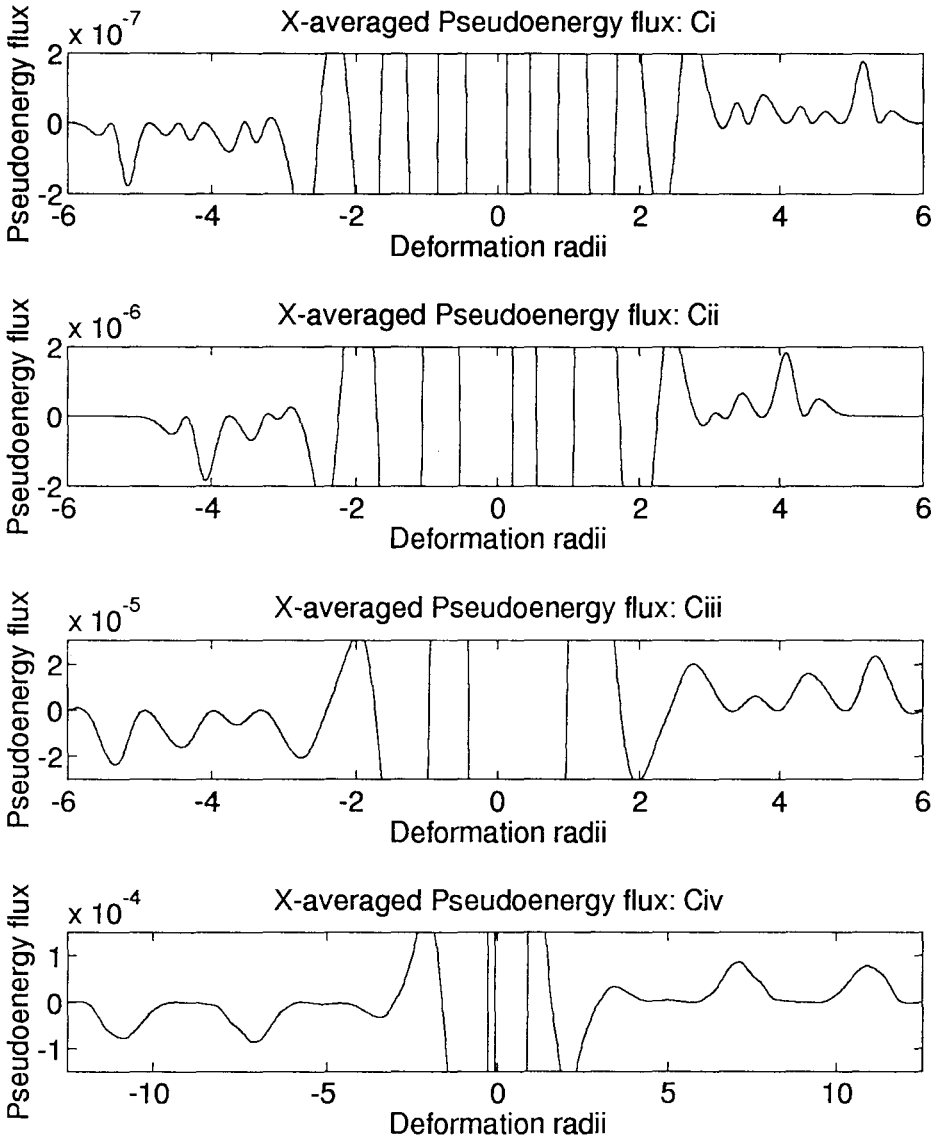


Figure 4.24: The pseudoenergy fulx at the termination of simulations Ci-iv

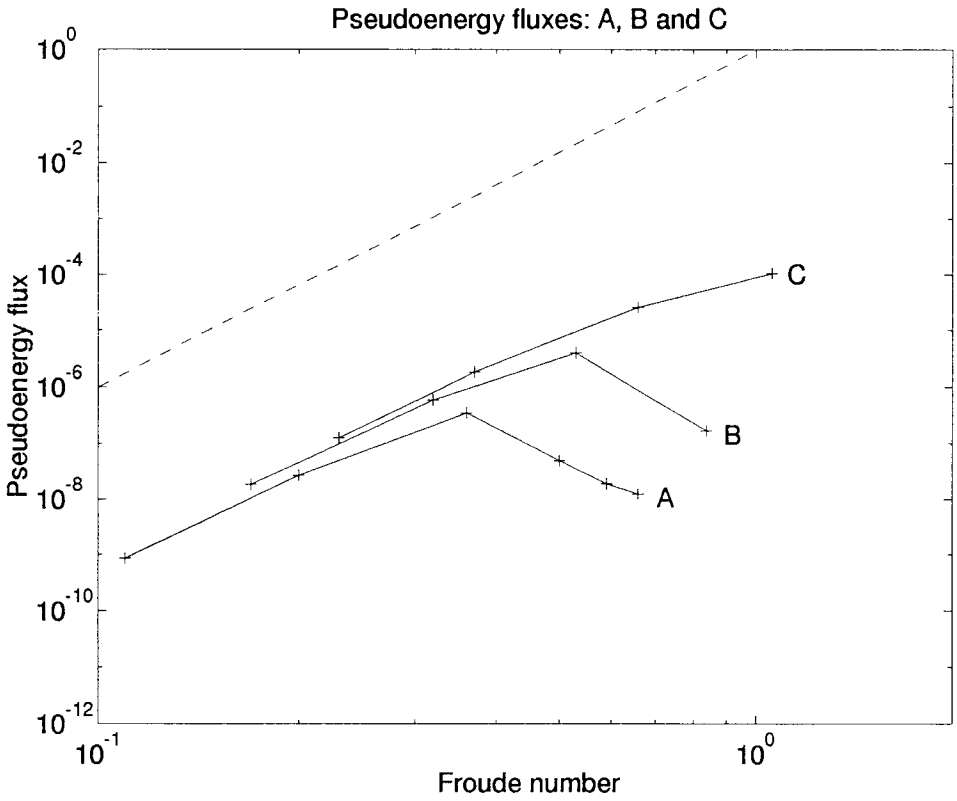


Figure 4.25: The peak pseudoenergy flux due to gravity wave radiation throughout simulations **A**, **B** and **C**

consider strips with non-negative potential vorticity, and then proceed in the following section, §4.9, to consider strips with negative potential vorticity.

4.8 Anticyclonic strips with non-negative potential vorticity

In this section, two sets of simulations are presented, with strip potential vorticity values of 0.1 and 0.0, to be labelled **D** and **E** respectively.

Now, at low Froude number, the vortical flow is identical for all Rossby numbers, except that the timescale is set by the magnitude of the potential vorticity jump, with strips with small jumps evolving more slowly than those with larger jumps. However, we know from simulations **A** – **C** that a potential vorticity jump of at least 5 is required if the pseudoenergy flux is to exceed 10^{-6} at a Froude number of 0.1. In this section we are concerned with anticyclonic but positive potential vorticity in the strip. We are therefore restricted to strip potential vorticity values between 1 and 0, and hence potential vorticity jumps between 0 and 1. It follows that we must expect very weak gravity wave radiation at low Froude numbers when the potential vorticity in the strip is between 0 and 1. Consequently, it was decided to concentrate on a strip potential vorticity value of 0.1, which is quite small (i.e. δQ is quite large for anticyclonic potential vorticity of the same sign as the background), and hence the flow will give rise to reasonably large radiated gravity wave amplitudes, without the strip potential vorticity being so close to zero that very large Froude numbers would be required to establish any significant differences between the two cases **D** and **E**.

Six simulations were performed with a strip potential vorticity of 0.1, with initial strip widths of 0.42, 0.84, 1.68. 3.36, 6.00 and 9.00, labelled **D**i–vi respectively. The vortical aspects of the dynamics in these simulations are similar to those found for simulations **A** – **C**. Figure 4.26 shows the dependence of the nutation time on the Froude number for experiments **A**, **B**, **C** and **D**. In common with the strips with cyclonic potential vorticity, the nutation times for experiments **D**i-vi increase super-exponentially with increasing Froude number.

However, agreement with the Lighthill theory is not found to be very good, even in the case **D**i, with a Froude number of only 0.2. Figure 4.27 shown the x–averaged $\partial h/\partial t$ field during simulations **D**i and **D**ii. Although the Lighthill reconstruction is capturing the essential features, on the whole it is performing rather poorly, when compared with cyclonic experiments

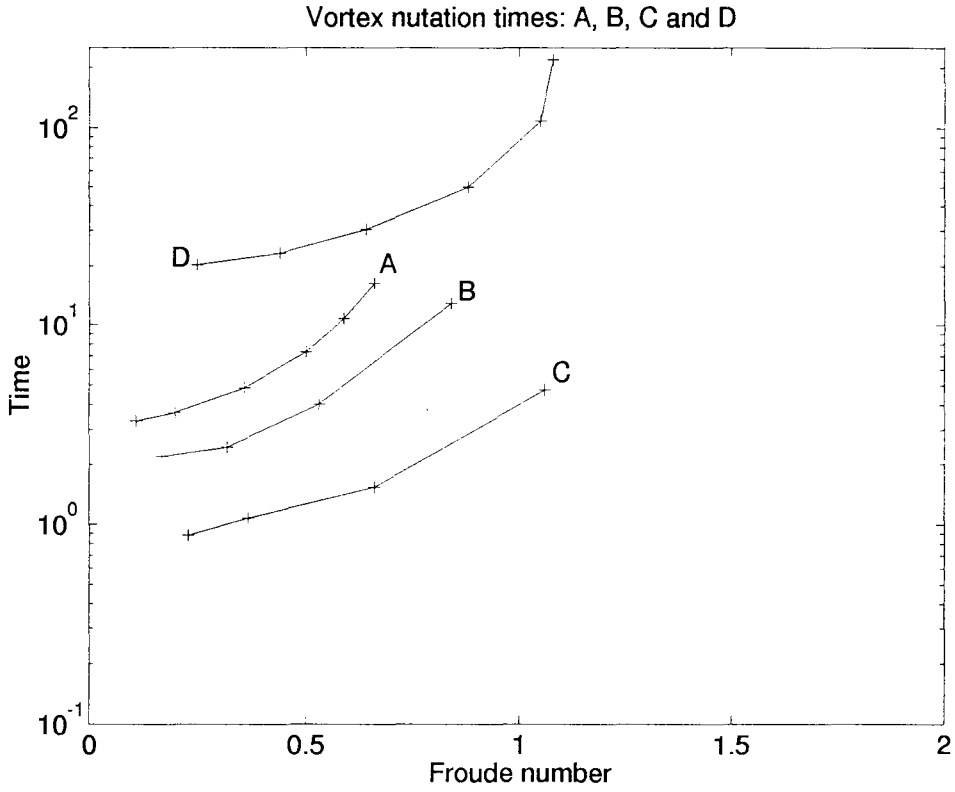


Figure 4.26: Vortex nutation times for simulations Di-vi

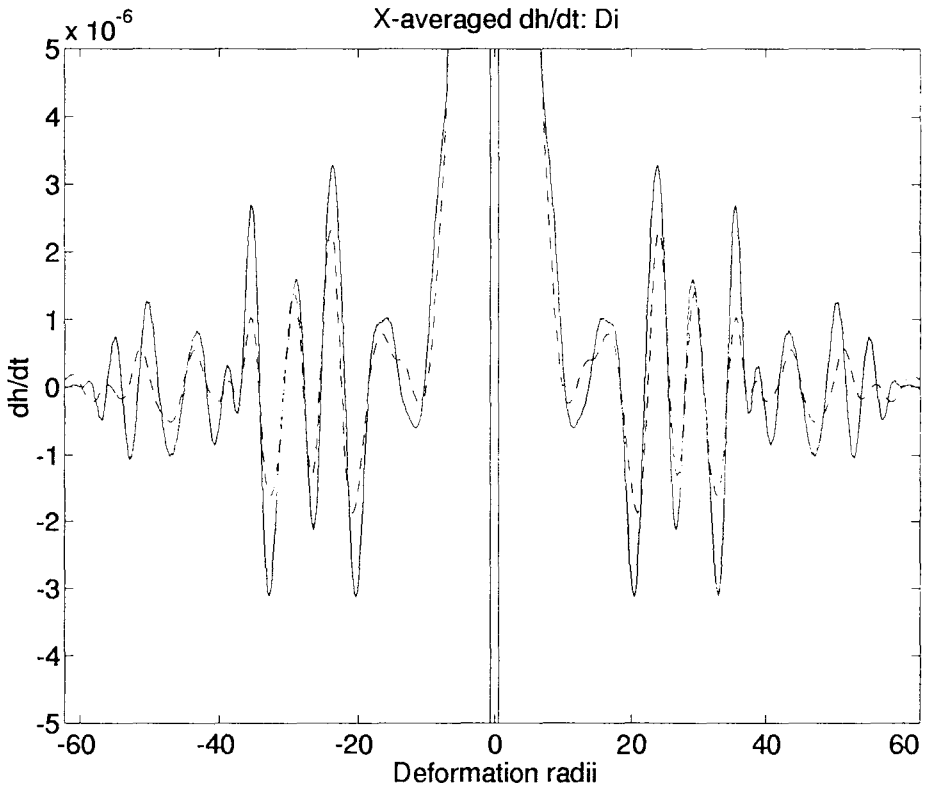


Figure 4.27: The x–averaged $\partial h/\partial t$ field at the termination of simulation **Di** (solid line), and its reconstruction using the Lighthill convolution integral (dashed line)

at similar Froude numbers.

Now, in classical aeroacoustics, only two length scales are present in the flow: the scale of the vortical motions, and the scale of the aeroacoustic waves which they generate. In the problem of gravity wave generation by vortical motions in a rotating frame, however, the Rossby deformation radius is an additional length scale, which is unity in the nondimensionalization used here. The Lighthill theory assumes that the quadrupole source may be concentrated at a single point. If this is to be a good approximation, the scale of the vortical flow must be small with respect to the wavelength of the waves it generates, which is the usual low Froude number assumption, and must also be small compared with a Rossby deformation radius. It seems likely that the reason for the rather poor agreement in figure 4.27 is that, although the Froude number is low, the vortices are not very small compared to a Rossby deformation radius, the initial width of the strip being 0.42.

In figure 4.28, the $\partial h/\partial t$ field is shown, at the time of most intense gravity wave emission, for simulations **D**i–iv. As the Froude number increases, the nature of the gravity wave generation appears to change. At low Froude numbers, the generation appears to be a rather long range effect, with gravity waves generated being of much longer wavelength than the scale of the vortical motions. At higher Froude numbers, however, the generation mechanism appears to be a rather more local effect, with gravity waves apparently being "launched" off the edges of the vortices as they rotate. Figure 4.29 shows a sequence of four frames from simulation **D**iv, during the period of generation of its most intense gravity waves, in which this process is illustrated, and figure 4.30 shows the potential vorticity field at the same times.

As the Froude number is increased still further, the gravity waves radiated by the vortex train lose their coherent x–independent structure. Figure 4.31 shows the gravity wave field during simulations **D**v and **D**vi. In comparison with simulations **D**iii and **D**iv, these gravity wave fields appear to be more arc-like, resembling radiation from point sources, rather than x–independent radiation from a line source exhibited at lower Froude numbers.

Figure 4.32 shows the dependence of the gravity wave pseudoenergy flux on the Froude number for experiments **A** – **D**. As in the case of experiments **A** – **C**, experiments **D**i-vi exhibit an optimal Froude number, above which the gravity wave pseudoenergy flux decreases as the Froude number is further increased. The existence of an optimal Froude number is almost certainly due to the fact that the vortex rotation times are increasing very rapidly with Froude

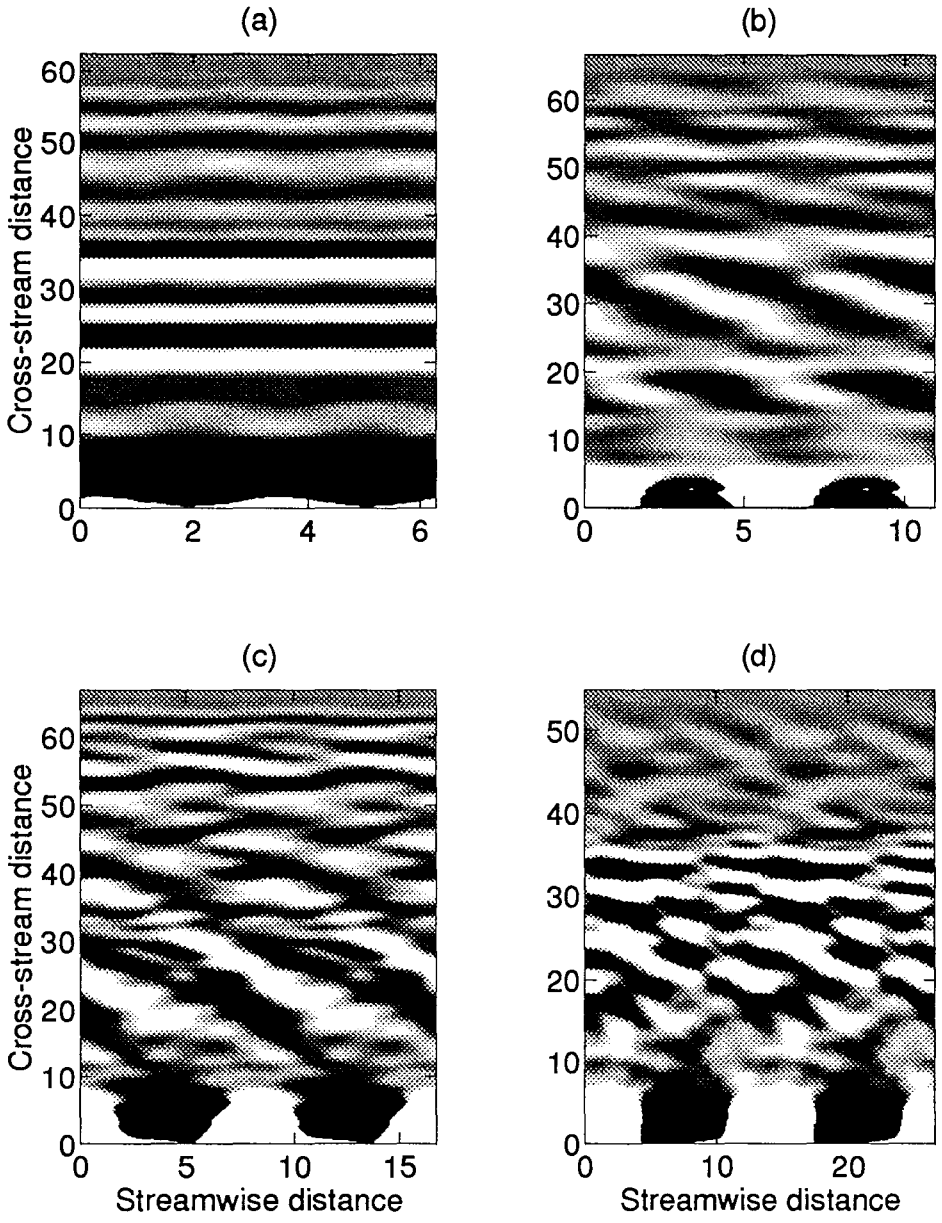


Figure 4.28: The $\partial h/\partial t$ field during the most active phase of gravity wave generation in simulations Di–iv. Two periods of the model domain are shown.

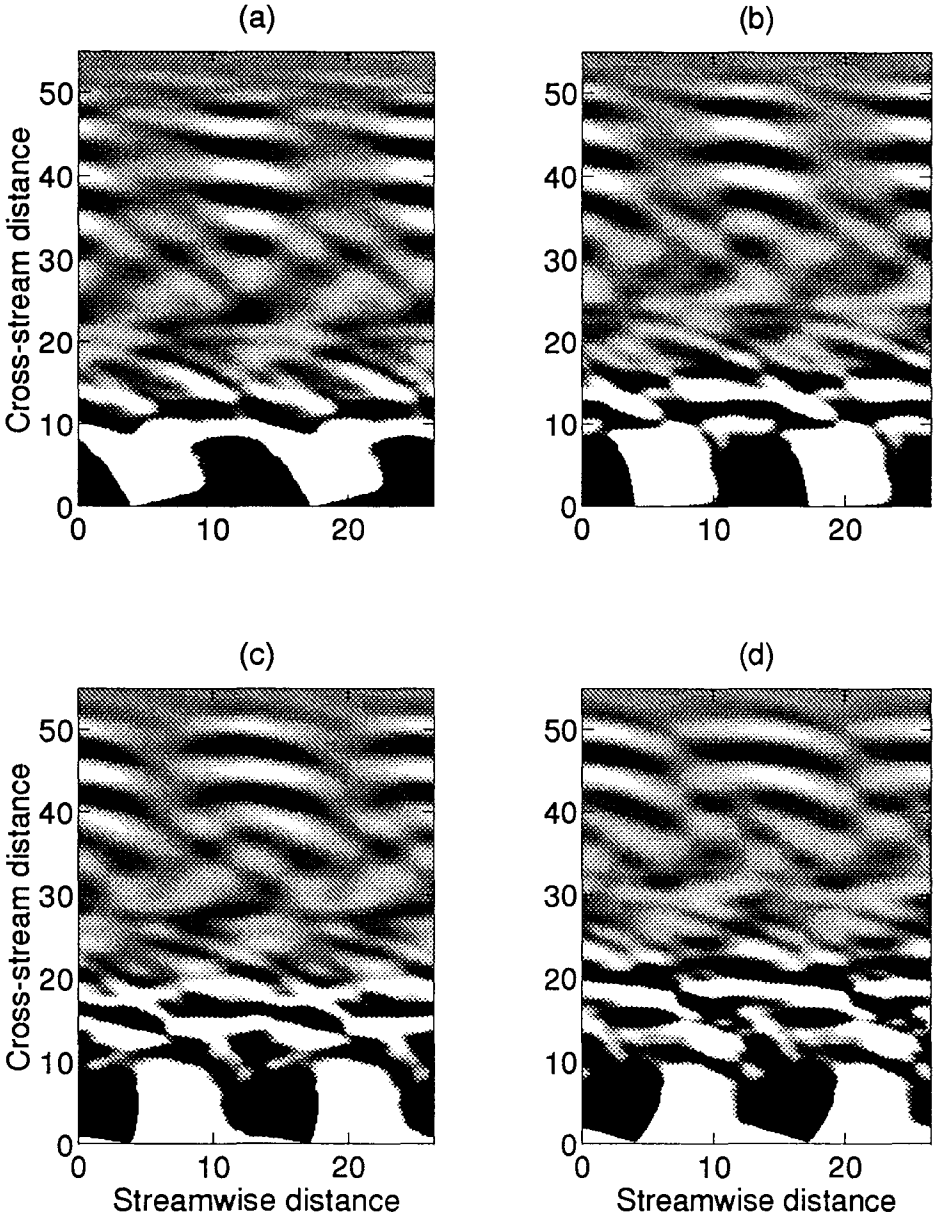


Figure 4.29: The stages of gravity wave emission during simulation Div. Figures a,b,c,d correspond to times 140, 143.75, 147.5 and 151.25 respectively. Two periods of the model domain are shown.

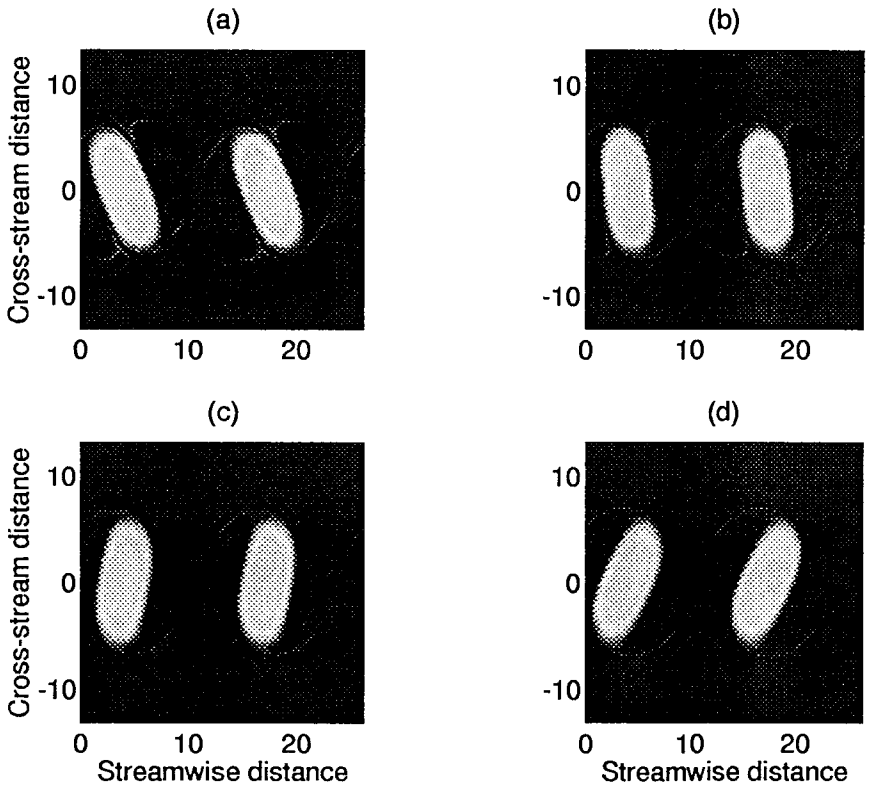


Figure 4.30: The potential vorticity field during the stages of gravity wave emission during simulation **D**iv. Figures a,b,c,d correspond to times 140, 143.75, 147.5 and 151.25 respectively. Two periods of the model domain are shown.

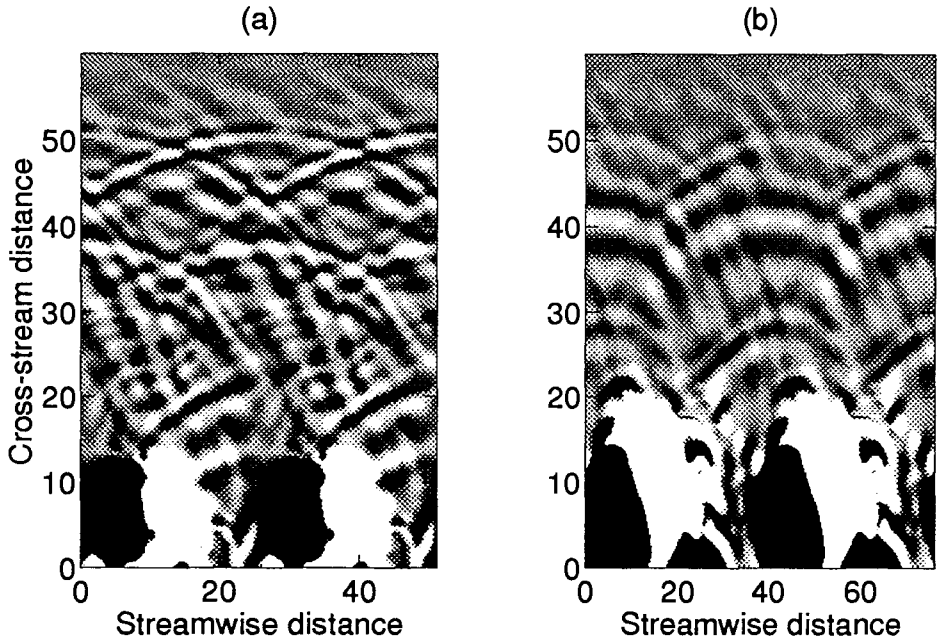


Figure 4.31: The $\partial h/\partial t$ field during simulations **D**v and **D**vi. Two periods of the model domain are shown.

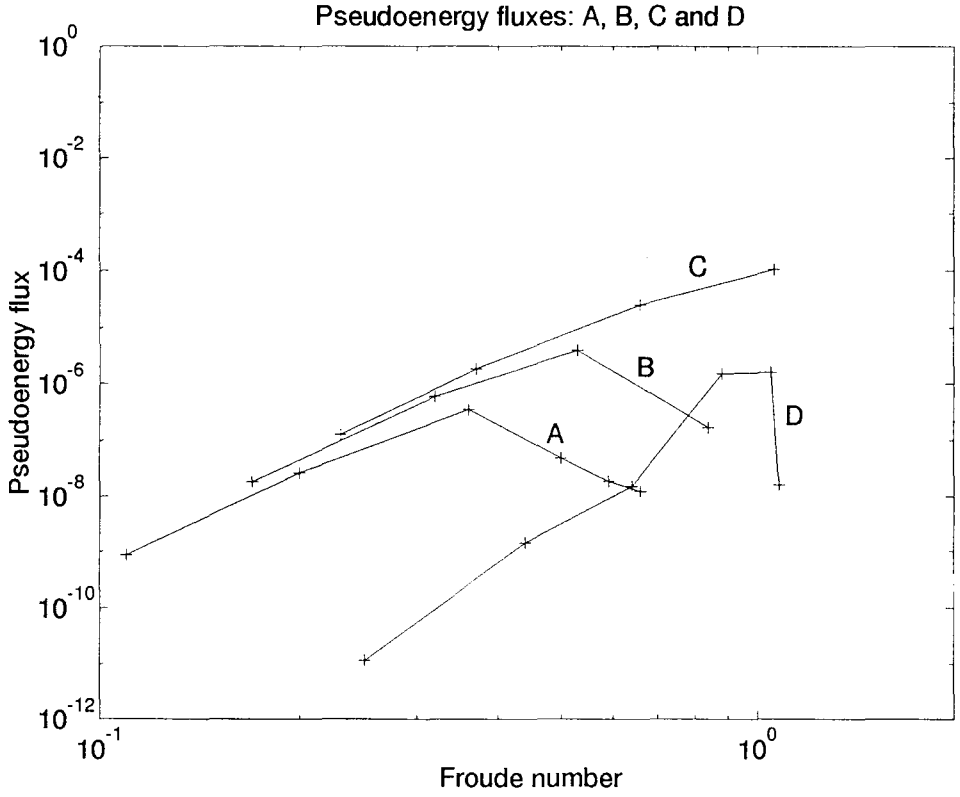


Figure 4.32: The maximum pseudoenergy flux plotted against Froude number for simulations Di-vi

number, and hence the gravity waves are severely inhibited by the presence of background rotation.

It seems clear that the nature of the gravity wave radiation, and its ultimate limitation, is somewhat different in simulations **D**i–vi from that found in simulations **A** – **C**. The gravity wave pseudoenergy flux appears to increase more rapidly with Froude number, at moderate Froude numbers, and the transition from increasing gravity wave amplitudes to decreasing amplitudes as the Froude number is increased is much more abrupt than is found for cases **A** – **C**, with cyclonic potential vorticity in the strip. However, analysis in this case is difficult because, unlike the cyclonic simulations **A** – **C**, the vortices are very much larger, when compared with a Rossby deformation radius, at the critical Froude number at which the transition from increasing to decreasing pseudoenergy fluxes occurs. This means that the Lighthill theory is inapplicable in this limit, and an alternative analytical description of the flow has not been attempted.

The experiments with small positive potential vorticity in the strip are to be contrasted with experiments in which the potential vorticity in the strip is exactly zero. Five experiments were conducted with a strip potential vorticity of zero, with initial strip widths of 0.42, 0.84, 1.26, 2.52 and 6.00. The experiments are labelled **E**i-v repectively.

Figure 4.33 shows the dependence of the vortex nutation times on the Froude number for experiments **E**i-v, with those for experiments **A** – **D** for comparison. Although there appears to be a very gradual increase in vortex nutation times with Froude number in this case, the behaviour is to be contrasted sharply with that found for a strip potential vorticity of just 0.1. A very significant difference in the nature of the vortical flows is thus observed between simulations **D** and **E**, although the general nature of the flow remains broadly similar: the strip rolls up into a periodic chain of vortices which proceed to nutate, albeit at markedly different rates between simulations **D** and **E**, without significant change of form.

The $\partial h/\partial t$ field during simulations **E**ii–v is shown in figure 4.34. At lower Froude numbers, it appears to behave in a similar way to the wave field in simulations **D**, with the gravity waves being launched from the edges of the vortices as they rotate. However, as the Froude number is increased, the radiated wave field does not become less coherent, but rather the wave crests of the launched waves become progressively sharper, almost resembling shock waves in the highest Froude number simulation **E**v.

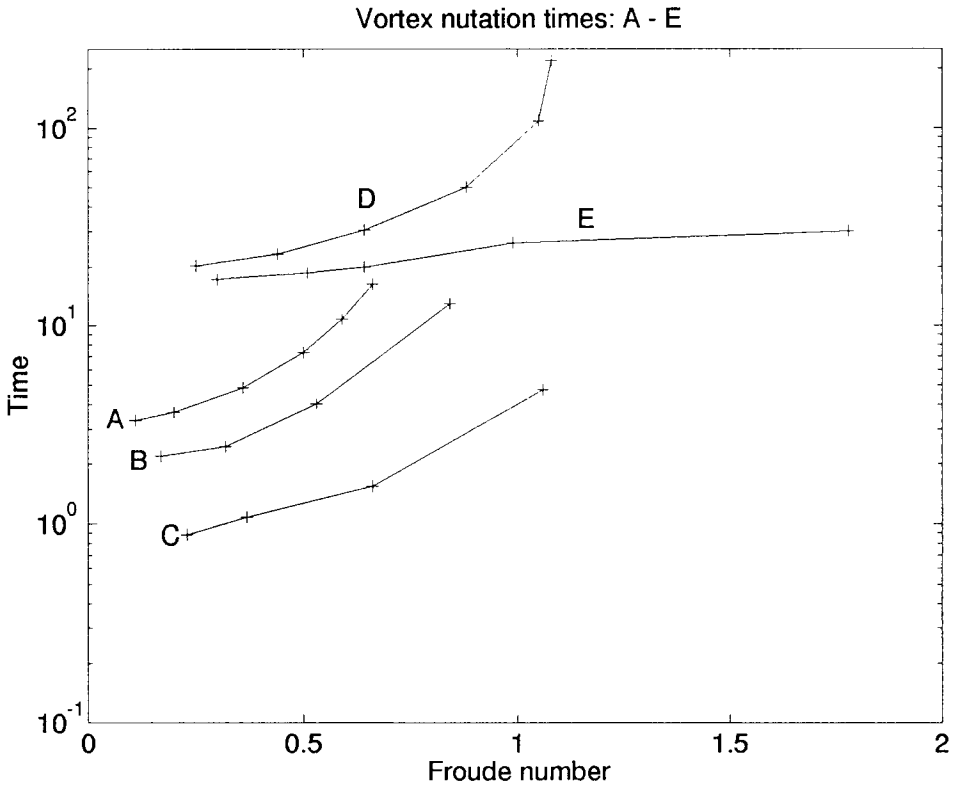


Figure 4.33: Vortex nutation times for simulations Ei-v

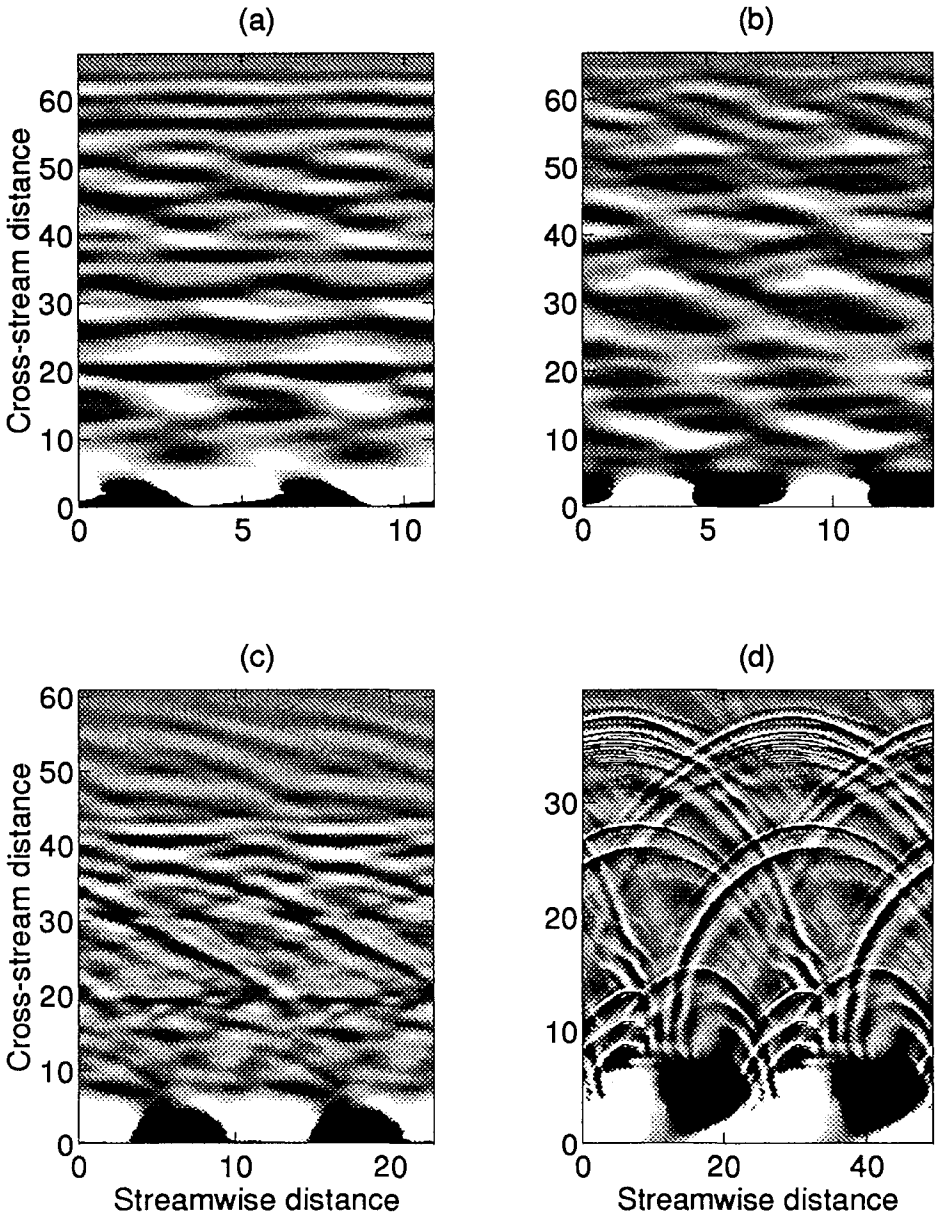


Figure 4.34: The $\partial h/\partial t$ field during simulations Eii-v. Two periods of the model domain are shown.

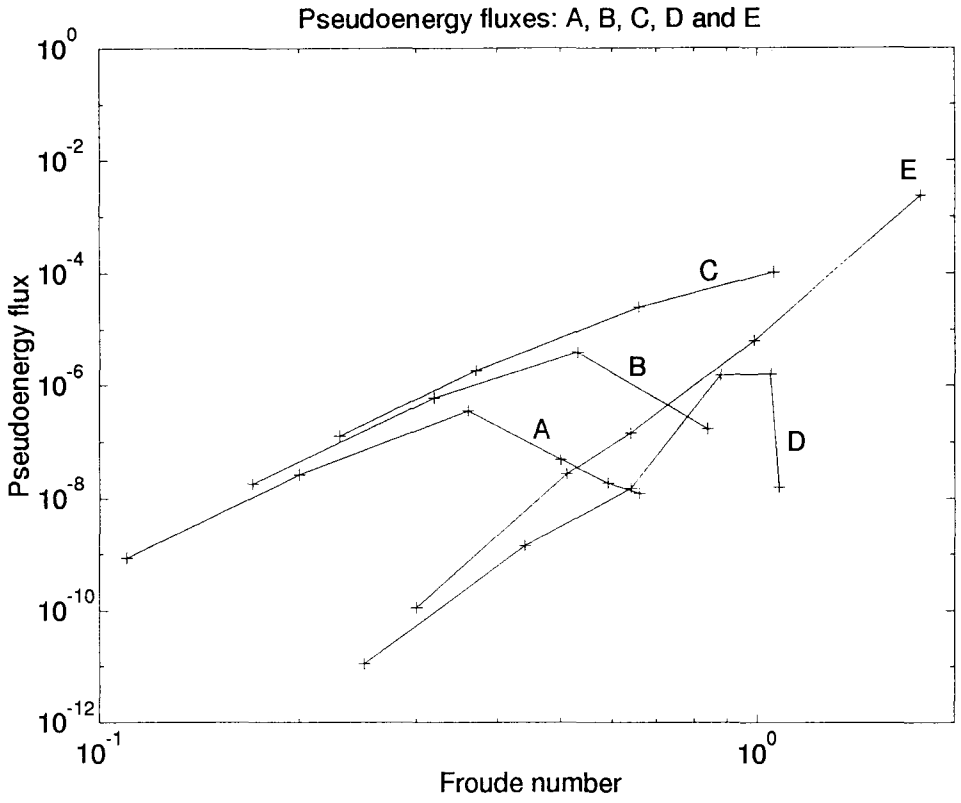


Figure 4.35: Maximum pseudoenergy flux plotted against Froude number for simulations Ei-v

The maximum pseudoenergy flux found in the gravity wave field in simulations Ei-v is shown in figure 4.35, with that found for experiments **A** – **D** for comparison. At moderate Froude numbers the pseudoenergy flux increases strongly with Froude number, in parallel with that observed for simulations Di-iv. However, in case **E** the pseudoenergy flux increases with increasing Froude number apparently without bound, with the gravity wave field apparently becoming highly nonlinear, and somewhat shock–like, at the largest Froude numbers.

4.9 Anticyclonic strips with negative potential vorticity

Having now classified the behaviour of the flows with non-negative potential vorticity, three further sets of simulations were undertaken to investigate gravity wave generation by vortex trains with negative potential vorticity. Strip potential vorticity values of -0.1, -1.0 and -19.0 were investigated, with a range of strip widths in each case.

The experiments with a strip potential vorticity of -0.1 were performed for comparison

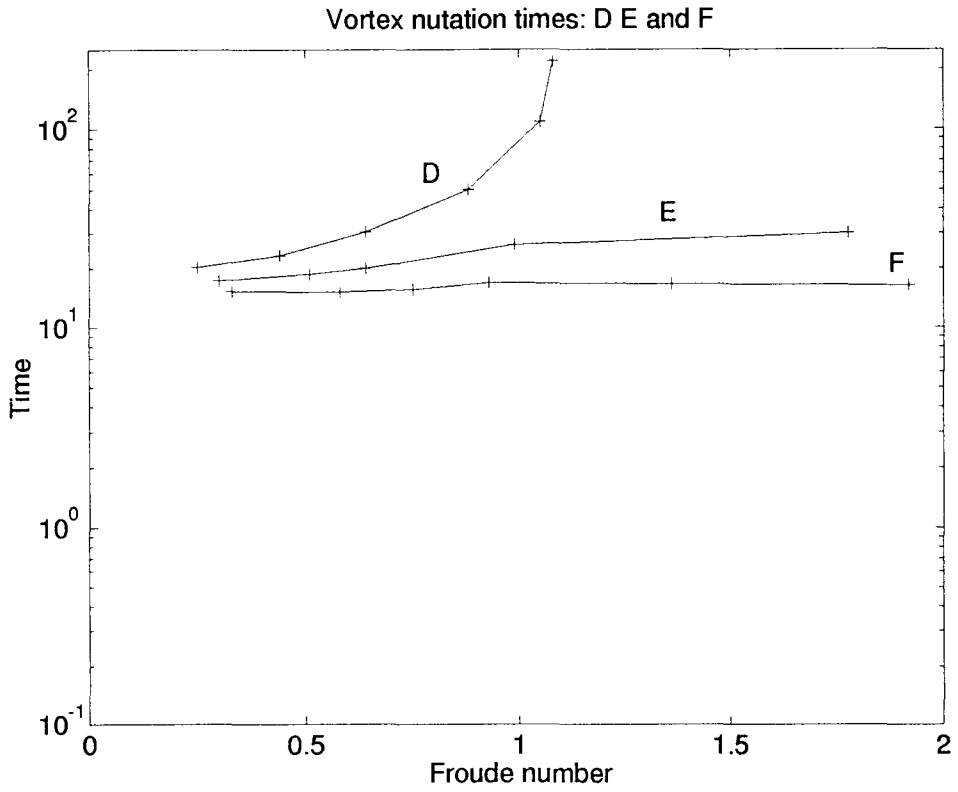


Figure 4.36: Vortex nutation times for simulations **F**i-v

with experiments **E**i-v, in which the strip potential vorticity was zero, to investigate whether any additional effects occurred when slightly negative potential vorticity was present in the simulations. Six simulations were performed with a strip potential vorticity of -0.1. Initial strip widths taken were 0.42, 0.84, 1.26, 1.68, 2.52 and 3.04. Figure 4.36 shows the vortex nutation times for simulations **D**, **E** and **F**. One can see that although there is a substantial difference between a strip potential vorticity of 0.1 and 0.0, there appears to be comparatively little difference between strip potential vorticity of 0.0 and -0.1.

In figure 4.37 the gravity wave pseudoenergy fluxes are shown for simulations **D**, **E** and **F**. Again, although there is a substantial difference between simulations **D** and **E**, there is almost no qualitative difference between simulations **E** and **F**.

The $\partial h/\partial t$ field for simulations **F**iii-vi is shown in figure 4.38. One can readily see that, at large Froude number, both sets of simulations **E** and **F** are tending to produce shock waves, and neither appears to be exhibiting any tendency for the pseudoenergy flux to stop increasing

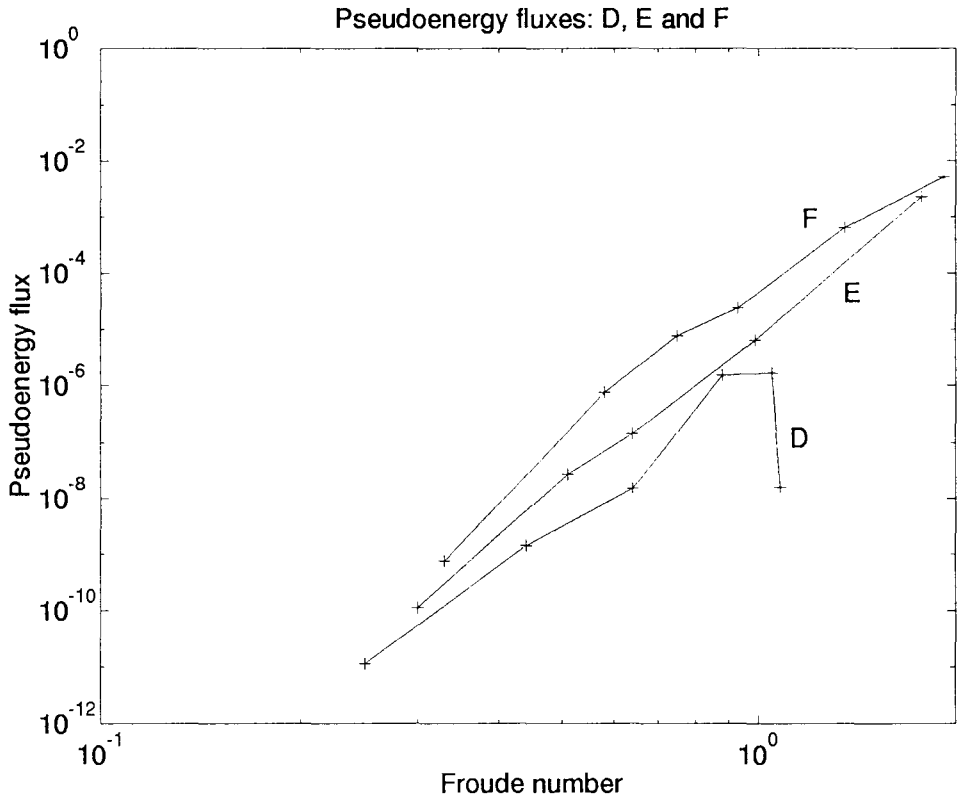


Figure 4.37: Maximum pseudoenergy flux against Froude numbers for simulation sets **D**, **E** and **F**

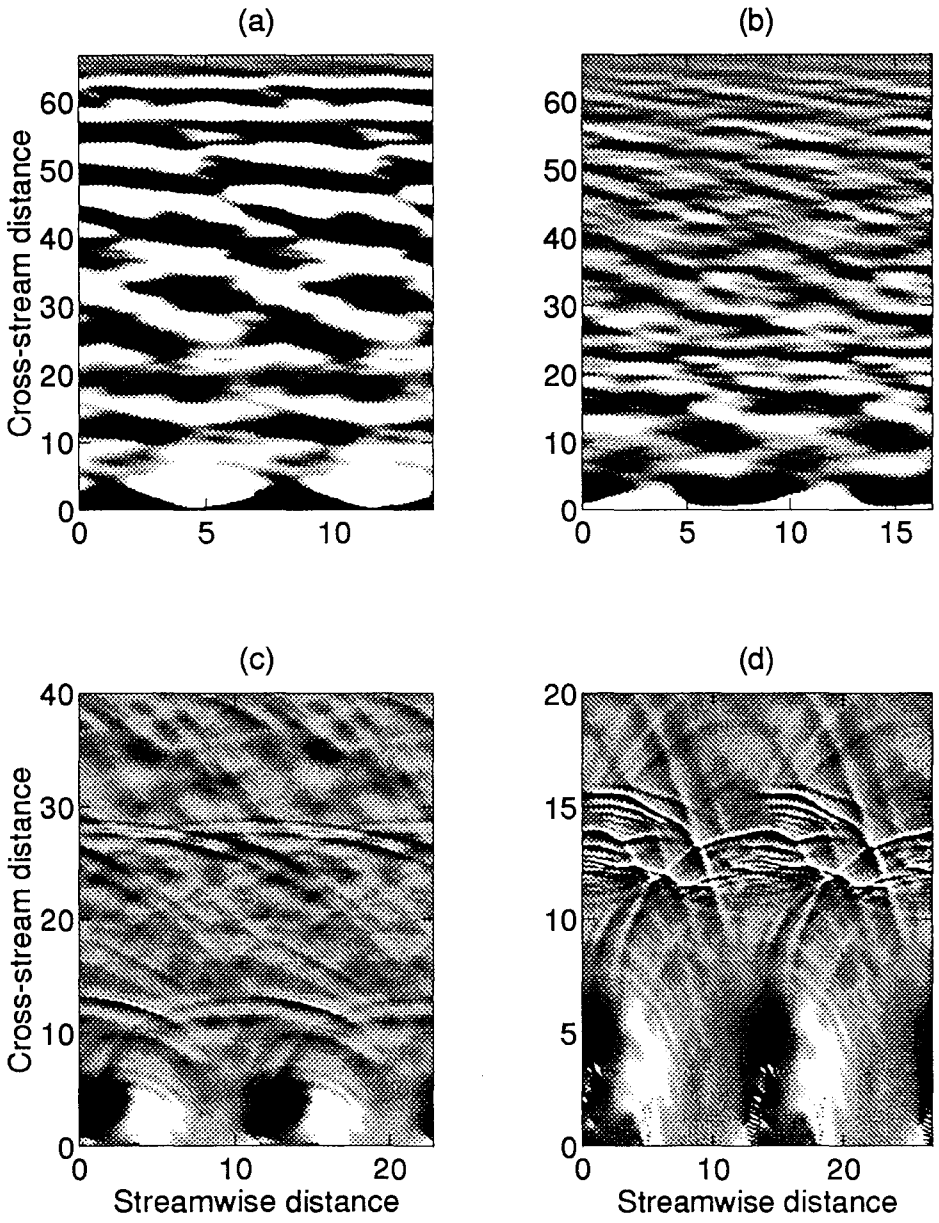


Figure 4.38: The $\partial h/\partial t$ field during simulations Fiii-vi. Two periods of the model domain are shown.

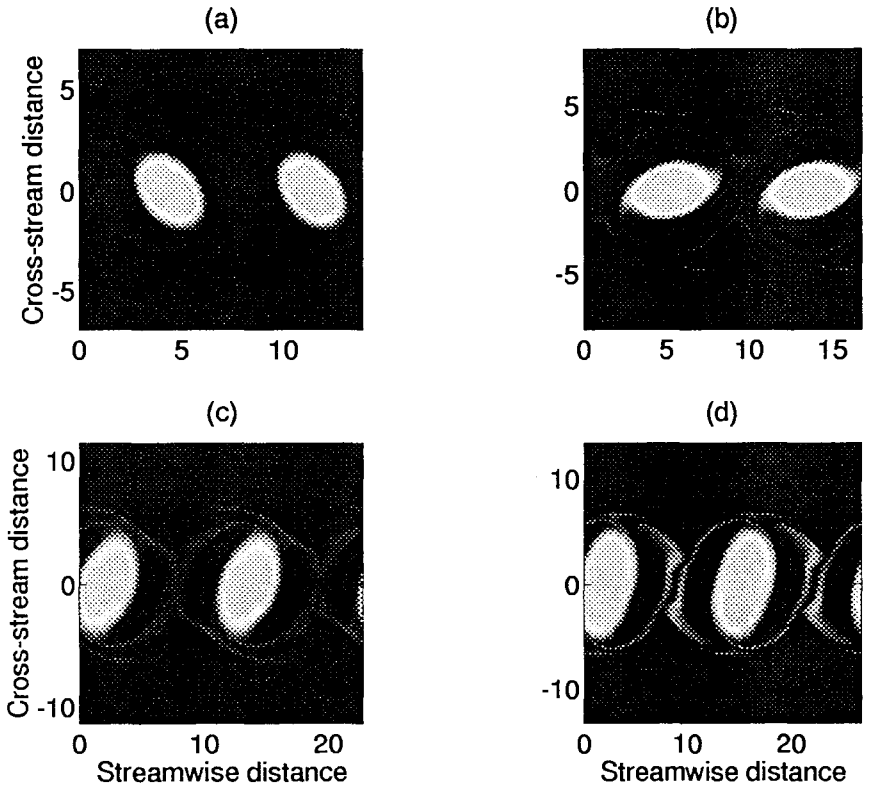


Figure 4.39: The potential vorticity field during simulations **F**iii-vi

with increasing Froude number. It seems reasonable to conclude that there is no appreciable difference between strips with zero potential vorticity, and strips with slightly negative potential vorticity.

The potential vorticity field corresponding to figures 4.38a–d is shown in figures 4.39a–d respectively. In simulation **F**vi, one can see in figure 4.38d that there is some small-scale noise in the $\partial h/\partial t$ field, in the vortical region, and hence the flow may be under–resolved. However, the potential vorticity in figure 4.39d does not exhibit significant small-scale noise, and so we can have some confidence that the performance of the model is at least reasonably good. even in this extreme case with a Froude number of almost 2.0.

Four simulations were then performed with a strip potential vorticity of -1.0. The strips had initial strip widths of 0.07, 0.21, 0.42 and 0.63, and are labelled **G**i-iv. The vortex nutations times do not appear to change significantly with Froude number, and are shown in figure 4.40, with those for simulations **D** – **F** for comparison.

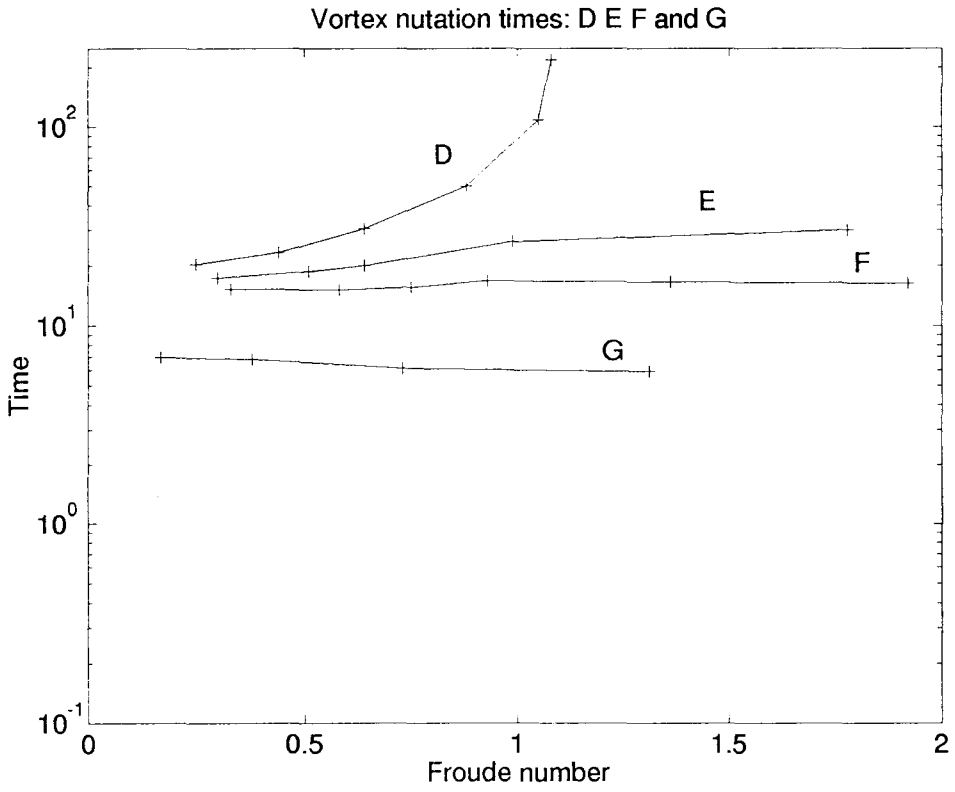


Figure 4.40: Vortex nutation times for simulations Gi-iv

The size of the vortices compared to a Rossby deformation radius is now much smaller than was the case in simulations **D**, **E** and **F**, with the largest initial strip width being 0.63, only 50% wider than the smallest initial strip with in simulations **D**, **E** and **F**. Consequently, the Lighthill theory is significantly more successful at reconstructing the wave field. The x–averaged $\partial h/\partial t$ field, and its reconstruction from the Lighthill source term, is shown for simulations **G**i-iii in figure 4.41, and good agreement is obtained for initial strip widths of 0.07 and 0.21.

However, in line with experiments **E** and **F**, the character of the wave field does depart significantly from that of linear waves as the Froude number is increased. In figure 4.42, the $\partial h/\partial t$ field is shown at the end of each simulation for the cases **G**i-iv. In the first three cases the wave field appears to be well described by linear waves. At low Froude numbers the wave field is dominated by x–independent waves, with x–dependence increasing with increasing Froude number in simulations **G**i-iii. Indeed, linear wave-like motions in the gravity wave field are a pre-requisite for good agreement with the Lighthill theory, observed in figures 4.41a&b, which assumes a linear radiating wave field. In simulation **G**iv, however, the character of the wave field differs markedly from linear wave-like motions, with shock waves appearing in the radiating gravity wave field. In common with simulations **E**v and **F**v, the shocks in simulation **G**iv also appear to originate in the vortical region itself, although they appear to be somewhat sharper in this case.

At this point it should be recognised that the discretization scheme is not designed to capture shocks, and any quantitative discussion of differences between shocks is inappropriate. However, it seems unlikely that a more sophisticated numerical scheme would significantly affect the existence of sharp shock structures in these high Froude number strongly anticyclonic simulations.

The maximum x–averaged pseudoenergy flux is shown against the Froude number in figure 4.43, with the flux for simulations **D** – **F** for comparison. At low Froude numbers, the pseudoenergy flux is increasing as the sixth power of the Froude number. The increase then appears to become more rapid as the Froude number increases, in common with with other anticyclonic simulations **D** – **F**, in which it appeared that the pseudoenergy flux was increasing more rapidly than the sixth power of the Froude number. These do, of course, differ from the cyclonic simulations, in which the pseudoenergy flux never increased more rapidly than the sixth power of the Froude number.

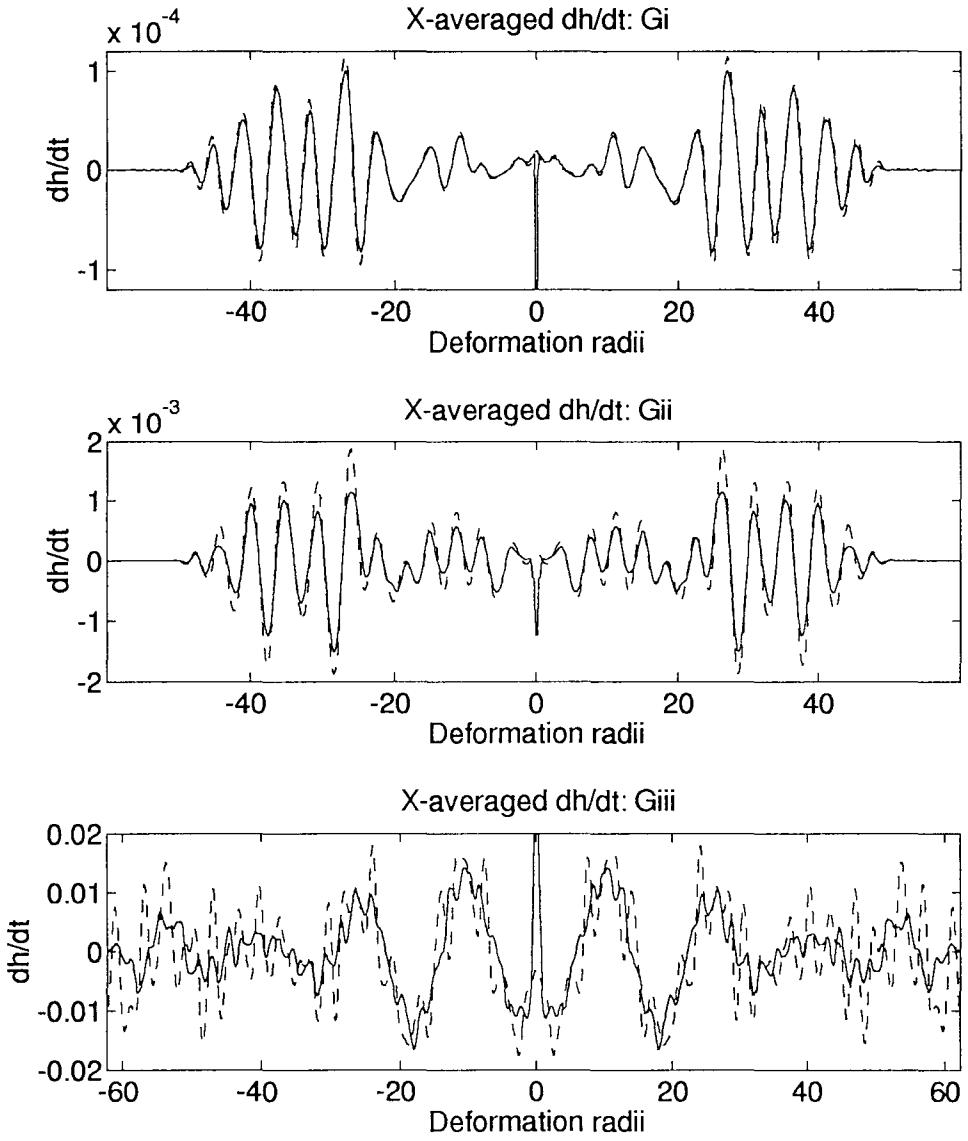


Figure 4.41: The x–averaged $\partial h/\partial t$ field at the termination of simulations **G**i-iii (solid line), and its reconstruction from the Lighthill convolution integral (dashed line)

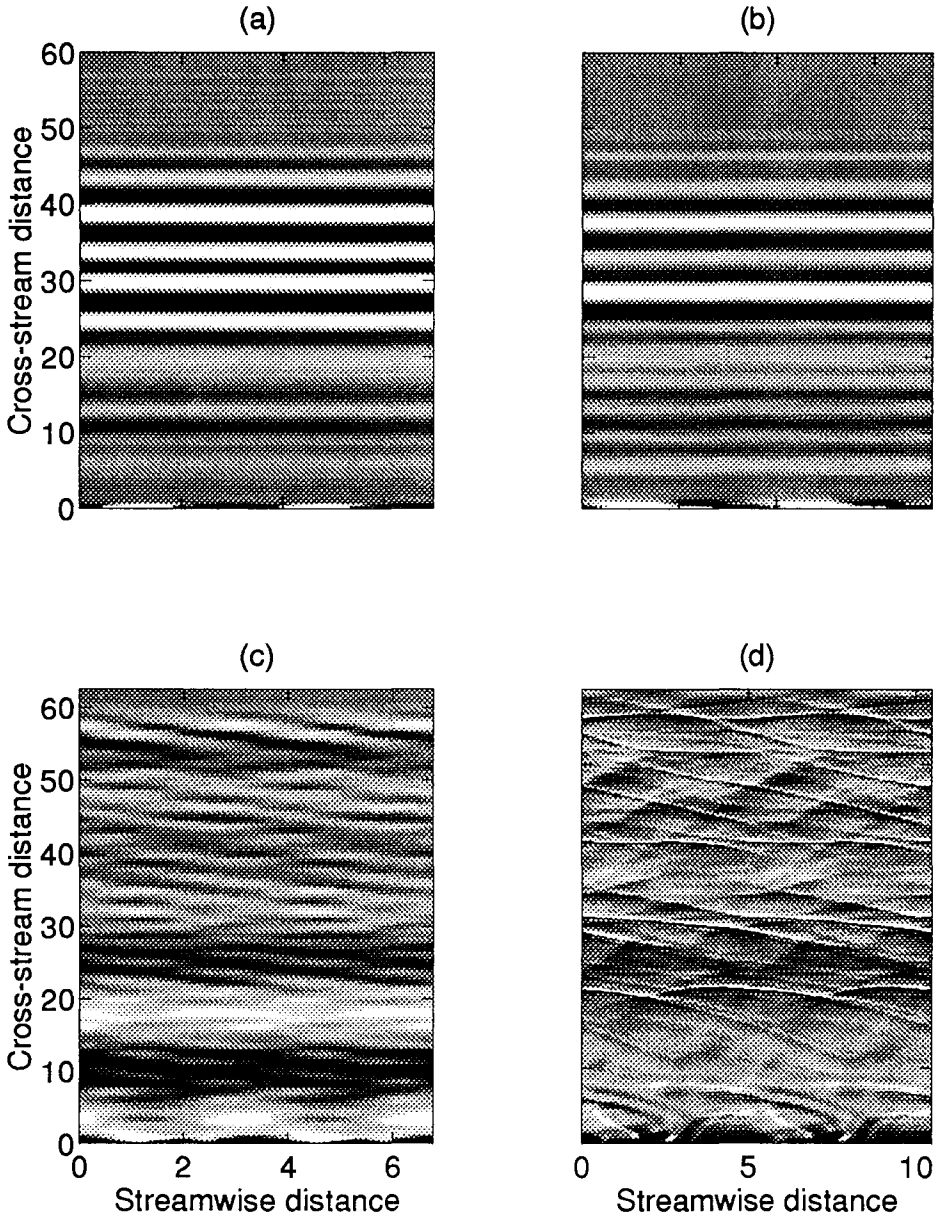


Figure 4.42: The $\partial h/\partial t$ field during the simulations Gi-iv. Two periods of the model domain are shown.

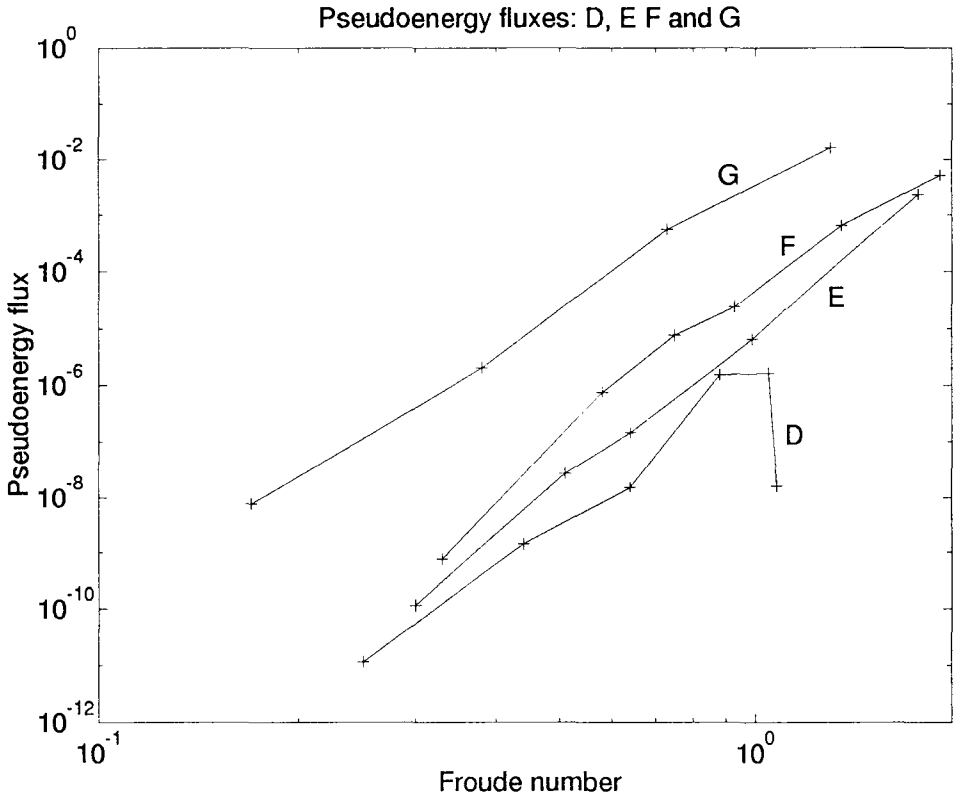


Figure 4.43: The maximum pseudoenergy flux against Froude number for simulations Gi-iv

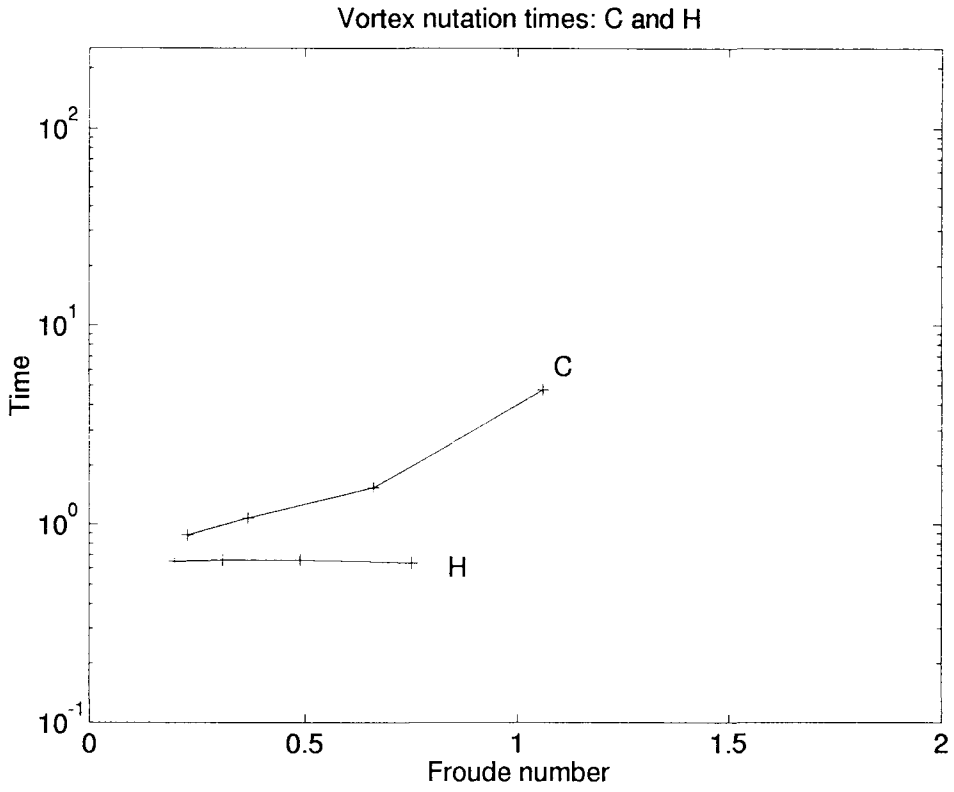


Figure 4.44: Vortex nutation times for simulations **Hi**-iv

Finally, four simulations were performed with a strip potential vorticity of -19.0. These simulations were performed for comparison with simulations **Ci**-iv, in which the strip potential vorticity was taken to be 21.0. Consequently, at low Froude numbers, the dynamics of the strips should be identical in the two cases, with the magnitude of the potential vorticity jump being equal to 20 in both cases. Strip widths of 0.007, 0.013, 0.021 and 0.031 were taken, and the simulations are labelled **Hi**-iv respectively.

Figure 4.44 shows the vortex nutation times for simulations **C** and **H** as a function of Froude number. It can clearly be seen that when there is negative potential vorticity in the strip the nutation times remain almost constant as the Froude number is increased, whereas in the case of positive potential vorticity in the strip the nutation times increase superexponentially.

In figure 4.45 the $\partial h/\partial t$ field is shown at the termination of simulations **Hi**-iv, and in figure 4.46 the comparison with the Lighthill theory is shown. The agreement seems quite good at low Froude numbers, but not as good as in simulations **Ci**-iv, even though the waves remain

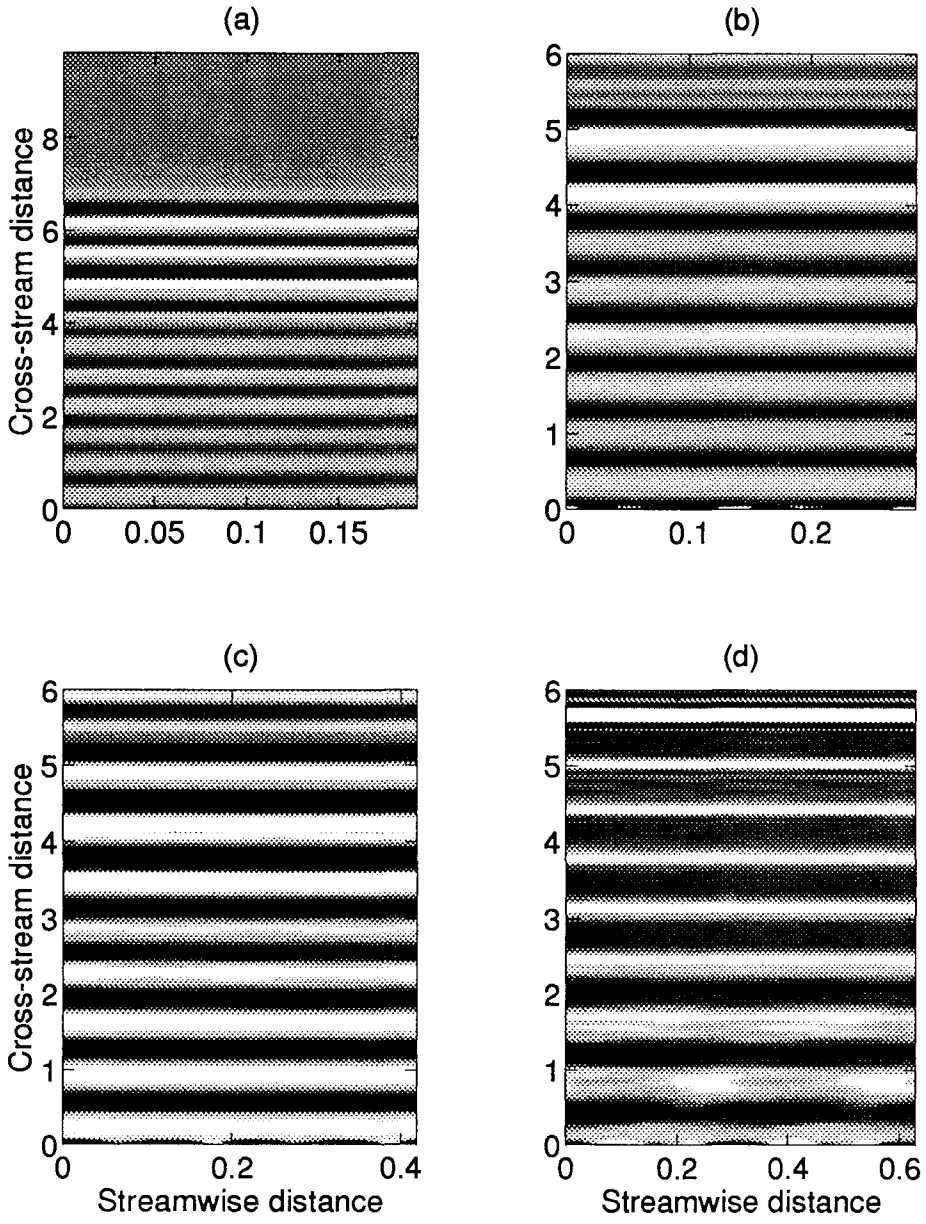


Figure 4.45: The $\partial h/\partial t$ field at the termination of simulations Hi-iv. Two periods of the model domain are shown.

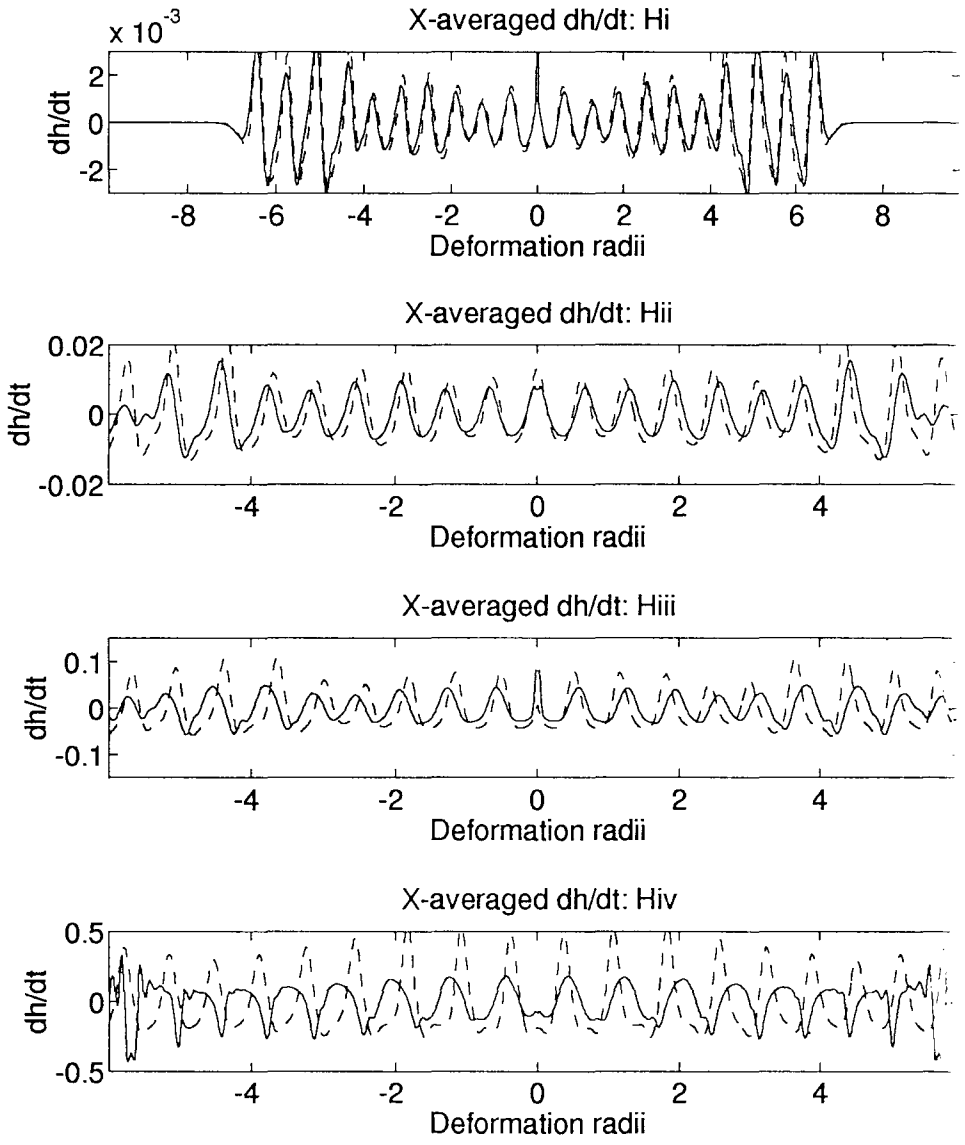


Figure 4.46: The x-averaged $\partial h/\partial t$ field at the termination of simulations H and its reconstruction from the Lighthill convolution integral (dashed lines)

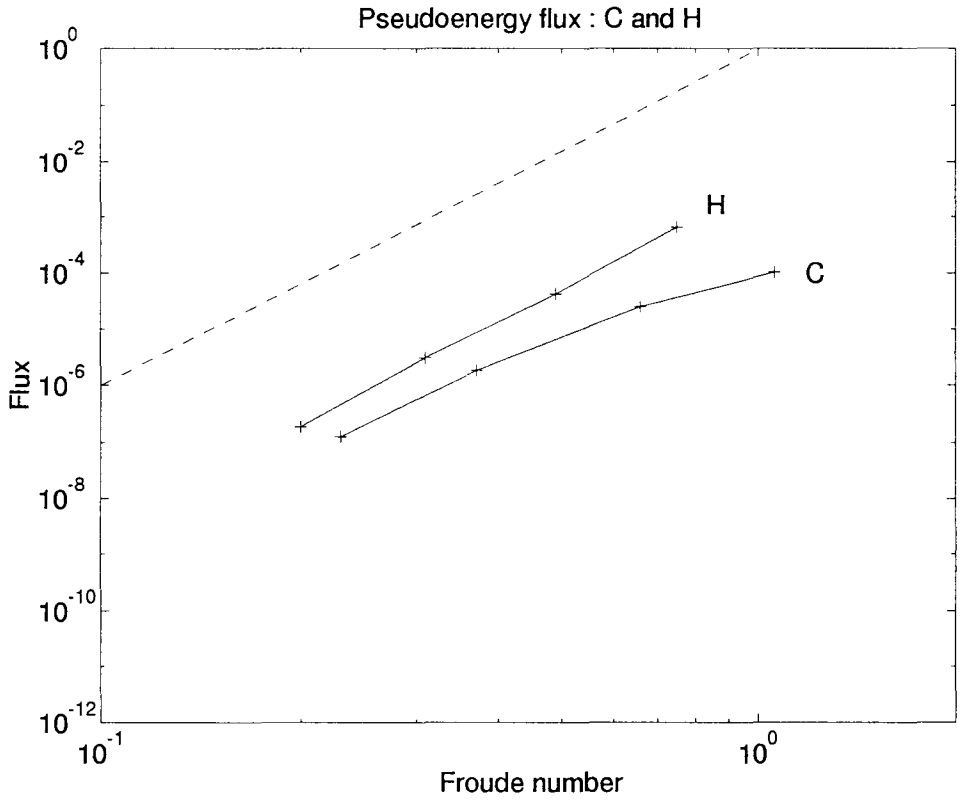


Figure 4.47: Maximum pseudoenergy flux plotted against Froude number for simulations **Hi-iv**

nearly independent of x.

The peak pseudoenergy flux in the gravity wave far field is shown against Froude number for simulations **Ci-iv** and **Hi-iv** in figure 4.47. One can see that, although both sets of simulations are initially quite similar, as the Froude number is increased the simulations **Hi-iv**, with negative potential vorticity in the strip, radiate gravity waves much more strongly that the corresponding cyclonic simulations with the same magnitude of potential vorticity difference between the strip and the background.

When comparing the x–averaged $\partial h/\partial t$ field with its reconstruction from the Lighthll convolution integral (4.7), it is of interest to note that, while in most cases the maxima of the actual and reconstructed field co-incide, in simulations **C** and **H**, shown in figures 4.23 and 4.47 respectively, the maxima appear to be offset. In figure 4.23, the maxima of the field obtained by reconstruction from (4.7) appear to lead those obtained by direct numerical simulation, whereas in 4.47, they tend to lag those obtained by direct numerical simulation. The reason is that the

Lighthill theory assumes that the wave phase speed c_0 is constant everywhere. However, the vortical flow adjusts to rest over the scale of a Rossby deformation radius. In most cases, the radiated waves are of long wavelength compared with a Rossby deformation radius. However, as the potential vorticity in the strip increases, so the frequency of the vortical motion increases, and hence the wavelength of the radiated waves decreases. In cases **C** and **H**, the wavelength of the radiated waves is of the order of a deformation radius, and therefore they may be significantly affected during their propagation through the region over which the flow relaxes adjusts to rest. In the cyclonic case **C**, the layer depth is less near the strip than at infinity, and the waves will tend to propagate more slowly there. Hence, the waves predicted by (4.7) will tend to lead those obtained by direct numerical simulation. Conversely, in the anticyclonic case **H**, the layer depth is greater near the strip than at infinity, and the waves will propagate more quickly there. Hence, the waves predicted by (4.7) will tend to lag those obtained by direct numerical simulation, in agreement with what is found in figures 4.23 and 4.47.

4.10 Discussion

Gravity wave radiation by a train of vortices in shallow water has been studied by direct numerical simulation. The dependence of the gravity waves radiated upon the Froude and Rossby numbers was investigated.

Arguably the most surprising result is that, for cyclonic vortices at moderate Rossby number, increasing the Froude number above 0.35 can lead to a decrease in the amplitude of the gravity waves radiated by the vortex train. This is in contrast to laboratory experiments, and the numerical experiments of Lele & Ho (1993) for a non-rotating gas, in which the acoustic wave amplitudes scaled according to the Lighthill theory for Mach numbers of up to 0.6 or so.

A summary of the results, showing the dependence of the radiated pseudoenergy flux on the Froude number for various Rossby numbers is shown in figure 4.48.

One can see that, in general, whereas positive potential vorticity trains tended to exhibit an "optimal" Froude number for gravity wave generation, the amplitude of gravity waves radiated by negative potential vorticity trains increase apparently without bound as the Froude number increases.

Typical vortex nutation times for all experiments are shown in figure 4.49. One can see

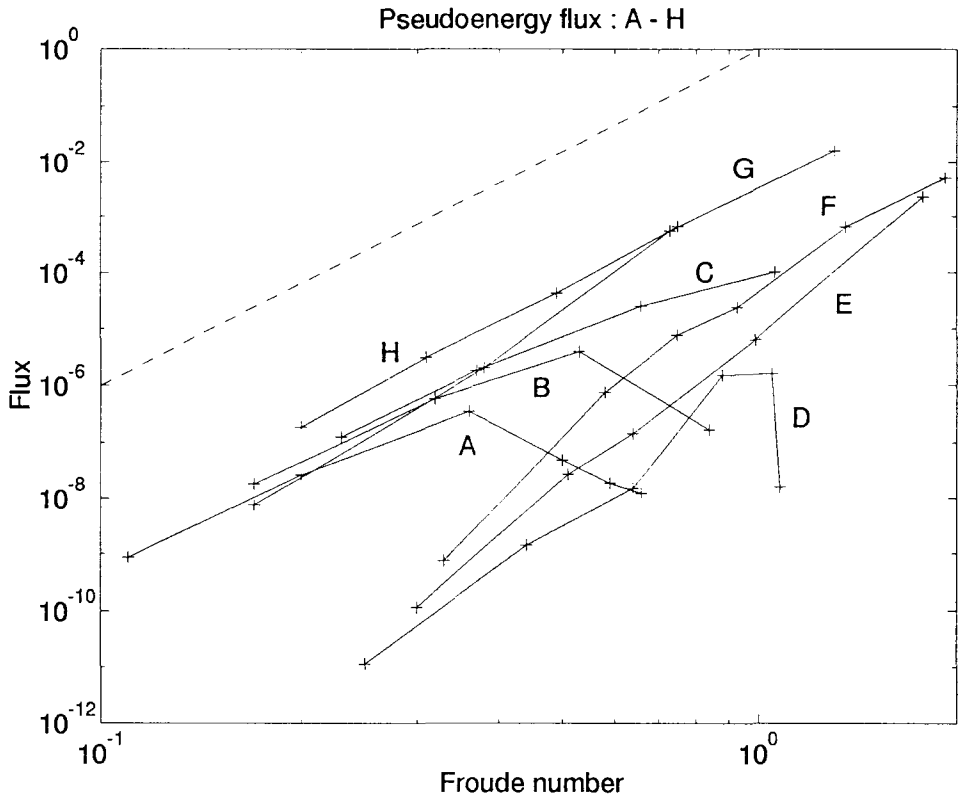


Figure 4.48: Pseudoenergy flux against Froude number for all simulations (solid lines). The gradient of the dashed line corresponds to pseudoenergy increasing as the sixth power of the Froude number

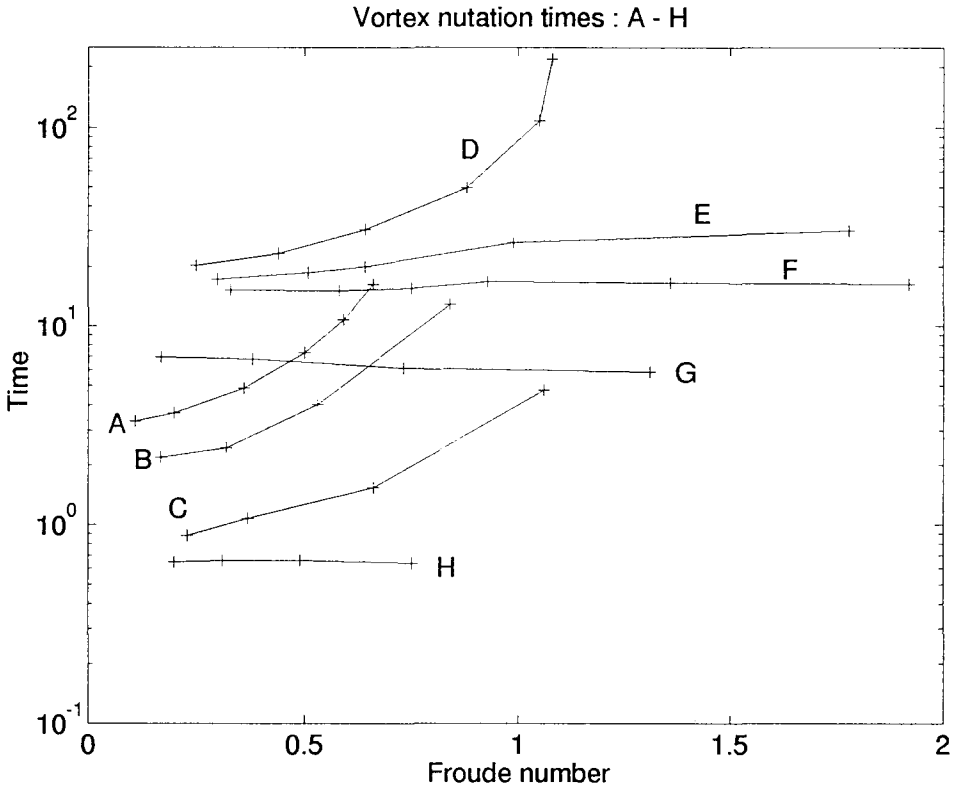


Figure 4.49: Vortex nutation times for all experiments, **A** – **H**

that nutation times increase super-exponentially with Froude number for all cyclonic vortex simulations, **A** – **C**, and also for anticyclonic simulations, when the potential vorticity in the strip is strictly positive. As a result, the intensity of the radiated gravity waves never exceeds about 10^{-4} in simulations **A** – **D**, as the inhibiting effect of the inertial frequency becomes increasingly significant as the nutation times increase. However, when the potential vorticity in the strip is zero or negative, there does not appear to be any significant dependence of the nutation time upon the Froude number, and the radiated gravity wave fluxes apparently increase without bound.

It might be suggested that symmetric instability might account for this difference. Symmetric instability, with wavenumber zero, will exist in parallel flow when $f(f - U_y)$ is negative in some part of the domain (Hoskins, 1974). Although his analysis is for a continuously stratified fluid, it can be readily extended to the shallow water equations (R. T. Pierrehumbert, personal communication). The condition $f(f - U_y) < 0$, which is both necessary and sufficient for symmetric instability, is met, at the initial instant, by flows **F**i-vi, but not by flows **E**i-vi.

However, there seems very little difference between simulations **E** and **F**, even though the initial states of **F** have an additional instability mechanism available to them, not present in the initial states of **E**. Therefore, it would appear that symmetric instability does not play a significant role in these simulations.

It should be noted, however, that all flows were initialized with an eigenmode at finite wavelength, corresponding to the fastest growing barotropic, or "Rossby wave–Rossby wave", instability mode. It may be of interest to repeat some of the experiments **F**, **G** and **H**, initializing with a symmetric instability mode of zero wavelength, or perhaps with a combination of the symmetric and barotropic instability modes, to investigate the characteristics of the gravity wave field which develops under their subsequent nonlinear evolution.

Some asymmetry between cyclones and anticyclones at Rossby and Froude numbers of order unity has been observed in numerical simulations by L. M. Polvani (personal communication). In his simulations, the geometry of the flow is doubly-periodic, and the initial vorticity field is random. The flow is balanced using the nonlinear balance equations (McWilliams, 1985). In time, large coherent vortices develop, and at small Froude number, there is symmetry between cyclonic and anticyclonic vortices. At larger Froude numbers, however, the flow is dominated by large coherent anticyclones, with much of the cyclonic potential vorticity smeared out into thin

filaments. In some simulations, shock waves have been observed, emanating from the anticyclones. Polvani's simulations thus broadly confirm the results presented here, that anticyclones are more vigorous features than cyclones in shallow water dynamics at order-one Froude numbers, capable of generating shock waves in the gravity wave field at sufficiently large Froude number.

Finally, we recall that this study was motivated, in part, by the observation by McIntyre & Norton (1993) that gravity waves of large amplitude were generally not generated by the vortical motions in their disturbed polar vortex simulations. Now, in polar vortex simulations, the majority of the vortical motions are cyclonic, and the potential vorticity is all of one sign throughout the hemispherical domain. Consequently, they are not able to access the regime of zero and negative potential vorticity, which has been demonstrated in this chapter to be the regime in which the strongest gravity wave generation might be expected. The results presented therefore support their findings, and offer some explanation for them.

Chapter 5

Discussion

5.1 Review of aims and objectives of this thesis

In the introduction to this thesis, it was stated that the aims of the thesis were:

- To investigate the generation of gravity waves in the shallow water system, and the dependence of their amplitudes and character on the Froude and Rossby numbers of the basic flow;
- To quantify the effect of the generation of gravity waves in the shallow water system on the potential vorticity distributions which generate them;
- To establish whether, without additional regard to flow geometry, there can or can not exist a slow manifold for the shallow water equations in any region of Froude and Rossby number space.

To be fair, it can not be said that these aims have been neatly divided up, with successive chapters each dealing with one particular aim. The investigation of gravity waves generated by vortical flows in the shallow water system has been an overall objective of the thesis, and each of the three substantive chapters of the thesis has made a significant contribution to it.

The general nature of gravity waves generated by vortical motions was investigated in chapter 2. The Lighthill theory, as extended to the rotating frame by Norton, was confirmed by a matched asymptotic analysis. The analysis enabled an expression to be derived for the effect of gravity wave radiation on the vortical flow. This expression would not be available through a "classical" application of the Lighthill theory, which would predict the amplitude and character of the gravity waves generated, but not their effect on the vortical flow which generates them.

The nature of the back reaction of the gravity waves on the vortical motions which generate them was then analysed with reference to balanced models. The back reaction takes the form of a large scale straining field, whose amplitude depends on time, and is given by a convolution integral involving the entire past history of the vortical flow. Hence it was shown that, without reference to an initial instant of time, the correction to the balanced dynamics which was required to account for the gravity wave radiation at low Froude number could not be regarded as an instantaneous function of the potential vorticity distribution, and could not therefore be regarded as part of the balanced dynamics.

The issue of the existence or non-existence of a slow manifold for the shallow water equations was pursued further in chapter 3, by analysing the character of the instability of an axisymmetric vortex with a potential vorticity which had a single discontinuity at $r = 1$. It was shown that the vortex was unstable to high wave number disturbances at arbitrary Froude and Rossby numbers. The instability is a coupled Rossby wave – gravity wave instability, and can not be part of a slow manifold which does not contain any gravity wave-like features. If a slow manifold were permitted to contain gravity wave-like features, it would then follow from the vortex instability analysis that it could not be unique. Thus, although the existence of a slow manifold can not be entirely ruled out, the notion of a unique slow manifold, entirely devoid of gravity wave activity, certainly can be.

Finally, nonlinear simulations of the shallow water equations were performed to investigate the nature of gravity wave generation by vortical flows at Froude and Rossby numbers of order unity. It was found that cyclonic vortices will tend to become progressively less unsteady as the Froude number is increased. This means that they become progressively less able to generate gravity waves, as typical frequencies in the vortical dynamics become smaller and smaller with respect to the inertial frequency. The result was that cyclonic vortices exhibited an optimal Froude number, above which the radiated gravity wave pseudoenergy flux actually decreased as the Froude number was further increased. It turned out that even for potential vorticity of nine times the background value, the optimal Froude number occurred at about 0.6, and the pseudoenergy flux did not exceed 10^{-6}.

Anticyclonic vortices appeared to behave somewhat differently at moderate Froude number, with the pseudoenergy flux increasing more rapidly than in the cyclonic cases. However, if the potential vorticity was strictly positive, their nutation times also increased super-exponentially with increasing Froude number, thus inhibiting gravity wave radiation. The maximum pseudoenergy flux from anticyclonic vortices with positive potential vorticity did not exceed 10^{-6}, and the Froude number at which the strongest gravity wave fluxes were observed was almost 1.0.

Vortices with zero and negative potential vorticity were not found to become significantly less unsteady as the Froude number increased. At large Froude numbers, it was found that such vortices tended to generate shock waves, which could then propagate through the fluid. The pseudoenergy flux continues to increase with increasing Froude number, at least as rapidly

as the sixth power of the Froude number, up to Froude number of order unity.

In the moderate Froude number regime, the gravity wave field generated was compared with that predicted by a suitably modified version of the Lighthill theory. In general, there seemed good agreement over a range of Froude numbers, provided the vortices were small compared with a deformation radius.

5.2 Suggestions for further research

This thesis has concentrated on specific features of flows which cannot be balanced, and there remains much work to be done in understanding and quantifying these phenomena.

In chapter 2, after the general theory was introduced, the response of an elliptical patch to radiating gravity waves was examined in some detail. The analysis supported the notion of "generalized adjustment" – i.e. the flow adjusts to a state where its gravity wave emission diminishes. This flow is, however, only one of many flows, and it would be of interest to investigate the nonlinear evolution of other vortical structures analytically, using the matched asymptotic expansions approach. A starting point for the study could be the flows classified by Abrashkin & Yakubovich (1984).

The nonlinear numerical experiments of chapter 4 illustrated a substantial difference between cyclonic and anticyclonic vortex motions, particularly when negative potential vorticity was present in the vortices. This phenomenon does not appear to have been studied in such extreme limits, and it would appear that there is much to be done in gaining a fully quantitative understanding of it.

One significant motivation for the numerical work in chapter 4 was to determine the range of validity of the Lighthill theory of aerodynamic sound generation, as applied to gravity wave generation in a rotating frame. The work presented in this thesis is that simple numerical models, at moderate resolution, can be expected to generate a radiated gravity wave field which is in good quantitative agreement with the wave field predicted by the Lighthill theory.

However, only a limited class of flows has been investigated. One serious omission is that there is no mean jet in the flow. Although not conceptually difficult, there is likely to be some technical difficulty in incorporating a mean jet into a modified version of the Lighthill theory.

The point is that, if there is to be a mean jet in rotating shallow water, the fluid will tend to different uniform heights on either side of the jet, far away from it. The Lighthill theory, as presented here, assumes that the background is homogeneous, and some work will be required to incorporate non-uniformity of the background into the theory. Only when this is done, however, will it be possible to test the Lighthill theory in flows with a mean jet in the quantitative way in which it was tested for flows without a mean jet in chapter 4.

Another development which will be required is the incorporation of vertical structure, so that the gravity waves generated by numerical simulations of baroclinic lifecycles can be analysed within the framework of the Lighthill theory.

Ultimately, we may wish to use the Lighthill theory as a quantitative tool to predict the intensity of gravity wave emission by jet streaks. To achieve this, we must first develop the theory to such a level that it can be applied to flows of practical meteorological interest. This aim, to be achieved through further analytical work with the shallow water and hydrostatic continuously stratified equations, is now of primary importance in furthering this work, and is being actively pursued. When complete, it should provide an excellent tool for analysing the performance of numerical models in which gravity wave generation is observed, and for assessing the significance of gravity waves generated by vortical motions in the global angular momentum budget, whose significance was first identified by Lindzen (1981).

Chapter 6

References

Abramowitz M. and Stegun I.A., 1965: *Handbook of mathematical functions*. Dover, 1046pp.

Abrashkin A.A. and Yakubovich E.I., 1984: Planar rotational flows for an ideal fluid. *Sov. Phys. Dokl.*, **29**, 370-371.

Allen, J. S., Barth, J. A. and Newberger, P. A., 1990: On intermediate models for barotropic continental shelf and slope flow fields. Part I: Formulation and comparison of exact solutions. *J. Phys. Oceanog.*, **20**, 1017–1042.

Allen, J. S., Barth, J. A. and Newberger, P. A., 1990: On intermediate models for barotropic continental shelf and slope flow fields. Part III: Comparison of numerical model solutions in periodic channels. *J. Phys. Oceanog.*, **20**, 1949–1973.

Arakawa A. and Lamb V.R., 1981: A potential enstrophy and energy conserving scheme for the shallow water equations *Mon. Wea. Rev.*, **109**, 18-36.

Asselin R., 1972: Frequency filter for time integrations. *Mon. Wea. Rev.*, **100**, 487-490.

Baer, F. and Tribbia, J. J., 1977: On complete filtering of gravity modes through nonlinear initialization. *Mon. Wea. Rev.*, **105**, 1536-1539.

Barsony-Nagy A., Er-El J. and Yungster S. 1987: Compressible flow past a contour and stationary vortices. *J Fluid Mech.*, **178**, 367-378.

Barth, J. A., Allen, J. S. and Newberger, P. A., 1990: On intermediate models for barotropic continental shelf and slope flow fields. Part II: Comparison of numerical model solutions in doubly periodic domains. *J. Phys. Oceanog.*, **20**, 1044–1076.

Batchelor G.K., 1967: *An introduction to fluid dynamics*, Cambridge University Press, 615pp.

Bourke W., 1972: An efficient, one-level, primitive-equation spectral model. *Mon. Wea. Rev.*, **100**, 683-689.

Bridges J., and Hussain A.K.M.F., 1992: Direct evaluation of aeroacoustic theory in a jet. *J. Fluid Mech.*, **240**, 469-501.

Broadbent E.G. and Moore D.W., 1979: Acoustic destabilisation of vortices. *Phil. Trans. Roy. Soc. Lond.*, **A 290**, 353-371.

Chan W.M., Shariff K. and Pulliam T.H., 1993: Instabilities of two-dimensional inviscid compressible vortices. *J. Fluid Mech.*, **253**, 173-209.

Charney J.G., 1947: The dynamics of long waves in a baroclinic westerly current. *J. Meteor.*, **4**, 135-162.

Charney J.G., 1948: On the scale of atmospheric motions. *Geof. Publ.*, **17, 2**, 17pp.

Charney J.G., 1949: On a physical basis for numerical prediction of large-scale of large-scale motions in the atmosphere. *J. Atmos. Sci.*, **6**, 371-385.

Charney J.G., 1951: Dynamic forecasting by numerical process. In *Compendium of Meteorology*, T. Malone, ed., American Meteorological Society, 1334pp.

Charney, J. G., 1955: The use of the primitive equations of motion in numerical prediction. *Tellus*, **7**, 22–26.

Crighton D.G., 1972a: Radiation from vortex filament motion near a half plane. *J. Fluid Mech.*, **51**, 357-362.

Crighton D.G., 1972b: Radiation properties of the semi-infinite vortex sheet. *Proc. Roy. Soc.*, **A 330**, 185-198.

Crighton D.G., 1975: Basic principles of aerodynamic noise generation. *Prog. Aerospace Sci.*, **16**, 31-96.

Crighton D.G. and Leppington F.G., 1974: Radiation properties of the semi-infinite vortex sheet - the initial value problem. *J. Fluid Mech.*, **64**, 393-414.

Crow S.C., 1970: Aerodynamic sound generation as a singular perturbation problem. *Stud. Appl. Math.*, **49**, 21-44.

Deem G.S. and Zabusky N.J., 1978: Stationaly V-states: interactions recurrence and breaking. *Phys. Rev. Lett.*, **40**, 859.

Dritschel, D. G., 1989: Contour dynamics and contour surgery: numerical algorithms for extended, high-resolution modelling of vortex dynamics in two-dimensional, inviscid, incompressible flows. *Computer Phys. Rep.*, **10**, 77–146.

Dritschel, D. G., 1993: A fast contour dynamics method for many-vortex calculations in two-dimensional flows. *Phys. Fluids*, **A 5**, 173-186.

Dritschel D.G. and Legras B., 1991: The elliptical model of two-dimensional vortex dynamics. II: Disturbance equations. *Phys. Fluids*, **A 3**, 855-869.

Dritschel, D.G. and Saravanan, R., 1993: Three-dimensional quasi-geostrophic contour dynamics, with an application to stratospheric vortex dynamics. *Q. J. Roy. Met. Soc.*, submitted.

Errico, R. M., 1982: Normal mode initialization and the generation of gravity waves by quasi-geostrophic forcing. *J. Atmos. Sci.*, **39**, 573–586.

Ertel, H., 1942: Ein Neuer hydrodynamischer Wirbelsatz. *Met. Z.*, **59**, 271–281.

Ffowcs-Williams J.E., 1969: Hydrodynamic noise. *Ann. Rev. Fluid Mech.*, **1**, 197-222.

Gent, P.R. and McWilliams, J.C., 1983a: Consistent balanced models in bounded and periodic domains. *Dyn. Atmos. Ocean.*, **7**, 67-93.

Gent, P.R., and McWilliams, J.C., 1983b: The equatorial waves balanced models. *J. Phys. Oceanog.*, **13**, 1179-1192.

Gill, A.E., 1982: *Atmosphere-Ocean Dynamics*, Academic Press, 602pp.

Griffiths R.W., Killworth P.D. and Stern M.E., 1982: Ageostrophic instability of ocean currents. *J. Fluid Mech.*, **117**, 343-377.

Gryanik V.M., 1983: Emission of sound by linear vortical filaments. *Izv. Atm. Oc. Phys.*, **19**, 150-152.

Haynes P.H., 1988: Forced, dissipative generalizations of finite-amplitude wave-activity conservation relations for zonal and non-zonal basic flows. *J. Atmos. Sci.*, **45**, 2352-2362.

Haynes P.H., 1989: The effect of barotropic instability on the nonlinear evolution of a Rossby wave critical layer. *J. Fluid Mech.*, **207**, 231-266.

Ho C.-M. and Huerre P., 1984: Perturbed free shear layers. *Ann. Rev. Fluid Mech.*, **16**, 365-424.

Howe M.S., 1975: Contributions to the theory of aerodynamic sound, with application to excess jet noise and the theory of the flute. *J. Fluid Mech.*, **71**, 625-673.

Hoskins B.J., 1974: The role of potential vorticity in symmetric stability and instability. *Q. J. Roy. Met. Soc.*, **100**, 480-482.

Hoskins, B. J., McIntyre, M. E. and Robertson, A. W., 1985: On the use and significance of isentropic potential-vorticity maps. *Q. J. Roy. Meteorol. Soc.*, **111**, 877–946.

Hayashi Y. and Young W.R., 1987: Stable and unstable shear modes on rotating parallel flows in shallow water. *J. Fluid Mech.*, **184**, 477-504.

Jacobs, S. J., 1991: Existence of a slow manifold in a model system of equations. *J. Atmos. Sci.*, **48**, 893–901.

Kambe T., 1986: Acoustic emissions by vortex motions. *J. Fluid Mech.*, **173**, 643-666.

Kambe T. and Minota T., 1981: Sound radiation from vortex systems. *J. Sound Vib.*, **74**, 61-72.

Kambe T. and Minota T., 1983: Acoustic wave radiated by head-on collision of two vortex rings. *Proc. Roy. Soc.*, **A 386**, 277-308.

Lord Kelvin, 1880: On the oscillations of a columnar vortex. *Phil. Mag.*, **10**, 155-168.

Kida S., 1981: Motion of an elliptic vortex in a uniform shear flow. *J. Phys. Soc. Japan*, **50**, 3517-3520

Killworth P.D. and Stern M.E., 1982: Instabilities on density-driven boundary currents and fronts. *Geophys. Astrophys. Fluid Dyn.*, **22**, 1-28.

Kimura Y., 1989: Transport properties of waves on a vortex filament. *Physica D*, **37**, 485-489.

Kleinschmidt, E., 1950: Über Aufbau und Entstehung von Zyklonen (1 Teil). *Met. Runds.*, **3**, 1–6.

Kleinschmidt, E., 1950: Über Aufbau und Entstehung von Zyklonen (2 Teil). *Met. Runds.*, **3**, 54–61.

Kleinschmidt, E., 1951: Über Aufbau und Entstehung von Zyklonen (3 Teil). *Met. Runds.*, **4**, 89–96.

Kleinschmidt, E., 1955: Die Entstehung einer Hohenzyklone über Nordamerika. *Tellus*, **7**, 96–110.

Kleinschmidt, E., 1957: In "dynamic meteorology", by Eliassen, A., and Kleinschmidt, E., Handbuck der Physik, 48, 1-154.

Klyatskin V.I., 1966: Sound radiation by a system of vortices. *Fluid Dyn.*, **1 6**, 54-57.

Knessl, C. and Keller, J. B., 1992: Stability of rotating shear flows in shallow water. *J. Fluid Mech.*, **244**, 605–614.

Koch S.E. and Dorian P.B., 1988: A mesoscale gravity wave event observed during CCOPE. Part III: Wave environment and possible source mechanisms. *Mon. Wea. Rev.*, **116**, 2570-2592.

Kop'ev V.F. and Leont'ev E.A., 1983: Acoustic instability of an axial vortex. *Sov. Phys.*

Acoustics, **29**, 111-115.

Kop'ev V.F. and Leont'ev E.A., 1985: Energy aspects of the acoustic instability of certain steady-state vortices. *Sov. Phys. Acoustics*, **31**, 205-207.

Kop'ev V.F. and Leont'ev E.A., 1986: Some remarks on Lighthill's theory in connection with sound radiation by compact vortices. *Sov. Phys. Acoustics*, **32**, 109-112.

Kubokawa A., 1986: Instability caused by the coalescence of two modes of a one-layer coastal current with a surface front. *J. Oceanogr. Soc. Japan*, **42**, 373-380.

Lamb, H.H., 1932: *Hydrodynamics*, Dover, New York.

Lauvstad, 1968: On non-uniform Mach number expansion of the Navier-Stokes equations and its relation to aerodynamically generated sound. *J. Sound Vib.*, **7**, 90-105.

Leith, C.E., 1980: Non-linear normal mode initialization and quasi-geostrophic theory. *J. Atmos Sci.*, **37**, 958-968.

Legras B. and Dritschel D.G., 1991: The elliptical model of two-dimensional vortex dynamics. I: The basic state. *Phys. Fluids*, **A 3**, 845-854.

Legras B. and Zeitlin V., 1992: Conformal dynamics for vortex motions. *Phys. Lett.*, **A 167**, 265-271.

Lele S.K. and Ho C.-M., 1993: Acoustic radiation from temporally evolving free shear layers. *J. Fluid Mech.*, submitted.

Lighthill M.J., 1952: On sound generated aerodynamically, I. General theory. *Proc. Roy. Soc.*, **A 211**, 564-587.

Lighthill M.J., 1954: On sound generated aerodynamically, II. Turbulence as a source of sound. *Proc. Roy. Soc. Lond.*, **A 222**, 1-32.

Lighthill M.J., 1958: *An introduction to Fourier analysis and generalized functions.* Cambridge

University Press, 79pp.

Lindzen, R. S., 1981: Turbulence and stress owing to gravity wave and tidal breakdown. *J. Geophys. Res.*, **86**, 9707–9714.

Lorenz, E.N., 1960: Energy and numerical weather prediction. *Tellus*, **12**, 364-373.

Lorenz, E. N., 1986: On the existence of a slow manifold. *J. Atmos. Sci.*, **43**, 1547–1557.

Lorenz, E. N., 1992: The slow manifold – what is it? *J. Atmos. Sci.*, **49**, 2449–2451.

Lorenz, E. N., Krishnamurthy, V., 1987: On the nonexistence of a slow manifold. *J. Atmos. Sci.*, **44**, 2940–2950.

Love, A.E.H., 1893: On the stability of certain vortex motions *Proc. Lond. Math. Soc.*, **25**, 18-34.

Lush P.A., 1971: Measurements of subsonic jet noise and comparison with theory. *J. Fluid Mech.*, **46**, 477-500.

Machenhauer, B., 1977: On the dynamics of gravity oscillations in a shallow water equation model with applications to normal mode initialization. *Contrib. Atmos. Phys.*, **50**, 253-271.

Mankbadi R. and Liu J. T. C., 1984: Sound generated aerodynamically revisited: large-scale structures in a turbulent jet as a source of sound. *Phil. Trans. Roy. Soc. Lond.*, **A 311**, 183-217.

McIntyre, M.E. and Norton, W.A., 1993: Potential vorticity inversion on a hemisphere. *J. Atmos. Sci.*, submitted.

McIntyre M.E. and Shepherd T.G., 1987: An exact local conservation theorem for finite amplitude disturbances to non-parallel shear flows, with remarks on Hamiltonian structure and Arnol'd's stability theorems. *J. Fluid Mech.*, **181**, 527-565.

McWilliams, J.C., 1985: A uniformly valid model spanning the regimes of geostrophic and

isotropic. stratified turbulence: balanced turbulence. *J. Atmos. Sci.*, **42**, 1773-1774.

McWilliams, J.C., Gent, P.R. and Norton, N.J., 1986: The evolution of balanced, low-mode vortices on the β-plane. *J. Phys. Oceanog.*, **16**, 838-855.

Margules M., 1904: Uber die Beziehung zwischen Barometerschwankungen und Kontinuitatsgleichung. *Festschrift Ludwig Boltzmann* pp. 585-589, Leipzig, J. A. Barth, 930pp.

Minota T. and Kambe T., 1986: Observations of acoustic emission from head-on collision of two vortex rings. *J. Sound Vib.*, **111**, 51-59.

Moore C.J., 1977: The role of shear-layer instability waves in jet exhaust noise. *J. Fluid Mech.*, **80**, 321-367.

Moore D.W. and Pullin D.I., 1987: The compressible vortex pair. *J. Fluid Mech.*, **185**, 171-204.

Morse P.M. and Feshbach H., 1953: *Methods of mathematical physics* II, McGraw - Hill, 997pp.

Norton, W.A., 1988: *Balance and potential vorticity inversion in atmospheric dynamics.* Ph.D. thesis, University of Cambridge. 167pp.

Norton, N.J., McWilliams, J.C. and Gent, P.R., 1986: A numerical model of the balance equations in a periodic domain and an example of balanced turbulence. *J. Comp. Phys.*, **67**, 439-471.

Obermeier F., 1967: Berechnung aerodynamisch erzeugter Schallfelder mittels der Methoder "Matched Asymptotic Expansions". *Acustica*, **18**, 238-240.

Paldor N., 1983: Linear stability and stable modes of geostrophic modes. *Geophys. Astrophys. Fluid Dyn.*, **24**, 299-326.

Platzman G.W., 1967: A retrospective view of Richardson's book on weather prediction. *Bull. Amer. Met. Soc.*, **48**, 514-550.

Platzman G.W., 1968: Richardson's weather prediction. *Bull. Amer. Met. Soc.*, **49**, 496-500.

Polvani L.M., Zabusky N.J. and Flierl G.R., 1989: Two-layer geostrophic vortex dynamics. Part I. Upper-layer V-states and merger. *J. Fluid Mech.*, **205**, 215-242.

Rahman S., 1971: Berechnung der Schallerzeugung beim frontalen Zusammenstoss zweier Wirbelpaare. *Acustica*, **24**, 50-54.

Richardson L.F., 1922: *Weather prediction by numerical processes.* Cambridge University Press, 236pp.

Ripa P., 1983: General stability conditions for zonal flows in a one-layer model on the beta-plane or the sphere. *J. Fluid Mech.*, **126**, 463-489.

Ripa P., 1987: On the stability of elliptical vortex solutions of the shallow water equations. *J. Fluid Mech.*, **183**, 343-363.

Pullin, D.I., 1992: Contour Dynamics Methods. *Ann. Rev. Fluid Mech.*, **24**, 89-115.

Robert A.J., 1966: The integration of a low order spectral form of the primitive meteorological equations. *J. Meteor. Soc. Japan*, 44, 237-245.

Robinson, A.R. and Stommel, H.M., 1959: The oceanic thermocline and the accociated thermocline circulation. *Tellus*, **3**, 295-308.

Rossby C.G., 1936: Dynamics of steady ocean currents in the light of experimental fluid mechanics. *Papers in physical oceanography and meteorology*, **5 1**, 43pp.

Rossby C.G. et al., 1939: Relation between variations in the intensity of the zonal circulation of the atmosphere and the displacements of the semi-permanent centers of action. *J. Mar. Res.*, **2**, 38-55.

Sakai S., 1989: Rossby-Kelvin instability: a new type of ageostrophic instability caused by resonance between Rossby waves and gravity waves. *J. Fluid Mech.*, **202**, 149-176.

Salmon, R., 1985: New equations for nearly geostrophic flow. *J. Fluid Mech.*, **153**, 461–477.

Salmon, R., 1988a: Semigeostrophic theory as a Dirac-bracket projection. *J. Fluid Mech.*, **196**, 345–358.

Salmon, R., 1988b: Hamiltonian fluid mechanics. *Ann. Rev. Fluid Mech.*, **20**, 225–256.

Satomura T., 1981: An investigation of shear instability in a shallow water. *J. Met. Soc. Japan*, **59**, 148-167.

Sozou C., 1987: New solutions representing adiabatic transverse waves in a Rankine vortex. *Proc. Roy. Soc. Lond.*, **A 413**, 225-234.

Spall, M.A. and McWilliams, J.C., 1992: Rotational and gravitational influences on the degree of balance in the shallow water equations. *Geophys. Astrophys. Fluid Dyn.*, **64**, 1-29.

Stuber B., 1970: Schallabstrahlung und Korperschallanregung durch Wirbel. *Acustica*, **23**, 82-92.

Tribbia, J. J., 1984: A simple scheme for high-order nonlinear normal mode initialization. *Mon. Wea. Rev.*, **112**, 278–284.

Vallis, G. K., 1992: Mechanisms and parameterizations of geostrophic adjustment and a variational approach to balanced flow. *J. Atmos. Sci.*, **49**, 1144–1160.

Vautard, R. and Legras, B., 1986: Invariant manifolds, quasi-geostrophy and initialization. *J. Atmos. Sci.*, **43**, 565–584.

Warn, T. and Ménard, R., 1986: Nonlinear balance and gravity-inertial wave saturation in a simple atmosperic model. *Tellus*, **38A**, 285–294.

Webster, R. B., 1970: Jet noise simulation on shallow water. *J. Fluid Mech.*, **40**, 423–432.

Williams J. S., 1992: *Nonlinear problems in vortex sound.* PhD Thesis, University of Leeds.

Zeitlin V., 1988: Acoustic radiation from distributed vortex structures. *Sov. Phys. Acoustics*, **34**, 188-191.

Zietlin V., 1991: On the backreaction of acoustic radiation for two-dimensional vortex strucures. *Phys. Fluids*, **A 3**, 1677-1680.

Experimental Research of the Magnetic-Gravity Effects.
Full Size SEG tests.

V. V. Roschin, E-mail: rochtchin@mail.ru
S. M. Godin, E-mail: serjio@glasnet.ru

Institute for High Temperatures, Russian Academy of Science, Izhorskaya 13/19, Moscow 127412, Russia

Abstract

In the present paper the results of the experimental research of magnetic-gravity effects are presented. Anomalous magnetic and thermal changes within a radius of 15 meters from the researched device were measured as well.

PACS: 41.20.-q; 44.60.+k; 76.50.+q

1. Introduction

There has been a great interest in examining non-linear effects in the system of rotating magnetic fields. Such effects have been observed in the device called Searl's generator or Searl Effect Generator (SEG) [1-4]. A SEG consists of a series of three concentric rings and rollers that circulate around the rings. All parts of SEG are based on the Law of the Squares [5]. The rollers revolve around the concentric rings, but they do not touch them. There's a primary north-pole and primary south pole on the rollers and a primary north-pole and primary south-pole on the concentric rings. Obviously, the north-pole of the roller is attracted to the south-pole of the concentric rings and vice versa.

The rollers have a layered structure similar to the concentric rings. The external layer is titanium, then iron, nylon and the last internal layer was made from neodymium. John R.R. Searl has supposed that electrons are given off from the central element (neodymium) and travel out through other elements. Dr. Searl contends that if nylon had not been used, the SEG would act like a laser and one pulse would go out and it would stop, build up, and another pulse would go out. The nylon acts as a control gate that yields an even flow of electrons throughout the SEG [4].

In [4] it was shown that in the process of magnetization of the plate and rollers, the combination of constant and variable magnetic fields for creating a special wave (sine wave) pattern on a plate surface and rollers surface was used. The basic effects consist of the rollers self-running around the ring plate with a concurrent reduction of weight and an increasing occurrence of propulsion. These effects come about because of a special geometry of experimental setup. It was shown that the operation of the device in the critical regime is accompanied by biological and real physical phenomena.

Other information where similar effects are be mentioned can be found in the books, *Unconventional Flying Objects* [6] and the *Homopolar Handbook* [7] which includes papers on magnetized dielectrics. In this paper we present the experimental device the results we have obtained.

2. Description of the Experimental Installation

The basic difficulty arises in choosing the materials and maintaining the necessary pattern imprinting on the plate and roller surfaces. To simplify the technology we decided to use a one-ring design with one-ring plate (stator) and one set of rollers (rotor). It is obvious, that it was necessary to strengthen the roller rotor near the bearings and balance the rollers well. In the suggested design, air bearings were used which provided the minimum losses due to friction.

From the available description [1-4] it was not clear how to build and magnetize a stator with a one-meter diameter. In order to make the stator, separate magnetized segments of rare earth magnets with a residual induction of 1T were used. The segments were magnetized in a usual way by discharging a

capacitor-battery system through a coil. Afterwards, the segments were assembled and glued together in a special iron armature, which reduced magnetic energy. To manufacture the stator, 110 kg of neodymium magnets were used and 115 kg of neodymium were used to manufacture the rotor. High-frequency field magnetization was not applied. It was decided to replace an imprinting technology described in [1-5] with cross-magnetic inserts having a flux vector directed at 90 degrees to the primary magnetization vector of the stator and rollers.

For the cross inserts, modified rare earth magnets with a residual magnetization of 1,2 T and coercive force a little bit greater than in a base material were used. In Fig.1 and Fig.2 the joint arrangement of stator 1 and rotor, made up of rollers 2, and a way of their mutual gearing or sprocketing by means of cross magnetic inserts 19, are shown. Between the stator and roller surfaces the air gap δ of 1-mm is maintained.

No layered structure was used except a continuous copper foil of 0.8 mm thickness, which wrapped up the stator and rollers. This foil has direct electrical contact to magnets of the stator and rollers. Distance between inserts in the rollers is equal to distance between inserts on the stator. In other words, $\mathbf{t_1} = \mathbf{t_2}$ in Fig.2.

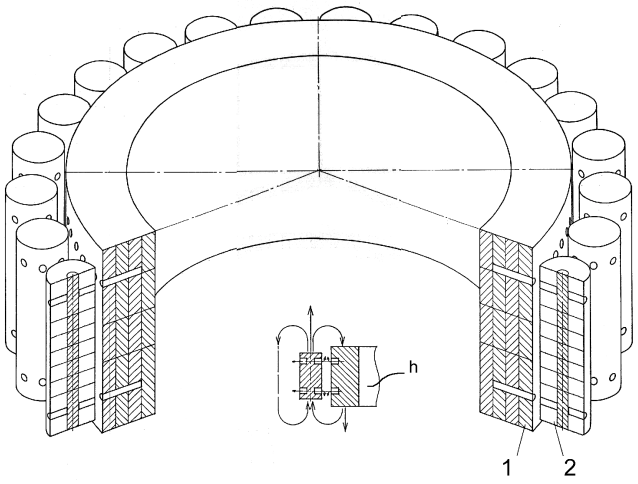


Fig.1. Variant of one-ring converter.

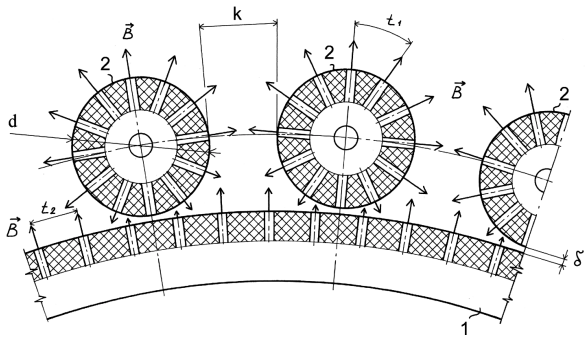


Fig.2. Sprocket effect of magnetic stator and roller inserts.

The ratio of parameters of the stator 1 and the rotor 2 in Fig.2 is chosen so that the relation of stator diameter D to the roller diameter d is an integer equal to or greater than 12. Choosing such ratio allowed us to achieve a "magnetic spin wave resonant mode" between elements of a working body of the device since the circumferences also maintained the same integer ratio.

The elements of magnetic system were assembled in a uniform design on an aluminum platform. In Fig. 3 the general view of the platform with the one-ring converter is displayed. This platform was supplied with springs and shock absorbers with limited ability to move vertically on three supports. The system has a maximum value of displacement of about 10 mm and was measured by the induction displacement meter, 14. Thus, the instantaneous change of the platform weight was defined during the experiment in real time. Gross weight of the platform with magnetic system in the initial condition was 350 kg.

The stator, 1, was mounted motionlessly, and the rollers, 2, were assembled on a mobile common separator, 3, also regarded as the rotor, connected with the basic shaft, 4, of the device. The rotary

moment was transferred through this shaft. The base of the shaft was connected through a friction clutch, 5, to a starting motor, 6, which accelerated the converter up to a mode of self-sustained rotation. The electrodynamics generator, 7, was connected to the basic shaft as a main loading of the converter. Adjacent to the rotor, electromagnetic inductors, 8, with open cores, 9, were located.

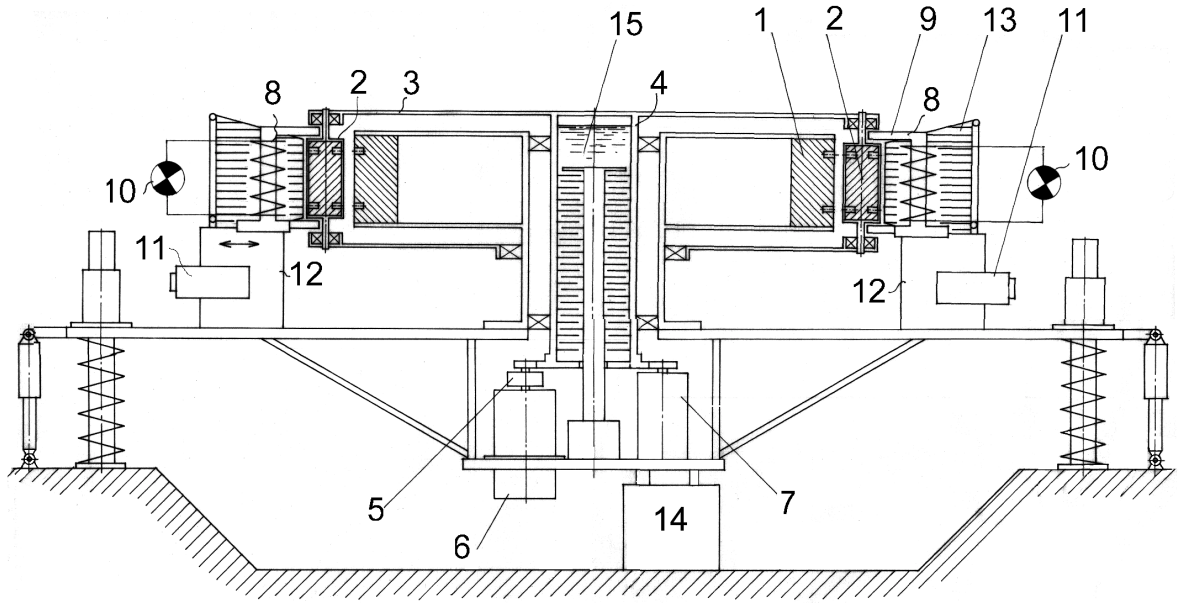


Fig.3. The general view of the one-ring converter and platform.

The magnetic rollers, 2, crossed the open cores of inductors and closed the magnetic flux circuit through electromagnetic inductors, 8, inducing an electromotive force emf in them, which acted directly on an active load, 10 (a set of inductive coils and incandescent lamps with a total power load of 1 kW). The electromagnetic inductor coils, 8, were equipped with an electrical drive, 11, on supports, 12. Driven coils for smooth stabilization of the rotor's rpm were used but the speed of the rotor could be adjusted by changing the main loading, 10.

To study the influence of high voltage on the characteristics of the converter, a system for radial electrical polarization was mounted. On the periphery of the rotor ring, electrodes, 13, were set between the electromagnetic inductors, 8, having an air gap of 10 mm with the rollers, 2. The electrodes are connected to a high-voltage source; the positive potential was connected to the stator, and the negative to the polarization electrodes. The polarizing voltage was adjusted in a range of 0-20 kV. In the experiments, a constant value of 20 kV was used.

In case of emergency braking, a friction disk from the ordinary car braking system was mounted on a basic shaft of the rotor. The electrodynamics generator, 7, was connected to an ordinary passive resistive load through a set of switches ensuring step connection of the load from 1 kW to 10 kW through a set of ten ordinary electric water heaters.

The converter undergoing testing had in its inner core the oil friction generator of thermal energy, 15, intended for tapping a superfluous power (more than 10 kW) into the thermo-exchange contour. But since the real output power of the converter in experiment has not exceeded 7 kW, the oil friction thermal generator was not used. The electromagnetic inductors were connected to an additional load,

which was set of incandescent lamps with total power 1 kW and facilitated complete stabilization of the rotor revolutions.

3. Experimental results

The magnetic-gravity converter was built in a laboratory room on three concrete supports at a ground level. The ceiling height the lab room was 3-meters, the common working area of the laboratory was about 100 sq. meters. Besides the presence of the iron-concrete ceiling, in the immediate proximity from the magnetic system there was a generator and electric motor, which contained some tens of kilograms of iron and could potentially deform the field's pattern.

The device was initially started by the electric motor that accelerated the rotation of the rotor. The revolutions were smoothly increased up to the moment the ammeter included in a circuit of the electric motor started to show zero or a negative value of consumed current. The negative value indicated a presence of back current. This back current was detected at approximately 550 rpm. The displacement meter, 14, starts to detect the change in weight of the whole installation at 200 rpm. Afterwards, the electric motor is completely disconnected by the electromagnetic clutch and the ordinary electrodynamics generator is connected to the switchable resistive load. The rotor of the converter continues to self-accelerate and approach the critical mode of 550 rpm where the weight of the device quickly changes.

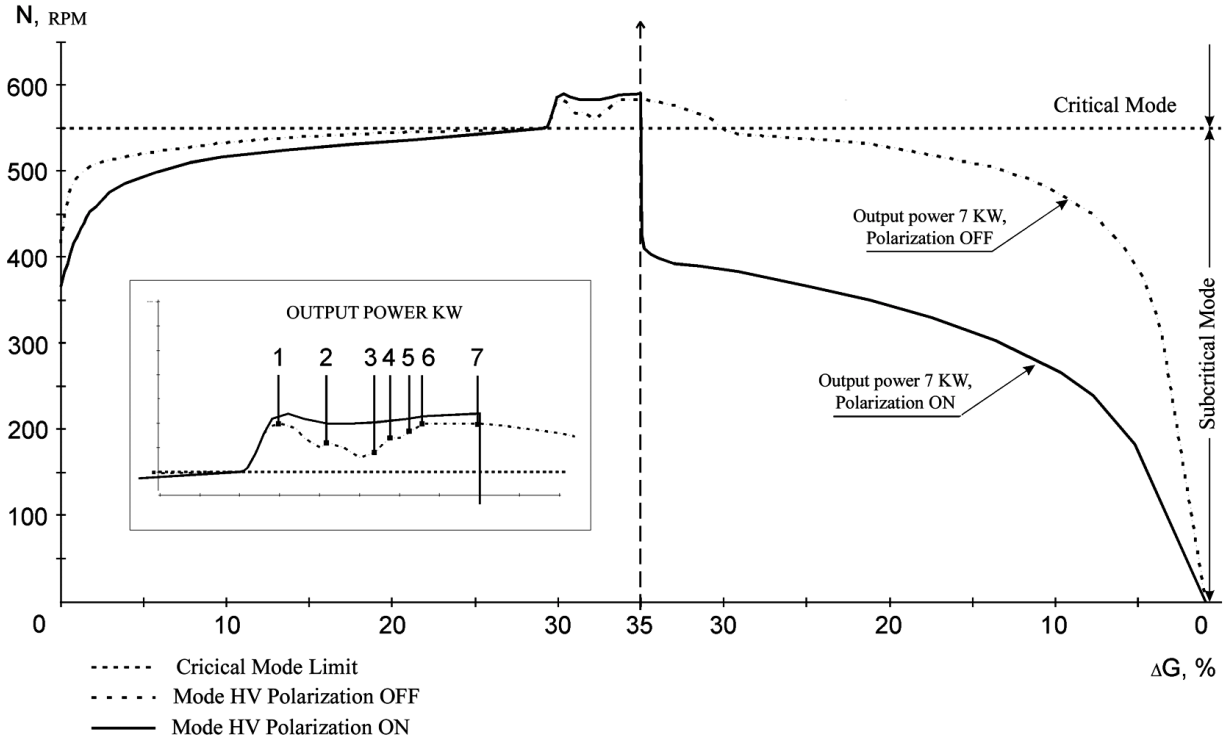


Fig. 4. -G, +G changes in weight of the platform vs. rpm

In addition to the dependence on the speed of rotation, the weight differential depends on the generated power through the load and on the applied polarizing voltage as well. As seen in Fig.4, at the maximum output power equal to 6-7 kW, the change of weight $\Delta\mathbf{G}$ of the whole platform (total weight is about 350 kg), reaches 35 % of the weight in an initial condition $\mathbf{G_i}$. Applying a load of more than 7 kW results in a gradual decrease in rotation speed and an exit from the mode of self-generation, with the rotor coming to a complete stop subsequently.

The net weight $\mathbf{G_n}$ of the platform can be controlled by applying high voltage to polarization ring electrodes located at a distance of 10 mm from external surfaces of the rollers. Under the high 20 kV voltage (electrodes having negative polarity) the increase of tapped power of the basic generator to more than 6 kW does not influence $\Delta\mathbf{G}$ if the rotation speed is kept above 400 rpm. "Tightening" of this effect is observed as well as the effect of hysteresis on $\Delta\mathbf{G}$ (a kind of "residual induction"). The experimental diagrams given on Fig.4 illustrate the +G and –G modes of the converter operations vs. rotor rpm.

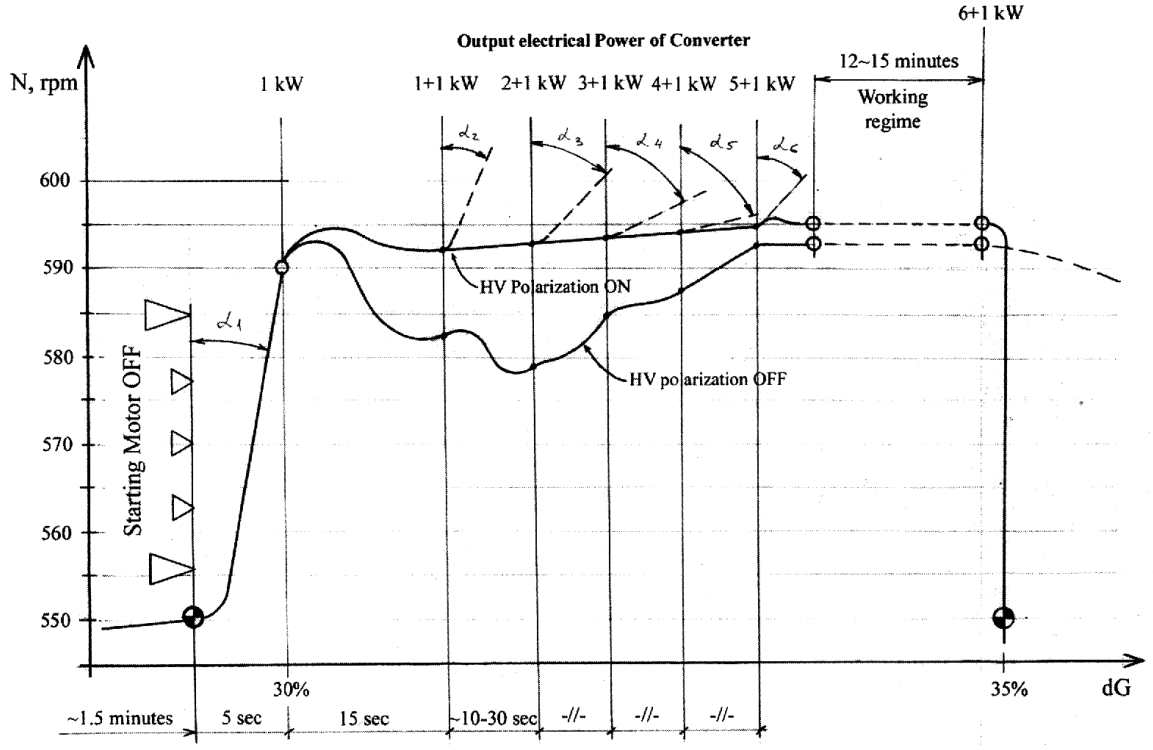


Fig.5 Diagrams of a rotor accelerating and loading of the converter.

The effect of a local change of the platform weight is reversible, relative to the direction of rotor turning, and has the same hysteresis. A clockwise rotation causes the critical mode to occur in the area of 550 rpm and the propulsion force against the direction of gravitation vector is created. Correspondingly, a counter-clockwise rotation causes the critical mode to occur the in area of 600 rpm and a force in the direction of gravitation vector is created. The difference in approach to a critical mode of 50 - 60 rpm was observed. It is necessary to mention that the most interesting region are situated above the critical area of 550 rpm, but due to of a number of circumstances the implementation of such research was not possible. It is necessary to note, that probably there are also other resonant modes appropriate to higher rpm of a rotor and to the significant levels of useful loading and weight changing. Proceeding from the theoretical assumptions, the dependence of tapped mechanical energy from the parameters of magnetic system of the converter and rpm of a rotor has a nonlinear character and the received effects are not optimum. From this point of view, the revealing of the maximal output power, of the maximal change of weight and resource of the converter represents the large practical and scientific interest. In tested sample of the converter the using of higher rpm was inadmissible because of unsufficient mechanical durability of the magnetic system, which was built from separate pieces.

In Fig.5 the dependence of weight of a platform and its output power, removed into resistive loading from rpm of a rotor of the converter is explained in detail. The diagrams are constructed for a case of HV polarization ON (top diagram) and HV polarization OFF (bottom diagram). The time from the start moment of the engine up to a mode of self-generation of the converter, at the rotation of a rotor clockwise, approximately is equal 1.5 minutes. (The power of starting engine was about of 2 KW with a reduction on the shaft of the converter is equal 1/10). At achievement of a critical mode (550 rpm.), the change of gross weight of a platform already achieves +/-30% from $\mathbf{G_i}$. In the point of transition to a resonant mode, the revolutions with the large acceleration has increases up to 590 rpm and weight has changes up to + /-35% from $\mathbf{G_i}$. In this time an unpleasant high frequency whistling sound became audible. This place of the diagram begins at once after a critical point (inclination of a curve $\boldsymbol{\alpha_1}$). At achievement of 590 rpm, the first stage of resistive loading in 1 kW is connected to the electrodynamics generator. The whistling sound at once stops, the revolutions are sharply reduced and ΔG also is changing. As soon as the revolutions begin to grow again, the second switchable load is connected and rotor's rpm are stabilized at a level of 590-595 rpm. ΔG continues to change. The increasing of switchable loading occurs by steps in 1 kW up to total power of 6 kW. All intervals in time are equal approximately of 10-30 sec. Afterwards, the short-term increasing of revolutions and then the full stabilization of a mode during of 12-15 min were observed.

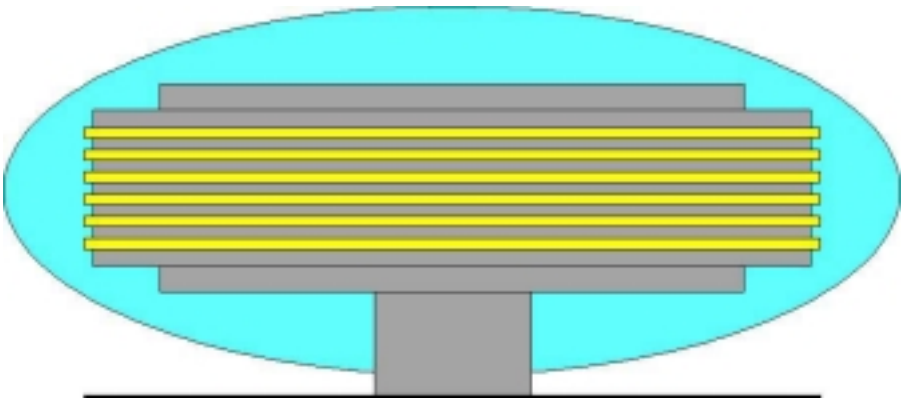
More than 50 launches of this converter with absolute repeatability within three months were carried out. It is necessary to note, that revolutions will grow with acceleration reflected on the Fig.5 by corners $\boldsymbol{\alpha_1}\ldots\boldsymbol{\alpha_5}$, if the generator is not switched to the next step of loading, the rpm will continue increasing. Twice as much loading was required to return to a previous rpm mode.

The above discussion concerns a mode with high voltage polarization of 20 kV "plus" on a "grounded" stator. Without the 20 kV polarization voltage (lower curve) the diagram is approximately the same, but indicates the more hard character of loading and faster change of weight of a platform due to decreasing of rpm.

Other interesting effects include the work of the converter in a dark room when corona discharges are observed around the converter's rotor as a blue-pink glowing luminescence and a characteristic ozone smell. In Fig.6 the cloud of ionization covers the area of a stator and a rotor and has accordingly a toroidal form.

Fig.6. Corona discharges around the converter.

On the background of luminescence glowing on rollers' surfaces, we distinguished a «wave picture.» A number of more vigorous strips of discharges around the rollers were observed. These discharges were of a white-yellow color but the characteristic sound for arc discharges was not audible. На поверхностях статора и роликов не имелось никаких видимых эрозионных повреждений as well as. Erosive damage induced by arc discharges were not present on any surfaces of the stator or the rollers .



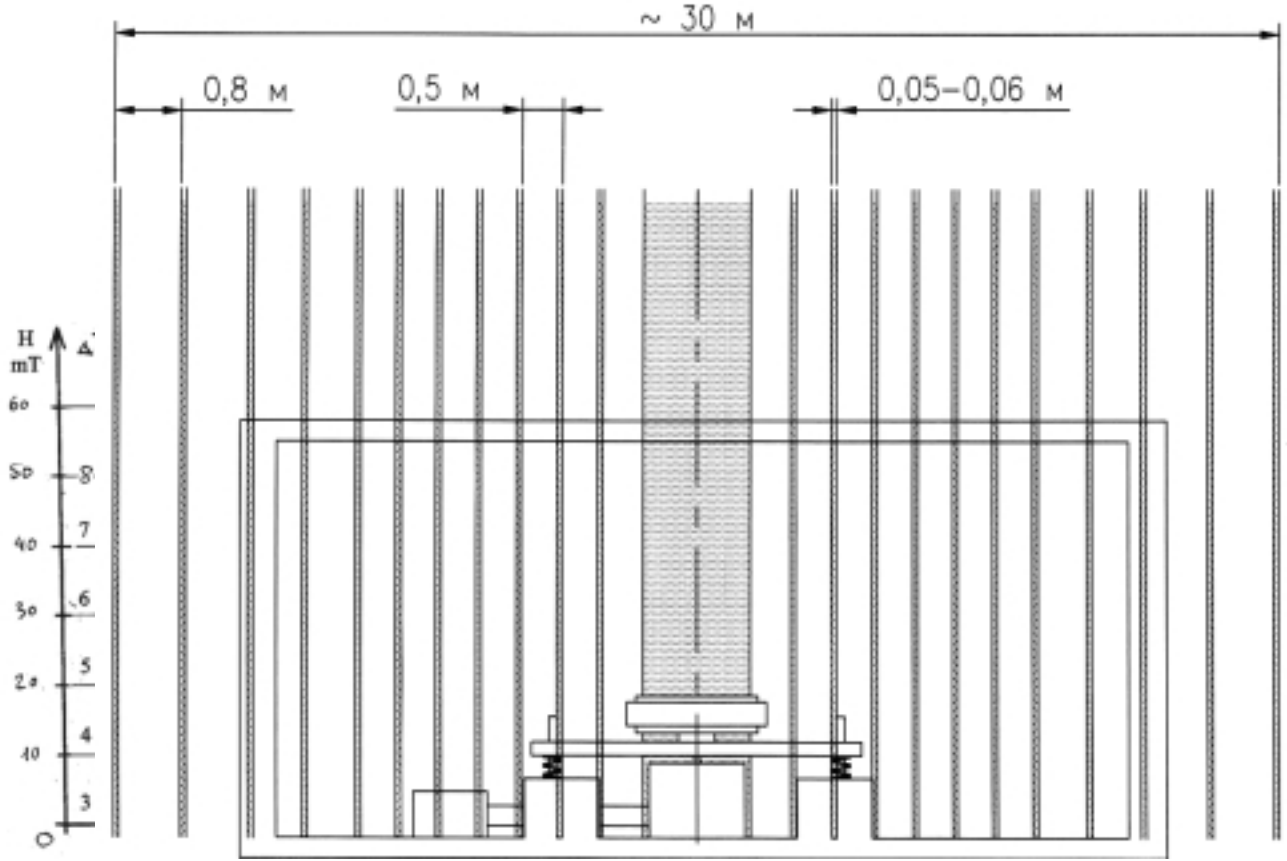
One more effect previously not mentioned was observed i.e. the vertical concentric magnetic "walls" around the installation. We noticed and measured the abnormal permanent magnetic field around the converter within the radius of 15 me-

ters. For the magnetic field measurement the Russian magnetometer F4354/1 was used. Magnetometer used the Hall-effect sensor in the copper shielding. The zones of an increased intensity of a magnetic flux 0.05 T located concentrically from the center of the installation were detected. The direction of the magnetic field vector in these walls coincided with the direction of the rollers' field vector. The structure of these zones reminded us of circles on water from a thrown stone. Between these zones, this portable magnetometer did not register abnormal magnetic fields. The layers of an increased intensity are distributed practically without losses up to a distance of about 15 meters from the center of the converter and quickly decrease at the border of this zone. The thickness of each layer is about 5 - 6 cm. The border of each layer has sharp shape, the distance between layers is about 50 - 80 cm where the upper limit is seen when moving from the center of the converter. A stable picture of this field was observed as well as at a height of 6 m above the installation (on the second floor above the lab). Above the second floor, measurements were not carried out. The similar picture was observed and outside of a room of laboratory, directly in the street, on the ground. The concentric walls are strictly vertical and no had appreciable distortions. In Fig.7 the schematic arrangement of the converter in a room of laboratory and arrangement of concentric magnetic fields are shown.

Fig.7. Schematic placing of the converter and pattern of magnetic fields in the lab's room.

An anomalous decrease temperature in the vicinity of the converter was also found. While the common temperature background in laboratory was + 22°C (±2°C) a fall of temperature equal to 6-8°C was noticed. The same phenomenon was observed in concentric vertical magnetic walls as well as. The measurements of temperature inside the magnetic walls were carried out by an ordinary alcohol thermometer with an inertia of indication about 1.5 min. In the magnetic walls the temperature changes can even be distinctly observed by hand. When placed into this magnetic wall the hand feels very cold at once. A similar picture was observed at the height above installation, i.e. on the second floor of the laboratory as well as despite the ferro-concrete blocking of the ceiling and also on open air outside of the laboratory.

Fig.8. Dependence of intensity of a magnetic



field and changes of temperature vs rotor's rpm of the converter.

Concentric magnetic walls and the accompanying thermal effects begin to show themselves by an appreciable image beginning approximately from of 200 rpm., and linearly grow with increasing of revolutions up to a critical mode. The measurements above 600 rpm were not made because of fear of destruction of magnetic system. In Fig.8 the curve of intensity of a magnetic field in мT and change of temperature in Celsius degrees due to rpm changing is represented.

4. Discussion

All the results we obtained are extremely unusual and require some theoretical explanation. Unfortunately, the interpretation of results within the framework of the conventional physical theory cannot explain all the observed phenomena besides the change of weight. It is possible to interpret the change of weight either as a local change of gravitational force or as an antigravity force repelling its own field. Direct experiment, confirming the presence of a draft force was not performed, but in any case both interpretations of the weight change do not correspond to the modern physics paradigm. A reconsideration of the standard theory of gravitation is possible if we take into consideration space-time curvature. For example, the Kerr metric usually represents the field exterior to an axially symmetric rotating body and distinguishes between positive and negative spin directions as well as forward and backward time directions [8]. An examination of the physical vacuum as a source of these phenomena may also lend itself to better interpretation since the Maxwell stress-energy tensor in the vicinity of the converter undergoes a complex evolution.

From the modern physics position, electrification and luminescence of the converter's magnetic system in the near zone is not completely clear. The phenomenon of the magnetic and thermal "walls" may be connected with Alphen's magnetic-sound waves raised in near zone in magnetized plasma induced by a variable magnetic field of a rotating rotor [9]. The energy exchange between ambient air molecules and the converter may be occurring. At the present time we can not give an exact description of the interactions mechanism and transformation of energy, but without a relativistic we are completely unable to give a physically substantial theory of these phenomena.

In conclusion, we emphasize that issues of the biological influence effects and especially of the variations of real time stream effects, which must be taking place in an operative zone of the converter, were not considered at all. These issues are extremely important and absolutely unexplored; though there are some mentions of J.R.R.Searl about healing action of the SEG's radiation. Our own experience allows us to make only cautious assumption that the short-term stay (dozen minutes) in a working zone of the converter with the fixed output power of 6 kW remains without observed consequences for those exposed. The present paper is only a beginning.

References

1. Schneider, Koeppl, & Ehlers. «Begegnung mit John R.R. Searl.» *Raum und Ziet*, #39, 1989, pp. 75-80.
2. Sandberg, Von S. Gunnar. «Was ist dran am Searl-effect.» *Raum und Ziet*, #40, 1989, pp. 67-75.
3. Schneider & Watt. «Dem Searl-effect auf der spur.» *Raum und Ziet*, # 42, 1989, pp.75-81; #43, pp.73-77.
4. Thomas Jr., John A. «ANTI-GRAVITY: The Dream Made Reality.» *Extraordinary Science*, vol.VI, Issue 2, 1994. (Also *ANTIGRAVITY: The Dream Made Reality*, Direct International Science Consortium, 1993, ISBN 1-898827-99-0)

5. Searl, Prof. John R. R., *The Law of the Squares, Books 1-8*, Direct International Science Consortium, 1993, ISBN 1-898827-00-1
6. Hill, Paul R., *Unconventional Flying Objects*, Hampton Roads, 1995
7. Valone, Thomas, *The Homopolar Handbook, A Definitive Guide to N-Machine and Faraday Disk Technologies*, Integrity Research Institute, 1994
8. Adler, Bazin, and Schiffer. *Introduction to General Relativity*, McGraw Hill, New York, 1965
9. Landau and Lifshitz. *Electrodynamics of continuous media*.- Moscow, Nauka, 1982. (in Russian).

PHYSICS OF GRAVITATIONAL WAVE DETECTION: RESONANT AND INTERFEROMETRIC DETECTORS

Peter R. Saulson*
Department of Physics
Syracuse University, Syracuse, NY 13244-1130

ABSTRACT

I review the physics of ground-based gravitational wave detectors, and summarize the history of their development and use. Special attention is paid to the historical roots of today's detectors.

*Supported by the National Science Foundation, under grant PHY-9602157.

© 1998 Peter R. Saulson.

1 The Nature of Gravitational Waves

1.1 What Is a Gravitational Wave?

There is a rich, if imperfect, analogy between gravitational waves and their more familiar electromagnetic counterparts. A nice heuristic derivation of the necessity for the existence of electromagnetic waves and of some of their properties can be found in, for example, Purcell's *Electricity and Magnetism*.[1] A key idea from special relativity is that no signal can travel more quickly than the speed of light. Consider the case of an electrically charged particle that is subject to a sudden acceleration of brief duration. It is clear that the electric field cannot be everywhere oriented in a precisely radial direction with respect to its present position; otherwise information could be transmitted instantaneously to arbitrary distances simply by modulating the position of the charge. What happens instead is that the field of a charge that has been suddenly accelerated has a kink in it that propagates away from the charge at the speed of light. Outside of the kink, the field is the one appropriate to the charge's position and velocity before the sudden acceleration, while inside it is the field appropriate to the charge subsequent to the acceleration. In each of those regions, the field looks like a radial field for the appropriate charge trajectory, but the kink itself has a transverse component, necessary to link the field lines outside and inside. The transverse field propagating at the speed of light away from the accelerating charge is the electromagnetic wave.

Heuristically, we can use similar reasoning to think about the case of gravitational fields. If the gravitational field of a mass were radial under all circumstances, then instantaneous communication would be possible. The transmitter could be a mass whose position could be modulated, and the receiver would be a device that determined the orientation of its gravitational field. In order that gravity not be able to violate relativity, it must be the case that the gravitational field of a recently accelerated body contains a transverse kink that propagates away from it at the speed of light. This transverse kink is the gravitational wave that carries the news that the body was recently accelerated.

It is helpful to consider the radiation from an extended system, a more realistic description than the point particle moved by prescription assumed in the foregoing paragraphs. In the electromagnetic case the lowest order moment of a charge distribution involved in radiation is the dipole moment, radiation by a time varying monopole moment (i.e., electric charge itself) being forbidden by the law of conservation of charge. The law of conservation of energy plays a similar role in forbidding monopole gravitational radiation, since that would require the total mass of an isolated system to vary. In addition, though, there can be no time-varying mass dipole moment of an isolated system; conservation of linear momentum forbids this. The difference between the gravitational case and the electromagnetic one is that the "gravitational charge-to-mass ratio" is the same for all bodies, by the Principle of Equivalence; the equal and opposite reaction that accompanies any action on a given body precisely cancels second time derivatives of the mass dipole moment of any isolated system. Similarly, the law of conservation of angular momentum forbids the gravitational equivalent of the magnetic dipole moment from varying in time. The lowest order moment of an isolated mass distribution that can accelerate is its quadrupole moment

$$I_{\mu\nu} \equiv \int dV \left(x_\mu x_\nu - \frac{1}{3}\delta_{\mu\nu} r^2 \right) \rho(\mathbf{r}) .$$

The second time derivative of the mass quadrupole moment plays the same role in gravitational radiation as does the first time derivative of the charge dipole moment in electromagnetic radiation, that of leading term in typical radiation problems.

But what is the analog of the electric field in gravitational wave problems, that is to say, the wave field amplitude itself? An answer comes from study of the Einstein field equations in the weak field limit. Specifically, one considers the case where the space-time metric can be approximated as $g^{\mu\nu} = \eta^{\mu\nu} + h^{\mu\nu}$, where $h^{\mu\nu}$ is a small perturbation of the Minkowski metric $\eta^{\mu\nu}$. Then, if one chooses a particular coordinate condition (the so-called "transverse traceless gauge"), the Einstein equations become a wave equation for the perturbation $h^{\mu\nu}$. Furthermore, the perturbation takes on a particular form: for a solution to the wave equation corresponding to a wave travelling along the z axis, h must be a linear combination of the two basis tensors

$$h_+ = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

and

$$h_\times = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} .$$

Notice that for this wave along the z axis, the effects of the wave are only in the x and y components. In other words, the wave is transverse.

1.2 The Effect of Gravitational Waves on Test Bodies

The statement that the gravitational wave amplitude is the metric perturbation tensor h is probably hard to visualize without considering some examples. Imagine a plane in space in which a square grid has been marked out by a set of infinitesimal test masses (so that their mutual gravitational interaction can be considered negligible compared to their response to the gravitational wave). This is a prescription for embodying a section of the transverse traceless coordinate system mentioned earlier, marking out coordinates by masses that are freely-falling (i.e., that feel no non-gravitational forces).

Now imagine that a gravitational wave is incident on the set of masses, along a direction normal to the plane. Take this direction to be the z axis, and the masses to be arranged along the x and y axes. Then, if the wave has the polarization called h_+, it will cause equal and opposite shifts in the formerly equal x and y separations between neighboring masses in the grid. That is, for one polarity of the wave, the separations of the masses along the x direction will decrease, while simultaneously the separations along the y direction will increase. When the wave oscillates to opposite polarity, the opposite effect occurs.

If, instead, a wave of polarization $h_\times$ is incident on the set of test masses, then there will be (to first order in the wave amplitude) no changes in the distances between any mass and its nearest neighbors along the x and y directions. However, $h_\times$ is responsible for a similar pattern of distance changes between a mass and its next-nearest neighbors along the diagonals of the grid.

There are several other aspects of the gravitational wave's deformation of the test system that are worth pondering. Firstly, the effect on any pair of neighbors in a given direction is identical to that on any other pair. The same *fractional* change occurs between other pairs oriented along the same direction, no matter how large their separation. This means that a larger *absolute* change in separation occurs, the larger is the original separation between two test masses. This property, which we can call "tidal" because of its similarity to the effect of ordinary gravitational tides, is exploited in the design of interferometric detectors of gravitational waves.

Another aspect of this pattern that is worthy of note is that the distortion is uniform throughout the coordinate grid. This means that any one of the test masses can be considered to be at rest, with the others moving in relation to it. In other words, a gravitational wave does not cause any absolute acceleration, only relative accelerations between masses. This, too, is fully consistent with other aspects of gravitation

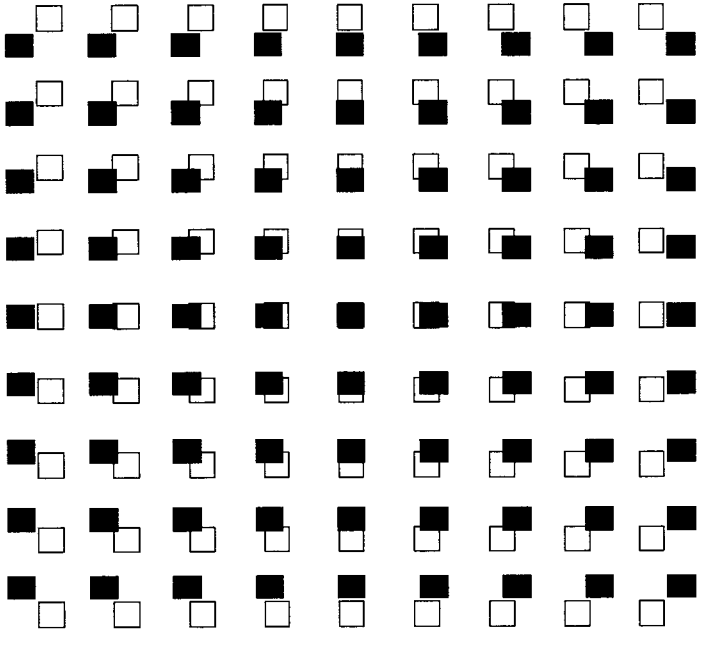


Fig. 1. An array of free test masses. The open squares show the positions of the masses before the arrival of the gravitational wave. The filled squares show the positions of the masses during the passage of a gravitational wave of the plus polarization.

as described by the general theory of relativity: a single freely-falling mass cannot tell whether it is subject to a gravitational force. Only a measurement of relative displacements between freely-falling test masses (the so-called "geodesic deviation") can reveal the presence of a gravitational field.

1.3 A *Gedanken* Experiment to Detect a Gravitational Wave

In the discussion in the preceding section, we took it for granted that the perturbations h_+ and $h_\times$ to the flat-space metric were, in some sense, real. But it is only by considering whether such effects are measurable that one can be convinced that a phenomenon like a gravitational wave is meaningful, rather than a mathematical artifact that could be transformed away by a suitable choice of coordinates.

To demonstrate the physical reality of gravitational waves, consider the example system of the previous section. We will concentrate our attention on three of the test masses, one chosen arbitrarily from the plane, along with its nearest neighbors in the $+x$ and $+y$ directions. Imagine that we have equipped the mass at the vertex of this "L" with a lamp that can be made to emit very brief pulses of light. Imagine also that the two masses at the ends of the "L" are fitted with mirrors aimed so that they will return the flashes of light back toward the vertex mass.

First, we will sketch how the apparatus can be properly set up, in the absence of a gravitational wave. Let the lamp emit a train of pulses, and observe when the reflected flashes of light are returned to the vertex mass by the mirrors on the two end masses. Adjust the distances from the vertex mass to the two end masses until the two reflected flashes arrive simultaneously.

Once the apparatus is nulled, let the lamp keep flashing, and wait for a burst of gravitational waves to arrive. When a wave of $\hat{h}_+$ polarization passes through the apparatus along the z axis, it will disturb the balance between the lengths of the two arms of the "L." Imagine that the gravitational wave has a waveform given by

$$h^{\mu\nu} = h(t)\hat{h}_+.$$

To see how this space-time perturbation changes the arrival times of the two returned flashes, let us carefully calculate the time it takes for light to travel along each of the two arms.

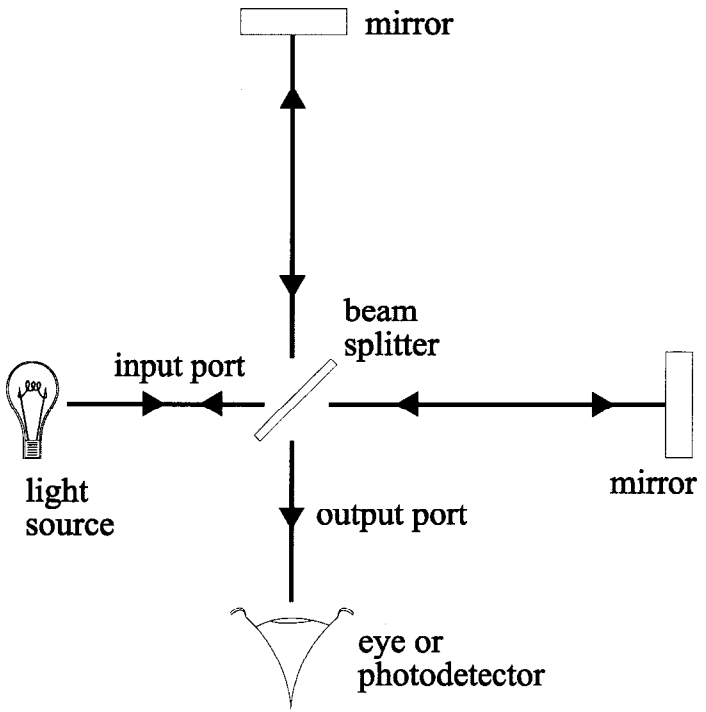


Fig. 2. A schematic diagram of an apparatus that can detect gravitational waves. It has the form of a Michelson interferometer.

First, consider light in the arm along the x axis. The interval between two neighboring space-time events linked by the light beam is given by

$$\begin{aligned} ds^2 = 0 &= g_{\mu\nu}dx^\mu dx^\nu \\ &= (\eta_{\mu\nu} + h_{\mu\nu})\, dx^\mu dx^\nu \\ &= -c^2dt^2 + (1 + h_{11}(2\pi ft - kz))\, dx^2. \end{aligned} \tag{1}$$

This says that the effect of the gravitational wave is to modulate the square of the distance between two neighboring points of fixed coordinate separation dx (as marked, in this gauge, by freely-falling test particles) by a fractional amount h_{11}.

We can evaluate the light travel time from the beam splitter to the end of the x arm by integrating the square root of Eq. (1)

$$\int_0^{\tau_{out}} dt = \frac{1}{c}\int_0^L \sqrt{1 + h_{11}}dx \approx \frac{1}{c}\int_0^L \left(1 + \frac{1}{2}h_{11}\,(2\pi ft - kz)\right) dx, \tag{2}$$

where, because we will only encounter situations in which $h \ll 1$, we've used the binomial expansion of the square root, and dropped the utterly negligible terms with more than one power of h. We can write a similar equation for the return trip

$$\int_{\tau_{out}}^{\tau_{rt}} dt = -\frac{1}{c}\int_L^0 \left(1 + \frac{1}{2}h_{11}(2\pi ft - kz)\right) dx. \tag{3}$$

The total round trip time is thus

$$\tau_{rt} = \frac{2L}{c} + \frac{1}{2c}\int_0^L h_{11}(2\pi ft - kz)dx - \frac{1}{2c}\int_L^0 h_{11}(2\pi ft - kz)dx. \tag{4}$$

The integrals are to be evaluated by expressing the arguments as a function just of the position of a particular wavefront (the one that left the beam-splitter at $t = 0$) as it propagates through the apparatus. That is, we should make the substitution $t = x/c$ for the outbound leg, and $t = (2L - x)/c$ for the return leg. Corrections to these relations due to the effect of the gravitational wave itself are negligible.

A similar expression can be written for the light that travels through the y arm. The only differences are that it will depend on h_{22} instead of h_{11} and will involve a different substitution for t.

If $2\pi f_{gw}\tau_{rt} \ll 1$, then we can treat the metric perturbation as approximately constant during the time any given flash is present in the apparatus. There will be equal and opposite perturbations to the light travel time in the two arms. The total travel time difference will therefore be

$$\Delta\tau(t) = h(t)\frac{2L}{c} = h(t)\tau_{rt0}, \tag{5}$$

where we have defined $\tau_{rt0} \equiv 2L/c$.

If we imagine replacing the flashing lamp with a laser that emits a coherent beam of light, we can express the travel time difference as a phase shift by comparing the travel time difference to the (reduced) period of oscillation of the light, or

$$\Delta\phi(t) = h(t)\tau_{rt0}\frac{2\pi c}{\lambda}. \tag{6}$$

Another way to say this is that the phase shift between the light that traveled in the two arms is equal to a fraction h of the total phase a light beam accumulates as it traverses the apparatus. This immediately says that the longer the optical path in the apparatus, the larger will be the phase shift due to the gravitational wave.

Thus, this *gedanken* experiment has demonstrated that gravitational waves do indeed have physical reality, since they can (at least in principle) be measured. Furthermore, it suggests a straightforward interpretation of the dimensionless metric perturbation h. The gravitational wave amplitude gives the fractional change in the difference in light travel times along two perpendicular paths whose endpoints are marked by freely-falling test masses.

1.4 Another Way to Picture the Effect of a Gravitational Wave on Test Bodies

In standard laboratory practice, it is not customary to define coordinates by the world-lines of freely-falling test masses. Instead, rigid rulers usually are used to do the job. The forces that make a rigid ruler rigid are something of a foreign concept in relativity, appearing ugly and awkward after the gravitational force has been made to disappear by expressing it as the curvature of space-time. On the other hand, non-gravitational forces are not only a fact of nature, but part of the familiar world of the laboratory. For many purposes, it is convenient to retreat from a purely relativistic picture and instead use a Newtonian picture in which gravity is treated as force on the same level as other forces.

What we are seeking is not a different theory of gravitational waves, but a translation of the theory discussed in the previous section into more familiar language. So let us reconsider the same *gedanken* experiment as before, but imagine that we have augmented the equipment with a rigid ruler along each axis. We saw that when a gravitational wave passed through our set of test masses, the amount of time it took for light to travel from the vertex mass to the end mass and back was made to vary. How can we

describe how this came about in the standard language of the laboratory? If we imagine (notwithstanding their fixed coordinates in the transverse traceless gauge) that the test masses have moved in response to the gravitational wave, we can have a consistent picture of the effect. What is necessary is that the gravitational wave give a tidal force across the pair of masses that will cause them to move apart by the amount necessary to account for the change in light travel time through the system. It is as if the far masses felt forces whose magnitude were given by

$$F_{gw} = \frac{1}{2} m L \frac{\partial^2 h_{11}}{\partial t^2}, \tag{7}$$

where m is the mass of each of the two test bodies, and L is the separation between their centers of mass.

There are several features of this expression that are worthy of note. The force is proportional to the mass of the test bodies, as required by the Principle of Equivalence. The force is also proportional to the separation between the two test masses, making it akin to conventional gravitational tidal forces. The dependence of the gravitational wave force on the second time derivative of h is reminiscent of Newton's Second Law $F = m\ddot{x}$. A natural interpretation follows: in a conventional laboratory coordinate system, free masses actually change their separation by an amount $\Delta L = hL/2$.

Note how in two different coordinate systems the same phenomenon (and particularly the same measurement) is described in completely different language. In transverse traceless coordinates, the free test masses still just fall freely, each marking out its own coordinate (by definition) under any gravitational influence, but the light travel time between them changes as the metric of space-time varies. In standard laboratory coordinates, light travel time changes because the test masses move. Neither of these pictures is more "correct" than the other. The laboratory-coordinate picture is markedly more convenient for seeing how to combine the effect of a gravitational wave with the effects of noise forces of various kinds. The transverse traceless coordinates offer the most clarity when one wants to consider extreme cases, such as test masses separated by distances comparable to or longer than the gravitational wave's wavelength.

2 Generating Gravitational Waves

As mentioned above, the second time derivative of the mass quadrupole moment I plays the same role in gravitational wave emission as does the first derivative of the charge dipole moment in electromagnetic radiation problems, that of strongest source term in most situations. More specifically, the expression for gravitational wave generation is

$$h_{\mu\nu} = \frac{2G}{Rc^4} \ddot{I}_{\mu\nu}, \tag{8}$$

usually referred to as the "quadrupole formula."[2]

Something about this expression should immediately give one pause—the prefactor of $2G/c^4$. In SI units, this has the value 1.6×10^{-44} sec^{2}kg^{-1}m^{-1}. It will take tremendously large values of $\ddot{I}/R$ in order for even modest values of h to be generated. *A priori*, one can think of two strategies that might work: make R small, or make $\ddot{I}$ large.

2.1 Laboratory Generators of Gravitational Waves

To construct a source of gravitational waves in the laboratory would allow one to have the benefit of placing it as close as possible to one's detector, thus exemplifying the first strategy in the previous paragraph. Of course, it would have other benefits as well. Control of the waveform, polarization, and other features would enable the detector to be carefully optimized to match the signal. At an even deeper level, confidence in the detection of gravitational waves could be assured by the requirement that they must be seen when, and only when, they were being emitted.

What one would really like to do is to replicate for gravitational waves what Hertz was able to accomplish for electromagnetic ones. His experiments of 1886–'91 not only conclusively demonstrated the existence of electromagnetic waves, they validated Maxwell's theory of electromagnetic radiation by exploring the rich phenomenology of polarization, reflection, and interference. They also began the process of harnessing the phenomenon for practical use. Marconi's work started by his following closely in Hertz's footsteps, and real long-distance communication via radio was not long in coming.

Unfortunately, no practicable way has been conceived to replicate Hertz's success in the gravitational domain. Assume we could construct a dumbbell consisting of two masses of one ton each, at either end of a rod two meters long. Spin this quadrupole about an axis orthogonal to the connecting rod passing through its midpoint, at an angular frequency $f_{rot} = 1$ kHz. Neglecting for simplicity the contribution of the connecting rod, we have a system very similar to a binary star system. The amplitude of the gravitational waves generated by this device will be

$$h_{lab} = 2.6 \times 10^{-33}\text{m} \times \frac{1}{R}. \tag{9}$$

Before we rush to plug in a distance R of a few meters, as Hertz was able to do for his experiment, we need to remember that wave phenomena are only distinguishable from near-field effects in the "wave zone," that is, at distances from the source comparable to or larger than one wavelength. With $\omega_{rot} = 2\pi \times 1$ kHz, we have $\lambda = 300$ km! The receiver for our Hertzian experiment must be at least that far away from the transmitter. Hertz's electromagnetic experiments involved waves of six meters down to 60 cm in length, so the distance across the lab was fine for him.

At a distance of one wavelength, our laboratory generator gives gravitational waves of amplitude

$$h_{lab} = 9 \times 10^{-39}. \tag{10}$$

This is pretty small.

Even creating such a strong source as this may not be practicable. Consider the stress in the connecting rod of the dumbbell. It must supply the centripetal force necessary for the masses to move in a circle. If the rod were made of good steel, it would need a cross-sectional area substantially greater than that of a one ton sphere in order not to fail under the stresses in a device with the parameters we have assumed. So we'd have to reduce the rotation frequency to keep the generator from flying apart, with a consequent reduction in the transmitted wave amplitude.

2.2 Astrophysical Sources of Gravitational Waves

Even if a gravitational version of the Hertz experiment is not feasible, all is not lost for the detection of gravitational waves. The best reason for optimism that detectable levels of gravitational radiation exist comes from the presence in the universe of objects with truly remarkable values of $\ddot{I}$. These systems are so extreme that even though their distances from our detectors are quite large, they still generate gravitational waves with amplitudes that exceed by almost 20 orders of magnitude the signal strengths from laboratory generators of the type described above.

It would be beyond the scope of this review to describe in detail all of the many astronomical objects that might be important sources of gravitational waves. Readers are urged to consult the article by Finn in these proceedings for further information on the variety of possible sources. But for the sake of a self-contained treatment, we show here how to estimate the magnitude of the strongest gravitational waves arriving at the Earth.

For the case of a binary star, there is an elegant way (due to Kafka[3]) of writing the amplitude of the quasi-sinusoidal gravitational wave strain. We can massage the quadrupole formula into a manifestly dimensionless form by recognizing that the mass dependence can be rewritten as a proportionality to the product of the Schwarzschild radii $R_s = 2GM/c^2$ of the stars. The frequency dependence and all of the stray factors remaining collect nicely as the separation r of the two stars. The gravitational wave amplitude is

$$h_{ns} = R_{s1}R_{s2}/rR. \tag{11}$$

If the binary consists of two neutron stars, then the Schwarzschild radii are both about 4 km. Astronomers estimate that within a sphere of radius 200 Mpc, roughly one of these sytems will coalesce each year. When the stars have a separation of ten diameters (or around 200 km), then the signal we would receive from that distance will have an amplitude of almost 10^{-23}. The stars can probably approach closer still before the system is destroyed.

A glance at this expression shows why a neutron star binary is a good choice as a strong source of gravitational waves. The substantial masses of the two stars make the numerator large. The fact that they are compact objects means that their separation r can be quite small. We could always wish that the distance R to the nearest example of such a system were smaller, but even so our estimated signal strength, while small in absolute terms, is certain dramatically larger than we were able to produce in our model laboratory generator.

Perhaps the only sort of astronomical system we can imagine that might generate stronger gravitational waves would be a binary system consisting of two black holes. Although it may be hazardous to treat such dramatically relativistic objects with the quasi-Newtonian physics used to derive Eq. (11), it will probably still give a good order of magnitude estimate. The possible advantages of black holes as sources of gravitational waves are twofold. Firstly, it is possible that the masses of black holes may be substantially in excess of the 1.4 $M_\odot$ typical of neutron stars. Secondly, black holes can approach to a separation r as close as their Schwarzchild radius R_s without disruption; instead the two will coalesce into a single larger black hole. Thus we guess that the gravitational signal from a black hole coalescence could be as large as

$$h_{bh} \sim R_s/R. \tag{12}$$

For a pair of 10 $M_\odot$ black holes at 200 Mpc, this expression would indicate a signal of $h \sim 5 \times 10^{-21}$.

Of course, we do not have nearly such secure knowledge of the existence of such binary systems as we do for the neutron star case. There is strong evidence for the existence of 10 $M_\odot$-class individual black holes in binary systems with main sequence stars.[4] The abundance of black hole binaries in this mass range is unknown, so the validity of our choice of 200 Mpc as a fiducial distance is uncertain at best.

One weakness of the elegant expressions Eqs. 11 and 12 is that they do not explicitly refer to the frequency of the gravitational wave. As we will see later, the various kinds of detectors of gravitational waves will each function best only in a certain band of frequencies. For reference, we anticipate the results of that discussion with the following summary: resonant-mass detectors work best in the vicinity of 1 kHz, interferometric detectors from around 10 Hz to a few kHz, and space interferometers between around 10^{-4} and 10^{-1} Hz. The strongest black hole signal will come at the frequency f_{qnm} of the lowest quasi-normal mode of the resulting combined black hole,

$$f_{qnm} \approx 0.7 \frac{c^3}{2GM}.$$

This makes resonant-mass detectors best suited for looking for black holes with masses of 10 $M_\odot$. Terrestrial interferometers will look for 3 $M_\odot$ to 1000 $M_\odot$ objects, while interferometers in space will search for black holes in the range near 10^6 $M_\odot$.

2.3 Summary

The contrast between terrestrial and astronomical sources of gravitational waves is striking. In spite of the fact that astronomical generators will only be found at distances far beyond the optimum, their other physical parameters are such that they will provide the strongest available gravitational wave signals. The contrast is so strong that this advantage alone must outweigh the ability to control the character and timing of the signal that would be inherent in a laboratory generator. Instead, even the most basic experiments on the nature of gravitational waves will necessarily involve astronomical observations.

3 Weber and the Birth of Gravitational Wave Detection

3.1 Weber's Original Vision

It is reasonable to ask the question whether even the astronomical signals are large enough to be detected by any conceivable device. On the face of it, the odds are daunting. A strain of 10^{-20}, for example, only generates relative motion of 10^{-20} meters between two test masses separated by one meter. Compare this with other characteristic length scales, and the challenge is clear: wavelength of visible light of 10^{-6} meter, atomic diameter of 10^{-10} meter, nuclear diameter of 10^{-15} meter. Nevertheless, it appears that gravitational wave signals from astronomical sources will probably soon be detected. It is the purpose of the rest of this review to show how this is possible.

There are at least two questions that might be raised by consideration of the numbers listed in the previous paragraph. The first is whether measurements of a macroscopic body can be capable of resolving motions substantially smaller than nuclear diameters. The second question that one might ask is whether success might be easier if the scale of the apparatus were made substantially larger than a meter, in order to take advantage of the fact that test masses separated by a larger distance will move by a proportionally larger amount in response to a gravitational wave of a given strength. Pondering the answers to these questions will lead to understanding of the most promising techniques for gravitational wave detection.

The postwar atmosphere of optimism about astronomical progress must have swept up Joseph Weber in the late 1950s. Weber became convinced that the time was right to try to extend the astronomical revolution beyond the electromagnetic spectrum. At that time, it was not obvious that strain sensitivities of 10^{-22} should be the goal. It was equally plausible that objects such as we discussed above might possibly be abundant enough that their typical distance might be the few kpc associated with galactic dimensions instead of 200 Mpc. So strains of 10^{-17} or perhaps even stronger might have been the proper goal to aim for. (Weber knew of a very optimistic estimate of wave strength by Wheeler,[5] which would allow an energy density of order of the cosmological closure density.) Weber's thinking showed the way to achieve such strain sensitivities; indeed, devices following directly in a line of development from his first instrument have in the past few years approached rms sensitivities of 10^{-19}, with the prospect of extension to a new generation of detectors sensitive to waves with amplitudes of order 10^{-20} or better.

Weber's early thinking is described in his *Physical Review* article of 1960,[6] and placed in the larger context of his thinking about general relativity in a small monograph published in 1961.[7] He describes a conceptual detector, in reciprocal relationship to a gravitational wave emitter, as a simple "mass quadrupole," sketched as two masses connected by a spring. Weber extends the general relativistic equation of geodesic deviation to include the non-gravitational forces applied by the elastic restoring force and the mechanical dissipation in the spring. The equation of motion of the system

then becomes that of a simple harmonic oscillator, with the driving term given by the effective force from the gravitational wave [our Eq. (7)].

Weber next shows how an extended elastic body behaves in such a way that each of its normal modes of vibration can be studied independently. (The gravest mode of a cylinder has a large quadrupole moment, and is the one that is usually used for detection.) He focuses attention on the use of a piezoelectric crystal as the detecting body, partly because he hopes that the electric field will make it a detector with effective size larger than half an acoustic wavelength, but also in large measure because the electric fields generated by the gravitational-wave-induced stress will give an integrated voltage between its ends that may be "large enough to be observed with a low-noise radio receiver." Weber calculates the amount of mechanical power that a sinusoidal gravitational wave can dissipate in the resonant detector as a function of frequency, then invokes the standard electrical network theorems to show what fraction of this power can be transferred to the input impedance of an amplifier.

A simple discussion of sensitivity follows. Weber first remarks that "in microwave spectroscopy it has been found that all spurious effects other than random fluctuations can be recognized." Then Weber states that the excitation of the detector must exceed the noise power associated with its thermal excitation.

Finally, Weber discusses possible practical experimental arrangements. In most of the discussion the devices are supposed to be made of large blocks of piezoelectric material. But in a footnote Weber states that the experimental work he is carrying out with David Zipoy and Robert L. Forward will probably make use of a large block of metal instead. (This is justified on the grounds that a half-wavelength at the 1 kHz frequency being contemplated is already large; thus the piezoelectric length-enhancement effect may not be necessary, and in any case such a large block of piezoelectric material "may not be obtainable as a single crystal".)

Two experimental strategies are foreseen: use of a single detector with examination of its output for a diurnal cycle associated with the scanning of its sensitivity pattern across the sky, and the cross-correlation of a pair of detectors so that external influences (presumably gravitational waves) can be distinguished from "internal fluctuations." He notes the necessity of preventing the excitation of the detector by "earth vibrations," and discusses an "ingenious" idea of Zipoy's for what is now called active vibration isolation.

Weber's very concise discussion is remarkable for the prescience with which it foreshadowed not only his own work, but that of so many others. It also marks a watershed in the history of general relativity. In a single blow, Weber wrested consideration of gravitational waves from theorists concerned about issues such as exact solutions, and appropriated the subject instead for experimentalists trained in issues of radio engineering. The boldness and brilliance of this move are remarkable.

3.2 The Logic of Weber's Idea

Weber sweeps quickly over a variety of issues that are worthy of more leisurely consideration. We'll give an overview of the important issues in this section, then devote the rest of this review to discussing their implications.

The detector Weber outlined can be divided into several subsystems: a set of test masses that respond to the gravitational wave, a transduction system that converts this mechanical response to a convenient electrical signal, a low-noise pre-amplifier, and a post-amplification averaging and recording mechanism. Notwithstanding the cleverness of Weber's original version, many variations on his basic scheme are possible, and indeed are responsible for much of the progress since he first announced the results of gravitational wave observations in 1969.[8]

Let's see how to analyze the original Weber design into these canonical subsystems. Weber explicitly pointed out how one could construct an analog of a pair of lumped test masses by monitoring an internal mode of vibration of an extended block of elastic material. In the version where this block is made of piezoelectric material, the same material serves both as test masses and as transducer from mechanical to electrical signal form. In the version in Weber's footnote (the one he actually built) a large aluminum cylinder serves as the set of test masses; piezoelectric strain gauges glued about the girth of the cylinder perform the transduction. The pre-amplifier is Weber's low-noise radio receiver. No averaging filter is shown in Weber's diagrams, but is implicit in his discussion.

Perhaps the most interesting choice that Weber made was to connect his test masses in a resonant system. It appears that Weber, at least in 1961, thought this was a necessity. In a footnote, he cites previous work by Pirani[9] in which the latter considered "measurement of the Riemann tensor by comparing accelerations of free test particles," but Weber continues, "The results of this chapter indicate that interacting particles must be used, in practice." In fact, it is not required either in principle or in practice, but it is interesting to consider why Weber may have thought so then, and what advantages still accrue to the use of resonant masses.

Weber couches a good deal of his discussion in terms of steady sinusoidal signals, still a common practice in much of engineering and even more so around 1960. If a gravitational wave did have this form, then masses connected by a spring give a resonant amplification of the response to a signal at the resonant frequency. The amount of this amplification is given by the resonator's quality factor Q, and can be substantial; Weber quotes an estimated $Q \sim 10^6$, still not a bad ballpark number.

On the other hand, there is essentially no resonant amplification if one has a sinusoidal signal whose frequency does not closely match the resonant frequency of the detector, or if the signal has a broad-band frequency content, as it would if it were a brief burst. Resonant amplification only comes about when the input force drives the resonant system with the proper phase for a substantial number of cycles; this can only occur when there is a good match between the signal frequency and the mechanical resonant frequency.

But even though the search for gravitational waves has come to focus mostly on burst-like signals, the resonant-mass configuration can still give a powerful advantage, albeit one not discussed by Weber in his 1961 book. A weak signal must compete for visibility against the noise in the pre-amplifier stage. This is why Weber made a point of calling for the lowest noise levels possible in this component. The noise in such amplifiers is generally of a broad-band character, best characterized by its power spectral density $S_v(f)$ which is typically roughly constant (or "white") over a wide range of frequencies. Usually there is an additional $1/f$ component that dominates at low frequencies.

The extent to which this noise competes with a signal depends in an essential way on the duration of the signal. We use the term "burst" to refer to a signal of limited duration in time; call its length τ_s. A fundamental theorem of signal detection states that the optimum contrast between a given signal and white noise can be attained when the time series containing the noise plus any possible signals is convolved with a template of the same form as the signal. This is called the *matched filter* when it is implemented in real time by an analog device. The heuristic idea behind such an optimum is that the matched filter rejects all components of the noise that do not "look like" the signal for which one is searching. Still, some noise passes through the matched filter. How much? If the Fourier transform of the signal waveform $v(t)$ is given by $V(f)$, then it passes noise power of

$$N^2 = \int_{-\infty}^{\infty} |V(f)|^2 S_v(f) df.$$

Another very general theorem of Fourier analysis takes the form of a classical "uncertainty relation." It states that there is an inverse relationship between the duration of a signal in the time domain and its width in the frequency domain:

$$\Delta f \Delta \tau \approx 1.$$

See what this implies for the question at hand. If we are looking for a brief signal, then its matched filter passes noise of a wide bandwidth. Thus, brief signals compete much less well against broad-band pre-amplifier noise than does a long-duration signal of the same amplitude.

Here is where a resonant detector of gravitational waves can make a difference even when one is looking for a broad-band signal. If the gravitational wave signal $h(t)$ contains substantial power in the vicinity of the detector's resonant frequency, then it will excite the motion of the detector's mode at an amplitude $\Delta L \sim hL$, not much different from if the test masses were free. But the subsequent behavior of the resonant detector is quite different from if the detector were made of free masses. The motions of free test masses only persist for the duration τ_s of the gravitational wave signal. But the resonant detector "rings" for a time of order of the damping time of the resonance,

$$\tau_d = \frac{Q}{\pi f_0} \gg \tau_s,$$

where f_0 is the resonant frequency of the detector.

It is the motion of this resonant system, converted to electrical form by the transducer, that is presented to the input terminals of the pre-amplifier. So it is an electrical signal of long duration τ_d that competes with the amplifier noise. As a long duration deterministic signal, its matched filter has a much narrower width in frequency than had the original signal $h(t)$, and so passes a much smaller proportion of the amplifier noise. Thus, the resonant response of the test mass system allows a weak signal to compete more effectively against amplifier noise than would be the case with free masses.

3.3 The Cost of Resonant Detection

This advantage of resonant-mass detectors is substantial; it is responsible for the continued vitality of the Weber style of detector over 30 years after it was first proposed. Still, it comes with a price that is not negligible. Implementing the matched filter described above, which is essential to attaining the advantage of the resonance, is tantamount to averaging the output of the amplifier for times of order τ_d. In the jargon of the field,

such a system has a low *post-detection bandwidth* (usually shortened simply to "bandwidth.") The averaging washes out any details of the waveform $h(t)$ on time scales short compared to τ_d. What one gains in signal-to-noise ratio, one gives up in temporal resolution. Whether this is a price one ought to be willing to pay or not depends on the stakes: if it is absolutely necessary even to detect the signal, averaging with a matched filter is certainly worthwhile. If the signal could be detected anyway, averaging simply throws away information, and should be avoided. In the high signal-to-noise case, the resonance does not help, but neither does it hurt much—a simple filtering operation could remove the resonant signature and allow reconstruction of the original signal waveform.

(N.B.: As we will show below, the actual choice of matched filter for a resonant detector is more subtle than that just described. Instead of τ_d, a shorter averaging time is almost always the optimum choice. Nevertheless, the qualitative thrust of the argument given in the previous paragraph still applies.)

3.4 Free-Mass Detectors as an Alternative

Given the trade-off between sensitivity and bandwidth that resonant systems tempt one to make, it is worth exploring whether there are other entirely non-resonant detection schemes that can achieve high sensitivity to gravitational waves without sacrificing signal bandwidth. In fact, such free-mass detectors have been developed by a variety of workers, including the same Robert Forward who worked with Weber on the original resonant detector.[10,11] The essential advantage of free-mass detectors comes from the fact that the farther apart their test masses are placed, the larger is the relative displacement between them caused by a given gravitational wave amplitude $h(t)$. (This scaling relation holds true up to the point that the light travel time between the masses becomes comparable to the period of the wave; that is when separation of the masses becomes comparable to the wavelength of the wave.) But the resonance in a resonant detector comes roughly when the sound travel time across the bar matches the period of the wave. That is to say, resonant detectors reach their optimum sensitivity when the separation of the test masses is of order of the acoustic wavelength at the gravitational wave frequency. Since the speed of sound in materials is of order 10^{-5} of the speed of light, a free-mass detector at its optimum length can have an advantage in signal size of 10^5 over a resonant-mass detector at its optimum length.

Another advantage is that no resonance is used to boost the signal. Thus, in principle a free-mass detector can have a completely white frequency response. This ideal can not be completely achieved in practice, since some of the noise sources discussed below have strong frequency dependences of their own. Still, it is possible to achieve useful bandwidths measured in decades rather than in fractions of an octave.

This signal size advantage would be a hollow one if there were no sensitive way to measure the relative displacement of test masses separated by many kilometers. Fortunately, there are such ways. As we saw above, the travel time of electromagnetic signals between the test masses can be measured with great precision. Interferometry using visible or near-infrared light to measure the separation of free masses has become a well-developed technology that now is completely competitive with the best resonant-mass detectors, and which is about to undergo a great leap in sensitivity as new instruments of multi-kilometer scale come on line in the next couple of years. Radio ranging between interplanetary space probes separated by many millions of kilometers has been used for some time; optical interferometers in solar orbit, with million kilometer baselines, are now being planned.

The conceptually simpler free-mass detectors are in practice substantially more complicated devices; the freedom of the test masses must be tamed by servo systems to keep them operating properly. This is in part what is responsible for the time lag in their development, even though they were conceived not much later than resonant-mass detectors. In the remainder of the review, we will discuss both styles of gravitational wave detector.

4 Noise Sources

In this section, we will focus our attention on understanding the most fundamental noise sources with which the practice of gravitational wave detection has to contend. Perhaps not surprisingly, the list will seem to have little to do with general relativity or with gravitational waves, as such. The chief concerns of gravitational wave detector designers are those that would confront anyone attempting to measure the effect of a very weak force on a mechanical system: Brownian motion (also known as thermal noise), and noise from the readout system (both in its direct influence on the output of the system and through its "back-reaction" on the mechanical front end). A ubiquitous but non-fundamental noise source, seismically-induced vibration, is treated as well.

It is pedagogically simpler to introduce the topics first in the context of interferometers. Then, we will describe how similar considerations apply to resonant-mass

detectors, where the signal processing issues are a bit more subtle.

4.1 Thermal Noise

The first recognition of a classical physical limit to measurement precision occurred when thermal noise was discovered in galvanometers, in the early 1930's.[12] The premier current-measuring instrument of their day, galvanometers typically consist of a coil of wire suspended from a fine fiber so that it rests between the poles of a strong permanent magnet. Leads from the coil are attached to the external source of current, which generates a torque about the vertical axis. The resulting angular displacement of the coil can be read through an optical lever arrangement that uses a small mirror fixed to the coil.

From our point of view as students of mechanical instruments, the components of a galvanometer that make it function as a current-measuring device are less interesting than its basic mechanical features as a single degree-of-freedom oscillator: an inertia element (characterized by the moment of inertia I of the coil, mirror, and connecting fixtures), an elastic element (represented by the torsional spring constant κ of the fiber from which the coil was suspended), mechanical damping (from some combination of air friction, electrical resistance, and internal friction in the suspension), and a means of observing the angular coordinate θ (the optical lever).

Attempts to make current measurements of the highest precision confronted the fact that, scrutinized carefully, the angle of the coil was not strictly fixed, but instead jittered about its mean position. At first, it seemed natural to attribute the motion to seismic excitation of the galvanometer and its surroundings. But careful mounting of galvanometers to rigid piers isolated from excess building vibrations only went so far in minimizing the noise. Furthermore, seismic noise is typically strongly variable with time, depending on the violence of the weather and on the diurnal cycle of human activity. But well-constructed galvanometers exhibit a noise whose amplitude does not vary in time.

It was not long after Einstein's explanation of Brownian motion as the result of random impacts upon the observed object by molecules from the surrounding fluid that physicists recognized that galvanometers were exhibiting the same phenomenon. According to the Equipartition Theorem, each degree of freedom of a system in thermodynamic equilibrium at temperature T should have an energy whose expectation value is $k_BT/2$. Applying this to the potential energy $\frac{1}{2}I\omega_0^2\theta^2$ associated with the angular displacement θ of the galvanometer, one finds

$$\theta_{rms} = \sqrt{\frac{k_BT}{I\omega_0^2}}, \tag{13}$$

where $\omega_0 = \sqrt{\kappa/I}$ is the resonant frequency of the galvanometer. The heuristic explanation of this theorem is the essentially atomic nature of all mechanisms of dissipation; in this case by the random impact of air molecules on the coil (in cases where air friction dominates the damping) or the random motion of electrons through the coil driving a noisy magnetic torque (if electrical dissipation dominates).

Because it is rooted in thermodynamics, this result has a generality far beyond the details of any particular system. There is a natural analogy between the galvanometer's torsional oscillator and the fundamental longitudinal mode of a Weber-style resonant-mass detector of gravitational waves. Note one striking feature of Eq. (13): even though the origin of the fluctuations lies in the mechanism that is responsible for the dissipation, the rms displacement does not depend on the magnitude (let alone on the mechanism) of the dissipation.

Subsequent work brought recognition of analogous phenomena in other systems. Perhaps the most important was the discovery of the electrical analog of Brownian motion by Johnson and Nyquist in 1928.[13,14] But full recognition of the essential unity of all thermodynamic fluctuation phenomena awaited the formulation of the Fluctuation-Dissipation Theorem. A particularly useful form was established in 1951 and 1952 by Callen and co-workers.[15] We will quote it in a form most directly applicable to mechanical systems, but the derivations in the original papers make it clear how it can be applied to any linear system in thermodynamic equilibrium.

In Callen's formulation, it is convenient to describe the dynamics of a physical system in terms of the network functions called the impedance and admittance, as measured at the point of interest in the system. These are defined in terms of the steady-state response of the system to sinusoidal excitation. The impedance Z is defined as the complex ratio of the force applied at the point of interest to the resulting velocity at that point. That is, if the system is driven with a force $F_0e^{i\omega t}$ and it responds in the steady state with $v = v_0e^{i(\omega t+\phi)}$, then the impedance is

$$Z(\omega) = \frac{F_0}{v_0}e^{-i\phi}.$$

A point mass has an impedance $Z_m(\omega) = i\omega m$, a Hooke's Law spring has $Z_k(\omega) = k/i\omega$, and a dashpot that supplies a force $F = bv$ thus has an impedance $Z_b = b$. The

related concept called the admittance Y is defined by

$$Y(\omega) \equiv Z^{-1}(\omega) = \frac{v_0}{F_0} e^{i\phi}.$$

With these preliminaries, Callen's Fluctuation-Dissipation Theorem can be succinctly stated. The thermodynamic fluctuations analogous to Brownian motion have a magnitude given by the application at the point of interest of a random force with a power spectrum

$$S_F(\omega) = 4k_B T Re(Z). \tag{14}$$

The strength of the applied force power spectrum is proportional to the dissipative (real) part of the impedance; hence the name "fluctation-dissipation" theorem. Note that this expression has the same form as the more familiar power spectrum for the Johnson noise voltage, $S_V(\omega) = 4k_B T R$, where the resistance R is the real part of the electrical impedance. The similarity is not accidental, but is only one example of the many phenomena unified by the theorem.

An alternative form of the theorem, more useful in some situations, directly gives the displacement fluctuation power spectrum instead of the equivalent applied noise force. It states

$$S_x(\omega) = \frac{4k_B T}{\omega^2} Re(Y). \tag{15}$$

Again, the power spectrum scales with the amount of dissipation in the system.

Clearly, this description of fluctuation phenomena is richer than the Equipartition Theorem, since here we have expressions for the entire power spectrum of the fluctuations, not just their rms amplitude. But are the two descriptions even consistent? The rms fluctuation, for example, the expression in Eq. (13), has no dependence on the magnitude of the dissipation. But Eq. (15) shows that the fluctuation power spectrum is proportional at each frequency to the amount of dissipation at that frequency. How can both be true? An oscillator with low dissipation shows a very pronounced peak in its response at the resonance frequency, while one with larger dissipation exhibits a less dramatic peak. So, although the driving noise force is smaller when the dissipation is smaller, the response on resonance is greater. The two effects precisely cancel, as can be verified by direct integration, thus guaranteeing that the integral of the power spectrum Eq. (15) is equal to what one would predict from the Equipartition Theorem.

These two faces of thermal noise, rms magnitude, and power spectrum, are each important in the appropriate context. In a broad-band gravitational wave detector, such as one using an interferometer, the power spectrum carries the most valuable information.

This insight is embodied in the universal choice to suspend the test masses as pendulums. Pendulums are chosen because they are the best way known to create a low frequency oscillator with very low dissipation. Heuristically, most of the restoring force in a pendulum comes from the tension in its wires (due in turn to the gravitational force on the mass); this process has no dissipation associated with it. The only unavoidable dissipation is that associated with the flexure of the wires, but in a properly designed pendulum the fraction of restoring force associated with flexure is small. Hence, the internal friction in the wires is "diluted" by a large factor (perhaps of order 10^3).

Similarly, one wants to minimize the thermal noise associated with internal vibrations of the test masses. This can be achieved only by making the masses out of a material with very low dissipation. Fortuitously, fused silica has very low mechanical dissipation at acoustic frequencies at room temperature.

A standard design rule in those devices is to attempt to place all resonances (such as those associated with the pendulum suspension of the test masses or those involving internal vibrations of the test masses themselves) outside of the frequency band in which signals will lie. When this is done, only the off-resonance amplitude of the power spectrum is important. The off-resonance transfer function of an oscillator to a given force is controlled by the compliance of the resonator in the low frequency limit, and by the inertia of the oscillator above resonance. If the dissipation that sets the driving force can be made low, so can the power spectrum of thermal noise at the frequencies of interest.

4.2 Readout Noise and the Quantum Limit

All experiments need readout and recording systems to register the effects for which we are searching. If the effect is large enough, then these functions can be carried out essentially perfectly. But in the case of the tiny mechanical effects we expect from gravitational waves, even to make the mechanical system's response large enough to record requires very carefully designed readout systems. It is not possible in all cases to ensure that the noise in the readout system is small compared to the mechanical noise in the test masses.

Readout noise has two faces, either one of which may dominate depending on the circumstances. The most familiar is additive noise that competes with a fair copy of the mechanical signal in the output of the measuring system. But measurement systems also unavoidably add mechanical noise to the front end; this "back reaction" noise

must also be kept small if the highest possible precision is to be attained.

The trade-off between additive noise and back-reaction noise is most familiar to physicists from discussions of the quantum mechanical Uncertainty Principle. And, indeed, the Uncertainty Principle governs the ultimate precision of a large class of measurements.

4.2.1 The Heisenberg Microscope as a Prototype Measuring Instrument

It is convenient to make a mental division of a gravitational wave interferometer into two parts. Call the nearly freely-falling mirrored test masses (and the space-time between them) the "system to be measured," and the laser, light beams, and photodetector the "measuring apparatus." There is a deep analogy here with the archetypal quantum mechanical measurement problem called the "Heisenberg microscope." Bohr gave a particularly clear description of it, using a semi-classical treatment. In his 1928 essay "The Quantum Postulate and the Recent Development of the Quantum Theory," Bohr wrote:[16]

> In using an optical instrument for determinations of position, it is necessary to remember that the formation of the image always requires a convergent beam of light. Denoting by λ the wave-length of the radiation used, and by ϵ the so-called numerical aperture, that is, the sine of half the angle of convergence, the resolving power of a microscope is given by the well-known expression $\lambda/2\epsilon$. Even if the object is illuminated by parallel light, so that the momentum h/λ of the incident light quantum is known both as regards magnitude and direction, the finite value of the aperture will prevent an exact knowledge of the recoil accompanying the scattering. Also, even if the momentum of the particle were accurately known before the scattering process, our knowledge of the component of momentum parallel to the focal plane after the observation would be affected by an uncertainty amounting to $2\epsilon h/\lambda$. The product of the least inaccuracies with which the positional co-ordinate and the component of momentum in a definite direction can be ascertained is just given by [the uncertainty relation].

In a gravitational wave interferometer, we are hardly dealing with a microscopic system: the test masses will have masses of 10 kg or more. Yet because we aspire to such extreme precision of measurement, it is crucial to consider the sort of quantum effects usually relevant only for processes on the atomic scale. Note that we are not satisfied to know our test masses' precise positions at one moment only; we want to know the history of the path length difference of the interferometer. Perturbations of the momenta of the masses cannot be ignored, therefore, since the value of the momentum at one time affects the position later.

In the Heisenberg microscope, the phenomenon conjugate to the registration of the arrival of a photon that has bounced off an atom is the recoil of the atom caused by the change in the photon's momentum upon reflection. In a gravitational wave interferometer we register an arrival rate of photons that depends on the difference in phase between electromagnetic fields returning from the two arms. We can recognize the conjugate phenomenon by looking for a fluctuating recoil that can affect the same degree of freedom that we measure. Fluctuating radiation pressure on the test masses causes them to move in a noisy way. The resulting fluctuation in the length difference between the two arms shows how this effect can alter the phase difference between light arriving from the two arms; this identifies it as the conjugate phenomenon.

4.2.2 Shot Noise in an Interferometer

The general principles discussed above can be made much clearer by consideration of specific cases. In an interferometer, all of the physics involved in the fundamental noise limits can be made explicit.

First, consider the question of the direct or "additive" noise, which sets the limit to how small a strain can be recognized. A Michelson interferometer with free mirrors can be described as a transducer from gravitational wave strain to output light power. Part of the reason for its good sensitivity comes from the fact that the output power changes from zero to its maximum and back again with a change in path length difference of only one optical wavelength, of order 1 μm. But it would clearly not be sufficient to only register the difference between one fringe and the next; with a precision of one optical wavelength, even if the arm lengths were as long as useful (of order of half the gravitational wavelength, or 150 km for waves of frequency 1 kHz) one could only register strains of about 10^{-11}.

The key to reaching strain sensitivities of 10^{-21} lies in determining the path length difference to a tiny fraction of a fringe, say 1 part in 10^{10}. Is this possible?

First, recall that the output power is given by

$$P_{out} = P_{in} \cos^2(k_x L_x - k_y L_y). \tag{16}$$

Consider in particular the behavior of the interferometer near an operating point at

which half of the maximum possible power exits the output port. At this point, the change in output power is maximized for a given change in path length difference. If we want to observe a very small change in arm length difference, then we must be able to recognize a very small change in the output power of the interferometer. In other words, the readout precision of an interferometer is limited by the precision with which we can measure optical power.

The fundamental limit to this power measurement is the so-called "shot noise" in the light. We can model the light flux at the photodetector as a set of discrete photons whose arrival times at the photodetector are statistically independent, although with a deterministic mean rate $\bar{n}$. Whenever we count a number of discrete independent events characterized by a mean number $\bar{N}$ per counting interval, the set of outcomes is characterized by a probability distribution $p(N)$ called the *Poisson distribution*,

$$p(N) = \frac{\bar{N}^N e^{-\bar{N}}}{N!}. \tag{17}$$

(This is also colloquially referred to as "counting statistics.") When $\bar{N} \gg 1$, the Poisson distribution can be approximated by a Gaussian distribution with a standard deviation σ equal to $\sqrt{\bar{N}}$.

We are trying to determine the rate of arrival of photons $\bar{n}$ (with units of $\sec^{-1}$), by making a set of measurements each lasting τ seconds. The mean number of photons in each measurement interval is $\bar{N} = \bar{n}\tau$. The Poisson fluctations of the measurement process mean that the fractional precision of a single measurement of the photon arrival rate (or, equivalently, of the power) is given by

$$\frac{\sigma_{\bar{N}}}{\bar{N}} = \frac{\sqrt{\bar{n}\tau}}{\bar{n}\tau} = \frac{1}{\sqrt{\bar{n}\tau}}. \tag{18}$$

This says that if we were to try to estimate $\bar{n}$ from measurements for which $\bar{n}\tau \sim 1$, then the fluctuations from instance to instance will be of order unity. If $\bar{n}\tau$ is very large, then the fractional fluctuations are small.

Let's carry through the calculation for the power fluctuations, and thence to the noise in measurements of h. Each photon carries an energy of $\hbar\omega = 2\pi\hbar c/\lambda$. If there is a power P_{out} at the output of the interferometer, then the mean photon flux at the output will be

$$\bar{n} = \frac{\lambda}{2\pi\hbar c} P_{out}. \tag{19}$$

At the half-power operating point,

$$\frac{dP_{out}}{dL} = \frac{2\pi}{\lambda} P_{in}. \tag{20}$$

We can also consider this to be the sensitivity to the test mass position *difference* δL, since the interferometer is equally sensitive (with opposite signs) to shifts in the length of either arm.

Now consider the fluctuations in the mean output power $P_{out} = P_{in}/2$, averaged over an interval τ. The mean number of photons per interval is $\overline{N} = (\lambda/4\pi\hbar c)P_{in}\tau$. Thus we expect a fractional photon number fluctuation of $\sigma_{\bar{N}}/\bar{N} = \sqrt{4\pi\hbar c/\lambda P_{in}\tau}$. Since we are using the output power as a monitor of test mass position difference, we would interpret such statistical power fluctuations as equivalent to position difference fluctuations of a magnitude given by the fractional photon number fluctuation divided by the fractional output power change per unit position difference, or

$$\sigma_{\delta L} = \frac{\sigma_N}{N} \Big/ \frac{1}{P_{out}} \frac{dP_{out}}{dL} = \sqrt{\frac{\hbar c\lambda}{4\pi P_{in}\tau}}. \tag{21}$$

Recall that we can describe the effect of a gravitational wave of amplitude h as equivalent to a fractional length change in one arm of $\Delta L/L = h/2$, along with an equal and opposite change in the orthogonal arm. The net change in test mass position difference is $\delta L = Lh$, so if we interpret brightness fluctuations in terms of the equivalent gravitational wave noise σ_h, we have $\sigma_h = \sigma_{\delta L}/L$, or

$$\sigma_h = \frac{1}{L}\sqrt{\frac{\hbar c\lambda}{4\pi P_{in}\tau}}. \tag{22}$$

There is no preferred frequency scale to this noise; the arrival of each photon is independent of the arrival of each of the others. Note also that the error in h scales inversely with the square root of the integration time. These facts can be summarized by rewriting Eq. (22) as the statement that the *photon shot noise* in h is described by a white amplitude spectral density of magnitude

$$h_{shot}(f) = \frac{1}{L}\sqrt{\frac{\hbar c\lambda}{2\pi P_{in}}}. \tag{23}$$

4.2.3 Radiation Pressure Noise in an Interferometer

A hint at where quantum mechanics might have some deep relevance comes when we consider how shot noise scales with the optical power used in the interferometer. As shown in Eq. 23 above, the shot noise readout precision improves as the square root of the optical power. Taken at face value, this would suggest that we could achieve arbitrarily good measurement precision, so long as we were able to use an arbitrarily powerful laser to illuminate the interferometer.

That conclusion is encouraging, but ought to cause some unease. After all, didn't Einstein fail to find a way to defeat the Uncertainty Principle's limit to the precision of mechanical measurements?[17] Or did he just blunder not to consider using an interferometer to make the measurements?

Of course not. An interferometer is in fact very much like the "Heisenberg microscope" of Bohr. The lesson we should draw from that example is that we must not neglect the effects of recoil in an interferometer. In effect, an interferometer with free masses is a Heisenberg "macroscope": the role of the small object whose position is to be determined is played by the set of macroscopic test masses. Their large size gives an obvious advantage against recoil, but can't make the effect vanish entirely.

The recoil effect does have one extra bit of subtlety, though. Recall that the measurement we make in a Michelson interferometer has to do with the difference in length of the two arms. So the recoil force we need to calculate is not the common mode force (the component that is equal in the two arms), but only the noise in the difference of the recoil forces in the two arms. If we use a naive picture in which photons independently choose which arm to enter, then the origin of a differential recoil force noise is clear. Each photon enters either one arm or the other; whenever one arm gains a photon, the other loses one.

To estimate the size of this effect, first recall that the force exerted by an electromagnetic wave of power P reflecting normally from a lossless mirror is

$$F_{rad} = \frac{P}{c}. \tag{24}$$

The fluctuation in this force is due to shot noise fluctuation in P. That is

$$\sigma_F = \frac{1}{c}\sigma_P, \tag{25}$$

or, in terms of an amplitude spectral density

$$F(f) = \sqrt{\frac{2\pi\hbar P_{in}}{c\lambda}} \tag{26}$$

independent of frequency.

This noisy force is applied to each mass in an arm. For now, let us consider a simple "one-bounce" interferometer. We will allow the mirrors at the ends of the arms to be free masses, but in this example assume that the beam splitter is much more massive than the other mirrors. The fluctuating radiation pressure from the power $P_{in}/2$ causes each mass to move with a spectrum

$$x(f) = \frac{1}{m(2\pi f)^2}F(f) = \frac{1}{mf^2}\sqrt{\frac{\hbar P_{in}}{8\pi^3 c\lambda}}. \tag{27}$$

The power fluctuations in the two arms will be anti-correlated. The radiation pressure noise is then

$$h_{rp}(f) = \frac{2}{L}x(f) = \frac{1}{mf^2L}\sqrt{\frac{\hbar P_{in}}{2\pi^3 c\lambda}}. \tag{28}$$

4.2.4 The Standard Quantum Limit

Thus we have two different sources of noise associated with the quantum nature of light. Note that they have opposite scaling with the light power-shot noise declines as the power grows, but radiation pressure noise grows with power.

If we choose to, we can consider these two noise sources to be two faces of a single noise that we can call optical readout noise, given by the quadrature sum

$$h_{o.r.o.}(f) = \sqrt{h^2_{shot}(f) + h^2_{rp}(f)}. \tag{29}$$

At low frequencies, the radiation pressure term (proportional to $1/f^2$) will dominate, while at high frequencies the shot noise (which is independent of frequency, or "white") is more important. We could improve the high frequency sensitivity by increasing P_{in}, at the expense of increased noise at low frequency. At any given frequency f_0, there is a minimum noise spectral density; clearly, this occurs when the power P_{in} is chosen to have the value P_{opt} that yields $h_{shot}(f_0) = h_{rp}(f_0)$.

When we solve for P_{opt} and insert it into our formula for $h_{o.r.o.}$ we find

$$h_{QL}(f) = \frac{1}{\pi fL}\sqrt{\frac{\hbar}{m}}. \tag{30}$$

We have renamed this locus of lowest possible noise $h_{QL}(f)$, for "quantum limit," to emphasize its fundamental relationship to quantum mechanical limits to the precision of measurements. Note that the expression does not depend on P_{in} or λ, or any other feature of the readout scheme, even though such details were useful for our derivation. Thus, this examination of the workings of our Heisenberg microscope provides an instrument-specific derivation of the Heisenberg Uncertainty Principle. And it reminds us of the truth Bohr's remarks expressed, that in any measurement the Uncertainty Principle emerges from the specific mechanism of the measurement.

There was a moment when some physicists believed, on seemingly sound physical grounds, that this picture of how photons interact with a beam splitter was so flawed that interferometers could perhaps evade the Uncertainty Principle.[18] The argument can be made based on quotation from quantum mechanical Scripture, Dirac's *The Principles of Quantum Mechanics*.[19] There one can read that photons in an interferometer travel down both arms simultaneously; furthermore, it is written that interference can only take place between a photon and itself, so the very existence of interference in a quantum mechanical world is proof of this picture. If this were taken as absolute and literal truth, then it would appear to rule out any differential radiation pressure at all, since the number of photons, and hence the recoil forces, would be identical in the two arms. Without the resulting differential recoil of the test masses, there is no quantum limit. Gravitational waves could in principle be measured with arbitrary precision. Some physicists defended this as gospel, despite the fact that the argument appeared to use quantum mechanical reasoning to disprove quantum mechanics.

The stubbornly naive were untroubled by this argument, and expected the Uncertainty Principle to hold. Some physicists read a few pages further in Dirac's book, to the passage explaining that allowing the possibility of energy measurements, say by observation of recoil of the mirrors, causes collapse of the wave function in such a way that photons end up either in one arm or the other. (Dirac's first discussion refers to an interferometer with rigidly fixed mirrors.) The learned were saved from error by the work of Caves,[20] who invoked the concept of vacuum fluctuations to explain the quantum mechanical behavior of photons at a beam splitter. A vacuum electromagnetic field with zero-point fluctuations enters the interferometer through the output port; its superposition with the field from the laser causes the light to behave in the way expected from semi-classical reasoning.

4.3 Seismic Noise

We have neglected to consider above another source of noise in gravitational wave detectors that is so common and important as to be essentially ubiquitous. This is what is commonly called seismic noise, the continual shaking of the terrestrial environment due to a variety of contingent causes, ranging from small earthquakes to ocean waves driven by large weather systems to automobiles striking potholes in poorly paved streets. Such a complex phenomenon can have no simple explanation from basic physics, yet dealing with it forms a substantial part of the challenge to designers of gravitational wave detectors. (Only moving the whole detector into space suffices to remove it entirely from consideration.)

At a reasonably quiet location, the spectrum of seismic noise from 1 Hz to several hundred Hz can be approximated as

$$x(f) = \begin{cases} 10^{-7}\mathrm{cm}/\sqrt{\mathrm{Hz}}, & \text{from 1 to 10 Hz} \\ 10^{-7}\mathrm{cm}/\sqrt{\mathrm{Hz}}(10\mathrm{Hz}/f)^2, & \text{for } f > 10 \text{ Hz}. \end{cases} \tag{31}$$

The magnitude of this mechanical noise background is distressingly large. The rms amplitude of the noise over this interval is of order 1 μm. The good news is that the spectrum falls with increasing frequency f. But even so, throughout the range of frequencies of interest to gravitational wave detectors, it involves motions many orders of magnitude larger than would be driven by any conceivable incident gravitational wave. There is no possibility of success unless the effects of seismic noise can be strongly attenuated.

It is straightforward to see the way in which seismic noise mimics a gravitational wave signal in an interferometer. As long as the separation between the mirrors is not very small, then the seismic inputs to each mirror are effectively independent; the difference in arm lengths is driven by the quadrature sum of the noise at all mirrors. The situation is a bit more subtle for a resonant mass detector. If it is suspended at its midpoint, it would appear that its internal modes should not be excited by any motion of the suspension point. However, this argument assumes perfect symmetry of the resonator about the suspension. The approximate symmetry of real systems may give several orders of magnitude of effective isolation, but the seismic spectrum is so large that additional isolation is always required.

Fortunately, the design of seismic isolators is a well-developed art. One can construct mechanical multi-pole low-pass filters that provide outstanding attenuation at frequencies well above those of the filter poles. The art of doing so was introduced to the field of gravitational wave detection by the founder, Weber.[21]

4.3.1 A Simple Two-Pole Isolator

The essence of vibration isolation can be understood using only ideas from freshman physics. Imagine that the object to be isolated has mass m. Assume that it is a rigid body, and that we are only interested in its motion x_m in a single direction. Then we can treat the object as a point mass. If it rests on the ground, it shares the ground's motion x_g, so $x_m = x_g$. To isolate the mass, replace the rigid connection to the ground

with a compliant connection, that is, attach it to the ground through a spring with spring constant k. Then the equation of motion of the mass is

$$m\ddot{x}_m = -k(x_m - x_g). \tag{32}$$

With the usual frequency domain ansatz $x_m(t) = x_m \exp(i\omega t)$, $x_g(t) = x_g \exp(i\omega t)$, Eq. (32) can be solved for the vibration transfer function

$$\frac{x_m}{x_g} = \frac{\omega_0^2}{\omega_0^2 - \omega^2}, \tag{33}$$

where $\omega_0^2 \equiv k/m$. The asymptotic behavior of Eq. (33) reveals the essential features of an isolator—although for low frequencies $\omega \ll \omega_0$, $x_m/x_g \approx 1$ (the mass moves rigidly with the ground), and at high frequencies $\omega \gg \omega_0$, $x_m/x_g \approx \omega_0^2/\omega^2$ (the amplitude of the mass's motion falls steeply with increasing frequency.)

(The transfer function of Eq. (33) has the unphysical feature that it predicts an infinite response at the resonant frequency ω_0. This problem is remedied if we add a damping force to our oversimplified model. Even so, an isolator does have the unfortunate feature that it gives a resonant amplification of the input noise for frequencies near ω_0.)

So, a strategy for isolation is as follows: to isolate an experiment for signal frequency ω, construct an isolator with a low enough resonant frequency ω_0 so that the isolation factor ω_0^2/ω^2 is sufficiently small to reduce the seismically driven value of x_m to a tolerable level.

4.3.2 Isolation Stacks

If it proves difficult to make ω_0 sufficiently low (it often does), another straightforward idea can often be made to work instead. Make two or more isolators with resonant frequencies as low as is convenient, then cascade them with the system you want to isolate farthest from the noisy ground. Solving for the resonant frequencies of the coupled chain of isolators is a normal mode problem that may get complicated. But the high frequency limit of a chain of N isolators, each with resonant frequency ω_0 by itself, has the simple form

$$\frac{x_m}{x_g} \approx \left(\frac{\omega_0^2}{\omega^2}\right)^N. \tag{34}$$

In Weber's original experiment, a vibration isolation "stack" of alternating steel plates and rubber sheets was used. No details are given in the text of his papers, but the idea was clear, and was adopted by those who followed Weber as well. Among the nice features of this isolator are: rough equality of the isolation in all degrees of freedom (since the rubber is compliant in both compression and in shear), and a small level of resonant amplification (since the rubber is rather lossy).

4.3.3 Isolation for Interferometers

The pendulum that provides a low-dissipation suspension for the test mass in an interferometer also provides two poles of seismic isolation. Typically, the resonant frequency is close to 1 Hz. This single stage might almost provide enough isolation for gravitational wave measurements in the vicinity of 1 kHz, but certainly requires substantial augmentation for good sensitivity at lower frequencies. This additional isolation needs to be at least roughly isotropic, since there are a variety of cross-coupling mechanisms that can make vertical noise couple to the phase of the light beam.

The simplest way to add isolation is to suspend the pendulum from a stack of the sort used to isolate a resonant-mass detector. Since interferometers are usually aimed at having good sensitivities down to the lowest possible frequencies, there is a strong incentive to make a stack with the lowest possible resonant frequencies.

More dramatic re-engineering of the basic stack concept can yield greater rewards. The strongest effort in this direction has been made by the VIRGO group in Pisa.[22] The basic idea is to build a stack based on many cascaded pendulums, but to build enough vertical compliance into each pendulum stage to give effective vertical resonant frequencies comparable to the 1 Hz resonances of the pendulum stages. In its original version, the vertical springs were air springs made of flexible bellows. This yielded impressive performance, which was improved even further when the vertical compliance was enhanced by magnetic anti-springs arranged to cancel some of the vertical stiffness. The original version would have made seismic noise negligible (compared with shot noise) for all frequencies above 10 Hz, while the enhanced version would work down to 4 Hz. The air springs did suffer from one serious drawback; the temperature coefficient of the spring force is determined by the ideal gas law, a much stronger sensitivity than elastic springs. Drifts proved very hard to control. To remedy this problem, the system was completely redesigned, with the air springs replaced by pre-stressed steel cantilevers in the form of narrow triangular blades, supplemented with magnetic anti-springs.

4.4 Noise in Resonant-Mass Detectors

Key features The essential complication in understanding resonant-mass detectors (as compared to interferometric detectors) is that the degree of freedom of interest is that of a simple harmonic oscillator (or a collection of them, as we'll see in the next section). So in addition to any intrinsic frequency dependence in the noise, there is a deliberately constructed resonant transfer function in the detector itself. As we saw earlier in this review, the resonance was introduced as part of a strategy for overcoming wide-band noise in the amplifier.

The use of this strategy involves different heuristic concepts than are appropriate for interferometers. In particular, optimizing the sensitivity of a resonant detector to short bursts almost always involves choosing to average the output over times that are long compared with the length of the burst itself. Then, the measurable quantity is no longer $h(t)$, but is instead net change in the vector amplitude (magnitude and phase) of the resonator's oscillation. This in turn can be expressed in terms of the energy that the wave would have deposited in a resonator at rest.[23] If the gravitational waveform $h(t)$ has a Fourier transform $H(f)$, then that excitation energy E is (for an optimal orientation between bar and wave)[24]

$$E = \frac{M v_s^4}{L^2} \left| H\left(f_0\right) \right|^2 , \tag{35}$$

where M is the total mass of the bar, v_s is the speed of sound in the material, L is the overall length of the bar, and f_0 is the resonant frequency.

The distinctive features are twofold: characterization of all candidate events by a single number (usually its "energy" or else $T \equiv E/k_B$), and a signal-to-noise optimization that involves choosing the right averaging time (or bandwidth).

Resonant transducers The second generation of resonant-mass detectors replaced Weber's piezoelectric transducer with a kind of a bridge circuit, in which the mechanical motion unbalanced the bridge by modulating the inductance or capacitance of one leg of the bridge. As with piezoelectric transducers, achieving a high level of coupling has proven difficult to achieve. A standard measure of the coupling is the Gibbons-Hawking parameter β, defined as "the proportion of elastic energy of the detector that can be extracted electrically from the transducer in one cycle."[25] In principle, the excitation of the bridge could be increased without limit, but in practice large fields usually lead to excess dissipation in the transducer even before electrical breakdown occurs. Transducers have been limited to working values of β of around 10^{-2}.

A heuristic way of understanding the design problem is to think of the issue as an attempt to design a transducer that makes a reasonable electrical impedance at its output appear to the mechanical system as a mechanical resistance sufficient to supply appreciable damping to the bar. With bar masses in excess of one ton, this may seem inordinately difficult. The good values mentioned in the previous paragraph avoided this problem by making use of so-called resonant transducers, which have been adopted almost universally since the idea was proposed by Paik in 1976.[26]

Paik's design called for a smaller mechanical resonator to be attached to the main resonant mass M. The resonant frequency of the smaller resonator itself (i.e., with the larger resonator "clamped") is chosen to match that of the main resonator. The actual coupled system then has two normal modes. If the mass ratio $m/M \equiv \alpha \ll 1$, then it is easy to show that the ratio of the amplitude of motion of the small mass, compared with that of the main resonator, is

$$\left| \frac{x}{X} \right| = \frac{2}{\sqrt{\alpha}}.$$

When a gravitational wave burst interacts with such a resonant system, it will at first mainly excite the vibration of the large bar. (The Paik resonator is a small device at one end of the bar, so the gravitational wave strain has only a small baseline for creating a stretch in its spring.) The free motion of this two-mode system then exhibits "beats," during which the mechanical energy of the main resonator's original motion is transferred into excitation of the small resonator. During this phase of the beat cycle, the effect of the gravitational wave has been transformed into a motion $2/\sqrt{\alpha}$ times larger than in a detector without the resonant transducer. The electro-mechanical transducer is mounted so as to measure the motion of the small mass with respect to the end of the main resonator, thus presenting this larger motion to the rest of the signal processing system.

The advantage this offers in detecting weak signals is probably obvious. The larger motion generates a comparably larger electrical output from the transducer, reducing the importance of a given level of electrical amplifier noise. Another way of seeing the advantage is to recognize the much smaller mechanical impedance required to damp the motion of the smaller mass, which means that β is increased by a factor of order α^{-1}. In present day designs, the mass ratio α is typically of order a few times 10^{-3}. The value is set in an optimization that involves not only thermal noise and additive amplifier noise but the back-action noise as well.

4.4.1 Thermal Noise in Resonant-Mass Detectors

Even though resonant-mass detectors work in a rather different way from free-mass detectors, they still benefit from low levels of dissipation. Here, one makes exactly the opposite design choice in regard to the resonant frequency—rather than trying to put all resonances outside the signal band of interest, the whole detection strategy is based on placing the resonance at a frequency where one expects signal power to be large. So, on the face of it, all of the $\frac{1}{2}k_BT$ of energy contributes to the rms scatter of the instrument's output. But it would be too hasty to conclude that there is no dependence of the effective thermal noise on the dissipation level. The typical resonant-mass detector can have a damping time of order 1000 sec, while the resonant frequency is near 1 kHz (i.e., a quality factor Q of order 10^6). The signal being sought might be a brief burst of radiation, lasting only a millisecond. In this brief interval, the bar's amplitude and/or phase of oscillation are altered. The characteristic time for such changes due to thermal noise is the damping time. So only a small fraction of the thermal noise power is relevant in the detection process. For this reason, low dissipation can be as valuable in resonant-mass detectors as in interferometers.

To make the analogy with free-mass detectors clearer, we can make a frequency domain version of the argument in the previous paragraph. It is most convenient to consider the comparison between the thermal noise driving force, Eq. (14), and the force exerted by the gravitational wave on the detector. A brief burst of gravitational waves will have a broad spectrum; its matched filter will admit power over a wide band. The thermal noise force power spectrum that may obscure the effect of the gravitational wave is broad-band as well; if the dissipation in the bar has the form of velocity damping, then the spectrum is white. The amplitude of this broad band noise spectrum is proportional to the amount of dissipation. So the signal to noise ratio will be better, the smaller is the amount of dissipation.

An expression for the minimum detectable noise $\Delta E_{\rm min}$ (at the 1 σ level) is[27]

$$\Delta E_{\rm min} \approx 2k_BT\frac{\tau_s}{\tau_d}, \tag{36}$$

where T is the temperature, τ_s is the averaging time, and τ_d is the damping time of the bar. In this case (where we assume only thermal noise is important), it is clear that the shorter the averaging time the better the sensitivity.

4.4.2 Readout Noise in a Resonant-Mass Detector

Most of the issues of readout noise discussed above for interferometers were first faced in the context of resonant-mass detectors. We have inverted the historical order solely for pedagogical reasons; the physical origin of readout noise in an interferometer is transparent, and nominally free masses make a more straightforward analogy to the Heisenberg microscope than does the complex amplitude of a resonant mode. A key role was played by Braginsky[28] in calling attention to the necessity of understanding these issues in resonant-mass gravitational wave detectors.

Transducer model The transducer is treated as a black box, in the form of a two-port network, emphasizing the fact that the system has a single input port and a single output port. Unlike the two-ports usually studied in electrical engineering classes, the transducer changes the dimensions of the signal: a mechanical motion at the input is converted to an electrical signal at the output.

One such device is a seismometer pickup: a permanent magnet attached to the end of the bar, arranged so that it moves back and forth in the vicinity of a conducting coil whenever the bar vibrates. Motion of the mechanical system generates an electromotive force in the coil that can be electrically amplified and recorded.

In gravitational wave detectors, the transducers usually involve a reactance (i.e., an inductance or a capacitance) whose value is modified by relative motion between the transducer mass and the end of the bar. This generates an electrical signal by virtue of the fact that the reactive component is placed in one leg of a bridge circuit. This kind of transducer is a passive device. It can add noise if its level of dissipation is large enough so that thermal noise is important.

There has also been development of parametric transducers, in which a balanced bridge (two counter-varying reactances) is excited at microwave frequencies. This is the kind of transducer used on the UWA detector.[29] There is a deep analogy between this kind of transducer and an interferometer.[30] This means that there are both extra benefits (gain in the transducer) and extra complications (such as noise in the oscillator that excites the bridge.)

Pre-amplifier The remaining essential part of the readout system is the low noise pre-amplifier, whose job is to transform the tiny electrical output of the transducer into a signal large enough so that it can be recorded. One needs the noise to be as low as

possible. To describe the noise, and to understand how it affects the detection process, it is valuable to consider a general black box representation of a noisy amplifier as a kind of a two-port network, just as we did for the transducer. Other than the trivial difference that this two-port is an all-electric device, there are two key differences between this kind of two-port and the model transducers discussed above. One is the fact that the amplifier has gain, i.e., it can supply more energy at its output than is supplied by its input. The other is the existence of two sources that represent the generation of noise. They are usually represented as a voltage source and a current source at the input of the two-port network; this is especially convenient, and has wide generality, but is only one of several equivalent ways of representing the noise. More on the general theory of noisy two-port networks can be found in the pioneering paper by Rothe and Dahlke.[31]

It is interesting to pause to inquire why two noise sources are necessary. Recall the Helmholtz theorem, often known in specialized forms as Thévenin's theorem or Norton's theorem.[32] The essence of the theorem is that an arbitrary network of sources and passive components can be represented, as far as its behavior at a given port is concerned, by a single source and a single impedance. But we are dealing here with a network in which two ports are relevant. At each port, one needs an impedance and a source, or their equivalents elsewhere in the circuit. It is traditional to replace the source at the output with an equivalent noise source at the input, whose strength is smaller than the output noise by a factor of the amplifier gain.

These two noise sources play different roles in the measurement process. There is one noise source that is physically present at the input, causing an influence on the system (here the electromechanical transducer) that is attached to the amplifier input. In the jargon of gravitational wave detection, this noise is responsible for "back action," since noise at the input of the pre-amp is thereby applied to the output end of the transducer,where it can cause a mechanical noise force at the transducer input; this is in turn attached to the resonant-mass detector proper. More on this below.

The second noise source (the one replacing the output noise source) is usually referred to as "additive noise": it is added to the amplified signal by the time it appears at the output, without causing any physical effect on the system hooked up to the input.

Amplifier noise in resonant-mass gravitational wave detectors, and the "amplifier limit" In contrast to our discussion of detection strategies in the case where thermal noise dominates, here we discuss the case when additive amplifier noise is the only important noise source. Then, we can best search for a brief burst of gravitational radiation by performing a cross-correlation between the system output and a template consisting of a sinusoid at the mechanical resonant frequency that is damped with the same time constant as the resonance itself. In other words, we look for responses that look like the test mass system suddenly set into resonance. The signal can arrive with any phase of course, so we need to keep track of both sine and cosine components with the bar's damping time. A two-phase lock-in amplifier can be set up to perform exactly this form of averaging.

In this case the energy sensitivity of the detector is given by[27]

$$\Delta E_{\rm min} \approx k_B T_n \frac{\lambda + \lambda^{-1}}{\pi \beta \tau_s f_0}, \tag{37}$$

where T_n is the noise temperature of the amplifier, λ is the ratio of the transducer output impedance to the amplifer noise impedance, β is the Gibbons-Hawking coupling parameter, and τ_s is the averaging time.

Consider the post-detection bandwidth implied by this prescription. The output of the cross-correlation described above is hardly affected if we displace the template with respect to the signal time series by one or even several cycles of oscillation. For there to be a substantial change in the value of the cross-correlation, the template must be displaced by a duration of order the damping time of the mechanical resonance. This means that, if the signal-to-noise ratio is not large, the arrival time of the impulsive gravitational wave signal will be uncertain by of order the damping time. In other words, the post-detection bandwidth Δf of such a signal extraction system is narrow, of order

$$\Delta f \approx 1/\tau_d.$$

For a quality factor of 10^6, this bandwidth is very narrow indeed. Increased bandwidth could be achieved, of course, at the expense of the signal-to-noise ratio, by averaging the output time series for a shorter time than the bar's damping time τ_d.

Combined optimum in the presence of thermal noise and additive amplifier noise When both kinds of noise are present at substantial levels, the best strategy is neither the rapid readout appropriate to thermal noise nor the long averaging time that would be best for amplifier noise. A broad-band output filter will admit too much amplifier noise. Using the narrow-band prescription appropriate to the amplifier-dominated case filters out much thermal noise, but also most of the signal power. Obviously, the optimum in the combined noise case lies somewhere between the extremes, where the net

noise contributions from thermal noise and amplifier noise are equal. Then, the energy sensitivity of the detector is given by[27]

$$\Delta E_{\min} \approx 2k_B T \frac{\tau_s}{\tau_d} + k_B T_n \frac{\lambda + \lambda^{-1}}{\pi \beta \tau_s f_0}, \tag{38}$$

where the averaging time is now to be chosen in such a way as to minimize the sum of the two terms.

The larger is the transducer's coupling β, the smaller is the importance of amplifier noise in units of gravitational wave strength. In the limit of large β, the bandwidth of the resonant-mass gravitational wave detector can approach that given by the thermal noise limit, where the matched template is the signal waveform itself.

For detectors with resonant transducers, bandwidth is also limited by the splitting between the coupled modes of the bar-transducer system. There is a competition between two effects: a small transducer mass gives a large mechanical amplification but a small mode splitting. In the time domain, this would be described as a long beat period; after a signal excites the bar one has to wait for a substantial fraction of the beat period for energy to be transferred to the motion of the transducer mass. This competition could be evaded by the construction of multi-mass transducers with a gradation from intermediate to tiny masses.[33]

Resonant-mass gravitational wave detectors built to date have almost always been substantially under-coupled to their amplifiers. First-generation bars of Weber's type, in which the transduction was performed by piezoelectric transducerss glued around the belly of the cylindrical mass, could only achieve coupling factors of order 5×10^{-6}. One variant constructed by the Glasgow group did achieve a large β.[34] It consisted of a cylindrical bar that had been split in half, then reassembled with a layer of piezoelectric material in effect forming the spring attaching the masses at either end. This design is the closest approach ever to Weber's original vision of a single massive piezoelectric mass as an antenna. In spite of these impressive parameters, the Glasgow bar was never even the most sensitive gravitational wave detector of its day, because the poor mechanical Q of the piezoelectric material gave a large amount of thermal noise.

4.4.3 The Quantum Limit in Resonant-Mass Detectors

One of the key features of Rothe and Dahlke's model of a noisy two-port network is the necessity to describe the noise with two generators, one directly limiting measurement precision and the other disturbing the system being measured. This is reminiscent of concepts most students of physics learn about in the context of quantum mechanics. In fact, quantum mechanics sets a minimum level of amplifier noise that can be expressed in terms of this pair of noise generators.

Quantum mechanical noise in an electrical amplifier A paper by Heffner in 1962 derived the quantum mechanical limit on amplifier noise;[35] its relevance for the gravitational wave detector problem was discovered by Giffard in 1976.[36] What Heffner showed was that, unless an electrical amplifier had a minimum level of noise, it could be used to make measurements of an electrical signal at frequency f that evaded the Uncertainty Principle

$$\Delta n \Delta \phi \geq \frac{1}{2},$$

where Δn is the uncertainty in the number of photons at the frequency f, and $\Delta\phi$ is the uncertainty in the measurement of the phase of the signal.

The amplifier noise that just barely allows measurements at the Uncertainty Principle limit is given by

$$\sqrt{S_v(f) S_i(f)} = \left(\ln \frac{2 - 1/G}{1 - 1/G} \right)^{-1} 2\pi\hbar f.$$

In the limit of large gain G, this limit becomes

$$\sqrt{S_v(f) S_i(f)} = \frac{1}{\ln 2} 2\pi\hbar f. \tag{39}$$

The combination of voltage and current noise power spectra represents yet another pair of conjugate variables governed by an uncertainty relation. This product is proportional to the noise temperature T_n of the amplifier, defined by $k_B T_n \equiv \sqrt{S_v(f) S_i(f)}$.

In the earliest days of the search for gravitational waves, the possible implications of such an amplifier sensitivity limit for the sensitivity of gravitational wave detectors were ignored. Neither Weber's papers nor the canonical theoretical analysis by Gibbons and Hawking mention such a limit. Eventually, cooler heads came to contemplate the fact that gravitational wave detectors ought to obey the Uncertainty Principle.

The electro-mechanical Uncertainty Principle Giffard gave the adaptation of this argument to the electro-mechanical case. His argument is so brief that it is almost hard to state. He points out that the functions of transducer and pre-amplifier together can be thought of as a so-called "mechanical amplifier," a two-port network whose input terminal receives a mechanical signal, but whose output is electrical.

From this, Giffard shows that the Uncertainty Principle requires that unless a gravitational wave signal has a minimum size, no linear gravitational wave detector will be able to register its arrival. He expressed the minimum size in terms of the energy U_s that the wave would deposit in a resonant detector initially at rest. The quantum mechanical limiting sensitivity is

$$U_s > 4\pi\hbar f.$$

As worded above, this argument can seem rather abstract. A heuristic description makes it as vivid as any of Bohr's *gedanken* experiments. In a mechanical amplifier, a crucial quantum mechanical role is played by the back action from the noise source that the Helmholtz Theorem places at the amplifier input. The force noise generator at the mechanical amplifier input (caused by electrical noise acting backwards through the transducer) perturbs the delicate mechanical system, here the resonant-mass system itself. Just as in the Heisenberg microscope, any design trade-off made in an attempt to reduce the position noise ends up increasing the momentum impulse applied to the system being measured.

5 History of Resonant-Mass Detectors

We now turn from a discussion of physics *per se* to a review of the way in which one assembles working gravitational wave detectors in light of the physical principles governing them. We will take a quasi-historical framework for this discussion, as a pedagogically sensible way of grappling with the issues involved. In this section, we will take the chronologically-motivated choice of treating resonant-mass detectors first; then we will start from scratch the overlapping history of interferometric detectors.

By 1966, Joseph Weber had constructed a complete working detector, and by 1968 was reporting coincident observations between detectors separated by 1000 km. The detector contained versions of every essential feature in resonant-mass gravitational wave detectors today, except for the facts that it operated at room temperature and that it used non-resonant strain transducers for its readout. The story of the development of the field since then can be seen as embodying a few key accomplishments: replication of Weber's detectors accompanied by a failure to confirm his claimed detection, clarification of the optimum way to detect gravitational wave signals in a noisy detector and of the sorts of technological developments that could lead to improved detector sensitivity, and the staged implementation of new generations of detectors embodying the improved technology.

5.1 Weber's Detectors as Gravitational-Wave Detection Systems

All of the important elements that make up a working gravitational wave detector are described in Weber's 1966 *Physical Review Letters*.[21] Looking at the signal chain from the front end, we see first the large $1\frac{1}{2}$ ton aluminum cylinder whose fundamental longitudinal mode at 1657 Hz interacts with any incoming gravitational wave. Around its midsection are glued the quartz transducers that give, through the piezoelectric effect, electrical signals proportional to the strain in the aluminum cylinder. Signal leads from those transducers pass through acoustic filters and through the wall of the vacuum chamber, then are connected via another acoustic filter to a superconducting inductance that serves as a "tank circuit" at the input of the low-noise pre-amplifier (whose noise temperature is 50 K.) The output of the pre-amp is connected to further amplification. There follows a rectifier for generating a positive-definite signal proportional to the power out of the amplifier. The end of the signal chain is a recording device, which in 1966 consisted of a pen-and-ink chart recorder.

Another aspect of the detector is the means used to prevent its being excited by influences other than gravitational waves. Isolation against mechanical influences in the form of acoustic or seismic noise is shown clearly in the diagram. Direct acoustic excitation is prevented by the placement of the key parts of the experiment inside a vacuum chamber. Transmission of vibration through the signal leads is attenuated by the acoustic filters mentioned above. The path for vibrations from the floor must pass through a pair of isolation stacks each consisting of three stages of rubber pad/steel block isolators; the aluminum bar is further isolated by a pendulum suspension consisting of a single wire sling that supports the bar about its middle. The top ends of the wire are attached to a beam that spans the space between the isolation stacks.

Further progress in reducing the sensitivity of his bar to spurious external influences enabled Weber to make his 1969 claim of "Evidence for discovery of gravitational radiation".[8] One aspect of this progress consisted of augmenting the electromagnetic shielding of his devices, after tests revealed some sensitivity. A seismometer array was also used to check for correlations between strong vibration of the ground and large detector outputs. But by far the most important system element in this regard was the construction of multiple copies of the complete detector system, and their deployment at spatially separated locations. The network in 1969 consisted of one bar of 66 cm

diameter and 153 cm length (for a resonant frequency of 1657 Hz) located at the Argonne National Laboratory in Illinois, and five other resonant-mass detectors located at the University of Maryland. One bar at Maryland, whose properties matched closely the Argonne bar, was used together with it to look for coincident excitation. Some use of the outputs of the additional bars was made, but the procedure is poorly described.

5.2 Weber's Observational Program

Weber's ongoing research program was concerned with looking for further evidence that the coincident events were real (i.e., not statistical accidents) and that they were in fact due to gravitational radiation. The *Physical Review Letters* of 9 February 1970[37] discusses the statistics from several interesting points of view. The claim that the events are real is buttressed by calculation of the rate of accidental coincidences. For the whole range of threshold levels considered, the number of coincidences exceeds the expected accidental rate by about an order of magnitude. Another even more convincing piece of evidence comes from carrying out a search for coincidences between the output of the two detectors using a scheme in which one of the detectors' outputs was delayed by two seconds, enough to remove any correlation due to excitation by gravitational wave impulses. Just as one would hope, only a small number of coincidences occurred between the offset data streams, at a level consistent with accidentals.

Still another convincing piece of evidence was revealed in the *Physical Review Letters* of 20 July 1970.[38] In it, Weber uses the classic astronomical test of testing for correlation of his event rate with sidereal time. Jansky's claim for the extraterrestrial origin of his observed radio noise had been buttressed in just this way; more recently, Jocelyn Bell had convinced her reluctant thesis advisor Hewish of the extraterrestrial nature of her pulsar signals by a similar test. Weber took advantage of the fact that the response of a cylindrical bar is anisotropic, and that the axes of his bars were oriented East-West, so that their sensitivity pattern was swept across the sky by the rotation of the Earth. The histogram of events versus sidereal time reveals a substantial anisotropy (of order of a factor of two) that peaks when the right ascension of the galactic center crosses the meridian at the mid-point between Maryland and Argonne, and again 12 hours later (as would be expected from the symmetry of the antenna pattern and the transparency of the Earth to gravitational waves). A similar histogram plotting coincidences against the solar time of their occurrence is flat within statistical errors.

One disturbing consequence follows from this observation, if one adopts the natural hypothesis that the sidereal anisotropy indicates that gravitational wave events are originating in the galactic center. Making other reasonable assumptions about the bandwidth of the signals, Weber is led to the conclusion that perhaps as much as 1000 $M_\odot$ per year is being converted into gravitational waves. It is not clear what to make of the inference that the galaxy would be entirely consumed in much less than a Hubble time at such a rate of gravitational luminosity. Weber mentions as possible ways out anisotropic emission from the sources (or focusing of the waves) in such a way that we see higher than average flux, or the possibility that his detector is not working in a linear regime, but is instead being stimulated by very weak gravitational waves to release energy stored in metastable states.

5.3 Early Response to Weber's Claims

Weber's claims attracted a great deal of attention. The experimental evidence appeared very strong. The problems with the energetics of the events could be taken as a strong counter-argument, but they could also be interpreted as a demonstration that a truly remarkable discovery was being made. The review "Gravitational-Wave Astronomy" by Press and Thorne in 1972's *Annual Reviews of Astronomy and Astrophysics*[39] gives testimony to the significance that was attributed to Weber's claims, with its extensive bibliography (completed in December 1971) of papers grappling both with the theoretical consequences of Weber's claims and with experimental ways to improve upon the sensitivity of Weber's detectors. The perceived stakes in the controversy are shown by the three-paragraph introductory section, which is mostly devoted to balancing the possibility of breakthroughs in the making against the alternative that Weber's results are mistaken. This first section concludes with a paragraph, remarkable in a review article, questioning the very urgency of its subject:

> We (the authors) find Weber's experimental evidence for gravitational waves fairly convincing. But we also recognize that there are as yet no plausible theoretical explanations of the waves' source and observed strength. Thus, we feel we must protect this review against being made irrelevant by a possible "disproof" of Weber's results. We have done this by relegating to the end of the article (Section 6) all ideas, issues, and discussions that hinge upon Weber's observations.

Both the theoretical speculation about sources and the quest for better ways to detect the waves are in evidence in one of the most influential papers of this or any other

epoch in the search for gravitational waves. Gibbons and Hawking, both renowned for work rather far removed from experimental physics, wrote in late 1970 the very practical "Theory of the Detection of Short Bursts of Gravitational Radiation."[25] The introduction presents Weber's results as established facts. Section II of the paper discusses possible sources, discussing with equanimity but at some length the extreme gravitational luminosities required. The rest of the paper is devoted to the theory of detecting weak gravitational wave bursts in the presence of noise, and to a clear and original set of proposals for maximizing the signal-to-noise ratio of detectors. Some of one's surprise at the authorship of this paper may be alleviated upon reading the thanks for insights and ideas given (at two places in the text and in the Acknowledgment) to P. Aplin of Bristol University, a very original experimentalist who published little on this subject under his own name.[40]

Gibbons and Hawking pointed out that Weber's own treatment of the theory of gravitational wave detectors, written before he started his observations, had been aimed at their response to steady sinusoidal signals. As such, it gives a misleading idea of the value of the high-Q resonance that characterizes Weber-style detectors. This lack was remedied by Gibbons and Hawking, who (with thanks to Aplin) point out that the low dissipation of a high-Q system means that the level of thermal noise power is low. They go on to show that the high-Q resonance is also of crucial importance in minimizing the effect of Johnson noise in the transducer, here playing the role we would generally call amplifier noise. (A weakness of this light treatment of amplifier noise is that it leaves out the back action effects that enforce the Uncertainty Principle.)

As we saw previously, Gibbons and Hawking noted that there is in fact a competition between these two benefits of low dissipation. To minimize thermal noise, one wants to integrate the output for as short a time as possible, to give the random walk of the resonator's complex amplitude the least opportunity to mask a signal. On the other hand, the importance of transducer/amplifier noise is minimized by integrating for as long as practicable, so that the gravitational wave signal competes with as small a bandwidth of the broadband noise as possible. Gibbons and Hawking showed how to derive the optimum averaging time that minimizes the total noise from these two sources. In so doing, they noted that the averaging time depends on the dimensionless coupling parameter they called β, the definition of which we gave above in Section 4. For Weber's detector they give an estimate of $\beta \approx 5 \times 10^{-6}$.

Gibbons and Hawking go on to note that a large value for β would have two benefits: improving the signal-to-noise ratio by making the gravitational wave signal appear as a larger electrical signal, while simultaneously changing the balance between thermal noise and electrical noise in the direction that causes the optimum sensitivity to be obtained with shorter integration times. In other words, a larger β would yield better sensitivity and high bandwidth. To obtain these benefits, they discuss a novel configuration, proposed by Aplin, that has come to be known as a "split bar." It consists of two large masses (two "ends" of a bar split in half) joined to each other by connection to either face of a layer of piezoelectric material. The benefit comes from the fact that, in this configuration, the piezo is actually functioning as the dominant spring in the system; by storing the bulk of the elastic energy, it is able to produce a larger amount of electrical energy. (In many ways this harks back to Weber's original proposal to make the bar entirely of piezoelectric material.)

Gibbons and Hawking sketch the details of a detector of this sort, in which lead zirconate titanate (PZT) is substituted for crystalline quartz (used by Weber) because of its larger piezoelectric coupling constant. Then, in spite of the fact that the thermal noise power is increased because the piezo is a rather lossy spring, the sensitivity should be increased by more than a factor of ten in energy compared with Weber's detector. At the same time, the optimum sampling time is shortened to 1 msec, so that more detailed information can be extracted from the signal.

In passing, Gibbons and Hawking also note that Weber uses a less than optimal way of searching for gravitational wave events. His definition of an event is a noticeable increase in the energy in the bar's fundamental mode. But a gravitational wave impulse will only increase the energy if it arrives with a particular phase relationship to the bar's previous excitation. If the wave arrives with a different phase, the bar's energy may be decreased, or the effect may instead primarily change the phase of the bar's vibration. They estimate that this means Weber saw only about 1/4 of the events exciting a given bar. And, since the two bars being used for a coincidence have independent phases, only $(1/4)^2 = 1/16$ of the detectable coincidences would have been registered by Weber's technique. This makes the question of the source of the gravitational luminosity that much more difficult to resolve. But it also means that, if Weber's results were real, even more events should be detectable.

Whatever mysteries there may have been regarding Weber's claims that he was detecting pulses of gravitational waves, if they were true they represented one of the most important astronomical and physical discoveries of the 20th century. So it is no surprise that a number of other workers decided to construct gravitational wave detectors. And quite naturally given Weber's apparent success, most of these detectors were built quite

a bit like Weber's. Success looked as if it should come quickly. Although it had taken Weber about a decade, working alone, to design, build, debug, and operate his detectors, those who wanted to follow the recipe he had developed needed much less time. As Weber himself described it:[41]

> The scheme that works is not very difficult to set up and costs about $25,000, excluding pay for the senior physicists involved. Once set up, the instrumentation has run without problems for a little longer than a year and a half. Furthermore, the instrumentation requires a skilled technician about a month to construct. It requires a physicist about two days, starting with an open vacuum chamber to put crystals on the apparatus and make the necessary adjustments to see coincidences. In view of the relative simplicity of the working instrumentation, it seems reasonable to suggest that others start in this way.

A group at Frascati[42] built a detector as close to Weber's as possible, including a close match with his resonant frequency of 1660 Hz. The Munich group of Billing and Winkler[43] also built a quite similar bar. At Moscow State University, a group led by Braginsky constructed another bar that matched Weber's rather closely. At Bell Labs, Tyson[44] began a program of construction along Weber's lines, with less concern to match all of his parameters exactly—the resonant frequency of his largest bar was 709 Hz. At IBM Labs, Garwin and Levine[45] built a small bar of Weber's style; they were convinced that they could improve upon Weber's statistical methods sufficiently to make up for poorer noise levels. Aplin,[40] on the other hand, began to construct a split bar along the lines he had suggested to Gibbons and Hawking. Although he did not succeed in carrying that work through to completion, the idea (and some of the equipment) was taken up by Drever's group in Glasgow.[34] There were also efforts started at Reading and Regina, and in addition, the second generation detectors of much greater sensitivity were planned by groups at Stanford, LSU, and Rome. (More about these later.)

As Weber had predicted, results began to become available very quickly. Unfortunately, it also soon became apparent that no one else was seeing results that looked anything like Weber's. By the time of the 6th Texas Symposium on Relativistic Astrophysics (held in New York December 18-22 1972),[41] several groups were ready to come forward with claims of null results at sensitivities comparable to or better than Weber's. Tyson made the strongest claims for the unreality of Weber's signals, based both on his own failure to find them in his single detector (at sensitivities better than Weber's) as well as his claims that Weber's electronics were too noisy even to record Brownian motion in his bar. Kafka, representing the Frascati and Munich groups (who performed observations in coincidence), was able to report only preliminary results (delivered to him by telegram) of five days of observations, but the failure to find any statistically significant number of coincidences was in strong contradiction to their expectations if Weber's claims were correct. Drever also reported negative results from the high-bandwidth split bar at Glasgow; he, however, tempered his claims of conflict with Weber by noting that his detector would only be expected to see Weber's events if they were of a few milliseconds' duration (an assumption that was plausible but by no means certain).

5.4 Resolution of the Weber Controversy

Over the next year and a half, there was an intensive concentration of effort to attempt to resolve the controversy. The groups listed above continued to accumulate data, none of which confirmed Weber's claim of a substantial flux of strong gravitational-wave pulses. New negative results from Garwin and Levine of IBM Labs were reported in a strongly-worded series of articles in *Physical Review Letters*.[45] Their work was carried out in the explicit belief that Weber's results were spurious, and could be shown to be so by the use of a single detector of well-characterized performance, even if that detector was small. (The titles of the three articles are miniature masterpieces of polemical style: "Absence of Gravity-Wave Signals in a Bar at 1695 Hz," "Single Gravity-Wave Detector Results Contrasted with Previous Coincidence Detections," and "New Negative Result for Gravitational Wave Detection, and Comparison with Reported Detection.") The first results to be reported came from a bar with a mass of 118 kg, the second from a 500 kg bar. Compare with the mass of Weber's principal detectors, 1500 kg.

The letters of the IBM group present the statistics of the noise output of their detectors, along with a careful tutorial on the theory of extracting impulsive signals from noise. This latter theme can be thought of as an explication of the signal processing scheme proposed (in rather telegraphic style) in Gibbons and Hawking's 1971 paper. The key idea was not to rely on net increase in the energy of the bar, but to look for sudden changes, of either sign, in either the magnitude or phase of the detector's complex excitation. The thrust of the papers is that if Weber's results were correct, then even their less massive single detector ought to show substantial departures from the

Gaussian distribution of excitation expected on the basis of thermal and amplifier noise alone. The near-perfect Gaussian fit to their data then constitutes an apparent contradiction of Weber's results.

An important argument used to buttress the claim that the IBM detector was well understood was the application of electrostatic calibration forces to one end of the bar, and the successful detection of those events (within the statistical limits set by the signal-to-noise ratio) by the data processing system. Tyson's 1972 presentation to the Texas meeting had previously emphasized the importance of this fundamental practice of experimental physics, as did the remarks of both Kafka and Drever. Levine and Garwin take Weber to task for having failed to use any calibration method, either as a check of his instruments' front ends or of his data analysis procedure.

A more subtle implicit argument against Weber's work is suggested to the reader of these papers by their admirable clarity, as contrasted with the rather Delphic pronouncements that fill Weber's own contributions to *Physical Review Letters*. Levine and Garwin make this explicit at one point when they compare their results with their best guess at how Weber's would be expressed in similar (sensible) units, complaining "We are thus forced to estimate these quantities, while noting that such information is easily obtained by the experimenter and is normally provided in the publication of a detection experiment."[46]

Garwin led a crusade against Weber's claims at the Fifth Cambridge Conference on Relativity (CCR-5), held at MIT on 10 June 1974.[47] Among the topics discussed was 1) an error in the computer program used by Weber to identify coincidences, shown to generate nearly all of the coincidences in the one data tape shared by Weber with other researchers, and 2) the puzzling feature of Weber's histogram of coincidences versus time delay showing a peak at zero delay in only the central 0.1 second wide bin, in spite of the fact that a 1.6 Hz wide bandpass filter was said to be part of the signal processing chain. But the most spectacular event of the discussion was what even those sympathetic to Garwin's cause might have felt was a trick that bordered on unsportsmanlike conduct. Weber had been given data from the detector of Douglass's group at the University of Rochester, to search for excitations in coincidence with Weber's own detectors; Weber reported at previous meetings that he had detected an excess of coincident events at a level of 2.6 standard deviations above the expected chance rate. According to Garwin's account in a letter to the editor of *Physics Today*,[47] "At CCR-5 Douglass revealed, and Weber agreed, that the Maryland Group had mistakenly assumed that the two antennas used the same time reference, whereas one was on Eastern Daylight Time and the other on Greenwich Mean Time." No stronger way can be imagined of impressing the community with the possibility that Weber was able, by some means, to find coincidences among any two data streams, whether the coincidences actually existed or not.

A panel discussion with almost precisely the same cast of characters as that of the 1972 Texas Symposium was staged at the 7th International Conference on General Relativity and Gravitation in Tel Aviv, June 23-28, 1974.[48] The plot, Weber's lonely claims of detections contradicted by the null results of the others, was also unchanged—the only substantial difference is that Weber's critics had had time to carry out more extensive searches and more careful data analysis. By this time, the Bell Labs group had carried out a coincidence run with an identical bar at the University of Rochester, operated by Douglass. The Munich group (which had by then incorporated the previously independent Frascati group) reported on the results of 150 days of coincident observations. Drever gave a report of a more extensive data run, seven months that had concluded in April 1973, yielding only one candidate coincidence; although this event could not be ruled out as a possible gravitational wave detection, neither could it be positively established as such in spite of the low probability that was estimated for it to have occurred by chance. (The detectors were only 50 m apart, and so may have both been driven by some other kind of influence.) In any event, Drever was able to show that the Glasgow experiment did not show the sort of event rate predicted by Weber's experiment, except under rather implausible assumptions about the nature of the individual gravitational wave pulses. Tyson also briefly reported on the negative results from Garwin and from Braginsky.

How did the physics community deal with these contradictory results? This is an almost classic example of attempted replication of an important claim, but with both opposing camps standing firm in their beliefs that their own results were correct. Valuable insight into the difficulties this situation posed to scientists can be found in the work of sociologist of science Harry Collins, who interviewed many of the principal actors during this period. His results are well worth consulting, even though the quotes from the interviews are reported without identifying the individual speakers.[49]

The other key resource in the written record is the transcripts of the open discussions at the 1972 and 1974 panels. Both Kafka and Tyson point out strongly that Weber (usually) uses a far-from-optimal statistical method to look for signals. Tyson also comes close to accusing Weber of fraud; the method by which Weber has deluded himself and others is said to be continual "tuning" of the statistics used to search for

coincidences, with choice of algorithm and threshold being chosen for each data set in such a way as to maximize the apparent statistical significance. (A similar suggestion was also made by Levine and Garwin.[47]) Weber denies this, but some of his remarks appear to support the accusation by using *ex post facto* reasoning to justify particular choices of pulse-detection algorithms: his test of which method is optimal is which gives a larger number of coincidences, and the choice can vary from one data set to another.

For his part, Weber appears to have believed that the results that contradicted his own did not constitute fair tests of his work, since they were not carried out in an identical way. This may strike some readers as disingenuous, since several of the other detectors were quite close copies indeed of Weber's apparatus. But it appears to be an honest reflection of Weber's belief that he was in fact detecting gravitational waves, and if others couldn't see them there must be something subtly wrong with their detectors. To this day, he continues to claim ongoing detections of gravitational wave pulses with his apparatus.

5.5 Beyond the Weber Era

One of the most interesting features of the Panel Discussion at GR7 was the time spent predicting the future of the field. Both Tyson and Drever gave optimistic predictions for future progress; both have been borne out faithfully, although at slower rates than either would have hoped to see. Tyson described the efforts already begun in 1974 (by groups at Stanford, Louisiana State University, and the University of Rome) to construct Weber-style detectors that were cooled by liquid helium to temperatures of a few degrees Kelvin. The obvious benefit, to reduce the Brownian motion noise (with power proportional to $k_B T$), is important. So, too, is the less obvious benefit that much quieter amplifiers are available at low temperature—these are the Superconducting QUantum Interference Devices, or SQUIDs.

Drever's prediction for future progress focused primarily on inteferometry, the other direction that subsequent history has in fact proven fruitful. It is a brief sketch, without any details or attribution, but still remarkably prescient. Drever starts from the premise that Weber's claimed detection of gravitational wave signals is most likely wrong. Then, he says, the point is no longer to try to verify or extend Weber's results, but instead to ask from first principles what might be the best way to probe the possible existence of gravitational waves, over the widest possible range of properties. Not only is improved sensitivity important, but the ability to look for a variety of signal types over a broad range of frequencies. Drever's own belief in this strategy can be inferred from his having adopted the wide-bandwidth Aplin-style split bar, and from having used the Glasgow detectors not only to search for brief pulses but also for a stochastic background of gravitational waves.

But it was difficult to push this strategy much farther, since the thermal noise in this kind of detector was very large; the losses came predominantly from the piezoelectric material that here serves not only as a transducer but as the primary "spring" for the resonant system, and hopes for reducing them would depend on a program of materials research with uncertain prospects at best. (Note that in the original Gibbons and Hawking paper, the split bar is argued for as the quickest way to give a substantial step in sensitivity, with the explicit assumption that since Weber was already detecting gravitational waves in a sub-optimal way, any sizable increment in sensitivity would immediately yield important dividends.) Instead, Drever recognized the kindred broadband sensitivity inherent in the proposal to build interferometric detectors and, with the prospect of placing the test masses very far apart, the possibility of substantially improved sensitivity as well.

Drever expanded on the themes mentioned briefly at GR7 in a graceful review first given as a talk to the Royal Astronomical Society in 1976, published as an article in 1977.[50] By 1978, when *Annual Reviews of Astronomy and Astrophysics* published its second review of the subject "Gravitational-Wave Astronomy," the authors Tyson and Giffard could write a much more mature piece than could Press and Thorne in 1972.[51] The style is more formal and less heuristic. There is less discussion of rather far-fetched possible sources of gravitational waves, but in its place a much more complete treatment of the theory of extracting weak impulsive signals from the output of a resonant detector. After a summary of the history of the work of Weber and of those who followed him, the authors write

> It must be concluded that the interpretation of the Weber events as gravitational wave pulses is erroneous, since there is no corroborated evidence to date either for an excess number of coincident events or any sidereal correlation.

Similar thoughts, more precisely focused, were expressed in a 1978 paper (written too late to be mentioned by Tyson and Giffard) by Kafka and Schnupp, giving the "Final Result of the Munich-Frascati Gravitational Radiation Experiment."[52] They write that

"Although the non-existence [of Weber's pulses] became obvious a long time ago, it still seems appropriate to publish our final negative result, because our experiment was as similar to Weber's as possible, whereas all other coincidence experiments deviated in one way or the other.... Moreover, we think we have set the lowest upper limits obtained by Weber-type experiments over a reasonable long period of observation," spanning 580 useful days of common observations of the two detectors. The main result of this paper is the null result that the statistics of the coincident excitation of the two detectors was just what would be expected from the laws of chance, given the noise levels in the detectors. Without mentioning Weber specifically, Kafka and Schnupp do remark that

> Scanning our whole data, we could, of course, find periods of a few days, for which at some pair of thresholds the number of coincidences was up to more than 3 standard deviations higher than the average over the various time delays. However, the same was true for arbitrary delays, and zero delay did not seem to be distinguished in any obvious way. However, one should not forget: If one searches long enough in our finite sample of data, one must find some complicated property which distinguishes zero delay signficantly from the others. (Again this is true for an arbitrary delay, but with a different property.)

The paper goes on to pay special attention to two periods, totaling 67 days in length, when the operation of the Munich-Frascati experiment overlapped with times for which the Weber group claimed to have detected substantial rates of coincidences with its own detectors. The authors write: "These results do not give the slightest hint of a simultaneous influence on both detectors. If the significant observations reported by Weber's group for these two periods had been due to gravitational radiation of any kind, they should have shown even more significantly in our experiment." The mention of "any kind" of signals refers to the fact that the present authors used not only the vector-difference algorithm that is optimal for short pulses, but also used for these 67 days the algorithm preferred by Weber, which would be more sensitive for very long wave trains that gradually excited the antenna. Kafka and Schnupp conclude this section by remarking that "we do not have an explanation for Weber's observations," although they suggest the possibility that there might have been some undiagnosed electrical feedback from signals on the telephone line from Argonne into the Maryland bar itself.

The final section of the paper compares the likely strengths and rates of gravitational wave signals from core collapse in supernovae with the then current and possible future sensitivities of gravitational wave detectors. In a dramatic figure, they superpose a model of the rate of supernovae at various distances from the Earth on the natural phase space for gravitational wave searches, event rate versus event strength. The authors point out that, even if one were able to improve the performance of gravitational wave detectors of the Weber type to the limit set by the Uncertainty Priniciple (by cooling, improving Q, or whatever other trick), one would still not have the sensitivity to detect events at the rate of several per year or greater. They conclude, "Because of the difficulties arising from this problem and because one would certainly like to measure more details than just the spectral density of pulses, the Munich group decided not to continue with (low temperature/high quality) Weber-type experiments, but rather with a Weiss-Forward type experiment, i.e., a laser-lighted Michelson interferometer."

In spite of the considerations that moved the Munich group to abandon resonant-mass detectors, the groups that had decided in the early 1970s to build cryogenic versions of the Weber bar pushed ahead. A number of strong reasons can be given to justify this strategy, including the dubious value of relying (as the Munich group did) on signal-strength predictions which necessarily must be ignorant of truly novel astronomical phenomena, as well as the belief that evolutionary development is often a more rapid and reliable strategy for progress than a revolutionary approach. And, although progress was slower than the hopes expressed for it in Tyson's 1974 remarks in Tel Aviv, this route did in fact lead to substantial increases in sensitivity well before the interferometric detectors began to catch up.

The first complete operating cryogenic resonant-mass detector was the one built at Stanford University by the group led by William Fairbank.[53] In addition to the obvious reduction of thermal noise by cooling with liquid helium to a temperature of 4.3 K, and the use of the Josephson junction SQUID as a low noise preamp, there was another technical innovation that helped the Stanford bar reach a new level of sensitivity. This was the introduction by Paik[26] of a resonant transducer, tuned to the same frequency as the bar's resonance, mounted on the end of the bar. Both Tyson and Garwin had used end-mounted transducers, but neither realized the advantages that would accrue to the tuned configuration—the ability to better "impedance-match" the mechanical excitation of the bar to the electrical system, thus increasing the coupling parameter β. (See the discussion above.) The Paik transducer represented a new generation in another sense—it made no use of the piezoelectric effect, but instead used the motion of the resonant proof mass in the transducer to modulate the value of an inductance in

the circuit that carried the persistent current through the Josephson junction.

As the first detector to operate at this new level of sensitivity, no coincidence observations were possible, so the first paper contains only the results of single-detector operation.[53] (The group at LSU, led by former Stanfordian W. O. Hamilton, was collaborating with the Stanford group, but was developing several subsystems independently, including an alternative transducer design.) The 1982 paper in *Astrophysical Journal Letters* included a histogram of excitations of the Stanford detector, showing that even without coincidences it set a new upper limit that, at least at large event rates (greater than of order 10^{-2} per day), was orders of magnitude more stringent than that achieved by the Munich group. An even better result was predicted once a second comparable detector would operate in coincidence.

The noise temperature of 20 mK was certainly a milestone, but doesn't entirely characterize the performance of the bar as a detector of rare events. The histogram of excitations recorded during the 74 days of data presented in the paper show a substantial excess of large excitations, beyond what would be expected from purely Gaussian statistics. The cause was unclear, but one possible explanation was acoustic emission within the bar from sudden stress relaxations. Without coincident operation with a comparable detector, gravitational wave pulses could not be ruled out either.

It wasn't until 1986 that coincident operation of all three cryogenic detectors started in the 1970s (Stanford, LSU, and Rome) was finally achieved. The results of that run were less than spectacular, because none of the detectors was operating as well as it might. Indeed, the event rate-event strength plot that characterized this coincidence run was hardly better than the single-detector plot from the Stanford 1981 solo paper.[24]

The problems with the LSU and Rome detectors were soon sorted out; the rms noise levels in these detectors have now fallen below 10^{-18}. Meanwhile, the Stanford bar met an untimely demise when it was irreparably damaged in the Loma Prieta earthquake of 1989. The LSU-Rome axis has carried out coincident observations since 1991. A first look gave a new upper limit on the flux of gravitational wave pulses that was presented in an otherwise unpublished conference report.[54] Another cryogenic detector that had been under development for some time, the niobium bar of Blair's group in Perth, recently achieved a comparable sensitivity. Coincidence observations continue as this review is being written.

The 1986 coincidence run was a consummation of years of work, whose occurence had been devoutly wished for by many friends of the field for some time. Aside from the fact that it may have been somewhat too late (from the point of view of Stanford,

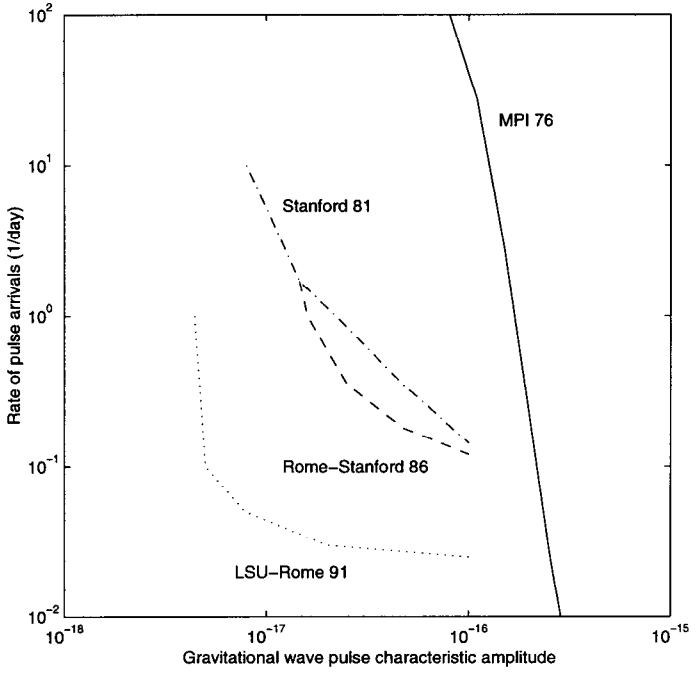


Fig. 3. Upper limits set by various experiments on the rate of gravitational wave impulse arrivals, as a function of the characteristic amplitude of the impulse.

whose detector wasn't working as well as in 1981) or somewhat too early (from the point of view of LSU or Rome, who had not yet shaken all of the bugs out of their systems), there was another accidental side effect of its timing. That was the complete coordination of the time when all three systems went off the air to fix the problems that running together had made obvious. This is the reason that none of the state-of-the-art detectors was on the air on February 23, 1987, when Supernova 1987A appeared. The closest observed supernova in centuries was a chance no one would have chosen to miss, although in fairness at a distance of order 50 kpc it is unlikely, according to standard estimates of the gravitational luminosity, that it would have been seen.

There were, however, non-state-of-the-art gravitational wave detectors observing at the time. Weber has kept a room temperature bar in operation nearly continuously since the '70s, as has the Rome group. An unusual chain of reasoning was constructed, involving a suspect time for the supernova collapse, an unorthodox signature for the gravitational wave event, *ad hoc* assumptions about neutrino physics, and tremendous gravitational luminosity, but leading to a claim of significant detections of a large flux of gravitational wave pulses.[55] This claim has attracted much less attention than did the original claims of Weber in the early 1970s. A serious effort has been made to demonstrate that the statistical significance of the analysis has been overstated, due to construction of the signature to match the data stream.[56]

6 History of Interferometers

6.1 The Work of Gertsenshtein and Pustovoit

Almost as soon as Weber had begun work on the first gravitational wave detector of the resonant-mass style, the idea arose to use interferometry to sense the motions induced by a gravitational wave. Weber and a student, Robert Forward, considered the idea in 1964.[10] We will discuss below how Forward later went about implementing the idea. But the first discussion of the idea is actually due to two Soviet physicists, M. E. Gertsenshtein and V. I. Pustovoit. They wrote in 1962[57] a criticism of Weber's 1960 *Physical Review* article, claiming (incorrectly) that resonant gravitational wave detectors would be very insensitive. Then, they make a remarkable statement justified only by intuition, that "Since the reception of gravitational waves is a relativistic effect, one should expect that the use of an ultrarelativistic body—light—can lead to a more effective indication of the field of the gravitational wave."

Gertsenshtein and Pustovoit followed up this imaginative leap by noting that a Michelson interferometer has the appropriate symmetry to be sensitive to the strain pattern produced by gravitational waves. They give a simple and clear derivation of the arm length difference caused by a wave of amplitude h. Next, they note that L. L. Bernstein had with ordinary light measured a path length differences of 10^{-11} cm in a 1 sec integration time. The newly invented laser, they claim, would "make it possible to decrease this factor by at least three orders of magnitude." (The concept of shot noise never appears explicitly here, so it is not clear what power levels are being anticipated.) They assume that one might make an interferometer with arm length of 10 m, thus leading to a sensitivity estimate of $10^{-14}/\sqrt{\text{Hz}}$ for "ordinary" light, or as good as $10^{-17}/\sqrt{\text{Hz}}$ for a laser-illuminated interferometer. This, Gertsenshtein and Pustovoit claim, is 10^7 to 10^{10} times better (it isn't clear whether they mean in amplitude or in power) than what would be possible with a Weber-style detector. Putting aside their unjustified pessimism about resonant-mass detectors, their arguments about interferometric sensing are right on the mark, even conservative.

For improvements beyond the quoted level, they make suggestions that are somewhat misguided. They say that observation time could be lengthened beyond 1 sec, which would be obvious for some sources (such as "monochromatic sinusoidal signals" or signals of long period) and hopeless for short bursts. Their other suggestion is to use "known methods for the separation of a weak signal from the noise background"; this suggestion is curious because known methods appear to be already built into their estimates that are referenced to a specific observing time. The other lack that is obvious in hindsight is any mention of mechanical noise sources. Still, the gist of the idea of interferometric detection of gravitational waves is clearly present, as is a demonstration that the idea can have interesting sensitivity.

6.2 The Origins of Today's Interferometric Detectors

For a variety of reasons, not least of which must have been the fact that it was written too early (before Weber's work had progressed beyond design studies), the proposal of Gertsenshtein and Pustovoit had little influence. The activity that began the by-now flourishing field of interferometric gravitational wave detection started independently in the West. In fact, it began semi-independently at several places in the United States at around the same time. The roots of this work can be seen in a pair of papers, written in 1971-2, by two teams linked in an unusual collaboration that is acknowledged in

the bodies of the papers, although not in the author lists. The first to be published was that of the Hughes Research Lab team, whose most committed member was Robert L. Forward, the former Weber student mentioned above. Later to appear, and not in a refereed journal, was the work of Rainer Weiss, an MIT physicist who had spent an influential postdoctoral stint with Robert H. Dicke at Princeton. Linking the two groups was someone who never published anything on the subject under his own name, but whose activity is mentioned in both papers—Philip K. Chapman, who had earned a doctorate in Instrumentation at MIT's Department of Aeronautics and Astronautics before joining NASA as a scientist-astronaut.

6.2.1 Interferometer Studies at Hughes Research Lab

An account of the idea for an interferometric detector of gravitational waves, and of the performance of an early-model prototype, is found in the 1971 paper of Moss, Miller, and Forward.[10] The authors cite a program to develop "long wideband" gravitational wave detectors that had started at Hughes in 1966, around the time of Weber's first account of a working resonant detector. The motivations for a wideband detector were 1) to allow detailed meaurement of waveforms which would in turn give insight into the nature of the sources, 2) "to allow the phasing of spaced antennas to form a phased array" (in other words, to allow good temporal resolution so that the direction of the wave can be determined by arrival time differences), and 3) to allow matched filters to be used to optimize the signal to noise ratio of a complex waveform "in addition to the use of standard narrowband frequency filtering for sinusoidal signals, which is the natural filtering action of a resonant antenna." The motivation for the use of a long detector is the larger test mass displacement, which, all else being equal, should directly translate into improved signal to noise ratio.

The authors credit the original idea for this way to achieve a long wideband detector to P. K. Chapman, and go on to state that "Our work has benefited from many discussions with Dr. Chapman as well as R. Weiss, who is involved in the design and construction of his own design at MIT."

The interferometer described by Moss *et al.* was "constructed to set experimental limits on the various noise sources in the laser transducer." It is a classic one-bounce Michelson interferometer, in which both output beams are detected "in a balanced bridge to reduce sensitivity to laser amplitude noise." The operating point for this arrangement was equal photocurrents from the photodetectors at the two output ports of the interferometer. (This is to be distinguished from the use of a photodetector at a single output port of the interferometer that is dithered about the dark fringe. See our discussion of this alternative below.) The interferometer is to be held at this balanced operating point by "slowly acting servo loops... so that the effects of laser amplitude and phase noise are minimized." The flat mirrors were rigidly mounted to an optical table, which was in turn supported on air-filled rubber tubes to give a resonant frequency of 2 Hz. (Other isolation schemes for the rigid interferometer were tried without success, causing the authors to lament that "vibration isolation is still an art rather than a science.") One of the mirrors was mounted on piezoelectric elements, which provided control of the operating point as well as a means of calibration.

The main result presented in the paper is the noise level that was achieved in this prototype interferometer, equivalent to a mirror displacement sensitivity of 1.3×10^{-14} $\mathrm{m}/\sqrt{\mathrm{Hz}}$ at a signal frequency of 5 kHz. This was about a factor of $\sqrt{2}$ larger than the calculated shot noise sensitivity, which the authors state is consistent with other indications that "the sensitivity limits were set by acoustic and ground noise." This noise level was "to date... the smallest vibrational displacement directly measured with a laser." No translation of the sensitivity into strain units was given, either directly or by specification of the arm length. This is perhaps appropriate, since the rigid mirror mounts made this test instrument ill-suited for actually searching for gravitational waves.

The last section of the paper lists the improvements intended to follow this initial work. Firstly, a more powerful laser was proposed; at the then impressive power level of 75 mW, the displacement sensitivity due to shot noise would be "close to that obtained in the present resonant antennas" that were at the date of writing appearing to give significant detections. A final paragraph listed the other proposed improvements: mirrors attached to masses that were large (to reduce thermal noise) and independently suspended from vibration isolation mounts, placed in a vacuum system whose initial length of several meters could be extended "to several kilometers by adding additional evacuated tubes." This section can be read as a telegraphic summary of the plans described at greater length by Weiss in his report written the following year.

6.2.2 The Vision of Rainer Weiss

The other paper that gave birth to the massive worldwide effort to detect gravitational waves using interferometers, Rainer Weiss' 1972 "Electromagnetically Coupled Broadband Gravitational Antenna," appeared only as an unpublished research progress report

of the organization at MIT that administered the umbrella research grant supporting his work.[11] Weber's claimed detection of gravitational waves was very much on Weiss' mind in 1972, reported as possibly correct but with the recognition that the energy flux the waves appeared to carry would dominate the luminosity of the Galaxy. Weiss states that he had been inspired by a 1956 paper by F. A. E. Pirani (that discussed the identification of measurable quantities in general relativity)[9] to consider the possibility that measurements of the light travel time between freely-falling test masses would make the best probes of spacetime structure. He further states that he had realized several years prior to writing (while teaching an undergraduate seminar) that the newly developed lasers could turn Pirani's *gedanken* experiment into a practical measurement strategy. Weiss also notes that the idea "has been independently discovered by Dr. Philip Chapman of the National Aeronautics and Space Administration, Houston."

Many of the ideas that appear in the breathless final paragraph of Moss *et al.* are elaborated at substantially greater length in Weiss' report, which should be considered the first serious design study of the concept of interferometric gravitational wave detection. After the review of Weber's claims, Weiss continues with a clear summary of the physical meaning of gravitational waves in general relativity, and an examination of the possible strength of gravitational waves from the then newly discovered pulsars. He then gives a summary of the key ideas of the proposed system:

- a Michelson interferometer used as a sensor of "differential strain induced in the arms,"
- operated "on a fixed fringe by a servo system" in a modulated system very much in the tradition of Dicke's improved Eötvös experiment,[58]
- "mirrors and beam splitter mounted on horizontal seismometer suspensions" that "must have resonant frequencies far below the frequencies in the gravitational wave" and "a high Q,"
- arms that "can be made as large as is consistent with the condition that the travel time of light in the arm is less than one-half the period of the gravitational wave," in part by being arranged as "optical delay lines"of the style described by Herriott.

Weiss is quite clear about the advantage that accrues from the last point. He says

> This points out the principal feature of electromagnetically coupled antennas relative to acoustically coupled ones such as bars; that an electromagnetic antenna can be longer than its acoustic counterpart in the ratio of the speed of light to the speed of sound in materials, a factor of 10^5. Since it is not the strain but rather the differential displacement that is measured in these gravitational antennas, the proposed antenna can offer a distinct advantage in sensitivity relative to detecting both broadband and single-frequency gravitational radiation. A significant improvement in thermal noise can also be realized.

This last sentence points out one of the key insights of this report, expanded upon at much greater length in the remainder of the text. As a sensitive mechanical measurement, the interferometric detection of gravitational waves is prey to a host of mechanical noise sources whose strengths need to be minimized if success is to be achieved. By far the largest section of the paper is devoted to estimates of the magnitudes of a long list of noise sources of various kinds. They include: amplitude noise in the laser (the only place where the work of the Hughes group is cited, as an example of a shot noise limited measurement), phase noise in the laser, mechanical thermal noise, radiation pressure noise, seismic noise, thermal gradient ("radiometer effect") noise, cosmic ray impacts, "gravitational-gradient" noise, and fluctuating forces from electric and magnetic fields. This looks almost (with a few omissions) like the list of noise sources that contemporary workers are grappling with as they strive to make the new kilometer scale interferometers work; by contrast, the other earlier treatments of the subject look myopic and unbalanced. And this insight is what led to the recognition that interferometers of the greatest practical length, with the resulting dilution of displacement noise terms as compared with a strain signal, would be the way to achieve the promise of good gravitational wave sensitivity, and would be worth the substantial investments needed to build them.

6.3 Interferometer as an Active Null Instrument

The agreement between Weiss and the Hughes group on the basic features of an interferometric detector must have something to do with the fact that they and Chapman were engaged in a remote three-way collaboration. But the fact that the key features of the design remain current to this day (with a few important additions) is evidence that they responded thoughtfully to an inherent logic of experimental design. Interferometric gravitational wave detection represents an extreme example of the application of design principles of wide validity in experimental physics. It is worthwhile to examine those principles here.

6.3.1 How to Maximize the Signal to Noise Ratio

Firstly, it is important to recognize where, in the spectrum of physics experiments, this one falls. To successfully detect gravitational waves, it will be necessary to attain AC strain sensitivity that is completely unprecedented. The concept of sensitivity is the key one—the signal to noise ratio for small strains is (essentially) the only figure of merit of interest here. Precision is primary, while accuracy of calibration or absence of other kinds of systematic errors definitely takes a back seat at this stage in the development of the field. Also, it is important to remember that one is perfectly content to measure signals only above some cut-off frequency; no absolute measurements of lengths or even of differences in length between two interferometer arms are necessary.

In experiments where precision is the primary desideratum, certain general principles apply. In particular, there are two fundamental strategies for maximizing a signal to noise ratio: one can make the response of the system large, and one can make the noise small. To make the response large, one can try to "capture" the largest possible effect from an external source. Here we want to make the largest possible apparatus, since relative displacements of two masses caused by a gravitational wave are proportional to the separation of the masses. This strategy is quite common—large telescopes can be more sensitive than small telescopes, for example. Another example is the improvement Michelson and Morley achieved in 1887 over Michelson's solo result of 1881, achieved by increasing the expected ether drift fringe shift through the addition of mirrors to lengthen the interferometer's optical path length.[59]

Another way to make a large response is to arrange for the external influence to drive an indicator that changes dramatically in response to a small effect. The implementation of that strategy here is to illuminate the interferometer with visible light, whose wavelength of order 1 μm sets the scale for motions that change the output port from dark to bright. This strategy too is common, as in such simple choices as making the indicator needle on an ammeter as narrow as practicable, or constructing a galvanometer with a fine fiber so that it will swing as much as possible or in using the longest possible lever arm for its optical lever.

Not all of the measures listed above are guaranteed to improve the signal to noise ratio; that depends on the nature of the noise. But one always improves the signal to noise ratio by reducing the magnitude of the noise. The noise sources in mechanical experiments can be loosely grouped in two classes. One class is noise in the measurement of the response of the test system. For interferometers, this is the shot noise in the fringe readout, or whatever other effect swamps it (excess amplifier noise, for example). In resonant mass detectors this class is represented by noise in the preamplifier that responds to the transducer output, or by excess transducer noise such as Johnson noise. We saw above that there may be quantum mechanical limits to the reduction of this kind of noise.

The other class of noise sources are those that also affect the test system in a way that mimics the effect being sought. This includes all of the noise forces on the test masses, like seismic noise and thermal noise. (This is the class of noise effects usually called, when it wouldn't cause confusion with the principle of an interferometer's operation, "interference"). Over a large range of frequencies, these are the dominant limits to sensitivity of a gravitational wave interferometer. Where that is true, measurement strategies that maximize the strength of the external effect on the test system (such as, here, the physical separation of test masses) can be helpful, but those that just maximize the response of the measuring instrument to any external effect just magnify the response to the noise as well. Reducing the strength of these kinds of noise is always a good idea, whenever it is possible.

6.3.2 What to Do when $1/f$ Noise Dominates

The considerations listed above are mostly obvious enough to be considered common sense. But there are some subtle aspects to consider as well. The most important non-obvious fact to consider is that almost every noise source has a pronounced $1/f$ character to its spectrum, at least at low frequencies. A laser whose power stability is limited at high frequencies by fundamental processes like shot noise almost invariably will show dramatically enhanced levels of noise at low frequencies. And, in many situations even such a fundamental process as thermal noise shows a $1/f$ or steeper power spectrum, as we saw above. Furthermore, seismic noise almost invariably has a very steep displacement power spectrum.

The difficulties caused by $1/f$ noise can come in several different forms. Sometimes the difficulty is simply that the noise has a larger magnitude than the minimum level it might have (laser power noise in excess of shot noise, for example), or that the noise is just inconveniently large (such as seismic noise at any frequency!). A more subtle kind of difficulty is the large integrated magnitude of the noise at frequencies that are low compared with the signal frequency. Many measuring devices have only a limited dynamic range over which they respond linearly to external influences; it seems a shame

to waste it dealing with noise outside the band in which one expects to find signals.

Measures to deal with $1/f$ noise often dominate the design of a high-sensitivity experiment. No one explored these measures more systematically than did Robert H. Dicke, who was for a few years in the early 1960s Weiss's mentor. Dicke was moved to think deeply about these problems while working on the development of microwave radar at the Radiation Laboratory at MIT during World War II. Devices called radiometers, receivers that measure the total power emitted by a broadband (often thermal) source of microwaves, would have had a variety of uses, if they hadn't been rendered so insensitive by the large amount of $1/f$ noise in the RF preamps. Dicke invented what came to be called the "Dicke radiometer" specifically to solve this problem.[60] The heart of the scheme was a device to periodically (at 30 Hz in the original case) interrupt the flow of RF power from the antenna to the preamp, replacing it instead with a thermal source of radiation. At the "back end" of the instrument, the electronics were arranged to give a measure of the difference between the detected power from the antenna and from the reference load. The reason this defeats the $1/f$ noise in the amplifier is that the comparison between measure and reference is made rapidly enough that the preamp's output can't wander much in the interval. Or, described in the frequency domain, the signal has been translated from DC up to a high enough frequency that the preamp's noise is not dominated by excess noise with a $1/f$ character.

Dicke's invention had the immediate effect that microwave radiometry became practical even for sources with rather low antenna temperatures; this made a substantial contribution to the Rad Lab's mission[61] as well as to the practice of radio astronomy.[62] But the greatest impact came from Dicke's realization that this modulation technique would have broad applicability, wherever $1/f$ noise was a problem. This insight led him to invent the lock-in amplifier, a universal back end that can control the chopping of an experiment, calculate the difference in output between the "on" and "off" states, and average the result to further reduce the noise. By now, "lock-in amplification" (also referred to as "phase sensitive detection") has become a nearly universal practice in the fight against $1/f$ noise.

A second classic measurement carried out by Dicke illustrates further insights in the battle against $1/f$ noise. The test of the equivalence of inertial and gravitational mass carried out by Roll, Krotkov, and Dicke[58] is considered one of the great examples of a *null experiment*. As championed by Dicke,[63] this term refers to a measurement where an answer of zero carries tremendous meaning. Precise equivalence of inertial and gravitational mass (or in other words a zero value for their difference) means that gravity can be described by a metric theory.

Null experiments play a special role for experimentalists as well as theorists, because an instrument that reads zero is immune to many of the sorts of problems that plague non-null measurements. Among these are calibration drifts and limited dynamic range of an instrument (whether from noise or from non-linear response.) Of course, turning a theoretical zero into an idea for an instrument that yields a zero output takes deep insight. One could argue that the torsion balance used in the improved Eötvös experiment, whose motion would track the Sun's if the aluminum and gold masses on opposite sides had differing ratios of inertial to gravitational mass, is among the most elegant instruments ever invented.

Maintaining the integrity of a null measurement takes insight that goes beyond the design of the front end of the experiment. For example, it would be a good idea for the null position of the test masses to be arranged to correspond to a null response from the sensor. Then, one can ignore (to first order) fluctuations in the drive level of the sensor (such as the light power in the optical lever), since zero is still zero even if it is multiplied by, say, 1.01 instead of 1.00. There are a variety of ways to create a null output from an optical lever at one particular operating point. One way would be to use a matched pair of photodetectors, placed so that the light beam falls equally on each detector when the balance is at the null position; as the beam moves to follow the balance's motion, one photodetector receives more light while the other receives less, and a differential amplifier will reveal the motion. This method is essentially a DC technique.

Dicke's team implemented a clever variation that let them make use of the advantages of a lock-in amplifier. A narrow light beam fell on a single photodetector, after passing by a wire of comparable width that cast a shadow on the photodetector. The wire was caused to vibrate from side to side by driving a current through it at the frequency of one of its "violin" resonances; as it did so, its shadow also moved from side to side across the light beam. If the beam were centered on the wire's position, then the light received by the photodetector would increase equally due to the wire's vibration to the left or the right. But if the beam were off center, then one direction of wire vibration lets more light pass than the other. So in the centered case the photocurrent varies only at twice the frequency of the wire's vibration, while in the off-center case the current contains a component at the wire's vibration frequency, whose amplitude and phase carries the information of the position of the light beam with respect to the wire. A lock-in amplifier converts the modulated signal into one at DC. The wire vibration can

easily be arranged to be at a frequency high enough to avoid $1/f$ noise.

Noise will always cause a sensitive instrument to depart from the precise null reading, and noise of a $1/f$ character will do so in an especially vexing way. Once it has drifted from the null, the instrument is once again sensitive to drive level fluctuations because its output has a non-zero value. The especially elegant feature of the Dicke torsion balance was the provision to apply a correction torque to hold it at the null position. This was accomplished by means of a servo loop that fed back a torque to null the output reading of the sensor. The greater the gain of this loop, the more tightly was the balance constrained near the null position, and the smaller was the sensitivity to light level fluctuations. Of course, the output has to be recalibrated for the effect of the loop, but the same information is present. One convenient place to read the output is to measure the control torque needed to hold the balance at null, which will just be equal to the external torque applied to the balance.

6.3.3 Application of these Principles in an Interferometer

Maximization of response in an interferometer is achieved by the obvious choices: placing the far mirrors as far away as possible from the beam splitter, and by using the shortest practicable wavelength for the illumination of the interferometer. Weiss also describes increasing the optical path length by use of a delay line that causes the light to make many trips through an arm; this is effective against sensing noise (like shot noise), but is just a wash for noise effects that cause actual mirror motions since N encounters with mirrors necessary import a factor N larger amount of these kinds of noise along with the extra signal.

Minimization of the noise starts with obvious choices as well: largest possible light power to make the shot noise small, and the minimization of the multitude of mechanical noises as discussed in Weiss' report. But it also involves replicating many of the features of the Roll *et al.* experiment to minimize the deleterious effects of $1/f$ noise of various kinds. This is the reason for the application of high frequency modulation of the fringe phase (by PZT transducer at Hughes, or by Pockels cell according to Weiss), for the arrangement of the measurement to give a null reading (by differential measurement with two photodetectors at the mid-point of a fringe in the Hughes case, and by modulation about the dark fringe at MIT), and by the provision for actively holding the instrument at the null state (by PZT again at Hughes, and by a combination of Pockels cell and force applied to test masses at MIT).

The fringe modulation and locking techniques mentioned just above have as their goal elimination of excess noise from low frequency fluctuations of the laser's output power. Other ways to make the instrument null are used to make it insensitive to other sorts of fluctuations. For example, laser wavelength fluctuations can be rendered harmless if the interferometer is operated at or near the "white light fringe," that is, with well-matched arm lengths. If the light travels through arms of equal length, then the phase relationship between them is hardly altered as the wavelength varies. But if the arms have a substantial difference in their lengths, wavelength fluctuations cause spurious relative phase shifts at the output, even without fluctuations in arm length difference. Note that this sets a different condition on declaring the instrument to be null—against laser power noise, it is the operating point within a fringe that matters, but for laser wavelength noise, it is the choice of which fringe on which to sit that is relevant.

6.4 Wrap-Up of the Hughes Program

A summary of the results eventually achieved by the Hughes group was reported in 1978 in a paper in *Physical Review D* whose sole author was R. L. Forward.[64] It represents a retrospective analysis of a completed experiment, somewhat in the same style as Kafka and Schnupp's paper of the same year. But there are also strong stylistic differences between the two papers—while Kafka and Schnupp want to make the strongest possible point with the data they analyzed (concerning whether Weber's events could be gravitational waves) with hardly a pause to describe the measurement apparatus, Forward's paper is primarily concerned with explicating the basic physics of the detector, and is somewhat cavalier about the observational data that is presented. The contrast is due in part, of course, to the distinction between publication about a mature technology (bars at room temperature) and one that was clearly immature, although with great promise. It is also a reflection of two facts about the Hughes interferometer as a scientific instrument: it was not as sensitive to bursts as were the bars of 1972, and its data stream was also much harder to analyze because of its broadband nature.

The first section of the 1978 paper is a detailed derivation of the antenna pattern (sensitivity versus angle) of an interferometer. Next comes a quite detailed description of the apparatus, with a strong emphasis on the optical aspects of the interferometer that determine the noise above 1 kHz: circuit diagrams of the photodiode biasing network and of the 1-20 kHz bandpass filter, part numbers for the photodiodes and the

front-end amplifier, and a painstaking derivation of the shot noise. Mechanical aspects of the interferometer, which mainly determine performance at lower frequencies, get shorter shrift: test mass suspensions are described in a single paragraph, as neoprene and brass stacks of "the desired height" with "typical frequency of 10 Hz," without any discussion of the mechanical transfer function or of the thermal noise of the suspension. The three paragraph section devoted to the "isolation system" gives information on both seismic isolation tables and the vacuum system enclosing the interferometer, and includes the following remark, almost in passing: "The vacuum system and isolation tables were designed so that after an initial checkout and operation with 2-m sections of aluminum irrigation pipe (8.5 m total interferometer pathlength), those sections could be replaced with longer sections (up to 1 km) with a substantial increase in interferometer-gravitational radiation-strain sensitivity for the same photon-noise-limited displacement sensitivity."

A careful discussion of the calibration of the instrument and of its linearity is provided, although without including any but the most cursory details about the servo system (whose actuator was a PZT stack on which one of the mirrors was mounted) used to keep the interferometer in lock. One presumes that ignoring the behavior of the servo was justifiable on the assumption that its bandwidth was smaller than 1 kHz. Indeed, given the strategic decision only to consider the output of the interferometer at frequencies above 1 kHz, most of the omissions that strike a modern eye as surprising can be seen to make sense.

The discussion of the operation of the interferometer as a gravitational wave detector begins with a paragraph that will evoke much sympathy from present-day readers. It repeats the "ultimate plan" of operation at a remote site with long arms, but concludes with the remark, "The funding for this next move proved to be unavailable so we concluded the program by operating the system as it was, despite the high level of acoustic, electromagnetic, and vibrational noise from the other activities in the building." Operation was evidently difficult, since as the Abstract notes, "The laser interferometer was operated as a detector for gravitational radiation for 150 h during the nights and weekends from the period 4 October through 3 December 1972," a duty cycle of a bit over 10%. Environmental noise was a serious problem, and was taken seriously: a set of monitors of seismic, acoustic, optical, and electrical noises was installed, and a measure of their outputs was recorded along with the interferometer output.

Part of what made taking data so difficult was the decision to take advantage of the high bandwidth available; although the band below 1 kHz was abandoned as useless because of high noise levels, the upper frequency cut-off was taken to be 20 kHz. This choice was never discussed in the paper—one might have expected some sort of argument on astrophysical grounds that such high frequencies might contain signals, but it is just as likely that the cut-off was chosen to match the bandwidth of the "high quality stereo tape recorder" that was used as the primary data storage medium. Dealing with this much data was a tremendous burden, given the state of computer technology in the early '70s. In fact, the data processing was performed almost entirely by listening to the audio tape—one section of the paper is called "Calibration of Ear." (One 10 msec digitized chunk of data is shown in the paper, both in the time and the frequency domain.)

This method of data analysis was a clever solution to a vexing problem, and indeed continues to be a model for qualitative analysis and debugging of interferometers today. But it showed its weaknesses in what might otherwise have been the most interesting section of the paper, "Comparison of Data with Other Observers." Here, Forward looks for coincidences between the unexplained events in his data set (not coincident with environmental signals in the monitor channel) and events in the resonant mass detectors that were in operation at the same time, at Frascati, Glasgow, and the Maryland group's detectors at College Park and Argonne. In every case, Forward found no event in his detector at the time of a candidate event from another detector. He notes ruefully that "The one 'distinctive signal' reported by the Glasgow group occurred at 13 h 07 min 29 sec GMT 5 September 1972, which was prior to the start of the Malibu data collection period."

The data were of course most interesting for their comparison to the results reported by Forward's former mentor Weber, since the latter was continuing to report coincident events between his various detectors. Forward notes that there were seven time blocks during which unexplained events in the interferometer occurred in close proximity to Weber coincidences. However, he further states, "Both raw power and derivative power-squared digitized data plots digitized to 0.1-sec accuracy were obtained from the Maryland group and compared with the 0.2-sec accuracy Malibu data. None of the audible Malibu signals fell within 0.6 sec of a Maryland-Argonne coincidence."

Fair enough, but what can be concluded from this lack of coincidences? Not much, according to Forward, since "It is difficult to compare the relative detection capabilities of the various antennas since their amplitude sensitivities, bandwidths, and signal processing techniques differ widely." He goes on to state that "at the time one of the bar-antenna systems produced an event or coincidence corresponding to a gravitational-

radiation signal with an amplitude of 0.1 fm/m [10^{-16}] due to spectral components in a narrow band around the bar resonance, the amplitude of the gravitational-radiation spectral components in the entire band from 1-20 kHz was definitely less than 10 fm/m and was probably less than 1 fm/m."

6.5 System Designs for More Sensitive Interferometers

6.5.1 The Magnitude of the Challenge

Forward's paper describing the state of the art of interferometric gravitational wave detectors in 1972 appeared in 1978. It was not until ten years later still that the next improvement in the state of the art was considered significant enough by its authors to warrant another comparable paper in the *Physical Review*. It is worth understanding what factors might have contributed to this slow pace of progress. Certainly the collapse of the credibility of Weber's "events" and the consequent redirecting of the effort at the much smaller signals predicted by astrophysical theory removed a lot of the motivation for trying to achieve incremental improvements in sensitivity. Instead, ultimate success came to be seen to require heroic improvements; neither the 10^{-16} of the room temperature bars nor the 10^{-18} sensitivity of the cryogenic bars would guarantee that gravitational wave astronomy could be established. It looked like frequent signals weren't likely to appear with strengths in excess of 10^{-21}, and perhaps not even much above 10^{-22}. If that were truly the case, then every possible advantage that could be squeezed out of the ideas detailed in Weiss' report would be required. And to achieve them, a great deal more effort was required than could be obtained from the Hughes interferometer, with its sensitivity in the range of 10^{-14} to 10^{-15}interferometer's arms from 2 meters to 1 km would hardly have begun to meet the need.

A sober look at the prospects for the field pushed its practitioners to even greater ambitions than simple improvement of sensitivity at or above 1 kHz. Much of the hope for detectable signals near 1 kHz had come to be focused on the radiation from the "bounce" at the end of the gravitational collapse of a stellar core at the onset of a supernova of Type II. But the theoretical predictions of the strength of such events were very uncertain. If one wanted to maximize the potential for discoveries, one had to imagine building instruments sensitive to the widest possible varieties of signals. And, since very few of the imaginable signals had frequencies above 1 kHz, the best way to improve one's chances was to improve the sensitivity to low frequency signals. This meant minimizing the magnitudes of many of the noise sources that Weiss had listed, a task that the Hughes group had chosen (for sensible reasons) to defer.

- Improved isolation from seismic noise is straightforward in principle, but very demanding in practice. The essential idea was employed by Weber: a cascade of mass-spring oscillators can provide a very steep slope to the frequency dependence of the isolation, making it acceptable for frequencies higher (by of order a decade or more) than the characteristic resonance frequency. It is a real trick to move the whole set of resonances to low enough frequencies (a few Hz) to extend the useful band down as low as 10 Hz or so.
- On the assumption that seismic noise can be controlled, another difficult noise source looms at low frequencies. The thermal noise of the test mass about its equilibrium position can become the dominant noise source, unless very specific measures are undertaken. In place of Forward's rubber mounts for the test masses, one needs to use what Weiss described as "horizontal seismometer suspensions." By this he meant suspensions that had two attributes: low resonant frequency and high Q. The latter was to be achieved by an arrangement whereby most of the restoring force on the test mass comes from gravity (which is nearly dissipationless), as in a pendulum. His system diagram is marked with the compliant directions for the seismometer suspensions, on the assumption that they would be relatively rigid in the orthogonal directions. Present thinking has moved to simple pendulums, partly because of worries about alignment but mostly because the rigid members of single d.o.f. seismometer suspensions tend to have internal resonant modes at inconveniently low frequencies.
- Shot noise minimization needs the whole bag of tricks, including a laser whose power was rated in tens of Watts (not tens of milliWatts), and an optical path length comparable to the gravitational wavelength. The latter would require Forward's desired kilometer-scale separation plus folding of the optical path, such as by the delay line described in Weiss' report. The Herriott delay line has served well in meter-scale prototypes, but most of the large projects have adopted long Fabry-Perot cavities, proposed for this purpose by Drever. (See the next section.) These have the great advantage of allowing the use of the smallest possible test masses, and taking up correspondingly small cross sections in the vacuum pipes. But their operation is substantially trickier than corner reflectors or delay lines, since each Fabry-Perot cavity is itself an interferometer that only performs like a beam-folding element when it is locked to the laser wavelength, using a servo of some

subtlety. Fabry-Perot cavities also appear to be less subject to excess noise from light scattered into unanticipated paths from mirror imperfections, a problem not suspected by either Weiss or Forward.

- Achieving the high sensitivities to which we now aspire requires vacuum of a quality much better than the Hughes interferometer. Pressures of 10^{-6} torr or better are required. The vacuum pipes are themselves much larger in diameter, due partly to the great care needed to keep scattered light effects at low levels. Scattered light also demands that baffles be properly placed in the interior of the pipes.

So it was probably going to take more than adding a few kilometers of irrigation pipe to the Hughes interferometer to detect gravitational waves with an interferometer. The realization that all of these features would be necessary was daunting, and caused the character of work on interferometers to change. Instead of quick demonstrations, it was considered necessary to try to engineer the variety of subsystems that high sensitivity would require. Instead of a device that Forward could honestly describe as a "gravitational-radiation experiment," workers conceived of their apparatus as "prototype gravitational-wave detectors."

6.5.2 Ron Drever's Bag of Tricks

Ronald W. P. Drever was one of the leaders of the generation of experimenters who followed Weber, only to find no signals that matched his claims. Rather than build a faithful copy of Weber's original bar, he chose to follow the path invented by Aplin (and publicized by Gibbons and Hawking) of the split bar, which maximized the bandwidth of the detector. When it became clear that much greater sensitivity would likely be required, he (like the German group) chose to switch to work on interferometers. His work in this period is again marked by an enthusiastic exploration of clever ideas. It is not marked, however, with many conventional papers in refereed journals. Instead, his most stimulating work is to be found in conference proceedings and lectures at physics summer schools.

Several of Drever's most important contributions are described in the text of the lectures he gave at the NATO Advanced Study Institute on Gravitational Waves held at the Les Houches Center of Physics in 1982.[65] Cast as an overview of the interferometric method of gravitational wave detection, it is dominated by an account of three crucial improvements on the basic scheme of Weiss, Forward *et al.* Each of these variations has come to play an important role in the design of the large detectors now under construction.

The Introduction gives an astoundingly brief account of the history of the field:

> An obvious way one might consider detecting gravity waves is through the changes in separation of free test particles, and the idea of using optical interferometers for observing this has certainly occurred to many physicists: indeed one might wonder why so few searches for gravity waves have been made this way.

The work of the Hughes group is mentioned in passing as demonstrating that a simple gravitational wave interferometer could achieve the shot noise limit. Weiss's work is referred to later as having contributed the idea of the delay line as an "important practical method for improving photon-noise limited sensitivity."

Drever goes on to describe the optimization of the parameters of a delay line, from the point of view of shot noise reduction. He then remarks on a "practical difficulty" that "became apparent in early experiments at Munich and at Glasgow—the potentially serious effect of incoherent scattering of light at the multireflection mirrors or elsewhere in the system." The interference between scattered light and light following the intended paths (which is non-stationary because the path followed by the scattered light can vary in length both on long and short time scales) proved to be a very troubling noise source in delay lines. The German group, Drever notes, proposed a way of modulating the laser light that would minimize the problem. Drever then suggests that "another approach would be to make the path traveled by scattered light equal to that of the main beam, and this may in fact be achieved if another type of optical system, a Fabry-Perot cavity, is used instead of a Michelson interferometer with many discrete reflections in each arm."

The basic idea was that light traveling between parallel mirrors can be, in effect, trapped for many round trips, until it is either absorbed, scattered, or leaked out by transmission through one of the mirrors. Thus, such a cavity can play the same role as a delay line with its many spatially separate reflections. The classic Fabry-Perot cavity used flat mirrors, usually equivalent to one another, usually closely spaced compared with their diameters or with the diameter of the beam of light, and usually operated in transmission (that is, with the interesting light emerging from the mirror opposite to the one into which the light was injected). What Drever proposed was rather different: a pair of small mirrors (no larger than necessary to keep diffraction losses small), spaced

apart by kilometers. The distant mirror has as nearly perfect a reflectivity as possible; a finite transmission is used on the near mirror, which functions both as an input and an output coupler for the light. Thus modified, a Fabry-Perot cavity can serve as an arm for a Michelson interferometer.

Drever is explicit about both the advantages and the drawbacks of this beam-folding scheme.

> The diameter of the cavity mirrors can be considerably smaller than that of delay-line mirrors.... This reduces the diameter of the vacuum pipe required, and also may make it easier to keep mechanical resonances in the mirrors and their mountings high compared with the frequency of the gravity waves, thus minimizing thermal noise. The Fabry-Perot system has however, some obvious disadvantages too—particularly the requirement for the very precise control of the wavelength of the laser and of the lengths of the cavities. Indeed with long cavities of the high finesse desirable here exceptional short-term wavelength stability is required from the laser.

The heart of the difficulty is that, unlike a delay line, a Fabry-Perot cavity stores light because it is in itself an interferometer—the trapping of the light for many round trips comes about only by careful adjustment of the phases of the superposed beams. This can only occur when the wavelength of the light and the length of the cavity are in resonance, that is, matched so that an integer number of waves fits into the cavity. Very near the resonance condition, the phase of the output light varies with mirror separation in the same way as the light that has traveled through a delay line. To achieve this condition, the light and the arm have to be locked together by a servo system. Drever's lecture goes on to describe the style of servo required, one that he and his group developed in conjuction with John Hall's group at the Joint Institute for Laboratory Astrophysics in Boulder, Colorado.[66] This servo design has its roots in an analogous microwave device developed by Robert Pound.[67]

While the essence of the difficulty was thus solved, in practice the use of Fabry-Perot cavities has additional complications. One is due to the fact that when a cavity is not very close to resonance, the phase of the output light has almost no dependence at all on the separation of the mirrors, thus making it very hard to generate the sort of signal necessary to acquire the lock on resonance in the first place. An additional level of complication comes when the arm cavities are assembled into a complete Michelson interferometer, since the interferometers within the interferometer need to be separately controlled without degrading the function of the main instrument. Solving these sorts of problems robustly has proven to be challenging work.

As a stopgap measure until those challenges could be met, Drever proposed a kind of interferometer that functioned without the light from the two arms being recombined at the beam splitter for the comparison between their phases. In the context of an ordinary Michelson interferometer this may seem impossible; but, since Fabry-Perot cavities are themselves interferometers, they can each be used to make an independent measurement of the difference between the wavelength of the light and the length of the cavity. A laser illuminates the two orthogonal arms of an interferometer containing Fabry-Perot cavities. Arm 1 is read out by the Pound-Drever-Hall method to generate an error signal reflecting the mismatch between the laser light and the length of the arm. This error signal is used in a servo loop to cause the wavelength of the laser to resonate with Arm 1. The other half of the laser's light is directed by the main beam splitter to Arm 2. There, the mismatch between the light and the arm generates another error signal, reflecting the difference in length (modulo N wavelengths) between Arm 2 and Arm 1. Since a fluctuating arm length difference is the signature of a gravitational wave, this second error signal constitutes the scientific output of the interferometer. (It also of course contains the various noise sources that cause or mimic arm length differences, but in that sense it is no different from the dark fringe output of a more conventional interferometer.) Drever included the use of that error signal being used to close a second servo loop that forces the length of Arm 2 to follow the wavelength of the light. This is necessary to keep the arm at a functional operating point; it means that the error signal has to be calibrated to take the servo action into account, which sounds more complicated at the heuristic level than it actually is in practice.

The scheme described in the previous paragraph formed the basis for almost all of the work on Fabry-Perot-based gravitational wave interferometers until well into the 1990s. But none of its proponents ever expected it to play more than a stopgap role. This is mainly because, without recombination of the light from the two arms to generate "dark" and "bright" output ports, the system cannot be used as the basis of what came to be known as *power recycling*, described for the first time in this lecture, featured there as one of the "possibilities for future enhancement in sensitivity."

Drever introduces the idea of recycling (modestly referred to as a "possibility for more efficient use of light") in the context of a delay-line Michelson interferometer, for pedagogical simplicity. The key idea is as follows. If an interferometer has long arms and if it is constructed from mirrors with high enough reflectivity, then the light exiting

from the bright port may be nearly as bright as the light entering the interferometer from the laser. (This insight represents a profoundly different "take" on the issue than can be found in Weiss' work; he instead worried about optimizing the shot noise versus number of bounces on the assumption, good for short arms and poor mirrors, that substantial losses would eventually occur.) The light exiting the bright port is every bit as good as "fresh" light from the laser, so it seems a shame to waste it. Drever's proposal is to arrange by an appropriate set of mirrors to redirect the used light into the interferometer, in coherent superposition with light arriving directly from the laser. (This has to be done using a beam splitter or other partially reflecting mirror of a carefully chosen reflectivity.) This arrangement has in effect made the whole interferometer into a single Fabry-Perot resonant cavity, whose back mirror is the Michelson interferometer, and whose input/output coupler is the partially-reflecting recycling mirror.

In principle, the advantages that could be achieved with this technique are quite large. Drever quotes rms shot noise in a search for 1 msec pulses of 10^{-22}, far superior to what could be achieved without recycling. He also gives a diagram showing how the technique could be applied to an interferometer whose arms were made of Fabry-Perot cavities. The components necessary to sense and control the various internal degrees of freedom are drawn in with dashed lines, as an indication of the provisional nature of the design. In fact, more subtle schemes have had to be developed to implement such as system. But, given the quality of the mirrors available today (and the lack of commensurate progress in the power levels available from stabilized lasers), power recycling has been adopted as an essential feature of every large interferometer under construction today.

The other "possibility" described here is one "for enhancing sensitivity for periodic signals." This one is again introduced in the pedagogically simpler delay line interferometer. And again, the aim is to find a way to make use of the fact that, with good mirrors, the light would not be significantly attenuated after it has spent one half of a gravitational wave period in an interferometer arm. A periodic gravitational wave persists (by definition) for much longer than one half period; why not find a way to accumulate a phase shift on the interferometer's light for a much longer interval? The scheme proposed here does just that, by arranging for light that exits one arm after one half cycle of the gravitational wave to enter the other arm, where it stays for another half cycle. The light changes arms at the same time that the gravitational wave changes sign, at least for the signal frequency that matches the length of the interferometer. A partially reflecting mirror governs how long the light repeats this cycle before finally exiting the interferometer. As with power recycling, Drever goes on to show how a similar effect can be achieved in an interferometer that uses Fabry-Perot cavities.

It has been shown recently that the scheme can actually be implemented in a much more elegant way, using a single partially reflecting mirror at the nominally dark port. In analogy with power recycling, this scheme (called *signal recycling*) can be thought of as forming a single large Fabry-Perot resonant cavity out of the interferometer, this one resonant at the frequency of the signal sidebands on the laser light that have been created by the action of the gravitational wave.[68] This version of the idea will almost certainly also find application in the next generation of large interferometers.

6.6 The Garching 30-Meter Prototype Gravitational-Wave Detector

The 1988 paper to which we referred above was the account by the group at the Max-Planck-Institut für Quantenoptik, the successor to the bar group of the early '70s. Shoemaker *et al.*[69] provided a beautifully detailed account of the best-characterized interferometer prototype yet built. It can be thought of as the work that brought to fruition, on the meter scale at least, the ideas embodied in Weiss's 1972 design study. Through the '70s and '80s, a number of groups (including Weiss's at MIT, Drever's at Glasgow and at Caltech, and Brillet's at Orsay) worked in parallel with the MPQ group to develop prototypes of kilometer-scale working interferometric detectors.[70] The MPQ paper makes a nice example, though, since it is an especially complete account of a well-functioning instrument. So for pedagogical purposes we let it here stand for the large body of work done worldwide through the 1980s.

The interferometer described here had test masses 30 meters from the beam splitter; light made 45 round trips, for a total light travel time in an arm of 9 μs. The folding of the optical path was achieved with a Herriott delay line. The interferometer was illuminated with an Argon-ion laser at $\lambda = 514.5$ nm, capable of supplying up to 0.23 W to the photodetector (at a bright fringe) after all optical losses in the interferometer are included. The test masses consisted of simple glass mirrors with a radius of curvature 31.6 m; they were suspended from single-wire slings of free length 0.72 m, giving a resonant frequency of about 0.6 Hz. Each of these was in turn suspended from a metal plate hung from coil springs. This upper level of the suspension not only added isolation along the optic axis, but gave isolation in the other degrees of freedom that might cross-couple into the sensitive direction.

The many compliant degrees of freedom of such a system needed damping, but damping of a sort that would not add substantial amounts of noise, in particular thermal noise. The solution was a set of optical shadow sensors that detected the positions of small vanes attached to the mirrors; the outputs of these sensors, suitably filtered, drove currents through coils that were in turn positioned near magnets integrally mounted with the vanes. This is a very robust kind of servo, and it can be made sufficiently noiseless, at least at intermediate and high frequencies. In the Garching 30-meter interferometer, 16 degrees of freedom were damped using servos of this kind.

The Garching group paid careful attention to the influence of fluctuations in laser power, wavelength, position, and angle. Laser power fluctuations are dealt with primarily by implementation of the modulated dark-fringe servo system. Sensitivity to variations in the input beam's transverse position and injection angle was minimized by using an optical fiber as a spatial filter for the laser, an idea the authors credit to Weiss. Frequency fluctuations enter in two ways:

- Mismatches in the curvature of the mirrors cause the two interferometer arms to differ in length by an amount ΔL, when the mirrors are spaced so as to give the same number of round trips N through the arms. The spectral density of effective arm length noise $x(f)$ due to a frequency fluctuation spectrum $\delta\nu(f)$ is

$$x(f) = \frac{\Delta L}{N}\frac{\delta\nu(f)}{\nu}.$$

- Light that scatters out of the intended path can nevertheless find its way to the beam splitter again, where it interferes with the rest of the light. Although good mirror surfaces and coatings minimize the fraction of the light scattered, the very large mismatches in path length that can result made this a comparable mechanism for conversion of laser frequency noise into apparent mirror motion.

In order to make laser frequency noise small enough, the Garching group stabilized the laser using a combination of two servo loops. In the first loop, a rigid Fabry-Perot resonant cavity serves as a length reference; the light's wavelength is held to resonance in the cavity, by inspection of the transmitted light intensity. A second error signal is derived by interfering the light that exits the bright port of the interferometer (whose phase depends on the average of the lengths of the two arms) with some of the new light entering. Feedback was applied to a mirror at the end of the laser cavity, to a Pockels cell within the laser cavity, and to the common mode positions of the two end mirrors of the interferometer.

The noise spectrum of the system was in reasonable agreement with a noise budget prepared from estimates of the noise sources described above. The rms noise in a 1 kHz bandwidth near 1 kHz was about 3×10^{-18}, only a factor of three poorer than the best that had been achieved by a resonant-mass detector. This was one of the great triumphs of the German work. The other was the robustness of the rather elaborate active system that this interferometer had become, of order two dozen servo loops. It typically stayed in lock for 30 minutes to an hour, and upon losing lock would shortly reacquire lock automatically.

The 30-meter interferometer was run for 100 hours in March 1989, in coincidence with the ten-meter interferometer at Glasgow that had achieved comparable sensitivity. Even then, the enthusiasm for looking for gravitational waves at "low" sensitivity was such that an analysis of the results of the run was not published until 1996.[71]

6.7 Designs for Kilometer-Class Interferometers

As noted above, laboratory work on interferometers was almost from the beginning considered an engineering exercise preparatory to the construction of instruments with arms of kilometer scale. With several such devices now under construction, it is worth reviewing their distinctive features. Here, even more so than for the other cases we have been discussing, the refereed literature is a poor source of information, and so are conference proceedings. For most projects the only detailed descriptions are those contained in funding proposals. (The one redeeming feature of this form of publication is that, since the reviewers of such documents were typically not expected to be experts in the field, they contain an abundance of carefully written tutorial material, and well-reasoned justifications for most design choices.)

The three largest approved projects today, (LIGO,[72] VIRGO,[73] and GEO[74]) all went through similar parallel processes of design study, proposal, and now construction. This was a fruitful period, with a rich interchange of ideas. For pedagogical purposes, we choose to focus in this review on a single line of development, that of the U.S. LIGO Project.

6.7.1 The "Blue Book"

In almost the same sense as the early table-top interferometers were prototypes of larger instruments, so too did the proposals for kilometer-scale interferometers have a prototype. This was the report called "A Study of a Long Baseline Gravitational Wave An-

tenna System," submitted to the U.S. National Science Foundation in October 1983.[75] (It has since its presentation been called the "Blue Book" because of the color of the cheap paper cover in which it was bound.) It was prepared primarily by Weiss and two colleagues at MIT (Paul S. Linsay and the present author), as the product of a planning exercise funded by the NSF starting in 1981. The report also contained a section by Stan Whitcomb of Caltech on Fabry-Perot systems (as a partial counter to Weiss' emphasis on Herriott delay lines), as well as extensive sections written by industrial consultants from Stone & Webster Engineering Corporation and from Arthur D. Little, Inc. These latter contributors were essential, because this document contains, for the first time anywhere, an extensive discussion of the engineering details specific to the problems of the construction and siting of a large interferometer. The report was presented, by both Weiss's MIT group and that of Drever at Caltech, at a meeting of the NSF's Advisory Council for Physics late in 1983. While not a formal proposal, it served as a sort of "white paper," suggesting the directions that subsequent proposals might (and in large measure did) take.

The first half of the report is devoted to the physics of gravitational wave interferometers. This section reads much like Weiss' 1972 design study, except that many issues only touched on briefly in the first paper are here discussed at substantially greater length. In the 11 years that elapsed between the two documents, there had been real progress on several fronts. There are chapters on sources of gravitational waves, the basic physics of the response of a free-mass interferometer to a gravitational wave, a discussion of beam-folding schemes and a summary of the current prototype interferometers, and another extensive discussion of noise sources. The report is bracketed by an introductory section outlining a history of the field to 1983 and by a pair of appendices, one of which compares the quantum limits of bars and interferometers and the other showing why the interferometer beams must travel through an evacuated space instead of through optical fibers.

The main emphasis of the Blue Book was less a discussion of physics *per se* than it was a consideration of the practical aspects of the experiment as an engineering and construction project. The completely new material appears in the second half of the Blue Book, in the chapters summarizing the work of the industrial consultants. Weiss believed that the only significant impediment to achieving astrophysically interesting sensitivity was the expense of building an interferometer with long arms (the issue that had brought the Hughes group's progress to a halt). The industrial study was undertaken with the aim of identifying what design trade-offs would allow for a large system to be built at minimum cost, and to establish a rough estimate of that cost (along with cost scaling laws) so that the NSF could consider whether it might be feasible to proceed with a full-scale project.

Before such an engineering exercise could be meaningful, though, it was necessary to define what was meant by "full-scale." The Blue Book approaches this question by first modeling the total noise budget as a function of frequency, then evaluating the model as a function of arm lengths ranging from 50 meters (not much longer than the Caltech prototype) to 50 km. The design space embodied in this model was then explored in a process guided by three principles:

- "The antenna should not be so small that the fundamental limits of performance cannot be attained with realistic estimates of technical capability." This was taken to mean that the length ought to be long enough that one could achieve shot noise limited performance for laser power of 100 W, without being limited instead by displacement noise sources, over a band of interesting frequencies. The length resulting from this criterion strongly depended on whether one took that band to begin around 1 kHz (in which case $L = 500$ m was adequate), 100 Hz (where $L = 5$ km was only approaching the required length), or lower still (in which case even $L = 50$ km would not suffice). Evidently, this strictly physics-based criterion was too elastic to be definitive.
- "The scale of the system should be large enough so that further improvement of the performance by a significant factor requires cost increments by a substantial factor." In other words, the system should be long enough so that the cost is not dominated by the length-independent costs of the remote installation.
- "Within reason no choice in external parameters of the present antenna design should preclude future internal design changes which, with advances in technology, will substantially improve performance." This was a justification for investing in a large-diameter beam tube, and for making sure that the vacuum system could achieve pressures as low as 10^{-8} torr.

In an iterative process, rough application of these principles was used to set the scope of options explored by the industrial consultants. Then at the end of the process, the principles were used again to select a preferred design. Arm lengths as long as 10 km were explored, and tube diameters as large as 48 inches. An extensive site survey was also carried out by the consultants. It was aimed at establishing that sites existed that were suitable for a trenched installation (which put stringent requirements

on flatness of the ground) of a 5 km interferometer. The survey covered federal land across the United States, and a study of maps of all land in the Northeastern United States, along with North Carolina, Colorado, and Nebraska. Thirteen "suitable" sites were identified. Evaluation criteria also included land use (specifically that the site not be crossed by roads, railroads, or oil and gas pipelines), earthquake risk, drainage, and accessibility.

The site survey also attempted to identify possibilities of locating an interferometer in a subsurface mine, which would give a more stable thermal environment and perhaps also reduced seismic noise (if it were located deep enough, and if it were inactive). No mines were found in the United States with two straight orthogonal tunnels even 2 km in length.

The conclusion of the exercise was a "proposed design" with the following features:

- Two interferometer installations separated by "continental" distances.
- Interferometer arm length of $L = 5$ km.
- Beam tubes of 48-inch diameter made of aluminum (chosen for an expected cost savings over stainless steel) pumped by a combination of Roots-blowers for roughing and ion-pumps for achieving and maintaining the high vacuum. A delay line interferometer would require a diameter of almost the proposed size. Drever's beam-interchange scheme for improved narrow-band sensitivity was also listed as one justification for preferring large tubes, as was the possibility of multiple interferometers (presumably based on Fabry-Perot cavities) side by side.
- The proposed installation method was to enclose the tube in a 7' by 12' cover constructed of a "multi-plate pipe-arch," in turn installed four feet below grade in a trench that was subsequently back-filled with soil.

The total estimated cost for such a system was given as $58M.

Note that there were no specific recommendations for the design of the interferometers themselves, beyond the "straw man" used for estimating the noise budget.

The Blue Book was received respectfully by the NSF's Advisory Committee for Physics. As a result, the MIT and Caltech research groups were encouraged to combine their forces to develop a complete specific design. Subsequently, both groups received funding with the eventual goal of a joint proposal for construction of a large interferometer system.

6.7.2 The LIGO Proposal of December 1987

By 1987, substantial progress had been made on lab-scale interferometers by research groups around the world. And, encouraged by the NSF's reception of the Blue Book, more thinking had gone into the best way to construct and exploit large interferometers. Much of this progress is evident in the proposal submitted to the U.S. National Science Foundation in December of 1987 by a formal Caltech-MIT collaboration that had adopted the name of LIGO, for Laser Interferometer Gravitational wave Observatory.[76] Since September 1987 it has been led by Rochus E. ("Robbie") Vogt as Project Director, with Drever and Weiss as science team leaders. The proposal requested funds for a three-year program of R&D and engineering studies, the outcome of which was intended to be another proposal (to be submitted in 1989) requesting authorization to build a pair of 4-km long interferometers.

Prototype interferometers were functioning at respectable sensitivities, after years of assembly, debugging, analysis, and redesign. The Caltech 40-meter interferometer was the showpiece of the proposal. It employed high-finesse Fabry-Perot cavities, arranged in the simplified non-recombined configuration for ease of testing. An impressive graph shows the improvement in strain sensitivity since its first operation in May 1983, by a factor of over 10^3 in amplitude. In that interval, ultra-low loss "supermirrors" were installed, first on an early optical-bench style of test mass and then on separated compact test masses (eventually ones made of fused silica), and a variety of improvements were made to the Argon laser and to the locking servos. The noise spectrum above 1 kHz was nearly white, with a level of $h(f) \approx 2 \times 10^{-19}/\sqrt{\text{Hz}}$. This was consistent with the expected level of shot noise.

Results were also presented from the table-top (1.5 meter arm length) prototype interferometer at MIT. It employed Herriott delay lines of 56 bounces between conventional high quality mirrors that were clamped to aluminum test masses. Unlike the Caltech instrument which could not function without sophisticated frequency stabilization of its laser, MIT used an unstabilized Argon laser to which phase noise was actually added to help suppress scattered light. Pointing, damping, and slow feedback were accomplished with electrostatic actuators. Strain sensitivity could of course not compete with the much longer Caltech instrument, but even the displacement noise was nearly an order of magnitude worse, given as 4.6×10^{-17} m$/\sqrt{\text{Hz}}$, for frequencies above 4 kHz. That level was a factor of two in excess of the expected shot noise in the 60 mW of light, diagnosed as insufficiently suppressed noise from scattered light. At lower fre-

quencies acoustic noise drove the interferometer via a variety of coupling paths through the injection optics as well as the test mass suspensions.

The proposal records substantial progress toward design of a full-scale interferometer. It states that the collaboration had adopted the Fabry-Perot beam-folding system. A preliminary design is presented in an appendix of the proposal. It envisioned use of 5 to 6 W of light at 514 nm from an Argon laser employed in a power recycled configuration. An elaborate schematic diagram gave a hint of the complexity of the servos necessary to control the large number degrees of freedom that need to be kept locked for such an instrument to function. These include lengths of the arm cavities, the separation of their input mirrors from the beam splitter, the location of the power recycling mirror, and the lengths of various "mode cleaning" resonant cavities used for spatial filtering of the laser beam. In addition to these lengths, control of a number of angular degrees of freedom also needs to be included. Four separate RF modulation frequencies are specified to drive these servos. Special features are designed into the main cavity locking servo so that the phase modulation can be injected with a small Pockels cell without the inevitable losses dominating the performance of the recycling system.

Another appendix describes an alternate optical configuration based on Herriott delay lines. It employed 86 cm diameter silicon test masses of 450 kg. The simplicity of the servos was listed as one of its major advantages. A "closed-path" variation of this design was also presented, in which light leaving one arm is injected into the other. This is like a single-interchange version of Drever's system for improving sensitivity to periodic waves. Here it was employed mainly to relax the tolerances on matching the curvature of the large mirrors. (The virtues of this design have recently been explored again by the Stanford group.[77])

On the engineering and site issues, there had also been some progress since the Blue Book study, mostly made by engineers at Caltech's Jet Propulsion Laboratory, but this was not considered complete enough to highlight in the proposal. Instead, one of the first proposed tasks was to complete a preliminary engineering design. Nevertheless, a mature understanding had been achieved of what LIGO ought to *be*. This insight was expressed in a list of "Essential Features of the LIGO":

1. "Two widely separated sites under common management." Two sites had been a feature of Weiss' earliest thinking, to allow coincidence observations to search for transient signals. The new feature was the commitment to truly have them managed as a single entity, "to guarantee that two receivers of nearly equal sensitivity are on line simultaneously at two sites, with a high live time."
2. "Arm lengths of order 4 kilometers at each site," a slight scaling back of the 5 km considered previously, but still long enough to strongly dilute the effects of displacement noise.
3. "The ability to operate simultaneously several receiver systems at each site." In a way, this was the most ambitious feature of the LIGO concept. In part it grew out of a kind of conservatism that was not clearly spelled out, but that was nevertheless real. The early LIGO interferometers, if they were not to be extremely risky extrapolations from known technology, were unlikely to have sufficient sensitivity to be assured of detecting astrophysical signals. Even if that weren't so, the project would have had to wrestle with the competition between time devoted to observation and time devoted to improving the performance of the instrument. This competition had bedeviled workers on resonant-mass detectors. The key new idea for LIGO was that the precious commodity, an evacuated beam pipe, might be available with abundant cross-sectional area since the Fabry-Perot geometry had been adopted. All that was required was an arrangement of tanks at the ends of the pipe to install the test masses of various interferometers, both operational and experimental. This actually called for substantial cleverness in developing an airlock system, so that installation and operation could take place with "a minimum of mutual interference."
4. "The capability for receivers of two different arm lengths." Drever urged the adoption of this feature, to allow a clean test of the gravitational origin of candidate signals, which should show up as the tidal signature that a longer interferometer sees twice the signal.
5. "A vacuum tube diameter of order 48 inches." This had the conservative justification that it would be necessary if one had to switch from Fabry-Perot cavities to delay lines, and the great benefit of allowing multiple Fabry-Perot interferometers, as mentioned above.
6. "The capability of a vacuum level of 10^{-8} torr." This would be needed, not for the first LIGO instrument, but to avoid having fluctuations in the index of refraction of the residual gas dominate shot noise in a more sensitive "advanced" receiver.
7. "A minimum lifetime of the facilities of 20 years." This was to be not just a one-shot discovery experiment, but a laboratory to exploit the gravitational wave

window in astronomy.

8. "Adequate support instrumentation."

The heritage of the Blue Book should be evident in the above list, but so too should be the progress in thinking beyond that point.

Because of a Federal budget crisis, this proposal was not funded. However, the two groups were encouraged to continue their work, and to submit a more complete proposal in a subsequent year.

6.7.3 The LIGO Proposal of December 1989

Many of the features only hinted at in the 1987 proposal are spelled out in much greater detail in the proposal for engineering design and construction of LIGO (now with a hyphen in Laser Interferometer Gravitational-Wave Observatory) that was submitted to the NSF in December 1989.[78] It fleshes out the design of a system that would embody the eight Essential Features of LIGO first described in the 1987 proposal. (The only differences are that the vacuum tube diameter is now specified as a "clear optical diameter of approximately 1 meter," and the vacuum spec is given as "10^{-9} torr of hydrogen and 10^{-10} of other gases.")

A single detector system would require three interferometers—a 4 km interferometer at each of two widely separated remote sites, plus a 2 km interferometer at one of the sites. This shorter interferometer played two related roles: a check that a candidate signal had the tidal signature of a gravitational wave, and the simpler but crucial requirement that a real signal should appear in all three interferometers. A calculation presented here shows that an accidental event rate of around 100/hour/interferometer can be tolerated without accidental three-way coincidences occurring more frequently than once in ten years, an improvement of about two orders of magnitude over what could be tolerated with only a single interferometer at each site.

The features described above are part of a plan that is aimed at accomplishing "three primary objectives":

- "observation," or a continuous "gravitational-wave watch,"
- "development," or "full functional testing of new and advanced interferometer-based detector concepts," and
- "special investigations" using detectors optimized for "particular phenomena."

These missions are "to be conducted without mutual interference."

It would take a substantial investment (of money, ideas, and energy of scientists) to accomplish all of these goals, and so for this reason full implementation of this strategy was to be accomplished in a series of phases.

- "Phase A, The Exploration/Discovery Phase," with one three-interferometer detector system, suitable for observation or development, but not both simultaneously.
- "Phase B, The Discovery/Observation Phase," with two three-interferometer detector systems, allowing "concurrent observation *and* development or specialized search."
- "Phase C, The Observatory Phase," room for three full detector systems, "allows concurrent observation, development, special investigations, and optimal access for the scientific community at large. It completes the LIGO evolution to its presently conceived full-design capability."

Because of the cost involved in elaborate vacuum chambers with airlocks, the 1989 proposal asked only for the funds to complete Phase A. The single important investment in the capability to upgrade to Phases B and C was the design of buildings large enough to accommodate all of the vacuum tanks that would be eventually required.

It should be emphasized that, to a large degree, this aggressive planning for an elaborate facility was a necessary consequence of a simple fact—that it looked difficult to build an interferometer that would have sufficient sensitivity to assure detection of gravitational waves. Hence the need to plan for ongoing interferometer development and specially optimized instruments. Of course, these activities would also be useful even if some signals proved easier to detect than expected, since carrying out gravitational wave astronomy would call for the highest achievable signal to noise ratios. For example, angular position errors are inversely proportional to the signal to noise ratio, and are as large as ten arcmin or more when the SNR = 10, even when observations are made with a three-detector U.S.-Europe network.

The 1989 proposal is much more explicit about the details of the design of the first interferometer. The design is based on the Fabry-Perot interferometer of the 1987 proposal's appendix. It has been fleshed out with specifications of laser power, finesse, test mass parameters, and vibration isolation performance, so that a specific noise budget could be presented. The 1989 proposal also contains a preliminary discussion of what sorts of improvements would be necessary to push the noise to levels low enough

to guarantee detection of signals. This includes laser power of 60 W recycled by a factor of 100, a much more aggressive vibration isolation system, and final pendulum suspensions with a quality factor of 10^9 carrying one-ton fused silica mirrors.

6.7.4 The Situation Today

Construction of LIGO was approved in 1991. By mid-1998 (the time of the writing of this review), construction of the two facilities in Hanford, WA and Livingston, LA was over three-quarters complete. The schedule calls for construction to be completed soon. Roughly speaking, 1999 is to be devoted to installation of the scientific equipment in the completed facilities, 2000 to shakedown of the interferometers, and 2001 to engineering activities to bring the performance up to the design specifications. Then, data will be collected during 2002-3. Beginning in 2004, upgrades to improve the performance will be carried out, interspersed with additional periods of observation.

The first instrument to be installed is expected to have a noise spectrum like that shown in Figure 4. The high frequency noise spectrum should be dominated by shot noise, as determined from an input power of 6 W, multiplied by a power recycling gain of 30. Thermal noise from the 1 Hz pendulum mode will dominate the intermediate frequency band; the oscillations of the 10 kg test masses should achieve a quality factor of 1.6×10^5. (Internal thermal noise will dominate the spectrum only in a narrow band, due to test mass modal quality factors of about 10^6.) Low frequency noise will be governed of course by seismic noise that passes through the multi-layer stack.

Performance of the VIRGO 3 km interferometer will be similar at medium and high frequencies. At low frequencies, seismic noise should be much lower in VIRGO than in LIGO, since a much more aggressive filter has been designed. This should allow the noise to be dominated by pendulum thermal noise down to 10 Hz or a bit lower.

The GEO 600 meter interferometer is not expected to reach quite such low levels, but it will be surprisingly close. To make up for the shorter length, advanced technologies (including signal recycling) will be pursued aggressively from an early date. Thus this instrument will play a dual role as part of the global network of interferometers and as a prototype for features that will later be incorporated into other interferometers.

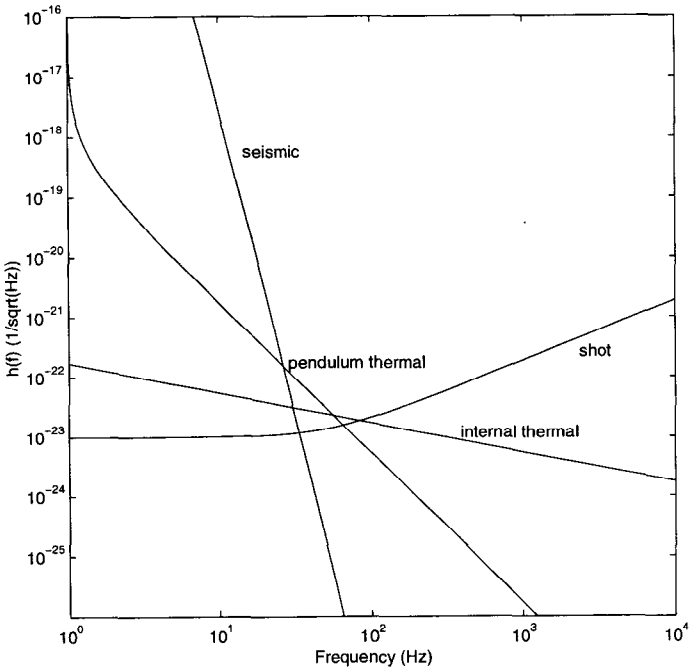


Fig. 4. An estimate of the noise spectrum of the LIGO I interferometers. The four most important noise sources are shown: seismic noise, pendulum mode thermal noise, thermal noise of internal vibrations of the mirror, and shot noise.

References

[1] E. M. Purcell, *Electricity and Magnetism* (McGraw-Hill, New York, 1965).

[2] L. D. Landau and E. M. Lifschitz, *The Classical Theory of Fields* (Pergamon, Oxford, 1975).

[3] P. Kafka, in *Space Science and Fundamental Physics, Proceedings of the Summer School*, Alpbach, Austria, 28 July - 7 August 1987 (ESA SP-283, May 1988).

[4] A. P. Cowley, *Ann. Rev. Astron. Astrophys.* **30**, 287 (1992).

[5] J. A. Wheeler, cited in Weber (1961).

[6] J. Weber, *Phys. Rev.* **117**, 306 (1960).

[7] J. Weber, *General Relativity and Gravitational Waves* (Interscience, New York, 1961).

[8] J. Weber, *Phys. Rev. Lett.* **22**, 1320 (1969).

[9] F. A. E. Pirani, *Acta. Phys. Polon.* **15**, 389 (1956).

[10] G. E. Moss, L. R. Miller, and R. L. Forward, *Appl. Opt.* **10**, 2495 (1971).

[11] R. Weiss, *Quarterly Progress Report, MIT Research Lab of Electronics* **105**, 54 (1972).

[12] E. Kappler, *Ann. Phys. (Leipzig)* **11**, 233 (1931); R. B. Barnes and S. Silverman, *Rev. Mod. Phys.* **6**, 162 (1934).

[13] J. B. Johnson, *Phys. Rev.* **32**, 97 (1928).

[14] H. Nyquist, *Phys. Rev.* **32**, 110 (1928).

[15] H. B. Callen and T. A. Welton, *Phys. Rev.* **83**, 34 (1951); H. B. Callen and R. F. Greene, *Phys. Rev.* **86**, 702 (1952).

[16] N. Bohr, *Nature* **121**, 580 (1928).

[17] N. Bohr, in *Albert Einstein: Philosopher-Scientist* (Library of Living Philosophers, Evanston, Ill., 1949).

[18] R. Weiss, in *Sources of Gravitational Radiation*, ed. L. Smarr (Cambridge University Press, Cambridge, 1979), p. 7.

[19] P. A. M. Dirac, *The Principles of Quantum Mechanics* (Clarendon Press, Oxford, 1958).

[20] C. M. Caves, *Phys. Rev. Lett.* **45**, 75 (1980).

[21] J. Weber, *Phys. Rev. Lett.* **17**, 1228 (1966).

[22] R. Del Fabbro, A. diVirgilio, A. Giazotto, H. Kautzky, V. Montelatici, and D. Passuello, *Phys. Lett. A* **132**, 237 (1988).

[23] J. L. Levine and R. L. Garwin, *Phys. Rev. Lett.* **31**, 173 (1973).

[24] E. Amaldi, O. Aguiar, M. Bassan, P. Bonifazi, P. Carelli, M. G. Castellano, G. Cavallari, E. Coccia, C. Cosmelli, W. M. Fairbank, S. Frasca, V. Foglietti, R. Habel, W. O. Hamilton, J. Henderson, W. Johnson, K. R. Lane, A. G. Mann, M. S. McAshan, P. F. Michelson, I. Modena, G. V. Pallottino, G. Pizzella, J. C. Price, R. Rapagnani, F. Ricci, N. Solomonson, T. R. Stevenson, R. C. Taber, and B.-X. Xu, *Astron. Astrophys.* **216**, 325 (1989).

[25] G. W. Gibbons and S. W. Hawking, *Phys. Rev. D* **4**, 2191 (1971).

[26] H. J. Paik, *J. Appl. Phys.* **47**, 1168 (1976).

[27] J. P. Richard and W. M. Folkner, in *The Detection of Gravitational Waves*, ed. D. G. Blair (Cambridge University Press, Cambridge, 1991).

[28] V. B. Braginsky and Yu. I. Vorontsov, *Sov. Phys. Uspekhi* **17**, 644 (1974).

[29] M. E. Tobar and D. G. Blair, *J. Phys. D: Appl. Phys.* **26**, 2276 (1993).

[30] P. R. Saulson, *Class. Quantum Grav.* **14** , 2435 (1997).

[31] H. Rothe and W. Dahlke, *Proc. IRE* **44**, 811 (1956).

[32] W. R. LePage, *Complex Variables and the Laplace Transform for Engineers* (McGraw-Hill, New York, 1961).

[33] J.-P. Richard, *Phys. Rev. Lett.* **52**, 165 (1984); J. C. Price, *Phys. Rev. D* **36**, 3555 (1987).

[34] R. W. P. Drever, J. Hough, R. Bland, and G. W. Lessnoff, *Nature* **246**, 340 (1973).

[35] H. Heffner, *Proc. IRE* **50**, 1604 (1962).

[36] R. P. Giffard, *Phys. Rev. D* **14**, 2478 (1976).

[37] J. Weber, *Phys. Rev. Lett.* **24**, 276 (1970).

[38] J. Weber, *Phys. Rev. Lett.* **25**, 180 (1970).

[39] W. H. Press and K. S. Thorne, *Ann. Rev. Astron. Astrophys.* **10**, 334 (1972).

[40] P. S. Aplin, *Gen. Rel. Grav.* **3**, 111 (1972).

[41] J. Weber, V. Hughes, P. Kafka, R. W. P. Drever, and C. W. Misner, *Ann. N.Y. Acad. Sci.* **224**, 101 (1973).

[42] D. Bramanti and K. Maischberger, *Lett. Nuovo Cimento* **4** , 1007 (1972); D. Bramanti, K. Maischberger, and D. Parkinson, *Lett. Nuovo Cimento* **7**, 665 (1973).

[43] H. Billing, P. Kafka, K. Maischberger, F. Meyer, and W. Winkler, *Lett. Nuovo Cimento* **12**, 111 (1975).

[44] J. A. Tyson, *Ann. N.Y. Acad. Sci.* **224**, 74 (1973).

[45] J. L. Levine and R. L. Garwin, *Phys. Rev. Lett.* **31**, 173 (1973); R. L. Garwin and J. L. Levine, *Phys. Rev. Lett.* **31**, 176 (1973); J. L. Levine and R. L. Garwin, *Phys. Rev. Lett.* **33**, 794 (1974).

[46] J. L. Levine and R. L. Garwin (1974), cited above.

[47] R. L. Garwin, *Physics Today* , December 1974, p. 9.

[48] D. Sciama, J. Weber, P. Kafka, R. W. P. Drever, J. A. Tyson, in *General Relativity and Gravitation*, Proceedings of GR7, Tel Aviv University, June 23-28, 1974, eds. G. Shaviv and J. Rosen (Wiley, New York 1974).

[49] H. M. Collins, *Changing Order* (Sage Publications, London, 1985). See also this exchange: A. Franklin, *Stud. Hist. Phil. Sci.* **25**, 463 (1994); H. M. Collins, *Stud. Hist. Phil. Sci.* **25**, 493 (1994).

[50] R. W. P. Drever, *Q. J. R. Astr. Soc.* **18**, 9 (1977).

[51] J. A. Tyson and R. P. Giffard, *Ann. Rev. Astron. Astrophys.* **16**, 521 (1978).

[52] P. Kafka and L. Schnupp, *Astron. Astrophys.* **70**, 97 (1978).

[53] S. P. Boughn, W. M. Fairbank, R. P. Giffard, J. N. Hollenhorst, E. R. Mapoles, M. S. McAshan, P. F. Michelson, H. J. Paik, and R. C. Taber, *Ap. J. Lett.* **261**, L19 (1982).

[54] P. Astone *et al.*, *Proceedings of the Xth Italian General Relativity Conference*, Bardonecchia, September 1-4, 1992, unpublished.

[55] M. Aglietta, A. Castellina, W. Fulgione, G. Trinchero, S. Vernetto, P. Astone, G. Badino, G. Bologna, M. Bassan, E. Coccia, I. Modena, P. Bonifazi, M. G. Castellano, M. Visco, C. Castagnoli, P. Galeotti, O. Saavedra, C. Cosmelli, S. Frasca, G. V. Pallottino, G. Pizzella, P. Rapagnani, F. Ricci, E. Majorana, D. Gretz, J. Weber, and G. Wilmot, *Nuovo Cimento* **106B**, 1257 (1991); G. Pizzella, in *Developments in General Relativity, Astrophysics and Quantum Theory: A Jubilee in Honour of Nathan Rosen*, eds. F. Cooperstock, L. P. Horwitz, and J. Rosen, *Annals of the Israel Physical Society* **9**, 195 (1990).

[56] C. A. Dickson and B. F. Schutz, *Phys. Rev. D* **51**, 2644 (1995).

[57] M. E. Gertsenshtein and V. I. Pustovoit, *Sov. Phys. JETP* **16**, 433 (1962).

[58] P. G. Roll, R. Krotkov, and R. H. Dicke, *Ann. Phys. (N.Y.)* **26**, 442 (1964).

[59] A. A. Michelson and E. W. Morley, *Am. J. Sci.* **34**, 333 (1887).

[60] R. H. Dicke, *Rev. Sci. Instrum.* **17**, 268 (1946).

[61] R. Buderi, *The Invention That Changed the World* (Simon and Schuster, New York, 1996).

[62] J. D. Kraus, *Radio Astronomy* (McGraw-Hill, New York, 1966).

[63] R. H. Dicke, *The Theoretical Significance of Experimental Relativity* (Gordon and Breach, New York, 1964).

[64] R. L. Forward, *Phys. Rev. D* **17**, 379 (1978).

[65] R. W. P. Drever, in *Gravitational Radiation*, eds. N. Deruelle and T. Piran (North Holland, Amsterdam, 1983), p. 321.

[66] R. W. Drever, J. L. Hall, F. V. Kowalski, J. Hough, G. M. Ford, A. J. Munley, and H. Ward, *Appl. Phys. B* **31**, 97 (1983); see also A. Schenzle, R. DeVoe, and G. Brewer, *Phys. Rev. A* **25**, 2606 (1982).

[67] R. V. Pound, *Rev. Sci. Instrum.* **17**, 490 (1946).

[68] B. J. Meers, *Phys. Rev. D* **38**, 2317 (1988); K. A. Strain and B. J. Meers, *Phys. Rev. Lett.* **66**, 1391 (1991).

[69] D. Shoemaker, R. Schilling, L. Schnupp, W. Winkler, K. Maischberger, and A. Rüdiger, *Phys. Rev. D* **38**, 423 (1988).

[70] P. F. Michelson, E.-K. Hu, and G. Pizzella, editors, *International Symposium on Experimental Gravitational Physics* (World Scientific, Singapore, 1988).

[71] D. Nicholson, C. A. Dickson, W. J. Watkins, B. F. Schutz, J. Shuttleworth, G. S. Jones, D. I. Robertson, N. L Mackenzie, K. A. Strain, B. J. Meers, G. P. Newton, H. Ward, C. A. Cantley, N. A. Robertson, J. Hough, K. Danzmann, T. M. Niebauer, A. Rüdiger, R. Schilling, L. Schnupp, and W. Winkler, *Phys. Lett. A* **218**, 175 (1996).

[72] A. Abramovici, W. E. Althouse, R. W. P. Drever, Y. Gürsel, S. Kawamura, F. J. Raab, D. Shoemaker, L. Sievers, R. E. Spero, K. S. Thorne, R. E. Vogt, R. Weiss, S. E. Whitcomb, and M. E. Zucker, *Science* **256**, 325 (1992).

[73] C. Bradaschia, R. Del Fabbro, A. Di Virgilio, A. Giazotto, H. Kautzky, V. Montelatici, D. Passuello, A. Brillet, O. Cregut, P. Hello, C. N. Man, P. T. Manh, A. Marraud, D. Shoemaker, J.-Y. Vinet, F. Barone, L. Di Fiore, L. Milano, G. Russo, S. Solimeno, J. M. Aguirregabiria, H. Bel, J.-P. Duruisseau, G. Le Denmat, P. Tourrenc, M. Capozzi, M. Longo, M. Lops, I. Pinto, G. Rotoli, T. Damour, S. Bonazzola, J. A. Marck, Y. Gourghoulon, L. E. Holloway, F. Fuligni, V. Iafolla, G. Natale, *Nucl. Instr. Meth. Phys. Res. A* **289**, 518 (1990).

[74] J. Hough, G. P. Newton, N. A. Robertson, H. Ward, A. M. Campbell, J. E. Logan, D. I. Robertson, K. A. Strain, K. Danzmann, H. Lück, A. Rüdiger, R. Schilling, M. Schrempel, W. Winkler, J. R. J. Bennett, V. Kose, M. Kühne, B. F. Schutz, D. Nicholson, J. Shuttleworth, H. Welling, P. Aufmuth, R. Rinkleff, A. Tünnermann, and B. Willke, in *Proceedings of the Seventh Marcel Grossmann Meeting on General Relativity*, eds. R. T. Jantzen and G. M. Keiser (World Scientific, Singapore, 1996), p. 1352.

[75] P. S. Linsay, P. R. Saulson, and R. Weiss, *A Study of a Long Baseline Gravitational Wave Antenna System* (MIT, Cambridge, Mass., 1983), unpublished.

[76] R. E. Vogt, R. W. P. Drever, K. S. Thorne, and R. Weiss, *Caltech/MIT Project for a Laser Interferometer Gravitational Wave Observatory*, proposal to the National Science Foundation (Caltech, Pasadena, 1987), unpublished.

[77] K.-X. Sun, M. M. Fejer, E. Gustafson, and R. L. Byer, *Phys. Rev. Lett.* **76**, 3053 (1996).

[78] R. E. Vogt, R. W. P. Drever, F. J. Raab, K. S. Thorne, and R. Weiss, *The Construction, Operation, and Supporting Research and Development of a Laser Interferometer Gravitational-Wave Observatory*, proposal to the National Science Foundation (Caltech, Pasadena, 1989), unpublished.

Beyond the first Generation: Extending the Science Range of the Gravitational Wave Detector GEO 600.

Von der Fakultät für Mathematik und Physik
der Gottfried Wilhelm Leibniz Universität Hannover
zur Erlangung des Grades

Doktor der Naturwissenschaften
– Dr. rer. nat. –

genehmigte Dissertation von

Dipl.-Phys. Stefan Hild

geboren am 1. April 1978 in Celle.

2007

Referent:	Prof. K Danzmann
Korreferent:	Prof. R Schnabel
Tag der Promotion:	2. Februar 2007

Summary

During recent years, the first generation of large-scale laser-interferometric gravitational wave (GW) detectors has been commissioned and is now in operation. This worldwide network of gravitational wave observatories collects the most GW-sensitive data to date.

The gravitational wave detector GEO 600 near Hannover, is the first large-scale instrument already now using second generation technology, such as electro-static actuators and signal recycling. The use of signal recycling allows to improve the sensitivity but at the cost of a significantly more complex detector. A new method was developed that allowed, for the first time, the realization of tuned signal recycling in a large interferometer. In Chapter 2 a comparison of tuned and detuned signal recycling operation is given and related problems are discussed. It is found that the combination of heterodyne readout and detuned signal recycling is unfavourable in many respects. This supports the decisions to operate the LCGT detector with tuned signal recycling, and advanced LIGO in detuned signal recycling, but with a homodyne readout. The concept of detuned signal recycling with a DC-readout scheme might actually also have advantages for GEO 600 as described in Chapter 6.

One of the main noise sources encountered during the commissioning of 1st generation detectors is stray light. The actual sensitivity of GEO 600 can already be degraded by stray light contributions of the order 10^{-20} W. Since second generation GW detectors will operate with significantly higher light powers, and aim for increased sensitivities, stray light could be even more problematic. Chapter 3 describes the experience that was gained by studying stray light in GEO 600. A guide is given to help avoid, identify and eliminate stray light noise.

GEO 600 is not only an excellent test facility for second generation technologies, but also provides sensitive data with a high duty cycle. Currently a peak strain sensitivity of $2.5 \cdot 10^{-22}/\sqrt{\mathrm{Hz}}$ is achieved. In order to allow the data from GEO 600 to be used for multi-detector analysis, a high calibration accuracy is required. In an attempt to validate the official calibration routines, photon pressure calibrators are used in GEO 600 and LIGO. As shown in Chapter 4, several problems have been encountered during the commissioning of the GEO photon pressure calibrator. In particular, at frequencies above 1 kHz, a large discrepancy between the official calibration and that derived from the photon pressure calibrator was observed. This can be explained by photon pressure induced test mass deformation.

In Chapter 5 a new statistical veto method is presented employing an amplitude consistency check. This technique allows the derivation of safe statistical vetoes from interferometer channels which can contain traces of GW signal. This veto was applied to S5 data of the GEO 600 detector and was found to give veto efficiencies between 5 % and 20 % and a use-percentage of up to 80 %. This new veto method can easily be applied to the data from other GW detectors.

Key words: *Gravitational wave detector, signal recycling, stray light, photon pressure calibrator, statistical veto, DC-readout*

Zusammenfassung

Die großen laser-interferometrischen Gravitationswellendetektoren (GWD) der ersten Generation haben ihren Betrieb aufgenommen und bilden ein weltweites Netzwerk von Gravitationswellen-Observatorien. Die aufgenommenen Daten haben die beste Empfindlichkeit, die bisher erreicht wurde.

Der Gravitationswellendetektor GEO 600 bei Hannover ist das erste große Instrument, das bereits jetzt Technologien der zweiten Generation einsetzt, wie z.B. elektrostatische Aktuatoren oder Signal-Recycling (SR). Der Gebrauch von SR erhöht die Empfindlichkeit des Interferometers, steigert aber gleichzeitig auch die Komplexität des Systems. Eine neue Methode wurde entwickelt, mit der es erstmals möglich ist, tuned SR in einem großen Interferometer zu realisieren. In Kapitel 2 werden detuned und tuned SR verglichen und die damit verbundenen Probleme diskutiert. Es zeigt sich, dass eine Kombination aus Heterodyndetektion und detuned SR in mehrfacher Hinsicht unvorteilhaft ist. Dies bestätigt die Entscheidung, LCGT im tuned SR modus und Advanced LIGO mit einer Homodyndetektion zu betreiben. Die Konfiguration, detuned SR mit Homodyndetektion, kann auch für GEO 600 von Vorteil sein (siehe Kapitel 6).

Eine der Hauptrauschquellen der GWD der ersten Generation stellt Streulicht dar. Die Empfindlichkeit von GEO 600 kann schon von Streulichtbeiträgen in der Größenordnung von 10^{-20} W limitiert werden. Da die Detektoren der zweiten Generation mit deutlich größeren Lichtleistungen und gleichzeitig besserer Empfindlichkeit betrieben werden, kann die Streulichtproblematik in Zukunft sogar noch verstärkt auftreten. Kapitel 3 beschreibt die Erfahrungen, die in GEO 600 mit Streulicht gemacht wurden. Es werden Methoden beschrieben, wie Streulicht gefunden und eliminiert, bzw. vermieden werden kann.

Mit GEO 600 können nicht nur Technologien der zweiten Generation hervorragend getestet werden, sondern GEO 600 nimmt auch Daten mit hoher Empfindlichkeit und großem Dutycycle auf. Aktuell wird eine Empfindlichkeit von $2.5 \cdot 10^{-22}/\sqrt{\text{Hz}}$ erreicht. Damit GEO-Daten für Multi-Detektor-Analysen eingesetzt werden können, ist eine hohe Kalibrationsgenauigkeit erforderlich. GEO 600 und LIGO versuchen, Strahlungsdruckkalibratoren zu entwickeln, um die offiziellen Kalibrationen zu überprüfen. Wie in Kapitel 4 dargestellt wird, treten dabei aber mehrere Probleme auf. Besonders bei Frequenzen oberhalb von 1 kHz wird eine große Abweichung der Strahlungsdruckkalibration von der offiziellen Kalibration gefunden, die durch eine vom Strahlungsdruck induzierte Verformung der Testmasse erklärt werden kann.

In Kapitel 5 wird eine neue Methode zur Gewinnung statistischer Vetos beschrieben, die einen Amplitudenschwellwert benutzt. Diese Technik erlaubt das Erstellen von zuverlässigen Vetos auch aus Interferometerkanälen, die Spuren von Gravitationswellensignalen enthalten können. Die Vetomethode wurde auf Daten vom GEO-Detektor mehrfach angewendet und hat sich als leistungsstark erwiesen.

Schlüsselwörter: *Gravitationswellendetektor, Signal-Recycling, Streulicht, Strahlungsdruckkalibrator, Statistisches Veto, Homodyndetektion*

Contents

Summary i

Zusammenfassung iii

Table of Contents v

List of Figures ix

List of Tables xv

Abbreviations and Symbols xvii

1. Introduction 1

1.1. Detection of Gravitational waves 1

1.2. A world wide network of large-scale gravitational wave detectors 3

1.3. A brief description of the GEO 600 GW detector 3

2. Demonstration and comparison of tuned and detuned signal recycling 7

2.1. Introduction 7

2.2. Control scheme for signal recycling in GEO 600 11

2.3. A new method for locking tuned signal recycling 15

2.3.1. Simulation of the control parameters for tuned signal recycling . 16

2.3.2. Calibration of the actuator strength 18

2.4. Measurements of optical gain 18

2.5. Reduction of the RF amplitude modulation 21

2.6. Comparison of noise transfer functions for tuned and detuned signal recycling 21

2.7. Calibrated detector sensitivity for tuned signal recycling 24

2.8. Comparison of detuned signal recycling for upper and lower operating point 25
2.9. Summary 29

3. Scattered light problems in interferometric gravitational wave detectors 31
3.1. Introduction 31
3.2. Controlled stray light injections 33
3.2.1. How to produce stray light sowing up in H at 1 kHz 33
3.2.2. A device for controlled stray light generation 34
3.2.3. Low frequency large amplitude scenario 37
3.2.4. High frequency low amplitude scenario 37
3.2.5. Combination of the two scenarios 38
3.2.6. Summary of the stray light injections 39
3.3. Indications for a stray light limited sensitivity 39
3.4. The filter experiment: A reliable method to identify the source of stray light 43
3.4.1. Limitations of the filter experiment 45
3.5. Experience of stray light problems in GEO 46
3.6. The cat's eye effect 49
3.7. A special class of stray light noise: the "ripples" 52
3.8. Summary 56

4. Photon pressure calibration 59
4.1. Introduction 59
4.2. Experimental setup of the photon pressure calibrator 60
4.2.1. Measurement of the modulated light power 62
4.3. Injections with different amplitudes 64
4.4. Injections with different frequencies 66
4.5. Validation of the official calibration 69
4.5.1. Transfer function from ESD to H 70
4.5.2. ESD injections up to high frequencies 70
4.6. The phase picture 73
4.7. Non-rigidity of the test mass 74
4.7.1. Rough estimation of photon pressure calibrator induced test mass deformation 75

4.7.2. Finite element analysis of photon pressure calibrator induced test mass deformation . . . 76
4.7.3. Effective test mass displacement . . . 76
4.7.4. Comparison of pendulum response and mirror deformation . . . 79
4.7.5. High frequency injections using the photon pressure calibrator . 81
4.7.6. Influence of beam jitter to the photon pressure calibration . . . 83
4.7.7. Conclusion and consequences . . . 85
4.8. Effects from test mass rotation . . . 86
4.8.1. Method to determine the off centering of the photon pressure calibrator . . . 90
4.9. Determining the sign of H . . . 91
4.10. Summary . . . 94

5. A statistical veto method employing an amplitude consistency check 95
5.1. Introduction . . . 95
5.2. Motivation for a standard statistical veto for GEO 600: Hourly mains glitches . . . 97
5.3. Motivation for a statistical veto with amplitude consistency check: Dust falling through laser beams . . . 101
5.4. Description of the event trigger generator used: mHACR . . . 104
5.4.1. Estimation of burst parameter with mHACR . . . 105
5.5. The standard statistical veto . . . 107
5.5.1. The method in general . . . 107
5.5.2. Application of a standard statistical veto to GEO S5 data . . . 109
5.6. A statistical veto for channels containing GW information . . . 112
5.6.1. The method in general . . . 113
5.6.2. A 'real-world' scenario . . . 114
5.7. Application of a statistical veto with a amplitude consistency-check to GEO 600 data . . . 114
5.7.1. Data set 1: Full September 2006 with low dust concentration . . 116
5.7.2. Data set 2: 8 hours from May 2006 with high dust concentration 120
5.7.3. Performance of the veto analysis . . . 120
5.8. Summary . . . 122

6. Towards DC-readout for GEO 600 123

6.1. Introduction . . . 123
6.2. Motivation . . . 125
6.2.1. Shot noise . . . 127
6.2.2. Oscillator phase noise . . . 128
6.2.3. Electronic noise of the main photo diode . . . 128
6.3. Determination of the optimal dark fringe offset . . . 128
6.4. Technical realization . . . 133
6.5. Summary and outlook . . . 135

A. The suspended output telescope of GEO 600 **137**

B. Light fields at the output port for detuned signal recycling **143**

C. Design for an AR coated window for quadrant diodes **147**

D. Optical layout of GEO 600 **149**

E. Matlab script: Photon pressure calibration **151**

F. A gravity calibrator for GEO 600 **159**

G. Matlab script: Statistical veto with amplitude consistency-check **163**

H. FINESSE input file of GEO 600 for DC-readout **169**

I. Measurement of the bulk-absorption of the GEO 600 beam splitter **177**
I.1. Introduction . . . 177
I.2. A new method for measuring bulk absorption . . . 178
I.3. Setting an upper limit for the bulk absorption in the GEO 600 beam splitter 182
I.4. Conclusion . . . 183

Bibliography **185**

Acknowledgement **193**

Curriculum vitae **195**

Publications **197**

List of Figures

1.1. Sensitivities of the large-scale GW detectors 2
1.2. Simplified optical layout of GEO 600. 4
1.3. Sensitivity improvement of GEO 600 over the last few years. 5

2.1. Shot noise limit of GEO 600 for different transmittances of the signal recycling mirror. 8
2.2. Shot noise limit of GEO 600 for different tunings of the signal recycling cavity. 9
2.3. Sideband picture for tuned and detuned SR 10
2.4. Simplified control scheme for two longitudinal degrees of freedom of GEO 600 12
2.5. Signal recycling error signals versus microscopic position of MSR for various modulation frequencies 14
2.6. Simulations of longitudinal error signals for MI and SR loop for detuned and tuned SR 17
2.7. Simplified diagram of the loop controlling the differential arm length of the Michelson 19
2.8. Measurement of the optical gains for tuned and detuned signal recycling for the two orthogonal quadratures P and Q. For the tuned case the demodulation phase was not optimized. Therefore the the differential armlength signal is still clearly present in the optical gain of the Q-signal. 20
2.9. Measurements of the transfer functions from oscillator phase noise (OPN) and oscillator amplitude noise (OAN) to the error signal of the differential arm length servo 22
2.10. Measurements of the transfer functions from laser frequency noise and laser power noise to the error signal of the differential arm length servo 23

2.11. Comparison of the sensitivity of GEO 600 for tuned and detuned signal recycling operation 26
2.12. Comparison of the sensitivity of GEO 600 for two different operating points of detuned signal recycling 28

3.1. Stray light shoulders showing up in various interferometer configurations. 32
3.2. Controlled forced scattering: driving the stray light source with 1 Hz and amplitudes in the millimeter range. 35
3.3. Controlled forced scattering: driving the stray light source with 810 Hz and an amplitude of less than an laser wavelength. 36
3.4. Controlled forced scattered light generation for a medium frequency and amplitude 38
3.5. Spectrograms of the detector sensitivity for enhanced scattering and no scattering. 40
3.6. Spectra of the detector sensitivity for times of clean data and times of enhanced scattering. 41
3.7. Time series of the mean ASD in the band from 904 to 1005 Hz of the uncalibrated detector sensitivity with and without enhanced scattering. 42
3.8. Sketch of the filter experiment 44
3.9. Progression of the optical layout of the signal recycling bench in order to reduce stray light. 47
3.10. Sensitivity progress of GEO 600 from autumn 2004 48
3.11. Schematic of the cat's eye effect 50
3.12. Spectrum of the ripple noise 52
3.13. Spectrograms of H and the seismometer in the north end station for a time when the n*45 Hz ripple noise is present. 54
3.14. Slight frequency drift of the n*45 Hz ripple noise seen in H and the north seismometer. 55

4.1. Response function of the lowest stage of a GEO 600 main pendulum. . . 61
4.2. OPTOCAD layout of the optics in the north end station of the GEO 600 detector, including the setup of the photon pressure calibrator. 63
4.3. photon pressure calibratorinjections with different amplitudes. 65
4.4. Photon pressure calibrator injections for various frequencies. 67
4.5. Ratio and phase of official and photon pressure calibration. 68

4.6. Transfer function ESD to H 71
4.7. Measurement of the relative calibration accuracy of the official calibration for high frequencies 72
4.8. Calibration discrepancy: The phase picture. 73
4.9. Finite element model (FEM) of the deformation of a GEO test mass caused by a the beam of the photon pressure calibrator. 77
4.10. Overlap of photon pressure calibrator induced test mass deformation and the main interferometer beam. 78
4.11. Radial contribution to the effective displacement for photon pressure calibrator induced test mass deformation. 78
4.12. Simple model for the photon pressure calibrator taking into account the responses from the pendulum and from the mirror deformation effect. . 80
4.13. Measurement of the photon pressure calibrator response at high frequencies. 82
4.14. Effective displacement for different spatial overlaps of the beams from photon pressure calibrator and the main interferometer. 84
4.15. Effective longitudinal mirror displacement from mirror rotation 87
4.16. Phase picture for photon pressure calibrator induced test mass rotation. 88
4.17. Magnitude of the transfer function from PCALmon to H versus the position of the main interferometer beam at MFn. 90
4.18. Time series of the photon pressure calibrator signal (PCALmon) 92
4.19. Transfer function from photon pressure calibrator monitor diode (PCALmon) to H 93

5.1. Time series of the ripple control signals. 97
5.2. Spectra of the ripples control signals. 98
5.3. Time coincidence analysis for the GW channel and the mains monitor. . 100
5.4. Time coincidence analysis for the GW channel (H) and $\mathcal{P}_{\mathrm{DC}}$ for times of high and low dust concentration at the GEO output bench. 102
5.5. Coherence between $\mathcal{P}_{\mathrm{DC}}$ and sinusoidal differential arm length injections 103
5.6. Schematic of mHACR pipeline 106
5.7. Exemplary application of a statistical veto to a 24 hour stretch of GEO data using a GW-free channel as veto channel. 110
5.8. Exemplary application of a statistical veto including an additional coincidence condition for the central frequency of the events. 111

5.9. A schematic view of two different scenarios in which coincident transient events can appear in both the main GW channel, H, and an auxilliary channel, X. 112
5.10. Efficiency, background and significance of a veto using $\mathcal{P}_{\mathrm{DC}}$as veto channel for various sizes of the time and frequency coincidence window. 115
5.11. Standard deviation of the distribution of the mHACR errors plotted as a function of the snr of the triggers. 117
5.12. Application of a statistical veto with a amplitude consistency check to GEO 600 data for the entire month of September 2006. 118
5.13. The application of a statistical veto with a amplitude consistency check to 8 hours of GEO data from May 119
5.14. Schematic of the veto pipeline used for the analysis of Data Set 1 (full September 2006). 121

6.1. Operating points for heterodyne and DC readout. 124
6.2. Noise projections from a time of S5 126
6.3. Comparison of shot noise limited sensitivities of heterodyne and DC readout of GEO 600 in S5 configuration. 129
6.4. Simulated displacement sensitivity of GEO 600 with DC-readout versus the offset from the dark fringe. 130
6.5. Sensitivity achievable with DC-readout versus the modelled shot noise of S5 132
6.6. Optical layout of the detection bench. 134

A.1. Optical layout of the GEO 600 output telescope. 138
A.2. Suspension of BDO2 139
A.3. Suspension of BDO1 and BDO3. 140
A.4. Drawing of a BDO mirror. 141

B.1. Power measurement of the light fields present at the output port of GEO 600. 144

C.1. Front plate of the new quadrant diode encapsulation. 148
C.2. Back plate of the new quadrant diode encapsulation. 148

D.1. Optical layout of GEO 600 150

F.1. Geometry of a 3-mass gravity calibrator. 160

I.1. Simplified optical layout with and without thermal lensing. 179
I.2. Influence of thermal lensing to the output mode 180

List of Tables

2.1. A summary of the resonance condition of the important light fields at the dark-port for a positive and negative detuning of the signal recycling cavity. The given values indicate the distance of the light field from the center of the nearest resonance peak. 27

3.1. Summary of the normalized backscattering efficiencies for the three scenarios depicted in Figure 3.11 52

5.1. A summary of the results of applying a statistical veto with a amplitude consistency check to two different data sets from GEO 600. 120

A.1. Optical and mechanical parameters of the mirrors used for the output telescope. 142

Abbreviations and Symbols

AR	anti reflective
AS	amplitude spectrum
ASD	amplitude spectral densitiy
BLRMS	band limited rms
DAQS	data acquisition system
DC	originally: "direct current". In this work also used for frequencies very close to 0 Hz.
DER_DATA_H	official channel name of the main gravitational wave channel of GEO 600
diff_CAL	signal injected into the differential arm length servo for calibration
DW	dewhitening filter
ESD	electro static drives
ETG	event trigger generator
FEA	finite element analysis
FEM	finite element method
FFT	fast Fourier transform or discrete Fourier transform (DFT)
FWHM	full width at half maximum
FSR	free spectral range
GPS	global positioning system
GW	gravitational wave
H	main gravitational wave channel, h(t)
HR	high reflective
HVA	high voltage amplifier
IFO	interferometer
LCGT	large-scale cryogenic gravitational wave telescope
LIGO	laser interferometric gravitational wave observatory

LPN	laser power noise
LSC	LIGO scientific collaboration
MCe	end mirror of east arm
MCn	end mirror of north arm
MFe	folding mirror in the east building
MFn	folding mirror in the north building
MI	Michelson interferometer
MPR	power recycling mirror
MSR	signal recycling mirror
NS	neutron star
OAN	oscillator amplitude noise
OMC	output mode cleaner
OP	operational amplifier
OPN	oscillator phase noise
PCALmon	Record of the monitor of the photon pressure calibrator
$\mathcal{P}_{\mathrm{DC}}$	DC light power hitting the main photodiode
PEM_CBCTR_PWRGRID	official channel name of the mains monitor
PS	power spectrum
PSD	power spectral densitiy
PR	power recycling
PRC	power recycling cavity
PTB	Physikalisch technische Bundesanstalt (equivalent to NIST in the US)
RF	radio frequency
rms	root mean square
rot	misalignment in horizontal direction (rotation, yaw)
S4	S4 LSC science run
S5	S5 LSC science run
snr	signal to noise ratio
SR	signal recyling
SRC	signal recycling cavity
std	standard deviation
UTC	coordinated universal time
tilt	misalignment in vertical direction (tilt, pitch)

Chapter 1.

Introduction

1.1. Detection of Gravitational waves

The existence of gravitational waves (GW) is a consequence of the General Theory of Relativity, which was discovered and published by Albert Einstein in 1916. Supernovae, coalescing compact binary systems and pulsars are only a few representatives of astrophysical source, which can emit gravitational waves. Gravitational waves are expected to cover a wide frequency range, starting at the lower end at about 10^{-17} Hz and going up to frequencies as high as a few kHz. The strength of even the strongest signals, for instance a supernova in our own galaxy, will be very small at the earth and only cause a relative length change of about 10^{-21}. A good overview of gravitational wave sources can be found in [Cutler/Thorne].

The first efforts towards detecting gravitational waves were carried out by Joseph Weber in the 1960s. He used large metal cylinders, so called bar detectors or resonant detectors, as antennas [Weber]. While Weber and his first followers used the bar detectors at room temperature, later experimenters cooled their bar detectors down to liquid Helium temperatures to suppress the intrinsic thermal noise. An overview of the present status of the resonant detectors can be found in [Astone02].

Indirect evidence of the existence of gravitational waves was supplied by R. A. Hulse and J. H. Taylor [Hulse], [Taylor]. They observed a pulsar, in the binary star system PSR 1913+16 over years, and found a continuous increase of its rotation frequency. The observed increase exactly matched the increase predicted from emission of gravitational waves. For this work, Hulse and Taylor were awarded the Nobel price in 1993.

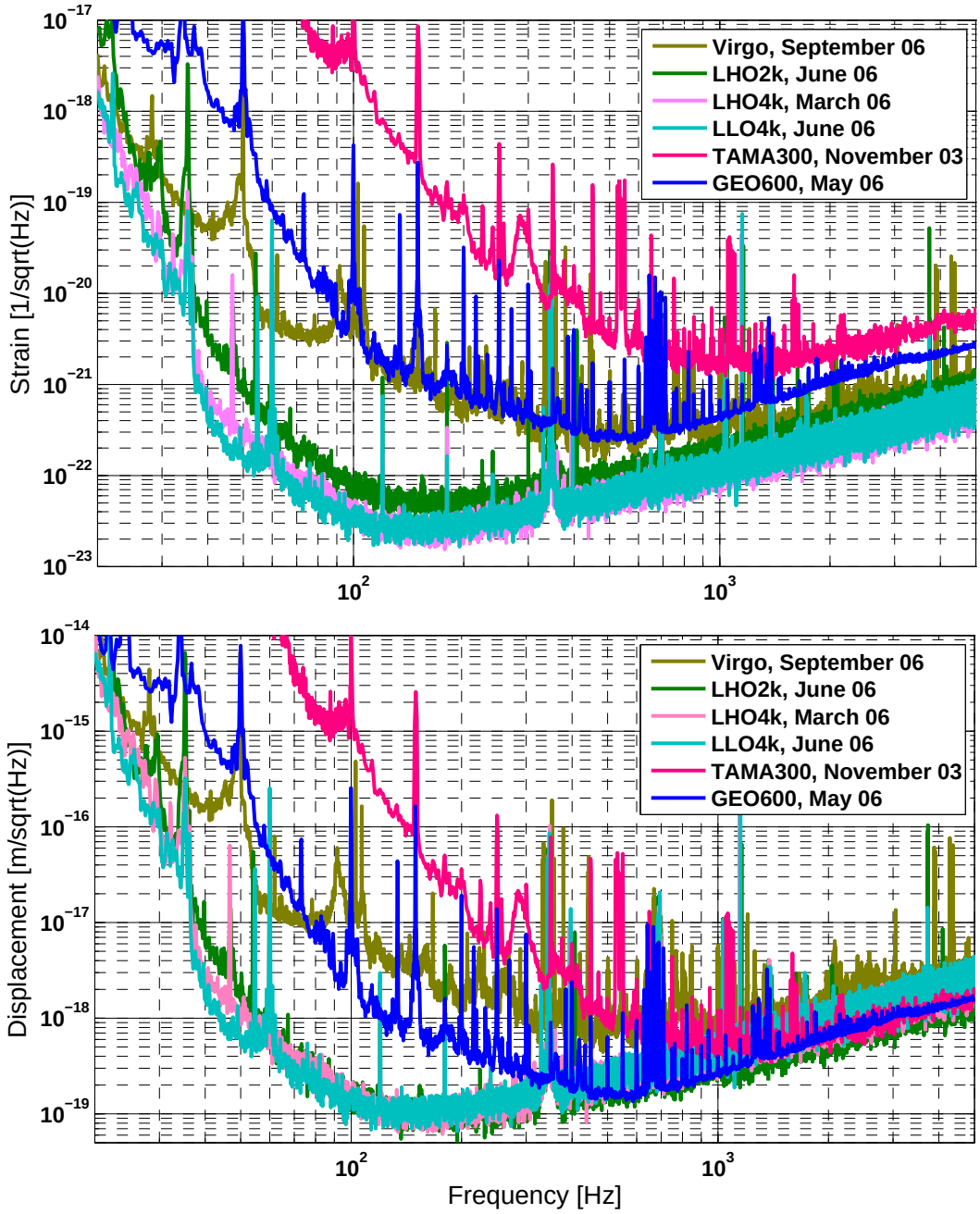


Figure 1.1.: Overview of the achieved sensitivities of the 6 large-scale laser-interferometric gravitational wave detectors. The upper subplot shows the strain sensitivities, while the lower plot depicts the corresponding displacement sensitivities, which are the sensitivities of the detectors to an absolute length change of the interferometer arms.

Today the direct search for gravitational waves concentrates on using large-scale laser interferometers, such as GEO 600. These extremely sensitive machines are based on kilometer long Michelson interferometers. Detailed descriptions of the measuring principle can be found [Rowan/Hough], [Saulson] and [Blair]. The currently operating generation of laser-interferometric GW detectors achieves sensitivities of the order $10^{-22}/\sqrt{\text{Hz}}$ and therefore it seems only to be a question of time until the first gravitational wave signal will be detected.

1.2. A world wide network of large-scale gravitational wave detectors

Currently six large-scale gravitational wave detectors are in operation. In the USA there are the three LIGO detectors [Waldman06]: One 4 km interferometer in Livingston (LLO), Louisiana and two detectors with arm lengths of 2 and 4 km in Hanford (LHO), Washington. An Italian-French Collaboration operates the 3 km long VIRGO detector [Acernese06]. The 300 m long Japananese detector TAMA300 is located in Tokio [Ando05]. Finally there is the 600 m long GEO 600 detector near Hannover, operated by a team of British and German scientists [Hild06c]. The sensitivities achieved by the detectors are shown in Figure 1.1.

Within the frame of LIGO Scientific Collaboration (LSC) GEO 600 and the three LIGO detectors have taken a huge amount of coincident data. These periods of simultaneous data taking are called *science runs*. The data of these 4 detectors are searched for gravitational wave signals by scientists of the LSC. Probably soon a memorandum of understanding (MOU) will be signed between the LSC and VIRGO collaboration. Then people will be able to analyze simultaneous data streams from the five most sensitive gravitational wave detectors built so far.

1.3. A brief description of the GE0 600 GW detector

The GEO 600 gravitational wave detector is a very complicated device, consisting of about 300 partly coupled control loops. The core of the detector comprises of about 30 optics, suspended by multi-stage vibration isolations and enclosed by an ultra high

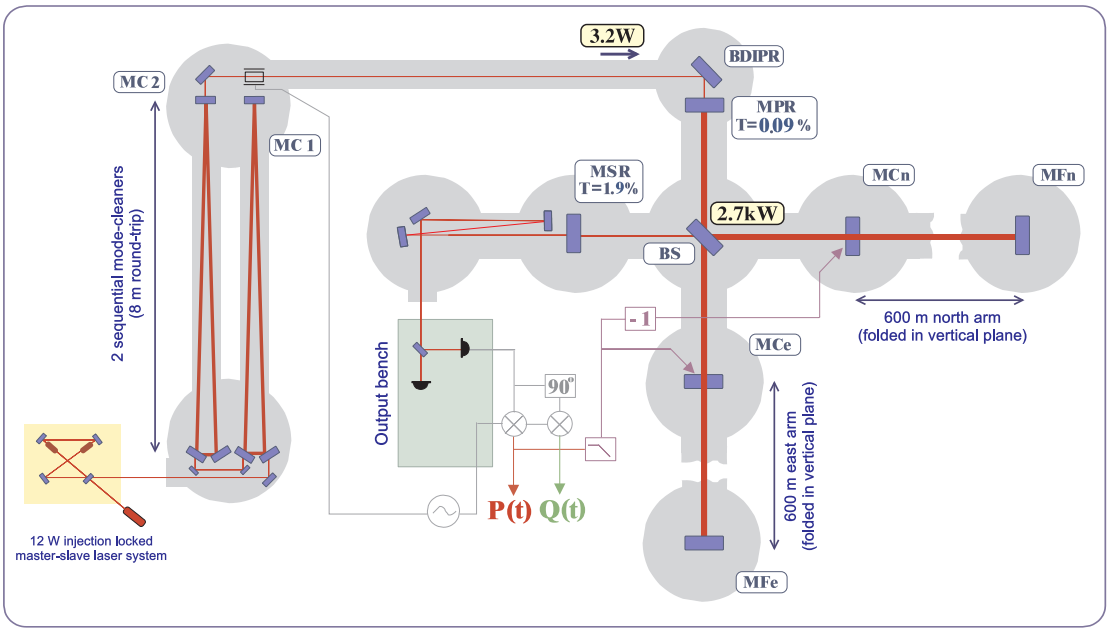


Figure 1.2.: The light of a 12 Watt master-slave laser system is injected into two sequential mode-cleaners, MC1 and MC2 of 8 and 8.1 meters round-trip length, and a finesse of 2700 and 1700, respectively. The stabilized and filtered light enters the main Michelson interferometer through the power-recycling mirror (MPR). The main interferometer consists of five optical components: the beam splitter (BS), the two end mirrors (MCe and MCn) and the two folding mirrors (MFe and MFn). In contrast to most of the other large scale interferometric gravitational-wave detectors GEO600 does not use Fabry-Perot resonators in the arms, but instead the simplest case of an optical delay line including one folding mirror per arm. The light containing potential gravitational-wave information leaves the Michelson interferometer at the antisymmetric port and the signal gets enhanced by usage of a second recycling mirror, the signal recycling mirror (MSR). The light passing the signal-recycling mirror is detected at the output bench, and the gravitational-wave information is derived from a RF-heterodyne method.

vacuum system of a volume of about 400 cubic meters. A simplified optical layout of the detector is shown in Figure 1.2. The alignment systems of GEO 600 automatically controls 38 angular degrees of freedom. At a frequency of 450 Hz a displacement sensitivity of $1.5 \cdot 10^{-19}\mathrm{m}/\sqrt{\mathrm{Hz}}$ and a strain sensitivity of better than $3 \cdot 10^{-22}/\sqrt{\mathrm{Hz}}$ are achieved. Figure 1.3 shows the improvement of the sensitivity from GEO 600 over the last few years. About 100 gigabytes of science data are collected and processed per day.

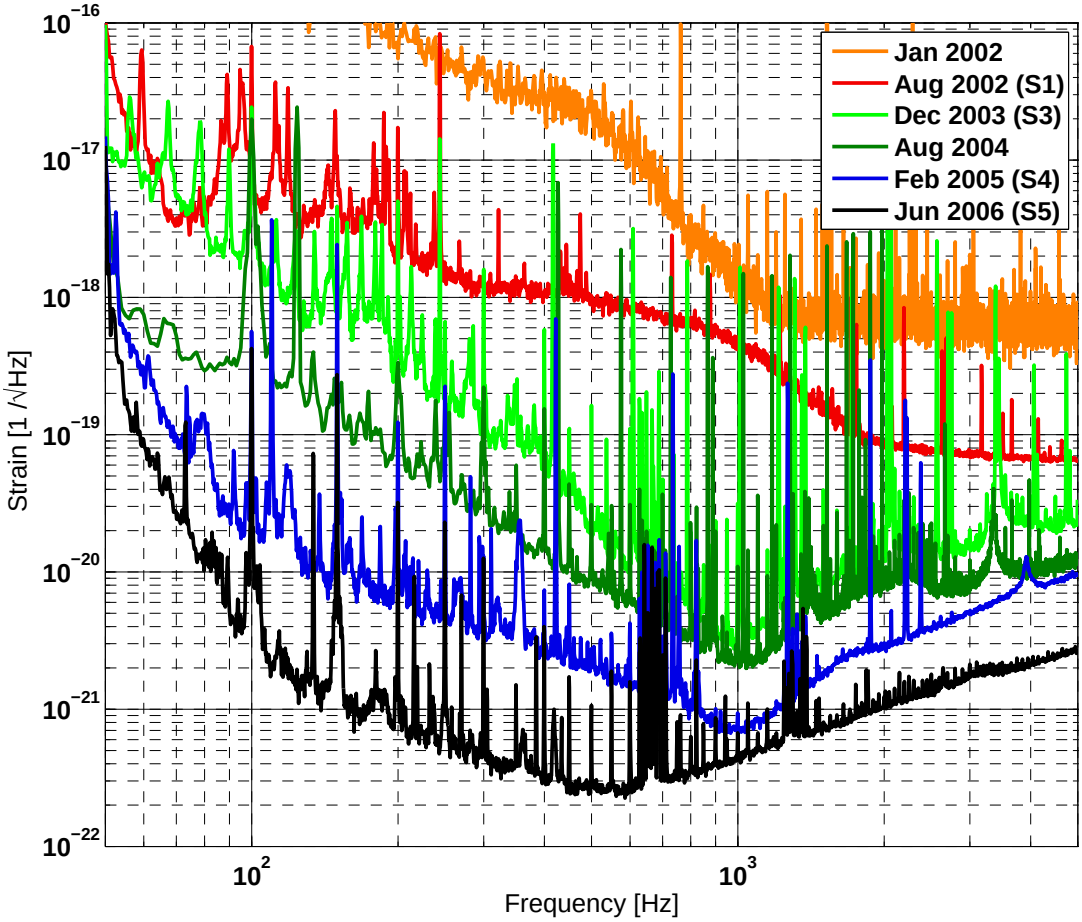


Figure 1.3.: Sensitivity improvement of GEO 600 over the last few years.

The preceding gives an impression of the complexity of GEO 600. A full description of GEO 600 would not fit into the frame of a single dissertation. Therefore the curious reader is asked to refer to the latest papers giving the status of GEO 600 [Willke04], [Grote05], [Lück06], [Hild06c] for a quick overview. More detailed descriptions of the most important subsystems of the GEO 600 detector can be found in the following five

dissertations: [Freise03], [Grote03], [Hewitson04a], [Gossler04] and [Smith06].

Chapter 2.

Demonstration and comparison of tuned and detuned signal recycling in a large-scale gravitational wave detector

2.1. Introduction

Placing a mirror (MSR in Figure 2.4) in the dark port of an interferometric gravitational wave detector can significantly increase its sensitivity in a certain frequency band. This technique, called signal recycling (SR), was proposed by Brian Meers [Meers] and first demonstrated in a table top experiment by Strain and Meers [Strain91]. The combination of power and signal recycling, called dual recycling, was realized 1998 in a fully suspended interferometer at the Garching prototype [Heinzel98]. The GEO 600 interferometer is the first and so far only large scale gravitational wave detector using signal recycling. However, nearly all projects plan to use signal recycling in their next generation of instruments, like for example Advanced LIGO [Giaime] or the Japanese LCGT project [Kuroda06].

The use of signal recycling allows a frequency dependent shaping of the detector response function, which is defined as the transfer function from to differential arm length changes to the output of the detector. The bandwidth of the signal recycling resonance is determined by the reflectivity of the signal recycling mirror. A high reflectivity gives a large increase of the response function in a narrow band (narrow-band operation), while a moderate reflectivity yields a medium improvement of the response function over a broader frequency range (broadband operation). Figure 2.1 shows exemplary the

shot noise limit of the GEO 600 detector for three different transmittances of the signal recycling mirror.

The frequency of maximum response, also called tuning frequency, is determined by the length of the signal recycling cavity and can be chosen by a change of the microscopic position of the signal recycling mirror. In this thesis we will refer to *tuned* signal recycling as the case when one of the resonances of the signal recycling cavity is centered at the frequency of the carrier, in contrast to *detuned* signal recycling where the signal recycling resonance is shifted to a frequency different from the carrier. The difference between the frequency of the signal recycling resonance and the carrier frequency is referred to as the detuning frequency. In Figure 2.2 the shot noise limit of GEO 600 is shown for various tuning frequencies.

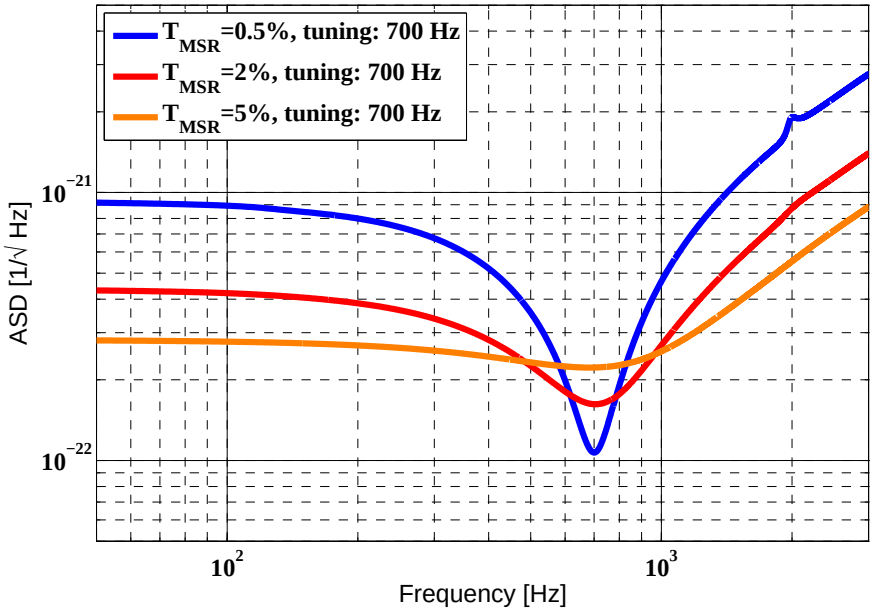


Figure 2.1.: Shot noise limit of GEO 600 for different transmittances of the signal recycling mirror.

Until recently, the GEO 600 detector has been operated with detuned signal recycling in order to shift the peak sensitivity to frequencies between 350 and 1000 Hz [Hild06c]. This was done to optimize the science contribution of the GEO 600 detector to the LSC (LIGO Scientific Collaboration) detector network. However, the use of detuned signal recycling brings some disadvantages. In contrast to a Michelson interferometer having

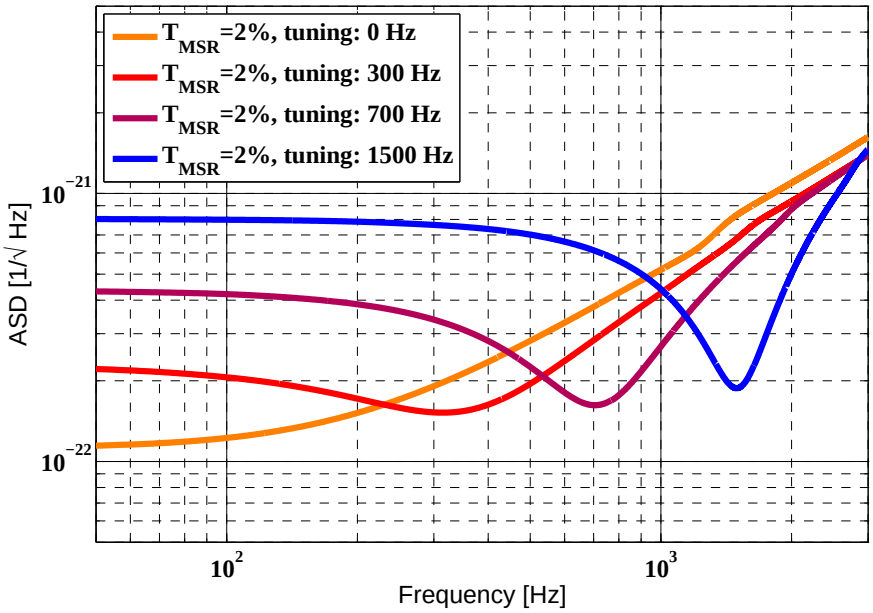


Figure 2.2.: Shot noise limit of GEO 600 for different tunings of the signal recycling cavity.

only power recycling, the signal recycling cavity changes amplitude and phase of almost any light component present at the dark port, for instance the gravitational wave signal sidebands or the radio frequency (RF) sidebands used for detector control.

Figure 2.3 gives a qualitative overview of the light fields in the signal recycling cavity for tuned and detuned operation. In the tuned case, the differential arm length information can be derived completely from demodulation of the photo current detected at the dark port in the in-phase quadrature (in the following referred to as P quadrature). In the detuned signal recycling case this information is in a frequency-dependent way spread over both quadratures, P and Q (out-of-phase quadrature) and both have to be analyzed to obtain the best signal-to-shotnoise ratio. A detailed description of this effect and its consequences for the calibration of the instrument can be found in [Hewitson04] and [Hewitson05].

Furthermore, the asymmetry of the RF control sidebands at the dark port causes a strong amplitude modulation of the light on the main photo detector, containing no gravitational wave signal. However, this strong RF amplitude modulation was found to

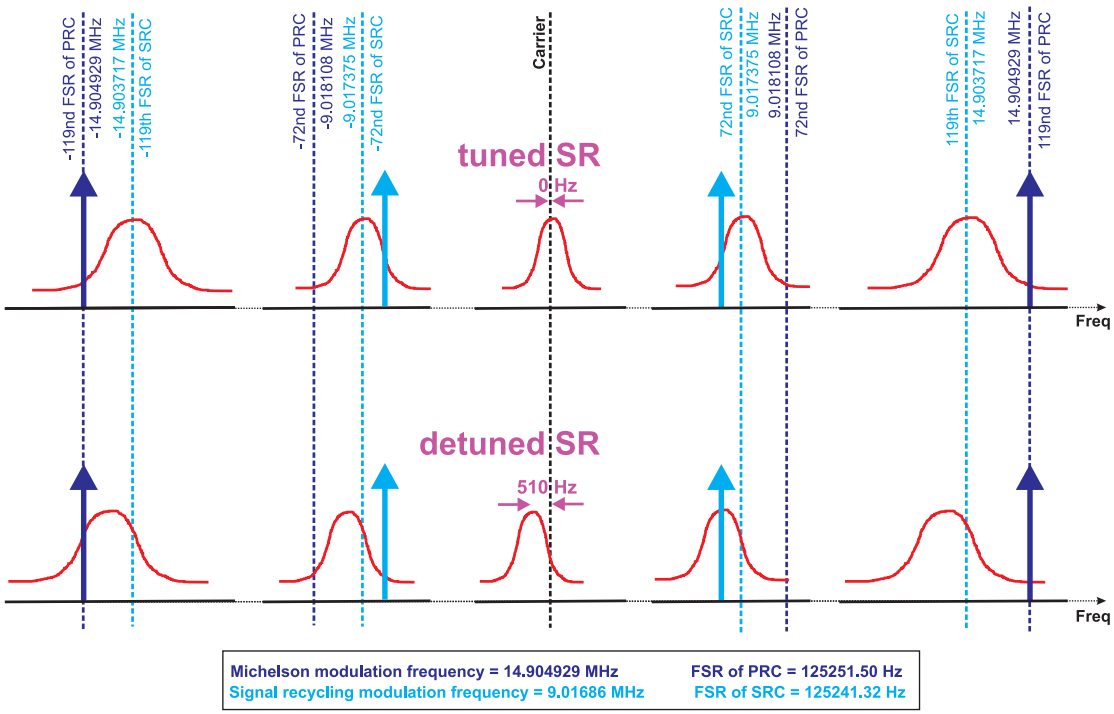


Figure 2.3.: Overview of the resonance conditions for the carrier (black dashed line), Michelson control sidebands (dark blue arrows) and signal recycling control sidebands (light blue arrows) inside the signal recycling cavity. The resonance conditions are shown for tuned and detuned (510 Hz) signal recycling in the upper and lower subplot, respectively. The comb of equidistant resonances of the signal recycling cavity (SRC) is indicated by the red Airy peaks. Due to the Schnupp asymmetry the shape of the resonance gets wider for frequencies far off the carrier. The gravitational wave signal is located around the carrier. The frequency of the Michelson modulation was chosen to be close to resonant in the 119th free spectral range (FSR) of the power recycling cavity (PRC). In detuned signal recycling the two Michelson sidebands see different resonance conditions inside the SRC: The lower sideband is nearly resonant, while the upper one is nearly off resonance. When going from detuned to tuned signal recycling using the method described in this thesis the frequency of all sidebands stays the same, but the comb of signal recycling resonances (red curves) is shifted by 510 Hz towards higher frequencies. In the tuned case the sidebands used for the Michelson control are balanced as well as the ones used for signal recycling control.

be a potential source of saturation of the main photo detector [Grote].[1]

It was also observed in the detuned case that various noise sources, for instance oscillator phase noise and laser power noise, couple to the gravitational wave channel in a more complex way than one would expect from simple models. A suspected cause for this are the imbalances of the sidebands in the interferometer and at the dark port.

An illustrative example of the complexity of the noise couplings in an interferometer with detuned signal recycling can be found found on the CD-ROM (Multimedia Appendix) at the end of this thesis. The time lapse video shows the evolution of the sensitivity of GEO 600 during an experiment in which a single parameter of the detector, namely the radius of curvature of MFE, was changed by about 1 %. The black trace indicates a reference sensitivity, while the blue trace represents the actual sensitivity.

The facts given above made it highly desirable for us to directly compare tuned and detuned signal recycling operation in GEO 600. In section 2.2 we describe the GEO control scheme implemented for detuned signal recycling, while in section 2.3 we introduce a new technique, which allows the operation of tuned signal recycling. In Sections 2.4-2.6 important aspects of tuned and detuned signal recycling (510 Hz) are compared: Section 2.4 shows measurements of the optical gain, in Section 2.5 a comparison of the size of the amplitude modulation at the main photodiode is given. In section 2.6 exemplary measurements of some noise coupling transfer functions to the gravitational wave channel are shown. Finally the sensitivity is compared for detuned and tuned signal recycling in Section 2.7 as well as for positive and negative detunings in Section 2.8.

2.2. Control scheme for signal recycling in GEO 600

The signal used to control the microscopic position of the signal recycling mirror (MSR) is derived from a radio frequency modulation/demodulation technique. Figure 2.4 shows a simplified diagram of the control of two degrees of freedom, the Michelson differential arm length and the length of the signal recycling cavity in blue and green respectively.

[1] At the LIGO detectors an additional feedback loop (called *I-servo*) is used which reduces the part of the RF signal that is not suppressed by the interferometer control loop. This is done by adding an appropriate RF-signal, essentially a sine wave at the modulation frequency, with correct amplitude and phase, directly to the PD-resonant circuit, thereby cancelling the RF-signal (at this frequency) in one quadrature [Sigg04].

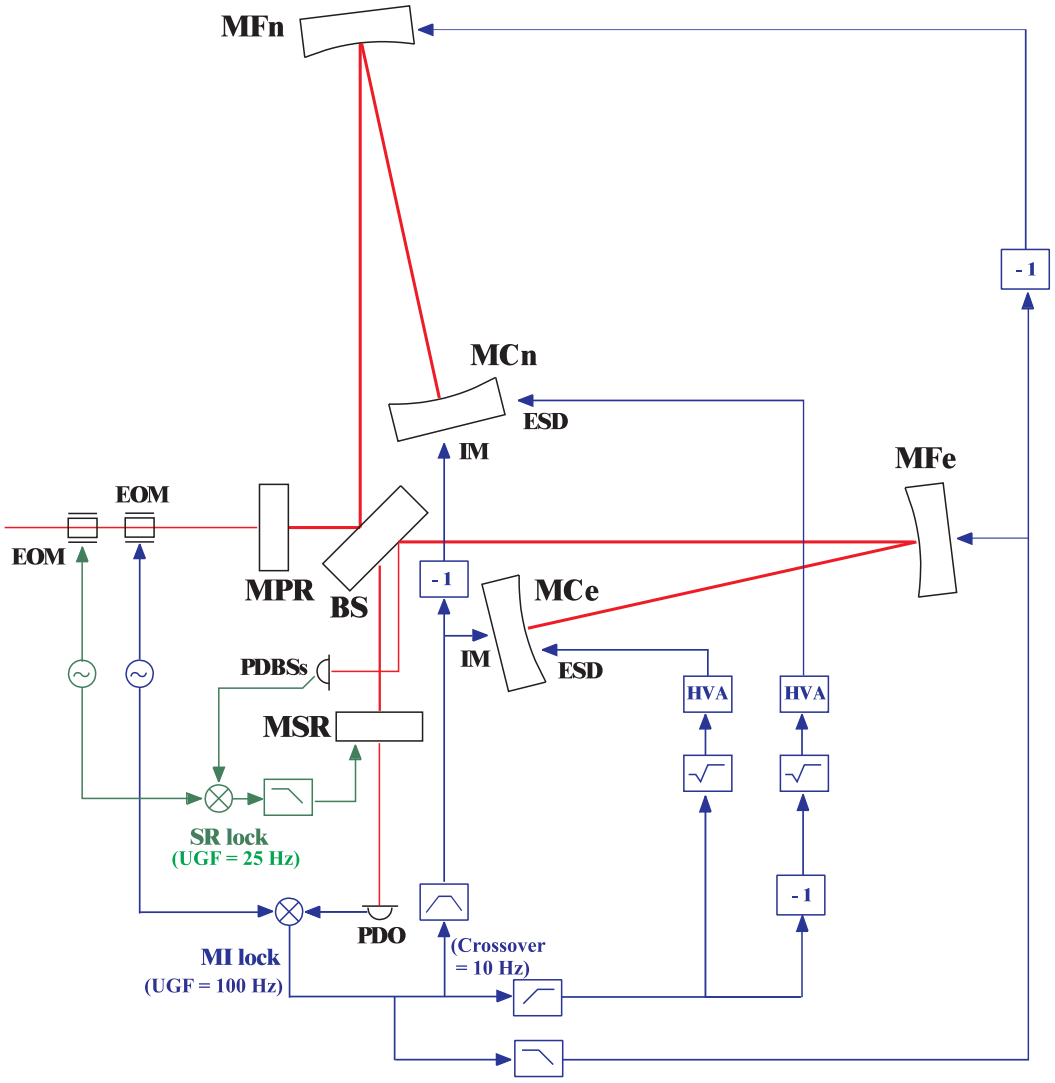


Figure 2.4.: Simplified control scheme for two longitudinal degrees of freedom of GEO 600. Shown are the Michelson differential loop and the signal recycling loop in blue and green, respectively. The control signals used for lock acquisition are not shown in the diagram.

Two sets of phase modulation sidebands are created in front of the power recycling mirror (MPR in Figure 2.4). Due to a asymmetry in interferometer arms of several cm (Schnupp asymmetry), a fraction of these sidebands leaves the interferometer at the dark port [Schnupp]. The error signal for controlling the differential arm length of the Michelson interferometer is derived from the photodiode PDO placed at the dark port of the interferometer. For the sensing of the signal recycling loop a pick-off beam from one of the interferometer arms is used (photodiode PDBSs, sensing the beam reflected at the beam splitter AR coated side).

For various reasons (described in detail in [Grote03]) it was so far not possible to realize lock acquisition for the tuned configuration of GEO 600. Therefore a procedure was developed to acquire lock at a high detuning (a few kHz) and then gradually tune the signal recycling cavity to lower frequencies. This so-called tuning process is done by changing the radio frequency of the signal recycling modulation and other relevant parameters [Grote05].

Figure 2.5 shows the error signal of the signal recycling loop for different modulation frequencies versus the position of the signal recycling mirror. These simulations were done using the FINESSE software [Freise04]. In the case of detuned signal recycling, the error signal structure shows three zero crossings. The two outer ones are referred to as the lower and upper sidebands corresponding to a negative or positive detuning, while the zero crossing in the center corresponds to the tuned case. So far the zero crossing corresponding to the upper sideband was chosen to be the nominal operating point. If the modulation frequency is increased, the zero crossings from the lower and upper signal recycling sideband are shifted towards the zero crossing corresponding to tuned signal recycling, i.e. the whole error signal structure gets squeezed, but keeps roughly its shape. As the signal recycling mirror is locked to the position at which the upper sideband crosses zero, the position of the mirror is shifted corresponding to the change of the modulation frequency of the SR control sideband. This tuning technique works for tunings as low as about 300 Hz.

As indicated in the second subplot of Figure 2.5, for tunings below 300 Hz the error signal structure is not only squeezed, but also changes its shape significantly. The slopes of the error signals of the upper and lower sideband become asymmetric around their zero crossings. For the extreme condition of a modulation frequency corresponding to the tuned case the zero crossings of the sidebands vanish completely. For frequencies

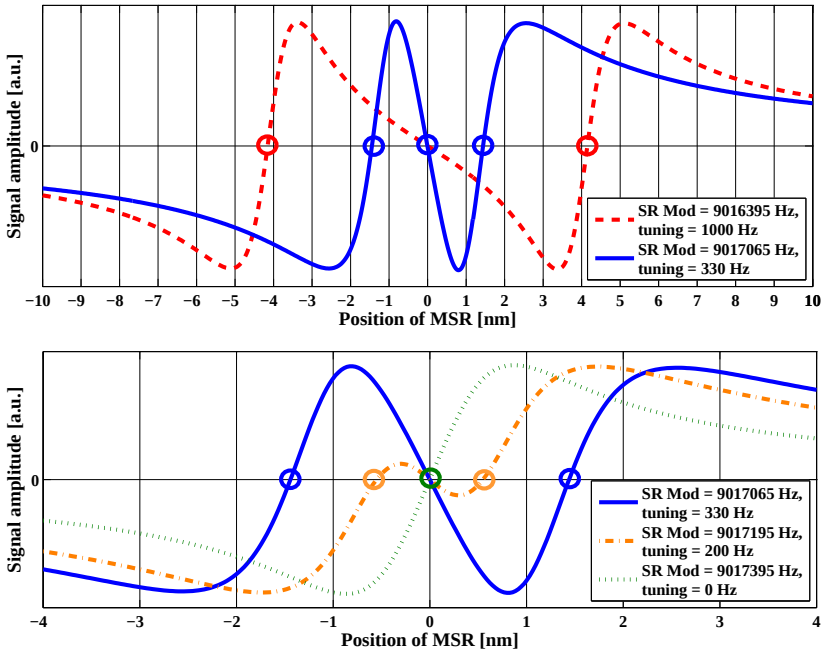


Figure 2.5.: Signal recycling error signals derived from an RF modulation demodulation technique versus microscopic position of the signal recycling mirror. The error signal is plotted for various modulation frequencies corresponding to detunings of the signal recycling cavity between 1000 Hz and 0 Hz. By increasing the modulation frequency the structure gets more and more narrow.

corresponding to tunings between 300 and 0 Hz in the presence of electronic or seismic noise, the error signals cannot be used for a stable control of the signal recycling mirror. Their range with a constant slope is too small to hold the mirror in a defined position, given that the control bandwidth of the signal recycling loop is limited. Therefore reaching the tuned case in small steps seems not a promising technique here.

2.3. A new method for locking tuned signal recycling

One possible way to reach the tuned case is to start from the lowest stable tuning, then make the microscopic position of the signal recycling mirror 'jump' over the region where the control signals are not valid, and 'catch' it again at the tuned state, where reasonable control signals can be derived. This can be done without changing the signal recycling modulation frequency by jumping from the zero crossing corresponding to the upper sideband to the zero crossing corresponding to the tuned case.

We realized this jumping by pushing the signal recycling mirror in the direction corresponding to a smaller detuning frequency and stopping it near the zero crossing of the tuned case. In order to not disturb the other control loops, for example the loop controlling the differential arm length of the Michelson interferometer, the signal recycling mirror has to be pushed over the region of non-valid error signals quickly.

In detail the procedure works as follows: We start from a modulation frequency corresponding to a tuning of about 350 Hz where the signal recycling mirror is locked at the position corresponding to the zero crossing of the error signal around the upper sideband. Then we switch off the signal recycling control loop to be able to push the signal recycling mirror away from its operating point. At the same time we start pushing the signal recycling mirror as hard as the coil-magnet actuators allow. After 4 milliseconds of pushing, the mirror has covered half the distance between the zero crossing of the upper sideband and the one from tuned case which is in this case a distance of 0.7 nm. Then the sign of the force applied to mirror is inverted to decelerate the mirror in 4 milliseconds from maximum speed to a velocity near zero when the mirror finally reaches the tuned position. After these 8 milliseconds the signal recycling feedback loop is closed again, but with opposite polarity, to account for the different sign of the slope around the zero crossing of tuned signal recycling.[2] The duration of this procedure is short

[2] Another possibility would be to change the modulation frequency while the mirror is being pushed,

enough, to not significantly disturb other control loops.

Using this technique it was also possible to jump twice the distance (about 3 nm) with the signal recycling mirror, and thereby go from the zero crossing of the upper sideband to that of the lower sideband, which turns out to be an equally stable operating point.

2.3.1. Simulation of the control parameters for tuned signal recycling

When jumping from detuned to tuned signal recycling a few control parameters need to be adjusted to account for the different operating point of the signal recycling mirror. The gain of the signal recycling loop needs to be adjusted for the different slope of the error signal. Furthermore in the case of tuned signal recycling the pole of the signal recycling cavity for carrier light is shifted towards lower frequencies, thus all signals generated from the dark port need to be adapted.

Figure 2.6 shows the results of simulations, done using the Finesse software [Freise04], for the longitudinal gains of the Michelson (upper set of subplots) and signal recycling (lower set of subplots) longitudinal error signals. In tuned signal recycling operation the Michelson error signal has at low frequencies about 9.5 dB and around the unity gain frequency (100 Hz) about 7.8 dB higher gain. Due to the shifted signal recycling cavity pole around the unity gain frequency about 30 degrees of phase are lost in tuned signal recycling operation compared to detuned operation. For tuned signal recycling operation this is compensated by switching an additional differentiator into the loop controlling the Michelson differential arm length and adjusting the overall gain of the loop.

The gain of the signal recycling error signal is about 2 dB higher in tuned case for all frequencies within the control bandwidth of 50 Hz. The phase is for all frequencies of interest 180 degrees different between tuned and detuned operation, i.e. the sign is swapped as indicated in Figure 2.5. A simple change of the overall loop gain of 2 dB is sufficient for stable operation of tuned signal recycling.

and then to keep the sign of the control loop.

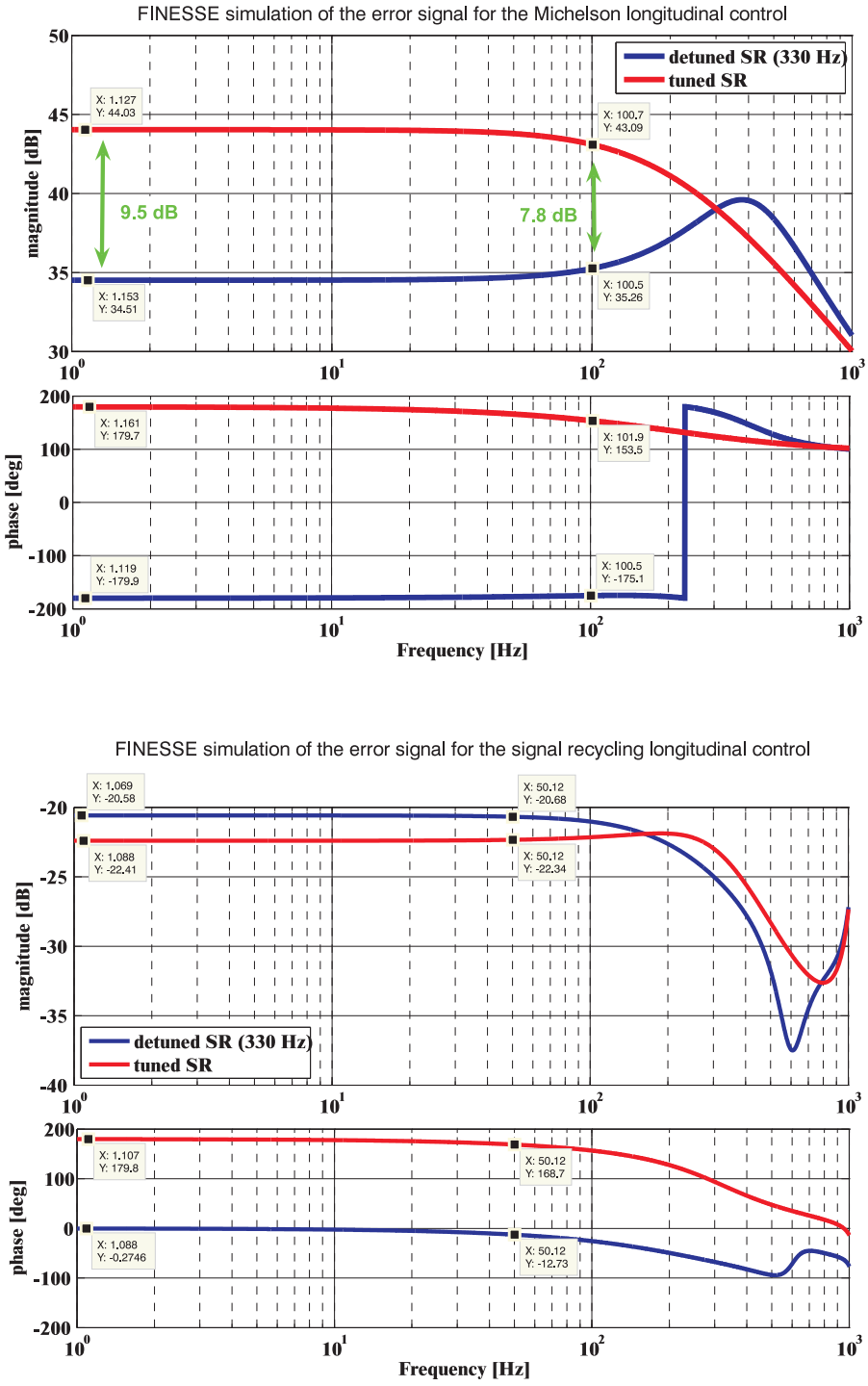


Figure 2.6.: Simulations of the error signal gain for the longitudinal control of the Michelson differential armlength (upper pair of plots) and the longitudinal control loop of the signal recycling mirror (lower pair of plots). The Michelson control loop has an unity gain frequency of about 100 Hz. The signal recycling control loop has unity gain frequencies of about 50 Hz in acquisition mode and 25 Hz in low-noise run mode.

2.3.2. Calibration of the actuator strength

In order to increase the accuracy of the jumping process it is desirable to calibrate the applied force to the signal recycling mirror in order to be able push the mirror as accurate as possible to its new operating point.

The idea of the calibration process is to use the known distance of the two positions of the signal recycling mirror for detuned and tuned operation. Starting from position P_{start} corresponding to detuned operation the mirror is pushed for Δt milliseconds with a force F_p towards the tuned operation point. Afterwards the mirror is decelerated by again Δt milliseconds, but with opposite force $-F_p$. At the end of this process the mirror reaches the position P_{end} and should, in absence of any disturbance, hang still again. We now vary Δt and F_p in order to stop the exactly at the position corresponding to tuned operating point. This position is in particular suitable as it corresponds to a zero crossing of the error signal.

Using the maximum force that can be driven by the coil driver electronics we get the following calibration of the actuator. The maximal acceleration a_{push} amounts to:

$$a_{\text{push}} = 1.53 \cdot 10^{-4} \frac{\text{m}}{\text{sec}^2} \tag{2.1}$$

2.4. Measurements of optical gain

The response function of the Michelson differential error signal to differential arm length fluctuations, the so called optical gain *G_opt*, is frequency dependent. In the case of detuned signal recycling the signal is also inherently spread between the two orthogonally demodulated signal quadratures, P and Q.

Figure 2.7 shows a simplified diagram of the loop controlling the differential arm length of the Michelson. The optical gain of the P quadrature can be described by the following expression

$$G_opt(P) = \frac{\text{EP-P} \cdot \text{P-dw}}{\text{FB} \cdot \text{FB-dw}} \cdot \frac{1}{\text{ESD}} \cdot \frac{1}{\text{HVA-dw}} \tag{2.2}$$

where EP-P (LSC_MID_EP-P_HP) is the error signal from the P quadrature recorded in the data acquisition system (DAQS) and FB (LSC_MID_FB-MCEMCN) is the record of

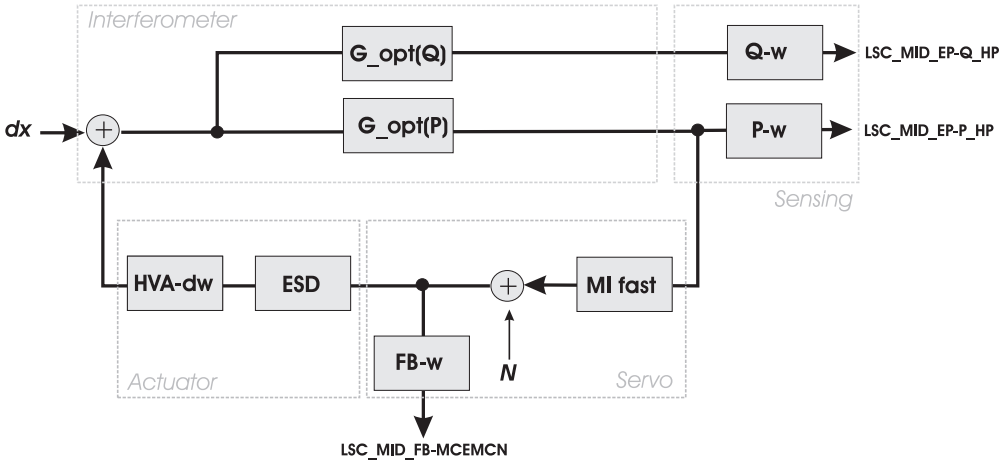


Figure 2.7.: A simplified diagram of the loop controlling the differential arm length of the Michelson. The diagram contains only components important for calibration and measuring the optical gains *G_opt* from the two orthogonal signal quadratures P and Q. Abbreviations used: MI fast = electronics of the servo, FB-w = whitening filter of feedback signal, ESD = high voltage amplifier (HVA) and electro static drives, HVA-dw = HVA dewhitening filter, P-w = whitening filter for P signal, P-w = whitening filter for Q signal, N = noise injected (for optical gain and loop transfer function measurements), dx = mirror displacement.

the feedback. ESD and HVA-dw represent the responses of the actuators, FB-dw is the response of the feedback whitening filter and P-dw is the inverse of the P error signal whitening filter.

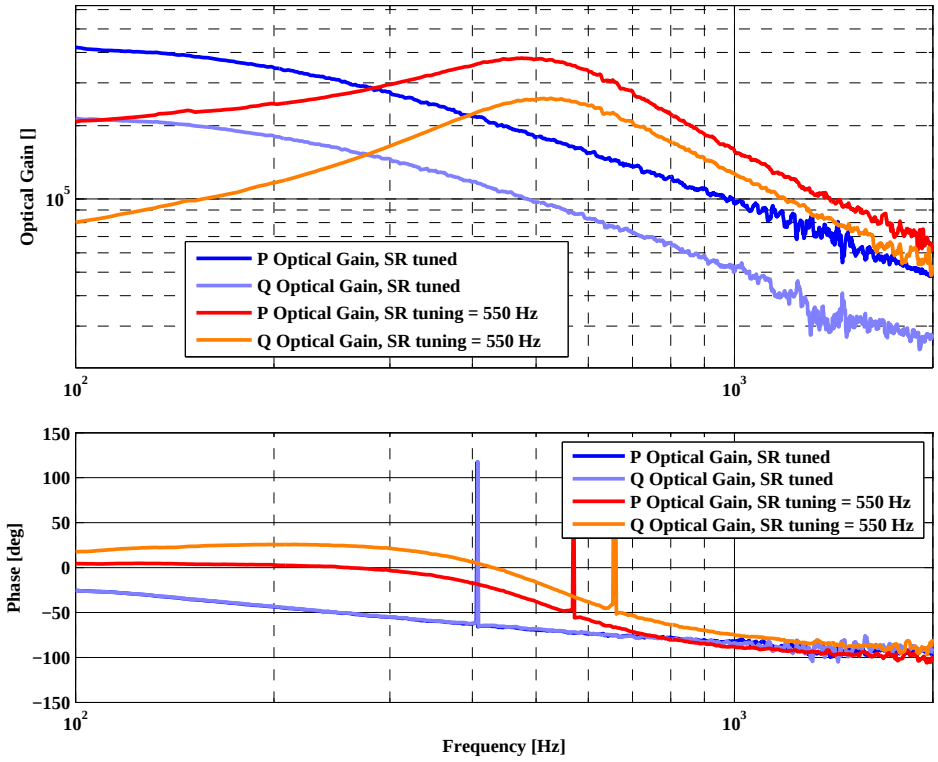


Figure 2.8.: Measurement of the optical gains for tuned and detuned signal recycling for the two orthogonal quadratures P and Q. For the tuned case the demodulation phase was not optimized. Therefore the the differential armlength signal is still clearly present in the optical gain of the Q-signal.

Figure 2.8 shows measurements of the optical gains for detuned and tuned signal recycling for the two quadratures. In the detuned case the optical gain shows a maximum at the frequency of the signal recycling detuning. The width of the maximum is given by the bandwidth of the signal recycling cavity, which is about 700 Hz for the currently installed signal recycling mirror with a transmission of about 2%. In the tuned case the maximum of the optical gain is centered around 0 Hz and the bandwidth of the response function is decreased to 350 Hz.

2.5. Reduction of the RF amplitude modulation

In detuned signal recycling the RF sidebands from the Schnupp modulation which originally represented a pure phase modulation get partly converted to amplitude modulation inside the signal recycling cavity. The light hitting the main photo diode at the dark port shows strong amplitude modulation at the frequency of the Michelson control sidebands, which are in case of GEO 600 at a frequency of about 15 MHz.

Some of our observations indicated that the strong amplitude modulation (in detuned SR) together with a high averaged photo current might cause saturation effects and nonlinearities in the photodiode. A detailed description of these problems can be found in [Grote]. However, it is important to mention that most of the spurious effects depend on the size of the RF amplitude modulation. Therefore it would be desirable to reduce the size of the amplitude modulation.

In the case of tuned signal recycling we have been able to reduce the rms of the signal in the Q quadrature, which is a good measure of the size of the amplitude modulation, by a factor of 12 compared to detuned operation.

2.6. Comparison of noise transfer functions for tuned and detuned signal recycling

Measurements of the transfer functions from various technical noise sources to the error signal of the differential arm length servo (and hence the gravitational wave channel) during detuned operation indicate complex couplings [Smith06]. Many of the measured transfer functions contain resonances and notch structures that could not be explained completely even by a complex frequency domain model of the interferometer [Malec06].

Asymmetry of the radio frequency sidebands in the detuned case was suspected to be the origin for some of the complexity observed in the noise couplings. To investigate this we performed measurements for four important technical noise sources for tuned and detuned signal recycling in order to check whether the couplings get simpler and perhaps less significant for tuned signal recycling.

- *Oscillator phase noise (OPN)*: The two sets of upper plots in Figure 2.9 show the measurements of the coupling from oscillator phase noise (of the modulation

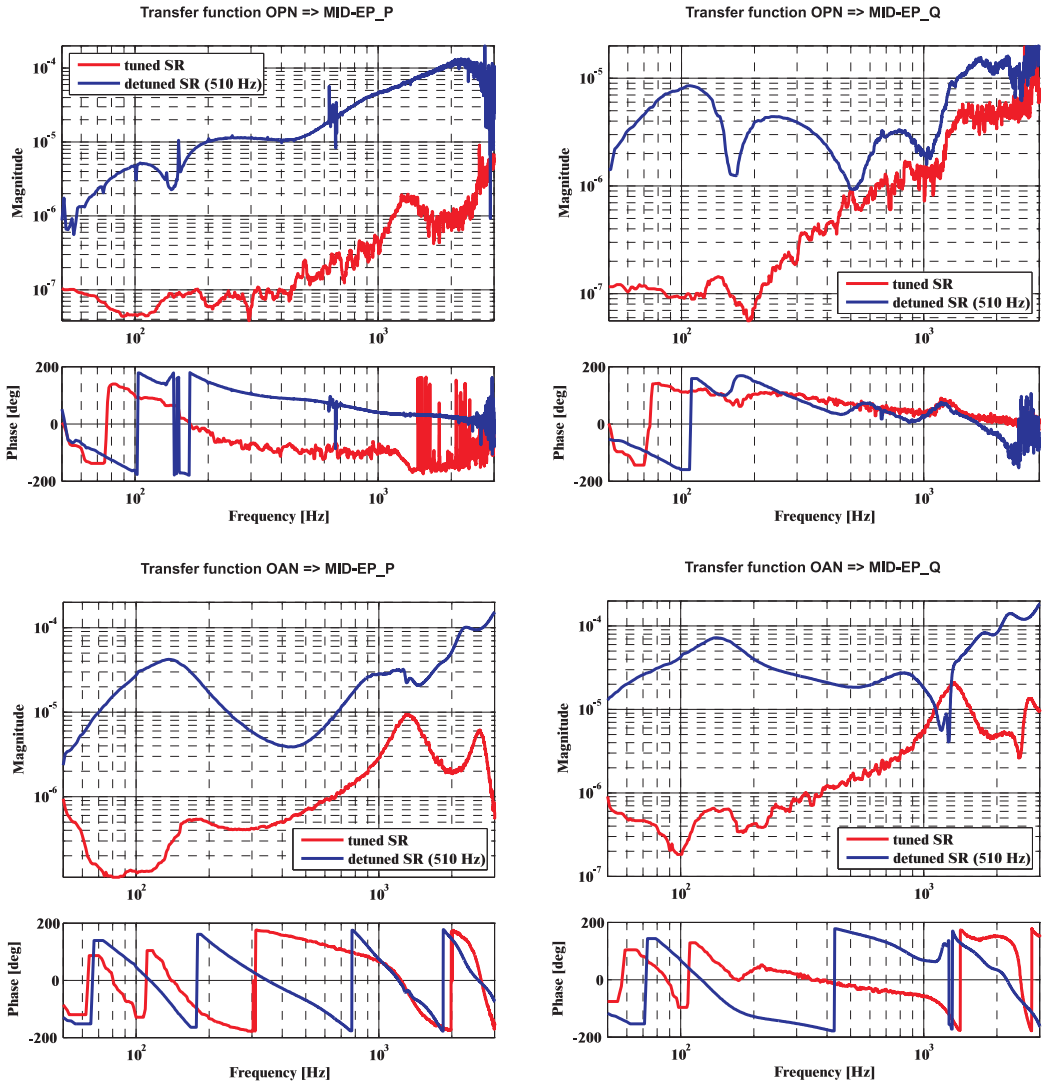


Figure 2.9.: Measurements of the transfer functions from oscillator phase noise (OPN) and oscillator amplitude noise (OAN) to the error signal of the differential arm length servo. The red trace represents tuned signal recycling, while blue indicates a detuning of 510 Hz. The magnitudes of transfer functions are normalized by magnitudes of the optical gains from Figure 2.8, which allows for a direct comparison of the transfer functions regarding the strain sensitivity of the detector.

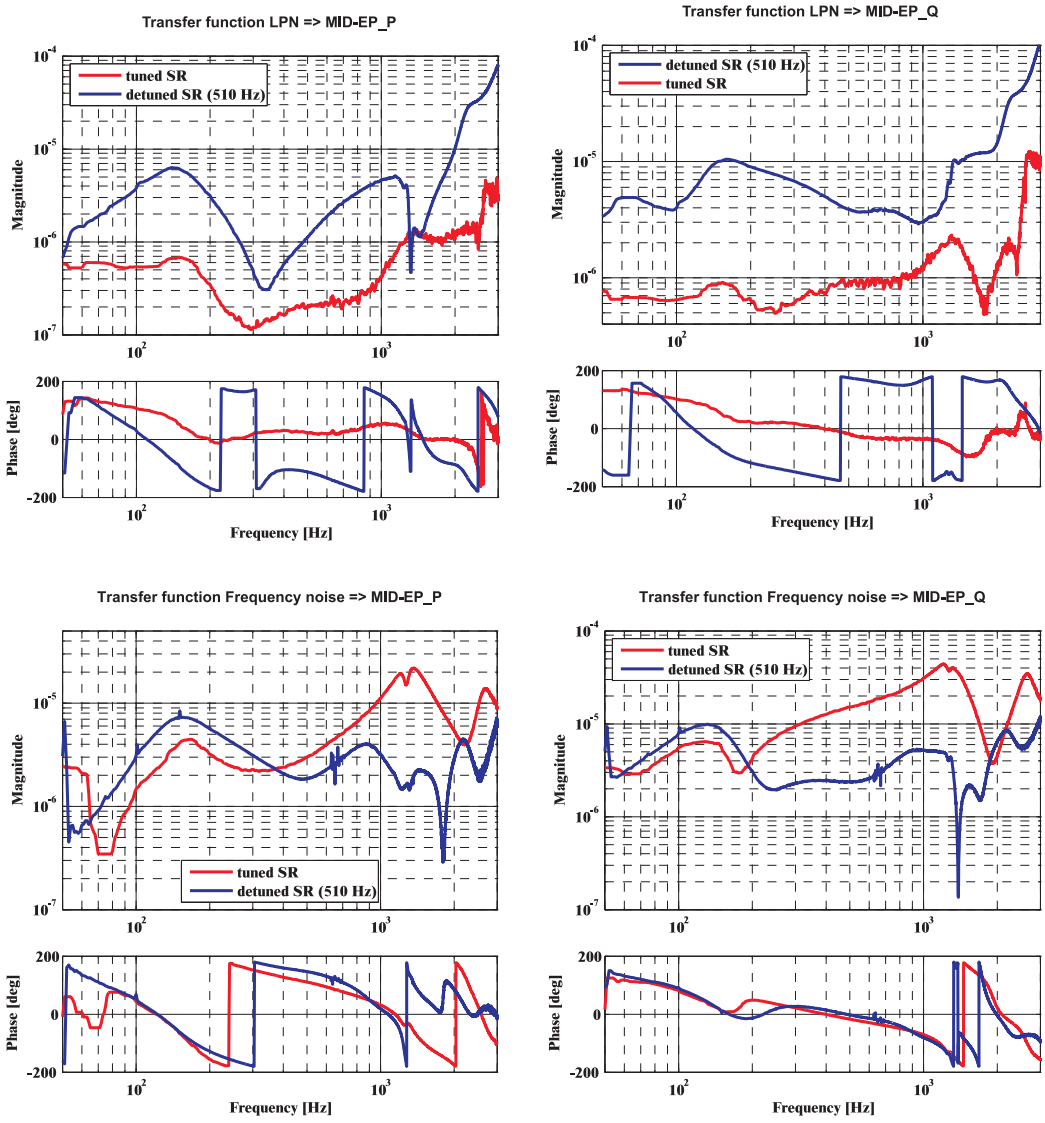


Figure 2.10.: Measurements of the transfer functions from laser frequency noise and laser power noise to the error signal of the differential arm length servo. The red trace represents tuned signal recycling, while blue indicates a detuning of 510 Hz. The magnitudes of transfer functions are normalized by magnitudes of the optical gains from Figure 2.8, which allows for a direct comparison of the transfer functions regarding the strain sensitivity of the detector.

used for creating control signals for the differential arm length of the Michelson interferometer and the local oscillator signal) to differential arm length. In the P quadrature the coupling in the tuned case is dramatically reduced over a wide frequency range by about two orders of magnitude. This can be explained by the reduction of the rms of the signal in the Q quadrature, (see section 2.5). The coupling of OPN to the Q quadrature is for frequencies below 500 Hz significantly decreased by up to a factor of 100 at 100 Hz. A detailed description of potential coupling mechanisms can be found in [Smith06].

- *Oscillator amplitude noise (OAN)*: The two sets of lower plots in Figure 2.9 show the measurements of the coupling from oscillator amplitude noise (of the modulatior used for creating control signals for the differential arm length of the Michelson interferometer) to differential arm length. The amplitude noise in the EOM path is eventually imparted on the carrier and control sidebands and couples to the detector output via asymmetries in the interferometer, as explained in [Malec06]. Again the coupling into both quadratures is strongly suppressed in tuned signal recycling due to more balanced sideband conditions at the dark-port.

- *Laser power noise (LPN)*: The two sets of upper plots in Figure 2.10 show the measurements of the coupling from laser power noise to differential arm length. The coupling to both quadratures is decreased on average by about a factor of 5 to 10 over the frequencies of interested. In addition at least in the P quadrature the structure of the coupling got slightly simpler. The strongly pronounced notch structure around 1.5 kHz vanishes for tuned signal recycling.

- *Laser frequency noise*: The two sets of lower plots in Figure 2.10 show the measurements of the coupling from laser frequency noise to differential arm length. At low frequencies we observed a moderate reduction of the coupling in the tuned case, while for higher frequencies the coupling is significantly increased. On the other hand the very distinct notch structures between 1 and 2 kHz seem to vanish completely.

2.7. Calibrated detector sensitivity for tuned signal recycling

Even though the performed investigations on comparing tuned and detuned signal recycling focus on the different propagation of optical sidebands through the interferometer,

the curious reader should get a chance to look at the GW sensitivity for tuned and detuned signal recycling, too.

The calibration of the detuned detector is done continuously using a time-domain method. A detailed description can be found in [Hewitson04] and [Hewitson05]. In order to calibrate the new case of tuned signal recycling operation we used a frequency domain method. With the aid of the closed loop transfer function ($CLTF$) the differential displacement dx (see Figure 2.7) can be expressed by:

$$dx = \frac{\text{EP-P} \cdot \text{P-dw}}{G_opt(P) \cdot CLTF}. \tag{2.3}$$

To convert the displacement dx to strain sensitivity we have to divide dx by the optical arm length of GEO 600:

$$HP = \frac{dx}{1200\,\text{m}} \tag{2.4}$$

The calibration of the orthogonal quadrature Q is done analogous.

Figure 2.11 shows the calibrated sensitivities of the two output quadrature signals. By going to tuned signal recycling no improvement in the calibrated P signal is observed. As shot noise is the main contribution to the noise level of GEO 600 for frequencies above 500 Hz, the sensitivity got worse in that frequency range when going to tuned signal recycling. At low frequencies we observed no increase in sensitivity because in that frequency range the shot noise is masked by the noise contributions of several technical noise sources.

For the sensitivity derived from the Q signal the situation is different: Between 90 and 200 Hz an improvement is achieved for tuned signal recycling operation. On the other hand, above 300 Hz the sensitivity is worse by about a factor of 2.

This is only a preliminary result as not much time during the science run could be spend to optimize the detector for tuned signal recycling operation yet.

2.8. Comparison of detuned signal recycling for upper and lower operating point

From the sensitivity point of view it might also be interesting to compare the sensitivities of the detector for locking the signal recycling mirror either at the position corresponding

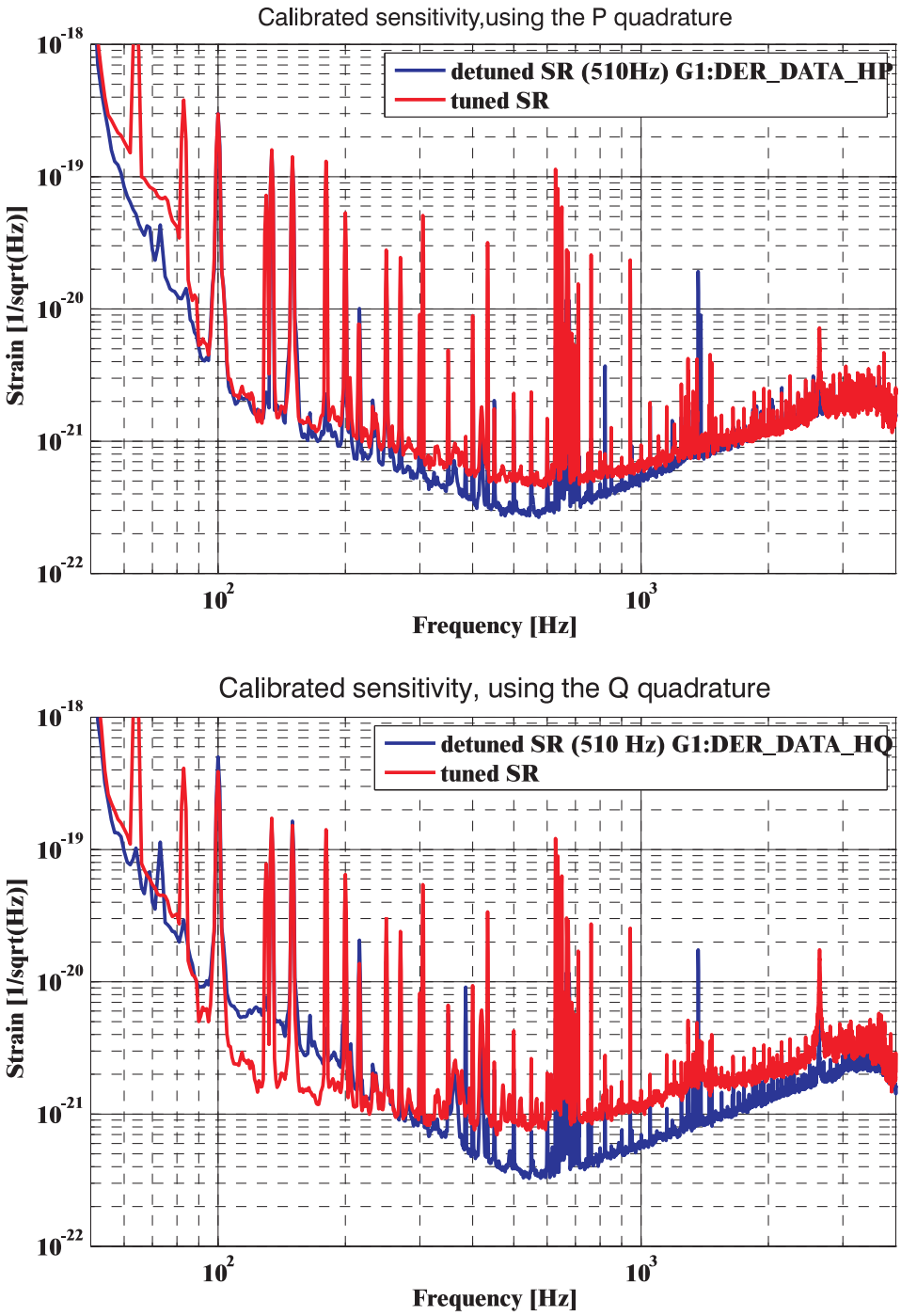


Figure 2.11.: Comparison of the sensitivity of GEO 600 for tuned and detuned signal recycling operation. The upper plot shows the calibration derived from using the in-phase signal (P), while the lower plot is derived from using the out-of-phase signal (Q).

to the lower sideband crossing or the position corresponding to the upper sideband crossing. In principle there should be no difference between the two cases: By going from the upper sideband lock to the lower sideband lock all the imbalances of the light fields in the dark-port are preserved. In the lower subplot of Figure 2.3 the resonance condition is plotted for a positive signal recycling detuning of 510 Hz. How does this picture change when going to the opposite detuning of -510 Hz. As the signal recycling mirror is locked in that case to the zero crossing of the lower signal recycling sideband, the comb of signal recycling resonances is shifted by 1020 Hz to the right, thus the -72nd signal recycling resonance is centered around lower signal recycling sideband. Consequently the complete picture in the lower plot of Figure 2.3 is just mirrored at the carrier frequency.

Light field	lower MI SB	lower SR SB	carrier	upper SR	sideband
SR tuning = +510 Hz	-702 Hz	+1020 Hz	+510 Hz	0 Hz	+1722 Hz
SR tuning = -510 Hz	-1722 Hz	0 Hz	-510 Hz	-1020 Hz	+702 Hz

Table 2.1.: A summary of the resonance condition of the important light fields at the dark-port for a positive and negative detuning of the signal recycling cavity. The given values indicate the distance of the light field from the center of the nearest resonance peak.

Table 2.1 gives an overview of the resonance condition of the light fields of interest inside the dual recycled interferometer. Provided that any kind of technical noise is impressed equally onto the lower and the upper sideband of each, MI and SR modulation, the detected signals at the dark-port, which orignate from the beat of the two sidebands with the carrier field, should be equivalent for positive and negative detuning. That is also the reason why all the gains of the control loops can stay the same for both cases and no adjustments of any control system is required.

Figure 2.12 shows a measurement of the sensitivity for a positive (blue trace) and a negative (red trace) detuning of the signal recycling cavity of 510 Hz. Apart from a few line features an astoundingly agreement is found.

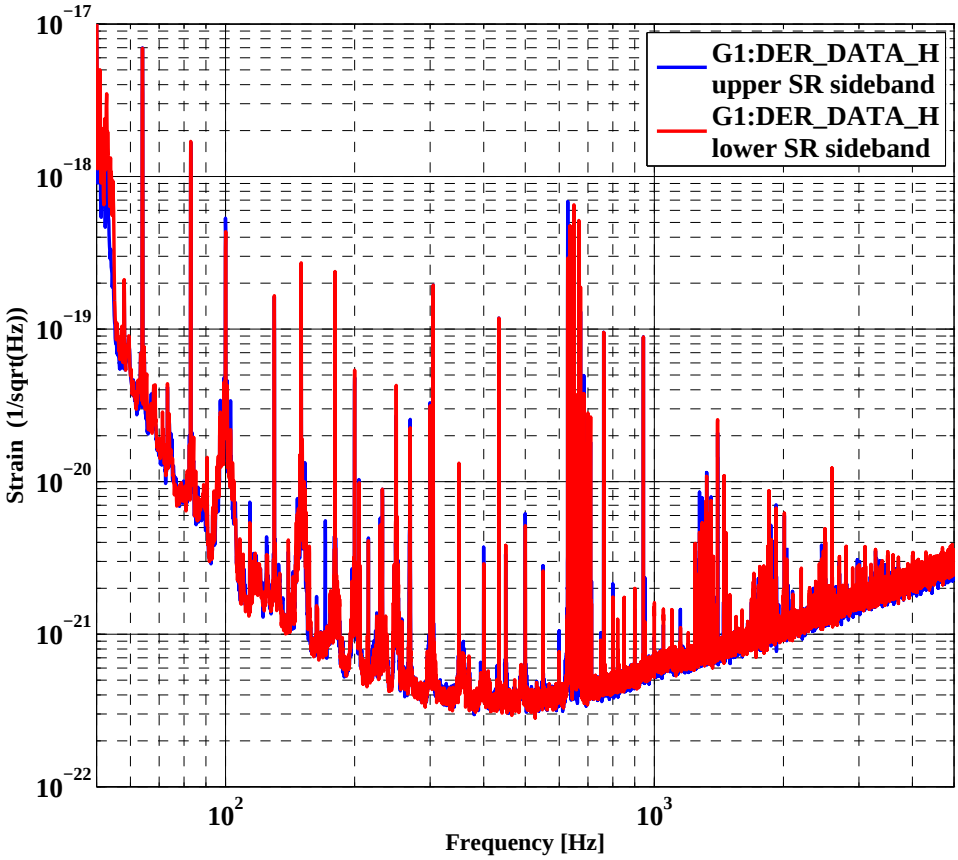


Figure 2.12.: Comparison of the sensitivity of GEO 600 for two different operating points of detuned signal recycling. The blue trace corresponds to a nominal tuning frequency of +510 Hz, while the red trace represents a negative detuning frequency of −510 Hz.

2.9. Summary

We developed and demonstrated a new technique for rapidly changing the signal recycled GEO 600 gravitational wave detector from the detuned to the tuned operating state. With this method, the signal recycling mirror is shifted accurately between two positions within milliseconds. This allowed for jumping with the signal recycling mirror from a position corresponding to a positive detuning of 510 Hz to a position corresponding to the tuned case and also from a positive detuning of 510 Hz to the corresponding negative detuning of −510 Hz. The method is generally applicable and only limited by the strength of the actuators used for longitudinal control of the signal recycling mirror.

A comparison of tuned and detuned (510 Hz) signal recycling was given. The measurements of the optical gain for the tuned case show that the maximum in the signal response is shifted to DC. In the tuned case the rms of the signal in the Q quadrature is reduced by a factor of 12. The measurement of various noise couplings to the gravitational wave channels were performed for tuned and detuned signal recycling. The transfer functions are shown to be significantly different. For most of the noise sources (except laser frequency noise) we found the magnitude of couplings being strongly decreased in tuned signal recycling. Furthermore in the tuned case the complexity of transfer functions seems to be, at least for some couplings, slightly reduced.

In the future we will try to improve the simulations of the noise couplings in order to reproduce the measured results. A deeper understanding of the complex noise couplings, originating from imbalanced sidebands, is not only important for commissioning of the current interferometers, but also essential for the design of next generation gravitational wave detectors.

Chapter 3.

Scattered light problems in interferometric gravitational wave detectors

3.1. Introduction

Stray light problems have been encountered during the commissioning of all currently operating large scale gravitational wave detectors. The underlying principle of all these interferometers is to make an extremely sensitive phase measurement. Therefore even tiny stray light contributions with a different or varying phase will harm the measurement. The currently achieved sensitivity of GEO 600 can already be spoiled by stray light of the order 10^{-20} W (!).[1]

A well-known manifestation of stray light noise is the so called *scattering shoulder* which has been observed in many detectors independent of their topology, as shown in Figure 3.1. A shoulder was observed in the VIRGO and LIGO instruments, which are power recycled Michelson interferometers with arm cavities, as well as in the folded arm GEO 600 configuration with and without signal recycling.

Stray light problems have already been observed in former interferometers like the 30 m-prototype in Garching. However, there the scattering was caused by the core optics of

[1] GEO 600 currently reaches a displacement sensitivity of $3 \cdot 10^{-19}\mathrm{m}/\sqrt{\mathrm{Hz}}$ for an end mirror (MCe, MCn), which corresponds to an accuracy of the phase readout of, $\phi = 4\pi \cdot 3 \cdot 10^{-19}\mathrm{m}/\sqrt{\mathrm{Hz}}/\lambda \approx 3.5 \cdot 10^{-12}\mathrm{rad}/\sqrt{\mathrm{Hz}}$. If we now assume the carrier light inside the Michelson interferometer to be represented by E_C with the amplitude A_C, we can calculate which minimal amplitude, A_S, of a potential stray light field E_S is required in order to shift the phase of the total field $E_T = E_C + E_S$ by ϕ. Assuming E_C and E_S to be 90 degrees out of phase we get $A_S = A_C \cdot tan(\phi)$. With the actual intra cavity power $A_C = 2.7\,\mathrm{kW}$ we get minimal stray light power of $A_S^2 \approx 3 \cdot 10^{-20}$ W.

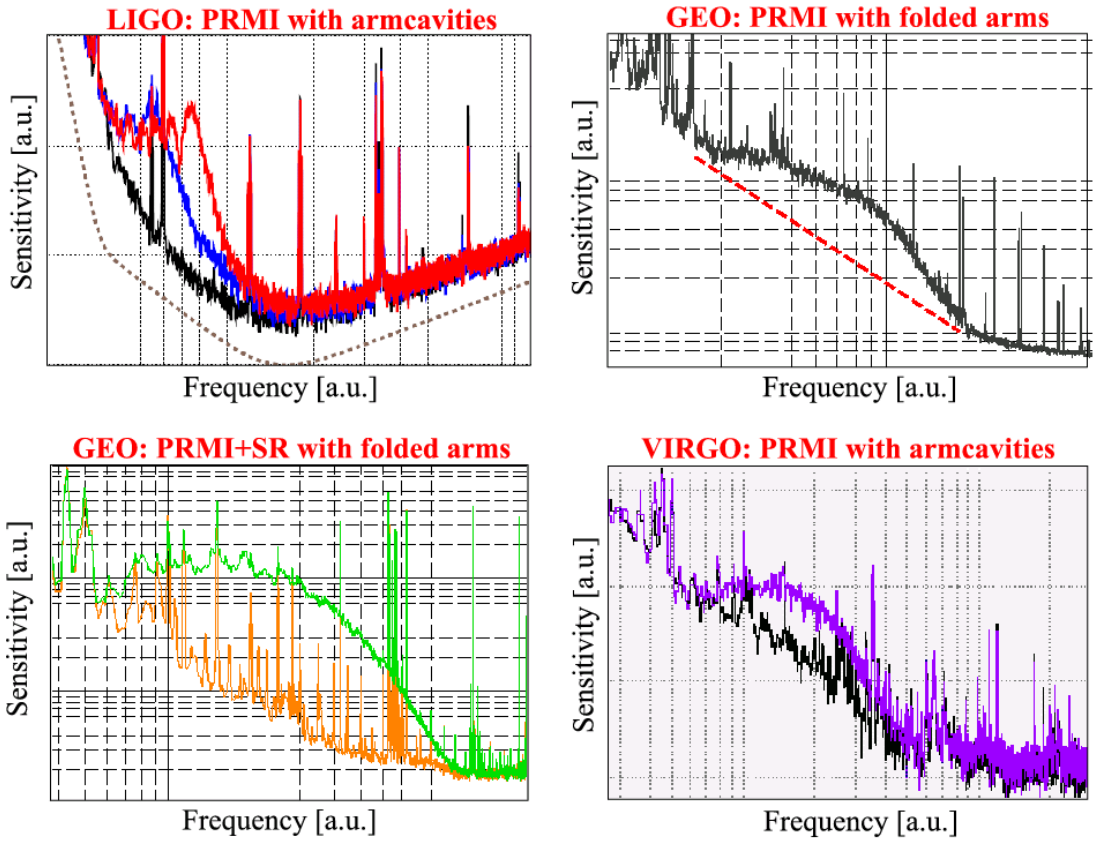


Figure 3.1.: In all interferometric gravitational wave detectors stray light is showing up in shape of a scattering shoulder. *UPPER LEFT PLOT:* Stray light shoulder in the LIGO Livingston detector, which is a power recycled Michelson interferometer of 4 km arm length using arm cavities [LIGO05]. *UPPER RIGHT PLOT:* Stray light shoulder in the GEO 600 detector, which was at the time shown here a power recycled Michelson interferometer using single folded arms of 600 m length. *LOWER LEFT PLOT:* Stray light shoulder in the dual recycled GEO 600 detector, which is a power and signal recycled Michelson interferometer with 600 m long folded arms. *LOWER RIGHT PLOT:* Stray light shoulder in the VIRGO detector, which is a power recycled Michelson interferometer of 3 km arm length using arm cavities [ILIAS06]. All four subplots have logarithmic x and y-axis.

the interferometer, i.e. the Herriot delay lines [Winkler]. The effect scattered light interacting with the walls of the vacuum vessel is theoretical described in [Vinet96]. In more recent times in the Japanese Tama300 detector stray light reflecting from the vacuum tubes was experimentally observed and characterized [Takahaschi04].

Learned from the experience collected in the prototypes, both the core optics and the beam tube of GEO 600 have been designed with great care with respect to stray light. So far there was no indication for stray light problems related to either the tube or the main optics.

The scattered light contributions observed in GEO 600 seem to originate from auxiliary beam paths where light is intentionally coupled out of the main interferometer for the purpose of sensing of the angular and longitudinal degrees of freedom. If some light from these paths is scattered back towards the interferometer it can reenter the main interferometer beam path. This light has, depending on the optical path length outside the interferometer, a shifted phase that varies in time with respect to the original light and is detected as noise in the output.

This chapter describes the experience with scattered light in the GEO 600 detector from the last two years. Powerful tools for observation, identification and elimination of stray light limiting the sensitivity of the GEO 600 detector have been developed and will be presented in this chapter.

3.2. Controlled stray light injections

This section describes experiments done with the goal to verify the general understanding of the stray lights effects observed in GEO 600. This is done by controlled generation of stray light in auxiliary beam pathes outside the vacuum system. For the experiments shown in this section the auxiliary path behind MFn and the one on the SR bench (a beam originating from a reflection on the AR coated side of the beam splitter) are used (see Appendix D for an optical layout of GEO 600).

3.2.1. How to produce stray light sowing up in H at 1 kHz

The scattering noise shoulders observed in GEO 600 extend to high frequencies of up to about 1 kHz (see for example Figure 3.6). In order to generate stray light effects inside

the GW detector at a frequency of 1 kHz one can image two ways:

- **The "high frequency small amplitude" scenario:** An optical component moving with a frequency of 1 kHz only needs to move by a fraction of the laser wavelength, in our case a micron. In presence of acoustical and seismic excitation most optical components, mounted in a standard way onto any optical table, show some movement at high frequencies, i.e. their resonance frequencies.
- **The "low frequency large amplitude" scenario:** An optical component moving with a very low frequency but with an large amplitude. Even though all suspension in GEO 600 are damped at their resonance frequencies they show a significant movement at low frequencies. Suppose a suspended optic with a resonance frequency of 1 Hz moves with an amplitude of 0.25 mm, then this will, as it will be shown in Equation 3.2.3, generate stray light with a phase shift of $2000 \cdot \pi$ showing up at 1 kHz in the sensitivity.

These two scenarios are extremes and of course an intermediate state, for example an component moving with an amplitude of 10 microns and a frequency of 25 Hz, would also be able to produce a stray light component showing up at a frequency of 1 kHz in the sensitivity. Often the stray light noise originates from number different vibrations at different frequencies. If for example the excess stray light is originating from a beam director on an optical table the surface of the mirror will at least to some amount vibrate with all kind of frequencies corresponding to the resonance frequencies of the optical table, the mirror mount and the mirror itself. Since the movement of the beam director has to be measured with respect to the main interferometer, which is moving up to several tens of microns at frequencies around 1 Hz in addition to the small amplitude high frequency movement we get also a low frequency large amplitude component.[2]

3.2.2. A device for controlled stray light generation

It is desirable to prove whether the scattered light we find in the GEO 600 detector can really be described fully by the picture of a moving optical component. This can be done by placing a scattering surface in an auxiliary beam path and moving the surface of the

[2] A third mechanism for stray light noise generation is known, but was so far not observed in GEO 600: Laser frequency noise in combination with stray light that travels a different path length with respect to the main beam can create stray light noise even without any moving component.

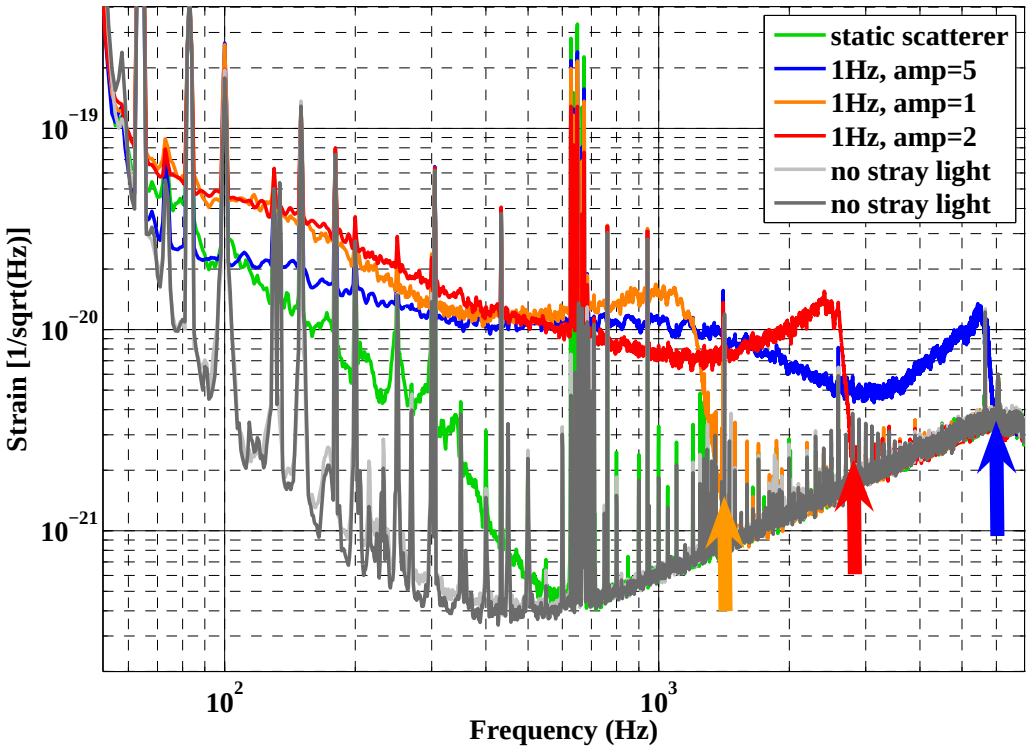


Figure 3.2.: Controlled forced scattering at the signal recycling bench: driving the stray light source with a triangular voltage of 1 Hz and amplitudes in the millimeter range. The two gray traces show the reference sensitivity for no stray light present. The green curve represents the scattered light contribution to the sensitivity for the speaker being installed but not driven with any signal. We see the usual scattering shoulder originating from the relative movement of the interferometer and the acoustically excited speaker diaphragm. The traces in orange, red and blue indicate the speaker driven with 1 Hz and different amplitudes. The cutoff frequency of the orange shoulder is around 1.3 kHz which corresponds to an amplitude of the movement of about 0.35 mm of the cone surface. When the amplitude of the triangular is increased by a factor of 2 (red trace) the cutoff frequency is doubled to 2.6 kHz as well. When the amplitude is further increased to a factor of 5 a cutoff frequency of about 5.5 kHz is observed which corresponds to a peak to peak movement of the scattering source of nearly 1.5 mm. This is already quite a large amplitude for the speaker and is probably the reason for the non-linearity of the observed cutoff frequency. Overall this experiment confirms our present understanding of the low frequency large amplitude scenario.

stray light source in a controlled way. We realized such a controllable scattering source

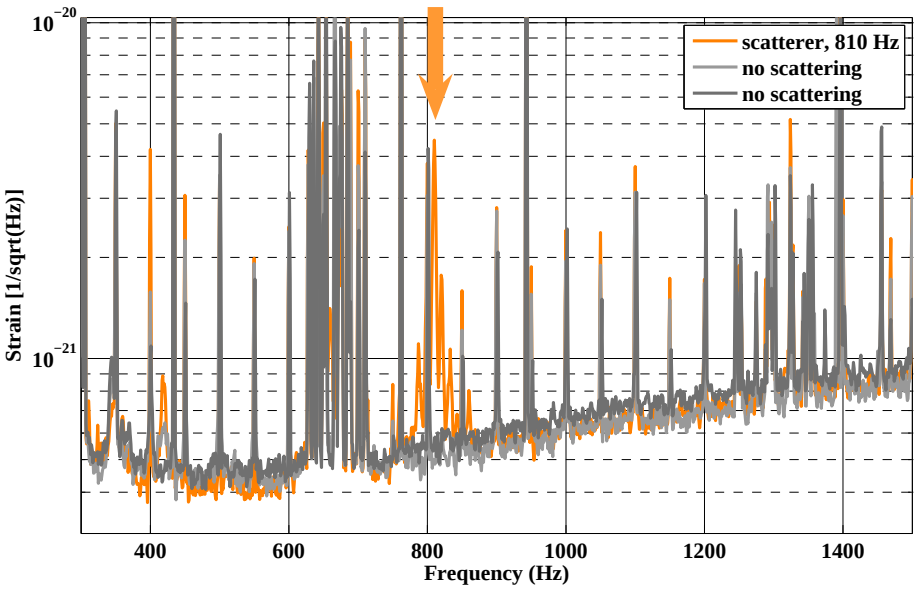


Figure 3.3.: Controlled forced scattering from the beam path transmitted through MFn: The two gray traces show reference spectra without any scattered light contribution. The orange curve represents a time when the stray light source (speaker) was driven with 810 Hz and an amplitude of less than a laser wavelength. A noise peak at 810 Hz surrounded by a strong sideband structure is observed.

by using an commercial low-cost loudspeaker driven by a commercial audio amplifier. In order to increase the backscattering amplitude on the one hand and to reduce the risk of thermal damage of the speaker caused by absorption of the laser light on the other hand we chose a speaker with a rough and silvery metal diaphragm (anodised aluminium). In addition we used the bare speaker without any housing to reduce the sound emission of the speaker. A speaker with high compliance was chosen to allow a total travel of several millimeters. The scattering efficiency was maximized by positioning the surface of the speaker close to a beam waist (see Section 3.6). The speaker was fixed into a strong metal frame with a mount to allow a rigid connection to the optical table.

3.2.3. Low frequency large amplitude scenario

First we want to verify that a low frequency movement of the speaker cone with a large amplitude produces a smooth scattering shoulder. Figure 3.2 shows the corresponding measurement, where the speaker was driven with triangular voltage of a frequency of 1 Hz. Shown are calibrated strain sensitivities for two reference times without any stray light contribution (gray), a reference for the static scattering source, where the speaker was installed but not driven (green) and three times when the speaker was driven in a controlled way with amplitudes in the mm-range. The static speaker already produces a strong scattering shoulder originating from the relative movement of the speaker surface and the suspended main interferometer. The cutoff frequency, f_{cutoff}, can roughly be estimated by the following relation

$$f_{\text{cutoff}} \approx 4 \cdot f_{\text{sp}} \cdot \frac{A_{\text{sp}}}{\lambda_{\text{laser}}}, \tag{3.1}$$

where f_{sp} is the frequency of the signal the speaker is driven by, A_{sp} is the peak to peak amplitude of the diaphragm and λ_{laser} is the wavelength of the laser. The factor 4 originates from a factor of 2, accounting for the fact that the light hitting the speaker and being reflected travels twice the distance the diaphragm moves, and a second factor of 2, originating from the fact that during a single period of the triangular signal 2 times the peak-peak movement is covered. For the measurement corresponding to the red trace of Figure 3.2 an amplitude twice as large as for the measurement corresponding to the orange trace was chosen. In good agreement the cutoff frequency of the red trace, 2.6 kHz, is found to be twice as large as the cutoff frequency of the orange curve, 1.3 kHz. However, caused by the non linear response of the speaker for large amplitudes, we find for the blue trace instead of the expected cutoff frequency of 6.5 kHz only 5.5 kHz. However, overall it can be stated that our understanding of the low frequency large amplitude scenario is confirmed by this measurement.

3.2.4. High frequency low amplitude scenario

The next measurement was done to prove the high frequency low amplitude scenario. The speaker was driven by a triangular signal with a frequency of 810 Hz and an amplitude much smaller than one laser wavelength[3]. Figure 3.3 shows this measurement

[3] That the amplitude was much smaller than the laser wavelength could be determined by the fact that no additional line in sensitivity was observed at 1620 Hz.

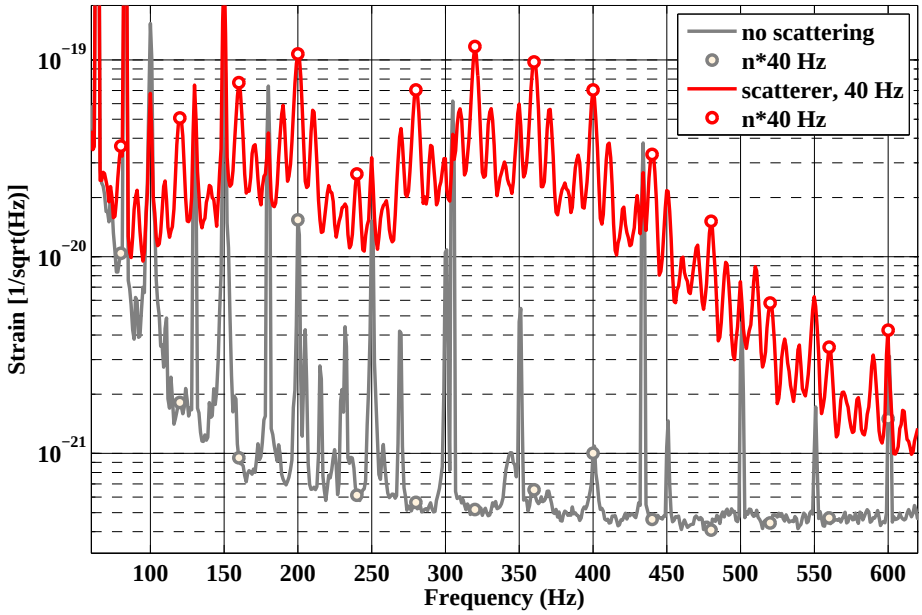


Figure 3.4.: Scattered light injection by controlled movement of the speaker surfaces. A triangle with a frequency of 40 Hz and an amplitude of a few ten microns.

in detail. The two gray traces represent reference spectra in the absence of any stray light contribution. The orange curve represents the time of the stray light injection. A large noise peak is generated with a strong sideband structure around it. The dominat sideband frequencies are 10 and 23 Hz with amplitudes of 50 and 25 % of the center peak, respectively. The sideband structure represents the up-converted low frequency movement of the speaker relative to the main interferometer. Comparisons of measurements for stray light injections at different frequency have shown identical sideband structures for all frequencies. The same is also true for measurements of injections with amplitudes of a few laser wavelengths, thus the first few harmonic frequencies with their sideband structures could be observed.

3.2.5. Combination of the two scenarios

Figure 3.4 shows a measurement where the speaker was driven with a frequency of 40 Hz and an amplitude of several tens of microns. As expected we find a comb of harmonics of 40 Hz (red circles). Each line of a frequency n*40 Hz is surrounded by the same

sideband structure that was already observed in Figure 3.3. That is why each line is accompanied by 10 Hz sidebands. Altogether this excitation of the speaker diaphragm gives a scattering shoulder containing strong peaks at all harmonic frequencies of 40 Hz. As we will see later (Section 3.7) this stray light injection shows some similarities to an observed stray light phenomen called *ripple noise*.

3.2.6. Summary of the stray light injections

We found a simple and very efficient way to generate stray light in an controlled way by using the cone of a loudspeaker as scattering surface. By carrying out stray light injections with various amplitudes and frequencies we found the excess noise in the sensitivity originating from stray light to qualitatively match our expectations. This means that we understand the basic principle of scattered light problems and there seems no "magic" about them.

3.3. Indications for a stray light limited sensitivity

During the commissioning of a gravitational wave detector one of the main problems regarding stray light is to find out whether the current detector sensitivity is in any frequency band limited by stray light effects. The tiny amounts of stray light that can already limit the sensitivity of GEO 600 are hardly detectable by independent sensors.

Therefore the most promising method to identify whether there is any stray light contribution is to use the specific signature of the stray light in the sensitivity of the GW detector. In GEO we found stray light always to show up in the sensitivity with three specific characteristics:

- **Non-stationarity on short time scales:** The noise originating from stray light is in many cases found to be highly non-stationary on second's time scales. In contrast to most of the technical noise, for instance feedback or electrical noises, and most of the fundamental noise, for instance thermal noise or shot noise, which are rather stationary. Figure 3.5 shows spectrograms of the uncalibrated error signal of the differential arm length servo for a time without any contribution from stray light and a time of intentionally increased scattered light contribution. The stray light shows up as strongly varying broadband excess noise. Not only the size of

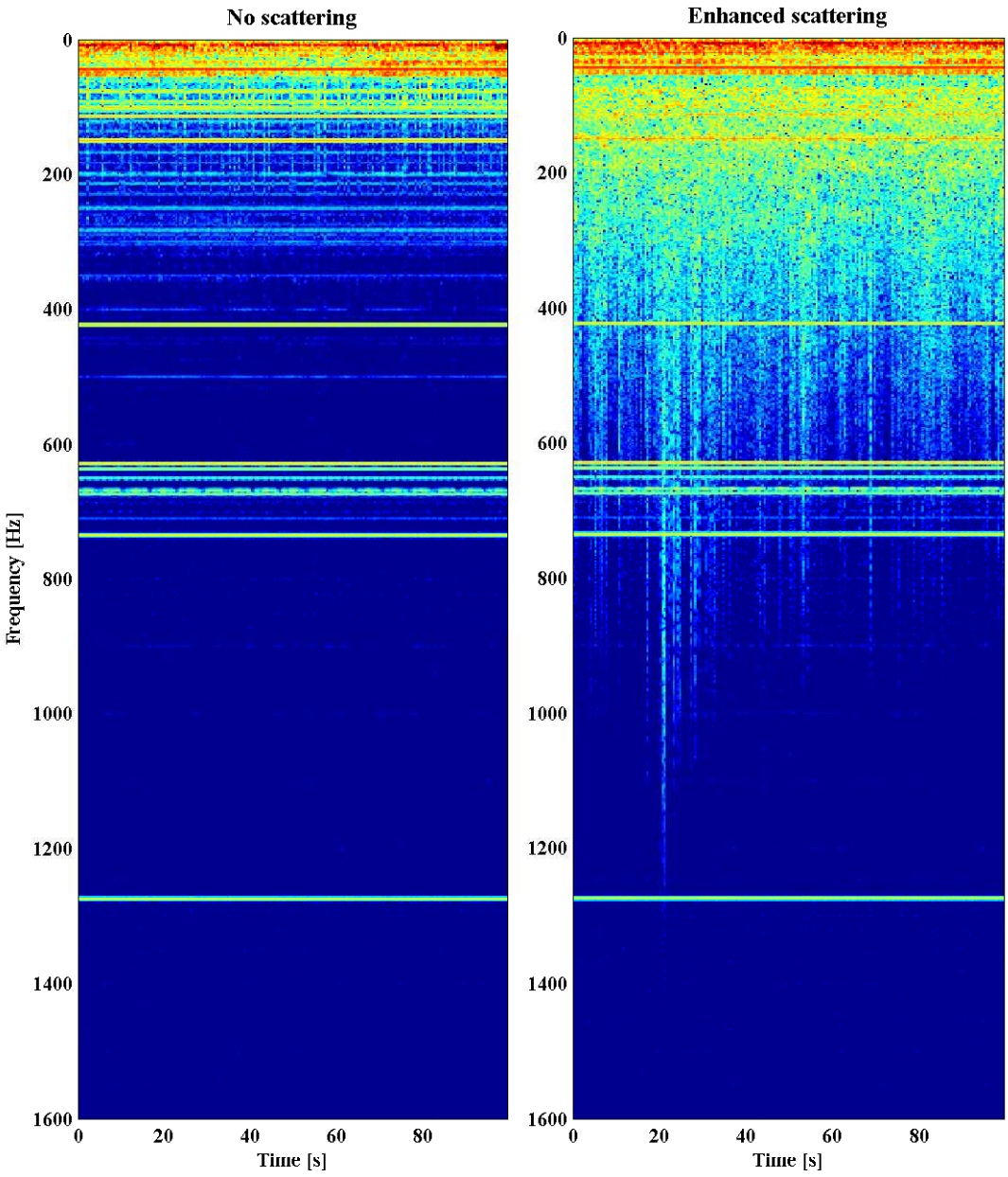


Figure 3.5.: Spectrograms of the uncalibrated error signal (P quadrature) of the differential arm length servo. The left hand plot shows 100 seconds of "clean" data without enhanced scattered light limiting the sensitivity. For frequencies above a few hundred the sensitivity is reasonable stationary. The right plot show 100 seconds of data from a time when the scattering was intentionally enhanced. Over most of the detection band the sensitivity varies strongly on the time scale of a second.

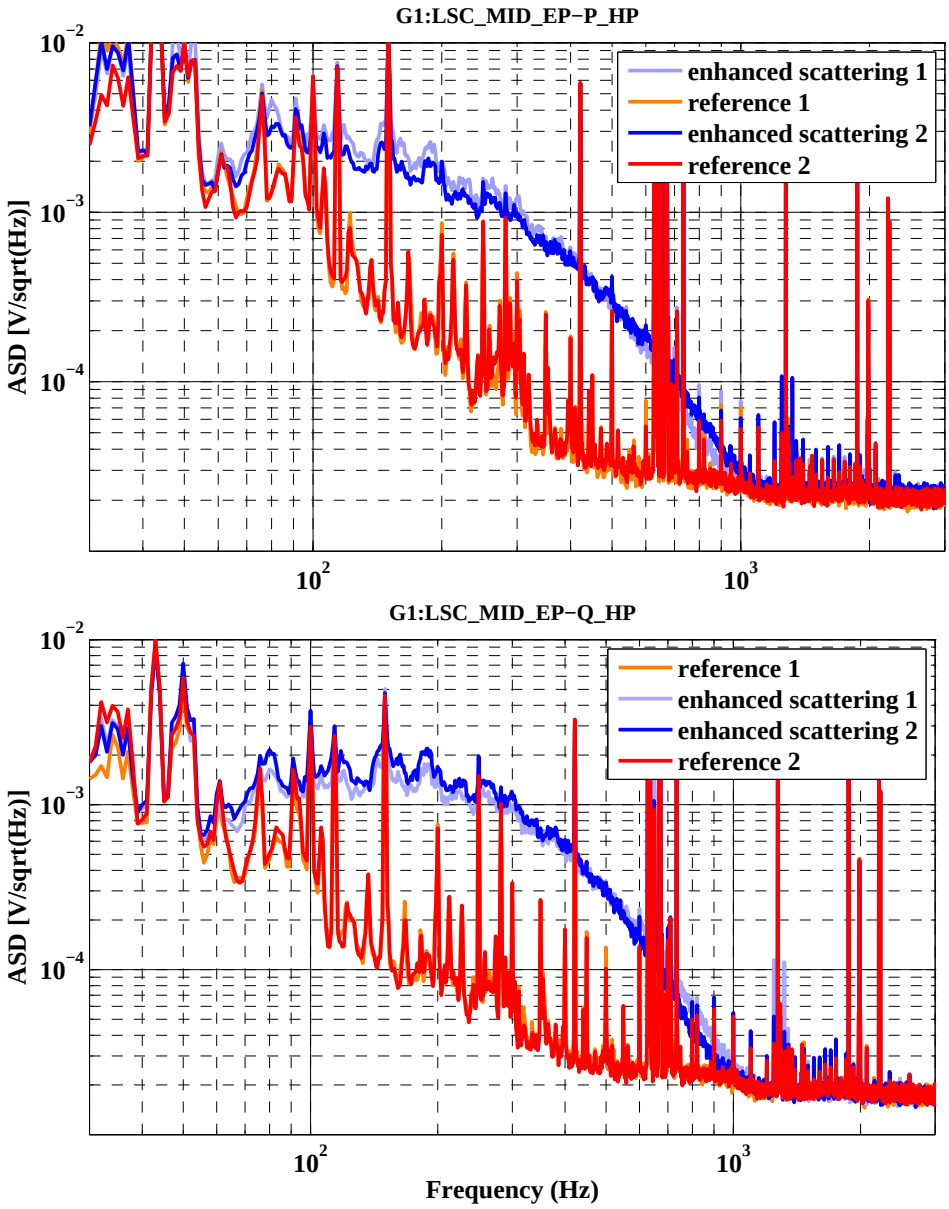


Figure 3.6.: Spectra of the uncalibrated error signal of the differential arm length servo. Shown are both output quadratures, P and Q, for each two times of "clean" data (minimal noise contribution from scattering) in red and orange and two times of enhanced scattering limiting the sensitivity from 60 to 1000 Hz (blue and light blue). The stray light causes excess noise in the typical shoulder shape. Below 100 Hz the stray light is suppressed by the loop gain of the servo controlling the differential arm length of the Michelson interferometer.

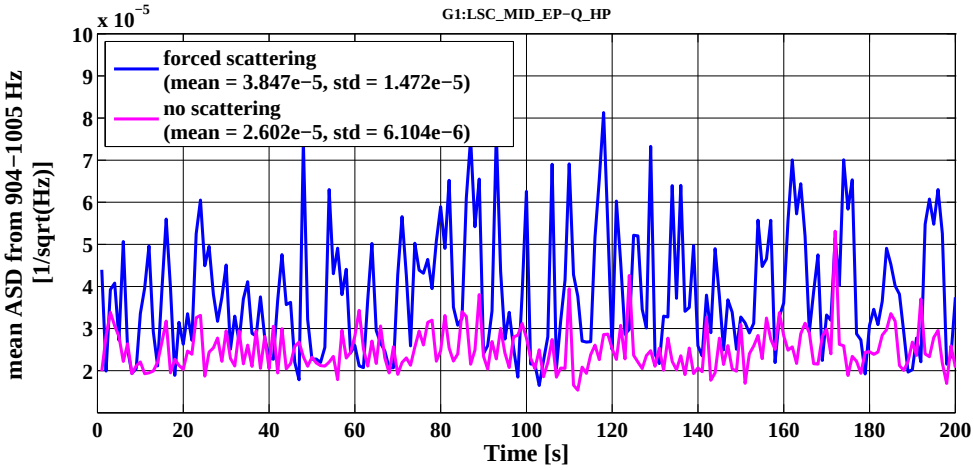


Figure 3.7.: Time series of the mean ASD of the uncalibrated detector sensitivity in the band from 904 to 1005 Hz. The pink trace represents a time when the frequency band was not limited by stray light, while the blue trace indicates a time of enhanced scattering. If stray light is present not only the mean of the time series is larger but also the standard deviation, σ, is strongly increased.

the scattered light noise changes but also the maximal effected frequency. Another way to look at the stationarity of the noise is plotted in Figure 3.7. Depicted are the time series of the mean ASD of the uncalibrated detector sensitivity in the band from 904 to 1005 Hz for two different times. The pink trace represents a time when the frequency band was not limited by stray light, while the blue trace indicates a time of enhanced scattering. If stray light is present not only the mean of the time series is larger but also the standard deviation, σ, is strongly increased.

- **The specific sound of stray light:** An excellent tool to judge the stationarity of a signal on short time scales is the human ear. Listening to a whitened signal representing the sensitivity was found to be the most powerful tool for identifying scattered light noise. At least in GEO stray light had a very specific sound. Two audio examples, of the sound of the sensitivity of GEO 600 for a time of clean data and a time when excess noise from stray light was present, can be found on the CD-ROM (Multimedia Appendix) at the end of this thesis.[4]

[4]The unique sound of stray light is hard to describe in words. Probably everybody of the GEO commissioning team would describe it in a different way. Some call it ”wooshing”, others compare it

- **The typical scattering shoulder:** As already mentioned in Section 3.1 noise from stray light is often found to show up as a shoulder of excess noise. As Figure 3.6 indicates the shoulder starts at low frequency and then extends to a certain frequencies called the *cutoff frequency*. In many cases the cutoff frequency scales with the size of the shoulder at low frequencies as one can see in the righthand spectrogram of Figure 3.5.

Even though none of these three points alone can give a reliable indication, all three together give a pretty strong indication for stray light limiting the sensitivity of the GW detector. If all three indicators suggest the presence of stray light, the next step is to find the source and coupling path of the stray light.

3.4. The filter experiment: A reliable method to identify the source of stray light

In the commissioning of the GEO 600 detector the so called *filter experiment* was found to be the best method to identify and locate limiting scattering paths [Hild06a]. Many potential scattering sources such as the auxiliary beam paths used for control and diagnosis of the detector are located outside the vacuum system. Auxiliary beam paths used for longitudinal and angular control of the interferometer usually contain many optical components like for instance lenses, telescopes or photodiodes, each of these a potential source of scattering.

Let us consider a simplified setup as it is depicted in Figure 3.8 to explain how the filter experiment works. A fraction, A_0, of the light circulating in the main interferometer is entering the auxiliary beam path of interest. A potential source of stray light scatters the light amplitude $A_0 \cdot k_1$ back towards the main interferometer, where k_1 is the scattering coefficient. Usually only a small fraction of $A_0 \cdot k_1 \cdot k_2$ then re-enters the main interferometer mode and is detected. If now an optical attenuator with the amplitude transmission of T_f is positioned between the main interferometer and the scattering source, first of all, the light incident on the scattering source is decreased and afterwards the scattered light going back into the interferometer is also attenuated. Overall the light amplitude re-entering the detection path is reduced from $A_0 \cdot k_1 \cdot k_2$ to $A_0 \cdot T_f^2 \cdot k_1 \cdot k_2$ by inserting the optical attenuator.

to an avalanche or a storm. I think the best way is to listen yourself.

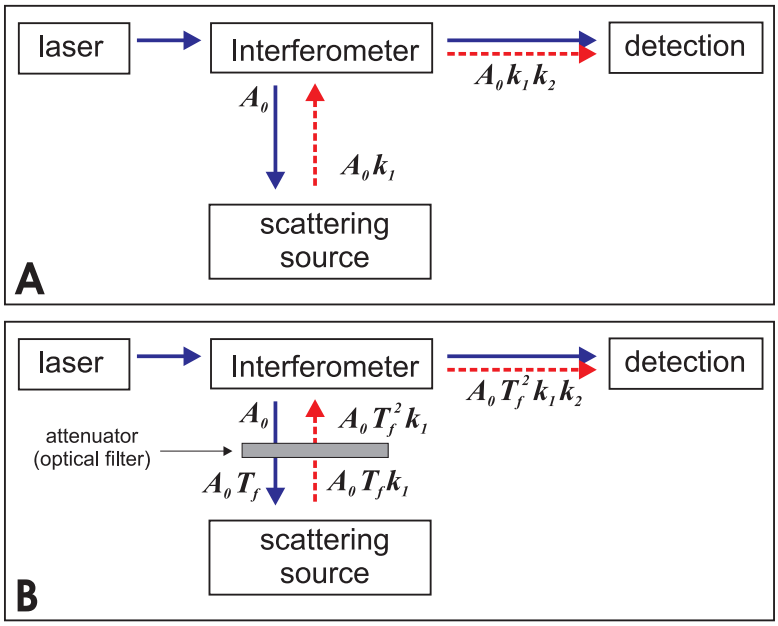


Figure 3.8.: Sketch of the filter experiment which provides an easy and reliable method to identify stray light limiting the sensitivity of the interferometer. **Figure A:** A fraction, A_0, of the light circulating in the interferometer leaves the main interferometer and enters an auxiliary beam path (for example used for alignment control). A scattering source in this auxiliary path would scatter $A_0 \cdot k_1$ back towards the interferometer, where k_1 is the scattering coefficient. Assuming the coefficient for the coupling back into the main interferometer to be k_2 a total light amplitude of the scattered light of $A_0 \cdot k_1 \cdot k_2$ is detected. **Figure B:** If an optical attenuator with the amplitude transmission of T_f is inserted between the main interferometer and the source of stray light the detected stray light amplitude is decreased by T_f^2.

If, for example, a neutral density filter is used as an attenuator, which has a power transmission of about 60 % at $\lambda = 1064\,\mathrm{nm}$ we expect the sensitivity limiting stray light contribution to go down by 60 % as well. Now by putting in the filter at different positions in the auxiliary path we can identify component by component their contribution to the detected scattering. However, as the inserted filter is an additional potential source for scattering it is necessary to take great care of the quality and cleanliness of the filter in order to not spoil the measurements. To decrease the chance of added scattering the filter must not be installed under normal incidence. In addition, as it is shown in Section 3.6 it is essential to avoid placing the filter near a beam waist.

Of course in the case the light in the auxiliary beam is used for automatic control of the interferometer it is necessary to adjust the gain of the servo in order to compensate for less light power hitting the photo detector.

3.4.1. Limitations of the filter experiment

Obviously it is not easily possible to insert an attenuator to any beam path inside the vacuum system. Therefore the filter experiment is limited to the parts of the auxiliary beams that are located outside the vacuum system. Another restriction can be given for scattering sources in the main detection path: For many conditions the potential benefit from reducing stray light in this case is compensated by less signal on the main photo detector. If we take for example a sensitivity which is at the frequencies of interest limited to equal shares by stray light and shot noise the stray light contribution[5] will decrease by T_f as well as the shot noise level will increase by T_f.

Finally the strongest limitation of the method is that only stray light contributions that actually limit the total detector noise can be identified. As long as the stray light does not contribute significantly to the detector noise we have no measure whether the stray light was reduced or not. This method does not allow for projecting the stray light noise.

[5] In opposite to the scenario described in Figure 3.8, when the attenuator is placed in the detection path the scattered light is only attenuated once.

3.5. Experience of stray light problems in GEO

In fall 2004 the sensitivity of GEO 600 was found to be limited by stray light induced noise over a wide range of the detection band. A typical scattering shoulder with a cutoff frequency of about 1.2 kHz was traced back to backscattering on the signal recycling bench. The steps involved in reducing the stray light on this optical table are a good example how to eliminate stray light problems in the optical setup of an auxiliary path.

On the signal recycling bench a pick-off beam, originating from the beam splitter AR coating in the east arm, is detected and used for longitudinal control and auto alignment of the signal recycling mirror (see Figure D.1). Figure 3.9A shows the original layout of the bench as it was present in late summer 2004. The beam leaves the vacuum system through an AR coated window and passes an AR coated focusing lens L1. Afterwards a set of two beam steering mirrors (bd1, bd2) is used to align the beam onto the second AR coated lens L2 of the beam telescope and a 2-axis galvo-scanner. Finally the beam is detected by a quadrant diode. With the aid of the filter experiment we identified L2 and the galvo scanner as sources of limiting stray light noise. A reduction of the stray light was achieved by replacing L2 and the galvo scanner by a high quality beam director[6]. In a next step we found bd3 still causing a measurable amount of stray light coupling to the GW channel. This was mainly caused by the fact that bd3 was located close to a beam waist, which is in terms of stray light in a very critical position for any kind of optical elements, regardless of their quality (see Section 3.6). Therefore the quadrant diode was moved in a next step closer towards L1 in order to avoid the presence of any beam waist on the SR bench. The corresponding layout which was used during the S4 science run is depicted in Figure 3.9C. Figure 3.10 shows the sensitivity improvements from autumn 2004 achieved by minimizing stray light from the signal recycling bench.

The curious reader might ask, why the optical layout of the bench was reorganized, instead of simply attenuating the stray light by installing several optical attenuators (permanent filter experiment). In general such an approach should have worked out, even though it always feels better to eliminate the source of a problem instead of suppressing the consequences. However, in the special case of the GEO SR bench any attenuation of the light level would have caused an increased feedback noise of the signal recycling

[6] In GEO galvo scanners are commonly used in front of all quadrant diodes in order to make sure the beams are well centered on the diodes all the time. However, after taking out the scanner from the signal recycling bench no excess noise was found in any of the related control loops.

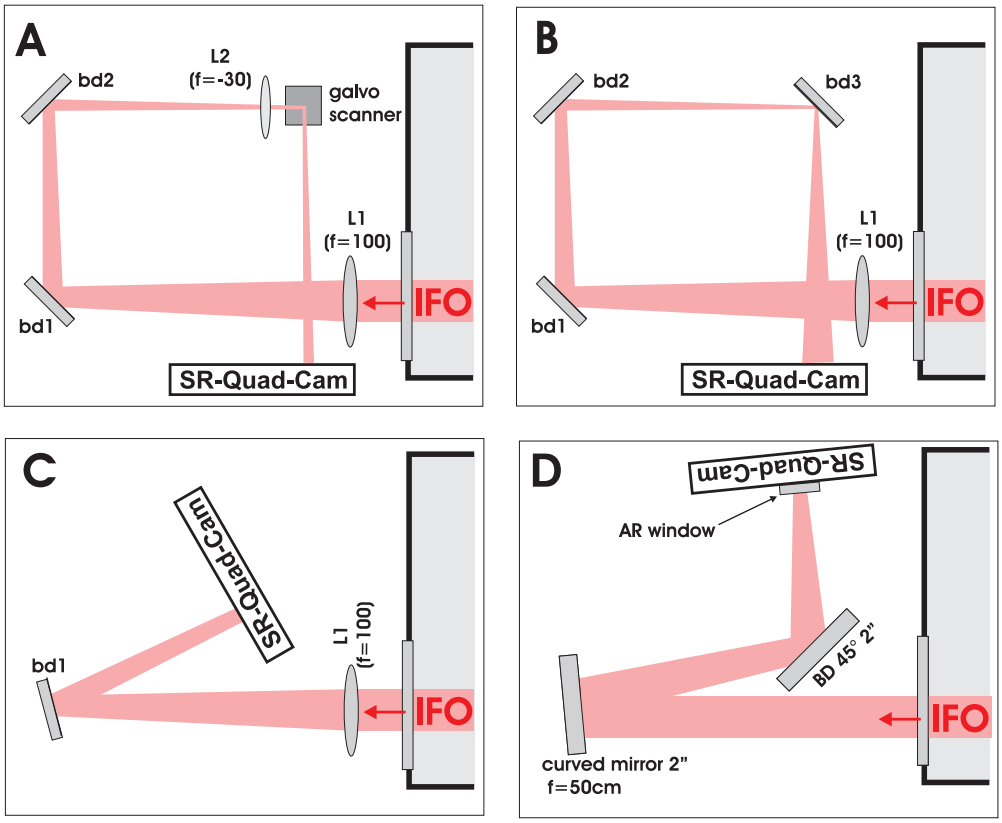


Figure 3.9.: Progression of the optical layout of the signal recycling bench in order to reduce stray light. **A:** Original layout of the bench from summer 2004. The beam leaves the vacuum system through a window, passes a focusing lens L1 and two stearing mirrors (bd1, bd2) which are used to align the beam to the second lens L2 of the beam telescope and 2 axis galvo-scanner. Finally the beam is detected by a quadrant diode. **B:** With the aid of the filter experiment we identified L2 and the galvo scanner as sources of limiting stray light noise. Both components were taken out and replaced by a single high quality mirror (bd3). **C:** S4 configuration. In a next step we found bd3 still causing a measurable amount of stray light. This was mainly caused by the fact that bd3 was located close to a beam waist, which is in terms of stray light in a very critical position for any kind of optical elements, regardless of their quality (see Sec.3.6). Therefore the quadrant diode was moved in a next step closer towards L1 in order to avoid the presence of a beam waist on the SR bench. **D:** S5 configuration. At the start of 2006 we found again stray light limiting the improved sensitivity of GEO 600. The un-coated window of the quadrat photodiode was replaced by an AR coated fused silica window. In a second step L1 was replaced by a curved 2" diameter mirror.

longitudinal loop which is limited by shot noise for frequencies of our interest. As the feedback noise of this loop is already close to limiting the overall sensitivity of GEO 600 in the few hundred Hz region, any reduction of the stray light noise by attenuating the beam path would have been compensated by larger feedback noise in the attenuated setup.

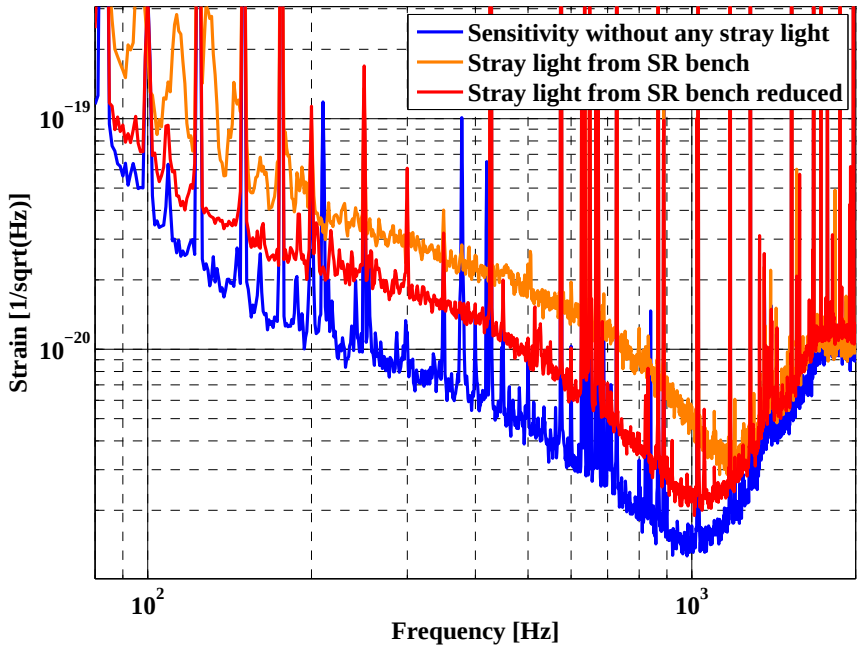


Figure 3.10.: Sensitivity progress of GEO 600 from autumn 2004. By eliminating stray light from the signal recycling bench the sensitivity of the GEO 600 detector was improved by about a factor of 3 for frequencies between 100 and 1000 Hz.

After S4 GEO's sensitivity was improved by about a factor of 10 in the few hundred Hz region at the beginning of 2006 again scattered light originating from the SR bench was found to limit the sensitivity. The dominating source of stray light was the low quality window of the quadrant photo detector. The original window was cut off and the diode was clamped in between two copper plates of which one was holding a 1" AR coated fused silica window. The AUTOCAD drawing of this diode holder can be found in Appendix C. The major difficulty of realizing this setup is to cleanly cut off the window of the quadrant diode. In contrast to small single element diodes where the window can easily be taken off, removing the window of a large quadrant diode poses difficulties. Due to the much larger ratio of diameter to thickness of the housing it turns

out to be nearly impossible to cut off the window without partly cracking it and spilling many small splinters over the active area of the diode. The only reliable way to get rid of the splinters was found to be repeated rinsing the whole diode with a few milliliters of ethanol and immediately afterwards drying it with compressed nitrogen.

In a second step of improvement the lens (L1) was replaced by a high quality 2" curved mirror. There was no strong indication that L1 was causing any additional stray light noise. However, in contrast to a curved mirror lenses always have the disadvantage that a fraction of their surface is always perpendicular with respect to the incidence beam. Finally it has to be mentioned that all secondary beams, such as the reflected beam from the photodiode itself, have to be properly dumped.

3.6. The cat's eye effect

As already stated in the previous section one important experimental finding of our stray light investigations was that stray light sources positioned close to a beam waist caused the largest stray light contributions to the GEO 600 sensitivity. Analogous to the cat's eye effect the light scattered directly at a beam waist has exactly the right geometrical mode to reenter the main interferometer.

Figure 3.11 shows an illustrative example of an optical configuration which is similar to the signal recycling bench described in the previous section. The subplot 1 shows the setup with no stray light present. Inside the central cluster the main interferometer beam has an radius of $\omega_{cc} \approx 1\,\mathrm{cm}$. A fraction of the beam leaves the interferometer at the beam splitter AR coating and is then focused down on the SR bench. With the drawn lens of 50 cm focal length an beam waist of $\omega_0 \approx 17\,\mu\mathrm{m}$ is generated. For the following analysis we assume the light amplitude incident to the lens is A_0 and all beams have an gaussian profile at all positions. For all scenarios described below we will use the same source of scattering which is a square area of $5 \times 5\,\mu\mathrm{m}$ which is assumed to scatter proportional to $cos(\varphi)$ into the solid angle, where $\varphi = 0$ corresponds to normal incidence.

Figure 3.11.2 shows the situation for the scattering square directly located in the waist. Since the beam is so small at the waist, A_2 which is the fraction of A_0 incident to the

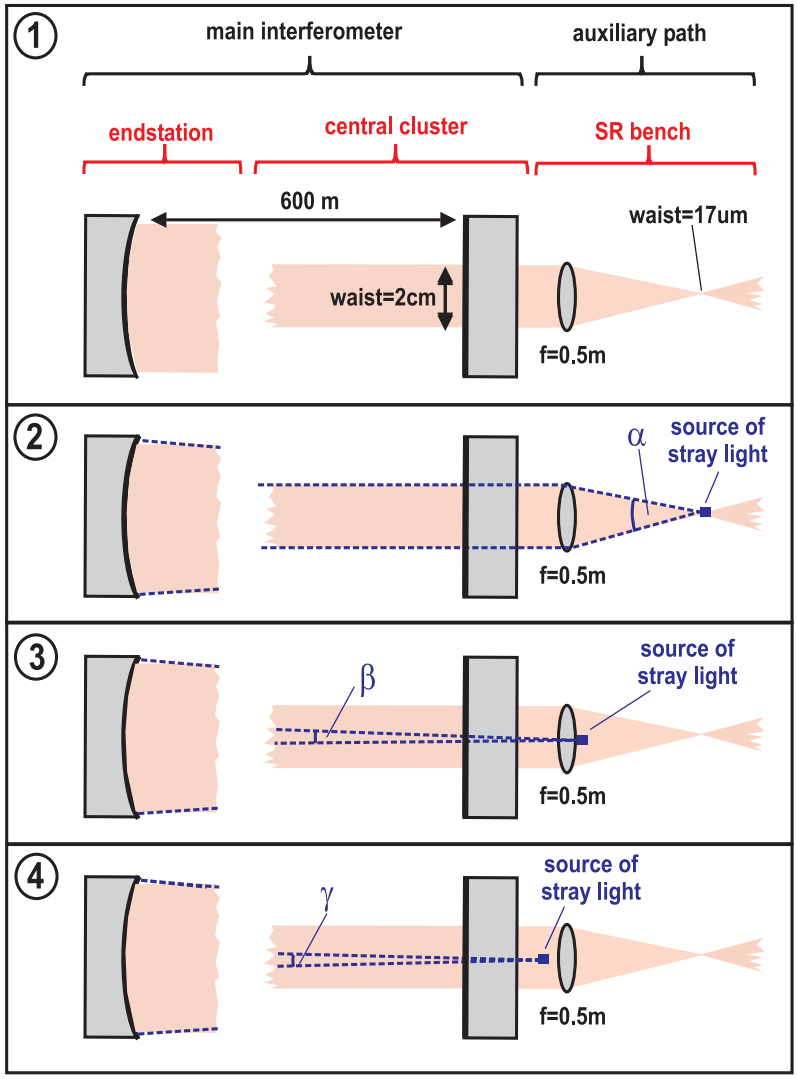


Figure 3.11.: Illustrating schematic of the cat's eye effect: The scattering efficiency of a scattering source strongly depends on its position in the auxiliary path. An extremely large back scattering efficiency is found for a stray light source sitting directly in the beam waist, because in that case first of all a large light power is reflected and secondly all light scattered back towards the lens reenters the main interferometer in the correct mode. A scattering source located right in the beam waist (panel 2) is found to have a $1.6 \cdot 10^9$ larger backscattering efficiency than the same scattering source sitting in the non diverging beam in front of the lens (panel 4). A detailed description of the depicted scenarios can be found in the text.

scattering square, is quite large:

$$A_2 = 0.2 \cdot A_0. \tag{3.2}$$

All of A_2 scattered back into α, which is 2.3°, reenters the main interferometer and is transferred to the output.

The situation when the same scattering source is placed right behind the lens is shown in Figure 3.11.3. In this scenario A_3 which is the fraction of A_0 incident to the scattering square is much smaller than before due to the larger beam radius.

$$A_3 = 1.6 \cdot 10^{-7} \cdot A_0 \tag{3.3}$$

Furthermore, the light scattered back is to a large fraction not in the correct interferometer mode. An upper limit for the amount of light that can in principle[7] reenter the interferometer and being transferred to the output is given by $\beta = 0.017$ which originates from the requirement that the stray light has at least to hit the far mirror (diameter = 18 cm) at a distance of 600 m. This means that the scattering efficiency will be very small in this case due to the small accepted angle β.

Finally in Figure 3.11.4 a scenario is depicted where the scattering square is placed on the left hand side of the lens in the parallel beam. The light amplitude A_4 incident onto the square is, due to equal beam sizes, identical to A_3.

$$A_4 = 1.6 \cdot 10^{-7} \cdot A_0 \tag{3.4}$$

Also the angle of acceptance γ is equal to β.

The overall amount of light backscattered into γ, $A_{4\text{back}}$, can be described by the following expression:

$$A_{4\text{back}} \propto A_4 \int_0^{\gamma/2} \cos(\varphi) d\varphi. \tag{3.5}$$

Analogously $A_{3\text{back}}$ and $A_{2\text{back}}$ can be calculated. If we normalize the results by $A_{4\text{back}}$ we get the following normalized backscattering efficiencies.

Exactly at the beam waist the backscattering efficiency is 1.6 billion times larger than for a scattering surface at the lens. From this result we can derive the rule that we have

[7] Assuming all light is converted by mode healing from the signal recycling back into the ground mode.

Scenario	Efficiency of backscattering
2	$1.6 \cdot 10^9$
3, 4	1

Table 3.1.: Summary of the normalized backscattering efficiencies for the three scenarios depicted in Figure 3.11

to avoid placing optics exactly in the beam waist of any auxiliary beam path directly connected to the main interferometer, or even better to avoid the presence of any waist at all if possible.

3.7. A special class of stray light noise: the "ripples"

In this section a class of stray light noise is described, the so called *"ripples"*. For two reasons these ripples are very interesting: First of all they have a noteworthy shape in the frequency domain and secondly this kind of stray light noise was limiting the sensitivity of the GEO 600 detector in the few hundred Hz range at least part time of S5.

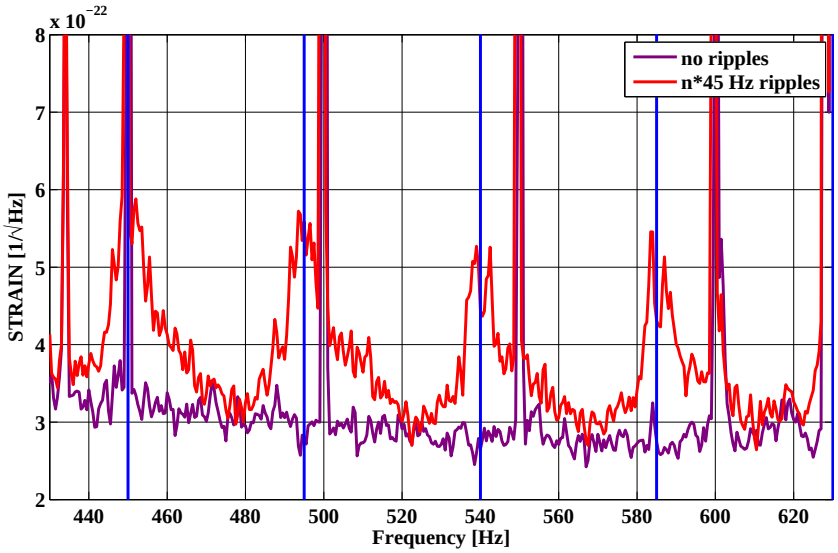


Figure 3.12.: Strain sensitivity of GEO 600 for a time when the n*45 Hz ripple noise is present (red) and reference time (purple). The vertical blue lines indicate harmonic frequencies of 45 Hz.

During S5 excess noise was frequently observed that showed up in the few hundred Hz region of the sensitivity as a series of equidistant bumps. Figure 3.12 shows spectra of a reference time (purple) when no ripples were present and a time with ripple noise (red). As indicated by the blue vertical lines which are drawn at harmonic frequencies of 45 Hz the ripples show up in this case at n*45 Hz. At the center frequency of the bumps the sensitivity is decreased by up to 80 %. The bumps have a FWHM of about 10 Hz and from the presence of their wings even in the middle between two bumps the sensitivity seems to be slightly decreased.

The n*45 Hz ripples showed up during S5 in irregular intervals of several hours and were then present for about 20 minutes. Figure 3.13 depicts how the n*45 Hz ripples shows up in the sensitivity. Plotted are spectrograms of H and the seismometer in the north end station. Both spectrograms are normalized by the overall average in each frequency bin. Reddish pixels indicate values larger than the average. Between hour 3 and 3.5 an excess noise is present at 45 Hz in the seismic. At the same time we see strong ripples at n*45 Hz in the sensitivity of GEO 600. This observation led to the suspicion that stray light is generated at a seismically driven component in the north end building. This suspicion was further confirmed by the observation of a slight drift in the frequency of the ripples, which was also seen in the frequency of the seismic excitation, as shown in Figure 3.14.

The excess noise in the seismic was tracked down to a cooling fan of a turbo-pump. The fan was automatically controlled and switched on only when a certain threshold of the turbo-pump temperature was crossed. For further experiments the cooling fan was set to continuous operation in order to continuously generate ripples. In that condition a filter experiment was carried out on the optical table in the north endstation. Even with the beam leaving the vacuum completely blocked the ripples stayed constant in size. This for the first time in GEO 600 gives strong evidence for a stray light problem originating from inside the vacuum system.

Potential stray light sources inside TFn are the catcher of MFn, the tank walls and the curved and partly un-coated rim of the view port behind MFn. A series of experiments where the beam position on MFn was changed gave the result that the n*45 Hz ripples are smallest when the beam is not centered on the mirror, but centered on the view port. This makes the view port to be the strongest suspect for being the stray light source. As shown in Figure 4.2 a significant fraction of the beam transmitted through

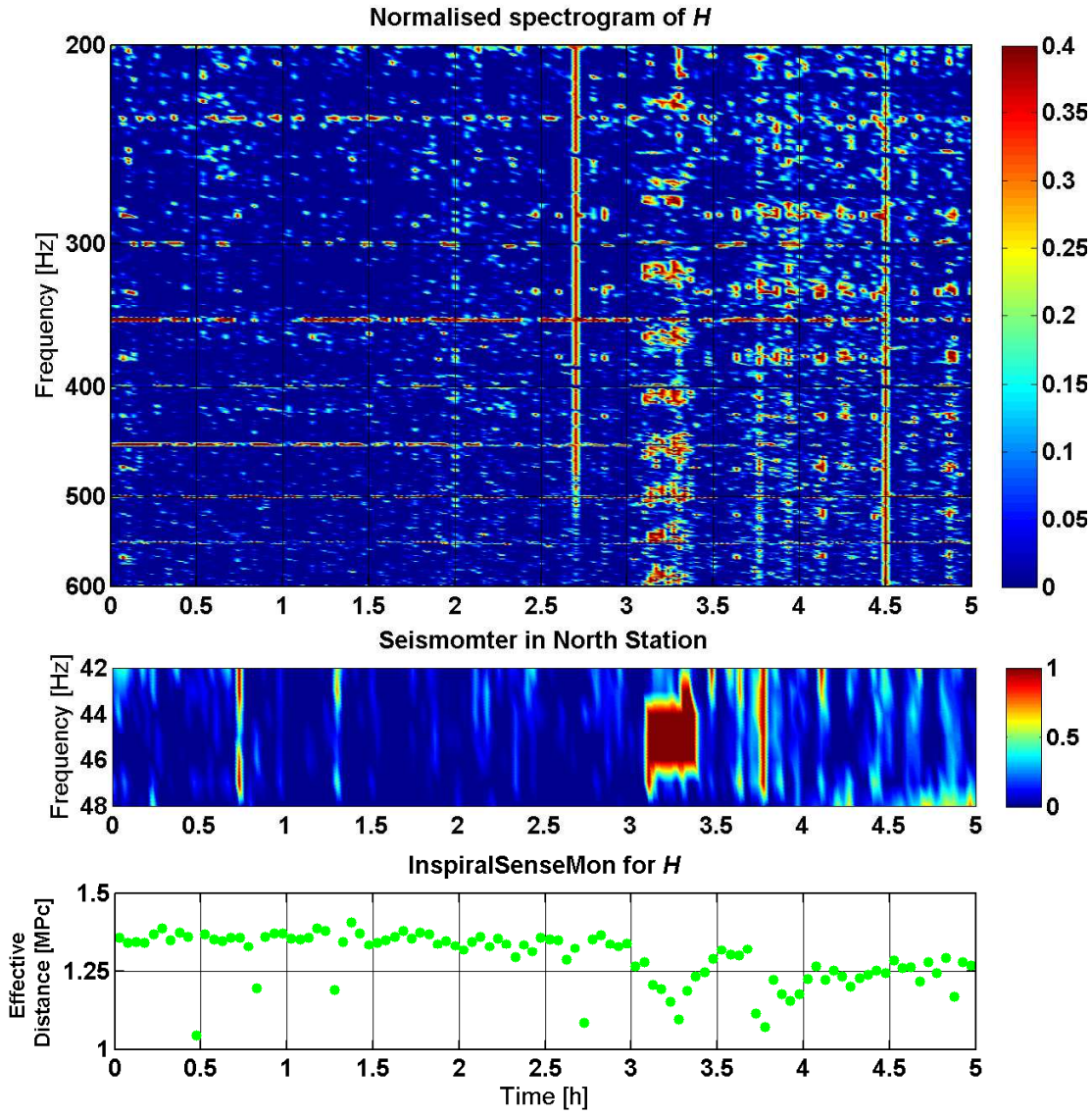


Figure 3.13.: Spectrograms of H and the seismometer in the north end station. Both spectrograms are normalized by the overall average in each frequency bin, thus the varying features are highlighted. Reddish pixels indicate values larger than the average. Between hour 3 and 3.5 an excess noise is present at 45 Hz in the seismic. At the same time we see strong ripples at n*45 Hz in the sensitivity of GEO 600. The decrease in sensitivity can also be seen in the neutron-star-neutron-star (NSNS) inspiral horizon which is decreased by about 10 % in the presence of the ripples. The NSNS inspiral horizon is calculated for neutron stars of 1.4 solar masses and optimal orientation for GEO 600.

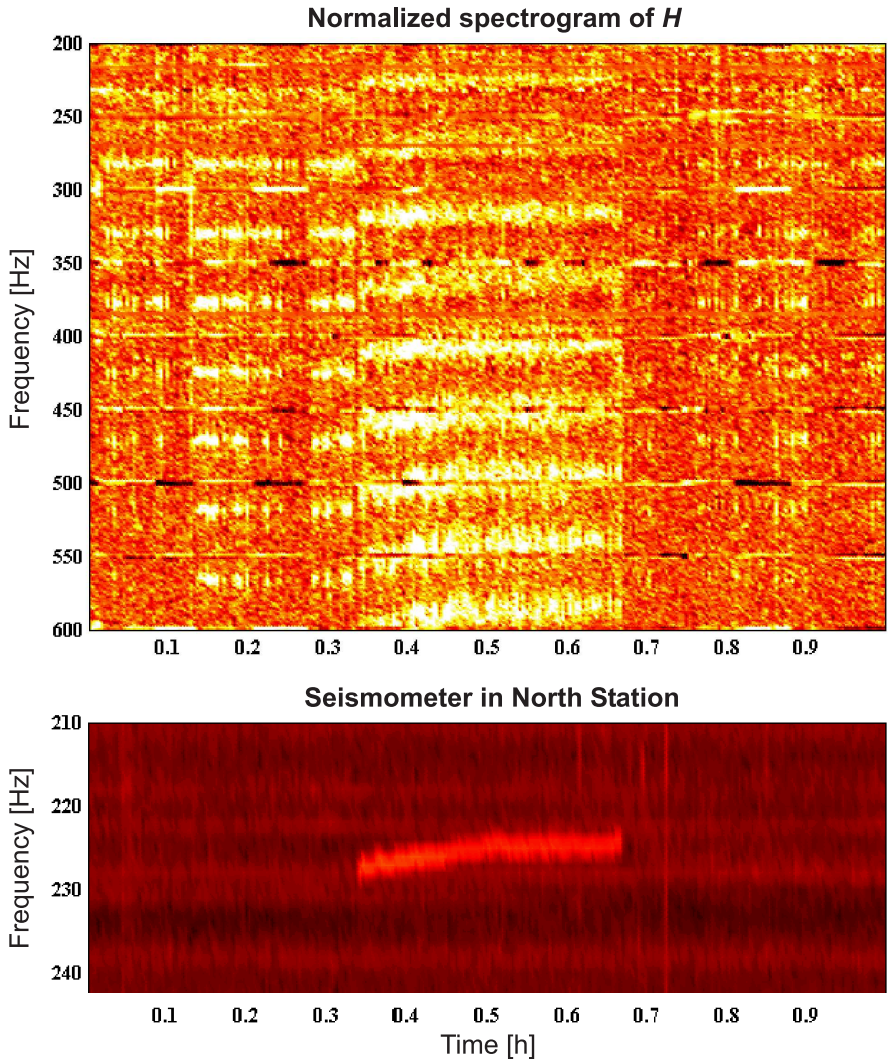


Figure 3.14.: Normalized spectrograms of H and the seismometer in the north end station. Orange pixels indicate the average value, black pixels show values below the average and white pixels indicate values above the average, i.e. excess noise. In H we find two different sets of ripple noise. The first class are weak ripples with frequencies corresponding to n*47 Hz, present for example at hour 0.2, 0.3 and 0.9. The second class of ripple noises are the strong n*45 Hz ripples, present in the period from hour 0.35 to 0.65. The frequency of this class changes slightly between hour 0.35 and hour 0.5. The same drift is observed in the fifth harmonic of 45 Hz in the seismometer signal from the north station. This gives a very strong indication for the n*45 Hz ripples be generated by seismically driven stray light generation in the north end station.

MFn is clipped at the too small view port.

In order to eliminate the n*45 Hz ripples we plan in a first step to reduce the seismic excitation in the north station. If that does not give enough reduction of the ripples we plan in a second step to open the vacuum system and exchange the view port by a larger one.

So far I only described the class of n*45 Hz ripples. As can be seen in Figure 3.14 we also observed a second class of ripples with a fundamental frequency of 47.2 Hz. The origin of these ripples is still under investigation.

3.8. Summary

Stray light is one of the main noise sources in gravitational wave detectors. Tiny amounts of phase shifted stray light can cause sensitivity limiting excess noise. A well known class of stray light noise is the scattering shoulder which was observed in all currently operating gravitational wave detectors (see Figure 3.1). The stray light problem might even become more important in the next generation of gravitational wave detectors operating with up to 1000 times larger light powers.

In order to widen the understanding of stray light generation and propagation through the interferometers, in this work a device was developed able to generate stray light with a controllable phase shift. With the aid of this device stray light injections have been carried out and it was proven that the stray light behavior and transfer can be explained in a purely linear coupling scheme.

Furthermore in Section 3.3 a rough guide is given which indicators point to stray light problems. A powerful tool to identify the origin of the scattered light in auxiliary beam paths, the so called filter experiment, was presented in Section 3.4.

In order to avoid stray light problems some simple rules have to be observed:

- Avoid the presence of a beam waist in auxiliary beam paths in air.
- If you cannot avoid the presence of a beam waist, do not place a optical component close to it.
- Only use low scatter optics, preferably with sub-angstrom surface finish and ion-beam-sputtered dielectric coatings.

- Always properly dump any secondary beams.
- Reduce movement of optical components by using stiffer mounts.
- Less seismic and acoustic excitation helps.

After two years of nearly continuous work on stray light problems in GEO 600 we now have a good understanding of scattering and know which aspects have to be taken into account for the design of next generation detectors.

Chapter 4.

Photon pressure calibration

4.1. Introduction

The production of a precisely calibrated output stream of a gravitational wave detector is one of the most challenging detector characterization tasks. An accurate calibration is essential for any multi-detector analysis. A lot of research effort is currently focussing on the use of null-stream construction [Wen05]. Null-streams combine the outputs of multiple detectors in order to produce a data stream containing no GW information, which can be used for veto analysis. However, it is clear that if the strain signals from the individual detectors are subject to systematic calibration errors then the resulting null-stream will not be truly null and could contain residual traces from any detected GW signal. Another class of multi-detector analysis, requiring high calibration accuracy are coherent analysis (for example, when searching for pulsar signals [Dupuis05]) where phase inaccuracies could potentially destroy a signal.

For the official calibration of the GEO 600 detector, we use an on-line time-domain technique to produce two calibrated signals, each potentially containing gravitational wave information, one for each output quadrature, P and Q [Hewitson04]. These two calibrated data streams are then optimally combined, using a maximum likelyhood method, in order to give a single signal with the best snr at all frequencies, H [Hewitson05].

The calibration of a gravitational wave detector can in general be split into two parts. The *absolute* calibration is given by a single overall factor, while the *relative* calibration represents a frequency dependent contribution. The absolute calibration of GEO 600 is a rather complicated procedure. In order to obtain a calibration for the ESD, the main actuator for the differential arm length of the interferometer, the ESD is first calibrated

back to the common mode error point of the frequency stabilization loop, which is then calibrated back through the mode cleaners to the frequency of the master laser, which is finally calibrated by using the well known free spectral range of the first mode-cleaner. Besides accumulating errors due to the many steps involved, this procedure implicates some drawbacks: some of the measurements can only be performed with low SNR and others are limited to be carried out in a certain frequency range. Furthermore, the procedure, which takes about on day on the whole, needs to be done partly with an unlocked interferometer, which means that the detector is not taking science data for the duration of the measurements.

Given the facts stated above, it is highly desirable to find new methods for providing an independent calibration of the gravitational wave detector. One approach is to use intentionally induced gravity gradients to apply a well known force to one of the test masses [Mantone06], [Hild04]. One proposal for such a "gravity calibrator" employing 3 masses rotating in a commercial spin dryer can be found in Appendix F. Another method is to use the photon pressure of a power modulated laser beam in order to apply a force to the test mass [Weiland04], [Mossavi06]. The mayor advantage of this method is that the displacement of the test mass, $x(\omega)$, is given in the ideal case by the simple relation:

$$x(\omega) = \frac{2 \cdot P}{M \cdot c \cdot \omega^2} \tag{4.1}$$

where P is the modulated light power, M is the mass of the test mass, c represents the speed of light and ω is the modulation frequency of the calibration signal. All values going into Equation 4.1 are, in principle easily measurable to a high precision. Figure 4.1 shows the corresponding (idealized) pendulum response of the lowest stage of a GEO main suspension. For frequencies much larger than the resonance frequency the magnitude of the response is analogous to Equation 4.1 proportional to $1/f^2$, and the pendulum follows the excitation with a phase shift of ϕ = -180 deg.

4.2. Experimental setup of the photon pressure calibrator

An optical layout of the north end station of GEO 600 where the photon pressure calibrator is currently installed is shown in Figure 4.2. The beams from the main interferometer are drawn in red color. A small fraction of the light hitting MFn from the south is transmitted through the test mass. It then leaves the vacuum system via a view port,

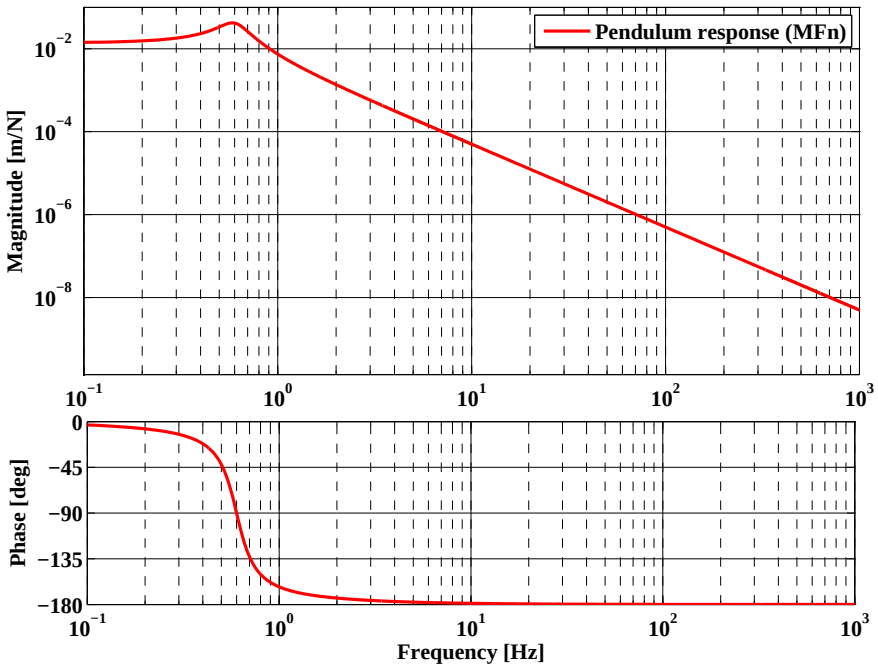


Figure 4.1.: Response function of the lowest stage of a GEO 600 main pendulum. For frequencies much larger than the resonance frequency the pendulum follows with a phase shift of ϕ = -180 deg.

is focused down by a lens (LNB) and is finally detected on the north bench by a spot position detector (PSDNB) and a CCD (CCDNB), both used for alignment control. As this auxiliary beam is essential for the alignment control of the main interferometer we cannot allow for any disturbance in its beam path, which sets some restrictions for the setup of the photon pressure calibrator.

The GEO photon pressure calibrator uses a fiber-coupled laser-diode with a wavelength of 1035 nm as a light source. After the light leaves the fiber it is collimated by a lens (collimator) of 60 mm focal length. The converging beam enters the vacuum via the view port and has a waist of about 5 mm diameter at the front surface of MFn. The PCAL beam (green in Figure 4.2) first passes the AR coating at the back surface of MFn and travels through the test mass substrate, made of fused silica, before it gets reflected at the HR coating. Then the PCAL beam passes the substrate and the AR coating a second time, leaves the vacuum system again via the view port and is finally dumped at the north bench. Some problems are connected to the setup described above. The beam profile provided by the laser diode is not a TEM_{00} and quickly diverges after being reflected at the test mass. Therefore, clipping of the outgoing beam at the small view port (radius = 4 cm) cannot be ruled out, thus a reliable measurement of the reflected light power is not accessible. Furthermore, the PCAL beam does not pass the view port and the mirror under normal incidence. Even though the angle is only about 2 degrees, effects depending on the polarization of the light may occur.

4.2.1. Measurement of the modulated light power

As described in the previous section the main disadvantage of the actual photon pressure calibrator setup is the fact that the light power reflected from the test mass is not accessible via any easy measurement. Therefore the light power is measured by an internal monitor photodiode connected to the laser diode. This signal is recorded in the DAQS and will be referred to as PCALmon in the following. PCALmon was calibrated at DC using a calibrated power meter. The measurement of the light power is then propagated through the collimator lens ($T_{\mathrm{lens}} = 0.9988$) and the view port ($T_{\mathrm{vp}} = 0.9986$)[1]. The reflection of the AR coating and the transmission of the HR coating can be neglected since both are, according to their specifications, clearly below a tenth of a

[1] As the transmission of the currently used view port cannot be measured in situ, the given value was measured at a spare view port with identically specified properties.

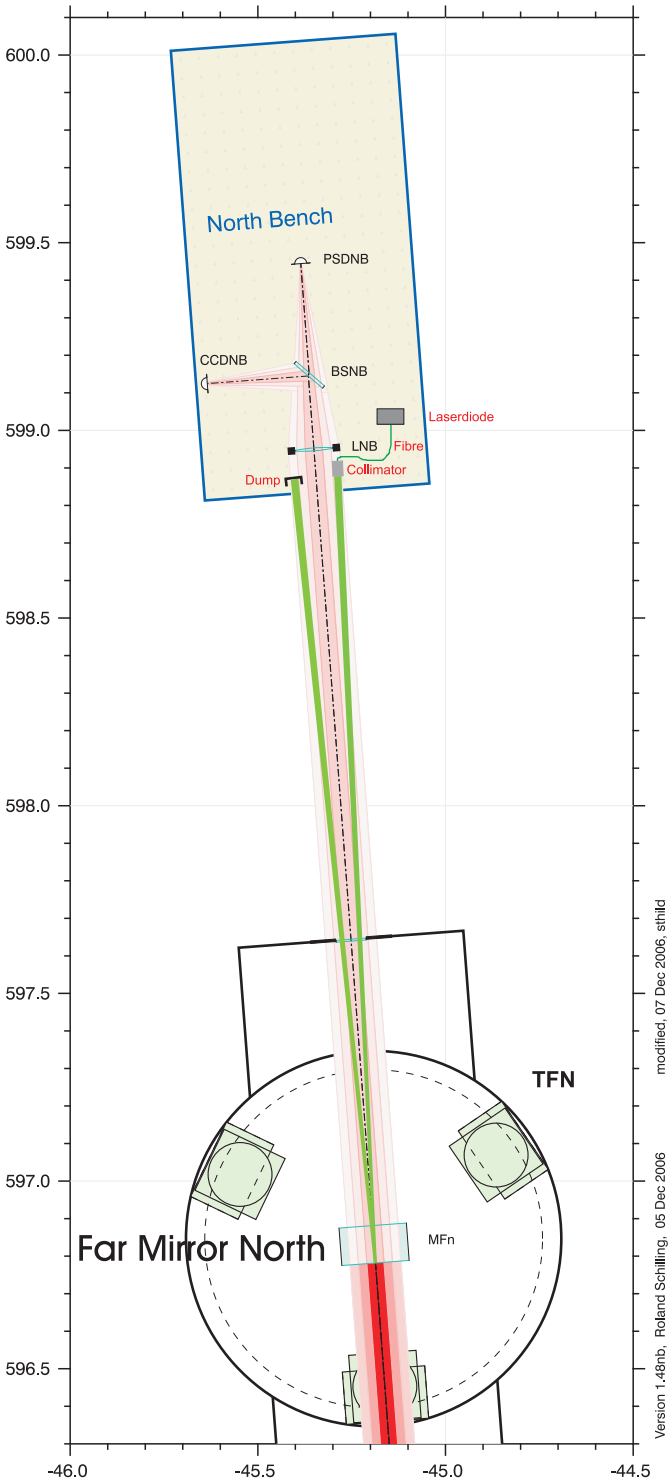


Figure 4.2.: OPTOCAD layout of the optics in the north end station of the GEO 600 detector, including the setup of the photon pressure calibrator.

percent. Again a direct in situ measurement of these two coatings is not possible.

However, a reliable measurement of the reflected DC light using a large lens (LNB on Figure 4.2) to collect all the light leaving the vacuum, showed that only 77 % of the light power entering the vacuum are leaving it afterwards. This strong power loss can only be explain by absorbtion due to dirty optics. Again an in situ measurement of the losses is not possible without opening the vacuum system. Venting the vacuum system was not possible due to the high risks involved, such as breaking a fused silica fibre of the quasi-monolith mirror suspension.

The unexplained loss of 23 % is the largest limitation of absolute accuracy of this setup. Depending on the position of the losses, any calibration derived from this photon pressure calibrator may be wrong by 23 % !! For all investigations in this chapter we will assume symmetrically distributed losses, meaning that 11.5 % of the light power are lost in front of the HR coating and 11.5 % are lost after the HR coating.

Altogether, the effective light power is estimated in the followin way

$$P[\mathrm{W}] = \mathrm{PCALmon}[\mathrm{V}] \cdot T_{\mathrm{lens}} \cdot T_{\mathrm{vp}} \cdot K_{\mathrm{pm}} \cdot (0.885 \pm 0.115) \tag{4.2}$$

where K_{pm} includes the calibration of the power meter and the calibration from PCALmon to the power meter.

4.3. Injections with different amplitudes

Injections using the photon pressure calibrator have been performed with different amplitudes in order to check that the response of the system is linear with respect to the amplitude of the modulated light power. A sine wave of 134 Hz has been injected with four different amplitudes, between 50 and 500 mW.

Figure 4.3 shows the analysis of this measurement. In the upper figure the magnitude of the transfer coefficient from PCALmon to H is plotted over the amplitude of the modulated light. The mean is indicated by red dashed line. The lower figure shows the deviation of each measurement point from the mean. We observe that the response function of the photon pressure calibrator is linear within ±4 % for different amplitudes of the modulated light power.

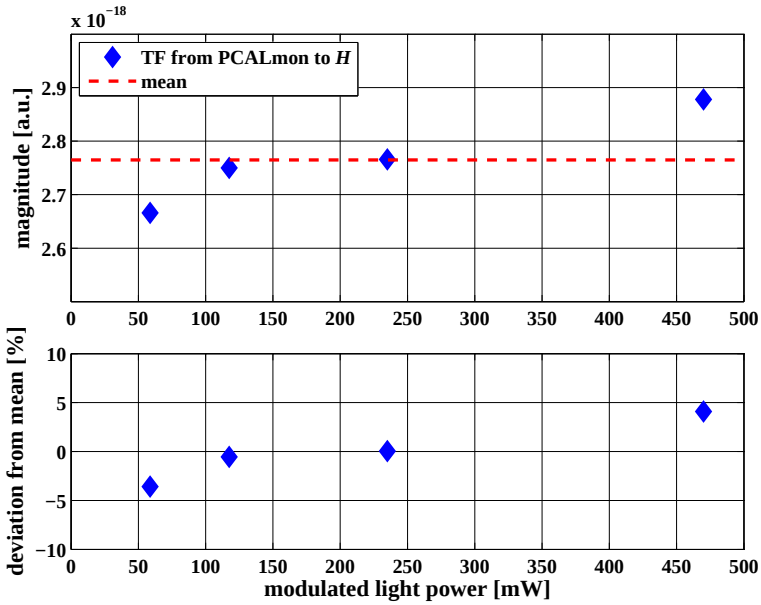


Figure 4.3.: Measurement of the transfer coefficient from PCALmon to H for various amplitudes of the modulated light. The upper plot shows the magnitude of the transfer coefficient (blue diamonds) and the mean (red dashed line). In the lower plot the deviation of each measurement from the mean is plotted. All measurements agree within $\pm 4\,\%$ with the mean.

4.4. Injections with different frequencies

One of the main goals of the experiments using the photon pressure calibrator is to check the relative accuracy (frequency dependence) of the official calibration. Therefore injections over the full detection band have been performed. Due to the strongly decreasing response ($1/f^2$) of the system it is not possible to get identical SNR at all frequencies. Especially towards high frequencies only small SNR could be achieved and the uncertainties of the measured value originating due the noise floor needs to be taken into account. Analogous to measuring the amplitude of an ESD line, described in [Koetter03], the amplitude of a photon pressure calibrator injection, A_{pcal}, can be approximated by

$$A_{\mathrm{pcal}} = \sqrt{V_{\mathrm{pcal}}^2 - \eta^2}, \tag{4.3}$$

where V_{pcal}^2 is the power measured in the frequency bin containing the injection and η^2 represents the averaged power of noise floor in the frequency bins around the injection.

Figure 4.4 shows the result of photon pressure calibrator injections at various frequencies. Each point is derived from a measurement that uses 2 minutes of data for computing a single FFT.[2] The blue diamonds are the amplitudes of the injections, A_{pcal}, as seen from the GW channel, which represent the official calibration. The red circles show the amplitude of the PCAL injection predicted from PCALmon, $A_{\mathrm{pcal}}^{\mathrm{theo}}$, using Equation 4.1 and represent the photon pressure calibration. Above 3 kHz the measurements are not accurate due to too low SNR as indicated by the measurement of the noise floor (green stars). At all frequencies the injections show up smaller in the official calibration than expected from the photon pressure calibration. Furthermore the discrepancy between the two calibration methods is observed to vary with frequency. Especially towards higher frequencies the deviation seems to increase strongly. The ratio of the calibration derived from the photon pressure calibrator and the official calibration is plotted over frequency in the upper subplot of Figure 4.5. The red circles are calculated assuming the observed power loss of 23 % (see Section 4.2.1) is caused half on the way in and half on the way out of the vacuum system. The two blue dashed lines indicate the extremes, i.e. all 23 % lost on the way in and all 23 % lost on the way out. The observed

[2] In order to get best SNR in this measurement we have to choose a high frequency resolution. The frequency bin to be analysed is determined automatically by the used MATLAB script by finding the frequency bin of maximum power in the amplitude spectrum of PCALmon. The full MATLAB code used for these investigations can be found in Appendix E.

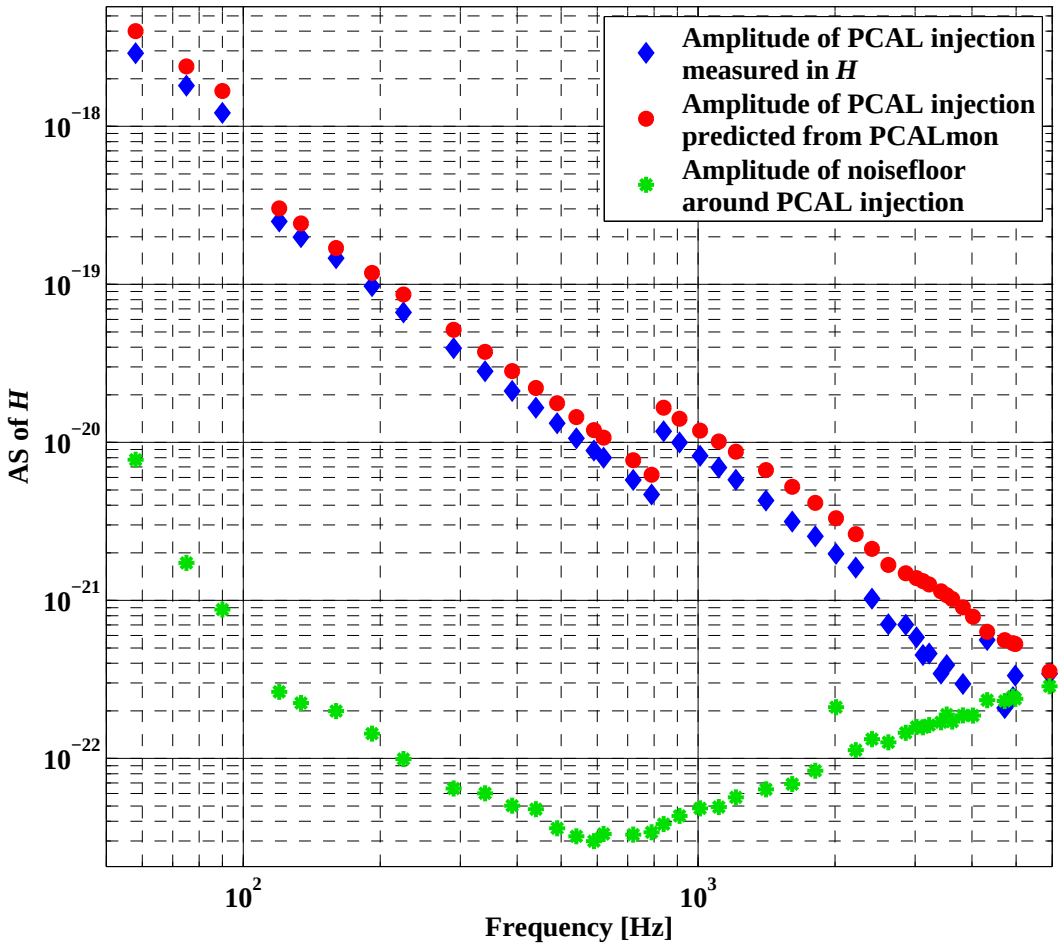


Figure 4.4.: Photon pressure calibrator injections for various frequencies. The blue diamonds represent the amplitude of the injections, A_{pcal}, already corrected for the noise floor using Equation 4.3. The red circles show the amplitude of the PCAL injection predicted from PCALmon, $A_{\mathrm{pcal}}^{\mathrm{theo}}$, using Equation 4.1. A discrepancy between the official calibration (blue diamonds) and the calibration derived from PCAL (red circles) is observed. Above 1 kHz the discrepancy seems to quickly increase. For frequencies higher than 3 kHz the measurements are not accurate due to too low SNR as indicated by the measurement of the noise floor (green stars). The injections in the frequency band from 100 to 800 Hz were performed using a smaller amplitude.

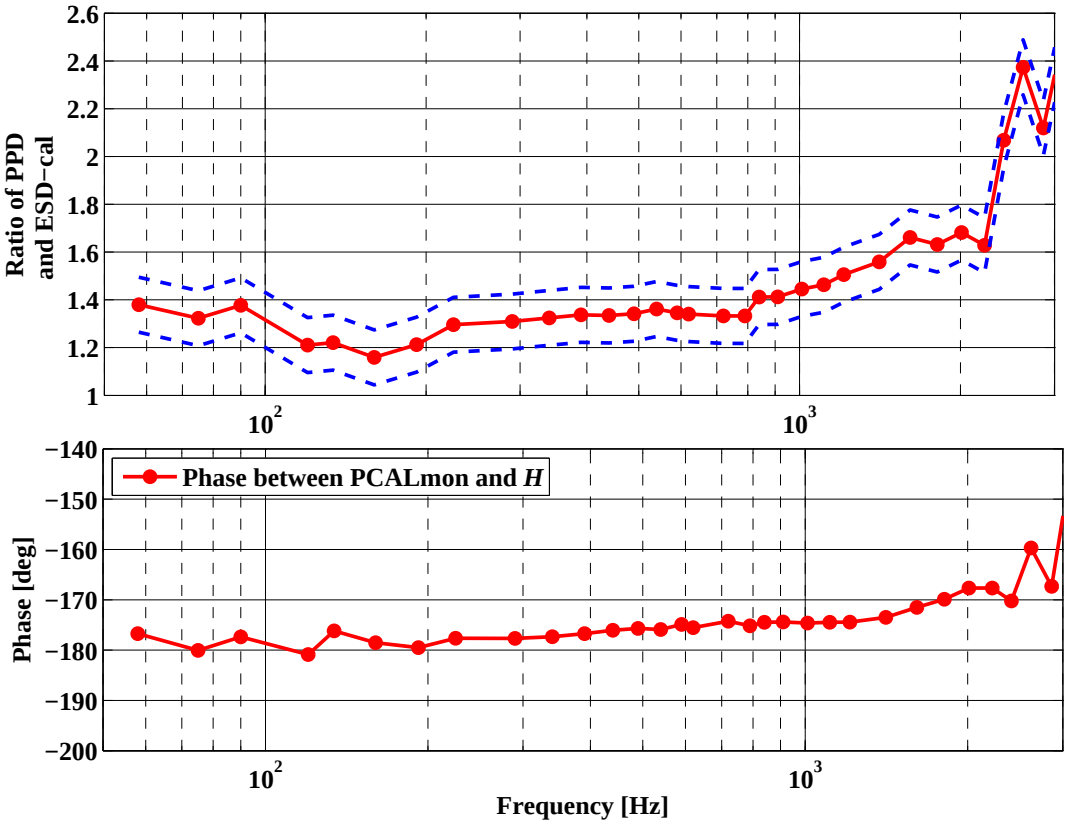


Figure 4.5.: *Upper plot*: Ratio of the photon pressure calibration and the official calibration versus frequency. The blue dashed lines indicate an uncertainty of ±11.5 %, caused by the observed power loss of 23 % (see Section 4.2.1). The observed discrepancy between the two calibration methods is on average about 20 to 40 % for frequencies below 1 kHz. Above 1 kHz the discrepancy quickly increases to about 120 % at 3 kHz. *Lower plot*: Phase relation between injected light power (PCALmon) and H. For all frequencies below 1 kHz the phase is -180 degrees within 5 degrees. Above 1 kHz the phase starts slowly to go off and reaches about -160 degrees at 3 kHz.

discrepancy between the two calibration methods is on average about 20 to 40 % for frequencies below 1 kHz. Above 1 kHz the discrepancy quickly increases to about 120 % at 3 kHz. The lower subplot of Figure 4.5 shows the phase relation between injected light monitored by PCALmon and H at the frequency of the injection. As we would expect from the pendulum response the phase is about -180 degrees. For all frequencies below 1 kHz the phase is -180 degrees within 5 degrees. Above 1 kHz the phase starts slowly to go off and reaches about -160 degrees at 3 kHz.

The discrepancy between the two calibration methods is surprisingly large, especially considering that the official calibration is, as we will see in Section 4.5, strongly believed to be correct. The observed deviation can roughly be separated into two phenomena: First of all there seems to be an absolute discrepancy of 20 to 40 % showing up equally at all frequencies. Secondly, at high frequencies the relative deviation increases strongly; the response of the photon pressure calibrator does not follow the $1/f^2$ expected from the pendulum response.

One effect that could explain the absolute mismatch is photon pressure calibrator induced test mass misalignment, which will be described in detail in Section 4.8. To find a mechanism that can explain the frequency dependent discrepancy seems to be harder. Any kind of potential beam clipping should occur frequency independent. Also the above mentioned photon pressure calibrator induced test mass misalignment or polarization related problems can only explain a frequency independent error. As will be shown in Section 4.7 the behavior of the photon pressure calibrator at high frequencies might be explainable by photon-pressure-induced non-rigidity of the test mass.

4.5. Validation of the official calibration

So far we compared the official calibration with the calibration derived from the photon pressure calibrator. The observed discrepancy can of course originate from either the official calibration being wrong, the photon pressure calibrator being wrong or both being wrong. There is no easy way to distinguish these three possibilities. However, the official calibration of GEO 600 was repeated and evaluated several times by several people over the last few years. Therefore the official calibration is strongly believed to be correct within 10 % in magnitude and 20 degrees in phase for frequencies between 50 and 2000 Hz. This statement includes the absolute calibration accuracy as well as the

relative calibration accuracy. Two validation measurements important for the context of this work are presented in the following.

4.5.1. Transfer function from ESD to H

The frequency dependent discrepancy between the official calibration observed in Section 4.4 could also be explained by a potential relative error of the official calibration. In order to rule out any relative inaccuracy in the official calibration we performed noise injections into the electro static drives (ESD) and measured the transfer function from the input of the ESD to H.

The electro static drives are believed to have a flat frequency response for all frequencies of interest. We would therefore expect to find the transfer function from the ESD to H to follow a $1/f^2$-law, originating from the pure pendulum response. The correctness of the ESD model was checked by white-noise injections in the power-recycled Michelson mode, where the optical response is known to be flat.

Figure 4.6 shows the measurement of the corresponding transfer function for frequencies between 50 and 2000 Hz. The signal was injected at the last OP of the electronics controlling the Michelson differential arm length. Then the transfer function from the feedback monitor of high voltage amplifiers (HVA) to H was computed and corrected for the responses of the dewhitening filters installed behind the high voltage amplifiers. As the upper subplot of Figure 4.6 indicates the transfer function is very close to an $1/f^2$-behavior. In center plot Figure 4.6 the deviation from an $1/f^2$-fit is shown. In the lower plot of Figure 4.6 the phase of the transfer function is plotted versus frequency.

Overall we can state that this measurement confirms a relative calibration accuracy of $\pm 10\,\%$ in magnitude and ± 5 degree in phase is achieved for frequencies between 50 and 2000 Hz.

4.5.2. ESD injections up to high frequencies

As already stated at the beginning of this chapter the main purpose of a calibration based on the photon pressure calibrator is to get an absolute and relative comparison to the official calibration. To get a measure to compare the photon pressure calibration to we have to make sure that we compare it to a meaningful measure. Therefore we have

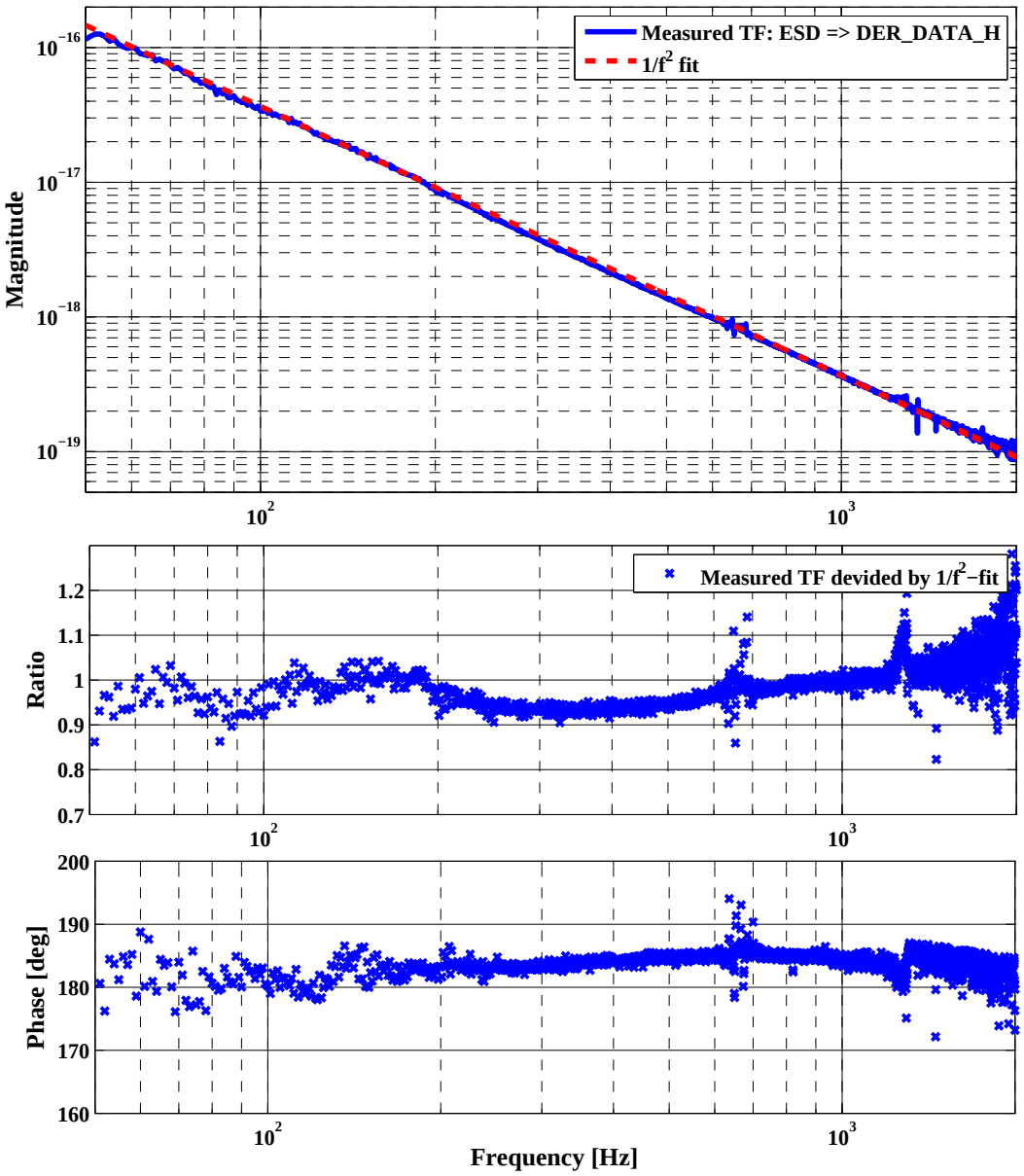


Figure 4.6.: Measurement of the transfer function from the input of the ESD to the calibrated output of GEO 600, H. *Upper plot*: Measurement of the transfer function (solid, blue line) and a corresponding $1/f^2$-fit (dashed, red line). Above 1.5 kHz the accuracy of the measurement goes slightly down due to low SNR of the measurement. *Center plot*: Ratio of the measured transfer function and the $1/f^2$-fit. *Lower plot*: Phase between the injected signal and H. Overall a relative calibration accuracy of $\pm 10\,\%$ in magnitude and ± 5 degree in phase is achieved for frequencies between 50 and 2000 Hz.

to validate the official calibration to the best of our knowledge before we can use it as an 'absolute standard'.

Most of the astrophysical analysis using data from GEO 600 consider only frequencies up to 2 kHz. That is why so far the calibration was examined only up to this frequency. However, some of the measurements presented in this chapter are performed at frequencies as high as 6 kHz. Therefore we have to make sure that a reasonable relative calibration accuracy is given also at higher frequency. The method we can use for this investigation is similar to the one described in Section 4.5.1, but at high frequencies we have to inject single discrete lines instead of broadband noise in order to get a sufficient SNR in the measurement.

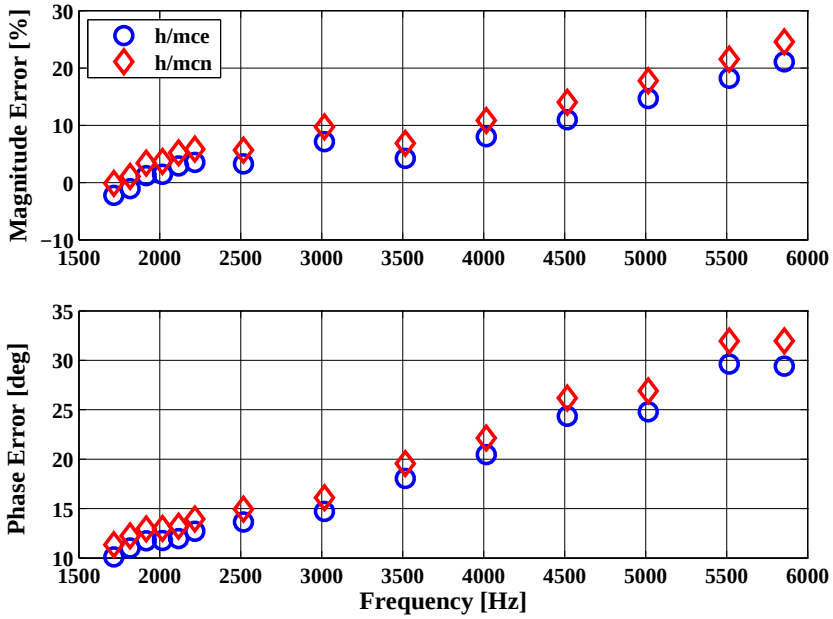


Figure 4.7.: Measurement of the relative calibration accuracy of the official calibration for high frequencies. Plotted are the differences between the official calibration (H) and the propagation of the injections through the loop model for two individual ESD at MCe and MCn. The analysis of this injections was provided by J.R.Smith.

Figure 4.7 shows the analysis of the high frequency ESD injections. The magnitude and phase of the injections measured in the officially calibrated data are compared to the expected values from propagating the injections through the loop model for each for the

two ESD, MCe and MCn. Compared to the measurements at 2 kHz for frequencies up to 6 kHz the relative error in the official calibration increases by 20 % in magnitude and 20 degrees in phase.

The relative errors of the official calibration are in summary:

- For frequencies below 2 kHz: 10 % in magnitude and 5 degree in phase.
- For frequencies between 2 and 6 kHz: 20 % in magnitude and 30 degree in phase.

4.6. The phase picture

The observed discrepancy of the official calibration and the photon pressure calibration can only be explained in the presence of an additional unknown effect, I will call "mystery effect" in this section, that adds to the pendulum response. As I will show in this section

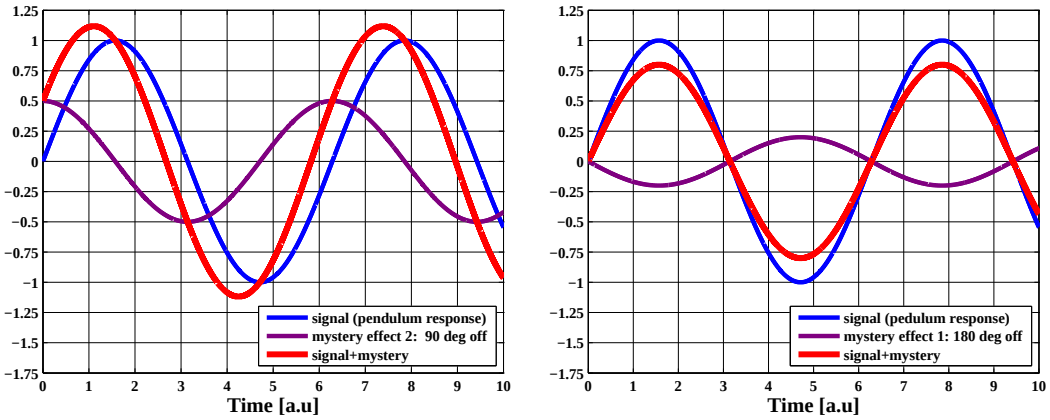


Figure 4.8.: The observed discrepancy of the official calibration and the photon pressure calibration might be explainable by an additional "mystery" effects adding to the pendulum response. *Left plot*: If the mystery effect is 90 deg off phase with the pendulum response the overall response would be increased (and the phase shifted). Therefore such n "mystery" effect cannot explain the observations of smaller overall response (and a constant phase of -180 degree). *Right plot*: Only a mystery effect 180 degrees off phase from the pendulum response can explain the observation of a reduced overall response (and a constant phase of -180 degrees).

the measurements shown in Figure 4.5 already contain very strong boundaries for the

behavior of the unknown "mystery" effect. In order to explain the absolute and the relative discrepancy of the two calibration methods, the "mystery" effect has to add to the pendulum response in a way that, first of all the overall response gets smaller in amplitude, and secondly the phase of the overall response still stays at -180 degrees. As indicated by Figure 4.8 these two conditions can be satisfied by an effect that is 180 degrees out of phase from the pendulum response, or in other words, an effect that is in phase with the modulated light of the photon pressure calibrator.

4.7. Non-rigidity of the test mass

The widely held assumption with in the GW-community so far has been that a test mass acts like a rigid body for frequencies below the first internal resonance. However, no quotable reference could be found that documents or proves this hypothesis.

In fact, in this section I will present some investigations disproving the theory of test masses being rigid bodies at low frequencies.

The idea of this traditional model is that a mirror for all frequencies below its first internal mode, which is for the GEO test masses around 11 kHz [Smith04], behaves equivalent to any other oscillator:

- It just follows the excitation without any phase lag (phase difference = 0 degree)
- The response is flat in frequency and equals the response at a frequency of 0 Hz (DC).

However, obviously at DC we can (to a tiny amount) compress any piece of glass by pushing from both sides, or stretch it by pulling on both ends. In a linear setup the amount of the compression (or elongation) ΔL only depends on the Young's module E_{young} of the material and the geometry of the test body

$$\Delta L = \frac{1}{E_{\text{young}}} \frac{L \cdot F}{A}, \tag{4.4}$$

where L is the length and A the cross section of the test body and F is the applied force.

The beam of the photon pressure calibrator hits the high reflective coating of the test mass and pushes at the surface of the mirror. Due to the inertia of the test mass the

force applied to the surface causes a deformation of the test mass. In order to find out whether this effect can be responsible for the strong discrepancy between the official and photon pressure calibration observed at frequencies around a few kHz, the main question we have to answer first is if the effect from force-induced mirror-deformation is roughly of the same size as the center of mass movement of the test mass induced by the photon pressure calibrator. In Section 4.7.1 a rough estimation is given using a very simple model, while in Section 4.7.2 a finite element analysis of the problem is presented.

4.7.1. Rough estimation of photon pressure calibrator induced test mass deformation

As we first just want to get an estimation of the order of magnitude of any effect related to photon pressure calibrator-induced mirror-deformation, we can use a simplified model using the following assumptions:

- The beam of the photon pressure calibrator has a flattop profile and illuminates the full mirror.
- The backplane of the test mass is fixed in position.

In this case we can simply apply Equation 4.7. With the parameters of the GEO test masses $L = 0.1\,\mathrm{m}$ and $A = \pi \cdot 90^2\,\mathrm{mm}^2$, Young's module for Suprasil of $E_{\mathrm{young}} = 7 \cdot 10^4 \mathrm{N/mm}^2$ [Heraeus] and a photon pressure induced force $F = 2.25 \cdot 10^{-9}\,\mathrm{N}$ (750 mW light power) we obtain a homogenous elongation of the test mass of:

$$\Delta L = 1.26 \cdot 10^{-19}\mathrm{m} \tag{4.5}$$

This corresponds to an apparent strain H_{def} of

$$H_{\mathrm{def}} = 2.1 \cdot 10^{-22}. \tag{4.6}$$

If we now compare this result with the measurement shown in Figure 4.4 we find that for high frequencies (a few kHz) this effect might have some non negligible influence. In conclusion that means that we have to evaluate this effect with a more realistic model, which is presented in the next section.

4.7.2. Finite element analysis of photon pressure calibrator induced test mass deformation

The finite element method (FEM) is a numerical technique to simulate complex mechanical and thermal setups. This type of analysis is widely used in engineering and has proven its reliability. One of the most powerful tools for finite element analysis (FEA) is the software ANSYS [Ansys], which was used for the analysis described below. The FEM simulation presented here was done by our colleagues from Glasgow: I.Martin, S.Reid and J.Hough.

The test mass (diameter = 18 cm, thickness = 10 cm, bulk material = suprasil) is represented by 87000 discrete nodes. A total force of 2.77 N (diameter = 5 mm, gaussian profile) is applied to the center of the front surface of the test mass.[3]

Figure 4.9 shows the setup and the result of the analysis. In the simulation the force is applied from the outside to the front surface of the test mass while in the photon pressure calibrator setup the test mass surface is pushed from inside. Since the deformation depends on the intermolecular forces, which can, for small deformations, be assumed to be linear, the simulated setup is equivalent to the photon pressure calibrator setup. The lower subplot of Figure 4.9 shows the resulting displacement in z-direction (along the axis of the main interferometer beam). The simulation and the corresponding result are rotationally symmetrical. The area of the front surfaces, illuminated by the photon pressure calibrator beam, is locally squeezed. The maximum displacement of $5 \cdot 10^{-9}$ m occurs in the center of the front surface.

4.7.3. Effective test mass displacement

The displacement measured by the interferometer is determined by the overlap of the photon pressure calibrator beam and the main interferometer beam, which can be described by a gaussian beam. The radial intensity, $I(r)$ is given by

$$I(r) = exp\left(\frac{-2r^2}{\omega^2}\right) \tag{4.7}$$

where ω is the radius of the beam. At the north end mirror the radius of the main interferometer beam is about 2.4 cm. Each point of the mirror surfaces contribution

[3] The test mass is held in place using a inertial relief function.

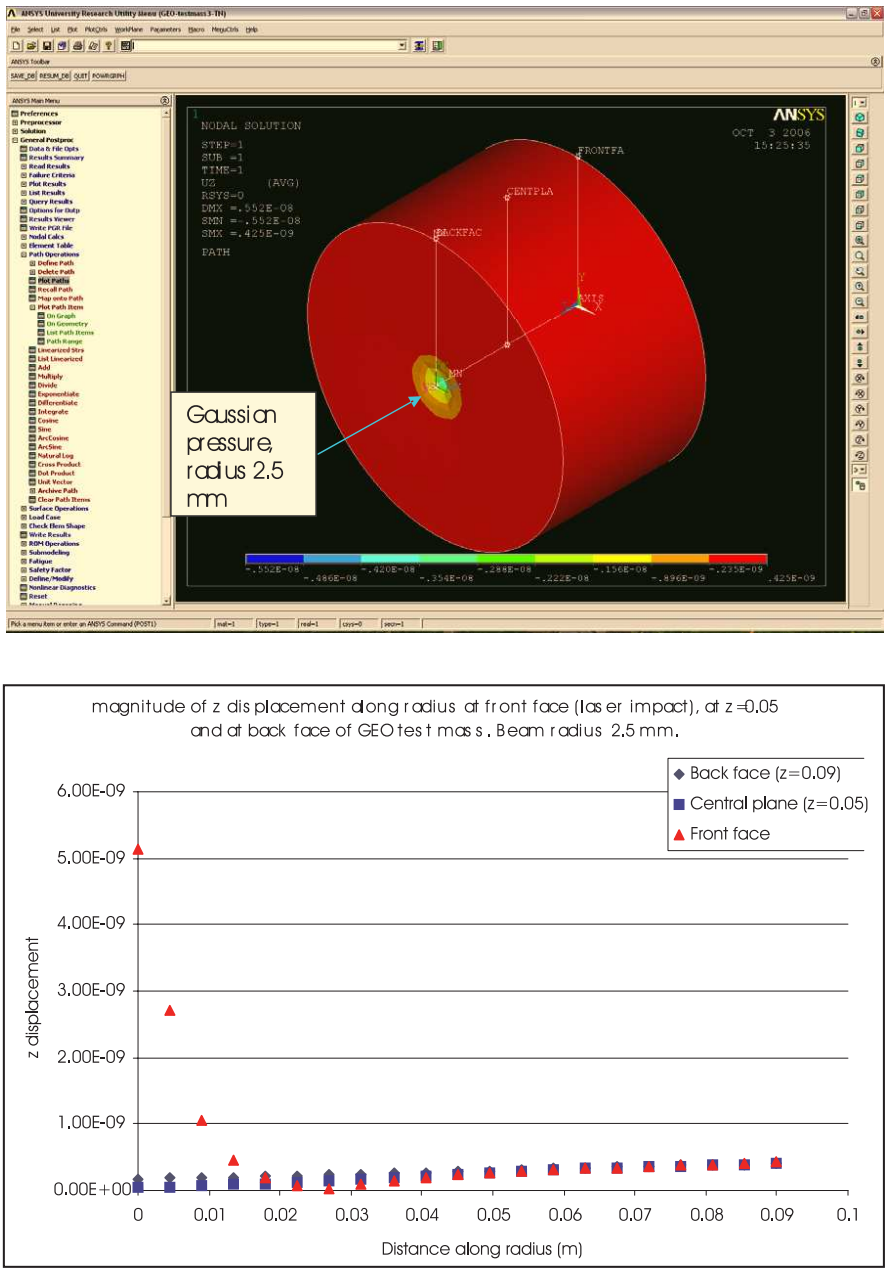


Figure 4.9.: *Upper plot*: Finite element model (FEM) of the deformation of a GEO test mass caused by the beam of the photon pressure calibrator. A total force of 2.77 N (diameter = 5 mm, gaussian profile) is applied to the center of the front surface of the test mass. The test mass is held in place using a inertial relief function. The analysis was done using the ANSYS software [Ansys]. The mirror was modelled by 87000 nodes. *Lower plot*: Result of this analysis. Plotted is the displacement over the radius of the test mass. The whole FEM analysis was provided by our colleagues from Glasgow: I.Martin, S.Reid and J.Hugh.

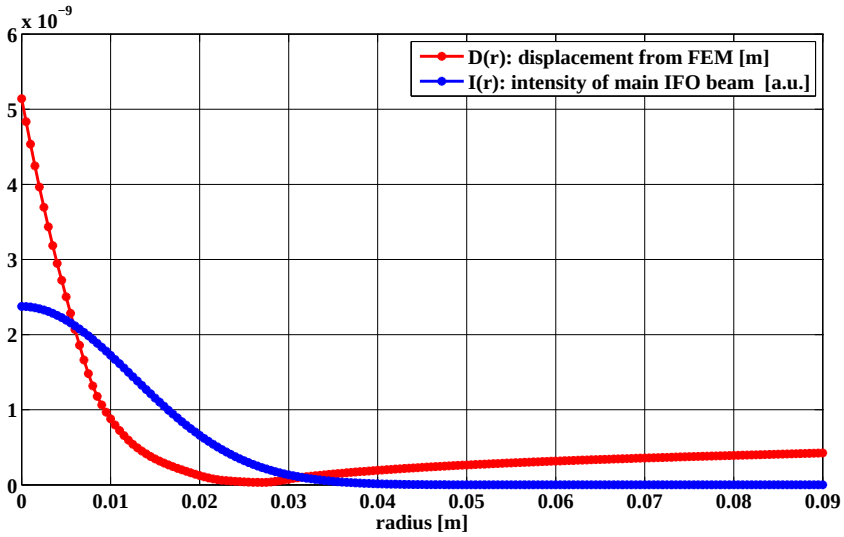


Figure 4.10.: Radial profile of the photon pressure calibrator induced test mass deformation, $D(r)$, and the intensity of the main interferometer beam, $I(r)$. The effective displacement of the test mass measured by the interferometer depends on the overlap of these two functions and is proportional to the $I(r) \cdot D(r)$.

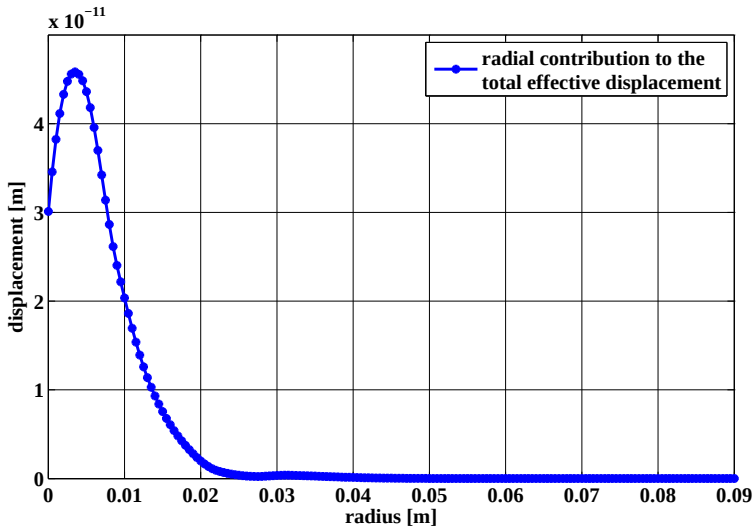


Figure 4.11.: Radial contribution to the effective displacement for photon pressure calibrator induced test mass deformation computed using Equation 4.8. The total effective displacement amounts to $7.57 \cdot 10^{-10}$ m.

to the total effective displacement, D_{total}, weighted by the product of the power of the main interferometer beam, $I(r)$, and the corresponding displacement $D(r)$. For a radially symmetric setup, as described above, the total effective displacement can be expressed by a single integral:

$$D_{\text{total}} = \int_0^{0.09\text{m}} \int_0^{2\pi} D_{\text{eff}} \cdot dr \cdot d\varphi = \int_0^{0.09\text{m}} 2\pi \cdot r \cdot k_I \cdot I(r) \cdot D(r) \cdot dr. \tag{4.8}$$

The factor k_I is a normalization factor for $I(r)$ in order to give

$$\int_0^{0.09\text{m}} 2\pi \cdot r \cdot k_I \cdot I(r) \cdot dr = 1. \tag{4.9}$$

Using Equation 4.8 the total effective displacement amounts to $7.57 \cdot 10^{-10}\,\text{m}$. Figure 4.11 shows the radial contribution to the effective displacement for photon pressure calibrator induced test mass deformation.

4.7.4. Comparison of pendulum response and mirror deformation

The last section showed that the test masses are not completely rigid and that the photon pressure calibrator beam really causes a non-negligible deformation. Next we have to evaluate how strong the displacement originating from the non-rigidity of the test mass is, compared to the displacement of the center of mass (originating from the pendulum response). Since both responses are linear for the applied light power of the photon pressure calibrator we can simply compare their responses.

The pendulum response, α_{pen}, follows a $1/f^2$-law for frequencies above 100 Hz, has a magnitude of $5 \cdot 10^{-7}\,\text{m/N}$ at 100 Hz and is 180 degrees out of phase with the power modulation (see Figure 4.1). The response of the mirror deformation, α_{def}, is assumed to be flat in frequency and in phase with the modulated light power for frequencies below the first internal resonances of the test mass. In Section 4.7.3 an effective displacement of $7.57 \cdot 10^{-10}\,\text{m}$ was found for an applied force of 2.77 N, which leads to a response of $2.73 \cdot 10^{-10}\,\text{m/N}$. The magnitude and phase of α_{pen} and α_{def} are shown in Figure 4.12, in blue (dashed) and red (dashed-dotted), respectively. The total response, α_{total}, plotted in green (solid) is the sum of the two individual responses:

$$\alpha_{\text{total}} = \alpha_{\text{pen}} + \alpha_{\text{def}}. \tag{4.10}$$

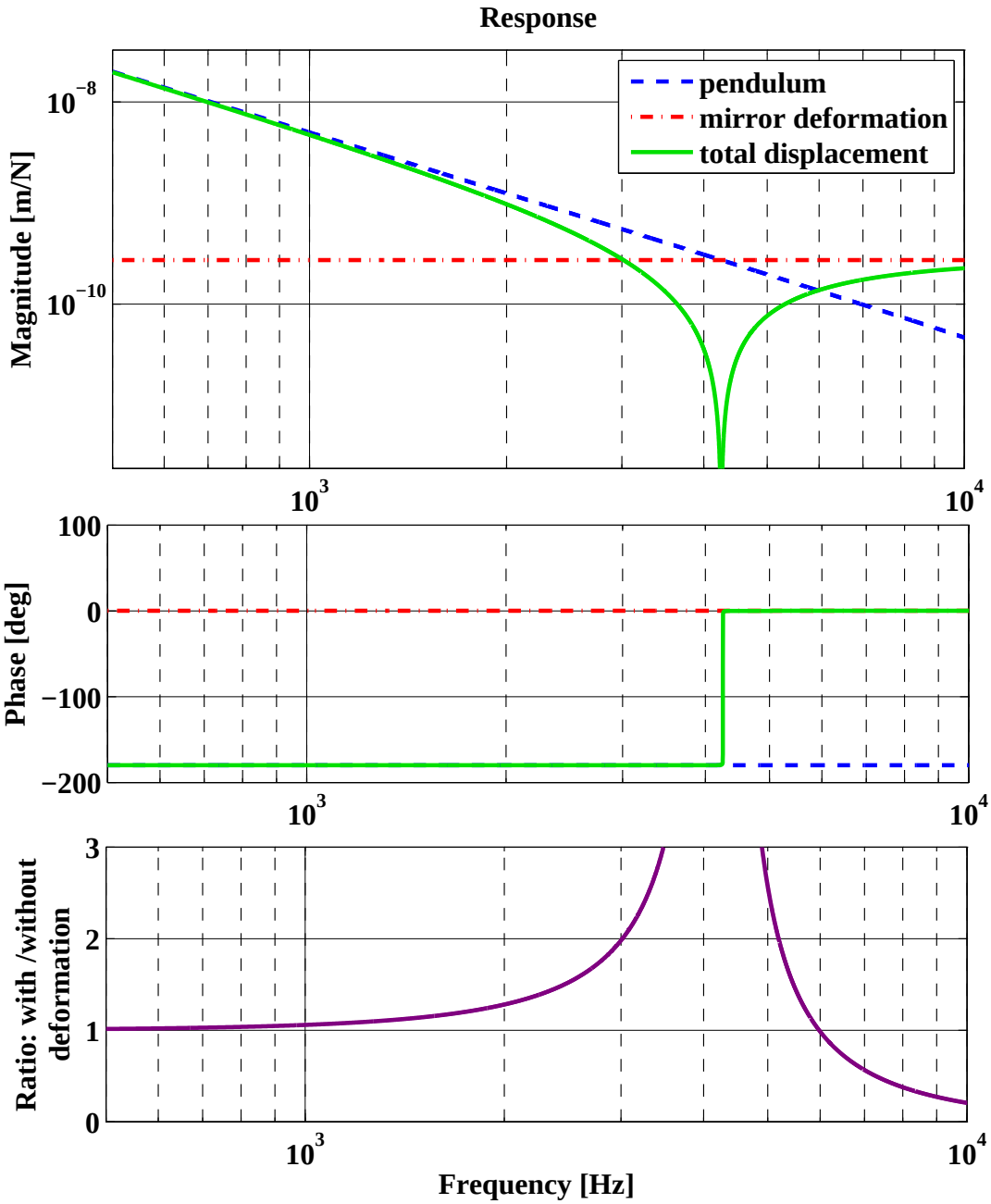


Figure 4.12.: Simple model for the photon pressure calibrator taking into account the responses from the pendulum and from the mirror deformation effect. The pendulum response follows a $1/f^2$-law and is 180 degrees out of phase from PCAL. The mirror deformation has a flat response and is in phase with PCAL. If both responses are added a notch appears at the frequency where both responses have equal size. The purple trace shows the expected discrepancy between the official calibration and photon pressure calibrator.

The total response shows a steep notch at the frequency where both individual responses have the same size and compensate each other completely due to having opposite phase. At the frequency of the notch the phase of α_{total} jumps from -180 to 0 degrees. The resulting discrepancy between the official calibration and the calibration derived from the photon pressure calibrator, including the effect from mirror deformation, is shown in lowest subplot of Figure 4.12.

In order to prove this model we have to do photon pressure calibrator injections up to higher frequencies. It will be difficult to find the notch in the magnitude of the response because at these frequencies the snr of any injected line is going the be very small, even for measurements of very long duration. However, the distinct change in phase should be measurable with reasonable effort.

4.7.5. High frequency injections using the photon pressure calibrator

Long duration injections using the photon pressure calibrator have been performed for frequencies in the range from 3 to 6 kHz. The snr of a periodic signal increases with the length of data used for a single FFT. This is due to the fact that by decreasing the width of the frequency bins, the noise contained in the signal-bin is decreased, while the signal stays constant. For the measurement presented in this section FFT containing between 2 and 10 hours of data were used. Such amounts of data[4] are difficult to handle with standard computers. This problem can be avoided by using a heterodyning technique. The time series of the data containing the signal of interest, $E_{\text{sig}} \cdot sin(\omega_{\text{sig}} t)$, is multiplied by a sine wave with a slightly lower frequency, ω_{het}:

$$E_{\text{sig}} \cdot sin(\omega_{\text{sig}} t) \cdot sin(\omega_{\text{het}} t) = \frac{1}{2} E_{\text{sig}} [cos(\omega_{\text{sig}} - \omega_{\text{het}})t - cos(\omega_{\text{sig}} + \omega_{\text{het}})t] \qquad (4.11)$$

The second term of the right hand side of Equation 4.11 still contains the signal, but shifted towards even higher frequencies. The signal component we are interested in is shifted to a very low frequency, $(\omega_{\text{sig}} - \omega_{\text{het}})$, which for our investigations was chosen to be 9 Hz. After heterodyning, the data stream is strongly low passed and down sampled to give a data stream that can be handled by desktop computers.

In order to compute the transfer function from PCALmon to H, both data streams are processed with the same heterodyning algorithm. Afterwards, the transfer function is calculated using the *tfe* function of MATLAB.

[4] 10 hours of data sampled with 16384 Hz consist of about 600 000 000 data points.

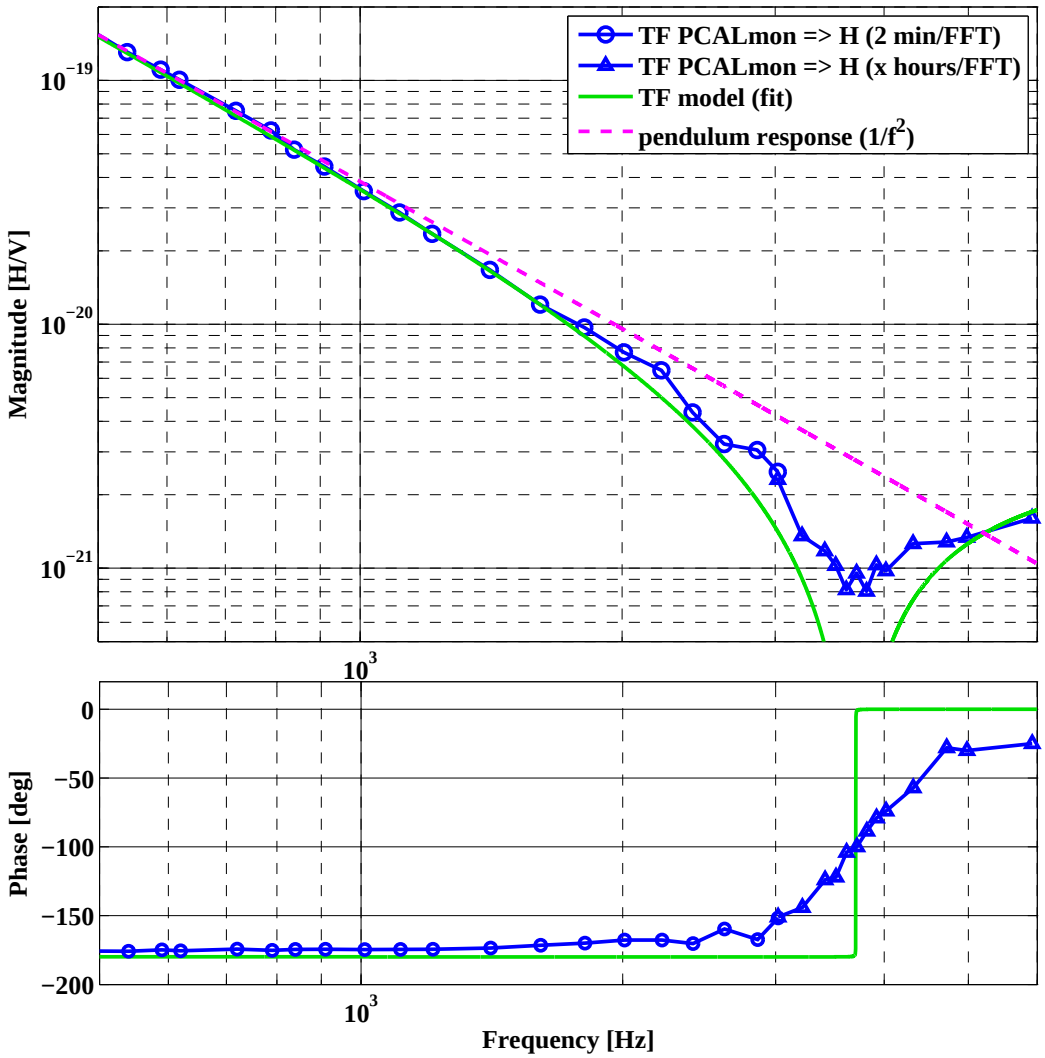


Figure 4.13.: Measurement of the photon pressure calibrator response at high frequencies. The measurements at frequencies below 3 kHz (blue circles) are the same as presented in Section 4.4. The measurement points above 3 kHz (blue triangles) are derived from long data stretches and contain up 10 hours of data in a single FFT. The green line represents the model described in Section 4.7.4. The model was fitted in order to shift the notch frequency from 4.25 to 3.8 kHz. The $1/f^2$ response from the pendulum is indicated by the pink dashed line. The presence of the expected notch structure is clearly confirmed by the measurement. However, the notch seems to be smeared out which might be explainable by beam jitter of the main interferometer beam (see Section 4.7.6).

Figure 4.13 shows the result of the photon pressure calibrator injections at high frequencies. The presence of the expected notch structure is clearly confirmed by the measurement. However, the notch seems to be smeared out. The magnitude of the response is about a factor of 3 below the pendulum response for frequencies between 3 to 4 kHz. At 5 kHz the measurement matches the pendulum response and finally at 6 kHz the measured response clearly exceeds the $1/f^2$ behavior of the pendulum response. The green line in Figure 4.13 represents the model including the photon pressure calibrator induced mirror deformation, described in Section 4.7.4. The model was fitted in order to shift the notch frequency from 4.25 to 3.8 kHz; this required an increase in the response, α_{def}, by about 25 %, from the originally estimated value of $2.73 \cdot 10^{-10}$ m/N to $3.41 \cdot 10^{-10}$ m/N. Around the notch frequency, the phase of the photon pressure calibrator response also changes significantly from about -165 degree at 2.8 kHz to about -30 degree at 4.8 kHz. A phase of nearly 0 degree at high frequencies clearly indicates that the response is no longer dominated by the pendulum response.

All together this result clearly confirms the hypothesis that the observed relative discrepancy between official and photon pressure calibration at high frequencies is caused by the influence of an additional effect with flat response and 0 degree phase lag compared to the modulated light power. The mirror deformation described above is obviously a strong suspect for this additional effect. However, we have to find a mechanism that can cause the smearing out of the notch structure. A candidate for such an effect is given in the next section.

4.7.6. Influence of beam jitter to the photon pressure calibration

So far we assumed a perfect spatial overlap of the beams from the photon pressure calibrator and the main interferometer. The beam of the photon pressure calibrator is expected to show no significant beam jitter because the whole setup is rigidly mounted on an optical table and the leaver of the beam is only a few meters (see Figure 4.2). However, the main interferometer is known to have some beam jitter. On long time scales its position on MFn is controlled by a low bandwidth servo. For frequencies above the unity gain frequency of this servo, which is around 0.01 Hz, the beam jitter is not suppressed.

If the two beams from the photon pressure calibrator and the interferometer move relative to each other, the effective displacement seen by the interferometer changes. A

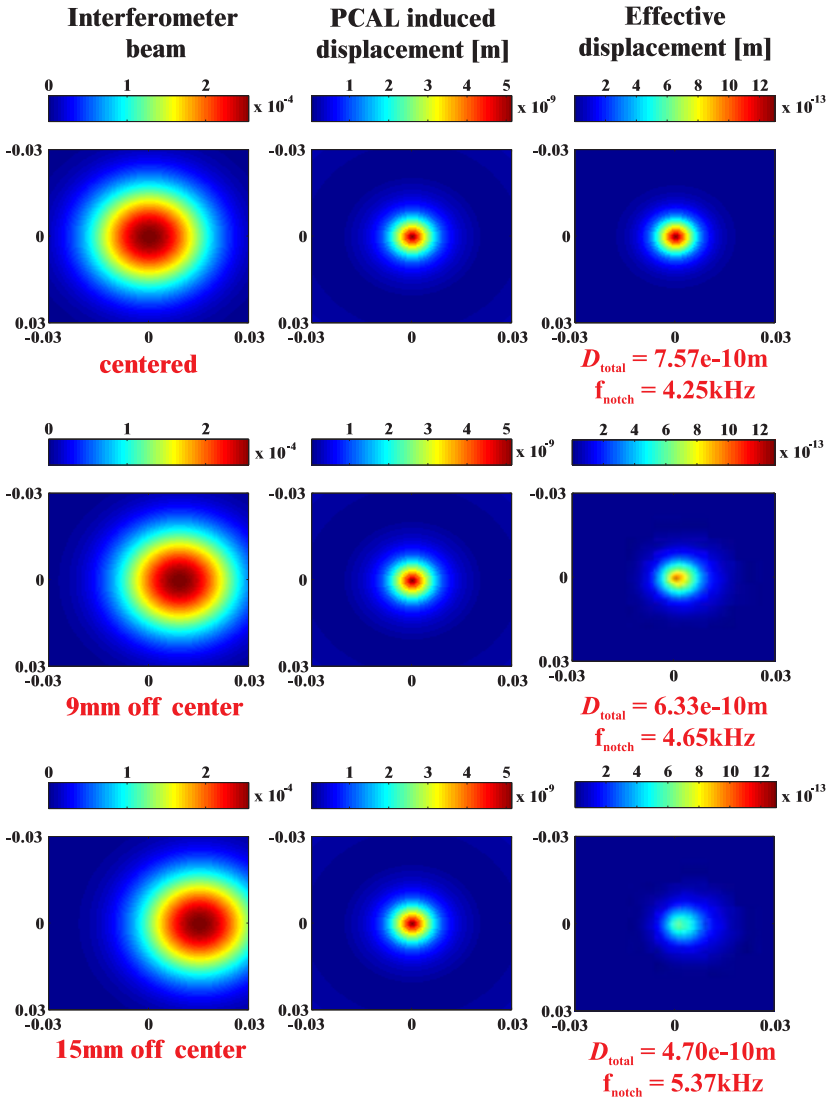


Figure 4.14.: The total effective displacement D_{total} caused by the photon pressure calibrator induced mirror deformation depends on the spatial overlap of the beams from the main interferometer and the photon pressure calibrator. The three plots in the left column show the main interferometer beam ($\text{TEM}_{00}, \omega = 2.4\,\text{cm}$). The plots in the center column display the test mass deformation derived from the FEM simulation (see Section 4.7.2). The right hand column shows the contribution to the effective displacement. All plots show an area of $6 \times 6\,\text{cm}$. A relative drift of the two beams of 15 mm would decrease the D_{total} by about 40 % from $7.57 \cdot 10^{-10}\,\text{m}$ to $4.7 \cdot 10^{-10}\,\text{m}$, corresponding to a change of more than 1 kHz of the notch frequency.

change in the relative size of the two displacement components would lead to a different cross over frequency of the responses from the pendulum and the mirror deformation and thus a shift in the notch frequency. The measurements, shown in Figure 4.13, are averaging over hours and thereby also averaging over different notch frequencies. This would lead to a smearing out of the notch as observed in the measurement.

The question we have to answer now is how much the main interferometer beam moves in reality and whether this movement gives a large enough shift of the notch frequency. Over a period of 10 hours a maximum variation of the position of 9 mm was observed in vertical direction. Analogous to Equation 4.8 we can now calculate the overall measured displacement, D_{total}. As this problem is not any longer rotationally symmetric we have to modify Equation 4.8 slightly by introducing a double integration.

$$D_{\text{total}} = k_I \int_0^{0.09\text{m}} \int_0^{360\,\text{deg}} \cdot I(r,\varphi) \cdot D(r,\varphi) \cdot dr \cdot d\varphi. \tag{4.12}$$

In Figure 4.14 the total effective displacement, D_{total}, caused by the photon pressure calibrator induced mirror deformation is shown for different spatial overlaps of the beams from the main interferometer and the photon pressure calibrator. The three plots in the left column show the amplitude of the main interferometer beam ($\text{TEM}_{00}, \omega = 2.4\,\text{cm}$). The plots in the center column display the test mass deformation derived from the FEM simulation (see Section 4.7.2). The right hand column shows the contribution to the effective displacement derived using Equation 4.12. A relative drift of the two beams of 15 mm would decrease D_{total} by about 40 % from $7.57 \cdot 10^{-10}$ m to $4.7 \cdot 10^{-10}$ m, corresponding to a change of 600 Hz of the notch frequency.

Overall the influence of beam jittering of the main interferometer beam seems to be sufficient to explain that the measured notch in Figure 4.13 is smeared out.

4.7.7. Conclusion and consequences

In the previous section a new idea was presented that can, at least to a large extent, explain the strong discrepancy between the official and the photon pressure calibration observed at high frequencies. In contrary to general belief, a test mass made of fused silica seems to show non rigidity also at frequencies below its first internal resonance.

The achievable accuracy of a calibration derived from a photon pressure actuator is at least to some extent limited by the non-rigidity of the test mass for frequencies in

the kHz range. However, careful design and the implementation of new ideas, like for instance minimizing the overlap of the PCAL beam and the interferometer beam by the use of two off center PCAL beams [Savage06], may help to improve the calibration accuracy. However, it needs to be taken into account that even off center PCAL beams can excite eigenmodes of the test mass.

Finally it has also to be stated that the above described deformation of a test mass can not only be induced by strong modulated light beams, but also by any other kind of inhomogeneous actuation, such as for example coil magnet actuators.

4.8. Effects from test mass rotation

The photon pressure calibrator might cause a non-negligible rotation or tilt of the test mass. One potential mechanism for that would be if the photon pressure calibrator is hitting the test mass off center. The frequencies of our interest are far above the resonance frequencies of the suspension, thus we consider the test mass as being a free mass for rotation or tilt. The rotation corresponding to an angle θ can be described by using the moment of inertia I

$$\theta(\omega) = \frac{2 \cdot P \cdot d_{PPD}}{c \cdot I \cdot \omega^2} \tag{4.13}$$

where P is the modulated light power, ω is modulation frequency and d_{PPD} the distance of the PPD beam center from the center of the testmass.

Figure 4.15 depicts the longitudinal effects from mirror rotation. If the beam from the main interferometer is centered on the testmass the longitudinal displacement is given by

$$x_d \approx \frac{d}{4}\theta^2 \qquad \text{for} \quad \theta \ll 1. \tag{4.14}$$

In the case that the beam of the main interferometer is off center, then the effective displacement, x_d also contains a factor linearly depending on θ [Weiland04]:

$$x_r \approx d_c\theta + \frac{d}{4}\theta^2 \qquad \text{for} \quad \theta \ll 1. \tag{4.15}$$

For very small angles, θ, as the ones we are interested in we can neglect the second part of Equation 4.15 and take only the first term. Together with Equation 4.13 this gives

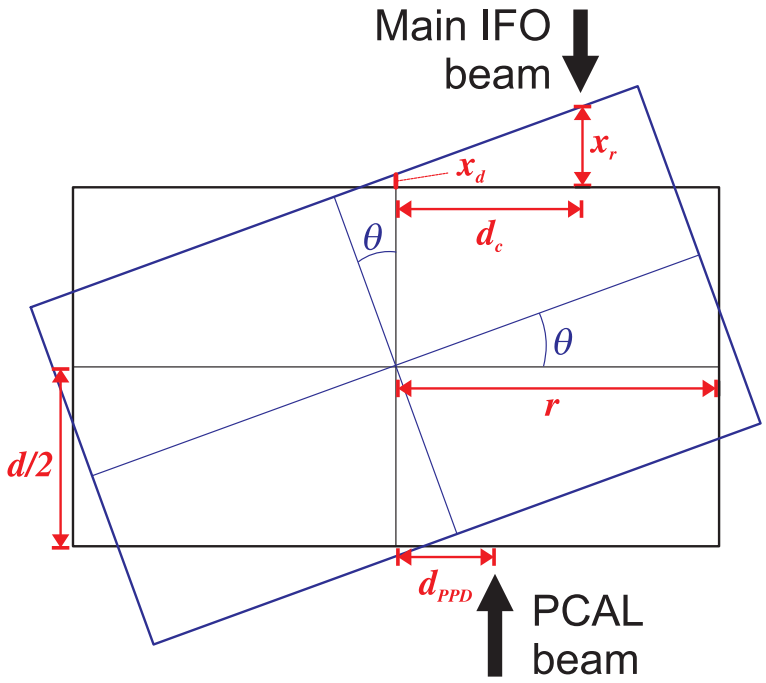


Figure 4.15.: Rotation of the mirror induced by the photon pressure drive will yield a longitudinal displacement of x_d in the case that the main interferometer beam is centered on the test mass. If the main interferometer beam is off center the longitudinal displacement is given by x_r.

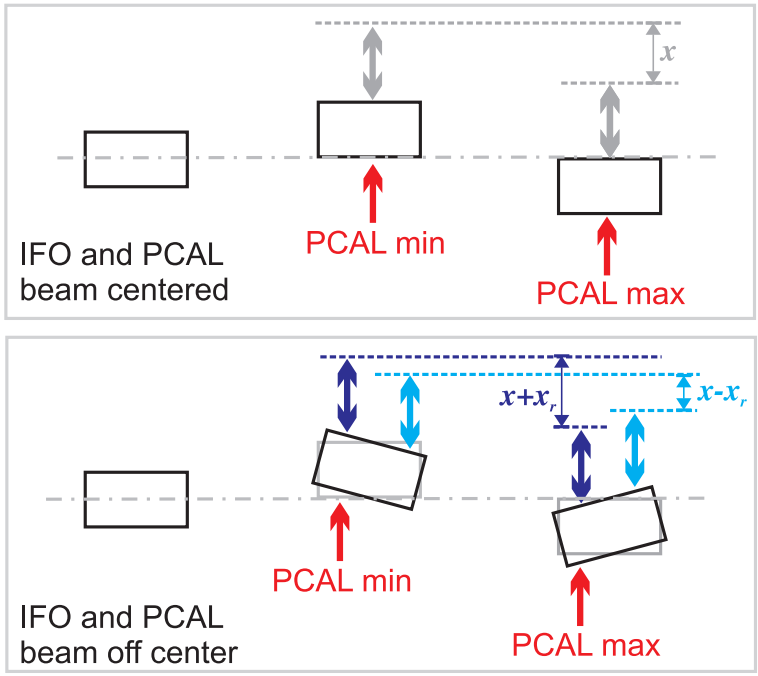


Figure 4.16.: *Upper plot*: When the PCAL beam (red) and the main interferometer beam (grey) are centered the amplitude of the photon pressure calibrator injection is x. *Lower plot*: When both beams are off center to the same side the measured amplitude of the injection will be increased to $x + x_r$ (blue). If both beams are offset to opposite sides the measured signal amplitude is decreased to $x - x_r$ (cyan).

an angle-independent displacement

$$x_r(\omega) \approx \frac{2P \cdot d_c \cdot d_{PPD}}{I \cdot c \cdot \omega^2} \tag{4.16}$$

which only depends on the modulation frequency and the distances of the laser beams from the center of the testmass.

If we want to check the influence of the rotational induced displacement we have to compare this with the longitudinal displacement from the pure pendulum transfer function

$$x(\omega) = \frac{2 \cdot P}{M \cdot c \cdot \omega^2}. \tag{4.17}$$

The ratio is

$$\frac{x_r(\omega)}{x(\omega)} = \frac{M d_c d_{PPD}}{I} \tag{4.18}$$

and does not depend on ω. Hence the rotation of the photon pressure calibrator induced test mass rotation results in a frequency independent error, i.e. an absolute error. Using $M = 5.3$ kg and $I = 0.01601\,\mathrm{kgm}^2$ we can calculate how far the beams have to be off center to explain the observed error of 30 % in Figure 4.5:

$$d_r d_{PPD} = 0.3\frac{I}{M} = 9.075 \cdot 10^{-4}\mathrm{m}^2 \tag{4.19}$$

Assuming both beams to be misaligned by the same distance we would have to take the square root which would give $d_r = d_{PPD} = 3.0\,\mathrm{cm}$.

If we were aiming for an 1 % accuracy we would have to make sure that both beams, the one from the photon pressure calibrator and the one from the interferometer are well centered on the test mass to within 0.5 cm.

Unfortunately with the currently installed setup of the photon pressure calibrator it is not possible to get an estimate better than a few centimeters of the position where the PCAL beam hits the mirror. This is mainly due the fact that the position of MFn in respect to the view port used for the PCAL beam is not accurately known. Furthermore the positions of the view ports available at MFn do not allow for a more accurate optical estimation of the spot positions. However, in the following section a method will be presented to accurately determine the spot position of the photon pressure calibrator beam by scanning the test mass with the position of the main interferometer beam.

4.8.1. Method to determine the off centering of the photon pressure calibrator

As given by Equation 4.15 the effective displacement x_r depends on one hand on the photon pressure calibrator induced misalignment of the test mass and on the other hand on the spot position of the main interferometer beam. If the beam from PCAL is exactly centered on MFn, no test mass misalignment is induced, and changing the spot position of the main interferometer beam should not change x_r. However, in the case the beam from PCAL is off center, x_r is a function of the spot position of the main interferometer beam, d_c. This relation can be used to accurately determine the position of the PCAL beam hitting MFn, d_{PPD}. Figure 4.17 shows the corresponding

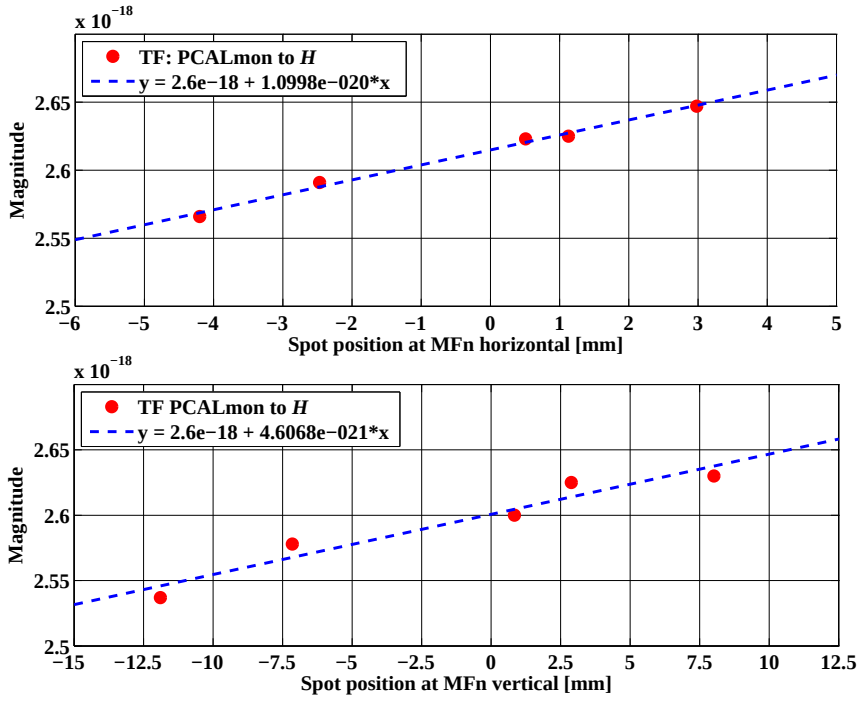


Figure 4.17.: Magnitude of the transfer function from PCALmon to H versus the position of the main interferometer beam at MFn. *Upper plot*: Changing the position of the main interferometer beam on MFn in horizontal direction changes the photon pressure calibration by 4.23 % / cm. *Lower plot*: Changing the position of the main interferometer beam on MFn in vertical direction changes the photon pressure calibration by 1.71 % / cm.

measurements. The position of the main interferometer beam was shifted on MFn and

the transfer function from PCALmon to H was measured. Changing the position of the main interferometer beam on MFn in horizontal direction changes the photon pressure calibration by 4.23 %/cm, while in vertical direction a change of 1.71 %/cm is found. Using Equation 4.18 these values can be used to calculate the horizontal and vertical off centering of the PCAL beam, d_{PPD}^{rot} and d_{PPD}^{tilt} respectively.

$$d_{PPD}^{\mathrm{rot}} = 12.8\,\mathrm{mm} \quad \wedge \quad d_{PPD}^{\mathrm{tilt}} = 5.2\,\mathrm{mm} \tag{4.20}$$

Assuming that the main interferometer beam and the beam from the photon pressure calibrator are off center in the same direction, the interferometer beam would have to be 44 mm off center to explain the smallest observed discrepancy of 20 % (see Figure 4.5) between the official and the photon pressure calibration. Such a large off centering cannot be completely excluded but seems fairly unlikely. Using the method described here d_{PPD} can iteratively be minimized. However, in order to keep the duty cycle of GEO 600 in S5 as high as possible, this work was postponed so far.

4.9. Determining the sign of H

For all kinds of network analysis that use data from two or more gravitational wave detectors (see, for example, [Ajith06a, Wen05, Candonati04, Rakhmanov05, Heng03]) it is indispensable to get well calibrated detector outputs. First of all a correct strain amplitude output is needed and secondly the sign of H must be determined.

In principle this information is easily accessible from the polarity of the fast actuators used for longitudinal control of the Michelson differential arm length which are in case of GEO 600 the ESDs. However, in reality the injected signals pass a complex system of electronics containing many stages with potential swaps of sign, like for instance inverted OPs and differential senders and receivers. Thus, one needs to very carefully measure the polarity of many pieces of electronics to find the correct sign of H.

A much less vulnerable procedure to determine the sign of H is provided by the photon pressure calibrator which only involves a very simple and clearly laid out system. The monitor photodiode built in the photon pressure calibrator laser diode and recorded in the DAQS (PCALmon) provides a measurement of the light hitting the suspended mirror. As shown in Figure 4.18 the amplifier of the photodiode gives a positive voltage

proportional to the detected light. The force F_{PCAL} acting on MFn is proportional to the voltage recorded in PCALmon.

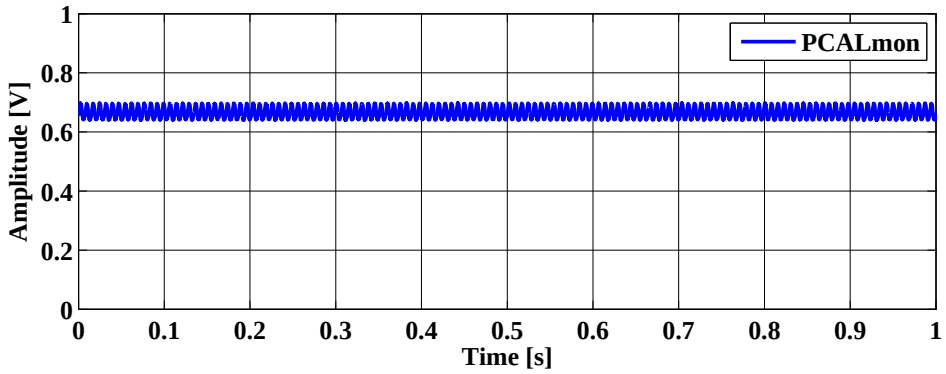


Figure 4.18.: Time series of the signal PCALmon, derived from the internal monitor photo diode of the photon pressure calibrator. The time series consists of a DC voltage of 0.67 V and an AC peak-peak voltage of ±25 mV corresponding to the injected calibration signal.

Next it is necessary to measure the phase relation between PCALmon and H by means of computing the transfer function at the frequency of the photon pressure calibration line f_{PCAL}= 134 Hz. Figure 4.19 shows this transfer function for a time during S5. The two signals have opposite phase, $\phi_{F_{\mathrm{PCAL}}\rightarrow H}$ = -180 deg out of phase. Finally we need to take the pendulum response of the MFn suspension into account. As indicated by Figure 4.1 the phase of the pendulum response is $\phi_{F_{\mathrm{PCAL}}\rightarrow x_{\mathrm{MFn}}}$ = -180 deg for a frequency $f \gg f_0$ where f_0 the resonance frequency of the pendulum and x_{MFn} the movement of MFn is.

Putting both relations together we get:

$$\phi_{x_{\mathrm{MFn}}\rightarrow H} = \phi_{F_{\mathrm{PCAL}}\rightarrow H} + \phi_{F_{\mathrm{PCAL}}\rightarrow x_{\mathrm{MFn}}} = -360\,\mathrm{deg} \tag{4.21}$$

Due to the fact that the photon pressure calibrator shines onto the back of MFn $x_{\mathrm{MFn}} > 0$ corresponds to a shortening of the north arm. Hence, all together we determine the sign of H as follows:

$$H > 0 \quad \text{for} \quad \delta l_{\mathrm{north}} < \delta l_{\mathrm{east}} \tag{4.22}$$

$$H < 0 \quad \text{for} \quad \delta l_{\mathrm{north}} > \delta l_{\mathrm{east}} \tag{4.23}$$

where l_{north} and l_{east} are the lengths of the north and the east arms, respectively.

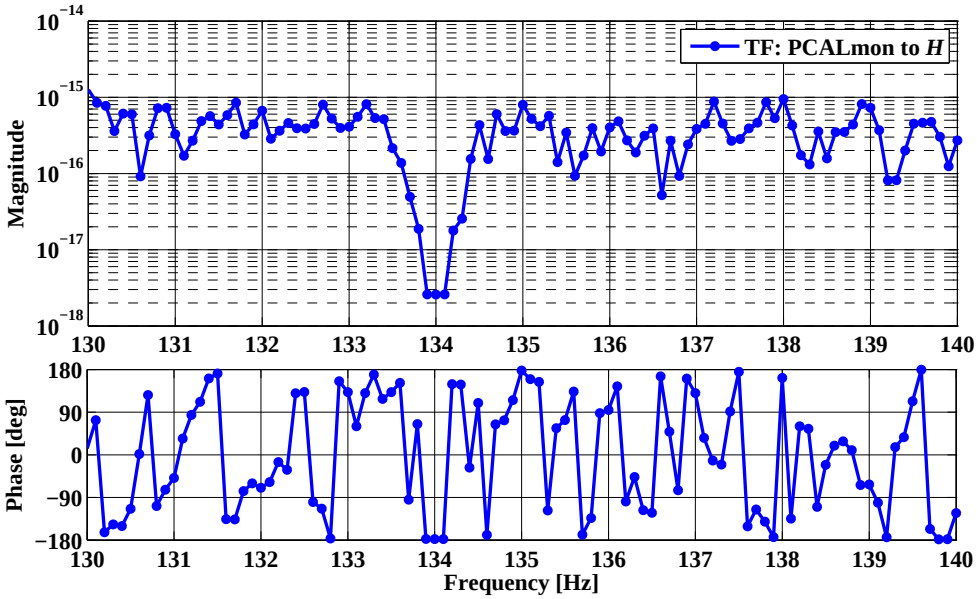


Figure 4.19.: Transfer function from the photon pressure calibrator monitor diode (PCALmon) to H . The photon pressure calibrator is used to apply a calibration line at $f_{\mathrm{PCAL}} = 134\,\mathrm{Hz}$. Only at this frequency is the transfer function is physically meaningful. The phase difference between both signals was measured to be $\phi_{F_{\mathrm{PCAL}} \to H} = $ -180 deg.

4.10. Summary

In this section experiments and investigations related to establishing a photon pressure calibration for GEO 600 are presented. The experimental setup and potential imperfections are described in detail in Section 4.2. A direct comparison of the relative and absolute values of the official GEO calibration and the calibration derived from the photon pressure calibrator is given in Section 4.4. The discrepancy between the two calibration methods was found to be unexpectedly large. The two resulting absolute calibrations seem to be different by about 20 to 40 %, which might be explained by effects from photon pressure induced test mass rotation as described in Section 4.8. The Section 4.7 focuses on the strong discrepancy of the relative calibration found at high frequencies, which can be explained by photon-pressure-induced deformation of the test mass. A finite element analysis is presented which predicts the presence of a notch structure in the response of the photon pressure calibrator. In long-duration measurements of the photon pressure calibrator at very high frequencies this notch structure was found. The The effect of non-rigidity of the test mass limits the accuracy of the photon pressure calibrator for frequencies above 1 kHz. Finally in Section 4.9 it is shown that a photon pressure calibrator can be used to reliably determine the polarity of H which is important for any kind of astrophysical analysis using data from more than one GW detector.

Chapter 5.

A statistical veto method employing an amplitude consistency check

5.1. Introduction

One of the most promising class of gravitational wave (GW) signatures that may be detected by the current array of ground-based laser interferometric GW detectors [Hild06c, Waldman06, Acernese06, Ando05] is un-modelled transient (burst) signals arising from short-duration, violent astronomical events. The search for such signals typically focusses on time-scales of the order of a few (or few tens of) milliseconds.

Due to their extremely complicated nature, GW detectors are themselves usually potent sources of transient signals. A subset of such signals arising from many different sub-systems within the detectors can (and do) couple to the main detector output, and hence appear as false GW triggers in any search effort. While a great amount of effort is afforded in the reduction of these instrumental and environmental glitches and/or their coupling to the main detector output (also termed the H channel), there is still a residual number that remain.

Once the glitch population of a particular sub-system (and its coupling to the GW channel), is (physically) reduced to a minimum, the residual must be identified and characterised so as to exclude those glitches from the search for GWs. The rate of false glitches in the GW channel ultimately sets a limit on the confidence with which a particular trigger can be identified as a GW. It is therefore important to try to reduce the number of glitches in the GW channel that are to be considered in a search for GWs. This is traditionally done by *vetoing* those events detected in H using events and

knowledge of events detected in auxiliary channels. Once an auxiliary channel has been identified as being a source of transient events which couple to H, it is termed a veto channel and is then studied and used to reduce the event list of H.

The application of a veto is done by identifying those events in the auxiliary channel which are, in some way, coincident with events detected in H. Detected glitches are typically characterised by a few parameters (time of occurrence, amplitude, central frequency, *etc*), such that saying that two events are coincident can be as simple as saying that they occur at the same time (within some error window) or as complicated as saying that many of the characteristics of the glitches are (within some defined windows) the same. When a parameterisation of glitch events is used to compare events between data-streams, we refer to the resulting veto as a *statistical* veto since its performance is based on the statistical properties of the distributions of glitches in the two data-streams and does not rely on any knowledge of the physical coupling mechanisms involved. Examples of the application of statistical veto methods to GW detector data can be found in [DiCredico05, Akutsu, Beauville05]. Another class of veto methods which relies on detailed knowledge of the coupling mechanism involved in transferring glitches from an auxiliary channel to the H channel, is not discussed here, but an example is given in [Ajith06]. There are also interferometer channels which, by definition, contain negligible GW information but can nevertheless remain highly correlated to the main GW channel. These 'null-stream' channels can also be used as effective veto channels (see [Hanna06, Koetter03, Hewitson05a] for examples). In addition to these single-detector veto methods, there is active research in the use of multiple detector outputs as a means of vetoing the events in a detector network (see, for example, [Ajith06a, Wen05, Candonati04, Rakhmanov05, Heng03]).

The approach of using a pure statistical correlation (such as time of occurence) can, of course, lead to *false vetoes*, that is, events in the auxiliary channel which are only accidentally coincident with events in the H channel. Requiring that more glitch characteristics be 'similar' when performing the coincident test can reduce the false-veto rate (usually at the cost of a reduced veto efficiency). The usefulness of a particular veto channel can be characterised according to the number of H events it can veto (its efficiency) for a given number of false-vetoes (false/accidental-veto rate).

This chapter investigates many aspects of the identification, characterisation and use of veto channels for data recorded from GEO 600, in particular for the case when

the simple coupling described above is further complicated by the presence of GW signal traces in the auxiliary channel. To motivate the application of statistical veto methods to data from GEO 600, in the Sections 5.2 and 5.3 two examples are given where events from the auxiliary channel X and the GW channel H are found to show a significant correlation. Section 5.4 describes the burst detection algorithm used in these studies. Section 5.5 goes on to discuss the *standard* statistical veto and shows an example application (with single and multiple coincidence windows) to data from GEO 600. Section 5.6 shows a possible extension to the classical statistical veto method for the case where the auxiliary channel can contain traces of GW signals, *i.e.*, the channel's sensitivity to GW signals is non-negligible. Section 5.7 shows the application of the extended classical statistical veto method to GEO 600 data. For a month of data the full GEO veto pipeline is applied, also including a statistical veto employing a amplitude consistency check.

5.2. Motivation for a standard statistical veto for GEO 600: Hourly mains glitches

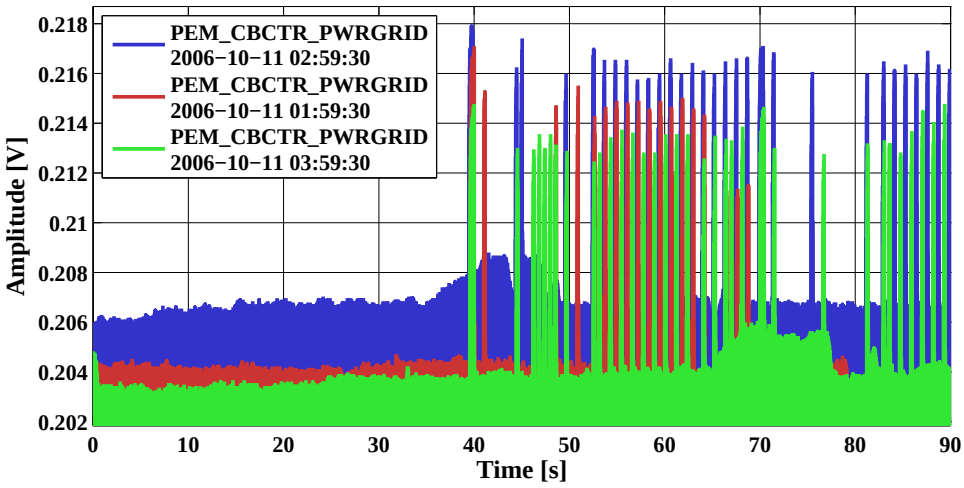


Figure 5.1.: Stacked time series of the mains monitor at the GEO site. Ripple control signals repeatedly show up 10 seconds after the beginning of a UTC-hour.

The power supply companies in Germany use a technique called *ripple control* to trans-

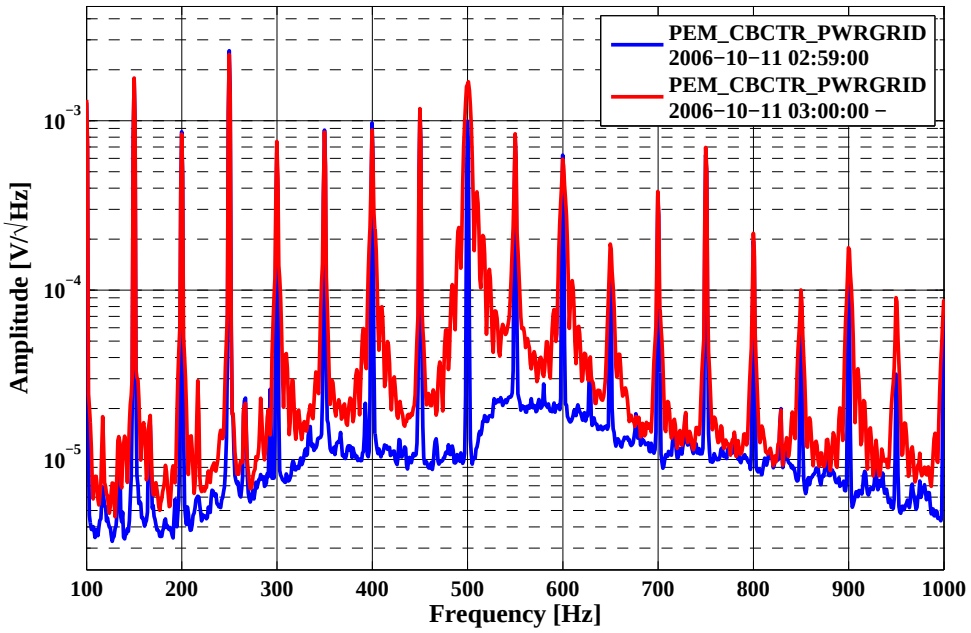


Figure 5.2.: Spectra of the mains monitor for the last (blue) and the first (red) minute of a UTC-hour. The ripple control signal has a nominal frequency of 500 Hz, but it is also strongly visible as sidebands around all harmonics of 50 Hz.

fer information via the mains from the power plant to the consumers. An amplitude modulation with a certain frequency (in our case 500 Hz) is superimposed on the usual 50 Hz power grid in order to remotely control lanterns and night storage heaters.

Figure 5.1 shows overlayed time series from a monitor of the main site power supply (mains). The ripple control signals repeatedly show up 10 seconds after the beginning of an UTC-hour. They consist of an up to 90 seconds long series of on-off-states with a maximum amplitude increase of about 6 %.

Spectra of a time when ripple control signals were present, compared to a time of their absence are shown in Figure 5.2. The ripple control signals have a nominal frequency of 500 Hz, but they are also strongly visible as sidebands around all other harmonics of 50 Hz, especially 400, 600 and 900 Hz.

It was found that, as indicated in Figure 5.3, some of the glitches in the mains monitor are time coincident with glitches in the main gravitational wave channel. A large fraction of the coincident events shows up shortly after the start of an hour. Furthermore nearly all of them are identified to have central frequencies of about 400, 500, 600 and 900 Hz, the same frequencies, where we found most of the excess power in the mains monitor caused by the ripple control (see Figure 5.2).

The coupling mechanism is not completely understood yet. Glitches in the mains can potentially couple in various ways to the GW channel. A glitch in the mains causes a glitch in the magnetic field surrounding the power line which can then interact with magnets glued onto the mirror [Gossler02]. Or the magnetic glitch can directly induce a voltage glitch in a piece of electronics used for control or readout of the detector.

However, as the coupling path for the ripple control signals to the gravitational wave channel is not known and therefore not accessible via measurements, no other veto method than a statistical veto can be applied. In Section 5.5.2 we will show some investigations done in order to find out whether a reasonably performing veto analysis can be based on the mains monitor.

In the long term we plan to filter the ripple control signals out of the local power grid by using some resonant LC-filter stages[1]. The installation of such a filter requires powering down of the full GEO 600 site which needs to be well prepared and is not free of risk.

[1] More detailed information about available filters can be found in http://www.eskap.de/tonfrequenzsperren.pdf .

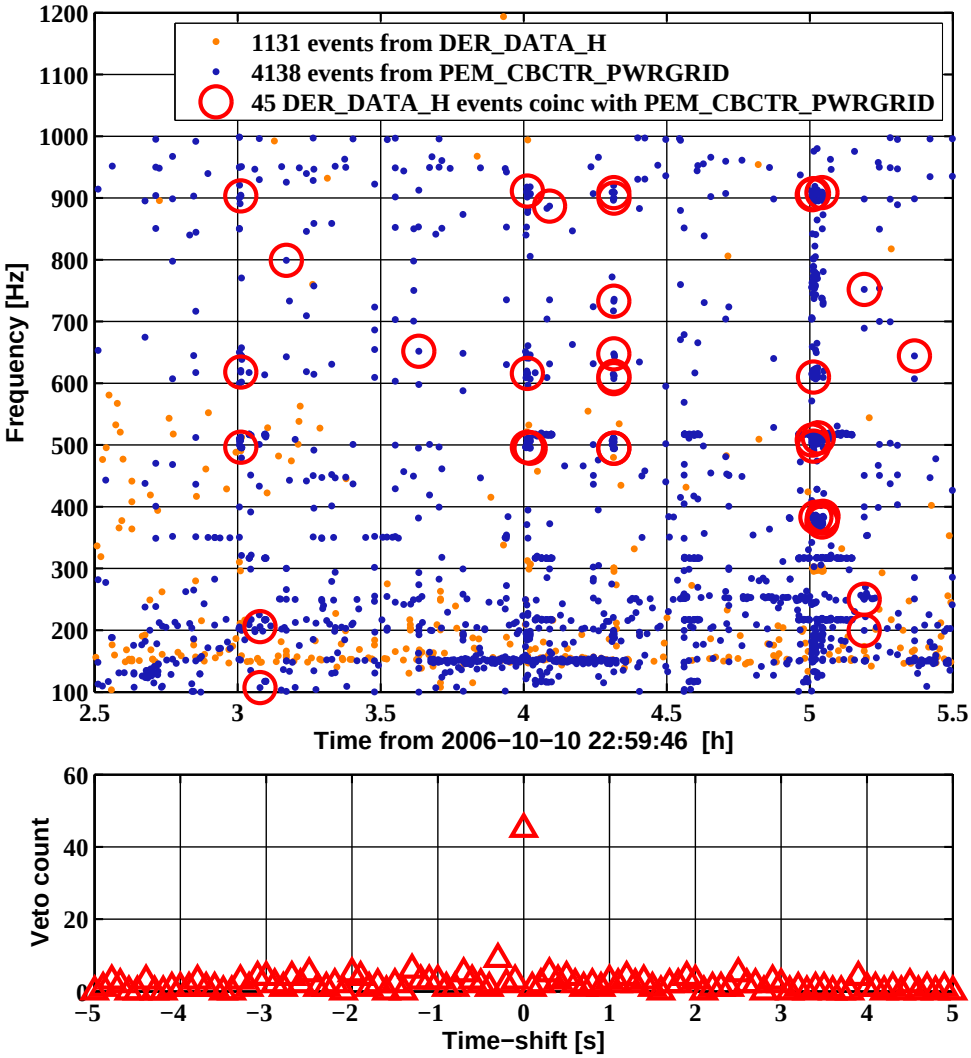


Figure 5.3.: *Upper plot*: Time frequency map of burst triggers in the mains monitor (blue dots) and the main gravitational wave channel (orange dots). The events in the mains monitor that are time coincident with an event in the GW channel are marked with a red circle. A time coincidence window of 10 milliseconds was used. No further coincident condition was imposed. *Lower plot*: Time shift analysis of the coincidence analysis from the plot above. A significant statistical correlation between the events from the mains monitor and the events from the GW channel is confirmed.

5.3. Motivation for a statistical veto with amplitude consistency check: Dust falling through laser beams

In May 2006, a significant increase in the glitch rate of the main GW channel (H) of GEO 600 was observed. A broken air conditioning system blew unfiltered air into the main clean room and increased the dust particle concentration by more than one order of magnitude. It turned out that the increase in the glitch rate of H originated from dust particles falling through the main output beam of the interferometer.

Many glitches in the recorded DC light power hitting the main photodiode (referred to as $\mathcal{P}_{\mathrm{DC}}$), were observed to be coincident with glitches in H. Figure 5.4 shows a time coincidence analysis of the two channels for two different times of the S5 LSC science run. The upper plot shows an 8 hour data stretch from May 2006, when the air conditioning system was broken. 1245 of the 1719 detected events in H (72 %) are time coincident with an event in $\mathcal{P}_{\mathrm{DC}}$. For this analysis a time coincidence window of 10 milliseconds was used. The lower subplot of Figure 5.4 shows the same analysis (with identical parameters) for a data segment from June 2006, when the air conditioning system was fixed again. Compared to the time of high dust concentration, in the low dust condition the total number of events from both signals is significantly reduced. Only about 5 % of the events in H are time coincident with an event from $\mathcal{P}_{\mathrm{DC}}$.

However, tests involving injecting signals (noise, sinusoidal, burst events) into the differential length control actuator for the Michelson Interferometer (to mimic the effect of a GW), showed that $\mathcal{P}_{\mathrm{DC}}$ can contain, to a non-negligible degree, some GW signal. Figure 5.5 shows sinusoidal GW-like hardware injections (calibration lines) using the electro static drives. Spectra of the injected signal (*diff_CAL*) and $\mathcal{P}_{\mathrm{DC}}$ are plotted in the two upper figures. The frequencies of the injected sine waves are indicated by the circled markers. The lower figure shows the coherence of the two signals, which is nearly one for all frequencies of the hardware injections.

Hence, using a standard statistical veto could, undesirably, also veto potential GW events. That is why we developed a method, described in the Sections 5.6–5.7, that extends the standard statistical veto methods by an additional amplitude consistency check.

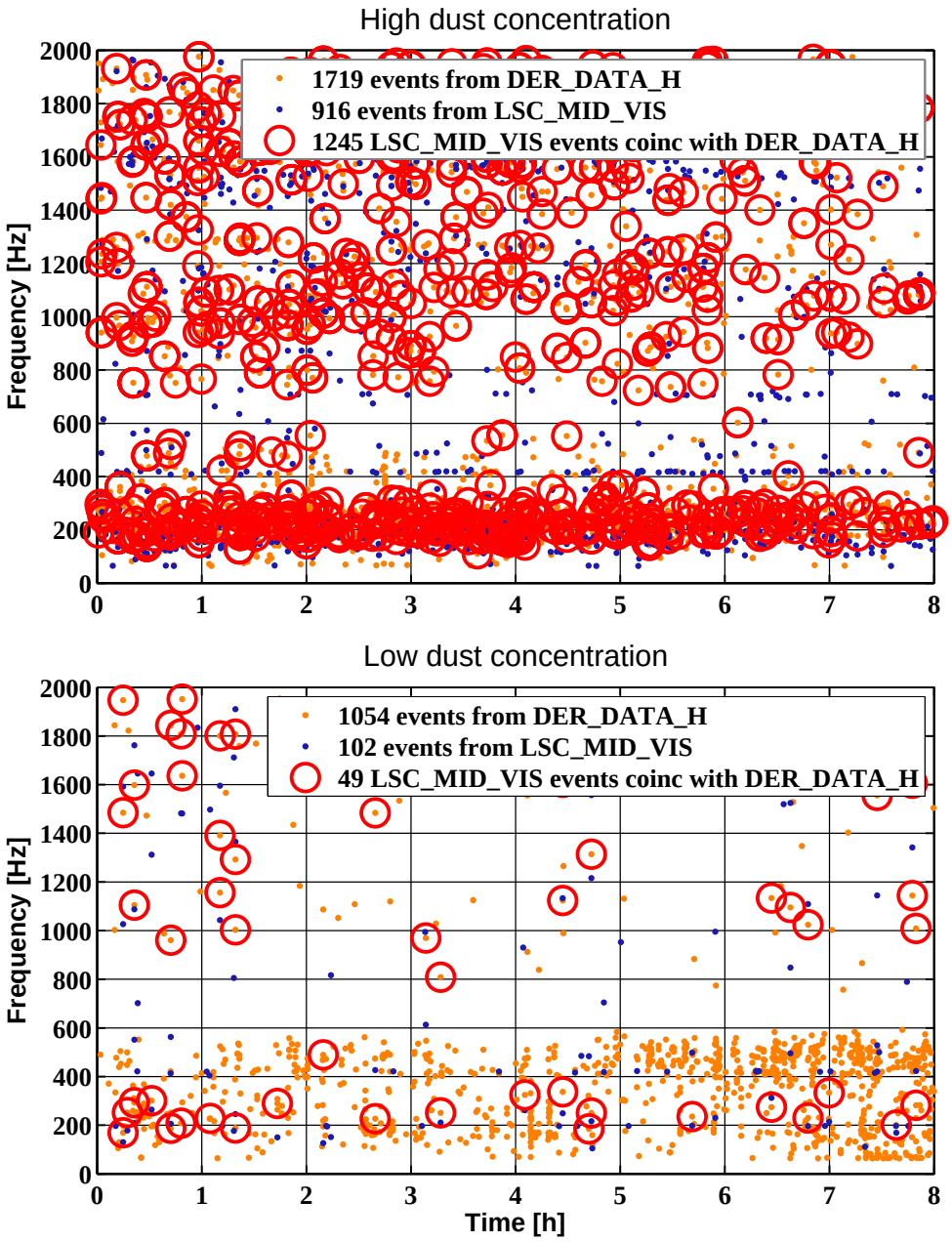


Figure 5.4.: *Upper Subplot*: Time coincidence analysis for the GW channel (H) and $\mathcal{P}_{\mathrm{DC}}$ for a time of high dust concentration at the detection bench, corresponding to a time when the air conditioning system was broken. *Lower Subplot*: The same analysis, but for a time of low dust concentration at the detection bench, corresponding to a time of nominal operation of the air conditioning system. For both analysis a time coincidence window of 10 milliseconds was applied.

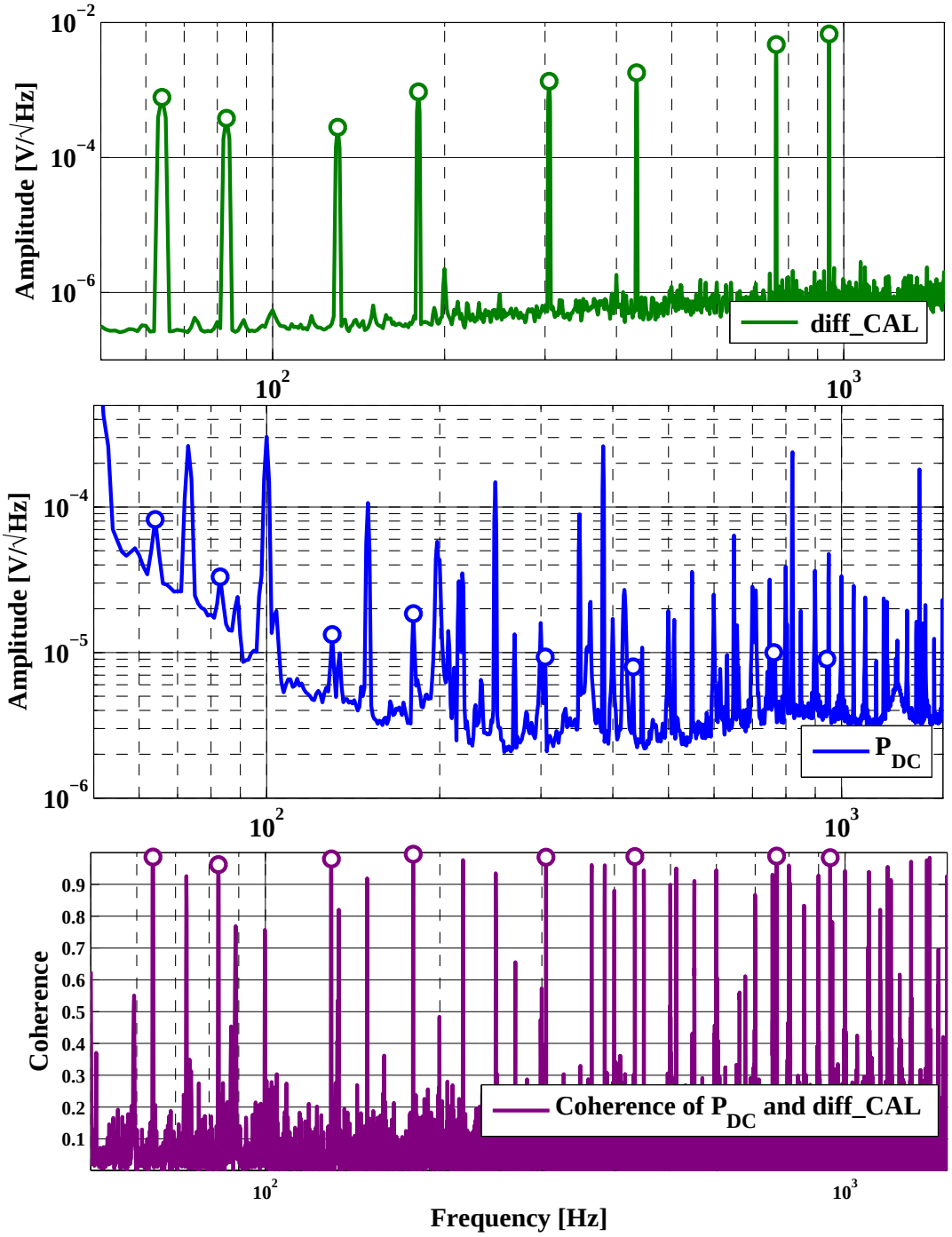


Figure 5.5.: Sinusoidal gravitational wave like hardware injections: The upper plot is a spectrum of the injected signal. The frequencies of the injections are marked with circles. The plot in the middle shows a spectrum of $\mathcal{P}_{\mathrm{DC}}$. The lower plot shows the coherence of the two signals. For all frequencies of the injections the coherence is found to be nearly one. This confirms that a non neglectabel amount of differential arm length signal couples to $\mathcal{P}_{\mathrm{DC}}$.

5.4. Description of the event trigger generator used: mHACR

Any potential *burst* gravitational wave signal is expected to be very weak and of very short duration, barley visible and hard to separate from the intrinsic noise of the instrument. Therefore sensitive detection algorithms have been developed in order to identify and parameterize any excess power in the main gravitational wave output. Examples of these event trigger generators (ETG) used for astrophysical searches are *Kleine Welle* [Chatterji04], *Excess Power* [Anderson01] and *Waveburst* [Klimenko04]. For the commissioning of the GEO 600 detector we use an ETG called *mHACR*, which is described in detail in [Balasubramanian05] [Heng03] and [Hild07a].

For the understanding of the following sections it is important to know how mHACR roughly works. That is why in this section I give a brief and simplified description of mHACR.

32 seconds of data from the channel to be analyzed are divided into short segments of 32 milliseconds length with an overlap of 28 milliseconds. From each of the 32 milliseconds-long subsegments a FFT is computed. Afterwards, all of the FFTs are joined together to give a time-frequency map, also known as spectrogram, with a very high time resolution of 4 milliseconds and a (poor) frequency resolution of 32 Hz. A normalization of the spectrogram is performed, such that the spectrogram has units of power spectral density.

This normalized spectrogram will be called ρ and each pixel ρ_{kl} can be identified by two indexes k and l, representing frequency bin and time, respectively. The main task of the ETG is to identify pixels which are statistically different from the background noise. This is done by computing the significance, s_{kl} of each pixel, given by

$$s_{kl} = \frac{\rho_{kl} - \mu_k}{\sigma_k}., \tag{5.1}$$

where μ_k is the mean and σ_k is the standard deviation of the k^{th} row of ρ. The significance is now used to sort the pixels into three groups of color, according to the following criteria:

$$\text{colour}(\rho_{kl}) = \begin{cases} \text{black}, & \text{if } s_{kl} \geq T_{\text{upp}} \\ \text{grey}, & \text{if } T_{\text{upp}} > s_{kl} \geq T_{\text{low}} \\ \text{white}, & \text{if } s_{kl} < T_{\text{low}}, \end{cases}$$

where T_{low} and T_{upp} are the chosen lower and the upper threshold. For the analysis described in the following the thresholds were chosen to be 5 and 25. In the next step mHACR identifies clusters of neighboring grey and black pixels. If a cluster consists of at least two pixels (black, grey) of which at least one is black, this cluster is identified as a burst event.

5.4.1. Estimation of burst parameter with mHACR

After the burst event is identified by mHACR we need to parameterize this event. Amplitude, time of occurrence, central frequency and duration are, amongst others, the most important parameters for the understanding of the following sections. The signal power of a single pixel, S_{kl}, is given by the power of the pixel minus the mean noise-power of the frequency bin:

$$S_{kl} = \rho_{kl} - \mu_k. \tag{5.2}$$

With this, the amplitude, a, of an event can be computed as

$$a = \sqrt{\frac{\sum_{k,l} S_{kl}}{r}}, \tag{5.3}$$

where r is the redundancy factor accounting for the 28 milliseconds overlap of each FFT. The estimation of the *central frequency*, f_0, and *central time*, t_0, is analogous to the calculation of the centre-of-mass of an extended body. Here, the *signal power*, S_{kl}, in each time-frequency pixel serves as the 'mass' term and the time/frequency associated with each pixel serves as the 'position' term. That is,

$$f_0 = \sum_{k,l} S_{kl}\, f_k \Big/ \sum_{k,l} S_{kl} \tag{5.4}$$

$$t_0 = \sum_{k,l} S_{kl}\, t_l \Big/ \sum_{k,l} S_{kl}. \tag{5.5}$$

Finally, the duration, d, of an event can be estimated by from the extend of the event in the time axis as

$$d = \max(t_l) - \min(t_l). \tag{5.6}$$

Usually there are more parameters available from mHACR, for instance bandwidth, peak power and SNR, but these are not used for any analysis presented in this work. In Figure 5.6 the whole algorithm of mHACR is briefly illustrated.

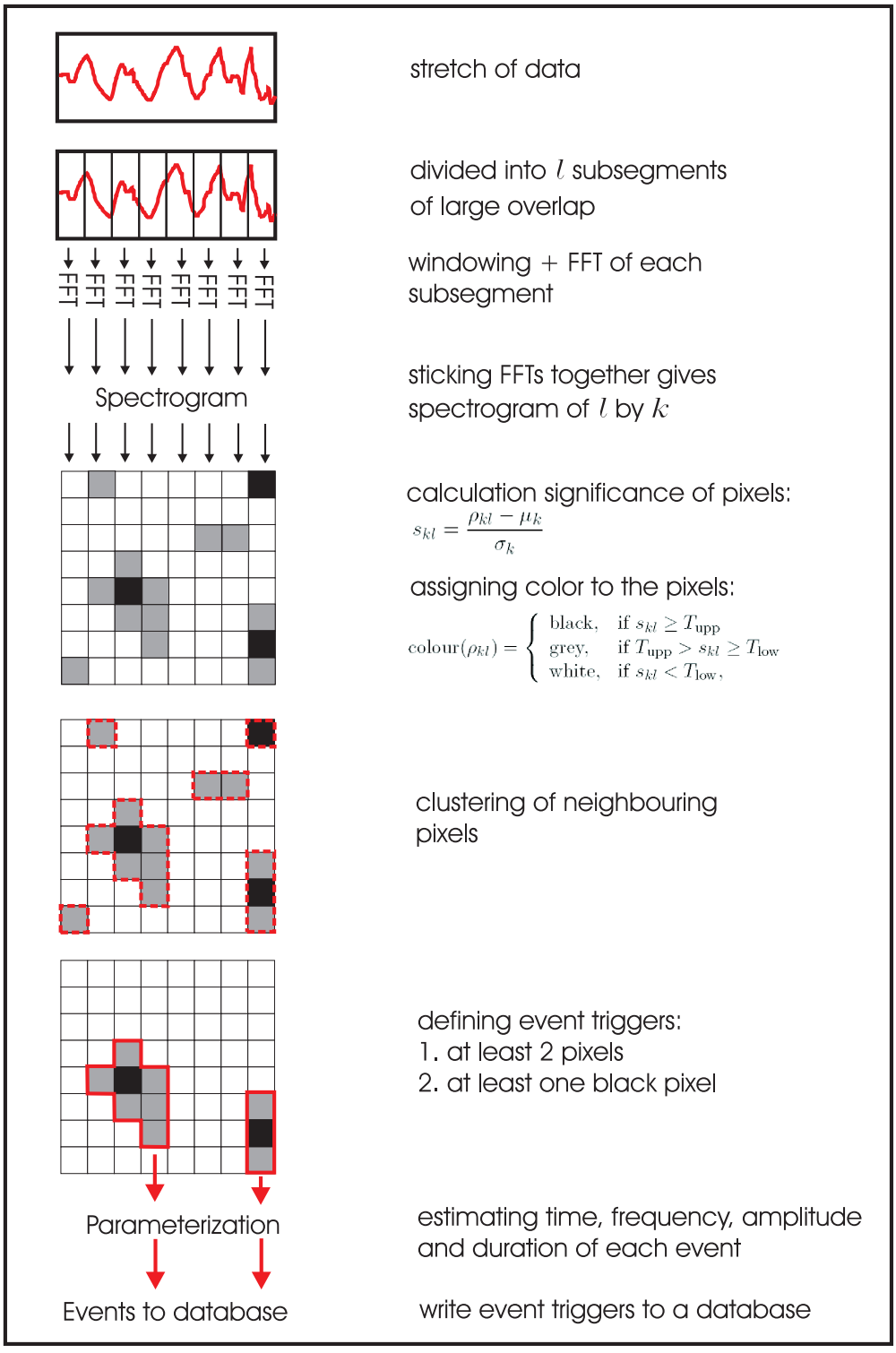


Figure 5.6.: Simplified illustration of the mHACR burst pipeline. A detailed description is given in the text.

5.5. The standard statistical veto

All currently operating GW detectors face the problem that several instrumental and environmental noises can couple to the GW channel. Therefore, it is desirable to identify in the GW stream, those events which can be shown to have a local instrumental or environmental source, and exclude them from any astrophysical analysis.

At each GW detector a multitude of sensors is installed to monitor instrumental noise sources (for instance, laser power noise) and environmental noise sources (such as magnetic fields). In the following, we will refer to both instrumental and environmental noise as technical noise sources.

If detected events in the GW channel are caused by a technical noise source, there will be a significant statistical correlation between the events of the GW channel and those detected in the recording of any device which monitors the noise source. (This assumes that both the main detector output and the auxiliary channel are properly recorded, and that the ETG does a good job of detecting and parameterising any glitch events in the data streams.) Even in the case when the coupling mechanism from the technical noise source to GW channel cannot be measured, we can use the statistical correlation to veto the coincident events.

In order that we end up with a sufficiently low false-veto rate, we must restrict ourselves (for the time-being) to auxiliary channels which cannot contain any GW information.

5.5.1. The method in general

Usually a simple statistical veto is based on the comparison of the time, t_0^H, of the event in the GW channel and the time, t_0^X, of the event in an auxiliary channel. Two events, $H[i]$ and $X[j]$, are defined as being coincident when they are separated in time by less than a chosen time window, $t_{\rm win}$:

$$\mid t_0^H[i] - t_0^X[j] \mid \quad < t_{\rm win}. \tag{5.7}$$

Every event, $H[i]$, that is time-coincident with at least one event from the auxiliary channel is vetoed.

The performance of a veto depends on many parameters (such as the event rates in the individual channels), and needs to be evaluated. The following measures can be useful

to judge the performance of a veto:

- The **Efficiency**, E_X, is the percentage of triggers in the GW channel that are vetoed by the use of the events of the auxiliary channel, X.

- The **Background**, B_X, is the percentage of triggers in the GW channel that are *accidentally* vetoed by the use of the events of the auxiliary channel, X. In general there are two possible groups of coincident events:

  1. The event $H[i]$ is either caused by, or originates from the same source as, the event $X[j]$.

  2. The event $H[i]$ is not caused by any event recorded in the auxiliary channel, but by accident there happens to be an independent event $X[j]$ that is time coincident with the event $H[i]$.

  By time-shifting the events in the GW channel with respect to the events in the auxiliary channel, it is possible to destroy the causal relationship between the events of the two channels and by that, to determine the accidental rate. Sometimes it is useful not to measure the background in percentage, but in units of accidentally-vetoed events per unit time, in order to determine how many potentially real GW events would be missed when applying the veto.

- The **Significance**, S_X, is defined as the ratio of efficiency and background and can be seen as the main figure of merit of a veto analysis. This can be illustrated by some examples. A veto that has a high efficiency of 50 % might also have a high background of, for example, 5 %, which would mean that we would falsely veto a large number of potential GW signals. On the other hand a very low background of 0.01 % does not necessarily guarantee a good veto performance, as it can still have a low efficiency of, say, 0.03 %. That is why we choose here the ratio of efficiency and background as a good way of judging the trade-off between efficiency and background. (When changing the width of a coincidence window, both the efficiency and the background are changed as well, but in general with different slopes.)

- The **Use-percentage**, U_X, is another means of the measuring the veto performance. It is defined as the percentage of the events in the auxiliary channel that can veto an event in the GW channel.

Additional coincident windows

In order to improve the significance of a veto, it might often be useful to demand that, not only are two events coincident in time, but also that other parameters of the events are similar. Assuming the event, $X[j]$, to be the origin of the event $H[i]$, for instance, also the central frequency of the two events, or their duration, might show a correlation. Of course, the strength of the correlation depends strongly on the coupling from channel X to the GW channel, H, and also on the noise level and the stationarity of noise in the two channels. In the worst case, the correlation might be completely destroyed. However, in GEO 600 we found some cases where the application of a second coincidence condition clearly improves the significance of a statistical veto.

In the following sections we will also apply an additional coincidence window, f_{win}, for the central frequency of the events. In that case an event, $H[i]$, in the GW channel is vetoed by the event, $X[j]$, from the auxiliary channel only when the following equation is satisfied,

$$\left|t_0^H[i] - t_0^X[j]\right| < t_{\mathrm{win}} \wedge \left|f_0^H[i] - f_0^X[j]\right| < f_{\mathrm{win}}, \tag{5.8}$$

where $f_0^H[i]$ and $f_0^X[j]$ are the central frequencies of the two events.

5.5.2. Application of a standard statistical veto to GEO S5 data

In this section, we will show some investigations done in order to find out whether it is reasonable to use a monitor of the main site power supply (mains) as a veto channel.

As described already in Section 5.2 glitches on the mains can easily couple to H, while on the other hand it is hard to imagine any way in which a GW could couple back to the mains monitor. This was confirmed by performing GW-like hardware injections. There we inject sine-Gaussian type waveforms into the electrostatic actuators used for the differential length control of the Michelson interferometer and no back coupling to the mains monitor was observed.

Figure 5.7 shows the application of a standard statistical veto to the GW channel using a mains monitor as veto channel. The efficiency and background rate are plotted for time windows between 0 and 20 ms. The shapes of both curves are roughly the same, *i.e.*, the significance stays roughly constant around a value of 20 for different sizes of the time coincidence window.

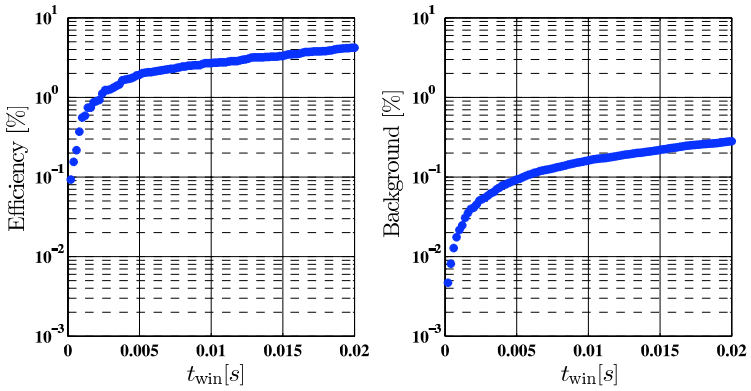


Figure 5.7.: Exemplary application of a statistical veto to a 24 hour stretch of GEO data using a GW-free channel (in this case a mains monitor), as veto channel. Efficiency and background are given as a function of the size of the coincidence time window. The data set used consisted of 3228 and 7725 events in the GW channel and the mains monitor, respectively.

If we choose a time-window of 6 milliseconds in order to get a background of 0.1 %, we achieve an efficiency of about 2 %. Figure 5.8 shows (for the same set of data and a fixed time-window of 6 milliseconds), the effect of including an additional coincidence window for the central frequency of the events.

The right-hand plot of Figure 5.8 shows the improvement in the significance for different sizes of the frequency coincidence window, compared to the case where only a coincidence window for the time of the events is applied. In this case, for all sizes of the frequency coincidence window, the significance is improved. A maximal improvement of about 50 % is achieved for a frequency window of 200 Hz.

If we choose windows, $t_{\mathrm{win}} = 6\,\mathrm{msec}$ and $f_{\mathrm{win}} = 200\,\mathrm{Hz}$, we get an efficiency of 1.5 %, a background of 0.05 %, a use-percentage of 0.6 % and a significance of 30. This example demonstrates the usefulness of additional coincidence windows. However, as the actual performance of this veto is not very impressive in itself, the veto should rather be viewed as an illustrative example.

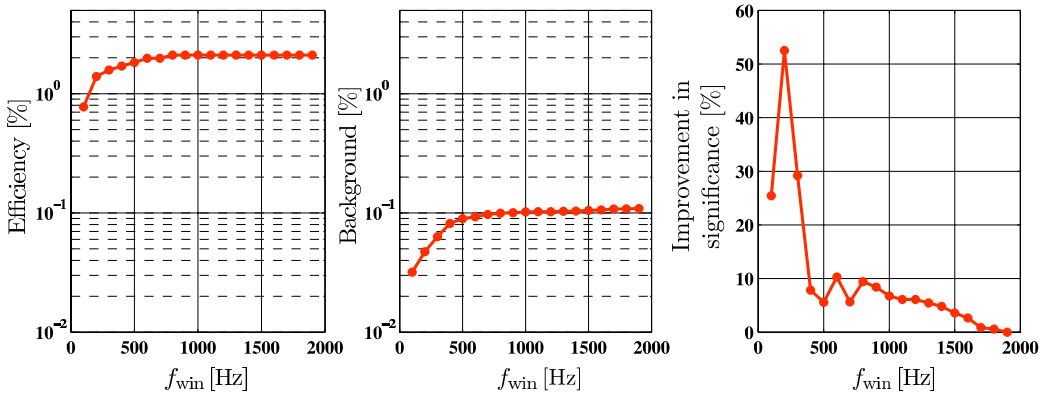


Figure 5.8.: Exemplary application of a statistical veto including an additional coincidence condition for the central frequency of the events. The same data set as for Figure 5.7 is used. A time coincidence window of 6 milliseconds is applied. The background and efficiency are plotted versus the size of the coincidence window for the central frequency. In addition, the right hand plot shows the improvement in significance versus the size of the frequency window. The improvement of significance is the relative improvement compared to the case where no frequency coincidence window is applied.

5.6. A statistical veto for channels containing GW information

In the previous section, we described the statistical veto method using an auxiliary channel containing no gravitational wave information. In this section we will show that, under certain conditions, and when handled with care, a reliable statistical veto can also be based on a veto channel which can contain gravitational wave information.

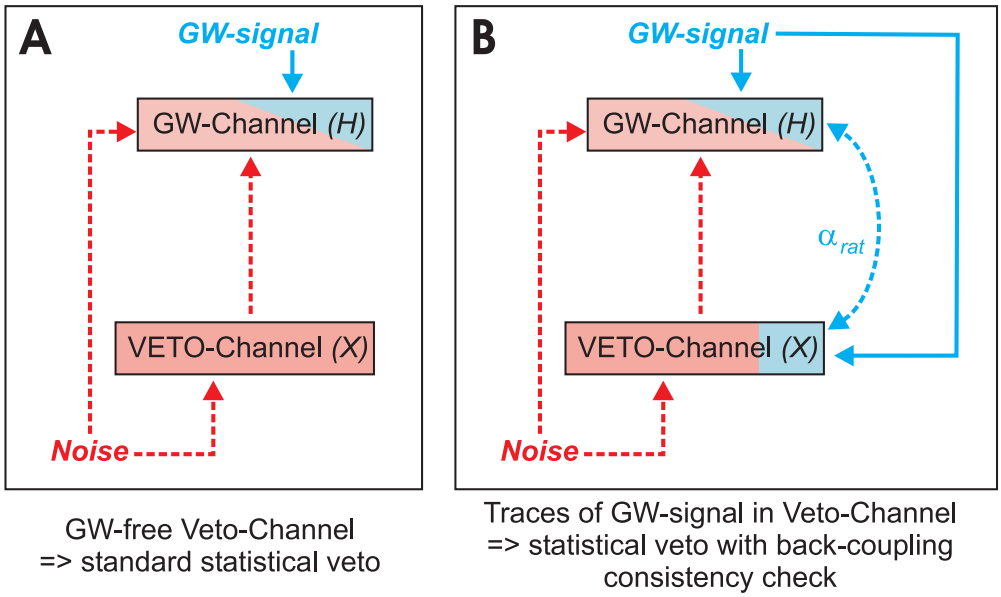


Figure 5.9.: A schematic view of two different scenarios in which coincident transient events can appear in both the main GW channel, H, and an auxilliary channel, X. The first scenario (Panel A) shows coincident events arising from a single noise source, which couples events to H either directly, or via subsystem X, or both. The second scenario (Panel B) shows the case where the events in X can originate, not only from the noise source, but also from the GW signal. The frequency dependent amplitude ratio of a GW-like event measured in both channels, α_{rat}, can be determined by injecting differential arm length noise (to mimic the effect of an GW signal) and measuring the transfer function from H to X.

Figure 5.9 shows two scenarios. Panel A describes the case where there is an (unknown) coupling of a noise source, N, into both the main GW channel, and an auxiliary channel, X; for this case, the standard statistical veto described in Section 5.5 can be applied. Panel B shows the case where the events in X can originate, not only from the

noise source, but also from the GW signal. In order to apply a statistical veto (with a sufficiently low false-veto rate), in such a situation, we have to introduce further coincidence conditions, such as a frequency-dependent amplitude cut. The application of this method requires a rough knowledge of the amplitude ratio, α_{rat}, and its stability.

5.6.1. The method in general

In the presence of gravitational wave signal in the veto channel, we will have two classes of coincidence events to consider when applying a standard statistical veto:

- Noise events that couple via the (red) dashed lines of Figure 5.9. These are the events we would like to veto.
- Gravitational-wave-like events that are showing up in the GW channel as well as in the veto channel.

It is essential to discriminate these two populations and exclude events from the latter class from being vetoed.

A possible way to do this is to compare the amplitudes $a^H[i]$ and $a^X[j]$ of the two coincident triggers $H[i]$ and $X[j]$. In the case that the event $X[j]$ originates from the same GW-like event as the event, $H[i]$, the following (ideal) amplitude ratio can be computed:

$$\frac{a^X[j]}{a^H[i]} = |\alpha_{\mathrm{rat}}[i]|, \tag{5.9}$$

where $|\alpha_{\mathrm{rat}}[i]|$ is amplitude ratio of the GW signal in H and $\mathcal{P}_{\mathrm{DC}}$, α_{rat}, evaluated at the central frequency of the event, $H[i]$, in the gravitational wave channel.

If all quantities in Equation 5.9 are known, the application of the statistical veto is rather simple. First of all, the coincident events need to be determined in the same way as for the standard statistical veto. Secondly, for each pair of events, the ratio of the two amplitudes is compared. If this ratio equals the magnitude of the amplitude ratio, α_{rat}, at the central frequency of the event, the event in the gravitational wave channel is not vetoed. If the amplitude ratio is not consistent with the α_{rat}, the event $H[i]$ is vetoed.

It is also possible that more than one event in the veto channel is coincident with the event $H[i]$. In this case, the H event is excluded from being vetoed if at least one of the

coincident pairs satisfies Equation 5.9.

5.6.2. A 'real-world' scenario

In reality, Equation 5.9 needs to be extended to account for several systematic and statistical errors. Probably the largest contribution to the error originates from the amplitude estimation of the burst events. The errors associated with the amplitude estimation of the events $H[i]$ and $X[j]$ are referred to as $\Delta a^H[i]$ and $\Delta a^X[j]$, respectively. Also, the measurement of the amplitude, $\alpha_{\rm rat}$, can be a source of error. More over, the amplitude ratio can also be non-stationary over time. We represent the cumulative errors due to these two effects by $\Delta\alpha_{\rm rat}$.

Considering these errors, we make a simple generalisation of Equation 5.9 into the 'real-life' situation. In order to veto an event, $H[i]$, in the gravitational wave channel, we first require that it be coincident with the another event, $X[j]$, in the veto channel in the sense of a standard statistical veto (in time and frequency), and secondly that it satisfy either of the following conditions:

$$\frac{a^X[j]}{a^H[i]} < \frac{|\alpha_{\rm rat}[i]|}{(1+\Delta a_{\rm tot})}, \tag{5.10}$$

or

$$\frac{a^X[j]}{a^H[i]} > |\alpha_{\rm rat}[i]|\,(1+\Delta a_{\rm tot}) \tag{5.11}$$

where $\Delta a_{\rm tot}$ is an upper-limit of the cumulative error from the amplitude estimation and the amplitude ratio measurement (also due to the non-stationarity of the amplitude ratio over time).

We will call the veto method based on the condition described above, a *frequency-dependent amplitude cut*, because we cut the events with a certain amplitude ratio out of the list of coincident vetoed events.

5.7. Application of a statistical veto with a amplitude consistency-check to GEO 600 data

In the following three subsections we will present the results of applying a statistical veto with a amplitude consistency check to GEO 600 data using $\mathcal{P}_{\rm DC}$ as the veto channel.

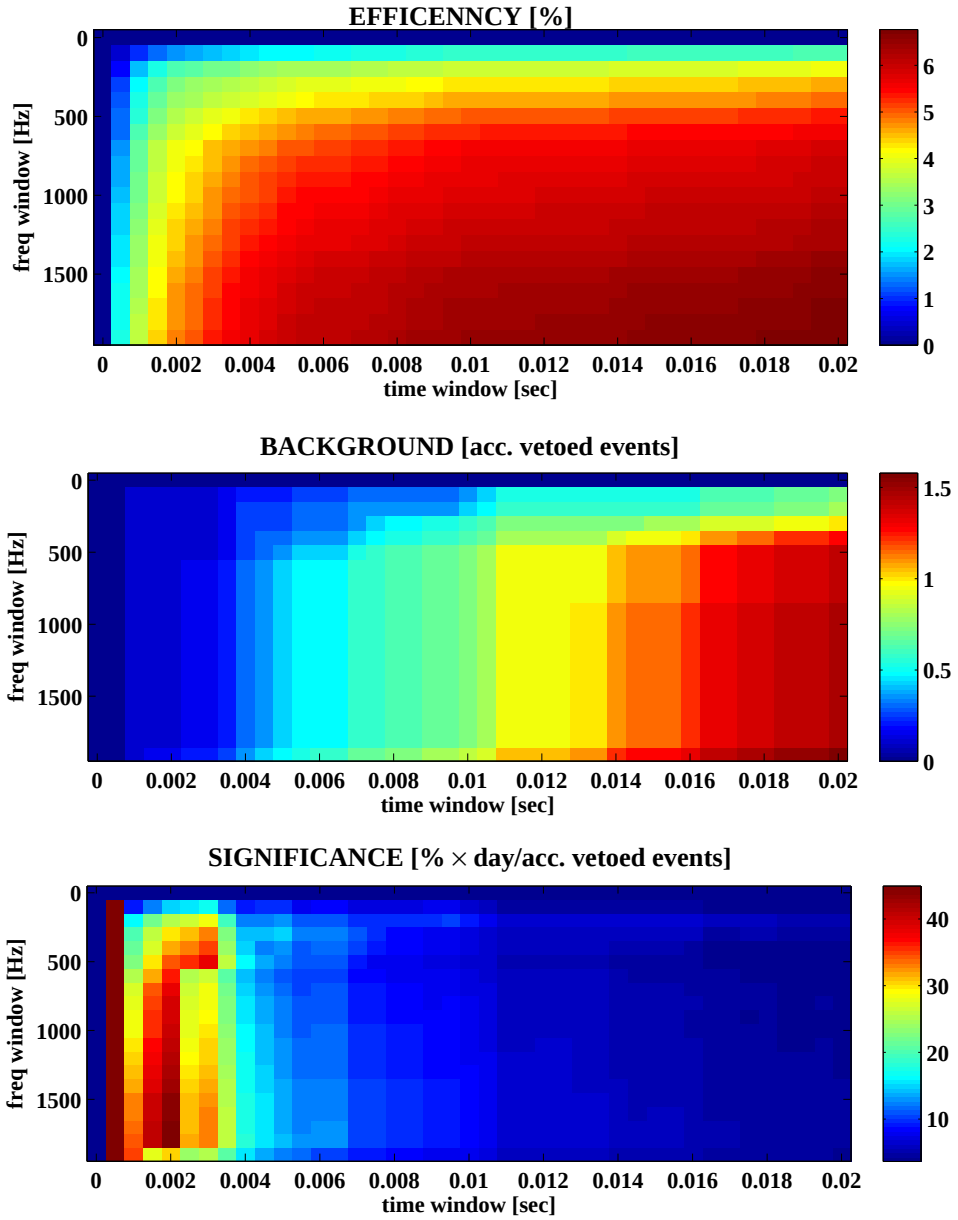


Figure 5.10.: Efficiency, background and Significance of a veto using $\mathcal{P}_{\mathrm{DC}}$ as veto channel for various sizes of the time and frequency coincidence window. To reduce computational resources required for this analysis a data stretch of 24 hours from September 2006 was used instead of the full month. This analysis is done to find good sizes for the coincidence windows. At this stage of the analysis no amplitude consistency-check is applied.

Two different periods of time are analysed. The analysis of a long stretch of data covering the entire month of September 2006 (where the dust concentration in the clean-room was nominal), is presented in subsection 5.7.1. An 8 hour stretch of science data from May 2006 (where a high dust concentration in the clean room was observed), is analysed in subsection 5.7.2. There are only short data stretches available for this period because the broken air conditioning system was fixed within a few days. However, even with the nominal dust concentration restored, still, dust glitches contribute to the glitch rate of H. Finally, the performance of the veto for both periods is compared in subsection 5.7.3.

5.7.1. Data set 1: Full September 2006 with low dust concentration

First we have to determine reasonable sizes for the time and frequency coincidence windows. This was done by computing efficiency, background and significance of the veto for various window sizes for a subset of 24 hours of data from September 2006. The result of this analysis is shown in Figure 5.10. The best significance can be achieved for very small time windows. However for these small time windows the efficiency is very small. That is why we have to trade off significance and efficiency. For a time window of 8 milliseconds and a frequency window of 1000 Hz we achieve a best efficiency for an acceptable background rate of about 0.5 events per day.

For both data sets, and the hardware injections, firstly a coincidence analysis is performed using a time coincidence window of 8 milliseconds and a frequency coincidence window of 1 kHz determined by the analysis described above. The upper plot of Figure 5.12 shows the ratio of the amplitude of the coincident events from $\mathcal{P}_{\mathrm{DC}}$ and H versus the central frequency of the event in H. The (blue) diamonds, corresponding to the GW-like hardware injections, are close to the measured amplitude ratio, α_{rat} (as expected); all the coincident events from the first data set (represented by the purple points), show either a similar amplitude ratio or a higher one. If the ratio is similar to the hardware injections, the points most probably correspond to GW-like events. If the amplitude ratio is higher, this means that $\mathcal{P}_{\mathrm{DC}}$ events show a higher amplitude than is consistent with events originating from a GW-like event. As we do not observe any pairs of events with an amplitude ratio much lower than from the hardware injections,

it is reasonable to simplify Equation 5.11 to the single condition,

$$\frac{a^X[j]}{a^H[i]} > |\alpha_{\mathrm{rat}}[i]|\,(1 + \Delta a_{\mathrm{tot}}). \tag{5.12}$$

Over one month, α_{rat} was measured a few times and $\Delta\alpha_{\mathrm{rat}}$ was found to be less than 0.5.

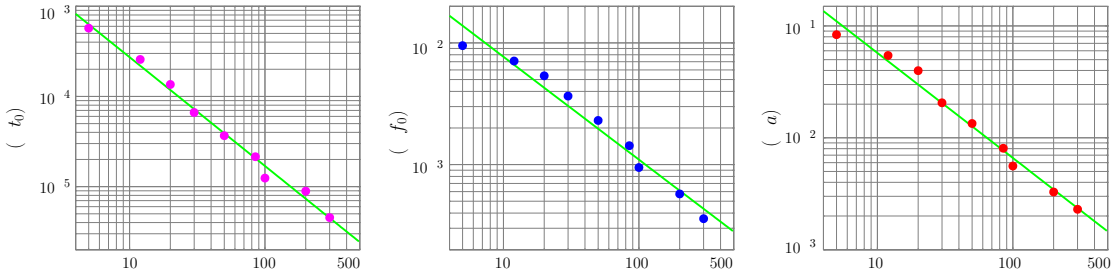


Figure 5.11.: Standard deviation of the distribution of the mHACR errors (in estimating the parameters of injected sine-Gaussian waveforms into gaussian noise) plotted as a function of the snr of the triggers. Also shown are linear-polynomial fits to the data. This plot is taken from [Hild07a] and was produced by P. Ajith.

The main error contribution of the amplitude estimation of the events can be described by three times the standard deviation (3σ) given in the right hand plot of Figure 5.11. The lowest SNR of an event contained in this analysis is about 4, which means that the maximum error in estimating the amplitude of an event is about 60 % (see Figure 5.11). Since the values in Figure 5.11 are estimated from ideal conditions [2], we will allow for a 200 % error in the amplitude estimation to get a safe upper limit of the error. All together, Δa_{tot} amounts to 2.

In the end we get the final set of three veto conditions:

$$\left|t_0^X[j] - t_0^H[i]\right| < 8\,\mathrm{ms}, \tag{5.13}$$

$$\left|f_0^X[j] - f_0^H[i]\right| < 1\mathrm{kHz}, \tag{5.14}$$

[2] The errors in parameter estimation given by a particular detection algorithm depend on many things, for example, the waveform of the transient signal and the characteristics of the underlying data-stream (noise, spectral lines, *etc*).

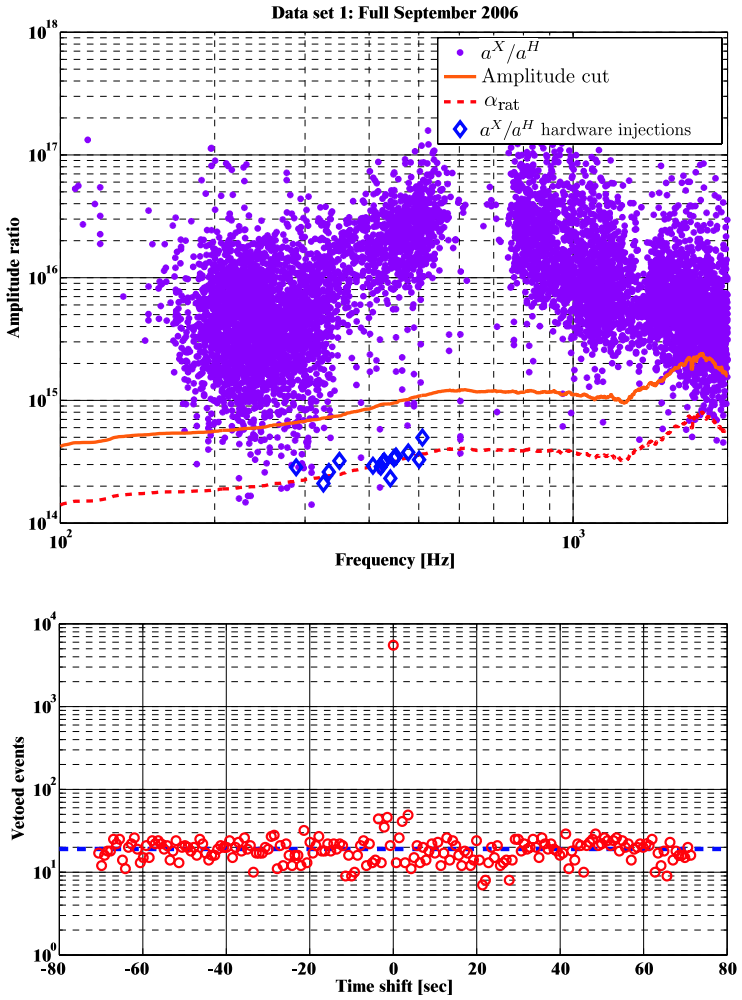


Figure 5.12.: *Upper plot*: Application of a statistical veto with a amplitude consistency check to GEO 600 data for the entire month of September 2006. The (red) dashed line is a measurement of the amplitude ratio, $\alpha_{\rm rat}$. The (blue) diamonds are the amplitude ratios of the coincident events in $\mathcal{P}_{\rm DC}$ and H from GW-like burst hardware injections. The hardware injections are consistent with $\alpha_{\rm rat}$. The solid (orange) line is the chosen amplitude cut, corresponding to $\Delta a_{\rm tot} = 2$. The (purple) points indicate the amplitude ratio of the coincident events from $\mathcal{P}_{\rm DC}$ and H for September. Each H event corresponding to a point above the solid line is vetoed, while each point below the solid line is taken as being consistent with a GW signal and is excluded from being vetoed. *Lower plot*: A time-shift analysis of the statistical veto with a amplitude consistency check for data set 1. 5517 events in H are vetoed, while the background of accidentally-vetoed events amounts to 19.1 events per month.

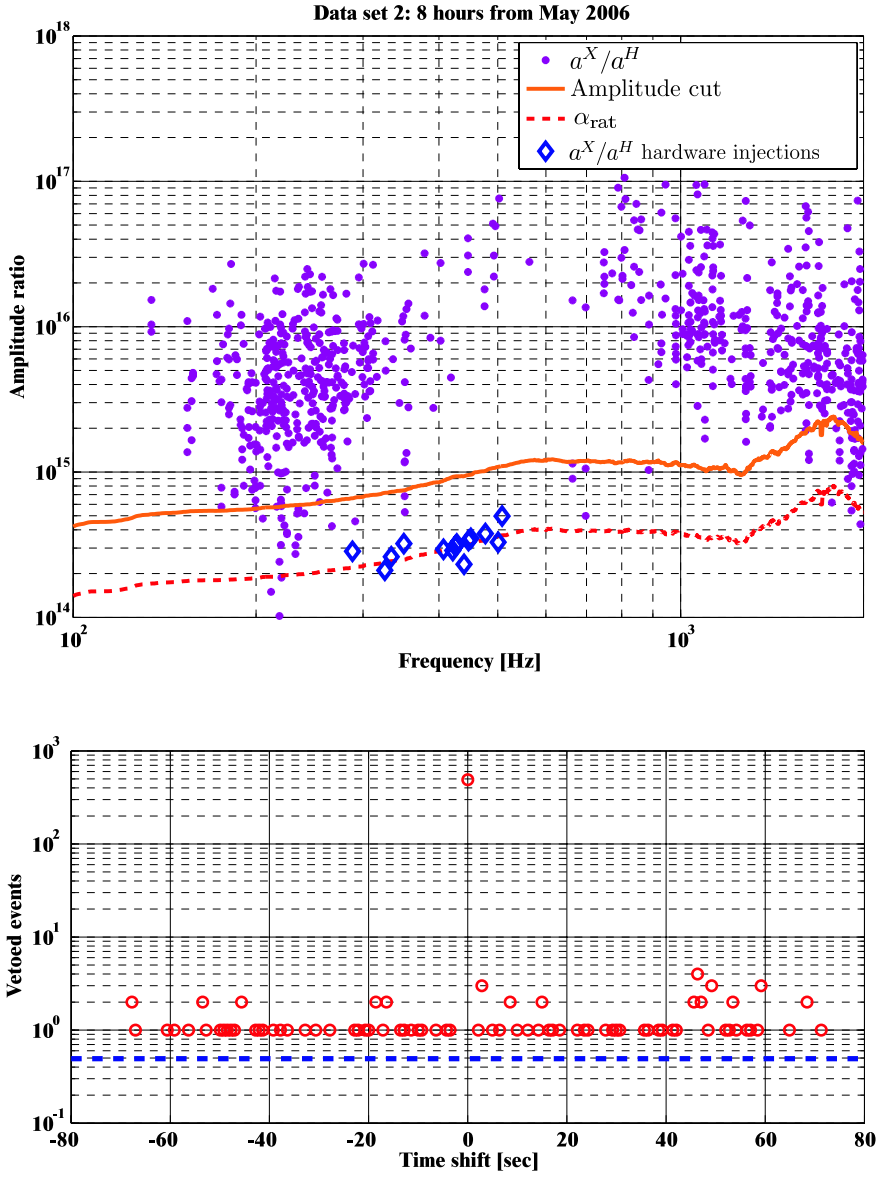


Figure 5.13.: *Upper plot*: The application of a statistical veto with a amplitude consistency check to 8 hours of GEO data from May. For further details, please see Figure 5.12. *Lower plot*: A time-shift analysis of the statistical veto with a amplitude consistency check for data set 2. 291 events in H are vetoed, while the background of accidentally vetoed events amounts to 0.49 events per 8 hours.

and

$$\frac{a^X[j]}{a^H[i]} > 3\,\alpha_{\mathrm{rat}}[i]. \tag{5.15}$$

The last condition can be seen as an amplitude cut. The solid (orange) line in Figure 5.12 indicates the level of this amplitude cut. Each H event corresponding to a (purple) point above the solid line is vetoed, while each point below the solid line is taken as being consistent with a potential GW signal and is excluded from being vetoed. By introducing the amplitude cut the veto efficiency is reduced from 5.94 to 5.72 %. The background rate of this veto is estimated by time-shifting the H events and is indicated by the (blue) dashed line in the lower plot of Figure 5.12.

5.7.2. Data set 2: 8 hours from May 2006 with high dust concentration

Data set 2 is from a time with a high dust concentration in the main clean room. For the analysis, identical veto conditions are applied as for data set 1. The result of the veto application is shown in Figure 5.13.

For this set of data, a high veto efficiency of greater than 20 % is obtained. The background rate of this veto, estimated from a time-shift analysis, is indicated by the (blue) dashed line in the lower subplot of Figure 5.13.

5.7.3. Performance of the veto analysis

The results of applying a statistical veto based on $\mathcal{P}_{\mathrm{DC}}$, together with a amplitude consistency check, are summarised in Table 5.1.

Data Set	1	2
Efficiency [%]	5.72	21.5
Background [%]	0.02	0.02
Significance	286	1075
Use-percentage [%]	20.7	79.8

Table 5.1.: A summary of the results of applying a statistical veto with a amplitude consistency check to two different data sets from GEO 600.

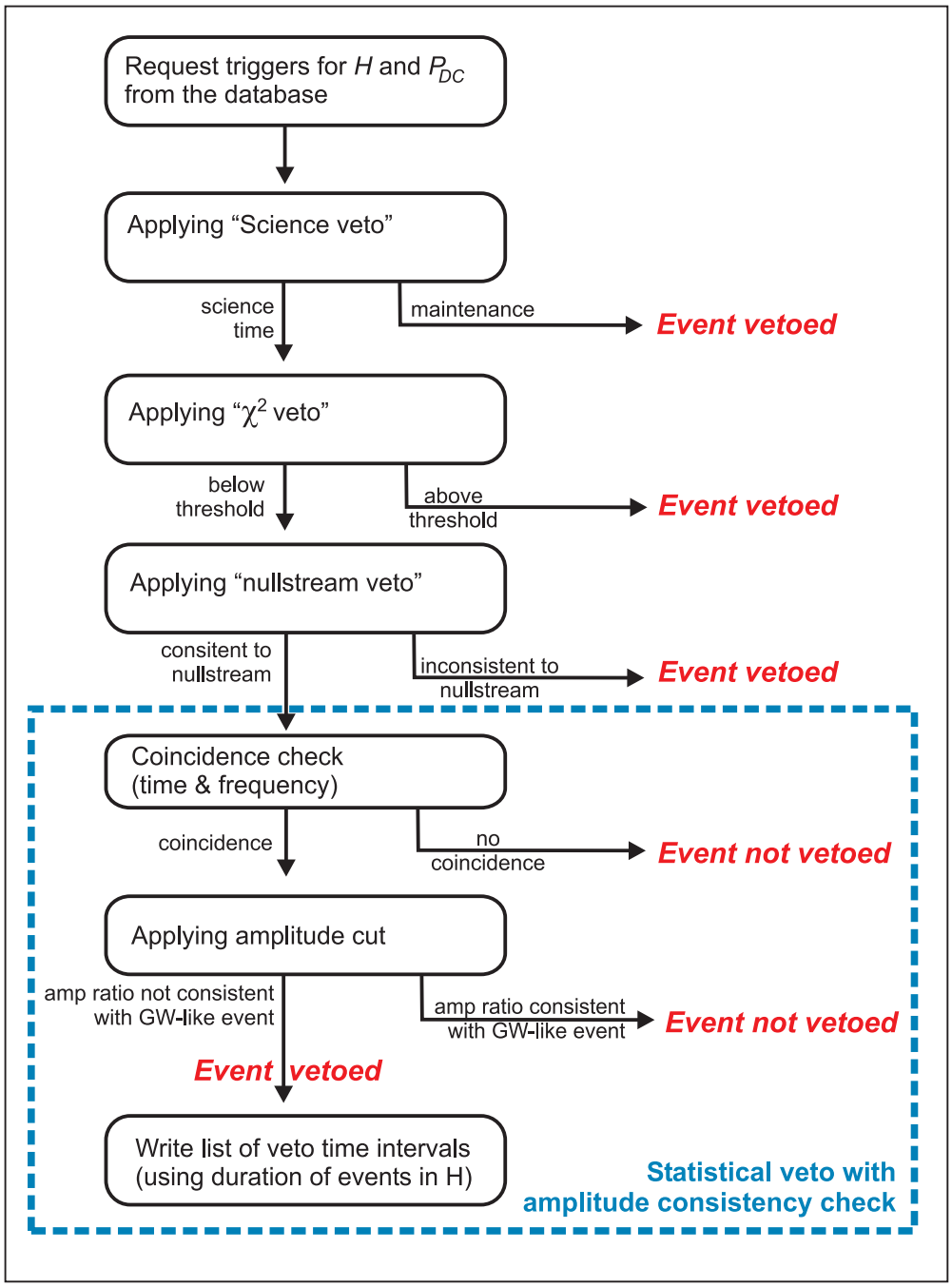


Figure 5.14.: Schematic of the veto pipeline used for the analysis of Data Set 1 (full September 2006). Before the statistical veto three other vetoes (science, χ^2 and null stream) are applied. The output of the whole pipeline is a list of veto intervals.

Figure 5.14 gives a schematic overview of veto analysis performed for the analysis of full September 2006. The corresponding Matlab script can be found in Appendix G.

5.8. Summary

It was shown that the performance of a standard statistical veto can be improved by application of more than one coincidence window. Furthermore, a new veto method was developed which allows the use of veto channels which can contain GW signal. By introducing an amplitude consistency check, safe statistical vetoes can be derived from interferometer channels. GW-like hardware injections have been performed to demonstrate the robustness of this veto method. Application of a statistical veto with amplitude consistency check to data from the GEO 600 detector was shown to perform well, giving a veto efficiency of up to 20 % and a use-percentage of up to 80 %. This new method is generally applicable and can also be used on the data from other GW detectors.

Chapter 6.

Towards DC-readout for GEO 600

6.1. Introduction

All of the currently running large-scale gravitational wave detectors are operated at the dark fringe and use a heterodyne readout. For a Michelson interferometer without signal recycling the choice of this operating point has two predominant advantages: First of all, the couplings of some technical noise, for example laser power noise, are suppressed. Secondly, the power recycling technique can easily be implemented.

However, the experience in the currently running interferometers reveal several problems connected to a heterodyne readout:

- As it will be shown in Section 6.2 a heterodyne readout always implies an increased shot noise level with respect to a homodyne readout.
- Phase noise of the applied modulation would probably limit the sensitivity of advanced detectors.
- Since the Michelson RF sidebands propagate differently through the interferometer than the carrier light the spatial overlap of the GW signal (carrier) and the local oscillator (Michelson sideband) might be imperfect.
- In an interferometer with detuned signal recycling the presence of imbalanced RF sidebands at the output port increases and complicates the technical noise contributions [Hild07b].

A way to get around these problems is to choose an operating point slightly off the dark fringe and to use a DC-readout (homodyne) scheme.

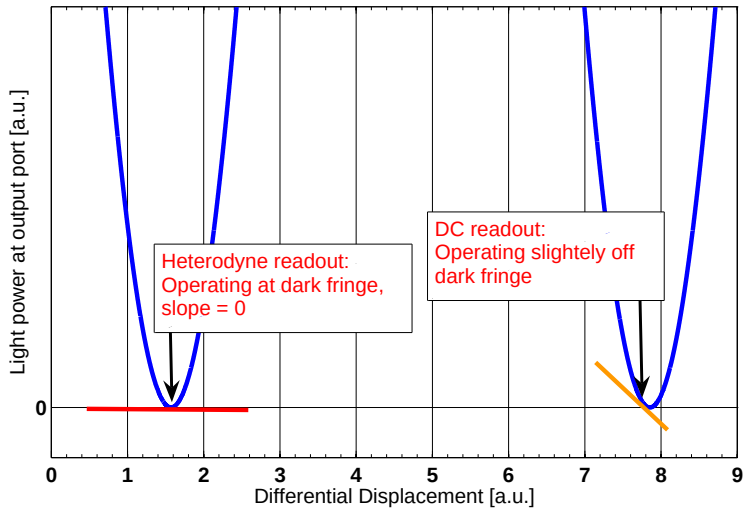


Figure 6.1.: Overview of the operating points of a heterodyne and DC-readout scheme. Currently GEO 600 uses heterodyne readout and is operated at the dark fringe. A DC-readout scheme uses an operating point slightly off the dark fringe. While in the heterodyne case RF sidebands are used as local oscillator, in the DC-readout scheme a fraction of the carrier light can leave the interferometer at the output port and serve as local oscillator.

Figure 6.1 schematically depicts the two different readout schemes. The currently in GEO 600 used heterodyne readout is shown on the left side: The Michelson interferometer is operated at the dark fringe, the slope of the light power is 0, and heterodyne sidebands at RF are used for the readout in order to produce a bipolar control signal. On the right side the situation used for DC-readout is shown. The operating point is slightly off the dark fringe, thus a fraction of the carrier light can leave the interferometer and serve, due to the large power recycling gain, as a highly stabilized local oscillator at the photodiode. Since the slope of the output light level versus differential arm length is not zero any more a bipolar control signal can be obtained from the level of DC light.

Due to the advantages of DC-readout with respect to heterodyne readout, for advanced LIGO [AdvLIGO] a DC-readout scheme will be used. Also for GEO 600 there are good reasons to change over to a DC-readout scheme as will be shown in the following sections.

6.2. Motivation

In the next few years LIGO as well as VIRGO are going to improve their detectors. Both plan for precursor projects, called enhanced LIGO [Adhikari06] and VIRGO+ with moderate sensitivity improvements before going to advanced LIGO [AdvLIGO], [Giaime] and advanced VIRGO [AdvVIRGO]. As shown in Figure 1.1 the current sensitivity of the only 600 m long GEO 600 detector is competitive to the much longer detectors only at frequencies above 500 Hz. Therefore in the near future any sensitivity improvement of GEO 600 at high frequencies is probably going to be more valuable than an improvement at low frequencies.[1]

Figure 6.2 shows the S5 noise projections of the GEO 600 detector [Smith06]. Above 400 Hz the uncorrelated sum of the individual noise contributions is mainly dominated by the modelled shot noise[2]. The first step in order to improve the sensitivity of GEO 600 at

[1] This statement is of course build on many assumptions, which might change in future. However, it bases on our currently best knowledge.

[2] Due to the asymmetric sidebands present at the output of a detuned signal recycled interferometer, the calculation of the shot noise limit of GEO 600 is more complicated than for a detector without signal recycling or tuned signal recycling and is therefore still a field of research. However, assuming completely unbalanced Michelson sidebands we can calculate an upper limit of the shot noise for the heterodyne readout. It was estimated by using the *pdS2*-command of FINESSE. The outcome of the FINESSE simulation was then multiplied by two correction factors: First a factor $1/\sqrt{2}$, accounting

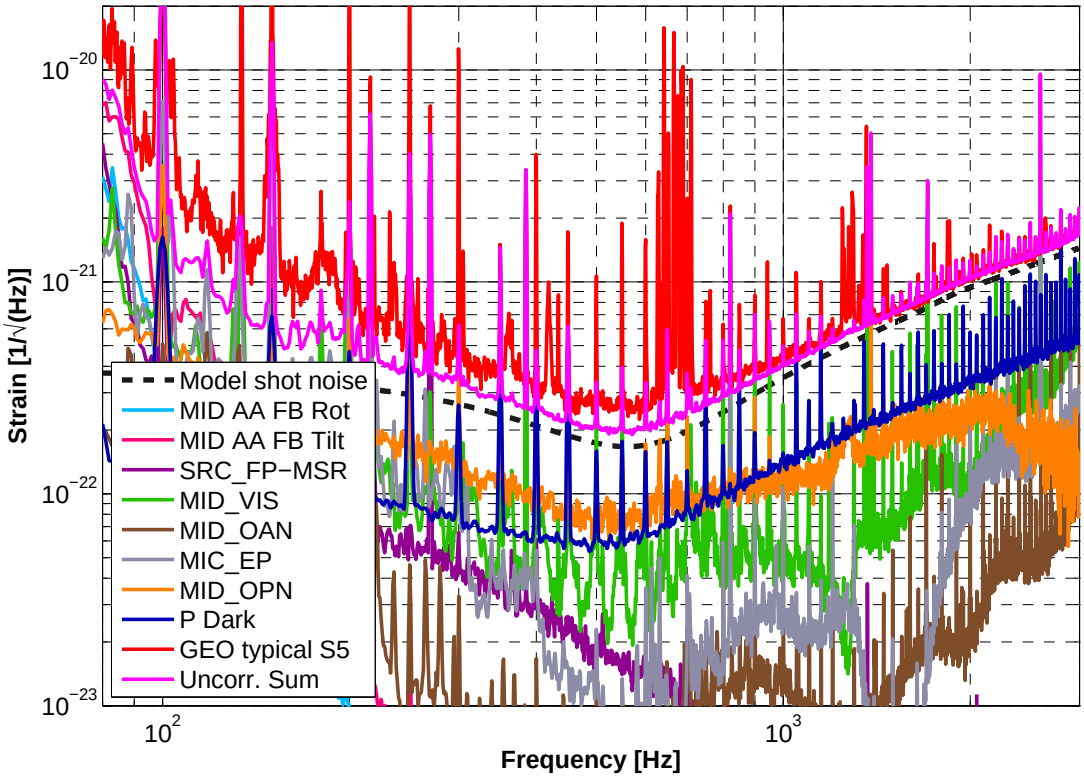


Figure 6.2.: Noise projections for a time of S5. The pink curve represents the uncorrelated sum of the individual noise contributions. The red trace shows a snapshot of the actual sensitivity. Above 400 Hz the uncorrelated sum is mainly dominated by the modelled shot noise (explanation is given in the text). The near-future plan is to increase the circulating light power by about a factor 4 in order to decrease shot noise by a factor of 2. Then the sensitivity between 500 and 2000 Hz would be limited to roughly equal shares by shot noise, oscillator phase noise (MID_OPN) and electronic noise of the main photodiode (P Dark). All three of these noise contributions will potentially be decreased by changing over from a heterodyne to a DC-readout scheme.

frequencies above 500 Hz is to decrease the shot noise contribution by means of increasing the circulating light power. As soon as the light power is increased by a factor of 4, which would bring the shot noise (black dashed line in Figure 6.2) down by a factor of 2, the sensitivity of GEO 600 would be limited above 500 Hz by roughly equal shares of three noise sources:

- Shot noise
- Oscillator phase noise of the modulation used for creating control signals for the differential arm length of the Michelson interferometer and the gravitational wave readout (MID_OPN, orange trace in Figure 6.2).
- Electronic noise of the main photo diode (P Dark, blue trace in Figure 6.2).

If we want to increase the sensitivity of GEO 600 by more than a factor of two at frequencies above 500 Hz we have to attack all three of the noises listed above. The concept of DC-readout implies the wonderful chance to reduce all three of these noises at once.

6.2.1. Shot noise

When going from the currently used heterodyne read out scheme to a DC-readout scheme the signal to shot noise ratio will be increased by a factor between $\sqrt{1.5}$ and $\sqrt{2}$ [Buonanno03], [Harms06]. This depends on the balancing of the two Michelson sidebands at the dark-port. The sensitivity will be increased by between a factor of $\sqrt{1.5}$ for balanced sidebands and a factor of $\sqrt{2}$ for completely unbalanced sidebands. This can in an easy picture be understood by looking at the individual shot noise contributions at the dark-port. Amongst others we find in the heterodyne case shot noise contributions from frequencies, 2ω, corresponding to two times the heterodyne modulation frequencies (see Appendix B). It is worth noting that this improvement in sensitivity is achieved by only changing the readout scheme and does not require any increase of light power.

for the different propagation of signal and shot noise in a FINESSE-mixer (for a detailed description see [Finesse], page 61). The second factor, $\sqrt{2}$, accounts for additional shot noise in a heterodyne readout with respect to a homodyne readout [Buonanno03], [Harms06].

6.2.2. Oscillator phase noise

The oscillator phase noise is a serious problem for all heterodyne readout techniques. Even though this noise is usually classified as a technical noise source we are now already close to the limits, even though a low phase noise crystal oscillator is currently used in GEO 600. One way to reduce the oscillator phase contribution would be to reduce the coupling factor by going to tuned signal recycling (see Section 2.6). However, when going from a heterodyne to a DC-readout scheme it would be possible to derive the GW signal without an RF demodulation of the main output photodetector. By that we would completely eliminate any oscillator phase noise contribution.[3]

6.2.3. Electronic noise of the main photo diode

The electronic noise of the main photodiode can also be reduced by changing over to DC-readout. At the moment the diode is required to have a large dynamic range and to be capable of detecting large RF signals at the same time. This requirement restricts the maximal possible conversion factor for the photo current to voltage transformation [Grote]. However, as in a DC-readout scheme the photodiode does not have to detect RF signals, the electronic noise of the photodiode can be significantly reduced by increasing the photocurrent to voltage conversion factor and an additional gain increase of the first amplification stage.

6.3. Determination of the optimal dark fringe offset

The most important question we have to answer is, what is the optimal offset from the dark fringe for a DC-readout scheme for GEO 600? The dark fringe offset, ξ_{df}, is defined as:

$$\xi_{\mathrm{df}} = \frac{|\Delta L_{\mathrm{n}}| + |\Delta L_{\mathrm{e}}|}{2}, \tag{6.1}$$

where $\Delta L_{\mathrm{e}} = -\Delta L_{\mathrm{n}}$ are the intentionally added offsets in the distance between the beam splitter and the end mirrors MCE and MCN, respectively. In the following, FINESSE simulations [Freise04] are used to find the optimal dark fringe offset. The

[3] It is a noteworthy fact, that in recent times the relative stability of the light in the interferometer is better than the relative phase noise achievable with excellent RF technique.

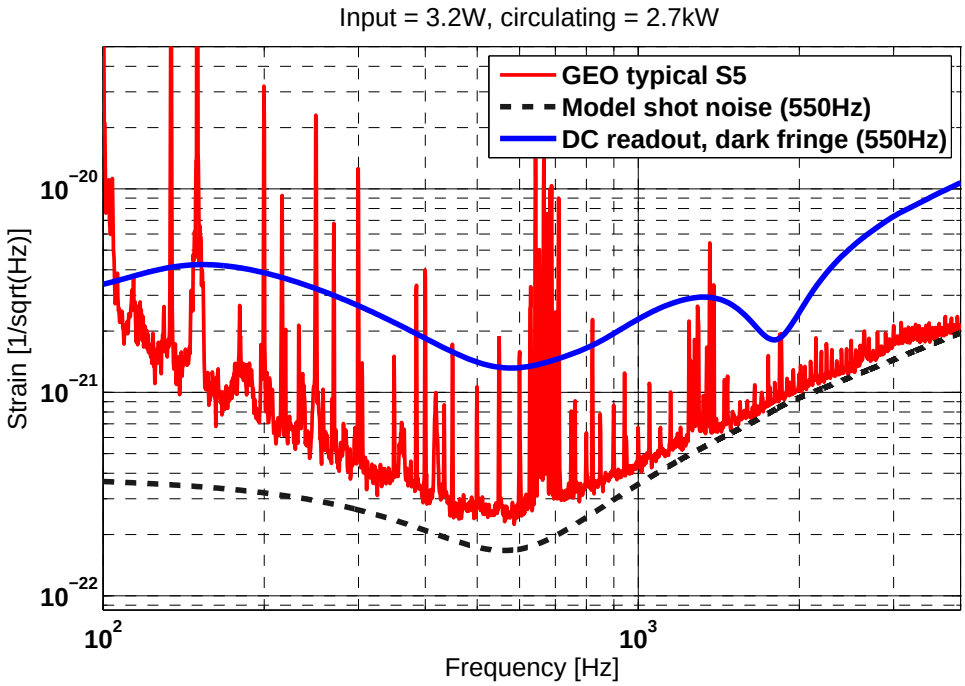


Figure 6.3.: Heterodyne versus DC-readout for GEO 600 in S5 configuration: The red trace shows the typical sensitivity of GEO 600 of S5. The blue trace corresponds to the shot noise limited sensitivity for using the DC light level of main photo detector, $\mathcal{P}_{\mathrm{DC}}$. For this simulation the nominal modulation indices of S5 have been used and the operating point was chosen to be the dark fringe. The black dashed curve represents the modelled shot noise.

first simulation that was carried out is shown in Figure 6.3. The blue line represents the sensitivity that can in principle be obtained in the current configuration of GEO 600 (nominal S5 parameters) by using the DC light of the main photo detector. As already presented in Chapter 5.3 the DC light of the main photo detector, $\mathcal{P}_{\mathrm{DC}}$ is sensitive to gravitational wave signal as well. This originates from the fact that for the Michelson modulation sidebands the Michelson interferometer is not at the dark fringe, i.e. the sideband strength at the output port is to some extent proportional to the differential arm length of the interferometer. With respect to the actual sensitivity of GEO 600 (red trace in Figure 6.3), the blue trace shows a maximal sensitivity around 1.8 kHz. The bump around 1.8 kHz can also be found in the measurement of the transfer function from differential arm length to $\mathcal{P}_{\mathrm{DC}}$ shown in Figure 5.12.

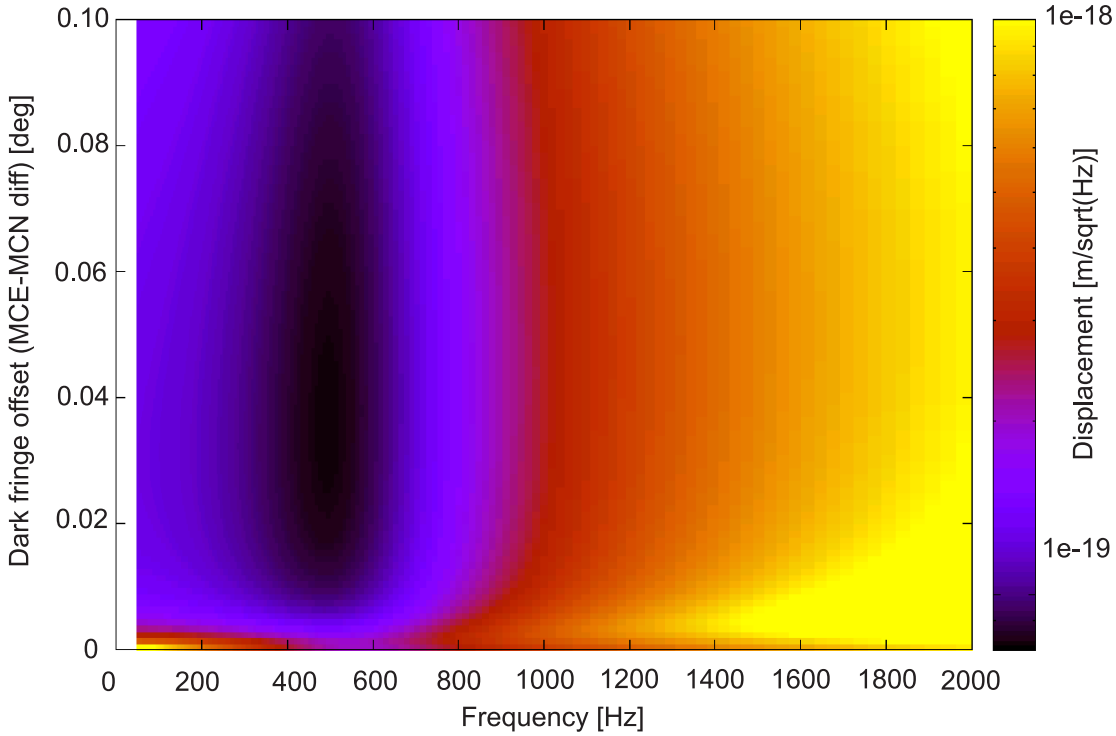


Figure 6.4.: Simulated displacement sensitivity of GEO 600 with DC-readout versus the offset from the dark fringe, ξ_{df}. The offset from the dark fringe was realized in this simulation by adding a differential phase offset to MCE and MCN. The modulation index of the Michelson sidebands was set to 5 % of the nominal S5 value. The best peak sensitivity is found for ξ_{df}= 0.036 deg.

Currently a modulation index for the Michelson modulation of 0.38 is used. For the DC-readout scheme it is desirable to get rid of all RF contributions at the output port light field in order to not spoil the measurement. The strength of the sidebands at the dark port is planned to be decreased in two steps. First it seems reasonable to decrease the applied modulation by a factor of two and still get reasonable control signals. In a second step it is planned to insert an output mode cleaner (OMC) into the main detection path which is assumed to suppress the Michelson sidebands by another factor of ten. Overall this would give a reduction by a factor of 20 of the Michelson sidebands at the main photo detector. In the following simulations this sideband suppression is realized by turning down the modulation in front of the interferometer from the nominal S5 value of 0.38 to 0.019[4].

With the reduced Michelson sideband strength we can now find the optimal offset from the dark fringe. Figure 6.4 shows the simulated displacement sensitivity versus the dark fringe offset, ξ_{df}. The offset from the dark fringe was implemented in this simulation by adding a differential phase offset (given in degree) to the end mirrors of the Michelson interferometer, MCE and MCN. ξ_{df} given in degrees can be converted into ξ_{df} given in meters by

$$\xi_{\mathrm{df}}\,[\mathrm{m}] = \frac{\lambda}{360\,\mathrm{deg}}\,\xi_{\mathrm{df}}\,[\mathrm{deg}], \tag{6.2}$$

where $\lambda = 1064\,\mathrm{nm}$ is the wavelength of the laser. The optimal dark fringe offset giving the best peak sensitivity is found for a ξ_{df} of 0.036 degrees, which is equivalent to ξ_{df}= $1.06 \cdot 10^{-10}\,\mathrm{m}$. The slope of the sensitivity improvement is found to be very steep for small offsets from the dark fringe, $\xi_{\mathrm{df}} < 0.01\,\mathrm{deg}$, and is nearly flat for larger offsets. That is why already for half the optimal dark fringe offset, $\xi_{\mathrm{df}} = 0.018$ degrees, which is equivalent to a dark fringe offset of about 50 pm, the achieved sensitivity is nearly optimal[5]. The FINESSE input file used for this simulation can be found in Appendix H.

Since many potential problems connected to a DC-readout scheme, such as noise couplings, increase with the dark fringe offset, a smaller offset, giving nearly the same sensitivity, is actually preferable. Figure 6.5 shows the spectra of the shot noise limit

[4] In contrast to an output mode cleaner this scenario gives a slightly increased circulating light power inside the interferometer, as now ratio of carrier and Michelson sidebands is increased. However, this effect influences the following simulations only on a scale of a few percent.

[5] This relation is also illustrated in a more obvious way in Figure 6.5.

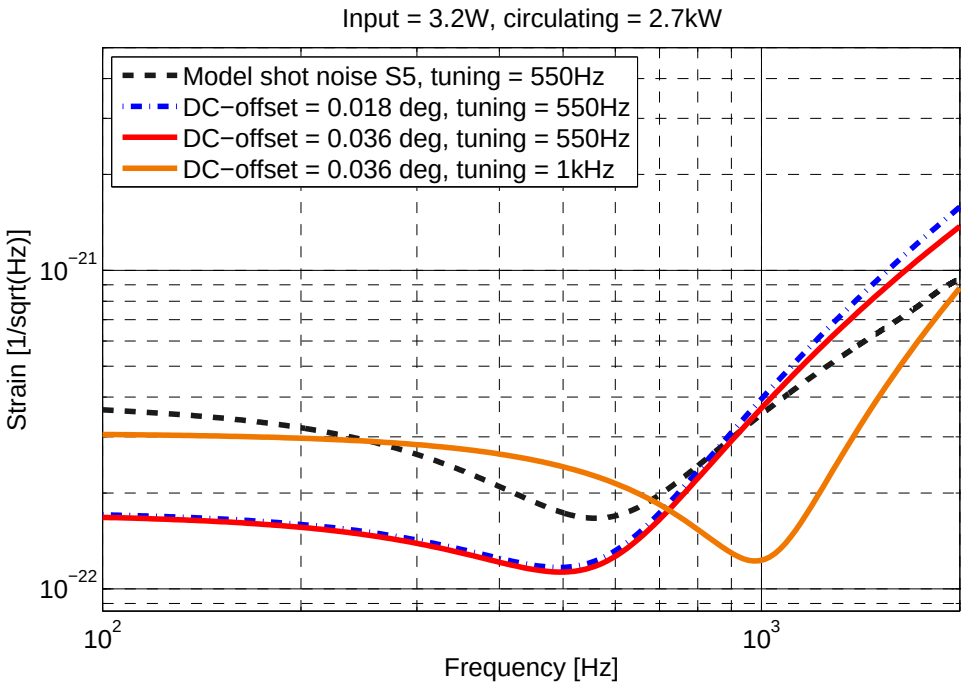


Figure 6.5.: Shot noise limited sensitivity achievable with DC-readout versus the modelled shot noise of S5: The red and the blue dashed-dotted trace represent the sensitivity with a DC-readout scheme for two different ξ_{df}. Even though an optimal ξ_{df}= 0.036 deg was determined from Figure 6.4, half of that dark fringe offset (ξ_{df}= 0.018 deg) already gives a sensitivity only marginally worse. An improvement of the peak sensitivity by about a factor 1.5 is achievable with respect to the modelled shot noise of S5. For the currently used input power of 3.2 W a peak sensitivity of about $1.1 \cdot 10^{-22}/\sqrt{\mathrm{Hz}}$ is obtained. The orange trace shows the sensitivity for a signal recycling tuning frequency of 1 kHz.

achievable with DC-readout versus the modelled shot noise of S5. The red and the blue dashed-dotted trace represent the sensitivity with a DC-readout scheme for two different ξ_{df}. Even though an optimal ξ_{df}= 0.036 deg was determined from Figure 6.4, half of that dark fringe offset (ξ_{df}= 0.018 deg) already gives a sensitivity only marginally worse. An improvement of the peak sensitivity by about a factor 1.5 is achievable with respect to the modelled shot noise of S5. For the currently used input power of 3.2 W a peak sensitivity of about $1.1 \cdot 10^{-22}/\sqrt{\mathrm{Hz}}$ is obtained.

The orange trace shows the simulated shot noise limited sensitivity with DC-readout for a signal recycling tuning frequency of 1 kHz. Again a peak sensitivity of about $1.2 \cdot 10^{-22}/\sqrt{\mathrm{Hz}}$ is obtained. At a frequency of 1 kHz this sensitivity is a factor of roughly 3 better than the S5 sensitivity of GEO 600 with the currently used signal recycling tuning of 550 Hz.

6.4. Technical realization

The technical implementation of a pure DC-readout scheme for GEO 600 would require some major changes, such as finding a new locking procedure or developing good in-vacuum photodiodes. Before the implementation of DC-readout into a running gravitational wave detector can be considered, a challenging R&D project needs to be carried out, which can probably be best done at a small prototype interferometer.

Nevertheless, a mix-scenario of a heterodyne and a DC-readout scheme for GEO 600 might be easy to implement and would give the chance get a sensitivity improvement at least at high frequencies. The idea is to keep the GEO detector as it is and only implement two changes: The first change would be to introduce an offset from the dark fringe of 50 pm. The second change would be to install an output mode cleaner (OMC) in front of the main photo detector (HPD), see Figure 6.6. Since in GEO 600 the Michelson sideband frequency is about 15 MHz, an OMC with a pole frequency of 1.5 MHz is required in order to give a suppression factor of 10 for the MI sidebands. The locking procedure and the full angular and longitudinal control of the interferometer would stay the same as it is now, except that the signal for the Michelson differential arm length is no longer derived from the main photo detector, but from an auxiliary detector (PDOQ) that is currently only used for lock acquisition and during the down-tuning procedure. For the DC-readout of the GW wave signal still the main photo

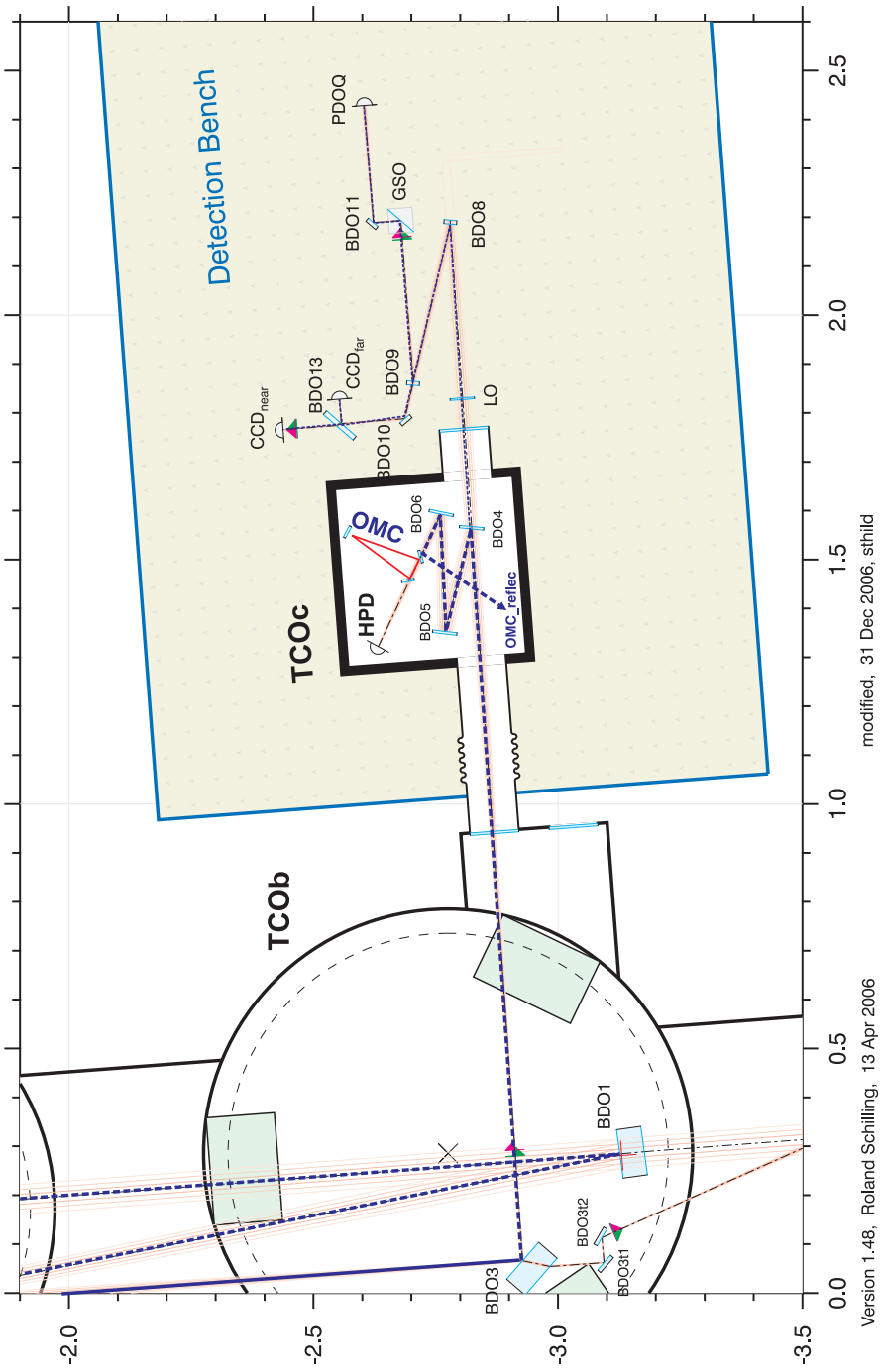


Figure 6.6.: Optical layout of the main detection bench of GEO 600. In front of the main photodiode (HPD) an output mode cleaner (OMC) needs to be added in order to suppress the Michelson sidebands (blue dashed lines) at the HPD. A more detailed explanation of a first realization of a DC-readout scheme can be found in the text.

detector would be used.

Figure 6.6 shows an optical layout of the detection bench. The output beam from the interferometer leaves TCOb and enters the small output vacuum chamber, TCOc, which is currently used at atmospheric pressure and only serves as acoustical enclosure. The beam is split by BDO4 (reflectivity = 0.98) into the so called *high power path* going to HPD and the so called *quadrant path* going to PDOQ. In the high power path an output mode cleaner is added. This filter cavity needs to be designed in a way that the carrier light can be transmitted, while the Michelson RF sidebands are reflected. The 2 % of the light in the quadrant path still contain the MI sidebands and can therefore be used for deriving control signals.

With the setup described above, at least at frequencies above 500 Hz, there is a chance to see the improvement in shot noise with respect to the heterodyne readout. Below 500 Hz the sensitivity will probably be limited by feedback noise introduced by the electronic noise of the quadrant diode. Hence in a next step it would be desirable to use the homodyne signal derived from the HPD for differential arm length control. In a DC-readout scheme a multitude of low frequency noise is expected to originate from acoustical and seismic excitation. That is why in a following step this setup could further be improvement by putting the output mode cleaner into vacuum (TCOc is already installed) and suspending the OMC.

6.5. Summary and outlook

In this chapter a possible new readout scheme of GEO 600 was introduced, combining a heterodyne scheme for control of the detector with a DC-readout scheme. With the aid of simulations a dark fringe offset of about 50 pm was found to be optimal. Without any change of the circulating light power a shot noise limited peak sensitivity can be obtained that is a factor of 1.5 improved with respect to currently modelled shot noise. However, the simulations shown in this chapter are not comprehensive and should be seen as an example of what we may find with GEO 600.

In case it will be decided to change the GEO 600 readout to the here described DC-readout scheme, further investigations have to be carried out in order to successfully enter this new and challenging field of GW detector design.

Appendix A.

The suspended output telescope of GEO 600

The output beam of GEO 600 transmitted through the signal recycling mirror has a diameter of about 2 cm. Such a large beam cannot easily be handled in the complex detection pathes (see Figure 6.6 for example). Therefore it is desirable to convert the large interferometer beam to a smaller beam by the use of an output telescope. Up to May 2004 the GEO 600 output telescope consisted of 3 mirrors rigidly connected to optical tables, with most of the beam path being in air.

In May 2004 a new output telescope, shown in Figure A.1, was installed. It consists of three suspended mirrors, BDO1, BDO2 and BDO3, of which the first one is curved in order to focus the beam. Each suspension consists of a vertical isolation stage, realized by cantilever springs, passively damped by eddy currents, and a horizontal stage, realized by a single pendulum stage. The mirrors are suspended in steel wire loops (127 microns) and the length of the pendulums was chosen to be 460 mm. The position of the mirror is read out by shadow sensors and feedback is applied via four coil magnet actuators each.

The new output telescope provides for the following benefits:

- **In-vacuum beam path:** Since the beam is inside the vacuum, the chance of stray light originating from dust is significantly reduced (see Section 5.3). In addition air induced beam distortion is avoided as well as acoustical excitation of the telescope mirrors.
- **Seismic isolation:** The seismic isolation of the three mirrors reduces the beam

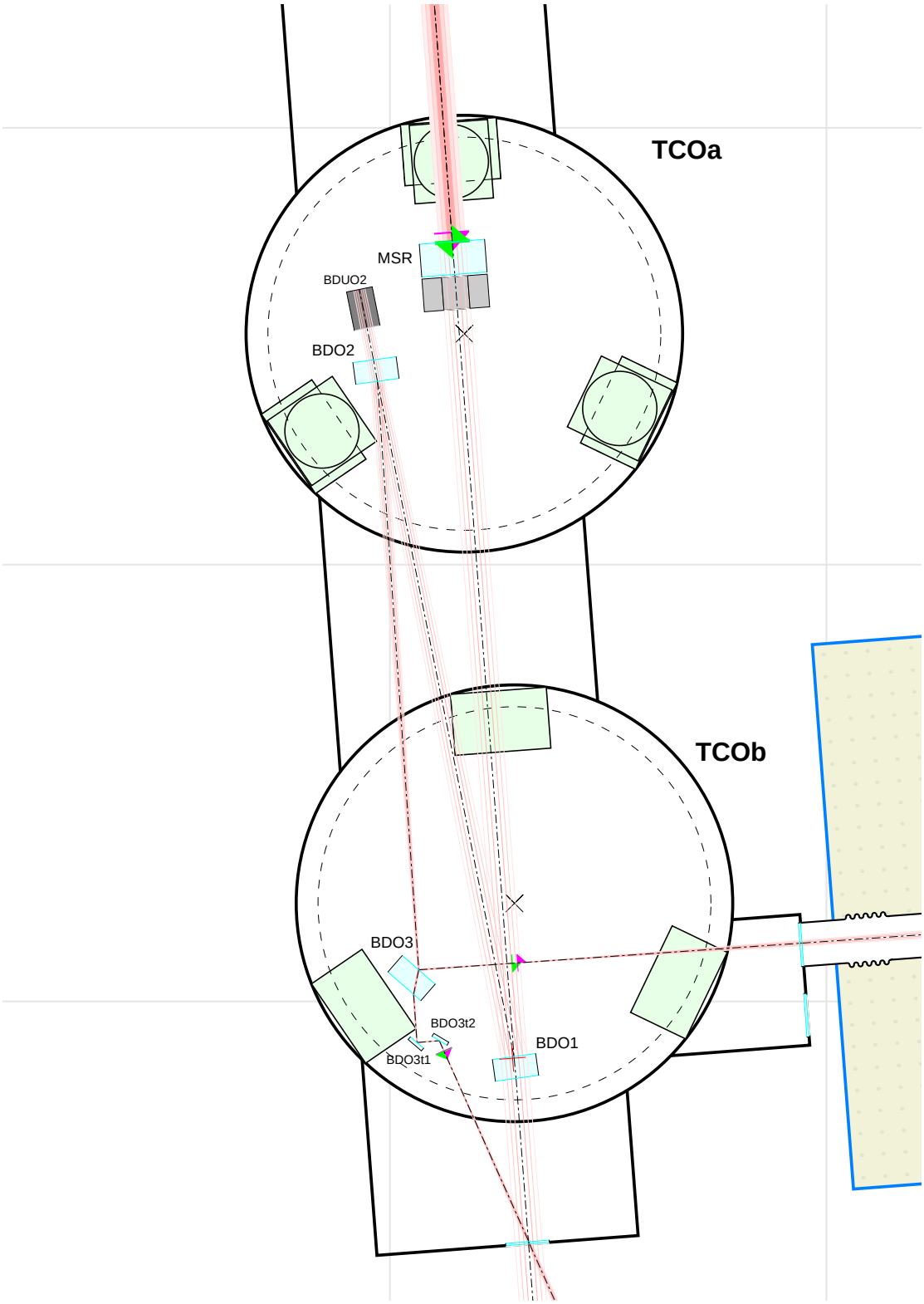


Figure A.1.: OPTOCAD layout of the GEO 600 output telescope, consisting of the three suspended mirrors BDO1, BDO2 and BDO3.

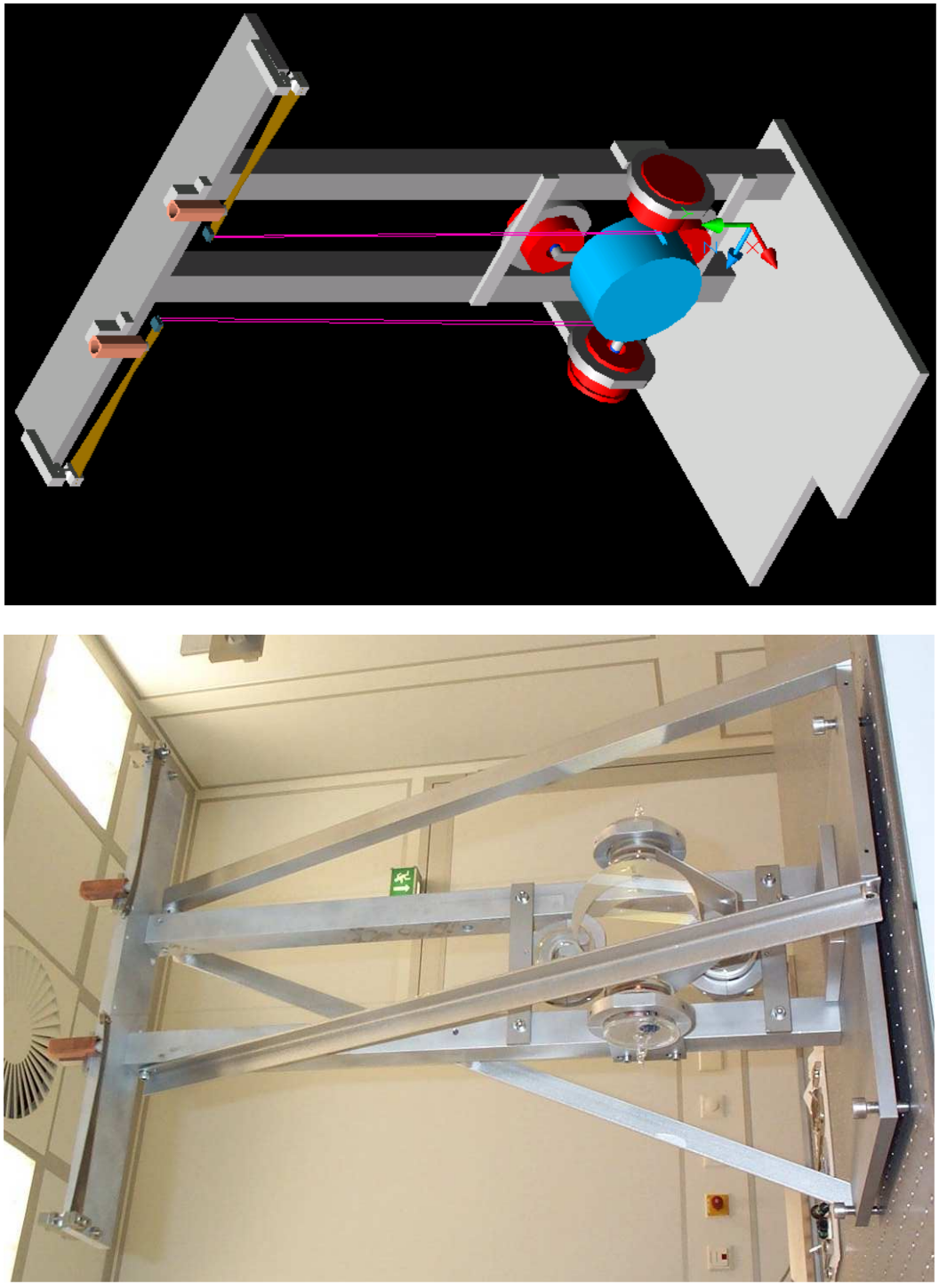


Figure A.2.: Suspension of BDO2. For BDO2 a stand alone Suspension is used which is clamped onto the bottom plate of TCOa.

Figure A.3.: The suspensions for BDO1 and BDO3

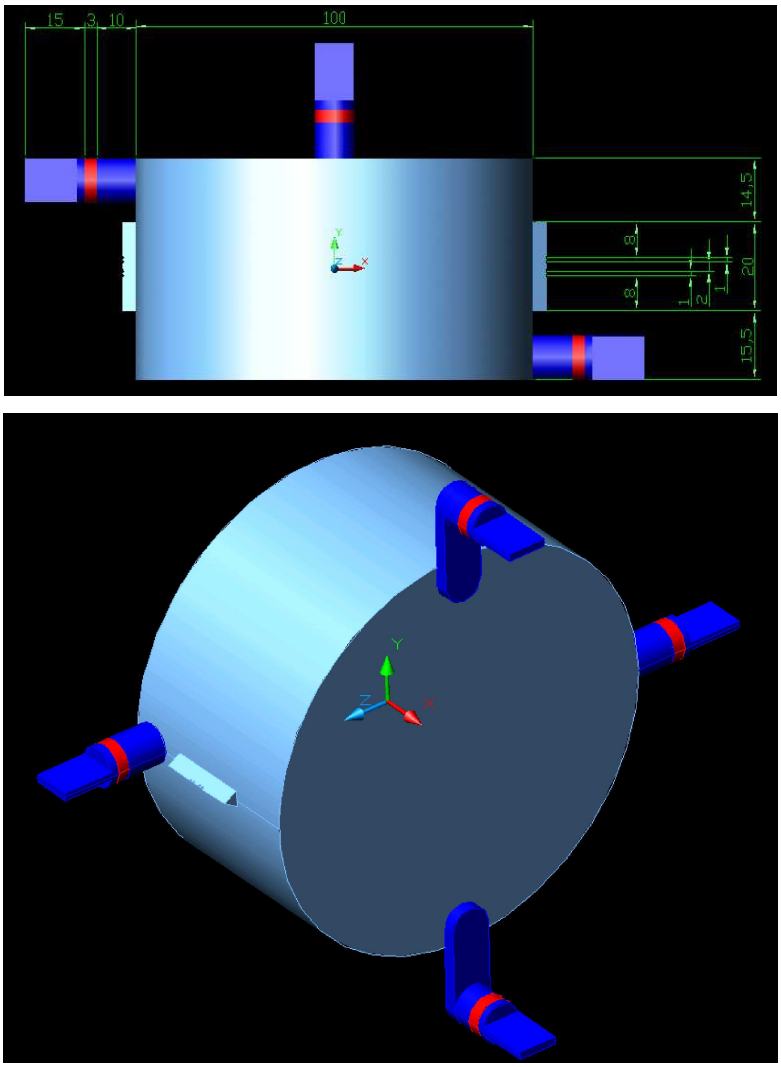


Figure A.4.: AutoCAD drawing of a BDO mirror with glued on break-off-bars, magnets (red) and flags, used for the local controls.

jittering at the main photo detector.

- **Auto-alignment of the output beam:** The suspended telescope mirrors can be used for auto-alignment of the main output beam. The position of the beam transmitted through BDO3 and the position of the beam close to the main photo detector are sensed with quadrant diodes and stabilized by feeding back to the local controls of the output telescope mirrors.

The following table gives an overview of the most important optical and mechanical parameters of the three telescope mirrors.

	BDO1	BDO2	BDO3
Transmission of HR coating (normal incidence)	**1.0 %**	**0.1 %**	14 %
Transmission of HR coating (45°)	7 %	1.3 %	**0.1 %**
Focal length	3.35 m	planar	planar
Thickness	49.8 mm	49.6 mm	49.7 mm
Diameter	99.7 mm	99.9 mm	99.8 mm
Separation of wire loops	2 mm	2 mm	3 mm

Table A.1.: Optical and mechanical parameters of the mirrors used for the output telescope.

Appendix B.

Light fields at the output port for detuned signal recycling

Figure B.1 shows the optical power of the different light fields present at the output of GEO 600, derived from measurements done with a scanning Fabry Perot cavity. Only three light fields are observed in this measurement: The two Michelson sidebands (MI+, MI-) and the carrier (C). As already shown in Figure 2.3 the two Michelson sidebands show a significant asymmetry, which contributes to the complex noise couplings described in Chapter 2.

For the reconstruction of the the total output field, E, we have to take three individual light fields into account, namely the lower Michelson sideband (E_-), the carrier (E_c) and the upper Michelson sideband (E_+). In general we can describe these fields in the following way,

$$E_- = a_- e^{i\omega_- t} \quad \text{with} \quad a_- = A_- e^{i\phi_-} \tag{B.1}$$

$$E_c = a_c e^{i\omega_c t} \quad \text{with} \quad a_c = A_c e^{i\phi_c} \tag{B.2}$$

$$E_+ = a_+ e^{i\omega_+ t} \quad \text{with} \quad a_+ = A_+ e^{i\phi_+} \tag{B.3}$$

where ω_i is the frequency of the light and a_i is complex and contains the amplitude A_i and the phase ϕ_i of the corresponding field. The total output field is just the sum of the three fields

$$E = E_- + E_c + E_+. \tag{B.4}$$

The actual by the photodiode detected property is the power, P of the output field

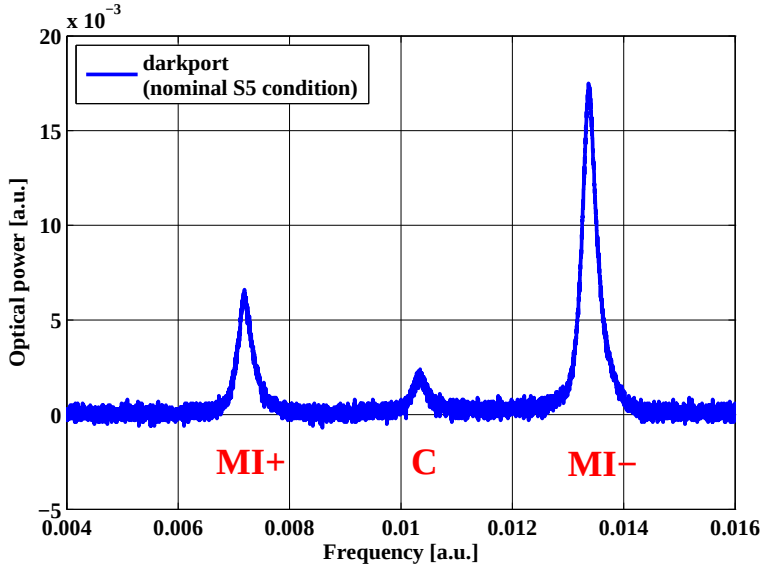


Figure B.1.: Power measurement of the light fields present at the output port of GEO 600. A strong asymmetry of the Schnupp modulation sidebands (MI+, MI-) used for the control of the differential arm length of the Michelson interferometer is observed. The measurement was done using a scanning Fabry Perot cavity.

given by

$$\begin{aligned} P &= EE^* \\ &= a_- a_-^* + a_- a_c^* e^{i(\omega_- - \omega_c)t} + a_- a_+^* e^{i(\omega_- - \omega_+)t} + a_c a_-^* e^{i(\omega_c - \omega_-)t} + a_c a_c^* \\ &\quad + a_c a_+^* e^{i(\omega_c - \omega_+)t} + a_+ a_-^* e^{i(\omega_+ - \omega_-)t} + a_+ a_c^* e^{i(\omega_+ - \omega_c)t} + a_+ a_+^* \end{aligned}$$

Now we can try to simplify this expression by introducing the following definitions. We use the frequency of the carrier light as reference which we define as zero, $\omega_c = 0$ and now redefine the frequencies of the sidebands to be:

$$\omega_- = -\omega \qquad \text{and} \qquad \omega_+ = +\omega \tag{B.5}$$

In addition we also use the phase of the carrier field as reference and define it to be zero, $\phi_c = 0$. With this the whole expression forms to

$$\begin{aligned} P &= A_-^2 + A_- A_c e^{i(-\omega t + \phi_-)} + A_- A_+ e^{i(-2\omega t + \phi_- - \phi_+)} + A_c A_- e^{i(\omega t - \phi_-)} \\ &\quad + A_c^2 + A_c A_+ e^{i(-\omega t - \phi_+)} + A_+ A_- e^{i(2\omega t + \phi_+ - \phi_-)} + A_+ A_c e^{i(\omega t + \phi_+)} + A_+^2 \end{aligned}$$

Now we group the terms by frequency. The DC component, P_{DC} is given by

$$P_{\mathrm{DC}} = A_-^2 + A_c^2 + A_+^2. \tag{B.6}$$

Second we get a contribution at the modulation frequency ω

$$P_{\mathrm{f}} = A_- A_c e^{i(-\omega t + \phi_-)} + A_c A_- e^{i(\omega t - \phi_-)} + A_c A_+ e^{i(-\omega t - \phi_+)} + A_+ A_c e^{i(\omega t + \phi_+)} \tag{B.7}$$

$$= 2A_- A_c cos(\omega t - \phi_-) + 2A_c A_+ cos(\omega t + \phi_+) \tag{B.8}$$

And finally from the beat of the lower and the upper Michelson sideband we also get a contribution at twice the modulation frequency:

$$P_{\mathrm{2f}} = A_- A_+ e^{i(-2\omega t + \phi_- - \phi_+)} + A_+ A_- e^{i(2\omega t + \phi_+ - \phi_-)} \tag{B.9}$$

$$= 2A_- A_+ cos(2\omega t - \phi_- + \phi_+) \tag{B.10}$$

Appendix C.

Design for an AR coated window for quadrant diodes

The standard windows of the quadrant diodes used in GEO 600 have been found to be of poor quality and a source of stray light (see Section 3.5). Therefore the housing of the diodes was cut and the window removed. A new encapsulation for the diodes was designed: The diode is clamped in between two copper plates (see Figures C.1 and C.2). In order to encapsulate the diode surface from dust an AR coated 1" window is glued onto the front-plate.

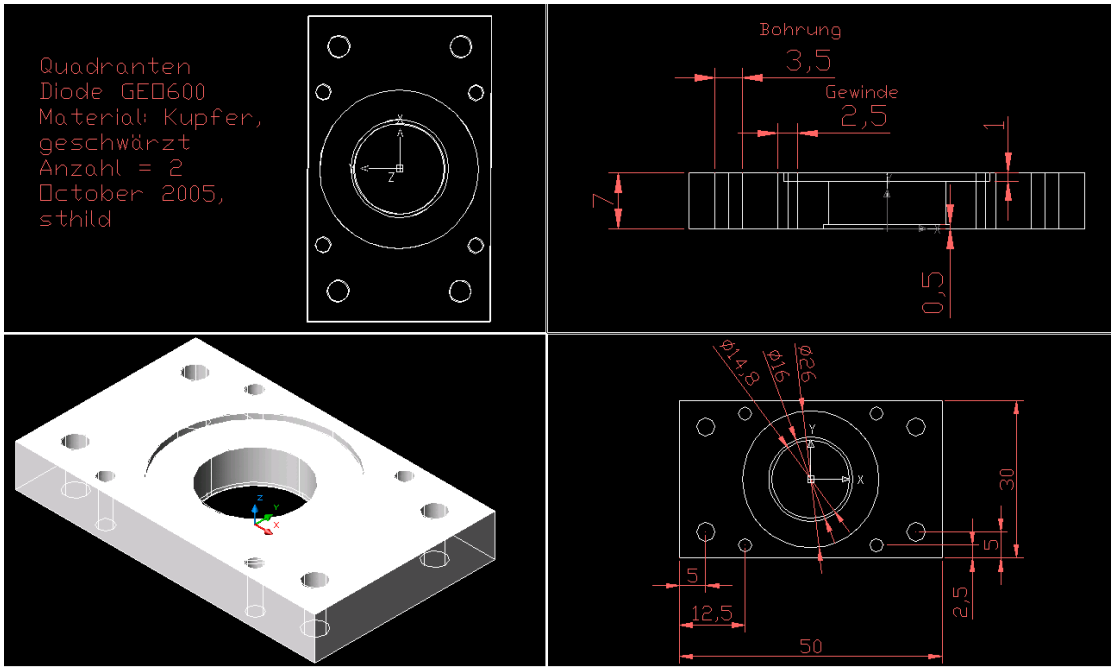


Figure C.1.: Front plate of the new quadrant diode encapsulation.

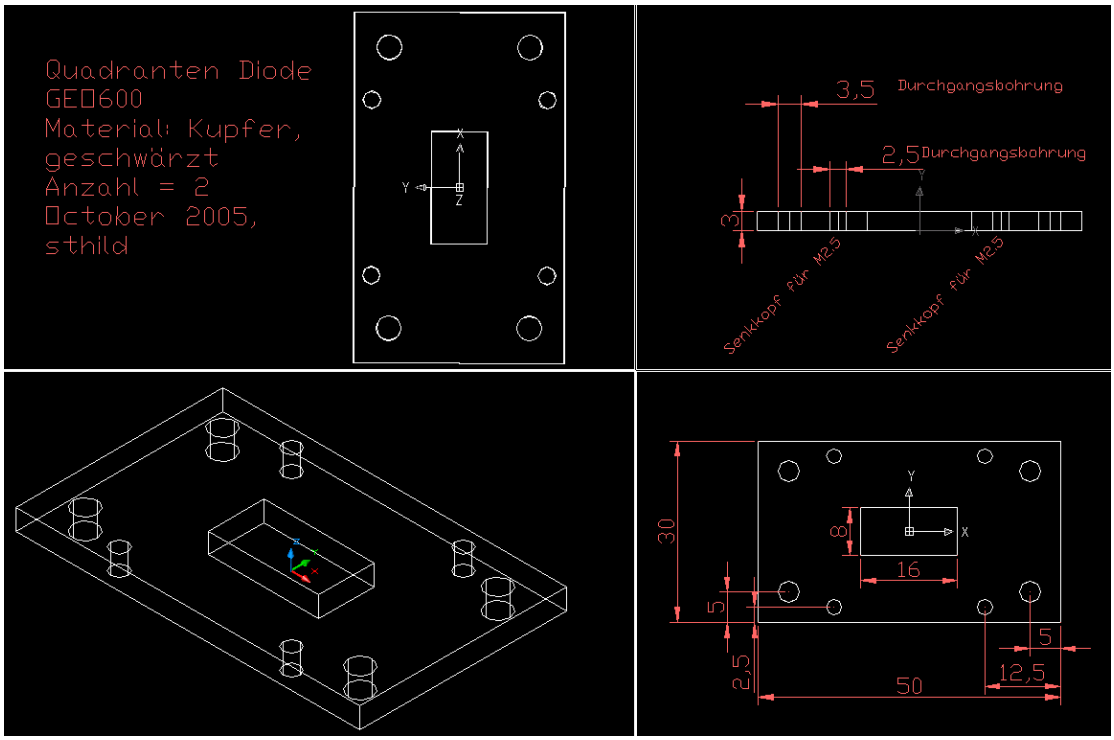


Figure C.2.: Back plate of the new quadrant diode encapsulation.

Appendix D.

Optical layout of GEO 600

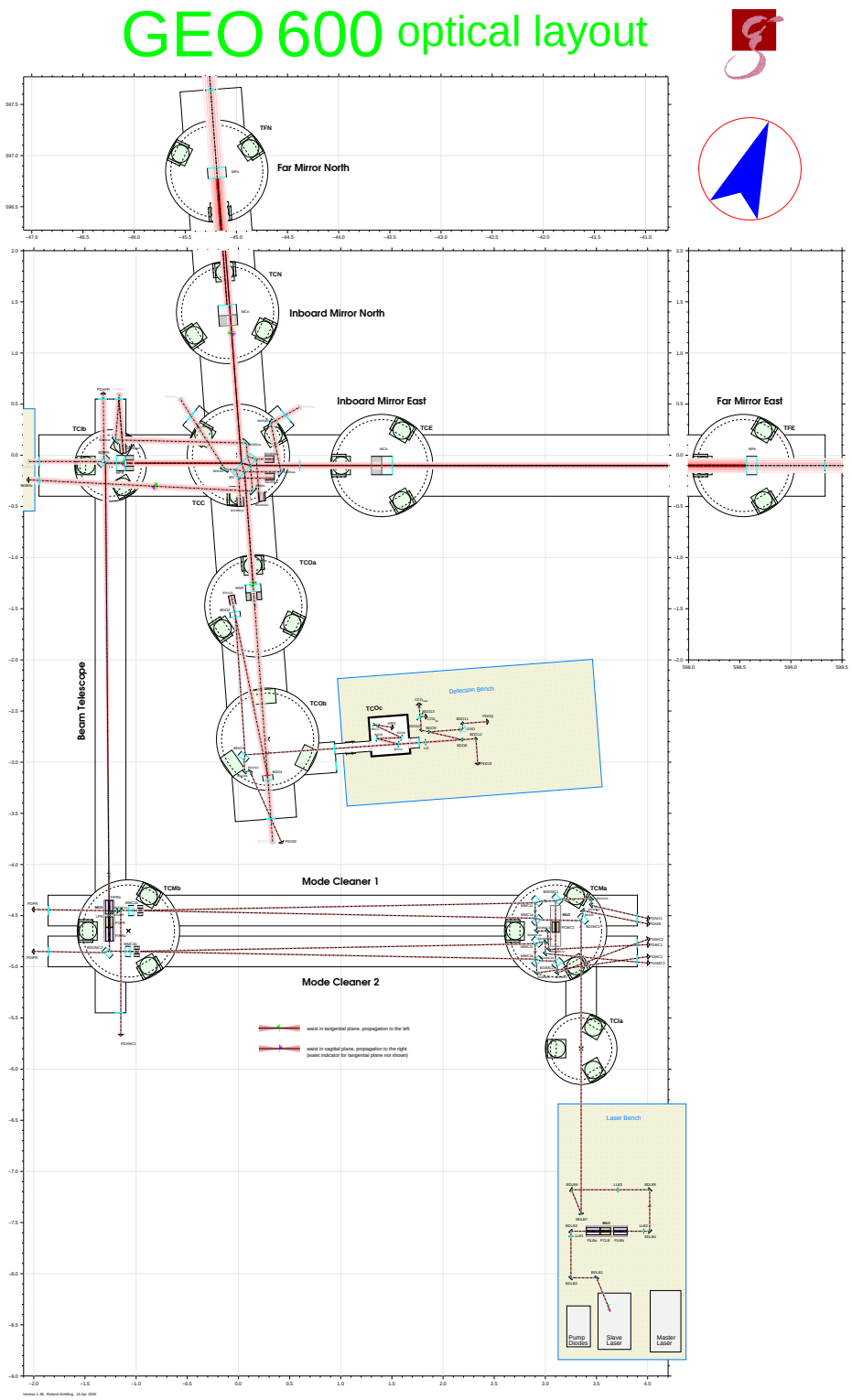


Figure D.1.: Optical layout of GEO 600.

Appendix E.

Matlab script: Photon pressure calibration

```
% Script for photon pressure calibration
% The script does the following steps
% - requesting data from H and PCALmon from the server
% - computation of the photon pressure calibration
% - comparing photon calibration with official calibration
% - computing the transfer function from PCAL to H (mag+phase)

clear;

savedir = ['disse\unsorted\PPD\final2006\nov_3rdrun\100sec_100secin\cal\'];
mkdir('D:\',savedir)%'GPS2UTC(start_time)')

savedir2 = ['disse\unsorted\PPD\final2006\nov_3rdrun\100sec_100secin\photon\'];
mkdir('D:\',savedir2)%'GPS2UTC(start_time)')

savedir3 = ['disse\unsorted\PPD\final2006\nov_3rdrun\100sec_100secin\TF\'];
mkdir('D:\',savedir3)%'GPS2UTC(start_time)')

time1 = UTC2GPS('2006-11-08 14:51:00');
f1 = 58;
time2 = UTC2GPS('2006-11-08 14:48:00');
f2 = 75;
time3 = UTC2GPS('2006-11-08 14:45:00');
f3 = 90;
time4 = UTC2GPS('2006-11-08 14:41:00');
f4 = 120;
```

```
time5 = UTC2GPS('2006-11-08 14:38:00');
f5 = 134;

times = [time1, time2, time3, time4, time5];
 f_phot = [f1, f2, f3,f4, f5];

nsecs = 120;    % total number odf seconds
n_sec_in =120;  % number of seconds in per FFT
loop_num = length(times);
%loop_num = 1;

convers = 1.078;                %conversion factor from V of PD to light power
                                %  in W
c_light = 3e8;                    % speed of light
m = 5.32 ;                       % mass of testmass

error_array = [];
esd_array = [];
photon_array = [];
noise_array = [];
noise_array_HP = [];
HP_array = [];
HP_array_unwh = [];
phase_ar = [];
phase_ar2 = [];
tf = [];
f_real = [];

for loopcounter = 1:1:length(times)
f = f_phot(loopcounter)
start_time = times(loopcounter);

%%%%%%% Getting data from photon_calibrator

server = '130.75.117.73';
port = 9000;
channel(2).name = 'G1:MISC_PHOTON_CAL';
c=2;
    [channel(c).t,channel(c).x,channel(c).fs] = m2fserv(server, port, start_time,
```

```
                                                start_time+nsecs-1, channel(c).name);
    nfft = n_sec_in*channel(c).fs
    S1 = sum(hann(nfft));
    S2 = sum(hann(nfft).^2);
    [channel(c).pxx,channel(c).f] = pwelch(channel(c).x, hann(nfft), 0, nfft,
                                                                  channel(c).fs);
    channel(c).info.enbw = channel(c).fs * S2 / (S1*S1);
    channel(c).info.nfft = nfft;
    channel(c).info.nsecs = nsecs;
    channel(c).info.type = 'AS';
    channel(c).info.ndata = nsecs * channel(c).fs;
    channel(c).pxx = sqrt(channel(c).pxx) * sqrt(channel(c).info.enbw);
   % channel(c).pxx2 = channel(c).pxx2/(sqrt(enbw));

   pho_t=channel(2).t;
   pho_x=channel(2).x;
   pho_fs=channel(2).fs;
phot_data = channel(2).pxx;
[indexmax, valuemax] = getmax(phot_data(10:length(channel(2).f)))
  indexmax = indexmax+9;

figure; semilogy(channel(2).f, phot_data, channel(2).f(indexmax),
phot_data(indexmax), 'or'); xaxis(f-1,f+4); legend('used AS of
photon', 'max value'); photon', 'max value'); grid on
xlabel('frequency [Hz]'); title('G1: MISC\_PHOTON\_CAL');

filename = ['D:\',savedir2, num2str(f)];
print(filename, '-dtiff');
delete(gcf)

%%%%%%%%%%% Getting h(t) data
server = '130.75.117.164';
port = 9008;
rds_level = 9;
cal_version =1;

% Set the channel names
channel(1).name = 'G1:DER_DATA_H';
c = 1;
```

```
% Get the data
    [channel(c).t,channel(c).x,channel(c).fs] = m2fserv(server, port, start_time,
                 start_time+nsecs-1, channel(c).name, cal_version, rds_level);;
    nfft = n_sec_in*channel(c).fs
    S1 = sum(hann(nfft));
    S2 = sum(hann(nfft).^2);
    [channel(c).pxx,channel(c).f] = pwelch(channel(c).x, hann(nfft), 0, nfft,
                                                          channel(c).fs);
    channel(c).info.enbw = channel(c).fs * S2 / (S1*S1);
    channel(c).info.nfft = nfft;
    channel(c).info.nsecs = nsecs;
    channel(c).info.type = 'AS';
    channel(c).info.ndata = nsecs * channel(c).fs;
    channel(c).pxx = sqrt(channel(c).pxx)* sqrt(channel(c).info.enbw);
    enbw = channel(c).info.enbw;

    h_t=channel(1).t;
    h_x=channel(1).x;
    h_fs=channel(1).fs;


h_data = channel(1).pxx;
photon_esd_uncor = h_data(indexmax)             %height of line in DER_DATA_H

noise = h_data((f+1)*n_sec_in+1:(f+3)*n_sec_in+1);   %estimating the noisefloor
                                                     %around the line
  noisefloor = mean(noise);

  photon_esd = sqrt((photon_esd_uncor^2)-(noisefloor^2));   % height of the ppd
                                                            %  line minus noisefloor
  noisefloor_f = [(f+1) (f+3)];                        % needed for plotting
                                                       % the noisefloor
  noisefloor_2 = [noisefloor noisefloor];              % needed for plotting
                                                       % the noisefloor


figure;
semilogy(channel(1).f, h_data, noisefloor_f, noisefloor_2,channel(1).f(indexmax) ,
```

```
                                                                photon_esd,'o');
xaxis(f-1,f+4);
legend('used AS of h(t)', 'used noise floor', 'PPD-line minus noisefloor');
grid on
xlabel('frequency [Hz]');
title('G1:DER\_DATA\_H');

filename = ['D:\',savedir, num2str(f)];
print(filename, '-dtiff');
delete(gcf)

photon_rms = phot_data(indexmax);  %AS 1 Hz frequency resolution from DAQS

photo_pp = photon_rms *2*sqrt(2);     %Gives modulation in Volts

mod_pwr = (photo_pp*convers)*1.029*0.9988*0.9986*0.885;
% photo_pp*convers = 1.078  // PD => PWR-meter
% 1.029  // PWR-meter => PTB
% 0.9988 // Collimator-lens loss
% 0.9986 // Viewport loss
% 0.885  // power loss

F = 2*mod_pwr /c_light ;                % Force to mirror [N]

x = F / (m * channel(1).f(indexmax)^2 * 4 * pi^2);       % actual displacement
                                                         %  of MFN

h_photon = 2*x / 1200
h_photon_rms = h_photon / (2*sqrt(2)); %prc peak height from photon calib
h_esd = photon_esd    %prc peak height from esd calib

error = (h_photon_rms-h_esd)*100/h_esd  %error of h_photon to h_esd in percent

loopcounter

error_array(loopcounter) = error;
esd_array(loopcounter) = h_esd;
photon_array(loopcounter) = h_photon_rms;
noise_array(loopcounter) = noisefloor;
```

```
f_real(loopcounter) = channel(2).f(indexmax)

%%%%%%% making the TF  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

[TF_cxy, TF_f] = tfe(pho_x, h_x, nfft, h_fs, hann(nfft), 0);

phas = angle(TF_cxy)*180/pi;

phase_ar(loopcounter)= phas(indexmax);
tf(loopcounter) = abs(TF_cxy(indexmax))
testphase=phase_ar(loopcounter);

figure;
subplot(2,1,1);
title('TF from MISC\_PHOTON\_CAL to DER\_DATA\_H');
semilogy(TF_f, abs(TF_cxy), TF_f(indexmax), abs(TF_cxy(indexmax)),'r*' );
ylabel('magnitude');
grid on;
subplot(2,1,2);
plot(TF_f, phas, TF_f(indexmax), phas(indexmax),'r*');
grid on;
ylabel('phase [deg]');
allxaxis(f-1,f+1);

filename = ['D:\',savedir3, num2str(f)];
print(filename, '-dtiff');
delete(gcf)

end

figure
subplot(3,1,2);
semilogx(f_real,error_array, '*r');
legend('error');
ylabel('relative error [%]');
%xlabel('Time (hours)');
grid on;
subplot(3,1,1);
loglog(f_real,esd_array, '*b', f_phot,photon_array, '*m');
```

```
legend('height measured in h(t)','predicted for h(t)');
ylabel('AS of h(t)');
grid on;
subplot(3,1,3);
loglog(f_real,noise_array, '*g');
legend('noisefloor around PPD-line');
ylabel('AS of h(t)');
xlabel('frequency of PPD [Hz]');
grid on;
%allxaxis(60,1100);

starttime = start_time-(nsecs*loopcounter);
msuptitle(sprintf('Time from %s (%d), freq = %d Hz', GPS2UTC(starttime),
                                                        starttime,f));

phase_ar2 = phase_ar;

for pz=1:length(phase_ar)
    if phase_ar(pz)>160
        phase_ar2(pz)= phase_ar(pz)-360
    end
end


ra_cal = photon_array./esd_array;
figure;
subplot(2,1,1);
plot(f_real, ra_cal, 'r*-')
xlabel('Frequency [Hz]')
ylabel('ratio of phot-cal/ead-cal')
grid on;
subplot(2,1,2)
plot(f_real, phase_ar2, 'r-*')
xlabel('frequency [Hz]');
ylabel('phase [deg]');
legend('Phase between PPD and G1:DER\_DATA\_H');
grid on;
allxaxis(0,2500);
```

```
figure;
subplot(2,1,1);
loglog(f_real, tf, 'r*-')
xlabel('Frequency [Hz]')
ylabel('TF from Phototn to H')
grid on;
subplot(2,1,2)
semilogx(f_real, phase_ar2, 'r-*')
xlabel('frequency [Hz]');
ylabel('phase [deg]');
legend('Phase between PPD and G1:DER\_DATA\_H');
grid on;
allxaxis(0,6000);
```

Appendix F.

A gravity calibrator for GEO 600

The gravity calibrator is a concept that could provide a calibration that is independent from any subsystem involved in the official calibration. The idea is to use a rotating system of masses, installed closely behind one of the test masses. This will yield, at the position of the testmass, a varying gravity gradient, which can be used as an independent actuator.

Figure F.1 shows the geometry of a possible 3-mass gravity calibrator. In order to find out which force can be applied to the testmass, we have to compare two different states of the gravity calibrator, which represent the two extreme cases of gravity force to the mirror. The first case is shown in the left part of the figure. We can simply calculate the force to the mirror by adding up three forces,one for each calibrator mass. The mirror is assumed to move only perpendicular to its surface.

The gravitational force F_G between two masses m_1 and m_2 separated by a distance of r is known as

$$F_G = G\frac{m_1 m_2}{r^2}, \tag{F.1}$$

where $G = 6,672 \cdot 10^{-11}\mathrm{m}^3/\mathrm{kg\ s}^2$ is the Gravity constant. Therefore the force F_1 on the test mass m_{tm} caused by calibration mass1 (m_{cal}) is given by

$$F_1 = G\frac{m_{tm} m_{cal}}{(r-R)^2}. \tag{F.2}$$

Due to the off axis position of calibration mass 2 and 3, their forces look slightly more complicated:

$$F_{2,3} = \cos(\beta) G\frac{m_{tm} m_{cal}}{(r + R\sin(\alpha))^2 + (R\cos(\alpha))^2}. \tag{F.3}$$

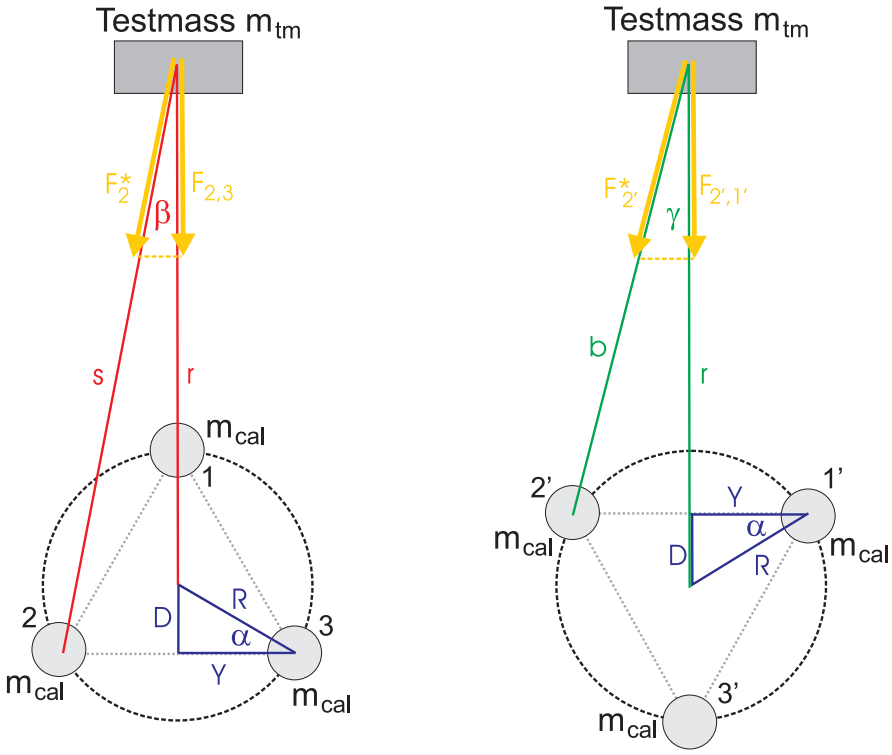


Figure F.1.: Geometry of a 3-mass gravity calibrator.

The cosine is present because we only allow the mirror to move perpendicular to its surface, leading to $F_2 * \cos\beta = F_{2,3}$. The sum of these three forces gives

$$F_A = G \cdot m_{tm} m_{cal} \left(\frac{1}{(r-R)^2} + \frac{2\cos(\beta)}{(r + R\sin(\alpha))^2 + (R\cos(\alpha))^2} \right) \tag{F.4}$$

Now we consider that the calibrator is rotated 60 degrees. Now the calibration masses 1 and 2 are off axis and mass 3 is placed behind the testmass. This configuration is shown in the right part of Figure F.1.

The force caused by the calibration masses $1'$ and $2'$ is given as

$$F_{1',2'} = \cos(\gamma) G \frac{m_{tm} m_{cal}}{(r - R\sin(\alpha))^2 + (R\cos(\alpha))^2} \tag{F.5}$$

and for mass $3'$

$$F_{3'} = G \frac{m_{tm} m_{cal}}{(r+R)^2}. \tag{F.6}$$

The resulting force, F_B is larger than F_A, because in this case two of the calibration masses are near to the mirror.

$$F_B = G \cdot m_{tm} m_{cal} \left(\frac{1}{(r+R)^2} + \frac{2\cos(\gamma)}{(r - R\sin(\alpha))^2 + (R\cos(\alpha))^2} \right) \tag{F.7}$$

The effective force seen by the mirror, ΔF, is given by the difference of F_A and F_B.

$$\Delta F = F_A - F_B \tag{F.8}$$

We simplify this expression by using the fact that α is 30 degrees in the current setup. This reduces the denominators as follows:

$$(r \pm R\sin\alpha)^2 + (R\cos\alpha)^2 = r^2 \pm rR + R^2. \tag{F.9}$$

With this we get

$$\Delta F = G \cdot m_{tm} \cdot m_{cal} \left[\frac{1}{(r+R)^2} + \frac{2\cos\beta}{r^2 - rR + R^2} - \frac{1}{(r-R)^2} - \frac{2\cos\gamma}{r^2 + rR + R^2} \right] \tag{F.10}$$

Now we can take some reasonable numbers for the different parameters:

- $r = 1\,\mathrm{m}$
- $R = 0.2\,\mathrm{m}$

- $m_{tm} = 5.6\,\mathrm{kg}$
- $m_{cal} = 2.0\,\mathrm{kg}$

which leads to a force:

$$\Delta F \approx 6.71 \cdot 10^{-10}\mathrm{N}. \tag{F.11}$$

This force corresponds to a strain amplitude of $h_{\text{grav-cal}} = 6 \cdot 10^{-20}/\sqrt{\mathrm{Hz}}$. Assuming the gravity calibrator rotates with a frequency of 30 Hz (that is the frequency a common washing machine motors provides), a signal at 90 Hz, $h_{\text{grav-cal}}$, would be generated. Comparing $h_{\text{grav-cal}}$ to the actual sensitivity of GEO 600 (see Figure 1.3), we find the gravity calibrator signal to show up in the detector sensitivity with a snr of about 5 for an integration time of one second.

Appendix G.

Matlab script: Statistical veto with amplitude consistency-check

```
%  Script for applying a statistical veto with amplitude
%  consistency check in the full GEO veto pipeline.
%  Actual case: MIDVIS_veto for Sep 2006.

%  Input:
%  - 2 Sets of triggers:  MIDVIS and H
%  - veto list for science, nullstream and CHi^2
%  - Amplitude ratio of H and MIDVIS
%  Output:
%  - List of veto intervals

%  sthild, November 2006

%-------------------------------------------------------
clear;

%% settings
% September 06
s1     = UTC2GPS('2006-09-01 00:00:00');
s2     = UTC2GPS('2006-10-01 00:00:00');
config = 'chacr_200609S5';
outdir = '0906';

%% load triggers
he  = load(sprintf('%s/DER_DATA_H_events_%s', outdir, config));
```

```
xe  = load(sprintf('%s/LSC_MID_VIS_events_%s', outdir, config));
he  = he.events;
xe  = xe.events;
t0 = min([he.gps_start(1) xe.gps_start(1)]);

he.a  = sqrt(10.^(he.totPower/10));
xe.a  = sqrt(10.^(xe.totPower/10));

%% Apply science veto

svi    = viload(sprintf('%s/science_veto_%s.txt', outdir, outdir));
disp(sprintf('+ applying non-science veto to h'));
[hes, hvidx] = applyvi(he, svi);
disp(sprintf('+ applying non-science veto to MID_VIS'));
[xes, xvidx] = applyvi(xe, svi);

%% Apply chi^2 veto

cvi    = viload(sprintf('%s/chi2_dqflag_%s.txt', outdir, outdir));
disp(sprintf('+ applying chi^2 to h'));
[hesc, hvidx] = applyvi(hes, cvi);
disp(sprintf('+ applying chi^2 to MID_VIS'));
[xesc, xvidx] = applyvi(xes, cvi);

%% Apply Nullstream veto

cvi    = viload(sprintf('%s/nullstream_%s.txt', outdir, outdir));
disp(sprintf('+ applying nullstream to h'));
[hescn, hvidx] = applyvi(hesc, cvi);
disp(sprintf('+ applying nullstream to MID_VIS'));
[xescn, xvidx] = applyvi(xesc, cvi);

%% Choose events
hidx  = find(hescn.gps_start > s1 & hescn.gps_start < s2);
xidx  = find(xescn.gps_start > s1 & xescn.gps_start < s2);

out.he = structidx(hescn, hidx, 'serverInfo');
 out.xe =structidx(xescn, xidx, 'serverInfo');
 save('hx_events', 'out');
```

```
out.he.time = out.he.gps_start+out.he.gps_offset;
out.xe.time = out.xe.gps_start+out.xe.gps_offset;

h_t = out.he.time;
vis_t = out.xe.time;
h_freq = out.he.freq_central;
vis_freq = out.xe.freq_central;
h_dur = out.he.duration;
vis_dur = out.xe.duration;
h_snr = out.he.totPower;
vis_snr = out.xe.totPower;
h_snr = undb(h_snr);
vis_snr = undb(vis_snr);

%read in TF data
TF = fig2data('TF_nfest_mag.fig');
TF_freq = TF.line(1).x
 TF_mag = TF.line(1).y*2;
 TF_mag_or = TF_mag/2;

twin = 0.008
fwin = 1000

snr_ratio_l = [];
snr_ratio_h = [];
vis_snr_all = [];
vis_freq_all = [];
h_snr_all = [];
h_freq_all = [];
snr_ratio_h_freq = [];
snr_ratio_l_freq = [];
snr_ratio_all = [];

result = 1:1:length(h_t);
result2 = [];

%% Applying the statistical veto with amplitude consistency-check
```

```
%     for n=1:400
%         timeshift = (n-200)*0.3
%         h_shift = h_t +timeshift;
%         fprintf('%d', n)

        for j=1:length(h_t)
        step = j;
        tdiff = vis_t - h_t(j);
        idx = find(abs(tdiff) < (twin));
            if sum(idx)>0.5
                 %   for u=1:length(idx)
                        freq_diff = h_freq(j)-vis_freq(idx) ;
                        idx2 = find(abs(freq_diff)<(fwin));
                        if sum(idx2)>0.5
                            result(j) =1;
                            snr_ratio = [];
                            snr_ratio = vis_snr(idx)./h_snr(j);
                            snr_ratio_all = [snr_ratio_all snr_ratio];
                            fre = round(h_freq(j));
                            comp=snr_ratio./TF_mag(fre);
                            idx3 = find(comp<1);

                            if idx3>0.5;
                               result2(j) =1;
                               snr_ratio_h = [snr_ratio_h snr_ratio];
                            else
                                result2(j) =0;
                                snr_ratio_l = [snr_ratio_l snr_ratio];
                            end

                        else
                        result(j) = 0;
                        end
            else
            result(j) = 0;
            end


        end
```

```
%% Plotting the results

    figure
    semilogy(snr_ratio_h_freq, snr_ratio_h, 'bx', snr_ratio_l_freq,
             snr_ratio_l, 'rx', TF_freq, TF_mag, 'r', TF_freq, TF_mag_or, 'r');
    xlabel ('Frequency [Hz]')
    ylabel('Ratio of total power in H and MID\_VIS')
    legend('TotPower\_H devided by TotPower\_MID\_VIS, fwin = 1000 Hz,
           twin = 0.008sec, FULL SEPTEMBER 2006', 'TotPower\_H devided
           by TotPower\_MID\_VIS, fwin = 1000 Hz, twin = 0.008sec, FULL
           SEPTEMBER 2006' )
    title(sprintf('triggers in h(t) = %d : triggers in auxiliary channel = %d
           \n  Efficiency = %d (no amp cut) : Efficiency = %d amp cut applied'
            , length(h_t) , length(vis_t), sum(result)/length(h_t)*100,
            (sum(result)-sum(result2))/length(h_t)*100));
    xaxis(50,2000);
    grid on;

    sum(result)
    sum(result2)
    veto = [];

%% Writing the list of veto intervals

    idx_veto = find(result>0.5);
    for i = 1:length(idx_veto)
        i
            veto(1,i)= hescn.gps_start(idx_veto(i))+hescn.gps_offset(idx_veto(i));
            veto(2,i)= hescn.duration(idx_veto(i));
    end
```

Appendix H.

FINESSE input file of GEO 600 for DC-readout

```
#----------------------------------------------------------------------
# File used to determine the optimal dark fringe offset for #
using a DC readout scheme. # sthild, 12/2006
#----------------------------------------------------------------------
# geo600-main.kat $Rev: 8 $ # Andreas Freise
(adf@star.sr.bham.ac.uk) # $Date: 2007/01/04 13:34:10 $ # # Input
File for FINESSE (www.rzg.mpg.de/~adf) # # Optical layout of GEO
600 with "real" parameters . #
#----------------------------------------------------------------------
# # History: # # 12.12.2006 by  Andreas Freise
(adf@star.sr.bham.ac.uk) # - changed distribution of losses to be
130ppm on each #   surface inside the DR MI # # 12.12.2006 by
Andreas Freise (adf@star.sr.bham.ac.uk) # # - changed MC mirror
parameters according to labbook page 4027 # - changed laser power
to be at (70deg ->) 4.82W (page 3984) # - changed node names of
MPR and MSR # - changed mirror specs of MCN, MFN, MCE. MFE
according to #   labbok page 4028, results are exactly as stated
there by Hartmut # - not yet done: mode-matching, curvature,
compensation check, .... #

# 21.11.2006 by  Andreas Freise (adf@star.sr.bham.ac.uk)
#                Stefan Hild    (stefan.hild@aei.mpg.de)
#
#   Restarting GEO 600 file maintenance, see Labbok page 4011
```

```
#   Staring from a file called power_evl_curr_8_car.kat
#   (26.06.2006, labbok page 3656)
#
# - changed syntax to Finesse version 0.99.4
# - put Mode cleaners back in
# - updated modulation indices (see page 4011)
# - added telescope in south output port (Hartmut)
# - added dymanic thermal lens (for BS) computation
#
#-------------------------------------------------------------


 l i1 3.2 0 nMU3in1                                # corresponds to 75 deg
                                                   # (nominal S5)
 gauss beam_in i1 nMU3in1 268u -550m               # beam parameter roughly
                                                   # matching PRC
## (old value used:  i1 nMU3in 248u -550m)          ** to be checked **
 mod  eom3 $fPR $midxPR 2 pm 0 nMU3in1 nMU3in  # PRC control
##

mod  eom4 $fSR $midxSR 2 pm 0 nMU3in nMU3_2      # Schnupp1 (SR control)
mod  eom5 $fMI $midxMI 2 pm 0 nMU3_2 nMU3_3      # Schnupp2 (MI control)

lens lpr 1.8 nMU3_3 nMU3_4
# some rather arbitrary thermal lense for the isolators and the EOMs:
lens therm 5.2 nMU3_4 nMU3_5                       # ** to be checked **
isol d2 120 nMU3_5 nMU3out                           # Faraday Isolator

# 070502 corrected length with respect to OptoCad (Roland Schilling)
s    smcpr3 4.391 nMU3out nBDIPR1
bs1  BDIPR 50u 30u 0 45 nBDIPR1 nBDIPR2 dump dump
s    smcpr4 0.11 nBDIPR2 nMPR1


##-------------------------------------------------------------
## main interferometer ##
##
## New MPR; values for MPR005 page 2264 (check with Harald)
## first (curved) surface of MPR
```

```
m    mPRo 0 1 0 nMPR1 nMPRi1
attr mPRo Rc -1.85
s    smpr 0.0718 1.44963 nMPRi1 nMPRi2
# second (inner) surface of MPR
m1   MPR 900u $LMPR 0. nMPRi2 nMPR2              # T=900 ppm, L=50 ppm

s    swest 1.1463 nMPR2 nBSwest     # new length with T_PR=900 ppm
                                    #  * to be checked **

##-------------------------------------------------------------
## BS
## basic data from old GEO files
##
##
##                        nBSnorth     ,'-.
##                              |     +    `.
##                              |   ,'       :'
##            nBSwest           |  +i1      +
##         ----------------->    ,:._  i2 ,'
##                              + \  `-. +       nBSeast
##                            ,' i3\   ,' ---------------
##                           +      \ +
##                         ,'     i4.'
##                        `._      ..
##                           `._ ,' |nBSsouth
##                              -   |
##                                  |
##                                  |
bs2   BS 0.485998 $LBS 0.0 42.834 nBSwest nBSnorth nBSi1 nBSi3
s    sBS1a 0.040 1.44963 nBSi1 nBSi1b
##-------------------------------------------------------------
# Thermal lense of beam splitter
lens bst 14.5k nBSi1b nBSi1c # static value for 1.9kW at BS
## -----------------------------------------------------------
# Alternative: dynamic thermal lens computation, as in:
# S. Hild et al, Applied Optics IP, vol. 45, Issue 28, pp.7269-7272
# assuming 0.25ppm/cm absorption, w=0.88cm, d=9cm
/*
pd prcpower nBSwest # we need 2* power in BS, so we measure power in
```

```
west arm
noplot prcpower
set bspow prcpower re
func flength = 1.6635E7 / ( $bspow + 0.0000000001)
#noplot flength
put bst f $flength
*/
s    sBS1 0.051 1.44963 nBSi1c nBSi2
s    sBS2 0.091 1.44963 nBSi3 nBSi4
bs2   BS2 60u $LBSAR 0  27.9694 nBSi2 dump nBSeast nBSAR   # R=60 ppm, L=30ppm
bs2   BS3 60u $LBSAR 0 -27.9694 nBSi4 dump nBSsouth dump     # R=60 ppm, L=30ppm
# two measured values for R_AR: labbook page 2418 (44ppm), 3996 (64ppm)

##-------------------------------------------------------------
## north arm
s snorth1 598.5682 nBSnorth nMFN1  # ** to be checked **

bs1 MFN 8.3u $LMFN 0.0 0.0 nMFN1 nMFN2 dump dump  # T=8.3 ppm
attr MFN Rc 666

s snorth2 597.0241 nMFN2 nMCN1 # ** to be checked **

m1 MCN 13u $LMCN -0.0 nMCN1 dump                  # T=13 ppm
attr MCN Rc 636

##-------------------------------------------------------------
## east arm
s seast1 598.4497 nBSeast nMFE1
bs1 MFE 8.3u $LMFE 0.0 0.0 nMFE1 nMFE2 dump dump  # T=8.3 ppm
# ** the Rc(T) below need to be checked, they certainly look wrong **
attr MFE Rcx 664    # 90 W
attr MFE Rcy 660    # 90 W
#attr MFE Rcx 665   # 71 W
#attr MFE Rcy 662   # 71 W
#attr MFE Rc 663.75 # perfect curvature
#attr MFE Rcx 666.41    # 66W
#attr MFE Rcy 663.75    # 66W
#attr MFE Rcy 660.75    # 75W
#attr MFE Rcx 663.75    # 75W
```

```
s seast2 597.0630 nMFE2 nMCE1
m1 MCE 13u $LMCE 0.0 nMCE1 dump                    # T=13ppm
attr MCE Rc 622

##-----------------------------------------------------------
## south arm
s ssouth 1.109 nBSsouth nMSR1

m MSR 0.9805 0.01945 0.7646 nMSR1 nMSR2     # R=0.9805, T=0.01945, L=50 ppm
# tuning = f_tune/FSR_SR * 180, FSR_SR=125241 Hz ** to be checked **
# e.g f_tune 532 Hz -> 0.7646 deg
##-----------------------------------------------------------
## output optics telescope

s sout1 1.8 nMSR2 nBDO1i
bs1 BDO1 0.01 0.0 0.0 5.0 nBDO1i nBDO1o dump dump    # T=1%
attr BDO1 Rc 6.72

s sout2 4.855 nBDO1o nLO1i # BDO2 and BDO3 are flat and omitted in this path
lens LO1 0.5 nLO1i nLO1o   # 1. lens on detection bench
s sout3 0.703 nLO1o nLO2i  # computed telescope length, 2 flat mirrors
                           # omitted in this path
lens LO2 -0.03 nLO2i nLO2o # 2. lens on detection bench
# actual lens -0.05m ?
s sout4 1.0 nLO2o nout     # length to quad. camera

##-----------------------------------------------------------
## further settings and commands

# Modulation frquencies
const fSR 9016865    ## corresponding to 532 Hz on tune.vi, (10/2006 S. Hild)
const fMI 14.904929M ## (10/2006 S. Hild)
const fPR  37.16M
const midxPR 0.13 # see page 4011
const midxSR 0.17 # see page 4011
#const midxMI 0.38 # see page 4011
const midxMI 0.019
```

```
# Michelson losses
const LMPR 130u
const LMCN 130u
const LMFN 130u
const LMCE 130u
const LMFE 130u
const LBS 130u
const LBSAR 130u

# PR cavity (north arm)
cav prc1 MPR nMPR2 MCN nMCN1
# PR cavity (east arm)
cav prc2 MPR nMPR2 MCE nMCE1
# SR cavity (north arm)
cav src1 MSR nMSR1 MCN nMCN1
# SR cavity (east arm)
cav src2 MSR nMSR1 MCE nMCE1

##-----------------------------------------------------------
## Simualtion commands

/*
# power detectors
#pd MC1out nMU2_2
#pd MC2out nMU3_5
#pd MPRrelf nMPR1
#pd MPRin nMPR1*
#pd PRC nMPR2
pd BSpow nBSwest
pd MSRpow nMSR1
#pd darport nBDO1i
xaxis BS phi lin -1 1 100
*/

maxtem 2
retrace off
time
phase 3
```

```
fsig sig1 MCN 10 0
fsig sig2 MCE 10 180

pdS1 hpd_DC 10 max nout
xaxis sig1 f log 50 2000 200
put hpd_DC f1 $x1
x2axis MCE phi lin 0 0.1 100
func t = 0-$x2
put MCN phi $t

yaxis log abs
scale meter hpd_DC

pause gnuterm windows
```

Appendix I.

Measurement of the bulk-absorption of the GEO 600 beam splitter

I.1. Introduction

The four large-scale laser-interferometric gravitational wave projects, LIGO [Waldman06], VIRGO [Acernese06], TAMA 300 [Ando05] and GEO 600 [Hild06c], have dedicated much effort to commissioning the detectors and improving their sensitivity. Soon the initial projects will have reached a sensitivity limited by fundamental noise sources like, for instance, shot noise. Second generation projects such as Advanced LIGO [AdvLIGO], Advanced VIRGO [AdvVIRGO] and GEO HF [Willke06] aiming at strain sensitivities in the region of 10^{-23} to $10^{-24}/\sqrt{\text{Hz}}$ will operate at much higher light powers than the initial detectors, in order to reduce the influence of shot noise. Accordingly, one major problem will be the absorption of laser light in the optical elements like beam splitters and test masses. Even though techniques for thermal correction and compensation have been developed (e.g., [Lück04, Lawrence02]), the use of low-absorption materials is a key point of future detector research.

The beam splitter substrate currently installed in GEO 600 was manufactured by Heraeus, Hanau [Heraeus] and consists of Suprasil 311 SV which provides extremely low bulk absorption due to an OH-content of less than 50 ppm [Loriette03]. Recent measurements of OH-reduced fused silica showed an absorption below 0.5 ppm/cm which was the lowest absorption of fused silica reported so far [Loriette03]. Therefore it is of great interest to obtain a lower absorption value, as presented in this work.

In Section I.2 I will explain the principle of a new method for a more sensitive estimation

of the bulk absorption in the beam splitter by using the GEO 600 interferometer itself as a measuring tool. Realization and details of this measurement are described in Section I.3 while the result is discussed in Section I.4.

I.2. A new method for measuring bulk absorption

Large Michelson interferometers such as GEO 600 working at high optical powers can be used to measure small effects of thermally induced distortion of their optical elements [Ottaway06]. One of the most sensitive optical components in GEO 600 is the beam splitter which is placed inside a high-finesse cavity: the power-recycling cavity (Figure I.1.1). The absorption of the bulk material causes a weak lens building up inside the beam splitter [Winkler91, Strain94]. This lens disturbs mainly the beam passing through the beam splitter substrate, whereas the beam reflected to the other arm (in case of GEO 600 this is the north arm) is only slightly influenced (Figure I.1.3). This means that depending on the strength of the lens we see a change in the interference pattern at the ports of the interferometer. Since GEO 600 is operated on a dark fringe at the output port, this effect can most easily be visualized there, using, for example, a CCD camera.

Due to the fact that the thermal lens does not build up instantaneously, we can compare the beam pattern at the output port for a cold state and a hot state of the beam splitter (as shown in Figure I.2). The high gain of the alignment control system guarantees a stable alignment and good suppression of first order higher modes [Grote04a]. Using GEO 600 parameters we can reproduce the pattern of the cold state using a FINESSE [Freise04] simulation. Now we can add to the simulation an additional lens inside the beam splitter and vary its focal length until it matches the experimentally observed beam pattern for the hot state. If the light power transmitted through the beam splitter is known, an estimate of the bulk absorption can be derived from the focal length of the simulated thermal lens as described below.

Absorption of light power from the transmitted beam heats the substrate non uniformly. Because of the temperature dependence of the index of refraction a path difference δs occurs between a light path measured along the beam axis and a light path measured

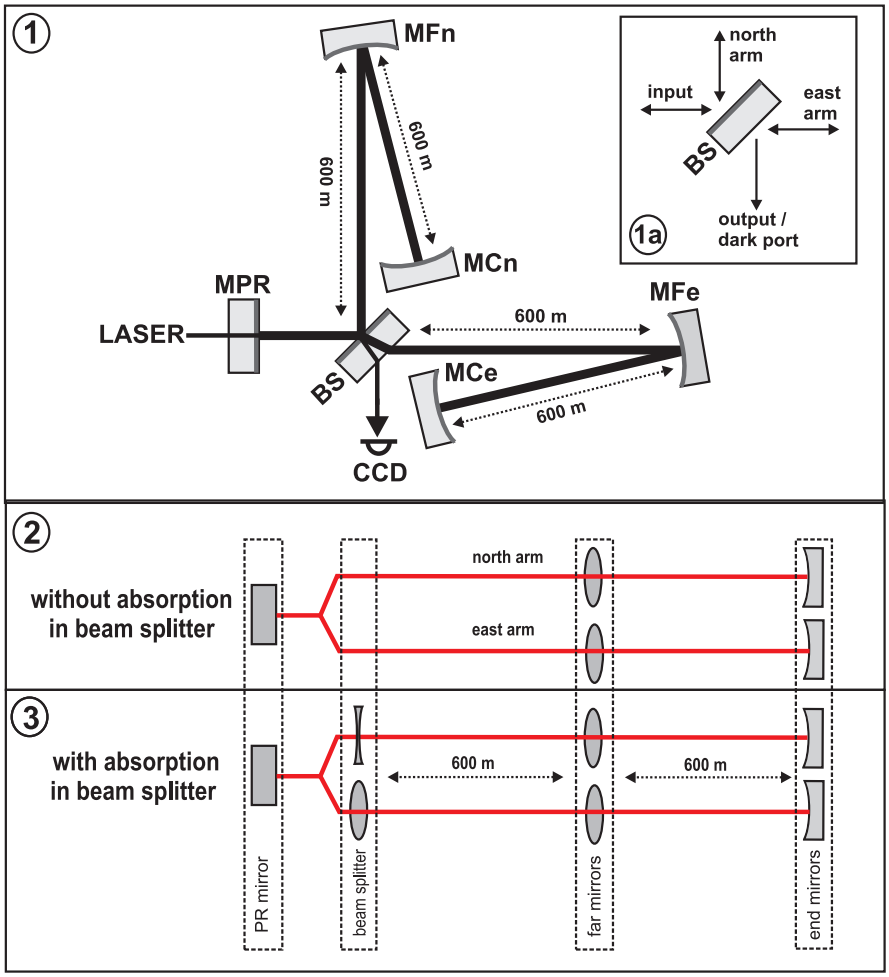


Figure I.1.: **1**. Simplified optical layout of the GEO 600 main interferometer with folded arms. The Laser beam enters the Michelson interferometer through the power-recycling mirror (MPR), gets split at the beam splitter and transverses the two folded arms, each of 2400 meter round trip length. At the output port of the interferometer the beam pattern is observed using a CCD camera. **2**. Optical imaging inside the interferometer arms expressed in an equivalent lens diagram for the case of no absorption inside the beam splitter substrate. **3**. Bulk absorption inside the beam splitter can be modelled by an additional convex lens inside the east arm (dominant effect) and an concave lens in the north arm.

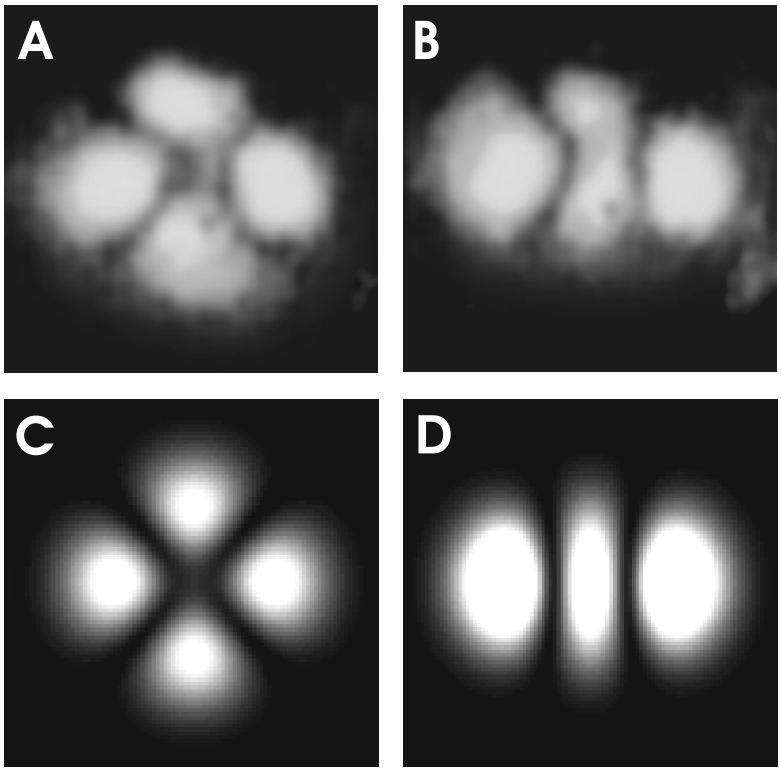


Figure I.2.: Beam pattern of the GEO 600 dark port for the case of a power-recycled Michelson interferometer. **A**: Observed dark port pattern for a cold beam splitter. **B**: Observed dark port pattern for a hot beam splitter, i.e., thermal lens present. **C**: Corresponding simulation of the cold beam splitter. **D**: Corresponding simulation of the hot beam splitter, i.e., thermal lens present.

along the $1/e^2$ point of the intensity distribution [Winkler91].

$$\delta s = 1.3 \cdot \frac{\beta}{4\pi\kappa} \cdot p_a \cdot d \cdot P \tag{I.1}$$

Here κ is the thermal conductivity, p_a is the absorption per unit length, P is the light power and d is the geometrical path length inside the substrate. The temperature dependence of the index of refraction is given by $\beta = dn/dT$. Expressing δs by the focal length of the thermal lens induced in the beam splitter as $f_{\mathrm{therm}} = w^2/2\delta s$, where w is the beam radius at the beam splitter, Equation (I.1) transforms to

$$p_a = \frac{4 \cdot \pi}{2.6} \cdot \frac{w^2 \cdot \kappa}{\beta \cdot d \cdot P \cdot f_{\mathrm{therm}}}. \tag{I.2}$$

Up to here we only took absorption inside the beam splitter and the corresponding thermo-refractive effect into account. Of course, in reality the optical imaging in GEO 600 is also slightly influenced by two other effects: the thermal expansion and the absorption of the dielectric coatings. Nevertheless ignoring these two effects still gives a valid upper limit for the bulk absorption of the beam splitter, as we will show in the following sections.

The principle of our absorption measurement relies on the thermally induced change in the difference of the wavefront curvature of the two interferometer arms. Therefore we can neglect absorption at any coating of the four main interferometer mirrors (MFe, MFn, MCn and MCe in Figure I.1) and the corresponding thermal deformation of the mirror surfaces, because this effect would influence the wavefront curvature similarly in both interferometer arms (assuming symmetric absorption in both arms of the interferometer). Nevertheless we can give a rough estimate of the change in radius of curvature of the mirror and the resulting change in the beam diameter caused by coating absorption of the four main interferometer mirrors. In that respect the far mirrors (MFe, MFn) are the most critical ones. Assuming a coating absorption at one of the far mirrors of 2 ppm, would increase the radius of curvature which is 666 m by about 10 cm. The beam diameter w which is about 8.8 mm would change by roughly 6 microns. Therefore the coating absorption of the main interferometer mirrors is totally negligible.

Assuming symmetrical absorption in the two arms of the interferometer the beam splitter is the only optical component that can change the wavefront curvatures of the two arms differentially. In the case of the bulk absorption it is reasonable to neglect the thermal

expansion because it is first of all ten times smaller than the thermo-refractive effect and second would enhance the strength of the thermal lens.

Finally we have to consider light absorption at the two dielectric coatings of the beam splitter and the related thermal effect. For the east arm these effects can be modelled by adding an additional convex lens, while in the north arm we have to insert a concave lens.

All of the effects discussed above would cause a differential change of the two wavefront curvatures, which would increase the strength of the observed thermal lens and by this lead to a smaller value of the bulk absorption p_a. Hence Equation I.2 gives a valid upper limit for the bulk absorption inside the beam splitter.

I.3. Setting an upper limit for the bulk absorption in the GEO 600 beam splitter

To investigate the thermal lensing of the beam splitter, the GEO 600 detector was used in the configuration of a power-recycled Michelson interferometer. To avoid the influence from mode healing [Grote04b], the signal-recycling mirror was misaligned, such that it can just be considered as an attenuator at the dark port. Figure I.2A shows the dark port image for a cold state beam splitter, i.e., immediately after lock acquisition. Figure I.2C shows the result from a FINESSE simulation for the same configuration using our best estimate of the parameters of the GEO 600 detector.

The cross shape of the dark port image is caused by an astigmatic mismatch of the radii of curvature from the two far mirrors (MFe and MFn in figure I.1.1) [Lück04]. Due to this astigmatism the dark port image changes quite strongly with the beam splitter's lensing, which makes it easier to match measurements and simulations. Furthermore this fact also makes our method more accurate in obtaining the focal length of the thermal lens.

Figure I.2B shows the dark port image for the hot state of the beam splitter, after the lens has fully developed and the beam splitter is in thermal equilibrium, which takes about 30 minutes. The result of the corresponding simulation is shown in Figure I.2D. An additional thermal lens with a focal length of $f_{\mathrm{therm}} = 13\,\mathrm{km}$ best matches the simulation to the observed dark port shape.

The light power inside the beam splitter for this measurement was $P = 1.4$ kW. The beam radius is $w = 0.88\,\mathrm{cm}$ and the geometrical path length inside the beam splitter is $d = 9\,\mathrm{cm}$. Using these parameters and $\beta/\kappa = 10^{-5}\mathrm{m/W}$ for Suprasil [Takke] Equation (I.2) gives an upper limit for the bulk absorption of the GEO 600 beam splitter of

$$p_\mathrm{a} = 0.25 \pm 0.1\,\mathrm{ppm/cm}. \tag{I.3}$$

The main contributions to the error budget are uncertainties in the measurement of the radii of curvature of the far mirrors, the intra-cavity power, P, and the beam radius w.

I.4. Conclusion

A new method was developed to estimate the bulk absorption of beam splitter substrates in a large scale power-recycled Michelson interferometer. Using this method we obtained an upper limit of the bulk absorption of the GEO 600 beam splitter of $p_\mathrm{a} = 0.25 \pm 0.1$ ppm/cm. This is, to the knowledge of the author, the lowest value ever measured for absorption in fused silica at a wavelength of 1064 nm.

Bibliography

[Acernese06] *"The Virgo status"*, F Acernese et al., Class. Quantum Grav. **23** No 19 (2006) S635–S642.

[Adhikari04] *"Sensitivity and Noise Analysis of 4 km Laser Interferometric Gravitational Wave Antennae"*, R Adhikari, PhD thesis (2004) Massachusetts Institute of Technology.

[Adhikari06] *"Enhanced LIGO"*, R Adhikari, P Fritschel, S Waldman, LIGO Technical Note, LIGO-T060156-01-I (2006)

[AdvLIGO] http://www.ligo.caltech.edu/advLIGO/

[AdvVIRGO] *"Advanced Virgo White Paper"*, R Flaminio, A Freise, A Gennai, P Hello, P La Penna, G Losurdo, H Lueck, N Man, A Masserot, B Mours, M Punturo, A Spallicci, A Vicer´e for the Virgo Collaboration, 2005

[Ajith06] *"Robust vetoes for gravitational-wave burst triggers using known instrumental couplings"*, P Ajith, M Hewitson, J R Smith and K A Strain, Class. Quantum Grav. **23** No 20 (2006) 5825–5837.

[Ajith06a] *"Null-stream veto for two co-located detectors: implementation issues"*, P Ajith, M Hewitson and I S Heng, Class. Quantum Grav. **23** No 19 (2006) S741–S749.

[Akutsu] *"Veto analysis for gravitational wave burst signals in TAMA300 data using an ALF filter"*, T Akutsu, M Ando, N Kanda, D Tatsumi, S Telada, S Miyoki, M Ohashi and the TAMA Collaboration, Class. Quantum Grav. **23** No 8 (2006) S23–S28.

[Anderson01] *"Excess power statistic for detection of burst sources of gravitational radiation"*, W G Anderson et al, Phys. Rev. D **63** No 4 (2001) 42003

[Ando05] *"Current status of the TAMA300 gravitational-wave detector"*, Masaki Ando and the TAMA Collaboration, Class. Quantum Grav. **22** No 18 (2005) S881–S889.

[Ansys] ANSYS, Inc., *www.ansys.com*

[Astone02] *"Resonant mass detectors: present status"*, P Astone, Class. Quantum Grav. **19** (2002) 1227–1235.

[Balasubramanian05] *"Results from the first burst hardware injections performed on GEO 600"*, R Balasubramanian, H Grote, I S Heng, M Hewitson, H Lück, J R Smith, K A Strain, H Ward and B Willke, Class. Quantum Grav. **22** No 14 (2005), 3015–3028.

[Beauville05] *"A first comparison of search methods for gravitational wave bursts using LIGO and Virgo simulated data"*, F Beauville, M-A Bizouard, L Blackburn, L Bosi, P Brady, L Brocco, D Brown, D Buskulic, S Chatterji, N Christensen, A-C Clapson, S Fairhurst, D Grosjean, G Guidi, P Hello, E Katsavounidis, M Knight, A Lazzarini, F Marion, B Mours, F Ricci, A Viceré and M Zanolin (The Joint LIGO/Virgo Working Group), Class. Quantum Grav. **22** No 18 (2005) S1293–S1301.

[Blair] *"The Detection of Gravitational Waves"*, D G Blair [Editor]:, Cambridge University Press 1991.

[Buonanno03] *" Quantum noise in laser-interferometer gravitational-wave detectors with a heterodyne readout scheme"*, A Buonanno, Y Chen, and N Mavalvala, Phys. Rev. D **67**, 122005, (2003)

[Candonati04] *"Coherent waveform consistency test for LIGO burst candidates"*, L Cadonati, Class. Quantum Grav. **21** No 20 (2004) S1695–S1703.

[Candonati05] *"CorrPower: a cross-correlation-based algorithm for triggered and untriggered gravitational-wave burst searches"*, L Cadonati and S Márka, Class. Quantum Grav. **22** No 18 (2005) S1159–S1167.

[Chatterji04] *"Multiresolution techniques for the detection of gravitational-wave bursts"*, S Chatterji et al, Class. Quantum Grav. **21** (2004) S1809.

[Clapson05] *"A gravitational wave burst search method based on the S transform"*, André-Claude Clapson, Matteo Barsuglia, Marie-Anne Bizouard, Violette Brisson,

Fabien Cavalier, Michel Davier, Patrice Hello, Stéphane Kreckelberg and Monica Varvella, Class. Quantum Grav. **22** No 18 (2005) S1381–S1390.

[Cutler/Thorne] *"AN OVERVIEW OF GRAVITATIONAL-WAVE SOURCES"*, C CUTLER and K S THORNE, submitted to World Scientific on May 17, 2006

[DiCredico05] *"Gravitational wave burst vetoes in the LIGO S2 and S3 data analyses"*, Alessandra Di Credico (for the LIGO Scientific Collaboration), Class. Quantum Grav. **22** No 18 (2005) S1051-S1058.

[Dupuis05] "Bayesian estimation of pulsarparameters from gravitational wave data", R J Dupuis and G Woan, Phys. Rev. D **72**, (2005) 102002.

[Finesse] *"FINESSE 0.99.4, Frequency domain INterferomEter Simulation SoftwarE"*, Andreas Freise, http://www.rzg.mpg.de/ adf/download/Finesse-0.99.4.pdf.

[Freise03] *"The Next Generation of Interferometry: Multi-Frequency Optical Modelling, Control Concepts and Implementation"*, A Freise, Dissertation, Universität Hannover (2003).

[Freise04] *"Frequency-domain interferometer simulation with higher-order spatial modes"*, A Freise et al, Class. Quantum Grav. **21** (2004) S1067–1074.

[Giaime] *"Adavnced LIGO"*, J Giaime et al, Amaldi6 Conference (2005).

[Gossler02] *"The modecleaner system and suspension aspects of GEO 600"*, S Gossler, M M Casey, A Freise, H Grote, H Luck, P McNamara, M V Plissi, D I Robertson, N A Robertson, K Skeldon, K A Strain, C I Torrie, H Ward, B Willke, J Hough and K Danzmann, Class. Quantum Grav. **19** No 7 (2002) 1835–1842.

[Gossler04] *"The suspension systems of the interferometric gravitational-wave detector GEO600"*, S Goßler, Dissertation, Universität Hannover (2004).

[Grote] *"High-power, low-noise, and multiply resonant photodetector for interferometric gravitational wave detectors."*, H Grote, submitted to Rev. of Sci. Instruments.

[Grote03] *"Making it Work: Second Generation Interferometry in GEO600 !"*, H Grote, PhD thesis, Hannover 2003.

[Grote04a] *"Alignment control of GEO 600"*, H Grote et al, 2004, Class. Quantum Grav. **21**, S441–449.

[Grote04b] *"Dual recycling for GEO 600"*, H Grore et al, 2004, Class. Quantum Grav. **21**, S473–480.

[Grote05] *"The Status of GEO 600"*, H Grote et al, Class. Quantum Grav. **22** (2005) S193–S198.

[Hanna06] *"Reducing gravitational wave false alarms using signals at the antisymmetric port in LIGO detectors"*, C R Hanna (for the LIGO Scientific Collaboration), Class. Quantum Grav. **23** No 8 (2006) S17-S22.

[Harms06] *"Shot Noise and Heterodyne Detectors"*, J Harms, internal document, 2006

[Heraeus] *"Quarzglas für die Optik, Daten und Eigenschaften"*, Heraeus Quarzglas GmbH & Co.KG, Hanau (1999).

[Heinzel98] *"Experimental Demonstration of a Suspended Dual Recycling Interferometer for Gravitational Wave Detection"*, G. Heinzel, K.A. Strain, J. Mizuno, K.D. Skeldon, B. Willke, W. Winkler, R. Schilling, A. Rüdiger, and K. Danzmann, Phys. Rev. Lett. **81** (1998) 5493–5496.

[Heng03] *"Methods for multi-detector burst gravitational wave search"*, I S Heng, F Salemi and A Ortolan, Class. Quantum Grav. **20** No 17 (2003) S617-S622.

[Heng04] *"First steps towards characterizing the hierarchical algorithm for curves and ridges pipeline"*, I S Heng, R Balasubramanian, B S Sathyaprakash and B F Schutz, Class. Quantum Grav. **21** No 5 (2004) S821-S826.

[Hewitson04] *"Principles of calibrating the dual-recycled GEO 600"*, M Hewitson et al, Rev. Sci. Instrum., **75** (2004) 4702

[Hewitson04a] *"On aspects of characterising and calibrating the interferometric gravitational wave detector, GEO 600"*, M Hewitson, PhD thesis, University of Glasgow, 2004.

[Hewitson05] *"Optimal time-domain combination of the two calibrated output quadratures of GEO 600"*, M Hewitson et al, Class. Quantum Grav. **22** (2005) 4253–4261

[Hewitson05a] *"Using the null-stream of GEO 600 to veto transient events in the detector output"*, M Hewitson and P Ajith, Class. Quantum Grav. **22** No 22 (2005) 4903–4912.

[Hild04] *"A new calibration method"*, S Hild, presentation at the GEO meeting, April 2004, Hannover.

[Hild06a] *"Towards gravitational wave astronomy: Commissioning and characterization of GEO600"*, S Hild, H Grote, J R Smith, M Hewitson, J. Phys.: Conf. Ser. **32** (2006) 66–73.

[Hild06b] *"Measurement of a low-absorption sample of OH-reduced fused silica"*, S Hild, H Lück, W Winkler, K Strain, H Grote, JR Smith, M Malec, M Hewitson, B Willke, J Hough and K Danzmann, Appl. Opt. 45, 7269-7272 (2006).

[Hild06c] *"The status of* GEO 600*"*, S Hild (for the LIGO Scientific Collaboration), Class. Quantum Grav. **23** No 19 (2006) S643-S651.

[Hild07a] *"A statistical veto method using auxiliary channels sensitive to gravitational wave signals"*, S Hild, P Ajith and M Hewitsion, submitted to Class. Quantum Grav.

[Hild07b] *Demonstration and comparison of tuned and detuned signal recycling in a large-scale gravitational wave detector*, S Hild, H Grote, M Hewtison, H Lück, J R Smith, K A Strain, B Willke and K Danzmann, submitted to Class. Quantum Grav.

[Hulse] *"The discovery of the binary pulsar (PSR 1913+16)"*,R A Hulse, Rev. Mod. Phys., 66, 699, (1994).

[ILIAS06] *"2nd Report on WG1: Antennas commissioning and characterization"*, M Barsuglia, A Freise, I Fiori, H Grote, H Heitmann, S Hild, P LaPenna, G Losurdo, H Lück, J Smith, L Taffarello, G Vajente, M Visco, B Willke, 2006.

[Klimenko04] *"A wavelet method for detection of gravitational wave bursts"*, S Klimenko and G Mitselmakher, Class. Quantum Grav. **21** (2004) S1819–S1830.

[Koetter03] *"PQMon: a powerful veto for burst events"*, K Kötter, I S Heng, M Hewitson, K A Strain, G Woan and H Ward, Class. Quantum Grav. **20** No 17 (2003) S895-S902.

[Kuroda06] *"The status of LCGT"*, K Kuroda (and the LCGT Collaboration), Class. Quantum Grav. **23** No 8 (2006) S215–S221

[Lawrence02] *"Adaptive thermal compensation of test masses in advanced LIGO"*, R Lawrence R al, 2002, Class. Quantum Grav. **19** S1803–1809.

[LIGO05] LLO electronic log 2005-08-19 (B O Reilly, personal communication).

[Loriette03] *"Absorption of low-loss optical materials measured at 1064 nm by a position-modulated collinear photothermal detection technique"*, V Loriette and C Boccara, 2003, Applied Optics, **42** 649–656.

[Lück04] *"Thermal correction of the radii of curvature of mirrors for GEO 600"*, H Lück et al, Class. Quantum Grav. **21** (2004), S985–989.

[Lück06] *"Status of the GEO600 detector"*, H Lück et al, Class. Quantum Grav. **23** S71-S78 (2006).

[Mantone06] *"Benefits of Artificially Generated Gravity Gradients for Interferometric Gravitational-Wave Detectors"*, L Mantone et al, in preparation.

[Malec06] *"Commissioning of advanced, dual-recycled gravitational-wave detectors: simulations of complex optical systems guided by the phasor picture"*, M Malec, PhD thesis, Hannover 2006.

[Meers] *"Recycling in laser-interferometric gravitational-wave detectors"*, B J Meers: , Phys. Rev. **D 38** (1988) 2317–2326.

[Mossavi06] *"A photon pressure calibrator for the GEO 600 gravitational wave detector"*, K Mossavi, M Hewitson, S Hild, F Seifert, U Weiland, J R Smith, H Luck, H Grote, B Willke, K Danzmann, Phys. Letters A **353** (2006) 1–3.

[Ottaway06] *"In-situ measurement of absorption in high power interferometers using beam diameter measurements"*, D Ottaway et al, Optics Letters, 31 (4) pp. 450–452, 2006.

[Rakhmanov05] *"A cross-correlation method for burst searches with networks of misaligned gravitational-wave detectors"*, M Rakhmanov and S Klimenko, Class. Quantum Grav. **22** No 18 (2005) S1311-S1320.

[Rowan/Hough] *"Gravitational Wave Detection by Interferometry (Ground and Space)"*, S Rowan, J Hough:, Living Rev. Relativity, 3, (2000), 3. [Online Article]: cited on 9.9.2003, http://www.livingreviews.org/Articles/Volumes3/2000-3hough/.

[Saulson] *"Fundamentals of Interferometric Gravitational Wave Detectors"*, P R Saulson, World Scientific Publishing, 1994.

[Savage06] R Savage, P Kalmus, E Goetz, personal communication.

[Schnupp] L Schnupp, Presentation at European Collaboration Meeting on Interferometric Detection of Gravitational Waves, Sorrent, 1988.

[Sigg04] *"Commissioning of LIGO detectors"*, D Sigg, Class. Quantum Grav. **21** (2004) S409–S415.

[Smith04] *"Mechanical quality factor measurements of monolithically suspended fused silica test masses of the GEO 600 gravitational wave detector"*, J R Smith, G Cagnoli, D R M Crooks, M M Fejer, S Goßler, H Lück, S Rowan, J Hough and K Danzmann, Class. Quantum Grav. **21** (2004) S1091–S1098.

[Smith06] *"Formulation of Instrument Noise Analysis Techniques and Their Use in the Commissioning of the Gravitational Wave Observatory, GEO600"*, J R Smith, PhD thesis, Hannover 2006.

[Strain91] *"Experimental Demonstration of Dual Recycling for Interferometric Gravitational-Wave Detectors"*, K. Strain, B. Meers, Phys. Rev. Lett. **66** (1991) 1391–1394.

[Strain94] *"Thermal lensing in recycling interferometric gravitational wave detectors"*, K Strain et al, 1994, Phys. Let. A, **194**, 124–132.

[Strain] K A Strain, personal communication.

[Stuver06] *"A first comparison of SLOPE and other LIGO burst event trigger generators"*, Amber L Stuver and Lee Samuel Finn, Class. Quantum Grav. **23** No 19 (2006) S733-S740.

[Takahaschi04] *"Direct measurement of the scattered light effect on the sensitivity in TAMA300"*, R Takahashi, K Arai, S Kawamura and M R Smith, Phys. rev. D, **(**70) No 6 (2004) 062003.1–062003.4

[Takke] Personal communication with R. Takke from Heraeus.

[Taylor] *"Binary pulsars and relativistic gravity"*, J H Taylor, Rev. Mod. Phys. 66, (1994) 711.

[Vinet96] *"Scattered light noise in gravitational wave interferometric detectors: Coherent effects"*, J Y Vinet, V Brisson, S Braccini, Phys. Rev. D **54**, (1996) 1276 – 1286

[Waldman06] *"Status of LIGO at the start of the fifth science run"*, Samuel J Waldman (for the LIGO Science Collaboration), Class. Quantum Grav. **23** No 19 (2006) S653-S660.

[Weber] *"Evidence for discovery of gravitational radiation"*, J. Weber Phys. Rev. Lett. **25** (1970) 180.

[Weiland04] *"Preparing for gravitational wave astronomy: A verification of the GEO600 detection chain by generation, injection, and extraction of continuous signals"*, U Weiland, PhD thesis, Hannover 2004.

[Wen05] *"Coherent network detection of gravitational waves: the redundancy veto"*, L Wen and B F Schutz, Class. Quantum Grav. **22** No 18 (2005) S1321-S1335.

[Willke04] *"Status of GEO 600"*, B Willke et al, Class. Quantum Grav. **21** S417-S423 (2004).

[Willke06] *"The GEO HF project"*, B Willke et al, Class. Quantum Grav. 23 S207–S214 (2006).

[Winkler] *"Eine optische Verzögerungsleitung für ein Breitband-Gravitationswellenexperiment"*, W Winkler, Doktorarbeit, Max-Planck-Institut für Quantenoptik, 1983.

[Winkler91] *"Heating by optical absorption and the performance of interferometric gravitational wave detectors"*, W Winkler et al, 1991, Phys. Rev. A, **44**, 7022.

Acknowledgement

Joining the GEO-project, then seeing and feeling it grow and proceed was an exciting experience for me and I would like to thank all the people who contributed to GEO 600.

First I would like to thank Karsten Danzman for giving me the chance to work right at the core of the GEO 600 detector. It is unusual for a nearly unexperienced student to get the opportunity to 'play'with such an exciting device as GEO 600.

The people who contributed most to this work are the members of the GEO commissioning team: Hartmut Grote, Harald Lück, Martin Hewitson, Joshua Smith, Ken Strain and Benno Willke have been for me an infinite supply of motivation, creativity and knowledge. I am happy and proud to be a member of such an great team. I'm very grateful to these people for sharing their knowledge and work with me, - nearly everything I know about building, commissioning and running a gravitational wave detector I learned from these bunch of scientists. Finally I want to add that working at GEO was not only scientifically interesting, but also always good fun and I enjoyed the familiar atmosphere in Ruthe very much. Thank you !!!

Furthermore I would like to thank many people from outside the Ruthe-team who contributed to this work: Keita Kawabe for introducing me to the 'real-live'interferometry at the Garching prototype interferometer. Stefan Goßler for teaching me how to build a proper suspension. Andreas Freise for many fruitful discussions and his help in doing simulations. Iain Martin, Stuard Reid and Jim Hough for providing the FEM analysis of the photon pressure induced mirror deformation. Peter Kalmus, Rick Savage and Brian O'Reilly for interesting discussions about the photon pressure calibrators used in GEO and LIGO. Simon Chelkowski for motivating me to use 3D AutoCAD. Ajith Parameswaran for giving me a background in veto analysis. Andreas Weidner and Heiko zur Mühlen for help in building electronics. Walter Winkler and Roland Schilling for sharing their knowledge with me and answering endless rows of questions. All members of the ILIAS Working group 1 for the interesting discussions and the nice atmosphere at all the meetings we had. Finally I thank Andre Thuering, not only for sharing an office in Hannover with me, but also for always being a friend.

For proof reading of this thesis I thank: Andreas Freise, Jan Harms, Hartmut Grote, Martin Hewitson, Harald Lück, Kasem Mossavi, Joshua Smith, Ken Strain and Benno Willke. However, I am of course responsible for all remaining mistakes in this thesis.

I am deeply grateful to Marion for all her patience and love.

Curriculum vitae

Stefan Hild
Meterstrasse 6
30169 Hannover

Born on April 1st, 1978 in Celle, Germany
Marital status: unmarried

12. 2003 – to date	Scientific Assistant, Max-Planck-Institut für Gravitationsphysik (AEI), Hannover
12. 2003 – 2. 2007	Doctoral Studies in Physics, University of Hannover
11. 2003	Physik Diplom, University of Hannover (Thermally tunable Signal-Recycling for GEO 600)
10. 2002 – 02.2003	Max-Planck Institute für Quantenoptik, Garching (Internship on the Garching prototype)
10. 1998 – 11.2003	Diploma Studies in Physics, University of Hannover
08. 1997 – 09.1998	Alternative civilian service
06. 1997	Abitur
08. 1990 – 06.1997	Hölty Gymnasium Celle
08. 1988 – 07.1990	Orientierungsstufe Ovelgönne
08. 1984 – 07.1988	Grundschule Hambühren

Publications

2007

- S Hild, H Grote, M Hewtison, H Lück, JR Smith, KA Strain, B Willke and K Danzmann, "Demonstration and comparison of tuned and detuned signal recycling in a large-scale gravitational wave detector", accepted for publication in Class. Quantum Grav.

2006

- S Hild for the LIGO Scientific Collaboration, "The Status of GEO600", Class. Quantum Grav. 23 S643-S651 (2006).
- S Hild, H Lück, W Winkler, K Strain, H Grote, J Smith, M Malec, M Hewitson, B Willke, J Hough and K Danzmann, "Measurement of a low-absorption sample of OH-reduced fused silica", Appl. Opt. 45, 7269-7272 (2006).
- JR Smith, P Ajith, H Grote, M Hewitson, S Hild, H Lück, KA Strain, B Willke, J Hough and K Danzmann, "Linear projection of technical noise for interferometric gravitational-wave detectors", Class. Quantum Grav. 23 527-537 (2006).
- K Mossavi, M Hewitson, S Hild, F Seifert, U Weiland, JR Smith, H Lück, H Grote, B Willke and K Danzmann, "A photon pressure calibrator for the GEO600 gravitational wave detector", Phys. Lett. A 353 1-3 (2006).
- S Hild, H Grote, JR Smith and M Hewitson, "Towards gravitational wave astronomy: Commissioning and characterization of GEO600", J. Phys.: Conf. Ser. 32 66-73 (2006).
- H Lück et al, "Status of the GEO600 detector", Class. Quantum Grav. 23 S71-S78 (2006).

- B Willke et al, "The GEO HF project", Class. Quantum Grav. 23 S207-S214 (2006).
- The LIGO Scientific Collaborration, "Joint LIGO and TAMA300 search for gravitational waves from inspiralling neutron star binaries", Phys. Rev. D 73, 102002 (2006).
- The LIGO Scientific Collaborration, "Search for gravitational waves from binary black hole inspirals in LIGO data", Phys. Rev. D 73, 062001 (2006).

2005

- JR Smith, H Grote, M Hewitson, S Hild, H Lück, M Parsons, KA Strain and B Willke, "Feedforward correction of mirror misalignment fluctuations for the GEO 600 gravitational wave detector", Class. Quantum Grav. 22 3093-3104 (2005).
- M Hewitson, H Grote, S Hild, H Lück, P Ajith, JR Smith, KA Strain, B Willke and G Woan, "Optimal time-domain combination of the two calibrated output quadratures of GEO600", Class. Quantum Grav. 22 4253-4261 (2005).
- M Hewitson for the LIGO Scientific Collaboration, "Preparing GEO 600 for gravitational wave astronomy: a status report", Class. Quantum Grav. 22 S891-S900 (2005).
- H Grote et al, "The status of GEO 600", Class. Quantum Grav. 22 S193-S198 (2005).
- The LIGO Scientific Collaborration, "Search for Gravitational Waves Associated with the Gamma Ray Burst GRB030329 Using the LIGO Detectors", Phys. Rev. D 72, 042002 (2005).
- The LIGO Scientific Collaborration, "First All-sky Upper Limits from LIGO on the Strength of Periodic Gravitational Waves Using the Hough Transform", Phys. Rev. D 72, 102004 (2005).
- The LIGO Scientific Collaborration, "Limits on Gravitational-Wave Emission from Selected Pulsars Using LIGO Data", Phys. Rev. Lett. 94, 181103 (2005).
- The LIGO Scientific Collaborration, "Search for Gravitational Waves from Galactic and Extra-galactic Binary Neutron Stars", Phys. Rev. D 72, 082001 (2005).

- The LIGO Scientific Collaborration, "Search for Gravitational Waves from Primordial Black Hole Binary Coalescences in the Galactic Halo", Phys. Rev. D 72, 082002 (2005).
- The LIGO Scientific Collaborration, "Upper Limits on Gravitational Wave Bursts in LIGO's Second Science Run", Phys. Rev. D 72, 062001 (2005).
- The LIGO Scientific Collaborration, "Upper limits from the LIGO and TAMA detectors on the rate of gravitational-wave bursts", Phys. Rev. D 72, 122004 (2005).

2004

- JR Smith et al, "Commissioning, characterization, and operation of the dual-recycled GEO 600", Class. Quantum Grav. 21 S1737-S1745 (2004).
- B Willke et al, "Status of GEO 600", Class. Quantum Grav. 21 S417-S423 (2004).
- KA Strain et al, "The status of GEO 600", Optical Fabrication, Metrology, and Material Advancements for Telescopes. Edited by Atad-Ettedgui, Eli; Dierickx, Philippe. Proceedings of the SPIE, Volume 5500, pp. 25-36 (2004).
- H Lück, A Freise, S Goßler, S Hild, K Kawabe and K Danzmann, "Thermal correction of the radii of curvature of mirrors for GEO 600 ", Class. Quantum Grav. 21 S985-S989 (2004)

2003

- S Hild, "Thermisch ducrhstimmbares Signal-Recycling für den Gravitationswellendetektor GEO600", Diploma Thesis, University Hannover, 2003.

Submitted

- S Hild, P Ajith and M Hewitson, "A statistical veto method using auxiliary channels sensitive to gravitational wave signals", submitted to Class. Quantum Grav.

US006417597B1

(12) **United States Patent**
Baker, Jr.

(10) **Patent No.: US 6,417,597 B1**
(45) **Date of Patent: Jul. 9, 2002**

(54) **GRAVITATIONAL WAVE GENERATOR**

(75) Inventor: **Robert M. L. Baker, Jr.**, 8123 Tuscany Ave., Playa del Rey, CA (US) 90293

(73) Assignees: **Robert M. L. Baker, Jr.**; **Bonnie Baker**, Playa Del Rey, CA (US); Trustees of the Dr. Robert and Bonnie Baker, Jr. Family Trust Dated Oct. 21, 1992 in part interest

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **09/616,683**

(22) Filed: **Jul. 14, 2000**

Related U.S. Application Data

(63) Continuation-in-part of application No. 09/443,527, filed on Nov. 19, 1999, now Pat. No. 6,160,336.

(51) **Int. Cl.**[7] **G21H 3/00**; G21H 1/00

(52) **U.S. Cl.** **310/300**; 310/311; 700/286; 976/DIG. 403; 976/DIG. 405

(58) **Field of Search** ..... 310/300, 311; 700/286; 976/DIG. 403, DIG. 405; 376/153

(56) **References Cited**

U.S. PATENT DOCUMENTS

2,719,944 A		10/1955	Brailsford	318/254
2,814,769 A		11/1957	Williams	318/171
3,612,630 A		10/1971	Rosensweig	308/10
3,667,019 A		5/1972	Elliott et al.	318/254
3,722,288 A	*	3/1973	Weber	73/382
3,903,463 A		9/1975	Kanamori	318/138
3,959,700 A		5/1976	Sugiura et al.	318/138
4,035,658 A		7/1977	Diggs	290/55
4,052,134 A		10/1977	Rumsey	416/119
4,086,505 A		4/1978	McDonald	310/74
4,546,264 A	*	10/1985	Pinson	290/54
4,874,942 A	*	10/1989	Clauser	250/251
5,398,571 A		3/1995	Lewis	74/572
5,446,018 A		8/1995	Takahata et al.	310/90.5
5,495,515 A	*	2/1996	Imasaki	378/119
5,513,530 A	*	5/1996	Ney et al.	73/382 R
5,514,923 A	*	5/1996	Gossler et al.	310/74
5,646,728 A		7/1997	Coutsomitros	356/352
5,721,461 A	*	2/1998	Taylor	310/268
5,831,362 A		11/1998	Chu et al.	310/90.5
5,929,579 A		7/1999	Hsu	318/439

FOREIGN PATENT DOCUMENTS

GB	1333343	10/1973	B01D/3/06

OTHER PUBLICATIONS

Chakrabarty, "Gravitational Waves: An Introduction," Aug. 21, 1999, Physics, pp. 1–21.*

J. Weber, "Gravitational Waves" in *Gravitation and Relativity*, Chapter 5, pp. 90–105 (W.A. Benjamin, Inc., New York, 1964).

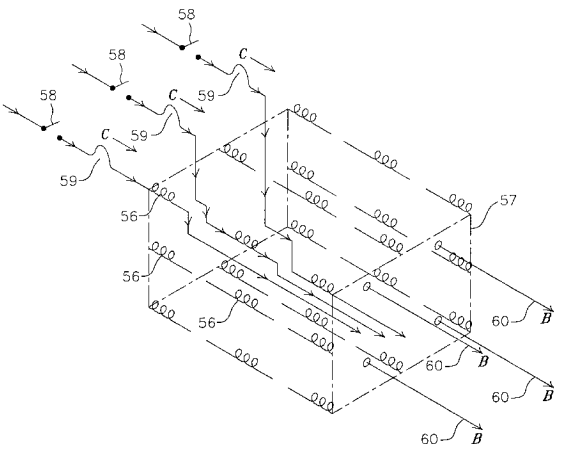
(List continued on next page.)

Primary Examiner—Burton S. Mullins
(74) *Attorney, Agent, or Firm*—Christie, Parker & Hale, LLP

(57) **ABSTRACT**

A gravitational wave generation device comprising an ensemble of energizable elements involving magnetic, electrical and electromechanical functions that are under the control of a computer and attendant computer software system. The magnetic and electrical force elements, when energized as directed by the computer, operate in concert to produce a rapid third-time-derivative motion of a mass. This action causes the generation of high-frequency gravitational waves that can be modulated and shaped in order to be utilized for communication, propulsion, and various physics experiments. The energizable elements can be very small coils or coil sets encased in a computer chip, current-carrying conductors, or small electromechanical devices. The mass acted upon by the coil elements can be a permanent magnet or magnets, or electromagnets. In the electromechanical-element configuration the device can be used both for the generation of gravitational waves and their detection.

50 Claims, 5 Drawing Sheets



OTHER PUBLICATIONS

J. Weber, "Detection and Generation of Gravitational Waves", *Physical Review*, (1960) vol. 117, No. 1, pp. 306–313.

J. Weber, "Gravitational Radiation from the Pulsars", *Physical Review Letters*, (1968) vol. 21, No. 6, pp. 395–396.

Robert L. Forward and Larry R. Miller "Generation and Detection of Dynamic Gravitational–Gradient Fields", Hughes Research Laboratories, Aug. 5, 1966, pp. 512–518.

F. Romero B., et al., Generation of Gravitational Radiation in the Laboratory, Fakultät für Physik der Universität Konstanz, Z. Naturforsch 36a, 948–955 (1981).

Richard, Jean–Paul, Recent developments in the measurement of space time curvature, Acta Astronautica, vol. 5, pp. 63–76. Pergamon Press 1978.

* cited by examiner-

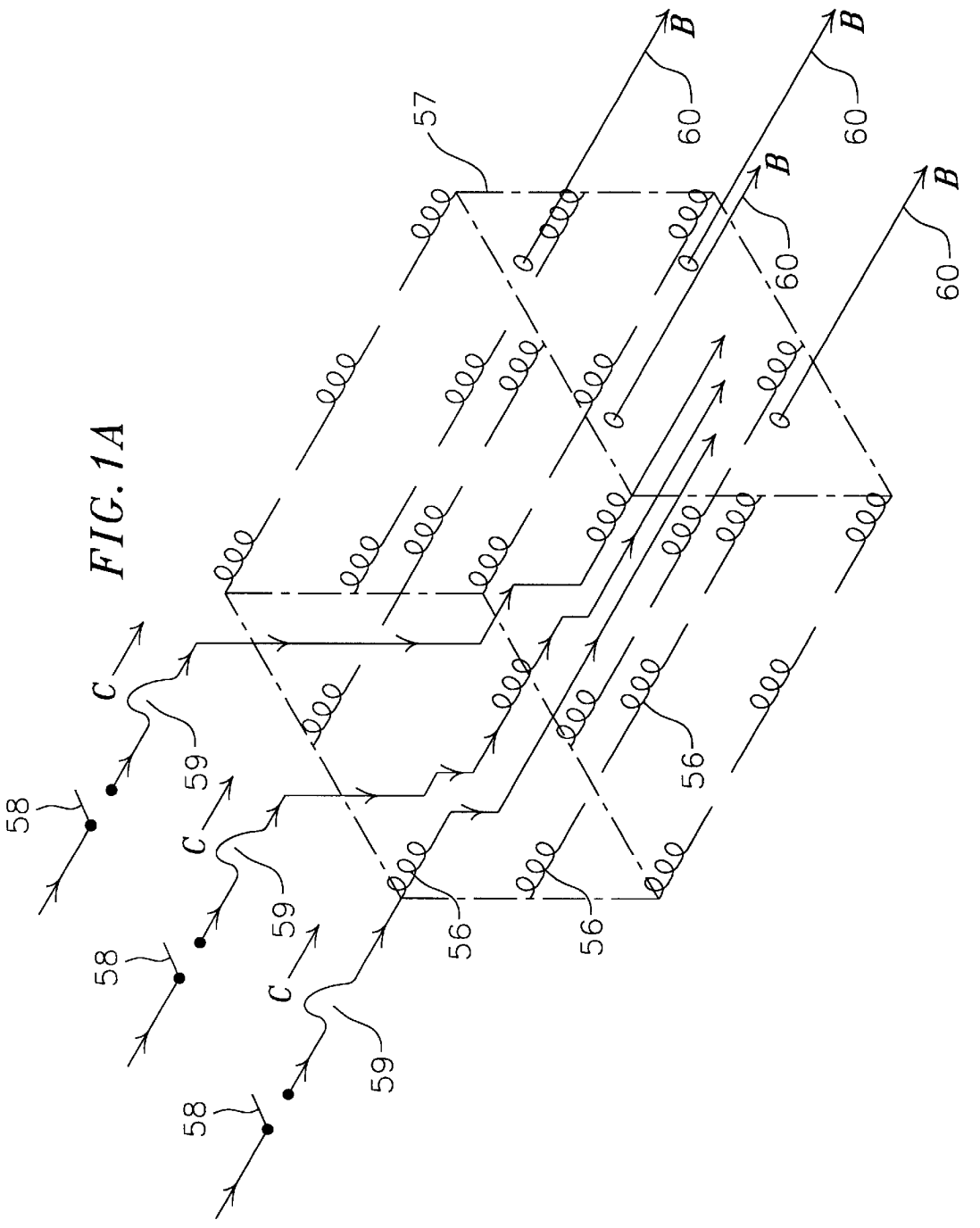


FIG. 1A

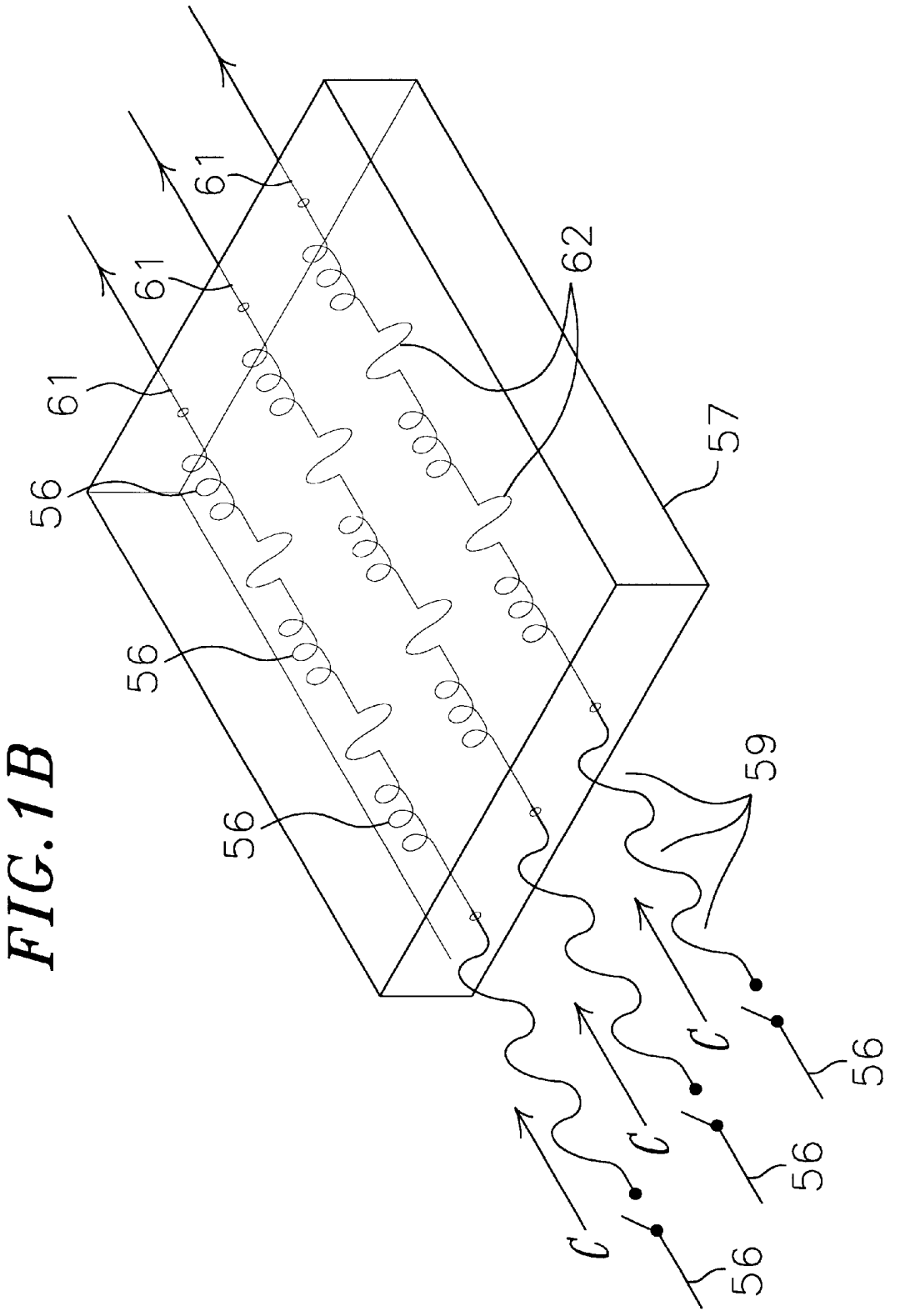


FIG. 1B

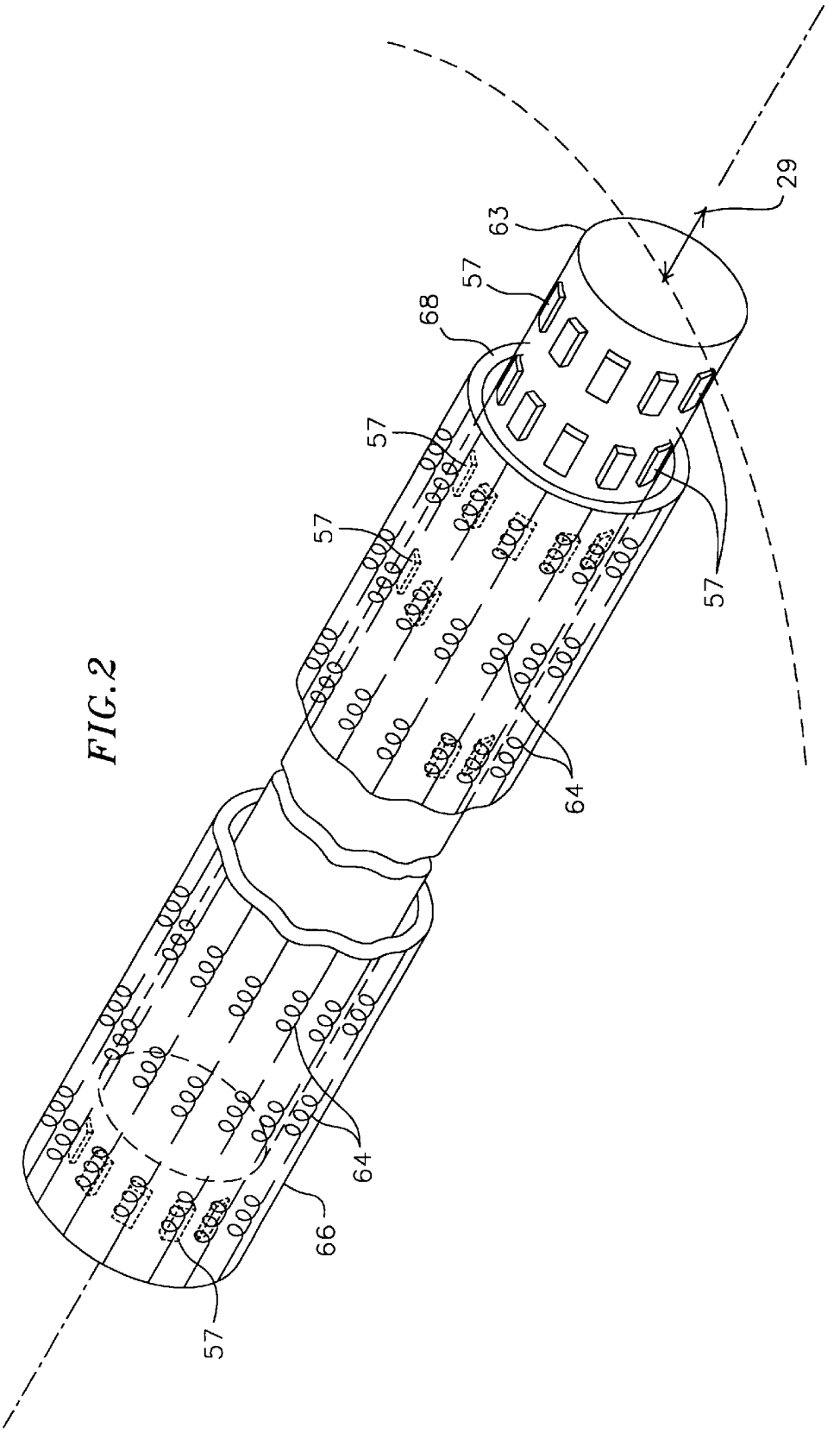


FIG. 2

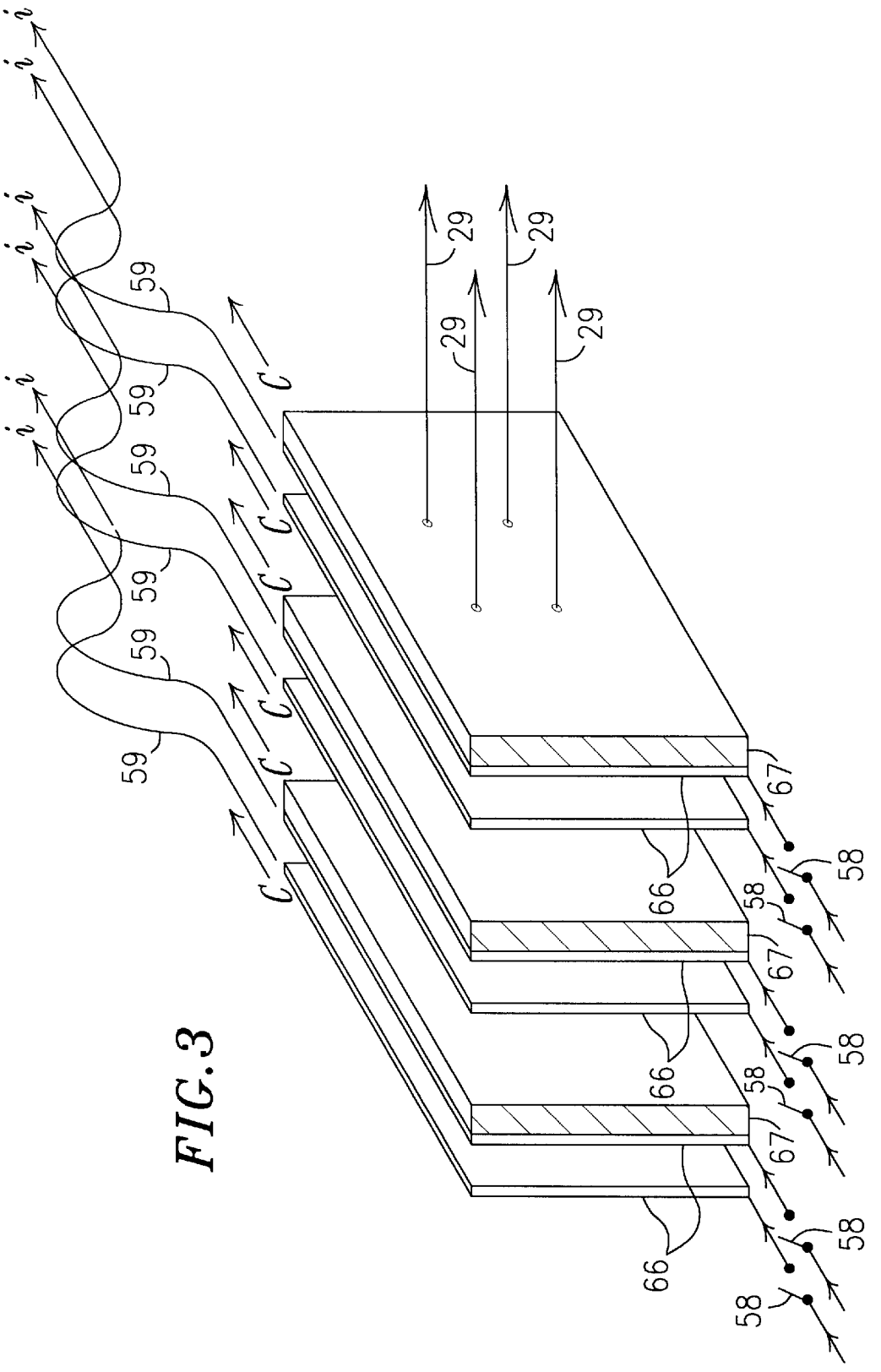


FIG. 3

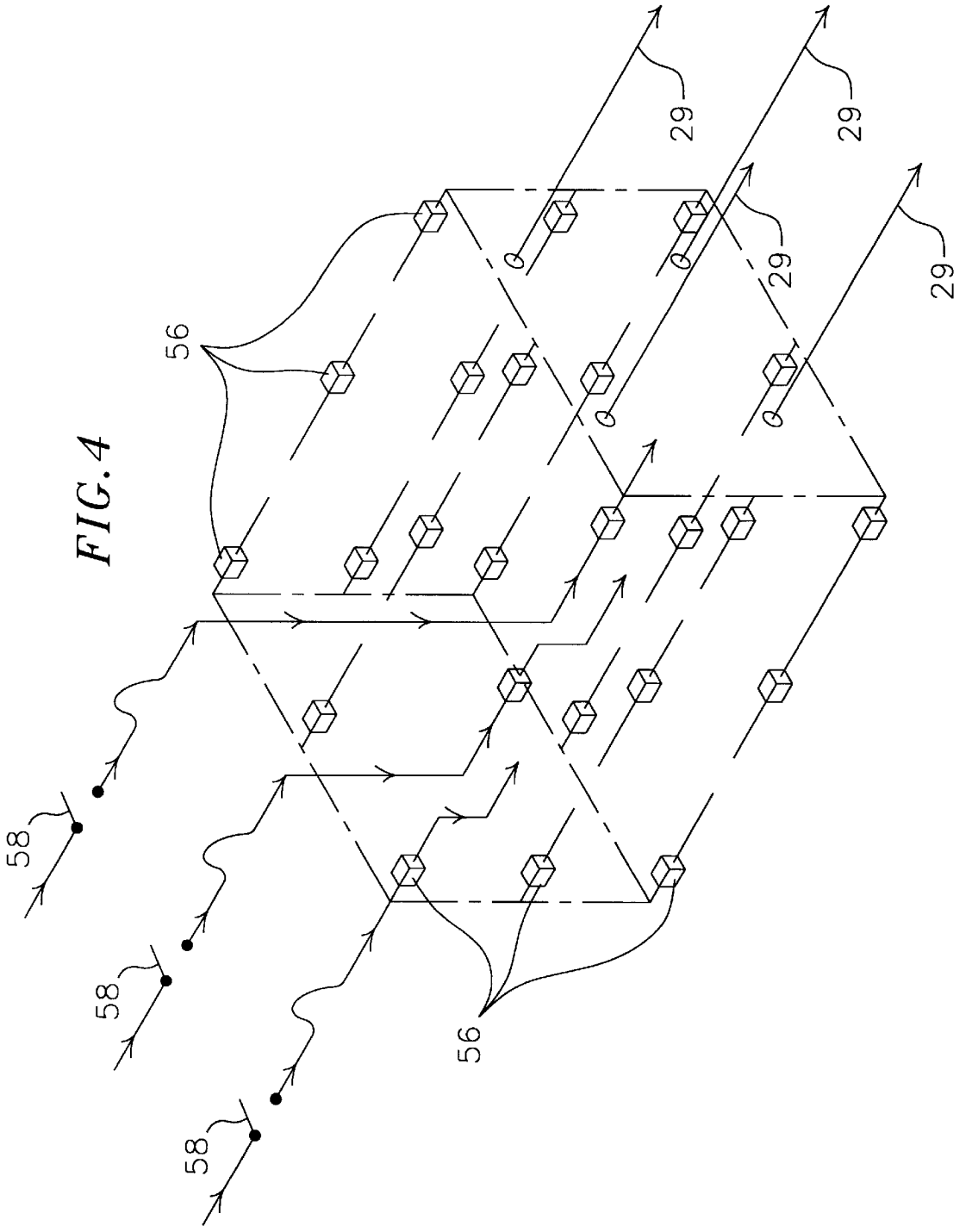


FIG. 4

GRAVITATIONAL WAVE GENERATOR

REFERENCE TO RELATED INVENTION

The present application is a continuation-in-part of application Ser. No. 09/443,527, filed Nov. 19, 1999, now U.S. Pat. No. 6,160,336 entitled Peak Power Energy Storage Device and Gravitational Wave Generator which is incorporated herein by reference.

BACKGROUND OF THE INVENTION

This invention relates to the generation of high-frequency gravitational waves that can be modulated and utilized for communication, propulsion and for the purpose of testing new physical theories, concepts, and conjectures. More particularly this invention relates to a combination of small energizable elements operatively connected to a computer working in concert to produce a third-time-derivative motion of a nearby mass or electromechanical elements that generates a high-frequency gravitational wave. The invention also relates to the detection of gravitational waves.

DESCRIPTION OF THE PRIOR ART

Albert Einstein in his General Theory of Relativity predicted gravitational waves (GW). Such waves have never been detected, but an extra-terrestrial source of low-frequency GW, namely a neutron double star pair, has been observed to coalesce at a rate exactly as predicted if it radiated GW. The production of GW having a high frequency and generated by relatively strong magnetic, electrical and electromechanical forces acting on relatively small masses rather than by the relatively weak gravitational forces acting on large celestial masses is not known.

The prior art indicates that gravitational-wave generators are theorized although not reduced to practice. Joseph Weber in "Detection and Generation of Gravitational Waves", *Physical Review*, Volume 117, Number 1, January, 1960, p. 313 has proposed electromechanical-force-produced GW by use of piezoelectric crystals: "Waves one meter long could be radiated by a crystal with dimensions about fifty centimeters on a side. If it is driven just below the breaking point, each crystal would radiate $\approx 10^{-20}$ [watts], assuming P_{max} to be its static published value." No one has, however, put such a device in practice.

According to Robert L. Forward a gravitational-wave generator could be constructed by means of a tube in which very dense Newtonium (element 127) is caused to move up-and-down the tube at high-speed (not really a dipole since as Joseph Weber indicated {"Gravitational Waves" in *Gravitation and Relativity*, Chapter 5,W. A. Benjamin, Inc., New York, 1964, p. 91} the lowest order of gravitational radiation from a system cannot be a dipole, but must be a quadrupole). However no drawings or other descriptions of such a generator are known that are sufficiently specific to enable a person skilled in the art to practice the generator.

As described in the parent patent, there is considerable prior art in the detection of gravitational waves including U.S. Pat. No. 5,646,728, which involves a very low amplitude interferometer instrument suitable for detecting gravitational waves. This prior art involves only the detection of low-frequency (below a MHz) gravitational waves generated by natural processes, such as astrophysical or celestial events. It is believed that there is no prior art in the detection of high-frequency gravitational waves that are artificially generated for communication or other purposes.

SUMMARY OF THE INVENTION

The present invention relies upon the fact that the rapid movement or "jerk" of a mass or the rapid change or "jerk" in angular momentum with time, over a period of time such as a picosecond caused by the operation of the present invention, will produce a quadrupole moment and generate a sequence or train of useful high-frequency, for example Tera Hertz (THz), gravitational waves (GW). The device described will accomplish this GW generation in several alternative ways based upon the device's rotating and non-rotating, symmetrical and non-symmetrical masses acted upon by means of relatively strong magnetic, electrical and electromechanical forces. Such forces are produced by an ensemble of very small, sub-millimeter, energizable elements operating in concert under the control of the device's computer. As noted, this process is substantially different from the extra-terrestrial generation of low-frequency GW by very large rotating and non-rotating celestial masses acted upon by relatively weak gravitational forces.

Gravitational waves are absorbed differently and propagate differently through matter and space than do electromagnetic waves. Thus gravitational waves may offer advantages over electromagnetic waves in that they can be transmitted through material opaque to electromagnetic waves and their intensity may fall off less rapidly with distance than electromagnetic waves. Up to now man-made machines in use do not generate significant or measurable GW.

BRIEF DESCRIPTION OF THE DRAWINGS

The foregoing features and advantages of the present invention will be more fully understood by reference to the following detailed description of the invention when considered in connection with the accompanying drawings in which:

FIG. **1**A and FIG. **1**B are schematic views of coil sets **56** embedded in or imprinted on a silicon chip **57** and connected to transistors or ultra-fast switches **58** in order to generate current pulses **59** in the magnetic field of the permanent magnets **60**;

FIG. **2** is a schematic view of a single-sector or linear-motor embodiment of a gravitational-wave-generator device. The sector refers to the sector of the annulus of the plan view of the rim of the spindle device shown in the parent patent, U.S. Pat. No. 6,160,336 and is, therefore, a single, isolated segment of the rim. A sheath **68** of coils, embedded in semiconductor chips **57** individually or collectively connected to computer controlled transistor or ultra-fast switches, surrounds a central magnetic-mass composed of magnetic sites **57** incorporated with core, piston, or barrel **63**;

FIG. **3** is a schematic view of infinite-radius coils evolved into parallel plate conductors **66**, which may have ballast **67** attached in order to vary their masses, connected to fast switches **58** in order to generate current pulses **59** acting in concert to generate a train of gravitational waves **29** (which propagate in both directions from the jerk since there is a square associated with the kernel of the quadrupole equation (3)—so there is no preferred direction along the axis of the jerk); and

FIG. **4** is a schematic view of the utilization of electromechanical elements **65** that act in concert to generate and detect gravitational waves **29**.

Analysis of Binary Pulsar PSR 1913+16

Since the observation of the binary pulsar PSR 1913+16 represents the only experimental confirmation of GW, the features and advantages of the present invention will be

better understood by the analyses of such a double-star system. This pair of neutron stars will coalesce in 3.5×10^8 years due to GW radiation and produce a rather continuous GW until that time. According to J. H. Taylor, Jr. in "Binary pulsars and relativistic gravity", *Reviews of Modern Physics*, Volume 66, Number 3, July, 1994, pp. 711–719, the period of their mutual rotation is 7.75 hours (or 2.79×10^4 [sec]), periastron is 1.1 solar radii (one solar radius is 6.965×10^8 [m]), and apastron is 4.8 solar radii. It's radius of gyration is essentially the semi-major axis=(1.1+4.8)/2=2.95 solar radii=(2.95) $(6.965\times10^8)=2.05\times10^9$ [m]. The neutron stars exhibit a mass of about 1.4 solar masses so that together their mass is (2) (1.4) $(1.987\times10^{30})=5.56\times10^{30}$ [kg]. Thus the moment of inertia of the binary-pulsar system is approximately (5.56×10^{30}) $(2.05\times10^9)^2=2.34\times10^{49}$ [kg-m^2]. The current angular rate of the system=$2\pi/2.79\times10^4=2.25\times10^{-4}$ [radians/sec]. Thus the angular momentum of the system is currently, $I\omega=(2.34\times10^{49})$ $(2.25\times10^{-4})=5.27\times10^{45}$ [kg-m^2/sec]. According to a perusal of binary-star catalogs by John Mosley of the Griffith Observatory, the binary pulsar PSR 1913+16 is at a distance from our Sun of 23,300 light years. If there was little or no GW diffraction, then the reference area is a circular ribbon or strip having a width of the diameter of a neutron star, 3×10^4 [m]. Thus the reference area equals (3×10^4) (2π) (2.33×10^4) $(9.5\times10^{15}$ meters per light year)=4.17×10^{25} [m^2].

In the case of a binary star pair such as PSR 1913+16 the GW power, P, is computed from the quadrupole moment, which for two masses on orbit about one another is given, for example, by an equation on p. 356 of L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields*, Forth Revised English Edition, Pergamon Press, 1975. They give the time-variable factor in P as a function of the true anomaly, v, and orbital eccentricity, e, as

$$(1+e\cos v)^4([1+\{e/12\}\cos v]^2+e^2\sin^2 v)/(a[1-e^2])^5. \quad (1)$$

In conventional astrodynamic/celestial-mechanics notation this factor is

$$p/r^6+(dr/d\tau)^2/12\mu r^4, \quad (2)$$

where p is the orbital "parameter" or semilatus rectum [AU], r is the radial distance between the two masses [AU], τ is the characteristic time measured in $k_s$days or in units of 5.0022×10^6 [s] for heliocentric-unit systems, and μ is the sum of the two masses [solar masses]. Note that one AU (astronomical unit)=1.496×10^{11} [m] and one solar mass=1.987×10^{30} [kg]. The $dr/d\tau$ term is related to $dI/d\tau$ $(=2\mu t[dr/d\tau])$, $d^3I/d\tau^3$ $(=-2\mu^2[dr/d\tau]/r^2)$, $d^2v/d\tau^2$ $(=-\sqrt{\mu p}[dr/d\tau]/r^3)$, and $d^3v/d\tau^3$ $(=-2\mu\sqrt{\mu p}[1/r-1/a-4\{dr/d\tau\}^2/\mu]/r^4)$, where a=the semi-major axis of the orbit [AU] and for a circular orbit $dr/d\tau=0$. These time derivatives are directly related to the embodiments of the invention.

The GW power radiated, P, which causes a perturbation in the semi-major axis, a, (and an attendant secular decrease in the orbital period) is obtained by integrating the time-variable factor, Eq. (2), over the orbital period using the mean anomaly, M, which is directly proportional to the time (that is, M=n [t-T], where n is the mean motion [=ω in Landau and Lifshitz's {ibid, p. 357} notation] and T is the time of periastron passage). The value of the average GW power, P, is computed from observations that define the eccentricity (based primarily upon Doppler-shift determination of the range rate at periastron and apastron), semi-major axis, and orbital orientation angles of PSR 1913+16. The error in the computed value of P is related to the observational error leading to the determination of the orbital elements as well as the determination of the masses of the pair of neutron stars, $\mu=m_1+m_2=1.4+1.4=2.8$ [solar masses]. For example, a 0.1 percent change in the measurement of range rate at periastron results in a 0.28 percent change in GW power, P, and a 0.1 percent change in the mass of the stars results in a 0.33 percent change in GW power.

The observed accumulated shift in the times of periastron passage, T, caused by the secular shortening of the orbital period of PSR 1913+16, compares closely, within observational error, to that predicted by General Relativity and confirms the existence of GW radiation. Likewise is confirmed the existence of a $dr/d\tau$ component, which is related to $d^2\omega/dt^2$ $(\approx d^3v/d\tau^3)$ and d^3I/dt^3, that are involved in the GW-generator embodiments of the invention

The average GW power, P, established by Landau and Lifshitz (ibid, p. 357) by analytically integrating and given as a function of eccentricity, e, is for e=0.641, 9.28×10^{24} [watts] so for the disk reference area the GW flux at the Sun's distance=$9.28\times10^{24}/4.17\times10^{25}=0.222$ [watts/m^2]. If the GW is totally diffracted and the propagation is spherically isotropic, then the GW flux at the Sun's distance=$9.28\times10^{24}/(4\pi[2.33\times10^4\times9.5\times10^{15}]^2)=1.51\times10^{-17}$ [watts/m^2]. By numerically integrating (see, for example, an algorithm found in R. L. M. Baker, Jr., *Astrodynamics, Applications and Advanced Topics*, Academic Press, New York , 1967, pp. 263–272) over the mean anomaly (directly proportional to time) the average GW power, P, is 9.296×10^{24} [watts]. The peak GW power, 1.73×10^{26} [watts] occurs at the time of periastron passage (every 7.75 hours) and at the Sun's distance would result in a GW flux of $1.73\times10^{26}/(4\pi[2.33\times10^4\times9.5\times10^{15}]^2)=2.81\times10^{-16}$ [watts/m^2] if the GW were totally diffracted and the propagation were spherically isotropic. If GW detectors were sensitive enough to detect such an intensity and they did not, then it would lend credence to the disk-like propagation of GW waves or at least to diffraction of GW less than 45 degrees from the orbit plane of PSR 1913+16 at the Sun's distance.

Individual Independently Programmable Coil System (IIPCS)

Of fundamental importance to the operation of the present invention is the Individual Independently Programmable Coil System (IIPCS) described and illustrated in parent patent U.S. Pat. No. 6,160,336. This system, is enabled by a computer and associated computer software, to control a system of either transistors or of ultra-fast switches. The switches rapidly turn off or on a myriad of sub-millimeter coils, current-carrying conductors and/or electromechanical elements and generate magnetic (or electromechanical) force to produce a third time derivative or "jerk" or, alternatively, a harmonic oscillation of a mass or masses.

For a very large number of ultra-small, sub-millimeter coil elements involved in some of the embodiments of this invention a miniaturized integrated circuit can be utilized. They are embedded in or imprinted on a silicon chip, organic material, or in connection with polymer-based devices. They consist of multiple layers, with appropriate sequencing time delays to ensure near simultaneity of the magnetic field's interaction with individual elements, as in FIG. **1A**, or as the direct-current train of approximately one-picosecond pulses, which traverse each coil set on the chip levels, as in FIG. **1B** and possibly integrated in the chip with the ultra-fast switches or transistors or other semi-conductors. Although switches or transistors having picosecond capability are utilized in the various numerical examples, much slower switches or transistors could be utilized to successfully practice the invention. A preferred embodiment utilizes

conventional computer chips, containing the IIPCS circuit elements about 18 micrometers or less apart, synchronizing clock, input/output ports, and sub-millimeter coils on 50 to 100 micrometer centers. The chips are about 6 mm to 9 mm square and are obtained from silicon wafers. These chips are sewn into a circuit-board roll with an approximately 25-micrometer-diameter gold thread. Several layers of this roll (for example, 25) are connected in a fixed location or band adjacent to the moving spindle's rim and form the IIPCS in the spindle rim's magnetic field. Such rolls are routinely fabricated by French-owned Oberthur Card Systems (plant in Rancho Dominguez, Calif.), French-based Gemplus, Schlumberger (Paris and New York), and California-based Frost & Sullivan.

In the miniaturized integrated circuit situation, as exhibited in FIG. **1**A, there will be a very large number of small, sub-millimeter coil sets or elements **56** embedded in or imprinted upon a silicon chip **57** in multiple layers. Ultra-fast switches or transistors of the IIPCS **58** will launch a series of current pulses **59** of approximately picosecond duration moving at the electron's mobility speed, c, that will be timed to reach the individual coil sets or elements almost simultaneously along several individual wire conductors, as in FIG. **1**A, or one single wire conductor per line, as in FIG. **1**B, and thereby interact with the magnetic field **60** in concert. This interaction will result in a third-time-derivative motion or jerk of the magnetic mass to generate a train of gravitational waves. The ultra-fast switches are preferably semiconductor-based, such as a semiconductor optical amplifier (SOA) or a semiconductor nonlinear interferometer such as a nonlinear Sagnac interferometer on a phosphide semiconductor chip, etc. (see, for example, D. Cotter, et al, "Non-linear Optics for High-speed Digital Information Processing", *Science*, Volume 286, Nov. 19, 1999, pp. 1523–1528). In FIG. **1**B, the IIPCS and its array of ultra-fast switches is programmed to launch a train of current pulses or intervals of approximately a picosecond duration **59** such that each member of the pulse train will reach each of the coils or coil sets at the same time. The pulse train can consist of stretches of constant or zero current. The duration of the pulses will be such to completely energize any given coil set as it passes through it in order to produce a magnetic field interaction. The interaction will result in a third-time-derivative lateral motion or jerk of a cylindrical, central magnetic core **63** shown in FIG. **2** and, as will be discussed, generate a GW train **29**, which propagates both in the direction of and opposite to the direction of the jerk as illustrated in FIGS. **8**A and **8**B of U.S. Pat. No. 6,160,336. This core, piston, or barrel **63** is surrounded by and immediately adjacent to a sheath **68** of IIPCS-controlled coil sets **64**. In the case of the current-pulse train on a single conductor wire interconnecting a line of coils or coil sets, there will be a build up of impulses to full value as the current-impulse train progresses down the line of coil sets. Use of a single conductor wire for each line of coils or coil sets reduces the resistive power loss. In order to transmit information, all pulses in the train may not be present or they may be at different amplitude thereby modulating the GW. Portions or stretches of the pulse train could also be intervals of constant current. In each line of coils set in series along one conductor wire **61** there will be time delays **62** between coil sets to ensure simultaneity of the current pulses reaching any given coil set.

In FIG. **3**, ultra-fast switches or transistors of the IIPCS **58** will launch a series of current pulses acting in either direction **59** of approximately picosecond duration moving at the electron's mobility speed, c, along individual conductors or single interconnecting conductor wires in order to produce current pulses **59**, acting in concert to generate a modulated gravitational wave **29**. The current pulses will be timed to reach individual parallel-plate conductors **66**, which may have different masses or may have ballast **67** attached and/or carry different current and/or have different modulus of elasticity and/or are constructed differently in their mounting for the purpose of exhibiting high-frequency asymmetrical mass displacements.

In FIG. **4**, ultra-fast switches or transistors of the IIPCS **58** will launch a series of current pulses **59** of approximately picosecond duration moving at the electron's mobility speed, c, along individual conductors or single interconnecting conductor wires that will be times to reach individual, sub-millimeter electrical or electromechanical-force elements **55** in sequence to reinforce the build up or modulation of a GW beam **29**. The ensemble of electrical or electromagnetic force elements will be embedded in or imprinted on a silicon chip **57** in multiple layers.

Quadrupole Moment

Although the specific relationship for GW generation will be an outcome of the experimental use of the invention; as an example of that relationship consider the standard GW quadrupole Eq. (110.16), p. 355 of L. D. Landau and E. M. Lifshitz (opcit) or Eq. (1), p. 463 of J. P. Ostriker, ("Astrophysical Sources of Gravitational Radiation" in *Sources of Gravitational Radiation*, Edited by L. L. Smarr, Cambridge University Press, 1979) which gives the GW radiated power [watts] as

$$P=-dE/dt=-(G/45c^5)\,(d^3D_{\alpha\beta}/dt^3)^2\ [\text{watts}] \qquad (3)$$

where

E=energy [joules],

t=time [s],

$G=6.67423\times10^{-11}$ [m^3/kg-s^2] (universal gravitational constant),

$c=3\times10^8$ [m/s] (the speed of light, approximately the electron's mobility speed in the conductor), and

$D_{\alpha\beta}$ [kg-m^2] is the quadrupole moment-of-inertia tensor of the mass of the device, and the

α and β subscripts signify the tensor components and directions.

Note especially the third time derivative in the squared term or kernel (that is, $(d^3D_{\alpha\beta/dt}{}^3)^2$) of Eq.)3). Such a time-rate-of-change of the second derivative ("acceleration") is often referred to as a "jerk". Because the factor of this kernel is so small, 1.76×10^{-52}, the kernel and hence the jerk must be very large. In the following examples of GW generation by various devices, we shall often cite astrophysical analyses of the same GW formulation. It should be recognized, however, that although kernels are analogous in the invention and in the celestial astrophysical systems (or events) their operation is quite different. In most cases the astrophysically generated GW rely on rather slow-moving, low-frequency events (a fraction of a Hertz to possibly a MHz) and weak gravitational forces. On the other hand, the various embodiments of the invention rely on vary fast-moving, high-frequency events (in the THz range) and relatively strong magnetic, electrical or electromechanical forces.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

In order to understand the various embodiments of the invention it is useful to refer to the historical roots of GW

generation. From Eq. (1), p. 90 of Joseph Weber (opcit, 1964) one has for Einstein's original formulation of the gravitational-wave (GW) radiated power of a rod spinning about an axis through its midpoint, having a time-constant moment of inertia, I [kg-m^2], and an angular rate, ω [radians/s]:

$$P=-(G/45c^5)\ ([\sqrt{\{2\times12\}}][I\omega^3])^{2=-G(I\omega3)^2}/5(c/2)^5[\text{watts}]. \quad (4)$$

Equation (4) is an approximation and is, strictly speaking, only valid for lengths very much less than the GW wavelength and speeds of the rod's rotation less than c.

This is the same equation as that given for two bodies on a circular orbit (also exhibiting a time-constant moment of inertia) on p. 356 of Landau and Lifshitz, opcit ($I=\mu r^2$ in their notation) where ω=n, the orbital mean motion, and similar to all of the equations associated with the various embodiments of the invention. Although Eqs. (3) and (4) result from relativistic mechanics, classical mechanics (such as the use of the conventional moment of inertia in Eq. (4)) will be utilized herein to provide useful results.

Spindle-Device GW Generation Embodiment

It is reasonable (by appealing to simplicity, that is, by Ockham's Razor) to suggest that for the spin up/down of a spindle device, as referred to in the parent application, (referred to herein as the "$(Id^2\omega/dt^2)^2$ formulation or component"):

$$P=-G\ \kappa_{I\omega2dot}(Id^2\omega/dt^2)^2/5(c/2)^5[\text{watts}] \quad (5)$$

where

$\kappa_{I\omega2dot}$=a dimensionless constant or function to be established by experiment and

$d^2\omega/dt^2$=second time derivative of the spindle's angular velocity, ω, or third time derivative of it's angle, termed, a "jerk". In fact, as noted by M. S. Turner and R. V. Wagoner "Gravitational Radiation from Slowly Rotating 'Supernova' Preliminary Results," in *Sources of Gravitational Radiation*, Edited by L. L. Smarr, Cambridge University Press , 1979, p. 383 that "If the angular velocity ω . . . is non-uniform, octupole (post-Newtonian) radiation is generated (in addition to the quadrupole (Newtonian) radiation . . . " (emphasis added) and on p. 385 they state "This radiation is generated not by non-spherical distribution of matter . . . , but by internal motions."

This third derivative, $d^2\omega/dt^2$, is computed by introducing the equation of motion for a rotating body

$$Id\omega/dt=rf \quad (6)$$

where

r=radius of the spindle's rim [m] and

f=force tangential to the rim [N].

The derivative is approximated by

$$Id^2\omega/dt^2\cong\Delta(Id\omega/dt)/\Delta t=\Delta(rf)/\Delta t=r\Delta f/\Delta t; \quad (7)$$

in which Δf is the nearly instantaneous increase in the force tangential to the rim or jerk caused by the magnetic field when it is turned on or turned off or pulsed by the transistors or ultra-fast switches of the Individual Independently Programmable Coil System (IIPCS), that is, a tangential jerk. Thus

$$P=-1.76\times10^{-52}(\kappa_{I\omega2dot}r\Delta f/\Delta t)^2[\text{watts}]. \quad (8)$$

(1) Numerical Example

As a numerical example, (for a spindle GW-generation device slightly different from the exemplar spindle shown in U.S. Pat. No. 6,160,336) set $\kappa_{I\omega2dot}=1$ (subject to experimental determination later), r=1000 [m], $\Delta f=1.8\times10^7$ [N], and $\Delta t=10^{-12}$ [s]. These numbers arise as follows: The rim is a thin (approximately one cm thick) band of Alnico 5 permanent magnets (or electromagnets) facing radially outward. In general, permanent magnets exhibit irregular magnetic fields and associated forces. As a rule of thumb such a band of juxtaposed magnets will produce in excess of 30 pounds per 1.75 inches (or 206 pounds per foot) of tangential rim force. Each 1.75-inch magnet has a flux density, B, of about 2,600 gauss or 0.26 [Tesla] developed every 4.4 cm. The kilometer-radius rim is a large hoop connected to a central spindle/hub as described in the parent patent. The IIPCS coil sets at the rim's periphery, when switched on generate a 0.26 [Tesla] flux density every 0.044 [m] and produce in excess of a 200 pound per foot or 3000 [N/m], which is defined as Δf/Δl, or impulse of tangential force every meter on the rim (that is, a force built-up almost to full value during spin up in approximately a picosecond and a similar build up of retarding force during spin down) . Since the rim's circumference is 2π (1000) (3.28 feet per meter)= 20,600 feet, the tangential rim force produced when the coils are fully energized is 4.1×10^6 pounds or 1.8×10^7 [N]. The 10^{-12} [s] intervals, with the coils turned turn on and then off, will generate a train of direct-current, approximately one-picosecond pulses. In each line of coils there will be an ultra-fast switch (such switches could be located near to the coils and each one energizing a large number of coil sets or else co-located with a central IIPCS control computer).

Inserting the numbers in Eq. (8) for the spindle's gravitational-wave (GW) power for the tangential jerk yields

$$P=-1.76\times10^{-52}(2\times10^3\times1.8\times10^7/10^{-12})^2=-2.3\times10^{-7}[\text{watts}]. \quad (9)$$

The reference area of the 1 cm thick rim is (0.01) 2π(1000)= 63 [m^2], so that the GW energy flux is 3×10^{-9} [watts/m^2]. For a time-constant value of I there may be a somewhat less simple, $\kappa_{I\omega d\omega\omega}([d\omega/dt]\omega)^2$, formulation or component of the GW power for spin up/down, but as will be seen, for most applications it is expected to result in a smaller power than the larger of the $\kappa_{I\omega2dot}(Id^2\omega/dt^2)^2$ or $(I\omega^3)^2$ formulations or components.

Note that the coil sets must be very close together. In order for the coils fields to interact with the whole rim's magnetic field and impart the mechanical impulse or jerk, they must be spaced no more than 0.3 mm or 300 [micrometers] apart (the distance light and, hence, the magnetic field and resulting impulse on the permanent magnets, travels in a picosecond). If all coil sets in a line of coil sets are connected in series by the same conductor, then each member of the pulse train traverses a 300-micrometer-length coil set, separated from the next coil set by a time delay circuit. Such a time-delay circuit could be simply a 300-micrometer-long jumper (see **62** in FIG. **1**B) between coil sets. In this connection it is noted that if each coil set is connected by its own unique conductors as in FIG. **1**A, instead of one single conductor wire along each line of coils, then the communications lines or conductors to all of the coil switches from the logic circuits of the control computer must be equal to better than 0.01 mm or 10 [micrometers] in order to ensure near simultaneity or proper timing. That is, the electrons must reach all of the coils sets at the appropriate time in less than approximately a fraction of a picosecond of time difference.

(2) Magnetic Field Build Up and Heat Loss

Although of little concern in most applications, the length of time to "build-up" the magnetic field of the coils is

important here. The electrons must complete sufficient coil turns (moving at the electron's mobility speed or about light speed) in approximately a picosecond to "launch" most of the magnetic field that produces the impulsive force, "hammer blow" or jerk when it interacts with the static magnetic field of the permanent or electromagnets carried around by the rim or other magnetic mass. Thus, they must be very tightly wound with each coil "set" having a total length of less than 0.3 mm (0.0003[m] or 300 micrometers). If each of the ultra-small, sub-millimeter coil sets consist of two coils or turns, as exhibited in FIGS. **7**A, **7**B, **7**C, **7**D, **7**E, **7**F, **7**G, and **7**F, of the parent patent, then their diameters are on the order of d=0.3/2π=0.05[mm]=50 [micrometers] or less. The coil wire could be made of gold having about a 0.015-mm or 15-micrometer diameter. The resistance for such wire at room temperature is about 135 [ohms/m]—high-temperature superconducting material would be useful here. In the spin-up mode the IIPCS will need to build up 0.26 [Tesla] flux density every 0.044 [m] (the requirements for the spin-down mode are essentially the same, but reversed). Thus

$$B=\mu_o ni/l[\text{Tesla}] \tag{10}$$

where $\mu_o=4\pi\times10^{-7}$ (permeability of free space), n is the number of coil turns, i is the current through the coils [amps], and l is the length of the coil conductors [m]. The double coil sets will be placed on 50 to 100 [micrometer] centers, so that there will be about $2\times100\times100=2\times10^4$ coil turns on each square-centimeter level of the stack of 25 coil levels or layers. With l=0.044 [m] and B=0.26 [Tesla], $ni=9.1\times10^3$ [amp turns]. For $n=25\times2\times10^4=5\times10^5$, $i=9.1\times10^3/5\times10^5=0.018$ [amps] or 18 milliamperes. The total length of 15-micrometer-diameter gold wire across any given layer or level is 100(rows)×100(coil & jumper/time-delays)×(600 micrometers)=6 [m]. For the 25 layers or levels there will 150 [m] of wire with a resistance of $150[m]\times135[ohms/m]=2.025\times10^4$ [ohms]. Since on average every other pulse or current interval across a conductor wire will carry no current, since in order to modulate the GW some pulses or current intervals in the train will be missing, the heat loss per centimeter of chip stack or semi-conductor layers is

$$(1/2)i^2R=3.28[\text{watts}]. \tag{11}$$

This heat loss can be reduced by 32% by using 25-micrometer-diameter wires for the time-delay jumpers, but high-temperature superconductors for this purpose are contemplated. In addition there may be some energy loss or resistance occasioned by electromagnetic radiation (EM) generated during the GW generation process. Such a loss can be reduced by the design of the energizing, for example coil, elements and controlling the direction of current pulses by the IIPCS. By the way, EM radiation can be easily screened out by means of interposing a conductor in the GW path, which is opaque to EM radiation but transparent to GW.

The spin-up/down of the entire rim is not instantaneous and is anticipated to progress at the speed of light in the rim from the juxtaposed permanent-magnet sites on the rim acted upon by the coil-magnetic fields. (Spin up/down does not progress at the local speed of sound, but rather it is expected to progress at the speed of light like a signal being transmitted by pushing a frictionless rod. That is, all of the ferromagnetic molecules comprising the magnets on the periphery of the rim move in concert, as the GW crest moves through the magnets at light speed, impelled by their magnetic fields. On the other hand, impulsive stresses in the spindle or dumbbell device are propagated inwardly at the speed of sound in the material of the device rather than at the speed of light.) Thus, for example, at a one picosecond cycle or switching rate an in-rim light speed of 3×10^8 [m/s] some $(10^{-12})(3\times10^8)=0.0003$ [m] or 0.3 [mm] or 300 micrometers of the rim on each side of the juxtaposed coil/magnet activity sites will respond (spin up or down) during each picosecond after coil activation. This process, in the case of a dumbbell-shaped rim, will generate an ever widening fan of gravitational waves, 51 as exhibited in FIG. **8**B of the parent patent. Many more than the 4,290 coil sites of the exemplar device of the parent patent could be distributed around the rim (perhaps over only a single sector or selected sectors adjacent to the rim; thereby greatly reducing the required number of coils and "focusing" the GW). A large number of ultra-fast switches, preferably semiconductor based, would be activated simultaneously by the IIPCS coil-control computer, with communication lines of nearly equal length to all switches.

(3) Rim Material Accelerations

A random series of positive and negative jerks tend to build up a positive acceleration over time by random walk. Consider an extreme case, however, in which the maximum jerk builds up rim acceleration (spin up, or deceleration, spin down) continuously over a length of time, $\delta t=10^{-7}$ [s] or 100 nanoseconds. The mass per unit length of the magnetic mass of the rim for the exemplar device of the parent application is $\Delta mass/\Delta l=3.83$ [kg/m]. The jerk is

$$da/dt=d^3S/dt^3=([\Delta f/\Delta l]/[\Delta mass/\Delta l]/\Delta t= (3000[N/m]/3.8[kg/m]/10^{-12}=7.89\times10^{14}[m/s^3], \tag{12}$$

where S is the displacement. Therefore, in δt=100 nanoseconds (10^{-7} [s]) of continuous and constant jerk, da/dt, the acceleration, a, would build up to

$$a(t)=\int_o^{\delta t}(da/dt)dt=(7.89\times10^{14})(10^{-7})=7.89\times10^7[m/s^2], \tag{13}$$

(of course, control of the jerks by the IIPCS would never allow such a high build up of acceleration) the speed would build up to

$$dS/dt=\int_o^{\delta t}(da/dt)dt=(da/dt)\delta t^2/2=(7.89/2)\times 10^{14}\times(10^{-7})^2=3.9[m/s], \tag{14}$$

and the displacement would build up to

$$S=\int_o^{\delta t}(ds/dt)dt=(da/dt)\delta t^3/6=(7.86/6)\times10^{14}\times (10^{-7})^3=1.315\times10^{-7}[m]. \tag{15}$$

The angular rate build up over 100 nanoseconds is 3.9[m/s]/1000[m]=3.9×10^{-3}[radians/s] versus, for example, the mean motion of double star PSR1913+16 of 2.25×10^{-4} [radians/s]. Thus there is considerable "motion" in the magnetic mass, but essentially the mass goes only a very short distance. In this regard, as already noted, the IIPCS can be programmed to ensure that there is not a secular increase or accumulation of magnetic-mass displacement beyond a certain prescribed limiting value. For example, if the jerk were reversed every two seconds (reciprocating) then the acceleration would build up to less than 200 [g's]. The stress moves away from the magnetic mass at sound speed (for example, 5000 [m/s]) or about 500 micrometers in 100 nanoseconds. They represent microscopic shock waves that will dissipate.

Radial-Jerk GW Generation Embodiment

In the case of the radially directed pulses or displacements of the rim or rim sector or sectors, they result in a time-

variable value of the moment of inertia, I. These displacements are built up in the radial direction by the sequential activation of radial arcs of coils in a given, single wedge-shaped sector or juxtaposed sectors of the rim or dumbbell. Under the control of the IIPCS, radially oriented strips of the aforementioned circuit-board or computer-chip rolls are sequentially activated to build up or generate a $(d^3I/dt^3)^2$ formulation or GW component as the GW moves outward at light speed. The radial displacements should be asymmetrical (as controlled by the IIPCS) in order to produce a quadrupole moment or so that the GW will not cancel out. The astrophysical analogy here is a vibrating white dwarf star emitting GW (see, for example, D. H. Douglas, p.491, of L. L. Smarr, opcit).

As is well known and noted in specifics by Geoff Burdge, Deputy Director for Technology and Systems of the National Security Agency (written communication dated Jan. 19, 2000) "Because of symmetry, the quadrupole moment can be related to a principal moment of inertia, I, of a three-dimensional tensor of the system and . . . can be approximated by

$$-dE/dt \approx G/5c^5(d^3I/dt^3)^2=5.5\times10^{-54}(d^3I/dt^3)^2." \quad (16)$$

In which k in Burdge's notation is G and the units are in the MKS system [watts] not the cgs. In this case

$$P=-G\kappa_{I3dot}(d^3I/dt^3)^2/5c^5[\text{watts}] \quad (17)$$

where $I=\delta m\ r^2$ [kg-m^2],

δm=mass of an individual rim sector or a number of sectors (or dumbbell) [kg], and

r=half of the distance between opposing δm [m]. Thus

$$d^3I/dt^3=\delta m\ d^3r^2/dt^3=2r\delta m\ d^3r/dt^3+\ldots \quad (18)$$

and d^3r/dt^3 is computed by noting that

$$2r\delta m\ d^2r/dt^2=2rf_r[\text{N-m}] \quad (19)$$

where f_r=radial force on a single rim sector, rim sectors, or dumbbell. (This single-sector embodiment of the invention can also be visualized as a linear motor.)

The derivative is approximated by

$$d^3I/dt^3=2r\Delta f_r/\Delta t, \quad (20)$$

in which Δf_r is the nearly instantaneous increase in the radial force on the rim caused by the magnetic field when it is turned on and off or pulsed by the transistors or ultra-fast switches of the IIPCS, that is, a radial jerk. In this regard the coils are sequenced radially outward by the IIPCS (at the speed of light) in order to generate or build up the high-frequency gravitational waves. Thus

$$P=-5.5\times10^{-54}\kappa_{I3dot}(2r\Delta f_r/\Delta t)^2[\text{watts}]. \quad (21)$$

Again, κ_{I3dot} will be a function determined experimentally to account for the fact that r may not be less than the GW wavelength for most high-frequency GW of interest.

As a numerical example, for a spindle similar to the one mentioned in the prior numerical example, but with a one-meter wide apron of peripheral magnets and IIPCS coil sets both top and bottom (thus 2×100 cm/m=200 times more force per meter along the rim's periphery), $\kappa_{I3dot}=32$, $\Delta f=3.6\times10^9$ [N], r=1000 [m] and $\Delta t=10^{-12}$ [s], so that (pending experimental verification)

$$P=-1.76\times10^{-52}(2\times1000\times3.6\times10^9/10^{-12})^2=-9.12\times10^{-3}[\text{watts}] \quad (22)$$

Again the reference area is 63 [m^2], so that the GW energy flux near the device is about 1.45×10^{-4} [watts/m^2].

Single-Sector or Linear-Motor, Tangential-Jerk GW Generation Embodiment

The single-sector or linear-motor embodiment of the invention (sometimes referred to as a linear induction motor or LIM) is the most preferred embodiment of the invention. It can be visualized to involve a single sector of the rim with the impulsive forces being tangential rather than radial. Alternatively, it can be conceptualized as the rim magnets and adjacent coils being peeled off from the rim and laid out flat. In this case an exemplar device would be 2000 [m] in length and 3 [m] in diameter. The approximately one-centimeter-wide chip rolls would be placed longitudinally along the sides of central, cylindrical, permanent- (or electro-) magnetic core, piston, or barrel as shown in FIG. **2**. Each meter-long segment of the roll would produce about 3000 [N] of longitudinal force, f_l, and all together they form a sheath of sub-millimeter coils surrounding the central magnetic core, piston, or barrel consisting of magnetic sites. Note that in this case the motion of the magnetic mass is asymmetrical (either "in" or "out") so that there is a quadrupole and the GW do not cancel and become null. The IIPCS controlled current can proceed in either direction and in the single-interconnecting-line-of-coils embodiment of the invention alternative coil lines can be energized by pulse trains moving in opposite directions to vibrate the magnetic mass. The magnetic mass itself can be composed of electromagnets, with or without cores, that can be controlled by the IIPCS to augment the GW generation.

(1) Numerical Example

As a numerical example, there would be about one roll or 25-layer strip of chips spaced around and adjacent to the barrel in a longitudinal direction (parallel to the barrel axis) every 2 centimeters forming the sheath. Thus there would be π×3[m]×100[cm/m]/2[cm]=471 strips, 2000 [m] long or

$$\Delta f_l=(471)\ (2000[\text{m}])\ (3000[\text{N/m}])=2.83\times10^9[\text{N}] \quad (23)$$

and with $\kappa_{mr3dot}=32$ (to be established experimentally),

$$P=-1.76\times10^{-52}(2\times2000\times2.83\times10^9/10^{-12})^2=-2.26\times10^{-2}[\text{watts}]. \quad (24)$$

Thus, with the reference area of the two 3 [m] diameter ends, $2\pi(1.5)^2=14$ [m^2], (GW propagating in both directions) the generated GW flux is about 1.6×10^{-3} [watts/m^2].

(2) Sector-Material Accelerations

In this case the jerk is obtained from

$$(da/dt)_{per\ unit\ area}=(\Delta f/\Delta t)/(\Delta\text{mass}/\Delta A) \quad (25)$$

where $\Delta f=\Delta f_i$ [N]/(2000[m]×3[m]π)=$2.83\times10^9/1.885\times10^4=1.5\times10^5$ [N/m^2], so that

$$\Delta f/\Delta t=1.5\times10^5/10^{-12}=1.5\times10^{17}[\text{N/m}^2\text{-s}] \text{ and} \quad (26)$$

Δmass/ΔA=mass per area (3.8 [kg/m] of strip) (471 strips per meter)=1.79×10^3 [kg/m^2], so that $da/dt=1.5\times10^{17}/1.79\times10^3=8.38\times10^{13}$ [m/s^3]. Therefore, in 100 nanoseconds of continuous jerk the acceleration would build up to

$$a=d^2S/dt^2=(da/dt)\delta t=(8.38\times10^{13})\ (10^{-7})=8.38\times10^6[\text{m/s}^2]. \quad (27)$$

As already noted, the IIPCS would be programmed so that accelerations would never approach this value! As an example, for a one THz alternating jerk the acceleration would only build up to $(8.38\times10^{13})\ (10^{-12})=83.8$ [m/s^2]=8.6 [g's] (alternating or reciprocating "hammer blows" acting on a single mass or masses (such as magnetic sites); not harmonic oscillation of two masses). In the extreme case of 100 nanoseconds of continuous jerk in the same direction, the speed would build up to

$$dS/dt=(da/dt)\delta t^2/2=(8.38\times10^{13})\times10^{-14}=0.42[m/s] \quad (28)$$

and the displacement of the magnetic mass (magnetic surface of single sector, piston, or barrel) is

$$S=(da/dt)\delta t^3/6=(8.38\times10^{13}/6)\times\mathbf{10}^{-12}=1.40\times10^{-8}[m]. \quad (29)$$

Again there is considerable "motion" of the magnetic mass, but it goes a very small distance before the IIPCS reverses the built-up acceleration, speed, and displacement.

Infinite-radius Coil GW Generation Embodiment

For comparison with the foregoing embodiments of the invention, consider the evolution of coil pairs or coil sets into flattened-out pairs of parallel wires (that is, infinite-radius coils) situated very close to each other and carrying a large current in the same direction (and, therefore, attracting each other). This current, which can go either way, is to be pulsed by a large number of ultra-fast switches or transistors **58** (FIG. **3**) about every picosecond by the IIPCS to produce pulses of electrical current through the wires. For simplicity, as an example consider the wires to be flat, one-meter square plates (therefore a one-square-meter GW reference area or smaller (e.g., in order to achieve $r<<\lambda_{GW}$) down to a current-pulse wavelength across or made larger by constructing a mosaic of individual plate pairs) as exhibited schematically in FIG. **3**. As a numerical example, let each plate carry a one-thousand ampere current and the plates in the pairs are situated one-micrometer (10^{-6} [m]) apart and the pairs are separated at a greater distance, say, 0.1 [mm]. In order to achieve asymmetrical mass displacement (to produce a quadrupole moment) one plate of each pair or of each "coil set" could be considerably more massive than the other, that is, exhibit a considerably larger cross section or be joined to a ballast **67** in FIG. **3** or carry a much larger current than the other or have different modulus of elasticity or be constrained differently in their mountings. If the IIPCS pulsed these conducting plates with picosecond-duration pulses, then during each cycle the attractive, impulsive force (lateral jerk) for each coil pair or set of coil pairs would be

$$\Delta f=(\mu_o/2\pi)\ (1000[amps]\times[1000[amps])/10^{-6}[m]=2\times10^5[N]. \quad (30)$$

The IIPCS-controlled switches **58** in FIG. **3**, will sequence the current pulses **59** moving approximately at light speed, c, to build up a gravitational wave **29**.

Let GW-radiated power be given by

$$P=G(md^3r^2/dt^3)^2/5(c/2)^5\ G(\Delta f/\Delta t)^2/5(c/2)^5=1.76\times10^{-52}(\Delta f/\Delta t)^2\ [watts], \quad (31)$$

where $\Delta f/\Delta t=2\times10^5/10^{-12}=2\times10^{17}$ [N/s]. Thus, pending experimental verification:

$$P\ 1.76\times10^{-52}(2\times10^{17})^2=7.04\times10^{-18}[watts] \quad (32)$$

and since the reference area is two square meters (GW propagates into and out of the plates) the GW flux=3.5×10^{-18} [watts/m^2]. The product of the amperage of, say, two plates would need to go up by a factor of about 10^5 (to about one-million amperes) or the distance between the plates reduced by the same factor (to ten picometers or 10^{-11} [m]), or the number of plate pairs increased by a factor of a thousand, or a mosaic of many plate pairs per level (and multiple levels), or some combination thereof in order to approach the GW-flux values for the other embodiments of the invention. Such a current is, however, exceeded by the eighteen-million-ampere current passed through the Sandia Laboratory Z-pinch machine (A. Wilson, "Z Mimics X-rays from Neutron Stars", *Science*, Volume 286, Dec. 10, 1999, p. 2059). The current-produced jerk of this machine would be expected to generate a GW pulse as its tungsten wires collapse on each other, if they do so in an asymmetrical manner.

Electromechanical-force GW Generation Embodiment

At one THz the GW wavelength is 3×10^{-4} [m] or 300 micrometers so that the half wavelength and the optimum crystal dimension according to Joseph Weber, p. 313, 1960 opcit, is 150 micrometers and, of course, even smaller for the approximate quadrupole equation to hold. If the ensemble of electromechanical elements, for example, piezoelectric crystals, were controlled by the IIPCS and replaced the coils, and were on 160 micrometer centers in the chips, then there would be about $60\times60=3.6\times10^3$ per square centimeter. If there were 25 chip levels or layers, then there would be about $25\times3.6\times10^3=9\times10^4$ crystals per square centimeter or 9×10^9 per [m^2] as shown schematically in FIG. **4**. The energy would be 10^{-20} [watts] per crystal (if each driven just below its breaking point as enhanced by low-temperature and high-frequency operation) multiplied by 9×10^9 (crystals)$\cong10^{-11}$ [watts] (subject to experimental verification) . With the crystals properly oriented and programmed by the IIPCS to propagate GW radiation out of the side of the centimeter-thick, one-meter-square crystal array (whose area is about one centimeter by one meter, that is a reference area of 2×10^{-2} [m^2]) the GW flux would be $1/2\times10^{-11}/10^{-2}=5\times10^{-1C}$ [watts/m^2]. As noted by Joseph Weber (ibid), such a system could be employed " . . . to generate and detect gravitational radiation." (Emphasis added.) With regard to detection, the crystals would represent very small resonators whose natural frequencies were in the terahertz range. Alternatives to the preferred embodiment of the invention using piezoelectric crystals (or piezoelectric polycrystalline ceramics), include, but are not limited to either P or N processed strain-gage silicon semiconductors, thin-film piezoelectric resonators, electromechanical nanomachines, capacitors, dielectric resonators, solenoids and piezoelectric polymers. Electromechanical nanomachines are such devices as piston actuators, motors, vibrators and pumps. For specific design details either G. L. Wojcik, et al, "Electromechanical Modeling Using Explicit Time-Domain Finite Elements", *IEEE* 1993 *Ultrasonics Symposium Proceedings*, Volume 2, pp. 1107–1112 or Jan Kocback's "Finite-Element-Modeling Analysis of Piezoelectric Disks,—Method and Testing", Master of Science Thesis, Department of Physics, University of Bergen, Bergen, Norway, can be utilized.

As discussed in detail in the 1960 and 1964 Joseph Weber articles referred to above, the passage of a gravitational wave deforms an object or set of objects as it passes through them. For example, a piezoelectric polymer, a silicon semiconductor, a thin-film piezoelectric resonator, a piezoelectric-crystal functioning as a collector element is deformed by a GW and produces a small electrical current. Likewise, the plates of a capacitor functioning as a collector element are slightly moved relative to each other and thereby produces a signal. In fact, these elements are both energizable and generate GW and also are collectors and detect GW through the same conductors. The nanomachine collectors operate in a similar fashion. A nanomachine is a microscopic or molecular sized machine, for example, a microscopic version of the dumbbell motor/generator of the parent patent. As a GW passes through the collector, the dumbbell moves slightly and submicroscopic coils respond to this motion and generate a small current. Likewise, energizing the microscopic coils in the motor mode will generate GW due to dumbbell motion. Electrical transducer or micro strain gauge nanomachines respond to the defor-

mation occasioned by the passage of a GW in exactly the same fashion as it does to a mechanically induced strain and thereby function as a GW collector. The nanomachine pressure transducer collector element responds to a slight change in pressure of a set of particles comprising a fluid as the GW passes through it. The location of the collector elements and their connection with ultra-fast switches or transistors is identical to the location of the energizer elements, shown in FIG. **4** and, as already noted may be one and the same element acting as either an energizer or a collector.

Communication Utilizing GW

As an approximate numerical example related to a possible gravitational-wave detector for a train of high-frequency, THz, gravitational waves, consider the absorption cross section, a [m^2], for such antennas as given by Joseph Weber (opcit, 1964, p.99)

$$\sigma=15\pi GIQ\beta^2N^2/8\omega c[m^2] \quad (33)$$

where $G=6.67423\times10^{-11}$ [m^3/kg-s^2] (universal gravitational constant),

I=moment of inertia or quadrupole moment of the detector element(s) [kg-m^2],

Q=π times the number of oscillations a free oscillator undergoes before its amplitude decays by a factor of e,

$\beta=2\pi/\lambda$ [1/m] (propagation constant),

$\lambda=c/\nu$ [m] (gravitational-wave wavelength),

N=the number of quadrupoles coupled together in the antenna (see Eq. (2.9A), p.62, of *Gravitational Radiation and Relativity*, Edited by J. Weber and T. M. Karade, World Scientific Publishing Co., Singapore, 1986),

$c=3\times10^8$ [m/s] (the speed of light),

ν=frequency of gravitational radiation [1/s or Hz], and

ω=angular frequency (or mean motion) [1/s].

For $Q=10^6$ (as noted by Joseph Weber opcit, 1960, p. 308, "A practical antenna might be expected to have $Q\approx10^6$." A large Q implies that a long time is required for the quadrupole element to reach thermal equilibrium. Also, the detection devices that Weber had in mind were large isolated aluminum cylinders, suspended and well isolated from the environment. The collector elements for the present device will probably be contained on a chip with damping constraints and a much smaller Q is likely).

$\nu=10^{12}$ [Hz] or one [THz], and

$\beta=2\pi\nu/c=2.09\times10^4$ [1/m], and

$\omega=\nu/2=5\times10^{11}$ [1/s]; see Weber, 1964, opcit p. 90, we have

$\sigma=1.15\times10^{-15}$ IN^2[m^2].

This value, depending upon I and the number of quadrupoles (with masses and characteristics nearly identical) coupled together in the antenna, N, compares favorably with $\sigma=10^{-20}$ [m^2] of the Weber Bar given on p. 102 of Weber , ibid.

An approximate estimate of what bandwidth a gravitational-wave (GW) communication system might achieve is obtained as follows: Suppose that the distance between the GW generating or transmitting device and the receiver or detector is about one Earth's radius, 7,000 [km] or 7000 rim radii. Also, suppose that we are transmitting through the Earth's mantle and that 10 percent of the GW energy gets through. Thus, for the tangential-jerk situation the "signal" obtained by modulating the current pulses by the IIPCS (some pulses missing and some forming a longer-duration pulse or pulses of different amplitudes) is using the power near the spindle device given by Eq. (9) and the average power flux of 2×10^{-9} [watts/m^2] there

$$S=(2\times10^{-9})\ (0.1)/7000=2.8\times10^{-14}[watts/m^2] \quad (34)$$

at the receiver or detector (assuming that the GW propagates in a single plane with little or no diffraction). For the radial-jerk spindle situation using the power near the device given by Eq. (22) and the average power flux of 1×10^{-4} [watts/m^2] there

$$S=(1\times10^{-4})\ (0.1)/7000=1.4\times10^{-9}[watts/m^2]. \quad (35)$$

For the longitudinal-jerk, linear-motor situation with average power from Eq. (24) (if there were a fall off with range in "rim" radii), with a 2000 [m] length or radius of gyration and a resulting average power of 1×10^{-3} [watts/m^2] we have,

$$S=(1\times10^{-3})\ (0.1)/7000=1.4\times10^{-8}[watts/m^2]. \quad (36)$$

Let us estimate the "noise" $N=10^{-20}$ [watts/m^2] in the THz band(probably not many GW sources there, but Brownian motion, thermal and quantum fluctuations, etc. may result in much more noise than this) and that the GW detector exhibits a sensitivity on this same order. It should be recognized, however, that the bandwidth of the long-baseline, interferometric GW detectors now under construction are at most about a few kHz and they are not designed for THz detection. Furthermore, a six to ten order-of-magnitude improvement of the sensitivity of the single-crystal detectors considered by Joseph Weber 42 years ago might need to be accomplished (sensitivity of about 10^{-10} [watts] as given on p. 313 of Weber opcit, 1960). More recently, however, Weber has speculated optimistically (opcit, 1986, p. 30) that there is " . . . no limit to the theoretical sensitivity of a (elastic solid) gravitational radiation antenna, and perhaps no limit to the number of novel methods for improving the sensitivity of existing antennas."

Using Shannon's classical equation (C. B. Shannon, *Bell Systems Technical Journal*, Volume 27, Number 379, p. 623, 1948), the maximum rate of information transfer, C, for the spindle's tangential-jerk GW embodiment is given by

$$C=B\log_2(1+2.8\times10^{-14}/10^{-20})\cong(10^{12})\ (20)=2\times10^{13}[bps], \quad (37)$$

for the radial-jerk GW embodiment

$$C=B\log_2(1+1.4\times10^{-9}/10^{-20})\cong(10^{12})\ (30)=3\times10^{13}[bps] \quad (38)$$

and for the longitudinal-jerk (linear-motor) GW preferred embodiment

$$C=B\log_2(1+1.4\times10^{-8}/10^{-20})\cong(10^{12})\ (40)=4\times10^{13}[bps]. \quad (39)$$

In each embodiment the bandwidth, B, is taken to be the IIPCS switch on-off or "chop" rate of about 10^{12} reciprocating "hammer blows" or jerks per second (one THz and multiple GW generators or "transmitters" could increase the bandwidth further).

There exists a useful figure-of-merit or trade-off function for the longitudinal-jerk (single-sector or linear-motor) preferred embodiment of the invention that relates to the received signal:

$$S\ ([2\pi rl\{l\Delta f_l/\Delta A\}/\Delta t]^2/[\pi r^2])\alpha \quad (40)$$

$$(l^2[\Delta f_l/\Delta A]/\Delta t)^2\alpha \quad (41)$$

(note that the radius of the single-sector cylindrical magnetic core, piston, or barrel, cancels out) where

S=signal at the detector (receiver) [watts/m^2],

r=radius of the single-sector magnetic core, piston, or barrel [m],

l=radius of the single-sector or length at magnetic core, piston, or barrel [m],

$\Delta f_l/\Delta A$=longitudinal force per unit area acting on the single-sector magnetic core, piston, or barrel [N/m^2],

Δt=impulse time [s],

α=attenuation due to intervening material between the GW generator (transmitter) and detector (receiver) [dimensionless],and assume that for the single-sector or linear-motor preferred embodiment of the invention, if there is little or no GW diffraction, then there is no range, ρ dependence (to be tested experimentally).

As a numerical example, consider the solution for the length, l,

$$l=\sqrt[4]{(S/\alpha)\times(\Delta t/[\Delta f_l/\Delta A])^2} \quad (42)$$

where

$S=4.5\times10^{-4}$ multiplied by the nominal=(4.5×10^{-4}) $(1.4\times10^{-8})=6.26\times10^{-12}$ [watts/m^2] (assume a $\{1/4.5\}\times10^4$ more sensitive detector or receiver),

α=10 multiplied by the nominal=(10) (0.1)=1 [dimensionless] (assume no attenuation),

$\Delta t=10^{-1}$ multiplied by the nominal=10^{-13} [s] (occasioned by the possible design of a 100 femtosecond ultra-fast switch and pulse duration), and

$\Delta f_l/\Delta A$=100 multiplied by the nominal=(100) $(5.64\times10^9/[2000\times3\pi])=(100)(3.0\times10^5)=3.0\times10^7$ [N/m^2] (assume increased magnet efficiencies due to, for example, use of high-temperature super conductors and electromagnets).

Thus, in this case the length of the GW generator would be (with the factor of 7000 rim radii removed; thus the factor of $10^{-3}/7$)

$$l=\sqrt[4]{([6.26\times10^{-12}/10][10^{-3}/7]/[100/0.1]^2)\ (2000)}=6.15\times10^{-3}[m]=6.15[mm]. \quad (43)$$

Equation (41) can be utilized by a person with average skill in the art to practice the inventions utilizing fast or ultra-fast switches or transistors having different capabilities, that is different Δt, different detection capabilities, S, different forces, Δf, and different lengths, l including lengths significantly smaller than GW wavelength.

Propulsion

No doubt high-frequency GW experiments will reveal many applications of GW to propel spacecraft by means of remote GW generators. In this regard, on p. 349 of Landau and Lifshitz (opcit), they comment: "Since it has a definite energy, the GW is itself the source of some additional gravitational field. Like the energy producing it, this field is a second-order effect in the h_{ik} (tensor describing a weak perturbation of the galilean metric). But in the case of high-frequency gravitational waves the effect is significantly strengthened . . . " (Emphasis added.)

The axis of rotation of a spindle GW-generation device defines a preferred, single, unique direction in space and also a preferred, single, unique plane. The axis of the single-sector, linear-motor GW generator device defines a preferred, unique direction in space as well. Thus there is an asphericity or pattern to the gravitational radiation, an anisotropy or focusing, that is analogous to a radio-antenna pattern of field strength. The concept that, as a part of this pattern, the gravitational waves are constrained to the "preferred" plane of the rim, or axis of the linear-motor's "preferred" line in space, without diffraction, will also be tested. These concepts have potential application to spacecraft propulsion either by remote "gravitational force field" generation or by placing anisotropic GW generators on board a spacecraft—a "Relativistic Rocket".

What is claimed is:

1. A gravitational wave generator utilizing a computer controlled logic system to produce gravitational waves by imparting a third time derivative to the motion of a mass or a system of masses.

2. A gravitational wave generation device comprising:

a plurality of energizable elements;

a plurality of electronic switches connected to the elements;

a computer-controlled logic system operatively connected to a power source for selectively energizing the elements; and

a transmitter operatively connected to the elements for selectively connecting the elements to an information processing device in a predetermined time sequence to generate gravitational waves.

3. A device according to claim **2** wherein the electronic switches are semiconductor based.

4. A device according to claim **2** wherein the plurality of elements are piezoelectric crystals.

5. A device according to claim **2** wherein the plurality of elements are silicon semiconductors.

6. A device according to claim **2** wherein the plurality of elements are thin-film piezoelectric resonators.

7. A device according to claim **2** wherein the plurality of energizable elements are piezoelectric polymers.

8. A device according to claim **2** wherein the plurality of elements are of a submillimeter size and incorporated into a semiconductor chip.

9. A device according to claim **2** wherein the plurality of energizable elements are of a submillimeter size and integrated with polymer-based devices.

10. A device according to claim **2** wherein a subset of the energizable elements are arranged in a line of such elements such that when energized by a train of current pulses, the duration of which allows the pulses to traverse a subsequent energizable element completely, and reaches the next energizable element in the line with time delays between the elements to ensure that the pulses can reach each of the elements in the line at an appropriate time to generate gravitational waves as the train of pulses progresses down the line.

11. A device according to claim **2** wherein the plurality of elements are of a submillimeter size and integrated with polymer-based devices.

12. A device according to claim **2** wherein the plurality of elements are nanomachines.

13. A device according to claim **12** wherein the nanomachines are piston actuators.

14. A device according to claim **12** wherein the nanomachines are motors.

15. A device according to claim **12** wherein the nanomachines are vibrators.

16. A device according to claim **12** wherein the nanomachines are pumps.

17. A device according to claim **2** wherein a mass or masses are is set in motion by the elements and exhibits a third-time-derivative motion of the mass or masses to produce a gravitational wave.

18. A device according to claim **17** wherein the mass or masses are a plurality of electromagnets.

19. A device according to claim **18** wherein a cylindrically shaped magnetic core mass is provided, surrounded by a

sheath composed of a plurality of very small coils or coil sets, each individually controlled by computer means.

20. A device according to claim **17** wherein the mass or masses are a plurality of permanent magnets located around the rim of a spindle.

21. A device according to claim **20** wherein to a plurality of conductive wire coils or coil sets adjacent to the spindle rim are segmented into one or more separate, juxtaposed sectors.

22. A device according to claim **20** wherein the plurality of magnets located around the periphery of the spindle rim are separated into one or more separate, juxtaposed magnets.

23. A device according to claim **2** wherein the plurality of elements are electrically energizable elements.

24. A device according to claim **23** wherein the plurality of electrically energizable elements are capacitors.

25. A device according to claim **23** wherein the plurality of electrically energizable elements are dielectric resonators.

26. A device according to claim **2** wherein the plurality of elements are small conductive wire coils or coil sets.

27. A device according to claim **23** wherein the plurality of conductive wire coils or coil sets are microscopic in size and integrated with polymer-based devices.

28. A device according to claim **26** wherein the plurality of small conductive coils or coil sets are individually sequenced radially outward by current pulses to generate gravitational waves resulting from a third-time-derivative motion or jerk of a time-variable moment of inertia or mass distribution of a magnetic mass.

29. A device according to claim **26** wherein the plurality of coils or coil sets along a predetermined line of such coils or coil sets are energized by a train of current pulses, the duration of which is controlled by computer means to cause the pulse to traverse a subsequent coil or coil set completely and reaches the coil set in the line via the same single conductor wire with predetermined time delays between the coils or coil sets to ensure that the pulses reach each of the coils or coil sets in the line at an appropriate time to generate gravitational waves as the train of pulses progresses down the line.

30. A device according to claim **23** wherein the plurality of conductive wire coils or coil sets are of a very small size and encased in or imprinted on a semiconductor chip.

31. A device according to claim **30** wherein the semiconductor chip, containing the plurality of very small coils or coil sets, is layered with circuit elements sequenced to launch a magnetic-field pulse of very brief duration which interacts almost simultaneously with the electromagnetic field of a magnetic mass to cause a sufficient third-time-derivative motion or jerk of the magnetic mass to generate gravitational waves.

32. A device according to claim **31** wherein the magnetic mass is a single magnet acted upon by the plurality of coils or coil sets adjacent to it.

33. A device according to claim **2** wherein the plurality of elements are electromagnetic-force elements.

34. A device according to claim **33** wherein the plurality of electromagnetic-force elements are solenoids.

35. A device according to claim **33** herein the electromagnetic force elements are coil sets that are flattened out into one or more parallel conductor pairs situated close to one another and are pulsed with a current by computer controlled switches to produce a gravitational wave.

36. A device according to claim **35** wherein the plurality of parallel conductor pairs are arranged in a mosaic pattern or in multiple layers of a mosaic pattern.

37. A device according to claim **35** wherein the plurality of parallel conductor pairs are of microscopic size and encased in or imprinted on a semiconductor chip or integrated with a polymer-based device.

38. A device according to claim **35** wherein the plurality of parallel conductor pairs along a predetermined line of such parallel conductor pairs are energized by a train of current pulses the duration of which is controlled by computer means to cause the pulse to traverse a subsequent pair of parallel conductors completely and reaches the parallel conductors in the line via the same single conductor wire with predetermined time delays between the pairs of parallel conductors to ensure that the pulses reach each of the pairs of parallel conductors in the line at an appropriate time to generate gravitational waves as the train of pulses progresses down the line.

39. A gravitational wave detection device comprising:

a plurality of collector elements;

a plurality of electronic switches connected to the elements;

a computer-controlled logic system operatively connected to an information processing device for selectively connecting the elements in a predetermined time sequence; and

a receiver operatively connected to the information processing device to indicate the detection of gravitational waves.

40. A device according to claim **39** wherein the plurality of elements are piezoelectric crystals.

41. A device according to claim **39** wherein the plurality of elements are silicon semiconductors.

42. A device according to claim **39** wherein the plurality of elements are thin-film piezoelectric resonators.

43. A device according to claim **39** wherein the plurality of elements are piezoelectric polymers.

44. A device according to claim **39** wherein the plurality of elements are capacitors.

45. A device according to claim **39** wherein the plurality of elements are of a submillimeter size and incorporated into a semiconductor chip.

46. A device according to claim **39** wherein the plurality of elements are nanomachines.

47. A device according to claim **46** wherein the nanomachines are piston actuators.

48. A device according to claim **46** wherein the nanomachines are generators.

49. A device according to claim **46** wherein the nanomachines are electrical transducers.

50. A device according to claim **46** wherein the nanomachines are pressure transducers.

* * * * *

Ninth International Conference on
Geometry, Integrability and Quantization
June 8–13, 2007, Varna, Bulgaria
Ivaïlo M. Mladenov, Editor
SOFTEX, Sofia 2008, pp 320–333

Geometry, Integrability and Quantization IX

ON THE POLARIZATION OF GRAVITATIONAL WAVES

GAETANO VILASI

Dipartimento di Fisica E. R. Caianiello, Università di Salerno
Istituto Nazionale di Fisica Nucleare, Sezione di Napoli
GC Salerno, Via S. Allende, I-84081 Baronissi (Salerno), Italy

Abstract. Physical properties of gravitational waves, belonging to the larger class of exact solutions of Einstein field equations which are invariant for a non-Abelian two-dimensional Lie algebra of Killing fields, are described. It is shown that in the would be quantum theory of gravity they correspond to spin $-$ 1 massless particles. The gravitational interaction of two pencils of light is analyzed.

Introduction

The aim of this talk is to illustrate some interesting and, in a sense, surprising physical properties of special solutions of Einstein field equations, belonging to the large class of Einstein metrics invariant for a non-Abelian two-dimensional Lie algebra of symmetries, which throw new light on the following problem.

A long time ago Tolman, Ehrenfest and Podolsky [30] and later Wheeler [33] analyzed the gravitational field of light beams and the corresponding geodesics in the linear approximation of Einstein equations. They discovered that null rays behave differently according whether they propagate parallel or antiparallel to a steady, long, straight beam of light, but they did not provide a physical explanation of this fact. The result was clarified in part by Faraoni and Dumse [14] using an approach based on a generalization to null rays of the gravitoelectromagnetic Lorentz force of linearized gravity. They also extended the analysis to the realm of exact *pp*-wave solutions of the Einstein equations.

Since the problem of the gravitational interaction of two photons is still unsolved, it appears necessary to take into full account the nonlinearity of Einstein's equations when studying the generation of gravitational waves from strong sources [12, 29].

On the other hand, some decades ago, Belinski and Zakharov have shown [4] that there exist metrics such that the corresponding vacuum Einstein field equations reduce to a system of non-linear partial differential equations, whose *generalized Lax form* is characteristic for integrable systems. Then, by using a suitable generalization of the *Inverse Scattering Transform*, they were able to find *solitary waves solutions*.

A geometric inspection of mentioned metrics shows that they are invariant under translations along the x, y-axes, i.e., they admit two Killing fields, ∂_x and ∂_y, closing on an Abelian two-dimensional Lie algebra $\mathcal{A}_2$. Moreover, the distribution $\mathcal{D}$ generated by ∂_x and ∂_y is two-dimensional and the distribution $\mathcal{D}^\perp$ orthogonal to $\mathcal{D}$ is integrable and transversal to $\mathcal{D}$.

Since a two-dimensional Lie algebra is either Abelian ($\mathcal{A}_2$) or non-Abelian ($\mathcal{G}_2$), it has been natural to consider [24–26] the problem of characterizing all gravitational fields g admitting a Lie algebra $\mathcal{G}$ of Killing fields such that

I the distribution $\mathcal{D}$, generated by vector fields of $\mathcal{G}$, is two-dimensional.

II the distribution $\mathcal{D}^\perp$, orthogonal to $\mathcal{D}$ is integrable and transversal to $\mathcal{D}$.

The condition of transversality can be relaxed [9, 10], so that in order to distinguish the different cases, the notation $(\mathcal{G}, r)$ is used. Metrics satisfying the conditions *I* and *II* are called *of* $(\mathcal{G}, 2)$*-type*. Metrics satisfying conditions *I* and *II*, except the transversality condition, are called *of* $(\mathcal{G}, 0)$*-type* or *of* $(\mathcal{G}, 1)$*-type* according to the rank of their restriction the leaves of $\mathcal{D}$ which are also called *Killing leaves*.

All the possible situations corresponding to a two-dimensional Lie algebra of isometries, are described by the following Table 1 in which the cases indicated with bold letters are essentially solved [2, 9, 10, 24–27] where a non integrable two-

Table 1

	$\mathcal{D}^\perp, r = 0$	$\mathcal{D}^\perp, r = 1$	$\mathcal{D}^\perp, r = 2$
$\mathcal{G}_2$	**integrable**	**integrable**	**integrable**
$\mathcal{G}_2$	**semi-integrable**	**semi-integrable**	**semi-integrable**
$\mathcal{G}_2$	**non-integrable**	**non-integrable**	non-integrable
$\mathcal{A}_2$	**integrable**	**integrable**	integrable
$\mathcal{A}_2$	semi-integrable	semi-integrable	semi-integrable
$\mathcal{A}_2$	non-integrable	non-integrable	non-integrable

dimensional distribution has been called *semi-integrable* if it is part (i.e., a suitable restriction) of a three-dimensional integrable distribution.

The study of $\mathcal{A}_2$-invariant Einstein metrics goes back to Einstein and Rosen [13], Kompaneyets [16], Geroch [15], Belinsky, Khalatnikov, Zakharov [3, 4], so that

some exact solutions already known in the literature [28] have been rediscovered. Nevertheless, the geometric approach allows to perform in a natural way the choice of coordinates, i.e., the coordinates adapted to the symmetries of the metrics, even if they do not admit integrable $\mathcal{D}^{\perp}$ distribution. Usually, the standard techniques to find exact solutions assume, from the very beginning, that there exist natural vector fields, surfaces forming, which simplify the choice of the coordinates system. These assumptions are strong topological constraints on the space-time. The method developed in [2, 9, 10, 24–27] can be applied also when such topological assumptions do not hold.

The paper is organized as follows. In Section 1 gravitational fields invariant for a non Abelian two-dimensional Lie algebra, when the commutator of generators of the Lie algebra is of *light-type*, are characterized from a geometric point of view. In Section 2, the canonical and the Landau energy-momentum pseudo-tensors are introduced and a comparison with the linearized theory is performed. The role of (realistic) sources for such gravitational waves is also described. Eventually, the analysis of the polarization leads to the conclusion that these fields are spin -1 gravitational waves. In Section 3, spin -1 and spin -2 gravitational waves are compared from a gravitoelectromagnetic perspective. Section 4 is devoted to the analysis of the photon-photon gravitational interaction.

1. Geometric Aspects

Let g be a metric on the space-time $\mathcal{M}$ and $\mathcal{G}_2$ one of its Killing algebras whose generators X, Y satisfy $[X, Y] = sY$, $s = 0, 1$. The Frobenius distribution $\mathcal{D}$ generated by $\mathcal{G}_2$ is two-dimensional and in the neighborhood of a non-singular point *adapted coordinates* (x, y, p, q) exist ([9, 10, 24–27]) such that

$$X = \frac{\partial}{\partial p}, \qquad Y = \exp(sp)\frac{\partial}{\partial q}\cdot$$

In these coordinates, the general solution of vacuum Einstein equations, in the case in which the Killing field Y is of *light type*, is given by

$$g = 2f(\mathrm{d}x^2 \pm \mathrm{d}y^2) + \mu[(w(x, y) - 2sq)\mathrm{d}p^2 + 2\mathrm{d}p\mathrm{d}q] \tag{1}$$

where $\mu = A\Phi + B$ with $A, B \in \mathbb{R}$, $\Phi(x, y)$ is a non-constant harmonic function, $f = (\nabla\Phi)^2\sqrt{|\mu|}/\mu$, $w(x, y)$ is the solution of the *Euler-Darboux equation*

$$\Delta_{\pm}w + (\partial_x \ln|\mu|)\partial_x w \pm (\partial_y \ln|\mu|)\partial_y w = 0$$

where $\Delta_{\pm}$ is the Laplace (d'Alembert) operator in the (x, y)-plane. Metrics (1) are Lorentzian if the orthogonal leaves are conformally Euclidean, i.e., the positive sign is chosen, and Kleinian otherwise. Only the Lorentzian case will be analyzed and these metrics will be called of $(\mathcal{G}_2, 2)$-*isotropic type*.

In the particular case $s = 1$, $f = 1/2$ and $\mu = 1$, the above (Lorentzian) metrics are locally diffeomorphic [7] to a subclass of the vacuum Peres solutions [20, 28], that for later purpose we rewrite in the form

$$g = \mathrm{d}x^2 + \mathrm{d}y^2 + 2\mathrm{d}u\mathrm{d}v + 2(\varphi_{,x}\mathrm{d}x + \varphi_{,y}\mathrm{d}y)\mathrm{d}u. \tag{2}$$

The correspondence between (1) and (2) depends on the special choice of the function $\varphi(x, y, u)$, which, in general, is harmonic in x and y. In our case

$$x \to x, \quad y \to y, \quad u \to u, \quad v \to v + \varphi(x, y, u), \quad h = \varphi_{,u}.$$

In the case $\mu = \mathrm{const}$, the Euler-Darboux equation reduces to the Laplace equation. For $\mu = 1$, in the harmonic coordinates system (x, y, z, t) defined in [6], the above Einstein metrics take the particularly simple form

$$g = 2f(\mathrm{d}x^2 + \mathrm{d}y^2) + \mathrm{d}z^2 - \mathrm{d}t^2 + \mathrm{d}(w)\mathrm{d}(\ln|z - t|). \tag{3}$$

This shows that, when w is constant, the Einstein metrics given by equation (3) are static and, under the further assumption $\Phi = x\sqrt{2}$, they reduce to the Minkowski one. Moreover, when w is not constant, gravitational fields (3) look like a *disturbance* propagating at light velocity along the z direction on the Killing leaves. In the following we will only consider the case $\Phi = x\sqrt{2}$.

More precisely, the wave character and the polarization of gravitational fields (2) can be checked by using the covariant Pirani's criterion. To use this criterion the Weyl scalars must be evaluated according to the Petrov-Penrose classification [19, 21].

To perform the Petrov-Penrose classification, one has to choose a *tetrad* basis with two real null vector fields and two real spacelike (or two complex null) vector fields. Then, according to the Pirani's criterion, if the metric belongs to type **N** of the Petrov classification, it is a gravitational wave propagating along one of the two real null vector fields. Since ∂_u and ∂_v are null real vector fields and ∂_x and ∂_y are spacelike real vector fields, the above set of coordinates is the right one to apply for the Pirani's criterion.

Since the only nonvanishing components of the Riemann tensor, corresponding to the metric (2), are

$$R_{iuju} = -\partial^2_{ij}\partial_u\varphi, \qquad i, j = x, y$$

this gravitational fields belong to Petrov type **N** [11, 34]. Then, according to the Pirani's criterion, the metric (2) does indeed represent a gravitational wave propagating along the null vector field ∂_u.

2. Physical Properties

In the following, physical properties of metrics (1) will be analyzed only in the case of Lorentzian signature. In previous section, the wavelike nature of gravitational

fields (1) has been checked [6] by using covariant criteria. Now, we will shortly review the most important properties of these waves which will turn out to have spin $-$ 1.

Let us remark that the definition and the meaning of spin or polarization for a theory, such as general relativity, which is non-linear and possesses a much bigger invariance than just the Poincaré one, deserve a careful analysis. It is well known that the concept of particle, together with its degrees of freedom like the spin, may be only introduced for linear theories. In these theories, when Poincaré invariant, the particles are classified in terms of the eigenvalues of two Casimir operators of the Poincaré group, P^2 and W^2 where P_μ are the translation generators and $W_\mu = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}P^\nu M^{\rho\sigma}$ is the **Pauli-Ljubanski polarization vector** with $M^{\mu\nu}$ Lorentz generators. Then, the total angular momentum $J = L + S$ is defined in terms of the generators $M_{\mu\nu}$ as $J^i = \frac{1}{2}\epsilon^{0ijk}M_{jk}$. The generators P_μ and $M_{\mu\nu}$ span the Poincaré algebra, ISO$(3,1)$. When $P^2 = 0$, $W^2 = 0$, W and P are linearly dependent of each other $W_\mu = \lambda P_\mu$; the constant of proportionality λ is given by $\lambda = \overrightarrow{P}\cdot\overrightarrow{J}/P_0$ and defines the *helicity* for massless particles like photons.

Let us turn now to the gravitational fields represented by equation (3). As it has been remarked, they represent gravitational waves moving at the velocity of light, that is, in the would be quantised theory, particles with zero rest mass. Thus, if a classification in terms of Poincaré group invariants could be performed, these waves would belong to the class of unitary (infinite-dimensional) representations of the Poincaré group characterized by $P^2 = 0$, $W^2 = 0$. Recall that, in order for such a classification to be meaningful P^2 and W^2 have to be invariants of the theory. This is not the case for general relativity, unless we restrict to a subset of transformations selected for example by some physical criterion or by experimental constraints. For the solutions of the linearized vacuum Einstein equations the choice of the harmonic gauge does the job [32]. There, the residual gauge freedom corresponds to the sole Lorentz transformations.

2.1. The Standard Linearized Theory

The standard analysis of linearized theory and the issue of the polarization will be analyzed. In particular, the usual transverse-traceless gauge in the linearized vacuum Einstein equations and the (usually implicit) assumptions needed to reduce to this gauge play an important role: the generality of the usual claim "the graviton has spin $-$ 2" (that, of course, is strictly related to the possibility of achieving this special gauge in any "reasonable" physical situation) is strictly related to these assumptions.

The gravitational field is said to be *weak* (in M') if there exists a (harmonic) coordinates system and a region $M' \subset M$ of space-time in which the following

conditions hold

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \qquad |h_{\mu\nu}| \ll 1, \qquad |h_{\mu\nu,\alpha}| \ll 1. \tag{4}$$

As it is known, in the weak field approximations in a harmonic coordinates system the Einstein equations read

$$\Box h_{\mu\nu} = 0. \tag{5}$$

The choice of the harmonic gauge plays a key role in deriving equation (5). No other special assumption either on the form or on the analytic properties of the perturbation h has been done. It is commonly believed that, with a suitable gauge transformation preserving the harmonicity of the coordinate system and the "weak character" of the field, one can always kill the "spin $-$ 1" components of the gravitational waves. However, even if not explicitly declared, the standard textbook analysis of the polarization is performed for globally *square integrable* solutions of the wave-equation (5) (that is, solutions which are square integrable everywhere on M) but, as we will see in the following, some very interesting solutions do not belong to this class.

What is lacking in our case is, obviously, the global *square integrability* due to the presence of the harmonic function solution of the two-dimensional Laplace equation. Therefore, non-globally square integrable spin $-$ 1 perturbations are not pure gauge because they cannot be killed by infinitesimal diffeomorphisms. Even if global *square integrability* is lacking, there exist solutions of this form that far away the singularities are perfectly well-behaved. In other words, spin $-$ 1 perturbations which are square integrable on a submanifold $M' \subset M$ of the whole spacetime can be found: thus, in order to exist, spin $-$ 1 perturbations necessarily need some singularities and/or some region with non trivial topology.

A transparent method to determine the spin of a gravitational wave is to look at its physical degrees of freedom, i.e., the components which contribute to the energy. One should use the Landau-Lifshitz (pseudo)-tensor t^{μ}_{ν} which, in the asymptotically flat case, agrees with the Bondi flux at infinity [8].

It is worth to remark that the canonical and the Landau-Lifchitz energy-momentum pseudo-tensors are tensors for Lorentz transformations. Thus, any Lorentz transformation will preserve the form of these tensor; this allows to perform the analysis according to the Dirac procedure. A globally square integrable solution $h_{\mu\nu}$ of the wave equation is a function of $r = k_{\mu}x^{\mu}$ with $k_{\mu}k^{\mu} = 0$. With the choice $k_{\mu} = (1, 0, 0, -1)$, we get for the energy density t^0_0 and the energy momentum t^3_0 the following result

$$16\pi t^0_0 = \frac{1}{4}(u_{11} - u_{22})^2 + u^2_{12}, \qquad t^0_0 = t^3_0$$

where $u_{\mu\nu} \equiv \mathrm{d}h_{\mu\nu}/\mathrm{d}r$. Thus, the physical components which contribute to the energy density are $h_{11} - h_{22}$ and h_{12}. These amplitudes are eigenvectors of the

infinitesimal rotation generator $\mathcal{R}$, in the plane x–y belonging to the eigenvalues $\pm 2\mathrm{i}$. Thus, the components of $h_{\mu\nu}$ which contribute to the energy correspond to spin $-$ 2.

In the case of the prototype of spin $-$ 1 gravitational waves (3), for $f = 1/2$, we have

$$\tau_0^0 \sim c_1 (h_{0x,x})^2 + c_2 (h_{0y,x})^2, \qquad t_0^0 = t_0^3$$

where c_1 and c_2 are constants, so that the physical components of the metric are h_{0x} and h_{0y}. Following the usual analysis one can see that these two components are eigenvectors of $\mathrm{i}\mathcal{R}$ belonging to the eigenvalues ± 1. In other words, metrics (3), which are not pure gauge since the Riemann tensor is not vanishing, represent spin $-$ 1 gravitational waves propagating along the z-axis at light velocity.

This is related to the harmonic function of the transverse coordinates: in order to have an asymptotically flat wave, singularities or some sort of non triviality in the spacetime topology are necessary. The question is, can reasonable sources be found to smooth out the singularities? The answer is positive as we will see in more details in the next sections. Now we will show a simple and interesting example of such solutions.

As a simple example let us consider perturbations, as in equation (3), of the form $h = \mathrm{d}w(x,y) \cdot \mathrm{d}f(z-t)$ which are not globally square integrable. The metric

$$g = \eta + \mathrm{d}w(x,y) \cdot \mathrm{d}f(u), \qquad u = z - t, \qquad \left(\partial_x^2 + \partial_y^2\right) w = 0 \qquad (6)$$

being spatially asymptotically flat for a wide choice of harmonic functions w. Indeed, it represents a physically interesting gravitational field: gravitational waves propagating along the z-axis at light velocity. Besides to be a solution of the linearized Einstein equations on flat background, it is an exact solution of Einstein equations too.

It is trivial to verify that metric (6) is written in harmonic coordinates and has an *off-diagonal* form, that is, the perturbation h has only one index in the plane x–y orthogonal to the propagation direction z. For this reason the above gravitational wave has spin equal to one and is not a pure gauge [6]. With a suitable transformation it is possible to bring the above gravitational wave in the standard *transverse-traceless* form, however one can check that the new coordinates are not harmonic anymore.

Summarizing: *globally square integrable spin*-1 *gravitational waves propagating on a flat background are always pure gauge. Spin* $-$ 1 *gravitational waves which are not globally square integrable are not pure gauge*.

2.2. Asymptotic Flatness and Matter Sources

In the vacuum case, the coordinates (x, y, z, t) of the metrics (6) are harmonic. Being z the propagation direction, the physical effects manifest themselves in the x–y planes orthogonal to the propagation direction. In order these metrics be *asymptotically Minkowski* for $x^2 + y^2 \to \infty$, the function w is required to satisfy the condition

$$\lim_{x^2+y^2\to\infty} (w - c_1 x + c_2 y - c_3) = 0$$

where c_1, c_2 and c_3 are arbitrary constants and the behaviour of w can be easily recognized by looking at the Riemann tensor of the metrics (6)

$$R_{uiuj} = f_{,u} w_{,ij} \tag{7}$$

which depends on the second derivatives of the harmonic function w.

Therefore, to have an asymptotically Minkowski metric, the function w must be asymptotically close to a linear functions. But, due to standard results in the theory of linear Partial Differential Equations, this is impossible unless w is a linear function everywhere and this would imply the flatness of the metrics (6). However, if we admit δ-like singularities in the x–y planes, non trivial spatially asymptotically Minkowski vacuum solutions with $w \neq \text{const}$ can exist [7]. Of course, it is not necessary to consider δ-like singularities: it is enough to take into account matter sources. For example, in the presence of an electromagnetic wave propagating along the z axis, with energy density equal to ρ which vanishes outside a compact region of the x–y planes, the exact non vacuum Einstein equations for metrics (6) read (see, for example [7])

$$f_{,u} \left(\partial_x^2 + \partial_y^2 \right) w = \kappa \rho$$

where κ is the gravitational coupling constant. Thus, one can have non-singular spin $-$ 1 gravitational waves by considering suitable matter sources which smooth out the singularities.

From the phenomenological point of view, it is worth to note that these kind of wave-like gravitational fields, unlike standard spin $-$ 2 gravitational waves which can be singularities free even in the vacuum case, have to be coupled to matter sources in order to represent reasonable gravitational fields. The observational consequence of this fact is that spin $-$ 1 gravitational waves are naturally weaker than spin $-$ 2 gravitational waves [18]. Typically, if the characteristic velocity of the matter source is v, the spin $-$ 1 wave is suppressed by factors $(v/c)^n$ with respect to a spin $-$ 2 wave. It is worth to note that a gravitational field may also have a repulsive character. For instance, a Kerr black hole is "more repulsive" than a Schwarzschild black hole with the same mass. This is obviously related to the

angular momentum. Roughly speaking, this effect may be attributed to the "gravitomagnetic" part of the Kerr metric which, in our terminology, is the "spin $-$ 1" part. On the other hand, the Kaluza-Klein mechanism allows to construct in pure five-dimensional gravity, solutions with spin $-$ 1 excitations (which in four dimensions may be interpreted as electromagnetic and, therefore, *repulsive*-degrees of freedom. Of course, the Kaluza-Klein mechanism also works when reducing from four to three dimensions. Solutions we are calling *spin $-$ 1 gravitational waves* when reduced to three dimensions (considering as extra dimension the propagation direction of the wave) give rise to purely electromagnetic fields.

3. A Gravitoelectromagnetic Perspective

A different point of view, which is useful in clarifying the nature of spin $-$ 1 gravitational waves is provided by the *gravitoelectromagnetism*, henceforth GEM (see, for example, [17]). In this scheme one tries to exploit as much as possible the similarities between the Maxwell and the linearized Einstein equations. To make this analogy evident it is enough to write a weak gravitational field fulfilling conditions (4) in the GEM form (see, for example, [17, 22])

$$\mathrm{d}s^2 = c^2\left(1+2\frac{\Phi_{(g)}}{c^2}\right)\mathrm{d}t^2 + \frac{4}{c}\left(\mathbf{A}_{(g)}\cdot\mathrm{d}\mathbf{x}\right)\mathrm{d}t - \left(1-2\frac{\Phi_{(g)}}{c^2}\right)\delta_{ij}\mathrm{d}x^i\mathrm{d}x^j \quad (8)$$

with

$$h_{00} = \frac{4\Phi_{(g)}}{c^2}, \qquad h_{0i} = -\frac{4A_{(g)i}}{c^2}$$

(in this section the speed of light c will be explicitly written). Hereafter, the spatial part of four-vectors will be denoted in bold and the standard symbols of three-dimensional vector calculus will be adopted. In terms of $\Phi_{(g)}$ and $\mathbf{A}$, the harmonic gauge condition reads

$$\frac{1}{c}\frac{\partial\Phi_{(g)}}{\partial t} + \frac{1}{2}\nabla\cdot\mathbf{A}_{(g)} = 0 \quad (9)$$

and, once the gravitoelectric and gravitomagnetic fields are defined in terms of GEM potentials, as

$$\mathbf{E}_{(g)} = -\nabla\Phi_{(g)} - \frac{1}{2c}\frac{\partial\mathbf{A}_{(g)}}{\partial t}, \qquad \mathbf{B}_{(g)} = \nabla\wedge\mathbf{A}_{(g)} \quad (10)$$

one finds that the linearized Einstein's equations resemble the Maxwell equations. Consequently, being the dynamics fully encoded in Maxwell-like equations, the GEM formalism describes the physical effects of the vector part of the gravitational field. The situations which are usually described in this formalism are, typically, static. In fact, when this assumption is dropped, GEM gravitational waves are also possible.

The energy-momentum content of gravitational fields of form in (8) is well described, in the asymptotically flat case, by the Landau-Lifshitz pseudo-tensor $t^{\mu\nu}$ [17].

Spin $-$ 1 gravitational waves, as one could expect on the basis of the Kaluza-Klein theory, can be put in an "almost" GEM form. This can be understood recalling that they have only one index in the plane transversal to the propagation direction. A generic spin $-$ 1 gravitational wave, propagating along the z axis on flat spacetimes, has the form in equation (2) with $u = z - t$ and $v = z + t$, and this is almost the same form of equation (8) provided one replaces the time-like index 0 with the light-like index u representing the propagation direction of the wave. The analogous of the gravitomagnetic potential reads in this case $\mathbf{A}_{(g)}(x, y, u) = (\varphi_{,x}, \varphi_{,y}, 0)$ and the harmonic gauge condition is

$$\nabla \cdot \mathbf{A}_{(g)} = \left(\partial_x^2 + \partial_y^2\right) \varphi = 0. \tag{11}$$

Thus, the analogue of the gravitoelectric and magnetic fields are

$$\mathbf{E}_{(g)} = -\frac{1}{2c}(\varphi_{,xu}, \varphi_{,yu}, 0), \qquad \mathbf{B}_{(g)} = (\varphi_{,yu}, -\varphi_{,xu}, 0) \tag{12}$$

and the Einstein equations reduce to

$$\nabla \cdot \mathbf{E}_{(g)} = -4\pi G\rho \tag{13}$$

so that, outside the matter sources, the harmonic gauge condition implies the vacuum field equations.

4. Back to Tolman-Ehrenfest-Podolsky-Wheeler Problem

A steady light beam lying along the z-axis is described by electromagnetic field $F_{\mu\nu}$ whose non vanishing components are

$$E_x = -F_{01} = E_0 \cos(kz - \omega t) = B_y = F_{31}.$$

The only non vanishing components of-the energy momentum tensor $T_{\mu\nu}$ are

$$T_{00} = T_{33} = -T_{03} = -T_{30} = E_0 2 \cos 2(kz - \omega t)/4\pi.$$

Taking the time average over a time grater than ω^{-1} and localizing the waves in a beam, we get

$$T_{00} = T_{33} = -T_{03} = -T_{30} = E_0 2\delta(x)\delta(y)/8\pi$$

where the Dirac delta-function δ has been introduced.

Thus, it is natural to consider a metric perturbation whose non vanishing components are

$$h_{00} = h_{33} = -h_{03} = -h_{30}, \qquad h^{\mu}_{\mu} = 0, \qquad \partial_t h_{\mu\nu} = \partial_z h_{\mu\nu} = 0.$$

Then, the gravitoelectric and the gravitomagnetic components of the metric are given by

$$E_{(g)\mu} = F_{(g)\mu 0}, \qquad B_{(g)_m} = -\varepsilon^{\mu ab} F_{(g)\alpha\beta}/2$$

where

$$F_{(g)}\mu\nu = \partial_\mu A_{(g)\nu} - \partial\nu A_{(g)\mu}, \qquad A_{(g)\mu} = -h_{0\mu}/4 = (\Phi_{(g)}, A_{(g)}).$$

It turns out that

- the first order geodesic motion for a *massive particle* in the light beam gravitational field is determined by the *force*

$$\mathbf{f}_{(g)} = -2\mathbf{E}_{(g)} - 4\mathbf{v} \wedge \mathbf{B}_{(g)}$$

  where $\mathbf{v}$ is the velocity of the particle.
- the first order geodesic motion for a *photon* propagating, in the light beam gravitational field, parallel(anti) to z-axis ($u_j = \pm\delta_{j3}$) is lightly different

$$\mathbf{f}_{(g)} = -4\left(\mathbf{E}_{(g)} + \mathbf{v} \wedge \mathbf{B}_{(g)}\right).$$

In previous section, we have seen that a gravitational wave generated by the light is described by the exact Einstein metric

$$g = \mathrm{d}x^2 + \mathrm{d}y^2 + 2\mathrm{d}u\mathrm{d}v + wu^{-2}\mathrm{d}u^2.$$

In that case, the perturbation is given by

$$h_{00} = h_{33} = -h_{03} = -h_{30} = wu^{-2}$$

and we have

$$\mathbf{E}_{(g)} = -\frac{1}{4}\left(w_x, w_y, \frac{w}{u}\right)u^{-2}, \qquad \mathbf{B}_{(g)} = \frac{1}{4}\left(w_y, -w_x, \frac{w}{u}\right)u^{-2}.$$

The gravitational force acting over a massless particle is given by

$$\mathbf{f}_{(g)} = -[w_x(1-v_z)\mathbf{i} + w_y(1-v_z)\mathbf{j} + (w_x v_x + w_y v_y)\mathbf{k}]/4u^2.$$

If the photon propagates parallel to the light beam, $\mathbf{v} = (0,0,1)$, then

$$\mathbf{f}_{(g)} = 0$$

and there is not attraction or repulsion. It is worth to address that this result holds at first order approximation; the analysis in the strong gravity regime will be exposed in a forthcoming paper [31].

Thus, if the photon propagates antiparallel to the light beam $\mathbf{v} = (0,0,-1)$, then

$$\mathbf{f}_{(g)} = -\nabla w/2u^2$$

and the force turns out to be attractive.

However, it is known from Quantum Field Theory that one consequence of spin $-$ 1 messengers is that particles with the same orientation repel and particles with opposite orientation attract.

Thus, the apparent lacking of attraction must be ascribed to the linear approximation since, according to our results, photons generate spin $-$ 1 gravitational waves and, as a consequence, two photons with same helicity must repel one another [31].

Are these effects observable?

It can be seen that the transversal acceleration per unit length for two laser beams in *VIRGO interferometer* ($W = 1$ watt, separation $d = 10\,\mathrm{cm}$) is only $\mathrm{d}v/\mathrm{d}l = 2.10^{-110}\,\mathrm{cm}^{-1}$ which is too small to be detected with the actual technology.

For a gravitational wave coming from *Virgo cluster* with dimensionless amplitude $h = 10^{-21}$ and frequency $\nu = 1\,\mathrm{KHz}$, it turns out that

$$\mathrm{d}v/\mathrm{d}l = h\nu/c = 3.3.10^{-29}\,\mathrm{cm}^{-1}.$$

Thus, even if the effects cannot be certainly observed in the Laboratory, they may be relevant at cosmic scale.

Acknowledgements

The author wishes to thank F. Canfora, L. Parisi, G. Sparano, S. Vilasi and A. Vinogradov for their contributions to the subject.

References

[1] Aichelburg P. and Sexl R., *On the Gravitational Field of a Massless Particle*, Gen. Rel. Grav. 2 (1971) 303–312.

[2] Bächtold M., *Ricci Flat Metrics with Bidimensional Null Orbits and Non-Integrable Orthogonal Distribution*, Diff. Geom. Appl. **25** (2007) 167–176.

[3] Belinsky V. and Khalatnikov I. , *General Solution of the Gravitational Equations with a Physical Singularities*, Sov. Phys. JETP **30** (1970) 1174–1183.

[4] Belinsky V. and Zakharov V., *Integration of the Einstein Equations by Means of the Inverse Scattering Problem Technique and Construction of the Exact Soliton Solutions*, Sov. Phys. JETP **48** (1978) 6; *Stationary Gravitational Solitons with Axial Symmetry*, Phys. JETP **50** (1979) 1–9.

[5] Bondi H., van der Burg M. and Metzner A., *Gravitational Waves in General Relativity. VII. Waves from Axi-symetric Isolated Systems*, Proc. Roy. Soc. 269, 21 (1962) 21–52.

[6] Canfora F., Vilasi G. and Vitale P., *Nonlinear Gravitational Waves and Their Polarization*, Phys. Lett. B **545** (2002) 373–378.

[7] Canfora F. and Vilasi G., *Spin $-$ 1 Gravitational Waves and Their Natural Sources*, Phys. Lett.B **585** (2004) 193–199.

[8] Canfora F., Vilasi G. and Vitale P., *Spin − 1 Gravitational Waves*, Int. J. Mod. Phys. B **18** (2004) 527–540.

[9] Catalano Ferraioli D. and Vinogradov A., *Ricci-Flat 4-Metrics with Bidimensional Null Orbits, Part I. General Aspects and Nonabelian Case*, Acta Appl. Math. 92 (2006) 209–225.

[10] Catalano Ferraioli D. and Vinogradov A., *Ricci-Flat 4-Metrics with Bidimensional Null Orbits, Part II. Abelian Case*, Acta Appl. Math. 92 (2006) 226–239.

[11] Chandrasekar S., *The Mathematical Theory of Black Holes*, Clarendon Press, Oxford, 1983.

[12] Christodoulou D., *Nonlinear Nature of Gravitation and Gravitational-Wave Experiments*, Phys. Rev. Lett. **67** (1991) 1486–1489.

[13] Einstein A. and Rosen N., *On Gravitational Waves*, J. Franklin Inst. **223** (1937) 43–54.

[14] Faraoni V. and Dumse R., *The Strong Interaction of Light: From Weak to Strong Fields*, Gen. Rel. Grav. **31** (1999) 91–105.

[15] Geroch R., *A Method for Generating New Solutions of Einstein's Equation*, II. J. Math. Phys. **13** (1972) 394–404.

[16] Kompaneyets A., *Strong Gravitational Waves in Free Space*, Sov. Phys. JETP **7** (1958) 659–668.

[17] Mashhoon B., *Gravitoelectromagnetism*, Talk at EREs2000, *gr-qc*/0011014.

[18] Neto E., *Geodesic Deviation in pp-Wave Spacetimes of Quadratic Curvature Gravity*, Phys. Rev. D **68** (2003) 124013–124020.

[19] Penrose R., *A Spinor Approach to General Relativity*, Ann. of Phys. **10** (1960) 171.

[20] Peres A., *Some Gravitational Waves*, Phys. Rev. Lett. **3** (1959) 571–572.

[21] Petrov A., *Einstein Spaces*, Pergamon Press, New York, 1969.

[22] Ruggiero M. and Tartaglia A., *Gravitomagnetic Effects*, Il Nuovo Cim. B **117** (2002) 743–768.

[23] Sakharov A., *Cosmological Transitions with Changes in the Signature of the Metric*, Sov. Phys. JETP **60** (1984) 214–218.

[24] Sparano G., Vilasi G. and Vinogradov A., *Gravitational Fields with a Non-Abelian, Bidimensional Lie Algebra of Symmetries*, Phys. Lett. B **513** (2001) 142–147.

[25] Sparano G., Vilasi G. and Vinogradov A., *Vacuum Einstein Metrics with Bidimensional Killing leaves. I. Local Aspects*, Diff. Geom. Appl. **16** (2002) 95–120.

[26] Sparano G., Vilasi G. and Vinogradov A., *Vacuum Einstein Metrics with Bidimensional Killing Leaves. II. Global Aspects*, Diff. Geom. Appl. **17** (2002) 15–35.

[27] Sparano G., Vilasi G. and Vinogradov A., *Vacuum Einstein Metrics with Bidimensional Killing Leaves. III. Semi-integrable Case* (in preparation).

[28] Stephani H., Kramer D., MacCallum M., Honselaers C. and Herlt E., *Exact Solutions of Einstein Field Equations*, Cambridge University Press, Cambridge, 2003.

[29] Thorne K., *Gravitational-Wave Bursts with Memory: The Christodoulou Effect*, Phys. Rev. D **45** (1992) 520–524.

[30] Tolman R., Ehrenfest P. and Podolsky B., *On the Gravitational Field Produced by Light*, Phys. Rev. **37** (1931) 602–605.

[31] Vilasi G. and Vilasi S., *Is the Light too Light?* (in preparation).

[32] Weinberg S., *Gravitation and Cosmology*, J. Wiley & Sons, N. Y., 1972.

[33] Wheeler J., Phys. Rev. *Geons* **97** (1955) 511–536.

[34] Zakharov V., *Gravitational Waves in Einstein's Theory*, Halsted Press, N. Y., 1973.

© Mehtapress 2012

Journal of
Space Exploration

WWW.MEHTAPRESS.COM

Full Paper

Giorgio Fontana

University of Trento, Via Sommarive 14,
38123 Trento, (ITALY)
E-mail: Giorgio.fontana@unitn.it

Received : July 4, 2012
Accepted : September 5, 2012
Published : September 13, 2012

*Corresponding author's Name & Add.

Giorgio Fontana
University of Trento, Via Sommarive 14,
38123 Trento, (ITALY)
E-mail: Giorgio.fontana@unitn.it

Directions for gravitational wave propulsion

Abstract

This paper summarizes results of past analyses, including proposed examples, in order to build a modern theoretical framework for Gravitational Wave Propulsion. The framework consists of families of generators of gravitational waves, which have been theorized but still require experimentation, and models of thrust generation. High efficiency generators are based on coherent sources, for instance synchronized MEMS oscillators, the HTSC Gaser, based on coherent spin-2 transitions in s-wave/d-wave superconductors, and the nuclear electromagnetic wave to gravitational wave up-converting transducer, based on dineutrons.
After gravitational wave generation is successfully proven in the laboratory, it will be possible to apply a concept developed in the field of cosmology. It was found that the background energy density may give mass to the graviton, which in turn may allow gravitons to produce thrust. Local background energy density can be increased by charging materials with high dielectric constant in close proximity to the wave generating elements. Focused Gravitational Waves may also produce singularities, where the radiation is converted into a coulomb-like gravitational field. Gravitational singularities will set an n-body gravitating system among themselves, the spacecraft, and the remaining bodies of the universe, with obvious propulsive effects. Applications of the present analysis will lead to a unique propulsion system capable of enabling the fast exploration of the solar system, the local star system, and possibly the whole galaxy.

Keywords

Space; Propulsion; Gravitational wave; Gaser; Dineutron; HFGW; HTSC; MEMS.

INTRODUCTION

On a general basis, a vehicle traveling in space requires energy and a reaction mass to accelerate and reach useful speeds. Usually the reaction mass is the mass of the propellant, which in most circumstances has also the role of energy source. Vehicles that are not required to carry reaction masses are more efficient and lightweight, but conventional ones are limited in scope. It is a fact that, after extraordinary developments, space travel by rocket technology has reached its limits and a new paradigm is required to make a big step forward in space propulsion; a step that should enable the exploration of nearby star systems and possibly the whole galaxy. These goals may seem unreachable with the current understanding of physics. Anyway with an open mind and a pragmatic approach, it is well known that we are dealing with opinions that are often suggested by the lack of interdisciplinary approaches to complex problems. It often happened that when so called theoretical limits were found wrong, accidental discoveries have shown why the good theory was erroneously applied the first time. An alternative to accidental discoveries are pieces of knowledge gathered from hundreds of research papers from different disciplines combined in an unusual way to create new concepts. They are normally rejected by experts of their single research field, thus painstaking efforts are required to simply communicate the new concept and let it grow in the laboratories.
At the and of the last century numerous theoretical efforts have started to show that Gravitational Waves (GWs) have not only astronomical and astrophysical relevance, but they also have technological applications[1,2]. Among them, several theories have approaches identified for telecommunication, imaging, material processing, and space propulsion. All of them elaborate on the fact that gravitational waves do exist and are emitted according to a theoretical

framework, namely General Relativity, which looks correct to a high accuracy[3]. Unfortunately successful experiments are yet to occur. However, the emission of GWs by star systems has been observed, and celebrated by a Nobel Prize[4,5].

For developing gravitational wave propulsion, multiple results of this framework have been adopted, combined, and applied to technological devices and methods. This same framework is also currently adopted for developing detectors of GWs from astronomical studies[6]. The detection difficulties suffer by today's detectors' accuracy although they have motivations that will be certainly discovered in the near future and will possibly reveal new information and a new paradigm about the universe and regarding specific approaches to propulsion.

Conceptually, gravitational wave propulsion is based on a generator of gravitational waves, and on one or more methods, that can theoretically produce motion. No physical reaction mass is used. Almost all generators are based on the Nobel Prize tested quadrupole formula[3,5] applied to various GW generating elements; among them there are MEMS oscillators, couples of non-electromagnetically radiating electrons and couples of neutrons.

DISCUSSION

Generators of gravitational waves for propulsion

This section introduces various generators of GWs that appear suitable for propulsion applications. According to the following elementary analysis, it is possible to prove that starting with given mass, volume and mechanical properties for the emitting solid body, the process of splitting the body into smaller parts will permit to increase the frequency and the power emitted.

By considering two equal revolving point masses each with mass m, distance between the masses r, and Universal constant of Gravity G, the power emitted in GWs can be computed by Eq. (10.5.25) at page 272 of Ref.[3], by considering two masses at half distance from the center of rotation, the two equal masses revolving about one another:

$$P = \frac{8G}{5c^5} m^2 r^4 \omega^6 \quad (1)$$

Equation (1) indicates that the power emitted is proportional to the 6th power of the frequency ω; obviously, the higher the frequency, the higher the power. Note that the denominator involving the fifth power of the speed of light, c, implies that the gravitational waves will be very weak. For real world materials a different analysis is required. In fact the actual limit is the ability to keep the emitting object in rotation (and/or vibration) at the highest possible frequency and avoiding that it breaks apart because of excessive internal stresses. The analysis is the following.

For rotating objects, the centripetal force is:

$$F = mr\omega^2 \quad (2)$$

If we choose to keep our system at the limit for the structural integrity of the selected material, F will be bounded by a limiting constant; let's choose force units, or just normalize the expression, according to the properties of the selected material in order to have for this limiting constant $F_{lim}=1$ so that $mr\omega^2 = 1$. Therefore splitting our system in two identical smaller systems each defined by:

$$m_{split} = m/2 \text{ and } r_{split} = r/2 \quad (3)$$

substituting into Eq. (2), we have from the centripetal force limiting formula:

$$\omega_{split} = 2\omega \quad (4)$$

Using these changes in the quadrupole formula for equal masses in Eq. (1) and considering that we obtain two systems, for each of them we obtain:

$$Psplit = \frac{8G}{5c^5}(m/2)^2 (r/2)^4 (2\omega)^6 \quad (5)$$

explicitly it becomes:

$$Psplit = \frac{\frac{8G}{5c^5}(m)^2 (r)^4 (\omega)^6}{4 \cdot 16/64} = P \quad (6)$$

Therefore starting with a given mass (and volume) of a preferred emitting material at the limit of structural integrity and cleverly rearranging the mass distribution in order to keep it always at the limit structural integrity by increasing the frequency, the power emitted by each "small piece" is equal to the power emitted by the "originating" element. If the "small pieces" do not interact with each other, and considering N elementary generators, incoherent summation gives $P_{splitN}=N \cdot P$. If the system of N elementary generators is coherent[7], amplitudes will add up instead of the powers and the power summation will give $P_{splitN}=N^2 \cdot P$, the quadratic growth will proceed till the conversion efficiency of the generator will be near 100%, after that it will saturate. The analysis clearly indicates that it is necessary to arrange the given mass into the largest number of discrete sources that technology allows; this approach leads to High Frequency Gravitational Waves – HFGWs[1]. Moving at high frequency coherently will increase the conversion efficiency of the generator. Equation (6) is the first direction to follow for increasing the power emitted by a GW generator with given mass and/or volume. It follows that the first technical and scientific development for achieving high power and high frequency is the MEMS generator[8]. The MEMS generator is a composite ensemble of micromechanical devices; it can be produced using available technology.

Further developments capable of increasing the frequency and the emitted power in GWs as a function of generator mass may come from particle based and nuclear-based generators. Specifically, in this paper we consider electro-

magnetically shielded electrons and neutral nucleons. Theoretical models have been developed even for those more advanced systems therefore programs for advanced developments are possible.

The MEMS generator

The MEMS generator represents the first significant improvement above the fundamental rotating bar emitter of GW. It is proposed by Dr. Robert M. L. Baker Jr.[9] Because of its very reasonable complexity it will probably be the first one to be tested.

Dr. Baker and collaborators[10], suggested in a number of very detailed papers that commercially available film-bulk acoustic resonators (FBAR), can be arranged in phased arrays and excited by 2.5 GHz electromagnetic microwave generated by magnetrons like those employed in microwave ovens. The phased array will convert EM radiation at 2.5 GHz to 5 GHz High Frequency Gravitational waves (HFGWs).

In Figure 1, a simple MEMS generator is schematized by an array of microscopic GW sources excited in parallel by an electromagnetic field at microwave frequencies that couples to the resonators.

Two electrodes are used to confine the field in a small volume to achieve very high energy densities. In the schematization of Figure 1, the electromagnetic energy in

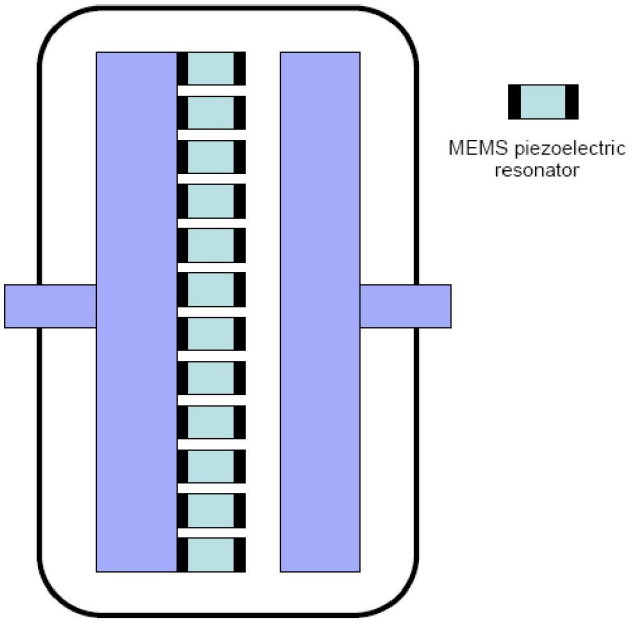


Figure 1 : Schematization of a MEMS generator of HFGWs.

the space surrounding the MEMS has a secondary purpose that will be discussed later.

More complex distributions are under development to improve the density of the microscopic emitters and the ability to drive them synchronously with microwaves and with the best possible use of the quantities appearing in Eq. (1) for each couple of associated masses of the elementary emitter. Among them there is a remarkable three-dimensional FBAR configuration recently proposed by Baker and Baker[9] shown in Figure 2.

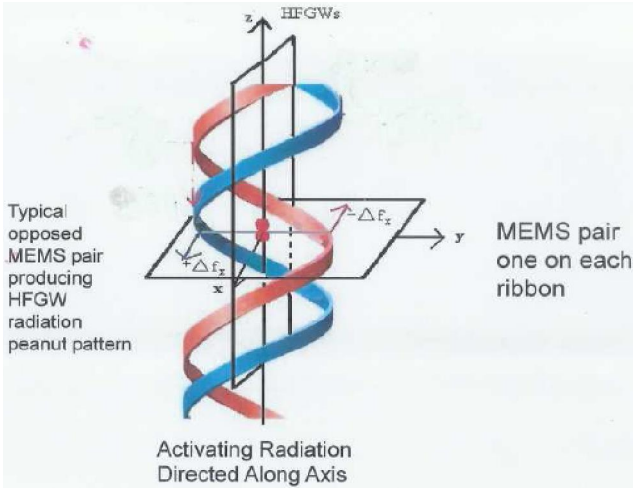


Figure 2 : The double helix MEMS generator of HFGWs.

The FBAR distribution is arranged on a double helix allowing dense packing, and efficient coupling to electromagnetic wave excitation. The configuration takes advantage of coherence and superradiance[9], by which the power emitted is proportional to N^2, where N is the number of elementary sources.

The HTSC gaser

The HTSC Gaser is a quantum source of GW in the THz range. First proposed by the author[11] in 1998, it was theoretically studied for more than a decade[12-15]. The HTSC Gaser promises to offer very high efficiency with directional and focusable GW beams.

The HTSC Gaser adopts electron couples for emitting HFGWs. Because electrons are charged, the whole HTSC Gaser technology is a method to make electron couples to start spinning coherently without emitting electromagnetic waves. The technology is based on early studies of L. Halpern and B. Laurent[16] on the emission of gravitational radiation from microscopic sources. The theory has to wait for the discovery of orthorombic cuprate high Tc superconductors, in which s-wave and d-wave cooper pairs were experimentally observed[17,18]. Transitions between s-wave and d-wave are gravitational *spin-2* transitions and may produce gravitational waves at THz frequencies. The HTSC Gaser depicted in Figure 3 operates by the injection of s-wave Cooper pairs generated in a low-Tc superconductor into the high-Tc superconductor. The injection process is well known and described in the literature with great detail[18]. The relatively new and relevant process is that transitions are linked to a gravitational phenomenon. The HTSC Gaser properly exploits the coherence of the order parameter of the superconductors and superrandiant effects in the very large num-

ber of Cooper pairs in the bulk superconductor.

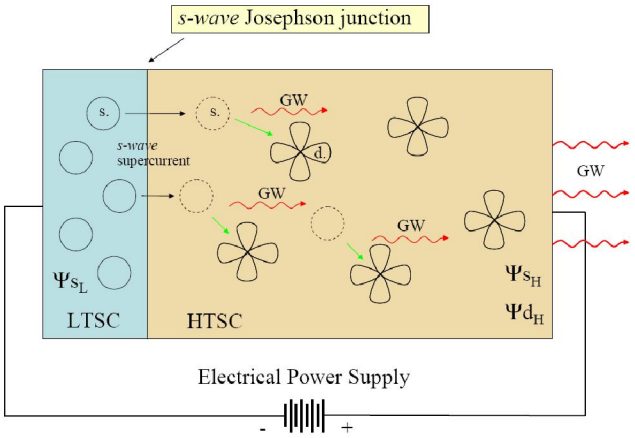


Figure 3 : Spin-2 transitions and HFGW emission in the HTSC gaser.

There are multiple advantages in the use of the HTSC Gaser. It converts DC current directly to HFGWs, it operates at submillimeter wavelengths, allowing easy productions of HFGWs beams. It is an efficient quantum device and allows the production of focused radiation because the junction can be constructed on a spherical dome with HFGWs focused at the center of the sphere. On the other hand it is a cryogenic device and requires special care and the use of very low temperature techniques.

Research should be directed towards the above mentioned potential discoveries because, in spite of the extremely large amount of data collected in more than twenty years, the real nature of high Tc superconductivity remains elusive; gravitational phenomena may be a key factor for a better understanding of the complete phenomenology.

To reach the ultimate goal of producing high intensity HFGWs and further increase the power, the frequency has to be increased as well as the mass of the emitting particles. Neutron mass is thousands of times higher than electron mass and they are electrically neutral. The clever point is how to induce controlled coherent motion to couples of neutrons and, most important, do stable couples of neutrons exist in nature?

It is possible to say yes to both questions.

The dineutron upconverting transducer

The Dineutron Upconverting Transducer (DUT) is the most complex and speculative device. First proposed by the author and Bernd Binder[19], it can convert nuclear excitations both from conventional X-ray sources and NMR signals to HFGWs at a much higher frequency respect to the excitation frequency. It is based on a nuclear model proposed by the late, two times Nobel laureate, Prof. Linus Pauling.

Pauling's nuclear model suggests a nuclear makeup of "spherons" that are protons, neutrons, alpha particles, and the more exotic dineutron[20-23]. Dineutrons are of great interest for this application. In fact dineutrons are ideal emitters of HFGWs. Dineutrons have nuclear densities, they are electrically neutral and the couple has a very small residual magnetic moment. According to Pauling, some nuclei, like for instance some stable isotopes of Gadolinium[21], may have one or two dineutrons floating almost in contact with the "surface" of the core nucleus as shown in Figure 4, under these conditions dineutrons are stable. According to detailed solutions of nuclear equations dineutrons do exist[24], and there are claims of their direct observation[25]. Dineutrons are dynamically coupled to the core of the nucleus by the highly "non linear" strong force. It is not necessary to know much about this dynamical system of nonlinearly coupled oscillators. Every system of nonlinearly coupled oscillators can be forced into a dynamic in which energy fully moves periodically from one oscillator to the other. Each oscillator oscillates at its own characteristic frequency. A well know example is the toy model "swinging spring"[26], which is indeed a simplified model of earth atmosphere. Another example is the multi pendulum marionette, a toy consisting of small pendulums appended to a bigger pendulum.

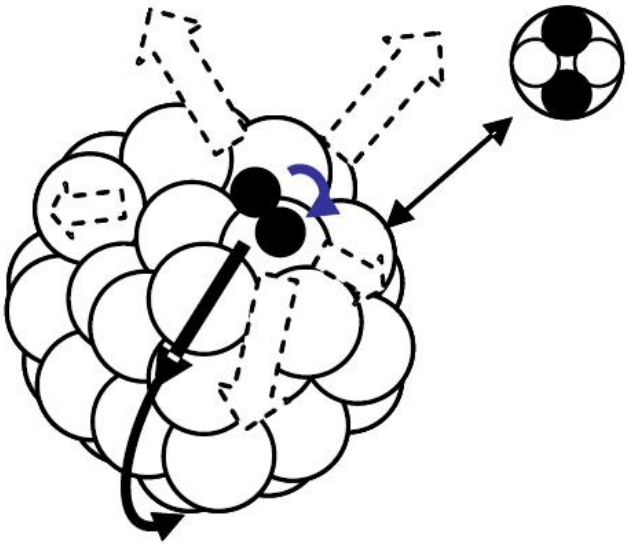


Figure 4 : The dineutron spins and orbits on the surface of the core nucleus, composed of alpha particles in this representation.

In our approach the system of coupled oscillators is composed by two or more spherons, one is the nucleus core, which being electrically charged can absorb electromagnetic waves. The others are the dineutrons that cannot absorb or emit electromagnetic waves; instead they can absorb and emit HFGWs. In Ref.[19], it is shown that according to the quadrupole formula the system can convert electromagnetic waves to HFGWs with high efficiency and at the same time coherently increase the frequency of HFGWs to high harmonics of the exciting electromagnetic radiation. It is the upconversion capability, mediated by the non-

linear "strong force" coupling that makes the efficient emission of HFGWs possible. The converter is estimated to absorb electromagnetic radiation at X-ray frequencies and emit HFGWs at gamma-ray frequencies.

In addition, according to a theory developed by Bernd Binder[27], dynamical instabilities may exist in some heavy nuclei capable of permitting the generation of energy by extracting it from the electromagnetic vacuum fluctuations. We observe that electromagnetic vacuum fluctuations will be converted to HFGWs of much higher frequency. The emitted HFGWs "should" be at the thermodynamic equilibrium with the gravitational quantum fluctuations, being this device an efficient transducer between the two radiation quantum fields. On the other hand, because our knowledge of gravitational quantum fluctuations is marginal at this time, electromagnetic quantum fluctuations might be converted by this device to HFGWs and allowed to propagate into an "almost empty and ultracold" gravitational radiation space, which is pure speculation but still remains a viable theoretical and experimental possibility and, certainly, create a direction to follow for future research. If this hypothesis is experimentally proved correct, a thermodynamic machine could extract useful work from the DUT, which could be therefore used as combined propulsor and useful energy generator.

Other nuclear generators

The DUT is the most complex, speculative, and potentially safe nuclear generator of HFGWs. Different generators were studied in past years and reported here for completeness, they may require nuclear reactions and may produce harmful radiation.

GW pulses could be emitted by the pulse of neutrons from a fission device[28], this is a one-shot concept that cannot be applied to GW propulsion. More recently Fontana and Baker[29] have proposed a HFGWs generating variant of the nuclear propulsion system in which neutrons or antiprotons may induce fission in a blanket of fissionable material. Layer by layer, in Rubbia's design[30], fission fragment may produce thrust.

We proposed an alternative approach in which the fissioning nucleus is prepared in a high rotational state (isomer), as shown in Figure 5; the rotational dynamics before the exact fission instant produces highly deformed nuclei that can emit HFGWs at X-ray and gamma-ray frequencies. By changing the angle between the wavefront of neutrons and the rotational axis of the isomers it is possible to match the speed of neutrons and the speed of HFGWs in order to produce high power pulses. This method is supported by theoretical papers on fission dynamics, on the other hand most energy is released as kinetic energy of fission products. Estimation of the power emitted in HFGWs is reported in Fontana and Baker[29].

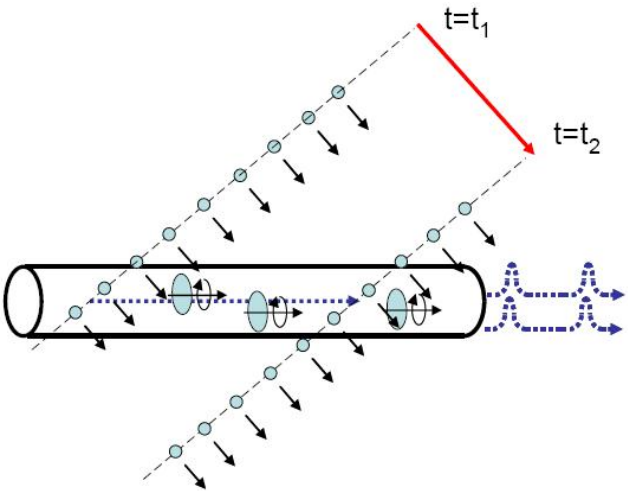


Figure 5 : Conceptual design of a traveling wave HFGW nuclear generator. Fissionable isomers are hit by wevefronts of neutrons. Pulses of HFGW are emitted.

ANALYSIS

Production of gravitons with mass

In the linearized approximation in flat space-time, GWs are purely transverse; therefore the interaction of GWs with matter is not useful for inducing longitudinal motion. In spite of this supposed inapplicability, the production of gravitons with mass is possible and it is the method that may permit propulsion by GWs[15,31]. If gravitons have mass while exiting the generator, propulsion by conservation of linear momentum is a straightforward possibility.

The existence of massive gravitons is predicted by General Relativity, which also gives some information on how to create the conditions required for giving mass to the graviton. The theory has been formulated by studying the propagation of *spin-2* fields in space-times with a background energy density.

While studying the propagation of GWs, it is normally assumed that space-time is flat and gravitational waves propagate as *spin-2* perturbation of flat space-time. Under these conditions gravitons are mass-less and travel at the speed of light. If there is a background energy density, like a cosmological constant, the background curvature and the graviton wave-packet curvature add non-linearly and the unbalance gives mass to the graviton (i.e. the positive perturbation "weights" more than the negative perturbation). Detailed analysis has provided a relationship between the mass of the graviton and the value of the background cosmological constant Λ in order to make propagation formally possible.

The relationship between the mass of the graviton m_g and background Λ, is, expressed in Planck units[32]:

$$mg^2 = \frac{2}{3}\Lambda \quad (7)$$

If the above relationship is not satisfied, the field is no longer a propagating radiation field. The theory was developed for large scale energy fields, like the cosmological constant, but it keeps validity at any scale above some wavelength of the selected GW. We promptly ask what happens if localized static high energy densities, for instance high electric energy densities in vacuum or in dielectrics, surround the GWs generator.

We expect that massive gravitons will be emitted, producing a back reaction on the generator, a thrust, and massive gravitons may possibly interact with distant matter, or maybe not very distant matter, because gravitons will leave the high energy density region where they are produced and will then propagate in "nearly" flat spacetime in which different propagation conditions apply[33]. Equation (7) can be rewritten in terms of energy density with the Einstein field equations:

$$mg^2 = \frac{2}{3}8\pi\rho \quad (8)$$

For engineering applications, the equation is converted to SI units by dividing the mass by the Planck mass. The energy density will be divided by the Planck energy and multiplied by the Planck volume:

$$\frac{m_g^{\,2}}{m_P^{\,2}} = \frac{16\pi}{3}\frac{\rho}{m_P c^2}\left(\frac{\hbar G}{c^3}\right)^{3/2} \Rightarrow m_g = \sqrt{\frac{16\pi m_P}{3}\frac{\rho}{c^2}\left(\frac{\hbar G}{c^3}\right)^{3/2}} \quad (9)$$

Substituting the SI values of the constants involved we have:

$$m_g = \sqrt{\frac{16\pi\cdot 2.17\cdot 10^{-8}}{3}\frac{\rho}{9\cdot 10^{16}}4.22\cdot 10^{-105}} = 1.3\cdot 10^{-48}\sqrt{\rho} \quad (10)$$

Focusing our attention only on the emission process, it is possible to adopt Eq. (10) to predict the amount of thrust that can be obtained from a GWs generator and the associated localized energy density. It is necessary to know the number of gravitons produced per second and the relationship between the energy density and the mass of the gravitons. According to Eq. (10) it is not necessary to know the frequency of the GW to define the mass of the graviton, the frequency must be sufficiently high for the gravitons to be generated in a beam that can resemble a rocket thrust, and therefore this section is developed following this specific approach.

To maximize the mass of the gravitons, it is necessary to employ dielectrics that may support high energy densities. Dielectrics are characterized by the dielectric constant ε higher than that of vacuum. Considering a plane capacitor with area A and armature distance d, the energy E of the electric field in the capacitor as a function of voltage V is:

$$E = 0.5CV^2 = 0.5\,\varepsilon\,V^2\,A/d \quad (11)$$

Dielectrics may suffer voltage induced breakdown if overstressed, and measurements can provide indications about the preferred materials and predicted performances. Table 1 of[34], suggests that $E_b = 10^6$ J/m^3 can be considered a standard high range value. Dielectrics are important for intra atmospheric travel, where air ionization must be reduced without reducing the background energy density. With $E_b = 10^6$ J/m^3, using SI units in the above formula we have:

$$m_g = 1.3\cdot 10^{-48}\sqrt{10^6} = 1.3\cdot 10^{-45}\,\text{kg} \quad (12)$$

It could be considered as a rest mass. To calculate the speed of the gravitons from an engineering point of view, it is possible to consider the Planck relationship applied, for instance, to high frequency gravitons. Let us choose $f = 5$ Ghz. The graviton energy is of the order of $hf = 3.3\cdot 10^{-24}$ J. The mass of the graviton is induced (produced) by the background energy density, being the particle massive and subluminal. It is possible to assume that its energy (most of it) is relativistic kinetic energy. With graviton mass of $1.3\cdot 10^{-45}$ kg, by using the relativistic kinetic energy E_k formula for massive objects, the graviton speed is 0.2 m/s less than the speed of light and the relativistic mass m_{gr} becomes of the order of $3.6\cdot 10^{-41}$ kg. If 10^6 W of GWs are produced by a generator, with $f = 5$ GHz, the number of gravitons per second is $3\cdot 10^{29}$. The amount of mass turns out to be about $1.1\cdot 10^{-11}$ kg/s expelled at nearly the speed of light. The momentum flow rate is $3\cdot 10^8$ m/s $\cdot 1.1\cdot 10^{-11}$ kg/s = $3.3\cdot 10^{-3}$ kg·m·s^{-2}, that is a measurable force.

With invariant GW power and invariant background energy density, reducing the frequency by adopting the generators discussed above, will increase the number of gravitons and reduce their relativistic kinetic energy. Relativistic effects combined with the background energy density that gives a "guaranteed rest mass to each graviton" turns out to be advantageous and with a frequency of 1 μHz, the thrust is 2000 kg·m·s^{-2}; 10^{-10} Hz will produce $2\cdot 10^5$ kg·m·s^{-2} thrust.

Gravitons with low frequency f_L can be produced by mixing two HFGW beams, with frequencies differing by f_L, within the the background energy density field. The method might be called gravitational wave mixing, and certainly will become a new research direction.

Gravitational wave focusing

The non-linearities of Einstein equations have also inspired a large number of theoretical papers on "colliding gravitational waves". Figure 6 shows a numerical solution of the collision of two impulsive gravitational waves[35].

The problem is usually described in u, v null coordinates (representing light propagation), expressed as: $v=t+r^*$; $u=t-r^*$; $r^*=r+2m\log((r-2m)/2m)$. The coordinate transformation is employed to represent space and time in a convenient

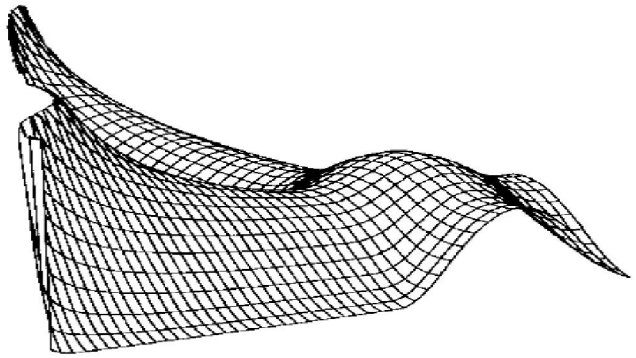


Figure 6 : Representation of the collision of two impulsive gravitational waves.

way, taking advantage of the symmetries of plane waves and "compressing" the space coordinates. The "bump" on the right of the picture represents the superimposing waves. The steep rise on the left of the picture is the effect of the non-linearity of Einstein equations and tends towards the singularity. The singularity does not appear by applying the linear approximation of Einstein equations.

Therefore, it has been shown theoretically and by numerical simulations, that, within the full theory of General Relativity, the singularity appears, accompanied by the diverging amplitude of the gravitational field.

The time (in Planck units) required for the creation of the singularity is a function of the amplitude of the waves A and the relative polarization α of the two waves[36,37].

$$\Delta t = \frac{1}{A^2}\sqrt{1+\sin\alpha} \qquad (13)$$

According to Eq. 13 focused HFGWs with amplitude $A=\Delta l/l=10^{-22}$ at the source may collapse to a singularity in about a second (SI units). This amplitude is relatively large but could even be generated by MEMS in double-helix configuration[9]. The result is equally valid with the more realistic beam-like gravitational waves[38], and the interaction of two graviton beams[39]. Colliding plane gravitational waves will produce a curvature singularity or a coordinate singularity, and the radiation is completely converted into a coulomb-like gravitational field[37].

The calculation of the gravitational potential of the resulting coulomb-like field as a function of the amplitude of the impinging waves would require an evaluation of the energy of the wave during the focusing process, the study of the accumulation process and the lifetime of the resulting singularity. Because of the curved space-time model, energy is not well defined in General Relativity and the lifetime of singularities is the very complex subject of quantum gravitational theories and Hawking radiation. Assuming that the energy of the wave is completely converted into a coulomb-like gravitational field and that no matter is created, all the energy will be converted into additional kinetic and potential energy of the n-body system composed by the spacecraft, the gravitational wave singularity and the remaining bodies of the universe[40]. Because n-body simulations[41] show that most energy exchange involves nearest bodies, if the gravitational singularity is produced near a spacecraft, the associated gravitational potential will mostly affect the kinetic energy of the spacecraft. This phenomenology can be intuitively described as a pulling effect, because the singularity co-moving with the spacecraft continuously pulls the spacecraft toward itself.

A large number of exact and numerical solutions are available for the n-body problem; therefore many solutions are possible, including librations, gravitational slingshots, complex orbital motions and chaotic motion.

It has been shown that a singularity appears also at the focus of a single, focused beam of gravitational radiation. The perfectly focused beam can be represented by a spherical gravitational wave, which has a singularity at the focus[42]. The focus may be the source of the wave or the absorber of the back-scattered wave, the second interpretation is the one of our interest. It is believed that the wave creates the singularity, which moves away at the speed of light. With the spherical gravitational wave, produced for instance, by a hemispherical array of synchronized microscopic sources or by a curved HTSC Gaser, the optical imperfections of the beam near the focus region could be spontaneously reduced by the intrinsic focusing behavior of the collision process. With focused gravitational waves the phenomena described by Eq. 13 will take place at the final stage of the process.

Gravitational wave generators may be external or onboard the spacecraft. Massless gravitons should be preferred in this case, because the direction of motion is reversed respect to the propulsion method described in the preceding two sections.

EXPERIMENTS

Various experiments were already made and patents already filed describing apparatuses and results that may fit the above framework. All of them are considered controversial, some of them were partially replicated, and some of them went completely ignored by physicists and rocket scientists. All of them report effects that are more intense than what can be theoretically predicted by the common elementary use of the theory. It is not the scope of this paper to evaluate the work of other scientists; on the other hand some of their results may fit to a high degree the present analysis. This is a good reason to reconsider them and plan for future replications with the present analysis at hand.

Podkletnov[43] described an apparatus with a weak capacity of gravity modification. The apparatus was a kind of asynchronous electric motor with a dual layered HTSC disc rotor. The operating disc allegedly shielded from earth grav-

ity objects put over it and located inside a cylindrical volume extending upwards for tens of meters. As an alternative interpretation it is possible to suggest that a repulsive force could have been applied to the target mass by an unidentified collimated radiation produced by the disc.

It is well known that the electric current in the rotor of asynchronous motors is higher than the current in the stator windings, and rotor current has maxima during acceleration and braking phases and under mechanical torque rotor load; this is a good reason to believe that the electric rotor current is what drives the phenomenon. In the second Podkletnov experiment[44], electric current with very high intensity was pulsed through the HTSC disc with a gas discharge apparatus. The current flowing through the disc produced similar effects, this time in a pulsed form and directed towards any desired direction, the intensity of the mechanical effects was more dramatic. This capability somehow confirms that the HTSC disc may emit a form of unusual radiation and the effects are very strong with higher voltages applied to the disc and the surrounding gas. According to the present analysis Podkletnov may have built and observed an HTSC Gaser, surrounded by electromagnetic fields capable of inducing the directional emission of massive gravitons. In the Podkletnov device the two layers, with different microscopic structures, but both made with YBCO, were probably characterized by different s-wave/d-wave Cooper pair densities. Because of the measured different critical temperatures, they were certainly characterized by different binding energies of cooper pairs, and transitions were possible. Unfortunately no direct measurement of the Cooper pair states was made on the Podkletnov devices.

The momentum back reaction on the source of massive gravitons was ignored in the Podkletnov apparatus, but it was possibly observed by Poher and described in Patent WO 200709369945. Poher experiment still adopts a dual layer HTSC both for propelling the disc itself and transferring linear momentum to nearby objects. Poher introduced a new hypothetical particle, the "Universon" to build a plausible theoretical framework for his experimental observations. Poher's analysis could be complementary to the one described here.

The coordinated use of concepts like a source of gravitational waves and localized energy densities can be adopted for the analysis and design of these new space propulsion systems.

Less dramatic but still measurable effects have been observed since the beginning of the space era on rotating and highly charged objects, like charged rotating gyroscopes or rotating and electrically charged ballistic missiles. All those objects may have directionally emitted massive gravitons and be consequently affected by linear momentum change. Most observations are still disputed and controversial, waiting for a generally accepted classification and understanding.

CONCLUSIONS

The theoretical foundations of gravitational wave propulsion with massive gravitons were summarized and research directions given. The possibility of producing thrust and induce motion can be the result of the technical use of gravitational waves in combination with high intensity static electromagnetic fields.

Coulomb like gravitational fields produced by colliding gravitational waves may also enable an additional, truly gravitational, propulsion technique, a gravitational pull.

This paper was developed on established principles and results of the astrophysical, the material science and the nuclear science communities applied to space travel engineering. If efficient generators of HFGWs will be constructed and operated at the MW power level, measurable effects should be obtained and it will be possible to start improving the technology. Some known experimental results do fit the presented model as order of magnitude estimates, nature of the phenomena and physical-technical requirements to obtain the claimed result. Therefore, I recommend to replicate known related experiments and to proceed in designing gravitational wave engines for the spacecraft of the future.

NOMENCLATURE AND UNITS

m	=	mass in kg
f	=	frequency in Hz (cycles·s^{-1})
ω	=	$2\pi f$ = angular velocity in rad·s^{-1}
r	=	radial distance in m
v	=	speed in m·s^{-1}
E	=	energy in J
$P=E\cdot s^{-1}$	=	power in W
ρ	=	energy density in J·m^{-3}
ε	=	dielectric constant in F·m^{-1}
$E_k = m_g c^2 \left(\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} - 1 \right)$	=	relativistic kinetic energy of the massive graviton
$m_{gr} = \frac{m_g}{\sqrt{1-\frac{v^2}{c^2}}}$	=	relativistic mass of the massive graviton

$c = 2.997924 \cdot 10^{-8}\ m \cdot s^{-1}$ = speed of light in vacuum

$G = 6.7384 \cdot 10^{-11}\ m^3 kg^{-1} s^{-2}$ = gravitational constant

$\hbar = \frac{h}{2\pi} = 1.054571726 \cdot 10^{-34}\ J \cdot s$ = reduced Planck constant

$m_P = \sqrt{\frac{\hbar c}{G}} = 2.17651 \cdot 10^{-8}\ kg$ = Planck mass

$l_P^3 = \left(\frac{\hbar G}{c^3}\right)^{3/2} = 4.22419 \cdot 10^{-105}\ m^3$ = Planck volume

$E_P = m_P c^2 = \sqrt{\frac{\hbar c^5}{G}} = 4.22419 \cdot 10^{9}\ J$ = Planck energy

$t_P = \sqrt{\frac{\hbar G}{c^5}} = 5.39106 \cdot 10^{-44}\ s$ = Planck time

$G_{\mu\nu} = 8\pi T_{\mu\nu}$ = Einstein field equations in Planck units

ACKNOWLEDGEMENTS

The author whishes to acknowledge Paul Murad for reviewing and cleaning the paper and Robert M. L. Baker jr. for suggestions, discussion and a picture included in this paper. The STAIF 2 organizing committee is kindly acknowledged.

REFERENCES

[1] R.M.L.Baker Jr.; AIP Conference Proceedings, **746**, 1306 **(2005)**.
[2] G.G.Cuadrado; AIP Conference Proceedings, **1103**, 553 **(2009)**.
[3] S.Weimberg; Gravitation and cosmology: Principles and applications of the general theory of relativity, John Wiley and Sons, New York, **(1972)**.
[4] Russel A.Hulse; Reviews of Modern Physics, **66**, 699 **(1994)**.
[5] J.H.Taylor Jr.; Reviews of Modern Physics, **66**, 711 **(1994)**.
[6] K.S.Thorne; Reviews of Modern Physics, **52**, 285 **(1980)**.
[7] R.H.Dicke; Physical Review, **93**, 99 **(1954)**.
[8] R.M.L.Baker Jr.; Proceedings of the gravitational-wave conference, P.Murad, R.Baker, (Ed); The MITRE Corporation, McLean, Virginia, Paper HFGW-03-117, **(2003)**.
[9] R.M.L.Baker, B.S.Baker; 'Gravitational wave generator apparatus' in the Proceedings of the space technology & applications international forum (STAIF II 2012), March 13-15, Albuquerque, New Mexico, USA.
[10] R.M.L.Baker Jr.; G.V.Stephenson, F.Li; AIP Conference Proceedings, **969**, 1036 **(2008)**.
[11] Cornell University Library, accessed on 23-06-2011; http://arxiv.org/abs/cond-mat/9812070
[12] G.Fontana; AIP Conference Proceedings, **504**, 1085 **(2000)**.
[13] G.Fontana, R.M.L.Baker, Jr.; Proceedings of the gravitational-wave conference, P.Murad, R.Baker, (Ed); The MITRE Corporation, Mclean, Virginia, May 6-9, Paper HFGW-03-107 **(2003)**.
[14] G.Fontana; AIP Conference Proceedings, **699**, 1114 **(2004)**.
[15] G.Fontana; in Gravity-superconductor interactions: Theory and experiment, Bentham science e-book, G.Modanese, G.Robertson, (Ed); **(2011)**.
[16] L.Halpern, B.Laurent; Il Nuovo Cimento, **33(3)**, 728 **(1964)**.
[17] C.C.Tsuei, J.R.Kirtley; Review of Modern Physics, **72(4)**, 969 **(2000)**.
[18] M.Moessle, R.Kleiner; Physical Review B, **59(6)**, 4486 **(1999)**.
[19] G.Fontana, B.Binder; AIP Conference Proceedings, **1103**, 524 **(2009)**.
[20] L.Pauling; Proc.Natl.Acad.Sci.USA, **54**, 989 **(1965)**.
[21] L.Pauling; 'Binding energy of dineutrons Z=64, N=82-98', accessed on 25-10-2011; http://osulibrary.orst.edu/specialcollections/rnb/32/32-143.html
[22] L.Pauling; 'Orbiting clusters in atomic nuclei', Proc.Natl.Acad.Sci.USA, **64**, 807 **(1969)**.
[23] L.Pauling; Proc.Natl.Acad.Sci.USA, **87**, 4435 **(1990)**.
[24] N.Itagaki, et al.; Physical Review C, **65**, 044302-1 **(2002)**.
[25] S.Kamal, B.Parker; Physical Review Letters, **66**, 2448 **(1991)**.
[26] P.Lynch; International Journal of Nonlinear Mechanics, **37**, 345-367 **(2002)**.
[27] B.Binder; AIP Conference Proceedings, **969**, 1103 **(2008)**.
[28] G.F.Chapline, J.Nuckolls, L.L.Woods; Physical Review D, **10(4)**, 1064 **(1974)**.
[29] G.Fontana, R.M.L.Baker Jr.; AIP Conference Proceedings, **813**, 1352 **(2006)**.
[30] C.Rubbia; Fission fragments heating for space propulsion, CERN SL-Note - 2000-036 EET, **(2000)**.
[31] G.Fontana; AIP Conference Proceedings, **746**, 1331 **(2005)**.
[32] M.Novello, R.P.Neves; Classical and Quantum Gravity, **20**, L67 **(2003)**.
[33] G.Modanese, G.Fontana; AIP Conference Proceedings, **699**, 1198 **(2004)**.
[34] D.R.Lohrmann, D.Ma, D.Wu; Report NRL/MR/5747—98-8197, August 26 **(1998)**.
[35] R.W.Corkill, J.M.Stewart; Proc.R.Soc.Lond.A, **386**, 373 **(1983)**.
[36] V.Ferrari; Physical Review D, **37(10)**, 3061 **(1998)**.
[37] V.Ferrari; Proceedings of the 8th Italian conference on general relativity and gravitational physics, M.Cerdonio, R.Cianci, M.Francaviglia, M.Toller, (Ed); World Scientific, Singapore, 171 **(1988)**.
[38] V.Ferrari, P.Pendenza, G.Veneziano; General Relativity and Gravitation, **20(11)**, 1185 **(1998)**.
[39] G.Veneziano; Modern Physics Letters A, **2(11)**, 899 **(1987)**.
[40] Q.D.Wang; Celestial Mechanics and Dynamical Astronomy, **50**, 73 **(1991)**.
[41] E.Buticov; Example 9 at this www page: accessed on 23 06 2011; http://faculty.ifmo.ru/butikov/Projects/Collection6.html#_applet
[42] G.A.Alekseev, J.B.Griffiths; Classical and Quantum Gravity, **13**, 2191 **(1996)**.
[43] E.Podkletnov, R.Nieminen; Physica C, **203**, 441 **(1992)**.
[44] E.Podkletnov, G.Modanese; J.Low.Tem.Phys., **132**, 239 **(2003)**.
[45] C.Poher; WO/2007/093699.